

# Python для Data Science

PRESS START

## Заняття 2.

NumPy. Лінійна алгебра і статистика

# План заняття

---



- Коротке введення до лінійної алгебри (коротке теоретичне визначення, практичне застосування, приклад в NumPy):
  - матриці
  - векторне множення
  - додавання, віднімання, множення
  - норма матриці
- Статистика: mean, median, std, histogram, percentile

# Disclaimer

---

Ми не збираємося охоплювати всю лінійну алгебру. Лише кілька понять, які допоможуть вам ознайомитися з матричними операціями та їх використанням.

Якщо ви хочете отримати фундаментальні знання з лінійної алгебри, зробіть це! Це дуже корисна і застосовувана сфера математики!

# Why we learn linear algebra?

---

- Ви повинні вміти читати та записувати векторні та матричні позначення. Алгоритми описані в книгах, статтях і на веб-сайтах з використанням векторної та матричної нотації
- Лінійна алгебра дозволяє представляти в компактному вигляді множини лінійних рівнянь та операцій над ними. Системи лінійних рівнянь в свою чергу лежать в основі методів оптимізації, що використовуються в машинному навчанні
- З точки зору імплементації обчислень та виконання їх комп'ютером, використання обрахунків у формі лінійної алгебри пришвидшує їх в десятки разів завдяки оптимізованим бібліотекам
- Розуміння лінійної алгебри допоможе розуміти, як працюють методи машинного навчання, покращити інтуїцію в машинному навчанні, а також розробляти власні алгоритми

# Нотація

- $A \in R^{m \times n}$  - матриця з  $m$  рядками та  $n$  стовпчиками, кожне значення якої є дійсне число.
- $x \in R^n$  - вектор з  $n$  значеннями, кожне з яких є дійсне число.  $n$  - розмірний вектор можна сприймати як матрицю. Матриця з  $n$  рядками і 1 колонкою називається **вектор-колонка**, а матриця з 1 рядком і  $n$  колонками – **вектор-рядок**. Вектор рядок ми позначаємо  $x^T$ .

# Нотація

- $i$ -й елемент вектору  $x$  ми позначаємо як  $x_i$ :

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{bmatrix}$$

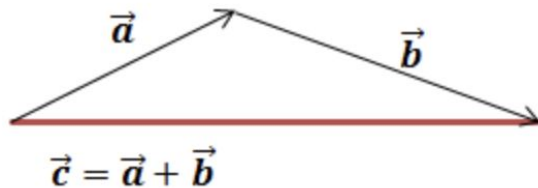
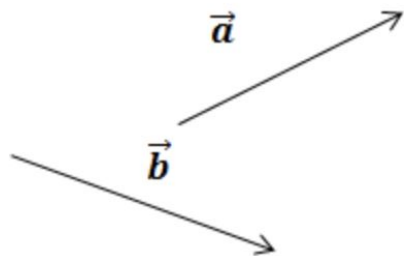
- $a_{ij}$  (або  $A_{ij}$ ,  $A_{i,j}$  і т.д.) – значення елемента матриці  $A$  у  $i$ -му рядку та  $j$ -й колонці:

# Нотація

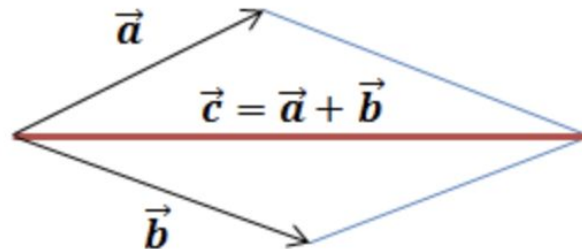
---

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

# Додавання векторів



Правило трикутника



Правило паралелограма

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} x_1 + y_1 \\ x_2 + y_2 \end{bmatrix}$$

# Додавання матриць

$$\begin{aligned} \mathbf{A} + \mathbf{B} &= \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{bmatrix} \\ &= \begin{bmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \cdots & a_{1n} + b_{1n} \\ a_{21} + b_{21} & a_{22} + b_{22} & \cdots & a_{2n} + b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} + b_{m1} & a_{m2} + b_{m2} & \cdots & a_{mn} + b_{mn} \end{bmatrix} \end{aligned}$$

# Множення вектора/матриці на число

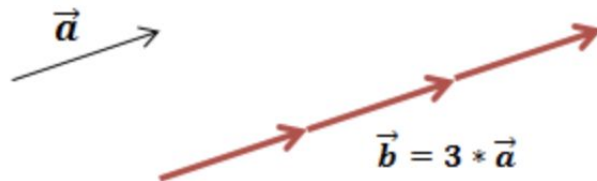
$$2A = 2 \cdot \begin{pmatrix} 3 & -1 \\ 2 & 4 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 2 \cdot 3 & 2 \cdot (-1) \\ 2 \cdot 2 & 2 \cdot 4 \\ 2 \cdot 1 & 2 \cdot 1 \end{pmatrix} = \begin{pmatrix} 6 & -2 \\ 4 & 8 \\ 2 & 2 \end{pmatrix},$$

$$\vec{a} (1;1) \rightarrow 2\vec{a} (2; 2)$$

$$\vec{a} (1;1) \rightarrow -3\vec{a} (-3; -3)$$

$$\vec{b} (3; 0) \rightarrow \frac{2}{3}\vec{b} (2; 0)$$

$$\vec{b} (3; 0) \rightarrow -\frac{1}{2}\vec{b} (-1,5; 0)$$



# Множення матриць

$$\begin{matrix} \mathbf{A} & \times & \mathbf{B} & = & \mathbf{AB} \\ m \times n & & n \times p & & m \times p \end{matrix}$$

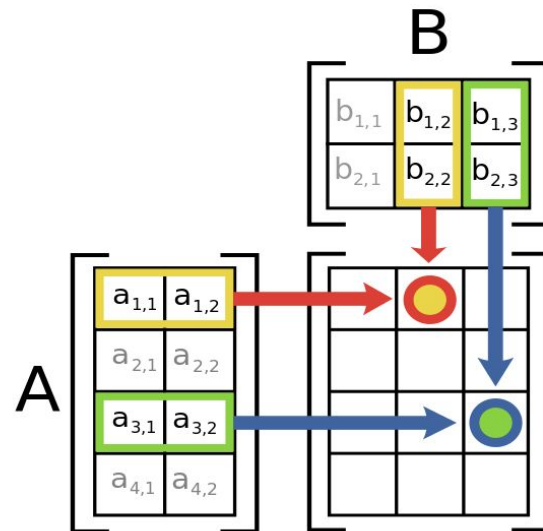
узгоджена форма матриць

$$x_{1,2} = (a_{1,1}, a_{1,2}) \cdot (b_{1,2}, b_{2,2})$$

$$= a_{1,1}b_{1,2} + a_{1,2}b_{2,2}$$

$$x_{3,3} = (a_{3,1}, a_{3,2}) \cdot (b_{1,3}, b_{2,3})$$

$$= a_{3,1}b_{1,3} + a_{3,2}b_{2,3}$$



# Множення матриць

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \begin{pmatrix} a & d \\ b & e \\ c & f \end{pmatrix} = \begin{pmatrix} 1 * a + 2 * b + 3 * c & 1 * d + 2 * e + 3 * f \\ 4 * a + 5 * b + 6 * c & 4 * d + 5 * e + 6 * f \\ 7 * a + 8 * b + 9 * c & 7 * d + 8 * e + 9 * f \end{pmatrix}$$

$$\begin{pmatrix} 1 & 2 & 3 \end{pmatrix} \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix} = (1 * 4 + 2 * 5 + 3 * 6) = (32)$$

$$\begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \end{pmatrix} = \begin{pmatrix} 4 * 1 & 4 * 2 & 4 * 3 \\ 5 * 1 & 5 * 2 & 5 * 3 \\ 6 * 1 & 6 * 2 & 6 * 3 \end{pmatrix}$$

# Множення матриць. Властивості

---

- Асоціативність  $(AB)C = A(BC)$
- Дистрибутивність щодо додавання  $A(B+C) = AB + AC$
- Некомутативність  $AB \neq BA$

# Одинична матриця

**Одинична матриця (identity matrix)**, що позначається  $I \in R^{n \times n}$ , - це квадратна матриця, всі елементи якої, крім елементів головної діагоналі, дорівнюють нулю, а елементи головної діагоналі дорівнюють одиниці:

$$I_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

Для будь-якої матриці  $A$  виконується властивість:

$$AI = IA = A$$

# Діагональна матриця

---

**Діагональна матриця** – квадратна матриця, всі елементи якої, крім елементів головної діагоналі, дорівнюють нулю.

$$D_{ij} = \begin{cases} d_i & i = j \\ 0 & i \neq j \end{cases}$$

Очевидно, що одинична матриця також є діагональною.

# Транспонована матриця

**Транспонована матриця** – матриця, в якій рядки та стовпці поміняли місцями.

$$A_{ij} = (A^T)_{ji}$$

Властивості транспонованої матриці:

- $(A^T)^T = A$
- $(AB)^T = B^T A^T$
- $(A + B)^T = A^T + B^T$

# Норма вектору

**Норма** вектору  $x$  записується як  $\|x\|$  і є, простими словами, мірою «довжини» цього вектору. Довжину вектору (і, відповідно, його норму) можна міряти різними способами. Найбільш популярним та таким, що є застосовним для векторів дійсних чисел, є Евклідова норма:

$$\|x\|_2 = \sqrt{\sum_{i=1}^n x_i^2}$$

Зверніть також увагу, що в векторній нотації це можна записати як:

$$\|x\|^2 = x^T x$$

# Обернена матриця

**Обернена матриця** до квадратної матриці  $A$  - це матриця, яка при множенні на матрицю  $A$  дає одиничну матрицю. Обернена матриця позначається як  $A^{-1}$ :

$$A^{-1}A = AA^{-1} = I$$

Не всі матриці мають обернену матрицю. Не квадратні матриця, наприклад, не мають обернених матриць взагалі. Проте навіть якщо матриця  $A$  є квадратною, може бути так, що до неї немає оберненої матриці (ми називаємо таку квадратну матрицю **виродженою (singular)**).

## Система лінійних рівнянь

---

$$\begin{array}{l} 4x_1 - 5x_2 = -13 \\ -2x_1 + 3x_2 = 9 \end{array} \longrightarrow Ax = b$$

$$A = \begin{bmatrix} 4 & -5 \\ -2 & 3 \end{bmatrix}, \quad b = \begin{bmatrix} -13 \\ 9 \end{bmatrix}$$

# Матричний метод розв'язання систем лінійних алгебраїчних рівнянь

---

$$AX = B.$$

$$A^{-1}AX = A^{-1}B \Rightarrow EX = A^{-1}B \Rightarrow X = A^{-1}B.$$

Обернена матриця до матриці A

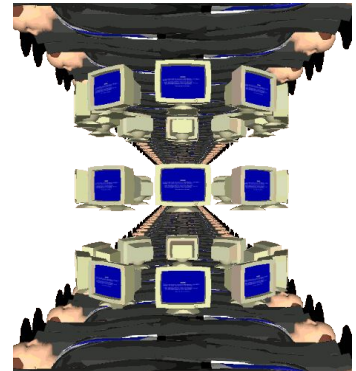
Одинична матриця до матриці A

# Поради до самостійного опрацювання

---

- Визначник (детермінант) матриці
- Ранг матриці
- Способи обчислення оберненої матриці
- Власні числа і власні вектори
- Методи розв'язання систем лінійних алгебраїчних рівнянь  
(Метод Гауса, Метод Крамера, ...)

# Live coding / Практика



# Описова статистика (descriptive statistics)

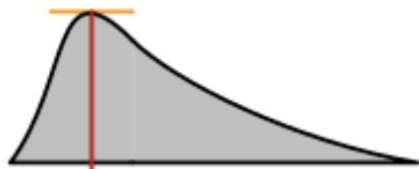
---

**Описова статистика** — це розділ статистики, який займається обробкою емпіричних даних, їх систематизацією, наочним представленням у вигляді графіків та таблиць, а також їх кількісним описом через основні статистичні показники.

**Генеральна сукупність** — вся множина однорідних за певною ознакою об'єктів чи подій, які є предметом інтересу або дослідження (всі можливі відхилення температури від норми, всі можливі результати футбольного матчу). Зазвичай, обсяг генеральної сукупності дуже великий, що робить прийняття до уваги всіх елементів непрактичним або неможливим.

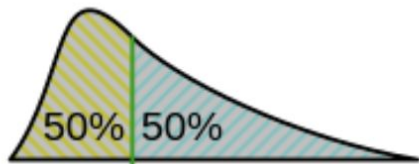
**Вибірка** — це множина об'єктів або подій, вибраних з генеральної сукупності для участі в дослідженні. Статистичні характеристики обчислюються таким чином, що в результаті можна зробити висновки або екстраполяцію із вибірки на всю генеральну сукупність.

# Центральні метрики



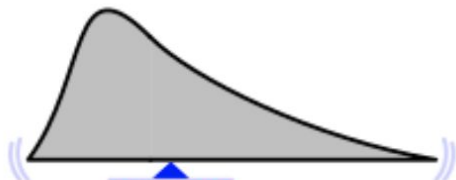
мода

**Мода** — значення, що зустрічається у вибірці найчастіше. Мода не є унікальною. Набір із більше ніж однією модою називають мультимодальним.



медіана

**Медіана** — значення, що розділяє впорядковану вибірку на дві рівні частини.



середнє

**Середнє** — середнє арифметичне вибірки.

# Центральні метрики (приклад)

---

Зарплати Data Scientists:

1000, 2000, 2300, 2300, 2300, 3500, 3600, 10000

Мода = 2300

Медіана =  $(2300 + 2300) / 2 = 2300$

Середнє =  $(1000+2000+2300+2300+2300+3500+3600+10000) / 8 = 3375$

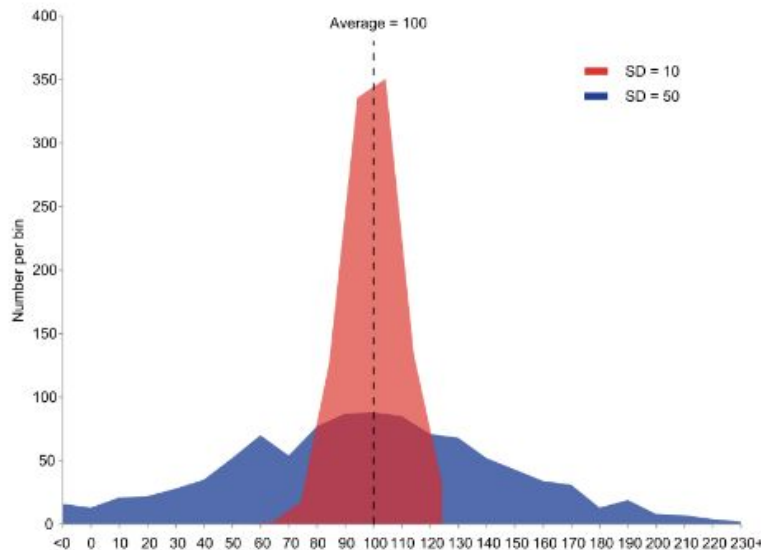
# Метрики варіативності

**Дисперсія** — це міра розсіювання значень випадкової величини відносно середнього значення розподілу. Більші значення дисперсії свідчать про більші відхилення значень випадкової величини від центру розподілу.

Незміщена оцінка дисперсії:

$$S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$$

\*Почитайте про поправку Бесселя

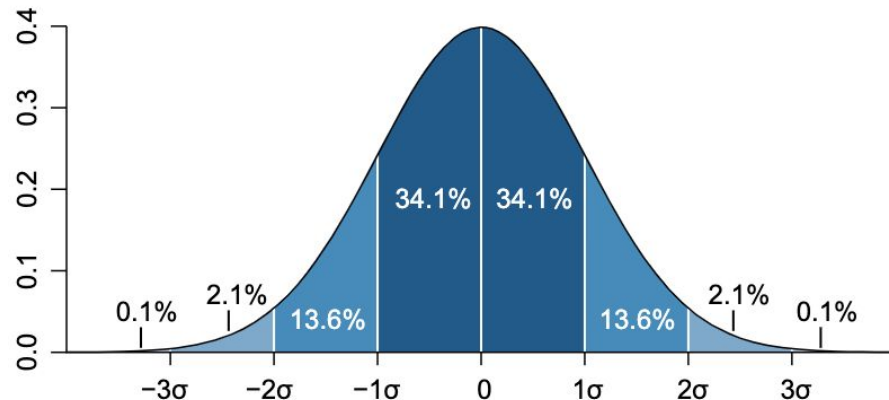


Дві вибірки з однаковим середнім,  
але різними дисперсіями

# Метрики варіативності

**Стандартне відхилення** — один із найпоширеніших показників розсіювання (розкиду) значень випадкової величини відносно її математичного сподівання, тобто центру розподілу. Має ту ж розмірність, що і випадкова величина.

$$s = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x}_B)^2}.$$



Частка вимірювань нормально розподіленої вибірки, які потраплять в інтервали величиною в 0, 1, 2, і 3 стандартних відхилень по обидві сторони від фактичного значення.

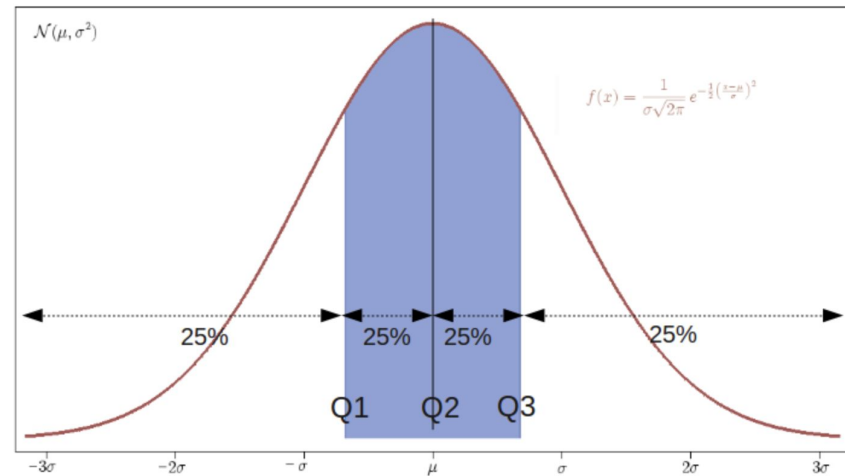
# Квантилі

**Квантиль** розподілення значень — це таке число  $x_p$ , що значення  $p$ -ї частини сукупності менше або рівне  $x_p$ . Наприклад, квантиль 0.25 (також називається 25-м процентилем або нижнім квантилем) змінної — це таке значення ( $x_p$ ), що 25 % ( $p$ ) значень змінної потрапляють нижче даного значення.

0.25-квантиль називається першим (або нижнім) квантилем

0.5-квантиль називається медіаною або другим квантилем

0.75-квантиль називається третім (або верхнім) квантилем



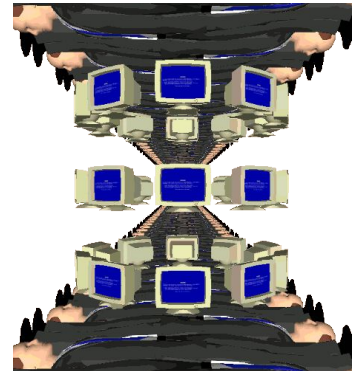
Квантилі нормального розподілу

# Розмах вибірки

**Розмах вибірки** — це різниця між найбільшим і найменшим значенням випадкової величини у вибірці.

$$R = x_{\max} - x_{\min}$$

# Live coding / Практика



# Computations with Arrays

---

- Правило 1: Операції між декількома об'єктами масиву спочатку перевіряються на відповідність форми
- Правило 2: Математичні операції (+ - \* / exp, log, ...) застосовуються елемент за елементом до значень
- Правило 3: Операції скорочення (mean, std, skew, kurt, sum, prod, ...) застосовуються до всього масиву, якщо не вказано вісь
- Правило 4: Відсутні значення поширюються, якщо вони явно не проігноровані (nanmean, nansum, ...)

Питання

???

???