

Ортогональные матрицы





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

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

$$U^TU=I_n,$$

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

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

$$U^TU=I_n,$$

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

ullet Квадратная матрица размера n imes n называется **ортогональной**, если

$$U^TU=UU^T=I_n, \\$$

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



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

$$U^TU=I_n,$$

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:

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

где U^* - эрмитово сопряжение матрицы U.



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

$$U^TU=I_n,$$

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

ullet Квадратная матрица размера n imes 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:

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

где U^* - эрмитово сопряжение матрицы U.

ullet Для прямоугольных матриц размера m imes n (n
eq m) только одно из равенств может выполняться

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

$$U^TU=I_n,$$

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

ullet Квадратная матрица размера n imes 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:

$$U^*U = UU^* = I_n$$

где U^* - эрмитово сопряжение матрицы U.

- ullet Для прямоугольных матриц размера m imes n (n
 eq m) только одно из равенств может выполняться
 - ullet $U^TU=I_n$ матрица U называется **матрицей с ортогональными столбцами** для m>n

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

$$U^TU=I_n,$$

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

ullet Квадратная матрица размера n imes 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:

$$U^*U = UU^* = I_n$$

где U^{st} - эрмитово сопряжение матрицы U.

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



Ортогональные матрицы

Важное свойство: произведение двух ортогональных матриц является ортогональной матрицей:

$$(UV)^TUV = V^TU^TUV = V^TV = I, \\$$

• 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



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

- ullet 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 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.



- ullet 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:

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

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



• Матрица вращения



- Матрица вращения
- Матрица перестановки



- Матрица вращения
- Матрица перестановки
- Матрица отражения Хаусхолдера



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



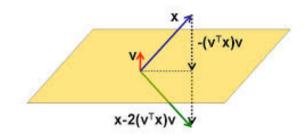
Матрица отражения Хаусхолдера

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 $v^Tv=1$.

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





Матрица отражения Хаусхолдера

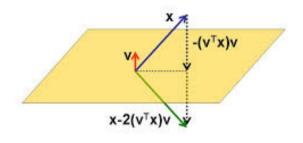
Матрица отражения Хаусхолдера - это ортогональная матрица, которая используется для отражения вектора относительно гиперплоскости. Она имеет следующий вид:

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

где v - вектор длины n и $v^Tv=1$.

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

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





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

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

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

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

где α - неизвестная константа.



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 $e_1 = (1, 0, ..., 0)^T$, тогда мы хотим найти v такой. что

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

где α - неизвестная константа.

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

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

и
$$\alpha = \pm \|x\|_2$$



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 α - 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 $x-2(v^Tx)v=\alpha e_1$:

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





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 $e_1 = (1, 0, ..., 0)^T$, тогда мы • Умножая последнее выражение на x^T мы хотим найти v такой. что

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

где α - неизвестная константа.

• Так как $\|\cdot\|_2$ является унитарно инвариантной. мы получаем

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

и
$$\alpha = \pm \|x\|_2$$

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

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

получаем

или

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

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

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 $e_1 = (1, 0, ..., 0)^T$, тогда мы • Умножая последнее выражение на x^T мы хотим найти v такой. что

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

где lpha - неизвестная константа.

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

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

и
$$\alpha = \pm \|x\|_2$$

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

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

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

получаем

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

или

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

• Следовательно.

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

Ортогональные матрицы

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 $e_1 = (1, 0, ..., 0)^T$, тогда мы • Умножая последнее выражение на x^T мы хотим найти v такой, что

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

где lpha - неизвестная константа.

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

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

и
$$\alpha = \pm \|x\|_2$$

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

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

получаем

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

или

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

• Следовательно.

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

• Таким образом, v существует и равно

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

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

$$H_2H_1A = \begin{bmatrix} \times & \times & \times & \times \\ 0 & \times & \times & \times \\ 0 & 0 & \times & \times \\ 0 & 0 & \times & \times \\ 0 & 0 & \times & \times \end{bmatrix}$$



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

$$H_2H_1A = \begin{bmatrix} \times & \times & \times & \times \\ 0 & \times & \times & \times \\ 0 & 0 & \times & \times \\ 0 & 0 & \times & \times \\ 0 & 0 & \times & \times \end{bmatrix}$$

ullet Затем находим $H_3 = egin{bmatrix} I_2 & & & \ \widetilde{H}_2 & & \widetilde{H}_2 \end{bmatrix}$ такую, что

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



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

$$H_2H_1A = \begin{bmatrix} \times & \times & \times & \times \\ 0 & \times & \times & \times \\ 0 & 0 & \times & \times \\ 0 & 0 & \times & \times \\ 0 & 0 & \times & \times \end{bmatrix}$$

ullet Затем находим $H_3 = egin{bmatrix} I_2 & & & \ \widetilde{H}_2 & & \ \widetilde{H}_2 & & \ \end{bmatrix}$ такую, что

$$\widetilde{H}_{3} \begin{bmatrix} \times \\ \times \\ \times \end{bmatrix} = \begin{bmatrix} \times \\ 0 \\ 0 \end{bmatrix}.$$

• Получаем

$$H_{3}H_{2}H_{1}A = \begin{bmatrix} \times & \times & \times & \times \\ 0 & \times & \times & \times \\ 0 & 0 & \times & \times \\ 0 & 0 & 0 & \times \\ 0 & 0 & 0 & \times \end{bmatrix}$$



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

$$H_2H_1A = \begin{bmatrix} \times & \times & \times & \times \\ 0 & \times & \times & \times \\ 0 & 0 & \times & \times \\ 0 & 0 & \times & \times \\ 0 & 0 & \times & \times \end{bmatrix}$$

ullet Затем находим $H_3 = egin{bmatrix} I_2 & & \ & \widetilde{H}_3 \end{bmatrix}$ 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 ${\cal H}_4$ и получаем верхнюю треугольную матрицу.



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

$$H_2H_1A = \begin{bmatrix} \times & \times & \times & \times \\ 0 & \times & \times & \times \\ 0 & 0 & \times & \times \\ 0 & 0 & \times & \times \\ 0 & 0 & \times & \times \end{bmatrix}$$

ullet Затем находим $H_3 = egin{bmatrix} I_2 & & \ \widetilde{H}_2 \end{bmatrix}$ 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 $H_{\scriptscriptstyle A}$ 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) - это ортогональная матрица, которая используется для вращения вектора в плоскости на угол α . Она имеет следующий вид:

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

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



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

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

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

ullet Для общего случая размерности n мы выбираем две координаты (i,j) и выполняем вращение вектора xтолько 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) - это ортогональная матрица, которая используется для вращения вектора в плоскости на угол α . Она имеет следующий вид:

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

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

ullet Для общего случая размерности n мы выбираем две координаты (i,j) и выполняем вращение вектора xтолько 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.

ullet Чтобы обнулить i-ю координату вектора, выбираем угол lpha 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) - это ортогональная матрица, которая используется для вращения вектора в плоскости на угол α . Она имеет следующий вид:

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

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

ullet Для общего случая размерности n мы выбираем две координаты (i,j) и выполняем вращение вектора xтолько 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.

ullet Чтобы обнулить i-ю координату вектора, выбираем угол lpha 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



Ранг матрицы

େ ପ 🕈

Скелетное разложение





Сингулярное разложение





Приложения





Image compression





Метод главных компонент





LoRa



Приложения

