ЛЕКЦІЇ З ЛІНІЙНОЇ АЛГЕБРИ

I КУРС   I СЕМЕСТР

СЕРГІЙ ОВСІЄНКО

28 грудня 2006 р.

Зміст

1 Елементи теорії множин 5

1.1 Множини, елементи множин . . . 5
1.2 Способи, якими задаються множини . . . 5
1.3 З якими множинами ми працюватимемо . . . 6
1.4 Операції над множинами . . . 6
1.5 Властивості операцій над множинами . . . 7
1.6 Сімейства множин . . . 7
1.7 Прямий або декартів добуток . . . 7
1.8 Відношення еквівалентності . . . 8
1.9 Поняття фактормножини . . . 9
1.10 Відображення . . . 9
1.11 Образ, прообраз відображення . . . 10
1.12 Композиція відображень . . . 10
1.13 Типи відображень . . . 11
1.14 Асоціативність операцій композиції відображень . . . 11
1.15 Тотожне відображення . . . 12

2 Комплексні числа 13

2.1 Теорема-означення . . . 13
2.2 Операції над комплексними числами . . . 13
2.3 Властивості операцій над комплексними числами . . . 14
2.4 Означення поля . . . 18
2.5 Операція віднімання . . . 20
2.6 Операція комплексного спряження . . . 21
2.7 Тригонометрична форма комплексного числа . . . 22
2.8 Полярні координати . . . 23
2.9 Операції над комплексними числами, заданими у тригонометричній формі 24
2.10 Корені з одиниці . . . 25

3 Системи лінійних рівнянь **29**
3.1 Лінійні рівняння 29
3.2 Системи лінійних рівнянь 30
3.3 Координатний числовий векторний простір 31
3.4 Елементарні перетворення рядків матриці 33
3.5 Метод Гауса (Жордана-Гауса) розв'язування систем лінійних рівнянь . 37
3.5.1 Приклади розв'язування систем лінійних рівнянь 37
3.5.2 Однорідні лінійні системи 40

4 Лінійне відображення, пов'язане з матрицею **42**
4.1 Основні властивості множення матриці на вектор. 42
4.2 Поняття лінійної комбінації. 44
4.2.1 Множення матриці на вектор та лінійні комбінації. 45
4.3 Лінійне відображення координатних векторних просторів, пов'язане з матрицею. 45
4.3.1 Властивості відображення L_A. 45

5 (Абстрактний) векторний простір **47**
5.1 Підпростори 50
5.1.1 Приклади підпросторів 50
5.1.2 Поняття лінійної оболонки 51
5.1.3 Терміни та позначення 51
5.2 Лінійні відображення векторних просторів 52
5.2.1 Властивості лінійних відображень 53
5.3 Базис та розмірність 58
5.3.1 Допоміжні твердження 60
5.3.2 Приклади-наслідки 61
5.4 Векторні простори типу $\mathbb{K}^n$ 62
5.5 Ранг системи векторів 66

6 Системи лінійних рівнянь **70**
6.1 Будова множини розв'язків сумісної системи 71
6.2 Базиси і лінійні відображення 72
6.3 Множення матриць 74
6.4 Теорема про ранг матриці 78
6.5 Елементарні матриці. Матричні одиниці 82

7 Детермінант (визначник) (полілінійна алгебра) **85**
7.1 Поняття групи 85
7.2 Приклади груп 85
7.3 Деякі властивості груп 86
7.4 Група підстановок – повна симетрична група 87
7.5 Детермінант 92

7.6 Поняття мінору та алгебраїчного доповнення 99

Вкажемо одразу деякі загальноприйняті математичні позначення, які спрощують запис математичних тверджень.

$\overset{df}{=}$ або $\overset{def}{=}$ означає: те, що стоїть з правого боку від знаку рівності, нам вже відомо і є означенням того, що стоїть зліва (*is defined*).

$\Longrightarrow$ - (достатності) : з того, що стоїть зліва, випливає те, що стоїть справа.

$\Longleftarrow$ - символ імплікації (необхідності) : з того, що стоїть справа, випливає те, що стоїть зліва.

$\Longleftrightarrow$ або $\equiv$ - символ еквівалентності (тоді і тільки тоді)

$\forall$ - квантор загальності

$\exists$ - квантор існування

$\neg$ - операція заперечення

$\wedge$ - логічне і

$\vee$ - логічне або

$\mathrm{id}_S : S \to S$ - тотожне відображення множини S в себе

1 Елементи теорії множин

Натуральні числа визначив Бог,
все решта - витвір рук людських.
Кронекер.

1.1 Множини, елементи множин

Множина - це поняття, яке не визначається. Його можна лише пояснити як сукупність елементів, які задовольняють деяким властивостям. Наприклад, множина всіх цілих чисел, всіх дійсних чисел.

Множина складається з *елементів*, причому вважається, що кожен елемент представлений в цій множині одному екземплярі. Запис $x \in X$ означає, що елемент x є елементом множини X, або ще кажуть, що елемент x належить до множини X.

Застереження 1.1. *Не кожну множину можна коректно визначити. Наприклад, перукар, який голить тих, хто не голиться сам, і не голить тих, хто сам голиться, має проблеми із собою. Множина всіх множин, які не є власними підмножинами, не існує, тому що припущення про її існування приводить до суперечності. Такі логічні суперечності усунені в більш детально розроблених аксіоматизаціях теорії множин. Ми не маємо змоги і часу заглибитись в такі теорії. Більш детально ці питання розглядаються в курсі математичної логіки.*

1.2 Способи, якими задаються множини

Множини можна вказувати безпосереднім, можливо нескінченним, перерахуванням їх елементів. Наприклад,

$$X = \{ \text{ стіл, за яким я зараз сиджу,} \tag{1}$$

$$\text{стілець, на якому я зараз сиджу } \}, \tag{2}$$

або $T = \{1, 2, \ldots\}$.

Ще один спосіб задання множин - за допомогою властивості. При цьому записують: $X = \{\, x \mid \mathcal{P}(x) \,\}$, де $\mathcal{P}(x)$ - це описання деякої властивості. Наприклад, $X = \{x \mid x$ - людина і вага її ≥ 83кг $\}$.

Означення 1.1. *Якщо задана деяка властивість $\mathcal{P}(x)$, де $x \in X$, то*

- $\forall x \in X\ \mathcal{P}(x) \iff$ *для кожного елемента x з множини X виконується $\mathcal{P}(x)$;*
- $\exists x \in X\ \mathcal{P}(x) \iff$ *існує $x \in X$, для якого $\mathcal{P}(x)$ виконується;*
- $\neg\mathcal{P}(x) \iff$ *для елемента x властивість $\mathcal{P}(x)$ не виконується.*

Приклади: $\forall x \in \mathbb{R}\ x^2 \geqslant 0$; $\exists x \in \mathbb{Z},\ \neg\,(x \in \mathbb{N})$.

1.3 З якими множинами ми працюватимемо

- $\mathbb{N} = \{1, 2, 3, \ldots\}$ - множина натуральних чисел.
- $\mathbb{N} = \{x \in \mathbb{Z} | x > 0\}$ - інше означення множини натуральних чисел.
- $\mathbb{Z} = \{0, \pm 1, \pm 2, \ldots\}$ або $\mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\}$ - множина цілих чисел.
- $\mathbb{Q}$ - множина раціональних чисел.
- $\mathbb{R}$ - множина дійсних чисел.
- $\mathbb{C}$ - множина комплексних чисел.
- $\varnothing$ - порожня множина, яка не містить жодного елемента.

1.4 Операції над множинами

- *Включення множин.* Говоримо, що $X \subset Y$ або X є підмножиною (в) Y, якщо кожен елемент $x \in X$ є елементом Y. В цьому випадку виконується така умова: $x \in X \Longrightarrow x \in Y$.

  Очевидно, що $X \subset X$.

  Включення $X \subset Y$ називається строгим, а X називається власною підмножиною Y, якщо існує $y \in Y$, такий що $y \notin X$.

  Іноді, щоб підкреслити, що включення $X \subset Y$ не є строгим, пишуть $X \subseteq Y$.

- *Об'єднання множин.* Говоримо, що елемент z належить об'єднанню $X \cup Y$ множин X і Y, якщо $z \in X$ або $z \in Y$. Це означення можна ще записати в скороченій формі:
$$z \in X \cup Y \iff z \in X \ \vee \ z \in Y.$$

- *Перетин множин.*
$$z \in X \cap Y \iff z \in X \ \wedge \ z \in Y.$$

  При цьому можливо, що $X \cap Y = \varnothing$, тобто множини X і Y не перетинаються.

- *Різниця множин.*
$$z \in X \setminus Y \iff z \in X \ \wedge \ z \notin Y.$$

  Зокрема, $X \setminus Y = \varnothing$ тоді і тільки тоді, коли $X \subset Y$.

- *Прямий (або декартів) добуток.* Прямий добуток двох множин задається як множина пар:
$$X \times Y = \{(x, y) \mid x \in X, y \in Y\}.$$

1.5 Властивості операцій над множинами

Асоціативність.

$$(X \cup Y) \cup Z = X \cup (Y \cup Z)$$
$$(X \cap Y) \cap Z = X \cap (Y \cap Z)$$

Ця властивість дає змогу поширити операцію на кілька елементів. Наприклад, операція віднімання на множині цілих чисел не є асоціативною, тому не можна визначити різницю трьох чисел. Маємо: $5 = 6-(3-2) \neq (6-3)-2 = 1$. Операції ж додавання та множення цілих чисел асоціативні, тому можна говорити про суму (добуток) кількох доданків (множників).

Комутативність.

$$X \cup Y = Y \cup X, \quad X \cap Y = Y \cap X.$$

Взаємна дистрибутивність. Дистрибутивність операцій об'єднання і перетину визначається формулами:

$$X \cap (Y \cup Z) = (X \cap Y) \cup (X \cap Z);$$
$$X \cup (Y \cap Z) = (X \cup Y) \cap (X \cup Z).$$

1.6 Сімейства множин

Деякі операції над множинами (наприклад, об'єднання, перетин) можна застосовувати до довільного сімейства множин.

Нехай I - довільна множина індексів, і для кожного $i \in I$ визначена множина X_i. В цьому випадку кажуть, що визначено сімейство множин X_i, $i \in I$.

Застосувавши операції об'єднання і перетину до сімейства множин, одержимо множини $\bigcup\limits_{i \in I} X_i$ та $\bigcap\limits_{i \in I} X_i$, які визначаються наступними правилами:

$$x \in \bigcup_{i \in I} X_i \iff \exists i \in I \;\; x \in X_i;$$
$$x \in \bigcap_{i \in I} X_i \iff \forall i \in I \;\; x \in X_i.$$

1.7 Прямий або декартів добуток

Прямий (або декартів) добуток $X = \prod\limits_{i \in I} X_i$ сімейства множин X_i, $i \in I$, визначається як множина, що складається з усіх можливих наборів $(x_i)_{i \in I}$, де $x_i \in X$.

Приклад 1.1. *Нехай* $I = \{1, 2, \ldots, n\} = \{i \in \mathbb{N} \,|\, i \leqslant n\}$, *де* $n \in \mathbb{N}$ *- деяке натуральне число. Прямий добуток сімейства множин* X_i, $i \in I$, *визначається як:*

$$X_1 \times X_2 \times \ldots \times X_n = \prod_{i \in I} X_i = \{(x_1, x_2, \ldots x_n) \,|\, x_1 \in X_1, \ldots,\ x_n \in X_n\}.$$

Іноді такий добуток позначається $\prod\limits_{i=1}^{n} X_i$.

Якщо всі множини-співмножники рівні між собою, тобто $X = X_1 = X_2 = \ldots = X_n$, то позначимо $X_1 \times X_2 \times \ldots \times X_n = X \times X \times \ldots \times X = X^n$. З цієї точки зору $X^1 = X$.

Зауваження 1.1. *Набори* $(x_1, x_2, \ldots x_n)$ *вважаються впорядкованими, тобто важливі не лише елементи, але і порядок слідування. Отже:* $(x_1, x_2, \ldots x_n) = (x'_1, x'_2, \ldots x'_n)$ *тоді і лише тоді, коли* $x_1 = x'_1$, $x_2 = x'_2$, $\ldots$ $x_n = x'_n$.

Приклад 1.2. *Нехай* $I = \{1, 2\}$, $X_1 = \mathbb{R}$, $X_2 = \mathbb{R}$, *де* $\mathbb{R}$ *- дійсна пряма. Тоді прямий добуток* $\mathbb{R} \times \mathbb{R}$ *- це множина пар* $\{(x_1, x_2) \,|\, x_1 \in \mathbb{R}, x_2 \in \mathbb{R}\}$. *Два елементи* $(0, 5)$, $(5, 0)$ $\in \mathbb{R} \times \mathbb{R}$ *- різні, тобто* $(0, 5) \neq (5, 0)$. *З геометричної точки зору* $\mathbb{R} \times \mathbb{R}$ *можна інтерпретувати як площину, де* x_1, x_2 *- координати. Аналогічно* $\mathbb{R} \times \mathbb{R} \times \mathbb{R}$ *- тривимірний простір. Питання: чи існує чотиривимірний простір?*

Властивість прямого добутку. Якщо в добутку хоча б одна множина порожня, то і весь добуток порожній, тобто $X \times \varnothing = \varnothing$. Якщо ж $X_1 \neq \varnothing$, $X_2 \neq \varnothing$, …, $X_n \neq \varnothing$, то $X_1 \times X_2 \times \ldots \times X_n \neq \varnothing$.

Вважається, що можна перемножувати довільну кількість множин.

Парадокс Тарського про сферу. Сферу радіуса 2 можна розрізати на скінченну кількість частин і скласти з них сферу радіуса 1.

1.8 Відношення еквівалентності

Введемо нове поняття, яким користуватимемось надалі.

Означення 1.2. *Припустимо, що* X *- деяка множина. Тоді* $n-$арним відношенням *на множині* X *називається деяка підмножина* $R \subset X^n = X \times X \times \ldots \times X$ *(n разів), де n - деяке натуральне число.*

Бінарним *відношенням (або* $2-$*арним) на множині* X *називається деяка підмножина* $R \subset X^2 = X \times X$, тернарним *(або* $3-$*арним) - підмножина* $R \subset X^3$. *Позначення: якщо* $(x, y) \in X \times X$ *така пара, що* $(x, y) \in R$, *то це позначається на письмі* $R(x, y)$ *або* xRy, $x \underset{R}{\sim} y$, $x \sim y$.

Зауваження 1.2. *Іноді кажуть, що задано відношення* $\sim$ *на множині* X, *маючи на увазі підмножину* $R = \{(x, y) \,|\, x \sim y\} \subset X^2$.

Приклад 1.3. *Нехай P - це множина людей. Визначимо таке тернарне відношення $R \subset P \times P \times P$:* $\quad R = \{(батько, мати, дитина)\}$*. Коли ми зафіксуємо третю координату, то перші дві визначені однозначно.*

Означення 1.3. *Бінарне відношення $R \subset X \times X$ називається* еквівалентністю *(або відношенням еквівалентності), якщо виконуються такі властивості:*

1. $\forall\, x \in X \;\; x \sim x$ *- рефлексивність;*

2. $\forall\, x, y \in X \;\; (x \sim y) \iff (y \sim x)$ *- симетричність;*

3. $\forall\, x, y, z \in X \;\; (x \sim y) \wedge (y \sim z) \implies (x \sim z)$ *- транзитивність (тобто еквівалентність можна перевірити локально, на невеликій віддалі).*

Множина $S_x = \{y \in X \,|\, y \sim x\}$ елементів, еквівалентних деякому $x \in X$, називається класом еквівалентності елемента x. Довільний елемент класу S_x називається представником цього класу.

Завдання 1.1. *Довести, що $y \in S_x \iff S_x = S_y$ для $x, y \in X$.*

1.9 Поняття фактормножини

Якщо $\sim$ - деяка еквівалентність на множині X, то через $X/_\sim$ позначимо множину класів еквівалентності, тобто $X/_\sim = \{S_x \,|\, x \in X\}$. Множина $X/_\sim$ називається *фактормножиною (quotient)* множини X (відносно еквівалентності $\sim$) .

Приклад 1.4. *Нехай X - деяка множина людей, які дотримуються принципів моногамності, і $\sim$ - відношення "бути в шлюбі". Фактормножиною на множині X відносно відношення $\sim$ є множина усіх подружь.*

1.10 Відображення

Нехай X, Y - дві множини. Відображенням F з множини X до множини Y називається правило, за яким кожному елементу $x \in X$ ставиться у відповідність елемент $F(x)$ множини Y. Позначається відображення: $F : X \to Y$ або $X \xrightarrow{F} Y$.

Приклади відображень.

1) Функції від дійсної змінної:

$$\left\{ \begin{array}{rccc} \sin: & \mathbb{R} & \to & \mathbb{R}; \\ & x & \longmapsto & \sin x. \end{array} \right. \qquad \left\{ \begin{array}{rccc} x^2: & \mathbb{R} & \to & \mathbb{R}; \\ & x & \longmapsto & x^2. \end{array} \right.$$

2) Ім'я: $P \to N$, де P - множина всіх людей, N - множина всіх мислимих імен. При цьому вважаємо, що у людини є головне ім'я.

Відображення можуть задаватися у вигляді таблиці або картинки.

1.11 Образ, прообраз відображення

Якщо задана функція або відображення $f : X \to Y$, то для елемента $x \in X$ елемент $f(x) \in Y$ - це *образ* x. Більш загально, коли $X' \subset X$ - деяка підмножина, тоді множина

$$f(X') \ = \ \{\, f(x) \,|\, x \in X' \,\} \subset Y$$

називається образом підмножини X'.

Для елемента $y \in Y$ елемент $x \in X$ такий, що $f(x) = y$, називається *прообразом.*

Зауваження 1.3. *1) Прообраз не завжди існує. Приклад: коли* $x \to x^2$, *то* $-1 \in \mathbb{R}$ *не має прообразу.*

2) Прообразів може бути кілька чи навіть нескінченно багато. Так, якщо $x \to x^2$, *то* $1 \in \mathbb{R}$ *має два прообрази* $1, -1 \in \mathbb{R}$.

Позначення. Для елемента $x \in X$ твердження, що $f(x) = y \in Y$ записується так: $x \mapsto y$ або $x \xrightarrow{f} y$, якщо треба підкреслити, яка саме функція.

Для кожного $y \in Y$ *повний прообраз* $f^{-1}(y) = \{x \in X \,|\, f(x) = y\}$ визначається як множина таких $x \in X$, що $f(x) = y$. Наприклад, $\sin^{-1}(0) = \{k\pi \,|\, k \in \mathbb{Z}\}$.

Повний прообраз підмножини $Y' \subset Y$ визначається як множина

$$f^{-1}(Y') = \{x \in X \,|\, f(x) \in Y'\} \subset X.$$

Визначимо *образ відображення* $f : X \to Y$ як таку підмножину в Y:

$$\textsc{Im}\, f = f(X) \ = \ \{y \in Y \mid \exists x \in X, f(x) = y\}.$$

Зауважимо, що поняття прообразу відображення не носить інформативного характеру, тому що це завжди буде вся множина X.

1.12 Композиція відображень

Припустимо, що задані відображення $f : X \longrightarrow Y$ і $g : Y \longrightarrow Z$. *Композицією* (або суперпозицією) цих відображень називається відображення $gf : X \longrightarrow Z$, задане за таким правилом: $(gf)(x) = g(f(x))$. Схематично це зображується так:

$$X \xrightarrow{f} Y \xrightarrow{g} Z.$$

Зауваження 1.4. *Суперпозиція функцій* f *і* g *записується в протилежному порядку як* gf, *оскільки історично сформувалася форма запису функції як* $f(x)$. *В деяких мовах програмування функція записується справа:* $(x)f$; *в цьому випадку маємо* $(x)(fg) = ((x)f)g$.

1.13 Типи відображень

- Відображення $f : X \to Y$ називається *ін'єктивним* (або *мономорфізмом*), якщо для $x_1, x_2 \in X$ таких, що $x_1 \neq x_2$, їх образи також різні, тобто $f(x_1) \neq f(x_2)$. Інакше кажучи, відображення ін'єктивне, якщо воно різні елементи переводить в різні.

- Відображення $f : X \to Y$ називається *сюр'єктивним* (або *епіморфізмом*), якщо $f(X) = Y$, або, що еквівалентно, $\forall y \in Y\ \exists x \in X\ f(x) = y$.

- Відображення $f : X \to Y$ називається *бієктивним* (або *ізоморфізмом*), якщо воно ін'єктивне і сюр'єктивне.

Приклад 1.5. *Розглянемо множину X, що складається з трьох елементів, і двоелементну множину Y. Тоді не існує ін'єктивного відображення $X \to Y$? І навпаки, не існує сюр'єктивного відображення $Y \to X$.*

Бієктивне відображення або бієкція $X \to Y$ називається *взаємно однозначною відповідністю* між множинами X, Y. Зауважимо, що в даному випадку множини X і Y не впорядковані і у нас фактично є два відображення: пряме і обернене, визначення якого дамо пізніше.

Приклад 1.6. *Якщо $\sim$ - деяка еквівалентність на множині X і через $X/_\sim$ позначена фактормножина множини X відносно еквівалентності $\sim$, то існує канонічне сюр'єктивне відображення $\pi : X \longrightarrow X/_\sim$, при якому $\pi(x) = C_x$, $x \in X$, тобто кожен елемент переходить в свій клас еквівалентності.*

1.14 Асоціативність операцій композиції відображень

Нехай для множин X, Y, Z, T задані відображення $f : X \to Y$, $g : Y \to Z$ і $h : Z \to T$, тобто визначена послідовність $X \xrightarrow{f} Y \xrightarrow{g} Z \xrightarrow{h} T$. Тоді має місце рівність:

$$h(gf) = (hg)f.$$

Доведемо цю властивість, незважаючи на її очевидність.

Доведення. Розглянемо довільний елемент $x \in X$ і подіємо на нього почергово послідовностями відображень, використовуючи означення композиції:

$$h((gf)(x)) = h(g(f(x))); \qquad (hg)f(x) = (hg)(f(x)) = h(g(f(x))).$$

Оскільки одержані праві частини збігаються, то на цьому доведення закінчується. $\square$

Ця властивість композицій відображень має назву асоціативності. Завдячуючи їй, ми можемо записувати композиції відображень без дужок, а саме, позначимо

$$hgf \overset{df}{=} h(gf) = (hg)f.$$

1.15 Тотожне відображення

Для будь-якої множини X існує так зване *тотожне відображення* $\mathtt{1}_X : X \to X$ (або $\mathtt{1}_X$, id_X). Це відображення задається за таким правилом: $\forall x \in X \quad \mathtt{1}_X(x) = x$. Очевидно, тотожне відображення бієктивне.

Для довільного відображення $f : X \to Y$ виконуються такі рівності:

$$\mathtt{1}_Y f = f, \qquad f\mathtt{1}_X = f.$$

Доведення очевидне.

Лемма 1.1. *Наступні твердження є рівносильними (або еквівалентними):*

1. *$f : X \to Y$ є бієкція;*
2. *для $f : X \to Y$ існує $g : Y \to X$ таке, що $gf = 1_X$ і $fg = 1_Y$, причому якщо таке g існує, то воно єдине.*

Зауваження 1.5. *Щоб довести рівносильність двох тверджень, припускають, що вірне перше твердження і доводять друге, позначається це $1 \Longrightarrow 2$. І навпаки, виводимо із другого твердження перше, тобто $2 \Longrightarrow 1$.*

Доведення. $1 \Longrightarrow 2$. Припустимо, що f - бієкція. Визначимо $g : Y \to X$ таке, щоб $fg = 1_Y$. Для цього розглянемо $y \in Y$. Тоді існує (бо f - сюр'єкція), причому єдиний (бо f - ін'єкція) елемент $x \in X$, такий що $f(x) = y$. Тоді покладемо $g(y) = x$. Це коректно визначене правило, бо такий елемент єдиний. Більше того, якщо ми хочемо, щоб виконувалось $fg = \mathtt{1}_Y$, то це єдиний спосіб визначити відображення g. Дійсно, $fg(y) = \mathtt{1}_Y(y) = y$ для всякого y. Тому, якщо $g(y) = x'$, то $f(x') = f(g(y)) = y$. Отже, вказане g визначено однозначно і властивість $fg = \mathtt{1}_Y$ виконана. Перевіримо, що $gf = \mathtt{1}_X$. Це еквівалентно тому, що $\forall x \in X \ gf(x) = \mathtt{1}_X(x)$. Насправді це те саме, що $gf(x) = x$, що еквівалентно умові $g(f(x)) = x$.

Нехай $f(x) = y$. Тоді, за означенням, $g(y)$ буде такий, що $\exists x' \in X \ f(x') = y$. Але ми такий елемент x знаємо, а саме $x' = x$. Отже, звідси $g(y) = x$, що еквівалентно умові $g(f(x)) = x$.

$2 \Longrightarrow 1$ Навпаки, припустимо, що існує g таке, що $fg = \mathtt{1}_y$ і $gf = \mathtt{1}_X$. Потрібно довести, що f - не ін'єктивне, тобто існують такі $x_1 \neq x_2$, що $f(x_1) = f(x_2)$. Позначимо цей елемент через y. Запишемо рівність: $x_1 = \mathtt{1}_X(x_1) = gf(x_1) = g(f(x_1)) = g(y)$. Отже, $x_1 = g(y)$. Далі застосовуємо цю саму рівність до x_2. Маємо: $x_2 = \mathtt{1}_X(x_2) = gf(x_2) = g(f(x_2)) = g(y)$. Отже, $x_1 = g(y) = x_2 \implies x_1 = x_2$. Це суперечить припущенню, що вони різні. Отже, f - ін'єктивне.

Чому f сюр'єктивне? Треба довести, що $\forall y \ \exists x \in X \ y = f(x)$. Запишемо рівність $\mathtt{1}_Y = f(y)$. Тоді $y = \mathtt{1}_Y(y) = (fg)(y) = f(g(y))$. Позначимо $g(y) = x$. Тобто, ми побудували такий елемент x, який переходить в y. □

Означення 1.4. *Відображення $g : Y \to X$, що задовольняє умовам попередньої леми, називається оберненим до відображення f і позначається $g = f^{-1}$. Отже, f^{-1} існує і едине тоді і тільки тоді, коли - f бієкція.*

Вправи. 1. $(f^{-1})^{-1} = f$.

2. Припустимо, що $X \xrightarrow{f} Y \xrightarrow{g} Z$ і існують f^{-1}, g^{-1}. Тоді існує $(gf)^{-1}$ і $(gf)^{-1} = g^{-1}f^{-1}$.

2 Комплексні числа

Теорія комплексних чисел розглядається як концептуальна база для подальшого просування в алгебру. На цьому прикладі застосуємо більшість мовних засобів сучасної алгебри.

В рамках натуральних чисел не можна платити борги, отже виникли цілі числа. Додали ділення - одержали раціональні числа, вимірювання відрізків - одержали дійсні числа. Комплексні числа виникли з необхідності розв'язувати квадратичні рівняння.

2.1 Теорема-означення

Позначимо через $\mathbb{R}$ множину дійсних чисел, вважаємо, що на них визначені операції додавання $+$ та множення $\times$ і що всі властивості дійсних чисел ми "знаємо".

Множина комплексних чисел $\mathbb{C}$ визначається як множина пар дійсних чисел, тобто $\mathbb{C} = \mathbb{R} \times \mathbb{R} = \{\,(a,b) \,|\, a,b \in \mathbb{R}\,\}$. З історичних причин та причин зручності числа (a,b) записуються у формі: $a + bi$, де i - це спеціальний символ, який називається *уявною одиницею*.

На множині $\mathbb{C}$ визначені два відображення: $\mathrm{Re} : \mathbb{C} \longrightarrow \mathbb{R}$, при якому $\mathrm{Re}\,(a+bi) = a$ - *дійсна частина*, і $\mathrm{Im} : \mathbb{C} \longrightarrow \mathbb{R}$, при якому $\mathrm{Im}\,(a+bi) = bi$ - *уявна частина* комплексного числа $a + bi$.

Приклад. $\sqrt{-2} + \pi i \in \mathbb{C}$.

Деякі особливості запису комплексних чисел. Вважаємо, що $a + 0 \cdot i = a$, а також $0+bi = bi$ - еквівалентні форми запису комплексного числа. Відповідно $0+0 \cdot i = 0$. Крім того, $1i = i$, $(-k)i = -ki$, тобто $1 + (-2)i = 1 - 2i$ (остання домовленість називається “правилом знаків”).

2.2 Операції над комплексними числами

На множині комплексних чисел ми розглядатимемо три основні операції:

- додавання $+ : \mathbb{C} \times \mathbb{C} \to \mathbb{C}$,
- множення $\times : \mathbb{C} \times \mathbb{C} \to \mathbb{C}$,
- комплексного спряження $\bar{\ } : \mathbb{C} \to \mathbb{C}$.

Операція *додавання* (або сума) визначається для двох комплексних чисел $a + bi$, $c + di \in \mathbb{C}$ наступним чином: $a + bi + c + di = (a + c) + (b + d)i$.

Множення двох комплексних чисел $a + bi$, $c + di \in \mathbb{C}$ визначається за формулою:

$$(a + bi) \times (c + di) = (ac - bd) + (ad + bc)i.$$

Комплексне спряження - це відображення $\mathbb{C} \to \mathbb{C}$, яке позначається σ або $\bar{}$ і задається таким правилом: $\sigma(a + bi) = \overline{a + bi} = a - bi$.

Зауваження 2.1. *Застосування правил додавання та множення до дійсних чисел (тобто до комплексних чисел, які мають нульову уявну частину) означає їх звичайне додавання та множення. Комплексне ж спряження не змінює дійсного числа. У цьому випадку кажуть, що операції над комплексними числами узгоджені з відомими операціями над дійсними числами.*

З правила множення випливає, зокрема, таке твердження стосовно символу i : $i^2 = -1$, тобто що i - корінь рівняння $z^2 + 1 = 0$.

Ці операції задають структуру поля комплексних чисел. Ми вжили слово поле і тепер мусимо дати пояснення. Отже, що таке поле на прикладі комплексних чисел?

2.3 Властивості операцій над комплексними числами

Сформулюємо теорему, яка одночасно є означенням поняття поля.

Теорема-означення 2.1. *Комплексні числа утворюють* поле.

(1) На множині $\mathbb{C}$ комплексних чисел введено операції:

a) $+ : \mathbb{C} \times \mathbb{C} \longrightarrow \mathbb{C}$ за таким правилом: коли $\alpha, \beta \in \mathbb{C}$, то $+ : (\alpha, \beta) \mapsto \alpha + \beta$ (операція додавання);

b) $\cdot : \mathbb{C} \times \mathbb{C} \longrightarrow \mathbb{C}$, яка для $\alpha, \beta \in \mathbb{C}$ визначається: $\cdot(\alpha, \beta) \mapsto \alpha \cdot \beta$ (операція множення).

Відносно введених операцій виконуються такі аксіоми поля.

1. *Асоціативність додавання.* Операція $+$ асоціативна, тобто для $\alpha, \beta, \gamma \in \mathbb{C}$ справедливе таке твердження:

$$\boxed{(\alpha + \beta) + \gamma \quad = \quad \alpha + (\beta + \gamma)}$$ [1]

Доведення. Нехай $\alpha = a + bi$, $\beta = c + di$, $\gamma = e + fi$. Тоді

(*) $((a + bi)+ (c + di))+ e + fi = ((a + c)+ (b + d)i)+ e + fi = ((a + c) + e)+ ((b + d) + f)i$.

[1] Формула в рамочці означатиме: "Аксіома поля".

Інша сума $\quad$ (**) $(a+bi)+((c+di)+(e+fi))=((a+bi)+((c+e)+(d+f)i)=(a+(c+e))+(b+(d+f))i$.

Два числа рівні, коли рівні відповідні координати. Оскільки для дійсних чисел $((a+c)+e)=(a+(c+e))$ і $((b+d)+f)=(b+(d+f))$ (аксіома асоціативності для дійсних чисел), то і комплексні числа в виразах (*) і (**) рівні між собою. Звідси висновок: операція + асоціативна. $\square$

2. *Існування нуля* або нейтрального елемента відносно додавання. В $\mathbb{C}$ існує елемент $0\in\mathbb{C}$ такий, що

$$\boxed{\forall\alpha\in\mathbb{C}\qquad 0+\alpha=\alpha+0=\alpha}$$

Такий елемент 0 називається *нейтральним відносно додавання* або нулем.

Доведення. Елемент $0=0+0i$ буде нейтральним. Дійсно, $(0+0i)+(a+bi)=(0+a)+(0+b)i=a+bi$ тому, що в дійсних числах $0+a=a$, $0+b=b$. Аналогічно, $(a+bi)+(0+0i)=a+bi$ тому, що для дійсних чисел $\mathbb{R}$ виконується $a+0=a$, $b+0=b$. $\square$

3. *Існування протилежного комплексного числа.* Для довільного $\alpha\in\mathbb{C}$ існує (і єдине) $(-\alpha)\in\mathbb{C}$ таке, що $\alpha+(-\alpha)=(-\alpha)+\alpha=0\in\mathbb{C}$.

$$\boxed{\alpha+(-\alpha)=\quad(-\alpha)+\alpha=0\quad\in\mathbb{C}}$$

Доведення. *(Типове саме для поля $\mathbb{C}$).* Розглянемо $\alpha=a+bi$. Тоді $(-\alpha)=(-a)+(-b)i$. Перевіримо. Порахуємо $(a+bi)+((-a)+(-b)i)=(a+(-a))+(b+(-b))i=0+0\cdot i$ тому, що для дійсних чисел $a+(-a)=0$, $b+(-b)=0$.

З іншого боку, за правилом додавання комплексних чисел $(-a)+(-b)i+(a+bi)=0$ тому, що $(-a)+a=0$ і $(-b)+b=0$ в $\mathbb{R}$. $\square$

4. *Комутативність додавання.* Тобто доведемо, що для $\alpha,\beta\in\mathbb{C}$ виконується

$$\boxed{\alpha+\beta\quad=\quad\beta+\alpha}$$

Доведення. Розглянемо два числа $a+bi$, $c+di\in\mathbb{C}$. Коли ми їх додамо, то, з одного боку, маємо:

$$(a+bi)\;+\;(c+di)\;=\;(a+c)\;+\;(b+d)i.$$

З іншого боку, в протилежному порядку:

$$(c+di)\;+\;(a+bi)\;=\;(c+a)\;+\;(d+b)i.$$

Одержані комплексні числа рівні, тому що в $\mathbb{R}$ виконуються рівності $a+c=c+a,\quad b+d=d+b$. $\square$ [2]

[2] Сукупність цих властивостей говорить про те, що комплексні числа утворюють комутативну абелеву групу, але це поняття буде введене пізніше.

5. *Асоціативність множення.* А саме, якщо α, β, $\gamma \in \mathbb{C}$, тоді

$$\boxed{(\alpha\,\beta)\,\gamma \;\;=\;\; \alpha\,(\beta\,\gamma)}$$

Зауваження 2.1. *Ця властивість вважається очевидною для $\mathbb{R}$. Але чому множення повинно вести себе так само і для $\mathbb{C}$? Коли виписати формулу потрійного добутку, властивість комутативності перестає бути очевидною.*

Доведення. (Для поля $\mathbb{C}$). Нехай $\alpha = a + bi$, $\beta = c + di$, $\gamma = e + fi$. Розглянемо $(\alpha\beta)\gamma = ((a + bi)(c + di))(e + fi) = ((ac - bd) + (ad + bc)i)(e + fi) = ((ac - bd)e - (ad + bc)f) + ((ac - bd)f + (ad + bc)e)i$.

З іншого боку, $\alpha(\beta\gamma) = (a + bi)((c + di)(e + fi)) = (a + bi)((ce - df) + (cf + de)i) = (a(ce - df) - b(cf + de)) + (a(cf + de) - b(ce - df))i$.

Порівняємо дійсні і уявні частини: $ace - bde - adf + bcf = ace - adf - b(cf + de))$; $acf - bdf + ade + bce = acf + ade - bce - bdf$. Неймовірно, але факт: це те саме. Отже, властивість доведена. $\square$

Зауваження 2.2. *При порівнянні ми користувалися такими властивостями: 1) асоціативністю додавання та множення в $\mathbb{R}$; 2) можливістю розкривати дужки (або дистрибутивністю); 3) комутативністю додавання та множення в $\mathbb{R}$.*

6. *Існування одиниці або нейтрального елемента* відносно множення. Це означає:

$$\boxed{\exists 1 \in \mathbb{C}, \quad 1 \neq 0, \quad \forall \alpha \in \mathbb{C} \quad 1 \cdot \alpha = \alpha \cdot 1}$$

Тобто існує елемент, який як-би не діє. Наявність такого елемента дуже важлива.

Доведення. Такий елемент існує, це $1 + 0i$. Дійсно, перевіримо: $(1 + 0i)(a + bi) = (1 \cdot a - 0 \cdot b) + (1 \cdot b + 0 \cdot a)i = a + bi$, тому що в $\mathbb{R}$ виконується: $1 \cdot a = a$, $1 \cdot b = b$, $0 \cdot a = 0$, $0 \cdot b = 0$.

Аналогічно, $(a + bi)(1 + 0i) = (a \cdot 1 - b \cdot 0) + (a \cdot 0 + b \cdot 1)i = a + bi$, тому що в $\mathbb{R}$ $a \cdot 1 = a$, $a \cdot 0 = 0$, $b \cdot 1 = b$, $b \cdot 0 = 0$. $\square$

7. *Існування оберненого.* А саме, для довільного комплексного числа $\alpha \in \mathbb{C}$, такого що $\alpha \neq 0$, існує комплексне число α^{-1}, таке що $\alpha \cdot \alpha^{-1} = \alpha^{-1} \cdot \alpha = 1$. Число α^{-1} називається оберненим. Така властивість дозволяє організувати операцію ділення.

Доведення. Випишемо явно цей елемент: $(a + bi)^{-1} = \dfrac{a}{a^2 + b^2} - \dfrac{b}{a^2 + b^2}i$. Доведення полягає в прямій перевірці властивостей.

$(a + bi)\,(\frac{a}{a^2+b^2} - \frac{b}{a^2+b^2}i) = (a\frac{a}{a^2+b^2} - b(-\frac{b}{a^2+b^2})) + (a(-\frac{b}{a^2+b^2}) + (b\frac{a}{a^2+b^2}))i =$
$(\frac{a^2+b^2}{a^2+b^2} + \frac{-ab+ba}{a^2+b^2}i) = 1 + 0i.$

Тут ми знов скористалися комутативністю множення в $\mathbb{R}$. Крім того, ділення на a^2+b^2 можливе, бо $\alpha \neq 0$. Аналогічно,

$(\frac{a}{a^2+b^2} - \frac{b}{a^2+b^2}i)\ (a+bi) = (\frac{a}{a^2+b^2}a - (-\frac{b}{a^2+b^2}b)) + (\frac{a}{a^2+b^2}b + (-\frac{b}{a^2+b^2})a)i = 1 + 0 \cdot i.$ $\square$

8. *Комутативність множення.* Доведемо, що для $\alpha, \beta \in \mathbb{C}$ виконується рівність

$$\boxed{\alpha\beta \quad = \quad \beta\alpha}$$

Доведення. Маємо: $\alpha\beta = (a+bi)(c+di) = (ac-bd) + (ad+bc)i.$

Перемножимо співмножники в протилежному порядку: $\beta\alpha = (c+di)(a+bi) = (ca-db) + (cb+da)i.$

Одержані вирази рівні, це випливає з комутативності множення та додавання в $\mathbb{R}$. $\square$

9. *Дистрибутивність* або розподільний закон (*білінійність*). Для довільних α, β, $\gamma \in \mathbb{C}$ справедливі такі властивості:

$$\boxed{\begin{array}{rcl} \alpha(\beta+\gamma) & = & \alpha\beta+\alpha\gamma \\ (\alpha+\beta)\gamma & = & \alpha\gamma+\beta\gamma \end{array}}$$

Доведення.

Нехай $\alpha = a+bi,\ \ \beta = c+di,\ \ \gamma = e+fi$. Розглянемо $(a+bi)((c+di)+(e+fi)) = (a+bi)((c+e)+(d+f)i) = (a(c+e)-b(d+f)) + (a(d+f)+b(c+e))i = (ac+ae-bd-bf) + (ad+af+bc+be)i.$

З іншого боку, $(a+bi)(c+di)+(a+bi)(e+fi) = ((ac-bd)+(ad+bc)i) + ((ae-bf)+(af+be)i) = (ac-bd+ae-bf) + (ad+bc+af+be)i.$

З рівності одержаних виразів випливає справедливість першого твердження. Ми скористалися тим, що в $\mathbb{R}$ дистрибутивний закон виконується. Друге твердження доводиться аналогічно.

Завдання 2.1. *Довести друге твердження дистрибутивності. Воно випливає з першого твердження та властивості комутативності, оскільки*

$$(\alpha+\beta)\gamma \overset{(ком)}{=} \gamma(\alpha+\beta) \overset{(I)}{=} \gamma\alpha+\gamma\beta \overset{(ком)}{=} \alpha\gamma+\beta\gamma.$$

Кінець теореми-означення. $\square$

2.4 Означення поля

Означення 2.1. *Нехай* $\mathbb{K}$ *- деяка множина, на якій визначено дві операції:*
$\left\{\begin{array}{rcl} "+": \mathbb{K}\times\mathbb{K} & \to & \mathbb{K} \\ (\alpha,\beta) & \mapsto & \alpha+\beta \end{array}\right.$ *(додавання) і* $\left\{\begin{array}{rcl} "\cdot": \mathbb{K}\times\mathbb{K} & \to & \mathbb{K} \\ (\alpha,\beta) & \mapsto & \alpha\beta \end{array}\right.$ *(множення)*
і виконуються наступні властивості ("аксіоми поля"*):*

1. *асоціативність додавання*:
$$\forall\alpha,\beta,\gamma\in\mathbb{K}\quad (\alpha+\beta)+\gamma=\alpha+(\beta+\gamma);$$
2. *існування нуля* (нейтрального елемента відносно додавання):
$$\exists 0\in\mathbb{K}\quad \forall\alpha\in\mathbb{K}\ \ 0+\alpha=\alpha+0=\alpha;$$
3. *існування протилежного елемента*:
$$\forall\,\alpha\in\mathbb{K}\ \exists(-\alpha)\in\mathbb{K}\quad \alpha+(-\alpha)=(-\alpha)+\alpha=0;$$
4. *комутативність додавання*:
$$\forall\,\alpha,\beta\in\mathbb{K}\quad \alpha+\beta=\beta+\alpha;$$
5. *асоціативність множення*:
$$\forall\,\alpha,\beta,\gamma\in\mathbb{K}\quad (\alpha\beta)\gamma=\alpha(\beta\gamma);$$
6. *існування одиниці* (нейтрального елемента відносно множення):
$$\exists\,1\in\mathbb{K}, 1\neq 0,\ \forall\,\alpha\in\mathbb{K}\quad 1\cdot\alpha=\alpha\cdot 1=\alpha;$$
7. *існування оберненого елемента*:
$$\forall\,\alpha\in\mathbb{K},\alpha\neq 0,\ \exists\alpha^{-1}\in\mathbb{K}\quad \alpha\cdot\alpha^{-1}=\alpha^{-1}\cdot\alpha=1;$$
8. *комутативність множення*:
$$\forall\,\alpha,\beta\in\mathbb{K}\quad \alpha\beta=\beta\alpha;$$
9. *дистрибутивність* (білінійність):
$$\forall\,\alpha,\beta,\gamma\in\mathbb{K}\qquad \alpha(\beta+\gamma)=\alpha\beta+\alpha\gamma,\quad (\alpha+\beta)\gamma=\alpha\gamma+\beta\gamma.$$

У цьому випадку говорять, що $\mathbb{K}$ *- це поле.*

Теорема-означення 2.2 (переформулювання)**.** *Множина* $\mathbb{C} = \{a + bi \,|\, a, b \in \mathbb{R}\}$ *з операціями* "$+$" *та* "$\cdot$"*, визначеними за правилами*

- $(a + bi) + (c + di) = (a + c) + (b + d)i,$
- $(a + bi) \cdot (c + di) = (ac - bd) + (ad + bc)i,$

утворює **поле**.

Приклад 2.1. *Приклади полів, відомих зі шкільного курсу.*

- $\mathbb{K} = \mathbb{Q}$ *- поле раціональних чисел.*
- $\mathbb{K} = \mathbb{R}$ *- поле дійсних чисел.*

Означення 2.2. *Нехай* $\mathbb{K}$ *- це поле,* $\mathbb{L} \subset \mathbb{K}$ *- підмножина з такими властивостями:*

- $0, 1 \in \mathbb{L}$*;*
- $\alpha, \beta \in \mathbb{L} \implies \alpha + \beta,\ \alpha \cdot \beta \in \mathbb{L}$*;*
- $\alpha \in \mathbb{L} \implies (-\alpha) \in \mathbb{L}$*;*
- $\alpha \in \mathbb{L} \implies \alpha^{-1} \in \mathbb{L}$*.*

В цьому випадку відносно тих самих операцій "$+$" *та* "$\cdot$" *множина* $\mathbb{L}$ *знову утворює поле. Говоримо, що* $\mathbb{L}$ *є підполе в* $\mathbb{K}$*.*

Приклад 2.2.

- $\mathbb{Q} \subset \mathbb{R}$ *- підполе.*
- *Поле* $\mathbb{R}$ *можна розглядати як підполе поля комплексних чисел* $\mathbb{C}$*. Існує відображення вкладення* $s : \mathbb{R} \longrightarrow \mathbb{C}$ *таке, що дійсному числу* $a \in \mathbb{R}$ *відповідає комплексне* $s(a) = a + 0 \cdot i$*, причому наступні властивості виконуються:*

  $s(a + b) = s(a) + s(b)$*;*

  $s(ab) = s(a)s(b)$*;*

  $s(1) = 1$*, тобто* $1_{\mathbb{R}} \overset{s}{=} 1_{\mathbb{C}}$*.*

Твердження 2.1. *В полі* $\mathbb{K}$ *існує єдиний нульовий елемент.*

Доведення. $0 \in \mathbb{K}$. Припустимо, $\exists 0' \in \mathbb{K}$, який також є нейтральним елементом, тобто $\forall \alpha \in \mathbb{K}\ \ 0' + \alpha = \alpha + 0' = \alpha$. Тоді $0 + 0' = 0'$, оскільки 0 - нейтральний і $0 + 0' = 0$, оскільки $0'$ - нейтральний. З рівності лівих частин випливає, що $0 = 0'$. $\square$

Завдання 2.2.

- *Довести єдиність протилежного елемента в полі* $\mathbb{K}$*, тобто що* $\forall \alpha \in \mathbb{K}\ \ \exists!(-\alpha)$*.*
- *Довести єдиність одиниці та єдиність оберненого.*
- *Довести, користуючись означенням поля, що* $\forall \alpha \in \mathbb{K}$ *виконується* $\alpha \cdot 0 = 0 \cdot \alpha = 0$*.*

2.5 Операція віднімання

Зауваження 2.3. *З аксіом поля можна вивести практично всі звичайні властивості арифметичних дій. Зокрема, з існування нуля, одиниці, протилежного та оберненого елементів випливає можливість означити і "обернені операції" – віднімання та ділення. Крім того, з аксіом комутативності та асоціативності випливає можливість у довільних сумах і добутках переставляти доданки або множники і розміщувати дужки в будь-який спосіб. З аксіоми дистрибутивності випливають звичні правила "розкриття дужок".*

Означення 2.3. *Говоримо, що $x = a - b$ для деяких $a, b \in \mathbb{K}$ тоді і тільки тоді, коли x задовольняє рівності $x + b = a$. Оскільки $x + b = a \equiv (x + b) + (-b) = a + (-b) \equiv x + 0 = a + (-b) \equiv x = a + (-b)$, то має місце таке означення дії* **віднімання** *в полі $\mathbb{K}$:*

$$a - b \overset{def}{=} a + (-b).$$

Означення 2.4. *Кажуть, що для елементів $a, b \in \mathbb{K}$, $b \neq 0$, визначена операція* **ділення** *a на b (або дріб $\frac{a}{b}$), якщо $\exists x \in \mathbb{K}$ такий, що $x \cdot b = a$. Знайдемо формулу для x: $x \cdot b = a \equiv (x \cdot b) \cdot b^{-1} = ab^{-1} \equiv (xb) \cdot b^{-1} = x(b \cdot b^{-1}) = x \cdot 1 = x \implies x = ab^{-1}$. Отже,*

$$\frac{a}{b} \overset{def}{=} a \cdot b^{-1}.$$

Приклад 2.3. *Раціональні функції над полем $\mathbb{K}$ - це множина*

$$\mathbb{K}(x) = \left\{ \frac{f(x)}{g(x)} \mid f, g - \text{ поліноми від } x \text{ з коефіцієнтами в полі } \mathbb{K}, g(x) \neq 0 \right\}.$$

Дроби в $\mathbb{R}(x)$ можна скорочувати. Приклад, $1 = \frac{x^2}{x^2} = \frac{x^3 + 7x}{x^3 + 7x}$.

Приклад 2.4.

- *Розглянемо таке поле $\mathbb{K}$, яке складається з дійсних чисел вигляду $a + b\sqrt{2}$ з раціональними a і b, тобто $\mathbb{K} = \{a + b\sqrt{2} \mid a, b \in \mathbb{Q}\}$. Тоді $\mathbb{K} \subset \mathbb{R}$ - підполе відносно операцій $+$ та $\times$.*

Завдання 2.3.

- *Перевірити, що обернений елемент в попередньому прикладі також належить $\mathbb{K}$, тобто $\mathbb{K}$ - поле.*
- *Довести, що множина дійсних чисел $\{a + b\sqrt[3]{5} + c\sqrt[3]{25} \mid a, b, c \in \mathbb{Q}\}$ також утворює поле.*

2.6 Операція комплексного спряження

Комплексне спряження σ (або $^{-}$) - це відображення $\mathbb{C} \longrightarrow \mathbb{C}$, задане за правилом: $a+bi \mapsto a-bi$, або для $z \in \mathbb{C}$ $\quad \mathrm{Re}\,(\bar{z}) = \mathrm{Re}\,(z)$, $\mathrm{Im}\,(\bar{z}) = -\mathrm{Im}\,(z)$.

Властивості комплексного спряження

1. *Лінійність відносно додавання*: $\sigma(x+y) = \sigma(x) + \sigma(y)$, $x, y \in \mathbb{C}$.

   Дійсно, якщо $x = a + bi$, $y = c + di$, то $\sigma(x+y) = \sigma((a+bi)+(c+di)) = \sigma((a+c)+(b+d)i) = ((a+c)-(b+d)i) = (a-bi)+(c-di) = \sigma(x)+\sigma(y)$.

2. *Лінійність відносно множення*: $\sigma(xy) = \sigma(x) \cdot \sigma(y)$, $x, y \in \mathbb{C}$.

   За означенням, $\sigma(xy) = \sigma((ac-bd)+(ad+bc)i) = (ac-bd) - (ad+bc)i$.

   З іншого боку, $\sigma(x) \cdot \sigma((y) = (a-bi)\ (c-di) = (ac-bd) - (ad+bc)i = \sigma(xy)$.

3. *Дія на нейтральні елементи*: $\sigma(0) = 0$, $\sigma(1) = 1$.

4. *Інволютивність* (або кажуть, що σ є інволюція). Це означає, що $\sigma(\sigma(x)) = x$. Це так, оскільки $\sigma(\sigma(x)) = \sigma(a-bi) = a+bi$ для будь-якого $x = a+bi \in \mathbb{C}$.

   Ця властивість мовою теорії множин означає, що в результаті суперпозиції відображень $\mathbb{C} \xrightarrow{\sigma} \mathbb{C} \xrightarrow{\sigma} \mathbb{C}$ кожен елемент переходить в себе, тобто $x \mapsto \sigma(x) \mapsto \sigma(\sigma(x)) = x$. Отже, $\sigma \cdot \sigma = 1_{\mathbb{C}}$ або $\sigma^2 = \mathbf{1}_{\mathbb{C}}$. Іноді можна зустріти запис: $\bar{\bar{x}} = x$.

5. Розглянемо поле $\mathbb{R}$ як підполе в $\mathbb{C}$ відносно введеного раніше відображення "вкладення":
$$s : \begin{cases} \mathbb{R} & \rightarrow & \mathbb{C} \\ a & \mapsto & a + 0 \cdot i \qquad \forall a \in \mathbb{R}. \end{cases}$$

   При цьому операції зберігаються.

   Комплексні числа вигляду $a + 0 \cdot i$ називають чисто дійсними, а вигляду $0 + bi$ - чисто уявними. Отже, $s(\mathbb{R})$ - це множина чисто дійсних комплексних чисел, а $i \cdot s(\mathbb{R})$ - множина чисто уявних комплексних чисел.

   Лемма 2.1. *Припустимо, $x \in \mathbb{C}$. Тоді*

   - $\sigma(x) = x \quad \equiv \ x$ *- чисто дійсне;*
   - $\sigma(x) = -x \ \equiv \ x$ *- чисто уявне.*

   Твердження очевидне.

6. *Модуль комплексного числа.* Для будь-якого комплексного числа $z = a + bi$ величина $z \cdot \bar{z} = (a+bi)(a-bi) = a^2 + b^2$ є дійсним числом. Число $|z| = \sqrt{a^2+b^2}$ називається *модулем* z. Формула
$$z \cdot \bar{z} = |z|^2$$
   виражає зв'язок між поняттями спряження комплексного числа та його модулем.

7. *Ділення комплексних чисел* $x = a + bi$ та $y = c + di$ здійснюється за правилом:

$$\frac{x}{y} = \frac{x \cdot \bar{y}}{|y|^2} = \frac{(a+bi)(c-di)}{(c+di)(c-di)} = \frac{(ac+bd)}{c^2+d^2} + \frac{(bc+ad)}{c^2+d^2}i.$$

Зауважимо, що при цьому в знаменнику утворюється дійсне число $|y| = c^2 + d^2$.

Зокрема, для оберненого числа маємо

$$\frac{1}{x} = \frac{\bar{x}}{|x|^2} = \frac{a}{a^2+b^2} - \frac{b}{a^2+b^2}i.$$

Лемма 2.2. *Припустимо, що* $f(x) = a_0x^n + a_1x^{n-1} + \ldots + a_n \in \mathbb{R}[x]$ *- поліном з дійсними коефіцієнтами, тобто такий, що* $a_i \in \mathbb{R}$, $i = 0, 1, \ldots, n$. *Якщо* $z \in \mathbb{C}$ *є коренем* $f(x)$, *тобто* $f(z) = a_0z^n + a_1z^{n-1} + \ldots + a_n = 0$ *в полі* $\mathbb{C}$, *то* $f(\bar{z}) = 0$, *тобто* $\bar{z}$ *теж є коренем рівняння* $f(x) = 0$.

Доведення. Запишемо поліном у вигляді: $f(x) = \sum\limits_{k=0}^{n} a_k x^{n-k}$. Тоді $f(\bar{z}) = \sum\limits_{k=0}^{n} a_k \bar{z}^{n-k} = \sum\limits_{k=0}^{n} a_k \overline{(z^{n-k})} \overset{a_k \in \mathbb{R}}{=} \sum\limits_{k=0}^{n} \overline{(a_k)}\ \overline{(z^{n-k})} = \sum\limits_{k=0}^{n} \overline{(a_k z^{n-k})} = \overline{(\sum\limits_{k=0}^{n} a_k z^{n-k})} = \overline{f(z)} = \bar{0} = 0$. (Ми скористалися тим, що $a_k = \bar{a_k}$, оскільки $a_k \in \mathbb{R}$, а також лінійністю функції спряження відносно множення). □

Задача 2.1. *Чому це доведення не годиться для довільних коефіцієнтів?*

2.7 Тригонометрична форма комплексного числа

Множина комплексних чисел $\mathbb{C} = \{x + yi \,|\, x, y \in \mathbb{R}\}$ - має доступну для сприйняття геометричну форму. Пари дійсних чисел (x, y) утворюють декартову площину $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R} = \{(x, y) \,|\, x, y \in \mathbb{R}\}$ (нагадаємо, що *декартовою площиною* називають площину разом з фіксованою системою координат).

Означення 2.5. Геометричним зображенням *комплексного числа* $z = x + yi \in \mathbb{C}$ *називається точка декартової площини, яку в такому разі звичайно звуть* комплексною площиною, $(x, y) \in \mathbb{R}^2$ *і відповідний цій точці радіус-вектор. При цьому дійсним числам відповідають точки осі абсцис, яку звуть також* дійсною віссю. *Точкам осі ординат відповідають комплексні числа вигляду* bi, *де* $b \in \mathbb{R}$. *Їх звуть* чисто уявними *і, відповідно, вісь ординат звуть також* уявною віссю.

Зауваження 2.4. *Додавання комплексних чисел в геометричній формі збігається з додаванням відповідних радіусів-векторів. Далі ми покажемо, що множення допускає таку само просту геометричну інтерпретацію.*

2.8 Полярні координати

В стандартній декартовій системі координат, яка задається двома взаємно перпендикулярними прямими - координатними осями, кожна точка визначається двома координатами.

Фіксуємо на площині точку O - полюс і вісь l з початком у цьому полюсі. Тоді кожній точці A, яка не дорівнює полюсу, ставиться у відповідність пара (r_A, φ_A), де $r_A = (OA)$ - довжина відрізка, а φ_A - це кут від осі до радіус-вектора OA, взятий проти годинникової стрілки і виміряний в радіанах з точністю до 2π.

Пояснення. Говорячи про кут, мають на увазі клас еквівалентності $\Phi_A = \{\varphi + 2k\pi \,|\, k \in \mathbb{Z}\}$, але для уникнення складнощів запису пишуть полярні координати у вигляді пари чисел (r, φ), де $r > 0$, $\varphi \in \mathbb{R}$. Такий запис означає, що ми маємо клас еквівалентності $\{(r, \varphi + 2k\pi) \,|\, k \in \mathbb{Z}\}$.

Можна розглядати полярні координати полюса, тобто коли $r = 0$, але в цьому випадку кут φ невизначений.

Нехай $z = a + bi$ - комплексне число, де x, y - декартові координати. Полярні координати знаходяться з такої системи рівнянь:

$$\left\{ \begin{array}{lcl} r^2 & = & a^2 + b^2; \\ \cos(\varphi) & = & \frac{a}{\sqrt{a^2+b^2}}; \\ \sin(\varphi) & = & \frac{b}{\sqrt{a^2+b^2}}. \end{array} \right. \tag{3}$$

З наведених умов r знаходиться однозначно, а кут φ з точністю до 2π. При цьому $r = |z| = \sqrt{a^2 + b^2}$ називається *модулем*, а множина $\{\varphi + 2k\pi \,|\, k \in \mathbb{Z}\}$ - *аргументом* комплексного числа z.

Навпаки, зворотний перехід, коли задані $r > 0$ і $\{\varphi + 2k\pi \,|\, k \in \mathbb{Z}\}$ (або, кажуть, просто φ), тоді координати x, y визначаються з тих же рівнянь. Маємо

$$\left\{ \begin{array}{lcl} x & = & r\cos(\varphi), \\ y & = & r\sin(\varphi). \end{array} \right. \tag{4}$$

Тригонометрична форма комплексного числа $z = x + yi$, $z \neq 0$, - це його запис у вигляді $z = r(\cos\varphi + i\sin\varphi)$, де r і φ знаходяться з рівнянь (3) або (4).

Зауваження 2.5. *1. В тригонометричній формі під знаками* $\sin$ *та* $\cos$ *записується один і той же кут і не пишеться $2k\pi$, $k \in \mathbb{Z}$. Тому тригонометричні форми $r(\cos\varphi + i\sin\varphi)$ і $r_1(\cos(\varphi_1) + i\sin(\varphi_1))$ визначають одне і те саме комплексне число, якщо $r = r_1$ і $\varphi = \varphi_1 + 2k\pi$ для деякого $k \in \mathbb{Z}$.*
2. Знання значення однієї *тригонометричної функції (косинуса чи синуса) недостатньо для визначення кута.*

2.9 Операції над комплексними числами, заданими у тригонометричній формі

1. *Множення комплексних чисел.* Тригонометрична форма комплексного числа має ту властивість, що вона спрощує множення. А саме, перемножимо два комплексних числа $z_1 = r_1(\cos\varphi_1 + i\sin\varphi_1)$, $z_2 = r_2(\cos\varphi_2 + i\sin\varphi_2)$, записані у тригонометричній формі. Тоді $z_1 z_2 = r_1 r_2((\cos\varphi_1\cos\varphi_2 - \sin\varphi_1\sin\varphi_2) + i(\cos\varphi_1\sin\varphi_2 + \sin\varphi_1\cos\varphi_2))$. Звідси, скориставшись відомими тригонометричними формулами додавання, одержимо:

$$z_1 z_1 = r_1 r_2(\cos(\varphi_1 + \varphi_2) + i\sin(\varphi_1 + \varphi_2)).$$

Одержаний добуток знову має тригонометричну форму. Маємо такі рівності:

$$\begin{aligned} |z_1 \cdot z_2| &= |z_1| \cdot |z_2|; \\ arg(z_1 \cdot z_2) &= arg(z_1) + arg(z_2), \end{aligned}$$

з яких випливає *правило множення* чисел в тригонометричній формі:

щоб перемножити два числа, кожне з яких записане в тригонометричній формі, треба їх модулі перемножити, а аргументи додати.

Зауваження 2.6. *В другій рівності додаються усі можливі значення аргументу.*

Множення на комплексне число $z = r(\cos\varphi + i\sin\varphi) \in \mathbb{C}$ є відображенням

$$m_z : \begin{cases} \mathbb{C} \to \mathbb{C} \\ z' \mapsto z \cdot z' \quad \forall z' \in \mathbb{C}. \end{cases}$$

Геометрично m_z є суперпозицією гомотетії із коефіцієнтом r (тобто розтягування в r разів), та повороту на кут φ (обидва перетворення відносно початку координат).

2. *Обернене комплексне число.* Якщо $z = r(\cos\varphi + i\sin\varphi)$, то

$$z^{-1} = r^{-1}(\cos(-\varphi) + i\sin(-\varphi)).$$

Число z^{-1} визначено однозначно. Перевіримо, що для наведеного виразу виконуються рівності: $z \cdot z^{-1} = z^{-1} \cdot z = 1$. Дійсно, $z \cdot z^{-1} = r(\cos\varphi + i\sin\varphi) \cdot r^{-1}(\cos(-\varphi) + i\sin(-\varphi)) = 1(\cos 0 + i\sin 0) = 1$.

3. *Ділення* чисел z_1 та $z_2 \neq 0$ у тригонометричній формі здійснюється за формулою:

$$\frac{z_1}{z_2} = \frac{r_1(\cos\varphi_1 + i\sin\varphi_1)}{r_2(\cos\varphi_2 + i\sin\varphi_2)} = \frac{r_1}{r_2}(\cos(\varphi_1 - \varphi_2) + i\sin(\varphi_1 - \varphi_2)).$$

4. Як наслідок з формули множення комплексних чисел випливає така формула *піднесення до степеня* у тригонометричній формі:

$$(r(\cos\varphi + i\sin\varphi))^n \;=\; r^n(\cos(n\varphi) + i\sin(n\varphi)).$$

Коли $r = 1$, ця формула називається *формулою Муавра*:

$$(\cos\varphi + i\sin\varphi)^n \;=\; \cos n\varphi + i\sin n\varphi.$$

З її допомогою можна знайти формули для $\cos n\varphi$, $\sin n\varphi$ через $\cos\varphi$, $\sin\varphi$.

Корені з комплексного числа

Нехай $z = r(\cos\varphi + i\sin\varphi)$. Розглянемо рівняння $t^n = z$, де t - невідоме. Така задача називається *задачею добування кореня $n-$того степеня.* Запишемо t у вигляді $\; = \rho(\cos\psi + i\sin\psi)$ з невідомими ρ, ψ. Тоді рівняння перепишеться у вигляді: $\rho^n(\cos n\psi + i\sin n\psi) = r(\cos\varphi + i\sin\varphi)$. Звідси одержимо, що $\rho^n = r$ і $n\psi - \varphi = 2\pi k$, $k \in \mathbb{Z}$. Отже маємо, що $\rho = \sqrt[n]{r}$ - в дійсних додатних числах, а $\psi = \dfrac{\varphi}{n} + \dfrac{2\pi k}{n}$, $k \in \mathbb{Z}$.

Ми одержали для ψ нескінченно багато значень. Але це не означає, що існує нескінченно багато кутів. Знайдемо, скільки є з наведених кутів різних з точністю до 2π. Для цього розглянемо кути $\psi_k = \frac{\varphi}{n} + \frac{2\pi k}{n}$, $k \in \mathbb{Z}$. Маємо: $\psi_k - \psi_{k+n} = (\frac{\varphi}{n} + 2\pi\frac{k}{n}) - (\frac{\varphi}{n} + 2\pi\frac{k+n}{n}) = -\frac{2\pi n}{n} = -2\pi$. Отже, коли k збільшується на n, то ψ_k змінюється на -2π. Таким чином, достатньо розглянути лише такі значення:

$$\psi_k \;=\; \frac{\varphi}{n} + \frac{2\pi k}{n}, \quad k = 0, 1, \ldots, n-1.$$

Всі інші значення k будуть приводити до кутів, які відрізняються від записаних на ціле кратне 2π.

Ми показали, що рівняння $t^n = z = r(\cos\varphi + i\sin\varphi)$ має n розв'язків:

$$t_k \;=\; \sqrt[n]{r}\left(\cos\left(\frac{\varphi}{n} + \frac{2\pi k}{n}\right) + i\sin\left(\frac{\varphi}{n} + \frac{2\pi k}{n}\right)\right), \quad k = 0, 1, \ldots, n-1.$$

2.10 Корені з одиниці

Розглянемо рівняння $z^n = 1$. Позначимо через $\mathbf{C}_n = \{z \in \mathbb{C} \mid z^n = 1\}$ множину розв'язків цього рівняння. Маємо n різних розв'язків:

$$\xi_k \;=\; \cos\left(\frac{2\pi k}{n}\right) + i\sin\left(\frac{2\pi k}{n}\right), \quad k = 0, 1, \ldots, n-1.$$

Отже $|\mathbf{C}_n| = n$ і $\mathbf{C}_n = \{\xi_0, \ldots, \xi_{n-1}\}$. Числа ξ_k називаються *коренями з* 1 степеня n.

Властивості коренів з 1.

Маємо $\mathbb{C}_n = \{\xi_0, \ldots, \xi_{n-1}\}$.

1. Добуток двох коренів з 1 степеня n є корінь з 1 степеня n, тобто виконується така умова:
$$\xi, \xi' \in \mathbb{C}_n \implies \xi \cdot \xi' \in \mathbb{C}_n.$$
Дійсно, $(\xi \cdot \xi')^n = \xi^n \cdot \xi'^n$.

2. Обернений до кореня з 1 степеня n також є коренем з 1 степеня n, або
$$\xi \in \mathbb{C}_n \implies \xi^{-1} \in \mathbb{C}_n.$$
Доведення випливає з такої рівності: $(\frac{1}{\xi})^n = \frac{1}{\xi^n} = 1$.

3. Одиниця є коренем з 1 степеня n, тобто $1 \in \mathbb{C}_n$ для довільного n. [3]

Зауваження 2.7. *Оскільки геометрично m_z є суперпозиція гомотетії із коефіцієнтом r (тобто розтягування в r разів), та повороту на кут φ (обидва перетворення відносно початку координат), то ξ_k утворюють вершини правильного $n-$кутника з центром в початку координат, одна з вершин якого є 1.*

Первісні корені з одиниці

Корінь з одиниці $\xi \in \mathbb{C}_n$ називається *первісним*, якщо $\xi^n = 1$, але $\xi^k \neq 1$ для довільного $1 \leqslant k < n$.

Властивості.

1. Первісний корінь існує, а саме $\xi_1 = \cos\left(\frac{2\pi}{n}\right) + i\sin\left(\frac{2\pi}{n}\right)$ - первісний. Очевидно, $\xi_1^k = \xi_k$ для будь-якого $k = 0, 1, \ldots, n-1$ і якщо $\xi_1^k = 1$, то $k = n$.

2. Якщо ξ - первісний корінь, то $\mathbb{C}_n = \{1, \xi, \xi^2, \ldots, \xi^{n-1}\}$.

   Доведення. З одного боку, $\{1, \xi, \xi^2, \ldots, \xi^{n-1}\} \subset \mathbb{C}_n$ тому, що добуток коренів з 1 є коренем.

   З іншого боку, всі елементи множини $\{1 = \xi^0, \xi, \xi^2, \ldots, \xi^{n-1}\}$ різні, тобто кількість елементів в цій множині дорівнює n. Дійсно, припустимо, $\xi^i = \xi^j$ для деяких $0 \leqslant i \neq j \leqslant n-1$. Звідси випливає, що $\xi^{j-i} = 1$. Оскільки при цьому $j - i < n$, то це є протиріччя з тим, що корінь ξ - первісний. Отже, $n-$елементна множина $\{1 = \xi^0, \xi, \xi^2, \ldots, \xi^{n-1}\}$ є підмножиною $n-$елементної множини $\mathbb{C}_n$, звідки випливає рівність. $\square$

Означення 2.6. *Через $\mathcal{P}_n \subset \mathbf{C}_n$ позначимо множину всіх первісних коренів з 1 степеня n. Функція Ейлера $\varphi(n) = |\mathcal{P}_n|$ визначається як кількість первісних коренів з 1 степеня n.*

[3] **Властивості 1, 2, 3 означають, що $\mathbb{C}_n$ є групою.**

Зауваження 2.8. $1 \leqslant \varphi(n) < n$ *для* $n > 1$ *(тому, що* ξ_1 *- завжди первісний).*

Визначимо деякі значення φ*:* $\varphi(1) = 1$*,* $\varphi(2) = 1$*,* $\varphi(3) = 2$*,* $\varphi(4) = 2$*,* $\varphi(5) = 4$*.*

Позначення. $d \,|\, n \iff d$ є дільником n або ж $\exists q \in \mathbb{Z} \;\; n = dq$. Якщо $m, n \in \mathbb{Z}$, то через $d = (m, n)$ позначається їх найбільший цілий додатній дільник.

Лемма 2.3. *Припустимо,* $\xi \in \mathbf{C}_n$ *і* $d > 0$ *- мінімальне ціле, таке що* $\xi^d = 1$*. Тоді* d *є дільником* n*.*

Доведення. Ясно, що $d \leqslant n$ ($\xi^n = 1 \implies n \geqslant d$). Ми можемо завжди розділити n на d із залишком, тобто представити $n = dq + r$, де $0 \leqslant r < d$. Тоді $1 = \xi^n = \xi^{dq+r} = \xi^{dq} \cdot \xi^r = (\xi^d)^q \cdot \xi^r = 1^q \cdot \xi^r = \xi^r \implies \xi^r = 1$, але $r < d$. Це суперечить тому, що d - мінімальне ціле > 0 таке, що $\xi^d = 1$. Отже, $r = 0$, тоді $n = dq$, звідки випливає, що $d \,|\, n$. $\square$

Доведену лему можна переформулювати таким чином:

$$\xi \in \mathbf{C}_n, \; \xi \in \mathcal{P}_d \implies d \,|\, n.$$

Лемма 2.4. *Якщо* $\xi \in \mathcal{P}_n$ *і* $\xi^N = 1$*, то* $n|N$*.*

Доведення леми аналогічне доведенню попередньої леми.

Наслідок 2.1. *Припустимо, що* m *і* n *взаємно прості. Тоді* $\mathbf{C}_m \cap \mathbf{C}_n = \{1\}$*.*

Доведення. Розглянемо $\xi \in \mathbf{C}_m \cap \mathbf{C}_n$. Нехай ξ є первісним степеня d. Звідси випливає, що $d|m$ і $d|n$. Отже d - спільний дільник m і $n \implies d = 1$ і $\xi = 1$. $\square$

Лемма 2.5. *Корінь* $\xi_k = \cos\left(\dfrac{2\pi k}{n}\right) + i \sin\left(\dfrac{2\pi k}{n}\right) \in \mathbf{C}_n$ *буде первісним (тобто* $\xi \in \mathcal{P}_n$*) тоді і тільки тоді, коли цілі* k *і* n *взаємно прості, тобто* $(k, n) = 1$*.*

Доведення. Очевидно, ξ_k є первісним деякого степеня d, при цьому, як було доведено, $d|n$. Доведемо, що $d = n \equiv (k, n) = 1$.

$\implies$ Нагадаємо, що $\xi_k = \xi_1^k$, де $\xi_1 = \cos(\frac{2\pi}{n}) + i\sin(\frac{2\pi}{n})$ - первісний. Припустимо, що k, n мають спільний дільник t, тобто $k = tk_1$, $n = tn_1$ і $t > 1$. Тоді $n_1 < n$, отже $(\xi_k)^{n_1} = (\xi_1^k)^{n_1} = \xi_1^{kn_1} = \xi_1^{k_1 t n_1} = \xi_1^{k_1 n} = (\xi_1^n)^{k_1} = 1$. Оскільки $n_1 < n$, то звідси випливає, що ξ_k не є первісним степеня n.

$\impliedby$ Навпаки, припустимо, що $(k, n) = 1$ і що $(\xi_k)^t = 1$. Тоді $\xi_1^{kt} = ((\xi_1)^k)^t = (\xi_k)^t = 1$. Але $\xi_1 \in \mathcal{P}_n$. Застосуємо попередню лему. $\xi_1^{kt} = 1 \implies n|kt \overset{(n,k)=1}{\implies} n|t \implies t \geqslant n$. Отже, якщо t - мінімальне, при якому $\xi_k^t = 1$, то $t = n$. $\square$

За доведеною лемою, $|\mathcal{P}_n| = \varphi(n)$, де φ - функція Ейлера. Отже,

$$\varphi(n) = \{k \in \mathbb{Z} \,|\, 0 \leqslant k < n, (k, n) = 1\}.$$

Приклад 2.5. $\mathcal{P}_{10} = \{\xi_1, \xi_3, \xi_7, \xi_9\}$.

Доведемо деякі прості властивості функції Ейлера.

Лемма 2.6. *Якщо p - просте ціле, то* $\varphi(p^n) = p^n\left(1 - \dfrac{1}{p}\right)$.

Доведення. Дійсно, $\varphi(p^n) = p^n - \{k \in \mathbb{Z} \,|\, 0 \leqslant k < p^n, p|k\} = p^n - \frac{p^n}{p} = p^n(1 - \frac{1}{p})$. $\square$

Лемма 2.7. *Якщо $(m,n) = 1$, то $\varphi(mn) = \varphi(m)\cdot\varphi(n)$ (мультиплікативність функції φ).*

Доведення. Розглянемо відображення

$$\begin{cases} p: & \mathbf{C}_m \times \mathbf{C}_n & \longrightarrow & \mathbf{C}_{mn} \\ & (x,y) & \mapsto & x \cdot y, \quad x \in \mathbf{C}_m,\ y \in \mathbf{C}_n. \end{cases}$$

Чому $x, y \in \mathbf{C}_{mn}$? $(xy)^{mn} = x^{mn} \cdot y^{mn} = (x^m)^n (y^n)^m = 1^n 1^m = 1$. Доведемо, що відображення p ін'єктивне. Припустимо, що $p(x_1, y_1) = p(x_2, y_2)$. Це еквівалентно $x_2^{-1}\cdot x_1 \cdot y_1 = y_2 \equiv x_2^{-1}\cdot x_1 = y_2 \cdot y_1^{-1}$. Тоді $x_2^{-1}\cdot x_1 \in \mathbf{C}_m$, $y_2 \cdot y_1^{-1} \in \mathbf{C}_n$ (ми використовували властивості коренів).

Позначимо $x = x_2^{-1} \cdot x_1 \in \mathbf{C}_m$, $y = y_2 \cdot y_1^{-1} \in \mathbf{C}_n$. Як показано вище, $x = y$, але ми довели, що коли $(m,n) = 1$, то звідси випливає, що $\mathbf{C}_m \cap \mathbf{C}_n = \{1\}$, а отже, оскільки $x = y \in \mathbf{C}_m \cap \mathbf{C}_n$, то $x = y = 1$, або ж $x_2^{-1}x_1 = 1$, $y_2 y_1^{-1} = 1$, звідки одержуємо $x_1 = x_2$, $y_2 = y_1$. Ми довели, що $(x_1, y_1) = (x_2, y_2)$, а це означає, що p ін'єктивне.

Щоб довести, що p також і сюр'єктивне, достатньо підрахувати кількість елементів в множині $\mathbf{C}_m \times \mathbf{C}_n$. Маємо $|\mathbf{C}_m \times \mathbf{C}_n| = |\mathbf{C}_m| \cdot |\mathbf{C}_n| = mn = \mathbf{C}_{mn}$. Отже, відображення p є бієкція.

Це твердження можна сформулювати так:

Кожен корінь степеня mn однозначно розкладається в добуток кореня степеня m і кореня степеня n (при умові, що $(m,n) = 1$).

Стверджуємо, що $p(\mathcal{P}_m \times \mathcal{P}_n) = \mathcal{P}_{mn}$.

Доведення. Припустимо, що в парі $(x,y) \in \mathbf{C}_m \times \mathbf{C}_n$ корінь x не первісний, тобто $x \notin \mathcal{P}_m$. Це означає, що існує таке $d < m$, що $x^d = 1$. Розглянемо число $dn < mn$. Тоді $(xy)^{dn} = (x^d)^n \cdot (y^n)^d = 1 \cdot 1 = 1$. Звідси випливає, що $p(x,y) = xy \notin \mathcal{P}_{mn}$, отже добуток xy також не первісний.

Припустимо тепер, що обидва корені x та y первісні, тобто $(x,y) \in \mathcal{P}_m \times \mathcal{P}_n$. Виберемо мінімальне додатне $N \in \mathbb{Z}$ таке, що $(xy)^N = 1$. Тоді $x^N y^N = 1 \implies x^N = y^{-N}$. Але $x^N \in \mathbf{C}_m$, $y^{-N} \in \mathbf{C}_n$. Звідси $x^N = y^{-N} \in \mathbf{C}_m \cap \mathbf{C}_n = \{1\} \implies x^N = 1 \wedge y^N = 1 \implies x \in \mathcal{P}_m \wedge y \in \mathcal{P}_n \implies m|N \wedge n|N$. А оскільки $(m,n) = 1$, то одержимо $mn|N$ і отже $N = mn$. Таким чином, $xy \in \mathcal{P}_{mn}$, отже добуток первісних також є первісним.

Отже, $p : \mathcal{P}_m \times \mathcal{P}_n \longrightarrow \mathcal{P}_{mn}$ - бієкція. Наслідок: $|\mathcal{P}_m \times \mathcal{P}_n| = \mathcal{P}_{mn}$. Але $|\mathcal{P}_m \times \mathcal{P}_n| = |\mathcal{P}_m| \cdot |\mathcal{P}_n| = |\mathcal{P}_{mn}|$. За означенням, це еквівалентно формулі добутку: $\varphi(m) \cdot \varphi(n) = \varphi(mn)$. Лему доведено. $\square$

Наслідок 2.2. *Нехай* $n = p_1^{k_1} \cdot p_2^{k_2} \cdots p_k^{k_t}$ *є розклад* $n > 0$ *в добуток простих. Тоді*

$$\varphi(n) = p_1^{k_1} \cdot p_2^{k_2} \cdot \cdot p_k^{k_t} \left(1 - \frac{1}{p_1}\right)\left(1 - \frac{1}{p_2}\right) \cdots \left(1 - \frac{1}{p_t}\right),$$

або перепишемо цю формулу еквівалентним способом:
$\varphi(n) = n(1 - \frac{1}{p_1})(1 - \frac{1}{p_2}) \cdots (1 - \frac{1}{p_t}) = p_1^{k_1 - 1} \cdot p_2^{k_2 - 1} \cdots p_k^{k_t - 1}\ (p_1 - 1) \cdot (p_2 - 1) \cdots (p_t - 1).$

Доведення. $\varphi(n) = \varphi(p_1^{k_1} \cdot p_2^{k_2} \cdots p_k^{k_t}) \overset{\text{лема 3}}{=} \varphi(p_1^{k_1})\varphi(p_2^{k_2}) \cdots \varphi(p_k^{k_t}) = p_1^{k_1}(1 - \frac{1}{p_1}) \cdot p_2^{k_2}(1 - \frac{1}{p_2}) \cdots p_k^{k_t}(1 - \frac{1}{p_k})$. □

Лемма 2.8 (Формула суми)**.** *Нехай* $n > 0$ *- ціле. Тоді*

$$n \ = \sum_{d>0,\, d\,|n} \varphi(d).$$

Доведення. Розглянемо $\mathbf{C}_n$. Кожен корінь $\xi \in \mathbf{C}_n$ є первісним, тобто $\xi \in \mathcal{P}_d$ і $d|n$.

Навпаки, якщо $\xi \in \mathcal{P}_d$, $d\,|n$, то $\xi \in \mathbf{C}_n$. Звідси $\mathbf{C}_n = \coprod_{d>0,\, d\,|n} \mathcal{P}_d$ (такий запис означає об'єднання множин, які не перетинаються). Тому $|\mathbf{C}_n| =$ або $n = \sum_{d>0,\, d\,|n} \varphi(d)$. Отже, ми довели формулу суми. □

3 Системи лінійних рівнянь

Інформація про розв'язки лінійних рівнянь містить в собі основи всієї подальшої математики, носить генетичний характер, як арифметика містить основи банківської справи.

3.1 Лінійні рівняння

Фіксуємо деяке поле $\mathbb{K}$.

Означення 3.1. Лінійним рівнянням *відносно невідомих* $x_1, x_2, \ldots,\ x_n$ *над полем* $\mathbb{K}$ *називаєтьcя рівняння вигляду*

$$a_1 x_1 + a_2 x_2 + \ldots + a_n x_n = b,$$

де $a_1, a_2, \ldots, a_n, b \in \mathbb{K}$ *- коефіцієнти, а* $x_1, x_2, \ldots, x_n$ *- деякі символи, які умовно називаються* невідомими. *Це рівняння від* n *невідомих.*

Домовленість 1. *При запису рівнянь ми дотримуємося стандартної угоди, у відповідності з якою*
1) якщо коефіцієнт при невідомому дорівнює одиниці, то він, як правило, не пишеться: замість $1x_i$ ми пишемо просто x_i;
2) якщо коефіцієнт при невідомому від'ємний, тобто $a_i = a$, то відповідний член записується так: $+(-a) \cdot x_i = -ax_i$;
3) якщо коефіцієнт при невідомому дорівнює нулю, то інколи відповідний доданок $0 \cdot x_i$ у запису пропускається.

3.2 Системи лінійних рівнянь

Система m лінійних рівнянь від n невідомих над полем $\mathbb{K}$ - це сукупність m рівнянь від n невідомих, яка записується у вигляді:

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

або, коротко, у вигляді

$$(**) \quad \left\{ \sum_{j=1}^{n} a_{ij}x_j = b_i, \qquad i = 1, \dots, m. \right.$$

Означення 3.2. *Розв'язком системи лінійних рівнянь $(*)$ (або $(**)$) називається вектор (тобто елемент декартового добутку $\mathbb{K}^n = \mathbb{K} \times \ldots \times \mathbb{K}$ (n разів), записаний в стовпчик*

$$\begin{pmatrix} l_1 \\ l_2 \\ \vdots \\ l_n \end{pmatrix}$$

де $l_1, l_2, \dots, l_n \in \mathbb{K}$, такий, що після підстановки в систему замість всіх x_i відповідних числових значень l_i, $i = 1, \dots, n$ всі рівняння системи $(*)$ перетворюються на правильні числові рівності в полі $\mathbb{K}$, тобто

$$\begin{cases} a_{11}l_1 + & a_{12}\,l_2 + & \dots & +a_{1n}l_n & = b_1 \\ a_{21}l_1 + & a_{22}\,l_2 + & \dots & +a_{2n}l_n & = b_2 \\ & \dots & \dots & \dots & \dots \\ a_{m1}l_1 + & a_{m2}\,l_2 + & \dots & +a_{mn}l_n & = b_m \end{cases} \tag{5}$$

Домовленість 2. *В чому різниця між записами $(*)$ та (5)? Запис $(*)$ - формальний, він ні до чого нас не зобов'язує, питання, чи вірно $(*)$, чи ні, не виникає. А запис (5) означає операції в полі $\mathbb{K}$.*

Означення 3.3. *Дві системи, які залежать від однакової кількості невідомих, називаються рівносильними, якщо кожен розв'язок першої системи є розв'язком другої системи та навпаки.*

Означення 3.4. *Позначимо через* $\mathfrak{L}_S \subset \mathbb{K}$ *множину всіх розв'язків системи* $(*)$. *Якщо* $\mathfrak{L}_S \neq \varnothing$, *то система* $(*)$ *називається* сумісною. *Якщо* $\mathfrak{L}_S = \varnothing$, *то система* $(*)$ - несумісна. *Коли* $|\mathfrak{L}_S| = 1$, *система* $(*)$ *називається* визначеною.

Приклад 3.1. *Дослідимо ці поняття у випадку* $n = 1$, $m = 1$. *Маємо одне рівняння з одним невідомим.*

$$\left\{ \begin{array}{rcl} a_{11}x_1 & = & b_1 \end{array} \right. , \qquad a_{11}, b_1 \in \mathbb{K}.$$

1. *Припустимо, що* $a_{11} \neq 0$. *Домножимо тоді рівняння на* a_{11}^{-1}, *одержимо, що* $x_1 = a_{11}^{-1}b_1$. *Це, очевидно, єдиний розв'язок системи. В цьому випадку* $|\mathfrak{L}_S| = 1$, *система визначена.*

2. $a_{11} = 0$. *Маємо два підвипадки:*

   (a) $b_1 \neq 0$, *тоді* $0 \cdot x_1 = b_1 \neq 0$, $\mathfrak{L}_S = \varnothing$, *система* $(*)$ *несумісна.*

   (b) $b_1 = 0$. *В цьому випадку довільне значення* x_1 *буде розв'язком,* $\mathfrak{L}_S = \mathbb{K}^1 = \mathbb{K}$, *система* $(*)$ *сумісна і має нескінченно багато розв'язків.*

3.3 Координатний числовий векторний простір

Корисно буде дослідити множину $\mathbb{K}^n$.

Означення 3.5. *Нехай* $\mathbb{K}$ *- деяке поле. Множина* $\mathbb{K}^n$ *разом з операціями:*

1) додавання $(+)$, *а саме* $\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} + \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} = \begin{pmatrix} a_1 + b_1 \\ a_2 + b_2 \\ \vdots \\ a_n + b_n \end{pmatrix}$

та 2) множення на елементи поля $\mathbb{K}$ *(їх ще називають скалярами) вигляду*

$$\lambda \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} \lambda a_1 \\ \lambda a_2 \\ \vdots \\ \lambda a_n \end{pmatrix}$$

називається координатним (числовим) векторним простором *над полем* $\mathbb{K}$. *Число* n *називається розмірністю простору* $\mathbb{K}^n$. *Елемент векторного простору* $a = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}$ *називають вектором (-стовпчиком) над полем* $\mathbb{K}$. *Елемент* a_i *вектора* a *називають*

i-тою координатою *вектора a. При цьому кажуть, що операції виконуються покоординатно.*

Інколи елементи з $\mathbb{K}^n$ записують в рядок $(a_1, a_2, \ldots, a_n) \in \mathbb{K}^n$. В цьому випадку говорять про вектори-рядки над полем $\mathbb{K}$.

Приклад 3.2. *Якщо $\mathbb{K} = \mathbb{R}$ - поле дійсних чисел, то над ним розглядаються в шкільному курсі математики числові векторні простори:*

- $\mathbb{R}^1 = \mathbb{R}$ - *пряма;*
- $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$ - *площина;*
- $\mathbb{R}^3 = \mathbb{R} \times \mathbb{R} \times \mathbb{R}$ - *тривимірний простір.*

Домовленість 3. *З системою $(*)$ пов'язують стовпчик невідомих $x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{K}^n$, стовпчик правих частин $b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix} \in \mathbb{K}^m$ та матрицю системи.*

Означення 3.6. *Нехай $m, n \geqslant 1$ - пара цілих чисел.* Матрицею *A розміру $(m \times n)$ (або $(m \times n)$-матрицею) над полем $\mathbb{K}$ називається набір mn елементів поля $\mathbb{K}$, записаний у вигляді прямокутної таблиці з m рядками та n стовпчиками:*

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix}, \quad \text{де } a_{ij} \in \mathbb{K},\ 1 \leqslant i \leqslant m,\ 1 \leqslant j \leqslant n. \tag{6}$$

Числа a_{ij} називаються елементами матриці.

У випадку $m = n$ матриця називається *квадратною.*

Вказану вище матрицю X інколи записують у вигляді

$$A = (a_{ij}), \quad 1 \leqslant i \leqslant m, 1 \leqslant j \leqslant n \quad \text{або} \quad A = (a_{ij})_{\left\{ \substack{1 \leqslant i \leqslant m, \\ 1 \leqslant j \leqslant n} \right\}}$$

Цей запис цілком еквівалентний попередньому.

Домовленість 4. *У запису елемента a_{ij} матриці A перший індекс i вказує номер рядка, а другий індекс j вказує номер стовпчика, в якому знаходиться цей елемент. У цьому випадку про a_{ij} говорять, що*

- він є $ij-$ тим елементом матриці A;
- він знаходиться на $ij-$ тому місці;
- він знаходиться на місці з номером ij;
- він є елемент з індексом ij.

Матрицю A можна розглядати, як складену з n векторів-стовпчиків висоти m:

$$A = \left(\mathbf{a_1} \,|\, \mathbf{a_2} \,|\, \ldots \,|\, \mathbf{a_n} \right)$$

де $\mathbf{a_j} = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix}$ - $j-$тий стовпчик матриці A, $j = 1, \ldots, n$. Вертикальні риски у вказаному запису матриці A відділяють стовпчики один від одного.

Аналогічно, матрицю $A = (a_{ij}), 1 \leqslant i \leqslant m, 1 \leqslant j \leqslant n$, можна розглядати, як складену з m векторів-рядків довжини n: $A = \begin{pmatrix} \underline{\mathbf{a^1}} \\ \underline{\mathbf{a^2}} \\ \vdots \\ \overline{\mathbf{a^m}} \end{pmatrix}$, де $\mathbf{a^i} = (a_{i1}, a_{i2}, \ldots, a_{in})$, $i = 1, 2, \ldots, m$. Горизонтальні риски у вказаному запису матриці A відділяють рядки один від одного.

З системою лінійних рівнянь пов'язують дві матриці: *матрицю A системи* $(*)$ порядку $(m \times n)$ вигляду (6) та *розширену матрицю системи* розміру $(m \times (n+1))$:

$$\bar{A} = \left(\begin{array}{cccc|c} a_{11} & a_{12} & \ldots & a_{1n} & b_1 \\ a_{21} & a_{22} & \ldots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} & b_m \end{array} \right). \tag{7}$$

Розширена матриця $\bar{A}$ системи $(*)$ записується у вигляді

$$\bar{A} = \left(\mathbf{a_1} \,|\, \mathbf{a_2} \,|\, \ldots \,|\, \mathbf{a_n} \,|\, \mathbf{b} \right).$$

Система лінійних рівнянь однозначно визначається своєю розширеною матрицею.

3.4 Елементарні перетворення рядків матриці

Припустимо, маємо матрицю A, записану у вигляді

$$A = \begin{pmatrix} \underline{\mathbf{a^1}} \\ \underline{\mathbf{a^2}} \\ \vdots \\ \overline{\mathbf{a^m}} \end{pmatrix}, \text{ де } \mathbf{a^i} = (a_{i1}, a_{i2}, \ldots, a_{in}) \text{ - } i-\text{тий рядок матриці } A,\ i = 1, 2, \ldots, m.$$

Означення 3.7. Елементарні перетворення I роду. *Фіксуємо пару індексів $1 \leqslant i \neq j \leqslant m$ і деякий параметр $\lambda \in \mathbb{K}$. Елементарним перетворенням I роду* $\mathbf{e_{ij}}(\lambda)$

матриці A називається перетворення вигляду:

$$A = \begin{pmatrix} \vdots \\ \overline{\underline{\mathbf{a^i}}} \\ \vdots \\ \overline{\underline{\mathbf{a^j}}} \\ \vdots \end{pmatrix} \quad \overset{\{[j]+\lambda*[i]\}}{\Longrightarrow} \quad \mathbf{e_{ij}}(\lambda)(A) = \begin{pmatrix} \vdots \\ \overline{\underline{\mathbf{a^i}}} \\ \vdots \\ \overline{\underline{\mathbf{a^j} + \lambda\mathbf{a^i}}} \\ \vdots \end{pmatrix}.$$

Тобто заміна матриці A на матрицю $\mathbf{e_{ij}}(\lambda)(\mathbf{A})$ розміру $(m \times n)$, у якої всі рядки, крім $j-$го, збігаються з відповідними рядками матриці A, а до $j-$того рядка додається i-тий, помножений на λ.

Елементарні перетворення II роду. *Фіксуємо індекс $1 \leqslant i \leqslant m$ і деякий параметр $\lambda \in \mathbb{K}$, $\lambda \neq 0$. Елементарним перетворенням II роду $\mathbf{e_i}(\lambda)$ матриці A називається перетворення вигляду:*

$$A = \begin{pmatrix} \vdots \\ \overline{\underline{\mathbf{a^i}}} \\ \vdots \end{pmatrix} \quad \overset{\{\lambda*[i]\}}{\Longrightarrow} \quad \mathbf{e_i}(\lambda)(A) = \begin{pmatrix} \vdots \\ \overline{\underline{\lambda\mathbf{a^i}}} \\ \vdots \end{pmatrix}.$$

при якому i-тий рядок матриці A на $\lambda \neq 0$.

Властивості елементарних перетворень.

1. $\mathbf{e_{ij}(0) = id}$;
2. $\mathbf{e_{ij}}(-\lambda)\,\mathbf{e_{ij}}(\lambda) = \mathbf{id} \quad \forall\lambda \in \mathbb{K}$;
3. $\mathbf{e_i(1) = id}$;
4. $\mathbf{e_i}(\lambda^{-1})\mathbf{e_{ij}}(\lambda) = \mathbf{id} \quad \forall\lambda \in \mathbb{K},\ \lambda \neq 0$.

Через $\mathbf{id}$ позначаємо тотожне перетворення $\mathbf{id} : A \longrightarrow A$.

Лемма 3.1. *Нехай маємо деяку систему $(*)$ із розширеною матрицею $\bar{A}$ вигляду (7).*

1. *Позначимо через $\mathbf{e_{ij}}(\lambda)(*)$ систему з розширеною матрицею $\mathbf{e_{ij}}(\lambda)(A)$, $1 \leqslant i \neq j \leqslant m$, $\lambda \in \mathbb{K}$. Тоді система $\mathbf{e_{ij}}(\lambda)(*)$ еквівалентна системі $(*)$, тобто $\mathfrak{L}_{(*)} = \mathfrak{L}_{\mathbf{e_{ij}}(\lambda)(*)}$ (множини їх розв'язків збігаються).*
2. *Якщо $\mathbf{e_i}(\lambda)(*)$ - система з розширеною матрицею $\mathbf{e_i}(\lambda)(A)$, $1 \leqslant i \neq j \leqslant m$, $\lambda \in \mathbb{K}$, $\lambda \neq 0$, то $\mathfrak{L}_{(*)} = \mathfrak{L}_{\mathbf{e_i}(\lambda)(*)}$, отже системи $(*)$ та $\mathbf{e_i}(\lambda)(*)$ еквівалентні.*

Доведення. Доведення містить цікавий типологічний момент. Розглянемо випадок перетворення вигляду $\mathbf{e_{ij}}(\lambda)$ (випадок $\mathbf{e_i}(\lambda)$ - аналогічний). Доведемо, що $\mathfrak{L}_{(*)} \subset \mathfrak{L}_{\mathbf{e_{ij}}(\lambda)(*)}$. Дійсно, розглянемо вектор

$$l = \begin{pmatrix} l_1 \\ l_2 \\ \vdots \\ l_n \end{pmatrix} \in \mathfrak{L}_{(*)} \subset \mathbb{K}^n.$$

Те, що цей вектор є розв'язком, означає, що $\forall i, \;\; 1 \leqslant i \leqslant m$ виконується рівність $a_{i1}l_1 + a_{i2}l_2 + \ldots + a_{in}l_n = b_i$. підставимо значення l в систему $\mathbf{e_{ij}}(\lambda)(*)$. Всі рівняння цієї системи, крім $j-$го, є рівняннями системи $(*)$. Тому після підстановки вони стануть вірними рівностями в $\mathbb{K}$.

Залишилось розглянути $j-$те рівняння і перевірити, чи воно виконується в системі $\mathbf{e_{ij}}(\lambda)(*)$. Розглянемо: $(a_{i1} + \lambda a_{i1})x_1 + \; (a_{i2} + \lambda a_{i2})x_2 + \ldots + (a_{in} + \lambda a_{in})x_n = b_j + \lambda b_i$. підставимо замість $x_1, x_2, \ldots, x_n$ значення $l_1, l_2, \ldots, l_n$. Тоді $(a_{i1} + \lambda a_{i1})l_1 + \; (a_{i2} + \lambda a_{i2})l_2 + \ldots + (a_{in} + \lambda a_{in})l_n = b_j + \lambda b_i$. Це еквівалентно $(a_{j1}l_1 + \; a_{j2}l_2 + \ldots + a_{jn}l_n) + \; \lambda(a_{i1}l_1 + a_{i2}l_2 + \ldots + a_{in}l_n) = b_i + \lambda b_j$. Отже, $j-$те рівняння системи $\mathbf{e_{ij}}(\lambda)(*)$ для вектора l теж виконується, тобто $l \in \mathfrak{L}_{\mathbf{e_{ij}}(\lambda)(*)}$.

Ми довели, що якщо $l \in \mathfrak{L}_{(*)}$, то звідси випливає, що $l \in \mathfrak{L}_{\mathbf{e_{ij}}(\lambda)(*)}$, тобто $\mathfrak{L}_{(*)} \subset \mathfrak{L}_{\mathbf{e_{ij}}(\lambda)(*)}$ - підмножина. Тоді, за доведеним,

$$\mathfrak{L}_{\mathbf{e_{ij}}(\lambda)(*)} \subset \mathfrak{L}_{\mathbf{e_{ij}}(-\lambda)(\mathbf{e_{ij}}(\lambda)(*))} = \mathfrak{L}_{(*)}.$$

Тому $\mathfrak{L}_{(*)} \subset \mathfrak{L}_{\mathbf{e_{ij}}(\lambda)(*)} \subset \mathfrak{L}_{(*)}$. Звідси випливає, що всі включення можна замінити на рівності, тобто $\mathfrak{L}_{(*)} = \mathfrak{L}_{\mathbf{e_{ij}}(\lambda)(*)}$. $\square$

В математиці типовою є ситуація, коли, щоб перейти від одного складного об'єкта до іншого, часто зовсім несхожого, необхідні перетворення робляться крок за кроком. Елементарні перетворення матриць дозволяють перевести довільну матрицю в деяку стандартну.

Лемма 3.2 (Гауса). *Довільну розширену матрицю $\bar{A}$ вигляду (7) елементарними перетвореннями рядків можна звести до стандартного вигляду*

$$(S') \qquad A = \begin{pmatrix} 0 & \ldots & 0 & 0 & 0 & \ldots & 1 & * & \ldots & 0 & * & \ldots & * \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \ldots & 0 & 0 & 0 & \ldots & 0 & 0 & \ldots & 1 & * & \ldots & * \\ 0 & \ldots & 0 & 0 & 0 & \ldots & 0 & 0 & \ldots & 0 & 0 & \ldots & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \ldots & 0 & 0 & 0 & \ldots & 0 & 0 & \ldots & 0 & 0 & \ldots & 0 \end{pmatrix}.$$

де зірочками позначаються довільні елементи поля.

Доведення . Застосуємо метод математичної індукції за кількістю рівнянь системи m.

База індукції. $m = 1$, тобто маємо систему з матрицею

$$\bar{A} = \left(\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & b_1 \end{array}\right).$$

Якщо $a_{11} = a_{12} = \ldots = a_{1n} = 0$, то система має вказаний вигляд (при $k = 0$), тобто $\bar{A} = \left(\begin{array}{cccc|c} 0 & 0 & \dots & 0 & b_1 \end{array}\right)$.

Якщо не всі елементи a_{1j} рівні 0, то нехай $a_{1j} \neq 0$ - перший ненульовий елемент, тобто $\bar{A} = \left(\begin{array}{cccccc|c} 0 & \dots & 0 & a_{1j_1} & a_{1j_1+1} & a_{1n} & b_1 \end{array}\right)$. Тоді нехай $\lambda = a_{1j_1}^{-1}$. Застосуємо перетворення $\mathbf{e_1}(\lambda)$, одержимо

$$\bar{A} = \left(\begin{array}{cccccc|c} 0 & \dots & 0 & 1 & a_{1j_1}^{-1}a_{1j_1+1} & a_{1j_1}^{-1}a_{1n} & a_{1j_1}^{-1}b_1 \end{array}\right).$$

Система набула потрібного вигляду.

Крок індукції. Розглянемо матрицю

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

Виберемо в ній перший зліва ненульовий стовпчик $\mathbf{a_{j_1}}$. При цьому або $\mathbf{a_{1j_1}} \neq \mathbf{0}$, або, якщо $\mathbf{a_{1j_1}} = \mathbf{0}$, то знаходимо ненульовий елемент $\mathbf{a_{ij_1}} \neq \mathbf{0}$ і виконуємо елементарне перетворення $\mathbf{e}_{i1}(1)(*)$, після якого на місці з індексом $1\,j_1$ теж знаходитиметься ненульовий елемент. Одержали матрицю $\bar{A}_1$. Виконаємо над $\bar{A}_1$ елементарне перетворення $\mathbf{e}_1(\lambda)$, $\lambda = a_{1j_1}^{-1}$. Одержана матриця $\bar{A}_2$ має вигляд

$$\bar{A}_2 = \begin{array}{c} \begin{array}{ccccccc} & & & j_1 & & & \end{array} \\ \left(\begin{array}{ccccccc|c} 0 & \dots & 0 & 1 & * & \dots & * & b_1 \\ 0 & \dots & 0 & c_2 & * & \dots & * & b_2 \\ 0 & \dots & 0 & c_3 & * & \dots & * & b_3 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & \dots & 0 & c_m & * & \dots & * & b_m \end{array}\right) \end{array}.$$

де c_2, c_3, ..., $c_m \in \mathbb{K}$ - деякі елементи. Тоді розглянемо таку матрицю $\bar{A}_3$:

$$\bar{A}_3 = \mathbf{e}_{1m}(-c_m)\ldots\mathbf{e}_{13}(-c_3)\mathbf{e}_{12}(-c_2)\bar{A}_3.$$

Тоді матриця $\bar{A}_3$ має вигляд:

$$A = \begin{array}{c} \begin{array}{ccccc} & & & j_1 & \phantom{\boxed{\bar{A}'}} \end{array} \\ \left(\begin{array}{cccc|c} 0 & \dots & 0 & 1 & \\ 0 & \dots & 0 & 0 & \\ 0 & \dots & 0 & 0 & \boxed{\bar{A}'} \\ \vdots & \ddots & \vdots & \vdots & \\ 0 & \dots & 0 & 0 & \end{array}\right) \end{array}$$

Розглянемо розширену матрицю $\bar{A}'$ меншого розміру і відповідну систему рівнянь. Елементарні перетворення над $\bar{A}'$ не змінюють першого рядка. Розмір матриці $\bar{A}'$ - $(m-1) \times (n+1-j_1)$. За припущенням індукції, $\bar{A}'$ зводиться до форми (S'). Тому разом з першим рядком вся матриця зведеться до вигляду (S'). $\square$

Означення 3.8. *Система з розширеною матрицею $\bar{A}$ вигляду (7) називається* **однорідною**, *якщо $b_1 = b_2 = \ldots = b_m = 0$. Однорідна система завжди сумісна, тому що має розв'язок*

$$\begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \in \mathbb{K}^n.$$

Наслідок 3.1 (з методу Гауса)**.** *Якщо в однорідній системі із розширеною матрицею $\bar{A}$ вигляду (7) кількість невідомих більша від кількості рівнянь, тобто $n > m$ (строго), то ця система має ненульовий розв'язок.*

3.5 Метод Гауса (Жордана-Гауса) розв'язування систем лінійних рівнянь

3.5.1 Приклади розв'язування систем лінійних рівнянь

Приклади 3.1. *1. Розглянемо систему*

$$\left\{ \begin{array}{rcl} 0 & = & 1 \end{array} \right.$$

При цьому необхідно вказати кількість невідомих. Множина розв'язків такої системи дорівнює $\varnothing$.

2. Розглянемо систему, що складається з одного рівняння спеціального вигляду:

$$\left\{ \begin{array}{llll} x_1 + & a_{12}x_2 + & \ldots & +a_{1n}x_n & = b_1 \end{array} \right. \tag{8}$$

Розв'язком такої системи є деякий вектор

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

в якому замість $x_2, \ldots, x_n$ можна підставити довільні елементи *поля $\mathbb{K}$, а x_1 покласти рівним $x_1 = -a_{12}x_2 - \ldots - a_{1n}x_n + b_1$. Таким способом одержимо всі розв'язки такої системи. В цьому випадку розв'язок зручно записувати у вигляді:*

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} -a_{12}x_2 - \ldots - a_{1n}x_n + b_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

Розпишемо цей вектор у вигляді суми

$$x = \begin{pmatrix} -a_{12}x_2 \\ x_2 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \begin{pmatrix} -a_{13}x_3 \\ 0 \\ x_3 \\ \vdots \\ 0 \end{pmatrix} + \ldots + \begin{pmatrix} -a_{1n}x_n \\ 0 \\ 0 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} b_1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} =$$

$$x_2 \begin{pmatrix} -a_{12} \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} -a_{13} \\ 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} + \ldots + x_n \begin{pmatrix} -a_{1n} \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} + \begin{pmatrix} b_1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} =$$

Розв'язок системи 8 можна записати у вигляді

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = x_2 \begin{pmatrix} -a_{12} \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} -a_{13} \\ 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} + \ldots + x_n \begin{pmatrix} -a_{1n} \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} + \begin{pmatrix} b_1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

де $x_2, x_3, \ldots, x_n$ *- довільні елементи поля* $\mathbb{K}$. *Це векторна форма запису розв'язку.*

3. Розв'язок системи спеціального вигляду. Припустимо, що розширена матриця системи розміру $(m \times (n+1))$ *має вигляд:*

$$(S) \qquad \bar{A} = \left(\begin{array}{cccccccc|c} 1 & 0 & \ldots & 0 & a_{1k+1} & \ldots & a_{1n} & & b_1 \\ 0 & 1 & \ldots & 0 & a_{2k+1} & \ldots & a_{2n} & & b_2 \\ \vdots & \vdots & \ddots & & \vdots & \ddots & \vdots & & \vdots \\ 0 & 0 & \ldots & 1 & a_{kk+1} & \ldots & a_{kn} & & b_k \\ 0 & 0 & \ldots & & 0 & \ldots & 0 & & b_{k+1} \\ \vdots & \vdots & \ddots & & \vdots & \ddots & \vdots & & \vdots \\ 0 & 0 & \ldots & & 0 & \ldots & 0 & & b_m \end{array}\right).$$

У зображеній матриці (S) *(від англійського standard) мають місце рівності:*

$$\begin{array}{rcll} a_{ii} & = & 1, & 1 \leqslant i \leqslant k; \\ a_{ij} & = & 0, & 1 \leqslant i, j \leqslant k, \quad i \neq j; \\ a_{ij} & = & 0, & i > k. \end{array}$$

Дослідимо цю систему.

1. Припустимо, що для деякого $i \geqslant k+1$ *виконується* $b_i \neq 0$. *В цьому випадку система несумісна.*

2. Вважаємо, що $b_{k+1} = b_{k+2} = \ldots = b_m = 0$. В цьому випадку невідомі x_{k+1}, x_{k+2}, $\ldots$, x_n можуть приймати довільні значення, а x_1, x_2, $\ldots$, x_k однозначно через них виражаються. Запишемо розв'язок системи у векторному вигляді:

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \\ x_{k+1} \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} -\sum\limits_{j=k+1}^{n} a_{1j}x_j + b_1 \\ -\sum\limits_{j=k+1}^{n} a_{2j}x_j + b_2 \\ \vdots \\ -\sum\limits_{j=k+1}^{n} a_{kj}x_j + b_k \\ x_{k+1} \\ \vdots \\ x_n \end{pmatrix} =$$

$$x = x_{k+1}\begin{pmatrix} -a_{1k+1} \\ -a_{2k+1} \\ \vdots \\ -a_{kk+1} \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + x_{k+2}\begin{pmatrix} -a_{1k+2} \\ -a_{2k+2} \\ \vdots \\ -a_{kk+2} \\ 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} + \ldots + x_n\begin{pmatrix} -a_{1n} \\ -a_{2n} \\ \vdots \\ -a_{kn} \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_k \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

де $x_{k+1}, x_{k+2}, \ldots, x_n$ - довільні елементи поля $\mathbb{K}$. Це є загальний розв'язок *нашої системи.*

Коли $k = n$, система визначена, тобто має рівно один розв'язок, який має такий вигляд: $x_1 = b_1$, $x_2 = b_2$, $\ldots$, $x_k = b_k$.

Домовленість 5. *Для спрощення запису ми інколи будемо перенумеровувати невідомі, наприклад, $x_i \longleftrightarrow x_j$ означає, що міняємо місцями невідомі x_i та x_j. при цьому в матриці системи $i-$тий та $j-$тий стовпчики поміняються місцями. Застереження: потрібно не забувати у запису остаточного розв'язку повернути початкову нумерацію.*

Наведене раніше правило розв'язування систем вигляду (S) дозволяє також записувати розв'язки системи, матриця якої має наступний вигляд (S'):

$$(S') \qquad A = \begin{array}{c} \begin{array}{ccccccccccccccc} & & & j_1 & & & j_2 & & & j_k & & & & & n \end{array} \\ \begin{pmatrix} 0 & \dots & 0 & 1 & * & \dots & 0 & * & \dots & 0 & * & \dots & * \\ 0 & \dots & 0 & 0 & 0 & \dots & 1 & * & \dots & 0 & * & \dots & * \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & \dots & 1 & * & \dots & * \\ 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & \dots & 0 & 0 & \dots & 0 \end{pmatrix} \end{array}.$$

Якщо в матриці вигляду (S') ми перемістимо стовпчик з номером j_1 на місце першого, j_2 - другого, а j_k ї на місце $k-$того, одержимо систему вигляду (S), яку ми вже розглядали раніше. Отже, розв'язок її ми також можемо виписати.

Задача. Записати розв'язок системи вигляду (S') як суму стовпчиків.

Метод Гауса полягає в зведенні матриці елементарними перетвореннями $\mathbf{e_{ij}}$, $\mathbf{e_i}$ до стандартного вигляду (S) за індукцією, тобто зводячи задачу до аналогічної задачі для системи меншого розміру.

Ми довели, що система при елементарних перетвореннях переходить в еквівалентну систему, тобто множина розв'язків не зміниться.

За припущенням індукції, матрицю $\bar{A}$ можна звести елементарними перетвореннями рядків до стандартного вигляду (S). Звідси випливає, що і всю матрицю можна звести до вигляду (S). Як? За припущенням індукції, вважаємо, що матриця $\bar{A}'$ має вигляд (S).

Чи буде при цьому вся матриця мати вигляд (S)? ні. Тому, що елементи на місцях з індексами $(1, j_2), (1, j_3), \dots, (1, j_k)$ можуть не дорівнювати нулю. Позначимо ці елементи через $d_2, d_3, \dots, d_k$ відповідно. [4] Тоді матриця, одержана з $\bar{A}_1$ застосуванням перетворень $e_{21}(-d_2)\ e_{31}(-d_3)\ \dots\ e_{k1}(-d_k)$ матиме стандартний вигляд (S), тобто на місцях з індексами $(1, j_2), (1, j_3), \dots, (1, j_k)$ також будуть нулі.

Доведення закінчено.

Це - доведення-алгоритм.

3.5.2 Однорідні лінійні системи

Означення 3.9. *Система з матрицею*

$$\bar{A} = \left(\; \mathbf{a_1} \,|\, \mathbf{a_2} \,|\, \dots \,|\, \mathbf{a_n} \,|\, \mathbf{b} \; \right)$$

називаеться однорідною, якщо $\mathbf{b} = 0$. *Однорідна система з m рівнянь відносно n невідомих, така що $n > m$, має* **ненульовий розв'язок**.

[4] Слово "відповідно"в даному контексті означає, що число d_2 знаходиться на місці j_2, … число d_k - на місці j_k (вважаємо, що у нас перераховані деякі індекси і деякі числа).

Термінологія. Коли матриця системи має стандартну форму (S), то елементи з індексами $(1, j_2)$, $(1, j_3), \ldots,$ $(1, j_k)$ називаються **провідними** елементами. Коли розв'язок системи записаний у вигляді: невідомі x_{j_1}, x_{j_2}, …, x_{j_k} представлені у вигляді лінійних комбінацій невідомих x_i з коефіцієнтами, де $i \neq j_1, j_2, \ldots, j_k$, та деяких вільних членів, причому x_{j_1}, x_{j_2}, …, x_{j_k} при $i \neq j_1, j_2, \ldots, j_k$ можуть приймати довільні значення, тоді говорять, що в такому записі x_{j_1}, x_{j_2}, …, x_{j_k} є **залежними** (або зв'язаними) невідомими, а решта - x_i при $i \neq j_1, j_2, \ldots, j_k$ є **вільними** невідомими або параметрами.

Приклад 3.3. *Розглянемо таку задачу:*

$$A = \left(\begin{array}{cccccccc} 1 & 0 & 0 & a_{1\,k+1} & \ldots & a_{1\,n} & | & b_1 \\ 0 & 1 & 0 & a_{2\,k+1} & \ldots & a_{2\,n} & | & b_2 \end{array} \right),$$

розв'язок якої

$$x_1 = \sum_{j=k+1}^{n} - a_{1j}x_j + b_1, \quad x_2 = \sum_{j=k+1}^{n} - a_{2j}x_j + b_2$$

можна записати скорочено:

$$x_i = \sum_{j=k+1}^{n} - a_{ij}x_j + b_i, \quad i = 1, 2.$$

Тоді вільні невідомі: x_{k+1}, …, x_n, зв'язані - це x_1, …, x_k. Цей розподіл неоднозначний. Наприклад, якщо $x_1 = x_2$, то ми можемо оголосити вільною невідомою або x_1, або x_2.

Критерій сумісності системи, записаної у вигляді (S). Система в стандартному вигляді (S) сумісна тоді і тільки тоді, коли вона не містить рядків вигляду:

$$A = \left(\begin{array}{cccccc} 0 & 0 & \ldots & 0 & | & b_i \end{array} \right), \qquad b_i \neq 0.$$

Критерій визначеності системи вигляду (S). Система в стандартному вигляді (S) є визначеною (має єдиний розв'язок!) тоді і тільки тоді, коли кількість зв'язаних невідомих дорівнює кількості всіх невідомих, або, що еквівалентно, коли кількість вільних невідомих дорівнює нулю, тобто $k = n$. Визначена система звичайно є сумісною, тобто не містить "поганих"рядків.

Доведемо наслідок про однорідні системи. З методу Гауса випливає, що в стандартній формі системи кількість зв'язаних невідомих k менша або дорівнює кількості рівнянь m, тобто $k \leqslant m$.

Оскільки $k \leqslant m$ і $m < n$, одержимо: $k < n$, тобто кількість зв'язаних невідомих строго менша від загальної кількості невідомих. Отже кількість вільних невідомих $n - k > 0$. Оскільки однорідна система сумісна, підставляючи замість вільних невідомих значення, не всі рівні нулю, одержуємо ненульовий розв'язок однорідної системи.

Приклад 3.4. *Припустимо, що* $x_1, \ldots, x_k$ *– зв'язані невідомі,* $x_{k+1}, \ldots, x_n$ *і* $x_i = \sum\limits_{j=k+1}^{n} (-a_{ij})x_j$ *– вільні невідомі. Підставивши* $x_{k+1} = \alpha_{k+1}$, …, $x_n = \alpha_i$ *і припустивши, що* $\exists \alpha_j : \alpha_j \neq 0$, *відразу одержимо ненульовий розв'язок.*

4 Лінійне відображення, пов'язане з матрицею

Візьмемо довільні вектори $\begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix}$, …, $\begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix} \in \mathbb{K}^m$ і розглянемо систему

$\{a_{ij}x_j = b_i$ з $(m \times n)-$матрицею

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

Результатом підстановки вектора $\begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}$ в цю систему буде вектор

$$\begin{pmatrix} a_{11}\alpha_1 + \ldots + a_{1n}\alpha_n \\ \ldots \\ a_{m1}\alpha_1 + \ldots + a_{mn}\alpha_n \end{pmatrix} = \begin{pmatrix} \sum\limits_{j=1}^{n} a_{1j}\alpha_j \\ \ldots \\ \sum\limits_{j=1}^{n} a_{mj}\alpha_j \end{pmatrix} \in \mathbb{K}^m.$$

Означення 4.1. *Добутком* $(m \times n)-$*матриці* $A = (a_{ij})_{\substack{1 \leqslant i \leqslant m, \\ 1 \leqslant j \leqslant n}}$ *над полем* $\mathbb{K}$ *на вектор* $\alpha = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} \in \mathbb{K}^n$ *називається вектор* $A\alpha = \begin{pmatrix} \sum\limits_{j=1}^{n} a_{1j}\alpha_j \\ \ldots \\ \sum\limits_{j=1}^{n} a_{mj}\alpha_j \end{pmatrix} \in \mathbb{K}^m$.

4.1 Основні властивості множення матриці на вектор.

Нагадаємо, що властивості операцій додавання векторів та множення на скаляр вже були вказані. Тобто, для векторів $\alpha = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}$, $\beta = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix} \in \mathbb{K}^n$ визначена сума

$\alpha + \beta = \begin{pmatrix} \alpha_1 + \beta_1 \\ \vdots \\ \alpha_n + \beta_n \end{pmatrix} \in \mathbb{K}^n$, і $\forall \lambda \in \mathbb{K}$ визначається добуток $\lambda\alpha = \begin{pmatrix} \lambda\alpha_1 \\ \vdots \\ \lambda\alpha_n \end{pmatrix} \in \mathbb{K}^n$.

Властивості множення матриці на вектор

1. $A(\alpha + \beta) = A\alpha + A\beta$;
2. $A(\lambda\alpha) = \lambda(A\alpha)$; де A - $(m \times n)-$матриця, α, $\beta \in \mathbb{K}^n$, $\lambda \in \mathbb{K}$.

Зауваження 4.1. *Якщо ввести вектор з невідомих* $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, *то за формулою* $(Ax)_i = \sum\limits_{j=1}^{n} a_{ij}x_j$, *де* $(Ax)_i$ - $i-$*та координата, можна формально означити добуток матриці* A *на вектор-стовпчик невідомих. В цьому випадку систему рівнянь* $(**)$ *можна переписати у вигляді* $Ax = b$, $b = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \in \mathbb{K}^m$. **Розв'язок системи** - *це такий вектор* $\alpha = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} \in \mathbb{K}^n$, *що виконується* $A \cdot \alpha = b$.

Доведення першої властивості.
Розглянемо $i-$ту координату в лівій і правій частинах рівності, яка доводиться.
$(A\alpha + A\beta)_i = (A\alpha)_i + (A\beta)_i = \sum\limits_{j=1}^{n} a_{ij}\alpha_j + \sum\limits_{j=1}^{n} a_{ij}\beta_j = \sum\limits_{j=1}^{n}(a_{ij}\alpha_j + a_{ij}\beta_j) = \sum\limits_{j=1}^{n} a_{ij}(\alpha_j + \beta_j) = \sum\limits_{j=1}^{n} a_{ij}(\alpha + \beta)_j = (A(\alpha + \beta))_i$ для довільного $i = 1, \ldots, m$.

Отже, ми одержали: $(A\alpha + A\beta)_i = (A(\alpha + \beta))_i$, $i = 1, \ldots, m$. Звідси випливає, що $A\alpha + A\beta = A(\alpha + \beta)$.

Приклад 4.1. *Нехай* $A = \begin{pmatrix} 1 & 0 & -1 \\ 2 & 3 & 5 \end{pmatrix}$, $\alpha = \begin{pmatrix} 1 \\ 2 \\ 4 \end{pmatrix}$. *Тоді*

$$A\alpha = \begin{pmatrix} 1_{11} & 0_{12} & -1_{13} \\ 2_{21} & 3_{22} & 5_{23} \end{pmatrix} \begin{pmatrix} 1_1 \\ 2_2 \\ 4_3 \end{pmatrix} = \begin{pmatrix} 1_{11} \cdot 1_1 + 0_{12} \cdot 2_2 + (-1)_{13} \cdot 4_3 \\ 2_{21} \cdot 1_1 + 3_{22} \cdot 2_2 + 5_{23} \cdot 4_3 \end{pmatrix} = \begin{pmatrix} -3 \\ 28 \end{pmatrix}.$$

Доведення другої властивості.

Нехай $A = (a_{ij})_{\left\{\substack{1\leqslant i\leqslant m,\\ 1\leqslant j\leqslant n}\right\}}$ - $(m \times n)-$матриця, $\alpha,\ \beta \in \mathbb{K}^n,\ \lambda \in \mathbb{K}$. Доведемо властивість: $A(\lambda\alpha) = \lambda(A\alpha),\ \lambda \in \mathbb{K}$. Розглянемо $i-$ту координату першого вектора, $1 \leqslant i \leqslant m$. Маємо:

$$(A(\lambda\alpha))_i = \sum_{j=1}^{n} a_{ij}(\lambda\alpha)_j = \sum_{j=1}^{n} a_{ij}(\lambda\alpha_j) = \sum_{j=1}^{n} (a_{ij}\lambda)\alpha_j = \lambda\sum_{j=1}^{n} a_{ij}\alpha_j = \lambda(A\alpha)_i = (\lambda(A\alpha))_i.$$

$\bowtie$

Зауваження 4.2. *В цьому доведенні ми винесли λ за знак суми, при цьому ми скористалися багатьма властивостями операцій в полі $\mathbb{K}$, а саме:*

$$\sum_{j=1}^{n} a_{ij}(\lambda\alpha)_j \overset{\text{асоц.}}{=} \sum_{j=1}^{n} (a_{ij}\lambda)\alpha_j \overset{\text{комут.}}{=} \sum_{j=1}^{n} (\lambda a_{ij})\alpha_j \overset{\text{асоц.}}{=} \sum_{j=1}^{n} \lambda(a_{ij})\alpha_j) \overset{\text{дистр.}}{=} \lambda\sum_{j=1}^{n} a_{ij}\alpha_j. \bowtie$$

Твердження 4.1. *Припустимо, що система (S): $Ax = b$, де $b = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \in \mathbb{K}^m$, сумісна. Розглянемо лінійну однорідну систему (S_0) з тією ж матрицею:*

$$(S_0) : \qquad Ax = 0.$$

Нехай $\alpha \in \mathbb{K}^n$ - довільний розв'язок системи (S). Тоді

$$\mathfrak{L}_S \ = \ \alpha \ + \ \mathfrak{L}_{S_0}, \quad \text{де} \ \ \alpha + \mathfrak{L}_{S_0} = \{\alpha \ + \ \beta \,|\, \beta \in \mathfrak{L}_{S_0}\}.$$

Доведення. α - розв'язок $(S) \iff A\alpha = b$. Доведемо включення $\mathfrak{L}_S \subset \alpha + \mathfrak{L}_{S_0}$. дійсно, якщо $\gamma \in \mathfrak{L}_S$, то $\gamma \in \alpha + \mathfrak{L}_{S_0}$.

Навпаки, покажемо, що $\mathfrak{L}_S \supset \alpha + \mathfrak{L}_{S_0}$. $\gamma \in \alpha + \mathfrak{L}_{S_0} \implies \gamma \in \mathfrak{L}_S$.

Припустимо, що $\gamma \in \mathfrak{L}_S \iff A\gamma = b \implies A(\gamma - \alpha) = A(\gamma + (-1)\alpha) = A\gamma + A((-1)\alpha) = A\gamma + (-1)A\alpha = A\gamma - A\alpha = b - b = 0$.

$A(\gamma - \alpha) = 0 \implies \gamma - \alpha \in \mathfrak{L}_S$.

Навпаки, нехай $\gamma \in \alpha + \mathfrak{L}_{S_0} \iff \gamma = \alpha + \beta,\ \beta \in \mathfrak{L}_{S_0} \iff \gamma = \alpha + \beta,\ A\beta = 0$.

$A\gamma = A(\alpha + \beta) = A\alpha + A\beta = b + 0 = b$.

$A\gamma = b \implies \gamma \in \mathfrak{L}_S$. $\square$

4.2 Поняття лінійної комбінації.

Нехай $(\mathbf{a_1}, \mathbf{a_2}, \ldots, \mathbf{a_n})$ - сімейство векторів, $\mathbf{a_i} \in \mathbb{K}^m$, $i = 1, \ldots, n$. Якщо $\alpha_1,\ \alpha_2,\ \ldots,$ $\alpha_n \in \mathbb{K}$, тоді вектор $\alpha_1\mathbf{a_1} + \alpha_2\mathbf{a_2} + \ldots + \alpha_n\mathbf{a_n} \in \mathbb{K}^m$ називається лінійною комбінацією векторів $\mathbf{a_1}, \mathbf{a_2}, \ldots, \mathbf{a_n}$ з коефіцієнтами $\alpha_1, \alpha_2, \ldots, \alpha_n$ (або просто лінійною комбінацією, якщо ясно, про які вектори та коефіцієнти іде мова).

Лінійна комбінація $\alpha_1\mathbf{a_1} + \alpha_2\mathbf{a_2} + \ldots + \alpha_n\mathbf{a_n}$ називається *тривіальною*, коли $\alpha_1 = \alpha_2 = \ldots = \alpha_n = 0$ і *нетривіальною* в протилежному випадку (нетривіальна $\iff \exists i$ $\alpha_i \neq 0$).

4.2.1 Множення матриці на вектор та лінійні комбінації.

Припустимо, що матриця $A = (a_{ij})_{\left\{\substack{1\leqslant i\leqslant m,\\ 1\leqslant j\leqslant n}\right\}}$ представлена у вигляді n векторів-стовпчиків

$$A = \left(\ \mathbf{a_1}\,|\,\mathbf{a_2}\,|\ \ldots\ |\,\mathbf{a_n}\ \right), \quad a_i \in \mathbb{K}^m,\ i = 1, \ldots, n.$$

Тоді утворити лінійну комбінацію $\alpha_1\mathbf{a_1}+\ \alpha_2\mathbf{a_2}+\ \ldots\ +\alpha_n\mathbf{a_n}$ цих векторів з коефіцієнтами α_1, α_2, ..., α_n - це те саме, що помножити матрицю A на вектор $\alpha = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix}$ $\in \mathbb{K}^n$. Маємо

$$A\alpha\ =\ \alpha_1\mathbf{a_1} + \alpha_\mathbf{2}\mathbf{a_2} + \ldots + \alpha_\mathbf{n}\mathbf{a_n}.$$

Доведення. За означенням,

$$A\alpha = \begin{pmatrix} \sum\limits_{j=1}^{n} a_{1j}\alpha_j \\ \ldots \\ \sum\limits_{j=1}^{n} a_{mj}\alpha_j \end{pmatrix} = \sum_{j=1}^{n} \begin{pmatrix} a_{1j}\alpha_j \\ \ldots \\ a_{mj}\alpha_j \end{pmatrix} = \sum_{j=1}^{n} \alpha_j \begin{pmatrix} a_{1j} \\ \ldots \\ a_{mj} \end{pmatrix} = \sum_{j=1}^{n} \alpha_j \mathbf{a_j}$$ - лінійна комбінація.

4.3 Лінійне відображення координатних векторних просторів, пов'язане з матрицею.

Нехай $A = (a_{ij})_{\left\{\substack{1\leqslant i\leqslant m,\\ 1\leqslant j\leqslant n}\right\}}$ - $(m \times n)-$матриця з коефіцієнтами з поля $\mathbb{K}$. Тоді через L_A позначимо відображення $\mathrm{L}_A : \mathbb{K}^n \to \mathbb{K}^m$ таке, що $\mathrm{L}_A(\alpha) = A\,\alpha$, $\alpha \in \mathbb{K}^n$. $\square$

4.3.1 Властивості відображення $\mathbf{L}_A$.

1. $\mathrm{L}_A(\alpha + \beta) = \mathrm{L}_A(\alpha) + \mathrm{L}_A(\beta), \quad \alpha, \beta \in \mathbb{K}^n.$
2. $\mathrm{L}_A(\lambda\alpha) = \lambda \mathrm{L}_A(\alpha), \quad \alpha \in \mathbb{K}^n,\ \lambda \in \mathbb{K}.$

Перевіримо ці властивості.

1. $\mathrm{L}_A(\alpha + \beta) = A(\alpha + \beta) = A\alpha +\ A\beta = \mathrm{L}_A(\alpha) + \mathrm{L}_A(\beta) =$ - вирази рівні.
2. $\mathrm{L}_A(\lambda\alpha) = A(\lambda\alpha) = \lambda(A\alpha) = \lambda \mathrm{L}_A(\alpha).$

Властивості 1, 2 об'єднують спільною назвою: кажуть, що L_A - лінійне відображення (над полем $\mathbb{K}$), або $\mathbb{K}-$лінійне.

Нульовий вектор.

Існує вектор $\mathbf{0} \in \mathbb{K}^{\mathbf{n}}$, $\mathbf{0} = \begin{pmatrix} 0 \\ 0 \\ \dots \\ 0 \end{pmatrix}$, такий, що $\mathrm{L}_A(\mathbf{0}) = \mathbf{0}$.

Доведення.

1. $A\mathbf{0} = A \begin{pmatrix} 0 \\ \cdots \\ 0 \end{pmatrix} = \sum\limits_{j=1}^{n} a_{ij} \cdot 0 = 0,\ i = 1, \dots, m \implies A \begin{pmatrix} 0 \\ \cdots \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ \cdots \\ 0 \end{pmatrix}.$

2. Нульовий вектор $\mathbf{0} \in \mathbb{K}^{\mathbf{n}}$ визначається властивістю $\mathbf{0} + \alpha = \alpha$ (для будь-якого $\alpha \in \mathbb{K}^n$). Аналогічно в $\mathbb{K}^m$.

$\mathrm{L}_A(\mathbf{0}) = \mathrm{L}_A(\mathbf{0} + \mathbf{0}) = \mathrm{L}_A(\mathbf{0}) + \mathrm{L}_A(\mathbf{0})$. Додамо до правої і лівої частин $(-\mathrm{L}_A(\mathbf{0}))$, маємо $\quad \mathrm{L}_A(\mathbf{0}) + (-\mathrm{L}_A(\mathbf{0})) = (\mathrm{L}_A(\mathbf{0}) + \mathrm{L}_A(\mathbf{0}) + (-\mathrm{L}_A(\mathbf{0})))$.

$\mathbb{K}^m \ni \mathbf{0} = \mathrm{L}_A(\mathbf{0})$. $\quad \mathrm{L}_A(\mathbf{0}) + (\mathrm{L}_A(\mathbf{0}) - \mathrm{L}_A(\mathbf{0})) = \mathrm{L}_A(\mathbf{0}) + \mathbf{0} = \mathrm{L}_{\mathbf{A}}(\mathbf{0})$. $\quad\square$

Домовленість 6. *Надалі вектори* $\mathbf{a}, \mathbf{0} \in \mathbb{K}^{\mathbf{n}}$ *часто, якщо це не призведе до непорозуміння, позначатимемо як* $a, 0 \in \mathbb{K}^n$ *(тобто нежирними буквами).*

Лемма 4.1. *Нехай* A *-* $(m \times n)$ *матриця над полем* $\mathbb{K}$. *Відображення* $\mathrm{L}_A : \mathbb{K}^n \to \mathbb{K}^m$ *буде мономорфізмом тоді і тільки тоді, коли з* $\mathrm{L}_A(\alpha) = 0$ *випливає, що* $\alpha = 0$.

Доведення. L_A - мономорфізм. Як доведено, $\mathrm{L}_A(0) = 0 \implies$ для $\alpha \neq 0$ $\mathrm{L}_A(\alpha) \neq \mathrm{L}_A(0)$, тобто $\mathrm{L}_A(\alpha) \neq 0$.

Припустимо, що L_A - не мономорфізм. Тоді побудуємо $\alpha \neq 0$ такий, що $\mathrm{L}_A(\alpha) = 0$.

L_A - не мономорфізм $\implies \exists \beta, \gamma \in \mathbb{K}^n, \beta \neq \gamma, \mathrm{L}_A(\beta) = \mathrm{L}_A(\gamma) \implies \mathrm{L}_A(\beta) - \mathrm{L}_A(\gamma) = 0 \implies \mathrm{L}_A(\beta) + (-1)\mathrm{L}_A(\gamma) = 0 \implies \mathrm{L}_A(\beta) + \mathrm{L}_A((-1)\gamma) = 0 \implies \mathrm{L}_A(\beta + (-1)\gamma) = 0 \implies \mathrm{L}_A(\beta - \gamma) = 0,\ \beta - \gamma \neq 0 \implies \alpha = \beta - \gamma,\ \mathrm{L}_A(\alpha) = 0,\ \alpha \neq 0.$ $\quad\square$

Завдання 4.1. *Перевірити такі властивості лінійного відображення* L_A *(на матрицях або як властивості лінійних відображень):*

$\mathrm{L}_A(\alpha - \beta) = \mathrm{L}_A(\alpha) - \mathrm{L}_A(\beta)$.

$\mathrm{L}_A(\lambda_1\alpha + \lambda_2\beta) = \lambda_1\mathrm{L}_A(\alpha) + \lambda_2\mathrm{L}_A(\beta)$, *де* $\lambda_1,\ \lambda_2 \in \mathbb{K}$.

Означення 4.2. *Нехай* $A = (a_{ij})_{\left\{\substack{1\leqslant i\leqslant m, \\ 1\leqslant j\leqslant n}\right\}}$. *Множина* $\alpha \in \mathbb{K}^n$ *таких, що* $\mathrm{L}_A(\alpha) = 0$, *називається* **ядром** *лінійного відображення* L_A *і позначається* $\mathrm{K{\small ER}}\ \mathrm{L}_A$. *Множина* $\alpha' \in \mathbb{K}^m$ *таких, що* $\mathrm{L}_A(\alpha) = \alpha'$ *для деякого* $\alpha \in \mathbb{K}^n$, *називається* **образом** *лінійного відображення* L_A *і позначається* $\mathrm{I{\small M}}\,\mathrm{L}_A$. *В скороченому запису маємо:*

$$\mathrm{K{\small ER}}\ \mathrm{L}_A \ = \ \{\alpha \in \mathbb{K}^n \,|\, \mathrm{L}_A(\alpha) = 0\} \ \subset \ \mathbb{K}^n.$$

$$\mathrm{I{\small M}}\,\mathrm{L}_A \ = \ \{\alpha' \in \mathbb{K}^m \,|\, \exists \alpha \in \mathbb{K}^n \ : \ \mathrm{L}_A(\alpha) = \alpha'\} \ \subset \ \mathbb{K}^m.$$

Образ ще іноді записують: $\operatorname{Im} \mathrm{L}_A = \mathrm{L}_A(\mathbb{K}^n)$. Поняття ядра, на відміну від образу, специфічне саме для лінійного відображення.

Наведемо зведену таблицю для системи лінійних рівнянь з $m \times n-$матрицею A та лінійним відображенням L_A:

$$(*) \quad Ax \ = \ b \qquad\qquad \begin{array}{rcl} \mathrm{L}_A : \ \mathbb{K}^n & \to & \mathbb{K}^m \\ \alpha & \mapsto & A\,\alpha \end{array}$$

1) α є розв'язком системи $(*)$	1) $\mathrm{L}_A(\alpha) = b$ - I спосіб $\alpha \in \mathrm{L}_A^{-1}(b)$ - II спосіб
2) Знайти всі розв'язки системи $Ax = b$	2) Знайти $\mathrm{L}_A^{-1}(b)$ - повний прообраз
3) $\forall\, b \in \mathbb{K}^m$ система $(*)$ має розв'язок	3) L_A - епіморфізм, або $\operatorname{Im} \mathrm{L}_A = \mathbb{K}^m$
4) $\forall\, b \in \mathbb{K}^m$ система $(*)$ має єдиний розв'язок	4) L_A - бієкція або ізоморфізм
5) Система $(*)$ має не більше одного розв'язку	5) L_A - мономорфізм
6) Розв'язати однорідну систему $Ax = 0$	6) Знайти ядро $\operatorname{Ker} \mathrm{L}_A$
7) Знайти ФСР системи $Ax = 0$	7) Знайти базис ядра $\operatorname{Ker} \mathrm{L}_A$
8) Знайти всі $b \in \mathbb{K}^m$, для яких $(*)$ сумісна	8) Знайти образ $\operatorname{Im} \mathrm{L}_A$
9) Знайти макс.лін.незал. систему стовпчиків	9) Знайти базис образа $\operatorname{Im} \mathrm{L}_A$

Тепер нам знадобляться деякі більш загальні поняття, ніж ті, які ми розглядали до цього часу.

5 (Абстрактний) векторний простір

Слово **абстрактний** вживають тільки тоді, коли його хочуть протиставити координатному векторному простору $\mathbb{K}^n$, $n \geqslant 1$.

Фіксуємо поле $\mathbb{K}$, елементи якого називаємо ще *скалярами*.

Означення 5.1. *Множина V називається* векторним простором *над полем $\mathbb{K}$, якщо на ній введено дві операції: додавання $(+)$ векторів та множення $(\times$ або $\cdot)$ вектора на скаляр:*

$$\begin{array}{rcl} V \times V & \to & V \\ (u, v) & \mapsto & u + v \end{array} \qquad\qquad \begin{array}{rcl} \mathbb{K} \times V & \to & V \\ (\lambda, v) & \mapsto & V \end{array}$$

причому ці операції задовольняють наведеним далі умовам (або аксіомам).

1. *Асоціативність додавання:*

$$\forall v_1, v_2, v_3 \ \in \ V \qquad (v_1 + v_2) + v_3 \ = \ v_1 + (v_2 + v_3).$$

2. *Існування нуля:*

$$\exists \text{ елемент } 0 \in V \text{ такий, що } \forall v \in V \qquad 0 + v \ = \ v + 0 \ = \ v.$$

3. *Існування протилежного елемента:*

$$\forall v \in V \ \exists \ -v \in V \qquad v + (-v) \ = \ (-v) + v \ = \ 0.$$

4. *Комутативність додавання:*

$$\forall v_1, v_2 \in V \qquad v_1 + v_2 \ = \ v_2 + v_1.$$

5. *Асоціативність множення на скаляр:*

$$\forall \lambda_1, \lambda_2 \in \mathbb{K}, \ \forall v \in V \qquad \lambda_1(\lambda_2 v) \ = \ (\lambda_1 \cdot \lambda_2)v.$$

6. *Унітарність:*

$$\forall v \in V \qquad 1 \cdot v \ = \ v.$$

7. *Два закони дистрибутивності:*

$$\forall \lambda_1, \lambda_2 \in \mathbb{K}, \ \forall v \in V (\lambda_1 + \lambda_2)v \ = \ \lambda_1 v + \lambda_2 v,$$

$$\forall \lambda \in \mathbb{K}, \ \forall v_1, v_2 \in V \lambda(v_1 + v_2) \ = \ \lambda v_1 + \lambda v_2.$$

Ці властивості ми вже згадували, розглядаючи координатний векторний простір. Дамо означення під-, под-, sub-, under- простору.

Означення 5.2. *Нехай V - векторний простір над полем $\mathbb{K}$ з операціями $+, \cdot$. Підмножина $W \subset V$ називається (векторним або лінійним) $\mathbb{K}-$підпростором в просторі V, якщо W є векторним простором відносно тих само операцій $+$ та $\cdot$.*

Це означає, що для W з тими ж операціями виконуються всі умови 1-7. Тобто, $W \subset V$ є підпростором, якщо для всіх $\lambda \in \mathbb{K}, w, w_1, w_2 \in W$ виконуються наступні умови:

1. $\forall \ w_1, w_2 \in W \qquad w_1 + w_2 \in W;$

2. $\forall \lambda \in \mathbb{K} \quad \forall w \in W \qquad \lambda w \in W;$

3. $0 \in W;$

4. $\forall w \in W \qquad -w \in W.$

Зауваження 5.1. *1) аксіоми векторного простору для W виконуються автоматично;*
2) як буде доведено, умови 3,4 в означенні підпростору є надлишковими, вони випливають з умов 1 та 2.

Приклад векторного простору. Координатний векторний простір $V = \mathbb{K}^n$, $n \geqslant 1$ є абстрактним векторним простором.

Деякі властивості

Властивість 1. В означенні нейтрального елемента $0 \in V$ не вимагається його єдиність. Ми доведемо, що нейтральний елемент єдиний.

Доведення. Припустимо, що $0, 0' \in V$ - два нейтральних елементи. Розглянемо тоді їх суму $0 + 0' : 0 + 0' = 0'$, оскільки 0 - нейтральний. Оскільки $0'$ - нейтральний, то $0 + 0' = 0$. Звідси випливає: $0 = 0'$. $\square$

Властивість 2. Протилежний елемент $(-v) \in V$ єдиний.

Доведення. Припустимо, для елемента $v \in V$ існує два протилежних $(-v)$ і $(-v)'$. Розглянемо суму $(-v) + v + (-v)'$. Скористаємося асоціативністю.

$((-v) + v) + (-v)' = 0 + (-v)' = (-v)'$

$(-v) + (v + (-v)') = (-v) + 0 = (-v)$

Отже, елементи $(-v)$, $(-v)'$ між собою рівні. $\square$

Властивість 3. Нехай $v_1, v_2 \in V$ такі, що $v_1 + v_2 = v_1$. Тоді $v_2 = 0$.

Доведення. Додамо до обох частин рівності $(-v_1)$. Одержимо $(-v_1) + v_1 + v_2 = (-v_1)+v_1$. Звідси, скориставшись асоціативністю, одержимо $((-v_1)+v_1)+v_2 = ((-v_1)+v_1) \iff 0 + v_2 = 0$ або $v_2 = 0$. Аналогічно, якщо $v_2 + v_1 = v_1$, то $v_2 = 0$. $\square$

Властивість 4. Для довільного $v \in V$ та $0 \in \mathbb{K}$ виконується $0 \cdot v = 0$.

Доведення. Перший спосіб підрахунку виразу $(0 + 1)v$: $(0 + 1)v = 1 \cdot v = v$.

Другий спосіб (скористаємося дистрибутивністю): $(0+1)v = 0 \cdot v + 1 \cdot v = 0 \cdot v + v$.

Отже, $v = 0 \cdot v + v$. Звідси, за попередньою властивістю, $0 \cdot v = 0$. $\square$

Властивість 5. Для довільного $v \in V \quad (-1) \cdot v = -v$.

Доведення. Ми знаємо, що протилежний елемент єдиний. Тоді перевіримо, що $v + (-1)v = 0$ і $(-1)v + v = 0$. Звідси випливатиме, що $(-1)v$ - протилежний.

$v+(-1) \cdot v = 1 \cdot v+(-1) \cdot v = (1+(-1)) \cdot v = 0 \cdot v = 0$, оскільки в полі $\mathbb{K}$ $1+(-1) = 0$. Скориставшись комутативністю, одержимо: $(-1)v + v = v + (-1) \cdot v = 0$. $\square$

Лемма 5.1. *Припустимо, що в векторному просторі V над $\mathbb{K}$ підмножина $W \subset V$ задовольняє таким властивостям:*

1. $\forall\ w_1, w_2 \in W \quad w_1 + w_2 \in W$; *2. $\forall\ \lambda \in \mathbb{K}$, $\forall\ w \in W\ \lambda w \in W$.*

Тоді W є підпростором.

Доведення. Необхідно довести, що $0 \in W$ і для будь-якого $w \in W$ $(-w) \in W$. Але це випливає з властивості 2.

Розглянемо $w \in W$. Тоді $0 = o \cdot w \in W$. Навпаки, $\forall\ w \in W\ (-1) \cdot w \in W$. Тоді $(-1) \cdot w = -w \in W$. $\square$

Отже, щоб перевірити, що W є підпростором, треба перевірити властивості 1, 2, а решта виконується автоматично.

Лемма 5.2. *Нехай A - це деяка $n \times m$-матриця над полем $\mathbb{K}$, $\mathrm{L}_A \in \mathbb{K}^n \to \mathbb{K}^m$ - відповідне їй лінійне відображення: $\mathrm{L}_A(\alpha) = A \cdot \alpha$. Тоді*

1. $\mathrm{Ker}\ \mathrm{L}_A \subset \mathbb{K}^n$ є підпростором.

2. $\mathrm{Im}\, \mathrm{L}_A \subset \mathbb{K}^m$ є підпростором.

Доведення. Нехай $\alpha, \beta \in \mathrm{Ker}\ \mathrm{L}_A$. Треба довести, що $\alpha + \beta \in \mathrm{Ker}\ \mathrm{L}_A$.

$\alpha, \beta \in \mathrm{Ker}\ \mathrm{L}_A \iff \mathrm{L}_A(\alpha) = 0, \mathrm{L}_A(\beta) = 0$. Обчислимо $\mathrm{L}_A(\alpha+\beta) = \mathrm{L}_A(\alpha)+\mathrm{L}_A(\beta) = 0 + 0 = 0$. Звідки випливає, що $\alpha + \beta \in \mathrm{Ker}\ \mathrm{L}_A$.

Аналогічно, якщо $\lambda \in \mathbb{K}$ і $\alpha \in \mathrm{Ker}\ \mathrm{L}_A$, то тоді $\mathrm{L}_A(\lambda\alpha) = \lambda \mathrm{L}_A(\alpha) = \lambda \cdot 0 = \lambda \cdot (0 \cdot 0) = (\lambda \cdot 0) \cdot 0 = 0$ (за визначенням лінійності).

Отже, $\mathrm{Ker}\ \mathrm{L}_A$ є підпростором. Доведемо, що $\mathrm{Im}\, \mathrm{L}_A$ буде підпростором в $\mathbb{K}^m$. Припустимо, $b_1, b_2 \in \mathrm{Im}\, \mathrm{L}_A$. Це означає: $\exists \alpha_1, \alpha_2 \in \mathbb{K}^n$, такі що $b_1 = \mathrm{L}_A(\alpha_1)$, $b_2 = \mathrm{L}_A(\alpha_2)$. Тоді $b_1 + b_2 = \mathrm{L}_A(\alpha_1)+ \mathrm{L}_A(\alpha_2) = \mathrm{L}_A(\alpha_1 + \alpha_2) \implies b_1 + b_2 \in \mathrm{Im}\, \mathrm{L}_A$.

Залишилось перевірити, що якщо $\lambda \in \mathbb{K}$ і $b \in \mathrm{Im}\, \mathrm{L}_A$, то $\lambda b \in \mathrm{Im}\, \mathrm{L}_A$. дійсно, $b \in \mathrm{Im}\, \mathrm{L}_A \implies \exists \alpha \in \mathbb{K}^n$: $\mathrm{L}_A(\alpha) = b$. Але тоді $\lambda \cdot v = \lambda \cdot \mathrm{L}_A(b) = \mathrm{L}_A(\lambda b)$, за властивістю лінійного відображення. Звідси випливає, що $\lambda v \in \mathrm{Im}\, \mathrm{L}_A$. $\square$

5.1 Підпростори

5.1.1 Приклади підпросторів

1. $\{0\}$ - підпростір (нульовий) в V.

2. V є власним підпростором.

Доведення. $0 + 0 = 0$, отже, властивість 1 в означенні підпростору виконується. Потрібно довести, що $\forall \lambda \in \mathbb{K}\ \lambda \cdot 0 = 0$, де $0 \in V$. Для $0 \in V$ виконується рівність: $0_{\mathbb{K}} \cdot 0_V = 0_V$, де $0_{\mathbb{K}}$ - це нуль в полі $\mathbb{K}$. Звідси $\lambda \cdot 0_V = \lambda \cdot (0_{\mathbb{K}} \cdot 0_V) = (\lambda \cdot 0_{\mathbb{K}}) \cdot 0_V = o_{\mathbb{K}} \cdot 0_V = 0_V$. $\square$

Припустимо, в просторі V вибрали деяку кількість векторів $v_1, \ldots, v_k \in V$ і $\lambda_1, \ldots, \lambda_k \in \mathbb{K}$. Тоді сума вигляду $\lambda_1 v_1 + \ldots + \lambda_k v_k = \sum_{i=1}^{k} \lambda_i v_i$ називається **лінійною комбінацією** векторів $v_1, \ldots, v_k$ (з коефіцієнтами $\lambda_1, \ldots, \lambda_k$).

Більш загально, для довільного сімейства векторів (можливо, нескінченного) $\{v_i\}_{i \in I}$ можемо розглядати лінійні комбінації $\sum_{i \in I} \lambda_i v_i$, $\lambda_i \in \mathbb{K}$; при цьому вважаємо, що лише скінченна кількість λ_i не дорівнює 0.

Нехай маємо вектори $v_1, \ldots, v_k$, $v \in V$. Говоримо, що v **лінійно виражається** через $v_1, \ldots, v_k$, якщо існують $\lambda_1, \ldots, \lambda_k \in \mathbb{K}$ такі, що $v = \sum_{i \in I} \lambda_i v_i$. (У випадку нескінченного сімейства векторів $\{v_i\}_{i \in I}$ вектор v лінійно виражається через сімейство

$\{v_i\}_{i\in I}$, що означає, що $v = \sum\limits_{i\in I}\lambda_i v_i$, причому лише скінчена кількість λ_i в цій сумі відмінна від нуля).

5.1.2 Поняття лінійної оболонки

Нехай $\{v_i\}_{i\in I}$, $v_i \in V$, - це деяке сімейство (або система) векторів. Тоді їх **лінійною оболонкою** в просторі V називається множина всіх лінійних комбінацій $\sum\limits_{i\in I}\lambda_i v_i$, де $\lambda_i \in \mathbb{K}$, $i \in I$ (і тільки скінченна кількість ненульових λ_i).

Лінійна оболонка позначається $(v_i)_{i\in I}$ для сімейства векторів $\{v_i\}_{i\in I}$, $v_i \in V$, або $(v_1, \dots, v_n)$ для скінченної системи векторів $\{v_1, \dots, v_n\}$.

Лемма 5.3. *Лінійна оболонка системи векторів $\{v_i\}_{i\in I}$ в просторі V є підпростором простору V.*

Доведення. Нехай $W = (v_i)_{i\in I}$ - це лінійна оболонка векторів $\{v_i\}_{i\in I}$. Тоді достатньо довести, що якщо $w_1, w_2 \in W$, то звідси випливає, що $w_1 + w_2 \in W$ і якщо $\lambda \in \mathbb{K}$, $w \in W$, то маємо $\lambda w \in W$.

Якщо $w_1 \in W$, то це еквівалентно умові, що $w_1 = \sum\limits_{i\in I}\lambda_i v_i$ (деяка скінченна сума). Аналогічно $w_2 = \sum\limits_{i\in I}\mu_i v_i$ (скінченна сума), де всі λ_i, $\mu_i \in \mathbb{K}$, $i \in I$, $j \in I$. Тоді $w_1 + w_2 = \sum\limits_{i\in I}\lambda_i v_i + \sum\limits_{i\in I}\mu_i v_i = \sum\limits_{i\in I}(\lambda_i + \mu_i)v_i =$ - теж скінченна сума. Отже, звідси випливає, що $w_1 + w_2 \in W$.

Аналогічно, якщо $w \in W$ має вигляд $w = \sum\limits_{i\in I}\lambda_i v_i =$ (скінченна сума), то звідси випливає, що $\lambda w = \sum\limits_{i\in I}(\lambda \cdot \lambda_i)v_i \in W$ (є скінченною сумою). □

Приклад 5.1. $V = \mathbb{R}^3$.

Якщо $v_1 = 0$, то $(v_1) = \{0\}$.

Якщо $v_1 \neq 0$, то $(v_1) = \mathbb{R}v_1$ - пряма, що проходить через початок координат.

Якщо $v_1, v_2 \in V$ не колінеарні, то $(v_1, v_2) = \mathbb{R}v_1 + \mathbb{R}v_2$ - площина, що проходить через початок координат.

5.1.3 Терміни та позначення

1. Нехай $\{v_i\}_{i\in I}$ - це деяка система векторів. Говоримо, що система $\{v_i\}_{i\in I}$ **лінійно залежна**, якщо існує набір таких $\{\lambda_i\}_{i\in I}$, серед яких знайдеться хоча б одне ненульове λ_i $(\lambda_i \neq 0)$ і скінченна кількість елементів системи не рівна 0, така що $\sum\limits_{i\in I}\lambda_i v_i = 0$.

Іншими словами, система $\{v_i\}_{i\in I}$ лінійно залежна $\iff$ коли існує лінійна комбінація $\sum\limits_{i\in I}\lambda_i v_i = 0$ така, що не всі $\lambda_i = 0$.

Лемма 5.4. *Нехай V - деякий простір, $W \subset V$ - підпростір і $\{v_i\}_{i\in I}$ - це деяка система векторів $v_i \in V$ таких, що для довільного $i \in I$ $v_i \in W$. Тоді лінійна оболонка $(v_i)_{i\in I} \subset V$ - теж належить W.*

Іншими словами, $\{v_i\}_{i\in I}$ є підпростором не лише у всьому V але і в W.

Доведення. Доведення очевидне. За означенням підпростору, якщо $v_i \in W$ для будь-якого $i \in I$, то довільна лінійна композиція $\sum\limits_{i\in I}\lambda_i v_i$ теж належить W (тобто операції суми і множення на скаляр не виводять з підпростору). □

Лемма 5.5. *Нехай $\{v_i\}_{i\in I}$ і $\{w_j\}_{j\in J}$ - дві системи векторів в просторі V. Тоді наступні твердження еквівалентні:*

1. *лінійна оболонка $(v_i)_{i\in I}$ - підмножина в $(w_j)_{j\in J}$, тобто $(v_i)_{i\in I} \subset (w_j)_{j\in J}$;*

2. *кожен вектор v_i, $i \in I$, лінійно виражається через $\{w_j\}_{j\in J}$.*

Доведення. **1 $\implies$ 2** Припустимо, що $(v_i)_{i\in I} \subset (w_j)_{j\in J}$. Тоді розглянемо довільний вектор $v_i \in (w_j)_{j\in J}$. Лінійна оболонка складається з лінійних комбінацій, отже, $v_i = \sum\limits_{j\in J}\lambda_j w_j$ для деяких $\lambda_j \in \mathbb{K}$.

2 $\implies$ 1 Навпаки, припустимо, що кожен $(v_i)_{i\in I}$, є лінійною комбінацією w_j, $j \in J$, тобто $v_i = \sum\limits_{j\in J}\lambda_j w_j$ для деяких $\lambda_j \in \mathbb{K}$. Тоді довільний елемент w з лінійної оболонки $(v_i)_{i\in I}$ має вигляд: $w = \sum\limits_{i\in I}\mu_i v_i$. Підставимо вираз для v_i. Тоді

$$w = \sum_{i\in I}\mu_i v_i = \sum_{i\in I}\mu_i\Big(\sum_{j\in J}\lambda_{ij} w_j\Big) = \sum_{j\in J}\Big(\sum_{i\in I}\mu_i\lambda_{ij}\Big)w_j.$$

Легко бачити, що одержана сума містить тільки скінченну кількість ненульових доданків. Звідси випливає, що w є лінійною комбінацією $\{w_j\}_{j\in J}$, а отже $w \in (w_j)_{j\in J}$. Тобто, кожен вектор з лінійної оболонки $(v_i)_{i\in I}$ потрапляє в (w). □

Наслідок 5.1. *Нехай $\{v_i\}_{i\in I}$, $\{w_j\}_{j\in J}$, $\{u_k\}_{k\in K}$ - системи векторів в просторі V такі, що $w_j \in (v_i)_{i\in I}$ (або, іншими словами, кожен w_j є лінійною комбінацією v_i), та $u_k \in (w_j)_{j\in J}$. Тоді $u_k \in (v_i)_{i\in I}$.*

5.2 Лінійні відображення векторних просторів

Нехай V, W – векторні простори над $\mathbb{K}$. Відображення $f : V \to W$ називається **лінійним**, якщо виконуються такі властивості:

1. $\forall\ v_1, v_2 \in V \quad f(v_1 + v_2) = f(v_1) + f(v_2)$;

2. $\forall\ \lambda \in \mathbb{K}$ та $v \in V \quad f(\lambda v) = \lambda f(v)$.

Приклади 5.1. *1) Якщо $V = \mathbb{K}^n$, $W = \mathbb{K}^m$ і A - це деяка $n \times m-$матриця, то відображення $\mathrm{L}_A : \mathbb{K}^n \to \mathbb{K}^m$, при якому $v \mapsto Av$, – лінійне.*

2) Нульове відображення $V \to W$, при якому $\forall\, v \mapsto 0$, є лінійним. Якщо $V = \mathbb{K}^n$, $W = \mathbb{K}^m$, то воно відповідає L_A для матриці $AA = (a_{ij}), \quad 1 \leqslant i \leqslant m, 1 \leqslant j \leqslant n$ або $A = (a_{ij})_{\left\{\substack{1\leqslant i\leqslant m,\\ 1\leqslant j\leqslant n}\right\}}$, у якої всі елементи a_{ij} нульові.

3) Для всякого простору V тотожне відображення $\mathtt{1}_V : V \to V$ таке, що $\mathtt{1}_V(v) = v \;\forall\; v \in V$ (його іноді позначають id_V) є лінійним.

Якщо $V = \mathbb{K}^n$, то яка матриця відповідає тотожному відображенню $\mathtt{1}_{\mathbb{K}^n}$? Така матриця розміру $n \times n$ існує і називається **одиничною**. Позначається

$$I_n = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}.$$

Якщо позначимо $I_n = (a_{ij}), \quad 1 \leqslant i \leqslant m, 1 \leqslant j \leqslant n$ або $A = (a_{ij})_{\left\{\substack{1\leqslant i\leqslant m,\\ 1\leqslant j\leqslant n}\right\}}$, то $a_{ij} = 1$, якщо $i = j$, і $a_{ij} = 1$, якщо $i \neq j$.

Позначення. Нехай M – довільна множина і $\alpha, \beta \in M$ – два елементи. Тоді *символ Кронекера* $\delta_{\alpha\beta}$ визначається умовою $\delta_{\alpha\beta} = \begin{cases} 1, & \alpha = \beta, \\ 0, & \alpha \neq \beta. \end{cases}$

Використовуючи цей символ, елементи одиничної матриці позначаються так:

$$a_{ij} = \delta_{ij}, \quad 1 \leqslant i, j \leqslant n.$$

Завдання 5.1. *Довести, що $\mathrm{L}_{I_n} = \mathbf{1}_{\mathbb{K}^n}$, тобто для довільного $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ виконується $I_n x = x$.*

5.2.1 Властивості лінійних відображень

Нехай $f : V \to W$ – лінійне відображення.

1. $f(0) = 0$.

   Дійсно, маємо $f(0) = f(0 \cdot v) = 0 \cdot f(v) = 0 \in W$.

2. $f(-v) = -f(v)$.

   Тому що $f(-v) = f(-1 \cdot v) = -1 \cdot f(v) = -f(v)$.

3. $f(v_1 - v_2) = f(v_1) - f(v_2)$.

   Це є наслідок з 2).

Як і у випадку відображень L_A вводиться поняття ядра та образу лінійного відображення.

Лемма 5.6. *Якщо $f : V \to W$ – лінійне відображення, то*

1. $\mathrm{K}\scriptstyle{\mathrm{ER}}$ $f \subset V$ *- підпростір в V;*
2. $\mathrm{I}\scriptstyle{\mathrm{M}}$ $f \subset W$ *- підпростір в W.*

Доведення таке саме, як для відображень L_A (потрібно L_A замінити на f). Проведіть це доведення самостійно.

Для скорочення записів введемо позначення:

▲ $\mathcal{L}(V, W)$ – множина лінійних відображень $f : V \to W$ ▼

Лемма 5.7. *Лінійне відображення $f \in \mathcal{L}(V, W)$ є мономорфним (ін'єктивним) тоді і лише тоді, коли* $\mathrm{Ker}\ f = 0$ *(це означає, що* $\mathrm{Ker}\ f = \{0\}$*, тобто* $\mathrm{Ker}\ f$ *є множина, яка складається з одного нуля).*

Доведення. Таке саме, як у випадку L_A. Якщо ядро $\mathrm{Ker}\ f \neq 0$, тобто $\exists v \neq 0$, $f(v) = 0$, f не ін'єктивне, тобто в 0 переходять два елементи.

Навпаки, якщо f – не ін'єктивне, то існують $v_1 \neq v_2$, для яких $f(v_1) = f(v_2) \iff f(v_1) - f(v_2) = 0 \iff f(v_1 - v_2) = 0$ і $v_1 - v_2 \neq 0 \implies v_1 - v_2 \in \mathrm{Ker}\ f \implies \mathrm{Ker}\ f \neq 0$. □

Наслідок 5.2. *Відображення $f \in \mathcal{L}(V, W)$ бієктивне тоді і лише тоді, коли або 1) f сюр'єктивне та ін'єктивне (за означенням); або 2)* $\mathrm{Ker}\ f = 0$, $\mathrm{Im}\, f = W$.

Має місце таке твердження.

Твердження 5.1. *Якщо $f \in \mathcal{L}(V, W)$ – лінійне і бієктивне, то обернене відображення $f^{-1} : W \to V$ теж буде лінійним (і, звичайно, бієктивним).*

Доведення. Як означається $f^{-1}(w)$? Вираз $f^{-1}(w) = v$ еквівалентний $f(v) = w$. Такий елемент $v \in V$ існує і єдиний, внаслідок бієктивності. Отже, нам необхідно перевірити такі властивості: $f^{-1}(w_1 + w_2) = f^{-1}(w_1) + f^{-1}(w_2)$ і $f^{-1}(\lambda w) = \lambda f^{-1}(w)$.

Два елементи v_1, $v_2 \in V$ рівні, $v_1 = v_2$ тоді і лише тоді, коли рівні їх образи $f(v_1) = f(v_2)$ (наслідок взаємної однозначності). Розглянемо $f^{-1}(w_1 + w_2) \in W$ і $f^{-1}(w_1) + f^{-1}(w_2) \in V$. Подіємо f: $f(f^{-1}(w_1 + w_2)) = w_1 + w_2$ (за означенням $+f^{-1}$). З лінійності f та з визначення оберненого відображення випливає, що $f(f^{-1}(w_1) + f^{-1}(w_2)) = f(f^{-1}(w_1)) + f(f^{-1}(w_2)) = w_1 + w_2$. під дією f обидва елементи стають рівними $w_1 + w_2$, отже вони рівні між собою, звідси $f^{-1}(w_1 + w_2) = f^{-1(w_1)} + f^{-1}(w_2)$.

Аналогічно доводиться твердження $f^{-1}(\lambda w) = \lambda f^{-1}(w)$. діємо відображенням f на праву і ліву частини. $f(f^{-1}(\lambda w)) \overset{df}{=} \lambda w$, $f(\lambda f^{-1}(w)) = (f$ - лін.$)\ \lambda\ f f^{-1}(w)) = \lambda w$. Елементи стали рівними, отже, вони були рівними до цього, тобто виконується рівність: $f^{-1}(\lambda w) = \lambda f^{-1}(w)$. □

Бієктивне лінійне відображення $f \in \mathcal{L}(V, W)$ ще називається *ізоморфізмом*. Якщо f – ізоморфізм, то, за доведеним, $f^{-1} \in \mathcal{L}(W, V)$ – теж ізоморфізм. В цьому випадку простори V та W називають *ізоморфними*. Отже, $V \sim W$, якщо існує ізоморфізм $f : V \to W$.

Твердження 5.2. *Композиція (суперпозиція) лінійних відображень є лінійним відображенням.*

Доведення. Нам потрібно довести, що для довільних $f \in \mathcal{L}(V, W)$ і $g \in \mathcal{L}(W, U)$ виконується: $g \circ f \in \mathcal{L}(V, U)$: $V \xrightarrow{f} W \xrightarrow{g} U$. Перевіримо, що для довільних v_1, $v_2 \in V$ виконується $(g \circ f)(v_1 + v_2) = (g \circ f)(v_1)+ (g \circ f)(v_2)$ і для $\lambda \in \mathbb{K}$, $v \in V$ виконується $(g \circ f)(\lambda v) = \lambda(g \circ f)(v)$. Розглянемо $(g \circ f)(v_1 + v_2) = g(f(v_1 + v_2)) \overset{f-\text{лін.}}{=}$ $g(f(v_1) + f(v_2)) \overset{g-\text{лін.}}{=} g(f(v_1))+ g(f(v_2)) \overset{\text{означ. комп.}}{=} (g \circ f)(v_1)+ (g \circ f)(v_2)$.

Перша рівність доведена, друга доводиться аналогічно. Залишається як вправа. □

Наслідок 5.3. *Якщо $f \in \mathcal{L}(V, W)$ і $g \in \mathcal{L}(W, U)$ – ізоморфізми, то $g \circ f \in \mathcal{L}(V, U)$ – теж ізоморфізм (векторних просторів).*

Тобто добуток бієкцій – знову бієкція.

Замість $g \circ f$ писатимемо gf, пропускаючи знак композиції.

Лемма-означення 5.1. *Сумою лінійних відображень $f, g \in \mathcal{L}(V, W)$ називається відображення $f + g \in \mathcal{L}(V, W)$, задане формулою*

$$\forall v \in V \quad (f + g)(v) \overset{df}{=} f(v) + g(v);$$

для $\alpha \in \mathbb{K}$, $f \in \mathcal{L}(V, W)$ добуток α на f визначається:

$$\forall v \in V \quad (\lambda f)(v) \overset{df}{=} \lambda(f(v)).$$

Перевіримо, що $(f + g)$ та (λf) є знов лінійними.

I. $(f + g)$ – лінійне відображення.

1. *Нехай $v_1, v_2 \in V$, $(f+g)(v_1+v_2) \overset{df}{=} f(v_1+v_2)+ g(v_1+v_2) \overset{f,g-\text{лін}}{=} (f(v_1)+f(v_2))+ (g(v_1) + g(v_2)) = f(v_1)+ g(v_1)+ f(v_2) + g(v_2) \overset{df}{=} (f + g)(v_1)+ (f + g)(v_2)$. Отже, $(f + g)(v_1 + v_2) = (f + g)(v_1)+ (f + g)(v_2)$.*

2. *Для $\lambda \in \mathbb{K}$ і $v \in V$ $(f+g)(\lambda v) \overset{df}{=} f(\lambda v)+ g(\lambda v) \overset{f,g-\text{лін}}{=} \lambda f(v)+ \lambda g(v) = \lambda(f(v) + g(v)) \overset{df}{=} \lambda((f + g)(v))$. Отже, $(f + g)(\lambda v) = \lambda((f + g)(v))$.*

Доведені властивості означають, що $f + g$ – лінійне відображення.

II. Аналогічно перевіримо умову:

$$f \in \mathcal{L}(V, W), \quad \lambda \in \mathbb{K} \implies \lambda f \in \mathcal{L}(V, W),$$

тобто λf – лінійне відображення. Для перевірки обчислимо

1. *Для* $v_1, v_2 \in V$, $(\lambda f)(v_1+v_2) \overset{df}{=} \lambda(f(v_1+v_2)) \overset{f-\text{лін.}}{=} \lambda(f(v_1)+f(v_2)) = \lambda(f(v_1))+\lambda(f(v_2)) \overset{df}{=} (\lambda f)(v_1)+ (\lambda f)(v_2)$. *Отже,* $(\lambda f)(v_1+v_2) = (\lambda f)(v_1)+ (\lambda f)(v_2)$.

2. *Для* $\lambda, \lambda' \in \mathbb{K}$ *і* $v \in V$ $(\lambda f)(\lambda' v) \overset{df}{=} \lambda f(\lambda' v)) \overset{f-\text{лін.}}{=} \lambda(\lambda' f(v)) = (\lambda\lambda')f(v) = (\lambda'\lambda)f(v) = \lambda'(\lambda f(v)) \overset{df}{=} \lambda'((\lambda f)(v))$. *Отже,* $(\lambda f)(\lambda' v) = \lambda'(\lambda f)(v))$.

Ми довели, що означені вище відображення $f+g$, $\lambda f : V \to W$ *знову будуть лінійними.*

Позначення. Нехай $f \in \mathcal{L}(V, W)$. Через $(-f)$ позначимо відображення

$$(-f) : \begin{cases} V & \longrightarrow & W \\ v & \mapsto & -f(v) \quad \forall v \in V. \end{cases}$$

Тоді $(-f) \in \mathcal{L}(V, W)$ тому, що $(-f) = (-1)f$. Дійсно, $(-f)(v) = -f(v)$. З іншого боку, $((-1)f)(v) = (-1)f(v) = -f(v)$.

Відображення 0 $(= 0_{V,W})$: $V \to W$ таке, що $\forall v \in V$ $0_{V,W}(v) = 0 \in W$, є лінійним.

Твердження 5.3. *Множина* $\mathcal{L}(V, W)$ *для довільних векторних просторів* V, W *над полем* $\mathbb{K}$ *утворює векторний простір над полем* $\mathbb{K}$ *відносно операцій* " + ", " · ", *визначених вище, тобто*

- $f, g \in \mathcal{L}(V, W) \implies f+g \in \mathcal{L}(V, W)$.
- $\lambda \in \mathbb{K}$, $f \in \mathcal{L}(V, W) \implies \lambda f \in \mathcal{L}(V, W)$.

Нулем в просторі $\mathcal{L}(V, W)$ *буде нульове відображення*

- $0 : V \to W :$ $\forall v \in V$ $0(v) = 0 \in W$.

Протилежне відображення до $f \in \mathcal{L}(V, W)$ *визначене як:* $(-f)$*:*

- $\forall\ v \in V$ $(-f)(v) = -f(v) \in W$.

Доведення. Щоб перевірити, що $\mathcal{L}(V, W)$ є векторним простором над $\mathbb{K}$, необхідно ввести операції додавання та множення на скаляр:

$$" + " : \begin{cases} \mathcal{L}(V, W) \times \mathcal{L}(V, W) & \longrightarrow & \mathcal{L}(V, W) \\ (f, g) & \mapsto & f+g \end{cases}$$

$$" \cdot " : \begin{cases} \mathbb{K} \times \mathcal{L}(V, W) & \longrightarrow & \mathcal{L}(V, W) \\ (\lambda, f) & \mapsto & \lambda f \end{cases}$$

Такі операції вже введені. Потрібно перевірити, що ці операції задовольняють аксіомам векторного простору 1–7.

В аксіомах векторного простору (коли ми будемо їх перевіряти для $\mathcal{L}(V, W)$) в правих і лівих частинах рівностей будуть знаходитись лінійні відображення (елементи $\mathcal{L}(V, W)$).

Два відображення, як відомо, рівні, коли вони однаково діють на однакові елементи. Тому для перевірки аксіом векторного простору для $\mathcal{L}(V, W)$ розглянемо довільний вектор $v \in V$ і застосуємо до нього відображення з правої і лівої частин відповідної аксіоми. Як результат одержимо деяку рівність в просторі W, яка виконується тому, що вона відповідає одночленній аксіомі векторного простору для W.

Виконаємо записану програму перевірки для конкретних аксіом:

1. *Аксіома асоціативності додавання.* Для $f, g, h \in \mathcal{L}(V, W)$ треба перевірити $(f + g) + h = f + (g + h)$.

   Застосуємо до $v \in V$ праву і ліву частини:

   $((f + g) + h)(v) = (f + g)(v) + h(v) = (f(v) + g(v) + h(v))$.

   $(f + (g + h))(v) = f(v) + (g + h)(v) = f(v) + g(v) + h(v)$.

   Одержані вектори рівні тому, що додавання векторів асоціативне.

2. *Аксіома нуля.* Для $0 \in \mathcal{L}(V, W)$, $\forall f \in \mathcal{L}(V, W)$ мають виконуватися рівності $0 + f = f + 0 = f$.

   Для довільного $v \in V$ маємо

   $(0 + f)(v) = 0(v) + f(v) = 0 + f(v) = f(v) = f(v) + 0 = f(v) + 0(v) = (f + 0)(v)$

   Тут ми використали аксіому нуля $0 + f(v) = f(v) + 0 = f(v)$ в просторі W.

3. *Аксіома протилежного елемента.* $\forall f \in \mathcal{L}(V, W) \quad \exists(-f) \in \mathcal{L}(V, W)$ таке, що виконується рівність: $f + (-f) = (-f) + f = 0$.

   Для довільного $v \in V$ маємо

   $f(v) + (-f)(v) = (-f)(v) + f(v) = 0(v) \quad \Longleftrightarrow \quad f(v) - f(v) = -f(v) + f(v) = 0.$

   Але остання рівність вірна в W.

4. *Аксіома комутативності додавання.* Для довільних $f, g \in \mathcal{L}(V, W) \quad f + g = g + f$, оскільки $f(v) + g(v) = g(v) + f(v)$ виконується в просторі W для довільного $v \in V$.

5. *Аксіома асоціативності множення на скаляри.* Для довільних $\lambda,\ \mu \in \mathbb{K}$ та $f \in \mathcal{L}(V, W)$ повинно виконуватися: $\lambda(\mu f) = (\lambda\mu)f$. Дійсно, $\forall\ v \in V$ маємо:

   $[\lambda(\mu f)](v) = ((\lambda\mu)f)(v) \iff \lambda((\mu f)(v)) = (\lambda\mu)f(v) \iff \lambda(\mu f(v)) = (\lambda\mu)f(v).$

   Це випливає з асоціативності множення векторів на скаляри в просторі W.

6. *Аксіома одиниці.* $\forall\ f \in \mathcal{L}(V, W) \quad 1 \cdot f = f$, де 1 – одиниця в $\mathbb{K}$.

   Але $(1 \cdot f)(v) = 1 \cdot f(v) = f(v)$ виконується в W.

7. *Аксіоми дистрибутивності.* Для довільних λ, $\mu \in \mathbb{K}$ та $f, g \in \mathcal{L}(V, W)$ треба перевірити, що $(\lambda+\mu)f = \lambda f + \mu f$ або $((\lambda+\mu)f)(v) = (\lambda f + \mu f)(v)\ \forall\ v \in V$.

   $((\lambda+\mu)f)(v) = (\lambda+\mu)f(v) = (\lambda f)(v) + (\mu f)(v) = \lambda \cdot f(v) + \mu \cdot f(v) = (\lambda f + \mu f)(v)$.

   Для доведення другої умови дистрибутивності $\lambda(f+g) = \lambda f + \lambda g$ застосуємо умову дистрибутивності в просторі W.

   $\forall v \in V\ \ (\lambda(f+g))(v) = (\lambda f + \lambda g)(v) \iff \lambda((f+g)(v)) = (\lambda f)(v) + (\lambda g)(v) \iff \lambda(f(v)+g(v)) = \lambda f(v) + \lambda g(v)$.

Отже, всі аксіоми векторного простору в $\mathcal{L}(V, W)$ виконуються. □

Наступна лема визначає деякі властивості композиції відображень.

Лемма 5.8. *1. Для довільних $f, g \in \mathcal{L}(V, W)$ та $h \in \mathcal{L}(W, U)$ має місце рівність $h(f+g) = hf + hg$, тобто* $V \xrightarrow{f+g} W \xrightarrow{h} U \iff V \xrightarrow{hf+hg} U$.

2. Для довільних $f \in \mathcal{L}(V, W)$, $h \in \mathcal{L}(W, U)$, $\lambda \in \mathbb{K}$ виконується:

$$\lambda(hf) = (\lambda h)f = h(\lambda f).$$

Доведення.

1. Розглянемо довільний $v \in V$ і перевіримо рівності на цьому елементі.

   $(h(f+g))(v) = h((f+g)(v)) = h(f(v)+g(v)) = h(f(v)) + h(g(v)) = (hf)(v) + (hg)(v)$ – перша рівність доведена.

2. Розглянемо $v \in V$, $\lambda \in \mathbb{K}$.

   $(\lambda(hf))(v) = \lambda((hf)(v)) = \lambda(h(f(v))) = (\lambda h)(f(v)) = ((\lambda h)f)(v)$.

   Тим самим ми довели $\lambda(hf) = (\lambda h)f$. З іншого боку,

   $(\lambda(hf))(v) = \lambda((hf)(v)) = \lambda(h(f(v)) = h(\lambda f(v)) = h(\lambda f(v)) = (h(\lambda f))(v)$.

   Ми довели: $\lambda(hf) = h(\lambda f)$.

□

Зауваження. Рівність $(g+h)f = gf + hf$ для $f \in \mathcal{L}(V, W)$ та $g, h \in \mathcal{L}(W, U)$ перевіряється аналогічно.

5.3 Базис та розмірність

Нехай $\mathcal{V} = \{v_i\}_{i \in I}$ є система векторів з простору V над полем $\mathbb{K}$. Система $\mathcal{V}$ називається *лінійно незалежною*, якщо для довільної лінійної комбінації $\sum\limits_{i \in I} \lambda_i v_i$, $\lambda_i \in \mathbb{K}$, з рівності $\sum\limits_{i \in I} \lambda_i v_i = 0$ випливає, що всі $\lambda_i = 0$, $i \in I$. Система $\mathcal{V}$ називається лінійно залежною в протилежному випадку, тобто, якщо існує деяка нетривіальна лінійна комбінація $\sum\limits_{i \in I} \lambda_i v_i = 0$ (де лише скінченна кількість коефіцієнтів відмінна від нуля).

Приклад 5.2. *Для простору $V = \mathbb{K}^n$, $n \geqslant 1$, розглянемо систему векторів $\{e_1, \ldots, e_n\}$,*

де $e_i = i \begin{pmatrix} 0 \\ \ldots \\ 1 \\ \ldots \\ 0 \end{pmatrix}$, $(e_i)_j = \delta_{ij}$, $j = 1, \ldots n$.

Ця система є лінійно незалежною. Дійсно:

$$\lambda_1 e_1 + \lambda_2 e_2 + \ldots + \lambda_n e_n = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \ldots \\ \lambda_n \end{pmatrix} = 0 \implies \lambda_1 = \lambda_2 = \ldots = \lambda_n = 0.$$

Нехай $J \subset I$. Система $\mathcal{V}_J = \{v_i\}_{i \in J}$ називається *підсистемою* системи $\mathcal{V}$. Згідно з цим означенням, $\mathcal{V} = \mathcal{V}_I$.

Якщо система $\mathcal{V}$ – лінійно незалежна, то довільна її підсистема теж лінійно незалежна.

Означення 5.3. *Система векторів $\mathcal{V} = \{v_i\}_{i \in I}$ називається базисом простору V, якщо для кожного $v \in V$ існує, причому єдиний набір коефіцієнтів $\lambda_i \in \mathbb{K}$, $i \in I$ такий, що v є лінійною комбінацією $v = \sum\limits_{i \in I} \lambda_i v_i$.*

Зауваження 5.2. *Звичайно, існує тільки скінченна кількість λ_i, відмінних від* 0.

Приклад 5.3. *В просторі $V = \mathbb{K}^n$ система векторів $\mathcal{E} = \{e_1, e_2, \ldots, e_n\}$, де*

$e_1 = \begin{pmatrix} 1 \\ 0 \\ \ldots \\ 0 \end{pmatrix}$, $e_2 = \begin{pmatrix} 0 \\ 1 \\ \ldots \\ 0 \end{pmatrix}$, …, $e_n = \begin{pmatrix} 0 \\ 0 \\ \ldots \\ 1 \end{pmatrix}$, *утворює базис простору $V = \mathbb{K}^n$, який називається* стандартний базисом *в $\mathbb{K}^n$.*

Зауваження 5.3. *Базис у векторному просторі не єдиний* $(n > 1)$.

Приклад 5.4. *В $\mathbb{K}^2$ маємо базиси: стандартний і, наприклад,* $\left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}$. *Перевірити, що це базис в $\mathbb{K}^2$.*

Означення 5.4. *Два простори V, W називаються* ізоморфними*, якщо існує ізоморфізм $f : V \to W$.*

Зауваження-лема. Відношення "бути ізоморфними" є відношенням еквівалентності.

Доведення . Позначимо ізоморфізм просторів через символ $\simeq$. Отже, $V \simeq W$, якщо існує ізоморфізм $f : V \to W$.

Перевіримо властивості еквівалентності.

1. *Рефлексивність.* $V \simeq V$ тому, що відображення $1_V : V \to V$, $v \mapsto v$ є лінійним і бієктивним.

2. *Симетричність.* Якщо $V \simeq W$, то $W \simeq V$.

Якщо $f : V \to W$ – деякий ізоморфізм, то обернене відображення $f^{-1} : W \to V$ теж ізоморфізм. Отже, $W \simeq V$.

3. *Транзитивність.* Якщо $V \simeq W$ і $W \simeq U$, то $V \simeq U$.

Чому? Нехай $f : V \to W$, $g : W \to U$ – ізоморфізми. Розглянемо композицію відображень $gf : V \to U$ $(V \xrightarrow{f} W \xrightarrow{g} U)$. Композиція ізоморфізмів – знову ізоморфізм тому, що gf – лінійне і бієктивне. Звідси випливає, що $V \simeq U$. $\square$

5.3.1 Допоміжні твердження

Лемма 5.9. *Нехай $f : V \to W$ – деякий ізоморфізм, $\{v_1, \ldots, v_k\}$ – деяка система векторів в V. Тоді ця система буде лінійно залежною тоді і тільки тоді, коли система $\{f(v_1), \ldots, f(v_k)\}$ буде лінійно залежною в W.*

Доведення. Якщо $\sum\limits_{i=1}^{k} \lambda_i v_i = 0$ – лінійна залежність між векторами $v_1, \ldots, v_k$, то, застосовуючи f, одержуємо $0 = f(0) = f(\sum\limits_{i=1}^{k} \lambda_i v_i) \overset{\text{лін.}}{=} \sum\limits_{i=1}^{k} \lambda_i f(v_i)$, що дає лінійну залежність між векторами $f(v_1)$, …, $f(v_k)$. Застосовуючи $f^{-1} : W \to V$, одержуємо

$$0 = f^{-1}\big(\sum_{i=1}^{k} \lambda_i f(v_i)\big) \; = \; \sum_{i=1}^{k} \lambda_i f^{-1} f(v_i) \overset{f^{-1}f = 1_V}{=} \sum_{i=1}^{k} \lambda_i v_i,$$

що дає лінійну залежність між векторами $v_1, \ldots, v_k$ в V. $\square$

Лемма 5.10 (про короткі вектори)**.** *Припустимо, що в просторі $\mathbb{K}^m$ є система векторів $\{a_1, \ldots, a_n\}$, причому $n > m$. Тоді ця система є лінійно залежною.*

Доведення. Запишемо вектори $a_1, \ldots, a_n$ як стовпчики у вигляді

$$a_1 = \begin{pmatrix} a_{11} \\ a_{21} \\ \ldots \\ a_{m1} \end{pmatrix}, \, a_2 = \begin{pmatrix} a_{12} \\ a_{22} \\ \ldots \\ a_{m2} \end{pmatrix}, \ldots, \, a_n = \begin{pmatrix} a_{1n} \\ a_{2n} \\ \ldots \\ a_{mn} \end{pmatrix},$$

Тоді існування лінійної залежності між $a_1, \ldots, a_n$ еквівалентно тому, що система рівнянь

$$x_1 a_1 + x_2 a_2 + \ldots + x_n a_n = 0 \qquad (*)$$

має ненульовий розв'язок в $\mathbb{K}^m$, тобто існують $x_1, \ldots, x_n \in \mathbb{K}$, не всі рівні 0, такі, що виконується $(*)$.

$(*)$ – це рівняння відносно векторів. Але очевидно, що ця векторна рівність є системою лінійних рівнянь відносно x_1, …, x_n. більше того, ця система однорідна і матриця її має вигляд:

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

В цій матриці кількість рядків менша кількості стовпчиків (або кількість рівнянь менше кількості невідомих). Ми довели (наслідок з методу Гауса), що така система має ненульовий розв'язок. Отже, вектори $a_1, \dots, a_m$ є лінійно залежні. □

Лемма 5.11 (про зайвий вектор)**.** *Припустимо, що система векторів* $\{v_1, \dots, v_k\}$ *з простору* V *є лінійно незалежною, а* $v \in V$ *– такий вектор, що система* $\{v_1, \dots, v_k, v\}$ *є лінійно залежною. Тоді вектор* v *єдиним способом лінійно виражається через* $\{v_1, \dots, v_k\}$*, тобто* $v = \sum\limits_{i=1}^{k} \mu_i v_i$*,* $\mu_i \in \mathbb{K}$*.*

Доведення. Система $\{v_1, \dots, v_k, v\}$ – лінійно залежна. Запишемо, цю лінійну залежність: $\sum\limits_{i=1}^{k} \lambda_i v_i + \lambda v = 0$, де λ_1, …, λ_k, $\lambda \in \mathbb{K}$ і не всі вони дорівнюють нулю. Тоді $\lambda \neq 0$. Чому? Припустимо, $\lambda = 0$. Звідси ми одержуємо, що $\sum\limits_{i=1}^{k} \lambda_i v_i = 0$ і не всі λ_i рівні 0. Але це дає лінійну залежність між v_1, …, v_k, що суперечить умові. Отже, $\lambda \neq 0$. Перепишемо рівність:

$$\lambda v \; = \; -\sum_{i=1}^{k} \lambda_i v_i \; = \; \sum_{i=1}^{k} (-\lambda_i) v_i.$$

Оскільки $\lambda \neq 0$, то існує $\lambda^{-1} \implies v = \sum\limits_{i=1}^{k} (-\lambda_i \cdot \lambda^{-1}) v_i$. Отже, ми представили v як лінійну комбінацію v_i.

Доведемо єдиність такого зображення. Припустимо, $v = \sum\limits_{i=1}^{k} \mu_i v_i$ і $v = \sum\limits_{i=1}^{k} \mu'_i v_i$. віднімемо ці рівності. Звідси одержимо $0 = v - v = \sum\limits_{i=1}^{k} (\mu_i - \mu'_i) v_i$. Але, оскільки v_1, …, v_k – лінійно незалежні, то $\mu_1 - \mu'_1 = 0$, …, $\mu_k - \mu'_k = 0$, тобто $\mu_1 = \mu'_1$, …, $\mu_k = \mu'_k$. □

5.3.2 Приклади-наслідки

1. Система векторів, в якій є два рівних вектори, лінійно залежна.

   Розглянемо систему $\{v_1, \dots, \underbrace{v}_{i}, \dots, \underbrace{v}_{j}, \dots, v_k\}$ $(v = v_i = v_j)$. Задамо $\lambda_1 = 0$, $\lambda_i = 1$, $\lambda_j = -1$, …, $\lambda_k = 0$.

2. Система векторів, яка містить нульовий вектор, лінійно залежна.

   Для системи $\{v_1, \ldots, \underbrace{0}_{i}, \ldots, v_k\}$ покладемо: $\lambda_1 = 0$, ..., $\lambda_i = 1$, ..., $\lambda_k = 0$. Тоді $\sum\limits_{i=1}^{k} v_i \lambda_i = 0$ і $\lambda_i \neq 0$.

3. Якщо в системі існують два пропорційні вектори, тобто для деяких i, j виконується $v_i = \lambda v_i$, то система лінійно залежна.

   Перевірити самостійно.

4. Система, що складається з одного вектора, є лінійно залежною лише в тому разі, якщо цей вектор нульовий.

5. Два вектори лінійно залежні, якщо вони пропорційні.

5.4 Векторні простори типу $\mathbb{K}^n$

Простір V називається скінченно вимірним, якщо в ньому існує скінченний базис $\mathcal{B} = \{v_1, \ldots, v_k\}$.

Лемма 5.12 (про координатизацію)**.** *Кожен базис $\mathcal{B} = \{v_1, \ldots, v_k\}$ визначає ізоморфізм $F_{\mathcal{B}} : V \to \mathbb{K}^n$.*

Навпаки, кожен ізоморфізм $F : V \to \mathbb{K}^n$ визначає деякий базис $\mathcal{B}$ (= $\mathcal{B}(F)$ – це означає залежність від F) такий, що $F = F_{\mathcal{B}}$.

Доведення. Припустимо, що дано базис $\mathcal{B} = \{v_1, \ldots, v_n\}$. Побудуємо $F_{\mathcal{B}} : V \to \mathbb{K}^n$. Відомо, що v однозначно зображується у вигляді $v = \lambda_1 v_1 + \ldots + \lambda_n v_n$. Тоді визначимо $F_{\mathcal{B}}(v) = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \ldots \\ \lambda_n \end{pmatrix} \in \mathbb{K}^n$. Таким чином, визначене відображення $F_{\mathcal{B}} : V \to \mathbb{K}^n$. Ясно, що це відображення бієктивне. Тому що по кожному набору $\lambda_1, \ldots, \lambda_n$ однозначно відновлюється вектор $v = \lambda_1 v_1 + \ldots + \lambda_n v_n$. Тобто $F_{\mathcal{B}}^{-1} \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \ldots \\ \lambda_n \end{pmatrix} = \sum\limits_{i=1}^{n} \lambda_i v_i$. Залишилось довести, що відображення $F_{\mathcal{B}}$ – лінійне, тобто перевірити, що $F_{\mathcal{B}}(v + v') = F_{\mathcal{B}}(v) + F_{\mathcal{B}}(v')$, $v, v' \in V$ і $F_{\mathcal{B}}(\lambda v) = \lambda F_{\mathcal{B}}(v)$, $\lambda \in \mathbb{K}$, $v \in V$. Припустимо, що $v = \sum\limits_{i=1}^{n} \lambda_i v_i$ – однозначний запис v, $v' = \sum\limits_{i=1}^{n} \lambda'_i v_i$ – однозначний запис v', де $\lambda_i \in \mathbb{K}$, $\lambda'_i \in \mathbb{K}$. Тоді як записується вектор $v + v'$? Цей запис однозначний, тому будь-який запис буде правильним. Можемо його знайти так: $v+v' = \sum\limits_{i=1}^{n}\lambda_i v_i + \sum\limits_{i=1}^{n}\lambda'_i v_i = \sum\limits_{i=1}^{n}(\lambda_i+\lambda'_i)v_i$. Отже, $v+v'$ ми представили у вигляді лінійної комбінації v_1, ..., v_n (зображення однозначне). Тепер, за означенням, $F_{\mathcal{B}}(v) = \begin{pmatrix} \lambda_1 \\ \dots \\ \lambda_n \end{pmatrix}$, $F_{\mathcal{B}}(v') = \begin{pmatrix} \lambda'_1 \\ \dots \\ \lambda'_n \end{pmatrix}$ і $F_{\mathcal{B}}(v+v') = \begin{pmatrix} \lambda_1+\lambda'_1 \\ \dots \\ \lambda_n+\lambda'_n \end{pmatrix}$. Тепер рівність $F_{\mathcal{B}}(v+v') = F_{\mathcal{B}}(v) + F_{\mathcal{B}}(v')$ випливає з правил додавання векторів.

Аналогічно перевіряється, що $F_{\mathcal{B}}(v) = \lambda F_{\mathcal{B}}(v)$. Коли $v = \sum\limits_{i=1}^{n}\lambda_i v_i$, то $\lambda v = \sum\limits_{i=1}^{n}(\lambda\lambda_i)v_i$. Звідси одержимо, що $F_{\mathcal{B}}(v) = \begin{pmatrix} \lambda_1 \\ \dots \\ \lambda_n \end{pmatrix}$, а $F_{\mathcal{B}}(\lambda v) = \begin{pmatrix} \lambda\lambda_1 \\ \dots \\ \lambda\lambda_n \end{pmatrix}$. Звідси одержуємо, що $F_{\mathcal{B}}(\lambda v) = \lambda F_{\mathcal{B}}(v)$ за правилом множення вектора на скаляр.

Отже, $F_{\mathcal{B}}$ – лінійне відображення і ізоморфізм.

Навпаки, припустимо, що заданий ізоморфізм $F : V \to \mathbb{K}^n$. Тоді $F' : \mathbb{K}^n \to V$ – теж ізоморфізм. Розглянемо в $\mathbb{K}^n$ стандартні базисні вектори $e_i = \begin{pmatrix} 0 \\ 1 \\ \dots \\ 0 \end{pmatrix}$, $i = 1, \dots, n$. Введемо в просторі V систему векторів $\{F^{-1}(e_1), \dots, F^{-1}(e_n)\}$, тобто запишемо $v_1 = F^{-1}(e_1)$, $v_2 = F^{-1}(e_2)$, ..., $v_n = F^{-1}(e_n)$. Тоді стверджується, що система векторів $v_1, \dots, v_n$ є базисом? Чому?

Розглянемо довільний вектор $v \in V$ і разом з ним $F(v) \in \mathbb{K}^n$. Маємо: $F(v) = \begin{pmatrix} \lambda_1 \\ \dots \\ \lambda_n \end{pmatrix} = \lambda_1 e_1 + \ldots + \lambda_n e_n$ – однозначний розклад. Застосуємо до цієї рівності ізоморфізм $F^{-1} : \mathbb{K}^n \to V$. Одержимо:

$$F^{-1}F(v) = v = F^{-1}(\sum_{i=1}^{n}\lambda_i e_i) = \sum_{i=1}^{n}\lambda_i F^{-1}(e_i) = \sum_{i=1}^{n}\lambda_i v_i.$$

Отже, довільний вектор $v \in V$ представляється як лінійна комбінація векторів $v_1, \dots, v_n$. Доведемо однозначність цього розкладу.

Припустимо, $v = \sum\limits_{i=1}^{n}\lambda_i v_i = \sum\limits_{i=1}^{n}\lambda'_i v_i$. Застосуємо $F : V \to \mathbb{K}^n$. Зауважимо, що $F(v_i) = F(F^{-1}(e_i)) = e_i$. Отже одержимо $F(v) = \sum\limits_{i=1}^{n}\lambda_i F(v_i) = \sum\limits_{i=1}^{n}\lambda'_i F(v_i)$, або ж $\sum\limits_{i=1}^{n}\lambda_i e_i = \sum\limits_{i=1}^{n}\lambda'_i e_i \implies \begin{pmatrix} \lambda_1 \\ \dots \\ \lambda_n \end{pmatrix} = \begin{pmatrix} \lambda'_1 \\ \dots \\ \lambda'_n \end{pmatrix} \implies$ розклад v однозначний. $\square$

Кажуть, що вибрати базис в просторі V – це прив'язати V до координатного простору $\mathbb{K}^n$.

Означення 5.5. *Якщо* $\mathcal{B} = \{v_1, \ldots, v_n\}$ *– базис в* V *і* $v = \sum\limits_{i=1}^{n} \lambda_i v_i$, *то числа* $\lambda_i \in \mathbb{K}$, $i = 1, \ldots, n$, *називаються* координатами *вектора* v *в базисі* $\mathcal{B}$. *Через* $[v]_{\mathcal{B}} = \begin{pmatrix} \lambda_1 \\ \ldots \\ \lambda_n \end{pmatrix}$ *позначається вектор-стовпчик координат.*

Лемма 5.13. *Нехай* V *– скінченно вимірний простір,* $\mathcal{B} = \{v_1, \ldots, v_n\}$ *– базис. Тоді довільне сімейство векторів* $a_1, \ldots, a_N \in V$, $N > n$, *лінійно залежне.*

Доведення. Нехай $F_{\mathcal{B}} = F : V \to \mathbb{K}^n$ – побудований ізоморфізм. Тоді $F(a_1), \ldots, F(a_N)$ лінійно залежні в $\mathbb{K}^n$, за лемою 5.10 про короткі вектори, оскільки $N > n$. Але, оскільки F – ізоморфізм, то вектори $a_1, \ldots, a_N$ – лінійно залежні $\iff$ $F(a_1), \ldots, F(a_N)$ – лінійно залежні. $\square$

Теорема 5.1 (про розмірність). *Припустимо, що простір* V *має базис* $\mathcal{B} = \{v_1, \ldots, v_n\}$. *Тоді довільний інший базис* $\mathcal{B}' = \{v_1', \ldots, v_m'\}$ *містить ту саму кількість векторів, тобто* $n = m$.

Доведення. Розглянемо ізоморфізм $F_{\mathcal{B}} : V \to \mathbb{K}^n$. Вектори $F_{\mathcal{B}}(v_1'), \ldots, F_{\mathcal{B}}(v_n')$ лінійно незалежні, тому що $v_1', \ldots, v_n'$ – лінійно незалежні. (Вектори базису лінійно незалежні тому, що нульовий вектор зображується єдиним способом як лінійна комбінація векторів базису $0 = 0 \cdot v_1' + \ldots + 0 \cdot v_m'$). Звідси випливає, що в $n-$вимірному просторі маємо m лінійно незалежних векторів $\implies$ $n \geqslant m$ (за лемою 5.10 про короткі вектори). Але $\mathcal{B}'$ – теж базис.

Аналогічно, розглядаючи ізоморфізм $F_{\mathcal{B}'} : V \to \mathbb{R}^m$, одержимо, що $F_{\mathcal{B}'}(v_1), \ldots, F_{\mathcal{B}'}(v_n)$ – лінійно незалежні $\implies$ $m \geqslant n$.

Звідси випливає, що $m = n$. $\square$

Отже, два довільних базиси містять однакову кількість векторів.

Означення 5.6. *Кількість векторів в базисі* $\mathcal{B}$ *не залежить від базису і називається* розмірністю *скінченно вимірного простору* V. *Розмірність простору позначається* $\operatorname{DIM}_{\mathbb{K}} V$. *Маємо*

$$n = \operatorname{DIM}_{\mathbb{K}} V \quad \iff \quad \exists \mathcal{B} = \{v_1, \ldots, v_n\} \ - \ \textit{базис.}$$

В просторі $\mathbb{R}^n$ множина $\mathcal{E} = \{e_1, \ldots, e_n\}$ є базисом $\implies$ $\operatorname{DIM}_{\mathbb{K}} \mathbb{K}^n = n$ (тому що стандартний базис містить рівно n елементів.) Отже, з цієї точки зору розмірність прямої дорівнює $\operatorname{DIM}_{\mathbb{R}} \mathbb{R}^1 = 1$, площини – $\operatorname{DIM}_{\mathbb{R}} \mathbb{R}^2 = 2$, простору $\operatorname{DIM}_{\mathbb{R}} \mathbb{R}^3 = 3$.

Домовленість 7. *Запис* $\{v_i\}_{i \in I}$ *еквівалентний запису* $\{v_1, \ldots, v_n\}$, *якщо* $I = \{1, \ldots, n\}$.

Лемма 5.14. *Нехай* V *– векторний простір над полем* $\mathbb{K}$ *(надалі опускатимемо це припущення), і* $\mathcal{V} = \{v_i\}_{i \in I}$ *– така система векторів, що лінійна оболонка її збігається з* V, *тобто* $(v_i)_{i \in I} = V$. *Тоді наступні твердження еквівалентні.*

1. Система $\mathcal{V} = \{v_i\}_{i \in I}$ утворює базис.

2. Система $\mathcal{V} = \{v_i\}_{i \in I}$ лінійно незалежна.

Доведення. 1) $\Longrightarrow$ 2). Припустимо, $\mathcal{V}$ – базис. Тоді нульовий вектор (як і усякий інший) можна єдиним способом зобразити у вигляді лінійної комбінації векторів базису: $v = \sum\limits_{i \in I} \lambda_i v_i$. Але такий спосіб ми знаємо: $\lambda_i = 0 \quad \forall i \in I$. Звідси випливає, що система $\mathcal{V}$ – лінійно незалежна.

2) $\Longrightarrow$ 1). Припустимо, ми маємо лінійно незалежну систему векторів $\mathcal{V} = \{v_i\}_{i \in I}$, причому $V = (v_i)_{i \in I}$ є лінійна оболонка системи $\mathcal{V}$. Доведемо, що для довільного $v \in V$ існує єдине зображення $v = \sum\limits_{i \in I} \lambda_i v_i$, $\lambda_i \in \mathbb{K}$. Таке зображення існує для довільного вектора $v \in V$ тому, що $v \in V = (v_i)_{i \in I}$ як лінійній оболонці. Треба довести єдиність. Припустимо, існують два різних зображення $v = \sum\limits_{i \in I} \lambda_i v_i$, $v = \sum\limits_{i \in I} \lambda'_i v_i$. Одержимо $0 = v - v = \sum\limits_{i \in I} (\lambda_i - \lambda'_i) v_i$. Але, оскільки система лінійно незалежна, звідси випливає, що $\lambda_i = \lambda'_i$ для довільного i. Отже, маємо протиріччя. $\square$

З точки зору однозначності розкладу достатньо перевірити, чи єдиним способом можна зобразити нуль.

Наслідок 5.4 (про розмірність). *Припустимо, що V – скінченно вимірний простір над $\mathbb{K}$ і W – інший простір над $\mathbb{K}$ такий, що $V \simeq W$. Тоді W – скінченно вимірний і $\operatorname{DIM}_{\mathbb{K}} W = \operatorname{DIM}_{\mathbb{K}} V$.*

Доведення. Нехай $f : V \to W$ – деякий ізоморфізм (тобто взаємно однозначне бієктивне лінійне відображення). Тоді має місце така допоміжна

Лемма 5.15. *Якщо $f : V \to W$ – ізоморфізм і $\mathcal{V} = \{v_1, \ldots, v_n\}$ – базис в V, то $f(\mathcal{V}) = \{f(v_1), \ldots, f(v_n)\}$ є базисом в просторі W.*

Доведення. Стверджується, що лінійна оболонка $(f(\mathcal{V}))$ системи $f(\mathcal{V})$ рівна простору W. Чому? Розглянемо довільний $w \in W$. Оскільки f – ізоморфізм, то f – епіморфізм (відображення "на") $\Longrightarrow \quad \exists v \in V : w = f(v)$. Оскільки $\mathcal{V}$ – базис в просторі V, то існують λ_1, …, $\lambda_n \in \mathbb{K}$ такі, що $v = \lambda_1 v_1 + \ldots + \lambda_n v_n$. Подіємо відображенням f на праву і ліву частини рівняння, одержимо $f(v) = f(\lambda_1 v_1 + \ldots + \lambda_n v_n) = \lambda_1 f(v_1) + \ldots + \lambda_n f(v_n) \Longrightarrow f(v) \in ((f(v))$, або, що те саме, $f(v)$ лінійно виражається через $f(v_1)$, …, $f(v_n)$. Залишилось довести, що вектори $f(v_1)$, …, $f(v_n)$ лінійно незалежні. За попередньою лемою, щоб довести, що $f(\mathcal{V})$ – базис, достатньо довести, що $f(\mathcal{V})$ – лінійно незалежна система. Припустимо протилежне. Розглянемо лінійну комбінацію $0 = \lambda_1 f(v_1) + \ldots + \lambda_n f(v_n) \overset{f-\text{лін.}}{\Longrightarrow} 0 = f(\lambda_1 v_1) + \ldots + f(\lambda_n v_n) \overset{f-\text{лін.}}{\Longrightarrow} 0 = f(\lambda_1 v_1 + \ldots + \lambda_n v_n) \overset{f-\text{ізом.}}{\Longrightarrow}$ існує обернене відображення $f^{-1} : W \to V$; застосуємо його, одержимо $f^{-1}(0) = f^{-1}(f(\lambda_1 v_1 + \ldots + \lambda_n v_n)) \overset{f^{-1}f = 1_V}{\Longrightarrow} \lambda_1 v_1 + \ldots + \lambda_n v_n$. Оскільки $\{v_1, \ldots, v_n\}$ –

базис, звідси одержимо, що $\lambda_1 = \lambda_2 = \ldots = \lambda_n = 0$. Тобто, ця залежність тривіальна, а отже, $f(v_1)$, ..., $f(v_n)$ – лінійно незалежні. $\square$

Доведемо наслідок. Розглянемо базис $\mathcal{B} = \{v_1, \ldots, v_n\} \subset V$. За означенням, розмірність $\text{DIM}_{\mathbb{K}} V = n$. Якщо $f : V \to W$ – ізоморфізм, то, як ми довели, $f(\mathcal{B}) = \{f(v_1), \ldots, f(v_n)\}$ є базисом в W. Але цей базис містить n векторів. Звідси одержуємо: $\text{DIM}_{\mathbb{K}} W = n$, отже $\text{DIM}_{\mathbb{K}} W = \text{DIM}_{\mathbb{K}} V$, і простір W є скінченно вимірним.

5.5 Ранг системи векторів

Пояснення: якщо є якесь поняття, яке може бути визначене кількома еквівалентними способами, то всі вони наводяться як означення, після чого доводиться їх еквівалентність.

Лемма-означення 5.2. *Нехай V – це скінченно вимірний простір над $\mathbb{K}$ і $\mathcal{V} = \{v_i\}_{i\in I}$ – система векторів в V. Рангом системи векторів $r = \text{RANK}\mathcal{V}$ називаеться одне з двох (рівних між собою) чисел:*

1. *розмірність лінійної оболонки системи $\mathcal{V}$;*
2. *кількість елементів в максимальній лінійно незалежній підсистемі $\{v_{i_1}, \ldots, v_{i_r}\} \subset \{v_i\}_{i\in I}$.*

Перше означення – більш теоретичного плану, друге – практичне.

Зауваження 5.4. *З коректності цього означення випливає, що дві максимальні лінійно незалежні підсистеми в системі $\mathcal{V}$ мають одну і ту саму кількість елементів, рівну розмірності над полем $\mathbb{K}$ лінійної оболонки $\text{DIM}_{\mathbb{K}}(v_i)_{i\in I}$.*

Доведення. Доведення коректності означення. Нехай $\{v_{i_1}, \ldots, v_{i_r}\}$ – це максимальна лінійно незалежна підсистема в системі $\mathcal{V}$. Зауважимо, що $r \leqslant \text{DIM}_{\mathbb{K}} V$ за наслідком. Тоді розмірність лінійної оболонки $(v_{i_1}, \ldots, v_{i_r})$ дорівнює r, тому що v_{i_1}, ..., v_{i_r} – лінійно незалежні, а отже утворюють базис в просторі V (в своїй лінійній оболонці).

З іншого боку, довільний $v_i \in \mathcal{V}$ виражається через $\{v_{i_1}, \ldots, v_{i_r}\}$. Чому? Тому що v_{i_1}, ..., v_{i_r}, v_i вже лінійно залежні $\implies v_i \in (v_{i_1}, \ldots, v_{i_r})$. Але оскільки v_i виражається через $\{v_{i_1}, \ldots, v_{i_r}\}$ (лема про лінійні оболонки систем векторів), $(v_i)_{i\in I} \subset (v_{i_1}, \ldots, v_{i_r})$. З іншого боку, очевидне протилежне включення, а саме $(v_{i_1}, \ldots, v_{i_r}) \subset (v_i)_{i\in I} \implies$ Звідси одержуємо як наслідок

$$(v_{i_1}, \ldots, v_{i_r}) \subset (v_i)_{i\in I} \subset (v_{i_1}, \ldots, v_{i_r}) \implies (v_{i_1}, \ldots, v_{i_r}) = (v_i)_{i\in I}$$

$$\implies r = \text{DIM}_{\mathbb{K}}(v_{i_1}, \ldots, v_{i_r}) = \text{DIM}_{\mathbb{K}}(v_i)_{i\in I}.$$

$\square$

Означення 5.7. *Нехай A – $m \times n-$матриця над полем $\mathbb{K}$, $A = (a_1|a_2|\ldots|a_n)$ – зображення у вигляді системи стовпчиків. Тоді ранг системи векторів $\{a_1, a_2, \ldots, a_n\}$ називається* **стовпчиковим** *рангом матриці A і позначається RANK_cL_A.*

Лемма 5.16 (про ранг та розмірність образу). *Нехай A – деяка $m \times n-$матриця над полем $\mathbb{K}$, $\mathrm{L}_A : V \to W$ – відповідне їй лінійне відображення. Тоді якщо матриця A розбита на стовпчики $A = (a_1|a_2|\ldots|a_n)$, то*

1) образ лінійного відображення є лінійною оболонкою векторів $a_1, a_2, \ldots, a_n$, тобто $\textsc{Im}\ \mathrm{L}_A = (a_1, a_2, \ldots, a_n)$;

2) $\textsc{dim}_{\mathbb{K}}\textsc{Im}\ \mathrm{L}_A = \textsc{rank}_c\mathrm{L}_A$.

Доведення. Розглянемо простір $\textsc{Im}\ \mathrm{L}_A \subset W$ ($W = \mathbb{K}^m$?) Тоді $w \in \textsc{Im}\ \mathrm{L}_A \iff \exists\ l \in \mathbb{K}^n : w = \mathrm{L}_A(l)$. Розглянемо l.

$$\exists l = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \ldots \\ \lambda_n \end{pmatrix}, \lambda_i \in \mathbb{K} \quad w = \mathrm{L}_A(l) = A \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \ldots \\ \lambda_n \end{pmatrix}, \iff$$

$$\exists l = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \ldots \\ \lambda_n \end{pmatrix}, \lambda_i \in \mathbb{K} \quad w = (a_1|a_2|\ldots|a_n) \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \ldots \\ \lambda_n \end{pmatrix}, \iff$$

$$\exists l = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \ldots \\ \lambda_n \end{pmatrix}, \lambda_i \in \mathbb{K} \quad w = \lambda_1 a_1 + \lambda_2 a_2 + \ldots + \lambda_n a_n),$$

тобто $w \in (a_1, \ldots, a_n)$ – належить лінійній оболонці. Отже, $\textsc{Im}\ \mathrm{L}_A = (a_1, a_2, \ldots, a_n)$.

Доведемо 2). $\textsc{dim}_{\mathbb{K}}\textsc{Im}\ \mathrm{L}_A \overset{\text{за доведеним}}{=} \textsc{dim}_{\mathbb{K}}(a_1, \ldots, a_n) \overset{\text{за означ.ст.рангу}}{=} \textsc{rank}_c A$. $\square$

фізичний зміст стовпчикового рангу. Ранг – це кількість правих частин, при яких система має розв'язок.

Лемма 5.17 (про монотонність розмірності). *Нехай V – скінченно вимірний простір над полем $\mathbb{K}$ і $W \subset V$ – підпростір. Тоді*

1) $\textsc{dim}_{\mathbb{K}}W \leqslant \textsc{dim}_{\mathbb{K}}V$;

2) якщо $\textsc{dim}_{\mathbb{K}}W = \textsc{dim}_{\mathbb{K}}V \implies W = V$.

Доведення. Позначимо $n = \textsc{dim}_{\mathbb{K}}V$. Нагадаємо, що довільні $n+1$ вектори v_1, v_2, …, $v_{n+1} \in V$ лінійно залежні. Припустимо, що $\textsc{dim}_{\mathbb{K}}W > \textsc{dim}_{\mathbb{K}}V = n$.

Розглянемо деякий базис в W: $\mathcal{B} = \{w_1, \ldots, w_N\}$, де $N = \textsc{dim}_{\mathbb{K}}W > n$. Розглянемо в базисі $\mathcal{B}$ перші $n+1$ векторів w_1, w_2, …, w_{n+1}. Вони лінійно незалежні, тому що $\mathcal{B}$ – базис. З іншого боку, w_1, w_2, …, $w_{n+1} \subset V \implies w_1, w_2$, …, w_{n+1} – лінійно залежні. Причина протиріччя полягає в припущенні, що розмірність більша.

Пункт 1) доведено.

Доведемо пункт 2. Використаємо наслідок з леми про зайвий вектор.

Припустимо, що для деякого $W \subset V$ $\textsc{dim}_{\mathbb{K}}W = \textsc{dim}_{\mathbb{K}}W = n$. Розглянемо базис W $\mathcal{B} = \{w_1, \ldots w_n\}$. Нехай $v \in V$ – довільний вектор.

Система векторів $\{w_1, \dots w_n\}$ лінійно незалежна (як базис в W). Система векторів $\{w_1, \dots w_n, v\}$ – лінійно залежна як система, що складається з $n+1$ вектора в $n-$вимірному просторі V. Застосовуємо лему про зайвий вектор. Маємо $v = \lambda_1 w_1 + \dots + \lambda_n w_n$, λ_1, …, $\lambda_n \in \mathbb{K}$, отже $v \in V$.

Отже, одержали твердження

$$\{v \in V \implies v \in W \iff V \subset W\}.$$

Навпаки, аналогічно, $W \subset V$. Тоді $V \subset W \subset V \iff V = W$. $\square$

Теорема 5.2 (про розмірність ядра та образу)**.** *Нехай $f : V \to W$ – лінійне відображення векторних просторів над полем $\mathbb{K}$ і припустимо, що простір V – скінченно вимірний. Тоді* $\textsc{Ker}\ f \subset V$*,* $\textsc{Im}\ f \subset W$ *– скінченно вимірні і*

$$\textsc{dim}_{\mathbb{K}}\textsc{Ker}\ f \ + \ \textsc{dim}_{\mathbb{K}}\textsc{Im}\ f \ = \ \textsc{dim}_{\mathbb{K}}V.$$

Який фізичний зміст цієї формули? Ядро вимірює те, що пропало при відображенні, образ – це те, що залишилось. Їх сума – інваріант.

Перш за все зауважимо, що $\textsc{Ker}\ f$, $\textsc{Im}\ f$ – скінченно вимірні простори. Це випливає з наступної леми.

Лемма 5.18. *1. Якщо U – підпростір V і V її скінченно вимірний, то U – також скінченно вимірний.*

2. Якщо $f : V \to W$ – лінійне відображення і V – скінченно вимірний, то $\textsc{Im}\ f$ *– теж скінченно вимірний.*

Доведення. Доведення 1 випливає з леми про монотонність.

Доведемо 2. Нехай $\textsc{dim}_{\mathbb{K}}V = n$. Тоді в просторі $\textsc{Im}\ f$ довільні $n+1$ вектори w_1, …, $w_{n+1} \in \textsc{Im}\ f$ лінійно залежні. Чому? Оскільки $w_i \in \textsc{Im}\ f$, то існують $v_1, \dots$, v_{n+1} $\in V$, такі що $w_i = f(v_i)$ для довільного $i = 1$, …, $n+1$. Розглянемо $v_1, \dots$, $v_{n+1} \in V$. Оскільки $\textsc{dim}_{\mathbb{K}}V = n$, то ці вектори лінійно залежні. Тобто $\lambda_1 v_1 + \dots + \lambda_{n+1} v_{n+1} = 0$ і не всі $\lambda_i = 0$. Подіємо f на праву і ліву частини. Одержимо $0 = f(0) = f(\lambda_1 v_1 + \dots + \lambda_{n+1} v_{n+1}) \overset{f-\text{лін}}{=} \lambda_1 f(v_1) + \dots + \lambda_{n+1} f(v_{n+1}) = \lambda_1 w_1 + \dots + \lambda_{n+1} v_{n+1}$. Ми показали, що w_1, …, w_{n+1} – лінійно залежні, а отже $\textsc{Im}\ f$ – скінченно вимірний. $\square$

Доведення теореми. Побудуємо спеціальний базис в просторі V.

Розглянемо деякий базис $\{v_1, \dots, v_s\}$ підпростору $\textsc{Ker}\ f$. При цьому $s = \textsc{dim}_{\mathbb{K}}\textsc{Ker}\ f$. Розглянемо відображення $f : V \to W$ і виберемо в $\textsc{Im}\ f$ деякий базис $\{w_1, \dots, w_r\}$ $\subset W$. Тоді кожен w_i є образом деякого елемента з простору V. Це означає: $w_i = f(v_{s+i})$, $i = 1, \dots, r$. Одержимо в V r елементів v_{s+1}, …, v_{s+r}. Розглянемо таку систему векторів в просторі V

$$\mathcal{V} \ = \ \{v_1, \dots, v_s,\, v_{s+1}, \dots, v_{s+r}\}.$$

Стверджується, що система $\mathcal{V} \subset V$ є базисом простору V. Для цього потрібно довести, що

1. система $\mathcal{V}$ – лінійно незалежна;

2. лінійна оболонка системи $(\mathcal{V})$ дорівнює всьому простору V.

Звідси випливатиме (за деякою попередньою лемою), що $\mathcal{V}$ – базис V.

Доведемо лінійну незалежність. Припустимо, що для $\lambda_1, \ldots, \lambda_{n+r} \in \mathbb{K}$ виконується така рівність:

$$\lambda_1 v_1 + \ldots + \lambda_s v_s + \lambda_{s+1} v_{s+1} + \ldots + \lambda_{s+r} v_{s+r} = 0.$$

Доведемо, що $\lambda_1 = \ldots = \lambda_s = \lambda_{s+1} = \ldots = \lambda_{s+r} = 0$. Подіємо на цю рівність f. Одержимо $\lambda_1 f(v_1)+ \ldots + \lambda_s f(v_s)+ \lambda_{s+1} f(v_{s+1})+ \lambda_{s+r} f(v_{s+r}) = 0$. Оскільки $\{v_1, \ldots, v_s\} \subset \textsc{Ker}\ f$, то $\lambda_1 f(v_1)+ \ldots +\lambda_s f(v_s) = 0 + \ldots + 0 = 0$. З іншого боку, $f(v_{s+i}) = w_i$. рівність перепишеться у вигляді $\lambda_{s+1} w_1+ \ldots +\lambda_{s+r} w_r = 0$. Але $\{w_1, \ldots, w_r\}$ – базис в $\textsc{Im}\ f$, звідси одержимо, що $\lambda_{s+1} = \ldots = \lambda_{s+r} = 0$. Тоді $\lambda_1 v_1+ \ldots +\lambda_s v_s = 0$. Але $\{v_1, \ldots, v_s\}$ – базис $\textsc{Ker}\ f$, звідси випливає, що $\lambda_1 = \ldots = \lambda_s = 0$.

Залишилося довести, що довільний вектор $v \in V$ є лінійною комбінацією векторів $v_1, \ldots, v_s, v_{s+1}, \ldots, v_{s+r}$. Або, що $(\mathcal{V}) = V$.

Розглянемо довільний вектор $v \in V$. Тоді $f(v) \in \textsc{Im}\ f \implies f(v)$ виражається через базис $w_1, \ldots, w_r$. Запишемо $f(v) = \mu_1 w_w+ \ldots +\mu_r w_r$, $\mu_1, \ldots, \mu_r \in \mathbb{K}$.

Розглянемо елемент $v' \in V$, $v' = \mu_1 v_{s+1}+ \ldots \mu_r v_{s+r}$. Тоді $v - v' \in \textsc{Ker}\ f$. Чому? $f(v-v') = f(v)- f(v') = \mu_1 w_1+ \ldots \mu_r w_r- f(\mu_1 v_{s+1}+\ldots+\mu_r v_{s+r}) \overset{f-\text{лін}}{=} \mu_1 w_1+ \ldots \mu_r w_r- \mu_1 f(v_{s+1})-\ldots- \mu_r f(v_{s+r}) = \mu_1 w_1+ \ldots \mu_r w_r- \mu_1 w_1- \ldots - \mu_r w_r = 0$. Отже, $f(v-v') = 0 \implies v - v' \in \textsc{Ker}\ f \implies v - v' = \lambda_1 v_1 + \ldots +\lambda_s v_s$ (зображення через базис ядра). Звідси одержуємо, що $v = \lambda_1 v_1 + \ldots +\lambda_s v_s + v' = \lambda_1 v_1 + \ldots +\lambda_s v_s+ \mu_1 v_{s+1}+ \ldots +\mu_r v_{s+r} \implies v \in (\mathcal{V}) = (v_1, \ldots, v_s, v_{s+1}, \ldots, v_{s+r})$. Ми довели, що $\{v_1, \ldots, v_s, v_{s+1}, \ldots, v_{s+r}\}$ – базис простору V. Звідси $\textsc{dim}_{\mathbb{K}} V = s + r$. З іншого боку, $\{v_1, \ldots, v_s\}$ – базис $\textsc{Ker}\ f$. Звідси $\textsc{dim}_{\mathbb{K}}\textsc{Ker}\ f = s$. Оскільки $\{w_1, \ldots, w_r\}$ – базис образу $\textsc{Im}\ f$, то $\textsc{dim}_{\mathbb{K}}\textsc{Im}\ f = r$. Ми одержали: $s + r = n$, тобто $\textsc{dim}_{\mathbb{K}}\textsc{Ker}\ f+ \textsc{dim}_{\mathbb{K}}\textsc{Im}\ f = \textsc{dim}_{\mathbb{K}} V$.

Доведення теореми закінчено.

Наслідки.

Застосуємо цю теорему у випадку, коли $f = \mathrm{L}_A$, де A – $m \times n$–матриця і $\mathrm{L}_A : \mathbb{K}^n \to \mathbb{K}^m$. В цьому випадку $V = \mathbb{K}^n$, $W = \mathbb{K}^m$. Що таке $\textsc{Ker}\ \mathrm{L}_A$? $\textsc{Ker}\ \mathrm{L}_A = \mathcal{L}$ – простір розв'язків однорідної системи $Ax = 0$. Його розмірність $\textsc{dim}_{\mathbb{K}}\textsc{Ker}\ \mathrm{L}_A = s$, де s – кількість лінійно незалежних розв'язків в ФСР системи $Ax = 0$.

Що таке $\textsc{dim}_{\mathbb{K}}\mathrm{L}_A$? $\textsc{dim}_{\mathbb{K}}\mathrm{L}_A = \textsc{rank}_c A = r$. В цьому випадку доведена теорема дає наслідок $\boxed{n = s + r}$, тобто, кількість невідомих дорівнює кількості лінійно незалежних векторів в $\textsc{Im}\ \mathrm{L}_A$ плюс ранг матриці. Зокрема, система $Ax = b$ буде визначеною (матиме єдиний розв'язок) $\iff \boxed{n = r}$, тобто коли стовпчиковий ранг матриці дорівнює кількості невідомих.

Дослідимо, коли система буде сумісною.

Теорема 5.3 (Кронекера-Капелі). *Нехай A – це $m \times n$–матриця над полем $\mathbb{K}$, $A = (\mathbf{a_1} \mid \ldots \mid \mathbf{a_n})$ – розклад на стовпчики. Система $Ax = b$ буде сумісною тоді і тільки тоді, коли*

$$\textsc{rank}(\mathbf{a_1} \mid \ldots \mid \mathbf{a_n}) \ = \ \textsc{rank}(\mathbf{a_1} \mid \ldots \mid \mathbf{a_n} \mid \mathbf{b})$$

або стовпчиковий ранг матриці системи дорівнює стовпчиковому рангу розширеної матриці системи: $\text{RANK}_c A = \text{RANK}_c \overline{A}$.

Доведення. Зауважимо, що, згідно з означенням, $\text{RANK}_c(\mathbf{a_1} \mid \ldots \mid \mathbf{a_n})$ дорівнює рангу системи векторів $(\mathbf{a_1}, \ldots, \mathbf{a_n})$.

Система $Ax = b$ сумісна $\iff \exists\ l = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \ldots \\ \lambda_n \end{pmatrix}$ такий, що $Al = b \iff \quad \iff \exists$ $\begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \ldots \\ \lambda_n \end{pmatrix}$ такий, що $\lambda_1 \mathbf{a_1} + \ldots + \lambda_n \mathbf{a_n} = \mathbf{b} \iff Al = \lambda_1 \mathbf{a_1} + \ldots + \lambda_n \mathbf{a_n} \iff$ $b \in (\mathbf{a_1}, \ldots, \mathbf{a_n}) \iff (\mathbf{a_1}, \ldots, \mathbf{a_n}) = (\mathbf{a_1}, \ldots, \mathbf{a_n}, \mathbf{b}) \iff (\mathbf{a_1} \mid \ldots \mid \mathbf{a_n}) = (\mathbf{a_1} \mid \ldots \mid \mathbf{a_n} \mid \mathbf{b})$. З рівності цих просторів одержимо, що $b \in (\mathbf{a_1}, \ldots, \mathbf{a_n})$.

Навпаки, якщо $b \in (\mathbf{a_1}, \ldots, \mathbf{a_n})$, то $(\mathbf{a_1}, \ldots, \mathbf{a_n}, \mathbf{b}) \subset (\mathbf{a_1}, \ldots, \mathbf{a_n})$ – підпростір. Отже, $(\mathbf{a_1}, \ldots, \mathbf{a_n}, \mathbf{b}) = (\mathbf{a_1}, \ldots, \mathbf{a_n})$. Але оскільки $(\mathbf{a_1}, \ldots, \mathbf{a_n})$ завжди є підпростором в просторі $(\mathbf{a_1}, \ldots, \mathbf{a_n}, \mathbf{b})$, вони будуть рівними тоді і лише тоді, якщо $\text{DIM}_{\mathbb{K}}(\mathbf{a_1}, \ldots, \mathbf{a_n}, \mathbf{b}) = \text{DIM}_{\mathbb{K}}(\mathbf{a_1}, \ldots, \mathbf{a_n})$. За означенням, звідси маємо $\text{RANK}(\mathbf{a_1} \mid \ldots \mid \mathbf{a_n} \mid \mathbf{b}) = \text{RANK}(\mathbf{a_1} \mid \ldots \mid \mathbf{a_n})$. □

6 Системи лінійних рівнянь

1. Розглянемо A – $m \times n$–матриця над $\mathbb{K}$. Розглянемо однорідну систему рівнянь

$$Ax \;=\; b.$$

Нехай $\mathcal{L} \subset \mathbb{K}^n$ – множина розв'язків системи $Ax = 0$. Ми довели, що $\mathcal{L} = \text{Ker } \mathrm{L}_A$, де $\mathrm{L}_A : \mathbb{K}^n \to \mathbb{K}^m$, таке що $\mathrm{L}_A(l) = Al$, $l \in \mathbb{K}^n$.

Наслідок. Множина $\mathcal{L}$ є підпростором в просторі $\mathbb{K}^n$.

Доведення. Якщо $f : V \to W$ – лінійне відображення, то $\text{Ker } f \subset V$ – підпростір. Але $\mathcal{L} = \text{Ker } \mathrm{L}_A$.

Означення 6.1. *Система* $\{l_1, \ldots, l_t\} \subset \mathfrak{L}$ *називається* **фундаментальною системою розв'язків** *або ФСР, якщо довільний розв'язок* $l \in \mathfrak{L}$ *однозначно представляється у вигляді лінійної комбінації* $l = \sum_{i=1}^{t} \lambda_i l_i$. *Еквівалентне означення:* $\{l_1, \ldots, l_t\}$ *є ФСР* $\iff$ $\{l_1, \ldots, l_t\}$ *є базисом в просторі розв'язків* $\mathcal{L}$.

1. Існування ФСР. Сформулюємо твердження (більш загального характеру, з якого випливає існування).

Твердження 6.1. *Наступні два твердження еквівалентні.*

1. *Простір V скінченно вимірний.*

2. *існує таке $N \in \mathbb{N}$, що довільні N векторів v_1, …, v_n є лінійно залежними.*

Доведення. 1 $\Longrightarrow$ 2. Якщо V – скінченно вимірний, то існує ізоморфізм $f : V \to \mathbb{K}^n$ ($n = \text{DIM}_{\mathbb{K}} V$). Тоді покладемо $N = n+1$. Якщо v_1, …, $v_{n+1} \in V$ довільні, подіємо на них f. $f(v_1), \ldots, f(v_{n+1}) \in \mathbb{K}^n$ є лінійно залежними в $\mathbb{K}^n$. Подіємо f^{-1}. Звідси випливає, що v_1, …, v_{n+1} теж лінійно залежні. Чому? Якщо $\sum\limits_{i=1}^{n-1} \lambda_i f(v_i) = 0 \overset{f^{-1}}{\Longrightarrow} \sum\limits_{i=1}^{n-1} \lambda_i v_i = 0.$

Навпаки, 2 $\Longrightarrow$ 1. Припустимо, що існує N таке, що v_1, …, $v_N \in V$ – завжди лінійно залежні. Тоді виберемо максимальну лінійно незалежну систему векторів v_1, …, $v_n \in V$ (Звичайно, $n < N$). Тоді система $\mathcal{V} = \{v_1, \ldots, v_n\}$ буде базисом. Чому? Розглянемо довільний вектор $v \in V$ і систему $\{v_1, \ldots, v_n, v\}$. Внаслідок максимальності $\{v_1, \ldots, v_n, v\}$ (за лемою про лінійний вектор) система $\{v_1, \ldots, v_n, v\}$ буде лінійно залежна і вектор v виражається через v_1, …, v_n.

Отже, для довільного вектора $v \in V$ одержали, що $v \in (v_1, \ldots, v_n)$, тобто $V = (v_1, \ldots, v_n)$, а оскільки, крім того, $\{v_1, \ldots, v_n\}$ – лінійно незалежна система, то $\mathcal{V} = \{v_1, \ldots, v_n\}$ є базисом простору V. □

Наслідок (з доведення). Якщо $\text{DIM}_{\mathbb{K}} V = n$ і v_1, …, $v_N \in V$, причому $N > n$, то система $\{v_1, \ldots, v_N\}$ – лінійно залежна.

Наслідок 6.1. *В множині $\mathcal{L}$ існує ФСР. Якщо $\{l_1, \ldots, l_t\}$ і $\{l'_1, \ldots, l'_s\}$ – дві такі системи, то $t = s$.*

Доведення. Як ми довели раніше, $\mathcal{L}$ є підпростором в просторі $\mathbb{K}^n \Longrightarrow e_1, e_2, \ldots, e_{n+1} \in \mathcal{L}$ – деякі вектори, то вони є лінійно залежними. Чому? Тому що $\mathcal{L} \subset \mathbb{K}^n$, а в $\mathbb{K}^n$ довільні $n+1$ векторів лінійно залежні. Звідси, за доведеною лемою, одержимо, що $\mathcal{L}$ – скінченно вимірний простір $\Longrightarrow$ в $\mathcal{L}$ існує скінченний базис l_1, …, l_t: якщо l'_1, …, l'_s – інший базис, то $t = s$. Але базис в $\mathcal{L}'$ – це те саме, що ФСР. □

Отже, t – це інваріант нашої системи – кількість векторів в ФСР. Слово інваріант означає те, шо не змінюється при різних перетвореннях системи.

В методі Гауса t – кількість вільних невідомих.

6.1 Будова множини розв'язків сумісної системи

Твердження 6.2. *Нехай $Ax = b$ – деяка сумісна лінійна система рівнянь, $\mathcal{L}_b$ – це множина всіх її розв'язків, $\mathcal{L} = \mathcal{L}_0$ – множина розв'язків однорідної системи $Ax = 0$, x_0 – довільний розв'язок (частковий). Тоді має місце така рівність*

$$\mathcal{L}_b \ = \ x_0 + \mathcal{L} \ = \ \{x_0 \in \mathcal{B} \| \, l \in \mathcal{L}\}.$$

Доведення. Розглянемо довільний вектор з $x_0 + \mathcal{L}$. Доведемо, що він є розв'язком.

$$x_0 + l \in x_0 + \mathcal{L} \implies A(x_0 + l) \overset{f-\text{дистр.}}{\Longrightarrow} Ax_0 + al0b + 0.$$

$$x_0 \in \mathcal{L}_b,\ l \in \mathcal{L}_0 \implies x_0 + l \in \mathcal{L}_b.$$

Навпаки, нехай $y \in \mathcal{L}_b$, тобто $Ay = b$. Розглянемо $l = y - x_0$. Тоді $l \in \mathcal{L}$. Чому? Тому, що $Al = A(y - x_0) = Ay - Ax_0 = b - b = 0 \implies l \in \mathcal{L}$. Тому y має вигляд $x_0 + \mathcal{L}$, де $l \in \mathcal{L} \implies y \in x_0 + \mathcal{L}$. $\square$

Отже, довільний вектор $x_0 + l$ є розв'язком, і навпаки, довільний розв'язок y має вигляд $x_0 + l$, $l \in \mathcal{L}$.

Означення 6.2. *Нехай $W \subset V$ – деякий підпростір і $v_0 \in V$ – деякий вектор. Множина векторів виду $v_0 + W = \{v_0 + w \mid w \in W\}$ називається* **афінним** *підпростором (який, взагалі кажучи, не є векторним підпростором). Отже, множина розв'язків $x_0 + \mathcal{L}$ – це афінний підпростір.*

Вправа 1. $w_0 + W$ буде векторним підпростором в $W \iff w_0 \in W$ і тоді $w_0 + W = W$. **2.** $v_0 + W = v_0' + W \iff v_0 + v_0' \in W$.

6.2 Базиси і лінійні відображення

Твердження 6.3 (про лінійне відображення і базис)**.** *Нехай $f : V \to W$ – лінійне відображення деякий векторних просторів і $\mathcal{B} = v_{i i \in I}$ – деякий базис з простору V. Тоді лінійне відображення $f : V \to W$ однозначно визначається своїм обмеженням на базис Φ $(= \Phi_{\mathcal{B}}) = f_{\mathcal{B}} : \mathcal{B} \to W$. Навпаки, кожне відображення множин $\Phi : \mathcal{B} \to W$ однозначно визначає лінійне відображення $f : V \to W$ таке, що $f_{\mathcal{B}} = \Phi : \mathcal{B} \to W$.*

Доведення. Припустимо, що ми маємо два лінійних відображення $f, f' : V \to W$, причому $f|_{\mathcal{B}} = f'|_{\mathcal{B}}$ (це означає, що $\forall\, i \in I \ \ f(v_i) = f'(v_i)$). Доведемо, що $f = f'$, тобто $\forall\, v \in V$ виконується $f(v) = f'(v)$. Як?

1. Вектор v розкладається за базисом $\mathcal{B}$, тобто $v = \sum\limits_{i \in I} \lambda_i v_i$. Тоді $f(v) = f(\sum\limits_{i \in I} \lambda_i v_i)$ $\overset{f-\text{лін.}}{=} \sum\limits_{i \in I} \lambda_i f(v_i) \overset{\text{за умовою}}{=} \sum\limits_{i \in I} \lambda_i f'(v_i) = f'(\sum\limits_{i \in I} \lambda_i v_i) = f'(v)$. Отже, $f(v) = f'(v)$.

Однозначно визначається.

Навпаки, припустимо, що задане деяке відображення $\Phi : \mathcal{B} \to W$. Побудуємо лінійне відображення $f \in \mathcal{L}(V, W)$ таке, що $f|_{\mathcal{B}} = \Phi$. За доведеним, таке f – єдине.

Нам необхідно визначити $f(v)$ для довільного $v \in V$, знаючи, що для довільного $v_i \in \mathcal{B}$ виконується $f(v_i) = \Phi(v_i)$. Як визначити $f(v)$? Для цього розкладемо v за базисом: $v = \sum\limits_{i \in I} \lambda_i v_i$. Якщо f – лінійне, то повинно було б виконуватись: $f(v) = \sum\limits_{i \in I} \lambda_i f(v_i)$. Тоді покладемо $f(v) = \sum\limits_{i \in I} \lambda_i f(v_i)$. Це коректне означення, тому що розклад $v = \sum\limits_{i \in I} \lambda_i v_i$ єдиний.

Залишилось перевірити, що побудоване f – лінійне відображення.

1. Необхідно перевірити: $f(v + v') = f(v) + f(v')$.

Запишемо розклад за базисом: $v = \sum_{i\in I}\lambda_i v_i$, $v' = \sum_{i\in I}\lambda'_i v_i \implies v + v' = \sum_{i\in I}\lambda_i v_i + \sum_{i\in I}\lambda_i v'_i = \sum_{i\in I}(\lambda_i + \lambda'_i)v'_i$. За означенням, $f(v) = \sum_{i\in I}\lambda_i\Phi(v_i)$, $f(v') = \sum_{i\in I}\lambda'_i\Phi(v_i)$. Тоді $f(v) + f(v') = \sum_{i\in I}\lambda_i\Phi(v_i) + \sum_{i\in I}\lambda'_i\Phi(v_i) = \sum_{i\in I}(\lambda_i + \lambda'_i)\Phi(v_i) = f(v + v')$. Отже, ми довели рівність $f(v) + f(v') = f(v + v')$.

Аналогічно перевіряється рівність $f(\lambda v) = \lambda f(v)$. Коли $v = \sum_{i\in I}\lambda_i v_i \implies \lambda v = \sum_{i\in I}(\lambda\lambda_i)v_i$ (враховуючи, що розклад за базисом однозначний). Тоді, за означенням, $f(\lambda v) = \sum_{i\in I}(\lambda\lambda_i)\Phi(v_i) = \lambda\sum_{i\in I}(\lambda_i)\Phi(v_i) = \lambda f(v)$.

Ще потрібна перевірка, чому $f(v_i) = \Phi(v_i)$. Коли $v = v_i$, то розклад v за базисом має вигляд $v_i = 1 \cdot v_i$ (коефіцієнти при решті базисних векторів дорівнюють нулю, а при v_i – 1). Тоді $v_i = 1 \cdot v_i + \ldots$. За означенням f, $f(v_i) = 1 \cdot \Phi(v_i) + 0 \ldots = 1 \cdot \Phi(v_i) = \Phi(v_i)$. $\square$

Отже, кожне лінійне відображення визначається своїм значенням за базисом, і навпаки, коли задане відображення на базисі, його можна продовжити на весь простір за формулою.

Наслідки (для скінченно вимірних просторів).

1. Нехай $f : V \to W$ і $\mathcal{B}$ – деякий базис в просторі V. Тоді $f = 1_V \iff f|_{\mathcal{B}} = 1_{\mathcal{B}}$, тобто для довільного $v_i \in \mathcal{B}$ $\Phi(v_i) = v_i$.

2. Нехай $f : V \to W$ – лінійне відображення і $\mathcal{B} = \{v_i \mid i \in I\}$ – базис в просторі V. Тоді f є ізоморфізмом $\iff f(\mathcal{B}) = \{f(v_i)\| \ i \in I\}$ є базисом в просторі W.

Доведення. Якщо f – ізоморфізм, тоді $f(\mathcal{B})$ є базисом, тому що лінійна оболонка $(f(v_i) \mid i \in I) = W$ і $f(\mathcal{B})$ – лінійно незалежна система.

Навпаки, припустимо, що $f(\mathcal{B})$ – базис простору W. Задамо відображення $\Gamma : f(\mathcal{B}) \to V$, а саме, $\Gamma(f(v_i)) = v_i$ для довільного $v_i \in \mathcal{B}$. Тоді відображення Γ однозначно визначає лінійне відображення $g : W \to V$, причому таке, що $g(f(v_i) = v_i$. Тоді стверджується, що $gf = 1_V$. Чому? Тому що на кожному базисному векторі $v_i \in \mathcal{B}$ виконується $(gf)v_i = g(f(v_i)) = v_i = 1_V \cdot v_i$ – за означенням, оскільки $f(v_i) \in f(\mathcal{B})$. Отже, $gf = 1_V$.

Аналогічно, $fg = 1_W$. Чому? Подіємо fg на базисний вектор $f(v_i) \in f(\mathcal{B})$.

$(fg)(f(v_i)) = f(g(f(v_i))) = f(v_i) \implies fg = \mathbf{1}_V$. Звідси робимо висновок, що f – бієкція, тобто ізоморфізм. $\square$

Зауваження 6.1. *Нехай A - $m \times n$-матриця над полем $\mathbb{K}$ і $A = (\mathbf{a_1}| \ldots |\mathbf{a_n})$, $\mathrm{L}_A : \mathbb{K}^n \to \mathbb{K}^m$ - відповідне цій матриці лінійне відображення, $\mathbf{e_i}$ - стандартний базисний вектор. Тоді* $\mathrm{L}_A(e_i) = A\mathbf{e_i} = (\mathbf{a_1} \ldots \mathbf{a_n})\begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \end{pmatrix} = \mathbf{a_i}$. *підсумок:* $\mathrm{L}_A(\mathbf{e_i}) = \mathbf{a_i}$.

Має місце таке твердження

Твердження 6.4. *Нехай $f \in \mathcal{L}(\mathbb{K}^n, \mathbb{K}^m)$ (тобто відображення між координатними векторними просторами). Тоді існує і єдина матриця $A = [f]$, така що $\mathrm{L}_A = f$. більше того, $[] : \mathcal{L}(\mathbb{K}^n, \mathbb{K}^m) \to Mat_{m\times n}(\mathbb{K})$ (простір матриць) є ізоморфізмом векторних просторів.*

Доведення. Нехай $f \in \mathcal{L}(\mathbb{K}^n, \mathbb{K}^m)$. Позначимо через $\mathbf{a_i} \in \mathbb{K}^\mathbf{m}$ вектор $\mathbf{a_i} = f(\mathbf{e_i})$, $i = 1, \dots, m$ і введемо матрицю

$$A \;=\; (\mathbf{a_1}|\mathbf{a_2}|\dots|\mathbf{a_n}).$$

Тоді $\mathrm{L}_A = f$. Чому? Обидва відображають $\mathbb{K}^n \to \mathbb{K}^m$. Досить перевірити, що вони збігаються на базисних векторах $\{\mathbf{e_i}\}_{\mathbf{i=1,\dots,n}}$. Перевіримо.

$f(\mathbf{e_i}) = \mathbf{a_i}$, за означенням $\mathbf{a_i}$.

$\mathrm{L}_A(\mathbf{e_i}) = A\mathbf{e_i}$, за попереднім зауваженням.

Отже, $\mathrm{L}_A = f$.

□

Матриці утворюють координатний векторний простір $Mat_{m\times n}(\mathbb{K})$ розмірності $m \times n$.

Позначимо побудовану матрицю A через $[f]$ і доведемо для $f, g \in \mathcal{L}(\mathbb{K}^n, \mathbb{K}^m)$, що

$$[f+g] \;=\; [f]+[g]; \quad [\lambda f] \;=\; \lambda[f].$$

Доведення. Позначимо $[f] = A$, $[g] = B$, тобто $f = \mathrm{L}_A$, $g = \mathrm{L}_B$. Нехай $A = (\mathbf{a_1}|\dots|\mathbf{a_n})$, $B = (\mathbf{b_1}|\dots|\mathbf{b_n})$. Тоді матриця $A+B$ запишеться у вигляді $(\mathbf{a_1}+\mathbf{b_1}|\dots|\mathbf{a_n}+\mathbf{b_n})$. Тоді $(f+g)(e_i) = f(e_i) + g(e_i) = a_i + b_i$. Але це є $i-$тим стовпчиком в матриці $A+B$, отже це дорівнює $\mathrm{L}_{A+B}(e_i)$.

Звідси випливає: $[f+g] = [f]+[g]$. □

Вправа. Другу рівність перевірити самостійно.

Остаточний висновок:

відображення з $\mathcal{L}(\mathbb{K}^n, \mathbb{K}^m)$ в $Mat_{m\times n}(\mathbb{K})$ є лінійне відображення. Воно бієктивне, тому що кожне відображення з $\mathcal{L}(\mathbb{K}^n, \mathbb{K}^m)$ однозначно визначається своєю матрицею.

6.3 Множення матриць

Нехай маємо простори $\mathbb{K}^n$, $\mathbb{K}^m$, $\mathbb{K}^l$. Розглянемо деякі лінійні відображення

$$\mathbb{K}^n \stackrel{\mathrm{L}_A}{\Longrightarrow} \mathbb{K}^m \stackrel{\mathrm{L}_B}{\Longrightarrow} \mathbb{K}^l$$

Композиція лінійного відображення $\mathrm{L}_B\mathrm{L}_A$ – це деяке лінійне відображення. За доведеним, $\mathrm{L}_B\mathrm{L}_A = \mathrm{L}_C : \mathbb{K}^n \to \mathbb{K}^l$. Отже, повинно існувати правило обчислення матриці C за матрицями B та A.

Нехай $A = (a_{ij})_{1\leqslant i\leqslant m, 1\leqslant j\leqslant n}$, $B = (b_{ij})_{1\leqslant i\leqslant l, 1\leqslant j\leqslant m}$.

Мнемонічне правило. На вході матриця B розміру $l \times m$ та матриця A розміру $m \times n$; на виході матриця розміру $l \times n$.

$$\begin{array}{ccccc} B & \cdot & A & \Longrightarrow & C \\ l\times m & & m\times n & & l\times n \end{array}$$

Лемма-означення 6.1. *В наведеній ситуації $l \times n-$матриця C така, що $\mathrm{L}_C = \mathrm{L}_B\mathrm{L}_A$, називається добутком матриць B і A і позначається $B \times A$. $(i,j)-$тий елемент матриці $C = (c_{ij})_{1\leqslant i\leqslant l, 1\leqslant j\leqslant n}$ визначається за формулою*

$$c_{ij} \ = \ \sum_{k=1}^{m} b_{ik}a_{ki}, \quad 1 \leqslant i \leqslant l, \ 1 \leqslant j \leqslant n.$$

Доведення (формули). Застосуємо L_C до e_j:

$\mathrm{L}_C(e_j) = \mathrm{L}_B\mathrm{L}_A(e_j) = \mathrm{L}_B(\mathrm{L}_A(e_j)) = \mathrm{L}_B(a_j)$.

Запишемо $B = (b_1 \dots b_m)$. Тоді $B(a_j) = Ba_j = (b_1 \dots b_m) \begin{pmatrix} a_{1j} \\ a_{2j} \\ \dots \\ a_{mj} \end{pmatrix} = a_{1j}b_1 + \ a_{2j}b_2 + \dots + a_{mj}b_m$.

візьмемо в цьому стовпчику елемент, який знаходиться на $i-$тому місці. Маємо $b_{i1}a_{1j} + \ b_{i2}a_{2j} + \dots + b_{im}a_{mj} = c_{ij}$.

Ми одержали формулу для множення матриць.

Означення 6.3. *$m \times n-$матриця A називається обертовною, якщо відображення L_A є ізоморфізм.*

Наслідок. 1. Якщо A – обертовна, то $m = n$. 2. A – обертовна $\iff \exists \ n \times n-$матриця B, така що $AB = BA = E_n$.

Зауваження 6.2. *Коли X та Y деякі матриці, то запис XY автоматично означає, що довжина рядка матриці X дорівнює висоті стовпчика матриці Y.*

Властивості

1) Асоціативність

$(CB)A = C(BA)$, A, B, C – деякі матриці $A \in \mathrm{Mat}_{m\times n}(\mathbb{K})$, $B \in \mathrm{Mat}_{l\times m}(\mathbb{K})$, $C \in \mathrm{Mat}_{k\times l}(\mathbb{K})$.

Доведення. Розглянемо лінійне відображення $\mathbb{K} \overset{nL_A}{\Longrightarrow} \mathbb{K} \overset{mL_B}{\Longrightarrow} \mathbb{K} \overset{lL_C}{\Longrightarrow} \mathbb{K}$

$(CB)A = C(BA) \iff \mathrm{L}_{(CB)A} = \mathrm{L}_{C(BA)}$

$\mathrm{L}_{(CB)A} \overset{def}{\Longrightarrow} \mathrm{L}_{CB} \cdot \mathrm{L}_A = (\mathrm{L}_C \cdot \mathrm{L}_B) \cdot \mathrm{L}_A$.

З іншого боку, $\mathrm{L}_{C(BA)} = \mathrm{L}_C \cdot \mathrm{L}_{BA} = \mathrm{L}_C(\mathrm{L}_B \cdot \mathrm{L}_A)$. Композиція відображень асоціативна, отже $(\mathrm{L}_C \cdot \mathrm{L}_B) \cdot \mathrm{L}_A = \mathrm{L}_C(\mathrm{L}_B \cdot \mathrm{L}_A) = \mathrm{L}_{CB} \cdot \mathrm{L}_A = (\mathrm{L}_C \cdot \mathrm{L}_B) \cdot \mathrm{L}_A \implies (CB)A = C(BA)$

□

2) $A(B + C) = AB + AC$, $(A + B)C = AC + BC$

Доведення. Доведемо першу властивість, друга доводиться аналогічно.

$A \in \mathrm{Mat}_{l\times n}(\mathbb{K})$ $B, C \in \mathrm{Mat}_{m\times n}(\mathbb{K})$

$A(B+C) = AB + AC \iff \mathrm{L}_{A(B+C)} = \mathrm{L}_{AB} + \mathrm{L}_{AC}$

$\mathbb{K}^n \overset{\mathrm{L}_{B+C}}{\Longrightarrow} \mathbb{K}^m \overset{\mathrm{L}_A}{\Longrightarrow} \mathbb{K}^l$

$\mathrm{L}_{A(B+C)} = \mathrm{L}_A \cdot \mathrm{L}_{B+C} = \mathrm{L}_A(\mathrm{L}_B + \mathrm{L}_C) = \mathrm{L}_A \cdot \mathrm{L}_B + \mathrm{L}_A \cdot \mathrm{L}_C = \mathrm{L}_{AB} + \mathrm{L}_{AC} = \mathrm{L}_{AB+AC}$ $\implies A(B+C) = AB + AC.$ $\square$

3. I_n – одинична матриця $I_n = \begin{pmatrix} 1 & \dots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \dots & 1 & 0 \\ 0 & \dots & 0 & 1 \end{pmatrix}$,

Для довільної $A \in \mathrm{Mat}_{m\times n}(\mathbb{K})$ має місце $I_m A = A I_n = A$.

Доведення. Зауважимо, що $\mathrm{L}_{I_m} = \mathrm{id}_{\mathbb{K}^m}$. Щоб довести, що $I_m A = A$, розглянемо $\mathrm{L}_{I_m A} = \mathrm{L}_{I_m} \cdot \mathrm{L}_A = \mathrm{id}_{\mathbb{K}^m} \cdot \mathrm{L}_A = \mathrm{L}_A \implies I_m A = A.$ $\square$

Завдання 6.1. *1. Доведіть рівність $AI_n = A$ (доводиться аналогічно попередній).*

2. Доведіть рівність $I_m A = AI_n = A$ за допомогою формули множення матриць.

Нехай $A \in \mathrm{Mat}_{m\times n}(\mathbb{K})$, $B \in \mathrm{Mat}_{l\times m}(\mathbb{K})$, $\lambda \in \mathbb{K}$. Тоді $\lambda(BA) = (\lambda B)A = B(\lambda A)$ – використовуючи властивість лінійного відображення з леми 5.8.

Означення 6.4. *Матриця A називається обертовною, якщо відображення L_A є ізоморфізмом.*

Зауважимо, що якщо для $A \in \mathrm{Mat}_{m\times n}(\mathbb{K})$ відображення $\mathrm{L}_A : \mathbb{K}^n \to \mathbb{K}^m$ є ізоморфізмом, то $n = m$ (матриця A квадратна).

Варіант означення.

Лемма 6.1. *$A \in \mathrm{Mat}_{n\times n}(\mathbb{K})$ є обертовною, тоді і лише тоді, коли існує матриця $A^{-1} \in \mathrm{Mat}_{n\times n}(\mathbb{K})$ така, що $AA^{-1} = A^{-1}A = I_n$.*

Доведення. Якщо $\mathrm{L}_A : \mathbb{K}^n \to \mathbb{K}^m$ ізоморфізм, то обернене відображення $(\mathrm{L}_A)^{-1} : \mathbb{K}^n \to \mathbb{K}^m$ теж лінійне, тому $(\mathrm{L}_A)^{-1} = \mathrm{L}_B$ для деякої матриці $B \in \mathrm{Mat}_{n\times n}(\mathbb{K})$. Тоді $\mathrm{L}_A\mathrm{L}_B = \mathrm{id}_{\mathbb{K}^n}$, $\mathrm{L}_B\mathrm{L}_A = \mathrm{id}_{\mathbb{K}^n}$, звідки одержимо $\mathrm{L}_{AB} = \mathrm{id}_{\mathbb{K}^n}$, $\mathrm{L}_{BA} = \mathrm{id}_{\mathbb{K}^n}$. Оскільки $\mathrm{id}_{\mathbb{K}^n} = \mathrm{L}_{I_n}$, то звідси одержимо $AB = I_n$, $BA = I_n$. Покладемо $A^{-1} = B$.

Навпаки, нехай існує така матриця A^{-1}, що $AA^{-1} = A^{-1}A = I_n$. Розглянувши відповідне лінійне відображення $\mathrm{L}_{AA^{-1}} = \mathrm{L}_{A^{-1}A} = \mathrm{L}_{I_n}$, одержимо, що $\mathrm{L}_A \cdot \mathrm{L}_{A^{-1}} = \mathrm{L}_{A^{-1}} \cdot \mathrm{L}_A = \mathrm{id}_{\mathbb{K}^n}$, звідки L_A – ізоморфізм. $\square$

Загальне зауваження.

З означення випливає, що A^{-1} – єдина така матриця для A (A^{-1} – матриця лінійного відображення $(\mathrm{L}_A)^{-1}$). Але єдиність A^{-1} можна вивести з властивості $AA^{-1} = A^{-1}A = I_n$. Дійсно, припустимо, що існують такі матриці $X, Y \in \mathrm{Mat}_{n\times n}(\mathbb{K})$, що $XA = I_n$, $AY = I_n$. Тоді $X = XI_n = X(AY) \overset{f-\text{асоц.}}{\Longrightarrow} (XA)Y = I_nY = Y$, тоді одержимо $X = Y$.

Множина усіх обертовних матриць порядку n позначається $\mathrm{GL}_n(\mathbb{K}) \subset \mathrm{Mat}_{n\times n}(\mathbb{K})$.

Лемма 6.2. *1. Якщо $A, B \in \mathrm{GL}_n(\mathbb{K})$, то $AB \in \mathrm{GL}_n(\mathbb{K})$. Більше того, $(AB)^{-1} = B^{-1}A^{-1}$.*

2. Якщо $A \in \mathrm{GL}_n(\mathbb{K})$, то $A^{-1} \in \mathrm{GL}_n(\mathbb{K})$, більше того, $(A^{-1})^{-1} = A$ тому, що для $(A^{-1})X = X(A^{-1}) = I_n$ виконується $X = A$.

Доведення. Достатньо знайти таку матрицю X, що $X(AB) = (AB)X = I_n$. Покладемо $X = B^{-1}A^{-1}$. Тоді $(B^{-1}A^{-1})(AB) = B^{-1}(A^{-1}A)B = B^{-1}I_nB = B^{-1}B = I_n$. □

Операція транспонування.

Транспонування – відображення $t : \mathrm{Mat}_{m\times n}(\mathbb{K}) \to \mathrm{Mat}_{n\times m}(\mathbb{K})$, яке задається для кожної пари $m, n \geq 1$ за правилом: $A \overset{t}{\Longrightarrow} A^t$,

$(A^t)_{ij} = A_{ji}$, $1 \leq i \leq n$, $1 \leq j \leq m$.

Приклад 6.1.

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} \quad \overset{t}{\Longrightarrow} \quad A^t = \begin{pmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{pmatrix}.$$

Приклад 6.2. *Операція транспонування матриць у випадку $n = 1$ або $m = 1$ співставляє вектори-рядки та вектори-стовпчики однакових вимірів. Маємо*

$$\begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \dots \\ \lambda_n \end{pmatrix}^t = \begin{pmatrix} \lambda_1 & \lambda_2 & \dots & \lambda_n \end{pmatrix}, \qquad \begin{pmatrix} \lambda_1 & \lambda_2 & \dots & \lambda_n \end{pmatrix}^t = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \dots \\ \lambda_n \end{pmatrix}.$$

Лемма 6.3 (властивості транспонування)**.** *1. Транспонування $t : \mathrm{Mat}_{m\times n}(\mathbb{K}) \to \mathrm{Mat}_{n\times m}(\mathbb{K})$ є лінійним відображенням, тобто $(A+B)^t = A^t + B^t$, $(\lambda A)^t = \lambda A^t$, $A, B \in \mathrm{Mat}_{m\times n}(\mathbb{K})$, $\lambda \in \mathbb{K}$.*

2. Якщо $A \in \mathrm{Mat}_{m\times n}(\mathbb{K})$, $B \in \mathrm{Mat}_{l\times n}(\mathbb{K})$, то $(BA)^t = A^tB^t$.

3. $(A^t)^t = A$, $I_n^t = I_n$.

4. $A \in \mathrm{GL}_n(\mathbb{K}) \implies A^t \in \mathrm{GL}_n(\mathbb{K})$, $(A^t)^{-1} = (A^{-1})^t$.

Доведення.

1. $((A+B)^t)_{ij} = (A+B)_{ji} = A_{ji} + B_{ji} = (A^t)_{ji} + (B^t)_{ji} = (A^t + B^t)_{ji}$. $(\lambda A)^t = \lambda A^t$ доводиться аналогічно.

2. Розглянемо $((BA)^t)_{ij} = (BA)_{ji} = \sum_k = 1^n B_{jk}A_{ki} = \sum_k = 1^n B^t_{kj}A^t_{ik} = \sum_k = 1^n A^t_{ik}B^t_{kj} = (A^tB^t)_{ij} \implies (BA)^t = (AB)$.

3. Доведення очевидне.

4. $A^t(A^{-1})^t = (A^{-1}A)^t = I_n^t = I_n$. $(A^{-1})^t = I_n$.

$\square$

Операція транспонування матриць у випадку $n = 1$ або $m = 1$ співставляє вектори-рядки та вектори-стовпчики однакових вимірів. Маємо

$$\begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \dots \\ \lambda_n \end{pmatrix}^t = \begin{pmatrix} \lambda_1 & \lambda_2 & \dots & \lambda_n \end{pmatrix}, \qquad \begin{pmatrix} \lambda_1 & \lambda_2 & \dots & \lambda_n \end{pmatrix}^t = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \dots \\ \lambda_n \end{pmatrix}.$$

6.4 Теорема про ранг матриці

Означення 6.5. *Довільну матрицю $A \in Mat_{m\times n}(\mathbb{K})$ можна представити у вигляді:*
$A = \begin{pmatrix} \underline{\mathbf{a^1}} \\ \underline{\mathbf{a^2}} \\ \underline{\vdots} \\ \mathbf{a^m} \end{pmatrix}$, *де $\mathbf{a^i}$ - вектори-рядки матриці A, $i = 1, 2, \dots, m$. Тоді*

$$A^t = \left((\mathbf{a^1})^\mathbf{t} \,|\, (\mathbf{a^2})^\mathbf{t} \,|\, \dots \,|\, (\mathbf{a^n})^\mathbf{t} \right)$$

- розклад матриці A^t на вектори-стовпчики. Рядковим рангом матриці A назвемо стовпчиковий ранг матриці A^t, тобто $\mathrm{RANK}_r A = \mathrm{RANK}_c A^t$. Оскільки $(A^t)^t = A$, то $\mathrm{RANK}_r A^t = \mathrm{RANK}_c A$.

Означення 6.6. *Позначимо через $I_{m\times n,r} \in Mat_{m\times n}(\mathbb{K})$ канонічну матрицю вигляду*

$$I_{m\times n,r} = \begin{pmatrix} 1 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 1 & 0 & \dots & 0 \\ 0 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & 0 & \dots & 0 \end{pmatrix},$$

де $0 \leqslant r \leqslant min(m, n)$. Очевидно, $(I_{m\times n,r})^t = I_{n\times m,r}$; $\mathrm{RANK}_c I_{m\times n,r} = \mathrm{RANK}_r I_{m\times n,r} = r$.

Теорема 6.1 (про ранг матриці)**.** *Для довільної матриці $A \in Mat_{m\times n}(\mathbb{K})$ має місце рівність*

$$\mathrm{RANK}_c A = \mathrm{RANK}_r A.$$

Тобто стовпчиковий та рядковий ранги матриці A рівні між собою. Це спільне значення називається рангом матриці A і позначається $\mathrm{RANK}\, A$.

Доведення теореми. З матрицею A зв'язане лінійне відображення $\mathrm{L}_A : \mathbb{K}^n \to \mathbb{K}^m$, з транспонованою матрицею A^t - лінійне відображення $\mathrm{L}_{A^t} : \mathbb{K}^m \to \mathbb{K}^n$. Оскільки, за означенням, $\operatorname{DIM}_{\mathbb{K}}\operatorname{IM} \mathrm{L}_A = \operatorname{RANK}_c A$, а $\operatorname{DIM}_{\mathbb{K}}\operatorname{IM} \mathrm{L}_{A^t} = \operatorname{RANK}_c A^t = \operatorname{RANK}_r A$, то теорема еквівалентна такому твердженню:

$$\operatorname{DIM}_{\mathbb{K}}\operatorname{IM} \mathrm{L}_A \;=\; \operatorname{DIM}_{\mathbb{K}}\operatorname{IM} \mathrm{L}_{A^t}. \tag{9}$$

Для доведення теореми сформулюємо та доведемо кілька лем.

Лемма 6.4. *Нехай ми маємо комутативну діаграму скінченно вимірних просторів та лінійних відображень*

$$\begin{array}{ccc} V & \xrightarrow{f} & W \\ {\scriptstyle g}\big\downarrow & & \big\downarrow{\scriptstyle h} \\ V_1 & \xrightarrow{f_1} & W_1 \end{array}$$

причому g, h - ізоморфізми, Тоді $\operatorname{DIM}_{\mathbb{K}}\operatorname{IM} f = \operatorname{DIM}_{\mathbb{K}}\operatorname{IM} f_1$. Більше того, існує ізоморфізм $F : \operatorname{IM} f \to \operatorname{IM} f_1$.

Зауваження 6.1. *Комутативність діаграми означає, що всі шляхи на діаграмі рівні, в даному випадку $hf = f_1 g$.*

Доведення.

Побудуємо ізоморфізм $F : \operatorname{IM}\ f \longrightarrow \operatorname{IM}\ f_1$. Для будь-якого $x \in \operatorname{IM}\ f \subset W$ покладемо $F(x) = h(x) \in W_1$.

1. Доведемо, що $F(x) \in \operatorname{IM}\ f_1$. Дійсно, $x \in \operatorname{IM}\ f \implies \exists v \;\; x = f(v) \implies F(x) = h(x) = h(f(v)) \overset{\text{ком.діагр.}}{=} f_1(g(v)) \in \operatorname{IM}\ f_1$.

2. Відображення F лінійне, оскільки лінійне h, а саме $F(x+y) = h(x+y) = h(x)+h(y) = F(x)+F(y)$, $x, y \in \operatorname{IM}\ f$; рівність $F(\lambda x) = \lambda F(x)$ доводиться аналогічно.

3. F є мономорфізмом. Дійсно, якщо $F(x) = 0$, то, за означенням, $h(x) = 0$, а оскільки h - ізоморфізм, то $x = 0$.

4. Доведемо, що F - епіморфізм. Розглянемо $y \in \operatorname{IM}\ f_1$. Тоді $y = f_1(v_1)$, де $v_1 \in V_1$. Оскільки g - ізоморфізм, то існує (і єдиний) $v \in V$, для якого $g(v) = v_1$. Розглянемо $x = f(v) \in \operatorname{IM}\ f$. Стверджується, що $F(x) = y$. Дійсно, $F(x) = h(x) = h(f(v)) = (hf)(v) = (f_1 g)(v) = f_1(g(v)) = f(v_1) = y \implies F$ - епіморфізм.

5. Оскільки F - мономорфізм і епіморфізм, то $F : \operatorname{IM}\ f \longrightarrow \operatorname{IM}\ f_1$ є ізоморфізмом, і отже, $\operatorname{DIM}_{\mathbb{K}}\operatorname{IM} f = \operatorname{DIM}_{\mathbb{K}}\operatorname{IM} f_1$.

$\square$

Завдання 6.2. *Довести, що $\operatorname{DIM}_{\mathbb{K}} Ker\, f = \operatorname{DIM}_{\mathbb{K}} Ker\, f_1$.*

Означення 6.7. *Нехай $f : V \to W$ - лінійне відображення скінченно вимірних просторів, $\mathcal{V} = \{v_1, \ldots, v_n\}$ - базис в V, $\mathcal{W} = \{w_1, \ldots, w_m\}$ - базис в W. Пару базисів $\mathcal{V}$ та $\mathcal{W}$ назвемо узгодженою з f парою базисів, якщо для деякого $1 \leqslant r \leqslant n$ має місце:*

$$\begin{cases} f(v_i) &= \;\; w_i, \quad 1 \leqslant i \leqslant r; \\ f(v_i) &= \;\; 0, \quad\;\; i \geqslant r+1. \end{cases}$$

Зауваження 6.2. *В цьому випадку* $r = \mathrm{DIM}_{\mathbb{K}}\mathrm{Im}\, f$.

Приклад. Нехай $V = \mathbb{K}^n$, $\mathcal{V} = \{e_1, \dots, e_n\}$ - стандартний базис в V, тобто $e_{ij} = \delta_{ij}$, $1 \leqslant i, j \leqslant n$ (тут e_{ij} - це $j-$та координата вектора e_i); $W = \mathbb{K}^m$, $\mathcal{W} = \{f_1, \dots, f_m\}$ - стандартний базис в W, тобто $f_{ij} = \delta_{ij}$, $1 \leqslant i, j \leqslant m$. І нехай $\mathrm{L}_A : \mathbb{K}^n \to \mathbb{K}^m$ - таке лінійне відображення, що $\mathcal{V}$, $\mathcal{W}$ - це узгоджена пара базисів. В цьому випадку $A = I_{m\times n,r}$ для деякого r.

Дійсно, розглянемо вектор e_i, тоді добуток Ae_i дорівнює

$$Ae_i = \begin{cases} f_i, & \text{якщо} \quad 1 \leqslant i \leqslant r; \\ 0, & \text{якщо} \quad i \geqslant r + 1. \end{cases}$$

Лемма 6.5. *Нехай U - векторний простір, $T \subset U$ - підпростір і $\{t_i\}_{i\in I}$ - деякий базис простору T. Тоді існує базис $\{u_j\}_{j\in J}$ простору U такий, що $\{t_i\}_{i\in I} \subset \{u_j\}_{j\in J}$.*

Альтернативне формулювання: Базис підпростору доповнюється до базису всього простору.

Доведення. Нехай $\mathcal{U} = \{u_j\}_{j\in J}$ - максимальна лінійно незалежна система векторів, яка містить $\{t_i\}_{i\in I}$ як підсистему. Тоді $\mathcal{U}$ - базис всього простору U. Чому? Оскільки $\mathcal{U}$ - максимальна лінійно незалежна система векторів, тоді, використовуючи лему про зайвий вектор, одержимо, що $(\mathcal{U}) = (\mathcal{U}) + v$ для будь-якого вектора $v \in U$. Звідси випливає, що $\mathcal{U}$ - це базис U. $\square$

Лемма 6.6. *Нехай $f : V \to W$ - довільне лінійне відображення скінченно вимірних просторів. Тоді існує пара базисів - відповідно $\mathcal{V}$ в просторі V та $\mathcal{W}$ в просторі W, яка узгоджена з відображенням f.*

Доведення. Виберемо базис $\{w_1, \dots, w_r\}$ в образі $\mathrm{Im}\, f$. Застосовуючи лему 2.4, доповнимо його до базису простору W і знайдений базис позначимо

$$\mathcal{W} = \{w_1, \dots, w_r, w_{r+1}, \dots, w_m\},$$

де $m = \mathrm{DIM}_{\mathbb{K}}W$. Розглянемо деяку систему векторів $\{v_1, \dots, v_r\} \subset V$ таких, що $f(v_1) = w_1$, ..., $f(v_r) = w_r$. Виберемо базис в ядрі $Ker\, f$ і позначимо його $\{v_{r+1}, \dots, v_{r+s}\}$, де $s = \mathrm{DIM}_{\mathbb{K}}Ker\, f$. Тоді $\mathcal{V} = \{v_1, \dots, v_r, v_{r+1}, \dots, v_{r+s}\}$ є базисом простору V (це було доведено раніше). Зокрема, $r + s = \mathrm{DIM}_{\mathbb{K}}V$. Тоді $\mathcal{V} = \{v_1, \dots, v_n\}$ і $\mathcal{W} = \{w_1, \dots, w_m\}$ - узгоджена пара базисів. А саме, $f(v_i) = w_i$, $1 \leqslant i \leqslant r$ - за конструкцією, і $f(v_i) = 0$, $i \geqslant r + 1$, тому що $v_i \in Ker\, f$. $\square$

Зауваження 6.3. *Така пара узгоджених базисів вибирається неоднозначно.*

Лемма 6.7. *Нехай $f : V \to W$ - довільне лінійне відображення скінченно вимірних просторів, $\mathcal{V}$, $\mathcal{W}$ - пара базисів, узгоджених з f. Розглянемо ізоморфізми*

$$f_V : V \longrightarrow \mathbb{K}^n, \qquad f_W : W \longrightarrow \mathbb{K}^m,$$

які вектору ставлять у відповідність його стовпчик координат у відповідному базисі. Нехай $r = \textsc{dim}_{\mathbb{K}}\textsc{Im}\ f$. Тоді має місце комутативна діаграма:

$$\begin{array}{ccc} V & \xrightarrow{f} & W \\ {\scriptstyle f_V}\big\downarrow & & \big\downarrow{\scriptstyle f_W} \\ \mathbb{K}^n & \xrightarrow{\mathrm{L}_{I_{m\times n,r}}} & \mathbb{K}^m \end{array}$$

Це можна переписати так: $f_W \cdot f = \mathrm{L}_{I_{m\times n,r}} \cdot f_V$.

Доведення. Рівність двох лінійних відображень достатньо перевірити на векторах з деякого базису. Нехай $\mathcal{V} = \{v_1, \ldots, v_n\}$, $\mathcal{W} = \{w_1, \ldots, w_m\}$. Перевірятимемо комутативність, визначаючи дії відображень діаграми на базисних векторах. Розглянемо довільний вектор v_i, $1 \leqslant i \leqslant r$. Маємо

$$\begin{array}{ccc} v_i & \xrightarrow{f} & w_i \\ {\scriptstyle f_V}\big\downarrow & & \big\downarrow{\scriptstyle f_W} \\ e_i \in \mathbb{K}^n & \xrightarrow{\mathrm{L}_{I_{m\times n,r}}} & f_i \in \mathbb{K}^m \end{array}$$

В цьому випадку діаграма буде комутативною. Правило множення матриці на вектор дає: $I_{m\times n,r} \cdot e_i = f_i$, $1 \leqslant i \leqslant r$.

Нехай $r + 1 \leqslant i \leqslant n$. Маємо діаграму

$$\begin{array}{ccc} v_i & \xrightarrow{f} & 0 \\ {\scriptstyle f_V}\big\downarrow & & \big\downarrow{\scriptstyle f_W} \\ e_i \in \mathbb{K}^n & \xrightarrow{\mathrm{L}_{I_{m\times n,r}}} & 0 \in \mathbb{K}^m \end{array}$$

Отже, діаграма комутативна, оскільки вона комутативна на кожному базисному векторі. $\square$

Доведення теореми про ранг. Нехай $V = \mathbb{K}^n$, $W = \mathbb{K}^m$ і $f = \mathrm{L}_A : \mathbb{K}^n \to \mathbb{K}^m$. За лемою 6.6, існує пара базисів $\mathcal{V}$ в просторі $\mathbb{K}^n$ та $\mathcal{W}$ в просторі $\mathbb{K}^m$, узгоджена з лінійним відображенням L_A. Тому, за лемою 6.7, існує відповідна комутативна діаграма, яка містить тільки координатні векторні простори. В цьому випадку існують ізоморфізм $f_V = \mathrm{L}_X : \mathbb{K}^n \to \mathbb{K}^n$ для деякої обертовної $n \times n$ матриці $X \in \mathrm{GL}_n(\mathbb{K})$ і ізоморфізм $f_W = \mathrm{L}_Y : \mathbb{K}^m \to \mathbb{K}^m$ для деякої обертовної $m \times m$ матриці $Y \in \mathrm{GL}_m(\mathbb{K})$. Діаграма тоді набуває вигляду

$$\begin{array}{ccc} \mathbb{K}^n & \xrightarrow{\mathrm{L}_A} & \mathbb{K}^m \\ {\scriptstyle \mathrm{L}_X}\big\downarrow & & \big\downarrow{\scriptstyle \mathrm{L}_Y} \\ \mathbb{K}^n & \xrightarrow{\mathrm{L}_{I_{m\times n,r}}} & \mathbb{K}^m \end{array}$$

де $r = \text{DIM}_{\mathbb{K}}\text{IM } \mathrm{L}_A$. За лемою 6.4, з комутативності діаграми випливає рівність

$$\text{DIM}_{\mathbb{K}}\text{IM } \mathrm{L}_A \; = \; \text{DIM}_{\mathbb{K}}\text{IM } \mathrm{L}_{I_{m\times n,r}} \; = \; \text{RANK}_c I_{m\times n,r} = r. \tag{10}$$

Транспонуємо всі матриці, які входять в цю діаграму, одержимо знову комутативну діаграму [5] вигляду

$$\begin{tikzcd}
\mathbb{K}^n & \mathbb{K}^m \arrow[l, "\mathrm{L}_{A^t}"'] \\
\mathbb{K}^n \arrow[u, "\mathrm{L}_{X^t}"] & \mathbb{K}^m \arrow[l, "\mathrm{L}_{I_{n\times m,r}}"] \arrow[u, "\mathrm{L}_{Y^t}"']
\end{tikzcd}
\quad \equiv \quad
\begin{tikzcd}
\mathbb{K}^m \arrow[r, "\mathrm{L}_{I_{n\times m,r}}"] \arrow[d, "\mathrm{L}_{Y^t}"'] & \mathbb{K}^n \arrow[d, "\mathrm{L}_{X^t}"] \\
\mathbb{K}^m \arrow[r, "\mathrm{L}_{A^t}"] & \mathbb{K}^n
\end{tikzcd}$$

Вертикальні стрілки в цій діаграмі - ізоморфізми. Дійсно, L_X - ізоморфізм $\Longrightarrow$ X - обертовна матриця $\Longrightarrow$ X^t - теж оборотна $\Longrightarrow$ L_{X^t} - ізоморфізм. Доведення для L_Y - аналогічно.

За лемою 6.4, маємо таку рівність

$$\text{DIM}_{\mathbb{K}}\text{IM } \mathrm{L}_{A^t} \; = \; \text{DIM}_{\mathbb{K}}\text{IM } \mathrm{L}_{I_{n\times m,r}} \; = \; \text{RANK}_c I_{n\times m,r} = r. \tag{11}$$

З рівностей (10) та (11) випливає рівність (9), а отже і твердження теореми.

Теорему про ранг матриці доведено.

Наслідок 6.2. *Для матриці $A \in Mat_{m\times n}(\mathbb{K})$ існують матриці $X \in \mathrm{GL}_n(\mathbb{K})$, $Y \in \mathrm{GL}_m(\mathbb{K})$ такі, що $YAX^{-1} = I_{m\times n,r}$, де $r = \text{RANK}_c A$.*

Доведення. Розглянемо побудовану діаграму. Маємо $\mathrm{L}_Y\mathrm{L}_A = \mathrm{L}_{I_{m\times n,r}}\mathrm{L}_X \Longrightarrow$ $YA = I_{m\times n,r}X \overset{\mathrm{L}_X-\text{ізом.}}{\Longrightarrow} YAX^{-1} = I_{m\times n,r}$ і $r = \text{RANK}_c A$. $\square$

Зауваження 6.4. *Операція транспонування не має простого аналога в теорії відображень.*

6.5 Елементарні матриці. Матричні одиниці

Позначимо через $e_{ij}^{m\times n} \in \mathrm{Mat}_{m\times n}(\mathbb{K})$, $1 \le i \le m$, $1 \le j \le n$, матрицю, у якої всі компоненти, крім $(i,j)-$тої, нульові, а $(i,j)-$та компонента рівна 1. Матриці $\{e_{ij}^{m\times n}\}_{1\le i\le m, 1\le j\le n}$ називаються матричними одиницями і утворюють стандартний базис в просторі матриць $\mathrm{Mat}_{m\times n}(\mathbb{K})$. Якщо $m = n$, пишемо просто e_{ij}^n (замість $e_{ij}^{m\times n}$). А коли m і n фіксовані, то пишуть просто e_{ij}.

Довільну матрицю $A = (a_{ij}) \in \mathrm{Mat}_{m\times n}(\mathbb{K})$ можна розкласти за базисом:

$A = \displaystyle\sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}e_{ij}$, де a_{ij} – коефіцієнти матриці A.

Правило множення матричних одиниць:

$$e_{ij}^{l\times m} \cdot e_{rs}^{m\times n} \; = \; \begin{cases} 0, & \text{коли} \quad j \ne r, \\ e_{is}^{l\times n}, & \text{коли} \quad j = r. \end{cases}$$

[5] Комутативність транспонованої діаграми випливає з комутативності вихідної діаграми і правила добутку транспонованих матриць: $(CD)^t = D^tC^t$.

Множення матриці на матричну одиницю.

Припустимо, що матриця A має вигляд $A \in \mathrm{Mat}_{m\times n}(\mathbb{K})$. Обчислимо

$$A \cdot e_{ij}^n \ = \ \sum_{k=1}^{m}\sum_{l=1}^{n} a_{kl}e_{kl} \cdot e_{ij}^n \ \overset{l=i}{=} \ \sum_{k=1}^{m} a_{ki}e_{kj}.$$

В одержаній матриці $i-$тий стовпчик записується на $j-$му місці, а решта стовпчиків нульові. Маємо: $A \cdot e_{ij}^n \ = \ \left(\ 0 \mid \ldots \mid \mathbf{a_i} \mid \ldots \mid \mathbf{0} \ \right)$.

Аналогічно, $e_{ij}^n \cdot A \ = \ i \begin{pmatrix} 0 \\ \ldots \\ \mathbf{a^j} \\ \ldots \\ 0 \end{pmatrix}$.

Перевірте самостійно.

Матриці вигляду $E_{ij}^n = I_n + \lambda e_{ij}^n$, $\lambda \in \mathbb{K}$, називаються **елементарними матрицями I роду**.

Матриці вигляду $E_i^n(\lambda) = \begin{pmatrix} 1 & \ldots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \ldots & \lambda & 0 \\ 0 & \ldots & 0 & 1 \end{pmatrix}$, де $\lambda \in \mathbb{K}^* = \mathbb{K} \setminus \{0\}$, називаються **елементарними матрицями II роду**. Отже, $E_i^n(\lambda)$ – це діагональна матриця, в якій на місці (i,i) стоїть λ, решта елементів рівні 0.

Факт 6.1. *Виконати елементарне перетворення над рядками (стовпчиками) матриці A – це те саме, що помножити матрицю A на деяку елементарну матрицю зліва (справа).*

Доведення полягає в перевірці. Дійсно,
$A \cdot E_{ij}^n(\lambda) = A(I_n + \lambda e_{ij}^n) = A = +\lambda A e_{ij}^n = \left(\ \mathbf{a^1} | \ldots | \mathbf{a^j} | \ldots | \mathbf{a^n} \ \right) + \lambda \left(\ 0 \mid \ldots \mid \mathbf{a^i} \mid \ldots \mid \mathbf{0} \ \right) =$

$$= \left(\ \mathbf{a^1} \mid \ldots \mid \mathbf{a^{j-1}} \mid \mathbf{a^j} + \lambda \mathbf{a^i} \mid \mathbf{a^{j+1}} \mid \ldots \mid \mathbf{a^n} \ \right).$$

$$A \cdot E_i^n(\lambda) = \left(\ \mathbf{a^1} \mid \ldots \mid \mathbf{a^{i-1}} \mid \lambda \mathbf{a^i} \mid \mathbf{a^{i+1}} \ldots \mid \mathbf{a^n} \ \right).$$

Відповідно, з іншого боку $E_{ij}^m(\lambda) \cdot A = \begin{pmatrix} \mathbf{a^1} \\ \hline \ldots \\ \hline \mathbf{a^{i-1}} \\ \hline \mathbf{a^i} + \lambda \mathbf{a^j} \\ \hline \mathbf{a^{i+1}} \\ \hline \ldots \\ \hline \mathbf{a^m} \end{pmatrix}$, $E_i^m(\lambda) \cdot A = \begin{pmatrix} \mathbf{a^1} \\ \hline \ldots \\ \hline \mathbf{a^{i-1}} \\ \hline \lambda \mathbf{a^i} \\ \hline \mathbf{a^{i+1}} \\ \hline \ldots \\ \hline \mathbf{a^m} \end{pmatrix}$.

Завдання 6.3. *Перевірте, що* $(E_{ij}^n(\lambda))^{-1} = E_{ij}^n(-\lambda)$, $(E_i^n(\lambda))^{-1} = E_i^n(\lambda^{-1})$.

Факт 6.2. *Перестановка рядків (стовпчиків) є композицією елементарних перетворень.*

Доведення. Розглянемо пару стовпчиків:

$$\left(\,\mathbf{a}\,|\,\mathbf{b}\,\right) \to \left(\,\mathbf{a}\,|\,\mathbf{b}+\mathbf{a}\,\right) \to \left(\,-\mathbf{b}\,|\,\mathbf{b}+\mathbf{a}\,\right) \to \left(\,-\mathbf{b}\,|\,\mathbf{a}\,\right) \to \left(\,\mathbf{b}\,|\,\mathbf{a}\,\right).$$

□

Твердження 6.5 (про елементарні перетворення). *1. Матрицю $A \in \mathrm{Mat}_{m\times n}(\mathbb{K})$ елементарними перетвореннями рядків і стовпчиків можна звести до вигляду $I_{n\times m,r}$.*

2. Для матриці A існують $r \leqslant \min\{n,m\}$ та елементарні матриці $X_1,\ldots,X_k \in \mathrm{GL}_m(\mathbb{K})$ і матриці $Y_1,\ldots,Y_l \in \mathrm{GL}_n(\mathbb{K})$ такі, що $X_k\ldots X_1AY_1\ldots Y_l = I_{m\times n,r}$.

3. Матриця A буде обертовною тоді і лише тоді, коли $n = m = r$.

Доведення.

1. Розглянемо $A \in \mathrm{Mat}_{m\times n}(\mathbb{K})$.

   1) Якщо $A = 0$, то $A = I_{m\times n,0}$.

   2) В матриці A є ненульовий елемент. Перестановкою рядків і стовпчиків переміщуємо його на місце $(1,1)$. Робимо його одиницею. Додаванням рядків та стовпчиків одержимо матрицю

$$A' = \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & A' \end{array}\right).$$

   Далі застосовуємо цю процедуру до матриці A'.

2. Це наслідок з першого пункту, оскільки кожне елементарне перетворення є множенням на елементарну матрицю.

3. Якщо A – обертовна, то A – квадратна, тобто $m = n$. З другого пункту одержимо, що $X_k\ldots X_1AY_1\ldots Y_l = I_{n\times n,r}$, де $X_1,\ldots,X_k,Y_1,\ldots,Y_l \in \mathrm{GL}_n(\mathbb{K})$ - елементарні матриці. Оскільки елементарні матриці обертовні і добуток обертовних матриць є знову обертовною матрицею, то одержимо, що $I_{m\times n,r}$ обертовна, звідки $m = n = r$.

□

7 Детермінант (визначник) (полілінійна алгебра)

7.1 Поняття групи

Нехай G – це деяка множина. *Бінарна операція* на G – це деяке відображення

$$\begin{array}{ccc} G \times G & \longrightarrow & G \\ (g_1, g_2) & \mapsto & m(g_1, g_2) \end{array}$$

В різних ситуаціях операції позначаються по-різному, наприклад, як $g_1 \cdot g_2$, $g_1\, g_2$, $g_1 * g_2$ або $g_1 + g_2$.

Означення 7.1. *Множина G з визначеною на ній бінарною операцією*

$$\begin{array}{ccccc} G \times G & \longrightarrow & G & & \\ (g_1, g_2) & \mapsto & m(g_1, g_2) & = g_1 \cdot g_2 & = g_1\, g_2 \end{array}$$

*називається **групою**, якщо виконуються наступні властивості:*

1. $\forall\ g_1, g_2, g_3 \in G \quad (g_1 g_2) g_3 = g_1 (g_2 g_3)$ *– властивість асоціативності;*

2. $\exists\ e \in G\ :\ \forall g \in G\ ge = eg = g$ *– існування нейтрального елемента;*

3. $\forall g \in G\ \exists g^{-1} \in G\ :\ gg^{-1} = g^{-1} g = e$ *– існування оберненого елемента.*

Якщо виконується ще четверта властивість:
4) $\forall g_1, g_2 \in G \quad g_1 g_2 = g_2 g_1$,
то говорять, що група G є комутативною або абелевою.

7.2 Приклади груп

1. Нехай $\mathbb{K}$ – поле. Розглянемо $\mathbb{K}^* = \mathbb{K} \setminus \{0\}$. Тоді $\mathbb{K}^*$ – група відносно операції множення (при цьому $e = 1 \in \mathbb{K}^*$).

2. $\mathbb{K}$ – поле, $n \geq 1$. Множина обертовних матриць $\mathrm{GL}_n(\mathbb{K})$ утворює групу відносно множення матриць ($e = I_n$, X^{-1} – обернена до $X \in \mathrm{GL}_n(\mathbb{K})$).

3. V – векторний простір над $\mathbb{K}$. Тоді ізоморфізми $f : V \to V$ утворюють групу $\mathrm{GL}(V)$ (операція fg для $f, g \in \mathrm{GL}(V)$ – це композиція $V \xrightarrow{g} V \xrightarrow{f} V$, нейтральний елемент $e = \mathtt{1}_V$, обернений – $f^{-1} : V \to V$).

4. Нехай X – довільна множина. Бієктивні відображення $f : X \to X$ утворюють групу (повну симетричну групу на множині X) $S(X)$, яка називається групою підстановок на множині X.

Операція в групі $S(X)$: $f, g \in S(X)$, тоді $fg : X \xrightarrow{g} X \xrightarrow{f} X$ – бієкція, тобто, $fg \in S(X)$; нейтральний елемент $e = \mathtt{1}_X$ і обернений до f – $f^{-1} : X \to X$ – теж бієкція.

5. На полі $\mathbb{K}$ визначена бінарна операція додавання

$$\begin{array}{ccc} \mathbb{K}\times\mathbb{K} & \longrightarrow & \mathbb{K} \\ (x,y) & \mapsto & x+y \end{array}$$

Множина елементів $\mathbb{K}$ утворює групу, яка називається адитивною групою поля $\mathbb{K}$. Нейтральний елемент відносно додавання – 0, обернений до $a \in \mathbb{K}$ – це протилежний, тобто $-a$.

6. Векторний простір V відносно операції додавання векторів утворює групу. Операція $v_1 + v_2$, нейтральний – 0, обернений (протилежний) до $v \in V$ – це $-v \in V$.

Зауваження 7.1. *В прикладах 1, 5, 6 групи абелеві.*

7.3 Деякі властивості груп

1. Нейтральний елемент (або одиниця групи, якщо операція записується у вигляді $g_1 \cdot g_2$ і називається множенням) визначений однозначно.

   Доведення. Якщо $e, e' \in G$ – два нейтральних елементи, то

$$e \cdot e' = \left\{ \begin{array}{llll} e, & \text{бо} & e' & - \text{ нейтральний} \\ e', & \text{бо} & e & - \text{ нейтральний} \end{array} \right. \Longrightarrow \quad e = e'.$$

   □

2. Обернений визначений однозначно. Дійсно, припустимо, що існують два обернених $g^{-1}, g^{-1'}$ до $g \in G$. Тоді

$$g^{-1} = e \cdot g^{-1} = (g^{-1'} \cdot g)g^{-1} \overset{\text{асоц.}}{=} g^{-1'}(g \cdot g^{-1}) = g^{-1'} \cdot e = g^{-1'} \implies g^{-1} = g^{-1'}.$$

3. $(g^{-1})^{-1} = g$, тому що $(g^{-1})^{-1} \cdot g^{-1} = g^{-1} \cdot (g^{-1})^{-1} = e$, але g таким рівностям задовольняє. Оскільки обернений $(g^{-1})^{-1}$ визначений однозначно, то звідси випливає, що $(g^{-1})^{-1} = g$.

4. $e^{-1} = e$.

5. $(g_1 \cdot g_2)^{-1} = g_2^{-1} \cdot g_1^{-1}$.

   Дійсно, $(g_1 g_2)(g_2^{-1} g_1^{-1}) \overset{\text{асоц.}}{=} ((g_1 g_2)g_2^{-1})g_1^{-1} = g_1(g_2 g_2^{-1})g_1^{-1} = g_1 \cdot g_1^{-1} = e$.

   Аналогічно, $(g_2^{-1} g_1^{-1})(g_1 g_2) = e$. Звідси, оскільки обернений визначений однозначно, одержимо: $(g_1 g_2)^{-1} = g_2^{-1} g_1^{-1}$.

6. Нехай $g_1, g_2, \ldots, g_n$ – деяка послідовність елементів G. При довільно розставлених дужках між $g_1, g_2, \ldots, g_n$ (без зміни їх порядку) і обчислення добутку завжди одержимо один і той же елемент групи G.

Для $n = 2$ існує єдиний спосіб розстановки дужок: $g_1 \cdot g_2 = (g_1 \cdot g_2)$. Для трьох елементів g_1, g_2, g_3 множення можна влаштувати двома способами: $g_1(g_2g_3)$ та $(g_1g_2)g_3$. Ці добутки будуть рівними за асоціативністю.

Розглянемо випадок $n = 4$. Маємо такі варіанти розстановки дужок:

$$(g_1 \cdot g_2)\cdot(g_3 \cdot g_4), \quad ((g_1 \cdot g_2)\cdot g_3)\cdot g_4, \quad (g_1 \cdot (g_2 \cdot g_3))\cdot g_4, \quad g_1 \cdot ((g_2 \cdot g_3)\cdot g_4), \quad g_1 \cdot (g_2 \cdot (g_3 \cdot g_4)).$$

Стверджується, що усі ці добутки рівні. Доведення дамо в загальному випадку.

Твердження 7.1. *За довільної розстановки дужок при множенні елементів $g_1, g_2, \dots, g_n$ групи G результат дії буде однаковим.*

Зауважимо, що, завдяки цьому твердженню, дужки в добутках елементів можна узагалі не розставляти. Єдиний добуток елементів $g_1, g_2, \dots, g_n$ групи G, взятих у фіксованому порядку, позначатимемо $g_1 \cdot g_2 \cdot \ldots \cdot g_n \in G$.

Доведення. Індукцією за n доведемо, що за довільної розстановки дужок одержимо добуток $g_1 \cdot g_2 \cdot \ldots \cdot g_n \in G$.

База індукції: $n = 3$. Є дві розстановки, які рівні, за аксіомою асоціативності. Позначимо $g_1 \cdot g_2 \cdot g_3 \overset{df}{=} (g_1 \cdot g_2) \cdot g_3 = g_1 \cdot (g_2 \cdot g_3)$.

Крок індукції: $(n-1) \to (n)$. Розглянемо довільну розстановку дужок між елементами $g_1, g_2, \dots, g_n$. Якась операція є останньою, тобто маємо $(\dots)\cdot(\dots)$. В кожній дужці менше, ніж n співмножників. Тоді перша дужка дорівнює, за припущенням індукції, $(g_1 \cdot g_2 \cdot \ldots \cdot g_i)$, а друга дужка – $(g_{i+1} \cdot \ldots g_{n-1} \cdot g_n)$ для деякого $1 \leq i < n$. Отже, наш довільний добуток дорівнює $(g_1 \cdot g_2 \cdot \ldots \cdot g_i)(g_{i+1} \cdot \ldots \cdot g_n) = (g_1 \cdot (g_2 \cdot \ldots \cdot g_i))(g_{i+1} \cdot \ldots \cdot g_n) \overset{\text{асоц.}}{=} g_1 \cdot ((g_2 \cdot \ldots \cdot g_i)(g_{i+1} \cdot \ldots \cdot g_n)) = g_1 \cdot (g_2 \cdot \ldots \cdot g_n)$ і не залежить від початкової розстановки дужок. □

Завдання 7.1. *Покажіть, що* $(g_1 \cdot g_2 \cdot \ldots \cdot g_n)^{-1} = (g_n^{-1} \cdot \ldots \cdot g_2^{-1} \cdot g_1^{-1})$.

7.4 Група підстановок – повна симетрична група

Якщо X – множина, підстановка на X – це деяка бієкція $X \to X$. Покажемо, що усі підстановки на множині X утворюють групу $S(X) = S_X$.

Факт 7.1. *Нехай* $|X| = n$. *Тоді* $|S_X| = n!$.

Доведення. Нехай $X = \{x_1, x_2, \dots, x_n\}$. Довільна підстановка $\pi \in S_X$ визначається за такою процедурою:

$$\begin{array}{lllcll}
x_1 & \mapsto & \pi(x_1) \in X & - & n & \text{варіантів,} \\
x_2 & \mapsto & \pi(x_2) \in X \setminus \{\pi(x_1)\} & - & (n-1) & \text{варіант,} \\
x_3 & \mapsto & \pi(x_3) \in X \setminus \{\pi(x_1), \pi(x_2)\} & - & (n-2) & \text{варіанти,} \\
\dots & \dots & \dots & & \dots & \dots \\
x_n & \mapsto & \pi(x_n) \in X \setminus \{\pi(x_1), \dots, \pi(x_{n-1})\} & - & 1 & \text{варіант,}
\end{array}$$

Усі ці варіанти незалежні, а їх загальна кількість дорівнює $n(n-1)(n-2)\dots 1 = n!$.
□

Означення 7.2. *Нехай $a, b \in X$, $a \neq b$. Означимо підстановку $\sigma_{ab} : X \to X$ наступним чином:*

$$\sigma_{ab}(a) = b,\ \sigma_{ab}(b) = a,\ \sigma_{ab}(x) = x, x \neq a, b.$$

Підстановки вигляду σ_{ab} називаються ***транспозиціями***.

Підстановку на скінченній множині можна записувати у вигляді таблиці. Якщо $X = \{x_1, x_2, \ldots, x_n\}$, то

$$\pi \ = \ \begin{pmatrix} x_1 & x_2 & \ldots & x_n \\ \pi(x_1) & \pi(x_2) & \ldots & \pi(x_n) \end{pmatrix} \in S_X.$$

Множення підстановок визначається як їх композиція. Нехай τ – ще одна підстановка з S_X:

$$\tau \ = \ \begin{pmatrix} x_1 & x_2 & \ldots & x_n \\ \tau(x_1) & \tau(x_2) & \ldots & \tau(x_n) \end{pmatrix}.$$

Тоді

$$\tau \cdot \pi \ = \ \begin{pmatrix} x_1 & x_2 & \ldots & x_n \\ \pi(x_1) & \pi(x_2) & \ldots & \pi(x_n) \\ \tau\pi(x_1) & \tau\pi(x_2) & \ldots & \tau\pi(x_n) \end{pmatrix} = \begin{pmatrix} x_1 & x_2 & \ldots & x_n \\ \tau\pi(x_1) & \tau\pi(x_2) & \ldots & \tau\pi(x_n) \end{pmatrix}.$$

Одинична підстановка (нейтральний елемент групи) $e \in S_X$ визначається тотожним відображенням $\mathtt{1}_X$, тобто

$$1 \ = \ e \ = \ \begin{pmatrix} x_1 & x_2 & \ldots & x_n \\ x_1 & x_2 & \ldots & x_n \end{pmatrix}.$$

Обернена підстановка $\pi^{-1} : X \to X$ обчислюється за таким правилом:

$$\pi^{-1} \ = \ \begin{pmatrix} \pi(x_1) & \pi(x_2) & \ldots & \pi(x_n) \\ x_1 & x_2 & \ldots & x_n \end{pmatrix}.$$

Після цього можна впорядкувати перший рядок.

Застереження 7.1. *Таблиця*

$$\pi^{-1} \ = \ \begin{pmatrix} x_1 & x_2 & \ldots & x_n \\ y_1 & y_2 & \ldots & y_n \end{pmatrix}$$

тоді і лише тоді визначає деяку підстановку з S_X, коли $\{y_1, y_2, \ldots, y_n\} = X$.

Якщо множина $X = \{1, 2, \ldots, n\} \subset \mathbb{Z}$ – інтервал натурального ряду, то в цьому випадку група S_X позначається S_n і називається *повною симетричною групою* на n елементах. В цьому випадку кожна підстановка $\pi \ = \ \begin{pmatrix} x_1 & x_2 & \ldots & x_n \\ \pi(x_1) & \pi(x_2) & \ldots & \pi(x_n) \end{pmatrix}$ однозначно визначена своїм другим рядком $(i_1, i_2, \ldots, i_n)$ де $i_k = \pi(k)$, $k = 1, \ldots, n$.

Позначимо через $\rho(\pi) = (i_1, i_2, \dots, i_n)$ другий рядок підстановки π. Тоді $\rho(\pi)$ утворює перестановку множини $\{1, 2, \dots, n\}$.

Транспозиції – це елементи $\sigma_{ij} \in S_n$, $1 \le i \ne j \le n$, для яких $\sigma_{ij}(i) = j$, $\sigma_{ij}(j) = i$ і $\sigma_{ij}(k) = k$ для довільного $k \in \{1, 2, \dots, n\} \setminus \{i, j\}$. При цьому a) $\sigma_{ij} = \sigma_{ji}$; b) $\sigma_{ij} \cdot \sigma_{ij} = \sigma_{ij}^2 = 1$.

Елементарною транспозицією $\sigma_i \in S_n$, $i = 1, \dots, n-1$, називається транспозиція вигляду $\sigma_i \overset{df}{=} \sigma_{i,i+1}$ (тобто $\sigma_1 = \sigma_{12}$, $\sigma_2 = \sigma_{23}$, ..., $\sigma_{n-1} = \sigma_{n-1,n}$).

Необхідні приклади груп.

$C_2 = \{1, -1\} \subset \mathbb{R}^*$ – група квадратних коренів з одиниці.

$\mathbb{Z}_2 = \{\bar{0}, \bar{1}\}$ – група лишків за модулем 2. (в $\mathbb{Z}_2$ маємо: $\bar{0} + \bar{0} = \bar{0}$, $\bar{0} + \bar{1} = \bar{1} + \bar{0} = \bar{1}$, $\bar{1} + \bar{1} = \bar{0}$).

Означення 7.3. *Нехай $(i_1, \dots, i_n)$ деяка перестановка на множині $\{i_1, \dots, i_n\}$. Говоримо, що i_k, i_l утворюють* ***інверсію****, якщо $k < l$, $i_k > i_l$.*

Через $T(i_1, \dots, i_n)$ позначимо кількість інверсій в перестановці $\rho(i_1, \dots, i_n)$. Якщо $\pi = \begin{pmatrix} 1 & 2 & \dots & n \\ i_1 & i_2 & \dots & i_n \end{pmatrix} \in S_n$, то, за означенням, $T(\pi) = T(\rho(\pi)) = T(i_1, \dots, i_n)$.

Приклад. Для $n = 3$ маємо $T(1, 2, 3) = 0$, $T(3, 1, 2) = 2$, $T(3, 2, 1) = 3$.

Лемма 7.1. *Якщо для деякої підстановки $\pi \in S_n$ $T(\pi) = 0$, то $\pi = 1$ – одинична підстановка.*

Доведення. Для випадку $n = 1$ доведення очевидне. Вважаємо $n > 1$. Розглянемо підстановку $\pi = \begin{pmatrix} 1 & 2 & \dots & n \\ i_1 & i_2 & \dots & i_n \end{pmatrix}$. Серед елементів $i_1, \dots, i_n$ є одиниця. Якщо $i_1 \ne 1$, то існує принаймні одна інверсія $(i_1, 1 = i_k)$, $k > 1$. Звідси одержимо, що $\pi = (1, i_2, \dots, i_n)$. Де знаходиться 2 (у випадку $n > 2$)? Якщо $i_2 \ne 2$, то $(i_2, 2 = i_k)$ – інверсія. Отже$\pi = (1, 2, i_3, \dots, i_n)$.

Повторюючи це міркування, одержимо твердження. $\square$

Лемма 7.2. *Довільна підстановка $\pi \in S_n$ є добутком елементарних транспозицій: $\pi = \sigma_{k_t} \dots \sigma_{k_1}$, $t \geqslant 0$, де $\sigma_{k_1}, \dots, \sigma_{k_t}$ – деякі транспозиції. Якщо t – мінімальна кількість співмножників в такому розкладі, то $t = T(\pi)$. Якщо $\pi = \sigma_{k_t} \dots \sigma_{k_1} = \sigma_{l_s} \dots \sigma_{l_1}$ – два таких добутки, то $t - s$ – парне число.*

Зауваження 7.2. *Запис $\pi = \sigma_{k_1} \dots \sigma_{k_t}$ у випадку $t = 0$ означає, що $\pi = 1$ – одинична підстановка.*

Доведення. Розглянемо довільну підстановку $\pi = \begin{pmatrix} 1 & 2 & \dots & n \\ t_1 & t_2 & \dots & t_n \end{pmatrix}$ і відповідний їй рядок $(t_1, t_2, \dots, t_n)$. Нехай σ_{ij} – деяка транспозиція. Запишемо добуток підстановок $\pi\sigma_{ij}$:

$$\begin{pmatrix} 1 & \dots & i & \dots & j & \dots & n \\ 1 & \dots & j & \dots & i & \dots & n \\ t_1 & \dots & t_j & \dots & t_i & \dots & t_n \end{pmatrix}.$$

Отже, другий рядок підстановки $\pi\sigma_{ij}$ одержується з другого рядка підстановки π перестановкою $i-$того та $j-$того елементів.

Доведемо, що для довільної підстановки $\pi \in S_n$, для якої $T(\pi) > 0$, існує такий розклад $\pi = \sigma_{k_t} \dots \sigma_{k_1}$, що $t = T(\pi)$. Для цього покажемо, що існує така елементарна транспозиція $\sigma_k = \sigma_{k,k+1}$, для якої $T(\pi\sigma_{k_1}) = T(\pi) - 1$.

Дійсно, у випадку $T(\pi) > 0$ існує $k \in \{1, \dots, n-1\}$, таке що $t_k > t_{k+1}$. Розглянемо тоді підстановку $\pi\sigma_k$. Другий рядок цієї підстановки дорівнює

$$\rho(\pi\sigma_k) = (t_1, \dots, t_{k-1}, t_{k+1}, t_k, t_{k+2}, \dots, t_n).$$

Кількість інверсій змінилася, а саме: $T(\pi\sigma_k) = T(\pi) - 1$. Це тому, що інверсія, яку утворювала пара t_k, t_{k+1}, зникла, а решта – збереглися.

Позначимо $k_1 = k$. Підсумуємо: ми знайшли k_1 таке, що $T(\pi\sigma_{k_1}) = T(\pi) - 1$. Якщо $T(\pi) - 1 = 0$, то ми зупиняємося. Інакше знаходимо k_2, таке що $T(\pi\sigma_{k_1}\sigma_{k_2}) = T(\pi\sigma_{k_1}) - 1 = T(\pi) - 2$. Через деяку кількість кроків одержимо підстановку $\pi\sigma_{k_1} \dots \sigma_{k_r}$ таку, що $T(\pi\sigma_{k_1} \dots \sigma_{k_r}) = 0$, тобто, за лемою 7.1, $\pi\sigma_{k_1} \dots \sigma_{k_r} = 1$. Звідси одержуємо, що $\pi = (\sigma_{k_1} \dots \sigma_{k_r})^{-1} = (\sigma_{k_r})^{-1} \dots (\sigma_{k_1})^{-1} = (\sigma_{k_r}) \dots \sigma_{k_1})$ (тому що $\sigma_{k_i}^2 = 1$ і $\sigma_{k_i}^{-1} = \sigma_{k_i}$).

Ми одержали розклад π в добуток елементарних транспозицій. Оскільки $0 = T(\pi\sigma_{k_1} \dots \sigma_{k_r}) = T(\pi) - r$ (за побудовою), то звідси $r = T(\pi)$, і отже, довжина розкладу дорівнює кількості інверсій в π.

З іншого боку, якщо τ – довільна підстановка і $\pi(\tau) = (t_1 \dots t_k)$ – другий рядок підстановки τ, тоді

$$T(\tau\sigma_i) = \begin{cases} T(\tau) + 1 & \text{якщо} \quad t_i < t_{i+1}, \\ T(\tau) - 1 & \text{якщо} \quad t_i > t_{i+1} \end{cases}$$

(це випливає з того, що $\pi(\tau\sigma_i) = (t_1 \dots t_{i+1} t_i \dots t_k)$).

Припустимо, що маємо довільний розклад π в добуток елементарних транспозицій: $\pi = \sigma_{l_s} \dots \sigma_{l_1}$. Звідси випливає, що $\pi\sigma_{l_1} \dots \sigma_{l_s} = 1$. Числа $T(\pi\sigma_{l_1} \dots \sigma_{i+1})$ і $T(\pi\sigma_{l_1} \dots \sigma_{i+1}\sigma_i)$ відрізняються між собою на 1 або -1. Звідси випливає, що ці числа мають різну парність. Тому

1) $s \geq T(\pi)$ (тому що при множенні на σ_i функція T від підстановки змінюється щонайбільше на 1);

2) парність цілих чисел $T(\pi)$ і s є однією і тією ж, тому що парності $T(\pi\sigma_{l_1} \dots \sigma_{i+1})$ і $T(\pi\sigma_{l_1} \dots \sigma_{i+1}\sigma_i)$ – різні, а число $T(\pi\sigma_{l_1} \dots \sigma_1)$ є парним. $\square$

Нагадаємо, що через $C_2 = \{1, -1\}$ ми позначаємо групу квадратних коренів з одиниці, а через $\mathbb{Z}_2 = \{\bar{0}, \bar{1}\}$ – групу лишків за модулем 2.

Означення 7.4. *Підстановка $\pi \in S_n$ називається **парною**, якщо виконується одна з двох умов:*

a) кількість інверсій $T(\pi) \in 2\mathbb{Z}$ – парна;

b) якщо $\pi = \sigma_1 \dots \sigma_t$ – деякий розклад в добуток елементарних транспозицій, то t – парне. Підстановка, що не є парною, називається непарною.

Розглянемо два відображення:

$$p : \begin{array}{ccc} S_n & \longrightarrow & \mathbb{Z}_2 \\ \pi & \mapsto & p(\pi) \end{array}, \quad \text{де} \quad p(\pi) = \begin{cases} \bar{0} & \text{якщо} \ \pi \ - \ \text{парна}, \\ \bar{1} & \text{якщо} \ \pi \ - \ \text{непарна}; \end{cases}$$

$$\begin{array}{ccccc} \varepsilon: & S_n & \longrightarrow & C_2 \\ & \pi & \mapsto & \varepsilon(\pi) \end{array}, \quad \text{де} \quad \varepsilon(\pi) = \begin{cases} 1 & \text{якщо } \pi - \text{парна}, \\ -1 & \text{якщо } \pi - \text{непарна}. \end{cases}$$

Парністю *підстановки* $\pi \in S_n$ *називається елемент* $p(\pi) \in \mathbb{Z}_2$. ***Знаком*** *підстановки* π *називається називається число* $\varepsilon(\pi)$.

Зауваження 7.1. $\varepsilon(\pi) = (-1)^{T(\pi)}$.

Лемма 7.3. *Відображення* $\varepsilon : S_n \to C_2$ *та* $p : S_n \to \mathbb{Z}_2$ *мають такі властивості:*
1) $\varepsilon(\pi\sigma) = \varepsilon(\pi)\varepsilon(\sigma)$*;*
2) $p(\pi\sigma) = p(\pi) + p(\sigma)$.

Доведення. Ці твердження еквівалентні. Доведемо перше. Розглянемо підстановки $\pi, \sigma \in S_n$ і розкладемо їх в добутки елементарних транспозицій:

$$\pi = \sigma_{k_t} \dots \sigma_{k_1}; \quad \sigma = \sigma_{l_s} \dots \sigma_{l_1}.$$

Тоді $\varepsilon(\pi) = (-1)^t$, $\varepsilon(\sigma) = (-1)^l$. Існує розклад $\pi\sigma = (\sigma_{k_t} \dots \sigma_{k_1}) \quad (\sigma_{l_s} \dots \sigma_{l_1}) = \sigma_{k_t} \dots \sigma_{k_1}\sigma_{l_s} \dots \sigma_{l_1}$ в добуток елементарних транспозиції довжини $t + l$, тому $\varepsilon(\pi\sigma) = (-1)^{t+l} = (-1)^t(-1)^l = \varepsilon(\pi)\varepsilon(\sigma)$.

Друге твердження доводиться аналогічно. Доведіть самостійно. □

Наведемо табличку парності добутку двох підстановок:

	п	н
п	п	н
н	н	п

Лемма 7.4. *Нехай* $A_n \subset S_n$, $n \geq 2$ *– множина усіх парних підстановок. Тоді*

1. A_n *утворює групу (відносно стандартних операцій на підстановках).*
2. *Нехай* s *– довільна непарна підстановка. Тоді* sA_n *є множиною всі непарних підстановок.*
3. $|A_n| = |sA_n| = \dfrac{n!}{2}$.

Доведення.

1. Якщо $g_1, g_2 \in A_n$, то $g_1 \cdot g_2 \in A_n$, оскільки добуток парних підстановок є парною підстановкою. Одинична підстановка 1 належить до A_n, тому що $T(1) = 0$. Якщо $\pi \in A_n$, то $\pi^{-1} \in A_n$. Покажемо це. Нехай $\pi = \sigma_1\sigma_2 \dots \sigma_t$ - розклад в добуток елементарних транспозицій. Тоді $\pi^{-1} = \sigma_t^{-1} \dots \sigma_2^{-1}\sigma_1^{-1} = \sigma_t \dots \sigma_2\sigma_1$. Тому $\varepsilon(\pi^{-1}) = (-1)^t = \varepsilon(\pi)$, а отже $\pi^{-1} \in A_n$.

2. Розглянемо відображення: $\begin{array}{cccc} \mathrm{s}: & A_n & \longrightarrow & sA_n \\ & \pi & \mapsto & s\pi \end{array}$

   За означенням, s є сюр'єкція. Перевіримо, що s є ін'єкцією, а отже і бієкцією. Дійсно, $\mathrm{s}(\pi_1) = \mathrm{s}(\pi_2) \iff s\pi_1 = s\pi_2 \iff s^{-1}(s\pi_1) = s^{-1}(s\pi_2) \iff \pi_1 = \pi_2$. Тому існує обернене відображення $\mathrm{s}^{-1} : sA_n \longrightarrow A_n$. Нехай τ – довільна непарна підстановка, тоді $\tau = \mathrm{s}(\mathrm{s}^{-1}\tau) \in sA_n$, оскільки $\mathrm{s}^{-1}\tau$ – парна.

3. Згідно з визначеннями, $S_n = A_n \cup \mathrm{s}A_n$, $A_n \cap \mathrm{s}A_n = \emptyset$. Тоді, оскільки s – бієкція, то $|A_n| = |\mathrm{s}A_n| = \dfrac{|S_n|}{2} = \dfrac{n!}{2}$.

□

Лемма 7.5. *Кожна транспозиція $\sigma_{ij} \in S_n$ є непарною підстановкою.*

Доведення. Припустимо $i < j$. Тоді другий рядок підстановки σ_{ij} дорівнює

$$\rho(\sigma_{ij}) = (1 \dots i-1 \overbrace{j}^{i} i+1 \dots j-1 \overbrace{i}^{j} j+1 \dots n).$$

Будемо домножувати σ_{ij} справа на елементарні транспозиції, щоб одержати 1. Переставимо послідовно j з $i+1$, $i+2$, ..., $j-1$, i. При цьому нам слід виконати $(j-i-1)+1$ елементарних транспозицій. Одержимо підстановку, другий рядок якої дорівнює:

$$(1 \dots i-1\, i+1 \dots j-1 \overbrace{i}^{j-1} \overbrace{j}^{j} j+1 \dots n).$$

Після цього переставимо послідовно i з $j-1$, $j-2$, ..., $i+1$ і одержимо одиничну підстановку. При цьому нам слід виконати $(j-i-1)$ елементарних транспозицій.

Загалом потрібно домножити σ_{ij} на $(j-i-1)+1+(j-i-1)$ елементарних транпозицій, щоб одержати одиничну підстановку. Звідси: $\varepsilon(\sigma_{ij}) = (-1)^{2(j-i-1)+1} = -1$.
□

Зауваження 7.2. *Якщо τ – підстановка, $\pi = \sigma_1 \dots \sigma_t$ – розклад в добуток транспозицій (не обов'язково елементарних), то парність t і $T(\pi)$ збігаються.*

Властивість $\varepsilon(\pi) = (-1)^t$ дає змогу визначити парність підстановки в довільній групі S_X, $|X| < \infty$.

7.5 Детермінант

Означення 7.5. *Нехай V, W – векторні простори над полем $\mathbb{K}$. Функція*

$$f : \underbrace{V \times \dots \times V}_{k} \longrightarrow W$$

називається ***полілінійною*** *над полем $\mathbb{K}$, якщо виконуються такі дві властивості:*

1. $f(v_1, \ldots, \underbrace{v_i + v_i^{'}}_{i}, \ldots, v_k) = f(v_1, \ldots, v_i, \ldots, v_k) + f(v_1, \ldots, v_i^{'}, \ldots, v_k)$, $i = 1, \ldots, k$;

2. $f(v_1, \ldots, \lambda v_i, \ldots, v_k) = \lambda f(v_1, \ldots, v_i, \ldots, v_k)$, $\lambda \in \mathbb{K}$, $i = 1, \ldots, k$.

Лемма 7.6. *Полілінійні функції* $\underbrace{V \times \ldots \times V}_{k} \to W$ *утворюють векторний простір відносно операцій*

$$(f+g)(v_1, \ldots, v_k) = f(v_1, \ldots, v_k) + g(v_1, \ldots, v_k);$$

$$(\lambda f)(v_1, \ldots, v_k) = \lambda(f(v_1, \ldots, v_k)).$$

Довести самостійно.

Означення 7.6. *Функція* $\underbrace{V \times \ldots \times V}_{k} \longrightarrow W$ *називається* ***кососиметричною***, *якщо* $f(\ldots, \underbrace{v}_{i}, \ldots, \underbrace{v}_{j}, \ldots) = 0$, *тобто якщо за наявності двох однакових компонент значення функції дорівнює нулю.*

Доведемо деякі властивості полілінійних та кососиметричних функцій.

1. Якщо $f : \underbrace{V \times \ldots \times V}_{k} \to W$ полілінійна та кососиметрична функція, то виконується рівність: $f(\ldots, \underbrace{x}_{i}, \ldots, \underbrace{y}_{j}, \ldots) = -f(\ldots, \underbrace{y}_{i}, \ldots, \underbrace{x}_{j}, \ldots)$.

   Це означає, що при перестановці аргументів в цих двох компонентах значення функції змінюється на протилежне.

   Доведення. $0 = f(\ldots, x+y, \ldots, x+y, \ldots) = f(\ldots, x, \ldots, x+y, \ldots) + f(\ldots, y, \ldots, x+y, \ldots) = f(\ldots, x, \ldots, x, \ldots) + f(\ldots, x, \ldots, y, \ldots) + f(\ldots, y, \ldots, x, \ldots) + f(\ldots, y, \ldots, y, \ldots) = f(\ldots, x, \ldots, y, \ldots) + f(\ldots, y, \ldots, x, \ldots) \Longrightarrow f(\ldots, x, \ldots, y, \ldots) = -f(\ldots, y, \ldots, x, \ldots)$.
   $\square$

2. Полілінійна та кососиметрична функція $f : \underbrace{V \times \ldots \times V}_{k} \to W$ не змінюється при елементарних перетвореннях 1 роду.

   Доведення. $f(v_1, \ldots, v_i, \ldots, v_j + \lambda v_i, \ldots, v_k) = f(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_k) + f(v_1, \ldots, v_i, \ldots, \lambda v_i, \ldots, v_k) = f(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_k) + \lambda f(v_1, \ldots, v_i, \ldots, v_i, \ldots, v_k) = f(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_k)$. $\square$

3. Якщо дві компоненти пропорційні, то значення полілінійної та кососиметричної функції $f : \underbrace{V \times \ldots \times V}_{k} \to W$ рівне нулю: $f(\ldots, v, \ldots, \lambda v, \ldots) = 0$.

Теорема 7.1. *Нехай $\mathbb{K}$ – поле, V – скінченно вимірний простір, $\mathrm{DIM}_{\mathbb{K}} V = n$. Тоді простір полілінійних кососиметричних відображень $f : V^n \to \mathbb{K}$, де $V^n = \underbrace{V \times \ldots \times V}_{n}$, є одновимірним (тобто для двох ненульових $f, f' : V^n \to \mathbb{K}$ існує $\lambda \in \mathbb{K}$ такий, що $f' = \lambda f$.*

Доведення. Фіксуємо базис $\mathcal{V} = \{v_1 \ldots, v_n\} \subset V$. Розглянемо довільні вектори $\alpha_1, \alpha_2, \ldots, \alpha_n \in V$. Розкладемо їх за базисом $\mathcal{V}$: $\alpha_j = \sum\limits_{i=1}^{n} a_{ij} v_i$. Обчислимо

$$f(\alpha_1, \ldots, \alpha_n) = f\Big(\sum_{i=1}^{n} a_{i1} v_i, \sum_{i=1}^{n} a_{i2} v_i, \ldots, \sum_{i=1}^{n} a_{in} v_i\Big) \overset{\text{полілін.}}{=} \sum_{i=1}^{n} f(a_{i1} v_i, \sum_{i=1}^{n} a_{i2} v_i, \ldots, \sum_{i=1}^{n} a_{in} v_i) =$$

$$= \ldots = \sum_{i_1=1}^{n} \ldots \sum_{i_n=1}^{n} f(a_{i_1 1} v_{i_1}, a_{i_2 2} v_{i_2}, \ldots, a_{i_n n} v_{i_n}) = \sum_{i_1, i_2, \ldots, i_n} f(a_{i_1 1} v_{i_1}, a_{i_2 2} v_{i_2}, \ldots, a_{i_n n} v_{i_n}) =$$

$$\overset{\text{кососим.}}{=} \sum_{(i_1, \ldots, i_n) - \text{перест.}} a_{i_1 1} a_{i_2 2} \ldots a_{i_n n} f(v_{i_1}, v_{i_2}, \ldots, v_{i_n}) \overset{(i_1, i_2, \ldots, i_n) = \rho(\sigma)}{=}$$

$$\sum_{\sigma \in S_n} a_{\sigma(1)1} a_{\sigma(2)2} \ldots a_{\sigma(n)n} f(v_{\sigma(1)}, v_{\sigma(2)}, \ldots, v_{\sigma(n)}) = \sum_{\sigma \in S_n} \varepsilon(\sigma) a_{\sigma(1)1} a_{\sigma(2)2} \ldots a_{\sigma(n)n} \varepsilon(\sigma) f(v_1, v_2, \ldots, v_n).$$

Дійсно, перестановку $(\sigma(1), \sigma(2), \ldots, \sigma(n))$ можна перетворити в $(1, 2, \ldots, n)$ за допомогою $T(\sigma)$ елементарних транспозицій, тому знак зміниться $T(\sigma)$ разів. Маємо

$$f(v_{\sigma(1)}, v_{\sigma(2)}, \ldots, v_{\sigma(n)}) = (-1)^{T(\sigma)} f(v_1, v_2, \ldots, v_n) = \varepsilon(\sigma) f(v_1, v_2, \ldots, v_n).$$

Отже, полілінійна та кососиметрична функція $f : V^n \to \mathbb{K}$ задається формулою

$$f(\alpha_1, \alpha_2, \ldots, \alpha_n) = \sum_{\sigma \in S_n} \varepsilon(\sigma) a_{\sigma(1)1} a_{\sigma(2)2} \ldots a_{\sigma(n)n} f(v_1, v_2, \ldots, v_n). \tag{12}$$

Перевіримо, що формула (12) для f дійсно визначає полілінійну і кососиметричну функцію $f : V^n \to \mathbb{K}$. Можемо вважати, що $f(v_1, \ldots, v_n) = 1$. Перевіримо полілінійність.

$$1) \quad f(\alpha_1, \ldots, \alpha_i' + \alpha_i'', \ldots, \alpha_n) = \sum_{\sigma \in S_n} \varepsilon(\sigma) a_{\sigma(1)1} \ldots (a'_{\sigma(i)i} + a''_{\sigma(i)i}) \ldots a_{\sigma(n)n} =$$

$$= \sum_{\sigma \in S_n} \varepsilon(\sigma) a_{\sigma(1)1} \ldots a'_{\sigma(i)i} \ldots a_{\sigma(n)n} + \sum_{\sigma \in S_n} \varepsilon(\sigma) a_{\sigma(1)1} \ldots a''_{\sigma(i)i} \ldots a_{\sigma(n)n} =$$

$$= f(\alpha_1, \ldots, \alpha_i', \ldots, \alpha_n) + f(\alpha_1, \ldots, \alpha_i'', \ldots, \alpha_n).$$

$$2) \quad f(\alpha_1, \ldots, \lambda\alpha_i, \ldots, \alpha_n) = \sum_{\sigma \in S_n} \varepsilon(\sigma) a_{\sigma(1)1} \ldots (\lambda a_{\sigma(i)i}) \ldots a_{\sigma(n)n} =$$

$$= \lambda \sum_{\sigma\in S_n} \varepsilon(\sigma) a_{\sigma(1)1} \dots a_{\sigma(i)i} \dots a_{\sigma(n)n} = \lambda f(\alpha_1, \dots, \alpha_i, \dots, \alpha_n).$$

Перевіримо полілінійність: $f(\alpha_1, \dots, \alpha_i, \dots, \alpha_j, \dots, \alpha_n) = 0$, якщо $\alpha_i = \alpha_j$. Введемо позначення: $A(\sigma) = \varepsilon(\sigma) a_{\sigma(1)1} a_{\sigma(2)2} \dots a_{\sigma(n)n}$. Тоді формула (12) набуває вигляду

$$f(\alpha_1, \alpha_2, \dots, \alpha_n) = \sum_{\sigma\in S_n} A(\sigma) = \sum_{\sigma\in A_n} A(\sigma) + \sum_{\sigma\in S_n \setminus A_n} A(\sigma).$$

Розглянемо транспозицію σ_{ij}. За лемою 7.5, це непарна підстановка. Тому, за лемою 7.4, маємо $S_n \setminus A_n = A_n \sigma_{ij}$. Отже,

$$f(\alpha_1, \alpha_2, \dots, \alpha_n) = \sum_{\sigma\in A_n} A(\sigma) + \sum_{\sigma\in S_n \setminus A_n} A(\sigma\sigma_{ij}) = \sum_{\sigma\in A_n} (A(\sigma) + A(\sigma\sigma_{ij})).$$

Стверджується, що кожен доданок в останній сумі рівний нулю. Дійсно,

$A(\sigma) + A(\sigma\sigma_{ij}) =$
$\varepsilon(\sigma) a_{\sigma(1)1} \dots a_{\sigma(i)i} \dots a_{\sigma(j)j} \dots a_{\sigma(n)n} + \varepsilon(\sigma\sigma_{ij}) a_{\sigma\sigma_{ij}(1)1} \dots a_{\sigma\sigma_{ij}(i)i} \dots a_{\sigma\sigma_{ij}(j)j} \dots a_{\sigma\sigma_{ij}(n)n}$.

При цьому $a_{\sigma\sigma_{ij}(i)i} = a_{\sigma(j)i}$, $a_{\sigma\sigma_{ij}(j)j} = a_{\sigma(i)j}$, а для решти індексів маємо: $a_{\sigma\sigma_{ij}(k)k} = a_{\sigma(k)k}$. Оскільки $i-$тий та $j-$тий стовпчики збігаються між собою, то $a_{\sigma(j)i} = a_{\sigma(j)j} \cdot a_{\sigma(i)i} = a_{\sigma(i)j}$.

Добутки $a_{\sigma(1)1} \dots a_{\sigma(n)n}$ і $a_{\sigma\sigma_{ij}(1)1} \dots a_{\sigma\sigma_{ij}(n)n}$ збігаються, а $\varepsilon(\sigma\sigma_{ij}) = -\varepsilon(\sigma)$, тому $A(\sigma) + A(\sigma\sigma_{ij}) = 0 \implies f(a_1, \dots, a_i, \dots, a_j, \dots, a_n) = 0$ при $a_i = a_j$.

Залишилось показати, що довільні дві такі функції вигляду (12) (якщо вони існують) пропорційні.

□

Означення 7.7. *Нехай $V = \mathbb{K}^n$, $\mathcal{E} = \{e_1, \dots, e_n\}$ – стандартний базис в просторі $\mathbb{K}^n$. Полілінійна і кососиметрична функція $\Delta : \underbrace{\mathbb{K}^n \times \ldots \times \mathbb{K}^n}_{n} \to \mathbb{K}$ така, що $\Delta(e_1, e_2, \dots, e_n) = 1$, називається* ***детермінантом*** *(визначником).*

Якщо $A = (a_1 \,|\, a_2 \,|\, \dots \,|\, a_n)$ – деяка $(n \times n)-$матриця, то детермінантом матриці називається $\Delta(A) \overset{df}{=} \Delta(a_1, a_2, \dots, a_n)$.

Наслідок 7.1. *Функція $\Delta : (\mathbb{K}^n)^n \to \mathbb{K}$ існує, визначена однозначно, і для матриці $A = (a_{ij})_{1\leq i,j\leq n}$ обчислюється за формулою:*

$$\Delta(A) = \Delta(a_1, a_2, \dots, a_n) = \sum_{\sigma\in S_n} a_{\sigma(1)1} \dots a_{\sigma(n)n}.$$

Прийнято вживати ще такі позначення: $\Delta(A) = |A| = \det(A)$.

Приклад. Якщо $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$, то $\Delta(A) = a_{11}a_{22} - a_{12}a_{21}$.

Твердження 7.2. *Нехай A – $(n \times n)-$матриця над $\mathbb{K}$, A^t – транспонована матриця. Тоді $\Delta(A) = \Delta(A^t)$.*

Доведення. Якщо $A = (a_{ij})_{1\le i,j\le n}$, то $A^t = (a^t_{ij})_{1\le i,j\le n}$, де $a^t_{ij} = a_{ji}$. Маємо:

$$\Delta(A^t) = \sum_{\sigma\in S_n} \varepsilon(\sigma) a^t_{\sigma(1)1}\dots a^t_{\sigma(n)n} = \sum_{\sigma\in S_n}\varepsilon(\sigma)a_{1\sigma(1)}\dots a_{n\sigma(n)}.$$

Тоді, оскільки $\begin{pmatrix} \sigma(1) & \sigma(2) & \dots & \sigma(n) \\ 1 & 2 & \dots & n \end{pmatrix} = \begin{pmatrix} 1 & 2 & \dots & n \\ \sigma(1)^{-1} & \sigma(2)^{-1} & \dots & \sigma(n)^{-1} \end{pmatrix} = \sigma^{-1}$, то $\Delta(A^t) = \sum_{\sigma\in S_n}\varepsilon(\sigma)a_{\sigma^{-1}(1)1}a_{\sigma^{-1}(2)2}\dots a_{\sigma^{-1}(n)n} = \sum_{\sigma\in S_n}\varepsilon(\sigma^{-1})a_{\sigma^{-1}(1)1}a_{\sigma^{-1}(2)2}\dots a_{\sigma^{-1}(n)n} = \sum_{\sigma\in S_n}\varepsilon(\sigma)a_{\sigma(1)1}a_{\sigma(2)2}\dots a_{\sigma(n)n} = \Delta(A)$. □

Наслідок 7.2. *Детермінант матриці A не змінюється при елементарних перетвореннях 1 роду над рядками матриці A.*

Доведення. Ми довели твердження про перетворення над стовпчиками. Перетворення над рядками зводяться до випадку стовпчиків транспонуванням матриці. □

Теорема 7.2. *Нехай A, B – дві $(n\times n)-$матриці над $\mathbb{K}$, тоді $\Delta(AB) = \Delta(A)\Delta(B)$.*

Доведення. Фіксуємо матрицю $A \in Mat_{n\times n}(\mathbb{K})$ і розглянемо $\Delta(AB)$ як функцію від стовпчиків матриці $B = (b_1\,|\,b_2\,|\,\dots\,|\,b_n)$, тобто визначимо функцію

$$f(b_1, b_2, \dots, b_n) = \Delta(AB) = \Delta(Ab_1\,|\,Ab_2\,|\,\dots\,|\,Ab_n), \qquad f : \underbrace{\mathbb{K}^n\times\dots\times\mathbb{K}^n}_{n} \to \mathbb{K}.$$

Покажемо, що f – полілінійна і кососиметрична.

$B = (b_1\,|\,b_2\,|\,\dots\,|\,b_n) \implies AB = (Ab_1\,|\,Ab_2\,|\,\dots\,|\,Ab_n)$, за правилом множення матриць. Тоді $f(b_1,\dots,b_i'+b_i''\dots b_n) = \Delta(A(b_1\,|\,\dots\,|\,b_i'+b_i''\,|\,\dots\,|\,b_n)) = \Delta((Ab_1\,|\,\dots\,|\,A(b_i'+b_i'')\,|\,\dots\,|\,Ab_n)) \overset{\text{дистр.}\Delta}{=} \Delta((Ab_1\,|\,\dots\,|\,Ab_i'+Ab_i''\,|\,\dots\,|\,Ab_n)) \overset{\text{полілін.}\Delta}{=} \Delta((Ab_1\,|\,\dots\,|\,Ab_i'\,|\,\dots\,|\,Ab_n)) + \Delta((Ab_1\,|\,\dots\,|\,Ab_i''\,|\,\dots\,|\,Ab_n)) = \Delta(A(b_1\,|\,\dots\,|\,b_i'\,|\,\dots\,|\,b_n)) + \Delta(A(b_1\,|\,\dots\,|\,b_i''\,|\,\dots\,|\,b_n)) = f(b_1,\dots,b_i'\dots b_n) + f(b_1,\dots,b_i''\dots b_n)$.

Аналогічно, $f(b_1,\dots,\lambda b_i\dots b_n) = \lambda f(b_1,\dots,b_i\dots b_n)$.

Кососиметричність: $f(b_1,\dots,b_i\dots,b_j\dots b_n) \overset{b_i=b_j}{=} \Delta(A(b_1\,|\,\dots\,|\,b_i\,|\,\dots\,|\,b_j\,|\,\dots\,|\,b_n)) = \Delta(Ab_1\,|\,\dots\,|\,Ab_i\,|\,\dots\,|\,Ab_j\,|\,\dots\,|\,Ab_n)) = 0$ тому, що з $b_i = b_j$ випливає $Ab_i = Ab_j$.

Оскільки f – полілінійна і симетрична, то в фіксованому базисі $(\{e_1, e_2, \dots e_n\})$ вона дорівнює $f(b_1,\dots,b_n) = \Delta(b_1\,|\,\dots\,|\,b_n)f(e_1,\dots,e_n)$. Але

$$f(e_1,\dots,e_n) = \Delta(A\underbrace{(e_1\,|\,\dots\,|\,e_n)}_{I_n}) = \Delta(AI_n) = \Delta(A).$$

Тому $\quad f(b_1,\dots,b_n) = \Delta(b_1\,|\,\dots\,|\,b_n)\Delta(A) = \Delta(B)\Delta(A) \implies \Delta(AB) = \Delta(A)\Delta(B)$.

□

Теорема 7.3. *Нехай $A = (A_{ij}) \in Mat_{n\times n}(\mathbb{K})$. Наступні умови є еквівалентними.*

1. *Відображення* $L_A : \mathbb{K}^n \to \mathbb{K}^n$ *– ізоморфізм.*

2. *Матриця* A *обертовна* ($\exists\ A^{-1} : AA^{-1} = A^{-1}A = I_n$).

3. L_A *– мономорфізм.*

4. L_B *– епіморфізм.*

5. *Система* $Ax = b$ *визначена (має єдиний розв'язок) для довільного* $b \in \mathbb{K}^n$.

6. *Існує* $(n \times n)-$*матриця* X *така, що* $XA = I$ *(ліва обернена).*

7. *Існує* $(n \times n)-$*матриця* Y *така, що* $AY = I$ *(права обернена).*

8. $\Delta(A) \neq 0$.

Термін. Матриця A, що задовольняє теоремі, називається **невиродженою**.

Доведення. Доведемо еквівалентність умов.

1) $\iff$ 2). За означенням обертовності матриці.

1) $\implies$ 3), 1) $\implies$ 4). Якщо L_A – ізоморфізм, то L_A є мономорфізмом і епіморфізмом.

3) $\implies$ 1) Якщо L_A – моно $\implies$ L_A – ізоморфізм (це вірно тільки для квадратичних матриць). Чому? Розглянемо $L_A : \mathbb{K}^n \to \mathbb{K}^n$ ми знаємо, що $n = \operatorname{DIM}_{\mathbb{K}}\mathbb{K}^n = \operatorname{DIM}_{\mathbb{K}}\operatorname{KER}\ L_A + +\operatorname{DIM}_{\mathbb{K}} \in L_A$.

L_A – моно $\implies$ $\operatorname{KER}\ L_A = 0 \implies \operatorname{DIM} \in L_A = n - \operatorname{DIM}\operatorname{KER}\ L_A$, отже $\operatorname{DIM} \in L_A = n$, але $\in L_A \subset \mathbb{K}^n$ – підпростір (коли маємо підпростір і розмірність підпростору дорівнює розмірності простору, то підпростір дорівнює простору) $\implies \in L_A = \mathbb{K}^n$ $\implies$ L_A – епіморфізм $\implies$ L_A – ізоморфізм.

4) $\implies$ 1) доводиться аналогічно. Якщо L_A – епіморфізм, то $\operatorname{DIM} \in L_A = n \implies \operatorname{DIM}\operatorname{KER}\ L_A = n - \operatorname{DIM} \in L_A = n - n = 0 \implies \operatorname{KER}\ L_A = 0$, тобто L_A – моно $\implies$ L_A – ізоморфізм.

1) $\implies$ 5). Припустимо, що L_A ізоморфізм. Тоді $\forall b \in \mathbb{K}^n\ \exists! a = L_A^{-1}(b)$ такий, що $L_A(a) = b$. Але умова $L_A(a) = b$ еквівалентна умові $Aa = b$, що еквівалентно тому, що a є розв'язком $Ax = b$.

5) $\implies$ 3). Припустимо, що для довільного $b \in \mathbb{K}^n$ система $Ax = b$ має єдиний розв'язок. Тоді підставимо $b = 0$, одержимо, що лінійна однорідна система $Ax = 0$ має єдиний розв'язок $x = 0 \implies \operatorname{KER}\ L_A = \{x \in \mathbb{K}^n \mid L_A(x) = 0\} = \emptyset \implies L_A$ – моно.

6) $\implies$ 3). Припустимо, існує $X \in \mathrm{Mat}_{n\times n}(\mathbb{K})$ така, що $X \cdot A = I$. Запишемо цю умову на лінійних відображеннях. Одержимо: $L_X \cdot L_A = L_I = 1_{\mathbb{K}^n}$. Розглянемо довільний вектор $a \in \operatorname{KER}\ L_A$, подіємо на нього правою і лівою частинами матричної рівності.

$= (L_X \cdot L_A)(a) = 1_{\mathbb{K}^n}(a)$

$= L_X(L_A(a)) = a$.

Отже, $L_X(L_A(a)) = 0 \implies a = 0$. Звідси випливає, що $\operatorname{KER}\ L_A = 0 \implies L_A$ – моно.

2) $\implies$ 6). Очевидно (оскільки обернена матриця є лівою оберненою).

2) $\implies$ 7) – аналогічно. Коли існує A^{-1}: $AA^{-1} = A^{-1}A = I$, то покладемо $Y = A^{-1}$.

7) $\implies$ 4). Існує $Y \in \mathrm{Mat}_{n\times n}(\mathbb{K})$: $A\cdot Y = I_n \implies L_A$ є epi. Запишемо умову $AY = I_n$ мовою лінійних відображень, одержимо $L_A L_Y = L_{I_n} = 1_{\mathbb{K}^n}$. Розглянемо довільний вектор $a \in \mathbb{K}^n$. Подіємо на нього правою і лівою частинами рівності, одержимо:

$(L_A \cdot L_Y(a) = 1_{\mathbb{K}^n}(a) \implies L_A(L_Y(a)) = a \implies a \in\in L_A \implies \in L_A = \mathbb{K}^n \implies L_A$ – епі.

2) $\implies$ 8). Припустимо, $A \in \mathrm{GL}_n(\mathbb{K})$ і запишемо рівність $A\cdot A^{-1} = I_n$. Детермінант добутку є добутком детермінантів. Застосуємо цю рівність.

$\Delta(AA^{-1}) = \Delta(I_n) = 1 \implies \Delta(A) \neq 0$.

8) $\implies$ 2). Нехай $\Delta(A) \neq 0$. Ми довели, що послідовністю елементарних перетворень рядків і стовпчиків довільна матриця, зокрема, матриця A,зводиться до вигляду $I_{n\times n,r}$. Елементарні перетворення виконуються за допомогою множення на елементарні матриці. Тому існують елементарні матриці $E_1,\ldots,E_k$, $\mathcal{E}_1,\ldots,\mathcal{E}_l$ такі, що $E_k,\ldots,E_1 A\mathcal{E}_1,\ldots,\mathcal{E}_l = I_{n\times n,r}$. Зауважимо, що детермінант елементарної матриці ненульовий (сказати раніше): $\Delta(\mathcal{E}_{ik}(\lambda)) = 1$, $\Delta(\mathcal{E}_k(\lambda)) = \lambda \neq 0$, або, оскільки елементарні матриці обертовні, застосуємо функцію det до правої та лівої частин: $\Delta(E_k,\ldots,E_1 A\mathcal{E}_1,\ldots,\mathcal{E}_l) = \Delta(I_{n\times n,r})$. Припустимо, що $r < n$, тоді $\Delta(I_{n\times n,r}) = 0$. Отже, якщо $r < n$, то $\Delta(I_{n\times n,r}) = 0 \implies \Delta(E_k)\ldots\Delta(E_k)\Delta(A)\Delta(\mathcal{E}_1)\ldots\Delta(\mathcal{E}_l) = 0$. Але всі співмножники ненульові. Одержимо протиріччя. Звідси маємо: $r = n$, тобто $I_{n\times n,n} = I_n$ $\implies E_k\ldots E_1 A\mathcal{E}_1\ldots\mathcal{E}_l = I_n \implies A = E_1^{-1}\ldots E_k^{-1} I_n \mathcal{E}_l^{-1}\ldots\mathcal{E}_1^{-1} \implies A$ – обертовна, як добуток обертовних матриць.

Теорему доведено.

□

Такі матриці називаються невиродженими.

Теорема 7.4. *(Правило) Крамера. Нехай $A \in \mathrm{Mat}_{m\times n}(\mathbb{K})$, $A = (a_1\,|\,a_2\,|\,\ldots\,|\,a_n)$, і маємо систему $Ax = b$. Позначимо через $\Delta = \Delta(A)$, а через $\Delta_i = \Delta(a_1\,|\,\ldots\,|\,a_{i-1}\,|\,b\,|\,a_{i+1}\,|\,\ldots\,|\,a_n)$, $i = 1,\ldots,n$. Якщо система $Ax = b$ сумісна і $\begin{pmatrix} l_1 \\ l_2 \\ \vdots \\ l_n \end{pmatrix}$ – її розв'язок, то має місце рівність $\Delta\cdot l_i = \Delta_i$, $i = 1,\ldots,n$.*

Доведення. $l \in \mathcal{L}$ є розв'язком тоді і лише тоді, коли виконується рівність $a_1 l_1 + \ldots + a_n l_n = b$, тобто $\sum_{i=1}^{n} l_i a_i = b$. Обчислимо $\Delta_i = \Delta(a_1\,|\,\ldots\,\underbrace{|\,b\,|}_{i}\ldots\,|\,a_n) = \Delta(a_1\,|\,\ldots\,|\sum_{j=1}^{n} l_j a_j\,|\ldots\,|\,a_n)$

$$\overset{\text{полілін.}}{=} \sum_{j=1}^{n}\Delta(a_1\,|\,\ldots\,|\,l_j a_j\,|\,\ldots\,|\,a_n) = \sum_{j=1}^{n} l_j\Delta(a_1\,|\,\ldots|\,a_j\,|\,\ldots\,|\,a_n) =$$

Коли $j \neq i$, то відповідний доданок в цій сумі є нульовим (оскільки є два однакових стовпчики, то визначник дорівнює нулю, це наслідок кососиметричності).

$= l_i\Delta(a_1 \,|\, \dots \underbrace{|\, a_i \,|}_{i} \dots a_n) = l_i\Delta(A).$

Отже, $\Delta_i = l_i\Delta(A)$. $\square$

Наслідок (правило Крамера). Нехай в системі $Ax = b$, $A \in \mathrm{Mat}_{n\times n}(\mathbb{K})$ $\Delta(A) \neq 0$ і $\Delta_i = \Delta(a_1 \,|\, \dots \underbrace{|\, b \,|}_{i} \dots a_n)$. Тоді система має єдиний розв'язок, що обчислюється за формулою: $l = \begin{pmatrix} l_1 \\ l_2 \\ \vdots \\ l_n \end{pmatrix}$, $l_i = \frac{\Delta_i}{\Delta}$, $i = 1, \dots, n$.

Доведення. Оскільки $\Delta(A) \neq 0$, у відповідності з пунктом 5 теореми, система $Ax = b$ визначена. Застосовуємо теорему Крамера і одержуємо, що $l_i\Delta = \Delta_i$ для довільного розв'язку $l = \begin{pmatrix} l_1 \\ l_2 \\ \vdots \\ l_n \end{pmatrix}$. Звідси, оскільки $\Delta \neq 0$, одержуємо формули: $l_i = \frac{\Delta_i}{\Delta}$, $i = 1, \dots, n$. $\square$

Зауваження 7.3. *Якщо $\Delta = 0$,і існує $i \in \{1, \dots, n\}$ таке, що $\Delta_i \neq 0$, то звідси випливає, що система несумісна.*

7.6 Поняття мінору та алгебраїчного доповнення

Припустимо, $A \in \mathrm{Mat}_{n\times n}(\mathbb{K})$. Нехай $1 \leq i_1 \leq \dots \leq i_k \leq n$ і $1 \leq j_1 \leq \dots \leq j_k \leq n$ – деякі два набори індексів. Розглянемо матрицю $(n-k) \times (n-k)$, яка одержана з A видаленням рядків з номерами $i_1, \dots, i_k$ і стовпчиків з номерами $j_1, \dots, j_k$. Детермінант одержаної матриці $A[i_1, \dots, i_k; j_1, \dots, j_k]$ називається **мінором** матриці A порядку $n - k$. Цей мінор позначається $A_{i_1,\dots,i_k;j_1,\dots,j_k}$.

Розглянемо випадок $k = 1$. Мінор $A_{i_1;j_1}$ позначатимемо через A_{i_1,j_1}. Елемент $(-1)^{i+j}A_{i_1,j_1}$ називається **алгебраїчним доповненням** $i, j-$того елемента матриці A.

Твердження 7.3. *Нехай $A \in \mathrm{Mat}_{n\times n}(\mathbb{K})$.* ***Правило розкладу за рядком.*** *Має місце така рівність для довільного $i = 1, \dots, n$:*

$$\Delta(A) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} A_{ij} \text{ або } \Delta(A) = \sum_{j=1}^{n} a_{ij}((-1)^{i+j} A_{ij})$$

(алгебраїчне доповнення).

Правило розкладу за стовпчиком. *Має місце така рівність для довільного $i = 1, \dots, n$:*

$$\Delta(A) = \sum_{i=1}^{n} a_{ij}(-1)^{i+j} A_{ij}.$$

Доведення. Доведемо перше твердження. Друга формула аналогічна першій (або ж одержується застосуванням першої формули до транспонованої матриці). Розглянемо

$$\Delta(A) = \left| \begin{array}{cc} i & \begin{pmatrix} \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots \\ a_{i1} & \dots & a_{in} \\ \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots \end{pmatrix} \end{array} \right| \overset{\text{полілін.}}{=} \sum_{j=1}^{n} \left| \begin{array}{cc} & \begin{matrix} & & & j & & & \end{matrix} \\ i & \begin{pmatrix} \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \dots & 0 & a_{ij} & 0 & \dots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \end{pmatrix} \end{array} \right| =$$

$$\sum_{j=1}^{n} a_{ij} \left| \begin{array}{cc} & \begin{matrix} & & & j & & & \end{matrix} \\ i & \begin{pmatrix} \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \dots & 0 & 1 & 0 & \dots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \end{pmatrix} \end{array} \right|$$

Визначник не змінюється при елементарних перетвореннях першого роду, тому за допомогою 1 робимо нулі в $j-$тому стовпчику. Одержимо

$$\Delta(A) = \sum_{j=1}^{n} a_{ij} \left| \begin{array}{cc} & \begin{matrix} & & & j & & & \end{matrix} \\ i & \begin{pmatrix} \cdots & \cdots & \cdots & 0 & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & 0 & \cdots & \cdots & \cdots \\ 0 & \dots & 0 & 1 & 0 & \dots & 0 \\ \cdots & \cdots & \cdots & 0 & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & 0 & \cdots & \cdots & \cdots \end{pmatrix} \end{array} \right|$$

В записаній матриці всі елементи, крім $i-$того рядка і $j-$того стовпчика такі ж, як і в матриці A.

Обчислимо визначник такої матриці. Переставимо $j-$ий та $1-$ий стовпчики, потім $i-$ий та $1-$ий рядки. Міняємо рядки та стовпчики за допомогою елементарних транспозицій. Виконуємо при цьому $j-1$ перестановку сусідніх стовпчиків та $i-1$ перестановку сусідніх рядків. Одержимо, що наш визначник рівний

$$(-1)^{j-1+i-1} \left| \begin{pmatrix} 1 & \vdots & 0 \\ \cdots & \cdots & \cdots \\ 0 & \vdots & A_{[ij]} \end{pmatrix} \right| = (-1)^{i+j} \left| \ A_{[ij]} \ \right| = (-1)^{i+j} A_{ij} = \sum_{j=1}^{n} a_{ij} (-1)^{i+j} A_{ij}.$$

$\square$

Розклад за рядком (стовпчиком)

$A = (a_{ij})$, $1 \leqslant i, j \leqslant n$.

$\Delta(A) = \sum\limits_{i=1}^{n}(-1)^{i+j}a_{ij}A_{ij}\ \forall j = 1, \ldots, n$

Ми використали таку рівність: $\begin{pmatrix} 1 & 0 \\ 0 & \boxed{X} \end{pmatrix} = \boxed{X}$

Наслідок (розклад за чужим рядком або стовпчиком).

$$\sum_{i=1}^{n}(-1)^{i+j}a_{ik}A_{ij} = 0 \text{ для довільного } k \neq j, k = 1, \ldots, n.$$

$$\sum_{j=1}^{n}(-1)^{i+j}a_{kj}A_{ij} = 0 \text{ для довільного } k \neq i, k = 1, \ldots, n.$$

Доведення. Доведемо першу формулу. Розглянемо матрицю A', яка одержана з A заміною $j-$того стовпчика на $k-$тий.

$A = (a_1 \,|\, \ldots \underbrace{|\, a_j \,|}_{j} \,|\, \ldots \underbrace{|\, a_k \,|}_{k} \ldots a_n)$

$A' = (a_1 \,|\, \ldots \underbrace{|\, a_k \,|}_{j} \,|\, \ldots \underbrace{|\, a_k \,|}_{k} \ldots a_n)$

$\Delta(A') = 0$ тому, що A' має два однакових стовпчики.

Розкладемо детермінант матриці A' за $j-$тим стовпчиком. Зауважимо, що всі мінори $A'_{ij} = A_{ij}$, $i = 1, \ldots, n$ (після викидання $j-$того стовпчика матриці A і A' перетворюються на рівні між собою матриці). Отже, $\Delta(A') = 0 = \sum\limits_{i=1}^{n}(-1)^{i+j} \underbrace{a_{ik}}_{k-\text{ий ст.}} A'_{ij} =$

$\sum\limits_{i=1}^{n}(-1)^{i+j}a_{ik}A_{ij} = 0.$

Друга формула доводиться аналогічно. $\square$

Формула для оберненої матриці

Ми довели, що коли $\Delta(A') \neq 0$, то звідси випливає, що $A \in \mathrm{GL}_n(\mathbb{K})$ (тобто існує A^{-1}). Як ми довели? Достатньо побудувати матрицю $X:\ AX + I_n$.

$X = (x_1 \,|\, x_2 \,|\, \ldots \,|\, x_n \,|\,) \implies AX = (Ax_1 \,|\, Ax_2 \,|\, \ldots \,|\, Ax_n \,|\,) = I_n$. Маємо n систем лінійних рівнянь з матрицею A. Розв'язуючи їх за правилом Крамера, знаходимо поліноміальні формули для елементів матриці X. Домашнє завдання: виконати обчислення.

Приєднана матриця

Нехай $A = (a_{ij}) \in \mathrm{Mat}_{n\times n}(\mathbb{K})$. Матриця $\widetilde{A} = (\widetilde{a}_{ij}) \in \mathrm{Mat}_{n\times n}(\mathbb{K})$ називається приєднаною до A, якщо $\widetilde{a}_{ij} = (-1)^{i+j}A_{ji}$, $1 \leqslant i, j \leqslant n$.

Лемма 7.7. (Про приєднану матрицю). *Нехай $A = (a_{ij}) \in \mathrm{Mat}_{n\times n}(\mathbb{K})$, $\widetilde{A}$ - приєднана до A. Тоді*

$$A \cdot \widetilde{A} \ = \ \widetilde{A} \cdot A \ = \ \Delta(A) \cdot I_n.$$

Доведення. $(A\widetilde{A})_{ij} = \sum\limits_{k=1}^{n} a_{ik}(\widetilde{A})_{kj} =$ (формула про множення матриць) $\sum\limits_{k=1}^{n} a_{ik}(-1)^{i+k}A_{jk} = \sum\limits_{k=1}^{n}(-1)^{i+k}a_{ik}A_{jk} = \begin{cases} 0, & \text{якщо} \quad i \neq j, \\ \Delta(A), & \text{якщо} \quad i = j. \end{cases}$

Тобто, в матриці $(A\widetilde{A})$ коефіцієнт $_{ij}$ дорівнює 0, коли $i \neq j$, і дорівнює $\Delta(A)$, коли $i = j$. Звідси одержимо $A \cdot \widetilde{A} \;=\; \Delta(A) \cdot I_n$. $\square$

Наслідок (формула для оберненої матриці). Якщо $\Delta(A) \neq 0$, то $A^{-1} = \dfrac{1}{\Delta(A)} \cdot \widetilde{A}$.

Доведення. Оскільки $\Delta(A) \neq 0$, то A^{-1} існує і однозначно визначається рівністю $AA^{-1} = I_n$. Перевіримо формулу:

$A(\dfrac{1}{\Delta(A)} \cdot \widetilde{A}) = \dfrac{1}{\Delta(A)} \cdot (A\widetilde{A}) = \dfrac{1}{\Delta(A)}\Delta(A)I_n = I_n$. $\square$

Приклад. $n = 2$

$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$, $A_{11} = a_{22}$, $A_{12} = a_{21}$, $A_{22} = a_{11}$, $A_{21} = a_{12}$. Тоді $\widetilde{A} = \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}$, $\Delta(A) = a_{11}a_{22} - a_{12}a_{21}$, $A^{-1} = \dfrac{1}{a_{11}a_{22} - a_{12}a_{21}} \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}$.

Мінорний ранг

Яка користь від визначників для неквадратних матриць?

$A \in \mathrm{Mat}_{m\times n}(\mathbb{K})$, $0 \leqslant k \leqslant \min(m, n)$

$1 \leqslant i_1 < \ldots < i_k \leqslant m$, $1 \leqslant j_1 < \ldots < j_k \leqslant n$

Позначимо через $A|_{[i_1,\ldots,i_k;j_1,\ldots,j_k]}$ $(k \times k)-$матрицю, утворену елементами, які знаходяться в рядках з номерами $i_1, \ldots, i_k$ та стовпчиками з номерами $j_1, \ldots, j_k$. Визначник цієї матриці ми називаємо мінором матриці A (мінором $k-$того порядку, що відповідає рядкам $i_1, \ldots, i_k$ та стовпчикам $j_1, \ldots, j_k$).

Мінорним рангом матриці A називається порядок (максимальний) k, за якого існує ненульовий мінор (порядок максимального ненульового мінора).

Зауваження 7.4. *Мінорний ранг вважається рівним* 0*, якщо A - нульова матриця.*

Тимчасове позначення: $\mathrm{RANK}_m(A)$.

Теорема 7.5. $A \in \mathrm{Mat}_{m\times n}(\mathbb{K})$. $\mathrm{RANK}_m(A) = \mathrm{RANK}(A)$.

Лемма 7.8. *Наступні твердження еквівалентні:*

1) $\mathrm{RANK}A \geqslant r$*;*

2) Існує мінор порядку r матриці A, відмінний від 0.

Зауваження 7.5. *Взявши максимально можливе r в 1), одержуємо $r = \mathrm{RANK}(A)$; максимально можливе r в 2) дає $r = \mathrm{RANK}_m(A) \implies$ теорема.*

Доведення. (леми).

$\mathrm{RANK}(A) \geqslant r \implies$ в матриці A існують rлінійно незалежних рядків.

Зауваження 7.6. *Ні* RANK(A), *ні* RANK$_m(A)$ *не змінюються при перестановках рядків і стовпчиків матриці* A.

Будемо вважати (з врахуванням зауваження), що перші r рядків матриці A - лінійно незалежні (інакше переставимо лінійно незалежні рядки на перші місця). Позначимо через A' $(r \times n)-$матрицю, утворену першими r рядками матриці A, $A = \begin{pmatrix} a^1 \\ \vdots \\ a^r \end{pmatrix}$.

Очевидно, RANK$A' = r$. Оскільки RANK$_r(A') = $ RANK$_C(A')$, то r існують r лінійно незалежних стовпчиків. (Вважаємо, що це перші r стовпчиків). Розглянемо мінор в A', утворений першими r рядками і r першими стовпчиками (M).

$M = \det A''$ (з точністю до знака M є мінором порядку r матриці A). Стверджується, що $M \neq 0$, тому що r стовпчиків матриці A', які утворюють матрицю A'', лінійно незалежні. Отже, з 1) $\implies$ 2). RANK$A \geqslant r \implies$ існує мінор порядку $r \neq 0$.

2) $\implies$ 1). Припустимо, що існує мінор порядку r, відмінний від 0. Переставивши рядки-стовпчики, вважаємо, що він утворений першими r рядками та стовпчиками.

Оскільки $\Delta(A') \neq 0$, то перші r рядків в A' - лінійно незалежні $\implies$ перші r рядків в A теж лінійно незалежні.

$a^1, a^2, \ldots, a^r$. Якщо $\sum\limits_{i=1}^{r} \lambda_i a^i = 0$, то, розглянувши лише перші r координат, одержимо $\sum\limits_{i=1}^{r} \lambda_i a^{i'} = 0$, де $a^{i'}$ - $i-$тий рядок матриці $A' \implies$ RANK$A \geqslant r$ (існує r лінійно незалежних рядків). $\square$

Корисні детермінанти

$$A = \begin{pmatrix} a_{11} & \ldots & 0 \\ \vdots & \vdots & \vdots \\ 0 & \ldots & a_{nn} \end{pmatrix}, \qquad \Delta(A) = a_{11} \ldots a_{nn}.$$

Доведення. $\Delta(A) = \sum\limits_{\sigma \in S_n} \varepsilon(\sigma) a_{\sigma(1)1} \ldots a_{\sigma(n)n}$

Для того, щоб доданок в цій сумі не був апріорно нульовим

$\sigma(1) \leqslant 1$, $\sigma(2) \leqslant 2$, ..., $\sigma(n) \leqslant n \implies \sigma : \{1, \ldots, n\} \to \{1, \ldots, n\}$ - бієкція $\implies$ $\sigma = \mathtt{1}$, $\sigma(i) = i$ для довільного $i \implies \Delta(A) = \sum\limits_{\sigma = \mathtt{1}} \varepsilon(\sigma) a_{\sigma(1)1} \ldots a_{\sigma(n)n} = a_{11} \ldots a_{nn}$. $\square$

Визначник блочної матриці

$$\begin{pmatrix} & k & & \\ k & \boxed{B} & \boxed{*} & \\ & \boxed{0} & \boxed{C} & l \\ & & l & \end{pmatrix}$$

$n = k + l$

$\Delta(A) = \Delta(B)\Delta(C)$.

Доведення. $\Delta(A) = \sum\limits_{\sigma \in S_n} \varepsilon(\sigma) a_{\sigma(1)1} \dots a_{\sigma(n)n}$

Якщо $i = 1, \dots, k$, $\sigma(i) > i$, то відповідний доданок дорівнює 0. Звідси випливає, що для σ, які визначають ненульовий доданок має місце: $\sigma : \{1, \dots, k\} \to \{1, \dots, k\}$, $\sigma : \{k+1, \dots, k+l\} \to \{k+1, \dots, k+l\}$, тобто σ є підстановкою також на множинах $\{1, \dots, k\}$ та $\{k+1, \dots, k+l\}$.

Введемо позначення $b_{ij} = a_{ij}$, $1 \leqslant i, j \leqslant k$, $c_{ij} = a_{k+i,k+j}$, $1 \leqslant i, j \leqslant l$, тоді $\Delta(A) = \sum\limits_{\sigma \in S_n} \varepsilon(\sigma) a_{\sigma(1)1} \dots a_{\sigma(k)k} a_{\sigma(k+1)k+1}\, a_{\sigma(k+l)k+l} =$

$\sum\limits_{\sigma_1 \in S_k, \sigma_2 \in S_l} \varepsilon(\sigma_1)\varepsilon(\sigma_2) b_{\sigma_1(1)1} \dots b_{\sigma_1(k)k} c_{\sigma_2(1)1}\, c_{\sigma_2(l)l} =$

$(\sum\limits_{\sigma_1 \in S_k} \varepsilon(\sigma_1) b_{\sigma_1(1)1} \dots b_{\sigma_1(k)k}) \times (\sum\limits_{\sigma_2 \in S_l} \varepsilon(\sigma_2) c_{\sigma_2(1)1}\, c_{\sigma_2(l)l}) = \Delta(B)\Delta(C)$. $\square$

Детермінант Вандермонда $\begin{vmatrix} 1 & 1 & \dots & 1 \\ x_1 & x_2 & \dots & x_n \\ \dots & \dots & \dots & \dots \\ x_1^{n-1} & x_2^{n-1} & \dots & x_n^{n-1} \end{vmatrix} = \prod\limits_{1 \leqslant i < j \leqslant n} (x_j - x_i).$