KHAI TRIỂN SVD

Tran Van Long

AI Academy Vietnam

20/7/2024

Nội dung

1. Khai triển SVD
2. Phân tích ảnh bằng khai triển SVD
3. Thực hành và Bài tập

Phương pháp khai triển SVD [1]

Cho ma trận A đối xứng, vuông cấp n thì A chéo hóa được bởi ma trận trực giao V,

$$A = V\Sigma V^T,$$

với $VV^T = V^TV = I$, Σ là ma trận đường chéo.
Nếu ta đặt $V = [v_1, v_2, \ldots, v_n]$, $\Sigma = diag[\lambda_1, \lambda_2, \ldots, \lambda_n]$ thì

$$A = \lambda_1 v_1 v_1^T + \lambda_2 v_2 v_2^T + \cdots + \lambda_n v_n v_n^T.$$

Phương pháp khai triển SVD áp dụng cho ma trận A tổng quát cỡ $m \times n$ dưới dạng

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

với U, V là các ma trận có các cột trực giao và Σ là ma trận đường chéo.

[1] (Tài liệu 1)

Phương pháp khai triển SVD

Cho ma trận A cấp $m \times n$. Khi đó các ma trận AA^T và A^TA là các ma trận đối xứng, xác định không âm.
Với $x \in \mathrm{R}^m$ và $y \in \mathrm{R}^n$ ta có:

$$x^T(AA^T)x = (A^Tx)^T(A^Tx) = ||A^Tx||^2 \geq 0,$$
$$y^T(A^TA)y = (Ay)^T(Ay) = ||Ay||^2 \geq 0.$$

Phương pháp khai triển SVD

Tính chất

- Nếu v là véc-tơ riêng với $||v|| = 1$ của A^TA ứng với giá trị riêng λ thì $u = Av$ là véc-tơ riêng của AA^T với cùng giá trị riêng λ. Hơn nữa, $||Av|| = \sqrt{\lambda}$.
- Nếu u là véc-tơ riêng với $||u|| = 1$ của AA^T ứng với giá trị riêng λ thì $v = A^Tu$ là véc-tơ riêng của A^TA với cùng giá trị riêng λ. Hơn nữa, $||A^Tu|| = \sqrt{\lambda}$.

$$AA^Tu = AA^T(Av) = A(A^TA)v = A\lambda v = \lambda Av = \lambda u.$$

$$||Av||^2 = (Av)^T(Av) = v^TA^TAv = v^T\lambda v = \lambda \Rightarrow ||Av|| = \sqrt{\lambda}$$

$$A^TAv = A^TA(A^Tu) = A^T(AA^T)u = A^T\lambda u = \lambda A^Tu = \lambda v.$$

$$||A^Tu||^2 = (A^Tu)^T(A^Tu) = u^TAA^Tu = u^T\lambda u = \lambda \Rightarrow ||A^Tu|| = \sqrt{\lambda}$$

Cặp véc-tơ riêng của AA^T và A^TA

Tính chất

- Các ma trận AA^T và A^TA có cùng các giá trị riêng $\lambda_1, \ldots, \lambda_k$.
- Nếu $v_1, \ldots, v_k$ là hệ trực chuẩn các véc-tơ riêng của A^TA ứng với các giá trị riêng $\lambda_1, \ldots, \lambda_k$ thì $u_i = Av_i/\sqrt{\lambda_i}, i = 1, \ldots, k$ là hệ trực chuẩn các véc-tơ riêng của AA^T.
- Nếu $u_1, \ldots, u_k$ là hệ trực chuẩn các véc-tơ riêng của AA^T ứng với các giá trị riêng $\lambda_1, \ldots, \lambda_k$ thì $v_i = A^Tu_i/\sqrt{\lambda_i}, i = 1, \ldots, k$ là hệ trực chuẩn các véc-tơ riêng của A^TA.

$$
\begin{aligned}
u_i^T u_j &= \frac{1}{\sqrt{\lambda_i\lambda_j}}(Av_i)^T(Av_j) = \frac{1}{\sqrt{\lambda_i\lambda_j}} v_i^T A^T A v_j \\
&= \frac{1}{\sqrt{\lambda_i\lambda_j}} v_i^T \lambda_j v_j = \frac{1}{\sqrt{\lambda_i}} v_i^T v_j = 0.
\end{aligned}
$$

Khai triển SVD

SVD

Cho A là ma trận cỡ $m \times n (m > n)$. Khi đó có khai triển

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

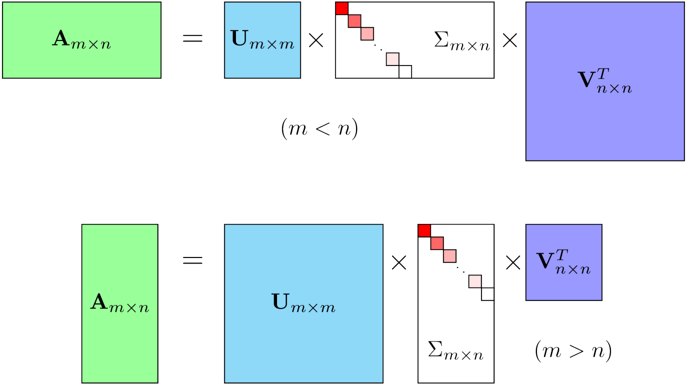
với ma trận U cấp $m \times n$ với các cột trực chuẩn, Σ ma trận $n \times n$ dạng đường chéo, V cấp $n \times n$ có các cột trực chuẩn.

Ma trận $A^T A$ là ma trận vuông cấp n, đối xứng, xác định không âm

$$A^T A = \sigma_1^2 v_1 v_1^T + \sigma_2^2 v_2 v_2^T + \cdots + \sigma_n^2 v_n v_n^T.$$

$$\begin{aligned} AV &= A[v_1, v_2, \ldots, v_n] = [\sigma_1 u_1, \sigma_2 u_2, \ldots, \sigma_n u_n] \\ &= [u_1, u_2, \ldots, u_n] \begin{pmatrix} \sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \sigma_n \end{pmatrix} = U\Sigma \Rightarrow A = U\Sigma V^T \end{aligned}$$

Phương pháp khai triển SVD



Khai triển SVD - dạng rút gọn

SVD rút gọn

Cho A là ma trận cỡ $m \times n$ với $r = rank(A)$. Khi đó có khai triển

$$A = U_r \Sigma_r V_r^T,$$

với ma trận U_r cấp $m \times r$ với các cột trực chuẩn, Σ_r ma trận $r \times r$ dạng đường chéo, V_r cấp $n \times r$ có các cột trực chuẩn.

$$A = \sigma_1 u_1 v_1^T + \sigma_2 u_2 v_2^T + \cdots + \sigma_r u_r v_r^T.$$

Với $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$.

Khai triển SVD và ma trận nghịch đảo suy rộng

SVD rút gọn

Cho A là ma trận cỡ $m \times n$ với $r = rank(A)$. Khi đó có khai triển

$$A = U_r \Sigma_r V_r^T,$$

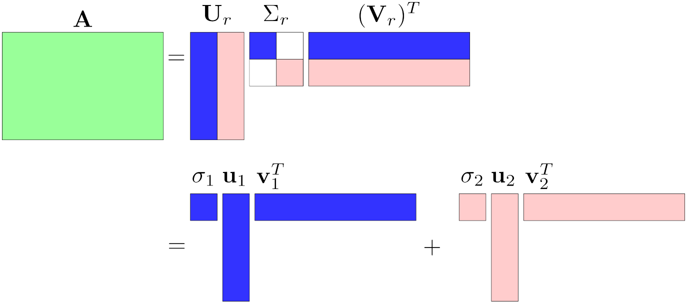
với ma trận U_r cấp $m \times r$ với các cột trực chuẩn, Σ_r ma trận $r \times r$ dạng đường chéo, V_r cấp $n \times r$ có các cột trực chuẩn.

$$A = \sigma_1 u_1 v_1^T + \sigma_2 u_2 v_2^T + \cdots + \sigma_r u_r v_r^T.$$

Nghịch đảo suy rộng của ma trận A là $A^\dagger = V_r D^{-1} U_r^T$

$$A^\dagger = 1/\sigma_1 v_1 u_1^T + 1/\sigma_2 v_2 u_2^T + \cdots + 1/\sigma_r v_r u_r^T.$$

Khai triển SVD



Khai triển SVD - dạng xấp xỉ

SVD xấp xỉ

$$A \approx A_k = U_k \Sigma_k V_k^T,$$

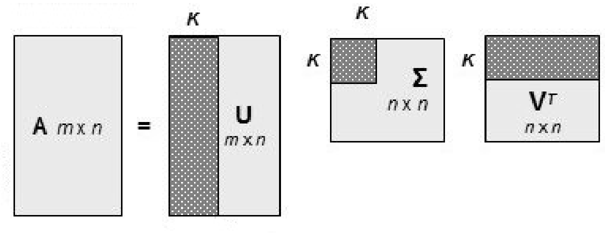
với ma trận U_k cấp $m \times k$ với các cột trực chuẩn, Σ_k ma trận $k \times k$ dạng đường chéo, V_k cấp $n \times k$ có các cột trực chuẩn.

$$A \approx A_k = \sigma_1 u_1 v_1^T + \sigma_2 u_2 v_2^T + \cdots + \sigma_k u_k v_k^T.$$

Sai số cho xấp xỉ là $||A - A_k||_F = \sqrt{\sigma_{k+1}^2 + \cdots + \sigma_r^2}$

Chọn k để phần giữ lại đủ lớn $\dfrac{||A_k||_F^2}{||A||_F^2} = \dfrac{\sum_{i=1}^{k} \sigma_i^2}{\sum_{i=1}^{r} \sigma_i^2}$

Khai triển SVD



Ví dụ: Khai triển SVD

Tìm khai triển SVD đối với ma trận $A = \begin{pmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{pmatrix}$

Tính ma trận $A^T A = \begin{bmatrix} 80 & 100 & 40 \\ 100 & 170 & 140 \\ 40 & 140 & 200 \end{bmatrix}$

$$V = [v_1, v_2, v_3] = \begin{bmatrix} 1/3 & -2/3 & 2/3 \\ 2/3 & -1/3 & -2/3 \\ 2/3 & 2/3 & 1/3 \end{bmatrix}$$

$\lambda_1 = 360, \lambda_2 = 90, \lambda_3 = 0$
$\sigma_1 = 6\sqrt{10}, \sigma_2 = 3\sqrt{10}, \sigma_3 = 0$

$$D = \begin{pmatrix} 6\sqrt{10} & 0 & 0 \\ 0 & 3\sqrt{10} & 0 \end{pmatrix}$$

$$u_1 = \frac{1}{\sigma_1} A v_1 = \frac{1}{6\sqrt{10}} \begin{pmatrix} 18 \\ 6 \end{pmatrix} = \begin{pmatrix} 3/\sqrt{10} \\ 1/\sqrt{10} \end{pmatrix},$$

$$u_2 = \frac{1}{\sigma_2} A v_2 = \frac{1}{3\sqrt{10}} \begin{pmatrix} 3 \\ -9 \end{pmatrix} = \begin{pmatrix} 1/\sqrt{10} \\ -3/\sqrt{10} \end{pmatrix}$$

$$U = [u_1, u_2] = \begin{pmatrix} 3/\sqrt{10} & 1/\sqrt{10} \\ 1/\sqrt{10} & -3/\sqrt{10} \end{pmatrix}$$

$$A = \begin{pmatrix} 3/\sqrt{10} & 1/\sqrt{10} \\ 1/\sqrt{10} & -3/\sqrt{10} \end{pmatrix} \begin{pmatrix} 6\sqrt{10} & 0 & 0 \\ 0 & 3\sqrt{10} & 0 \end{pmatrix} \begin{pmatrix} 1/3 & 2/3 & 2/3 \\ -2/3 & -1/3 & 2/3 \\ 2/3 & -2/3 & 1/3 \end{pmatrix}$$

$$= UDV^T$$

VINBIGDATA VINGROUP AI Academy Vietnam

Khai triển ma trận

Trong khai triển SVD ta phân tích ma trận A thành tích của ba ma trận U, D, V, dạng

$$S = UDV^T.$$

Bài toán xấp xỉ ma trận thành tích của hai ma trận có hạng bằng k, dạng

$$A \approx WH$$

Sử dụng khai triển SVD thì

$$W = U_k D_k; \ \ H = V_k^T$$

Bài toán phân tích ma trận

$$\min ||A - WH||^2.$$

Ví dụ: Khai triển SVD

Cho 6 văn bản, với 5 từ khóa. Ma trận A đếm số lần suất hiện các từ trong các văn bản.

$$\mathbf{A} = \left(\begin{array}{lccccc} & \text{lion} & \text{tiger} & \text{jaguar} & \text{porsche} & \text{ferrari} \\ \text{Document-1} & 2 & 2 & 2 & 0 & 0 \\ \text{Document-2} & 2 & 3 & 3 & 0 & 0 \\ \text{Document-3} & 1 & 1 & 1 & 0 & 0 \\ \text{Document-4} & 2 & 2 & 3 & 1 & 1 \\ \text{Document-5} & 0 & 0 & 1 & 1 & 1 \\ \text{Document-6} & 0 & 0 & 2 & 1 & 2 \end{array}\right)$$

Ví dụ: Khai triển SVD

$$A = \begin{bmatrix} 2 & 2 & 2 & 0 & 0 \\ 2 & 3 & 3 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 2 & 2 & 3 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 2 & 1 & 2 \end{bmatrix}, \quad A^T A = \begin{bmatrix} 13.0 & 15.0 & 17.0 & 2.0 & 2.0 \\ 15.0 & 18 & 20 & 2 & 2 \\ 17.0 & 20 & 28 & 6 & 8 \\ 2.0 & 2 & 6 & 3 & 4 \\ 2.0 & 2 & 8 & 4 & 6 \end{bmatrix}$$

Ví dụ: Khai triển SVD

Các giá trị riêng của $A^T A$ là
$\sigma_1^2 = 57.6; \sigma_2^2 = 9.8; \sigma_3^2 = 0.426; \sigma_4^2 = 0.175; \sigma_5^2 = 0$

$$U = \begin{bmatrix} 0.4408 & -0.2729 & 0.2983 & -0.4621 & 0.5161 \\ 0.6015 & -0.3121 & -0.6541 & 0.33598 & 0 \\ 0.2204 & -0.1364 & 0.1492 & -0.2311 & -0.8463 \\ 0.5679 & 0.1771 & 0.5157 & 0.1518 & -0.0930 \\ 0.1270 & 0.4500 & 0.2173 & 0.6139 & 0.09304 \\ 0.2378 & 0.7586 & -0.38429 & -0.4692 & 0.0000 \end{bmatrix}$$

Khai triển SVD

$$\Sigma = \begin{bmatrix} 7.5894 & 0 & 0 & 0 & 0 \\ 0 & 3.130 & 0 & 0 & 0 \\ 0 & 0 & 0.6525 & 0 & 0 \\ 0 & 0 & 0 & 0.41802 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

$$V = \begin{bmatrix} 0.4533 & -0.3042 & 0.7189 & -0.43017 & 0 \\ 0.5326 & -0.40394 & -0.2835 & 0.3733 & 0.5773 \\ 0.6869 & 0.2810 & -0.3380 & -0.0397 & -0.5774 \\ 0.1229 & 0.442 & 0.5343 & 0.7098 & 0 \\ 0.1542 & 0.685 & -0.05451 & -0.4131 & 0.5773 \end{bmatrix}$$

Ví dụ

$$A_2 = \begin{bmatrix} 1.7769 & 2.1273 & 2.0581 & 0.0330 & -0.0691 \\ 2.3672 & 2.8265 & 2.8612 & 0.1284 & 0.0347 \\ 0.8885 & 1.0637 & 1.0290 & 0.0165 & -0.0346 \\ 1.7854 & 2.0717 & 3.1163 & 0.7751 & 1.0446 \\ 0.008 & -0.0556 & 1.0581 & 0.7421 & 1.113 \\ 0.095 & 0.0021 & 1.9074 & 1.273 & 1.9052 \end{bmatrix}$$

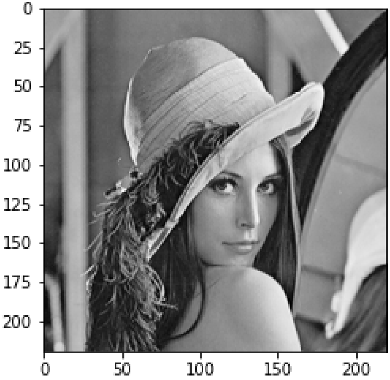
Ví dụ

$$AV_2^T = U_2 S_2 = \begin{bmatrix} 3.345874514 & -0.8543387501 \\ 4.565420247 & -0.9772577443 \\ 1.672937257 & -0.4271693747 \\ 4.309910415 & 0.5544380636 \\ 0.9640358974 & 1.408776813 \\ 1.805170452 & 2.374867311 \end{bmatrix}$$

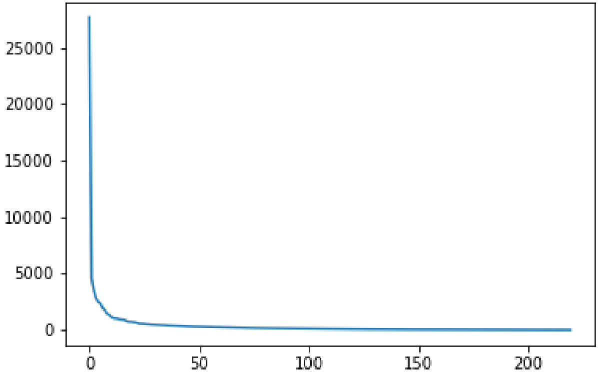
Phân tích ảnh bằng khai triển SVD [2]

[2] (Tài liệu 1)

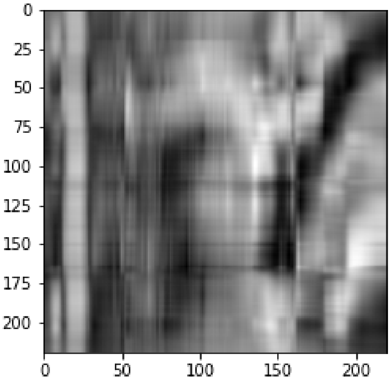
Phân tích ảnh bằng khai triển SVD



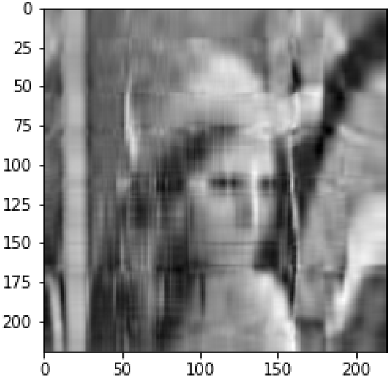
Phân tích ảnh bằng khai triển SVD



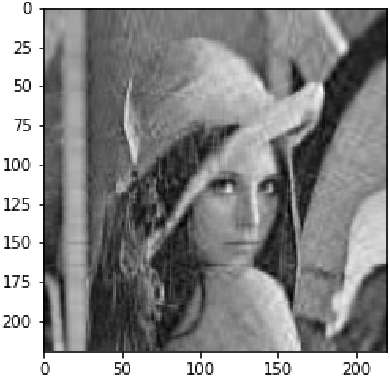
Xấp xỉ $A \approx A_k = U_k D_k V_k^T, k = 5$



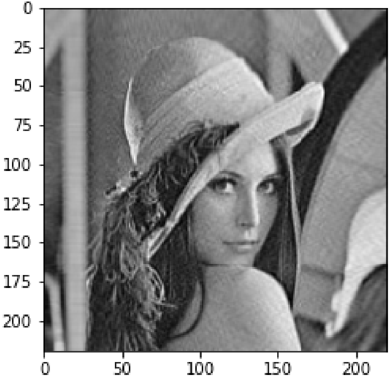
Xấp xỉ $A \approx A_k = U_k D_k V_k^T, k = 10$



Xấp xỉ $A \approx A_k = U_k D_k V_k^T, k = 25$



Xấp xỉ $A \approx A_k = U_k D_k V_k^T, k = 50$



Thực hành 1

Cho ma trận

$$A = \begin{pmatrix} 2 & 1 & 3 \\ 1 & 2 & 0 \end{pmatrix}$$

a) Thực hiện phép phân tích SVD.

b) Tính ma trận nghịch đảo suy rộng của A.

Thực hành 2

Sử dụng thư viện *numpy.linalg.svd* tìm khai triển SVD cho ma trận

$$A = \begin{pmatrix} 1 & 0 & 0 & 0 & 2 \\ 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \end{pmatrix}$$

Bài tập 1

Giả sử ma trận A cấp $m \times n$ có khai triển SVD $A = UDV^T$. Tìm nghiệm của bài toán hiệu chỉnh

$$\min\{||Ax - b||^2 + \lambda||x||^2\}$$

Bài tập 2

Sử dụng phương pháp khai triển SVD để giảm biểu diễn dữ liệu iris với $k = 2$.

Bài tập 3

Thực hành phân tích dữ liệu MNIST
Ma trận X cỡ 70000×784. Tách ma trận A cỡ 5000×784.
Thực hiện khai triển SVD cho ma trận A.

Tài liệu tham khảo

1. Charu C. Aggarwal; Linear Algebra and Optimization for Machine Learning, Springer, 2020.
2. David C. Lay, Steven R. Lay, Judi J. McDonald; Linear Algebra and Its Applications, Fifth edition, Pearson, 2016
3. Gilbert Strang; Linear Algebra and Learning from Data, Wellesley- Cambridge Press, 2019.
4. Stephen Boyd, Lieven Vandenberghe; Introduction to Applied Linear Algebra: Vectors, Matrices, and Least Squares, Cambridge University Press, 2018.
5. Tom Lyche; Numerical Linear Algebra and Matrix Factorizations, Springer, 2020.
6. Hyun-Seok Son; Linear Algebra Coding with Python, 2020.