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.

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

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

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

Reduced SVD

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

1) $\boldsymbol{m > n}$

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

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

Reduced SVD

2) $\boldsymbol{n > m}$

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

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

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.

$m > n$

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

$n > m$

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

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

In a similar way,

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.

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

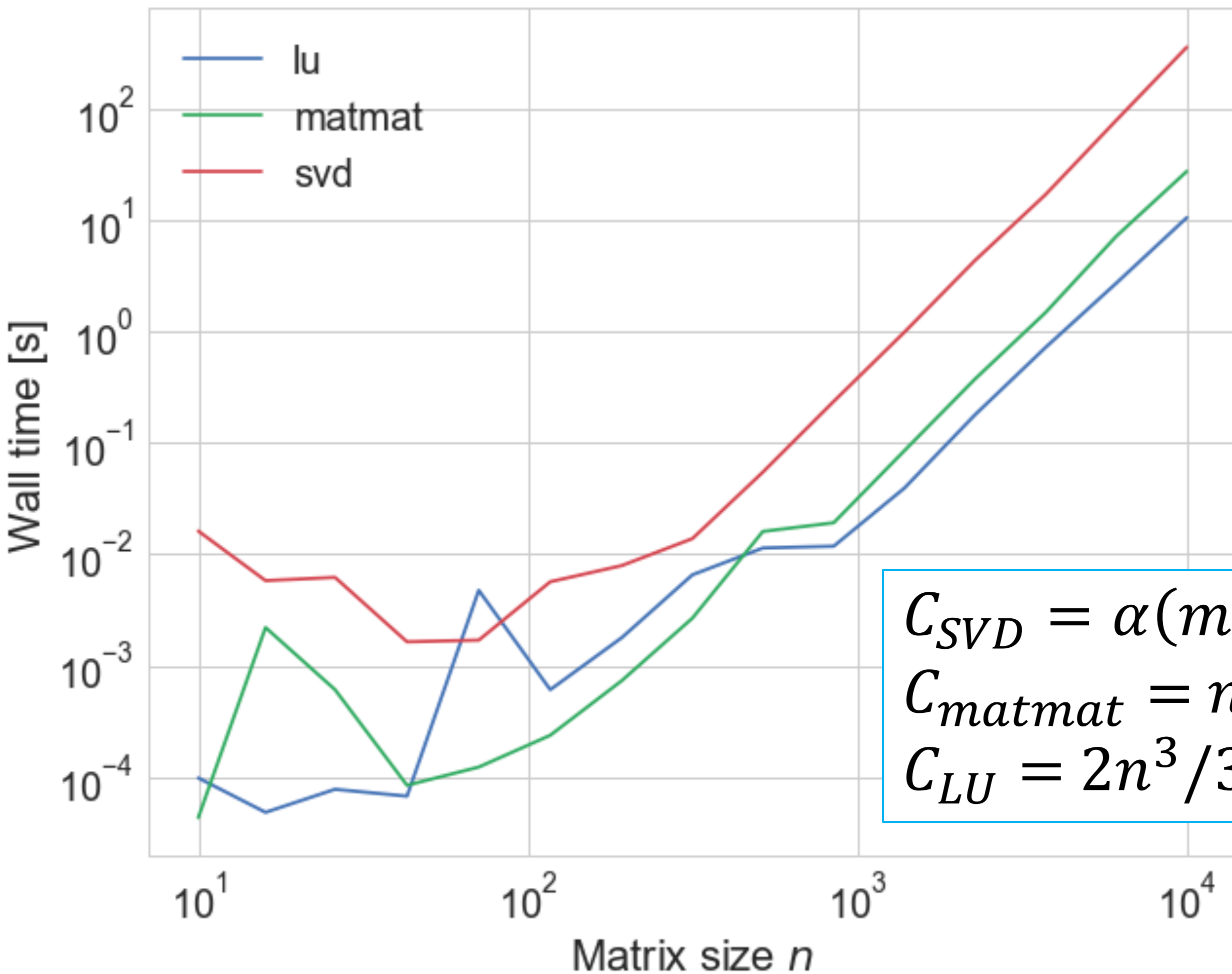
4. Find $\boldsymbol{U}$: $\boldsymbol{A} = \boldsymbol{U}\,\boldsymbol{\Sigma}\,\boldsymbol{V}^T \Longrightarrow \boldsymbol{U}\,\boldsymbol{\Sigma} = \boldsymbol{A}\,\boldsymbol{V}$. The columns are called the "left singular vectors".

Singular values are always non-negative

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

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: $\boldsymbol{A} = \boldsymbol{U_R}\boldsymbol{\Sigma_R}\boldsymbol{V_R}^{\boldsymbol{T}}$, 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 $\boldsymbol{A}^{\boldsymbol{T}}\boldsymbol{A}$, 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 $\lambda \geq 0$)