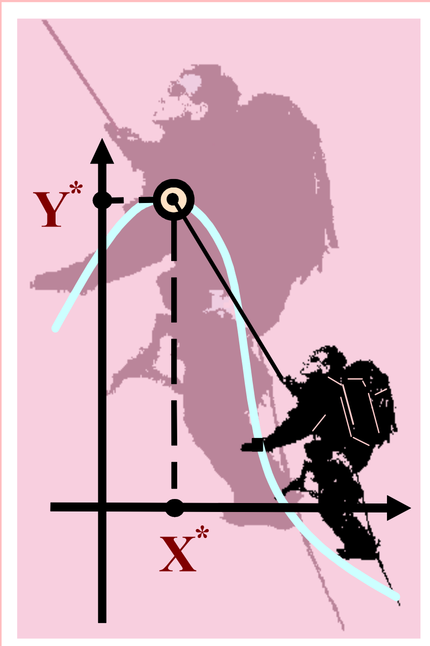
М. І. Самойленко



МАТЕМАТИЧНЕ ПРОГРАМУВАННЯ

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ХАРКІВСЬКА ДЕРЖАВНА АКАДЕМІЯ
МІСЬКОГО ГОСПОДАРСТВА

М.І. Самойленко

МАТЕМАТИЧНЕ ПРОГРАМУВАННЯ

*Рекомендовано Міністерством освіти і науки України
як навчальний посібник
для студентів економічних спеціальностей
вищих навчальних закладів*

Харків
ОСНОВА
2001

ББК 22.11я73
С17
УДК 717 (07)+519.6

Гриф надано Міністерством освіти і науки України, лист №14/18.2-788 від 06.06.2001.

Р е ц е н з е н т и :

директор Інституту комп'ютерних і інформаційних технологій Харківського державного технічного університету радіоелектроніки, д-р. техн. наук, проф. *В.М.Левикін*;

зав. кафедри вищої математики Харківської державної академії міського господарства, д-р техн. наук, проф. *А.І.Колосов*

Самойленко М.І.

С17 Математичне програмування: Навч. посібник. – Харків: Основа, 2001. – 424 с.

ISBN 5–7768–0760–3.

Посібник містить теоретичні відомості про основи математичного програмування та поширені методи розв'язання екстремальних задач. Наведені методи ілюструються типовими прикладами. Пропонуються набори завдань економічного змісту для самостійного рішення та комплекси індивідуальних завдань для перевірки знань студентів, розраховані на академічну групу. Посібник також містить необхідний довідковий матеріал, предметний покажчик і відповіді.

Для студентів економічних спеціальностей вищих навчальних закладів і економістів-практиків.

Табл. – 36, іл. – 30, бібліогр. – 11 назв.

$$С\ \frac{1602010000-42}{226-2001}$$

ББК 22.11я73

ISBN 5–7768–0760–3

Присвячується 80-річчю
Харківської державної академії
міського господарства

ПЕРЕДМОВА

Цей навчальний посібник призначений для студентів економічних спеціальностей вузів денної і заочної форм навчання, які прослухали загальний курс вищої математики відповідно до програми бакалаврської підготовки.

Мета посібника – сприяти подальшому підвищенню рівня фундаментальної математичної підготовки студентів. Особливостями його є економічний ухил у викладі матеріалу, індивідуалізація навчання і можливість самостійного вивчення курсу. З цією метою по кожній темі курсу до посібника включені добірки задач економічного характеру, розраховані на індивідуальне вирішення студентами. Для забезпечення можливості самостійного вивчення курсу посібник містить приклади розв'язаних типових задач по кожній темі, набори завдань для самостійного вирішення з відповідями і контрольні запитання для самоперевірки.

Наведені наприкінці книги додатки містять довідкові відомості, необхідні для виконання індивідуальних завдань.

Завдання для самостійного вирішення мають порядкову нумерацію в межах розділу. Даються відповіді до задач по розділах.

Завдання для індивідуального (аудиторного або самостійного) вирішення мають по 30 варіантів. Вибір номера варіанта повинен відповідати порядковому номеру студента в академічному журналі групи.

Таким чином, навчальним посібником можна користуватися як довідником, підручником і в той же час задачником, що особливо важливо і зручно для студентів і читачів, які самостійно вивчають курс.

В основу посібника покладено курс лекцій з прикладної математики, що читається автором протягом багатьох років.

Користуюсь приємною можливістю висловити щиру подяку рецензентам за їхню копітку роботу по резензуванню навчального посібника й суттєві зауваження, що сприяли поліпшенню змісту книги і методики її викладання.

ВСТУП

Переведення економіки країни на шлях інтенсивного розвитку безпосередньо пов’язано з необхідністю підвищення ефективності використання математичної теорії у прикладній сфері діяльності людини. Вирішальну роль у досягненні поставленій мети повинні відіграти фахівці, сякі добре володіють математичними методами і мають достатній досвід їх використання при вирішенні практичних задач. Теоретичну підготовку таких фахівців здійснює вища школа. Математичне програмування як прикладний розділ вищої математики відіграє винятково важливу роль у підготовці фахівців економічного профілю. Використання математичних методів в інженерно-економічній діяльності дозволяє вирішувати оптимальним способом багато економічних і організаційних задач. Іншими словами, інженер-економіст стає власником надійного інструменту для одержання найвищого економічного ефекту в конкретних виробничих умовах.

Прикладами можливих економічних задач, що наочно ілюструють корисність і необхідність знання запропонованої дисципліни, можуть бути наступні задачі (надаються в змістовній постановці):

- одержання максимального випуску продукції або максимального прибутку при заданих матеріальних, трудових або тимчасових витратах;
- забезпечення планових показників підприємства при мінімальних фінансових витратах;
- досягнення максимально короткого терміну виготовлення продукції, будівництва об'єкта, товарообігу, виробничого циклу і т.п. при існуючих або заданих виробничих ресурсах (матеріальних, трудових, енергетичних, тимчасових та ін.).

> У наведених прикладах максимальний випуск продукції, максимальний прибуток, мінімальні фінансові вкладення, максимально короткий термін є шукані *оптимуми* (*максимуми* або *мінімуми*).

У математиці максимум і мінімуми мають ще одну назву – *екстремум*, а задачі пошуку екстремуму називають екстремальними задачами.

У наведених прикладах умови, що накладаються на вирішення задачі (задані матеріальні, трудові й тимчасові витрати; планові показники; виробничі ресурси), називають *обмеженнями* задачі.

Обмеження задач визначають *область припустимих рішень.*

Ті припустимі рішення, при яких досягається оптимум, називають *оптимальними,* або *екстремальними рішеннями.*

У загальному випадку екстремальна задача може мати одне, декілька, безліч, нескінчену безліч або жодне оптимальне рішення.

У практиці інженера-економіста оптимальне рішення прийнято називати *оптимальним планом*.

Змістовна постановка задачі повинна дозволяти переходити до строгої *математичної моделі*. У протилежному випадку необхідно пройти досить трудомісткі й копіткі процеси математичного моделювання та ідентифікації, але останні тут не розглядаються.

У загальному вигляді екстремальна задача формулюється наступним чином: визначити найбільше (максимальне) або найменше (мінімальне) значення деякої функції $Y(x_1, x_2, \ldots, x_n)$ при умовах $f_i(x_1, x_2, \ldots, x_n) \le b_i$ $(i = \overline{1, m})$, де y і f_i – задані функції, а b_i – дійсні числа.

Наведене формулювання є узагальненням постановок ряду приватних задач математичного програмування, що можуть розрізнятися між собою як видом функцій y і f_i, так і характером перемінних (дискретний, безупинний).

Функцію $Y(x_1, x_2, \ldots, x_n)$, яку оптимізують (мінімізують або максимізують), називають *цільовою* функцією.

Залежно від особливостей функцій y і f_i математичне програмування можна розділити на ряд самостійних дисциплін, що вивчають і розробляють методи вирішення окремих класів задач.

Насамперед, задачі математичного програмування поділяють на задачі *лінійного* і *нелінійного програмування*. При цьому, якщо всі функції y і f_i є лінійними, то відповідні задачі відносяться до класу задач лінійного програмування. Якщо ж хоча б одна із зазначених функцій є нелінійною, то відповідна задача відноситься до класу задач нелінійного програмування.

Найбільш вивченим розділом математичного програмування є лінійне програмування. Для вирішення задач лінійного програмування розроблено цілий ряд ефективних методів, алгоритмів і програм.

Серед задач нелінійного програмування найбільш глибоко вивчені задачі *опуклого програмування*. Це задачі, в результаті вирішення яких визначається екстремум опуклої функції, заданої на опуклій замкнутій множині.

У свою чергу, серед задач опуклого програмування більш докладно досліджені задачі *квадратичного програмування*. У результаті вирішення таких задач потрібно в загальному випадку знайти екстремум квадратичної функції при обмеженнях на змінні у вигляді системи лінійних рівнянь або лінійних нерівностей.

Визначеними класами задач математичного програмування є задачі *цілочислового, параметричного* і *дрібно-лінійного програмування*.

У задачах *цілочислового*, або *дискретного* програмування невідомі можуть приймати тільки цілочислові значення.

У задачах параметричного програмування цільова функція, або функції, що визначають область можливих змін перемінних, або то і інше залежать від деяких параметрів.

У задачах дрібно-лінійного програмування цільова функція являє собою відношення двох лінійних функцій, а функції, що визначають область припустимих рішень, також є лінійними.

Особливі класи становлять задачі *стохастичного* і *динамічного програмування.* Якщо в цільовій функції або функціях-обмеженнях, містяться випадкові розміри, то така задача відноситься до класу задач стохастичного програмування. Задача, процес знаходження якої є багатоетапним, належить до класу задач динамічного програмування.

Таким чином, *математичне програмування* є математичною дисципліною, що досліджує екстремальні задачі і розробляє методи їх вирішення. Математичне програмування як наука знаходиться у процесі постійного розвитку. Вченими всього світу розроблено багато методів для вирішення різних класів задач математичного програмування. Разом з тим багато задач ще не мають ефективних методів розв'язання і чекають своїх дослідників. Таким дослідником може стати будь-хто. Автор буде щиро радий, якщо цей навчальний посібник виявиться тією відправною точкою, що приведе читача до нових результатів в області математичного програмування.

Навчальний посібник має такі методичні особливості, що, на думку автора, вигідно відрізняють його серед інших посібників аналогічного типу. Він орієнтований не тільки на підготовку інженерів-економістів, а також і економістів-дослідників, тобто на прищеплення студентам дослідницької жилки.

Основною науковою і методичною особливістю посібника є єдиний підхід до вирішення задач математичного програмування, заснований на використанні *диференціального алгоритму* розв'язання екстремальних задач різних класів. Такий підхід дозволяє розглядати задачі лінійного і квадратичного програмування як часткові щодо загальної задачі математичного програмування (нелінійного програмування).

Оскільки матеріали з розвитку й узагальнення диференціального алгоритму публікуються вперше, вони наводяться в розширеному обсязі. При цьому ставиться мета не стільки навчити студентів новим методам математичного програмування, скільки звернути їхню увагу на динамічний розвиток прикладних наук і необхідність критичного аналізу до їх опанування.

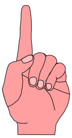
Інша методична особливість посібника стосується початкового етапу вивчення дисципліни. На перших практичних заняттях, коли безпосередній матеріал з математичного програмування ще не викладений у достатньому обсязі, розглядаються вже знайомі для аудиторії розділи вищої математики (елементи матричного числення, лінійна алгебра, теорія множин), але з введенням нових елементів, необхідних для розуміння курсу й спрощення методів вирішення задач математичного програмування. Це дозволяє здійснити органічний перехід від загального курсу вищої математики до математичного програмування, робить його більш цілісним і незалежним від інших дисциплін.

Розділ 1. ЕЛЕМЕНТИ МАТРИЧНОЇ АЛГЕБРИ

Тут розглядаються елементи матричного числення, необхідні для вивчення і розуміння наступних розділів курсу математичного програмування, а також багато математичних викладок , що сприяють компактному поданню навчального матеріалу. Матричне числення - зручний і ефективний спосіб скорочення навчального часу, що надається для вивчення дисципліни, виконання індивідуальних завдань і вирішення задач на практичних заняттях або під час самопідготовки.

1.1. Матриця. Вектор

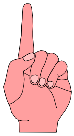
Матриця є основним елементом матричного числення.



Визначення 1.1. Матрицею називають прямокутну таблицю елементів, розташованих по рядках і стовпцях:

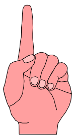
$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}.$$

Якщо матриця містить mn елементів, що утворять m рядків і n стовпців, то говорять, що вона має розмір m на n, і записують це як $m \times n$.



Визначення 1.2. Матрицю розміру $n \times n$ називають *квадратною порядку* n.

Елемент, що знаходиться на перетині i-го рядка і j-го стовпця, позначається рядковою латинською літерою a_{ij} з подвійною індексацією, набраною нежирним курсивним шрифтом. Перший індекс завжди відповідає порядковому номеру рядка, в якому розташовується елемент, а другий – порядковому номеру стовпця.



Визначення 1.3. Елементи квадратної матриці a_{ij}, що мають рівні індекси *(i=j)*, називаються *діагональними*, а лінія, на якій розташовані діагональні елементи, – *головною діагоналлю*.

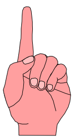
Для матриць використовують також позначення:

$$\begin{Vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{Vmatrix}; \quad \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}; \quad \left[a_{ij}\right], \ (i = \overline{1,m}, \ j = \overline{1,n}).$$

У математичних виразах для стислості матриці позначають великими латинськими літерами, набраними жирним шрифтом (не курсивом). Наприклад: **A**, **X.**

Окремим випадком матриці є *вектор*. Розрізняють *вектор-стовпець* і *вектор-рядок*. Вектор-стовпець має розмір $m \times 1$ і позначається рядковими латинськими літерами, набраними жирним шрифтом (не курсивом). Елементи вектора позначаються рядковими курсивними літерами з простою (одинарною) індексацією. Наприклад, $\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \dots \\ b_m \end{bmatrix}$. Вектор-рядок має розмір $1 \times n$ і позначається також рядковими латинськими літерами, але з верхнім індексом «т». Наприклад, $\mathbf{c}^{\text{т}} = \begin{bmatrix} c_1 & c_2 & \cdots & c_n \end{bmatrix}$.

1.2. Математичні дії над матрицями



Визначення 1.4. Дві матриці $\mathbf{A}$ і $\mathbf{B}$ називаються *рівними,* якщо вони мають однаковий розмір і $a_{ij} = b_{ij}$ для всіх i і j. Це означає, що рівні матриці збігаються по кожному елементу.

Додавання матриць. Якщо $\mathbf{A}$ і $\mathbf{B}$ – матриці одного розміру, їх можна скласти. При цьому утворюється нова матриця **C**=**A**+**B**, що має той же розмір, і для всіх i та j справедливо

$$\boxed{c_{ij} = a_{ij} + b_{ij}},$$

тобто щоб одержати елементи матриці $\mathbf{C}$, потрібно скласти відповідні елементи матриць $\mathbf{A}$ і $\mathbf{B}$. Наприклад, якщо $\mathbf{A} = \begin{bmatrix} 2 & 0 \\ -5 & 6 \end{bmatrix}$ і $\mathbf{B} = \begin{bmatrix} -3 & 6 \\ 4 & 1 \end{bmatrix}$, то $\mathbf{C} = \mathbf{A} + \mathbf{B} = \begin{bmatrix} -1 & 6 \\ -1 & 7 \end{bmatrix}$.

Множення матриці на скаляр. Якщо λ - скаляр, то добуток матриці на скаляр визначається як

$$\boxed{\lambda\mathbf{A} = \lfloor \lambda a_{ij} \rfloor},$$

тобто кожний елемент матриці **A** збільшується на λ. Наприклад, якщо $\mathbf{A} = \begin{bmatrix} 2 & 0 \\ -5 & 6 \end{bmatrix}$ і λ=–5, то $\lambda\mathbf{A} = \begin{bmatrix} -10 & 0 \\ 25 & -30 \end{bmatrix}$.

З правил додавання матриць і множення матриці на скаляр випливає

$$\mathbf{A} - \mathbf{B} = \mathbf{A} + (-1)\mathbf{B} = \lfloor a_{ij} \rfloor + \lfloor -b_{ij} \rfloor = \lfloor a_{ij} - b_{ij} \rfloor .$$

Множення матриць. Якщо матриця **A** має розмір $m \times n$, а матриця **B** має розмір $n \times p$, то добуток матриць **AB** визначається як нова матриця **D** розміру $m \times p$, в якій елемент, що знаходиться на перетині i-го рядка і j-го стовпця, дорівнює

$$\boxed{d_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}} .$$

Отже, елемент добутку матриць, що має індекси ij, визначається як сума попарних добутків елементів i–го рядка першої матриці на відповідні елементи j–го стовпця другої матриці. Для того, щоб це було можливо, необхідна рівність числа елементів у рядку першої матриці і числа елементів у стовпці другої матриці, тобто необхідна рівність числа *стовпців* першої матриці числу *рядків* другої матриці. Матриці, що мають таку властивість, називають *відповідними* щодо операції множення. Тому завжди треба вказувати розмір матриць при їхньому множенні.

У добутку **AB** матриця **A** розглядається як перший (попередній) співмножник стосовно **У**, а **B** – як другий (наступний) співмножник стосовно **A.** Добуток **BA** звичайно відрізняється від **AB**, а може і зовсім не існувати. Обидва добутки **AB** і **BA** існують тільки в тому випадку, коли матриці мають розміри $m \times n$ і $n \times m$. У цьому випадку перший добуток є матриця розміру $m \times m$, а другий – матриця розміру $n \times n$.

Розглянемо приклад. Нехай дані матриці $\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix}$ і $\mathbf{B} = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \end{bmatrix}$. Тоді їхній добуток $\mathbf{AB}$ є матриця порядку 2×2:

$$\mathbf{AB} = \begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} + a_{13}b_{31} & a_{11}b_{12} + a_{12}b_{22} + a_{13}b_{32} \\ a_{21}b_{11} + a_{22}b_{21} + a_{23}b_{31} & a_{21}b_{12} + a_{22}b_{22} + a_{23}b_{32} \end{bmatrix},$$

а добуток $\mathbf{BA}$ є матриця розміру 3×3:

$$\mathbf{AB} = \begin{bmatrix} b_{11}a_{11} + b_{12}a_{21} & b_{11}a_{12} + b_{12}a_{22} & b_{11}a_{13} + b_{12}a_{23} \\ b_{21}a_{11} + b_{22}a_{21} & b_{21}a_{12} + b_{22}a_{22} & b_{21}a_{13} + b_{22}a_{23} \\ b_{31}a_{11} + b_{32}a_{21} & b_{31}a_{12} + b_{32}a_{22} & b_{31}a_{13} + b_{32}a_{23} \end{bmatrix}.$$

Слід бути особливо уважним при перенесенні законів скалярної алгебри на алгебру матриць. Усі вони мають бути виведені заново. При цьому одні матимуть місце для алгебри матриць, а інші – ні. Перерахуємо найбільш важливі закони алгебри матриць, що будуть потрібні для здійснення послідовних матричних операцій.

Комутативний закон. Для матричних операцій додавання справедливий комутативний закон $\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$. Матриці $\mathbf{A}$ і $\mathbf{B}$ повинні, звичайно, мати однаковий розмір. Цей закон виводиться безпосередньо з визначення дії додавання матриць.

Приклад 1.1.

$$\begin{bmatrix} 2 & 3 \\ -5 & 0 \end{bmatrix} + \begin{bmatrix} 0 & -1 \\ -2 & 6 \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ -2 & 6 \end{bmatrix} + \begin{bmatrix} 2 & 3 \\ -5 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 2 \\ -7 & 6 \end{bmatrix}.$$

Для операції множення комутативний закон несправедливий $\mathbf{AB} \neq \mathbf{BA}$. Якщо дві матриці мають відповідно розміри $m\times n$ і $n\times m$, то обидва добутки $\mathbf{AB}$ і $\mathbf{BA}$ існують, однак вони матимуть різні розміри. Якщо ж обидві матриці квадратні й одного розміру, то обидва добутки матимуть той же розмір, але не обов'язково вони будуть рівні, що показує такий приклад.

Якщо $\mathbf{A} = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$; $\mathbf{B} = \begin{bmatrix} 3 & 0 \\ 1 & 2 \end{bmatrix}$, то $\mathbf{AB} = \begin{bmatrix} 7 & 2 \\ 4 & 2 \end{bmatrix}$, $\mathbf{BA} = \begin{bmatrix} 6 & 3 \\ 4 & 3 \end{bmatrix}$.

Якщо ж $\mathbf{A} = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$; $\mathbf{B} = \begin{bmatrix} 1 & -1 \\ 1 & 2 \end{bmatrix}$, то $\mathbf{AB} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \mathbf{BA}$.

Асоціативний закон. Закон має місце і для операції додавання матриць $(\mathbf{A} + \mathbf{B}) + \mathbf{C} = \; = \mathbf{A} + (\mathbf{B} + \mathbf{C})$. Справедливість цього закону випливає з того, що додавання матриць виконується поелементно, а для елементів не істотно, в якому порядку вони складаються.

Асоціативний закон має місце і для множення матриць $(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC})$. Можна послідовно обчислити вираз, що стоїть зліва, виконавши спочатку множення $\mathbf{AB}$, а потім помножити результат на $\mathbf{C}$ справа. Той же результат буде отриманий, якщо спочатку виконати множення $\mathbf{BC}$, а потім результат помножити на $\mathbf{A}$ справа.

Нехай $\mathbf{A} = [a_{ir}]$ – матриця розміру $m \times n$; $\mathbf{B} = [b_{rs}]$ – матриця розміру $n \times p$; $\mathbf{C} = [c_{sj}]$ – матриця розміру $p \times q$. Тоді добуток $\mathbf{AB}$ буде матрицею розміру $m \times p$, в якій на перетині i–го рядка і s–го стовпця знаходиться елемент, рівний $\sum_{r=1}^{n} a_{ir} b_{rs}$. Матриця $\mathbf{ABC}$ створюється множенням $\mathbf{AB}$ на $\mathbf{C}$ справа. Це буде матриця розміру $m \times q$, елемент ij якої є $\sum_{s=1}^{p} \left(\sum_{r=1}^{n} a_{ir} b_{rs} \right) c_{sj} = \sum_{s=1}^{p} \sum_{r=1}^{n} a_{ir} b_{rs} c_{sj}$. Аналогічно елементом rj матриці $\mathbf{BP}$ буде $\sum_{s=1}^{p} b_{rs} c_{sj}$, а елемент ij матриці $\mathbf{A}(\mathbf{BP})$ стане рівним

$\sum\limits_{s=1}^{n} a_{ir}\left(\sum\limits_{r=1}^{p} b_{rs} c_{sj}\right) = \sum\limits_{s=1}^{p}\sum\limits_{r=1}^{n} a_{ir} b_{rs} c_{sj}$, тобто збігається з елементом ij матриці **(AB)C.**

Дистрибутивний закон. Для матричних операцій справедливий дистрибутивний закон множення щодо додавання, який записується $\mathbf{A}(\mathbf{B}+\mathbf{C}) = \mathbf{AB}+\mathbf{AC}$ або $(\mathbf{B}+\mathbf{C})\mathbf{A} = \mathbf{BA}+\mathbf{CA}$. Елемент ij матриці $\mathbf{A}(\mathbf{B}+\mathbf{C})$ дорівнює:

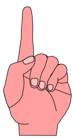
$$\sum_{k} a_{ik}\left(b_{kj} + c_{kj}\right) = \sum_{k} a_{ik} b_{kj} + \sum_{s=1} a_{ik} c_{kj} \quad ,$$

тобто дорівнює елементу ij матриці **AB** плюс елемент ij матриці **AC**.

Для матричних операцій справедливий також *дистрибутивний закон щодо множення на скаляр,* що записується так: $\lambda(\mathbf{A}+\mathbf{B}) = \lambda\mathbf{A}+\lambda\mathbf{B}$ або $(\lambda+\mu)\mathbf{A} = \lambda\mathbf{A}+\mu\mathbf{A}$.

Наведені закони важливі для правильного виконання послідовних дій над матрицями і для алгебраїчних перетворень матричних виразів.

Одинична, або тотожна матриця



Визначення 1.5. Квадратну матрицю n-го порядку

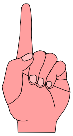
$$\mathbf{I}_n = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 \end{bmatrix},$$

на головній діагоналі якої розташовані одиниці, а на всіх інших місцях – нулі, називають *одиничною,* або *тотожною*

Множення одиничної матриці n–го порядку на матрицю А того ж порядку зліва або справа залишає останню без змін:

$$\mathbf{IA} = \mathbf{AI} = \mathbf{A} \; .$$

Матриця-скаляр



Визначення 1.6. Квадратну матрицю

$$\lambda\mathbf{I} = \begin{bmatrix} \lambda & 0 & \ldots & 0 \\ 0 & \lambda & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & \lambda \end{bmatrix},$$

в якої на головній діагоналі стоїть одне і теж число, а на всіх інших місцях – нулі, називають *матрицею-скаляром*.

Скалярну величину можна завжди перетворити в матрицю-скаляр, помноживши її на одиничну матрицю необхідного порядку. Множення матриці на скаляр, що було визначено раніше, аналогічно множенню матриць і може бути надано і як множення справа, і як множення зліва:

$$\lambda\mathbf{A} = (\lambda\mathbf{I})\mathbf{A} = \mathbf{A}(\lambda\mathbf{I}) = \mathbf{A}\lambda \ .$$

Наприклад,

$$\lambda\mathbf{A} = \lambda\mathbf{IA} = \begin{bmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} =$$

$$= \begin{bmatrix} \lambda a_{11} & \lambda a_{12} & \lambda a_{13} \\ \lambda a_{21} & \lambda a_{22} & \lambda a_{23} \\ \lambda a_{31} & \lambda a_{32} & \lambda a_{33} \end{bmatrix} = \mathbf{A}\lambda\mathbf{I} = \mathbf{A}\lambda \ .$$

Таким чином, одиничну матрицю можна додатково включити в матричний вираз або виключити з нього, не змінюючи значення цього виразу.

Діагональна матриця

Діагональна матриця має на головній діагоналі скалярні елементи, не обов'язково рівні один одному, і нулі в якості всіх інших елементів, тобто

$$\mathbf{A} = \lfloor a_{ij} \rfloor, \quad i, j = 1, 2, \ldots, n; \quad a_{ij} = 0, i \neq j \ ,$$

або $$\mathbf{A} = \begin{bmatrix} a_{11} & 0 & \ldots & 0 \\ 0 & a_{22} & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & a_{nn} \end{bmatrix} = \mathrm{diag}\left[a_{11} \quad a_{22} \quad \cdots \quad a_{nn}\right].$$

Матриця-скаляр являє собою окремий випадок діагональної матриці.

Транспонування матриці

Транспонування матриці $\mathbf{A}$ визначається як дія з перетворення матриці $\mathbf{A}$ в нову матрицю, рядками якого служать стовпці матриці $\mathbf{A}$, а стовпцями – рядки матриці $\mathbf{A}$**.** Таким чином, перший рядок матриці $\mathbf{A}$ стає першим стовпцем транспонованої матриці, другий рядок матриці $\mathbf{A}$ стає другим стовпцем і взагалі елемент ji транспонованої матриці є елемент ij вихідної матриці $\mathbf{A}$**.** Транспонована матриця позначається так само, як вихідна, але з додаванням спеціального верхнього індексу «т».

Наприклад,

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix}, \quad \mathbf{A}^{\text{т}} = \begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \\ a_{13} & a_{23} \end{bmatrix}.$$

Якщо $\mathbf{A}^{\text{т}} = \mathbf{A}$, то матриця називається *симетричною*. Очевидно, що симетрична матриця неодмінно повинна бути квадратною. Для неї має виконуватися рівність $a_{ij} = a_{ji}, \quad i, j = 1,2,\ldots,n$.

Нехай, як і в попередньому прикладі, $\mathbf{A}=\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix}$. Тоді

$$\mathbf{AA}^{\mathrm{T}}=\begin{bmatrix} a_{11}^2+a_{12}^2+a_{13}^2 & a_{11}a_{21}+a_{12}a_{22}+a_{13}a_{23} \\ a_{21}a_{11}+a_{22}a_{12}+a_{23}a_{13} & a_{21}^2+a_{22}^2+a_{23}^2 \end{bmatrix},\quad \text{і}$$

$$\mathbf{A}^{\mathrm{T}}\mathbf{A}=\begin{bmatrix} a_{11}^2+a_{12}^2 & a_{11}a_{12}+a_{21}a_{22} & a_{11}a_{13}+a_{21}a_{23} \\ a_{21}a_{11}+a_{22}a_{21} & a_{12}^2+a_{22}^2 & a_{12}a_{13}+a_{22}a_{23} \\ a_{13}a_{11}+a_{23}a_{21} & a_{13}a_{12}+a_{23}a_{22} & a_{13}^2+a_{23}^2 \end{bmatrix}.$$

Зазначимо, що обидві матриці $\mathbf{AA}^{\mathrm{T}}$ і $\mathbf{A}^{\mathrm{T}}\mathbf{A}$ симетричні. Оскільки $\mathbf{A}$ не є квадратною матрицею, ці добутки мають різний розмір. Однак суми елементів матриць $\mathbf{AA}^{\mathrm{T}}$ і $\mathbf{A}^{\mathrm{T}}\mathbf{A}$, що стоять на головній діагоналі (таку суму називають *слідом* матриці), рівні. Слід матриці $\mathbf{AA}^{\mathrm{T}}$, який позначають $\operatorname{tr}\left(\mathbf{AA}^{\mathrm{T}}\right)$, дорівнює сумі квадратів елементів матриці **A.**

Звернемо увагу на спеціальний випадок, коли $\mathbf{x}$ – вектор-стовпець, що складається з n елементів, а $\mathbf{x}^{\mathrm{T}}$ – вектор-рядок із тих же елементів. Тоді

$$\mathbf{x}^{\mathrm{T}}\mathbf{x}=\sum_{i=1}^{n}x_i^2 \text{ і. } \mathrm{xx}^{\mathrm{T}}=\begin{bmatrix} \mathrm{x}_1^2 & \mathrm{x}_1\mathrm{x}_2 & \cdots & \mathrm{x}_1\mathrm{x}_n \\ \mathrm{x}_2\mathrm{x}_1 & \mathrm{x}_1^2 & \cdots & \mathrm{x}_2\mathrm{x}_n \\ \cdots & \cdots & \cdots & \cdots \\ \mathrm{x}_n\mathrm{x}_1 & \mathrm{x}_n\mathrm{x}_2 & \cdots & \mathrm{x}_n^2 \end{bmatrix}$$

Теореми про транспоновані матриці

Теорема 1.1. *Двічі транспонована матриця дорівнює початковій*

$$\boxed{\left(\mathbf{A}^{\mathrm{T}}\right)^{\mathrm{T}}=\mathbf{A}}\,. \tag{1.1}$$

Теорема 1.2. *Транспонована сума матриць дорівнює сумі транспонованих матриць*

$$\boxed{\left(\mathbf{A}+\mathbf{B}\right)^{\mathrm{T}}=\mathbf{A}^{\mathrm{T}}+\mathbf{B}^{\mathrm{T}}}\,. \tag{1.2}$$

Теорема 1.3. *Транспонований добуток двох матриць дорівнює добутку транспонованої другої матриці на транспоновану першу*

$$\boxed{(\mathbf{AB})^{\mathrm{T}} = \mathbf{B}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}}\,. \tag{1.3}$$

Перші дві теореми очевидним способом випливають із визначення операції транспонування. Для доведення теореми 1.3 зазначимо, що елемент *ji* матриці $\mathbf{B}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}$ отриманий при множенні *j*–й рядка матриці $\mathbf{B}^{\mathrm{T}}$ на *i*–й стовпець матриці $\mathbf{A}^{\mathrm{T}}$, тобто він дорівнює добутку *j*–го стовпця матриці $\mathbf{B}$ на *i*–й рядок матриці $\mathbf{A}$, або дорівнює елементу *ji* матриці $\mathbf{AB}$.

Тоді можна сказати, що

$$(\mathbf{ABC})^{\mathrm{T}} = \mathbf{C}^{\mathrm{T}}\mathbf{B}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}\ .$$

Дійсно,

$$(\mathbf{ABC})^{\mathrm{T}} = [(\mathbf{AB})\mathbf{C}]^{\mathrm{T}} = \mathbf{C}^{\mathrm{T}}(\mathbf{AB})^{\mathrm{T}} = \mathbf{C}^{\mathrm{T}}\mathbf{B}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}\ .$$

Обернена матриця

Математична операція обертання застосовується тільки для квадратних матриць.

Визначення 1.7. *Оберненою* називають таку квадратну матрицю, яка, будучи помноженою на початкову, дає одиничну матрицю.

Якщо початкова матриця позначається як $\mathbf{A}$, то обернена – як $\mathbf{A}^{-1}$. Не кожна квадратна матриця має обернену, але якщо вона існує, то $\mathbf{AA}^{-1} = \mathbf{I}$. Наприклад, матриця $\mathbf{A}^{-1} = \begin{bmatrix} 1 & -1 \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}$ є оберненою для матриці $\mathbf{A} = \begin{bmatrix} 1 & 2 \\ -1 & 0 \end{bmatrix}$, оскільки їхній добуток дає одиничну матрицю $\mathbf{I} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$.

Комутативний закон для добутку матриць у загальному випадку не виконується, однак він має місце, як виняток, для добутку оберненої матриці на початкову $\mathbf{A}^{-1}\mathbf{A} = \mathbf{AA}^{-1}$.

Обертанням матриці називають таку математичну дію над матрицею $\mathbf{A}$, що приводить до одержання оберненої матриці $\mathbf{A}^{-1}$. Обертання матриці – складна операція, що більш детально буде розглянута в наступних розділах. Зазначимо тільки, що результатом обертання оберненої матриці є початкова матриця $\left(\mathbf{A}^{-1}\right)^{-1} = \mathbf{A}$, що очевидним способом випливає з визначення операції обертання.

1.3. Лінійні й квадратичні вирази

Лінійна функція і лінійна форма

З курсу алгебри відомо, що функція вигляду $y(x_1, x_2, \ldots, x_n) = \sum_{i=1}^{n} a_i x_i + a_0$ є *лінійною функцією*.

Лінійна функція складається з двох частин: *лінійної форми* $\sum_{i=1}^{n} a_i x_i$ і *вільного члена* a_0.

У матричному вигляді лінійна функція записується таким чином:

$$y(\mathbf{x}) = \mathbf{a}^{\mathrm{T}}\mathbf{x} + a_0 = \begin{bmatrix} a_1 \\ a_2 \\ \cdots \\ a_n \end{bmatrix} \begin{bmatrix} x_1 & x_2 & \ldots & x_n \end{bmatrix} + a_0 \,,$$

де $\mathbf{a}^{\mathrm{T}}\mathbf{x}$ – лінійна форма як скалярний добуток вектора-рядка коефіцієнтів функції і вектора-стовпця перемінних; a_0 – вільний член.

Приклад 1.2. Лінійна функція $y(x_1, x_2) = -4x_1 + 2x_2 - 3$ в матричному вигляді записується таким чином:

$$y(\mathbf{x}) = [-4 \quad 2]\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} - 3 \ .$$

Приклад 1.3. Лінійна функція $y(x_1, x_2, x_3) = x_1 - 2x_2 + 3x_3 + 4$ в матричному вигляді записується як

$$y(\mathbf{x}) = [1 - 2 \quad 3]\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + 4$$

Система лінійних рівнянь

Система лінійних рівнянь

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n - b_1 &= 0\ ; \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n - b_2 &= 0\ ; \\ \ldots \\ a_{m1}x_1 + a_{m12}x_2 + \ldots + a_{mn}x_n - b_m &= 0 \end{aligned} \qquad (1.4)$$

у матричному вигляді записується таким чином:

$$\mathrm{Ax - b = 0}, \quad \text{або} \quad \begin{bmatrix} \mathrm{a}_{11} & \mathrm{a}_{12} & \cdots & \mathrm{a}_{1n} \\ \mathrm{a}_{12} & \mathrm{a}_{22} & \cdots & \mathrm{a}_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ \mathrm{a}_{m1} & \mathrm{a}_{m2} & \cdots & \mathrm{a}_{mn} \end{bmatrix} \begin{bmatrix} \mathrm{x}_1 \\ \mathrm{x}_2 \\ \ldots \\ \mathrm{x}_n \end{bmatrix} - \begin{bmatrix} \mathrm{b}_1 \\ \mathrm{b}_2 \\ \ldots \\ \mathrm{b}_m \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \ldots \\ 0 \end{bmatrix} ,$$

де $\mathbf{A}$ – матриця коефіцієнтів розміру $m \times n$; $\mathbf{x}$ – вектор-стовпець перемінних, що складається з n елементів; $\mathbf{b}$ – вектор-стовпець вільних членів, що складається з m елементів.

Система лінійних рівнянь (1.4) може бути також записана у вигляді $\mathbf{Ax} = \mathbf{b}$.

Квадратична функція і квадратична форма

З курсу алгебри відомо, що функція $y(x_1, x_2, \ldots, x_n) = \sum_{i=1}^{n}\sum_{j=1}^{n} b_{ij} x_i x_j + \sum_{i=1}^{n} a_i x_i + a_0$ є *квадратичною функцією*.

Квадратична функція складається з двох частин: *квадратичної форми* $\sum_{i=1}^{n}\sum_{j=1}^{n} b_{ij} x_i x_j$ і лінійної функції $\sum_{i=1}^{n} a_i x_i + a_0$.

У матричному вигляді квадратична функція записується таким чином:

$$y(\mathbf{x}) = \frac{1}{2}\mathbf{x}^{\mathrm{т}}\mathbf{A}\,\mathbf{x} + \mathbf{a}^{\mathrm{т}}\mathbf{x} + a_0 \ ,$$

де $\frac{1}{2}\mathbf{x}^{\mathrm{т}}\mathbf{A}\,\mathbf{x}$ – квадратична форма як скалярний добуток константи $\frac{1}{2}$, вектора-рядка змінних $\mathbf{x}^{\mathrm{т}}$, симетричної матриці коефіцієнтів при квадратичних членах $\mathbf{A}$ розмірності *(n×n)* і вектора-стовпця $\mathbf{x}$; $(\mathbf{a}^{\mathrm{т}}\mathbf{x} + a_0)$ – лінійна функція.

Кожний діагональний елемент a_{ii} матриці **A** дорівнює відповідному подвоєному коефіцієнту b_{ii} квадратичної функції, а інші елементи матриці **A** в точності дорівнюють відповідним коефіцієнтам функції: $a_{ij} = b_{ij}$.

Квадратична форма може бути записана і без константи $\frac{1}{2}$, тобто як $\mathbf{x}^{\mathrm{т}}\mathbf{A}^{*}\,\mathbf{x}$. У цьому випадку значення всіх елементів матриці коефіцієнтів A^{*} повинні бути у два рази менше значень відповідних елементів матриці **A.**

Приклад 1.4. Квадратична функція $y(x_1,x_2)=x_1^2+5x_2^2-3x_1x_2-3x_1+5$ в матричному виді може бути записана як

$$y(\bar{x})=\frac{1}{2}\begin{bmatrix}x_1 & x_2\end{bmatrix}\begin{bmatrix}2 & -2\\ -2 & 10\end{bmatrix}\begin{bmatrix}x_1\\ x_2\end{bmatrix}+\begin{bmatrix}-3 & 0\end{bmatrix}\begin{bmatrix}x_1\\ x_2\end{bmatrix}+5$$

або

$$y(\bar{x})=\begin{bmatrix}x_1 & x_2\end{bmatrix}\begin{bmatrix}1 & -1\\ -1 & 5\end{bmatrix}\begin{bmatrix}x_1\\ x_2\end{bmatrix}+\begin{bmatrix}-3 & 0\end{bmatrix}\begin{bmatrix}x_1\\ x_2\end{bmatrix}+5\ .$$

Для перевірки матричного запису квадратичної функції треба знайти добуток $\mathbf{x}^{\mathrm{т}}\mathbf{A}\,\mathbf{x}$ (або $\mathbf{x}^{\mathrm{т}}\mathbf{A}^{*}\,\mathbf{x}$) і $\mathbf{a}^{\mathrm{т}}\mathbf{x}$ за правилами множення, що були розглянуті в підрозділі 1.2 і скласти їх з вільним членом a_0. Кінцевий результат повинен збігатися з алгебраїчним записом квадратичної функції.

Сума квадратів

Сума квадратів $x_1^2+x_2^2+\ldots+x_n^2$ може бути подана в матричному записі таким чином:

$$\sum_{i=1}^{n}x_i^2=\mathbf{x}^{\mathrm{т}}\mathbf{x}=\begin{bmatrix}x_1 & x_2 & \cdots & x_n\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ \cdots\\ x_n\end{bmatrix}.$$

Зважена сума квадратів

Алгебраїчний вираз вигляду $a_{11}x_1^2+a_{22}x_2^2+\ldots+a_{nn}x_n^2$ в матричному записі подається як

$$\sum_{i=1}^{n}a_{ii}x_i^2=\mathbf{x}^{\mathrm{т}}\mathbf{A}\mathbf{x}=\begin{bmatrix}x_1 & x_2 & \cdots & x_n\end{bmatrix}\begin{bmatrix}a_{11} & 0 & \cdots & 0\\ 0 & a_{22} & \cdots & 0\\ \cdots & \cdots & \cdots & \cdots\\ 0 & 0 & \cdots & a_{nn}\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ \cdots\\ x_n\end{bmatrix}.$$

Зважена сума квадратів є частковий випадок квадратичної форми.

1.4. Визначники

> Будь-якій квадратній матриці **A** відповідає деяке число, що називається *визначником* або *детермінантом* і позначається $\det \mathbf{A}$ або $|\mathbf{A}|$.

Детермінант матриці обчислюється підсумовуванням визначених добутків елементів матриці. Так, детермінант матриці другого порядку, тобто матриці розміру (2×2), визначається таким чином:

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21} \ ;$$

а детермінант матриця третього порядку –

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{21} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33} - a_{12}a_{21}a_{33} + a_{12}a_{23}a_{31} - \\ - a_{13}a_{22}a_{31} + a_{13}a_{21}a_{32} - a_{11}a_{23}a_{32}.$$

Визначник матриці першого порядку дорівнює єдиному елементу матриці $|a_{11}| = a_{11}$.

Відзначимо ряд особливостей цих визначників:

- Кожний член у правій частині є добуток такої ж кількості елементів, який порядок матриці, а саме: одного елемента для матриці порядку 1; двох елементів для матриці порядку 2 і трьох елементів для матриці порядку 3. У загальному випадку для матриці порядку n кожний член алгебраїчної суми містить n співмножників.
- Кожний член алгебраїчної суми містить як співмножник один і тільки один елемент з кожного рядка і один і тільки один елемент з кожного стовпця матриці. Таким чином, жодний елемент не зустрічається двічі в тому самому добутку.
- Кількість членів (добутків) у виразі, що розкриває визначник матриці n–го порядку, дорівнює $n! = (n\ \text{-}1)(\ n\ \text{-}2)\ldots 3\cdot 2\cdot 1$... Так, для визначника

матриці першого порядку це число дорівнює $1!=1$, для другого порядку – $2!=2$, третього – $3!=6$ і т.д.

- Кожний член алгебраїчної суми містить елементи, розташовані в матриці так, що їхні перші індекси утворюють частину натурального ряду чисел $1,2,3,\ldots,n$.. Другі індекси являють собою різні перестановки з цих елементів. Оскільки таких перестановок із n різних об'єктів існує $n!$, утворюється $n!$ різних членів, як уже було відзначено в попередньому пункті.

- Половина всіх членів додатна, а половина від’ємна. Знак залежить від розташування других індексів. Говорять, що в розташуванні натуральних чисел виникає інверсія природного порядку, якщо серед двох натуральних чисел більше передує меншому. Кількість інверсій у перестановці n натуральних чисел є число пар елементів, не обов'язково сусідніх, у яких більше число передує меншому. Перестановка називається парною, якщо число інверсій у ній парне, і непарною, якщо число інверсій непарне. Парним перестановкам відповідає знак плюс, а непарним – мінус. Наприклад, другий член у поданні визначника третього порядку складається з елементів, другі індекси яких розташовані в порядку 2–1–3 , де є тільки одна інверсія, тому що число 2 стоїть раніше числа 1. Отже, ця перестановка непарна, і перед другим членом повинен стояти від’ємний знак. У наступного члена другі індекси розташовані в порядку 2–3–1, де є дві інверсії, тому що 2 стоїть раніше 1 і 3 розташовано також раніше 1. Отже, цей член повинний бути взятий із знаком плюс.

Ці п'ять особливостей приводять до загального визначення детермінанту матриці $\mathbf{A}$ n–го порядку як

$$\boxed{|\mathbf{A}| = \sum\left(\pm\, a_{1i}a_{2j}\ldots a_{nr}\right)},$$

де сума береться по всіх перестановках других індексів співмножників, причому зі знаком плюс беруться члени з парними перестановками, а зі знаком мінус – члени з непарними перестановками.

Властивості визначників

Властивість 1. *Детермінант матриці збігається з детермінантом транспонованої матриці*

$$|\mathbf{A}| = |\mathbf{A}^{\text{т}}|, \tag{1.5}$$

тобто операція транспонування матриці не змінює значення її визначника.

Для доведення (1.5) вважатимемо $\mathbf{B} = \mathbf{A}^{\text{т}}$. Виразимо визначник матриці **B** через визначник матриці **A.** Проілюструємо доведення на прикладі матриця третього порядку. У цьому випадку член $b_{13}b_{21}b_{32}$ у розкладанні визначника матриці **B** утвориться як добуток трьох елементів a_{31}, a_{12} і a_{23} матриці **A.** Знак цього члена у визначнику матриці **B** визначається кількістю інверсій у перестановці 3-1-2 (другі індекси при елементах) і тому додатний. Знак члена $a_{31}a_{12}a_{23}$ у визначнику матриці **A** можна встановити, якщо розташувати співмножники в порядку зростання перших індексів. Однак таке перегрупування приведе в точності до того ж числа інверсій у других індексах, що було на початку для перших індексів. Наприклад, якщо перший елемент перемістити на третє місце, то дві інверсії в перших індексах пропадуть, оскільки число 3 буде стояти після 1 і після 2, однак виникнуть дві інверсії в других індексах, тому що число 1 стоятиме праворуч від чисел 2 і 3. Таким чином, $a_{31}a_{12}a_{23}$ буде мати той же знак у розкладанні визначника матриці **A**, що й у розкладанні визначника матриці **B.** Ці міркування можна навести для будь-якого члена розкладання, що і доводить властивість (1.5).

Властивість 2. *Перестановка будь-яких стовпців (або рядків) матриці* **A** *змінює знак її визначника.*

Для доведення цієї властивості зазначимо, що транспозиція будь-яких елементів переводить перестановку з парного класу в непарний або навпаки. Припустимо, наприклад, що поміняли місцями числа j і k, між якими розташовані m інших чисел (рис.1.1). Щоб досягти цього, можна перемістити число k на m місць уперед, а число j – на $m+1$ місць назад. Коли міняються місцями два сусідніх елементи, загальне число інверсій або збільшується на одиницю, або зменшується на одиницю. Отже, здійснюючи транспозицію двох чисел j і k, відокремлених одне від одного m іншими числами, ми змінимо число інверсій $2m+1$ разів. Оскільки це число непарне, то перестановка належить протилежному класові.

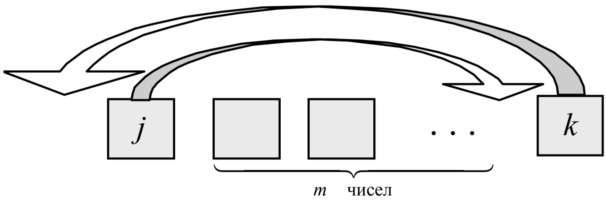


Рис.1.1.

Отже, у результаті перестановки будь-яких стовпців матриці $\mathbf{A}$ перші індекси у всіх елементах $|\mathbf{A}|$ залишаться незмінними, а знак кожного елемента розкладання зміниться на протилежний, тобто зміниться знак визначника $|\mathbf{A}|$.

Нехай $\mathbf{B}$ позначає матрицю, отриману з $\mathbf{A}$ перестановкою j-го і k-го стовпців. Уже було показано, що $|\mathbf{B}| = -|\mathbf{A}|$. І нехай тепер $|\mathbf{B}|^{\text{т}}$ – матриця, отримана з транспонованої матриці $\mathbf{A}^{\text{т}}$ перестановкою рядків j і k. Тоді за властивістю (1.5)

$$\left|\mathbf{B}^{\text{т}}\right| = |\mathbf{A}| .$$

Відповідно до другої властивості

$$|\mathbf{B}| = -|\mathbf{A}| .$$

І знову внаслідок (1.5)

$$-|\mathbf{A}| = -\left|\mathbf{A}^{\text{т}}\right| ,$$

тобто

$$\left|\mathbf{B}^{\text{т}}\right| = -\left|\mathbf{A}^{\text{т}}\right|.$$

Отже, перестановка будь-яких двох стовпців (або рядків) матриці змінює знак її визначника.

Властивість 3. *Визначник матриці з двома рівними рядками (або стовпцями) дорівнює нулю.*

Ця властивість виходить з другої властивості, оскільки перестановка двох однакових рядків (або) стовпців, з одного боку, приводить до визначника, рівного $-|\mathbf{A}|$, а з другого – залишає визначник незмінним. Зазначені суперечності перестають бути такими тільки за умови, що $|\mathbf{A}| = 0$.

Властивість 4. *Множення кожного елемента рядка (або стовпця) матриці* $\boldsymbol{A}$ *на число* λ *приводить до виразу*

$$\lambda|\mathbf{A}| \ . \tag{1.6}$$

Властивість 5. Множення кожного елемента матриці $\mathbf{A}$ на число λ приводить до виразу

$$|\lambda\mathbf{A}| = \lambda^n |\mathbf{A}|. \tag{1.7}$$

Ці дві властивості безпосередньо випливають з того факту, що кожний член у розкладанні визначника містить тільки один елемент кожного рядка (або стовпця) матриці $\mathbf{A}$.

Якщо ми розглянемо $|\mathbf{A}| = \sum\left(\pm\, a_{1\alpha} a_{2\beta} \ldots a_{n\nu}\right)$ і зберемо разом усі члени розкладання, що містять елемент a_{11}, то коефіцієнтом при a_{11} буде

$$|\mathbf{A}| = \sum\left(\pm\, a_{2\beta} a_{3\gamma} \ldots a_{n\nu}\right) , \tag{1.8}$$

де підсумовування робиться по всіх $(n\text{-}1)!$ чисел $2,3,\ldots,n,$. Оскільки розглянутий елемент a_{11} знаходиться в рядку і стовпці з номерами 1 і 1, то його додавання як співмножника до будь-якого члена (1.8) не впливає на знак цього члена. Тому цей знак визначається перестановкою $(\beta, \lambda, \ldots, \nu)$ натуральних чисел $2,\ 3,\ \ldots\ ,\ n$. Таким чином, сума $|\mathbf{A}| = \sum\left(\pm\, a_{2\beta} a_{3\gamma} \ldots a_{n\nu}\right)$ є визначник порядку $n-1$

$$\begin{vmatrix} a_{22} & a_{23} & \cdots & a_{2n} \\ a_{32} & a_{33} & \cdots & a_{3n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n2} & a_{n3} & \cdots & a_{nn} \end{vmatrix} \tag{1.9}$$

матриці, отриманої з матриці **A** викреслюванням з неї першого рядка і першого стовпця.

Мінор

Отриманий визначник (1.9) називається мінором. Позначають його $\left|\mathbf{A}_{11}\right|$. У загальному випадку

$$\left|\mathbf{A}_{ij}\right| = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1,j-1} & a_{1,j+1} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2,j-1} & a_{2,j+1} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{i-1,1} & a_{i-1,2} & \cdots & a_{i-1,j-1} & a_{i-1,j+1} & \cdots & a_{i-1,n} \\ a_{i+1,1} & a_{i+1,2} & \cdots & a_{i+1,j-1} & a_{i+1,j+1} & \cdots & a_{i+1,n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{n,J-1} & a_{n,j+1} & \cdots & a_{nn} \end{vmatrix}.$$

Раніше було показано, що сума всіх членів розкладання визначника $\left|\mathbf{A}\right|$, що містять як співмножник елемент a_{11}, може бути записана у вигляді $a_{11}\left|\mathbf{A}_{11}\right|$.

Виникає запитання, чому дорівнює сума всіх членів розкладання визначника $\left|\mathbf{A}\right|$, що містять як співмножник елемент a_{ij}.

Розглянемо матрицю **B**, що утворюється з матриці **A** таким чином. Спочатку міняються місцями рядок з номером i і попередній йому рядок матриці **A,** після чого продовжується процедура транспозиції рядків доти, поки i–й рядок матриці **A** не стане першим. Усі інші рядки зберігають свій звичайний порядок. Потім міняють місцями стовпець з номером j і попередній йому стовпець матриці **A.** Процедуру транспозиції стовпців продовжують доти, поки j–й стовпець не займе місце першого стовпця. Всього буде здійснено $i+j-2$ транспозицій, і детермінант результуючої матриці **B** буде визначений таким чином:

$$\left|\mathbf{B}\right| = \left(-1\right)^{i+j-2}\left|\mathbf{A}\right| = \left(-1\right)^{i+j}\left|\mathbf{A}\right|. \tag{1.10}$$

При цьому $b_{11} = a_{ij}$ і $\left|\mathbf{B}_{11}\right| = \left|\mathbf{A}_{ij}\right|$, тоді

$$\left|\mathbf{A}_{ij}\right| = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1,j-1} & a_{1,j+1} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2,j-1} & a_{2,j+1} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{i-1,1} & a_{i-1,2} & \cdots & a_{i-1,j-1} & a_{i-1,j+1} & \cdots & a_{i-1,n} \\ a_{i+1,1} & a_{i+1,2} & \cdots & a_{i+1,j-1} & a_{i+1,j+1} & \cdots & a_{i+1,n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{n,J-1} & a_{n,j+1} & \cdots & a_{nn} \end{vmatrix}.$$

Отже, сума всіх членів розкладання визначника $\left|\mathbf{B}\right|$, що містять як співмножник елемент a_{ij}, рівний b_{11}, обчислюється як $b_{11}\left|\mathbf{B}_{11}\right|$. Відповідно до співвідношення (1.10) множення членів розкладання визначника $\left|\mathbf{B}\right|$ на $(-1)^{i+j}$ перетворює їх у члени розкладання визначника $\left|\mathbf{A}\right|$. Таким чином, сума всіх членів розкладання визначника $\left|\mathbf{A}\right|$, що містять як співмножник елемент a_{ij}, дорівнює $(-1)^{i+j} a_{ij}\left|\mathbf{A}_{ij}\right|$.

Алгебраїчне доповнення

Назвемо коефіцієнт при a_{ij} в попередньому виразі *алгебраїчним доповненням* елемента a_{ij} і будемо позначати його

$$c_{ij} = (-1)^{i+j} a_{ij}\left|\mathbf{A}_{ij}\right|, \tag{1.11}$$

так що $c_{11} = \left|\mathbf{A}_{11}\right|$; $c_{12} = -\left|\mathbf{A}_{12}\right|$; $c_{22} = \left|\mathbf{A}_{22}\right|$ і т.д.

Розглянемо тепер вираз $a_{i1}c_{i1} + a_{i2}c_{i2} + \ldots + a_{in}c_{in}$. Він містить $(n-1)!\cdot n = n!$ членів розкладання визначника $\left|\mathbf{A}\right|$. Усі члени різні і мають ті ж знаки, що й у розкладанні визначника $|\mathbf{A}|$. Але розкладання визначника $|\mathbf{A}|$ має всього $n!$ різних членів. Отже,

$$|\mathbf{A}| = a_{i1}c_{i1} + a_{i2}c_{i2} + \ldots + a_{in}c_{in} \; . \tag{1.12}$$

Формула (1.12) є розкладанням визначника $|\mathbf{A}|$ по елементах i–го рядка, кожний з яких помножений на своє алгебраїчне доповнення. Аналогічним способом визначник $|\mathbf{A}|$ може бути розкладений по елементах будь-якого *рядка* (або *стовпця*).

Властивість 6. *Значення визначника не зміниться, якщо до будь-якого його рядка (або стовпця) додати будь-який інший рядок (або стовпець), помножений на постійний множник.*

Доведення. Розглянемо визначник

$$\begin{vmatrix} a_{11}+\lambda\alpha & a_{12}+\lambda\beta & a_{13}+\lambda\gamma \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} .$$

Скористаємося співвідношенням (1.12) і розкладемо розглянутий визначник по елементах першого рядка:

$$\Delta = (a_{11}+\lambda\alpha)\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - (a_{12}+\lambda\beta)\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + (a_{13}+\lambda\gamma)\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} =$$

$$= \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} + \lambda \begin{vmatrix} \alpha & \beta & \gamma \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} .$$

Якщо $[\alpha \quad \beta \quad \gamma] = [a_{21} \quad a_{22} \quad a_{23}]$ або $[\alpha \quad \beta \quad \gamma] = [a_{31} \quad a_{32} \quad a_{33}]$, то останній визначник у правій частині рівності перетворюється в нуль, оскільки в нього з'являються два однакові рядки.

Таким чином, права частина рівності доводиться до значення початкового визначника.

Властивістю 6 часто користуються при обчисленні визначників.

Приклад 1.5. Обчислити $|\mathbf{A}| = \begin{vmatrix} 2 & 4 & 6 \\ 3 & 2 & 3 \\ 1 & 4 & 9 \end{vmatrix}$.

Розв'язання. Розкладемо визначник по елементах першого рядка:

$$|\mathbf{A}| = 2\begin{vmatrix} 2 & 3 \\ 4 & 9 \end{vmatrix} - 4\begin{vmatrix} 3 & 3 \\ 1 & 9 \end{vmatrix} + 6\begin{vmatrix} 3 & 2 \\ 1 & 4 \end{vmatrix} = 2(6) - 4(24) + 6(10) = -24 \ .$$

Можна скористатися другим способом: відняти з першого рядка подвоєний другий. Тоді $|\mathbf{A}| = \begin{vmatrix} -4 & 0 & 0 \\ 3 & 2 & 3 \\ 1 & 4 & 9 \end{vmatrix}$. Розкладання цього визначника по елементах першого рядка дає $|\mathbf{A}| = -4 \cdot 6 = -24$.

Із властивості 6 випливає, що значення визначника не зміниться, якщо до будь-якого його рядка (або стовпця) додати будь-яку лінійну комбінацію інших рядків (або стовпців) цього визначника.

Властивість *7. Розкладання по елементах рядка (або стовпця) із коефіцієнтами, що є алгебраїчними доповненнями іншого рядка (або стовпця), дорівнює нулю.*

Утворимо формально вираз

$$a_{i1}c_{j1} + a_{i2}c_{j2} + \ldots + a_{in}c_{jn} \ , \quad i \neq j \ ,$$

в якому елементи належать рядку з номером i, а алгебраїчні доповнення відповідають елементам рядка j. Такий вираз може бути отриманий для визначника матриці, в котрого рядки i і j рівні. Отже, визначник цієї матриці в наслідок властивості 3 дорівнює нулю.

Таким чином,

$$\sum_{r=1}^{n} a_{ir} c_{jr} = 0\,, \quad i \neq j\,;$$
$$\sum_{r=1}^{n} a_{ir} c_{jr} = 0\,, \quad i \neq j\,. \tag{1.13}$$

Обернена матриця

В алгебрі чисел для $x \neq 0$ справедливо співвідношення $xx^{-1} = x^{-1}x = 1$ (одиниця). Аналогічно в матричному численні має місце співвідношення $\mathbf{AA}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$ (одинична матриця).

Якщо матриця $\mathbf{A}^{-1}$ існує, її називають оберненою матрицею щодо початкової матриці $\mathbf{A}$ (див. операцію обертання матриці).

Обернену матрицю можна отримати за допомогою такої процедури. З матриці $\mathbf{A}$ порядку n формують нову матрицю, в котрої кожний елемент a_{ij} замінений його алгебраїчним доповненням c_{ij}. Цю нову матрицю транспонують. Отриману таким чином матрицю *називають приєднаною матрицею* $\mathbf{A}$ і позначають:

$$(\text{aid } \mathbf{A}) = \begin{bmatrix} c_{11} & c_{21} & \cdots & c_{n1} \\ c_{12} & c_{22} & \cdots & c_{n2} \\ \cdots & \cdots & \cdots & \cdots \\ c_{1n} & c_{2n} & \cdots & c_{nn} \end{bmatrix}.$$

За допомогою властивості 7 можна переконатися, що

$$\mathbf{A}(\text{aid } \mathbf{A}) = (\text{aid } \mathbf{A})\mathbf{A} = \begin{bmatrix} |\mathbf{A}| & 0 & \dots & 0 \\ 0 & |\mathbf{A}| & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & |\mathbf{A}| \end{bmatrix} = |\mathbf{A}|\mathbf{I}\,.$$

Визначимо матрицю $\mathbf{A}^{-1}$ таким чином:

$$\mathbf{A}^{-1} = \frac{1}{|\mathbf{A}|}(\text{aid } \mathbf{A}) = \begin{bmatrix} \frac{c_{11}}{|\mathbf{A}|} & \frac{c_{21}}{|\mathbf{A}|} & \cdots & \frac{c_{n1}}{|\mathbf{A}|} \\ \frac{c_{12}}{|\mathbf{A}|} & \frac{c_{22}}{|\mathbf{A}|} & \cdots & \frac{c_{n2}}{|\mathbf{A}|} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{c_{1n}}{|\mathbf{A}|} & \frac{c_{2n}}{|\mathbf{A}|} & \cdots & \frac{c_{nn}}{|\mathbf{A}|} \end{bmatrix}. \tag{1.14}$$

Остання операція можлива тільки при $|\mathbf{A}| \neq 0$. У цьому випадку матриця **A** називається несингулярною або невиродженою. Якщо ж $|\mathbf{A}| = 0$, то матриця **A** називається сингулярною або виродженою.

Матриця $\mathbf{A}^{-1}$ єдина. Справді, припустивши, що існує матриця **B** така, що **AB**=**I** , одержимо

$$\mathbf{A}^{-1} = \mathbf{A}^{-1}\mathbf{I} = \mathbf{A}^{-1}\mathbf{AB} = \mathbf{B} .$$

Отже, якщо існує матриця **C**, що задовольняє умові **CA=I,** то $\mathbf{C} = \mathbf{A}^{-1}$.

Приклад 1.6. Обчислити обернену до **A** матрицю:

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 3 & 5 \\ 1 & 5 & 12 \end{bmatrix}.$$

Замінимо елементи заданої матриці їхніми алгебраїчними доповненнями :

$$\begin{bmatrix} \begin{vmatrix} 3 & 5 \\ 5 & 12 \end{vmatrix} & -\begin{vmatrix} 1 & 5 \\ 1 & 12 \end{vmatrix} & \begin{vmatrix} 1 & 3 \\ 1 & 5 \end{vmatrix} \\ -\begin{vmatrix} 2 & 3 \\ 5 & 12 \end{vmatrix} & \begin{vmatrix} 1 & 3 \\ 1 & 12 \end{vmatrix} & -\begin{vmatrix} 1 & 2 \\ 1 & 5 \end{vmatrix} \\ \begin{vmatrix} 2 & 3 \\ 3 & 5 \end{vmatrix} & -\begin{vmatrix} 1 & 3 \\ 1 & 5 \end{vmatrix} & \begin{vmatrix} 1 & 2 \\ 1 & 3 \end{vmatrix} \end{bmatrix} = \begin{bmatrix} 11 & -7 & 2 \\ -9 & 9 & -3 \\ 1 & -2 & 1 \end{bmatrix}.$$

Отже, $(\text{adj } \mathbf{A}) = \begin{bmatrix} 11 & -9 & 1 \\ -7 & 9 & -2 \\ 2 & -3 & 1 \end{bmatrix}$.

Обчислимо визначник $|\mathbf{A}|$. Для цього віднімемо з другого і третього його рядків перший: $|\mathbf{A}| = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 3 & 9 \end{bmatrix} = 3 \cdot$

Відповідно до (1.14) маємо

$$\mathbf{A}^{-1} = \begin{bmatrix} \frac{11}{3} & -3 & \frac{1}{3} \\ -\frac{7}{3} & 3 & -\frac{2}{3} \\ \frac{2}{3} & -1 & \frac{1}{3} \end{bmatrix}.$$

Перевірка: $\mathbf{AA}^{-1} = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 3 & 5 \\ 1 & 5 & 12 \end{bmatrix} \begin{bmatrix} \frac{11}{3} & -3 & \frac{1}{3} \\ -\frac{7}{3} & 3 & -\frac{2}{3} \\ \frac{2}{3} & -1 & \frac{1}{3} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$.

Приклад 1.7. *Статична модель витрати-випуск Леонтьєва.* Нехай a_{ij} - витрати продукту i на одиницю випуску продукту j; X_i – загальний випуск продукту i; C_i – кінцевий попит на продукт i. Розглядається модель для двох продуктів:

$$a_{11}X_1 + a_{12}X_2 + C_1 = X_1 \;;$$
$$a_{21}X_1 + a_{22}X_2 + C_2 = X_2$$

у припущенні, що немає ні втрат, ні надлишків продуктів. Після приведення подібних членів одержимо.

$$\begin{aligned} &(1 - a_{11})X_1 - a_{12}X_2 = C_1 \;; \\ &- a_{21}X_1 + (1 - a_{22})X_2 = C_2, \end{aligned} \qquad (1.15)$$

або $\mathbf{Ax} = \mathbf{c}$, де $\mathbf{A} = \begin{bmatrix} 1 - a_{11} & -a_{12} \\ -a_{21} & 1 - a_{22} \end{bmatrix}$; $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$; $\mathbf{c} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$.

Якщо розв'язати систему рівнянь (1.15) щодо **x**, то знайдемо, яка загальна кількість продуктів X_1 і X_2 повинна бути зроблена, щоб забезпечити виробниче споживання і задовольнити кінцевий попит C_1 і C_2:

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{c} \ ,$$

де $\mathbf{A}^{-1} = \dfrac{1}{(1 - a_{11})(1 - a_{22}) - a_{12}a_{21}} \begin{bmatrix} 1 - a_{22} & -a_{21} \\ -a_{12} & 1 - a_{11} \end{bmatrix}$, так що

$$X_1 = \frac{1 - a_{22}}{\Delta} C_1 + \frac{a_{21}}{\Delta} C_2;$$

$$X_2 = \frac{a_{12}}{\Delta} C_1 + \frac{1 - a_{11}}{\Delta} C_2,$$

де $\Delta = (1 - a_{11})(1 - a_{22}) - a_{12}a_{21}$.

Правило Крамера. *Якщо* $\mathbf{Ax} = \mathbf{h}$, *де* $\boldsymbol{A}$ *– невироджена матриця, то множення обох частин зліва на* $\mathbf{A}^{-1}$ *приводить до* $\mathbf{x} = \mathbf{A}^{-1}\mathbf{h}$.

Для випадку матриці $\mathbf{A}$ третього порядку маємо:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \frac{1}{|\mathbf{A}|} \begin{bmatrix} c_{11} & c_{21} & c_{31} \\ c_{12} & c_{22} & c_{32} \\ c_{13} & c_{23} & c_{33} \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \\ h_3 \end{bmatrix},$$

де c_{ij} – алгебраїчні доповнення. Отже,

$$x_1 = \frac{h_1c_{11} + h_2c_{21} + h_3c_{31}}{|\mathbf{A}|} \ . \tag{1.16}$$

Чисельником виразу (1.16) служить визначник, розкладений по елементах першого стовпця, рівного $\begin{bmatrix} h_1 \\ h_2 \\ h_3 \end{bmatrix}$. Два інших стовпців цього визначника співпадають з другим і третім стовпцями матриці **A.**

У загальному випадку, щоб у системі лінійних рівнянь визначити x_i, треба скласти відношення двох визначників, перший з яких – визначник матриці коефіцієнтів, у якому *i*–й стовпець замінений вектором $\mathbf{h}$, а другий – визначник матриці коефіцієнтів системи:

$$x_1 = \frac{\begin{bmatrix} h_1 & a_{12} & a_{13} \\ h_2 & a_{22} & a_{23} \\ h_3 & a_{32} & a_{33} \end{bmatrix}}{|\mathbf{A}|}; \quad x_2 = \frac{\begin{bmatrix} a_{11} & h_1 & a_{13} \\ a_{21} & h_2 & a_{23} \\ a_{31} & h_2 & a_{33} \end{bmatrix}}{|\mathbf{A}|}; \quad x_3 = \frac{\begin{bmatrix} a_{11} & a_{12} & h_1 \\ a_{21} & a_{22} & h_2 \\ a_{31} & a_{32} & h_3 \end{bmatrix}}{|\mathbf{A}|}.$$

Властивості оберненої матриці

Теорема 1.4. *Обернена матриця добутку двох матриць* $\mathbf{A}$ *і* $\mathbf{B}$ *дорівнює добутку оберненої матриці* $\mathbf{B}$ *на обернену матрицю* $\mathbf{A}$

$$\boxed{(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}}. \qquad (1.17)$$

Доведення.

$$(\mathbf{AB})\left(\mathbf{B}^{-1}\mathbf{A}^{-1}\right) = \mathbf{A}\left(\mathbf{BB}^{-1}\right)\mathbf{A}^{-1} = \mathbf{AIA}^{-1} = \mathbf{I}\ .$$

Аналогічно

$$\left(\mathbf{B}^{-1}\mathbf{A}^{-1}\right)(\mathbf{AB}) = \mathbf{I}\ .$$

Легко узагальнити цей результат і одержати

$$\boxed{(\mathbf{ABC})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}\mathbf{C}^{-1}}. \qquad (1.18)$$

За визначенням оберненої матриці

$$\left(\mathbf{A}^{-1}\right)\left(\mathbf{A}^{-1}\right)^{-1} = \mathbf{I}\ . \tag{1.19}$$

Помножимо обидві частини (1.19) зліва на **A**:

$$\mathbf{A}\left(\mathbf{A}^{-1}\right)\left(\mathbf{A}^{-1}\right)^{-1} = \mathbf{AI}\ ;$$

$$\mathbf{I}\left(\mathbf{A}^{-1}\right)^{-1} = \mathbf{A}\ ;$$

$$\boxed{\left(\mathbf{A}^{-1}\right)^{-1} = \mathbf{A}} \tag{1.20}$$

Таким чином, *зворотна матриця щодо оберненої відтворює початкову матрицю*.

Теорема 1.5. *Обернена матриця до транспонованої дорівнює транспонованої від оберненої*

$$\boxed{\left(\mathbf{A}^{\mathrm{т}}\right)^{-1} = \left(\mathbf{A}^{-1}\right)^{\mathrm{т}}}, \tag{1.21}$$

Доведення. Транспонуємо рівність $\mathbf{AA}^{-1} = \mathbf{I}$. Одержимо $\left(\mathbf{A}^{-1}\right)^{\mathrm{т}}\mathbf{A}^{\mathrm{т}} = \mathbf{I}$.

Помножимо справа на $\left(\mathbf{A}^{\mathrm{т}}\right)^{-1}$:

$$\left(\mathbf{A}^{-1}\right)^{\mathrm{т}}\mathbf{A}^{\mathrm{т}}\left(\mathbf{A}^{\mathrm{т}}\right)^{-1} = \left(\mathbf{A}^{\mathrm{т}}\right)^{-1}\ .$$

Отже,

$$\left(\mathbf{A}^{-1}\right)^{\mathrm{т}} = \left(\mathbf{A}^{\mathrm{т}}\right)^{-1}\ .$$

Маємо, нарешті, що *визначник оберненої матриці дорівнює оберненому визначнику початкової матриці*

$$\boxed{\left|\mathbf{A}^{-1}\right| = \frac{1}{|\mathbf{A}|}}\ . \tag{1.22}$$

Оскільки $\mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$, то $|\mathbf{I}| = \mathbf{I} = \left|\mathbf{A}^{-1}\mathbf{A}\right| = \left|\mathbf{A}^{-1}\right|\cdot|\mathbf{A}|$. Отже, $\left|\mathbf{A}^{-1}\right| = \dfrac{1}{|\mathbf{A}|}$.

Одним з найважливіших результатів теорії визначників є властивість, що визначник добутку матриць **A** і **B** дорівнює добутку визначників **A** і **B**

$$\boxed{|\mathbf{AB}| = |\mathbf{A}|\cdot|\mathbf{B}|}\ . \tag{1.23}$$

1.5. Розбивка матриць

Оскільки матриця являє собою прямокутну таблицю елементів, то її можна розділити за допомогою вертикальної і горизонтальної ліній на менші підматриці Наприклад, матриця

$$\mathbf{A} = \left[\begin{array}{ccc|c} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ \hline a_{31} & a_{32} & a_{33} & a_{34} \end{array}\right] \tag{1.24}$$

може бути розбита двома проведеними усередині матриці лініями на чотири підматриці:

$$\mathbf{A}_{11} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix}; \quad \mathbf{A}_{12} = \begin{bmatrix} a_{14} \\ a_{24} \end{bmatrix};$$
$$\mathbf{A}_{21} = \begin{bmatrix} a_{31} & a_{32} & a_{33} \end{bmatrix}; \quad \mathbf{A}_{22} = \begin{bmatrix} a_{34} \end{bmatrix}. \tag{1.25}$$

і матриця $\mathbf{A}$ може бути записана у виді $\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix}$.

Лінії розбивки можуть проходити як усередині початковій матриці, так і над (під) ній.

Основні операції додавання і множення поширюються також і на розбиті матриці, якщо процес розбивки здійснений відповідним способом. Наприклад, якщо розбивка матриці $\mathbf{B}$ розміру 3×4 має вид:

$$\mathbf{B} = \begin{bmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \mathbf{B}_{21} & \mathbf{B}_{22} \end{bmatrix},$$

де матриці $\mathbf{B}_{ij}$ того ж порядку, що і матриці $\mathbf{A}_{ij}$, то сума $\mathbf{A}+\mathbf{B}$ може бути записана як $\mathbf{A} + \mathbf{B} = \begin{bmatrix} \mathbf{A}_{11} + \mathbf{B}_{11} & \mathbf{A}_{12} + \mathbf{B}_{12} \\ \mathbf{A}_{21} + \mathbf{B}_{21} & \mathbf{A}_{22} + \mathbf{B}_{22} \end{bmatrix}$.

Розглянемо операцію множення на прикладі множення матриць **A** і **B,** якщо матриця **A** визначається як (1.24), а $\mathbf{B}=\begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \\ b_{41} & b_{42} \end{bmatrix}$. Добуток **AB** існує**,** при цьому розмір результуючої матриці дорівнює 3×2.

Щоб дати вираз добутку через матриці, отримані при розбивці, необхідно виконання єдиної умови – *розбивка рядків матриці* **B** *повинна відповідати розбивці стовпців матриці* **A**.

Наприклад, нехай матриця **A** має розбивку (1.25), а розбивка матриці **B** має вид:

$$\mathbf{B}=\begin{bmatrix} \mathbf{B}_{11} \\ \mathbf{B}_{21} \end{bmatrix}, \text{ де } \mathbf{B}_{11}=\begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \end{bmatrix}; \quad \mathbf{B}_{21}=\begin{bmatrix} b_{41} & b_{42} \end{bmatrix}.$$

Оперуючи підматрицями як звичайними елементами, одержимо

$$\mathbf{AB}=\begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix}\begin{bmatrix} \mathbf{B}_{11} \\ \mathbf{B}_{21} \end{bmatrix}=\begin{bmatrix} \mathbf{A}_{11}\mathbf{B}_{11}+\mathbf{A}_{12}\mathbf{B}_{21} \\ \mathbf{A}_{21}\mathbf{B}_{11}+\mathbf{A}_{22}\mathbf{B}_{21} \end{bmatrix}. \qquad (1.26)$$

Перегляд елементів у добутках матриць показує, що (1.26) дає той же результат, що безпосередньо утворюється при множенні матриць без розбивки. Отже, підматриці можна сприймати як звичайні елементи матриць, розбитих відповідним способом.

Умова відповідної розбивки матриці **B**, що дозволяє обчислити **AB**, буде задоволена також і при $\mathbf{B}=\begin{bmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \mathbf{B}_{21} & \mathbf{B}_{22} \end{bmatrix}$, де $\mathbf{B}_{11}=\begin{bmatrix} b_{11} \\ b_{21} \\ b_{31} \end{bmatrix}$;

$$\mathbf{B}_{12} = \begin{bmatrix} b_{12} \\ b_{22} \\ b_{32} \end{bmatrix}; \ \mathbf{B}_{21} = [b_{41}]; \ \mathbf{B}_{22} = [b_{42}].$$

Тоді добуток, виражений через підматриці, прийме вид:

$$\mathbf{AB} = \begin{bmatrix} \mathbf{A}_{11}\mathbf{B}_{11} + \mathbf{A}_{12}\mathbf{B}_{21} & \mathbf{A}_{11}\mathbf{B}_{12} + \mathbf{A}_{12}\mathbf{B}_{22} \\ \mathbf{A}_{21}\mathbf{B}_{11} + \mathbf{A}_{22}\mathbf{B}_{21} & \mathbf{A}_{21}\mathbf{B}_{12} + \mathbf{A}_{22}\mathbf{B}_{22} \end{bmatrix}.$$

Існує інша важлива схема множення матриць – *пряма,* або *кронеккеров* добуток двох матриць. Якщо матриця $\mathbf{A}$ має розмір $n{\times}m$, а матриця $\mathbf{B}$ – розмір $p{\times}q$, то їхній прямий (кронеккеров) добуток $\mathbf{A} \otimes \mathbf{B}$ визначається так:

$$\mathbf{A} \otimes \mathbf{B} = \begin{bmatrix} a_{11}\mathbf{B} & a_{12}\mathbf{B} & \cdots & a_{1n}\mathbf{B} \\ a_{21}\mathbf{B} & a_{22}\mathbf{B} & \cdots & a_{2n}\mathbf{B} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1}\mathbf{B} & a_{m2}\mathbf{B} & \cdots & a_{mn}\mathbf{B} \end{bmatrix}.$$

Прямий добуток $\mathbf{A} \otimes \mathbf{B}$ має, отже, порядок $mp{\times}nq$. Якщо $\mathbf{A}$ і $\mathbf{B}$ – квадратні невироджені матриці одного порядку, то легко показати, що

$$\boxed{(\mathbf{A} \otimes \mathbf{B})^{-1} = \mathbf{A}^{-1} \otimes \mathbf{B}^{-1}}. \qquad (1.27)$$

Варто звернути увагу на відмінність цього результату від властивості обертання матриць (1.17), де порядок множення обернених матриць змінюється, у той час як у (1.27) порядок залишається колишнім. Читачу пропонується довести (1.27) самостійно в якості вправи.

Матрицю після розбивки часто доводиться обертати, і нерідко обчислення обернених матриць істотно спрощується, якщо розбивка початкової матриці обрано відповідним способом.

Нехай A – невироджена матриця, для якої задана розбивка

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix},$$

де $\mathbf{A}_{11}$ і $\mathbf{A}_{22}$ – квадратні невироджені матриці.

Нехай обернена матриця розбита відповідним способом і записана у виді $\mathbf{A}^{-1} = \begin{bmatrix} \mathbf{E} & \mathbf{F} \\ \mathbf{G} & \mathbf{H} \end{bmatrix}$. За допомогою рівняння $\mathbf{A}\mathbf{A}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, де одинична матриця також розбита відповідним способом на підматриці, легко показати, що обернена матриця $\mathbf{A}^{-1}$ може бути виражена наступним чином:

$$\mathbf{A}^{-1} = \begin{bmatrix} \mathbf{E} & -\mathbf{E}\mathbf{A}_{12}\mathbf{A}_{22}^{-1} \\ -\mathbf{A}_{22}^{-1}\mathbf{A}_{21}\mathbf{E} & \mathbf{A}_{22}^{-1} + \mathbf{A}_{22}^{-1}\mathbf{A}_{21}\mathbf{E}\mathbf{A}_{12}\mathbf{A}_{22}^{-1} \end{bmatrix}, \tag{1.28}$$

де

$$\mathbf{E} = \left(\mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21}\right)^{-1}. \tag{1.29}$$

Це лише одна з можливих форм, у яких може бути подана обернена матриця, якщо вихідна піддалася розбивці. Така форма виявляється корисною для обчислень, якщо $\mathbf{A}_{22}$ має особливості, що дозволяють легко обчислити обернену матрицю $\mathbf{A}_{22}^{-1}$.

Альтернативна форма представлення оберненої матриці може обпиратися на обчислення матриці $\mathbf{A}_{11}^{-1}$. Її висновок пропонується читачу в якості вправи.

Корисне застосування формули (1.28) має при знаходженні матриці, оберненої до $\mathbf{A} = \begin{bmatrix} \mathbf{X}^{\mathrm{T}}\mathbf{X} & \mathbf{X}^{\mathrm{T}}\mathbf{D} \\ \mathbf{D}^{\mathrm{T}}\mathbf{X} & \mathbf{D}^{\mathrm{T}}\mathbf{D} \end{bmatrix}$, де $\mathbf{X}$ – матриця (розміру $n \times k$) спостережень за деякими змінними, а $\mathbf{D}$ – матриця (розміру $n \times s$) відповідним способом обраних фіктивних змінних. За допомогою (1.28) і (1.29) можна записати:

$$\mathbf{A}^{-1} = \begin{bmatrix} \mathbf{E} & -\mathbf{E}(\mathbf{D}^{\mathrm{T}}\mathbf{X})(\mathbf{D}^{\mathrm{T}}\mathbf{D})^{-1} \\ -(\mathbf{D}^{\mathrm{T}}\mathbf{D})^{-1}(\mathbf{D}^{\mathrm{T}}\mathbf{X})\mathbf{E} & (\mathbf{D}^{\mathrm{T}}\mathbf{D})^{-1} + (\mathbf{D}^{\mathrm{T}}\mathbf{D})^{-1}(\mathbf{D}^{\mathrm{T}}\mathbf{X})\mathbf{E}(\mathbf{D}^{\mathrm{T}}\mathbf{X})(\mathbf{D}^{\mathrm{T}}\mathbf{D})^{-1} \end{bmatrix}, \quad (1.30)$$

де

$$\mathbf{E} = \left(\mathbf{X}^{\mathrm{T}}\mathbf{M}\mathbf{X}\right)^{-1} \text{ і } \mathbf{M} = 1 - \mathbf{D}\left(\mathbf{D}^{\mathrm{T}}\mathbf{D}\right)^{-1}\mathbf{D}^{\mathrm{T}} \quad (1.31)$$

З (1.28) і (1.29) безпосередньо випливає, якщо $\mathbf{A}$ – невироджена матриця, що має вид

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & 0 \\ 0 & \mathbf{A}_{22} \end{bmatrix}, \quad (1.32)$$

то

$$\mathbf{A}^{-1} = \begin{bmatrix} \mathbf{A}_{11}^{-1} & 0 \\ 0 & \mathbf{A}_{22}^{-1} \end{bmatrix}. \quad (1.33)$$

Матриці виду (1.32) – (1.33) називаються блочно-діагональними. Результат (1.33) має місце і для матриць, що допускають розбивку більш ніж на два блоки, розташованих уздовж головної діагоналі.

У ряді випадків доводиться обчислювати визначники матриць, що були розбиті. Почнемо з простого зауваження, що

$$\begin{vmatrix} \mathbf{A}_{11} & 0 \\ 0 & \mathbf{I} \end{vmatrix} = \left|\mathbf{A}_{11}\right| . \quad (1.34)$$

Рівність (1.34) обґрунтовується наступнім чином. Якщо обчислити визначник, що стоїть в лівій частині, розкладанням по елементах останнього рядка, у котрої лише останній елемент відмінний від нуля і дорівнює одиниці, то в результаті одержимо одиницю, помножену на визначник такої ж форми, за винятком того, що порядок одиничної матриці $\mathbf{I}$ став на одиницю менше. Продовжуючи цю процедуру крок за кроком, прийдемо до результату (1.34).

Визначник блочно-діагональної матриці можна представити у виді

$$\begin{vmatrix} \mathbf{A}_{11} & 0 \\ 0 & \mathbf{A}_{22} \end{vmatrix} = \begin{vmatrix} \mathbf{A}_{11} & 0 \\ 0 & \mathbf{I} \end{vmatrix} \cdot \begin{vmatrix} \mathbf{I} & 0 \\ 0 & \mathbf{A}_{22} \end{vmatrix} = |\mathbf{A}_{11}| \cdot |\mathbf{A}_{22}|.$$

Розглянемо тепер визначник $\begin{vmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ 0 & \mathbf{I} \end{vmatrix}$. За допомогою аргументів, використаних при доказі (1.34), можна показати, що він дорівнює $|\mathbf{A}_{11}|$. Це дозволяє обчислити визначник *блочно-трикутної* матриці:

$$\begin{vmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ 0 & \mathbf{A}_{22} \end{vmatrix} = \begin{vmatrix} \mathbf{I} & 0 \\ 0 & \mathbf{A}_{22} \end{vmatrix} \cdot \begin{vmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ 0 & \mathbf{I} \end{vmatrix} = |\mathbf{A}_{11}| \cdot |\mathbf{A}_{22}| \ . \tag{1.35}$$

Тим самим одержуємо узагальнення результату (1.35), у силу якого визначник *трикутної* матриці дорівнює добутку його діагональних елементів (див. вправу 1.13).

Установимо, нарешті, вираз для визначника матриці, що має загальну розбивку:

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix},$$

де, як і колись, $\mathbf{A}_{11}$ і $\mathbf{A}_{22}$ – квадратні матриці., при цьому матриця $\mathbf{A}_{22}$ – невироджена. Визначимо матриці $\mathbf{B}_1$ й $\mathbf{B}_2$ наступним чином:

$$\mathbf{B}_1 = \begin{bmatrix} \mathbf{I} & -\mathbf{A}_{12}\mathbf{A}_{22}^{-1} \\ 0 & \mathbf{I} \end{bmatrix}; \quad \mathbf{B}_2 = \begin{bmatrix} \mathbf{I} & \\ -\mathbf{A}_{22}^{-1}\mathbf{A}_{21} & \mathbf{I} \end{bmatrix},$$

тоді

$$\mathbf{B}_1\mathbf{A}\mathbf{B}_2 = \begin{bmatrix} \mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21} & 0 \\ 0 & \mathbf{A}_{22} \end{bmatrix} .$$

І оскільки $|\mathbf{B}_1| = |\mathbf{B}_2| = \mathbf{I}$, то

$$\boxed{\left|\mathbf{A}\right| = \left|\mathbf{A}_{22}\right| \cdot \left|\mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21}\right|} \,. \tag{1.36}$$

Аналогічним способом у припущенні, що $\mathbf{A}_{22}$ - невироджена матриця, можна одержати

$$\boxed{\left|\mathbf{A}\right| = \left|\mathbf{A}_{11}\right| \cdot \left|\mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12}\right|} \,. \tag{1.37}$$

1.6. Лінійна залежність. Ранг матриці. Вирішення системи однорідних лінійних рівнянь

Розглянемо систему однорідних лінійних рівнянь

$$\mathbf{Ax} = 0 \;, \tag{1.38}$$

де $\mathbf{A}$ – матриця (розміру $m{\times}n$) коефіцієнтів, а $\mathbf{x}$ – вектор-стовпець із n невідомих. Якщо позначити стовпці матриці $\mathbf{A}$ через $\mathbf{a}_1,\ \mathbf{a}_2, \dots, \mathbf{a}_n,$ то (1.38) можна буде записати так:

$$x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \dots + x_n\mathbf{a}_n = 0 \;, \tag{1.39}$$

де вирази, що стоять в лівій частині рівності, являють собою лінійну комбінацію стовпців матриці $\mathbf{A}$**.** Якщо рівняння (1.39) має тільки одне тривіальне рішення $\mathbf{x}$=0 (тобто *кожний* елемент вектора дорівнює нулю), то вектори $\mathbf{a}_1,\ \mathbf{a}_2, \dots, \mathbf{a}_n,$ називаються *лінійно-незалежними*. Якщо ж існує деякий вектор $\mathbf{x} \neq 0$ (тобто не всі елементи вектора $\mathbf{x}$ рівні нулю), що задовольняє рівнянню (1.39), то говорять, що стовпці матриці $\mathbf{A}$ *линійно залежні*. Ранг як *максимальне число лінійно-незалежних її стовпців*.

Обмежимося розглядом тільки квадратних матриць (m=n). Якщо матриця $\mathbf{A}$ невироджена, то $\left|\mathbf{A}\right| \neq 0$, і єдиним рішенням рівняння (1.38) буде $\mathbf{x}$=0. Таким чином, n стовпців матриці А являються лінійно-незалежними і ранг матриці А дорівнює *n*, що записується в такий спосіб:

$$\rho(\mathbf{A}) = n \;.$$

Очевидно, що n рядків матриці $\mathbf{A}$ теж лінійно-незалежні, оскільки $\mathbf{y}^{\mathrm{T}}\mathbf{A}$ (де $\mathbf{y}^{\mathrm{T}}$ – вектор-рядок із n елементів) являє собою лінійну комбінацію рядків матриці $\mathbf{A}$ і існує тільки тривіальне рішення $\mathbf{y}=0$ рівняння $\mathbf{y}^{\mathrm{T}}\mathbf{A} = 0$.

Отже, у випадку квадратної невиродженої матриці $\mathbf{A}$ $\left(|\mathbf{A}| \neq 0\right)$ ранг матриць $\mathbf{A}$ визначається її порядком і тому називається повним рангом. Якщо ж рівняння $\mathbf{A}$ має нетривіальне рішення, то визначник $|\mathbf{A}|$ дорівнюватиме нулю і у цьому випадку n стовпців (або рядків) матриці повинні бути лінійно залежними.

Для ілюстрації цих моментів розглянемо систему рівнянь

$$\begin{bmatrix} 1 & 2 & -3 \\ 2 & 0 & 2 \\ 4 & 1 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}. \qquad (1.40)$$

Оскільки $|\mathbf{A}| = -4$, тобто $|\mathbf{A}| \neq 0$, існує єдине рішення $x_1 = x_2 = x_3 = 0$. Стовпці (і рядки) матриці лінійно незалежні, і $\rho(\mathbf{A}) = 0$. Розглянемо тепер іншу систему рівнянь

$$\begin{bmatrix} 1 & 2 & -3 \\ 2 & 0 & 2 \\ 4 & 1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}. \qquad (1.41)$$

Визначник матриці коефіцієнтів системи (1.40) дорівнює нулю $|\mathbf{A}| = 0$**, а** це означає, що можна знайти безліч значень невідомих, не всі з яких дорівнюють нулю, таких, що

$$x_1 \begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix} + x_2 \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix} + x_3 \begin{bmatrix} -3 \\ 2 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}. \qquad (1.42)$$

При перевірці переконуємося, що $x_1 = 1, \quad x_2 = -2, \quad x_1 = -1$, або ці значення, помножені на будь-яке число, задовольняють рівнянню (1.42). Позначимо вектори-стовпці в (1.42) через $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$, при цьому одержимо

$$\mathbf{a}_1 - 2\mathbf{a}_2 - \mathbf{a}_3 = 0 \tag{1.43}$$

Ці вектори є лінійно-залежними. Рівняння (1.43) означає, що будь-який з цих трьох векторів може бути виражений як лінійна комбінація двох, що залишилися, наприклад,

$$\mathbf{a}_2 = 0{,}5\,\mathbf{a}_1 - 0{,}5\,\mathbf{a}_3 \ .$$

При перевірці переконуємося, що кожна пара векторів із (1.42) лінійно незалежна. Це означає, що ранг матриці $\mathbf{A}$ системи (1.41) дорівнює двом. Вкажемо також, що всі мінори другого порядку матриці $\mathbf{A}$ відмінні від нуля.

Як тільки виявиться лінійна залежність стовпців квадратної матриці $\mathbf{A}$, можна зробити висновок, що і рядки її теж лінійно залежні. Можна відшукати ненульовий вектор $\mathbf{y}$, такий що

$$y_1 \begin{bmatrix} 1 \\ 2 \\ -3 \end{bmatrix} + y_2 \begin{bmatrix} 2 \\ 0 \\ 2 \end{bmatrix} + y_3 \begin{bmatrix} 4 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} . \tag{1.44}$$

Ним буде або вектор $\mathbf{y}^{\mathrm{т}} = \begin{bmatrix} -0{,}5 & -1{,}75 & 1 \end{bmatrix}$, або його добуток на будь-який скаляр. Очевидно, що будь-які пари векторів із (1.44) лінійно незалежні.

Розглянемо тепер систему рівнянь

$$\begin{bmatrix} 1 & 2 & 2 \\ 2 & 0 & 4 \\ 1 & 1 & 8 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} . \tag{1.45}$$

Знову $|\mathbf{A}| = 0$. Лінійну залежність між стовпцями можна легко виявити і записати у вигляді

$$2\mathbf{a}_1 + 0 \cdot \mathbf{a}_2 - \mathbf{a}_3 = 0 \ ,$$

або, простіше,

$$2\mathbf{a}_1 - \mathbf{a}_3 = 0 \ .$$

Як бачимо, вектор $\mathbf{a}_2$ не може бути у вигляді лінійної комбінації векторів $\mathbf{a}_1$ і $\mathbf{a}_3$, або, іншими словами, не всі пари векторів-стовпців лінійно незалежні. Вектори $\mathbf{a}_1$ і $\mathbf{a}_2$ лінійно незалежні, те саме стосується і векторів $\mathbf{a}_2$ і $\mathbf{a}_3$, але не векторів $\mathbf{a}_1$ і $\mathbf{a}_3$. Той же факт має місце і для мінорів другого порядку. Заміна всіх елементів матриці $\mathbf{A}$ їхніх мінорів приводить до матриці

$$\begin{bmatrix} -4 & 0 & 2 \\ 14 & 0 & -7 \\ 8 & 0 & 1 \end{bmatrix},$$

де лінійна залежність між $\mathbf{a}_1$ і $\mathbf{a}_3$ виявилася в появі на місці $\mathbf{a}_2$ стовпця нульових мінорів. Ранг матриці $\mathbf{A}$ в цьому прикладі дорівнює двом, оскільки серед трьох лінійно залежних стовпців можна відшукати два лінійно-незалежних.

$$\begin{bmatrix} 1 & 2 & 6 \\ 2 & 4 & 12 \\ 4 & 8 & 24 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Тут теж $|\mathbf{A}| = 0$, однак на додаток до цього всі мінори другого порядку рівні нулю, тому неможливо знайти жодної пари лінійно-незалежних векторів-стовпців. Ранг цієї матриці $\mathbf{A}$ дорівнює одиниці.

Перейдемо тепер до більш загального випадку, коли матриця $\mathbf{A}$ має розмір $m \times n$. Нехай число r визначає максимальна кількість лінійно-незалежних стовпців цієї матриці. Можна безпосередньо переконатися в тому, що r не в змозі перевищити число рядків m.

Припустимо, що $r > m$ і що перші r стовпців матриці $\mathbf{A}$ лінійно незалежні. Тоді перші m стовпців теж лінійно незалежні. Якщо через $\mathbf{A}_1$ позначити квадратну матрицю порядку m, утворену першими стовпцями матриці $\mathbf{A}$, то одержимо $\left|\mathbf{A}_1\right| = 0$.

Таким чином, для будь-якого з лінійно-незалежних векторів $\mathbf{a}_i \quad (i = m+1, m+2, \cdots, r)$ рівняння

$$\mathbf{A}_1\mathbf{b} = \mathbf{a}_i$$

має ненульове рішення

$$\mathbf{b} = \mathbf{A}_1^{-1}\mathbf{a}_i \ ,$$

що суперечить припущенню про існування більш ніж m лінійно-незалежних стовпців.

Отже,

$$\boxed{\rho(\mathbf{A}) \le \min\{m, n\}} \ . \qquad (1.46)$$

Якщо $\rho(\mathbf{A}) = r$ і якщо знову припустити, що стовпці матриці $\mathbf{A}$ пронумеровані таким чином, що перші r лінійно незалежні, то єдиним вирішенням рівняння

$$x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + ... + x_r\mathbf{a}_r = 0 \ , \qquad (1.47)$$

буде $x_1 = x_2 = ... + x_r = 0$.

Якщо $r = m$, то вектори $\mathbf{a}_1$, $\mathbf{a}_2$, ... , $\mathbf{a}_r$ утворять квадратну матрицю порядку з відмінним від нуля визначником. Якщо $r < m$, то можна виключити із системи (1.47) $m - r$ рівнянь, лінійно-залежних від тих, що залишилися, і одержати рівняння

$$x_1\mathbf{a}_1^* + x_2\mathbf{a}_2^* + ... + x_r\mathbf{a}_r^* = 0 \ , \qquad (1.48)$$

де через $\mathbf{a}_i^*$ позначений вектор-стовпець, що містить r елементів. Єдиним вирішенням рівняння (1.48) є $x_1 = x_2 = ... = x_n$. Тому існує принаймні один мінор порядку r матриці $\mathbf{A}$, не рівний нулю. Оскільки будь-яка множина, що містить $r+1$ стовпців матриці $\mathbf{A}$, являє собою множину лінійно-залежних векторів, усі мінори матриці $\mathbf{A}$ порядку $r+1$ повинні бути рівні нулю. Те саме відноситься до мінорів порядку $r+2, \quad r+3$ і т.д. Оскільки існування в матриці ненульового мінору рівнозначно лінійній залежності існуючого розміру мінору кількості її рядків або стовпців, то одержуємо три ідентичних способи визначення рангу матриці, що дорівнює або максимальному числу лінійно-незалежних стовпців, або максимальному числу лінійно-незалежних рядків, або максимальному порядку відмінного від нуля мінору.

Наприклад, ранг матриці

$$\mathbf{A} = \begin{bmatrix} 4 & 8 & 2 \\ 2 & 4 & 4 \\ 2 & 4 & -2 \end{bmatrix}$$

дорівнює двом, тому що $|\mathbf{A}| = 0$**,** а деякі мінори другого порядку рівні нулю, тоді як інші відмінні від нуля. З визначення рангу очевидним способом випливає, що $\rho(\mathbf{I}_n) = n$ і що ранг діагональної матриці $\mathbf{A}$ дорівнює числу відмінних від нуля діагональних елементів.

Як уже було показано в (1.22) для невиродженої матриці $\mathbf{A}$

$$\left|\mathbf{A}^{-1}\right| = \frac{1}{|\mathbf{A}|} \ .$$

Отже, матриця $\mathbf{A}^{-1}$ має повний ранг.

Сформулюємо без доведення важливу теорему про ранг добутку двох матриць

$$\boxed{\rho(\mathbf{AB}) \leq min \, \{\rho(\mathbf{A}), \rho(\mathbf{B})\}} \, , \tag{1.49}$$

тобто ранг добутку матриць $\mathbf{AB}$ не може перевищувати менший з рангів матриць $\mathbf{A}$ і $\mathbf{B}$. З цієї теореми легко одержати висновок, за яким множення матриці $\mathbf{A}$ справа або зліва на невироджену матрицю $\mathbf{B}$ приводить до матриці, ранг якої дорівнює рангу матриці $\mathbf{A}$**.** Нехай матриця $\mathbf{A}$ має розмір

$m \times n$ і нехай $\rho(\mathbf{A}) = r$, а $\mathbf{B}$ – невироджена матриця порядку n, так що $\rho(\mathbf{B}) = r$. Нехай далі $\rho(\mathbf{AB}) = k$. Тоді за виразом (1.49)

$$k \le \min \{r, n\} ,$$

тобто $k \le r$, оскільки $r \le n$. Але $\mathbf{A} = (\mathbf{AB})\mathbf{B}^{-1}$. Отже,

$$r \le \min \{k, n\} .$$

Звідки $r \le k$, а виходить $k = r$, тоді маємо

$$\rho(\mathbf{AB}) = \rho(\mathbf{A}) , \tag{1.50}$$

де $\mathbf{B}$ – невироджена матриця

Аналогічно доводиться, що

$$\rho(\mathbf{CA}) = \rho(\mathbf{A}) , \tag{1.51}$$

де $\mathbf{C}$ – невироджена матриця, і що

$$\rho(\mathbf{CAB}) = \rho(\mathbf{A}) , \tag{1.52}$$

де $\mathbf{C}$ і $\mathbf{B}$ – невироджені матриці.

Важливим результатом для математичного програмування є розв'язання системи лінійних рівнянь. Розглянемо знову систему рівнянь (1.41):

$$\begin{bmatrix} 1 & 2 & -3 \\ 2 & 0 & 2 \\ 4 & 1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} .$$

Як уже зазначалося, $\rho(\mathbf{A}) = 2$ і кожний рядок матриці $\mathbf{A}$ може бути виражений лінійною комбінацією двох інших. Тому тільки одне рівняння системи можна відкинути без збитку для інформації, що міститься в системі. Відкинемо третє рівняння і перепишемо перші два у вигляді

$$\begin{bmatrix} 1 & 2 \\ 2 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = -\begin{bmatrix} -3 \\ 2 \end{bmatrix} x_3 \ .$$

Розв'язання цієї системи можна записати в такий спосіб

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = -\begin{bmatrix} -x_3 \\ 2x_3 \end{bmatrix} \ .$$

Вектор рішень системи (1.41) запишемо так:

$$\mathbf{x}^{\mathrm{T}} = \begin{bmatrix} -x_3 & 2x_3 & x_3 \end{bmatrix} = \begin{bmatrix} -1 & 2 & 1 \end{bmatrix} x_3 \ . \qquad (1.53)$$

Значення x_3 може бути довільним, тому існує нескінченна множина рішень системи (1.41), де для кожного заданого x_3 значення x_1 й x_2 одержують множенням x_3 на фіксовані коефіцієнти. Іншими словами, *відношення* елементів вектора рішення залишаються *незмінними,* і всі можливі рішення лежать на однієї прямої, що проходить через початок координат у тривимірному просторі. Геометрично це означає, що *вимірність* підпростору рішень системи (1.41) дорівнює одиниці і можна помітити, що в даному прикладі ця вимірність дорівнює різниці між числом невідомих (n = 3) і рангом матриці $\mathbf{A}$ $\left(\rho\left(\mathbf{A}\right) = 2\right)$.

Якщо тепер розглянути систему $\begin{bmatrix} 1 & 2 & 6 \\ 2 & 4 & 12 \\ 4 & 8 & 24 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$, то для $\rho\left(\mathbf{A}\right) = 1$ є тільки одне незалежне рівняння, що можна записати таким чином:

$$x_1 = -2x_2 - 6x_3 \ . \qquad (1.54)$$

Тепер уже два елементи вектора рішень можуть бути обрані довільно і множина рішень системи (1.54) являє собою площину в тривимірному просторі, що проходить через початок координат. Вимірність підпростору рішень знову дорівнює різниці між числом невідомих (n = 3) і рангом матриці $\mathbf{A}$ $\left(\rho\left(\mathbf{A}\right) = 2\right)$.

Ці приклади підказують загальний результат, внаслідок якого для матриці $\mathbf{A}$ розміру $m \times n$, ранг якої $\rho(\mathbf{A}) = r$, вимірність простору рішень $\mathbf{x}_N$ дорівнює $n - \rho(\mathbf{A})$. Це можна довести.

Розіб'ємо систему $\mathbf{Ax} = 0$:

$$\begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix} = 0 \ , \qquad (1.55)$$

де $\mathbf{A}_{11}$ – квадратна і невироджена матриця порядку r; $\mathbf{A}_{12}$ – матриця розміром $r \times (n\text{-}r)$ і т.д. Оскільки останні $m\text{-}r$ рядків у (1.55) надлишкові, можна обмежитися співвідношенням $\mathbf{A}_{11}\mathbf{x}_1 + \mathbf{A}_{12}\mathbf{x}_2 = 0$, що дає вираз для $\mathbf{x}_1$:

$$\mathbf{x}_1 = -\mathbf{A}_{11}^{-1}\mathbf{A}_{12}\mathbf{x}_2 0 \ .$$

Отже, усі вирішення системи $\mathbf{Ax} = 0$ набувають вигляду

$$\mathbf{x} = \begin{bmatrix} -\mathbf{A}_{11}^{-1}\mathbf{A}_{12}\mathbf{x}_2 \\ \mathbf{x}_2 \end{bmatrix} = \begin{bmatrix} -\mathbf{A}_{11}^{-1}\mathbf{A}_{12} \\ \mathbf{I}_{m-r} \end{bmatrix} \mathbf{x}_2 \ . \qquad (1.56)$$

Матриця, що стоїть у правій частині рівності (1.56), має розмір $n \times (n\ \text{-}r)$ і ранг $n\text{-}r$, а $\mathbf{x}_2$ позначає довільний вектор з $n – r$ елементів. Тому всі вектори простору рішень є лінійними комбінаціями $n\text{-}r$ лінійно-незалежних векторів і, отже, мірність простору рішень системи $\mathbf{Ax} = 0$ є $n - \rho(\mathbf{A})$.

Із сказаного випливає, що рішення системи рівнянь $\mathbf{Ax} = 0$, де $\mathbf{A}$ – матриця розміру $m \times n$, а $\mathbf{x}$ – $n \times 1$-вимірний вектор, лежать на однієї прямої тоді і тільки тоді, коли $\rho(\mathbf{A}) = n - 1$.

1.7. Власні значення і вектори

Проблема відшукання *власних значень* (*характеристичних чисел*) полягає у встановленні таких чисел λ і відповідних їм власних *векторів* $\mathbf{x} \neq 0$, що задовольняють рівності

$$\mathbf{Ax} = \lambda\mathbf{x} \ , \tag{1.57}$$

де $\mathbf{A}$ – деяка матриця розміру $m \times n$. Число λ називають *власним (характеристичним) значенням* матриці $\mathbf{A}$**,** а $\mathbf{x}$ – її *власним (характеристичним) вектором.* Якщо візьмемо для ілюстрації випадок матриці 2×2, то з (1.57) випливає система рівнянь

$$\begin{aligned} (a_{11} - \lambda)x_1 + a_{12}x_2 &= 0; \\ a_{21}x_1 + (a_{22} - \lambda)x_2 &= 0, \end{aligned}$$

яку можна переписати у матричній формі як

$$(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = 0 \ . \tag{1.58}$$

Рівняння (1.58) має нетривіальне рішення $\mathbf{x} \neq 0$ тільки в тому випадку, якщо матриця $(\mathbf{A} - \lambda\mathbf{I})$ вироджена, тобто якщо

$$|\mathbf{A} - \lambda\mathbf{I}| = 0 \ . \tag{1.59}$$

Рівняння (1.59) – алгебраїчне щодо невідомої λ. Його потрібно спочатку вирішити відносно λ, а потім знайти відповідні власні вектори. Для нашого прикладу рівняння (1.59) має вигляд

$$(a_{11} - \lambda)(a_{22} - \lambda) - a_{12}a_{21} = 0 \ ,$$

тобто ми одержали квадратне рівняння

$$\lambda^2 - (a_{11} + a_{22})\lambda + a_{11}a_{22} - a_{12}a_{21} = 0$$

з коренями

$$\lambda_1 = \tfrac{1}{2}\left[(a_{11}+a_{22}) + \sqrt{(a_{11}+a_{22})^2 - 4(a_{11}a_{22} - a_{12}a_{21})}\right];$$

$$\lambda_1 = \tfrac{1}{2}\left[(a_{11}+a_{22}) - \sqrt{(a_{11}+a_{22})^2 - 4(a_{11}a_{22} - a_{12}a_{21})}\right].$$

У спеціальному випадку симетричної матриці, коли $a_{12} = a_{21}$, власними значеннями будуть

$$\lambda_{1,2} = \tfrac{1}{2}\left[(a_{11}+a_{22}) \pm \sqrt{(a_{11}-a_{22})^2 - 4a_{12}^2}\right] , \qquad (1.60)$$

і оскільки вирази, що стоять під знаком квадратного кореня, є сумою двох квадратів, корені λ_1 і λ_2 – неодмінно дійсні, якщо матриця $\mathbf{A}$ симетрична.

Розглянемо чисельний приклад з матрицею $\mathbf{A} = \begin{bmatrix} 4 & 2 \\ 2 & 1 \end{bmatrix}$. Підставимо значення матричних коефіцієнтів у (1.60) і одержимо, що $\lambda_1 = 5$ і $\lambda_2 = 0$. Щоб знайти власні вектори, які відповідають кожному з цих власних значень, вирішимо для конкретного λ систему (1.58). У даному прикладі для $\lambda_1 = 5$

$$\mathbf{A} - \lambda\mathbf{I} = \begin{bmatrix} -1 & 2 \\ 2 & -4 \end{bmatrix},$$

і система (1.58) має вигляд

$$-x_1 + 2x_2 = 0;$$
$$2x_1 - 4x_2 = 0,$$

звідки $x_1 = 2x_2$.

Як бачимо, один елемент власного вектора довільний і тому, якщо $\mathbf{x}$ задовольняє системі (1.58) при даному значенні λ, то $k\mathbf{x}$, де k – довільна константа, теж задовольняє цій системі. Ми можемо нормалізувати вектор $\mathbf{x}$, привівши його довжину до одиниці, тобто забезпечивши $x_1^2 + x_2^2 = 1$. Це при $x_1 = 2x_2$ дає $x_2 = \frac{1}{\sqrt{5}}$, $x_1 = \frac{2}{\sqrt{5}}$. Власний вектор, що відповідає власному значенню $\lambda_1 = 5$, визначиться таким чином:

$$\mathbf{x}_1^{\mathrm{T}} = \begin{bmatrix} \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{bmatrix}.$$

Аналогічно можна одержати власний вектор , що відповідає значенню $\lambda_2 = 0$:

$$\mathbf{x}_2^{\mathrm{T}} = \begin{bmatrix} \frac{1}{\sqrt{5}} & \frac{-2}{\sqrt{5}} \end{bmatrix}.$$

Варто звернути увагу на те що $\mathbf{x}_1^{\mathrm{T}} \mathbf{x}_2 = 0$, тобто власні вектори симетричної матриці ортогональні[1].

Надалі будемо розглядати тільки дійсні симетричні матриці.

Відзначені дві властивості – існування дійсних коренів і ортогональність власних векторів – мають місце й у загальному випадку дійсної симетричної матриці порядку n. Більше того, якщо характеристичний корінь λ має кратність k (тобто повторюється k разів), то є k відповідних йому ортогональних власних векторів.

Симетрична матриця $\mathbf{A}$ порядку n має власні значення $\lambda_1, \lambda_2, \ldots, \lambda_n$, не обов'язково різні між собою. Цим значенням відповідає система ортогональних власних векторів $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$, таких що

$$\mathbf{x}_i^{\mathrm{T}} \mathbf{x}_j = \delta_{ij}, \quad \delta_{ij} = \begin{cases} 0, & i \neq j; \\ 1, & i = j, \end{cases} \tag{1.61}$$

[1] Два вектори $\mathbf{x}$ і $\mathbf{y}$ називаються ортогональними, якщо $\mathbf{x}^{\mathrm{T}} \mathbf{y} = \mathbf{y}^{\mathrm{T}} \mathbf{x} = 0$ и $\mathbf{x}^{\mathrm{T}} \mathbf{x} \neq 0$, $\mathbf{y}^{\mathrm{T}} \mathbf{y} \neq 0$.

де δ_{ij} – символ Кронеккера. Якщо $\mathbf{X}$ – матриця порядку n, стовпці якої є вектори $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$, то

$$\mathbf{X}^{\mathrm{т}}\mathbf{X} = \boldsymbol{Й}_n \ , \tag{1.62}$$

або

$$\mathbf{X}^{\mathrm{т}} = \mathbf{X}^{-1} \ , \tag{1.63}$$

тобто матриця, транспонована до $\mathbf{X}$, дорівнює оберненій матриці $\mathbf{X}^{-1}$. Така матриця називається *ортогональною*. Для неї також безпосередньо встановлюємо, що $\mathbf{X}\mathbf{X}^{\mathrm{т}} = \boldsymbol{Й}$. Якщо тепер утворити добуток $\mathbf{X}^{\mathrm{т}}\mathbf{A}\mathbf{X}$ і обчислити елемент ij отриманої в результаті матриці n–го порядку, то, скориставшись співвідношенням (1.57), одержимо

$$\mathbf{x}_i^{\mathrm{т}}\mathbf{A}\mathbf{x}_j = \lambda_j \mathbf{x}_i^{\mathrm{т}}\mathbf{x}_j \ ,$$

де $\mathbf{x}_j$ – власний вектор, що відповідає значенню λ_j.

Замінимо тепер $\mathbf{x}_i^{\mathrm{т}}\mathbf{x}_j$ згідно з (1.61) на δ_{ij}, тобто $\mathbf{x}_i^{\mathrm{т}}\mathbf{A}\mathbf{x}_j = \lambda_j \delta_{ij}$. Остаточно маємо:

$$\mathbf{X}^{\mathrm{т}}\mathbf{A}\mathbf{X} = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}. \tag{1.64}$$

Якщо власні вектори матриці $\mathbf{A}$ розташовані у вигляді стовпців матриці $\mathbf{X}$, то добуток $\mathbf{X}^{\mathrm{т}}\mathbf{A}\mathbf{X}$ перетворить матрицю $\mathbf{A}$ в діагональну матрицю, що має власне значення $\mathbf{A}$ на своїй головній діагоналі. Співвідношення (1.64) можна розглядати як рівняння щодо перетворення $\mathbf{X}$, що приводить симетричну матрицю до діагонального вигляду.

1.8. Позитивна визначеність матриці

Нагадаємо, що квадратична форма в матричному вигляді записується як $\mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x}$, де $\mathbf{x}^{\mathrm{T}}$ і $\mathbf{x}$ – відповідно вектор-рядок і вектор-стовпець перемінних, що перебувають із n елементів; $\mathbf{A}$ – симетрична матриця коефіцієнтів n–го порядку.

Визначення 1.8. Квадратична форма і відповідна їй матриця називаються додатно визначеними тоді і тільки тоді, коли $\mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x} > 0$ для усіх $\mathbf{x} \neq 0$.

Визначення 1.9. Квадратична форма і відповідна їй матриця називаються додатно напіввизначеними, якщо $\mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x} \geq 0$ для всіх $\mathbf{x}$.

З наведених визначень випливає, що позитивно визначена матриця повинна бути невиродженою. Якби вона була вироджена, то рівняння $\mathbf{A}\mathbf{x}=0$ мало б нетривіальне рішення і тоді $\mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x} = 0$ для $\mathbf{x} \neq 0$, що суперечить припущенню про позитивну визначеність матриці **A.**

Надалі вважатимеме, якщо тільки не існують інші умови, що $\mathbf{A}$ – симетрична додатно визначена матриця порядку n.

Якщо матриця $\mathbf{B}$ має розмір $n \times s$ $(s \leq n)$ і $\rho(\mathbf{B}) = s$, то матриця $\mathbf{B}^{\mathrm{T}}\mathbf{A}\mathbf{B}$ додатно визначена. (1.65)

Очевидно, що матриця $\mathbf{B}^{\mathrm{T}}\mathbf{A}\mathbf{B}$ є симетричною і має порядок s. Розглянемо деякий відмінний від нульового вектор $\mathbf{y}$, що містить s елементів. Нехай $\mathbf{x} = \mathbf{B}\mathbf{y}$. Звідси випливає, що $\mathbf{x} \neq 0$. Справді, якщо $s<n$, то можна записати це рівняння у вигляді

$$\begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{B}_1 \\ \mathbf{B}_2 \end{bmatrix} \mathbf{y} \ ,$$

де матриця **B** розбита на дві підматриці: $\mathbf{B}_1$, що містить перші $s<n$ рядків, і $\mathbf{B}_2$, куди входять інші n-s рядків. Оскільки ранг матриці **B** дорівнює s, можна припустити, що $\mathbf{B}_1$ – невироджена матриця і тому має обернену матрицю. Тоді $\mathbf{y} = \mathbf{B}_1^{-1}\mathbf{x}_1$. Якщо $\mathbf{x} = 0$, то $\mathbf{x}_1$ теж дорівнює нулю, звідки $\mathbf{y} = 0$, що суперечить початковому твердженню.

Таким чином, ми довели, що $\mathbf{y} \neq 0$. Якщо ж s=n, то цей результат випливає негайно без усякої розбивки матриці . Тепер очевидно, що для довільного вектора $\mathbf{y} \neq 0$

$$\mathbf{y}^{\mathrm{т}}\left(\mathbf{B}^{\mathrm{т}}\mathbf{AB}\right)\mathbf{y} = \mathbf{x}^{\mathrm{т}}\mathbf{Ax} > 0 \ ,$$

і, отже, матриця $\mathbf{B}^{\mathrm{т}}\mathbf{AB}$ позитивно визначена.

☞ Спеціальний випадок твердження (1.65) маємо, коли s=n і матриця **B** невироджена. Умови теореми при цьому задовольняються, тому матриця $\mathbf{B}^{\mathrm{т}}\mathbf{AB}$ додатно визначена для невиродженої **B.** (1.66)

☞ Якщо **B** в (1.66) замінити матрицею $\mathbf{A}^{\mathrm{т}}$, то $\mathbf{B}^{\mathrm{т}}\mathbf{AB} = \left(\mathbf{A}^{-1}\right)^{\mathrm{т}} = \mathbf{A}^{-1}$, тобто матриця, обернена до симетричної, теж симетрична. Крім того, матриця $\mathbf{A}^{-1}$ додатно визначена. (1.67)

З визначення позитивної визначеності матриці, випливає, що одинична матриця **I** теж додатно визначена, тому що

$$\mathbf{x}^{\mathrm{т}}\mathbf{Ix} = \mathbf{x}^{\mathrm{т}}\mathbf{x} > 0 \ .$$

Замінивши матрицю $\mathbf{A}$ в (1.65) матрицею $\mathbf{I}$, одержимо, що матриця $\mathbf{B}^{\mathrm{T}}\mathbf{B}$ додатно визначена і, отже, $\mathbf{B}^{\mathrm{T}}\mathbf{B}$ – невироджена матриця. (1.68)

Іншими словами,

$$\rho\left(\mathbf{B}^{\mathrm{T}}\mathbf{B}\right) = s = \rho\left(\mathbf{B}\right) . \tag{1.69}$$

Важлива властивість додатно визначених матриць міститься в наступному твердженні.

Матриця $\mathbf{A}$ додатно визначена тоді і тільки тоді, коли всі її власні значення додатні. (1.70)

Нехай $\mathbf{X}$ – ортогональна матриця, що перетворює матрицю $\mathbf{A}$ до діагонального вигляду, тобто

$$\mathbf{X}^{\mathrm{T}}\mathbf{A}\mathbf{X} = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} = \mathbf{Л} , \tag{1.71}$$

де λ_i – власні значення матриці $\mathbf{A}$. Нехай $\mathbf{z}$ – відмінний від нуля вектор-стовпець із n елементів. Визначимо $\mathbf{y} = \mathbf{X}^{\mathrm{T}}\mathbf{z}$. Тоді $\mathbf{z} = \mathbf{y}\mathbf{X}$. Оскільки $\mathbf{X}$ – ортогональна матриця, то

$$\mathbf{z}^{\mathrm{T}}\mathbf{A}\mathbf{z} = \mathbf{y}^{\mathrm{T}}\mathbf{X}^{\mathrm{T}}\mathbf{A}\mathbf{X}\mathbf{y} = \sum_{i=1}^{n}\lambda_i y_i^2 .$$

Для достатності припустимо, що всі λ_i додатні. Тоді і матриця $\mathbf{A}$ позитивно визначена, тобто

$$\mathbf{z}^{\mathrm{T}}\mathbf{A}\mathbf{z} = \sum_{i=1}^{n} \lambda_i y_i^2 > 0.$$

Нехай, наприклад, $\lambda_i \le 0$. Виберемо вектор стовпець $\mathbf{y}$ такий, що $\mathbf{y}^{\mathrm{T}} = \begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix}$. З $\mathbf{z} = \mathbf{y}\mathbf{X}$ випливає, що $\mathbf{z}$ – ненульовий вектор, оскільки він фактично є першим стовпцем матриці $\mathbf{X}$. Тому маємо

$$\mathbf{z}^{\mathrm{T}}\mathbf{A}\mathbf{z} = \lambda_1 \le 0 \ ,$$

що суперечить припущенню про додатну визначеність матриці $\mathbf{A}$. Таким чином завершується доведення твердження (1.70).

Оскільки $\mathbf{X}^{\mathrm{T}}\mathbf{X} = \mathbf{I}$ і $\left|\mathbf{X}^{\mathrm{T}}\right| = \left|\mathbf{X}\right|$, то $\left|\mathbf{X}\right|^2 = 1$ і $\left|\mathbf{X}\right| = \pm 1$, тобто визначник ортогональної матриці дорівнює або одиниці, або мінус одиниці. Застосуємо цей результат до (1.71), і одержимо $\left|\mathbf{A}\right| = \lambda_1 \lambda_2 \ldots \lambda_n$, тоді

$$\left|\mathbf{A}\right| > 0 \ . \tag{1.72}$$

Отже, додатно визначена матриця невироджена, має додатні власні значення і додатний визначник. Звідси випливає, що всі *головні мінори додатно визначеної матриці додатні*.

Головні мінори матриці – визначники різних підматриць, утворених у результаті викреслювання в початковій матриці рядків і стовпців з однаковими номерами. Головними мінорами першого порядку є діагональні елементи a_{ii} $\left(i = 1, 2, ..., n\right)$, отримані в результаті викреслювання всіх рядків і стовпців, крім тих, на перетині яких стоїть елемент a_{ii}. Головний мінор другого порядку має вигляд: $\begin{vmatrix} a_{ii} & a_{ij} \\ a_{ji} & a_{jj} \end{vmatrix}$. Він отриманий шляхом викреслювання всіх рядків і стовпців матриці $\mathbf{A}$, за винятком тих, в яких номери i і j .

Отриманий результат про головні мінори випливає з (1.65), якщо взяти одиничну матрицю порядку $n \times n$, викреслити з неї ті ж стовпці, що повинні бути викреслені з матриці $\mathbf{A}$ для одержання відповідного головного мінору, і позначити отриману підматрицю одиничної матриці через **B.** Наприклад, якщо $\mathbf{B}=\begin{bmatrix}1 & 0\\ 0 & 1\\ 0 & 0\\ \cdots & \cdots\\ 0 & 0\end{bmatrix}$, то $\mathbf{B}^{\mathrm{т}}\mathbf{AB}=\begin{bmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{bmatrix}$. Згідно з (1.65) $\mathbf{B}^{\mathrm{т}}\mathbf{AB}$ – позитивно визначена матриця. Отже, з урахуванням (1.72) $\begin{bmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{bmatrix}>0$.

Аналогічні міркування мають місце і для будь-яких інших головних мінорів.

Рівняння (1.71) можемо застосувати для одержання іншого важливого результату. Оскільки усі λ_i додатні, можна задати діагональну матрицю $\mathbf{D}$ таким чином:

$$\mathbf{D}=\begin{bmatrix}\frac{1}{\sqrt{\lambda_1}} & 0 & \cdots & 0\\ 0 & \frac{1}{\sqrt{\lambda_2}} & \cdots & 0\\ \cdots & \cdots & \cdots & \cdots\\ 0 & 0 & \cdots & \frac{1}{\sqrt{\lambda_n}}\end{bmatrix}.$$

Множення (1.71) на $\mathbf{D}$ зліва і справа дає $(\mathbf{XD})^{\mathrm{т}}\mathbf{A}(\mathbf{XD})=\mathbf{I}_n$, або

$$\mathbf{Q}^{\mathrm{T}}\mathbf{AQ}=\mathbf{I}_n\,, \tag{1.73}$$

де $\mathbf{Q}=\mathbf{XD}$.

Оскільки матриці $\mathbf{X}$ і $\mathbf{D}$ – невироджені, $\mathbf{Q}$ – теж невироджена матриця. Згідно з (1.73)

$$\mathbf{A}=\left(\mathbf{Q}^{-1}\right)^{\mathrm{т}}\mathbf{Q}^{-1}\,. \tag{1.74}$$

Таким чином, якщо $\mathbf{A}$ – позитивно визначена матриця, то можна знайти таку невироджену матрицю $\mathbf{P} = \left(\mathbf{Q}^{-1}\right)^{\mathrm{т}}$, що

$$\mathbf{A} = \mathbf{P}\mathbf{P}^{\mathrm{т}} . \tag{1.75}$$

Інший корисний результат, що випливає безпосередньо з (1.71), полягає в тому, що слідом матриці $\mathbf{A}$ є сума її власних значень:

$$\sum_{i=1}^{n} \lambda_i = \mathrm{tr}\left(\mathbf{X}^{\mathrm{т}}\mathbf{A}\mathbf{X}\right) = \mathrm{tr}\left(\mathbf{X}^{\mathrm{т}}\mathbf{A}\mathbf{X}\right) = \mathrm{tr}\left(\mathbf{A}\right) . \tag{1.76}$$

При доведенні (1.70) було розглянуте перетворення $\mathbf{y} = \mathbf{X}^{\mathrm{т}}\mathbf{z}$ і показано, що

$$\mathbf{z}^{\mathrm{т}}\mathbf{A}\mathbf{z} = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \ldots + \lambda_n y_n^2 . \tag{1.77}$$

Становить особливий інтерес випадок, коли у співвідношенні (1.77) матриця $\mathbf{A}$ є *ідемпотентною*. Надалі будуть розглядатися тільки симетричні ідемпотентні матриці, визначені як

$$\mathbf{A} = \mathbf{A}^{\mathrm{т}} , \quad \mathbf{A} = \mathbf{A}^2 , \tag{1.78}$$

тобто ідемпотентна матриця – це матриця, що не змінюється при множенні на себе.

Усі власні значення ідемпотентної матриці – або нулі, або одиниці. Дійсно, якщо $\mathbf{A}\mathbf{z} = \lambda\mathbf{z}$, то $\mathbf{A}^2\mathbf{x} = \lambda\mathbf{A}\mathbf{x} = \lambda^2\mathbf{x}$. Однак $\mathbf{A}^2\mathbf{z} = \mathbf{A}\mathbf{z} = \lambda\ \mathbf{z}$, тобто $\lambda^2\mathbf{z} = \lambda\mathbf{z}$. Оскільки $\mathbf{z} \neq 0$, то $\lambda^2 - \lambda = 0$. Звідки або $\lambda = 0$, або $\lambda = 1$.

Тепер із (1.71) випливає, що коли ідемпотентна матриця порядку n і рангу n-k $\left(0 \leq k \leq n\right)$ приведена до діагонального вигляду, то вона буде містити n-k одиниць і k нулів на головній діагоналі. У протилежному випадку ранг отриманої матриці не буде дорівнювати n-k, а він не міг змінитися при множенні $\mathbf{A}$ на невироджену матрицю.

Отже, отримано важливий результат, за яким квадратична форма $\mathbf{z}^{\mathrm{T}}\mathbf{A}\mathbf{z}$ приводиться до вигляду

$$\mathbf{z}^{\mathrm{T}}\mathbf{A}\mathbf{z} = y_1^2 + y_2^2 + \ldots + y_{n-k}^2 \ , \tag{1.79}$$

якщо $\mathbf{A}$ – ідемпотентна матриця рангу n-k.

Приклад 1.8. Розглянемо наступний приклад докладніше, оскільки він ілюструє багато важливих моментів матричної алгебри.

Нехай дано матрицю $\mathbf{Z} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix}$.

Тоді $\mathbf{Z}^{\mathrm{T}}\mathbf{Z} = \begin{bmatrix} 3 & 6 \\ 6 & 14 \end{bmatrix}$, $\left|\mathbf{Z}^{\mathrm{T}}\mathbf{Z}\right| = 6,$ $\left(\mathbf{Z}^{\mathrm{T}}\mathbf{Z}\right)^{-1} = \begin{bmatrix} \frac{14}{6} & -1 \\ -1 & \frac{3}{6} \end{bmatrix}$ і

$$\mathbf{Z}\left(\mathbf{Z}^{\mathrm{T}}\mathbf{Z}\right)^{-1}\mathbf{Z}^{\mathrm{T}} = \begin{bmatrix} \frac{5}{6} & \frac{2}{6} & -\frac{1}{6} \\ \frac{2}{6} & \frac{2}{6} & \frac{2}{6} \\ -\frac{1}{6} & \frac{21}{6} & \frac{5}{6} \end{bmatrix}.$$

Визначимо нову матрицю $\mathbf{A}$ таким чином:

$$\mathbf{A} = \mathbf{I}_3 - \mathbf{Z}\left(\mathbf{Z}^{\mathrm{T}}\mathbf{Z}\right)^{-1}\mathbf{Z}^{\mathrm{T}} = \begin{bmatrix} \frac{1}{6} & -\frac{2}{6} & \frac{1}{6} \\ -\frac{2}{6} & \frac{4}{6} & -\frac{2}{6} \\ \frac{1}{6} & -\frac{2}{6} & \frac{1}{6} \end{bmatrix}.$$

Як бачимо, $\mathbf{A}$ – симетрична матриця. Вона також являє собою ідемпотентну матрицю, тому що $\mathbf{A}^2 = \mathbf{A}$. Зазначимо, що ранг матриці $\mathbf{A}$ дорівнює одиниці.

Знайдемо тепер власні значення і власні вектори матриці $\mathbf{A}$. Її характеристичне рівняння $|\mathbf{A}-\lambda\mathbf{I}|=0$ при розкладанні визначника $|\mathbf{A}-\lambda\mathbf{I}|$ по першому рядку (або стовпцю) дає наступний результат:

$$\left(\frac{1}{6}-\lambda\right)\left[\left(\frac{4}{6}-\lambda\right)\left(\frac{1}{6}-\lambda\right)-\frac{4}{36}\right]+\frac{2}{6}\left[-\frac{2}{6}\left(\frac{1}{6}-\lambda\right)+\frac{2}{36}\right]+$$
$$+\frac{1}{6}\left[\frac{4}{36}-\frac{1}{6}\left(\frac{4}{6}-\lambda\right)\right]=0,$$

тобто

$$\lambda^3+\lambda^2=0\ . \tag{1.80}$$

Рівняння (1.80) має простий корінь $\lambda_1=1$ і корінь $\lambda_2=0$ кратності 2, що ілюструє раніше отриманий результат, згідно з яким власні значення ідемпотентної матриці можуть дорівнювати або нулю, або одиниці, а число одиничних власних значень збігається з рангом цієї матриці. Для власного значення $\lambda_1=1$ одержимо

$$\mathbf{A}-\lambda\mathbf{I}=\begin{bmatrix}-\frac{5}{6} & -\frac{2}{6} & \frac{1}{6}\\ -\frac{2}{6} & -\frac{2}{6} & -\frac{2}{6}\\ \frac{1}{6} & -\frac{2}{6} & -\frac{5}{6}\end{bmatrix},$$

і матричне рівняння $(\mathbf{A}-\lambda\mathbf{I})\mathbf{x}=0$ набуває вигляду

$$\begin{aligned}-5x_1-2x_2+x_3&=0\,;\\ -2x_1-2x_2-2x_3&=0\,;\\ x_1-2x_2-5x_3&=0\,;\end{aligned} \tag{1.81}$$

До цих рівнянь додамо умову

$$x_1^2+x_2^2+x_3^2=1\ , \tag{1.82}$$

що дозволяє нормалізувати вектор .

Рівнянням (1.81) і (1.82) задовольняє вектор

$$x^{\text{т}} = \begin{bmatrix} \frac{1}{\sqrt{6}} & \frac{-2}{\sqrt{6}} & \frac{1}{\sqrt{6}} \end{bmatrix} . \qquad (1.83)$$

Для $\lambda_2 = 0$ матричне рівняння $(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = 0$ приводить тільки до одного незалежного рівняння

$$x_1 - 2x_2 + x_3 = 0 \ , \qquad (1.84)$$

яке разом з умовою нормалізації (1.82) означає, що один з елементів вектора $\mathbf{x}$ може бути обраний довільно. Наприклад, нехай $x_3 = 0$. Тоді спільне вирішення рівняння (1.82) і (1.84) дає вектор

$$\left(\mathbf{x}^*\right)^{\text{т}} = \begin{bmatrix} \frac{2}{5} & \frac{1}{\sqrt{5}} & 0 \end{bmatrix} , \qquad (1.85)$$

тобто власний вектор, що відповідає власному значенню $\lambda = 0$. Оскільки це власне значення має кратність 2, правомірно очікувати одержання іншого власного вектора $\mathbf{x}^{**}$, що відповідає тому ж власному значенню й ортогонального вектора $\mathbf{x}^*$.

Умова ортогональності векторів $\left(\mathbf{x}^*\right)^{\text{т}} \mathbf{x}^{**} = 0$ дає

$$2x_1^{**} + x_2^{**} = 0 \ . \qquad (1.86)$$

Оскільки $\mathbf{x}^{**}$ відповідає $\lambda = 0$, цей вектор повинен також задовольняти рівнянню (1.84) і, звичайно, умові ортогональності (1.82). Ці три рівняння визначають вектор

$$\left(\mathbf{x}^{**}\right)^{\text{т}} = \begin{bmatrix} \frac{1}{\sqrt{30}} & \frac{-2}{\sqrt{30}} & \frac{5}{\sqrt{30}} \end{bmatrix} . \qquad (1.87)$$

Утворимо тепер матрицю $\mathbf{X}$, стовпці якої складені з трьох власних векторів (1.83), (1.85), (1.87):

$$\mathbf{X} = \begin{bmatrix} \frac{1}{\sqrt{6}} & \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{30}} \\ \frac{-2}{\sqrt{6}} & \frac{1}{\sqrt{5}} & \frac{-2}{\sqrt{30}} \\ \frac{1}{\sqrt{6}} & 0 & \frac{-5}{\sqrt{30}} \end{bmatrix}. \qquad (1.88)$$

Визначивши через $\mathbf{x}_i$ i-й стовпець матриці $\mathbf{X}$, можна записати, що

$$\mathbf{x}_i^{\mathrm{т}} \mathbf{x}_j = \begin{cases} 0, & i \neq j; \\ 1, & i = j, \end{cases}$$

тобто $\mathbf{X}$ – ортогональна матриця і $\mathbf{X}^{\mathrm{т}} = \mathbf{X}^{-1}$.

Якщо тепер утворити добуток $\mathbf{X}^{\mathrm{т}}\mathbf{AX}$, то в силу (1.71) одержимо

$$\mathbf{X}^{\mathrm{т}}\mathbf{AX} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \qquad (1.89)$$

Якщо розглянути вектор-стовпець **z**, що складається з трьох елементів, і ортогональне перетворення $\mathbf{y} = \mathbf{X}^{\mathrm{т}}\mathbf{z}$, то з використанням (1.89) одержимо $\mathbf{zA}^{\mathrm{т}}\mathbf{z} = \mathbf{y}^{\mathrm{т}}\mathbf{X}^{\mathrm{т}}\mathbf{AXz} = y_1^2$.

Таким чином, показано, що за допомогою ідемпотентної матриці $\mathbf{A}$ рангу 1 можна задати таке ортогональне перетворення $\mathbf{y} = \mathbf{X}^{\mathrm{т}}\mathbf{z}$, що $\mathbf{zA}^{\mathrm{т}}\mathbf{z} = y_1^2$.

1.9. Диференціальне числення в матричних позначеннях

У наступних розділах нам прийдеться диференціювати деякі вирази, що містять вектори і матриці. Тому встановимо декілька основних результатів. Розглянемо спочатку простий лінійний вираз

$$\mathbf{a}^{\mathrm{т}}\mathbf{x} = \begin{bmatrix} a_1 & a_2 & \ldots & a_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{bmatrix} = a_1 x_1 + a_2 x_2 + \ldots + a_n x_n . \quad (1.90)$$

Якщо обчислити частинні похідні від виразу (1.90) по скалярним змінним x_s $\left(i = \overline{1, n}\right)$, то одержимо

$$\frac{\partial\left(\mathbf{a}^{\mathrm{т}}\mathbf{x}\right)}{\partial x_1} = a_1, \quad \frac{\partial\left(\mathbf{a}^{\mathrm{т}}\mathbf{x}\right)}{\partial x_2} = a_2, \quad \ldots \quad , \quad \frac{\partial\left(\mathbf{a}^{\mathrm{т}}\mathbf{x}\right)}{\partial x_{\mathrm{n}}} = a_n .$$

Як бачимо, частинні похідні є елементами вектора $\mathbf{a}$. Тому, якщо взяти частинні похідні від виразу (1.90), а потім скласти з них вектор $\mathbf{a}$, то можна вважати цей процес як *операцію диференціювання по вектору* $\mathbf{x}$:

$$\frac{\partial\left(\mathbf{a}^{\mathrm{т}}\mathbf{x}\right)}{\partial \mathbf{x}} = \mathbf{a} , \quad (4.91)$$

де ліва частина означає операцію диференціювання.

Розглянемо тепер квадратичну форму:

$$\mathbf{x}^{\mathrm{т}}\mathbf{A}\mathbf{x} = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{bmatrix} =$$

$$\begin{aligned} &= a_{11}x_1^2 + 2a_{12}x_1x_2 + 2a_{13}x_1x_3 + \ldots + 2a_{1n}x_1x_n + \\ &\qquad + a_{22}x_2^2 + 2a_{23}x_2x_3 + \ldots + 2a_{2n}x_2x_n + \\ &\qquad\qquad \cdots \\ &\qquad\qquad + a_{22}x_2^2 . \end{aligned}$$

Візьмемо частині похідні по елементах вектора **x** і одержимо:

$$\frac{\partial}{\partial x_1}\left(\mathbf{x}^{\mathrm{т}}\mathbf{A}\mathbf{x}\right)=2\left(a_{11}x_1+a_{12}x_2+a_{13}x_3\ldots+a_{1n}x_n\right);$$

$$\frac{\partial}{\partial x}\left(\mathbf{x}^{\mathrm{т}}\mathbf{A}\mathbf{x}\right)=2\left(a_{12}x_1+a_{22}x_2+a_{23}x_3\ldots+a_{2n}x_n\right);$$

$$\ldots$$

$$\frac{\partial}{\partial x}\left(\mathbf{x}^{\mathrm{т}}\mathbf{A}\mathbf{x}\right)=2\left(a_{1n}x_1+a_{2n}x_2+a_{3n}x_3\ldots+a_{nn}x_n\right);$$

За винятком множника 2, праві частини цих рівнянь містять елементи матричного добутку $\mathbf{Ax}$, що утворюють *n-мірний* вектор-стовпець. З іншого боку, праві частини цих рівнянь можна розглядати як елементи матричного добутку $\mathbf{x}^{\mathrm{т}}\mathbf{A}$, що утворюють вектор-рядок із n елементів.

Отже,

$$\frac{\partial}{\partial x_1}\left(\mathbf{x}^{\mathrm{т}}\mathbf{A}\mathbf{x}\right)=2\mathbf{A}\mathbf{x}\;; \tag{1.92}$$

$$\frac{\partial}{\partial x_1}\left(\mathbf{x}^{\mathrm{т}}\mathbf{A}\mathbf{x}\right)=2\mathbf{x}^{\mathrm{т}}\mathbf{A}\;. \tag{1.93}$$

Вибір між (1.92) і (1.93) звичайно визначається контекстом, у рамках якого здійснюється диференціювання, оскільки в матричних рівняннях можна порівнювати тільки матриці одного розміру і не можна вектор-рядок порівняти з вектором-стовпцем.

Розглянемо випадок, коли **y** позначає *n-мірний* вектор-стовпець, кожний з елементів якого є функція m елементів вектора **x**, тобто

$$y_i=f_i\left(x_1,x_2,...,x_n,\right),\quad i=1,2,...,n\,.$$

Кожну з функцій y_i можна продиференціювати частинним способом по кожній змінній x_i, що приведе до mn частинних похідних. Якщо ці частинні похідні розташувати у вигляді матриці розміру $m \times n$, то результат виглядатиме таким чином:

$$\frac{\partial \mathbf{y}}{\partial \mathbf{x}} = \begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_2}{\partial x_1} & \cdots & \frac{\partial y_n}{\partial x_1} \\ \frac{\partial y_1}{\partial x_2} & \frac{\partial y_2}{\partial x_2} & \cdots & \frac{\partial y_n}{\partial x_2} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial y_1}{\partial x_m} & \frac{\partial y_2}{\partial x_m} & \cdots & \frac{\partial y_n}{\partial x_m} \end{bmatrix}. \tag{1..94}$$

Для функції $y = f(x_1, x_2, ..., x_n,)$ можна скласти квадратну матрицю других частинних похідних

$$\frac{\partial^2 y}{\partial \mathbf{x}^2} = \frac{\partial(\partial y / \partial \mathbf{x})}{\partial \mathbf{x}} = \begin{bmatrix} \frac{\partial^2 y}{\partial x_1^2} & \frac{\partial^2 y}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 y}{\partial x_1 \partial x_n} \\ \frac{\partial^2 y}{\partial x_2 \partial x_1} & \frac{\partial^2 y}{\partial x_2^2} & \cdots & \frac{\partial^2 y}{\partial x_2 \partial x_n} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial^2 y}{\partial x_n \partial x_1} & \frac{\partial^2 y}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 y}{\partial x_n^2} \end{bmatrix}. \tag{1..95}$$

Згідно з тим, для змішаних частинних похідних має місце рівність

$$\frac{\partial^2 y}{\partial x_i \partial x_j} = \frac{\partial^2 y}{\partial x_j \partial x_i} ,$$

матриця других частинних похідних (1.95) є симетричною.

1.10. Контрольні запитання, приклади і вправи

1. Дати визначення поняттям «матриця» і «елемент матриці».

2. Яку матрицю називають квадратною? Діагональною? Си–метричною? Одиничною?

3. Дати визначення вектору-рядку і вектору-стовпцю як окремим випадкам матриці.

4. У чому полягає операція транспонування матриці і як вона позначається?

5. Транспонувати матрицю $\mathbf{A} = \begin{bmatrix} 7 & 5 \\ 0 & -2 \\ 3 & -2 \end{bmatrix}$.

Розв'язання. У цьому завданні необхідно одержати матрицю $\mathbf{A}^{\mathrm{T}}$, елементи якої a_{ij} $(i = 1,2;\ \ j = 1,2,3)$ рівні відповідним елементам a_{ij} початкової матриці. Для цього досить усі рядки початкової матриці послідовно подати у вигляді стовпців шуканої матриці:

$$\mathbf{A}^{-1} = \begin{bmatrix} 7 & 0 & 3 \\ 5 & -2 & -2 \end{bmatrix}.$$

6. Транспонувати матриці:

$$\mathbf{K} = \begin{bmatrix} 3 & 7 & -2 \\ 0 & 8 & 1 \end{bmatrix};\quad \mathbf{L} = \begin{bmatrix} 7 & 8 \\ 9 & 1 \end{bmatrix};\quad \mathbf{M} = \begin{bmatrix} 12 & 0 & 5 \end{bmatrix};\quad \mathbf{N} = \begin{bmatrix} 3 \\ 4 \end{bmatrix};\quad \mathbf{E} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

8. Які матриці є рівними? Навести приклад рівних матриць.

9. У чому полягає операція додавання матриць?

10. Знайти суму матриць $\mathbf{A} = \begin{bmatrix} 1 & 5 \\ 6 & 7 \\ 0 & -2 \end{bmatrix}$ і $\mathbf{B} = \begin{bmatrix} 2 & 4 \\ 3 & 0 \\ 0 & -1 \end{bmatrix}$.

Розв'язання. За виначенням скласти дві матриці **A** и **B** – це значить знайти таку третю матрицю **C**, елементи якої визначаються рівністю $c_{ij} = a_{ij} + b_{ij}$. Для цього прикладу шуканою буде матриця $\mathbf{C} = \begin{bmatrix} 3 & 3 \\ 6 & 5 \\ 3 & 1 \end{bmatrix}$.

10. Обчислити суму матриць:

$$\mathbf{C} = \begin{bmatrix} 2 & -1 & 3 \\ 0 & -1 & 2 \end{bmatrix}; \quad \mathbf{D} = \begin{bmatrix} 2 & -1 & 3 \\ 0 & -1 & 2 \end{bmatrix}; \quad \mathbf{F} = \begin{bmatrix} 2 & -1 & 3 \\ 0 & -1 & 2 \end{bmatrix}.$$

11. Знайти різницю матриць **C** і **B**, наведених у попередній вправі.

12. Визначити (**T**+**C**), якщо $\mathbf{T} = \begin{bmatrix} 1 & 0 \\ 2 & -1 \end{bmatrix}; \quad \mathbf{S} = \begin{bmatrix} -1 & 1 \\ 1 & 0 \end{bmatrix}$.

13. У чому полягає операція множення матриць?

14. Матриці якої розмірності можуть бути перемножені? Яким способом визначається розмірність результуючої матриці?

15. Чи має місце операція множення між транспонованою і початкової матрицями?

16. Яку розмірність має добуток транспонованої і початкової матриць?

17. Добуток транспонованої і початкової матриці зобов'язаний являти собою квадратну матрицю чи ні?

18. Чи справедливий комутативний закон для операції множення матриць? Навести приклад, що підтверджує відповідь.

19. Знайти добуток матриць $\mathbf{A} = \begin{bmatrix} 1 & 0 & 2 \\ -2 & 3 & 0 \end{bmatrix}$ і $\mathbf{B} = \begin{bmatrix} 2 & 4 \\ 3 & 0 \\ 0 & -1 \end{bmatrix}$.

Розв'язання. Дві матриці **A** і **B** можуть бути перемножені, якщо число стовпців першої матриці збігається з числом рядків другої. У даному випадку ця умова дотримується: у матриці **A** – три стовпці, а в матриці **B** – стільки ж рядків.

За визначенням операції множення елементи результуючої матриці **D** обчислюється за формулою , де в умовах задачі $n=3$. Елементи шуканої матриці обчислюються таким чином:

$$d_{11} = a_{11}b_{11} + a_{12}b_{21} + a_{13}b_{31} = 1 \cdot 2 + 0 \cdot 3 + 2 \cdot 0 = 2\,;$$
$$d_{12} = a_{11}b_{12} + a_{12}b_{22} + a_{13}b_{32} = 1 \cdot 4 + 0 \cdot 0 + 2 \cdot (-1) = 2\,;$$
$$d_{21} = a_{21}b_{11} + a_{22}b_{21} + a_{23}b_{31} = (-2) \cdot 2 + 3 \cdot 3 + 0 \cdot 0 = 5\,;$$
$$d_{22} = a_{21}b_{12} + a_{22}b_{22} + a_{23}b_{32} = (-2) \cdot 4 + 3 \cdot 0 + 0 \cdot (-1) = -8\,,$$

а сама результуюча матриця має вигляд: $\mathbf{D} = \begin{bmatrix} 2 & 2 \\ 5 & -8 \end{bmatrix}$.

20. Перемножити матриці $\mathbf{A} = \begin{bmatrix} 2 & -2 & -2 \\ 1 & 0 & 3 \end{bmatrix}$ і $\mathbf{B} = \begin{bmatrix} -5 \\ -2 \\ 0 \end{bmatrix}$.

21. Знайти добуток $\mathbf{A} \times \mathbf{У} \times \mathbf{C}$, якщо $\mathbf{A} = \begin{bmatrix} 2 & -1 \\ 0 & 3 \end{bmatrix}$, $\mathbf{B} = \begin{bmatrix} -4 & -1 & 0 \\ -2 & 1 & -2 \end{bmatrix}$, $\mathbf{C} = \begin{bmatrix} 1 & 0 \\ -1 & 2 \\ 0 & 1 \end{bmatrix}$.

22. Чи рівні матричні вирази $(\mathbf{A} + \mathbf{B})(\mathbf{A} - \mathbf{B})$ і $(\mathbf{A} - \mathbf{B})(\mathbf{A} + \mathbf{B})$? Якщо ні, то чому? Скільки членів міститься в кожному з них?

23. Дано: $\mathbf{A} = \begin{bmatrix} 1 & 0 & 2 \\ 2 & -1 & 1 \end{bmatrix}$; $\mathbf{B} = \begin{bmatrix} 3 & 0 & 1 \\ 0 & 1 & -1 \\ 1 & 2 & 0 \end{bmatrix}$; $\mathbf{C} = \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix}$.

Обчислити $(\mathbf{AB})^{\mathrm{т}}$, $\mathbf{B}^{\mathrm{т}}\mathbf{A}^{\mathrm{т}}$, $(\mathbf{AC})^{\mathrm{т}}$ і $\mathbf{C}^{\mathrm{т}}\mathbf{A}^{\mathrm{т}}$.

24. Знайти всі матриці $\mathbf{B}$, що задовольняють рівнянню $\begin{bmatrix} 0 & 1 \\ 0 & 2 \end{bmatrix} \times \mathbf{B} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 2 \end{bmatrix}$.

25. Знайти всі матриці $\mathbf{B}$, комутативні щодо множення матриці $\mathbf{A}=\begin{bmatrix}0 & 1\\ 0 & 2\end{bmatrix}$, тобто $(\mathbf{AB})=(\mathbf{BA})$.

26. Знайти квадрат матриці $\mathbf{A}=\begin{bmatrix}3 & 0 & 1\\ 0 & 1 & -1\\ 1 & 2 & 0\end{bmatrix}$. Знайти куб матриці $\mathbf{A}$ двома способами: $\mathbf{A}(\mathbf{A})^2$ і $(\mathbf{A})^2\mathbf{A}$.

27. Довести, що діагональна матриця комутативна щодо множення будь-якої матриці того ж порядку.

28. Нехай $\mathbf{J}=\begin{bmatrix}0 & 0 & 1\\ 0 & 1 & 0\\ 1 & 0 & 0\end{bmatrix}$. Записати декілька добутків, де $\mathbf{A}$ – прямокутна матриця. Описати в словесній формі вплив, який справляє матриця $\mathbf{J}$ на матрицю $\mathbf{A}$. Знайти $\mathbf{J}^2$.

29. Знайти $\mathbf{V}^2$ і $\mathbf{V}^3$, якщо $\mathbf{V}=\begin{bmatrix}0 & 1 & 0\\ 0 & 0 & 1\\ 0 & 0 & 0\end{bmatrix}$. Розглянути добутки типу $\mathbf{VA}, \mathbf{V}^2\mathbf{A}$ і $\mathbf{V}^{\mathrm{T}}\mathbf{A}$.

30. Що називають слідом матриці?

31. Довести, що коли матриці $\mathbf{A}$ і $\mathbf{B}$ такі, що $\mathbf{AB}$ і $\mathbf{BA}$ існують одночасно, то $\mathbf{AB}$ і $\mathbf{BA}$ мають однакову суму діагональних елементів $(\mathrm{tr}(\mathbf{AB})=\mathrm{tr}(\mathbf{BA}))$.

32. Яку функцію називають лінійною? Квадратичною? Навести приклади.

33. Що називають лінійною формою? Квадратичною формою?

34. Навести загальну форму запису в матричному вигляді лінійної функції.

35. Чи може лінійна функція не містити лінійну форму?

36. Навести загальну форму запису в матричному вигляді квадратичної функції.

37. Чи може квадратична функція не містити лінійну форму?

38. Як обчислюються діагональні й недиагональні елементи матриці коефіцієнтів квадратичної форми?

39. Подати квадратичну функцію $y(x_1, x_2, x_3) =$ $= x_1^2 - 2x_2^2 + 4x_1x_3 + x_2 + 1$ в матричному вигляді.

Розв'язання. Виділимо у заданій квадратичній функції квадратичну форму, лінійну форму і вільний член: $x_1^2 - 2x_2^2 + 4x_1x_3$ – квадратична форма; x_2 - лінійна форма; 1 – вільний член.

У матричному вигляді квадратична функція записується таким чином: , де перший додаток – квадратична форма; другий – лінійна форма; третій – вільний член.

Почнемо перетворення заданої квадратичної функції у матричний вигляд із квадратичної форми. Оскільки функція залежить від трьох змінних, то вектор-рядок $\overline{x}^{\mathbf{T}}$ у квадратичній формі являє собою матрицю $\begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix}$, тобто включає три елементи. З тієї ж причини вектор-стовпець $\overline{x}$ у квадратичній і лінійній формах також буде містити три елементи: $\overline{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$. Матриця $\mathbf{A}$ – симетрична матриця, тобто $a_{ij} = a_{ji}$. Її елементи a_{ij} при $i \neq j$ рівні відповідним коефіцієнтам b_{ij} при квадратичному члені $x_i x_j$ у заданій функції, а при $i = j$ рівні подвоєному коефіцієнту при квадратичному члені x_i^2. Перетворимо квадратичну форму початкової функції у форму, зручну для формування матриці $\mathbf{A}$:

$$x_1^2 - 2x_2^2 + 4x_1x_3 = x_1^2 - 2x_2^2 + 0x_2^2 + 0x_1x_2 + 4x_1x_3 + 0x_2x_3 \ .$$

Відповідно до отриманої форми визначимо елементи матриці $\mathbf{A}$:

$a_{11} = 2 \cdot 1 = 2$; $\quad a_{22} = 2 \cdot (-2) = -4$; $\quad a_{33} = 2 \cdot 0 = 0$;

$a_{12} = a_{21} = 0$; $\quad a_{13} = a_{31} = 4$; $\quad a_{23} = a_{32} = 0$.

Сформуємо шукану матрицю квадратичної форми: $\mathbf{A} = \begin{bmatrix} 2 & 0 & 4 \\ 0 & -4 & 0 \\ 4 & 0 & 0 \end{bmatrix}$.

Тепер переходимо до перетворення лінійної форми заданої функції. Подамо її у вигляді, зручному для формування вектора-рядка $\overline{a}^{\text{т}} = [a_1 \quad a_2 \quad a_3]$: $\quad x_2 = 0x_1 + 1x_2 + 0x_3$. Звідси $a_1 = 0$; $a_2 = 1$; $a_3 = 0$. Отже, $\overline{a}^{\text{т}} = [0 \quad a_2 \quad 0]$.

Вільний член залишається без перетворень: $a_0 = 1$.

Отже, всі компоненти квадратичної функції, необхідні для її подання в матричному вигляді, отримані. Запишемо остаточний результат:

$$y(\overline{x}) = \frac{1}{2}[x_1 \quad x_2 \quad x_3] \begin{bmatrix} 2 & 0 & 4 \\ 0 & -4 & 0 \\ 4 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + [x_1 \quad x_2 \quad x_3] \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + 1.$$

40. Подати **у** матричному вигляді лінійну функцію

$y(x_1, x_2) = x_1 - 2x_2 + 3$ *юююю* .

41. Записати в матричному вигляді квадратичну функцію $y(x_1, x_2)$ $= x_1^2 - x_1 x_2 + 4x_2 - 3$.

42. Подати квадратичну функцію $y(x_1, x_2, x_3) =$ $= 2x_2^2 - x_3^2 - 4x_1 x_3 - x_2 x_3 + x_1 - 2x_3 + 7$ в матричному вигляді.

43. Надані матриці $\mathbf{A}=\begin{bmatrix}1&3&2\\2&6&9\\7&6&1\end{bmatrix}$ і $\mathbf{E}=\begin{bmatrix}0&1&0\\1&0&0\\0&0&1\end{bmatrix}$. Обчислити $|\mathbf{A}|$, $|\mathbf{E}|$ і $|\mathbf{B}|$, де $\mathbf{B}=\mathbf{EA}$. Перевірити рівність $|\mathbf{B}|=|\mathbf{E}|\,|\mathbf{A}|$.

44. Показати, що $\begin{bmatrix}1&1&1\\a&b&c\\a^2&b^2&c^2\end{bmatrix}=(c-b)(c-a)(b-a)$.

45. Показати, що коли (x_1,y_1) і (x_2,y_2) – крапки на площині x,y, то рівняння $\begin{vmatrix}x&y&1\\x_1&y_1&1\\x_2&y_2&1\end{vmatrix}=0$ визначає пряму лінію, яка проходить через ці дві крапки.

46. Довести, що $\begin{bmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\0&a_{22}&\cdots&a_{2n}\\\cdots&\cdots&\cdots&\cdots\\0&0&\cdots&a_{nn}\end{bmatrix}=a_{11}a_{22}\cdots a_{nn}$.

47. Довести, що визначник кососиметричної матриці непарного порядку дорівнює нулю. (Кососиметричною називається матриця $\mathbf{A}$, якщо $\mathbf{A}^{\mathrm{т}}=-\mathbf{A}$).

48. Показати, що матриця $\mathbf{Q} = \begin{bmatrix} \frac{1}{\sqrt{6}} & \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{30}} \\ \frac{-2}{\sqrt{6}} & \frac{1}{\sqrt{5}} & \frac{-2}{\sqrt{30}} \\ \frac{1}{\sqrt{6}} & 0 & \frac{-5}{\sqrt{30}} \end{bmatrix}$ ортогональна, тобто $\mathbf{Q}^{\mathrm{т}} = \mathbf{Q}^{-1}$.

49. Дано: $\mathbf{X} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 1 \\ 1 & 3 \end{bmatrix}$. Обчислити $\mathbf{A} = \left(\mathbf{I}_4 - \mathbf{X}\left(\mathbf{X}^{\mathbf{т}}\mathbf{X}\right)^{\mathbf{-1}}\mathbf{X}^{\mathbf{т}}\right)$. Показати, що матриця $\mathbf{A}$ ідемпотентна і визначити її ранг. Знайти власні значення і відповідні їм власні вектори матриці $\mathbf{A}$, а потім побудувати діагональну матрицю, що приводить $\mathbf{A}$ до діагонального вигляду.

50. Поширити результат вправи 9 на випадок трьох матриць і показати, що $\operatorname{tr}(\mathbf{ABC}) = \operatorname{tr}(\mathbf{BCA}) = \operatorname{tr}(\mathbf{CAB})$, де матриці підібрані так, що всі добутки існують.

51. Показати, що $\rho\left(\mathbf{BB}^{\mathrm{т}}\right) = s$, якщо $\mathbf{B}$ – матриця розміру $n \times s$ і рангу s, де $s \le n$.

52. Довести, що $\mathbf{AB}$ і $\mathbf{BA}$ мають однакові власні значення, якщо $\mathbf{A}$ и $\mathbf{B}$ – невироджені матриці одного порядку. Показати також, що не існує матриць, які задовольняли рівнянню $\mathbf{AB} - \mathbf{BA} = \mathbf{I}$.

53. Знайти власні значення і власні вектори матриці $\mathbf{A} = \begin{bmatrix} 5 & -6 & -6 \\ -1 & 4 & 2 \\ 3 & -6 & -4 \end{bmatrix}$.

54. Перевірити наступні квадратичні форми на позитивну визначеність:

а) $6x_1^2 + 49x_2^2 + 51x_3^2 - 82x_1x_2 + 20x_1x_3 - 4x_1x_2$;

б) $4x_1^2 + 9x_2^2 + 2x_3^2 + 8x_2x_3 + 6x_1x_3 + 6x_2x_1$.

55. Знайти власні значення матриці $\begin{bmatrix} 3 & 1 & 1 \\ 0 & -2 & 1 \\ 0 & 0 & 2 \end{bmatrix}$.

56. Довести, що коли власні вектори $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m$ матриці $\mathbf{A}$ відповідають різним власним значенням, то вектори $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m$ лінійно незалежні.

57. Довести, що коли $\mathbf{x}$ – власний вектор невиродженої матриці $\mathbf{A}$, то він є власним вектором і матриці $\mathbf{A}^{-1}$.

58. Дано: $\mathbf{A} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & -1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$; $\mathbf{B} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & -1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$. Визначити $\mathbf{A}^n$ і $\mathbf{B}^n$ для $n > 1$.

59. Показати, що матриця $\mathbf{A} = \begin{bmatrix} \frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{1}{3} & -\frac{2}{3} \\ \frac{2}{3} & -\frac{2}{3} & \frac{1}{3} \end{bmatrix}$ ортогональна.

60. Довести, що добуток двох ортогональних матриць одного порядку є ортогональна матриця.

1.11. Фонд індивідуальних завдань

Індивідуальне завдання №1. Знайти транспоновану матрицю $\mathbf{A}^{\mathrm{т}}$, якщо початкова матриця визначається таким чином:

1. $\mathbf{A}=\begin{bmatrix}1 & 0 & 2\\ 3 & -1 & 1\\ 0 & 2 & 1\end{bmatrix}$;

2. $\mathbf{A}=\begin{bmatrix}2 & 0 & 1\\ 1 & 4 & 0\\ 0 & -1 & 1\end{bmatrix}$;

3. $\mathbf{A}=\begin{bmatrix}1 & -1 & 0\\ 0 & 2 & 1\\ -2 & 0 & 2\end{bmatrix}$;

4. $\mathbf{A}=\begin{bmatrix}-1 & 1 & 2\\ 2 & 2 & 0\\ 0 & -1 & 0\end{bmatrix}$;

5. $\mathbf{A}=\begin{bmatrix}1 & 3 & 0\\ 0 & 1 & -1\\ 1 & 3 & -1\end{bmatrix}$;

6. $\mathbf{A}=\begin{bmatrix}-1 & 1 & 0\\ 0 & 1 & -2\\ 2 & 0 & -3\end{bmatrix}$;

7. $\mathbf{A}=\begin{bmatrix}-1 & -1 & 0\\ 0 & 1 & -2\\ 2 & 0 & -3\end{bmatrix}$;

8. $\mathbf{A}=\begin{bmatrix}1 & -1 & 0\\ 1 & 0 & -2\\ 0 & 2 & -3\end{bmatrix}$;

9. $\mathbf{A}=\begin{bmatrix}1 & 0 & -2\\ 1 & -1 & 0\\ 0 & 2 & -3\end{bmatrix}$.

10. $\mathbf{A}=\begin{bmatrix}-2 & 1 & 0\\ 0 & 1 & -1\\ -3 & 0 & 2\end{bmatrix}$.

11. $\mathbf{A}=\begin{bmatrix}3 & 3 & 1\\ 7 & 6 & 2\\ 7 & 9 & 2\end{bmatrix}$;

12. $\mathbf{A}=\begin{bmatrix}1 & 1 & 1\\ 3 & 4 & 3\\ 9 & 8 & 5\end{bmatrix}$;

13. $\mathbf{A}=\begin{bmatrix}3&5&4\\8&9&7\\5&6&4\end{bmatrix}$;

14. $\mathbf{A}=\begin{bmatrix}1&0&3\\3&1&7\\2&1&8\end{bmatrix}$;

15. $\mathbf{A}=\begin{bmatrix}3&5&5\\7&10&12\\9&6&7\end{bmatrix}$;

16. $\mathbf{A}=\begin{bmatrix}1&4&6\\5&9&9\\1&7&10\end{bmatrix}$;

17. $\mathbf{A}=\begin{bmatrix}1&-1&3\\2&3&3\\1&-2&-1\end{bmatrix}$;

18. $\mathbf{A}=\begin{bmatrix}5&4&2\\8&9&4\\13&9&5\end{bmatrix}$;

19. $\mathbf{A}=\begin{bmatrix}5&4&5\\8&7&9\\1&2&2\end{bmatrix}$;

20. $\mathbf{A}=\begin{bmatrix}4&3&2\\4&5&2\\3&2&3\end{bmatrix}$;

21. $\mathbf{A}=\begin{bmatrix}-2&3&4\\3&-1&-3\\-1&2&2\end{bmatrix}$;

22. $\mathbf{A}=\begin{bmatrix}2&2&5\\3&3&6\\4&3&4\end{bmatrix}$;

23. $\mathbf{A}=\begin{bmatrix}4&3&6\\4&5&10\\5&6&10\end{bmatrix}$;

24. $\mathbf{A}=\begin{bmatrix}5&1&-2\\9&3&-1\\0&4&-1\end{bmatrix}$;

25. $\mathbf{A}=\begin{bmatrix}6&0&3\\2&1&0\\2&2&1\end{bmatrix}$;

26. $\mathbf{A}=\begin{bmatrix}2&5&1\\1&0&5\\0&-1&1\end{bmatrix}$;

27. $\mathbf{A} = \begin{bmatrix} 1 & 8 & 2 \\ 3 & -1 & 1 \\ 0 & 2 & 1 \end{bmatrix}$; **28.** $\mathbf{A} = \begin{bmatrix} 2 & 4 & 1 \\ 1 & 4 & 0 \\ 0 & 4 & -1 \end{bmatrix}$;

29. $\mathbf{A} = \begin{bmatrix} 1 & 0 & 2 \\ -3 & -3 & 3 \\ 0 & 2 & 1 \end{bmatrix}$; **30.** $\mathbf{A} = \begin{bmatrix} 2 & 0 & -1 \\ 1 & 4 & -1 \\ 0 & -1 & -1 \end{bmatrix}$.

Індивідуальне завдання №2. Знайти суму матриць $\mathbf{A}$ і $\mathbf{B}$, де матриця $\mathbf{A}$ відповідає однойменній матриці індивідуального завдання №1, а матриця $\mathbf{B} = \begin{bmatrix} -2 & 0 & 1 \\ 1 & -1 & 2 \\ 0 & 3 & 2 \end{bmatrix}$.

Індивідуальне завдання №3. Знайти добутки матриць $\mathbf{AB}$, $\mathbf{AC}$, $\mathbf{ABC}$, $\mathbf{BA}$, $\mathbf{BAC}$, де матриця $\mathbf{A}$ відповідає однойменній матриці індивідуального завдання №1, матриця $\mathbf{B}$ – індивідуального завдання №2, $\mathbf{C} = \begin{bmatrix} 1 & -2 & 0 \\ 0 & 1 & -1 \end{bmatrix}$.

Індивідуальне завдання №4. Подати квадратичну функцію y у матричному вигляді, якщо функція визначається таким чином:

4.1. $y = 2x_1^2 + 2x_2^2 + x_1 + 1$;

4.2. $y = -2x_1^2 + x_1x_2 + x_2$;

4.3. $y = 2x_1^2 - 2x_1x_2 - 2x_2 - 1$;

4.4. $y = x_1^2 + x_2^2 - 2x_3^2 - 2x_1x_3 + x_2$;

4.5. $y = x_1^2 - 2x_2^2 - 3x_3^2 - 2x_2x_3 - x_1$;

4.6. $y = x_2^2 - x_3^2 - 6x_1x_2 - 6x_1x_3 + 1$;

4.7. $y = x_1^2 - 2x_2x_3 - 8x_2$;

4.8. $y = 2x_2x_3 + x_1$;

4.9. $y = -x_4^2 + 2x_1x_2 - 4x_3$;

4.10. $y = x_2^2 - 3x_3^2 - 4x_1x_2 + 6x_1x_3 + 2$;

4.11. $y = x_1^2 - 8x_2x_3 + 5x_3 + 7$;

4.12. $y = x_2^2 + x_1 + 2$;

4.13. $y = 2x_1x_2$;

4.14. $y = x_3^2 - 3x_1x_2 - 2x_1$;

4.15. $y = x_2^2 - 5x_3^2 - 4x_1x_3 + 7x_2 - x_3$;

4.16. $y = 2x_1x_2 + 4x_1x_3 - 3x_2x_3 + 5x_2$;

4.17. $y = 3{,}5x_1^2 - 6x_1x_2 + x_3$;

4.18. $y = 2x_3^2 - 2x_1x_2 + x_3 - 3$;

4.19. $y = x_1^2 - 2x_1x_2 + 3x_1x_3 - x_2$;

4.20. $y = x_1^2 + x_2 + 3$;

4.21. $y = x_1x_2 - 4x_1$;

4.22. $y = x_3^2 + x_1x_2 - x_2$;

4.23. $y = x_2^2 + x_3^2 + x_1x_2 + 2x_1x_3 + x_3$;

4.24. $y = x_1^2 - 10x_3^2 + 2x_2x_3 + 4x_1 - x_3 - 1$;

4.25. $y = 2x_1^2 - 3x_1x_2 + 2x_2x_3 - 6x_2$;

4.26. $y = 2x_1^2 - 2x_2^2 - 2x_3^2 - x_1x_2 - x_1x_3 - x_2x_3 - x_1 - x_3 - 2$;

4.27. $y = 3x_1x_2 - 3x_1x_3 - x_2x_3 - x_1 - 2x_2 - x_3$;

4.28. $y = x_1^2 - x_2^2 - x_3^2 - 2x_1x_2 - 2x_1x_3 - 2x_2x_3 - x_1 - x_2 - x_3$;

4.29. $y = 2x_1^2 - 2x_2^2 - 2x_3^2 + x_1x_3 - x_2x_3 - 5$;

.30. $y = -x_1^2 - 2x_2^2 - 3x_3^2 - 3x_1x_2 + 2x_1x_3 - x_2 - x_3 + 3$.

Індивідуальне завдання №5. Знайти визначник матриці $\mathbf{A}$, де матриця $\mathbf{A}$ відповідає однойменній матриці індивідуального завдання №1. Сформувати матрицю $\mathbf{A}^*$, елементами якої є мінори матриці $\mathbf{A}$, тобто $a_{ij}^* = \left|\mathbf{A}_{ij}\right|$.

Індивідуальне завдання №6. Знайти обернену матрицю $\mathbf{A}^{-1}$ для матриці **А, що** відповідає однойменній матриці індивідуального завдання №1. Перевірити рівність добутку $\mathbf{A}^{-1}\mathbf{A}$ одиничній матриці $\mathbf{I}_3$.

Індивідуальне завдання №7. Знайти власні значення і власні вектори для матриці **А,** що відповідає однойменній матриці індивідуального завдання №1. Привести матрицю $\mathbf{A}$ до діагонального вигляду.

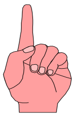
Розділ 2. ЕЛЕМЕНТИ ТЕОРІЇ МНОЖИН

Розглядаються основні поняття теорії множин, способи надання множин і математичні дії над ними. Найбільша увага приділяється точковим множинам як елементарним об'єктам математичного програмування. Даються визначення основним типам точкових множин.

Теорія множин разом із диференціальним і матричним численням являє собою базовий математичний апарат для математичного програмування.

2.1. Поняття множини

При вивченні дисципліни часто доводиться мати справу з деякими словосполученнями (наприклад: множина значень функції, множина припустимих рішень, множина всіх підмножин, непорожня множина, множина точок і т.п.), що тісно зв'язані з поняттям *множини* або *підмножини.*



Визначення 2.1. Множиною називають сукупність об'єктів довільної природи, об'єднаних за якими-небудь ознаками.

Приклади: множина книг у бібліотеці, множина студентів в академічній групі, множина розділів у підручнику, множина будинків у місті.

Множини, як і матриці, позначають жирними прописними латинськими літерами: **A, B, C** і т.д. Об'єкти, з яких складається множина, називають *елементами множини* і позначають малими літерами: *a, b, c*. При позначенні елемента *множини* може бути використана проста або складна індексація.

Множини діляться на *скінченні множини* і *нескінченні.*

Скінченна множина складається із скінченного числа елементів, які можна перелічити за кінцеве число кроків. У протилежному випадку множина вважається нескінченною.

Нескінченні множини, у свою чергу, діляться на *лічені* (*дискретні*) і *нелічені* (*неперервні*).

Лічена множина – це така нескінченна множина, елементи якої можна “перенумерувати”.

Множина задається перерахуванням усіх її елементів або зазначенням характерних властивостей усіх її (і тільки її) елементів.

Якщо множина $\mathbf{A}$ складається з n елементів $a_1, a_2,..., a_n$, то надання множини $\mathbf{A}$ шляхом перерахування всіх її елементів робиться таким чином: $\mathbf{A} = \{a_1, a_2,..., a_n\}$ або $\mathbf{A} = \{a_i\}, \quad i = \overline{1,n}$. При цьому розмір n (загальне число елементів) називають *потужністю* множини і записують як $n = card\,\mathbf{A}$.

Нескінченна множина **B** дійсних чисел з інтервалу чисел (3, 7) може бути задана за допомогою зазначення властивостей усіх його елементів. При цьому математичний запис має вигляд: **B:** $3<x<7$. Тут **B** - позначення множини; двокрапка – символ, що передує зазначенню властивості (у загальному випадку – властивостей) елементів цієї множини; x – довільний елемент множини.

Альтернативне надання множини через зазначення властивостей її елементів може мати вид: $\mathbf{B} = \{x \mid 3 < x < 7\}$. Тут символ “вертикальна риса” розділяє довільний елемент множини і властивість, якою він володіє.

Приналежність елемента *a* множині **A** позначається як . Якщо елемент не належить множині, то .

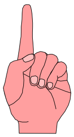
Будь-яка сукупність елементів множин являє собою *підмножину* цієї множини. Підмножини і їхні елементи позначаються так само, як і множини та їхні елементи. Наприклад, множина всіх однозначних чисел натурального ряду **A** має підмножину **B** = { 3, 6, 9}, всі елементи якої кратні 3.

Якщо **B** – підмножина множини **A**, то говорять, що підмножина **B** входить у множину **A** або множину **A** включає підмножина **B**. Для приведених висловлень прийняті відповідні математичні вирази: і .

Якщо $\mathbf{A} \subset \mathbf{B}$ і $\mathbf{B} \subset \mathbf{A}$, то множини **A** і **B** називають рівними і пишуть **A** = **B**. Рівні множини складаються із тих самих елементів.

2.2. Математичні дії над множинами. Порожня множина

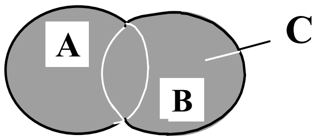
Над множинами можна робити різні математичні операції. Найпростіші з них: *об'єднання, переріз, різниця і декартовий добуток.*



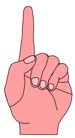
Визначення 2.2. Об'єднанням двох множин **A** і **B** називають таку третю множину **C**, що створена з усіх елементів, які належать хоча б одній з множин **A** і **B,** тобто $\mathbf{C} = \left\{ x \mid x \in \mathbf{A} \text{ или } x \in \mathbf{B} \right\}$.

Математичний запис операції об'єднання двох множин **A** і **B** і графічна інтепретація цієї операції показані на рис.2.1.

$$\mathbf{C} = \mathbf{A} \cup \mathbf{B}$$



Ри.2.1.



Визначення 2.3. Перерізом двох множин **A** і **B** називають таку третю множину **D**, що складається з елементів, які одночасно входять в обидві множини **A** і **B**, тобто $\mathbf{D} = \left\{ x \mid x \in \mathbf{A} \text{ и } x \in \mathbf{B} \right\}$.

Математичний запис операції об’єднання двох множин **A** і **B** та графічна інтепретація операції показані на рис.2.2.

$$\boxed{\mathbf{D} = \mathbf{A} \bigcap \mathbf{B}}$$

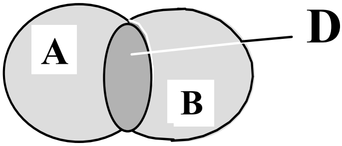


Рис.2.2.

Поняття об'єднання і перерізу можуть бути узагальнені на випадок будь-якого числа множин (скінченного або нескінченного). Якщо дано множини $\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_k, \ldots$, то символічний запис $\bigcup\limits_{k=1}^{\infty} \mathbf{A}_k$ означає об'єднання цих множин, тобто визначає множину, кожний елемент якої належить хоча б одній з даних множин. Символічний запис $\bigcap\limits_{k=1}^{\infty} \mathbf{A}_k$ означає переріз цих множин, тобто визначає множину, кожний елемент якої належить усім даним множинам.

Операції об'єднання і перерізу множин мають властивості *комутативності, асоційованості* і *дистрибутивності.*

Властивість комутативності визначається рівностями:

$$\mathbf{A} \bigcup \mathbf{B} = \mathbf{B} \bigcup \mathbf{A} \ ;$$

$$\mathbf{A} \bigcap \mathbf{B} = \mathbf{B} \bigcap \mathbf{A} \ .$$

Властивість асоційованості визначається рівностями:

$$(\mathbf{A} \bigcup \mathbf{B}) \bigcup \mathbf{C} = \mathbf{A} \bigcup (\mathbf{B} \bigcup \mathbf{C}) \ ;$$

$$(\mathbf{A} \bigcap \mathbf{B}) \bigcap \mathbf{C} = \mathbf{A} \bigcap (\mathbf{B} \bigcap \mathbf{C}) \ .$$

Властивість дистрибутивності визначається рівностями:

$$(\mathbf{A} \bigcup \mathbf{B}) \bigcap \mathbf{C} = (\mathbf{A} \bigcap \mathbf{C}) \bigcup (\mathbf{B} \bigcap \mathbf{C}) \ ;$$

$$(\mathbf{A} \bigcap \mathbf{B}) \bigcup \mathbf{C} = (\mathbf{A} \bigcup \mathbf{C}) \bigcap (\mathbf{B} \bigcup \mathbf{C}) \ .$$

Визначення 2.4. Різницею множин $\mathbf{A}$ і $\mathbf{B}$ називають таку третю множину $\mathbf{F}$, що складається з елементів множини $\mathbf{A}$, які не належать множині **B.**

Математичний запис операції різниці множин та графічна інтепретація показані на рис.2.3.

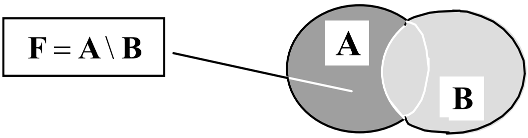


Рис.2.3.

Визначення 2.5. Якщо $\mathbf{B} \subset \mathbf{A}$, то різниця $\mathbf{F} = \mathbf{A} \setminus \mathbf{B}$ називають доповненням множини $\mathbf{B}$ до множини $\mathbf{A}$ і позначають $\mathbf{C_A B}$.

Графічна інтепретація доповнення множини $\mathbf{B}$ до множини $\mathbf{A}$ показана на рис.2.4.

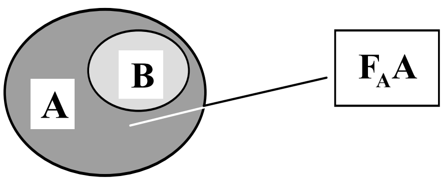


Рис.2.4.

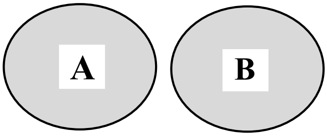
Для операції доповнення справедливі закони подвійності. Якщо і $\mathbf{B} \subset \mathbf{E}$, то

$$\mathbf{E} \setminus (\mathbf{A} \cup \mathbf{B}) = (\mathbf{E} \setminus \mathbf{A}) \cap (\mathbf{E} \setminus \mathbf{B}) \text{ або } \mathbf{C_E}(\mathbf{A} \cup \mathbf{B}) = \mathbf{C_E A} \cap \mathbf{C_E B};$$

$$\mathbf{E} \setminus (\mathbf{A} \cap \mathbf{B}) = (\mathbf{E} \setminus \mathbf{A}) \cup (\mathbf{E} \setminus \mathbf{B}) \text{ або } \mathbf{C_E}(\mathbf{A} \cap \mathbf{B}) = \mathbf{C_E A} \cup \mathbf{C_E B}.$$

Визначення 2.6. Множину, що не включає жодного елемента, називають порожньою і позначають символом $\varnothing$.

На рис.2.5. показаний приклад порожньої множини.



$$\mathbf{A} \cap \mathbf{B} = \varnothing$$

Рис.2.5.

Будь-яка множина містить $\varnothing$ як підмножини. Очевидно, $\mathbf{A} \subset \mathbf{A}$, тобто множина **A** включає рівні собі підмножину **A** і порожню підмножину $\varnothing$. Ці підмножини називають *невласними підмножинами* множини **A.** Всі інші підмножини називають *власними*.

Визначення 2.7. Декартовим добутком множин **A** і **B** називають таку третю множину **G**, що складається з упорядкованих пар елементів *(a, b)*, де $a \in \mathbf{A}$, $b \in \mathbf{B}$. Елементи a і b називають при цьому компонентами або координатами пари *(a, b)*.

Декартовий добуток $\mathbf{A}_1 \times \mathbf{A}_2 \times \ldots \times \mathbf{A}_n$ множин $\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_n$ являє собою множини всіх упорядкованих підмножин з елементів $(a_1, a_2, \ldots, a_n)$, де $a_1 \in \mathbf{A}_1$, $a_2 \in \mathbf{A}_2$, … , $a_n \in \mathbf{A}_n$... Зокрема, декартовий добуток $\underbrace{\mathbf{R} \times \mathbf{R} \times \ldots \times \mathbf{R}}_{n}$, де **R** – множини дійсних чисел, визначає n-вимірний евклідовий простір $\mathbf{R}^n$.

Приклади.

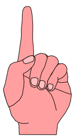
Якщо $\mathbf{A}$ – множина цілих парних позитивних чисел, а $\mathbf{B}$ – множина цілих непарних додатних чисел, то $\mathbf{A} \bigcup \mathbf{B}$ визначає множину усіх натуральних чисел, тобто множину $\mathbf{N} = \{1, 2, \ldots, n, \ldots\}$.

Якщо $\mathbf{A}$ – множина усіх цілих чисел, що діляться на 2, а $\mathbf{B}$ – множина усіх цілих чисел, що діляться на 5, то $\mathbf{A} \bigcap \mathbf{B}$ визначає множину усіх цілих чисел, що діляться і на 2, і на 5, тобто що діляться на 10.

Якщо $\mathbf{A} = \{1, 2, 3, 4, 5\}$, а, $\mathbf{B} = \{3, 5\}$, то $\mathbf{C}_{\mathbf{A}}\mathbf{B} = \mathbf{A} \setminus \mathbf{B} = \{1, 2, 4\}$, а $\mathbf{B} \setminus \mathbf{A} = \varnothing$.

Якщо $\mathbf{A} = \{1, 2\}$, а, $\mathbf{B} = \{3, -1, 0\}$, то декартовий добуток $\mathbf{A} \times \mathbf{B} = \{(1, 2), (1, -1), (1, 0), (2, 3), (2, -1), (2, 0)\}$.

2.3. Точка й основні типи точкових множин



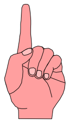
Визначення 2.8. Послідовність n дійсних чисел $x_1, x_2, \ldots, x_n$ називають n-вимірною *точкою,* а самі числа $x_1, x_2, \ldots, x_n$ – *координатами* цієї точки.

Та сама точка може бути позначена і відповідно задана двома способами:

- у вигляді вектора-стовпця: $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{bmatrix}$;
- у вигляді вектора-рядка: $\mathbf{x}^{\mathrm{т}} = [x_1 \;\; x_2 \;\; \ldots \;\; x_n]$.

Вибір способу позначення визначається контекстними умовами відповідно до правил матричного числення.

Для позначення точок і їхніх координат може бути використана проста і складна індексація. Наприклад, точки, що відповідають вершинам трикутника, можуть бути позначені $\mathbf{x}_A$, $\mathbf{x}_B$ і $\mathbf{x}_C$; граничні точки відрізка – $\mathbf{x}^+$ і $\mathbf{x}^{++}$; довільна сукупність точок – $\{\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_m\}$. Координати точки, що ітераційно змінює своє положення, можуть бути позначені як $x_i^{(k)}$, де індекс i визначає порядковий номер координати, а k – номер ітерації.



Визначення 2.9. Точковою множиною називають довільну сукупність n-вимірних точок.

Точкові множини є окремими випадками звичайних множин. Точкові множини, як і будь-які інші, задаються перерахуванням усіх елементів (точок) або зазначенням властивостей, якими володіють тільки точки даної множини. Для точкових множин також існують поняття підмножини і приналежності. Мають місце ті ж математичні операції, властивості операцій і позначення, що і для звичайних множин.

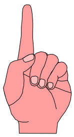
Точкові множини теж діляться на скінченні й нескінченні.

Скінченна точкова множина складається із скінченного числа точок, які можна перелічити за скінченне число кроків. Прикладом скінченної точкової множини є вже згадана множина точок, що відповідають вершинам трикутника.

Нескінченні очкові множини, у свою чергу, діляться на лічені (дискретні), незлічені (неперервні) і порожні. Прикладом ліченої точкової множини може служити множина точок із цілими координатами або множина точок координатної сітки з будь-яким відмінним від нуля кроком дискретності.

Серед неперервних точок з множин виділимо такі типи множин: *простір; гіперпаралелепіпед; гіперкулю.* Дані типи множин найбільш часто фігурують у курсі математичного програмування.

2.3.1. Простір



Визначення 2.10. *Простір* – це множина, задана як декартовий добуток $\underbrace{\mathbf{R}\times\mathbf{R}\times\ldots\times\mathbf{R}}_{n}$, де $\mathbf{R}$ – множина дійсних чисел.

Простір, або арифметичний простір, або n-вимірний евклідовий простір, являє собою множину, що складається з *усіх* n-вимірних точок. Простір прийнято позначати як $\mathbf{R}^n$, де n – вимірність простору, що співпадає з кількістю координат точок.

Приклади просторів: $\mathbf{R}^1$ – одновимірний простір (лінія); $\mathbf{R}^2$ – двовимірний простір (площина); $\mathbf{R}^3$ – звичайний тривимірний евклідовий простір.

Точковий простір $\mathbf{R}^4$ і простори більшої вимірності є абстрактними просторами, що не мають геометричної інтерпретації.

Приклад **надання точки.** Якщо точка задана вектором-рядком $\mathbf{x}^{\mathrm{т}} = \begin{bmatrix}-3 & 2\end{bmatrix}$, то за кількістю координат визначають вимірність простору. У даному випадку дві координати визначають двовимірний простір . Простір умовно зображується двома ортогональними координатними осями. Геометрична інтепретація простору $\mathbf{R}^2$ і двовимірної точки $\mathbf{x}^{\mathrm{т}} = \begin{bmatrix}-3 & 2\end{bmatrix}$ показана на рис.2.6.

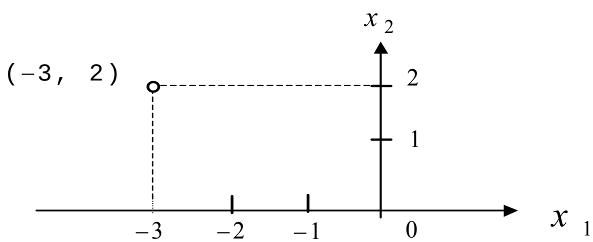


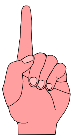
Рис.2.6.

Відстань між точками. Нехай і $\mathbf{y}^{\text{т}} = [y_1 \;\; y_2 \;\; \ldots \;\; y_n]$ - дві n-вимірні точки. Тоді розмір

$$\rho(\mathbf{x},\mathbf{y}) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \ldots + (x_n - y_n)^2} = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}$$

є відстанню між точками **x** і **y**.

2.3.2. Гіперпаралелепіпед



Визначення 2.11. Гіперпаралелепіпедом називають точкову множину, що задається:

$$\mathbf{G}: \;\; \mathbf{x}^{+} \le \mathbf{x} \le \mathbf{x}^{++} \quad \text{або} \quad \mathbf{G}: \;\; \mathbf{x}^{+\text{т}} \le \mathbf{x}^{\text{т}} \le \mathbf{x}^{++\text{т}},$$

де $\mathbf{x}^{+}$ – вектор нижніх границь координат точок (змінної);

$\mathbf{x}^{++}$ – вектор верхніх границь координат точок (змінної).

Гіперпаралелепіпед у просторі являє собою *відрізок*, у просторі $\mathbf{R}^2$ – прямокутник; у просторі $\mathbf{R}^3$ -паралелепіпед, у чотиривимірному просторі і просторах більшої вимірності – власне гіперпаралелепіпед (абстрактний n-вимірний паралелепіпед).

Приклад 2.1. Нехай гіперпаралелепіпед заданий таким чином:

$$\mathbf{G}: \begin{bmatrix} -3 \\ 1 \end{bmatrix} \le \mathbf{x} \le \begin{bmatrix} 2 \\ 1 \end{bmatrix},$$ тоді

затінений прямокутник на рис.2.7 визначає всі точки множини $\mathbf{G}$

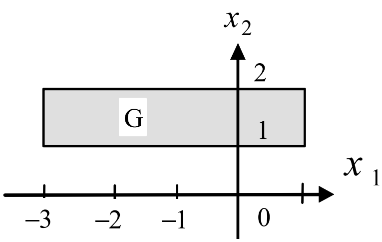


Рис.2.7.

2.3.3. Гіперкуля

> **Визначення 2.12.** Гіперкулею називають точкову множину, що задається таким чином:
>
> $$\mathbf{F}: \ \left\| \mathbf{x}\text{-}\mathbf{a} \right\| \le \varepsilon \ ,$$
>
> де ε – радіус гіперкуля; **a** – центр гіперкулі; $\left\| \mathbf{x}\ \text{-}\ \mathbf{a} \right\| = \sqrt{\sum_{i=1}^{n} \left(x_i - a_i \right)^2}$ - евклідова норма вектора.

Гіперкуля у просторі являє собою *відрізок,* середня точка якого відповідає центру гіперкулі, а відстань від середньої точки до граничних точок відрізка відповідає радіусу гіперкулі. У просторі $\mathbf{R}^2$ гіперкуля трансформується в коло, радіус і центр якого співпадають відповідно з радіусом і центром гіперкулі. У просторі $\mathbf{R}^3$ гіперкуля – це звичайна куля, центр і радіус якої відповідають центру і радіусу гіперкулі. У чотиривимірному просторі і просторах більшої вимірності гіперкулею є абстрактна n-вимірна куля.

Приклад 2.2. Нехай гіперкуля задана таким чином:

$$\mathbf{F}: \ \left\| \mathbf{x} - \begin{bmatrix} 2 \\ 3 \end{bmatrix} \right\| \le 0{,}5 \ ,$$

тоді затінене коло на рис.2.8 визначає всі точки множини $\mathbf{F}$.

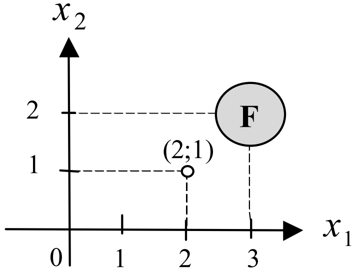


Рис.2.8.

Приклад 2.3. Перевірити приналежність точки $\mathbf{b}^{\mathrm{T}} = \begin{bmatrix} 1 & 2 \end{bmatrix}$ гіперкулі, зображеній на рис..2.8.

Розв'язання. Для вирішення задачі необхідно спочатку визначити норму вектора $\mathbf{b}$ відносно центра гіперкулі $\mathbf{a}$ за формулою

$$\|\mathbf{b} - \mathbf{a}\| = \sqrt{\sum_{i=1}^{n} \left(b_i - a_i\right)^2} \ ,$$

а потім порівняти з радіусом гіперкулі. Якщо норма вектора не перевершує радіус, то точка $\mathbf{b}$, що перевіряється , належить гіперкулі, у протилежному випадку – не належить.

Обчислимо норму вектора для точки, що перевіряється:

$$\|\mathbf{b} - \mathbf{a}\| = \sqrt{\left(b_1 - a_1\right)^2 + \left(b_2 - a_2\right)^2} = \sqrt{\left(3-2\right)^2 + \left(2-1\right)^2} = \sqrt{2} \approx 1{,}4 \ .$$

Оскільки 1,4>ε, то точка $\mathbf{b}$ не належить гіперкулі $\mathbf{F}$, зображеній на рис.2.8: $\mathbf{b} \notin \mathbf{F}$.

2.4. Явне і неявне задання точкових множин

Точкова множина може бути надана у вигляді системи рівностей або нерівностей. При цьому всі змінні (невідомі) системи розглядаються як координати точок, а сукупність усіх коренів (рішень) системи – як безпосередньо точкова множина.

Приклад. Нехай точкова множина задана таким чином:

$$\mathbf{W}\text{:} \ \begin{cases} x_1 + 2x_2^2 - 5 = 0; \\ x_1 |x_2| - 2 = 0. \end{cases} \qquad (2.1)$$

Тоді множина $\mathbf{W}$ складається з двох двовимірних точок $\mathbf{x}_1 = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$ і $\mathbf{x}_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$, що є коренями системи рівнянь (2.1).

Розглянуті раніше способи надання точкових множин безпосередньо визначали точкові множини. Тому ці способи вважаються явними. Надання ж точкових множин через системи рівностей або нерівностей може бути *явним* і *неявним*. Так, система рівностей у наведеному прикладі є неявною формою надання, оскільки для визначення точок, що складають обумовлену множину, необхідно попереднє вирішення цієї системи.

Якщо система з m рівнянь типу рівностей або нерівностей має n невідомих і розв'язана відносно будь-яких m змінних, то ці змінні називаються *залежними,* оскільки їхні значення залежать від значень інших $p=n-m$ змінних, що, у свою чергу, називаються *незалежними*. Незалежні змінні можуть набувати будь-які значення: вони визначені на всій координатній осі. Значення залежних перемінних обчислюються на основі конкретно заданих значень незалежних змінних. Тому таке надання точкових множин можна вважати явним.

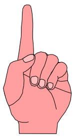
Приклад 2.4. Нехай точкова множина задана системою рівностей

$$\mathbf{W}\text{:}\ \begin{cases} x_1 = 5x_3^2 + x_4 - 3; \\ x_2 = 3x_3 x_5 - 8\ . \end{cases} \tag{2.2}$$

Тоді множина $\mathbf{W}$ являє собою неперервну множину 5-вимірних точок. Як значення незалежних координат x_3, x_4 і x_5 можуть бути обрані будь-які дійсні числа. Значення інших двох координат обчислюються відповідно до рівностей (2.2) з урахуванням обраних значень незалежних координат. Оскільки для одержання конкретних точок заданої множини немає необхідності в попередньому вирішенні системи рівнянь, то надання множини $\mathbf{W}$ у формі (2.2) вважається явним.

2.5. Окіл точки. Обмежені точкові множини Внутрішні й граничні точки

2.5.1. Окіл точки



Визначення 2.13. Якщо ε – деяке позитивне число, то ε-околом $\mathbf{S}_\varepsilon(\mathbf{x}_0)$ точки $\mathbf{x}_0$ в n-вимірному просторі $\mathbf{R}^n$ називають множину усіх точок $\mathbf{x} \in \mathbf{R}^n$ таких, що $\rho(\mathbf{x}, \mathbf{x}_0) < \varepsilon$, тобто

$$\mathbf{S}_\varepsilon(\mathbf{x}_0) = \{\mathbf{x} \in \mathbf{R}^n \mid \rho(\mathbf{x}, \mathbf{x}_0) < \varepsilon\}.$$

Наведемо приклади аналізу приналежності конкретних точок ε-околу заданої точки.

Приклад 2.5.

$\mathbf{x}_1 \in \mathbf{S}_2(\mathbf{x}_0)$, де $\mathbf{x}_1^{\mathrm{т}} = [2 \quad 3 \quad -1 \quad 3]$, $\mathbf{x}_0^{\mathrm{т}} = [1 \quad 2 \quad -1 \quad 2]$, тому що $\rho(\mathbf{x}_1, \mathbf{x}_0) = \sqrt{3} < 2$

Приклад 2.6.

$\mathbf{x}_2 \notin \mathbf{S}_2(\mathbf{x}_0)$, де $\mathbf{x}_2^{\mathrm{т}} = [3 \quad 3 \quad -1 \quad 3]$, $\mathbf{x}_0^{\mathrm{т}} = [1 \quad 2 \quad -1 \quad 2]$, тому що $\rho(\mathbf{x}_2, \mathbf{x}_0) = \sqrt{6} > 2$.

У просторі $\mathbf{R}^1$ ε-окіл точки $\mathbf{x}_0 = [a]$ - це інтервал $]a - \varepsilon, a + \varepsilon[$.

У просторі $\mathbf{R}^2$ ε-окіл точки $\mathbf{x}_0^{\mathrm{т}} = [a \quad b]$ - це внутрішність кола радіуса ε із центром у точці $\mathbf{x}_0$.

У просторі $\mathbf{R}^3$ точка $\mathbf{x}_0^{\mathrm{т}} = [a \quad b \quad c]$ - це внутрішність кулі радіуса ε із центром у точці $\mathbf{x}_0$.

У просторі $\mathbf{R}^n$ ε-окіл точки $\mathbf{x}_0$ - це внутрішність гіперкулі радіуса ε із центром у точці $\mathbf{x}_0$.

2.5.2. Обмежені множини

Визначення 2.14. Множина $\mathbf{W}$ точок n-вимірного простору $\mathbf{R}^n$ називають обмеженим, якщо існує число $X>0$ таке, що для будь-якої точки $\mathbf{x} \in \mathbf{W}$ виконується співвідношення $|x_i| \le X,\ i = \overline{1,n}$.

Прикладами обмежених множин можуть служити: ε-окіл будь-якої точки n-вимірного простору, гіперпаралелепіпед, гіперкуля.

Переріз і об'єднання обмежених множин також є обмеженою множиною.

2.5.3. Внутрішні й граничні точки

Точку $\mathbf{x}_0 \in \mathbf{W}$ називають *внутрішньою*, якщо будь-яка точка з її ε-окола також належить множині $\mathbf{W}$ (рис.2.8).

Точку $\mathbf{x}_0 \in \mathbf{W}$ називають *граничною*, якщо деякі точки з її ε-окола не належать множині $\mathbf{W}$ (рис.2.8).

Визначення 2.15. Точку $\mathbf{x}_0 \in \mathbf{W}$ називають *внутрішньою*, якщо будь-яка точка з її ε-окола також належить множині $\mathbf{W}$ (рис.2.8).

Точку $\mathbf{x}_0 \in \mathbf{W}$ називають *граничною*, якщо деякі точки з її ε-окола не належать множині $\mathbf{W}$ (рис.2.8).

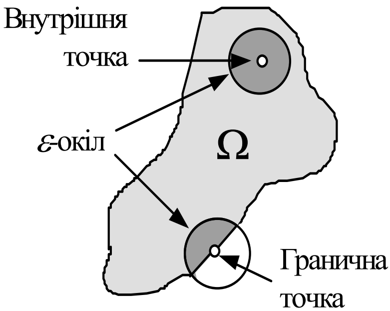


Рис.2.8.

Множину усіх граничних точок множини W називають *границею* цієї множини.

Приклад 2.7.. Якщо $\mathbf{W} = \left\{\mathbf{x} \in \mathbf{R}^2 \mid x_1^2 + x_2^2 < 1\right\}$, то всі точки цієї множини внутрішні, границя збігається з колом $x_1^2 + x_2^2 = 1$.

2.6. Граничні точки. Замкнуті й відкриті точкові множини

Визначення 2.16. Точку $\mathbf{x}_0 \in \mathbf{R}^n$ називають граничною точкою множини W n-вимірних точок, якщо кожний окіл точки $\mathbf{x}_0$ містить нескінченно багато точок множини W.

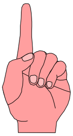
Наприклад, точка $\mathbf{x}_0 = [0]$ з простору $\mathbf{R}^1$ є граничною точкою множини $\mathbf{W} = \left\{\frac{1}{k} \mid k = 1, 2, \ldots\right\}$ (рис.2.9).

Властивості граничних точок

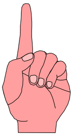
Будь-яка внутрішня точка множини $\mathbf{W}$ є граничною точкою цієї множини.

Якщо гранична точка множини $\mathbf{W}$ не належить цій множині, то вона є граничною точкою множини $\mathbf{W}$.

Замкнуті й відкриті множини



Визначення 2.17. Множина $\mathbf{W}$ у просторі $\mathbf{R}^n$ називається *замкнутою,* якщо вона містить усі свої граничні точки.



Визначення 2.18. Множина $\mathbf{W}$ у просторі $\mathbf{R}^n$ називається *відкритою,* якщо всі точки множини $\mathbf{W}$ є внутрішніми.

Приклади:

гіперпаралелепіпед – замкнута множина у просторі $\mathbf{R}^n$;

гіперперкуля – замкнута множина у просторі $\mathbf{R}^n$;

ε-окіл n-мірної точки $\mathbf{x}_0$ – відкрита множина у просторі $\mathbf{R}^n$.

Властивості відкритих і замкнутих точкових множин

Якщо множина $\mathbf{W}$ має свою границю, то вона замкнута.

Переріз будь-якого числа замкнутих множин є множина замкнута.

Об'єднання скінченного числа замкнутих множин є множина замкнута.

Переріз будь-якого числа відкритих множин є множина відкрита.

Об'єднання будь-якого числа відкритих множин є множина відкрита.

Множина $\mathbf{W}$ відкрита тоді і тільки тоді, коли її доповнення замкнуте.

Обмежені замкнуті множини в просторі $\mathbf{R}^n$ називають *компактними.*

Далі будуть розглядатися тільки замкнуті множини, якщо вони не стосуються всього n-вимірного евклідового простору $\mathbf{R}^n$ або не буде оговорене протилежне.

2.7. Контрольні запитання, приклади і вправи

1. Дати визначення поняттям «множина» і «точкова множина».

2. Назвати способи надання множин. Навести приклади.

3. Які множини називаються скінченними? Нескінченними? Ліченими? Незліченними?

4. Навести приклади надання точки у вигляді вектора-рядка і вектора-стовпця.

5. Яким способом вказується входження підмножини в множину? Включення множиною підмножини?

6. Які операції над множинами називаються об'єднанням? Перерізом? Різницею? Доповненням?

7. Дати визначення операції об'єднання множин. Навести приклад.

8. Дати визначення операції перерізу множин. Навести приклад.

9. Дати визначення операції різниці множин. Навести приклад.

10. Дати визначення операції доповнення множин. Навести приклад.

11. Яку множину називають порожньою? Як вона позначається?

12. Задані числові множини: $\mathbf{A} = \{2, 5, 8, 11\}$, $\mathbf{B} = \{2, 4, 8, 9\}$, $\mathbf{C} = \{1, 5, 6, 9\}$. Знайти множини: $\mathbf{A} \bigcup \mathbf{B}$, $\mathbf{A} \bigcup \mathbf{C}$, $\mathbf{B} \bigcup \mathbf{C}$, $\mathbf{A} \bigcap \mathbf{B} \bigcap \mathbf{C}$, $\mathbf{A} \bigcap \mathbf{B}$, $\mathbf{A} \bigcap \mathbf{C}$, $\mathbf{B} \bigcap \mathbf{C}$, $\mathbf{A} \setminus \mathbf{B}$, $\mathbf{A} \setminus \mathbf{C}$, $\mathbf{B} \setminus \mathbf{C}$, $\mathbf{B} \setminus \mathbf{A}$, $\mathbf{C} \setminus \mathbf{A}$, $\mathbf{C} \setminus \mathbf{B}$, $(\mathbf{A} \bigcup \mathbf{B}) \setminus \mathbf{C}$, $(\mathbf{A} \bigcap \mathbf{B}) \setminus \mathbf{C}$, $\mathbf{A} \setminus (\mathbf{B} \bigcup \mathbf{C})$, $\mathbf{A} \setminus (\mathbf{B} \bigcap \mathbf{C})$, $\mathbf{A} \setminus \mathbf{B} \setminus \mathbf{C}$, $\mathbf{B} \setminus \mathbf{C} \setminus \mathbf{A}$.

13. Довести, що $\bigcup \mathbf{M}_i$ – найменша множина, що містить всі множини $\mathbf{M}_i$.

14. Довести, що $\bigcap \mathbf{M}_i$ – найбільша множина, що міститься у всіх множинах $\mathbf{M}_i$.

15. Розв'язати систему рівнянь

$$\begin{cases} \mathbf{A} \setminus \mathbf{X} = \mathbf{B}\,; \\ \mathbf{A} \bigcup \mathbf{X} = \mathbf{C}\,. \end{cases}$$

16. Відомо, що з 100 студентів живописом захоплюються 28, спортом – 42, музикою – 30, живописом і спортом – 10, живописом і музикою – 8, спортом і музикою – 5, живописом, спортом і музикою – 5. Визначити: а) кількість студентів, причетних до спорту; б) кількість студентів, що захоплюються тільки музикою; в) кількість студентів, що захоплюються тільки спортом і музикою.

17. Нехай $\mathbf{B}$ – множина усіх паралелограмів на площині, $\mathbf{A}_1$ – множина квадратів, $\mathbf{A}_2$ – множина прямокутників, $\mathbf{A}_3$ – множина ромбів. Знайти результат таких операцій: $\mathbf{A}_i \bigcup \mathbf{A}_j$, $\mathbf{A}_i \bigcap \mathbf{A}_j$, $\mathbf{A}_i \setminus \mathbf{A}_j$, $i, j = 1, 2, 3$.

18. Як позначається нескінченна множина усіх точок n-вимірного простору? Навести приклад надання двовимірного простору.

19. Дати визначення гіперпаралелепіпеду як точковій множині. Навести приклади гіперпаралелепіпедів в одно-, дво- і тривимірному просторах.

20. Дати геометричну інтерпретацію гіперпаралелепіпеду

$$\mathbf{G}: \begin{bmatrix} -4 \\ -2 \end{bmatrix} \le \mathbf{x} \le \begin{bmatrix} 2 \\ -1 \end{bmatrix}.$$

Розв'язання. У векторів $\mathbf{x}^{+}$ і $\mathbf{x}^{++}$ є по дві складові x_1 і x_2. Отже, точкова множина задана в двовимірному просторі $\mathbf{R}^2$, тобто на площині. У декартовій прямокутній системі координат заданий гіперпаралелепіпед являтиме собою прямокутник. При побудові прямокутника необхідно послідовно для кожної компоненти вектора $\mathbf{x}$ на відповідній осі системи координат відзначити точки, що відповідають її нижній і верхній границям x_i^{+} і x_i^{++}, $i = 1, 2$. У наведеному прикладі $x_1^{+} = -4$; $x_2^{+} = -2$; $x_1^{++} = 2$; $x_2^{++} = -1$ (рис.2.9). Відрізки $\left[x_1^{+}, x_1^{++}\right]$ і $\left[x_2^{+}, x_2^{++}\right]$ являють собою проекції шуканого прямокутника на відповідні осі координат. Відновлюючи перпендикуляри від осей у точках x_1^{+}, x_2^{++} і x_1^{++}, x_2^{++} до їхнього взаємного перетинання, одержуємо шуканий прямокутник.

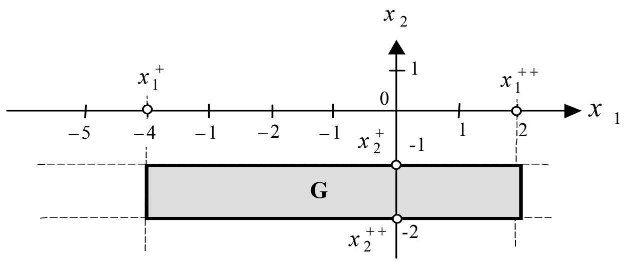


Рис.2.9

21. Побудувати гіперпаралелепіпед $\mathbf{G}$: $\begin{bmatrix} 1 \\ -3 \end{bmatrix} \le \mathbf{x} \le \begin{bmatrix} 3 \\ 1 \end{bmatrix}$.

22. Побудувати гіперпаралелепіпед $\mathbf{G}$: $\mathbf{x}^{+} \le \mathbf{x} \le \mathbf{x}^{++}$, якщо $\mathbf{x}^{+\mathbf{T}} = [1 \quad 3]$, $\mathbf{x}^{++\mathbf{T}} = [-2 \quad -1]$.

23. Побудувати гіперпаралелепіпед $\mathbf{G}$: $\begin{bmatrix} 1 \\ 1 \end{bmatrix} \le \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \le \begin{bmatrix} 3 \\ 3 \end{bmatrix}$.

24. Дати визначення гіперкулі як точкової множини. Привести приклад гіперкулі.

25. Побудувати точкову множину, подану у вигляді гіперкулі $\mathbf{F}: \ \left\| \mathbf{x} - \begin{bmatrix} 2 \\ 2 \end{bmatrix} \right\| \le 1$.

26. Яка умова приналежності довільної n-вимірної точки гіперкулі? Границі гіперкулі?

27. Перевірити приналежність точки з координатами $\left(\mathbf{x}^{+}\right)^{\mathrm{т}} = \begin{bmatrix} 1 & -2 & 0 \end{bmatrix}$ гіперкулі $\mathbf{F}$ із центром у точці $\left(\mathbf{a}^{+}\right)^{\mathrm{т}} = \begin{bmatrix} 2 & 3 & 1 \end{bmatrix}$ і радіусом $\varepsilon = 2$.

Розв'язання. Для вирішення задачі необхідно визначити норму вектора $\mathbf{x} - \mathbf{a}$, тобто відстань від точки $\mathbf{x}$ до точки $\mathbf{a}$, і порівняти її з радіусом ε. Якщо норма буде більше значення ε, то точка $\mathbf{x}$ не належить наданої гіперкулі, у протилежному випадку – належить.

Для обчислення норми вектора використовують формулу $\left\| \mathbf{x} - \mathbf{a} \right\| = \sqrt{\sum_{i=1}^{n} \left(x_i - a_i\right)^2}$. Для наведеного приклада n=3. Отже, $\left\| \bar{x} - \bar{a} \right\| =$ $= \sqrt{\left(x_1 - a_1\right)^2 + \left(x_2 - a_2\right)^2 + \left(x_3 - a_3\right)^2} = = \sqrt{\left(1-2\right)^2 + \left(-2-3\right)^2 + \left(0-1\right)^2} = 3\sqrt{3} > 2$. Оскільки $\left\| \mathbf{x} - \mathbf{a} \right\| > \varepsilon$, то $\mathbf{x} \notin \mathbf{F}$.

28. Визначити, які з точок $\mathbf{b}^{\mathrm{т}} = \begin{bmatrix} 4 & 3 \end{bmatrix}$, $\mathbf{c}^{\mathrm{т}} = \begin{bmatrix} 4 & 3 \end{bmatrix}$, $\mathbf{d}^{\mathrm{т}} = \begin{bmatrix} 1 & 1 \end{bmatrix}$ належать гіперкулі $\mathbf{Q}$ із центром у точці $\mathbf{a}^{\mathrm{т}} = \begin{bmatrix} 2 & 1 \end{bmatrix}$ і радіусом $\varepsilon = 1{,}8$.

29. Перевірити приналежність точки $\mathbf{z}^{\mathrm{т}} = \begin{bmatrix} 4 & 6 & 7 & 5 \end{bmatrix}$ гіперкулі $\mathbf{S}$ радіуса $\varepsilon = 0{,}5$ із центром у точці $\mathbf{a}^{\mathrm{т}} = \begin{bmatrix} 4 & 5 & 8 & 4 \end{bmatrix}$.

2.8. Фонд індивідуальних завдань

Індивідуальне завдання №8. Побудувати гіперпаралелепіпед у двовимірному просторі, якщо вектори нижніх і верхніх границь координат точок задані таким чином:

1. $\left(\mathbf{x}^{+}\right)^{\mathrm{T}} = [-1 \quad 0]$,
$\left(\mathbf{x}^{++}\right)^{\mathrm{T}} = [1 \quad 2]$;

2. $\left(\mathbf{x}^{+}\right)^{\mathrm{T}} = [1 \quad -2]$,
$\left(\mathbf{x}^{++}\right)^{\mathrm{T}} = [2 \quad 0]$;

3. $\left(\mathbf{x}^{+}\right)^{\mathrm{T}} = [1 \quad 1]$,
$\left(\mathbf{x}^{++}\right)^{\mathrm{T}} = [2 \quad 2]$;

4. $\left(\mathbf{x}^{+}\right)^{\mathrm{T}} = [0 \quad 0]$,
$\left(\mathbf{x}^{++}\right)^{\mathrm{T}} = [2 \quad 1]$;

5. $\left(\mathbf{x}\right)^{+\mathbf{T}} = [-1 \quad -1]$,
$\left(\mathbf{x}\right)^{++\mathbf{T}} = [1 \quad 2]$;

6. $\left(\mathbf{x}^{+}\right)^{\mathrm{T}} = [0 \quad 1]$,
$\left(\mathbf{x}^{++}\right)^{\mathrm{T}} = [3 \quad 2]$;

7. $\mathbf{x}^{+\mathbf{T}} = [1 \quad 0]$,
$\mathbf{x}^{++\mathbf{T}} = [2 \quad 2]$;

8. $\mathbf{x}^{+\mathbf{T}} = [-1 \quad -1]$,
$\mathbf{x}^{++\mathbf{T}} = [2 \quad 1]$;

9. $\mathbf{x}^{+\mathbf{T}} = [0 \quad 1]$,
$\mathbf{x}^{++\mathbf{T}} = [2 \quad 3]$;

10. $\mathbf{x}^{+\mathbf{T}} = [0 \quad -1]$,
$\mathbf{x}^{++\mathbf{T}} = [1 \quad 2]$;

11. $\mathbf{x}^{+\mathbf{T}} = [-1 \quad -1]$,
$\mathbf{x}^{++\mathbf{T}} = [2 \quad 3]$;

12. $\mathbf{x}^{+\mathbf{T}} = [-1 \quad 1]$,
$\mathbf{x}^{++\mathbf{T}} = [1 \quad 2]$;

13. $\mathbf{x}^{+\mathbf{T}} = [1 \quad 1]$,
$\mathbf{x}^{++\mathbf{T}} = [2 \quad 3]$;

14. $\mathbf{x}^{+\mathbf{T}} = [0 \quad 1]$,
$\mathbf{x}^{++\mathbf{T}} = [2 \quad 2]$;

15. $\mathbf{x}^{+\mathbf{T}} = [-1 \quad -2]$,
$\mathbf{x}^{++\mathbf{T}} = [1 \quad 0]$;

16. $\mathbf{x}^{+\mathbf{T}} = [-2 \quad -2]$,
$\mathbf{x}^{++\mathbf{T}} = [2 \quad -1]$;

17. $\mathbf{x}^{+\mathrm{T}} = [-2 \quad -1]$,

$\mathbf{x}^{++\mathrm{T}} = [-1 \quad 1]$;

18. $\mathbf{x}^{+\mathrm{T}} = [1 \quad -2]$,

$\mathbf{x}^{++\mathrm{T}} = [2 \quad 2]$;

19. $\mathbf{x}^{+\mathrm{T}} = [-1 \quad 0]$,

$\mathbf{x}^{++\mathrm{T}} = [0 \quad 2]$;

20. $\mathbf{x}^{+\mathrm{T}} = [1 \quad 0]$,

$\mathbf{x}^{++\mathrm{T}} = [2 \quad 3]$;

21. $\mathbf{x}^{+\mathrm{T}} = [-1 \quad 1]$,

$\mathbf{x}^{++\mathrm{T}} = [2 \quad 2]$;

22. $\mathbf{x}^{+\mathrm{T}} = [-2 \quad 0]$,

$\mathbf{x}^{++\mathrm{T}} = [0 \quad 2]$;

23. $\mathbf{x}^{+\mathrm{T}} = [-1 \quad -1]$,

$\mathbf{x}^{++\mathrm{T}} = [1 \quad 1]$;

24. $\mathbf{x}^{+\mathrm{T}} = [-1 \quad 1]$,

$\mathbf{x}^{++\mathrm{T}} = [1 \quad 2]$;

25. $\mathbf{x}^{+\mathrm{T}} = [-1 \quad -1]$,

$\mathbf{x}^{++\mathrm{T}} = [2 \quad 2]$;

26. $\mathbf{x}^{+\mathrm{T}} = [-2 \quad -1]$,

$\mathbf{x}^{++\mathrm{T}} = [0 \quad 3]$;

27. $\mathbf{x}^{+\mathrm{T}} = [1 \quad 2]$,

$\mathbf{x}^{++\mathrm{T}} = [3 \quad 3]$;

28. $\mathbf{x}^{+\mathrm{T}} = [-2 \quad -1]$,

$\mathbf{x}^{++\mathrm{T}} = [3 \quad 3]$;

29. $\mathbf{x}^{+\mathrm{T}} = [-1 \quad 1]$,

$\mathbf{x}^{++\mathrm{T}} = [3 \quad 2]$;

30. $\mathbf{x}^{+\mathrm{T}} = [-2 \quad 0]$,

$\mathbf{x}^{++\mathrm{T}} = [0 \quad 3]$;

Індивідуальне завдання №9. Визначити, які з точок **b, c, d** належать гіперкулі **G** радіуса ε і з центром у точці **a**, якщо:

1. $\mathbf{b}^{\mathrm{T}} = [1 \quad 1]$, $\mathbf{c}^{\mathrm{T}} = [1{,}5 \quad 2]$, $\mathbf{d}^{\mathrm{T}} = [2{,}5 \quad 1{,}5]$,

$\varepsilon = 1$, $\mathbf{a}^{\mathrm{T}} = [2 \quad 2]$;

2. $\mathbf{b}^{\mathrm{T}} = [2 \quad 3]$, $\mathbf{c}^{\mathrm{T}} = [1{,}5 \quad 2{,}5]$, $\mathbf{d}^{\mathrm{T}} = [1{,}5 \quad 3]$,
$\varepsilon = 1$, $\mathbf{a}^{\mathrm{T}} = [1 \quad 2]$;

3. $\mathbf{b}^{\mathrm{T}} = [3 \quad 3]$, $\mathbf{c}^{\mathrm{T}} = [3 \quad 2]$, $\mathbf{d}^{\mathrm{T}} = [2 \quad 2]$,
$\varepsilon = 2$, $\mathbf{a}^{\mathrm{T}} = [2 \quad 1]$;

4. $\mathbf{b}^{\mathrm{T}} = [2 \quad 2]$, $\mathbf{c}^{\mathrm{T}} = [2 \quad 1]$, $\mathbf{d}^{\mathrm{T}} = [3 \quad 1]$,
$\varepsilon = 2$, $\mathbf{a}^{\mathrm{T}} = [1 \quad 1]$;

5. $\mathbf{b}^{\mathrm{T}} = [0{,}5 \quad 2]$, $\mathbf{c}^{\mathrm{T}} = [1 \quad 0]$, $\mathbf{d}^{\mathrm{T}} = [0 \quad 2]$,
$\varepsilon = 1$, $\mathbf{a}^{\mathrm{T}} = [0 \quad 2]$;

6. $\mathbf{b}^{\mathrm{T}} = [1 \quad 1]$, $\mathbf{c}^{\mathrm{T}} = [1 \quad 0]$, $\mathbf{d}^{\mathrm{T}} = [1 \quad 2]$,
$\varepsilon = 2$, $\mathbf{a}^{\mathrm{T}} = [0 \quad 2]$;

7. $\mathbf{b}^{\mathrm{T}} = [6 \quad 1]$, $\mathbf{c}^{\mathrm{T}} = [5 \quad 2]$, $\mathbf{d}^{\mathrm{T}} = [7 \quad 10]$,
$\varepsilon = 3$, $\mathbf{a}^{\mathrm{T}} = [7 \quad 2]$;

8. $\mathbf{b}^{\mathrm{T}} = [3 \quad 4]$, $\mathbf{c}^{\mathrm{T}} = [3 \quad 3]$, $\mathbf{d}^{\mathrm{T}} = [2 \quad 4]$,
$\varepsilon = 2$, $\mathbf{a}^{\mathrm{T}} = [4 \quad 5]$;

9. $\mathbf{b}^{\mathrm{T}} = [4 \quad 6]$, $\mathbf{c}^{\mathrm{T}} = [3 \quad 5]$, $\mathbf{d}^{\mathrm{T}} = [2 \quad 5]$,
$\varepsilon = 2$, $\mathbf{a}^{\mathrm{T}} = [4 \quad 4]$;

10. $\mathbf{b}^{\mathrm{T}} = [-1 \quad 1]$, $\mathbf{c}^{\mathrm{T}} = [0 \quad 0]$, $\mathbf{d}^{\mathrm{T}} = [1 \quad 1]$,
$\varepsilon = 2$, $\mathbf{a}^{\mathrm{T}} = [-1 \quad -1]$;

11. $\mathbf{b}^{\mathrm{T}} = [8 \quad 2]$, $\mathbf{c}^{\mathrm{T}} = [7 \quad 1]$, $\mathbf{d}^{\mathrm{T}} = [6 \quad 0]$,
$\varepsilon = 3$, $\mathbf{a}^{\mathrm{T}} = [5 \quad 1]$;

12. $\mathbf{b}^{\mathrm{T}} = [3 \quad 1]$, $\mathbf{c}^{\mathrm{T}} = [4 \quad 0]$, $\mathbf{d}^{\mathrm{T}} = [5 \quad 2]$,
$\varepsilon = 1$, $\mathbf{a}^{\mathrm{T}} = [4 \quad 1]$;

13. $\mathbf{b}^{\mathrm{T}} = [4 \quad 2]$, $\mathbf{c}^{\mathrm{T}} = [3 \quad 1]$, $\mathbf{d}^{\mathrm{T}} = [2 \quad 2]$,
$\varepsilon = 4$, $\mathbf{a}^{\mathrm{T}} = [0 \quad 0]$;

14. $\mathbf{b}^{\mathrm{T}} = [3 \quad 7]$, $\mathbf{c}^{\mathrm{T}} = [2 \quad 8]$, $\mathbf{d}^{\mathrm{T}} = [2 \quad 9]$,
$\varepsilon = 2$, $\mathbf{a}^{\mathrm{T}} = [4 \quad 8]$;

15. $\mathbf{b}^{\mathrm{T}} = [-3 \quad 3]$, $\mathbf{c}^{\mathrm{T}} = [4 \quad 4]$, $\mathbf{d}^{\mathrm{T}} = [0 \quad 2]$,
$\varepsilon = 3$, $\mathbf{a}^{\mathrm{T}} = [-2 \quad 2]$;

16. $\mathbf{b}^{\mathrm{T}} = [3 \quad 4]$, $\mathbf{c}^{\mathrm{T}} = [2 \quad 5]$, $\mathbf{d}^{\mathrm{T}} = [1 \quad 6]$,
$\varepsilon = 3$, $\mathbf{a}^{\mathrm{T}} = [1 \quad 8]$;

17. $\mathbf{b}^{\mathrm{T}} = [3 \quad 3]$, $\mathbf{c}^{\mathrm{T}} = [1 \quad 8]$, $\mathbf{d}^{\mathrm{T}} = [3 \quad 6]$,
$\varepsilon = 2$, $\mathbf{a}^{\mathrm{T}} = [2 \quad 7]$;

18. $\mathbf{b}^{\mathrm{T}} = [2 \quad 1]$, $\mathbf{c}^{\mathrm{T}} = [4 \quad 1]$, $\mathbf{d}^{\mathrm{T}} = [3 \quad 3]$,
$\varepsilon = 2$, $\mathbf{a}^{\mathrm{T}} = [3 \quad 0]$;

19. $\mathbf{b}^{\mathrm{T}} = [4 \quad 6]$, $\mathbf{c}^{\mathrm{T}} = [3 \quad 3]$, $\mathbf{d}^{\mathrm{T}} = [7 \quad 6]$,
$\varepsilon = 2$, $\mathbf{a}^{\mathrm{T}} = [5 \quad 5]$;

20. $\mathbf{b}^{\mathrm{T}} = [4 \quad 3]$, $\mathbf{c}^{\mathrm{T}} = [5 \quad 1]$, $\mathbf{d}^{\mathrm{T}} = [4 \quad 2]$,
$\varepsilon = 2$, $\mathbf{a}^{\mathrm{T}} = [6 \quad 2]$;

21. $\mathbf{b}^{\mathrm{T}} = [0 \quad 0]$, $\mathbf{c}^{\mathrm{T}} = [-2 \quad 2]$, $\mathbf{d}^{\mathrm{T}} = [3 \quad 0]$,
$\varepsilon = 3$, $\mathbf{a}^{\mathrm{T}} = [2 \quad -2]$;

22. $\mathbf{b}^{\mathrm{T}} = [3 \quad 7]$, $\mathbf{c}^{\mathrm{T}} = [1 \quad 8]$, $\mathbf{d}^{\mathrm{T}} = [1 \quad 5]$,
$\varepsilon = 3$, $\mathbf{a}^{\mathrm{T}} = [1 \quad 9]$;

23. $\mathbf{b}^{\mathrm{T}} = [5 \quad 4]$, $\mathbf{c}^{\mathrm{T}} = [7 \quad 2]$, $\mathbf{d}^{\mathrm{T}} = [9 \quad 4]$,
$\varepsilon = 2$, $\mathbf{a}^{\mathrm{T}} = [8 \quad 3]$;

24. $\mathbf{b}^{\mathrm{T}} = [3 \quad 3]$, $\mathbf{c}^{\mathrm{T}} = [5 \quad 5]$, $\mathbf{d}^{\mathrm{T}} = [3 \quad -2]$,
$\varepsilon = 3$, $\mathbf{a}^{\mathrm{T}} = [4 \quad 2]$;

25. $\mathbf{b}^{\mathrm{T}} = [0 \quad 4]$, $\mathbf{c}^{\mathrm{T}} = [-3 \quad 5]$, $\mathbf{d}^{\mathrm{T}} = [-2 \quad 5]$,
$\varepsilon = 2$, $\mathbf{a}^{\mathrm{T}} = [-4 \quad 4]$;

26. $\mathbf{b}^{\mathrm{T}} = [4 \quad 3]$, $\mathbf{c}^{\mathrm{T}} = [4 \quad 1]$, $\mathbf{d}^{\mathrm{T}} = [4 \quad 2]$,
$\varepsilon = 2$, $\mathbf{a}^{\mathrm{T}} = [5 \quad 2]$;

27. $\mathbf{b}^{\mathrm{T}} = [-1 \quad 0]$, $\mathbf{c}^{\mathrm{T}} = [-2 \quad 7]$, $\mathbf{d}^{\mathrm{T}} = [3 \quad 0]$,
$\varepsilon = 3$, $\mathbf{a}^{\mathrm{T}} = [2 \quad -1]$;

28. $\mathbf{b}^{\mathrm{T}} = [3 \quad 5]$, $\mathbf{c}^{\mathrm{T}} = [1 \quad 7]$, $\mathbf{d}^{\mathrm{T}} = [1 \quad 6]$,
$\varepsilon = 3$, $\mathbf{a}^{\mathrm{T}} = [1 \quad 8]$;

29. $\mathbf{b}^{\mathrm{T}} = [5 \quad 4]$, $\mathbf{c}^{\mathrm{T}} = [6 \quad 2]$, $\mathbf{d}^{\mathrm{T}} = [6 \quad 4]$,
$\varepsilon = 2$, $\mathbf{a}^{\mathrm{T}} = [7 \quad 3]$;

30. $\mathbf{b}^{\mathrm{T}} = [3 \quad 3]$, $\mathbf{c}^{\mathrm{T}} = [5 \quad 5]$, $\mathbf{d}^{\mathrm{T}} = [5 \quad -2]$,
$\varepsilon = 3$, $\mathbf{a}^{\mathrm{T}} = [3 \quad 2]$.

Розділ 3. ЖОРДАНОВІ ВИКЛЮЧЕННЯ

Розглядається метод, що дозволяє робити транспозицію залежних і незалежних змінних у системі лінійних форм. Метод широко використовується в лінійній алгебрі для обертання матриць і вирішення систем лінійних рівнянь як при числі змінних n, рівних числу рівнянь m, так і при $n>m$. У математичному програмуванні жорданові виключення мають ще одне застосовування. Метод лежить в основі диференціального алгоритму розв'язання екстремальних задач при лінійних обмеженнях. Це дозволяє з єдиних позицій розглядати різні типи задач математичного програмування. Так, задачі лінійного і квадратичного програмування як за постановкою, так і за методом їх вирішення стають окремими випадками загальної задачі математичного програмування.

3.1. Призначення жорданових виключень. Заміна залежної змінної на незалежну за допомогою алгебраїчних перетворень

Жорданові виключення – це метод, що робить у системі лінійних форм

$$s_i = \sum_{j=1}^{p} a_{ij} t_j \quad , \quad i = \overline{1, m} \,. \tag{3.1}$$

транспозицію (взаємну заміну) залежної змінної s_k і незалежної t_r без яких-небудь попередніх алгебраїчних перетворень.

Для полегшення розуміння суті питання і наступних висновків правил жорданових виключень зробимо спочатку транспозицію змінних за допомогою алгебраїчних перетворень на конкретному прикладі.

Нехай задана система двох лінійних форм, для котрої необхідно виконати транспозицію залежної змінної s_1 і незалежної t_2 :

$$\begin{cases} s_1 = t_1 + 2t_2 - t_3; \\ s_2 = -t_1 - 3t_2 + 2t_3. \end{cases} \tag{3.2}$$

Поставлене завдання можна виконати за допомогою таких алгебраїчних перетворень:

початкова система $\begin{cases} s_1 = \ t_1 + 2\ t_2 - \ t_3; \\ s_2 = -t_1 - 3\ t_2 + 2t_3. \end{cases}$

проміжна система $\begin{cases} t_2 = -\dfrac{1}{2}t_1 + \dfrac{1}{2}\ s_1 + \dfrac{1}{2}t_3; \\ s_2 = -t_1 - 3\ s_1 + t_3; \end{cases}$

проміжна система $\begin{cases} t_2 = -\dfrac{1}{2}t_1 + \dfrac{1}{2}\ s_1 + \dfrac{1}{2}t_3; \\ s_2 = -t_1 - 3\left(-\dfrac{1}{2}t_1 + \dfrac{1}{2}\ s_1 + \dfrac{1}{2}t_3\right) + t_3; \end{cases}$

результуюча система $\begin{cases} t_2 = -\dfrac{1}{2}t_1 + \dfrac{1}{2}s_1 + \dfrac{1}{2}t_3; \\ s_2 = \dfrac{1}{2}t_1 - \dfrac{3}{2}s_1 + \dfrac{1}{2}t_3. \end{cases}$

При великому числі змінних і великому числі лінійних форм у системі (3.1) алгебраїчні перетворення вимагають значних витрат сил і часу. Жорданові виключення прискорюють процедуру транспозиції змінних у системі лінійних форм. Вони забезпечують взаємну заміну змінних шляхом перетворення вихідної системи лінійних форм безпосередньо в результуючу без одержання яких-небудь інших проміжних систем. Відсутність проміжних систем лінійних форм у жорданових виключеннях – це їхня принципова відмінність від транспозиції змінних за допомогою звичайних алгебраїчних перетворень.

3.2. Таблиця жорданових виключень

Для виконання жорданових виключень використовуються спеціальні таблиці, що називаються таблицями жорданових виключень. Ці таблиці дозволяють у процесі транспозиції змінних подавати в зручній формі як вихідну, так і результуючу системи лінійних форм.

Подамо систему лінійних форм (3.1) у матричному вигляді:

$$\begin{bmatrix} s_1 \\ s_2 \\ \cdots \\ s_k \\ \cdots \\ s_m \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1r} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2r} & \cdots & a_{2p} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{k1} & a_{k2} & \cdots & a_{kr} & \cdots & a_{kp} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mr} & \cdots & a_{mp} \end{bmatrix} \begin{bmatrix} t_1 \\ t_2 \\ \cdots \\ t_r \\ \cdots \\ t_p \end{bmatrix}.$$

Таблиці жорданових виключень багато в чому нагадують матричне надання системи лінійних форм (3.1). Так, початкова таблиця має вигляд:

$$
\begin{array}{r|ccccccc|}
 & t_1 & t_2 & \cdots & t_r & \cdots & t_p \\
\hline
s_1 = & a_{11} & a_{12} & \cdots & a_{1r} & \cdots & a_{1p} \\
s_2 = & a_{21} & a_{22} & \cdots & a_{2r} & \cdots & a_{2p} \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
s_k = & a_{k1} & a_{k2} & \cdots & a_{kr} & \cdots & a_{kp} \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
s_m = & a_{m1} & a_{m2} & \cdots & a_{mr} & \cdots & a_{mp} \\
\hline
\end{array}
$$

Домовимося елемент a_{kr} називати *головним,* або *розв'язуючим* елементом; k-й рядок матриці – *направляючим,* або *розв'язуючим* рядком; r-й стовпець – *направляючим,* або *розв'язуючим* стовпцем.

Результуюча таблиця жорданових виключень набуває вигляду:

$$
\begin{array}{r|ccccccc|}
 & t_1^{\text{рез}} & t_2^{\text{рез}} & \cdots & t_r^{\text{рез}} & \cdots & t_p^{\text{рез}} \\
\hline
s_1^{\text{рез}} = & b_{11} & b_{12} & \cdots & b_{1r} & \cdots & b_{1p} \\
s_2^{\text{рез}} = & b_{21} & b_{22} & \cdots & b_{2r} & \cdots & b_{2p} \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
s_k^{\text{рез}} = & b_{k1} & b_{k2} & \cdots & b_{kr} & \cdots & b_{kp} \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
s_m^{\text{рез}} = & b_{m1} & b_{m2} & \cdots & b_{mr} & \cdots & b_{mp} \\
\hline
\end{array}
$$

У результуючій таблиці як незалежна змінна $s_k^{\text{рез}}$ виступає та змінна, яка в початковій таблиці відповідала незалежній змінній t_r, а як незалежна $t_r^{\text{рез}}$ – s_k. Інші змінні в результуючій таблиці співпадають з відповідними змінними початкової таблиці.

В умовах прикладу (3.2) початкова і результуюча таблиці мають вигляд:

	t_1	t_2	t_3
$s_1 =$	1	2	-1
$s_2 =$	-1	-3	2

	t_1	s_1	t_3
$t_2 =$	$-\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$
$s_2 =$	$\frac{1}{2}$	$-\frac{3}{2}$	$\frac{1}{2}$

Зворотний перехід від результуючої таблиці жорданових виключень прикладу (3.2) до матричного вигляду відповідає запису

$$\begin{bmatrix} t_2 \\ s_2 \end{bmatrix} = \begin{bmatrix} -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{3}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} t_1 \\ s_1 \\ t_3 \end{bmatrix}.$$

Загальний вигляд системи лінійних форм (3.1) у матричному вигляді відповідає запису $\bar{s} = \mathbf{A}\bar{t}$ **.**

3.3. Виведення правил жорданових виключень

Правила жорданових виключень визначають порядок обчислення елементів нової (результуючої) таблиці. Існує чотири правила жорданових виключень, що формулюються наступним чином.

Правило 1. Головний елемент нової таблиці дорівнює оберненому головному елементові старої таблиці:

$$b_{kr} = \frac{1}{a_{kr}}, \quad k \in \{i\}_1^m, \quad r \in \{j\}_1^p \ . \tag{3.3}$$

Правило 2. Нові елементи направляючого рядка дорівнюють відповідним старим, узятим із зворотним знаком і поділеним на головний елемент:

$$b_{kj} = -\frac{a_{kj}}{a_{kr}}, \quad j = \overline{1,p}, \quad j \neq r \ . \tag{3.4}$$

Правило 3.. Нові елементи направляючого стовпця рівні відповідним старим, поділеним на головний елемент:

$$b_{ir} = \frac{a_{ir}}{a_{kr}}, \quad i = \overline{1,m}, \quad i \neq k \ . \tag{3.5}$$

Правило 4.. Інші елементи нової таблиці, не розташовані на направляючому рядку або стовпці, визначаються за схемою «чотирикутника»:

$$b_{ij} = a_{ij} - \frac{a_{ir}a_{kj}}{a_{kr}}, \quad i = \overline{1,m}, i \neq k, \quad j = \overline{1,p}, j \neq r \ . \tag{3.6}$$

У правилі (3.6) всі елементи старої таблиці, що беруть участь в обчисленні нового елемента b_{ij}, розташовуються у вершинах чотирикутника. Останній утвориться в результаті взаємного перетину направляючого рядка і направляючого стовпця з двома перпендикулярами, опущеними з головного елемента a_{ij} на направляючий стовпець і направляючий рядок.

Виведення правил жорданових виключень здійснюються в повній відповідності з алгебраїчними перетвореннями, які робилися над системою (3.2) у підрозділі 3.1.

Отже, для того щоб у загальному вигляді здійснити транспозицію довільної залежної змінної з довільною незалежною , необхідно спочатку розв'язати *k*-е рівняння в системі лінійних форм (3.1) відносно незалежної змінної :

. (3.7)

Потім у всі інші рівняння системи (3.1), крім *k*-го, замість змінної підставити її вираз (3.7):

.

У цьому виразі наведемо подібні:

. (3.8)

Співвідношення (3.7) і (3.8) відповідають результуючій системі лінійних форм, що утворюється в результаті транспозиції змінних і у початковій системі (3.1). Тому коефіцієнти при змінних у виразах (3.7) і (3.8) відповідають новим коефіцієнтам у результуючій таблиці жорданових виключень

$$
\begin{cases}
t_r = b_{kr}s_k + \sum\limits_{\substack{j=1;\\ j\neq r}}^{p} b_{kj}t_j; \\
s_i = b_{ir}s_k + \sum\limits_{\substack{j=1;\\ j\neq r}}^{p} b_{ij}t_j, \quad \text{где} \quad i = \overline{1,m}, \quad i \neq k,
\end{cases}
$$

і знаходяться в повній відповідності з раніше наведеними правилами (3.3) – (3.6):

$$
\boxed{b_{kr} = \frac{1}{a_{kr}}, \quad k \in \{i\}_1^m, \quad r \in \{j\}_1^p};
$$

$$
\boxed{b_{kj} = -\frac{a_{kj}}{a_{kr}}, \quad j = \overline{1,p}, \quad j \neq r};
$$

$$
\boxed{b_{ir} = \frac{a_{ir}}{a_{kr}}, \quad i = \overline{1,m}, \quad i \neq k};
$$

$$
\boxed{b_{ij} = a_{ij} - \frac{a_{ir}a_{kj}}{a_{kr}}, \quad i = \overline{1,m}, i \neq k, \quad j = \overline{1,p}, j \neq r}.
$$

Легко перевірити, що в умовах прикладу (3.2) наведені правила жорданових виключень приводять до тих самих результатів, що й відповідні алгебраїчні перетворення (див. вихідну і результуючі таблиці жорданових виключень для розглянутого прикладу).

> Обчислювальну процедуру, зв'язану з перетворенням початкової таблиці жорданових виключень у результуючу називають *кроком жорданових виключень*.

3.4. Обертання матриць за допомогою жорданових виключень

Нехай задана система лінійних форм у матричному вигляді:

, (3.9)

де **A** – *невироджена квадратна* матриця, тобто визначник матриці не дорівнює нулю і число залежних перемінних дорівнює числу незалежних змінних,

Помножимо обидві частини рівності (3.9) на обернену матрицю *«зліва»:*

,

або

.

Поміняємо місцями ліву і праву частини в отриманій рівності:

$$\bar{t} = \mathbf{A}^{-1}\,\bar{s}\ . \tag{3.10}$$

Якщо тепер порівняти (3.9) і (3.10), то можна дійти до висновку, що транспозиція векторів залежних і незалежних змінних у початковій системі лінійних форм (3.9) автоматично перетворює матрицю коефіцієнтів **A** в обернену . Для транспозиції векторів треба здійснити n кроків жорданових виключень.

> Будь-яка невироджена матриця розмірності $(n \times n)$, якщо з нею здійснити n кроків жорданових виключень, автоматично перетвориться в обернену матрицю.

3.5. Вирішення системи лінійних рівнянь при числі змінних, рівному числу рівнянь

Нехай задана в матричному вигляді система m лінійних рівнянь із числом невідомих (змінних) n, рівним кількості рівнянь, тобто n=m :

$$\mathbf{A}\,\overline{x} = \overline{b} \ , \tag{3.11}$$

де **A** – неособлива матриця коефіцієнтів системи; $\overline{x}$ – вектор-стовпець змінних; $\overline{b}$ – вектор-стовпець вільних членів.

Помножимо обидві частини рівності (3.1) на обернену матрицю $\mathbf{A}^{-1}$ *«зліва»:*

$$\mathbf{A}^{-1}\mathbf{A}\,\overline{x} = \mathbf{A}^{-1}\,\overline{b} \ ,$$

або

$$\overline{x} = \mathbf{A}^{-1}\,\overline{b} \ . \tag{3.12}$$

З порівняння (3.11) і (3.12) стає очевидним, що рішенням системи лінійних рівнянь є добуток оберненої матриці коефіцієнтів системи на вектор-стовпець вільних членів.

Для вирішення невиродженої системи лінійних рівнянь досить знайти обернену матрицю коефіцієнтів цієї системи і помножити її на вектор вільних членів.

3.6. Вирішення системи лінійних рівнянь при числі змінних, що перевищує число рівнянь

У випадку, коли система включає лінійні рівняння в кількості m, що менше числа невідомих n, вирішення задачі може бути також знайдене за допомогою жорданових виключень.

Вирішити систему лінійних рівнянь при $n > m$ – це значить виразити m змінних задачі через ті p змінних, що залишилися, де $p = n - m$. Отримане рішення називають *загальним* рішенням системи лінійних рівнянь щодо m обраних змінних.

Розіб'ємо змінні задачі на залежні й незалежні. Тоді (3.11) набуде вигляду:

$$\mathbf{A}_s\mathbf{s} + \mathbf{A}_t\mathbf{t} = \mathbf{b}, \tag{3.13}$$

де $\mathbf{A}_s$ – матриця коефіцієнтів при залежних змінних розміру $m \times m$; $\mathbf{A}_t$ – матриця коефіцієнтів при залежних змінних розміру $m \times p$; $\mathbf{s}$ – вектор залежних змінних; $\mathbf{t}$ – вектор незалежних змінних.

Розв’яжемо (3.13) щодо вектора залежних змінних:

$$\mathbf{s} = -\mathbf{A}_s^{-1}\mathbf{A}_t\mathbf{t} + \mathbf{A}_s^{-1}\mathbf{b} . \tag{3.14}$$

Отриманий вираз є шуканим рішенням.

Із загального рішення (3.14) завжди можна одержати нескінченну множину *часткових* рішень. Для цього незалежним змінним, що стоять у правій частині загального рішення, привласнюють довільні значення. Потім підставляють ці значення в загальне рішення і обчислюють значення залежних змінних.

Часткові рішення дозволяють перевірити правильність знаходження загального рішення. Якщо часткове рішення при його підстановці у початкову систему лінійних рівнянь (3.13) перетворює останню в тотожність, то загальне рішення знайдено вірно.

Приклад. Нехай задана система двох лінійних рівнянь:

$$\begin{cases} x_1 - x_2 + x_3 - x_4 + 2 = 0\ ; \\ 2x_1 + x_2 - x_3 + x_4 - 1 = 0, \end{cases} \tag{3.15}$$

для якої n=4, m=p=2, n>m. Необхідно вирішити її щодо змінних x_1 і x_3.

Розв'язання. Приведемо (3.15) до вигляду (3.13):

$$\begin{bmatrix} 1 & 1 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_3 \end{bmatrix} + \begin{bmatrix} -1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_2 \\ x_4 \end{bmatrix} = \begin{bmatrix} -2 \\ 1 \end{bmatrix}.$$

Знайдемо рішення відповідно до (3.14):

$$\begin{bmatrix} x_1 \\ x_3 \end{bmatrix} = -\begin{bmatrix} \frac{1}{3} & \frac{1}{3} \\ \frac{2}{3} & -\frac{1}{3} \end{bmatrix} \begin{bmatrix} -1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_2 \\ x_4 \end{bmatrix} + \begin{bmatrix} \frac{1}{3} & \frac{1}{3} \\ \frac{2}{3} & -\frac{1}{3} \end{bmatrix} \begin{bmatrix} -2 \\ 1 \end{bmatrix} =$$

$$= \begin{bmatrix} 0 & 0 \\ -1 & -1 \end{bmatrix} \begin{bmatrix} x_2 \\ x_4 \end{bmatrix} + \begin{bmatrix} -\frac{1}{3} \\ -\frac{5}{3} \end{bmatrix},$$

або

$$\begin{cases} x_1 = -\frac{1}{3}\ ; \\ x_3 = x_2 + x_4 - \frac{5}{3}. \end{cases} \tag{3.16}$$

Будь-яке часткове рішення, наприклад, $\mathbf{x}^{\mathrm{T}} = \begin{bmatrix} -\frac{1}{3} & 1 & \frac{1}{3} & 1 \end{bmatrix}$, перетворює рівняння системи (3.15) у тотожність. Отже, загальне рішення (3.16) знайдено вірно.

Розв'язання систем лінійних рівнянь при числі змінних, що перевищує число рівнянь, у формі (3.14) не є раціональним. На практиці вирішення системи одержують іншим шляхом. Покажемо, як це робиться на прикладі вирішення системи (3.15).

Розв'язання системи лінійних рівнянь шляхом приведення її до системи лінійних форм. Введемо в систему лінійних рівнянь (3.15) m фіктивних залежних змінних (x_5 і x_6 в умовах прикладу), що завжди рівні нулю і заміняють нульові частини в рівняннях системи, і одну фіктивну незалежну змінну (x_7 в умовах прикладу), що завжди дорівнює одиниці і перетворює вільні члени рівнянь у коефіцієнти:

$$\begin{cases} x_5 = x_1 - x_2 + x_3 - x_4 + 2x_7\ ; \\ x_6 = 2x_1 + x_2 - x_3 + x_4 - 1x_7 . \end{cases}$$

Введення фіктивних змінних дозволяє систему лінійних рівнянь подати у виді системи лінійних форм типу (3.1) і побудувати для неї таблицю жорданових виключень:

	x_1	x_2	x_3	x_4	$x_7 = 1$
$x_5 = 0 =$	1	−1	1	−1	2
$x_6 = 0 =$	2	1	−1	1	−1

Послідовно замінюючи за допомогою жорданових виключень фіктивні залежні змінні x_5 і x_6 незалежними змінними x_1 і x_3 (у будь-якому порядку, аби головний елемент не дорівнював нулю), одержимо шукане рішення. У загальному випадку рішення досягається за m кроків, в умовах прикладу достатньо двох кроків.

Нехай на першому кроку в транспозиції беруть участь перемінні x_1 і x_6. Тоді перший крок жорданових виключень приведе до таблиці

	$x_6 = 0$	x_2	x_3	x_4	$x_7 = 1$
$x_5 = 0 =$	$\frac{1}{2}$	$-\frac{3}{2}$	$\frac{3}{2}$	$-\frac{3}{2}$	$\frac{5}{2}$
$x_1 =$	$\frac{1}{2}$	$-\frac{1}{2}$	$\frac{1}{2}$	$-\frac{1}{2}$	$\frac{1}{2}$

,

перший стовпець якої відповідає фіктивній змінній. Якщо тепер від таблиці перейти до алгебраїчного запису системи рівнянь, то можна побачити, що система ні якою мірою не залежатиме від коефіцієнтів у першому стовпці, оскільки завжди $x_6 = 0$. Як наслідок, перший стовпець з фіктивною змінною x_6 доцільно вилучити з таблиці. Надалі в таблиці зовсім не будемо відводити місце для стовпців, що містять нульові незалежні фіктивні змінні, і не будемо гаяти час на обчислення значень відповідних коефіцієнтів.

Другий крок жорданових виключень приводить до остаточної таблиці

	x_2	x_4	$x_7 = 1$
$x_3 =$	1	1	$-\frac{5}{3}$
$x_1 =$	0	0	$-\frac{1}{3}$

,

яка дає шукане рішення:

$$\begin{cases} x_3 = x_2 + x_4 - \dfrac{5}{3} \ ; \\ x_1 = -\dfrac{1}{3} \ . \end{cases} \qquad (3.17)$$

Відмінність рішення (3.17) від попереднього (3.16) не має принципового значення. Воно полягає тільки в порядку проходження рівностей і пояснюється вибором змінних x_1 і x_6 на першому кроці жорданових виключень. Вибір же змінних x_1 і x_5 (або x_3 і x_6) привів би до результату (3.16).

Розв'язання системи лінійних рівнянь шляхом приведення її до системи лінійних форм має місце і при числі змінних n, рівному числу рівнянь m. У цьому випадку m кроків жорданових виключень приводять початкову систему лінійно незалежних рівнянь типу (3.11) до рішення (3.12). При цьому рішення не буде містити незалежних перемінних, тобто загальне рішення буде співпадати з єдиним частковим. При лінійній залежності системи (3.11) жорданові виключення в повному обсязі здійснити не вдасться через появу нульових елементів як головних.

3.7. Контрольні запитання, приклади і вправи

1. Привести загальну форму запису системи лінійних форм у матричному виді.

2. Привести загальну форму запису системи лінійних рівнянь у матричному виді.

3. Привести загальну форму запису рішення системи лінійних рівнянь у матричному виді при числі змінних, рівному числу рівнянь.

4. Привести загальну форму запису рішення системи лінійних рівнянь у матричному виді при числі змінних, що перевищує число рівнянь.

5. Яке призначення жорданових виключень?

6. Сформулювати правила жорданових виключень.

7. Який стовпець і який рядок називають розв'язуючими (направляючими)?

8. Який елемент таблиці жорданових виключень називають розв'язуючим (головним)?

9. Здійснити заміну залежної змінної s_2 на незалежну t_3 в системі лінійних форм

$$\begin{bmatrix} s_1 \\ s_2 \end{bmatrix} = \begin{bmatrix} 2 & 1 & 6 \\ 3 & 4 & -2 \end{bmatrix} \begin{bmatrix} t_1 \\ t_2 \\ t_3 \end{bmatrix}.$$

Розв'язання. Вирішення задачі досягається за допомогою одного кроку жорданових виключень.

Початкова таблиця жорданових виключень в умовах прикладу має вигляд:

	t_1	t_2	t_3
$s_1 =$	2	1	6
$s_2 =$	3	4	-2

Змінні, що беруть участь у транспозиції, визначають направляючий рядок і направляючий стовпець. Вони в таблиці виділені сірим фоном. Головний елемент виділений більш насиченим сірим фоном.

Для здійснення кроку жорданових виключень треба обчислити елементи нової таблиці жорданових виключень відповідно до правил (3.3) – (3.6).

Так, відповідно до (3.3), новий головний елемент $b_{23} = \frac{1}{a_{23}} =$ $= \frac{1}{-2} = -0{,}5$. Нові елементи направляючого рядка, крім головного, визначаються співвідношенням (3.4), тобто $b_{21} = -\frac{a_{21}}{a_{23}} = -\frac{3}{-2} = 1{,}5;$ $b_{21} = -\frac{a_{21}}{a_{23}} = -\frac{3}{-2} = 1{,}5;$ $b_{22} = -\frac{a_{22}}{a_{23}} = -\frac{4}{-2} = 2.$ Новий елемент направляючого стовпця відповідно до правила (3.5) $b_{13} = \frac{a_{13}}{a_{23}} = \frac{6}{-2} = -3$.

Інші нові елементи обчислюються за четвертим правилом жорданових виключень (3.6):

$$b_{11} = a_{11} - \frac{a_{21}a_{13}}{a_{23}} = 2 - \frac{3 \cdot 6}{-2} = 11 \ ;$$

$$b_{12} = a_{12} - \frac{a_{22}a_{12}}{a_{23}} = 1 - \frac{4 \cdot 6}{-2} = 13 \ .$$

Результуюча таблиця жорданових виключень для розглянутого приклада має вигляд:

	t_1	t_2	s_2
$s_1 =$	11	–13	-3
$t_3 =$	1,5	2	-0,5

Результуючій таблиці відповідає нова (шукана) система лінійних форм:

$$\begin{bmatrix} s_1 \\ t_3 \end{bmatrix} = \begin{bmatrix} 11 & 13 & -3 \\ 1{,}5 & 2 & -0{,}5 \end{bmatrix} \begin{bmatrix} t_1 \\ t_2 \\ s_2 \end{bmatrix} .$$

10. За допомогою жорданових виключень замінити в системі лінійних форм

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 & -1 & 0 & 1 \\ 0 & -1 & 1 & -1 \\ 1 & 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} x_4 \\ x_5 \\ x_6 \\ x_7 \end{bmatrix}$$

залежні змінні x_1 і x_1 на незалежні x_5 і x_6 .

11. Розв’язати систему лінійних форм

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 & 2 & 1 \\ -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_3 \\ x_4 \\ x_5 \end{bmatrix}$$

змінних x_3 і x_5.

12. В умовах вправи 11 замінити змінну x_2 на x_4.

13. Яке застосування в лінійній алгебрі знаходять жорданові виключення.

14. Вирішити систему лінійних рівнянь

$$\begin{cases} x_1 - x_2 + x_3 - 2x_4 + 3 = 0; \\ \quad 2x_1 - x_2 + x_4 + 2 = 0 \end{cases} \tag{3.18}$$

щодо перемінних x_1 і x_3.

Розв'язання. Вирішити систему лінійних рівнянь (3.18) щодо змінних x_1 і x_3 – це значить виявити змінні x_1 і x_3 через змінні x_2 і x_4. Це можна зробити за допомогою жорданових виключень. Для цього введемо вектор незалежних змінних $\mathbf{s} = \begin{bmatrix} s_1 \\ s_2 \end{bmatrix}$, кожна складова якого приймає тільки нульове значення, і незалежну змінну t, що приймає єдине значення, а саме 1. Тоді задану систему лінійних рівнянь можна подати у вигляді системи лінійних форм

$$\begin{bmatrix} s_1 \\ s_2 \end{bmatrix} = \begin{bmatrix} 1 & -1 & 1 & -2 & 3 \\ 2 & -1 & 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ t \end{bmatrix}.$$

Здійснимо два кроки жорданових виключень, в яких залежну змінну s_1 замінимо на незалежну x_1, а s_2 – на x_3. Одержимо шукане *загальне* рішення

$$\begin{bmatrix} x_1 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0,5 & -0,5 & -1 \\ 0,5 & 2,5 & -2 \end{bmatrix} \begin{bmatrix} x_2 \\ x_4 \\ 1 \end{bmatrix}. \tag{3.19}$$

Перевіримо правильність рішення за допомогою підстановки у вихідну систему (3.18) якогось часткового рішення. Для одержання часткового рішення надамо незалежним змінним x_2 і x_4 які-небудь значення (нульові значення не рекомендуються) і підставимо їх у (3.19) для обчислення залежних змінних x_1 і x_3. Наприклад, нехай $x_2 = 2$ і $x_4 = 2$. Тоді відповідно до (3.19) $x_1 = -1$ і $x_3 = 4$. Знайдене часткове рішення $\bar{x}^T = \begin{bmatrix} -1 & 2 & 4 & 2 \end{bmatrix}$ перетворює рівняння системи (3.18) в тотожності. Отже, загальне рішення (3.19) знайдено правильно.

15. Вирішити систему лінійних рівнянь

$$\begin{cases} x_1 - x_2 + x_4 - 1 = 0; \\ x_2 - 2x_3 + 1 = 0; \\ 2x_1 + x_3 - x_4 + 2 = 0 \end{cases}$$

щодо змінних x_1, x_2 , x_4 .

16. Вирішити систему лінійних рівнянь

$$\begin{cases} x_1 + x_2 + x_4 - 2 = 0; \\ x_2 + 2x_3 - x_4 + 1 = 0 \end{cases}$$

щодо змінних x_2 , x_3 .

17. Знайти розв'язок системи

$$\begin{cases} 2x_1 - x_3 - x_4 = 0; \\ x_1 + x_2 - 2x_4 - 1 = 0; \\ -x_2 + 2x_3 + 2 = 0 \end{cases}$$

щодо змінних x_1, x_3 , x_4 .

18. Вирішити систему лінійних рівнянь

$$\begin{cases} x_1 + 3x_2 = 9; \\ 2x_1 + 7x_2 = 20 \end{cases}$$

за допомогою жорданових виключень шляхом приведення її до системи лінійних форм.

19. Яку матрицю називають виродженою?

20. Яку матрицю називають оберненою? Як позначається обернена матриця?

21. Які матриці не мають обернених?

22. За допомогою жорданових виключень знайти матрицю $\mathbf{A}^{-1}$, що буде оберненою відносно матриці $\mathbf{A} = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 3 & 1 \\ 1 & 0 & 1 \end{bmatrix}$.

Розв'язання. Для вирішення задачі надамо матрицю $\mathbf{A}$ як матрицю коефіцієнтів системи лінійних форм $\mathbf{s} = \mathbf{A}\mathbf{t}$:

$$\begin{bmatrix} s_1 \\ s_2 \\ s_3 \end{bmatrix} = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 3 & 1 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} t_1 \\ t_2 \\ t_3 \end{bmatrix}. \tag{3.20}$$

Перетворимо (3.20) у систему $\mathbf{t} = \mathbf{A}^{-1}\mathbf{s}$ з шуканою оберненою матрицею $\mathbf{A}^{-1}$. Для цього в системі (3.20) здійснимо послідовно три кроки жорданових виключень (у загальному випадку n кроків), щораз обираючи як головній елемент один із діагональних, і тільки діагональних. Мета жорданових виключень – замінити кожну залежну змінну s_i незалежної t_i, $(i = \overline{1,3})$. Послідовність замін несуттєва.

Після проведення зазначених кроків жорданових виключень одержимо нову систему лінійних форм

$$
\begin{bmatrix} t_1 \\ t_2 \\ t_3 \end{bmatrix} = \begin{bmatrix} \frac{3}{7} & -\frac{1}{7} & \frac{1}{7} \\ \frac{1}{7} & \frac{2}{7} & -\frac{2}{7} \\ -\frac{3}{7} & \frac{1}{7} & \frac{6}{7} \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \\ s_3 \end{bmatrix},
$$

в якій матриця коефіцієнтів являє собою шукану матрицю, обернену щодо заданої.

Для перевірки правильності знайденого рішення необхідно перемножити вихідну і шукану матриці

$$
\mathbf{A} \cdot \mathbf{A}^{-1} = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 3 & 1 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} \frac{3}{7} & -\frac{1}{7} & \frac{1}{7} \\ \frac{1}{7} & \frac{2}{7} & -\frac{2}{7} \\ -\frac{3}{7} & \frac{1}{7} & \frac{6}{7} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \mathbf{I} .
$$

Оскільки добутком цих матриць є одинична матриця, знайдена матриця згідно з визначенням є оберненою, тобто рішення знайдено вірно.

23. Визначити матрицю $\mathbf{A}^{-1}$, якщо $\mathbf{A} = \begin{bmatrix} 2 & 1 \\ 1 & -1 \end{bmatrix}$.

24. Визначити матрицю $\mathbf{A}^{-1}$, обернену відносно матриці

$$
\mathbf{A} = \begin{bmatrix} 1 & 2 & 2 \\ 0 & 3 & 1 \\ 1 & 1 & 0 \end{bmatrix} .
$$

25. Дати порівняльну оцінку методам вирішення системи лінійних рівнянь при числі змінних, що співпадає з числом рівнянь, за допомогою приведення її до системи лінійних форм і шляхом обчислення оберненої матриці коефіцієнтів. Які недоліки і переваги кожного методу?

26. Вирішити систему лінійних рівнянь, задану умовами вправи 18, шляхом обчислення оберненої матриці коефіцієнтів.

3.8. Фонд індивідуальних завдань

Індивідуальне завдання №10. Вирішити систему лінійних рівнянь за допомогою жорданових виключень методом перетворення початкової системи в систему лінійних форм, якщо початкова система включає такі рівняння:

1.
$$\begin{cases} x_1 - 2x_2 + x_4 - 1 = 0; \\ x_2 - x_3 + 2 = 0; \\ 2x_1 + x_3 - x_4 + 1 = 0. \end{cases}$$
Вирішити щодо змінних x_1, x_2, x_4.

2.
$$\begin{cases} 2x_1 + x_3 - 2x_4 - x_5 + 1 = 0; \\ x_2 + x_4 + 2x_5 = 0. \end{cases}$$
Вирішити щодо змінних x_1, x_2.

3.
$$\begin{cases} x_1 - 2x_2 + x_4 - 1 = 0; \\ x_2 - x_3 + 2 = 0; \\ 2x_1 + x_3 - x_4 + 1 = 0. \end{cases}$$
Вирішити щодо змінних x_2, x_3, x_4.

4.
$$\begin{cases} 2x_1 + x_3 - 2x_4 - x_5 + 1 = 0; \\ x_2 + x_4 + 2x_5 = 0. \end{cases}$$
Вирішити щодо змінних x_1, x_4.

5.
$$\begin{cases} x_1 - 2x_2 + x_4 - 1 = 0; \\ x_2 - x_3 + 2 = 0; \\ 2x_1 + x_3 - x_4 + 1 = 0. \end{cases}$$
Вирішити щодо змінних x_1, x_3, x_4.

6.
$$\begin{cases} 2x_1 + x_3 - 2x_4 - x_5 + 1 = 0; \\ x_2 + x_4 + 2x_5 = 0. \end{cases}$$
Вирішити щодо змінних x_1, x_5.

7.

$$\begin{cases} x_1 - 2x_2 + x_4 - 1 = 0; \\ x_2 - x_3 + 2 = 0; \\ 2x_1 + x_3 - x_4 + 1 = 0. \end{cases}$$

Вирішити щодо змінних x_1, x_2, x_3.

8.

$$\begin{cases} 2x_1 + x_3 - 2x_4 - x_5 + 1 = 0; \\ x_2 + x_4 + 2x_5 = 0. \end{cases}$$

Вирішити щодо змінних x_2, x_4.

9.

$$\begin{cases} 2x_1 - x_2 + x_5 - 1 = 0; \\ 2x_3 - x_4 - x_5 = 0; \\ x_1 + x_2 + x_3 - x_4 - 2 = 0. \end{cases}$$

Вирішити щодо змінних x_1, x_2, x_3.

10.

$$\begin{cases} 2x_1 + x_3 - 2x_4 - x_5 + 1 = 0; \\ x_2 + x_4 + 2x_5 = 0. \end{cases}$$

Вирішити щодо змінних x_2, x_5.

11.

$$\begin{cases} 2x_1 - x_2 + x_5 - 1 = 0; \\ 2x_3 - x_4 - x_5 = 0; \\ x_1 + x_2 + x_3 - x_4 - 2 = 0. \end{cases}$$

Вирішити щодо змінних x_1, x_2, x_4.

12.

$$\begin{cases} 2x_1 + x_3 - 2x_4 - x_5 + 1 = 0; \\ x_2 + x_4 + 2x_5 = 0. \end{cases}$$

Вирішити щодо змінних x_2, x_4.

13.

$$\begin{cases} 2x_1 - x_2 + x_5 - 1 = 0; \\ 2x_3 - x_4 - x_5 = 0; \\ x_1 + x_2 + x_3 - x_4 - 2 = 0. \end{cases}$$

Вирішити щодо змінних x_1, x_2, x_5.

14.

$$\begin{cases} 2x_1 + x_3 - 2x_4 - x_5 + 1 = 0; \\ x_2 + x_4 + 2x_5 = 0. \end{cases}$$

Вирішити щодо змінних x_3, x_5.

15.
$$\begin{cases} 2x_1 - x_2 + x_5 - 1 = 0; \\ 2x_3 - x_4 - x_5 = 0; \\ x_1 + x_2 + x_3 - x_4 - 2 = 0. \end{cases}$$
Вирішити щодо змінних x_1, x_3, x_4.

16.
$$\begin{cases} 2x_1 + x_3 - 2x_4 - x_5 + 1 = 0; \\ x_2 + x_4 + 2x_5 = 0. \end{cases}$$
Вирішити щодо змінних x_4, x_5.

17.
$$\begin{cases} -2x_1 + x_2 + 2x_4 - 1 = 0; \\ -x_2 - x_3 - x_4 + 2 = 0; \\ x_1 - 2x_2 + 2x_3 + 1 = 0. \end{cases}$$
Вирішити щодо змінних x_1, x_2, x_3.

18.
$$\begin{cases} -x_1 + 2x_2 - x_3 + 2x_4 - x_5 - 1 = 0; \\ 2x_2 + 2x_3 + x_5 + 1 = 0. \end{cases}$$
Вирішити щодо перемінних x_2, x_5.

19.
$$\begin{cases} x_1 - x_4 + 2 = 0; \\ x_2 + x_3 - 1 = 0; \\ x_1 - x_2 + x_3 + x_4 = 0. \end{cases}$$
Вирішити щодо змінних x_1, x_2, x_3.

20.
$$\begin{cases} -2x_1 + x_2 + 2x_4 - 1 = 0; \\ -x_2 - x_3 - x_4 + 2 = 0; \\ x_1 - 2x_2 + 2x_3 + 1 = 0. \end{cases}$$
Вирішити щодо змінних x_1, x_2, x_4.

21.
$$\begin{cases} -x_1 + 2x_2 - x_3 + 2x_4 - x_5 - 1 = 0; \\ 2x_2 + 2x_3 + x_5 + 1 = 0. \end{cases}$$
Вирішити щодо змінних x_1, x_3.

22.
$$\begin{cases} -2x_1 + x_2 + 2x_4 - 1 = 0; \\ -x_2 - x_3 - x_4 + 2 = 0; \\ x_1 - 2x_2 + 2x_3 + 1 = 0. \end{cases}$$
Вирішити щодо змінних x_1, x_3, x_4.

23.
$$\begin{cases} -x_1 + 2x_2 - x_3 + 2x_4 - x_5 - 1 = 0; \\ 2x_2 + 2x_3 + x_5 + 1 = 0. \end{cases}$$
Вирішити щодо змінних x_1, x_5.

24.
$$\begin{cases} -2x_1 + x_2 + 2x_4 - 1 = 0; \\ -x_2 - x_3 - x_4 + 2 = 0; \\ x_1 - 2x_2 + 2x_3 + 1 = 0. \end{cases}$$
Вирішити щодо змінних x_2, x_3, x_4.

25.
$$\begin{cases} 2x_1 + x_3 + 2x_4 + 1 = 0; \\ x_2 + 2x_3 - 1 = 0; \\ x_1 - x_2 - x_3 + x_4 = 0. \end{cases}$$
Вирішити щодо змінних x_1, x_2, x_3.

26.
$$\begin{cases} 2x_1 + x_2 + x_3 + 3 = 0; \\ -x_1 - x_3 + 2x_4 + 1 = 0; \\ x_2 + x_3 - x_4 - 2 = 0. \end{cases}$$
Вирішити щодо змінних x_1, x_2, x_3.

27.
$$\begin{cases} 2x_1 - x_2 + 3x_3 - x_5 + 2 = 0; \\ x_2 + 2x_3 - x_4 - 1 = 0. \end{cases}$$
Вирішити щодо змінних x_1, x_4.

28.
$$\begin{cases} -x_1 + 2x_2 + x_3 + 2x_5 + 3 = 0; \\ x_1 - x_2 + 2x_4 - 1 = 0. \end{cases}$$
Вирішити щодо змінних x_1, x_3.

29.

$$\begin{cases} -2x_1 - x_2 - x_3 - 3 = 0; \\ -x_1 - x_3 + 2x_4 + 1 = 0; \\ x_2 + x_3 - x_4 - 2 = 0. \end{cases}$$

Вирішити щодо змінних x_1, x_2, x_4.

30.

$$\begin{cases} 2x_1 + x_3 - 2x_4 - x_5 + 1 = 0; \\ x_2 + x_4 + 2x_5 + 2 = 0. \end{cases}$$

Вирішити щодо змінних x_1, x_4.

Індивідуальне завдання №11. Вирішити систему лінійних рівнянь за допомогою жорданових виключень методом перетворення початкової системи в систему лінійних форм, якщо початкова система включає такі рівняння:

1.

$$\begin{cases} 2x_1 + 3x_2 + 4x_3 + 9 = 0; \\ 4x_2 + 11x_3 + 1 = 0; \\ 7x_1 - 5x_2 - 1 = 0. \end{cases}$$

2.

$$\begin{cases} 3x_1 - 5x_2 + 3x_3 + 46 = 0; \\ 1x_1 + 2x_2 + x_3 + 8 = 0; \\ x_1 - 7x_2 - 2x_3 + 5 = 0. \end{cases}$$

3.

$$\begin{cases} 2x_1 - x_2 + 5x_3 + 27 = 0; \\ 5x_1 + 2x_2 + 13x_3 + 70 = 0; \\ 3x_1 - x_3 - 2 = 0. \end{cases}$$

4.

$$\begin{cases} 4x_1 + x_2 - 3x_3 - 1 = 0; \\ 8x_1 + 3x_2 - 6x_3 - 1 = 0; \\ x_1 + x_2 - x_3 - 1 = 0. \end{cases}$$

5.

$$\begin{cases} x_1 - 4x_2 - 2x_3 = 0; \\ 3x_1 - 5x_2 - 6x_3 - 21 = 0; \\ 3x_1 + x_2 + x_3 - 4 = 0. \end{cases}$$

6.

$$\begin{cases} 3x_1 + x_2 - 2x_3 + 6 = 0; \\ 5x_1 - 3x_2 + x_3 - 4 = 0; \\ 4x_1 - 2x_2 - 3x_3 - 2 = 0. \end{cases}$$

7. $\begin{cases} 3x_1 + x_2 + 2x_3 + 11 = 0; \\ 2x_1 + 2x_2 - 3x_3 + 9 = 0; \\ x_1 - 5x_2 - 8x_3 + 23 = 0. \end{cases}$

8. $\begin{cases} 5x_1 + 6x_2 - 2x_3 + 12 = 0; \\ 2x_1 + 5x_2 - 3x_3 + 9 = 0; \\ 4x_1 - 3x_2 + 2x_3 - 15 = 0. \end{cases}$

9. $\begin{cases} 2x_1 + 3x_2 + 4x_3 + 15 = 0; \\ x_1 + x_2 + 5x_3 + 16 = 0; \\ 3x_1 - 2x_2 + x_3 + 1 = 0. \end{cases}$

10. $\begin{cases} 12x_1 - 13x_2 - 4x_3 - 10 = 0; \\ 7x_1 - 9x_2 - 11x_3 = 0; \\ 12x_1 - 17x_2 - 15x_3 - 7 = 0. \end{cases}$

11. $\begin{cases} 5x_1 - 19x_2 - x_3 + 26 = 0; \\ 2x_1 - 5x_2 - x_3 + 6 = 0; \\ 8x_1 - 31x_2 - 4x_3 + 35 = 0. \end{cases}$

12. $\begin{cases} 8x_1 - x_2 + 3x_3 + 22 = 0; \\ 4x_1 + x_2 + 6x_3 - 1 = 0; \\ 13x_1 + x_2 + 16x_3 + 5 = 0. \end{cases}$

13. $\begin{cases} 6x_1 + 3x_2 - 5x_3 = 0; \\ 9x_1 + 4x_2 - 7x_3 + 3 = 0; \\ 14x_1 + 6x_2 - 11x_3 + 6 = 0. \end{cases}$

14. $\begin{cases} 7x_1 - 2x_2 - 24 = 0; \\ 6x_1 - 4x_2 - 5x_3 - 37 = 0; \\ 4x_1 + 3x_2 + 6x_3 + 13 = 0. \end{cases}$

15. $\begin{cases} -5x_2 + x_3 + 23 = 0; \\ 3x_1 + 2x_2 - x_3 + 7 = 0; \\ x_1 + 6x_2 - 2x_3 - 21 = 0. \end{cases}$

16. $\begin{cases} 5x_1 - 5x_2 - 4x_3 - 3 = 0; \\ x_1 - x_2 + 5x_3 + 11 = 0; \\ 4x_1 - 3x_2 - 6x_3 - 9 = 0. \end{cases}$

17. $\begin{cases} 5x_1 - 9x_2 - 14x_3 + 6 = 0; \\ x_1 + 7x_2 + 5x_3 + 11 = 0; \\ 5x_1 - 21x_2 - 27x_3 - 5 = 0. \end{cases}$

18. $\begin{cases} x_1 - 5x_3 + 8 = 0; \\ x_1 - 3x_2 - 2x_3 - 1 = 0; \\ 2x_1 - x_2 - 6x_3 + 7 = 0. \end{cases}$

19. $\begin{cases} 12x_1 + 2x_3 + 22 = 0; \\ 2x_1 - 8x_2 + 5x_3 - 25 = 0; \\ 19x_1 - 11x_2 + 9x_3 - 4 = 0. \end{cases}$

20. $\begin{cases} 4x_1 - 7x_2 - 2x_3 = 0; \\ 2x_1 - 3x_2 - 4x_3 + 16 = 0; \\ 3x_1 - 8x_2 - 7x_3 + 22 = 0. \end{cases}$

21. $\begin{cases} 2x_1 + 3x_2 + 4x_3 + 8 = 0; \\ 4x_2 + 11x_3 + 19 = 0; \\ 7x_1 - 5x_2 - 17 = 0. \end{cases}$

22. $\begin{cases} 4x_1 - 3x_2 + x_3 + 43 = 0; \\ x_1 + x_2 - x_3 + 3 = 0; \\ 2x_1 + x_2 + 13 = 0. \end{cases}$

23. $\begin{cases} 2x_1 - 4x_2 + 7x_3 + 45 = 0; \\ 4x_1 - 2x_2 + 3x_3 + 23 = 0; \\ 7x_1 - 3x_2 + 8x_3 + 53 = 0. \end{cases}$

24. $\begin{cases} 2x_1 + 13x_2 + 37 = 0; \\ 5x_1 + 2x_2 - 8x_3 + 17 = 0; \\ 9x_1 + 19x_2 - 11x_3 + 70 = 0. \end{cases}$

25. $\begin{cases} 14x_1 - 5x_2 + 9x_3 + 38 = 0; \\ 5x_1 + x_2 + 7x_3 + 1 = 0; \\ 27x_1 - 5x_2 + 21x_3 + 62 = 0. \end{cases}$

26. $\begin{cases} 4x_1 - 5x_2 + 5x_3 + 18 = 0; \\ 5x_1 + x_2 - x_3 + 8 = 0; \\ 6x_1 - 4x_2 + 3x_3 + 2 = 0. \end{cases}$

27. $\begin{cases} x_1 - 5x_3 + 8 = 0; \\ x_1 - 3x_2 - 2x_3 - 1 = 0; \\ 2x_1 - x_2 - 6x_3 + 7 = 0. \end{cases}$

28. $\begin{cases} 7x_2 - 2x_3 - 8 = 0; \\ 5x_1 - 6x_2 + 4x_3 + 20 = 0; \\ 6x_1 + 4x_2 + 3x_3 + 7 = 0. \end{cases}$

29. $\begin{cases} 5x_1 - 6x_2 - 3x_3 + 16 = 0; \\ 7x_1 - 9x_2 - 4x_3 + 22 = 0; \\ 11x_1 - 14x_2 - 6x_3 + 35 = 0. \end{cases}$

30. $\begin{cases} 3x_1 + 8x_2 - x_3 + 34 = 0; \\ 6x_1 + 4x_2 + x_3 + 17 = 0; \\ 16x_1 + 13x_2 + x_3 + 52 = 0. \end{cases}$

Індивідуальне завдання №12. За допомогою жорданових виключень знайти обернену матрицю $\mathbf{A}^{-1}$ для матриці **A,** що відповідає однойменній матриці індивідуального завдання №1. Перевірити рівність добутку $\mathbf{A}^{-1}\mathbf{A}$ одиничній матриці $\mathbf{I}_3$.

Індивідуальне завдання №13. Виконати індивідуальне завдання №11 шляхом обчислення оберненої матриці коефіцієнтів методом жорданових виключень

Розділ 4. БЕЗУМОВНА ОПТИМІЗАЦІЯ

Безумовна оптимізація являє собою розділ вищої математики, в якому розглядаються задачі пошуку екстремальних значень неперервної функції і методи їх вирішення. При цьому на змінні функції не накладаються ніякі обмеження (умови). У цьому випадку область визначення функції збігається з областю припустимих рішень задачі. Якщо функція визначена в кожній точці n-мірного евклідового простору $\mathbf{R}^n$, де n – число змінних у завданні, то будь-яка точка простору може стати точкою екстремума.

Незважаючи на те, що задача безумовної оптимізації порівняно проста і добре вивчена, вона залишається центральною в теорії оптимізації, оскільки методи вирішення будь-яких задач умовної оптимізації в тій чи іншій мірі зв'язані з методами безумовної оптимізації.

У цьому розділі главі розглядаються аксіоматичні поняття теорії безумовної оптимізації, необхідні й достатні умови існування локальних екстремальних точок неперервних функцій і способи практичної перевірки виконання умов екстремальності в заданих точках.

4.1. Аксіоматика і формулювання задачі безумовної оптимізації

Нехай $y(\mathbf{x})$ – функція, визначена в n-мірному евклідовому просторі $\mathbf{R}^{\mathrm{n}}$, а $\mathbf{V}$ – область існування функції, $\mathbf{V} \subseteq \mathbf{R}^n$.

Задачі безумовної оптимізації бувають двох типів: задачі *мінімізації* і задачі *максимізації*.

Задачі мінімізації (максимізації) полягають у відшуканні найменшого (найбільшого) значення функції $y(\mathbf{x})$.

Для того, щоб вирішити задачу безумовної мінімізації досить знайти її мінімальне рішення $\mathbf{x}^{**}$, тобто зазначити таку точку $\mathbf{x}^{**}$ із простору $\mathbf{R}^{\mathrm{n}}$, для котрої справедлива нерівність $y\left(\mathbf{x}^{**}\right) \leq y(\mathbf{x})$ при будь-якій точці $\mathbf{x} \in \mathbf{V}$.

Для вирішення задачі безумовної максимізації досить знайти її максимальне рішення $\mathbf{x}^{**}$, тобто визначити таку точку $\mathbf{x}^{**} \in \mathbf{R}^{n}$, для котрої справедлива нерівність $y\left(\mathbf{x}^{**}\right) \geq y(\mathbf{x})$ при будь-якій точці $\mathbf{x} \in \mathbf{V}$.

Оптимізаційна задача називається нерозв'язною, якщо вона не має оптимального рішення. Зокрема, задача безумовної мінімізації (максимізації) буде нерозв'язною для монотонно спадної (зростаючої) функції, оскільки не представляється можливо визначити конкретну точку $\mathbf{x}^{**}$ і відповідне їй значення функції $y\left(\mathbf{x}^{**}\right)$.

Вирішити задачу безумовної оптимізації – значить або знайти її оптимальне рішення, або встановити її нерозв'язність.

Оптимальне значення функції $y\left(\mathbf{x}^{**}\right)$ (спрощено: y^{**}), якщо воно є, завжди є єдиним і називається глобальним або абсолютним оптимумом (максимумом або мінімумом). Для оптимального вирішення $\mathbf{x}^{**}$ таке твердження несправедливе.

Безумовне оптимальне рішення шукається серед множини екстремальних точок функції, тобто безумовне оптимальне рішення завжди є екстремальною точкою функції. Зворотне твердження невірно. Множина екстремальних точок може бути скінченною або нескінченною. Так, параболічні функції мають один екстремум, мінімум або максимум, а тригонометричні функції – нескінченну множину. Скінченна множина може бути порожньою, зокрема для монотонних функцій.

У загальному випадку функція $y(\mathbf{x})$ може мати не більше одного оптимального значення y^{**}, а різних екстремальних значень – у кількості не більше числа екстремальних точок цієї функції.

Екстремальні значення функції будемо позначати y^{*} і називати локальними або частинними оптимумами (мінімумами або максимумами), а екстремальні точки відповідно позначати $\mathbf{x}^{*}$ і називати локальними або частинними оптимальними рішеннями.

Як локальні оптимальні рішення функції виступають *точки локальних мінімумів* і *точки локальних максимумів*.

Визначення 4.1. Точка $\mathbf{x}^* \in \mathbf{R}^n$ називається точкою локального мінімуму функції, якщо існує окіл $\varepsilon\left(\mathbf{x}^*\right)$ точки $\mathbf{x}^*$ такий, що при всіх $\mathbf{x} \in \varepsilon\left(\mathbf{x}^*\right)$ виконується нерівність $y\left(\mathbf{x}^*\right) \le y(\mathbf{x})$, де $\mathbf{x}$ – n-мірний вектор змінних задачі.

Визначення 4.2. Точка $\mathbf{x}^* \in \mathbf{R}^n$ називається точкою локального максимуму функції $y(\mathbf{x})$, якщо існує окіл $\varepsilon\left(\mathbf{x}^*\right)$ точки $\mathbf{x}^*$ такий, що при всіх $\mathbf{x} \in \varepsilon\left(\mathbf{x}^*\right)$ виконується нерівність $y\left(\mathbf{x}^*\right) \ge y(\mathbf{x})$.

З визначень 4.1 – 4.2 випливає, що точки екстремумів функції завжди є внутрішніми точками області визначення цих функцій.

З визначення 4.1 виходить, що і при додатних, і при від'ємних приростах вектора змінних $\Delta\mathbf{x}$ приріст функції Δy^* в точці локального мінімуму $\mathbf{x}^*$ не може бути негативним (рис. 4.1):

$$\Delta y^* = y\left(\mathbf{x}^* + \Delta\mathbf{x}\right) - y\left(\mathbf{x}^*\right) \ge 0 \ . \tag{4.1}$$

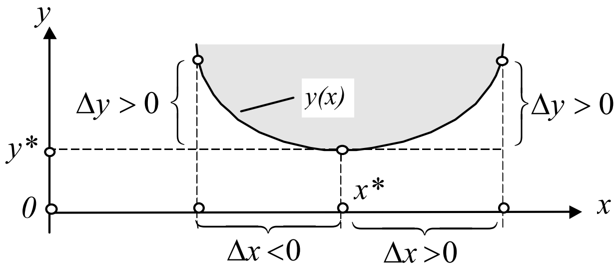


Рис.4.1 Графічна інтепретація властивості точки локального мінімуму для функції однієї змінної

З визначення 4.2 випливає, що і при будь-яких приростах вектора змінних $\Delta\mathbf{x}$ приріст функції Δy^* в точці локального максимуму $\mathbf{x}^*$ не може бути позитивним (рис.4.2):

$$\Delta y^* = y\left(\mathbf{x}^* + \Delta\mathbf{x}\right) - y\left(\mathbf{x}^*\right) \le 0 \ . \tag{4.2}$$

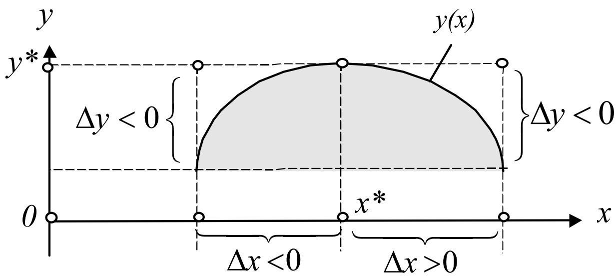


Рис.4.2 Графічна інтепретація властивості точки локального максимуму для функції однієї змінної

Співвідношення (4.1) і (4.2) відіграють ключову роль при виведенні необхідних і достатніх умов локального екстремуму функції.

Сучасні методи вирішення задач безумовної оптимізації, як правило, орієнтовані на відшукання тільки однієї локальної екстремальної точки. Тому

$\mathbf{x}^{**}$ і y^{**} – це то, що ми прагнемо знайти, а і y^{*} – це то, що ми звичайно знаходимо. Отже, для одержання вірного рішення необхідно або переконатися в тому, що функція має один екстремум, або визначити всі локальні екстремальні точки, серед яких шляхом порівняння відповідних значень функції вибрати глобальну. У першому випадку локальний екстремум одночасно є і глобальним $y^{**}=y^{*}$, у другому – оптимум визначається такими співвідношеннями:

при пошуку мінімуму –

$$y^{**} = y(\mathbf{x}^{**}) = \min\left(y(x)\right) = \min\left\{y_i^{*}\right\} = \min\left\{y\left(\mathbf{x}_i^{*}\right)\right\}, \quad i = \overline{1,\mathbf{M}} \,, \quad (4.3)$$

де $\mathbf{M}$ – потужність множини всіх локальних мінімумів функції;

при пошуку максимуму –

$$y^{**} = y(\mathbf{x}^{**}) = \max\left(y(x)\right) = \max\left\{y_i^{*}\right\} = \max\left\{y\left(\mathbf{x}_i^{*}\right)\right\}, \quad i = \overline{1,\mathbf{N}} \,, \quad (4.4)$$

де $\mathbf{N}$ – потужність множини всіх локальних максимумів функції.

Максимуми і мінімуми функцій мають важливі властивості. Так, якщо оптимальне рішення $\mathbf{x}^{**}$ мінімізує $y(\mathbf{x})$, воно мінімізує і функцію $a + by(\mathbf{x})$, де a - довільна константа , а b - позитивний коефіцієнт,

$$a + by\left(\mathbf{x}^{**}\right) = \min\left[a + by(\mathbf{x})\right]. \qquad (4.5)$$

Крім того, воно максимізує функцію $a - by\left(\mathbf{x}\right)$:

$$a - by\left(\mathbf{x}^{**}\right) = \max\left[a - by(\mathbf{x})\right]. \qquad (4.6)$$

З властивостей (4.5) – (4.6) випливає, що додавання до функції константи або позитивного коефіцієнта не впливає на розташування екстремумів функції. Іншим важливим наслідком властивості (4.6) є можливість простою зміною знака функції $y(\mathbf{x})$ задачу пошуку мінімуму звести до задачі максимізації і, навпаки, задачу пошуку максимуму звести до задачі мінімізації. А це значить, що ті самі методи відшукання екстремумів прийнятні як для вирішення задач мінімізації, так і для задач максимізації.

Надалі, якщо не буде вказано інше, мова йде про задачі мінімізації, хоча все, що викладається, однаковою мірою стосується і задач максимізації.

Зі співвідношень (4.3) і (4.4) видно, що задача пошуку глобального екстремуму зв'язана з пошуком локальних екстремумів. Більше того, задачі пошуку глобального екстремуму унімодальної (одноекстремальної) функції зводяться до задач пошуку локального екстремуму.

Задача пошуку безумовного глобального мінімуму формулюється таким чином: знайти мінімум функції $y(\mathbf{x})$, заданої в n-мірному евклідовому просторі $\mathbf{R}^n$. Аналітичний запис цієї задачі має вигляд:

$$y(\mathbf{x}) \to \min_{\mathbf{x} \in \mathbf{V} \subseteq \mathbf{R}^n} . \tag{4.7}$$

де $\mathbf{V}$ – область існування функції $y(\mathbf{x})$.

Задача пошуку безумовного локального мінімуму формулюється таким чином: знайти мінімум функції $y(\mathbf{x})$ в ε-околі деякої точки $\mathbf{x}^{(\circ)} \in \mathbf{V}$. Аналітичний запис цієї задачі має вигляд:

$$y(\mathbf{x}) \to \min_{\mathbf{x} \in \varepsilon\left(\mathbf{x}^{(\circ)}\right) \subset \mathbf{V} \subseteq \mathbf{R}^n} , \tag{4.8}$$

де $\varepsilon\left(\mathbf{x}^{(\circ)}\right)$ – ε-окіл деякої точки $\mathbf{x}^{(o)}$.

Пошук мінімуму функції в ε-околі точки $\mathbf{x}^{(o)}$ приводить до пошуку екстремальної точки, в даному випадку – точки мінімуму. Такі точки в теорії оптимізації часто називають точками *явного мінімуму* (*явного максимуму*).

У математичній літературі постановка задачі безумовної оптимізації дається у спрощеному вигляді – без вказівки приналежності точки $\mathbf{x}$ до області існування функції $\mathbf{V}$, оскільки цей факт має значення тільки для вибору початкової точки наближення при пошуку рішення некласичними методами. Надалі ми також будемо користуватися спрощеною постановкою, тобто припускати, що функція $y(\mathbf{x})$ існує в кожній точці простору $\mathbf{R}^n$.

4.2. Необхідні умови локального екстремуму

Найбільш важливим питанням безумовної мінімізації (максимізації) функції є установлення факту наявності екстремуму в заданій точці або, що те саме, установлення факту приналежності довільної точки з області визначення функції множині локальних екстремальних точок даної функції. З цією метою доведемо необхідні умови існування локального мінімуму функції.

Розкладемо функцію $y(\mathbf{x})$ в околі точки мінімуму $\mathbf{x}^*$ в ряд Тейлора:

$$y(\mathbf{x}^* + \Delta\mathbf{x}) = y(\mathbf{x}^*) + \left(\frac{\partial y}{\partial x_1}\right)^* \Delta x_1 + \left(\frac{\partial y}{\partial x_2}\right)^* \Delta x_2 + \ldots + \left(\frac{\partial y}{\partial x_n}\right)^* \Delta x_n +$$
$$+ R^*\left(\Delta\mathbf{x}^2, \Delta\mathbf{x}^3, \ldots\right), \tag{4.9}$$

де R – частина ряду розкладання, що залежить від приросту вектора змінних другого і більш високого порядку; індекс * при похідних є ознака їх обчислення в точці локального мінімуму.

Перепишемо (4.9) у матричному вигляді:

$$y(\mathbf{x}^* + \Delta\mathbf{x}) = y(\mathbf{x}^*) + \left(\frac{\partial y}{\partial \mathbf{x}}\right)^{*\mathrm{T}} \Delta\mathbf{x} + R^*\left(\Delta\mathbf{x}^2, \Delta\mathbf{x}^3, \ldots\right),$$

або

$$y(\mathbf{x}^* + \Delta\mathbf{x}) - y(\mathbf{x}^*) = \left(\frac{\partial y}{\partial \mathbf{x}}\right)^{*\mathrm{T}} \Delta\mathbf{x} + R^*\left(\Delta\mathbf{x}^2, \Delta\mathbf{x}^3, \ldots\right). \tag{4.10}$$

Права частина виразу (4.10) являє собою збільшення функції в точці локального мінімуму, тобто

$$\Delta y^* = \left(\frac{\partial\, y}{\partial\, \mathbf{x}} \right)^{*\mathrm{T}} \Delta\mathbf{x} + R^*\left(\Delta\mathbf{x}^2, \Delta\mathbf{x}^3, \ldots\right). \tag{4.11}$$

З урахуванням властивості точки локального мінімуму (4.1) приходимо до ключового співвідношення

$$\Delta y^* = \left(\frac{\partial\, y}{\partial\, \mathbf{x}} \right)^{*} \Delta\mathbf{x} + R^*\left(\Delta\mathbf{x}^2, \Delta\mathbf{x}^3, \ldots\right) \geq 0\ . \tag{4.12}$$

Припустимо, що змінюється одна змінна x_r, а інші не змінюються: $x_r = var;\ x_i = const\ \ (i=\overline{1,n};\ i \neq r)$. Тоді

$$\Delta y^* = \left(\frac{\partial\, y}{\partial\, x_r} \right)^{*} \Delta x_r + R^*\left(\Delta x_r^2, \Delta x_r^3, \ldots\right) \geq 0\ . \tag{4.13}$$

З (4.13) випливає, що в точці локального мінімуму похідна

$$\left(\frac{\partial y}{\partial\, x_r} \right)^{*} = 0\,. \tag{4.14}$$

Доведемо це твердження від супротивного. Нехай $\left(\frac{\partial y}{\partial\, x_r} \right)^{*} \neq 0$. Тоді межа

$$\lim_{\Delta x_r \to 0} \frac{R^*\left(\Delta\, x_r^2, \Delta\, x_r^3, \ldots\right)}{\left(\frac{\partial y}{\partial x_r} \right)^{*} \Delta\, x_r} = 0\ ,$$

і залишком $R^*\left(\Delta\, x_r^2, \Delta\, x_r^3, \ldots\right)$ у (4.13) можна зневажати:

Обираючи приріст Δx_r, протилежним за знаком похідної $\left(\frac{\partial y}{\partial x_r}\right)^*$, одержуємо

$$\Delta y^* = \left(\frac{\partial y}{\partial x_r}\right)^* \Delta x_r < 0 \,, \tag{4.15}$$

що суперечить виразу (4.13), а отже, і передумові, що визначає точку $\mathbf{x}^*$ як точку локального мінімуму. Таким чином, справедливість твердження (4.14) доведена.

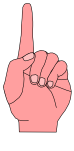
Оскільки змінна з вектора $\mathbf{x}$ вибирається довільно, то твердження (4.14) в однаковій мірі стосується кожної змінної , із вектора $\mathbf{x}$. З урахуванням цього зауваження необхідні умови для точки локального мінімуму набувають остаточного вигляду:

$$\boxed{\left(\frac{\partial y}{\partial \mathbf{x}}\right)^* = 0}\,, \tag{4.16}$$

або, іншими словами, в точці локального мінімуму градієнт функції дорівнює нулю.

Аналогічним способом умову (4.16) доводять як необхідну умову існування локального максимуму. Отже, рівність нулю градієнта функції є необхідною умовою будь-якого локального екстремуму функції.

З геометричної точки зору умова (4.16) означає, що гіперплощина, дотична до функції в точці оптимуму, паралельна гіперплощині визначення цієї функції. Так, для функції однієї змінної – це дотична лінія, паралельна осі x; для функції двох змінних – це площина, паралельна площині $x_1 0 x_2$.



Визначення 4.3. Точка, для якої виконується рівність $\left(\frac{\partial y}{\partial \mathbf{x}}\right)^* = 0$ називається *стаціонарною* точкою функції $y(\mathbf{x})$.

На рис.4.3 наведені різні приклади стаціонарних точок для функції однієї змінної. Стаціонарна точка не обов'язково повинна бути екстремальною. Прикладом такої точки може служити точка перетину функції на рис.4.3,*б*.

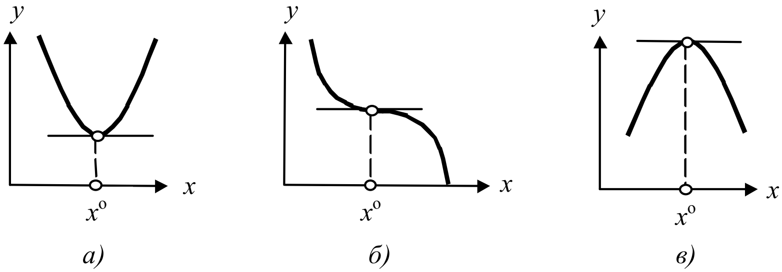


Рис.4.3 – Приклади стаціонарних точок функції однієї змінної

На рис.4.4 наведена функція двох змінних

$$y(\mathbf{x}) = 2x_1^3 + x_2^3 - 3x_1^2 x_2 + 3x_1 x_2^2 - 15x_1 \ . \tag{4.17}$$

Функція (4.17) зображена у вигляді еквіпотенційних кривих на координатній площині $x_1 0 x_2$. Кожна еквіпотенційна крива функції $y(\mathbf{x}) = y_\text{i}$ являє собою проекцію на площину $x_1 0 x_2$ лінії перетину поверхні функції $y(\mathbf{x})$ із горизонтальною площиною, що відстоїть від площини на відстані y_i . Таким чином, усі точки еквіпотенційної кривої відповідають тому самому значенню y_i.

Як видно з рисунка, функція (4.17) має чотири стаціонарні точки, в кожній з яких її частинні похідні рівні нулю, тобто виконується необхідна умова локального екстремуму (4.16). При цьому одна стаціонарна точка, розташована в першому квадранті, відповідає точці локального мінімуму функції. Інша, симетрична до першої і розташована в третьому квадранті, відповідає точці локального максимуму функції. Дві стаціонарні точки, що залишилися, розташовані в другому і четвертому квадрантах і симетричні між собою, являють собою типові *приклади* сідлових точок, у яких по одній змінній функція $y(\mathbf{x})$ досягає максимуму, а по другій – мінімуму.

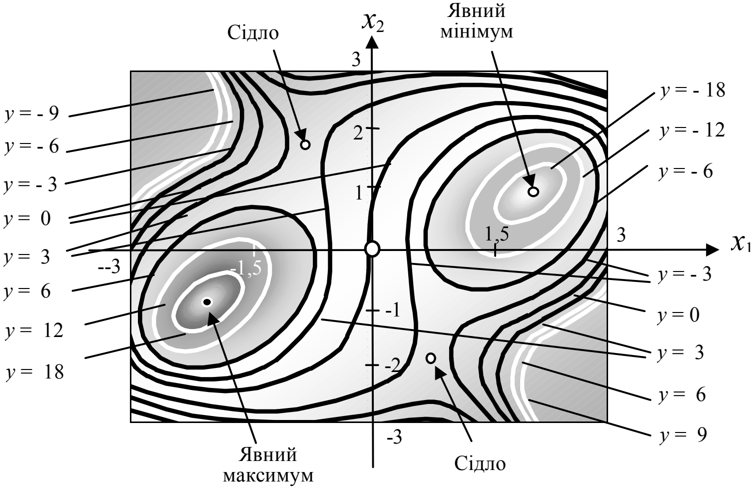


Рис. 4.4 – Еквіпотенціальні криві і стаціонарні точки функції двох змінних (4.17)

Як бачимо, виконання умови (4.16) у деякій точці $\mathbf{x}^{\circ}$ не гарантує наявності екстремуму функції в цій точці.

4.3. Достатні умови локального екстремуму

Рівність нулю перших частинних похідних (градієнта функції) – це необхідна, але недостатня умова для існування локального екстремуму, а тим більше – для локального мінімуму або максимуму. Визначимо достатні умови для точки локального мінімуму.

Екстремальна точка, як було доведено раніше, завжди є стаціонарною. Тому для доведення достатніх умов точки локального екстремуму будемо оперувати з функцією в околі її стаціонарної точки.

Розкладемо функцію $y(\mathbf{x})$ в околі стаціонарної точки $\mathbf{x}^{o}$ в ряд Тейлора:

$$y(\mathbf{x}^\circ+\Delta\mathbf{x})=y(\mathbf{x}^\circ)+\left(\frac{\partial y}{\partial x_1}\right)^\circ\Delta x_1+\left(\frac{\partial y}{\partial x_2}\right)^\circ\Delta x_2+\ldots+\left(\frac{\partial y}{\partial x_n}\right)^\circ\Delta x_n+$$

$$+\frac{1}{2!}\left[\left(\frac{\partial^2 y}{\partial x_1^2}\right)^\circ\Delta x_1^2+\left(\frac{\partial^2 y}{\partial x_2^2}\right)^\circ\Delta x_2^2+\ldots+\left(\frac{\partial^2 y}{\partial x_n^2}\right)^\circ\Delta x_n^2+\right.$$

$$\left.+\left(\frac{\partial^2 y}{\partial x_1\partial x_2}\right)^\circ\Delta x_1\Delta x_2+\ldots+\left(\frac{\partial^2 y}{\partial x_{n-1}\partial x_n}\right)^\circ\Delta x_{n-1}\Delta x_n\right]+$$

$$+R^\circ\left(\Delta\mathbf{x}^3,\Delta\mathbf{x}^4,\ldots\right),\tag{4.18}$$

де R – частина ряду розкладання, що залежить від приростів вектора змінних третього і більшого порядку; індекс «$\circ$» при похідних є ознакою їхнього обчислення в стаціонарній точці.

Перепишемо (4.18) у більш компактному матричному вигляді:

$$y(\mathbf{x}^\circ+\Delta\mathbf{x})=y(\mathbf{x}^\circ)+\left(\frac{\partial y}{\partial\mathbf{x}}\right)^{\circ\text{т}}\Delta\mathbf{x}+\frac{1}{2}\Delta\mathbf{x}^\text{т}\mathbf{H}^\circ\Delta\mathbf{x}+R^\circ(\Delta\mathbf{x}^3,\Delta\mathbf{x}^4,\ldots),\tag{4.19}$$

де $\mathbf{H}$ – симетрична матриця, що називають матрицею Гесса (гессіаном) і складається з других частинних і змішаних похідних функції цілі

$$\mathbf{H}=\begin{bmatrix}\dfrac{\partial^2 y}{\partial x_1^2} & \dfrac{\partial^2 y}{\partial x_1\partial x_2} & \ldots & \dfrac{\partial^2 y}{\partial x_1\partial x_n}\\ \dfrac{\partial^2 y}{\partial x_2\partial x_1} & \dfrac{\partial^2 y}{\partial x_2^2} & \ldots & \dfrac{\partial^2 y}{\partial x_2\partial x_n}\\ \ldots & \ldots & \ldots & \ldots\\ \dfrac{\partial^2 y}{\partial x_n\partial x_1} & \dfrac{\partial^2 y}{\partial x_n\partial x_2} & \ldots & \dfrac{\partial^2 y}{\partial x_n^2}\end{bmatrix};$$

$\mathbf{H}^{\circ}$ – матриця Гесса з елементами, обчисленими в стаціонарній точці;

$\frac{1}{2}\Delta\mathbf{x}^{\mathrm{T}}\mathbf{H}^{\circ}\Delta\mathbf{x}$ – диференціальна квадратична форма в стаціонарній точці.

Перекидаючи $y(\mathbf{x}^{\circ})$ в (4.19) з лівої частини в праву і з урахуванням того, що в стаціонарній точці градієнт функції дорівнює нулю, розкладання (4.19) набуває вигляду:

$$y\left(\mathbf{x}^{\circ}+\Delta\mathbf{x}\right)-y\left(\mathbf{x}^{\circ}\right)=\frac{1}{2}\Delta\mathbf{x}\ \ \mathbf{H}^{\circ}\Delta\mathbf{x}+R\left(\Delta\mathbf{x}^{3},\Delta\mathbf{x}^{4},...\right),$$

або

$$\Delta y^{\circ}=\frac{1}{2}\Delta\mathbf{x}^{\mathrm{T}}\mathbf{H}^{\circ}\Delta\mathbf{x}+R\left(\Delta\mathbf{x}^{3},\Delta\mathbf{x}^{4},...\right). \tag{4.20}$$

Якщо вважати

$$\Delta\mathbf{x}^{\mathrm{T}}\mathbf{H}^{\circ}\Delta\mathbf{x}\neq 0\ , \tag{4.21}$$

то частиною $R\left(\Delta\mathbf{x}^{3},\Delta\mathbf{x}^{4},...\right)$ в розкладанні (4.20) можна зневажати, оскільки

$$\lim_{\Delta\mathbf{x}\to 0}\frac{R^{*}(\Delta\mathbf{x}^{3},\Delta\mathbf{x}^{4},..)}{\frac{1}{2}\Delta\mathbf{x}^{\mathrm{T}}\mathbf{H}^{*}\Delta\mathbf{x}}=0\ .$$

Тоді приріст функції у стаціонарній точці визначається спрощеним виразом

$$\Delta y^{\circ}=\frac{1}{2}\Delta\mathbf{x}^{\mathrm{T}}\mathbf{H}^{\circ}\Delta\mathbf{x}\ . \tag{4.22}$$

Співвідношення (4.22) є ключовим для визначення достатніх умов існування локальних екстремумів.

Якщо стаціонарна точка $\mathbf{x}^{\mathrm{o}}$ є точкою локального мінімуму, то відповідно до (4.1) приріст функції в цій точці не може бути від'ємним, а відповідно до припущення (4.21) він і не дорівнює нулю, тобто

$$\Delta y^{\circ} = \Delta y^{*} = \frac{1}{2}\Delta \mathbf{x}^{\mathrm{T}}\mathbf{H}^{*}\Delta \mathbf{x} > 0 \ . \tag{4.23}$$

Аналогічно одержуємо умову наявності локального максимуму в стаціонарній точці $\mathbf{x}^{\mathrm{o}}$:

$$\Delta y^{\circ} = \Delta y^{*} = \frac{1}{2}\Delta \mathbf{x}^{\mathrm{T}}\mathbf{H}^{*}\Delta \mathbf{x} < 0 \ . \tag{4.24}$$

Знак квадратичних форм у співвідношеннях (4.23) і (4.24) цілком визначається характером гессіана $\mathbf{H}^{*}$.

Повторимо визначення 1.8 і 1.9 щодо матриці Гесса і додамо відсутні визначення для охоплення всіх можливих характерів квадратичних форм.

Визначення 4.4. Якщо квадратична форма $\Delta \mathbf{x}^{\mathrm{T}}\mathbf{H}\Delta \mathbf{x}$ додатна для всіх $\Delta \mathbf{x} \neq 0$, то вона (і матриця $\mathbf{H}$ відповідно) називається *додатно визначеною*.

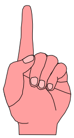
Визначення 4.5. Якщо квадратична форма $\Delta \mathbf{x}^{\mathrm{T}}\mathbf{H}\Delta \mathbf{x}$ від'ємна для всіх $\Delta \mathbf{x} \neq 0$, то вона (і матриця $\mathbf{H}$ відповідно) називається *від'ємно визначеною*.

Визначення 4.6. Якщо квадратична форма $\Delta \mathbf{x}^{\mathrm{T}}\mathbf{H}\Delta \mathbf{x}$ при різних приростах $\Delta \mathbf{x}$ може приймати будь-які значення (додатні, від'ємні або нульові), то вона (і матриця $\mathbf{H}$ відповідно) називається *невизначеною*.

Таким чином, достатня умова для точки локального мінімуму полягає (при виконанні необхідних) у додатній визначеності матриці Гесса, обчисленій в цій точці.

Аналогічно, достатня умова для точки локального максимуму (при виконанні необхідних) полягає у від’ємній визначеності матриці Гесса в цій точці.

Нарешті, функція не має екстремуму в стаціонарній точці, якщо матриця Гесса в цій точці є невизначеною. Це характерно для всіх сідлових точок функції.



Визначення 4.7. Якщо квадратична форми $\Delta\mathbf{x}^{\mathrm{T}}\mathbf{H}\Delta\mathbf{x}$ невід’ємна (недодатна) для всіх $\Delta\mathbf{x} \neq 0$, то вона (і матриця $\mathbf{H}$ відповідно) називається *напіввизначеною*.

Напіввизначеність матриці Гесса має місце в стаціонарних точках функції, що утворюють *долини* або *гребені* функції.

У загальному випадку при напіввизначеності гессіана не можна дати гарантовану відповідь про характер стаціонарної точки. Відповідь досить проста, коли є тільки одна незалежна змінна. Якщо друга похідна $\dfrac{\partial^2 y}{\partial x^2}$ (для одномірного гессіана) дорівнює нулю, то третя похідна також дорівнює нулю в точці $\mathbf{x}^{\mathrm{o}}$ у випадку її приналежності оптимуму. При цьому якщо четверта похідна додатна, $\mathbf{x}^{\mathrm{o}}$ - мінімум; якщо від’ємна, $\mathbf{x}^{\mathrm{o}}$ - максимум; якщо дорівнює нулю, питання залишається відкритим і необхідно аналізувати п'яту похідну і т.д. Маклорен у 1742 р. довів, що $\mathbf{x}^{\mathrm{o}}$ - мінімум (максимум), якщо відмінні від нуля похідні найнижчого порядку мають парний степінь.

У 1797 р. Лагранж поширив аналогічний принцип на функції багатьох змінних, а саме: якщо диференціали, що не приймають нульового значення, найбільш низького порядку додатно (від’ємно) визначені, то $\mathbf{x}^{\mathrm{o}}$ - мінімум (максимум). Це твердження було спростовано майже через сторіччя. Це було зроблено за допомогою елементарних прикладів, у яких функції мали криві долини (гребені). Помилковий принцип Лагранжа, заснований на зневажанні членами більш високого порядку, виявився спростованим, а на зміну йому прийшла строга і більш складна теорія Шифера і Штольца.

Надалі вважатимемо, що точка $\mathbf{x}^*$ задовольняє необхідним і достатнім умовам локального мінімуму, якщо в цій точці справедливе співвідношення (4.16) і має місце додатна визначеність матриці других похідних функції $y(\mathbf{x})$, обчислених у цій точці.

Існує великий клас неперервних функцій, для яких необхідні умови локального екстремуму одночасно є і достатніми. Це клас *строго опуклих* і *строго увігнутих* функцій.

Строго опукла функція має таку властивість: якщо вона в якійсь точці має другі похідні, то її матриця Гесса додатно визначена.

Аналогічну властивість має строго увігнута функція, а саме: якщо функція в якійсь точці має другі похідні, то її матриця Гесса від'ємна.

Строго опукла функція як стаціонарна точка може мати тільки точку мінімуму і при цьому єдину.

Строго увігнута функція як стаціонарна точка може мати тільки точку максимуму і при цьому єдину.

Отже, якщо функція $y(\mathbf{x})$ є строго опукла (увігнута) і двічі диференціюється, то виконання необхідних умов автоматично спричиняє і виконання достатніх, тобто $\mathbf{x}^o = \mathbf{x}^*$. Більше того, $\mathbf{x}^o = \mathbf{x}^{**}$, тобто стаціонарна крапка функції є не тільки локальним екстремумом, а й абсолютним (глобальним).

4.4. Чисельні методи визначення характеру квадратичної форми

Основою визначення виду стаціонарної точки є встановлення характеру диференціальної квадратичної форми (4.22), або, що те саме, встановлення характеру матриці (гессіана). З математичним аспектом цього питання ми вже частково зустрічалися раніше (див. підрозділ 1.8). Оскільки визначення характеру квадратичної форми має виняткове значення в теорії оптимізації, розглянемо постановку цієї задачі і методи її вирішення більш докладно.

Нехай дана квадратична форма

$$X = \mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x} = \sum_{i=1}^{n}\sum_{j=i}^{n} x_i a_{ij} x_j \quad , \qquad (4.25)$$

де $\mathbf{x}$ – n-компонентний вектор змінних; $\mathbf{A}$ – неособлива симетрична матриця коефіцієнтів розмірності $n \times n$, діагональні елементи якої a_{ii} рівні відповідним коефіцієнтам a_{ii} квадратичної форми, записаної в алгебраїчній формі, а недіагональні визначаються як $\mathrm{a}_{ij} = \frac{1}{2} a_{ij}$.

> Визначити характер квадратичної форми – це встановити знак усього виразу (4.25) залежно від будь-яких можливих значень змінних x_i и x_j , $i, j = \overline{1,n}$.

Існують такі лінійні перетворення координат (змінних x_i, $i,= \overline{1,n}$), що дозволяють виразити квадратичну форму (4.25) через нові координати (змінні v_i, $i,= \overline{1,n}$) у вигляді суми повних квадратів:

$$X = \mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x} = \sum_{i=1}^{n}\sum_{j=i}^{n} x_i a_{ij} x_j = \sum_{i=1}^{n} b_{ii} v_i^2 = \mathbf{v}^{\mathrm{T}}\mathbf{B}\mathbf{v} = V \quad , \qquad (4.26)$$

де X – вихідна квадратична форма; V – перетворена квадратична форма; x_i – початкові змінні, $i = \overline{1,n}$; v_i – перетворені змінні, лінійно зв'язані з початковими $v_i = \sum_{j=i}^{n} \text{€}_{ij} x_j$, $i = \overline{1,n}$; €_{ij} – коефіцієнти лінійних перетворень, $i = \overline{1,n}$; $\mathbf{B}$ – діагональна матриця з діагональними елементами b_{ii} .

Важливою особливістю перетворень (4.26) є еквівалентність характеру квадратичної форми

$$V = \mathbf{v}^{\mathrm{T}}\mathbf{B}\mathbf{v} = \sum_{i=1}^{n} b_{ii} v_i^2 \qquad (4.27)$$

характерові початкової квадратичної форми X. Тому характер останньої визначають за допомогою перетвореної квадратичної форми (4.27).

Очевидно, якщо всі коефіцієнти квадратичної форми (4.27) додатні (від'ємні), то квадратична форма додатно (від'ємно) визначена. Якщо ж частина цих коефіцієнтів додатна, а частина від'ємна, квадратична форма є невизначеною. Якщо всі коефіцієнти недодатні (невід'ємні), тобто є коефіцієнти b_{ii}, рівні нулю, квадратична форма є додатно (від'ємно) напіввизначеною.

Отже, пробивши над квадратичною формою (4.25) лінійні перетворення (4.26), можна легко визначити характер стаціонарної точки. Ці перетворення називаються перетвореннями Лагранжа, оскільки вони еквівалентні процедурі доповнення до повного квадрата, запропонованої Лагранжем у 1759 р. Покажемо перетворення Лагранжа на конкретному прикладі.

Приклад 4.1. Перетворити квадратичну форму

$$X = \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} \begin{bmatrix} 2 & 4 & -2 \\ 4 & 11 & 2 \\ -2 & 2 & 19 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} =$$

$$= 2x_1^2 + 11x_2^2 + 19x_3^2 + 8x_1x_2 - 4x_1x_3 + 4x_2x_3 \qquad (4.28)$$

у квадратичну форму вигляду (4.27).

Розв'язання. Перетворення Лагранжа здійснюють за допомогою формули розкладання квадрата суми

$$\left(z_1 + z_2 + \ldots + z_n\right)^2 = z_1^2 + z_2^2 + z_3^2 + 2z_1z_2 + 2z_1z_3 + \ldots + 2z_{n-1}z_n \ . \qquad (4.29)$$

Процес перетворення починають з угруповання всіх членів початкової квадратичної форми, що містять змінну x_1. Потім виносять за дужку коефіцієнт при цій змінній і доповнюють вираз у дужках до повного квадрата за формулою (4.29). У даному випадку її використовують для трьох доданків (за числом змінних):

$$2x_1^2 + 11x_2^2 + 19x_3^2 + 8x_1x_2 - 4x_1x_3 + 4x_2x_3 =$$

$$= 2(x_1^2 + 4x_1x_2 - 2x_1x_3) + 11x_2^2 + 19x_3^2 + 4x_2x_3 =$$

$$= 2(x_1^2 + 4x_1x_2 - 2x_1x_3 + 4x_2^2 + x_3^2 - 4x_2x_3) +$$

$$- 8x_2^2 - 2x_3^2 + 8x_2x_3 + 11x_2^2 + 19x_3^2 + 4x_2x_3 =$$

$$= 2(x_1 + 2x_2 - x_3)^2 + 3x_2^2 + 17x_3^2 + 12x_2x_3 .$$

На наступному кроці проводяться аналогічні процедури для змінної x_2 у частині квадратичної форми, що залишилася. Змінна x_1 в ній уже відсутня. При цьому формулу (4.29) використовують для випадку двох доданків (на одиницю менше числа змінних), і т.д. У загальному випадку перетворення Лапласа складається із (n-1)-го кроку.

Наступна послідовність алгебраїчних рівностей відображає завершення перетворень Лапласа в умовах прикладу:

$$2(x_1 + 2x_2 - x_3)^2 + 3x_2^2 + 17x_3^2 + 12x_2x_3 =$$

$$= 2(x_1 + 2x_2 - x_3)^2 + 3(x_2^2 + 4x_2x_3) + 17x_3^2 =$$

$$= 2(x_1 + 2x_2 - x_3)^2 + 3(x_2^2 + 4x_2x_3 + 4x_3^2) - 12x_3^2 + 17x_3^2 =$$

$$= 2(x_1 + 2x_2 - x_3)^2 + 3(x_2 + 2x_3)^2 + 5x_3^2 . \qquad (4.30)$$

Переходячи до нових змінних

$$\begin{aligned} v_1 &= x_1 + 2x_2 - x_3; \\ v_2 &= \phantom{x_1 + {}} x_2 + 2x_3; \\ v_3 &= \phantom{x_1 + 2x_2 + {}} x_3 , \end{aligned} \qquad 4.31)$$

одержимо

$$V = \sum_{i=1}^{3} b_{ii} v_i^2 = 2v_1^2 + 3v_2^2 + 5v_3^2 =$$

$$= \begin{bmatrix} v_1 & v_2 & v_3 \end{bmatrix} \begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 5 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \mathbf{v}^{\mathrm{T}} \mathbf{B} \mathbf{v} \, . \tag{4.32}$$

Оскільки всі коефіцієнти b_{ii} в перетвореній квадратичній формі (4.32) додатні, то початкова квадратична форма (4.28) також додатно визначена.

У процесі перетворення одержуємо трикутну матрицю лінійних перетворень **P**, що дозволяє здійснювати перехід до нових змінних від старих: $\mathbf{v} = \mathbf{P}\mathbf{x}$. В умовах прикладу $\mathbf{P} = \begin{bmatrix} 1 & 2 & -1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}$. Матриця являє собою побічний результат, що при оцінці характеру квадратичної форми не бере участі.

Перетворення Лагранжа може бути реалізоване за допомогою стандартної процедури гауссових виключень, широко використовуваних для вирішення систем лінійних рівнянь.

Алгоритм Гаусса являє собою ітераційну процедуру. Розглянемо послідовність ітерацій для довільної матриці **A** розмірності $n \times n$.

На першій ітерації всі елементи першого рядка a_{1j} матриці **A** діляться на її діагональний елемент a_{11}, який називають *головним* на даній ітерації, якщо він не дорівнює нулю. У противному разі робиться така перестановка стовпців і рядків у результаті перенумерації змінних, щоб новий діагональний елемент першого рядка був відмінний від нуля. Всі елементи інших рядків, починаючи з 2-го, перетворюють таким чином, щоб елементи першого стовпця (крім 1-го) стали рівними нулю.

На другій ітерації процедура повторюється щодо другого діагонального елемента a_{22}. Всі елементи другого рядка діляться на головний a_{22}, якщо він не дорівнює нулю. Всі елементи інших рядків, починаючи з 3-го, перетворюють так, щоб елементи другого стовпця (крім 1-го і 2-го) стали рівними нулю.

На третій ітерації процедура повторюється щодо діагонального елемента a_{33} і т.д.

Після кожного k–го кроку одержуємо проміжну матрицю перетворень $\mathbf{A}^{(k\text{-}1)}$, елементи якої зв'язані з елементами попередньої матриці $\mathbf{A}^{(k)}$ таким рекурентним співвідношенням:

$$a_{ij}^{(k)} = a_{ij}^{(k-1)} \quad \left(i = \overline{1, k-1}, \quad j = \overline{1, n}\right); \tag{4.33}$$

$$a_{ij}^{(k)} = 0 \qquad \left(i = \overline{k+1, n}, \quad j = \overline{1, k}\right); \tag{4.34}$$

$$a_{kj}^{(k)} = \frac{a_{kj}^{(k-1)}}{a_{kk}^{(k-1)}} \qquad \left(j = \overline{1, n}\right); \tag{4.36}$$

$$a_{ij}^{(k)} = a_{ij}^{(k-1)} - \frac{a_{ik}^{k-1} a_{kj}^{(k-1)}}{a_{kk}^{(k-1)}} \quad \left(i, j = \overline{k+1,\ n}\right). \tag{4.37}$$

У результаті n кроків таких перетворень матриця $\mathbf{A}$ буде перетворена в трикутну матрицю з одиничною діагоналлю $\mathbf{P}$, а n головних елементів, що фігурують у процесі гауссових виключень, утворять діагональну матрицю $\mathbf{B}$.

Продемонструємо розглянуту процедуру гауссових виключень на конкретному прикладі.

Приклад 4.2. Перетворити матрицю $\mathbf{A}^{(0)} = \begin{bmatrix} 2 & 4 & -2 \\ 4 & 11 & 2 \\ -2 & 2 & 19 \end{bmatrix}$ за допомогою гауссових виключень (4.33) – (4.37).

Розв'язання Використання виключень Гаусса дає таку послідовність проміжних матриць $\mathbf{A}^{(k)}$ і діагональних елементів b_{kk}:

$$a_{11}^{(0)} = b_{11} = 2; \qquad \mathbf{A}^{(1)} = \begin{bmatrix} 1 & 2 & -1 \\ 0 & 3 & 6 \\ 0 & 6 & 17 \end{bmatrix};$$

$$a_{22}^{(1)} = b_{22} = 3; \qquad \mathbf{A}^{(2)} = \begin{bmatrix} 1 & 2 & -1 \\ 0 & 1 & 2 \\ 0 & 0 & 5 \end{bmatrix}$$

$$a_{33}^{(2)} = b_{33} = 5; \qquad \mathbf{A}^{(2)} = \begin{bmatrix} 1 & 2 & -1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix} = \mathbf{P}\ .$$

Трійка головних елементів, отриманих у процесі перетворення, утворить діагональну матрицю $\mathbf{B} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 5 \end{bmatrix}$. Оскільки всі діагональні елементи є додатними, то матриця $\mathbf{A}$ додатно визначена.

Початкова матриця **A** збігається з матрицею коефіцієнтів квадратичної форми (4.28). Тому результати перетворень у прикладах 4.1 і 4.2 цілком збігаються.

Одержання матриці **B** за допомогою гауссових виключень легко реалізується на ЕОМ.

В інженерній практиці найбільш часто характер квадратичної форми визначають за допомогою *критерію Сільвестра*.

Введемо поняття *головних визначників матриці.*

Головні визначники матриць – це визначники матриць, що утворюються уздовж головної діагоналі (із верхнього лівого кута в правий нижній).

Для матриці $\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}$ головними визначниками

є:

$$\Delta_1 = \det\left[a_{11}\right]\ ;\ \Delta_2 = \det\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}\ ;\ \dots\ ;\ \Delta_n = \det\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}.$$

Відповідно до критерію Сільвестра:

матриця **A** позитивно визначена в тому і тільки в тому випадку, якщо всі головні визначники матриці додатні;

матриця **A** визначена в тому і тільки в тому випадку, якщо всі непарні головні визначники матриці від'ємні, а всі парні – додатні;

якщо $\Delta_1 \geq 0,\ \Delta_2 \geq 0,\ \ldots\ ,\ \Delta_n \geq 0$ або $\Delta_1 \leq 0,\ \ \Delta_2 \geq 0, \ldots,$ $(-1)^n \Delta_n \geq 0$ і є j, для якого $\Delta_j = 0$, то матриця нестрого визначена (локальний екстремум у стаціонарній точці може бути або не бути);

у всіх інших випадках матриця вважається невизначеною.

Наприклад, матриця коефіцієнтів квадратичної форми (4.28) позитивно визначена, тому що всі її головні визначники є додатними: $\Delta_1 = 2 > 0;\ \ \Delta_2 = 6 > 0;\ \ \Delta_3 = 374 > 0.$

4.5. Вирішення задачі безумовної оптимізації методом Ейлера

Усі методи вирішення задачі безумовної оптимізації можна розділити на два класи: *класичні* й *пошукові*.

Класичні методи, які в літературі ще називають *непрямими, точними,* дозволяють знайти оптимальну точку непрямим шляхом – через вирішення системи рівнянь, у загальному випадку нелінійної. При цьому якщо вирішення системи здійснюється ненаближеними методами, то одержують точні значення координат екстремальних точок функції.

Пошукові методи, які також називають *прямими*, *ітераційними, чисельними, наближеними* або *некласичними,* вирішують задачу безумовної оптимізації шляхом поступового поетапного наближення до точки екстремуму. Рішення одержують наближеним, але з наперед заданою точністю.

На відміну від класичних прямі методи складають відносно велику групу методів. Надалі будуть розглянуті тільки найбільш важливі для практичного використання. Тут же зупинимо увагу на класичному методі вирішення задачі безумовної оптимізації, що зветься *методом Ейлера.*

Метод Ейлера заснований на необхідних і достатніх умовах існування екстремуму. Метод дозволяє виявити всі екстремальні точки цільової функції (як локальні мінімуми, так і локальні максимуми) і таким чином дає загальне уявлення про поведінку гіперповерхні функції $y(\mathbf{x})$ в гіперпросторі $\mathbf{R}^n$.

Алгоритм методу полягає в наступному.

- Беруть частинні похідні функції $y(\mathbf{x})$ по кожній змінній x_i і відповідно до необхідних умов для точки локального екстремуму дорівнюють нулю
$$\frac{\partial y}{\partial x_i} = 0, \quad i = \overline{1, n}.$$
- Розв'язують будь-яким відомим методом одержану систему, що складається у загальному випадку з n нелінійних рівнянь. Корені системи, якщо вони існують, являють собою стаціонарні точки функції $y(\mathbf{x})$, оскільки в них усі частинні похідні рівні нулю.
- Беруть усі другі частинні й змішані похідні від функції $y(\mathbf{x})$ й обчислюють їх у кожній стаціонарній точці. За обчисленими похідними складають матрицю Гесса для кожної стаціонарної точки.
- Досліджують характер отриманих гессіанів. За характером матриці Гесса визначають вид відповідних екстремальних точок:

  додатна визначеність матриці відповідає точці локального мінімуму;

  від'ємна – точці локального максимуму;

  напіввизначеність залишає стаціонарну точку для додаткових досліджень;

  невизначеність вилучає точку із подальшого розгляду.
- Обчислюють значення функції $y(\mathbf{x})$ в кожному локальному мінімумі, якщо вирішується задача мінімізації, або в кожному локальному максимумі – якщо максимізації. Потім шляхом порівняння обчислених значень знаходять абсолютний екстремум.

Розглянемо метод Ейлера на конкретному прикладі.

Приклад 4.3. Дослідити на наявність екстремумів функцію двох змінних (4.17):

$$y(\mathbf{x}) = 2x_1^3 + x_2^3 - 3x_1^2 x_2 + 3x_1 x_2^2 - 15x_1 \qquad (4.39)$$

і визначити глобальні екстремуми (мінімум і максимум) функції за умови їхньої наявності.

Розв’язання. Взявши частинні похідні функції (4.17) по змінних x_1 і x_2 і прирівнявши згідно з (4.16) нулю, одержимо систему двох рівнянь з двома невідомими:

$$\begin{cases} \dfrac{\partial y}{\partial x_1} = 6x_1^2 + 3x_2^2 - 6x_1 x_2 - 15 = 0 & (4.40) \\ \dfrac{\partial y}{\partial x_2} = -3x_1^2 + 3x_2^2 + 6x_1 x_2 = 0. & (4.41) \end{cases}$$

Знайдемо корені цієї системи. Поділивши обидва рівняння (4.40) і (4.41) на 3 і віднімемо з першого друге, маємо

$$3x_1^2 - 4x_1 x_2 - 5 = 0 \ .$$

Визначаючи x_2 через x_1

$$x_2 = \frac{3x_1^2 - 5}{4x_1} \ , \tag{4.42}$$

підставляючи отриманий вираз в (4.41) із зауваженням, що $x_1 \neq 0$, і здійснюючи ряд спрощень, одержуємо

$$17x_1^4 - 70x_1^2 + 25 = 0 \ . \tag{4.43}$$

У результаті розв’язання біквадратного рівняння (4.43) маємо чотири корені: $x_{1\text{A}} = -0{,}6285;\ x_{1\text{B}} = 0{,}6285;\ x_{1\text{C}} = -1{,}9294;\ x_{1\text{D}} = 1{,}9294$. За допомогою співвідношення (4.42) знаходимо відповідні координати для другої змінної: $x_{2\text{A}} = 1{,}5175;\ x_{2\text{B}} = -1{,}5175;\ x_{2\text{C}} = -0{,}7992;$ $x_{2\text{D}} = 0{,}7992$.

Оскільки всі корені після їх підстановки в початкову систему рівнянь (4.40) – (4.41) перетворюють її в систему тотожностей, то сторонніх коренів немає. Отже, функція (4.39) має чотири стаціонарні точки (див. рис. 4.4):

$$\overset{\circ}{\mathbf{x}}_{\mathrm{A}} = \begin{bmatrix} -0{,}6285 \\ 1{,}5175 \end{bmatrix}; \ \overset{\circ}{\mathbf{x}}_{\mathrm{B}} = \begin{bmatrix} 0{,}6285 \\ -1{,}5175 \end{bmatrix}; \ \overset{\circ}{\mathbf{x}}_{\mathrm{C}} = \begin{bmatrix} -1{,}9294 \\ -0{,}7992 \end{bmatrix}; \ \overset{\circ}{\mathbf{x}}_{\mathrm{D}} = \begin{bmatrix} 1{,}9294 \\ 0{,}7992 \end{bmatrix}$$

.

Для визначення виду стаціонарних точок сформуємо чотири матриці Гесса. Для цього спочатку візьмемо всі другі частинні і змішані похідні функції (4.39):

$$\frac{\partial^2 y}{\partial x_1^2} = 12x_1 - 6x_2 \quad ; \tag{4.44}$$

$$\frac{\partial^2 y}{\partial x_1^2} = 6x_1 + 6x_2 \quad ; \tag{4.45}$$

$$\frac{\partial^2 y}{\partial x_1 \partial x_2} = \frac{\partial^2 y}{\partial x_2 \partial x_1} = -6x_1 + 6x_2 \ . \tag{4.46}$$

Потім обчислимо їх у кожній стаціонарній точці (див. табл. 4.1).

Таблиця 4.1

$\frac{\partial^2 y}{\partial x_i \partial x_j}, \quad i, j = \overline{1{,}2}$	Стаціонарна точка			
	$\overset{\circ}{\mathbf{x}}_{\mathrm{A}}$	$\overset{\circ}{\mathbf{x}}_{\mathrm{B}}$	$\overset{\circ}{\mathbf{x}}_{\mathrm{C}}$	$\overset{\circ}{\mathbf{x}}_{\mathrm{D}}$
$\frac{\partial^2 y}{\partial x_1^2}$	-16,647	16,647	-18,3576	18,3576
$\frac{\partial^2 y}{\partial x_2^2}$	5,334	-5,334	-16,3716	16,3716
$\frac{\partial^2 y}{\partial x_1 \partial x_2} = \frac{\partial^2 y}{\partial x_2 \partial x_1}$	12,876	-12,876	6,7812	-6,7812

Складемо матриці Гесса для кожної стаціонарної точки:

$$\mathbf{H}_{\mathrm{A}} = \begin{bmatrix} -16{,}647 & 12{,}876 \\ 12{,}876 & 5{,}334 \end{bmatrix}; \qquad \mathbf{H}_{\mathrm{B}} = \begin{bmatrix} -16{,}647 & -12{,}876 \\ -12{,}876 & 5{,}334 \end{bmatrix}$$

$$\mathbf{H}_{\mathrm{C}} = \begin{bmatrix} -18{,}3576 & 6{,}7812 \\ 6{,}7812 & -16{,}3716 \end{bmatrix}; \qquad \mathbf{H}_{\mathrm{D}} = \begin{bmatrix} 18{,}3576 & -6{,}7812 \\ -6{,}7812 & 16{,}3716 \end{bmatrix}.$$

Далі підрахуємо значення головних визначників одержаних гессіанів (див. табл. 4.2).

Таблиця 4.2

Головний визначник $\Delta_i, \quad i=\overline{1,2}$	Стаціонарна точка			
	$\overset{\circ}{\mathbf{x}}_{\mathrm{A}}$	$\overset{\circ}{\mathbf{x}}_{\mathrm{B}}$	$\overset{\circ}{\mathbf{x}}_{\mathrm{C}}$	$\overset{\circ}{\mathbf{x}}_{\mathrm{D}}$
Δ_1	-16,647	16,647	-18,3576	18,3576
Δ_2	-254,5865	-254,5865	254,5586	254,5586

З табл.3 відповідно до критерію Сільвестра випливає, що стаціонарні точки $\overset{\circ}{\mathbf{x}}_{\mathrm{A}}$ і $\overset{\circ}{\mathbf{x}}_{\mathrm{B}}$ не є екстремальними, $\overset{\circ}{\mathbf{x}}_{\mathrm{D}}$ – точка локального мінімуму, $\overset{\circ}{\mathbf{x}}_{\mathrm{C}}$ – точка локального максимуму. Оскільки функція має один явний мінімум і один явний максимум, то вони одночасно є і глобальними екстремумами.

Отже, рішеннями приклада 4.3 є дві точки: точка мінімуму $\mathbf{x}_{\mathrm{D}}^{**\mathrm{T}} = \left[1{,}9294 \quad 0{,}7992\right]$, в якій $y_{\min} = -19{,}293$; і точка максимуму $\mathbf{x}_{\mathrm{C}}^{**\mathrm{T}} = \left[-1{,}9294 \quad -0{,}7992\right]$, в якій $y_{\max} = 19{,}293$.

Метод Ейлера є найбільш ефективним методом для пошуку глобальних екстремумів. Однак його реалізація пов'язана з розв’язанням системи рівнянь. При великому числі змінних системи нелінійних рівнянь, як правило, алгебраїчно нерозв'язні. Для вирішення таких систем доводиться застосовувати обчислювальні методи, що орієнтовані на пошук тільки локальних екстремумів.

4.6. Контрольні запитання і вправи

1. Які квадратичні форми і матриці називають додатно визначеними? Від'ємно визначеними? Напіввизначеними? Невизначеними?

2. Як виконується аналіз характеру квадратичної форми за допомогою перетворень Лагранжа?

3. Дослідити квадратичну форму

$$y(\mathbf{x}) = 2x_1^2 + x_2^2 + 3x_3^2 - x_1x_2 + 2x_1x_3 - x_2x_3$$

за допомогою перетворень Лагранжа.

4. Визначити характер квадратичної форми $3x_1^2 - x_1x_2$ методом Лагранжа.

5. Викласти алгоритм приведення матриці до діагонального вигляду за допомогою гауссових виключень.

6. Як виконується аналіз характеру квадратичної форми за допомогою виключень Гаусса?

7. Дослідити квадратичну форму із вправи 3 за допомогою гауссових виключень.

8. Які визначники називають головними?

9. Як обчислюються детермінанти матриць?

10. Для чого призначений і в чому полягає критерій Сільвестра?

11. Дослідити квадратичну форму із впр.3 за допомогою критерію Сільвестра.

12. За допомогою критерію Сільвестра визначити характер квадратичної форми із впр.4.

13. За допомогою критерію Сильвестра визначити характер квадратичної форми

$$y(\mathbf{x}) = -2x_1^2 - x_2^2 - 3x_3^2 + 2x_1x_2 - x_2x_3 .$$

14. Сформулювати задачу багатовимірної оптимізації? У чому полягає задача *безумовної* оптимізації? Безумовної мінімізації? Безумовної максимизації?

15. Дати визначення локальному мінімуму функції.

16. Дати визначення локальному максимуму функції.

17. Що називають глобальним екстремумом?

18. Що називають рішенням задачі безумовної оптимізації?

19. Як зміниться екстремальне значення функції при додаванні до неї (функції) додатної константи? Як за тих самих умов зміниться оптимальне рішення?

20. Як зміниться екстремальне значення функції при множенні останньої на додатній (від'ємний) коефіцієнт? Як при зазначених умовах зміниться оптимальне рішення?

21. Як задачу мінімізації звести до задачі максимізації і навпаки?

22. Розкласти функцію

$$y(\mathbf{x}) = 2x_1^3 + x_2^2 - x_1 x_2 - x_2 + 5$$

в ряд Тейлора в околі точки $\mathbf{x}^{\mathrm{T}} = [\,1 \;\; 2\,]$.

23. Сформулювати необхідні умови для точки локального екстремуму?

24. Які точки називають стаціонарними?

25. Сформулювати достатні умови для точки локального мінімуму.

26. Яку матрицю називають гессіаном або матрицею Гесса?

27. Що собою становить диференціальна квадратична форма?

28. Яка властивість у строго опуклої функції? Строго увігнутої?

29. Які методи вирішення оптимізаційних задач називають класичними?

30. Які методи вирішення задач оптимізації називають прямими?

31. У чому принципова відмінність між класичними і прямими методами пошуку екстремумів неперервних функцій?

32. У чому полягає пошук екстремальних точок функції методом Ейлера?

33. Знайти мінімум функції

$$y(\mathbf{x}) = -x_1^3 + 6x_1x_2 + 3x_2^2 + 6x_1 - 18x_2 .$$

34. Вирішити задачу безумовної мінімізації

$$y(\mathbf{x}) = x_1^3 + x_2^3 + \frac{3}{x_1x_2} \to \min_{\mathbf{x}\in\mathbf{R}^2} .$$

35. Визначити мінімальне рішення і явний мінімум функції

$$y(\mathbf{x}) = 2x_1x_2 + \frac{1}{x_1^2} + \frac{1}{x_2^2}$$

методом Ейлера.

4.7. Фонд індивідуальних завдань

Індивідуальне завдання №14. За допомогою алгебраїчних перетворень Лапласа, використовуючи формулу (4.29), визначити характер квадратичної форми:

1. $y = 2x_1^2 + 2x_2^2 + x_1x_3 - x_2x_3$;

2. $y = -x_1^2 - 3x_2^2 + 3x_3^2 + 2x_1x_2 - x_1x_3 + 2x_2x_3$;

3. $y = -6x_1^2 - 3x_2^2 + 3x_3^2 - 2x_1x_2 - 3x_1x_3 + 5x_2x_3$;

4. $y = x_1^2 + x_2^2 - 2x_3^2 - 2x_1x_3 + x_2x_3$;

5. $y = x_1^2 - 2x_2^2 - 3x_3^2 - 2x_2x_3$;

6. $y = -5x_1^2 - 3x_2^2 - 2x_3^2 + 10x_1x_2 + x_1x_3 - 5x_2x_3$;

7. $y = x_1^2 + x_2^2 + 3x_3^2 + 2x_1x_2 + 7x_1x_3 + 2x_2x_3$;

8. $y = -5x_1^2 - 3x_2^2 + 3x_3^2 + 2x_1x_2 - x_1x_3 - 3x_2x_3$;

9. $y = -x_1^2 + 2x_3^2 + 4x_3^2 + 2x_1x_2 - 4x_2x_3$;

10. $y = 4x_1^2 + x_2^2 - 3x_3^2 - 4x_1x_2 + 6x_1x_3$;

11. $y = x_1^2 - 3x_2^2 + 3x_3^2 + 8x_1x_3 + 5x_2x_3$;

12. $y = x_1^2 - 3x_2^2 + 3x_3^2 + 2x_1x_2 + 8x_1x_3 + 5x_2x_3$;

13. $y = -2x_1^2 - 3x_2^2 - 2x_3^2 + x_1x_2 - 6x_1x_3 + 3x_2x_3$;

14. $y = 5x_1^2 - x_2^2 + 4x_3^2 + 2x_1x_2 + 3x_1x_3 + 2x_2x_3$;

15. $y = 2x_1^2 - 3x_2^2 - 3x_3^2 + 2x_1x_2 - 2x_1x_3 + 5x_2x_3$;

16. $y = x_1^2 - 3x_2^2 + 3x_3^2 - 2x_1x_2 - x_1x_3 + 3x_2x_3$;

17. $y = 2x_1^2 - 2x_2^2 + 3x_3^2 + 2x_1x_2 + 5x_1x_3 + 6x_2x_3$;

18. $y = 3x_1^2 - 3x_2^2 + 3x_3^2 + 3x_1x_2 - 8x_1x_3 + 4x_2x_3$;

19. $y = x_1^2 + 3x_2^2 + 3x_3^2 + 4x_1x_2 + 2x_1x_3 + 7x_2x_3$;

20. $y = 4x_1^2 - 2x_2^2 + 5x_3^2 + 2x_1x_2 + 3x_1x_3 + 6x_2x_3$;

21. $y = x_1^2 + x_2^2 + 3x_3^2 + 2x_1x_2 + 4x_1x_3 + 2x_2x_3$;

22. $y = x_1^2 - 3x_2^2 + 3x_3^2 + 2x_1x_2 + 5x_2x_3$;

23. $y = x_1^2 - 3x_2^2 + 3x_3^2 + 2x_1x_2 + 6x_1x_3$;

24. $y = x_1^2 - 3x_2^2 + 5x_3^2 + 2x_1x_2 + 2x_1x_3 + 5x_2x_3$;

25. $y = -x_1^2 - 3x_2^2 - 3x_3^2 + 2x_1x_2 + 2x_1x_3 + 5x_2x_3$;

26. $y = x_1^2 - 3x_2^2 + 3x_3^2 + 2x_1x_2 + 8x_1x_3 + 5x_2x_3$;

27. $y = -4x_1^2 - 3x_2^2 - 3x_3^2 + 5x_1x_2 + 2x_1x_3 + 5x_2x_3$;

28. $y = x_1^2 - 3x_2^2 + 3x_3^2 + 2x_1x_2 - 5x_2x_3$;

29. $y = 3x_1^2 + x_2^2 + 3x_3^2 + 2x_1x_2 + x_1x_3 + 4x_2x_3$;

30. $y = 2x_1^2 - 4x_2^2 + 6x_3^2 + 2x_1x_2 + 2x_1x_3 + 4x_2x_3$.

Індивідуальне завдання №15. Визначити характер квадратичної форми, поданої в індивідуальному завданні №17 , за допомогою гауссових виключень.

Індивідуальне завдання №16. Визначити характер квадратичної форми, поданої в індивідуальному завданні №17, за допомогою критерію Сільвестра.

Індивідуальне завдання №17. Вирішити задачу безумовної мінімізації

$$y(\mathbf{x}) \to \min_{\mathbf{x}\in\mathbf{R}^n}$$

класичним методом (методом Ейлера), якщо функція надана таким чином:

1. $y = 2x_1^2 - 2x_2^2 - x_1x_2 - x_1x_3 - x_2x_3 - x_1 - x_3 - 2$;

2. $y = x_1^2 - x_2^2 + 2x_3^2 - 2x_1x_2 - 2x_1x_3 - 5x_2 - x_3$;

3. $y = 3x_1^2 - x_3^2 + 4x_1x_2 - 2x_1x_3 - 2x_2x_3 - x_1 - 3x_2 - 2x_3$;

4. $y = 2x_1^2 + x_2^2 + 5x_1x_2 - 2x_1x_3 - 3x_2x_3 - x_1 - 3x_2$;

5. $y = 3x_1^2 - x_3^2 + x_1x_2 - 2x_1x_3 - 4x_2x_3 - 3x_2 - 2x_3$;

6. $y = 3x_1x_2 - 3x_1x_3 + x_2x_3 - x_1 - 2x_2 - x_3$;

7. $y = x_1^2 - x_2^2 - x_3^2 - 2x_1x_2 - 2x_1x_3 - 2x_2x_3 - x_1 - x_2 - x_3$;

8. $y = 2x_1^2 - 2x_2^2 - 2x_3^2 + x_1x_3 - x_2x_3 - 5$;

9. $y = -x_1^2 - 2x_2^2 - 3x_3^2 - 3x_1x_2 + 2x_1x_3 - x_2 - x_3 + 3$.

10. $y = x_1^2 - x_2^2 - x_3^2 - 2x_1x_2 - 2x_1x_3 - 2x_2x_3 - x_1 - x_2 - x_3$;

11. $y = x_1^2 + x_2^2 - x_3^2 - 2x_1x_2 - 3x_1x_3 - 2x_2x_3 - x_1 - 4_2 - x_3$;

12. $y = 4x_1^2 - x_2^2 + x_3^2 - 2x_1x_2 - 2x_1x_3 - 2x_2x_3 - x_2 - x_3$;

13. $y = x_1^2 - x_2^2 - 3x_1x_2 - 2x_1x_3 - 7x_2x_3 - x_3$;

14. $y = 2x_1^2 + 2x_2^2 + x_1 + 1$;

15. $y = -2x_1^2 + x_1x_2 + x_2$;

16. $y = 2x_1^2 - 2x_1x_2 - 2x_2 - 1$;

17. $y = x_1^2 + x_2^2 - 2x_3^2 - 2x_1x_3 + x_2$;

18. $y = x_1^2 - 2x_2^2 - 3x_3^2 - 2x_2x_3 - x_1$;

19. $y = x_2^2 - x_3^2 - 6x_1x_2 - 6x_1x_3 + 1$;

20. $y = x_1^2 - 2x_2x_3 - 8x_2$;

21. $y = -x_4^2 + 2x_1x_2 - 4x_3$;

22. $y = x_2^2 - 3x_3^2 - 4x_1x_2 + 6x_1x_3 + 2$;

23. $y = x_1^2 - 8x_2x_3 + 5x_3 + 7$;

24. $y = x_3^2 - 3x_1x_2 - 2x_1$;

25. $y = x_2^2 - 5x_3^2 - 4x_1x_3 + 7x_2 - x_3$;

26. $y = 2x_3^2 - 2x_1x_2 + x_3 - 3$;

27. $y = x_1^2 - 2x_1x_2 + 3x_1x_3 - x_2$;

28. $y = x_2^2 + x_3^2 + x_1x_2 + 2x_1x_3 + x_3$;

29. $y = x_1^2 - 10x_3^2 + 2x_2x_3 + 4x_1 - x_3 - 1$;

30. $y = 2x_1^2 - 3x_1x_2 + 2x_2x_3 - 6x_2$;

Індивідуальне завдання №18. Знайти всі екстремальні точки функції $y(\mathbf{x})$, даної в попередньому індивідуальному завданні.

Розділ 5. ПРЯМІ МЕТОДИ БЕЗУМОВНОЇ ОПТИМІЗАЦІЇ

Незважаючи на те, що класичні методи розв’язання задач безумовної оптимізації дозволяють знаходити точне рішення, причому абсолютне, вони, проте, в інженерній практиці поступово витісняються прямими методами. Одна з причин такого становища – відсутність універсального методу розв’язання систем нелінійних рівнянь. Інша, більш серйозна – складність машинної реалізації класичних методів, пов'язана з взяттям табличних похідних для укладання системи рівнянь типу (5.16).

У цьому розділі будуть розглянуті тільки ті прямі методи розв’язання задач безумовної оптимізації, що найчастіше від інших використовуються в інженерній практиці і визначають окремі напрямки методів або стратегії пошуку екстремальних точок.

Нагадаємо, що прямі методи оптимізації називають також *пошуковими*, *покроковими, ітераційними, наближеними, послідовними, некласичними, чисельними.*

5.1. Особливості прямих методів безумовної оптимізації

Розв’язання оптимізаційних задач прямими методами являє собою ітераційний процес, що припускає цілеспрямований рух по гіперповерхні функції цілі до точки локального екстремуму.

Рух до точки екстремуму починають від деякої точки $\mathbf{x}_0 \in \mathbf{R}^n$, яку називають *точкою початкового наближення*. Така точка повинна бути відома до початку процесу оптимізації. Вона може бути задана умовою задачі або вибиратися з розумних міркувань – якнайближче до передбачуваного місця знаходження екстремуму. Якщо апріорна інформація про знаходження локального екстремуму відсутня і не можна зробити розумних припущень про її місцезнаходження, то як початкову точку наближення доцільно вибрати початок системи координат. У цьому випадку спростяться обчислювальні процедури на першому кроці руху. Однак завжди слід пам'ятати, що початкова точка наближення повинна належати області існування функції, тобто $\mathbf{x}_0 \in \mathbf{V} \subseteq \mathbf{R}^n$.

На кожному наступному $(k+1)$-м кроці оптимізації за деякою (заздалегідь визначеною) стратегією обчислюється нова точка наближення $\mathbf{x}^{(k+1)}$, що розташовується ближче до шуканого екстремуму, ніж попередня. Іншими словами, при пошуку локального мінімуму значення функції в новій точці наближення повинно бути менше, ніж у попередньої

$$y\left(\mathbf{x}^{(k+1)}\right) < y\left(\mathbf{x}^{(k)}\right) , \tag{5.1}$$

а при пошуку максимуму – більше:

$$y\left(\mathbf{x}^{(k+1)}\right) > y\left(\mathbf{x}^{(k)}\right) . \tag{5.2}$$

У більшості прямих методів безумовної оптимізації наступна точка наближення $\mathbf{x}^{(k+1)}$ зв'язана рекурентним[1] співвідношенням із попередньою $\mathbf{x}^{(k)}$

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \lambda^{(k)} \Delta\mathbf{x}^{(k)} , \tag{5.3}$$

де $\Delta\mathbf{x}^{(k)}$ – направляючий вектор, що визначає напрямок руху; $\lambda^{(k)}$ – скалярна величина, що регулює розмір кроку в напрямку руху $\Delta\mathbf{x}^{(k)}$; $k = 0, 1, 2, \ldots$

Побудова послідовного ланцюга точок наближення (5.3), таких що виконується умова (5.1) при пошуку мінімуму або умова (5.2) при пошуку максимуму, називається релаксійним процесом [8], а сама послідовність – оптимізуючою (мінімізуючою або максимізуючою) .

[1] Рекурентне [від лат. rekurrents – той, що повертається] співвідношення – формула зв'язку члена математичної послідовності, починаючи з деякого, з попередніми членами цієї послідовності. Такою, наприклад, є послідовність чисел Фібоначчі 1,1, 2, 3, 5, 8, 13, … , де кожний наступний член, починаючи з третього, дорівнює сумі двох попередніх.

Релаксійний процес називається збіжним, якщо

$$\lim_{k\to\infty} \mathbf{x}^{(k)} = \mathbf{x}^* .$$

Збіжний релаксійний процес не гарантує точного відшукання екстремуму, однак дозволяє за кінцеве число кроків (ітерацій) одержати рішення, як завгодно мало відмінне від екстремального.

На перших кроках оптимізації наближення до екстремуму відбувається більш стрімко, чим на наступних. Чим ближче до точки екстремуму, тим коротше крок. Послідовність точок, обумовлена співвідношенням (5.3), і, отже, сам процес наближення до екстремуму нескінченний. Збільшувати ступінь наближення до екстремуму можна безмежно. Однак у реальних задачах це втрачає сенс. Безглуздо шукати нове рішення $\mathbf{x}^{(k+1)}$, якщо вектор приростів змінних $\left(\lambda^{(k)}\Delta\mathbf{x}^{(k)}\right)$ не приводить до зміни функції $\Delta y^{(k)}$, яку можна «відчути» в реальних умовах. Наприклад, якщо функція цілі – це деяка вартісна залежність, то немає рації домагатися її поліпшення на тисячні або мільйонні частки копійки. Однаково вона буде округлена до копійок або гривень. Також немає рації поліпшувати функцію, що є математичною моделлю деякої фізичної величини, якщо поліпшення перевершує точність приладів, за допомогою яких можна виміряти значення цієї фізичної величини. Те саме стосується і вектора змінних $\mathbf{x}$. Якщо реальні технічні умови такі, що значення складових вектора $\mathbf{x}$ встановлюються з деякою відомою точністю ε, то безглуздо намагатися поліпшити значення функції цілі зо допомогою кроків, довжина яких менше за точність ε.

Постає запитання, як довго варто наближатися до точки локального екстремуму, тобто коли доцільно припинити процес оптимізації?

Якщо продовжити початі міркування, то можна прийти до наступного висновку. Процес оптимізації повинен закінчуватися тоді, коли або збільшення функції, або збільшення змінних, або те й інше одночасно стануть менше деякого розміру ε. Оскільки приріст функції і приріст вектора змінних пов'язані з похідними, то закінчувати процес оптимізації треба в момент настання нечутливості приросту функції цілі до приростів аргументів, тобто коли вектор перших частинних похідних за модулем стане менше деякого порогу чутливості ε (точності обчислення) :

$$\left|\frac{\Delta y^{(k)}}{\Delta \mathbf{x}^{(k)}}\right| \approx \left|\left(\frac{\partial y}{\partial \mathbf{x}}\right)^{(k)}\right| < \varepsilon \ . \tag{5.4}$$

Умову (5.4) інтепретують таким чином: процес оптимізації завершують тоді, коли всі частинні похідні за модулем одночасно стануть менше точності обчислення

$$\left|\left(\frac{\partial y}{\partial x_i}\right)^{(k)}\right| < \varepsilon \ , \quad i = \overline{1,n} \ . \tag{5.5}$$

Точність обчислення в оптимізаційних задачах може бути задана умовою задачі. У протилежному разі необхідно її вибрати самостійно з розумних міркувань. При виборі варто керуватися змістовною постановкою задачі. Як правило, точність обчислення вибирається на 2-4 порядки менше одиниць виміру функції цілі.

На практиці, коли релакційний процес не припускає обчислення частинних похідних функції цілі, часто використовують інші критерії зупинки процесу оптимізації. Наприклад, використовують умову, коли модуль збільшення функції на черговому кроці оптимізації не перевищуватиме заданої величини ε_y

$$\left| y\left(\mathbf{x}^{k+1}\right) - y\left(\mathbf{x}^{k}\right) \right| \le \varepsilon_y \ , \tag{5.6}$$

або умову, коли відстань між новою точкою наближення і попередньою зменшиться до деякої величини $\varepsilon_{\mathbf{x}}$

$$\left\|\left(\mathbf{x}^{k+1}-\mathbf{x}^{k}\right)\right\| \le \varepsilon_{\mathbf{x}} \quad . \tag{5.7}$$

При виборі точності обчислення, як і при виборі точки початкового наближення, досить чітко виявляється прикладний характер теорії оптимізації, її взаємозв'язок з фізичними явищами, процесами, об'єктами та їхніми відношеннями. У випадку використання прямих методів оптимізація цей зв'язок, з одного боку, визначає обов'язкову наявність додаткових початкових даних (точка початкового наближення і точність обчислення), а з другого – дозволяє за результатами оптимізації в розумних межах поліпшувати параметри або показники фізичних об'єктів або процесів.

Рекурентне співвідношення (5.3) за інформацією, що визначає місце розташування k-й точки наближення до екстремуму, дозволяє визначити місце розташування $(k + 1)$-й точки.

Існує багато різних методів визначення направляючого вектора $\Delta\mathbf{x}^{(k)}$ і параметра довжини кроку $\lambda^{(k)}$ у виразі (5.3). Залежно від методу рух до точки екстремуму здійснюється за різними траєкторіями за різне число кроків і закінчується в різних точках. Але в будь-якому випадку досягається задана точність визначення екстремуму.

Більшість існуючих методів безумовної оптимізації засновані на обчисленні частинних похідних функції цілі. Залежно від порядку похідних, що обчислюються у процесі визначення направляючого вектора $\Delta\mathbf{x}^{(k)}$ і параметра довжини кроку $\lambda^{(k)}$, вони діляться на *методи* першого *і* другого порядків.

Використання в методах другого порядку додаткової інформації про функцію цілі у вигляді других частинних похідних дозволяє, як правило, зменшити число кроків наближення до екстремуму.

Скорочення числа кроків зовсім не означає скорочення реального часу досягнення екстремальної точки. Реальний час оптимізації визначається витратою часу на всі обчислювальні операції, які необхідно виконати в процесі руху від початкової точки наближення до кінцевої. У методах оптимізації другого порядку процедура обчислень на кожному кроці більш складна, ніж методах першого порядку. За відносне скорочення числа кроків доводиться розплачуватися складністю обчислювальних процедур.

Порівняння різних методів безумовної оптимізації за часом знаходження екстремуму являє собою складну задачу. Більше того, така задача в деякому розумінні є некоректною. Порівняння методів на різних типах ЕОМ може дати суперечливі результати через різні співвідношення часу виконання математичних операцій (додавання, множення, зведення в ступінь та ін.) і математичних процедур (обертання матриць, взяття похідних, обчислення стандартних функцій і т.п.).

Крім того, різні початкові точки наближення дають різні результати порівняння. Кожний конкретний метод має свої початкові точки, в яких використання саме цього методу більш ефективно, ніж інших. У деяких методах таких точок більше, в інших – менше.

Важливою характеристикою прямих методів оптимізації є *збіжність*. Говорять, що метод (5.3) збігається, якщо послідовність оптимізації $\mathbf{x}^k \to \mathbf{x}^*$, при $k \to \infty$, де $\mathbf{x}^*$ – вирішення задачі безумовної оптимізації.

Ефективність збіжного методу умовно характеризують *швидкістю збіжності*. Говорять, що послідовність $\mathbf{x}^k$ сходиться до точки $\mathbf{x}^*$ *лінійно* (із лінійною швидкістю, із швидкістю геометричної прогресії), якщо існують такі константи $q \in (0,1)$ і k_0, що

$$\left\| x^{k+1} - x^* \right\| \le q \left\| x^k - x^* \right\| \quad \text{при} \quad k \ge k_0 \ . \tag{5.8}$$

Говорять, що $\mathbf{x}^k$ збігається до $\mathbf{x}^*$ *надлінійно* (із надлінійною швидкістю), якщо

$$\left\| x^{k+1} - x^* \right\| \le q_{k+1} \left\| x^k - x^* \right\|, \quad q_k \to 0^+ \quad \text{при} \quad k \to \infty \ . \tag{5.9}$$

Нарешті, говорять, що послідовність $\mathbf{x}^k$ збігається до $\mathbf{x}^*$ із *квадратичною швидкістю*, якщо існують такі константи $C \geq 0$ і k_0, що

$$\left\| x^{k+1} - x^* \right\| \leq C \left\| x^k - x^* \right\|^2 \quad \text{при} \quad k \geq k_0 \,. \tag{5.10}$$

Іноді, зберігаючи ту ж термінологію, нерівності (5.8)–(5.10) замінюють відповідно на нерівності

$$\left\| x^{k+1} - x^* \right\| \leq C_1 q^{k+1} \quad \text{при} \quad k \geq k_0 \,; \tag{5.11}$$

$$\left\| x^{k+1} - x^* \right\| \leq C_2 q_{k+1} q_k \ldots q_1 \,; \tag{5.12}$$

$$\left\| x^{k_0+1} - x^* \right\| \leq C_3 q^{2^{k_0+1}} \quad \text{при} \quad k \geq k_0 \,. \quad 0 < q < 1 \,. \tag{5.13}$$

Для характеристики збіжності послідовності $y\left(\mathbf{x}^k\right)$ до y^* в ситуаціях, аналогічних (5.8), (5.11); (5.9), (5.12) і (5.10), (5.13), використовують аналогічні терміни: лінійна, надлінійна і квадратична збіжність.

Більшість теорем про збіжність методів оптимізації доводяться в припущенні опуклості цільової функції, оцінки ж швидкості збіжності часто встановлюються при ще більш обмежувальному припущенні сильної опуклості. У задачах неопуклого програмування чисельні методи оптимізації дозволяють відшукувати лише локальні рішення, а точніше кажучи, стаціонарні точки. Задачі відшукання глобального рішення в загальному випадку надзвичайно складні. Алгоритми глобальної оптимізації вимагають великих витрат обчислювальних ресурсів навіть для функцій однієї змінної. Одержання ж досить точного рішення багатовимірних задач глобальної оптимізації за допомогою існуючих у даний час чисельних методів часто виявляється взагалі неможливим [9].

Процес оптимізації будь-яким методом являє собою тривалу і стомливу процедуру навіть для функцій малого числа аргументів, тому в інженерній практиці для розв’язання оптимізаційних задач слід використовувати комп'ютерну техніку. Оскільки порівняльні характеристики методів дуже умовні, то вибір методу оптимізації для вирішення тій або іншої задачі в першу чергу залежить від наявності відповідного програмного забезпечення. Якщо програмне забезпечення включає декілька методів, то вибір залежить від типу задачі оптимізації і наявності досвіду їх використання.

У наступних підрозділах розглядаються найбільш поширені методи безумовної оптимізації першого і другого порядків.

5.2. Метод найшвидшого спуску

Одним з найбільш поширених методів безумовної оптимізації є метод, запропонований ще в 1847 році французьким математиком О.Коші. Метод ґрунтується на визначенні градієнта функції цілі. Рух по гіперповерхні функції до точки локального екстремуму на кожному кроці починається в напрямку найшвидшого зростання функції (за градієнтом) у випадку пошуку максимуму або в напрямку найшвидшого убування функції (за антиградієнтом) у випадку пошуку мінімуму. Цей метод добре відомий у математиці як метод найшвидшого спуску.

На рис. 5.1 подана графічна інтепретація стратегії руху по гіперповерхні функції до точки локального екстремуму при використанні методу найшвидшого спуску на прикладі функції двох змінних.

На рисунку функція цілі зображена у вигляді еквіпотенціалей (ліній рівних значень). Сірим фоном виділений окіл точки екстремуму. Як видно з рисунка, рух із початкової точки наближення починається за нормаллю (за градієнтом при пошуку максимуму й антиградієнтом при пошуку мінімуму) до еквіпотенціалі, що проходить через цю точку. Рух продовжується доти, поки функція зменшується в даному напрямку, тобто до еквіпотенціалі, для якої напрямок руху є дотичним. Точка такого дотику являє собою нову точку наближення. Стратегія руху з кожної нової точки наближення та сама. Оптимізація закінчується, коли ε-окіл нової точки наближення «захопить» екстремум функції, тобто коли виконується умова (5.5).

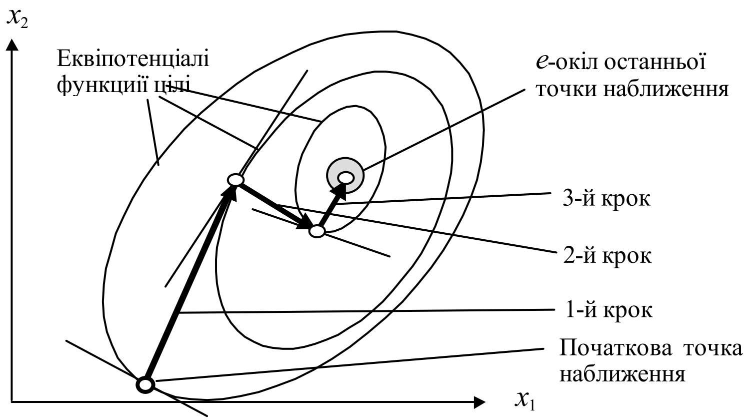


Рис.5.1– Графічна інтепретація методу найшвидшого спуску

Якщо довжину кроку в методі найшвидшого спуску спрямувати до 0, то відповідна траєкторія руху до локального мінімуму двовимірної функції буде являти собою плавну криву, що збігається зі шляхом, який вибирає вода, коли стікає по поверхні функції з точки початкового наближення.

У методі найшвидшого спуску направляючий вектор математично визначається таким чином:

при пошуку локального максимуму як градієнт функції

$$\Delta \mathbf{x}^{(k)} = \left(\frac{\partial y}{\partial \mathbf{x}} \right)^{(k)} ; \tag{5.14}$$

при пошуку локального мінімуму як антиградієнт функції

$$\Delta \mathbf{x}^{(k)} = -\left(\frac{\partial y}{\partial \mathbf{x}} \right)^{(k)} . \tag{5.15}$$

Отже, рекурентне співвідношення (5.3) для визначення нових точок наближення відповідно набуває вигляду:

при пошуку локального максимуму

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \lambda^{(k)} \left(\frac{\partial y}{\partial \mathbf{x}} \right)^{(k)} ; \tag{5.16}$$

при пошуку локального мінімуму

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} - \lambda^{(k)} \left(\frac{\partial y}{\partial \mathbf{x}} \right)^{(k)} . \tag{5.17}$$

Параметр $\lambda^{(k)}$ у виразі (5.16) визначається в результаті вирішення одновимірної задачі максимізації функції

$$y(\lambda^{(k)}) = y\left(\mathbf{x}^{(k)} + \lambda^{(\mathrm{k})} \left(\frac{\partial y}{\partial \mathbf{x}} \right)^{(k)} \right) \to \max_{\lambda^{(k)} \in R^1} , \tag{5.18}$$

а у виразі (5.17) – в результаті вирішення одновимірної задачі мінімізації функції

$$y(\lambda^{(k)}) = y\left(\mathbf{x}^{(k)} - \lambda^{(\mathrm{k})} \left(\frac{\partial y}{\partial \mathbf{x}} \right)^{(k)} \right) \to \min_{\lambda^{(k)} \in R^1} . \tag{5.19}$$

Процес закінчення оптимізації закінчується в повній відповідності з умовою (5.5).

Розглянемо метод найшвидшого спуску на конкретному прикладі.

Приклад 5.1 Вирішити двовимірну задачу безумовної максимізації функції (4.17)

$$y = 2x_1^3 + x_2^3 - 3x_1^2 x_2 + 3x_1 x_2^2 - 15x_1 .$$

Як початкове наближення взяти точку $\mathbf{x}^{(0)^{\mathrm{T}}} = [-1{,}8 \quad -1]$. Точність обчислення визначити як $\varepsilon = 0{,}01$.

Розв'язання. Початкове значення функції $y\left(\mathbf{x}^{(0)}\right)=18{,}0976$. Це значення знадобиться для контролю правильності руху на першому кроці. Обчислимо перші частинні похідні функції в початковій точці:

$$\left(\frac{\partial y}{\partial \mathbf{x}}\right)^{(0)}=\begin{bmatrix}\frac{\partial y}{\partial x_1}\\ \frac{\partial y}{\partial x_2}\end{bmatrix}^{(0)}=\begin{bmatrix}6x_1^2-6x_1x_2+3x_2^2-15\\ 3x_2^2-3x_1^2+6x_1x_2\end{bmatrix}^{(0)}=\begin{bmatrix}-3{,}36\\ 4{,}08\end{bmatrix}.$$

Частинні похідні за модулем в початковій точці наближення значно перевершують задану точність обчислення:

$$\left|\left(\frac{\partial y}{\partial x_1}\right)^{(0)}\right|=\left|-3{,}36\right|>0{,}01=\varepsilon;\quad \left|\left(\frac{\partial y}{\partial x_2}\right)^{(0)}\right|=\left|4{,}08\right|>0{,}01=\varepsilon\ .$$

Отже, відповідно до критерію (5.5) необхідно зробити, принаймні, один крок процесу максимізації.

Визначимо направляючий вектор на першому кроці як градієнт функції:

$$\Delta\mathbf{x}^{(0)}=\left(\frac{\partial y}{\partial \mathbf{x}}\right)^{(0)}=\begin{bmatrix}-3{,}36\\ 4{,}08\end{bmatrix}.$$

Для визначення параметра $\lambda^{(0)}$ сформуємо функцію $y(\lambda^{(0)})$ таким чином:

$$y(\lambda^{(0)})=y\left(\mathbf{x}^{(0)}+\lambda^{(0)}\left(\frac{\partial y}{\partial \mathbf{x}}\right)^{(0)}\right)=2\left(x_1^{(0)}+\lambda^{(0)}\left(\frac{\partial y}{\partial x_1}\right)^{(0)}\right)^3+$$

$$+\left(x_2^{(0)}+\lambda^{(0)}\left(\frac{\partial y}{\partial x_2}\right)^{(0)}\right)^3-3\left(x_1^{(0)}+\lambda^{(0)}\left(\frac{\partial y}{\partial x_1}\right)^{(0)}\right)^2\left(x_2^{(0)}+\lambda^{(0)}\left(\frac{\partial y}{\partial x_2}\right)^{(0)}\right)+$$

$$+3\left(x_1^{(0)}+\lambda^{(0)}\left(\frac{\partial y}{\partial x_1}\right)^{(0)}\right)\left(x_2^{(0)}+\lambda^{(0)}\left(\frac{\partial y}{\partial x_2}\right)^{(0)}\right)^2-15\left(x_1^{(0)}+\lambda^{(0)}\left(\frac{\partial y}{\partial x_1}\right)^{(0)}\right)=$$

$$= 2\left(-1{,}8 - 3{,}36\lambda^{(0)}\right)^3 + \left(-1 + 4{,}08\lambda^{(0)}\right)^3 -$$

$$-3\left(-1{,}8 - 3{,}36\lambda^{(0)}\right)^2\left(-1 + 4{,}08\lambda^{(0)}\right) +$$

$$+3\left(-1{,}8 - 3{,}36\lambda^{(0)}\right)\left(-1 + 4{,}08\lambda^{(0)}\right)^2 - 15\left(-1{,}8 - 3{,}36\lambda^{(0)}\right) =$$

$$= -313{,}9292\left(\lambda^{(0)}\right)^3 - 293{,}6909\left(\lambda^{(0)}\right)^2 + 27{,}936\lambda^{(0)} - 18{,}656. \quad (5.20)$$

Знайдемо мінімум функції (5.20), використовуючи метод Ейлера,

$$\left(\frac{dy}{d\lambda}\right)^{(0)} = -941{,}7876\left(\lambda^{(0)}\right)^2 - 587{,}3818\lambda^{(0)} + 27{,}936 = 0 \ . \quad (5.21)$$

Визначимо корені рівняння (5.21):

$$\lambda^{(0)} = \frac{293{,}6909 \pm \sqrt{862543{,}3447 + 26309{,}7784}}{-941{,}7876} == \begin{cases} -0{,}6681 = \lambda_1^{(0)}; \\ 0{,}0444 = \lambda_2^{(0)}. \end{cases}$$

Як бачимо, функція (5.20) має дві стаціонарні точки. Для встановлення характеру точок знайдемо чисельні значення другою похідною в цих точках:

$$\left(\frac{d^2y}{d\lambda^2}\right)^{(0)} = -1883{,}5752\lambda^{(0)} - 587{,}3818 = \begin{cases} 671{,}0348\Big|_{\lambda_1^{(0)} = -0{,}6681}; \\ -671{,}0125\Big|_{\lambda_2^{(0)} = 0{,}0444} \end{cases}$$

Перша стаціонарна точка $\lambda_1^{(0)} = -0{,}6681$ відповідає мінімуму функції (5.20), друга $\lambda_2^{(0)} = 0{,}0444$ – максимуму. Отже, як параметр довжини кроку вибираємо величину $\lambda^{(0)} = 0{,}0444$.

Нове значення вектора змінних (нова точка наближення) у результаті виконання першого кроку визначиться відповідно до рекурентного співвідношення (5.16):

$$\mathbf{x}^{(1)} = \mathbf{x}^{(0)} + \lambda^{(0)}\left(\frac{\partial y}{\partial \mathbf{x}}\right)^{(0)} = \begin{bmatrix} -1{,}8 \\ -1 \end{bmatrix} + 0{,}0444\begin{bmatrix} -3{,}36 \\ 4{,}08 \end{bmatrix} = \begin{bmatrix} -1{,}9492 \\ -0{,}8188 \end{bmatrix}.$$

Значення функції цілі в новій точці наближення $y\left(\mathbf{x}^{(1)}\right) = 19{,}2899$. Оскільки воно перевищує початкове значення $y\left(\mathbf{x}^{(0)}\right) = 18{,}0976$, то рух відбувається в напрямку максимуму функції, тобто у правильному напрямку.

На другому кроці повторюємо обчислювальну процедуру, яка мала місце на першому кроці максимізації, але вже щодо нової точки наближення $\mathbf{x}^{(1)}$:

$$\Delta\mathbf{x}^{(1)} = \left(\frac{\partial y}{\partial \mathbf{x}}\right)^{(1)} = \begin{bmatrix} 0{,}2315 \\ 0{,}1893 \end{bmatrix};$$

$$\left|\left(\frac{\partial y}{\partial x_1}\right)^{(1)}\right| = |0{,}2315| > 0{,}01 = \varepsilon, \quad \left|\left(\frac{\partial y}{\partial x_2}\right)^{(1)}\right| = |0{,}1893| > 0{,}01 = \varepsilon;$$

$$y\left(\lambda^{(1)}\right) = y\left(\mathbf{x}^{(1)} + \lambda^{(1)}\left(\frac{\partial y}{\partial \mathbf{x}}\right)^{(1)}\right) = 0{,}0261\left(\lambda^{(1)}\right)^3 - 0{,}4956\left(\lambda^{(1)}\right)^2 + 0{,}0948\lambda^{(1)} + 19{,}2671;$$

$$\left(\frac{dy}{d\lambda}\right)^{(1)} = 0{,}0783\left(\lambda^{(1)}\right)^2 - 0{,}9912\lambda^{(1)} + 0{,}0948 = 0\ ;$$

$$\lambda^{(1)} = \frac{0{,}4956 \pm \sqrt{0{,}2456 - 0{,}0074}}{0{,}0783} == \begin{cases} 12{,}5626 = \lambda_1^{(1)}, \\ 0{,}0958 = \lambda_2^{(1)}; \end{cases}$$

$$\left(\frac{d^2y}{d\lambda^2}\right)^{(1)} = 0{,}1566\lambda^{(1)} - 0{,}9912 = \begin{cases} 0{,}9761\Big|_{\lambda_1^{(1)}=12{,}5626}, \\ -0{,}9761\Big|_{\lambda_2^{(1)}=0{,}0958}; \end{cases}$$

$$\lambda^{(1)} = 0{,}0958 \; ;$$

$$\mathbf{x}^{(2)} = \mathbf{x}^{(1)} + \lambda^{(1)}\left(\frac{\partial y}{\partial \mathbf{x}}\right)^{(1)} = \begin{bmatrix} -1{,}9492 \\ -0{,}8188 \end{bmatrix} + 0{,}0958\begin{bmatrix} 0{,}2315 \\ 0{,}1893 \end{bmatrix} = \begin{bmatrix} -1{,}927 \\ -0{,}8007 \end{bmatrix};$$

$$y\left(\mathbf{x}^{(2)}\right) = 19{,}294 \; .$$

Третій крок:

$$\Delta\mathbf{x}^{(2)} = \left(\frac{\partial y}{\partial \mathbf{x}}\right)^{(2)} = \begin{bmatrix} -0{,}0543 \\ 0{,}0411 \end{bmatrix} ;$$

$$\left|\left(\frac{\partial y}{\partial x_1}\right)^{(2}\right| = \left|-0{,}0543\right| > 0{,}01 = \varepsilon \; , \quad \left|\left(\frac{\partial y}{\partial x_2}\right)^{(2)}\right| = \left|0{,}0411\right| > 0{,}01 = \varepsilon \; ;$$

$$y\left(\lambda^{(2)}\right) = y\left(\mathbf{x}^{(2)} + \lambda^{(2)}\left(\frac{\partial y}{\partial \mathbf{x}}\right)^{(2)}\right) = -0{,}0009\left(\lambda^{(2)}\right)^3 - 0{,}0561\left(\lambda^{(2)}\right)^2 + 0{,}0047\lambda^{(2)} + 19{,}294;$$

$$\left(\frac{dy}{d\lambda}\right)^{(2)} = -0{,}0027\left(\lambda^{(2)}\right)^2 - 0{,}1122\lambda^{(2)} + 0{,}0047 = 0 \; ;$$

$$\lambda^{(2)} = \frac{0{,}0561 \pm \sqrt{0{,}0561^2 + 0{,}0027 \cdot 0{,}0047}}{-0{,}0027} == \begin{cases} -41{,}5956 = \lambda^{(2)}_1, \\ 0{,}0418 = \lambda^{(2)}_2; \end{cases}$$

$$\left(\frac{d^2 y}{d\lambda^2}\right)^{(2)} = -0{,}0054\lambda^{(2)} - 0{,}1122 = \begin{cases} 0{,}1124\Big|_{\lambda_1^{(2)}=-41{,}5956}, \\ -0{,}1124\Big|_{\lambda_2^{(2)}=0{,}0418}; \end{cases}$$

$$\lambda^{(2)} = 0{,}0418 \ ;$$

$$\mathbf{x}^{(3)} = \mathbf{x}^{(2)} + \lambda^{(2)}\left(\frac{\partial y}{\partial \mathbf{x}}\right)^{(2)} = \begin{bmatrix} -1{,}927 \\ -0{,}8007 \end{bmatrix} + 0{,}0418\begin{bmatrix} -0{,}0543 \\ 0{,}0411 \end{bmatrix} = \begin{bmatrix} -1{,}9293 \\ -0{,}799 \end{bmatrix};$$

$$y\left(\mathbf{x}^{(2)}\right) = 19{,}29405 \ .$$

У новій точці наближення $\mathbf{x}^{(3)^{\mathrm{T}}} = \begin{bmatrix} -1{,}9293 & -0{,}799 \end{bmatrix}$ вектор перших частинних похідних $\left(\frac{\partial y}{\partial \mathbf{x}}\right)^{(3)} = \begin{bmatrix} -0{,}0007 \\ -0{,}0023 \end{bmatrix}$. Обидві частинні похідні за модулем виявляються менше заданої точності обчислення:

$$\left|\left(\frac{\partial y}{\partial x_1}\right)^{(3}\right| = \left|-0{,}0007\right| < 0{,}01 = \varepsilon \ , \ \left|\left(\frac{\partial y}{\partial x_2}\right)^{(3)}\right| = \left|-0{,}0023\right| < 0{,}01 = \varepsilon \ .$$

Це свідчить про те, що задана точність наближення досягнута. Отже, локальний максимум $y^* = 19{,}29405$ розташовується в точці $\mathbf{x}^{*^{\mathrm{T}}} = \begin{bmatrix} -1{,}9293 & -0{,}799 \end{bmatrix}$. Задачу вирішено.

Порівнюючи отримані результати з результатами розв’язання цієї задачі методом Ейлера (див. підрозділ 4.5), можна переконатися в тому, що результати збігаються в межах заданої точності.

При використанні методу найшвидшого спуску слід пам'ятати, що параметр довжини кроку λ повинен бути обов'язково додатним. У протилежному разі він змінить напрямок руху до екстремуму, який визначається направляючим вектором $\Delta\mathbf{x}$, на протилежний. Цей факт може служити додатковим засобом контролю правильності обчислювального процесу на поточному кроці. Якщо параметр λ має від'ємний знак, треба перервати процес обчислень до виявлення і усунення помилкових дій.

5.3. Метод Ньютона

Найбільш поширеним методом другого порядку є метод Ньютон. Метод має ряд модифікацій, що поліпшують характеристики збіжності методу.

Метод Ньютона уявляє собою ітераційну обчислювальну процедуру, в основі якої лежить апроксимація в точці $\Delta\mathbf{x}^{(k)}$ функції цілі $y(\mathbf{x})$ квадратичною функцією з наступним вибором направляючого вектора $\Delta\mathbf{x}^{(k)}$, що відразу приводить в екстремальну точку цієї допоміжної квадратичної (апроксимуючої) функції.

На рис.5.2 подана графічна інтепретація процесу оптимізації методом Ньютона на прикладі пошуку мінімуму функції двох змінних. Жирними стрілками позначені два кроки руху в області припустимих рішень, тонкими стрілками – один крок по поверхні функції цілі.

Якщо $y(\mathbf{x})$ – від'ємно або додатно визначена квадратична функція, то метод Ньютона дозволяє знайти екстремальну точку за один крок. Якщо $y(\mathbf{x})$ – увігнута або опукла функція, то метод гарантує відшукання екстремуму функції за кінцеве число кроків. У загальному випадку цей процес може привести до будь-якої стаціонарної точки

У класичному методі Ньютона параметр довжини кроку $\lambda^{(k)}$, що характеризує довжину кроку в напрямку рухи до екстремуму функції, приймається рівним 1. У цьому випадку основне рекурентне співвідношення (5.3) для методу Ньютона має вигляд

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \Delta\mathbf{x}^{(k)} . \tag{5.22}$$

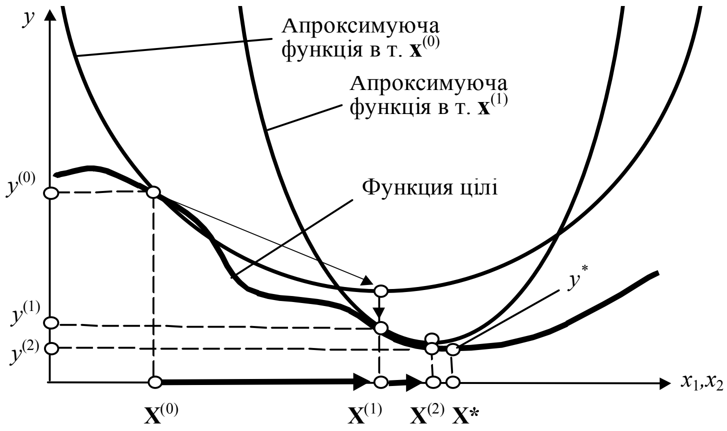


Рис.5.2 – Графічна інтепретація методу Ньютона

Направляючий вектор $\Delta\mathbf{x}^{(k)}$ у методі Ньютона на k-му кроці оптимізації і при пошуку локального максимуму, і при пошуку локального мінімуму визначається співвідношенням

$$\Delta\mathbf{x}^{(k)} = -\left(\mathbf{H}^{-1}\right)^{(k)} \left(\frac{\partial y}{\partial \mathbf{x}}\right)^{(k)} . \tag{5.23}$$

Підставляючи (5.23) у (5.22), одержуємо рекурентне співвідношення для визначення нової точки наближення на k-му кроці оптимізації методом Ньютона

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} - \left(\mathbf{H}^{-1}\right)^{(k)} \left(\frac{\partial y}{\partial \mathbf{x}}\right)^{(k)} . \tag{5.24}$$

Вираз (5.24) визначає величину приросту вектора змінних $\Delta\mathbf{x}^{(k)}$ із k-й точки наближення $\mathbf{x}^{(k)}$ до точки екстремуму квадратичної функції, що апроксимує задану функцію $y(\mathbf{x})$. Покажемо це, для чого розкладемо функцію цілі $y(\mathbf{x})$ у ряд Тейлора в околі точки наближення $\mathbf{x}^{(k)}$ у приростах, обмежуючи членами розкладання першого і другого порядків:

$$\Delta y^{(k)} = \left(\frac{\partial y}{\partial \mathbf{x}}\right)^{(k)\mathrm{T}} \Delta \mathbf{x}^{(k)} + \frac{1}{2}\Delta \mathbf{x}^{(k)\mathrm{T}} \mathbf{H}^{(k)} \Delta \mathbf{x}^{(k)} . \tag{5.25}$$

Подивимося на (5.25) як на функцію Δy від приросту аргументу $\Delta \mathbf{x}$:

$$\Delta y(\Delta \mathbf{x}) = \left(\frac{\partial y}{\partial \mathbf{x}}\right)^{(k)\mathrm{T}} \Delta \mathbf{x} + \frac{1}{2}\Delta \mathbf{x}^{\mathrm{T}} \mathbf{H}^{(k)} \Delta \mathbf{x} . \tag{5.26}$$

Керуючись тим, що приріст функції є незначним у точці екстремуму, спробуємо знайти такий приріст $\Delta \mathbf{x}$, що приведе в точку з мінімальним приростом Δy або, що те саме, в точку мінімуму апроксимуючої функції (5.26). Найбільш просто цю задачу можна вирішити методом Ейлера, для чого необхідно взяти похідну від функції (5.26) і прирівняти її нулю:

$$\left(\frac{\partial y}{\partial \mathbf{x}}\right)^{(k)} + \mathbf{H}^{(k)} \Delta \mathbf{x} = 0 \tag{5.27}$$

Вирішення рівняння (5.27) у випадку неособливості матриці $\mathbf{H}^{(k)}$ являє собою стаціонарну точку функції (5.26) і одночасно приріст вектора змінних на k-му кроці оптимізації методом Ньютона:

$$\Delta \mathbf{x} = -\left(\mathbf{H}^{-1}\right)^{(k)} \left(\frac{\partial y}{\partial \mathbf{x}}\right)^{(k)} = \Delta \mathbf{x}^{(k)} . \tag{5.28}$$

Порівняй з (5.23).

Розглянемо метод Ньютона на конкретному прикладі.

Приклад 5.2. Знайти методом Ньютона максимум функції (5.17)

$$y = 2x_1^3 + x_2^3 - 3x_1^2 x_2 + 3x_1 x_2^2 - 15x_1 \ .$$

при початковому наближенні $\mathbf{x}^{(0)\,\mathrm{T}} = \begin{bmatrix} -1{,}8 & -1 \end{bmatrix}$ і точності обчислення $\varepsilon = 0{,}01$.

Розв'язання. Початкове значення функції $y\left(\mathbf{x}^{(0)}\right) = 18{,}0976$. Це значення знадобиться для контролю правильності руху на першому кроці. Обчислимо перші частинні похідні функції в початковій точці:

$$\left(\frac{\partial y}{\partial \mathbf{x}}\right)^{(0)} = \begin{bmatrix} \frac{\partial y}{\partial x_1} \\ \frac{\partial y}{\partial x_2} \end{bmatrix}^{(0)} = \begin{bmatrix} 6x_1^2 - 6x_1 x_2 + 3x_2^2 - 15 \\ 3x_2^2 - 3x_1^2 + 6x_1 x_2 \end{bmatrix}^{(0)} = \begin{bmatrix} -3{,}36 \\ 4{,}08 \end{bmatrix}.$$

Частинні похідні за модулем в початковій точці наближення значно перевершують задану точність обчислення:

$$\left|\left(\frac{\partial y}{\partial x_1}\right)^{(0)}\right| = \left|-3{,}36\right| > 0{,}01 = \varepsilon; \quad \left|\left(\frac{\partial y}{\partial x_2}\right)^{(0)}\right| = \left|4{,}08\right| > 0{,}01 = \varepsilon \ .$$

Отже, відповідно до критерію (5.5) необхідно зробити, принаймні, один крок процесу максимізації.

Визначимо відповідно до виразів (5.43), (5.44) і (5.45) всі другі частині і змішані похідні в початковій точці наближення:

$$\left(\frac{\partial^2 y}{\partial x_1^2}\right)^{(0)} = 12x_1^{(0)} - 6x_2^{(0)} = 12 \cdot \left(-1{,}8\right) - 6 \cdot \left(-1\right) = -15{,}6 \ ;$$

$$\left(\frac{\partial^2 y}{\partial x_2^2}\right)^{(0)} = 6x_1^{(0)} + 6x_2^{(0)} = 6 \cdot \left(-1{,}8\right) + 6 \cdot (-1) = \text{-}16{,}8 \ ;$$

$$\left(\frac{\partial^2 y}{\partial x_1 \partial x_2}\right)^{(0)} = \left(\frac{\partial^2 y}{\partial x_2 \partial x_1}\right)^{(0)} = -6x_1^{(0)} + 6x_2^{(0)} = -6 \cdot (-1{,}8) + 6(-1) = 4{,}8 \ .$$

Сформуємо матрицю Гесса для початкової точки наближення і обернемо її:

$$\mathbf{H}^{(0)} = \begin{bmatrix} -15{,}6 & 4{,}8 \\ 4{,}8 & -16{,}8 \end{bmatrix}; \qquad \left(\mathbf{H}^{-1}\right)^{(0)} = \begin{bmatrix} -0{,}0703 & -0{,}0201 \\ -0{,}0201 & -0{,}0653 \end{bmatrix}$$

Визначимо направляючий вектор відповідно до виразу (5.23):

$$\Delta\mathbf{x}^{(0)} = -\left(\mathbf{H}^{-1}\right)^{(0)} \left(\frac{\partial y}{\partial \mathbf{x}}\right)^{(0)} =$$

$$= -\begin{bmatrix} -0{,}0703 & -0{,}0201 \\ -0{,}0201 & -0{,}0653 \end{bmatrix} \begin{bmatrix} -3{,}36 \\ 4{,}08 \end{bmatrix} = \begin{bmatrix} -0{,}1542 \\ 0{,}1989 \end{bmatrix}.$$

Обчислимо нову точку наближення відповідно до рекурентного співвідношення (5.22) :

$$\mathbf{x}^{(1)} = \mathbf{x}^{(0)} + \Delta\mathbf{x}^{(0)} = \begin{bmatrix} -1{,}8 \\ -1 \end{bmatrix} + \begin{bmatrix} -0{,}1542 \\ 0{,}1989 \end{bmatrix} = \begin{bmatrix} -1{,}9542 \\ -0{,}8011 \end{bmatrix}.$$

Значення функції цілі у новій точці наближення $y\left(\mathbf{x}^{(1)}\right) = 19{,}2887$. Оскільки воно перевищує початкове значення $y\left(\mathbf{x}^{(0)}\right) = 18{,}0976$, то рух відбувається в напрямку максимуму функції, тобто у правильному напрямку.

На другому кроці повторюємо обчислювальну процедуру, що мала місце на першому кроці максимізації, але вже щодо нової точки наближення $\mathbf{x}^{(1)}$:

$$\left(\frac{\partial y}{\partial \mathbf{x}}\right)^{(1)} = \begin{bmatrix} 6x_1^2 - 6x_1x_2 + 3x_2^2 - 15 \\ 3x_2^2 - 3x_1^2 + 6x_1x_2 \end{bmatrix}^{(1)} = \begin{bmatrix} 0{,}4456 \\ -0{,}1383 \end{bmatrix};$$

$$\left|\left(\frac{\partial y}{\partial x_1}\right)^{(1)}\right| = |0{,}4456| > 0{,}01 = \varepsilon; \quad \left|\left(\frac{\partial y}{\partial x_2}\right)^{(1)}\right| = |-0{,}1383| > 0{,}01 = \varepsilon;.$$

$$\left(\frac{\partial^2 y}{\partial x_1^2}\right)^{(1)} = 12x_1^{(1)} - 6x_2^{(1)} = 12\cdot(-1{,}9542) - 6\cdot(-0{,}8011) = -18{,}6438 \ ;$$

$$\left(\frac{\partial^2 y}{\partial x_2^2}\right)^{(1)} = 6x_1^{(1)} + 6x_2^{(1)} = 6\cdot(-1{,}9542) + 6\cdot(-0{,}8011) = -16{,}5318 \ ;$$

$$\left(\frac{\partial^2 y}{\partial x_1 \partial x_2}\right)^{(1)} = \left(\frac{\partial^2 y}{\partial x_2 \partial x_1}\right)^{(1)} = -6x_1^{(1)} + 6x_2^{(1)} =$$

$$= -6\cdot(-1{,}9542) + 6\cdot(-0{,}8011) = 6{,}9186 \ ;$$

$$\mathbf{H}^{(1)} = \begin{bmatrix} -18{,}6438 & 6{,}9186 \\ 6{,}9186 & -16{,}5318 \end{bmatrix}; \quad \left(\mathbf{H}^{-1}\right)^{(1)} = \begin{bmatrix} -0{,}0635 & -0{,}0266 \\ -0{,}0266 & -0{,}0716 \end{bmatrix} \ ;$$

$$\Delta\mathbf{x}^{(1)} = -\left(\mathbf{H}^{-1}\right)^{(1)} \left(\frac{\partial y}{\partial \mathbf{x}}\right)^{(1)} =$$

$$= -\begin{bmatrix} -0{,}0635 & -0{,}0266 \\ -0{,}0266 & -0{,}0716 \end{bmatrix}\begin{bmatrix} 0{,}4456 \\ -0{,}1383 \end{bmatrix} = \begin{bmatrix} 0{,}0246 \\ 0{,}002 \end{bmatrix};$$

$$\mathbf{x}^{(2)} = \mathbf{x}^{(1)} + \Delta\mathbf{x}^{(1)} = \begin{bmatrix} -1{,}9542 \\ -0{,}8011 \end{bmatrix} + \begin{bmatrix} 0{,}0246 \\ 0{,}002 \end{bmatrix} = \begin{bmatrix} -1{,}9296 \\ -0{,}7991 \end{bmatrix};$$

$$y\left(\mathbf{x}^{(2)}\right) = 19{,}294 \ .$$

У новій точці наближення $\mathbf{x}^{(2)^{\mathrm{T}}} = \left[-1{,}9296 \quad \text{-}\,0{,}7991\right]$ вектор перших частинних похідних $\left(\dfrac{\partial y}{\partial \mathbf{x}}\right)^{(2)} = \begin{bmatrix} 0{,}0041 \\ 0{,}0027 \end{bmatrix}$. Обидві частинні похідні за модулем виявляються менше заданої точності обчислення:

$$\left|\left(\frac{\partial y}{\partial x_1}\right)^{(2}\right| = \left|0{,}0041\right| < 0{,}01 = \varepsilon \ , \ \left|\left(\frac{\partial y}{\partial x_2}\right)^{(2)}\right| = \left|0{,}0027\right| < 0{,}01 = \varepsilon \, .$$

Це свідчить про те, що задана точність наближення досягнута.

Таким чином, локальний максимум $y^{*} = 19{,}294$ розташовується в точці $\mathbf{x}^{*^{\mathrm{T}}} = \left[-1{,}9296 \quad \text{-}\,0{,}7991\right]$. Задачу вирішено.

Порівнюючи отримані результати з результатами вирішення цієї задачі методом Ейлера (див. підрозділ 4.5) і методом найшвидшого спуску (див. приклад 5.1), можна переконатися в тому, що результати із заданою точністю збігаються.

Слід звернути увагу на те, що при однакових умовах прикладів 5.1 і 5.2 рішення в прикладі 5.2 отримано за два кроки, а в прикладі 5.1 – за три. Для методу Ньютона характерна швидка збіжність.

5.4. Модифікації методу Ньютона

Як уже згадувалося, метод Ньютона має ряд модифікацій, що поліпшують характеристики збіжності методу. Розглянемо деякі з них.

Узагальнений метод Ньютона. Метод заснований на квадратичній апроксимації і, як наслідок, потребує для своєї реалізації обчислення похідних тільки першого і другого порядків. При оптимізації нелінійних неквадратичних функцій така апроксимація не є точною, оскільки ігноруються члени третього і більш високого порядків. Причому відхилення значення апроксимуючої функції від початкової буде тим більше, чим більше $\Delta\mathbf{x}^{(k)}$. Тому доцільно процес оптимізації за формулою (5.24) модифікувати, ввівши додатний параметр . Введення в рекурентне співвідношення (5.24) параметра

$$\mathbf{x}^{(k+1)}=\mathbf{x}^{(k)}-\lambda^{(k)}\left(\mathbf{H}^{-1}\right)^{(k)}\left(\frac{\partial y}{\partial \mathbf{x}}\right)^{(k)}, \tag{5.29}$$

що змінюється в процесі оптимізації, приводить до узагальненого методу Ньютона, або методу Ньютона з регулюванням кроку. Одна з процедур по регулюванню кроку [9] включає такі етапи:

1. Точка $\mathbf{x}^{(k+1)}$ обчислюється відповідно до виразу (5.29), тобто за умови, що = 1, і обчислюється значення функції цілі $y_1^{(k+1)}$ в цій точці.

2. Здійснюється перевірка нерівності

$$y_1^{(k+1)}-y^{(k)}\leq \varepsilon\lambda^{(k)}\left(\frac{\partial y}{\partial \mathbf{x}}\right)^{(k)\mathrm{T}}\left(\mathbf{H}^{-1}\right)^{(k)}\left(\frac{\partial y}{\partial \mathbf{x}}\right)^{(k)}, \tag{5.30}$$

де $0 < \varepsilon < \frac{1}{2}$.

3. Якщо нерівність (5.30) виконується, то $\mathbf{x}^{(k+1)} = \mathbf{x}_1^{(k+1)}$, якщо порушується, то зменшується до тих пір, поки умова (5.30) не буде виконуватися.

Інша процедура регулювання кроку полягає у виборі такого параметра , що доставляє мінімум функції однієї змінної, тобто в розв’язанні задачі

$$y\left(\mathbf{x}^{(k)} - \lambda^{(k)}\left(\mathbf{H}^{-1}\right)^{(k)}\left(\frac{\partial y}{\partial \mathbf{x}}\right)^{(k)}\right) \to \min_{\mathbf{x}^{(k)} \in R^1} . \tag{5.31}$$

Модифікований метод Ньютона. Метод Ньютона, володіючи швидкою збіжністю, вимагає великої кількості елементарних операцій на кожному кроці оптимізації, пов'язаних з обчисленням, формуванням і обертанням матриці Гесса. Кількість цих операцій можна скоротити, якщо застосувати модифікований метод Ньютона, що дозволяє для визначення направляючого вектора $\Delta\mathbf{x}^{(k)}$ на другому і наступних кроках оптимізації використовувати обернену матрицю Гесса, отриману на першому кроці обчислень. При цьому вираз (5.29) набуває вигляду

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} - \lambda^{(k)}\left(\mathbf{H}^{-1}\right)^{(0)}\left(\frac{\partial y}{\partial \mathbf{x}}\right)^{(k)} . \tag{5.32}$$

Модифікований узагальнений метод Ньютона, заснований на приведенні матриці Гесса до діагонального вигляду. Введення в ньютонівську процедуру параметра дозволило не тільки підвищити збіжність методу, але і визначити випадок помилкового напрямку, що приводить не до екстремуму функції $y(\mathbf{x})$, а від нього [4]. Однак узагальненому методу Ньютона теж властиві два серйозних недоліки:

- якщо матриця других похідних $\mathbf{H}^{(k)}$ не є від'ємно визначеною у випадку пошуку максимуму (додатно визначеною у випадку пошуку мінімуму), обчислювальний процес (5.29) може привести не до поліпшення функції $y(\mathbf{x})$, а до погіршення; в результаті буде отримано , і процес перерветься в точці $\mathbf{x}^{(k)}$;
- у точці $\mathbf{x}^{(k)}$ взагалі може не існувати матриці, оберненої до $\mathbf{H}^{(k)}$.

Фіакко і Мак-Кормік [10] запропонували модифікацію узагальненого методу Ньютона, що дає можливість обминути названі труднощі. Спробуємо дати дещо інше трактування цієї модифікації, використовуючи ідею приведення матриці $\mathbf{H}^{(k)}$ до діагонального вигляду.

Нехай у виразі (5.25) $\mathbf{H}^{(k)}$ є неособлива матриця. Перейдемо до нової системи координат $\Delta\mathbf{v}$, в якій квадратична форма з (5.25) зазнає таке перетворення:

$$\Delta\mathbf{x}^{(k)\mathrm{T}}\mathbf{H}^{(k)}\Delta\mathbf{x}^{(k)} = \Delta\mathbf{v}^{(k)\mathrm{T}}\mathbf{B}^{(k)}\Delta\mathbf{v}^{(k)}, \tag{5.33}$$

де $\mathbf{B}^{(k)}$ – трикутна матриця; $\Delta\mathbf{v}^{(k)}=\mathbf{P}^{(k)}\Delta\mathbf{x}^{(k)}$; $\mathbf{P}^{(k)}$ – матриця перетворення. При цьому зворотне перетворення змінних має вигляд:

$$\Delta\mathbf{x}^{(k)}=\left(\mathbf{P}^{-1}\right)^{(k)}\Delta\mathbf{v}^{(k)}. \tag{5.34}$$

Підставляючи (5.34) у (5.25) , одержимо

$$\Delta y^{(k)}=\sum_{j=1}^{n}\Delta y_j^{(k)}=\sum_{j=1}^{n}\left[\left(\frac{\partial y}{\partial v_j}\right)^{(k)}\Delta v_j^{(k)}+\left(\frac{\partial^2 y}{\partial v_j^2}\right)^{(k)}\left(\Delta v_j^{(k)}\right)^2\right], \tag{5.35}$$

де $\left(\dfrac{\partial y}{\partial v_j}\right)^{(k)}$ – j-а складова n-вимірного вектора

$$\left(\frac{\partial y}{\partial \mathbf{v}}\right)^{(k)\mathrm{T}} = \left(\frac{\partial y}{\partial \mathbf{x}}\right)^{(k)\mathrm{T}} \left(\mathbf{P}^{-1}\right)^{(k)}, \tag{5.36}$$

$\dfrac{\partial^2 y}{\partial v_j^2}$ - значення j-го діагонального елемента діагональної матриці n-го порядку

$$\mathbf{B}^{(k)} = \frac{\partial^2 y}{\partial \mathbf{v}^2} = (\mathbf{P}^{-1})^{(k)^T} \mathbf{H}^{(k)} (\mathbf{P}^{-1})^{(k)} \ . \tag{5.37}$$

Квадратична функція (5.35) подана в сепарабельному вигляді. Отже, вибір напрямку по кожній з нових координат є незалежним і зводиться до безумовної оптимізації квадратичної функції однієї змінної. При цьому можуть бути два випадки.

У першому, коли

$$\left(\frac{\partial^2 y}{\partial v_j^2}\right)^{(k)} \neq 0 \quad \text{и} \quad \left(\frac{\partial y}{\partial v_j}\right)^{(k)} \neq 0 \ , \tag{5.38}$$

оптимальне рішення визначається відношенням

$$\Delta v^{(k)}_j = -\frac{\left(\dfrac{\partial y}{\partial v_j}\right)^{(k)}}{\left|\left(\dfrac{\partial^2 y}{\partial v_j^2}\right)^{(k)}\right|} \ . \tag{5.39}$$

У другому випадку, коли

$$\left(\frac{\partial^2 y}{\partial v_j^2}\right)^{(k)} \neq 0, \quad \text{а} \quad \left(\frac{\partial y}{\partial v_j}\right)^{(k)} = 0 \ , \tag{5.40}$$

процес оптимізації опинився безпосередньо в екстремальній точці функції (5.35). Це випадок, характерний для сідлової точки функції $y(\mathbf{x})$.

Умова (5.40) свідчить про те, що рух у будь-який бік уздовж координати v_j приводить до поліпшення Δy_j. Зокрема, можна взяти

$$\Delta v_j^{(k)} = \pm\sqrt{\sum_{i\in\mathbf{T}}\left(\Delta v_i^{(k)}\right)^2} \ , \tag{5.41}$$

де $\mathbf{T}$ - підмножина індексів змінних v_i, прирости яких $\Delta v_i^{(k)}$ визначаються співвідношенням (5.39), тобто змінних щодо першого випадку. Знак визначається випадково або в результаті аналізу зміни функції $y(\mathbf{x})$ при протилежних приростах змінної $v_j^{(k)}$.

Після вибору напрямку руху $\Delta\mathbf{v}^{(k)}$ згідно з формулами (5.38) – (5.41) можна перейти до цікавлячого нас напрямку $\Delta\mathbf{x}^{(k)}$. Для цього треба скористатися співвідношенням (5.34) і далі зайнятися звичайною процедурою вибору параметра , характерного для узагальненого методу Ньютона.

При неопуклості й неугнутості функції $y(\mathbf{x})$ умови (5.38) – (5.41) дозволяють знайти напрямок досить швидкого убування функції $y(\mathbf{x})$, якщо процес оптимізації знаходиться в околі або навіть безпосередньо в точці екстремуму або в сідловій точці, що становить досить складну задачу для більшості методів безумовної оптимізації.

Таким чином, запропонований алгоритм не дозволяє визначити ненульовий направляючий вектор $\Delta\mathbf{x}^{(k)}$ тільки в єдиному випадку, коли процес оптимізації знаходиться в точці $\mathbf{x}^{(k)}$, в якій матриця $\mathbf{H}^{(k)}$ додатно напіввизначена, а $\left(\dfrac{\partial y}{\partial \mathbf{x}}\right)^{(k)} = 0$. Для з'ясування характеру цієї точки і вибору направляючого вектора $\Delta\mathbf{x}^{(k)}$, якщо ця точка не є екстремумом, необхідний аналіз частинних похідних третього і більш високих порядків.

Градієнтно-ньютонівський метод. Найбільш істотним недоліком класичного методу Ньютона є визначення помилкового напрямку оптимізації, пов'язаного з вирішенням рівняння (5.27). Знайдене з його допомогою рішення визначає переміщення по гіперповерхні функції цілі в бік найближчої стаціонарної точки. Тому при пошуку максимуму або мінімуму обраний напрямок може привести до протилежного екстремуму або сідлової точки.

Проста модифікація методу Ньютона, що враховує особливість методу найшвидшого спуску визначати напрямок гарантованого поліпшення функції (зростання при максимізації і убування при мінімізації), дозволяє позбутися помилкового вибору напрямку.

Основне рекурентне співвідношення для градієнтно-ньютонівського методу при пошуку максимуму має вигляд

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \mathbf{sign}\left(\frac{\partial y}{\partial \mathbf{x}}\right)^{(k)\mathrm{T}} \left|\Delta\mathbf{x}^{(k)}\right|, \tag{5.42}$$

де

$$\Delta\mathbf{x}^{(k)} = -\left(\mathbf{H}^{-1}\right)^{(k)} \left(\frac{\partial y}{\partial \mathbf{x}}\right)^{(k)} \text{ – див. метод Ньютона;} \tag{5.43}$$

$$\mathbf{sign}\left(\frac{\partial y}{\partial \mathbf{x}}\right)^{(k)\mathrm{T}} = \left[sign\left(\frac{\partial y}{\partial x_1}\right)^{(k)} \quad sign\left(\frac{\partial y}{\partial x_2}\right)^{(k)} \quad \cdots \quad sign\left(\frac{\partial y}{\partial x_n}\right)^{(k)} \right] –$$

вектор знака градієнта.

Підставляючи (5.43) у (5.42), маємо

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \mathbf{sign}\left(\frac{\partial\, y}{\partial\, \mathbf{x}}\right)^{(k)\mathrm{T}} \left| \left(\mathbf{H}^{-1}\right)^{(k)} \left(\frac{\partial\, y}{\partial\, \mathbf{x}}\right)^{(k)} \right| . \qquad (5.44)$$

При пошуку мінімуму основне рекурентне співвідношення має вигляд

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} - \mathbf{sign}\left(\frac{\partial\, y}{\partial\, \mathbf{x}}\right)^{(k)\mathrm{T}} \left|\Delta\mathbf{x}^{(k)}\right| . \qquad (5.45)$$

Як бачимо, абсолютне значення вектора приросту змінних $\Delta\mathbf{x}^{(k)}$ визначається за методом Ньютона. Знайдений приріст коректується за допомогою вектора знака градієнта. Рух по кожній координаті x_i буде визначатися тільки знаком відповідної частинної похідної.

Приклад 5.3. Продемонструємо роботу методу в умовах прикладу 5.1, тобто при пошуку максимуму функції (4.17):

$$y = 2x_1^3 + x_2^3 - 3x_1^2 x_2 + 3x_1 x_2^2 - 15x_1 \ .$$

при початковому наближенні $\mathbf{x}^{(0)\mathrm{T}} = \left[-1{,}8 \quad -1\right]$ і точності обчислення $\varepsilon = 0{,}01$.

У даному прикладі результати, одержані градієнтно-ньютонівським методом, збігаються з результатами класичного методу (див. приклад 5.1):

$$\mathbf{sign}\left(\frac{\partial\, y}{\partial\, \mathbf{x}}\right)^{(0)\mathrm{T}} = \left[sign\left(\frac{\partial y}{\partial x_1}\right)^{(0)} \quad sign\left(\frac{\partial y}{\partial x_2}\right)^{(0)} \right] =$$

$$= \left[sign(-3{,}36) \quad sign(4{,}08)\right] = \left[-1 \quad 1\right] ;$$

$$\mathbf{x}^{(1)} = \mathbf{x}^{(0)} + \mathbf{sign}\left(\frac{\partial y}{\partial \mathbf{x}}\right)^{(0)\mathrm{T}} \left|\Delta\mathbf{x}^{(0)}\right| =$$

$$=\begin{bmatrix}-1{,}8\\-1\end{bmatrix}+\begin{bmatrix}-1 & 1\end{bmatrix}\left|\begin{bmatrix}-0{,}1542\\0{,}1989\end{bmatrix}\right|=\begin{bmatrix}-1{,}9542\\-0{,}8011\end{bmatrix}.$$

$$\mathbf{sign}\left(\frac{\partial y}{\partial \mathbf{x}}\right)^{(1)\mathrm{T}}=\left[sign(0{,}2315)\quad sign(0{,}1893)\right]=\begin{bmatrix}1 & 1\end{bmatrix};$$

$$\mathbf{x}^{(2)}=\mathbf{x}^{(1)}+\mathbf{sign}\left(\frac{\partial y}{\partial \mathbf{x}}\right)^{(1)\mathrm{T}}\left|\Delta\mathbf{x}^{(1)}\right|=$$

$$=\begin{bmatrix}-1{,}9542\\-0{,}8011\end{bmatrix}+\begin{bmatrix}1 & 1\end{bmatrix}\begin{bmatrix}0{,}0246\\0{,}002\end{bmatrix}=\begin{bmatrix}-1{,}9296\\-0{,}7991\end{bmatrix};$$

$$y\left(\mathbf{x}^{(2)}\right)=19{,}294.$$

У точці наближення $\mathbf{x}^{(2)\mathrm{T}}=\left[-1{,}9296\quad -0{,}7991\right]$ значення функції $y^{(2)}=19{,}294$. Оскільки $\left|\left(\frac{\partial y}{\partial x_1}\right)^{(2}\right|=\left|0{,}0041\right|<0{,}01=\varepsilon$ і $\left|\left(\frac{\partial y}{\partial x_2}\right)^{(2)}\right|=\left|0{,}0027\right|<0{,}01=\varepsilon$, то процес оптимізації закінчуємо. Причому локальний максимум $y^{*}=19{,}294$ розташовується в точці $\mathbf{x}^{*\mathrm{T}}=\left[-1{,}9296\quad -0{,}7991\right]$, тобто результати цілком збігаються з результатами класичного методу Ньютона.

Модифікований градієнтно-ньютонівський метод. Градієнтно-ньютонівський метод, володіючи швидкою збіжністю, вимагає великої кількості елементарних операцій на кожному кроці оптимізації, пов'язаних з обчисленням, формуванням і обертанням матриці Гесса. Кількість цих операцій можна скоротити, якщо застосувати модифікований градієнтно-ньютонівський метод, що дозволяє для визначення направляючого вектора $\Delta\mathbf{x}^{(k)}$ на другому і наступних кроках оптимізації використовувати обернену матрицю Гесса, отриману на першому кроці обчислень. При цьому рекурентне співвідношення (5.44) для визначення нової точки наближення у випадку пошуку максимуму матиме вигляд

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \mathbf{sign}\left(\frac{\partial y}{\partial \mathbf{x}}\right)^{(k)\mathrm{T}} \left| \left(\mathbf{H}^{-1}\right)^{(0)} \left(\frac{\partial y}{\partial \mathbf{x}}\right)^{(k)} \right| . \qquad (5.46)$$

Аналогічне рекурентне співвідношення при пошуку мінімуму матиме вигляд

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} - \mathbf{sign}\left(\frac{\partial y}{\partial \mathbf{x}}\right)^{(k)\mathrm{T}} \left| \left(\mathbf{H}^{-1}\right)^{(0)} \left(\frac{\partial y}{\partial \mathbf{x}}\right)^{(k)} \right| . \qquad (5.47)$$

Приклад 5.4. Продемонструємо роботу модифікованого градієнтно-ньютонівського методу в умовах прикладу 5.1, тобто при пошуку максимуму функції (4.17)

$$y = 2x_1^3 + x_2^3 - 3x_1^2 x_2 + 3x_1 x_2^2 - 15x_1$$

при тих же початковому наближенні $\mathbf{x}^{(0)^{\mathrm{T}}} = \left[-1{,}8 \quad -1\right]$ і точності обчислення $\varepsilon = 0{,}01$.

У даному прикладі на першому кроці результати, одержані модифікованим градієнтно-ньютонівским методом, збігаються з результатами класичного методу і результатами немодифікованого градієнтно-ньютонівського методу:

$$\mathbf{x}^{(1)} = \mathbf{x}^{(0)} + \mathbf{sign}\left(\frac{\partial\, y}{\partial\, \mathbf{x}}\right)^{(0)\,\mathrm{T}} \left|\left(\mathbf{H}^{-1}\right)^{(0)} \left(\frac{\partial\, y}{\partial\, \mathbf{x}}\right)^{(0)}\right| =$$

$$= \begin{bmatrix} -1{,}8 \\ -1 \end{bmatrix} + \begin{bmatrix} -1 & 1 \end{bmatrix} \left\| \begin{bmatrix} -0{,}0703 & -0{,}0201 \\ -0{,}0201 & -0{,}0653 \end{bmatrix} \begin{bmatrix} -3{,}36 \\ 4{,}08 \end{bmatrix} \right\| = \begin{bmatrix} -1{,}9542 \\ -0{,}8011 \end{bmatrix} .$$

$$y\left(\mathbf{x}^{(1)}\right) = 19{,}2887 \ .$$

Результати другого кроку:

$$\mathbf{x}^{(2)} = \mathbf{x}^{(1)} + \mathbf{sign}\left(\frac{\partial\, y}{\partial\, \mathbf{x}}\right)^{(1)\,\mathrm{T}} \left|\left(\mathbf{H}^{-1}\right)^{(0)} \left(\frac{\partial\, y}{\partial\, \mathbf{x}}\right)^{(1)}\right| =$$

$$= \begin{bmatrix} -1{,}9542 \\ -0{,}8011 \end{bmatrix} + \begin{bmatrix} 1 & -1 \end{bmatrix} \left\| \begin{bmatrix} -0{,}0703 & -0{,}0201 \\ -0{,}0201 & -0{,}0653 \end{bmatrix} \begin{bmatrix} 0{,}4456 \\ -0{,}1383 \end{bmatrix} \right\| =$$

$$= \begin{bmatrix} -1{,}9542 \\ -0{,}8011 \end{bmatrix} + \begin{bmatrix} 1 & -1 \end{bmatrix} \left\| \begin{bmatrix} -0{,}0285 \\ 0{,}0001 \end{bmatrix} \right\| = \begin{bmatrix} -1{,}9257 \\ -0{,}801 \end{bmatrix} ;$$

$$y\left(\mathbf{x}^{(2)}\right) = 19{,}2938 \ .$$

Результати третього кроку:

$$\mathbf{x}^{(3)} = \mathbf{x}^{(2)} + \mathbf{sign}\left(\frac{\partial\, y}{\partial\, \mathbf{x}}\right)^{(2)\,\mathrm{T}} \left|\left(\mathbf{H}^{-1}\right)^{(0)} \left(\frac{\partial\, y}{\partial\, \mathbf{x}}\right)^{(2)}\right| =$$

$$= \begin{bmatrix} -1{,}9257 \\ -0{,}801 \end{bmatrix} + \begin{bmatrix} -1 & 1 \end{bmatrix} \left\| \begin{bmatrix} -0{,}0703 & -0{,}0201 \\ -0{,}0201 & -0{,}0653 \end{bmatrix} \begin{bmatrix} -0{,}0802 \\ 0{,}0547 \end{bmatrix} \right\| =$$

$$= \begin{bmatrix} -1{,}9257 \\ -0{,}801 \end{bmatrix} + \begin{bmatrix} -1 & 1 \end{bmatrix} \left| \begin{bmatrix} 0{,}0045 \\ -0{,}0020 \end{bmatrix} \right| = \begin{bmatrix} 1{,}9302 \\ -0{,}799 \end{bmatrix};$$

$$y\left(\mathbf{x}^{(3)}\right) = 19{,}29404 \ .$$

Результати четвертого кроку:

$$\mathbf{x}^{(4)} = \mathbf{x}^{(3)} + \mathbf{sign}\left(\frac{\partial y}{\partial \mathbf{x}}\right)^{(3)\,\mathrm{T}} \left| \left(\mathbf{H}^{-1}\right)^{(0)} \left(\frac{\partial y}{\partial \mathbf{x}}\right)^{(3)} \right| =$$

$$= \begin{bmatrix} -1{,}9302 \\ -0{,}799 \end{bmatrix} + \begin{bmatrix} 1 & -1 \end{bmatrix} \left| \begin{bmatrix} -0{,}0703 & -0{,}0201 \\ -0{,}0201 & -0{,}0653 \end{bmatrix} \begin{bmatrix} 0{,}0158 \\ -0{,}0084 \end{bmatrix} \right| =$$

$$= \begin{bmatrix} -1{,}9302 \\ -0{,}799 \end{bmatrix} + \begin{bmatrix} 1 & -1 \end{bmatrix} \left| \begin{bmatrix} -0{,}0009 \\ 0{,}0002 \end{bmatrix} \right| = \begin{bmatrix} 1{,}9293 \\ -0{,}7992 \end{bmatrix};$$

$$y\left(\mathbf{x}^{(4)}\right) = 19{,}29405 \ .$$

У точці $\mathbf{x}^{(4)\,\mathrm{T}} = \begin{bmatrix} -1{,}9293 & \text{-}\,0{,}7992 \end{bmatrix}$ значення функції $\left|\left(\frac{\partial y}{\partial x_1}\right)^{(4}\right| = \left|-0{,}002\right| < 0{,}1 = \varepsilon$ і $\left|\left(\frac{\partial y}{\partial x_2}\right)^{(4)}\right| = \left|0{,}001\right| < 0{,}1 = \varepsilon$. Тому процес оптимізації зупиняється. Локальний максимум $y^{*} = 19{,}29405$ досягається в точці $\mathbf{x}^{*\mathrm{T}} = \begin{bmatrix} -1{,}9293 & \text{-}\,0{,}7992 \end{bmatrix}$.

За швидкістю збіжності модифікований градієнтно-ньютонівський метод поступається раніше розглянутим, однак має більш простий процес обчислення.

5.5. Метод покоординатного спуску

Метод покоординатного спуску, відомий також як метод Гаусса-Зейделя, має різку відмінність від розглянутих раніше.

Метод покоординатного спуску уявляє собою ітераційну процедуру. Кожна ітерація складається з n кроків (за числом змінних задачі). Як і у всіх методах безумовної оптимізації задаються або вибираються точка початкового наближення і точність обчислення. Однак, на відміну від інших методів, на кожному k-му кроці змінюється тільки одна змінна x_r, за якою порушена необхідна умова (5.14) із точністю до ε:

$$\left| \left(\frac{\partial y}{\partial x_r} \right)^* \right| \leq \varepsilon \; . \tag{5.48}$$

Рух до екстремуму по гіперповерхні функції цілі на кожному кроці здійснюється паралельно осі змінної x_r. Рис.5.3 демонструє стратегію наближення до точки екстремуму функції двох змінних методом покоординатного спуску.

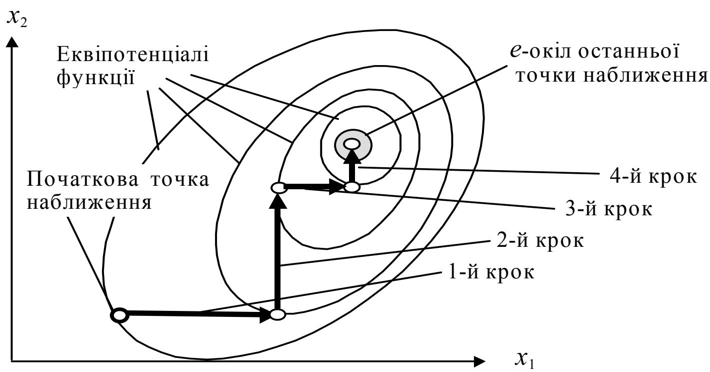


Рис.5.3 – Приклад безумовної оптимізації функції двох змінних методом покоординатного спуску

Рекурентне співвідношення (5.3) можна подати у вигляді

$$\mathbf{x}^{(k+1)\mathrm{T}} = \mathbf{x}^{(k)\mathrm{T}} + \lambda^{(k)}\left[0 \ldots 0 \;\; \Delta x_r^{(k)} \;\; 0 \ldots 0\right]. \tag{5.49}$$

Цей метод випливає з методу найшвидшого спуску, якщо останній обмежити вимогою змінювати на кожному кроці тільки одну змінну, тобто направляючий вектор при пошуку максимуму функції визначається як градієнт

$$\Delta x_r^{(k)} = \left(\frac{\partial y}{\partial x_r}\right)^{(k)}, \tag{5.50}$$

а при пошуку мінімуму – як антиградієнт

$$\Delta x_r^{(k)} = -\left(\frac{\partial y}{\partial x_r}\right)^{(k)}. \tag{5.51}$$

Для функції однієї змінної обидва методи збігаються.

Залежно від стратегії, застосовуваної при виборі змінної $x_r^{(k)}$, і методу мінімізації функції однієї змінної можна генерувати досить велику кількість модифікацій методу, які відрізняються за швидкістю збіжності.

Розглянемо метод покоординатного спуску, що відрізняється найбільшою простотою організації обчислювального процесу. У цьому методі кожна ітерація складається з n кроків, вибір кроку здійснюється послідовно по кожній із змінних, а визначення довжини кроку в обраному напрямку проводиться в результаті розв'язання задачі оптимізації функцій однієї змінної за методом Ньютона, відомим як метод дотичних.

Нехай $x_r^{(k)}$ – значення r-ї складової вектора $\mathbf{x}^{(k)}$, отримане при зміні цієї змінної на r-му кроці k-ї ітерації. Тоді при пошуку максимуму

$$x_r^{(k+1)} = x_r^{(k)} + \lambda^{(k)} \left(\frac{\partial y}{\partial x_r} \right)^{(k)} . \tag{5.52}$$

Визначимо значення параметра за умовами (5.14) обертання в нуль частинної похідної , попередньо розклавши цю похідну в ряд Тейлора, обмежений лінійними членами і враховуючи зміну тільки однієї координати x_r при незмінних інших:

$$\frac{\partial y}{\partial x_r} \left(\mathbf{x}^{(k)} + \lambda^{(k)} \Delta \mathbf{x}^{(k)} \right) = \left(\frac{\partial y}{\partial x_r} \right)^{(k)} + \left(\frac{\partial^2 y}{\partial x_r^2} \right)^{(k)} \lambda^{(k)} \Delta x_r^{(k)} = 0 . \tag{5.53}$$

Підставляючи значення з рівнянь (5.50) у (5.53) і розв'язуючи останнє відносно , одержимо його значення

$$\lambda^{(k)} = \frac{1}{\left(\frac{\partial^2 y}{\partial x_r^2} \right)^{(k)}} . \tag{5.54}$$

З огляду на те, що параметр довжини кроку не може бути від'ємним, перепишемо (5.54) у вигляді

$$\lambda^{(k)} = \frac{1}{\left| \left(\frac{\partial^2 y}{\partial x_r^2} \right)^{(k)} \right|} . \tag{5.55}$$

Із виразу (5.52) приходимо до основного рекурентного співвідношення розглянутої модифікації покоординатного спуску при пошуку локального максимуму

$$\boxed{x_r^{(k+1)} = x_r^{(k)} + \frac{\left(\dfrac{\partial y}{\partial x_r}\right)^{(k)}}{\left|\left(\dfrac{\partial^2 y}{\partial x_r^2}\right)^{(k)}\right|}.} \tag{5.55}$$

Аналогічно одержуємо основне рекурентне співвідношення для безумовної мінімізації методом покоординатного спуску

$$\boxed{x_r^{(k+1)} = x_r^{(k)} - \frac{\left(\dfrac{\partial y}{\partial x_r}\right)^{(k)}}{\left|\left(\dfrac{\partial^2 y}{\partial x_r^2}\right)^{(k)}\right|}.} \tag{5.56}$$

Послідовність обчислювальних операцій при безумовній оптимізації методом покоординатного спуску така:

Перша ітерація.

1-й крок. Обчислюємо значення функції в початковій точці наближення (для перевірки правильності руху до точки екстремуму) і перевіряємо виконання необхідної умови локального екстремуму (5.48) за змінною x_1. Якщо умови виконуються, переходимо до аналізу необхідних умов за наступною змінною x_2. При порушенні умов визначаємо нову точку наближення за рекурентною формулою (5.55) у випадку максимізації або (5.56) у випадку мінімізації.

2-й крок. Обчислюємо значення функції в новій точці наближення і порівнюємо з попереднім. Якщо функція не поліпшується, шукаємо й усуваємо помилку першого кроку. Потім повторюємо п.1 для змінної x_2.

І так далі, поки не буде виконаний повний цикл із n кроків.

Друга ітерація. Повторюємо процедуру першої ітерації щодо останньої точки наближення.

І так далі, поки не будуть виконані необхідні умови по всім змінним одночасно.

Приклад 5.5. Продемонструємо роботу методу покоординатного спуску в умовах прикладу 5.1, тобто при пошуку максимуму функції (4.17)

$$y = 2x_1^3 + x_2^3 - 3x_1^2 x_2 + 3x_1 x_2^2 - 15x_1 \ .$$

при тих же початковому наближенні $\mathbf{x}^{(0)^{\mathrm{T}}} = \begin{bmatrix} -1{,}8 & -1 \end{bmatrix}$ і точності обчислення $\varepsilon = 0{,}01$.

1-ша ітерація. 1-й крок. Змінюється тільки змінна x_1. У точці $\mathbf{x}^{(0)^{\mathrm{T}}} = \begin{bmatrix} -1{,}8 & -1 \end{bmatrix}$ значення функції цілі $y^{(0)} = 18{,}0976$, а частинна похідна $\left(\dfrac{\partial y}{\partial x_1}\right)^{(0)} = -3{,}36$. Необхідна умова локального максимуму (5.48) за даною змінною не виконується: $\left|\left(\dfrac{\partial y}{\partial x_1}\right)^{(0)}\right| = 3{,}36 > 0{,}01 = \varepsilon$.

Використовуючи рекурентне співвідношення (5.55), одержуємо для змінної x_1 нове значення

$$x_1^{(1)} = x_1^{(0)} + \frac{\left(\dfrac{\partial y}{\partial x_1}\right)^{(0)}}{\left|\left(\dfrac{\partial^2 y}{\partial x_1^2}\right)^{(0)}\right|} = -1{,}8 + \frac{-3{,}36}{|-15{,}6|} = -2{,}0154 .$$

У новій точці наближення $\mathbf{x}^{(1)^{\mathrm{T}}} = \begin{bmatrix} -2{,}0154 & -1 \end{bmatrix}$ значення функції $y^{(1)} = 18{,}9978 > 18{,}0976 = y^{(0)}$, тобто має місце поліпшення функції.

Переходимо до 2-го кроку.

1-ша ітерація. 2-й крок.. У новій точці $\mathbf{x}^{(1)\,\mathrm{T}} = \left[-2{,}0154 \quad -1\right]$ змінюється тільки змінна x_2. Оскільки частинна похідна $\left(\dfrac{\partial y}{\partial x_2}\right)^{(1)} = 3{,}0931$, то необхідна умова локального максимуму за даною змінною не виконується: $\left|\left(\dfrac{\partial y}{\partial x_2}\right)^{(1)}\right| = 3{,}36 > 0{,}01 = \varepsilon$. Використовуючи рекурентне співвідношення (5.55), одержуємо для змінної x_2 нове значення

$$. \; x_2^{(2)} = x_2^{(1)} + \frac{\left(\dfrac{\partial y}{\partial x_1}\right)^{(1)}}{\left|\left(\dfrac{\partial^2 y}{\partial x_1^2}\right)^{(1)}\right|} = -1 + \frac{3{,}0931}{\left|-18{,}0924\right|} = -0{,}829$$

У новій точці наближення $\mathbf{x}^{(2)\,\mathrm{T}} = \left[-2{,}0154 \quad -0{,}829\right]$ значення функції $y^{(2)} = 19{,}2354 > 18{,}9978 = y^{(1)}$, тобто має місце поліпшення функції.

Переходимо до другої ітерації (третьому кроку).

2-га ітерація. 3-й крок.

$$\left|\left(\frac{\partial y}{\partial x_1}\right)^{(2)}\right| = 1{,}4081 > 0{,}01 \;; \qquad x_1^{(3)} = -2{,}0154 + \frac{1{,}4081}{\left|-19{,}2108\right|} = -1{,}9421\,;$$

$$\mathbf{x}^{(3)\,\mathrm{T}} = \left[-1{,}9421 \quad -0{,}829\right]; \qquad y^{(3)} = 19{,}2878 > 19{,}2354 = y^{(2)}.$$

2-га ітерація. 4-й крок.

$$\left|\left(\frac{\partial y}{\partial x_2}\right)^{(3)}\right| = 0,4064 > 0,01 \;; \quad x_2^{(4)} = -0,829 + \frac{0,4064}{|-16,6266|} = -0,8046\,;$$

$$\mathbf{x}^{(4)^{\mathrm{T}}} = \begin{bmatrix}-1,9421 & -0,8046\end{bmatrix}; \quad y^{(4)} = 19,2929 > 19,2878 = y^{(3)}.$$

3-я ітерація. 5-й крок.

$$\left|\left(\frac{\partial y}{\partial x_1}\right)^{(4)}\right| = 0,1969 > 0,01 \;; \quad x_1^{(5)} = -1,9421 + \frac{0,1969}{|-18,4776|} = -1,9314\,;$$

$$\mathbf{x}^{(5)^{\mathrm{T}}} = \begin{bmatrix}-1,9314 & -0,8046\end{bmatrix}; \quad y^{(5)} = 19,2929 > 19,2938 = y^{(4)}.$$

3-я ітерація. 6-й крок.

$$\left|\left(\frac{\partial y}{\partial x_2}\right)^{(5)}\right| = 0,0752 > 0,01 \;; \quad x_2^{(6)} = -0,8046 + \frac{0,0752}{|-16,416|} = -0,8\,;$$

$$\mathbf{x}^{(6)^{\mathrm{T}}} = \begin{bmatrix}-1,9314 & -0,8\end{bmatrix}; \quad y^{(6)} = 19,294 > 19,2938 = y^{(5)}.$$

4-а ітерація. 7-й крок.

$$\left|\left(\frac{\partial y}{\partial x_1}\right)^{(6)}\right| = 0,0311 > 0,01 \;; \quad x_1^{(7)} = -1,9314 + \frac{0,0311}{|-18,3768|} = -1,9297\,;$$

$$\mathbf{x}^{(7)\,\mathrm{T}} = \begin{bmatrix} -1{,}9297 & -0{,}8046 \end{bmatrix}; \qquad y^{(7)} = 19{,}29415 > 19{,}294 = y^{(6)}.$$

4-а ітерація. 8-й крок.

У новій точці наближення необхідна умова (5.48) виконується: $\left| \left(\frac{\partial y}{\partial x_2} \right)^{(7)} \right| = 0{,}0086 < 0{,}01$. Тому переходимо до наступного кроку. Але оскільки функція має тільки дві змінні, то можна вважати, що екстремум уже досягнутий, тому що попередній крок робився в точку з нульової частинною похідною за першою змінною.

Задачу вирішено. Остання точка наближення із заданою точністю є локальним екстремумом $\mathbf{x}^* = \mathbf{x}^{(7)\,\mathrm{T}} = \begin{bmatrix} -1{,}9297 & -0{,}8 \end{bmatrix}$. У цій точці функція приймає максимальне значення $y^{(7)} = 19{,}29415$.

У розглянутому прикладі задана точність рішення досягнута в результаті семи кроків. Приклад добре ілюструє можливості методу, який відрізняється від усіх раніше розглянутих винятковою простотою обчислювального процесу на кожному кроці, але потребує великої кількості кроків для досягнення необхідної точності.

Таблиця порівняння прямих методів безумовної оптимізації. У даному розділі були розглянуті п'ять прикладів безумовної оптимізації функції п'ятьма різними методами, причому умови оптимізації були абсолютно ідентичними. Табл.5.1 дозволяє дати порівняльну оцінку швидкості збіжності методів. Але зазначимо, що порівняння методів дуже умовне.

При необхідності машинної реалізації безумовної оптимізації функції рекомендується використовувати метод покоординатного спуску, оскільки машинне (не табличне) взяття перших, других і змішаних похідних функції багатьох змінних як відношення збільшення функції до малого збільшення аргументів не складає труднощів.

Таблиця 5.1 – Таблиця порівняння прямих методів безумовної оптимізації функції двох змінних (4.17)

№ п/п	Метод	Трудомісткість кроку	Можливість помилкового напрямку	Кількість кроків
1	Найшвидший спуск	Найбільш значна	Відсутня	4
2	Метод Ньютона	Значна	Має місце	2
3	Градієнтно-ньютонівський	Значна	Відсутня	2
4	Модифікований градієнтно-ньютонівський	Незначна	Відсутня	4
5	Покоординатний спуск	Наднезначна	Відсутня	7

5.6. Контрольні запитання і вправи

1. Які додаткові дані треба мати для застосування прямих методів безумовної оптимізації на відміну від класичних методів?

2. Як вибирається початкова точка наближення у прямих методах оптимізації?

3. Як вибирається точність обчислення в прямих методах оптимізації?

4. Навести основне рекурентне співвідношення, що пов'язує дві послідовні точки наближення до екстремуму багатовимірних функцій.

5. Що визначає направляючий вектор?

6. Чи може параметр довжини кроку бути від'ємним?

7. Яка умова закінчення процесу оптимізації?

8. Що розуміють під лінійною швидкістю збіжності прямих методів оптимізації?

9. Коли процес оптимізації збігається з надлінійною швидкістю?

10. Дати визначення квадратичній збіжності прямих методів оптимізації.

11. Які прямі методи відносяться до методів першого порядку? Другого порядку?

12. Яким способом визначається направляючий вектор при оптимізації функцій багатьох змінних методом найшвидшого спуску?

13. Яким способом визначається параметр довжини кроку при оптимізації методом найшвидшого спуску?

14. За допомогою методу найшвидшого спуску знайти мінімум функції $y = x_1^3 + x_2^3 - 6x_1x_2$ з точністю обчислення $\varepsilon = 0{,}15$, прийнявши як початкову точку наближення $\mathbf{x}^{(0)^{\mathrm{T}}} = \begin{bmatrix}2{,}7 & 2{,}5\end{bmatrix}$.

15. Знайти максимум функції $y = -3x_1^3 - \frac{1}{4}x_2^3 + x_1x_2$ методом найшвидшого спуску, якщо початкова точка наближення $\mathbf{x}^{(0)^{\mathrm{T}}} = \begin{bmatrix}0{,}2 & 0{,}5\end{bmatrix}$, а точність обчислення $\varepsilon = 0{,}1$.

16. Методом найшвидшого спуску знайти мінімум функції $y = \frac{2}{3}x_1^3 + \frac{1}{3}x_2^3 - x_1^2x_2 + x_1x_2^2 - 5x_1$, якщо точка початкового наближення $\mathbf{x}^{(0)^{\mathrm{T}}} = \begin{bmatrix}3 & 2\end{bmatrix}$, а точність обчислення $\varepsilon = 0{,}1$.

17. Яким способом визначаються направляючий вектор і параметр довжини кроку при оптимізації функцій багатьох змінних методом Ньютона?

18. За допомогою методу Ньютона знайти мінімум функції $y = x_1^3 + x_2^3 - 6x_1x_2$ з точністю обчислення $\varepsilon = 0{,}05$, прийнявши як точку початкового наближення $\mathbf{x}^{(0)^{\mathrm{T}}} = \begin{bmatrix}2{,}4 & 2{,}3\end{bmatrix}$.

19. Знайти максимум функції $y = -\frac{1}{x_1} - \frac{1}{2x_2} - 2x_1x_2$ методом Ньютона, якщо початкова точка наближення $\mathbf{x}^{(0)^{\mathrm{T}}} = [0{,}7 \quad 0{,}7]$, а точність обчислення $\varepsilon = 0{,}25$.

20. Методом Ньютона знайти мінімум функції , якщо точка початкового наближення $\mathbf{x}^{(0)^{\mathrm{T}}} = [3 \quad 2]$, а точність обчислення $\varepsilon = 0{,}01$.

21. Яким способом визначаються направляючий вектор і параметр довжини кроку при оптимізації функцій багатьох змінних узагальненим методом Ньютона?

22. За допомогою узагальненого методу Ньютона знайти мінімум функції $y = \frac{1}{x_1^2} + 4\sqrt{x_1x_2} + \frac{1}{x_2^2}$ з точністю обчислення $\varepsilon = 0{,}25$, прийнявши за точку початкового наближення $\mathbf{x}^{(0)^{\mathrm{T}}} = [1{,}5 \quad 1{,}7]$.

23. Знайти максимум функції $y = -\frac{1}{x_1} - \frac{1}{2x_2} - 2x_1x_2$ узагальненим методом Ньютона. Початкову точку наближення вибрати самостійно

24. Яким способом визначаються направляючий вектор і параметр довжини кроку при оптимізації функцій багатьох змінних модифікованим методом Ньютона?

25. Яким способом визначаються направляючий вектор і параметр довжини кроку при оптимізації функцій багатьох змінних модифікованим узагальненим методом Ньютона?

26. Яким способом визначається направляючий вектор при оптимізації функцій багатьох змінних градієнтно-ньютонівським методом?

27. У чому відмінність градієнтно-ньютонівського методу від методу Ньютона? Від методу найшвидшого спуску?

28. Яким способом визначається направляючий вектор при оптимізації функцій багатьох змінних модифікованим узагальненим градієнтно-ньютонівським методом?

29. У чому відмінність узагальненого градієнтно-ньютонівського методу від методу Ньютона? Узагальненого методу Ньютона? Градієнтно-ньютонівського методу?

30. Навести основне рекурентне співвідношення, що зв'язує дві послідовні точці наближення при оптимізації функцій багатьох змінних методом покоординатного спуску.

31. У чому основна відмінність методу покоординатного спуску від інших прямих методів оптимізації?

32. Показати, що основне рекурентне співвідношення (5.55) для методу покоординатного спуску збігається з основним рекурентним співвідношенням градієнтно-ньютонівського методу (5.46) при пошуку локального максимуму функції однієї змінної.

33. Показати, що основне рекурентне співвідношення (5.56) для методу покоординатного спуску збігається з основним рекурентним співвідношенням градієнтно-ньютонівського методу (5.47) при пошуку локального мінімуму функції однієї змінної.

34. За допомогою методу покоординатного спуску знайти мінімум функції $y = x_1^3 + x_2^3 - 6x_1x_2$ з точністю обчислення $\varepsilon = 0{,}05$, прийнявши за початкове наближення точку $\mathbf{x}^{(0)^{\mathrm{T}}} = \left[2{,}4 \quad 2{,}3\right]$.

35. Знайти максимум функції $y = -\dfrac{1}{x_1} - \dfrac{1}{2x_2} - 2x_1x_2$ методом покоординатного спуску, якщо початкова точка наближення $\mathbf{x}^{(0)^{\mathrm{T}}} = \left[0{,}7 \quad 0{,}7\right]$, а точність обчислення $\varepsilon = 0{,}25$.

36. Методом покоординатного спуску знайти мінімум функції , якщо точка початкового наближення $\mathbf{x}^{(0)^{\mathrm{T}}} = \left[3 \quad 2\right]$, а точність обчислення $\varepsilon =$ $0{,}01$.

37. За допомогою методу покоординатного спуску визначити мінімальні витрати y на укладання бетону при будівництві критого басейну об'ємом 1000 *м*^{3} у формі паралелепіпеда, якщо вартість бетонування 1 *м*^{2} складає для основи басейну 14 ; для стін – 23; для перекриття – 41 *грн*.

Примітка. При вирішенні задачі ширину басейну поставити у відповідність змінній x_1 ; довжину – x_2 ; висоту – виразу $\dfrac{1}{x_1 x_2}$. За точку початкового наближення взяти $\mathbf{x}^{(0)^{\mathrm{T}}} = \begin{bmatrix} 10 & 10 \end{bmatrix}$. Задачу вирішити з точністю до копійки, тобто $\varepsilon = 0{,}01$.

5.7. Фонд індивідуальних завдань

Індивідуальне завдання №19. За допомогою методу найшвидшого спуску знайти мінімум функції $y = \dfrac{1}{2x_1^2} + 2\sqrt{x_1 x_2} + \dfrac{1}{2x_2^2}$ з точністю обчислення $\varepsilon = 0{,}01$, прийнявши за початкове наближення точку:

1. $\mathbf{x}^{(0)^{\mathrm{T}}} = \begin{bmatrix} -0{,}55 & -0{,}4 \end{bmatrix}$

2. $\mathbf{x}^{(0)^{\mathrm{T}}} = \begin{bmatrix} -0{,}3 & -0{,}7 \end{bmatrix}$

3. $\mathbf{x}^{(0)^{\mathrm{T}}} = \begin{bmatrix} -0{,}35 & 0{,}5 \end{bmatrix}$

4. $\mathbf{x}^{(0)^{\mathrm{T}}} = \begin{bmatrix} -0{,}3 & -0{,}4 \end{bmatrix}$

5. $\mathbf{x}^{(0)^{\mathrm{T}}} = \begin{bmatrix} -0{,}35 & -0{,}6 \end{bmatrix}$

6. $\mathbf{x}^{(0)^{\mathrm{T}}} = \begin{bmatrix} -0{,}4 & -0{,}6 \end{bmatrix}$

7. $\mathbf{x}^{(0)^{\mathrm{T}}} = \begin{bmatrix} -0{,}5 & -0{,}3 \end{bmatrix}$

8. $\mathbf{x}^{(0)^{\mathrm{T}}} = \begin{bmatrix} -0{,}5 & -0{,}4 \end{bmatrix}$

9. $\mathbf{x}^{(0)^{\mathrm{T}}} = \begin{bmatrix} -0{,}5 & -0{,}65 \end{bmatrix}$

10. $\mathbf{x}^{(0)^{\mathrm{T}}} = \begin{bmatrix} -0{,}5 & -0{,}6 \end{bmatrix}$

11. $\mathbf{x}^{(0)^{\mathrm{T}}} = \begin{bmatrix} -0{,}6 & 0{,}35 \end{bmatrix}$

12. $\mathbf{x}^{(0)^{\mathrm{T}}} = \begin{bmatrix} -0{,}6 & -0{,}7 \end{bmatrix}$

13. $\mathbf{x}^{(0)^{\mathrm{T}}} = \begin{bmatrix} 1{,}55 & 1{,}7 \end{bmatrix}$

14. $\mathbf{x}^{(0)^{\mathrm{T}}} = \begin{bmatrix} 1{,}5 & 1{,}75 \end{bmatrix}$

15. $\mathbf{x}^{(0)^{\mathrm{T}}} = \begin{bmatrix} 1{,}5 & 1{,}45 \end{bmatrix}$

16. $\mathbf{x}^{(0)^{\mathrm{T}}} = \begin{bmatrix} 1{,}5 & 1{,}6 \end{bmatrix}$

17. $\mathbf{x}^{(0)\,\mathrm{T}} = [1{,}5 \quad 1{,}65]$

18. $\mathbf{x}^{(0)\,\mathrm{T}} = [1{,}5 \quad 1{,}4]$

19. $\mathbf{x}^{(0)\,\mathrm{T}} = [1{,}4 \quad 1{,}65]$

20. $\mathbf{x}^{(0)\,\mathrm{T}} = [1{,}4 \quad 1{,}7]$

21. $\mathbf{x}^{(0)\,\mathrm{T}} = [1{,}4 \quad 1{,}55]$

22. $\mathbf{x}^{(0)\,\mathrm{T}} = [1{,}4 \quad 1{,}3]$

23. $\mathbf{x}^{(0)\,\mathrm{T}} = [1{,}35 \quad 1{,}7]$

24. $\mathbf{x}^{(0)\,\mathrm{T}} = [1{,}7 \quad 1{,}3]$

25. $\mathbf{x}^{(0)\,\mathrm{T}} = [1{,}35 \quad 1{,}6]$

26. $\mathbf{x}^{(0)\,\mathrm{T}} = [1{,}6 \quad 1{,}3]$

27. $\mathbf{x}^{(0)\,\mathrm{T}} = [1{,}35 \quad 1{,}25]$

28. $\mathbf{x}^{(0)\,\mathrm{T}} = [1{,}65 \quad 1{,}3]$

29. $\mathbf{x}^{(0)\,\mathrm{T}} = [1{,}5 \quad 1{,}7]$

30. $\mathbf{x}^{(0)\,\mathrm{T}} = [1{,}45 \quad 1{,}35]$

Індивідуальне завдання №20. Виконати індивідуальне завдання №19 за допомогою методу Ньютона.

Індивідуальне завдання №21. Виконати індивідуальне завдання №19 за допомогою методу покоординатного спуску.

Розділ 6. ОПТИМІЗАЦІЯ ПРИ ОБМЕЖЕННЯХ У ВИГЛЯДІ РІВНОСТЕЙ

Розглянута в розділах 4 і 5 безумовна оптимізація припускає відсутність яких-небудь умов на область визначення функції цілі. Вважається, що будь-яка крапка n-мірного евклідового простору, в якій функція існує, може бути шуканим рішенням. У задачі безумовної оптимізації на область зміни кожної змінної не накладаються додаткові обмеження (ні односторонні, ні двосторонні), а самі змінні ніяк не зв'язані між собою, тобто взаємно незалежні. Однак у більшості практичних задач ми натрапляємо на різні обмеження (умови), що роблять область задання функції цілі відмінної від області існування функції. У цьому випадку методи визначення оптимальних значень функцій помітно відрізняються від методів безумовної оптимізації.

У цьому розділі розглядається клас задач умовної оптимізації, коли на змінні задачі накладаються обмеження у вигляді рівностей. Такі задачі називають класичними задачами пошуку умовного екстремуму. Вони є окремим випадком задач умовної оптимізації. У загальному випадку на змінні задач можуть накладатися обмеження як у вигляді рівностей, так і нерівностей.

6.1. Постановка задачі оптимізації при обмеженнях у вигляді рівностей

Прикладом умовної оптимізації при обмеженнях у вигляді рівностей може бути задача проектування циліндричної місткості заданого об'єму з мінімальною боковою поверхнею. Така задача виникає в реальному виробництві при виготовленні місткостей для збереження рідких і сипучих матеріалів, коли виготовлювач хоче домогтися мінімальної витрати матеріалу.

Розглянемо постановку цієї задачі докладніше.

Нехай треба спроектувати циліндричну місткість заданого об'єму a. Спроектувати циліндричну місткість – це значить визначити її основні розміри: радіус основи і висоту. Позначимо через x_1 радіус основи циліндра, через x_2 – висоту циліндра. Тоді математична модель бокової поверхні буде визначатися функцією

$$y(\mathbf{x}) = 2\pi x_1 x_2 + 2\pi x_1^2 \ , \tag{6.1}$$

а математична модель об'єму циліндра – рівністю

$$\pi x_1^2 x_2 - a = 0 \ . \tag{6.2}$$

Ми перейшли від змістовної, або вербальної, постановки задачі до математичної, а саме: знайти мінімум функції (6.2), якщо на змінні задачі накладаються обмеження у вигляді рівності (6.2).

У цій задачі є дві змінні (x_1 і x_2) і одне обмеження-рівність (6.2), що строго зв'язує значення цих змінних. У загальному випадку класична задача оптимізації має n змінних і m обмежень-рівностей. При цьому $n > m$. У противному разі область припустимих рішень являтиме собою або порожню множину при $n < m$, або множину відособлених точок. У другому випадку задача умовної оптимізації зводиться до обчислення функції цілі в цих точках і вибору серед цих значень шляхом порівняння оптимального.

Загальна постановка задачі мінімізації при обмеженнях у вигляді рівностей формулюється таким чином: знайти мінімум функції $y(\mathbf{x})$, визначеної на області $\mathbf{W} \subset \mathbf{R}^n$, де область $\mathbf{W}$ (область припустимих рішень) формується як переріз області існування функції $\mathbf{V}$ і області рішень системи з m рівнянь.

Аналітичний запис цієї задачі такий:

$$y(\mathbf{x}) \to \min_{\mathbf{x} \in \mathbf{W} \subset \mathbf{R}^n} \ , \tag{6.3}$$

$$\mathbf{W}: \quad \mathbf{V} \cup \mathbf{\Omega} \ ; \tag{6.4}$$

$$\mathbf{\Omega}: \quad f_i(\mathbf{x}) = 0, \quad i = \overline{1, m}, \quad n < m \ . \tag{6.5}$$

У літературі з математики постановка задачі безумовної оптимізації надається у спрощеному вигляді – без вказівки приналежності точки $\mathbf{x}$ області існування функції $\mathbf{V}$:

$$y(\mathbf{x}) \to \min_{\mathbf{x} \in \Omega \subset \mathbf{R}^n}, \tag{6.6}$$

$$\mathbf{\Omega}: \quad f_i(\mathbf{x}) = 0, \quad i = \overline{1,m}, \quad n < m \; . \tag{6.7}$$

Факт приналежності точки $\mathbf{x}$ області $\mathbf{V}$ має значення тільки для вибору початкової точки наближення при пошуку рішення розглянутого класу задач некласичними (прямими) методами. Надалі будемо користуватися постановкою (6.6), (6.7), тобто будемо припускати, що функція $y(\mathbf{x})$ існує в кожній точці простору $\mathbf{R}^n$, або множина $\mathbf{\Omega}$ не містить точок, в яких функція $y(\mathbf{x})$ не існує.

Шукане рішення задачі мінімізації при обмеженнях у виді рівностей, на відміну від задачі безумовної оптимізації, завжди знаходиться на границі області припустимих рішень $\mathbf{\Omega}$, оскільки остання взагалі не має внутрішніх точок.

6.2. Залежні і незалежні змінні

Процес умовної оптимізації при обмеженнях у вигляді системи рівностей ґрунтується на концепції *залежних* і *незалежних* змінних, що уявляє собою розбивку вектора змінних на два підвектори:

$$\mathbf{x} = \begin{bmatrix} \mathbf{s} \\ \mathbf{t} \end{bmatrix}, \tag{6.8}$$

де $\mathbf{s}$ – підвектор незалежних змінних; $\mathbf{t}$ – підвектор залежних змінних.

Підвектор $\mathbf{s}$ має m компонент, саме стільки, скільки рівнянь входить у систему обмежень (6.7), а підвектор $\mathbf{t}$ має p компонент, де $p = n - m$.

Будемо вважати, що розбивка змінних задачі робиться таким чином:

$$\mathbf{s} = \begin{bmatrix} s_1 \\ s_2 \\ \cdots \\ s_m \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ \cdots \\ x_m \end{bmatrix}; \quad \mathbf{t} = \begin{bmatrix} t_1 \\ t_2 \\ \cdots \\ t_p \end{bmatrix} = \begin{bmatrix} x_{m+1} \\ x_{m+2} \\ \cdots \\ x_{m+p} \end{bmatrix}; \quad m + p = n. \tag{6.9}$$

Відзначимо, що будь-яку іншу розбивку завжди можна звести до (6.9) за рахунок відповідної перенумерації початкових змінних. Відзначимо також, що залежні змінні часто називають змінними *стану*, а незалежні – змінними *рішення*.

6.3. Метод підстановки

Найпростішим методом рішення задачі (6.6) – (6.7) є метод підстановки. Однак, відразу слід зазначити, що він не є універсальним і може бути застосований тільки за певних умов, про які буде сказано наприкінці даного підрозділу.

Задача мінімізації (6.6) – (6.7) з урахуванням розбивки змінних на залежні і незалежні набуває вигляду:

$$y(\mathbf{s}, \mathbf{t}) \to \min_{\begin{bmatrix} \mathbf{s} \\ \mathbf{t} \end{bmatrix} \in \mathbf{\Omega} \subset \mathbf{R}^n}, \tag{6.10}$$

$$\mathbf{\Omega}: \quad f_i(\mathbf{s}, \mathbf{t}) = 0, \quad i = \overline{1, m}, \quad n < m . \tag{6.11}$$

Метод підстановки полягає в перетворенні системи рівностей (6.11) у систему

$$s_i = \varphi_i(\mathbf{t}), \quad i = \overline{1, m} , \tag{6.12}$$

і підстановці у функцію цілі (6.10) замість залежних змінних s_i їхні вирази через підвектор незалежних змінних **t,** тобто (6.12). Така процедура підстановки приводить до того, що складна задача умовної мінімізації (6.6) – (6.7) зводиться до задачі безумовної мінімізації:

$$y_{\text{пр}}(\mathbf{t}) \to \min_{\mathbf{t}\in\mathbf{R}^{n-m}} \tag{6.13}$$

с числом змінних меншим, чим у початковій постановці задачі.

Покажемо використання методу підстановки на прикладі рішення задачі про проектування циліндричної місткості заданого об'єму з мінімальною боковою поверхнею.

Приклад 6.1.

$$y(\mathbf{x}) = 2\pi x_1 x_2 + 2\pi x_1^2 \to \min_{\mathbf{x}\in\mathbf{\Omega}\subset\mathbf{R}^2} , \tag{6.14}$$

$$\mathbf{\Omega}: \quad \pi x_1^2 x_2 - a = 0 . \tag{6.15}$$

Розв’язання. Розв’яжемо рівність (6.15) щодо змінної x_2 , тобто приведемо (6.15) до вигляду (6.12):

$$x_2 = \frac{a}{\pi x_1^2} . \tag{6.16}$$

Підставляючи (6.16) у (6.14), одержимо задачу безумовної мінімізації

$$y(x_1) = 2\pi\left(\frac{a}{\pi x_1} + x_1^2\right) \to \min_{x_1\in\mathbf{R}^1} , \tag{6.17}$$

яку можна вирішити за допомогою вже знайомого методу Ейлера.

Оптимальним рішенням задачі (6.17) є

$$x_1{}^* = \sqrt[3]{\frac{a}{2\pi}} . \tag{6.18}$$

Здійснюючи ще одну підстановку: (6.18) у (6.16), визначимо оптимальне значення другої змінної:

$$x_2{}^* = 2\sqrt[3]{\frac{a}{2\pi}} \;. \tag{6.19}$$

Шукане значення мінімальної поверхні циліндричної місткості заданого об'єму a, визначиться як

$$y^* = \frac{3}{2\pi}\sqrt[3]{2\pi\ a^2} \;. \tag{6.20}$$

Таким чином, відповіддю на поставлену задачу (6.14) – (6.15) буде вектор оптимальних розмірів циліндра $\overline{x}\,^{*\text{т}} = \left[\sqrt[3]{\frac{a}{2\pi}} \quad 2\sqrt[3]{\frac{a}{2\pi}}\right]$, при яких повна поверхня складе $y^* = \frac{3\sqrt[3]{2\pi\ a^2}}{2\pi}$. Знайдене рішення забезпечує максимальну економію матеріалу при виробництві циліндричних місткостей заданого об'єму.

Метод підстановки є дуже ефективним методом вирішення задачі оптимізації при обмеженнях у вигляді рівностей. Він не тільки зводить задачу умовної оптимізації до задачі безумовної оптимізації, а також і знижує порядок задачі. При цьому, чим більше обмежень, тим відчутніше зниження порядку, а отже, і відчутніше зниження складності рішення задачі.

Якщо система (6.11) являє собою систему лінійних рівнянь, то вона завжди може бути перетворена в систему (6.12), наприклад, за допомогою жорданових виключень. У загальному випадку система рівностей (6.11) є нелінійною. Її вирішення відносно окремих змінних, тобто зведення до вигляду (6.12), не завжди представляється можливим або вимагає надмірних зусиль. У таких випадках метод підстановки або не може бути застосованим, або не є ефективним. Як бачимо, при очевидній ефективності даний метод не має універсальність.

З метою одержання універсального класичного методу рішення задачі оптимізації при обмеженнях у вигляді рівностей (6.10) – (6.11) виведемо достатні і необхідні умови для точки локального мінімуму функції (6.10), які враховують обмеження задачі (6.11). Така точка називається умовним мінімумом, а задача типу (6.10) – (6.11) – задачею пошуку умовного екстремуму. У даному випадку – умовного мінімуму.

6.4. Необхідні умови для точки умовного локального мінімуму

Доведення необхідних і достатніх умов у задачі безумовної мінімізації дозволив розробити класичний метод її рішення – метод Ейлера. Вчинимо аналогічним способом і в задачі пошуку мінімуму функції при обмеженнях у вигляді рівностей. Доведення необхідних і достатніх умов також буде засновано на властивості невід’ємності приросту функції цілі в точці умовного екстремуму.

Дамо визначення умовним локальним екстремумам.

Визначення 6.1. Точка $\mathbf{x}^* \in \mathbf{\Omega} \subset \mathbf{R}^n$ називається точкою умовного локального мінімуму функції $y(\mathbf{x})$, якщо існує окіл $\varepsilon(\mathbf{x}^*) \in \mathbf{\Omega}$ точки $\mathbf{x}^*$ такий, що при усіх $\mathbf{x} \in \varepsilon(\mathbf{x}^*) \cap \mathbf{\Omega}$ виконується нерівність $y(\mathbf{x}^*) \leq y(\mathbf{x})$, де $\mathbf{x}$ – n-вимірний вектор змінних задачі.

Визначення 6.2. Точка $\mathbf{x}^* \in \mathbf{\Omega} \subset \mathbf{R}^n$ називається точкою умовного локального максимуму функції $y(\mathbf{x})$, якщо існує окіл $\varepsilon(\mathbf{x}^*) \in \mathbf{\Omega}$ точки $\mathbf{x}^*$ такий, що при усіх $\mathbf{x} \in \varepsilon(\mathbf{x}^*) \cap \mathbf{\Omega}$ виконується нерівність $y(\mathbf{x}^*) \geq y(\mathbf{x})$.

З визначень 6.1 – 6.2 випливає, що точки локальних екстремумів функції завжди є граничними точками області припустимих рішень $\boldsymbol{\Omega}$, оскільки остання взагалі не має жодної внутрішньої точки.

З визначення 6.1 випливає, що і при від'ємних, і при додатних приростах вектора змінних $\Delta\mathbf{x}$, що не виводять із області припустимих рішень $\boldsymbol{\Omega}$, приріст функції Δy^* в точці умовного локального мінімуму $\mathbf{x}^*$ не може бути від'ємним:

$$\Delta y^* = y\left(\mathbf{x}^* + \Delta\mathbf{x}\right) - y\left(\mathbf{x}^*\right) \geq 0 \ . \tag{6.21}$$

Отже, розкладемо функцію $y(\mathbf{x})$ в ряд Тейлора в приростах в околі точки $\mathbf{x} \in \boldsymbol{\Omega}$, враховуючи розбиття змінних на залежні і незалежні (6.8) і обмежуючи розкладання лінійними членами:

$$\Delta y = \left(\frac{\partial\ y}{\partial\ \mathbf{s}}\right)^{\mathrm{т}} \Delta\mathbf{s} + \left(\frac{\partial\ y}{\partial\ \mathbf{t}}\right)^{\mathrm{т}} \Delta\mathbf{t} \ . \tag{6.22}$$

Звернемо особливу увагу на приналежність точки розкладання області припустимих рішень

$$\mathbf{x} \in \boldsymbol{\Omega} \ . \tag{6.23}$$

У кожному околі точки $\mathbf{x} \in \boldsymbol{\Omega}$ є два типи точок: точки, що не належать області $\boldsymbol{\Omega}$, бо не виконується умова $\mathbf{f}(\mathbf{x}) = \mathbf{0}$, і точки, що належать їй, для яких остання умова виконується.

Розкладемо аналогічним способом у ряд Тейлора обмеження задачі:

$$\Delta\,\mathbf{f} = \left(\frac{\partial\ \mathbf{f}}{\partial\ \mathbf{s}}\right)^{\mathrm{т}} \Delta\mathbf{s} + \left(\frac{\partial\ \mathbf{f}}{\partial\ \mathbf{t}}\right)^{\mathrm{т}} \Delta\mathbf{t} = 0 \ . \tag{6.24}$$

Розкладання (6.24) дорівнює нулю, оскільки нас цікавлять тільки зміни функції цілі, що не виводять її з області припустимих рішень $\boldsymbol{\Omega}$, тобто не порушують умови (6.23).

Позначимо $\mathbf{W} = \left(\dfrac{\partial\ \mathbf{f}}{\partial\ \mathbf{s}} \right)$; $\mathbf{C} = \left(\dfrac{\partial\ \mathbf{f}}{\partial\ \mathbf{t}} \right)$.

> Матриця $\mathbf{W}$ має розмір $(m \times m)$ і називається матрицею стану. Її також називають матрицею Якобі або якобіаном.
>
> Матриця $\mathbf{C}$ має розмір $(m \times p)$ і називається матрицею управління.

У загальному випадку матриця Якобі і матриця управління мають відповідно вигляд:

$$\mathbf{W} = \begin{bmatrix} \dfrac{\partial f_1}{\partial s_1} & \dfrac{\partial f_1}{\partial s_2} & \dots & \dfrac{\partial f_1}{\partial s_m} \\ \dfrac{\partial f_2}{\partial s_1} & \dfrac{\partial f_2}{\partial s_2} & \dots & \dfrac{\partial f_2}{\partial s_m} \\ \dots & \dots & \dots & \dots \\ \dfrac{\partial f_m}{\partial s_1} & \dfrac{\partial f_m}{\partial s_2} & \dots & \dfrac{\partial f_m}{\partial s_m} \end{bmatrix}; \tag{6.25}$$

$$\mathbf{C} = \begin{bmatrix} \dfrac{\partial f_1}{\partial t_1} & \dfrac{\partial f_1}{\partial t_2} & \dots & \dfrac{\partial f_1}{\partial t_p} \\ \dfrac{\partial f_2}{\partial t_1} & \dfrac{\partial f_2}{\partial t_{21}} & \dots & \dfrac{\partial f_2}{\partial t_p} \\ \dots & \dots & \dots & \dots \\ \dfrac{\partial f_m}{\partial t_1} & \dfrac{\partial f_m}{\partial t_2} & \dots & \dfrac{\partial f_m}{\partial t_p} \end{bmatrix}. \tag{6.26}$$

Перша матриця складається з перших частинних похідних функцій обмежень по змінних стану (залежні перемінні), а друга – по змінних рішення (незалежні перемінні).

З урахуванням введених позначень розкладання (6.24) набуває вигляду:

$$\mathbf{W}\Delta\mathbf{s} + \mathbf{C}\Delta\mathbf{t} = 0 \ . \tag{6.27}$$

Будемо вважати, що розбиття змінних задачі зроблено таким чином, що матриця Якобі є неособливою. Ця умова необхідна для існування оберненої матриці Якобі, оскільки тільки неособлива матриця має обернену.

Помножимо (6.27) на обернену матрицю Якобі зліва:

$$\left(\mathbf{W}\right)^{-1}\mathbf{W}\Delta\mathbf{s} + \mathbf{W}^{-1}\mathbf{C}\Delta\mathbf{t} = 0 \ , \tag{6.28}$$

або

$$\Delta\mathbf{s} + \mathbf{W}^{-1}\mathbf{C}\Delta\mathbf{t} = 0 \ . \tag{6.29}$$

Розв’яжемо (6.29) відносно приросту вектора залежних змінних:

$$\Delta\mathbf{s} = -\mathbf{W}^{-1}\mathbf{C}\Delta\mathbf{t} \ . \tag{6.30}$$

Тут слід зазначити, що вираз (6.30) завжди може бути отриманий з (6.27) за допомогою жорданових виключень, оскільки (6.27) являє собою систему лінійних форм вигляду (4.1), де залежними змінними виступають прирости змінних стану, а незалежними – прирости змінних рішення. Це значить, що за допомогою жорданових виключень можна зробити такий вибір m залежних і p незалежних змінних $\left(m + p = n\right)$, що забезпечить неособливість матриці Якобі.

Треба також зазначити, що вираз (6.30) при довільній зміні p змінних рішення дозволяє єдиним способом визначити прирости змінних стану, що не виводять точку $\mathbf{x}$ за межі області припустимих рішень Ω. Якщо довільним способом змінювати більше, ніж p змінних, то гарантувати приналежність нової точки області Ω не можна. Число p називається числом ступенів свободи задачі. Кожне додаткове обмеження типу «рівність» зменшує ступінь свободи і число незалежних змінних на одиницю, спрощуючи тим самим проблему оптимізації.

Слід також зазначити, що вираз (6.30) позначає лінійну залежність приросту змінних стану від приросту змінних рішення. Якщо кожну i–у компоненту вектора $\mathbf{s}$ продиференціювати по кожній j–й незалежній змінній t_j, то в результаті прийдемо до матриці $\mathbf{B}$ розмірності $m \times p$:

$$\mathbf{B}=\left[\frac{\delta\, s_i}{\delta\, t_j}\right]=\begin{bmatrix}\frac{\delta\, s_1}{\delta\, t_1} & \frac{\delta\, s_1}{\delta\, t_2} & \cdots & \frac{\delta\, s_1}{\delta\, t_p}\\ \frac{\delta\, s_2}{\delta\, t_1} & \frac{\delta\, s_2}{\delta\, t_{21}} & \cdots & \frac{\delta\, s_2}{\delta\, t_p}\\ \cdots & \cdots & \cdots & \cdots\\ \frac{\delta\, s_m}{\delta\, t_1} & \frac{\delta\, s_m}{\delta\, t_2} & \cdots & \frac{\delta\, s_m}{\delta\, t_p}\end{bmatrix}=-\mathbf{W}^{-1}\mathbf{C},\quad i=\overline{1,m},\ j=\overline{1,p}\ . \tag{6.31}$$

Її елемент $\frac{\delta\, s_i}{\delta\, t_j}$ називається умовною похідною i–ї залежної змінної по j–й незалежній змінній або i–ю похідною стану по j–й змінній рішення. Вона визначає (при лінійній апроксимації) величину зміни i–ї залежної змінної, при зміні тільки однієї j–ї незалежної змінної на одиницю і являє собою частинну похідну неявної функції s_i по змінній t_j, якщо розглядати змінні $s_1, s_2, ..., s_m$ як неявні функції від змінних рішення $t_1, t_2, ..., t_m$, що задані системою рівнянь (6.11).

Підставимо (6.30) у (6.22):

$$\Delta y=-\left(\frac{\partial\, y}{\partial\, \mathbf{s}}\right)^{\mathrm{T}}\mathbf{W}^{-1}\mathbf{C}\Delta\mathbf{t}+\left(\frac{\partial\, y}{\partial\, \bar{t}}\right)^{\mathrm{T}}\Delta\mathbf{t}\ . \tag{6.32}$$

Приведемо подібні:

$$\Delta y=\left[\left(\frac{\partial\, y}{\partial\, \mathbf{t}}\right)^{\mathrm{T}}-\left(\frac{\partial\, y}{\partial\, \mathbf{s}}\right)^{\mathrm{T}}\mathbf{W}^{-1}\mathbf{C}\right]\Delta\mathbf{t}\ . \tag{6.33}$$

Вирази в квадратних дужках називають вектором-рядком *перших умовних похідних* функції y і позначають $\left(\frac{\delta\, y}{\delta\, \mathbf{t}}\right)^{\mathrm{T}}$:

$$\left(\frac{\delta\ y}{\delta\ \mathbf{t}}\right)^{\mathrm{T}} = \left(\frac{\partial\ y}{\partial\ \mathbf{t}}\right)^{\mathrm{T}} - \left(\frac{\partial\ y}{\partial\ \mathbf{s}}\right)^{\mathrm{T}} (\mathbf{W})^{-1}\mathbf{C}\ . \qquad (6.34)$$

Складову $\left(\frac{\delta\ y}{\delta\ t_j}\right)$ p-вимірного вектора (6.34) називають *першою умовною похідною* функції цілі y по j-й незалежній змінній. Вона показує (при лінійній апроксимації), як змінюється функція y при зміні незалежної змінної на одиницю.

Враховуючи позначення (6.34), розкладання (6.33) набуває вигляду

$$\Delta y = \left(\frac{\delta\ y}{\delta\ \mathbf{t}}\right)^{\mathrm{T}} \Delta\mathbf{t}\ . \qquad (6.35)$$

Відповідно до властивості точки умовного локального мінімуму (6.21), що випливає з визначення 6.1, приріст функції в ній не може бути від’ємним. Отже, в точці умовного локального мінімуму розкладання (6.35) повинно задовольняти нерівності

$$\Delta y^* = \left(\frac{\delta\ y}{\delta\ \mathbf{t}}\right)^{*\mathrm{T}} \Delta\mathbf{t} \ge 0\ . \qquad (6.36)$$

За умови зміни тільки однієї змінної останнє співвідношення набуває вигляду

$$\Delta y^* = \left(\frac{\delta\ y}{\delta\ t_r}\right)^{*} \Delta t_r \ge 0\ . \qquad (6.37)$$

Нерівність (6.37) повинна виконуватися при будь-яких приростах Δt_r як додатних, так і від'ємних. Це можливо тільки в єдиному випадку, коли $\left(\frac{\delta\ y}{\delta\ t_r}\right)^* = 0$ (див. висновок необхідних умов у задачі безумовної мінімізації).

Змінна вибиралася довільно, тому те саме можна сказати про будь-яку незалежну змінну . З урахуванням цього узагальнення, можна сказати, що в точці умовного локального мінімуму функції при обмеженнях у вигляді рівностей вектор перших умовних похідних повинен дорівнювати нулю

$$\boxed{\left(\frac{\delta\ y}{\delta\ \mathbf{t}}\right)^* = 0}\ . \qquad (6.38.)$$

Вираз (6.38) являє собою необхідну умову для точки локального мінімуму в задачі мінімізації функції при обмеженнях у вигляді рівностей.

Частинні похідні функції (6.13) по змінних рішення (незалежні змінні) збігаються з умовними похідними (6.34) і можуть бути отримані аналітично як похідні складної функції $y(\mathbf{t}, \varphi(\mathbf{t}))$:

$$\frac{\delta\ y}{\delta\ t_j} = \frac{\partial\ y}{\partial\ t_j} + \sum_{i=1}^{m} \frac{\partial\ y}{\partial \varphi_i} \frac{\partial \varphi_{\text{i}}}{\partial\ t_j} \qquad (j = \overline{1, p})\ , \qquad (6.39)$$

де функції $\varphi_i(\mathbf{t}), \quad i = \overline{1, m}$ визначаються системою виразів для залежних змінних (6.12).

Проілюструємо останнє твердження конкретним прикладом. Так, у задачі (6.14) – (6.15) частина похідна функції цілі (6.17) по змінній x_1 визначається виразом

$$\frac{\partial\ y(x_1)}{\partial\ x_1} = -\frac{2\pi}{x_1^2} + 4\pi x_1\ .$$

Такий же вираз ми одержуємо при визначенні умовної похідної функції $y(\mathbf{x}) = 2\pi x_1 x_2 + 2\pi x_1^2$ по змінній x_1 згідно з формулою (6.39). Тут

$$x_2 = \frac{a}{\pi\ x_1^2}\,;\quad \frac{\partial\ y}{\partial\ x_1} = 2\pi x_2 + 4\pi x_1\,;\qquad \frac{\partial\ y}{\partial\ x_1} = 2\pi x_1\,;$$

$$\frac{\partial\ x_2}{\partial\ x_1} = -\frac{2a}{\pi x_1^3}\,.$$

6.5. Достатні умови для точки умовного локального мінімуму

Необхідні умови (6.37) були отримані при лінійній апроксимації розкладання функції цілі $y(\mathbf{s},\mathbf{t})$ і обмежень $\mathbf{f}(\mathbf{s},\mathbf{t})$ у ряд Тейлора. Така апроксимація припустима, якщо члени розкладання другого і вище порядків занадто малі в порівнянні з лінійними членами. Однак у результаті виведення необхідних умов для точки локального мінімуму з'ясувалося, що лінійна складова розкладання дорівнює нулю. Отже, зневажати нелінійними членами розкладання при строгому доведенні умов наявності умовного екстремуму не можна.

Необхідна умова (6.37) дозволяє визначати умовні стаціонарні крапки. Вони можуть бути точками умовних локальних екстремумів, а можуть і не бути. Щоб перевірити умовну стаціонарну точку на наявність у ній умовного екстремуму, необхідно мати достатні умови екстремуму.

Для одержання достатніх умов введемо нове поняття *умовної похідної* функції цілі $\left(\dfrac{\delta\ y}{\delta\ f_j}\right)$ по i-му обмеженню-рівності, $i = \overline{1,m}$.

Припустимо, що кожне i-е обмеження-рівність у системі (6.7) може незначно відрізнятися від нуля. Цій відмінності поставимо у відповідність нову змінну f_i, тобто систему рівностей (6.7) подамо у вигляді

$$\mathbf{f}(\mathbf{x}) = \mathbf{f}, \tag{6.40}$$

де $\mathbf{f}(\mathbf{x})$ – вектор-функція обмежень, визначений лівою частиною системи (6.7), а $\mathbf{f}$ – вектор нових змінних, які можуть мати довільні значення.

Нехай $\mathbf{x}$ – точка в області припустимих рішень, яка визначена вектором-функцією $\mathbf{f}(\mathbf{x}) = 0$, а $\Delta\mathbf{f}$ – приріст вектора-функції $\mathbf{f}(\mathbf{x})$. Тоді рівняння (6.27) запишеться так:

$$\mathbf{W}\Delta\mathbf{s} + \mathbf{C}\Delta\mathbf{t} = \Delta\mathbf{f} \ . \tag{6.41}$$

Розв'яжемо (6.41) щодо вектора приросту залежних змінних

$$\Delta\mathbf{s} = \mathbf{W}^{-1}\Delta\mathbf{f} - \mathbf{W}^{-1}\mathbf{C}\Delta\mathbf{t} \ . \tag{6.42}$$

Підставимо (6.42) у (6.22):

$$\Delta y = \left(\frac{\partial y}{\partial \mathbf{s}}\right)^{\mathrm{T}} \mathbf{W}^{-1}\Delta\mathbf{f} + \left[\left(\frac{\partial y}{\partial \bar{t}}\right)^{\mathrm{T}} - \left(\frac{\partial y}{\partial \mathbf{s}}\right)^{\mathrm{T}} \mathbf{W}^{-1}\mathbf{C}\right]\Delta\mathbf{t} =$$

$$= \left(\frac{\delta y}{\delta \mathbf{f}}\right)^{\mathrm{T}} \Delta\mathbf{f} + \left(\frac{\delta y}{\delta \mathbf{t}}\right)^{\mathrm{T}} \Delta\mathbf{t} \ , \tag{6.43}$$

де

$$\left(\frac{\delta y}{\delta \mathbf{f}}\right)^{\mathrm{T}} = \left(\frac{\delta y}{\delta \mathbf{s}}\right)^{\mathrm{T}} \mathbf{W}^{-1} \tag{6.44}$$

являє собою m-вимірний вектор частинних похідних функції цілі по обмеженнях задачі. Кожна i-я складова цього вектора характеризує зміну функції y залежно від зміни i-го обмеження задачі (при незмінності незалежних змінних та інших обмежень) і є частинною похідною складної функції $y(\mathbf{t},\mathbf{f},\mathbf{s}(\mathbf{t},\mathbf{f}))$ щодо незалежній змінній f_i. Тут $\mathbf{s}(\mathbf{t},\mathbf{f})$ – вектор-функція, що неявно задана системою рівнянь (6.7), поданою у вигляді (6.40).

У точці умовного оптимуму $\mathbf{x}^*$ вектор умовних похідних $\left(\dfrac{\delta\, y}{\delta\, \mathbf{t}}\right)^* = 0$, і, отже, вираз (6.43) набуває більш простого вигляду:

$$\Delta y^* = \left(\frac{\delta\, y}{\delta\, \mathbf{f}}\right)^{*\mathrm{T}} \Delta\mathbf{f}\,. \tag{6.45}$$

Умовна похідна $\left(\dfrac{\delta\, y}{\delta\, f_i}\right)^*$, взята в точці мінімуму $\mathbf{x}^*$, визначає зміну оптимального значення функції залежно від зміни i-го обмеження і називається коефіцієнтом чутливості функції y по цьому обмеженню.

Коефіцієнти чутливості мають важливу властивість, яка полягає в тому, що їхні значення не залежать від вибору системи залежних і незалежних змінних. Незважаючи на те, що значення елементів оберненої матриці Якобі в рівнянні (6.44) частково визначаються характером розбивки змінних на залежні й незалежні, кожний коефіцієнт чутливості єдиний.

Повернемося до виведення достатніх умов наявності умовного локального мінімуму. Розкладемо функцію $y(\mathbf{x})$ в ряд Тейлора в приростах в околиці $\varepsilon(\mathbf{x})$ точки $\mathbf{x} \in \mathbf{\Omega}$:

$$\Delta y = \left(\frac{\partial\, y}{\partial\, \mathbf{x}}\right)^{\mathrm{T}} \Delta\mathbf{x} + \frac{1}{2}\Delta\mathbf{x}^{\mathrm{T}}\mathbf{H}\Delta\mathbf{x} + R\ , \tag{6.46}$$

де $\frac{1}{2}\Delta\mathbf{x}^{\mathrm{T}}\mathbf{H}\Delta\mathbf{x}$ – квадратичний член розкладання, записаний через матрицю $\mathbf{H}$ других частинних похідних функції $y(\mathbf{x})$; R – частина ряду розкладання, що складається з членів розкладання третього і більш високих порядків.

Аналогічним способом розкладемо кожне обмеження задачі:

$$\Delta f_i = \left(\frac{\partial f_i}{\partial \mathbf{x}}\right)^{\mathrm{T}} \Delta\mathbf{x} + \frac{1}{2}\Delta\mathbf{x}^{\mathrm{T}}\mathbf{H}_i\Delta\mathbf{x} + R_i = 0, \quad i = \overline{1,m}\ , \tag{6.47}$$

де $\mathbf{H}_i = \begin{bmatrix} \frac{\partial^2 f_i}{\partial x_1^2} & \frac{\partial^2 f_i}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f_i}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f_i}{\partial x_2 \partial x_1} & \frac{\partial^2 f_i}{\partial x_2^2} & \cdots & \frac{\partial^2 f_i}{\partial x_2 \partial x_n} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial^2 f_i}{\partial x_n \partial x_1} & \frac{\partial^2 f_i}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f_i}{\partial x_n^2} \end{bmatrix}$ – матриця других частинних похідних i-го обмеження задачі $f_i(\mathbf{x})$; R_i – частина ряду розкладання i-го обмеження, що складається з членів розкладання третього і більш високих порядків.

Помножимо приріст кожного i-го обмеження (6.47) на відповідну умовну похідну $\left(\frac{\delta y}{\delta f_i}\right)$, позначену для скорочення через λ_i, і віднімемо послідовно з рівняння (6.46):

$$\Delta y = \left[\frac{\partial}{\partial \mathbf{x}}\left(y - \boldsymbol{\lambda}^{\mathrm{T}}\mathbf{f}\right)\right]^{\mathrm{T}} \Delta\mathbf{x} + \frac{1}{2}\Delta\mathbf{x}^{\mathrm{T}}\left(\mathbf{H} - \sum_{i}^{m}\lambda_i\mathbf{H}_i\right)\Delta\mathbf{x} + R - \sum_{i=1}^{m}\lambda_i R_i\,. \tag{6.48}$$

Вважаючи, що

$$\lim_{\Delta x \to 0} \frac{\left[\frac{\partial}{\partial \mathbf{x}}\left(y - \boldsymbol{\lambda}^{\mathrm{T}}\mathbf{f}\right)\right]^{\mathrm{T}} \Delta\mathbf{x} + \frac{1}{2}\Delta\mathbf{x}^{\mathrm{T}}\left(\mathbf{H} - \sum_{i}^{m}\lambda_i \mathbf{H}_i\right)\Delta\mathbf{x}}{R - \sum_{i=1}^{m}\lambda_i R_i} = 0,$$

і вводячи позначення

$$\mathbf{P} = \mathbf{H} - \sum_{i}^{m}\lambda_i \mathbf{H}_i \quad , \tag{6.49}$$

одержимо компактну форму запису розкладання (6.48)

$$\Delta y = \left[\frac{\partial}{\partial \mathbf{x}}\left(y - \boldsymbol{\lambda}^{\mathrm{T}}\mathbf{f}\right)\right]^{\mathrm{T}} \Delta\mathbf{x} + \frac{1}{2}\Delta\mathbf{x}^{\mathrm{T}}\mathbf{P}\Delta\mathbf{x} \ . \tag{6.50}$$

Необхідні в процесі обчислення матриці $\mathbf{P}$ коефіцієнти λ_i підраховуємо за формулою (6.44), тобто вони є складовими вектора

$$\boldsymbol{\lambda}^{\mathrm{T}} = \left(\frac{\delta\ y}{\delta\ \mathbf{f}}\right)^{\mathrm{T}} = \left(\frac{\partial\ y}{\partial\ \mathbf{s}}\right)^{\mathrm{T}} \mathbf{W}^{-1} \quad . \tag{6.51}$$

Враховуючі розбивку змінних на залежні й незалежні $\mathbf{x} = \begin{bmatrix} \mathbf{s} \\ \mathbf{t} \end{bmatrix}$ і розкладаючи матрицю P на чотири підматриці відповідно до розбивки змінних, перетворимо (6.50)

$$\Delta y = \left\{\left[\frac{\partial}{\partial \mathbf{s}}\left(y - \boldsymbol{\lambda}^{\mathrm{T}}\mathbf{f}\right)\right]^{\mathrm{T}} + \left[\frac{\partial}{\partial \mathbf{t}}\left(y - \boldsymbol{\lambda}^{\mathrm{T}}\mathbf{f}\right)\right]^{\mathrm{T}}\right\}\begin{bmatrix} \Delta\mathbf{s} \\ \Delta\mathbf{t} \end{bmatrix} +$$

$$+ \frac{1}{2}\begin{bmatrix} \Delta\mathbf{s}^{\mathrm{T}} & \Delta\mathbf{t}^{\mathrm{T}} \end{bmatrix}\begin{bmatrix} \mathbf{P}_{ss} & \mathbf{P}_{st} \\ \mathbf{P}_{ts} & \mathbf{P}_{tt} \end{bmatrix}\begin{bmatrix} \Delta\mathbf{s} \\ \Delta\mathbf{t} \end{bmatrix} =$$

$$=\left[\frac{\partial y}{\partial \mathbf{s}}-\boldsymbol{\lambda}^{\mathrm{T}}\frac{\partial \mathbf{f}}{\partial \mathbf{s}}\right]^{\mathrm{T}}\Delta\mathbf{s}+\left[\frac{\partial y}{\partial \mathbf{t}}-\boldsymbol{\lambda}^{\mathrm{T}}\frac{\partial \mathbf{f}}{\partial \mathbf{t}}\right]^{\mathrm{T}}\Delta\mathbf{t}+$$

$$+\frac{1}{2}\left[\Delta\mathbf{s}^{\mathrm{T}}\mathbf{P}_{ss}\Delta\mathbf{s}+\Delta\mathbf{s}^{\mathrm{T}}\mathbf{P}_{st}\Delta\mathbf{t}+\Delta\mathbf{t}^{\mathrm{T}}\mathbf{P}_{ts}\Delta\mathbf{s}+\Delta\mathbf{t}^{\mathrm{T}}\mathbf{P}_{tt}\Delta\mathbf{t}\right]. \qquad (6.52)$$

Використовуючи раніше введені позначення $\mathbf{W}$ і $\mathbf{C}$ для матриць похідних функцій обмежень за залежними і незалежними змінними, одержуємо

$$\Delta y=\left[\frac{\partial y}{\partial \mathbf{s}}-\boldsymbol{\lambda}^{\mathrm{T}}W\right]^{\mathrm{T}}\Delta\mathbf{s}+\left[\frac{\partial y}{\partial \mathbf{t}}-\boldsymbol{\lambda}^{\mathrm{T}}\mathbf{C}\right]^{\mathrm{T}}\Delta\mathbf{t}+$$

$$+\frac{1}{2}\left[\Delta\mathbf{s}^{\mathrm{T}}\mathbf{P}_{ss}\Delta\mathbf{s}+\Delta\mathbf{s}^{\mathrm{T}}\mathbf{P}_{st}\Delta\mathbf{t}+\Delta\mathbf{t}^{\mathrm{T}}\mathbf{P}_{ts}\Delta\mathbf{s}+\Delta\mathbf{t}^{\mathrm{T}}\mathbf{P}_{tt}\Delta\mathbf{t}\right]=$$

$$=\left[\frac{\partial y}{\partial \mathbf{s}}\right]^{\mathrm{T}}\Delta\mathbf{s}+\left[\frac{\partial y}{\partial \mathbf{t}}\right]^{\mathrm{T}}\Delta\mathbf{t}-\boldsymbol{\lambda}^{\mathrm{T}}\left(\mathbf{W}\Delta\mathbf{s}+\mathbf{C}\Delta\mathbf{t}\right)+$$

$$+\frac{1}{2}\left[\Delta\mathbf{s}^{\mathrm{T}}\mathbf{P}_{ss}\Delta\mathbf{s}+\Delta\mathbf{s}^{\mathrm{T}}\mathbf{P}_{st}\Delta\mathbf{t}+\Delta\mathbf{t}^{\mathrm{T}}\mathbf{P}_{ts}\Delta\mathbf{s}+\Delta\mathbf{t}^{\mathrm{T}}\mathbf{P}_{tt}\Delta\mathbf{t}\right]. \qquad (6.53)$$

Враховуючі рівність (6.27) і підставляючи значення вектора $\Delta\mathbf{s}$ з рівняння (6.30) у (6.53), маємо

$$\Delta y=\left[\left(\frac{\partial y}{\partial \mathbf{s}}\right)^{\mathrm{T}}-\left(\frac{\partial y}{\partial \mathbf{t}}\right)^{\mathrm{T}}\mathbf{W}^{-1}\mathbf{C}\right]\Delta\mathbf{t}+$$

$$+\frac{1}{2}\Delta\mathbf{t}^{\mathrm{T}}\left[\left(\mathbf{W}^{-1}\mathbf{C}\right)^{\mathrm{T}}\mathbf{P}_{ss}\mathbf{W}^{-1}\mathbf{C}-\left(\mathbf{W}^{-1}\mathbf{C}\right)^{\mathrm{T}}\mathbf{P}_{st}-\mathbf{P}_{ts}\mathbf{W}^{-1}\mathbf{C}+\mathbf{P}_{tt}\right]\Delta\mathbf{t}\ . \qquad (6.54)$$

Внаслідок симетричності матриць $\mathbf{H}$ і $\mathbf{H}_i$ з (6.49) випливає симетричність матриці $\mathbf{P}$, це значить, що

$$\mathbf{P}_{st} = \left(\mathbf{P}_{ts}\right)^{\mathrm{T}} \ . \qquad (6.55).$$

З урахуванням (6.34) і (6.55) розкладання (6.54) набуває вигляду

$$\Delta y = \left(\frac{\delta y}{\delta \mathbf{t}}\right)^{\mathrm{T}} \Delta\mathbf{t} +$$

$$+\frac{1}{2}\Delta\mathbf{t}^{\mathrm{T}}\left[\mathbf{P}_{tt} - \mathbf{P}_{ts}\mathbf{W}^{-1}\mathbf{C} - \left(\mathbf{P}_{ts}\mathbf{W}^{-1}\mathbf{C}\right)^{\mathrm{T}} + \left(\mathbf{W}^{-1}\mathbf{C}\right)^{\mathrm{T}}\mathbf{P}_{ss}\mathbf{W}^{-1}\mathbf{C}\right]\Delta\mathbf{t} \ . \qquad (6.56)$$

Вираз у квадратних дужках в (6.56) являє собою матрицю *других умовних похідних* цільової функції по незалежних змінних $\left[\dfrac{\delta^2 y}{\delta t_i \delta t_j}\right]$, $i, j = \overline{1, p}$. Введемо позначення

$$\mathbf{S} = \mathbf{P}_{tt} - \mathbf{P}_{ts}\mathbf{W}^{-1}\mathbf{C} - \left(\mathbf{P}_{ts}\mathbf{W}^{-1}\mathbf{C}\right)^{\mathrm{T}} + \left(\mathbf{W}^{-1}\mathbf{C}\right)^{\mathrm{T}}\mathbf{P}_{ss}\mathbf{W}^{-1}\mathbf{C}. \qquad (6.57)$$

Тоді розкладання функції цілі в точці, що належить області припустимих рішень, набуває остаточного вигляду

$$\Delta y = \left(\frac{\delta y}{\delta \mathbf{t}}\right)^{\mathrm{T}} \Delta\mathbf{t} + \frac{1}{2}\Delta\mathbf{t}^{\mathrm{T}}\mathbf{S}\Delta\mathbf{t} \ . \qquad (6.58)$$

У точці умовного локального мінімуму $\mathbf{x}^*$ вектор перших умовних похідних відповідно до необхідних умов наявності екстремуму, дорівнює нулю. Отже, розкладання (6.58) в цьому випадку набуває більш простого вигляду

$$\Delta y^* = \frac{1}{2} \Delta \mathbf{t}^{\mathrm{T}} \mathbf{S}^* \Delta \mathbf{t} \ . \qquad (6.59)$$

Відповідно до умови локального мінімуму (6.21)

$$\Delta y^* = \frac{1}{2} \Delta \mathbf{t}^{\mathrm{T}} \mathbf{S}^* \Delta \mathbf{t} \geq 0 \ . \qquad (6.60)$$

З (6.60) безпосередньо випливає, що достатньою умовою для точки умовного локального мінімуму $\mathbf{x}^*$ є додатна визначеність у ній матриці других умовних похідних $\mathbf{S}^*$, тобто коли умовна диференціальна квадратична форма (6.60) строго більше нуля. Тільки в цьому випадку можна зневажати залишком $R - \sum_{i=1}^{m} \lambda_i R_i$ у (6.48).

☞

> Отже, необхідні умови локального мінімуму полягають у додатній визначеності матриці других умовних похідних $\mathbf{S}^*$, обумовленої виразом (6.57).

Другі частинні похідні функції (6.13) за змінними рішення (незалежні змінні) збігаються з другими умовними похідними, визначеними матрицею (6.57) і можуть бути отримані аналітично як похідні складної функції $y(\mathbf{t}, \varphi(\mathbf{t}))$:

$$\frac{\delta^2 y}{\delta t_r \delta t_p} = \frac{\partial^2 y}{\partial t_r \partial t_p} + \sum_{i=1}^{m} \left[\frac{\partial^2 y}{\partial t_r \partial s_i} \frac{\partial s_i}{\partial t_p} + \frac{\partial^2 y}{\partial t_p \partial s_i} \frac{\partial s_i}{\partial t_r} + \frac{\partial y}{\partial s_i} \frac{\partial^2 s_i}{\partial t_r \partial t_p} \right] +$$

$$+ \sum_{i=1}^{m} \sum_{j=1}^{m} \left[\frac{\partial^2 y}{\partial s_i \partial s_j} \frac{\partial s_i}{\partial t_r} \frac{\partial s_j}{\partial t_p} \right], \quad r, p = \overline{1, p} \, . \qquad (6.61)$$

6.6. Метод умовних похідних

Вирішення задачі (6.10) – (6.11) може бути отримане як за допомогою класичних, так і наближених методів.

Процес рішення класичними методами припускає два етапи. На першому етапі визначають усі умовно стаціонарні точки, функції цілі, на другому – перевіряють ці точки на наявність у них умовного екстремуму.

Одним із класичних методів визначення умовно стаціонарних точок у задачі (6.10) – (6.11) є метод, заснований на визначенні вектора умовних похідних. Метод полягає в розв’язанні системи (у загальному випадку нелінійної), що складається з n рівнянь, серед яких m визначаються системою обмежень задачі (6.11) , а інші $p = n - m$ – необхідними умовами наявності умовного екстремуму:

$$\begin{cases} f_i(\mathbf{s},\mathbf{t}) = 0, \quad i = \overline{1,m}; \\ \dfrac{\delta \ y}{\delta \ t_j} = 0, \quad j = \overline{1,p} \end{cases} \tag{6.62}$$

Розглянемо метод умовних похідних на прикладі вирішення задачі про мінімальну циліндричну поверхню

$$y(x_1, x_2) = 2\pi\, x_1 x_2 + 2\pi\, x_1^2 \to \min_{\mathbf{x}\in\mathbf{\Omega}},$$

$$\mathbf{\Omega}: \ \pi\, x_1^2 x_2 - a = 0.$$

Через малу розмірність задачі основні вектори й матриці, що беруть участь у розв’язанні, будуть складатися з одного елемента.

Одержання системи рівнянь (6.62) в умовах прикладу і її розв’язання здійснюється в нижчеприведеній послідовності.

➭ Розбивка змінних на залежні й незалежні:

$$\mathbf{s} = [s_1] = [x_1]\ ; \quad \mathbf{t} = [t_1] = [x_2].$$

➯ Формування вектора частинних похідних функції, що мінімізується , за змінними стану:

$$\frac{\partial y}{\partial \mathbf{s}} = \left[\frac{\partial y}{\partial s_1}\right] = [4\pi\, x_1 + 2\pi\, x_2]\ .$$

➯ Формування вектора частинних похідних функції, що мінімізується , за змінними рішення:

$$\frac{\partial y}{\partial \mathbf{t}} = \left[\frac{\partial y}{\partial t_1}\right] = [2\pi\, x_1]\ .$$

➯ Формування матриці Якобі:

$$\mathbf{W} = \left[\frac{\partial \mathbf{f}}{\partial \mathbf{s}}\right] = \left[\frac{\partial f_1}{\partial s_1}\right] = [2\pi\, x_1 x_2]\ .$$

➯ Обертання якобіана:

$$\mathbf{W}^{-1} = \left[\frac{1}{2\pi x_1 x_2}\right].$$

➯ Формування матриці керування:

$$\mathbf{C} = \left[\frac{\partial f}{\partial \bar{t}}\right] = \left[\frac{\partial f_1}{\partial t_1}\right] = \left[\pi x_1^2\right]\ .$$

➯ Визначення вектора умовних похідних**:**

$$\left(\frac{\delta\ y}{\delta\ \mathbf{t}}\right)^{\mathrm{T}} = \left(\frac{\partial\, y}{\partial\ \mathbf{t}}\right)^{\mathrm{T}} - \left(\frac{\partial\, y}{\partial\ \mathbf{s}}\right)^{\mathrm{T}} (\mathbf{W})^{-1}\mathbf{C} =$$

$$= \left[2\pi x_1\right] - \left[4\pi x_1 + 2\pi x_2\right] \left[\frac{1}{2\pi x_1 x_2}\right] \left[\pi x_1^2\right] =$$

$$2\pi\, x_1 - 2\pi\,(2x_1 + x_2)\frac{x_1}{x_2}.$$

Укладання системи рівнянь вигляду (6.62) в умовах прикладу:

$$\begin{cases} \pi\, x_1^2 x_2 - a = 0; \\ 2\pi\, x_1 - \pi\,(2x_1 + x_2)\dfrac{x_1}{x_2} = 0. \end{cases}$$

Вирішення системи рівнянь:

$$x_1 = \sqrt[3]{\frac{a}{2\pi}}\ ; \quad x_2 = 2\sqrt[3]{\frac{a}{2\pi}}\,.$$

Як бачимо, рішення задачі, отримане методом умовних похідних, збігається з рішенням її, отриманим методом підстановки. Звичайно, вирішення даної задачі на основі визначення умовних похідних не є доцільним, оскільки шукане рішення досягається набагато простіше методом підстановки. Розгляд методу умовних похідних у цьому підрозділі носить чисто ілюстративний характер.

Метод умовних похідних часто називають методом якобіана або методом Якобі.

У процесі вирішення задачі мінімізації (6.10) – (6.11) методом умовних похідних необхідно робити багато обчислень. Великий обсяг обчислювальних операцій – основний недолік цього методу. В інженерній практиці для вирішення задачі (6.10) – (6.11) користуються іншим класичним методом – методом *невизначених множників Лагранжа.*

6.7. Метод невизначених множників

Відповідно до виразу (6.45) приріст функції цілі Δy^* поблизу точки оптимуму можна виразити через вектор коефіцієнтів $\left(\dfrac{\delta\, y}{\delta\, \mathbf{f}}\right)$ чутливості і вектор приросту $\Delta \mathbf{f}^*$ обмежень. Якщо скористатися введеним раніше позначенням $\boldsymbol{\lambda} = \left(\dfrac{\delta\, y}{\delta\, \mathbf{f}}\right)$, то співвідношення (6.45) набуває вигляду

$$\Delta y^* = \boldsymbol{\lambda}^{\mathrm{T}} \Delta \mathbf{f}^* . \qquad (6.63)$$

Тут вектор коефіцієнтів чутливості $\boldsymbol{\lambda}$ розглядається як вектор із m нових змінних, що називають невизначеними множниками або множниками Лагранжа.

Диференціюючи (6.63) по кожній з n незалежних змінних x_{j} $\left(j = \overline{1,n}\right)$ і переносячи праву частину рівності в ліву, одержимо

$$\left(\frac{\partial\, y}{\partial\, \mathbf{x}}\right)^{*\mathrm{T}} - \boldsymbol{\lambda}^{\mathrm{T}} \left(\frac{\partial\, \mathbf{f}}{\partial\, \mathbf{x}}\right)^{*} = 0 \ ,$$

або

$$\frac{\partial}{\partial\, \mathbf{x}}\left(y^* - \boldsymbol{\lambda}^{\mathrm{T}} \mathbf{f}\right) = 0 \ . \qquad (6.64)$$

Тут всередині дужок – функція Лагранжа

$$L\left(\mathbf{x}, \boldsymbol{\lambda}\right) = y\left(\mathbf{x}\right) - \boldsymbol{\lambda}^{\mathrm{T}} \mathbf{f}\left(\mathbf{x}\right) , \qquad (6.65)$$

Вектор похідних функції Лагранжа за змінними задачі x_j $\left(j = \overline{1,n}\right)$ в точці умовного екстремуму $\mathbf{x}^*$ відповідно до (6.64) дорівнює нулю

$$\left(\frac{\partial L}{\partial \mathbf{x}}\right)^* = 0 \ . \tag{6.66}$$

Якщо взяти похідні від функції Лагранжа по новим змінним λ_i $\left(i = \overline{1,m}\right)$, то вони дорівнюють відповідним обмежуючим функціям

$$\frac{\partial L}{\partial \lambda_i} = -f_i(\mathbf{x}), \quad i = \overline{1,m} \ ,$$

або

$$\frac{\partial L}{\partial \boldsymbol{\lambda}} = -\mathbf{f}(\mathbf{x}). \tag{6.67}$$

В області припустимих рішень $\boldsymbol{\Omega}$, відповідно до виразу (6.7), останнє співвідношення повинно дорівнювати нулю

$$\frac{\partial L}{\partial \boldsymbol{\lambda}} = -\mathbf{f}(\mathbf{x}) = 0 \ . \tag{6.68}$$

З рівнянь (6.66) і (6.68) випливає, що будь-яка умовно стаціонарна точка задачі (6.6) – (6.7) є стаціонарною точкою функції Лагранжа (6.65). Спільна система (6.66) і (6.68) має $n+m$ рівнянь із n невідомими x_j і m невідомими λ_i, яка, у принципі, розв'язна щодо $\mathbf{x}^*$ і $\boldsymbol{\lambda}$. Отже, вирішення спільної системи (6.66) і (6.68) дозволяє визначити умовно стаціонарні точки $\mathbf{x}^*$ задачі (6.6) – (6.7). Попутно в процесі вирішення методом Лагранжа ми одержуємо вектори обчислених невизначених множників $\boldsymbol{\lambda}$ (множників Лагранжа) для кожної умовно стаціонарної точки $\mathbf{x}^*$. Однак ці вектори є побічними результатами, які в кінцевому рішенні можуть бути опущені, оскільки не є шуканими величинами в початковій задачі (6.6) – (6.7).

Отже вирішення задачі мінімізації функції $y(x)$ при обмеженнях у вигляді рівностей $\mathbf{f(x)} = 0$ методом невизначених множників, або методом Лагранжа, полягає в побудові функції Лагранжа $L(\mathbf{x},\boldsymbol{\lambda}) = y - \lambda^{\mathrm{T}}\mathbf{f(x)}$ за допомогою введення вектора додаткових змінних $\boldsymbol{\lambda}$ і вирішенні системи рівнянь

$$\begin{cases} \dfrac{\partial\, L(\mathbf{x},\boldsymbol{\lambda})}{\partial \mathbf{x}} = 0; \\ \mathbf{f}(\mathbf{x}) = 0\ . \end{cases} \tag{6.69}$$

Продемонструємо метод на прикладі вирішення вже добре знайомої задачі про мінімальну циліндричну поверхню:

$$y(x_1, x_2) = 2\pi\, x_1 x_2 + 2\pi\, x_1^2 \to \min_{\mathbf{x}\in\boldsymbol{\Omega}},$$
$$\boldsymbol{\Omega}:\quad f(\mathbf{x}) = \pi\, x_1^2 x_2 - a = 0.$$

В умовах цієї задачі її вирішення методом Лагранжа здійснюється в нижче наведеній послідовності.

Насамперед, формуємо функцію Лагранжа вигляду (6.65):

$$L(\mathbf{x},\boldsymbol{\lambda}) = 2\pi\, x_1 x_2 + 2\pi\, x_1^2 - \lambda\left(\pi\, x_1^2 x_2 - a\right).$$

Потім складаємо систему рівнянь вигляду (6.69):

$$\begin{cases} \dfrac{\partial L(\mathbf{x},\boldsymbol{\lambda})}{x_1} = 2\pi\, x_2 + 4\pi\, x_1 - \lambda\, 2\pi\, x_1 x_2 = 0\,; \\ \dfrac{\partial L(\mathbf{x},\boldsymbol{\lambda})}{\partial\, x_2} = 2\pi\, x_1 - \lambda\, 2\pi\, x_1^2 = 0\,; \\ \mathbf{f}(\mathbf{x}) = \left(\pi\, x_1^2 x_2 - a\right) = 0. \end{cases} \tag{6.70}$$

Вирішуємо систему (6.70) за допомогою алгебраїчних перетворень:

$$x_1 = \sqrt[3]{\frac{a}{2\pi}}\ ; \quad x_2 = 2\sqrt[3]{\frac{a}{2\pi}}\ ; \quad \lambda = 2\sqrt[3]{\frac{2\pi}{a}}\ .$$

Таким чином, задача має одну умовно стаціонарну точку $\mathbf{x}^{\circ\,\mathrm{T}} = \left[\sqrt[3]{\frac{a}{2\pi}} \quad 2\sqrt[3]{\frac{a}{2\pi}} \right]$.

Визначення умовно стаціонарних точок за допомогою класичних методів має істотний недолік, пов'язаний із вирішенням систем нелінійних рівнянь. При великій розмірності задачі і неможливості подання рішення у явному вигляді використовують чисельні методи вирішення.

6.8. Аналіз умовно стаціонарних точок на наявність умовного екстремуму

Розглянуті класичні методи вирішення задачі (6.6) – (6.7) дозволяють визначати тільки стаціонарні точки. Щоб гарантовано відібрати серед умовно стаціонарних точок умовно екстремальні, необхідно перевірити достатні умови наявності умовного екстремуму в кожній стаціонарній точці. Як ми встановили, достатні умови полягають у додатній визначеності матриці других умовних похідних функції цілі по змінним рішення.

Проведемо аналіз умовно стаціонарної точки на виконання достатніх умов у задачі визначення мінімальної поверхні циліндричної місткості. Нагадаємо, що умовно стаціонарна точка в цій задачі єдина і відповідає вектору $\mathbf{x}^{\circ\,\mathrm{T}} = \left[\sqrt[3]{\frac{a}{2\pi}} \quad 2\sqrt[3]{\frac{a}{2\pi}} \right]$.

Аналіз умовно стаціонарній точки на наявність умовного екстремуму проводиться в такій послідовності:

↬ Розбивка змінних на залежні й незалежні[2] :

$$\mathbf{s} = [s_1] = [x_1] \ ; \quad \mathbf{t} = [t_1] = [x_2].$$

↬ Формування векторів перших частинних похідних функції цілі за змінними стану і рішення:

$$\frac{\partial y}{\partial \mathbf{s}} = \left[\frac{\partial y}{\partial s_1}\right] = \left[\frac{\partial y}{\partial x_1}\right] = [4\pi\, x_1 + 2\pi\, x_2] \ ;$$

$$\frac{\partial y}{\partial \mathbf{t}} = \left[\frac{\partial y}{\partial t_1}\right] = \left[\frac{\partial y}{\partial x_2}\right] = [2\pi\, x_1] \ .$$

↬ Обчислення перших частинних похідних у стаціонарній точці $\mathbf{x}^\circ$:

$$\left(\frac{\partial y}{\partial \mathbf{s}}\right)^\circ = \left[4\sqrt[3]{4\pi^2 a}\right]; \qquad \left(\frac{\partial y}{\partial \mathbf{t}}\right)^\circ = \left[\sqrt[3]{4\pi^2 a}\right].$$

↬ Формування матриці Якобі і матриці управління:

$$\mathbf{W} = \left[\frac{\partial \mathbf{f}}{\partial \mathbf{s}}\right] = \left[\frac{\partial f_1}{\partial s_1}\right] = \left[\frac{\partial f_1}{\partial x_1}\right] = [2\pi x_1 x_2] \ ;$$

$$\mathbf{C} = \left[\frac{\partial \mathbf{f}}{\partial \mathbf{t}}\right] = \left[\frac{\partial f_1}{\partial t_1}\right] = \left[\frac{\partial f_1}{\partial x_2}\right] = [\pi\, x_1^2] \ .$$

↬ Обчислення матриці Якобі і матриці управління в точці $\mathbf{x}^\circ$:

[2] У зв'язку з малою розмірністю задачі вектори залежних і незалежних змінних складаються з однієї компоненти.

$$\mathbf{W} = \left[2\sqrt[3]{2\pi a^2}\right] \quad ; \qquad \mathbf{C} = \left[\frac{1}{2}\sqrt[3]{2\pi a^2}\right].$$

↳ Обертання якобіана й обчислення добутку у стаціонарній точці $\mathbf{x}^{\circ}$:

$$\mathbf{W}^{-1} = \left[\frac{1}{2\sqrt[3]{2\pi a^2}}\right]; \quad .$$

↳ Обчислення коефіцієнтів чутливості відповідно до виразу (6.51)

$$\lambda = [\lambda_1] = \left(\frac{\partial y}{\partial \mathbf{s}}\right)^{\text{т}} \mathbf{W}^{-1} = 2\sqrt[3]{2\frac{\pi}{a}} \quad .$$

↳ Формування матриць других частинних похідних цільової та обмежуючої функцій задачі:

$$\mathbf{H} = \begin{bmatrix} 4\pi & 2\pi \\ 2\pi & 0 \end{bmatrix}; \quad \mathbf{H}_1 = \begin{bmatrix} 2\pi x_2 & 2\pi x_1 \\ 2\pi x_1 & 0 \end{bmatrix}.$$

↳ Обчислення елементів матриць других частинних похідних цільової та обмежуючої функцій задачі в стаціонарній точці $\mathbf{x}^{\circ}$:

$$\mathbf{H} = \begin{bmatrix} 4\pi & 2\pi \\ 2\pi & 0 \end{bmatrix}; \quad \mathbf{H}_1 = \begin{bmatrix} 2\sqrt[3]{4\pi^2 a} & \sqrt[3]{4\pi^2 a} \\ \sqrt[3]{4\pi^2 a} & 0 \end{bmatrix}.$$

↳ Формування матриці **P**:

$$\mathbf{P} = \mathbf{H} - \sum_{i=1}^{m} \lambda_i \mathbf{H}_i = \mathbf{H} - \lambda_1 \mathbf{H}_1 = \begin{bmatrix} -4\pi & -2\pi \\ -2\pi & 0 \end{bmatrix}.$$

↬ Формування матриці $\mathbf{S}$:

$$\mathbf{S} = \mathbf{P}_{tt} - \mathbf{P}_{ts}\mathbf{W}^{-1}\mathbf{C} - \left(\mathbf{P}_{ts}\mathbf{W}^{-1}\mathbf{C}\right)^{\mathrm{T}} + \left(\mathbf{W}^{-1}\mathbf{C}\right)^{\mathrm{T}}\mathbf{P}_{ss}\mathbf{W}^{-1}\mathbf{C} =$$

$$= 0 + 2\pi\frac{1}{4} + 2\pi\frac{1}{4} - \frac{1}{4}4\pi\frac{1}{4} = \frac{3}{4}\pi \ .$$

↬ Аналіз матриці **S** на додатну визначеність:

$$\Delta_1 = \frac{3}{4}\pi > 0 \, .$$

Отже, стаціонарна точка є точкою мінімуму.

6.9. Прямі методи в класичній задачі визначення умовного екстремуму

Введення концепції перших і других умовних похідних дозволяє не тільки перевірити в будь-якій припустимій невиродженій точці виконання необхідних і достатніх умов наявності умовного локального екстремуму, але й використовувати прямі методи безумовної оптимізації першого і другого порядків для вирішення класичної задачі пошуку умовного екстремуму. При цьому процес оптимізації протікає в p-вимірному просторі незалежних змінних $t_j, \quad j = \overline{1, p}, \quad p = n - m$, який значно менший початкового n-вимірного простора. Проте, трудомісткість обчислень на кожному кроці оптимізації значно зростає через необхідність визначення умовних похідних (особливо, других умовних похідних). Крім того, після визначення нової точки наближення $\mathbf{t}^{(k+1)}$, яка визначається згідно з рекурентним співвідношенням

$$\mathbf{t}^{(k+1)} = \mathbf{t}^{(k)} + \lambda^{(k)}\Delta\mathbf{t}^{(k)} \ , \tag{6.71}$$

треба вирішувати систему нелінійних рівнянь (6.7) відносно вектора $\mathbf{s}$. Останнє пов’язано з тим, що навіть для незначного приросту вектора незалежних змінних $\Delta\mathbf{t}^{(k)}$ при обчисленні вектори залежних змінних згідно з виразом

$$\mathbf{s}^{(k+1)} = \mathbf{s}^{(k)} + \Delta\mathbf{s}^{(k)} = \mathbf{s}^{(k)} - \mathbf{W}^{-1}\mathbf{C}\Delta\mathbf{t}^{(k)} \qquad (6.72)$$

знайдене рішення може вийти з області припустимих рішень Ω, оскільки вираз (6.30) і похідний з нього (6.72) отримані в результаті лінійної апроксимації (лінійного розкладання цільової функції і системи обмежень у ряд Тейлора). У випадку розв'язання системи рівнянь (6.7) з метою уточнення вектора $\mathbf{s}$ прямими методами значення вектора $\mathbf{s}^{(k+1)}$, підраховане згідно з виразом (6.72), варто взяти як початкове наближення.

Найбільш наочним алгоритмом прямої оптимізації володіє метод покоординатного спуску, який на кожному кроці припускає приріст тільки однієї змінної. Щодо класичної задачі умовної оптимізації такою змінною є одна з незалежних. У випадку умовної мінімізації основне рекурентне співвідношення (6.55) набуває вигляду

$$\boxed{t_r^{(k+1)} = t_r^{(k)} - \frac{\left(\dfrac{\delta y}{\delta t_r}\right)^{(k)}}{\left|\left(\dfrac{\delta^2 y}{\delta t_r^2}\right)^{(k)}\right|}}, \qquad (6.73)$$

а у випадку умовної максимізації –

$$\boxed{t_r^{(k+1)} = t_r^{(k)} + \frac{\left(\dfrac{\delta y}{\delta t_r}\right)^{(k)}}{\left|\left(\dfrac{\delta^2 y}{\delta t_r^2}\right)^{(k)}\right|}} \qquad (6.74)$$

Складність визначення умовних похідних є істотним недоліком прямих методів, у тому числі методу покоординатного спуску. Щоб уникнути зайвих обчислень, що мають місце при визначенні вектора умовних похідних (6.34) і матриці других умовних похідних (6.57), доцільно умовні похідні, що входять у (6.73) і (6.74), обчислювати за формулами

$$\frac{\delta y}{\delta t_r} = \frac{\partial y}{\partial t_r} + \sum_{i=1}^{m} \frac{\partial y}{\partial s_i} \frac{\delta s_\mathrm{i}}{\delta t_r}, \quad r \in \{1,2,...,p\} \quad , \tag{6.75}$$

$$\frac{\delta^2 y}{\delta t_r^2} = \frac{\partial^2 y}{\partial t_r^2} + 2\sum_{i=1}^{m} \frac{\partial^2 y}{\partial t_r \partial s_i} \frac{\delta s_i}{\delta t_r} + \sum_{i-1}^{m} \frac{\partial y}{\partial s_i} \frac{\delta^2 s_i}{\delta t_r^2} +$$

$$+ \sum_{i=1}^{m} \sum_{j=1}^{m} \left[\frac{\partial^2 y}{\partial s_i \partial s_j} \frac{\delta s_i}{\delta t_r} \frac{\delta s_j}{\delta t_r} \right], \quad r \in \{1,2,...,p\}. \tag{6.76}$$

Формули (6.75), (6.76) є модифікаціями відповідно формул (6.39) і (6.61). У цих формулах $\frac{\delta s_\mathrm{i}}{\delta t_r}$, $\frac{\delta s_\mathrm{j}}{\delta t_r}$, $\frac{\delta^2 s_\mathrm{i}}{\delta t_r^2}$ – це умовні похідні неявно заданої функції $s_i(\mathbf{t})$ за r–ю незалежною змінною. Обчислення другої похідної $\frac{\delta^2 s_\mathrm{i}}{\delta t_r^2}$ за формулою (6.76) не потребує попереднього визначення матриці других умовних похідних **S** і пов'язаної з нею матриці **P.** Але формулами (6.75), (6.76) можна скористатися тільки в тих випадках, коли змінні стану не входять явно у функцію цілі, тобто $\frac{\partial y}{\partial \mathbf{s}} = 0$, або рівняння (6.11) розв'язні щодо залежних змінних, тобто є можливість обчислити умовні похідні $\frac{\delta s_\mathrm{i}}{\delta t_r}$, $\frac{\delta s_\mathrm{j}}{\delta t_r}$, $\frac{\delta^2 s_\mathrm{i}}{\delta t_r^2}$.

Якщо функція цілі сепарабельна, тобто

$$y = \sum_{ki=1}^{n} y_k(x_k) , \tag{6.77}$$

то

$$\frac{\partial^2 y}{\partial x_i \partial x_j} = 0, \quad i, j = \overline{1, n}, \quad i \neq j \ . \tag{6.78}$$

Тому вираз (6.76) спрощується і набуває вигляду

$$\frac{\delta^2 y}{\delta t_r^2} = \frac{\partial^2 y}{\partial t_r^2} + \sum_{i-1}^{m} \frac{\partial y}{\partial s_i} \frac{\delta^2 s_i}{\delta t_r^2} + \sum_{i-1}^{m} \frac{\partial^2 y}{\partial s_i^2} \left(\frac{\delta s_i}{\delta t_r} \right)^2 \ . \tag{6.79}$$

Іншим шляхом зменшення трудомісткості оптимізації прямими методами, що базуються на обчисленні других умовних похідних $\dfrac{\delta^2 y}{\delta t_i \delta t_j}$ $\left(i, j = \overline{1, p}\right)$, є заміна їх більш грубими умовними похідними v_{ij}. Другі умовні похідні v_{ij} є елементами матриці $\mathbf{V}$, що обчислюється за формулою

$$\mathbf{V} = \mathbf{H}_{tt} - \mathbf{H}_{ts} \mathbf{W}^{-1} \mathbf{C} - \left(\mathbf{H}_{ts} \mathbf{W}^{-1} \mathbf{C} \right)^{\text{т}} + \left(\mathbf{W}^{-1} \mathbf{C} \right)^{\text{т}} \mathbf{H}_{ss} \mathbf{W}^{-1} \mathbf{C} \ , \tag{6.80}$$

де $\mathbf{H}_{tt}, \mathbf{H}_{ts}, \mathbf{H}_{ss}$ – підматриця матриці Гесса цільової функції

$$\mathbf{H} = \begin{bmatrix} \mathbf{H}_{ss} & \mathbf{H}_{st} \\ \mathbf{H}_{ts} & \mathbf{H}_{tt} \end{bmatrix} . \tag{6.81}$$

Відмінність матриці $\mathbf{V}$ від матриці $\mathbf{S}$ полягає в наступному. Матриця визначається з умов квадратичної апроксимації цільової функції (6.6) і лінійної апроксимації системи обмежень (6.7). Більш грубе визначення других умовних похідних за формулою (6.79), яке веде до збільшення числа кроків процесу оптимізації, компенсується зниженням трудомісткості обчислень на кожному кроці.

При розв'язанні задачі умовної оптимізації прямими методами, як і задачі безумовної оптимізації, для забезпечення початкового кроку необхідно попередньо вибрати з припустимої множини рішень Ω початкову точку наближення $\mathbf{x}^{\circ \mathrm{T}} = \begin{bmatrix} \mathbf{s}^{\circ \mathrm{T}} & \mathbf{t}^{\circ \mathrm{T}} \end{bmatrix}$. Вдалий вибір початкової точки наближення, коли систему обмежень (6.7) не вдасться звести до вигляду (6.12), має велике значення, оскільки дозволяє значно знизити обсяг обчислень. При виборі $\mathbf{x}^{\circ}$ необхідно використовувати всю наявну інформацію про передбачуване положення умовного локального екстремуму і прагнути максимальним способом наблизити до нього початкову точку оптимізації. І тільки за умови повної невизначеності положення шуканого оптимуму початкова точка вибирається довільно з припустимої множини Ω. Довільний вибір вектора $\mathbf{x}^{\circ}$ полягає в довільному задані підвектора $\mathbf{t}^{\circ}$ (по можливості такого, що спрощує обчислення хоча б на першому кроці оптимізації) та у вирішенні системи (6.7) щодо підвектора $\mathbf{s}^{\circ}$.

Для забезпечення завершального кроку оптимізації необхідно також задатися ступенем наближення до точки локального екстремуму ε (точністю обчислення). Ця величина не повинна перевищувати ступінь адекватності математичних моделей, використовуваних у постановці задачі, точності виміру значень y і x_i $\left(i=\overline{1,n}\right)$ в реальних умовах, а також точності обчислення ЕОМ, за допомогою якої може здійснюватися вирішення задачі.

Проілюструємо ідею застосування прямих методів безумовної мінімізації для вирішення задач пошуку умовного екстремуму функції на конкретному прикладі.

Приклад 6.1. З допомогою методу покоординатного спуску в просторі незалежних змінних знайти умовний екстремум у задачі мінімізації

$$y = x_1^2 + 2x_2^2 + 3x_1x_2 + x_3^2 + x_4^2 + 2x_3x_4 \to \min_{\mathbf{x}\in\Omega}, \tag{6.82}$$

$$\Omega: \quad f_1 = x_1|x_1| + x_2|x_2| - 2x_3|x_3| - x_4|x_4| - 1 = 0; \tag{6.83}$$

$$f_2 = x_1|x_1| - x_2|x_2| + x_3|x_3| - 2x_4|x_4| + 2 = 0 \tag{6.84}$$

при початковій точці наближення до умовному екстремуму функції $\mathbf{x}^{(0)^{\text{T}}} = \begin{bmatrix}0{,}7071 & -0{,}7071 & -1 & 1\end{bmatrix}$ і точності рішення $\varepsilon = 0{,}1$.

Розв'язання. Розіб'ємо компоненти вектора $\mathbf{x}$ на залежні й незалежні змінні. Оскільки число обмежень $m = 2$, тому число залежних змінних теж дорівнює двом. Нехай змінна x_1 відповідає залежній змінній s_1; x_2 – змінній s_1; x_3 – незалежній змінній t_1; x_4 – змінній t_1. У результаті зробленої розбивки змінних задача умовної мінімізації (6.82) – (6.84) сформулюється таким чином:

$$y = s_1^2 + 2s_2^2 + 3s_1 s_2 + t_1^2 + t_2^2 + 2t_1 t_2 \to \min_{\mathbf{s},\mathbf{t}\in\boldsymbol{\Omega}} , \tag{6.85}$$

$$\boldsymbol{\Omega}: \quad f_1 = s_1|s_1| + s_2|s_2| - 2t_1|t_1| - t_2|t_2| - 1 = 0 \ ; \tag{6.86}$$

$$f_2 = s_1|s_1| - s_2|s_2| + t_1|t_1| - 2t_2|t_2| + 2 = 0 \ . \tag{6.87}$$

при початковій точці наближення до умовного мінімуму $\begin{bmatrix}\mathbf{s}^{(0)^{\text{T}}} & \mathbf{t}^{(0)^{\text{T}}}\end{bmatrix} =$ $= \begin{bmatrix}0{,}7071 & -0{,}7071 & -1 & 1\end{bmatrix}$ і точності обчислення $\varepsilon = 0{,}1$.

Послідовність обчислювальних операцій відповідає процедурі безумовної мінімізації методом покоординатного спуску з тією лише різницею, що на кожному кроці приріст одержує одна з *незалежних* змінних і після кожного чергового кроку робиться корекція значень залежних змінних шляхом вирішення системи рівнянь (6.7) при відомих незалежних змінних. Таку корекцію доцільно робити за допомогою чисельних методів вирішення систем нелінійних рівнянь, взявши як початкове наближення вектор (6.72). Однак в умовах даного прикладу розв'язання системи (6.86) – (6.87) легко досягається за допомогою простих алгебраїчних перетворень.

У процесі мінімізації на кожному k-му кроці згідно із співвідношенням (6.73) треба знати чисельні значення першої і другої умовних похідних за змінною t_r. Оскільки обмеження задачі (6.86) – (6.87) явно нерозв'язні щодо залежних змінних, перша умовна похідна визначається як r-та складова вектора $\dfrac{\delta y}{\delta \mathbf{t}}$, обчисленого за формулою (6.34), а друга умовна похідна – як діагональний елемент матриці $\mathbf{S}$, обчисленої за виразом (6.57).

Перед початком мінімізації визначимо у загальному вигляді ряд векторів і матриць, що входять у формули (6.34), (6.57). Причому при диференціюванні модульних виразів будемо керуватися такими правилами:

$$\frac{d|u|}{dx} = \operatorname{sgn} u \frac{du}{dx} \;; \tag{6.88}$$

$$\frac{d[(\operatorname{sgn} u)v]}{dx} = \operatorname{sgn} u \frac{dv}{dx} \;; \tag{6.89}$$

де оператор знака $\operatorname{sgn} u = \begin{cases} -1, & \text{если} \quad u < 0; \\ 0, & \text{если} \quad u = 0; \\ 1, & \text{если} \quad u > 0. \end{cases}$

Наприклад, похідну $\dfrac{\partial f_1}{\partial s_1}$ в умовах задачі можна одержати двома способами:

а) $$\frac{\partial f_1}{\partial s_1} = \frac{\partial(s_1|s_1|)}{\partial s_1} = \frac{\partial(s_1^2 \operatorname{sgn} s_1)}{\partial s_1} = 2s_1 \operatorname{sgn} s_1 = 2|s_1| \;;$$

б) $$\frac{\partial f_1}{\partial s_1} = \frac{\partial(s_1|s_1|)}{\partial s_1} = \frac{\partial(uv)\big|_{u=s_1,\, v=|s_1|}}{\partial s_1} = \frac{\partial s_1}{\partial s_1}|s_1| + s_1 \frac{\partial|s_1|}{\partial s_1} =$$

$$= |s_1| + s_1 \operatorname{sgn} s_1 = |s_1| + |s_1| = 2|s_1| \,.$$

Отже, вектор умовних похідних функції (6.85) за незалежними змінними подається вектором

$$\left[\frac{\delta y}{\delta \mathbf{t}}\right]^{\mathrm{T}} = \left[\frac{\partial y}{\partial \mathbf{t}}\right]^{\mathrm{T}} - \left[\frac{\partial y}{\partial \mathbf{s}}\right]^{\mathrm{T}} \mathbf{W}^{-1}\mathbf{C} \ , \tag{6.90}$$

де

$$\left[\frac{\partial y}{\partial \mathbf{t}}\right]^{\mathrm{T}} = \begin{bmatrix} 2t_1 + 2t_2 & 2t_1 + 2t_2 \end{bmatrix} ; \tag{6.91}$$

$$\left[\frac{\partial y}{\partial \mathbf{s}}\right]^{\mathrm{T}} = \begin{bmatrix} 2s_1 + 3s_2 & 3s_1 + 4s_2 \end{bmatrix}; \tag{6.92}$$

$$\mathbf{W} = \begin{bmatrix} 2|s_1| & 2|s_2| \\ 2|s_1| & -2|s_2| \end{bmatrix}; \quad \mathbf{W}^{-1} = \begin{bmatrix} \frac{1}{4|s_1|} & \frac{1}{4|s_1|} \\ \frac{1}{4|s_2|} & \frac{-1}{4|s_2|} \end{bmatrix}, \quad s_1, s_2 \neq 0, \tag{6.93}$$

$$\mathbf{C} = \begin{bmatrix} -4|t_1| & -2|t_2| \\ 2|t_1| & -4|t_2| \end{bmatrix} . \tag{6.94}$$

Матриця других умовних похідних цільової функції за змінними рішення подається матрицею

$$\mathbf{S} = \mathbf{P}_{tt} - \mathbf{P}_{ts}\mathbf{W}^{-1}\mathbf{C} - \left(\mathbf{P}_{ts}\mathbf{W}^{-1}\mathbf{C}\right)^{\mathrm{T}} + \left(\mathbf{W}^{-1}\mathbf{C}\right)^{\mathrm{T}} \mathbf{P}_{ss}\mathbf{W}^{-1}\mathbf{C}, \tag{6.95}$$

де $\mathbf{P}_{tt}$, $\mathbf{P}_{ts}$, $\mathbf{P}_{ss}$ – підматриці матриці $\mathbf{P}$**.** Остання визначається за допомогою виразу (6.49)

$$\mathbf{P} = \mathbf{H} - \sum_{i=1}^{m} \lambda_i \mathbf{H}_i = \mathbf{H} + \lambda_1 \mathbf{H}_1 + \lambda_2 \mathbf{H}_2 \ . \tag{6.96}$$

Тут $\mathbf{H}$ – гессіан функції (6.85); $\mathbf{H}_1$ – гессіан функції обмеження (6.86); $\mathbf{H}_2$ – гессіан функції обмеження (6.87); λ_1, λ_2 – складові вектора

$$\boldsymbol{\lambda}^{\mathrm{T}} = \left(\frac{\partial \ y}{\partial \ \mathbf{s}}\right)^{\mathrm{T}} \mathbf{W}^{-1} \ , \tag{6.97}$$

де співмножники правої частини визначаються формулами (6.92), (6.93). Гессіани, що входять у праву частину виразу (6.96), в умовах задачі визначаються таким чином:

$$\mathbf{H} = \begin{bmatrix} 2 & 3 & 0 & 0 \\ 3 & 4 & 0 & 0 \\ 0 & 0 & 2 & 2 \\ 0 & 0 & 2 & 2 \end{bmatrix}; \tag{6.98}$$

$$\mathbf{H}_1 = \begin{bmatrix} 2\operatorname{sgn} s_1 & 0 & 0 & 0 \\ 0 & 2\operatorname{sgn} s_2 & 0 & 0 \\ 0 & 0 & -4\operatorname{sgn} t_1 & 0 \\ 0 & 0 & 0 & -2\operatorname{sgn} t_2 \end{bmatrix}; \tag{6.99}$$

$$\mathbf{H}_2 = \begin{bmatrix} 2\operatorname{sgn} s_1 & 0 & 0 & 0 \\ 0 & -2\operatorname{sgn} s_2 & 0 & 0 \\ 0 & 0 & 2\operatorname{sgn} t_1 & 0 \\ 0 & 0 & 0 & -4\operatorname{sgn} t_2 \end{bmatrix}. \tag{6.100}$$

З (6.98) випливає, що гессіан функції цілі не залежить від змінних задачі. Крім того, із (6.96) – (6.100) виходить, що всі елементи підматриці $\mathbf{P}_{ts}$ рівні нулю, внаслідок чого вираз (6.95) спрощується:

$$\mathbf{S} = \mathbf{P}_{tt} + \left(\mathbf{W}^{-1}\mathbf{C}\right)^{\mathrm{T}} \mathbf{P}_{ss} \mathbf{W}^{-1}\mathbf{C}. \tag{6.101}$$

Визначимо також до початку мінімізації в початковій точці наближення $\left[\mathbf{s}^{(0)^{\mathrm{T}}} \quad \mathbf{t}^{(0)^{\mathrm{T}}}\right] = \left[0{,}7071 \quad -0{,}7071 \quad 1 \quad 1\right]$ значення функції цілі: $y^{(0)} = 0$. Знайдене значення цільової функції знадобиться для контролю правильності руху до умовного мінімуму. Обчислення будемо здійснювати з точністю до четвертого десяткового знака, щоб гарантовано забезпечити точність, задану умовами задачі.

Тепер перейдемо безпосередньо до процедури покрокової мінімізації.

Перший крок. На першому кроку мінімізації здійснюємо рух у просторі незалежних змінних уздовж осі t_1. Згідно з формулами (6.90) – (6.94) знаходимо вектор перших умовних похідних функції цілі за змінними рішення:

$$\left(\frac{\delta y}{\delta \mathbf{t}}\right)^{(0)^{\mathrm{T}}} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}^{\mathrm{T}} - \begin{bmatrix} -0{,}7071 \\ 0{,}7071 \end{bmatrix}^{\mathrm{T}} \begin{bmatrix} 0{,}3536 & 0{,}3536 \\ 0{,}3536 & 0{,}3536 \end{bmatrix} \begin{bmatrix} -4 & -2 \\ 2 & -4 \end{bmatrix} = \begin{bmatrix} -2 \\ -1 \end{bmatrix}^{\mathrm{T}} .$$

Звідки $\left(\dfrac{\partial y}{\partial t_1}\right)^{(0)} = -2$. Той факт, що складові вектора умовних похідних за абсолютним значенням перевершують задану точність $\varepsilon = 0{,}1$, свідчить про необхідність проведення процесу мінімізації.

Обчислимо гессіани $\mathbf{H}_1$, $\mathbf{H}_2$ у початковій точці наближення:

$$\mathbf{H}_1 = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & -2 & 0 & 0 \\ 0 & 0 & -4 & 0 \\ 0 & 0 & 0 & -2 \end{bmatrix}; \quad \mathbf{H}_2 = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & -4 \end{bmatrix}. \qquad (6.102)$$

Підрахуємо вектор коефіцієнтів чутливості згідно з формулою (6.97):

$$\mathbf{л}^{(0)^{\mathrm{T}}} = \begin{bmatrix} -0{,}7071 & -0{,}7071 \end{bmatrix} \begin{bmatrix} 0{,}3536 & 0{,}3536 \\ 0{,}3536 & -0{,}3536 \end{bmatrix} = \begin{bmatrix} -0{,}5 & 0 \end{bmatrix} .$$

Звідки $\lambda_1^{(0)} = -0{,}5; \quad \lambda_2^{(0)} = 0.$

Після підстановки в (6.96) знайдених значень коефіцієнтів чутливості й обчислених гессіанів, одержимо матрицю

$$\mathbf{P}^{(0)} = \begin{bmatrix} 3 & 3 & 0 & 0 \\ 3 & 3 & 0 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Звідки $\mathbf{P}_{ss}^{(0)} = \begin{bmatrix} 3 & 3 \\ 3 & 3 \end{bmatrix}; \quad \mathbf{P}_{tt}^{(0)} = \begin{bmatrix} 4 & 2 \\ 2 & 1 \end{bmatrix}.$

Підставляючи $\mathbf{P}_{ss}^{(0)}, \mathbf{P}_{tt}^{(0)}$ і раніше обчислені матриці $\mathbf{W}^{-1}$ і C в (6.101), знаходимо матрицю других умовних похідних у початковій точці наближення: $\mathbf{S}^{(0)} = \begin{bmatrix} 28 & 14 \\ 14 & 7 \end{bmatrix}$. Звідки $\left(\dfrac{\partial^2 y}{\partial t_1^2} \right)^{(0)} = 28$.

Обчислимо приріст змінної t_1:

$$\Delta t_1^{(1)} = -\frac{\left(\dfrac{\delta y}{\delta t_1} \right)^{(0)}}{\left| \left(\dfrac{\delta^2 y}{\delta t_1^2} \right)^{(0)} \right|} = -\frac{(-2)}{|28|} = 0{,}0714 \ .$$

Змінюючи змінну t_1 на отриманий приріст переміщуємося уздовж осі t_1 в нову точку з координатами

$$\begin{bmatrix} t_1^{(1)} \\ t_2^{(1)} \end{bmatrix} = \begin{bmatrix} t_1^{(0)} + \Delta t_1^{(0)} \\ t_2^{(0)} \end{bmatrix} = \begin{bmatrix} -0.9286 \\ 1 \end{bmatrix}.$$

Скоригуємо значення залежних змінних шляхом вирішення системи рівнянь (6.86) – (6.87) щодо цих змінних при новому значенні незалежної змінної t_1:

$$s_1 |s_1| = \frac{t_1 |t_1| + 3t_2 |t_2| - 1}{2} = 0{,}5689 \ ;$$

$$s_2 |s_2| = \frac{3t_1 |t_1| - t_2 |t_2| + 3}{2} = 0{,}293 \ .$$

Звідки $s_1^{(1)} = 0{,}7542; \quad s_2^{(1)} = -0{,}5416$.

У новій точці $\mathbf{x}^{(1)} = \left[\mathbf{s}^{(0)^{\text{т}}} \quad \mathbf{t}^{(0)^{\text{т}}}\right] = \left[0{,}7071 \quad -0{,}7071 \quad -1 \quad 1\right]$ значення функції в порівнянні з початковим помітно зменшилося і досягло величини $-0{,}0649$, що свідчить про правильність зробленого кроку.

Другий крок. Послідовність обчислення аналогічна послідовності першого кроку з тією лише різницею, що приріст одержує друга незалежна змінна:

$$t_2^{(2)} = t_2^{(1)} + \Delta t_2^{(1)} = 1 - \frac{\left(\dfrac{\delta y}{\delta t_2}\right)^{(1)}}{\left|\left(\dfrac{\delta^2 y}{\delta t_2^2}\right)^{(1)}\right|} = 1 - \frac{(-0{,}1775)}{|2{,}74538|} = 1{,}0647 \ .$$

У новій точці наближення до екстремуму $\mathbf{x}^{(1)} = \left[\mathbf{s}^{(0)^{\text{т}}} \quad \mathbf{t}^{(0)^{\text{т}}}\right] = \left[0{,}7071 \quad -0{,}7071 \quad -1 \quad 1\right]$ значення функції цілі зменшилося до значення $-0{,}0710$.

На *третьому кроку* знову варіюємо першою незалежною змінною, на *четвертому* – другою і т.д., поки не буде досягнута задана точність. Основні проміжні результати мінімізації наведені в табл.6.1.

Після одинадцятого кроку одержимо вектор перших умовних похідних $\left(\dfrac{\delta y}{\delta \mathbf{t}}\right)^{(11)} = \begin{bmatrix} 0{,}0018 \\ -0{,}0832 \end{bmatrix}$, складові якого за абсолютним значенням не перевершують заданої точності обчислення $\varepsilon = 0{,}1$. Отже, із заданою точністю умовний екстремум досягнутий. Тому процес мінімізації зупиняємо.

В останній точці наближення перевіряємо достатні умови точки локального мінімуму. Додатна визначеність матриці других умовних похідних $\mathbf{s}^{(11)} = \begin{bmatrix} 17{,}7499 & 6{,}511 \\ 6{,}511 & 3{,}0917 \end{bmatrix}$ у цій точці свідчить про виконання достатніх умов. Таким чином, точка $\mathbf{x}^{(11)} = \left[0{,}7071 \ -0{,}7071 \ -1 \ 1\right]$ є точкою локального умовного мінімуму функції (6.84) при обмеженнях (6.86) – (6.87).

Таблиця 6.1.

k	$\mathbf{t}^{(k)}$	$\mathbf{s}^{(k)}$	$y^{(k)}$	$\left[\frac{\delta y}{\delta t_r}\right]^{(k)}$	$\left[\frac{\delta^2 y}{\delta t_r^2}\right]^{(k)}$	$\Delta t_r^{(k)}$
0	$\begin{bmatrix}-1\\1\end{bmatrix}$	$\begin{bmatrix}0,7071\\-0,7071\end{bmatrix}$	0	-2	28	0,0714
1	$\begin{bmatrix}-0,9286\\1\end{bmatrix}$	$\begin{bmatrix}0,7542\\-0,5416\end{bmatrix}$	-0,0649	-0,1775	2,7453	0,0647
2	$\begin{bmatrix}-0,9286\\1,0647\end{bmatrix}$	$\begin{bmatrix}0,8770\\-0,6002\end{bmatrix}$	-0,0710	0,7820	33,0121	-0,0237
3	$\begin{bmatrix}-0,9523\\1,0647\end{bmatrix}$	$\begin{bmatrix}0,8643\\-0,6535\end{bmatrix}$	-0,0807	-0,1974	2,9605	0,0668
4	$\begin{bmatrix}-0,9523\\1,1314\end{bmatrix}$	$\begin{bmatrix}0,9832\\-0,7073\end{bmatrix}$	-0,0870	0,5253	25,2048	-0,0208
5	$\begin{bmatrix}-0,9731\\1,1314\end{bmatrix}$	$\begin{bmatrix}0,9729\\-0,7487\end{bmatrix}$	-0,0926	-0,1497	2,9300	0,0511
6	$\begin{bmatrix}-0,9731\\1,1825\end{bmatrix}$	$\begin{bmatrix}1,0602\\-0,7871\end{bmatrix}$	-0,0965	0,3678	21,5243	-0,0171
7	$\begin{bmatrix}-0,9902\\1,1825\end{bmatrix}$	$\begin{bmatrix}1,0522\\-0,8184\end{bmatrix}$	-0,0997	-0,1222	2,9081	0,0420
8	$\begin{bmatrix}-0,9902\\1,2245\end{bmatrix}$	$\begin{bmatrix}1,1220\\-0,8488\end{bmatrix}$	-0,1024	0,2842	19,4041	-0,0146
9	$\begin{bmatrix}-1,0048\\1,2245\end{bmatrix}$	$\begin{bmatrix}1,1155\\-0,8742\end{bmatrix}$	-0,1044	-0,1004	2,8929	0,0347
10	$\begin{bmatrix}-1,0048\\1,2592\end{bmatrix}$	$\begin{bmatrix}1,1720\\-0,8985\end{bmatrix}$	-0,1062	0,2276	18,0298	-0,0126
11	$\begin{bmatrix}-1,0174\\1,2592\end{bmatrix}$	$\begin{bmatrix}1,1665\\-0,8985\end{bmatrix}$	-0,1077	-0,0832		

Приклад 6.2. Вирішити задачу мінімізації (6.86) – (6.88) за допомогою методу покоординатного спуску в умовах квадратично-лінійної апроксимації, тобто обчислюючи на кожному кроці мінімізації замість других умовних похідних $\dfrac{\partial^2 y}{\partial t_r^2}$ умовні похідні v_{rr}, одержані при квадратичній апроксимації функції цілі і лінійній апроксимації функцій обмежень.

Розв'язання. За умовами задачі точка початкового наближення і точність обчислення не задані. Виберемо як початкове наближення колишню точку, тобто $\mathbf{x}^{(0)^{\mathrm{T}}} = \begin{bmatrix}0{,}7071 & -0{,}7071 & -1 & 1\end{bmatrix}$, і точність обчислення також залишимо колишньою, тобто $\varepsilon = 0{,}1$.

На кожному кроці умовної мінімізації зазначеним за умовами задачі методом поточне значення варійованої змінної буде визначатися виразом

$$t_r^{(k+1)} = t_r^{(k)} - \frac{\left(\dfrac{\delta y}{\delta t_r}\right)^{(k)}}{\left|v_{rr}^{(k)}\right|}, \tag{6.103}$$

яке є модифікацією рекурентного співвідношення (6.73), що враховує квадратично-лінійну апроксимацію розв'язуваної задачі.

Для визначення першої умовної похідної $\dfrac{\delta y}{\delta t_r}$ доцільно використовувати раніше наведені формули (6.90) – (6.94)., а при визначенні другою умовною похідною – формулу (6.80). Остання потребує визначення матриці Гесса функції цілі. За умовами задачі гессіан не залежить від вектора $\mathbf{x}$:

$$\mathbf{H} = \left[\begin{array}{cc|cc} 2 & 3 & 0 & 0 \\ 3 & 4 & 0 & 0 \\ \hline 0 & 0 & 2 & 2 \\ 0 & 0 & 2 & 2 \end{array}\right],$$

що спрощує процес мінімізації, оскільки матриця Гесса визначається тільки на першому кроці. Крім того, підматриця $\mathbf{H}_{ts} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = 0$ і, отже, вираз (6.80) допускає подальше спрощення:

$$\mathbf{V} = \mathbf{H}_{tt} + \left(\mathbf{W}^{-1}\mathbf{C}\right)^{\mathrm{T}} \mathbf{H}_{ss} \mathbf{W}^{-1}\mathbf{C} \ , \qquad (6.104)$$

де $\mathbf{H}_{ss} = \begin{bmatrix} 2 & 3 \\ 3 & 4 \end{bmatrix}$; $\mathbf{H}_{tt} = \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix}$,

а основне рекурентне співвідношення (6.103) набуває вигляду

$$t_r^{(k+1)} = t_r^{(k)} - \frac{\left(\dfrac{\delta y}{\delta t_r}\right)^{(k)}}{\left|v_{rr}^{(k)}\right|} \ , \qquad (6.105).$$

Перейдемо безпосередньо до процедури покрокової мінімізації.

Перший крок. Визначимо значення функції цілі у початковій точці наближення: $y^{(0)} = 0$. Воно знадобиться для підтвердження наближення до умовного мінімуму функції після першого кроку. Обчислимо вектор умовних похідних функції цілі, за незалежними змінними за формулами (6.90) – (6.94). Шуканий вектор, а він вже визначався в попередньому прикладі, $\left(\dfrac{\delta y}{\delta \mathbf{t}}\right)^{(0)\mathrm{T}} = \begin{bmatrix} -2 & -1 \end{bmatrix}$ говорить про те, що необхідна точність не забезпечується: складові вектора за абсолютним значенням перевершують задану точність $\varepsilon = 0{,}1$, а це значить, що необхідно здійснити, принаймні, один крок мінімізації.

Для здійснення першого кроку, згідно з основним рекурентним співвідношенням (6.103), не дістає тільки значення $v_{11}^{(0)}$, яким є перший діагональний елемент матриці $\mathbf{V}$. Остання визначається за формулами (6.104) із залученням раніше отриманих результатів (6.93) – (6.94):

$$\mathbf{V}^{(0)} = \mathbf{H}_{tt} + \left(\mathbf{W}^{(0)^{-1}}\mathbf{C}^{(0)}\right)^{\mathrm{т}} \mathbf{H}_{ss}\mathbf{W}^{(0)^{-1}}\mathbf{C}^{(0)} = \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix} +$$

$$+\left(\begin{bmatrix} 0{,}3536 & 0{,}3536 \\ 0{,}3536 & -0{,}3536 \end{bmatrix}\begin{bmatrix} -4 & -2 \\ 2 & -4 \end{bmatrix}\right)^{\mathrm{т}} \begin{bmatrix} 2 & 3 \\ 3 & 4 \end{bmatrix}\begin{bmatrix} 0{,}3536 & 0{,}3536 \\ 0{,}3536 & -0{,}3536 \end{bmatrix} \times$$

$$\times \begin{bmatrix} -4 & -2 \\ 2 & -4 \end{bmatrix} = \begin{bmatrix} 30 & 11 \\ 11 & 4 \end{bmatrix}.$$

Звідки $v_{11}^{(0)} = 30$. Тоді згідно з (6.105)

$$t_1^{(1)} = t_1^{(1)} - \frac{\left(\dfrac{\delta y}{\delta t_1}\right)^{(0)}}{\left|v_{11}^{(0)}\right|} = -1 - \frac{(-2)}{|30|} = -0{,}9333 \ .$$

Значення змінної t_2 залишається колишнім $\left(t_2^{(1)} = t_2^{(1)} = 1\right)$. Скоригуємо значення залежних змінних шляхом вирішення системи рівнянь (6.86) – (6.87) щодо цих змінних при новому значенні незалежної змінної t_1:

$$s_1|s_1| = \frac{t_1|t_1| + 3t_2|t_2| - 1}{2} = 0{,}5645 \ ;$$

$$s_2\left|s_2\right| = \frac{3t_1\left|t_1\right| - t_2\left|t_2\right| + 3}{2} = -0{,}3066\,.$$

Звідки $s_1^{(1)} = 0{,}7542; \quad s_2^{(1)} = -0{,}5538\,.$

У новій точці наближення значення функції зменшилося до величини $y^{(1)} = -0{,}0659$, тобто рух відбувається у правильному напрямку.

Другий крок. На другому кроці послідовність обчислень аналогічна послідовності першого кроку з тієї лише різницею, що немає необхідності обчислювати підматриці $\mathbf{H}_{ss}$ і $\mathbf{H}_{tt}$ (вони були визначені на першому кроці) і приріст вже одержує незалежна змінна t_2:

$$\Delta t_2^{(1)} = -\frac{\left(\dfrac{\delta y}{\delta t_2}\right)^{(1)}}{\left|v_{22}^{(1)}\right|} = -\frac{(-0{,}2186)}{\left|2{,}4172\right|} = 0{,}0904\ ;$$

$$\begin{bmatrix} t_1^{(2)} \\ t_2^{(2)} \end{bmatrix} = \begin{bmatrix} t_1^{(1)} \\ t_2^{(1)} + \Delta t_2^{(1)} \end{bmatrix} = \begin{bmatrix} -0.9286 \\ 1{,}0904 \end{bmatrix}.$$

У новій точці наближення значення функції зменшується до величини $y^{(2)} = -0{,}0747$, тобто відбувається «спуск» по поверхні функції у просторі незалежних змінних паралельно осі t_2.

На третьому кроці знову змінюємо незалежну змінну t_1; на четвертому – t_2 і т.д. Основні проміжні результати покрокового процесу мінімізації наведені в табл.6.2

Таблиця 6.2

k	$\mathbf{t}^{(k)}$	$\mathbf{s}^{(k)}$	$y^{(k)}$	$\left[\frac{\delta y}{\delta t_r}\right]^{(k)}$	$v_{rr}^{(k)}$	$\Delta t_r^{(k)}$
0	$\begin{bmatrix}-1\\1\end{bmatrix}$	$\begin{bmatrix}0{,}7071\\-0{,}7071\end{bmatrix}$	0	-2	30	0,0667
1	$\begin{bmatrix}-0{,}9333\\1\end{bmatrix}$	$\begin{bmatrix}0{,}7513\\-0{,}5538\end{bmatrix}$	-0,0659	-0,2186	2,4172	0,0904
2	$\begin{bmatrix}-0{,}933\\1{,}0904\end{bmatrix}$	$\begin{bmatrix}0{,}9209\\-0{,}6333\end{bmatrix}$	-0,0747	0,7918	28,7794	-0,0275
3	$\begin{bmatrix}-0{,}9608\\1{,}0904\end{bmatrix}$	$\begin{bmatrix}0{,}9066\\-0{,}6923\end{bmatrix}$	-0,0856	-0,1776	2,4656	0,0720
4	$\begin{bmatrix}-0{,}9608\\1{,}1624\end{bmatrix}$	$\begin{bmatrix}1{,}0321\\-0{,}7485\end{bmatrix}$	-0,0912	0,5154	22,6389	-0,0227
5	$\begin{bmatrix}-0{,}9836\\1{,}1624\end{bmatrix}$	$\begin{bmatrix}1{,}0213\\-0{,}7916\end{bmatrix}$	-0,0971	-0,1729	2,4648	0,0701
6	$\begin{bmatrix}-0{,}9836\\1{,}2325\end{bmatrix}$	$\begin{bmatrix}1{,}1380\\-0{,}8431\end{bmatrix}$	-0,0997	0,4613	19,1613	-0,0241
7	$\begin{bmatrix}-1{,}0077\\1{,}2325\end{bmatrix}$	$\begin{bmatrix}1{,}1273\\-0{,}8847\end{bmatrix}$	-0,1052			

На сьомому кроці мінімізації потрапляємо в точку $\mathbf{x}^{(7)} = [0{,}7071 \quad -0{,}7071 \quad -1 \quad 1]$, в якій складові вектора перших умовних похідних $\left(\dfrac{\delta y}{\delta \mathbf{t}}\right)^{(7)\mathrm{T}} = [0{,}0035 \quad -0{,}0961]$ за абсолютним значенням не перевершують задану точність $\varepsilon = 0{,}1$. Оскільки задана точність досягнута, процес мінімізації припиняємо.

Ідея квадратично-лінійної апроксимації може бути узагальнена на всі прямі методи умовної оптимізації другого порядку. Механічна заміна других умовних похідних $\dfrac{\partial^2 y}{\partial t_{ij}^2}$ $\left(i, j = \overline{1, p}\right)$ умовними похідними v_{ij} або відповідно матриці $\mathbf{S}$ більш грубою матрицею $\mathbf{V}$ у методах другого порядку значно спрощує обчислювальні процедури процесу оптимізації. Наприклад, в умовній мінімізації методом Ньютона транспозиція зазначених матриць приводить до основного рекурентного співвідношення у вигляді

$$\mathbf{t}^{(k+1)} = \mathbf{t}^{(k)} + \Delta\mathbf{t}^{(k)} = \mathbf{t}^{(k)} - \mathbf{V}^{(k)^{-1}}{}^{-1} \left(\frac{\delta y}{\delta \mathbf{t}}\right)^{(k)}. \qquad (6.106)$$

Теорія умовної оптимізації, незважаючи на строгу математичну основу, продовжує розвиватися: розробляються нові методи пошукової оптимізації, модифікуються старі. Програмування методів на машинних мовах є чудовим способом як більш глибокого вивчення і розуміння особливостей умовної оптимізації, так і оволодіння основами алгоритмізації і програмування. Читач, який захопиться програмною реалізацією методів умовної оптимізації, безумовно, одержить неоціненний практичний досвід, а, можливо, і додасть свій внесок у розвиток теорії оптимізації.

6.10. Контрольні запитання і вправи

1. У чому суттєва відмінність задач пошуку умовного екстремуму від задач безумовної оптимізації?

2. Які задачі на умовний екстремум належать до класичних задач умовної оптимізації?

3. Сформулюйте класичну задачу умовної мінімізації.

4. Сформулюйте класичну задачу умовної максимізації.

5. Чи може точка локального екстремуму бути внутрішньою точкою області припустимих рішень Ω?

6. У чому полягає концепція залежних і незалежних змінних?

7. Які змінні в задачі пошуку умовного екстремуму називають змінними рішення? Змінними стану?

8. На чому засноване вирішення класичної задачі оптимізації методом підстановки?

9. Чи володіє метод підстановки універсальністю?

10. При яких умовах може бути застосований метод підстановки?

11. Чи можна використовувати метод підстановки при пошуку локального екстремуму нелінійної функції при лінійних обмеженнях на змінні задачі?

12. Дати визначення точці локального умовного мінімуму, максимуму.

13. Знайти мінімум функції

$$y(\mathbf{x}) = 7x_1^3 + 9x_2^3 + 15\left(x_1^2 x_2 + x_1 x_2^2\right) - 150\left(x_1 + x_2\right),$$

якщо змінні функції пов'язані співвідношенням

$$x_1 x_2 - 1 = 0 .$$

Вирішення здійснити за допомогою методу підстановки.

14. Методом підстановки розв’язати задачу умовної мінімізації:

$$y(\mathbf{x}) = \frac{1}{2}x_1^2 + \frac{1}{2x_2} + \frac{x_3}{x_2} - \frac{3}{4} \to \min_{\mathbf{x}\in\Omega},$$

$$\Omega: \quad \begin{aligned} & x_1^2 x_2 - 1 = 0; \\ & x_1 x_3 - 1 = 0. \end{aligned}$$

15. Знайти мінімум функції $y(\mathbf{x}) = x_1^2 - 2x_1 x_2$, якщо аргументи функції зв'язані залежністю $x_1^2 x_2 - 8 = 0$.

16. Яку матрицю називають матрицею стану? Яка розмірність матриці? Запишіть матрицю стану в загальному вигляді.

17. Яку матрицю називають матрицею управління? Яка розмірність матриці? Запишіть матрицю управління в загальному вигляді.

18. Дати визначення першій умовній похідній. За якою формулою вона визначається?

19. Записати вектор-рядок умовних похідних у вигляді вектора-стовпця.

20. Що називають умовною похідною i–й залежної змінної за j–й незалежній змінній рішення? Що вона показує?

21. Як визначається матриця умовних похідних змінних стану за змінними рішення?

22. Яка необхідна умова існування умовного екстремуму функції?

23. Які точки називають стаціонарними у завданнях пошуку умовного екстремуму?

24. На чому засноване визначення умовних стаціонарних точок методом умовних похідних?

25. Як формується система рівнянь для визначення умовних стаціонарних точок за методом якобіана?

26. Навести послідовність етапів визначення умовних стаціонарних точок за методом Якобі.

27. Знайти стаціонарні точки задачі умовної мінімізації

$$y(\mathbf{x}) = 2x_1^2 - \frac{x_3}{x_2} \to \min_{\mathbf{x}\in\Omega},$$

$$\Omega: \quad f_1 = x_1^2 x_2 - 1 = 0;$$

$$f_2 = x_1 x_3 - 2 = 0$$

за методом якобіана.

28. За допомогою методу Якобі знайти стаціонарні точки в задачі умовної мінімізації

$$y(\mathbf{x}) = x_1^2 + 2x_1x_2 \to \min_{\mathbf{x}\in\Omega},$$
$$\Omega: \quad f_1 = x_1^2 x_2 - 8 = 0.$$

29. Записати функцію Лагранжа в загальному вигляді.

30. Як формується система рівнянь для визначення умовних стаціонарних точок за методом невизначених множників Лагранжа?

31. Навести послідовність етапів визначення умовних стаціонарних точок за методом невизначених множників Лагранжа .

32. За допомогою методу невизначених множників Лагранжа знайти стаціонарні точки в задачі умовної мінімізації

$$y(\mathbf{x}) = 2x_1^2 - \frac{x_3}{x_2} \to \min_{\mathbf{x}\in\Omega},$$
$$\Omega: \quad \begin{aligned} f_1 &= x_1^2 x_2 - 1 = 0; \\ f_2 &= x_1 x_3 - 2 = 0 \end{aligned}$$

33. За методом невизначених множників Лагранжа знайти стаціонарні точки в задачі умовної мінімізації

$$y(\mathbf{x}) = \frac{1}{2}x_1^2 + \frac{1}{2x_2} + \frac{x_3}{x_2} - \frac{3}{4} \to \min_{\mathbf{x}\in\Omega},$$
$$\Omega: \quad \begin{aligned} x_1^2 x_2 - 1 &= 0; \\ x_1 x_3 - 1 &= 0. \end{aligned}$$

34. За допомогою методу невизначених множників Лагранжа знайти стаціонарні точки в задачі умовної мінімізації

$$y(\mathbf{x}) = x_1^2 + 2x_1x_2 \to \min_{\mathbf{x}\in\Omega},$$
$$\Omega: \quad f_1 = x_1^2 x_2 - 8 = 0.$$

35. Використовуючи метод невизначених множників Лагранжа, знайти стаціонарні точки в задачі умовної мінімізації

$$y(\mathbf{x}) = 7x_1^3 + 9x_2^3 + 15\left(x_1^2 x_2 + x_1 x_2^2\right) - 150(x_1 + x_2) \to \min_{\mathbf{x}\in\Omega},$$

$$\Omega: \quad x_1 x_2 - 1 = 0.$$

36. Яка достатня умова існування умовного екстремуму функції?

37. Як визначається матриця других умовних похідних функції цілі за незалежними змінними? Яка розмірність цієї матриці?

38. Перевірити виконання достатніх умов локального мінімуму в умовно стаціонарній точці $\mathbf{x}^{\circ\,\mathrm{T}} = \begin{bmatrix} \frac{1}{2} & 4 & 4 \end{bmatrix}$ задачі умовної мінімізації

$$y(\mathbf{x}) = \frac{1}{2}x_1^2 + \frac{1}{2x_2} + \frac{x_3}{x_2} - \frac{3}{4} \to \min_{\mathbf{x}\in\Omega},$$

$$\Omega: \quad x_1^2 x_2 - 1 = 0;$$
$$x_1 x_3 - 1 = 0.$$

39. Здійснити один крок мінімізації в умовах прикладу 6.1 методом покоординатного спуску в просторі незалежних змінних, взявши як початкове наближення точку

$$\mathbf{x}^{(0)^{\mathrm{T}}} = \begin{bmatrix} 1{,}9060 & -1{,}6168 & -1{,}2813 & 1{,}8173 \end{bmatrix}.$$

Порівняти результати мінімізації з рішенням прикладу 6.1. Чи є нова точка наближення, отримана в результаті першого кроку, шуканим умовним мінімумом, якщо точність рішення залишається колишньою, тобто $\varepsilon = 0{,}1$?

40. Здійснити два кроки (одну ітерацію) процесу мінімізації в умовах прикладу 6.1 за методом покоординатного спуску в просторі незалежних змінних, взявши як початкове наближення точку

$$\mathbf{x}^{(0)^{\mathrm{T}}} = \begin{bmatrix} -1{,}6231 & 1{,}6647 & 0{,}5796 & -1{,}2389 \end{bmatrix},$$

і дати висновок про те, чи треба продовжувати процес мінімізації.

41. Знайти мінімальне рішення і мінімум функції в задачі мінімізації (6.85) – (6.87) при початковому наближенні

$$\mathbf{x}^{(0)^{\mathrm{T}}} = \begin{bmatrix} 0{,}8030 & -0{,}5378 & -0{,}9190 & -1{,}0221 \end{bmatrix}$$

і точності обчислення $\varepsilon = 0{,}15$.

42. Як обчислюються другі умовні похідні в умовах квадратичної апроксимації функції цілі і лінійної апроксимації функцій обмежень?

43. Здійснити одну ітерацію (два кроки) в задачі умовної мінімізації (6.85) – (6.87) за методом покоординатного спуску в просторі незалежних змінних, починаючи із точки

$$\mathbf{x}^{(0)^{\mathrm{T}}} = \begin{bmatrix} 1{,}9538 & -1{,}6356 & -1{,}2813 & 1{,}8508 \end{bmatrix},$$

обчислюючи другі умовні похідні за допомогою формули (6.104).

44. Знайти мінімальне рішення і мінімум функції в задачі (6.85) – (6.87) за методом покоординатного спуску в умовах квадратично-лінійної апроксимації при точності обчислення $\varepsilon = 0{,}15$, здійснюючи перший крок із точки

$$\mathbf{x}^{(0)^{\mathrm{T}}} = \begin{bmatrix} 1{,}8030 & -0{,}5378 & -0{,}9190 & 1{,}1780 \end{bmatrix}.$$

45. Як вибирається початкова точка наближення і точність обчислення в задачах умовної оптимізації прямими методами?

6.11. Фонд індивідуальних завдань

Індивідуальне завдання №22. Методом підстановки знайти умовні стаціонарні точки в задачі мінімізації

$$y(\mathbf{x}) = a_1 x_1^3 + a_2 x_2^3 + a_3\left(x_3^2 + x_4^2\right) + a_4\left(x_1 + x_2 + x_3\right) \to \min_{\mathbf{x}\in\Omega}, \quad (6.102)$$

$$\Omega: \quad x_1 + x_2 + x_3 + x_4 - 1 = 0 \;; \quad (6.103)$$

$$2x_1 - x_2 + 2x_3 - x_4 - 1 = 0 \ , \qquad (6.104)$$

де коефіцієнти a_1, a_2, a_3 і a_4 вибираються з табл.6.3.

Таблиця 6.3.

Варіант	a_1	a_2	a_3	a_4
1	4	5	9	-150
2	6	7	12	-120
3	8	9	15	-90
4	10	11	18	-60
5	12	13	21	-30
6	3	5	6	-150
7	5	7	9	-120
8	7	9	12	-90
9	9	11	15	-60
10	11	13	18	-30
11	4	3	3	-30
12	6	5	6	-60
13	8	7	9	-90
14	10	9	12	-120
15	12	11	15	-150
16	9	10	21	-30
17	3	8	18	-60
18	9	7	15	-90
19	8	6	9	-120
20	8	5	6	-150
21	4	9	14	-30
22	5	7	12	-60
23	7	5	10	-90
24	2	8	9	-120
25	3	2	8	-150
26	4	3	2	-30
27	5	4	3	-60
28	6	5	4	-90
29	7	6	5	-120
30	8	7	6	-150

Індивідуальне завдання №23. Виконати індивідуальне завдання ***№22*** методом умовних похідних (методом Якобі).

Індивідуальне завдання №24. Виконати індивідуальне завдання ***№22*** методом невизначених множників Лагранжа.

Індивідуальне завдання №25. Виконати індивідуальне завдання ***№22*** методом покоординатного спуску в просторі незалежних змінних із точністю обчислення $\varepsilon = 0{,}5$. Як початкову точку наближення взяти $\mathbf{x}^{(0)\,\mathrm{T}} = \begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix}$, а як залежні змінні – x_1 і x_2. Процес мінімізації обмежити двома кроками. Завдання завершити висновком, в якому зазначити, чи є необхідність у третьому кроці мінімізації.

Розділ 7. ОПТИМІЗАЦІЯ ПРИ ДВОСТОРОННІЙ ОБМЕЖЕНОСТІ ЗМІННИХ

У даному розділі розглядається ще одна задача умовної оптимізації – задача оптимізації при двосторонній обмеженості змінних. Як і задача оптимізації при обмеженнях у вигляді рівностей, вона є окремим випадком загальної задачі математичного програмування. Таку задачу часто називають задачею оптимізації на гіперпаралелепіпеді, оскільки областю припустимих рішень є точкова множина у вигляді гіперпаралелепіпеда. З методологічної точки зору розгляд такої задачі, як і розгляд класичної задачі оптимізації, є дуже важливим етапом у процесі вивчення дисципліни «Математичне програмування». Цей розділ, як і попередній, з одного боку, дозволяє виявити особливості і методи вирішення цілого класу задач умовної оптимізації, а з другого – закладає фундаментальні основи для розуміння більш складних задач, відомих як загальні задачі математичного програмування.

7.1. Постановка задачі оптимізації при двосторонній обмеженості змінних

Задача оптимізації при двосторонній обмеженості змінних формулюється в такий спосіб: *знайти мінімум функції* $y(\mathbf{x})$*, визначеної на області* $\mathbf{\Omega} \subset \mathbf{R}^n$*, де областю припустимих рішень* $\mathbf{\Omega}$ *є гіперпаралелепіпед.*

Аналітичний запис цієї задачі має вигляд

$$y(\mathbf{x}) \to \underset{\mathbf{x} \in \mathbf{\Omega} \subset \mathbf{R}^n}{\text{opt}} , \tag{7.1}$$

$$\mathbf{\Omega}: \quad \mathbf{x}^+ \le \mathbf{x} \le \mathbf{x}^{++} . \tag{7.2}$$

Вважатимемо, що $y(\mathbf{x})$ – неперервна нелінійна функція дійсних змінних, що має перші й другі похідні по всіх змінних x_i, $i = \overline{1,n}$, а припустима множина рішень задачі (7.2) формується з вектора нижніх границь змінних $\mathbf{x}^{+}$ і вектора верхніх границь змінних $\mathbf{x}^{++}$.

Нагадаємо, що гіперпаралелепіпедом (7.2) у задачі одновимірної оптимізації виступає відрізок прямої, двовимірної – ортогональний прямокутник, тривимірної – ортогональний паралелепіпед.

Якщо області рішень $\mathbf{R}^{\mathrm{n}}$ (n-вимірний евклідовий простір) у задачі безумовної оптимізації і область $\mathbf{R}^{\mathrm{p}}$ (p-вимірний простір незалежних змінних) у класичній задачі оптимізації були необмеженими, то область припустимих рішень $\mathbf{\Omega}$ у задачі оптимізації на гіперпаралелепіпеді є замкнутою через замкнутість гіперпаралелепіпеда. Оптимальне рішення в такій задачі може знаходитися як всередині припустимої області, так і на її границі. При цьому стаціонарні точки, а з ними і явні екстремуми функції $y(\mathbf{x})$, можуть не потрапляти на область $\mathbf{\Omega}$ або взагалі бути відсутніми. Саме з цієї причини всі класичні методи, засновані на попередньому визначенні стаціонарних точок, в загальному випадку непридатні для розв'язання задачі оптимізації при двосторонніх обмеженнях змінних. Класичні методи дозволяють правильно вирішити її тільки в одному випадку, коли явний шуканий екстремум знаходиться на області $\mathbf{\Omega}$, всередині його або на границі. Однак таку ситуацію заздалегідь передбачати неможливо.

На рис.7.1 побудована у вигляді еквіпотенціальних рівнів деяка функція $y(\mathbf{x})$ і чотири різні області припустимих рішень $\mathbf{\Omega}_i$, $i=\overline{1,4}$. З рисунку видно, що явний екстремум функції y^* збігається із шуканим тільки на припустимій області $\mathbf{\Omega}_2$. Для інших областей явний екстремум знаходиться за їхніми межами, тому шукані оптимальні рішення лежать на границях цих областей. При цьому для області $\mathbf{\Omega}_1$ оптимальне рішення по обох змінних є граничним; для $\mathbf{\Omega}_3$ – тільки по змінній x_2; для області $\mathbf{\Omega}_4$ – тільки по змінній x_1.

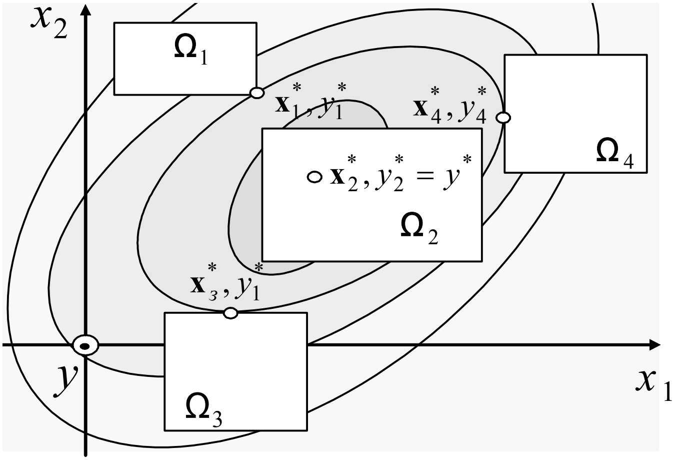


Рис.7.1 – Приклади задач оптимізації при двосторонній обмеженості змінних із позначенням оптимальних рішень

Оскільки заздалегідь не можна передбачити, чи буде оптимальне рішення внутрішньою точкою області припустимих рішень (точка $\mathbf{x}_2$ на рис.7.1) або граничною (точки $\mathbf{x}_1, \mathbf{x}_3, \mathbf{x}_4$ на рис.7.1), при виведенні необхідних і достатніх умов локального оптимуму треба враховувати всі можливі ситуації.

7.2. Необхідні умови для точки локального оптимуму

Виведення необхідних умов для точки локального оптимуму в задачі оптимізації при двосторонній обмеженості змінних, за аналогією з виведенням у попередніх задачах, почнемо з розкладання цільової функції $y(\mathbf{x})$ в ряд Тейлора в околі точки локального мінімуму. При цьому розкладання відразу запишемо в приростах і обмежимося тільки лінійними членами розкладання, вважаючи інші нелінійні члени занадто малими в порівнянні з першими:

$$\Delta y^* = \left(\frac{\partial y}{\partial \mathbf{x}}\right)^{*\mathrm{T}} \Delta\mathbf{x} \quad . \tag{7.3}$$

За визначенням точки локального мінімуму приріст функції в ній не може бути від'ємним:

$$\Delta y^* = \left(\frac{\partial y}{\partial \mathbf{x}}\right)^{*\mathrm{T}} \Delta\mathbf{x} \geq 0 \quad . \tag{7.4}$$

Припустимо, що змінюється тільки одна змінна x_r, яка є однієї зі складових вектора x_r, а інші залишаються незмінними. Тоді приріст (7.4) запишеться у такому вигляді:

$$\Delta y^* = \left(\frac{\partial y}{\partial x_r}\right)^{*\mathrm{T}} \Delta x_r \geq 0 \quad . \tag{7.5}$$

Для змінної x_r існують три варіанти розташування локального мінімуму щодо області її визначення:

- ♦ на своїй нижній границі $x_r^* = x_r^+$ (наприклад, на рис.7.1 змінна x_2 у випадку припустимої області $\mathbf{\Omega}_1$ або x_1 у випадку $\mathbf{\Omega}_4$) ;
- ♦ на своїй верхній границі $x_r^* = x_r^{++}$ (наприклад, на рис.7.1 змінна x_1 у випадку припустимої області $\mathbf{\Omega}_1$ або x_2 у випадку $\mathbf{\Omega}_3$) ;
- ♦ усередині області визначення $x_r^+ < x_r^* < x_r^{++}$(наприклад, на рис.7.1 змінні x_1 й x_2 у випадку припустимої області $\mathbf{\Omega}_2$, змінна x_1 у випадку $\mathbf{\Omega}_3$ або x_2 в $\mathbf{\Omega}_4$).

Розглянемо кожний варіант окремо.

Нехай мінімум функції по змінній x_r знаходиться на її нижній границі (рис.7.2), тобто $x_r^* = x_r^+$. У цьому випадку приріст аргументу може бути тільки додатним $\Delta x_r > 0$, а перша частинна похідна функції по цій змінній – невід'ємною:

$$\left(\frac{\partial y}{\partial x_r}\right)^* \geq 0 \ . \tag{7.6}$$

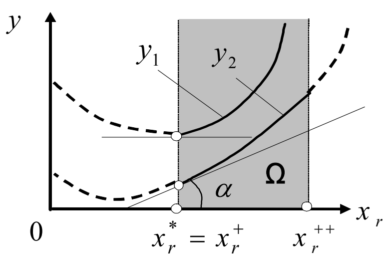


Рис. 7.2.

У протилежному разі буде порушена умова (7.5).

Дотична до графіка функції в точці локального мінімуму або паралельна осі абсцис (рис.7.2, функція y_1), або має додатний нахил $tg\alpha \geq 0$ (рис.7.2, функція y_2).

Нехай мінімум функції по змінній x_r знаходиться на її верхній границі (рис.7.3), тобто $x_r^* = x_r^{++}$. У цьому випадку збільшення аргументу може бути тільки від'ємним $\Delta x_r < 0$, а перша приватна похідна функції $y(\mathbf{x})$ по цій змінній недодатна:

$$\left(\frac{\partial y}{\partial x_r}\right)^* \leq 0 \ . \tag{7.7}$$

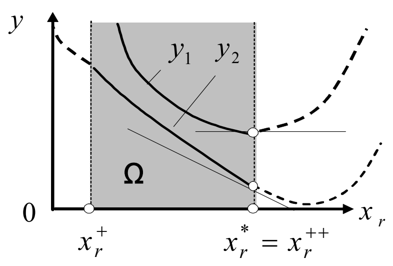


Рис. 7.3.

Дотична до графіка функції в точці локального мінімуму або паралельна осі ординат (рис. 7.3, функція y_1), або має додатний нахил (рис.7.3, функція y_2).

Нехай мінімум функції по змінній x_r знаходиться усередині області її визначення (рис.7.4), тобто $x_r^+ < x_r^* < x_r^{++}$. У цьому випадку (за аналогією з виведенням необхідних умов у задачі безумовної оптимізації) приріст цієї змінної може бути як від'ємним $\Delta x_r < 0$, так і додатним $\Delta x_r > 0$, а перша частинна похідна функції $y(\mathbf{x})$ по цій змінній повинна дорівнювати нулю:

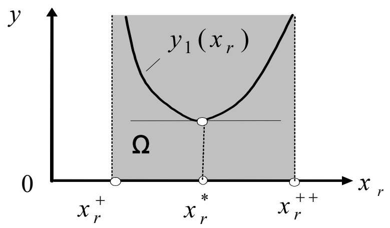


Рис. 7.4.

$$\left(\frac{\partial y}{\partial x_r}\right)^* = 0 \; . \qquad (7.8)$$

У протилежному разі завжди можна підібрати приріст аргументу зі знаком, протилежним знаку похідної, що приведе до порушення умови (7.5).

Оскільки змінна x_r вибиралася довільно, то наведені міркування в однаковій мірі стосуються будь-якої іншої змінної задачі. Узагальнюючи ці міркування на всі змінні, можна стверджувати, що необхідні умови локального мінімуму функції при двосторонній обмеженості змінних полягають у:

- невід'ємності тих складових вектора частинних похідних $\left(\frac{\partial y}{\partial \mathbf{x}}\right)^*$, які відповідають змінним, що знаходяться на своїх нижніх границях;
- недодатності тих складових вектора частинних похідних, які відповідають змінним, що знаходяться на своїх верхніх границях;
- рівності нулю тих компонент вектора частинних похідних, що відповідають змінним, які лежать усередині області визначення.

Отже, необхідні умови локального мінімуму формулюються таким чином:

$$\boxed{\begin{aligned} &\left(\frac{\partial y}{\partial x_r}\right)^{*} \geq 0, \quad \text{если} \quad x_r^{*} = x_r^{+}; \qquad (7.9)\\ &\left(\frac{\partial y}{\partial x_r}\right)^{*} \leq 0, \quad \text{если} \quad x_r^{*} = x_r^{++}; \qquad (7.10)\\ &\left(\frac{\partial y}{\partial x_r}\right)^{*} = 0, \quad \text{если} \quad x_r^{+} < x_r^{*} < x_r^{++}. \qquad (7.11)\end{aligned}}$$

Необхідні умови для точки локального максимуму відрізняються від (7.9) – (7.11) тільки знаками нерівностей: знак «$\geq$» у першому виразі стає знаком «$\leq$», а в другому виразі, навпаки, – знак «$\leq$» стає знаком «$\geq$».

7.3. Достатні умови для точки локального оптимуму

Залежно від того, як розташовується область припустимих рішень (гіперпаралелепіпед) відносно явного екстремуму функції, достатні умови формулюються по-різному.

Перша група достатніх умов для точки локального мінімуму формулюється таким чином: якщо всі складові вектора $\mathbf{x}^{\circ}$ знаходяться на своїх границях (однаково, на яких: на нижніх або на верхніх), жодна з похідних $\left(\frac{\partial y}{\partial x_i}\right)^{\circ}$ не дорівнює нулю і при цьому виконуються необхідні умови мінімуму (7.9):

$$\left(\frac{\partial y}{\partial x_i}\right)^{\circ} > 0, \quad \text{якщо} \quad x_i^{\circ} = x_i^{+} \ ; \tag{7.12}$$

$$\left(\frac{\partial y}{\partial x_i}\right)^{*} < 0, \quad \text{якщо} \quad x_i^{\circ} = x_i^{++} \ , \tag{7.13}$$

то ці умови не тільки необхідні, але і достатні. Тому $\mathbf{x}^{\circ} = \mathbf{x}^{*}$.

Для доведення першої групи достатніх умов доповнимо ряд розкладання функції цілі $y(\mathbf{x})$ (7.1) членами другого і більш високого порядків у вигляді залишку $R\left(\Delta\mathbf{x}^2, \Delta\mathbf{x}^3, \ldots\right)$. Тоді розкладання функції $y(\mathbf{x})$ в точці $\mathbf{x}^{\circ}$ набуває вигляду

$$\Delta y^{\bullet} = \left(\frac{\partial y}{\partial \mathbf{x}}\right)^{\circ\mathrm{T}} \Delta\mathbf{x} + R^{\circ}\left(\Delta\mathbf{x}^2, \Delta\mathbf{x}^3, \ldots\right). \tag{7.14}$$

Тут $R^{\circ}\left(\Delta\mathbf{x}^2, \Delta\mathbf{x}^3, \ldots\right)$ являє собою суму нелінійних членів розкладання, кожний з яких залежить від добутку двох або більше приростів Δx_i .

Якщо необхідні умови виконуються, то лінійний член розкладання $\left(\frac{\partial y}{\partial \mathbf{x}}\right)^{\circ\mathrm{T}} \Delta\mathbf{x}$ являє собою суму позитивних доданків $\left(\frac{\partial y}{\partial x_i}\right)^{\circ} \Delta x_i$. Це дійсно так, оскільки додатним похідним $\left(\frac{\partial y}{\partial x_i}\right)^{\circ}$ можуть відповідати тільки додатні прирости Δx_i, а від'ємним – від'ємні. Природно, сума додатних значень не може бути рівної нулю. А оскільки $\left(\frac{\partial y}{\partial \mathbf{x}}\right)^{\circ} \neq 0$, то

$\lim\limits_{\Delta\mathbf{x}\to 0} \dfrac{R^{\circ}\left(\Delta\mathbf{x}^2, \Delta\mathbf{x}^3, \ldots\right)}{\left(\dfrac{\partial\ y}{\partial\ \mathbf{x}}\right)^{\circ \mathrm{T}} \Delta\mathbf{x}} = 0$, і залишком у (7.14) можна зневажати. У результаті одержимо розкладання

$$\Delta y^{\bullet} = \left(\frac{\partial\ y}{\partial\ \mathbf{x}}\right)^{\circ \mathrm{T}} \Delta\mathbf{x}\ , \tag{7.15}$$

з якого випливає, що в будь-якому малому припустимому околі точки $\mathbf{x}^{\circ}$ приріст функції $\Delta y^{\bullet}$ при дотриманні умов (7.12) – (7.13) завжди більше нуля. Отже, умови (7.12) – (7.13) є як необхідними, так і достатніми.

Друга група достатніх умов має місце, коли в точці шуканого локального мінімуму $\mathbf{x}^{*}$ частинні похідні цільової функції по деяких (або по всіх) змінним рівні нулю. У цьому випадку другим членом розкладання (7.14) при виведенні достатніх умов зневажати не можна.

Для доведення другої групи достатніх умов розіб'ємо вектор $\mathbf{x}^{*}$ на два складових підвектора $\mathbf{x}_1^{*}$ і $\mathbf{x}_2^{*}$ так, щоб складовим x_{1i}^{*} підвектора $\mathbf{x}_1^{*}$ відповідали похідні $\left(\dfrac{\partial\, y}{\partial\, x_{1i}}\right)^{*} \neq 0$, а складовим x_{2i}^{*} підвектора $\mathbf{x}_2^{*}$ – похідні $\left(\dfrac{\partial\, y}{\partial\, x_{2i}}\right)^{*} = 0$. Вважатимемо, що підвектор $\mathbf{x}_1^{*}$ містить перші n_1 компонент вектора $\mathbf{x}^{*}$, а вектор $\mathbf{x}_2^{*}$ – інші $n_2 = n - n_1$. У протилежному разі відповідна перенумерація змінних завжди дозволить домогтися зазначеної розбивки.

Розкладемо функцію $y(\mathbf{x})$ в околі точки мінімуму в ряд Тейлора з урахуванням розбивки змінних і обмежуючи розкладання лінійними і квадратичними членами:

$$\Delta y^{*}=\left(\frac{\partial y}{\partial \mathbf{x}_1}\right)^{*\mathrm{T}}\Delta\mathbf{x}_1+\left(\frac{\partial y}{\partial \mathbf{x}_2}\right)^{*\mathrm{T}}\Delta\mathbf{x}_2+\frac{1}{2}\Delta\mathbf{x}_1^{\mathrm{T}}\mathbf{H}_{11}^{*}\Delta\mathbf{x}_1+$$
$$+\frac{1}{2}\Delta\mathbf{x}_1^{\mathrm{T}}\mathbf{H}_{12}^{*}\Delta\mathbf{x}_2+\frac{1}{2}\Delta\mathbf{x}_2^{\mathrm{T}}\mathbf{H}_{21}^{*}\Delta\mathbf{x}_1+\frac{1}{2}\Delta\mathbf{x}_2^{\mathrm{T}}\mathbf{H}_{22}^{*}\Delta\mathbf{x}_2\ , \qquad (7.16)$$

де $\mathbf{H}_{11},\mathbf{H}_{12},\mathbf{H}_{21},\mathbf{H}_{22}$ – підматриці матриці Гесса відповідно до розбивки змінних.

Нехай в околі $\mathbf{x}^{*}$ змінюються тільки складові вектора $\mathbf{x}_1^{*}$, тобто $\Delta\mathbf{x}_2=0$. Тоді

$$\Delta y^{*}=\left(\frac{\partial y}{\partial \mathbf{x}_1}\right)^{*\mathrm{T}}\Delta\mathbf{x}_1+\frac{1}{2}\Delta\mathbf{x}_1^{\mathrm{T}}\mathbf{H}_{11}^{*}\Delta\mathbf{x}_1\ . \qquad (7.17)$$

Оскільки $\left(\frac{\partial y}{\partial \mathbf{x}_1}\right)^{*}\neq 0$, то другим членом розкладання (7.17) можна зневажати. Якщо тепер міркувати таким же способом, як і при виведенні першої групи достатніх умов, то прийдемо до аналогічного висновку: *необхідні умови щодо складових підвектора* $\mathbf{x}_1^{*}$ *одночасно є достатніми.*

Нехай в околі $\mathbf{x}^{*}$ змінюються тільки складові підвектора $\mathbf{x}_2^{*}$, тобто $\Delta\mathbf{x}_1=0$. Тоді

$$\Delta y^{*}=\frac{1}{2}\Delta\mathbf{x}_2^{\mathrm{T}}\mathbf{H}_{22}^{*}\Delta\mathbf{x}_2\ . \qquad (7.18)$$

Тут лінійний член розкладання відсутній, оскільки за умовою розбивки всі складові вектори $\left(\frac{\partial y}{\partial \mathbf{x}_2}\right)^{*}$ рівні нулю.

Оскільки розкладання робилося в точці мінімуму, останній вираз (7.18) повинен бути строго більше нуля. Якщо допустити, що він дорівнює нулю, то в цьому випадку не можна зневажати членами розкладання більш високого порядку. Вираз (7.18) буде більше нуля тільки у випадку, якщо підматриця

$\mathbf{H}_{22}^{*}$ додатно визначена.

Таким чином, *друга група достатніх умов для точки локального мінімуму в задачі мінімізації при двосторонній обмеженості змінних полягає у виконанні необхідних умов і додатній визначеності підматриці Гесса* $\mathbf{H}_{22}^{*}$*, складеної для тих змінних* x_i^{*}*, відповідні похідні яких* $\left(\dfrac{\partial y}{\partial x_i}\right)^{*}$ *рівні нулю.*

7.4. Диференціальний алгоритм

Як було встановлено раніше, класичні методи розв’язання оптимізаційних задач, які засновані на властивості явних екстремумів мати нульові частині похідні, стають непридатними для пошуку оптимальних рішень в задачах, у яких явний екстремум на припустимій області $\boldsymbol{\Omega}$ може бути відсутнім. У цьому випадку хоча б одна частинна похідна відмінна від нуля, і пошук рішення здійснюється за допомогою прямих методів. Найбільш простим і досить ефективним методом розв’язання задачі оптимізації при обмеженнях на змінні задачі у вигляді гіперпаралелепіпеда є метод покоординатного спуску, який в даному випадку має більш складний алгоритм пошуку рішення. Останній одержав у математичній літературі назву «диференціального алгоритму».

Розв’язання задачі мінімізації при двосторонній обмеженості змінних за диференціальним алгоритмом робиться за тією схемою, що й в задачі безумовної мінімізації. Перед початком пошуку оптимального рішення задаються початковою точкою наближення $\mathbf{x}^{(0)}$, яка тепер обов'язково повинна належати області припустимих рішень $\boldsymbol{\Omega}$, і точністю обчислення ε. На кожному кроці пошуку рішення змінюється тільки одна змінна x_r, а інші залишаються незмінними. Змінна x_r вибирається в результаті аналізу виконання необхідних умов (7.9) – (7.11) у поточній k-й точці наближення, $k = 0,1,2,...$ Якщо за змінною x_r необхідні умови (7.9) – (7.11) яким-небудь чином порушені, то, змінюючи її у певному напрямку, можна поліпшити (зменшити) значення функції цілі, залишаючись у припустимій $\boldsymbol{\Omega}$ області .

Аналіз необхідних умов (7.9) – (7.11) показує, що в поточній точці наближення $\mathbf{x}^{(k)}$ можуть мати місце тільки два випадки їхніх порушень:

$$\left(\frac{\partial y}{\partial x_r}\right)^{(k)} > 0, \qquad x_r^{(k)} > x_r^{+}; \tag{7.19}$$

$$\left(\frac{\partial y}{\partial x_r}\right)^{(k)} < 0, \qquad x_r^{(k)} < x_r^{++}. \tag{7.20}$$

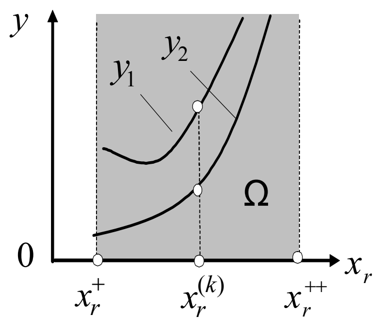


Рис.7.5.

На рис.7.5 наведені графіки функцій y_1 і y_2 однієї змінної x_r, для котрих в точці $x_r^{(k)}$ має місце порушення необхідних умов першого типу (7.19). З рисунку видно, що, надаючи змінній $x_r^{(k)}$ від'ємний приріст Δx_r, можна зменшити значення функцій y_1 і y_2.

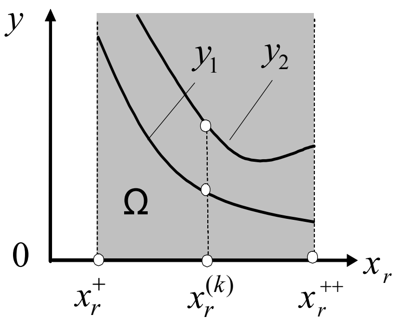


Рис. 7.6.

На рис.7.6 наведені приклади порушення необхідних умов другого типу (7.20). З рисунку видно, що, надаючи тепер змінній $x_r^{(k)}$ додатний приріст Δx_r, можна також зменшити значення функцій y_1 і y_2.

У кожному з розглянутих прикладів на рис.7.5 і рис.7.6 можна зробити гарантоване зменшення функції. При цьому знаки похідних і знаки збільшення аргументу протилежні.

Таким чином, порушення необхідних умов (7.19) і (7.20) дозволяють вибрати змінну і напрямок руху уздовж відповідної осі в області Ω, що приводить до зменшення функції, а саме:

$$\text{якщо } \left(\frac{\partial y}{\partial x_r}\right)^{(k)} > 0 \text{ и } x_r^{(k)} > x_r^{+}\text{, то } \Delta x_r < 0; \tag{7.21}$$

$$\text{якщо } \left(\frac{\partial y}{\partial x_r}\right)^{(k)} < 0 \text{ и } x_r^{(k)} < x_r^{++}\text{, то } \Delta x_r > 0. \tag{7.22}$$

Після вибору напрямку руху необхідно визначити довжину кроку в цьому напрямку. При рухові до точки мінімуму можуть відбутися дві різні події: зустріч з точкою явного мінімуму по змінній x_r (див. функцію y_1 на рис.7.5 і функцію y_2 на рис.7.6) або зустріч з границею (див. функцію y_2 на рис.7.5 і функцію y_1 на рис.7.6). Для правильно вибору довжини кроку треба визначити, яка з двох розглянутих подій настане раніше. Тому шукана довжина кроку (приріст варійованої змінної x_r) визначається за допомогою спеціальних критеріїв.

Якщо має місце перший тип порушень (7.19), то шуканий приріст від'ємний, і рух по змінній x_r відбувається убік її нижньої границі x_r^{+}. У цьому випадку критерій вибору довжини кроку має вигляд

$$\Delta x_r^{(k)} = \max_{\Delta x<0}\left\{ x_r^{+} - x_r^{(k)};\ \left. x_r^{(k)}\right|_{\frac{\partial y}{\partial x_r}=0} \right\}. \tag{7.23}$$

Критерій (7.23) потребує визначення двох довжин. Перша $\left(x_r^{+} - x_r^{(k)}\right)$ – це відстань до нижньої границі, друга $\left. x_r^{(k)}\right|_{\frac{\partial y}{\partial x_r}=0}$ – це довжина, що обчислюється з умови обертання в нуль першої частинної похідної, тобто відстань до явного мінімуму. Обидві довжини від'ємні. Максимальна з них (менша за абсолютним значенням) визначає ту подію, яка відбувається раніше і, отже, відповідає шуканому приросту $\Delta x_r^{(k)}$.

Якщо має місце другий тип порушень (7.20), то шуканий приріст додатний, і рух по змінній x_r відбувається убік її верхньої границі x_r^{++}. У цьому випадку критерій вибору довжини кроку має вигляд:

$$\Delta x_r^{(k)} = \min_{\Delta x > 0} \left\{ x_r^{++} - x_r^{(k)};\ \ x_r^{(k)} \Big|_{\frac{\partial y}{\partial x_r} = 0} \right\}. \tag{7.24}$$

Розмір $x_r^{(k)} \Big|_{\frac{\partial y}{\partial x_r} = 0}$ в обох виразах (7.23) і (7.24) варто обчислювати за аналогією з методом покоординатного спуску, який розглядався в задачі безумовної мінімізації. Тоді критерії вибору приросту на *k*-му кроці наближення набуватимуть вигляду:

$$\Delta x_r^{(k)} = \max_{\Delta x < 0} \left\{ x_r^{+} - x_r^{(k)};\quad -\frac{\left(\frac{\partial y}{\partial x_r}\right)^{(k)}}{\left|\left(\frac{\partial^2 y}{\partial x_r^2}\right)^{(k)}\right|} \right\}; \tag{7.25}$$

$$\Delta x_r^{(k)} = \min_{\Delta x > 0} \left\{ x_r^{++} - x_r^{(k)};\quad -\frac{\left(\frac{\partial y}{\partial x_r}\right)^{(k)}}{\left|\left(\frac{\partial^2 y}{\partial x_r^2}\right)^{(k)}\right|} \right\}. \tag{7.26}$$

7.5. Приклади вирішення задач оптимізації за диференціальним алгоритмом

Розглянемо диференціальний алгоритм вирішення задачі мінімізації при двосторонній обмеженості змінних на конкретному прикладі. Цикл із n кроків, що включає аналіз необхідних умов по кожній змінній і обчислення при необхідності відповідного приросту, будемо називати ітерацією.

Приклад 7.1. Знайти мінімум функції $y = x_1 x_2 + x_1^2 + x_2^2$, якщо початкова точка наближення $\mathbf{x}^{(\circ)\mathrm{T}} = \left[70 \quad -40\right]$, точність обчислення ε=0,1, а область припустимих рішень подана точковою множиною у вигляді двовимірного гіперпаралелепіпеда $\mathbf{\Omega}: \begin{bmatrix} 50 \\ -50 \end{bmatrix} \le \mathbf{x} \le \begin{bmatrix} 90 \\ 20 \end{bmatrix}$.

Розв'язання. *1-та ітерація. 1-й крок.* Обчислимо першу похідну функції y за змінною x_1 в початковій точці наближення: $\left(\dfrac{\partial y}{\partial x_1}\right)^{(\circ)} = x_2 + 2x_1 = 150 > 0$. Оскільки похідна менше заданої точності і додатна, а відповідна змінна не лежить на нижній границі, то має місце порушення необхідних умов мінімуму типу (7.19). Тому приріст змінної x_1, відповідно до (7.21), повинен бути від'ємним і обчислюватися за виразом (7.25):

$$\Delta x_1^{(0)} = \max_{\Delta x < 0}\left\{50 - 70; \quad -\tfrac{150}{2}\right\} = -20 .$$

Надамо змінній x_1 обчислений приріст: $x_1^{(1)} = x_1^{(0)} + \Delta x_1^{(0)} = = 70 - 20 = 50$. Переходимо в нову точку наближення $\bar{x}^{(1)\square} = \left[50 \quad -40\right]$, в якій значення функції $y^{(1)} = 3100$ менше, ніж у початковій точці наближення $y^{(1)} = 3700$.

1-ша ітерація. 2-й крок. Обчислимо першу похідну цільової функції по змінній x_2 в новій точці наближення: $\left(\dfrac{\partial y}{\partial x_2}\right)^{(1)} = x_1 + 2x_2 = -30 < 0.$

Оскільки похідна по змінній x_2 від'ємна і за абсолютним значенням перевершує задану точність, а сама змінна не лежить на верхній границі, то має місце порушення необхідних умов типу (7.20). Отже, відповідно до (7.22) приріст змінної x_2 повинен бути додатним і обчислюватися за критерієм (7.26):

$$\Delta x_2^{(1)} = \min_{\Delta x > 0}\left\{20 - (-40); \quad -\tfrac{30}{2}\right\} = -35 \ .$$

Нове значення змінної $x_2^{(2)} = x_2^{(1)} + \Delta x_2^{(1)} = -40 + 15 = -25$. Нова точка наближення $\mathbf{x}^{(2)^{\mathrm{T}}} = \begin{bmatrix} 50 & -25 \end{bmatrix}$. Функція стала ще менше $y^{(2)} = 1875$.

2-а ітерація. 3-й крок. Необхідні умови в точці наближення $\mathbf{x}^{(2)^{\mathrm{T}}} = \begin{bmatrix} 50 & -25 \end{bmatrix}$ по змінній x_1 виконуються: $\left(\dfrac{\partial y}{\partial x_1}\right)^{(2)} = x_2 + 2x_1 = 75 > 0,$

а сама змінна знаходиться на нижній границі $x_1^{(2)} = x_1^{+} = 50$. Тому переходимо до наступного кроку, залишаючись у точці наближення $\mathbf{x}^{(2)}$, тобто нова точка наближення $\mathbf{x}^{(3)} = \mathbf{x}^{(2)}$.

2-а ітерація. 4-й крок. Необхідні умови в точці наближення $\mathbf{x}^{(3)^{\mathrm{T}}} = \begin{bmatrix} 50 & -25 \end{bmatrix}$ по змінній x_2 також виконуються: $\left(\dfrac{\partial y}{\partial x_2}\right)^{(2)} = x_1 + 2x_2 = 0.$

Звідси випливає, що $\mathbf{x}^{(3)^{\mathrm{T}}} = \begin{bmatrix} 50 & -25 \end{bmatrix} = \mathbf{x}^*$. У цій точці функція y на заданому гіперпаралелепіпеді має найменше значення: y^*=1875. Цей висновок підтверджується достатніми умовами, що в умовах прикладу вироджуються в невід'ємність другої частинної похідної функції по змінній x_2: $\left(\frac{\partial^2 y}{\partial x_2^2}\right)^{(2)} = 2 > 0$. Задачу вирішено.

Приклад 7.2. Знайти мінімальне рішення і мінімум функції $y(\mathbf{x}) = 4x_1^3 + 2x_1^2 x_2^2 + x_2^2 x_3 + 2x_3^2$, якщо область припустимих рішень подана як Ω: $\begin{bmatrix} 1 \\ -1 \\ -2 \end{bmatrix} \le \mathbf{x} \le \begin{bmatrix} 2 \\ 2 \\ 0 \end{bmatrix}$. Розв'язання здійснити за допомогою диференціального алгоритму при точності обчислення ε=0,1.

Розв'язання. Початкова точка наближення умовою задачі не задана. Виберемо її довільно з області припустимих рішень **Ш**. Нехай $\mathbf{x}^{(\circ)\mathrm{T}} = [70 \;\; \text{-}40]$.

1-ша ітерація. 1-й крок. У початковій точці наближення $\mathbf{x}^{(\circ)}$ функція цілі приймає значення $y^{(\circ)} = 64$. На першому кроці як варійована змінна виступає x_1. Перевіряємо виконання необхідних умов (7.9 (7.11) по цій змінній в точці $\mathbf{x}^{(\circ)}$. Оскільки $\left(\frac{\partial y}{\partial x_1}\right)^{(\circ)} = 80 > 0$ і $\left|\left(\frac{\partial y}{\partial x_1}\right)^{(\circ)}\right| = 80 > \varepsilon = 0{,}1$, а $x_1^{(0)} > x^+$, має місце порушення типу (7.19). Отже, значення функції можна поліпшити (зменшити), якщо змінній x_1 надати від'ємний приріст. Згідно з критерієм (7.25)

$$\Delta x_1^{(0)} = \max_{\Delta x<0}\left\{ x_1^+ - x_1^{(0)}; \quad -\frac{\left(\frac{\partial y}{\partial x_1}\right)^{(0)}}{\left|\left(\frac{\partial^2 y}{\partial x_1^2}\right)^{(0)}\right|} \right\} = \max_{\Delta x<0}\left\{1-2; \quad -\frac{80}{64}\right\} = 1 \ .$$

Знаходимо нове значення змінної x_1 з урахуванням обчисленого приросту: $x_1^{(1)} = x_1^{(0)} - \Delta x_1^{(0)} = 2-1=1$. Значення змінних x_2 і x_3 залишилися колишніми. Значення функції в новій точці наближення $\mathbf{x}^{(1)^{\mathrm{T}}} = \begin{bmatrix}1 & 2 & -2\end{bmatrix}$ зменшилося: $y^{(1)} = 12 < y^{(0)} = 64$, що свідчить про правильність зробленого кроку.

2-й крок. У новій точці наближення необхідна умова (7.11) по змінній x_2 дотримується: $\left(\dfrac{\partial y}{\partial x_2}\right)^{(1)} = 0, \quad x_2^{(1)} = x_2^+ = 2$. Тому даний крок мінімізації пропускається. Нова точка наближення зберігає координати попередньої: $\mathbf{x}^{(2)} = \mathbf{x}^{(1)}$.

3-й крок. Перевіряємо в точці $\mathbf{x}^{(2)}$ виконання необхідних умов (7.9) – (7.11) по змінній x_3 . Виявляємо порушення умов другого типу (7.20):

$$\left(\frac{\partial y}{\partial x_3}\right)^{(2)} = -4 < 0, \ \left|\left(\frac{\partial y}{\partial x_3}\right)^{(2)}\right| = 4 > \varepsilon = 0{,}1, \ \ x_3^{(2)} = -2 < x_3^{++} = 0 \quad .$$

Отже, функцію можна зменшити, якщо змінній x_3 надати додатний приріст, який обчислюється за критерієм (7.26):

$$\Delta x_3^{(2)} = \min_{\Delta x>0}\left\{ x_3^{++} - x_3^{(2)}; \quad -\frac{\left(\frac{\partial y}{\partial x_3}\right)^{(2)}}{\left|\left(\frac{\partial^2 y}{\partial x_3^2}\right)^{(2)}\right|} \right\} = \max_{\Delta x>0}\left\{ 0-(-2); \quad -\frac{-4}{4} \right\} = 1 \ .$$

У новій точці наближення $\mathbf{x}^{(3)^{\mathrm{T}}} = [1 \quad 2 \quad -1]$ спостерігаємо подальше зменшення функції: $y^{(2)} = 10$. На цьому перша ітерація закінчується.

На наступних ітераціях циклічна процедура варіювання змінними повторюється. Основні проміжні результати наведені в табл.7.1.

На восьмому кроці одержуємо точку $\mathbf{x}^{(8)^{\mathrm{T}}} = [1 \quad 0 \quad 0]$, в якої необхідні умови (7.9) – (7.11) виконуються по всіх змінних. Так, по змінній x_1 виконується умова (7.9) :

$$x_1^{(8)} = 1 = x_1^{+}, \quad \left(\frac{\partial y}{\partial x_1}\right)^{(8)} = 12 > 0 \ ;$$

по змінній x_2 виконується умова (7.11):

$$x_2^{++} < \left(x_2^{(8)} = 0\right) < x_2^{++}, \quad \left(\frac{\partial y}{\partial x_1}\right)^{(8)} = 0 \ ;$$

по змінній x_3 виконується умова (7.10):

$$x_3^{(8)} = 0 = x_3^{++}, \quad \left(\frac{\partial y}{\partial x_3}\right)^{(8)} = 0 \ .$$

Таблиця 7.1

k	i	$x_1^{(k)}$	$x_2^{(k)}$	$x_3^{(k)}$	$y^{(k)}$	$\left(\frac{\partial y}{\partial x_i}\right)^{(k)}$	$\left(\frac{\partial^2 y}{\partial x_i^2}\right)^{(k)}$	$\Delta x_i^{(k)}$
0	1	2	2	-2	64	80	64	-1
1	2	1	2	-2	12	0	–	–
2	3	1	2	-2	12	-4	4	1
3	1	1	2	-1	10	28	–	–
4	2	1	2	-1	10	4	2	-2
5	3	1	0	-1	6	-4	4	1
6	1	1	0	0	4	12	–	–
7	2	1	0	0	4	0	–	–
8	3	1	0	0	4	0	–	–

На заключному етапі вирішення задачі в отриманій точці $\mathbf{x}^{(8)^{\mathrm{т}}} = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}$ перевіряємо виконання достатніх умов. Оскільки по двох змінних у точці $\mathbf{x}^{(8)}$ частинні похідні дорівнюють нулю, то має місце друга група достатніх умов, що полягає в додатній визначеності матриці других частинних похідних по цих змінних, тобто по x_2 і x_3. Сформуємо зазначену матрицю:

$$\mathbf{H} = \begin{bmatrix} \left(\frac{\partial^2 y}{\partial x_1^2}\right)^{(8)} & \left(\frac{\partial^2 y}{\partial x_1 \partial x_2}\right)^{(8)} \\ \left(\frac{\partial^2 y}{\partial x_2 \partial x_1}\right)^{(8)} & \left(\frac{\partial^2 y}{\partial x_2^2}\right)^{(8)} \end{bmatrix} = \begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix}. \qquad (7.27)$$

Згідно з критерієм Сільвестра (головні визначники аналізованої матриці повинні бути строго додатні) матриця (7.27) додатно визначена. А це значить, що точка $\mathbf{x}^{(8)} = \mathbf{x}^*$ дійсно є шуканим рішенням:, а значення $y^{(8)} = y^* = 4$ – шуканим мінімумом функції.

7.6. Контрольні запитання і вправи

1. Сформулювати задачу мінімізації при двосторонній обмеженості змінних?

2. У чому полягають необхідні умови для точки локального мінімуму в задачі мінімізації функції при обмеженнях на змінні у вигляді гіперпаралелепіпеда?

3. У чому полягають необхідні умови для точки локального максимуму в задачі максимізації функції при обмеженнях на змінні у вигляді гіперпаралелепіпеда?

4. Який знак може мати приріст змінної x_r, якщо остання знаходиться на своїй:

а) нижній границі;

б) верхній границі?

5. У чому полягає перша група достатніх умов для точки локального оптимуму в задачі оптимізації при двосторонній обмеженості змінних?

6. У чому полягає друга група достатніх умов в задачі оптимізації при двосторонній обмеженості змінних? Для точки локального максимуму.?

7. Вивести достатні умови для точки локального максимуму в задачі максимізації при обмеженнях у вигляді гіперпаралелепіпеда.

8. При яких умовах значення функції $y(\mathbf{x})$, заданої на гіперпаралелепіпеді, може бути поліпшено, якщо поточна точка її визначення $\mathbf{x}^{(k)}$ є внутрішньою точкою гіперпаралелепіпеда?

9. Які типи порушень необхідних умов для точки локального мінімуму можуть мати місце при вирішенні задач оптимізації на гіперпаралелепіпеді?

10. Як вибирається напрямок поліпшення функції при порушенні необхідних умов локального мінімуму? Максимуму?

11. Сформулювати правила (критерії) вибору приросту змінної при пошуку локального мінімуму за диференціальним алгоритмом.

12. Знайти за допомогою диференціального алгоритму локальний мінімум функції в задачі

$$y(\mathbf{x}) = 4x_1^3 + 2x_1^2 x_2^2 + x_2^2 x_3 + 2x_3^2 \to \min_{\mathbf{x}\in\boldsymbol{\Omega}},$$

$$\boldsymbol{\Omega}: \begin{bmatrix} -1 \\ -1 \\ -1 \end{bmatrix} \le \mathbf{x} \le \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

при точності обчислення $\varepsilon = 0{,}01$, взявши як початкове наближення точку $\mathbf{x}^{(0)\,\mathrm{T}} = [1 \quad 1 \quad 1]$.

13. Знайти оптимальне рішення і мінімум функції в задачі

$$y(\mathbf{x}) = x_1 x_2 + \frac{1}{2x_1^2} + \frac{1}{2x_2^2} \to \min_{\mathbf{x}\in\boldsymbol{\Omega}},$$

$$\boldsymbol{\Omega}: \begin{bmatrix} 0{,}5 \\ 5 \end{bmatrix} \le \mathbf{x} \le \begin{bmatrix} 10 \\ 10 \end{bmatrix}$$

з точністю обчислення $\varepsilon = 0{,}1$ і точкою початкового наближення $\mathbf{x}^{(0)\,\mathrm{T}} = [10 \quad 10]$.

14. Вирішити задачу пошуку локального мінімуму

$$y(\mathbf{x}) = x_1^3 x_2 - \frac{2x_2^2}{x_1} \to \min_{\mathbf{x}\in\boldsymbol{\Omega}},$$

$$\boldsymbol{\Omega}: \begin{bmatrix} -3 \\ -5 \end{bmatrix} \le \mathbf{x} \le \begin{bmatrix} 5 \\ -2 \end{bmatrix}.$$

Пошук почати з точки $\mathbf{x}^{(0)^{\mathrm{T}}} = [2 \quad -2]$.

15. Знайти оптимальне рішення і локальний мінімум функції в задачі

$$y(\mathbf{x}) = 3{,}5x_1^3 + 4{,}5x_2^3 + 6(x_1^2 x_2 + x_1 x_2^2) - 45(x_1 + x_2) \to \min_{\mathbf{x}\in\mathbf{\Omega}}, \quad (7.28)$$

$$\mathbf{\Omega}: \quad \begin{bmatrix} 1 \\ 1{,}5 \end{bmatrix} \le \mathbf{x} \le \begin{bmatrix} 2 \\ 2 \end{bmatrix}$$

при точності обчислення $\varepsilon = 0{,}01$, взявши як початкове наближення точку $\mathbf{x}^{(0)^{\mathrm{T}}} = [2 \quad 2]$.

16. Мінімізувати функцію (7.28) із точністю обчислення $\varepsilon = 0{,}01$ і початкового наближення $\mathbf{x}^{(0)^{\mathrm{T}}} = [0{,}5 \quad 0{,}5]$, якщо область припустимих рішень визначена таким чином:

$$\mathbf{\Omega}: \quad \begin{bmatrix} 0{,}5 \\ 0{,}5 \end{bmatrix} \le \mathbf{x} \le \begin{bmatrix} 2 \\ 3 \end{bmatrix}.$$

17. Вивести необхідні умови для точки локального мінімуму функції при односторонній обмеженості змінних знизу.

18. Вивести необхідні умови для точки локального мінімуму функції при односторонній обмеженості змінних зверху.

19. Вивести необхідні умови для точки локального максимуму функції при односторонній обмеженості змінних знизу.

20. Вивести необхідні умови для точки локального максимуму функції при односторонній обмеженості змінних зверху.

21. Яке порушення необхідних умов може мати місце в задачі пошуку локального мінімуму при односторонній обмеженості змінних знизу?

22. Яке порушення необхідних умов може мати місце в задачі пошуку локального мінімуму при односторонній обмеженості змінних зверху?

23. Яке порушення необхідних умов може мати місце в задачі пошуку локального максимуму при односторонній обмеженості змінних знизу?

24. Яке порушення необхідних умов може мати місце в задачі пошуку локального максимуму при односторонній обмеженості перемінних зверху?

25. Скласти блок-схему диференціального алгоритму вирішення задачі оптимізації при двосторонній обмеженості змінних.

7.7. Фонд індивідуальних завдань

Індивідуальне завдання №26. За допомогою диференціального алгоритму знайти локальний мінімум функції

$$y(\mathbf{x}) = a_1 x_1^3 + a_2 x_2^3 + a_3\left(x_1^2 x_2 + x_1 x_2^2\right) + a_4\left(x_1 + x_2\right) \to \min_{\mathbf{x}\in\mathbf{\Omega}} ,$$

де коефіцієнти a_1, a_2, a_3 і a_4 вибираються з табл.6.3, а область припустимих рішень $\mathbf{\Omega}$:

30.1. $\begin{bmatrix}2\\2\end{bmatrix} \le \mathbf{x} \le \begin{bmatrix}3\\3\end{bmatrix}$;

30.2. $\begin{bmatrix}1\\2\end{bmatrix} \le \mathbf{x} \le \begin{bmatrix}2\\3\end{bmatrix}$;

30.3. $\begin{bmatrix}2\\1\end{bmatrix} \le \mathbf{x} \le \begin{bmatrix}3\\2\end{bmatrix}$;

30.4. $\begin{bmatrix}1\\0{,}5\end{bmatrix} \le \mathbf{x} \le \begin{bmatrix}2\\2\end{bmatrix}$;

30.5. $\begin{bmatrix}0{,}5\\1\end{bmatrix} \le \mathbf{x} \le \begin{bmatrix}1\\1{,}5\end{bmatrix}$;

30.6. $\begin{bmatrix}2\\1\end{bmatrix} \le \mathbf{x} \le \begin{bmatrix}3\\1{,}5\end{bmatrix}$;

30.7. $\begin{bmatrix}1\\1{,}5\end{bmatrix} \le \mathbf{x} \le \begin{bmatrix}2\\2\end{bmatrix}$;

30.8. $\begin{bmatrix}1{,}5\\1\end{bmatrix} \le \mathbf{x} \le \begin{bmatrix}2\\2\end{bmatrix}$;

30.9. $\begin{bmatrix}1\\0{,}5\end{bmatrix} \le \mathbf{x} \le \begin{bmatrix}2\\1\end{bmatrix}$;

30.10. $\begin{bmatrix}0{,}5\\1\end{bmatrix} \le \mathbf{x} \le \begin{bmatrix}1\\1{,}5\end{bmatrix}$;

30.11. $\begin{bmatrix}1\\0{,}5\end{bmatrix} \le \mathbf{x} \le \begin{bmatrix}2\\2\end{bmatrix}$;

30.12. $\begin{bmatrix}1\\1{,}5\end{bmatrix} \le \mathbf{x} \le \begin{bmatrix}1{,}5\\2\end{bmatrix}$;

30.13. $\begin{bmatrix}1\\1{,}5\end{bmatrix} \le \mathbf{x} \le \begin{bmatrix}1{,}5\\2{,}5\end{bmatrix}$;

30.14. $\begin{bmatrix}1{,}5\\1\end{bmatrix} \le \mathbf{x} \le \begin{bmatrix}2\\1{,}5\end{bmatrix}$;

30.15. $\begin{bmatrix}1\\1{,}5\end{bmatrix} \le \mathbf{x} \le \begin{bmatrix}1{,}5\\2\end{bmatrix}$;

30.16. $\begin{bmatrix}0{,}5\\0{,}5\end{bmatrix} \le \mathbf{x} \le \begin{bmatrix}1\\1\end{bmatrix}$;

30.17. $\begin{bmatrix}0{,}5\\1\end{bmatrix} \le \mathbf{x} \le \begin{bmatrix}1\\1{,}5\end{bmatrix}$;

30.18. $\begin{bmatrix}1\\1{,}5\end{bmatrix} \le \mathbf{x} \le \begin{bmatrix}2\\2\end{bmatrix}$;

30.19. $\begin{bmatrix}1{,}5\\1{,}5\end{bmatrix} \le \mathbf{x} \le \begin{bmatrix}2\\2\end{bmatrix}$;

30.20. $\begin{bmatrix}1\\1\end{bmatrix} \le \mathbf{x} \le \begin{bmatrix}2\\2\end{bmatrix}$;

30.21. $\begin{bmatrix}1\\0{,}5\end{bmatrix} \le \mathbf{x} \le \begin{bmatrix}1{,}5\\1\end{bmatrix}$;

30.22. $\begin{bmatrix}1\\0{,}8\end{bmatrix} \le \mathbf{x} \le \begin{bmatrix}1{,}5\\1\end{bmatrix}$;

30.23. $\begin{bmatrix}1{,}5\\1\end{bmatrix} \le \mathbf{x} \le \begin{bmatrix}2\\1{,}5\end{bmatrix}$;

30.24. $\begin{bmatrix}1\\1{,}8\end{bmatrix} \le \mathbf{x} \le \begin{bmatrix}2\\2{,}5\end{bmatrix}$;

30.25. $\begin{bmatrix}1{,}2\\1\end{bmatrix} \le \mathbf{x} \le \begin{bmatrix}2\\1{,}5\end{bmatrix}$;

30.26. $\begin{bmatrix}1\\1{,}4\end{bmatrix} \le \mathbf{x} \le \begin{bmatrix}1{,}8\\2\end{bmatrix}$;

30.27. $\begin{bmatrix}1{,}5\\0{,}5\end{bmatrix} \le \mathbf{x} \le \begin{bmatrix}2\\1{,}2\end{bmatrix}$;

30.28. $\begin{bmatrix}1\\1{,}2\end{bmatrix} \le \mathbf{x} \le \begin{bmatrix}1{,}5\\1{,}8\end{bmatrix}$;

30.29. $\begin{bmatrix}0{,}5\\0{,}8\end{bmatrix} \le \mathbf{x} \le \begin{bmatrix}1{,}8\\1{,}5\end{bmatrix}$;

30.30. $\begin{bmatrix}0{,}8\\1{,}2\end{bmatrix} \le \mathbf{x} \le \begin{bmatrix}2\\1{,}8\end{bmatrix}$.

Точність обчислення прийняти $\varepsilon = 0{,}1$. Точку початкового наближення вибрати самостійно.

Індивідуальне завдання №27. В умовах попереднього індивідуального завдання знайти локальний максимум функції $y(\mathbf{x})$.

Розділ 8. ЗАГАЛЬНА ЗАДАЧА МАТЕМАТИЧНОГО ПРОГРАМУВАННЯ

У цьому розділі розглядається загальна задача математичного програмування і чисельний метод її вирішення, заснований на пошуку локального оптимуму в просторі незалежних змінних з урахуванням двосторонньої обмеженості залежних і незалежних змінних.

8.1. Постановка загальної задачі математичного програмування

Загальна задача математичного програмування формулюється таким чином: *знайти оптимум функції* $y(\mathbf{x})$*, визначеної на області припустимих рішень* $\boldsymbol{\Omega} \subset \mathbf{R}^n$ *, де* $\boldsymbol{\Omega}$ *формується системою рівностей і нерівностей із накладенням на змінні задачі двосторонньої обмеженості.*

Аналітичний запис цієї задачі має вигляд

$$y(\mathbf{x}) \to \underset{\mathbf{x} \in \boldsymbol{\Omega} \subset \mathbf{R}^n}{\text{opt}} , \tag{8.1}$$

$$\boldsymbol{\Omega}: \quad f_i(\mathbf{x}) = 0, \quad i = \overline{1, m_1} ; \tag{8.2}$$

$$f_k(\mathbf{x}) \geq 0, \quad k = \overline{1, m_2} ; \tag{8.3}$$

$$f_l(\mathbf{x}) \leq 0, \quad l = \overline{1, m_3} ; \tag{8.4}$$

$$x_j^+ \leq x_j \leq x_j^{++}, \quad j = \overline{1, n}, \quad n \geq m_1 + m_2 + m_3 , \tag{8.5}$$

де $y(\mathbf{x}),\, f_j(\mathbf{x}),\, f_k(\mathbf{x})$ і $f_j(\mathbf{x})$ – неперервні нелінійні функції дійсних змінних, що мають перші й другі похідні по всіх змінних x_i, $i=\overline{1,n}$; x_i^{+}, x_i^{++} – відповідно нижня і верхня границі змінної x_i.

Очевидно, що обмеження-нерівність типу "$\le$" можна перетворити на обмеження-рівність додаванням до його лівої частини додаткової додатної змінної, а кожне обмеження-нерівність типу "$\ge$" – в обмеження-рівність відніманням із його лівої частини додаткової додатної змінної. Таким чином, задачі оптимізації (8.1) – (8.5) завжди можна звести до задачі

$$y(\mathbf{x}) \to \underset{\mathbf{x}\in\boldsymbol{\Omega}\subset\mathbf{R}^n}{\mathrm{opt}}\ , \tag{8.6}$$

$$\boldsymbol{\Omega}:\quad f_i(\mathbf{x})=0, \quad i=\overline{1,m}\ ; \tag{8.7}$$

$$x_j^{+} \le x_j \le x_j^{++}, \quad j=\overline{1,n}, \quad n \ge m\ . \tag{8.8}$$

Останнє задача являє собою композицію задач оптимізації при обмеженнях на змінні у вигляді рівностей і гіперпаралелепіпеда. У даній задачі, як і задачі оптимізації на гіперпаралелепіпеді, не можна заздалегідь передбачити, чи буде оптимальне рішення внутрішньою точкою області припустимих рішень або граничною. Оскільки оптимум може розташовуватися у граничній точці, то класичні методи вирішення, засновані на рівності нулю умовних похідних, стають непридатними.

Вирішення загальної задачі математичного програмування здійснюється за допомогою прямих методів. Для обґрунтування одного з таких методів спочатку розглянемо необхідні й достатні умови локального оптимуму. При цьому будемо використовувати досвід, набутий у процесі виведення необхідних і достатніх умов у задачі оптимізації при обмеженнях на змінні у вигляді системи рівностей (6.6) – (6.7) і в задачі оптимізації на гіперпаралелепіпеді (7.1) – (7.2).

8.2. Необхідні умови для точки локального оптимуму

Початок процедури висновку необхідних умов для точки локального оптимуму в загальній задачі математичного програмування будемо здійснювати за тією схемою, що й в задачі оптимізації при обмеженнях у вигляді рівностей. Такий початок дозволить загальну задачу математичного програмування (8.6) – (8.8) звести (шляхом розбивки змінних на залежні й незалежні змінні з наступним вираженням перших через другі) до задачі оптимізації в просторі незалежних змінних, на які накладені умови двосторонньої обмеженості, тобто задачу (8.6) – (8.8) звести до задачі

$$y(\mathbf{t}) \to \underset{\mathbf{t}\in\mathbf{\Omega}^{\mathrm{p}}\subset\mathbf{R}^{\mathrm{p}}}{\mathrm{opt}} , \tag{8.9}$$

$$\mathbf{\Omega}^{\mathrm{p}} : \quad t_j^{+} \le t_j \le t_j^{++}, \quad j = \overline{1, p}, \quad p = n - m \ . \tag{8.10}$$

Отже, зробимо розбивку вектора змінних $\mathbf{x}$ на підвектор залежних $\mathbf{s}$ і підвектор незалежних змінних $\mathbf{t}$ таким чином, щоб точка локального оптимуму $\mathbf{x}^*$ була не тільки неособливою (детермінант матриці Якобі не дорівнює нулю: $\det \mathbf{W} \neq 0$), але і невиродженою. Перше означає, що існує обернена матриця Якобі $\mathbf{W}^{-1}$, останнє – що всі залежні змінні $s_i \ \ (i = \overline{1, m})$ знаходяться строго всередині інтервалу $\left[s_i^{+}, s_i^{++}\right]$, тобто не набувають граничного значення. У цьому випадку досить малі довільні зміни вектора $\mathbf{t}^*$ не приводять до виходу з області припустимих рішень $\mathbf{\Omega}$.

Розкладаючи функцію (8.6) і обмеження задачі (8.7) в ряд Тейлора, як це робилося в розділі 6) і замінюючи приріст підвектора залежних змінних його виразом через приріст підвектора незалежних (6.30), одержимо Δy як функцію тільки вектора Δt (див. вираз (6.35)):

$$\Delta y = \left(\frac{\delta \ y}{\delta \ \mathbf{t}}\right)^{\mathrm{T}} \Delta\mathbf{t} \ . \tag{8.11}$$

Такий приріст має функція при пошуку умовного локального оптимуму в просторі необмежених незалежних змінних, розглянутого в розділі 6. Однак, у випадку загальної задачі математичного програмування область припустимих рішень обмежена гіперпаралелепіпедом (8.10), тому оптимум може знаходитися не тільки всередині інтервалу визначення змінної $\Delta t_r \quad \left(r \in \{1,2,...p\}\right)$, але і на його границі. З огляду на те, що при пошуку локального *мінімуму* приріст $\Delta y^* \geq 0$ для будь-яких малих припустимих змін Δt_r^*, і використовуючи ті ж міркування, що й у задачі мінімізації при двосторонній обмеженості змінних, але тільки щодо незалежних змінних і умовних похідних по цих змінних, одержуємо необхідні умови для точки умовного локального мінімуму в задачі (8.6) – (8.8).

Необхідні умови для точки локального мінімуму $\mathbf{x}^*$ полягають у наступному:

$$\text{если} \quad t_r^* = t_r^+, \quad \text{то} \quad \left(\frac{\delta\, y}{\delta\, t_r}\right)^* \geq 0\,; \tag{8.12}$$

$$\text{если} \quad t_r^* = t_r^{++}, \quad \text{то} \quad \left(\frac{\delta\, y}{\delta\, t_r}\right)^* \leq 0\,; \tag{8.13}$$

$$\text{если} \quad t_r^+ < t_r^* < t_r^+, \quad \text{то} \quad \left(\frac{\delta\, y}{\delta\, t_r}\right)^* = 0 \quad , \tag{8.14}$$

тобто необхідні умови для точки локального мінімуму в загальній задачі математичного програмування (8.6) – (8.8) полягають у невід’ємності тих складових вектора перших умовних похідних $\dfrac{\delta\, y}{\delta \mathbf{t}}$, які відповідають змінним, розташованим на своїх нижніх границях $t_r = t_r^+$, у недодатності тих складових вектори $\dfrac{\delta\, y}{\delta \mathbf{t}}$, які відповідають змінним, розташованим на своїх верхніх границях $t_r = t_r^{++}$, і в рівності нулю тих компонент вектора $\frac{\delta y}{\delta \mathbf{t}}$, які відповідають змінним, розташованим усередині інтервалу $\left[t_i^{+},\, t_i^{++}\right]$.

Якщо задача математичного програмування формулюється у вигляді (8.1), (8.3) із вимогою невід’ємності змінних, то введення додаткових додатних змінних для кожної з нерівностей (8.3) приводить її до часткового випадку задачі (8.6) – (8.8), коли двостороння обмеженість змінних заміняється на односторонню (знизу):

$$y(\mathbf{x}) \to \underset{\mathbf{x}\in\mathbf{\Omega}\subset\mathbf{R}^n}{\mathrm{opt}}, \tag{8.15}$$

$$\mathbf{\Omega}: \quad f_i(\mathbf{x}) = 0, \quad i = \overline{1,m}\ ; \tag{8.16}$$

$$x_j \ge 0, \quad j = \overline{1,n}, \quad n \ge m\ . \tag{8.17}$$

У цьому випадку необхідні умови (8.12) – (8.14) набувають більш простого вигляду:

$$\left(\frac{\delta y}{\partial \mathbf{t}}\right)^{*} \ge 0\,; \tag{8.18}$$

$$\left(\frac{\delta y}{\partial \mathbf{t}}\right)^{*\mathrm{T}} \mathbf{t} \ge 0\ , \tag{8.19}$$

тобто необхідні умови полягають у невід’ємності вектора перших умовних похідних по незалежних змінних $\frac{\delta y}{\partial \mathbf{t}}$ і його ортогональності щодо вектора незалежних змінних $\mathbf{t}$.

8.3. Достатні умови для точки локального оптимуму

У загальній задачі математичного програмування, як і в задачі оптимізації на гіперпаралелепіпеді, мають місце дві групи достатніх умов для точки локального оптимуму. До того ж, як ми побачимо пізніше, вони дуже схожі одна на одну. Відмінність полягає в тому, що в загальній задачі математичного програмування оптимум шукається на гіперпаралелепіпеді (8.10) у просторі незалежних змінних $\mathbf{R}^{\mathrm{p}}$ і частинні похідні поступаються місцем умовним похідним.

Залежно від того, як розташовується область припустимих рішень (гіперпаралелепіпед) щодо умовного екстремуму функції достатні умови формулюються по-різному.

Перша група достатніх умов для точки умовного локального мінімуму за аналогією із задачею при двосторонній обмеженості змінних полягає у виконанні необхідних при відсутності у вектора умовних похідних $\dfrac{\delta y}{\partial \mathbf{t}}$ нульових складових. Іншими словами, якщо умовний мінімум розташовується на границі гіперпаралелепіпеда (8.10) і виконуються необхідні умови:

$$\text{якщо } t_i^* = t_i^+, \quad \text{то } \left(\frac{\delta y}{\delta t_i}\right)^* > 0; \tag{8.20}$$

$$\text{якщо } t_i^* = t_i^{++}, \quad \text{то } \left(\frac{\delta y}{\delta t_i}\right)^* < 0, \tag{8.21}$$

і при цьому

$$\left(\frac{\delta y}{\delta t_i}\right)^* \neq 0, \ i = \overline{1, p}, \tag{8.22}$$

то вони одночасно є і достатніми.

Для доведення першої групи достатніх умов доповнимо ряд розкладання функції $y(\mathbf{t})$ (8.11) членами другого і більш високих порядків у вигляді залишку $R\left(\Delta\mathbf{t}^2, \Delta\mathbf{t}^3, ...\right)$. Тоді розкладання функції $y(\mathbf{t})$ в точці $\mathbf{t}^*$ набуде вигляду

$$\Delta y^* = \left(\frac{\delta y}{\delta \mathbf{t}}\right)^{*\mathrm{T}} \Delta\mathbf{t} + R^*\left(\Delta\mathbf{t}^2, \Delta\mathbf{t}^3, ...\right). \tag{8.23}$$

Тут $R^*\left(\Delta\mathbf{t}^2, \Delta\mathbf{t}^3, ...\right)$ являє собою суму нелінійних членів розкладання, кожний з яких залежить від добутку двох або більше приростів Δt_i .

Якщо необхідні умови виконуються, то лінійний член розкладання $\left(\frac{\delta y}{\delta \mathbf{t}}\right)^{*\mathrm{T}} \Delta\mathbf{t}$ являє собою суму додатних додатків $\left(\frac{\delta y}{\delta t_i}\right)^{*} \Delta t_i$. Це дійсно так, оскільки додатним похідним $\left(\frac{\delta y}{\delta t_i}\right)^{*}$ можуть відповідати тільки додатні прирости Δt_i, а від'ємним – від'ємні. Природно, сума додатних величин не може бути рівної нулю. А оскільки $\left(\frac{\delta y}{\delta \mathbf{t}}\right)^{*} \neq 0$, то $\lim\limits_{\Delta\mathbf{t}\to 0} \frac{R^*\left(\Delta\mathbf{t}^2, \Delta\mathbf{t}^3, ...\right)}{\left(\frac{\delta y}{\delta \mathbf{t}}\right)^{*\mathrm{T}} \Delta\mathbf{t}} = 0$, і залишком у (8.23) можна зневажати. У результаті одержимо розкладання

$$\Delta y^{\bullet} = \left(\frac{\delta y}{\delta \mathbf{t}}\right)^{*\mathrm{T}} \Delta\mathbf{t} , \tag{8.24}$$

з якого випливає, що в будь-якому малому припустимому околі точки $\mathbf{t}^*$ приріст функції $\Delta y^{\bullet}$ при дотриманні умов (8.20) – (8.22) завжди більше нуля, а це значить, що точка $\mathbf{t}^*$ дійсно є точкою умовного мінімуму. Отже, умови (8.20) – (8.22) є як необхідними, так і достатніми.

Друга група достатніх умов має місце, коли в точці шуканого умовного локального мінімуму $\mathbf{t}^*$ частинні умовні похідні цільової функції по деяких (або по всіх) незалежних змінних дорівнюють нулю. У цьому випадку другим членом розкладання (8.23) при виведенні достатніх умов зневажати не можна.

Для доведення другої групи достатніх умов розіб'ємо вектор $\mathbf{t}^*$ на два складових підвектора $\mathbf{t}_1^*$ і $\mathbf{t}_2^*$ так, щоб складовим t_{1i}^* підвектора $\mathbf{t}_1^*$ відповідали похідні $\left(\dfrac{\delta y}{\delta t_{1i}}\right)^* \neq 0$, а складовим t_{2i}^* підвектора $\mathbf{t}_2^*$ – похідні $\left(\dfrac{\delta y}{\delta t_{2i}}\right)^* = 0$. Вважатимемо, що підвектор $\mathbf{t}_1^*$ містить перші p_1 компонент вектора $\mathbf{t}^*$, а підвектор $\mathbf{t}_2^*$ – інші $p_2 = p - p_1$. У протилежному разі відповідна перенумерація змінних завжди дозволить домогтися зазначеної розбивки.

Розкладемо функцію $y(\mathbf{t})$ в околі точки $\mathbf{t}^*$ у ряд Тейлора (6.58) з урахуванням розбивки вектора незалежних змінних:

$$\Delta y^* = \left(\frac{\delta y}{\delta \mathbf{t}_1}\right)^{*\mathrm{T}} \Delta\mathbf{t}_1 + \left(\frac{\delta y}{\delta \mathbf{t}_2}\right)^{*\mathrm{T}} \Delta\mathbf{t}_2 + \frac{1}{2}\Delta\mathbf{t}_1^{\mathrm{T}}\mathbf{S}_{11}^*\Delta\mathbf{t}_1 + \\ + \frac{1}{2}\Delta\mathbf{t}_1^{\mathrm{T}}\mathbf{S}_{12}^*\Delta\mathbf{t}_2 + \frac{1}{2}\Delta\mathbf{t}_2^{\mathrm{T}}\mathbf{S}_{21}^*\Delta\mathbf{t}_1 + \frac{1}{2}\Delta\mathbf{t}_2^{\mathrm{T}}\mathbf{S}_{22}^*\Delta\mathbf{t}_2 , \qquad (8.25)$$

де $\mathbf{S}_{11}, \mathbf{S}_{12}, \mathbf{S}_{21}, \mathbf{S}_{22}$ – підматриці матриці других умовних похідних по незалежним змінним $\mathbf{S}$, обчисленої за формулою (6.57).

Нехай в околі $\mathbf{t}^*$ змінюються тільки складові підвектора, $\mathbf{t}_1^*$, тобто

$\Delta\mathbf{t}_2 = 0$. Тоді

$$\Delta y^* = \left(\frac{\delta y}{\partial \mathbf{t}_1}\right)^{*\mathrm{T}} \Delta\mathbf{t}_1 + \frac{1}{2}\Delta\mathbf{t}_1^{\mathrm{T}}\mathbf{S}_{11}^*\mathbf{t}_1 \ . \qquad (8.26)$$

Оскільки $\left(\frac{\delta y}{\partial \mathbf{t}_1}\right)^* \neq 0$, то другим членом розкладання (8.26) можна зневажати. Якщо тепер міркувати так, як і при виведенні першої групи достатніх умов, то приходимо до аналогічного висновку: *необхідні умови щодо складових вектора* $\mathbf{t}_1^*$ *одночасно є і достатніми.*

Нехай тепер в околиці $\mathbf{t}^*$ змінюються тільки складові вектора $\mathbf{t}_2^*$, тобто $\Delta\mathbf{t}_1 = 0$. Тоді

$$\Delta y^* = \frac{1}{2}\Delta\mathbf{t}_2^{\mathrm{T}}\mathbf{S}_{22}^*\Delta\mathbf{t}_2 \ . \qquad (8.27)$$

Тут лінійний член розкладання відсутній, оскільки за умовою розбивки всі складові вектори $\left(\frac{\delta y}{\partial \mathbf{t}_2}\right)^*$ дорівнюють нулю.

Оскільки розкладання робилося в точці мінімуму, останній вираз повинний бути строго більше нуля. Коли допустити, що він дорівнює нулю, то в цьому випадку не можна зневажати членами розкладання більш високого порядку. Вираз (8.27) буде більше нуля тільки у випадку, якщо підматриця $\mathbf{S}_{22}^*$ додатно визначена.

Таким чином, *друга група достатніх умов для точки умовного локального мінімуму в загальній задачі математичного програмування (8.6) – (8.8) полягає у виконанні необхідних і додатній визначеності підматриці* $\mathbf{S}_{22}^*$*, складеної для тих незалежних змінних* t_i^**, відповідні похідні яких* $\left(\frac{\partial y}{\partial t_i}\right)^*$ *дорівнюють нулю.*

Доведені достатні умови другої групи орієнтовані на загальний випадок задачі математичного програмування, коли функція цілі може мати декілька умовних екстремумів. Коли відомо, що функція (8.1) строго опукла (увігнута) на припустимій множині рішень Ω, то необхідні умови для умовного екстремуму автоматично стають і достатніми, незалежно від того, що деякі (або всі) складові вектори умовних похідних можуть стати рівними нулю. Пояснити це можна тим, що така функція має тільки одну умовну стаціонарну точку, яка у випадку опуклості функції цілі зобов'язана бути точкою умовного мінімуму, а у випадку увігнутості – максимуму. Тому, якщо знайдена умовна стаціонарна точка опуклої або увігнутої функції, то немає необхідності перевіряти її на екстремальність.

8.4. Диференціальний алгоритм вирішення загальної задачі математичного програмування

Диференціальний алгоритм вирішення загальної задачі математичного програмування багато в чому нагадує раніше розглянутий алгоритм для вирішення задачі оптимізації на гіперпаралелепіпеді. Останній був заснований на методі покоординатного спуску по гіперповерхні цільової функції в просторі змінних $\mathbf{R}^n$. На кожному кроці спуску робилося точне обчислення приросту варійованої змінної до явного мінімуму і до границі інтервалу її визначення. Вибір мінімального за абсолютним значенням приросту не допускав виходу процесу оптимізації з області припустимих рішень **Ш**. Остання являла собою гіперпаралелепіпед (7.2).

Розв'язання загальної задачі математичного програмування за диференціальним алгоритмом виконується за тією схемою, що і вирішення задачі оптимізації на гіперпаралелепіпеді. Перед початком пошуку оптимального рішення задаються точністю обчислення ε, вибирають початкову точку наближення $\mathbf{x}^{(0)} \in \Omega$, але, крім того, розбивають вектор змінних $\mathbf{x}$ на підвектор залежних $\mathbf{s}$ і підвектор незалежних змінних $\mathbf{t}$. Причому, розбивку роблять таким чином, щоб точка початкового наближення $\mathbf{x}^{(0)}$ була не тільки неособливою, але і невиродженою. Це дозволяє задачу оптимізації (8.6) – (8.8) звести до задачі оптимізації функції (8.9) у просторі незалежних змінних при двосторонній обмеженості останніх (8.10).

При пошуку умовного мінімуму за диференціальним алгоритмом безпосередня процедура оптимізації полягає в наступному. На кожному кроці оптимізації змінюється одна незалежна змінна t_r, а інші залишаються постійними, тобто функція одержує приріст

$$\Delta y^{(k)} = \left(\frac{\delta y}{\delta t_r}\right)^{(k)} \Delta t_r^{(k)} \,, \tag{8.28}$$

де k – індекс, що позначає крок оптимізації, $k = 0,1,\ldots$ При цьому змінна змінюється тільки в тому випадку, якщо по ній порушені необхідні умови (8.12) – (8.14). Для перевірки необхідних умов обчислюють умовну похідну $\left(\frac{\delta y}{\delta t_r}\right)^{(k)}$. Якщо обчислена похідна за абсолютним значенням не перевершує задану точність ε, то необхідні умови вважаються непорушеними.

Якщо в поточній точці наближення $\mathbf{t}^{(k)}$ виконуються умови (8.20) – (8.22), то вона відповідає точці умовного локального мінімуму, оскільки ці умови такою мірою достатні, як і необхідні.

Якщо умова (8.22) порушена, а функція y або область завдання $\boldsymbol{\Omega}$ не є опуклими, але при цьому виконуються необхідні умови (8.12) – (8.14) , то аналізують матрицю других умовних похідних $\mathbf{S}_{22}^{(k)}$, яка визначає зміну функції (8.27) у точці $\mathbf{t}^{(k)}$. Позитивна визначеність матриці $\mathbf{S}_{22}^{(k)}$ вказує на наявність умовного мінімуму в точці $\mathbf{t}^{(k)}$.

Якщо похідна за абсолютним значенням перевершує задану точність ε, має додатний знак, а відповідна їй змінна t_r не дорівнює значенню нижній границі t_r^{+} (перший тип порушень), то значення функції y можна зменшити за рахунок зменшення змінної t_r. Зменшувати змінну t_r можна доти, поки вона або не досягне нижньої границі t_r^+, або похідна $\dfrac{\delta y}{\delta t_r}$ не обернеться на нуль, або будь-яка залежна змінна не стане рівної своєму граничному значенню (нижньому або верхньому).

Якщо похідна за абсолютним значенням перевершує задану точність ε, має від'ємний знак, а відповідна їй змінна t_r не дорівнює своєму верхньому значенню t_r^{++} (другий тип порушень), то значення функції y також можна зменшити за рахунок збільшення змінної t_r. Збільшувати змінну t_r можна доти, поки вона або не досягне верхньої границі t_r^+, або похідна $\dfrac{\delta y}{\delta t_r}$ не обернеться на нуль, або будь-яка залежна змінна не стане рівною одному із своїх граничних значень.

Як бачимо, рух по змінній t_r здійснюється в двох випадках: коли маємо порушення необхідних умов першого або другого типу. При цьому чітко визначається напрямок руху: при першому типі порушення $\Delta t_r > 0$, при другому $\Delta t_r < 0$.

Набагато складніша справа з вибором довжини кроку в обраному напрямку. При зміні змінної t_r інші незалежні змінні залишаються постійними, але всі залежні змінні також одержують прирости

$$\Delta s_i = \left(\frac{\delta s_i}{\delta t_r}\right)\Delta t_r \quad i = \overline{1,m} \ , \tag{8.29}$$

які при досить великому Δt_r виводять залежні змінні з області припустимих рішень Ω. Щоб уникнути виходу залежних змінних з області Ω, необхідно стежити за їхньою зміною при зміні змінної t_r. У випадку, якщо будь-яка залежна змінна вийде на свою границю раніше, ніж змінна t_r досягне явного мінімуму або своєї границі, то таку залежну змінну треба зробити незалежною замість змінної t_r. Для організації контролю за зміною залежних змінних необхідно на кожному кроці аналізувати вектор умовних похідних $\dfrac{\delta \mathbf{s}}{\delta t_r}$, складові якого дозволяють визначити, при якому прирості змінної t_r та або інша залежна змінна вийде на свою границю:

$$\Delta t_r = \frac{\Delta s_i}{\left(\dfrac{\delta s_i}{\delta t_r}\right)}, \quad i = \overline{1,m} \tag{8.30}$$

де Δs_i – відстань від поточного значення залежної змінної s_i до її границі (верхньої або нижньої в залежності від знака $\dfrac{\delta s_i}{\delta t_r}$).

Аналітично вибір напрямку і приріст варійованої змінної на k-му кроці оптимізації виражається за допомогою наступних критеріїв.

Вибір напрямку:

$$\text{якщо} \quad \left(\frac{\delta y}{\delta t_r}\right)^{(k)} > 0 \quad \text{и} \quad t_r^{(k)} \neq t_r^{+}, \quad \text{то} \quad \Delta t_r^{(k)} < 0; \tag{8.31}$$

$$\text{якщо} \quad \left(\frac{\delta y}{\delta t_r}\right)^{(k)} < 0 \quad \text{и} \quad t_r^{(k)} \neq t_r^{++}, \quad \text{то} \quad \Delta t_r^{(k)} > 0. \tag{8.32}$$

Вибір довжини кроку:

якщо $\Delta t_r^{(k)} < 0$, то

$$\Delta t_r^{(k)} = \max_{\Delta t_r < 0} \left[t_r^{+} - t_r^{(k)};\ \Delta t_r^{(k)} \Big|_{\frac{\delta y}{\delta t_r} = 0};\ \max_{\frac{\delta s_i}{\delta t_r} < 0} \frac{s_i^{++} - s_i^{(k)}}{\left(\frac{\delta s_i}{\delta t_r}\right)^{(k)}};\ \max_{\frac{\delta s_i}{\delta t_r} > 0} \frac{s_i^{+} - s_i^{(k)}}{\left(\frac{\delta s_i}{\delta t_r}\right)^{(k)}} \right]; \tag{8.33}$$

якщо $\Delta t_r^{(k)} > 0,$ то

$$\Delta t_r^{(k)} = \min_{\Delta t_r > 0}\left[t_r^{++} - t_r^{(k)};\ \Delta t_r^{(k)}\Big|_{\frac{\delta y}{\delta t_r}=0};\ \min_{\frac{\delta s_i}{\delta t_r}<0}\frac{s_i^{+} - s_i^{(k)}}{\left(\frac{\delta s_i}{\delta t_r}\right)^{(k)}};\ \min_{\frac{\delta s_i}{\delta t_r}>0}\frac{s_i^{++} - s_i^{(k)}}{\left(\frac{\delta s_i}{\delta t_r}\right)^{(k)}}\right].(8.34)$$

Якщо в результаті кроку варійована змінна t_r приймає граничне значення або похідна $\left(\frac{\delta y}{\delta t_r}\right)^{(k)}$ обертається на нуль, то поточна система залежних змінних залишається колишньою, а нові значення залежних змінних визначаються вирішенням системи рівнянь (8.7) у припущенні, що $t_r^{(k)}$ змінилося на величину Д$t_r^{(k)}$ відповідно до критеріїв (8.33) або (8.34), а інші незалежні змінні залишилися постійними.

Якщо в результаті зробленого кроку будь-яка залежна змінна $s_k^{(k)}$ приймає граничне значення, то цю змінну переводять у систему незалежних змінних замість незалежної змінної t_r , а нові поточні значення залежних змінних визначаються вирішенням системи рівнянь (8.7) за умовами, що нова незалежна змінна дорівнює відповідному граничному значенню, а інші незалежні змінні залишилися постійними.

Через нелінійність рівнянь (8.7) при досить великій довжині кроку (8.33) або (8.34) може мати місце помилкове визначення залежної змінної, яка перша прийме граничне значення. Тому після вибору такої змінної і вирішення системи рівнянь (8.7) щодо залежних змінних останні повинні перевірятися на приналежність області припустимих рішень. Якщо деякі з них виходять із заданого інтервалу, обчислення приросту за формулою (8.33) або (8.34) робиться наново, але тільки для тих змінних, які опинилися за межами області Ω .

Найбільш трудомісткою процедурою при виборі довжини кроку за критеріями (8.33) і (8.34) є обчислення приросту $\Delta t_r^{(k)}\Big|_{\frac{\delta y}{\delta t_r}=0}$, тобто зміна t_r , яка пов'язана з обертанням на нуль умовною похідною $\dfrac{\delta y}{\delta t_r}$. Найбільш простий спосіб визначення наближеного значення $\Delta t_r^{(k)}\Big|_{\frac{\delta y}{\delta t_r}=0}$ полягає в його обчисленні по вже добре відомій формулі:

$$Дt_r^{(k)}\Bigg|_{\frac{\delta y}{\delta t_r}=0} = -\frac{\left(\dfrac{\delta y}{\delta t_r}\right)^{(k)}}{\left|\left(\dfrac{\delta^2 y}{\delta t_r^2}\right)^{(k)}\right|}. \qquad (8.35)$$

Перша умовна похідна в (8.35) може бути обчислена за допомогою формули (6.34), а друга – за допомогою (6.57).

Проміжні результати вирішення задачі (8.6) – (8.8) за диференціальним алгоритмом зручно надавати у вигляді спеціальної таблиці k –го кроку алгоритму. Така таблиця (див. табл.8.1) складається із двох частин. Одна з них містить значення, які обчислюються на кожному кроці, друга містить позначення або постійні значення. Остання на табл.8.1 виділена сірим фоном.

Перший рядок таблиці містить позначення t_j, $j=\overline{1,p}$, що закріплюють відповідні стовпці за цими змінними. Аналогічно перший стовпець таблиці містить позначення s_i, $i=\overline{1,m}$, що закріплюють відповідні рядки за цими змінними.

Перший рядок частини таблиці, яка обчислюється, містить поточні значення умовних похідних по незалежним змінним $\left(\dfrac{\delta y}{\delta t_j}\right)^{(k)}$ і функції цілі

$y^{(k)}$. Інші рядки цієї частини, крім останньої, містять поточні значення умовних $\left(\frac{\delta s_i}{\delta t_j}\right)^{(k)}$ похідних і залежної перемінної $s_i^{(k)}$. Останній рядок частини, яка обчислюється , призначена для поточних значень незалежних змінних $t_j^{(k)}$.

Останні два стовпці таблиці містять значення верхніх і нижніх границь залежних змінних s_i^{+} і s_i^{++}, а два останні рядки – незалежних t_j^{+} і t_j^{++} .

Таблиця 8.1 – Загальний вид таблиці k–го кроку диференціального алгоритму для загальної задачі математичного програмування

	t_1	…	t_j	…	t_p			
	$\left(\frac{\delta y}{\delta t_1}\right)^{(k)}$	…	$\left(\frac{\delta y}{\delta t_j}\right)^{(k)}$	…	$\left(\frac{\delta y}{\delta t_p}\right)^{(k)}$	$y^{(k)}$		
s_1	$\left(\frac{\delta s_1}{\delta t_1}\right)^{(k)}$	…	$\left(\frac{\delta s_1}{\delta t_j}\right)^{(k)}$	…	$\left(\frac{\delta s_1}{\delta t_p}\right)^{(k)}$	$s_1^{(k)}$	s_1^{+}	s_1^{++}
…	…	…	…	…	…	…	…	…
s_i	$\left(\frac{\delta s_i}{\delta t_1}\right)^{(k)}$	…	$\left(\frac{\delta s_i}{\delta t_j}\right)^{(k)}$	…	$\left(\frac{\delta s_i}{\delta t_p}\right)^{(k)}$	$s_i^{(k)}$	s_i^{+}	s_i^{++}
…	…	…	…	…	…	…	…	…
s_m	$\left(\frac{\delta s_m}{\delta t_1}\right)^{(k)}$	…	$\left(\frac{\delta s_m}{\delta t_j}\right)^{(k)}$	…	$\left(\frac{\delta s_m}{\delta t_h}\right)^{(k)}$	$s_m^{(k)}$	s_m^{+}	s_m^{++}
	$t_1^{(k)}$	…	$t_j^{(k)}$	…	$t_p^{(k)}$			
	t_1^{+}	…	t_j^{+}	…	t_p^{+}			
	t_1^{++}	…	t_j^{++}	…	t_p^{++}			

8.5. Диференціальний алгоритм при лінійних обмеженнях

Окремим випадком загальної задачі математичного програмування є задача оптимізації при лінійних обмеженнях у вигляді рівностей і двосторонньої обмеженості змінних:

$$y(\mathbf{x}) \to \min_{\mathbf{x}\in\mathbf{\Omega}\subset\mathbf{R}^n} \ , \qquad (8.36)$$

$$\mathbf{\Omega}: \quad \mathbf{Ax}+\mathbf{b}=0 \ ; \qquad (8.37)$$

$$\mathbf{x}^{+} < \mathbf{x} < \mathbf{x}^{++} \ . \qquad (8.38)$$

Тут $y(\mathbf{x})$ – нелінійна функція n змінних; (8.37) – система лінійних рівнянь у матрично-векторній формі запису; $\mathbf{A}$ – матриця коефіцієнтів розміру $m\times n, \quad m<n$; $\mathbf{b}$ – вектор-стовпець вільних членів.

При розбивці змінних задачі на залежні й незалежні $\mathbf{x}^{\mathrm{T}} = \begin{bmatrix}\mathbf{s}^{\mathrm{T}} & \mathbf{t}^{\mathrm{T}}\end{bmatrix}$ матриця $\mathbf{A}$ розпадеться на дві підматриці $\begin{bmatrix}\mathbf{W} & \mathbf{C}\end{bmatrix}$, де $\mathbf{W}$ – матриця Якобі; $\mathbf{C}$ – матриця управління. На відміну від аналогічних матриць загальної задачі математичного програмування, ці матриці не залежать від змінних задачі через лінійність системи обмежень (8.37).

З урахуванням розбивки змінних система рівнянь (8.37) набуває вигляду:

$$\begin{bmatrix}\mathbf{W} & \mathbf{C}\end{bmatrix}\begin{bmatrix}\mathbf{s}\\ \mathbf{t}\end{bmatrix}+\mathbf{b}=0 \ ,$$

або

$$\mathbf{Ws}+\mathbf{Ct}+\mathbf{b}=0 \ . \qquad (8.39)$$

Якщо $\mathbf{W}$ є неособливою матрицею ($\det\mathbf{W}\neq 0$), із (8.39) випливає

$$\mathbf{s}=-\mathbf{W}^{-1}\mathbf{Ct}-\mathbf{W}^{-1}\mathbf{b} \ . \qquad (8.40)$$

Вводячи позначення

$$\mathbf{B}=-\mathbf{W}^{-1}\mathbf{C} \ , \qquad (8.41)$$

$$\mathbf{в} = -\mathbf{W}^{-1}\mathbf{b} \tag{8.42}$$

і підставляючи їх у (8.40), одержимо

$$\mathbf{s} = \mathbf{Bt} + \mathbf{в} \ . \tag{8.43}$$

Тут $\mathbf{s}$ – n-вимірний вектор залежних змінних; $\mathbf{t}$ – p-вимірний вектор незалежних змінних; $\mathbf{B}$ – матриця умовних похідних, $b_{ij} = \dfrac{\delta s_i}{\delta t_j}$, $i = \overline{1, m}$, $j = \overline{1, p}$; $\mathbf{в}$ – p-вимірний вектор вільних членів. Вираз (8.43) завжди можна одержати з (8.37) за допомогою звичайних жорданових виключень.

Оскільки всі величини в правій частині (8.41) не залежать від змінних задачі, то і ліва частині також не залежить від них. Елементи b_{ij} матриці $\mathbf{B}$ є константами. Якщо в (8.43) змінюється одна незалежна змінна t_r, то $\Delta s_i = b_{ir}\Delta t_r$, або

$$\Delta t_r = \frac{\Delta s_i}{b_{ir}} \ . \tag{8.44}$$

Лінійна залежність між змінними t_r і s_i дозволяє точно визначити, при якому прирості змінної t_r залежна змінна s_i зі свого поточного положення $s_i^{(k)}$ переміститься на одну із своїх границь, а саме:

на нижню

$$\Delta t_r^{(k)} = \frac{s_i^{+} - s_i^{(k)}}{b_{ir}^{(k)}} \ ; \tag{8.45}$$

на верхню

$$\Delta t_r^{(k)} = \frac{s_i^{++} - s_i^{(k)}}{b_{ir}^{(k)}} \ . \tag{8.46}$$

Таким чином, вибір на кожному кроці диференціального алгоритму приросту змінної відповідно до критерію

$$\Delta t_r^{(k)} = \max_{\Delta t_r < 0} \left[t_r^{+} - t_r^{(k)}; \Delta t_r^{(k)} \Big|_{\frac{\delta y}{\delta t_r} = 0}; \max_{b_{ir} < 0} \frac{s_i^{++} - s_i^{(k)}}{b_{ir}^{(k)}}; \max_{b_{ir} > 0} \frac{s_i^{+} - s_i^{(k)}}{b_{ir}^{(k)}} \right]; \tag{8.47}$$

або

$$\Delta t_r^{(k)} = \min_{\Delta t_r > 0} \left[t_r^{++} - t_r^{(k)}; \Delta t_r^{(k)} \Big|_{\frac{\delta y}{\delta t_r} = 0}; \min_{b_{ir} < 0} \frac{s_i^{+} - s_i^{(k)}}{b_{ir}^{(k)}}; \min_{b_{ir} > 0} \frac{s_i^{++} - s_i^{(k)}}{b_{ir}^{(k)}} \right]. \tag{8.48}$$

не приведе до виходу з області припустимих рішень Ω при як завгодно великому кроці.

Якщо на черговому кроці приріст змінної вибирається з умови виходу якоїсь залежної змінної s_i на границю, то вирішення системи рівнянь (8.37) з метою визначення поточних значень залежних змінних $s_i^{(k)}$ також зводиться до кроку жорданових виключень щодо системи (8.43).

Умовні похідні, необхідні для обчислення $\Delta t_r^{(k)} \Big|_{\frac{\delta y}{\delta t_r} = 0}$, доцільно розраховувати за формулами

$$\frac{\delta y}{\delta t_r} = \frac{\partial y}{\partial t_r} + \sum_{i}^{m} b_{ir} \frac{\partial y}{\partial s_i} \ ; \tag{8.49}$$

$$\frac{\delta^2 y}{\delta t_r^2} = \frac{\partial^2 y}{\partial t_r^2} + 2\sum_{i}^{m} b_{ir} \frac{\partial^2 y}{\partial s_i \partial t_r} + \sum_{i}^{m} \sum_{j}^{m} b_{ir} b_{jr} \frac{\partial^2 y}{\partial s_i \partial s_j} \ . \tag{8.50}$$

При перевірці достатніх умов наявності умовного екстремуму другі частинні й змішані умовні похідні для формування матриці $\mathbf{S}_{22}$ можна обчислювати за формулою

$$\frac{\delta^2 y}{\delta t_k \delta t_l} = \frac{\partial^2 y}{\partial t_k \partial t_l} + \sum_i^m \left(b_{ik} \frac{\partial^2 y}{\partial s_i \partial t_l} + b_{il} \frac{\partial^2 y}{\partial s_i \partial t_k} \right) +$$

$$+ \sum_i^m \sum_j^m b_{ik} b_{jl} \frac{\partial^2 y}{\partial s_i \partial s_j}, \quad k,l \in \{1,2,..., p\} \ . \quad (8.51)$$

Таблиця 8.2 – Загальний вид таблиці k–го кроку диференціального алгоритму для загальній задачі математичного програмування при лінійних обмеженнях

	t_1	…	t_j	…	t_p				
	$\left(\frac{\delta y}{\delta t_1}\right)^{(k)}$	…	$\left(\frac{\delta y}{\delta t_j}\right)^{(k)}$	…	$\left(\frac{\delta y}{\delta t_p}\right)^{(k)}$	$y^{(k)}$			
s_1	$b_{11}^{(k)}$	…	$b_{1j}^{(k)}$	…	$b_{1p}^{(k)}$	$\beta_1^{(k)}$	$s_1^{(k)}$	s_1^{+}	s_1^{++}
…	…	…	…	…	…	…	…	…	…
s_i	$b_{i1}^{(k)}$	…	$b_{ij}^{(k)}$	…	$b_{ip}^{(k)}$	$\beta_i^{(k)}$	$s_i^{(k)}$	s_i^{+}	s_i^{++}
…	…	…	…	…	…	…	…	…	…
s_m	$b_{m1}^{(k)}$	…	$b_{mj}^{(k)}$	…	$b_{mp}^{(k)}$	$\beta_m^{(k)}$	$s_m^{(k)}$	s_m^{+}	s_m^{++}
	$t_1^{(k)}$	…	$t_j^{(k)}$	…	$t_p^{(k)}$				
	t_1^{+}	…	t_j^{+}	…	t_p^{+}				
	t_1^{++}	…	t_j^{++}	…	t_p^{++}				

Таблиця k -го кроку диференціального алгоритму задачі (8.36) – (8.38) відрізняється від аналогічної таблиці загальної задачі математичного програмування (8.6) – (8.8). У ній (див. табл.8.2) з'являється додатковий стовпець для поточних значень вільних членів $\beta_i^{(k)}$ із системи (8.43). Введення додаткового стовпця дозволяє робити транспозицію залежної і незалежної змінних (якщо в ній виникає необхідність) за допомогою жорданових виключень. Останні значно полегшують обчислювальні процедури, необхідні для заповнення чергової таблиці диференціального алгоритму, а саме для обчислення нових значень залежних змінних і системи умовних похідних залежних змінних за незалежними.

8.6. Контрольні запитання і вправи

1. Які особливості має загальна задача математичного програмування? Які її принципові відмінності в постановці від задачі безумовної оптимізації? Задачі оптимізації при обмеженнях у вигляді рівностей? Задачі оптимізації при двосторонній обмеженості змінних?

2. Привести до вигляду (8.6) – (8.8) задачу пошуку екстремуму

$$y(\mathbf{x}) = \frac{x_1}{x_2} - \frac{x_3 x_4}{x_5} \to \min_{\mathbf{x}\in\mathbf{\Omega}} ,$$

$$\mathbf{\Omega}: \quad x_1 + x_2 + x_3 + x_4 + x_5 - 100 \le 0 \;;$$

$$x_1^2 + x_2^2 + x_3^2 + x_4^2 + x_5^2 - 100 \ge 0 \;;$$

$$x_1 x_2 x_3 x_4 x_5 = 100 \;;$$

$$|x_i| \le 10, \quad i = \overline{1,3} \;; \quad x_5 \ge 0 \,.$$

3. Знайти рішення в задачі пошуку умовного мінімуму:

$$y(\mathbf{x}) = x_1^2 + x_2^2 \to \min_{\mathbf{x}\in\mathbf{\Omega}} \;;$$

$$\mathbf{\Omega}: \quad x_1 x_2 - 1 = 0 \;;$$

$$\begin{bmatrix} 2 \\ 2 \end{bmatrix} \le \mathbf{x} \le \begin{bmatrix} 3 \\ 3 \end{bmatrix}.$$

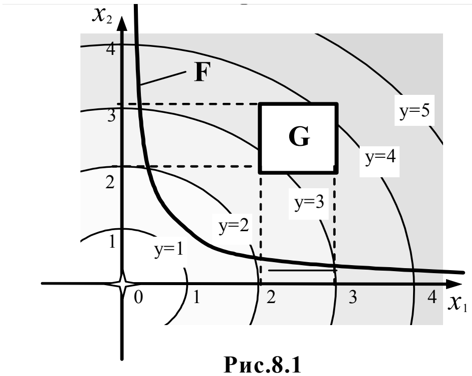


Рис.8.1

Розв'язання. Множина припустимих рішень $\mathbf{\Omega}$ являє собою переріз двох множин $\mathbf{F}: x_1x_2 - 1 = 0$ і $\mathbf{G}: \begin{bmatrix} 2 \\ 2 \end{bmatrix} \le \mathbf{x} \le \begin{bmatrix} 3 \\ 3 \end{bmatrix}$. Оскільки $\mathbf{\Omega} = \mathbf{F} \cup \mathbf{G} = \varnothing$ (рис.8.1), то задача не має рішення.

4. Знайти оптимальне рішення в задачі:

$$y(\mathbf{x}) = 2x_1x_2 + x_3^2 + x_4 \to \min_{\mathbf{x}\in\mathbf{\Omega}\subset\mathbf{R}^4} \; ;$$

$$\mathbf{\Omega}: \quad x_1x_3 + x_4 - 2 = 0 \; ;$$

$$x_1^2 - 2x_3 + x_2 - 4 = 0 \, ;$$

$$0{,}5 \le x_1 \le 1 \; ; \qquad 1 \le x_3 \le 2 \, .$$

Розв'язання. Задача вирішується шляхом приведення її до задачі мінімізації при двосторонній обмеженості змінних. Дійсно, система обмежень, що формує область припустимих рішень $\mathbf{\Omega}$, така, що дозволяє виразити необмежені змінні x_2 і x_4 через змінні x_1 і x_3: $x_2 = 4 - x_1^2 + 2x_3$; $x_4 = 2 - x_1x_3$. Підставляючи отримані вирази у початкову функцію $y(\mathbf{x})$, одержимо задачу мінімізації при двосторонній обмеженості змінних:

$$y(\mathbf{x}) = -2x_1^3 + x_3^2 + 3x_1x_3 + 8x_1 + 2 \to \min_{\mathbf{x} \in \mathbf{H} \subset \mathbf{R}^2} \; ;$$

$$\mathbf{H}: \quad 0{,}5 \le x_1 \le 1 \; ;$$
$$1 \le x_3 \le 2 \; .$$

Розв’язання задач такого класу було докладно розглянуто в попередньому розділі.

5. Сформулювати необхідні умови для точки умовного локального мінімуму в загальному завданні математичного програмування.

6. Сформулювати необхідні умови для точки умовного локального максимуму в загальній задачі математичного програмування.

7. Сформулювати першу групу достатніх умов для точки умовного локального мінімуму (максимуму) в загальній задачі математичного програмування.

8. Сформулювати другу групу достатніх умов для точки умовного локального мінімуму в загальній задачі математичного програмування.

9. Сформулювати другу групу достатніх умов для точки умовного локального максимуму в загальній задачі математичного програмування.

10. Які порушення необхідних умов для точки умовного локального мінімуму (максимуму) можуть мати місце в довільній точці, що належить області припустимих рішень загальної задачі математичного програмування?

11. Як вибирається приріст варійованої змінної t_r в загальній задачі математичного програмування при її розв’язуванні за диференціальним алгоритмом, коли вона (змінна) знаходиться строго всередині області припустимих рішень, а частинна умовна похідна функції цілі цей змінній має додатний знак? Від’ємний знак? Нульове значення?

12. Як вибирається приріст варійованої змінної t_r в загальній задачі математичного програмування при її розв’язуванні за диференціальним алгоритмом, якщо вона (змінна) знаходиться на своїй нижній границі, а частинна умовна похідна функції цілі по цій змінній має додатний знак? Від’ємний знак? Нульове значення?

13. . Як вибирається приріст варійованої змінної t_r в загальній задачі математичного програмування при її розв’язуванні за диференціальним алгоритмом, якщо вона (змінна) знаходиться на своїй верхній границі, а частинна умовна похідна функції цілі по цей змінній має додатний знак? Від'ємний знак? Нульове значення?

14. Який вигляд матиме таблиця k –го кроку диференціального алгоритму вирішення загальної задачі математичного програмування, якщо всі змінні не обмежені зверху?

15. Вирішити завдання

$$y(\mathbf{x}) = x_1 x_2 + 2x_3^2 \to \min_{\mathbf{x}\in\mathbf{\Omega}} \; ; \qquad (8.52)$$

$$\mathbf{\Omega}: \quad x_2 x_4 + x_3 - 2 = 0 \; ; \qquad (8.53)$$

$$x_1 + x_2^2 + 2x_4 - 4 = 0 \, ; \qquad (9{,}54)$$

$$0 \le x_1 \le 1{,}2 \; ; \qquad 0 \le x_2 \le 1{,}5 \, ; \qquad (8.55)$$

$$0 \le x_3 \le 2 \; ; \qquad 0 \le x_4 \le 3 \, . \qquad (8.56)$$

за точністю обчислення $\varepsilon = 0{,}1$ і початковою точкою наближення $\mathbf{x}^{\mathrm{T}} = \begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix}$.

Розв'язання. За постановкою маємо загальну задачу математичного програмування. Здійснимо її вирішення за допомогою диференціального алгоритму.

Сформуємо початкову таблицю диференціального алгоритму, для чого попередньо розіб'ємо змінні задачі на залежні й незалежні і визначимо значення всіх її елементів. Нехай x_1 і x_2 будуть залежними змінними, а x_3 і x_4 – незалежними. Обчислимо в початковій точці наближення $\mathbf{x}^{\mathrm{T}} = \begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix}$ значення умовних похідних залежних змінних за незалежними. Згідно з (6.31)

$$\left(\frac{\delta \mathbf{s}}{\delta \mathbf{t}}\right)^{(0)} = \left[\left(\frac{\delta s_i}{\delta t_j}\right)^{(0)}\right] = \begin{bmatrix} \left(\frac{\delta x_1}{\delta x_3}\right)^{(0)} & \left(\frac{\delta x_1}{\delta x_4}\right)^{(0)} \\ \left(\frac{\delta x_2}{\delta x_3}\right)^{(0)} & \left(\frac{\delta x_2}{\delta x_4}\right)^{(0)} \end{bmatrix} = -\left(\mathbf{W}^{-1}\right)^{(0)} \mathbf{C}^{(0)} =$$

$$= -\begin{bmatrix} -2 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ -1 & -1 \end{bmatrix} .$$

Іншим способом обчислення матриці $\left(\dfrac{\delta \mathbf{s}}{\delta \mathbf{t}}\right)^{(0)}$ є приведення системи обмежень (8.53) – (8.54) , якщо тільки вона дозволяє, до вигляду

$$\mathbf{s} = \varphi(\mathbf{t}) \tag{8.57}$$

з подальшим звичайним диференціюванням за незалежними змінними. Перетворимо (8.53) – (8.54) до вигляду (8.57):

$$\mathrm{x}_1 = 4 - 2\mathrm{x}_4 - \left(\frac{2 - \mathrm{x}_3}{\mathrm{x}_4}\right); \qquad \mathrm{x}_2 = \frac{2 - \mathrm{x}_3}{\mathrm{x}_4} \ .$$

Тоді

$$\frac{\delta \mathrm{x}_1}{\delta \mathrm{x}_3} = -\frac{2(2 - \mathrm{x}_3)}{\mathrm{x}_4} \cdot \frac{(-1)}{\mathrm{x}_4}, \qquad \left(\frac{\delta \mathrm{x}_1}{\delta \mathrm{x}_3}\right)^{(0)} = 2;$$

$$\frac{\delta \mathrm{x}_1}{\delta \mathrm{x}_4} = -2 + \frac{2(2 - \mathrm{x}_3)}{\mathrm{x}_4} \cdot \frac{(2 - \mathrm{x}_3) \cdot 1}{\mathrm{x}_4}, \quad \left(\frac{\delta \mathrm{x}_1}{\delta \mathrm{x}_4}\right)^{(0)} = 0;$$

$$\frac{\delta \mathrm{x}_2}{\delta \mathrm{x}_3} = \frac{-1}{\mathrm{x}_4}, \qquad \left(\frac{\delta \mathrm{x}_2}{\delta \mathrm{x}_3}\right)^{(0)} = -1;$$

$$\frac{\delta \mathrm{x}_2}{\delta \mathrm{x}_4} = -\frac{2(2 - \mathrm{x}_3)}{\mathrm{x}_4^2}, \qquad \left(\frac{\delta \mathrm{x}_2}{\delta \mathrm{x}_4}\right)^{(0)} = -1 \ .$$

Недоліком першого способу обчислення матриці $\dfrac{\delta \mathbf{s}}{\delta \mathbf{t}}$ є потреба в трудомісткій процедурі обертання якобіана.

Елементи таблиці, що відповідають умовним похідним $\dfrac{\delta y}{\delta t_j}$, можуть бути визначені за формулами (6.34) або (6.39). Скористаємося останньою, тоді

$$\frac{\delta y}{\delta t_1}=\frac{\delta y}{\delta x_3}=\frac{\partial y}{\partial x_3}+\frac{\partial y}{\partial x_1}\frac{\delta x_1}{\delta x_3}+\frac{\partial y}{\partial x_2}\frac{\delta x_2}{\delta x_3}; \quad \frac{\partial y}{\partial x_3}=4x_3; \quad \frac{\partial y}{\partial x_1}=x_2; \quad \frac{\partial y}{\partial x_2}=x_1;$$

$$\left(\frac{\delta y}{\delta t_1}\right)^{(0)}=4+1\cdot 2+1\cdot(-1)=5\ ;$$

$$\frac{\delta y}{\delta t_2}=\frac{\delta y}{\delta x_4}=\frac{\partial y}{\partial x_4}+\frac{\partial y}{\partial x_1}\frac{\delta x_1}{\delta x_4}+\frac{\partial y}{\partial x_2}\frac{\delta x_2}{\delta x_4}; \quad \frac{\partial y}{\partial x_4}=0;$$

$$\left(\frac{\delta y}{\delta t_2}\right)^{(0)}=0+1\cdot 0+1\cdot(-1)=-1\ .$$

Значення функції в початковій точці $y^{(0)}=3$. Інші елементи таблиці визначаються за умовами задачі: поточні значення залежних і незалежних змінних відповідають координатам точки початкового наближення $\mathbf{x}^{(0)}$, значення нижніх і верхніх границь залежних змінних добуваються з обмеження (8.55), а незалежних – із (8.56). Початкова таблиця диференціального алгоритму на першому кроці оптимізації відповідає табл.8.3.

Таблиця 8.3

	x_3	x_4			
	5	-1	3		
x_1	2	0	1	0	1.2
x_2	-1	-1	1	0	1.5
	1	1			
	0	0			
	2	3			

З табл.8.3. видно, що умовна похідна за незалежною змінною x_3 додатна: $\left(\frac{\delta y}{\delta x_3}\right)^{(0)}=5>0$, а поточне значення змінної $x_3^{(0)}=1$ не лежить на нижній границі $x_3^{+}=0$. Отже, за змінною x_3 має місце порушення необхідних умов першого типу. У цьому разі напрямок руху визначається умовою (8.31), а необхідний приріст змінної x_3 – за критерієм вибору (8.33).

Для обчислення приросту $\left.\Delta t_r^{(k)}\right|_{\frac{\delta y}{\delta t_r}=0}$ за формулою (8.35) необхідно

обчислити другу умовну похідну функції цілі за незалежною змінною x_3. Щоб уникнути зайвих обчислень, що мають місце при формуванні матриці других умовних похідних $\mathbf{S}$ за формулою (6.57), обчислимо другу умовну похідну за незалежною змінною за формулою

$$\frac{\delta^2 y}{\delta t_r^2} = \frac{\partial^2 y}{\partial t_r^2} + 2\sum_i^m \frac{\partial^2 y}{\partial s_i \partial t_r}\frac{\delta s_i}{\delta t_r} + \sum_i^m \frac{\partial y}{\partial s_i}\frac{\delta^2 s_i}{\delta t_r^2} + .$$

$$+\sum_i^m \sum_j^m \frac{\partial^2 y}{\partial s_i \partial s_j}\frac{\delta s_i}{\delta t_r}\frac{\delta s_j}{\delta t_r} . \qquad (8.58)$$

Наведену формулу, як і формулу (6.39) для перших умовних похідних, можна використовувати тільки у випадку, коли залежні змінні розв'язні явно щодо незалежних змінних, тобто відповідають вигляду (8.57). Оскільки в задачі ця умова виконується, обчислимо необхідну похідну $\frac{\delta^2 y}{\delta t_1^2}$ за формулою (8.58):

$$\frac{\delta^2 y}{\delta t_1^2} = \frac{\delta^2 y}{\delta x_3^2} = \frac{\partial^2 y}{\partial x_3^2} + 2\left(\frac{\partial^2 y}{\partial x_1 \partial x_3}\frac{\delta x_1}{\delta x_3} + \frac{\partial^2 y}{\partial x_2 \partial x_3}\frac{\delta x_2}{\delta x_3}\right) +$$

$$+\left(\frac{\partial y}{\partial x_1}\frac{\delta^2 x_1}{\delta x_3^2} + \frac{\partial y}{\partial x_2}\frac{\delta^2 x_2}{\delta x_3^2}\right) + \frac{\partial^2 y}{\partial x_1^2}\frac{\delta x_1}{\delta x_3}\frac{\delta x_1}{\delta x_3} +$$

$$. + 2\frac{\partial^2 y}{\partial x_1 \partial x_2}\frac{\delta x_1}{\delta x_3}\frac{\delta x_2}{\delta x_3} + \frac{\partial^2 y}{\partial x_2^2}\frac{\delta x_2}{\delta x_3}\frac{\delta x_2}{\delta x_3} .$$

$$\frac{\partial^2 y}{\partial x_3^2} = 4; \quad \frac{\partial^2 y}{\partial x_1 \partial x_3} = 0; \quad \frac{\partial^2 y}{\partial x_2 \partial x_3} = 0; \quad \frac{\delta^2 x_1}{\delta x_3^2} = \frac{-2}{x_4^2};$$

$$\frac{\delta^2 x_2}{\delta x_3^2} = 0; \quad \frac{\partial^2 y}{\partial x_1^2} = 0; \quad \frac{\partial^2 y}{\partial x_1 \partial x_2} = 1; \quad \frac{\partial^2 y}{\partial x_2^2} = 0;$$

$$\left(\frac{\delta^2 y}{\delta x^2}\right)^{(0)} = 4 + 2(0 \cdot 2 + 0(-1)) + (1(-2) + 1 \cdot 2) +$$
$$+ 0 \cdot 2 \cdot 2 + 2 \cdot 1 \cdot 2(-1) + 0(-1)(-1) = 2.$$

Згідно з (8.33)

$$\Delta t_1^{(0)} = \Delta x_3^{(0)} = \max_{\Delta t_r < 0} \left\{ 0 - 1; \quad -\frac{5}{2}; \quad \max_{\frac{\delta s_i}{\delta t_r} < 0} \frac{1{,}5 - 1}{-1}; \quad \max_{\frac{\delta s_i}{\delta t_r} > 0} \frac{0 - 1}{2} \right\} = -\frac{1}{2}.$$

Першою у процесі зменшення незалежної змінної x_3 (при лінійній апроксимації) свого нижньограничного значення досягає залежна змінна x_1. Переводячи змінну x_1 в розряд незалежних замість x_3, прирівнюючи її своєму нижньому значенню (нулю) і вирішуючи систему нелінійних рівнянь (8.53) – (8.54) при старих значеннях інших незалежних змінних щодо залежних змінних, одержимо нове припустиме рішення $\mathbf{x}^{(1)^{\mathrm{T}}} = [0 \quad 1{,}4142 \quad 0{,}5858 \quad 1]$.

У новій точці значення функції $y^{(1)} = 0{,}6863$ менше початкового $y^{(0)} = 3$. Отже, перший крок зроблено у бік умовного мінімуму.

На наступному кроці мінімізації складаємо нову таблицю диференціального алгоритму (табл.8.4) щодо точки $\mathbf{x}^{(1)}$ і нової системи залежних і незалежних змінних. Правила укладання таблиці ті самі, що і на першому кроці.

Таблиця 8.4

	x_1	x_4			
	2,2426	-1,6566	0,6863		
x_2	-0,3536	-0,7072	1,4142	0	1.5
x_3	0,3536	--0,707	0,5858	0	2
	0	1			
	0	0			
	1,2	3			

Аналіз табл.8.4 показує, що мінімум поки не досягнутий, оскільки за змінною x_4 порушена необхідна умова (8.13) – другий тип порушень (8.32). За критерієм (8.34) для поліпшення функції змінну x_4 треба збільшити на величину

$$\Delta x_4^{(1)} = \min_{\Delta t_r >}\left\{3-1;\ -\frac{-1{,}6566}{6{,}1426};\ \min_{\frac{\delta s_i}{\delta t_r}<0}\left\{\frac{0-1{,}4142}{-0{,}7072};\ \frac{0-0{,}5858}{-0{,}707}\right\}\right\} = 0{,}2696.$$

Приріст змінної обраний з умови обертання на нуль умовної похідної $\frac{\delta y}{\delta x_4}$. Отже, система розбивки змінних на залежні й незалежні залишається колишньою. Вирішуючи систему рівнянь (8.53) – (8.54) при новому значенні $x_4 = 1{,}2696$ і старому значенні $x_1 = 0$ щодо залежних змінних x_2 і x_3, одержимо нову точку наближення до умовного мінімуму $\mathbf{x}^{(2)^{\mathrm{T}}} = \begin{bmatrix}0 & 1{,}2086 & 0{,}4656 & 1{,}4069\end{bmatrix}$. Таблиця диференціального алгоритму для чергового кроку мінімізації відповідає табл.8.5.

Таблиця 8.5

	x_1	x_4			
	2,2925	-0,083	0,4336		
x_2	-0,4137	-0,8274	1,2086	0	1.5
x_3	0,582	0,0446	0,4336	0	2
	0	1,2696			
	0	0			
	1,2	3			

У новій точці $\mathbf{x}^{(2)}$ необхідні умови умовного мінімуму виконуються із заданою точністю по всіх змінних. Так, по змінній x_1 виконується умова (8.12), а по змінній x_4 – умова (8.14), оскільки $\left|\left(\dfrac{\delta y}{\delta x_4}\right)\right| = 0{,}083 < \varepsilon = 0{,}1$.

Для остаточного вирішення задачі залишається тільки перевірити виконання достатніх умов мінімуму в точці $\mathbf{x}^{(2)}$, причому виконання другої групи достатніх умов, тому що не всі незалежні змінні лежать на границі області Ω. Перевірка другої групи достатніх умов полягає в перевірці додатної визначеності матриці других умовних похідних $\mathbf{S}_{22}$, що в розв'язуваній задачі складається з одного додатного елемента $\left(\dfrac{\delta^2 y}{\delta x_4^2}\right)^{(2)}$, який дорівнює 0,1784. Оскільки достатні умови виконуються, то точка $\mathbf{x}^{(2)}$ є умовним мінімальним рішенням, а значення функції в цій точці $y^{(2)} = 0{,}4336$ – шуканим умовним мінімумом.

16. Знайти локальний умовний мінімум за допомогою диференціального алгоритму в задачі (докладне вирішення задачі міститься в [3] на стор. 91 – 101)

$$y = x_7 x_8 + 2x_8 \to \min_{x \in \mathbf{\Omega}} \ ;$$

$$\mathbf{\Omega}: \quad x_1 x_7^2 + x_7^2 + 4x_1 + 4x_1 x_7 + 4x_3 + x_4 + 4x_7 - x_8 + 11 = 0 \ ;$$

$$x_1 x_7^2 + x_2 x_7^2 + 2x_7^2 + 4x_1 + 4x_1 x_7 + x_2 +$$
$$+ 2x_2 x_7 + 4x_3 + x_5 + 6x_7 - x_8 + 10 = 0 \ ;$$

$$x_1 x_7^2 + x_2 x_7^2 + 3x_7^2 + 4x_1 + 4x_1 x_7 + x_2 +$$
$$+ 2x_2 x_7 + x_6 x_7^2 + 6x_7 - x_8 + 5 = 0 \ ;$$

$$0 \le x_i \le 20, \quad i = \overline{1,6} \ ; \qquad 20 \le x_8 \le 70 \, ;$$

$$\mathbf{x}^{(0)\,\mathrm{T}} = \begin{bmatrix} 0 & 0 & 0 & 18 & 0 & 0 & 3 & 50 \end{bmatrix}.$$

17. Вказати відмінні риси загальної задачі математичного програмування при лінійних обмеженнях.

18. Чому спрощується процедура обчислення похідних при лінійних обмеженнях?

19. Скільки явних екстремальних точок в області припустимих рішень має строго опукла або строго увігнута функція при обмеженнях (8.37) – (8.38)?

20. Знайти умовний мінімум функції $y = e^{(x_1 - 0,5)^2 + (x_2 - 1,5)^2}$, якщо змінні x_1 і x_2 повинні задовольняти системі нерівностей

$$0,5x_1 + x_2 \le 1;$$

$$x_1 + 0,5x_2 \le 1;$$

$$1,5x_1 + x_2 \ge 1;$$

$$0,2 \le x_i \le 0,8, \quad i = 1,2 \ .$$

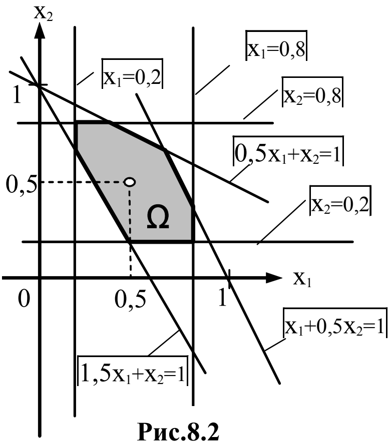


Рис.8.2

Розв'язання. Точність обчислень ε і початкова точка $\mathbf{x}^{(0)}$ за умовами задачі не задані. Виберемо їх самостійно. Нехай $\varepsilon = 0{,}001$. Вибір точки $\mathbf{x}^{(0)}$ повинен здійснюватися відповідно до вимоги $\mathbf{x} \in \mathbf{\Omega}$. Оскільки число змінних завдання дорівнює двом, виберемо її графічно, попередньо побудувавши на площині область припустимих рішень (рис.8.2). Будь-яка точка з області $\mathbf{\Omega}$ може служити початковим наближенням. Виберемо як початкову точку $x^{(0)^{\mathrm{T}}} = \begin{bmatrix} 0{,}5 & 0{,}5 \end{bmatrix}$. Приведемо обмеження–нерівності задачі до обмежень–рівностей вигляду (8.36) – (8.38):

$$y = e^{(x_1 - 0{,}5)^2 + (x_2 - 1{,}5)^2} \to \min_{\mathbf{x} \in \mathbf{\Omega}} \; ; \tag{8.59}$$

$$\mathbf{\Omega}: \quad 0{,}5x_1 + x_2 + x_2 - 1 = 0; \tag{8.60}$$

$$x_1 + 0{,}5x_2 + x_4 - 1 = 0 \; ; \tag{8.61}$$

$$1{,}5x_1 + x_2 - x_5 - 1 = 0 \; ; \tag{8.62}$$

$$0{,}2 \le x_i \le 0{,}8, \quad i = 1{,}2 \; ; \tag{8.63}$$

$$0 \le x_j, \quad i = 3{,}4{,}5 \; . \tag{8.64}$$

Як видно з (8.63) – (8.64), одні змінні обмежені з двох боків, інші – з одного. Віднесемо останні змінні x_3, x_4, x_5 до залежних, а x_1, x_2 – до незалежних.

Приведемо систему обмежень (8.60) – (8.62) до вигляду (8.43)

$$\begin{bmatrix} x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} -0,5 & -1 \\ -1 & -0,5 \\ 1,5 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} . \qquad (8.65)$$

При невиродженості системи лінійних рівнянь (8.60) – (8.62) пророблену процедуру завжди можна здійснити, наприклад, за допомогою жорданових виключень.

Отриманий вираз (8.65) дозволяє визначити поточні значення незалежних змінних при відомих значеннях незалежних. Так, при початкових значеннях $x_1 = x_2 = 0,5$ значення залежних змінних стануть рівними $x_3 = x_4 = x_5 = 0,25$.

Залежні змінні, будучи додатковими, не входять до функції (8.59). А це значить, що в початковій точці всі частинні похідні по залежних змінних рівні нулю. Внаслідок цього всі умовні похідні по незалежних змінних рівні частинним похідним по цих же змінних:

$$\left(\frac{\delta y}{\delta x_1}\right)^{(0)} = \left(\frac{\partial y}{\partial x_1}\right)^{(0)} = e^{(x_1-0,5)^2+(x_2-1,5)^2} \cdot 2(x_1 - 0,5)\Big|_{\substack{x_1=0,5 \\ x_2=0,5}} = 0 \ ;$$

$$\left(\frac{\delta y}{\delta x_2}\right)^{(0)} = \left(\frac{\partial y}{\partial x_2}\right)^{(0)} = e^{(x_1-0,5)^2+(x_2-1,5)^2} \cdot 2(x_2 - 1,5)\Big|_{\substack{x_1=0,5 \\ x_2=0,5}} = -5,4366$$

Значення функції в початковій точці наближення $y^{(0)} = e \approx 2,7183$.

Отримані значення залежних змінних, умовних похідних по незалежних змінних, функції в початковій точці наближення разом із системою (8.65) і обмеженнями (8.63) – (8.64) дозволяють сформувати початкову таблицю диференціального алгоритму (табл.8.6).

Таблиця 8.6

	x_1	x_2				
	0	-5,4366	2,7183			
x_3	-0.5	-1	1	0.25	0	∞
x_4	-1	-0,5	1	0,25	0	∞
x_5	1,5	1	-1	0,25	0	∞
	0,5	0,5				
	0,2	0,2				
	0,8	0,8				

З аналізу табл.8.6 випливає, що по змінній x_1 виконується необхідна умова умовного локального мінімуму (8.14), а по змінній x_2 є порушення необхідних умов другого типу (8.32): $\left(\dfrac{\delta y}{\delta x_2}\right)^{(0)} < 0$, а $x_2^{(0)} \neq x_2^{++}$. Збільшуючи змінну x_2, ми можемо поліпшити (зменшити) функцію y. Вибір приросту змінної здійснимо за критерієм (8.48), обчислення його елемента $\text{Д}x_2^{(0)}\Big|_{\frac{\delta y}{\delta t_r}=0}$ здійснимо за формулою $\text{Д}x_2^{(0)}\Big|_{\frac{\delta y}{\delta t_r}=0} = -\dfrac{\left(\dfrac{\delta y}{\delta x_2}\right)^{(0)}}{\left|\left(\dfrac{\delta^2 y}{\delta x_2^2}\right)^{(0)}\right|}$.

Друга умовна похідна, необхідна для розрахунку $\text{Д}x_2^{(0)}\Big|_{\frac{\delta y}{\delta t_r}=0}$, як і перша, і з тієї причини в точності відповідає другій частинній похідній, тобто $\left(\dfrac{\delta^2 y}{\delta x_2^2}\right)^{(0)} = \left(\dfrac{\partial^2 y}{\partial x_2^2}\right)^{(0)} = 2e^{(x_1-0,5)^2+(x_2-1,5)^2} + 4e^{(x_1-0,5)^2+(x_2-1,5)^2}(x_2-1,5)^2 = 6e\cdot$

Критерій (8.48) внаслідок односторонньої обмеженості залежних змінних спроститься і для поточного кроку набуде вигляду

$$\Delta x_2^{(0)} = \min_{\Delta x_2 > 0}\left\{ x_2^{++} - x_2^{(0)};\ \Delta x_2^{(0)}\Big|_{\frac{\delta y}{\delta x_2}=0};\ \min_{b_{i2}<0}\frac{s_i^{+} - s_i^{(0)}}{b_{i2}^{(0)}} \right\},$$

або, після підстановки відповідних значень,

$$\Delta x_2^{(0)} = \min_{\Delta x_2 > 0}\left\{ 0,8 - 0,5;\ -\frac{-2e}{|6e|};\ \min_{b_{i2}<0}\left\{\frac{0-0,25}{-1};\frac{0-0,25}{-0,5}\right\}\right\} = 0,25 \ .$$

Таким чином, при збільшенні змінної x_2 першою своєї нижньої границі досягає залежна змінна x_3. Отже, її треба перевести в розряд незалежних замість змінної x_2. Транспозицію змінних здійснимо за допомогою жорданових виключень. Жордановим виключенням піддаються тільки елементи табл.8.6, обмежені подвійною рамкою. Щоб визначити значення умовних похідних для табл.8.7, треба затратити трохи більше зусиль, ніж на першому кроці. Відповідно до (8.49)

$$\left(\frac{\delta y}{\delta x_1}\right)^{(1)} = \left(\frac{\partial y}{\partial x_1}\right)^{(1)} + b_{11}^{(1)}\left(\frac{\partial y}{\partial x_2}\right)^{(1)} + b_{21}^{(1)}\left(\frac{\partial y}{\partial x_4}\right)^{(1)} + b_{31}^{(1)}\left(\frac{\partial y}{\partial x_5}\right)^{(1)} =$$

$$= 0 - (-0,5)\cdot(-1,5)e^{0,75^2} + 0,75\cdot 0 + 1\cdot 0 = 1,3163\ ;$$

$$\left(\frac{\delta y}{\delta x_3}\right)^{(1)} = \left(\frac{\partial y}{\partial x_3}\right)^{(1)} + b_{12}^{(1)}\left(\frac{\partial y}{\partial x_2}\right)^{(1)} + b_{22}^{(1)}\left(\frac{\partial y}{\partial x_4}\right)^{(1)} + b_{32}^{(1)}\left(\frac{\partial y}{\partial x_5}\right)^{(1)} =$$

$$= 0 + (-1)\cdot(-1,5)e^{0,75^2} + 0,5\cdot 0 + (-1)\cdot 0 = 2,6326\ .$$

Поточні значення залежних змінних визначаються відповідно до системи (8.43). Зменшення поточного значення функції ($y^{(1)} = 1{,}7551$) у новій точці наближення $\mathbf{x}^{(1)}$ свідчить про правильність протікання процесу мінімізації. Інші елементи табл.8.7 заповнюються за умовами задачі або передаються з попередньої таблиці.

Таблиця 8.7

	x_1	x_3				
	1,3163	2,6326	1,7551			
x_2	-0.5	-1	1	0.75	0,2	0,8
x_4	-0,75	0,5	0,5	0,125	0	∞
x_5	1	-1	0	0,5	0	∞
	0,5	0				
	0,2	0				
	0,8	∞				

На наступному кроці знову аналізуємо таблицю диференціального алгоритму з метою перевірки виконання необхідних умов локального мінімуму в новій точці наближення. Тепер умови не виконуються по змінній x_1 : $\left(\dfrac{\delta y}{\delta x_1}\right)^{(1)} > 0$, а $x_1^{(1)} \neq x_1^{+}$. Оскільки порушення відноситься до першого типу (8.31), то для обчислення приросту $\Delta x_1^{(1)}$ використовуємо критерій (8.47), але попередньо визначимо згідно з (8.50) другу умовну похідну функції по змінній x_1 :

$$\left(\frac{\delta^2 y}{\delta x_1^2}\right)^{(1)} = \left(\frac{\partial^2 y}{\partial x_1^2}\right)^{(1)} + 2b_{11}^{(1)}\left(\frac{\partial^2 y}{\partial x_1 \partial x_2}\right)^{(1)} + \left(b_{21}^{(1)}\right)^2\left(\frac{\partial^2 y}{\partial x_2^2}\right)^{(1)} =$$

$$= 2e^{0,75^2} + 2\cdot(-0,5)\cdot 0 + (-0,5)^2 \cdot 2e^{0,75^2}\left(1 + 2\cdot 0,75^2\right) = 4,3876 \;.$$

Тоді $\text{Д}x_1^{(1)}\Big|_{\frac{\delta y}{\delta t_r}=0} = -\dfrac{\left(\dfrac{\delta y}{\delta x_1}\right)^{(1)}}{\left|\left(\dfrac{\delta^2 y}{\delta x_1^2}\right)^{(1)}\right|} = -\dfrac{1,3163}{4,3876} = -0,3$, а критерій (8.47) набуде вигляду

$$\Delta x_1^{(1)} = \max_{\Delta x_2<0}\left\{0,2-0,5;\, -0,3;\, \max_{b_{i2}<0}\left\{\frac{0,8-0,75}{-0,5}\right\};\max_{b_{i2}>0}\left\{\frac{0-0,5}{1}\right\}\right\} = -0,1.$$

Обчислення показали, що першою досягає своєї верхньої границі залежна змінна x_2, тому її слід повернути в розряд незалежних замість змінної x_1. Чергові жорданові перетворення і обчислення умовних похідних по незалежних змінних приводять до нової таблиці диференціального алгоритму (табл.8.8)

Таблиця 8.8

	x_2	x_3				
	-1,648	0,6592	1,648			
x_1	-2	-2	2	0.4	0,2	0,8
x_4	1,5	2	-1	0,2	0	∞
x_5	-2	-3	2	0,4	0	∞
	0,8	0				
	0,2	0				
	0,8	∞				

У новій точці наближення $\mathbf{x}^{(2)\mathrm{T}} = \left[0,4 \quad 0,8 \quad 0 \quad 0,2 \quad 0,4\right]$ необхідні умови по всіх незалежних змінних виконуються:

$$\left(\frac{\delta y}{\delta x_2}\right)^{(2)} < 0, \quad x_2^{(2)} = x_2^{++}; \qquad \left(\frac{\delta y}{\delta x_3}\right)^{(2)} > 0, \quad x_3^{(2)} = x_3^{+} .$$

Оскільки необхідні умови для даної точки одночасно є і достатніми (немає нульових умовних похідних), то $\mathbf{x}^{(2)}$ – шукана точка умовного мінімуму. Вилучивши із знайденого вектора рішення складові, що відповідають додатковим змінним, одержуємо остаточний розв'язок $\mathbf{x}^{*\mathrm{T}} = [0{,}4 \quad 0{,}8]$, $y^* = 1{,}648$.

8.7. Фонд індивідуальних завдань

Індивідуальне завдання №28. За допомогою диференціального алгоритму знайти локальний мінімум функції

$$y(\mathbf{x}) = a_1 x_1^3 + a_2 x_2^3 + a_3\left(x_1^2 x_2 + x_1 x_2^2\right) + a_4\left(x_1 + x_2\right) \to \min_{\mathbf{x}\in\mathbf{\Omega}}$$

з точністю $\varepsilon = 0{,}1$, якщо область припустимих рішень $\mathbf{\Omega}$ формується з обмеження типу рівність $x_1 x_2 - 1 = 0$ і двомірного гіперпаралелепіпеда з індивідуального завдання №27. Коефіцієнти a_1, a_2, a_3 і a_4 функції y вибрати з табл.6.3. Початкову точку наближення вибрати самостійно.

Індивідуальне завдання №29. За допомогою диференціального алгоритму знайти локальний мінімум функції

$$y(\mathbf{x}) = a_1 x_1^3 + a_2 x_2^3 + a_3\left(x_1^2 x_2 + x_1 x_2^2\right) + a_4\left(x_1 + x_2\right) \to \min_{\mathbf{x}\in\mathbf{\Omega}}$$

з точністю $\varepsilon = 0{,}1$. Коефіцієнти a_1, a_2, a_3 і a_4 функції y вибрати з табл.6.3. Початкову точку наближення вибрати самостійно. Область припустимих рішень $\mathbf{\Omega}$ формується з двовимірного гіперпаралелепіпеда індивідуального завдання №27 і системи нерівностей

1.
$$\begin{aligned} 2x_1 + x_2 &\le 1; \\ 2x_1 + x_2 &\ge -1; \\ x_1 - x_2 &\le 1; \\ x_1 - x_2 &\ge -1. \end{aligned}$$

2.
$$\begin{aligned} x_1 + x_2 &\le 2; \\ x_1 + 2x_2 &\ge -1; \\ x_1 - x_2 &\le 1; \\ x_1 - x_2 &\ge -2. \end{aligned}$$

3.
$$\begin{aligned} x_1 + 2x_2 &\le 2; \\ 2x_1 + 2x_2 &\ge -1; \\ x_1 - x_2 &\le 1. \end{aligned}$$

4.
$$\begin{aligned} x_1 + x_2 &\ge 1; \\ 2x_1 + x_2 &\le 4; \\ -x_1 - x_2 &\ge 1. \end{aligned}$$

5.
$$\begin{aligned} x_1 + x_2 &\le 2; \\ x_1 + 2x_2 &\ge -1; \\ 2x_1 - x_2 &\le 2; \\ x_1 - x_2 &\ge -1. \end{aligned}$$

6.
$$\begin{aligned} x_1 + x_2 &\le 3; \\ -x_1 - x_2 &\le 2; \\ -x_1 + x_2 &\ge 2; \\ x_1 - x_2 &\ge -1. \end{aligned}$$

7.
$$\begin{aligned} x_1 + 4x_2 &\ge 1; \\ x_1 + 2x_2 &\le 3; \\ x_1 - x_2 &\ge -1. \end{aligned}$$

8.
$$\begin{aligned} -x_1 - 3x_2 &\le 1; \\ x_1 + x_2 &\le 3; \\ 2x_1 - x_2 &\ge 3. \end{aligned}$$

9.
$$\begin{aligned} 2x_1 + x_2 &\ge -1; \\ x_1 + x_2 &\le 2; \\ x_1 - x_2 &\ge -1; \\ 2x_1 - x_2 &\le 3. \end{aligned}$$

10.
$$\begin{aligned} x_1 + 2x_2 &\ge 1; \\ 2x_1 + x_2 &\le 3; \\ -x_1 + x_2 &\ge -2; \\ x_1 - 2x_2 &\le 1. \end{aligned}$$

11.
$$\begin{aligned} x_1 + x_2 &\le 4; \\ -x_1 - x_2 &\le 2; \\ x_1 - x_2 &\le 2. \end{aligned}$$

12.
$$\begin{aligned} -x_1 - x_2 &\le 0; \\ 2x_1 + x_2 &\le 3; \\ -x_1 - x_2 &\le 1. \end{aligned}$$

13.
$$\begin{aligned} x_1 + x_2 &\le 4; \\ -x_1 - x_2 &\le -3; \\ 2x_1 - x_2 &\ge -1; \\ x_1 - 2x_2 &\le -2. \end{aligned}$$

14.
$$\begin{aligned} x_1 + x_2 &\ge 1; \\ 2x_1 + x_2 &\le 3; \\ x_1 - x_2 &\ge 3; \\ -x_1 + 2x_2 &\ge 2. \end{aligned}$$

15.
$$\begin{aligned} -x_1 - 2x_2 &\ge -2; \\ 2x_1 + 2x_2 &\ge 1; \\ x_1 - x_2 &\le 1. \end{aligned}$$

16.
$$\begin{aligned} -x_1 - x_2 &\ge -1; \\ 2x_1 + x_2 &\le 4; \\ -x_1 + x_2 &\ge 1. \end{aligned}$$

17.
$$\begin{aligned} -2x_1 - x_2 &\ge -1; \\ x_1 + x_2 &\ge -1; \\ x_1 - x_2 &\le 1; \\ x_1 - x_2 &\ge -1. \end{aligned}$$

18.
$$\begin{aligned} -x_1 - x_2 &\ge -2; \\ -x_1 - x_2 &\le 1; \\ 2x_1 - x_2 &\ge 1; \\ x_1 - 2x_2 &\le -2. \end{aligned}$$

19.
$$\begin{aligned} -x_1 - 4x_2 &\le -1; \\ x_1 + 2x_2 &\le 3; \\ x_1 - x_2 &\ge -1. \end{aligned}$$

20.
$$\begin{aligned} -x_1 - 3x_2 &\ge -1; \\ x_1 + x_2 &\le 3; \\ 2x_1 - x_2 &\le 3. \end{aligned}$$

21.
$$\begin{aligned} -x_1 - x_2 &\ge -3; \\ x_1 + 2x_2 &\ge -1; \\ 2x_1 - x_2 &\le 2; \\ x_1 - x_2 &\ge -1. \end{aligned}$$

22.
$$\begin{aligned} -x_1 - 2x_2 &\ge -3; \\ -x_1 - x_2 &\le 2; \\ -x_1 + x_2 &\ge 2; \\ x_1 - x_2 &\ge -1. \end{aligned}$$

23.
$$\begin{aligned} -2x_1 - 4x_2 &\ge -4; \\ -x_1 - x_2 &\le 2; \\ x_1 - x_2 &\le 2. \end{aligned}$$

24.
$$\begin{aligned} x_1 + x_2 &\ge 0; \\ 2x_1 + x_2 &\le 3; \\ -x_1 + x_2 &\le 1. \end{aligned}$$

25.
$$\begin{aligned} -2x_1 - x_2 &\le -1; \\ x_1 + x_2 &\le 2; \\ x_1 - x_2 &\ge -1; \\ 2x_1 - x_2 &\le 3. \end{aligned}$$

26.
$$\begin{aligned} -x_1 - 2x_2 &\ge -2; \\ -x_1 - x_2 &\le 2; \\ -x_1 + x_2 &\ge 2; \\ x_1 - x_2 &\ge -2. \end{aligned}$$

27.
$$\begin{aligned} -2x_1 - 4x_2 &\ge -4; \\ -x_1 - x_2 &\le 0; \\ x_1 - x_2 &\le 2. \end{aligned}$$

28.
$$\begin{aligned} x_1 + x_2 &\ge 0; \\ 2x_1 + x_2 &\le 4; \\ -x_1 + x_2 &\le 2. \end{aligned}$$

29.
$$\begin{aligned} 2x_1 + x_2 &\ge 0; \\ x_1 + x_2 &\le 2; \\ x_1 - x_2 &\ge -1; \\ 2x_1 - x_2 &\le 3. \end{aligned}$$

30.
$$\begin{aligned} x_1 + 2x_2 &\ge 1; \\ 2x_1 + x_2 &\le 3; \\ -x_1 + x_2 &\ge -2; \\ x_1 - 2x_2 &\le 0. \end{aligned}$$

Розділ 9. ЛІНІЙНЕ ПРОГРАМУВАННЯ

Лінійне програмування – найбільш вивчений розділ математичного програмування. Задача лінійного програмування є частинним випадком загальної задачі математичного програмування. Така задача досить часто зустрічається в інженерно-економічній і фінансовій діяльності. Диференціальний алгоритм вирішення загальної задачі математичного програмування може бути з успіхом використаний для розв'язання задач лінійного програмування. Через значне спрощення алгоритму оптимізації внаслідок лінійності цільової функції і обмежень він викликає особливий інтерес і заслуговує окремого детального розгляду.

9.1. Постановка задачі лінійного програмування

Загальна задача лінійного програмування формулюється таким чином: *знайти оптимум лінійної функції* $y(\mathbf{x})$*, якщо на змінні задачі накладені лінійні обмеження у вигляді рівностей і нерівностей, а також обмеження у вигляді гіперпаралелепіпеда..*

Аналітичний запис цієї задачі такий:

$$y(\mathbf{x}) = \mathbf{c}^{\mathrm{T}}\mathbf{x} + c_0 \to \underset{\mathbf{x}\in\mathbf{\Omega}\subset\mathbf{R}^n}{\mathrm{opt}} \quad , \tag{9.1}$$

$$\Omega: \quad \mathbf{A}_1\mathbf{x} + \mathbf{b}_1 \le 0 \;; \tag{9.2}$$

$$\mathbf{A}_2\mathbf{x} + \mathbf{b}_2 = 0 \;; \tag{9.3}$$

$$\mathbf{A}_3\mathbf{x} + \mathbf{b}_3 \ge 0 \;; \tag{9.4}$$

$$\mathbf{x} \ge 0 \;, \tag{9.5}$$

де $\mathbf{x}$ – n-вимірний вектор дійсних змінних; $\mathbf{c}$ – n-вимірний вектор коефіцієнтів функції цілі; c_0 – вільний член функції цілі; $\mathbf{A}_1$, $\mathbf{A}_2$, $\mathbf{A}_3$ –

матриці коефіцієнтів лінійних систем розмірності $m_1 \times n$, $m_2 \times n$, $m_3 \times n$ відповідно, $m_2 < n$; $\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3$ – вектори вільних членів обмежень розмірності $m_1 \times 1$, $m_2 \times 1$, $m_3 \times 1$ відповідно.

Частинні задачі лінійного програмування можуть не містити однієї або двох систем обмежень типу (9.2) – (9.5), однаково яких. Крім того, замість умови невід'ємності (9.5) може мати місце двостороння або одностороння обмеженість змінних.

Задачу, складену з (9.1), (9.2) і (9.5), називають *стандартною* задачею лінійного програмування.

Канонічна, або *основна* задача лінійного програмування має вигляд [1]:

$$y(\mathbf{x}) = \mathbf{c}^{\mathrm{T}} \mathbf{x} + c_0 \to \max_{\mathbf{x} \in \Omega \subset \mathbf{R}^n}, \tag{9.6}$$

$$\mathbf{\Omega}: \quad \mathbf{A}\mathbf{x} + \mathbf{b} = 0; \tag{9.7}$$

$$\mathbf{x} \geq 0, \tag{9.8}$$

де $\mathbf{A}$ – матриця коефіцієнтів розмірності $m \times n$, $m < n$; $\mathbf{b}$ – вектор вільних членів розмірності $m \times 1$.

Очевидно, що обмеження-нерівність типу "$\leq$" можна перетворити на обмеження-рівність *додаванням* до його лівої частини додаткової невід'ємної змінної, а кожне обмеження-нерівність типу "$\geq$" – в обмеження-рівність *відніманням* з його лівої частини додаткової невід'ємної змінної. Задачу мінімізації лінійної функції y множенням останньої на -1 можна звести до задачі максимізації. Таким чином, задачу лінійної оптимізації (9.1) – (9.5) завжди можна перетворити на задачу (9.6) – (9.8) і навпаки.

Укладання математичної моделі загальної задачі математичного програмування або її канонічної форми вимагає певних зусиль. Особливий інтерес являє собою векторне-матричне подання канонічної задачі лінійного програмування, що дозволяє легко переходити до таблиці диференціального алгоритму.

Розглянемо декілька прикладів одержання математичних моделей задач лінійного програмування поки без їхнього остаточного вирішення.

Приклад 9.1. Скласти математичну модель задачі лінійного програмування такого змісту:

Для виготовлення трьох видів виробів *A, B, C* використовується токарне, фрезерне, зварювальне і шліфувальне устаткування. Витрати часу на обробку одного виробу для кожного з типів устаткування зазначені в табл.9.1. У ній же поданий загальний фонд робочого часу кожного з типів використовуваного устаткування, а також прибуток від реалізації одного виробу кожного виду.

Треба визначити, скільки виробів і якого виду слід виготовити підприємству, щоб прибуток від їхньої реалізації був максимальним.

Таблиця 9.1

Тип Устаткування	Витрати часу на обробку одного виробу, *верстато-год.*			Загальний фонд робочого часу устаткування, *год.*
	A	B	C	
Фрезерне	2	4	5	120
Токарське	1	8	6	280
Зварювальне	7	4	5	240
Шліфувальне	4	6	7	360
Прибуток, *грн.*	10	14	12	–

Розв'язання. Припустимо, що буде виготовлено x_1 одиниць виробів виду *A*, x_2 – виду *B* і x_3 – виду *C*. Тоді для виробництва такої кількості виробів потрібно затратити $2x_1 + 4x_2 + 5x_3$ верстато-годин фрезерного устаткування. Оскільки загальний фонд робочого часу верстатів даного типу не може перевищувати 120, повинна виконуватися нерівність $2x_1 + 4x_2 + 5x_3 \leq 120$.

Аналогічні міркування щодо можливого використання токарського, зварювального і шліфувального устаткування приведуть до таких нерівностей:

$$x_1 + 8x_2 + 6x_3 \leq 280\ ;$$

$$7x_1 + 4x_2 + 5x_3 \leq 240\ ;$$

$$4x_1 + 6x_2 + 7x_3 \le 360 \ .$$

При цьому кількість виготовлених виробів не може бути від'ємною:

$$x_1 \ge 0; \qquad x_2 \ge 0; \qquad x_3 \ge 0 \ . \tag{9.9}$$

Далі, якщо буде виготовлено x_1 одиниць виробів виду *A*, x_2 – виду *B* і x_3 – виду *C*, то прибуток становитиме

$$y = 10x_1 + 14x_2 + 12x_3 \ . \tag{9.10}$$

Підприємство зацікавлене одержувати максимальний прибуток від реалізації своєї продукції. Отже, необхідно серед усіх невід'ємних рішень системи нерівностей

$$\begin{gathered} 2x_1 + 4x_2 + 5x_3 \le 120; \\ x_1 + 8x_2 + 6x_3 \le 280; \\ 7x_1 + 4x_2 + 5x_3 \le 240; \\ 4x_1 + 6x_2 + 7x_3 \le 360 \end{gathered} \tag{9.11}$$

знайти таке, при якому функція (9.10) набуває максимального значення.

Оскільки функція (9.10) лінійна і система (9.11) містить тільки лінійні нерівності, розглянута задача є задача лінійного програмування, яка в алгебраїчній формі запишеться таким чином:

$$y = 10x_1 + 14x_2 + 12x_3 \to \max_{x_1, x_2, x_3 \in \Omega},$$

$$\Omega: \quad \begin{gathered} 2x_1 + 4x_2 + 5x_3 \le 120; \\ x_1 + 8x_2 + 6x_3 \le 280; \\ 7x_1 + 4x_2 + 5x_3 \le 240; \\ 4x_1 + 6x_2 + 7x_3 \le 360; \\ x_1, x_2, x_3 \ge 0 \ . \end{gathered}$$

Приклад 9.2. Скласти математичну модель задачі лінійного програмування такого змісту. Продукцією міського молочного заводу є молоко, кефір і сметана, розфасовані в пляшки. На виробництво 1 *т* молока, кефіру і сметани потрібно відповідно 1010, 1010 і 9450 *кг* молока. При цьому витрати робочого часу при розливі 1 *т* молока і кефіру складають 0,18 і 0,19 верстато-годин, на розфасовці 1 *т* сметани зайняті спеціальні автомати протягом 3,25 години. Усього для виробництва суцільномолочної продукції завод може використовувати 136000 *кг* молока. Основне устаткування може бути зайняте протягом 21,4 верстато-годин, а автомати по розфасовці сметани – протягом 16,25 годин. Прибуток від реалізації 1 *т* молока, кефіру і сметани відповідно дорівнює 30, 22, 136 *грн.* Завод повинен щодня виготовляти не менше 100 *т* молока, розфасованого в пляшки. На виробництво іншої продукції обмежень немає.

Треба визначити, яку продукцію і в якій кількості слід щодня виготовляти заводу, щоб прибуток від її реалізації був максимальний.

Розв'язання. Припустимо, що молочний завод буде щодня виробляти x_1 тонн молока, x_2 тонн кефіру і x_3 тонн сметані. Тоді йому для виготовлення цієї продукції необхідно $1010x_1 + 1010x_2 + 9450x_3$ тонн молока. Оскільки завод може використовувати щодня не більше 136000 *т* молока, повинна виконуватися нерівність $1010x_1 + 1010x_2 + 9450x_3 \le 136000$.

Аналогічні міркування, проведені щодо можливого використання ліній розливу суцільномолочної продукції і автоматів з розфасовки сметани, дозволяють записати наступні нерівності: $0{,}18x_1 + 0{,}19x_2 \le 21{,}4$; $3{,}25x_3 \le 16{,}25$. Оскільки щодня повинно вироблятися не менше 100 *т* молока, то $x_1 \ge 100$. Далі, за своїм економічним змістом змінні x_2, x_3 можуть приймати тільки невід'ємні значення: $x_2 \ge 0; \quad x_3 \ge 0$. Загальний прибуток від реалізації x_1 тонн молока, x_2 тонн кефіру і x_3 тонн сметани дорівнює $y = 30x_1 + 22x_2 + 136x_3$ *гр.* Таким чином, приходимо до такої математичної постановки задачі:

$$y = 30x_1 + 22x_2 + 136x_3 \to \max_{x_1, x_2, x_3 \in \Omega} \; ;$$

$$\Omega: \quad 1010x_1 + 1010x_2 + 9450x_3 \le 136000 \; ;$$

$$0{,}18x_1 + 0{,}19x_2 \le 21{,}4 \ ;$$

$$x_1 \ge 100 \ ; \ x_2 \ge 0; \quad 0 \le x_3 \le 5 \, .$$

Приклад 9.3. На швейній фабриці тканина може бути розкроєна декількома способами для виготовлення потрібних деталей швейних виробів. Нехай при j-му $\left(j = \overline{1,n}\right)$ варіанті розкрою 100 *м²* тканини виготовляється a_{ij} деталей i-го $\left(i = \overline{1,m}\right)$ виду, а відходи при даному варіанті розкрою дорівнюють c_j *м²*. Знаючи, що деталей i-го виду слід виготовляти b_i штук, потрібно розкроїти тканину так, щоб була отримана необхідна кількість деталей кожного виду при мінімальних загальних відходах.

Скласти математичну модель задачі.

Розв'язання. Припустимо, що j-й варіант розкрою становить x_j сотень *м²* тканини. Оскільки при розкрої 100 м² тканини за j-м варіантом утворюється a_{ij} деталей i-го виду, за всіма варіантами розкрою з використовуваних тканин буде отримано $a_{i1}x_1 + a_{i2}x_2 + \ldots + a_{in}x_n$ деталей i-го виду. Оскільки повинно бути виготовлене b_i деталей даного виду, то $a_{i1}x_1 + a_{i1}x_1 + \ldots + a_{in}x_1 = b_i, \quad i = \overline{1,m}$. Загальний розмір відходів по всіх варіантах розкрою складе $y = c_1x_1 + c_2x_2 + \ldots + c_nx_n$.

Таким чином, приходимо до такої задачі: знайти мінімум функції $y = \sum_{j=1}^{n} c_j x_j$ за умови, що її змінні задовольняють системі рівнянь $\sum_{j=1}^{n} a_{ij}x_j - b_i = 0 \quad \left(i = \overline{1,m}\right)$ і невід'ємні за значенням, тобто $x_j \ge 0 \quad \left(j = \overline{1,n}\right)$. Така модель одночасно є канонічною, векторно-матрична форма якої записується таким чином:

$$y(\mathbf{x}) = \mathbf{c}^{\mathrm{T}}\mathbf{x} + c_0 \to \min_{\mathbf{x}\in\mathbf{\Omega}\subset\mathbf{R}^n} ,$$

$$\mathbf{\Omega:} \quad \mathbf{Ax} + \mathbf{b} = 0 ;$$
$$\mathbf{x} \geq 0 ,$$

Приклад 9.4. Перетворити в канонічну і записати у векторно-матричній формі таку задачі лінійного програмування:

$$y = 3x_1 - 2x_2 - 5x_4 + x_5 \to \max_{x_j\in\mathbf{\Omega}\subset\mathbf{R}^5} ,$$

$$\mathbf{\Omega:} \quad 2x_1 + x_3 - x_4 + x_5 \leq 2 ;$$
$$x_1 - x_3 + 2x_4 + x_5 \leq 3 ;$$
$$2x_2 + x_3 - x_4 + 2x_5 \leq 6 ;$$
$$x_1 + x_4 - 5x_5 \geq 8 ;$$
$$x_j \geq 0, \quad j = \overline{1,5} .$$

Розв'язання. Щоб записати задачу у формі основної задачі (9.6) – (9.8), необхідно вихідну функцію цілі помножити на -1, а також замінити обмеження-нерівності типу (9.2), (9.4) обмеженнями-рівностями типу (9.3). Оскільки число нерівностей, що входять у систему обмежень задачі, дорівнює чотирьом, така заміна може бути здійснена введенням чотирьох додаткових невід'ємних змінних. При цьому лівим частинам кожного з нерівностей виду «$\leq$ » додається відповідна невід'ємна змінна, а з нерівностей виду «$\geq$ » – віднімається. У результаті введення додаткових змінних задача набуває канонічного вигляду:

$$y = 3x_1 - 2x_2 - 5x_4 + x_5 \to \min_{x_j\in\Omega\subset\mathbf{R}^9}$$

$$\mathbf{\Omega:} \quad 2x_1 + x_3 - x_4 + x_5 + x_6 = 2 ;$$
$$x_1 - x_3 + 2x_4 + x_5 + x_7 = 3 ;$$
$$2x_2 + x_3 - x_4 + 2x_5 + x_8 = 6 ;$$
$$x_1 + x_4 - 5x_5 - x_9 = 8 ;$$
$$x_j \geq 0, \quad j = \overline{1,9} .$$

Перетворимо отриману математичну модель канонічної задачі лінійного програмування у векторно-матричний вигляд:

$$y = \begin{bmatrix} 3 & -2 & 0 & -5 & 1 & 0 & 0 & 0 & 0 \end{bmatrix} \mathbf{x} \rightarrow \min_{\mathbf{x} \in \mathbf{\Omega} \subset \mathbf{R}^9}$$

$$\mathbf{\Omega}: \quad \begin{bmatrix} 2 & 0 & 1 & -1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & -1 & 2 & 1 & 0 & 1 & 0 & 0 \\ 0 & 2 & 1 & -1 & 2 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & -2 & 0 & 0 & 0 & 1 \end{bmatrix} \mathbf{x} + \begin{bmatrix} -2 \\ -3 \\ -6 \\ -8 \end{bmatrix} = 2 \ ;$$

$$\mathbf{x} \geq 0 \ .$$

Тут $\mathbf{x}^{\text{т}} = \begin{bmatrix} x_1 & x_2 & x_3 & x_4 & x_5 & x_6 & x_7 & x_8 & x_9 \end{bmatrix}$.

Приклад 9.5. Записати в канонічній формі задачу лінійного програмування:

$$y = -x_1 - x_2 \rightarrow \min_{x_{1,2} \in \mathbf{\Omega} \subset \mathbf{R}^2} \ ,$$

$$\mathbf{\Omega}: \quad x_1 + 2x_2 \geq \tfrac{1}{2} \ ;$$

$$x_1 \leq -\tfrac{1}{4} \ ;$$

$$-\tfrac{1}{2} \leq x_2 \leq \tfrac{1}{2} \ .$$

Розв'язання. Щоб дану задачу математичного програмування привести до канонічної форми, треба спочатку позбутися від'ємності змінної x_1. і двосторонньої обмеженості змінної x_2.

Якщо змінна обмежена тільки зверху $x_j \leq x_j^{++}, \quad j \in \left\{\overline{1,n}\right\}$, то для усунення від'ємності змінної досить замінити її двома невід'ємними змінними, прийнявши $x_j = x_{n+1} + x_{n+2}$, і відкоригувати функцію y й обмеження задачі з урахуванням зазначеної заміни. Так, заперечність змінної

x_1 в розглянутому прикладі усувається заміною $x_2 = x_3 + x_4$ з наступною корекцією математичної моделі:

$$y = -x_3 + x_4 - x_2 \rightarrow \min_{x_{2,3,4} \in \mathbf{\Omega} \subset \mathbf{R}^3},$$

$$\mathbf{\Omega}: \quad \begin{aligned} & x_3 - x_4 + 2x_2 \geq \tfrac{1}{2} ; \\ & x_3 - x_4 \leq -\tfrac{1}{4} ; \\ & -\tfrac{1}{2} \leq x_2 \leq \tfrac{1}{2} ; \qquad x_{3,4} \geq 0 . \end{aligned}$$

Двостороння обмеженість змінних (від'ємності або невід'ємна) усувається таким чином. Спочатку вихідне обмеження $x_j^{+} \leq x_j \leq x_j^{++}$ перетворюють до вигляду

$$0 \leq x_j - x_j^{+} \leq x_j^{++} - x_j^{+} \quad . \tag{9.12}$$

Потім, замінивши в (9.12) $x_j - x_j^{+}$ на нову змінну x_{n+1}, позбуваються від від'ємності змінної x_j. Якщо змінна входить до функції y або обмеження, то останні повинні бути відкориговані з урахуванням заміни: замість змінної x_j необхідно підставити вираз $x_{n=1} + x_j^{+}$. Для усунення двосторонньої обмеженості достатньо умови (9.12) подати у вигляді системи двох нерівностей:

$$\begin{gathered} x_{n+1} \leq x_j^{++} - x_j^{+}; \\ x_{n+1} \geq 0. \end{gathered}$$

Приведемо двосторонню обмеженість змінної x_2 в поточному прикладі до вигляду (9.12): $0 \leq x_2 - \frac{1}{2} \leq 1$. Введемо заміну $x_5 = x_2 + \frac{1}{2}$. Відкоригуємо математичну модель шляхом введення замість змінної x_2 її еквівалента $x_3 - \frac{1}{2}$:

$$y = -x_3 + x_4 - x_5 + \tfrac{1}{2} \to \min_{x_{3,4,5} \in \mathbf{\Omega} \subset \mathbf{R}^3},$$

$$\mathbf{\Omega}: \quad x_3 - x_4 + 2x_5 - 1 \geq \tfrac{1}{2};$$

$$x_3 - x_4 \leq -\tfrac{1}{4};$$

$$x_5 \leq 1;$$

$$x_{3,4,5} \geq 0.$$

Отриману математичну модель приведемо до канонічної форми, усунувши нерівності за раніше розглянутою методикою:

$$y = -x_3 + x_4 - x_5 + \tfrac{1}{2} \to \min_{x_{3,4,5,6,7,8} \in \mathbf{\Omega} \subset \mathbf{R}^6},$$

$$\mathbf{\Omega}: \quad x_3 - x_4 + 2x_5 - x_6 - \tfrac{3}{2} = 0;$$

$$x_3 - x_4 + x_7 + \tfrac{1}{4} = 0;$$

$$x_5 + x_8 - 1 = 0;$$

$$x_{3,4,5,6.7,8} \geq 0.$$

9.2. Графічне вирішення задачі лінійного програмування

Припустима множина рішень задачі лінійного програмування Ω утворює опуклий гіпермногогранник, на границі якого знаходиться оптимум функції цілі $y(\mathbf{x})$.

Якщо задача лінійного програмування, записана у формі стандартної (9.1), (9.2) і (9.5), містить не більше двох змінних, або задача, записана в канонічній формі (9.6) – (9.8), містить не більше $m+2$ змінних, де m – число лінійно незалежних обмежень-рівностей, то оптимальну граничну точку $\mathbf{x}^*$ із Ω легко знайти графічним способом. Якщо задача задана в канонічній формі, її треба попередньо перетворити в стандартну, тобто в задачу лінійного програмування такого вигляду:

$$y(\mathbf{x}) = \begin{bmatrix} c_1 & c_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \to \underset{\mathbf{x} \in \mathbf{\Omega} \subset \mathbf{R}^2}{\text{opt}} , \tag{9.13}$$

$$\mathbf{\Omega}: \begin{bmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \\ \cdots & \cdots \\ a_{m1} & a_{m2} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \\ \cdots \\ b_m \end{bmatrix} \le 0 ; \tag{9.14}$$

$$\mathbf{x} \ge 0 . \tag{9.15}$$

Через двовимірний простір припустимих рішень гіпермногогранник $\mathbf{\Omega}$ задачі (9.13) – (9.15) трансформується у многокутник. Сторони многокутника утворюються прямими $a_{i1}x_1 + a_{i2}x_2 + b_i = 0, \quad i = \overline{1, m}$; $x_1 = 0$ і $x_2 = 0$, відповідні рівняння для яких утворюються з початкової системи обмежень (9.14) – (9.15) шляхом заміни знаків нерівностей «$\le$ » і «$\ge$ » на знак точної рівності «=».

Оскільки оптимальне рішення є граничною точкою многокутника рішень, графічне вирішення задачі лінійного програмування полягає у графічній побудові многокутника рішень і відшукання на його границі точки $\mathbf{x}^*$, що оптимізує функцію цілі. Останнє досягається побудовою лінії рівня (еквіпотенціалі) $c_1x_1 + c_2x_2 = y_0$ (y_0 – константа), що проходить через многокутник рішень, і її переміщенням паралельно самої собі в напрямку вектора $\mathbf{c}^{\mathrm{T}} = \begin{bmatrix} c_1 & c_2 \end{bmatrix}$ доти, поки вона не досягне крайньої граничної точки торкання з многокутником. У досягнутій точці цільова функція має максимальне значення. При переміщенні лінії рівня у протилежному напрямку координати крайньої граничної точки торкання визначають точку мінімуму функції y. При цьому можливі різні випадки, зображені на рис.9.1 – 9.4. Область припустимих рішень на кожному рисунку виділена сірим фоном. На рис.9.1 показано випадок, коли функція приймає максимальне значення в єдиній точці, на рис.9.2 – у будь-якій точці на відрізку *AB*. На рис.9.3 зображено випадок, коли функція, що максимізується, не обмежена зверху на області припустимих рішень, а на рис.9.4 – випадок, коли система обмежень задачі несумісна.

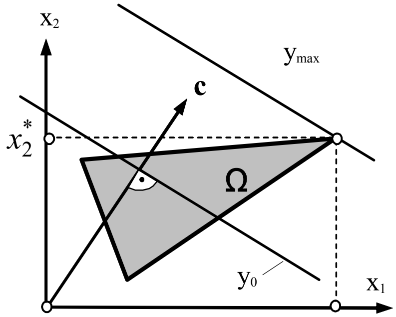


Рис.9.1

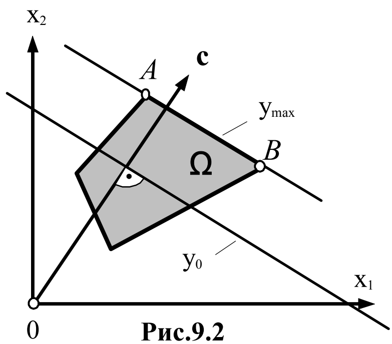


Рис.9.2

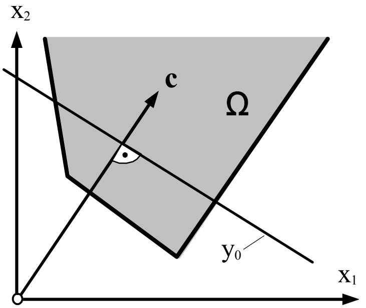


Рис.9.3

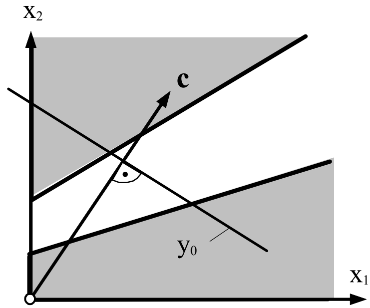


Рис.9.4

На рисунках розглянуті можливі випадки при пошуку максимуму цільової функції. Аналогічні ситуації можуть виникнути і при пошуку мінімуму.

Послідовність розв'язання задачі лінійного програмування графічним способом, як правило, включає такі етапи.

1. Приведення математичної моделі задачі до вигляду (9.13) – (9.15).

2. Побудова прямих, визначених рівняннями

$$a_{i1}x_1 + a_{i2}x_2 + b_i = 0, \quad i = \overline{1, m} \; ; \; x_1 = 0 \text{ і } x_2 = 0 .$$

3. Знаходження напівплощин, визначених кожним з обмежень задачі.

4. Виділення многокутника рішень (області припустимих рішень Ω).

5. Побудова прямої $y_0 = c_1 x_1 + c_2 x_2$, що проходить через многокутник рішень.

6. Побудова вектора $\mathbf{c}^{\mathrm{T}} = \begin{bmatrix} c_1 & c_2 \end{bmatrix}$.

7. Переміщення прямої $y_0 = c_1 x_1 + c_2 x_2$ в напрямку (у зворотному напрямку) вектора $\mathbf{c}$ до границі області Ω.

8. Визначення координат граничної точки й обчислення значення цільової функції в цій точці, що є максимальною (мінімальною).

Приклад 9.6. Для виробництва двох видів виробів *A* і *B* підприємство використовує три види сировини. Норми витрати сировини кожного виду наведені в табл.9.2. У ній же зазначено прибуток від реалізації одного виробу кожного виду і загальна кількість сировини, яка може бути використана підприємством.

Таблиця 9.2

Вид сировини	Норми витрати сировини на один виріб, *кг*		Загальна кількість сировини, *кг*
	A	*B*	
I	12	4	300
II	4	4	120
III	3	12	252
Прибуток від реалізації одного виробу, *гр.*	30	40	–

Вироби *A* і *B* можуть виготовлятися в будь-яких співвідношеннях (збут забезпечений). Треба скласти такий план їхнього випуску, при якому прибуток підприємства від реалізації усіх виробів буде максимальним.

Розв'язання. Припустимо, що буде виготовлено x_1 одиниць виробів виду *A*, x_2 – виду *B*. Оскільки виробництво продукції обмежено наявністю у розпорядженні підприємства сировиною кожного виду і кількість виготовлених виробів не може бути від'ємною, повинні виконуватися нерівності

$$12x_1 + 4x_2 \le 300 \;;$$

$$4x_1 + 4x_2 \le 120 \;;$$

$$3x_1 + 12x_2 \le 252 \;;$$

$$x_{1,2} \ge 0 \;.$$

Загальний прибуток від реалізації x_1 виробів виду *A* і x_2 – виду *B* складе $y = 30x_1 + 40x_2$.

Таким чином, приходимо до такої математичної задачі: серед усіх невід'ємних рішень складеної системи лінійних нерівностей треба знайти таке, при якому функція y набуває максимального значення.

Вирішимо сформульовану задачу лінійного програмування графічним способом. Оскільки задача подана в стандартній формі, відпадає необхідність у першому етапі вирішення, Почнемо розв'язання задачі з побудови многокутника рішень, для чого всі знаки нерівностей в системі обмежень замінимо точними рівностями і побудуємо відповідні прямі:

$$12x_1 + 4x_2 = 300 \;;$$

$$4x_1 + 4x_2 = 120 \;;$$

$$3x_1 + 12x_2 = 252 \;;$$

$$x_1 = 0 \;;$$

$$x_2 = 0 \;.$$

Ці прямі зображені на рис.9.5. Кожна з побудованих прямих поділяє площину на дві напівплощини. Координати точок однієї напівплощини задовольняють початковій нерівності, а іншою – ні. Щоб визначити шукану напівплощину, потрібно взяти точку, що належить однієї з площин, і перевірити, чи задовольняють її координати цій нерівності. Якщо координати точки задовольняють даній нерівності, шуканою є та напівплощина, якій належить ця точка, у протилежному разі – інша напівплощина.

Знайдемо, наприклад, напівплощину, визначену нерівністю $12x_1 + 4x_2 \le 300$. Для цього, побудувавши пряму $12x_1 + 4x_2 = 300$, візьмемо будь-яку точку, що належить однієї з двох отриманих напівплощин. Нехай це буде точка початку координат $\mathbf{x}_0^{\mathrm{T}} = \begin{bmatrix} 0 & 0 \end{bmatrix}$. Координати цій точці задовольняють нерівності $12 \cdot 0 + 4 \cdot 0 < 300$. Виходить, напівплощина, якій належить точка початку координат $\mathbf{x}_0$, відповідає нерівності $12x_1 + 4x_2 \le 300$.

Аналогічно робимо при пошуку інших напівплощин. Переріз знайдених напівплощин визначає опуклий многокутник припустимих рішень. На рисунку многокутник виділений сірим фоном. Будь-яка точка з області задовольняє системі обмежень задачі.

Поставлена задача буде вирішена, якщо ми зможемо знайти таку точку з області $\mathbf{\Omega}$, в якій функція y приймає максимальне значення. Щоб знайти зазначену точку, побудуємо пряму $y_0 = 30x_1 + 40x_2$ (нехай $y_0 = 600$) і вектор $\mathbf{c}^{\mathrm{T}} = \begin{bmatrix} 30 & 40 \end{bmatrix}$.

Якщо тепер взяти будь-яку точку, що належить побудованої прямої і многокутнику $\mathbf{\Omega}$, то її координати визначають такий план виготовлення виробів *А* і *В*, при якому прибуток від їхньої реалізації складе 600 *гр.* Далі, переміщуючи побудовану пряму в напрямку вектора $\mathbf{c}$ паралельно самої собі, будемо збільшувати праву частину рівняння прямої. Останньою спільною точкою прямої і многокутника є точка *М.* Координати цієї точки і визначають оптимальний план випуску виробів *А* і *В*, при якому прибуток від їхньої реалізації буде максимальним.

Знайдемо координати точки *М* як точки перетину двох прямих $4x_1 + 4x_2 = 120$, $3x_1 + 12x_2 = 252$. Шукані координати є рішенням системи цих двох рівнянь. Вирішивши систему, одержимо $x_1^* = 12$, $x_2^* = 18$. Отже, якщо підприємство виготовить 12 виробів виду *А* і 18 виробів виду *В*, воно одержить максимальний прибуток $y^* = 30 \cdot 12 + 40 \cdot 18 = 1080$ *грн.*

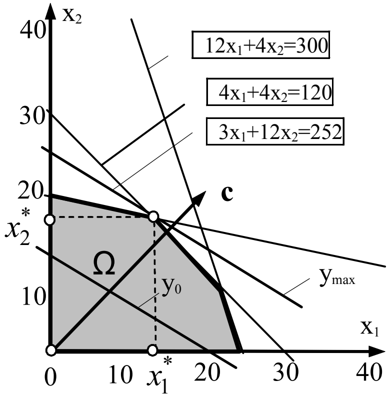


Рис.9.5

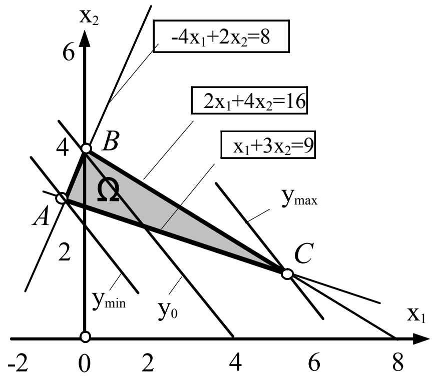


Рис.9.6

Приклад 9.7. Знайти максимум і мінімум функції $y = x_1 + x_2$ при умовах:

$$2x_1 + 4x_2 \le 16 \ ;$$

$$-4x_1 + 2x_2 \le 8 \ ;$$

$$x_1 + 3x_2 \ge 9 \ ;$$

$$x_{1,2} \ge 0 \ .$$

Розв'язання. При вирішенні задач лінійного програмування графічним способом необов'язково її математичну модель приводити до стандартної форми, головне, щоб вони не містили обмежень-рівностей. Так, в умовах цієї задачі немає необхідності в перетворенні математичної моделі, хоча вона містить елементи нестандартної задачі: обмеження-нерівність типу «$\ge$ » і одночасно оператори $\min$ і $\max$.

Побудуємо многокутник рішень. Для цього в нерівностях системи обмежень і умов невід'ємності змінних знаки нерівностей замінимо знаками точних рівностей:

$$2x_1 + 4x_2 = 16 \ ;$$

$$-4x_1 + 2x_2 = 8 \ ;$$

$$x_1 + 3x_2 = 9 \ ;$$

$$x_1 = 0 \ ; \quad x_2 = 0 \ .$$

Побудувавши отримані прямі, знайдемо відповідні напівплощини і їхній перетини – виділений сірим фоном трикутник *ABC* на рис.9.6. Координати точок цього трикутника задовольняють умові невід’ємності і нерівностям системи обмежень. Отже, задача буде вирішена, якщо серед точок трикутника *ABC* знайдемо такі, в яких функція $y = x_1 + x_2$ приймає максимальне і мінімальне значення. Для знаходження цих точок у побудові вектори $\mathbf{c}$ немає ніякої потреби. Досить побудувати пряму $y_0 = x_1 + x_2$ при довільному значенні y_0, яка проходить або не проходить через область припустимих рішень Ω, і переміщати її паралельно самій собі з метою виявлення найближчої і самої віддаленої спільних точок із многокутником рішень. Одна з цих точок відповідає мінімальному рішенню, інша – максимальному. На рис.9.6 цим точкам відповідають точки *C* і *A*.

Знайдемо координати цих точок із вирішення відповідних систем рівнянь:

$$\begin{cases} 2x_1 + 4x_2 = 16, \\ x_1 + 3x_2 = 9; \end{cases} \qquad \begin{cases} x_1 + 3x_2 = 9, \\ x_1 = 0. \end{cases}$$

Вирішивши по черзі кожну систему, знайдемо шукані рішення: $x^*_{1C} = 6$, $x^*_{2C} = 1$ і $x^*_{1A} = 0$, $x^*_{2A} = 3$. Після підстановки знайдених значень змінних у функцію цілі одержимо відповідно $y^*_C = 7$ і $y^*_A = 3$. Порівняння обчислених значень функцій говорить про те, що точка *C* – точка максимуму, а *A* – точка мінімуму.

9.3. Особливості задачі лінійного програмування

Диференціальний алгоритм розв'язання загальної задачі математичного програмування може бути з успіхом використаний для вирішення задач лінійного програмування. Це особливо важливо, оскільки не всяка задача лінійного програмування допускає геометричну інтепретацію. Отже, вирішення не всякої задачі можна отримати графічним методом. Диференціальний алгоритм у цьому змісті має універсальність.

Диференціальний алгоритм розв'язання задачі лінійного програмування логічно випливає з диференціального алгоритму загальної задачі математичного програмування при лінійних обмеженнях і, в остаточному підсумку, зводиться до широко відомого симплексного методу.

Диференціальний алгоритм вирішення загальної задачі математичного програмування при лінійних обмеженнях у постановці (8.6) – (8.8) розглядався раніше. Ближче за інші постановки до (8.6) – (8.8) наближається канонічна форма задання задачі лінійного програмування (9.6) – (9.8). Тому розгляд задачі лінійного програмування будемо розглядати в цій постановці. Для збереження логічної послідовності навчального посібника оператор $\max$ у канонічній моделі (9.6) – (9.8) замінимо оператором $\min$. Тоді канонічна форма запису задачі лінійного програмування набуватиме вигляду

$$y(\mathbf{x}) = \mathbf{c}^{\mathrm{T}}\mathbf{x} + c_0 \to \min_{\mathbf{x}\in\Omega\subset\mathbf{R}^n}, \tag{9.16}$$

$$\Omega: \quad \mathbf{A}\mathbf{x} + \mathbf{b} = 0; \tag{9.17}$$

$$\mathbf{x} \geq 0, \tag{9.18}$$

Дана модель відрізняється від (8.6) – (8.8) лінійністю функції $y(\mathbf{x})$ і односторонньою невід'ємною обмеженістю змінних. Такі відмінності додають задачі лінійного програмування (9.16) – (9.18) нові властивості і є причиною значного спрощення диференціального алгоритму її вирішення.

Однією з істотних властивостей задачі (9.16) – (9.18) є незалежність частинних похідних функції цілі та функцій обмежень від змінних задачі. Всі частинні похідні дорівнюють константам. Результатом цієї рівності є незалежність *умовних* похідних функції цілі від змінних рішень.

Необхідні умови для точки умовного локального мінімуму через односторонню обмеженість (невід'ємність) змінних трансформуються із системи трьох умов (8.12) – (8.14) у систему двох умов:

$$\text{якщо} \quad t_r^* = 0, \quad \text{то} \quad \left(\frac{\delta y}{\delta t_r}\right)^* \geq 0; \tag{9.19}$$

$$\text{якщо} \quad t_r^* > 0, \quad \text{то} \quad \left(\frac{\delta y}{\delta t_r}\right)^* = 0 \ . \tag{9.20}$$

Але якщо похідна лінійної функції $\left(\frac{\delta y}{\delta t_r}\right)^* = 0$ при деякому значенні t_r, більшому її нижньограничного значення, то внаслідок незалежності похідною від змінних вона дорівнюватиме нулю і при нижньограничному значенні, тобто з (9.20) випливає

$$\left(\frac{\delta y}{\delta t_r}\right)^* = 0 \quad \text{при} \quad t_r = 0. \tag{9.21}$$

Але умова (9.21) є окремим випадком умови (9.19). Таким чином, необхідні умови мінімуму в задачі лінійного програмування зводяться до виконання однієї умови (9.19), тобто

$$\left(\frac{\delta y}{\delta t_r}\right)^* \geq 0 \quad \text{при} \quad t_r^* = 0. \tag{9.22}$$

Ще одною важливою властивістю задачі лінійного програмування є те, що функція цілі може мати один мінімум (максимум) або ні одного. Це пояснюється опуклістю (угнутістю) області припустимих рішень **Ш** і лінійністю функції цілі. Завдяки цій властивості необхідні умови (9.22) одночасно є достатніми.

Отже, якщо для задачі лінійного програмування (9.16) – (9.18) в якийсь точці виконуються необхідні умови по всіх незалежних змінних $t_j \ \left(j = \overline{1, p}\right)$, то вона є глобальним мінімумом (максимумом).

З умови (9.9) випливає ще одна властивість задачі лінійного програмування: оптимальне рішення завжди є *опорним,* тобто таким вектором $\mathbf{x}^*$, в якого складові, відповідні поточним незалежним змінним, дорівнюють нулю.

> Опорне рішення задачі лінійного програмування завжди задовольняє системі обмежень (9.17).

Серед опорних рішень виділяють *припустимі опорні* рішення і *оптимальне опорне* рішення.

Припустимим опорним називають таке опорне рішення, в якому відсутні від'ємні складові. У припустимому опорному рішенні всі складові, що відповідають поточним незалежним змінним t_j $\left(j = \overline{1, p}\right)$, дорівнюють нулю, а інші складові, відповідні до залежних змінних s_i $\left(i = \overline{1, m}\right)$, не мають від'ємних значень.

> Припустиме опорне рішення завжди задовольняє системі рівностей (9.17) і умові невід'ємності (9.18).
>
> Оптимальне опорне рішення $\mathbf{x}^*$ – це таке припустиме опорне рішення, по якому задовольняється необхідна умова (9.22).

9.4. Диференціальний алгоритм вирішення задачі лінійного програмування

Задача лінійного програмування достатньо вивчена. У математичній літературі добре відомі різні методи її вирішення, наприклад, *симплексний метод, метод штучного базису, модифікований симплексний метод.* Вказані методи розглядаються відособлено – без їхньої прив'язки до загальної задачі математичного програмування. Менше в навчальній літературі можна зустріти вирішення задачі лінійного програмування за диференціальним алгоритмом. У відомих публікаціях [3,5] вирішення за диференціальним алгоритмом розпадається на три етапи: пошук опорного рішення, пошук припустимого опорного рішення і пошук оптимального рішення. Вибір довжини кроку за диференціальним алгоритмом на кожному етапі обчислюється за різними критеріями.

У цьому підрозділі розкривається одноетапний диференціальний алгоритм вирішення задачі лінійного програмування, розроблений автором. Даний метод викладався студентам економічних спеціальностей Харківської державної академії міського господарства протягом багатьох років і здобув широке визнання з методичної точки зору як простий метод, що логічно випливає з методу вирішення загальної задачі математичного програмування.

Єдиний етап, з якого складається метод, – це етап пошуку оптимального опорного рішення. Початкове рішення одержують простим прийомом з умов задачі.

Значну роль у вирішенні задачі лінійного програмування відіграє таблиця диференціального алгоритму, яка через особливості лінійного програмування відрізняється від аналогічної таблиці диференціального алгоритму загальної задачі математичного програмування при лінійних обмеженнях (табл. 8.2). Оскільки в задачі лінійного програмування відсутня верхня границя змінних, то відпадає необхідність у стовпці для значень s_i^{++} $\left(i=\overline{1,m}\right)$ і рядку для значень t_j^{++} $\left(j=\overline{1,p}\right)$. Крім того, немає необхідності в стовпі та в рядку для значень нижніх границь змінних s_i^{+}, t_j^{+}, оскільки вони (значення) у всіх змінних однакові і дорівнюють нулю. Немає також необхідності в рядку для поточних значень незалежних змінних $t_j^{(k)}$, оскільки для опорних рішень вони завжди дорівнюють нулю. Нарешті, немає необхідності в стовпці для поточних значень залежних змінних $s_i^{(k)}$, оскільки при рівності $\mathbf{t}^{(k)}=0$ вектор залежних змінних автоматично стає рівним вектору вільних членів $\mathbf{в}$.

Диференціальний алгоритм вирішення задачі математичного програмування припускає, що початкова точка наближення обов’язково належить області припустимих рішень Ω. Довільна розбивка змінних на залежні s_i $\left(i=\overline{1,m}\right)$ й незалежні t_i $\left(j=\overline{1,p}\right)$ не гарантує одержання припустимого опорного рішення. Однак нескладні перетворення початкової системи обмежень в канонічній задачі лінійного програмування дозволяють забезпечувати невід'ємність рішень на кожному кроці алгоритму, але при цьому порушується вимога опірності.

Перетворення полягають у наступному. Кожне обмеження-рівність, що містить негативний вільної член, інвертується, тобто помножається на -1. Крім того, вводяться штучні, або фіктивні, невід'ємні залежні змінні по одному в кожну праву частину обмежень-рівностей (9.17) замість нуля. Коса лінія в позначенні фіктивної змінної використовується для її відмінності від реальної залежної змінної. Всі початкові реальні змінні задачі на першому кроці вважаються незалежними, а тому рівними нулю.

Враховуючи зазначені перетворення канонічна задача лінійного програмування (9.16) – (9.18) набуває вигляду

$$y(\mathbf{t}) = \mathbf{c}^{\mathrm{T}}\mathbf{t} + c_0 \to \min_{\mathbf{t}\in\Omega\subset\mathbf{R}^{n+m}} , \tag{9.23}$$

$$\Omega : \quad \not{\mathbf{s}} = \mathbf{A}\mathbf{t} + \mathbf{b} \ ; \tag{9.24}$$

$$\not{\mathbf{s}}, \mathbf{t} \geq 0 , \tag{9.25}$$

Початкова таблиця диференціального алгоритму з урахуванням особливостей задачі лінійного програмування приймає вигляду табл.9.3. Коефіцієнти $a_{ij}^{(0)}$ в табл.9.3 рівні відповідним елементам матриці $\mathbf{A}$ в системі рівностей (9.24).

Прирівнюючи всі незалежні змінні нулю, одержуємо початкові значення штучних залежних змінних $s_i = b_i$, серед яких немає від'ємних через раніше проведені перетворення. Таким чином, одержуємо перше опорне рішення задачі (9.23) – (9.25) :

$$\mathbf{x}^{(0)\,\mathrm{T}} = \begin{bmatrix} \underbrace{0 \quad 0 \quad \cdots \quad 0}_{n} & b_1 & b_2 & \cdots & b_m \end{bmatrix} . \tag{9.26}$$

Рішення (9.26) є припустимим і опорним для задачі (9.23) – (9.25), але не припустимим – для задачі (9.16) – (9.18). Рішення (9.26) має n складових, рівних нулю, а опорне рішення задачі (9.16) – (9.18). повинно мати p таких складових. Тому диференціальний алгоритм вирішення задачі (9.23) – (9.25) буде, насамперед, спрямований на виключення штучних змінних $\$_i \left(i = \overline{1,m}\right)$, точніше, на переведення їх у систему незалежних змінних із наступним автоматичним виключенням із вирішення завдання. Процедура переведення змінних здійснюється, як і в загальній задачі математичного програмування, за допомогою вже добре знайомих жорданових виключень.

Таблиця 9.3

	t_1	…	t_j	…	t_n	
	$\left(\frac{\delta y}{\delta t_1}\right)^{(0)} = c_1$	…	$\left(\frac{\delta y}{\delta t_j}\right)^{(0)} = c_j$	…	$\left(\frac{\delta y}{\delta t_n}\right)^{(0)} = c_n$	$y^{(0)} = c_0$
$\$_1 =$	$a_{11}^{(0)}$	…	$a_{1j}^{(0)}$	…	$a_{1n}^{(0)}$	b_1
…	…	…	…	…	…	…
$\$_i =$	$a_{i1}^{(0)}$	…	$a_{ij}^{(0)}$	…	$a_{in}^{(0)}$	b_i
…	…	…	…	…	…	…
$\$_{m=}$	$a_{m1}^{(0)}$	…	$a_{mj}^{(0)}$	…	$a_{mn}^{(0)}$	b_m

Якщо серед коефіцієнтів c_j функції y є від'ємні складові, то це свідчить про порушення необхідних умов мінімуму (9.22). Для проведення на кожному кроці оптимізації процедури транспозиції змінних вибирають незалежну змінну, по якій порушена необхідна умова (9.22). Від'ємна похідна $\left(c_j \le 0\right)$ в другому рядку таблиці визначить ту незалежну змінну t_r, яка повинна брати участь у черговій транспозиції змінних. Оскільки всі незалежні змінні знаходяться на своїх нижніх границях $\left(t_j = 0\right)$, то приріст варійованої змінної на кожному кроці повинен бути тільки додатним:

$$\Delta t_r > 0 \,. \qquad (9.27)$$

Умова (9.27) визначає (8.48) як єдиний критерій вибору величини (довжини кроку) приросту варійованої змінної в задачі лінійного програмування. З урахуванням лінійності задачі, односторонньої обмеженості змінних і їхньої невід'ємності критерій спрощується і набуває вигляду

$$\Delta t_r^{(k)} = \min_{\frac{\delta s_i}{\delta t_r}<0} \left\{ \frac{s_i^+ - s_i^{(k)}}{\left(\frac{\delta s_i}{\delta t_r}\right)^{(k)}} \right\} = \min_{a_{ir}<0} \left\{ \frac{-b_i^{(k)}}{a_{ir}^{(k)}} \right\} = -\frac{b_k^{(k)}}{a_{kr}^{(k)}} \ . \tag{9.28}$$

Нижній індекс вільного члена $b_k^{(k)}$ вказує на залежну змінну s_k $\left(k \in \left\{\overline{1,m}\right\}\right)$, яка повинна разом із змінною t_r брати участь у черговій транспозиції змінних.

Транспозиція змінних здійснюється за допомогою звичайних жорданових виключень. При цьому елементи другого рядка також піддаються обчислювальним процедурам жорданових виключень на загальних підставах з іншими елементами таблиці. Останній елемент у цьому рядку показує значення функції $y^{(k)}$ в поточній точці наближення до мінімуму. Коли необхідні умови по всіх змінних будуть виконані, поточне значення функції $y^{(k)}$ відповідатиме шуканому мінімуму функції y^*.

Якщо в процесі вирішення задачі всі умовні похідні по незалежних змінних стануть одночасно більшими нуля, а серед залежних змінних буде знаходитися хоча б одна штучна змінна, то така задача не має жодного припустимого рішення.

Якщо в процесі рішення задачі в стовпці з від'ємною умовною похідною $\left(\frac{\delta y}{\delta t_r}\right)^{(k)}$ не буде жодного негативно елемента $a_{ir}^{(k)}$, то така задача також не має рішення, оскільки функція цілі необмежена знизу: її можна зменшувати до нескінченності.

Приклад 9.8. Знайти рішення задачі лінійного програмування:

$$y = -x_1 - x_2 \to \min_{x_{1,2} \in \Omega} \ ;$$

$$\Omega: \quad 3x_1 + 2x_2 \le 60 \ ;$$

$$-x_1 + x_2 \le 5 \ ;$$

$$x_1 + 4x_2 \ge 40 \ ;$$

$$x_{1,2} \ge 0 \ .$$

Розв’язання. Приведемо задачу до канонічної форми

$$y = -x_1 - x_2 \to \min_{x_{1,2} \in \Omega} \ ;$$

$$\Omega: \quad 3x_1 + 2x_2 + x_3 - 60 = 0 \ ;$$

$$-x_1 + x_2 + x_4 - 5 = 0 \ ;$$

$$x_1 + 4x_2 - x_5 - 40 = 0 \ ;$$

$$x_{1,2,3,4,5} \ge 0 \ .$$

Перетворимо обмеження задачі відповідно до вимог диференціального алгоритму про допустимість опорного рішення: всі вільні члени в обмеженнях-рівностях повинні бути додатними. У результаті перетворення одержимо канонічну задачу лінійного програмування у формі, зручній для укладання початкової таблиці диференціального алгоритму:

$$y = -x_1 - x_2 \to \min_{x_{1,2} \in \Omega} \ ;$$

$$\Omega: \quad -3x_1 - 2x_2 - x_3 + 60 = 0 \ ;$$

$$x_1 - x_2 - x_4 + 5 = 0 \ ;$$

$$-x_1 - 4x_2 + x_5 + 40 = 0 \ ;$$

$$x_{1,2,3,4,5} \ge 0 \ .$$

Сформуємо початкову таблицю диференціального алгоритму (табл.9.3).

Таблиця 9.3 – Початкова таблиця диференціального алгоритму

	x_1	x_2	x_3	x_4	x_5	<1>
$y =$	–1	–1	0	0	0	0
$\$_1 =$	–3	–2	–1	0	0	60
$\$_2 =$	1	–1	0	–1	0	5
$\$_3 =$	–1	–4	0	0	1	40

З аналізу таблиці випливає, що по незалежних змінних x_1 і x_2 порушені необхідні умови: відповідні умовні похідні $\left(\frac{\delta y}{\delta x_1}\right)^{(0)} = -1$ і $\left(\frac{\delta y}{\delta x_2}\right)^{(0)} = -1$ мають від'ємний знак. Треба на першому кроці оптимізації одну з цих змінних (однаково яку) перевести в систему залежних змінних. Нехай нею буде змінна x_1. Критерій вибору довжини кроку (9.28) вказує на те, що змінну x_1 слід ввести в систему залежних змінних замість штучної змінної $\$_1$:

$$\Delta t_r^{(0)} = \Delta x_1^{(0)} = \min_{a_{i1}<0}\left\{\frac{-b_i^{(0)}}{a_{i1}^{(0)}}\right\} == \min_{a_{i1}<0}\left\{\frac{-60}{-3}; \frac{-40}{-1}\right\} = \frac{-60}{-3} = -\frac{b_1^{(0)}}{a_{11}^{(0)}} \ .$$

Після кроку жорданових виключень, спрямованого на транспозицію змінних x_1 і $\$_1$, одержимо нову табл.9.4. У цій є присутнім стовпець (виділений сірим фоном) із незалежної штучною змінною $\$_1$, що виконала свою особисту роль у справі пошуку опорного рішення і в подальших обчисленнях тепер буде тільки “баластом”. Тому його (стовпець) можна вилучити з таблиці без якоїсь шкоди для подальшого вирішення задачі. Більш того, його можна було зовсім не включити до нової таблиці. Надалі стовпці з черговими штучними (вже незалежними) змінними в нову таблицю включати не будемо, щоб не виконувати зайвих обчислень їхніх елементів.

Таблиця 9.4

	$\$_1$	x_2	x_3	x_4	x_5	<1>
$y =$	1/3	–1/3	1/3	0	0	–20
$x_1 =$	–1/3	–2/3	–1/3	0	0	20
$\$_2 =$	1/3	–5/3	–1/3	–1	0	25
$\$_3 =$	–1/3	–10/3	1/3	0	1	20

З нової таблиці видно, що необхідні умови порушені по незалежній змінній x_2: тільки вона має від'ємну умовну похідну $\left(\frac{\delta y}{\delta x_2}\right)^{(1)} = -\frac{1}{3}$. Критерій вибору довжини кроку (9.28) вказує на штучну змінну $\$_3$, яку слід замінити на змінну x_2:

$$\Delta x_2^{(1)} = \min_{a_{i2}<0}\left\{\frac{-b_i^{(1)}}{a_{i2}^{(1)}}\right\} = \min_{a_{i2}<0}\left\{\frac{-20}{-2/3}; \frac{-25}{-5/3}; \frac{-20}{-10/3}\right\} = \frac{-20}{-10/3} = -\frac{b_3^{(1)}}{a_{32}^{(1)}} .$$

Другий крок жорданових виключень, що здійснює транспозицію змінних x_2 і $\$_3$, приводить до табл.9.5, в якої вже стовпці зі штучними незалежними змінними не присутні.

У новій таблиці функція цілі має від'ємну похідну $\left(\frac{\delta y}{\delta x_5}\right)^{(1)} = -\frac{1}{3}$ по незалежній змінній x_5. Критерій (9.28) визначає штучну змінну $\$_2$, яка на черговому кроці повинна поступитися своїм місцем додатковій змінній x_2 :

$$\Delta x_5^{(2)} = \min_{a_{i3}<0}\left\{\frac{-b_i^{(2)}}{a_{i3}^{(2)}}\right\} == \min_{a_{i3}<0}\left\{\frac{-16}{-1/5}; \frac{-15}{-1/2}\right\} = \frac{-15}{-1/2} = -\frac{b_2^{(2)}}{a_{23}^{(2)}} .$$

Після третього кроку жорданових виключень (див. табл.9.6) у задачі не залишилося ні однієї штучної змінної і по всіх незалежних змінних виконуються необхідні умови мінімуму. Шукане оптимальне опорне рішення $\mathbf{x}^{\mathrm{T}} = \begin{bmatrix} 10 & 15 & 0 & 0 & 30 \end{bmatrix}$ знайдено. У ньому, як цього вимагає властивість опорності задачі, рівно дві $\left(p = n - m = 5 - 3 = 2\right)$ незалежні змінні мають нульові значення. Мінімальне значення функції цілі, відповідно до результуючої табл.9.6, визначається як $y^{*} = y^{(3)} = -25$.

Таблиця 9.5

	x_3	x_4	x_5	<1>
$y =$	3/10	0	–1/10	–22
$x_1 =$	–2/5	0	–1/5	16
$\$_2 =$	–1/2	–1	–-1/2	15
$x_2 =$	1/10	0	3/10	6

Таблиця 9.6

	x_3	x_4	<1>
$y =$	2/5	1/5	–25
$x_1 =$	–1/5	2/5	10
$x_5 =$	–1	–2	30
$x_2 =$	–1/5	–3/5	15

Приклад 9.9. Знайти рішення задачі лінійного програмування:

$$y = -x_1 - x_2 \to \min_{x_{1,2} \in \Omega} \;;$$

$$\Omega: \quad 2x_1 + x_2 \le 2 \;;$$

$$3x_1 + 4x_2 \ge 12 \;;$$

$$x_{1,2} \ge 0 \;.$$

Розв'язання. Приведемо задачу до канонічної форми, зручної для формування таблиці диференціального алгоритму:

$$y = -x_1 - x_2 \to \min_{x_{1,2} \in \Omega} \;;$$

$$\Omega: \quad -2x_1 - x_2 - x_3 + 2 = 0 \;;$$

$$-3x_1 - 4x_2 + x_4 + 12 = 0 \;;$$

$$x_{1,2,3,4} \ge 0 \;.$$

Початкова таблиця диференціального алгоритму відповідає табл.9.7. Для спрощення процесу вирішення задачі штучні змінні не вводитимо, але їхня наявність передбачається.

Таблиця 9.7

	x_1	x_2	x_3	x_4	<1>
$y =$	–1	–1	0	0	0
$0 =$	–2	–1	–1	0	2
$\$_2 =$	–3	–4	0	1	12

Наступні дві таблиці відбивають кроки оптимізації за диференціальним алгоритмом. Результуюча таблиця диференціального алгоритму (табл.9.9) свідчить про відсутність рішення в задачі, оскільки вектор умовних похідних містить тільки позитивні складові, а умова опорності, відповідно до якої дві змінні задачі повинні дорівнювати нулю, не виконується.

Таблиця 9.8

	x_2	x_3	x_4	<1>
$y =$	–1/2	1	0	–1
$x_1 =$	–1/2	–1/2	0	1
$0 =$	–5/2	–3/2	1	15

Таблиця 9.9

	x_1	x_3	x_4	<1>
$y =$	1	3/2	0	–2
$x_2 =$	–2	1	0	2
$0 =$	5	4	1	4

Приклад 9.10. Знайти рішення задачі лінійного програмування:

$$y = -x_1 - x_2 \to \min_{x_{1,2} \in \Omega} \; ;$$

$$\Omega: \quad 3x_1 + x_2 \geq 3 \; ;$$

$$x_1 + 2x_2 \geq 2 \; ;$$

$$x_{1,2} \geq 0 \ .$$

Розв'язання. Приведемо задачу до канонічної форми, зручної для формування таблиці диференціального алгоритму:

$$y = -x_1 - x_2 \to \min_{x_{1,2} \in \Omega} \ ;$$

$$\Omega: \quad -3x_1 - x_2 + x_3 + 3 = 0 \ ;$$

$$-x_1 - 2x_2 + x_4 + 2 = 0 \ ;$$

$$x_{1,2,3,4} \geq 0 \ .$$

Початкова таблиця диференціального алгоритму відповідає табл.9.10.

Таблиця 9.10

	x_1	x_2	x_3	x_4	<1>
$y =$	–1	–1	0	0	0
$0 =$	–3	–1	1	0	3
$0 =$	–1	–2	0	1	2

Наступні три таблиці відбивають кроки оптимізації за диференціальним алгоритмом. Результуюча таблиця диференціального алгоритму (табл.9.13) свідчить про відсутність рішення в задачі, оскільки стовпець із від'ємною умовною похідною $\left(\frac{\delta y}{\delta x_4}\right)^{(1)} = -1$ не містить жодного від'ємного елемента $a_{i2}^{(2)}$.

Таблиця 9.11

	x_2	x_3	x_4	<1>
$y =$	–2/3	–1/3	0	–1
$x_1 =$	–1/3	1/3	0	1
$0 =$	–5/3	–1/3	1	1

Таблиця 9.12

	x_3	x_4	<1>
$y =$	–1/5	–2/5	–7/5
$x_1 =$	2/5	–1/5	4/5
$x_2 =$	–1/5	3/5	3/5

Таблиця 9.13

	x_2	x_4	<1>
$y =$	1	–1	–2
$x_1 =$	–2	1	2
$x_3 =$	-5	3	3

9.5. Контрольні запитання, приклади і вправи

1. Сформулювати загальну задачу лінійного програмування і навести її математичну модель

2. У чому полягають відмінності канонічної задачі лінійного програмування від загальної? Від стандартної?

3. Чому в задачі лінійного програмування число змінних повинно бути строго більше числа обмежень-рівностей?

4. Кондитерська фабрика для виробництва трьох видів карамелі *A, B, C* використовує три види основної сировини: цукровий пісок, патоку і фруктове пюре. Норми витрати сировини кожного виду на виробництво 1 *т* карамелі даного виду наведені в табл.9.14. У ній же зазначені загальна кількість сировини кожного виду, що використовується фабрикою, а також прибуток від реалізації 1 *т* карамелі кожного виду.

Таблиця 9.14

Вид сировини	Норми витрати сировини на 1 *т* карамелі, *т*			Загальна кількість сировини, *т*
	A	*B*	*C*	
Цукровий пісок	0,8	0,5	0,6	800
Патока	0,4	0,4	0,3	600
Фруктове пюре	–	0,1	0,1	120
Прибуток від реалізації 1 *т* продукції, *гр.*	108	112	126	–

За наведеними умовами скласти математичну модель задачі лінійного програмування, якщо функцією цілі виступає прибуток від реалізації продукції.

5. Скласти математичну модель задачі.

При відгодівлі кожна тварина щодня повинна одержувати не менше 60 одиниць живильної речовини *A*, не менше 50 одиниць речовини *B* і не менше 12 одиниць речовини *C*. Всі живильні речовини містяться в трьох видах корму. Кількість одиниць живильних речовин на 1 *кг* кожного з видів корму і їхні ціни наведені в табл.9.15.

Таблиця 9.15

Живильні речовини	Кількість одиниць живильної речовини на 1 *кг* корму		
	I	*II*	*III*
A	1	3	4
B	2	4	2
C	1	4	3
Ціна 1 *кг* корму, *гр.*	0,09	0,12	0,1

Скласти денний раціон, що забезпечує одержання необхідної кількості живильних речовин при мінімальних грошових витратах на придбання корму.

6. Привести задачу лінійного програмування

$$y = -2x_1 + x_2 + 5x_3 \to \max_{x_{1,2,3} \in \Omega},$$

$$\Omega: \quad 4x_1 + 2x_2 + 5x_4 \le 12 ;$$

$$6x_1 - 3x_2 + 4x_3 = 18 ;$$

$$3x_1 + 3x_2 - 2x_3 \ge 16 ;$$

$$x_{1,2,3} \ge 0$$

до канонічної форми і записати її у векторному-матричному вигляді.

7. Записати в канонічній формі задачу лінійного програмування

$$y = 2x_1 - 5x_2 - 3x_3 \to \max_{x_{1,2,3} \in \Omega},$$

$$\Omega: \quad -x_1 + x_2 + x_3 \ge 4 ;$$

$$2x_1 - x_2 + x_3 \le 16 ;$$

$$3x_1 + x_2 + x_3 \ge 18 ;$$

$$x_{1,2,3,} \ge 0.$$

8. Перетворити задачу

$$y = 2x_1 - x_2 + x_3 \to \max_{x_{1,2,3} \in \Omega} ,$$

$$\Omega: \quad 2x_1 - x_2 \le 3 ;$$
$$x_1 + x_2 - 0{,}5x_3 = 1 ;$$
$$x_1 + 2x_2 \ge 5 ;$$
$$x_{1,2,3} \ge 0$$

у стандартну задачу лінійного програмування.

9. Вирішити графічним методом задачу

$$y = x_1 + x_2 \to \max_{x_{1,2} \in \Omega} ,$$

$$\Omega: \quad x_1 + 2x_2 \le 14 ;$$
$$-5x_1 + 3x_2 \le 15 ;$$
$$4x_1 + 6x_2 \ge 24 ;$$
$$x_{1,2} \ge 0$$

10. Вирішити графічним методом задачу

$$y = x_1 + 2x_2 \to \max_{x_{1,2} \in \Omega} ,$$

$$\Omega: \quad 4x_1 - 2x_2 \le 12 ;$$
$$-x_1 + 3x_2 \le 6 ;$$
$$2x_1 + 4x_2 \ge 16 ;$$
$$x_{1,2} \ge 0$$

11. Вирішити графічним методом задачу

$$y = \begin{bmatrix} -2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \to \min_{\mathbf{x} \in \Omega} ,$$

$$\Omega: \quad \begin{bmatrix} 3 & -2 \\ -1 & 2 \\ -2 & -3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} -12 \\ -8 \\ -6 \end{bmatrix} \le 0 ;$$

$$\mathbf{x} \ge 0 .$$

12. Вирішити графічним методом задачу лінійного програмування, задану в канонічній формі:

$$y = -16x_1 - x_2 + x_3 + 5x_4 + 5x_5 \to \max_{x_j \in \Omega} ,$$

$$\Omega: \quad 2x_1 + x_2 + x_3 = 10 ;$$

$$-2x_1 + 3x_2 + x_4 = 6 ;$$

$$2x_1 + 4x_2 - x_5 = 8 ;$$

$$x_j \ge 0, \quad j = \overline{1,5} .$$

Розв’язання. Задача лінійного програмування в канонічній формі може бути вирішена графічним методом, якщо загальне число змінних не перевищує величини $m + 2$, де m – число обмежень-рівностей у математичній моделі задачі. Оскільки в даній задачі $m + 2 = 5 \quad (m = 3)$ і число змінних $n = 5$, вказана вимога до загального числа змінних виконується і задача може бути вирішена графічно.

Приведемо канонічну форму задачі до стандартної, для чого спочатку розв’яжемо систему обмежень-рівностей щодо будь-яких трьох змінних, наприклад x_3, x_4, x_5:

$$x_3 = -2x_1 - x_2 + 10 ;$$

$$x_4 = 2x_1 - 3x_2 + 6 ;$$

$$x_5 = 2x_1 + 4x_2 - 8 .$$

Після усунення залежних змінних в отриманій системі рівностей і підстановки у функцію цілі замість залежних змінних їхні вирази через незалежні x_1, x_2 одержимо математичну модель початкової задачі у стандартній формі:

$$y = 2x_1 + 3x_2 \to \max_{x_{1,2} \in \Omega'} ,$$

$$\Omega': \quad 2x_1 + x_2 \le 10 \ ;$$
$$-2x_1 + 3x_2 \le 6 \ ;$$
$$-2x_1 - 4x_2 \le -8 \ ;$$
$$x_{1,2} \ge 0 \ .$$

Дану задачу без труднощів можна вирішити графічним методом. Подальшу процедуру розв'язання задачі виконати самостійно.

13. Вирішити графічно методом задачу

$$y = \begin{bmatrix} -1 & 4 & 0 & 2 & -1 \end{bmatrix} \mathbf{x} \to \max_{\mathbf{x} \in \Omega \subset \mathbf{R}^5} ,$$

$$\Omega: \quad \begin{bmatrix} 1 & -5 & 1 & 0 & 0 \\ -1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 \end{bmatrix} \mathbf{x} + \begin{bmatrix} -5 \\ -4 \\ -8 \end{bmatrix} = 0 \ ;$$

$$\mathbf{x} \ge 0 \ .$$

14. Вирішити графічним методом задачу

$$y = 6{,}5x_1 - 7{,}5x_3 + 23{,}5x_4 - 5x_5 \to \max_{x_j \in \Omega} ,$$

$$\Omega: \quad x_1 + 3x_2 + x_3 + 4x_4 - x_5 = 12 \ ;$$
$$2x_1 - x_3 + 12x_4 - x_5 = 14 \ ;$$
$$x_1 + 2x_2 + 3x_4 - x_5 = 6 \ ;$$
$$x_j \ge 0, \quad j = \overline{1{,}5} \ .$$

15. Вирішити графічним методом таку задачу.

Для виробництва столів і шаф меблева фабрика використовує необхідні ресурси. Норми витрат ресурсів на один виріб даного виду, прибуток від реалізації одного виробу і загальна кількість наявних ресурсів кожного виду наведені в табл.9.16.

Таблиця 9.16

Вид сировини	Норми витрат ресурсів на один виріб		Загальна кількість ресурсів
	Стіл	Шафа	
Деревина виду *А*,, *м*^{2}	0,2	0,1	40
Деревина виду *В*, *м*^{2}	0,1	0,3	60
Трудомісткість, *чол-год.*	1,2	1,5	371,4
Прибуток від реалізації одного виробу, *грн.*	6	8	

Визначити, скільки столів і шаф фабриці треба виготовити, щоб прибуток від їхньої реалізації був максимальним.

16. Вирішити графічним методом таку задачу.

Для виробництва двох видів виробів *А* і *В* використовується токарне, фрезерне і шліфувальне устаткування. Норми витрат часу для кожного з типів устаткування на один виріб кожного виду наведені в табл.9.17. У ній же зазначено загальний фонд робочого часу кожного з типів устаткування, а також прибуток від реалізації одного виробу.

Таблиця 9.17

Тип устаткування	Витрати часу на обробку одного виробу, *верстато-год*		Загальний фонд робочого часу, *г*
	А	*В*	
Фрезерне	10	8	168
Токарне	5	10	180
Шліфувальне	6	12	144
Прибуток від реалізації одного виробу, *грн.*	14	18	

Знайти план випуску виробів А і В, що забезпечує максимальний прибуток.

17. Вирішити графічним методом таку задачу.

Меблева фабрика із стандартних листів фанери повинна вирізати заготівлі трьох видів – відповідно 24, 31, 18 *шт.* Кожний лист фанери може бути розрізаний на заготівлі двома способами. Кількість одержаних заготівель при даному способі розкрою наведена в табл.9.18. У ній же вказана площа відходів, що утворюються при даному способі розкрою одного листа фанери.

Таблиця 9.18

Вид заготівлі	Кількість заготівель при розкрою, *шт*	
	1-й спосіб	2-й спосіб
I	2	6
II	5	4
II	2	3
Відходи, *см*^{2}	12	16

Визначити, скільки листів фанери і яким способом треба розкроїти, щоб отримати не менше потрібної кількості заготівель при мінімальних загальних відходах.

18. Вирішити графічним методом таку задачу.

На звірофермі можуть вирощуватися чорно-білі лисиці і песці. Для забезпечення нормальних умов їхнього вирощування використовуються три види кормів. Кількість корму кожного виду, які повинні одержувати лисиці й песці, наведена у табл.9.19. У ній же зазначені загальна кількість корму кожного виду, що може бути використана звірофермою, і прибуток від реалізації однієї шкурки лисиці або песця.

Таблиця 9.19

Вид корму	Кількість одиниць корму на одну тварину		Загальна кількість корму
	Лисиця	Песець	
I	2	3	180
II	4	1	240
II	6	7	426
Прибуток від реалізації однієї шкурки, *грн.*	16	12	

Визначити, скільки лисиць і песців треба вирощувати на звірофермі, щоб прибуток від реалізації шкурок був максимальним.

19. У чому полягають необхідні умови для точки мінімуму в задачі лінійного програмування?

20. Яке рішення задачі лінійного програмування називають опорним? Припустимим опорним? Оптимальним?

21. У чому відмінність таблиці диференціального алгоритму для задачі лінійного програмування від таблиці для загальної задачі математичного програмування?

22. Яким способом повинна бути перетворена математична модель задачі лінійного програмування для її наступного вирішення за одноетапним диференціальним алгоритмом?

23. У чому полягає одноетапний диференціальний алгоритм вирішення задачі лінійного програмування?

24. При якій умові задача лінійного програмування, що розв'язується за одноетапним диференціальним алгоритмом, не має жодного припустимого рішення?

25. При якій умові цільова функція задачі лінійного програмування, що розв'язується за одноетапним диференціальним алгоритмом, може мати як завгодно мале значення (нескінченно зменшуватися)?

26. Записати у загальному вигляді критерій вибору приросту варійованої змінної в задачі лінійного програмування при її вирішенні за одноетапним диференціальним алгоритмом?

27, Як вибираються змінні для транспозиції в одноэтапном диференціальному алгоритмі вирішення задачі лінійного програмування?

28. Вирішити задачу лінійного програмування за допомогою одноетапного диференціального алгоритму:

$$y = 5x_1 + x_2 - x_3 - x_4 \to \min_{x_{1,2,3,4} \in \Omega} x \ ,$$

$$\Omega : \quad -3x_1 - 3x_2 - x_3 - x_4 + 3 = 0 \ ;$$

$$-2x_1 - 3x_3 + x_4 + 4 = 0 \ ;$$

$$\Omega: \begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 2 & 1 & 0 & 1 & 0 \\ 1 & 2 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} + \begin{bmatrix} -5 \\ -9 \\ -7 \end{bmatrix} = 0 \ ;$$

$$\mathbf{x} \geq 0 \ .$$

29. Для виготовлення різних виробів *A, B, C* підприємство використовує три різних види сировини. Норми витрати сировини на виробництво кожного виду, ціна одного виробу *A, B, C*, а також загальна кількість сировини кожного виду, що використовується підприємством, наведені в табл.9.20.

Таблиця 9.20

Вид сировини	Норми витрати сировини на один виріб, *кг*			Загальна кількість сировини, *кг*
	A	B	C	
I	18	15	12	360
II	6	4	8	192
III	5	3	3	180
Ціна одного виробу, *грн.*	9	10	16	

Вироби *A, B, C* можуть випускатися в будь-яких співвідношеннях (збут забезпечений), але виробництво обмежене виділеною підприємству сировиною кожного виду. Скласти план виробництва виробів, при якому загальна вартість усієї зробленої підприємством продукції є максимальною.

30. Перевірити вирішення попередньої задачі за допомогою графічного методу.

31. Знайти мінімум функції $y = -2x_1 + 6x_2 - 5x_5$ при умовах

$$-2x_1 + x_2 + x_3 + x_5 = 20 \ ;$$

$$-x_1 - 2x_2 + x_4 + 3x_5 = 24 \ ;$$

$$3x_1 - x_2 - 12x_5 + x_6 = 18 \ ;$$
$$x_j \geq 0, \quad j = \overline{1,6} \ .$$

32. Вирішити задачу лінійного програмування

$$y = \begin{bmatrix} -2 & -1 & 1 & -1 & 1 \end{bmatrix} \begin{bmatrix} x_1 & x_2 & x_3 & x_4 & x_5 \end{bmatrix}^{\text{т}} \to \min_{\mathbf{x} \in \Omega} \ ,$$

$$\Omega : \begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 2 & 1 & 0 & 1 & 0 \\ 1 & 2 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} + \begin{bmatrix} -5 \\ -9 \\ -7 \end{bmatrix} = 0 \ ;$$

$$\mathbf{x} \geq 0 \ .$$

Дати геометричну інтепретацію процесу вирішення.

33. Скільки різних опорних рішень має основна задача лінійного програмування при n змінних і m обмеженнях-рівностей? Перерахувати всі можливі опорні рішення в попередній задачі.

34. На швейній фабриці для виготовлення чотирьох видів виробів може бути використана тканина трьох артикулів. Норми витрату тканин всіх артикулів на пошиття одного виробу наведені в табл.9.21. У ній же зазначені загальні кількості тканин кожного артикулу, які є в розпорядженні фабрики, і ціни одного виробу кожного виду. Визначити, скільки виробів кожного виду повинна виробляти фабрика, щоб вартість виготовлення продукції була максимальною.

Таблиця 9.21

Артикул тканини	Норми витрати тканини на один виріб, *м*				Загальна кількість тканини, *м*
	1	2	3	4	
I	1	–	2	1	180
II	–	1	3	2	210
III	4	2	–	4	800
Ціна виробу, *грн.*	9	6	4	7	

35. Підприємство випускає чотири види продукції і використовує три типи основного устаткування: токарне, фрезерне і шліфувальне. Витрати часу на виготовлення одиниці продукції для кожного з типів устаткування наведені в табл.9.22. У ній же вказано загальний фонд робочого часу кожного з типів устаткування, а також прибуток від реалізації одного виробу даного виду. Визначити такий обсяг випуску кожного з виробів, при якому загальний прибуток від їхньої реалізації є максимальним.

Таблиця 9.22

Тип устаткування	Витрати часу, *верстато-год.*				Загальний фонд робочого часу, *верстато-год*
	1	2	3	4	
Токарне	2	1	1	3	300
Фрезерне	1	–	2	1	70
Шліфувальне	1	2	1	–	340
Прибуток від реалізації одиниці продукції, *грн.*	8	3	2	1	

36. Для перевезень вантажу на трьох лініях можуть бути використані судна трьох типів. Продуктивність суден при використанні їх на різних лініях характеризується даними, наведеними в табл.9.23. У ній же зазначені загальний час, протягом якого судна кожного типу знаходяться в експлуатації, і мінімально необхідні обсяги перевезень на кожній лінії. Визначити, які судна, на якій лінії і протягом якого часу слід використовувати, щоб забезпечити максимальне завантаження судів з урахуванням можливого часу їхньої експлуатації.

Таблиця 9.23

Тип судна	Продуктивність суден на лінії, *млн.тонно-миль на добу*			Загальний час експлуатації суден, *доба*
	1	2	3	
I	8	14	11	300
II	6	15	13	300
III	12	12	4	300
Об'єм перевезень, *млн.тонно-миль*	3000	5400	3300	

37. Визначити оптимальний план продуктивності кондитерської фабрики за умовою вправи 4.

38. Вирішити задачу лінійного програмування, подану у векторно-матричному вигляді:

$$y = \begin{bmatrix} 2 & -3 & 6 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 & x_2 & x_3 & x_4 & x_5 & x_6 \end{bmatrix}^{\text{т}} \to \min_{\mathbf{x}\in\Omega} ,$$

$$\Omega: \begin{bmatrix} 2 & 1 & -2 & 1 & 0 & 0 \\ 1 & 2 & 4 & 0 & 1 & 0 \\ 1 & -1 & 2 & 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_1 \\ x_1 \\ x_1 \\ x_1 \\ x_1 \end{bmatrix} + \begin{bmatrix} -24 \\ -22 \\ -10 \end{bmatrix} = 0 ;$$

$$\mathbf{x} \geq 0$$

39. Вирішити задачу лінійного програмування

$$y = 2x_1 - x_2 - x_4 \to \max_{x_j\in\Omega} ,$$

$$\Omega: \quad -2x_1 - x_2 - 2x_4 \geq 18 ;$$

$$x_1 - 2x_2 + x_3 = 10 ;$$

$$-3x_1 - 2x_2 - x_4 \leq -36 ;$$

$$x_j \geq 0, \quad j = \overline{1,4} .$$

40. Вирішити задачу лінійного програмування

$$y = x_1 - 4x_2 + 2x_3 + 26 \to \max_{x_j\in\Omega} ,$$

$$\Omega: \quad x_1 - 3x_2 - 3x_3 - x_4 - 2 = 0 ;$$

$$-x_1 + x_2 - x_3 + 6 \geq 0 ;$$

$$x_j \geq 0, \quad j = \overline{1,4} .$$

9.6. Фонд індивідуальних завдань

Індивідуальне завдання №30. За допомогою диференціального алгоритму знайти рішення задачі лінійного програмування:

1. $$y = 2x_1 + 2x_2 - 3x_3 + 1 \to \min_{x_{1,2,3} \in \Omega},$$

$$\Omega: \quad -2x_1 - x_2 - x_3 - x_4 + 4 = 0;$$
$$-x_1 + x_2 - x_3 - x_5 - 2 = 0;$$
$$x_{1,2,3,4,5} \geq 0$$

2. $$y = 2x_1 + x_2 - 3x_3 + 2 \to \min_{x_{1,2,3} \in \Omega},$$

$$\Omega: \quad -x_1 + x_2 - x_3 - x_4 - 2 = 0;$$
$$-2x_1 - x_2 - x_3 - x_5 + 4 = 0;$$
$$x_{1,2,3,4,5} \geq 0$$

3. $$y = x_1 + 2x_2 - 3x_3 + 3 \to \min_{x_{1,2,3} \in \Omega},$$

$$\Omega: \quad -x_1 - 2x_2 - x_3 - x_4 + 4 = 0;$$
$$x_1 - x_2 - x_3 - x_5 - 2 = 0;$$
$$x_{1,2,3,4,5} \geq 0$$

4. $$y = 2x_1 + 2x_2 - 3x_3 + 4 \to \min_{x_{1,2,3} \in \Omega},$$

$$\Omega: \quad x_1 - x_2 - x_3 - x_4 - 2 = 0;$$
$$-x_1 - 2x_2 - x_3 - x_5 + 4 = 0;$$
$$x_{1,2,3,4,5} \geq 0$$

5. $$y = 2x_1 - 3x_2 + x_3 + 5 \to \min_{x_{1,2,3} \in \Omega},$$

$$\Omega: \quad -2x_1 - x_2 - x_3 - x_4 + 8 = 0;$$
$$-x_1 - x_2 + x_3 - x_5 - 4 = 0;$$
$$x_{1,2,3,4,5} \geq 0$$

6. $$y = 2x_1 - 3x_2 + x_3 + 6 \to \min_{x_{1,2,3} \in \Omega},$$

$$\Omega: \quad -x_1 - x_2 + x_3 - x_4 - 4 = 0;$$
$$-2x_1 - x_2 - x_3 - x_5 + 8 = 0;$$
$$x_{1,2,3,4,5} \geq 0$$

7. $$y = -3x_1 + 2x_2 + x_3 + 7 \to \min_{x_{1,2,3} \in \Omega},$$

$$\Omega: \quad -x_1 - 2x_2 - x_3 - x_4 + 8 = 0;$$
$$x_1 - x_2 + x_3 - x_5 - 4 = 0;$$
$$x_{1,2,3,4,5} \geq 0$$

8. $$y = -3x_1 + 2x_2 + x_3 + 8 \to \min_{x_{1,2,3} \in \Omega},$$

$$\Omega: \quad -x_1 - x_2 + x_3 - x_4 - 4 = 0;$$
$$-x_1 - 2x_2 - x_3 - x_5 + 8 = 0;$$
$$x_{1,2,3,4,5} \geq 0$$

9. $$y = x_1 - 3x_2 + 2x_3 + 9 \to \min_{x_{1,2,3} \in \Omega},$$

$$\Omega: \quad -x_1 - x_2 - 2x_3 - x_4 + 2 = 0;$$
$$-x_1 - x_2 - x_3 - x_5 - 1 = 0;$$
$$x_{1,2,3,4,5} \geq 0$$

10. $$y = x_1 - 3x_2 + 2x_3 + 10 \to \min_{x_{1,2,3} \in \Omega},$$

$$\Omega: \quad x_1 - x_2 - x_3 - x_4 - 1 = 0;$$
$$-x_1 - x_2 - 2x_3 - x_5 + 2 = 0;$$
$$x_{1,2,3,4,5} \geq 0$$

11. $$y = -3x_1 + x_2 + 2x_3 + 11 \to \min_{x_{1,2,3} \in \Omega},$$

$$\Omega: \quad -x_1 - x_2 - 2x_3 - x_4 + 6 = 0;$$
$$-x_1 + x_2 - x_3 - x_5 - 3 = 0;$$
$$x_{1,2,3,4,5} \geq 0$$

12. $$y = -3x_1 + x_2 + 2x_3 + 12 \to \min_{x_{1,2,3} \in \Omega},$$

$$\Omega: \quad -x_1 + x_2 - x_3 - x_4 - 3 = 0;$$
$$-x_1 - x_2 - 2x_3 - x_5 + 6 = 0;$$
$$x_{1,2,3,4,5} \geq 0$$

13. $$y = 4x_1 + x_2 - 2x_3 + 13 \to \min_{x_{1,2,3} \in \Omega},$$

$$\Omega: \quad -2x_1 + x_2 - x_3 - x_4 - 1 = 0;$$
$$-2x_1 - x_2 - 3x_3 - x_5 + 5 = 0;$$
$$x_{1,2,3,4,5} \geq 0$$

14. $$y = 4x_1 + x_2 - 2x_3 + 14 \to \min_{x_{1,2,3} \in \Omega},$$

$$\Omega: \quad -2x_1 - x_2 - 3x_3 - x_4 + 5 = 0;$$
$$-2x_1 + x_2 - x_3 - x_5 - 1 = 0;$$
$$x_{1,2,3,4,5} \geq 0$$

15. $$y = x_1 + 4x_2 - 2x_3 + 15 \to \min_{x_{1,2,3} \in \Omega},$$

$$\Omega: \quad x_1 - 2x_2 - x_3 - x_4 - 2 = 0;$$
$$-x_1 - 2x_2 - 3x_3 - x_5 + 10 = 0;$$
$$x_{1,2,3,4,5} \geq 0$$

16. $$y = x_1 + 4x_2 - 2x_3 + 16 \to \min_{x_{1,2,3} \in \Omega},$$

$$\Omega: \quad -x_1 - 2x_2 - 3x_3 - x_4 + 10 = 0;$$
$$x_1 - 2x_2 - x_3 - x_5 - 2 = 0;$$
$$x_{1,2,3,4,5} \geq 0$$

17. $$y = -2x_1 + 4x_2 + x_3 + 17 \to \min_{x_{1,2,3} \in \Omega},$$

$$\Omega: \quad -x_1 - 2x_2 + x_3 - x_4 - 1 = 0;$$
$$-3x_1 - x_2 - x_3 - x_5 + 5 = 0;$$
$$x_{1,2,3,4,5} \geq 0$$

18. $$y = -2x_1 + 4x_2 + x_3 + 18 \to \min_{x_{1,2,3} \in \Omega},$$

$$\Omega: \quad -3x_1 - x_2 - x_3 - x_4 + 5 = 0;$$
$$-x_1 - 2x_2 + x_3 - x_5 - 1 = 0;$$
$$x_{1,2,3,4,5} \geq 0$$

19. $$y = 4x_1 - 2x_2 + x_3 + 19 \to \min_{x_{1,2,3} \in \Omega},$$

$$\Omega: \quad -2x_1 - x_2 + x_3 - x_4 - 2 = 0;$$
$$-x_1 - 3x_2 - x_3 - x_5 + 10 = 0;$$
$$x_{1,2,3,4,5} \geq 0$$

20. $$y = 4x_1 - 2x_2 + x_3 + 20 \to \min_{x_{1,2,3} \in \Omega},$$

$$\Omega: \quad -x_1 - 3x_2 - x_3 - x_4 + 10 = 0;$$
$$-2x_1 - x_2 + x_3 - x_5 - 2 = 0;$$
$$x_{1,2,3,4,5} \geq 0$$

21. $$y = x_1 - 2x_2 + 4x_3 + 21 \to \min_{x_{1,2,3} \in \Omega},$$

$$\Omega: \quad x_1 - x_2 - 2x_3 - x_4 - 3 = 0;$$
$$-x_1 - 3x_2 - x_3 - x_5 + 15 = 0;$$
$$x_{1,2,3,4,5} \ge 0$$

22. $$y = x_1 - 2x_2 + 4x_3 + 22 \to \min_{x_{1,2,3} \in \Omega},$$

$$\Omega: \quad -x_1 - 3x_2 - x_3 - x_4 + 15 = 0;$$
$$x_1 - x_2 - 2x_3 - x_5 - 3 = 0;$$
$$x_{1,2,3,4,5} \ge 0$$

23. $$y = -2x_1 + x_2 + 4x_3 + 23 \to \min_{x_{1,2,3} \in \Omega},$$

$$\Omega: \quad -x_1 + x_2 - 2x_3 - x_4 - 3 = 0;$$
$$-3x_1 - x_2 - x_3 - x_5 + 15 = 0;$$
$$x_{1,2,3,4,5} \ge 0$$

24. $$y = -2x_1 + x_2 + 4x_3 + 24 \to \min_{x_{1,2,3} \in \Omega},$$

$$\Omega: \quad -3x_1 - x_2 - x_3 - x_4 + 15 = 0;$$
$$-x_1 + x_2 - 2x_3 - x_5 - 3 = 0;$$
$$x_{1,2,3,4,5} \ge 0$$

25. $$y = x_1 - 4x_2 + 2x_3 + 25 \to \min_{x_{1,2,3} \in \Omega},$$

$$\Omega: \quad -x_1 + x_2 - x_3 - x_4 + 6 = 0;$$
$$x_1 - 3x_2 - 3x_3 - x_5 - 2 = 0;$$
$$x_{1,2,3,4,5} \ge 0$$

26. $$y = x_1 - 4x_2 + 2x_3 + 26 \to \min_{x_{1,2,3} \in \Omega},$$

$$\Omega: \quad x_1 - 3x_2 - 3x_3 - x_4 - 2 = 0;$$
$$-x_1 + x_2 - x_3 - x_5 + 6 = 0;$$
$$x_{1,2,3,4,5} \geq 0$$

27. $$y = 2x_1 + 2x_2 - 3x_3 + 27 \to \min_{x_{1,2,3} \in \Omega},$$

$$\Omega: \quad 2x_1 - x_2 - x_3 - x_4 + 4 = 0;$$
$$-3x_1 + x_2 - x_3 - x_5 - 3 = 0;$$
$$x_{1,2,3,4,5} \geq 0$$

28. $$y = 2x_1 + 2x_2 - 3x_3 + 28 \to \min_{x_{1,2,3} \in \Omega},$$

$$\Omega: \quad -x_1 + x_2 - x_3 - x_4 - 2 = 0;$$
$$-4x_1 - x_2 - x_3 - x_5 + 5 = 0;$$
$$x_{1,2,3,4,5} \geq 0$$

29. $$y = 4x_1 + 2x_2 - 3x_3 + 29 \to \min_{x_{1,2,3} \in \Omega},$$

$$\Omega: \quad -x_1 - 2x_2 - x_3 - x_4 + 4 = 0;$$
$$x_1 - x_2 - x_3 - x_5 - 2 = 0;$$
$$x_{1,2,3,4,5} \geq 0$$

30. $$y = 4x_1 + 2x_2 - 5x_3 + 30 \to \min_{x_{1,2,3} \in \Omega},$$

$$\Omega: \quad x_1 - 2x_2 - x_3 - x_4 - 1 = 0;$$
$$-3x_1 - 2x_2 - x_3 - x_5 + 5 = 0;$$
$$x_{1,2,3,4,5} \geq 0$$

Індивідуальне завдання №31. Виконати індивідуальне завдання №30 методом перебору всіх опорних рішень.

Індивідуальне завдання №32. За допомогою диференціального алгоритму знайти рішення задачі лінійного програмування:

1. $$y = 3x_1 + 2x_3 - 6x_5 \to \max_{x_{1,3,5} \in \Omega},$$
$$\Omega: \quad \begin{aligned} & -2x_1 + x_2 - 3x_3 + 6x_4 = 18; \\ & -3x_1 + 2x_3 - 2x_4 + x_5 = 24; \\ & x_1 + 3x_3 - 2x_4 \ge 36 \\ & x_{1,2,3,4,5} \ge 0. \end{aligned}$$

2. $$y = x_1 - 2x_2 + 5x_3 \to \max_{x_{1,2,3} \in \Omega},$$
$$\Omega: \quad \begin{aligned} & 2x_1 + 2x_2 + 4x_3 \le 18; \\ & 2x_1 + x_3 - 3x_4 \le 20; \\ & 5x_1 - 3x_2 + 6x_3 \ge 19 \\ & x_{1,2,3,4} \ge 0. \end{aligned}$$

3. $$y = 2x_1 + 3x_2 - x_4 \to \max_{x_{1,2,4} \in \Omega},$$
$$\Omega: \quad \begin{aligned} & 2x_1 - x_2 - 2x_4 \ge 16; \\ & 3x_1 + 2x_2 + x_3 - 3x_4 = 18; \\ & -x_1 + 3x_2 + 4x_4 \ge 24 \\ & x_{1,2,3,4} \ge 0. \end{aligned}$$

4. $$y = 3x_1 + 3x_2 - 4x_3 \to \max_{x_{1,2,3} \in \Omega},$$
$$\Omega: \quad \begin{aligned} & 2x_1 + x_2 - 3x_3 - 18 \le 0; \\ & 4x_1 - 5x_3 - 12 \le 0; \end{aligned}$$

$$3x_1 - 2x_2 + x_3 - 14 \ge 0$$
$$x_{1,2,3} \ge 0.$$

5. $$y = 8x_2 + 7x_4 + x_5 \to \max_{x_{2,4,5} \in \Omega},$$
$$\text{Щ}: \quad x_1 - 2x_2 - 3x_4 - 2x_5 - 12 = 0;$$
$$4x_2 + x_3 - 4x_4 - 3x_5 - 12 = 0;$$
$$5x_2 + 5x_4 + x_6 - 25 \ge 0$$
$$x_{1,2,3,4,5} \ge 0.$$

6. $$y = 6x_1 - x_2 + 3x_3 \to \max_{x_{1,2,3} \in \text{Щ}},$$
$$\Omega: \quad 3x_1 - 7x_2 + 5x_3 \le 15;$$
$$6x_1 + 5x_2 - 8x_3 \le 12;$$
$$2x_1 + 3x_2 - 4x_3 = 16;$$
$$x_{1,2,3} \ge 0.$$

7. $$y = x_1 + 3x_2 - 5x_4 \to \max_{x_{1,2,4} \in \Omega},$$
$$\Omega: \quad 2x_1 + 4x_2 + x_3 + 2x_4 - 28 = 0;$$
$$4x_1 - 2x_2 + 8x_4 + 32 \ge 0;$$
$$-3x_1 + 5x_2 - 3x_4 - 30 \ge 0$$
$$x_{1,2,3,4} \ge 0.$$

8. $$y = -2x_1 + 5x_2 - 4x_3 \to \min_{x_{1,2,3} \in \Omega},$$
$$\Omega: \quad 4x_1 + 2x_2 - 3x_3 \ge 9;$$
$$3x_1 - 2x_2 + 5x_3 \ge 8;$$
$$x_1 + 3x_2 + 4x_3 \ge 12;$$
$$x_{1,2,3} \ge 0.$$

9. $$y = -x_1 - 2x_2 + 5x_3 \to \min_{x_{1,2,3} \in \Omega},$$

$$\Omega: \quad \begin{aligned} -x_1 + 4x_2 - 2x_3 &\le 6; \\ x_1 + x_2 + 2x_3 &\ge 6; \\ 2x_1 - x_2 + 2x_3 &= 4; \\ x_{1,2,3} &\ge 0. \end{aligned}$$

10. $$y = -3x_1 + 4x_2 - 6x_3 \to \min_{x_{1,2,3} \in \Omega},$$

$$\Omega: \quad \begin{aligned} 2x_1 + 3x_2 - x_3 &\ge 8; \\ -3x_1 + 2x_2 - 2x_3 &= 10; \\ 5x_1 - 4x_2 + x_3 &\ge 7; \\ x_{1,2,3} &\ge 0. \end{aligned}$$

11. $$y = 3x_1 + 2x_4 - 5x_5 \to \max_{x_{1,4,5} \in \Omega}$$

$$\Omega: \quad \begin{aligned} 2x_1 - 3x_4 + 5x_5 - 34 &\ge 0; \\ 4x_1 + x_2 + 2x_4 - 4x_5 - 28 &= 0; \\ -3x_1 + x_3 - 3x_4 + 6x_5 - 24 &= 0; \\ x_{1,2,3,4,5} &\ge 0. \end{aligned}$$

12. $$y = 2x_1 + x_2 + 3x_3 \to \max_{x_{1,2,3} \in \Omega},$$

$$\Omega: \quad \begin{aligned} -x_1 + 3x_2 - 5x_3 - 12 &\le 0; \\ 2x_1 - x_2 + 4x_3 - 24 &\le 0; \\ 3x_1 + x_2 + x_3 - 18 &\le 0; \\ x_{1,2,3} &\ge 0. \end{aligned}$$

13. $$y = 8x_1 - 3x_2 + x_3 + 6x_4 - 5x_5 \to \max_{x_{1,2,3,4,5} \in \Omega}$$

$$\Omega: \quad 2x_1 + 4x_4 + x_5 + x_4 - 2x_5 \ge 28;$$

$$x_1 - 2x_2 + x_4 + x_5 = 31;$$
$$-x_1 + 3x_3 + 5x_3 + 4x_4 - 8x_5 = 118;$$
$$x_{1,2,3,4,5} \geq 0.$$

14. $$y = -4x_1 - x_2 + 4x_3 \to \min_{x_{1,2,3} \in \Omega},$$
$$\Omega: \quad 2x_1 - x_2 + 4x_3 - 12 \leq 0;$$
$$x_1 + 3x_2 - 2x_3 - 13 = 0;$$
$$2x_1 + 5x_2 - 6x_3 - 11 \leq 0;$$
$$x_{1,2,3} \geq 0.$$

15. $$y = 2x_1 - 3x_2 + 4x_3 + 5x_4 - x_5 + 8x_6 \to \max_{x_{1,2,3,4,5,6} \in \Omega}$$
$$\Omega: \quad x_1 + 5x_2 - 3x_3 - 4x_4 + 2x_5 + x_6 = 28;$$
$$2x_1 + 9x_2 - 5x_3 - 7x_4 + 4x_5 + 2x_6 = 118;$$
$$x_{1,2,3,4,5,6} \geq 0.$$

16. $$y = 12x_1 + 24x_2 + 18x_3 \to \min_{x_{1,2,3} \in \Omega},$$
$$\Omega: \quad -x_1 + 2x_2 + 3x_3 \geq 2;$$
$$3x_1 - x_2 + x_3 \geq 1;$$
$$-5x_1 + 4x_2 + x_3 \geq 3;$$
$$x_{1,2,3} \geq 0.$$

17. $$y = 3x_1 + 5x_2 - 3x_3 + x_4 + x_5 + 8x_6 \to \max_{x_{1,2,3,4,5,6} \in \Omega}$$
$$\Omega: \quad x_1 - 3x_2 + 4x_3 + 5x_4 - 6x_5 + x_6 = 60;$$
$$7x_1 - 17x_2 + 26x_3 + 31x_4 - 35x_5 + 6x_6 = 420;$$
$$x_{1,2,3,4,5,6} \geq 0.$$

18. $$y = 12x_1 + 13x_2 + 11x_3 \to \min_{x_{1,2,3} \in \Omega},$$

$$\Omega:\quad 2x_1 + x_2 + 2x_3 \geq 4;$$
$$-x_1 + 3x_2 + 5x_3 \geq 1;$$
$$4x_1 - 2x_2 - 6x_3 \geq -1;$$
$$x_{1,2,3} \geq 0.$$

19. $$y = -6x_1 - 8x_2 - x_3 \to \min_{x_{1,2,3} \in \Omega},$$
$$\Omega:\quad x_1 + x_2 + x_3 - 3 \leq 0;$$
$$x_1 + 2x_2 - 4 \leq 0;$$
$$x_{1,2,3} \geq 0.$$

20. $$y = 27x_1 + 10x_2 + 15x_3 + 28x_4 \to \max_{x_{1,2,3,4} \in \Omega},$$
$$\Omega:\quad 3x_1 + 2x_2 + x_3 + 2x_4 \leq 2;$$
$$3x_1 + x_2 + 3x_3 + 4x_4 \leq 5;$$
$$x_{1,2,3,4} \geq 0.$$

21. $$y = 3x_1 + 2x_5 - 5x_6 \to \max_{x_{1,5,6} \in \Omega},$$
$$\Omega:\quad 2x_1 + x_2 - 3x_5 + 5x_6 - 30 = 0;$$
$$4x_1 + x_3 + 2x_5 - 4x_6 - 28 = 0;$$
$$-3x_1 + x_4 - 3x_5 + 6x_6 - 24 = 0;$$
$$x_{1,2,3,4,5,6} \geq 0.$$

22. $$y = 5x_1 - x_2 + 3x_4 + x_5 - x_6 \to \max_{x_{1,2,4,5,6} \in \Omega},$$
$$\Omega:\quad 2x_1 - x_2 + 3x_4 + x_5 - x_6 = 36;$$
$$-x_1 + 2x_2 + x_3 + 2x_4 + 2x_6 = 20;$$
$$3x_2 - x_2 + 2x_5 - x_4 + 3x_5 + x_6 = 30;$$
$$x_{1,2,3,4,5,6} \geq 0.$$

23. $$y = -x_1 - 3x_2 + 5x_4 \to \min_{x_{1,2,4} \in \Omega},$$
$$\Omega: \quad 2x_1 + 4x_2 + x_3 + 2x_4 - 28 = 0;$$
$$2x_1 - x_2 - 4x_4 - 16 \le 0;$$
$$-3x_2 + 5x_3 - 3x_4 \le 0;$$
$$x_{1,2,3,4} \ge 0.$$

24. $$y = -3x_1 + 7x_2 + 4x_3 \to \max_{x_{1,2,3} \in \Omega},$$
$$\Omega: \quad -2x_1 - 3x_2 - 2x_3 \le 12;$$
$$-4x_1 - 4x_2 - 3x_3 \le 24;$$
$$5x_1 + 5x_3 + 3x_3 \le 15;$$
$$x_{1,2,3} \ge 0.$$

25. $$y = 3x_1 - 2x_2 + 6x_3 \to \min_{x_{1,2,3} \in \Omega},$$
$$\Omega: \quad 2x_1 - 3x_2 + x_3 \le 18;$$
$$-3x_1 + 2x_2 - 2x_3 \le 24;$$
$$x_1 + 3x_2 - 4x_3 \le 36;$$
$$x_{1,2,3} \ge 0.$$

26. $$y = 2x_1 + 2x_2 - x_4 \to \max_{x_{1,2,4} \in \Omega},$$
$$\Omega: \quad 2x_1 - x_2 - 2x_4 \ge 16;$$
$$3x_1 + x_2 + x_3 - 3x_4 = 18;$$
$$-x_1 + 3x_2 + 2x_4 \ge 24$$
$$x_{1,2,3,4} \ge 0.$$

27. $$y = -x_1 + 3x_2 - 4x_3 \to \max_{x_{1,2,3} \in \Omega},$$
$$\Omega: \quad 2x_1 + x_2 - 4x_3 - 18 \le 0;$$

$$2x_1 - 5x_3 - 12 \le 0;$$
$$3x_1 - 2x_2 + 3_3 - 14 \ge 0$$
$$x_{1,2,3} \ge 0.$$

28. $$y = x_1 + 2x_2 + 5x_3 \to \max_{x_{1,2,3} \in \Omega},$$
$$\Omega: \quad 6x_1 + 2x_2 + 4x_3 \le 18;$$
$$2x_1 + x_3 - 3x_4 \le 15;$$
$$5x_1 - 3x_2 + 6x_3 \ge 12$$
$$x_{1,2,3,4} \ge 0.$$

29. $$y = 2x_1 + 3x_2 + x_4 \to \max_{x_{1,2,4} \in \Omega},$$
$$\Omega: \quad 4x_1 - x_2 - 2x_4 \ge 16;$$
$$3x_1 + 2x_2 + x_3 - 3x_4 = 6;$$
$$-x_1 + 3x_2 + 4x_4 \ge 24$$
$$x_{1,2,3,4} \ge 0.$$

30. $$y = 3x_1 + 3x_2 - 4x_3 \to \max_{x_{1,2,3} \in \Omega},$$
$$\Omega: \quad 2x_1 + x_2 - 3x_3 - 6 \le 0;$$
$$4x_1 - 3x_3 - 12 \le 0;$$
$$3x_1 - 2x_2 + x_3 - 6 \ge 0$$
$$x_{1,2,3} \ge 0.$$

ВІДПОВІДІ

Розділ 1.

$\mathbf{K}^{\mathrm{T}} = \begin{bmatrix} 3 & 0 \\ 7 & 8 \\ -2 & 1 \end{bmatrix}$; $\mathbf{L}^{\mathrm{T}} = \begin{bmatrix} 7 & 9 \\ 8 & 1 \end{bmatrix}$; $\mathbf{M}^{\mathrm{T}} = \begin{bmatrix} 12 \\ 0 \\ 5 \end{bmatrix}$; $\mathbf{N}^{\mathrm{т}} = \begin{bmatrix} 3 & 4 \end{bmatrix}$; $\mathbf{E}^{\mathrm{т}} = \mathbf{E} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$. **1.8.**

$\mathbf{C} + \mathbf{D} + \mathbf{F} = \begin{bmatrix} 0 & -2 & 1 \\ 1 & 3 & 3 \end{bmatrix}$. **1.9.** $\mathbf{C} - \mathbf{D} = \begin{bmatrix} 6 & 0 & 6 \\ -2 & -2 & 1 \end{bmatrix}$. **1.10.**

$\mathbf{T} + \mathbf{S} = \begin{bmatrix} 0 & 1 \\ 3 & -1 \end{bmatrix}$. **1.15.** Так. **1.16.** Ні. **1.18.** $\mathbf{A} \times \mathbf{B} = \begin{bmatrix} -6 \\ -5 \end{bmatrix}$. **1.19.**

$\mathbf{A} \times \mathbf{B} \times \mathbf{C} = \begin{bmatrix} -3 & -4 \\ -9 & 0 \end{bmatrix}$. **1.20**. Ні. **1.21.** $(\mathbf{AB})^{\mathrm{т}} = \mathbf{B}^{\mathrm{т}} \mathbf{A}^{\mathrm{т}} = \begin{bmatrix} 5 & 7 \\ 4 & 1 \\ 1 & 3 \end{bmatrix}$;

$(\mathbf{AC})^{\mathrm{т}} = \mathbf{C}^{\mathrm{т}} \mathbf{A}^{\mathrm{т}} = \begin{bmatrix} 4 & 4 \end{bmatrix}$. **1.22.** $\mathbf{B} = \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & 1 \end{bmatrix}$.

1.23. $\mathbf{B} = \begin{bmatrix} b_{11} & \dfrac{b_{22} - b_{11}}{2} \\ 0 & b_{22} \end{bmatrix}$. **1.24.** $(\mathbf{A})^2 = \begin{bmatrix} 10 & 2 & 3 \\ -1 & -1 & -1 \\ 3 & 2 & -1 \end{bmatrix}$;

$(\mathbf{A})^3 = \begin{bmatrix} 33 & 8 & 8 \\ -4 & -3 & 0 \\ 8 & 0 & 1 \end{bmatrix}$. **1.26.** $\mathbf{J}^2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$.

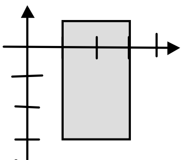
1.26. $(\mathbf{V})^2 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$; $(\mathbf{V})^3 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$. **1.33.** Ні. **1.35.** Так.

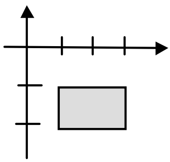
1.38. $y(\mathbf{x}) = [x_1 \quad x_2]\begin{bmatrix} -6 \\ -5 \end{bmatrix} + 3.$

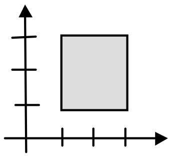
1.39. $y(\mathbf{x}) = \frac{1}{2}[x_1 \quad x_2]\begin{bmatrix} 2 & -1 \\ -1 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + [0 \quad 4]\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} - 3\,.$

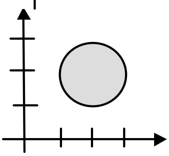
1.40. $y(\mathbf{x}) = \; = \frac{1}{2}[x_1 \quad x_2 \quad x_3]\begin{bmatrix} 0 & 0 & 4 \\ 0 & 4 & -1 \\ 4 & -1 & -2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + [1 \quad 0 \quad -2]\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + 7\,\cdot$

Розділ 2.

2.14. 

2.15. 

2.16. 

2.18. 

2.19. $a)\ \|\overline{x} - \overline{a}\| \le \varepsilon\,,\quad б)\ \|\overline{x} - \overline{a}\| \le \varepsilon\,.$

2.21. $\overline{b} \notin \mathbf{Q}\,,\quad \overline{c} \in \mathbf{Q}\,,\quad \overline{d} \in \mathbf{Q}\,.$ **2.22.** $\overline{z} \notin \mathbf{S}\,.$

Розділ 3.

3.1. $\bar{s} = \mathbf{A}\bar{t}$ **.** **3.2.** $\mathbf{A}\,\bar{x} = \bar{b}$ або $\mathbf{A}\,\bar{x} + \bar{b} = 0$. **3.3.** $\bar{x} = \mathbf{A}^{-1}\,\bar{b}$.

3.4. $\bar{s} = -\mathbf{A}_s^{-1}\mathbf{A}_t\bar{t} + \mathbf{A}_s^{-1}\bar{b}$.

3.10. $\begin{bmatrix} x_5 \\ x_2 \\ x_6 \end{bmatrix} = \begin{bmatrix} 2 & -1 & 0 & 1 \\ -2{,}5 & 1 & 0{,}5 & -2{,}5 \\ 0{,}5 & 0 & 0{,}5 & -0{,}5 \end{bmatrix} \begin{bmatrix} x_4 \\ x_1 \\ x_3 \\ x_7 \end{bmatrix}.$

3.11. $\begin{bmatrix} x_5 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 & 2 & 1 \\ -1 & -2 & 1 \end{bmatrix} \begin{bmatrix} x_2 \\ x_4 \\ x_1 \end{bmatrix}.$ **3.12.** Рішення не існує.

3.15. $\begin{bmatrix} x_1 \\ x_2 \\ x_4 \end{bmatrix} = \begin{bmatrix} \frac{1}{3} & -\frac{2}{3} \\ 2 & 1 \\ \frac{5}{3} & \frac{2}{3} \end{bmatrix} \begin{bmatrix} x_3 \\ 1 \end{bmatrix}.$ **3.16.** $\begin{bmatrix} x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -1 & -1 & 2 \\ \frac{1}{2} & 1 & -\frac{3}{2} \end{bmatrix} \begin{bmatrix} x_1 \\ x_4 \\ 1 \end{bmatrix}.$

3.17. $\begin{cases} x_1 = \frac{2}{3}x_2 - 1; \\ x_3 = \frac{1}{2}x_2 - 1; \\ x_4 = \frac{5}{6}x_2 - 1. \end{cases}$ **3.18.** $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \frac{6}{5} \\ \frac{1}{5} \end{bmatrix}.$ **3.23.** $\mathbf{A}^{-1} = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & -\frac{2}{3} \end{bmatrix}$ **.**

3.24. $\mathbf{A}^{-1} = \begin{bmatrix} \frac{1}{5} & -\frac{2}{5} & \frac{4}{5} \\ -\frac{1}{5} & \frac{2}{5} & \frac{1}{5} \\ \frac{3}{5} & -\frac{1}{5} & -\frac{3}{5} \end{bmatrix}$ **.** **3.26.** Див. 3.18.

Розділ 4.

4.3. Додатно визначена. **4.4.** Не визначена. **5.13.** Від'ємно визначена.

4.22. $y(\mathbf{x} + \Delta\mathbf{x}) = 7 + 4\Delta x_1 + 2\Delta x_2 + 6\Delta x_1^2 + \Delta x_2^2 - \Delta x_1\Delta x_2 + 2\Delta x_1^3$.

4.33. $\mathbf{x}^{*\mathrm{T}} = \mathbf{x}^{**\mathrm{T}} = [-4 \quad 7]; \quad y^* = y^{**} = 107.$

4.34. $\mathbf{x}^{*\mathrm{T}} = \mathbf{x}^{**\mathrm{T}} = [1 \quad 1]; \quad y^* = y^{**} = 5.$

4.35. $\mathbf{x}_{\mathrm{A}}^{*\mathrm{T}} = [1 \quad 1]; \mathbf{x}_{\mathrm{B}}^{*\mathrm{T}} = [-1 \quad -1]; \quad y^* = 4.$

Розділ 5.

5.1. Початкова точка наближення та точність обчислення.

5.4. $\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \lambda^{(k)} \Delta \mathbf{x}^{(k)}$. **5.5.** Напрямок руху.

5.6. Ні. **5.7.** $\left| \left(\dfrac{\partial y}{\partial x_i} \right)^{(k)} \right| < \varepsilon, \quad i = \overline{1, n}.$

5.11. $\mathbf{x}^{*\mathrm{T}} = \mathbf{x}^{(4)\mathrm{T}} = [2{,}0156 \quad 2{,}0114]; \quad y^* = -7{,}9988.$

5.12. $\mathbf{x}^{*\mathrm{T}} = \mathbf{x}^{(2)\mathrm{T}} = [0{,}2101 \quad 0{,}64964]; \quad y^* = 0{,}0149.$

5.13. $\mathbf{x}^{*\mathrm{T}} = \mathbf{x}^{(2)\mathrm{T}} = [1{,}94 \quad 0{,}812]; \quad y^* = -6{,}431.$

5.15. $\mathbf{x}^{*\mathrm{T}} = \mathbf{x}^{(2)\mathrm{T}} = [2{,}0011 \quad 2{,}0010]; \quad y^* = -7{,}999993.$

5.16. $\mathbf{x}^{*\mathrm{T}} = \mathbf{x}^{(2)\mathrm{T}} = [0{,}9521 \quad 0{,}4944]; \quad y^* = -3{,}0064.$

5.17. $\mathbf{x}^{*\mathrm{T}} = \mathbf{x}^{(3)\mathrm{T}} = [1{,}9298 \quad 0{,}7998]; \quad y^* = -6{,}4314.$

5.19. $\mathbf{x}^{*\mathrm{T}} = \mathbf{x}^{(5)\mathrm{T}} = [0{,}99999 \quad 1{,}0000]; \quad y^* = -6{,}0000.$

5.31. $\mathbf{x}^{*\mathrm{T}} = \mathbf{x}^{(8)\mathrm{T}} = [2{,}0032 \quad 2{,}0016]; \quad y^* = -8{,}0000.$

5.32. $\mathbf{x}^{*\mathrm{T}} = \mathbf{x}^{(2)\mathrm{T}} = [0{,}9521 \quad 0{,}5115]; \quad y^* = -3{,}0018.$

5.33. $\mathbf{x}^{*\mathrm{T}} = \mathbf{x}^{(8)\mathrm{T}} = [0{,}9231 \quad 0{,}7966]; \quad y^* = -6{,}431.$

5.34. $y^* = y^{(16)} = 14646{,}96\ \textit{грн.}$ при $x_1^* = 9{,}4217\textit{м}$; $x_2^* = 9{,}4218\textit{м}$.

Розділ 6.

6.5. Ні. **6.10.** Ні. **6.11.** Так. **6.13.** $\mathbf{x}^{*\mathrm{T}} = \begin{bmatrix} 2{,}2915 & 0{,}4364 \end{bmatrix}$; $y^* = -283{,}2897$. **6.14.** $\mathbf{x}^{*\mathrm{T}} = \begin{bmatrix} -\frac{1}{2} & 4 & -2 \end{bmatrix}$; $y^* = -1$.

6.15. $\mathbf{x}^{*\mathrm{T}} = \begin{bmatrix} 2 & 2 \end{bmatrix}$; $y^* = 12$. **6.16.** $(m \times m)$. **6.17.** $(m \times p)$.

6.19. $\dfrac{\delta y}{\delta \mathbf{t}} = \dfrac{\partial y}{\partial \mathbf{t}} - \mathbf{C}^{\mathrm{T}} \left(\mathbf{W}^{-1} \right)^{\mathrm{T}} \dfrac{\partial y}{\partial \mathbf{s}}$. **6.21.** $\mathbf{B} = \left[\dfrac{\delta s_i}{\delta t_j} \right] = -\mathbf{W}^{-1}\mathbf{C}$.

6.22. $\left(\dfrac{\delta\ y}{\delta\ \mathbf{t}} \right)^* = 0$. **6.27.** $\mathbf{x}^{\circ\mathrm{T}} = \begin{bmatrix} \frac{1}{2} & 4 & 4 \end{bmatrix}$. **6.28.** $\mathbf{x}^{\circ\mathrm{T}} = \begin{bmatrix} 2 & 2 \end{bmatrix}$.

6.29. $L(\mathbf{x}, \boldsymbol{л}) = y(\mathbf{x}) - \boldsymbol{л}^{\mathrm{T}} \mathbf{f}(\mathbf{x})$. **6.32.** $\mathbf{x}^{\circ\mathrm{T}} = \begin{bmatrix} \frac{1}{2} & 4 & 4 \end{bmatrix}$.

6.33. $\mathbf{x}^{\circ\mathrm{T}} = \begin{bmatrix} -\frac{1}{2} & 4 & -2 \end{bmatrix}$. **6.34.** $\mathbf{x}^{\circ\mathrm{T}} = \begin{bmatrix} 2 & 2 \end{bmatrix}$.

6.35. $\mathbf{x}^{\circ\mathrm{T}} = \begin{bmatrix} 2{,}2915 & 0{,}4364 & -15{,}7514 \end{bmatrix}$. **6.37.** $(p \times p)$.

6.39. $\mathbf{x}^{(1)\mathrm{T}} = \mathbf{x}^{*\mathrm{T}} = \begin{bmatrix} 1{,}9113 & -1{,}5980 & -1{,}2656 & 1{,}8173 \end{bmatrix}$.

6.40. $\mathbf{x}^{(2)\mathrm{T}} = \begin{bmatrix} -1{,}6231 & 1{,}6641 & 0{,}5796 & -1{,}2389 \end{bmatrix}$. Процес мінімізації слід продовжити.

6.41. $\mathbf{x}^{*\mathrm{T}} = \mathbf{x}^{(5)\mathrm{T}} = \begin{bmatrix} 1{,}0077 & -0{,}7684 & -0{,}9753 & 1{,}1521 \end{bmatrix}$; $y^* = -0{,}0953$. **6.43.** $\mathbf{x}^{(1)\mathrm{T}} = \begin{bmatrix} 1{,}0077 & -0{,}7684 & -0{,}9753 & 1{,}1521 \end{bmatrix}$; $y^{(1)} = -0{,}0967$.

6.44. $\mathbf{x}^{*\mathrm{T}} = \mathbf{x}^{(8)\mathrm{T}} = \begin{bmatrix} 1{,}0450 & -0{,}8139 & -0{,}9895 & 1{,}1780 \end{bmatrix}$;

$y^{(8)} = -0{,}0991$.

Розділ 7.

7.4,а. $\Delta x_r > 0$. **7.4,б.** $\Delta x_r < 0$. **7.8.** $\left(\frac{\partial y}{\partial \mathbf{x}}\right)^{(k)} \neq 0$.

7.12. $\mathbf{x}^{*\mathrm{T}} = \mathbf{x}^{(13)\mathrm{T}} = [0{,}0189 \quad 0{,}2368 \quad -0{,}0140]$; $y^* = 0{,}0018$.

7.13. $\mathbf{x}^{*\mathrm{T}} = [0{,}5831 \quad 5]$; $y^* = 4{,}4061$.

7.14. $\mathbf{x}^{*\mathrm{T}} = [-5 \quad 5]$; $y^* = -635$.

7.15. $\mathbf{x}^{*\mathrm{T}} = \mathbf{x}^{(5)\mathrm{T}} = [1{,}0755 \quad 1{,}5]$; $y^* = 71{,}4263$.

7.16. $\mathbf{x}^{*\mathrm{T}} = \mathbf{x}^{(15)\mathrm{T}} = [1{,}3896 \quad 1{,}0704]$; $y^* = 73{,}8335$.

7.21. $\left(\frac{\partial y}{\partial x_r}\right)^{(k)} > 0, \quad x_r^{(k)} \neq x_r^{+}$. **7.22.** $\left(\frac{\partial y}{\partial x_r}\right)^{(k)} < 0, \quad x_r^{(k)} \neq x_r^{++}$.

7.23. $\left(\frac{\partial y}{\partial x_r}\right)^{(k)} < 0, \quad x_r^{(k)} \neq x_r^{+}$. **7.24.** $\left(\frac{\partial y}{\partial x_r}\right)^{(k)} > 0, \quad x_r^{(k)} \neq x_r^{++}$.

Розділ 8.

8.2.

$$y(\mathbf{x}) = \frac{x_1}{x_2} - \frac{x_3 x_4}{x_5} \to \min_{\mathbf{x}\in\Omega},$$

$$\begin{aligned} \Omega: \quad & x_1 + x_2 + x_3 + x_4 + x_5 + x_6 - 100 = 0; \\ & x_1^2 + x_2^2 + x_3^2 + x_4^2 + x_5^2 - x_7 - 100 = 0; \\ & x_1 x_2 x_3 x_4 x_5 = 100; \\ & -10 \le x_i \le 10, \quad i = \overline{1,3}; \quad x_{5,6,7} \ge 0. \end{aligned}$$

8.16. $\mathbf{x}^* = \mathbf{x}^{(5)} = \begin{bmatrix}1{,}1 & 0 & 0 & 0{,}23 & 0 & 20 & 0{,}49 & 20\end{bmatrix}$; $\quad y^* = 49{,}8$.

8.19. Один або жодного. **8.20.** $\mathbf{x}^* = \mathbf{x}^{(5)} = \begin{bmatrix}0{,}4 & 0{,}8\end{bmatrix}$; $\quad y^* = 1{,}648$.

Розділ 9.

9.4.

$$y(\mathbf{x}) = 108x_1 + 112x_2 + 126x_3 \to \min_{x_{1,2,3,} \in \Omega},$$

$$\begin{aligned} \Omega:\quad & 0{,}8x_1 + 0{,}5x_2 + 0{,}6x_3 - 800 \le 0; \\ & 0{,}4x_1 + 0{,}4x_2 + 0{,}3x_3 - 600 \le 0; \\ & 0{,}1x_2 + 0{,}1x_3 - 120 \le 0; \\ & \mathbf{x} \ge 0. \end{aligned}$$

9.5.

$$y(\mathbf{x}) = 0{,}09x_1 + 0{,}12x_2 + 0{,}1x_3 \to \min_{x_{1,2,3,} \in \Omega},$$

$$\begin{aligned} \Omega:\quad & x_1 + 3x_2 + 4x_3 - 60 \ge 0; \\ & 2x_1 + 4x_2 + 2x_3 - 5 \ge 0; \\ & x_1 + 4x_2 + 3x_3 - 12 \ge 0; \\ & \mathbf{x} \ge 0. \end{aligned}$$

9.6.

$$y(\mathbf{x}) = \begin{bmatrix}-2 & 1 & 5 & 0 & 0\end{bmatrix}\mathbf{x} \to \max_{\mathbf{x}\in\Omega},$$

$$\Omega: \begin{bmatrix} 4 & 2 & 5 & 1 & 0 \\ 6 & -3 & 4 & 0 & 0 \\ 3 & 3 & -2 & 0 & -1 \end{bmatrix} \mathbf{x}^{\mathrm{T}} + \begin{bmatrix} -12 \\ -18 \\ -16 \end{bmatrix} = 0;$$

$$\mathbf{x} \ge 0, \quad \mathbf{x}^{\mathrm{T}} = \begin{bmatrix} x_1 & x_2 & x_3 & x_4 & x_5 \end{bmatrix}.$$

9.7.
$$y(\mathbf{x}) = [2 \quad 5 \quad -3 \quad 0 \quad 0 \quad 0]\,\mathbf{x} \to \max_{\mathbf{x}\in\Omega},$$
$$\Omega: \begin{bmatrix} -1 & 1 & 1 & -1 & 0 & 0 \\ 2 & -1 & 1 & 0 & 1 & 0 \\ 3 & 1 & 1 & 0 & 0 & -1 \end{bmatrix} \mathbf{x}^{\mathrm{T}} + \begin{bmatrix} -4 \\ -16 \\ -18 \end{bmatrix} = 0;$$
$$\mathbf{x} \ge 0, \quad \mathbf{x}^{\mathrm{T}} = [x_1 \quad x_2 \quad x_3 \quad x_4 \quad x_5 \quad x_6].$$

9.8.
$$y(\mathbf{x}) = [-4 \quad 1] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \to \max_{\mathbf{x}\in\Omega},$$
$$\Omega: \begin{bmatrix} 2 & -1 \\ -2 & -2 \\ -1 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} -4 \\ -16 \\ -18 \end{bmatrix} \le 0;$$
$$\mathbf{x} \ge 0.$$

9.9. $\mathbf{x}^{*\mathrm{T}} = [14 \quad 0]; \quad y^* = 14$. **9.10.** $\mathbf{x}^{*\mathrm{T}} = [4{,}8 \quad 3{,}6]; \quad y^* = 12$.

9.11. $\mathbf{x}^{*\mathrm{T}} = [10 \quad 9]; \quad y^* = -11$. **9.13.** $\mathbf{x}^{*\mathrm{T}} = [2 \quad 6 \quad 33 \quad 0 \quad 0];$. $y^* = 22$. **9.14.** $\mathbf{x}^{*\mathrm{T}} = [\frac{18}{7} \quad 0 \quad \frac{34}{7} \quad \frac{8}{7} \quad 0]; \quad y^* = \frac{50}{7}$.

9.15. $\mathbf{x}^{*\mathrm{T}} = [102 \quad 166]; \quad y^* = 1940$. **9.16.** $\mathbf{x}^{*\mathrm{T}} = [12 \quad 6]; \quad y^* = 276$.

9.17. $\mathbf{x}^{*\mathrm{T}} = [3 \quad 4]; \quad y^* = 100$. **9.18.** $\mathbf{x}^{*\mathrm{T}} = [57 \quad 12]; \quad y^* = 1056$.

9.28. $\mathbf{x}^{*\mathrm{T}} = [0 \quad 0 \quad \frac{7}{4} \quad \frac{5}{4}]; \quad y^* = -3$. **9.29.** $\mathbf{x}^{*\mathrm{T}} = [0 \quad 8 \quad 20]; \quad y^* = 400.$

9.31. Рішення не існує. **9.32.** $\mathbf{x}^{*\mathrm{T}} = [3 \quad 2 \quad 0 \quad 1 \quad 0]; \quad y^* = -9$.

Не більше C_n^{n-m} (або C_n^{m}); $\mathbf{x}^{*\mathrm{T}} = [0 \quad 0 \quad 5 \quad 9 \quad 7]$;

$\mathbf{x}^{*\mathrm{T}} = [9/2 \quad 0 \quad 1/2 \quad 0 \quad 5/2]$; $\mathbf{x}^{*\mathrm{T}} = [0 \quad 5 \quad 0 \quad 4 \quad -3]$;

$\mathbf{x}^{*\mathrm{T}} = [7 \quad 0 \quad -2 \quad -5 \quad 0]$; $\mathbf{x}^{*\mathrm{T}} = [0 \quad 9 \quad -4 \quad 0 \quad -11]$; $\mathbf{x}^{*\mathrm{T}} = [4 \quad 1 \quad 0 \quad 0 \quad 1]$;

$\mathbf{x}^{*\mathrm{T}} = [0 \quad 7/2 \quad 3/2 \quad 11/2 \quad 0]$; $\mathbf{x}^{*\mathrm{T}} = [3 \quad 2 \quad 0 \quad 1 \quad 0]$;

$\mathbf{x}^{*\mathrm{T}} = [5 \quad 0 \quad 0 \quad -1 \quad 2]$; $\mathbf{x}^{*\mathrm{T}} = [11/3 \quad 5/3 \quad -1/3 \quad 0 \quad 0]$.

$\mathbf{x}^{*\mathrm{T}} = [95 \quad 210 \quad 0 \quad 0]$; $y^* = 2115$. **9.35.** $\mathbf{x}^{*\mathrm{T}} = [70 \quad 135 \quad 0 \quad 0]$; $y^* = 965$.

9.36. $\mathbf{x}^{*\mathrm{T}} = [24 \quad 18 \quad 0]$; $y^* = 492$.

9.37.. $\mathbf{x}^{*\mathrm{T}} = [100 \quad 0 \quad 1200]$; $y^* = 16200$

$\mathbf{x}^{*\mathrm{T}} = [0 \quad 0 \quad 11/2 \quad 35 \quad 0 \quad 1]$; $y^* = 68$.

Д О Д А Т К И

Додаток А. ОСНОВНІ ФОРМУЛИ ЕЛЕМЕНТАРНОЇ МАТЕМАТИКИ

А.1. Алгебраїчні функції

Особливості степенів

Для будь-яких x, y і додатних a, b справедливі такі рівності:

$$a^0 = 1; \qquad (ab)^x = a^x b^x; \qquad a^x a^y = a^{x+y}$$

$$\left(\frac{a}{b}\right)^x = \frac{a^x}{b^x}; \qquad \frac{a^x}{a^y} = a^{x-y}; \qquad a^{-x} = \frac{1}{a^x};$$

$$\left(a^x\right)^y = a^{xy}.$$

Многочлени

Для будь-яких a, b, c справедливі такі рівності:

$$a^2 - b^2 = (a-b)(a+b);$$

$$(a+b)^2 = a^2 + 2ab + b^2;$$

$$(a-b)^2 = a^2 - 2ab + b^2;$$

$$(a+b)^3 = a^3 + 3a^2b + 3ab^2 + b^3;$$

$$(a-b)^3 = a^3 - 3a^2b + 3ab^2 - b^3;$$

$$a^3 + b^3 = (a+b)\left(a^2 - ab + b^2\right);$$

$$a^3 - b^3 = (a-b)(a^2 + ab + b^2);$$

$$ax^2 + bx + c = a(x - x_1)(x - x_2)$$

де x_1, x_2 – корені рівняння $ax^2 + bx + c = 0$.

Особливості арифметичних коренів

Для будь-яких натуральних m, k, більших за 1, і будь-яких a, b справедливі такі рівності:

$$\sqrt[n]{ab} = \sqrt[n]{a} \cdot \sqrt[n]{b};$$

$$\sqrt[n]{\frac{a}{b}} = \frac{\sqrt[n]{a}}{\sqrt[n]{b}} \quad (b \neq 0);$$

$$\left(\sqrt[n]{a}\right)^k = \sqrt[n]{a^k}$$

$$\sqrt[n]{a} = \sqrt[nk]{a^k};$$

$$\sqrt[2n]{a^{2n}} = |a|;$$

$$\left(\sqrt[n]{a}\right)^n = a \quad (a \geq 0);$$

$$\sqrt[n]{a} < \sqrt[n]{b}, \quad \text{якщо} \quad 0 \leq a < b;$$

$$\sqrt[n]{\sqrt[k]{a}} = \sqrt[kn]{a};$$

$$\sqrt{a^2} = |a| = \begin{cases} a & \text{при} \quad a \geq 0, \\ -a & \text{при} \quad a < 0; \end{cases}$$

$$\sqrt[2n+1]{-a} = -\sqrt[2n+1]{a} \quad (a \geq 0).$$

А.2. Тригонометричні функції

Співвідношення між тригонометричними функціями одного і того ж аргументу:

$$\sin^2 x + \cos^2 x = 1;$$

$$\operatorname{tg} x = \frac{\sin x}{\cos x}, \quad x \neq \frac{\pi n}{2}, \quad n \in \mathbf{Z};$$

$$\operatorname{ctg} x = \frac{\cos x}{\sin x}, \quad x \neq \pi n, \quad n \in \mathbf{Z};$$

$$\operatorname{tg} x \cdot \operatorname{ctg} x = 1, \quad x \neq \frac{\pi n}{2}, \quad n \in \mathbf{Z};$$

$$1 + \operatorname{tg}^2 x = \frac{1}{\cos^2 x}, \quad x \neq \frac{\pi}{2}(2n+1), \quad n \in \mathbf{Z};$$

$$1 + \operatorname{ctg}^2 x = \frac{1}{\sin^2 x}, \quad x \neq \pi n, \quad n \in \mathbf{Z}.$$

Формули додавання:

$$\sin(x+y) = \sin x \cdot \cos y + \cos x \cdot \sin y;$$

$$\sin(x-y) = \sin x \cdot \cos y - \cos x \cdot \sin y;$$

$$\cos(x+y) = \cos x \cdot \cos y - \sin x \cdot \sin y;$$

$$\cos(x-y) = \cos x \cdot \cos y + \sin x \cdot \sin y;$$

$$\operatorname{tg}(x+y) = \frac{\operatorname{tg} x + \operatorname{tg} y}{1 - \operatorname{tg} x \cdot \operatorname{tg} \mathrm{y}}, \quad x, y, x+y \neq \frac{\pi}{2} + \pi n, \quad n \in \mathbf{Z};$$

$$\operatorname{tg}(x-y) = \frac{\operatorname{tg} x - \operatorname{tg} y}{1 + \operatorname{tg} x \cdot \operatorname{tg} \mathrm{y}}, \quad x, y, x-y \neq \frac{\pi}{2} + \pi n, \quad n \in \mathbf{Z}.$$

Формули подвійного аргументу:

$$\sin 2x = 2 \sin x \cdot \cos x;$$

$$\cos 2x = \cos^2 x - \sin^2 x = 2\cos^2 x - 1 = 1 - 2\sin^2 x;$$

$$\operatorname{tg} 2x = \frac{2 \operatorname{tg} x}{1 - \operatorname{tg}^2 x}, \quad x \neq \frac{\pi}{4} + \frac{\pi}{2} k, \quad k \in Z, \quad x \neq \frac{\pi}{2} + \pi n, \quad n \in \mathbf{Z}.$$

Формули половинного аргументу:

$$\sin^2\frac{x}{2}=\frac{1-\cos x}{2};$$

$$\cos^2\frac{x}{2}=\frac{1+\cos x}{2};$$

$$tg\frac{x}{2}=\frac{\sin x}{1+\cos x}=\frac{1-\cos x}{\sin x},\quad x\neq\pi+2\pi n,\quad n\in\mathbf{Z}.$$

Формули перетворення:

$$\sin x+\sin y=2\sin\frac{x+y}{2}\cdot\cos\frac{x-y}{2};$$

$$\sin x-\sin y=2\cos\frac{x+y}{2}\cdot\sin\frac{x-y}{2};$$

$$\cos x+\cos y=2\cos\frac{x+y}{2}\cdot\cos\frac{x-y}{2};$$

$$\cos x-\cos y=-2\sin\frac{x+y}{2}\cdot\sin\frac{x-y}{2};$$

$$\operatorname{tg}x-\operatorname{tg}y=\frac{\sin(x-y)}{\cos x\cdot\cos y},\quad x,y\neq\frac{\pi}{2}+\pi n,\quad n\in\mathbf{Z};$$

$$2\sin x\cdot\sin y=\cos(x-y)-\cos(x+y);$$

$$2\cos x\cdot\cos y=\cos(x-y)+\cos(x+y);$$

$$2\sin x\cdot\cos y=\sin(x-y)+\sin(x+y).$$

Співвідношення між $\sin x, \cos x, tg\frac{x}{2}$:

$$\sin x = \frac{2\mathrm{tg}\,\frac{x}{2}}{1+\mathrm{tg}^2\,\frac{x}{2}}, \quad x \neq (2n+1)\pi, \quad n \in \mathbf{Z};$$

$$\cos = \frac{1-\mathrm{tg}^2\,\frac{x}{2}}{1+\mathrm{tg}^2\,\frac{x}{2}}, \quad x \neq (2n+1)\pi, \quad n \in \mathbf{Z}.$$

А.3. Особливості логарифмів

$$x = a^{\log_a x}, \quad x > 0.$$

$$\log_a a = 1; \qquad \log_a 1 = 0;$$

$$\log_a (x_1 x_2) = \log_a x_1 + \log_a x_2, \quad x_{1,2} > 0;$$

$$\log_a \frac{x_1}{x_2} = \log_a x_1 - \log_a x_2, \quad x_{1,2} > 0;$$

$$\log_a x^p = p\log_a x, \quad x > 0, \quad p \in \mathbf{R};$$

$$\log_a x = \frac{\log_a x}{\log_b a}, \quad x > 0, \quad b > 0, \quad b \neq 1.$$

Додаток В. ОСНОВНІ ФОРМУЛИ ДИФЕРЕНЦІАЛЬНОГО ЧИСЛЕННЯ

В.1. Похідна:

$$y' = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x}.$$

В.2. Правила диференціювання:

$$c' = 0;$$

$$(cU)' = cU';$$

$$(U \pm V)' = U' \pm V';$$

$$(U \cdot V)' = U'V + UV';$$

$$\left(\frac{U}{V}\right)' = \frac{U'V - UV'}{V^2};$$

$$\left(f(U)\right)_x' = f_U' \cdot U_x'.$$

Тут c – константа, $U = U(x), \quad V = V(x)$.

В2. Формули диференціювання:

$$\left(U^a\right)' = aU^{a-1}U', \quad a \in \mathbf{R};$$

$$\left(\log_a U\right)' = \frac{U'}{U \ln a};$$

$$(\ln U)' = \frac{U'}{U};$$

$$\left(a^U\right)' = a^U \ln a \cdot U';$$

$$(\sin U)' = \cos U \cdot U';$$

$$(\cos U)' = -\sin U \cdot U';$$

$$(\mathrm{tg} U)' = \frac{1}{\cos^2 U} U';$$

$$(\mathrm{ctg} U)' = -\frac{1}{\sin^2 U} U';$$

$$(\arcsin U)' = \frac{1}{\sqrt{1-U^2}} U';$$

$$(\arccos U)' = -\frac{1}{\sqrt{1-U^2}} U';$$

$$(\mathrm{arctg} U)' = \frac{1}{1+U^2} U';$$

$$(\mathrm{arcctg} U)' = -\frac{1}{1+U^2} U'.$$

Тут $U = U(x)$. Якщо $U(x) = x$, то $U'(x) = x' = 1$.

СПИСОК ЛІТЕРАТУРИ

Акулич И.А. Математическое программирование в примерах и задачах. – М.: Высш. шк., 1986. – 319 с.

Горбатов В.А. Основы дискретной математики. – М.: Высш. шк., 1986. – 311 с.

Евдокимов А.Г. Минимизация функций. – Харьков: Изд-во при Харьк. ун-те, 1977. – 160 с.

Евдокимов А.Г. Минимизация функций и ее приложения к задачам автоматизированного управления инженерными сетями. – Харьков: Изд-во при Харьк. ун-те, 1985. – 288 с

Евдокимов А.Г., Самойленко Н.И., Пальченко Л.А, Рябченко И.Н. Минимизация функций с применением микро- и мини-ЭВМ. Сборник задач и упражнений. – Харьков: Основа, 1993. – 256 с.

Корбут А.А., Финкельштейн Ю.Ю. Дискретное программирование. – М.: Наука, 1969. – 368 с.

Пшеничный Б.И., Данилин Ю.М. Численные методы в экстремальных задачах. – М.: Наука, 1975. – 319 с.

Самойленко М.І. Курс лекцій з математичного програмування.– Харків: ХДАМГ, 1997. –103 с.

Справочник по математике для экономистов / Под ред. *В.И.Ермакова*. – М.: Высш. шк., 1987. – 336 с.

Сухарев А.Г., Тимохов А.В., Фёдоров В.В. Курс методов оптимизации. – М.: Наука, 1988. –328 с.

Фиакко А., *Мак-Кормик Г.* Нелинейное программирование. Методы последовательной безусловной оптимизации. – М.: Мир, 1972. – 240 с.

ПРЕДМЕТНИЙ ПОКАЖЧИК

З М І С Т

Навчальне видання

САМОЙЛЕНКО Микола Іванович

М А Т Е М А Т И Ч Н Е   П Р О Г Р А М У В А Н Н Я

Навчальний посібник

Редактор *М.З.Аляб’єв*
Художник обкладинки *М.І.Самойленко*
Комп’ютерна верстка *М.І.Самойленка*

Підп. до друку оригінал-макет 14.11.2001. Формат 60х84/16.
Папір офсетний. Гарнітура Таймс. Друк на ризографі.
Умов.-друк. арк. 24,64. Обл.-вид. арк 25,9. Тираж 500 прим. Зам. № .

Державне спеціалізоване видавництво “ОСНОВА”
при Харківському університеті.
Україна, 61052, Харків, вул. Маршала Конєва, 10/2.

Сектор оперативної поліграфії ІОЦ ХДАМГ.
Харків, вул. Революції, 24.