

Векторы и матрицы



Вектор и операции над ним. Матрицы и матричные операции. Типы матриц. Собственные векторы и собственные значения. Приближение матрицей меньшего ранга. Сингулярное разложение и низкоранговое приближение.

Даниил Корбут

Специалист по Анализу Данных



Даниил Корбут
DL Researcher
Insilico Medicine, Inc

Окончил магистратуру ФИБТ
МФТИ в 2020.

Работал в Statsbot и Яндекс.
Алиса.

Сейчас в Insilico Medicine, Inc,
занимаюсь генерацией
активных молекул и
исследованиями старения с
помощью DL.

Скаляр

Скаляр — число.

Число может являться целым натуральным и записывается как:

$$s \in \mathbb{N}$$

Число может быть десятичным и записываться как:

$$s \in \mathbb{R}$$

Векторы

Вектор — упорядоченный конечный список чисел.
Вектора обычно записываются как вертикальный список,
например:

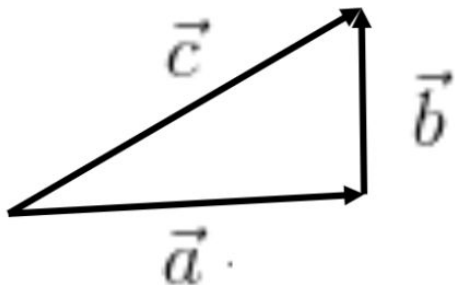
$$\begin{bmatrix} -1.1 \\ 0.0 \\ 3.6 \\ -7.2 \end{bmatrix} \quad \begin{pmatrix} -1.1 \\ 0.0 \\ 3.6 \\ -7.2 \end{pmatrix}$$

Вектор может быть записан также в следующем виде:

$$(-1.1, 0.0, 3.6, -7.2)$$

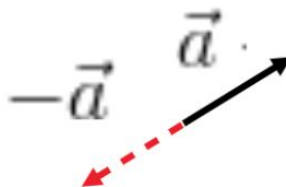
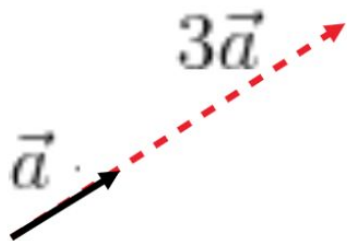
Операции с векторами

Сложение векторов:



$$\vec{a} + \vec{b} = \vec{c}$$

Умножение вектора на число:



Длина вектора

Длина вектора может быть подсчитана по формуле евклидова расстояния:

$$|x| = \sqrt{\sum_i (|x_i|^2)}$$

где p — размерность вектора.

Для двумерного вектора данная формула становится следующего вида и называется L^2 нормой:

$$|a| = \sqrt{(x^2 + y^2)}$$

где a — вектор с координатами (x, y) .

Скалярное произведение векторов

Скалярное произведение векторов (dot product по англ.) - это скаляр (число), полученное в результате перемножения длин векторов на косинус угла между ними.

$$\vec{a} \cdot \vec{b} = |\vec{a}| \cdot |\vec{b}| \cdot \cos(\alpha)$$

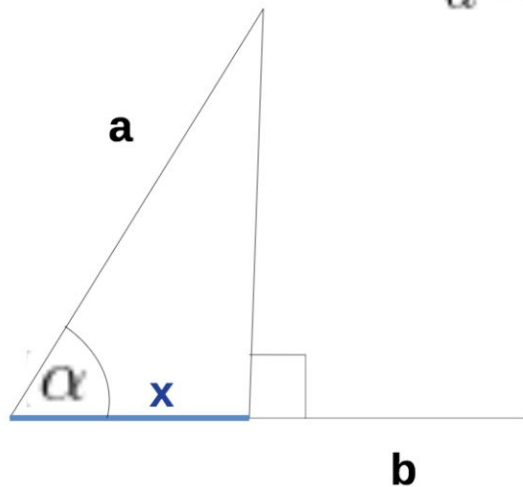
Если известны координаты векторов, то скалярное произведение можно посчитать по формуле:

$$\vec{a} \cdot \vec{b} = x_a \cdot x_b + y_a \cdot y_b$$

где $\vec{a}(x_a; y_a)$ и $\vec{b}(x_b; y_b)$ вектора в двумерном пространстве

Проекция одного вектора на другой

Длина вектора x , полученного в результате проекции вектора a на вектор b , равна делению скалярного произведения вектора \mathbf{a} на вектор \mathbf{b} на длину b .



$$\vec{a} \cdot \vec{b} = |\vec{a}| \cdot |\vec{b}| \cdot \cos(\alpha)$$

$$\cos(\alpha) = \frac{|x|}{|a|}$$

$$|x| = \cos(\alpha) \cdot |a|$$

$$\cos(\alpha) \cdot |a| = \frac{\vec{a} \cdot \vec{b}}{|b|}$$

$$|x| = \frac{\vec{a} \cdot \vec{b}}{|b|}$$

Матрица

$$M = \begin{bmatrix} A_{1,1} & A_{1,2} & A_{1,3} \\ A_{2,1} & A_{2,2} & A_{2,3} \end{bmatrix}$$

Матрица — математический объект, записываемый в виде прямоугольной таблицы.

Диагональная матрица — квадратная матрица, все элементы которой, кроме диагональных, равны 0.

$$M = \begin{bmatrix} A_{1,1} & 0 & 0 \\ 0 & A_{2,2} & 0 \\ 0 & 0 & A_{3,3} \end{bmatrix}$$

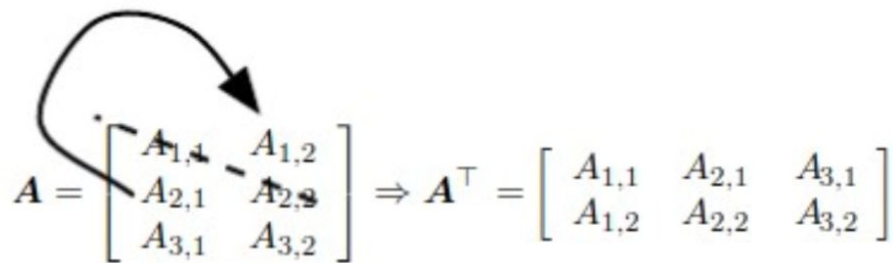
$$M = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Единичная матрица — квадратная матрица, все диагональные элементы которой равны 1, остальные элементы равны 0.

Транспонирование матрицы

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}^T = \begin{bmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{bmatrix}$$

Транспонирование матрицы
— это замена строк на
столбцы.


$$A = \begin{bmatrix} A_{1,1} & A_{1,2} \\ A_{2,1} & A_{2,2} \\ A_{3,1} & A_{3,2} \end{bmatrix} \Rightarrow A^T = \begin{bmatrix} A_{1,1} & A_{2,1} & A_{3,1} \\ A_{1,2} & A_{2,2} & A_{3,2} \end{bmatrix}$$

Транспонирование матрицы можно рассматривать как отображение матрицы относительно главной диагонали.

Сложение и умножение матрицы на скаляр

Сложение.

$$\begin{bmatrix} 2 & 3 \\ 2 & 3 \\ 2 & 3 \end{bmatrix} + 1 = \begin{bmatrix} 3 & 4 \\ 3 & 4 \\ 3 & 4 \end{bmatrix}$$

Умножение на скаляр.

$$\begin{bmatrix} 2 & 3 \\ 2 & 3 \\ 2 & 3 \end{bmatrix} \cdot 5 = \begin{bmatrix} 10 & 15 \\ 10 & 15 \\ 10 & 15 \end{bmatrix}$$

Обратная матрица

Обратная матрица к данной — это матрица при перемножении которой с текущей матрицей получается единичная матрица.

$$\mathbf{A}\mathbf{A}^{-1} = \mathbf{I}$$

Например:

$$B = \begin{bmatrix} 2 & 5 & 7 \\ 6 & 3 & 4 \\ 5 & -2 & -3 \end{bmatrix} \quad B^{-1} = \begin{bmatrix} 1 & -1 & 1 \\ -38 & 41 & -34 \\ 27 & -29 & 24 \end{bmatrix}$$
$$B \cdot B^{-1} = \begin{bmatrix} 2 & 5 & 7 \\ 6 & 3 & 4 \\ 5 & -2 & -3 \end{bmatrix} \cdot \begin{bmatrix} 1 & -1 & 1 \\ -38 & 41 & -34 \\ 27 & -29 & 24 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \mathbf{I}$$

Перемножение матриц

Даны 2 матрицы: A и B. Умножение матрицы A на B можно выполнить, если количество столбцов матрицы A равно количеству строк матрицы B.

$$A \cdot B = \begin{pmatrix} 1 & 2 & 2 \\ 3 & 1 & 1 \end{pmatrix} \begin{pmatrix} 4 & 2 \\ 3 & 1 \\ 1 & 5 \end{pmatrix} = \begin{pmatrix} 1 \cdot 4 + 2 \cdot 3 + 2 \cdot 1 & 1 \cdot 2 + 2 \cdot 1 + 2 \cdot 5 \\ 3 \cdot 4 + 1 \cdot 3 + 1 \cdot 1 & 3 \cdot 2 + 1 \cdot 1 + 1 \cdot 5 \end{pmatrix} = \begin{pmatrix} 12 & 14 \\ 16 & 12 \end{pmatrix}$$

$$B \cdot A = \begin{pmatrix} 4 & 2 \\ 3 & 1 \\ 1 & 5 \end{pmatrix} \begin{pmatrix} 1 & 2 & 2 \\ 3 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 4 \cdot 1 + 2 \cdot 3 & 4 \cdot 2 + 2 \cdot 1 & 4 \cdot 2 + 2 \cdot 1 \\ 3 \cdot 1 + 1 \cdot 3 & 3 \cdot 2 + 1 \cdot 1 & 3 \cdot 2 + 1 \cdot 1 \\ 1 \cdot 1 + 5 \cdot 3 & 1 \cdot 2 + 5 \cdot 1 & 1 \cdot 2 + 5 \cdot 1 \end{pmatrix} = \begin{pmatrix} 10 & 10 & 10 \\ 6 & 7 & 7 \\ 16 & 7 & 7 \end{pmatrix}$$

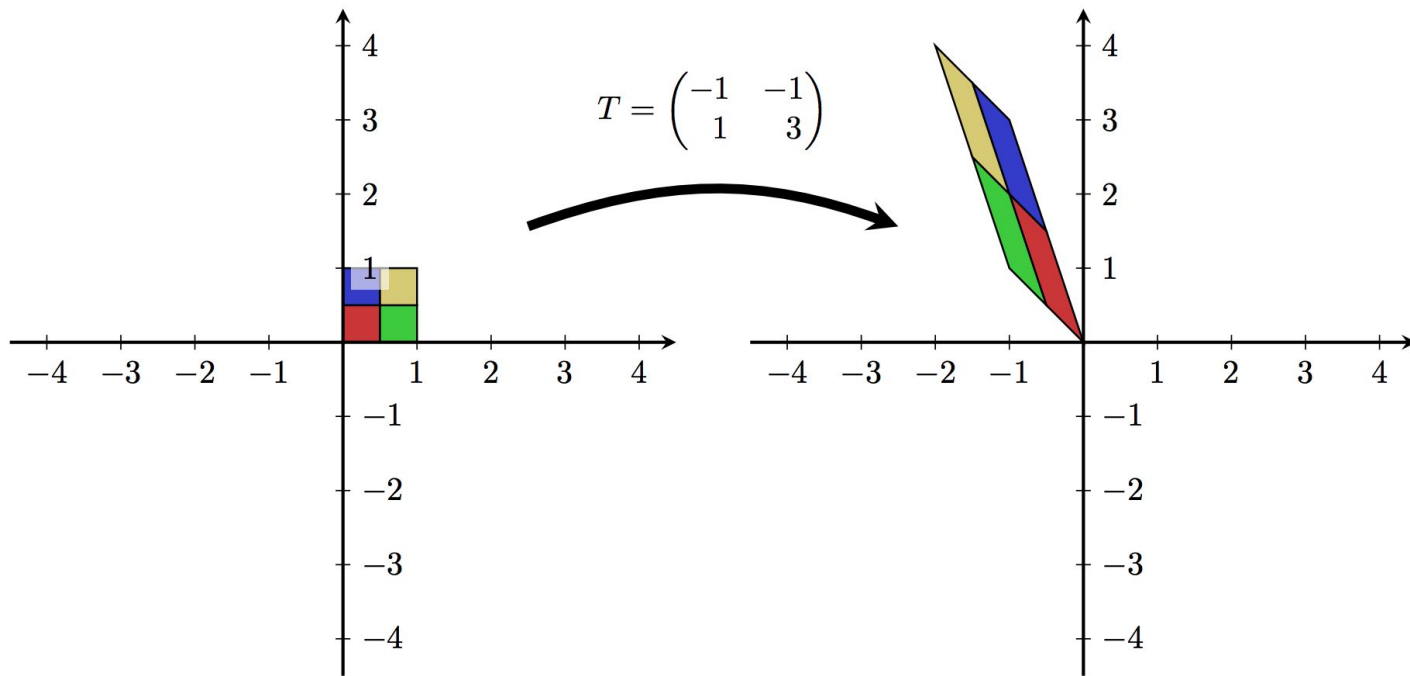
Типы матриц

$$A_{m \times n} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

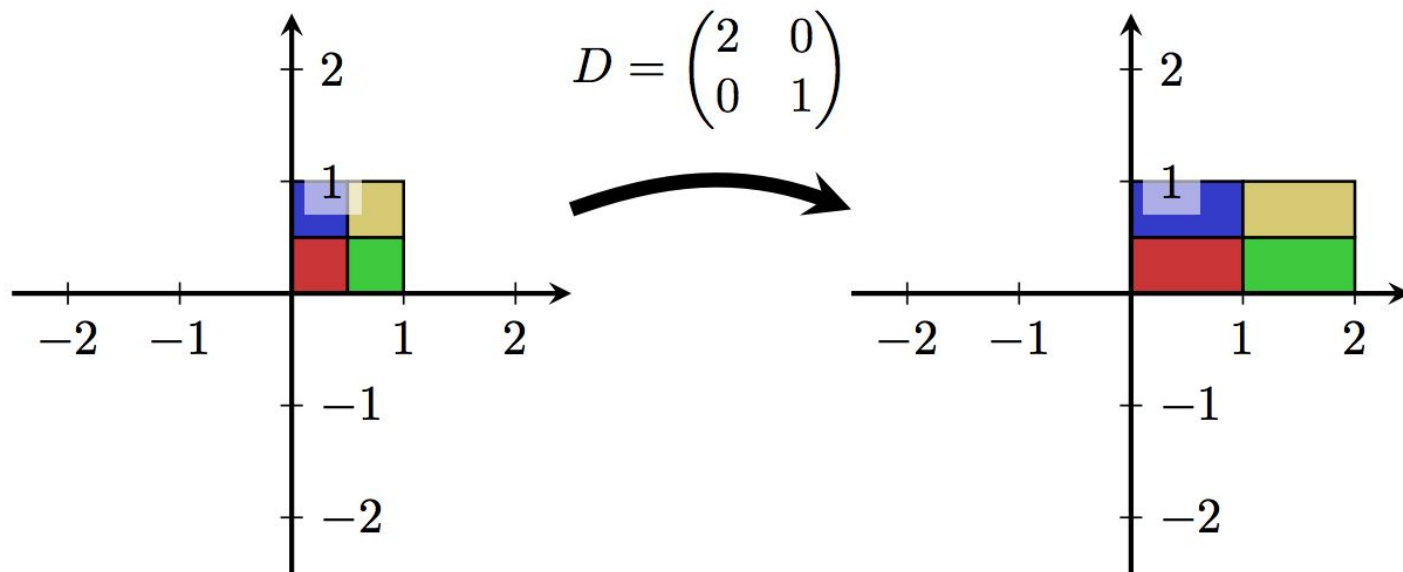
1. $m=n \Rightarrow$ квадратная, иначе прямоугольная
2. $m = 1 \Rightarrow$ матрица-строка
3. $n=1 \Rightarrow$ матрица-столбец
4. нулевая матрица, если все элементы = 0
5. диагональная (единичная)
6. треугольная(нижнетреугольная, верхнетреугольная)
7. ортогональная

$$AA^T = A^T A = E$$

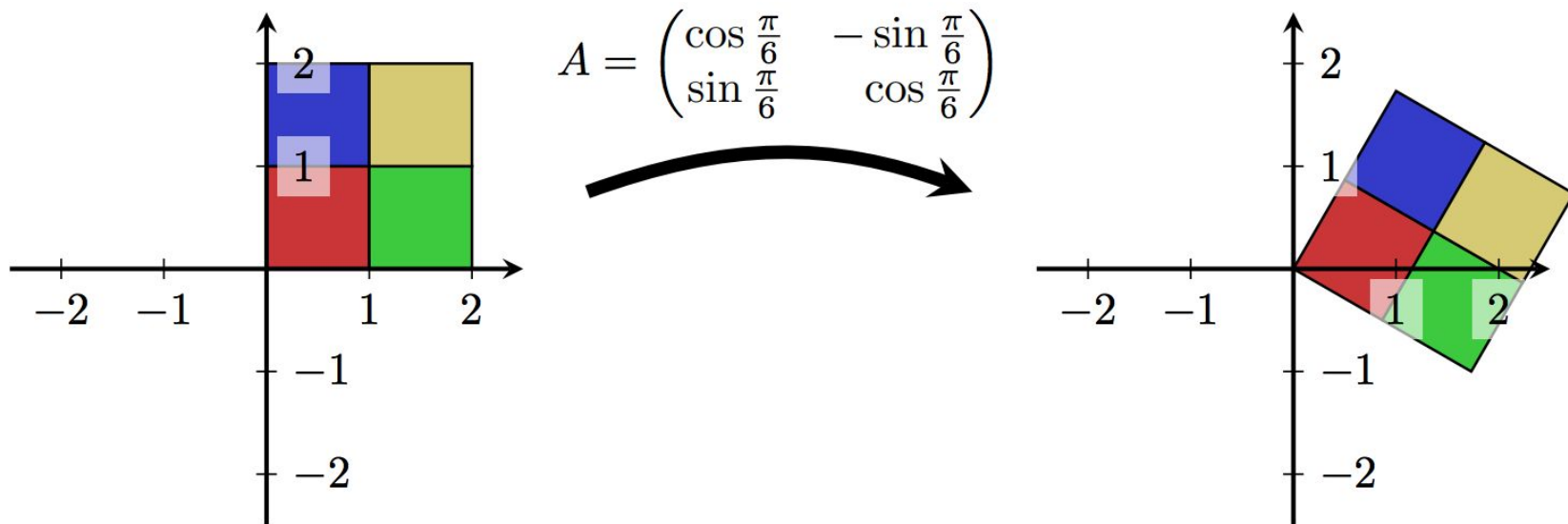
Типы матриц (преобразование пространства)



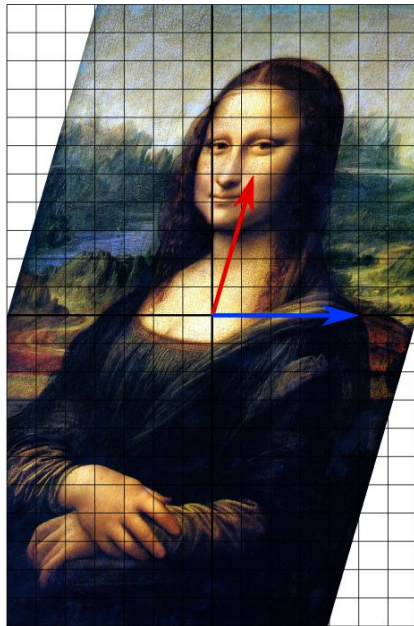
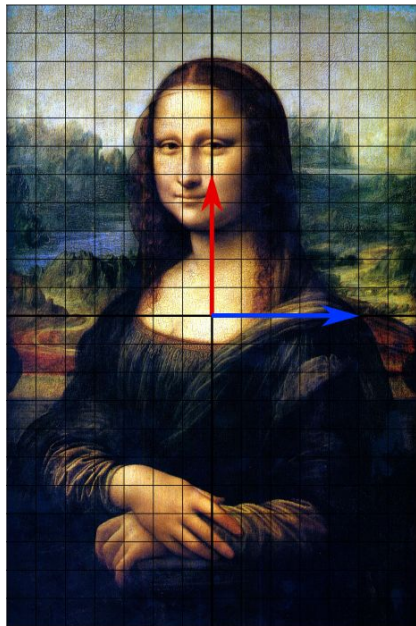
Типы матриц (преобразование пространства)



Типы матриц (преобразование пространства)



Собственные векторы и собственные значения



<https://ru.wikipedia.org>

Собственный вектор преобразования A

$$AX = \lambda X$$

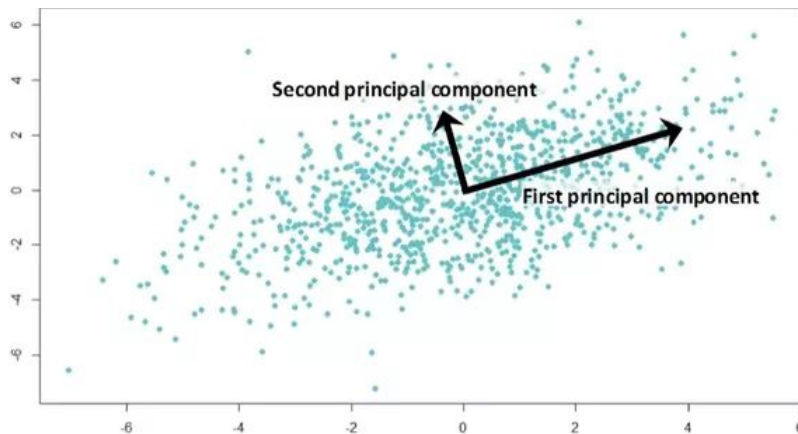
X - собственный вектор (ненулевой!)
 λ - собственное значение

!

У матрицы $n \times n$ не более n
собственных значений

Собственные векторы и собственные значения: применение

1. Собственные векторы - направления, в которых матрица лишь растягивает или сжимает векторы, но не поворачивает
2. Показывают направления наибольшего изменения
3. Возникают при уменьшении размера матрицы



<https://medium.com/analytics-vidhya/principal-component-analysis-pca-with-code-on-mnist-dataset-da7de0d07c22>

Собственные вектора и собственные значения (поиск)

$$A - \lambda E$$

характеристическая матрица матрицы **A**

$$|A - \lambda E|$$

характеристический многочлен матрицы **A**

$$|A - \lambda E| = 0$$

характеристическое уравнение матрицы **A**



Теорема. Собственными числами матрицы **A** являются корни характеристического уравнения матрицы **A** и только они.

Собственные вектора и собственные значения (пример)

$$A = \begin{pmatrix} -1 & -6 \\ 2 & 6 \end{pmatrix} \rightarrow \begin{vmatrix} -1-\lambda & -6 \\ 2 & 6-\lambda \end{vmatrix} = 0$$



$$\lambda^2 - 5\lambda + 6 = 0$$



$$\lambda_1 = \frac{5-1}{2} = 2$$

$$\lambda_2 = \frac{5+1}{2} = 3$$



$$\begin{cases} -4x - 6y = 0 \\ 2x + 3y = 0 \end{cases} \quad x = -\frac{3}{2}y$$

$$\begin{cases} -3x - 6y = 0 \\ 2x + 4y = 0 \end{cases} \Rightarrow x = -2y$$

Ответ: собственные числа: $\lambda_1 = 2$, $\lambda_2 = 3$, собственные векторы: $\vec{u}_1 = \begin{pmatrix} 2 \\ -1 \end{pmatrix}$, $\vec{u}_2 = \begin{pmatrix} 3 \\ -2 \end{pmatrix}$

Матричные разложения (спектральное разложение)

Разложение матрицы - представление в виде произведения некоторых других, обладающих интересными свойствами.

Пример: спектральное разложение X

$$X = S^T \cdot D \cdot S$$



X - **симметричная**, S - **ортогональная**, D - **диагональная** из собственных значений X .

Часто встречаются квадратичные формы

$$f(y) = y^T X y$$

с помощью спектрального разложения приводим к более простому виду:

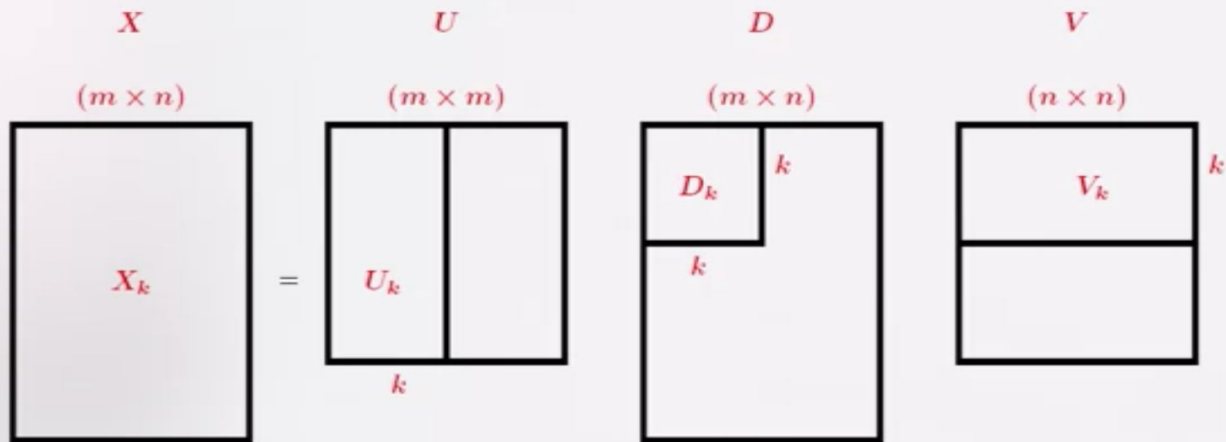
$$f(y) = y^T \cdot S^T \cdot D \cdot S \cdot y = (S \cdot y)^T \cdot D \cdot (S \cdot y) = z^T \cdot D \cdot z = \sum_{i=1}^n \lambda_i z_i^2,$$

Матричные разложения (сингулярное разложение)

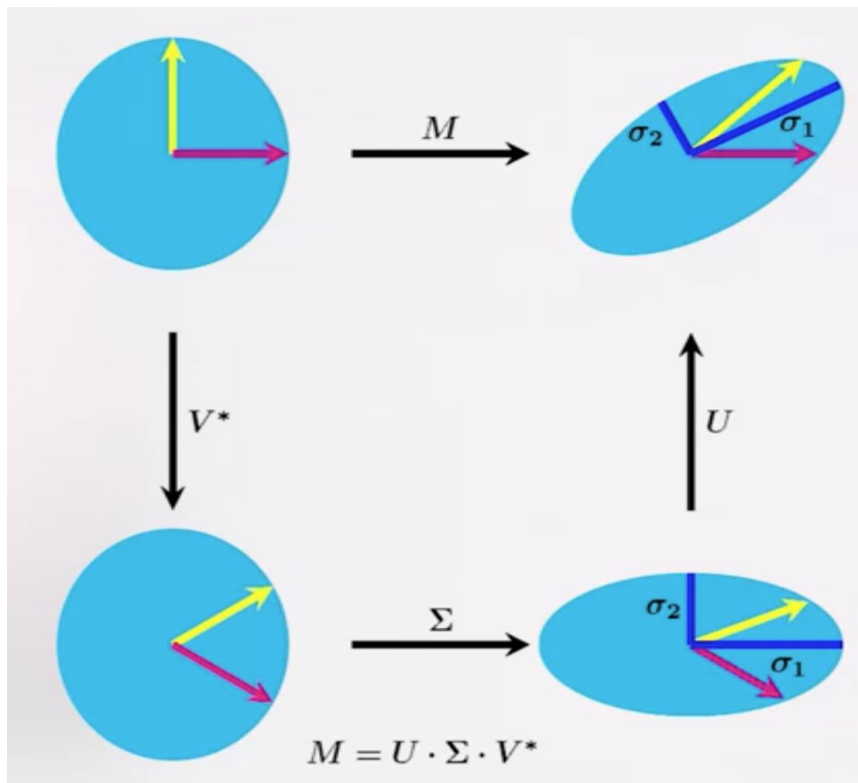
Но это была симметричная матрица, что в случае произвольной?

$$\triangleright X = UDV$$

$\triangleright U, V$ — ортогональные, D —
диагональная



Матричные разложения (сингулярное разложение)



Сингулярное разложение представляет линейное преобразование в виде композиции: вращения, растяжения по осям, вращения.

Матричные разложения (приближение матрицей меньшего ранга)

Что такое ранг матрицы?

Матричные разложения (приближение матрицей меньшего ранга)

- Матрица задаёт отображение, ранг в какой-то степени мера “сложности” отображения
- Ранг - максимальное количество линейно независимых столбцов или строк
- Ранг - максимальный размер подматрицы с ненулевым определителем



$\text{rank}(X) \leq \min(n, m)$, если X - матрица $m \times n$

Пусть $X = AB$, A размера (m, k) , B размера (k, n)

Пусть также $k < m$, $k < n$

Что можно сказать о ранге X ?

Матричные разложения (приближение матрицей меньшего ранга)

- Матрица задаёт отображение, ранг в какой-то степени мера “сложности” отображения
- Ранг - максимальное количество линейно независимых столбцов или строк
- Ранг - максимальный размер подматрицы с ненулевым определителем



$\text{rank}(X) \leq \min(n, m)$, если X - матрица $m \times n$

Пусть $X = AB$, A размера (m, k) , B размера (k, n)

Пусть также $k < m$, $k < n$

Что можно сказать о ранге X ?

Верно! $\text{rank}(X) \leq k$

Матричные разложения (приближение матрицей меньшего ранга)

Зачем приближать матрицу матрицей меньшего ранга?

Мы предполагаем, что матрица преобразования X на самом деле более простая.

Что значит приблизить

$$\triangleright X \approx X' = UV^T$$

$$\triangleright U - m \times k, V - n \times k$$

Просто наилучшее приближение по
норме: $\|X - UV^T\| \rightarrow \min$

Матричные разложения (приближение матрицей меньшего ранга)

Что значит приблизить

$$\triangleright X \approx X' = UV^T$$

$$\triangleright U - m \times k, V - n \times k$$

Просто наилучшее приближение по
норме: $\|X - UV^T\| \rightarrow \min$

$$\|X\|_F = \sqrt{\sum_{i,j} x_{ij}^2}$$

Итоговая задача выглядит так:

$$U, V = \underset{U \in \mathbb{R}^{m \times k}, V \in \mathbb{R}^{n \times k}}{\operatorname{argmin}} \sum_{i,j} (x_{ij} - u_i^T v_j)^2$$

Матричные разложения (пример применения)

- › Пусть X – матрица признаков объектов
- › Тогда U – матрица новых признаков
- › При $k < n$ преобразование признаков понижает размерность пространства
- › По U с максимальной возможной точностью восстанавливаются исходные признаки X

Матричные разложения (пример применения)

- › Пусть X – матрица с оценками x_{ij} , поставленными пользователем i фильму j
- › Некоторые значения матрицы неизвестны
- › $x_{ij} \approx \widehat{x_{ij}} = u_i v_j$, где u_i отражает интересы пользователя, а v_j – признаковое описание фильма
- › Идея: настроим u_i и v_j на известных x_{ij} , а неизвестные спрогнозируем
- › Будем рекомендовать фильмы, для которых спрогнозирована высокая оценка

Что делать с пропущенными значениями?

Матричные разложения (пример применения)

- › Пусть X – матрица с оценками x_{ij} , поставленными пользователем i фильму j
- › Некоторые значения матрицы неизвестны
- › $x_{ij} \approx \widehat{x_{ij}} = u_i v_j$, где u_i отражает интересы пользователя, а v_j – признаковое описание фильма
- › Идея: настроим u_i и v_j на известных x_{ij} , а неизвестные спрогнозируем
- › Будем рекомендовать фильмы, для которых спрогнозирована высокая оценка

$$U, V = \underset{U \in \mathbb{R}^{m \times k}, V \in \mathbb{R}^{n \times k}}{\operatorname{argmin}} \sum_{i,j: x_{ij} \neq 0} (x_{ij} - u_i^T v_j)^2$$

Спасибо за внимание!