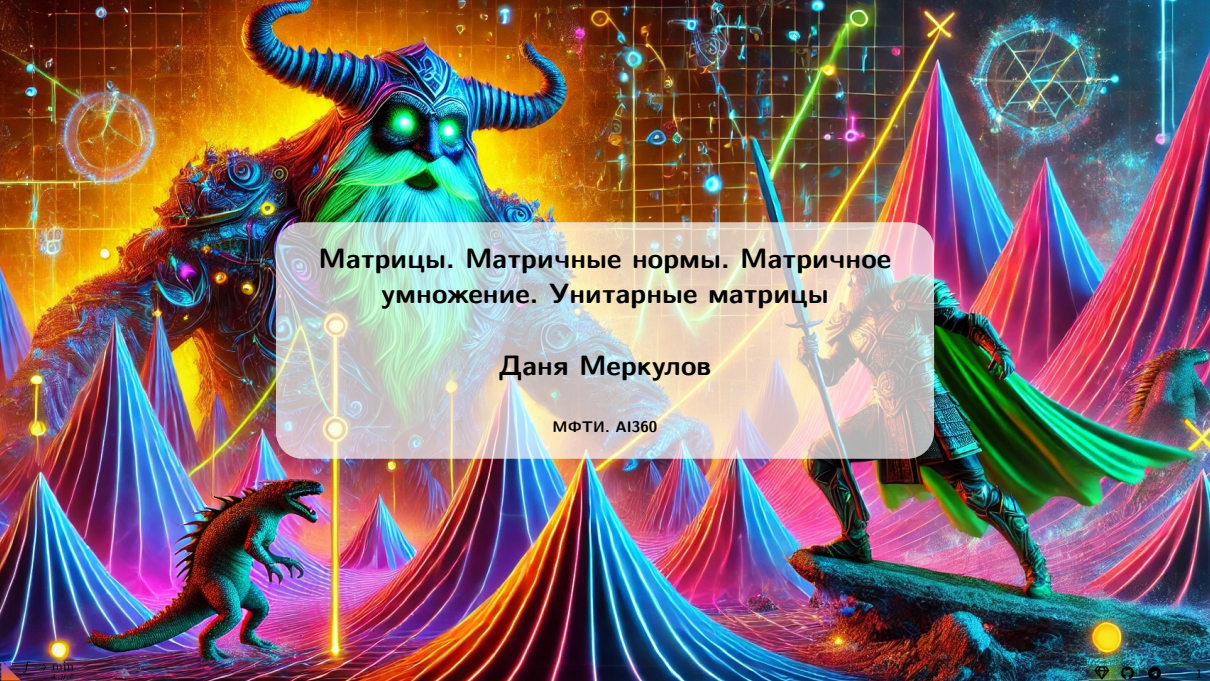


# Матрицы. Матричные нормы. Матричное умножение. Унитарные матрицы

Даня Меркулов

МФТИ. AI360

# Базовые понятия линейной алгебры

# Векторы и матрицы

Мы будем считать, что все векторы являются столбцами по умолчанию. Пространство векторов длины $n$ обозначается $\mathbb{R}^n$, а пространство вещественных $m \times n$ матриц обозначается $\mathbb{R}^{m \times n}$. [1]

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \quad x^T = \begin{bmatrix} x_1 & x_2 & \dots & x_n \end{bmatrix} \quad x \in \mathbb{R}^n, x_i \in \mathbb{R} \tag{1}$$

[1] Подробный вводный курс по прикладной линейной алгебре можно найти в книге Introduction to Applied Linear Algebra -- Vectors, Matrices, and Least Squares - книга Стивена Бойда и Ливена Ванденбергена, которая указана в источнике. Также полезной книгой по линейной алгебре является приложение A в книге Numerical Optimization Джорджа Носедаля и Стивена Райта.

# Векторы и матрицы

Мы будем считать, что все векторы являются столбцами по умолчанию. Пространство векторов длины $n$ обозначается $\mathbb{R}^n$, а пространство вещественных $m \times n$ матриц обозначается $\mathbb{R}^{m \times n}$. [1]

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \quad x^T = \begin{bmatrix} x_1 & x_2 & \dots & x_n \end{bmatrix} \quad x \in \mathbb{R}^n, x_i \in \mathbb{R} \tag{1}$$

Аналогично, если $A \in \mathbb{R}^{m \times n}$ мы обозначаем транспонирование как $A^T \in \mathbb{R}^{n \times m}$:

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \quad A^T = \begin{bmatrix} a_{11} & a_{21} & \dots & a_{m1} \\ a_{12} & a_{22} & \dots & a_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \dots & a_{mn} \end{bmatrix} \quad A \in \mathbb{R}^{m \times n}, a_{ij} \in \mathbb{R}$$

Мы будем писать $x \geq 0$ и $x \neq 0$ для обозначения покомпонентного неравенства.

[1] Подробный вводный курс по прикладной линейной алгебре можно найти в книге Introduction to Applied Linear Algebra -- Vectors, Matrices, and Least Squares - книга Стивена Бойда и Ливена Ванденбергена, которая указана в источнике. Также полезной книгой по линейной алгебре является приложение A в книге Numerical Optimization Джорджа Носедаля и Стивена Райта.

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

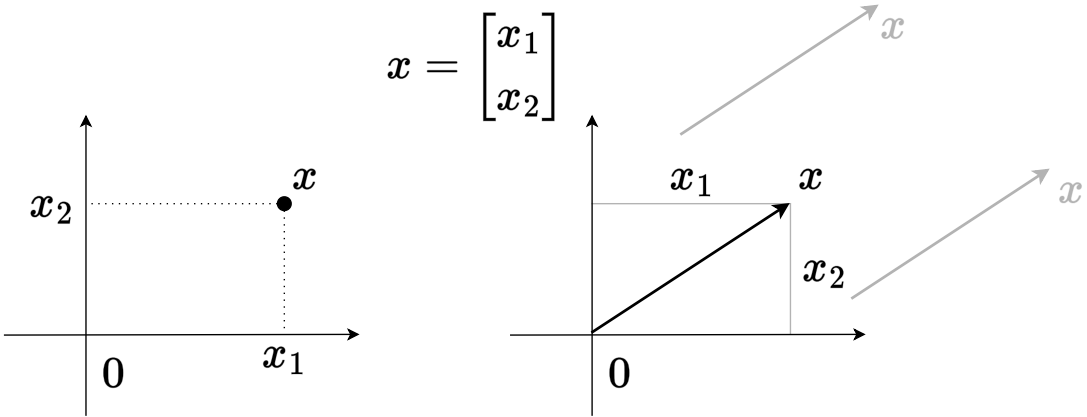



Рис. 1: Эквивалентные представления вектора

Матрица $A$ называется симметричной, если $A = A^T$. Обозначается как $A \in \mathbb{S}^n$ (множество квадратных симметричных матриц размерности $n$). Заметим, что только квадратная матрица может быть симметричной по определению.

Матрица $A$ называется симметричной, если $A = A^T$. Обозначается как $A \in \mathbb{S}^n$ (множество квадратных симметричных матриц размерности $n$). Заметим, что только квадратная матрица может быть симметричной по определению.

Матрица $A \in \mathbb{S}^n$ называется **положительно (отрицательно) определенной**, если для всех $x \neq 0 : x^T Ax > (<)0$. Обозначается как $A \succ (\prec)0$. Множество таких матриц обозначается как $\mathbb{S}^n_{++}(\mathbb{S}^n_{--})$.

Матрица $A$ называется симметричной, если $A = A^T$. Обозначается как $A \in \mathbb{S}^n$ (множество квадратных симметричных матриц размерности $n$). Заметим, что только квадратная матрица может быть симметричной по определению.

Матрица $A \in \mathbb{S}^n$ называется **положительно (отрицательно) определенной**, если для всех $x \neq 0 : x^T Ax > (<)0$. Обозначается как $A \succ (\prec)0$. Множество таких матриц обозначается как $\mathbb{S}^n_{++}(\mathbb{S}^n_{--})$.

Матрица $A \in \mathbb{S}^n$ называется **положительно (отрицательно) полуопределенной**, если для всех $x : x^T Ax \geq (\leq)0$. Обозначается как $A \succeq (\preceq)0$. Множество таких матриц обозначается как $\mathbb{S}^n_{+}(\mathbb{S}^n_{-})$.

> **Question**
>
> Верно ли, что положительно определенная матрица имеет все положительные элементы?

Матрица $A$ называется симметричной, если $A = A^T$. Обозначается как $A \in \mathbb{S}^n$ (множество квадратных симметричных матриц размерности $n$). Заметим, что только квадратная матрица может быть симметричной по определению.

Матрица $A \in \mathbb{S}^n$ называется **положительно (отрицательно) определенной**, если для всех $x \neq 0 : x^T Ax > (<)0$. Обозначается как $A \succ (\prec)0$. Множество таких матриц обозначается как $\mathbb{S}^n_{++}(\mathbb{S}^n_{--})$.

Матрица $A \in \mathbb{S}^n$ называется **положительно (отрицательно) полуопределенной**, если для всех $x : x^T Ax \geq (\leq)0$. Обозначается как $A \succeq (\preceq)0$. Множество таких матриц обозначается как $\mathbb{S}^n_{+}(\mathbb{S}^n_{-})$.

**Question**

Верно ли, что положительно определенная матрица имеет все положительные элементы?

**Question**

Верно ли, что если матрица симметрична, то она положительно определена?

Матрица $A$ называется симметричной, если $A = A^T$. Обозначается как $A \in \mathbb{S}^n$ (множество квадратных симметричных матриц размерности $n$). Заметим, что только квадратная матрица может быть симметричной по определению.

Матрица $A \in \mathbb{S}^n$ называется **положительно (отрицательно) определенной**, если для всех $x \neq 0 : x^T Ax > (<)0$. Обозначается как $A \succ (\prec)0$. Множество таких матриц обозначается как $\mathbb{S}^n_{++}(\mathbb{S}^n_{--})$.

Матрица $A \in \mathbb{S}^n$ называется **положительно (отрицательно) полуопределенной**, если для всех $x : x^T Ax \geq (\leq)0$. Обозначается как $A \succeq (\preceq)0$. Множество таких матриц обозначается как $\mathbb{S}^n_{+}(\mathbb{S}^n_{-})$.

**Question**

Верно ли, что положительно определенная матрица имеет все положительные элементы?

**Question**

Верно ли, что если матрица симметрична, то она положительно определена?

**Question**

Верно ли, что если матрица положительно определена, то она симметрична?

# Умножение матриц

Пусть $A$ - матрица размера $m \times n$, и $B$ - матрица размера $n \times p$, и пусть произведение $AB$ задается как:

$$C = AB$$

тогда $C$ - матрица размера $m \times p$, с элементом $(i, j)$ задаваемым как:

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

Эта операция в наивной реализации требует $\mathcal{O}(n^3)$ арифметических операций, где $n$ - наибольший размер матриц.

# Умножение матриц

Пусть $A$ - матрица размера $m \times n$, и $B$ - матрица размера $n \times p$, и пусть произведение $AB$ задается как:

$$C = AB$$

тогда $C$ - матрица размера $m \times p$, с элементом $(i, j)$ задаваемым как:

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

Эта операция в наивной реализации требует $\mathcal{O}(n^3)$ арифметических операций, где $n$ - наибольший размер матриц.

**Question**

Можно ли умножить две матрицы быстрее, чем за $\mathcal{O}(n^3)$? Как насчет $\mathcal{O}(n^2)$, $\mathcal{O}(n)$?

# Умножение матрицы на вектор

Пусть $A$ - матрица размера $m \times n$, и $x$ - вектор длины $n$, тогда $i$-й компонент произведения:

$$z = Ax$$

определяется как:

$$z_i = \sum_{k=1}^{n} a_{ik} x_k$$

Эта операция в наивной реализации требует $\mathcal{O}(n^2)$ арифметических операций, где $n$ - наибольший размер матриц.

Заметим, что:

- $C = AB \quad C^T = B^T A^T$

# Умножение матрицы на вектор

Пусть $A$ - матрица размера $m \times n$, и $x$ - вектор длины $n$, тогда $i$-й компонент произведения:

$$z = Ax$$

определяется как:

$$z_i = \sum_{k=1}^{n} a_{ik} x_k$$

Эта операция в наивной реализации требует $\mathcal{O}(n^2)$ арифметических операций, где $n$ - наибольший размер матриц.

Заметим, что:

- $C = AB \quad C^T = B^T A^T$
- $AB \neq BA$

# Умножение матрицы на вектор

Пусть $A$ - матрица размера $m \times n$, и $x$ - вектор длины $n$, тогда $i$-й компонент произведения:

$$z = Ax$$

определяется как:

$$z_i = \sum_{k=1}^{n} a_{ik} x_k$$

Эта операция в наивной реализации требует $\mathcal{O}(n^2)$ арифметических операций, где $n$ - наибольший размер матриц.

Заметим, что:

- $C = AB \quad C^T = B^T A^T$
- $AB \neq BA$
- $e^A = \sum\limits_{k=0}^{\infty} \frac{1}{k!} A^k$

# Умножение матрицы на вектор

Пусть $A$ - матрица размера $m \times n$, и $x$ - вектор длины $n$, тогда $i$-й компонент произведения:

$$z = Ax$$

определяется как:

$$z_i = \sum_{k=1}^{n} a_{ik} x_k$$

Эта операция в наивной реализации требует $\mathcal{O}(n^2)$ арифметических операций, где $n$ - наибольший размер матриц.

Заметим, что:

- $C = AB \quad C^T = B^T A^T$
- $AB \neq BA$
- $e^A = \sum\limits_{k=0}^{\infty} \frac{1}{k!} A^k$
- $e^{A+B} \neq e^A e^B$ (но если $A$ и $B$ - коммутирующие матрицы, то есть $AB = BA$, то $e^{A+B} = e^A e^B$)

# Умножение матрицы на вектор

Пусть $A$ - матрица размера $m \times n$, и $x$ - вектор длины $n$, тогда $i$-й компонент произведения:

$$z = Ax$$

определяется как:

$$z_i = \sum_{k=1}^{n} a_{ik} x_k$$

Эта операция в наивной реализации требует $\mathcal{O}(n^2)$ арифметических операций, где $n$ - наибольший размер матриц.

Заметим, что:

- $C = AB \quad C^T = B^T A^T$
- $AB \neq BA$
- $e^A = \sum\limits_{k=0}^{\infty} \frac{1}{k!} A^k$
- $e^{A+B} \neq e^A e^B$ (но если $A$ и $B$ - коммутирующие матрицы, то есть $AB = BA$, то $e^{A+B} = e^A e^B$)
- $\langle x, Ay \rangle = \langle A^T x, y \rangle$

# Нормы

Норма - это **количественная мера маленькости вектора** и обычно обозначается как $\|x\|$.

Норма должна удовлетворять определенным свойствам:

1. $\|\alpha x\| = |\alpha|\|x\|$, $\alpha \in \mathbb{R}$

# Нормы

Норма - это **количественная мера маленькости вектора** и обычно обозначается как $\|x\|$.

Норма должна удовлетворять определенным свойствам:

1. $\|\alpha x\| = |\alpha|\|x\|$, $\alpha \in \mathbb{R}$
2. $\|x + y\| \leq \|x\| + \|y\|$ (неравенство треугольника)

# Нормы

Норма - это **количественная мера маленькости вектора** и обычно обозначается как $\|x\|$.

Норма должна удовлетворять определенным свойствам:

1. $\|\alpha x\| = |\alpha|\|x\|$, $\alpha \in \mathbb{R}$
2. $\|x + y\| \leq \|x\| + \|y\|$ (неравенство треугольника)
3. Если $\|x\| = 0$ то $x = 0$

# Нормы

Норма - это **количественная мера маленькости вектора** и обычно обозначается как $\|x\|$.

Норма должна удовлетворять определенным свойствам:

1. $\|\alpha x\| = |\alpha|\|x\|$, $\alpha \in \mathbb{R}$
2. $\|x + y\| \leq \|x\| + \|y\|$ (неравенство треугольника)
3. Если $\|x\| = 0$ то $x = 0$

# Нормы

Норма - это **количественная мера маленькости вектора** и обычно обозначается как $\|x\|$.

Норма должна удовлетворять определенным свойствам:

1. $\|\alpha x\| = |\alpha|\|x\|$, $\alpha \in \mathbb{R}$
2. $\|x + y\| \leq \|x\| + \|y\|$ (неравенство треугольника)
3. Если $\|x\| = 0$ то $x = 0$

Расстояние между двумя векторами определяется как:

$$d(x, y) = \|x - y\|.$$

Наиболее широко используемой нормой является **евклидова норма**:

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

которая соответствует расстоянию в нашей повседневной жизни. Если векторы имеют комплексные элементы, мы используем их модуль. Евклидова норма, или 2-норма, является подклассом важного класса $p$-норм:

$$\|x\|_p = \left( \sum_{i=1}^{n} |x_i|^p \right)^{1/p}.$$

## $p$-норма вектора

Существуют два очень важных частных случая. Бесконечность норма, или норма Чебышёва, определяется как максимальный абсолютный элемент:

$$\|x\|_\infty = \max_i |x_i|$$

## $p$-норма вектора

Существуют два очень важных частных случая. Бесконечность норма, или норма Чебышёва, определяется как максимальный абсолютный элемент:

$$\|x\|_\infty = \max_i |x_i|$$

$L_1$ норма (или **манхэттенское расстояние**) определяется как сумма модулей элементов вектора:

$$\|x\|_1 = \sum_i |x_i|$$

## $p$-норма вектора

Существуют два очень важных частных случая. Бесконечность норма, или норма Чебышёва, определяется как максимальный абсолютный элемент:

$$\|x\|_\infty = \max_i |x_i|$$

$L_1$ норма (или **манхэттенское расстояние**) определяется как сумма модулей элементов вектора:

$$\|x\|_1 = \sum_i |x_i|$$

$L_1$ норма играет очень важную роль: она связана с методами **compressed sensing**, которые стали одной из популярных тем исследований в середине 00-х. Код для изображения ниже доступен *здесь:*. Также можно посмотреть *этот* видео.

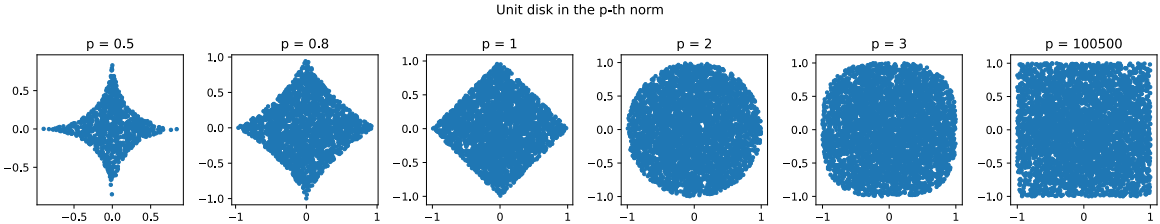



Рис. 2: Шары в разных нормах на плоскости

# Матричные нормы

# Матричные нормы

$\|\cdot\|$ называется **матричной нормой**, если она является векторной нормой на векторном пространстве $n \times m$ матриц:

1. $\|A\| \geq 0$ и если $\|A\| = 0$, то $A = O$

# Матричные нормы

$\|\cdot\|$ называется **матричной нормой**, если она является векторной нормой на векторном пространстве $n \times m$ матриц:

1. $\|A\| \geq 0$ и если $\|A\| = 0$, то $A = O$
2. $\|\alpha A\| = |\alpha| \|A\|$

# Матричные нормы

$\|\cdot\|$ называется **матричной нормой**, если она является векторной нормой на векторном пространстве $n \times m$ матриц:

1. $\|A\| \geq 0$ и если $\|A\| = 0$, то $A = O$
2. $\|\alpha A\| = |\alpha| \|A\|$
3. $\|A + B\| \leq \|A\| + \|B\|$ (неравенство треугольника)

# Матричные нормы

$\|\cdot\|$ называется **матричной нормой**, если она является векторной нормой на векторном пространстве $n \times m$ матриц:

1. $\|A\| \geq 0$ и если $\|A\| = 0$, то $A = O$
2. $\|\alpha A\| = |\alpha| \|A\|$
3. $\|A + B\| \leq \|A\| + \|B\|$ (неравенство треугольника)

- Дополнительно, некоторые нормы могут удовлетворять **свойству субмультипликативности**:

$$\|AB\| \leq \|A\| \|B\|$$

Эти нормы называются **субмультипликативными нормами**.

# Матричные нормы

$\|\cdot\|$ называется **матричной нормой**, если она является векторной нормой на векторном пространстве $n \times m$ матриц:

1. $\|A\| \geq 0$ и если $\|A\| = 0$, то $A = O$
2. $\|\alpha A\| = |\alpha| \|A\|$
3. $\|A + B\| \leq \|A\| + \|B\|$ (неравенство треугольника)

- Дополнительно, некоторые нормы могут удовлетворять **свойству субмультипликативности**:

$$\|AB\| \leq \|A\| \|B\|$$

  Эти нормы называются **субмультипликативными нормами**.

- Пример несубмультипликативной нормы - норма Чебышёва:

$$\|A\|_C = \max_{i,j} |a_{ij}|$$

# Пример

Придумайте контрпример

Покажите, что норма Чебышёва не является субмультипликативной.

# Пример

**Придумайте контрпример**

Покажите, что норма Чебышёва не является субмультипликативной.

Рассмотрим матрицы

$$A = B = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}.$$

Для каждой из них

$$\|A\|_C = \|B\|_C = 1.$$

Вычислим произведение:

$$AB = \begin{bmatrix} 1 \cdot 1 + 1 \cdot 1 & 1 \cdot 1 + 1 \cdot 1 \\ 1 \cdot 1 + 1 \cdot 1 & 1 \cdot 1 + 1 \cdot 1 \end{bmatrix} = \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix}.$$

Если норма была бы субмультипликативной, то должно выполняться

$$2 = \|AB\|_C \leq \|A\|_C \cdot \|B\|_C = 1 \cdot 1 = 1,$$

но $2 > 1$.

# Нормы матриц

В некотором смысле нет большой разницы между матрицами и векторами (можно векторизовать матрицу), таким образом вводится простейшая норма матрицы **Фробениуса**:

$$\|A\|_F = \left( \sum_{i=1}^{m} \sum_{j=1}^{n} |a_{ij}|^2 \right)^{1/2}$$

# Нормы матриц

В некотором смысле нет большой разницы между матрицами и векторами (можно векторизовать матрицу), таким образом вводится простейшая норма матрицы **Фробениуса**:

$$\|A\|_F = \left( \sum_{i=1}^{m} \sum_{j=1}^{n} |a_{ij}|^2 \right)^{1/2}$$

Спектральная норма, $\|A\|_2$ является одной из наиболее широко используемых норм матриц (вместе с нормой Фробениуса).

$$\|A\|_2 = \sup_{x \neq 0} \frac{\|Ax\|_2}{\|x\|_2},$$

Её нельзя вычислить непосредственно из элементов, используя простую формулу, как норму Фробениуса, однако существуют эффективные алгоритмы для ее вычисления. Она непосредственно связана с **сингулярным разложением** (SVD) матрицы. Она удовлетворяет следующему свойству:

$$\|A\|_2 = \sigma_1(A) = \sqrt{\lambda_{\max}(A^T A)}$$

где $\sigma_1(A)$ является наибольшим сингулярным значением матрицы $A$.

# Операторные нормы

- Наиболее важный класс матричных норм - это класс **операторных норм**. Они определяются как

$$\|A\|_{\alpha\to\beta} = \sup_{x\neq 0} \frac{\|Ax\|_\alpha}{\|x\|_\beta},$$

где $\|\cdot\|_\alpha$ и $\|\cdot\|_\beta$ являются **векторными нормами**.

# Операторные нормы

- Наиболее важный класс матричных норм - это класс **операторных норм**. Они определяются как

$$\|A\|_{\alpha\to\beta} = \sup_{x\neq 0} \frac{\|Ax\|_\alpha}{\|x\|_\beta},$$

  где $\|\cdot\|_\alpha$ и $\|\cdot\|_\beta$ являются **векторными нормами**.

- Легко проверить, что они субмультипликативны, если $\|\cdot\|_\alpha = \|\cdot\|_\beta$. Иначе они не субмультипликативны, придумайте пример.

# Операторные нормы

- Наиболее важный класс матричных норм - это класс **операторных норм**. Они определяются как

$$\|A\|_{\alpha\to\beta} = \sup_{x\neq 0} \frac{\|Ax\|_\alpha}{\|x\|_\beta},$$

  где $\|\cdot\|_\alpha$ и $\|\cdot\|_\beta$ являются **векторными нормами**.

- Легко проверить, что они субмультипликативны, если $\|\cdot\|_\alpha = \|\cdot\|_\beta$. Иначе они не субмультипликативны, придумайте пример.

- **Норма Фробениуса** является матричной нормой, но не является операторной нормой, т.е. нельзя найти $\|\cdot\|_\alpha$ и $\|\cdot\|_\beta$ такие, чтобы они её порождали.

# Операторные нормы

- Наиболее важный класс матричных норм - это класс **операторных норм**. Они определяются как

$$\|A\|_{\alpha\to\beta} = \sup_{x\neq 0} \frac{\|Ax\|_\alpha}{\|x\|_\beta},$$

  где $\|\cdot\|_\alpha$ и $\|\cdot\|_\beta$ являются **векторными нормами**.

- Легко проверить, что они субмультипликативны, если $\|\cdot\|_\alpha = \|\cdot\|_\beta$. Иначе они не субмультипликативны, придумайте пример.

- **Норма Фробениуса** является матричной нормой, но не является операторной нормой, т.е. нельзя найти $\|\cdot\|_\alpha$ и $\|\cdot\|_\beta$ такие, чтобы они её порождали.

- Это нетривиальный факт и общий критерий матричной нормы для того, чтобы быть операторной нормой можно найти в Теореме 6 и следствии 4.

# Пример

**Question**

Покажите, что операторная норма $\|\cdot\|_{\alpha\to\alpha}$ является субмультипликативной.

# Пример

**Question**

Покажите, что операторная норма $\|\cdot\|_{\alpha\to\alpha}$ является субмультипликативной.

Определим операторную норму для матрицы $A$ как

$$\|A\|_{\alpha\to\alpha} = \sup_{x\neq 0} \frac{\|Ax\|_\alpha}{\|x\|_\alpha}.$$

Для двух матриц $A$ и $B$ рассмотрим для любого $x \neq 0$:

$$\|ABx\|_\alpha \leq \|A\|_{\alpha\to\alpha}\|Bx\|_\alpha \leq \|A\|_{\alpha\to\alpha}\|B\|_{\alpha\to\alpha}\|x\|_\alpha.$$

Делим на $\|x\|_\alpha$ и берем супремум по $x \neq 0$:

$$\|AB\|_{\alpha\to\alpha} \leq \|A\|_{\alpha\to\alpha}\|B\|_{\alpha\to\alpha}.$$

Таким образом, операторная норма с одинаковыми векторными нормами является субмультипликативной.

# Матричные $p$-нормы

Важный частный случай операторных норм - это матричные $p$-нормы, которые определяются для $\|\cdot\|_\alpha = \|\cdot\|_\beta = \|\cdot\|_p$.

Среди всех $p$-норм три наиболее часто используемые нормы:

- $p = 1, \quad \|A\|_1 = \max\limits_j \sum\limits_{i=1}^{n} |a_{ij}|.$

# Матричные $p$-нормы

Важный частный случай операторных норм - это матричные $p$-нормы, которые определяются для $\|\cdot\|_\alpha = \|\cdot\|_\beta = \|\cdot\|_p$.

Среди всех $p$-норм три наиболее часто используемые нормы:

- $p = 1, \quad \|A\|_1 = \max\limits_j \sum\limits_{i=1}^{n} |a_{ij}|.$
- $p = 2, \quad$ спектральная норма, обозначаемая как $\|A\|_2$.

# Матричные $p$-нормы

Важный частный случай операторных норм - это матричные $p$-нормы, которые определяются для $\|\cdot\|_\alpha = \|\cdot\|_\beta = \|\cdot\|_p$.

Среди всех $p$-норм три наиболее часто используемые нормы:

- $p = 1, \quad \|A\|_1 = \max\limits_j \sum\limits_{i=1}^{n} |a_{ij}|.$
- $p = 2, \quad$ спектральная норма, обозначаемая как $\|A\|_2$.
- $p = \infty, \quad \|A\|_\infty = \max\limits_i \sum\limits_{j=1}^{m} |a_{ij}|.$

# Пример

Покажите, что $\|A\|_1 = \max_j \sum_{i=1}^{n} |a_{ij}|$.

# Пример

Покажите, что $\|A\|_1 = \max_j \sum_{i=1}^n |a_{ij}|$.

Решениеможно разбить на две части: показать, что норма не превосходит максимальной суммы модулей столбца, и построить вектор, для которого достигается это значение. Норма $\|A\|_1$ определяется как операторная норма, индуцированная векторной 1-нормой:

$$\|A\|_1 = \sup_{x \neq 0} \frac{\|Ax\|_1}{\|x\|_1}.$$

При этом можно показать, что

$$\|Ax\|_1 = \sum_i \left| \sum_j a_{ij} x_j \right| \leq \sum_{i=1}^n \sum_{j=1}^n |a_{ij}||x_j| = \sum_{j=1}^n |x_j| \left( \sum_{i=1}^n |a_{ij}| \right).$$

Если выбрать $x$ так, чтобы вся масса $\|x\|_1 = 1$ приходилась на столбец с максимальной суммой модулей, получим, что супремум равен

$$\max_j \sum_{i=1}^n |a_{ij}|.$$

# Пример

Покажите, что $\|A\|_\infty = \max_i \sum_{j=1}^m |a_{ij}|$.

# Пример

Покажите, что $\|A\|_\infty = \max_i \sum_{j=1}^m |a_{ij}|$.

Аналогично, операторная норма, индуцированная $\infty$-нормой, определяется как

$$\|A\|_\infty = \sup_{x \neq 0} \frac{\|Ax\|_\infty}{\|x\|_\infty}.$$

При этом легко доказать, что

$$\|Ax\|_\infty = \max_i \left| \sum_j a_{ij} x_j \right| \leq \max_i \sum_j |a_{ij}| \, \|x\|_\infty.$$

Отсюда следует, что

$$\|A\|_\infty = \max_i \sum_{j=1}^m |a_{ij}|.$$

# Скалярное произведение

Стандартное **скалярное (внутреннее) произведение** между векторами $x$ и $y$ из $\mathbb{R}^n$ определяется как:

$$\langle x, y \rangle = x^T y = \sum_{i=1}^{n} x_i y_i = y^T x = \langle y, x \rangle$$

Здесь $x_i$ и $y_i$ являются $i$-ми компонентами соответствующих векторов.

**Example**

Докажите, что можно перемещать матрицу внутри скалярного произведения с транспонированием: $\langle x, Ay \rangle = \langle A^T x, y \rangle$ и $\langle x, yB \rangle = \langle xB^T, y \rangle$

# Скалярное произведение матриц

Стандартное **скалярное (внутреннее) произведение** между матрицами $X$ и $Y$ из $\mathbb{R}^{m \times n}$ определяется как:

$$\langle X, Y \rangle = \text{tr}(X^T Y) = \sum_{i=1}^{m} \sum_{j=1}^{n} X_{ij} Y_{ij} = \text{tr}(Y^T X) = \langle Y, X \rangle$$

**Question**

Существует ли связь между нормой Фробениуса $\|\cdot\|_F$ и скалярным произведением между матрицами $\langle \cdot, \cdot \rangle$?

# Пример

**Question**

Посчитайте скалярное произведение матриц

$$\left\langle \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \right\rangle$$

# Пример

**Question**

Упростите следующее выражение:

$$\sum_{i=1}^{n} \langle S^{-1} a_i, a_i \rangle,$$

где $S = \sum_{i=1}^{n} a_i a_i^T, a_i \in \mathbb{R}^n, \det(S) \neq 0$

# Ортогональные матрицы

# Ортогональные (унитарные) матрицы

- Пусть $U$ - матрица размера $n \times n$, и $\|Uz\|_2 = \|z\|_2$ для всех $z$.

# Ортогональные (унитарные) матрицы

- Пусть $U$ - матрица размера $n \times n$, и $\|Uz\|_2 = \|z\|_2$ для всех $z$.
- Это может произойти **тогда и только тогда**, когда

$$U^T U = I_n,$$

где $I_n$ - единичная матрица $n \times n$.

# Ортогональные (унитарные) матрицы

- Пусть $U$ - матрица размера $n \times n$, и $\|Uz\|_2 = \|z\|_2$ для всех $z$.
- Это может произойти **тогда и только тогда**, когда

$$U^T U = I_n,$$

где $I_n$ - единичная матрица $n \times n$.

- Квадратная матрица размера $n \times n$ называется **ортогональной**, если

$$U^T U = U U^T = I_n,$$

что означает, что столбцы и строки ортогональной матрицы образуют ортонормированный базис в $\mathbb{R}^n$.

# Ортогональные (унитарные) матрицы

- Пусть $U$ - матрица размера $n \times n$, и $\|Uz\|_2 = \|z\|_2$ для всех $z$.
- Это может произойти **тогда и только тогда**, когда

$$U^T U = I_n,$$

где $I_n$ - единичная матрица $n \times n$.

- Квадратная матрица размера $n \times n$ называется **ортогональной**, если

$$U^T U = U U^T = I_n,$$

что означает, что столбцы и строки ортогональной матрицы образуют ортонормированный базис в $\mathbb{R}^n$.

- Для матрицы из комплексных чисел вводится аналогичное определение **унитарной матрицы**:

$$U^* U = U U^* = I_n,$$

где $U^*$ - эрмитово сопряжение матрицы $U$.

# Ортогональные (унитарные) матрицы

- Пусть $U$ - матрица размера $n \times n$, и $\|Uz\|_2 = \|z\|_2$ для всех $z$.
- Это может произойти **тогда и только тогда**, когда

$$U^T U = I_n,$$

где $I_n$ - единичная матрица $n \times n$.

- Квадратная матрица размера $n \times n$ называется **ортогональной**, если

$$U^T U = U U^T = I_n,$$

что означает, что столбцы и строки ортогональной матрицы образуют ортонормированный базис в $\mathbb{R}^n$.

- Для матрицы из комплексных чисел вводится аналогичное определение **унитарной матрицы**:

$$U^* U = U U^* = I_n,$$

где $U^*$ - эрмитово сопряжение матрицы $U$.

- Для прямоугольных матриц размера $m \times n$ $(n \neq m)$ только одно из равенств может выполняться

# Ортогональные (унитарные) матрицы

- Пусть $U$ - матрица размера $n \times n$, и $\|Uz\|_2 = \|z\|_2$ для всех $z$.
- Это может произойти **тогда и только тогда**, когда

$$U^T U = I_n,$$

где $I_n$ - единичная матрица $n \times n$.

- Квадратная матрица размера $n \times n$ называется **ортогональной**, если

$$U^T U = U U^T = I_n,$$

что означает, что столбцы и строки ортогональной матрицы образуют ортонормированный базис в $\mathbb{R}^n$.

- Для матрицы из комплексных чисел вводится аналогичное определение **унитарной матрицы**:

$$U^* U = U U^* = I_n,$$

где $U^*$ - эрмитово сопряжение матрицы $U$.

- Для прямоугольных матриц размера $m \times n$ $(n \neq m)$ только одно из равенств может выполняться
  - $U^T U = I_n$ - матрица $U$ называется **матрицей с ортогональными столбцами** для $m > n$

# Ортогональные (унитарные) матрицы

- Пусть $U$ - матрица размера $n \times n$, и $\|Uz\|_2 = \|z\|_2$ для всех $z$.
- Это может произойти **тогда и только тогда**, когда

$$U^T U = I_n,$$

где $I_n$ - единичная матрица $n \times n$.

- Квадратная матрица размера $n \times n$ называется **ортогональной**, если

$$U^T U = U U^T = I_n,$$

что означает, что столбцы и строки ортогональной матрицы образуют ортонормированный базис в $\mathbb{R}^n$.

- Для матрицы из комплексных чисел вводится аналогичное определение **унитарной матрицы**:

$$U^* U = U U^* = I_n,$$

где $U^*$ - эрмитово сопряжение матрицы $U$.

- Для прямоугольных матриц размера $m \times n$ $(n \neq m)$ только одно из равенств может выполняться
  - $U^T U = I_n$ - матрица $U$ называется **матрицей с ортогональными столбцами** для $m > n$
  - $U U^T = I_m$ - матрица $U$ называется **матрицей с ортогональными строками** для $m < n$

# Ортогональные матрицы

Важное свойство: **произведение двух ортогональных матриц является ортогональной матрицей:**

$$(UV)^T UV = V^T U^T UV = V^T V = I,$$

- Позже мы покажем, что существуют типы матриц (**отражения Хаусхолдера** и **вращения Гивенса**) композиция которых способна произвести любую унитарную матрицу

# Ортогональные матрицы

Важное свойство: **произведение двух ортогональных матриц является ортогональной матрицей:**

$$(UV)^T UV = V^T U^T UV = V^T V = I,$$

- Позже мы покажем, что существуют типы матриц (**отражения Хаусхолдера** и **вращения Гивенса**) композиция которых способна произвести любую унитарную матрицу
- Эта идея является основой некоторых алгоритмов, например, QR-разложения

# Ортогональная инвариантность $\|\cdot\|_2$ и $\|\cdot\|_F$ норм

- Для векторной 2-нормы мы уже видели, что $\|Uz\|_2 = \|z\|_2$ для любой ортогональной $U$.

# Ортогональная инвариантность $\|\cdot\|_2$ и $\|\cdot\|_F$ норм

- Для векторной 2-нормы мы уже видели, что $\|Uz\|_2 = \|z\|_2$ для любой ортогональной $U$.
- Можно показать, что ортогональные матрицы не изменяют матричные нормы $\|\cdot\|_2$ и $\|\cdot\|_F$, т.е. для любой квадратной $A$ и ортогональной $U$,$V$:

$$\|UAV\|_2 = \|A\|_2 \qquad \|UAV\|_F = \|A\|_F.$$

# Ортогональная инвариантность $\|\cdot\|_2$ и $\|\cdot\|_F$ норм

- Для векторной 2-нормы мы уже видели, что $\|Uz\|_2 = \|z\|_2$ для любой ортогональной $U$.
- Можно показать, что ортогональные матрицы не изменяют матричные нормы $\|\cdot\|_2$ и $\|\cdot\|_F$, т.е. для любой квадратной $A$ и ортогональной $U$,$V$:

$$\|UAV\|_2 = \|A\|_2 \qquad \|UAV\|_F = \|A\|_F.$$

- Для $\|\cdot\|_2$ это следует из определения операторной нормы и того факта, что векторная 2-норма является ортогонально инвариантной.

# Ортогональная инвариантность $\|\cdot\|_2$ и $\|\cdot\|_F$ норм

- Для векторной 2-нормы мы уже видели, что $\|Uz\|_2 = \|z\|_2$ для любой ортогональной $U$.
- Можно показать, что ортогональные матрицы не изменяют матричные нормы $\|\cdot\|_2$ и $\|\cdot\|_F$, т.е. для любой квадратной $A$ и ортогональной $U$,$V$:

$$\|UAV\|_2 = \|A\|_2 \qquad \|UAV\|_F = \|A\|_F.$$

- Для $\|\cdot\|_2$ это следует из определения операторной нормы и того факта, что векторная 2-норма является ортогонально инвариантной.
- Для $\|\cdot\|_F$ это следует из $\|A\|_F^2 = \mathrm{trace}(A^T A)$ и того факта, что $\mathrm{trace}(BC) = \mathrm{trace}(CB)$.

# Примеры ортогональных матриц

- Матрица вращения

# Примеры ортогональных матриц

- Матрица вращения
- Матрица перестановки

# Примеры ортогональных матриц

- Матрица вращения
- Матрица перестановки
- Матрица отражения Хаусхолдера

# Примеры ортогональных матриц

- Матрица вращения
- Матрица перестановки
- Матрица отражения Хаусхолдера
- Матрица вращения Гивенса

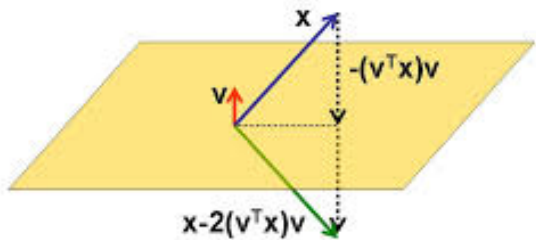
# Матрица отражения Хаусхолдера

Матрица отражения Хаусхолдера - это ортогональная матрица, которая используется для отражения вектора относительно гиперплоскости. Она имеет следующий вид:

$$H \equiv H(v) = I - 2vv^T,$$

где $v$ - вектор длины $n$ и $v^T v = 1$.

- Покажите, что $H$ - ортогональная матрица и $H^T = H$.



# Матрица отражения Хаусхолдера

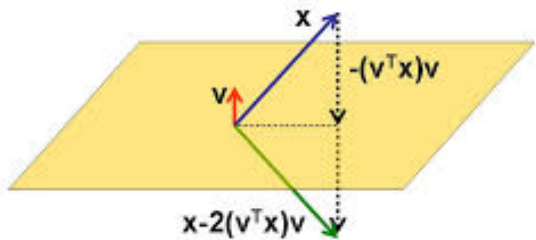
Матрица отражения Хаусхолдера - это ортогональная матрица, которая используется для отражения вектора относительно гиперплоскости. Она имеет следующий вид:

$$H \equiv H(v) = I - 2vv^T,$$

где $v$ - вектор длины $n$ и $v^T v = 1$.

- Покажите, что $H$ - ортогональная матрица и $H^T = H$.
- Покажите, что $H$ - отражение:

$$Hx = x - 2(v^T x)v$$



# Важное свойство матрицы отражения Хаусхолдера

Хорошее свойство матрицы отражения Хаусхолдера состоит в том, что она может обнулить все элементы вектора, кроме первого:

$$H \begin{bmatrix} \times \\ \times \\ \times \end{bmatrix} = \begin{bmatrix} \times \\ 0 \\ 0 \end{bmatrix}.$$

- Доказательство. Пусть $e_1 = (1, 0, \dots, 0)^T$, тогда мы хотим найти $v$ такой, что

$$Hx = x - 2(v^T x)v = \alpha e_1,$$

где $\alpha$ - неизвестная константа.

# Важное свойство матрицы отражения Хаусхолдера

Хорошее свойство матрицы отражения Хаусхолдера состоит в том, что она может обнулить все элементы вектора, кроме первого:

$$H \begin{bmatrix} \times \\ \times \\ \times \end{bmatrix} = \begin{bmatrix} \times \\ 0 \\ 0 \end{bmatrix}.$$

- Доказательство. Пусть $e_1 = (1, 0, \dots, 0)^T$, тогда мы хотим найти $v$ такой, что

$$Hx = x - 2(v^T x)v = \alpha e_1,$$

где $\alpha$ - неизвестная константа.

- Так как $\|\cdot\|_2$ является унитарно инвариантной, мы получаем

$$\|x\|_2 = \|Hx\|_2 = \|\alpha e_1\|_2 = |\alpha|.$$

и $\alpha = \pm\|x\|_2$

# Важное свойство матрицы отражения Хаусхолдера

Хорошее свойство матрицы отражения Хаусхолдера состоит в том, что она может обнулить все элементы вектора, кроме первого:

$$H \begin{bmatrix} \times \\ \times \\ \times \end{bmatrix} = \begin{bmatrix} \times \\ 0 \\ 0 \end{bmatrix}.$$

- Доказательство. Пусть $e_1 = (1, 0, \dots, 0)^T$, тогда мы хотим найти $v$ такой, что

$$Hx = x - 2(v^T x)v = \alpha e_1,$$

где $\alpha$ - неизвестная константа.

- Так как $\|\cdot\|_2$ является унитарно инвариантной, мы получаем

$$\|x\|_2 = \|Hx\|_2 = \|\alpha e_1\|_2 = |\alpha|.$$

и $\alpha = \pm\|x\|_2$

- Также мы можем выразить $v$ из $x - 2(v^T x)v = \alpha e_1$:

$$v = \frac{x - \alpha e_1}{2v^T x}$$

# Важное свойство матрицы отражения Хаусхолдера

Хорошее свойство матрицы отражения Хаусхолдера состоит в том, что она может обнулить все элементы вектора, кроме первого:

$$H \begin{bmatrix} \times \\ \times \\ \times \end{bmatrix} = \begin{bmatrix} \times \\ 0 \\ 0 \end{bmatrix}.$$

- Доказательство. Пусть $e_1 = (1, 0, \dots, 0)^T$, тогда мы хотим найти $v$ такой, что

$$Hx = x - 2(v^T x)v = \alpha e_1,$$

где $\alpha$ - неизвестная константа.

- Так как $\|\cdot\|_2$ является унитарно инвариантной, мы получаем

$$\|x\|_2 = \|Hx\|_2 = \|\alpha e_1\|_2 = |\alpha|.$$

и $\alpha = \pm\|x\|_2$

- Также мы можем выразить $v$ из $x - 2(v^T x)v = \alpha e_1$:

$$v = \frac{x - \alpha e_1}{2v^T x}$$

- Умножая последнее выражение на $x^T$ мы получаем

$$x^T x - 2(v^T x)x^T v = \alpha x_1;$$

или

$$\|x\|_2^2 - 2(v^T x)^2 = \alpha x_1.$$

# Важное свойство матрицы отражения Хаусхолдера

Хорошее свойство матрицы отражения Хаусхолдера состоит в том, что она может обнулить все элементы вектора, кроме первого:

$$H \begin{bmatrix} \times \\ \times \\ \times \end{bmatrix} = \begin{bmatrix} \times \\ 0 \\ 0 \end{bmatrix}.$$

- Доказательство. Пусть $e_1 = (1, 0, \dots, 0)^T$, тогда мы хотим найти $v$ такой, что

$$Hx = x - 2(v^T x)v = \alpha e_1,$$

где $\alpha$ - неизвестная константа.

- Так как $\|\cdot\|_2$ является унитарно инвариантной, мы получаем

$$\|x\|_2 = \|Hx\|_2 = \|\alpha e_1\|_2 = |\alpha|.$$

и $\alpha = \pm\|x\|_2$

- Также мы можем выразить $v$ из $x - 2(v^T x)v = \alpha e_1$:

$$v = \frac{x - \alpha e_1}{2v^T x}$$

- Умножая последнее выражение на $x^T$ мы получаем

$$x^T x - 2(v^T x)x^T v = \alpha x_1;$$

или

$$\|x\|_2^2 - 2(v^T x)^2 = \alpha x_1.$$

- Следовательно,

$$(v^T x)^2 = \frac{\|x\|_2^2 - \alpha x_1}{2}.$$

# Важное свойство матрицы отражения Хаусхолдера

Хорошее свойство матрицы отражения Хаусхолдера состоит в том, что она может обнулить все элементы вектора, кроме первого:

$$H \begin{bmatrix} \times \\ \times \\ \times \end{bmatrix} = \begin{bmatrix} \times \\ 0 \\ 0 \end{bmatrix}.$$

- Доказательство. Пусть $e_1 = (1, 0, \dots, 0)^T$, тогда мы хотим найти $v$ такой, что

$$Hx = x - 2(v^T x)v = \alpha e_1,$$

где $\alpha$ - неизвестная константа.

- Так как $\|\cdot\|_2$ является унитарно инвариантной, мы получаем

$$\|x\|_2 = \|Hx\|_2 = \|\alpha e_1\|_2 = |\alpha|.$$

и $\alpha = \pm\|x\|_2$

- Также мы можем выразить $v$ из $x - 2(v^T x)v = \alpha e_1$:

$$v = \frac{x - \alpha e_1}{2v^T x}$$

- Умножая последнее выражение на $x^T$ мы получаем

$$x^T x - 2(v^T x)x^T v = \alpha x_1;$$

или

$$\|x\|_2^2 - 2(v^T x)^2 = \alpha x_1.$$

- Следовательно,

$$(v^T x)^2 = \frac{\|x\|_2^2 - \alpha x_1}{2}.$$

- Таким образом, $v$ существует и равно

$$v = \frac{x \mp \|x\|_2 e_1}{2v^T x} = \frac{x \mp \|x\|_2 e_1}{\pm\sqrt{2(\|x\|_2^2 \mp \|x\|_2 x_1)}}.$$

# Алгоритм Хаусхолдера для построения QR-разложения

- Используя полученное свойство мы можем сделать произвольную матрицу $A$ нижней треугольной:

$$H_2 H_1 A = \begin{bmatrix} \times & \times & \times & \times \\ 0 & \times & \times & \times \\ 0 & 0 & \times & \times \\ 0 & 0 & \times & \times \\ 0 & 0 & \times & \times \end{bmatrix}$$

# Алгоритм Хаусхолдера для построения QR-разложения

- Используя полученное свойство мы можем сделать произвольную матрицу $A$ нижней треугольной:

$$H_2 H_1 A = \begin{bmatrix} \times & \times & \times & \times \\ 0 & \times & \times & \times \\ 0 & 0 & \times & \times \\ 0 & 0 & \times & \times \\ 0 & 0 & \times & \times \end{bmatrix}$$

- Затем находим $H_3 = \begin{bmatrix} I_2 & \\ & \widetilde{H}_3 \end{bmatrix}$ такую, что

$$\widetilde{H}_3 \begin{bmatrix} \times \\ \times \\ \times \end{bmatrix} = \begin{bmatrix} \times \\ 0 \\ 0 \end{bmatrix}.$$

# Алгоритм Хаусхолдера для построения QR-разложения

- Используя полученное свойство мы можем сделать произвольную матрицу $A$ нижней треугольной:

$$H_2 H_1 A = \begin{bmatrix} \times & \times & \times & \times \\ 0 & \times & \times & \times \\ 0 & 0 & \times & \times \\ 0 & 0 & \times & \times \\ 0 & 0 & \times & \times \end{bmatrix}$$

- Затем находим $H_3 = \begin{bmatrix} I_2 & \\ & \widetilde{H}_3 \end{bmatrix}$ такую, что

$$\widetilde{H}_3 \begin{bmatrix} \times \\ \times \\ \times \end{bmatrix} = \begin{bmatrix} \times \\ 0 \\ 0 \end{bmatrix}.$$

- Получаем

$$H_3 H_2 H_1 A = \begin{bmatrix} \times & \times & \times & \times \\ 0 & \times & \times & \times \\ 0 & 0 & \times & \times \\ 0 & 0 & 0 & \times \\ 0 & 0 & 0 & \times \end{bmatrix}$$

# Алгоритм Хаусхолдера для построения QR-разложения

- Используя полученное свойство мы можем сделать произвольную матрицу $A$ нижней треугольной:

$$H_2 H_1 A = \begin{bmatrix} \times & \times & \times & \times \\ 0 & \times & \times & \times \\ 0 & 0 & \times & \times \\ 0 & 0 & \times & \times \\ 0 & 0 & \times & \times \end{bmatrix}$$

- Затем находим $H_3 = \begin{bmatrix} I_2 & \\ & \widetilde{H}_3 \end{bmatrix}$ такую, что

$$\widetilde{H}_3 \begin{bmatrix} \times \\ \times \\ \times \end{bmatrix} = \begin{bmatrix} \times \\ 0 \\ 0 \end{bmatrix}.$$

- Получаем

$$H_3 H_2 H_1 A = \begin{bmatrix} \times & \times & \times & \times \\ 0 & \times & \times & \times \\ 0 & 0 & \times & \times \\ 0 & 0 & 0 & \times \\ 0 & 0 & 0 & \times \end{bmatrix}$$

- Аналогично находим $H_4$ и получаем верхнюю треугольную матрицу.

# Алгоритм Хаусхолдера для построения QR-разложения

- Используя полученное свойство мы можем сделать произвольную матрицу $A$ нижней треугольной:

$$H_2 H_1 A = \begin{bmatrix} \times & \times & \times & \times \\ 0 & \times & \times & \times \\ 0 & 0 & \times & \times \\ 0 & 0 & \times & \times \\ 0 & 0 & \times & \times \end{bmatrix}$$

- Затем находим $H_3 = \begin{bmatrix} I_2 & \\ & \widetilde{H}_3 \end{bmatrix}$ такую, что

$$\widetilde{H}_3 \begin{bmatrix} \times \\ \times \\ \times \end{bmatrix} = \begin{bmatrix} \times \\ 0 \\ 0 \end{bmatrix}.$$

- Получаем

$$H_3 H_2 H_1 A = \begin{bmatrix} \times & \times & \times & \times \\ 0 & \times & \times & \times \\ 0 & 0 & \times & \times \\ 0 & 0 & 0 & \times \\ 0 & 0 & 0 & \times \end{bmatrix}$$

- Аналогично находим $H_4$ и получаем верхнюю треугольную матрицу.
- Так как произведение ортогональных и обратная к ортогональной матрицы являются ортогональными матрицами, мы получаем **следствие:** (QR-разложение) Любая $A \in \mathbb{R}^{n \times m}$ может быть представлена как

$$A = QR,$$

где $Q$ - ортогональная и $R$ - верхняя треугольная. См. постер, каковы размеры $Q$ и $R$ для $n > m$ и $n < m$.

# Матрица вращения Гивенса (Якоби)

- Матрица вращения Гивенса (Якоби) - это ортогональная матрица, которая используется для вращения вектора в плоскости на угол $\alpha$. Она имеет следующий вид:

$$G = \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix}$$

Легко проверить, что $G^T G = GG^T = I$, то есть матрица является ортогональной.

# Матрица вращения Гивенса (Якоби)

- Матрица вращения Гивенса (Якоби) - это ортогональная матрица, которая используется для вращения вектора в плоскости на угол $\alpha$. Она имеет следующий вид:

$$G = \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix}$$

Легко проверить, что $G^T G = GG^T = I$, то есть матрица является ортогональной.

- Для общего случая размерности $n$ мы выбираем две координаты $(i, j)$ и выполняем вращение вектора $x$ только в этой плоскости:

$$x' = Gx,$$

где изменяются только $i$-я и $j$-я координаты:

$$x'_i = x_i \cos\alpha - x_j \sin\alpha, \quad x'_j = x_i \sin\alpha + x_j \cos\alpha,$$

при этом остальные $x_k$ остаются неизменными.

# Матрица вращения Гивенса (Якоби)

- Матрица вращения Гивенса (Якоби) - это ортогональная матрица, которая используется для вращения вектора в плоскости на угол $\alpha$. Она имеет следующий вид:

$$G = \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix}$$

Легко проверить, что $G^T G = GG^T = I$, то есть матрица является ортогональной.

- Для общего случая размерности $n$ мы выбираем две координаты $(i, j)$ и выполняем вращение вектора $x$ только в этой плоскости:

$$x' = Gx,$$

где изменяются только $i$-я и $j$-я координаты:

$$x'_i = x_i \cos\alpha - x_j \sin\alpha, \quad x'_j = x_i \sin\alpha + x_j \cos\alpha,$$

при этом остальные $x_k$ остаются неизменными.

- Чтобы обнулить $j$-ю координату вектора, выбираем угол $\alpha$ так, что:

$$\cos\alpha = \frac{x_i}{\sqrt{x_i^2 + x_j^2}}, \quad \sin\alpha = -\frac{x_j}{\sqrt{x_i^2 + x_j^2}}$$

# Матрица вращения Гивенса (Якоби)

- Матрица вращения Гивенса (Якоби) - это ортогональная матрица, которая используется для вращения вектора в плоскости на угол $\alpha$. Она имеет следующий вид:

$$G = \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix}$$

Легко проверить, что $G^T G = GG^T = I$, то есть матрица является ортогональной.

- Для общего случая размерности $n$ мы выбираем две координаты $(i, j)$ и выполняем вращение вектора $x$ только в этой плоскости:

$$x' = Gx,$$

где изменяются только $i$-я и $j$-я координаты:

$$x'_i = x_i \cos\alpha - x_j \sin\alpha, \quad x'_j = x_i \sin\alpha + x_j \cos\alpha,$$

при этом остальные $x_k$ остаются неизменными.

- Чтобы обнулить $j$-ю координату вектора, выбираем угол $\alpha$ так, что:

$$\cos\alpha = \frac{x_i}{\sqrt{x_i^2 + x_j^2}}, \quad \sin\alpha = -\frac{x_j}{\sqrt{x_i^2 + x_j^2}}$$

- Применяя последовательно матрицы Гивенса, можно привести матрицу к верхнетреугольному виду: для этого нужно $n - 1$ вращений, чтобы обнулить элементы под главной диагональю в каждом столбце.

# QR через вращения Гивенса

Также мы можем сделать матрицу верхнетреугольной с помощью вращений Гивенса:

$$\begin{bmatrix} \times & \times & \times \\ * & \times & \times \\ * & \times & \times \end{bmatrix} \to \begin{bmatrix} * & \times & \times \\ * & \times & \times \\ 0 & \times & \times \end{bmatrix} \to \begin{bmatrix} \times & \times & \times \\ 0 & * & \times \\ 0 & * & \times \end{bmatrix} \to \begin{bmatrix} \times & \times & \times \\ 0 & \times & \times \\ 0 & 0 & \times \end{bmatrix}$$

# Вращения Гивенса vs. отражения Хаусхолдера

- Отражения Хаусхолдера полезны для плотных матриц (сложность приблизительно в два раза меньше, чем для Якоби) и мы должны обнулить большое количество элементов.

# Вращения Гивенса vs. отражения Хаусхолдера

- Отражения Хаусхолдера полезны для плотных матриц (сложность приблизительно в два раза меньше, чем для Якоби) и мы должны обнулить большое количество элементов.
- Вращения Гивенса более подходят для разреженных матриц или параллельных вычислений, так как они действуют локально на элементах.