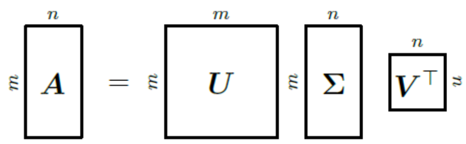
Matrix Decomposition and Approximation

3. Singular Value Decomposition

3.1 Introduction

Theorem 4.22 (SVD Theorem). *Let $\boldsymbol{A}^{m\times n}$ be a rectangular matrix of rank $r \in [0, \min(m, n)]$. The SVD of $\boldsymbol{A}$ is a decomposition of the form*



(4.64)

with an orthogonal matrix $\boldsymbol{U} \in \mathbb{R}^{m\times m}$ with column vectors $\boldsymbol{u}_i$, $i = 1, \dots, m$, and an orthogonal matrix $\boldsymbol{V} \in \mathbb{R}^{n\times n}$ with column vectors $\boldsymbol{v}_j$, $j = 1, \dots, n$. Moreover, $\boldsymbol{\Sigma}$ is an $m \times n$ matrix with $\Sigma_{ii} = \sigma_i \geqslant 0$ and $\Sigma_{ij} = 0$, $i \neq j$.

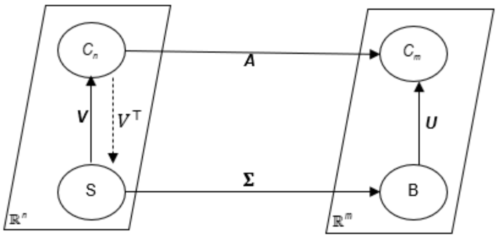
The diagonal entries σ_i, $i = 1, \dots, r$, of $\boldsymbol{\Sigma}$ are called the *singular values*, $\boldsymbol{u}_i$ are called the *left-singular vectors*, and $\boldsymbol{v}_j$ are called the *right-singular vectors*. By convention, the singular values are ordered, i.e., $\sigma_1 \geqslant \sigma_2 \geqslant \sigma_r \geqslant 0$.

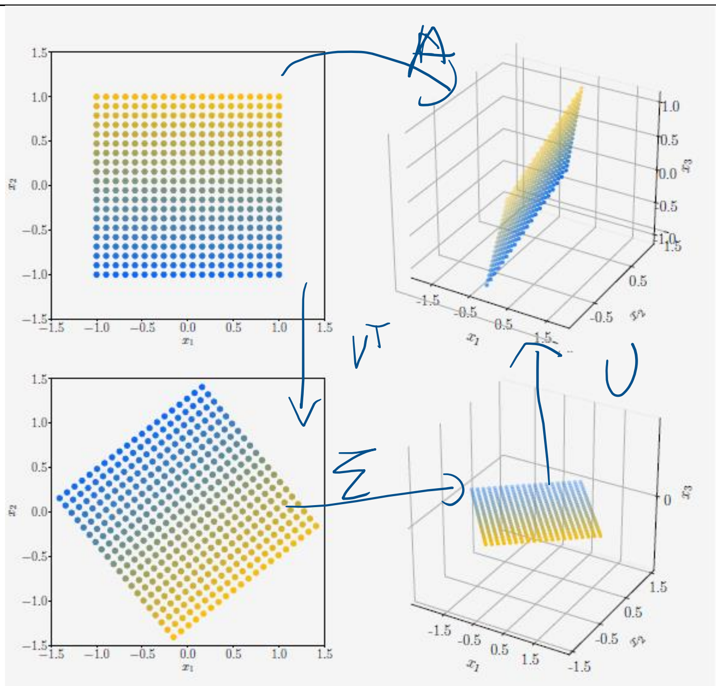
Remark

- The matrix Σ is unique
- The matrix Σ have the same size of A, this mean Σ can be rectangular

Several types of Σ.

$m = n$	$m > n$	$m < n$
$\begin{bmatrix} \sigma_1 & 0 & \dots & 0 \\ 0 & \sigma_2 & \dots & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & \dots & \sigma_n \end{bmatrix}$	$\begin{bmatrix} \sigma_1 & 0 & \dots & 0 \\ 0 & \sigma_2 & \dots & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & \dots & \sigma_m \\ 0 & 0 & \dots & 0 \\ \vdots & & & \\ 0 & 0 & \dots & 0 \end{bmatrix}$	$\begin{bmatrix} \sigma_1 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & \sigma_2 & \dots & 0 & 0 & & 0 \\ 0 & 0 & \ddots & 0 & 0 & & 0 \\ 0 & 0 & \dots & \sigma_n & 0 & \dots & 0 \end{bmatrix}$ $m - m$
	$m - m$	





3.2 Construction of the SVD

3.2.1 $\boldsymbol{m = n}$

1. Construct $\boldsymbol{V}$

1.1 Determine the eigenvalues of $A^\top A$.

1.2 Order the eigenvalues such that $\lambda_1 \geq \ \dots \geq \ \lambda_n$ and their respective eigenvectors.

1.3 Normalize the eigenvectors.

1.4 $\boldsymbol{V = [v_1\, v_2 \dots .\, v_n],\ v_i\ eigenvector.}$

2. Construct $\boldsymbol{\Sigma}$

- $D = diag(\lambda_1, \dots, \lambda_m);$
- $V\mathrm{D}\, V^\top = A^\top A = (U\Sigma V^\top)^\top (U\Sigma V^\top) = V\Sigma \underbrace{U^\top U}_{I} \Sigma V^\top = V\Sigma^2\, V^\top;$
- $VDV^\top = V\Sigma^2\, V^\top \Leftrightarrow D = \Sigma^2;$

2.1 $\mathbf{\Sigma} = \begin{bmatrix} \sqrt{\boldsymbol{\lambda_1}} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \sqrt{\boldsymbol{\lambda_2}} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \ddots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \dots & \sqrt{\boldsymbol{\lambda_n}} \end{bmatrix}$.

3. **Construct $\boldsymbol{U}$**

3.1 Determine $\Sigma^{-1} = \begin{bmatrix} \frac{1}{\sqrt{\lambda_1}} & 0 & \dots & 0 \\ 0 & \frac{1}{\sqrt{\lambda_2}} & \dots & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & \dots & \frac{1}{\sqrt{\lambda_n}} \end{bmatrix}$, if for some i, $\lambda_i = 0$, then replace $\frac{1}{\sqrt{\lambda_i}}$ by 0.

3.2 If for every i $\lambda_i \neq 0$ then:

3.2.1 $\boldsymbol{U = AV\Sigma^{-1}}$

- $U\Sigma V^\top = A \Leftrightarrow U\Sigma = AV \Leftrightarrow U = AV\Sigma^{-1}$;
- U is orthogonal $U^\top U = (AV\Sigma^{-1})^\top(AV\Sigma^{-1}) = \Sigma^{-1} \underbrace{V^\top A^\top AV}_{\Sigma^2} \Sigma^{-1} = I_n$;
- $A^\top A = VDV^\top = V\Sigma^2 V^\top \Leftrightarrow V^\top A^\top AV = \Sigma^2$;

3.3 If for some i, $\lambda_i = 0$, then:

3.3.1 $W1 = AV\Sigma_{\mathrm{I}}$.

Suppose n_1 is the number $\lambda_i = 0$, then W_1 has n_1 null columns.

1.3.2 Choose the submatrix W, constitute by the $n - n_1$ non null columns of W_1 .

1.3.3 Find the subspace $Z = (W)^\perp$ (orthogonal complement);

1.3.3.1 Choose $u_1 .. u_{n_1}, n_1$ vectors in $\mathbb{R}^n$, such that $rank([W, u_1 .. u_{n_1}]) = n$.

The vectors and the columns of W must be linearly independent.

3.3.3.2 Using the Gram-Schmidt process find vectors $z_1, \dots . z_{n_1}$, such that $\{w_1 \dots, w_{m-n_1}, z_1, \dots . z_{n_1}\}$ is an orthonormal base.

3.3.4 $\boldsymbol{U = [w_1 \dots, w_{m-n_1}, z_1, \dots . z_{n_1}]}$.

3.2.2 $\boldsymbol{m < n}$

1. **Construct $\boldsymbol{V}$**

1.1 Determine the eigenvalues of $A^{\top}A$.

1.2 Order the eigenvaluessuch that $\lambda_1 \geq \ldots \geq \lambda_n$ and their respective eigenvectors .($\underbrace{\lambda_{m+1} = \cdots = \lambda_n}_{n-m\ eigenvalues\ are\ zero} = 0)$.

1.3 Normalize the eigenvectors

$$V = [v_1\ v_2 \ldots .\, v_n]\ \ v_i\ eigenvector\ .$$

2. **Construct Σ**

$$\Sigma = \begin{bmatrix} \sqrt{\lambda_1} & 0 & \ldots & 0 & 0 & \ldots & 0 \\ 0 & \sqrt{\lambda_2} & \ldots & 0 & 0 & & 0 \\ 0 & 0 & \ddots & 0 & 0 & & 0 \\ 0 & 0 & \ldots & \sqrt{\lambda_m} & 0 & \ldots & 0 \end{bmatrix}.$$

3. **Construct $\boldsymbol{U}$**

3.1 Determine $\boldsymbol{\Sigma_I} = \begin{bmatrix} \frac{1}{\sqrt{\lambda_1}} & 0 & \ldots & 0 \\ 0 & \frac{1}{\sqrt{\lambda_2}} & \ldots & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & \ldots & \frac{1}{\sqrt{\lambda_m}} \\ 0 & 0 & \ldots & 0 \\ \vdots & & & \vdots \\ 0 & 0 & \ldots & 0 \end{bmatrix}$, if for some i $\lambda_i = 0$, then replace $\frac{1}{\sqrt{\lambda_i}}$ by 0.

3.2 If for every i $\lambda_i \neq 0$ then then $\boldsymbol{U = AV\Sigma_I}$**.**

3.3 If for some i, $\lambda_i = 0$ then

3.3.1 $W1 = AV\Sigma_I$

Suppose n_1 is the number $\lambda_i = 0$, then W_1 has n_1 null columns.

3.3.2 Choose the submatrix W, constitute by the $m - n_1$ non null columns of W_1 .

3.3.3 Find the subspace $Z = (W)^{\perp}$ (orthogonal complement).

1.3.3.2 Choose $u_1 .. u_{n_1}, n_1$ vectors in $\mathbb{R}^m$, such that $rank([W, u_1 .. u_{n_1}]) = m$

The vectors and the columns of W must be linearly independent.

3.3.3.2 Using the Gram-Schmidt process find vectors $z_1, \ldots .\, z_{n_1}$, such that $\{w_1 \ldots, w_{m-n_1}, z_1, \ldots .\, z_{n_1}\}$ is a orthonormal base.

3.3. 4 $\boldsymbol{U} = [\boldsymbol{w_1} \ldots, \boldsymbol{w_{m-n_1}}, \boldsymbol{z_1}, \ldots .\, \boldsymbol{z_{n_1}}]$

3.2.3 $\boldsymbol{m > n}$

1. Construct $\boldsymbol{V}$

1.1 Determine the eigenvalues of $A^{\top}A$.

1.2 Order the eigenvalues such that $\lambda_1 \geq \ ... \geq \ \lambda_n$ and their respective eigenvectors.

1.3 Normalize the eigenvectors.

$$V = [v_1\ v_2\ ....\ v_n]\ \ v_i\ eigenvector\ \ .$$

2 Construct $\boldsymbol{\Sigma}$

$$\boldsymbol{\Sigma} = \begin{bmatrix} \sqrt{\lambda_1} & 0 & \dots & 0 \\ 0 & \sqrt{\lambda_2} & \dots & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & \dots & \sqrt{\lambda_n} \\ 0 & 0 & \dots & 0 \\ \vdots & & & \vdots \\ 0 & 0 & \dots & 0 \end{bmatrix}.$$

3 Construct $\boldsymbol{U}$

3.1 Determine $\Sigma_I = \begin{bmatrix} \frac{1}{\sqrt{\lambda_1}} & 0 & \dots & 0 \\ 0 & \frac{1}{\sqrt{\lambda_2}} & \dots & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & \dots & \frac{1}{\sqrt{\lambda_n}} \\ 0 & 0 & \dots & 0 \\ \vdots & & & \vdots \\ 0 & 0 & \dots & 0 \end{bmatrix}$, if for some i $\lambda_i = 0$, then replace $\frac{1}{\sqrt{\lambda_i}}$ by 0.

3.2 $\boldsymbol{W1 = AV\Sigma_I}$

$W_1 =$ has at least $m - n$ null columns, suppose n_1 is the number of null columns.

3.3 Choose W , the submatrix constitute by the $m - n_1$, non null columns of W_1 .

3.4 Find the subspace $Z = (W)^{\perp}$ (orthogonal complement).

3.4.1 Choose $u_1 .. u_{n_1}, n_1$ vectors in $\mathbb{R}^m$, such that $rank\left([W, u_1 .. u_{n_1}]\right) = m$.

The vectors and the columns of W must be linearly independent.

3.4.2 Using the Gram-Schmidt process find vectors $z_1, ...z_{n_1}$, such that $\{w_1 \dots, w_{m-n_1}, z_1, ...z_{n_1}\ \}$ is a orthonormal base.

3.5 $\boldsymbol{U = [w_1 \dots, w_{m-n_1}, z_1, ...z_{n_1}]}$

4. Matrix Approximation

Consider $A \in \mathbb{R}^{m\times n}$ with the factorization $A = U\Sigma V^\top \in \mathbb{R}^{m\times n}$ with $U \in \mathbb{R}^{m\times m}$, $V \in \mathbb{R}^{n\times n}$ and $\Sigma \in \mathbb{R}^{m\times n}$.

We construct a rank-1 matrix $A_i \in \mathbb{R}^{m\times n}$ as

$$A_i := u_i v_i^\top$$

u_i is the ith column of U, and v_i is the ith column of V.

A matrix $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ of rank r can be written as a sum of rank-1 matrices $\boldsymbol{A}_i$ so that

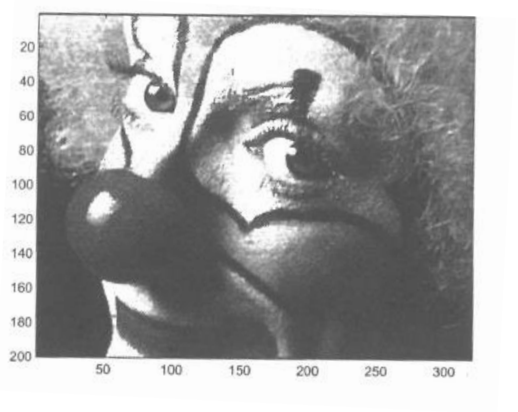
$$\boldsymbol{A} = \sum_{i=1}^{r} \sigma_i \boldsymbol{u}_i \boldsymbol{v}_i^\top = \sum_{i=1}^{r} \sigma_i \boldsymbol{A}_i\,, \tag{4.91}$$

Definition

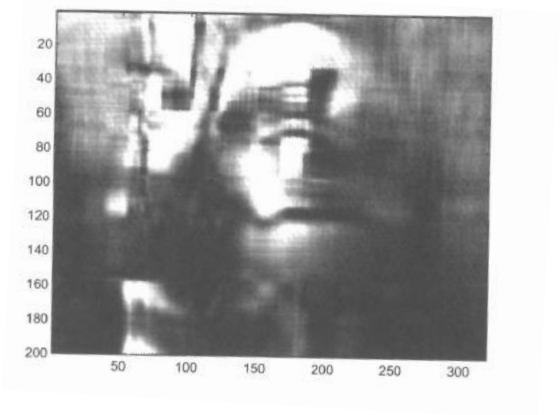
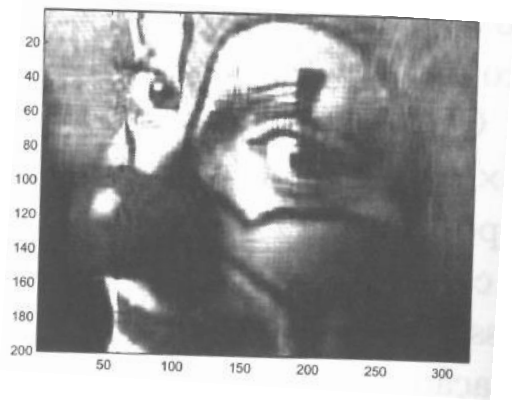
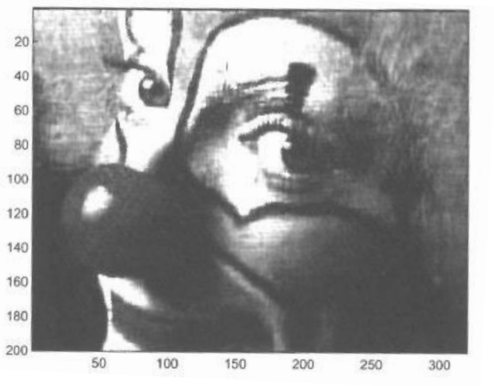
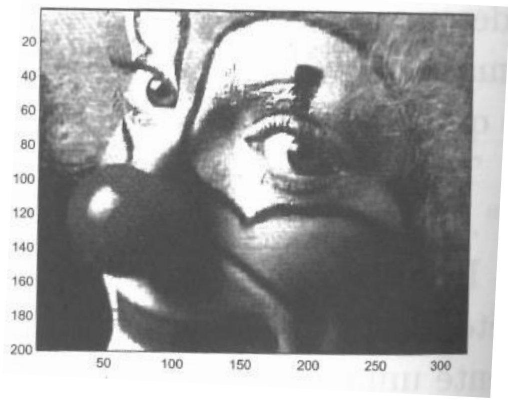
$$\widehat{\boldsymbol{A}}(k) := \sum_{i=1}^{k} \sigma_i \boldsymbol{u}_i \boldsymbol{v}_i^\top = \sum_{i=1}^{k} \sigma_i \boldsymbol{A}_i \tag{4.92}$$

The matrix $\hat{A}(k)$ is called $rank - k$ approximation of A with $rank\left(\hat{A}(k)\right) = k$.

Example



$rank(A) = 200$

	
Approximation with rank-10 matrix	Approximation with rank-20 matrix
	
Approximation with rank-30 matrix	Approximation with rank-50 matrix $\boldsymbol{\sigma_{50} = 0.02\sigma_1}$

Definition 4.23 (Spectral Norm of a Matrix)**.** For $\boldsymbol{x} \in \mathbb{R}^n \backslash \{\boldsymbol{0}\}$, the *spectral norm* of a matrix $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ is defined as

$$\|\boldsymbol{A}\|_2 := \max_{\boldsymbol{x}} \frac{\|\boldsymbol{A}\boldsymbol{x}\|_2}{\|\boldsymbol{x}\|_2} . \tag{4.93}$$

We introduce the notation of a subscript in the matrix norm (left-hand side), similar to the Euclidean norm for vectors (right-hand side), which has subscript 2. The spectral norm (4.93) determines how long any vector $\boldsymbol{x}$ can at most become when multiplied by $\boldsymbol{A}$.

Theorem 4.24. *The spectral norm of $\boldsymbol{A}$ is its largest singular value σ_1.*

Theorem 4.25 (Eckart-Young Theorem (Eckart and Young, 1936))**.** *Consider a matrix $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ of rank r and let $\boldsymbol{B} \in \mathbb{R}^{m\times n}$ be a matrix of rank k. For any $k \leqslant r$ with $\widehat{\boldsymbol{A}}(k) = \sum_{i=1}^{k} \sigma_i \boldsymbol{u}_i \boldsymbol{v}_i^\top$ it holds that*

$$\widehat{\boldsymbol{A}}(k) = \operatorname{argmin}_{\operatorname{rk}(\boldsymbol{B})=k} \left\|\boldsymbol{A} - \boldsymbol{B}\right\|_2 , \tag{4.94}$$

$$\left\|\boldsymbol{A} - \widehat{\boldsymbol{A}}(k)\right\|_2 = \sigma_{k+1} . \tag{4.95}$$