

Lecture 25

Singular Value I: Forms and Properties of SVD

Instructor: Ruoyu Sun



香港中文大學(深圳)

The Chinese University of Hong Kong, Shenzhen

数据科学学院

School of Data Science

Roadmap of MATH2041

Segment 1

Solving linear systems (Lec 3-15);
Solving least squares problem (Lect 16,17)

Segment 2

[Lec 18-21]: Two relatively independent parts:
—Determinant. [Important tool!]
—Linear transformation. [Advanced math perspective of matrix]

Segment 3

Next: Lec 22-27: Eigenvalues and related.

Schedule:

- Eigenvalues. Lec 22-24
- Singular values. Lec 25-26
- Quadratic forms. Lec 27.

Outline of SVD

- 1) **Forms of SVD.**
- 2) **Computation of SVD.**
- 3) **Properties of SVD.**
- 4) PCA and image/data compression (if time permitting)

Review

Eigenvalues and Eigenvectors

Definition 21.1 (Eigenvalues and Eigenvectors)

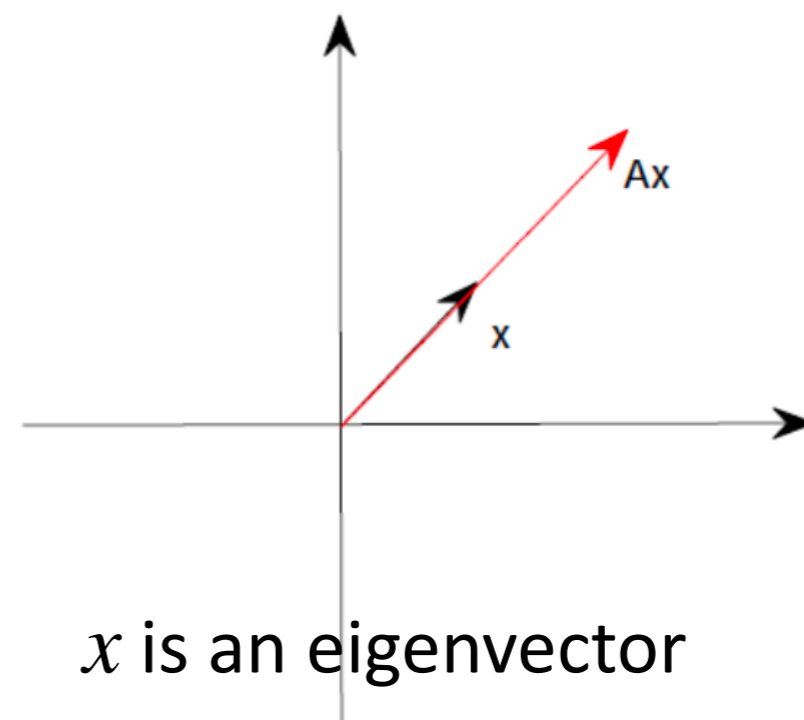
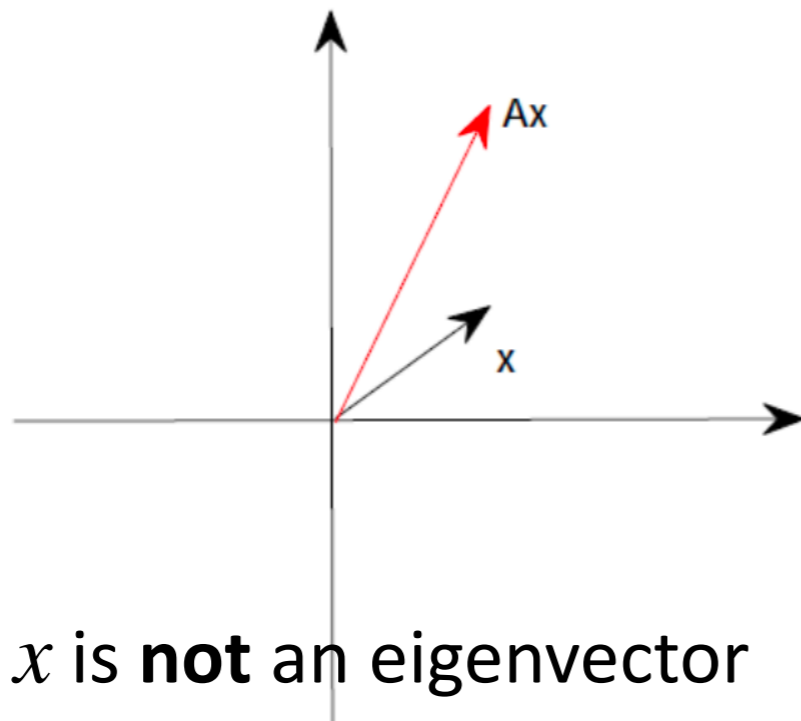
Let $A \in \mathbb{C}^{n \times n}$ be a square matrix.

If there exists a scalar λ ($\in \mathbb{R}$ or \mathbb{C}) and a **nonzero** vector x such that

$$Ax = \lambda x,$$

then λ is called a (real or complex) **eigenvalue** and x is called an **eigenvector** with respect to (or associated with; corresponding to) λ .

\mathbb{C} := set of complex numbers



General Procedure to Find Eigenvalues/Eigenvectors

Step 1: Solve $\det(A - \lambda I) = 0$ and get n roots $\lambda_1, \dots, \lambda_n$

Solving single-var polynomial.

Step 2: For each λ_i , find the eigenspace $\text{Null}(\lambda_i I - A)$

Solving up to n linear systems.

Any nonzero vector in $\text{Null}(\lambda_i I - A)$ is an eigenvector corresponding to λ_i

$$\text{Null}(\lambda_i I - A) = \left\{ \begin{array}{l} \text{eigenvectors of } A \text{ with} \\ \text{respect to the eigenvalue } \lambda_i \end{array} \right\} \cup \{0\}$$

Spectral Theorem

Theorem 23.1 [Spectral Theorem]

Any real symmetric matrix A can be written as

$$A = VDV^T = \sum_{j=1}^n \lambda_j \mathbf{v}_j \mathbf{v}_j^T. \quad (*)$$

where $D = \text{diag}(\lambda_1, \dots, \lambda_n)$ is a real diagonal matrix,

$V = [\mathbf{v}_1, \dots, \mathbf{v}_n]$ is a real orthogonal matrix.

(*): Eigenvalue decomposition (EVD) or eigendecomposition of A .

Judgement Question

Judgement: Any real square matrix A can be written as $A = SDS^T$ where D is a diagonal matrix, and S is an orthogonal matrix.

Judgement: If $A, B \in \mathbb{R}^{n \times n}$ are similar, then A and B have the same multiset of eigenvalues.

Judgement: If $A, B \in \mathbb{R}^{n \times n}$ have the same multiset of eigenvalues, then they are similar.

Judgement: A real square matrix is similar to a real diagonal matrix.

Judgement: If $A, B \in \mathbb{R}^{n \times n}$ are similar, then $\text{tr}(A) = \text{tr}(B)$.

Part I SVD

Judgement Questions

Q1: Any rectangular matrix has at least one complex eigenvalue.

Q2: Eigenvectors of an $n \times n$ matrix can form a basis \mathbb{C}^n .

Q3: Eigenvectors of an $n \times n$ matrix corresponding to different eigenvalues are orthogonal.

Facts About Eigenvalues

Fact 1: Eigenvalues are defined for _____ matrices, NOT _____ matrices.

Fact 2: The eigenvectors of a general matrix _____ form a basis of the whole space.

Fact 3: The eigenvectors of a general matrix _____ orthogonal.

Textbook: Singular vectors solve these problems in a perfect way.
SVD is a highlight of linear algebra.

Full SVD

Rectangular diagonal matrix $\Sigma \in \mathbb{R}^{m \times n}$ is a matrix of the form

$$\Sigma = \begin{bmatrix} \sigma_1 & 0 & \dots & 0 \\ 0 & \sigma_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_n \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{bmatrix} \quad \text{or} \quad \Sigma = \begin{bmatrix} \sigma_1 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & \sigma_2 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & 0 & \dots & 0 \\ 0 & 0 & \dots & \sigma_m & 0 & \dots & 0 \end{bmatrix}$$

Denote as $\Sigma = \text{rec-diag}(\sigma_1, \dots, \sigma_{\min(m,n)})_{m \times n}$

Theorem 25.1 [Full SVD]

Any matrix $A \in \mathbb{R}^{m \times n}$ can be written as

$$A = U\Sigma V^T = \sum_{j=1}^{\min(m,n)} \sigma_j \mathbf{u}_j \mathbf{v}_j^T, \quad (*)$$

where $U \in \mathbb{R}^{m \times m}$, $V \in \mathbb{R}^{n \times n}$ are two orthogonal matrices,

$\Sigma = \text{rec-diag}(\sigma_1, \dots, \sigma_{\min(m,n)})_{m \times n} \in \mathbb{R}^{m \times n}$ is a rectangular diagonal matrix.

Full SVD

Rectangular diagonal matrix $\Sigma \in \mathbb{R}^{m \times n}$ is a matrix of the form

$$\Sigma = \begin{bmatrix} \sigma_1 & 0 & \dots & 0 \\ 0 & \sigma_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_n \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{bmatrix} \quad \text{or} \quad \Sigma = \begin{bmatrix} \sigma_1 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & \sigma_2 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & 0 & \dots & 0 \\ 0 & 0 & \dots & \sigma_m & 0 & \dots & 0 \end{bmatrix}$$

Denote as $\Sigma = \text{rec-diag}(\sigma_1, \dots, \sigma_{\min(m,n)})_{m \times n}$

Theorem 25.1 [Full SVD]

Any matrix $A \in \mathbb{R}^{m \times n}$ can be written as

$$A = U\Sigma V^T = \sum_{j=1}^{\min(m,n)} \sigma_j \mathbf{u}_j \mathbf{v}_j^T, \quad (*)$$

where $U \in \mathbb{R}^{m \times m}$, $V \in \mathbb{R}^{n \times n}$ are two orthogonal matrices,

$\Sigma = \text{rec-diag}(\sigma_1, \dots, \sigma_{\min(m,n)})_{m \times n} \in \mathbb{R}_+^{m \times n}$ is a rectangular diagonal matrix.

Corollary 25.1:

If (*) holds, then columns of U are **eigenvectors** of $AA^T \in \mathbb{R}^{m \times m}$, called **left singular vectors of A**;

columns of V are **eigenvectors** of $A^T A \in \mathbb{R}^{n \times n}$, called **right singular vectors of A**;

σ_j 's are called **singular values of A**, and $\sigma_j = \sqrt{\lambda_j}$ where $\lambda_1, \dots, \lambda_{\min(m,n)}$ are **eigenvalues** of both AA^T and $A^T A$.

Proof Preparation: Two Lemmas

Lemma 25.1 [eigenvectors and singular vectors]

Suppose $A^T Av = \lambda v$, where $\lambda \neq 0, v \in \mathbb{R}^n \setminus \{\mathbf{0}_n\}$, then $Av \neq \mathbf{0}_m$.

Suppose $A^T Av = \mathbf{0}_n$, then $Av = \mathbf{0}_m$.

Lemma 25.2 [singular-equation-pairs in matrix form]

Suppose $A \in \mathbb{R}^{m \times n}$, $Av_1 = \sigma_1 u_1, \dots, Av_k = \sigma_k u_k$, where $v_j \in \mathbb{R}^{n \times 1}, u_j \in \mathbb{R}^{m \times 1}, \sigma_j \in \mathbb{R}$.

Then $AV = UD$, (*)

where $D = \text{diag}(\sigma_1, \dots, \sigma_k)$ is a diagonal matrix, $V = [v_1, \dots, v_k], U = [u_1, \dots, u_k]$.

Proof Analysis (II): What's Next after $AV = UD$?

Lemma 25.1 [eigenvectors and singular vectors]

Suppose $A^T Av = \lambda v$, where $\lambda \neq 0, v \in \mathbb{R}^n \setminus \{\mathbf{0}_n\}$, then $Av \neq \mathbf{0}_m$.

Suppose $A^T Av = \mathbf{0}_n$, where $v \in \mathbb{R}^n \setminus \{\mathbf{0}_n\}$, then $Av = \mathbf{0}_m$.

From Lemma 25.1, we can construct a few equations

Lemma 25.2 [singular-equation-pairs in matrix form]

Suppose $A \in \mathbb{R}^{m \times n}$, $Av_1 = \sigma_1 u_1, \dots, Av_k = \sigma_k u_k$, where $v_j \in \mathbb{R}^{n \times 1}, u_j \in \mathbb{R}^{m \times 1}, \sigma_j \in \mathbb{R}$.

Then $AV = UD$, (*)

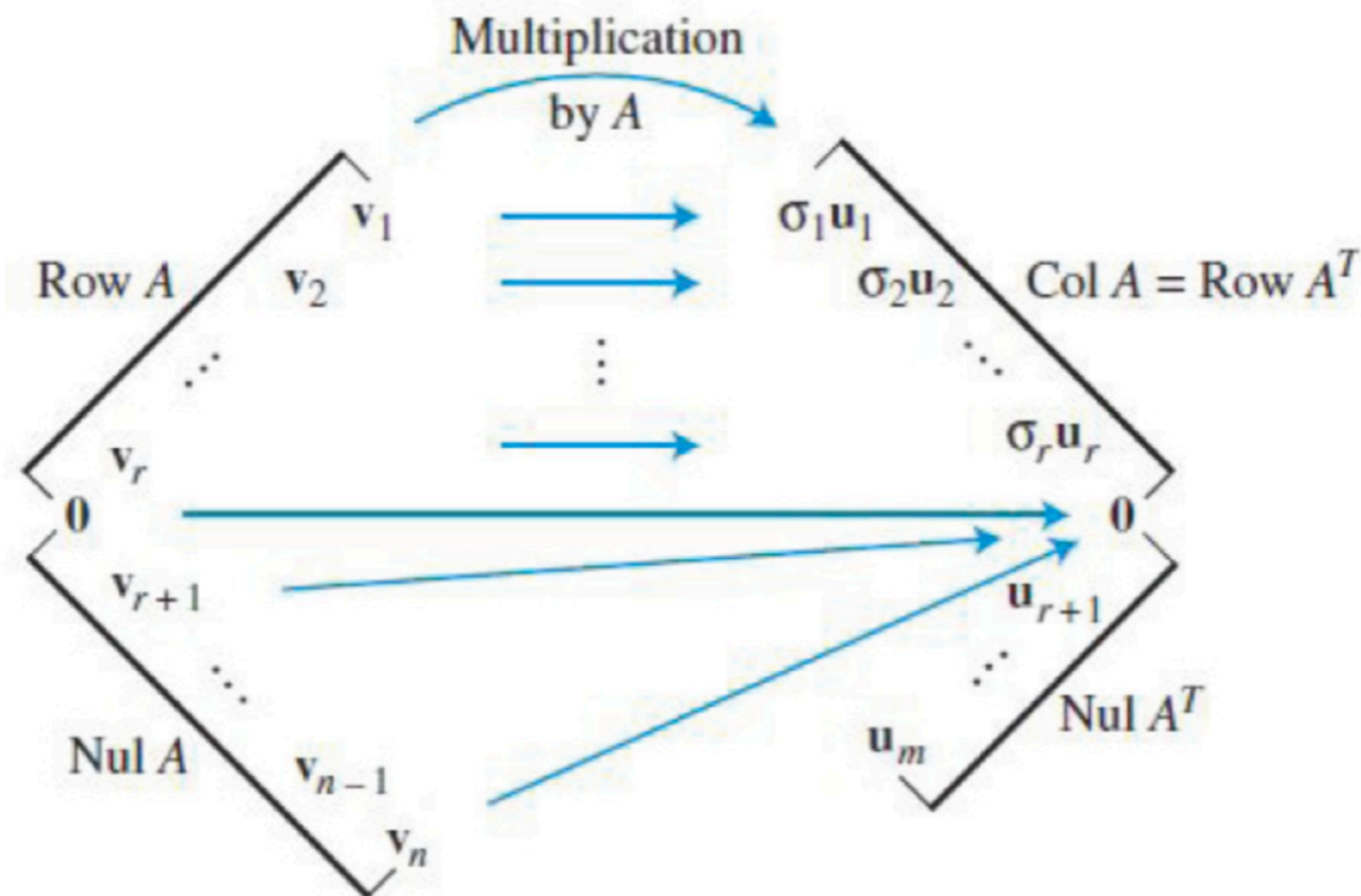
where $D = \text{diag}(\sigma_1, \dots, \sigma_k)$ is a diagonal matrix, $V = [v_1, \dots, v_k], U = [u_1, \dots, u_k]$.

From $AV = UD$, can we get an expression of A?

Proof Analysis (III): Completing Basis

$$V = [v_1, \dots, v_r], U = [u_1, \dots, u_r].$$

From $AV = UD$, can we get an expression of A ?



Wrap-up: Proof Sketch of full SVD Theorem

Assume $m \geq n$ (for case $m > n$, consider $\tilde{A} = A^\top$).

Step 1: Suppose $A^\top A$ has eigenvalues $\lambda_1 \geq \dots \geq \lambda_r > 0 = \lambda_{r+1} = \dots = \lambda_n$ (1).

By spectral theorem, there exists **orthonormal eigenbasis** of $A^\top A$: $\{\mathbf{v}_1, \dots, \mathbf{v}_r, \mathbf{v}_{r+1}, \dots, \mathbf{v}_n\}$,

$$\text{s.t. } A^\top A \mathbf{v}_j = \lambda_j \mathbf{v}_j.$$

Step 2: We have $A \mathbf{v}_n = \mathbf{0}$ because $\|A \mathbf{v}_n\|^2 = \mathbf{v}_n^\top A^\top A \mathbf{v}_n = \mathbf{v}_n^\top (\lambda_n \mathbf{v}_n) = 0$.

[Essentially proved: $A^\top A \mathbf{v} = \mathbf{0} \Rightarrow A \mathbf{v} = \mathbf{0}$.]

By same reasoning, $A \mathbf{v}_{r+1} = \dots = A \mathbf{v}_n = \mathbf{0}$ (2).

Step 3: Let $\sqrt{\lambda_1} \mathbf{u}_1 = A \mathbf{v}_1, \dots, \sqrt{\lambda_r} \mathbf{u}_r = A \mathbf{v}_r$ (3)

Easy to show: they are unit-norm, and orthogonal.

Let $\{\mathbf{u}_{r+1}, \dots, \mathbf{u}_m\} \subseteq \mathbb{R}^m$ be orthonormal basis of the span $(\{\mathbf{u}_1, \dots, \mathbf{u}_r\})^\perp$.

By (1),(2): $\sqrt{\lambda_{r+1}} \mathbf{u}_{r+1} = \mathbf{0} = A \mathbf{v}_1, \dots, \sqrt{\lambda_n} \mathbf{u}_n = \mathbf{0} = A \mathbf{v}_n$ (4)

Step 4: Verify $AV = U\Sigma$ using (3),(4). Since $V^\top V = I_n$, get $A = U\Sigma V^\top$.

Remark: Directly put (3), (4) in matrix form, can get reduced SVD, not full SVD.

Compact SVD

Theorem 25.2 [Compact SVD and vector form]

Any matrix $A \in \mathbb{R}^{m \times n}$ with rank r can be written as

$$A = U_r \Sigma_r V_r^\top = \sum_{j=1}^r \sigma_j \mathbf{u}_j \mathbf{v}_j^\top, \quad (*)$$

where $U_r = [\mathbf{u}_1, \dots, \mathbf{u}_r] \in \mathbb{R}^{m \times r}$ has orthogonal unit-norm columns,
 $V_r = [\mathbf{v}_1, \dots, \mathbf{v}_r] \in \mathbb{R}^{n \times r}$ has orthogonal unit-norm columns,
 $\Sigma_r = \text{diag}(\sigma_1, \dots, \sigma_r) \in \mathbb{R}_+^{r \times r}$ is a diagonal matrix.

Proof: Direct Corollary of Theorem 25.1

Number of singular values

For $A \in \mathbb{R}^{m \times n}$, $A = U\Sigma V^T$ and $\text{rank}(A)=r$, one has

$\lambda_1 \geq \dots \geq \lambda_r > \lambda_{r+1} = \dots = \lambda_n = 0$ are the eigenvalues of $A^T A$.

$\lambda_1 \geq \dots \geq \lambda_r > \lambda_{r+1} = \dots = \lambda_m = 0$ are the eigenvalues of AA^T .

If $m \geq n$, singular values are

If $m \leq n$, singular values are

In summary, the number of singular values is _____.

Part II Computation of SVD

Example 1: Compute SVD

Problem 1: Find the SVD of the matrix $A = \begin{bmatrix} 3 & 0 \\ 4 & 5 \end{bmatrix}$.

Solution:

Step 1: Find eigenvalues of $A^T A$.

$$A^T A = \begin{bmatrix} 25 & 20 \\ 20 & 25 \end{bmatrix}$$

$$A A^T = \begin{bmatrix} 9 & 12 \\ 12 & 41 \end{bmatrix}$$

Eigenvalues of $A^T A$ are 45 and 5.

Problem 1: Step 1 and Step 2

Problem 1: Find the SVD of the matrix $A = \begin{bmatrix} 3 & 0 \\ 4 & 5 \end{bmatrix}$.

Solution:

Step 1: Find eigenvalues of AA' and $A'A$.

$$A^T A = \begin{bmatrix} 25 & 20 \\ 20 & 25 \end{bmatrix} \quad AA^T = \begin{bmatrix} 9 & 12 \\ 12 & 41 \end{bmatrix}$$

Eigenvalues of $A^T A$ are 45 and 5.

Step 2: Find (unit) eigenvectors of $A'A$, i.e. right singular vectors of A .

$$\begin{bmatrix} 25 & 20 \\ 20 & 25 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = 45 \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad \begin{bmatrix} 25 & 20 \\ 20 & 25 \end{bmatrix} \begin{bmatrix} -1 \\ 1 \end{bmatrix} = 5 \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

Right singular vectors $v_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ $v_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 1 \end{bmatrix}$

Problem 1: Step 3 and Remark

Step 3: Find left singular vectors of A. Left singular vectors $u_i = \frac{Av_i}{\sigma_i}$

Now compute Av_1 and Av_2 which will be $\sigma_1 u_1 = \sqrt{45} u_1$ and $\sigma_2 u_2 = \sqrt{5} u_2$:

$$Av_1 = \frac{3}{\sqrt{2}} \begin{bmatrix} 1 \\ 3 \end{bmatrix} = \sqrt{45} \frac{1}{\sqrt{10}} \begin{bmatrix} 1 \\ 3 \end{bmatrix} = \sigma_1 u_1$$

$$Av_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} -3 \\ 1 \end{bmatrix} = \sqrt{5} \frac{1}{\sqrt{10}} \begin{bmatrix} -3 \\ 1 \end{bmatrix} = \sigma_2 u_2$$

Conclusion: The SVD of A is $A = U\Sigma V^T$ where

$$U = \frac{1}{\sqrt{10}} \begin{bmatrix} 1 & -3 \\ 3 & 1 \end{bmatrix}$$

$$\Sigma = \begin{bmatrix} \sqrt{45} & \\ & \sqrt{5} \end{bmatrix}$$

$$V = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$$

Remark: Verify your SVD by checking _____

Example 2: Compute SVD

Problem 2: Find the (full) SVD of the matrix $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 0 & 0 \end{bmatrix}$.

Solution:

Step 1: Find eigenvalues of $A^T A$ (which is 2 by 2, where $2 = \min(m,n)$)

$$A^T A = \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix}$$

Eigenvalues of $A^T A$ are 4 and 0.

Example 2: Compute SVD

Problem 1: Find the SVD of the matrix $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 0 & 0 \end{bmatrix}$.

Solution:

Step 1: Find eigenvalues of $A^T A$ (which is 2 by 2, where $2 = \min(m,n)$)

$$A^T A = \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix}$$

Eigenvalues of $A^T A$ are 4 and 0.

Step 2: Find (unit) eigenvectors of $A^T A$, i.e. right singular vectors of A .

The unit eigenvector w.r.t. $\lambda_1 = 4$ is $\mathbf{v}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

The unit eigenvector w.r.t. $\lambda_2 = 0$ is $\mathbf{v}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$

Problem 1: Step 3 and Remark

Step 3: Find left singular vectors of A. **Left singular vectors** $u_i = \frac{Av_i}{\sigma_i}$

Now $\sigma_1 = \sqrt{\lambda_1} = 2$ so

$$u_1 = \frac{1}{\sigma_1} Av_1 = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 0 & 0 \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$$

An orthonormal basis for $\text{Null}(A^T) = \text{Null}(AA^T)$ is

$$\{u_2, u_3\} = \left\{ \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$$

Conclusion: The full SVD of A is

$$\begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 0 & 0 \end{bmatrix} = A = U\Sigma V^T = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}$$

Remark: Verify your SVD by checking _____

Part III Properties of SVD

Four Fundamental Subspaces

$$A = [U_1, \dots, U_r, U_{r+1}, \dots, U_m] \begin{bmatrix} \sigma_1 & & & & \\ & \sigma_2 & & & \\ & & \ddots & & \\ & & & \sigma_r & \\ 0 & \dots & & 0 & \\ \vdots & & & & \ddots \end{bmatrix} \begin{bmatrix} V_1^T \\ \vdots \\ V_r^T \\ V_{r+1}^T \\ \vdots \\ V_n \end{bmatrix}$$

$\mathcal{R}(A)$ column space : dim r , orthonormal basis U_1, U_2, \dots, U_r ,
 null space : dim $n-r$, — — — basis V_{r+1}, \dots, V_n .
 row space : dim r , — — — basis V_1, \dots, V_r .
 left null space : dim $m-r$, — — — basis U_{r+1}, \dots, U_m .

Row Space and Columns Space

Lemma 25.1 Suppose $A = \sum_{j=1}^k \mathbf{x}_j \mathbf{y}_j^\top$,

where

\mathbf{x}_j 's are linearly independent, \mathbf{y}_j 's are linearly independent.

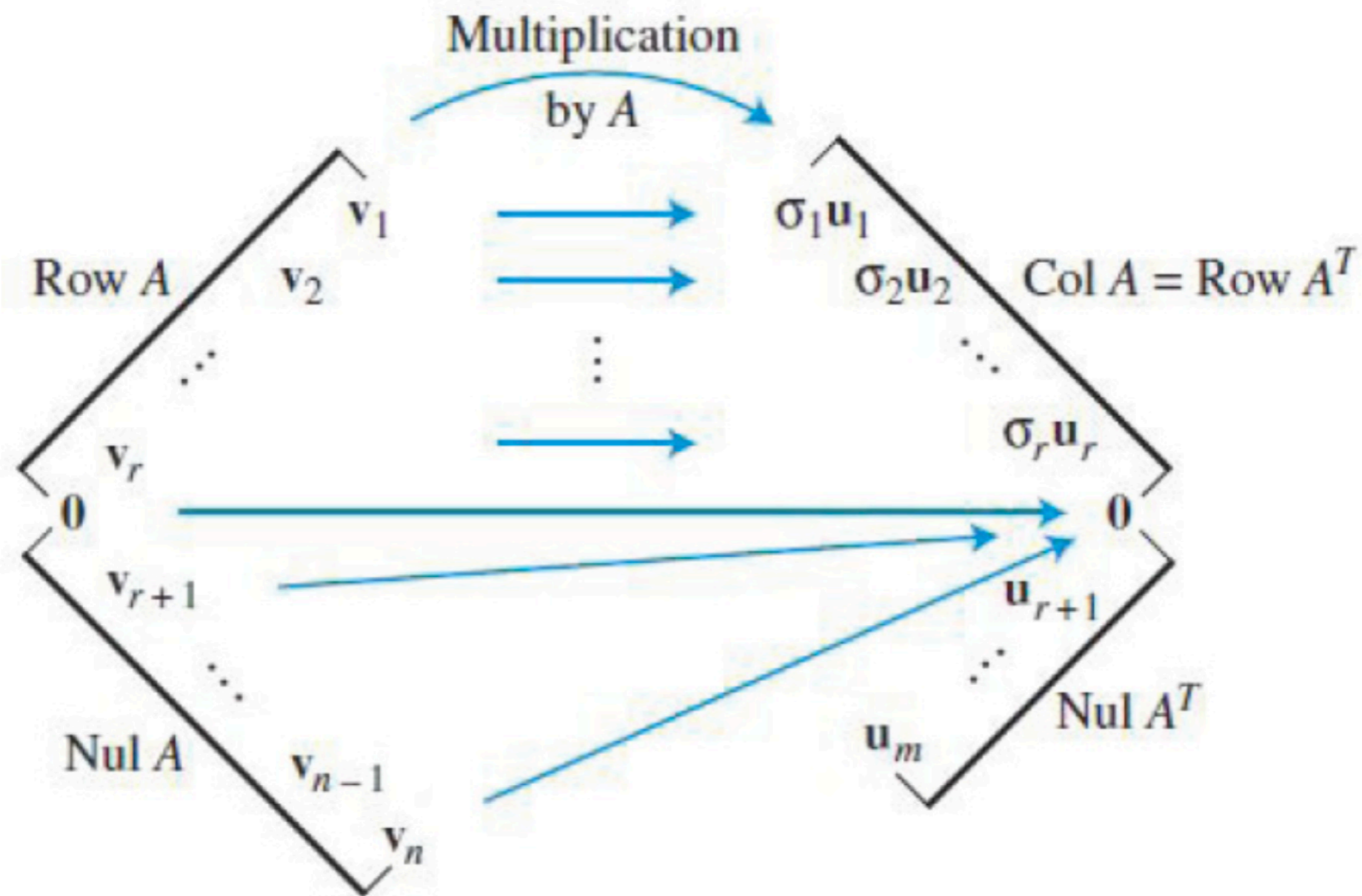
Then $\text{Row}(A) = \text{span}\{\mathbf{y}_1, \dots, \mathbf{y}_k\}$, $C(A) = \text{span}\{\mathbf{x}_1, \dots, \mathbf{x}_k\}$.

Lemma 25.2 (equivalent) Suppose $A = XY^\top$,

where $X \in \mathbb{R}^{m \times r}$, $Y \in \mathbb{R}^{n \times r}$ have full rank.

Then $C(A) = C(X)$, $\text{Row}(A) = \text{Row}(Y^\top)$.

Four Fundamental Subspaces



Rank = # of Nonzero Singular Values

rank(A) = # of nonzero **singular values** (counting multiplicity)

rank(A) _____ # of nonzero eigenvalues.

Example:

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

Eigenvalues are _____,

of nonzero eigenvalue is _____

rank(A) = _____.

Judgement

The number of singular values is the same as the rank of the matrix.

The number of singular values of an $m \times n$ matrix is m or n .

The rank of an $m \times n$ matrix is the number of nonzero eigenvalues.

Part IV Geometry

Facts About Eigenvalues

In geometrical interpretation of eigenvectors, we view the matrix as a linear transformation between _____ and _____.

Why?

Two geometrical views:

- 1st view: “Unchanged” during “change”; fixed point.
This motivation only applies to “**self-transformation**”.
- 2nd view: “Tiles” \rightarrow similar tiles with stretching.

Two-Space Transformation: Standard Basis

Consider $A \in \mathbb{R}^{3 \times 2}$.

A can be viewed as a linear transformation: $\mathbb{R}^2 \rightarrow \mathbb{R}^3$

Geometrical view:

Changing tiles: _____ to _____

Two-Space Transformation: Arbitrary Basis

Consider an arbitrary basis.

Changing tiles: _____ to _____


Two-Space Transformation: Arbitrary Basis

Wish: Find basis s.t. shape of “tiles” do NOT change (too much).

Changing tiles: _____ to _____

Mathematically:

Find orthonormal basis $\{\mathbf{v}_1, \mathbf{v}_2\}$ s.t. $\{A\mathbf{v}_1, A\mathbf{v}_2\}$ is _____



Summary Today (Write Your Own)

One sentence summary:

Detailed summary:

Summary Today (Write Your Own)

One sentence summary:

We learn SVD today.

Detailed summary:

1) SVD is a decomposition of any real matrix A (can be rectangular) s.t.

$$A = UDV^T,$$

where

1b) Compact SVD:

2) Relation of singular values/vectors and eigenvalues/eigenvectors:

3) Number of singular values is _____.

4) Steps of computing SVD: