Матричные разложения и их применения в анализе данных

Руслан Хайдуров, Анастасия Иовлева

16 октября 2017

Содержание

1. Постановка задачи
2. Сингулярное разложение
3. Свойства сингулярного разложения
4. Вероятностный взгляд на SVD
5. SVD в рекомендательных системах
6. Разложение Холецкого

Постановка задачи

Пусть есть некая матрица $M \in \mathbb{R}^{m \times n}$.

$$M \approx U \times V^T$$

Постановка задачи

$$||\mathrm{A} - \mathrm{U}\mathrm{V}^{\mathrm{T}}|| \to \min_{\mathrm{U},\mathrm{V}}$$

Если $\mathrm{M} \in \mathbb{R}^{\mathrm{m}\times\mathrm{n}}$

$$||\mathrm{M}|| = \sum_{\mathrm{i}=1}^{\mathrm{m}} \sum_{\mathrm{j}=1}^{\mathrm{n}} \mathrm{m}_{\mathrm{ij}}^2$$

— норма Фробениуса.

Зачем нужны матричные разложения

- В машинном обучении данные представляются в матричном виде. (матрица объекты-признаки, картинка, матрица частот слов и т.д.)

	T1	T2	T3	T4	T5	T6	T7	T8
Doc1	2	0	4	3	0	1	0	2
Doc2	0	2	4	0	2	3	0	0
Doc3	4	0	1	3	0	1	0	1
Doc4	0	1	0	2	0	0	1	0
Doc5	0	0	2	0	0	4	0	0
Doc6	1	1	0	2	0	1	1	3
Doc7	2	1	3	4	0	2	0	2

Зачем нужны матричные разложения

Популярные направления использования

- Понижение размерности данных и визуализация
- Рекомендательные системы
- Поиск структур в данных

Сингулярное разложение матрицы (SVD)

Произвольная матрица $\mathrm{M} \in \mathbb{R}^{\mathrm{m}\times\mathrm{n}}$ может быть представлена в виде произведение трёх матриц

$$\mathrm{M} = \mathrm{U}\Sigma\mathrm{V}^{\mathrm{T}}$$

где матрицы $\mathrm{U}, \mathrm{V}^{\mathrm{T}}$ квадратные ортогональные, а $\Sigma \in \mathbb{R}^{\mathrm{m}\times\mathrm{n}}$ — диагональная с числами $\sigma_1 \geqslant \sigma_2 \geqslant \ldots \geqslant \sigma_{\min(\mathrm{m},\mathrm{n})} \geqslant 0$. Эти числа называются сингулярными числами матрицы.

Сингулярное разложение матрицы (SVD)

Усечённым сингулярным разложением называется представление произвольной матрицы M в виде

$$M = U\Sigma V^T$$

Где матрицы U и V^T — прямоугольные ортогональные, а Σ — квадратная с сингулярными значениями на диагонали.

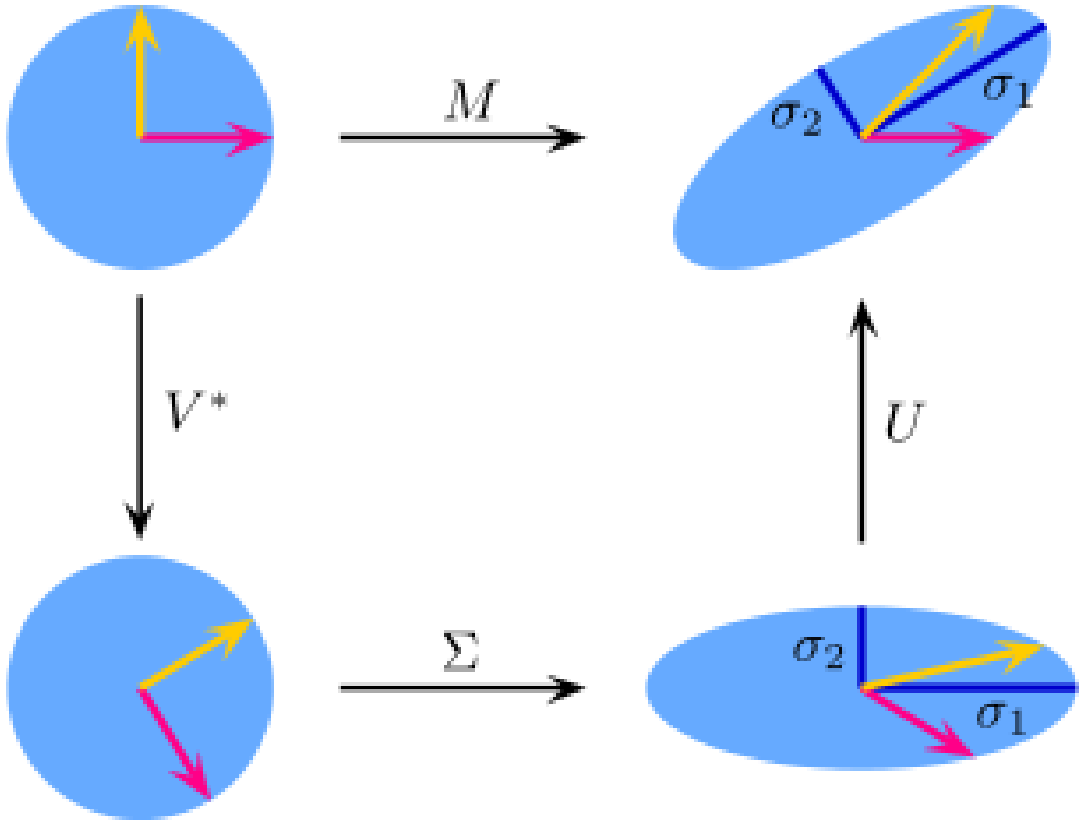
Построение разложения

- Ищутся левые и правые сингулярные векторы. Это собственныее векторы матриц MM^T и M^TM соответственно.
- В базис из сингулярных векторов можно перейти при помощи ортогональных преобразований U и V.
- В базисе из сингулярных векторов матрица имеет искомый вид.

Из построения видно, что сингулярное разложение сложно вычислять.

Геометрическая интерпретация

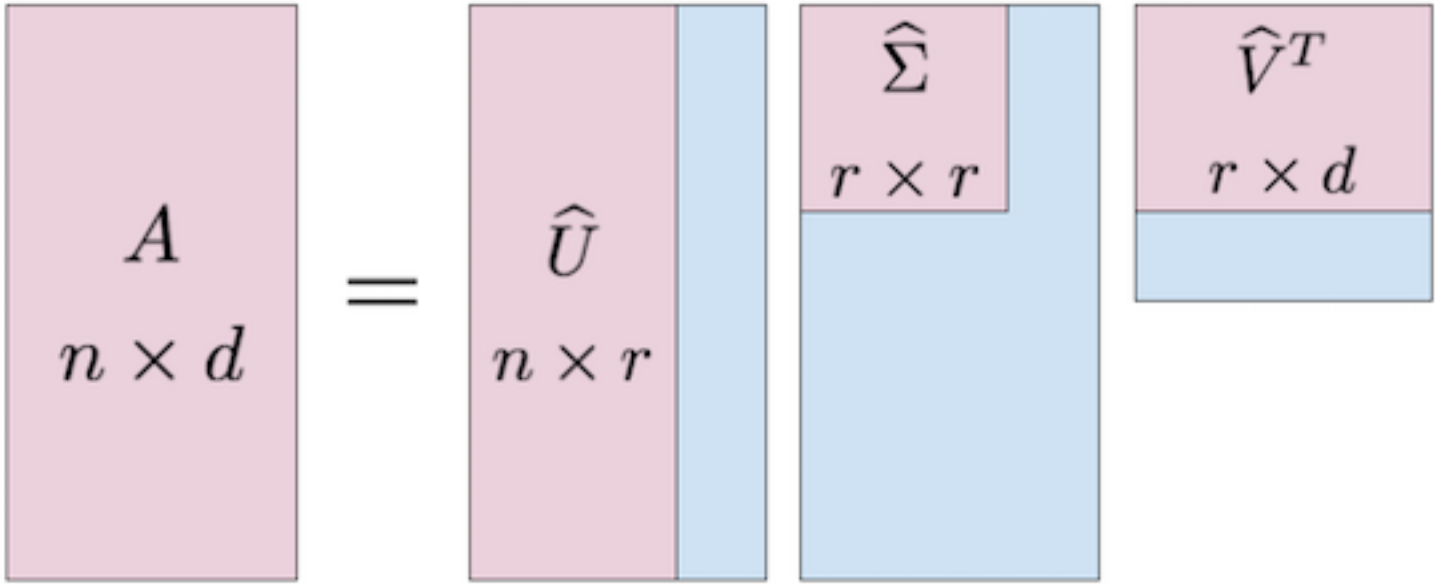
Выбрать такие базисы в $\mathbb{R}^m$ и в $\mathbb{R}^n$, что пребразование M будет просто растяжением вдоль осей.



$$M = U \cdot \Sigma \cdot V^*$$

Свойства сингулярного разложения

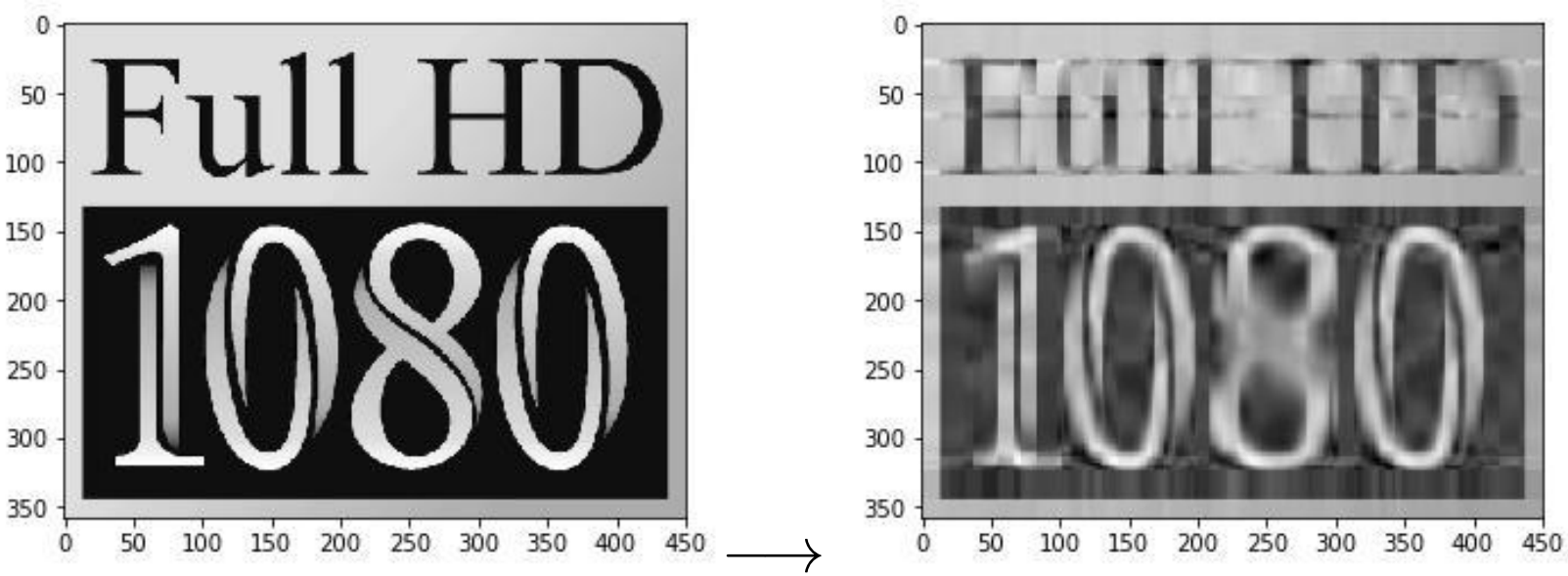
Даёт наилучшее низкоранговое приближение матрицы



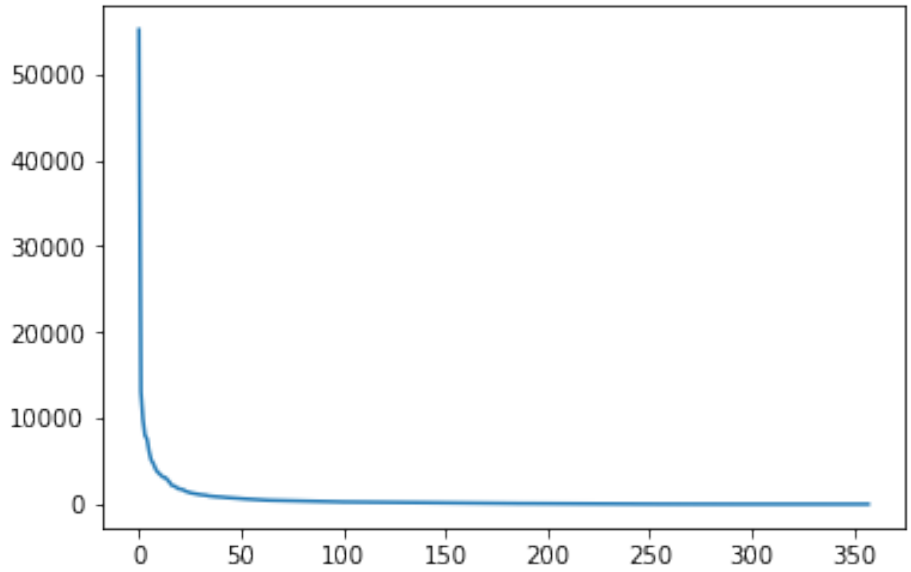
$$||\mathrm{A} - \hat{\mathrm{U}}\hat{\Sigma}\hat{\mathrm{V}}^{\mathrm{T}}|| \to \min$$

Свойства сингулярного разложения

Сингулярное разложение может быть использовано для сжатия данных.



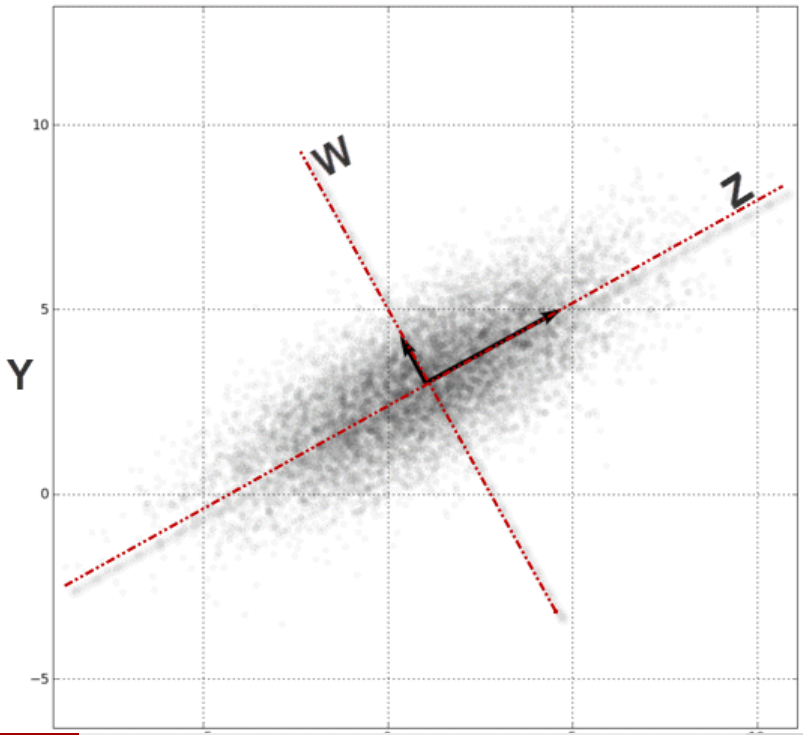
Свойства сингулярного разложения




```
U, S, VT = svd(img)
k = 10
Us = U[:, :k]
Vs = VT[:k, :]
Ss = np.diag(S)[:k, :k]
yup = (Us.dot(Ss)).dot(Vs)
plt.imshow(yup, cmap='gray')
```

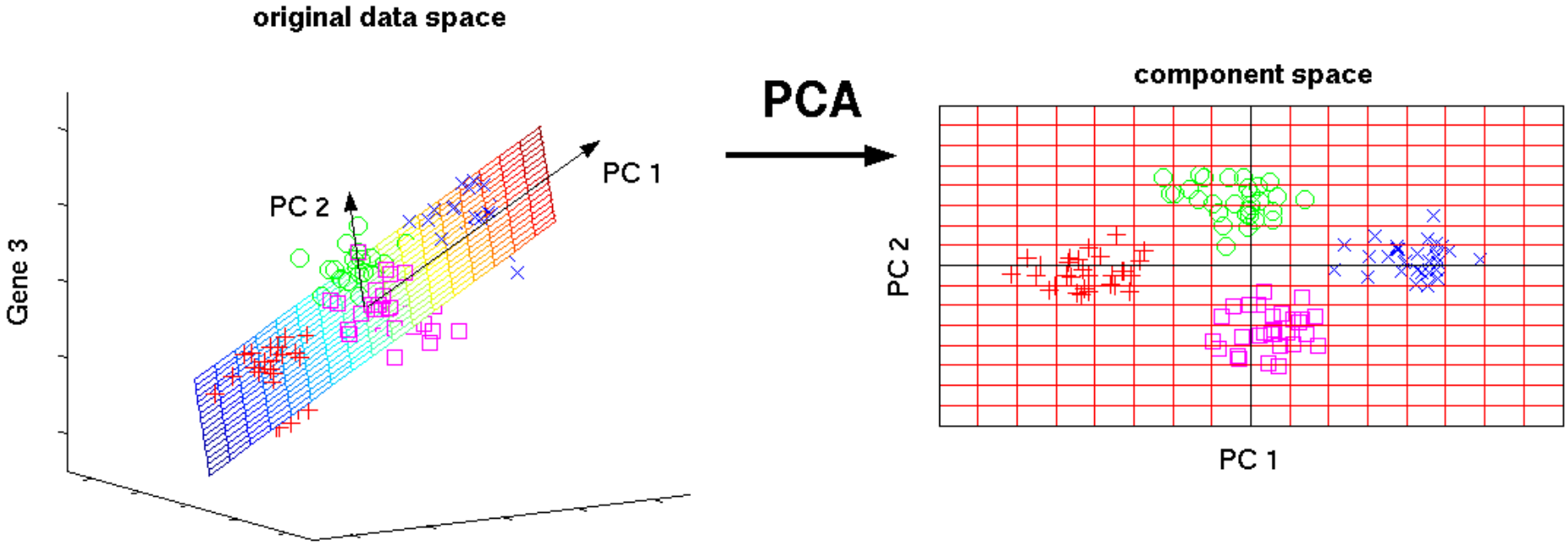
Вероятностный взгляд на SVD

- Будем рассматривать точки выборки как многомерные случайные величины.
- Наибольшее значение дисперсии проекции случайного вектора достигается при проекции на собственный вектор матрицы XX^T с наибольшим собственным значением.



Метод главных компонент (PCA)

- Найти сингулярное разложение матрицы.
- Выбрать направления сингулярных векторов, соответствующие наибольшим сингулярным числам (главные компоненты)
- Прибавить к векторам среднее значение по выборке и спроецировать на аффинное подпространство, натянутое на главные компоненты.



SVD в рекомендательных системах

Name	Jazz	Rock	Dubstep	Rap
Steve	1	3	2	1
Sam	4	2	3	4
Max	0	5	1	1
Helene	5	1	2	0
Michael	0	1	5	0

$$M \approx U_k \hat{V}_k^T, \text{ где } \hat{V}_k^T = \Sigma_k V_k^T.$$

Можно интерпретировать это разложение как две новые матрицы признаков — одна содержит характеристики жанра, другая — предпочтения слушателя.

Разложение Холецкого

Пусть есть матрица A, квадратная и положительно определённая. Тогда найдётся такое её разложение

$$A = LL^T$$

Что L — нижнетреугольная матрица со строго положительными элементами на диагонали.

Построение разложения

Итеративно для каждого элемента матрицы L.

$$L_{ii} = \sqrt{A_{ii} - \sum_{k=1}^{i-1} L_{ik}^2}, \tag{1}$$

$$L_{ij} = \frac{1}{L_{jj}} \left(A_{ij} - \sum_{k=1}^{j-1} L_{ik} L_{jk} \right). \tag{2}$$

Применение

- Для быстрого решения систем линейных уравнений и обращения матриц;
- Для генерации многомерных случайных величин.