Продвинутая линейная алгебра

Типы матриц. Собственные вектора и собственные значения. Матричные разложения. Приближение матрицей меньшего ранга. Сингулярное разложение и низкоранговое приближение.

Даниил Корбут
Специалист по Анализу Данных

НЕТОЛОГИЯ

Часть 1: Лекция 3

Даниил Корбут
DL Researcher
Insilico Medicine, Inc

Окончил бакалавриат ФИВТ МФТИ (Анализ данных) в 2018г
Учусь на 2-м курсе магистратуры ФИВТ МФТИ
Работал в Statsbot и Яндекс. Алиса.
Сейчас в Insilico Medicine, Inc, занимаюсь генерацией активных молекул и исследованиеми старения с помощью DL.

НЕТОЛОГИЯ

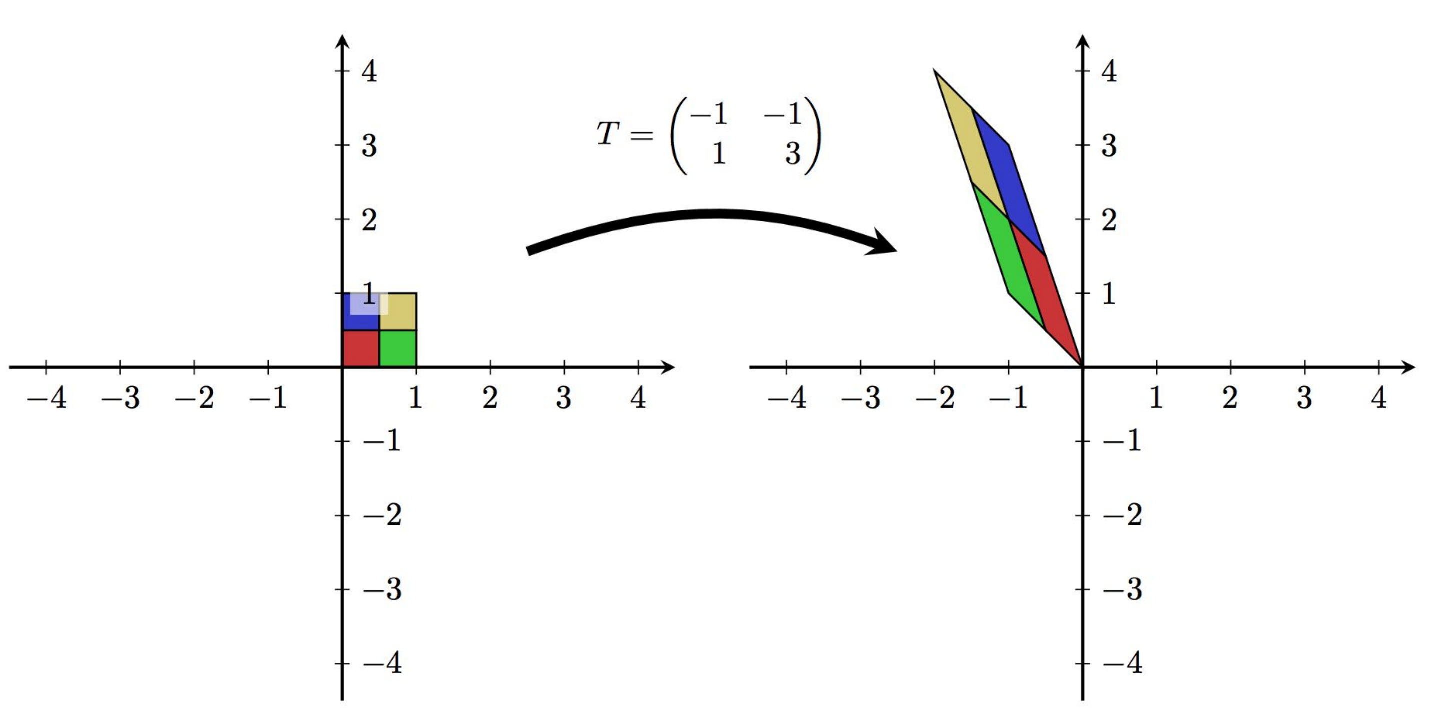
Типы матриц

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

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

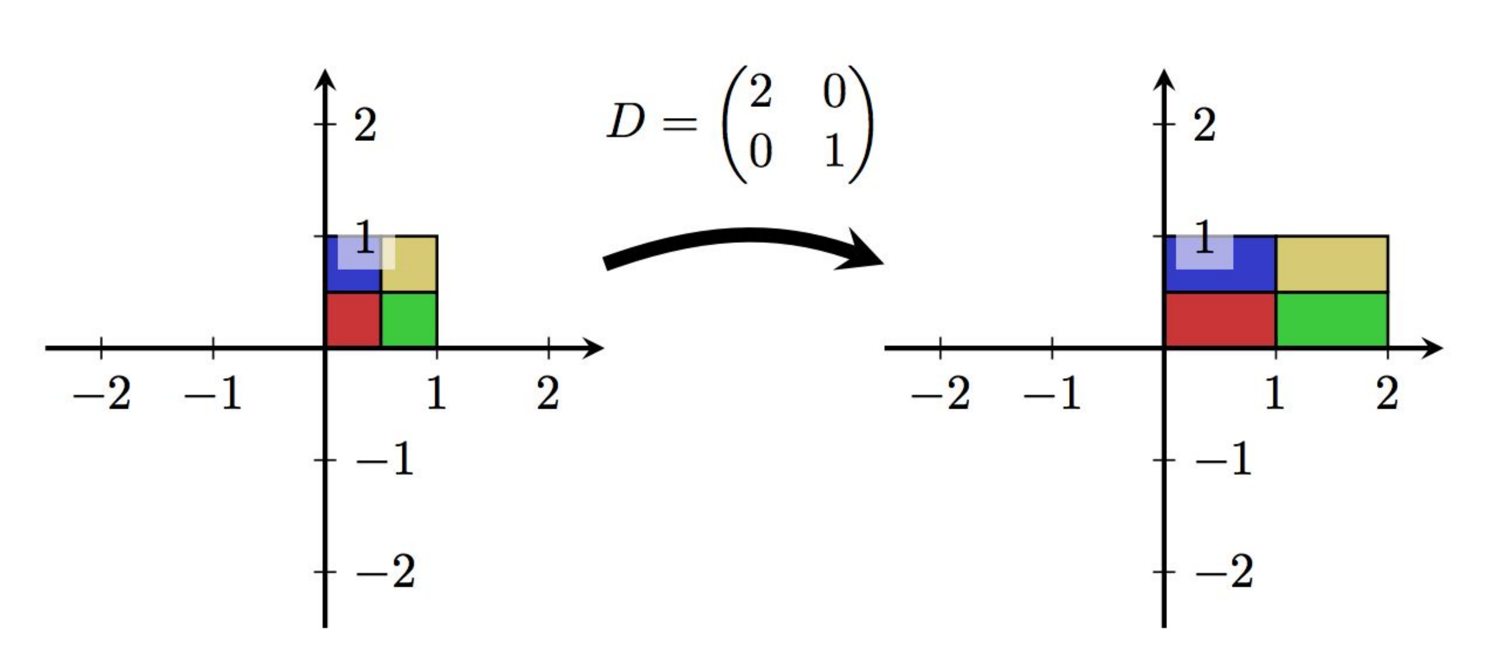
$$AA^T = A^T A = E$$

Типы матриц (преобразование пространства)



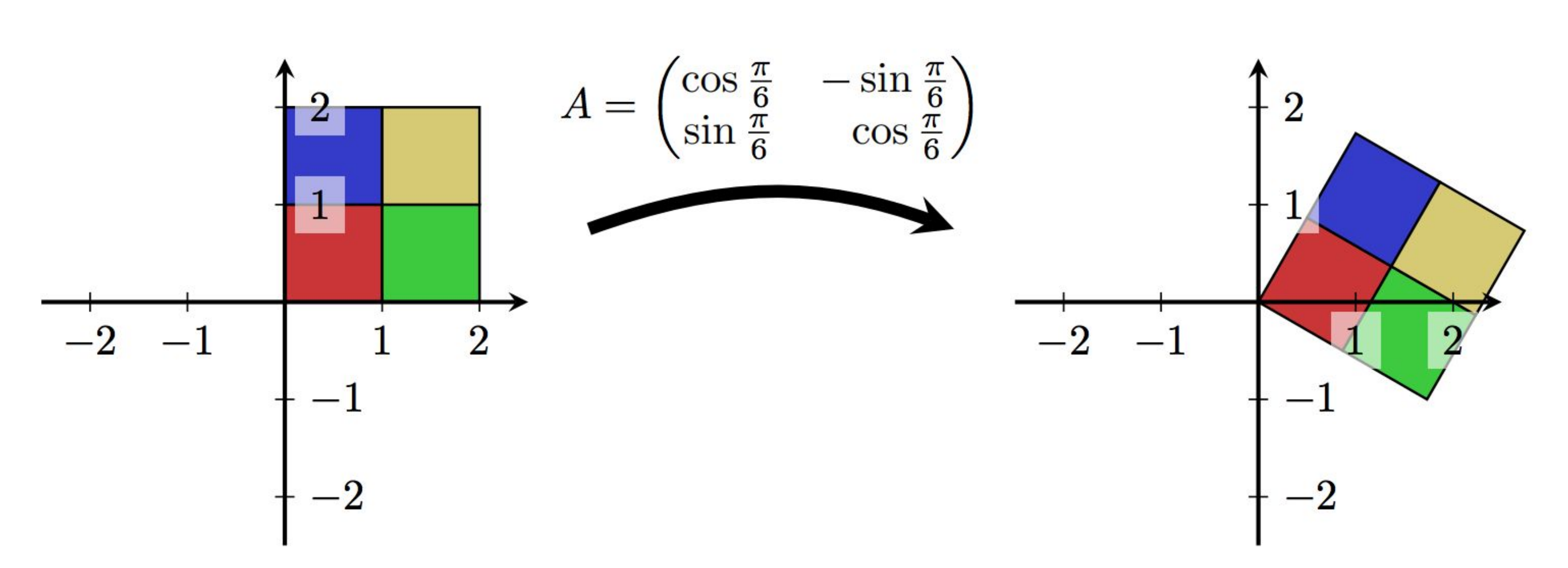
https://www.coursera.org/learn/mathematics-and-python

Типы матриц (преобразование пространства)



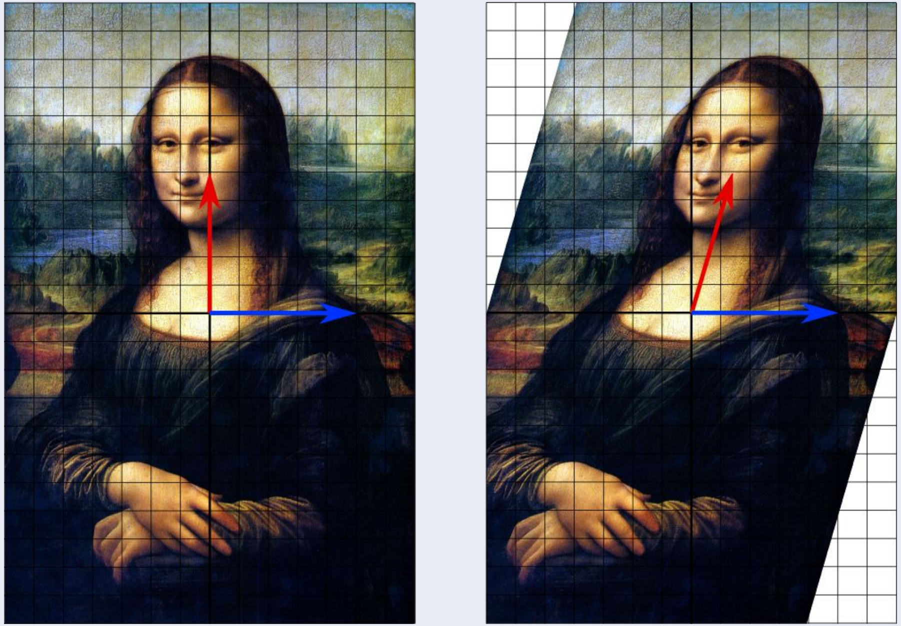
https://www.coursera.org/learn/mathematics-and-python

Типы матриц (преобразование пространства)



https://www.coursera.org/learn/mathematics-and-python

Собственные векторы и собственные значения



https://ru.wikipedia.org

Собственный вектор преобразования A

$$AX = \lambda X.$$

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

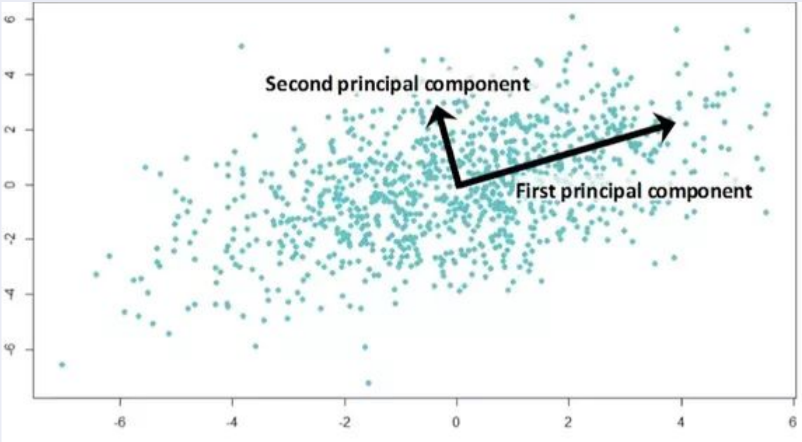
!

У матрицы **n x n** не более **n** собственных значений

НЕТОЛОГИЯ

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)

$$A=\begin{pmatrix}-1 & -6\\ 2 & 6\end{pmatrix}$$

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

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

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

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

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

$$x=-\frac{3}{2}y$$

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

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

http://mathprofi.ru/sobstvennye_znachenija_i_sobstvennye_vektory.html

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:

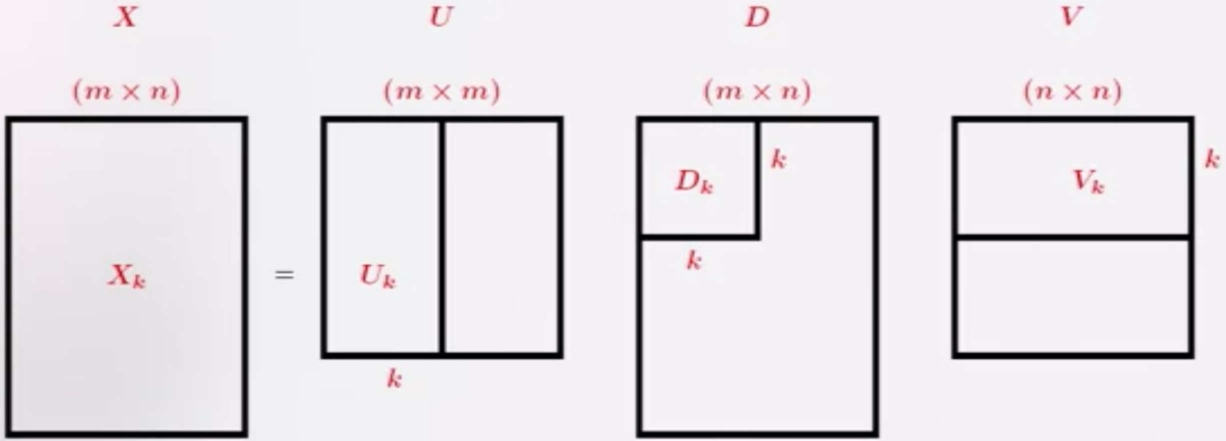
$$f(y) = y^T \cdot S^T \cdot D \cdot S \cdot y = (S \cdot y)^T \cdot D \cdot (S \cdot y) = z^T \cdot D \cdot z = \sum_{i=1}^{n} \lambda_i z_i^2,$$

https://www.coursera.org/learn/mathematics-and-python

Матричные разложения (сингулярное разложение)

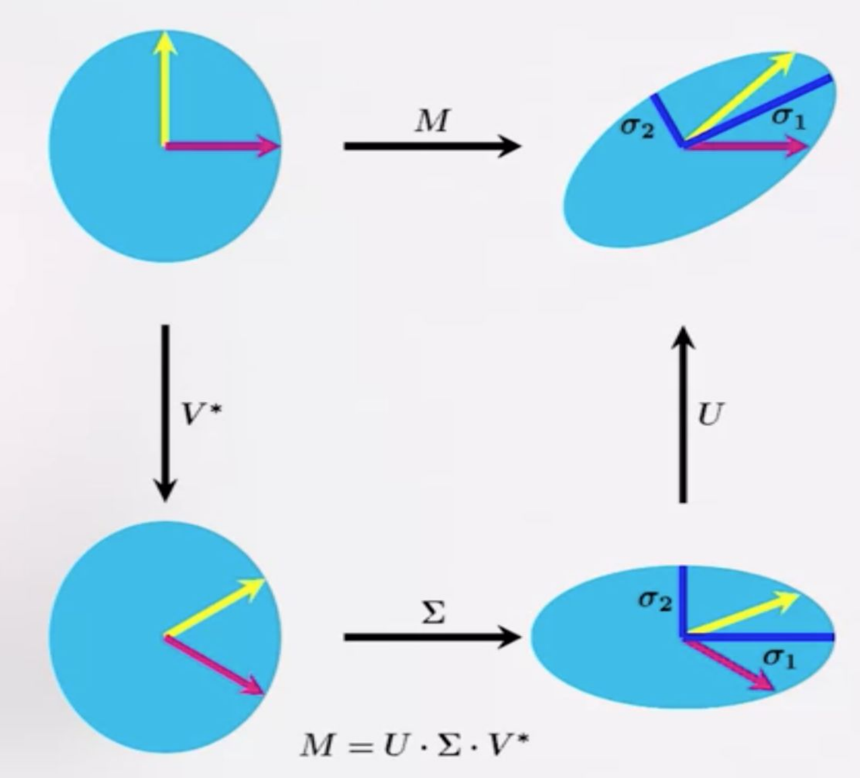
Но это была симметричная матрица, что в случае произвольной?

- $X = UDV$
- U, V — ортогональные, D — диагональная



https://www.coursera.org/learn/mathematics-and-python

Матричные разложения (сингулярное разложение)



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

НЕТОЛОГИЯ

https://www.coursera.org/learn/mathematics-and-python

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

! rank(X) <= min(n, m), если X - матрица m x 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

! rank(X) <= min(n, m), если X - матрица m x 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! rank(X) <= 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

- $X \approx X' = UV^T$
- $U - m \times k, V - n \times k$

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

https://www.coursera.org/learn/mathematics-and-python

Матричные разложения (приближение матрицей меньшего ранга)

Что значит приблизить

- $X \approx X' = UV^T$
- $U - m \times k, V - n \times k$

Просто найлучшее приближение по норме: $||X - UV^T|| \to min$

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

Итоговая задача выглядит так:

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

НЕТОЛОГИЯ

https://www.coursera.org/learn/mathematics-and-python

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

https://www.coursera.org/learn/mathematics-and-python

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?

https://www.coursera.org/learn/mathematics-and-python

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

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

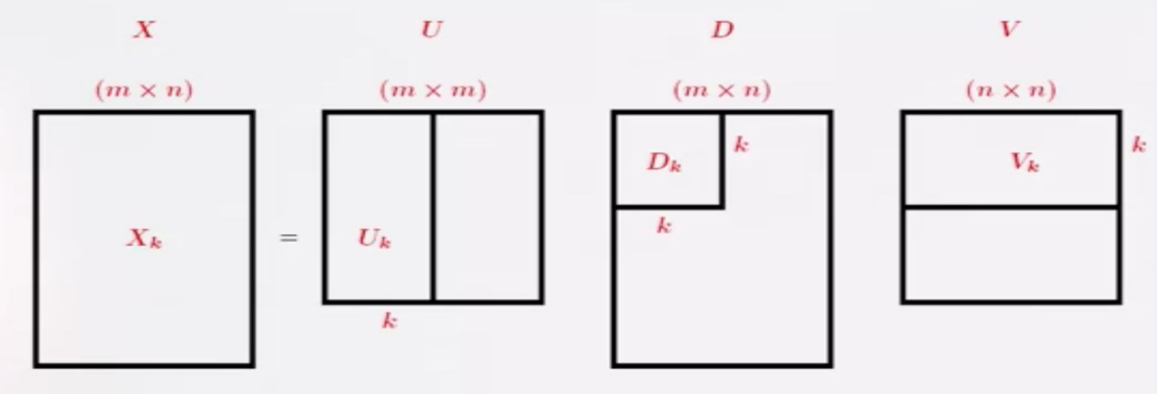
НЕТОЛОГИЯ

https://www.coursera.org/learn/mathematics-and-python

Матричные разложения (связь SVD и низкорангового приближения)

$$\hat{X} = \underset{\operatorname{rg} \hat{X} \leq k}{\operatorname{argmin}} \left\| X - \hat{X} \right\|$$

$$X = U \cdot D \cdot V^T$$



$$X_k = U_k D_k V_k^T$$

Усечённый SVD

https://www.coursera.org/learn/mathematics-and-python

Матричные разложения (связь SVD и низкорангового приближения)

Оказывается, Xk - наилучшее приближение матрицы X матрицей ранга <= k по норме Фробениуса!

$$X_k = U_k D_k V_k^T$$

$$\hat{X}_k = \underset{rg(\hat{X}) \leq k}{\operatorname{argmin}} ||X - \hat{X}||_F$$

$$||X - \hat{X}||_F = \sqrt{\sum_{i,j} (x_{ij} - \hat{x}_{ij})^2}$$

https://www.coursera.org/learn/mathematics-and-python

Матричные разложения (рекомендательные системы)

- Вариант 1 (не очень правильно, но просто): сделать SVD, матрицу $U_k D_k$ использовать как матрицу профилей пользователей, а матрицу V_k как матрицу профилей фильмов, произведение профилей – прогноз оценки фильма

- Вариант 2 (более правильно, но нужно более глубоко вникнуть в метод): Не будем никак использовать SVD, а просто подберем U и V, минимизируя функционал

НЕТОЛОГИЯ

https://www.coursera.org/learn/mathematics-and-python

Спасибо за внимание!

НЕТОЛОГИЯ