SVD computations

Tuesday, May 30, 2023 11:32 AM

$A \in \mathbb{R}^{m \times n}$ $\quad A = [a_1 \; a_2 \; \dots \; a_n]$

Matrix factorizations: 1) $QR = A$; $Q = [q_1 \; q_2 \; \dots \; q_n]$

$Q \in \mathbb{R}^{m \times n} \quad R \in \mathbb{R}^{n \times n}$

(Gram-Schmidt)

$$Q^T Q = \begin{bmatrix} q_1^T \\ q_2^T \\ \vdots \\ q_n^T \end{bmatrix} [q_1 \; q_2 \; \dots \; q_n] = I_n \in \mathbb{R}^{n \times n}$$

FLOPs: $O(mn^2)$; $O(m^3)$ $\underset{m=n}{}$

2) $LU = A \qquad L \in \mathbb{R}^{m \times n} \quad U \in \mathbb{R}^{n \times n}$ (Gaussian elimination)

L lower triangular, U upper triangular

FLOPs: $O(m^3/3)$

3) $AX = X\Lambda$ eigenproblem $\quad A \in \mathbb{R}^{m \times m}$

if X spans the entire space; $C(X) = \mathbb{C}^m$; X is not singular

$\Rightarrow \quad A = X \Lambda X^{-1}, \quad X^{-1} \Lambda X = A$ (eigendecomposition)

if A is normal ($A^T A = A A^T$; A symmetric, $A = A^T$; A anti-symmetric, $A = -A^T$)

$\Rightarrow \quad A = Q \Lambda Q^T \quad Q$ orthogonal eigenvector matrix

4) S.V.D. $\quad A = U \Sigma V^T \qquad A \in \mathbb{R}^{m \times n}$

$U \in \mathbb{R}^{m \times m} \quad \Sigma \in \mathbb{R}_+^{m \times n} \quad V \in \mathbb{R}^{n \times n}$

$U^T U = I_m \quad \Sigma = \mathrm{diag}(\sigma_1, \dots, \sigma_r, 0, \dots, 0) \quad V^T V = I_n$

$$\Sigma = \begin{bmatrix} \sigma_1 & & & & & \\ & \sigma_2 & & & & \\ & & \ddots & & & \\ & & & \sigma_r & & \\ & & & & 0 & \\ & & & & & \ddots \; 0 \end{bmatrix} \updownarrow m, \quad \leftarrow n \rightarrow$$

$\sigma_1, \dots, \sigma_r$ singular values

$\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > 0$

$$A = U \Sigma V^T = [u_1 \; u_2 \; \dots \; u_m] \begin{bmatrix} \sigma_1 & & & & \\ & \sigma_2 & & & \\ & & \sigma_r & & \\ & & & 0 & \\ & & & & \ddots \; 0 \end{bmatrix} \begin{bmatrix} v_1^T \\ v_2^T \\ \vdots \\ v_n^T \end{bmatrix}$$

$$A = [u_1 \; u_2 \; \dots \; u_n] \begin{bmatrix} \sigma_1 v_1^T \\ \sigma_2 v_2^T \\ \vdots \\ \sigma_r v_r^T \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \sum_{i=1}^{r} \sigma_i u_i v_i^T$$

$$u_i v_i^T = \begin{bmatrix} \; \end{bmatrix} \begin{bmatrix} \; \end{bmatrix} = \begin{bmatrix} \; \end{bmatrix}$$ rank-1 update

$(m \times 1) \cdot (1 \times n) \qquad (m \times n)$

scalar

$(m\times 1)\cdot(1\times n)$ $\qquad\qquad (m\times n)$

scal

$$= [v_{1i}u_i \quad v_{2i}u_i \quad \dots \quad v_{ni}u_i]$$

$$A = \sum_{i=1}^{r} \sigma_i u_i v_i^T \qquad \sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > 0$$

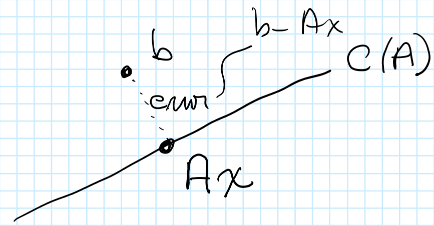
often

$$\sigma_1 >> \sigma_2 >> \sigma_3 \dots$$

such that $A \cong \sum_{i=1}^{k} \sigma_i u_i v_i^T$ (rank-k approximation).

Solution to common lin. alg. problems:

1) $b \in \mathbb{R}^m$ find approx in $C(A)$ by lin. comb. $\qquad A \in \mathbb{R}^{m\times n} \quad n << m$



$$\min_{x\in\mathbb{R}^n} \|b - Ax\|$$

Solve LSQ by

QR	SVD
1) $QR = A$	1) $U\Sigma V^T = A$ $\hat{U} = [u_1 \ u_2 \dots u_r] \in \mathbb{R}^{m\times r}$ $C(A) = C(\hat{U})$
2) $P_{C(A)} = P_{C(Q)} = QQ^T$ $QQ^Tb = Ax$	2) $P_{C(A)} = P_{C(U)} = \hat{U}\hat{U}^T$ $\hat{U}\hat{U}^Tb = Ax$
3) $QQ^Tb = QRx \Rightarrow$ $Rx = Q^Tb$ linear system with upper triangular matrix R find solution by backsubstitution	3) $\hat{U}\hat{U}^Tb = \hat{U}\hat{\Sigma}\hat{V}^Tx$

$A = \tilde{Q}R$

$C(A) = C(Q)$

$A = U\Sigma V^T$

$C(A) = C([u_1 \dots u_r])$

$\hat{U} \in \mathbb{R}^{m\times r}$

$\hat{\Sigma} \in \mathbb{R}^{r\times r}_m$

$\hat{V}^T \in \mathbb{R}^{r\times n}$

$$\Rightarrow \hat{\Sigma}\hat{V}^Tx = \hat{U}^Tb$$

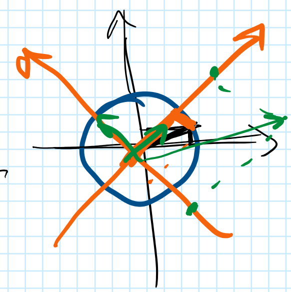
$$\hat{V}^Tx = \hat{\Sigma}^{-1}\hat{U}^Tb$$

$$x = \underbrace{\hat{V}\hat{\Sigma}^{-1}\hat{U}^T}_{\text{pseudo-inverse of } A}b = A^+b$$

When, in Octave, $x = A \backslash b$

if $A \in \mathbb{R}^{m \times m}$ & non-singular $x = A^{-1} b$

if $A \in \mathbb{R}^{m \times n}$ or singular $x = A^{+} b$



Simple exercises

1) SVD of $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \in \mathbb{R}^{2 \times 2}$

$$A = U \Sigma V^T = \underset{2\times 2}{\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}} \underset{2\times 2}{\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}} \underset{2\times 2}{\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}}$$

$\sigma_1 = 1$ $\sigma_2 = 1$ $r = \text{rank}(2)$

2) SVD of $A = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \in \mathbb{R}^{2 \times 2}$

$A = U \Sigma V^T$

$A \in \mathbb{R}^{m \times n}$; $U \in \mathbb{R}^{m \times m}$

$V \in \mathbb{R}^{n \times n}$

$\Sigma \in \mathbb{R}^{m \times n}$

$$A = \underset{2\times 2}{\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}} \underset{2\times 2}{\begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}} \underset{2\times 2}{\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}}$$

3) $A = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \in \mathbb{R}^{2 \times 2}$ $m = n = 2$ $\text{rank}(A) = 2$

$\sigma_1 =$

$$A = \underset{2\times 2}{\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}} \underset{2\times 2}{\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}} \underset{2\times 2}{\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}}$$

✓

Remember: $\sigma_1 \geq \sigma_2 \geq \sigma_3 \geq \ldots \geq \sigma_r > 0$ $\boxed{r = 2}$

$$\underline{\sigma_1 \geq \sigma_2 > 0}$$

$\sigma_1 = 1$

$\underline{\sigma_2 = 1}$

4) $$A = \begin{bmatrix} 2 & 0 \\ 0 & 3 \\ 0 & 0 \end{bmatrix} \in \mathbb{R}^{3\times 2}$$

$\sigma_1 = 3$

$\sigma_2 = 2$

$J = [e_1\ e_2]$

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

3x3   3x2   2x2

$U = [e_2\ e_1\ e_3]$

$J = [e_1\ e_2\ e_3]$

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

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

5) $$A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

rotation $\rightarrow$ does not change lengths

$\Rightarrow \sigma_1 = \sigma_2 = 1$

$$A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

6) Exploring idea of SVD:

"transform coordinates, scale, transform back"

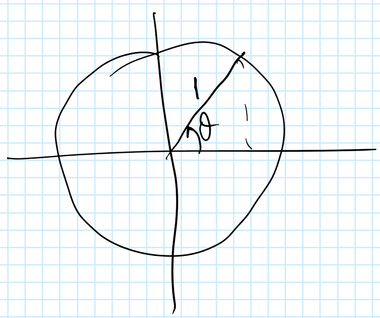
$$Ax = U\Sigma V^T x = U\Sigma(\underbrace{V^T x}_{1^{st}\text{ coord trans.}}) = U(\underbrace{\Sigma(V^T x)}_{\text{scaling}})$$

$$= U(\Sigma(V^T x))$$

↳ 2^{nd} coord transf.

$A \in \mathbb{R}^{2\times 2}$; Ask: how do vectors in $\mathbb{R}^2$ get "scaled" by A.

1) Choose vectors of unit norm & various directions



$$x = \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}$$

$$A = \begin{bmatrix} 2 & -1 \\ 3 & 1 \end{bmatrix}$$

$f: \mathbb{R}^2 \to \mathbb{R}^2 \quad f(x) = Ax = y$

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 2 & -1 \\ 3 & 1 \end{bmatrix}\begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix} \Rightarrow$$

$$y_1 = 2\cos\theta - \sin\theta$$

$$y_2 = 3\cos\theta + \sin\theta$$

$\cancel{(AB)^T = A^T B^T}$

$(AB)^T = B^T A^T$

$A = U\Sigma V^T \qquad A^T = (U\Sigma V^T)^T = V\Sigma^T U^T$

$$B = AA^T = U\Sigma \underbrace{V^T V}\Sigma^T U^T = U\Sigma I \Sigma^T U^T = U\Sigma\Sigma^T U^T$$

$$C = A^T A = V\Sigma^T \underbrace{U^T U}\Sigma V^T = V\Sigma^T\Sigma V^T$$

$$B = U \Lambda U^T \qquad \Lambda = \Sigma \Sigma^T \text{ diagonal matrix}$$

$$\Sigma \in \mathbb{R}^{m \times n} \qquad \Sigma^T \in \mathbb{R}^{n \times m} \qquad \Lambda \in \mathbb{R}^{m \times m}$$

eigendecomposition

$$\Lambda = \begin{bmatrix} \sigma_1 & & \\ & \sigma_2 & \\ & & \ddots \end{bmatrix} \begin{bmatrix} \sigma_1 & \sigma_2 & \dots \end{bmatrix} = \begin{bmatrix} \sigma_1^2 & & & \\ & \sigma_2^2 & & \\ & & \ddots & \\ & & & \sigma_r^2 \end{bmatrix}$$

$$B = AA^T = \begin{bmatrix} 2 & -1 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} 2 & 3 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 5 & 5 \\ 5 & 10 \end{bmatrix}$$

$$p_B(\lambda) = \det(\lambda I - B) = \begin{vmatrix} \lambda - 5 & -5 \\ -5 & \lambda - 10 \end{vmatrix} = (\lambda - 5)(\lambda - 10) - 25$$

$$p_B(\lambda) = \lambda^2 - 15\lambda + 25$$

$$\lambda_{1,2} = \frac{15 \pm \sqrt{225 - 4 \cdot 25}}{2} = \frac{15 \pm \sqrt{125}}{2} = \frac{15 \pm 5\sqrt{5}}{2}$$

(15

$$\lambda_1 \cong 13.09 \qquad \lambda_1 = \sigma_1^2 \Rightarrow \sigma_1 = 3.618$$

$$\cancel{\sigma_2 = -3.618}$$

$$\lambda_2 \cong 1.91 = \sigma_2^2 \Rightarrow \sigma_2 = 1.382$$

$$N(B - \lambda_1 I):$$

$$B - \lambda_1 I = \begin{bmatrix} 5 - 13.09 & 5 \\ 5 & 10 - 13.09 \end{bmatrix} \sim \begin{bmatrix} -8.09 & 5 \\ 0 & 0 \end{bmatrix}$$

$$-8.09\, x_1 + 5 x_2 = 0 \Rightarrow \begin{bmatrix} \frac{5}{8.09} \end{bmatrix} = \begin{bmatrix} 0.618 \end{bmatrix}$$

$$-8.09\,x_1 + 5x_2 = 0 \Rightarrow \begin{bmatrix} \frac{5}{8.09} \\ 1 \end{bmatrix} = \begin{bmatrix} 0.618 \\ 1 \end{bmatrix}$$

$$u_1 = \frac{1}{\sqrt{1+0.618^2}} \begin{bmatrix} 0.618 \\ 1 \end{bmatrix}$$