**КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ**
**імені ТАРАСА ШЕВЧЕНКА**

# Алгебра.

**Методичний посібник**
**для студентів радіофізичного факультету**
**напряму підготовки «Прикладна фізика»**

**Київ – 2015**

УДК 514.74(075.8)

Рецензент:

**Загальна алгебра. Методичний посібник для студентів радіофізичного факультету напрям підготовки «Прикладна фізика» / С.В. Єфіменко – К.: КНУ, 2015 – 103 с.**

Методичний посібник поданий у вигляді конспекту лекцій нормативного курсу «Загальна алгебра», що викладається студентам напряму підготовки «Прикладна фізика» на радіофізичному факультеті в об'ємі 34 лекційні години.

Слід зазначити, що донедавна цей курс (який фактично об'єднує дві математичні дисципліни – аналітична геометрія та основи лінійної алгебри) викладався в об'ємі 51 лекційна година.

Намагаючись не втратити наповненість курсу, обмеженого дуже скромною кількістю годин, крім мінімально обов'язкового матеріалу тут викладаються і додаткові теми, які будуть безумовно корисними для більш глибокого розуміння матеріалу. Відповідні розділи, набрані більш дрібними шрифтом, виходять за межі нинішнього навчального плану та призначені для самостійної роботи студентів.

Затверджено
Радою радіофізичного факультету,
Протокол № 7 від 2 березня 2015 року

## Лекція 1. Простір геометричних векторів. Векторний простір.

### 1. Вектори. Операції з векторами.

Геометричні вектори та дії з ними вивчаються в школі. Тут коротко нагадуються основні означення та поняття.

**<u>Означення 1.</u>** ***Геометричним вектором*** називатимемо напрямлений відрізок. Отже, вектор характеризується напрямком та довжиною і цілком визначається двома точками: одна задає початок вектора, друга – його кінець. Аналітично вектори можна позначати цими двома точками, наприклад, вектор $\mathbf{AB}$.

**<u>Означення 2.</u>** Два геометричних вектори називаються ***рівними***, якщо вони мають однакову довжину та напрямок – тобто лежать на паралельних прямих. (Це так звані ***вільні*** вектори з довільною точкою прикладення. У фізиці, крім них розглядаються ще ковзні вектори, тобто напрямлені вздовж однієї прямої).

Отже, рівні вектори можуть бути суміщені паралельним перенесенням, тобто точка-початок вектора не є визначальною, тому інколи вектор позначають однією літерою, наприклад, вектор $\mathbf{a}$ або $\mathbf{b}$, вважаючи, що всі вектори починаються в одній і тій самій точці.

Довжину вектора $\mathbf{a}$ позначатимемо таким чином: $|\mathbf{a}|$.

**<u>Означення 3.</u>** Нуль-вектором називають вектор $\mathbf{0}$, який має нульову довжину та невизначений напрямок.

**<u>Означення 4.</u>** Протилежним вектором до вектора $\mathbf{a}$ називають вектор $-\mathbf{a}$, такий що $\mathbf{a}+(-\mathbf{a})=\mathbf{0}$.

**<u>Означення 5.</u>** Вектори $\mathbf{a}$ та $\mathbf{b}$ називаються ***колінеарними*** (позначається цей факт так: $\mathbf{a} \Updownarrow \mathbf{b}$ ), якщо вони лежать на паралельних прямих.

У множині геометричних векторів визначені лінійні операції над ними – додавання векторів та множення вектора на скаляр (дійсне число). Правила визначення результуючого вектора при цих операціях добре відомі зі шкільної програми, тому не будемо тут їх повторювати. Зауважимо лише властивості цих операцій над геометричними векторами. Операція додавання векторів асоціативна та комутативна, існують нульовий вектор та вектор протилежний для довільного вектора. Операція множення на скаляр асоціативна відносно кількох множників. Виконуються також дистрибутивні закони.

Виявляється, що множина геометричних векторів є лише однією з реалізацій, моделлю більш загального математичного поняття – лінійного або векторного (на цей час для нас ці слова будуть синонімами!) простору. Дамо його точне означення.

**<u>Означення 6.</u>** Деяку множину елементів $L$ будемо називати ***векторним*** або ***лінійним простором*** над полем $\Re$ дійсних чисел, якщо в цій множині визначені операції додавання (+) елементів та множення (·) елементу на число:

1. $\forall \mathbf{a},\mathbf{b} \in L \;\; \exists\; \mathbf{c} \in L: \;\; \mathbf{c}=\mathbf{a}+\mathbf{b}$
2. $\forall \mathbf{a} \in L \;\; \exists\; \mathbf{c} \in L: \;\; \mathbf{c}=\lambda \cdot \mathbf{a}$

причому ці операції задовольняють наступні 8 аксіом:

1. $\forall \mathbf{a},\mathbf{b} \in L \quad \mathbf{a}+\mathbf{b}=\mathbf{b}+\mathbf{a}$;
2. $\forall \mathbf{a},\mathbf{b},\mathbf{c} \in L \quad (\mathbf{a}+\mathbf{b})+\mathbf{c}=\mathbf{a}+(\mathbf{b}+\mathbf{c})$;
3. $\exists \mathbf{0} \in L: \forall \mathbf{a} \in L \quad \mathbf{a}+\mathbf{0}=\mathbf{0}+\mathbf{a}=\mathbf{a}$;
4. $\forall \mathbf{a} \in L \;\; \exists(-\mathbf{a}) \in \mathrm{L}: \;\; \mathbf{a}+(-\mathbf{a})=\mathbf{0}$;
5. $\forall \mathbf{a} \in L \;\; \forall \lambda,\mu \in \Re \quad \lambda \cdot (\mu \cdot \mathbf{a})=(\lambda\mu)\cdot \mathbf{a}$;
6. $\forall \mathbf{a} \in L \quad 1 \cdot \mathbf{a}=\mathbf{a}$;
7. $\forall \mathbf{a} \in L \;\; \forall \lambda,\mu \in \Re \quad (\lambda+\mu)\cdot \mathbf{a}=\lambda\mathbf{a}+\mu\mathbf{a}$;
8. $\forall \mathbf{a},\mathbf{b} \in L \;\; \forall \lambda \in \Re \;\; \lambda \cdot (\mathbf{a}+\mathbf{b})=\lambda\mathbf{a}+\lambda\mathbf{b}$.

**Зауваження**. Якщо замість поля дійсних чисел всі вимоги сформульовані для поля комплексних чисел $C$, то маємо векторний простір над полем комплексних чисел, або комплексний векторний простір.

Всі елементи векторного простору називають ***векторами*** (інколи всупереч математичній або фізичній природі цих елементів).

Наведемо деякі приклади векторних просторів.

1. Множина дійсних чисел $\Re$ сама утворює векторний простір над полем дійсних чисел.
2. Множина комплексних чисел $C$ – комплексний векторний простір.
3. Множина квадратних матриць порядку $n$.
4. Множина поліномів (многочленів) степеню, що не перевищує $n$.
5. Нуль-вектор $\mathbf{0}$ утворює векторний простір і над полем дійсних і над полем комплексних чисел.

**Зауваження**. Із аксіом векторного простору випливає низка наслідків (наприклад, ***єдиність*** нульового та протилежного елементів та ряд інших).

## 2. Лінійно залежні та лінійно незалежні системи векторів.

Розглянемо довільний векторний простір $L$. Зафіксуємо в ньому деяку підмножину (систему) векторів $\mathbf{a}_k, k = 1,\dots, n$. З означення операцій над векторами цього простору випливає можливість утворювати ***лінійні комбінації векторів***, тобто розглядати вектори виду; $\sum\limits_{k=1}^{n}\alpha^k\mathbf{a}_k \in L$ – лінійна комбінація векторів $\mathbf{a}_k, k = 1,\dots,n$ простору $L$. Числа $\alpha^k, k = 1,\dots,n$ називають ***коефіцієнтами*** лінійної комбінації.

**Означення 7.** Лінійна комбінація називається ***тривіальною***, якщо всі її коефіцієнти рівні нулю.

**Зауваження**. Очевидно, що тривіальна лінійна комбінація рівна нуль-вектору.

**Означення 8.** Система векторів $\{\mathbf{a}_k\}_{k=1,\dots n}$ називається ***лінійно незалежною***, якщо лише тривіальна лінійна комбінація векторів цієї системи рівна нуль-вектору.

**Зауваження**. Іншими словами, система векторів $\{\mathbf{a}_k\}_{k=1,\dots n}$ лінійно незалежна, якщо з того, що $\sum\limits_{k=1}^{n}\alpha^k\mathbf{a}_k = \mathbf{0}$ випливає: $\alpha^1 = \alpha^2 = \dots = \alpha^n$.

**Означення 9.** Система векторів $\{\mathbf{a}_k\}_{k=1,\dots n}$ називається ***лінійно залежною***, якщо існує нетривіальна лінійна комбінація векторів цієї системи, рівна нуль-вектору.

**Приклади**.

1. $\{\mathbf{0}\}$ – лінійно залежна система, оскільки існує нетривіальна лінійна комбінація $1\cdot\mathbf{0} = \mathbf{0}$.
2. Нехай $\mathbf{a} \neq \mathbf{0}$. Тоді $\{\mathbf{a}\}$ – лінійно незалежна система, оскільки з рівності $\alpha\mathbf{a} = \mathbf{0}$ обов'язково випливає, що $\alpha = \mathbf{0}$.
3. Нехай $\mathbf{a}_1$ та $\mathbf{a}_2$ два ненульових не колінеарних геометричних вектора. Тоді $\{\mathbf{a}_1, \mathbf{a}_2\}$ – лінійно незалежна система векторів. Доведемо це від супротивного, тобто припустимо існування деякої нетривіальної лінійної комбінації цих векторів, яка рівна нулю: $\alpha^1\mathbf{a}_1 + \alpha^2\mathbf{a}_2 = \mathbf{0}$. Оскільки дана лінійна комбінація нетривіальна, то нехай приміром $\alpha^1 \neq 0$. Тоді $\mathbf{a}_1 = -\frac{\alpha^2}{\alpha^1}\mathbf{a}_2$, що означає колінеарність векторів $\mathbf{a}_1$ та $\mathbf{a}_2$. Одержана супе-

речність свідчить про лінійну незалежність системи векторів $\{\mathbf{a}_1,\mathbf{a}_2\}$.

**Вправа**. Показати, що серед геометричних векторів у просторі лінійно незалежними будуть лише трійки векторів, які не паралельні одній площині (іншими словами, не компланарні вектори).

**3. Властивості систем лінійно залежних та лінійно незалежних векторів.**

1. Система векторів $\{\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_k\}$, серед яких є нуль-вектор – лінійно залежна.
   Доведення. Дійсно, нехай маємо систему векторів $\{\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_k\}$, причому $\mathbf{a}_i=\mathbf{0}\left(1\le i\le k\right)$. Розглянемо лінійну комбінацію з коефіцієнтами $\alpha^1=\ldots=\alpha^{i-1}=\alpha^{i+1}=\ldots=\alpha^n,\alpha^i=1$. Вона нетривіальна, проте, вочевидь, рівна нуль-вектору.
2. ***Критерій лінійної залежності векторів***. Для того, щоб система векторів $\{\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_k\}$ була лінійно залежною, необхідно та достатньо, щоб принаймні один із векторів системи був лінійною комбінацією інших.
   ►(Необхідність). Нехай система векторів $\{\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_k\}$ лінійно залежна. Розглянемо деяку нетривіальну нульову лінійну комбінацію цих векторів: $\sum_{k=1}^{n}\alpha^k\mathbf{a}_k=\mathbf{0}$. Нехай коефіцієнт цієї комбінації, наприклад, $\alpha^1\ne 0$ (якщо ненульовим є інший коефіцієнт – перенумеруємо вектори та коефіцієнти відповідним чином ). Тоді маємо $\mathbf{a}_1=\sum_{k=2}^{n}\left(-\frac{\alpha^k}{\alpha^1}\right)\mathbf{a}_k$ – вектор $\mathbf{a}_1$ є лінійною комбінацією інших векторів системи.
   (Достатність). Нехай, наприклад, $\mathbf{a}_1=\sum_{k=2}^{n}\alpha^k\mathbf{a}_k$. Тоді маємо нетривіальну лінійну комбінацію цих векторів, рівну нуль-вектору: $1\cdot\mathbf{a}_1+\sum_{k=2}^{n}\left(-\alpha^k\right)\mathbf{a}_k=\mathbf{0}$.◄
3. Якщо серед $k$ векторів системи $\{\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_k\}$ є лінійно залежна підсистема із яких-небудь $m$ $(m\le k)$ векторів, то і вся система $\{\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_k\}$ – лінійно залежна.
   Доведення. Вважатиме, що перші $m$ векторів системи $\{\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_k\}$ утворюють лінійно залежну систему і розглянемо їх нульову нетривіальну лінійну комбінацію $\sum_{k=1}^{m}\alpha^k\mathbf{a}_k=\mathbf{0}$. Тоді лінійна комбінація векторів всієї системи $\sum_{k=1}^{m}\alpha^k\mathbf{a}_k+\sum_{k=m+1}^{n}0\cdot\mathbf{a}_k=\mathbf{0}$ є також нетривіальною, проте рівною нуль-вектору.
4. Будь-яка підсистема векторів лінійно незалежної системи $\{\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_k\}$ – лінійно незалежна.
   Доведення. Від супротивного припустимо протилежне: Нехай деякі $m$ векторів системи $\{\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_k\}$ є лінійно залежними. Тоді, згідно властивості 3, вся система $\{\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_k\}$ є також лінійно залежною, що суперечить умові.

**Означення 10.** Векторний простір $L$ називається ***$n$-вимірним***, якщо в ньому існує лінійно незалежна система із $n$ векторів, а будь-які $n+1$ векторів утворюють лінійно залежну систему. Таким чином, число $n$ визначає вимірність простору.

Для підкреслення вимірності векторного простору будемо позначати його із відповідним індексом $L_n$.

**Означення 11.** Векторний простір $L_\infty$ називається ***нескінченно вимірним***, якщо в ньому можна вибрати довільну кількість векторів, які утворюють лінійно незалежну систему.

**Приклади**.

1. У просторі, що складається лише з нуль вектора, не існує лінійно незалежних векторів, тому вимірність цього простору рівна нулю.
2. Множина геометричних векторів, колінеарних фіксованому вектору $\mathbf{a} \neq \mathbf{0}$, разом з нуль-вектором утворює одновимірний простір. Доведіть це самостійно.
3. Вимірність простору всіх геометричних векторів рівна 3. Адже вектори $\{\mathbf{i}, \mathbf{j}, \mathbf{k}\}$ прямокутної декартової системи координат (ПДСК) лінійно незалежні, а будь-який вектор, як відомо, може бути записаний у вигляді їх лінійної комбінації.
4. Множина поліномів степеню, що не перевищує $n$, утворює простір вимірності $n+1$. Дійсно $n$ поліномів $\{1, x, x^2, \ldots, x^n\}$ утворюють лінійно-незалежну систему, а будь-який інший поліном степеню не більшого за $n$, є їх лінійною комбінацією.

### 4. Базис скінчено вимірного векторного простору. Координати векторів.

**Означення 12.** ***Базисом*** $n$-вимірного векторного простору $L_n$ називається довільна ***впорядкована*** лінійно незалежна система із $n$ векторів цього простору.

**Зауваження 1.** З означення випливає, що у векторному просторі існує безліч базисів.

**Зауваження 2.** Базис називають ще впорядкованою ***максимально лінійно незалежною системою*** векторів у просторі. Слово «максимально» тут означає, що до системи базисних векторів неможливо приєднати жодного вектору простору так, щоб система залишалась лінійно незалежною.

**Теорема**. Кожний вектор $\mathbf{a} \neq \mathbf{0}$ $n$-вимірного векторного простору $L_n$ може бути поданий у вигляді лінійної комбінації векторів базису, причому таке подання єдине.

►Нехай $\{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_n\}$ – деякий базис векторного простору $L_n$. Нехай $\mathbf{a} \in L_n, \mathbf{a} \neq \mathbf{0}$. Тоді система векторів $\{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_n, \mathbf{a}\}$ є лінійно залежною, тобто існує нетривіальна лінійна комбінація: $\sum\limits_{k=1}^{n} \alpha^k \mathbf{f}_k + \alpha^0 \mathbf{a} = \mathbf{0}$. В цій лінійній комбінації коефіцієнт $\alpha^0 \neq 0$, інакше $\alpha^1 = \ldots = \alpha^n = 0$ і тоді система векторів $\{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_n, \mathbf{a}\}$– лінійно залежна, що суперечить умові. Отже, маємо: $\mathbf{a} = \sum\limits_{k=1}^{n} \left( -\dfrac{\alpha^k}{\alpha^0} \right) \mathbf{f}_k$, тобто вектор $\mathbf{a}$ є лінійною комбінацією базисних векторів. Покажемо, що таке подання єдине. Від супротивного припустимо, що існує два різних розклади вектора $\mathbf{a}$ по системі базисних векторів: $\mathbf{a} = \sum\limits_{k=1}^{n} \beta^k \mathbf{f}_k$ та $\mathbf{a} = \sum\limits_{k=1}^{n} \gamma^k \mathbf{f}_k$. Звідси випливає, що $\sum\limits_{k=1}^{n} \left(\beta^k - \gamma^k\right) \mathbf{f}_k = \mathbf{0}$, а оскільки вектори $\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_n$ лінійно незалежні, то $\beta^k - \gamma^k = 0 \ \ \forall k = \overline{1,n}$.◄

**Означення 13.** Якщо $\{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_n\}$ – базис векторного простору $L_n$ і $\mathbf{a} = \sum_{k=1}^{n} \alpha^k \mathbf{f}_k$ – розклад деякого вектору $\mathbf{a} \in L_n$ по базису $\{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_n\}$, то коефіцієнти цього розкладу $\alpha^1, \alpha^2, \ldots, \alpha^n$ називаються ***координатами вектора*** $\mathbf{a} \in L_n$ в базисі $\{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_n\}$.

З доведеної вище теореми випливає, що будь-який вектор простору однозначно визначається своїм набором координат у вибраному базисі. Це дозволяє повністю абстрагуватись від самої природи векторного простору $L_n$ і мати справу лише з наборами $n$ координат замість векторів. Крім того, це вказує на певну «схожість» всіх векторних просторів однакової розмірності. Пізніше ми доведемо теорему, в якій ця схожість називатиметься ***ізоморфізмом*** векторних просторів.

**Зауваження.** Координати вектора $\mathbf{a} \in L_n$ будемо записувати як вектор-стовпчик, позначаючи їх наступним чином: $|\mathbf{a}\rangle = \begin{pmatrix} \alpha^1 \\ \alpha^2 \\ \vdots \\ \alpha^n \end{pmatrix}$.

**Наслідки** з теореми.

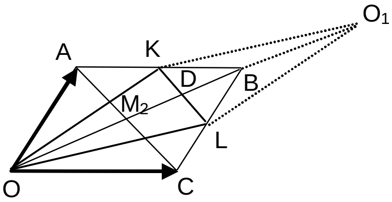
1. Координати будь-якого вектору простору у фіксованому базисі визначаються однозначно.
2. Два вектори рівні тоді і тільки тоді, коли рівні їх координати у фіксованому базисі.
3. Якщо $\mathbf{a} = \sum_{k=1}^{n} \alpha^k \mathbf{f}_k$ та $\mathbf{b} = \sum_{k=1}^{n} \beta^k \mathbf{f}_k$ розклади довільних векторів простру $L_n$ по базису $\{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_n\}$, то вектор $\mathbf{a} + \mathbf{b}$ в цьому базисі матиме координати $\alpha^1 + \beta^1, \ldots, \alpha^n + \beta^n$, а вектор $\lambda\mathbf{a}$ – координати $\lambda\alpha^1, \ldots, \lambda\alpha^n$. Доведіть цей факт самостійно.
4. Система векторів $\{\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_k\}$ простору $L_n$ лінійно незалежна тоді і тільки тоді, коли лінійно незалежна система вектор-стовпчиків їх координат. Доведіть це самостійно.

## 5. Системи координат в просторі геометричних векторів.

Повернемось у тривимірний простір геометричних векторів. Очевидно, базис в ньому утворюють довільні три не компланарні вектори (нагадаємо, що ***компланарними*** називаються три вектори, які паралельні одній площині). Отже, виберемо три не компланарних вектори та зведемо їх до спільного початку – деякої точки O. Одержимо ***загальну афінну систему координат***. Якщо три базисних вектори взаємно перпендикулярні, система координат називається ***прямокутною***. І нарешті, якщо у прямокутній системі координат базисні вектори мають одиничну довжину, маємо знайому із школи ПДСК – ***прямокутну декартову систему координат***. Базисні вектори в ній позначаються, як вже згадувалось вище, $\{\mathbf{i}, \mathbf{j}, \mathbf{k}\}$, а координати вектора називаються відповідно абсциса, ордината та апліката. Будемо позначати їх наступним чином: $\mathbf{a} = \{a^1, a^2, a^3\}$.

**Приклад**. Точки K та L – середини сторін AB та BC паралелограма OABC. Довести, що точка перетину діагоналей паралелограму співпадає з точкою перетину медіан трикутника OKL.

Ведемо систему координат на площині: точка O – початок, $\mathbf{OA}=\mathbf{a}$ та $\mathbf{OC}=\mathbf{b}$ – базисні вектори.




Позначимо $M_1$ – точку перетину діагоналей паралелограма OABC. Тоді за правилами додавання векторів маємо $\mathbf{OB}=\mathbf{OA}+\mathbf{OC}=\mathbf{a}+\mathbf{b}$, тому $\mathbf{OM_1}=\frac{1}{2}(\mathbf{a}+\mathbf{b})$. Тепер розглянемо трикутник OKL. Позначимо $M_2$ – точку перетину медіан цього трикутника (див. мал.). Маємо $\mathbf{OK}=\mathbf{OA}+\mathbf{AK}=\mathbf{a}+\frac{1}{2}\mathbf{b}$, $\mathbf{OL}=\mathbf{OC}+\mathbf{CL}=\frac{1}{2}\mathbf{a}+\mathbf{b}$.

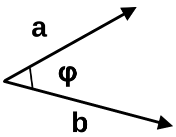
Для паралелограма $OKO_1L$ діагональ $OO_1$ визначається як сума сторін-векторів: $\mathbf{OO_1}=\mathbf{OK}+\mathbf{OL}=\frac{3}{2}(\mathbf{a}+\mathbf{b})$. Тоді медіана OD трикутника OKL – це половина даної діагоналі: $\mathbf{OD}=\frac{1}{2}\mathbf{OO_1}=\frac{3}{4}(\mathbf{a}+\mathbf{b})$ . Як відомо медіани трикутника в точці перетину діляться у співвідношенні 2:1, тому для точки їх перетину справедлива рівність: $\mathbf{OM_2}=\frac{2}{3}\mathbf{OD}=\frac{1}{2}(\mathbf{a}+\mathbf{b})$, що означає векторну рівність $\mathbf{OM}_1=\mathbf{OM_2}$, тому точки $M_1$ та $M_2$ співпадають.

## Лекція 2. Добутки геометричних векторів.

В даному розділі будемо розглядати геометричні вектори із простору $L=\Re^3$.

### 1. Скалярний добуток векторів.

<u>**Означення 1.**</u> ***Скалярним добутком*** двох векторів $\mathbf{a},\mathbf{b}\in\Re^3$ називається число $\mathbf{a}\cdot\mathbf{b}$, рівне: $\mathbf{a}\cdot\mathbf{b}=|\mathbf{a}|\cdot|\mathbf{b}|\cdot\cos\varphi$, де $\varphi$ – кут утворений векторами $\mathbf{a}$ та $\mathbf{b}$, а через $|\mathbf{a}|$ позначено довжину вектора $\mathbf{a}$.




Неважко зрозуміти, що скалярний добуток двох ненульових векторів рівний 0 тоді і тільки тоді, коли ***вектори ортогональні***: $\mathbf{a}\cdot\mathbf{b}=0 \Leftrightarrow \varphi=\pi/2$

<u>**Означення 2.**</u> ***Ортом*** вектора $\mathbf{a}\in\Re^3$ називається вектор $\mathbf{e_a}$ одиничної довжини співнапрямлений з вектором $\mathbf{a}$. Тобто $\mathbf{e_a}=\frac{1}{|\mathbf{a}|}\mathbf{a}$.

<u>**Означення 3.**</u> ***Проекцією*** вектора $\mathbf{a}\in\Re^3$ на напрямок вектора $\mathbf{b}\in\Re^3,\mathbf{b}\neq\mathbf{0}$ називається число $np_{\mathbf{b}}\mathbf{a}\cdot=\mathbf{a}\cdot\mathbf{e_b}=|\mathbf{a}|\cdot\cos\varphi$. (тут $\varphi$ – кут між векторами $\mathbf{a}$ та $\mathbf{b}$) Звідси легко одержати, що $\mathbf{a}\cdot\mathbf{b}=np_{\mathbf{b}}\mathbf{a}\cdot|\mathbf{b}|=np_{\mathbf{a}}\mathbf{b}\cdot|\mathbf{a}|$

<u>**Зауваження**</u>. З означення випливає, що проекція додатна, якщо кут $\varphi$ між векторами $\mathbf{a}$ та $\mathbf{b}$ гострий, та від'ємна, у разі, коли кут $\varphi$ тупий.

***Властивості скалярного добутку векторів***.

1. $\mathbf{a}\cdot\mathbf{b}=\mathbf{b}\cdot\mathbf{a}\ \forall\mathbf{a},\mathbf{b}\in\Re^3$ – комутативність скалярного добутку безпосередньо випливає з його означення;

2. $(\lambda\mathbf{a})\cdot\mathbf{b}=\lambda(\mathbf{a}\cdot\mathbf{b})\ \forall\mathbf{a},\mathbf{b}\in\Re^3\ \ \forall\lambda\in\Re$ – скалярний множник можна виносити за знак скалярного добутку (доведіть це самостійно);
3. $(\mathbf{a}+\mathbf{b})\cdot\mathbf{c}=\mathbf{a}\cdot\mathbf{c}+\mathbf{b}\cdot\mathbf{c}\ \forall\mathbf{a},\mathbf{b},\mathbf{c}\in\Re^3$ – дистрибутивність скалярного добутку випливає із властивостей проекцій, які вивчались у курсі шкільної програми: $(\mathbf{a}+\mathbf{b})\cdot\mathbf{c}=np_{\mathbf{c}}(\mathbf{a}+\mathbf{b})\cdot|\mathbf{c}|=(np_{\mathbf{c}}\mathbf{a}+np_{\mathbf{c}}\mathbf{b})\cdot|\mathbf{c}|=np_{\mathbf{c}}\mathbf{a}\cdot|\mathbf{c}|+np_{\mathbf{c}}\mathbf{b}\cdot|\mathbf{c}|=\mathbf{a}\cdot\mathbf{c}+\mathbf{b}\cdot\mathbf{c}$
4. $|\mathbf{a}|=\sqrt{\mathbf{a}\cdot\mathbf{a}}\ \ \forall\mathbf{a}\in\Re^3$ – скалярний добуток визначає квадрат довжини вектора.

Нехай у просторі геометричних векторів вибрана деяка ПДСК. Як визначити скалярний добуток векторів $\mathbf{a}$ та $\mathbf{b}$, якщо відомі їх координати? Отже, нехай маємо вектори $\mathbf{a}=\{a^1,a^2,a^3\}$ та $\mathbf{b}=\{b^1,b^2,b^3\}$. Визначимо скалярний добуток цих векторів, скориставшись властивостями скалярного добутку:

$$\mathbf{a}\cdot\mathbf{b}=\left(a^1\mathbf{i}+a^2\mathbf{j}+a^3\mathbf{k}\right)\cdot\left(b^1\mathbf{i}+b^2\mathbf{j}+b^3\mathbf{k}\right)=$$
$$=a^1b^1\mathbf{i}\cdot\mathbf{i}+a^2b^2\mathbf{j}\cdot\mathbf{j}+a^3b^3\mathbf{k}\cdot\mathbf{k}=a^1b^1+a^2b^2+a^3b^3$$

**Зауваження**. Зверніть увагу, що одержана формула має місце ***лише для ПДСК***. У загальній афінній системі координат ми мали б враховувати кути між базисними векторами.

Таким, чином ***у прямокутній декартовій системі координат*** справедлива наступна формула скалярного добутку двох векторів:

$$\boxed{\mathbf{a}\cdot\mathbf{b}=a^1b^1+a^2b^2+a^3b^3} \qquad (1)$$

Формула (1) дає важливі наслідки, зокрема, довжина вектора визначається через його координати наступним чином: $|\mathbf{a}|=\sqrt{\left(a^1\right)^2+\left(a^2\right)^2+\left(a^2\right)^2}\ \ \forall\mathbf{a}\in\Re^3$ (2)

Для кута між двома векторами маємо таку формулу:

$$\cos\varphi=\frac{\mathbf{a}\cdot\mathbf{b}}{|\mathbf{a}|\cdot|\mathbf{b}|}=\frac{a^1b^1+a^2b^2+a^3b^3}{\sqrt{\left(a^1\right)^2+\left(a^2\right)^2+\left(a^2\right)^2}\sqrt{\left(b^1\right)^2+\left(b^2\right)^2+\left(b^2\right)^2}} \qquad (3)$$

**Означення 4.** ***Напрямними косинусами*** вектора $\mathbf{a}\in\Re^3$ називаються косинуси кутів, утворених вектором $\mathbf{a}$ з ортами ПДСК. Ці кути позначаються відповідно $\alpha$, $\beta$ та $\gamma$, отже, з формули (3) випливає: $\cos\alpha=\frac{\mathbf{a}\cdot\mathbf{i}}{|\mathbf{a}|}=\frac{a^1}{|\mathbf{a}|}$, аналогічно $\cos\beta=\frac{\mathbf{a}\cdot\mathbf{j}}{|\mathbf{a}|}=\frac{a^2}{|\mathbf{a}|}$ та $\cos\gamma=\frac{\mathbf{a}\cdot\mathbf{k}}{|\mathbf{a}|}=\frac{a^3}{|\mathbf{a}|}$. Звідси маємо основну властивість напрямних косинусів:

$$\cos^2\alpha+\cos^2\beta+\cos^2\gamma=1 \qquad (4)$$

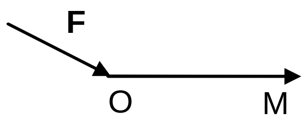
Таким чином, якщо $|\mathbf{a}|=1$, то $\cos\alpha=a^1,\cos\beta=a^2,\cos\gamma=a^3$ – тобто, напрямними косинусами орта є його координати.

**Приклади**.

1. Знайдемо орт вектора $\mathbf{a}=\{1,1,1\}$. Маємо $|\mathbf{a}|=\sqrt{1+1+1}=\sqrt{3}$, тому $\mathbf{e}_{\mathbf{a}}=\left\{1/\sqrt{3},1/\sqrt{3},1/\sqrt{3}\right\}$.
2. Кут між ортами $\mathbf{a}$ і $\mathbf{b}$ рівний $\pi/3$. Знайти скалярний добуток цих векторів. Знаходимо за означенням $\mathbf{a}\cdot\mathbf{b}=|\mathbf{a}|\cdot|\mathbf{b}|\cdot\cos\varphi=1\cdot1\cdot\cos\pi/3=\sqrt{3}/2$.

3. Нехай $\mathbf{a}=\{-1,2,3\}$, $\mathbf{b}=\{2,0,1\}$. Знайдемо кут між цими векторами. З формули (3) одержимо $\cos\varphi=\dfrac{-2+0+3}{\sqrt{1+4+9}\sqrt{4+1}}=\dfrac{1}{\sqrt{70}}$ , отже $\varphi=\arccos\dfrac{1}{\sqrt{70}}$ .
4. Знайдемо скалярний добуток $(2\mathbf{a}-3\mathbf{b})\cdot(3\mathbf{a}+2\mathbf{b})$ для векторів з прикладу 3. Виконаємо спочатку дії над векторами:
$(2\mathbf{a}-3\mathbf{b})\cdot(3\mathbf{a}+2\mathbf{b})=6\mathbf{a}^2-6\mathbf{b}^2-5\mathbf{a}\cdot\mathbf{b}=64-30-5=29$.

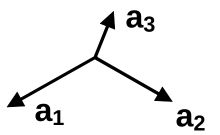
**<u>Фізичний зміст скалярного добутку</u>.**



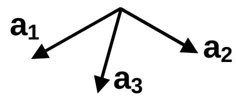

Якщо під впливом сили $\mathbf{F}$ деяка матеріальна точка переміщається з положення $O$ у положення $M$ , то робота $A$ цієї сили по переміщенню даної точки рівна скалярному добутку сили та вектора переміщення точки: $A=\mathbf{F}\cdot\mathbf{OM}$ (див мал.).

**2. Векторний добуток векторів.**

**<u>Означення 5.</u>** Впорядкована трійка некомпланарних векторів $\mathbf{a}_1,\mathbf{a}_2,\mathbf{a}_3\in\Re^3$ , приведених до спільного початку, називається ***правою (лівою)***, якщо по цим векторам можна спрямувати відповідно великий, вказівний та середній пальці ***правої (лівої) руки***. Або ж при повороті на найменший кут від вектору $\mathbf{a}_1$ до вектору $\mathbf{a}_2$ напрямок вектору $\mathbf{a}_3$ відповідатиме руху ***правого (лівого)*** гвинта. (див. мал. нижче). Зрозуміло, що при зміні напрямку одного з векторів, або при зміні порядку нумерації двох з векторів трійка міняє свою орієнтацію на протилежну. Завжди вважатимемо, що орти ПДСК мають ***праву*** орієнтацію.




Права трійка векторів




Ліва трійка векторів

**<u>Означення 6.</u>** Векторним добутком двох векторів $\mathbf{a},\mathbf{b}\in\Re^3$ називається вектор $\mathbf{c}$, що визначається трьома умовами:

1. $|\mathbf{c}|=|\mathbf{a}|\cdot|\mathbf{b}|\cdot\sin\varphi$ де $\varphi$ – кут утворений векторами $\mathbf{a}$ та $\mathbf{b}$;
2. вектор ортогональний до обох векторів $\mathbf{a}$ та $\mathbf{b}$;
3. вектори $\mathbf{a},\mathbf{b},\mathbf{c}$ утворюють праву трійку.

Для векторного добутку використовуватимемо позначення $\mathbf{c}=\mathbf{a}\times\mathbf{b}$.

***Властивості векторного добутку векторів*.**

1. $\mathbf{a}\times\mathbf{b}=-\mathbf{b}\times\mathbf{a}\ \forall\mathbf{a},\mathbf{b}\in\Re^3$ – антикомутативність векторного добутку є наслідком зміни орієнтації трійки векторів при зміні їх порядку;
2. $(\lambda\mathbf{a})\times\mathbf{b}=\mathbf{a}\times(\lambda\mathbf{b})=\lambda(\mathbf{a}\times\mathbf{b})\ \forall\mathbf{a},\mathbf{b}\in\Re^3\ \forall\lambda\in\Re$ – скалярний множник можна виносити за знак векторного добутку (доведіть це самостійно);
3. $(\mathbf{a}+\mathbf{b})\times\mathbf{c}=\mathbf{a}\times\mathbf{c}+\mathbf{b}\times\mathbf{c}\ \forall\mathbf{a},\mathbf{b},\mathbf{c}\in\Re^3$ – дистрибутивність векторного добутку <u>буде доведена дещо пізніше</u>;
4. $|\mathbf{a}\times\mathbf{b}|$ дорівнює площі паралелограма, побудованого на векторах $\mathbf{a}$ і $\mathbf{b}$;
5. $\mathbf{a}\times\mathbf{b}=\mathbf{0}$ тоді і тільки тоді, коли вектори $\mathbf{a}$ і $\mathbf{b}$ колінеарні або принаймні один з них нульовий (доведіть це самостійно).

Знайдемо вираз векторного добутку через координати векторів-множників. Розглянемо вектори $\mathbf{a}=\{a^1,a^2,a^3\}$ та $\mathbf{b}=\{b^1,b^2,b^3\}$. Визначимо їх векторний добуток, скориставшись властивостями векторного добутку та врахувавши, що $\mathbf{i}\times\mathbf{j}=\mathbf{k},\ \mathbf{j}\times\mathbf{k}=\mathbf{i},\ \mathbf{k}\times\mathbf{i}=\mathbf{j}$: $\mathbf{a}\times\mathbf{b}=\left(a^1\mathbf{i}+a^2\mathbf{j}+a^3\mathbf{k}\right)\times\left(b^1\mathbf{i}+b^2\mathbf{j}+b^3\mathbf{k}\right)=$

$$=a^2b^3\mathbf{i}-a^3b^2\mathbf{i}-a^1b^3\mathbf{j}+a^3b^1\mathbf{j}+a^1b^2\mathbf{k}-a^2b^1\mathbf{k} \qquad (5)$$

Якщо пригадати вигляд формули для обчислення визначника матриці розмірності 3, то можна записати:

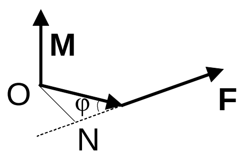
$$\boxed{\mathbf{a}\times\mathbf{b}=\begin{vmatrix}\mathbf{i} & \mathbf{j} & \mathbf{k}\\ a^1 & a^2 & a^3\\ b^1 & b^2 & b^3\end{vmatrix}} \qquad (6)$$

**<u>Приклади</u>**.

5. Знайдемо площу трикутника, побудованого на векторах з прикладу 3.
Для цього визначимо спершу векторний добуток даних векторів:
$$\mathbf{a}\times\mathbf{b}=\begin{vmatrix}\mathbf{i} & \mathbf{j} & \mathbf{k}\\ -1 & 2 & 3\\ 2 & 0 & 1\end{vmatrix}=2\mathbf{i}+7\mathbf{j}-4\mathbf{k}=\{2,7,-4\}$$
Модуль цього вектора визначає площу паралелограма, побудованого на векторах $\mathbf{a}$ і $\mathbf{b}$, отже шукана площа трикутника складає половину цієї величини: $S=\frac{1}{2}\left|\{2,7,-4\}\right|=\frac{1}{2}\sqrt{4+49+16}=\frac{\sqrt{69}}{2}$ кв. од.

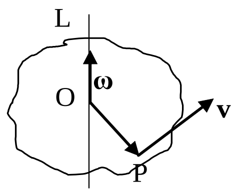
## <u>Фізичний зміст векторного добутку</u>.

1. Одним з основних понять статики є момент сили $\mathbf{F}$, прикладеної до деякої точки P відносно фіксованої точки O.




***Моментом сили*** $\mathbf{F}$, прикладеної до точки P відносно фіксованої точки O, називається вектор $\mathbf{M}=\mathbf{r}\times\mathbf{F}$, де $\mathbf{r}=\mathbf{OP}$ – радіус-вектор точки P. За величиною момент сили рівний $|\mathbf{M}|=|\mathbf{F}|\cdot|\mathbf{OP}|\sin\varphi=|\mathbf{F}|\cdot|\mathbf{ON}|$ – добуток величини сили на плече.

2. Якщо тверде тіло обертається навколо нерухомої осі $L$ зі сталою кутовою швидкістю $\boldsymbol{\omega}$, то миттєва швидкість $v$ довільної точки P цього тіла, як відомо, визначається ***формулою Ейлера***: $\mathbf{v}=\boldsymbol{\omega}\times\mathbf{r}$, де $\mathbf{r}=\mathbf{OP}$ – радіус-вектор точки P відносно полюса O на осі обертання $L$. Напрямлений вектор швидкості по дотичній до кола, по якому рухається точка P у бік обертання.

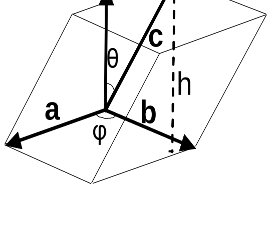



## 3. Мішаний добуток трьох векторів.

**<u>Означення 7.</u>** ***Мішаним*** добутком впорядкованої трійки векторів $\mathbf{a},\mathbf{b},\mathbf{c}$ називається число, рівне $(\mathbf{a}\times\mathbf{b})\cdot\mathbf{c}$.

Якщо кут між векторами $\mathbf{a}$ та $\mathbf{b}$ позначити $\varphi$, а кут між векторним добутком $\mathbf{a}\times\mathbf{b}$ та вектором $\mathbf{c}$ – $\theta$, то значення мішаного добутку можна обрахувати наступним чином: $$(\mathbf{a}\times\mathbf{b})\cdot\mathbf{c}=|\mathbf{a}|\cdot|\mathbf{b}|\cdot|\mathbf{c}|\sin\varphi\cos\theta \qquad (7)$$

***Властивості мішаного добутку векторів***.




1. Мішаний добуток ***правої трійки*** векторів $\mathbf{a},\mathbf{b},\mathbf{c}$ дорівнює об’єму паралелепіпеда, побудованого на цих векторах. Це очевидно (див. мал.), оскільки площа

основи такого паралелепіпеда рівна $|\mathbf{a}|\cdot|\mathbf{b}|\sin\varphi$, а висота $h$ рівна $|\mathbf{c}|\cos\theta$. Отже, $V=|\mathbf{a}|\cdot|\mathbf{b}|\cdot|\mathbf{c}|\sin\varphi\cos\theta=(\mathbf{a}\times\mathbf{b})\cdot\mathbf{c}$. Проте, у випадку лівої трійки векторів кут $\theta$ виявиться тупим і тоді $V=-(\mathbf{a}\times\mathbf{b})\cdot\mathbf{c}$, отже в загальному випадку, $V=|(\mathbf{a}\times\mathbf{b})\cdot\mathbf{c}|$. Крім того, очевидно, що знак мішаного добутку визначає орієнтацію трійки векторів: якщо $(\mathbf{a}\times\mathbf{b})\cdot\mathbf{c}>0$, то трійка векторів права, інакше трійка ліва. Якщо ж вектори компланарні, тобто паралельні одній площині, то висота такого «паралелепіпеда» рівна 0, отже нулю рівний і його об'єм.

2. $(\mathbf{a}\times\mathbf{b})\cdot\mathbf{c}=\mathbf{a}\cdot(\mathbf{b}\times\mathbf{c})$ (8)

Дійсно, площу паралелограма можна визначити, взявши за основу іншу грань, наприклад, утворену векторами $\mathbf{b}$ та $\mathbf{c}$ . Тоді маємо $V=|(\mathbf{b}\times\mathbf{c})\cdot\mathbf{a}|=|\mathbf{a}\cdot(\mathbf{b}\times\mathbf{c})|=|(\mathbf{a}\times\mathbf{b})\cdot\mathbf{c}|$. А оскільки орієнтація трійки $\mathbf{a},\mathbf{b},\mathbf{c}$ така сама, як і трійки $\mathbf{b},\mathbf{c},\mathbf{a}$, то таким чином, бачимо, що у мішаному добутку векторів $\mathbf{a},\mathbf{b},\mathbf{c}$ не має значення, до якої пари множників $\mathbf{a}$ і $\mathbf{b}$ чи $\mathbf{b}$ і $\mathbf{c}$ застосовувати векторне множення. Власне, тому мішаний добуток позначають просто $\mathbf{abc}$, оскільки важливим виявляється лише порядок множників.

3. $\mathbf{abc}=\mathbf{bca}=\mathbf{cab}=-\mathbf{acb}=-\mathbf{cba}=-\mathbf{bac}$ (9)

Дана властивість легко випливає із узагальнення властивості 2 та антикомутативності векторного добутку.

4. $\mathbf{abc}=0$ тоді і тільки тоді, коли вектори $\mathbf{a},\mathbf{b},\mathbf{c}$ компланарні.

Достатність цього твердження була обґрунтована при доведенні першої властивості, тому доведемо тут необхідність. Отже, нехай $\mathbf{abc}=0$. Якщо принаймні один з множників є нуль-вектором, то трійка напевне компланарна, тому далі вважатимемо, що всі вектори-множникі ненульові. У такому випадку із формули (7) випливає, що кут $\theta=\dfrac{\pi}{2}$ – це означає компланарність векторів, або ж кут $\varphi=0$. В цьому разі вектори $\mathbf{a}$ і $\mathbf{b}$ колінеарні, що також означає компланарність векторів $\mathbf{a},\mathbf{b},\mathbf{c}$.

5. $(\lambda\mathbf{a})\mathbf{bc}=\mathbf{a}(\lambda\mathbf{b})\mathbf{c}=\mathbf{ab}(\lambda\mathbf{c})=\lambda\mathbf{abc}\ \ \forall\mathbf{a},\mathbf{b},\mathbf{c}\in\Re^3\ \ \forall\lambda\in\Re$

Дана властивість є наслідком відповідних властивостей скалярного та векторного добутків.

6. $(\mathbf{a}_1+\mathbf{a}_2)\mathbf{bc}=\mathbf{a}_1\mathbf{bc}+\mathbf{a}_2\mathbf{bc}\ \ \forall\mathbf{a}_1,\mathbf{a}_2,\mathbf{b},\mathbf{c}\in\Re^3$

$\mathbf{a}(\mathbf{b}_1+\mathbf{b}_2)\mathbf{c}=\mathbf{ab}_1\mathbf{c}+\mathbf{ab}_2\mathbf{c}\ \ \forall\mathbf{a},\mathbf{b}_1,\mathbf{b}_2,\mathbf{c}\in\Re^3$

$\mathbf{ab}(\mathbf{c}_1+\mathbf{c}_2)=\mathbf{abc}_1+\mathbf{abc}_2\ \ \forall\mathbf{a},\mathbf{b},\mathbf{c}_1,\mathbf{c}_2\in\Re^3$ – дистрибутивність мішаного добутку. Очевидно, що доведення потребує дистрибутивність лише відносно одного з множників, на інші множники вона перенесеться автоматично завдяки властивості 3. Отже, розглянемо

$(\mathbf{a}_1+\mathbf{a}_2)\mathbf{bc}=((\mathbf{a}_1+\mathbf{a}_2)\times\mathbf{b})\cdot\mathbf{c}=(\mathbf{a}_1+\mathbf{a}_2)\cdot(\mathbf{b}\times\mathbf{c})$

Скористаємось далі доведеною вже дистрибутивністю скалярного добутку:

$(\mathbf{a}_1+\mathbf{a}_2)\cdot(\mathbf{b}\times\mathbf{c})=\mathbf{a}_1\cdot(\mathbf{b}\times\mathbf{c})+\mathbf{a}_2\cdot(\mathbf{b}\times\mathbf{c})=\mathbf{a}_1\mathbf{bc}+\mathbf{a}_2\mathbf{bc}$ – що й треба було довести.

Повернемось тепер до доведення дистрибутивності векторного добутку. Розглянемо вектор

$\mathbf{d}=(\mathbf{a}+\mathbf{b})\times\mathbf{c}=\{d^1,d^2,d^3\}$. Його координати – це проекції цього вектора на відповідні орти ПДСК, тому $d^1=\mathbf{d}\cdot\mathbf{i}=((\mathbf{a}+\mathbf{b})\times\mathbf{c})\cdot\mathbf{i}=(\mathbf{a}+\mathbf{b})\mathbf{ci}=\mathbf{aci}+\mathbf{bci}=$ $=(\mathbf{a}\times\mathbf{c})\cdot\mathbf{i}+(\mathbf{b}\times\mathbf{c})\cdot\mathbf{i}=(\mathbf{a}\times\mathbf{c})^1+(\mathbf{b}\times\mathbf{c})^1$

Отже, маємо рівність координат $d^1=(\mathbf{a}+\mathbf{b})\times\mathbf{c}^1=(\mathbf{a}\times\mathbf{c})^1+(\mathbf{b}\times\mathbf{c})^1$. Аналогічно доводиться і рівність двох інших координат векторних добутків $(\mathbf{a}+\mathbf{b})\times\mathbf{c}$ та $\mathbf{a}\times\mathbf{c}+\mathbf{b}\times\mathbf{c}$ . Тим самим рівність $(\mathbf{a}+\mathbf{b})\times\mathbf{c}=\mathbf{a}\times\mathbf{c}+\mathbf{b}\times\mathbf{c}$ повністю доведено.

Знайдемо тепер вираз мішаного добутку через координати векторів-множників. Розглянемо вектори $\mathbf{a}=\{a^1,a^2,a^3\}$, $\mathbf{b}=\{b^1,b^2,b^3\}$ та $\mathbf{c}=\{c^1,c^2,c^3\}$. Визначимо їх мішаний добуток.

Оскільки $\mathbf{a}\times\mathbf{b}=\begin{vmatrix}\mathbf{i}&\mathbf{j}&\mathbf{k}\\a^1&a^2&a^3\\b^1&b^2&b^3\end{vmatrix}=\{(\mathbf{a}\times\mathbf{b})^1,(\mathbf{a}\times\mathbf{b})^2,(\mathbf{a}\times\mathbf{b})^3\}$, а

$\mathbf{abc}=(\mathbf{a}\times\mathbf{b})\cdot\mathbf{c}=(\mathbf{a}\times\mathbf{b})^1c^1+(\mathbf{a}\times\mathbf{b})^2c^2+(\mathbf{a}\times\mathbf{b})^3c^3$, то

$$\mathbf{abc}=\begin{vmatrix}c^1&c^2&c^3\\a^1&a^2&a^3\\b^1&b^2&b^3\end{vmatrix}=\mathbf{cab}=\begin{vmatrix}a^1&a^2&a^3\\b^1&b^2&b^3\\c^1&c^2&c^3\end{vmatrix}.$$

Остаточно одержуємо формулу, яку легко запам'ятати: $$\mathbf{abc}=\begin{vmatrix}a^1&a^2&a^3\\b^1&b^2&b^3\\c^1&c^2&c^3\end{vmatrix}\qquad(10)$$

**Зауваження**. З формули (10) та властивості 3 мішаного добутку випливає, що визначник матриці розмірності 3 не зміниться при циклічній перестановці рядків і змінить знак на протилежний при перестановці двох рядків.

**Приклади**.

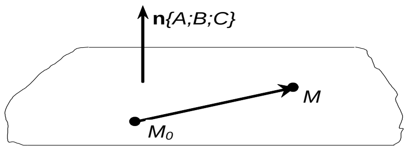
6. Знайдемо об'єм тетраедра з вершинами $A=(-4,-3,9); B(3,4,2); C(5,2,-1); D(7,4,8)$.

Даний тетраедр утворений векторами $\mathbf{AB}=\{7,7,-7\}; \mathbf{AC}=\{9,5,-10\}; \mathbf{AD}=\{11,7,-1\}$. Оскільки орієнтація цієї трійки нам невідома, об'єм тетраедра визначатимемо як $1/6$ модуля мішаного добутку векторів $\mathbf{AB},\mathbf{AC},\mathbf{AD}$:

$$V=\frac{1}{6}\left|\begin{vmatrix}7&7&-7\\9&5&-10\\11&7&-1\end{vmatrix}\right|=\frac{1}{6}\left|-35-441-770+385+63+490\right|=\frac{308}{6}=\frac{154}{3}\,\text{кв. од.}$$

## Лекція 3. Площина.

Поняття площини є одним із первісних геометричних понять. Вважатимемо, що у просторі обрана та зафіксована деяка ПДСК. Якщо визначити напрямок з допомогою довільного вектора, наприклад, $\mathbf{n}=\{A;B;C\}$, то тим самим визначена множина площин, перпендикулярних даному напрямку. Вектор $\mathbf{n}$ будемо називати ***нормаллю*** цих площин. Щоб визначити серед цієї множини конкретну площину,




досить вказати точку $M_0(x_0, y_0, z_0)$ на ній (див рис.). Позначимо $M(x, y, z)$ довільну точку цієї ж площини. Тоді вектори $\mathbf{M_0M}$ та $\mathbf{n}$ будуть ортогональними, а отже, $(\mathbf{r}-\mathbf{r_0})\cdot\mathbf{n}=0$ (1)

Це векторне рівняння площини. Тут $\mathbf{r}$ та $\mathbf{r_0}$ – відповідно радіус-вектори точок $M$ та $M_0$. Записавши рівність (1) у координатній формі, матимемо ***рівняння площини, заданої точкою та нормаллю***:

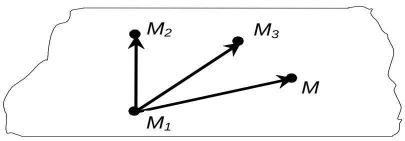
$$A(x-x_0)+B(y-y_0)+C(z-z_0)=0 \qquad (2)$$

Поклавши $-Ax_0-By_0-Cz_0=D$, зведемо рівняння (2), до рівняння виду

$$Ax+By+Cz+D=0 \qquad (3)$$

Покажемо, що й навпаки, довільне рівняння першого порядку (3) визначає деяку площину у просторі. З цією метою розглянемо точку $M_0(x_0, y_0, z_0)$, яка задовольняє рівняння (3). Для неї виконана рівність $Ax_0+By_0+Cz_0+D=0$. Віднявши цю рівність від рівняння (3), одержимо рівність (2), яка виражає умову ортогональності вектора $\mathbf{M_0M}$, що належить розглядуваному геометричному об'єкту, та деякого фіксованого вектора $\mathbf{n}=\{A;B;C\}$. Отже, рівняння (3) задає деяку площину. Воно називається ***загальним рівнянням площини***. Дослідимо частинні випадки цього рівняння, коли деякі коефіцієнти в ньому рівні нулеві.

1. Припустимо, що $A=0$. Тоді вектор нормалі площини, очевидно, ортогональний до осі OX, а сама площина їй паралельна.
2. Припустимо, що $A=B=0$. Тоді площина паралельна обом осям і OX, і OY, тобто паралельна координатній площині XOY і перпендикулярна осі OZ.
3. Нехай $D=0$. Дана площина проходить через початок координат – точку $O(0,0,0)$.
4. Припустимо, що $A=D=0$ – площина паралельна осі OX та проходить через початок координат, отже, вісь OX належить площині.
5. Якщо $A=B=D=0$, то маємо рівняння $z=0$, яке задає координатну площину XOY.




Площина також може бути визначена трьома своїми точками. Нехай $M_1(x_1, y_1, z_1)$, $M_2(x_2, y_2, z_2)$ та $M_3(x_3, y_3, z_3)$ – три точки відповідно з радіус-векторами $\mathbf{r_1}$, $\mathbf{r_2}$ та $\mathbf{r_3}$, які визначають деяку площину (див. рис.). Щоб записати її рівняння розглянемо довільну точку $M(x, y, z)$ цієї ж площини. Тоді вектори $\mathbf{r}-\mathbf{r_1}$, $\mathbf{r_2}-\mathbf{r_1}$ та $\mathbf{r_3}-\mathbf{r_1}$ – компланарні, отже, їх мішаний добуток нульовий, тобто

$$(\mathbf{r}-\mathbf{r_1})(\mathbf{r_2}-\mathbf{r_1})(\mathbf{r_3}-\mathbf{r_1})=0 \qquad (4)$$

є векторним ***рівнянням площини, заданої трьома точками***. Відповідне рівняння у координатній формі має вигляд:
$$\begin{vmatrix} x-x_1 & y-y_1 & z-z_1 \\ x_2-x_1 & y_2-y_1 & z_2-z_1 \\ x_3-x_1 & y_3-y_1 & z_3-z_1 \end{vmatrix}=0 \qquad (5)$$

<u>Приклад 1</u>. Знайти рівняння площини, що проходить через початок координат перпендикулярно до вектора $\mathbf{n}=(1;-1;2)$.

Заданий вектор $\mathbf{n}$ є нормаллю даної площини, а оскільки початок координат належить площині, в загальному рівнянні (3) слід покласти $D=0$. Таким чином, шукане рівняння має вигляд $x-y+2z=0$.

Приклад 2. Скласти рівняння площини, що проходить через три дані точки $A(1,2,3)$, $B(0,-1,2)$ та $C(-2,1,0)$.

Запишемо рівняння площини у вигляді (5): $\begin{vmatrix} x-1 & y-2 & z-3 \\ -1 & -3 & -1 \\ -3 & -1 & -3 \end{vmatrix}=0$, або $8(x-1)-8(z-3)=0$, або нарешті, $x-z+2=0$.

Припустимо, що площина (3) не проходить через початок координат, тобто коефіцієнт $D\neq 0$, тоді рівняння площини можна подати у вигляді:

$$\frac{Ax}{-D}+\frac{By}{-D}+\frac{Cz}{-D}=1 \text{, або } \qquad \frac{x}{a}+\frac{y}{b}+\frac{z}{c}=1, \tag{6}$$

де $a=\dfrac{-D}{A}, b=\dfrac{-D}{B}, c=\dfrac{-D}{C}$ – відповідно точок перетину площини (3) з осями OX, OY та OZ. Зверніть увагу, якщо, наприклад, $A=0$, то координату точки перетину площини із віссю OX визначити неможливо – площина їй паралельна. Рівняння (6) носить назву рівняння площини ***у відрізках на осях***.

Двогранний кут $\alpha$ між двома площинами, заданими рівняннями $A_1x+B_1y+C_1z+D_1=0$ $A_2x+B_2y+C_2z+D_2=0$, можна визначити як кут між їх нормалями $\mathbf{n}_1=(A_1,B_1,C_1)$ та $\mathbf{n}_2=(A_2,B_2,C_2)$:

$$\cos\alpha=\frac{\mathbf{n}_1\cdot\mathbf{n}_2}{|\mathbf{n}_1|\cdot|\mathbf{n}_2|}=\frac{A_1A_2+B_1B_2+C_1C_2}{\sqrt{A_1^2+B_1^2+C_1^2}\cdot\sqrt{A_2^2+B_2^2+C_2^2}} \tag{7}$$

Зауваження. Зрозуміло, що таким чином буде визначений лише один з двох двогранних кутів, інший рівний $\pi-\alpha$.

З формули (7) випливає ***умова перпендикулярності двох площин***, а саме: дві площини перпендикулярні тоді і тільки тоді, коли

$$A_1A_2+B_1B_2+C_1C_2=0. \tag{8}$$

***Умова паралельності двох площин*** еквівалентна умові колінеарності їх нормалей:

$$\frac{A_1}{A_2}=\frac{B_1}{B_2}=\frac{C_1}{C_2}. \tag{9}$$

Приклад 3. Визначити кути між площинами $x-3y+7z-5=0$ та $2x-4y-2z+1=0$.

Нормалями до заданих площин є вектори $\mathbf{n}_1=(1,-3,7)$ та $\mathbf{n}_2=(2,-4,-2)$. Легко переконатись, що скалярний добуток цих векторів рівний нулю, тобто виконана умова (8): $1\cdot 2+(-3)\cdot(-4)+7\cdot(-2)=0$ – площини перпендикулярні.

Приклад 4. Скласти рівняння площини, що проходить через точку $M_1(2,1,0)$ перпендикулярно до площин $x+y-2z-3=0$ та $2x-y-z+4=0$.

Оскільки шукана площина перпендикулярна до площин з нормалями $\mathbf{n}_1=(1,1,-2)$ та $\mathbf{n}_2=(2,-1,-1)$, то її нормаль $\mathbf{n}$ є перпендикулярною до обох нормалей $\mathbf{n}_1$ та $\mathbf{n}_2$, а отже, колінеарна їх векторному добутку $\mathbf{n}_1\times\mathbf{n}_2$. Визначимо $\mathbf{n}_1\times\mathbf{n}_2=\begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 1 & 1 & -2 \\ 2 & -1 & -1 \end{vmatrix}=\{-3;-3;-3\}$. Таким чином, можна вважати, що $\mathbf{n}=\{1;1;1\}$.

Шукана площина тепер визначається рівнянням (2): $(x-2)+(y-1)+z=0$, або $x+y+z-3=0$.

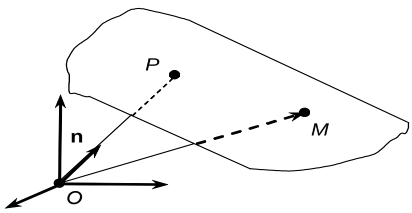
Приклад 5. Знайти об'єм тетраедра, утвореного координатними площинами та площиною $2x+3y+4z-12=0$.

Запишемо рівняння даної площини у вигляді рівняння (6), перенісши вільний член у праву частину та поділивши обидві частини рівняння на -12. Одержимо рівняння $\frac{x}{6}+\frac{y}{4}+\frac{z}{3}=1$. Отже, прямокутний тетраедр має ребра з довжинами $6$, $4$ та $3$. Його об'єм рівний 12 кубічним одиницям.

Приклад 6. Скласти рівняння площини, що проходить через точку $M_1(3,-1,5)$ та відтинає рівні відрізки від координатних осей.

З умови випливає, що шукана площина може бути описана рівнянням виду (4), в якому $a=b=c$. Залишилось підставити в це рівняння координати точки $M_1$.Одержимо рівняння: $\frac{3}{a}-\frac{1}{a}+\frac{5}{a}=1$, звідки $a=7$. Отже, шукана площина визначається рівнянням $x+y+z-7=0$.

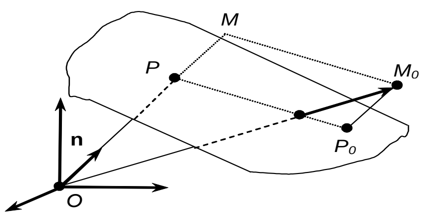
Розглянемо орт $\mathbf{n}=(\cos\alpha,\cos\beta,\cos\gamma)$ нормалі, проведеної із початку координат




до площини (див. рис.). Якщо позначити через $|\mathbf{OP}|=p\geq 0$ відстань від початку координат до площини, то для довільної точки $M(x,y,z)$ площини маємо рівність $np_{\mathbf{n}}\mathbf{OM}=p$, або

$$x\cos\alpha+y\cos\beta+z\cos\gamma-p=0 \qquad (10)$$

Дане рівняння називається ***нормальним рівнянням площини***, оскільки в ньому фігурує орт особливої нормалі площини – проведеної із початку координат. Щоб записати рівняння (3) у ***нормальній формі***, необхідно помножити обидві частини рівняння




на коефіцієнт $\mu=\frac{1}{\pm\sqrt{A^2+B^2+C^2}}$, підібравши знак так, щоб виконувалось $D\mu\leq 0$.

Нехай $M_0(x_0,y_0,z_0)$ – довільна точка простору, яка розташована з протилежного боку ніж початок координат відносно площини. Тоді для неї справджується очевидна рівність (див рис.): $np_{\mathbf{n}}\mathbf{OM_0}=|\mathbf{OP}|+|\mathbf{PM}|=p+|\mathbf{P_0M_0}|=p+d$, де $d=M_0P_0$ – відстань від точки $M_0(x_0,y_0,z_0)$ до площини. Отже, в цьому випадку маємо $d=x_0\cos\alpha+y_0\cos\beta+z_0\cos\gamma-p$. Якщо точка $M_0(x_0,y_0,z_0)$ розташована по той самий бік від площини, що й початок координат, то аналогічними міркуваннями (зробіть малюнок самостійно) доходимо висновку, що $np_{\mathbf{n}}\mathbf{OM_0}=p-d$, отже, $d=-(x_0\cos\alpha+y_0\cos\beta+z_0\cos\gamma-p)$. Введемо у розгляд величину

$$\delta=x_0\cos\alpha+y_0\cos\beta+z_0\cos\gamma-p. \qquad (11)$$

Вона зветься ***відхиленням*** точки $M_0(x_0,y_0,z_0)$ від площини і є результатом підстановки її координат у нормальне рівняння площини (10). Відхилення точки $M_0$ ***додатне***, якщо точка та початок координат розташовані по різні боки від площини і ***від'ємне***, якщо точка $M_0$ та початок координат розташовані з одного боку площини. У будь-якому випадку, $d = |\delta|$.

<u>Приклад 7</u>. Знайти відстань та відхилення точки $M_0(-1,2,5)$ від площини $x - 2y + 3z + 1 = 0$.

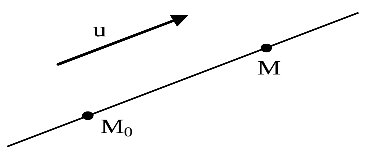
Довжина нормалі заданої площини дорівнює $\sqrt{1^2 + (-2)^2 + 3^2} = \sqrt{14}$. Щоб записати нормальне рівняння цієї площини, досить поділити обидві частини заданого рівняння на коефіцієнт $-\sqrt{14}$. Отже, нормальне рівняння матиме вигляд: $\dfrac{-x + 2y - 3z - 1}{\sqrt{14}} = 0$. Визначимо відхилення точки $M_0$ : $\delta = \dfrac{1 + 4 - 15 - 1}{\sqrt{14}} = \dfrac{-11}{\sqrt{14}}$ – точка та початок координат знаходяться з одного боку площини; відстань точки $M_0$ від площини рівна $d = \dfrac{11}{\sqrt{14}}$.

<u>Приклад 8</u>. Скласти рівняння площин, паралельних площині $4x - 2y - 4z - 5 = 0$ та розташованих на відстані $2$ одиниць від неї.

Щоб записати нормальне рівняння цієї площини, досить поділити обидві частини заданого рівняння на коефіцієнт $\pm\sqrt{4^2 + 2^2 + 4^2} = \pm 6$, причому знак треба вибрати так, щоб вільний член нормального рівняння був від'ємним. Отже, нормальне рівняння заданої площини матиме вигляд: $\dfrac{4x - 2y - 4z - 5}{6} = 0$. Тому рівняння паралельних площин, які розташовані на відстані $2$ одиниць від неї, матимуть вигляд: $\dfrac{4x - 2y - 4z - 5}{6} \pm 2 = 0$. Таким чином, шуканими є площини $4x - 2y - 4z + 7 = 0$ та $4x - 2y - 4z - 17 = 0$.

## Лекція 4. Прямі в просторі.

Якщо зафіксувати деякий вектор $\mathbf{u} = \{l, m, n\}$, то він задає напрямок у просторі. Тим самим визначена множина прямих паралельних даному напрямку. Для визначення конкретної прямої $L$ з цієї множини, досить вказати точку $M_0(x_0, y_0, z_0)$ на ній (див рис.). Щоб визначити цю пряму аналітично, тобто вказати рівняння, яке пов'язує координати довільної точки прямої $M(x, y, z)$, скористаємось колінеарністю векторів $\mathbf{M_0M}$ та $\mathbf{u}$ (вектор $\mathbf{u}$ називається ***напрямним вектором*** прямої) : $(\mathbf{r} - \mathbf{r})_\mathbf{0} \times \mathbf{u} = \mathbf{0}$ або $\mathbf{r} - \mathbf{r_0} = \lambda\mathbf{u}$, де $\lambda$ – довільне число. Таким чином одержане ***векторне рівняння прямої*** $L$.




Ця ж рівність для кожної координати вектора $\mathbf{M_0M}$ дає ***параметричні рівняння прямої*** $L$:

$$\begin{cases} x = x_\mathbf{0} + \lambda l \\ y = y_\mathbf{0} + \lambda m. \\ x = z_\mathbf{0} + \lambda n \end{cases} \qquad (1)$$

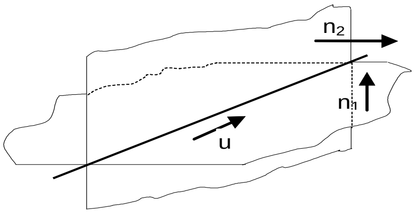
Тут $\lambda \in \mathbf{R}$ є параметром – кожна точка прямої визначена деяким його значенням. Якщо з рівностей (1) виключити параметр $\lambda$, то одержимо ***канонічні рівняння прямої***:

$$\frac{x - x_\mathbf{0}}{l} = \frac{y - y_\mathbf{0}}{m} = \frac{z - z_\mathbf{0}}{n} \qquad (2)$$

Розглянемо деякі частинні випадки. Припустимо, що одна з координат напрямного вектора прямої (2) рівна нулю, наприклад, $l = 0$. Тоді пряма, очевидно, перпендикулярна осі абсцис. Якщо ж $l = m = 0$, то пряма перпендикулярна до площини $XOY$ , тобто паралельна осі аплікат.

Пряму також можна задати, вказавши дві точки $M_1 = (x_1, y_1, z_1)$ та $M_2(x_2, y_2, z_2)$ на ній. Нехай $M(x, y, z)$ – довільна точка шуканої прямої. Тоді вектор $\mathbf{M_1M_2}$ буде напрямним і можемо записати канонічні ***рівняння прямої, заданої двома точками***: $\dfrac{x - x_1}{x_2 - x_1} = \dfrac{y - y_1}{y_2 - y_1} = \dfrac{z - z_1}{z_2 - z_1}$.

Пряма у просторі може також бути визначена як ***лінія перетину двох непаралельних площин:***




$$\begin{cases} A_1x + B_1y + C_1z + D_1 = 0 \\ A_2x + B_2y + C_2z + D_2 = 0 \end{cases} \qquad (3)$$

Ці рівняння описують площини, проте дають мало уявлення про власне пряму. Щоб записати пряму, задану рівнянням (3), у канонічному вигляді, необхідно визначити напрямний вектор прямої та деяку точку на ній. Напрямний вектор, очевидно, має бути паралельним кожній з площин, а отже, перпендикулярним до нормалей $\mathbf{n_1}$ та $\mathbf{n}_2$ обох площин, тому можна вважати, що $\mathbf{u} = \mathbf{n_1} \times \mathbf{n}_2 = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ A_1 & B_1 & C_1 \\ A_2 & B_2 & C_2 \end{vmatrix}$. Для того, щоб визначити точку на прямій (3), покладемо одну із змінних, наприклад $z$, рівною $z_0$ і розв'яжемо систему $\begin{cases} A_1x + B_1y = -D_1 - C_1z_0 \\ A_2x + B_2y = -D_2 - C_2z_0 \end{cases}$ відносно змінних $x$ та $y$.

<u>Приклад 1</u>. Скласти рівняння прямої, що проходить через точку $M_1(3{,}1{,}-2)$ паралельно прямій $x = 2 - \lambda;\ y = -3\lambda;\ z = -1 + 2\lambda$.

Напрямний вектор $\mathbf{u}$ заданої прямої визначаємо із її параметричних рівнянь: $\mathbf{u} = \{-1;-3;2\}$. Шукана пряма паралельна заданій, отже, вектор $\mathbf{u}$ буде напрямним і для неї. За формулами (2) визначаємо канонічні рівняння нашої прямої: $\dfrac{x-3}{-1} = \dfrac{y-1}{-3} = \dfrac{z+2}{2}$.

<u>Приклад 2</u>. Скласти канонічні рівняння прямої $x - 2y + 3z = 0$, $x + y - 6 = 0$.

Напрямний вектор $\mathbf{u}$ шуканої прямої визначимо як векторний добуток нормалей площин: $\mathbf{u} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 1 & -2 & 3 \\ 1 & 1 & 0 \end{vmatrix} = \{-3;3;3\}$, або $\mathbf{u} = \{-1;1;1\}$. Визначимо тепер точку, що належить шуканій прямій. Покладемо, наприклад, $z = 0$ та розв'яжемо систему $\begin{cases} x - 2y = 0 \\ x + y - 6 = 0 \end{cases}$. Помноживши перше рівняння на $-1$ та додавши до другого, знайдемо $y=2$. Отже, $x=4$, тому точка $(4,2,0)$ належить прямій. Запишемо канонічні рівняння цієї прямої: $\frac{x-4}{-1}=\frac{y-2}{1}=\frac{z}{1}$.

Зауваження. Очевидна неоднозначність канонічних рівнянь (2) – можна було використати іншу точку на прямій та взяти за напрямний будь-який вектор, колінеарний до $\mathbf{u}=\{-1;1;1\}$.

**Означення 1**. ***Пучком площин***, що проходять через вісь $L$, називається вся сукупність площин, що проходять через пряму $L$.

**Зауваження.** Пучок площин англійською мовою звучить як: *pencil of planes* або *sheaf of planes*.

**Теорема.** Нехай вісь пучка $L$ задана як лінія перетину двох непаралельних площин $\pi_1$: $A_1x+B_1y+C_1z+D_1=0$ та $\pi_2$: $A_2x+B_2y+C_2z+D_2=0$. Тоді при довільних $\alpha$ та $\beta$, таких, що $\alpha^2+\beta^2\neq 0$, пучок задається рівнянням:

$$\alpha\left(A_1x+B_1y+C_1z+D_1\right)+\beta\left(A_2x+B_2y+C_2z+D_2\right)=0 \qquad (4)$$

►Покажемо, що по-перше, рівняння (4) – не тотожність. Дійсно, з припущення про те, що $\alpha A_1+\beta A_2\equiv\alpha B_1+\beta B_2\equiv\alpha C_1+\beta C_2\equiv\alpha D_1+\beta D_2\equiv 0$ випливає $\frac{A_1}{A_2}=\frac{B_1}{B_2}=\frac{C_1}{C_2}=\frac{D_1}{D_2}=-\frac{\beta}{\alpha}$, тобто $\pi_1$ та $\pi_2$ збігаються. По-друге, вісь $L$ належить пучку. Це видно з того, що для кожної точки прямої $M_0(x_0,y_0,z_0)\in L$ виконуються рівності $A_1x_0+B_1y_0+C_1z_0+D_1=0$ та $A_2x_0+B_2y_0+C_2z_0+D_2=0$, а отже, виконано рівняння (4). По-третє, залишилось показати, що кожна площина $\pi$, що проходить через вісь $L$, описується рівнянням (4) при деяких значеннях параметрів $\alpha$ та $\beta$. Для цього розглянемо деяку точку $M_0(x_0,y_0,z_0)$, яка належить шуканій площині $\pi$, але не лежить на осі $L$. Точка $M_0$ та вісь $L$ повністю визначають площину $\pi$. Підставимо координати точки $M_0$ в рівняння осі пучка $L$. Величини $A_1x_0+B_1y_0+C_1z_0+D_1$ та $A_2x_0+B_2y_0+C_2z_0+D_2$ не дорівнюють нулю водночас (адже $M_0$ не належить осі $L$), а тому визначене відношення $\frac{A_1x_0+B_1y_0+C_1z_0+D_1}{A_2x_0+B_2y_0+C_2z_0+D_2}$, або $\frac{A_2x_0+B_2y_0+C_2z_0+D_2}{A_1x_0+B_1y_0+C_1z_0+D_1}$. Якщо існує перше, то позначимо його через $-\frac{\beta}{\alpha}$, якщо існує друге, то через $-\frac{\alpha}{\beta}$. У першому випадку маємо рівність $\frac{A_1x_0+B_1y_0+C_1z_0+D_1}{A_2x_0+B_2y_0+C_2z_0+D_2}=-\frac{\beta}{\alpha}$, у другому – $\frac{A_2x_0+B_2y_0+C_2z_0+D_2}{A_1x_0+B_1y_0+C_1z_0+D_1}=-\frac{\alpha}{\beta}$. Таким чином, для довільної площини $\pi$ пучка можна визначити коефіцієнти $\alpha$ та $\beta$, при яких площина описується рівнянням (4). ◄

**Зауваження.** Рівняння пучка у вигляді (4) описує і обидві площини $\pi_1$ та $\pi_2$. Якщо ж покласти у рівнянні (4) $\beta=\lambda\alpha$, то одержимо рівняння, яке описує всі площини пучка за винятком площини $\pi_2$.

Приклад 3. Записати рівняння площини, що проходить через пряму $\frac{x-4}{1}=\frac{y-4}{2}=\frac{z-1}{1}$ та точку $M(1,1,1)$. Шукана площина належить пучку з віссю

$\begin{cases} 2(x-4) = y-4 \\ y-4 = 2(z-1) \end{cases}$, або $\begin{cases} 2x - y - 4 = 0 \\ y - 2z - 2 = 0 \end{cases}$ і описується рівнянням $\lambda(2x - y - 4) + y - 2z - 2 = 0$ при деякому значенні $\lambda$. Щоб визначити $\lambda$, підставимо в це рівняння координати заданої точки $M$: $\lambda(-3) - 3 = 0$, отже $\lambda = -1$. Тому шуканою є площина $x - y + z - 1 = 0$.

### *Взаємне розташування прямої та площини.*

Розглянемо основні задачі про взаємне розташування прямої та площини.

1. Знайти точку перетину прямої $\frac{x - x_0}{l} = \frac{y - y_0}{m} = \frac{z - z_0}{n}$ з площиною $Ax + By + Cz + D = 0$.

Використаємо рівняння прямої у параметричній формі (1) та підставимо їх у рівняння площини. Таким чином з'ясуємо, при якому значенні параметру $\lambda$ має місце перетин прямої з площиною. Визначивши таким чином коефіцієнт $\lambda$, знайдемо координати точки перетину прямої та площини.

<u>Приклад 4</u>. Знайти точку перетину прямої $\frac{x-3}{-1} = \frac{y-1}{-2} = \frac{z+2}{2}$ з площиною $3x - y - 2z - 7 = 0$.

Підставимо параметричні рівняння прямої $x = 3 - \lambda;\ y = 1 - 2\lambda;\ z = -2 + 2\lambda$ у рівняння площини: $3(3-\lambda) - (1 - 2\lambda) - 2(-2 + 2\lambda) - 7 = 0$, звідки $\lambda = 1$. Отже, $x = 2;\ y = -1;\ z = 0$ – точка перетину даної прямої з площиною.

2. Знайти кут $\varphi$ між прямою $\frac{x - x_0}{l} = \frac{y - y_0}{m} = \frac{z - z_0}{n}$ та площиною $Ax + By + Cz + D = 0$. Сформулювати умови їх паралельності та ортогональності.

Неважко зрозуміти, що $\varphi = \pi/2 - \alpha$ або $\varphi = \pi/2 + \alpha$, де $\alpha$ – це кут між нормаллю площини та напрямним вектором прямої. Тому $\sin\varphi = \cos\alpha = \frac{\mathbf{n}\cdot\mathbf{u}}{|\mathbf{n}|\cdot|\mathbf{u}|} = \frac{Al + Bm + Cn}{\sqrt{A^2 + B^2 + C^2}\sqrt{l^2 + m^2 + n^2}}$. Пряма та площина паралельні, коли $Al + Bm + Cn = 0$, а перпендикулярні – при $\mathbf{n}\parallel\mathbf{u}$, тобто $\frac{A}{l} = \frac{B}{m} = \frac{C}{n}$.

<u>Приклад 5</u>. Знайти кут $\varphi$ між прямою $\frac{x-5}{6} = \frac{y-1}{-3} = \frac{z-2}{1}$ та площиною $7x + 2y - 3z + 5 = 0$.

Тут $\mathbf{u} = \{6;-3;1\}$ – напрямний вектор прямої, а $\mathbf{n} = \{7;2;-3\}$ – нормаль площини. Тоді $\sin\varphi = \frac{6\cdot 7 + (-3)\cdot 2 + 1\cdot(-3)}{\sqrt{7^2 + 2^2 + (-3)^2}\sqrt{6^2 + (-3)^2 + 1^2}} = \frac{33}{\sqrt{46}\sqrt{62}}$

3. Сформулювати умови належності прямої $\frac{x - x_0}{l} = \frac{y - y_0}{m} = \frac{z - z_0}{n}$ до площини $Ax + By + Cz + D = 0$.

Для того, щоб пряма лежала у площині необхідно і достатньо виконання двох умов: пряма паралельна площині і одна точка прямої належить площині. Запишемо ці умови аналітично: $\begin{cases} Al + Bm + Cn = 0 \\ Ax_0 + By_0 + Cz_0 + D = 0 \end{cases}$.

4. Сформулювати умови перетину двох непаралельних прямих $\dfrac{x-x_1}{l_1} = \dfrac{y-y_1}{m_1} = \dfrac{z-z_1}{n_1}$ та $\dfrac{x-x_2}{l_2} = \dfrac{y-y_2}{m_2} = \dfrac{z-z_2}{n_2}$.

Переконайтесь самостійно, що умова перетину двох непаралельних прямих еквівалентна умові компланарності векторів $\mathbf{u_1} = \{l_1, m_1, n_1\}$, $\mathbf{u_2} = \{l_2, m_2, n_2\}$ та $\mathbf{M_1 M_2}$, де $M_1 = (x_1, y_1, z_1)$ та $M_2(x_2, y_2, z_2)$.

5. Записати рівняння площини, що проходить через дві задані паралельні прямі.

6. Записати рівняння площини, що проходить через дві задані прямі, що перетинаються.

7. Знайти точку симетричну заданій відносно заданої площини.

Приклад 6. Знайти точку, симетричну точці $P(2,7,1)$ відносно площини $x - 4y + z + 7 = 0$.

Опустимо перпендикуляр із точки $P$ на площину та знайдемо точку $E$ перетину його з площиною. Точка $E$ буде серединою відрізку $PQ$, де $Q$ – шукана симетрична точка. Отже, перпендикуляр має проходити через точку $P$, а його напрямним вектором буде нормаль заданої площини. Таким чином одержимо рівняння цього перпендикуляра: $\dfrac{x-2}{1} = \dfrac{y-7}{-4} = \dfrac{z-1}{1}$. Точку $E$ знайдемо, як описано в прикладі 4: $E(3,3,2)$. Далі, оскільки точка $E$ – середина відрізку $PQ$, то її координати є напівсумою відповідних координат точок $P$ та $Q$, звідки й знаходимо $Q(4,-1,3)$.

8. Знайти відстань між двома заданими паралельними прямими.

9. Знайти відстань до заданої прямої від точки, що не лежить на даній прямій.

Приклад 7. Знайти відстань точки $P(1,3,5)$ від прямої $\dfrac{x-1}{1} = \dfrac{y+2}{-1} = \dfrac{z-1}{-1}$.

Побудуємо площину, що проходить через точку $P$ перпендикулярно до заданої прямої: $1\cdot(x-1) - 1\cdot(y-3) - 1\cdot(z-5) = 0$, або $x - y - z + 7 = 0$. Відшукаємо точку $E$ перетину цієї площини із заданою прямою: $E(-2,1,4)$. Шукана відстань дорівнює довжині відрізку $PE$: $d = |PE| = \sqrt{3^2 + 2^2 + 1^2} = \sqrt{14}$.

10. Знайти найкоротшу відстань $d$ між двома мимобіжними прямими.

Приклад 9. Знайти найкоротшу відстань між двома мимобіжними прямими $L_1: \dfrac{x-3}{1} = \dfrac{y-1}{-1} = \dfrac{z-2}{2}$ та $L_2: \dfrac{x}{-1} = \dfrac{y-2}{3} = \dfrac{z}{3}$.

Побудуємо площину $\sigma$, якій належить одна з прямих, приміром, $L_2$, паралельну до іншої – $L_1$. Тоді шукана відстань $d$ буде відстанню від довільної точки прямої $L_1$ (нам відома одна з її точок – точка $M_1(3,1,2)$) до площини $\sigma$. Позначимо напрямні вектори заданих прямих відповідно $\mathbf{u}_1 = \{1;-1;2\}$ та $\mathbf{u}_2 = \{-1;3;3\}$, відому точку на $L_2$ – $M_2(0,2,0)$. Нехай $M(x, y, z)$ буде довільна точка на шуканій площині $\sigma$. Для визначення рівняння площини використаємо компланарність векторів $\mathbf{u}_1$, $\mathbf{u}_2$ та $\mathbf{M}_2\mathbf{M}$ : $\begin{vmatrix} x & y-2 & z \\ 1 & -1 & 2 \\ -1 & 3 & 3 \end{vmatrix} = 0$, звідки одержимо рівняння площини $\sigma$ – $9x+5y-2z-10=0$. Нормальне рівняння цієї площини матиме вигляд $\frac{9x+5y-2z-10}{\sqrt{110}}=0$, тому $d=\frac{|9\cdot 3+5\cdot 1-2\cdot 2-10|}{\sqrt{110}}=\frac{18}{\sqrt{110}}$.

Другий спосіб: $\mathbf{M}_1\mathbf{M}_2=\{-3;1;-2\}$, вектор $\mathbf{u}=\mathbf{u}_1\times\mathbf{u}_2=\begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 1 & -1 & 2 \\ -1 & 3 & 3 \end{vmatrix}=\{-9;-5;2\}$ буде ортогональним до обох прямих, тоді $d=|np_{\mathbf{u}}\mathbf{M}_1\mathbf{M}_2|=\frac{|\mathbf{u}\cdot\mathbf{M}_1\mathbf{M}_2|}{|\mathbf{u}|}=\frac{18}{\sqrt{110}}$.

11. Записати рівняння спільного перпендикуляра двох мимобіжних прямих. Переконайтесь, що шукана пряма є лінією перетину двох площин: перша проходить через першу пряму паралельно вектору $\mathbf{u}=\mathbf{u}_1\times\mathbf{u}_2$ (тут $\mathbf{u}_1,\mathbf{u}_2$ – напрямні вектори заданих прямих), а друга площина – через другу пряму паралельно вектору $\mathbf{u}$.

Приклад 10. Записати рівняння спільного перпендикуляру $L$ до прямих $L_1:\ \frac{x-5}{1}=\frac{y-3}{-1}=\frac{z-13}{1}$ та $L_2:\ \frac{x-6}{1}=\frac{y-1}{2}=\frac{z-10}{-1}$; обчислити відстань між прямими та знайти точки перетину з ними спільного перпендикуляру.

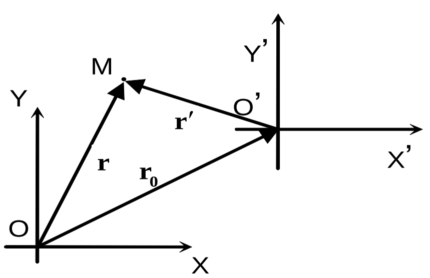
Знайдемо $\mathbf{u}=\mathbf{u}_1\times\mathbf{u}_2=\begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 1 & -1 & 1 \\ 1 & 2 & -1 \end{vmatrix}=\{-1;2;3\}$. Очевидно, що цей вектор буде напрямним для шуканої прямої $L$. А саму цю пряму можна знайти як лінію перетину двох площин, кожна з яких проходить через пряму $L$ та одну з прямих $L_1$ та $L_2$. Оскільки ми плануємо використовувати рівняння площин, що проходить через кожну з прямих $L_1$ та $L_2$, запишемо рівняння цих прямих як лінію перетину площин, щоб використати потім рівняння пучка площин. Для прямої $L_1$ маємо: $x+y-8=0$ та $y+z-16=0$. Для прямої $L_2$ відповідно $2x-y-11=0$ та $x+z-16=0$. Рівняння пучка площин з віссю $L_1$ буде таким: $x+y-8+\lambda(y+z-16)=0$, тоді для нормалі площини $\mathbf{n}_1=\{1,\lambda+1,\lambda\}$ з цього пучка запишемо умову ортогональності з вектором $\mathbf{u}$: $\mathbf{u}\cdot\mathbf{n}_1=-1+2\lambda+2+3\lambda=0$ (нагадаємо, що ця площина проходить через пряму $L_1$ паралельно до вектора $\mathbf{u}$). Отже, знайдемо $\lambda=-{}^1\!/_5$, тому шукана площина задається рівнянням $5x+4y-z-24=0$. Аналогічно, пучок з віссю $L_2$ описується рівнянням $2x-y-11+\lambda(x+z-16)=0$, $\mathbf{n}_2=\{2+\lambda,-1,\lambda\}$, $\mathbf{u}\cdot\mathbf{n}_2=-2-\lambda-2+3\lambda=0$, звідки $\lambda=2$, тому для другої площини маємо рівняння $4x-y+2z-43=0$. Таким чином, спільний перпендикуляр до прямих $L_1$ та $L_2$ задається рівняннями $5x+4y-z-24=0$, $4x-y+2z-43=0$. Щоб знайти точку перетину спільного перпендикуляру з кожною з прямих $L_1$, $L_2$, підставимо параметричні рівняння прямих у рівняння спільного перпендикуляру. Неважко визначити, що цими точками будуть точки $P_1=(5{,}3{,}13)$ та $P_2=(6{,}1{,}10)$. Знайдемо відстань між ними $d=\left|\mathbf{P_1P}_2\right|=\sqrt{14}$.

## Лекція 5. Основи теорії кривих другого порядку.

Знамениті криві – еліпс, гіпербола та парабола відомі математикам уже кілька тисячоліть. Їх відкриття приписують одному з учнів Платонівської Академії Менехму (IV ст. до н.е.). Він розглядав переріз прямого кругового конуса площинами і з'ясував, що в залежності від кута нахилу твірних конуса та розташування площини в перерізі з'являються криві з характерними геометричними властивостями. На його честь ці криві звалися тріадою Менехма. Сторіччям пізніше інший грецький математик Аполлоній присвятив цим кривим цілу монографію з восьми книг «Про конічні перерізи». Власне, саме Аполлоній дав назви «еліпс», «гіпербола» та «парабола» елементам тріади Менехма та відкрив багато залежностей, що й понині є предметом вивчення аналітичної геометрії. Сучасні математики, на відміну від своїх славнозвісних попередників, озброєні потужним координатним методом. При його використанні значного спрощення рівнянь можна досягти вдалим вибором системи координат.

### 1. Перетворення координат

Перехід до нової системи координат, очевидно, призводить до зміни рівнянь, що описують досліджувані геометричні об'єкти. При розгляді прямих та площин заміна однієї ПДСК на іншу принципово не змінює нічого – ми так само матимемо рівняння першого порядку з, можливо, іншими коефіцієнтами. У випадку більш складних геометричних об'єктів така заміна може призвести до суттєвого спрощення чи навпаки ускладнення рівняння. Приміром, всім відоме рівняння кола радіуса $R$ із центром у початку координат: $x^2+y^2=R^2$. Якщо перенести початок координат у точку $(x_0,y_0)$, то рівняння кола набуде вигляду $(x-x_0)^2+(y-y_0)^2=R^2$, коли ж у ньому розкрити дужки, то впізнати коло у одержаному рівняння буде не так просто. Отже, перетворення або заміна координат – це не самомета, а прагнення спростити для подальшого дослідження вигляд рівняння геометричного об'єкту. Обмежимось поки що розглядом перетворень систем координат на площині. Можна показати, що без зміни орієнтації ПДСК (коли найменший поворот від осі ОХ до осі ОУ відбувається у напрямку проти руху годинникової стрілки) всі можливі перетворення можна звести до послідовного виконання лише двох перетворень, а саме: паралельне перенесення системи координат (її осей) у деяку точку $(x_0,y_0)$ та поворот системи координат на деякий кут $\varphi$. Розглянемо їх.




#### 1.1. Паралельне перенесення системи координат у деяку точку $(x_0,y_0)$

Розглянемо ПДСК із початком у точці $O(0{,}0)$ і деяку точку $O'(x_0,y_0)$. Позначимо її радіус вектор через $\mathbf{r_0}$. Вважатимемо тепер точку $O'$ за початок нової ПДСК, вісі якої $O'X'$ та $O'Y'$ паралельні відповідним осям початкової системи координат (див. рис). Нехай $M$ – довільна точка площини. Її радіус-вектором відносно початкової системи координат є вектор $\mathbf{r}\{x.y\}$, а відносно нової системи координат – вектор $\mathbf{r}'\{x'.y'\}$. Очевидно, що $\mathbf{r}=\mathbf{r}'+\mathbf{r_0}$, що для координат точки $M$

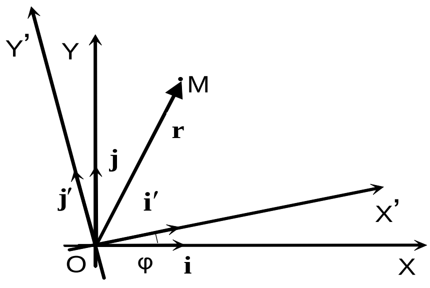
означає $\begin{cases} x = x' + x_0 \\ y = y' + y_0 \end{cases}$ або $\begin{cases} x' = x - x_0 \\ y' = y - y_0 \end{cases}$

(1)

Ця формула виражає зв'язок старих та нових координат довільної точки при паралельному перенесення системи координат.

**1.2. Поворот системи координат на деякий кут** $\varphi$

Розглянемо, як і раніше деяку ПДСК із початком у точці $O(0,0)$. Повернемо її осі на деякий кут $\varphi$, відраховуючи кут у напрямку проти руху годинникової стрілки (див. рис.). Позначимо орти старої та нової систем координат відповідно $\{\mathbf{i};\mathbf{j}\}$ та $\{\mathbf{i}';\mathbf{j}'\}$.




Тоді для будь-якої точки $M$ її радіус-вектор $\mathbf{r}\{x.y\}$ є незмінним в обох системах координат: $\mathbf{r} = x\mathbf{i} + y\mathbf{j} = x'\mathbf{i}' + y\mathbf{j}'$. З іншого боку, як неважко переконатись, для гострого кута $\varphi$, який зображено на малюнку,

$$\begin{cases} \mathbf{i}' = \cos\varphi\,\mathbf{i} + \sin\varphi\,\mathbf{j} \\ \mathbf{j}' = -\sin\varphi\,\mathbf{i} + \cos\varphi\,\mathbf{j} \end{cases}. \qquad (2)$$

Перевірте самостійно, що ця формула справедлива для довільного кута $\varphi$. З формул (2) неважко одержати формули

$$\begin{cases} \mathbf{i} = \cos\varphi\,\mathbf{i}' - \sin\varphi\,\mathbf{j}' \\ \mathbf{j} = \sin\varphi\,\mathbf{i}' + \cos\varphi\,\mathbf{j}' \end{cases} \qquad (3)$$

З виразу для радіус-вектора довільної точки $M$, одержимо: $x\mathbf{i} + y\mathbf{j} = x'(\cos\varphi\,\mathbf{i} + \sin\varphi\,\mathbf{j}) + y'(-\sin\varphi\,\mathbf{i} + \cos\varphi\,\mathbf{j})$. Оскільки розклад по базису єдиний, то одержимо формули переходу від старої системи координат при повороті системи на кут $\varphi$:

$$\begin{cases} x = x'\cos\varphi - y'\sin\varphi \\ y = y'\sin\varphi + y'\cos\varphi \end{cases} \qquad (4)$$

звідси неважко одержати, що

$$\begin{cases} x' = x\cos\varphi + y\sin\varphi \\ y' = -x\sin\varphi + y\cos\varphi \end{cases} \qquad (5)$$

**Зауваження.** Вираз від координат довільного вектора, який залишається незмінним при перетвореннях координат, називається ***інваріантом*** даного перетворення. Очевидно, що ***довжина вектора*** є інваріантом повороту системи координат так само, як і ***скалярний добуток*** векторів, оскільки він визначається довжинами векторів та кутом між ними. Які ще інваріанти перетворень координат можете назвати ви?

Приклад 1. Знайти координати точки $M(1,2)$ при повороті системи координат на кут $\varphi = \pi/4$ та паралельному перенесенні осей у точку $O'(-1,2)$.

Координати точки $M$ у новій системі координат знайдемо, послідовно застосувавши формули (5) та (1):

$x' = 1\cdot\cos\frac{\pi}{4} + 2\sin\frac{\pi}{4} + 1;\ y' = -1\cdot\sin\frac{\pi}{4} + 2\cos\frac{\pi}{4} - 2$. Отже, в новій системі координат положення точки $M$ визначається координатами: $M\left(\frac{3\sqrt{2}}{2} - 1, \frac{\sqrt{2}}{2} + 2\right)$.

## 2. Криві другого порядку.

### *Дослідження загального рівняння другого порядку двох змінних.*

Із шкільної геометрії добре відомо, що коло радіуса $R$ з центром в точці з координатами $(x_0, y_0)$ описується наступним рівнянням другого порядку відносно змінних $x, y$. Спробуємо знайти відповідь на питання, які ще криві можуть бути описані загальним рівнянням другого порядку:

$$a_{11}x^2 + 2a_{12}xy + a_{22}y^2 + 2a_{13}x + 2a_{23}y + a_{33} = 0 \qquad (6)$$

Очевидно, що рівняння (1) при певних значеннях коефіцієнтів $a_{ij}, (i, j = 1,2,3)$ може визначати коло. Тому спробуємо спростити загальне рівняння (6) так, щоб на передній план вийшли характерні особливості визначеної їм кривої.

Основою нашого аналітичного дослідження стане «вдалий» вибір ПДСК. Ступінь вдалості визначатиметься кількістю коефіцієнтів рівняння (1), які перетворяться на нуль в новій системі координат. Почнемо з повороту координатних осей на кут $\varphi$ проти руху годинникової стрілки. Як відомо, зв'язок старих координат $(x, y)$ точки з новими $(x', y')$ задається формулами (4). Підставимо їх в рівняння (6), щоб одержати рівняння того ж геометричного об'єкту в новій системі координат:

$$\tilde{a}_{11}x'^2 + 2\tilde{a}_{12}x'y' + \tilde{a}_{22}y'^2 + 2\tilde{a}_{13}x' + 2\tilde{a}_{23}y' + a_{33} = 0, \qquad (7)$$

де $\tilde{a}_{11} = a_{11}\cos^2\varphi + 2a_{12}\sin\varphi\cos\varphi + a_{22}\sin^2\varphi;$

$$\tilde{a}_{22} = a_{11}\sin^2\varphi - 2a_{12}\sin\varphi\cos\varphi + a_{22}\cos^2\varphi;$$

$$\tilde{a}_{12} = (a_{22} - a_{11})\sin\varphi\cos\varphi + a_{12}(\cos^2\varphi - \sin^2\varphi), \qquad \tilde{a}_{13} = a_{13}\cos\varphi + a_{23}\sin\varphi,$$

$$\tilde{a}_{23} = -a_{13}\sin\varphi + a_{23}\cos\varphi.$$

Зауважимо, що коефіцієнт $a_{33}$ залишився незмінним після перетворення (4) так само, як і сума коефіцієнтів $a_{11} + a_{22}$. Ці вирази є ***інваріантами*** повороту координатних осей. Придивимось уважніше до коефіцієнту $\tilde{a}_{12}$ – права частина виразу для $\tilde{a}_{12}$ є однорідним виразом другого порядку відносно тригонометричних функцій $\sin\varphi$ та $\cos\varphi$, а отже, рівняння

$$(a_{22} - a_{11})\sin\varphi\cos\varphi + a_{12}(\cos^2\varphi - \sin^2\varphi) = 0 \qquad (8)$$

завжди має розв'язки. Таким чином, має місце наступна ***теорема.***

**Теорема.** Завжди існує такий кут $\varphi$ повороту координатних осей ПДСК, що в новій системі координат коефіцієнт $\tilde{a}_{12}$ рівняння (7) дорівнює 0.

В цій системі координат рівняння (7) таким чином матиме вигляд

$$\tilde{a}_{11}x'^2 + \tilde{a}_{22}y'^2 + 2\tilde{a}_{13}x' + 2\tilde{a}_{23}y' + a_{33} = 0, \qquad (7^*)$$

На практиці, для відшукання кута $\varphi$ замість рівняння (8) зручніше користуватись еквівалентним йому квадратним рівнянням відносно $t = \operatorname{tg}\varphi$:

$$a_{12}t^2 + (a_{11} - a_{22})t - a_{12} = 0 \qquad (9)$$

Дискримінант цього рівняння додатний, а два його корені $t_1, t_2$ задають два взаємно перпендикулярних напрямки, тому за кут $\varphi$ обираємо арктангенс будь-якого з коренів $t_1, t_2$.

Наша подальша мета на шляху спрощення рівняння (7*) – виконати паралельне перенесення осей в деяку точку $(x_0, y_0)$: $\begin{cases} x' = x'' + x_0 \\ y' = y'' + y_0 \end{cases}$ (10)

Припустимо спочатку, що $\tilde{a}_{11} \cdot \tilde{a}_{22} \neq 0$. Тоді рівняння (7*) після заміни (10) перетвориться на таке:

$$\tilde{a}_{11}x''^2 + \tilde{a}_{22}y''^2 + 2(\tilde{a}_{13} + \tilde{a}_{11}x_0)x'' + 2(\tilde{a}_{23} + \tilde{a}_{22}y_0)y'' + \tilde{a}_{33} = 0,$$

де $\tilde{a}_{33} = a_{33} + 2\tilde{a}_{13}x_0 + 2\tilde{a}_{23}y_0 + \tilde{a}_{11}x_0{}^2 + \tilde{a}_{22}y_0{}^2$. (Коефіцієнти $\tilde{a}_{11}$ та $\tilde{a}_{22}$, очевидно, є інваріантами перетворення (10)). Поклавши $x_0 = -\dfrac{\tilde{a}_{13}}{\tilde{a}_{11}}$, $y_0 = -\dfrac{\tilde{a}_{23}}{\tilde{a}_{22}}$,

остаточно одержимо:
$$\tilde{a}_{11}x''^2 + \tilde{a}_{22}y''^2 + \tilde{a}_{33} = 0 \qquad (11)$$

Рівняння (11) задає криву другого порядку ***в канонічній формі***. Дослідимо можливі варіанти одержаної кривої, припустивши спочатку, що $\tilde{a}_{33} \neq 0$ (просто для зручності повернемось до початкових позначень координат $(x, y)$ ).

1. Нехай коефіцієнти $\tilde{a}_{11}, \tilde{a}_{22}$ та $\tilde{a}_{33}$ мають один і той самий знак. Тоді рівняння (11) не визначає жодної точки з дійсними координатами. Відповідний геометричний образ носить назву ***«уявного» еліпса***. Його канонічне рівняння має вид: $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = -1$. Тут $a^2 = \tilde{a}_{33}/\tilde{a}_{11}$, $b^2 = \tilde{a}_{33}/\tilde{a}_{22}$.

2. Нехай $\tilde{a}_{11}, \tilde{a}_{22}$ мають один і той самий знак, а $\tilde{a}_{33}$ – протилежний. Тоді рівняння (11) визначає дійсний ***еліпс***. Його канонічне рівняння має вид: $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = 1$.
Тут $a^2 = -\tilde{a}_{33}/\tilde{a}_{11}$, $b^2 = -\tilde{a}_{33}/\tilde{a}_{22}$.

3. Нехай $\tilde{a}_{11}$ та $\tilde{a}_{22}$ мають протилежні знаки, наприклад, $\tilde{a}_{11} \cdot \tilde{a}_{33} < 0$, а $\tilde{a}_{22} \cdot \tilde{a}_{33} > 0$. Тоді рівняння (11) визначає ***гіперболу***. Її канонічним рівнянням є наступне: $\dfrac{x^2}{a^2} - \dfrac{y^2}{b^2} = 1$. Тут $a^2 = -\tilde{a}_{33}/\tilde{a}_{11}$, $b^2 = \tilde{a}_{33}/\tilde{a}_{22}$. Якщо ж навпаки, $\tilde{a}_{11} \cdot \tilde{a}_{33} > 0$, а $\tilde{a}_{22} \cdot \tilde{a}_{33} < 0$, то одержимо ***спряжену гіперболу*** $\dfrac{x^2}{a^2} - \dfrac{y^2}{b^2} = -1$
з $a^2 = \tilde{a}_{33}/\tilde{a}_{11}$, $b^2 = -\tilde{a}_{33}/\tilde{a}_{22}$.

Якщо ж в рівнянні (11) $\tilde{a}_{33} = 0$, то можливий один з наступних варіантів:

4. Нехай $\tilde{a}_{11}$ та $\tilde{a}_{22}$ мають один і той самий знак. Тоді рівняння (11) є канонічним рівнянням ***пари комплексних прямих****, що перетинаються в дійсній точці (0;0)*:
$\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = 0$. Тут $a^2 = 1/\tilde{a}_{11}$, $b^2 = 1/\tilde{a}_{22}$.

5. Нехай $\tilde{a}_{11}$ та $\tilde{a}_{22}$ мають протилежні знаки. Тоді рівняння (8) є канонічним рівнянням ***пари дійсних прямих****, що перетинаються в початку координат (0;0)*:

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 0, \text{ або } \quad y = \pm\frac{b}{a}x.$$

Повернемось до рівняння (7*) та розглянемо випадок, коли один із коефіцієнтів $\tilde{a}_{11}$ або $\tilde{a}_{22}$ рівний 0. Припустимо, що $\tilde{a}_{11} = 0$, а $\tilde{a}_{22} \neq 0$. Вважатимемо поки, що

$\tilde{a}_{13} \neq 0$. Виконаємо паралельне перенесення осей (10) в точку $(x_0, y_0) = \left(0; -\dfrac{\tilde{a}_{23}}{\tilde{a}_{22}}\right)$. Одержимо: $\tilde{a}_{22} y''^2 + 2\tilde{a}_{13} x'' + \tilde{a}_{33} = 0$, (12)

де $\tilde{a}_{33} = a_{33} + 2\tilde{a}_{23} y_0$.

Якщо $\tilde{a}_{13} = 0$ , то матимемо один з наступних випадків:

6. Припустимо, що $\tilde{a}_{33} \neq 0$ і знак $\tilde{a}_{22}$ збігається із знаком $\tilde{a}_{33}$. Тоді рівняння (12) визначає ***пару уявних паралельних прямих:*** $y^2 = -b^2$, де $b^2 = \dfrac{\tilde{a}_{33}}{\tilde{a}_{22}}$.

7. Нехай $\tilde{a}_{33} \neq 0$, а знаки $\tilde{a}_{22}$ та $\tilde{a}_{33}$ протилежні. В цьому разі рівняння (12) визначає ***пару дійсних паралельних прямих:*** $y^2 = b^2$, де $b^2 = -\dfrac{\tilde{a}_{33}}{\tilde{a}_{22}}$, або ж $y = \pm b$

8. При $\tilde{a}_{13} = \tilde{a}_{33} = 0$ рівняння (12) визначає ***пару дійсних паралельних прямих, що співпали:***: $y^2 = 0$.

Якщо в рівнянні (12) $\tilde{a}_{13} \neq 0$, то виконаємо додаткове перетворення координат: $\begin{cases} x'' = x''' - \dfrac{\tilde{a}_{33}}{2\tilde{a}_{13}} \\ y'' = y''' \end{cases}$. Рівняння (12) перетвориться на наступне: $\tilde{a}_{22} y'''^2 = -2\tilde{a}_{13} x'''$.

(Можна вважати, що знаки $\tilde{a}_{22}$ та $\tilde{a}_{13}$ протилежні (в супротивному разі покласти $x'' = -x''' - \dfrac{\tilde{a}_{33}}{2\tilde{a}_{13}}$ ). Остаточно одержимо наступний висновок.

9. Рівняння (12) в цій ситуації перетворюється на канонічне рівняння ***параболи:*** $y^2 = 2px$, де $p = -\dfrac{\tilde{a}_{13}}{\tilde{a}_{22}} > 0$.

Отже, загальне рівняння (6) при всіх можливих значеннях коефіцієнтів слушним вибором нової ПДСК визначає одну і лише одну з 9 наведених вище кривих другого порядку. Основні з них – це еліпс, гіпербола та парабола. Крім того, можливі 3 випадки, коли криві вироджуються в пари прямих (таких, що перетинаються, паралельних або співпадаючих). Ще 3 випадки описують уявні криві. Одержані вище результати зведені у таблиці.

| **Номер** | **Рівняння кривої** | **Назва кривої** |
|---|---|---|
| 1 | $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = 1$ | еліпс |
| 2 | $\dfrac{x^2}{a^2} - \dfrac{y^2}{b^2} = 1$ | гіпербола |
| 3 | $y^2 = 2px$ | парабола |
| 4 | $\dfrac{x^2}{a^2} - \dfrac{y^2}{b^2} = 0$ | пара дійсних прямих, що перетинаються |
| 5 | $y^2 = b^2$ | пара дійсних паралельних прямих |
| 6 | $y^2 = 0$ | пара дійсних паралельних прямих, що співпали |

| | | |
|---|---|---|
| 7 | $\frac{x^2}{a^2}+\frac{y^2}{b^2}=-1$ | уявний еліпс |
| 8 | $\frac{x^2}{a^2}+\frac{y^2}{b^2}=0$ | пара уявних прямих, що перетинаються |
| 9 | $y^2=-b^2$ | пара уявних паралельних прямих |

**Зауваження.** Вище вже говорилось про те що сума коефіцієнтів $a_{11}+a_{22}$ є інваріантом повороту системи координат. Можна показати також, що інваріантом обох видів перетворень є і дискримінант $\delta$ старших членів рівняння (6): $\delta=\begin{vmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{vmatrix}=a_{11}a_{22}-a_{12}{}^2$ . Знак дискримінанту $\delta$ визначає тип кривої – при $\delta>0$ маємо криву еліптичного типу, при $\delta>0$ маємо криву гіперболічного типу, при $\delta=0$ – параболічного типу.

Приклад. Скласти канонічне рівняння кривої, заданої загальним рівнянням:

$$32x^2+52xy-7y^2+180=0$$

Складемо за коефіцієнтами цього рівняння тригонометричне рівняння (9) для визначення кута $\varphi$ повороту декартової системи координат: $2t^2+3t-2=0$, Звідки визначаємо, що $\operatorname{tg}\varphi=1/2$ або $\operatorname{tg}\varphi=-2$. Вибираємо додатне значення кута та знаходимо: $\sin\varphi=1/\sqrt{5}$, $\cos\varphi=2/\sqrt{5}$. Формули (4) для даного прикладу набувають вигляду: $x=\frac{2x'-y'}{\sqrt{5}}$, $y=\frac{x'+2y'}{\sqrt{5}}$. Отже початкове рівняння набуває вигляду: $32\left(\frac{2x'-y'}{\sqrt{5}}\right)^2+52\left(\frac{2x'-y'}{\sqrt{5}}\right)\left(\frac{x'+2y'}{\sqrt{5}}\right)-7\left(\frac{x'+2y'}{\sqrt{5}}\right)^2+180=0$, або $225x'^2-100y'^2=900$, звідки остаточно одержуємо рівняння гіперболи:

$$\frac{x'^2}{4}-\frac{y'^2}{9}=-1$$

Приклад. Скласти канонічне рівняння кривої, заданої загальним рівнянням:

$$25x^2-14xy+25y^2+64x-64y-224=0$$

Рівняння для визначення кута $\varphi$ в цьому випадку набуває вигляду: $7t^2-7=0$, звідки $t=\pm1$. Поклавши $\operatorname{tg}\varphi=1$, одержимо: $x=\frac{x'-y'}{\sqrt{2}}$, $y=\frac{x'+y'}{\sqrt{2}}$. Таким чином, початкове рівняння перетвориться на наступне:

$$25\left(\frac{x'-y'}{\sqrt{2}}\right)^2-14\left(\frac{x'-y'}{\sqrt{2}}\right)\left(\frac{x'+y'}{\sqrt{2}}\right)+25\left(\frac{x'+y'}{\sqrt{2}}\right)^2+64\left(\frac{x'-y'}{\sqrt{2}}\right)-64\left(\frac{x'+y'}{\sqrt{2}}\right)-224=0$$

або $18x'^2+32y'^2-64\sqrt{2}y'-224=0$, або ж $18x'^2+32\left(y'-\sqrt{2}\right)^2-288=0$. Очевидно, що дане рівняння додатково вимагає паралельного перенесення осей (формули (1)): $x'=x''$, $y'=y''+\sqrt{2}$. Таким чином одержимо $18x''^2+32y''^2=288$ і остаточно одержуємо: $\frac{x''^2}{16}+\frac{y''^2}{9}=1$ – канонічне рівняння еліпса.

## Лекція 6. Геометричні властивості основних кривих другого порядку.

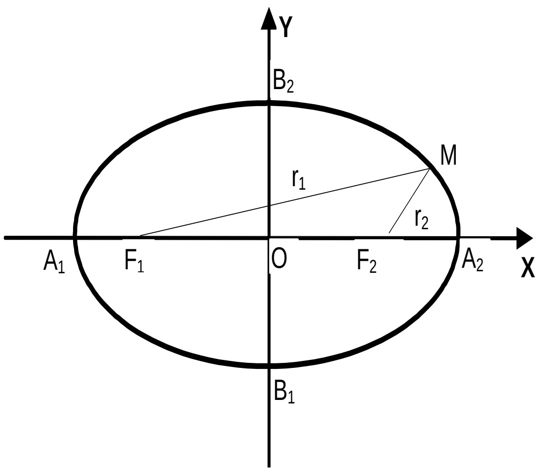
### 1. Еліпс

На цей момент нам відомо, що еліпс визначається канонічним рівнянням:

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=1 \qquad (1)$$

Деякі геометричні властивості еліпса випливають безпосередньо з нього. Наприклад, очевидно, що еліпс є симетричним відносно обох координатних осей, оскільки разом з точкою $(x,y)$ він обов'язково містить і точки $(-x,y)$, $(x,-y)$, $(-x,-y)$, тому всі властивості точок досить розглядати лише в першій чверті. Очевидно також, що всі точки еліпса знаходяться всередині прямокутника $|x|\le a,\ |y|\le b$ Точки $A_1(-a;0)$, $A_2(a;0)$, $B_1(0;-b)$ та $B_2(0;b)$ називаються вершинами еліпса. Оскільки $y=\pm b\sqrt{1-\frac{x^2}{a^2}}$ , то при зростанні $x$ від 0 до $a$ $y$ спадає від $b$ до 0.

Дослідження першої та другої похідних $y'$ та $y''$ дозволяють встановити гладкість еліпса та напрямки опуклості (це ми залишаємо для самостійної роботи читача). Таким чином, форму еліпса повністю з'ясовано – її зображено нижче на рис.




Розглянемо далі геометричне місце точок, що мають таку властивість: сума відстаней будь-якої точки від двох фіксованих точок, які називаються фокусами, є величиною сталою. Покажемо, що цим геометричним місцем точок є саме еліпс.

Виберемо прямокутну систему координат так, щоб вісь ОХ співпала з фокальною віссю (прямою, якій належать фокуси), а фокуси знаходилися б симетрично відносно початку координат. Розглянемо довільну точку $M$ з координатами $(x,y)$, яка має вказану вище геометричну властивість. Координати фокусів позначимо відповідно $F_1(-c;0)$ та $F_2(c;0)$. Позначимо $r_1=|F_1M|$, $r_2=|F_2M|$.

**<u>Означення 1.</u>** $r_1$ та $r_2$ називаються ***фокальними радіусами*** точки $M$.

Покладемо (у відповідності з властивістю ГМТ) $\quad r_1+r_2=2a \qquad (2)$

(З очевидних геометричних міркувань випливає, що $2a>2c$). Отже, маємо: $\sqrt{(x+c)^2+y^2}+\sqrt{(x-c)^2+y^2}=2a$, або $\sqrt{(x+c)^2+y^2}=2a-\sqrt{(x-c)^2+y^2}$. Оскільки обидві частини рівності додатні, піднесемо рівність до квадрату і виконаємо очевидні скорочення: $a\sqrt{(x-c)^2+y^2}=a^2-xc$. Повторне піднесення до квадрату приводить до рівності: $(a^2-c^2)x^2+a^2y^2=a^2(a^2-c^2)$.

Позначимо $b^2=a^2-c^2>0$ та поділимо обидві частини останньої рівності на $a^2b^2$. Остаточно одержимо $\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$ – це і є канонічне рівняння еліпса,

***центр*** якого знаходиться в початку координат. Залишаємо читачеві можливість самостійно переконатись, що проведені вище перетворення не привносять сторонніх коренів в рівняння (1). Параметри канонічного рівняння еліпса носять цілком природні назви: $A_1A_2 = 2a$ – ***велика вісь*** ($a$ – велика піввісь) $B_1B_2 = 2b$ – ***мала вісь*** ($b$ – мала піввісь).

<u>**Означення 2.**</u> ***Ексцентриситетом еліпса*** називають відношення відстані між фокусами до великої осі: $\varepsilon = \dfrac{c}{a}$. Для еліпса очевидно, що $\varepsilon \in [0,1)$.

При розв'язанні задач корисними є формули, що зв'язують параметри еліпса з його ексцентриситетом: $\varepsilon = \sqrt{1 - \dfrac{b^2}{a^2}}$, або $b = a\sqrt{1-\varepsilon^2}$.

Ексцентриситет є мірою «сплюснутості» еліпса – зокрема при $\varepsilon = 0$ одержимо $a = b$, тобто рівняння (1) задає коло. Цікавим виявляється і «шлях у зустрічному напрямку» – вивести з канонічного рівняння (1) геометричну властивість (2). Дійсно, для довільної точки еліпса $M(x, y)$ маємо:

$$r_1^2 = (x+c)^2 + y^2 = x^2 + 2cx + c^2 + b^2\left(1 - \frac{x^2}{a^2}\right) =$$

$$x^2\left(1 - \frac{b^2}{a^2}\right) + 2ax\varepsilon + (c^2 + b^2) = x^2\varepsilon^2 + 2ax\varepsilon + a^2 = (a + x\varepsilon)^2,$$

а отже, $$r_1 = a + x\varepsilon. \tag{3}$$

Цілком аналогічно можна одержати, що $r_2 = a - x\varepsilon$, (4)

звідки й випливає властивість (2). Формули (3) та (4) дають раціональні вирази для фокальних радіусів $r_1$ та $r_2$.

Рівняння еліпса інколи розглядається в параметричній формі: $x = a\cos t, y = b\sin t$, де $t \in [0, 2\pi]$.

**2. Гіпербола**

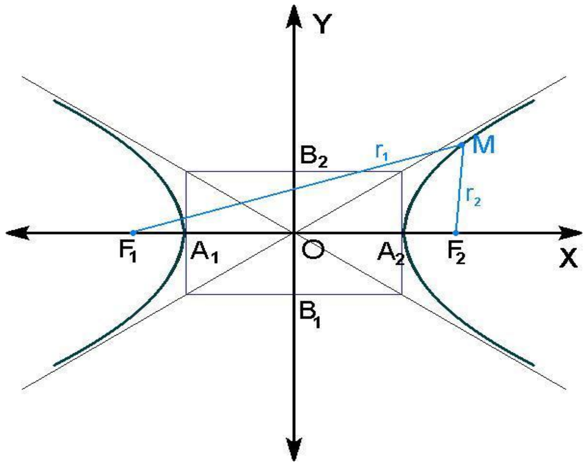
Канонічне рівняння гіперболи $$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1 \tag{5}$$

дає підстави для наступних висновків: по-перше, гіпербола, так само як і еліпс, є кривою симетричною відносно обох координатних осей, по-друге, гіпербола перетинає лише вісь ОХ в точках $A_1(-a;0)$ та $A_2(a;0)$ – вони називаються ***вершинами*** гіперболи, а відрізок $A_1A_2 = 2a$ – ***дійсною*** віссю. З віссю OY гіпербола спільних точок не має, тому відрізок $B_1B_2 = 2b$ називається ***уявною*** віссю. Для всіх точок гіперболи справедлива нерівність $|x| \geq a$ – отже, гіпербола складається із двох симетричних гілок. Оскільки $y = \pm b\sqrt{\dfrac{x^2}{a^2} - 1}$, то при зростанні $x$ від $a$ до $+\infty$ $y$ в першій чверті також зростає від 0 до $+\infty$. Неважко пересвідчитись, що $\lim\limits_{x \to +\infty}\left(\dfrac{b}{a}x - b\sqrt{\dfrac{x^2}{a^2} - 1}\right) = +0$, тобто гілка гіперболи нескінченно наближається знизу до прямої $y = \dfrac{b}{a}x$, але не перетинає її.

**Означення 3.** Прямі $y = \pm \frac{b}{a} x$ називаються ***асимптотами гіперболи***.

Дослідженням першої та другої похідних $y'$ та $y''$ читачі самостійно можуть переконатись у гладкості гіперболи та встановити напрямки опуклості її графіка. Графік гіперболи зображено на малюнку (рис.).

Покажемо, що точки гіперболи мають видатну геометричну властивість – модуль різниці відстаней будь-якої точки гіперболи від двох фіксованих точок, що називаються фокусами, є сталою величиною.




З цією метою розглянемо ПДСК, початок якої знаходиться посередині відрізку, кінцями якого є фокуси, а фокальна вісь збігається з віссю ОХ. Нехай точка $M$ має координати $(x, y)$. Координати фокусів позначимо відповідно $F_1(-c;0)$ та $F_2(c;0)$, а фокальні радіуси точки $M$ – $r_1 = |F_1M|$ та $r_2 = |F_2M|$.

Покладемо відповідно до вказаної властивості ГМТ: $|r_1 - r_2| = 2a$ (6)

Припустимо спочатку, що $x > 0$, тоді $r_1 > r_2$. Отже з рівності (6) маємо: $\sqrt{(x+c)^2 + y^2} - \sqrt{(x-c)^2 + y^2} = 2a$, або $\sqrt{(x+c)^2 + y^2} = 2a + \sqrt{(x-c)^2 + y^2}$. Оскільки обидві частини рівності додатні, піднесемо її до квадрату і виконаємо очевидні скорочення: $a\sqrt{(x-c)^2 + y^2} = xc - a^2$, звідки випливає $(c^2 - a^2)x^2 - a^2y^2 = a^2(c^2 - a^2)$.

Позначивши $b^2 = c^2 - a^2 > 0$ остаточно одержимо: $\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$ – канонічне рівняння гіперболи. Йому задовольняють і точки, для яких $x < 0$. Початок координат є ***центром гіперболи***, так само як і у випадку з еліпсом.

**Означення 4.** ***Ексцентриситетом гіперболи*** називається відношення відстані між фокусами до дійсної осі, таким чином для гіперболи (5) ексцентриситет рівний $\varepsilon = \frac{c}{a}$. Очевидно, що для гіперболи $\varepsilon \in (1, +\infty)$. Легко одержати також формули зв'язку параметрів гіперболи з її ексцентриситетом: $\varepsilon = \sqrt{1 + \frac{b^2}{a^2}}$, або $b = a\sqrt{\varepsilon^2 - 1}$.

Пропонуємо переконатись самостійно, що для точок правої гілки гіперболи справедливі формули раціональних виразів фокальних радіусів точки:

$r_1 = x\varepsilon + a$; $r_2 = x\varepsilon - a$, (7)

а для точок лівої гілки – $r_1 = -x\varepsilon - a$; $r_2 = -x\varepsilon + a$ (8)

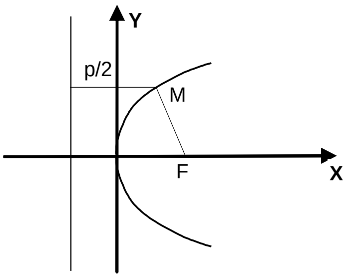
Гіпербола, яка задається рівнянням $\frac{x^2}{a^2} - \frac{y^2}{b^2} = -1$, має за дійсну вісь відрізок $2b$ на осі OY, а за уявну – відрізок $2a$ на осі OX. Вона називається ***спряженою*** до гіперболи (5).

### 3. Парабола

Повернемось до рівняння параболи: $y^2 = 2px$ (9)

Оскільки $p > 0$, то очевидно, що параболі належать пари точок симетрично розташованих відносно осі ОХ: $(x; y)$ та $(x; -y)$, для яких $x > 0$. Кажуть , що вісь ОХ є віссю параболи, а точка $O(0;0)$ – її вершиною.

Покажемо, що точки параболи мають наступну властивість: відстань будь-якої точки до деякої прямої, що зветься ***директрисою***, рівна відстані цієї точки до фіксованої точки – фокуса.




Введемо в розгляд ПДСК, вісь ОХ якої перпендикулярна директрисі, а початок координат є серединою відрізка проведеного від фокуса до директриси (див. рис.).
Нехай точка $M$ має координати $(x, y)$. Відстань від директриси до фокуса позначимо $p\ (p > 0)$. Тоді координати фокуса – точки $F\left(\frac{p}{2};0\right)$. Позначимо $r = |FM|$ і одержимо:

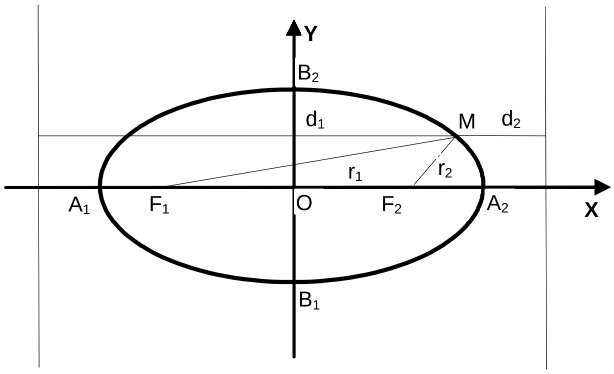
$$\frac{p}{2} + x = r\,, \qquad (10)$$

або $\frac{p}{2} + x = \sqrt{\left(\frac{p}{2} - x\right)^2 + y^2}$, звідки і випливає рівняння (9). Рівняння директриси параболи: $x = -\frac{p}{2}$.

**Фокальні властивості кривих другого порядку.**

Спорідненість всіх трьох кривих проявляється і в спільних так званих ***фокальних*** властивостях. Введемо спершу поняття директрис для еліпса та гіперболи.

<u>**Означення 5.**</u> ***Директрисами*** еліпса або гіперболи називаються прямі, перпендикулярні фокальній осі кривої, симетрично розташовані на відстані $\frac{a}{\varepsilon}$ від центру кривої (випадок кола, коли $\varepsilon = 0$, тут не розглядається).
Таким чином і еліпс, і гіпербола мають по дві директриси (як і по два фокуси), які задаються рівняннями $x = \pm\frac{a}{\varepsilon}$. Поняття директриси параболи і її рівняння були означені вище. Неважко помітити, що директриси всіх трьох кривих не перетинають самих кривих. Дійсно, для еліпса $\varepsilon < 1$, отже, $\frac{a}{\varepsilon} > a$, тобто директриси еліпса розташовані зовні його (див. рис.).




Для гіперболи ж $\varepsilon > 1$, отже, $\frac{a}{\varepsilon} < a$, тобто директриси розміщені між гілками гіперболи. Для точок параболи $x \geq 0$, директриса знаходиться в півплощині $x \leq 0$, а ексцентриситет параболи вважають рівним 1: $\varepsilon = 1$.

Основна фокальна властивість кривих другого порядку визначена наступною теоремою.

**Теорема** Відношення довжини фокального радіуса кожної точки довільної кривої другого порядку до відстані цієї точки до односторонньої з фокусом директриси є величиною сталою, рівною ексцентриситету кривої тобто $\frac{r}{d}=\varepsilon$ (11)

►Позначимо, як і раніше, фокальні радіуси довільної точки еліпса або гіперболи $r_1$ та $r_2$, а відстані цієї точки до відповідних директрис – $d_1$ та $d_2$. Отже, для лівого фокуса і лівої директриси еліпса маємо: $\frac{r_1}{d_1}=\frac{a+x\varepsilon}{\frac{a}{\varepsilon}+x}=\varepsilon$.

Цілком аналогічно для правого фокуса та правої директриси еліпса : $\frac{r_2}{d_2}=\frac{a-x\varepsilon}{\frac{a}{\varepsilon}-x}=\varepsilon$.

Для гіперболи необхідно окремо розглянути точки лівої та правої гілок. Отже, для точок лівої гілки маємо:

$$\frac{r_1}{d_1}=\frac{-x\varepsilon-a}{|x|-\frac{a}{\varepsilon}}=\frac{-x\varepsilon-a}{-x-\frac{a}{\varepsilon}}=\varepsilon; \qquad \frac{r_2}{d_2}=\frac{-x\varepsilon+a}{|x|+\frac{a}{\varepsilon}}=\frac{-x\varepsilon+a}{-x+\frac{a}{\varepsilon}}=\varepsilon.$$

Аналогічно, для точок правої гілки маємо: $\frac{r_1}{d_1}=\frac{x\varepsilon+a}{x+\frac{a}{\varepsilon}}=\varepsilon$; $\frac{r_2}{d_2}=\frac{x\varepsilon-a}{x-\frac{a}{\varepsilon}}=\varepsilon$.

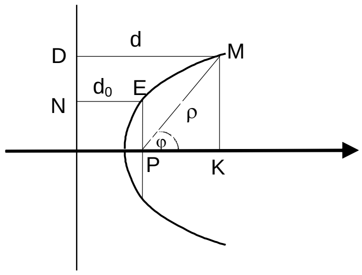
Для параболи безпосередньо з рівняння (10) випливає $\frac{r}{d}=1=\varepsilon$, що цілком виправдовує визначення ексцентриситету параболи. ◄

**Полярне рівняння кривих другого порядку.**

Нагадаємо, що дослідження загального рівняння кривої другого порядку визначило три основних канонічних рівняння, які описують не вироджені дійсні криві другого порядку – еліпс, гіперболу та параболу. З іншого боку, використання історично відомих геометричних властивостей саме цих кривих та вдало вибрана декартова система координат привели нас до тих самих канонічних рівнянь. Розглянемо ці криві тепер з іншого боку – визначимо їх рівняння в полярній системі координат. Справа в тому, що дуже багато кривих в давній геометрії вивчались саме в полярній системі координат тому, що і рівняння виглядали, як правило, просто та привабливо, і властивості кривих, описаних цими рівняннями, було видно, як на долоні. Згадаємо як приклад хоча б рівняння відомої спіралі Архімеда: $\rho=\varphi$ .

Отже, наша цікавість до полярних рівнянь кривих другого порядку має законне підґрунтя. Перейдемо до вибору полярної системи координат. Випереджаючи події, зазначимо, що всі три криві будуть описуватись одним і тим самим полярним рівнянням, що ще раз підкреслить їх спорідненість. Тому і розташування полюсу вибиратимемо таким чином, щоб форма всіх трьох кривих, принаймні в околі полюса, була схожою.

Отже, розглянемо параболу, еліпс та лише праву (поки що) гілку гіперболи, направивши полярну вісь вздовж фокальної осі кривих (в додатному напрямку декартової осі OX) і розташувавши полюс відповідно в фокусі параболи, лівому фокусі еліпса та правому фокусі гіперболи. Таким чином фрагменти всіх кривих матимуть вигляд, наведений на рис:




Тут $P$ – полюс кривої, точка $M$ – довільна точка однієї з названих кривих. Позначимо довжину її полярного радіусу $\rho$, а полярний кут – $\varphi$. Вертикальна пряма на малюнку – це директриса кривої (найближча до даного фокусу). Відрізок хорди кривої, що проходить через полюс перпендикулярно до полярної осі вважатимемо параметром полярного рівняння і позначимо через $2p$. Скористаємось основною фокальною властивістю кривих другого порядку для точок $E$ та $M$, позначивши їх відстані до директриси відповідно $d_0$ та $d$: $\frac{p}{d_0} = \varepsilon$, $\frac{\rho}{d} = \varepsilon$. Оскільки $d = d_0 + |PK| = d_0 + \rho\cos\varphi = \frac{p}{\varepsilon} + \rho\cos\varphi$

одержимо $\frac{\rho}{\varepsilon} = \frac{p}{\varepsilon} + \rho\cos\varphi$, або $\rho = \frac{p}{1-\varepsilon\cos\varphi}$ (12)

Це і є полярне рівняння кривої другого порядку, у випадку коли полюс обраний саме так, як зазначено вище. Якої саме кривої з трьох можливих – визначає величина ексцентриситету $\varepsilon$.

Пропонуємо самостійно переконатись, що полярним рівнянням еліпса з полюсом, розташованим у його правому фокусі буде наступне: $\rho = \frac{p}{1+\varepsilon\cos\varphi}$, а також вивести рівняння правої і лівої (окремо) гілок гіперболи з полюсом у кожному з фокусів.

Повернемось до розгляду рівняння (12), а саме: з'ясуємо зв'язок його параметрів $p$ та $\varepsilon$ з параметрами канонічних рівнянь кривих. Найпростіша ситуація у випадку параболи. З її означення безпосередньо випливає, що при $\varepsilon = 1$, $p$ – це параметр канонічного рівняння параболи. Якщо ж в рівнянні (12) $\varepsilon > 1$, то маємо гіперболу. Тоді декартові координати точки $E$ $(c; p)$. Підставимо їх в рівняння (5): $\frac{c^2}{a^2} - \frac{p^2}{b^2} = 1$, звідки $p = \frac{b}{a}\sqrt{c^2 - a^2} = \frac{b^2}{a}$ . Аналогічно при $\varepsilon < 1$ рівняння (12) описує еліпс. Підставимо точку $E(-c; p)$ в канонічне рівняння (1): $\frac{c^2}{a^2} + \frac{p^2}{b^2} = 1$. Легко одержати, що і в цьому випадку $p = \frac{b^2}{a}$.

<u>Приклад.</u> Записати канонічне рівняння кривої, заданої полярним рівнянням:

$$\rho = \frac{21}{5 - 2\cos\varphi} \tag{13}$$

Запишемо це рівняння у вигляді $\rho = \frac{21/5}{1 - 2/5\cos\varphi}$. Отже, $p = 21/5$, $\varepsilon = 2/5$, тобто полярне рівняння задає еліпс. Параметри його канонічного рівняння $a$ та $b$ можна одержати з виразів для $p$ та $\varepsilon$, ми ж поступимо іншим чином. Підставимо в рівняння (13) два значення полярного кута 0 та $\pi$: $\rho(0) = 7$, $\rho(\pi) = 3$. Тоді неважко побачити, що тоді $a+c=7$, $a-c=3$, тобто $a=5$, $c=2$. Отже, канонічне рівняння даного еліпса: $\frac{x^2}{25}+\frac{y^2}{21}=1$.

Приклад. Дано рівняння гіперболи $\frac{x^2}{16}-\frac{y^2}{9}=1$. Скласти рівняння її правої гілки, вважаючи, що напрямок її полярної осі збігається з додатним напрямком осі OX, а полюс розташований у правому фокусі.

З рівняння гіперболи випливає, що $c=\sqrt{a^2+b^2}=\sqrt{16+9}=5$. Отже, $\varepsilon=5/4$, $p=b^2/a=9/4$. Таким чином, полярне рівняння даної гіперболи є таким: $\rho=\frac{9/4}{1-5/4\cos\varphi}$ або $\rho=\frac{9}{4-5\cos\varphi}$ .

**Дотичні до кривих другого порядку.**

Гладкість кривих другого порядку забезпечує існування дотичних до них в усіх точках. Будемо спиратись на відомий з курсу математичного аналізу факт: якщо деяка точка $M_0(x_0;y_0)$ належить гладкій кривій, що задається рівнянням $y=y(x)$, то дотична до цієї кривої в даній точці задається рівнянням:

$$y-y_0=y'(x_0)(x-x_0) \qquad (14)$$

(За винятком дотичних, паралельних осі OY).

**1.** Нехай точка $M_0(x_0;y_0)$ належить еліпсу. Диференціюючи рівняння еліпса (1), маємо $\frac{2x}{a^2}+\frac{2y\cdot y'}{b^2}=0$, звідки $y'=-\frac{b^2}{a^2}\cdot\frac{x}{y}$ (для тих точок еліпса, для яких $y\neq 0$). Отже, дотична до еліпса в точці $M_0$ задаватиметься рівнянням: $y-y_0=-\frac{b^2}{a^2}\cdot\frac{x_0}{y_0}(x-x_0)$, або $\frac{x_0(x-x_0)}{a^2}+\frac{y_0(y-y_0)}{b^2}=0$. Враховуючи, що координати точки $M_0$ задовольняють рівняння еліпса (10), остаточно отримаємо ***рівняння дотичної до еліпса в точці*** $M_0(x_0;y_0)$:

$$\frac{xx_0}{a^2}+\frac{yy_0}{b^2}=1 \qquad (15)$$

**2.** Розглянемо точку $M_0(x_0;y_0)$, що належить одній з гілок гіперболи. З рівняння гіперболи (5) одержимо вираз для похідної гіперболи : $y'=\frac{b^2}{a^2}\cdot\frac{x}{y}$. Тоді рівняння дотичної (14) в цьому випадку матиме вигляд: $y-y_0=\frac{b^2}{a^2}\cdot\frac{x_0}{y_0}(x-x_0)$, або $\frac{x_0(x-x_0)}{a^2}-\frac{y_0(y-y_0)}{b^2}=0$. Звідси приходимо до ***рівняння дотичної до гіперболи в точці*** $M_0(x_0;y_0)$:

$$\frac{xx_0}{a^2}-\frac{yy_0}{b^2}=1 \qquad (16)$$

**3.** І нарешті, у випадку параболи (9) рівняння її похідної буде таким: $y'=\frac{p}{y}$ (для тих точок параболи, для яких $y\neq 0$). Рівняння (14) дотичної до параболи в точці $M_0(x_0;y_0)$ набуває вигляду: $y-y_0=\frac{p}{y_0}(x-x_0)$, або $y_0(y-y_0)=p(x-x_0)$, звідки, беручи до уваги, що $y_0{}^2=2px_0$, отримаємо ***рівняння дотичної до параболи:***

$$yy_0=p(x+x_0) \qquad (17)$$

## Лекція 7. Поверхні другого порядку.

Загальна теорія поверхонь другого порядку значно складніша порівняно з теорією кривих. Ми не будемо викладати її тут детально. Познайомимось лише з основними видами поверхонь та їх властивостями.

### 1. Загальне рівняння другого порядку трьох змінних.

Загальне рівняння другого порядку, що описує деяку поверхню в тривимірному просторі, залежить від трьох змінних і має вид:

$$a_{11}x^2 + 2a_{12}xy + 2a_{13}xz + a_{22}y^2 + 2a_{23}yz + a_{33}z^2 + 2a_{14}x + 2a_{24}y + 2a_{34}z + a_{44} = 0 \quad (1)$$

Перетвореннями системи координат можна звести це рівняння до одного з канонічних рівнянь поверхонь другого порядку.

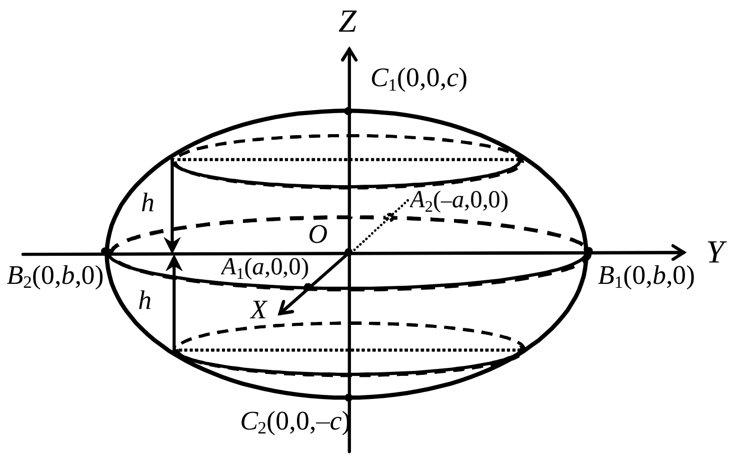
Розглянемо основні з них.

#### 1.1. Еліпсоїд.

Канонічним рівнянням еліпсоїду є рівняння виду:

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1 \qquad (a > 0, b > 0, c > 0) \qquad (2)$$

З рівняння (2) безпосередньо випливає, що по-перше, дана поверхня симетрична відносно всіх трьох координатних площин, і по-друге, $|x| \le a,\ |y| \le b,\ |z| \le c$.




Точки $A_1(a,0,0)$, $A_1(-a,0,0)$, $B_1(0,b,0)$, $B_2(0,-b,0)$, $C_1(0,0,c)$ та $C_2(0,0,-c)$ називаються вершинами еліпсоїда. Покажемо, що якщо існує переріз еліпсоїда поверхнею паралельною до однієї з координатних площин, то таким перерізом завжди є еліпс (власне, цій обставині він і завдячує своєю назвою). Дійсно, розглянемо переріз площиною $z = \pm h$, де $0 \le h \le c$. Маємо:

$$\begin{cases} \dfrac{x^2}{a^2} + \dfrac{t^2}{b^2} = 1 - \dfrac{h^2}{c^2} \\ z = \pm h \end{cases}.$$

Ці рівняння визначають еліпс з півосями $a\sqrt{1 - \frac{h^2}{c^2}}$ та $b\sqrt{1 - \frac{h^2}{c^2}}$ (див. рис.). .

Максимальні півосі, рівні $a$ і $b$, матиме еліпс, одержаний в перерізі координатною площиною $XOY$ при $h = 0$. При $h = c$ перерізи перетворюється на точки – *вершини еліпсоїда*. Аналогічні міркування справедливі і для перерізів еліпсоїда площинами $y = \pm h$ та $z = \pm h$ .

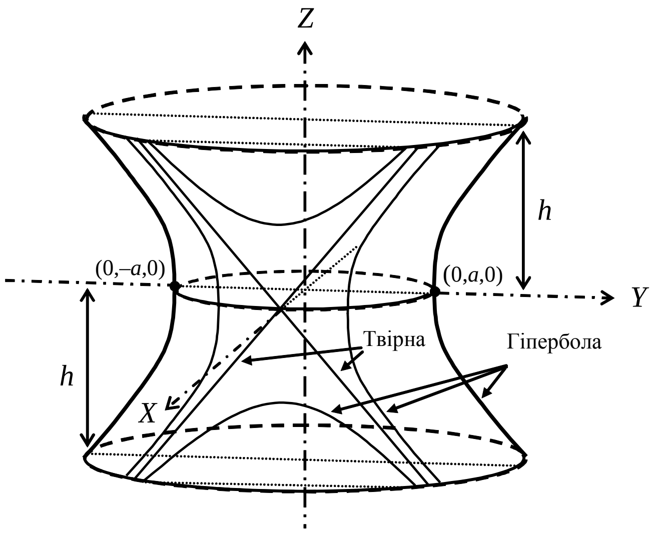
#### 1.2. Однопорожнинний гіперболоїд.

Однопорожнинний гіперболоїд задається канонічним рівнянням виду:

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1 \qquad (a > 0, b > 0, c > 0) \qquad (3)$$

Очевидна симетрія точок поверхні відносно всіх трьох координатних площин. Дослідимо форму поверхні з допомогою перерізів площинами паралельними координатним. Розглянемо спочатку переріз поверхні площиною $z=\pm h$. Маємо:

$$\begin{cases} \dfrac{x^2}{a^2}+\dfrac{y^2}{b^2}=1+\dfrac{h^2}{c^2} \\ z=\pm h \end{cases}$$




Ці рівняння при будь-якому значенні $h$ визначають еліпс з півосями $a\sqrt{1+\dfrac{h^2}{c^2}}$ та $b\sqrt{1+\dfrac{h^2}{c^2}}$. Мінімальні півосі, рівні $a$ і $b$, матиме еліпс, одержаний в перерізі координатною площиною $XOY$ при $h=0$.

В перерізі однопорожнинного гіперболоїда будь-якою з площин $x=\pm h$ матимемо гіперболу:

$$\begin{cases} \dfrac{y^2}{b^2}-\dfrac{z^2}{c^2}=1-\dfrac{h^2}{a^2} \\ x=\pm h \end{cases}$$

Її дійсною віссю при $0<h<a$ буде $b\sqrt{1-\dfrac{h^2}{a^2}}$, а уявною – $c\sqrt{1-\dfrac{h^2}{a^2}}$. Якщо ж $h>a$, то дійсна вісь буде рівна $c\sqrt{\dfrac{h^2}{a^2}-1}$, а уявна – рівна $b\sqrt{\dfrac{h^2}{a^2}-1}$. Зауважимо, що при $h=a$ перерізи являють собою дві прямі, що перетинаються в точці $(\pm a,0,0)$. Аналогічно можна переконатись, що і в перерізі будь-якою з площин $y=\pm h$ також отримаємо гіперболу. Рівняння (3) приховує одну дуже важливу особливість однопорожнинного гіперболоїда, а саме; зазначена вище пара прямих є лише однією з двох сімейств таких пар прямих, цілком розташованих на цій поверхні. Це так звані ***прямолінійні твірні*** **однопорожнинного** гіперболоїда. Визначимо їх. Рівняння (4) можна подати у вигляді:

$$\frac{x^2}{a^2}-\frac{z^2}{c^2}=1-\frac{y^2}{b^2}, \quad \text{або} \quad \left(\frac{x}{a}+\frac{z}{c}\right)\left(\frac{x}{a}-\frac{z}{c}\right)=\left(1+\frac{y}{b}\right)\left(1-\frac{y}{b}\right) \qquad (4)$$

Розглянемо тепер пряму лінію, утворену перетином двох площин, визначених при деяких значеннях числових параметрів $\lambda$ та $\mu$:

$$\begin{cases} \lambda\left(\dfrac{x}{a}+\dfrac{z}{c}\right)=\mu\left(1+\dfrac{y}{b}\right) \\ \mu\left(\dfrac{x}{a}-\dfrac{z}{c}\right)=\lambda\left(1-\dfrac{y}{b}\right) \end{cases} \qquad (5)$$

Координати кожної точки прямої (5), очевидно задовольняють рівняння (4), а отже, і рівняння (3). Тому всі точки прямих (5) лежать на гіперболоїді (3). Аналогічними міркуванням переконайтесь, що і сімейство прямих

$$\begin{cases} \alpha\left(\dfrac{x}{a}+\dfrac{z}{c}\right)=\beta\left(1-\dfrac{y}{b}\right) \\ \beta\left(\dfrac{x}{a}-\dfrac{z}{c}\right)=\alpha\left(1+\dfrac{y}{b}\right) \end{cases} \qquad (6)$$

також належить однопорожнинному гіперболоїду (3) при всіх значеннях параметрів $\alpha$ та $\beta$. Якщо розглянути довільну точку на гіперболоїді (3) і підставити її координати в одне з рівнянь сімейства (5) або (6), то визначимо значення параметрів $\lambda, \mu$ та $\alpha, \beta$ , які відповідають парі прямих, що проходять саме через цю точку однопорожнинного гіперболоїда. Зрозуміло, що значення цих пар параметрів можуть бути знайдені лише з точністю до спільного множника. Однопорожнинний гіперболоїд зображена на рис. вище**.**

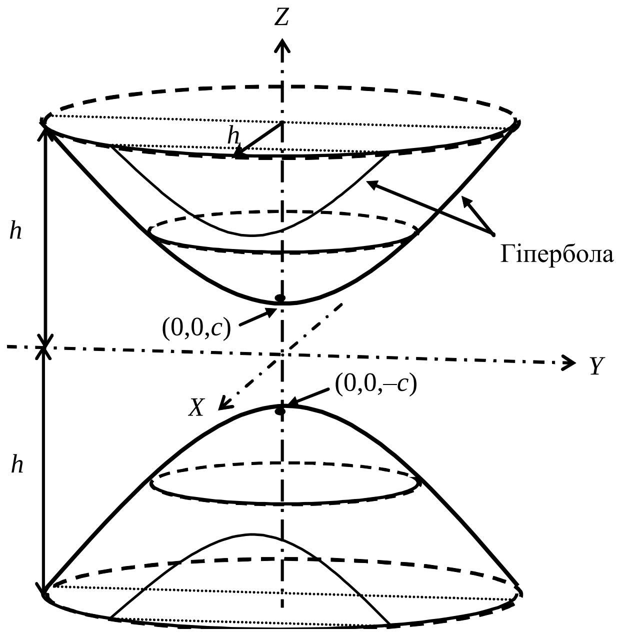
### 1.3. Двопорожнинний гіперболоїд.

Двопорожнинний гіперболоїд задається канонічним рівнянням виду:

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}-\frac{z^2}{c^2}=-1 \qquad (a>0, b>0, c>0) \qquad (7)$$

Ця поверхня, очевидно, також симетрична відносно всіх трьох координатних площин. Форму поверхні знову встановимо з допомогою перерізів. Розглянемо спочатку переріз площиною $z=\pm h$. Маємо:

$$\begin{cases} \dfrac{x^2}{a^2}+\dfrac{y^2}{b^2}=\dfrac{h^2}{c^2}-1 \\ z=\pm h \end{cases}$$




Бачимо, що при $h>c$ дані рівняння визначають еліпс з півосями $a\sqrt{\dfrac{h^2}{c^2}-1}$ та $b\sqrt{\dfrac{h^2}{c^2}-1}$. Переріз взагалі не існує при $0<h<c$, а при $h=c$ стає точкою $(0,0,\pm c)$. Ці точки є вершинами двопорожнинного гіперболоїда – ця поверхня складається з двох окремих частин (рис.).

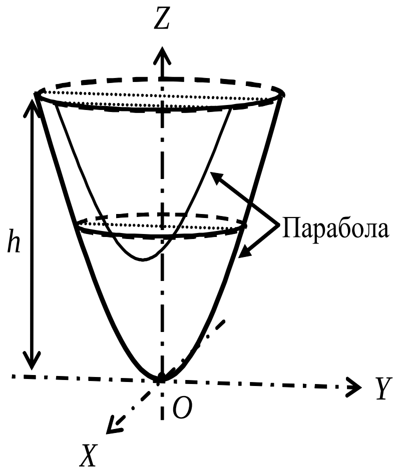
Дослідимо тепер переріз цієї поверхні площиною $x=\pm h$:

$$\begin{cases} \dfrac{y^2}{b^2}-\dfrac{z^2}{c^2}=-1-\dfrac{h^2}{a^2} \\ x=\pm h \end{cases}$$

В перерізі при будь-якому значенні $h$ маємо гіперболу, дійсна вісь якої рівна $c\sqrt{1+\dfrac{h^2}{a^2}}$ і паралельна осі $OZ$, а уявна, рівна $b\sqrt{1+\dfrac{h^2}{a^2}}$ , паралельна осі $OY$.

Аналогічно переконайтесь, що і в перерізі площиною $y=\pm h$ при довільному значенні $h$ матимемо гіперболу з дійсною віссю, паралельною осі $OZ$ і уявною – паралельною осі $OX$.

## 1.4. Еліптичний параболоїд.




Канонічним рівнянням еліптичного параболоїду є наступне: $\frac{x^2}{p}+\frac{y^2}{q}=2z \qquad (p>0, q>0)$ (8)

Ця поверхня є симетричною відносно координатних площин $XOZ$ та $YOZ$ і розташована над площиною $XOY$. В перерізі площиною $z=h$ при $h\geq 0$ одержимо еліпс (при $h=0$ перерізу належить лише вершина еліптичного параболоїда – точка $O(0,0,0)$):

$$\begin{cases}\frac{x^2}{p}+\frac{y^2}{q}=2h\\ z=h\end{cases}$$

Півосі цього еліпса рівні $\sqrt{2ph}$ та $\sqrt{2qh}$.

Якщо ж розглянути переріз еліптичного параболоїду наприклад площиною $x=\pm h$, то одержимо параболу: $\begin{cases}\frac{y^2}{q}=2z-\frac{h^2}{p}\\ x=\pm h\end{cases}$, вершиною якої є точка з координатами $(\pm h, 0, \frac{h^2}{2p})$, а вісь паралельна осі $OZ$. Так само і в перерізі площиною $y=\pm h$ для довільного значення параметра $h>0$ матимемо параболу. Цей факт і пояснює назву цієї поверхні. Її вигляд показаний на рис. вище.

**<u>Зауваження.</u>** Слід зазначити, що будь-яка з розглянутих вище поверхонь може бути одержана при обертанні відповідної кривої другого порядку навколо деякої осі.

Розглянемо деяку криву другого порядку, розташовану в площині $XOZ$. Її рівняння в загальному випадку має вигляд: $\begin{cases}F(x,z)=0\\ y=0\end{cases}$. Нехай тепер дана крива (вона називається меридіан) обертається навколо осі $OZ$. Щоб записати рівняння утвореної обертанням меридіану поверхні, виберемо на ній довільну точку $M$ з координатами $M(x,y,z)$. Позначимо відповідну точку на меридіані через $M_0(x_0,0,z_0)$. Для неї виконана рівність $F(x_0,z_0)=0$. (9) Оскільки точки $M$ та $M_0$ лежать на одному колі (паралель обертання), то неважко зрозуміти, що $x_0{}^2=x^2+y^2$. Крім того, очевидно, що при обертанні навколо осі $OZ$ має місце рівність $z_0=z$. Підставивши одержані співвідношення у (9), одержимо рівність $F(\sqrt{x^2+y^2},z)=0$ – це і є рівняння поверхні, утвореної обертанням заданого меридіану навколо осі $OZ$.

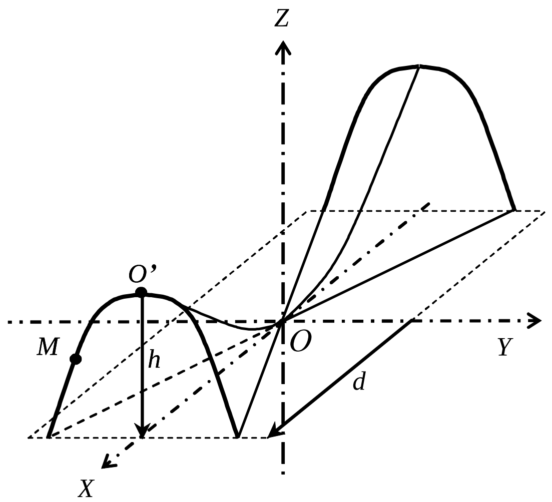
Наприклад, еліпс $\frac{x^2}{a^2}+\frac{z^2}{c^2}=1$, розташований в площині $XOZ$, при обертанні навколо осі $OZ$ опише еліпсоїд обертання: $\frac{x^2+y^2}{a^2}+\frac{z^2}{c^2}=1$.

Перетворенням координат $x'=x, y'=\lambda y, z'=z$, при якому вздовж осі $OY$ відбувається «стискання», якщо $0<\lambda<1$, або «розтягування», якщо $\lambda>1$, можна одержати еліпсоїд (2), де $b=\lambda a$.

Одно- та двопорожнинний гіперболоїди обертання можна одержати обертанням гіперболи відповідно навколо уявної або дійсної осей, а параболоїд обертання – обертанням параболи навколо її осі симетрії. Описане вище перетворення координат, застосоване до вказаних поверхонь обертання, приведе до рівнянь (3), (7), (8).

Поверхня, до розгляду якої ми переходимо далі, в цьому розумінні є унікальною, тому що не може бути одержана обертанням жодної кривої другого порядку.

### 1.5. Гіперболічний параболоїд.




Гіперболічний параболоїд описується канонічним рівнянням:

$$\frac{x^2}{p} - \frac{y^2}{q} = 2z \qquad (p > 0, q > 0) \qquad (10)$$

**<u>Теорема.</u>** Гіперболічний параболоїд (10) є поверхнею, утвореною рухом параболи $y^2 = -2qz,\ x = 0$, вершина якої ковзає вздовж нерухомої параболи $x^2 = 2pz,\ y = 0$ так, що площини обох парабол залишаються взаємно перпендикулярними, а осі – протилежно напрямленими (див. рис.).

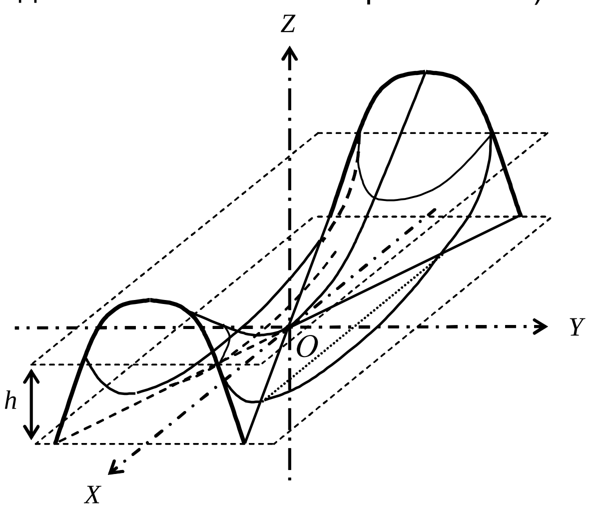
► Розглянемо довільну точку $M(x, y, z)$, яка належить поверхні, одержаній вказаним рухом параболи. Нехай точка $M$ належить рухомій параболі в той момент, коли її вершина знаходиться в точці $O'$ з координатами $(d, 0, h)$, де $d$ – відстань від площини рухомої параболи до координатної площини $YOZ$. Оскільки точка $O'$ належить водночас і нерухомій параболі, то $d^2 = 2ph$, а координати точки $M$ задовольняють рівняння:

$$\begin{cases} y^2 = 2q(h - z) \\ x = d \end{cases}$$

Оскільки $h = \dfrac{d^2}{2p}$, а $d^2 = x^2$, то одержуємо рівняння $y^2 = 2q\left(\dfrac{x^2}{2p} - z\right)$, звідки й випливає рівняння (10). ◄

**<u>Зауваження.</u>** Якщо осі обох парабол, що рухаються описаним вище способом, напрямлені в один бік, то одержимо еліптичний параболоїд (8).

Дослідимо форму параболічного гіперболоїда (поки що законним в його назві здається лише слово параболічний) за допомогою перерізів.




Розглянемо спочатку переріз поверхні площиною, паралельною $XOY$:

$$\begin{cases} \dfrac{x^2}{p} - \dfrac{y^2}{q} = 2h \\ z = h \end{cases}$$

Маємо в перерізі гіперболу (виняток складає випадок, коли $h = 0$, тоді переріз являє собою пару прямих в площині $XOY$, що перетинаються в початку координат: $y = \pm\sqrt{\dfrac{q}{p}}x$). Дійсна вісь такої гіперболи при $h > 0$ паралельна осі $OX$, а уявна – осі $OY$. При $h < 0$ – навпаки, дійсна вісь паралельна осі $OY$, а уявна – осі $OX$.

В перерізі площиною $x=\pm h$ одержимо одне з положень рухомої параболи, яка утворила своїм рухом дану поверхню: $\begin{cases} y^2 = q\left(\dfrac{h^2}{p} - 2z\right) \\ x = \pm h \end{cases}$. В перерізі площиною $y=\pm h$ також утворюється парабола: $\begin{cases} x^2 = p\left(\dfrac{h^2}{q} + 2z\right) \\ y = \pm h \end{cases}$

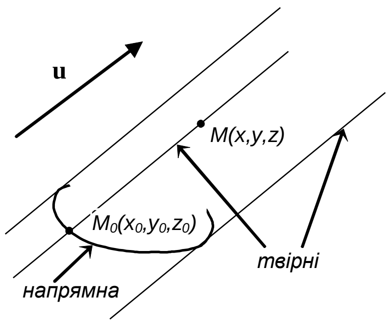
Вісь цієї параболи паралельна осі $OZ$ та співнапрямлена з нею, а вершина знаходиться нижче координатної площини $XOZ$. Загальний вигляд гіперболічного параболоїда на рис. вище нагадує «сідло».

**<u>Зауваження.</u>** Гіперболічний параболоїд так само, як і однопорожнинний гіперболоїд має два сімейства прямолінійних твірних. Пропонуємо самостійно переконатись в тому, що їх рівняннями є наступні:

$$\begin{cases} \lambda\left(\dfrac{x}{\sqrt{p}} + \dfrac{y}{\sqrt{q}}\right) = \mu \\ \mu\left(\dfrac{x}{\sqrt{p}} - \dfrac{y}{\sqrt{q}}\right) = 2\lambda z \end{cases} \quad \text{та} \quad \begin{cases} \alpha\left(\dfrac{x}{\sqrt{p}} - \dfrac{y}{\sqrt{q}}\right) = \beta \\ \beta\left(\dfrac{x}{\sqrt{p}} + \dfrac{y}{\sqrt{q}}\right) = 2\alpha z \end{cases}$$

### 1.6. Конуси та циліндри.

Окремим класом поверхонь є так звані ***конічні*** та ***циліндричні поверхні***. З точки зору загальної теорії поверхонь другого порядку, дані поверхні є виродженими (наприклад, у випадку кривих другого порядку гіпербола вироджується в пару прямих, що перетинаються, а парабола – в пару паралельних прямих). Дамо точні означення та одержимо рівняння цих поверхонь.




**<u>Означення 1.</u>** Циліндричною поверхнею (або просто циліндром) називається поверхня, утворена рухом прямої лінії, що зберігає деякий сталий напрямок (ця пряма називається ***твірною*** циліндра) і перетинає деяку криву другого порядку – ***напрямну*** циліндра (див. рис. ).

Будемо вважати, що твірна циліндра має напрямний вектор $\mathbf{u}$ з координатами $(l,m,n)$, а напрямна задана як лінія перетину двох поверхонь:

$$\begin{cases} F_1(x,y,z) = 0 \\ F_2(x,y,z) = 0 \end{cases} \qquad (10)$$

(Одна з цих поверхонь, як правило є площиною, а інша – одна з поверхонь, описаних вище).

Розглянемо довільну точку $M(x,y,z)$, що належить циліндру. Нехай $M_0(x_0,y_0,z_0)$ – та точка напрямної (1), яка належить тій самій твірній, що і точка $M$. Тоді координати точок $M$ і $M_0$ пов'язані наступним співвідношенням $\dfrac{x-x_0}{l} = \dfrac{y-y_0}{m} = \dfrac{z-z_0}{n} = \lambda$, звідки маємо:

$$x_0 = x - \lambda l,\ y_0 = y - \lambda m,\ z_0 = z - \lambda n \qquad (11)$$

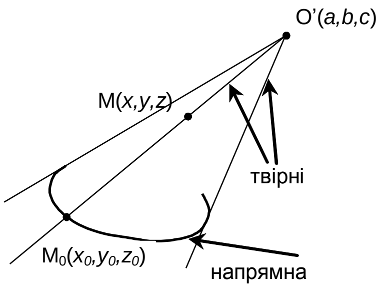
Далі підставимо вирази (11) в одне з рівнянь напрямної (10), адже точка $M_0$ має задовольняти рівняння напрямної. Виразимо звідси $\lambda$ через $x,y,z$ і знову підставимо знайдений вираз для $\lambda$ у формули (11). Залишилось тепер знайдені координати точки $M_0$ підставити у друге з рівнянь напрямної (10). Таким чином одержимо рівняння циліндра.

Реалізуємо цей алгоритм на прикладі.

<u>Приклад 1</u>. Скласти рівняння циліндра, напрямна якого задана рівняннями $x^2 - y^2 = z$; $x+y+z=0$, а твірні перпендикулярні до площини напрямної.

Оскільки напрямна лежить в площині $x+y+z=0$, то твірні цього циліндру мають напрямок нормалі цієї площини: $\mathbf{u}=(1,1,1)$. Нехай точки $M$ та $M_0$ вибрані як описано вище. Тоді рівняння твірної, що через них проходить є таким: $\frac{x-x_0}{1}=\frac{y-y_0}{1}=\frac{z-z_0}{1}=\lambda$, звідки $x_0=x-\lambda$, $y_0=y-\lambda$, $z_0=z-\lambda$. З другого рівняння напрямної $x_0+y_0+z_0=0$ знаходимо, що $\lambda=\frac{x+y+z}{3}$, а тому $x_0=x-\frac{x+y+z}{3}=\frac{2x-y-z}{3}$, $y_0=y-\frac{x+y+z}{3}=\frac{2y-x-z}{3}$, $z_0=z-\frac{x+y+z}{3}=\frac{2z-x-y}{3}$. Підставимо координати точки $M_0$ тепер у перше рівняння напрямної та спростимо одержаний вираз: $\left(\frac{2x-y-z}{3}\right)^2-\left(\frac{2y-x-z}{3}\right)^2=\frac{2z-x-y}{3}$, звідки одержимо рівняння циліндра $x^2-y^2-2xz+2yz+x+y-2z=0$.

**<u>Означення 2.</u>** Конічною поверхнею (або просто конусом) називається поверхня, утворена рухом прямої лінії, що проходить через деяку фіксовану точку – вершину конуса (ця пряма називається ***твірною*** конуса) і перетинає деяку криву другого порядку – ***напрямну*** конуса (див. рис.).




Нехай вершиною конуса є точка $O'(a,b,c)$, а напрямна як і у випадку з циліндром, задана як лінія перетину двох поверхонь: $\begin{cases}F_1(x,y,z)=0\\F_2(x,y,z)=0\end{cases}$ (12)

(Одна з цих поверхонь, як правило є площиною, а інша – одна з поверхонь, описаних вище).

Розглянемо довільну точку $M(x,y,z)$, що належить циліндру. Нехай $M_0(x_0,y_0,z_0)$ – та точка напрямної (12), яка належить тій самій твірній, що і точка $M$. Тоді координати точок $M$ і $M_0$ пов'язані наступним співвідношенням $\frac{x-a}{x_0-a}=\frac{y-b}{y_0-b}=\frac{z-c}{z_0-c}=\lambda$, звідки маємо: $x_0=\frac{x-a+a\lambda}{\lambda},\ y_0=\frac{y-b+b\lambda}{\lambda},\ z_0=\frac{z-c+c\lambda}{\lambda}$ (13)

Підставимо вирази (13) в одне з рівнянь напрямної (12), оскільки точка $M_0$ має задовольняти рівняння напрямної. Виразимо звідси $\lambda$ через $x,y,z$ і знову підставимо знайдений вираз для $\lambda$ у формули (13). Щоб одержати рівняння конуса, підставимо знайдені координати точки $M_0$ у друге з рівнянь напрямної (12).

Розглянемо приклад.

**<u>Приклад 2</u>.** Скласти рівняння конуса з вершиною в початку координат, напрямна якого дана рівняннями $x^2-2z+1=0$; $y-z+1=0$.

Нехай $M(x,y,z)$ – довільна точка циліндра, $M_0(x_0,y_0,z_0)$ – відповідна точка напрямної. Тоді їх координати задовольняють рівність: $\frac{x}{x_0}=\frac{y}{y_0}=\frac{z}{z_0}=\lambda$, звідки $x_0=\frac{x}{\lambda},\ y_0=\frac{y}{\lambda},\ z_0=\frac{z}{\lambda}$.

Оскільки $y_0-z_0+1=0$, то $\lambda=z-y$, а отже $x_0=\frac{x}{z-y},\ y_0=\frac{y}{z-y},\ z_0=\frac{z}{z-y}$. З першого рівняння напрямної, якому також задовольняє точка $M_0(x_0,y_0,z_0)$, одержимо рівняння конуса: $\left(\frac{x}{z-y}\right)^2-2\left(\frac{z}{z-y}\right)+1=0$, або $x^2+y^2-z^2=0$.

**Зауваження 1.** Циліндр, напрямні якого перпендикулярні площині твірної, називається ***прямим циліндром***. В залежності від форми напрямної його називають круговим, еліптичним, гіперболічним або параболічним.

**Зауваження 2.** Конус, твірною якого є коло, а вершина конуса проектується в центр цього кола, називається ***прямим круговим конусом***. Дослідженням перерізів можна з'ясувати, що одержана прикладі 2 поверхня – саме такий конус.

## Лекція 8. Матриці та операції з ними.
### 1. Основні означення.

**Означення 1**. ***Матрицею*** (числовою) розмірностей $m \times n$ називатимемо прямокутну таблицю дійсних чисел, що складається із $m$ рядків та $n$ стовпчиків. Елементи матриці будемо позначати $a_j^i \in \mathbf{R}$ (інколи використовують позначення з двома нижніми індексами: $a_{ij}$), де $i = \overline{1,m}$ – номер рядка, $j = \overline{1,n}$ – номер стовпчика. Таким чином,

$$\mathbf{A} = \begin{pmatrix} a_1^1 & a_2^1 & \cdots & a_n^1 \\ a_1^2 & a_2^2 & \cdots & a_n^2 \\ \vdots & \vdots & \cdots & \vdots \\ a_1^m & a_2^m & \cdots & a_n^m \end{pmatrix} = \left(a_j^i\right)_{j=\overline{1,n}}^{i=\overline{1,m}}$$

Дві матриці називатимемо ***рівними***, якщо вони мають однакові розмірності та рівні елементи з однаковими індексами. Матрицю, всі елементи якої нульові, називатимемо нуль-матрицею і позначатимемо через $\mathbf{0}$.

Матриця, яка складається з одного рядка, називається вектор-рядком. Для неї використовують наступне позначення: $\langle \mathbf{A} | = (a_1, a_2, ...., a_n)$. Матриця, яка складається з одного стовпчика, називається вектор-стовпчиком. Його позначають, як $|\mathbf{A}\rangle = \begin{pmatrix} a^1 \\ a^2 \\ \vdots \\ a^n \end{pmatrix}$.

Очевидно, матрицю можна інтерпретувати як вектор-рядок вектор-стовпчиків, або вектор-стовпчик вектор-рядків: $\mathbf{A} = \left( |\mathbf{A}_1\rangle \quad |\mathbf{A}_2\rangle \quad \cdots \quad |\mathbf{A}_n\rangle \right) = \begin{pmatrix} \langle \mathbf{A}^1 | \\ \langle \mathbf{A}^2 | \\ \vdots \\ \langle \mathbf{A}^m | \end{pmatrix}$.

**Означення 2**. Матриця, у якої рівні розмірності, тобто $m = n$, називається ***квадратною*** матрицею розмірності (порядку) $n$.

**Означення 3**. Сукупність елементів $a_1^1, a_2^2, \ldots, a_n^n \in \mathbf{R}$ називають ***головною діагоналлю*** квадратної матриці.

Серед квадратних матриць виділяють ***діагональну*** матрицю – всі елементи її рівні нулю, крім елементів головної діагоналі: $a_j^i = 0$ при $i \neq j,\ i, j = \overline{1,n}$; ***верхню*** та ***нижню трикутні*** матриці – матриці з рівними нулю елементами відповідно під головною діагоналлю: $a_j^i = 0$ при $i < j$ та над головною діагоналлю: $i > j,\ i, j = \overline{1,n}$, а також ***одиничну*** матрицю – діагональна матриця, у якої $a_i^i = 1,\ i = \overline{1,n}$. Одиничну матрицю розмірності $n$ позначатимемо $\mathbf{E}_n$:

$$\mathbf{E}_n = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}.$$

## 2. Додавання матриць та множення на скаляр.

Нехай $\mathbf{A}$ і $\mathbf{B}$ – матриці рівних розмірностей $m \times n$ з елементами $a_j^i$ та $b_j^i$ відповідно, де $\forall\, i = \overline{1,m},\ j = \overline{1,n}$.

**Означення 4**. ***Сумою*** матриць $\mathbf{A}$ і $\mathbf{B}$ називається матриця $\mathbf{C}$ розмірностей $m \times n$ з елементами $c_j^i = a_j^i + b_j^i \quad \forall\, i = \overline{1,m}, \quad j = \overline{1,n}$.

**Означення 5**. ***Добутком*** матриці $\mathbf{A}$ на число $\lambda$ називається матриця $\mathbf{C}$ з елементами $c_j^i = \lambda a_j^i\ \forall\, i = \overline{1,m},\ j = \overline{1,n}$.

Матрицею, **протилежною** до матриці $\mathbf{A}$, називатимемо матрицю $(-1)\mathbf{A}$ і позначатимемо її через $-\mathbf{A}$. Таким чином, $\mathbf{A} + (-\mathbf{A}) = \mathbf{0}$ для будь-якої матриці $\mathbf{A}$.
**Різницею** матриць $\mathbf{A}$ і $\mathbf{B}$ називатимемо матрицю $\mathbf{A} - \mathbf{B} = \mathbf{A} + (-\mathbf{B})$.

Наступна теорема є прямим наслідком властивостей додавання та множення чисел.

**Теорема 1**. Для довільних матриць $\mathbf{A}$, $\mathbf{B}$ та $\mathbf{C}$ рівних розмірностей та довільних чисел $\lambda, \mu$ справедливі наступні рівності :

1. $\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$; (комутативність)
2. $(\mathbf{A} + \mathbf{B}) + \mathbf{C} = \mathbf{A} + (\mathbf{B} + \mathbf{C})$; (асоціативність)
3. $\lambda(\mathbf{A} + \mathbf{B}) = \lambda\mathbf{A} + \lambda\mathbf{B}$; (дистрибутивність відносно додавання матриць)
4. $(\lambda + \mu)\mathbf{A} = \lambda\mathbf{A} + \mu\mathbf{A}$; (дистрибутивність відносно додавання скаляра)
5. $\lambda(\mu\mathbf{A}) = (\lambda\mu)\mathbf{A}$; (асоціативність множення на скаляр)
6. $\mathbf{A} + \mathbf{0} = \mathbf{A}$; (існування нуля)
7. $1 \cdot \mathbf{A} = \mathbf{A}$

**Зауваження**. Дана теорема означає, що множина матриць розмірностей $m \times n$ утворює ***векторний (лінійний) простір*** над полем дійсних чисел.

## 3. Транспонована матриця та її властивості.

Нехай $\mathbf{A}$ – деяка матриці розмірностей $m \times n$ з елементами $a_j^i$, де $\forall\, i = \overline{1,m},\ j = \overline{1,n}$.

**Означення 6**. Матрицею ***транспонованою*** до матриці $\mathbf{A}$ називається матриця $\mathbf{A}^T$ розмірностей $m \times n$: $\mathbf{A}^T = \left(a_i^j\right)_{i=\overline{1,m}}^{j=\overline{1,n}}$.

Іншими словами, $i$-тим стовпчиком транспонованої матриці є $i$-тий рядок вихідної і навпаки: $j$-тим рядком транспонованої матриці є $j$-тий стовпчик вихідної матриці.

Приклад 1. Нехай $\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix}$, тоді $\mathbf{A}^T = \begin{pmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{pmatrix}$;

Нехай маємо вектор-стовпчик $\left|\mathbf{b}\right\rangle = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$, тоді при транспонуванні одержимо вектор-рядок $\left|\mathbf{b}\right\rangle^T = \begin{pmatrix} 1 & 0 & -1 \end{pmatrix}$.

Очевидні властивості транспонованих матриць зібрані в наступні твердження.

**<u>Теорема 2</u>.** $\left(\mathbf{A}^T\right)^T = \mathbf{A}$

**<u>Теорема 3</u>.** $(\mathbf{A}+\mathbf{B})^T = \mathbf{A}^T + \mathbf{B}^T$

**<u>Означення 7</u>.** Квадратна матриця $\mathbf{A} = \left(a_j^i\right)_{i,\,j=\overline{1,n}}$ називається ***симетричною***, якщо $a_j^i = a_i^j \ \forall i, j = \overline{1,n}$, тобто $\mathbf{A} = \mathbf{A}^T$.

**<u>Означення 8</u>.** Квадратна матриця $\mathbf{A} = \left(a_j^i\right)_{i,\,j=\overline{1,n}}$ називається ***кососиметричною*** (або ***антисиметричною***), якщо $a_j^i = -a_i^j \ \ \forall i, j = \overline{1,n}$, тобто $\mathbf{A} = -\mathbf{A}^T$.

**<u>Зауваження</u>.** Оскільки для діагональних елементів кососиметричної матриці виконується рівність $a_i^i = -a_i^i$ , звідси випливає, що всі елементи головної діагоналі антисиметричної матриці рівні нулю: $a_i^i = 0 \ \forall i = \overline{1,n}$.

**<u>Теорема 4</u>.** Будь-яка квадратна матриця може бути подана у вигляді суми симетричної та кососиметричної матриць.

►Очевидною є наступна рівність: $\mathbf{A} = \overline{\mathbf{A}} + \overline{\overline{\mathbf{A}}} = \frac{1}{2}\left(\mathbf{A}+\mathbf{A}^T\right) + \frac{1}{2}\left(\mathbf{A}-\mathbf{A}^T\right)$, де матриця $\overline{\mathbf{A}} = \frac{1}{2}\left(\mathbf{A}+\mathbf{A}^T\right)$ є симетричною, а матриця $\overline{\overline{\mathbf{A}}} = \frac{1}{2}\left(\mathbf{A}-\mathbf{A}^T\right)$ є кососиметричною. Отже, рівність $\mathbf{A} = \overline{\mathbf{A}} + \overline{\overline{\mathbf{A}}}$ є розкладом, про який йдеться в теоремі. ◄

<u>Приклад 2</u> Матриця $\mathbf{A} = \begin{pmatrix} 1 & -1 & 0 \\ -1 & 2 & 4 \\ 0 & 4 & 3 \end{pmatrix}$ є, очевидно, симетричною згідно з означенням **7**, а матриця $\mathbf{B} = \begin{pmatrix} 0 & 1 & 5 \\ -1 & 0 & -2 \\ -5 & 2 & 0 \end{pmatrix}$ – кососиметричною, згідно з означенням **8**.

## 4. Добуток матриць.

Розглянемо дві матриці $\mathbf{A} = \left(a_j^i\right)_{j=\overline{1,n}}^{i=\overline{1,m}}$ та $\mathbf{B} = \left(b_j^i\right)_{j=\overline{1,p}}^{i=\overline{1,n}}$ .

**<u>Означення 9</u>.** Добутком матриць $\mathbf{A}$ і $\mathbf{B}$ називається матриця

$$\mathbf{C} = \mathbf{A}\cdot\mathbf{B} = \left(c_j^i\right)_{j=\overline{1,p}}^{i=\overline{1,m}}, \text{ де } c_j^i = \sum_{k=1}^{n} a_k^i b_j^k. \qquad (1)$$

**<u>Зауваження</u>.** Формулу (1), враховуючи угоду Ейнштейна про сумування по індексу, який в виразі зустрічається вгорі і внизу (він називається «німим»), можна записати також у вигляді: $c_j^i = a_k^i b_j^k,\ k = 1, n$. Надалі здебільшого будемо використовувати саме такі позначення.

Таким чином, добутком двох матриць є матриця, кількість рядків якої рівна кількості рядків першої матриці, а кількість стовпчиків – кількості стовпчиків другої. Елементами матриці-добутку є суми, складені з добутків відповідних елементів рядків першого множника та стовпчиків другого. Саме тому добуток двох матриць визначений лише тоді, коли «довжини» рядків першої матриці та стовпчиків другої матриці збігаються.

Приклад 3 Нехай маємо вектор-рядок $\langle \mathbf{a}| = \begin{pmatrix} a_1 & a_2 & \cdots & a_n \end{pmatrix}$

та вектор-стовпчик $|\mathbf{b}\rangle = \begin{pmatrix} b^1 \\ b^2 \\ \vdots \\ b^n \end{pmatrix}$. Тоді добутком $\langle \mathbf{a}| \cdot |\mathbf{b}\rangle = \sum_{k=1}^{n} a^k b_k$ є число, а добутком

$|\mathbf{b}\rangle \cdot \langle \mathbf{a}|$ – квадратна матриця розмірності $n$: $|\mathbf{b}\rangle \cdot \langle \mathbf{a}| = \begin{pmatrix} b^1 a_1 & b^1 a_2 & \cdots & b^1 a_n \\ b^2 a_1 & b^2 a_2 & \cdots & b^2 a_n \\ \vdots & \vdots & \cdots & \vdots \\ b^n a_1 & b^n a_2 & \cdots & b^n a_n \end{pmatrix}$

**Зауваження.** Цей приклад підкреслює той факт, що множення матриць не є комутативною операцією, тобто $\mathbf{A} \cdot \mathbf{B} \neq \mathbf{B} \cdot \mathbf{A}$.

Приклад 4. Нехай $\mathbf{A}$ – квадратна матриця: $\mathbf{A} = \left( a_j^i \right)_{i,j=\overline{1,n}}$, а $|\mathbf{x}\rangle = \begin{pmatrix} x^1 \\ x^2 \\ \vdots \\ x^n \end{pmatrix}$ та

$|\mathbf{b}\rangle = \begin{pmatrix} b^1 \\ b^2 \\ \vdots \\ b^n \end{pmatrix}$ – вектор-стовпчики. Тоді визначений добуток

$\mathbf{A} \cdot |\mathbf{x}\rangle = \begin{pmatrix} a_1^1 x^1 + a_2^1 x^2 + \cdots + a_n^1 x^n \\ a_1^2 x^1 + a_2^2 x^2 + \cdots + a_n^2 x^n \\ \vdots \\ a_1^n x^1 + a_2^n x^2 + \cdots + a_n^n x^n \end{pmatrix}$, отже $\mathbf{A} \cdot |\mathbf{x}\rangle = |\mathbf{b}\rangle$ є матричним записом системи $n$ лінійних рівнянь з $n$ невідомими.

Приклад 5. Нехай $\mathbf{A}$ – квадратна матриця: $\mathbf{A} = \left( a_j^i \right)_{i,j=\overline{1,n}}$, $\mathbf{E}$ – одинична матриця розмірності $n$, $\langle \mathbf{E}^i |$ та $| \mathbf{E}_j \rangle$ – відповідно її $i$-тий рядок та $j$-тий стовпчик. Тоді $\langle \mathbf{E}^i | \cdot \mathbf{A} \cdot | \mathbf{E}_j \rangle = a_j^i \quad \forall\, i, j = \overline{1,n}$.

Властивості добутку матриць описані наступною теоремою.

**Теорема 5.** Для довільних матриць $\mathbf{A}$, $\mathbf{B}$ та $\mathbf{C}$, розмірності яких допускають їх множення, справедливі рівності:

1. $(\mathbf{A} + \mathbf{B}) \cdot \mathbf{C} = \mathbf{A} \cdot \mathbf{C} + \mathbf{B} \cdot \mathbf{C}$ (дистрибутивність)
2. $\mathbf{A} \cdot (\mathbf{B} + \mathbf{C}) = \mathbf{A} \cdot \mathbf{B} + \mathbf{A} \cdot \mathbf{C}$ (дистрибутивність)

3. $(\mathbf{A}\cdot\mathbf{B})\cdot\mathbf{C}=\mathbf{A}\cdot(\mathbf{B}\cdot\mathbf{C})$ (асоціативність)

4. $(\mathbf{A}\cdot\mathbf{B})^T=\mathbf{B}^T\cdot\mathbf{A}^T$

►Доведемо першу властивість. Отже, нехай $\mathbf{A}=\left(a_j^i\right)_{j=\overline{1,n}}^{i=\overline{1,m}}$, $\mathbf{B}=\left(b_j^i\right)_{j=\overline{1,n}}^{i=\overline{1,m}}$, $\mathbf{C}=\left(c_j^i\right)_{j=\overline{1,p}}^{i=\overline{1,n}}$. Тоді, згідно з означенням **9**, елемент матриці-добутку $(\mathbf{A}+\mathbf{B})\cdot\mathbf{C}$ з індексами $i$ та $j$ є наступним:

$$[(\mathbf{A}+\mathbf{B})\cdot\mathbf{C}]_j^i=\left(a_k^i+b_k^i\right)c_j^k=a_k^i c_j^k+b_k^i c_j^k=[\mathbf{A}\cdot\mathbf{C}]_j^i+[\mathbf{B}\cdot\mathbf{C}]_j^i$$

Друге та третє твердження доводяться цілком аналогічно, тому розглянемо відразу четверту властивість. Нехай тут $\mathbf{A}=\left(a_j^i\right)_{j=\overline{1,n}}^{i=\overline{1,m}}$ та $\mathbf{B}=\left(b_j^i\right)_{j=\overline{1,p}}^{i=\overline{1,n}}$. Тоді при $\forall\, i=\overline{1,m},\ j=\overline{1,n}$: $\left[(\mathbf{A}\cdot\mathbf{B})^T\right]_j^i=[\mathbf{A}\cdot\mathbf{B}]_i^j=a_k^j b_i^k=\left[\mathbf{A}^T\right]_j^k\cdot\left[\mathbf{B}^T\right]_k^i=\left[\mathbf{B}^T\cdot\mathbf{A}^T\right]_j^i$, ◄ таким чином четверте твердження доведене.

**<u>Означення 10</u>.** Перехід від будь-якої квадратної матриці $\mathbf{A}$ до матриці $\overline{\mathbf{A}}=\frac{1}{2}\left(\mathbf{A}+\mathbf{A}^T\right)$ називається операцією симетризації матриці $\mathbf{A}$, а до матриці $\overline{\overline{\mathbf{A}}}=\frac{1}{2}\left(\mathbf{A}-\mathbf{A}^T\right)$ – операцією кососиметризації.

**<u>Зауваження</u>.** Матриця вигляду $\begin{pmatrix}\mathbf{A} & \cdots & \mathbf{B}\\ \vdots & & \vdots\\ \mathbf{C} & \cdots & \mathbf{D}\end{pmatrix}$, елементами якої є матриці, називається ***блочною***. Неважко зрозуміти, що результатом додавання двох матриць з блоками рівних розмінностей буде блочна матриці, блоки яких є сумами відповідних блоків матриць-доданків.

## Лекція 9. Визначник (детермінант) матриці.

### 1. Перестановки, транспозиції, інверсії.

Розглянемо послідовність чисел $1,2,\ldots,n$. Числа в цій послідовності можна переставити $n!$ способами – дійсно, при виборі першого числа є $n$ варіантів, при виборі другого – $n-1$ варіант, і так далі, для останнього – лише один..

**<u>Означення 1</u>.** Кожне можливе розташування цих чисел $(j_1\ j_2\cdots j_n)$, де $j_k\in\{\overline{1,n}\}$, $j_i\neq j_k$, називається ***перестановкою*** $n$ чисел $1,2,\ldots,n$.

<u>Приклад 1</u>**.** $(2143)$, $(3142)$ – перестановки 4-х чисел.

**<u>Зауваження</u>.** Очевидно, що кількість перестановок рівна $n!$. Крім того, можна розглядати перестановки елементів довільної природи, не обов'язково лише чисел. Визначальним є порядковий номер елементу в перестановці.

**<u>Означення 2</u>.** ***Транспозицією*** в перестановці $n$ чисел називається дія, внаслідок якої два елементи перестановки міняються місцями. В результаті виникає нова перестановка.

В попередньому прикладі друга перестановка одержана з першої транспозицією першого та останнього елементів.

**<u>Теорема 1</u>.** Будь-яку перестановку можна одержати з довільної перестановки за допомогою скінченої кількості транспозицій.

► Проведемо доведення методом математичної індукції по $n$.

Нехай $n=2$ . (Розглядати випадок $n=1$ немає сенсу). Тоді можливих перестановок всього дві: $(12)$, $(21)$, отже, твердження є очевидним.

Припустимо далі, що твердження справедливе для деякого значення $n=k$. Доведемо його для $n=k+1$. Отже, нехай $(i_1\, i_2 \ldots i_{k+1})$ та $(j_1\, j_2 \ldots j_{k+1})$ – дві перестановки $(k+1)$ чисел. Припустимо спочатку, що $j_1 = i_1$. Тоді перестановку $(j_2 \ldots j_{k+1})$ можна одержати з перестановки $(i_2 \ldots i_{k+1})$ за скінчену кількість транспозицій згідно з припущенням індукції, а отже, те саме стосується і перестановок $(i_1\, i_2 \ldots i_{k+1})$ та $(j_1\, j_2 \ldots j_{k+1})$. Розглянемо тепер дві перестановки $n$ чисел з різними першими елементами: $(i_1\, i_2 \ldots i_{k+1})$ та $(j_1\, j_2 \ldots i_1 \ldots j_{k+1})$. Виконавши в другій перестановці рівно одну транспозицію: $(j_1\, j_2 \ldots i_1 \ldots j_{k+1}) \to (i_1\, j_2 \ldots j_1 \ldots j_{k+1})$, зведемо задачу до попередньої, що й завершує доведення цієї теореми. ◄

**Означення 3.** В перестановці $(j_1\ j_2\ \ldots\ j_i\ \ldots\ j_k\ \ldots\ j_n)$ елементи $j_i$ та $j_k$ утворюють *інверсію*, якщо $j_i > j_k$.

Вважатимемо, що кількість інверсій в перестановці – це кількість пар, що утворюють інверсію.

**Означення 4.** Перестановка називається *парною*, якщо кількість інверсій в ній парна або рівна нулю, і *непарною* у супротивному випадку.

Приклад 2. Перестановка $(1\,2\,3 \ldots n)$ – парна, бо в ній жодної інверсії.

Приклад 3. Перестановка $(2\,1\,3\,4 \ldots n)$ – непарна, в ній одна інверсія, яку утворюють елементи $1$ і $2$.

Приклад 4. Кількість інверсій в перестановці $(n\ n\text{-}1\ \ldots 2\,1)$ рівна $n(n-1)/2$, тому її парність визначається парністю числа $n$.

**Зауваження.** Неважко зрозуміти, що кількість парних та непарних перестановок рівна відповідно $\frac{1}{2}n!$.

**Теорема 2.** Будь-яка транспозиція змінює парність перестановки на протилежну.

► Нехай маємо довільне $n>2$ (при $n=2$ все очевидно). З'ясуємо, як змінює парність перестановки $(j_1\, j_2 \ldots j_i\ j_{i+1} \ldots j_n)$ транспозиція двох сусідніх елементів, наприклад, $j_i$ та $j_{i+1}$. Отже, після транспозиції вказаних елементів одержимо перестановку $(\underbrace{j_1\, j_2 \ldots}_{1}\, j_{i+1}\, j_i\, \underbrace{\ldots j_n}_{2})$. Кількість інверсій в ній рівна сумі кількості інверсій в частинах 1, 2 та, власне, між елементами $j_{i+1}$ та $j_i$. Очевидно, що кількість інверсій в 1-ій та 2-ій частинах другої перестановки залишилась незмінною порівняно з першою перестановкою. Елементи ж $j_{i+1}$ та $j_i$ утворюють інверсію, якщо її не було в парі $j_i$ та $j_{i+1}$ і навпаки – не утворюють інверсії, якщо вона була в парі $j_i$ та $j_{i+1}$. Отже, транспозиція сусідніх елементів в перестановці збільшує або зменшує на одиницю кількість інверсій в перестановці, а отже, змінює її парність на протилежну.

Розглянемо тепер транспозицію двох елементів $j_i$ та $j_{i+s+1}$ в перестановці, між якими знаходиться рівно $s$ елементів: $(j_1 \cdots \underline{j_i}\ j_{i+1}\ j_{i+2}\ \cdots j_{i+s}\ \underline{j_{i+s+1}} \cdots j_n) \to (j_1 \cdots \underline{j_{i+s+1}}\ j_{i+1}\ j_{i+2}\ \cdots j_{i+s}\ \underline{j_i} \cdots j_n)$. Зауважимо, що той самий результат можна досягти з допомогою послідовних транспозицій сусідніх елементів, а саме: $j_i$ та $j_{i+1}$, а потім $j_i$ та $j_{i+2}$ і далі. Останньою буде транспозиція елементів $j_i$ та $j_{i+s}$. Таким чином виконаємо $s$ транспозицій сусідніх елементів і одержимо перестановку $(j_1 \ldots j_{i-1}\ j_{i+1}\ j_{i+2}\ \cdots j_{i+s}\ \underline{j_i}\ \underline{j_{i+s+1}} \cdots j_n)$. Тепер будемо просувати елемент $j_{i+s+1}$ на колишнє місце елемента $j_i$, виконуючи транспозиції сусідніх елементів: спочатку $j_{i+s+1}$ та $j_i$, потім $j_{i+s+1}$ і $j_{i+s}$, і далі до транспозиції пари $j_{i+s+1}$ та $j_{i+1}$. Кількість таких транспозицій буде $s+1$. Отже, транспозиція двох елементів $j_i$ та $j_{i+s+1}$ в перестановці еквівалентна виконанню $2s+1$ транспозицій сусідніх елементів, кожна з яких змінить парність перестановки на протилежну. Таким чином, при транспозиції двох довільних елементів перестановки її парність змінюється на протилежну. ◄

## 2. Означення визначника матриці.

***Визначник*** або ***детермінант*** *матриці* – це скаляр, який визначений для квадратної матриці $\mathbf{A} = \left(a_j^i\right)_{i,j=\overline{1,n}}$, він позначається $|\mathbf{A}| = \det \mathbf{A}$. Дамо означення цього поняття.

**<u>Означення 5</u>.** Визначником матриці $\mathbf{A} = \left(a_1^1\right) \equiv (a)$ порядку 1 називається значення єдиного елементу цієї матриці $|\mathbf{A}| = \det \mathbf{A} = a$. Визначником матриці порядку $n > 1$ називається число, яке задається наступною формулою:

$$|\mathbf{A}| = \det \mathbf{A} = \begin{vmatrix} a_1^1 & a_2^1 & \cdots & a_n^1 \\ a_1^2 & a_2^2 & \cdots & a_n^2 \\ \vdots & \vdots & & \vdots \\ a_1^n & a_2^n & \cdots & a_n^n \end{vmatrix} = \sum_{(i_1\, i_2 \ldots i_n)} a_1^{i_1} a_2^{i_2} \cdot \ldots \cdot a_n^{i_n} \cdot (-1)^t \qquad (1)$$

Тут сумування відбувається по всім можливим перестановкам $(i_1\, i_2 \ldots i_n)$ індексів рядків матриці $\mathbf{A}$, а $t$ – кількість інверсій в цій перестановці, тобто $(-1)^t$ це 1 для парної та -1 відповідно для непарної перестановки $(i_1\, i_2 \ldots i_n)$.

Таким чином, ***визначник квадратної матриці розмірності*** $n$ ***– це алгебраїчна сума*** $n!$ ***доданків, утворених всіма можливими добутками, елементів матриці, взятих по одному з кожного стовпчика та рядка із знаками плюс або мінус в залежності від парності чи непарності перестановки індексів вибраних рядків та стовпчиків.***

При цьому порядок матриці називається порядком її визначника.

<u>Приклад 5</u>**.** Для матриці розмірності 2 формула (5) набуває вигляду:

$$|\mathbf{A}| = \begin{vmatrix} a_1^1 & a_2^1 \\ a_1^2 & a_2^2 \end{vmatrix} = a_1^1 a_2^2 - a_1^2 a_2^1,$$

тобто виражає відоме раніше правило «хреста».

<u>Приклад 6</u>**.** Обчислимо визначник для матриці розмірності 3:

$$|\mathbf{A}| = \det \mathbf{A} = \begin{vmatrix} a_1^1 & a_2^1 & a_3^1 \\ a_1^2 & a_2^2 & a_3^2 \\ a_1^3 & a_2^3 & a_3^3 \end{vmatrix} =$$

$$= a_1^1 a_2^2 a_3^3 + a_1^3 a_2^1 a_3^2 + a_1^2 a_2^3 a_3^1 - a_1^3 a_2^2 a_3^1 - a_1^1 a_2^3 a_3^2 - a_1^2 a_2^1 a_3^3.$$

Таким чином, це відома раніше формула, що виражає правило «зірочки».

### 3. Властивості визначника матриці.

Нижче, якщо не буде вказано інше, розмірність всіх матриць вважатимемо рівною $n$, отже індекси елементів матриць змінюються від 1 до $n$.

**<u>Теорема 3</u>.** Визначник транспонованої матриці збігається з визначником самої матриці: $\left|\mathbf{A}^T\right| = |\mathbf{A}|$.

► $|\mathbf{A}| = \sum\limits_{(i_1\, i_2 \ldots i_n)} a_1^{i_1} a_2^{i_2} \cdot \ldots \cdot a_n^{i_n} \cdot (-1)^t$. Перегруповуємо тепер множники в кожному доданку так, щоб верхні індекси в елементах матриці були впорядковані: $a_1^{i_1} a_2^{i_2} \cdot \ldots \cdot a_n^{i_n} = a_{j_1}^1 a_{j_2}^2 \cdot \ldots \cdot a_{j_n}^n$ (очевидно, що тоді нижні індекси елементів матриці утворять деяку нову перестановку $(j_1\ j_2 \ldots j_n)$). Оскільки кількість транспозицій при переході від перестановки $(i_1\, i_2 \ldots i_n)$ до перестановки $(1 2 \ldots n)$ така сама, як і при переході від перестановки $(1 2 \ldots n)$ до перестановки $(j_1\ j_2 \ldots j_n)$, то кількість інверсій $t$ в перестановці $(i_1\, i_2 \ldots i_n)$ така сама, як і кількість інверсій $s$ в перестановці $(j_1\ j_2 \ldots j_n)$. Якщо позначити елементи транспонованої матриці через $\tilde{a}^i_j$ (тобто $\tilde{a}^i_j = a^j_i \quad \forall i, j = \overline{1, n}$) то, одержимо:

$$|\mathbf{A}| = \sum_{(i_1\, i_2 \ldots i_n)} a_1^{i_1} a_2^{i_2} \cdot \ldots \cdot a_n^{i_n} \cdot (-1)^t = \sum_{(j_1\, j_2 \ldots j_n)} a^1_{j_1} a^2_{j_2} \cdot \ldots \cdot a^n_{j_n} \cdot (-1)^t =$$

$$= \sum_{(j_1\, j_2 \ldots j_n)} \tilde{a}_1^{j_1} \tilde{a}_2^{j_2} \cdot \ldots \cdot \tilde{a}_n^{j_n} \cdot (-1)^s = \left|\mathbf{A}^T\right|$$

Отже, маємо $|\mathbf{A}| = \left|\mathbf{A}^T\right|$.◄

**Зауваження.** З попереднього твердження стає зрозумілим, що всі твердження щодо визначника, пов'язані із стовпчиками матриці, виконуються і для рядків матриці.

**Теорема 4.** Визначник діагональної матриці $\mathbf{D}$ рівний добутку елементів головної діагоналі: $|\mathbf{D}| = a_1^1 a_2^2 \cdot \ldots \cdot a_n^n$. Звідси також випливає, що $|\mathbf{E}| = 1$, $|\mathbf{0}| = 0$.

► Дійсно, в сумі, заданій формулою (1), від нуля відмінним залишиться лише один доданок, в якому $i_1 = 1,\ i_2 = 2, \ldots,\ i_n = n$, оскільки решта елементів матриці рівна нулю. ◄

**Теорема 5.** Якщо в матриці один із стовпчиків (рядків) є нульовим, то визначник такої матриці рівний нулю.

► Це випливає з того, що в кожному доданку формули (1) має міститись елемент з нульового рядка, тобто кожний доданок цієї суми, а отже і вона сама рівні нулю. ◄

**Теорема 6.** Визначник трикутної матриці (верхньої або нижньої) $|\mathbf{A}| = \left|a^i_j\right|$, де $a^i_j = 0$ відповідно при $i > j$ або $i < j$, рівний добутку елементів головної діагоналі: $|\mathbf{A}| = a_1^1 a_2^2 \cdot \ldots \cdot a_n^n$.

► Твердження досить довести лише для нижньої трикутної матриці, оскільки верхню трикутну можна одержати з неї переходом до транспонованої матриці, а ця операція не змінює визначника. Отже, розглянемо доданки в формулі (1) у випадку нижньої трикутної матриці. Оскільки $a^i_j = 0$ при $i < j$, індекс $i_n$ має бути рівним $n$, після цього індекс $i_{n-1}$ у ненульового доданку має бути рівним $n-1$, і так далі. Врешті решт зрозуміло, що ненульовий доданок в формулі (1) єдиний – рівний $(-1)^0 \cdot a_1^1 a_2^2 \cdot \ldots \cdot a_n^n = a_1^1 a_2^2 \cdot \ldots \cdot a_n^n$, що й завершує доведення. ◄

**Теорема 7.** Якщо два стовпчики (рядки) матриці поміняти місцями, визначник матриці змінить знак на протилежний, тобто:

$$|\mathbf{A}| = \left| |\mathbf{A}_1\rangle \quad \cdots \quad |\mathbf{A}_j\rangle \quad \cdots \quad |\mathbf{A}_k\rangle \quad \cdots \quad |\mathbf{A}_n\rangle \right| = -\left| |\mathbf{A}_1\rangle \quad \cdots \quad |\mathbf{A}_k\rangle \quad \cdots \quad |\mathbf{A}_j\rangle \quad \cdots \quad |\mathbf{A}_n\rangle \right|$$

► Доведемо дану властивість у випадку, коли переставляються два сусідніх стовпчики, наприклад, з індексами $k$ та $k+1$. Позначимо через $\mathbf{B}$ матрицю, одержану із матриці $\mathbf{A}$ вказаною перестановкою стовпчиків. Тоді

$$|\mathbf{B}| = \sum_{(i_1\, i_2 \ldots i_n)} a_1^{i_1} a_2^{i_2} \cdot \ldots \cdot a_{k+1}^{i_{k+1}} \cdot a_k^{i_k} \cdot \ldots \cdot a_n^{i_n} \cdot (-1)^t =$$

$$\sum_{(i_1\, i_2 \ldots i_n)} a_1^{i_1} a_2^{i_2} \cdot \ldots \cdot a_k^{i_k} \cdot a_{k+1}^{i_{k+1}} \cdot \ldots \cdot a_n^{i_n} \cdot (-1)^{t+1} = -|\mathbf{A}|$$

(Якщо через $(-1)^t$ позначити парність перестановки $(i_1\, i_2 \ldots i_{k+1}\, i_k \ldots i_n)$, то в перестановці $(i_1\, i_2 \ldots i_k\ i_{k+1} \ldots i_n)$ парність зміниться на протилежну згідно теоремі **2**, тому вона рівна $(-1)^{t+1}$).

Якщо ж переставляються стовпчики з довільними індексами $j$ та $k$ (нехай $k > j$), тоді між ними знаходиться $k - j - 1$ стовпчик, і дана перестановка стовпчиків може бути здійснена $2(k - j) - 1$ перестановок сусідніх стовпчиків. І знову по теоремі **2** знак визначника зміниться на протилежний. ◄

**Теорема 8.** Якщо матриця містить два однакових стовпчики (рядки), то її визначник рівний нулю.

► Дійсно, для такої матриці перестановка двох рівних стовпчиків приведе до рівності: $|\mathbf{A}| = -|\mathbf{A}|$, звідки й випливає, що $|\mathbf{A}| = 0$. ◄

**Теорема 9.** Якщо один із стовпчиків (рядків) матриці помножити на деяке число $\lambda$, то визначник матриці помножиться на це число:

$$\left| |\mathbf{A}_1\rangle \quad \cdots \quad |\lambda\mathbf{A}_k\rangle \quad \cdots \quad |\mathbf{A}_n\rangle \right| = \lambda|\mathbf{A}|.$$

► Очевидно, що в формулі (1) кожний з доданків суми міститиме множник $\lambda$, який можна винести за знак суми:

$$|\mathbf{A}| = \left| |\mathbf{A}_1\rangle \quad \cdots \quad |\lambda\mathbf{A}_k\rangle \quad \cdots \quad |\mathbf{A}_n\rangle \right| = \sum_{(i_1\, i_2 \ldots i_n)} a_1^{i_1} a_2^{i_2} \cdot \ldots \cdot \left(\lambda a_k^{i_k}\right) \cdot \ldots \cdot a_n^{i_n} \cdot (-1)^t =$$

$$= \lambda \cdot \sum_{(i_1\, i_2 \ldots i_n)} a_1^{i_1} a_2^{i_2} \cdot \ldots \cdot a_k^{i_k} \cdot \ldots \cdot a_n^{i_n} \cdot (-1)^t = \lambda|\mathbf{A}|$$

Іншими словами за знак визначника можна виносити будь-який ненульовий множник із довільного стовпчика або рядка. Ця властивість називається властивістю лінійності визначника. ◄

**Наслідок.** З попереднього твердження випливає, що $|\lambda\mathbf{A}| = \lambda^n \mathbf{A}$.

**Теорема 10.** Якщо в матриці два стовпчики (рядки) пропорційні, то визначник матриці рівний нулеві.

► Це прямий наслідок двох попередніх тверджень. ◄

**Теорема 11.** Якщо кожний елемент одного із стовпчиків (рядків) матриці, наприклад, з індексом $k$, можна подати у вигляді суми двох елементів: $a_k^i = b^i + c^i,\ i = \overline{1,n}$, то визначник такої матриці є сумою визначників двох матриць, всі стовпчики яких за винятком $k$-го, збігаються із стовпчиками самої матриці, а $k$-ті стовпчики утворені відповідно елементами $b^i$ та $c^i$.

► Позначимо $|\mathbf{b}\rangle$ та $|\mathbf{c}\rangle$ вектор-стовпчики, утворені елементами $b^i$ та $c^i$, відповідно. Тоді маємо:

$$|\mathbf{A}| = \left| |\mathbf{A}_1\rangle \quad \cdots \quad |\mathbf{A}_k\rangle \quad \cdots \quad |\mathbf{A}_n\rangle \right| = \left| |\mathbf{A}_1\rangle \quad \cdots \quad |\mathbf{b}+\mathbf{c}\rangle \quad \cdots \quad |\mathbf{A}_n\rangle \right| =$$

$$= \sum_{(i_1\, i_2 \ldots i_n)} a_1^{i_1} a_2^{i_2} \cdot \ldots \cdot \left(b^{i_k} + c^{i_k}\right) \cdot \ldots \cdot a_n^{i_n} \cdot (-1)^t =$$

$$= \sum_{(i_1\, i_2 \ldots i_n)} a_1^{i_1} a_2^{i_2} \cdot \ldots \cdot b^{i_k} \cdot \ldots \cdot a_n^{i_n} \cdot (-1)^t + \sum_{(i_1\, i_2 \ldots i_n)} a_1^{i_1} a_2^{i_2} \cdot \ldots \cdot c^{i_k} \cdot \ldots \cdot a_n^{i_n} \cdot (-1)^t =$$

$$= \left| |\mathbf{A}_1\rangle \quad \cdots \quad |\mathbf{b}\rangle \quad \cdots \quad |\mathbf{A}_n\rangle \right| + \left| |\mathbf{A}_1\rangle \quad \cdots \quad |\mathbf{c}\rangle \quad \cdots \quad |\mathbf{A}_n\rangle \right|,$$

що й треба було довести. ◄

**Теорема 12.** Якщо один із стовпчиків (рядків) матриці помножити на число $\lambda$ та додати до іншого стовпчика (рядка), то визначник такої матриці не зміниться.

► Дане твердження випливає безпосередньо із двох тверджень **10** та **11**:

$$\left| |\mathbf{A}_1\rangle \quad \cdots |\mathbf{A}_i\rangle \quad \cdots \quad |\lambda\mathbf{A}_i\rangle + |\mathbf{A}_j\rangle \quad \cdots \quad |\mathbf{A}_n\rangle \right| =$$

$$= \left| |\mathbf{A}_1\rangle \quad \cdots |\mathbf{A}_i\rangle \quad \cdots \quad |\lambda\mathbf{A}_i\rangle \quad \cdots \quad |\mathbf{A}_n\rangle \right| +$$

$$+\left| |\mathbf{A}_1\rangle \ \cdots \ |\mathbf{A}_i\rangle \ \cdots \ |\mathbf{A}_j\rangle \ \cdots \ |\mathbf{A}_n\rangle \right| = 0 + |\mathbf{A}| = |\mathbf{A}| \blacktriangleleft$$

Приклад 7. Розглянемо визначники матриць другого порядку:

а) $\begin{vmatrix} b^1 + c^1 & a_2^1 \\ b^2 + c^2 & a_2^2 \end{vmatrix} = \begin{vmatrix} b^1 & a_2^1 \\ b^2 & a_2^2 \end{vmatrix} + \begin{vmatrix} c^1 & a_2^1 \\ c^2 & a_2^2 \end{vmatrix}$  б) $\begin{vmatrix} \lambda a_1^1 & a_2^1 \\ \lambda a_1^2 & a_2^2 \end{vmatrix} = \lambda \begin{vmatrix} a_1^1 & a_2^1 \\ a_1^2 & a_2^2 \end{vmatrix}$

в) $\begin{vmatrix} \lambda a_1^1 & \lambda a_2^1 \\ \mu a_1^2 & \mu a_2^2 \end{vmatrix} = \lambda\mu \begin{vmatrix} a_1^1 & a_2^1 \\ a_1^2 & a_2^2 \end{vmatrix}$

**Наслідок.** Визначник матриці не зміниться, якщо до деякого стовпчика (рядка) матриці додати деяку лінійну комбінацію інших стовпчиків (рядків).

**Теорема 13.** Якщо стовпчики (рядки) матриці лінійно залежні, то визначник такої матриці рівний нулю.

► Дійсно, лінійна залежність стовпчиків матриці означає, що принаймні один з них є лінійною комбінацією інших стовпчиків. Тоді відповідно з теоремою **12** визначник даної матриці може бути поданий у вигляді лінійної комбінації визначників, кожен з яких матиме у своєму складі по два рівних стовпчики, а отже рівних нулю. ◄

**Теорема 14. (Теорема про визначник добутку матриць).** Якщо $\mathbf{A}$ і $\mathbf{B}$ квадратні матриці порядку $n$, то визначник добутку цих матриць рівний добутку їх визначників: $|\mathbf{A}\cdot\mathbf{B}| = |\mathbf{B}\cdot\mathbf{A}| = |\mathbf{A}|\cdot|\mathbf{B}|$ (Без доведення)

## 4. Обчислення визначника матриці.

**Означення 6.** Доповняльним мінором до елемента $a_j^i$ матриці $\mathbf{A}$ порядку $n$ називається визначник порядку $(n-1)$, одержаний з матриці $\mathbf{A}$ викреслюванням $i$-го рядка та $j$-го стовпчика. Одержаний визначник позначається $\overline{\mathbf{M}}_i^j$.

**Теорема 15. (Правило Лапласа обчислення визначника).** Для визначника будь-якої квадратної матриці $\mathbf{A}$ мають місце формули:

$$|\mathbf{A}| = \sum_{k=1}^{n} (-1)^{k+j} a_j^k \overline{\mathbf{M}}_k^j, \text{ де } k = \overline{1,n} \qquad (2)$$

$$|\mathbf{A}| = \sum_{k=1}^{n} (-1)^{k+i} a_k^i \overline{\mathbf{M}}_i^k, \text{ де } i = \overline{1,n} \qquad (3)$$

Формула (2) називається розкладом визначника за елементами $j$-го стовпчика, а формула (3) – розкладом визначника за елементами $i$-го рядка. Ці співвідношення дають можливість обчислювати визначники порядку $n$ рекурентно через визначники порядку $n-1$, а ті, в свою чергу, через визначники ще нижчих порядків.

**Означення 7.** Алгебраїчним доповненням до елемента $a_j^i$ матриці $\mathbf{A}$ називається число $\mathbf{A}_i^j = (-1)^{i+j}\overline{\mathbf{M}}_i^j$, тобто доповняльний мінор разом із знаком елементу.

**Зауваження.** Наведене вище означення дозволяє інакше записати формули (2) та (3):

$$|\mathbf{A}| = a_j^k \mathbf{A}_k^j \qquad \text{(сумування тільки по } k\text{)} \qquad (2^*)$$

$$|\mathbf{A}| = a_k^i \mathbf{A}_i^k \qquad \text{(сумування тільки по } k\text{)} \qquad (3^*)$$

Для доведення теореми Лапласа спершу встановимо справедливість такої леми.

**<u>Лема</u>**. Визначник матриці виду $\mathbf{A}=\begin{pmatrix} a_1^1 & 0 & \cdots & 0 \\ a_1^2 & a_2^2 & & a_n^2 \\ \vdots & & & \\ a_1^n & a_2^n & \cdots & a_n^n \end{pmatrix}$ визначається формулою:

$|\mathbf{A}| = a_1^1 \cdot \mathbf{A}_1^1$

►За означенням визначника, $|\mathbf{A}| = \sum\limits_{(i_1 i_2 \dots i_n)} a_1^{i_1} a_2^{i_2} \cdot \ldots \cdot a_n^{i_n} \cdot (-1)^t$. Очевидно, що відмінним від нуля в цій сумі буде лише доданок з $i_2 \neq 1, \ldots, i_n \neq 1$. Тому $i_1 = 1$, і оскільки $\mathbf{A}_1^1 = \sum\limits_{(i_2 \dots i_n)} a_2^{i_2} \cdot \ldots \cdot a_n^{i_n} \cdot (-1)^s$, маємо встановленою формулу леми.◄

Доведемо тепер формулу (3*), формула ж (2*) випливає з неї та теореми **3**.

►Виберемо в матриці $\mathbf{A} = \left(a_j^i\right)_{i,j=\overline{1,n}}$ рядок з індексом $i$ $(i=\overline{1,n})$ та запишемо її у вигляді суми матриць:

$\mathbf{A} = \sum\limits_{k=1}^{n} \tilde{\mathbf{A}}_k^i$, де $\tilde{\mathbf{A}}_k^i = \begin{pmatrix} a_1^1 & \cdots & a_k^1 & \cdots & a_n^1 \\ \vdots & & \vdots & & \vdots \\ 0 & \cdots & a_k^i & \cdots & 0 \\ \vdots & & & & \vdots \\ a_1^n & \cdots & a_k^n & \cdots & a_n^n \end{pmatrix}$. З теореми **11** випливає, що $|\mathbf{A}| = \sum\limits_{k=1}^{n} \left|\tilde{\mathbf{A}}_k^i\right|$. У визначниках $\left|\tilde{\mathbf{A}}_k^i\right|$ переставимо $i-1$ раз рядки та $k-1$ разів стовпчики, так, щоб елемент $a_k^i$ опинився у першому стовпчику першого рядка. На підставі теореми **7** одержимо:

$$\left|\tilde{\mathbf{A}}_k^i\right| = (-1)^{i-1+k-1} \begin{vmatrix} a_k^i & 0 & \cdots & 0 & 0 & \cdots & 0 \\ a_1^1 & a_2^1 & \cdots & a_{k-1}^1 & a_{k+1}^1 & \cdots & a_n^1 \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ a_1^{i-1} & a_2^{i-1} & \cdots & a_{k-1}^{i-1} & a_{k+1}^{i-1} & \cdots & a_n^{i-1} \\ a_1^{i+1} & a_2^{i+1} & \cdots & a_{k-1}^{i+1} & a_{k+1}^{i+1} & \cdots & a_n^{i+1} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ a_1^n & a_2^n & \cdots & a_{k-1}^n & a_{k+1}^n & \cdots & a_{n1}^n \end{vmatrix} = (-1)^{i+k}\overline{\mathbf{M}}_i^k = (-1)^{i+k}\mathbf{A}_i^k .$$

Отже, $|\mathbf{A}| = \sum\limits_{k=1}^{n} a_k^i \mathbf{A}_i^k$ ◄

**<u>Зауваження</u>.** Правило Лапласа може бути узагальнене на довільну кількість стовпчиків або рядків.

<u>Приклад 8</u>. Обчислимо визначник четвертого порядку, скориставшись методом зведення матриці до трикутного вигляду операціями, які не змінюють визначник:

$$\begin{vmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \\ 1 & 3 & 6 & 10 \\ 1 & 4 & 10 & 20 \end{vmatrix} = \begin{vmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \\ 0 & 2 & 5 & 9 \\ 0 & 3 & 9 & 19 \end{vmatrix} = \begin{vmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 3 & 10 \end{vmatrix} = \begin{vmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 1 \end{vmatrix} = 1\cdot 1\cdot 1\cdot 1 = 1$$

Перший крок полягав у тому, щоб множенням першого рядка матриці на $(-1)$ і додаванням результату по черзі до кожного наступного рядка, одержати нулі під першим елементом першого стовпчика. Другим кроком аналогічними діями одержимо нулі під другим елементом другого стовпчика, потім – під третім елементом третього стовпчика. Таким чином, залишилось обчислити визначник одержаної верхньої трикутної матриці, він рівний добутку елементів на головній діагоналі.

Приклад 9. Обчислити визначник четвертого порядку: $\Delta = \begin{vmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & a & b \\ 1 & a & 0 & c \\ 1 & b & c & 0 \end{vmatrix}$.

Скористаємось правилом Лапласа, розкриваючи визначник по першому рядку. На користь нього працює наявність одного нульового елементу в цьому рядку, що зменшує кількість подальших обчислень:

$$\Delta = 1\cdot\begin{vmatrix} 1 & a & b \\ 1 & 0 & c \\ 1 & c & 0 \end{vmatrix}\cdot(-1)^{1+2} + 1\cdot\begin{vmatrix} 1 & 0 & b \\ 1 & a & c \\ 1 & b & 0 \end{vmatrix}\cdot(-1)^{1+3} + 1\cdot\begin{vmatrix} 1 & 0 & a \\ 1 & a & 0 \\ 1 & b & c \end{vmatrix}\cdot(-1)^{1+4}.$$

Кожний з визначників третього порядку обчислимо за правилом «зірочки»:

$$\Delta = -(bc + ac - c^2) + (b^2 - ab - bc) - (ac + ab - a^2) = a^2 + b^2 + c^2 - 2(ab + bc + ac)$$

## Лекція 10. Системи $n$ лінійних рівнянь з $n$ невідомими (системи Крамерівського типу). Обернена матриця.

### 1. Теорема Крамера.

**Означення 1.** Системою $m$ лінійних рівнянь з $n$ невідомими називається система рівнянь виду:

$$\begin{cases} a_1^1 x^1 + a_2^1 x^2 + \cdots + a_n^1 x^n = b^1 \\ a_1^2 x^1 + a_2^2 x^2 + \cdots + a_n^2 x^n = b^2 \\ \cdots \\ a_1^m x^1 + a_2^m x^2 + \cdots + a_n^m x^n = b^m \end{cases}, \text{ або } \qquad \{a_j^i x^j = b^i,\ \ i = \overline{1,m};\ j = \overline{1,n} \qquad (1)$$

Матриця $\mathbf{A} = \begin{pmatrix} a_1^1 & a_2^1 & \cdots & a_n^1 \\ a_1^2 & a_2^2 & \cdots & a_n^2 \\ \vdots & \vdots & \cdots & \vdots \\ a_1^m & a_2^m & \cdots & a_n^m \end{pmatrix} = \left(a_j^i\right)_{j=\overline{1,n}}^{i=\overline{1,m}}$ називається матрицею системи, а

вектор-стовпчик $|\mathbf{b}\rangle = \begin{pmatrix} b^1 \\ b^2 \\ \vdots \\ b^m \end{pmatrix}$ – стовпчиком вільних членів. Елементи матриці $a_j^i$ є коефіцієнтами системи. Якщо вектор-стовпчик невідомих позначити через $|\mathbf{x}\rangle = \begin{pmatrix} x^1 \\ x^2 \\ \vdots \\ x^n \end{pmatrix}$, то у матричному вигляді запис системи буде таким: $\mathbf{A}\cdot|\mathbf{x}\rangle = |\mathbf{b}\rangle$, або просто $\mathbf{A}\cdot\mathbf{x} = \mathbf{b}$. Позначимо також через $(\mathbf{A}|\,\mathbf{b})$ так звану ***розширену матрицю*** системи: $(\mathbf{A}|\,\mathbf{b}) = \left(\begin{array}{cccc|c} a_1^1 & a_2^1 & \cdots & a_n^1 & b^1 \\ a_1^2 & a_2^2 & \cdots & a_n^2 & b^2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_1^m & a_2^m & \cdots & a_n^m & b^m \end{array}\right)$.

**Означення 2.** ***Розв'язком*** системи (1) називатимемо такий вектор $|\boldsymbol{\alpha}\rangle = \begin{pmatrix} \alpha^1 \\ \alpha^2 \\ \vdots \\ \alpha^n \end{pmatrix}$, який перетворює кожне рівняння системи на тотожність: $\mathbf{A}\cdot\boldsymbol{\alpha} \equiv \mathbf{b}$.

**Означення 3.** Система (1) називається *сумісною*, якщо вона має розв'язок.

Нижче розглядатимемо системи, у яких $m = n$, тобто матриця системи є квадратною. Системи такого виду будемо називати системами «крамерівського» типу. Визначник матриці цієї системи $\Delta = |\mathbf{A}|$ називається ***визначником системи***.

**Теорема (теорема Крамера).** Якщо визначник матриці системи (1) $\Delta \neq 0$, то існує **єдиний** розв'язок $|\boldsymbol{\alpha}\rangle$ цієї системи. Він визначається ***формулами Крамера***:

$$\alpha^i = \frac{\Delta^i}{\Delta},\ i = \overline{1,n}, \tag{2}$$

де $\Delta^i$ – визначник, утворений із визначника $\Delta$ заміною $i$–го стовпчика на вектор стовпчик вільних членів $|\mathbf{b}\rangle$.

►Для доведення виберемо довільний індекс $k \in \{\overline{1.n}\}$ і допишемо в розширену матрицю системи першим рядком її $k$-тий рядок. Тоді визначник одержаної квадратної матриці (позначимо його $\tilde{\Delta}$) буде рівним нулю (згадайте відповідну властивість визначника). З іншого боку, значення цього визначника можна обчислити, розкриваючи його по першому рядку згідно правилу Лапласа:

$$\tilde{\Delta} = \left|\begin{array}{cccc|c} a_1^k & a_2^k & \cdots & a_n^k & b^k \\ a_1^1 & a_2^1 & \cdots & a_n^1 & b^1 \\ a_1^2 & a_2^2 & \cdots & a_n^2 & b^2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_1^n & a_2^n & \cdots & a_n^n & b^n \end{array}\right| = \sum_{k=1}^{n}(-1)^{1+i}\cdot a_i^k\cdot \overline{\mathbf{M}}_1^i + (-1)^{1+(n+1)}\cdot b^k\cdot\Delta$$

Тут через $\overline{\mathbf{M}}_1^i, i = \overline{1,n}$ позначений доповняльний мінор елемента $a_i^k$ першого рядка цього визначника. Оскільки за припущенням теореми $\Delta \neq 0$, то з рівності $\tilde{\Delta} = 0$ випливає, що $b^k = \dfrac{\sum_{k=1}^{n}(-1)^{n+i}\cdot a_i^k\cdot\overline{\mathbf{M}}_1^i}{\Delta}$. Розглянемо уважніше значення мінору $\overline{\mathbf{M}}_1^i, i = \overline{1,n}$:

$$\overline{\mathbf{M}}_1^i = \begin{vmatrix} a_1^1 & \cdots & a_{i-1}^1 & a_{i+1}^1 & \cdots & a_n^1 & b^1 \\ a_1^2 & \cdots & a_{i-1}^2 & a_{i+1}^2 & \cdots & a_n^2 & b^2 \\ \vdots & & \vdots & \vdots & & \vdots & \vdots \\ a_1^n & \cdots & a_{i-1}^n & a_{i+1}^n & \cdots & a_n^n & b^n \end{vmatrix} = \begin{vmatrix} a_1^1 & \cdots & a_{i-1}^1 & b^1 & a_{i+1}^1 & \cdots & a_n^1 \\ a_1^2 & \cdots & a_{i-1}^2 & b^2 & a_{i+1}^2 & \cdots & a_n^2 \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_1^n & \cdots & a_{i-1}^n & b^n & a_{i+1}^n & \cdots & a_n^n \end{vmatrix} \cdot (-1)^{n-i}$$

Остання рівність є наслідком пересування стовпчика $|\mathbf{b}\rangle$ на місце між стовпчиками $|\mathbf{A}_{i-1}\rangle$ та $|\mathbf{A}_{i+1}\rangle$. На це потрібно $n-i$ кроків. Отже, $\overline{\mathbf{M}}_1^i = \Delta^i \cdot (-1)^{n-i}$. Підставимо одержане значення мінору у вираз для $b^k$: $b^k = \dfrac{\sum\limits_{k=1}^{n}(-1)^{n+i} \cdot a_i^k \cdot \Delta^i \cdot (-1)^{n-i}}{\Delta} = \sum\limits_{k=1}^{n} a_i^k \cdot \dfrac{\Delta^i}{\Delta}$, що й означає, що вектор стовпчик

$|\boldsymbol{\alpha}\rangle = \begin{pmatrix} \alpha^1 \\ \alpha^2 \\ \vdots \\ \alpha^n \end{pmatrix}$, де $\alpha^i = \dfrac{\Delta^i}{\Delta}, i = \overline{1,n}$ є розв'язком системи (1).

Покажемо, що даний розв'язок єдиний. Припустимо супротивне – нехай $|\boldsymbol{\alpha}\rangle = \begin{pmatrix} \alpha^1 \\ \alpha^2 \\ \vdots \\ \alpha^n \end{pmatrix}$ та $|\boldsymbol{\beta}\rangle = \begin{pmatrix} \beta^1 \\ \beta^2 \\ \vdots \\ \beta^n \end{pmatrix}$ різні розв'язки системи (1). Тоді із рівностей $\begin{cases} \alpha^1 \cdot |\mathbf{A}_1\rangle + \alpha^2 \cdot |\mathbf{A}_2\rangle + \cdots + \alpha^n \cdot |\mathbf{A}_n\rangle \equiv |\mathbf{b}\rangle \\ \beta^1 \cdot |\mathbf{A}_1\rangle + \beta^2 \cdot |\mathbf{A}_2\rangle + \cdots + \beta^n \cdot |\mathbf{A}_n\rangle \equiv |\mathbf{b}\rangle \end{cases}$ випливає, що

$(\alpha^1 - \beta^1) \cdot |\mathbf{A}_1\rangle + (\alpha^2 - \beta^2) \cdot |\mathbf{A}_2\rangle + \cdots + (\alpha^n - \beta^n) \cdot |\mathbf{A}_n\rangle = |\mathbf{0}\rangle$.

Це означає лінійну залежність вектор-стовпчиків матриці $\mathbf{A}$, що в свою чергу приводить до висновку, що $\Delta = 0$. Одержана суперечність з умовою теореми означає єдиність розв'язку (2). ◄

**Зауваження 1.** Випадок $\Delta = 0$ обов'язково буде досліджений пізніше.

**Зауваження 2.** Перестановка рядків та множення довільного рівняння системи (1) з додаванням результату до іншого рівняння, очевидно не змінює розв'язку (2) системи. Виконання цих операцій з рівняннями системи еквівалентне діям з рядками розширеної матриці системи. Як відомо, ці операції не перетворять визначник системи на нуль, отже, називатимемо їх допустимими перетвореннями системи (1).

## 2. Метод Гаусса (метод послідовних виключень невідомих).

Враховуючи зауваження **2**, замість перетворень рівнянь системи (1) будемо перетворювати рядки розширеної матриці системи.

Ідея методу Гаусса розв'язування системи (1) полягає в тому, щоб шляхом допустимих перетворень звести матрицю системи (1) спочатку до вигляду верхньої трикутної матриці (прямий хід методу Гаусса), а потім – до діагонального виду (обернений хід). Після цього розв'язок системи записується легко (або встановлюється факт неможливості це зробити). Отже, розглянемо розширену матрицю системи (1):

$$(\mathbf{A}|\,\mathbf{b}) = \left(\begin{array}{cccc|c} a_1^1 & a_2^1 & \cdots & a_n^1 & b^1 \\ a_1^2 & a_2^2 & \cdots & a_n^2 & b^2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_1^n & a_2^n & \cdots & a_n^n & b^n \end{array}\right)$$

1. Припустимо, що коефіцієнт $a_1^1 \neq 0$ (якщо це не так, то переставимо відповідним чином рядки матриці).
2. Множитимемо перший рядок матриці $(\mathbf{A}|\,\mathbf{b})$ на $\left(-\dfrac{a_1^i}{a_1^1}\right)$, $i=\overline{2,n}$ та додаватимемо по черзі до кожного наступного рядка. Наша мета – одержати нулі в першому стовпчику матриці під елементом $a_1^1$.
3. Розширена матриця набуде вигляду: $(\mathbf{A}|\,\mathbf{b}) = \left(\begin{array}{cccc|c} a_1^1 & a_2^1 & \cdots & a_n^1 & b^1 \\ 0 & \tilde{a}_2^2 & \cdots & \tilde{a}_n^2 & \tilde{b}^2 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & \tilde{a}_2^n & \cdots & \tilde{a}_n^n & \tilde{b}^n \end{array}\right)$

   Тут $\tilde{a}_j^k = a_j^k + \left(-\dfrac{a_1^k}{a_1^1}\right)\cdot a_j^1$, $\tilde{b}^k = b^k + \left(-\dfrac{a_1^k}{a_1^1}\right)\cdot b^1$, $k, j=\overline{2,n}$.
4. Потім розглянемо коефіцієнт $\tilde{a}_2^2$, вважаючи, що він не рівний нулю, і одержимо нулі в другому стовпчику матриці під ним аналогічним чином.
   Врешті решт розширена матриця системи набуде вигляду, в якому основна матриця система є верхньою трикутною матрицею:
   $$(\mathbf{A}|\,\mathbf{b}) = \left(\begin{array}{cccc|c} a_1^1 & a_2^1 & \cdots & a_n^1 & b^1 \\ 0 & \tilde{a}_2^2 & \cdots & \tilde{a}_n^2 & \tilde{b}^2 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & \tilde{\tilde{a}}_n^n & \tilde{\tilde{b}}^n \end{array}\right)$$

Якщо визначник системи $\Delta \neq 0$, то жоден з елементів на головній діагоналі матриці $\mathbf{A}$, очевидно, не рівний нулю, тому розв'язок системи єдиний. Перейдемо до оберненого ходу методу Гаусса.

5. Отже, помножимо останній рядок матриці на $\left(-\dfrac{\tilde{\tilde{a}}_n^i}{\tilde{\tilde{a}}_n^n}\right)$, $i=\overline{1,n-1}$ та додамо до всіх рядків по черзі від $(n-1)$-го до першого. Таким чином одержимо нулі в останньому стовпчику основної матриці над елементом $\tilde{\tilde{a}}_n^n$.
6. Повторимо цю процедуру для всіх стовпчиків від $(n-1)$-го до другого. Одержимо розширену матрицю системи, зведену до діагонального вигляду:
7. $(\mathbf{A}|\,\mathbf{b}) = \left(\begin{array}{cccc|c} d_1^1 & 0 & \cdots & 0 & \tilde{\tilde{b}}^1 \\ 0 & d_2^2 & \cdots & 0 & \tilde{\tilde{b}}^2 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & d_n^n & \tilde{\tilde{b}}^n \end{array}\right)$

Останній крок полягає в тому, щоб одержати зліва одиничну матрицю (оскільки $\Delta \neq 0$, то жоден з елементів $d_i^i \neq 0, i = \overline{1,n}$): $\left(\mathbf{A}\middle|\,\mathbf{b}\right) = \left(\begin{array}{cccc|c} 1 & 0 & \cdots & 0 & \frac{\tilde{\tilde{b}}^1}{d_1^1} \\ 0 & 1 & \cdots & 0 & \frac{\tilde{\tilde{b}}^2}{d_2^2} \\ \vdots & \vdots & & \vdots & \\ 0 & 0 & \cdots & 1 & \frac{\tilde{\tilde{b}}^n}{d_n^n} \end{array}\right)$. Ця розширена матриця визначає розв'язок системи (1): $\begin{cases} x^1 = \dfrac{\tilde{\tilde{b}}^1}{d_1^1} \\ x^2 = \dfrac{\tilde{\tilde{b}}^2}{d_2^2} \\ \cdots \\ x^n = \dfrac{\tilde{\tilde{b}}^n}{d_n^n} \end{cases}$

**Зауваження.** Метод Гаусса полягає фактично в приведенні допустимими перетвореннями рядків розширеної матриці системи від вигляду $\left(\mathbf{A}\middle|\,\mathbf{b}\right)$ до вигляду $\left(\mathbf{E}\middle|\,\boldsymbol{\alpha}\right)$, де вектор стовпчик $\left|\boldsymbol{\alpha}\right\rangle$ і буде розв'язком системи.

Приклад. Розв'язати систему рівнянь: $\begin{cases} x^1 + 2x^2 + 5x^3 = -9 \\ x^1 - x^2 + 3x^3 = 2 \\ 3x^1 - 6x^2 - x^3 = 25 \end{cases}$

Піддамо розширену матрицю системи допустимим перетворенням згідно алгоритму методу Гаусса:

$$\left(\begin{array}{ccc|c} 1 & 2 & 5 & -9 \\ 1 & -1 & 3 & 2 \\ 3 & -6 & -1 & 25 \end{array}\right) = \left(\begin{array}{ccc|c} 1 & 2 & 5 & -9 \\ 0 & -3 & -2 & 11 \\ 0 & -12 & -16 & 52 \end{array}\right) = \left(\begin{array}{ccc|c} 1 & 2 & 5 & -9 \\ 0 & -3 & -2 & 11 \\ 0 & 0 & -8 & 8 \end{array}\right) =$$

$$= \left(\begin{array}{ccc|c} 1 & 2 & 5 & -9 \\ 0 & -3 & -2 & 11 \\ 0 & 0 & -8 & 8 \end{array}\right) = \left(\begin{array}{ccc|c} 1 & 2 & 5 & -9 \\ 0 & -3 & -2 & 11 \\ 0 & 0 & 1 & -1 \end{array}\right) = \left(\begin{array}{ccc|c} 1 & 2 & 0 & -4 \\ 0 & -3 & 0 & 9 \\ 0 & 0 & 1 & -1 \end{array}\right) =$$

$$= \left(\begin{array}{ccc|c} 1 & 2 & 0 & -4 \\ 0 & 1 & 0 & -3 \\ 0 & 0 & 1 & -1 \end{array}\right) = \left(\begin{array}{ccc|c} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & -3 \\ 0 & 0 & 1 & -1 \end{array}\right),$$

звідки знаходимо: $\begin{cases} x^1 = 2 \\ x^2 = -3 \\ x^3 = -1 \end{cases}$ – розв'язок даної системи.

### 3. Означення оберненої матриці.

Матриці, які розглядатимуться в цьому розділі, вважаються квадратними матрицями порядку $n$.

**Означення 1.** Квадратна матриця $\mathbf{A} = \left(a^i_j\right)_{i,j=\overline{1,n}}$ називається ***виродженою***, якщо $|\mathbf{A}| = 0$ і ***невиродженою***, якщо $|\mathbf{A}| \neq 0$.

**Означення 2.** Квадратна матриця, що позначатиметься $\mathbf{A}^{-1}$, називається *правою оберненою* до невиродженої квадратної матриці $\mathbf{A}$, якщо $\mathbf{A} \cdot \mathbf{A}^{-1} = \mathbf{E}$ (3)

**Зауваження.** Права обернена матриця виникає із-за не комутативності множення матриць. Якщо позначити елементи правої оберненої матриці як $\left(\tilde{a}^i_j\right)_{i,j=\overline{1,n}} = \mathbf{A}^{-1}$ , то рівність (1) можна записати у вигляді:

$a^i_j \tilde{a}^j_k = \delta^i_k$, $j = \overline{1,n}$, де $\delta^i_k = \begin{cases} 1, i = k \\ 0, i \neq k \end{cases}$ – символ Кронекера.

З означення **2.** випливають наступні наслідки.

**Теорема 1.** Якщо матриця $\mathbf{A}$ невироджена , то і матриця $\mathbf{A}^{-1}$ невироджена.

► Дійсно, $\left|\mathbf{A} \cdot \mathbf{A}^{-1}\right| = |\mathbf{A}| \cdot \left|\mathbf{A}^{-1}\right| = |\mathbf{E}| = 1$, тому з $|\mathbf{A}| \neq 0$ випливає $\left|\mathbf{A}^{-1}\right| \neq 0$. ◄

**Теорема 2.** Правою оберненою матрицею до $\mathbf{A}^{-1}$ є сама матриця $\mathbf{A}$, тобто $\left(\mathbf{A}^{-1}\right)^{-1} = \mathbf{A}$.

► Помножимо рівність $\left(\mathbf{A}^{-1}\right) \cdot \left(\mathbf{A}^{-1}\right)^{-1} = \mathbf{E}$ на $\mathbf{A}$ зліва: $\mathbf{A} \cdot \left(\mathbf{A}^{-1}\right) \cdot \left(\mathbf{A}^{-1}\right)^{-1} = \mathbf{A} \cdot \mathbf{E}$, або $\mathbf{E} \cdot \left(\mathbf{A}^{-1}\right)^{-1} = \mathbf{A}$, що й треба було показати. ◄

**Теорема 3.** Права обернена до $\mathbf{A}$ матриця є й лівою оберненою, тобто: $\mathbf{A}^{-1} \cdot \mathbf{A} = \mathbf{E}$.

► Дійсно, $\left(\mathbf{A}^{-1}\right) \cdot \mathbf{A} = \left(\mathbf{A}^{-1}\right) \cdot \left(\mathbf{A}^{-1}\right)^{-1} = \mathbf{E}$, або $\tilde{a}^i_j a^j_k = \delta^i_k$, $j = \overline{1,n}$. ◄

Ця властивість дозволяє сформулювати наступне означення.

**Означення 3.** Квадратна матриця, $\mathbf{A}^{-1}$, називається ***оберненою*** до невиродженої квадратної матриці $\mathbf{A}$, якщо $\mathbf{A} \cdot \mathbf{A}^{-1} = \mathbf{A}^{-1} \cdot \mathbf{A} = \mathbf{E}$. (4)

**Теорема 4.** Обернена матриця єдина.

► Від супротивного припустимо, що $\tilde{\mathbf{A}}^{-1}$ та $\tilde{\tilde{\mathbf{A}}}^{-1}$ дві обернені матриці до матриці $\mathbf{A}$. Тоді

$\tilde{\mathbf{A}}^{-1} = \tilde{\mathbf{A}}^{-1} \cdot \mathbf{E} = \tilde{\mathbf{A}}^{-1} \cdot \left(\mathbf{A} \cdot \tilde{\tilde{\mathbf{A}}}^{-1}\right) = \left(\tilde{\mathbf{A}}^{-1} \cdot \mathbf{A}\right) \cdot \tilde{\tilde{\mathbf{A}}}^{-1} = \mathbf{E} \cdot \tilde{\tilde{\mathbf{A}}}^{-1} = \tilde{\tilde{\mathbf{A}}}^{-1}$. ◄

**Теорема 5.** $(\mathbf{A} \cdot \mathbf{B})^{-1} = \mathbf{B}^{-1} \cdot \mathbf{A}^{-1}$

► Зауважимо спочатку, що для невиродженості добутку матриць $\mathbf{A} \cdot \mathbf{B}$ необхідно, що були не виродженими множники, тому обернені матриці до матриць $\mathbf{A}$ та $\mathbf{B}$ мають існувати. Розглянемо далі добуток матриць $(\mathbf{A} \cdot \mathbf{B}) \cdot \left(\mathbf{B}^{-1} \cdot \mathbf{A}^{-1}\right) = \mathbf{A} \cdot \left(\mathbf{B} \cdot \mathbf{B}^{-1}\right) \cdot \mathbf{A}^{-1} = \mathbf{A} \cdot \mathbf{A}^{-1} = \mathbf{E}$. Звідси випливає, що $\mathbf{B}^{-1} \cdot \mathbf{A}^{-1}$ є оберненою матрицею до $\mathbf{A} \cdot \mathbf{B}$. ◄

**Теорема 6.** $\left(\mathbf{A}^T\right)^{-1} = \left(\mathbf{A}^{-1}\right)^T$

► Оскільки матриці $\mathbf{A}$ та $\mathbf{A}^T$ невироджені одночасно, то обернена до $\mathbf{A}^T$ існує. Розглянемо $\mathbf{A}^T \cdot \left(\mathbf{A}^{-1}\right)^T = \left(\mathbf{A}^{-1} \cdot \mathbf{A}\right)^T = \mathbf{E}^T = \mathbf{E}$, отже, $\left(\mathbf{A}^{-1}\right)^T$ є оберненою до $\mathbf{A}^T$. ◄

## 4. Обчислення оберненої матриці.

Знайдемо тепер вираз для елементів $\tilde{a}^i_j, i, j = \overline{1,n}$ оберненої матриці $\mathbf{A}^{-1}$. Будемо розглядати рівність (1) у вигляді $n$ систем лінійних рівнянь з $n$ невідомими, якими є елементи стовпчиків $\left|\tilde{\mathbf{A}}_i\right\rangle = \begin{pmatrix} \tilde{a}^1_i \\ \tilde{a}^2_i \\ \vdots \\ \tilde{a}^n_i \end{pmatrix}$, $i = \overline{1,n}$ матриці $\mathbf{A}^{-1}$, тобто $\mathbf{A} \cdot \left|\tilde{\mathbf{A}}_i\right\rangle = \left|\mathbf{E}_i\right\rangle,\ i = \overline{1,n}$ (5)

Зафіксуємо індекс $i$ в системі (3), визначник якої $|\mathbf{A}| \neq 0$. Згідно з теоремою Крамера існує єдиний розв'язок системи $\left|\tilde{\mathbf{A}}_i\right\rangle$, що задається формулами: $\tilde{a}^j_i = \dfrac{\Delta^j\left(\left|\mathbf{E}_i\right\rangle\right)}{|\mathbf{A}|},\ j = \overline{1,n}$. Тут через $\Delta^j\left(\left|\mathbf{E}_i\right\rangle\right)$ позначено визначник $|\mathbf{A}|$, у якого $j$-тий стовпчик замінено на стовпчик вільних членів $\left|\mathbf{E}_i\right\rangle$ системи (3). Розглянемо уважніше цей визначник:

$$\Delta^j\left(\left|\mathbf{E}_i\right\rangle\right) = \begin{vmatrix} a^1_1 & \cdots & a^1_{j-1} & 0 & a^1_{j+1} & \cdots & a^1_n \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a^i_1 & \cdots & a^i_{j-1} & 1 & a^i_{j+1} & \cdots & a^i_n \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a^n_1 & \cdots & a^n_{j-1} & 0 & a^n_{j+1} & \cdots & a^n_n \end{vmatrix} = 1 \cdot (-1)^{i+j} \cdot \overline{\mathbf{M}}^j_i = \mathbf{A}^j_i$$

Тут $\overline{\mathbf{M}}^j_i$ та $\mathbf{A}^j_i$ стандартні позначення відповідно доповняльного мінору та алгебраїчного доповнення елементу $a^i_j$ матриці $\mathbf{A}$. Остаточна формула для елементів оберненої матриці набуває вигляду: $\tilde{a}^j_i = \dfrac{\mathbf{A}^j_i}{|\mathbf{A}|},\ i,j = \overline{1,n}$ (6)

<u>**Теорема 7**</u>**.** З теорем **1** та **3** і формули (4) випливає ***критерій існування оберненої матриці***: обернена до матриці $\mathbf{A}$ існує тоді і тільки тоді, коли матриця $\mathbf{A}$ не вироджена.

<u>**Правило знаходження оберненої матриці**</u>**.** Отже, щоб визначити елементи матриці, оберненої до деякої матриці $\mathbf{A}$, необхідно, перш за все, обчислити визначник матриці $\mathbf{A}$. Якщо $|\mathbf{A}| \neq 0$, то обернена матриці існує. ***Елементи її $j$-го рядка дорівнюють алгебраїчним доповненням до елементів $j$-го стовпчика матриці $\mathbf{A}$, поділеним на визначник матриці***.

<u>**Метод Гаусса знаходження оберненої матриці**</u>**.**

Відшукання елементів оберненої матриці зводиться до розв'язування $n$ систем (3), де вектор-стовпчик $\left|\tilde{\mathbf{A}}_i\right\rangle$ є невідомим. Якщо застосувати для пошуку розв'язків даних систем метод Гаусса, то необхідно буде перетворювати розширені матриці цих систем, які мають вигляд: $\left(\mathbf{A}\middle|\,\mathbf{E}_i\right), i = \overline{1,n}$, до вигляду $\left(\mathbf{E}\middle|\,\tilde{\mathbf{A}}_i\right), i = \overline{1,n}$. Оскільки при цьому будуть виконуватись ті самі перетворення, то чому б не застосувати їх одночасно до всіх стовпчиків одиничної матриці? Отже, метод Гаусса відшукання оберненої матриці зводиться до виконання допустимих перетворень з «надрозширеною» матрицею $\left(\mathbf{A}\middle|\,\mathbf{E}\right)$. Звівши її до вигляду $\left(\mathbf{E}\middle|\,\tilde{\mathbf{A}}\right) \equiv \left(\mathbf{E}\middle|\,\mathbf{A}^{-1}\right)$, одержимо обернену матрицю.

Приклад 1. Знайти обернену матрицю для невиродженої матриці розмірності 2: $\mathbf{A}=\begin{pmatrix} a & b \\ c & d \end{pmatrix}$.

За умовою, $|\mathbf{A}|=ad-bc\neq 0$. Отже, оскільки алгебраїчним доповненням до елемента $a$ є елемент $d$, і навпаки, а для елемента $b$– $(-c)$, так само, як і для елемента $c$ – $(-b)$, то оберненою є матриця: $\mathbf{A}^{-1}=\dfrac{1}{ad-bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$.

Приклад 2. Знайти обернену матрицю для матриці $\mathbf{A}=\begin{pmatrix} 1 & -1 & 2 \\ 0 & -3 & 1 \\ 1 & -2 & 0 \end{pmatrix}$.

1 спосіб (метод алгебраїчних доповнень – правило знаходження оберненої матриці).

Знайдемо визначник матриці $|\mathbf{A}|=-1+6+2=7$. Отже, матриця невироджена. Шукатимемо елементи оберненої матриці, які знаходяться в першому її стовпчику. Для цього обчислимо алгебраїчні доповнення елементів першого рядка матриці $\mathbf{A}$: $\mathbf{A}_1^1=(-1)^{1+1}\cdot 2=2$, $\mathbf{A}_1^2=(-1)^{1+2}\cdot(-1)=1$, $\mathbf{A}_1^3=(-1)^{1+3}\cdot 3=3$. Аналогічно знаходимо алгебраїчні доповнення елементів другого рядка: $\mathbf{A}_2^1=(-1)^{1+2}\cdot 4=-4$, $\mathbf{A}_2^2=(-1)^{2+2}\cdot(-2)=-2$, $\mathbf{A}_2^3=(-1)^{2+3}\cdot(-1)=1$ та третього рядка: $\mathbf{A}_3^1=(-1)^{1+3}\cdot 5=5$, $\mathbf{A}_3^2=(-1)^{2+3}\cdot 1=-1$, $\mathbf{A}_3^3=(-1)^{3+3}\cdot(-3)=-3$. Таким чином, обернена матриця є такою:

$$\mathbf{A}^{-1}=\frac{1}{7}\begin{pmatrix} 2 & -4 & 5 \\ 1 & -2 & -1 \\ 3 & 1 & -3 \end{pmatrix}=\begin{pmatrix} \frac{2}{7} & \frac{-4}{7} & \frac{5}{7} \\ \frac{1}{7} & \frac{-2}{7} & \frac{-1}{7} \\ \frac{3}{7} & \frac{1}{7} & \frac{-3}{7} \end{pmatrix}$$

2 спосіб (метод Гаусса).

Запишемо розширену матрицю і будемо виконувати очевидні перетворення над її рядками:

$$(\mathbf{A}|\,\mathbf{E})=\left(\begin{array}{ccc|ccc} 1 & -1 & 2 & 1 & 0 & 0 \\ 0 & -3 & 1 & 0 & 1 & 0 \\ 1 & -2 & 0 & 0 & 0 & 1 \end{array}\right)\to\left(\begin{array}{ccc|ccc} 1 & -1 & 2 & 1 & 0 & 0 \\ 0 & -3 & 1 & 0 & 1 & 0 \\ 0 & -1 & -2 & -1 & 0 & 1 \end{array}\right)\to\left(\begin{array}{ccc|ccc} 1 & -1 & 2 & 1 & 0 & 0 \\ 0 & 1 & -\frac{1}{3} & 0 & -\frac{1}{3} & 0 \\ 0 & -1 & -2 & -1 & 0 & 1 \end{array}\right)\to$$

$$\to\left(\begin{array}{ccc|ccc} 1 & -1 & 2 & 1 & 0 & 0 \\ 0 & 1 & -\frac{1}{3} & 0 & -\frac{1}{3} & 0 \\ 0 & 0 & -\frac{7}{3} & -1 & -\frac{1}{3} & 1 \end{array}\right)\to\left(\begin{array}{ccc|ccc} 1 & -1 & 2 & 1 & 0 & 0 \\ 0 & 1 & -\frac{1}{3} & 0 & -\frac{1}{3} & 0 \\ 0 & 0 & 1 & \frac{3}{7} & \frac{1}{7} & -\frac{3}{7} \end{array}\right)\to\left(\begin{array}{ccc|ccc} 1 & -1 & 0 & \frac{1}{7} & -\frac{2}{7} & \frac{6}{7} \\ 0 & 1 & 0 & \frac{1}{7} & -\frac{2}{7} & -\frac{1}{7} \\ 0 & 0 & 1 & \frac{3}{7} & \frac{1}{7} & -\frac{3}{7} \end{array}\right)\to$$

$$\to\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{2}{7} & -\frac{4}{7} & \frac{5}{7} \\ 0 & 1 & 0 & \frac{1}{7} & -\frac{2}{7} & -\frac{1}{7} \\ 0 & 0 & 1 & \frac{3}{7} & \frac{1}{7} & -\frac{3}{7} \end{array}\right)=(\mathbf{E}|\,\mathbf{A}^{-1}),\text{ отже, } \mathbf{A}^{-1}=\begin{pmatrix} \frac{2}{7} & \frac{-4}{7} & \frac{5}{7} \\ \frac{1}{7} & \frac{-2}{7} & \frac{-1}{7} \\ \frac{3}{7} & \frac{1}{7} & \frac{-3}{7} \end{pmatrix}.$$

### 5. Матричні рівняння.

Нехай $\mathbf{A}$ і $\mathbf{B}$ задані квадратні матриці порядку $n$.

**Означення 4.** Рівняння виду $\mathbf{A}\cdot\mathbf{X}=\mathbf{B}$, або (7)

$\mathbf{X}\cdot\mathbf{A}=\mathbf{B}$, (8)

де $\mathbf{X}$ – невідома квадратна матриця порядку $n$, називаються *матричними рівняннями*.

Рівняння (5) та (6) будемо розв'язувати в припущення, що матриця $\mathbf{A}$ – невироджена. Для розв'язку рівняння (5) помножимо його обидві частини зліва на $\mathbf{A}^{-1}$. Одержимо:

$\mathbf{A}^{-1}\cdot\mathbf{A}\cdot\mathbf{X}=\mathbf{A}^{-1}\cdot\mathbf{B}$, тобто $\mathbf{X}=\mathbf{A}^{-1}\cdot\mathbf{B}$. Для розв'язку рівняння (6) помножимо його обидві частини справа на $\mathbf{A}^{-1}$. Одержимо: $\mathbf{X}\cdot\mathbf{A}\cdot\mathbf{A}^{-1}=\mathbf{B}\cdot\mathbf{A}^{-1}$, звідки $\mathbf{X}=\mathbf{B}\cdot\mathbf{A}^{-1}$.

**Зауваження.** Міркування, подібні наведеним вище, дозволяють розв'язувати рівняння виду (5) методом Гаусса. Якщо розширену матрицю $(\mathbf{A}|\,\mathbf{B})$ перетворити допустимими перетвореннями рядків так, щоб ліва матриця стала одиничною, то права матриця буде шуканою матрицею $\mathbf{X}$, тобто $(\mathbf{A}|\,\mathbf{B})\to\left(\mathbf{E}\middle|\,\mathbf{A}^{-1}\cdot\mathbf{B}\right)\equiv(\mathbf{E}|\,\mathbf{X})$. Цей спосіб безпосередньо не може бути застосований до рівняння (6). Проте, якщо замість цього рівняння розглянути еквівалентне йому рівняння $\mathbf{A}^T\cdot\mathbf{X}^T=\mathbf{B}^T$, то метод Гаусса тут спрацьовує таким чином: $\left(\mathbf{A}^T|\,\mathbf{B}^T\right)\to\left(\mathbf{E}\middle|\left(\mathbf{A}^T\right)^{-1}\cdot\mathbf{B}^T\right)\equiv\left(\mathbf{E}\middle|\,\mathbf{X}^T\right)$.

Отже, шуканою буде матриця, транспонована до правої частини перетвореної розширеної матриці.

Приклад 3. Розв'язати матричне рівняння: $\mathbf{X}\cdot\begin{pmatrix}1&-1\\3&4\end{pmatrix}=\begin{pmatrix}-1&1\\1&2\end{pmatrix}$.

Це рівняння виду (6). Тут $\mathbf{A}=\begin{pmatrix}1&-1\\3&4\end{pmatrix}$, $\mathbf{B}=\begin{pmatrix}-1&1\\1&2\end{pmatrix}$. Використаємо метод Гаусса:

$$\left(\mathbf{A}^T|\,\mathbf{B}^T\right)=\left(\begin{array}{cc|cc}1&3&-1&1\\-1&4&1&2\end{array}\right)\to\left(\begin{array}{cc|cc}1&3&-1&1\\0&7&0&3\end{array}\right)\to\left(\begin{array}{cc|cc}1&3&-1&1\\0&1&0&\frac{3}{7}\end{array}\right)\to\to\left(\begin{array}{cc|cc}1&0&-1&\frac{-2}{7}\\0&1&0&\frac{3}{7}\end{array}\right)$$

Отже, $\mathbf{X}=\begin{pmatrix}-1&0\\\frac{-2}{7}&\frac{3}{7}\end{pmatrix}$. У вірності відповіді можна переконатись безпосередньо.

**Завдання для самостійної роботи.** Довести, що для блочних матриць виконуються рівності:

1. $\begin{pmatrix}\mathbf{A}&\mathbf{B}\\\mathbf{0}&\mathbf{D}\end{pmatrix}^{-1}=\begin{pmatrix}\mathbf{A}^{-1}&-\mathbf{A}^{-1}\mathbf{B}\mathbf{D}^{-1}\\\mathbf{0}&\mathbf{D}^{-1}\end{pmatrix}$  2. $\begin{pmatrix}\mathbf{A}&\mathbf{0}\\\mathbf{C}&\mathbf{D}\end{pmatrix}^{-1}=\begin{pmatrix}\mathbf{A}^{-1}&\mathbf{0}\\-\mathbf{D}^{-1}\mathbf{C}\mathbf{A}^{-1}&\mathbf{D}^{-1}\end{pmatrix}$

## Лекція 11. Системи $n$ лінійних рівнянь з $m$ невідомими.

Нижче розглядається загальна теорія систем з лінійних рівнянь. Вище вже згадувались системи лінійних рівнянь (див. лекцію 10). Одним з визначних понять в цій теорії є ранг матриці.

### 1. Ранг матриці.

Розглянемо матрицю $\mathbf{A}=\left(a_j^i\right)_{j=\overline{1,n}}^{i=\overline{1,m}}$ розмірностей $m\times n$.

**Означення 1.** Мінором $\mathbf{M}_{j_1j_2\ldots j_r}^{i_1i_2\ldots i_r}$ порядку $r$ матриці $\mathbf{A}$ називається визначник, розташований на перехресті $r$ рядків з індексами $i_1,i_2,\ldots,i_r$ та $r$ стовпчиків з індексами $j_1,j_2,\ldots,j_r$.

**Зауваження.** Раніше (див. лекції 9-10) визначався доповняльний мінор до елементу матриці – порівняйте позначення.

**Означення 2.** Мінор $\mathbf{M}$ матриці $\mathbf{A}$ порядку $r$ називається ***базисним***, якщо $\mathbf{M}\neq 0$, а всі мінори порядків $r+1$ рівні нулю, або $r=\min(m,n)$ (тобто мінорів порядку $r+1$ не існує).

**Наслідок 1.** З теореми Лапласа випливає, що якщо порядок базисного мінору рівний $r$, то всі мінори порядків вищих за $r+1$ теж рівні нулю.

**Наслідок 2.** Базисний мінор, очевидно, визначений неоднозначно.

**Наслідок 3.** Будь-яка матриця, крім нуль-матриці, має базисний мінор.

**Означення 3.** Стовпчики і рядки матриці $\mathbf{A}$, які містять базисний мінор, називаються ***базисними***.

**Означення 4.** ***Рангом*** матриці $\mathbf{A}$ (позначається $\operatorname{Rank}\mathbf{A}$ або $\operatorname{Rg}\mathbf{A}$) називається ***порядок її базисного мінору***.

Наступна теорема не тільки описує вплив перетворень матриці на її ранг, але й дає спосіб визначення базисного мінору.

**<u>Теорема 1</u>.** Ранг матриці не змінюються при виконанні наступних елементарних перетворень рядків або стовпчиків матриці (будемо називати ці перетворення допустимими):

– множення рядка (стовпчика) на число не рівне нулю;

– додавання до рядка (стовпчика) іншого рядка (стовпчика);

– перестановка двох довільних рядків (стовпчиків).

►Нехай ранг матриці $\mathbf{A}$ рівний $r$. Покажемо, що при жодному елементарному перетворенні $r$ не зміниться. Зауважимо, що досить довести твердження теореми лише для рядків матриці, оскільки транспонування матриці не змінює значень її мінорів, а отже, не змінить рангу матриці. Розглянемо по черзі кожне з перетворень.

1. Нехай довільний рядок матриці помножили на деяке число $\lambda \neq 0$. Якщо базисний мінор містив даний рядок, то значення базисного мінору помножилось на $\lambda$, якщо ж ні – то значення базисного мінору не змінилося. Жоден мінор матриці, який був рівним нулю, не стане відмінним від нуля, а жоден відмінний від нуля не став нульовим.

2. Нехай в матриці до $i$-го рядка додали $j$-ий рядок. Покажемо, що ранг матриці при цьому не міг збільшитись. Розглянемо довільний мінор $\mathbf{M}=\mathbf{0}$ порядку $r+1$(якщо він існує). Якщо мінор $\mathbf{M}$ не містив рядків з індексами $i$ та $j$, то цей мінор не змінить свого значення. Якщо обидва рядки, що додаються, входили в $\mathbf{M}$, то згідно з властивістю визначників, нове значення мінору $\mathbf{M}$ буде рівним сумі вихідного значення мінору $\mathbf{M}$ та визначника з двома однаковими рядками, а отже дорівнюватиме нулю. Якщо ж $i$-тий рядок входив в $\mathbf{M}$, а $j$-ий – ні, то нове значення $\mathbf{M}$ дорівнюватиме сумі мінору порядку $r+1$ вихідної матриці, який рівний нулю (власне, початкового значення $\mathbf{M}$), та деякого мінору порядку $r+1$, який містить $j$-ий рядок на місці $i$-го. Оскільки це мінор порядку $r+1$ вихідної матриці, то він рівний нулю.

Таким чином, ранг матриці при даному перетворенні не збільшився, бо всі мінори порядку $r+1$, як були, так і лишились нульовими.

Припущення про те, що ранг міг зменшитись є також хибним, оскільки тоді при виконанні зворотних перетворень (відніманні рядків) він мав би збільшуватись, що, як було показано, неможливо.

3. Можна показати, що перестановка двох рядків зводиться до послідовності перетворень двох уже розглянутих типів, а отже, теж не змінює ранг. Дійсно:

$$\mathbf{A}=\begin{pmatrix}\vdots\\ \langle\mathbf{A}^i|\\ \vdots\\ \langle\mathbf{A}^j|\\ \vdots\end{pmatrix}\to\begin{pmatrix}\vdots\\ \langle\mathbf{A}^i|+\langle\mathbf{A}^j|\\ \vdots\\ \langle\mathbf{A}^j|\\ \vdots\end{pmatrix}\to\begin{pmatrix}\vdots\\ -\langle\mathbf{A}^i|-\langle\mathbf{A}^j|\\ \vdots\\ \langle\mathbf{A}^j|\\ \vdots\end{pmatrix}\to\begin{pmatrix}\vdots\\ -\langle\mathbf{A}^i|-\langle\mathbf{A}^j|\\ \vdots\\ -\langle\mathbf{A}^i|\\ \vdots\end{pmatrix}\to\begin{pmatrix}\vdots\\ -\langle\mathbf{A}^i|-\langle\mathbf{A}^j|\\ \vdots\\ \langle\mathbf{A}^i|\\ \vdots\end{pmatrix}\to\begin{pmatrix}\vdots\\ -\langle\mathbf{A}^j|\\ \vdots\\ \langle\mathbf{A}^i|\\ \vdots\end{pmatrix}\to\begin{pmatrix}\vdots\\ \langle\mathbf{A}^j|\\ \vdots\\ \langle\mathbf{A}^i|\\ \vdots\end{pmatrix}$$ ◄

Дана теорема дає метод обчислення рангу довільної матриці – необхідно виконувати допустимі елементарні перетворення над рядками або стовпчиками матриці, намагаючись звести її до якомога простішого вигляду, коли матриця містить достатню кількість нулів. Проілюструємо цей метод на прикладах, а потім сформулюємо остаточний висновок.

<u>Приклад 1</u>**.** Обчислити ранг матриці $\mathbf{A}$: $\mathbf{A}=\begin{pmatrix}1 & 2 & -1 & 4\\ 4 & -1 & -5 & -6\\ -1 & 7 & 2 & 18\\ 2 & -5 & -3 & -14\end{pmatrix}$

Використаємо перший рядок матриці з першим елементом, рівним одиниці, для того, щоб одержати нульові значення під ним так само, як ми це робили розв'язуючи систему методом Гаусса. Перетворення рядків при цьому не змінюватимуть ранг матриці:

$$\mathbf{A} \to \begin{pmatrix} 1 & 2 & -1 & 4 \\ 0 & -9 & -1 & -22 \\ 0 & 9 & 1 & 22 \\ 0 & -9 & -1 & -22 \end{pmatrix} \to \begin{pmatrix} 1 & 2 & -1 & 4 \\ 0 & -9 & -1 & -22 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Нулі в останніх двох рядках одержані внаслідок додавання до них другого рядка безпосередньо і помноженого на (-1). Отже,

$\mathrm{Rg}\,\mathbf{A} = \mathrm{Rg}\begin{pmatrix} 1 & 2 & -1 & 4 \\ 0 & -9 & -1 & -22 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = 2$, оскільки очевидно, що базисним мінором є, наприклад, мінор $\mathbf{M}_{1,2}^{1,2} = \begin{vmatrix} 1 & 2 \\ 0 & -9 \end{vmatrix} = -9$. Будь-який мінор третього порядку обов'язково включатиме нульовий рядок, а отже, буде рівним нулю.

<u>Приклад 2</u>**.** Обчислити ранг матриці $\mathbf{A}$: $\mathbf{A} = \begin{pmatrix} 2 & -4 & 3 & 1 & 0 \\ 1 & -2 & 1 & -4 & 2 \\ 0 & 1 & -1 & 3 & 1 \\ 4 & -7 & 4 & -4 & 5 \end{pmatrix}$

Спочатку переставимо перший і другий рядки і елементарними перетвореннями одержимо нулі в першому стовпчику матриці під елементом, рівним 1:

$$\mathbf{A} \to \begin{pmatrix} 1 & -2 & 1 & -4 & 2 \\ 0 & 0 & 1 & 9 & -4 \\ 0 & 1 & -1 & 3 & 1 \\ 0 & 1 & 0 & 12 & -3 \end{pmatrix}$$

Тепер переставимо другий та третій рядки і одержимо нулі під другим елементом другого рядка:

$$\begin{pmatrix} 1 & -2 & 1 & -4 & 2 \\ 0 & 0 & 1 & 9 & -4 \\ 0 & 1 & -1 & 3 & 1 \\ 0 & 1 & 0 & 12 & -3 \end{pmatrix} \to \begin{pmatrix} 1 & -2 & 1 & -4 & 2 \\ 0 & 1 & -1 & 3 & 1 \\ 0 & 0 & 1 & 9 & -4 \\ 0 & 1 & 0 & 12 & -3 \end{pmatrix} \to \begin{pmatrix} 1 & -2 & 1 & -4 & 2 \\ 0 & 1 & -1 & 3 & 1 \\ 0 & 0 & 1 & 9 & -4 \\ 0 & 0 & 1 & 9 & -4 \end{pmatrix}$$

І нарешті третій рядок помножимо на (-1) та додамо до четвертого:

$$\begin{pmatrix} 1 & -2 & 1 & -4 & 2 \\ 0 & 1 & -1 & 3 & 1 \\ 0 & 0 & 1 & 9 & -4 \\ 0 & 0 & 1 & 9 & -4 \end{pmatrix} \to \begin{pmatrix} 1 & -2 & 1 & -4 & 2 \\ 0 & 1 & -1 & 3 & 1 \\ 0 & 0 & 1 & 9 & -4 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Матриці такого вигляду будемо називати ***трапецієподібними***. Тепер можна сказати, що

$$\mathrm{Rg}\,\mathbf{A} = \mathrm{Rg}\begin{pmatrix} 1 & -2 & 1 & -4 & 2 \\ 0 & 1 & -1 & 3 & 1 \\ 0 & 0 & 1 & 9 & -4 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} = \mathrm{Rg}\begin{pmatrix} 1 & -2 & 1 & -4 & 2 \\ 0 & 1 & -1 & 3 & 1 \\ 0 & 0 & 1 & 9 & -4 \end{pmatrix} = 3.$$

Базисним можна вважати, наприклад, мінор $\mathbf{M}_{1,2,3}^{1,2,3} = 1$.

Ми скористались тут очевидним фактом, що якщо до будь-якої матриці дописати або викреслити з матриці рядок (стовпчик), повністю складений з нулів, – це не змінить рангу матриці.

**Висновок.** Для обчислення рангу матриці необхідно звести її допустимими елементарними перетвореннями до трапецієподібного вигляду.

**Теорема 2. (про базисний мінор).** В довільній матриці кожний рядок (стовпчик) є лінійною комбінацією базисних рядків (стовпчиків).

►З міркувань, наведених вище, досить доводити цю теорему лише для рядків довільної матриці.

Припустимо, що ранг матриці $\mathbf{A} = \left(a_j^i\right)_{j=\overline{1,n}}^{i=\overline{1,m}}$ рівний $r$, та розглянемо довільний рядок матриці, наприклад, $\left\langle \mathbf{A}^i \right|$, де $i = \overline{1,m}$ – фіксований індекс. Щодо цього рядка можна зробити два припущення:

**1.** $\left\langle \mathbf{A}^i \right|$ входить в базисний мінор, тоді він є лінійною комбінацією базисних рядків, бо $\left\langle \mathbf{A}^i \right| = 1 \cdot \left\langle \mathbf{A}^i \right| + \sum\limits_{k \neq i} 0 \cdot \left\langle \mathbf{A}^k \right|$.

**2.** $\left\langle \mathbf{A}^i \right|$ не входить в базисний мінор, при цьому зрозуміло, слід вважати, що $r < \min(m,n)$. Припустимо також, що базисний мінор $\mathbf{M}$ є верхнім лівим блоком матриці (інакше допустимими перетвореннями переставимо стовпчики та рядки матриці так, щоб це припущення було виконаним), тобто $\mathbf{M} = \begin{vmatrix} a_1^1 & \cdots & a_r^1 \\ \vdots & & \vdots \\ a_1^r & \cdots & a_r^r \end{vmatrix} \neq 0$. Розглянемо довільний мінор $\mathbf{M}_{r+1}$ порядку $r+1$, який містить базисний мінор і включає вибраний рядок $\left\langle \mathbf{A}^i \right|$ та довільний стовпчик матриці з індексом $j$:

$$\mathbf{M}_{r+1} = \begin{vmatrix} a_1^1 & \cdots & a_r^1 & a_j^1 \\ \vdots & & \vdots & \vdots \\ a_1^r & \cdots & a_r^r & a_j^r \\ a_1^i & \cdots & a_r^i & a_j^i \end{vmatrix} = 0, \text{ оскільки } \operatorname{Rg} \mathbf{A} = r.$$

З іншого боку мінор $\mathbf{M}_{r+1}$ може бути обчислений через мінор $\mathbf{M}$, якщо його розкласти по останньому стовпчику: $\mathbf{M}_{r+1} = \sum\limits_{k=1}^{r} a_j^k \cdot \mathbf{A}_k^j(i) + a_j^i \cdot \mathbf{M}$ (тут $\mathbf{A}_k^j(i)$ – це алгебраїчне доповнення до елемента $a_j^k$, яке містить частину $i$-го рядка). Тоді з рівності нулю мінору $\mathbf{M}_{r+1}$ випливає: $a_j^i = \sum\limits_{k=1}^{r} a_j^k \cdot \left( -\frac{\mathbf{A}_k^j(i)}{\mathbf{M}} \right)$, або у векторній формі: $\left\langle \mathbf{A}^i \right| = \sum\limits_{k=1}^{r} \left\langle \mathbf{A}^k \right| \cdot \alpha_k(i)$. Тут $\alpha_k(i) = -\frac{\mathbf{A}_k^j(i)}{\mathbf{M}}$ насправді не залежить від індексу $j$, оскільки алгебраїчні доповнення до елементів стовпчика $\left| \mathbf{A}_j \right\rangle$ не включають елементів самого стовпчика. ◄

**Наслідок 1.** Ранг матриці, визначений як порядок базисного мінору матриці, може бути еквівалентним чином визначений як максимальна кількість лінійно незалежних рядків (стовпчиків) матриці.

**Наслідок 2. (критерії виродженості квадратної матриці).** Наступні твердження є еквівалентними:

1. Визначник матриці рівний нулю.
2. Ранг $r$ квадратної матриці менший за її порядок $n$.
3. Рядки або стовпчика квадратної матриці лінійно залежні.

**Наслідок 3. (критерії невиродженості матриці).** Наступні твердження є еквівалентними:

1. Визначник матриці не дорівнює нулю.

1. Ранг $r$ квадратної матриці рівний її порядку $n$.

3. Рядки та стовпчика квадратної матриці лінійно незалежні.

**Теорема 3. (про ранг добутку матриць).** Ранг добутку матриць не перевищує рангу кожного з множників: $\operatorname{Rg}(\mathbf{A}\cdot\mathbf{B})\le\min(\operatorname{Rg}\mathbf{A},\operatorname{Rg}\mathbf{B})$.

Дійсно, очевидно, що $\operatorname{Rg}(\mathbf{A}\cdot\mathbf{B})\le\operatorname{Rg}(\mathbf{A}|\mathbf{A}\cdot\mathbf{B})$, де $(\mathbf{A}|\mathbf{A}\cdot\mathbf{B})$ – розширена матриця, складена зі стовпчиків матриць $\mathbf{A}$ та $\mathbf{A}\cdot\mathbf{B}$ (кількість рядків у них однакова).

Нехай розмірності матриць $\mathbf{A}$ та $\mathbf{B}$ рівні відповідно $m\times n$ та $n\times p$. Якщо позначити $\mathbf{A}\cdot\mathbf{B}=\mathbf{C}$, то з правила множення матриць випливає, що стовпчики матриці $\mathbf{A}\cdot\mathbf{B}$ є лінійними комбінаціями стовпчиків матриці $\mathbf{A}$: $\left|\mathbf{C}_j\right\rangle=\sum\limits_{k=1}^{n}\left|\mathbf{A}_k\right\rangle\cdot b_j^k,\ j=\overline{1,p}$, тому $\operatorname{Rg}(\mathbf{A}|\mathbf{A}\cdot\mathbf{B})=\operatorname{Rg}(\mathbf{A}|\mathbf{C})=\operatorname{Rg}\mathbf{A}$. Отже, $\operatorname{Rg}(\mathbf{A}\cdot\mathbf{B})\le\operatorname{Rg}\mathbf{A}$.

З іншого боку, $\left\langle\mathbf{C}^i\right|=\sum\limits_{k=1}^{n}a_k^i\cdot\left\langle\mathbf{B}^k\right|,i=\overline{1,m}$, тобто рядки матриці $\mathbf{A}\cdot\mathbf{B}$ є лінійними комбінаціями рядків матриці $\mathbf{B}$, тому аналогічно можна показати, що $\operatorname{Rg}(\mathbf{A}\cdot\mathbf{B})\le\operatorname{Rg}\mathbf{B}$, тобто $\operatorname{Rg}(\mathbf{A}\cdot\mathbf{B})\le\min(\operatorname{Rg}\mathbf{A},\operatorname{Rg}\mathbf{B})$, що й треба було довести.

## 2. Загальна теорія систем $n$ лінійних рівнянь з $m$ невідомими.

### 2.1. Критерій сумісності системи.

Повернемось до розгляду систем $m$ лінійних рівнянь з $n$ невідомими загального виду:

$$\mathbf{A}\cdot\left|\mathbf{x}\right\rangle=\left|\mathbf{b}\right\rangle,\ \text{або}\qquad \{a_j^i x^j=b^i,\ \ i=\overline{1,m};\ j=\overline{1,n}\qquad (1)$$

В цьому розділі дослідимо умови сумісності даної системи та правила знаходження розв'язків.

**Теорема 4. (Кронекера-Капеллі – критерій сумісності системи лінійних рівнянь)**
Система (1) сумісна тоді і тільки тоді, коли ранг розширеної матриці системи рівний рангу матриці системи, тобто $\operatorname{Rg}(\mathbf{A}|\,\mathbf{b})=\operatorname{Rg}\mathbf{A}$.

► Необхідність. Нехай система (1) сумісна. Тоді існує принаймні один її розв'язок $\left|\boldsymbol{\alpha}\right\rangle^T=\begin{pmatrix}\alpha^1 & \alpha^2 & \cdots & \alpha^n\end{pmatrix}$: $\mathbf{A}\cdot\boldsymbol{\alpha}=\mathbf{b}$, або $\alpha^1\cdot\left|\mathbf{A}_1\right\rangle+\alpha^2\cdot\left|\mathbf{A}_2\right\rangle+\cdots+\alpha^n\cdot\left|\mathbf{A}_n\right\rangle=\left|\mathbf{b}\right\rangle$. Остання рівність означає, що вектор-стовпчик $\left|\mathbf{b}\right\rangle$ є лінійною комбінацією стовпчиків матриці $\mathbf{A}$, тобто включення цього стовпчика в матрицю системи не змінює її ранг, що власне, і означає рівність $\operatorname{Rg}(\mathbf{A}|\,\mathbf{b})=\operatorname{Rg}\mathbf{A}$.

Достатність. Нехай $\operatorname{Rg}(\mathbf{A}|\,\mathbf{b})=\operatorname{Rg}\mathbf{A}$. Тоді базисний мінор матриці $\mathbf{A}|\,\mathbf{b}$ розташований в стовпчиках матриці $\mathbf{A}$, отже, стовпчик $\left|\mathbf{b}\right\rangle$ є лінійною комбінацією з деякими коефіцієнтами $\alpha^1,\alpha^2,\ldots,\alpha^n$ стовпчиків матриці $\mathbf{A}$: $\alpha^1\cdot\left|\mathbf{A}_1\right\rangle+\alpha^2\cdot\left|\mathbf{A}_2\right\rangle+\cdots+\alpha^n\cdot\left|\mathbf{A}_n\right\rangle=\left|\mathbf{b}\right\rangle$. Дана рівність означає, що $\left|\boldsymbol{\alpha}\right\rangle^T=\begin{pmatrix}\alpha^1 & \alpha^2 & \cdots & \alpha^n\end{pmatrix}$ є розв'язком системи (1). ◄

**Означення 5.** Загальним розв'язком системи (1) називається ***множина всіх її розв'язків***.

### 2.2. Однорідна система лінійних рівнянь.

**Означення 6.** Якщо в системі (1) вектор-стовпчик $\left|\mathbf{b}\right\rangle$ нульовий:

$$\mathbf{A}\cdot\left|\mathbf{x}\right\rangle=\left|\mathbf{0}\right\rangle,\qquad (\text{надалі для простоти, просто } \mathbf{A}\cdot\mathbf{x}=\mathbf{0})\qquad (2)$$

то така система називається ***однорідною***.

**Зауваження.** Однорідна система (2), очевидно, завжди сумісна. Вона має принаймні один розв'язок $\mathbf{x}=\mathbf{0}$, який називається ***тривіальним***. Цікавість являють нетривіальні розв'язки цієї системи, якщо вони існують.

**<u>Наслідок з теореми Крамера</u>.** Однорідна система $n$ лінійних рівнянь з $n$ невідомими має єдиний (тривіальний) розв'язок, тоді і тільки тоді, коли $|\mathbf{A}| \neq 0$. Отже, для існування нетривіальних розв'язків необхідно і достатньо, щоб $|\mathbf{A}| = 0$, або $\mathrm{Rg}\,\mathbf{A} = r < n$.

Позначимо через $H$ множину всіх розв'язків однорідної системи (2) та дослідимо властивості цих розв'язків.

**<u>Теорема 5</u>.** Якщо $\mathbf{x}_1$ та $\mathbf{x}_2$ – розв'язки системи (2) , то і $\mathbf{x}_1 + \mathbf{x}_2$ є роз в'язком, або $\forall \mathbf{x}_1, \mathbf{x}_2 \in H \Rightarrow \mathbf{x}_1 + \mathbf{x}_2 \in H$.

► Дійсно, з дистрибутивності множення матриць випливає, що $\mathbf{A}\cdot(\mathbf{x}_1 + \mathbf{x}_2) = \mathbf{A}\cdot\mathbf{x}_1 + \mathbf{A}\cdot\mathbf{x}_2 = \mathbf{0}$. ◄

**<u>Теорема 6</u>.** Якщо $\mathbf{x}_1$ – розв'язок системи (2) , то і $\lambda\mathbf{x}_1$ є розв'язком $\forall\lambda\in\mathbf{R}$, або $\forall\lambda\in\mathbf{R}\ \forall\mathbf{x}_1 \in H \Rightarrow \lambda\mathbf{x}_1 \in H$.

► Справді, $\mathbf{A}\cdot(\lambda\mathbf{x}) = \lambda\cdot\mathbf{A}\cdot\mathbf{x} = \mathbf{0}$. ◄

**<u>Наслідок</u>.** Оскільки $\mathbf{0}\in H$, то множина $H$ розв'язків системи (2) – лінійний простір, який є частиною простору $L_n$, тобто є його ***підпростором***.

З'ясуємо, яка розмірність цього підпростору. Позначимо ранг матриці $\mathbf{A}$ через $r$: $\mathrm{Rg}\,\mathbf{A} = r \le \min(m, n)$.

Не втрачаючи загальності, можна вважати, що базисний мінор матриці $\mathbf{A}$ розташований в перших $r$ стовпчиках та $r$ рядках (інакше переставимо відповідним чином рівняння та позначимо змінні в потрібному порядку).

Базисний мінор позначимо $\mathbf{M}_r = \begin{vmatrix} a_1^1 & a_2^1 & \cdots & a_r^1 \\ a_1^2 & a_2^2 & \cdots & a_r^2 \\ \vdots & \vdots & \cdots & \vdots \\ a_1^r & a_2^r & \cdots & a_r^r \end{vmatrix}$. Саму матрицю системи будемо розглядати як блочну: $\mathbf{A} = \begin{pmatrix} \mathbf{A}^* & \mathbf{A}^{**} \\ \mathbf{B} & \mathbf{C} \end{pmatrix}$, де $|\mathbf{A}^*| = \mathbf{M}_r$, $\mathbf{A}^{**} = \begin{pmatrix} a_{r+1}^1 & \cdots & a_n^1 \\ \vdots & & \vdots \\ a_{r+1}^r & \cdots & a_n^r \end{pmatrix}$, $\mathbf{B} = \begin{pmatrix} a_1^{r+1} & \cdots & a_r^{r+1} \\ \vdots & & \vdots \\ a_1^m & \cdots & a_r^m \end{pmatrix}$, $\mathbf{C} = \begin{pmatrix} a_{r+1}^{r+1} & \cdots & a_n^{r+1} \\ \vdots & & \vdots \\ a_{r+1}^m & \cdots & a_n^m \end{pmatrix}$. Вектор-стовпчик невідомих теж складається з блоків: $|\mathbf{x}\rangle = \begin{pmatrix} x^1 \\ x^2 \\ \vdots \\ x^n \end{pmatrix} = \begin{pmatrix} \mathbf{x}^* \\ \mathbf{x}^{**} \end{pmatrix}$, де $|\mathbf{x}^*\rangle = \begin{pmatrix} x^1 \\ \vdots \\ x^r \end{pmatrix}$ – вектор-стовпчик базисних (залежних) змінних, а $|\mathbf{x}^{**}\rangle = \begin{pmatrix} x^{r+1} \\ \vdots \\ x^n \end{pmatrix}$ – вільних (параметричних) змінних.

Такі позначення дозволяють замість системи (2) розглядати еквівалентну їй систему:

$$\mathbf{A}\cdot\mathbf{x} = \mathbf{0} \Leftrightarrow \begin{pmatrix} \mathbf{A}^* & \mathbf{A}^{**} \end{pmatrix}\cdot\begin{pmatrix} \mathbf{x}^* \\ \mathbf{x}^{**} \end{pmatrix} = \mathbf{0} \qquad (3)$$

(з $m$ рівнянь системи (2) лише $r$ будуть лінійно незалежними, решта $m-r$ рівнянь, коефіцієнти яких записані в матрицях $\mathbf{B}$ і $\mathbf{C}$, – лінійно виражаються через них). Останню матричну рівність можна записати у вигляді: $\mathbf{A}^*\cdot\mathbf{x}^* + \mathbf{A}^{**}\cdot\mathbf{x}^{**} = \mathbf{0}$, або $\mathbf{A}^*\cdot\mathbf{x}^* = -\mathbf{A}^{**}\cdot\mathbf{x}^{**}$, звідки $\left(\mathbf{A}^*\right)^{-1}\cdot\mathbf{A}^*\cdot\mathbf{x}^* = -\left(\mathbf{A}^*\right)^{-1}\cdot\mathbf{A}^{**}\cdot\mathbf{x}^{**}$, тобто

$$\mathbf{x}^* = -\left(\mathbf{A}^*\right)^{-1}\cdot\mathbf{A}^{**}\cdot\mathbf{x}^{**} \qquad (4)$$

Таким чином, $r$ базисних змінних лінійно виражаються через $n-r$ незалежних змінних, яким можна надавати довільних значень. Враховуючи, що $\mathbf{x}^{**} = \mathbf{E}_{n-r} \cdot \mathbf{x}^{**}$, можна записати вираз для розв'язку системи (2) через довільні значення незалежних змінних:

$$\mathbf{x} = \begin{pmatrix} \mathbf{x}^{*} \\ \mathbf{x}^{**} \end{pmatrix} = \begin{pmatrix} -\left(\mathbf{A}^{*}\right)^{-1} \cdot \mathbf{A}^{**} \\ \mathbf{E}_{n-r} \end{pmatrix} \cdot \mathbf{x}^{**}, \qquad \text{або} \qquad \mathbf{x} = \mathbf{H} \cdot \mathbf{x}^{**}, \text{ де} \tag{5}$$

$$\mathbf{H} = \begin{pmatrix} -\left(\mathbf{A}^{*}\right)^{-1} \cdot \mathbf{A}^{**} \\ \mathbf{E}_{n-r} \end{pmatrix} = \begin{pmatrix} h_1^1 & h_2^1 & \cdots & h_{n-r}^1 \\ \vdots & \vdots & & \vdots \\ h_1^r & h_2^r & \cdots & h_{n-r}^r \\ 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

Розглянемо уважніше матрицю $\mathbf{H}$. Кожен з її $n-r$ вектор-стовпчиків $|\mathbf{H}_1\rangle, |\mathbf{H}_2\rangle, \ldots, |\mathbf{H}_{n-r}\rangle$ є розв'язком системи (2). Дійсно, поклавши вектор-стовпчики вільних змінних рівними стовпчикам матриці $\mathbf{E}_{n-r}$, одержимо, згідно з (4), значення базисних змінних, а згідно з (5) – розв'язок системи (2). Крім того, $\operatorname{Rg}\mathbf{H} = n-r$, оскільки $\mathbf{E}_{n-r}$ – базисний мінор матриці $\mathbf{H}$. Отже, матриця $\mathbf{H}$ утворена $n-r$ лінійно незалежними розв'язками системи (2). Щоб стверджувати, що $\dim(H) = n-r$, залишився лише один крок – достатньо показати, що будь-який розв'язок системи (2), є лінійною комбінацією вектор-стовпчиків $|\mathbf{H}_1\rangle, |\mathbf{H}_2\rangle, \ldots, |\mathbf{H}_{n-r}\rangle$, тобто вони утворюють базис простору $H$.

Нехай $|\boldsymbol{\alpha}\rangle$ – довільний розв'язок системи (2). Тоді $|\boldsymbol{\alpha}\rangle$ є розв'язком еквівалентної їй системи (3), а отже, виражається співвідношенням (5), тобто є лінійною комбінацією вектор-стовпчиків $|\mathbf{H}_1\rangle, |\mathbf{H}_2\rangle, \ldots, |\mathbf{H}_{n-r}\rangle$ з деякими коефіцієнтами $\left|\mathbf{x}^{**}\right\rangle^{T} = \left((x^{**})^1, \ldots, (x^{**})^{n-r}\right)$. Таким чином, доведено наступне твердження.

**<u>Теорема 7</u>.** Розмірність простору $H$ розв'язків однорідної системи (2) рівна $n-r$, де $n$ – кількість невідомих, $r$ – ранг матриці системи.

**<u>Означення 7</u>.** Довільний базис простору $H$ розв'язків однорідної системи (2) називається ***фундаментальною системою розв'язків***.

**<u>Означення 8</u>.** Фундаментальна система розв'язків однорідної системи (2), в якій вектор-стовпчики вільних змінних рівні стовпчикам одиничної матриці, називається ***нормальною фундаментальною системою розв'язків***.

З теореми **7** випливає наслідок, який визначає структуру загального розв'язку однорідної системи (2).

**<u>Наслідок</u>.** Загальний розв'язок однорідної системи (2) $|\mathbf{x}\rangle_{заг.одн.}$ є лінійною комбінацією векторів фундаментальної системи розв'язків:

$$|\mathbf{x}\rangle_{заг.одн.} = \sum_{k=1}^{n-r} c^k \cdot |\mathbf{H}_k\rangle. \tag{6}$$

Тут $c^k, k = \overline{1, n-r}$ – довільні сталі.

<u>Приклад 3</u>**.** Знайти загальний розв'язок системи:

$$\begin{cases} 3x^1 + x^2 - 8x^3 + 2x^4 + x^5 = 0 \\ 2x^1 - 2x^2 - 3x^3 - 7x^4 + 2x^5 = 0 \\ x^1 + 11x^2 - 12x^3 + 34x^4 - 5x^5 = 0 \\ x^1 - 5x^2 + 2x^3 - 16x^4 + 3x^5 = 0 \end{cases}$$

Запишемо матрицю системи та використаємо допустимі перетворення, щоб визначити її базисний мінор (першим рядком відразу запишемо третій, бо в ньому перший коефіцієнт рівний 1):

$$\mathbf{A}=\begin{pmatrix}1 & 11 & -12 & 34 & -5\\ 2 & -2 & -3 & -7 & 2\\ 3 & 1 & -8 & 2 & 1\\ 1 & -5 & 2 & -16 & 3\end{pmatrix}\to\begin{pmatrix}1 & 11 & -12 & 34 & -5\\ 0 & -24 & 21 & -75 & 12\\ 0 & -32 & 28 & -100 & 16\\ 0 & -16 & -14 & -50 & 8\end{pmatrix}\to$$

$$\to\begin{pmatrix}1 & 11 & -12 & 34 & -5\\ 0 & 8 & -7 & 25 & -4\\ 0 & 8 & -7 & 25 & -4\\ 0 & 8 & -7 & 25 & -4\end{pmatrix}\to\begin{pmatrix}1 & 11 & -12 & 34 & -5\\ 0 & 8 & -7 & 25 & -4\\ 0 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 0\end{pmatrix}$$

Таким чином $r=\mathrm{Rg}\,\mathbf{A}=2$, $\dim(H)=n-r=5-2=3$. За базисні змінні вибираємо $x^1, x^2$, тоді $x^3, x^4, x^5$ – незалежні змінні. Продовжимо допустимі перетворення над матрицею системи таким чином, щоб базисний мінор $\mathbf{M}_{1,2}^{1,2}$ (через $\mathbf{M}_{i_1,\ldots,i_k}^{j_1,\ldots,j_l}$ позначений мінор, утворений елементами матриці, що належать одночасно рядкам з індексами $i_1,\ldots,i_k$ та стовпчикам з індексами $j_1,\ldots,j_l$) зробити одиничним:

$$\begin{pmatrix}1 & 11 & -12 & 34 & -5\\ 0 & 8 & -7 & 25 & -4\\ 0 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 0\end{pmatrix}\to\begin{pmatrix}1 & 11 & -12 & 34 & -5\\ 0 & 1 & -\frac{7}{8} & \frac{25}{8} & -\frac{1}{2}\\ 0 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 0\end{pmatrix}\to\begin{pmatrix}1 & 0 & -\frac{19}{8} & -\frac{3}{8} & \frac{1}{2}\\ 0 & 1 & -\frac{7}{8} & \frac{25}{8} & -\frac{1}{2}\\ 0 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 0\end{pmatrix}$$

Тепер базисні змінні легко виражаються через вільні: $\begin{cases}x^1=\frac{19}{8}x^3+\frac{3}{8}x^4-\frac{1}{2}x^5\\ x^2=\frac{7}{8}x^3-\frac{25}{8}x^4+\frac{1}{2}x^5\end{cases}$

(це формула (4)). Формула ж (5) в даному прикладі набуває вигляду:

$$\begin{cases}x^1=\frac{19}{8}x^3+\frac{3}{8}x^4-\frac{1}{2}x^5\\ x^2=\frac{7}{8}x^3-\frac{25}{8}x^4+\frac{1}{2}x^5\\ x^3=x^3\\ x^4=x^4\\ x^5=x^5\end{cases},$$

або з допомогою матрицю $\mathbf{H}$: $$|\mathbf{x}\rangle=\begin{pmatrix}\frac{19}{8} & \frac{3}{8} & -\frac{1}{2}\\ \frac{7}{8} & -\frac{25}{8} & \frac{1}{2}\\ 1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}\cdot\begin{pmatrix}x^3\\ x^4\\ x^5\end{pmatrix}$$

Вектор-стовпчики цієї матриці утворюють нормальну фундаментальну систему розв'язків даної системи: $|\mathbf{H}_1\rangle^T=\left(\frac{19}{8};\frac{7}{8};1;0;0\right)$, $|\mathbf{H}_2\rangle^T=\left(\frac{3}{8};-\frac{25}{8};0;1;0\right)$, $|\mathbf{H}_3\rangle^T=\left(-\frac{1}{1};\frac{1}{2};0;0;1\right)$ (ці вектори можна було одержати і безпосередньо із формули (4) для даної системи, надавши відповідні значення вільним змінним). Загальний розв'язок даної системи виражається формулою (6), яка в координатній формі набуває вигляду:

$$\begin{cases} x^1 = \frac{19}{8}c^1 + \frac{3}{8}c^2 - \frac{1}{2}c3 \\ x^2 = \frac{7}{8}c^1 - \frac{25}{8}c^2 + \frac{1}{2}c^3 \\ x^3 = c^1 \\ x^4 = c^2 \\ x^5 = c^3 \end{cases}$$, де $c^1, c^2, c^3$ – довільні сталі.

### 2.3. Неоднорідна система лінійних рівнянь.

Повернемось до розгляду неоднорідної системи (1): $\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$.

**<u>Означення 9</u>.** Однорідна система $\mathbf{A} \cdot \mathbf{x} = \mathbf{0}$ з тою самою матрицею системи називається *зведеною* системою для системи (1).

Покажемо, що загальний розв'язок зведеної однорідної системи (2) істотно пов'язаний з розв'язками неоднорідної системи (1). Припущення про існування розв'язків системи (1) згідно з теоремою Кронекера-Капеллі означає рівність рангів: $\mathrm{Rg}(\mathbf{A}|\,\mathbf{b}) = \mathrm{Rg}\,\mathbf{A}$. Вважатимемо цю умову виконаною.

**<u>Теорема 8</u>.** Нехай $|\boldsymbol{\alpha}_1\rangle$ та $|\boldsymbol{\alpha}_2\rangle$ – розв'язки системи (1). Тоді вектор-стовпчики $|\boldsymbol{\alpha}_1\rangle - |\boldsymbol{\alpha}_2\rangle$ та $|\boldsymbol{\alpha}_2\rangle - |\boldsymbol{\alpha}_1\rangle$ будуть розв'язками зведеної системи (2).

► Підставимо в ліву частину системи (1) наприклад, вектор-стовпчик $|\boldsymbol{\alpha}_1\rangle - |\boldsymbol{\alpha}_2\rangle$ та використаємо дистрибутивність множення матриць:

$$\mathbf{A} \cdot \big(|\boldsymbol{\alpha}_1\rangle - |\boldsymbol{\alpha}_2\rangle\big) = \mathbf{A} \cdot |\boldsymbol{\alpha}_1\rangle - \mathbf{A} \cdot |\boldsymbol{\alpha}_2\rangle = |\mathbf{b}\rangle - |\mathbf{b}\rangle = \mathbf{0}, \text{ отже, } \quad \mathbf{A} \cdot \big(|\boldsymbol{\alpha}_1\rangle - |\boldsymbol{\alpha}_2\rangle\big) = \mathbf{0},$$

що і треба було показати. ◄

**<u>Теорема 9</u>.** Нехай $|\boldsymbol{\alpha}_1\rangle$ та $|\boldsymbol{\alpha}_2\rangle$ – розв'язки системи (1) з стовпчиками вільних членів, рівними відповідно $|\mathbf{b}_1\rangle$ та $|\mathbf{b}_2\rangle$. Тоді вектор-стовпчик $|\boldsymbol{\alpha}_1\rangle + |\boldsymbol{\alpha}_2\rangle$ буде розв'язком системи (1) з правою частиною $|\mathbf{b}_1\rangle + |\mathbf{b}_2\rangle$.

Доведення повністю аналогічне попередньому.

**<u>Теорема 10</u>.** Нехай $|\boldsymbol{\alpha}\rangle$ – деякий розв'язок системи (1). Тоді довільний вектор $|\boldsymbol{\beta}\rangle$ буде розв'язком цієї ж системи тоді і тільки тоді, коли знайдеться такий розв'язок $|\boldsymbol{\alpha}_0\rangle$ зведеної системи (2), що $|\boldsymbol{\beta}\rangle = |\boldsymbol{\alpha}\rangle + |\boldsymbol{\alpha}_0\rangle$.

► <u>Необхідність</u>. Нехай $|\boldsymbol{\beta}\rangle$ є розв'язком системи (1). Позначимо $|\boldsymbol{\beta}\rangle - |\boldsymbol{\alpha}\rangle = |\boldsymbol{\alpha}_0\rangle$. Тоді з теореми **8** випливає, що $|\boldsymbol{\alpha}_0\rangle$ є розв'язком зведеної системи (2).

<u>Достатність</u>. Нехай $|\boldsymbol{\alpha}\rangle$ є розв'язком системи (1), $|\boldsymbol{\alpha}_0\rangle$ – розв'язок зведеної системи (2). Тоді з теореми **9** маємо, що вектор-стовпчик $|\boldsymbol{\beta}\rangle = |\boldsymbol{\alpha}\rangle + |\boldsymbol{\alpha}_0\rangle$ також буде розв'язком системи (1) з тою самою правою частиною. ◄

**<u>Наслідок</u>.** Нехай $|\mathbf{H}_1\rangle, |\mathbf{H}_2\rangle, \ldots, |\mathbf{H}_{n-r}\rangle$ – фундаментальна система розв'язків однорідної системи (2), а $|\mathbf{x}\rangle_{\text{част. неодн.}}$ – деякий розв'язок системи (1) (його називають частинним розв'язком). Тоді загальний розв'язок неоднорідної системи (1) визначається наступною формулою:

$$|\mathbf{x}\rangle_{\text{заг. неодн.}} = |\mathbf{x}\rangle_{\text{част. неодн.}} + \sum_{k=1}^{n-r} c^k \cdot |\mathbf{H}_k\rangle. \qquad (7)$$

Тут $c^k, k = \overline{1, n-r}$ – довільні сталі.

Приклад 4. Знайти загальний розв'язок системи: $\begin{cases} 4x^1 + x^2 - 2x^3 + x^4 = 3 \\ x^1 - 2x^2 - x^3 + 2x^4 = 2 \\ 2x^1 + 5x^2 - x^4 = -1 \\ 3x^1 + 3x^2 - x^3 - 3x^4 = 1 \end{cases}$

Будемо виконувати допустимі перетворення над рядками розширеної матриці системи, зводячи її до спрощеного вигляду:

$$(\mathbf{A}|\,\mathbf{b}) = \left(\begin{array}{cccc|c} 4 & 1 & -2 & 1 & 3 \\ 1 & -2 & -1 & 2 & 2 \\ 2 & 5 & 0 & -1 & -1 \\ 3 & 3 & -1 & -3 & 1 \end{array}\right) \to \left(\begin{array}{cccc|c} 1 & -2 & -1 & 2 & 2 \\ 4 & 1 & -2 & 1 & 3 \\ 2 & 5 & 0 & -1 & -1 \\ 3 & 3 & -1 & -3 & 1 \end{array}\right) \to$$

$$\to \left(\begin{array}{cccc|c} 1 & -2 & -1 & 2 & 2 \\ 0 & 9 & 2 & -7 & -5 \\ 0 & 9 & 2 & -5 & -5 \\ 0 & 9 & 2 & -9 & -5 \end{array}\right) \to \left(\begin{array}{cccc|c} 1 & -2 & -1 & 2 & 2 \\ 0 & 9 & 2 & -6 & -5 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -3 & 0 \end{array}\right) \to \left(\begin{array}{cccc|c} 1 & -2 & -1 & 2 & 2 \\ 0 & 9 & 2 & -6 & -5 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

Отже, $\operatorname{Rg}(\mathbf{A}|\,\mathbf{b}) = \operatorname{Rg}\mathbf{A} = 3$. За базисні змінні виберемо $x^1, x^3, x^4$, тоді $x^2$ – незалежна змінна. Перетворимо тепер базисний мінор $\mathbf{M}^{1,3,4}_{1,2,3}$ допустимими перетвореннями рядків на одиничну матрицю:

$$\left(\begin{array}{cccc|c} 1 & -2 & -1 & 2 & 2 \\ 0 & 9 & 2 & -6 & -5 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right) \to \left(\begin{array}{cccc|c} 1 & -2 & -1 & 0 & 2 \\ 0 & 9 & 2 & 0 & -5 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right) \to \left(\begin{array}{cccc|c} 1 & -2 & -1 & 0 & 2 \\ 0 & \frac{9}{2} & 1 & 0 & -\frac{5}{2} \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right) \to$$

$$\to \left(\begin{array}{cccc|c} 1 & \frac{5}{2} & 0 & 0 & -\frac{1}{2} \\ 0 & \frac{9}{2} & 1 & 0 & -\frac{5}{2} \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

Запишемо рівняння, що відповідають спрощеній системі, та знайдемо залежність базисних змінних від незалежної змінної:

$$\begin{cases} x^1 + \frac{5}{2}x^2 = -\frac{1}{2} \\ \frac{9}{2}x^2 + x^3 = -\frac{5}{2} \\ x^4 = 0 \end{cases} \Rightarrow \begin{cases} x^1 = -\frac{1}{2} - \frac{5}{2}x^2 \\ x^3 = -\frac{5}{2} - \frac{9}{2}x^2 \\ x^4 = 0 \end{cases}$$

Частинний розв'язок неоднорідної системи (1) визначимо, поклавши, наприклад, $x^2 = 0$. Тоді $|\mathbf{x}\rangle^T_{част.неодн.} = \left(-\frac{1}{2}; 0; -\frac{5}{2}; 0\right)$. Оскільки $n = 4,\ r = 3$, фундаментальна система розв'язків зведеної однорідної системи (2) складається з одного вектора, наприклад, при $x^2 = 2$: $|\mathbf{H}_1\rangle^T = (-5; 2; -9; 0)$.

Отже, $|\mathbf{x}\rangle_{заг.неодн.} = |\mathbf{x}\rangle_{част.неодн.} + c\cdot|\mathbf{H}_1\rangle$, тобто

$|\mathbf{x}\rangle_{заг.неодн.} = \begin{pmatrix} x^1 & x^2 & x^3 & x^4 \end{pmatrix}^T = \begin{pmatrix} -\frac{1}{2} & 0 & -\frac{5}{2} & 0 \end{pmatrix}^T + c\cdot\begin{pmatrix} -5 & 2 & -9 & 0 \end{pmatrix}^T$, де $c$ – довільна стала.

Приклад 5. Знайти загальний розв'язок системи: $\begin{cases} x+y-2z=1 \\ -2x-2y+4z=-2 \end{cases}$

Неважко зрозуміти, що ранг цієї системи рівний 1, отже, простір розв'язків цієї системи має розмірність 2, і власне, описується рівнянням площини $x+y-2z=1$. Використаємо очевидність цього прикладу, щоб зробити більш наочною структуру загального розв'язку. Маємо: $\left| x, y, z \right\rangle^T_{част.неодн.} = (2; 1;\ 1)$, фундаментальна система розв'язків складається з векторів $\left|\mathbf{H}_1\right\rangle^T = (1;\ \text{-}1; 0)$ та $\left|\mathbf{H}_2\right\rangle^T = (2;\ 0; 1)$, отже, $\begin{pmatrix} x \\ y \\ z \end{pmatrix}_{заг.неодн.} = \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} + c_1 \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} + c_2 \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}$ . Зверtіть увагу, вектори $\left|\mathbf{H}_1\right\rangle$ та $\left|\mathbf{H}_2\right\rangle$ є базисом двовимірного підпростору – площини $x+y-2z=0$, паралельної площині $x+y-2z=1$, яка визначена системою. Вектор же $(2; 1;\ 1)$ забезпечує зсув цієї площини відносно першої. Таким чином, розв'язки неоднорідної системи не утворюють лінійного простору (їх множина – лінійний многовид), проте можуть бути одержані з лінійного простору, породженого зведеною однорідною системою, зсувом на деякий вектор, що належить множині розв'язків неоднорідної системи.

## Лекції 12-13. Векторні (лінійні) простори та підпростори. Евклідів простір

## Ізоморфізм просторів.

В першій лекції було дане означення векторного простору, а також введене поняття базису векторного простору. Нижче будуть досліджуватись властивості таких базисів, а також зв'язки між різними базисами векторного простору.

### 1. Зв'язок між різними базисами векторного простору.

Нехай $L$ – векторний простір розмірності $n$. Розглянемо два довільних базиси цього простору: $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ та $\{\mathbf{e}'_1, \mathbf{e}'_2, \ldots, \mathbf{e}'_n\}$. Нас цікавитиме зв'язок цих базисів та координат одного й того самого вектора у різних базисах. Зрозуміло, що вектори другого базису (називатимемо його *новим* базисом на відміну від першого, який вважатимемо *старим*) виражаються через базисні вектори першого. Позначимо це таким чином:

$$\mathbf{e}'_j = \sum_{i=1}^{n} t^i_j \mathbf{e}_i,\ j = \overline{1, n}. \tag{1}$$

Якщо записати коефіцієнти розкладу (1) у вигляді матриці $\mathbf{T} = \left(t^i_j\right),\ i, j = \overline{1, n}$, то очевидно, що її ***стовпчики є координатами базисних векторів нового базису*** у старому базисі, а рівність (1) також можна записати таким чином:

$$\mathbf{e}'_j = \left\langle \mathbf{e}_1 \ldots \mathbf{e}_n \right| \cdot \left| \mathbf{T}_j \right\rangle,\ j = \overline{1, n}, \quad \text{або ж} \quad \left\langle \mathbf{e}'_1 \ldots \mathbf{e}'_n \right| = \left\langle \mathbf{e}_1 \ldots \mathbf{e}_n \right| \cdot \mathbf{T}. \tag{2}$$

Введена таким чином матриця $\mathbf{T}$ називається ***матрицею переходу*** від базису $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ до базису $\{\mathbf{e}'_1, \mathbf{e}'_2, \ldots, \mathbf{e}'_n\}$. Очевидно, що матриця переходу невироджена: $|\mathbf{T}| \neq 0$, оскільки її стовпчики, що містять координати базисних векторів, лінійно незалежні. Тому існує обернена матриця $\mathbf{T}^{-1}$, яка вочевидь буде матрицею переходу від нового базису $\{\mathbf{e}'_1, \mathbf{e}'_2, \ldots, \mathbf{e}'_n\}$ до старого $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$: $\left\langle \mathbf{e}_1 \ldots \mathbf{e}_n \right| = \left\langle \mathbf{e}'_1 \ldots \mathbf{e}'_n \right| \cdot \mathbf{T}^{-1}$ – це прямий наслідок з рівності (2). З'ясуємо, як змінюються при цьому координати довільного вектора $\mathbf{x}\in L$. Припустимо, що $\mathbf{x}=\alpha^i\mathbf{e}_i=\beta^i\mathbf{e}'_i$, тобто $|\boldsymbol{\alpha}\rangle=\left(\alpha^1\ldots\alpha^n\right)^T$ та $|\boldsymbol{\beta}\rangle=\left(\beta^1\ldots\beta^n\right)^T$ – вектори-стовпчики координат цього вектора у базисах $\{\mathbf{e}_1,\mathbf{e}_2,\ldots,\mathbf{e}_n\}$ та $\{\mathbf{e}'_1,\mathbf{e}'_2,\ldots,\mathbf{e}'_n\}$ відповідно. Тоді $\mathbf{x}=\langle\mathbf{e}_1\ldots\mathbf{e}_n|\cdot|\boldsymbol{\alpha}\rangle$. З іншого ж боку, $\mathbf{x}=\langle\mathbf{e}'_1\ldots\mathbf{e}'_n|\cdot|\boldsymbol{\beta}\rangle=\langle\mathbf{e}_1\ldots\mathbf{e}_n|\cdot\mathbf{T}\cdot|\boldsymbol{\beta}\rangle$. З однозначності розкладу по базису випливає рівність $|\boldsymbol{\alpha}\rangle=\mathbf{T}\cdot|\boldsymbol{\beta}\rangle$, або $\alpha^i=\langle\mathbf{T}^i\|\beta\rangle=t^i_j\beta^j$ (3)

Порівнюючи рівності (1) та (3), бачимо, що координати вектора у старому базисі є лінійною комбінацією нових координат, причому коефіцієнти цієї лінійної комбінації знаходяться ***у рядках матриці переходу*** $\mathbf{T}$.

**Теорема 1.** Будь-яка невироджена матриця є матрицею переходу від деякого фіксованого базису $\{\mathbf{e}_1,\mathbf{e}_2,\ldots,\mathbf{e}_n\}$ простору $L$ до деякого нового базису.

► Дійсно, нехай $\mathbf{T}$ – довільна матриця, причому $|\mathbf{T}|\neq 0$, а $\{\mathbf{e}_1,\mathbf{e}_2,\ldots,\mathbf{e}_n\}$ – деякий фіксований базис простору $L$. Розглянемо систему векторів $\{\mathbf{e}'_1,\mathbf{e}'_2,\ldots,\mathbf{e}'_n\}$, визначену формулами (1). Очевидно, що ці вектори є лінійно незалежними, оскільки такими є стовпчики матриці $\mathbf{T}$. А оскільки їх кількість рівна розмірності простору $L$, то ці вектори утворюють базис. Матриця $\mathbf{T}$ буде матрицею переходу до цього базису. ◄

Приклад 1. Розглянемо матрицю $\mathbf{T}=\begin{pmatrix}\cos\alpha & -\sin\alpha\\ \sin\alpha & \cos\alpha\end{pmatrix}$. Оскільки $|\mathbf{T}|=1$, ця матриця є матрицею переходу від довільного базису $\{\mathbf{i},\mathbf{j}\}$ площини до деякого базису $\{\mathbf{i}',\mathbf{j}'\}$, де $\mathbf{i}'=\cos\alpha\cdot\mathbf{i}+\sin\alpha\cdot\mathbf{j}$, $\mathbf{j}'=\mathbf{e}'_2=-\sin\alpha\cdot\mathbf{i}+\cos\alpha\cdot\mathbf{j}$. Координати в старому базисі довільного вектора $\mathbf{a}=\begin{pmatrix}x\\ y\end{pmatrix}$ через координати у новому базисі задаватимуться формулами: $\begin{pmatrix}x\\ y\end{pmatrix}=\mathbf{T}\cdot\begin{pmatrix}x'\\ y'\end{pmatrix}$, або $x=x'\cos\alpha-y'\sin\alpha$, $y=x'\sin\alpha+y'\cos\alpha$. Як відомо, цими формулами визначається поворот системи координат на кут $\alpha$.

## 2. Ізоморфізм векторних просторів.

**Означення 1.** Векторні простори $L$ та $L'$ називаються ***ізоморфними***, якщо існує взаємно-однозначна відповідність між векторами просторів, коли кожному вектору $\mathbf{x}\in L$ ставиться у відповідність вектор $\mathbf{x}'\in L'$ $(\mathbf{x}\leftrightarrow\mathbf{x}')$, причому, якщо $\mathbf{x}\leftrightarrow\mathbf{x}'$ та $\mathbf{y}\leftrightarrow\mathbf{y}'$, то $\mathbf{x}+\mathbf{y}\leftrightarrow\mathbf{x}'+\mathbf{y}'$ та $\lambda\mathbf{x}\leftrightarrow\lambda\mathbf{x}'$ $\forall\mathbf{x},\mathbf{y}\in L$ $\forall\lambda\in\mathbf{R}$.

**Зауваження 1.** Термін «ізоморфізм» означає схожість, подібність. Цю подібність ми використовували, коли ставили у відповідність геометричним векторам вектори-стовпчики їх координат та навіть ототожнювали їх.

**Теорема 2.** Нехай простори $L$ та $L'$ ізоморфні. Системі лінійно незалежних векторів в просторі $L$ відповідає система лінійно незалежних векторів в просторі $L'$.

► Нехай системі лінійно незалежних векторів $\mathbf{x}_1,\mathbf{x}_2,\ldots,\mathbf{x}_k$ із $L$ відповідає система векторів $\mathbf{x}_1{}',\mathbf{x}_2{}',\ldots,\mathbf{x}_k{}'$ із $L'$. Доведемо їх лінійну незалежність. Розглянемо довільну нульову лінійну комбінацію векторів $\mathbf{x}_1{}',\mathbf{x}_2{}',\ldots,\mathbf{x}_k{}'$: $\lambda^1\mathbf{x}_1{}'+\lambda^2\mathbf{x}_2{}'+\ldots+\lambda^k\mathbf{x}_k{}'=\mathbf{0}'$. Їй відповідає також нульова лінійна комбінація векторів із $L$: $\lambda^1\mathbf{x}_1+\lambda^2\mathbf{x}_2+\ldots+\lambda^k\mathbf{x}_k=\mathbf{0}$. Із-за лінійної незалежності цих векторів вона тривіальна: $\lambda^1 = \lambda^2 = \ldots = \lambda^k = 0$. Таким чином, вектори $\mathbf{x}_1{}', \mathbf{x}_2{}', \ldots, \mathbf{x}_k{}'$ також лінійно незалежні. ◄

**<u>Теорема 3</u>.** Розглянемо векторні простори $L$ та $L'$. Нехай $\dim L = n$, а $\dim L' = m \neq n$. Тоді простори $L$ та $L'$ не ізоморфні.

► Від супротивного припустимо, що простори $L$ та $L'$ ізоморфні. Не втрачаючи загальності, можна вважати, що $n > m$. Тоді системі із $n$ базисних векторів простору $L$ має відповідати система із $n$ лінійно незалежних векторів простору $L'$, що неможливо, оскільки розмірність цього простору менша за $n$. ◄

**<u>Теорема 4</u>.** Всі векторні простори рівної розмірності ізоморфні між собою.

► Доведемо цю теорему, встановивши безпосередньо ізоморфізм між елементами двох довільно вибраних просторів $L$ та $L'$ розмірності $n$.

Отже, нехай $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ – деякий базис простору $L$. Виберемо також довільний базис $\{\mathbf{e}'_1, \mathbf{e}'_2, \ldots, \mathbf{e}'_n\}$ в просторі $L'$. Розглянемо довільний вектор $\mathbf{x} \in L:\ x = \alpha^i \mathbf{e}_i, i = \overline{1,n}$. Поставимо йому у відповідність вектор $\mathbf{x}' \in L:\ \mathbf{x}' = \alpha^i \mathbf{e}'_i, i = \overline{1,n}$. В силу єдиності розкладу будь-якого вектору по базису ця відповідність взаємно-однозначна. Покажемо, що таким чином встановлений ізоморфізм просторів $L$ та $L'$. Дійсно розглянемо два довільних вектори $\mathbf{x}, \mathbf{y} \in L:\ \mathbf{x} = \alpha^i \mathbf{e}_i, \mathbf{y} = \beta^i \mathbf{e}_i, i = \overline{1,n}$. Їм відповідають вектори $\mathbf{x}', \mathbf{y}' \in L:\ \mathbf{x}' = \alpha^i \mathbf{e}'_i, \mathbf{y}' = \beta^i \mathbf{e}'_i, i = \overline{1,n}$. Але тоді сумі векторів $\mathbf{x} + \mathbf{y} = (\alpha^i + \beta^i)\mathbf{e}_i,\ i = \overline{1,n}$ відповідає вектор $(\alpha^i + \beta^i)\mathbf{e}'_i,\ i = \overline{1,n}$, рівний сумі $\mathbf{x}' + \mathbf{y}'$. Аналогічно встановлюємо, що вектору $\lambda\mathbf{x} \in L$ відповідає вектор $\lambda\mathbf{x}' \in L'$.◄

Таким чином, єдиною істотною характеристикою векторного (лінійного) простору є його розмірність.

### 3. Підпростори векторних просторів.

**<u>Означення 2</u>.** Непорожня підмножина $L'$ елементів векторного простору $L$ називається його підпростором, якщо $\forall \mathbf{x}, \mathbf{y} \in L' \Rightarrow \mathbf{x} + \mathbf{y} \in L';\ \lambda\mathbf{x} \in L' \forall \lambda \in \mathbf{R}$.

**<u>Наслідок 1</u>.** Будь-яка лінійна комбінація векторів підпростору належить цьому підпростору.

**<u>Наслідок 2</u>.** Будь-який підпростір векторного простору містить нульовий елемент.

► Нехай $L'$ – підпростір векторного простору $L$. Для довільного вектора $\mathbf{x} \in L'$ лінійна комбінація $0 \cdot \mathbf{x} = \mathbf{0} \in L'$ ◄

**<u>Наслідок 3</u>.** Будь-який підпростір векторного простору разом з будь-яким вектором містить його протилежний вектор.

► Дійсно, нехай $L'$ – деякий підпростір векторного простору $L$. Тоді $\forall \mathbf{x} \in L' \Rightarrow (-1) \cdot \mathbf{x} = -\mathbf{x} \in L'$ ◄

<u>Приклад 2</u>. Для будь-якого векторного простору $L$: $\{\mathbf{0}\}$ та $L$ – його підпростори.

<u>Приклад 3</u>. Пряма та площина, що проходять через початок координат, є підпросторами простору геометричних векторів.

**<u>Теорема 5</u>.** Нехай $L'$ – підпростір векторного простору $L$. Тоді $\dim L' \leq \dim L$, причому знак «=» можливий тоді і лише тоді, коли $L' = L$.

► Випадок, коли $L' = \{0\}$ не розглядатимемо в силу тривіальності. Нехай $\dim L' = m \neq 0$. Розглянемо деякий базис $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_m\}$ підпростору $L'$. Якщо $L' \neq L$, то знайдеться ненульовий вектор $\mathbf{f} \in L, \mathbf{f} \notin L'$. Якщо розглянути його разом з базисними векторами $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_m\}$, одержимо знову систему лінійно незалежних векторів простору $L$. А це і означає, що $m \leq \dim L$. У випадку, коли $L' = L$, такого вектора $\mathbf{f} \in L, \mathbf{f} \notin L'$ не існує, а тому $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_m\}$ є базисом всього простору. Отже, $\dim L' = \dim L$.

Якщо ж $\dim L' = \dim L$, то в цих підпросторах базиси складаються з рівної кількості векторів, а отже, будь-який вектор із $L$ лінійно виражається через вектори базису підпростору $L'$ і навпаки. Це і означає, що кожний вектор належить водночас і $L$, і $L'$, тобто $L' = L$. ◄

**<u>Наслідок</u>.** Нехай $L'$ – підпростір векторного простору $L$, причому $\dim L' = m, \dim L = n, m < n$. Якщо $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_m\}$ – базис підпростору $L'$, а $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_m, \mathbf{e}_{m+1}, \ldots, \mathbf{e}_n\}$ – базис всього простору $L$, то будь-який вектор $\mathbf{x} \in L': \mathbf{x} = \alpha^i \mathbf{e}_i, i = \overline{1,n}$ і лише він має координати $\alpha_{m+1} = \alpha_{m+2} = \cdots = \alpha_n = 0$.

► <u>Необхідність</u>. Нехай $\mathbf{x} \in L': x = \beta^i \mathbf{e}_i, i = \overline{1,m}$. З іншого боку, як вектор простору $L$, він може бути розкладений по базису $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_m, \mathbf{e}_{m+1}, \ldots, \mathbf{e}_n\}$: $\mathbf{x} = \alpha^i \mathbf{e}_i, i = \overline{1,n}$. З єдиності розкладу по базису випливає, що $\alpha^i = \beta^i, i = \overline{1,m}$, а $\alpha_{m+1} = \alpha_{m+2} = \cdots = \alpha_n = 0$.

Достатність. Нехай для деякого вектора $\mathbf{x} \in L : x = \alpha^i \mathbf{e}_i, i = \overline{1, m}$. Тоді цей вектор є лінійною комбінацією базисних векторів підпростору $L'$, а отже, є елементом цього підпростору.

### 4. Пряма сума підпросторів.

Нехай $L'$ та $L''$ – підпростори векторного простору $L$.

**Означення 3.** Якщо будь-який вектор $\mathbf{x} \in L$ може бути поданий у вигляді суми $\mathbf{x} = \mathbf{x}' + \mathbf{x}''$, де $\mathbf{x}' \in L', \mathbf{x}'' \in L''$, причому таке представлення єдине, то простір $L$ розкладається в ***пряму суму підпросторів*** $L'$ та $L''$. Цей факт позначається формулою: $L = L' \oplus L''$.

**Теорема 6.** Для того, щоб простір $L$ був прямою сумою своїх підпросторів $L'$ та $L''$ достатньо виконання наступних умов: 1) єдиним спільним елементом підпросторів $L'$ та $L''$ є нульовий вектор; 2) $\dim L' + \dim L'' = \dim L$.

► Нехай виконані умови 1) та 2). Позначимо $\dim L' = k$, $\dim L'' = m$. Тоді $k + m = n$. Нехай $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_k\}$ та $\{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_m\}$ – відповідно базиси підпросторів $L'$ та $L''$. Розглянемо систему векторів $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_k, \mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_m\}$. Складемо довільну нульову лінійну комбінацію векторів цієї системи: $\alpha^i \mathbf{e}_i + \beta^j \mathbf{f}_j = 0,\ i = \overline{1, k},\ j = \overline{1, m}$. Запишемо її у вигляді рівності $\alpha^i \mathbf{e}_i = (-\beta^j) \mathbf{f}_j$. Вектор у лівій частині належить підпростору $L'$, а вектор у правій – підпростору $L''$. Згідно умові 1), такий вектор є нульовим: $\alpha^i \mathbf{e}_i = 0, i = \overline{1, k}$, $\beta^j \mathbf{f}_j = 0, j = \overline{1, m}$. Оскільки вектори $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_k\}$ та $\{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_m\}$ – лінійно незалежні, то $\alpha^i = 0, i = \overline{1, k}$, $\beta^j = 0, j = \overline{1, m}$, що означає що вектори $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_k, \mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_m\}$ також лінійно незалежні, а отже, утворюють базис простору $L$. Далі, розглянемо довільний вектор $\mathbf{x} \in L :\ \mathbf{x} = \alpha^i \mathbf{e}_i + \beta^j \mathbf{f}_j,\ i = \overline{1, k},\ j = \overline{1, m}$. Його можна подати у вигляді $\mathbf{x} = \mathbf{x}' + \mathbf{x}''$, де $\mathbf{x}' = \alpha^i \mathbf{e}_i \in L'$, а $\mathbf{x}'' = \beta^j \mathbf{f}_j \in L''$. Залишилось переконатись у єдиності такого подання. Від супротивного припустимо, що деякий вектор $\mathbf{x}$ має різні представлення: $\mathbf{x} = \mathbf{x}' + \mathbf{x}''$ та $\mathbf{x} = \tilde{\mathbf{x}}' + \tilde{\mathbf{x}}''$. Тоді матимемо: $\tilde{\mathbf{x}}' + \tilde{\mathbf{x}}'' = \mathbf{x}' + \mathbf{x}''$, або $\tilde{\mathbf{x}}' - \mathbf{x}' = \mathbf{x}'' - \tilde{\mathbf{x}}''$, де вектор у лівій частині цієї рівності належить $L'$, а у правій – $L''$. Отже, $\tilde{\mathbf{x}}' - \mathbf{x}' = \mathbf{0}$, $\mathbf{x}'' - \tilde{\mathbf{x}}'' = \mathbf{0}$ тобто $\tilde{\mathbf{x}}' = \mathbf{x}', \mathbf{x}'' = \tilde{\mathbf{x}}''$, що суперечить припущенню про різні представлення вектора $\mathbf{x}$. Таким чином, простір $L$ є прямою сумою підпросторів $L'$ та $L''$.◄

Приклад 4. Розглянемо простір $\mathbf{R}^3$ геометричних векторів. Нехай $\pi$ та $l$ – відповідно деякі площина та пряма ($l \not\in \pi$), що проходять через початок координат. Тоді $\pi$ та $l$ є підпросторами простору $\mathbf{R}^3$, для яких виконані умови теореми: по-перше, єдиним спільним елементом для $\pi$ та $l$ є нульовий вектор; по-друге, $\dim l = 1, \dim \pi = 2$. Отже, $\mathbf{R}^3 = \pi \oplus l$.

**Означення 4.** Розглянемо підмножину векторів $M \subseteq L$ .Якщо будь-який вектор $\mathbf{x} \in M$ може бути поданий у вигляді суми $\mathbf{x} = \mathbf{x}' + \mathbf{x}''$, де $\mathbf{x}' \in L', \mathbf{x}'' \in L''$, то $M$ називається ***сумою підпросторів*** $L'$ та $L''$: $M = L' + L''$.

**Означення 5.** Розглянемо підмножину векторів $M \subseteq L$ .Якщо для будь-якого вектору $\mathbf{x} \in M$ виконується $x \in L', x \in L''$, то $M$ називається ***перетином підпросторів*** $L'$ та $L''$: $M = L' \cap L''$.

**Теорема 7.** Сума та перетин будь-яких лінійних підпросторів є лінійними підпросторами того ж простору, причому

$$\dim(L' \cup L'') = \dim L' + \dim L'' - \dim(L' \cap L'').$$

Пропонуємо переконатись в цьому самостійно.

### 5. Скалярний добуток векторів лінійного простору.

Поняття геометричного вектора та простору геометричних векторів ми узагальнили визначенням лінійного (векторного) простору, аксіоматично ввівши властивості додавання векторів та множення на скаляр таким чином, щоб були виконані відповідні властивості цих операцій для геометричних векторів. Введемо тепер поняття скалярного добутку векторів довільного лінійного простору. Властивості скалярного добутку векторів визначимо з допомогою аксіом таким чином, щоб збереглися звичайні властивості скалярного добутку.

**Означення 6.** Будемо говорити що у дійсному векторному просторі $L$ задана операція ***скалярного добутку*** векторів, якщо будь-якій парі векторів $\mathbf{x}, \mathbf{y} \in L$

цього простору ставиться у відповідність дійсне число $\mathbf{x}\cdot\mathbf{y}$ таким чином, що виконані наступні аксіоми:

1. $\mathbf{x}\cdot\mathbf{y}=\mathbf{y}\cdot\mathbf{x}\quad\forall\mathbf{x},\mathbf{y}\in L$ (комутативність скалярного добутку);
2. $(\lambda\mathbf{x})\cdot\mathbf{y}=\lambda(\mathbf{x}\cdot\mathbf{y})\quad\forall\mathbf{x},\mathbf{y}\in L\,\forall\lambda\in\mathbf{R}$;
3. $(\mathbf{x}_1+\mathbf{x}_2)\cdot\mathbf{y}=\mathbf{x}_1\cdot\mathbf{y}+\mathbf{x}_2\cdot\mathbf{y}\quad\forall\mathbf{x}_1,\mathbf{x}_2,\mathbf{y}\in L$ (дистрибутивність скалярного добутку);
4. $\mathbf{x}\cdot\mathbf{x}\geq 0\quad\forall\mathbf{x}\in L$, причому $\mathbf{x}\cdot\mathbf{x}=0\iff\mathbf{x}=\mathbf{0}$

**<u>Означення 7</u>.** Векторний простір $L_n$ зі скалярним добутком називається простором Евкліда, або ***евклідовим*** простором і позначається $E_n$.

<u>Приклад 5</u>**.** Простір геометричних векторів зі звичайним скалярним добутком векторів – тривимірний евклідів простір.

<u>Приклад 6</u>**.** Простір векторів-стовпчиків координат векторів $\mathbf{x}=(x^1,x^2,\ldots,x^n)$ зі скалярним добутком $\mathbf{x}\cdot\mathbf{y}=\mathbf{x}^T\cdot\mathbf{y}=\sum\limits_{i=1}^{n}x^iy^i$, визначеним через звичайне множення матриць є $n$-вимірним евклідовим простором, в чому легко переконатись безпосередньо.

<u>Приклад 7</u>**.** Простір $C([a,b])$ неперервних на відрізку $[a,b]$ функцій є евклідовим простором зі скалярним добутком $f\cdot g=\int\limits_a^b f(x)g(x)dx\ \forall f,g\in C([a,b])$. Виконання аксіом 1 – 4 безпосередньо випливає із властивостей інтеграла.

**<u>Означення 8</u>.** Довжиною вектора $\mathbf{x}\in E_n$ називається число $|\mathbf{x}|=\sqrt{\mathbf{x}\cdot\mathbf{x}}$.

**<u>Означення 9</u>.** Кутом між ненульовими векторами $\mathbf{x},\mathbf{y}\in E_n$ називається число $\varphi=\arccos\left(\dfrac{\mathbf{x}\cdot\mathbf{y}}{|\mathbf{x}||\mathbf{y}|}\right)$.

**<u>Означення 10</u>.** Вектори $\mathbf{x},\mathbf{y}\in E_n$ називаються ***ортогональними***, якщо $\varphi=\dfrac{\pi}{2}$, тобто $\mathbf{x}\cdot\mathbf{y}=0$.

**<u>Зауваження 1</u>.** Нуль-вектор і лише він ортогональний до кожного вектора $\mathbf{x}\in E_n$.

► Дійсно, нехай $\mathbf{y}\in E_n$ – довільний вектор. Тоді $\mathbf{0}\cdot\mathbf{y}=0$. Отже, довільний вектор $\mathbf{y}\in E_n$ – ортогональний до нуль-вектора. З іншого боку, якщо для деякого вектора $\mathbf{x}\in E_n$ виконується рівність $\mathbf{x}\cdot\mathbf{y}=0\ \forall\mathbf{y}\in E_n$, то зокрема при $\mathbf{y}=\mathbf{x}$ маємо $\mathbf{x}\cdot\mathbf{x}=0$, тоді з четвертої аксіоми скалярного добутку випливає, що $\mathbf{x}=0$ ◄

**<u>Зауваження 2</u>.** Для того, щоб означення **9** було змістовним, необхідно, щоб $-1\leq\dfrac{\mathbf{x}\cdot\mathbf{y}}{|\mathbf{x}||\mathbf{y}|}\leq 1$, або $|\mathbf{x}\cdot\mathbf{y}|\leq|\mathbf{x}||\mathbf{y}|\ \forall\mathbf{x},\mathbf{y}\in E_n$. Покажемо, що це дійсно так. Розглянемо вектор $\mathbf{x}-\lambda\mathbf{y}$, де $\mathbf{x},\mathbf{y}\in E_n$, а $\lambda$ – довільне число. Тоді з аксіоми 4 означення скалярного добутку випливає, що $(\mathbf{x}-\lambda\mathbf{y})^2=(\mathbf{x}-\lambda\mathbf{y})\cdot(\mathbf{x}-\lambda\mathbf{y})\geq 0$. Тобто $\mathbf{x}\cdot\mathbf{x}-2\lambda\mathbf{x}\cdot\mathbf{y}+\lambda^2\mathbf{y}\cdot\mathbf{y}\geq 0$, або $|\mathbf{x}|^2-2\lambda\mathbf{x}\cdot\mathbf{y}+\lambda^2|\mathbf{y}|^2\geq 0\,\forall\lambda\in\mathbf{R}$. Квадратний тричлен відносно числа $\lambda$ невід’ємний при всіх $\lambda$ тоді і лише тоді, коли він має від’ємний дискримінант: $(\mathbf{x}\cdot\mathbf{y})^2-|\mathbf{x}|^2|\mathbf{y}|^2\leq 0\ \ \forall\mathbf{x},\mathbf{y}\in E_n$, або $|\mathbf{x}\cdot\mathbf{y}|\leq|\mathbf{x}||\mathbf{y}|$. Одержана нерівність носить назву ***нерівності Коші-Буняковського***. Дуже важливо зауважити також, що знак «=» можливий тоді і лише тоді, коли вектори $\mathbf{x},\mathbf{y}\in E_n$ – лінійно залежні. Переконайтесь у цьому самостійно.

**Наслідок 1.** З нерівності Коші-Буняковського для довільних векторів просторів прикладу **6** випливає наступна нерівність: $\sum_{i=1}^{n} x^i y^i \le \sqrt{\sum_{i=1}^{n}\left(x^i\right)^2}\sqrt{\sum_{i=1}^{n}\left(y^i\right)^2}$. Для функцій простору $C([a,b])$ (приклад **7**) – нерівність Коші-Буняковського має вигляд: $\int_a^b f(x)g(x)dx \le \sqrt{\int_a^b f^2(x)dx}\sqrt{\int_a^b g^2(x)dx} \;\; \forall f, g \in C([a,b])$.

**Наслідок 2.** З нерівності Коші-Буняковського випливає ***нерівність трикутника***: $|\mathbf{x}+\mathbf{y}| \le |\mathbf{x}| + |\mathbf{y}| \;\; \forall \mathbf{x}, \mathbf{y} \in E_n$.

► Назва цієї нерівності пов'язана з її очевидною наочністю для геометричних векторів. Отже, розглянемо для довільних векторів $\mathbf{x}, \mathbf{y} \in E_n$

$$|\mathbf{x}+\mathbf{y}|^2 = (\mathbf{x}+\mathbf{y})\cdot(\mathbf{x}+\mathbf{y}) = |\mathbf{x}|^2 + |\mathbf{y}|^2 + 2\mathbf{x}\cdot\mathbf{y} \le |\mathbf{x}|^2 + |\mathbf{y}|^2 + 2|\mathbf{x}\cdot\mathbf{y}| \le |\mathbf{x}|^2 + |\mathbf{y}|^2 + 2|\mathbf{x}||\mathbf{y}| = (|\mathbf{x}|+|\mathbf{y}|)^2$$

тобто $|\mathbf{x}+\mathbf{y}| \le |\mathbf{x}| + |\mathbf{y}|$. ◄

Завдання для самостійної роботи. Записати нерівність трикутника для векторів просторів прикладів **6** та **7**.

**Означення 11.** Відстанню між векторами $\mathbf{x}, \mathbf{y} \in E_n$ називатимемо число $d = |\mathbf{x}-\mathbf{y}|$.

## 6. Ортонормований базис евклідового простору.

**Означення 12.** Система векторів $\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_m \in E_n$ називається ***ортогональною***, якщо $\mathbf{f}_i \cdot \mathbf{f}_j = 0$ при $i \ne j \; i, j = \overline{1,m}$ та ***ортонормованою***, якщо $\mathbf{f}_i \cdot \mathbf{f}_j = \delta_{ij} = \begin{cases} 1, i = j \\ 0, i \ne j \end{cases}$ при $i, j = \overline{1,m}$.

**Теорема 8.** Ортогональна система векторів, яка не містить нуль-вектора, – лінійно незалежна.

► Нехай $\{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_m\} \in E_n$ – ортогональна система векторів. Розглянемо довільну лінійну комбінацію цих векторів, яка рівна нулю: $\alpha^1\mathbf{f}_1 + \alpha^2\mathbf{f}_2 + \cdots + \alpha^m\mathbf{f}_m = \mathbf{0}$. Помножимо цю рівність скалярно на вектор $\mathbf{f}_1$: $\alpha^1\mathbf{f}_1\cdot\mathbf{f}_1 + \sum_{i=2}^{m}\alpha^i\mathbf{f}_i\cdot\mathbf{f}_1 = \alpha^1\mathbf{f}_1\cdot\mathbf{f}_1 = \mathbf{0}$. Оскільки $\mathbf{f}_1 \ne \mathbf{0}$, то $\alpha^1 = 0$. Аналогічно переконаємось, що $\alpha^2 = \cdots = \alpha^m = 0$, тобто довільна нульова лінійна комбінація цих векторів – тривіальна. ◄

**Наслідок.** Будь-яка ортонормована система векторів – лінійно незалежна.

**Теорема 9.** В евклідовому просторі $E_n$ завжди існує ортонормовна система $n$ векторів $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n \in E_n$, тобто ***існує ортонормований базис*** $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$.

► Для доведення цієї теореми розглянемо довільний базис $\{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_n\} \in E_n$ евклідового простору $E_n$ та побудуємо шуканий ортонормований базис, скориставшись ***процедурою ортогоналізації*** Грама-Шмідта.

1. Покладемо $\mathbf{a}_1 = \mathbf{f}_1, \; \mathbf{e}_1 = \dfrac{\mathbf{a}_1}{|\mathbf{a}_1|}$. Таким чином, побудований перший вектор майбутнього ортонормованого базису.
2. Вектор $\mathbf{a}_2$ шукаємо у вигляді $\mathbf{a}_2 = \mathbf{f}_2 + \alpha^1\mathbf{a}_1$, підбираючи значення параметру $\alpha^1$ з умови ортогональності $\mathbf{a}_1$ та $\mathbf{a}_2$, тобто: $\mathbf{a}_2\cdot\mathbf{a}_1 = (\mathbf{f}_2 + \alpha^1\mathbf{a}_1)\cdot\mathbf{a}_1 = \mathbf{f}_2\cdot\mathbf{a}_1 + \alpha^1\mathbf{a}_1\cdot\mathbf{a}_1 = 0$ (У випадку геометричних векторів,

наприклад, це означає вибір вектора $\mathbf{a}_2$ із площини векторів $\mathbf{f}_1, \mathbf{f}_2$, ортогонального до $\mathbf{a}_1$). Отже, покладемо $\alpha^1 = -\frac{\mathbf{f}_2 \cdot \mathbf{a}_1}{\mathbf{a}_1 \cdot \mathbf{a}_1}$. Далі визначимо $\mathbf{e}_2 = \frac{\mathbf{a}_2}{|\mathbf{a}_2|}$. Зрозуміло, що система векторів $\{\mathbf{e}_1, \mathbf{e}_2\}$ – ортонормована.

3. Припустимо, що таким чином вже побудовано систему з $k$ $(k < n)$ попарно ортогональних векторів $\{\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_k\}$ і відповідно ортонормовану систему векторів $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_k\}$ . Побудуємо наступні вектори, а саме. Вектор $\mathbf{a}_{k+1}$ шукатимемо у вигляді лінійної комбінації $\mathbf{a}_{k+1} = \mathbf{f}_{k+1} + \sum_{i=1}^{k} \alpha^i \mathbf{a}_i$ , визначаючи параметри $\alpha^i$ таким чином, щоб вектор $\mathbf{a}_{k+1}$ був ортогональним до всіх векторів $\mathbf{a}_i, i = \overline{1,k}$:

$$\begin{cases} \left(\mathbf{f}_{k+1} + \sum_{i=1}^{k} \alpha^i \mathbf{a}_i\right) \cdot \mathbf{a}_1 = 0 \\ \left(\mathbf{f}_{k+1} + \sum_{i=1}^{k} \alpha^i \mathbf{a}_i\right) \cdot \mathbf{a}_2 = 0 \\ \ldots \\ \left(\mathbf{f}_{k+1} + \sum_{i=1}^{k} \alpha^i \mathbf{a}_i\right) \cdot \mathbf{a}_k = 0 \end{cases}, \text{ звідки одержимо } \begin{cases} \mathbf{f}_{k+1} \cdot \mathbf{a}_1 + \alpha^1 \mathbf{a}_1 \cdot \mathbf{a}_1 = 0 \\ \mathbf{f}_{k+1} \cdot \mathbf{a}_2 + \alpha^2 \mathbf{a}_2 \cdot \mathbf{a}_2 = 0 \\ \ldots \\ \mathbf{f}_{k+1} \cdot \mathbf{a}_k + \alpha^k \mathbf{a}_k \cdot \mathbf{a}_k = 0 \end{cases}, \text{ тобто}$$

$$\alpha^i = -\frac{\mathbf{f}_{k+1} \cdot \mathbf{a}_i}{\mathbf{a}_i \cdot \mathbf{a}_i}, \; i = \overline{1,k}$$

Визначивши таким чином вектор $\mathbf{a}_{k+1}$, покладемо $\mathbf{e}_{k+1} = \frac{\mathbf{a}_{k+1}}{|\mathbf{a}_{k+1}|}$. Очевидно, що включення цього вектора в систему $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_k\}$ залишає її ортонормованою.

4. Процедуру ортогоналізації продовжимо, поки не побудуємо $n$ векторів $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$, що утворюють ортогональний базис евклідового простору $E_n$.

Доведення теореми можна вважати завершеним, якщо покажемо, що жоден з так побудованих векторів $\mathbf{a}_k$ (а, отже, і $\mathbf{e}_k$) не є нульовим. Дійсно, вектор $\mathbf{a}_{k+1} = \mathbf{f}_{k+1} + \sum_{i=1}^{k} \alpha^i \mathbf{a}_i$ – є лінійною комбінацією векторів $\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_{k+1}$ : $\mathbf{a}_{k+1} = \beta^i \mathbf{f}_i, \; i = \overline{1, k+1}$. Якщо припустити, що $\mathbf{a}_{k+1} = \mathbf{0}$, то маємо нульову лінійну комбінацію лінійно незалежних векторів $\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_{k+1}$. Тому всі коефіцієнти $\beta^i = 0, \; i = \overline{1, k+1}$, що неможливо, оскільки принаймні $\beta^{k+1} = 1$. Дана суперечність доводить, що серед одержаних векторів $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$ немає нульового вектора.◄

**7. Вираз скалярного добутку через координати векторів в ортонормованому базисі.**

Нехай $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ – ортонормований базис евклідового простору $E_n$. Розглянемо два довільних вектори $\mathbf{x} = x^i \mathbf{e}_i, \mathbf{y} = y^i \mathbf{e}_i,\ i = \overline{1,n}$. Тоді

$$\mathbf{x} \cdot \mathbf{y} = \left(x^i \mathbf{e}_i\right) \cdot \left(y^j \mathbf{e}_j\right) = x^i y^j \mathbf{e}_i \cdot \mathbf{e}_j = x^i y^j \delta_{ij} = \sum_{i=1}^{n} x^i y^i = \left|\mathbf{x}\right\rangle^T \cdot \left|\mathbf{y}\right\rangle \qquad (4)$$

Формула (4) задає ***вираз скалярного добутку векторів через їх координати в ортонормованому базисі***. Знайдемо в цьому ж базисі координати довільного вектора: $\mathbf{x} \cdot \mathbf{e}_k = \left(x^i \mathbf{e}_i\right) \cdot \mathbf{e}_k = x^k,\ i = \overline{1,n},\ k = \overline{1,n}$ (5)

Приклад 8**.** Побудувати ортонормований базис лінійної оболонки векторів $\mathbf{f}_1 = (2{,}1{,}3{,}-1), \mathbf{f}_2 = (7{,}4{,}3{,}-3),\ \mathbf{f}_3 = (1{,}1{,}-6{,}0), \mathbf{f}_4 = (5{,}7{,}7{,}8)$.

**Зауваження.** Лінійною оболонкою векторів $\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_k$ називається множина векторів $M = \left\{ \sum_{i=1}^{k} \alpha^i \mathbf{f}_i \right\}, \forall \alpha^i \in R$.

► Неважко безпосередньо переконатись, що вектори лінійно залежні, оскільки $\mathbf{f}_3 = \mathbf{f}_2 - 3\mathbf{f}_1$. Тому ортогоналізувати будемо систему векторів $\{\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_4\}$. Отже,

покладемо $\mathbf{a}_1 = \mathbf{f}_1 = (2{,}1{,}3{,}-1),\ \mathbf{e}_1 = \dfrac{\mathbf{a}_1}{|\mathbf{a}_1|} = \frac{1}{\sqrt{15}}(2{,}1{,}3{,}-1)$,

$\mathbf{a}_2 = \mathbf{f}_2 + \lambda \mathbf{a}_1$, де $\lambda = -\dfrac{\mathbf{f}_2 \cdot \mathbf{a}_1}{\mathbf{a}_1 \cdot \mathbf{a}_1} = -\dfrac{(7{,}4{,}3{,}-3) \cdot (2{,}1{,}3{,}-1)}{(2{,}1{,}3{,}-1) \cdot (2{,}1{,}3{,}-1)} = -\dfrac{30}{15} = -2$.

Тому $\mathbf{a}_2 = (7{,}4{,}3{,}-3) - 2(2{,}1{,}3{,}-1) = (3{,}2{,}-3{,}-1)$; $\mathbf{e}_2 = \dfrac{\mathbf{a}_2}{|\mathbf{a}_2|} = \frac{1}{\sqrt{23}}(3{,}2{,}-3{,}-1)$.

Далі, шукаємо $\mathbf{a}_3 = \mathbf{f}_4 + \lambda \mathbf{a}_1 + \mu \mathbf{a}_2$, де

$$\lambda = -\frac{\mathbf{f}_4 \cdot \mathbf{a}_1}{\mathbf{a}_1 \cdot \mathbf{a}_1} = -\frac{(5{,}7{,}7{,}8) \cdot (2{,}1{,}3{,}-1)}{(2{,}1{,}3{,}-1) \cdot (2{,}1{,}3{,}-1)} = -\frac{30}{15} = -2,$$

$$\mu = -\frac{\mathbf{f}_4 \cdot \mathbf{a}_2}{\mathbf{a}_2 \cdot \mathbf{a}_2} = -\frac{(5{,}7{,}7{,}8) \cdot (3{,}2{,}-3{,}-1)}{(3{,}2{,}-3{,}-1) \cdot (3{,}2{,}-3{,}-1)} = -\frac{0}{23} = 0.$$

Отже, $\mathbf{a}_3 = \mathbf{f}_4 - 2\mathbf{a}_1 = (5{,}7{,}7{,}8) - 2(2{,}1{,}3{,}-1) = (1{,}5{,}1{,}10)$; $\mathbf{e}_3 = \dfrac{\mathbf{a}_3}{|\mathbf{a}_3|} = \frac{1}{\sqrt{127}}(1{,}5{,}1{,}10)$ ◄

## 8. Матриця Грама та її властивості.

Нехай $\{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_n\}$ – довільний базис в евклідовому просторі $E_n$. Розглянемо довільні вектори $\mathbf{x}, \mathbf{y} \in E_n$: $\mathbf{x} = x^i \mathbf{f}_i, \mathbf{y} = y^i \mathbf{f}_i,\ i = \overline{1,n}$. Тоді скалярний добуток цих векторів, враховуючи аксіоми 1-3 скалярного добутку, виражається наступним чином:

$\mathbf{x} \cdot \mathbf{y} = \left(x^i \mathbf{f}_i\right) \cdot \left(y^j \mathbf{f}_j\right) = x^i y^j \mathbf{f}_i \cdot \mathbf{f}_j,\ i,j = \overline{1,n}$, або $\mathbf{x} \cdot \mathbf{y} = x^i y^j g_{ij},\ i, j = \overline{1,n}$, (6)

де $g_{ij} = \mathbf{f}_i \cdot \mathbf{f}_j$ . Якщо записати ці елементи у вигляді матриці, одержимо ***матрицю Грама*** базису $\{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_n\}$(позначимо її через $\mathbf{G} = \left(g_{ij}\right), i, j = \overline{1,n}$). Таким чином, для кожного базису евклідового простору визначена матриця Грама, яка містить скалярні добутки базисних векторів. Зазначимо її властивості.

**Теорема 10. (про властивості матриці Грама)**

1. Матриця Грама – симетрична матриця: $\mathbf{G}^T = \mathbf{G}$.

2. Формула (3) скалярного добутку двох векторів може бути записана через матрицю Грама: $\mathbf{x}\cdot\mathbf{y} = |\mathbf{x}\rangle^T \mathbf{G}|\mathbf{y}\rangle$ (7)
3. Матриця Грама будь-якого ортонормованого базису – одинична: $\mathbf{G} = \mathbf{E}_n$
4. Нехай $\mathbf{T}$ – матриця переходу від деякого базису $\{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_n\}$ до базису $\{\mathbf{f}'_1, \mathbf{f}'_2, \ldots, \mathbf{f}'_n\}$. Позначимо відповідно через $\mathbf{G}$ та $\mathbf{G}'$ матриці Грама базисів $\{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_n\}$ та $\{\mathbf{f}'_1, \mathbf{f}'_2, \ldots, \mathbf{f}'_n\}$. Тоді має місце формула перетворення матриці Грама при заміні базису евклідового простору $\mathbf{G}' = \mathbf{T}^T\mathbf{G}\mathbf{T}$ (8)
5. Матриця Грама будь-якого базису – невироджена.
6. **(Критерії лінійної залежності та лінійної незалежності довільної системи векторів в евклідовому просторі)** Нехай $\{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_m\}$ – довільна система векторів в $E_n$. Складемо матрицю Грама цієї системи: $\mathbf{G} = (g_{ij})_{i,j=\overline{1,m}}$, де $g_{ij} = \mathbf{f}_i \cdot \mathbf{f}_j$. Тоді 1) система векторів $\{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_m\}$ – лінійно незалежна, тоді і тільки тоді, коли $|\mathbf{G}| > 0$; 2) система векторів $\{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_m\}$ – лінійно залежна, тоді і тільки тоді, коли $|\mathbf{G}| = 0$.

► Доведемо ці властивості.

1. Дійсно, $g_{ij} = \mathbf{f}_i \cdot \mathbf{f}_j = \mathbf{f}_j \cdot \mathbf{f}_i = g_{ji}\ \ \forall i, j = \overline{1,n}$
2. Доведення очевидне – це формула (6):
$$\mathbf{x}\cdot\mathbf{y} = x^i y^j g_{ij} = \sum_{i=1}^{n} x^i \sum_{j=1}^{n} g_{ij} y^j = \sum_{i=1}^{n} x^i \langle \mathbf{G}_i | \cdot | \mathbf{y}\rangle = |\mathbf{x}\rangle^T \mathbf{G}|\mathbf{y}\rangle$$
3. Дійсно, в цьому випадку $g_{ij} = \mathbf{e}_i \cdot \mathbf{e}_j = \delta_{ij}\ \forall i, j = \overline{1,n}$.
4. Для доведення розглянемо довільний елемент $g'_{ij}$ матриці Грама $\mathbf{G}'$:
$$g'_{ij} = \mathbf{f}'_i \cdot \mathbf{f}'_j = \left(\sum_{k=1}^{n} t_i^k \mathbf{f}_k\right)\cdot\left(\sum_{m=1}^{n} t_j^m \mathbf{f}_m\right) = \sum_{k,m=1}^{n} t_i^k t_j^m g_{km} = |\mathbf{T}_i\rangle^T \cdot \mathbf{G} \cdot |\mathbf{T}_j\rangle,$$
або $\mathbf{G}' = \mathbf{T}^T\mathbf{G}\mathbf{T}$
5. Розглянемо довільний базис $\{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_n\}$ евклідового простору $E_n$. Оскільки в $E_n$ завжди існує ортонормований базис $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$, можна вважати, що базис $\{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_n\}$ одержано саме з нього за допомогою деякої матриці переходу $\mathbf{T}$, тобто $\mathbf{G} = \mathbf{T}^T\mathbf{E}\mathbf{T}$. Тоді $|\mathbf{G}| = |\mathbf{T}^T\mathbf{E}\mathbf{T}| = |\mathbf{T}^T||\mathbf{E}||\mathbf{T}| = |\mathbf{T}^T||\mathbf{T}| = |\mathbf{T}|^2 > 0$, оскільки матриця переходу $\mathbf{T}$ завжди невироджена.
6. Доведемо <u>необхідність</u> першого твердження. Нехай система векторів $\{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_m\}$ – лінійно незалежна. Тоді вона утворює базис підпростору, який є лінійною оболонкою цих векторів, і згідно з п'ятим твердженням має матрицю Грама з додатним визначником, тобто $|\mathbf{G}| > 0$.

   Доведемо <u>необхідність</u> другого твердження. Припустимо, що система векторів $\{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_m\}$ – лінійно залежна. Тоді принаймні один з векторів системи (не втрачаючи загальності, можна вважати, що це $\mathbf{f}_1$) лінійно виражається через решту: $\mathbf{f}_1 = \sum_{k=2}^{m} \lambda^k \mathbf{f}_k$. Звідси одержимо:

$g_{1j} = \mathbf{f}_1 \cdot \mathbf{f}_j = \left( \sum_{k=2}^{m} \lambda^k \mathbf{f}_k \right) \cdot \mathbf{f}_j = \sum_{k=2}^{m} \lambda^k (\mathbf{f}_k \cdot \mathbf{f}_j) = \sum_{k=2}^{m} \lambda^k g_{kj}$, тобто перший рядок матриці Грама системи $\{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_m\}$ лінійно виражається через решту рядків, тому $|\mathbf{G}| = 0$.

Достатність першого твердження доведемо, припустивши супротивне – припустимо, що $|\mathbf{G}| > 0$, але система векторів $\{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_m\}$ – лінійно залежна. Тоді з необхідності другого твердження дістанемо суперечність – $|\mathbf{G}| = 0$.

Аналогічно встановимо достатність другого твердження. Нехай $|\mathbf{G}| = 0$, але система векторів $\{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_m\}$ – лінійно незалежна. Тоді з необхідності другого твердження має виконуватись нерівність $|\mathbf{G}| > 0$. Тим самим критерій повністю доведений.

◄

Приклад 9. Нехай $\{\mathbf{f}_1, \mathbf{f}_2\}$ – базис деякої площини геометричних векторів $\mathbf{R}^2$. Нехай $|\mathbf{f}_1| = 1, |\mathbf{f}_2| = 2$, кут $\varphi$ між базисними векторами рівний $45°$. Визначити матрицю Грама даного базису та записати з її допомогою скалярний добуток довільних векторів $\mathbf{x}, \mathbf{y} \in \mathbf{R}^2$.

Визначимо елементи матриці Грама заданого базису: $g_{11} = \mathbf{f}_1 \cdot \mathbf{f}_1 = 1$, $g_{22} = \mathbf{f}_2 \cdot \mathbf{f}_2 = 4$, $g_{12} = g_{21} = \mathbf{f}_1 \cdot \mathbf{f}_2 = |\mathbf{f}_1||\mathbf{f}_2| \cos\varphi = 2\frac{1}{\sqrt{2}} = \sqrt{2}$. Отже, $\mathbf{G} = \begin{pmatrix} 1 & \sqrt{2} \\ \sqrt{2} & 4 \end{pmatrix}$.

Тому $\forall \mathbf{x}, \mathbf{y} \in \mathbf{R}^2$: $\quad \mathbf{x} \cdot \mathbf{y} = |\mathbf{x}\rangle^T \mathbf{G} |\mathbf{y}\rangle = \begin{pmatrix} x^1 & x^2 \end{pmatrix} \cdot \begin{pmatrix} 1 & \sqrt{2} \\ \sqrt{2} & 4 \end{pmatrix} \cdot \begin{pmatrix} y^1 \\ y^2 \end{pmatrix} =$

$\begin{pmatrix} x^1 + \sqrt{2}x^2 & \sqrt{2}x^1 + 4x^2 \end{pmatrix} \cdot \begin{pmatrix} y^1 \\ y^2 \end{pmatrix} = x^1 y^1 + \sqrt{2}(x^2 y^1 + x^1 y^2) + 4x^2 y^2$ ◄

**Наслідок.** Для будь-якої системи векторів евклідового простору її визначник Грама $\mathbf{G}$ – невід'ємний, причому $|\mathbf{G}| = 0$ тоді і тільки тоді, коли система векторів лінійно залежна.

**Зауваження 1.** Розглянемо матрицю Грама для пари довільних векторів $\mathbf{x}, \mathbf{y} \in E_n$:

$|\mathbf{G}| = \left| \begin{pmatrix} \mathbf{x} \cdot \mathbf{x} & \mathbf{x} \cdot \mathbf{y} \\ \mathbf{x} \cdot \mathbf{y} & \mathbf{y} \cdot \mathbf{y} \end{pmatrix} \right| = |\mathbf{x}|^2 |\mathbf{y}|^2 - (\mathbf{x} \cdot \mathbf{y})^2 \geq 0$, отже, даний наслідок узагальнює нерівність Коші-Буняковського.

**Зауваження 2.** Нехай $\mathbf{T}$ – матриця переходу від довільного ортонормованого базису $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ простору до іншого ортонормованого базису $\{\mathbf{e}'_1, \mathbf{e}'_2, \ldots, \mathbf{e}'_n\}$. Тоді матриці Грама цих базисів – одиничні матриці. Але, згідно з формулою (8) маємо $\mathbf{E}_n = \mathbf{T}^T \mathbf{E}_n \mathbf{T} = \mathbf{T}^T \mathbf{T}$. Оскільки матриця переходу не вироджена, то існує її обернена, а отже, $\mathbf{T}^{-1} = \mathbf{T}^T$.

**Означення 13.** Матриця $\mathbf{Q}$, для якої має місце властивість $\mathbf{Q}^T \mathbf{Q} = \mathbf{E}$, називається ***ортогональною***.

**Теорема 11. (про властивість матриці переходу між двома ортонормованими базисами)**

Матриця переходу від будь-якого ортонормованого базису евклідового простору до іншого ортонормованого базису – ортогональна.

### 9. Ортогональне доповнення підпростору евклідового простору.

Розглянемо $E_k$ – $k$-вимірний підпростір евклідового простору $E_n$.

**Означення 14.** Ортогональним доповненням підпростору $E_k$ називається множина $E_k^{\perp}$ всіх векторів, ортогональних до кожного з векторів $E_k$: $E_k^{\perp} = \{\mathbf{x} \in E_n : \mathbf{x} \cdot \mathbf{y} = \mathbf{0} \ \ \forall \mathbf{y} \in E_k\}$.

**Теорема 12. (про властивості ортогонального доповнення)**

1. Ортогональне доповнення $E_k^{\perp}$ є лінійним підпростором простору $E_n$.
2. Розмірність підпростору $E_k^{\perp}$ рівна $n-k$: $\dim E_k^{\perp} = n-k$.
3. Перетином підпросторів $E_k$ та $E_k^{\perp}$ є лише нуль-вектор: $E_k \cap E_k^{\perp} = \{\mathbf{0}\}$.
4. Ортогональним доповненням до ортогонального доповнення $E_k^{\perp}$ є підпростір $E_k$: $\left(E_k^{\perp}\right)^{\perp} = E_k$.

►

1. Для доведення розглянемо довільні вектори $\mathbf{x}_1, \mathbf{x}_2 \in E_k^{\perp}$ та довільне число $\lambda \in \mathbf{R}$. Візьмемо будь-який вектор $\mathbf{y} \in E_k$ та обчислимо скалярні добутки: $(\mathbf{x}_1 + \mathbf{x}_2) \cdot \mathbf{y} = \mathbf{x}_1 \cdot \mathbf{y} + \mathbf{x}_2 \cdot \mathbf{y} = \mathbf{0}$; $(\lambda \mathbf{x}_1) \cdot \mathbf{y} = \lambda(\mathbf{x}_1 \cdot \mathbf{y}) = \lambda \mathbf{0} = \mathbf{0}$. Таким чином, вектори $\mathbf{x}_1 + \mathbf{x}_2$ та $\lambda \mathbf{x}_1$ (при довільному значенні $\lambda$) належать $E_k^{\perp}$ разом з $\mathbf{x}_1, \mathbf{x}_2$. Отже, $E_k^{\perp}$ – підпростір простору $E_n$.
2. Нехай $\{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_k\}$ – деякий базис підпростору $E_k$. Якщо $\mathbf{x}$ – довільний вектор з ортогонального доповнення $E_k^{\perp}$, то очевидно, $\quad \mathbf{x} \cdot \mathbf{f}_i = 0 \ \ \forall i = \overline{1,k}$. (7)
   Припустимо, що нам відомі координати всіх вказаних векторів в деякому ортонормованому базисі простору $E_n$: $\mathbf{f}_i = \left| f_i^1 \ \ f_i^2 \ \ \ldots \ \ f_i^n \right\rangle^T, \mathbf{x} = \left| x^1 \ \ x^2 \ \ \ldots \ \ x^n \right\rangle^T$. Тоді рівності (7) визначають систему $k$ рівнянь з $n$ невідомими: $\begin{cases} f_1^1 x^1 + f_1^2 x^2 + \cdots + f_1^n x^n = 0 \\ f_2^1 x^1 + f_2^2 x^2 + \cdots + f_2^n x^n = 0 \\ \vdots \\ f_k^1 x^1 + f_k^2 x^2 + \cdots + f_k^n x^n = 0 \end{cases}$ Ранг матриці цієї системи рівний $k$, оскільки вектори $\{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_k\}$ – лінійно незалежні. Тому розмірність простору розв'язків, який, очевидно збігається з $E_k^{\perp}$, рівна $n-k$.
3. Дійсно, припустимо, що існує вектор $\mathbf{x} \in E_n$, такий що $\mathbf{x} \in E_k, \mathbf{x} \in E_k^{\perp}$. Тоді маємо: $\mathbf{x} \cdot \mathbf{x} = \mathbf{0}$, отже, $\mathbf{x} = \mathbf{0}$.
4. З означення ортогонального доповнення випливає, що $E_k \subseteq \left(E_k^{\perp}\right)^{\perp}$. Але оскільки $\dim\left(E_k^{\perp}\right)^{\perp} = n - \dim E_k^{\perp} = k$, то з теореми про розмірність підпростору векторного простору випливає, що $E_k = \left(E_k^{\perp}\right)^{\perp}$.

◄

**Наслідок.** Із другого та третього тверджень випливає, що виконані достатні умови теореми про пряму суму підпросторів, отже $E_k \oplus E_k^{\perp} = E_n$, тобто будь-який вектор $\mathbf{x} \in E_n$ єдиним чином представляється у вигляді $\mathbf{x} = \mathbf{x}' + \mathbf{x}''$, де $\mathbf{x}' \in E_k$, а $\mathbf{x}'' \in E_k^{\perp}$. Вектор $\mathbf{x}'$ називається ***ортогональною проекцією*** вектора $\mathbf{x}$ на підпростір $E_k$, а вектор $\mathbf{x}''$ – ***ортогональною складовою***. З четвертого твердження випливає, що вектор $\mathbf{x}''$ є ортогональною проекцією вектора $\mathbf{x}$ на підпростір $E_k^{\perp}$.

## Лекція 14-15. Лінійні відображення лінійних просторів.

### 1. Основні означення та теореми.

Нехай $L_n$ та $L_m$ – два лінійних простори над полем дійсних (або комплексних) чисел.

**Означення 1.** Відображення $A: L_n \to L_m$, яке кожному вектору $\mathbf{x} \in L_n$ ставить у відповідність вектор $\mathbf{y} \in L_m$ за правилом $\mathbf{y} = A\mathbf{x}$, називається ***лінійним відображенням*** простору $L_n$ в простір $L_n$, якщо

1. $A(\mathbf{x}+\mathbf{y}) = A\mathbf{x} + A\mathbf{y} \quad \forall \mathbf{x} \in L_n, \forall \mathbf{y} \in L_n$; (1)
2. $A\lambda\mathbf{x} = \lambda A\mathbf{x} \quad \forall \mathbf{x} \in L_n, \forall \lambda \in \mathbf{R}$. (2)

Якщо $L_m = L_n$, то лінійне відображення $A$ називається ***лінійним перетворенням*** простору $L_n$ або ***лінійним оператором*** в $L_n$.

Розглянемо деякі приклади.

Приклад 1. Нехай $A\mathbf{x} = \mathbf{0} \ \forall \mathbf{x} \in L_n$. Неважко переконатись, що це лінійний оператор в просторі $L_n$. Його називають ***нуль-оператором***.

Приклад 2. Нехай $A\mathbf{x} = \mathbf{x} \ \forall \mathbf{x} \in L_n$. Тоді $A(\mathbf{x}+\mathbf{y}) = \mathbf{x}+\mathbf{y} = A\mathbf{x} + A\mathbf{y}$; $A(\lambda\mathbf{x}) = \lambda\mathbf{x} = \lambda A\mathbf{x} \ \forall \lambda \in \mathbf{R}$. Цей лінійний оператор називають ***тотожним*** і позначають $E$.

Приклад 3. Визначимо для деякого числа $\alpha$ перетворення $A\mathbf{x} = \alpha\mathbf{x} \ \forall \mathbf{x} \in L_n$. Очевидно, що обидві умови лінійності виконані. Цей лінійний оператор називають оператором гомотетії.

Приклад 4. Визначимо відображення $A: L_n \to L_m \ (m < n)$, для $\mathbf{x} = \begin{pmatrix} x^1 \\ \vdots \\ x^n \end{pmatrix} \in L_n$ за правилом: $A\mathbf{x} = \mathbf{y}$, де $\mathbf{y} = \begin{pmatrix} x^1 \\ \vdots \\ x^m \end{pmatrix} \in L_m \subset L_n$. Легко переконатись, що $A$ – це лінійне відображення.

**Означення 2.** ***Образом (множиною значень)*** лінійного відображення $A: L_n \to L_m$ називається множина $\operatorname{Im} A = \{\mathbf{y} \in L_m : \exists \mathbf{x} \in L_n, \ A\mathbf{x} = \mathbf{y}\}$.

**Означення 3.** ***Ядром*** лінійного відображення $A: L_n \to L_m$ називається множина $\operatorname{Ker} A = \{\mathbf{x} \in L_n : A\mathbf{x} = \mathbf{0}\}$.

**Теорема 1.** Образ лінійного відображення $\operatorname{Im} A$ є лінійним підпростором простору $L_m$.

► Покажемо, що множина $\operatorname{Im} A$ замкнена відносно лінійних операцій над векторами простору $L_m$. Нехай довільні вектори $\mathbf{y}_1, \mathbf{y}_2 \in \operatorname{Im} A$. Тоді знайдуться такі вектори $\mathbf{x}_1, \mathbf{x}_2 \in L_n$, що $\mathbf{y}_k = A\mathbf{x}_k, \ k = 1,2$. Звідси випливає, що $\mathbf{y}_1 + \mathbf{y}_2 = A\mathbf{x}_1 + A\mathbf{x}_2 = A(\mathbf{x}_1 + \mathbf{x}_2) \in \operatorname{Im} A$, оскільки $\mathbf{x}_1 + \mathbf{x}_2 \in L_n$. Цілком аналогічно одержимо, що $\lambda\mathbf{y}_1 = \lambda A\mathbf{x}_1 = A(\lambda\mathbf{x}_1) \in \operatorname{Im} A \ \forall \lambda \in \mathbf{R}$. Отже, $\operatorname{Im} A$– лінійний підпростір $L_m$.◄

**<u>Теорема 2</u>.** Ядро лінійного відображення $\mathrm{Ker}A$ є лінійним підпростором простору $L_n$.

► Нехай довільні вектори $\mathbf{x}_1, \mathbf{x}_2 \in \mathrm{Ker}\ A$. Це означає, що $A\mathbf{x}_k = \mathbf{0},\ k = 1,2$. В силу лінійності відображення $A$ маємо: $A(\mathbf{x}_1 + \mathbf{x}_2) = A\mathbf{x}_1 + A\mathbf{x}_2 = \mathbf{0}$ і $A(\lambda\mathbf{x}_1) = \lambda A\mathbf{x}_1 = \mathbf{0}\ \forall\lambda \in \mathbf{R}$. Тобто множина $\mathrm{Ker}A$ замкнена відносно лінійних операцій над векторами простору $L_n$. Отже, $\mathrm{Ker}A$ – лінійний підпростір $L_n$.◄

**<u>Теорема 3</u>.** Образом довільного підпростору простору $L_n$ при лінійному відображенні $A: L_n \to L_m$ є деякий підпростір простору $L_m$.

► Нехай $L \subseteq L_n$ – довільний лінійний підпростір. Позначимо $M = \{\mathbf{y} \in L_m : \exists\mathbf{x} \in L, A\mathbf{x} = \mathbf{y}\}$. Необхідно показати, що $M \subseteq L_m$ є лінійним підпростором. Розглянемо довільні вектори $\mathbf{y}_1, \mathbf{y}_2 \in M$, а також відповідні їм вектори $\mathbf{x}_1, \mathbf{x}_2 \in L$, для яких $A\mathbf{x}_k = \mathbf{y}_k,\ k = 1,2$. Тоді $\mathbf{y}_1 + \mathbf{y}_2 = A\mathbf{x}_1 + A\mathbf{x}_2 = A(\mathbf{x}_1 + \mathbf{x}_2)$, звідки випливає, що $\mathbf{y}_1 + \mathbf{y}_2 \in M$, оскільки $\mathbf{x}_1 + \mathbf{x}_2 \in L$. Аналогічно, яким би не було число $\lambda \in \mathbf{R}$, $\lambda\mathbf{y}_1 = \lambda A\mathbf{x}_1 = A(\lambda\mathbf{x}_1) \in M$, бо $\lambda\mathbf{x}_1 \in L$. Таким чином, $M$ – лінійний підпростір $L_m$.◄

**Зауваження.** Якщо так зване ***звуження лінійного відображення***, тобто відображення, яке діє не в усьому просторі $L_n$, а лише на деякому його підпросторі $L$, позначити через $A|_L$, то доведена теорема означає, що $\mathrm{Im}\left(A|_L\right)$ є підпростором $L_m$.

**<u>Означення 4</u>.** Нехай $A: L_n \to L_n$ деякий лінійний оператор. Підпростір $L \subseteq L_n$ називається ***інваріантним*** відносно оператора $A$, якщо $\mathrm{Im}\left(A|_L\right) \subseteq L$.

<u>Приклад 5</u>. Очевидно, що $\{\mathbf{0}\}$ і $L_n$ – тривіальні інваріантні підпростори будь-якого лінійного оператора. Вони називаються власними підпросторами простора $L_n$.

Визначимо образ та ядро деяких лінійних операторів.

<u>Приклад 6</u>. Розглянемо нуль-оператор, тобто оператор $A: L_n \to L_n$, для якого $A\mathbf{x} = \mathbf{0}\ \forall\mathbf{x} \in L_n$. Очевидно, що $\mathrm{Im}\,A = \{\mathbf{0}\}$, $\mathrm{Ker}\ A = L_n$.

<u>Приклад 7</u>. Нехай $E: L_n \to L_n$ – тотожний оператор. Легко бачити, що $\mathrm{Im}\,A = L_n$, $\mathrm{Ker}\,A = \{\mathbf{0}\}$.

<u>Приклад 8</u>. Позначимо $P_n(\mathbf{R})$ – лінійний простір поліномів від $x \in \mathbf{R}$ степеню не вищого за $n$. Нехай $A$ – оператор диференціювання поліномів. В силу властивостей похідної, $A$ є лінійним оператором в просторі $P_n(\mathbf{R})$. Тоді, очевидно, $\mathrm{Im}\,A = P_{n-1}(\mathbf{R})$, $\mathrm{Ker}\ A = P_0(\mathbf{R})$.

**<u>Означення 5</u>.** ***Рангом*** лінійного відображення $A$ називається розмірність його образу: $\mathrm{Rg}\,A = \dim(\mathrm{Im}\,A)$.

**<u>Означення 6</u>.** Розмірність ядра лінійного відображення $A$ називається ***дефектом*** цього лінійного відображення.

**<u>Зауваження</u>.** Якщо підрахувати в останніх прикладах суму розмірностей підпросторів $\mathrm{Im}A$ та $\mathrm{Ker}A$, то вона вийде рівною розмірності простору, в якому діє лінійний оператор. Що цей факт не є випадковим, доводить наступна теорема.

**<u>Теорема 4</u>.** Нехай $A: L_n \to L_m$ – лінійне відображення. Тоді сума рангу відображення $A$ та розмірності його ядра рівна розмірності простору, в якому діє відображення $A$:

$$\dim(\mathrm{Im}\ A) + \dim(\mathrm{Ker}\ A) = \dim(L_n) = n \qquad (3)$$

► Позначимо $\dim(\mathrm{Ker}\,A) = k \ge 0$ ( $\mathrm{Ker}\,A$ не порожня множина, оскільки їй належить принаймні вектор $\mathbf{0}$ ). Нехай $\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_k$ – базис $\mathrm{Ker}\,A$. Доповнимо його $n-k$ векторами $\mathbf{f}_{k+1}, \mathbf{f}_{k+2}, \ldots, \mathbf{f}_n$ до повного базису простору $L_n$ . Визначимо вектори $\mathbf{g}_i = A\mathbf{f}_i \in \mathrm{Im}\,A,\ i = k+1, \ldots, n$. Розглянемо довільний вектор $\mathbf{y} \in \mathrm{Im}\,A$, а також його прообраз: $\mathbf{x} \in L_n$ : $A\mathbf{x} = \mathbf{y}$ . Вектор $\mathbf{x}$ може бути розкладений по базису $\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_n$ : $\mathbf{x} = \sum\limits_{i=1}^{n} x^i \mathbf{f}_i$ .

Тоді $\mathbf{y} = A\left(\sum\limits_{i=1}^{n} x^i \mathbf{f}_i\right) = \sum\limits_{i=1}^{n} x^i A\mathbf{f}_i = \sum\limits_{i=k+1}^{n} x^i A\mathbf{f}_i = \sum\limits_{i=k+1}^{n} x^i \mathbf{g}_i$ . Це означає, що $\mathrm{Im}\,A$ є лінійною оболонкою векторів $\mathbf{g}_i,\ i = k+1, \ldots, n$ , тобто: $\mathrm{Im}\,A = \mathrm{Len}(\mathbf{g}_{k+1}, \ldots, \mathbf{g}_n)$. Покажемо, що вектори $\mathbf{g}_i,\ i = k+1, \ldots, n$ – лінійно незалежні, тобто насправді вони утворюють базис в $\mathrm{Im}\,A$. Від супротивного припустимо, що існує нульова нетривіальна лінійна комбінація цих векторів:

$$\lambda^1 \mathbf{g}_{k+1} + \lambda^2 \mathbf{g}_{k+2} + \cdots + \lambda^{n-k} \mathbf{g}_n = \mathbf{0}$$

Розглянемо вектор $\mathbf{q} = \lambda^1 \mathbf{f}_{k+1} + \lambda^2 \mathbf{f}_{k+2} + \cdots + \lambda^{n-k} \mathbf{f}_n$. Оскільки не всі числа $\lambda^i,\ i = k+1, \ldots, n$ рівні нулеві, а вектори $\mathbf{f}_{k+1}, \mathbf{f}_{k+2}, \ldots, \mathbf{f}_n$ – лінійно незалежні як підсистема базису простору $L_n$ , вектор $\mathbf{q} \ne \mathbf{0}$ . Проте $A\mathbf{q} = A\left(\sum\limits_{i=k+1}^{n} \lambda^{i-k} \mathbf{f}_i\right) = \sum\limits_{i=k+1}^{n} \lambda^{i-k} A\mathbf{f}_i = \sum\limits_{i=k+1}^{n} \lambda^{i-k} \mathbf{g}_i = \mathbf{0}$ за припущенням про лінійну залежність векторів $\mathbf{g}_i,\ i = k+1, \ldots, n$ . Отже, $\mathbf{q} \in \mathrm{Ker}\,A$ , а тому розкладається по базису цього підпростору: $\mathbf{q} = \sum\limits_{i=1}^{k} \mu^i \mathbf{f}_i$ . Таким чином, вектор $\mathbf{q} \in L_n$ має різні координати в базисі $\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_n$ : $\mathbf{q}^{\mathrm{T}} = \left(\mu^1 \ldots \mu^k\ 0 \ldots 0\right)$ та $\mathbf{q}^{\mathrm{T}} = \left(0 \ldots 0\ \lambda^1 \ldots \lambda^{n-k}\right)$. Це можливо лише за умови, коли $\mathbf{q} = \mathbf{0}$ , тобто $\mu^i = 0,\ i = 1, \ldots k$ і $\lambda^i = 0,\ i = 1, \ldots, n-k$ , що однак суперечить припущенню про лінійну залежність векторів $\mathbf{g}_i,\ i = k+1, \ldots, n$ . Таким чином, ці вектори утворюють базис підпростору $\mathrm{Im}\,A$, звідки випливає, що ранг відображення $A$ рівний $n-k$ , що й завершує доведення теореми. ◄

## 2. Матриця лінійного відображення.

Нехай, як і раніше, $A : L_n \to L_m$ – лінійне відображення.

Виберемо в лінійних просторах $L_n$ і $L_m$ відповідно базиси $\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_n$ та $\mathbf{g}_1, \mathbf{g}_2, \ldots, \mathbf{g}_m$. Вектори $A\mathbf{f}_k,\ k = 1, \ldots, n$ можуть бути розкладені по базису простору $L_m$:

$$\begin{cases} A\mathbf{f}_1 = a_1^1 \mathbf{g}_1 + a_1^2 \mathbf{g}_2 + \cdots + a_1^m \mathbf{g}_m \\ A\mathbf{f}_2 = a_2^1 \mathbf{g}_1 + a_2^2 \mathbf{g}_2 + \cdots + a_2^m \mathbf{g}_m \\ \qquad\vdots \\ A\mathbf{f}_n = a_n^1 \mathbf{g}_1 + a_n^2 \mathbf{g}_2 + \cdots + a_n^m \mathbf{g}_m \end{cases} \tag{4}$$

або

$$A\mathbf{f}_i = \sum_{k=1}^{m} a_i^k \mathbf{g}_k,\ i = 1, \ldots, n \tag{5}$$

З вектор-стовпчиків координат векторів $A\mathbf{f}_i,\ i = 1, \ldots, n$ можна утворити матрицю розмірностей $m \times n$: $\mathbf{A} = \left(a_i^k\right),\ {\scriptstyle k=\overline{1,m} \atop \scriptstyle i=\overline{1,n}}$ . Вона називається ***матрицею лінійного відображення*** $A$ ***в базисах*** $\{\mathbf{f}_1, \ldots, \mathbf{f}_n\}$, $\{\mathbf{g}_1, \ldots, \mathbf{g}_m\}$. Таким чином, у вибраних базисах кожному лінійному відображенню відповідає певна матриця. Покажемо, що і навпаки: для будь-якої матриці розмірностей $m \times n$ існує лінійне відображення простору $L_n$ в $L_m$, якому у фіксованих базисах цих просторів відповідає задана матриця.

Отже, нехай маємо матрицю $\mathbf{A} = \left(a_i^k\right),\ {\scriptstyle k=\overline{1,m} \atop \scriptstyle i=\overline{1,n}}$ та базиси $\{\mathbf{f}_1, \ldots, \mathbf{f}_n\}$ і $\{\mathbf{g}_1, \ldots, \mathbf{g}_m\}$ в просторах $L_n$ та $L_m$ відповідно. Відображення $A$ задамо на базисних векторах формулою (5). Тоді для довільного вектора $\mathbf{x} \in L_n$: $\mathbf{x} = \sum\limits_{i=1}^{n} x^i \mathbf{f}_i$ визначимо його образ наступним чином:

$$A\mathbf{x} = A\left(\sum_{i=1}^{n} x^i \mathbf{f}_i\right) = \sum_{i=1}^{n} x^i A\mathbf{f}_i\ , \tag{6}$$

Покажемо, що відображення $A$, задане формулою (6), – лінійне. Дійсно, для довільних векторів $\mathbf{x}=\sum_{i=1}^{n}x^i\mathbf{f}_i$ та $\mathbf{y}=\sum_{i=1}^{n}y^i\mathbf{f}_i$ із простору $L_n$ матимемо:

$$A(\mathbf{x}+\mathbf{y})=A\left(\sum_{i=1}^{n}x^i\mathbf{f}_i+\sum_{i=1}^{n}y^i\mathbf{f}_i\right)=A\left(\sum_{i=1}^{n}(x^i+y^i)\mathbf{f}_i\right)=\sum_{i=1}^{n}(x^i+y^i)A\mathbf{f}_i=$$

$$=\sum_{i=1}^{n}x^iA\mathbf{f}_i+\sum_{i=1}^{n}y^iA\mathbf{f}_i=A\mathbf{x}+A\mathbf{y}$$

І аналогічно, для довільного числа $\lambda$:

$$A(\lambda\mathbf{x})=A\left(\sum_{i=1}^{n}\lambda x^i\mathbf{f}_i\right)=\sum_{i=1}^{n}(\lambda x^i)A\mathbf{f}_i=\lambda\sum_{i=1}^{n}x^iA\mathbf{f}_i=\lambda A\mathbf{x}$$

Таким чином, доведено наступну теорему.

**Теорема 5.** Існує взаємно-однозначна відповідність між множиною дійсних матриць розмірностей $m\times n$ та множиною лінійних відображень простору $L_n$ в простір $L_m$.

**Наслідок.** Якщо $\mathbf{y}=A\mathbf{x}$, де $\mathbf{x}=\sum_{i=1}^{n}x^i\mathbf{f}_i$, а $\mathbf{y}=\sum_{k=1}^{m}y^k\mathbf{g}_k$, то, знаючи матрицю $\mathbf{A}=\left(a_i^k\right),\ k=\overline{1,m},\ i=\overline{1,n}$ лінійного відображення $A$ в базисах $\{\mathbf{f}_1,\ldots,\mathbf{f}_n\}$ і $\{\mathbf{g}_1,\ldots,\mathbf{g}_m\}$, можемо записати:

$$\mathbf{y}=A\mathbf{x}=\sum_{i=1}^{n}x^iA\mathbf{f}_i=\sum_{i=1}^{n}x^i\sum_{k=1}^{m}a_i^k\mathbf{g}_k=\sum_{k=1}^{m}\sum_{i=1}^{n}a_i^kx^i\mathbf{g}_k,$$

а враховуючи однозначність розкладу вектору $\mathbf{y}$ по базису $\{\mathbf{g}_1,\ldots,\mathbf{g}_m\}$, одержимо

$$y^k=\sum_{i=1}^{n}a_i^kx^i,\ k=1,\ldots,m, \text{ або: } \quad |\mathbf{y}\rangle=\mathbf{A}\cdot|\mathbf{x}\rangle, \qquad (7)$$

тобто дія відображення $A$ на довільний вектор $\mathbf{x}$ зводиться до множення матриці $\mathbf{A}$ лінійного відображення на координатний стовпчик цього вектора.

**Зауваження 1.** Неважко зрозуміти, що і навпаки, якщо має місце формула (7), де $\mathbf{A}$ – матриця деякого лінійного відображення $A$ в тому базисі, в якому виражені координати векторів $\mathbf{x}$ та $\mathbf{y}$, то $\mathbf{y}$ є образом вектора $\mathbf{x}$ при відображенні $A$, тобто

$$\mathbf{y}=A\mathbf{x} \qquad (8)$$

**Зауваження 2.** Нагадаємо, що стовпчиками матриці $\mathbf{A}$ є координати образів базисних векторів $\{\mathbf{f}_1,\ldots,\mathbf{f}_n\}$ при відображенні $A$, а згідно з формулами (7) та (8), рядки цієї ж матриці задають перетворення координат довільного вектора $\mathbf{x}$ під дією відображення $A$.

**Теорема 6.** Ранг відображення $A$ рівний рангу його матриці $\mathbf{A}$ в деякому базисі $\{\mathbf{f}_1,\ldots,\mathbf{f}_n\}$, тобто $\operatorname{Rg}A=\operatorname{Rg}\mathbf{A}$

► Позначимо ранг матриці через $r$: $\operatorname{Rg}\mathbf{A}=r$. Нехай $i_1,i_2,\ldots,i_r$ – індекси базисних стовпчиків матриці $\mathbf{A}$. Тоді згідно з формулою (5), $A(\mathbf{f}_{i_1}),A(\mathbf{f}_{i_2}),\ldots,A(\mathbf{f}_{i_r})$ – лінійно незалежні вектори, а кожний вектор $A(\mathbf{f}_i),i=1,\ldots,n$ є їх лінійною комбінацією. Це означає, що для довільного вектора $\mathbf{x}\in L_n$ його образ $A(\mathbf{x})\in L_m$ є лінійною комбінацією саме векторів $A(\mathbf{f}_{i_1}),A(\mathbf{f}_{i_2}),\ldots,A(\mathbf{f}_{i_r})$, причому коефіцієнти цієї лінійної комбінації визначені однозначно, тобто вони утворюють базис в підпросторі $\operatorname{Im}A$. Таким чином, $\dim(\operatorname{Im}A)=r=\operatorname{Rg}A$, що і треба було довести. ◄

**Зауваження.** Ця теорема показує, що матриці лінійного відображення в різних базисах мають один і той самий ранг.

Приклад 9. Розглянемо нуль-оператор в просторі $L_n$. Його матрицею в будь-якому базисі, очевидно, буде нульова матриця розміру $n \times n$.

Приклад 10. Розглянемо тотожний оператор $E$ в просторі $L_n$. Оскільки для довільного вектора: $E\mathbf{x} = \mathbf{x}$, то його матрицею в будь-якому базисі, очевидно, буде одинична матриця $\mathbf{E}_n$ розміру $n \times n$: $\mathbf{E}_n = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$, оскільки $|\mathbf{x}\rangle = \mathbf{E}_n \cdot |\mathbf{x}\rangle$

Приклад 11. Розглянемо оператор $A$ лінійного проектування геометричних векторів простору $\mathbf{R}^3$ на площину XOY. Якщо використати за базисні вектори $\mathbf{i}, \mathbf{j}, \mathbf{k}$ ПДСК, то дію оператора $A$ на базисні вектори, очевидно, маємо визначити наступним чином: $A(\mathbf{i}) = \mathbf{i}, A(\mathbf{j}) = \mathbf{j}, A(\mathbf{k}) = \mathbf{0}$. Отже, матрицею лінійного проектування в цьому базисі буде матриця $\mathbf{A} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$, $\operatorname{Rg} A = 2$. Ядром цього оператора є вісь OZ, тобто $\dim(\operatorname{Ker} A) = 1$, як і має бути згідно з теоремою 4.

Приклад 12. Знайти матрицю цього ж оператора в базисі $\mathbf{f}_1 = \mathbf{i}, \mathbf{f}_2 = \mathbf{j}, \mathbf{f}_3 = \mathbf{i} + \mathbf{j} + \mathbf{k}$.

Розглянемо образи базисних векторів: $A(\mathbf{f}_1) = A(\mathbf{i}) = \mathbf{i} = \mathbf{f}_1$, $A(\mathbf{f}_2) = A(\mathbf{j}) = \mathbf{j} = \mathbf{f}_2$, $A(\mathbf{f}_3) = A(\mathbf{i} + \mathbf{j} + \mathbf{k}) = \mathbf{f}_1 + \mathbf{f}_2$. Таким чином матриця цього перетворення в базисі $\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3$ є такою: $\mathbf{A}' = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}$

Приклад 13. Знайти ядро та базис образу лінійного оператора $A$, який в деякому базисі простору геометричних векторів задається матрицею $\mathbf{A} = \begin{pmatrix} 1 & 1 & 1 \\ -1 & 1 & 1 \\ -1 & -1 & 1 \end{pmatrix}$.

► Вектор $\mathbf{x} \in \operatorname{Ker} A \Leftrightarrow \mathbf{A} \cdot |\mathbf{x}\rangle = |\mathbf{0}\rangle$, тобто $\mathbf{x}$ – це розв'язок однорідної системи рівнянь з матрицею $\mathbf{A}$. Знайдемо всі розв'язки цієї системи методом Гауса:

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

Оскільки $|\mathbf{A}| = 1$, дана однорідна система рівнянь має лише тривіальний розв'язок, отже $\operatorname{Ker} A = \{\mathbf{0}\}$. З теореми **4** випливає, що $\dim(\operatorname{Im} A) = 3$, а тому базисом образу заданого лінійного оператора $A$ може бути довільний базис простору, наприклад $\mathbf{i}, \mathbf{j}, \mathbf{k}$.◄

Приклад 14. Знайти базиси ядра та образу лінійного відображення $A$, який в деяких базисах відповідних просторів задається матрицею $\mathbf{A}=\begin{pmatrix}1 & 1 & -1\\ -2 & -2 & 2\\ -3 & -2 & 2\\ 4 & -1 & 1\\ 6 & 5 & -5\end{pmatrix}$.

► З розмірностей матриці $\mathbf{A}$ зрозуміло, що $A:L_3\to L_5$. Далі, нехай вектор $\mathbf{x}\in\operatorname{Ker}A$, шукаємо його як розв'язок однорідної системи рівнянь $\mathbf{A}\cdot|\mathbf{x}\rangle=|\mathbf{0}\rangle$:

$$\mathbf{A}=\begin{pmatrix}1 & 1 & -1\\ -2 & -2 & 2\\ -3 & -2 & 2\\ 4 & -1 & 1\\ 6 & 5 & -5\end{pmatrix}\to\begin{pmatrix}1 & 1 & -1\\ 0 & 0 & 0\\ 0 & 1 & -1\\ 0 & -5 & 5\\ 0 & -1 & 1\end{pmatrix}\to\begin{pmatrix}1 & 1 & -1\\ 0 & 1 & -1\\ 0 & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & 0\end{pmatrix}\to\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & -1\\ 0 & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & 0\end{pmatrix}$$

Отже, ранг системи рівний 2, тому розмірність ядра рівна $1\ (3-2=1)$. Базис ядра задається рівностями: $x_1=0; x_2=x_3$, отже, базисом ядра можна взяти, наприклад, вектор $\mathbf{f}_1=(0,1,1)^T$. З теореми **4** випливає, що ранг відображення $A$ рівний $\dim(\operatorname{Im}A)=3-1=2$. Саме таким є ранг матриці $\mathbf{A}$. За базис образу відображення $A$ тому можна взяти будь-які два лінійно незалежні стовпчики матриці $\mathbf{A}$, наприклад, $\mathbf{g}_1=(1,-2,-3,4,6)^T;\mathbf{g}_2=(1,-2,-2,-1,5)^T$ ◄

### 3. Сума та добуток лінійних відображень.

Означення 7. ***Сумою лінійних відображень*** $A,B:L_n\to L_m$ називається відображення $C=A+B:L_n\to L_m$, визначене за правилом: $C(\mathbf{x})=A(\mathbf{x})+B(\mathbf{x})$.

Легко бачити, що $C$ – лінійне відображення. Нехай $\mathbf{A}=\left(a_i^k\right),\mathbf{B}=\left(b_i^k\right)$, де $k=\overline{1,m};\ i=\overline{1,n}$ – матриці відображень $A$ і $B$ в довільних базисах $\{\mathbf{f}_1,\ldots,\mathbf{f}_n\}$ і $\{\mathbf{g}_1,\ldots,\mathbf{g}_m\}$. Знайдемо матрицю відображення $C$ в тих самих базисах $A$ і $B$: $C(\mathbf{f}_i)=A(\mathbf{f}_i)+B(\mathbf{f}_i)=\sum\limits_{k=1}^{m}\left(a_i^k+b_i^k\right)\mathbf{g}_k=\sum\limits_{k=1}^{m}c_i^k\mathbf{g}_k$, де $c_i^k$ – елементи матриці $\mathbf{C}$, яка, очевидно, є матрицею відображення $C$. Отже, $\mathbf{C}=\mathbf{A}+\mathbf{B}$, тобто сума відображень має матрицю, яка в будь-якому базисі є сумою матриць цих відображень.

Означення 8. ***Добутком лінійного відображення*** $A:L_n\to L_m$ ***на довільне число*** $\lambda$ називається відображення $(\lambda A):L_n\to L_m$, визначене за правилом: $(\lambda A)(\mathbf{x})=\lambda\cdot A(\mathbf{x})$.

Легко встановити лінійність цього відображення, також визначити, що його матрицею в деякому базисі буде матриця $\lambda\mathbf{A}$, де $\mathbf{A}$ – матриця відображення $A$ в цьому базисі.

Означення 9. ***Добутком лінійних відображень*** $B:L_n\to L_m$ і $A:L_m\to L_p$ називається відображення $C=A\cdot B\ :\ L_n\to L_p$, визначене за правилом: $C(\mathbf{x})=A\big(B(\mathbf{x})\big)$.

Неважко переконатись у лінійності відображення $A\cdot B$. Розглянемо деякі базиси $\{\mathbf{f}_1,\ldots,\mathbf{f}_n\}$, $\{\mathbf{g}_1,\ldots,\mathbf{g}_m\}$ та $\{\mathbf{q}_1,\ldots,\mathbf{q}_p\}$ у просторах $L_n$, $L_m$ та $L_p$ відповідно. Покажемо також, що матрицею $\mathbf{C}=\left(c_i^j\right)_{i=\overline{1,n}}^{j=\overline{1,p}}$ відображення $C$ у базисах $\{\mathbf{f}_1,\ldots,\mathbf{f}_n\}$ і $\{\mathbf{q}_1,\ldots,\mathbf{q}_p\}$ є добуток матриць $\mathbf{A}$ і $\mathbf{B}$ відображень $A$ і $B$ у вибраних базисах просторів $L_m$, $L_p$ та $L_n$.

Отже, розглянемо у базисі $\{\mathbf{q}_1,\ldots,\mathbf{q}_p\}$ простору $L_p$ образи базисних векторів $\{\mathbf{f}_1,\ldots,\mathbf{f}_n\}$ при відображенні $C$: $C(\mathbf{f}_i)=\sum\limits_{j=1}^{p}c_i^j\mathbf{q}_j$. З іншого боку – маємо:

$$C(\mathbf{f}_i)=(A\cdot B)(\mathbf{f}_i)=A\big(B(\mathbf{f}_i)\big)=A\left(\sum_{k=1}^{m}b_i^k\mathbf{g}_k\right)=\sum_{k=1}^{m}b_i^kA(\mathbf{g}_k)=\sum_{k=1}^{m}b_i^k\sum_{j=1}^{p}a_k^j\mathbf{q}_j=$$

$= \sum_{j=1}^{p} \sum_{k=1}^{m} a_k^j b_i^k \mathbf{q}_j$, звідки випливає, що $c_i^j = \sum_{k=1}^{m} a_k^j b_i^k$, $i = 1, \ldots, n;\ j = 1, \ldots, p$, тобто $\mathbf{C} = \mathbf{A} \cdot \mathbf{B}$.

Встановлений факт, крім того, означає, що всі властивості множення матриць переносяться на добуток лінійних відображень. Зокрема, якщо мова йде про лінійний оператор $A$, ранг якого рівний розмірності лінійного простору, то існує обернений лінійний оператор $A^{-1}$, матрицею якого є $\mathbf{A}^{-1}$ – матриця обернена до матриці $\mathbf{A}$ оператора $A$. Більше того, неважко зрозуміти, що лінійні відображення лінійного простору $L_n$ в лінійний простір $L_m$ утворюють лінійний же простір, який ізоморфний простору матриць розмірностей $m \times n$.

## 4. Перетворення матриці лінійного відображення при заміні базисів просторів.

Нехай $A : L_n \to L_m$ – лінійне відображення, а $\mathbf{A} = \left(a_j^i\right),\ i = \overline{1, m},\ j = \overline{1, n}$ – матриця цього оператора в деяких базисах $\{\mathbf{f}_1, \ldots, \mathbf{f}_n\}$ та $\{\mathbf{g}_1, \ldots, \mathbf{g}_m\}$. Розглянемо деякий інший базис $\{\mathbf{f}'_1, \ldots, \mathbf{f}'_n\}$ простору $L_n$ з матрицею $\mathbf{T} = \left(t_j^i\right),\ i, j = \overline{1, n}$ переходу до нього, а також базис $\{\mathbf{g}'_1, \ldots, \mathbf{g}'_m\}$ простору $L_m$ з матрицею переходу до нього $\mathbf{S} = \left(s_j^i\right),\ i, j = \overline{1, m}$. Знайдемо вираз для матриці $\tilde{\mathbf{A}} = \left(\tilde{a}_i^k\right),\ i = \overline{1, m},\ j = \overline{1, n}$ лінійного оператора $A$ в базисах $\{\mathbf{f}'_1, \ldots, \mathbf{f}'_n\}$ та $\{\mathbf{g}'_1, \ldots, \mathbf{g}'_m\}$. Для цього розглянемо довільний вектор $\mathbf{x} \in L_n$. Його координати відповідно у базисах $\{\mathbf{f}_1, \ldots, \mathbf{f}_n\}$ та $\{\mathbf{f}'_1, \ldots, \mathbf{f}'_n\}$ це вектор-стовпчики $\left|\mathbf{x}\right\rangle = \begin{pmatrix} x^1 \\ \vdots \\ x^n \end{pmatrix}$ та $\left|\mathbf{x}'\right\rangle = \begin{pmatrix} x'^1 \\ \vdots \\ x'^n \end{pmatrix}$. Нехай образом вектора $\mathbf{x}$ є вектор $\mathbf{y} = A\mathbf{x}$ із простору $L_m$ з координатами $\left|\mathbf{y}\right\rangle = \begin{pmatrix} y^1 \\ \vdots \\ y^m \end{pmatrix}$ та $\left|\mathbf{y}'\right\rangle = \begin{pmatrix} y'^1 \\ \vdots \\ y'^m \end{pmatrix}$ відповідно у базисах $\{\mathbf{g}_1, \ldots, \mathbf{g}_m\}$ та $\{\mathbf{g}'_1, \ldots, \mathbf{g}'_m\}$. Зауважимо, що зв'язок старих та нових координат векторів, що розглядаються, виражається наступним чином: $\left|\mathbf{x}\right\rangle = \mathbf{T}\left|\mathbf{x}'\right\rangle$, $\left|\mathbf{y}\right\rangle = \mathbf{S}\left|\mathbf{y}'\right\rangle$. Підставивши ці вирази у формулу (7), одержимо: $\mathbf{S}\left|\mathbf{y}'\right\rangle = \mathbf{AT}\left|\mathbf{x}'\right\rangle$, або $\left|\mathbf{y}'\right\rangle = \mathbf{S}^{-1}\mathbf{AT}\left|\mathbf{x}'\right\rangle$, оскільки матриця переходу завжди має обернену. Але за означенням матриці $\tilde{\mathbf{A}}$, маємо $\left|\mathbf{y}'\right\rangle = \tilde{\mathbf{A}}\left|\mathbf{x}'\right\rangle$, тобто
$$\tilde{\mathbf{A}} = \mathbf{S}^{-1}\mathbf{AT}, \tag{9}$$
оскільки матриця лінійного відображення для будь-якої пари базисів визначається однозначно. Формула (9) як раз і виражає шуканий зв'язок між матрицями лінійних відображень в різних базисах просторів $L_n$ та $L_m$.

Зауважимо, що якщо $A : L_n \to L_n$ – лінійний оператор, а $\mathbf{A} = \left(a_j^i\right),\ i, j = \overline{1, n}$ – його матриця в базисі $\{\mathbf{f}_1, \ldots, \mathbf{f}_n\}$, то при заміні базису на новий $\{\mathbf{f}'_1, \ldots, \mathbf{f}'_n\}$ з матрицею переходу $\mathbf{T} = \left(t_j^i\right),\ i, j = \overline{1, n}$ формула (9) набуває вигляду:
$$\tilde{\mathbf{A}} = \mathbf{T}^{-1} \cdot \mathbf{A} \cdot \mathbf{T}. \tag{10}$$
Таким чином, при переході до нового базису матриця лінійного оператора змінюється за формулою (10).

**<u>Означення 10</u>.** Матриці, пов’язані співвідношенням (10), де $\mathbf{T}$ – деяка невироджена матриця, називаються ***подібними*** з матрицею подібності $\mathbf{T}$.

**<u>Наслідок</u>.** Матриці лінійного оператора в різних базисах – подібні між собою.

## Лекція 16. Власні числа та власні вектори лінійних операторів.

### 1. Спектр лінійного оператора.

Нехай $L_n$ – лінійний простір, в якому діє лінійний оператор $A$: $A: L_n \to L_n$. Припустимо, що існує ***одновимірний*** підпростір $L \subseteq L_n$, інваріантний відносно оператора $A$, тобто $\forall \mathbf{x} \in L : A\mathbf{x} \in L$, а отже, знайдеться таке число $\lambda$, що $A\mathbf{x} = \lambda \mathbf{x}$. Іншими словами, дія оператора $A$ на будь-який вектор із підпростору $L$ є дуже простою – вона зводиться до множення цього вектора на деяке число. Спробуємо дослідити подібні підпростори.

**Означення 1.** Ненульовий вектор $\mathbf{x} \in L_n$ називається ***власним вектором****, який відповідає (належить)* ***власному числу (значенню)*** $\lambda$, лінійного оператора $A$, якщо
$$A\mathbf{x} = \lambda \mathbf{x} \qquad (1)$$
Повний набір власних чисел $\lambda_1, \lambda_2, \ldots, \lambda_p$ лінійного оператора з урахуванням їх кратності називається його ***спектром***, а задача по відшуканню власних чисел та власних векторів – ***спектральною задачею*** для лінійного оператора. Якщо всі власні числа мають одиничну кратність, то спектр лінійного оператора називається ***простим***.

### 2. Характеристичне рівняння та його корені.

Наступне питання, яке нам належить вирішити, – як відшукати власні числа та вектори лінійного оператора. Рівняння (1) можна записати у вигляді: $A\mathbf{x} = \lambda E\mathbf{x}$, або
$$\left(A - \lambda E\right)\mathbf{x} = \mathbf{0} \qquad (2)$$
Тобто множина власних векторів, що відповідають даному власному числу $\lambda$, є ядром оператора $A - \lambda E$. Рівняння (2) еквівалентне, як було показано раніше матричному рівнянню
$$\left(\mathbf{A} - \lambda \mathbf{E}\right) \cdot \left| \mathbf{x} \right\rangle = \mathbf{0}, \qquad (3)$$
де $\mathbf{A}$ – матриця лінійного оператора $A$ в тому ж базисі $\{\mathbf{f}_1, \ldots, \mathbf{f}_n\}$, в якому дані координати вектора $\mathbf{x}$. Рівність (3) є однорідною системою $n$ лінійних рівнянь з $n$ невідомими координатами власного вектора $\mathbf{x} \in L_n$, заданими в тому ж базисі :
$$\begin{cases} (a_1^1 - \lambda)x^1 + a_2^1 x^2 + \cdots + a_n^1 x^n = 0 \\ a_1^2 x^1 + (a_2^2 - \lambda)x^2 + \cdots + a_n^2 x^n = 0 \\ \vdots \\ a_1^n x^1 + a_2^n x^2 + \cdots + (a_n^n - \lambda)x^n = 0 \end{cases} \qquad (3^*)$$
Згідно з означенням 1, шуканий вектор $\mathbf{x}$ не рівний нулю. Такий вектор існує тоді і тільки тоді, коли визначник системи (3) рівний нулю:
$$\left|\mathbf{A} - \lambda \mathbf{E}\right| = \begin{vmatrix} a_1^1 - \lambda & a_2^1 & \cdots & a_n^1 \\ a_1^2 & a_2^2 - \lambda & \cdots & a_n^2 \\ \vdots & \vdots & & \vdots \\ a_1^n & a_2^n & \cdots & a_n^n - \lambda \end{vmatrix} = \left(-1\right)^n \lambda^n + p_1 \lambda^{n-1} + \cdots + p_n = 0 \quad (4)$$
**Означення 2.** Рівняння (4) носить назву ***характеристичного рівняння***, а його ліва частина називається ***характеристичним поліномом*** матриці $\mathbf{A}$. Неважко визначити коефіцієнти характеристичного полінома $P_n(\lambda)$, зокрема,

$p_n = \left|\mathbf{A}\right|$, $p_1 = \left(-1\right)^{n-1} \operatorname{Sp} \mathbf{A}$, де $\operatorname{Sp} \mathbf{A} = \sum_{i=1}^{n} a_i^i$ – так званий слід матриці $\mathbf{A}$.

**Зауваження 1.** Рівняння (4) є рівнянням $n$–го порядку відносно $\lambda$, а отже, не завжди має корені над полем дійсних чисел, проте над полем комплексних чисел завжли визначені $n$ його коренів.

**<u>Зауваження 2</u>.** Характеристичне рівняння та характеристичний поліном визначені через матрицю $\mathbf{A}$ лінійного оператора, вигляд якої, як нам відомо, залежить від вибору базису простору $L_n$. Наступна теорема демонструє, що насправді і характеристичне рівняння, і характеристичний поліном повністю визначаються ***лише самим оператором*** $A$.

**<u>Теорема 1</u>.** Характеристичне рівняння (4) ***не залежить від вибору базису***, в якому визначена матриця $\mathbf{A}$ лінійного оператора $A$.

► Нехай $\widetilde{\mathbf{A}}$ – матриця лінійного оператора $A$ в базисі $\{\mathbf{f}_1',\ldots,\mathbf{f}_n'\}$, перехід до якого від базису $\{\mathbf{f}_1,\ldots,\mathbf{f}_n\}$ здійснюється за допомогою матриці $\mathbf{T}$. Тоді характеристичне рівняння (4) в базисі $\{\mathbf{f}_1',\ldots,\mathbf{f}_n'\}$ запишеться у вигляді: $\left|\widetilde{\mathbf{A}}-\lambda\mathbf{E}\right|=0$. Розглянемо його ліву частину, тобто характеристичний поліном:

$$\left|\widetilde{\mathbf{A}}-\lambda\mathbf{E}\right|=\left|\mathbf{T}^{-1}\cdot\mathbf{A}\cdot\mathbf{T}-\lambda\mathbf{E}\right|=\left|\mathbf{T}^{-1}\cdot\mathbf{A}\cdot\mathbf{T}-\mathbf{T}^{-1}\cdot\lambda\mathbf{E}\cdot\mathbf{T}\right|=\left|\mathbf{T}^{-1}\cdot(\mathbf{A}-\lambda\mathbf{E})\cdot\mathbf{T}\right|=$$

$$\left|\mathbf{T}^{-1}\right|\cdot\left|\mathbf{A}-\lambda\mathbf{E}\right|\cdot\left|\mathbf{T}\right|=\left|\mathbf{T}^{-1}\cdot\mathbf{T}\right|\cdot\left|\mathbf{A}-\lambda\mathbf{E}\right|=\left|\mathbf{A}-\lambda\mathbf{E}\right| \text{ ◄}$$

### 3. Властивості спектра оператора та його власних векторів.

**<u>Означення 3</u>.** Якщо $P_n(\lambda)=(\lambda_1-\lambda)^{m_1}(\lambda_2-\lambda)^{m_2}\cdot\ldots\cdot(\lambda_p-\lambda)^{m_p}$, де $m_1+m_2+\ldots+m_p=n$, то $m_i$ називається ***алгебраїчною кратністю*** власного числа $\lambda_i$, $i=\overline{1,p}$. Максимальна кількість лінійно незалежних власних векторів, що належать власному числу $\lambda_i$, називається ***геометричною кратністю*** власного числа $\lambda_i$ і позначається $s_i$, $i=\overline{1,p}$.

Неважко показати, що $s_i\le m_i$, тобто геометрична кратність власного числа не перевищує його алгебраїчної кратності (доведіть це самостійно!).

**<u>Теорема 2</u>.** Власні вектори, що належать одному й тому самому власному числу $\lambda$ разом з вектором $\mathbf{x}=\mathbf{0}$, утворюють інваріантний відносно лінійного оператора $A$ підпростір.

► Дійсно, нехай вектори $\mathbf{x}_1$ та $\mathbf{x}_2$ належать власному числу $\lambda$. Тоді $A(\mathbf{x}_1+\mathbf{x}_2)=A\mathbf{x}_1+A\mathbf{x}_2=\lambda\mathbf{x}_1+\lambda\mathbf{x}_2=\lambda(\mathbf{x}_1+\mathbf{x}_2)$, тобто вектор $\mathbf{x}_1+\mathbf{x}_2$ є власним вектором, що належить тому самому власному числу $\lambda$, аналогічно можна переконатись, що для будь-якого числа $\mu$ вектор $\mu\mathbf{x}_1$ також є власним вектором, що належить числу $\lambda$. ◄

**<u>Теорема 3</u>.** Якщо власні вектори належать попарно різним власним числам, то вони лінійно незалежні.

► Доведення проведемо методом математичної індукції по кількості $k$ власних векторів.

Нехай $k=2$, власні вектори $\mathbf{x}_1$ та $\mathbf{x}_2$ належать відповідно власним числам $\lambda_1\neq\lambda_2$. Від супротивного припустимо, що вектори $\mathbf{x}_1$ та $\mathbf{x}_2$ лінійно залежні, наприклад, $\mathbf{x}_1=\mu\mathbf{x}_2\,(\mu\neq 0)$. Тоді $A\mathbf{x}_1=\lambda_1\mathbf{x}_1=\lambda_1\mu\mathbf{x}_2$, а з іншого боку – $A\mathbf{x}_1=A(\mu\mathbf{x}_2)=\mu\lambda_2\mathbf{x}_2$, звідки випливає, що $(\lambda_1-\lambda_2)\cdot\mu\mathbf{x}_2=\mathbf{0}$, але це неможливо при $\lambda_1\neq\lambda_2$ та $\mathbf{x}_1\neq\mathbf{0},\mathbf{x}_2\neq\mathbf{0}$. Отже, припущення про лінійну залежність векторів $\mathbf{x}_1$ та $\mathbf{x}_2$ – хибне.

Припустимо, що система із будь-яких $k$ власних векторів $\mathbf{x}_1,\mathbf{x}_2,\ldots,\mathbf{x}_k$, що належать відповідно власним числам $\lambda_1,\lambda_2,\ldots,\lambda_k$, де $\lambda_i\neq\lambda_j$ при $i\neq j$, – лінійно незалежна.

Розглянемо тепер систему із $k+1$ власних векторів, що задовольняють умовам теореми.
Від супротивного припустимо, що вектори цієї системи лінійно залежні, тобто існує нетривіальна нульова лінійна комбінація: $\alpha^1\mathbf{x}_1+\alpha^2\mathbf{x}_2+\cdots+\alpha^{k+1}\mathbf{x}_{k+1}=\mathbf{0}$ (5)

Застосуємо до даної рівності оператор $A$: $A\left(\sum_{i=1}^{k+1}\alpha^i\mathbf{x}_i\right)=\sum_{i=1}^{k+1}\alpha^i A\mathbf{x}_i=\mathbf{0}$, або $\sum_{i=1}^{k+1}\alpha^i\lambda_i\mathbf{x}_i=\mathbf{0}$.

Віднімемо від останньої рівності рівність (5), помножену на $\lambda_{k+1}\neq 0$ (якщо ж $\lambda_{k+1}=0$, то вибиремо в якості множника інше власне число та перенумеруємо відповідним чином власні вектори і числа). В результаті одержимо рівність: $\sum_{i=1}^{k}\alpha^i\left(\lambda_i-\lambda_{k+1}\right)\mathbf{x}_i=\mathbf{0}$. Оскільки система із $k$ власних векторів $\mathbf{x}_1,\mathbf{x}_2,\ldots,\mathbf{x}_k$ є лінійно незалежною за припущенням індукції, то всі коефіцієнти останньої лінійної комбінації мають бути рівними нулю, а із умов теореми випливає, що $\alpha^i=0\ \forall i=\overline{1,k}$. Але тоді з рівності (5) робимо висновок, що і $\alpha^{k+1}=0$ . Отже, лінійна комбінація (5) виявляється тривіальною, що суперечить припущенню про лінійну залежність векторів $\mathbf{x}_1,\mathbf{x}_2,\ldots,\mathbf{x}_{k+1}$. ◄

**<u>Теорема 4</u>.** Нехай лінійний оператор $A:L_n\to L_n$ має $n$ попарно різних власних чисел $\lambda_1,\lambda_2,\ldots,\lambda_n$. Тоді існує базис лінійного простору, що складається із власних векторів оператора $A$, в якому матриця оператора має діагональний вид. (Кажуть, що матриця лінійного оператора діагоналізується). Іншими словами, матриця лінійного оператора у цьому випадку подібна діагональній матриці.

► Власне, доведення є прямим наслідком теореми 3, бо $n$ попарно різним власним числам, згідно неї, відповідає $n$ лінійно незалежних власних векторів $\mathbf{f}_1,\mathbf{f}_2,\ldots,\mathbf{f}_n$, які утворюють базис простору $L_n$. І оскільки $A\mathbf{f}_k=\lambda_k\mathbf{f}_k\ \ \forall k=\overline{1,n}$, матрицею лінійного оператора $A$ в базисі $\{\mathbf{f}_1,\mathbf{f}_2,\ldots,\mathbf{f}_n\}$ буде діагональна матриця

$$\mathbf{D}=\begin{pmatrix}\lambda_1 & 0 & \cdots & 0\\ 0 & \lambda_2 & \cdots & 0\\ \vdots & \vdots & & \vdots\\ 0 & 0 & \cdots & \lambda_n\end{pmatrix}. ◄$$

**<u>Зауваження 1</u>.** Очевидно, що теорема 4 виконується, тоді і тільки тоді, коли $s_i=m_i,\ i=\overline{1,p}$.

**<u>Зауваження 2</u>.** Ще одне формулювання теореми 4 звучить так: лінійний оператор із простим спектром має матрицю, що діагоналізується.

## 4. Алгоритм відшукання власних чисел та власних векторів лінійного оператора.

Нехай $\mathbf{A}$ – матриця лінійного оператора $A$, що діє в лінійному просторі $L_n$ над полем дійсних чисел, в деякому базисі цього простору. Доведені факти дозволяють сформулювати наступний алгоритм відшукання власних чисел та власних векторів даного лінійного оператора.

1. Складаємо характеристичне рівняння $|\mathbf{A}-\lambda\mathbf{E}|=0$.

2. Шукаємо спектр оператора – дійсні розв'язки цього рівняння, якщо вони

існують. (Якщо дійсних розв'язків немає, то не існуватиме і власних векторів з дійсними координатами.)

3. Для кожного значення спектру розв'язуємо систему (3), яка визначає координати власних векторів, що належать даному власному числу.

4. Якщо кількість так визначених лінійно незалежних власних векторів рівна розмірності простору, то можна стверджувати, що матриця $\mathbf{A}$ лінійного оператора приводиться до діагонального виду в новому базисі із власних векторів (іншими словами вона подібна діагональній матриці $\mathbf{D}$, у якої на діагоналі розташовані власні числа). Матрицею переходу $\mathbf{T}$, до цього нового базису є матриця, стовпчики якої складені із координат власних векторів, розташованих у порядку, що відповідає власним числам.

Розглянемо приклади.

Приклад 1. Знайти власні числа та власні вектори лінійного оператора, який в деякому базисі задається матрицею $\mathbf{A}=\begin{pmatrix}4 & 2\\ 2 & 4\end{pmatrix}$.

► Складаємо характеристичне рівняння: $|\mathbf{A}-\lambda\mathbf{E}|=\begin{vmatrix}4-\lambda & 2\\ 2 & 4-\lambda\end{vmatrix}=(4-\lambda)^2-4=$ $\lambda^2-8\lambda+12=0$, звідки знаходимо власні числа: $\lambda_1=2$, $\lambda_2=6$.

Шукаємо власні вектори, що належать цим власним числам.

Система (3) для $\lambda_1=2$ набуває вигляду: $\begin{pmatrix}2 & 2\\ 2 & 2\end{pmatrix}\begin{pmatrix}x^1\\ x^2\end{pmatrix}=\begin{pmatrix}0\\ 0\end{pmatrix}$, звідки маємо $x^1+x^2=0$, а тому першим власним вектором є, наприклад, вектор $\mathbf{f}_1=\begin{pmatrix}1\\ -1\end{pmatrix}$.

Далі, система (3) для $\lambda_2=6$ набуває вигляду: $\begin{pmatrix}-2 & 2\\ 2 & -2\end{pmatrix}\begin{pmatrix}x^1\\ x^2\end{pmatrix}=\begin{pmatrix}0\\ 0\end{pmatrix}$, звідки маємо $x^1-x^2=0$, отже, другим власним вектором є вектор $\mathbf{f}_2=\begin{pmatrix}1\\ 1\end{pmatrix}$.◄

**Зауваження 1.** Даний лінійний оператор має 2 одновимірні інваріантні підпростір, що містять вектори відповідно виду $c_1\mathbf{f}_1$ та $c_2\mathbf{f}_2$, де $c_1, c_2$ – довільні дійсні числа.

Таким чином, доходимо висновку, що матриця даного оператора в базисі $\{\mathbf{f}_1,\mathbf{f}_2\}$ подібна матриці $\mathbf{D}=\begin{pmatrix}2 & 0\\ 0 & 6\end{pmatrix}$ з матрицею подібності $\mathbf{T}=\begin{pmatrix}1 & 1\\ -1 & 1\end{pmatrix}$ – матриця переходу від вихідного базису до базису $\{\mathbf{f}_1,\mathbf{f}_2\}$: $\mathbf{D}=\mathbf{T}\cdot\mathbf{A}\cdot\mathbf{T}^{-1}$.

**Зауваження 2.** Обов'язково переконайтесь у справедливості останньої формули – це гарантує відсутність помилок у ході розв'язування. Правда, зручніше перевіряти еквіваленту їй формулу: $\mathbf{A}=\mathbf{T}^{-1}\cdot\mathbf{D}\cdot\mathbf{T}$, оскільки множення діагональної матриці значно простіше.

Приклад 2. З'ясувати, чи приводиться до діагонального вигляду матриця $\mathbf{A}$ деякого лінійного оператора, де $\mathbf{A}=\begin{pmatrix}3 & 0 & 8\\ 3 & -1 & 6\\ -2 & 0 & -5\end{pmatrix}$. Знайти цю діагональну матрицю, якщо вона існує.

► Складаємо характеристичне рівняння:

$$|\mathbf{A}-\lambda\mathbf{E}| = \begin{vmatrix} 3-\lambda & 0 & 8 \\ 3 & -1-\lambda & 6 \\ -2 & 0 & -5-\lambda \end{vmatrix} = (3-\lambda)(1+\lambda)(5+\lambda) - 16(1+\lambda) = -(1+\lambda)^3 = 0,$$

звідки знаходимо єдине власне число: $\lambda = -1$, його алгебраїчна кратність рівна 3.

Із системи (3) знайдемо власні вектори, що належать цьому власному числу: $\begin{pmatrix} 4 & 0 & 8 \\ 3 & 0 & 6 \\ -2 & 0 & -4 \end{pmatrix}\begin{pmatrix} x^1 \\ x^2 \\ x^3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$, звідки маємо: $x^1 + 2x^3 = 0$. Ця система має два лінійно незалежних розв'язки: $\mathbf{f}_1 = \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix}$ та $\mathbf{f}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$, що означає, що геометрична кратність власного числа $\lambda = -1$ рівна 2. Вектори $\mathbf{f}_1$ та $\mathbf{f}_2$ утворюють базис двовимірного інваріантного підпростору власних векторів, але в жодному базисі матриця даного лінійного оператора не є подібною до діагональної, оскільки власних векторів просто не вистачає для вибору такого базису. Іншими словами, задана матриця не діагоналузується, або не є подібною діагональній матриці в жодному базисі. ◄

Приклад 3. Визначити власні вектори лінійного оператора, який задається матрицею $\mathbf{A} = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 2 & 0 \\ -2 & -2 & -1 \end{pmatrix}$

► Складаємо характеристичне рівняння:

$|\mathbf{A}-\lambda\mathbf{E}| = \begin{vmatrix} 1-\lambda & 2 & 0 \\ 0 & 2-\lambda & 0 \\ -2 & -2 & -1-\lambda \end{vmatrix} = -(1-\lambda)(2-\lambda)(1+\lambda) = 0$, звідки знаходимо власні числа: $\lambda_1 = -1, \lambda_2 = 1, \lambda_3 = 2$.

Оскільки маємо три різних власних числа, то згідно з теоремою **4**, матриця лінійного оператора подібна діагональній. Розв'яжемо систему (3) для кожного значення спектру.

Нехай $\lambda_1 = -1$: $(\mathbf{A} - \lambda_1\mathbf{E}) = \begin{pmatrix} 2 & 2 & 0 \\ 0 & 3 & 0 \\ -2 & -2 & 0 \end{pmatrix} \to \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$, звідки маємо: $x^1 = x^2 = 0$. Отже, система має один розв'язок, наприклад: $\mathbf{f}_1 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$.

Нехай $\lambda_2 = 1$: $(\mathbf{A} - \lambda_2\mathbf{E}) = \begin{pmatrix} 0 & 2 & 0 \\ 0 & 1 & 0 \\ -2 & -2 & -2 \end{pmatrix} \to \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 1 \end{pmatrix}$, звідки маємо:

$x^1 = -x^3, x^2 = 0$. Значенню $\lambda_2 = 1$ відповідає один вектор, наприклад: $\mathbf{f}_2 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$.

Нехай $\lambda_3 = 2$: $(\mathbf{A} - \lambda_3 \mathbf{E}) = \begin{pmatrix} -1 & 2 & 0 \\ 0 & 0 & 0 \\ -2 & -2 & -3 \end{pmatrix} \to \begin{pmatrix} 1 & -2 & 0 \\ 0 & -6 & -3 \\ 0 & 0 & 0 \end{pmatrix} \to \begin{pmatrix} 1 & -2 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 0 \end{pmatrix}$,

звідки маємо: $x^1 = 2x^2, x^3 = -2x^2$.

Одержимо третій власний вектор: $\mathbf{f}_3 = \begin{pmatrix} 2 \\ 1 \\ -2 \end{pmatrix}$.

Отже, в базисі з власних векторів $\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3$, матриця заданого лінійного оператора – діагональна: $\mathbf{D} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$, матрицею переходу до цього власного базису є $\mathbf{T} = \begin{pmatrix} 0 & 1 & 2 \\ 0 & 0 & 1 \\ 1 & -1 & -2 \end{pmatrix}$. Безпосередньою перевіркою можна переконатись, що $\mathbf{D} = \mathbf{T}^{-1}\mathbf{AT}$, або (що простіше перевірити) $\mathbf{A} = \mathbf{TDT}^{-1}$.◄

**5. Деякі властивості характеристичного полінома. Теорема Гамільтона-Келі.**

Розглянемо деякий поліном степені $m$ з дійсними коефіцієнтами: $P_m(x) = a_m x^m + a_{m-1}x^{m-1} + \dots + a_1 x + a_0$.

Оскільки визначені операції з матрицями, можна розглядати значення такого поліному від деякої квадратної матриці $\mathbf{A}$ довільного порядку $n$, так званий матричний поліном:

$$P_m(\mathbf{A}) = a_m \cdot \mathbf{A}^m + a_{m-1} \cdot \mathbf{A}^{m-1} + \dots + a_1 \cdot \mathbf{A} + a_0 \cdot \mathbf{E}_n$$

<u>Приклад 4</u>. Нехай $\mathbf{A} = \mathbf{D}$, де $\mathbf{D} = \begin{pmatrix} d_1^1 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & d_n^n \end{pmatrix}$ – діагональна матриця, тоді $P_m(\mathbf{D}) = \begin{pmatrix} P_m(d_1^1) & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & P_m(d_n^n) \end{pmatrix}$ – теж діагональна матриця.

<u>**Зауваження.**</u> Визначають і більш складні функції від матриць, наприклад, матричну експоненту $e^{\mathbf{A}}$, або $\sin \mathbf{A}$ через розклад в ряд Тейлора: $e^{\mathbf{A}} = \mathbf{E} + \mathbf{A} + \frac{\mathbf{A}^2}{2!} + \dots + \frac{\mathbf{A}^n}{n!} + \cdots$ . Залишимо поза увагою питання збіжності таких матричних рядів.

Відомо, що квадратні матриці порядку $n$ утворюють лінійний простір розмірності $n^2$, тобто будь-які $n^2 + 1$ квадратних матриці порядку $n$ є лінійно залежними, а тому утворюють нетривіальну лінійну комбінацію рівну нулю. Зокрема, можна скласти таку лінійну комбінацію з матриць $\mathbf{E}, \mathbf{A}, \mathbf{A}^2, \dots, \mathbf{A}^{n^2}$. Тобто для довільної матриці $\mathbf{A}$ лінійного оператора $A$ завжди існує поліном порядку не вищого за $n^2$, такий що його значенням від матриці $\mathbf{A}$ є нуль-матриця. Такий поліном $P(x)$ називається *анулюючим* для матриці $\mathbf{A}$, а сама матриця є матричним коренем рівняння $P(x) = 0$.

**Означення 4.** Поліном із старшим коефіцієнтом рівним 1 *мінімального порядку* серед анулюючих поліномів називається *мінімальним поліномом* і позначається $\mu_{\mathbf{A}}(x)$.

**Зауваження.** Мінімальний поліном визначається однозначно, бо якби існувало два різних анулюючих поліноми однакового порядку, то їх різниця також була б анулюючим поліномом, проте, ще меншого порядку.

**Теорема Гамільтона-Келі.** Характеристичний поліном лінійного оператора $A$ є анулюючим, або іншими словами, будь-яка матриця є матричним коренем свого характеристичного рівняння.

► **Лема.** Нехай поліном $P_m(\lambda)=a_m\lambda^m+a_{m-1}\lambda^{m-1}+\cdots+a_1\lambda+a_0$ та квадратна матриця $\mathbf{A}$ порядку $n$ пов'язані співвідношенням: $P_m(\lambda)\cdot\mathbf{E}=(\mathbf{A}-\lambda\mathbf{E})\mathbf{Q}(\lambda)$, де $\mathbf{Q}(\lambda)=\mathbf{B}_{m-1}\lambda^{m-1}+\cdots+\mathbf{B}_1\lambda+\mathbf{B}_0$ – поліном від $\lambda$ з матричними коефіцієнтами ($\mathbf{B}_{m-1},\ldots,\mathbf{B}_1,\mathbf{B}_0$ – матриці порядку $n$ ). Тоді $P(\mathbf{A})=\mathbf{0}$.

► Зауважимо, що дана лема просто узагальнює теорему Безу. Дійсно, розглянемо

$$(\mathbf{A}-\lambda\mathbf{E})\mathbf{Q}(\lambda)=(\mathbf{A}-\lambda\mathbf{E})\left(\mathbf{B}_{m-1}\lambda^{m-1}+\cdots+\mathbf{B}_1\lambda+\mathbf{B}_0\right)=\mathbf{AB}_0+(\mathbf{AB}_1-\mathbf{B}_0)\lambda+(\mathbf{AB}_2-\mathbf{B}_1)\lambda^2+\cdots-\mathbf{B}_{m-1}\lambda^m$$

Порівняємо матричні коефіцієнти цього полінома при степенях $\lambda$ з відповідними коефіцієнтами $P_m(\lambda)\cdot\mathbf{E}$:

$\mathbf{AB}_0=a_0\mathbf{E};\ \mathbf{AB}_1-\mathbf{B}_0=a_1\mathbf{E};\ \mathbf{AB}_2-\mathbf{B}_1=a_2\mathbf{E};\ -\mathbf{B}_{m-1}=a_m\mathbf{E}$. Тепер помножимо кожну рівність зліва відповідно на $\mathbf{E},\mathbf{A},\mathbf{A}^2,\ldots,\mathbf{A}^m$ та результати множень додамо. Одержимо: $\mathbf{0}=a_0\mathbf{E}+a_1\mathbf{A}+\cdots+a_m\mathbf{A}^m=P_m(\mathbf{A})$ ◄

Повернемось до доведення теореми. Розглянемо обернену матрицю $(\mathbf{A}-\lambda\mathbf{E})^{-1}$. Вона визначається через мінори $(n-1)$-го порядку матриці $(\mathbf{A}-\lambda\mathbf{E})$, тому $(\mathbf{A}-\lambda\mathbf{E})^{-1}=\dfrac{\mathbf{Q}(\lambda)}{|\mathbf{A}-\lambda\mathbf{E}|}$, де $\mathbf{Q}(\lambda)$ – матриця порядку $n$ з елементами, які є поліномами порядку не вищого за $n-1$. Отже, $\mathbf{E}=(\mathbf{A}-\lambda\mathbf{E})\cdot(\mathbf{A}-\lambda\mathbf{E})^{-1}=(\mathbf{A}-\lambda\mathbf{E})\cdot\dfrac{\mathbf{Q}(\lambda)}{|\mathbf{A}-\lambda\mathbf{E}|}=(\mathbf{A}-\lambda\mathbf{E})\cdot\dfrac{\mathbf{Q}(\lambda)}{P_n(\lambda)}$, де $P_n(\lambda)$ – характеристичний поліном матриці $\mathbf{A}$. Таким чином, $P_n(\lambda)\mathbf{E}=(\mathbf{A}-\lambda\mathbf{E})\cdot\mathbf{Q}(\lambda)$, тобто виконані умови леми, тому $P_n(\mathbf{A})=\mathbf{0}$.◄

## Лекція 17 Лінійні оператори в евклідовому просторі.

### 1. Оператор, спряжений даному.

Зрозуміло, що все сказане вище щодо операторів у лінійному просторі справедливо і для евклідового простору. Проте наявність скалярного добутку дає можливість виділити ряд важливих видів лінійних операторів. Розглянемо деякий евклідів простір $E_n$ розмірності $n$, в якому діє лінійний оператор $A$: $A:E_n\to E_n$.

**Означення 1.** Оператор $A^*:E_n\to E_n$ називається ***спряженим*** даному оператору $A$, якщо $\forall\mathbf{x},\mathbf{y}\in E_n$:

$$A\mathbf{x}\cdot\mathbf{y}=\mathbf{x}\cdot A^*\mathbf{y} \tag{1}$$

Неважко переконатись, що так визначений оператор є лінійним, тобто мають місце рівності $\mathbf{x}\cdot A^*(\lambda\mathbf{y})=\lambda\mathbf{x}\cdot A^*\mathbf{y}$ $\forall\mathbf{x},\mathbf{y}\in E_n$ $\forall\lambda\in R$ та $\forall\mathbf{x},\mathbf{y}_1,\mathbf{y}_2\in E_n$ $\mathbf{x}\cdot A^*(\mathbf{y}_1+\mathbf{y}_2)=\mathbf{x}\cdot A^*\mathbf{y}_1+\mathbf{x}\cdot A^*\mathbf{y}_2$ .

Приклад 1. Нехай $E_n$ – простір геометричних векторів $\mathbf{R}^3$ зі звичайним скалярним добутком. Визначимо оператор $A$ в $\mathbf{R}^3$ за правилом: $A\mathbf{x}=\mathbf{a}\times\mathbf{x}$ (тобто векторний добуток деякого фіксованого вектора $\mathbf{a}$) . Із-за властивостей векторного множення лінійність цього оператора очевидна. Визначимо оператор, спряжений до оператора $A$: $A\mathbf{x}\cdot\mathbf{y}=(\mathbf{a}\times\mathbf{x})\cdot\mathbf{y}=\mathbf{axy}=-\mathbf{xay}=-\mathbf{x}\cdot(\mathbf{a}\times\mathbf{y})=\mathbf{x}\cdot(-\mathbf{a}\times\mathbf{y})=\mathbf{x}\cdot A^*\mathbf{y}$. Таким чином, $A^*=-A$.

**Теорема 1.** Матриця спряженого оператора в будь-якому базисі подібна ***транспонованій матриці даного лінійного оператора*** з матрицею подібності $\mathbf{G}$ – матриця Грама даного базису.

► Нехай $\{\mathbf{f}_1,\mathbf{f}_2,\ldots,\mathbf{f}_n\}$ – довільний базис евклідового простору, а $\mathbf{G}$ – його

матриця Грама. Позначимо через $\mathbf{A}$ та $\mathbf{A}^*$ відповідно матриці операторів $A$ та $A^*$ у цьому базисі. Розглянемо довільні вектори $\mathbf{x}, \mathbf{y} \in E_n$ з координатними стовпчиками $|\mathbf{x}\rangle$ та $|\mathbf{y}\rangle$ у вибраному базисі $\{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_n\}$. Тоді, згідно формулі (4) з лекції 19, маємо:
$$A\mathbf{x} \cdot \mathbf{y} = |A\mathbf{x}\rangle^T \mathbf{G}|\mathbf{y}\rangle = (\mathbf{A} \cdot |\mathbf{x}\rangle)^T \mathbf{G}|\mathbf{y}\rangle = |\mathbf{x}\rangle^T \mathbf{A}^T \mathbf{G}|\mathbf{y}\rangle \qquad (2)$$

З іншого боку, $$A\mathbf{x} \cdot \mathbf{y} = \mathbf{x} \cdot A^*\mathbf{y} = |\mathbf{x}\rangle^T \mathbf{G}\left|A^*\mathbf{y}\right\rangle = |\mathbf{x}\rangle^T \mathbf{G}\mathbf{A}^*|\mathbf{y}\rangle \qquad (3)$$

Порівнюючи формули (2) та (3), одержимо: $|\mathbf{x}\rangle^T \left(\mathbf{A}^T\mathbf{G} - \mathbf{G}\mathbf{A}^*\right)|\mathbf{y}\rangle = 0$. Оскільки дана рівність виконана для довільних $\mathbf{x}, \mathbf{y} \in E_n$, робимо висновок, що $\mathbf{A}^T\mathbf{G} - \mathbf{G}\mathbf{A}^* = \mathbf{0}$ (покласти $|\mathbf{x}\rangle = |\mathbf{E}_i\rangle$, а $|\mathbf{y}\rangle = |\mathbf{E}_j\rangle$ та пригадати приклад 5 з лекції 11 – одержимо рівність елементів $\left(\mathbf{A}^T\mathbf{G} - \mathbf{G}\mathbf{A}^*\right)^i_j = 0, i, j = \overline{1, n}$), тобто $\mathbf{A}^* = \mathbf{G}^{-1}\mathbf{A}^T\mathbf{G}$ ◄ (4)

**Наслідок.** В ортонормованому базисі $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ формула (4) набуває вигляду
$$\mathbf{A}^* = \mathbf{A}^T \qquad (5)$$

**Теорема 2.** Кожний лінійний оператор в евклідовому просторі має спряжений, причому єдиний.

► Позначимо через $\mathbf{A}$ – матрицю довільного лінійного оператора $A$ простору $E_n$ в деякому ортонормованому базисі $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$. Розглянемо матрицю $\mathbf{A}^T$. Їй відповідає деякий лінійний оператор $\tilde{A}$, що діє за правилом: $\tilde{A}\mathbf{x} = \mathbf{A}^T|\mathbf{x}\rangle \ \forall \mathbf{x} \in E_n$. Нехай $\mathbf{x}, \mathbf{y} \in E_n$ – довільні вектори з координатами $|\mathbf{x}\rangle$ та $|\mathbf{y}\rangle$ в ортонормованому базисі $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$. Для них виконується рівність: $A\mathbf{x} \cdot \mathbf{y} = |A\mathbf{x}\rangle^T |\mathbf{y}\rangle = (\mathbf{A}|\mathbf{x}\rangle)^T |\mathbf{y}\rangle = |\mathbf{x}\rangle^T \mathbf{A}^T |\mathbf{y}\rangle = \mathbf{x} \cdot \tilde{A}\mathbf{y}$. Отже, $\tilde{A} = A^*$, тобто спряжений оператор визначений для довільного лінійного оператора. Покажемо, що спряжений оператор визначається однозначно. Від супротивного припустимо, що деякий лінійний оператор $A$ має два різних спряжених оператора $A_1^*, A_2^*$. В деякому ортонормованому базисі їм відповідають дві різні матриці $\mathbf{A}_1^*, \mathbf{A}_2^*$. Проте, з наслідку теореми **2**: $\mathbf{A}_1^* = \mathbf{A}^T$ та $\mathbf{A}_2^* = \mathbf{A}^T$, тобто $\mathbf{A}_1^* = \mathbf{A}_2^*$. Одержана суперечність означає, що спряжений оператор – єдиний ◄

**Наслідок.** Оператором, спряженим до спряженого, є даний: $\left(A^*\right)^* = A$ (6)

► Дана рівність випливає з того, що $\left(\mathbf{A}^T\right)^T = \mathbf{A}$. ◄

Завдання для самостійної роботи. Довести, що:

1. $(A+B)^* = A^* + B^*$ 2. $(A \cdot B)^* = B^* \cdot A^*$ 3. $(\lambda A)^* = \lambda A^*$ 4. $E^* = E;\ 0^* = 0$

## 2. Самоспряжений оператор.

**Означення 2.** Лінійний оператор $A: E_n \to E_n$ називається ***самоспряженим***, якщо $\forall \mathbf{x}, \mathbf{y} \in E_n$:
$$A\mathbf{x} \cdot \mathbf{y} = \mathbf{x} \cdot A\mathbf{y} \qquad (7)$$

Із формули (5) безпосередньо випливає, що оператор $A$ є самоспряженим тоді і тільки тоді, коли його матриця $\mathbf{A}$ в будь-якому ортонормованому базисі є симетричною: $\mathbf{A} = \mathbf{A}^T$.

Завдання для самостійної роботи. Нехай $A$ і $B$ – два самоспряжених оператори. Довести, що:

1. $\forall \lambda, \mu \in \mathbf{R}: \lambda A + \mu B$ – самоспряжений оператор.
2. Необхідною та достатньою умовою, щоб оператор $A \cdot B$ був самоспряженим, є комутативність операторів: $A \cdot B = B \cdot A$
3. Якщо $A$– довільний оператор, то $A \cdot A^*$ та $A^* \cdot A$ – самоспряжені оператори.

**Зауваження.** Операція переходу до спряженого оператора до певної міри подібна переходу до спряженого числа в комплексному просторі (більш наочно ця аналогія проявиться в унітарному просторі), а серед комплексних чисел виділяються дійсні числа – їх аналогом в множині лінійних операторів є самоспряжені оператори.

**Теорема 3.** Всі власні значення самоспряженого оператора – дійсні.

► Припустимо супротивне. Нехай деяке власне число $\lambda_0$ самоспряженого оператора $A$ – комплексне. Позначимо через $\mathbf{A}$ матрицю цього оператора в деякому ортонормованому базисі. Власний вектор $\mathbf{x}_0$, що належить власному числу $\lambda_0$, – це нетривіальний розв'язок однорідної системи $(\mathbf{A}-\lambda_0\mathbf{E})|\mathbf{x}_0\rangle=\mathbf{0}$, тобто $\mathbf{x}_0$ – взагалі кажучи, вектор з комплексними координатами. Помножимо рівність $\mathbf{A}|\mathbf{x}_0\rangle=\lambda_0|\mathbf{x}_0\rangle$ зліва на комплексно спряжений вектор $\overline{|\mathbf{x}_0\rangle}^T$ : $\overline{|\mathbf{x}_0\rangle}^T\mathbf{A}|\mathbf{x}_0\rangle=\lambda_0\overline{|\mathbf{x}_0\rangle}^T|\mathbf{x}_0\rangle=\lambda_0\sum_{i=1}^{n}\left|x_0^i\right|^2$ Права частина цього виразу – комплексне число, а отже, $\omega=\overline{|\mathbf{x}_0\rangle}^T\mathbf{A}|\mathbf{x}_0\rangle$ – теж комплексне число. Виходячи з очевидної рівності $\omega=\omega^T$, маємо:

$$\omega=\left(\overline{|\mathbf{x}_0\rangle}^T\mathbf{A}|\mathbf{x}_0\rangle\right)^T=|\mathbf{x}_0\rangle^T\mathbf{A}^T\overline{|\mathbf{x}_0\rangle}=|\mathbf{x}_0\rangle^T\mathbf{A}\overline{|\mathbf{x}_0\rangle}$$

Ця рівність означає, що $\omega=\overline{\omega}$, тобто $\omega$ – дійсне число, що неможливо. ◄

**Теорема 4.** Власні вектори самоспряженого оператора , які належать попарно різним власним числам, – ортогональні.

► Нехай власні вектори самоспряженого оператора $A$ $\mathbf{x}_1$ та $\mathbf{x}_2$ належать відповідно власним числам $\lambda_1\neq\lambda_2$. Розглянемо скалярний добуток $A\mathbf{x}_1\cdot\mathbf{x}_2=\lambda_1\mathbf{x}_1\cdot\mathbf{x}_2$. З іншого боку, $A\mathbf{x}_1\cdot\mathbf{x}_2=\mathbf{x}_1\cdot A\mathbf{x}_2=\lambda_2\mathbf{x}_1\cdot\mathbf{x}_2$. Тобто $(\lambda_1-\lambda_2)\mathbf{x}_1\cdot\mathbf{x}_2=0$, отже, вектори $\mathbf{x}_1$ та $\mathbf{x}_2$ – ортогональні. ◄

**Теорема 5. (про існування ортонормованого базису евклідового простору з власних векторів самоспряженого оператора).**
В евклідовому просторі завжди існує ортонормований базис із власних векторів будь-якого самоспряженого оператора.

► **Лема.** Нехай підпростір $E\subseteq E_n$ – інваріантний відносно самоспряженого оператора $A:E_n\to E_n$. Тоді підпростір $E^{\perp}$ також інваріантний відносно оператора $A$.

► Розглянемо довільний вектор $\mathbf{x}\in E$, тоді образ $A\mathbf{x}\in E$. Нехай тепер $\mathbf{y}$ – довільний вектор із $E^{\perp}$. Отже, $0=A\mathbf{x}\cdot\mathbf{y}=\mathbf{x}\cdot A\mathbf{y}$, тобто $A\mathbf{y}\in E^{\perp}$. ◄

Для доведення теореми використаємо індукцію по $n$ – розмірність $E_n$.

1. Нехай $A:E_1\to E_1$ – самоспряжений оператор. Згідно теоремі **3**, власне число цього оператора – дійсне, а отже, існує власний вектор, який йому належить. Із-за розмірності простора $E_1$ кожний ненульовий вектор буде власним. Виберемо за базисний одиничний вектор.
2. Припустимо, що в просторі $E_n$, де $n=k>1$ існує вказаний ортонормований базис.
3. Покажемо, що такий базис існує і в $E_n$, де $n=k+1$. Дійсно, характеристичне рівняння оператора $A$ має принаймні 1 дійсний корінь (можливо, кратний), отже, існує принаймні одновимірний підпростір $E_n$ інваріантний відносно оператора $A$. Позначимо цей підпростір $E_1$. Нехай $\mathbf{l}$ – орт в $E_1$. Позначимо через $\tilde{E}$ ортогональне доповнення $E_1^{\perp}$. Цей підпростір інваріантний відносно оператора $A$ згідно лемі, причому $\dim(\tilde{E})=k$. Розглянемо звуження $\tilde{A}$ оператора $A$ на підпростір $\tilde{E}$: $\tilde{A}=A\big|_{\tilde{E}}$, тобто $\tilde{A}\mathbf{x}=A\mathbf{x}$, якщо $\mathbf{x}\in\tilde{E}$. Очевидно, що оператор $\tilde{A}$ – самоспряжений оператор на $\tilde{E}$. Крім того, якщо власний вектор оператора $A$ належить підпростору $\tilde{E}$, то він же є власним вектором оператора $\tilde{A}$, що належить тому самому ж власному числу. Згідно припущенню індукції, в підпросторі $\tilde{E}$ існує ортонормований базис $\{\mathbf{e}_1,\mathbf{e}_2,\ldots,\mathbf{e}_k\}$ з власних векторів оператора $\tilde{A}$, а отже і $A$. Тоді $\{\mathbf{e}_1,\mathbf{e}_2,\ldots,\mathbf{e}_k,\mathbf{l}\}$ – шуканий ортонормований базис в просторі $E_n$. ◄

**Теорема 6. (про зведення симетричної матриці до діагонального вигляду).** Симетрична матриця завжди подібна деякій діагональній матриці з ортогональною матрицею подібності.

► Нехай $\mathbf{A}$ – симетрична матриця. Тоді в деякому ортонормованому базисі простору вона задає самоспряжений оператор $A$. Згідно з теоремою **5**, у нього існує $n$ взаємно ортогональних власних векторів $\mathbf{f}_1,\mathbf{f}_2,\ldots,\mathbf{f}_n\in E_n$. В базисі з цих власних векторів матриця $\mathbf{A}$ набуває діагонального вигляду: $\mathbf{D}=\mathbf{Q}^T\mathbf{AQ}$ (діагональними елементами матриці $\mathbf{D}$ є власні числа матриці $\mathbf{A}$). Причому матриця $\mathbf{Q}$ – ортогональна як матриця переходу між двома ортонормованими базисами. ◄

Приклад 2. Знайти ортонормований базис із власних векторів та вказати діагональний вид матриці самоспряженого оператора, заданого в деякому ортонормованому базисі матрицею $\mathbf{A}=\begin{pmatrix}0&0&1\\0&1&0\\1&0&0\end{pmatrix}$.

► Складаємо характеристичне рівняння:

$|\mathbf{A}-\lambda\mathbf{E}|=\begin{vmatrix}-\lambda&0&1\\0&1-\lambda&0\\1&0&-\lambda\end{vmatrix}=(1-\lambda)\lambda^2-(1-\lambda)=-(1-\lambda)^2(1+\lambda)=0$, звідки знаходимо власні числа: $\lambda_1=-1$, $\lambda_2=1$.

Шукаємо власні вектори, що належать цим власним числам.

$\lambda_1=-1$: із системи $\begin{pmatrix}1&0&1\\0&2&0\\1&0&1\end{pmatrix}\begin{pmatrix}x^1\\x^2\\x^3\end{pmatrix}=\begin{pmatrix}0\\0\\0\end{pmatrix}$ маємо $x^1+x^3=0;\ x^2=0$. Отже, нехай $\mathbf{f}_1=\begin{pmatrix}1\\0\\-1\end{pmatrix}$. Відповідним ортом буде власний вектор $\mathbf{e}_1=\begin{pmatrix}\frac{1}{\sqrt{2}}\\0\\-\frac{1}{\sqrt{2}}\end{pmatrix}$.

$\lambda_2=1$: із системи $\begin{pmatrix}-1&0&1\\0&0&0\\1&0&-1\end{pmatrix}\begin{pmatrix}x^1\\x^2\\x^3\end{pmatrix}=\begin{pmatrix}0\\0\\0\end{pmatrix}$ маємо $x^1=x^3$. Ця система рангу 1 дозволяє визначити два власних вектори. Один з них виберемо довільним ортом, а другий шукатимемо ортогональним до першого. Отже, нехай $\mathbf{f}_2=\begin{pmatrix}1\\0\\1\end{pmatrix}$. Відповідним ортом буде власний вектор $\mathbf{e}_2=\begin{pmatrix}\frac{1}{\sqrt{2}}\\0\\\frac{1}{\sqrt{2}}\end{pmatrix}$. Шукатимемо тепер вектор $\mathbf{f}_3=\begin{pmatrix}\alpha\\\beta\\\alpha\end{pmatrix}$ такий, що $\mathbf{f}_2\cdot\mathbf{f}_3=0$: $\alpha+\alpha=0$, тобто $\mathbf{f}_3=\begin{pmatrix}0\\1\\0\end{pmatrix}=\mathbf{e}_3$. Таким чином, $\mathbf{D}=\mathbf{Q}^T\mathbf{AQ}$, де $\mathbf{D}=\begin{pmatrix}-1&0&0\\0&1&0\\0&0&1\end{pmatrix}$, а матриця $\mathbf{Q}=\begin{pmatrix}\frac{1}{\sqrt{2}}&\frac{1}{\sqrt{2}}&0\\0&0&1\\-\frac{1}{\sqrt{2}}&\frac{1}{\sqrt{2}}&0\end{pmatrix}$ – ортогональна, в чому легко переконатись безпосередньо. ◄

### 3. Ортогональний оператор

Ще один важливий клас лінійних операторів евклідового простору – так звані ортогональні оператори.

**Означення 3.** Лінійний оператор $A: E_n \to E_n$ називається ***ортогональним***, якщо він зберігає скалярний добуток векторів, тобто $\forall \mathbf{x}, \mathbf{y} \in E_n$:

$$A\mathbf{x} \cdot A\mathbf{y} = \mathbf{x} \cdot \mathbf{y} \tag{8}$$

Очевидно , що при ортогональному перетворенні зберігаються і довжини векторів, а отже, і кути між ними. Можна сказати, що ортогональний оператор зберігає «геометрію» простору. Тому ортогональний оператор переведе ортонормований базис знову в ортонормований. Неважко зрозуміти, що в ортонормованому базисі матрицею ортогонального оператора буде ортогональна матриця. Дійсно, з рівності (8) випливає, що $\forall \mathbf{x}, \mathbf{y} \in E_n : A\mathbf{x} \cdot A\mathbf{y} = \mathbf{x} \cdot A^* A\mathbf{y} = \mathbf{x} \cdot \mathbf{y}$, тобто ортогональне перетворення має властивість, що $A^* A = E$, а оскільки добутку лінійних операторів відповідає добуток їх матриць, то в ортонормованому базисі (і лише в ньому!) матриця ортогонального оператора має властивість: $\mathbf{A}^T \mathbf{A} = \mathbf{E}$, тобто є ортогональною матрицею. Звідси також випливає, що визначник ортогональної матриці по модулю рівний 1: $\left|\mathbf{A}^T \mathbf{A}\right| = \left|\mathbf{A}^T\right| \cdot |\mathbf{A}| = |\mathbf{A}|^2 = |\mathbf{E}| = 1$, а з рівності (8) для власного числа $\lambda$, якому належить власний вектор $\mathbf{x}$, випливає, що: $\mathbf{x} \cdot \mathbf{x} = A\mathbf{x} \cdot A\mathbf{x} = \lambda\mathbf{x} \cdot \lambda\mathbf{x} = \lambda^2 \mathbf{x}$. Це означає, що власні числа ортогонального оператора, якщо вони дійсні, рівні по модулю 1.

Дослідимо більш уважно структуру ортогонального оператора в просторі $\mathrm{R}^2$. Нехай в ортонормованому базисі $\{\mathbf{e}_1, \mathbf{e}_2\}$ матрицею лінійного оператора є матриця $\mathbf{A} = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$. Припустимо спочатку, що ортогональний оператор ***власний***, тобто $|\mathbf{A}| = 1$. Тоді $\mathbf{A}^{-1} = \begin{pmatrix} \delta & -\beta \\ -\gamma & \alpha \end{pmatrix}$. З іншого боку, $\mathbf{A}^{-1} = \mathbf{A}^T = \begin{pmatrix} \alpha & \gamma \\ \beta & \delta \end{pmatrix}$, отже, $\beta = -\gamma,\ \alpha = \delta$, а з умови $\alpha^2 + \beta^2 = 1$ випливає, що $\mathbf{A} = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}$. Це означає, що власний ортогональний оператор в двовимірному просторі задає поворот на деякий кут $\varphi$.

Розглянемо тепер ***невласний*** ортогональний оператор, тобто такий, що $|\mathbf{A}| = -1$. З аналогічних міркувань випливає, що $\beta = \gamma,\ \alpha = -\delta$. В цьому випадку характеристичне рівняння оператора має вигляд: $|\mathbf{A} - \lambda\mathbf{E}| = \left|\begin{pmatrix} \alpha - \lambda & \beta \\ \beta & -\alpha - \lambda \end{pmatrix}\right| = \lambda^2 - (\alpha^2 + \beta^2) = \lambda^2 - 1$, звідки маємо можливий вигляд матриці невласного ортогонального оператору $\mathbf{A} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ або $\mathbf{A} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$ (дзеркальні повороти відносно однієї з осей).

Можна показати, що загалом матриця довільного ортогонального оператора в $E_n$ є блочно діагональною матрицею, у якої на діагоналі знаходяться $\pm 1$ або блоки виду $\begin{pmatrix} \cos\varphi_j & -\sin\varphi_j \\ \sin\varphi_j & \cos\varphi_j \end{pmatrix}$.

### 4. Унітарний простір та лінійні оператори в ньому.

**Означення 4.** Будемо говорити що у комплексному векторному просторі $L$ задана операція ***скалярного добутку*** векторів, якщо будь-якій парі векторів $\mathbf{x}, \mathbf{y} \in L$ цього простору ставиться у відповідність комплексне число $\mathbf{x} \cdot \mathbf{y}$ таким чином, що виконані наступні аксіоми:

1. $\mathbf{x} \cdot \mathbf{y} = \overline{\mathbf{y} \cdot \mathbf{x}} \quad \forall \mathbf{x}, \mathbf{y} \in L$;
2. $(\lambda\mathbf{x}) \cdot \mathbf{y} = \lambda(\mathbf{x} \cdot \mathbf{y}) \quad \forall \mathbf{x}, \mathbf{y} \in L \forall \lambda \in \mathbf{C}$;
3. $(\mathbf{x}_1 + \mathbf{x}_2) \cdot \mathbf{y} = \mathbf{x}_1 \cdot \mathbf{y} + \mathbf{x}_2 \cdot \mathbf{y} \quad \forall \mathbf{x}_1, \mathbf{x}_2, \mathbf{y} \in L$ (дистрибутивність скалярного добутку);
4. $\mathbf{x} \cdot \mathbf{x} \ge 0 \quad \forall \mathbf{x} \in L$, причому $\mathbf{x} \cdot \mathbf{x} = 0 \iff \mathbf{x} = \mathbf{0}$

**Означення 5.** Комплексний векторний простір $L_n$ зі скалярним добутком називається ***унітарним*** простором і позначається $U_n$.

**Зауваження 1.** Зверніть увагу, четверта аксіома гарантує, що скалярний добуток комплексного вектора на себе – дійсне число, а отже, поняття довжини вектора, кута між векторами залишаються таким самими як у евклідовому просторі. Так само не змінюються визначення ортогонального та ортонормованого базисів.

**Зауваження 2.** Наслідком першої та другої аксіом є формула: $\mathbf{x}\cdot\lambda\mathbf{y}=\overline{\lambda}(\mathbf{x}\cdot\mathbf{y})\quad\forall\mathbf{x},\mathbf{y}\in L\,\forall\lambda\in\mathbf{C}$

Розглянемо матрицю Грама деякого базису $\{\mathbf{f}_1,\mathbf{f}_2,\ldots,\mathbf{f}_n\}$ унітарного простору $U_n$: $\mathbf{G}=(g_{ij})\ i,j=\overline{1,n}$, де $g_{ij}=\mathbf{f}_i\cdot\mathbf{f}_j$. Оскільки $\mathbf{f}_i\cdot\mathbf{f}_j=\overline{\mathbf{f}_j\cdot\mathbf{f}_i}$, то $g_{ij}=\overline{g_{ji}}\ \forall i,j=\overline{1,n}$.

**Означення 6.** Квадратна комплексна матриця $\mathbf{A}=(a_j^i)\ i,j=\overline{1,n}$, елементи якої мають властивість $a_j^i=\overline{a_i^j}$, називається ***ермітовою*** матрицею. Тобто для ермітової матриці справедливо, що $\mathbf{A}=\overline{\mathbf{A}^T}$.

Очевидно, що діагональні елементи ермітової матриці – дійсні, бо для них виконується $a_i^i=\overline{a_i^i}$.
Ця властивість узагальнює властивість симетричності матриці для комплексного простору матриць.

Приклад 3. Матриця $\mathbf{A}=\begin{pmatrix}1 & i\\ -i & 2\end{pmatrix}$ є ермітовою, оскільки $\mathbf{A}^T=\begin{pmatrix}1 & -i\\ i & 2\end{pmatrix}$, тобто $\mathbf{A}=\overline{\mathbf{A}^T}$.

**Висновок.** Матриця Грама довільного базису унітарного простору – ермітова: $\mathbf{G}=\overline{\mathbf{G}^T}$. Якщо ж базис ортонормований, то його матриця Грама і в унітарному просторі – одинична.
Неважко одержати, як це було зроблено в лекції 19, вираз скалярного добутку двох векторів через їх координати у довільному базисі:

$$\mathbf{x}\cdot\mathbf{y}=|\mathbf{x}\rangle^T\mathbf{G}\overline{|\mathbf{y}\rangle},\tag{9}$$

а отже, в ортонормованому базисі скалярний добуток двох векторів рівний: $\mathbf{x}\cdot\mathbf{y}=|\mathbf{x}\rangle^T\overline{|\mathbf{y}\rangle}=\sum\limits_{i=1}^{n}x^i\overline{y^i}$.

Матриця Грама при переході до іншого базису міняється за формулою $\mathbf{G}'=\mathbf{T}^T\mathbf{G}\overline{\mathbf{T}}$ (переконайтесь в цьому самостійно), отже, якщо обидва

базиси – ортонормовані, то $\mathbf{E}_n=\mathbf{T}^T\mathbf{E}_n\overline{\mathbf{T}}$, $\mathbf{T}^T\overline{\mathbf{T}}=\mathbf{E}_n$, або $\mathbf{T}^{-1}=\overline{\mathbf{T}}^T$ (10)

**Означення 7.** Квадратна комплексна матриця $\mathbf{T}$, яка має властивість (10), називається ***унітарною*** матрицею.

**Висновок.** Матриця переходу між двома ортонормованими базисами унітарного простору є унітарною.
Спряжений та самоспряжений оператори визначаються в унітарному просторі так само, як і у евклідовому. Проте, властивості їх матриць узагальнюються наступним чином. Матриця $\mathbf{A}^*$ спряженого оператора $A^*$ пов'язана з матрицею оператора $A$ в довільному базисі з матрицею Грама $\mathbf{G}$ формулою: $\overline{\mathbf{A}^*}=\mathbf{G}^{-1}\mathbf{A}^T\mathbf{G}$, а в ортонормованому базисі – $\overline{\mathbf{A}^*}=\mathbf{A}^T$, тобто $\mathbf{A}^*=\overline{\mathbf{A}^T}$. Звідси очевидно випливає, що матриця самоспряженого оператора в ортонормовному базисі є ермітовою.

**Означення 8.** Лінійний оператор $A:U_n\to U_n$ в унітарному просторі називається ***унітарним***, якщо, якщо він зберігає скалярний добуток векторів, тобто $\forall\mathbf{x},\mathbf{y}\in E_n$: $\quad A\mathbf{x}\cdot A\mathbf{y}=\mathbf{x}\cdot\mathbf{y}$ (11)

Неважко переконатись, що його матриця унітарна в ортонормовному базисі, а власні числа, які можуть бути комплексними, рівні по модулю одиниці.

Зведемо всі зв'язки між лінійними операторами евклідових та унітарних просторів та їх матрицями в ортонормованих базисах у наступну таблицю (тут – лінійний оператор, - його матриця в ортонормованому базисі).

| **Дійсний евклідів простір** $E_n$ | | **Комплексний унітарний простір** $U_n$ | |
|---|---|---|---|
| Лінійний оператор $A:E_n\to E_n$ | Його матриця $\mathbf{A}$ **(в ортонормованому базисі!)** | Лінійний оператор $A:U_n\to U_n$ | Його матриця $\mathbf{A}$ **(в ортонормованому базисі!)** |
| Спряжений $A^*$ | $\mathbf{A}^*=\mathbf{A}^T$ | Спряжений $A^*$ | $\mathbf{A}^*=\overline{\mathbf{A}^T}$ |
| Самоспряжений $A=A^*$ | Симетрична: $\mathbf{A}=\mathbf{A}^T$ | Самоспряжений $A=A^*$ | Ермітова: $\mathbf{A}=\overline{\mathbf{A}^T}$ |
| Ортогональний $A^*A=E$ | Ортогональна: $\mathbf{A}^{-1}=\mathbf{A}^T$ | Унітарний $A^*A=E$ | Унітарна: $\mathbf{A}^{-1}=\overline{\mathbf{A}^T}$ |
| Матриця Грама $\mathbf{G}$ довільного базису – симетрична: $\mathbf{G}=\mathbf{G}^T$ | | Матриця Грама $\mathbf{G}$ довільного базису – ермітова: $\mathbf{G}=\overline{\mathbf{G}^T}$ | |
| Матриця переходу між двома ортонормованими базисами – ортогональна: $\mathbf{T}^{-1}=\mathbf{T}^T$ | | Матриця переходу між двома ортонормованими базисами – унітарна: $\mathbf{T}^{-1}=\overline{\mathbf{T}^T}$ | |

## Список літератури.

1. Придатченко Ю.В., Львов В.А., Єфіменко С.В. Векторна алгебра та аналітична геометрія. – Видавничо - поліграфічний центр “Київський університет”, 2010. – 96 с.
2. Білоусова В.П. та ін. Аналітична геометрія. – К.: Вища школа, 1973. – 328 с.
3. Беклемишев Д.В. Курс аналитической геометрии и линейной алгебры. – М.: Наука, 1987. – 320 с.
4. Клетеник Д.В. Сборник задач по аналитической геометрии.– М.: Наука, 1972.–240с.
5. Ильин В.А., Позняк Э.Г. Аналитическая геометрия. – М.: Наука, 1981. – 340с.
6. Курош А. Г. Курс высшей алгебры. – М.: Наука, 1970. – 432 с.
7. Проскуряков И.В. Сборник задач по линейной алгебре. – М.: Бином, 2005. – 384 с.
8. Гельфанд И.М. Лекции по линейной алгебре. – М.: Наука, 1971. – 272 с.
9. Ильин В.А., Позняк Э.Г. Линейная алгебра. – М.: Наука, 1986. – 360 с.

## Зміст.