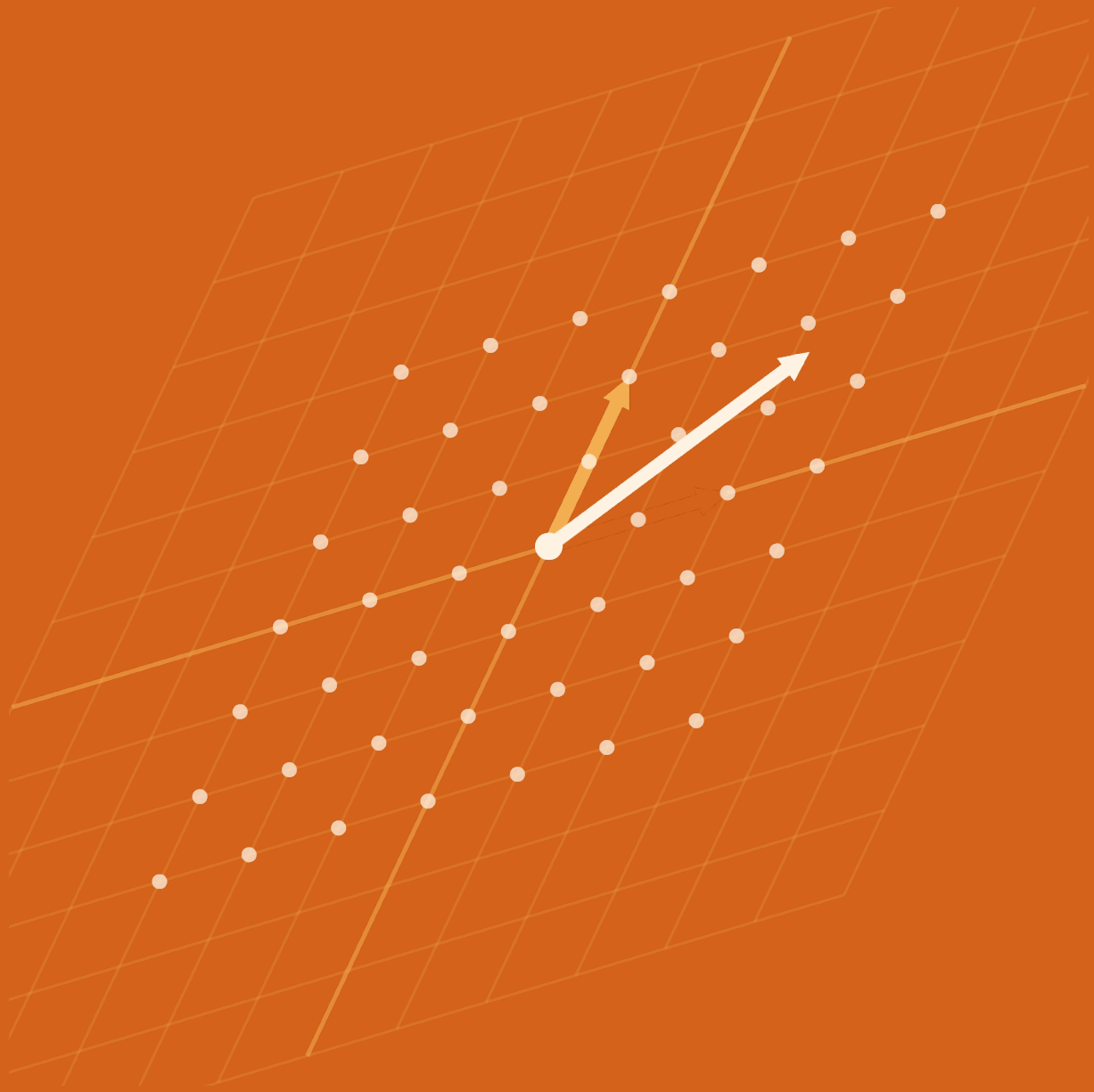
LINEAR ALGEBRA FOR AIML AND 6G

Singular Value Decomposition



Lecture 18

Prof. Aditya K. Jagannatham
IIT Kanpur

Outline

1. **Singular value decomposition**
   The factorization $A = U\Sigma V^T$ and its properties
2. **SVD example**
   Computing the SVD of a 2×2 matrix
3. **SVD and rank**
   Rank and the four fundamental subspaces
4. **Condition number**
   Ill-conditioned matrices and practical rank

01

Singular value decomposition

The factorization $A = U\Sigma V^T$ and its properties



SVD

- The **Singular Value Decomposition (SVD)** of any matrix is given as

$$\mathbf{A} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^T$$

$$= \begin{bmatrix} | & | & & | \\ \bar{\mathbf{u}}_1 & \bar{\mathbf{u}}_2 & \dots & \bar{\mathbf{u}}_m \\ | & | & & | \end{bmatrix} \begin{bmatrix} \sigma_1 & & \\ & \sigma_2 & \\ & & \dots \end{bmatrix} \begin{bmatrix} - & \bar{\mathbf{v}}_1^T & - \\ - & \bar{\mathbf{v}}_2^T & - \\ & \vdots & \\ - & \bar{\mathbf{v}}_n^T & - \end{bmatrix}$$

SVD properties

$$\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$$

- The matrix $\mathbf{U}$ is of size $m \times m$
- The matrix $\mathbf{V}$ is of size $n \times n$
- These are **orthogonal**

$$\mathbf{U}\mathbf{U}^T = \mathbf{U}^T\mathbf{U} = \mathbf{I}$$
$$\mathbf{V}\mathbf{V}^T = \mathbf{V}^T\mathbf{V} = \mathbf{I}$$

SVD properties

$$\mathbf{U} = \begin{bmatrix} | & | & & | \\ \bar{\mathbf{u}}_1 & \bar{\mathbf{u}}_2 & \dots & \bar{\mathbf{u}}_m \\ | & | & & | \end{bmatrix}$$

- Therefore, columns $\bar{\mathbf{u}}_i$ of $\mathbf{U}$ form an **orthonormal set**
  - These are termed **left singular vectors**.
  - $\mathbf{U}$ is termed the **left singular matrix**

SVD properties

$$\mathbf{V}^T = \begin{bmatrix} - & \bar{\mathbf{v}}_1^T & - \\ - & \bar{\mathbf{v}}_2^T & - \\ & \vdots & \\ - & \bar{\mathbf{v}}_n^T & - \end{bmatrix}$$

- Similarly, columns $\bar{\mathbf{v}}_i$ of $\mathbf{V}$ form an **orthonormal set**
  - These are termed **right singular vectors**.
  - $\mathbf{V}$ is termed the **right singular matrix**

SVD properties

- The matrix Σ is of size $m \times n$
- It has σ_i, $i = 1,2,\ldots,l$ along the **principal diagonal** where $l = \min\{m,n\}$

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

SVD properties

- The **singular values** are non-negative and arranged in decreasing order

$$\sigma_1 \geq \sigma_2 \geq \sigma_3 \geq \dots$$

- Furthermore,

$$\mathbf{A}\bar{\mathbf{v}}_i = \sigma_i \bar{\mathbf{u}}_i$$

SVD properties

- This can be seen as follows.
- Consider $i = 1$

$$\mathbf{A}\bar{\mathbf{v}}_1 = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T\bar{\mathbf{v}}_1$$

$$= \mathbf{U}\mathbf{\Sigma}\begin{bmatrix} - & \bar{\mathbf{v}}_1^T & - \\ - & \bar{\mathbf{v}}_2^T & - \\ & \vdots & \\ - & \bar{\mathbf{v}}_n^T & - \end{bmatrix}\bar{\mathbf{v}}_1 = \mathbf{U}\mathbf{\Sigma}\begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

SVD properties

- This can be simplified as

$$\mathbf{A}\bar{\mathbf{v}}_1 = \mathbf{U}\boldsymbol{\Sigma}\begin{bmatrix}1\\0\\\vdots\\0\end{bmatrix}$$

$$= \begin{bmatrix} | & | & & | \\ \bar{\mathbf{u}}_1 & \bar{\mathbf{u}}_2 & \dots & \bar{\mathbf{u}}_m \\ | & | & & | \end{bmatrix} \begin{bmatrix} \sigma_1 & & \\ & \sigma_2 & \\ & & \dots \end{bmatrix} \begin{bmatrix}1\\0\\\vdots\\0\end{bmatrix}$$

SVD properties

- which reduces to

$$\mathbf{A}\bar{\mathbf{v}}_1 = \begin{bmatrix} | & | & & | \\ \bar{\mathbf{u}}_1 & \bar{\mathbf{u}}_2 & \dots & \bar{\mathbf{u}}_m \\ | & | & & | \end{bmatrix} \begin{bmatrix} \sigma_1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \sigma_1 \bar{\mathbf{u}}_1$$

- Generalizing, it is easy to see that

$$\mathbf{A}\bar{\mathbf{v}}_i = \sigma_i \bar{\mathbf{u}}_i$$

SVD properties

- Observe also that

$$\mathbf{A} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^T$$
$$\mathbf{A}^T\mathbf{A} = \mathbf{V}\boldsymbol{\Sigma}^T\boldsymbol{\Sigma}\mathbf{V}^T$$

- Note that $\boldsymbol{\Sigma}^T\boldsymbol{\Sigma}$ is **diagonal**
- Therefore, $\bar{\mathbf{v}}_i$ form the **eigenvectors** of the matrix $\mathbf{A}^T\mathbf{A}$

SVD properties

- Note that $\mathbf{\Sigma}^T\mathbf{\Sigma}$ is

$$\mathbf{\Sigma}^T\mathbf{\Sigma} = \begin{bmatrix} \sigma_1 & & \\ & \sigma_2 & \\ & & \ddots \end{bmatrix} \begin{bmatrix} \sigma_1 & & \\ & \sigma_2 & \\ & & \ddots \end{bmatrix}$$

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

SVD properties

$$\mathbf{A}^T\mathbf{A} = \mathbf{V}\boldsymbol{\Sigma}^T\boldsymbol{\Sigma}\mathbf{V}^T$$

- Corresponding **eigenvalues** of $\mathbf{A}^T\mathbf{A}$ are $\lambda_i = \sigma_i^2$
- In other words, $\sigma_i = \sqrt{\lambda}_i$, where λ_i are eigenvalues of $\mathbf{A}^T\mathbf{A}$

SVD properties

- Similarly,

$$\begin{aligned}\mathbf{A}\mathbf{A}^T &= \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^T\mathbf{V}\boldsymbol{\Sigma}^T\mathbf{U}^T \\ &= \mathbf{U}\boldsymbol{\Sigma}\boldsymbol{\Sigma}^T\mathbf{U}^T\end{aligned}$$

- $\bar{\mathbf{u}}_i$ are **eigenvectors** of $\mathbf{A}\mathbf{A}^T$
- Once again, its **eigenvalues** are σ_i^2

02

SVD example

Computing the SVD of a 2 × 2 matrix



SVD example

- Consider the matrix below

$$\mathbf{A} = \begin{bmatrix} 1 & 5 \\ 7 & 5 \end{bmatrix}$$

$$\mathbf{A}^T\mathbf{A} = \begin{bmatrix} 50 & 40 \\ 40 & 50 \end{bmatrix}$$

SVD example

- Its eigenvalues are

$$\mathbf{A}^T\mathbf{A} - \lambda\mathbf{I} = \begin{bmatrix} 50-\lambda & 40 \\ 40 & 50-\lambda \end{bmatrix}$$

$$(50-\lambda)^2 = 40^2$$

$$50-\lambda = \pm 40$$

$$\lambda = 90, 10$$

$$\Rightarrow \sigma_1 = \sqrt{90}, \sigma_2 = \sqrt{10}$$

SVD example

- Its **eigenvector** of $\lambda = 90$ is

$$\mathbf{A}^T\mathbf{A} - 90\mathbf{I} = \begin{bmatrix} -40 & 40 \\ 40 & -40 \end{bmatrix}$$

$$\begin{bmatrix} -40 & 40 \\ 40 & -40 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \mathbf{0}$$

$$\Rightarrow v_1 = v_2$$

SVD example

- Eigenvector is

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix} \sim \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$$

SVD example

- Its eigenvector of $\lambda = 10$ is

$$\mathbf{A}^T\mathbf{A} - 10\mathbf{I} = \begin{bmatrix} 40 & 40 \\ 40 & 40 \end{bmatrix}$$

$$\begin{bmatrix} 40 & 40 \\ 40 & 40 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \mathbf{0}$$

$$\Rightarrow v_1 + v_2 = 0 \Rightarrow v_1 = -v_2$$

SVD example

- Eigenvector is

$$\begin{bmatrix} 1 \\ -1 \end{bmatrix} \sim \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{bmatrix}$$

SVD example

- Furthermore $\bar{\mathbf{u}}_1$ can be computed as,

$$\mathbf{A}\bar{\mathbf{v}}_1 = \sigma_1 \bar{\mathbf{u}}_1$$

$$\Rightarrow \bar{\mathbf{u}}_1 = \frac{1}{\sigma_1}\mathbf{A}\bar{\mathbf{v}}_1$$

$$\bar{\mathbf{u}}_1 = \frac{1}{\sqrt{90}}\begin{bmatrix}1 & 5\\ 7 & 5\end{bmatrix}\begin{bmatrix}\frac{1}{\sqrt{2}}\\ \frac{1}{\sqrt{2}}\end{bmatrix} = \frac{1}{\sqrt{180}}\begin{bmatrix}6\\ 12\end{bmatrix} = \begin{bmatrix}\frac{1}{\sqrt{5}}\\ \frac{2}{\sqrt{5}}\end{bmatrix}$$

SVD example

- Next $\bar{\mathbf{u}}_2$ can be determined as,

$$\mathbf{A}\bar{\mathbf{v}}_2 = \sigma_2 \bar{\mathbf{u}}_2$$

$$\Rightarrow \bar{\mathbf{u}}_2 = \frac{1}{\sigma_2}\mathbf{A}\bar{\mathbf{v}}_2$$

$$\bar{\mathbf{u}}_2 = \frac{1}{\sqrt{10}}\begin{bmatrix} 1 & 5 \\ 7 & 5 \end{bmatrix}\begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{bmatrix} = \frac{1}{\sqrt{20}}\begin{bmatrix} -4 \\ 2 \end{bmatrix} = \begin{bmatrix} -\frac{2}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} \end{bmatrix}$$

SVD example

- Therefore, the SVD of $\mathbf{A}$ is given as

$$\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$$

$$= \begin{bmatrix} \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{bmatrix} \begin{bmatrix} \sqrt{90} & 0 \\ 0 & \sqrt{10} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}$$

03

SVD and rank

Rank and the four fundamental subspaces



SVD and rank

- Let $\sigma_1, \sigma_2, \ldots, \sigma_r > 0$

$$\sigma_{r+1} = \sigma_{r+2} = \ldots = 0$$

- Then

$$r = \text{rank}(\mathbf{A})$$

- i.e. rank of matrix equals number of **non-zero singular values**.

SVD and rank

- Then $\bar{\mathbf{u}}_1, \bar{\mathbf{u}}_2, \ldots, \bar{\mathbf{u}}_r$ form an orthonormal basis for $\mathrm{colsp}(\mathbf{A})$
- Moreover, $\bar{\mathbf{u}}_{r+1}, \bar{\mathbf{u}}_{r+2}, \ldots, \bar{\mathbf{u}}_m$ form an orthonormal basis for $\mathcal{N}(\mathbf{A}^T)$

SVD and rank

- Similarly $\bar{\mathbf{v}}_1, \bar{\mathbf{v}}_2, \ldots, \bar{\mathbf{v}}_r$ form an orthonormal basis for $\mathrm{rowsp}(\mathbf{A})$
- Furthermore, $\bar{\mathbf{v}}_{r+1}, \bar{\mathbf{v}}_{r+2}, \ldots, \bar{\mathbf{v}}_n$ form an orthonormal basis for $\mathcal{N}(\mathbf{A})$
- Thus the SVD can be used to determine the bases of the **four fundamental subspaces** of any matrix $\mathbf{A}$.

04

Condition number

Ill-conditioned matrices and practical rank



Condition number

- The **condition number** of an $n \times n$ matrix is the ratio between **largest and smallest singular value**

$$\kappa = \frac{\sigma_1}{\sigma_l}$$

Condition number

- A matrix with a very large condition number is said to be **ill-conditioned;**
- In practice, this occurs when the condition number is larger than the reciprocal of the **machine's precision**,
  - e.g., 10^7 for most single precision arithmetic.

Condition number

- In the extreme case when $\mathbf{A}$ has one or more **zero singular values**,
- so $\sigma_l = 0$, its **condition number is infinite**, and
  - the linear system is singular, with either **no solution or infinitely many solutions**.

Condition number – example

- Consider

$$\mathbf{A} = \begin{bmatrix} 1 & -2 & 2 \\ 2 & -4 & 4 \\ -1 & 2 & -2 \end{bmatrix}$$

- This matrix has rank $r = 1$

Condition number – example

- Now consider

$$\tilde{\mathbf{A}} = \begin{bmatrix} 1.00001 & -2 & 2 \\ 2 & -4.00001 & 4 \\ -1 & 2 & -2.00001 \end{bmatrix}$$

- Note $\mathbf{A} \approx \tilde{\mathbf{A}}$.
- However, this matrix has rank $r = 3$

Condition number – example

- The singular values of $\tilde{\mathbf{A}}$ are below

$$\sigma_1 = 7.34848, \sigma_2 \approx \sigma_3 \approx 10^{-5}$$

$$\kappa = \frac{\sigma_1}{\sigma_3} \approx 7.7 \times 10^5$$

- Thus, **condition number** is very high!
  - Hence, although its rank is 3, one should treat it as a rank 1 matrix

Practical rank

- Practical method for computing the rank of a matrix is to first assign a threshold, e.g., 10^{-5}, for **singular values**,
  - and then treat any small **singular value** lying below the threshold as if it were zero.

Summary

- Every matrix has an SVD $\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$ with orthogonal $\mathbf{U}$, $\mathbf{V}$ and diagonal $\mathbf{\Sigma}$
- Singular values satisfy $\sigma_1 \geq \sigma_2 \geq \ldots \geq 0$ and $\mathbf{A}\bar{\mathbf{v}}_i = \sigma_i \bar{\mathbf{u}}_i$
- $\bar{\mathbf{v}}_i$ and $\bar{\mathbf{u}}_i$ are eigenvectors of $\mathbf{A}^T\mathbf{A}$ and $\mathbf{A}\mathbf{A}^T$, with eigenvalues σ_i^2
- The rank is the number of non-zero singular values; $\mathbf{U}$ and $\mathbf{V}$ give bases of the four subspaces
- The **condition number** $\kappa = \sigma_1/\sigma_l$ is large for nearly singular matrices



Thank You!

Questions?