Singular Value Decomposition

Jiman Hwang

This document is for self-study only.
Source: [Lay11a]

1 Statement

Let A denote a $m \times n$ matrix with rank r. Then there exist decompositions such that

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

where U is the $m \times m$ orthogonal matrix, V is the $n \times n$ orthogonal matrix, and Σ is $m \times n$ matrix such that

$$\Sigma = \begin{bmatrix} D & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix},\ D = \operatorname{diag}(\sigma_1, \cdots, \sigma_r),\ \sigma_1 \geq \cdots \geq \sigma_r > 0$$

where $\sigma_1, \cdots, \sigma_r$ are singular values of A, i.e., the square roots of the eigenvalues of $A^T A$. Note that A may have two or more decomposition of this form.

2 Derivation

2.1 Resolving the rank of A

Let A denote a $m \times n$ matrix. By Spectral Theorem[Lay11b],

$$\exists V\ A^T A = VDV^T,\ V^{-1} = V^T,\ D = \operatorname{diag}(\lambda_1, \cdots, \lambda_n)$$

If $V = \begin{bmatrix} \mathbf{v}_1 & \cdots & \mathbf{v}_n \end{bmatrix}$,

$$\|A\mathbf{v}_i\| = \sqrt{\mathbf{v}_i^T A^T A \mathbf{v}_i} = \sqrt{\lambda_i \mathbf{v}_i^T \mathbf{v}_i} = \sqrt{\lambda_i \|\mathbf{v}_i\|^2} = \sqrt{\lambda_i} = \sigma_i > 0 \tag{1}$$

Renumber $\lambda_1, \cdots, \lambda_n$ so that

$$\sigma_1 \geq \cdots \geq \sigma_r > 0,\ \sigma_{r+1} = \cdots = 0 \tag{2}$$

where $1 \leq r \leq n$.

Meanwhile, $\{A\mathbf{v}_1, \cdots, A\mathbf{v}_r\}$ is an orthogonal set.

$$\because (A\mathbf{v}_i) \bullet (A\mathbf{v}_j) = \mathbf{v}_i^T A^T A \mathbf{v}_j = \lambda_j \mathbf{v}_i^T \mathbf{v}_j = \begin{cases} \lambda_j & i = j \\ 0 & i \neq j \end{cases}$$

Also, $\operatorname{span}\{\mathbf{v}_1, \cdots, \mathbf{v}_n\} = \mathbb{R}^n$. Thus,

$$\text{col } A = \{A\mathbf{x} : \mathbf{x} \in \mathbb{R}^n\}$$

where

$$\mathbf{x} = c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n,\ \{c_1, \cdots, c_n\} \subset \mathbb{R}$$

But

$$\begin{aligned} A\mathbf{x} =& c_1 A\mathbf{v}_1 + \cdots + c_r A\mathbf{v}_r + c_{r+1} A\mathbf{v}_{r+1} + \cdots + c_n A\mathbf{v}_n \\ =& c_1 A\mathbf{v}_1 + \cdots + c_r A\mathbf{v}_r + 0 + \cdots + 0 && \text{from (1), (2)} \end{aligned}$$

Hence

$$\text{col } A = \operatorname{span}\{A\mathbf{v}_1, \cdots, A\mathbf{v}_r\}$$

Also, rank $A = r$

2.2 Decomposition

$$
\begin{aligned}
AV =& A\begin{bmatrix}\mathbf{v}_1 & \cdots & \mathbf{v}_n\end{bmatrix} \\
=& \begin{bmatrix}A\mathbf{v}_1 & \cdots & A\mathbf{v}_r & A\mathbf{v}_{r+1} & \cdots & A\mathbf{v}_n\end{bmatrix} \\
=& \begin{bmatrix}\sigma_1\frac{A\mathbf{v}_1}{\sigma_1} & \cdots & \sigma_r\frac{A\mathbf{v}_r}{\sigma_r} & 0 & \cdots & 0\end{bmatrix} \\
=& \begin{bmatrix}\frac{A\mathbf{v}_1}{\sigma_1} & \cdots & \frac{A\mathbf{v}_r}{\sigma_r} & 0 & \cdots & 0\end{bmatrix}\begin{bmatrix}\sigma_1 & & & \\ & \ddots & & \\ & & \sigma_r & \\ & & & \end{bmatrix} \quad \text{dim: } (m\times n)\times(n\times n) \\
=& \begin{bmatrix}\mathbf{u}_1 & \cdots & \mathbf{u}_r & 0 & \cdots & 0\end{bmatrix}\begin{bmatrix}\sigma_1 & & & \\ & \ddots & & \\ & & \sigma_r & \\ & & & \end{bmatrix} \quad \text{dim: } (m\times n)\times(n\times n)
\end{aligned}
$$

where the empty entries in the second matrix at the last line are filled with 0's.

Noticing $\mathbf{u}_{r+1}, \cdots, \mathbf{u}_m$ is already orthonormal set, use Gram-Schmidt to find $\mathbf{u}_{r+1}, \cdots, \mathbf{u}_m$ in a way that $\{\mathbf{u}_1, \cdots, \mathbf{u}_m\}$ becomes an orthonormal set. Then, replace the trailing 0's with $\mathbf{u}_{r+1}, \cdots, \mathbf{u}_m$. This is legitimate because the corresponding entries in the second matrix are all 0's.

$$
AV = \begin{bmatrix}\mathbf{u}_1 & \cdots & \mathbf{u}_r & \mathbf{u}_{r+1} & \cdots & \mathbf{u}_m\end{bmatrix}\begin{bmatrix}\sigma_1 & & & \\ & \ddots & & \\ & & \sigma_r & \\ & & & \end{bmatrix} = U\Sigma
$$

Hence,

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

References

[Lay11a] D.C. Lay. *Linear Algebra and Its Applications*, page 417. Pearson College Division, 4th edition, 2011.

[Lay11b] D.C. Lay. *Linear Algebra and Its Applications*, page 396. Pearson College Division, 4th edition, 2011.