

Notes on SVD

Contents

- ① Properties of SVD
- ② Reduction of Dimensions
- ③ Appendix: Unitary Diagonalization and SVD

Recap SVD.

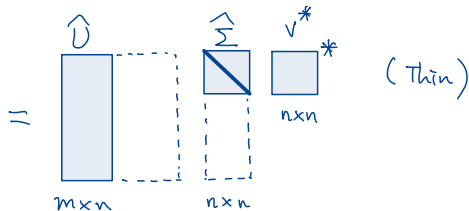
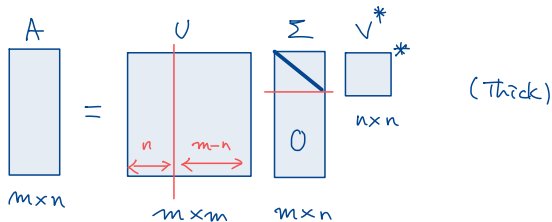
Let $A \in \mathbb{C}^{m \times n}$. Then

$$A = U \Sigma V^*$$

where

- $U \in \mathbb{C}^{m \times m}$ unitary
- $V \in \mathbb{C}^{n \times n}$ unitary
- $\Sigma \in \mathbb{R}^{m \times n}$ diagonal.

Illustration for $m > n$



Note:

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

When $m > n$.

$$\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n \geq 0$$

linearly dependent columns. \Rightarrow not orthogonal.

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

Not an SVD

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

An SVD

Properties of SVD

SVD and the 2-Norm

largest singular value because $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n \geq 0$.

Theorem 1

Let $A \in \mathbb{C}^{m \times n}$ have an SVD $A = U\Sigma V^*$. Then

- 1 $\|A\|_2 = \sigma_1$ and $\|A\|_F = \sqrt{\sigma_1^2 + \sigma_2^2 + \dots + \sigma_r^2}$.
- 2 The rank of A is the number of nonzero singular values.
- 3 Let $r = \min\{m, n\}$. Then

$$\kappa_2(A) = \|A\|_2 \|A^+\|_2 = \frac{\sigma_1}{\sigma_r}.$$

of linearly independent columns.

e.g.

$$\Sigma = \begin{bmatrix} * & & & \\ & * & & \\ & & * & \\ & \text{non zero} & & \dots & 0 \end{bmatrix} \Rightarrow \text{rank}(A) = 3$$

Connection to EVD

$$B \in \mathbb{C}^{n \times n}$$

Let $A = U\Sigma V^* \in \mathbb{C}^{m \times n}$ and $B = A^*A$. Observe that

- $B \in \mathbb{C}^{n \times n}$ is a *hermitian matrix*¹, i.e., $B^* = B$. (Pf: $B^* = (A^*A)^* = A^*(A^*)^* = A^*A = B$)
- B has an EVD:

$$B = \underbrace{(V\Sigma^*U^*)}_{A^*} \underbrace{(U\Sigma V^*)}_A = V\Sigma^*\Sigma V^* = V(\Sigma^*\Sigma)V^{-1}.$$

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

- The squares of singular values of A are eigenvalues of B .
- An EVD of $B = A^*A$ reveals the singular values and a set of right singular vectors of A .

Assume $m \geq n$

$$\Sigma^T \Sigma = \begin{bmatrix} \sigma_1 & & 0 \\ & \sigma_2 & \\ 0 & & \ddots \\ & & & \sigma_n \\ & & & & 0 \end{bmatrix} \begin{bmatrix} \sigma_1 & & 0 \\ & \sigma_2 & \\ & & \ddots \\ & & & \sigma_n \\ & & & & 0 \end{bmatrix} = \begin{bmatrix} \sigma_1^2 & & \\ & \sigma_2^2 & \\ & & \ddots \\ & & & \sigma_n^2 \end{bmatrix} \quad (\text{diagonal})$$

σ_j^2 eigenvalue of B
 σ_j : singular value of A .

¹This is the \mathbb{C} -extension of real symmetric matrices. ($A \in \mathbb{R}^{n \times n}$ is symmetric if $A^T = A$.)

Q. How would you find left-singular vectors of A ?

A. Form $C = AA^*$ and find its eigenvectors.

Note:

$$C = \underbrace{(U \Sigma V^*)}_A \underbrace{(V \Sigma^* U^*)}_{A^*}$$

$\Sigma \parallel I$ $\Sigma^* \parallel I$

$$= U \underbrace{\Sigma \Sigma^T}_{\text{diagonal matrix}} U^{-1}$$

$$\begin{bmatrix} \sigma_1^2 & & & \\ & \sigma_2^2 & & \\ & & \ddots & \\ & & & \sigma_n^2 & & \\ & & & & 0 & \\ & & & & & \ddots & \\ & & & & & & 0 \end{bmatrix}$$

$m \times m$

Theorem 2

*The nonzero singular values of $A \in \mathbb{C}^{m \times n}$ are the square roots of the nonzero eigenvalues of A^*A or AA^* .*

Reduction of Dimensions

Low-Rank Approximations

Let $A \in \mathbb{C}^{m \times n}$ with $m \geq n$. Its thin SVD $A = \hat{U} \hat{\Sigma} V^*$ can be written as

$$A = \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_n \end{bmatrix} \begin{bmatrix} \sigma_1 & & & \\ & \ddots & & \\ & & \sigma_n & \end{bmatrix} \begin{bmatrix} \mathbf{v}_1^* \\ \vdots \\ \mathbf{v}_n^* \end{bmatrix}$$

$$= \begin{bmatrix} \sigma_1 \mathbf{u}_1 & \cdots & \sigma_n \mathbf{u}_n \end{bmatrix} \begin{bmatrix} \mathbf{v}_1^* \\ \vdots \\ \mathbf{v}_n^* \end{bmatrix} = \sum_{j=1}^r \sigma_j \underbrace{\mathbf{u}_j \mathbf{v}_j^*}_{\text{outer product}}$$

$m \times n$ $n \times n$ $\in \mathbb{C}^{m \times n}$

If $\text{rank}(A) = r$,
then

$$\sigma_{r+1} = \cdots = \sigma_n = 0.$$

where r is the rank of A .

(outer product expansion of A)

- Each outer product $\mathbf{u}_j \mathbf{v}_j^*$ is a rank-1 matrix.
- Since $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$, important contributions to A come from terms with small j .

$$\vec{u} \vec{v}^* = \vec{u} [\vec{v}_1 \ \vec{v}_2 \ \dots \ \vec{v}_n]$$

$$= \underbrace{[\vec{v}_1(\vec{u}) \ \vec{v}_2 \vec{u} \ \dots \ \vec{v}_n \vec{u}]}_{\text{linearly dependent on } \vec{u}}$$

linearly dependent on \vec{u}

$$= \text{rank } 1,,$$

Low-Rank Approximations (cont')

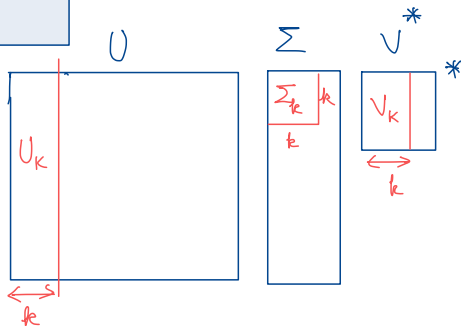
For $1 \leq k \leq \overset{\text{rank}(A)}{r}$, define

$$A_k = \sum_{j=1}^k \sigma_j \mathbf{u}_j \mathbf{v}_j^* = U_k \Sigma_k V_k^*,$$

where

- U_k is the first k columns of U ;
- V_k is the first k columns of V ;
- Σ_k is the upper-left $k \times k$ submatrix of Σ .

This is a rank- k approximation of A .



Best Rank- k Approximation

Theorem 3 (Eckart-Young)

Let $A \in \mathbb{C}^{m \times n}$. Suppose A has rank r and let $A = U\Sigma V^*$ be an SVD. Then

- $\|A - A_k\|_2 = \sigma_{k+1}$, for $k = 1, \dots, r-1$.
- For any matrix B with $\text{rank}(B) \leq k$, $\|A - B\|_2 \geq \sigma_{k+1} = \|A - A_k\|_2$

the first omitted singular value
in construction of A_k .

Appendix: Unitary Diagonalization and SVD

Unitary Diagonalization of Hermitian Matrices

The previous discussion is relevant to hermitian matrices constructed in a specific manner. For a generic hermitian matrix, we have the following result.

Theorem 4 (Spectral Decomposition)

Let $A \in \mathbb{C}^{n \times n}$ be hermitian. Then A has a unitary diagonalization

$$A = VDV^{-1},$$

where $V \in \mathbb{C}^{n \times n}$ is unitary and $D \in \mathbb{R}^{n \times n}$ is diagonal.

In words, a hermitian matrix (or symmetric matrix) has a complete set of orthonormal eigenvectors and all its eigenvalues are real.

Notes on Unitary Diagonalization and Normal Matrices

- A unitarily diagonalizable matrix $A = VDV^{-1}$ with $D \in \mathbb{C}^{n \times n}$, is called a **normal matrix**². All hermitian matrices are normal.
- Let $A = VDV^{-1} \in \mathbb{C}^{n \times n}$ be normal. Since $\kappa_2(V) = 1$ (why?), Bauer-Fike implies that eigenvalues of A can be changed by no more than $\|\delta A\|_2$.

²Usual definition: $A \in \mathbb{C}^{n \times n}$ is normal if $AA^* = A^*A$.

Unitary Diagonalization and SVD

Theorem 5

Let $A \in \mathbb{C}^{n \times n}$ be hermitian. Then the singular values of A are the absolute values of the eigenvalues of A .

Precisely, if $A = VDV^{-1}$ is a unitary diagonalization of A , then

$$A = (V \operatorname{sign}(D)) |D| V^*$$

is an SVD, where

$$\operatorname{sign}(D) = \begin{bmatrix} \operatorname{sign}(d_1) & & \\ & \ddots & \\ & & \operatorname{sign}(d_n) \end{bmatrix}, \quad |D| = \begin{bmatrix} |d_1| & & \\ & \ddots & \\ & & |d_n| \end{bmatrix}.$$

When Do Unitary EVD and SVD Coincide?

Theorem 6

If $A = A^*$, then the following statements are equivalent:

- 1 Any unitary EVD of A is also an SVD of A .
- 2 The eigenvalues of A are positive numbers.
- 3 $\mathbf{x}^* A \mathbf{x} > 0$ for all nonzero $\mathbf{x} \in \mathbb{C}^n$. (HPD)

- The equivalence of 1 and 2 is immediate from Theorem 5
- The property in 3 is called the **hermitian positive definiteness**, c.f., symmetric positive definiteness.
- The equivalence of 2 and 3 can be shown conveniently using **Rayleigh quotient**; see next slide.

Note: Rayleigh Quotient

Let $A \in \mathbb{R}^{n \times n}$ be fixed. The **Rayleigh quotient** is the map $R_A : \mathbb{R}^n \rightarrow \mathbb{R}$ given by

$$R_A(\mathbf{x}) = \frac{\mathbf{x}^T A \mathbf{x}}{\mathbf{x}^T \mathbf{x}}.$$

- R_A maps an eigenvector of A into its associated eigenvalue, *i.e.*, if $A\mathbf{v} = \lambda\mathbf{v}$, then $R_A(\mathbf{v}) = \lambda$.
- If $A = A^T$, then $\nabla R_A(\mathbf{v}) = \mathbf{0}$ for an eigenvector \mathbf{v} , and so

$$R_A(\mathbf{v} + \epsilon \mathbf{z}) = R_A(\mathbf{v}) + 0 + O(\epsilon^2) = \lambda + O(\epsilon^2), \quad \text{as } \epsilon \rightarrow 0.$$

The Rayleigh quotient is a quadratic approximation of an eigenvalue.