SVD Draft 2

Sunday, August 1, 2021 2:06 AM

Singular Value Decomposition

Bài toán 1 (Linear Regression) Cho các điểm data $x_1, x_2, \dots, x_n \in \mathbb{R}^d$ và vector nhãn $y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \in \mathbb{R}^n$

Tìm vector $a = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_d \end{pmatrix} \in \mathbb{R}^d$ sao cho:

$$Err(a) = \sum_{i=1}^{n} (a^T x_i - y_i)^2 \text{ đạt GTNN.}$$

Homework 1 Hãy so sánh bài toán này với bài toán Linear Regression mà bạn đã học (nếu có) và chứng minh 2 bài toán tương đương.

Thử giải bài toán 1

Ta có: $X = \begin{bmatrix} - x_1^T - \\ - x_2^T - \\ \vdots \\ - x_n^T - \end{bmatrix} \in \mathbb{R}^{n\times d}$ (data matrix)

$$\Rightarrow Xa - y = \begin{pmatrix} a^T x_1 - y_1 \\ \vdots \\ a^T x_n - y_n \end{pmatrix} \in \mathbb{R}^n \Rightarrow \underbrace{\sum_{i=1}^{n} (a^T x_i - y_i)^2}_{= Err(a)} = \|Xa - y\|_2^2$$

$\Rightarrow$ Phát biểu lại bài toán: Tìm $a \in \mathbb{R}^d$: $\|Xa-y\|_2^2$ đạt GTNN

$$\|Xa-y\|_2^2 = (Xa-y)^T(Xa-y)$$

$$= \underbrace{a^TX^TXa - y^TXa - a^TX^Ty}_{\text{hàm theo } a} + \underbrace{y^Ty}_{\text{hằng số}}$$

Giả sử $a^* = \arg\min_a \|Xa-y\|_2^2$

$$a^* = \arg\min_a \left(a^TX^TXa - 2a^TX^Ty\right)$$

$$\Rightarrow \nabla_a Err(a^*) = 0 \Leftrightarrow 2X^TXa - 2X^Ty = 0$$

$$\Leftrightarrow \underbrace{X^TX}_{\in \mathbb{R}^{d\times d}}a = X^Ty \overset{?}{\Leftrightarrow} a = (X^TX)^{-1}X^Ty \; ?$$

$\rightarrow$ Nếu X^TX không khả nghịch? $\rightarrow$ Thử cách khác?

Bài toán 2 (Low-rank approximation) Cho $A \in \mathbb{R}^{m\times n}$ (VD: ảnh)

Lưu trữ A: mn bytes (giả sử 1 số chiếm 1 bytes) $\rightarrow$ Nén A?

Cách nén Tìm A_k: $\left.\begin{cases} \text{rank}(A_k) \le k & (1) \\ \|A - A_k\|_2 \text{ đạt GTNN với đk (1)} \end{cases}\right\} \Rightarrow$ Bài toán 2

$(1) \Rightarrow \exists U \in \mathbb{R}^{m\times k}, \exists W \in \mathbb{R}^{n\times k}$ … $A_k = UW$ … Lưu trữ A_k = Lưu trữ U, W

$\{ \|A - A_k\|_2$ đạt GTNN với đk (1) $\}$

(1) $\Rightarrow \exists U \in \mathbb{R}^{m\times k}, \exists W \in \mathbb{R}^{n \times k}$ sao cho $A_k = UW$. Lưu trữ A_k = Lưu trữ U, W $= k(m+n)$ bytes

Compression rate: $\frac{k(m+n)}{mn} = \frac{k}{m} + \frac{k}{n}$

Bài toán 3 Cho $A \in \mathbb{R}^{m\times n}$. $S = \{x \in \mathbb{R}^n \mid \|x\|_2 \leq 1\}$.

Mô tả $A_S = \{Ax \mid x \in S\}$.

Giải bài toán 3 $\{Ax \mid x \in S\} \subset \mathbb{R}^m$ là 1 hình ellipsoid

- Ellipsoid r chiều: $\left\{ \begin{pmatrix} x_1 \\ \vdots \\ x_r \end{pmatrix} \middle| \sum_{i=1}^{r} \frac{x_i^2}{\sigma_i^2} \leq 1 \right\} \xrightarrow{\text{Phép biến hình}}$ Ellipsoid

Ví dụ: $A = \begin{bmatrix} 4 & 0 \\ 3 & -5 \end{bmatrix}$ $\quad S = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \middle| x^2 + y^2 \leq 1 \right\}$ $\quad A_S = \left\{ \begin{pmatrix} 4x \\ 3x - 5y \end{pmatrix} \middle| x^2 + y^2 \leq 1 \right\}$

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} 4x \\ 3x - 5y \end{pmatrix} \Leftrightarrow \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \frac{1}{4}x' \\ \frac{3}{20}x' - \frac{1}{5}y' \end{pmatrix}$$

$$A_S = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \middle| \left(\frac{1}{4}x\right)^2 + \left(\frac{3}{20}x - \frac{1}{5}y\right)^2 \leq 1 \right\}$$

Ký hiệu:

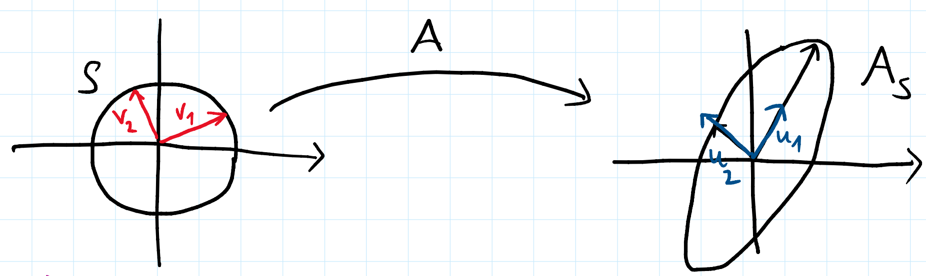
Trục chính của $A_S = \sigma_1 u_1$

$\sigma_1 = \|\text{Trục chính}\|_2$

$\|u_1\| = 1$

Trục phụ của $A_S = \sigma_2 u_2$

$\sigma_2 = \|\text{Trục phụ}\|_2$, $\|u_2\| = 1$



Tìm $u_1, u_2, \sigma_1, \sigma_2$ và v_1, v_2 sao cho $Av_i = \sigma_i u_i$

Nhận xét: $\begin{cases} u_1 \perp u_2 \Rightarrow \sigma_2 = \max\{\|u\|_2 \mid u \in A_S, u \perp u_1\} & (1) \\ \sigma_1 = \max\{\|u\|_2 \mid u \in A_S\} = \max\{\|Av\|_2 \mid v \in S\} & (2) \end{cases}$

(2) $\Rightarrow \sigma_1^2 = \max(\text{Eigenvalues}(A^TA))$, v_1 = Eigenvector của A^TA ứng với σ_1^2

$\sigma_2^2 = 2^{nd}$ eigenvalue, v_2 = Eigenvector của A^TA ứng với σ_2^2

$$\Rightarrow A^TAv_i = \sigma_i^2 v_i \Rightarrow \sigma_i^2 v_i = A^T \sigma_i u_i \Rightarrow A^T u_i = \sigma_i v_i \Rightarrow u_i^T A = \sigma_i v_i^T$$

$$\Rightarrow \begin{bmatrix} u_1^T \\ u_2^T \end{bmatrix} A = \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{bmatrix} \begin{bmatrix} v_1^T \\ v_2^T \end{bmatrix} \Rightarrow A = \underbrace{\begin{bmatrix} u_1 & u_2 \end{bmatrix}}_{U} \underbrace{\begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{bmatrix}}_{\Sigma} \underbrace{\begin{bmatrix} v_1^T \\ v_2^T \end{bmatrix}}_{V^T} = U\Sigma V^T$$

Hơn nữa: $v_1 A^T A v_2 = \sigma_2^2 v_1^T v_2 = \sigma_1^2 v_1^T v_2$

$\Rightarrow v_1^T v_2 = 0 \Rightarrow v_1 \perp v_2$ | $\Rightarrow U^T U = I = V^T V$

Tổng quát $\forall A \in \mathbb{R}^{m\times n}$: $\begin{cases} \text{Trục chính } i^{th} \text{ của } A_s: \ \sigma_i u_i, \quad \|u_i\| = 1, \quad \sigma_i > 0 \\ A v_i = \sigma_i u_i \end{cases}$

$$\|Av_i\|_2^2 = \max\{\|u\|_2^2 \mid u \perp u_1, \dots, u_{i-1}\} \Rightarrow \begin{cases} \sigma_i^2 = i^{th} \text{ Eigenvalue của } A^TA \\ V_i = i^{th} \text{ Eigenvector} \end{cases} \quad (*)$$

$$\Rightarrow v_i A^T A v_j = \sigma_i^2 v_i^T v_j = \sigma_j^2 v_i^T v_j \Rightarrow \begin{cases} v_i \perp v_j \text{ nếu } \sigma_i \neq \sigma_j \\ (+\text{ Nhiều chứng minh}) \ v_i \perp v_j \text{ nếu } \sigma_i = \sigma_j \end{cases}$$

$A v_i = \sigma_i u_i \ \forall i$ (1) $A^T u_i = \sigma_i v_i$ (2)

Nếu rank $A = r \Rightarrow$ Có r eigenvalues > 0 của A^TA. (Nhiều chứng minh)

$\Rightarrow$ Đặt $U = [u_1 \dots u_r] \in \mathbb{R}^{m\times r}$, $V = [v_1, \dots, v_r] \in \mathbb{R}^{n\times r}$, $\Sigma = \begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r \end{pmatrix} \in \mathbb{R}^{r\times r}$

$$\Rightarrow U^T U = V^T V = I_r. \quad (*) \Rightarrow \begin{cases} A^T A = V_n \begin{bmatrix} \sigma_1^2 & & & & & \\ & \ddots & & & & \\ & & \sigma_r^2 & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{bmatrix} V_n^T & \lambda_i = \sigma_i^2, \ i \leq r; \quad \lambda_{r+1} = \dots = \lambda_n = 0 \\ V_n = [v_1 \dots v_r, \underbrace{v_{r+1}, \dots, v_n}] \end{cases}$$

$\hookrightarrow$ Thêm vector vào thành Cơ sở trực chuẩn của $\mathbb{R}^n$

$$\Rightarrow A^T A = \sum_{i=1}^n \lambda_i v_i v_i^T = \sum_{i=1}^r \sigma_i^2 v_i v_i^T$$

$$\Rightarrow A^T A = \sum_{i=1}^r \sigma_i A^T u_i v_i^T = \sum_{i=1}^r A^T A v_i v_i^T \Rightarrow A^T A E = 0$$

với $E = I_n - \sum_{i=1}^r v_i v_i^T$

$\Rightarrow AE = 0$

$\Rightarrow A = A \sum_{i=1}^r v_i v_i^T$

[Bổ đề: $\forall E$: $A^TAE = 0 \Leftrightarrow AE = 0$
CM: ($\Leftarrow$): Hiển nhiên ($\Rightarrow$): $(EA)^T AE = 0$ ✓]

$\Rightarrow A = \sum_{i=1}^r \sigma_i u_i v_i^T$

$$\Rightarrow A = [u_1 \dots u_r] \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r \end{bmatrix} \begin{bmatrix} v_1^T \\ \vdots \\ v_r^T \end{bmatrix} = U\Sigma V^T$$

Định nghĩa $A = U\Sigma V^T$ là singular value decomposition của A.

Singular values $= \{\sigma_1, \dots, \sigma_r\}$

Định lý 1

- $U^TU = I_r = V^TV$. $r = \text{rank}(A)$
- $\sigma_1^2 \geq \sigma_2^2 \geq \dots \geq \sigma_r^2$ là toàn bộ Eigenvalues $\neq 0$ của A^TA
- $v_1 \quad v_2 \quad \dots \quad v_r$ = Eigenvectors
- $u_1 \quad u_2 \quad \dots \quad u_r$ = Eigenvectors (Same) của AA^T

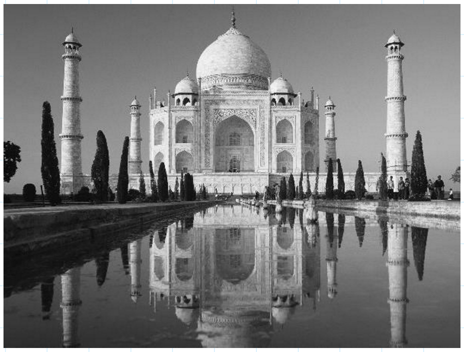
Chứng minh trên

Định lý 2 (Low-rank approximation) $A \in \mathbb{R}^{m\times n}$, $\text{rank}(A) = r$, $k \leq r$

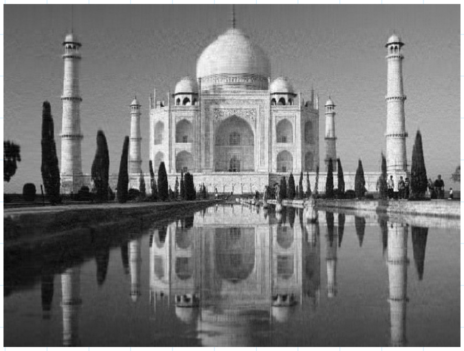
$$\Rightarrow A_k = \text{argmin}_B \left\{ \|A - B\|_F^2 \;\middle|\; \text{rank}(B) \leq k \right\} \text{ với:}$$

$$A_k = [u_1 \dots u_k] \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_k \end{bmatrix} \begin{bmatrix} v_1^T \\ \vdots \\ v_k^T \end{bmatrix} = U_k \Sigma_k V_k^T$$

Trả lời bài toán 2: $A_k = U_k \Sigma_k V_k^T$ là xấp xỉ tối ưu



Ảnh gốc: 450 x 600



SVD của ảnh gốc với rank 80:
RMSE = 8.2, Compression rate = 0.3

Homework 2 Chứng minh rằng mọi vector a thoả $\|Xa - y\|_2^2$ đạt GTNN trong Bài toán 1 đều có dạng $a = VU^Ty + c$ với $c \in \mathbb{R}^d$ bất kỳ thỏa mãn $V^Tc = 0$

(Trong đó $X = U\Sigma V^T$)