Singular Value Decomposition

MATH 271.1: Statistical Methods

Jakov Ivan S. Dumbrique

Ateneo de Manila University

Singular Value Decomposition

Consider a matrix $\mathbf{A} \in \mathbb{R}^{n \times p}$. Singular Value Decomposition (SVD) is one of the most well-known and effective matrix decomposition methods. It has two different forms, i.e. complete and incomplete.

Singular Value Decomposition

Consider a matrix $\mathbf{A} \in \mathbb{R}^{n\times p}$. Singular Value Decomposition (SVD) is one of the most well-known and effective matrix decomposition methods. It has two different forms, i.e. complete and incomplete.
The "complete SVD" decomposes the matrix as:

$$\underset{n\times p}{\mathbf{A}} = \underset{n\times n}{\mathbf{U}}\,\underset{n\times p}{\mathbf{\Sigma}}\,\underset{p\times p}{\mathbf{V}^{\mathbf{T}}},$$

where

- $\mathbf{U} \in \mathbb{R}^{n\times n}$ is orthonormal
  ① orthogonal
  ② the columns of U are unit vectors $\|\vec{u}_i\| = 1$
- $\mathbf{\Sigma} \in \mathbb{R}^{n\times p}$ is diagonal
- $\mathbf{V} \in \mathbb{R}^{p\times p}$ is orthonormal

Singular Value Decomposition

Consider a matrix $\mathbf{A} \in \mathbb{R}^{n \times p}$. Singular Value Decomposition (SVD) is one of the most well-known and effective matrix decomposition methods. It has two different forms, i.e. complete and incomplete.

The "complete SVD" decomposes the matrix as:

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

where

$\mathbf{U} \in \mathbb{R}^{n \times n}$ is orthonormal

$\mathbf{\Sigma} \in \mathbb{R}^{n \times p}$ is diagonal

$\mathbf{V} \in \mathbb{R}^{p \times p}$ is orthonormal

The columns of $\mathbf{U}$ are called the *left singular vectors* of $\mathbf{A}$.

Singular Value Decomposition

Consider a matrix $\mathbf{A} \in \mathbb{R}^{n \times p}$. Singular Value Decomposition (SVD) is one of the most well-known and effective matrix decomposition methods. It has two different forms, i.e. complete and incomplete.
The "complete SVD" decomposes the matrix as:

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

where

- $\mathbf{U} \in \mathbb{R}^{n \times n}$ is orthonormal
- $\mathbf{\Sigma} \in \mathbb{R}^{n \times p}$ is diagonal
- $\mathbf{V} \in \mathbb{R}^{p \times p}$ is orthonormal

The columns of $\mathbf{U}$ are called the *left singular vectors* of $\mathbf{A}$.
The columns of $\mathbf{V}$ are called the *right singular vectors* of $\mathbf{A}$.

Complete SVD

In complete SVD, the Σ is a rectangular diagonal matrix. In cases $n > p$ and $n < p$, this matrix is in the forms:

$n>p$ $\qquad\qquad\qquad\qquad\qquad$ $n<p$

$$\Sigma = \begin{bmatrix} \sigma_1 & 0 & \dots & 0 \\ 0 & \sigma_2 & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_p \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 0 \end{bmatrix}, \text{ and } \begin{bmatrix} \sigma_1 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & \sigma_2 & \dots & 0 & 0 & \dots & 0 \\ \vdots & & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \sigma_n & 0 & \dots & 0 \end{bmatrix},$$

n

$n=p$

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

respectively.

Complete SVD

$$\mathbf{\Sigma} = \begin{bmatrix} \sigma_1 & 0 & \dots & 0 \\ 0 & \sigma_2 & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_p \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 0 \end{bmatrix}$$

The diagonal entries of $\mathbf{\Sigma}$, denoted by σ_i are nonnegative and are called the *singular values* of $\mathbf{A}$.

Complete SVD

$$\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_1 & 0 & \dots & 0 \\ 0 & \sigma_2 & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_p \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 0 \end{bmatrix}$$

$\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_p \geq 0$

The diagonal entries of $\boldsymbol{\Sigma}$, denoted by σ_i are nonnegative and are called the *singular values* of $\mathbf{A}$. They are ordered such that the largest singular value, σ_1 is placed in the $(1, 1)$ entry of $\boldsymbol{\Sigma}$, and the other singular values are placed down the diagonal, and satisfy $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_l \geq 0$ such that $l = \min(n, p)$.

Complete SVD

An example of a complete SVD:

incomplete SVD

$$\Sigma = \begin{bmatrix} \sqrt{2} & & & \\ & \sqrt{2} & & \\ & & 1 & \\ & & & \ddots \\ 0 & 0 & \cdots & 0 \quad 0 \end{bmatrix}$$

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

$$\underset{3\times 2}{\begin{bmatrix} 1 & -1 \\ 0 & 1 \\ -1 & 0 \end{bmatrix}} = \underset{3\times 3}{\begin{bmatrix} -\frac{2}{\sqrt{6}} & 0 & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \end{bmatrix}} \underset{3\times 2}{\overset{\Sigma}{\begin{bmatrix} \sqrt{3} & 0 \\ 0 & 1 \\ \cancel{0} & \cancel{0} \end{bmatrix}}} \underset{2\times 2}{\begin{bmatrix} -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}^T}$$

rank of A = no. of nonzero singular values of A
(multiplicities are accounted for)
= 2

Incomplete SVD

The "incomplete SVD" decomposes the matrix as:

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

left singular

right singular

where

- $\mathbf{U} \in \mathbb{R}^{n\times l}$ is orthonormal

square matrix

- $\mathbf{\Sigma} \in \mathbb{R}^{l\times l}$ is diagonal

$\mathbf{V} \in \mathbb{R}^{p\times l}$ is orthonormal,

such that $l := \min(n, p)$. The columns of $\mathbf{U}$ and the columns of $\mathbf{V}$ are still called the *left singular vectors* and the *right singular vectors* of $\mathbf{A}$, respectively.

Incomplete SVD

In incomplete SVD, the Σ is a square diagonal matrix whose main diagonal includes the *singular values* of $\mathbf{A}$. The matrix Σ is in the form:

$$\mathbf{\Sigma} = \begin{bmatrix} \sigma_1 & 0 & \dots & 0 \\ 0 & \sigma_2 & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_l \end{bmatrix}.$$

≥ 0

Incomplete SVD

An example of an incomplete SVD:

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

$$\underbrace{\begin{bmatrix} 1 & -1 \\ 0 & 1 \\ -1 & 0 \end{bmatrix}}_{3\times 2} = \underbrace{\begin{bmatrix} -\frac{2}{\sqrt{6}} & 0 \\ \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} \end{bmatrix}}_{3\times 2} \underbrace{\begin{bmatrix} \sqrt{3} & 0 \\ 0 & 1 \end{bmatrix}}_{2\times 2} \underbrace{\begin{bmatrix} -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}^T}_{2\times 2}$$

Singular vectors in SVD

Note that in both complete and incomplete SVD, the left singular vectors are orthonormal and the right singular vectors are also orthonormal;

Singular vectors in SVD

Note that in both complete and incomplete SVD, the left singular vectors are orthonormal and the right singular vectors are also orthonormal; therefore, $\mathbf{U}$ and $\mathbf{V}$ are both orthogonal matrices so:

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

Singular vectors in SVD

Note that in both complete and incomplete SVD, the left singular vectors are orthonormal and the right singular vectors are also orthonormal; therefore, $\mathbf{U}$ and $\mathbf{V}$ are both orthogonal matrices so:

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

If these orthogonal matrices are not truncated and thus are square matrices, e.g., for complete SVD, we also have:

complete SVD

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

Singular Values and Rank of A

If $\mathbf{A}$ has r non-zero singular values, counted according to their multiplicity (i.e. if $\mathbf{\Sigma}$ has r non-zero diagonal entries), then $\mathbf{A}$ has rank r.

Singular Values and Rank of A

If $\mathbf{A}$ has r non-zero singular values, counted according to their multiplicity (i.e. if $\mathbf{\Sigma}$ has r non-zero diagonal entries), then $\mathbf{A}$ has rank r. That is, the rank of $\mathbf{A}$ is equal to the number of non-zero singular values, when their multiplicities are accounted for.

Singular Values and Rank of $\mathbf{A}$

If $\mathbf{A}$ has r non-zero singular values, counted according to their multiplicity (i.e. if $\mathbf{\Sigma}$ has r non-zero diagonal entries), then $\mathbf{A}$ has rank r. That is, the rank of $\mathbf{A}$ is equal to the number of non-zero singular values, when their multiplicities are accounted for. Also, for $j > r$, the singular value σ_j has a value of 0.

Theorem 1

In both complete and incomplete SVD, ① the singular values of an $n \times p$ matrix $\mathbf{A}$ are the square roots of the eigenvalues of $\mathbf{A}^{\mathbf{T}}\mathbf{A}$.

② The right singular vectors are the eigenvectors of $\mathbf{A}^{\mathbf{T}}\mathbf{A}$.

Proof:

Suppose that $n \times p$ matrix A has the fll. complete SVD:

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

WLOG $n > p$: $\underset{n\times p}{\Sigma} = \begin{bmatrix} \sigma_1 & & 0 \\ & \sigma_2 & \\ 0 & & \ddots \ \sigma_p \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

$$A^TA = \underbrace{(U\Sigma V^T)^T}(U\Sigma V^T)$$

$$= V\Sigma^T \underbrace{U^T U}\Sigma V^T$$

$$\underset{\text{symmetrical matrix}}{A^TA} = \underset{\text{orthogonal}}{V}\,\underbrace{\Sigma^T\Sigma}_{\text{diagonal matrix}}\,V^T$$

orthogonal

$\underbrace{\hspace{1em}}_{\text{orthogonal}}$ diagonalization of A^TA

$$\underset{p\times p}{\Sigma^T\Sigma} = \begin{bmatrix} \sigma_1 & & 0 & 0 & 0 \\ & \sigma_2 & & 0 \cdots & 0 \\ 0 & & \ddots \ \sigma_p & 0 & 0 \end{bmatrix}\begin{bmatrix} \sigma_1 & & 0 \\ & \sigma_2 & \\ 0 & & \ddots \ \sigma_p \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$= \underset{\text{diagonal matrix}}{\begin{bmatrix} \sigma_1^2 & & 0 \\ & \sigma_2^2 & \\ 0 & & \ddots \ \sigma_p^2 \end{bmatrix}}$$

① the eigenvalues of A^TA are the diagonal elements of $\Sigma^T\Sigma$, which are the squares of the singular values of A.

② the eigenvectors of A^TA are the columns of V, which are the right singular vectors of A.

Eigenvalues & Eigenvectors of $\mathbf{A}^\mathbf{T}\mathbf{A}$

Given the complete SVD of $\mathbf{A}$: $\longrightarrow$ eigenvalue decomposition of A^TA

$A^TA = V \cdot (\Sigma^T\Sigma)\, V^T$ ✓

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

$$\begin{bmatrix} 1 & -1 \\ 0 & 1 \\ -1 & 0 \end{bmatrix} = \begin{bmatrix} -\frac{2}{\sqrt{6}} & 0 & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \end{bmatrix} \begin{bmatrix} \sqrt{3} & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}^T$$

What are the eigenvalues and eigenvectors of $\mathbf{A}^\mathbf{T}\mathbf{A}$?

eigenvalues

$\lambda_1 = \sigma_1^2 = 3 \longrightarrow \vec{v}_1 = \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$

$\lambda_2 = \sigma_2^2 = 1 \longrightarrow \vec{v}_2 = \begin{bmatrix} -1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix}$

✓ $A^TA\,\vec{v}_i = \lambda_i \vec{v}_i$

Theorem 2

Given an $n \times p$ matrix $\mathbf{A}$ of rank r, then the left and right singular vectors in its complete singular value decomposition

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

form orthonormal bases for the four fundamental subspaces.

Theorem 2

Given an $n \times p$ matrix $\mathbf{A}$ of rank r, then the left and right singular vectors in its complete singular value decomposition

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

form orthonormal bases for the four fundamental subspaces. In particular,

- $\mathbf{u_1}, \ldots, \mathbf{u_r}$ is an orthonormal basis for the **column space** $Col(\mathbf{A})$
- $\mathbf{u_{r+1}}, \ldots, \mathbf{u_n}$ is an orthonormal basis for the **left nullspace** $N(\mathbf{A^T})$
- $\mathbf{v_1}, \ldots, \mathbf{v_r}$ is an orthonormal basis for the **row space** $Col(\mathbf{A^T})$
- $\mathbf{v_{r+1}}, \ldots, \mathbf{v_p}$ is an orthonormal basis for the **nullspace** $N(\mathbf{A})$

This is the third part of the Fundamental Theorem of LinAlg!

Finding bases for the subspaces of $\mathbf{A}$

Given the complete SVD of $\mathbf{A}$: $r=2$

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

$$\begin{bmatrix} 1 & -1 \\ 0 & 1 \\ -1 & 0 \end{bmatrix} = \begin{bmatrix} -\frac{2}{\sqrt{6}} & 0 & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \end{bmatrix} \begin{bmatrix} \sqrt{3} & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}^T$$

What is an orthonormal basis for each fundamental subspace of $\mathbf{A}$?

① $Col(A)$
$$\left\{ \begin{bmatrix} -2/\sqrt{6} \\ 1/\sqrt{6} \\ 1/\sqrt{6} \end{bmatrix}, \begin{bmatrix} 0 \\ -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} \right\}$$

② $N(A^T)$
$$\left\{ \begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix} \right\}$$

③ $Col(A^T)$
$$\left\{ \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}, \begin{bmatrix} -1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix} \right\}$$

④ $N(A)$.
$$\{ \}$$
i.e. A has zero nullspace.

Proof: Theorem 2 Part 1

$\mathbf{u_1}, \ldots, \mathbf{u_r}$ is an orthonormal basis for the **column space** $Col(\mathbf{A})$

→ set of linearly ind. vectors that span the $Col(A)$

$\{$ all $\vec{b}: n\times 1$ $A\vec{x} = \vec{b}\}$

for $1 \le i \le r$,

$$A\vec{v_i} = (U\Sigma \underbrace{V^T)\vec{v_i}}$$

$$= U\underbrace{\Sigma \vec{e_i}}_{} \rightarrow p\times 1$$

$$= U\sigma_i \vec{e_i} \rightarrow n\times 1$$

$$= \sigma_i \underbrace{U\vec{e_i}}$$

$$\underbrace{A\vec{v_i}}_{A\vec{x}} = \underbrace{\sigma_i \vec{u_i}}_{n\times 1,\ \vec{b}}$$

$$V = \begin{bmatrix} \vec{v_1} \ldots \vec{v_2} & \cdots & \vec{v_p} \end{bmatrix}$$

$$V^T\vec{v_i} = \begin{bmatrix} \leftarrow \vec{v_1}^T \rightarrow \\ \leftarrow \vec{v_2}^T \rightarrow \\ \vdots \\ \leftarrow \vec{v_p}^T \rightarrow \end{bmatrix} \vec{v_i}$$

$$= \begin{bmatrix} \vec{v_1}^T\vec{v_i} \\ \vec{v_2}^T\vec{v_i} \\ \vdots \\ \vec{v_i}^T\vec{v_i} \\ \vdots \\ \vec{v_p}^T\vec{v_i} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ 0 \end{bmatrix} = \vec{e_i}$$

canonical unit vector

The vectors $\sigma_i\vec{u_i}$ for $1 \le i \le r$ are in the $Col(A)$.

The vectors $\vec{u_i}$ for $1 \le i \le r$ are in the $Col(A)$.

[① These r vectors are in $Col(A)$.
② Because the $\vec{u_i}$'s are orthogonal, thus they are lin. ind.]

$\{\vec{u_i}\}$ $1 \le i \le r$ form a basis of $Col(A)$.

Since $\vec{u_i}$ are the left singular vectors of A, they are orthonormal.

$n > p$

$$\underset{\substack{n\times p\\ n\times 1}}{\Sigma}\ \underset{p\times 1}{\vec{e}_i} = \begin{bmatrix} \sigma_1 & & & 0 \\ 0 & \sigma_2 & & \\ & & \sigma_i & \\ & & & \sigma_p \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ \sigma_i \\ \vdots \\ 0 \end{bmatrix} = \sigma_i \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix}$$

$$= \sigma_i \underset{n\times 1}{\vec{e}_i}.$$

$\to \{\vec{u}_1, \vec{u}_2, \dots, \vec{u}_r\}$ form an ORTHONORMAL basis of $\mathrm{Col}(A)$.

$$\underset{n\times n}{U}\ \underset{n\times 1}{\vec{e}_i} = [\vec{u}_1\ \ u_2 \dots \vec{u}_i \dots \vec{u}_n] \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix}$$

$$= \vec{u}_i$$