Singular Value Decomposition
(matrix factorization)

Singular Value Decomposition

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.

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

$$\sigma_1 \geq \sigma_2 \geq \sigma_3 \ldots$$

Reduced SVD

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

1) $\boldsymbol{m > n}$

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

$m \times m$ $\quad$ $m \times n$ $\quad$ $n \times n$

We can instead re-write the above as:

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

Where $\boldsymbol{U_R}$ is a $m \times n$ matrix and $\boldsymbol{\Sigma_R}$ is a $n \times n$ matrix

Reduced SVD

2) $\boldsymbol{n > m}$

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

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

We can instead re-write the above as:

$$\boldsymbol{A} = \boldsymbol{U}\,\boldsymbol{\Sigma}_{\boldsymbol{R}}\boldsymbol{V}_{\boldsymbol{R}}{}^{\boldsymbol{T}}$$

where $\boldsymbol{V}_{\boldsymbol{R}}$ is a $n\times m$ matrix and $\boldsymbol{\Sigma}_{\boldsymbol{R}}$ is a $m\times m$ matrix

In general:

$$\boldsymbol{A} = \boldsymbol{U}_{\boldsymbol{R}}\boldsymbol{\Sigma}_{\boldsymbol{R}}\boldsymbol{V}_{\boldsymbol{R}}{}^{\boldsymbol{T}}$$

$\boldsymbol{U}_{\boldsymbol{R}}$ is a $m\times k$ matrix

$\boldsymbol{\Sigma}_{\boldsymbol{R}}$ is a $k\times k$ matrix

$\boldsymbol{V}_{\boldsymbol{R}}$ is a $n\times k$ matrix

$k = \min(m, n)$

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.

$m > n$

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

$n > m$

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

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

$$\boldsymbol{A^T A = \left(U\,\Sigma\,V^T\right)^T \left(U\,\Sigma\,V^T\right)}$$

$$\boldsymbol{\left(V^T\right)^T (\,\Sigma\,)^T U^T \left(U\,\Sigma\,V^T\right) = V \Sigma^T U^T\, U\,\Sigma\,V^T = V\,\Sigma^T \Sigma\,V^T}$$

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

Recall that columns of $\boldsymbol{V}$ are all linear independent (orthogonal matrix), then from diagonalization ($\boldsymbol{B = XDX^{-1}}$), we get:

- the columns of $\boldsymbol{V}$ are the eigenvectors of the matrix $\boldsymbol{A^T A}$
- The diagonal entries of $\boldsymbol{\Sigma^2}$ are the eigenvalues of $\boldsymbol{A^T A}$

Let's call λ the eigenvalues of $\boldsymbol{A^T A}$, then ${\sigma_i}^2 = \lambda_i$

In a similar way,

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

$$(\boldsymbol{U\,\Sigma\,V^T})(\boldsymbol{V^T})^{\boldsymbol{T}}(\,\boldsymbol{\Sigma}\,)^{\boldsymbol{T}}\boldsymbol{U^T} = \boldsymbol{U\,\Sigma\,}{\color{red}\boldsymbol{V^TV}}\boldsymbol{\Sigma^T U^T} = \boldsymbol{U\Sigma\,\Sigma^T U^T}$$

Hence $\boldsymbol{AA^T} = \boldsymbol{U\,\Sigma^2\,U^T}$

Recall that columns of $\boldsymbol{U}$ are all linear independent (orthogonal matrices), then from diagonalization ($\boldsymbol{B} = \boldsymbol{XDX^{-1}}$), we get:

- The columns of $\boldsymbol{U}$ are the eigenvectors of the matrix $\boldsymbol{AA^T}$

How can we compute an SVD of a matrix A ?

1. Evaluate the n eigenvectors $\mathbf{v}_i$ and eigenvalues λ_i of $\boldsymbol{A}^T\boldsymbol{A}$
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.

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

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".

Demo "Computing the SVD"

True or False?

$\boldsymbol{A}$ has the singular value decomposition $\boldsymbol{A} = \boldsymbol{U}\,\boldsymbol{\Sigma}\,\boldsymbol{V}^{\boldsymbol{T}}$.

- The matrices $\boldsymbol{U}$ and $\boldsymbol{V}$ are not singular
- The matrix $\boldsymbol{\Sigma}$ can have zero diagonal entries
- $\|\boldsymbol{U}\|_2 = 1$
- The SVD exists when the matrix $\boldsymbol{A}$ is singular
- The algorithm to evaluate SVD will fail when taking the square root of a negative eigenvalue

Singular values are always non-negative

Singular values cannot be negative since $\boldsymbol{A^T A}$ **is a positive semi-definite matrix** (for real matrices $\boldsymbol{A}$)

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

- What do we know about the matrix $\boldsymbol{A^T A}$?

$$\boldsymbol{x^T\ A^T A\ x = (Ax)^T Ax = \|Ax\|_2^2} \geq 0$$

- Hence we know that $\boldsymbol{A^T A}$ is a positive semi-definite matrix

- A positive semi-definite matrix has non-negative eigenvalues

$$\boldsymbol{Bx} = \lambda \boldsymbol{x} \Longrightarrow \boldsymbol{x^T B x} = \boldsymbol{x^T}\ \lambda\ \boldsymbol{x} = \lambda\ \|\boldsymbol{x}\|_{\boldsymbol{2}}^{\boldsymbol{2}} \geq 0 \Longrightarrow \lambda \geq 0$$

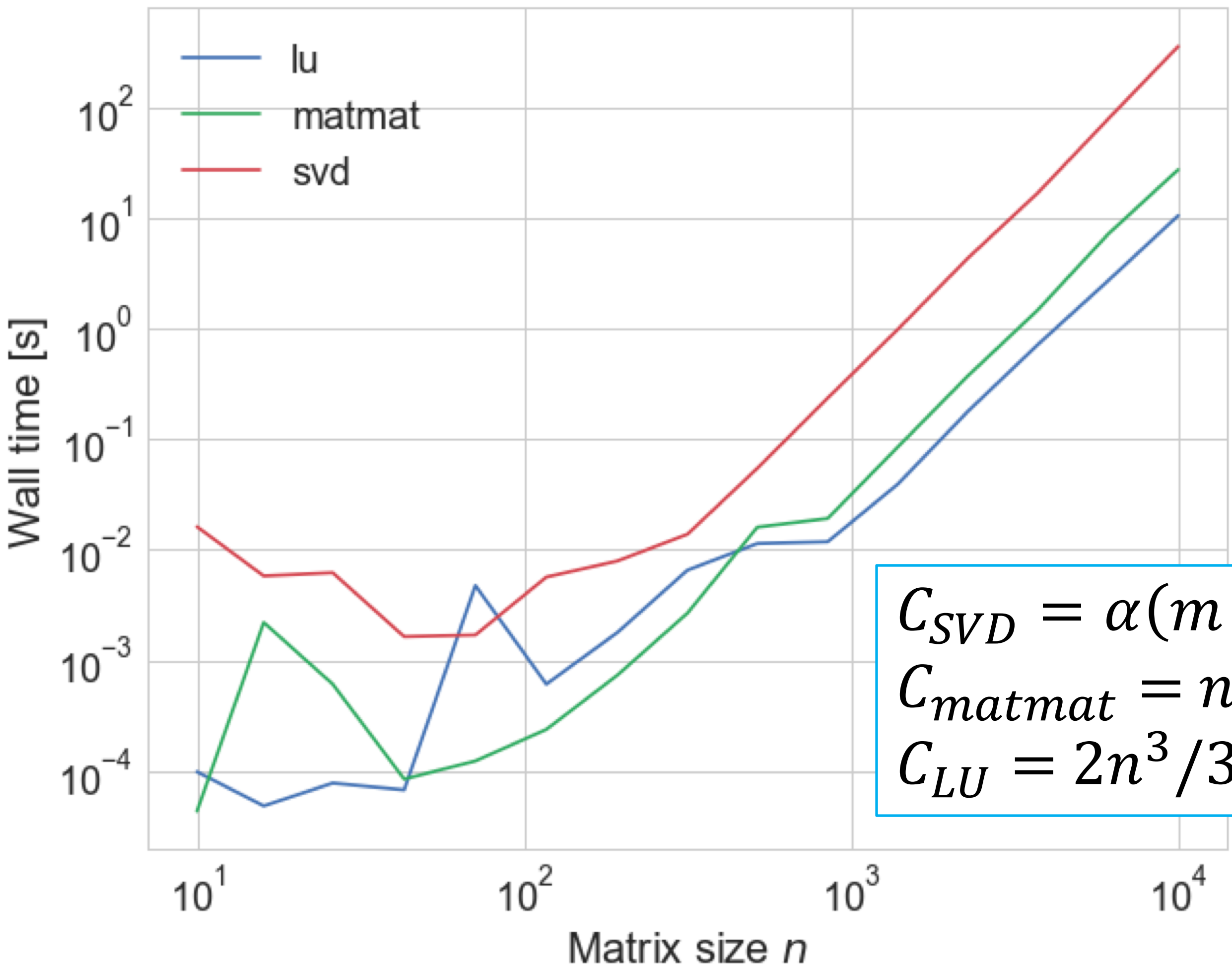
Euclidean norm of orthogonal matrices:

$$\|\boldsymbol{U}\|_2 = \max_{\|\boldsymbol{x}\|_2=1} \|\boldsymbol{Ux}\|_2 = \max_{\|\boldsymbol{x}\|_2=1} \sqrt{(\boldsymbol{Ux})^{\boldsymbol{T}}(\boldsymbol{Ux})}$$

$$= \max_{\|\boldsymbol{x}\|_2=1} \sqrt{\boldsymbol{x}^{\boldsymbol{T}}\boldsymbol{x}} = \max_{\|\boldsymbol{x}\|_2=1} \|\boldsymbol{x}\|_2 = 1$$

Cost of SVD

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.



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

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: $\boldsymbol{A} = \boldsymbol{U_R \Sigma_R V_R}^{\boldsymbol{T}}$, where $\boldsymbol{U_R}$ is a $m \times k$ matrix, $\boldsymbol{\Sigma_R}$ is a $k \times k$ matrix, and $\boldsymbol{V_R}$ is a $n \times k$ matrix, and $k = \min(m, n)$.

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

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

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

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