Singular Value Decomposition

Contents

Singular Value Decomposition:

Overview

Singular Value Decomposition

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

Theorem 1 (SVD)

Let $A \in \mathbb{C}^{m \times n}$. Then A can be written as

$(m \geqslant n)$

$\mathbb{C}$ analog of T

$$A = U\Sigma V^{\boxed{*}}, \qquad \text{(SVD)}$$

where $U \in \mathbb{C}^{m \times m}$ and $V \in \mathbb{C}^{n \times n}$ are unitary and $\Sigma \in \mathbb{R}^{m \times n}$ is diagonal. If A is real, then so are U and V.

$\mathbb{C}$ analog of orthogonal

- The columns of U are called the **left singular vectors** of A;
- The columns of V are called the **right singular vectors** of A;
- The diagonal entries of Σ, written as $\sigma_1, \sigma_2, \ldots, \sigma_r$, for $r = \min\{m, n\}$, are called the **singular values** of A and they are nonnegative numbers ordered as

$$\sigma_1 \geqslant \sigma_2 \geqslant \cdots \geqslant \sigma_r \geqslant 0.$$

Singular Value Decomposition (cont')

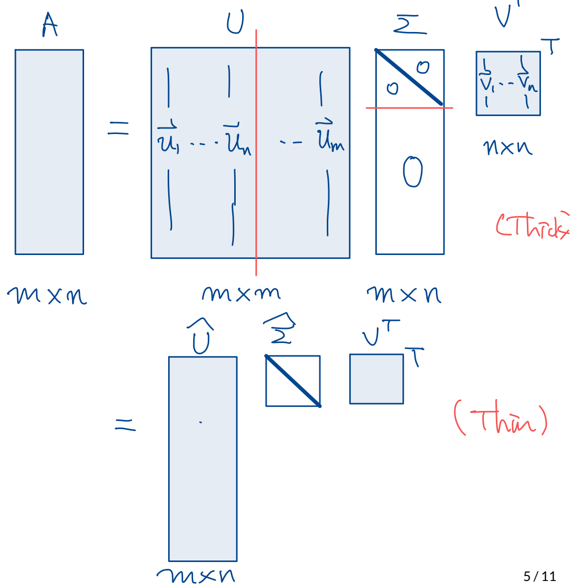
SVD for real matrices

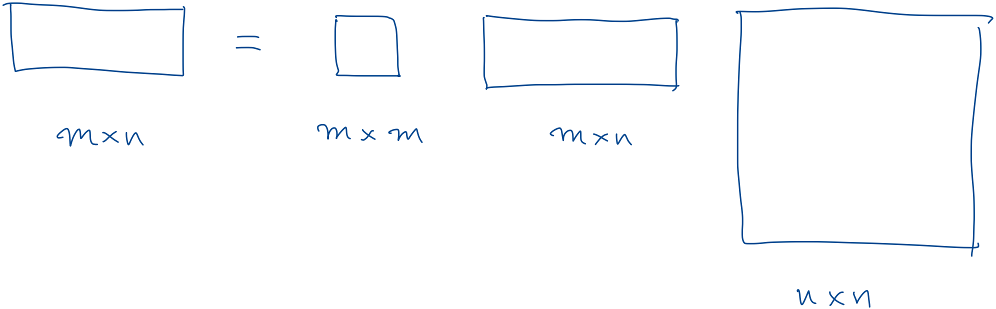
Let $A \in \mathbb{R}^{m\times n}$. Then

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

where $U \in \mathbb{R}^{m\times m}$ and $V \in \mathbb{R}^{n\times n}$ are orthogonal, and $\Sigma \in \mathbb{R}^{m\times n}$ is diagonal.

Cartoon view of SVD $(m \geq n)$





=
m×n
m×m
m×n
n×n

Thick vs Thin SVD

Suppose that $m > n$ and observe that:

$$U\Sigma = \left[\begin{array}{ccc|ccc} \mathbf{u}_1 & \cdots & \mathbf{u}_{n-1} & \mathbf{u}_n & \cdots & \mathbf{u}_m \end{array}\right] \left[\begin{array}{ccc} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \\ \hline & \mathbf{0} & \end{array}\right]$$

$$= \begin{bmatrix} \mathbf{u}_1 & \cdots & \mathbf{u}_{n-1} \end{bmatrix} \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \end{bmatrix} = \widehat{U}\widehat{\Sigma}.$$

SVD in MATLAB

- Thick SVD: `[U,S,V] = svd(A);` $\nearrow \Sigma$
- Thin SVD: `[U,S,V] = svd(A, 0);`

- If only want singular values:

  `>> S = svd(A);     % S is a col. vec.`

- If only want, say, V:

  `>> [~,~,V]=svd(A);`

To confirm $A = U\Sigma V^*$:

`>> norm(A - U*S*V');`

Understanding SVD

Geometric Perspective

Write $A = U\Sigma V^*$ as $AV = U\Sigma$:

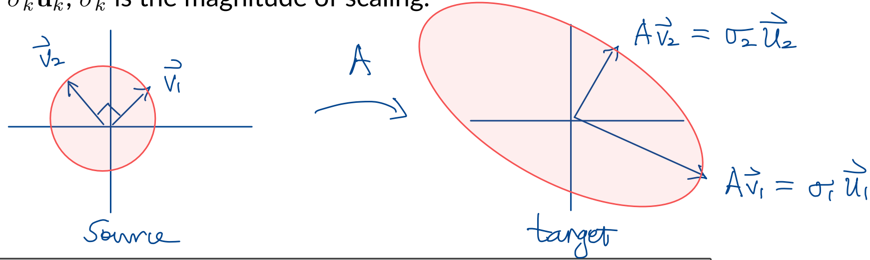
$$A\mathbf{v}_k = \sigma_k \mathbf{u}_k, \quad k = 1, \ldots, r = \min\{m, n\}.$$

$\sigma_1 \geq \sigma_2$

- Each right singular vector $\mathbf{v}_k$ is mapped by A to a scaled left singular vector $\sigma_k \mathbf{u}_k$; σ_k is the magnitude of scaling.

$\mathbb{R}^2$

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



> The image of the unit sphere under any $m \times n$ matrix is a hyperellipse.

Algebraic Perspective

$A \in \mathbb{C}^{m \times n}$, $A = U\Sigma V^*$

Alternately, note that $\mathbf{y} = A\mathbf{z} \in \mathbb{C}^m$ for any $\mathbf{z} \in \mathbb{C}^n$ can be written as

$\vec{y} = A\vec{z}$

$\vec{y} = U\Sigma V^* \vec{z}$

$U^* \vec{y} = \boxed{U^* U} \Sigma V^* \vec{z}$

$\vec{y}$

I because U is unitary

$$\underbrace{(\overset{=U^{-1}}{U^*}\mathbf{y})}_{\text{Coord. of } \vec{y} \text{ w.r.t. } U\text{-basis}} = \Sigma \underbrace{(\overset{=V^{-1}}{V^*}\mathbf{z})}_{\text{Coord. of } \vec{z} \text{ w.r.t. } V\text{-basis.}}.$$

- Since U and V are unitary, $U^* = U^{-1}$ and $V^* = V^{-1}$.
- $U^*\mathbf{y}$ is the coordinates of $\mathbf{y} \in \mathbb{C}^m$ with respect to the basis consisting of columns of U, which is an ONB.
- $V^*\mathbf{z}$ is the coordinates of $\mathbf{z} \in \mathbb{C}^n$ with respect to the basis consisting of columns of V, which is an ONB.

Any matrix $A \in \mathbb{C}^{m \times n}$ can be viewed as a diagonal transformation from $\mathbb{C}^n$ (source space) to $\mathbb{C}^m$ (target space) with respect to suitably chosen orthonormal bases for both spaces.

SVD vs. EVD

Recall that a diagonalizable $A = VDV^{-1} \in \mathbb{C}^{n\times n}$ satisfies

$$\mathbf{y} = A\mathbf{z} \quad \longrightarrow \quad \left(V^{-1}\mathbf{y}\right) = D\left(V^{-1}\mathbf{z}\right).$$

This allowed us to view any diagonalizable square matrix $A \in \mathbb{C}^{n\times n}$ as a diagonal transformation from $\mathbb{C}^n$ to itself[1] with respect to the basis formed by a set of eigenvector of A.

Differences.

- **Basis:** SVD uses two ONBs (left and right singular vectors); EVD uses one, usually non-orthogonal basis (eigenvectors).
- **Universality:** all matrices have an SVD; not all matrices have an EVD.
- **Utility:** SVD is useful in problems involving the behavior of A or A^+; EVD is relevant to problems involving A^k.