1a

1. Загальні поняття системи лінійних рівнянь

Розв'язання багатьох економічних задач зводиться до дослідження й розв'язання систем лінійних рівнянь.

Система m лінійних рівнянь із n змінними має вигляд

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots, \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m, \end{cases} \qquad (1)$$

або в скороченому записі

$$\sum_{j=1}^{n} a_{ij}x_j = b_i, \quad i = 1, 2, \ldots, m.$$

Тут роль змінних, які підлягають визначенню, грають величини x_j, що називаються *невідомими*. Параметрами є змінні a_{ij} і b_i, які можуть набувати будь-яких дійсних значень. Всі a_{ij} називаються *коефіцієнтами* при змінних, всі b_i — *вільними членами* рівнянь.

Розв'язком системи називається впорядкована сукупність чисел $\alpha_1, \alpha_2, \ldots, \alpha_n$, що при підстановці в систему перетворює кожне рівняння в тотожність. Лінійна система рівнянь називається *сумісною*, якщо вона має хоча б один розв'язок, і *несумісною*, якщо вона не має розв'язків. Сумісна система рівнянь називається визначеною, якщо вона має єдиний розв'язок, і *невизначеною*, якщо вона має більше одного р*озв'язку*. Дві системи рівнянь називаються рівносильними, якщо їх множини розв'язків співпадають.

Системи рівнянь зручно розв'язувати в матричній формі. Запишемо систему (1) у вигляді матричного рівняння. Введемо позначення:

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix}; \quad X = \begin{pmatrix} x_1 \\ x_2 \\ \ldots \\ x_n \end{pmatrix}; \quad B = \begin{pmatrix} b_1 \\ b_2 \\ \ldots \\ b_m \end{pmatrix},$$

2a

2. Знаходження єдиного розв'язання системи лінійних рівнянь

Система рівнянь може бути представлена в матричному вигляді

$$\begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ \ldots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \ldots \\ b_m \end{pmatrix}$$

або в компактній матричній формі

$$A \cdot X = B.$$

Метод оберненої матриці

Розв'яжемо це рівняння у випадку $m = n$ і $|A| \neq 0$. Для матриці A існує обернена матриця A^{-1}. Помножимо ліворуч обидві частини матричного рівняння на A^{-1}:

$$A^{-1} \cdot A \cdot X = A^{-1} \cdot B$$

або

$$X = A^{-1} \cdot B.$$

Рівняння в матричному вигляді розв'язане. Для знаходження елементів матриці невідомих X варто знайти обернену матрицю коефіцієнтів і помножити її на стовпець вільних членів B.

ПРИКЛАД. Розв'язати систему рівнянь

$$\begin{cases} x_1 - x_2 + 2x_3 = 1, \\ -2x_1 + x_2 + x_3 = 2, \\ x_1 - x_2 + x_3 = -1. \end{cases} \qquad (1)$$

Розв'язання. Запишемо систему в матричному вигляді

$$\begin{pmatrix} 1 & -1 & 2 \\ -2 & 1 & 1 \\ 1 & -1 & 1 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}.$$

Матриця коефіцієнтів

$$A = \begin{pmatrix} 1 & -1 & 2 \\ -2 & 1 & 1 \\ 1 & -1 & 1 \end{pmatrix}$$

не вироджена (її визначник $|A| = 1$). Тому існує обернена матриця

3a

3. Загальний підхід до розв'язання систем рівнянь

ТЕОРЕМА (про рівносильність систем при елементарних перетвореннях)

При елементарних перетвореннях рядків системи перших чотирьох типів лінійні системи залишаються рівносильними.

Доведення. Розглянемо елементарні перетворення кожного типу окремо.

1. Якщо система лінійних рівнянь отримана з початкової системи елементарними перетвореннями 1-го типу, тобто змінений порядок рівнянь у системі, то розв'язки системи не зміняться.

2. Якщо система лінійних рівнянь отримана з початкової системи елементарними перетвореннями 2-го типу, тобто одне з рівнянь помножене на число $a \neq 0$, це не призведе до зміни розв'язків системи.

3. Якщо система лінійних рівнянь отримана з початкової системи елементарними перетвореннями 3-го типу, тобто одне з рівнянь являє собою суму двох рівнянь, одне з яких попередньо було помножене на число λ, це також не призведе до зміни розв'язків системи.

Дійсно, нехай $x_1^0, x_2^0, \ldots, x_n^0$ — розв'язки системи рівнянь. Тоді рівняння з довільними номерами i і j —

$$\begin{aligned} a_{i1}x_1 + \ldots + a_{in}x_n &= b_i \\ a_{j1}x_1 + \ldots + a_{jn}x_n &= b_j, \end{aligned}$$

при підстановці чисел $x_1^0, x_2^0, \ldots, x_n^0$ перетворяться в тотожності. Додамо обидва рівняння, попередньо помноживши перше з них на число λ:

$$\lambda(a_{i1}x_1 + \ldots + a_{in}x_n - b_i) + (a_{j1}x_1 + \ldots + a_{jn}x_n - b_j) = 0.$$

Підставивши сюди числа $x_1^0, x_2^0, \ldots, x_n^0$, отримаємо $\lambda \cdot 0 + 0 = 0$.

4. Якщо система лінійних рівнянь отримана з початкової системи елементарними перетвореннями 4-го типу, тобто одне з рівнянь, що містить нульові коефіцієнти й нульовий вільний член, викреслене, це, вочевидь, не змінить розв'язків системи.

4a

4. Базисні розв'язки системи рівнянь

Розв'язування багатьох економічних задач зводиться до дослідження й розв'язання систем лінійних рівнянь.

Система m лінійних рівнянь із n змінними має вигляд

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots, \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m, \end{cases} \qquad (1)$$

або в скороченому записі

$$\sum_{j=1}^{n} a_{ij}x_j = b_i, \quad i = 1, 2, \ldots, m.$$

Нехай число лінійно незалежних рівнянь менше числа змінних, отже, $r < n$. Назвемо r змінних *основними,* або *базисними*, якщо визначник матриці з коефіцієнтів при них відрізняється від нуля. Припустимо, що це змінні $x_1, x_2, \ldots, x_r$. Тоді змінні $x_1, x_2, \ldots, x_r$ виражаються через змінні $x_{r+1}, x_{r+2}, \ldots, x_n$.

$$\begin{cases} x_1 = \ddot{a}_{1,r+1}x_{r+1} + \ddot{a}_{1,r+2}x_{r+2} + \ldots + \ddot{a}_{1n}x_n + \ddot{b}_1, \\ x_2 = \ddot{a}_{2,r+1}x_{r+1} + \ddot{a}_{2,r+2}x_{r+2} + \ldots + \ddot{a}_{2n}x_n + \ddot{b}_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots, \\ x_r = \ddot{a}_{r,r+1}x_{r+1} + \ddot{a}_{r,r+2}x_{r+2} + \ldots + \ddot{a}_{rn}x_n + \ddot{b}_r. \end{cases} \qquad (1')$$

Надаючи невідомим $x_{r+1}, \ldots, x_n$ довільних значень, одержуємо нескінченну множину розв'язків системи рівнянь.

Назвемо останні змінні *неосновними,* або *вільними*. Кількість цих змінних дорівнює $n - r$.

Означення. Розв'язок системи (1), у якому всі $n - r$ вільних змінних припускаються рівними нулю, називається *базисним*. У системі (1') базисним розв'язком є

$$\begin{cases} x_1 = \ddot{b}_1, \\ x_2 = \ddot{b}_2, \\ \ldots\ldots\ldots, \\ x_r = \ddot{b}_r, \\ x_{r+1} = 0, \\ \ldots\ldots\ldots, \\ x_n = 0 \end{cases} \quad \text{або в матричній формі} \quad \begin{pmatrix} x_1 \\ x_2 \\ \ldots \\ x_r \\ x_{r+1} \\ \ldots \\ x_n \end{pmatrix} = \begin{pmatrix} \ddot{b}_1 \\ \ddot{b}_2 \\ \ldots \\ \ddot{b}_r \\ 0 \\ \ldots \\ 0 \end{pmatrix}.$$

5а

5. Однорідні системи лінійних рівнянь

Система лінійних рівнянь називається **однорідною**, якщо всі вільні члени системи дорівнюють нулю:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = 0, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots, \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = 0. \end{cases} \quad (1)$$

У матричному вигляді систему можна записати в такий спосіб:

$$A \cdot X = 0,$$

де

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix}; \quad X = \begin{pmatrix} x_1 \\ x_2 \\ \ldots \\ x_n \end{pmatrix}; \quad 0 = \begin{pmatrix} 0 \\ 0 \\ \ldots \\ 0 \end{pmatrix}.$$

Властивості однорідної системи лінійних рівнянь

1) Однорідна система завжди сумісна, тому що завжди має, щонайменше, нульовий розв'язок.

2) Для існування ненульових розв'язків ранг матриці коефіцієнтів повинен бути менше числа змінних $r < n$, тобто число лінійно незалежних рівнянь повинне бути менше числа змінних. У цьому випадку $|A| = 0$.

3) Якщо матриця-стовпець (вектор) $\boldsymbol{e} = \begin{pmatrix} x_{11} \\ x_{21} \\ \ldots \\ x_{n1} \end{pmatrix}$ є розв'язком системи (1), то і стовпець $\boldsymbol{e}' = \lambda \cdot \boldsymbol{e} = \begin{pmatrix} \lambda x_{11} \\ \lambda x_{21} \\ \ldots \\ \lambda x_{n1} \end{pmatrix}$ також є розв'язком системи.

Нехай $\boldsymbol{e}$ — розв'язок системи. Тоді матричне рівняння при підстановці $x = \boldsymbol{e}'$ перетворюється в тотожність $A \cdot \boldsymbol{e}' = 0$.

Дійсно, знайдемо добуток матриць A і $\lambda \cdot \boldsymbol{e}$:

$$A \cdot (\lambda \boldsymbol{e}) = \lambda \cdot (A \cdot \boldsymbol{e}) = \lambda \cdot 0 = 0.$$

6а

6. Загальний розв'язок системи неоднорідних лінійних рівнянь

ТЕОРЕМА (про загальний розв'язок системи неоднорідних рівнянь)

Загальний розв'язок системи лінійних рівнянь із n змінними дорівнює сумі загального розв'язку відповідної їй системи однорідних лінійних рівнянь і деякого часткового розв'язку початкової системи.

Доведення. Для системи m лінійних рівнянь із n змінними

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots, \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m, \end{cases}$$

запишемо розширену матрицю

$$\left(\begin{array}{cccc|c} a_{11} & a_{12} & \ldots & a_{1n} & b_1 \\ a_{21} & a_{22} & \ldots & a_{2n} & b_2 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & \ldots & a_{mn} & b_m \end{array}\right).$$

Використовуючи елементарні перетворення, приведемо її до східчастого вигляду:

$$\left(\begin{array}{cccccc|c} \dot{a}_{11} & \dot{a}_{12} & \ldots & \dot{a}_{1r} & \ldots & \dot{a}_{1n} & \dot{b}_1 \\ 0 & \dot{a}_{22} & \ldots & \dot{a}_{2r} & \ldots & \dot{a}_{2n} & \dot{b}_2 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & \dot{a}_{rr} & \ldots & \dot{a}_{rn} & \dot{b}_{rr} \\ 0 & 0 & \ldots & 0 & \ldots & 0 & \dot{b}_{r+1} \\ 0 & 0 & \ldots & 0 & \ldots & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & 0 & \ldots & 0 & 0 \end{array}\right).$$

Зрозумівши, що система має розв'язки, методом Гаусса знайдемо послідовно змінні $x_r, x_{r-1}, \ldots, x_1$, виражені через вільні змінні $x_{r+1}, x_{r+2}, \ldots, x_n$. Вільні змінні позначимо $x_{r+1} = c_1, x_{r+2} = c_2, \ldots, x_n = c_{n-r}$. У результаті одержимо розв'язок системи в загальному вигляді:

7а

7. Модель Леонтьєва багатогалузевої економіки

Багатогалузеве господарство вимагає балансу між окремими галузями. Виникає задача: узгодити обсяги виробництва кожної з галузей, щоб задовольнити всі потреби в продукті кожної галузі.

Розглянемо для визначеності виробничу сферу з n галузей, кожна з яких робить один (свій) продукт. Виділимо певний період часу, наприклад рік, протягом якого всі коефіцієнти залишаються постійними.

Нехай

x_i — загальний (валовий) випуск i-ї галузі ($x_i \geq 0$), $i = 1, 2, \ldots, n$;

x_{ij} — об'єм продукції i-ї галузі, що поставляється для j-ї галузі в процесі виробництва;

y_i — кінцевий попит на продукцію i-ї галузі ($y_i \geq 0$). Сюди належать особисте споживання громадян, утримання державних і суспільних інститутів, чистий експорт, виробниче накопичування й т. д. Тоді балансові співвідношення матимуть вигляд

$$\begin{cases} x_1 = x_{11} + x_{12} + \ldots + x_{1n} + y_1, \\ x_2 = x_{21} + x_{22} + \ldots + x_{2n} + y_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots, \\ x_n = x_{n1} + x_{n2} + \ldots + x_{nn} + y_n. \end{cases} \quad (1)$$

Введемо коефіцієнти прямих витрат $a_{ij} = \dfrac{x_{ij}}{x_j} \geq 0$; $i, j = 1, 2, \ldots, n$, що показують витрати i-ї галузі на випуск однієї одиниці продукції для j-ї галузі. Замінивши в (1) $x_{ij} = a_{ij}x_j$, одержимо систему n лінійних рівнянь із n змінними:

$$\begin{cases} x_1 = a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n + y_1, \\ x_2 = a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n + y_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots, \\ x_n = a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n + y_n \end{cases}$$

або в матричній формі

$$X = A \cdot X + Y, \quad (2)$$

де X — матриця (вектор) випусків галузей;

A — матриця прямих витрат;

Y — матриця (вектор) кінцевого попиту.

8а

8. Метод Гаусса

Метод Гаусса полягає в послідовному виключенні невідомих за допомогою елементарних перетворень.

Для системи рівнянь

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots, \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m, \end{cases}$$

утворимо розширену матрицю:

$$\left(\begin{array}{cccc|c} a_{11} & a_{12} & \ldots & a_{1n} & b_1 \\ a_{21} & a_{22} & \ldots & a_{2n} & b_2 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & \ldots & a_{mn} & b_m \end{array}\right).$$

За допомогою елементарних перетворень приведемо розширену матрицю до східчастого вигляду

$$\left(\begin{array}{cccccc|c} \dot{a}_{11} & \dot{a}_{12} & \ldots & \dot{a}_{1r} & \ldots & \dot{a}_{1n} & \dot{b}_1 \\ 0 & \dot{a}_{22} & \ldots & \dot{a}_{2r} & \ldots & \dot{a}_{2n} & \dot{b}_2 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & \dot{a}_{rr} & \ldots & \dot{a}_{rn} & \dot{b}_{rr} \\ 0 & 0 & \ldots & 0 & \ldots & 0 & \dot{b}_{r+1} \\ 0 & 0 & \ldots & 0 & \ldots & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & 0 & \ldots & 0 & 0 \end{array}\right).$$

Розглянемо різні випадки:

1) $\dot{b}_{r+1} \neq 0$. Система рівнянь несумісна. Дійсно, рівняння з номером $r + 1$ містить нульові коефіцієнти перед невідомими, в той час як вільний член відмінний від нуля. Нехай далі $\dot{b}_{r+1} = 0$.

2) Число невідомих n і число рівнянь r співпадають. Розширена матриця прийме трикутний вигляд:

$$\left(\begin{array}{cccc|c} \dot{a}_{11} & \dot{a}_{12} & \ldots & \dot{a}_{1n} & \dot{b}_1 \\ 0 & \dot{a}_{22} & \ldots & \dot{a}_{2n} & \dot{b}_2 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & \dot{a}_{nn} & \dot{b}_n \end{array}\right).$$

9а

9. Обчислення визначника

Існує кілька способів обчислення величини визначника. Вибір способу диктується виглядом і порядком визначника. Вдало обраний спосіб дозволяє істотно скоротити обчислення. Розглянемо їх на прикладі визначника матриці третього порядку.

ПРИКЛАД. Обчислити визначник матриці

$$A=\begin{pmatrix} 2 & 2 & -1 \\ 2 & -4 & 1 \\ -2 & -2 & 2 \end{pmatrix}.$$

1-й спосіб. Використання теореми про розкладання визначника по будь-якому рядку або стовпцю.

Розкладемо визначник, наприклад, по третьому стовпцю:

$$\begin{vmatrix} 2 & 2 & -1 \\ 2 & -4 & 1 \\ -2 & -2 & 2 \end{vmatrix} = -1(-1)^{1+3}\begin{vmatrix} 2 & -4 \\ -2 & -2 \end{vmatrix} + 1(-1)^{2+3}\begin{vmatrix} 2 & 2 \\ -2 & -2 \end{vmatrix} + 2(-1)^{3+3}\begin{vmatrix} 2 & 2 \\ 2 & -4 \end{vmatrix} =$$

$$= -(-4-8)-(-4+4)+2(-8-4) = -12.$$

2-й спосіб. Використання правила трикутників:

$$\begin{vmatrix} 2 & 2 & -1 \\ 2 & -4 & 1 \\ -2 & -2 & 2 \end{vmatrix} = 2\cdot(-4)\cdot 2 + 2\cdot 1\cdot(-2) + 2\cdot(-2)\cdot(-1) - (-1)\cdot(-4)\cdot(-2) -$$

$$-2\cdot 2\cdot 2 - (-2)\cdot 1\cdot 2 = -16-4+4+8-8+4 = -12.$$

3-й спосіб. Використання властивостей визначника для перетворення його до вигляду, коли він містить рядок або стовпець із максимальною кількістю нулів. Розкладання визначника по цьому рядку (стовпцю):

$$\begin{vmatrix} 2 & 2 & -1 \\ 2 & -4 & 1 \\ -2 & -2 & 2 \end{vmatrix} = \{\text{додамо до першого рядка третій}\} = \begin{vmatrix} 0 & 0 & 1 \\ 2 & -4 & 1 \\ -2 & -2 & 2 \end{vmatrix} =$$

$= \{$розкладемо визначник по першому рядку$\} =$

$$= 1\cdot(-1)^{1+3}\begin{vmatrix} 2 & -4 \\ -2 & -2 \end{vmatrix} = -4-8 = -12.$$

10а

10. Матриці

Матрицею з розмірами $m \times n$ називається прямокутна таблиця чисел, що містить m рядків і n стовпців. Числа, що утворюють матрицю, називаються **елементами матриці**. Матриці звичайно позначають великими літерами латинського алфавіту, а для позначення елементів матриці використовуються малі літери з подвійною індексацією: a_{ij}, де i — номер рядка, j — номер стовпця. Числа i і j визначають розташування елемента a_{ij} у матриці A та відіграють роль координат цього елемента в прямокутній таблиці чисел.

Наприклад, матриця

$$A=\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

має m рядків і n стовпців.

Набір $(a_{i1}, a_{i2}, \dots, a_{in})$ називається i-м рядком матриці A, а набір

$$\begin{pmatrix} a_{1j} \\ a_{2j} \\ \dots \\ a_{mj} \end{pmatrix}$$

називається j-м стовпцем матриці A. Будь-які рядки й стовпці матриці A, у свою чергу, також є матрицями.

Дві матриці A та B однакового розміру називаються **рівними**, якщо вони співпадають поелементно. Рівність записується як $A = B$.

Види матриць

Матриця довільного розміру, всі елементи якої дорівнюють нулю, називається **нульовою** і позначається O.

Матриця, що складається з одного рядка, наприклад, $A = (a_{11}, a_{12}, \dots, a_{1n})$, називається **матрицею-рядком**, або **вектором**.

Матриця, що складається з одного стовпця, наприклад, $A=\begin{pmatrix} a_{11} \\ a_{21} \\ \dots \\ a_{n1} \end{pmatrix}$, називається **матрицею-стовпцем** або також **вектором**.

Матриця називається **квадратною n-го порядку**, якщо число її рядків дорівнює числу стовпців і дорівнює n.

11а

11. Операції над матрицями

Множення числа на матрицю

Ця операція виконується за наступним правилом: число необхідно помножити на кожний елемент матриці.

Добутком числа на матрицю $A = (a_{ij})$ називається матриця $B = (b_{ij})$, така, що $B = \lambda \cdot A$.

Елементи матриці B обчислюються за формулою $b_{ij} = \lambda \cdot a_{ij}$, де $i = 1, 2, \dots, m$; $j = 1, 2, \dots, n$.

Зауваження. Спільний множник всіх елементів матриці можна виносити за знак матриці.

Додавання матриць однакового розміру

Відповідні елементи матриць додаються.

Сумою матриць $A = (a_{ij})$ і $B = (b_{ij})$ називається матриця $C = (c_{ij})$, така, що $C = A + B$. Елементи матриці C обчислюються за формулою $c_{ij} = a_{ij} + b_{ij}$, де $i = 1, 2, \dots; m$, $j = 1, 2, \dots, n$.

Віднімання матриць однакового розміру

Відповідні елементи матриць віднімаються.

Різницею матриць $A = (a_{ij})$ і $B = (b_{ij})$ називається матриця $C = (c_{ij})$, така, що $C = A + (-1) \times B$. Елементи матриці C обчислюються за формулою $c_{ij} = a_{ij} - b_{ij}$, де $i = 1, 2, \dots, m$; $j = 1, 2, \dots, n$.

Множення матриці на матрицю

Елемент нової матриці, що стоїть на перетині i-го рядка й j-го стовпця, дорівнює сумі добутків елементів i-го рядка першої матриці і відповідних елементів j-го стовпця другої матриці. Операція визначена за умови, що число стовпців першої матриці дорівнює числу рядків другої.

Добутком матриці $A = (a_{ij})$ і матриці $B = (b_{ij})$ називається матриця $C = (c_{ij})$, така, що $C = A \cdot B$. Елементи матриці C обчислюються за формулою

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \dots + a_{in}b_{nj}. \qquad (1)$$

Використовуючи знак суми, формулу (1) можна записати у вигляді

$$c_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj}, \quad i = 1, 2, \dots, m, \quad j = 1, 2, \dots, n.$$

12а

12. Визначники квадратних матриць

Зв'яжемо з кожною квадратною матрицею A деяке число, що вводиться за певним правилом. Назвемо це число **визначником матриці** й позначимо його $|A|$.

Визначником матриці першого порядку $A = (a_{11})$ назвемо число $|A| = (a_{11})$.

Визначником матриці другого порядку $A=\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ назвемо число, що дорівнює

$$|A| = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}M_{11} - a_{12}M_{12},$$

де M_{1j} (індекс j дорівнює 1 або 2) — визначник матриці першого порядку, отриманий викреслюванням з матриці A 1-го рядка й j-го стовпця.

Тоді $|A| = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}M_{11} - a_{12}M_{12} = a_{11}a_{22} - a_{12}a_{21}.$

Визначником матриці третього порядку $A=\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$ назвемо число, що дорівнює

$$|A| = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}M_{11} - a_{12}M_{12} + a_{13}M_{13}, \qquad (1)$$

де M_{1j} (індекс j дорівнює 1, 2 або 3) — визначник матриці другого порядку, отриманий викреслюванням з матриці A 1-го рядка й j-го стовпця. Наприклад, визначник M_{11} отриманий з матриці A викреслюванням 1-го рядка й 1-го стовпця:

$$M_{11} = \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} = a_{22}a_{33} - a_{23}a_{32}, \quad M_{12} = \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} = a_{21}a_{33} - a_{23}a_{31},$$

$$M_{13} = \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} = a_{21}a_{32} - a_{22}a_{31}.$$

Підставимо отримані співвідношення в (1):

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

$$+a_{13}(a_{21}a_{32} - a_{22}a_{31}) = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + \qquad (2)$$

$$+a_{21}a_{32}a_{13} - a_{13}a_{22}a_{31} - a_{11}a_{32}a_{23} - a_{21}a_{12}a_{33}.$$

Зі структури формули видно, що в кожний доданок у правій частині рівності входить по одному елементу з кожного рядка й кожного стовпця матриці. Формулу (2) нескладно запам'ятати, якщо

13а

13. Розв'язування систем з квадратною матрицею за правилом Крамера

Спосіб розв'язування системи з n рівнянь із n невідомими дає теорема швейцарського математика Габріеля Крамера (1704–1752).

ТЕОРЕМА КРАМЕРА (про розв'язування системи рівнянь за допомогою визначників)

Нехай у квадратній матриці коефіцієнтів при невідомих у системі з n лінійних рівнянь із n невідомими визначник $|A| \neq 0$. Нехай $|A|_j$ — визначник матриці, отриманої з матриці A заміною j-го стовпця стовпцем вільних членів. Тоді система має єдиний розв'язок, який має вигляд

$$x_1 = \frac{|A_1|}{|A|},\ x_2 = \frac{|A_2|}{|A|},\ \ldots,\ x_n = \frac{|A_n|}{|A|}.$$

Доведення. Розгорнемо матричне рівняння $X = A^{-1} \cdot B$ і запишемо обернену матрицю через алгебраїчні доповнення:

$$\begin{pmatrix} x_1 \\ x_2 \\ \ldots \\ x_n \end{pmatrix} = \frac{1}{|A|} \begin{pmatrix} A_{11} & A_{21} & \ldots & A_{n1} \\ A_{12} & A_{22} & \ldots & A_{n2} \\ \ldots & \ldots & \ldots & \ldots \\ A_{1n} & A_{2n} & \ldots & A_{nn} \end{pmatrix} \cdot \begin{pmatrix} b_1 \\ b_2 \\ \ldots \\ b_n \end{pmatrix}.$$

Перемноживши матриці, маємо

$$\begin{pmatrix} x_1 \\ x_2 \\ \ldots \\ x_n \end{pmatrix} = \frac{1}{|A|} \begin{pmatrix} A_{11}b_1 + A_{21}b_2 + \ldots + A_{n1}b_n \\ A_{12}b_1 + A_{22}b_2 + \ldots + A_{n2}b_n \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ A_{1n}b_1 + A_{2n}b_2 + \ldots + A_{nn}b_n \end{pmatrix}.$$

Сума $A_{11}b_1 + A_{21}b_2 + \ldots + A_{n1}b_n$ є добутком чисел $b_1, b_2, \ldots, b_n$ і алгебраїчних доповнень елементів 1-го стовпця. Вона дорівнює визначнику матриці, отриманої з даної заміною елементів цього стовпця на числа $b_1, b_2, \ldots, b_n$ (властивість 9 визначників). Отже,

14а

14. Критерій Сильвестра

ТЕОРЕМА (про визначення знака форми за кутовими мінорами)

Для того щоб квадратична форма з матрицею

$$\begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{12} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{1n} & a_{2n} & \ldots & a_{nn} \end{pmatrix}$$

була додатно (від'ємно) визначеною, необхідно й достатньо, визначивши знаки кутових мінорів її матриці

$$M_1 = a_{11},\quad M_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{vmatrix},\ \ldots,\ M_n = \begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{12} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{1n} & a_{2n} & \ldots & a_{nn} \end{vmatrix},$$

керуватися наступним:

1) якщо всі кутові мінори додатні, квадратична форма буде додатно визначеною;

2) якщо мінори мають переміжні знаки, починаючи з від'ємного, квадратична форма буде від'ємно визначеною.

Доведення. Із чотирьох сформульованих у критерії тверджень доведемо, наприклад, що при $\hat{L}(\boldsymbol{x}) > 0$ кутові мінори $M_1 > 0, M_2 > 0, \ldots, M_n > 0$. Нехай додатно визначена квадратична форма задана в базисі $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n$. Розглянемо підпростір V_k, побудований на базисі $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_k$, де $1 \le k \le n$, і квадратичну форму $\hat{L}(\boldsymbol{x})$, що на підпросторі V_k матиме вигляд $\hat{L}_k(\boldsymbol{x})$. Квадратична форма $\hat{L}_k(\boldsymbol{x}) > 0$ для будь-якого ненульового вектора $\boldsymbol{x} = (x_1, x_2, \ldots, x_k)$ з підпростору V_k, а значить, і всі власні числа $\lambda_1, \lambda_2, \ldots, \lambda_k$ додатні. Звідси визначник $|L_k|$ матриці квадратичної форми

$$L_k = \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ 0 & 0 & \ldots & 0 \\ 0 & 0 & \ldots & \lambda_k \end{pmatrix}$$

також додатний: $|L_k| > 0$. Але $|L_k| = M_k$. Виходить, $M_k > 0$. Перебираючи всі $k = 1, 2, \ldots, n$, прийдемо до нерівностей $M_1 > 0, M_2 > 0, \ldots, M_n > 0$.

Приведемо таблицю визначення знаку квадратичної форми.

15а

15. Властивості визначників

1) Визначник з нульовим рядком або нульовим стовпцем дорівнює нулю.

2) Множення визначника на число рівносильне множенню будь-якого рядка або стовпця визначника на це число.

Помножимо будь-який рядок або стовпець початкового визначника на число, розкладемо визначник по цьому рядку або стовпцю, винесемо це число за дужки й згорнемо вираз, що залишився в дужках, у початковий визначник.

3) При транспонуванні матриці величина її визначника не змінюється: $|A| = |A^T|$.

Розкладемо визначник $|A|$ по першому рядку, транспонуємо його. Розкладемо отриманий визначник $|A^T|$ по першому стовпцю. З теореми про визначник матриці випливає, що результат буде однаковий.

4) При перестановці двох рядків або стовпців визначник змінює знак.

5) Визначник матриці із двома однаковими рядками (стовпцями) дорівнює нулю.

При перестановці двох рядків визначник змінить знак. Переставимо місцями однакові рядки. Визначник залишиться таким самим. Виходить, $-|A| = |A|$. Звідси випливає, що $|A| = 0$.

6) Визначник, що містить два пропорційні рядки (стовпця), дорівнює нулю.

Винесемо коефіцієнт пропорційності за знак визначника. У ньому утворяться два однакові рядки. Тому такий визначник дорівнює нулю.

7) Визначник можна розкласти на суму визначників.

Представимо елементи i-го рядка визначника у вигляді суми двох доданків. Маємо

$$|A| = \begin{vmatrix} a_{11} & \ldots & a_{1n} \\ \ldots & \ldots & \ldots \\ \alpha b_1 + \beta c_1 & \ldots & \alpha b_n + \beta c_n \\ \ldots & \ldots & \ldots \\ a_{n1} & \ldots & a_{nn} \end{vmatrix},$$

16а

16. Відображення базису

ТЕОРЕМА (про відображення базису)

Нехай $\boldsymbol{e} = (\boldsymbol{e}_1\ \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n)$ — довільний базис простору R_1^n. Тоді, яким би не був набір елементів $\boldsymbol{a} = (\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n)$ лінійного простору R_2^n, існує відображення $\tilde{P}$, що переводить вектори $\boldsymbol{e}_i \in R_1^n$ у вектори $\boldsymbol{a}_i \in R_2^n$, де $i = 1, 2, \ldots, n$, тобто $\tilde{P}(\boldsymbol{e}_i) = \boldsymbol{a}_i$, $i = 1, 2, \ldots, n$. Це відображення лінійне і єдине.

Доведення. Розкладемо довільний елемент $\boldsymbol{x}$ із простору R_1^n за базисом $\boldsymbol{e}$:

$$\boldsymbol{x} = x_1\boldsymbol{e}_1 + x_2\boldsymbol{e}_2 + \ldots + x_n\boldsymbol{e}_n.$$

Побудуємо елемент $\boldsymbol{a} \in R_2^n$ за наступним правилом:

$$\boldsymbol{a} = x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2 + \ldots + x_n\boldsymbol{a}_n.$$

Введемо відображення $\tilde{P}$:

$$\boldsymbol{a} = \tilde{P}(\boldsymbol{x}).$$

Запишемо рівність у розгорнутій формі:

$$\tilde{P}(x_1\boldsymbol{e}_1 + x_2\boldsymbol{e}_2 + \ldots + x_n\boldsymbol{e}_n) = x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2 + \ldots + x_n\boldsymbol{a}_n,$$

або

$$x_1\tilde{P}(\boldsymbol{e}_1) + x_2\tilde{P}(\boldsymbol{e}_2) + \ldots + x_n\tilde{P}(\boldsymbol{e}_n) = x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2 + \ldots + x_n\boldsymbol{a}_n.$$

Перенесемо всі доданки вліво й згрупуємо їх:

$$x_1(\tilde{P}(\boldsymbol{e}_1) - \boldsymbol{a}_1) + x_2(\tilde{P}(\boldsymbol{e}_2) - \boldsymbol{a}_2) + \ldots + x_n(\tilde{P}(\boldsymbol{e}_n) - \boldsymbol{a}_n) = 0.$$

Рівність справедлива для довільних $x_1, x_2, \ldots, x_n$ за умови

$$\tilde{P}(\boldsymbol{e}_1) = \boldsymbol{a}_1,\ \tilde{P}(\boldsymbol{e}_2) = \boldsymbol{a}_2, \ldots, \tilde{P}(\boldsymbol{e}_n) = \boldsymbol{a}_n.$$

Лінійність. Скористаємося скороченою формою запису. Нехай $\boldsymbol{x} = \sum_{i=1}^{n} x_i\boldsymbol{e}_i$, $\boldsymbol{y} = \sum_{i=1}^{n} y_i\boldsymbol{e}_i$. Складемо лінійну комбінацію цих елементів і знайдемо її відображення:

$$\tilde{P}(\lambda\boldsymbol{x} + \mu\boldsymbol{y}) = \sum_{i=1}^{n}(\lambda x_i + \mu y_i)\boldsymbol{a}_i = \lambda\sum_{i=1}^{n} x_i\boldsymbol{a}_i + \mu\sum_{i=1}^{n} y_i\boldsymbol{a}_i = \lambda\tilde{P}(\boldsymbol{x}) + \mu\tilde{P}(\boldsymbol{y}).$$

Умова лінійності відображення виконана.

Однозначність. Припустимо існування ще одного відображення Q, такого, що

17а

17. Структура лінійного оператора

Дія лінійного оператора на вектор $\boldsymbol{x}$ зводиться до множення деякої матриці $P = (a_{ij})$ на матрицю-стовпець X, утворений з координат вектора $\boldsymbol{x}$. Матриця P називається **матрицею лінійного оператора** $\tilde{P}$ в базисі $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n$.

Будемо досліджувати ранг лінійного оператора $\tilde{P}$. Оператор $\tilde{P}$ — це частковий випадок лінійного відображення, тому має ранг, що дорівнює розмірності образу $\operatorname{im}\tilde{P}$ цього відображення, а отже, розмірності відповідного підпростору. Підпростір утворюється з векторів $\tilde{P}(\boldsymbol{e}_1)$, $\tilde{P}(\boldsymbol{e}_2), \ldots, \tilde{P}(\boldsymbol{e}_n)$ і їхніх лінійних комбінацій. Число лінійно незалежних векторів серед них становить розмірність підпростору. Координати векторів $\tilde{P}(\boldsymbol{e}_1), \tilde{P}(\boldsymbol{e}_2), \ldots, \tilde{P}(\boldsymbol{e}_n)$ утворюють стовпці матриці P. Тому число лінійно незалежних стовпців матриці P (їм відповідають лінійно незалежні вектори) і є рангом оператора $\tilde{P}$. Висновок: ранг матриці P лінійного оператора $\tilde{P}$ дорівнює рангу оператора.

У тому випадку, коли ранг r лінійного оператора дорівнює n ($r = n$, тобто матриця оператора не вироджена ($|P| \neq 0$), тільки нульовий вектор перетвориться оператором у нульовий вектор. Дійсно, матричне рівняння $Y = P \cdot X$ має єдиний розв'язок, що забезпечує взаємно однозначну відповідність між векторами $\boldsymbol{x}$ та $\boldsymbol{y}$, причому нульовому вектору відповідає нульовий вектор.

Якщо матриця лінійного оператора є виродженою ($r < n$), то деякі вектори, відмінні від нульового, такий оператор переводять у нульові вектори. Виникає дефект лінійного оператора. Відповідний підпростір ядра оператора перестає бути нульовим. За видом матриці оператора можна знайти ядро, а саме базис ядра і розмір дефекту.

ПРИКЛАД. Лінійний оператор $\tilde{P}$ заданий матрицею $P = \begin{pmatrix} 1 & -1 & 2 \\ -2 & 1 & 1 \\ -1 & 0 & 3 \end{pmatrix}$ в деякому базисі векторного простору R^3. Знайти базис ядра й розмір дефекту оператора.

Розв'язання. Нехай вектор $\boldsymbol{x} = (x_1, x_2, x_3)$ належить ядру лінійного оператора $\tilde{P}$, тобто оператор переводить вектор x у нульовий вектор $\tilde{P}(\boldsymbol{x}) = O$. Оскільки $\tilde{P}(\boldsymbol{x}) = P \cdot X$, запишемо матричне рівняння $P \cdot X = O$ у розгорнутому вигляді:

18а

18. Теорема про доповнення до базису

ТЕОРЕМА (про доповнення до базису)

Нехай вектори $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_k$ лінійного простору W розмірності n лінійно незалежні, причому $k < n$. Тоді в просторі W знайдуться вектори $\boldsymbol{a}_{k+1}, \boldsymbol{a}_{k+2}, \ldots, \boldsymbol{a}_n$, такі, що сукупність n векторів $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_k, \boldsymbol{a}_{k+1}, \boldsymbol{a}_{k+2}, \ldots, \boldsymbol{a}_n$ утворить базис цього простору.

Доведення. Нехай $\boldsymbol{x}$ — довільний вектор лінійного векторного простору W. Записати будь-який вектор $\boldsymbol{x}$ у вигляді лінійної комбінації векторів $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_k$ не можна, тому що в противному випадку сукупність векторів $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_k$ була б базисом. Однак, враховуючи умову $k < n$, це неможливо. Тому повинен існувати вектор $\boldsymbol{a}_{k+1} \in W$ такий, що доповнена система векторів $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_k, \boldsymbol{a}_{k+1}$ буде лінійно незалежною.

Якщо $k + 1 = n$, то ця система є базисом простору W.

Якщо $k + 1 < n$, варто повторити міркування з векторами $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_k, \boldsymbol{a}_{k+1}$.

Отже, будь-яка задана сукупність лінійно незалежних векторів може бути доповнена до базису векторного простору.

Перейдемо до методу доповнення k лінійно незалежних векторів до базису. Нехай дані вектори $\boldsymbol{a}_1 = (x_{11}, \ldots, x_{1n}), \ldots, \boldsymbol{a}_k = (x_{k1}, \ldots, x_{kn})$. Утворимо з координат векторів матрицю, розташовуючи для зручності координати векторів по рядках:

$$\begin{pmatrix} x_{11} & \ldots & x_{1n} \\ \ldots & \ldots & \ldots \\ x_{k1} & \ldots & x_{kn} \end{pmatrix}.$$

Оскільки вектори лінійно незалежні й $k < n$, елементарними перетвореннями рядків матриця зводиться до східчастого вигляду. Доповнимо отриману матрицю $n - k$ рядками вигляду $(0, \ldots, 1, \ldots, 0)$ так, щоб ранг нової матриці, наприклад, такої:

19а

19. Зв'язок рангу з числом незалежних рядків (стовпців) матриці

У матриці

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix}$$

введемо позначення рядків:

$$\boldsymbol{e}_1 = (a_{11} \quad a_{12} \quad \ldots \quad a_{1n})$$
$$\boldsymbol{e}_2 = (a_{21} \quad a_{22} \quad \ldots \quad a_{2n})$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$\boldsymbol{e}_m = (a_{m1} \quad a_{m2} \quad \ldots \quad a_{mn}).$$

Ці рядки $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_m$ є n-мірними і являють собою матриці розмірністю $1 \times n$.

Рядок $\boldsymbol{e} = (b_1 \quad b_2 \quad \ldots \quad b_n)$, що визначається рівністю

$$\boldsymbol{e} = \lambda_1\boldsymbol{e}_1 + \lambda_2\boldsymbol{e}_2 + \ldots + \lambda_m\boldsymbol{e}_m,$$

називається **лінійною комбінацією** рядків $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_m$, де $\lambda_1, \lambda_2, \ldots, \lambda_m$ — будь-які дійсні числа.

У розгорнутому матричному вигляді остання рівність виглядає так:

$$(b_1 \quad b_2 \quad \ldots \quad b_n) = \lambda_1 \cdot (a_{11} \quad a_{12} \quad \ldots \quad a_{1n}) + \lambda_2 \cdot (a_{21} \quad a_{22} \quad \ldots \quad a_{2n}) + \ldots + \lambda_m \cdot (a_{m1} \quad a_{m2} \quad \ldots \quad a_{mn}).$$

Для елементів рядка $\boldsymbol{e}$ маємо систему рівнянь

$$\begin{cases} b_1 = \lambda_1 a_{11} + \lambda_2 a_{21} + \ldots + \lambda_1 a_{m1}, \\ b_2 = \lambda_1 a_{12} + \lambda_2 a_{22} + \ldots + \lambda_m a_{m2}, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots, \\ b_n = \lambda_1 a_{1n} + \lambda_2 a_{2n} + \ldots + \lambda_m a_{mn}. \end{cases}$$

Рядки $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_m$ називаються **лінійно залежними**, якщо існують такі $\lambda_1, \lambda_2, \ldots, \lambda_m$, не рівні нулю одночасно, що лінійна комбінація цих рядків дорівнює нульовому рядку

$$\lambda_1 \boldsymbol{e}_1 + \lambda_2 \boldsymbol{e}_2 + \ldots + \lambda_m \boldsymbol{e}_m = \boldsymbol{o}.$$

Рядки $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_m$ називаються **лінійно незалежними**, якщо лінійна комбінація цих рядків дорівнює нульовому рядку тільки при $\lambda_1 = \lambda_2 = \ldots = \lambda_m = 0$.

ТЕОРЕМА (про лінійну комбінацію рядків матриці)

Якщо рядки матриці лінійно залежні, то один з них є лінійною комбінацією інших.

20а

20. Обернена матриця

Матриця A називається **невиродженою**, якщо її визначник відрізняється від нуля, тобто $|A| \neq 0$. У протилежному випадку вона називається **виродженою**.

Визначення. Матриця A^{-1} називається **оберненою** до квадратної матриці A, якщо виконується рівність

$$A^{-1} \cdot A = A \cdot A^{-1} = E.$$

ТЕОРЕМА (про існування оберненої матриці)

Обернена матриця A^{-1} існує і єдина тоді й тільки тоді, коли початкова матриця A *невироджена*.

Властивості обернених матриць

1) $(A^{-1})^{-1} = A$.

Помножимо обидві частини рівності ліворуч на A^{-1}:

$$A^{-1} \cdot (A^{-1})^{-1} = A^{-1} \cdot A.$$

Ліворуч стоїть добуток матриці A^{-1} на обернену до неї $(A^{-1})^{-1}$, що дорівнює одиничній матриці, праворуч — добуток оберненої матриці на дану, що також дорівнює одиничній матриці.

2) $(A^{-1})^T = (A^T)^{-1}$.

Помножимо обидві частини рівності ліворуч на A^T:

$$A^T \cdot (A^{-1})^T = A^T \cdot (A^T)^{-1}.$$

Далі скористаємося властивістю 4 транспонування матриці й перепишемо ліву частину співвідношення так: $(A^{-1} \cdot A)^{T.}$ Права частина рівності є добуток матриці A^T на обернену до неї. Маємо $E^T = E$. Звідси випливає $E = E$.

3) $(A^{-1})^m = (A^m)^{-1}$.

Помножимо рівність ліворуч на A^m:

$$A^m \cdot (A^{-1})^m = A^m \cdot (A^m)^{-1}.$$

Ліву частину рівності запишемо у вигляді добутку $2m$ співмножників:

$$A \cdot A \ldots \underbrace{A \cdot A^{-1}}_{E} \cdot A^{-1} \ldots A^{-1} = A^m \cdot (A^m)^{-1}.$$

Ліва частина рівності згортається до матриці E, права частина рівності є добутком матриці A^m на обернену до неї. Рівність стає тотожністю $E = E$.

4) $|A^{-1}| = \dfrac{1}{|A|}$.

21a

21. Спосіб побудови оберненої матриці

Нехай $E_1, E_2, \ldots, E_k$ — матриці елементарних перетворень.

Помножимо обидві частини рівності

$$E_k \cdot \ldots \cdot E_2 \cdot E_1 \cdot A = E \qquad (1)$$

на матрицю A^{-1} праворуч. Тоді

$$E_k \cdot \ldots \cdot E_2 \cdot E_1 \cdot A \cdot A^{-1} = E \cdot A^{-1}.$$

Після перетворень маємо

$$E_k \cdot \ldots \cdot E_2 \cdot E_1 = A^{-1}. \qquad (2)$$

Ця рівність лежить в основі способу побудови оберненої матриці. Нехай A — невироджена матриця n-го порядку. Утворимо нову матрицю, назвемо її розширеною: $(A|E)$.

Нехай одинична матриця E має також порядок n. Будемо послідовно здійснювати з розширеною матрицею такі елементарні перетворення, які рівносильні множенню цієї матриці ліворуч на матриці елементарних перетворень $E_1, E_2, \ldots, E_n$. Одержимо

$$(E_k \cdot \ldots \cdot E_2 \cdot E_1 \cdot A | E_k \cdot \ldots \cdot E_2 \cdot E_1 \cdot E).$$

Підставляючи (1) і (2) в отриману розширену матрицю, маємо

$$(E_k \cdot \ldots \cdot E_2 \cdot E_1 \cdot A | E_k \cdot \ldots \cdot E_2 \cdot E_1 \cdot E) = (E | A^{-1}).$$

Таким чином, якщо шляхом елементарних перетворень з розширеною матрицею ліворуч від риски одержати одиничну матрицю, то праворуч від риски утвориться обернена матриця.

Зауваження. Застосовуючи елементарні перетворення до розширеної матриці $(A|B)$, можна одержати матрицю $(E|A^{-1} \cdot B)$. Матриця $A^{-1} \cdot B$ широко використовується при розв'язуванні систем лінійних рівнянь. Для одержання цієї матриці варто проводити елементарні перетворення тільки з *рядками* розширеної матриці, тому що ці дії рівносильні множенню матриць елементарних перетворень *ліворуч* на розширену матрицю.

ПРИКЛАД. Знайти матрицю, обернену до матриці

$$A = \begin{pmatrix} 2 & 0 & 1 \\ 1 & -1 & 0 \\ 0 & 1 & 1 \end{pmatrix}.$$

Розв'язання. Утворимо розширену матрицю:

22a

22. Матриці елементарних перетворень

Типи матриць елементарних перетворень

Матрицями елементарних перетворень називаються матриці наступних трьох типів.

1-й тип. Матрицею елементарних перетворень 1-го типу називається будь-яка матриця, отримана з одиничної матриці перестановкою двох будь-яких рядків або стовпців. Наприклад, якщо в одиничній матриці п'ятого порядку

$$E = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

поміняти місцями другий та третій рядки, виходить матриця елементарних перетворень 1-го типу:

$$I_{23} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

У матриці I_{23} всі елементи поза головною діагоналлю дорівнюють нулю, за винятком тих, які стоять у позиціях (2, 3) і (3, 2).

2-й тип. Матрицею елементарних перетворень 2-го типу називається будь-яка матриця, отримана з одиничної заміною діагонального елемента на будь-яке дійсне число, що не дорівнює нулю. Наприклад, матрицею елементарних перетворень 2-го типу є матриця

$$II_{44} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & a & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

у якої в позиції (4, 4) знаходиться число $a \neq 0$.

3-й тип. Матрицею елементарних перетворень 3-го типу називається будь-яка матриця, що відрізняється від одиничної наявністю одного недіагонального елемента, що не дорівнює нулю. Наприклад,

23a

23. Ранг матриці

У матриці A розміром $m \times n$ викреслюванням яких-небудь рядків або стовпців можна утворити квадратну матрицю k-го порядку $(k \times k)$. Визначник M_k такої матриці називається мінором k-го порядку. У матриці розміру $m \times n$ є мінори першого порядку, другого порядку й так далі до k-го порядку, де $k = \min(m, n)$. Наприклад, у матриці $\underset{5\times 3}{A}$ є мінори першого, другого й третього порядків.

Рангом матриці $\underset{m\times n}{A}$ називається найвищий порядок відмінних від нуля мінорів цієї матриці.

Позначення: $\operatorname{rang} A$ або $r(A)$.

Властивості рангу:

1) Вважається, що ранг нульової матриці дорівнює нулю.

2) $r(A) \leq \min(m, n)$.

3) $r(A) = n$ у матриці n-го порядку тоді й тільки тоді, коли $|A| \neq 0$.

До трьох типів елементарних перетворень матриці додамо ще два:

4-й тип. Відкидання нульового рядка або стовпця.

5-й тип. Транспонування матриці.

ТЕОРЕМА (про ранг матриці при елементарних перетвореннях)

Елементарні перетворення не змінюють рангу матриці.

Доведення.

Елементарні перетворення 1-го типу змінюють рядки або стовпці в матриці. У цьому випадку визначник матриці змінює знак, але не може стати нулем.

Елементарні перетворення 2-го типу множать рядок або стовпець на число, що не дорівнює нулю. Але тоді визначник матриці помножиться на це число, тобто не може стати нулем.

Елементарні перетворення 3-го типу додають до i-го рядка матриці A її j-й рядок, що не змінює величини визначника.

Елементарні перетворення 4-го типу дозволяють відкинути всі мінори k-го порядку, що дорівнюють нулю, й перейти до розгляду мінорів $k-1$ порядку. На величину рангу це, вочевидь, не вплине.

24a

24. Теорема Кронекера—Капеллі

ТЕОРЕМА (про сумісність системи рівнянь)

Лінійна система сумісна тоді й тільки тоді, коли ранг матриці системи дорівнює рангу розширеної матриці цієї системи.

Доведення.

Система m лінійних рівнянь із n змінними має вигляд

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots, \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m. \end{cases} \qquad (1)$$

У системі лінійних рівнянь (1) введемо позначення матриці коефіцієнтів

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix}$$

і розширеної матриці

$$A^P = \left(\begin{array}{cccc|c} a_{11} & a_{12} & \ldots & a_{1n} & b_1 \\ a_{21} & a_{22} & \ldots & a_{2n} & b_2 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & \ldots & a_{mn} & b_m \end{array}\right),$$

а також позначення стовпців, що складаються з коефіцієнтів при невідомих

$$\begin{pmatrix} a_{11} \\ a_{21} \\ \ldots \\ a_{m1} \end{pmatrix} = \boldsymbol{a}_1, \begin{pmatrix} a_{12} \\ a_{22} \\ \ldots \\ a_{m2} \end{pmatrix} = \boldsymbol{a}_2, \begin{pmatrix} a_{1n} \\ a_{2n} \\ \ldots \\ a_{mn} \end{pmatrix} = \boldsymbol{a}_n, \begin{pmatrix} b_1 \\ b_2 \\ \ldots \\ b_m \end{pmatrix} = \boldsymbol{b}.$$

У нових позначеннях система виглядає так:

$$x_1 \boldsymbol{a}_1 + x_2 \boldsymbol{a}_2 + \ldots + x_n \boldsymbol{a}_n = \boldsymbol{b}.$$

Необхідність. Нехай система сумісна, тобто існують $\lambda_1, \lambda_2, \ldots, \lambda_n$ такі, що виконується рівність

$$\lambda_1 \boldsymbol{a}_1 + \lambda_2 \boldsymbol{a}_2 + \ldots + \lambda_n \boldsymbol{a}_n = \boldsymbol{b}.$$

25a

25. Вектори на площині

Означення. *Вектором* на площині або у просторі називається спрямований відрізок, що має початкову й кінцеву точки. Звичайно вектор позначається малою літерою з рискою $\overline{a}$ або ж двома великими літерами, що позначають початок і кінець вектора $\overline{AB}$.

Довжиною, або *модулем* $|\boldsymbol{a}|$ вектора називається число, що дорівнює довжині спрямованого відрізка. Вектор, довжина якого дорівнює одиниці, називається *одиничним* вектором, або *ортом*. Вектор, початок і кінець якого співпадають, називається *нульовим* і позначається $\boldsymbol{o}$. Вектори, що лежать на паралельних прямих, називаються *колінеарними*.

Лінійні операції над векторами

1. *Сумою* векторів $\boldsymbol{a}$ й $\boldsymbol{b}$ називається вектор $\boldsymbol{c} = \boldsymbol{a} + \boldsymbol{b}$, початок якого співпадає з початком вектора $\boldsymbol{a}$, а кінець — з кінцем вектора $\boldsymbol{b}$ (рис. 1). Цей спосіб побудови суми векторів називається *правилом трикутника*.

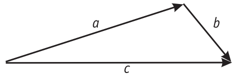

Рис. 1

Якщо на векторах $\boldsymbol{a}$ і $\boldsymbol{b}$, як на сторонах, побудувати паралелограм, то більша діагональ паралелограма буде сумою векторів $\boldsymbol{a}$ і $\boldsymbol{b}$. Цей спосіб побудови суми векторів називається *правилом паралелограма* (рис. 2).

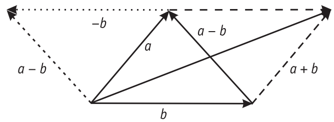

Рис. 2

2. *Добутком* числа λ на вектор $\boldsymbol{a}$ називається вектор $\boldsymbol{b}$, що має довжину $|\boldsymbol{b}| = |\lambda||\boldsymbol{a}|$. Напрямок вектора $\boldsymbol{b}$ співпадає з напрямком $\boldsymbol{a}$, якщо $\lambda > 0$, і протилежний за напрямком, якщо $\lambda < 0$.

26a

26. Лінійні векторні простори

Означення. Множина W елементів $x, y, z, \ldots$ називається *лінійним простором*, якщо за деяким правилом:

1. Будь-яким двом елементам x і y з W ставиться у відповідність елемент із W, що позначається $x + y$ та називається сумою елементів x та y.

2. Будь-якому елементу x з W і кожному числу λ ставиться у відповідність елемент із W, що позначається $\lambda \cdot x$ та називається добутком числа λ на елемент x, причому справедливі наступні аксіоми:

1) $x + y = y + x$;
2) $(x + y) + z = x + (y + z)$;
3) $\lambda \cdot (x + y) = \lambda \cdot x + \lambda \cdot y$;
4) $(\lambda + \mu) \cdot x = \lambda \cdot x + \lambda \cdot x$;
5) $\lambda \cdot (\mu \cdot x) = (\lambda \cdot \mu) \cdot x$;
6) $1 \cdot x = x$;
7) існує нульовий елемент 0 такий, що $x + 0 = x$ для будь-якого $x \in W$;
8) для кожного елемента x існує протилежний елемент $-x$, такий, що $x + (-x) = 0$.

Елементи будь-якої природи, що задовольняють двом правилам і восьми аксіомам, за визначенням, утворюють лінійний простір. Наприклад, сукупність будь-яких матриць розміру $m \times n$ утворюють лінійний простір, оскільки для них виконані обидва правила й всі аксіоми. Легко перевірити, що сукупність геометричних векторів, наприклад тривимірного простору, також є лінійним простором.

Лінійний простір називається *порожнім*, якщо він складається з нульового елемента.

27a

27. Вектор в *n*-мірному просторі

Означення. *n-мірним вектором* називається математичний об'єкт, що складається з упорядкованої сукупності дійсних чисел, які називаються *компонентами* вектора, і записується у вигляді $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$. Назва «$n$-мірний вектор» пов'язана з тим, що при $n = 2$ або $n = 3$ сукупність чисел можна інтерпретувати як сукупність координат вектора на площині або в просторі.

Два n-мірних вектори $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$ і $\boldsymbol{y} = (y_1, y_2, \ldots, y_n)$ називаються *рівними*, якщо рівні всі компоненти векторів, тобто $x_i = y_j$, де $i = 1, 2, \ldots, n$.

Нехай для n-мірних векторів виконані правила додавання й множення на число, що задовольняють аксіомам лінійного простору. Тоді множина всіх n-мірних векторів називається *лінійним векторним простором* і позначається W.

Декларуючи правила додавання й множення для n-мірних векторів, ми маємо пояснити, як варто робити ці дії над сукупностями n дійсних чисел. Інакше кажучи, введемо операції над n-мірними векторами.

Сумою двох векторів $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$ і $\boldsymbol{y} = (y_1, y_2, \ldots, y_n)$ назвемо вектор $\boldsymbol{z} = \boldsymbol{x} + \boldsymbol{y}$, такий, що

$$\boldsymbol{z} = (z_1, z_2, \ldots, z_n) = (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n).$$

Добутком дійсного числа λ на вектор $\boldsymbol{x}$ назвемо вектор $\boldsymbol{y} = \lambda \cdot \boldsymbol{x}$, такий, що

$$\boldsymbol{y} = (y_1, y_2, \ldots, y_n) = (\lambda x_1, \lambda x_2, \ldots, \lambda x_n).$$

Зауваження. Введені відповідно до визначення операції над n-мірними векторами аналогічні операціям над прямокутними матрицями. Тому n-мірні вектори можна розглядати як матриці-рядки $X = (x_1, x_2, \ldots, x_n)$ або як матриці-стовпці $X = \begin{pmatrix} x_1 \\ x_2 \\ \ldots \\ x_n \end{pmatrix}$ та здійснювати над векторами матричні операції.

28a

28. Лінійна залежність і незалежність векторів

Нехай кожний з векторів у наборі $\boldsymbol{x}, \boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_m$ є n-мірний вектор.

Означення. Вектор $\boldsymbol{x}$ називається *лінійною комбінацією* векторів $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_m$, якщо знайдуться такі дійсні числа $\lambda_1, \lambda_2, \ldots, \lambda_m$, що

$$\boldsymbol{x} = \lambda_1\boldsymbol{a}_1 + \lambda_2\boldsymbol{a}_2 + \ldots + \lambda_m\boldsymbol{a}_m.$$

Вектори $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_m$ називаються *лінійно залежними*, якщо існують такі числа $\lambda_1, \lambda_2, \ldots, \lambda_m$, які не всі одночасно дорівнюють нулю, що

$$\lambda_1\boldsymbol{a}_1 + \lambda_2\boldsymbol{a}_2 + \ldots + \lambda_m\boldsymbol{a}_m = \boldsymbol{o}. \qquad (1)$$

Якщо рівність (1) виконується тільки при $\lambda_1 = \lambda_2 = \ldots = \lambda_m = 0$, то вектори називаються *лінійно незалежними*.

Властивості лінійної залежності векторів

1) Якщо серед декількох векторів (набору векторів) один з них є лінійною комбінацією частини інших, то весь набір векторів лінійно залежний.

Нехай є вектори $\boldsymbol{x}, \boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_m$, причому вектор $\boldsymbol{x} = \lambda_1\boldsymbol{a}_1 + \lambda_2\boldsymbol{a}_2 + \ldots + \lambda_k\boldsymbol{a}_k$, де $k \le m$. Перенесемо всі члени в одну частину й доповнимо доданками $0 \cdot \boldsymbol{a}_{k+1}, \ldots, 0 \cdot \boldsymbol{a}_m$. Одержимо лінійну комбінацію

$$\lambda_1\boldsymbol{a}_1 + \lambda_2\boldsymbol{a}_2 + \ldots + \lambda_k\boldsymbol{a}_k + 0 \cdot \boldsymbol{a}_{k+1} + \ldots + 0 \cdot \boldsymbol{a}_m + (-1)\boldsymbol{x} = \boldsymbol{o},$$

у якій знайшлися $\lambda_1, \lambda_2, \ldots, \lambda_m$, такі, що не всі одночасно дорівнюють нулю. Виходить, що вектори $\boldsymbol{x}, \boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_m$ лінійно залежні.

2) Якщо серед набору векторів $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_m$ є нульовий вектор, то весь набір векторів лінійно залежний.

Нехай, наприклад, нульовим є вектор $\boldsymbol{a}_1$. Тоді рівність (1) залишиться справедливою при $\lambda_1 = 1, \lambda_2 = \lambda_3 = \ldots = \lambda_m = 0$.

3) Якщо вектори $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_m$ лінійно незалежні й існує вектор $\boldsymbol{x}$, що є лінійною комбінацією $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_m$, тобто $\boldsymbol{x} = \lambda_1\boldsymbol{a}_1 + \lambda_2\boldsymbol{a}_2 + \ldots + \lambda_m\boldsymbol{a}_m$, то коефіцієнти $\lambda_1, \lambda_2, \ldots, \lambda_m$ визначаються за вектором $\boldsymbol{x}$ єдиним способом.

Нехай вектор $\boldsymbol{x}$ можна представити як дві лінійні комбінації з різними коефіцієнтами:

$$\boldsymbol{x} = \lambda_1\boldsymbol{a}_1 + \lambda_2\boldsymbol{a}_2 + \ldots + \lambda_k\boldsymbol{a}_k \text{ та } \boldsymbol{x} = \mu_1\boldsymbol{a}_1 + \mu_2\boldsymbol{a}_2 + \ldots + \mu_k\boldsymbol{a}_k.$$

Тоді

$$\boldsymbol{x} = \lambda_1\boldsymbol{a}_1 + \lambda_2\boldsymbol{a}_2 + \ldots + \lambda_k\boldsymbol{a}_k = \mu_1\boldsymbol{a}_1 + \mu_2\boldsymbol{a}_2 + \ldots + \mu_k\boldsymbol{a}_k,$$

звідки

$$(\lambda_1 - \mu_1)\boldsymbol{a}_1 + (\lambda_2 - \mu_2)\boldsymbol{a}_2 + \ldots + (\lambda_k - \mu_k)\boldsymbol{a}_k = 0.$$

29а

29. Розмірність. Базис векторного простору

Означення. Векторний простір називається n-мірним, якщо серед множини його векторів знайдуться n лінійно незалежних векторів, а будь-які $n + 1$ векторів вже виявляться залежними. Число n називається *розмірністю* векторного простору.

Наприклад, серед нескінченної множини векторів, розташованих в одній площині, будь-які два неколінеарних вектори є лінійно незалежними. Виберемо які-небудь два неколінеарних вектори $\boldsymbol{a}_1 = (x_1, x_2)$ і $\boldsymbol{a}_2 = (y_1, y_2)$. Додавання третього вектора $\boldsymbol{a}_3 = (z_1, z_2)$ до обраних двох робить їх лінійно залежними. Дійсно, система рівнянь, отримана з векторної рівності $\lambda_1\boldsymbol{a}_1 + \lambda_2\boldsymbol{a}_2 + \lambda_3\boldsymbol{a}_3 = \boldsymbol{o}$, буде мати вигляд

$$\lambda_1\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \lambda_2\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} + \lambda_3\begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Матриця коефіцієнтів $\begin{pmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \end{pmatrix}$ має ранг, що дорівнює двом, звідки випливає, що один із трьох стовпців є лінійною комбінацією двох інших. Отже, розмірність такого лінійного векторного простору дорівнює двом.

Означення. Упорядкована сукупність n лінійно незалежних векторів n-мірного векторного простору називається *базисом* цього простору.

Обрані нами в розглянутому вище прикладі в певному порядку два неколінеарних вектори становлять базис у двовимірному просторі. Якщо вектори поміняти місцями, вони також утворять базис цього простору, але інший. Якщо вибрати два інших неколінеарних вектори в певному порядку, з них можна побудувати свій базис.

Якщо у векторному просторі визначений базис, інші вектори можуть бути виражені через цей базис.

30а

30. Перехід до нового базису

Матриця переходу до нового базису

Нехай $B = (\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n)$ і $B' = (\boldsymbol{e}_1', \boldsymbol{e}_2', \ldots, \boldsymbol{e}_n')$ — старий та новий базиси лінійного n-мірного векторного простору W. Кожний з векторів нового базису можна виразити через вектори старого базису:

$$\begin{cases} \boldsymbol{e}_1' = a_{11}\boldsymbol{e}_1 + a_{21}\boldsymbol{e}_2 + \ldots + a_{n1}\boldsymbol{e}_n, \\ \boldsymbol{e}_2' = a_{12}\boldsymbol{e}_1 + a_{22}\boldsymbol{e}_2 + \ldots + a_{n2}\boldsymbol{e}_n, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots, \\ \boldsymbol{e}_n' = a_{1n}\boldsymbol{e}_1 + a_{2n}\boldsymbol{e}_2 +,,, + a_{nn}\boldsymbol{e}_n, \end{cases}$$

а також представити в матричній формі:

$$(\boldsymbol{e}_1', \boldsymbol{e}_2', \ldots, \boldsymbol{e}_n') = (\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n)\begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{pmatrix} \qquad (1)$$

або в скороченій матричній формі:

$$B' = B \cdot T.$$

Матриця

$$T = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{pmatrix}$$

називається матрицею переходу від старого базису $\boldsymbol{e}$ до нового базису $\boldsymbol{e}'$.

Координати розкладання векторів нового базису за старим базисом розташовуються в матриці переходу по стовпцях.

Властивості матриці переходу

1) Матриця переходу є невиродженою, тобто $|T| \neq 0$. Дійсно, з рівності $|T| = 0$ випливає, що один зі стовпців матриці T є лінійною комбінацією інших стовпців. Тоді один з векторів є лінійною комбінацією інших векторів цього базису, що неможливо.

2) Матриця переходу від нового базису до старого має вигляд T^{-1}. Помноживши рівність (1) на T^{-1} праворуч, одержимо

31а

31. Теорема про множення матриці на матриці елементарних перетворень

ТЕОРЕМА.

Будь-яка невироджена матриця A шляхом множення на матрицю елементарних перетворень $E_1, E_2, \ldots, E_k$ може бути зведена до одиничної, тобто знайдуться такі матриці елементарних перетворень $E_1, E_2, \ldots, E_k$, послідовне множення яких на матрицю A ліворуч перетворить дану матрицю A в одиничну:

$$E_k \cdot \ldots \cdot E_2 \cdot E_1 \cdot A = E.$$

Доведення. Нехай матриця A невироджена. Приведемо матрицю A за допомогою елементарних перетворень до матриці трикутного вигляду. Оскільки матриця A невироджена, вона ненульова. Знайдемо в 1-му стовпці ненульовий елемент і, міняючи рядки місцями, поставимо цей елемент у позицію (1, 1), якщо раніше там стояв нульовий елемент. Отже, $a_{11} \neq 0$. Додамо до 2-го рядка матриці 1-й рядок, попередньо помножений на $-a_{21}/a_{11}$. У позиції (2, 1) з'являється нуль. Додамо до 3-го рядка матриці 1-й рядок, попередньо помножений на $-a_{31}/a_{11}$. Тоді в позиції (3, 1) також з'явиться нуль. Здійснивши ці елементарні перетворення $(n - 1)$ раз, одержимо матрицю

$$A_1 = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ 0 & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ 0 & a_{n2} & \ldots & a_{nn} \end{pmatrix}.$$

У подальших перетвореннях 1-й рядок не бере участі. Знайдемо в 2-му стовпці ненульовий елемент і, міняючи рядки місцями, поставимо цей елемент у позицію (2, 2), якщо раніше там стояв нульовий елемент. Маємо $a'_{22} \neq 0$. Додамо до 3-го рядка матриці 2-й рядок, попередньо помножений на $-a_{32}/a_{22}$. У позиції (3, 2) з'являється нуль. Продовживши ці елементарні перетворення $(n - 1)$ разів, одержимо матрицю

$$A_2 = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ 0 & a_{22}' & \ldots & a_{2n}' \\ 0 & 0 & \ldots & a_{3n}' \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & a_{nn}' \end{pmatrix}.$$

Тепер у подальших перетвореннях уже не беруть участі 1-й та 2-й рядки. Продовживши цей процес (зробивши t раз елементарні перетворення), прийдемо до трикутної матриці

32а

32. Лінійні підпростори

Із множини векторів лінійного простору W виберемо деяку сукупність векторів і позначимо її V. Нехай для будь-яких векторів $\boldsymbol{x}$ і $\boldsymbol{y}$ з V і будь-якого числа $\lambda \in R$ виконуються наступні умови:

1) $\boldsymbol{x} + \boldsymbol{y} \in V$; 2) $\lambda \cdot \boldsymbol{x} \in V$.

Тоді множина векторів V називається **лінійним підпростором** простору W.

Приклади лінійних підпросторів:

1. Кожний лінійний простір володіє двома підпросторами: нульовим підпростором і самим простором. Ці підпростори називають **тривіальними**.

2. Лінійний простір W_1 векторів на прямій, що проходить через початок координат, має два тривіальних підпростори.

3. Лінійний простір W_2 векторів на площині (рис. 1) має, окрім двох тривіальних підпросторів, нескінченну множину підпросторів $V_1, V_2, \ldots$. Кожен з них складається з векторів, які лежать на прямій, що проходить через початок координат (вважається, що всі вектори відкладені від початку координат).

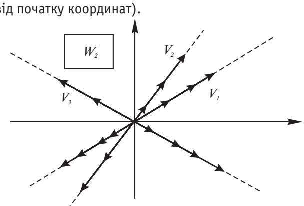

Рис. 1

4. У геометричному просторі W_3 векторів простору кожна пряма й кожна площина, що проходять через початок координат, визначають лінійний підпростір.

33a

33. Фундаментальний набір розв'язків однорідної системи лінійних рівнянь

Розв'язки $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_p$ однорідної системи називаються *лінійно незалежними*, якщо лінійна комбінація цих розв'язків $\lambda_1\boldsymbol{e}_1 + \lambda_2\boldsymbol{e}_2 + \ldots + \lambda_p\boldsymbol{e}_p$ дорівнює нульовому стовпцю тільки за умови $\lambda_1 = \lambda_2 = \ldots = \lambda_m = 0$. Побудуємо матрицю розв'язків, розташувавши матриці-стовпці розв'язків по стовпцях нової матриці. Відповідно до теореми про ранг матриці, ранг нової матриці буде дорівнювати числу стовпців нової матриці, тобто числу лінійно незалежних розв'язків системи.

Означення. Сукупність лінійно незалежних розв'язків $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_p$ однорідної системи рівнянь називається *фундаментальною,* якщо загальний розв'язок системи є лінійною комбінацією розв'язків $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_p$.

ТЕОРЕМА (про фундаментальні розв'язки однорідної системи)

Якщо ранг r матриці коефіцієнтів при змінних однорідної системи рівнянь менше числа змінних n, то:

1) існує сукупність лінійно незалежних розв'язків системи;

2) число лінійно незалежних розв'язків дорівнює $n - r$;

3) будь-який розв'язок системи можна представити у вигляді сукупності цих незалежних розв'язків, тобто у вигляді лінійної комбінації фундаментального набору розв'язків.

Зауваження 1. Як базисний можна вибрати будь-який набір з r змінних за умови, що визначник матриці з коефіцієнтів при них відмінний від нуля. Кількість способів вибору таких r змінних з їхнього загального числа n не перевищує величини

$$C_n^r = \frac{n!}{r!(n-r)!}.$$

Звідси випливає, що число фундаментальних наборів розв'язків (ФНР) обмежене.

Зауваження 2. Для знаходження множини розв'язків однорідної системи достатньо знайти який-небудь ФНР системи і скласти його лінійну комбінацію.

Зауваження 3. Будь-яка однорідна система рівнянь, що має ненульові розв'язки, має ФНР.

34a

34. Ортонормований базис

Система векторів $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n$ називається *ортогональною*, якщо скалярний добуток $(\boldsymbol{e}_i, \boldsymbol{e}_j) = 0$ при $i \neq j$, і *нормованою*, якщо $|\boldsymbol{e}_i| = 1$ для всіх $i = 1, 2, \ldots, n$. Якщо вектори системи ортогональні та нормовані, вони називаються *ортонормованими* .

Зауваження. Щоб нормувати ненульовий вектор, необхідно поділити його на норму. Нехай заданий вектор $\boldsymbol{x} = (1, -1, 2, 0)$. Його норма $|\boldsymbol{x}| = \sqrt{1^2 + (-1)^2 + 2^2} = \sqrt{6}$. Нормований вектор має вигляд $\boldsymbol{e} = \frac{\boldsymbol{x}}{|\boldsymbol{x}|} = \left(\frac{1}{\sqrt{6}}, \frac{-1}{\sqrt{6}}, \frac{2}{\sqrt{6}}, 0\right)$. Його довжина $|\boldsymbol{e}| = 1$.

ТЕОРЕМА (про незалежність ортонормованої системи векторів)

Ортонормована система векторів лінійно незалежна.

Доведення. Доведемо, що ортогональні й нормовані вектори $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n$ лінійно незалежні, тобто доведемо, що рівність

$$\lambda_1\boldsymbol{e}_1 + \lambda_2\boldsymbol{e}_2 + \ldots + \lambda_n\boldsymbol{e}_n = \boldsymbol{o}$$

справедлива лише при $\lambda_1 = \lambda_2 = \ldots = \lambda_n = 0$. Помножимо обидві частини рівності скалярно на вектор $\boldsymbol{e}_1$, маємо

$$\lambda_1(\boldsymbol{e}_1, \boldsymbol{e}_1) + \lambda_2(\boldsymbol{e}_1, \boldsymbol{e}_2) + \ldots + \lambda_n(\boldsymbol{e}_1, \boldsymbol{e}_n) = (\boldsymbol{e}_1, \boldsymbol{o})$$

або $\lambda_1 \cdot 1 = \lambda_2 \cdot 0 = \ldots = \lambda_n \cdot 0 = 0$.

Звідси випливає $\lambda_1 = 0$. Якщо помножити послідовно рівність скалярно на $\boldsymbol{e}_2, \boldsymbol{e}_3, \ldots, \boldsymbol{e}_n$, будемо мати $\lambda_2 = \ldots = \lambda_n = 0$.

ТЕОРЕМА (про існування ортобазису)

У будь-якому n-вимірному евклідовому просторі існує ортонормований базис.

Доведення. Доведення теореми — це алгоритм послідовної побудови ортонормованого базису по заданому базису $\boldsymbol{f}_1, \boldsymbol{f}_2, \ldots, \boldsymbol{f}_n$, що називається *методом ортогоналізації.*

Покладемо вектор $\boldsymbol{g}_1 = \boldsymbol{f}_1$ і нормуємо його: $\boldsymbol{e}_1 = \frac{\boldsymbol{g}_1}{|\boldsymbol{g}_1|}$, одержуючи перший вектор ортонормованого базису. Побудуємо вектор

$$\boldsymbol{g}_2 = \boldsymbol{f}_2 - \alpha \cdot \boldsymbol{e}_1 \qquad (1)$$

35a

35. Ортогональне доповнення

Нехай заданий евклідовий простір W і нехай $V^=$ — деякий лінійний підпростір евклідового простору W.

Означення. Сукупність $V^{\perp}$ векторів $\boldsymbol{y}$ простору W, що мають властивість

$$(\boldsymbol{y}, \boldsymbol{x}) = 0,$$

де $\boldsymbol{x}$ — довільний вектор з $V^=$, називається *ортогональним доповненням* до підпростору $V^=$.

Властивості ортогонального доповнення

1) Ортогональне доповнення $V^{\perp}$ є лінійним підпростором евклідового простору W.

Нехай вектори $\boldsymbol{y}_1$ і $\boldsymbol{y}_2$ належать $V^{\perp}$, вектор $\boldsymbol{x}$ — довільний вектор з $V^=$. Тоді $(\boldsymbol{y}_1, \boldsymbol{x}) = (\boldsymbol{y}_2, \boldsymbol{x}) = 0$. Додавши скалярні добутки, одержимо

$$(\boldsymbol{y}_1, \boldsymbol{x}) = (\boldsymbol{y}_2, \boldsymbol{x}) = (\boldsymbol{y}_1 + \boldsymbol{y}_2, \boldsymbol{x}) = 0.$$

Отже, $\boldsymbol{y}_1 + \boldsymbol{y}_2 \in V^{\perp}$.

Якщо $(\boldsymbol{y}, \boldsymbol{x}) = 0$, то й $(\alpha \cdot \boldsymbol{y}, \boldsymbol{x}) = 0$, де $\alpha \in R$, тобто з $\boldsymbol{y} \in V^{\perp}$ випливає $\alpha \cdot \boldsymbol{y} \in V^{\perp}$.

2) Лінійний простір W є прямою сумою підпросторів $V^=$ і $V^{\perp}$, тобто не містить перетинів $V^=$ і $V^{\perp}$.

Нехай $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_k$ — ортонормований базис $V^=$: $\boldsymbol{e}_l, \boldsymbol{e}_{l+1}, \ldots, \boldsymbol{e}_p$ — ортонормований базис $V^{\perp}$. Серед системи векторів $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_k, \boldsymbol{e}_l, \boldsymbol{e}_{l+1}, \ldots, \boldsymbol{e}_p$ немає однакових, і ця система ортонормована, отже, лінійно незалежна. Доведемо, що вона утворює базис n-мірного евклідового простору. Припустимо, що це не так, що число $(1+2+\ldots+k)+(l+(l+1)+\ldots+p)<n$. Тоді існує вектор $\boldsymbol{x}$ простору, такий, що система векторів $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_k, \boldsymbol{e}_l, \boldsymbol{e}_{l+1}, \ldots, \boldsymbol{e}_p, \boldsymbol{x}$ лінійно незалежна. Застосуємо до цієї системи процес ортогоналізації. Побудований на основі $\boldsymbol{x}$ вектор $\boldsymbol{e}_{p+1}$ буде ортогональним до $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_k$, виходить, $\boldsymbol{e}_{p+1} \in V^{\perp}$. З іншого боку, вектор $\boldsymbol{e}_{p+1}$ ортогональний до $\boldsymbol{e}_l, \boldsymbol{e}_{l+1}, \ldots, \boldsymbol{e}_p$, виходить, $\boldsymbol{e}_{p+1} \perp V^=$. Таку властивість має тільки нульовий вектор. Тому $\boldsymbol{e}_{p+1} = \boldsymbol{o}$. Звідси випливає лінійна залежність $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_k, \boldsymbol{e}_l, \boldsymbol{e}_{l+1}, \ldots, \boldsymbol{e}_p, \boldsymbol{x}$, що суперечить припущенню. Таким чином, система векторів $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_k, \boldsymbol{e}_l, \boldsymbol{e}_{l+1}, \ldots, \boldsymbol{e}_p$ є базисом n-мірного евклідового простору, тобто $1+2+\ldots+k+l+(l+1)+\ldots+p=n$.

36a

36. Евклідові простори

Означення. Лінійний векторний простір W називається *евклідовим*, якщо будь-яким двом векторам $\boldsymbol{x}$ і $\boldsymbol{y}$ з W ставиться у відповідність число, що позначається як $(\boldsymbol{x}, \boldsymbol{y})$, причому виконуються наступні умови:

1) $(\boldsymbol{x}, \boldsymbol{y}) = (\boldsymbol{y}, \boldsymbol{x})$;

2) $(\boldsymbol{x} + \boldsymbol{y}, \boldsymbol{z}) = (\boldsymbol{x}, \boldsymbol{z}) + (\boldsymbol{y}, \boldsymbol{z})$;

3) $(\lambda \cdot \boldsymbol{x}, \boldsymbol{y}) = \lambda \cdot (\boldsymbol{x}, \boldsymbol{y})$, де $\lambda \in R$;

4) $(\boldsymbol{x}, \boldsymbol{x}) > 0$, якщо $\boldsymbol{x}$ — ненульовий вектор; $(\boldsymbol{x}, \boldsymbol{x}) = 0$, якщо $\boldsymbol{x}$ — нульовий вектор.

Евклідовим лінійний простір названий на честь давньогрецького математика Евкліда, що створив в III ст. до н. е. евклідову геометрію.

Означення. Число $(\boldsymbol{x}, \boldsymbol{y})$ називається *скалярним добутком* векторів $\boldsymbol{x}$ та $\boldsymbol{y}$. Оскільки вектори задаються набором своїх координат, треба визначити, які дії варто зробити з координатами векторів, щоб одержати число $(\boldsymbol{x}, \boldsymbol{y})$. Визначимо скалярний добуток векторів $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$ і $\boldsymbol{y} = (y_1, y_2, \ldots, y_n)$ формулою

$$(\boldsymbol{x}, \boldsymbol{y}) = x_1 y_1 + x_2 y_2 + \ldots + x_n y_n.$$

Заданий в такий спосіб скалярний добуток n-мірних векторів для випадку $n = 2$ або $n = 3$ перетворюється в розглянутий раніше скалярний добуток геометричних векторів. Тепер з'являється можливість визначити довжини й кути в просторі.

Зауваження. Очевидно, що з рівності $\boldsymbol{x} = \boldsymbol{y}$ випливає рівність $(\boldsymbol{x}, \boldsymbol{z}) = (\boldsymbol{y}, \boldsymbol{z})$. Перехід від рівності векторів до рівності скалярних добутків назвемо *скалярним множенням*.

Означення. *Довжиною (нормою)* вектора $\boldsymbol{x}$ в евклідовому просторі називається величина

$$|\boldsymbol{x}| = \sqrt{(\boldsymbol{x}, \boldsymbol{x})} = \sqrt{x_1^2 + x_2^2 + \ldots + x_n^2}.$$

У двох- або тривимірному евклідовому просторі довжина вектора має чіткий геометричний зміст. Однак у чотирьох-, п'яти- або n-мірному евклідовому просторі зміст довжини вектора втрачається. Тому замість довжини вектора часто вводять поняття норми вектора.

37a

37. Диференціювання векторної функції скалярного аргументу

Припустимо, що кожному значенню скалярного аргументу x ставиться у відповідність n функцій $y_1(x), y_2(x), \ldots, y_n(x)$. Набір цих функцій покладемо координатами вектора $\boldsymbol{y}$:

$$\boldsymbol{y}(x) = (y_1(x), y_2(x), \ldots, y_n(x))$$

і будемо вважати, що *задано векторну функцію (вектор-функція)* $\boldsymbol{y}$ *скалярного аргументу* x.

Нехай для векторної функції $\boldsymbol{y}(x)$ існує границя:

$$\lim_{\Delta x \to 0} \frac{\boldsymbol{y}(x_0 + \Delta x) - \boldsymbol{y}(x_0)}{\Delta x} = \lim_{\Delta x \to 0} \frac{\Delta \boldsymbol{y}}{\Delta x}.$$

Вона називається похідною від $\boldsymbol{y}(x)$ по змінний x у точці x_0 і позначається $\frac{d\boldsymbol{y}}{dx}$. Для координат вектора $\boldsymbol{y}(x)$ це означає, за визначенням, що

$$\frac{d\boldsymbol{y}(x)}{dx} = \lim_{\Delta x \to 0} \frac{\Delta \boldsymbol{y}}{\Delta x} = \left(\lim_{\Delta x \to 0} \frac{\Delta y_1(x)}{\Delta x}, \quad \lim_{\Delta x \to 0} \frac{\Delta y_2}{\Delta x}, \quad \ldots, \quad \lim_{\Delta x \to 0} \frac{\Delta y_n}{\Delta x} \right) = (y_1'(x), \quad y_2'(x), \quad \ldots, \quad y_n'(x)).$$

Властивості похідної векторної функції скалярного аргументу

1) Похідна $\frac{d\boldsymbol{y}(x)}{dx} \equiv \boldsymbol{y}'(x)$ є вектором (вектор-похідною) з координатами

$$(y_1'(x), y_2'(x), \ldots, y_n'(x)).$$

2) Вектор $\frac{d\boldsymbol{y}(x)}{dx}$ лежить на дотичній до кривої $\boldsymbol{y}(x)$. Зокрема, для вектора $\boldsymbol{r}(x)$ вектор $\frac{d\boldsymbol{r}(x)}{dx}$ співпадає з дотичною до годографа в точці M (рис. 1).

3) Скалярний добуток вектор-функції і вектор-похідної дорівнює нулю:

$$\left(\boldsymbol{y}(x), \quad \frac{d\boldsymbol{y}(x)}{dx} \right) = 0.$$

38a

38. Достатні умови локального екстремуму для функцій двох змінних

ТЕОРЕМА (про дослідження на екстремум за кутовими мінорами)
Нехай всюди у околі точки (x_0, y_0):
1) визначена функція $z = z(x, y)$;
2) z_x, z_y неперервні, причому $z_x'(x_0, y_0) = 0$, $z_y'(x_0, y_0) = 0$;
3) $z_{xx}'', z_{xy}'', z_{yy}''$ неперервні.
У цьому випадку кутові мінори

$$M_1 = z_{xx}''(x_0, y_0), \quad M_2 = \begin{vmatrix} z_{xx}''(x_0, y_0) & z_{xy}''(x_0, y_0) \\ z_{xy}''(x_0, y_0) & z_{yy}''(x_0, y_0) \end{vmatrix}$$

матриці

$$\begin{pmatrix} z_{xx}''(x_0, y_0) & z_{xy}''(x_0, y_0) \\ z_{xy}''(x_0, y_0) & z_{yy}''(x_0, y_0) \end{pmatrix}$$

визначають поведінку функції $z = z(x, y)$ у точці (x_0, y_0), причому:

1) якщо кутові мінори $M_1 > 0$, $M_2 > 0$, функція $z = z(x, y)$ має в точці (x_0, y_0) мінімум;

2) якщо кутові мінори $M_1 < 0$, $M_2 > 0$, функція $z = z(x, y)$ має в точці (x_0, y_0) максимум;

3) якщо кутовий мінор $M_2 < 0$, функція $z = z(x, y)$ не має в точці (x_0, y_0) екстремуму.

Доведення. Розкладемо приріст функції в точці (x_0, y_0) у ряд Тейлора до другого диференціала включно:

$$\Delta z|_{(x_0, y_0)} = dz|_{(x_0, y_0)} + \frac{1}{2!} d^2 z \Big|_{(x_0, y_0)} + \boldsymbol{o}(\rho^2).$$

У точці (x_0, y_0) перший диференціал $dz|_{(x_0, y_0)} = 0$. Поведінка функції (її приріст) визначається другим диференціалом. Доданок $\boldsymbol{o}(\rho^2)$ є нескінченно малим по відношенню до $\frac{1}{2!} d^2 z \Big|_{(x_0, y_0)}$. Тому можна вибрати настільки малий окіл точки (x_0, y_0), що знак Δz буде визначатися знаком $d^2 z$ у цій точці. Отже, якщо $d^2 z|_{(x_0, y_0)} > 0$, то $\Delta z|_{(x_0, y_0)} > 0$, тобто при виході із точки (x_0, y_0) функція зростає. У протилежному випадку при $d^2 z|_{(x_0, y_0)} < 0$ будемо мати $\Delta z|_{(x_0, y_0)} < 0$. Функція буде спадати при русі із точки (x_0, y_0).

Розкриємо другий диференціал:

$$d^2 z = z_{xx}'' dx^2 + 2 z_{xy}'' dx dy + z_{yy}'' dy^2$$

39a

39. Властивості довжини вектора

1) Якщо вектор $\boldsymbol{x}$ нульовий, тобто $\boldsymbol{x} = (0, \ldots, 0)$, то його довжина $|\boldsymbol{x}| = \sqrt{(\boldsymbol{x}, \boldsymbol{x})} = \sqrt{0 + \ldots + 0}$, і навпаки: якщо $|\boldsymbol{x}| = 0$, то вектор $\boldsymbol{x}$ нульовий.

2) $|\lambda \boldsymbol{x}| = |\lambda| \cdot |\boldsymbol{x}|$, де $\lambda \in R$.

Дійсно,

$$|\lambda \boldsymbol{x}| = \sqrt{(\lambda x_1)^2 + (\lambda x_2)^2 + \ldots + (\lambda x_n)^2} = |\lambda| \cdot \sqrt{x_1^2 + x_2^2 + \ldots + x_n^2} = |\lambda| \cdot |\boldsymbol{x}|.$$

3) $|(\boldsymbol{x}, \boldsymbol{y})| \le |\boldsymbol{x}| \cdot |\boldsymbol{y}|$ — нерівність Коші—Буняковського.

Для доведення введемо параметр $t \in R$. За означенням скалярного добутку, нерівність

$$(t\boldsymbol{x} - \boldsymbol{y}, t\boldsymbol{x} - \boldsymbol{y}) \ge 0$$

вірна для будь-яких векторів $\boldsymbol{x}$ і $\boldsymbol{y}$. Перетворимо нерівність на квадратну відносно t:

$$t^2 (\boldsymbol{x}, \boldsymbol{x}) - 2t (\boldsymbol{x}, \boldsymbol{y}) + (\boldsymbol{y}, \boldsymbol{y}) \ge 0.$$

При $t \in R$ нерівність вірна. Отже, дискримінант квадратного тричлена повинен бути менше нуля або дорівнювати нулю:

$$D = (\boldsymbol{x}, \boldsymbol{y})^2 - (\boldsymbol{x}, \boldsymbol{x})(\boldsymbol{y}, \boldsymbol{y}) \le 0.$$

Звідси

$$(\boldsymbol{x}, \boldsymbol{y})^2 \le (\boldsymbol{x}, \boldsymbol{x})\ (\boldsymbol{y}, \boldsymbol{y}) \text{ або } |(\boldsymbol{x}, \boldsymbol{y})| \le |\boldsymbol{x}| \cdot |\boldsymbol{y}|.$$

4) $|\boldsymbol{x} + \boldsymbol{y}| \le |\boldsymbol{x}| + |\boldsymbol{y}|$ — нерівність трикутника.

З одного боку,

$$(\boldsymbol{x} + \boldsymbol{y}, \boldsymbol{x} + \boldsymbol{y}) = |\boldsymbol{x} + \boldsymbol{y}|^2.$$

З іншого боку,

$$(\boldsymbol{x} + \boldsymbol{y}, \boldsymbol{x} + \boldsymbol{y}) = (\boldsymbol{x}, \boldsymbol{x}) + 2(\boldsymbol{x}, \boldsymbol{y}) + (\boldsymbol{y}, \boldsymbol{y}) = |\boldsymbol{x}|^2 + 2(\boldsymbol{x}, \boldsymbol{y}) + |\boldsymbol{y}|^2.$$

Скористаємося нерівністю

$$2(\boldsymbol{x}, \boldsymbol{y}) \le 2|(\boldsymbol{x}, \boldsymbol{y})| \le 2|\boldsymbol{x}| \cdot |\boldsymbol{y}|.$$

Тоді

$$|\boldsymbol{x}|^2 + 2(\boldsymbol{x}, \boldsymbol{y}) + |\boldsymbol{y}|^2 \le |\boldsymbol{x}|^2 + 2|\boldsymbol{x}| \cdot |\boldsymbol{y}| + |\boldsymbol{y}|^2 = (|\boldsymbol{x}| + |\boldsymbol{y}|)^2.$$

Виходить,

$$|\boldsymbol{x} + \boldsymbol{y}|^2 \le (|\boldsymbol{x}| + |\boldsymbol{y}|)^2.$$

Добудемо квадратний корінь і в результаті одержимо нерівність трикутника.

Два ненульових вектори $\boldsymbol{x}$ та $\boldsymbol{y}$ називаються *ортогональними*, якщо їхній скалярний добуток дорівнює нулю, тобто $(\boldsymbol{x}, \boldsymbol{y}) = 0$. З рівності

40a

40. Диференціювання векторної функції векторного аргументу

Припустимо, що всі функції можуть бути продиференційовані в області D. *Похідною* $\frac{d\boldsymbol{y}}{d\boldsymbol{x}}$ *вектор-функції* $\boldsymbol{y}$ *по вектор-аргументу* $\boldsymbol{x}$ у точці $\boldsymbol{x}_0$ називається границя

$$\lim_{\Delta \boldsymbol{x} \to 0} \frac{\boldsymbol{y}(\boldsymbol{x}_0 + \Delta \boldsymbol{x}) - \boldsymbol{y}(\boldsymbol{x}_0)}{\Delta \boldsymbol{x}} = \lim_{\Delta \boldsymbol{x} \to 0} \frac{\Delta \boldsymbol{y}}{\Delta \boldsymbol{x}}.$$

Для координат вектора $\boldsymbol{y}(x)$ це означає, за означенням, що

$$\frac{d\boldsymbol{y}(\boldsymbol{x})}{d\boldsymbol{x}} = \lim_{\Delta \boldsymbol{x} \to 0} \frac{\Delta \boldsymbol{y}}{\Delta \boldsymbol{x}} = \begin{pmatrix} \lim\limits_{\Delta \boldsymbol{x} \to 0} \frac{\Delta y_1(\boldsymbol{x})}{\Delta \boldsymbol{x}} \\ \lim\limits_{\Delta \boldsymbol{x} \to 0} \frac{\Delta y_2}{\Delta \boldsymbol{x}} \\ \ldots \\ \lim\limits_{\Delta \boldsymbol{x} \to 0} \frac{\Delta y_n}{\Delta \boldsymbol{x}} \end{pmatrix} = \begin{pmatrix} \frac{dy_1}{d\boldsymbol{x}} \\ \frac{dy_2}{d\boldsymbol{x}} \\ \ldots \\ \frac{dy_n}{d\boldsymbol{x}} \end{pmatrix}.$$

У свою чергу, похідна скалярної функції векторного аргументу визначається так:

$$\lim_{\Delta \boldsymbol{x} \to 0} \frac{\Delta y_k}{\Delta \boldsymbol{x}} = \left(\lim_{\Delta x_1 \to 0} \frac{\Delta y_k}{\Delta x_1}, \quad \lim_{\Delta x_2 \to 0} \frac{\Delta y_k}{\Delta x_2}, \quad \ldots, \quad \lim_{\Delta x_m \to 0} \frac{\Delta y_k}{\Delta x_m} \right),$$

де $k = 1, 2, \ldots, n$.

Таким чином, похідною від $\boldsymbol{y}(x)$ по змінній $\boldsymbol{x}$ є матриця розмірністю $n \times m$, утворена із часткових похідних функцій за наступним правилом:

$$\frac{d\boldsymbol{y}}{d\boldsymbol{x}} = \begin{pmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_1}{\partial x_2} & \cdots & \frac{\partial y_1}{\partial x_m} \\ \frac{\partial y_2}{\partial x_1} & \frac{\partial y_2}{\partial x_2} & \cdots & \frac{\partial y_2}{\partial x_m} \\ \ldots & \ldots & \ldots & \ldots \\ \frac{\partial y_n}{\partial x_1} & \frac{\partial y_n}{\partial x_2} & \cdots & \frac{\partial y_n}{\partial x_m} \end{pmatrix}.$$

Ця матриця одержала назву матриці Якобі.

41а

41. Векторні функції скалярного аргументу

Припустимо, що кожному значенню скалярного аргументу x ставиться у відповідність n функцій $y_1(x), y_2(x), \ldots, y_n(x)$. Набір цих функцій позначимо координатами вектора $\boldsymbol{y}$

$$\boldsymbol{y}(x) = (y_1(x), y_2(x), \ldots, y_n(x))$$

і будемо вважати, що *задано векторну функцію (вектор-функцію)* $\boldsymbol{y}$ *скалярного аргументу* x.

Якщо вектор-функцію $\boldsymbol{y}(x)$ представити як радіус-вектор $\boldsymbol{r}(x)$, початок якого співпадає з початком координат, то кінець радіуса-вектора $\boldsymbol{r}(x)$ буде описувати деяку криву, що називається *годографом* векторної функції. На рис. 1 для евклідового векторного простору з розмірністю, що дорівнює трьом, побудована векторна функція

$$\boldsymbol{r}(x) = \left(\frac{\sin 6x}{x}, \quad \frac{\cos 6x}{x}, \quad x\right),$$

де $x \in [0{,}1;\, 2\pi]$. Перед нами приклад годографа, що зображує спіраль. Для чотирьох послідовних значень аргументу $t_0 < t_1 < t_2 < t_3$ показані радіуси-вектори, що виходять із початку координат.

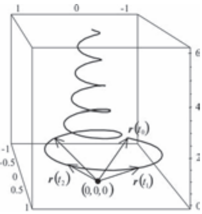

Рис. 1

Нехай $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n$ — базис n-мірного векторного простору. Можна розкласти вектор-функцію за векторами базису:

$$\boldsymbol{y}(x) = y_1(x)\boldsymbol{e}_1 + y_2(x)\boldsymbol{e}_2 + \ldots + y_n(x)\boldsymbol{e}_n.$$

42а

42. Векторні функції векторного аргументу

Нехай у деякій області D задані n функцій від m змінних:

$$y_1 = y_1(x_1, x_2, \ldots, x_m),\ y_2 = y_2(x_1, x_2, \ldots, x_m), \ldots, \quad y_n = y_n(x_1,\ x_2,\ \ldots,\ x_m). \qquad (1)$$

Інакше кажучи, нехай задані n функцій від векторного аргументу $\boldsymbol{x} = (x_1, x_2, \ldots, x_m)$:

$$y_1 = y_1(\boldsymbol{x}),\ y_2 = y_2(\boldsymbol{x}), \ldots,\ y_n = y_n(\boldsymbol{x}).$$

Набір змінних $y_1, y_2, \ldots, y_n$ можна представити як координати деякого нового вектора $\boldsymbol{y}$:

$$\boldsymbol{y} = (y_1, y_2, \ldots, y_n).$$

Тоді будемо говорити, що задано *n-мірну векторну функцію* $\boldsymbol{y}$ *від m-мірного вектор-аргументу* $\boldsymbol{x}$:

$$\boldsymbol{y} = \boldsymbol{y}(\boldsymbol{x}) = (y_1(\boldsymbol{x}), y_2(\boldsymbol{x}), \ldots, y_n(\boldsymbol{x})).$$

Якщо в кожній точці $M(x_1, x_2)$ евклідового векторного простору R^2 визначена векторна величина

$$\boldsymbol{y} = \boldsymbol{y}(M) = y_1(x_1, x_2)\cdot\boldsymbol{e}_1 + y_2(x_1, x_2)\cdot\boldsymbol{e}_2,$$

то говорять, що задано *векторне поле* $\boldsymbol{y}$. Координати $y_1(x_1, x_2), y_2(x_1, x_2)$ утворюють дві скалярні функції.

Для геометричної характеристики векторного поля служать векторні лінії. *Векторною лінією* називається крива, дотична до якої в будь-якій точці M має той самий напрямок, що й вектор поля $\boldsymbol{y}$ у цій точці. На рис. 1 пунктиром зображена векторна лінія на площині.

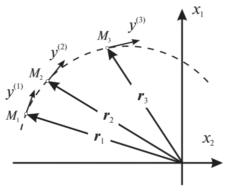

Рис. 1

43а

43. Лінійні оператори і їхні властивості

Нехай R_1^n та R_2^m — лінійні простори розмірності n та m.

Означення. *Відображенням* лінійного простору R_1^n в лінійний простір R_2^m називається правило $\tilde{P}$, за яким кожному елементу $\boldsymbol{x} \in R_1^n$ ставиться у відповідність єдиний елемент $\boldsymbol{y} \in R_2^m$. Запишемо відображення у вигляді

$$\boldsymbol{y} = \tilde{P}(\boldsymbol{x}).$$

Під елементом простору будемо розуміти не обов'язково вектор. Це може бути скалярний елемент, матриця й т. д. Частковим випадком відображення є функція $y = f(x)$, оскільки кожному значенню аргументу x ставиться у відповідність за певним правилом єдиний елемент y. Наприклад, $y = \sin(x)$.

Відображення називається *лінійним*, якщо для будь-якого елемента x_i простору й будь-якого числа λ виконуються співвідношення:

1) $\tilde{P}(\boldsymbol{x}_1 + \boldsymbol{x}_2) = \tilde{P}(\boldsymbol{x}_1) + \tilde{P}(\boldsymbol{x}_2)$;
2) $\tilde{P}(\lambda\cdot\boldsymbol{x}_2) = \lambda\cdot\tilde{P}(\boldsymbol{x})$.

Означення. Будемо розглядати лінійні відображення, що діють із векторного простору R^n у цей же простір R^n. Подібні відображення називаються *лінійними операторами*.

Введемо арифметичні операції з лінійними операторами.

1. *Сумою* операторів $\tilde{P}_1$ та $\tilde{P}_2$ називається оператор $\tilde{P} = \tilde{P}_1 + \tilde{P}_2$, який діє за наступним правилом: $\tilde{P}(\boldsymbol{x}) = \tilde{P}_1(\boldsymbol{x}) + \tilde{P}_2(\boldsymbol{x})$.

2. *Добутком оператора $\tilde{P}_1$ на число* λ називається оператор $\tilde{P}_2 = \lambda\tilde{P}_1$, що діє за правилом: $\tilde{P}_2(\boldsymbol{x}) = \lambda\cdot\tilde{P}_1(\boldsymbol{x})$.

3. *Добутком оператора $\tilde{P}_2$ на оператор $\tilde{P}_1$* називається оператор $\tilde{P} = \tilde{P}_2\tilde{P}_1$, що визначається з рівності: $\tilde{P}(\boldsymbol{x}) = \tilde{P}_2(\tilde{P}_1(\boldsymbol{x}))$.

Оператор $\tilde{E}$ називається *тотожним*, або *одиничним*, якщо $\tilde{E}(\boldsymbol{x}) = \boldsymbol{x}$.

Всі введені оператори є лінійними. Доведемо, наприклад, що оператор добутку лінійний:

$$\tilde{P}_2(\tilde{P}_1(\lambda\boldsymbol{x} + \mu\boldsymbol{y})) = \tilde{P}_2(\tilde{P}_1(\lambda\boldsymbol{x}) + \tilde{P}_1(\mu\boldsymbol{y})) = \tilde{P}_2(\lambda\tilde{P}_1(\boldsymbol{x}) + \mu\tilde{P}_1(\boldsymbol{y})) = $$
$$= \lambda\tilde{P}_2(\tilde{P}_1(\boldsymbol{x})) + \mu\tilde{P}_2(\tilde{P}_1(\boldsymbol{y})) = \lambda\tilde{P}(\boldsymbol{x}) + \mu\tilde{P}(\boldsymbol{y}).$$

44а

44. Матриці оператора в різних базисах

При переході від старого базису до нового в просторі R^n матриця P лінійного оператора $\tilde{P}$ змінюється, однак визначник матриці оператора зберігає свою величину.

ТЕОРЕМА (про зв'язок матриць оператора в різних базисах)

Матриці P і P' лінійного оператора $\tilde{P}$ в старому базисі $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n$ і новому базисі $\boldsymbol{e}_1', \boldsymbol{e}_2', \ldots, \boldsymbol{e}_n'$ зв'язані співвідношенням $P' = T^{-1}\cdot P\cdot T$, де T — матриця переходу від старого базису до нового.

Доведення. Нехай у просторі R^n задані два базиси: $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n$ та $\boldsymbol{e}_1', \boldsymbol{e}_2', \ldots, \boldsymbol{e}_n'$. Дія оператора $\tilde{P}$ на вектор $\boldsymbol{x}$ породжує вектор $\boldsymbol{y}$. У старому базисі маємо $Y = P\cdot X$, у новому базисі — $Y' = P'\cdot X'$. Нехай T — матриця переходу від старого базису до нового. Тоді $X = T\cdot X'$, $Y = T\cdot Y'$.

Помножимо рівність $X = T\cdot X'$ ліворуч на матрицю P: $P\cdot X = P\cdot X'$, $Y = T\cdot Y'$ або $Y = P\cdot T\cdot X'$. Помножимо обидві частини останньої рівності на T^{-1}: $T^{-1}\cdot Y = T^{-1}\cdot P\cdot T\cdot X'$ або $Y' = T^{-1}\cdot P\cdot T\cdot X'$. Порівнюючи отриману рівність із $Y' = T'\cdot X'$, маємо

$$P' = T^{-1}\cdot P\cdot T.$$

ТЕОРЕМА (про визначник оператора в різних базисах)

Визначник матриці лінійного оператора не залежить від вибору базису.

Доведення. Обчислимо визначник матриці P', використовуючи властивість визначника добутку матриць, а також властивість визначників обернених матриць:

$$|P'| = |T^{-1}\cdot P\cdot T| = |T^{-1}|\cdot|P|\cdot|T| = |T^{-1}|\cdot|T|\cdot|P| = |E|\cdot|P| = |P|.$$

Наслідок. Ранг матриці лінійного оператора не залежить від вибору базису.

ПРИКЛАД. У базисі $\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3$ лінійний оператор заданий матрицею $P = \begin{pmatrix} 2 & -1 & -1 \\ 1 & 2 & -2 \\ 3 & 0 & -1 \end{pmatrix}$. Знайти матрицю оператора в новому базисі $\boldsymbol{e}_1', \boldsymbol{e}_2', \boldsymbol{e}_3'$, що пов'язаний зі старим базисом матрицею переходу

45а

45. Власні вектори й власні значення лінійного оператора

Означення. Ненульовий вектор $\boldsymbol{x}$ називається *власним вектором* лінійного оператора $\tilde{P}$, якщо знайдеться таке число λ, яке називається *власним значенням* лінійного оператора, що виконується

$$\tilde{P}(\boldsymbol{x}) = \lambda \cdot \boldsymbol{x}. \qquad (1)$$

Рівність (1) означає, що вектор $\boldsymbol{x}$ під дією лінійного оператора множиться на число λ. З'являється колінеарний вектор. Серед векторів лінійного векторного простору можуть існувати такі, дія оператора на які переводить ці вектори в колінеарні собі. Якщо на таких векторах побудувати базис, перетворення лінійної алгебри значно спростяться.

Не всякий лінійний оператор має власні вектори. Наприклад, у геометричній площині R^2 оператор повороту на кут, не кратний π, не має жодного власного вектора, оскільки жоден ненульовий вектор після повороту не залишиться колінеарним самому собі.

Розв'яжемо задачу знаходження власних векторів оператора. Запишемо рівність (1) у матричній формі:

$$P \cdot X = \lambda \cdot X.$$

Перетворимо матричне рівняння:

$$P \cdot X - \lambda \cdot X = 0 \text{ або } (P - \lambda \cdot E)X = 0.$$

Матричне рівняння завжди має нульовий *розв'язок*:

$$X = 0.$$

Для існування ненульових розв'язків ранг матриці коефіцієнтів повинен бути менше числа змінних $r < n$, тобто число лінійно незалежних рівнянь повинне бути менше числа змінних. У цьому випадку має виконуватися умова

$$|P - \lambda \cdot E| = 0. \qquad (2)$$

Розписавши рівняння (2) відносно λ докладніше, одержимо

$$|P-\lambda\cdot E| = \begin{vmatrix} a_{11}-\lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22}-\lambda & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn}-\lambda \end{vmatrix} = 0.$$

46а

46. Алгебраїчна й геометрична кратність коренів характеристичного многочлена

Розв'яжемо задачу знаходження власних векторів оператора. У цьому випадку повинна бути виконана умова

$$|P - \lambda \cdot E| = 0. \qquad (1)$$

Розписавши рівняння (1) відносно λ докладніше, одержимо

$$|P-\lambda\cdot E| = \begin{vmatrix} a_{11}-\lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22}-\lambda & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn}-\lambda \end{vmatrix} = 0.$$

Розкривши визначник, одержимо рівняння n-го степеня відносно λ:

$$(-1)^n\lambda^n + \alpha_{n-1}\lambda^{n-1} + \ldots + \alpha_1\lambda + \alpha_0 = 0,$$

яке називається *характеристичним рівнянням* оператора $\tilde{P}$. Корені рівняння називаються *характеристичними,* або власними *числами* оператора. Множина всіх власних чисел оператора $\tilde{P}$ називається *спектром* цього оператора. Многочлен лівої частини рівняння називається *характеристичним многочленом*.

Розв'язавши характеристичне рівняння, одержуємо власні числа $\lambda_1, \lambda_2, \ldots, \lambda_n$. Для кожного знайденого власного значення λ_i знайдемо ненульові вектори ядра оператора $P - \lambda_i E$. Саме вони будуть власними векторами, що відповідають власному значенню λ_i. Інакше кажучи, необхідно розв'язати однорідну систему рівнянь $(P - \lambda_i \cdot E)X = 0$. Її загальний розв'язок дає всю сукупність власних векторів, що відповідають λ_i.

Загальний *розв'язок* однорідної системи, як відомо, структурований. Він являє собою лінійну комбінацію фундаментального набору лінійно незалежних *розв'язків* (векторів). Число лінійно незалежних векторів у фундаментальному наборі називається *геометричною кратністю* власного значення λ_i. Вводиться також *алгебраїчна кратність* — кратність λ_i як кореня характеристичного многочлена.

47а

47. Знаковизначеність квадратичної форми

Означення. Квадратична форма $L(x)$ називається *додатно визначеною,* або знакододатною, якщо для будь-якого ненульового вектора $\boldsymbol{x}$ виконується нерівність $L(x) > 0$.

Означення. Квадратична форма $L(x)$ називається *додатно напіввизначеною*, якщо для будь-якого ненульового вектора $\boldsymbol{x}$ виконується нерівність $L(x) \geq 0$.

Аналогічно визначаються *від'ємно визначена* й *недодатно визначена* (*від'ємно напіввизначена*) квадратична форма.

Означення. Мінори, що приєднуються до лівого верхнього кута матриці, називаються *кутовими*.

Означення. Мінори, що мають своєю діагоналлю головну діагональ матриці, називаються *головними*.

ТЕОРЕМА (про визначення знака форми за власними числами)

Для того щоб квадратична форма була додатно (від'ємно) *визначеною*, необхідно й достатньо, щоб всі власні числа матриці квадратичної форми були додатними (від'ємними).

Доведення. Нехай для визначеності квадратична форма додатно визначена. Представимо квадратичну форму в канонічному вигляді $L(\boldsymbol{x}) = \lambda_1 x_1^2 + \lambda_2 x_2^2 + \ldots + \lambda_n x_n^2 > 0$. Тут $\lambda_1, \lambda_2, \ldots, \lambda_n$ — власні числа матриці в базисі $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n$.

Необхідність. Нехай $\hat{L}(\boldsymbol{x}) > 0$. Квадратична форма пов'язана з оператором: $\hat{L}(\boldsymbol{x}) = (\tilde{P}(\boldsymbol{x}), \boldsymbol{x})$.

Зокрема, $L(\boldsymbol{e}_i) = (\tilde{P}(\boldsymbol{e}_i), \boldsymbol{e}_i) = (\lambda_i \boldsymbol{e}_i, \boldsymbol{e}_i) = \lambda_i$, де $i = 1, 2, \ldots, n$. З умови $L(\boldsymbol{e}_i) > 0$ випливає $\lambda_i > 0$.

Достатність. Якщо всі $\lambda_1, \lambda_2, \ldots, \lambda_n$ додатні, то

$$L(\boldsymbol{x}) = \lambda_1 x_1^2 + \lambda_2 x_2^2 + \ldots + \lambda_n x_n^2 > 0.$$

48а

48. Відображення. Образ, ранг, ядро

Нехай R_1^n та R_2^m — лінійні простори розмірності n та m.

Означення. *Відображенням* лінійного простору R_1^n у лінійний простір R_2^m називається правило $\tilde{P}$, за яким кожному елементу $\boldsymbol{x} \in R_1^n$ ставиться у відповідність єдиний елемент $\boldsymbol{y} \in R_2^m$. Запишемо відображення у вигляді

$$\boldsymbol{y} = \tilde{P}(\boldsymbol{x}).$$

Під елементом простору будемо розуміти не обов'язково вектор. Це може бути скалярний елемент, матриця й т. д. Частковим випадком відображення є функція $y = f(x)$, оскільки кожному значенню аргумента x ставиться у відповідності за певним правилом єдиний елемент y. Наприклад, $y = \sin(x)$.

Відображення називається *лінійним*, якщо для будь-якого елемента x_i простору й будь-якого числа λ виконуються співвідношення:

1) $\tilde{P}(\boldsymbol{x}_1 + \boldsymbol{x}_2) = \tilde{P}(\boldsymbol{x}_1) + \tilde{P}(\boldsymbol{x}_2)$;
2) $\tilde{P}(\lambda \cdot \boldsymbol{x}_2) = \lambda \cdot \tilde{P}(\boldsymbol{x})$.

Приклади лінійних відображень

1. Якщо кожному вектору $\boldsymbol{x} \in R^n$ ставиться у відповідність вектор $\boldsymbol{y} = \alpha \cdot \boldsymbol{x} \in R^n$, де $\alpha \neq 0$, то кажуть, що задано відображення (перетворення) *подібності*. Процес відображення — це множення кожного вектора на число.

2. Якщо кожній матриці-стовпцю X розмірності $n \times 1$ із простору матриць $R^{n\times 1}$ ставиться у відповідність матриця-стовпець Y розмірності $m \times 1$ із простору матриць $R^{m\times 1}$, то задано відображення простору стовпців $\underset{n\times 1}{X}$ у простір стовпців $\underset{m\times 1}{Y}$. Процес відображення — це множення матриці $\underset{m\times n}{P}$ на стовпці $\underset{n\times 1}{X}$:

$$\underset{m\times 1}{Y} = \underset{m\times n}{P} \cdot \underset{n\times 1}{X}.$$

Рівними називаються лінійні відображення $\tilde{P}_1$ та $\tilde{P}_2$, якщо для будь-якого елемента $\boldsymbol{x} \in R^m$ виконується рівність

$$\tilde{P}_1(\boldsymbol{x}) = \tilde{P}_2(\boldsymbol{x}).$$

49a

49. Симетричний оператор

Означення. Лінійний оператор $\tilde{P}$ в евклідовому просторі R^n називається *симетричним*, якщо для будь-яких векторів $\boldsymbol{x}$ і $\boldsymbol{y}$ із простору R^n виконується рівність

$$(\tilde{P}(\boldsymbol{x}),\ \boldsymbol{y}) = (\boldsymbol{x},\ \tilde{P}(\boldsymbol{y})).$$

ТЕОРЕМА (про умову симетричності оператора)

Для того щоб лінійний оператор був симетричним, необхідно й достатньо, щоб його матриця в ортонормованому базисі була симетричною.

Доведення. Розглянемо для простоти евклідовий простір R^2. Нехай в ортобазисі $\boldsymbol{e}_1, \boldsymbol{e}_2$ задані вектори $\boldsymbol{x} = (x_1, x_2)$, $\boldsymbol{y} = (y_1, y_2)$. Лінійні оператори $\tilde{P}_1$ та $\tilde{P}_2$ визначені своїми матрицями: $P_1 = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ та $P_2 = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}$. Обчислимо вектори $\tilde{P}_1(\boldsymbol{x})$ та $\tilde{P}_2(\boldsymbol{y})$:

$$\tilde{P}_1(\boldsymbol{x}) = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + a_{12}x_2 \\ a_{21}x_1 + a_{22}x_2 \end{pmatrix},$$

$$\tilde{P}_2(\boldsymbol{y}) = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \cdot \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} b_{11}y_1 + b_{12}y_2 \\ b_{21}y_1 + b_{22}y_2 \end{pmatrix}.$$

Знайдемо скалярні добутки $(\tilde{P}(\boldsymbol{x}),\ \boldsymbol{y})$ та $(\boldsymbol{x},\ \tilde{P}(\boldsymbol{y}))$:

$$(\tilde{P}_1(\boldsymbol{x}),\ \boldsymbol{y}) = (a_{11}x_1 + a_{12}x_2)y_1 + (a_{21}x_1 + a_{22}x_2)y_2 = \\ = a_{11}y_1x_1 + a_{12}y_1x_2 + a_{21}y_2x_1 + a_{22}y_2x_2,$$

$$(\boldsymbol{x},\ \tilde{P}_2(\boldsymbol{y})) = (b_{11}y_1 + b_{12}y_2)x_1 + (b_{21}y_1 + b_{22}y_2)x_2 = \\ = b_{11}x_1y_1 + b_{12}x_1y_2 + b_{21}x_2y_1 + b_{22}x_2y_2.$$

Знайдемо різницю скалярних добутків:

$$(\tilde{P}_1(\boldsymbol{x}),\ \boldsymbol{y}) - (\boldsymbol{x},\ \tilde{P}_2(\boldsymbol{y})) = (a_{11} - b_{11})x_1y_1 + (a_{21} - b_{12})x_1y_2 + \\ + (a_{12} - b_{21})x_2y_1 + (a_{22} - b_{22})x_2y_2.$$

Якщо для будь-яких векторів x і y із простору R^2 рівність

$$(\tilde{P}_1(\boldsymbol{x}),\ \boldsymbol{y}) - (\boldsymbol{x},\ \tilde{P}_2(\boldsymbol{y})) = 0 \qquad (3)$$

виконується (необхідність), то вірна система

50a

50. Рядок матриці як лінійна комбінація незалежних рядків матриці

Рядок $\boldsymbol{e}$, що визначений рівністю

$$\boldsymbol{e} = \lambda_1\boldsymbol{e}_1 + \lambda_2\boldsymbol{e}_2 + \ldots + \lambda_m\boldsymbol{e}_m,$$

називається *лінійною комбінацією* рядків $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_m$, де $\lambda_1, \lambda_2, \ldots, \lambda_m$ — будь-які дійсні числа.

Рядки $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_m$ називаються *лінійно залежними*, якщо існують такі $\lambda_1, \lambda_2, \ldots, \lambda_m$, що не дорівнюють нулю одночасно, і при цьому лінійна комбінація цих рядків дорівнює нульовому рядку

$$\lambda_1\boldsymbol{e}_1 + \lambda_2\boldsymbol{e}_2 + \ldots + \lambda_m\boldsymbol{e}_m = \boldsymbol{o}.$$

Рядки $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_m$ називаються *лінійно незалежними*, якщо лінійна комбінація цих рядків дорівнює нульовому рядку тільки при $\lambda_1 = \lambda_2 = \ldots = \lambda_m = 0$.

ТЕОРЕМА (про запис рядка у вигляді лінійної комбінації незалежних рядків)

Кожний рядок матриці A може бути представлений у вигляді лінійної комбінації незалежних рядків матриці.

Доведення. Нехай матриця $\underset{m\times n}{A}$ має ранг r. За означенням рангу матриці, існує мінор порядку r, відмінний від нуля. Нехай для визначеності це мінор

$$M_r = \begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1r} \\ a_{21} & a_{22} & \ldots & a_{2r} \\ \ldots & \ldots & \ldots & \ldots \\ a_{r1} & a_{r2} & \ldots & a_{rr} \end{vmatrix} \neq 0.$$

Розглянемо мінор $(r + 1)$-го порядку матриці A, який можна одержати, додавши до мінору M_r i-й рядок та j-й стовпець матриці $(r < i \le m,\ r < j \le n)$:

$$M_{r+1} = \left|\begin{array}{cccc|c} a_{11} & a_{12} & \ldots & a_{1r} & a_{1j} \\ a_{21} & a_{22} & \ldots & a_{2r} & a_{2j} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ a_{r1} & a_{r2} & \ldots & a_{rr} & a_{rj} \\ \hline a_{i1} & a_{i2} & \ldots & a_{ir} & a_{ij} \end{array}\right|.$$

51a

51. Квадратичні форми

Нехай $L = (a_{ij})$ — симетрична матриця n-го порядку, тобто $a_{ij} = a_{ji}$.

Означення. Вираз

$$\hat{L} = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_ix_j \qquad (1)$$

називається квадратичною формою змінних $x_1, x_2, \ldots, x_n$. Вираз (1) є сумою всіх квадратів змінних плюс сума всіх подвоєних добутків різних змінних, причому кожний член суми взятий з деяким коефіцієнтом. Матриця L називається *матрицею квадратичної форми*.

Побудуємо квадратичну форму. Введемо матрицю-стовпець змінних $X = \begin{pmatrix} x_1 \\ x_2 \\ \ldots \\ x_n \end{pmatrix}$, матрицю-рядок цих змінних $X^T = (x_1, x_2, \ldots, x_n)$ і знайдемо добуток матриць:

$$(x_1, x_2, \ldots, x_n)\begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ \ldots \\ x_n \end{pmatrix} = (x_1, x_2, \ldots, x_n)\begin{pmatrix} \sum\limits_{j=1}^{n} a_{1j}x_j \\ \sum\limits_{j=1}^{n} a_{2j}x_j \\ \ldots \\ \sum\limits_{j=1}^{n} a_{nj}x_j \end{pmatrix}.$$

Після множення маємо

$$\sum_{j=1}^{n} a_{1j}x_1x_j + \sum_{j=1}^{n} a_{1j}x_2x_j + \ldots + \sum_{j=1}^{n} a_{1j}x_nx_j = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_ix_j.$$

Отже, у матричній формі квадратична форма може бути представлена у вигляді

$$\hat{L} = X^T \cdot L \cdot X.$$

Матриці-стовпцю змінних можна поставити у відповідність вектор $\boldsymbol{x}$, координатами якого в ортобазисі $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n$ будуть елементи матриці-стовпця. Тоді вираз (1) можна інтерпретувати як числову функцію векторного аргументу $\boldsymbol{x}$: $\hat{L}(\mathbf{x})$.

ПРИКЛАД. Знайти матрицю квадратичної форми

$$\hat{L}(\mathbf{x}) = -x_1^2 + 6x_1x_2 - 3x_1x_3 + 4x_2x_3 + x_2^2 - 3x_3^2.$$

52a

52. Приведення квадратичної форми до канонічного вигляду ортогональним перетворенням

У лінійній алгебрі часто виникає необхідність приведення квадратичної форми до найбільш простого вигляду. У загальному випадку таким виглядом є діагональний вигляд квадратичної форми.

Означення. Квадратична форма називається *канонічною*, якщо всі коефіцієнти при добутках різних змінних дорівнюють нулю, тобто $a_{ij} = 0$ при $i \neq j$:

$$L(\boldsymbol{x}) = a_{11}x_1^2 + a_{22}x_2^2 + \ldots + a_{nn}x_n^2 = \sum_{i=1}^{n} a_{ii}x_i^2.$$

Матриця канонічної квадратичної форми є діагональною:

$$\begin{pmatrix} a_{11} & 0 & \ldots & 0 \\ 0 & a_{22} & \ldots & 0 \\ 0 & 0 & \ldots & 0 \\ 0 & 0 & \ldots & a_{nn} \end{pmatrix}.$$

ТЕОРЕМА (про зв'язок між формою й оператором)

Нехай матриці симетричного лінійного оператора й квадратичної форми, що задані в ортонормованому базисі, співпадають. Тоді квадратична форма пов'язана з оператором в евклідовому просторі формулою

$$\hat{L}(\boldsymbol{x}) = (\tilde{P}(\boldsymbol{x}),\ \boldsymbol{x}). \qquad (1)$$

ТЕОРЕМА (про приведення квадратичної форми до канонічного вигляду)

Будь-яка квадратична форма, задана в евклідовому просторі, може бути приведена до канонічного вигляду.

Доведення. Нехай матриця симетричного оператора $\tilde{P}$ в евклідовому просторі співпадає з матрицею квадратичної форми та є ортонормований базис $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n$, що складається із власних векторів лінійного симетричного оператора $\tilde{P}$ (враховуючи одну із властивостей симетричного оператора, це завжди можливо):

$$\tilde{P}(\boldsymbol{e}_i) = \lambda_i\boldsymbol{e}_i.$$

53а

53. Локальний екстремум

Нехай функція $z = z(x, y)$ існує в деякій області, і точка $M(x_0, x_0)$ — внутрішня точка цієї області.

Означення. Точка $M(x_0, x_0)$ називається точкою *локального максимуму (мінімуму)* функції $z = z(x, y)$, якщо існує окіл точки $M(x_0, x_0)$, такий, що для всіх точок $M(x, y)$ із цього околу $z(x_0, y_0) \ge z(x, y)$ $(z(x_0, y_0) \le z(x, y))$.

Визначення можна інтерпретувати в термінах приростів. Якщо приріст функції $\Delta z = z(x_0 + \Delta x, y_0 + \Delta y) - z(x_0, y_0)$, що відповідає приросту аргументів Δx і Δy, у деякому околі точки $M(x_0, x_0)$ зберігає знак, то точка $M(x_0, x_0)$ є точкою локального максимуму (мінімуму).

Значення функції в точці $M(x_0, x_0)$ називається *локальним максимумом* або *локальним мінімумом* функції. Точки максимуму й точки мінімуму функції називаються точками *екстремуму* функції, а самі максимуми й мінімуми функції — її *екстремумами*.

Будемо розрізняти *строгий екстремум* (його умова $z(x_0, y_0) > z(x, y)$ або $z(x_0, y_0) < z(x, y)$) і *нестрогий екстремум* (умова $z(x_0, y_0) \ge z(x, y)$ або $z(x_0, y_0) \le z(x, y)$).

ТЕОРЕМА (про необхідну умову екстремуму)

Якщо $M(x_0, x_0)$ — точка екстремуму функції, то всі часткові похідні в цій точці дорівнюють нулю або не існують.

Доведення. Нехай у точці $M(x_0, x_0)$ функція $z = z(x, y)$ має екстремум. Зафіксуємо змінну y, поклавши її значення $y = y_0$. Тоді функція $z = z(x, y_0)$ буде функцією однієї змінної по x. Оскільки вона має екстремум у точці x_0, її похідна по x, що дорівнює $\left.\dfrac{\partial z}{\partial x}\right|_{(x_0,\, y_0)}$, буде нулем або не буде існувати. Подібним чином переконуємося, що $\left.\dfrac{\partial z}{\partial y}\right|_{(x_0,\, y_0)}$ також дорівнює нулю або не існує.

Наслідок. Якщо $M(x_0, x_0)$ — точка екстремуму функції $z = z(x, y)$, то перший диференціал функції дорівнює нулю.

54а

54. Умовний екстремум

Означення. *Умовним екстремумом* функції $z = z(x, y)$ називається максимум або мінімум функції, який досягається за умови, що її аргументи зв'язані деяким рівнянням $g(x, y) = 0$:

$$\begin{cases} z = z(x,\, y), \\ g(x,\, y) = 0. \end{cases}$$

Надалі екстремуми, що не є умовними, будемо називати *безумовними*.

При знаходженні умовних екстремумів функції $z = z(x, y)$ аргументи x і y вже не можна розглядати як незалежні змінні. Вони зв'язані між собою співвідношенням $g(x, y) = 0$, що називається *рівнянням зв'язку*.

Для пояснення відмінності між локальним (безумовним) і умовним екстремумом розглянемо функцію

$$z = z(x, y) = x^2 + y^2.$$

Вона описує так званий параболоїд обертання й має безумовний мінімум у точці, що зазначена темною стрілкою. Додамо рівняння зв'язку (обмеження у вигляді рівності)

$$g(x, y) = y - x,$$

що описує площину. Задача формулюється тепер так: знайти екстремум функції $z = x^2 + y^2$, розглядаючи серед всіх значень (x, y) тільки ті, які в сукупності утворять площину P. Інакше кажучи, екстремум варто шукати серед точок, що належать одночасно обом поверхням, зображеним на рис. 1. Ці точки утворюють білу лінію. Мінімальне значення (умовний мінімум) досягається в точці, що зазначена світлою стрілкою.

Рис. 1

55а

55. Екстремум неявної функції

Нехай функція $z = z(x, y)$ визначається рівнянням $F(x, y, z) = 0$. А також нехай функція $z = z(x, y)$ буде двічі неперервно диференційована та $F'_z \ne 0$.

Перший диференціал у стаціонарній точці дорівнює нулю:

$$dz = -\frac{F'_x}{F'_z}dx - \frac{F'_y}{F'_z}dy = 0,$$

що може мати місце, враховуючи незалежність диференціалів dx і dy, тільки при

$$\begin{cases} F'_x = 0, \\ F'_y = 0. \end{cases} \tag{1}$$

Рівняння (1) виражають необхідні умови існування екстремуму, визначаючи точку (x_0, y_0) можливого екстремуму. Для знаходження достатніх умов знайдемо другий диференціал у стаціонарній точці (x_0, y_0):

$$\begin{aligned} d^2z = d(dz) = d\left(-\frac{F'_x}{F'_z}dx - \frac{F'_y}{F'_z}dy\right)_{(x_0,\, y_0)} &= \\ = -dx \cdot d\left(\frac{F'_x}{F'_z}\right)_{(x_0,\, y_0)} - dy \cdot d\left(\frac{F'_y}{F'_z}\right)_{(x_0,\, y_0)}&. \end{aligned} \tag{2}$$

Знайдемо окремо кожен з доданків у правій частині рівності (2):

$$dx \cdot d\left(\frac{F'_x}{F'_z}\right) = dx \cdot \left(\frac{F'_x}{F'_z}\right)'_x \cdot dx + dx \cdot \left(\frac{F'_x}{F'_z}\right)'_y \cdot dy =$$

$$= \left(\frac{F''_{xx}F'_z - F'_xF''_{zx}}{(F'_z)^2}\right)_{(x_0,\, y_0)} dx^2 + \left(\frac{F''_{xy}F'_z - F'_xF''_{zy}}{(F'_z)^2}\right)_{(x_0,\, y_0)} dxdy.$$

У стаціонарній точці $F'_x = 0$, тому рівність значно спроститься:

$$dx \cdot d\left(\frac{F'_x}{F'_z}\right) = \left(\frac{F''_{xx}}{F'_z}\right)_{(x_0,\, y_0)} dx^2 + \left(\frac{F''_{xy}}{F'_z}\right)_{(x_0,\, y_0)} dxdy.$$

Аналогічно можна довести, що

$$dy \cdot d\left(\frac{F'_x}{F'_z}\right) = \left(\frac{F''_{yx}}{F'_z}\right)_{(x_0,\, y_0)} dxdy + \left(\frac{F''_{yy}}{F'_z}\right)_{(x_0,\, y_0)} dy^2.$$

56а

56. Глобальний екстремум

Глобальним екстремумом називаються найбільше й найменше значення функції в замкненій обмеженій області D. Нехай функція $z = z(x, y)$ неперервна в цій області. Тоді знайдеться точка (x_0, y_0), у якій функція має найбільше (найменше) значення. Якщо точка (x_0, y_0) лежить усередині області D, вона є стаціонарною і в ній може досягатися локальний максимум або мінімум. Найбільшого або найменшого значення функція може набувати також на границі області. Отже, задачу дослідження функції $z = z(x, y)$ на глобальний (global) екстремум в області $g(x, y) \le 0$ можна сформулювати так:

$$\begin{cases} z = z(x,\, y) \to \text{gl. extr}, \\ g(x,\, y) \le 0. \end{cases}$$

Розв'язання задачі розбивається на дві частини:

1) дослідження функції $z = z(x, y)$ на локальний (local) екстремум в області $g(x, y) < 0$:

$$\begin{cases} z = z(x,\, y) \to \text{loc. extr}, \\ g(x,\, y) < 0. \end{cases}$$

2) дослідження функції $z = z(x, y)$ на умовний (conditional) екстремум на границі області $g(x, y) = 0$:

$$\begin{cases} z = z(x,\, y) \to \text{cond. extr}, \\ g(x,\, y) = 0. \end{cases}$$

Найбільше (найменше) із всіх цих чисел і буде шуканим найбільшим (найменшим) значенням функції $z = z(x, y)$ в області D.

ПРИКЛАД. Дослідити на глобальний екстремум функцію $z = 2 - x^2 - y^2$ в області $x^2 + y^2 \ge 1$, $|x| \le 2$, $|y| \le 2$.

Розв'язання. Для розв'язання задачі

$$\begin{cases} z = 2 - x^2 - y^2 \to \text{gl. extr}, \\ x^2 + y^2 \ge 1, \\ |x| \le 2, |y \le 2| \end{cases}$$

треба розглянути кілька випадків:

$$1)\ \begin{cases} z = 2 - x^2 - y^2 \to \text{loc. extr}, \\ x^2 + y^2 > 1, \\ |x| < 2, |y| < 2. \end{cases}$$

57a

57. Екстремум у системах функцій

У загальному випадку задачу можна сформулювати так: знайти екстремум функції $z = z(x, y)$, що задана системою функцій

$$\begin{cases} z = f(x, y), \\ z = g(x, y). \end{cases} \qquad (1)$$

Пошук розв'язку почнемо зі знаходження множини таких точок на координатній площині xOy, у яких значення функцій системи рівні

$$f(x, y) = g(x, y).$$

Розв'язок цього рівняння $y = y(x)$ підставимо в одну з функцій $z = f(x, y(x)) = z(x)$ і дослідимо на екстремум функцію $z = z(x)$. Критична точка x_0 для функції $z = z(x)$ задає стаціонарну точку (x_0, y_0) на площині xOy, у якій можливий екстремум $z_{\text{extr}} = f(x_0, y_0)$. Геометрично це означає, що в просторі знаходиться сукупність ліній перетину поверхонь, що описуються заданими функціями. На лініях перетину знаходимо найбільше або найменше значення z_0.

Одержимо необхідні умови екстремуму системи функцій. Вважаючи в системі (1) змінні x і y функціями змінної z, продиференціюємо обидва рівняння по z:

$$\begin{cases} f'_x \cdot x'_z + f'_y \cdot y'_z = 1, \\ g'_x \cdot x'_z + g'_y \cdot y'_z = 1. \end{cases}$$

Віднімемо від першого рівняння друге:

$$(f'_x - g'_x)x'_z + (f'_y - g'_y)y'_z = 0$$

і, поклавши , що $x'_z \neq 0$ та $(f'_y - g'_y) \neq 0$, поділимо обидві частини рівності на ці величини. Маємо

$$\frac{f'_x - g'_x}{f'_y - g'_y} = -\frac{y'_z}{x'_z}.$$

Очевидно, $\dfrac{y'_z}{x'_z} = \dfrac{\frac{dy}{dz}}{\frac{dx}{dz}} = \dfrac{dy}{dx} \equiv y'_x$.

Оскільки $z = f(x, y(x))$, знайдемо похідну по x, що у критичній точці z_0 буде дорівнювати нулю:

$$z'_x\big|_{x=x_0} = f'_x + f'_y \cdot y'_x = 0.$$

58a

58. Екстремум у системах нерівностей

Математичними об'єктами, що задаються системами нерівностей, у тривимірному випадку є об'ємні геометричні фігури, границі яких описуються деякими функціями. Формулювання задачі виглядає так: знайти екстремум функції $z = z(x, y)$, що задана системою нерівностей

$$\begin{cases} F(x, y, z) \leq 0, \\ G(x, y, z) \leq 0. \end{cases}$$

Задача пошуку подібного екстремуму математичного об'єкта зводиться до розв'язку декількох систем.

1. Дослідження функції $z = z(x, y)$, що визначається з рівняння $F(x, y, z) = 0$, на локальний екстремум за умови справедливості нерівності $G(x, y, z) < 0$, тобто

$$\begin{cases} F(x, y, z) = 0 \to z = z(x, y) \to \text{loc. extr}, \\ G(x, y, z) < 0. \end{cases}$$

2. Дослідження функції $z = z(x, y)$, що визначається з рівняння $G(x, y, z) = 0$, на локальний екстремум за умови справедливості нерівності $F(x, y, z) < 0$, тобто

$$\begin{cases} G(x, y, z) = 0 \to z = z(x, y) \to \text{loc. extr}, \\ F(x, y, z) < 0. \end{cases}$$

3. Дослідження функції $z = z(x, y)$, що визначається з рівняння $F(x, y, z) = G(x, y, z)$, яке описує лінію перетину поверхонь.

ПРИКЛАД. Дослідити на екстремум функцію $z = x(x, y)$, що задана системою нерівностей

$$\begin{cases} z \geq x^2 - xy + y^2, \\ z \leq x + y + 1. \end{cases}$$

Розв'язання. 1. Дослідимо функцію $z = x^2 - xy + y^2$ на локальний екстремум в області $z < x + y + 1$, тобто розглянемо систему

$$\begin{cases} z = x^2 - xy + y^2, \\ z < x + y + 1. \end{cases}$$

59a

59. Максимізація прибутку в проектному аналізі

У процесі прийняття інвестиційних рішень доводиться вирішувати задачу максимізації прибутку, який планується. Одна з функцій, за допомогою якої в проектному аналізі приймається інвестиційне рішення (критерій), називається *чистим дисконтованим доходом* NPV (Net Present Value). Економічний зміст цієї функції полягає в максимізації різниці між всіма проектними доходами й витратами, тобто прибутку. Однак, враховуючи теорію проектного аналізу, у цьому розрахунку необхідно враховувати час проведення витрат і час одержання доходів, а також процентну ставку. Все це знаходить висвітлення в критерії NPV завдяки застосуванню теорії зміни цінності грошей у часі.

Поклавши гроші в банк, ми розраховуємо на грошовий приріст. Він визначається за формулою складних відсотків

$$X(t) = X(0)\left(1 + \frac{r}{100}\right)^t,$$

де $X(0)$ — первісний внесок, покладений у банк на t років під r річних відсотків. Поставимо тепер задачу іншим чином. Яку суму треба покласти в банк під r % у рік, щоб через t років одержати $X(t)$ грошових одиниць? Інакше кажучи, знайти вартість суми грошей, отриманої через t років, але наведеної на теперішній час. Вочевидь, вона дорівнює

$$X(0) = \frac{X(t)}{\left(1 + \frac{r}{100}\right)^t}.$$

Величину r можна назвати *ставкою дисконту*, що показує зміну вартості грошей у часі.

Припустимо, розраховується проект, що принесе через рік дохід у розмірі $b(1)$ грошових одиниць, через два роки — $b(2)$ грошових одиниць і т. д. Тоді, порахований на теперішній час, він складе протягом t років величину, що дорівнює

$$B_{enefits} = \frac{b(1)}{1 + \frac{r}{100}} + \frac{b(2)}{\left(1 + \frac{r}{100}\right)^2} + \ldots + \frac{b(t)}{\left(1 + \frac{r}{100}\right)^t} = \sum_{n=1}^{t} \frac{b(n)}{\left(1 + \frac{r}{100}\right)^n}.$$

60a

60. Глобальний екстремум у задачах математичного програмування

У загальному вигляді *задача математичного програмування* для двох змінних записується так:

$$1) \begin{cases} z = f(x, y) \to \max, \\ g_1(x, y) \leq 0, \\ \ldots\ldots\ldots\ldots\ldots, \\ g_n(x, y) \leq 0, \\ x \geq 0, \\ y \geq 0. \end{cases} \quad \text{або } 2) \begin{cases} z = f(x, y) \to \min, \\ g_1(x, y) \geq 0, \\ \ldots\ldots\ldots\ldots\ldots, \\ g_n(x, y) \geq 0, \\ x \geq 0, \\ y \geq 0, \end{cases}$$

де $f(x, y)$ — *цільова функція*, що не має локальних екстремумів, $g_1(x, y), \ldots, g_n(x, y)$ — функції, що обмежують пошук розв'язку. Обмеження $x \geq 0$, $y \geq 0$ вказують на пошук *розв'язку* тільки в першій чверті координатної площини. Точка (x_0, y_0) називається *оптимальним розв'язком* системи рівнянь і нерівностей, якщо в цій точці цільова функція досягає максимального значення в першому випадку й мінімального — у другому. Максимум або мінімум цільової функції є глобальними екстремумами. У цьому відмінність задачі математичного програмування від задач на умовний або локальний і навіть на глобальний екстремум, що може досягатися як на границі області, так і у внутрішніх точках. Задачі 1 і 2 називаються *спряженими* й утворюють двоїсту пару.

До необхідності розв'язувати задачу математичного програмування (ЗМП) приводять:

1. *Проблема планування виробництва*, тобто планування виробництва певних видів продукції так, щоб було забезпечено найбільш раціональне використання матеріальних, фінансових і інших ресурсів. Повинен досягатись максимум або мінімум (в економіці — оптимум) деякої функції, що описує прибуток, обсяг виробництва й т. д.

2. *Проблема оптимального змішування*. Потрібно вибрати кількість кожного з початкових інгредієнтів для складання суміші, якщо відома вартість одиниці інгредієнта. Суміш треба одержати із заданими властивостями, причому з найменшими витратами. Такі задачі оптимального

61a

61. Оптимізація споживчої поведінки (функція попиту)

Розглянемо задачу залежності об'єму особистого споживання або попиту від доходів, цін і соціально-демографічних факторів. Нехай $U(x)$ — цільова функція споживання, що характеризує переваги споживача, $\boldsymbol{x} = (x_1, x_2, \dots, x_n)$ — вектор об'ємів споживання благ x_i, n — кількість розглянутих благ. Купуючи для особистого споживання товари x_i за цінами p_i, покупець обмежений розміром свого доходу I. Ставиться задача оптимізації споживчого поводження з бюджетним обмеженням

$$\begin{cases} U(\boldsymbol{x}) \to \max, \\ \sum\limits_{i=1}^{n} p_i x_i = I. \end{cases}$$

Розв'язком задачі споживчого вибору є функція попиту — вектор $\boldsymbol{x}^0 = \boldsymbol{x}^0(\boldsymbol{p}, I) = (x_1^0(\boldsymbol{p}, I), x_2^0(\boldsymbol{p}, I), \dots, x_n^0(\boldsymbol{p}, I))$, кожна координата якого залежить від вектора цін $\boldsymbol{p} = (p_1, p_2, \dots, p_n)$ і доходу I.

Розглянемо задачу на умовний екстремум для функції двох змінних

$$\begin{cases} U(\boldsymbol{x}) = \frac{1}{2} x_1^{1/2} x_2^{1/2} \to \max, \\ p_1 x_1 + p_2 x_2 = I. \end{cases}$$

Цільова функція є опуклою. Тому, як вказувалося раніше, при дослідженні функції на екстремум достатні умови можна не використовувати. Складемо функцію Лагранжа:

$$L(x_1, x_2) = \frac{1}{2} x_1^{1/2} x_2^{1/2} - \lambda(p_1 x_1 + p_2 x_2 - I)$$

і запишемо систему рівнянь із перших похідних і рівняння зв'язку:

$$\begin{cases} L'_x = \frac{1}{4} x_1^{-1/2} x_2^{1/2} = \lambda p_1, \\ L'_y = \frac{1}{4} x_1^{1/2} x_2^{-1/2} = \lambda p_2, \\ p_1 x_1 + p_2 x_2 = I. \end{cases}$$

Поділивши перше рівняння на друге, прийдемо до системи

62a

62. Вектори в тривимірному просторі

У тривимірному просторі вектор задається трьома числами: $\boldsymbol{a} = (x_1, y_1, z_1)$. Лінійні операції над векторами $\boldsymbol{a} = (x_1, y_1, z_1)$ і $\boldsymbol{b} = (x_2, y_2, z_2)$ у просторі аналогічні операціям над векторами на площині:

1) $|\boldsymbol{a}| = \sqrt{x_1^2 + y_1^2 + z_1^2}$;
2) $\boldsymbol{a} \pm \boldsymbol{b} = (x_1 \pm x_2, y_1 \pm y_2, z_1 \pm z_2)$;
3) $\lambda \cdot \boldsymbol{a} = (\lambda x_1, \lambda y_1, \lambda z_1)$;
4) $(\boldsymbol{a}, \boldsymbol{b}) = |\boldsymbol{a}||\boldsymbol{b}|\cos\varphi = x_1 x_2 + y_1 y_2 + z_1 z_2$;
5) $\cos\varphi = \frac{(\boldsymbol{a}, \boldsymbol{b})}{|\boldsymbol{a}||\boldsymbol{b}|} = \frac{x_1 x_2 + y_1 y_2 + z_1 z_2}{\sqrt{x_1^2 + y_1^2 + z_1^2} \cdot \sqrt{x_2^2 + y_2^2 + z_2^2}}$;
6) направляючі косинуси вектора $\boldsymbol{a}$ дорівнюють:

$$\cos\alpha = \frac{x_1}{\sqrt{x_1^2 + y_1^2 + z_1^2}},$$

$$\cos\beta = \frac{y_1}{\sqrt{x_1^2 + y_1^2 + z_1^2}}, \quad \cos\gamma = \frac{z_1}{\sqrt{x_1^2 + y_1^2 + z_1^2}},$$

де α, β, γ — кути між вектором $\boldsymbol{a}$ та додатними напрямами осей Ox, Oy, Oz відповідно (рис. 1), причому $\cos^2\alpha + \cos^2\beta + \cos^2\gamma = 1$.

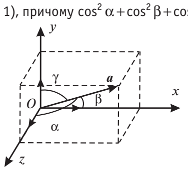

Рис. 1

Зауваження. Нехай вектор $\boldsymbol{a} = (x_1, y_1, z_1)$ є одиничним, тобто $\sqrt{x_1^2 + y_1^2 + z_1^2} = 1$. В цьому випадку $\cos\alpha = x_1, \cos\beta = y_1, \cos\gamma = z_1$. Отже, одиничний вектор повністю задається своїми направляючими косинусами $\boldsymbol{a} = (\cos\alpha, \cos\beta, \cos\gamma)$.

Нехай задані вектори *одиничної* довжини $\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}$, спрямовані уздовж осей x, y, z відповідно (рис. 1). Вони мають наступні властивості.

63a

63. Метод Лагранжа знаходження умовного екстремуму

Означення. *Умовним екстремумом* функції $z = z(x, y)$ називається максимум або мінімум функції, який досягається за умови, що її аргументи зв'язані деяким рівнянням $g(x, y) = 0$:

$$\begin{cases} z = z(x, y), \\ g(x, y) = 0. \end{cases}$$

Рівняння зв'язку $g(x, y) = 0$ визначає величину y як функцію $y = y(x)$ від змінної x. При підстановці $y = y(x)$ у досліджувану функцію $z = z(x, y)$ одержимо функцію однієї змінної ($z = z(x, y(x))$, похідна якої в точці можливого умовного екстремуму $M_0(x_0, y_0)$ дорівнює нулю, або, що рівносильно цьому, повинен дорівнювати нулю диференціал функції:

$$dz\big|_{M_0} = z'_x\big|_{M_0} dx + z'_y\big|_{M_0} dy = 0. \qquad (1)$$

З рівняння зв'язку $g(x, y) = 0$ співвідношення між диференціалами аргументу й функції в будь-якій точці, а отже, і в точці M_0, визначається рівністю

$$g'_x\big|_{M_0} dx + g'_y\big|_{M_0} dy = 0. \qquad (2)$$

Помножимо рівність (2) на деяке число λ (множник Лагранжа), яке буде визначено нижче, і додамо до рівності (1):

$$(z'_x + \lambda g'_x)_{M_0} dx + (z'_y + \lambda g'_y)_{M_0} dy = 0. \qquad (3)$$

Ліворуч у рівності (3) стоїть диференціал $d(x, y)$ функції $L(x, y) = z(x, y) + \lambda g(x, y)$, що називають функцією Лагранжа. Рівність $d(x, y) = 0$ при вивченні нами локального екстремуму була отримана, виходячи з того, що $L'_x = 0, L'_y = 0$. Це давало необхідні умови локального екстремуму. У рамках логічного ланцюжка міркувань про локальний екстремум функції $L(x, y)$ будемо вимагати виконання умов $L'_x = 0, L'_y = 0$ у точці M_0. З цією метою виберемо множник λ так, щоб виконувалася рівність

$$L'_y\big|_{M_0} = (z'_y + \lambda g'_y)_{M_0} = 0. \qquad (4)$$

Тоді зі співвідношення $(z'_x + \lambda g'_x)_{M_0} dx = 0$ одержимо

$$L'_y\big|_{M_0} = (z'_x + \lambda g'_x)_{M_0} = 0. \qquad (5)$$

64a

64. Використання квадратичної форми для дослідження на локальний екстремум

Другий диференціал функції декількох змінних можна розглядати як симетричну квадратичну форму відносно диференціалів змінних. Її частковим випадком є другий диференціал від функції двох змінних, розглянутий вище. Другий диференціал функції n змінних $u = u(x_1, x_2, \dots, x_n)$ має вигляд

$$d^2 u = \left(\frac{\partial}{\partial x_1} dx_1 + \frac{\partial}{\partial x_2} dx_2 + \ldots + \frac{\partial}{\partial x_n} dx_n\right)^2 u.$$

Матриця цієї квадратичної форми, складена із других часткових похідних функції $u = u(x_1, x_2, \dots, x_n)$, називається *матрицею Гессе**.

$$H = \begin{pmatrix} u''_{x_1x_1} & u''_{x_1x_2} & \dots & u''_{x_1x_n} \\ u''_{x_2x_1} & u''_{x_2x_2} & \dots & u''_{x_2x_n} \\ \dots & \dots & \dots & \dots \\ u''_{x_nx_1} & u''_{x_nx_2} & \dots & u''_{x_nx_n} \end{pmatrix}.$$

Як і у випадку двох змінних, приріст функції Δu у точці можливого екстремуму $M_0(x_1^0, x_2^0, \dots, x_n^0)$ розкладемо в ряд Тейлора:

$$\Delta z\big|_{M_0} = \underbrace{dz\big|_{M_0}}_{=0} + \frac{1}{2!} d^2 z\Big|_{M_0} + \boldsymbol{o}(\rho^2).$$

Знак приросту $\Delta z\big|_{M_0}$ при неперервності часткових похідних до другого порядку включно в точці M_0 і її деякого околу буде визначатися знаком другого диференціала, що є квадратичною формою. Умови знаковизначеності квадратичної форми відомі як критерій Сильвестра. Нагадаємо, що критерій Сильвестра ґрунтується на вивченні знаків кутових мінорів матриці квадратичної форми. На основі цих міркувань сформулюємо теорему.

ТЕОРЕМА (про дослідження на екстремум за другим диференціалом)

Нехай всюди в околі точки $M_0(x_1^0, x_2^0, \dots, x_n^0)$:

1) визначена функція $u = u(x_1, x_2, \dots, x_n)$;

* Насправді, розглядаючи теорію алгебраїчних ліній на площині й алгебраїчних поверхонь, професор Гессе ввів поняття визначника подібної матриці, названого пізніше *гессианом*.