



Отражения Хаусхолдера. Вращения  
Гивенса. Ранг матрицы. Скелетон. SVD

Даня Меркулов

МФТИ. AI360

# 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  $U, V$ :

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

## 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  $U, V$ :

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

- 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  $U, V$ :

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

- 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  $\|A\|_F^2 = \text{trace}(A^T A)$  и того факта, что  $\text{trace}(BC) = \text{trace}(CB)$ .

# 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

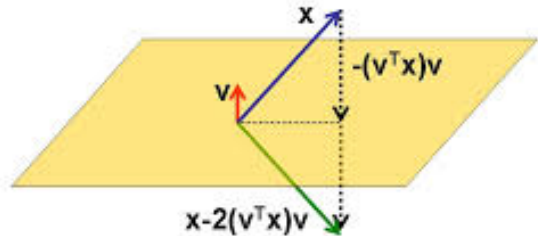
# Матрица отражения Хаусхолдера

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

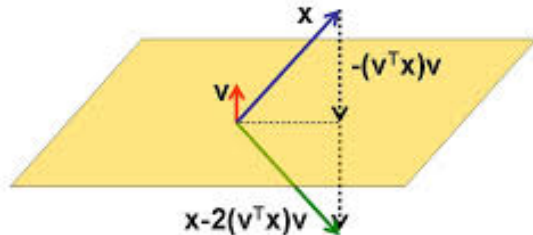
Матрица отражения Хаусхолдера - это ортогональная матрица, которая используется для отражения вектора относительно гиперплоскости. Она имеет следующий вид:

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

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

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

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

где  $\alpha$  - неизвестная константа.

или

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

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

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

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

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

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

где  $\alpha$  - неизвестная константа.

или

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

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

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

- Следовательно,

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

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

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

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

где  $\alpha$  - неизвестная константа.

или

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

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

- Следовательно,

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

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

и  $\alpha = \pm \|x\|_2$

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

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

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

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

## Алгоритм Хаусхолдера для построения QR-разложения

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

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

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

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

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

## Алгоритм Хаусхолдера для построения QR-разложения

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

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

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

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

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

- 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:
- Получаем

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

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

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

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

- 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  $n < m$ .



## QR-разложение матрицы



Для произвольной матрицы  $A \in \mathbb{R}^{m \times n}$  существует разложение вида:

$$A = QR,$$

где  $Q \in \mathbb{R}^{m \times m}$  - ортогональная матрица ( $Q^T Q = Q Q^T = I$ ), а  $R \in \mathbb{R}^{m \times n}$  - верхнетреугольная матрица. В случае, когда  $m > n$ , матрица  $Q$  может быть усечена до размера  $m \times n$  с сохранением ортогональности столбцов. Если зафиксировать все диагональные элементы матрицы  $R$ , то разложение становится единственным.

## 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  $G^T G = G G^T = I$ , 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  $G^T G = G G^T = I$ , 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  $G^T G = G G^T = I$ , 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  $G^T G = G G^T = I$ , 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:

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

# 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.



## Ранг матрицы

## Скелетное разложение матрицы

Простое, но очень интересное разложение - скелетное, которое может быть записано в двух формах:

$$A = UV^T \quad A = \hat{C}\hat{A}^{-1}\hat{R}$$

## Скелетное разложение матрицы

Простое, но очень интересное разложение - скелетное, которое может быть записано в двух формах:

$$A = UV^T \quad A = \hat{C}\hat{A}^{-1}\hat{R}$$

Последнее выражение отсылает к любопытному факту: вы можете случайным образом выбрать  $r$  линейно независимых столбцов матрицы и любых  $r$  линейно независимых строк матрицы и хранить только их, при этом имея возможность точно восстановить всю матрицу.

# Скелетное разложение матрицы

Простое, но очень интересное разложение - скелетное, которое может быть записано в двух формах:

$$A = UV^T \quad A = \hat{C}\hat{A}^{-1}\hat{R}$$

Последнее выражение отсылает к любопытному факту: вы можете случайным образом выбрать  $r$  линейно независимых столбцов матрицы и любых  $r$  линейно независимых строк матрицы и хранить только их, при этом имея возможность точно восстановить всю матрицу.

Применения скелетного разложения:

- Упрощение модели, сжатие данных и ускорение вычислений в численных методах: для матрицы ранга  $r$  с  $r \ll n, m$  достаточно хранить  $\mathcal{O}((n+m)r) \ll nm$  элементов.



Рис. 1: Иллюстрация скелетного разложения

# Скелетное разложение матрицы

Простое, но очень интересное разложение - скелетное, которое может быть записано в двух формах:

$$A = UV^T \quad A = \hat{C}\hat{A}^{-1}\hat{R}$$

Последнее выражение отсылает к любопытному факту: вы можете случайным образом выбрать  $r$  линейно независимых столбцов матрицы и любых  $r$  линейно независимых строк матрицы и хранить только их, при этом имея возможность точно восстановить всю матрицу.

Применения скелетного разложения:

- Упрощение модели, сжатие данных и ускорение вычислений в численных методах: для матрицы ранга  $r$  с  $r \ll n, m$  достаточно хранить  $\mathcal{O}((n+m)r) \ll nm$  элементов.
- Извлечение признаков в машинном обучении, где также известно как матричное разложение.

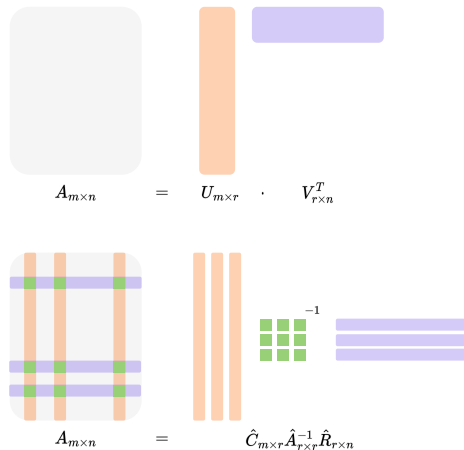


Рис. 1: Иллюстрация скелетного разложения

## LoRA: Low-Rank Adaptation of Large Language Models (arXiv:2106.09685)

Так как современные LLM слишком большие, чтобы поместиться в память обычной GPU, мы используем некоторые трюки, чтобы сделать потребление памяти меньше. Один из самых популярных трюков - LoRA (Low-Rank Adaptation of Large Language Models).

Предположим, что у нас есть матрица  $W \in \mathbb{R}^{d \times k}$  и мы хотим выполнить следующее обновление:

$$W = W_0 + \Delta W.$$

Основная идея LoRA - разложить обновление  $\Delta W$  на две низкоранговые матрицы:

$$W = W_0 + \Delta W = W_0 + BA, \quad B \in \mathbb{R}^{d \times r}, A \in \mathbb{R}^{r \times k}, \\ \text{rank}(A) = \text{rank}(B) = r \ll \min\{d, k\}.$$

Проверьте 📄 ноутбук для примера реализации LoRA.

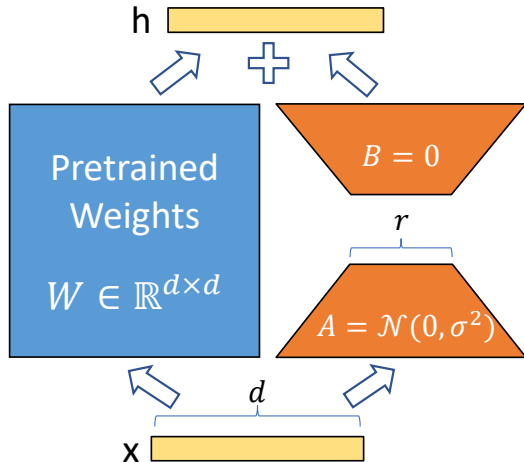


Рис. 2: Иллюстрация LoRA

# Сингулярное разложение матрицы

The diagram illustrates the Singular Value Decomposition (SVD) of a matrix  $A_{m \times n}$ . It is represented as the product of three matrices:  $U_{m \times m}$  (orange box),  $\Sigma_{m \times n}$  (grey box), and  $V_{n \times n}^*$  (purple box). The matrix  $\Sigma_{m \times n}$  is depicted with its diagonal elements  $\sigma_1, \sigma_2, \sigma_3, \dots, \sigma_r$  along the main diagonal.

$$A_{m \times n} = U_{m \times m} \Sigma_{m \times n} V_{n \times n}^*$$

Для любой матрицы  $A \in \mathbb{R}^{m \times n}$  существует разложение:

$$A = U \Sigma V^*,$$

где

- $U \in \mathbb{R}^{m \times m}$  - унитарная матрица левых сингулярных векторов

# Сингулярное разложение матрицы

The diagram illustrates the Singular Value Decomposition (SVD) of a matrix  $A$ . It consists of three colored boxes connected by an equals sign. The first box is light gray and contains  $A_{m \times n}$ . The second box is orange and contains  $U_{m \times m}$ . The third box is light gray and contains  $\Sigma_{m \times n}$ , which is represented by a diagonal matrix with singular values  $\sigma_1, \sigma_2, \sigma_3, \dots, \sigma_r$  on the diagonal. To the right of the third box is a purple box containing  $V_{n \times n}^*$ .

Для любой матрицы  $A \in \mathbb{R}^{m \times n}$  существует разложение:

$$A = U \Sigma V^*,$$

где

- $U \in \mathbb{R}^{m \times m}$  - унитарная матрица левых сингулярных векторов
  - $\Sigma \in \mathbb{R}^{m \times n}$  - диагональная матрица сингулярных чисел
- $$\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_{\min(m,n)} \geq 0$$



# Сингулярное разложение матрицы

The diagram illustrates the Singular Value Decomposition (SVD) of a matrix  $A_{m \times n}$ . It is represented as the product of three matrices:  $U_{m \times m}$  (orange box),  $\Sigma_{m \times n}$  (light gray box with diagonal elements  $\sigma_1, \sigma_2, \sigma_3, \dots, \sigma_r$  shown), and  $V_{n \times n}^*$  (purple box). The equation is shown as  $A_{m \times n} = U_{m \times m} \Sigma_{m \times n} V_{n \times n}^*$ .

Для любой матрицы  $A \in \mathbb{R}^{m \times n}$  существует разложение:

$$A = U \Sigma V^*,$$

где

- $U \in \mathbb{R}^{m \times m}$  - унитарная матрица левых сингулярных векторов
- $\Sigma \in \mathbb{R}^{m \times n}$  - диагональная матрица сингулярных чисел  
 $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_{\min(m,n)} \geq 0$
- $V \in \mathbb{R}^{n \times n}$  - унитарная матрица правых сингулярных векторов

# Сингулярное разложение матрицы

The diagram illustrates the SVD of a matrix  $A_{m \times n}$ . It is represented as the product of three matrices:  $U_{m \times m}$  (orange box),  $\Sigma_{m \times n}$  (light gray box), and  $V_{n \times n}^*$  (purple box). The matrix  $\Sigma_{m \times n}$  is shown with its diagonal elements  $\sigma_1, \sigma_2, \sigma_3, \dots, \sigma_r$  along the main diagonal.

Для любой матрицы  $A \in \mathbb{R}^{m \times n}$  существует разложение:

$$A = U \Sigma V^*,$$

где

- $U \in \mathbb{R}^{m \times m}$  - унитарная матрица левых сингулярных векторов
- $\Sigma \in \mathbb{R}^{m \times n}$  - диагональная матрица сингулярных чисел  
 $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_{\min(m,n)} \geq 0$
- $V \in \mathbb{R}^{n \times n}$  - унитарная матрица правых сингулярных векторов
- Сингулярные числа единственны. Если все сингулярные числа различны, то разложение единственно с точностью до унитарной диагональной матрицы  $D$ :  $U \Sigma V^* = U D \Sigma (V D)^* = U \Sigma V^*$ .