Singular Value Decomposition (matrix factorization)

Singular Value Decomposition

The SVD is a factorization of a $m\times n$ matrix into

$$\underbrace{A}_{m\times n} = \underbrace{U}_{m\times m}\,\underbrace{\Sigma}_{m\times n}\,\underbrace{V^T}_{n\times n}$$

where $\boldsymbol{U}$ is a $m\times m$ orthogonal matrix, $\boldsymbol{V}^T$ is a $n\times n$ orthogonal matrix and $\boldsymbol{\Sigma}$ is a $m\times n$ diagonal matrix.

$\sigma_1 \geq \sigma_2 \geq \sigma_3 \dots$

u_i (left singular vectors)

singular values

right singular vector

For a square matrix ($\boldsymbol{m} = \boldsymbol{n}$):

$$\boldsymbol{A} = \underbrace{\begin{pmatrix} \vdots & \dots & \vdots \\ \boldsymbol{u}_1 & \dots & \boldsymbol{u}_n \\ \vdots & \dots & \vdots \end{pmatrix}}_{U}\underbrace{\begin{pmatrix} \sigma_1 & & \phi \\ & \ddots & \\ \phi & & \sigma_n \end{pmatrix}}_{\Sigma}\underbrace{\begin{pmatrix} \vdots & \dots & \vdots \\ \boldsymbol{v}_1 & \dots & \boldsymbol{v}_n \\ \vdots & \dots & \vdots \end{pmatrix}^T}_{V}$$

$\sigma_1 \geq \sigma_2 \geq \sigma_3 \dots$

$$A = \begin{pmatrix} | & | & & | \\ u_1 & u_2 & \cdots & u_n \\ | & | & & | \end{pmatrix}\begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \end{pmatrix}\begin{pmatrix} — & v_1^T & — \\ — & v_2^T & — \\ & \vdots & \\ — & v_n^T & — \end{pmatrix}$$

Reduced SVD

What happens when $\boldsymbol{A}$ is not a square matrix?

1) $\boldsymbol{m > n}$

$$\boldsymbol{A} = \boldsymbol{U}\,\boldsymbol{\Sigma}\,\boldsymbol{V}^{T} = \begin{pmatrix} \vdots & \cdots & \vdots & \cdots & \vdots \\ \boldsymbol{u}_1 & \cdots & \boldsymbol{u}_n & \cdots & \boldsymbol{u}_m \\ \vdots & \cdots & \vdots & \cdots & \vdots \end{pmatrix} \begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \\ & 0 & \\ & \vdots & \\ & 0 & \end{pmatrix} \begin{pmatrix} \cdots & \mathbf{v}_1^T & \cdots \\ \vdots & \vdots & \vdots \\ \cdots & \mathbf{v}_n^T & \cdots \end{pmatrix}$$

m×n n×n m×n n n n n m A

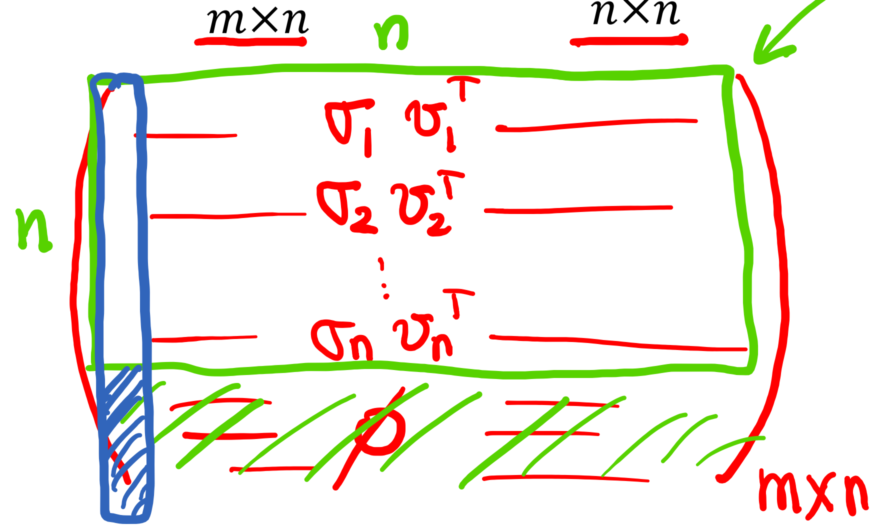
$m{\times}m$ $\qquad m{\times}n \qquad n{\times}n$

m×m

m > n

$$A = \underset{=}{U_R}\, \underset{=}{\Sigma_R}\, V^T$$

m×n n×n n×n



Reduced SVD

2) $n > m$ $\quad \begin{array}{c} n \\ \left[\quad \underset{=}{A} \quad\right] m \end{array}$

$$\underset{m\times n}{A} = \underset{m\times m}{U}\,\underset{m\times n}{\Sigma}\,\underset{n\times n}{V^T} = \underset{m\times m}{\begin{pmatrix} \vdots & \cdots & \vdots \\ u_1 & \cdots & u_m \\ \vdots & & \vdots \end{pmatrix}} \underset{m\times n}{\begin{pmatrix} \sigma_1 & & & 0 & \\ & \ddots & & & \ddots \\ & & \sigma_m & & & 0 \end{pmatrix}} \underset{n\times n}{\begin{pmatrix} \cdots & \mathbf{v}_1^T & \cdots \\ \vdots & \vdots & \vdots \\ \cdots & \mathbf{v}_m^T & \cdots \\ \vdots & \vdots & \vdots \\ \cdots & \mathbf{v}_n^T & \cdots \end{pmatrix}}$$

$m \times m$ (blue, under U); $m \times m$ (blue, under the $\sigma_1 \ldots \sigma_m$ block); $m \times n$ (blue, under the $\mathbf{v}_1^T \ldots \mathbf{v}_m^T$ rows)

$m\times m$ $\qquad$ $m\times n$ $\qquad$ $n\times n$

$\boxed{\begin{array}{l} \underline{\underline{m<n}} \\ \underset{=}{A} = \underset{m\times m}{\underset{=}{U}}\ \underset{m\times m}{\underset{=}{\Sigma_R}}\ \underset{m\times n}{V_R^T} \end{array}}$

General:

$$\underset{m\times n}{A} = \underset{=}{U_R}\ \underset{=}{\Sigma_R}\ \underset{=}{V_R^T}$$

	U_R	Σ_R	V_R^T
$m<n$	$m\times m$	$m\times m$	$m\times n$
$m>n$	$m\times n$	$n\times n$	$n\times n$

U_R : $m\times k$

Σ_R : $k\times k$ $\qquad k = \min(m, n)$

V_R^T : $k\times n$

Let's take a look at the product $\boldsymbol{\Sigma}^T\boldsymbol{\Sigma}$, where $\boldsymbol{\Sigma}$ has the singular values of a $\boldsymbol{A}$, a $m \times n$ matrix.

$m > n$

$$\boldsymbol{\Sigma}^T\boldsymbol{\Sigma} = \underbrace{\begin{pmatrix} \sigma_1 & & & 0 & & \\ & \ddots & & & \ddots & \\ & & \sigma_n & & & 0 \end{pmatrix}}_{n\times m} \underbrace{\begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \\ & & 0 \\ & & \vdots \\ & & 0 \end{pmatrix}}_{m\times n} = \underbrace{\begin{pmatrix} \sigma_1{}^2 & & \\ & \ddots & \\ & & \sigma_n{}^2 \end{pmatrix}}_{n\times n} \rightarrow \Sigma_R^2$$

$n > m$

$$\boldsymbol{\Sigma}^T\boldsymbol{\Sigma} = \underbrace{\begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_m \\ & & 0 \\ & & \vdots \\ & & 0 \end{pmatrix}}_{n\times m} \underbrace{\begin{pmatrix} \sigma_1 & & & 0 & & \\ & \ddots & & & \ddots & \\ & & \sigma_m & & & 0 \end{pmatrix}}_{m\times n} = \underbrace{\begin{pmatrix} \sigma_1{}^2 & & & 0 & & \\ & \ddots & & & \ddots & \\ & & \sigma_m{}^2 & & & 0 \\ 0 & & 0 & & & \\ & \ddots & & & \ddots & \\ & & 0 & & & 0 \end{pmatrix}}_{n\times n}$$

Σ_R^2

Assume $\boldsymbol{A}$ with the singular value decomposition $\boldsymbol{A} = \boldsymbol{U}\,\boldsymbol{\Sigma}\,\boldsymbol{V}^{T}$. Let's take a look at the eigenpairs corresponding to $\boldsymbol{A}^{T}\boldsymbol{A}$:

(x, λ) $\boxed{A^T A x = \lambda x}$

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

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

$$= V \Sigma^T I \Sigma V^T$$

$$= V \Sigma^T \Sigma V^T$$

$$\boxed{A^T A = \boxed{V}\,\boxed{\Sigma^2}\, V^T}$$

$(ABC)^T = C^T B^T A^T$

$\underline{\underline{U}}^T = \underline{\underline{U}}^{-1}$

$\underline{\underline{U}}^{-1}\underline{\underline{U}} = \underline{\underline{I}}$

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

Diagonalization:

$\longleftrightarrow$ $\underline{\underline{B}} = \underline{\underline{\underline{X}\,\underline{D}\,\underline{X}^{-1}}}$

$\Rightarrow$ columns of $\underline{\underline{V}}$ are the eigenvectors of $\underline{\underline{A}}^T\underline{\underline{A}}$

$\Rightarrow$ diagonal entries of $\underline{\underline{\Sigma}}^2$ are the eigenvalues of $A^T A$

$x, \lambda = \text{eig}(A^T A)$ $\quad \boxed{\lambda_i = \sigma_i^2}$

In a similar way,

$\underline{A^TA}$

$$AA^T = (U\Sigma V^T)(U\Sigma V^T)^T$$

$$= U\Sigma V^T (V^T)^T \Sigma^T U^T$$

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

$$V^{-1} = V^T$$

$$= U\Sigma\Sigma^T U^T$$

$$\boxed{AA^T = U\Sigma^2 U^T}$$

$$B = XDX^{-1}$$

→ columns of U are the eigenvectors of AA^T

How can we compute an SVD of a matrix A ?

1. Evaluate the n eigenvectors $\mathbf{v}_i$ and eigenvalues λ_i of $\boldsymbol{A}^T\boldsymbol{A}$ la.eig(A^TA)
2. Make a matrix $\boldsymbol{V}$ from the normalized vectors $\mathbf{v}_i$. The columns are called "right singular vectors".

$$\boldsymbol{V} = \begin{pmatrix} \vdots & \cdots & \vdots \\ \mathbf{v}_1 & \cdots & \mathbf{v}_n \\ \vdots & \cdots & \vdots \end{pmatrix}$$

3. Make a diagonal matrix from the square roots of the eigenvalues.

singular values $\sigma_i^2 = \lambda_i$

$$\boldsymbol{\Sigma} = \begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \end{pmatrix} \quad \sigma_i = \sqrt{\lambda_i} \ \text{ and } \ \sigma_1 \geq \sigma_2 \geq \sigma_3 \ldots$$

4. Find $\boldsymbol{U}$: $\boldsymbol{A} = \boldsymbol{U}\,\boldsymbol{\Sigma}\,\boldsymbol{V}^T \Longrightarrow \boldsymbol{U}\,\boldsymbol{\Sigma} = \boldsymbol{A}\,\boldsymbol{V}$. The columns are called the "left singular vectors".

$\boxed{U = AV\Sigma^{-1}}$

Singular values are always non-negative

A^TA

- A matrix is positive definite if $\boldsymbol{x}^T\boldsymbol{B}\boldsymbol{x} > \mathbf{0}$ for $\forall \boldsymbol{x} \neq \mathbf{0}$
- A matrix is positive semi-definite if $\boldsymbol{x}^T\boldsymbol{B}\boldsymbol{x} \geq \mathbf{0}$ for $\forall \boldsymbol{x} \neq \mathbf{0}$

$$x^T(A^TA)x = \underbrace{(Ax)^T}_{y}\underbrace{Ax}_{y} = \|Ax\|_2^2 \geq 0$$

$$y \cdot y = \|y\|_2^2$$

A^TA is positive semi-definite

$$A^TAx = \lambda x \longrightarrow (x, \lambda)$$

$$x^TA^TAx = \lambda x^Tx = \lambda \underbrace{\|x\|_2^2}_{\geq 0}$$

$$\lambda = \frac{x^TA^TAx}{\|x\|_2^2} = \frac{\|Ax\|_2^2}{\|x\|^2} \geq 0$$

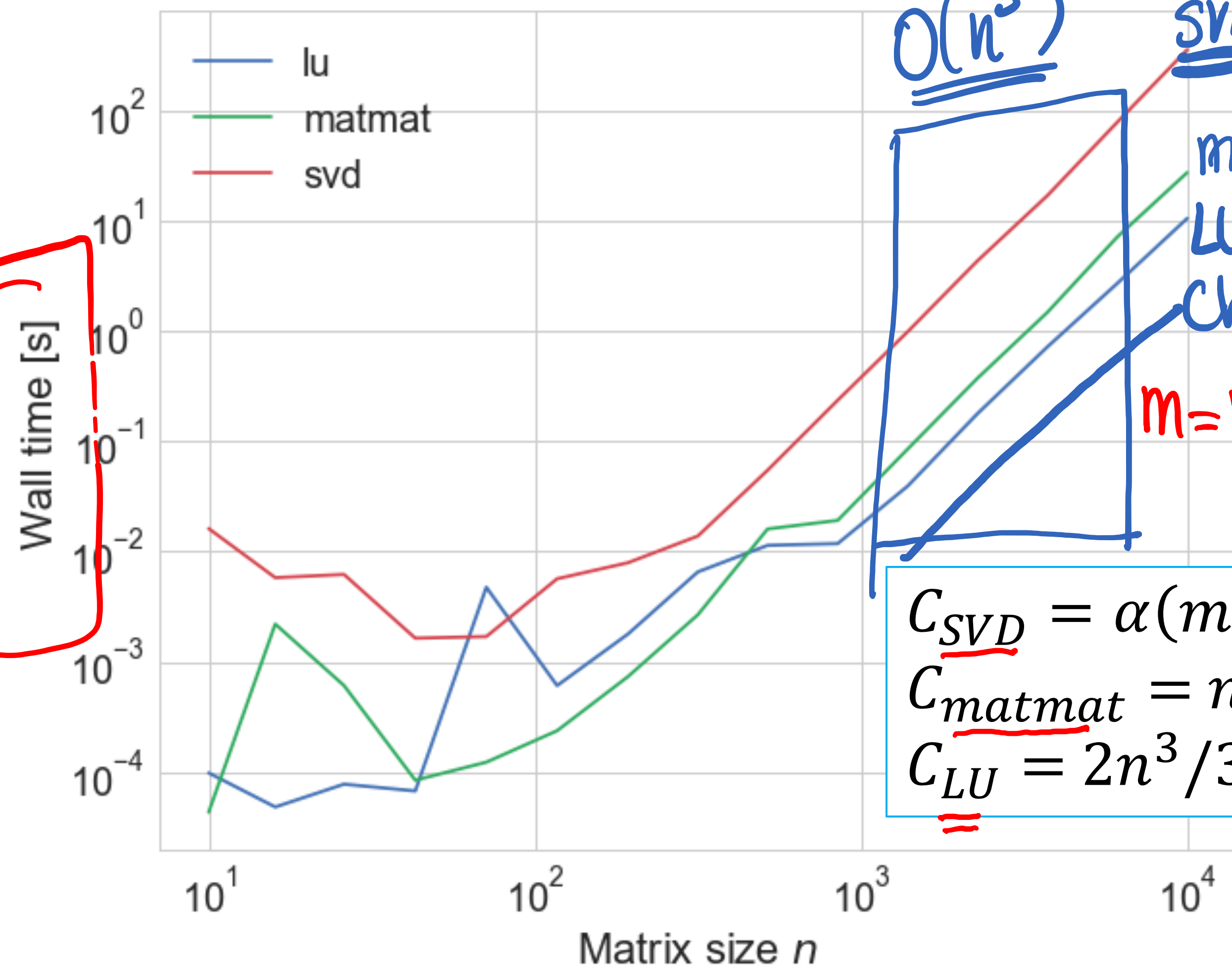
A^TA

$$\boxed{\lambda_i \geq 0}$$

Cost of SVD

$A_{m \times n}$

The cost of an SVD is proportional to $\boldsymbol{m\,n^2 + n^3}$ where the constant of proportionality constant ranging from 4 to 10 (or more) depending on the algorithm.



$O(n^3)$

SVD cost: $\alpha\,(mn^2+n^3)$

mat.mat

LU

Chol

$(4, 10)$

$m=n \rightarrow \text{cost} = \alpha\, n^3$

$$C_{SVD} = \alpha(m\,n^2 + n^3) = O(\,n^3)$$
$$C_{matmat} = n^3 = O(\,n^3)$$
$$C_{LU} = 2n^3/3 = O(\,n^3)$$

$\frac{2}{3} n^3$

$A_{n \times n}$ Chol. $\frac{1}{3} n^3$

SVD summary:

- The SVD is a factorization of a $m \times n$ matrix into $\boldsymbol{A} = \boldsymbol{U}\,\boldsymbol{\Sigma}\,\boldsymbol{V}^{\boldsymbol{T}}$ where $\boldsymbol{U}$ is a $m \times m$ orthogonal matrix, $\boldsymbol{V}^{\boldsymbol{T}}$ is a $n \times n$ orthogonal matrix and $\boldsymbol{\Sigma}$ is a $m \times n$ diagonal matrix.

- In reduced form: $\boldsymbol{A} = \boldsymbol{U}_{\boldsymbol{R}}\boldsymbol{\Sigma}_{\boldsymbol{R}}{\boldsymbol{V}_{\boldsymbol{R}}}^{\boldsymbol{T}}$, where $\boldsymbol{U}_{\boldsymbol{R}}$ is a $m \times k$ matrix, $\boldsymbol{\Sigma}_{\boldsymbol{R}}$ is a $k \times k$ matrix, and $\boldsymbol{V}_{\boldsymbol{R}}$ is a $n \times k$ matrix, and $k = \min(m, n)$.

- The columns of $\boldsymbol{V}$ are the eigenvectors of the matrix $\boldsymbol{A}^{\boldsymbol{T}}\boldsymbol{A}$, denoted the right singular vectors.

- The columns of $\boldsymbol{U}$ are the eigenvectors of the matrix $\boldsymbol{A}\boldsymbol{A}^{\boldsymbol{T}}$, denoted the left singular vectors.

- The diagonal entries of $\boldsymbol{\Sigma}^{\boldsymbol{2}}$ are the eigenvalues of $\boldsymbol{A}^{\boldsymbol{T}}\boldsymbol{A}$. $\sigma_i = \sqrt{\lambda_i}$ are called the singular values.

- The singular values are always non-negative (since $\boldsymbol{A}^{\boldsymbol{T}}\boldsymbol{A}$ is a positive semi-definite matrix, the eigenvalues are always $\lambda \geq 0$)