Northeastern University, Department of Mathematics

MATH G5110: Applied Linear Algebra and Matrix Analysis.

• Instructor: **He Wang** Email: **he.wang@northeastern.edu**

§13 Singular Value Decomposition

$D = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_m \end{bmatrix}$

1. Singular Value Decomposition

$= VDV^{-1}$

Recall the spectral decomposition for symmetric matrices:

$A = VDV^T$

Theorem 1 (Spectral Decomposition for Symmetric Matrices). *A is an $m \times m$* ***symmetric*** *matrix if an only if $A = VDV^{-1}$ such that D is diagonal and V is an orthogonal matrix.*
Let $\lambda_1, \ldots, \lambda_m$ be the diagonal entries of D, and let $\vec{v}_1, \ldots, \vec{v}_m$ be the column vectors of V. Then $A = VDV^T$ can be written as

$$A = \lambda_1 \Big(\vec{v}_1 \cdot (\vec{v}_1)^T \Big) + \cdots + \lambda_m \Big(\vec{v}_m \cdot (\vec{v}_m)^T \Big)$$

We want to find a similar decomposition for **any** $n \times m$ matrix M.

- $\text{rank}\, M = r$
- $\text{rank}\, M = \text{rank}\, M^T M = \text{rank}\, A$

- $A = M^T M$ symmetric, positive semi-definite.
- $A = VDV^T = VDV^{-1}$

$V^T = V^{-1}$ orthogonal

$D = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_m \end{bmatrix}$

$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_r > \lambda_{r+1} = \cdots = \lambda_m = 0$

$V = [\vec{v}_1 \cdots \vec{v}_m]$

- Goal: S.V.D

$\underset{n\times m}{M} = \underset{n\times n}{U} \underset{n\times m}{\Sigma} \underset{m\times m}{V^T} = \sigma_1 \vec{u}_1 \vec{v}_1^T + \cdots + \sigma_r \vec{u}_r \vec{v}_r^T$

orthogonal

"diagonal"

$\Sigma = \begin{bmatrix} \sigma_1 & & & & 0 \\ & \sigma_2 & & & \\ & & \ddots & & \\ & & & \sigma_r & \\ 0 & & & & 0 \\ & & & & & \ddots & 0 \end{bmatrix}$

- Def: The singular values of M are

$\sigma_i = \sqrt{\lambda_i}$ for $i = 1, \ldots, m$.

$$M^T M = V D V^T$$

$$V^T M^T M V = D$$

$$\boxed{(MV)^T MV = D} \iff \begin{bmatrix} (M\vec{v}_1)\cdot(M\vec{v}_1) & & & \\ & \ddots & & \\ & & (M\vec{v}_r)(M\vec{v}_r) & \\ & & & \ddots \\ & & & & (M\vec{v}_m)(M\vec{v}_m) \end{bmatrix} = \begin{bmatrix} \lambda_1 & & & & & \\ & \lambda_2 & & & 0 & \\ & & \ddots & & & \\ & & & \lambda_r & & \\ & 0 & & & 0 & \\ & & & & & \ddots \\ & & & & & & 0 \end{bmatrix}$$

$$MV = [M\vec{v}_1 \cdots M\vec{v}_m]$$

$$\boxed{\|M\vec{v}_i\| = \sigma_i}$$

① $\|M\vec{v}_i\|^2 = \lambda_i$ for $i = 1, \cdots, r$.

② $\|M\vec{v}_j\| = 0$ for $j = r+1, \cdots m$

③ $(M\vec{v}_i)\cdot(M\vec{v}_j) = 0$ for $i \neq j$

$$\underset{m\times m}{M}\ \underset{m\times m}{V} = [\underline{M\vec{v}_1} \cdots \underline{M\vec{v}_r}\ \ \underset{0}{\underset{\shortparallel}{M\vec{v}_{r+1}}} \cdots \underset{0}{\underset{\shortparallel}{M\vec{v}_m}}]$$

$$= [\sigma_1 \vec{u}_1 \cdots \sigma_r \vec{u}_r\ \boxed{0\vec{u}_{r+1} \cdots 0\vec{u}_m}]$$

<u>Denote</u> $\vec{u}_i := \dfrac{M\vec{v}_i}{\underbrace{\|M\vec{v}_i\|}_{=\sigma_i}}$ $i = 1, \cdots, r.$

$\iff M\vec{v}_i = \sigma_i \vec{u}_i$

$$= \underset{m\times m}{[\underline{\vec{u}_1 \cdots \vec{u}_r}\ \vec{u}_{r+1} \cdots \vec{u}_m]} \underset{m\times m}{\begin{bmatrix} \sigma_1 & & & & & \\ & \sigma_2 & & & 0 & \\ & & \ddots & & & \\ & & & \sigma_r & & \\ & 0 & & & 0 & \\ & & & & & 0 \\ & & & & & & 0 \end{bmatrix}}$$

- $\{\vec{u}_{r+1}, \cdots, \vec{u}_m\}$ <u>orthonormal basis</u> for $\mathrm{Span}\{\vec{u}_1 \cdots \vec{u}_r\}^\perp = (\mathrm{im}\, M)^\perp = \ker M^T$

$$= U\Sigma$$

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

$$= [\vec{u}_1 \cdots \vec{u}_m] \begin{bmatrix} \sigma_1 & & & 0 \\ & \ddots & & \\ & & \sigma_r & \\ 0 & & & 0 \end{bmatrix} \begin{bmatrix} \vec{v}_1^T \\ \vdots \\ \vec{v}_n^T \end{bmatrix}$$

Theorem 2 (Singular Value Decomposition(SVD))**.** *And $n \times m$ matrix M can be decomposed as*

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

or as

$$M = \sigma_1 \vec{u}_1 \vec{v}_1^T + \cdots + \sigma_r \vec{u}_r \vec{v}_r^T$$

$\text{rank}(\vec{u}_i \vec{v}_i^T) = 1$

Example 3. Find an SVD decomposition for the matrix

$$M = \begin{bmatrix} 1 & 4 \\ 2 & 2 \\ 2 & -4 \end{bmatrix}_{3\times 2} = \underset{3\times 3}{U}\ \underset{3\times 2}{\Sigma}\ \underset{2\times 2}{V^T}$$

Example 4. Find an SVD decomposition for the matrix

$$M = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix}$$

Step 1. $A = M^T M = \begin{bmatrix} 9 & 0 \\ 0 & 36 \end{bmatrix}$ $\qquad \sigma_1 = \sqrt{\lambda_1} = 6 \quad \sigma_2 = \sqrt{\lambda_2} = 3$

2. Eigenvalue of A $\quad \lambda_1 = 36 \quad \lambda_2 = 9.$ $\quad \Longrightarrow \Sigma = \begin{bmatrix} 6 & 0 \\ 0 & 3 \\ 0 & 0 \end{bmatrix}_{3\times 2}$

3. eigenvector of A $\quad \vec{v}_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad \vec{v}_2 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ $\quad \Longrightarrow V = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$

orthonormal!

4. $M\vec{v}_1 = \begin{bmatrix} 4 \\ 2 \\ -4 \end{bmatrix} \qquad \vec{u}_1 = \dfrac{M\vec{v}_1}{\|M\vec{v}_1\|} = \dfrac{1}{3}\begin{bmatrix} 2 \\ 1 \\ -2 \end{bmatrix}$

$M\vec{v}_2 = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix} \qquad \vec{u}_2 = \qquad = \dfrac{1}{3}\begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}$

$U = \begin{bmatrix} \vec{u}_1 & \vec{u}_2 & \vec{u}_3 \end{bmatrix}_{3\times 3}$ $\qquad \vec{u}_3$ is basis for $\ker(M^T) = \operatorname{im}(M)^\perp$

$U^T = U^{-1}$

$$\boxed{M = \sigma_1 \vec{u}_1 \vec{v}_1^T + \sigma_2 \vec{u}_2 \vec{v}_2^T}$$

$$\underset{3\times 2}{M} = \underset{3\times 3}{U}\ \underset{3\times 2}{\Sigma}\ \underset{2\times 2}{V^T}$$

$$= \underset{3\times 2}{\tilde{U}}\ \underset{2\times 2}{\Sigma}\ \underset{2\times 2}{V^T}$$

Example 5.

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

(1). Calculate $M^T M$ and MM^T.

(2). Find all eigenvalues and an eigenbasis of $M^T M$.

(3). Find all eigenvalues and an eigenbasis of MM^T.

(4). Find an SVD decomposition for the matrix M.

Applications.

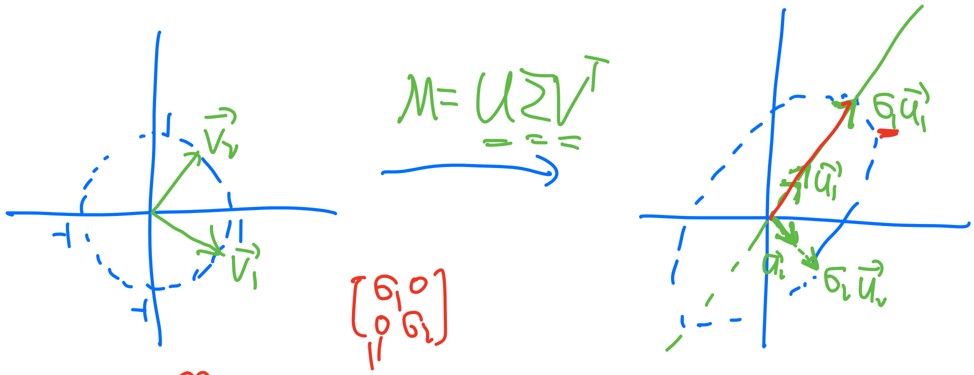
1. Geometric meaning in $\mathbb{R}^2$.

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

> **Theorem 6.** *Let M be an 2×2 invertible matrix. The image of M of the unit circle is an ellipse. The lengths of the semimajor and the semiminor axes of the ellipse are the singular values of M.*

$$\mathbb{R}^2 \xrightarrow{M} \mathbb{R}^2$$

$$\vec{x} \longrightarrow M\vec{x}$$



$$\Sigma = \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{bmatrix}$$

$$\mathbb{R}^2 \xrightarrow{V^T} \mathbb{R}^2 \xrightarrow{\Sigma} \mathbb{R}^2 \xrightarrow{U} \mathbb{R}^2$$

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

$$\mathbb{R}^2 \longrightarrow \mathbb{R}^2$$

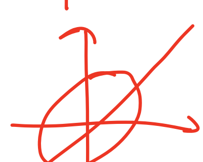
$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \longrightarrow \begin{bmatrix} 3x_1 \\ 2x_2 \end{bmatrix}$$

$$M = \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} = \underset{2\times 2}{U}\ \underset{2\times 3}{\Sigma}\ \underset{3\times 3}{V^T}$$

$$\mathbb{R}^3 \xrightarrow{2\times 3} \mathbb{R}^2$$

$$\xrightarrow{V^T} \mathbb{R}^3 \longrightarrow \mathbb{R}^2 \xrightarrow{U}$$

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



2. Solving least-squares problems.

3. Principal component analysis.

4. Digital image compressing. ✓

$$\underset{n\times n}{A} \rightsquigarrow A^{-1} \quad \text{s.t.} \quad A^{-1}A = I_n = AA^{-1}$$

<u>Def</u>: The $\boxed{\text{pseudo-inverse}}$ of $A \in \mathbb{R}^{m\times n}$ is

a <u>$n\times m$</u> matrix A^+ s.t.

① $(AA^+)A = A$

② $A^+(AA^+) = A^+$

③ AA^+ and A^+A symmetric

Ex ① If A is $n\times n$ invertible, then $A^+ := A^{-1}$

② If A is $m\times n$ with $\boxed{\operatorname{rank} A = n}$ ($m \geq n$), then define $A^+ := (A^TA)^{-1}A^T$

<u>In general</u>: $A \in \mathbb{R}^{m\times n}$

<u>S.V.D</u>: $A = U\Sigma V^T$

$$\Sigma = \begin{bmatrix} \sigma_1 & & & \\ & \ddots & & 0 \\ & & \sigma_r & \\ 0 & & & \ddots \end{bmatrix}_{m\times n}$$

<u>Define</u> $A^+ := V\Sigma^+U^T$

where $\Sigma^+ = \begin{bmatrix} \frac{1}{\sigma_1} & & & 0 \\ & \ddots & \frac{1}{\sigma_r} & \\ 0 & & & 0 \end{bmatrix}$

$_{n\times m}$

$$A^+A = V\Sigma^+U^TU\Sigma V$$

$$A(A^+A) = \cdots = A$$

$$\Sigma\Sigma^+ = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}_{m\times m}$$

$$\Sigma^+\Sigma = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}_{n\times n}$$

- Application: Estimate eigenvalues $\Rightarrow$ PCA

- App. Norm of a matrix M $m\times n$ matrix

magnifying power

Def: $$\|M\|_p := \sup_{\vec{x}\neq\vec{0}} \left\{ \frac{\|M\vec{x}\|_p}{\|\vec{x}\|_p} \;\middle|\; \text{for } \vec{x}\neq\vec{0} \right\}$$

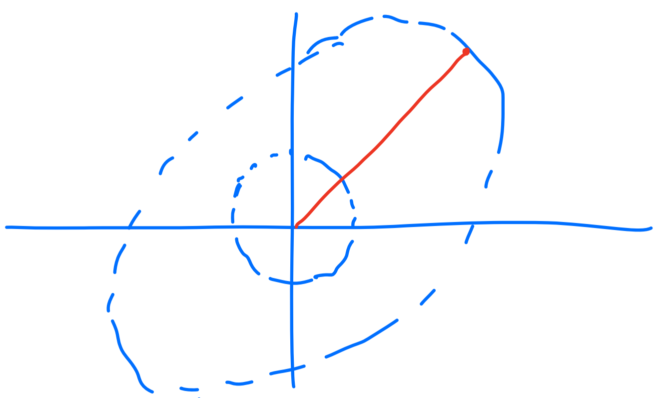
$\vec{x}, \vec{y} \in \mathbb{R}^n$

l_p-norm

$$= \sup_{\vec{y}=1} \left\{ \|M\vec{y}\|_p \;\middle|\; \|\vec{y}\|_p = 1 \right\}$$

e.g.

$M \in \mathbb{R}^{2\times 2}$



- $p=1$ $\quad \|M\|_1 = \max_j \left\{ \sum_i |m_{ij}| \right\}$

$p=\infty$ $\quad \|M\|_\infty = \max_i \left\{ \sum_j |m_{ij}| \right\}$

$p=2$ $\quad \|M\|_2 = \sigma_1$

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

$\|\vec{y}\| = 1$

$\|M\vec{y}\| = \|U\Sigma V^T \vec{y}\|$

U is orthogonal $= \|\Sigma V^T \vec{y}\|$

Denote $V^T\vec{y} = \vec{w}$

$\|\vec{U}\vec{x}\| = \|\vec{x}\|$

$= \|\Sigma \vec{w}\|$

$\|\vec{w}\| = 1$

$= \left\| \begin{bmatrix} \sigma_1 w_1 \\ \sigma_2 w_2 \\ \vdots \\ \sigma_r w_r \\ 0 \\ \vdots \\ 0 \end{bmatrix} \right\|$

$\Sigma \vec{w} = \begin{bmatrix} \sigma_1 & & & & 0 \\ & \sigma_2 & & & \\ & & \ddots & & \\ & & & \sigma_r & \\ 0 & & & & 0 \end{bmatrix} \begin{bmatrix} w_1 \\ \vdots \\ w_n \end{bmatrix}$

$m \times n$

$= \sqrt{\sigma_1^2 w_1^2 + \cdots + \sigma_r^2 w_r^2}$

$w_1^2 + \cdots + w_r^2 + \cdots + w_n^2 = 1$

$\leq \sqrt{\sigma_1^2 (w_1^2 + \cdots + w_r^2)}$

$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r$

$\leq \sigma_1$

$\vec{w} = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$

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

$A = M^T M$ $\quad$ $n \times n$ symmetric

$= V D V^{-1} = V D V^T$ $\quad$ $D = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}$

$\lambda_1 > \lambda_2 \geq \cdots \lambda_n$

- A irreducible
- $\lim_{k \to \infty} \left(\frac{1}{\lambda_1} A \right)^k = \frac{\vec{u} \vec{v}^T}{\vec{u} \cdot \vec{v}}$ $\quad$ here $A\vec{v} = \lambda_1 \vec{v}$, $\vec{u}^T A = \lambda_1 \vec{u}^T$

- $\lim_{k \to \infty} \left(\frac{1}{\lambda_1} A \right)^k \vec{x} = \frac{\vec{u} \vec{v}^T \vec{x}}{\vec{u} \cdot \vec{v}} = c \cdot \vec{u}$

We know A and $\vec{x}$

$$\left(\frac{1}{\lambda_1} A \right)^k \vec{x} \approx c \cdot \vec{u}$$

$$\left(\frac{1}{\lambda_1} \right)^k A^k \vec{x} \approx c \cdot \vec{u}$$

power method ; $\quad A^k \vec{x} \approx a \cdot \vec{u}$

Rayleigh quotient ; $\quad \frac{\vec{u}^T A \vec{u}}{\vec{u}^T \vec{u}} =: \lambda_1$