Singular Value Decomposition

Singular Value Decomposition (SVD)

- singular value decomposition of a $M \times N$ matrix $\underline{A}$ with rank R:

$$\underline{A} = \underbrace{\left(\underline{u}_1, \ldots, \underline{u}_M \right)}_{\substack{\underline{U} \\ \text{unitary} \\ M \times M \text{ matrix}}} \cdot \underbrace{\left(\begin{array}{ccc|c} \sigma_1 & & 0 & \\ & \ddots & & 0 \\ 0 & & \sigma_R & \\ \hline & 0 & & 0 \end{array} \right)}_{\substack{\Sigma \\ \text{nonnegative} \\ M \times N \text{ diagonal} \\ \text{matrix}}} \cdot \underbrace{\begin{pmatrix} \underline{v}_1^{*T} \\ \vdots \\ \underline{v}_N^{*T} \end{pmatrix}}_{\substack{\underline{V}^{*T} \\ \text{unitary} \\ N \times N \text{ matrix}}}$$

- exists for any matrix $\underline{A}$
- singular values in nonincreasing order:

$$\sigma_1 \geq \ldots \geq \sigma_R > \sigma_{R+1} = \ldots = \sigma_{\min(M,N)} = 0$$

Remarks

- The left singular vectors $\underline{u}_1, \dots, \underline{u}_M$ are the eigenvectors of the Hermitian $M \times M$ matrix

$$\underline{A} \cdot \underline{A}^{*T} = \underline{U} \cdot \Sigma \cdot \underline{V}^{*T} \cdot \underline{V} \cdot \Sigma^{*T} \cdot \underline{U}^{*T}$$

$$= \underline{U} \cdot (\Sigma \cdot \Sigma^T) \cdot \underline{U}^{*T} .$$

- The right singular vectors $\underline{v}_1, \dots, \underline{v}_N$ are the eigenvectors of the Hermitian $N \times N$ matrix

$$\underline{A}^{*T} \cdot \underline{A} = \underline{V} \cdot \Sigma^{*T} \cdot \underline{U}^{*T} \cdot \underline{U} \cdot \Sigma \cdot \underline{V}^{*T}$$

$$= \underline{V} \cdot (\Sigma^T \cdot \Sigma) \cdot \underline{V}^{*T} .$$

- The nonzero singular values $\sigma_1, \dots, \sigma_R$ are the square roots of the positive eigenvalues of $\underline{A} \cdot \underline{A}^{*T}$ or $\underline{A}^{*T} \cdot \underline{A}$.

Four Subspaces

- The left singular vectors $\underline{u}_1, \dots, \underline{u}_R$ corresponding to the nonzero singular values span the column space of dimension R.

- The remaining left singular vectors $\underline{u}_{R+1}, \dots, \underline{u}_M$ span the left nullspace of dimension $M-R$.

- The right singular vectors $\underline{v}_1, \dots, \underline{v}_R$ corresponding to the nonzero singular values span the row space of dimension R.

- The remaining right singular vectors $\underline{v}_{R+1}, \dots, \underline{v}_N$ span the nullspace of dimension $N-R$.

Frobenius Norm

- norm of a matrix:

$$\|\underline{A}\|_F = \sqrt{\sum_m \sum_n |[\underline{A}]_{m,n}|^2}$$

$$= \sqrt{\mathrm{trace}\,(\underline{A} \cdot \underline{A}^{*T})}$$

$$= \sqrt{\mathrm{trace}\,(\underline{A}^{*T} \cdot \underline{A})}$$

$$= \sqrt{\sum_r \underbrace{\sigma_r^2}}$$

- squares of the singular values of $\underline{A}$
- eigenvalues of $\underline{A} \cdot \underline{A}^{*T}$
- eigenvalues of $\underline{A}^{*T} \cdot \underline{A}$

- multiplication with a unitary matrix $\underline{U}^{*T}$ or $\underline{V}$ does not change the norm:

$$\|\underline{U}^{*T} \cdot \underline{A} \cdot \underline{V}\|_F = \|\underline{A}\|_F$$

Matrix Approximation

- Find a rank Q approximation $\underline{B}$ of $\underline{A}$ such that the error $\|\underline{A}-\underline{B}\|_F^2$ is minimized.
- singular value decomposition:

$$\underline{A} = \underbrace{\left(\overset{\substack{\text{first } Q\\ \text{columns}\\ \uparrow}}{\underline{U}_1}, \ \underline{U}_2 \right)}_{\underline{U}} \cdot \underbrace{\begin{pmatrix} \overset{\substack{Q \text{ largest}\\ \text{singular values}\\ Q\times Q\\ \uparrow}}{\Sigma_1} & 0 \\ 0 & \Sigma_2 \end{pmatrix}}_{\Sigma} \cdot \underbrace{\begin{pmatrix} \overset{\substack{\text{first } Q\\ \text{rows}\\ \uparrow}}{\underline{V}_1^{*T}} \\ \underline{V}_2^{*T} \end{pmatrix}}_{\underline{V}^{*T}}$$

$$\Rightarrow \|\underline{A}-\underline{B}\|_F^2 = \|\underline{U}^{*T}\cdot(\underline{A}-\underline{B})\cdot\underline{V}\|_F^2$$

$$= \|\Sigma - \underline{U}^{*T}\cdot\underline{B}\cdot\underline{V}\|_F^2$$

- The best rank Q approximation of the diagonal matrix

$$\Sigma = \begin{pmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{pmatrix}$$

is

$$\underline{U}^{*T}\cdot\underline{B}\cdot\underline{V} = \begin{pmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{pmatrix}.$$

$$\Rightarrow \quad \underline{B} = \underbrace{(\underline{U}_1, \underline{U}_2)}_{\underline{U}} \cdot \begin{pmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{pmatrix} \cdot \underbrace{\begin{pmatrix} \underline{V}_1^{*T} \\ \underline{V}_2^{*T} \end{pmatrix}}_{\underline{V}^{*T}}$$

$$= \underline{U}_1 \cdot \Sigma_1 \cdot \underline{V}_1^{*T}$$

"keep the Q largest singular values"

Least Squares (LS)

- singular value decomposition:

$$\underline{A} = \underline{U} \cdot \underbrace{\left(\begin{array}{ccc|c} \sigma_1 & & 0 & \\ & \ddots & & 0 \\ 0 & & \sigma_R & \\ \hline & 0 & & 0 \end{array}\right)}_{\Sigma,\ M \times N \text{ matrix}} \cdot \underline{V}^{*T}$$

- pseudoinverse of a diagonal matrix:

$$\Sigma^{+} = \left(\begin{array}{ccc|c} 1/\sigma_1 & & 0 & \\ & \ddots & & 0 \\ 0 & & 1/\sigma_R & \\ \hline & 0 & & 0 \end{array}\right) \qquad N \times M \text{ matrix}$$

- linear system of equations:

$$\underset{M\times N}{\underline{A}} \cdot \underset{N\times L}{\underline{X}} = \underset{M\times L}{\underline{B}}$$

has in general no exact solution

- least squares pseudosolution:

$$\hat{\underline{X}} = \underset{\underline{X}}{\operatorname{argmin}} \| \underline{A} \cdot \underline{X} - \underline{B} \|_F^2$$

$$= \underset{\underline{X}}{\operatorname{argmin}} \| \underline{U} \cdot \Sigma \cdot \underline{V}^{*T} \cdot \underline{X} - \underline{B} \|_F^2$$

$$= \underset{\underline{X}}{\operatorname{argmin}} \| \Sigma \cdot \underline{V}^{*T} \cdot \underline{X} - \underline{U}^{*T} \cdot \underline{B} \|_F^2$$

$$\Rightarrow \underline{V}^{*T} \cdot \hat{\underline{X}} = \Sigma^{+} \cdot \underline{U}^{*T} \cdot \underline{B}$$

$$\hat{\underline{X}} = \underbrace{\underline{V} \cdot \Sigma^{+} \cdot \underline{U}^{*T}}_{\substack{\underline{A}^{+} \\ \text{pseudoinverse}}} \cdot \underline{B}$$

Left Pseudoinverse

- full rank $M \times N$ matrix $\underline{A}$:

  $\operatorname{rank}(\underline{A}) = N \leq M$

- pseudoinverse:

$$\underline{A}^{+} = \underline{V} \cdot \Sigma^{+} \cdot \underline{U}^{*T}$$

$$= \underline{V} \cdot \underbrace{\left(\Sigma^{*T} \cdot \Sigma \right)^{-1}}_{\begin{pmatrix} 1/\sigma_1^2 & & 0 \\ & \ddots & \\ 0 & & 1/\sigma_N^2 \end{pmatrix}} \cdot \Sigma^{*T} \cdot \underline{U}^{*T}$$

$$= \underline{V} \cdot \left(\Sigma^{*T} \cdot \underline{U}^{*T} \cdot \underline{U} \cdot \Sigma \right)^{-1} \cdot \underline{V}^{*T} \cdot \underline{V} \cdot \Sigma^{*T} \cdot \underline{U}^{*T}$$

$$= \left(\underbrace{\underline{V} \cdot \Sigma^{*T} \cdot \underline{U}^{*T}}_{\underline{A}^{*T}} \cdot \underbrace{\underline{U} \cdot \Sigma \cdot \underline{V}^{*T}}_{\underline{A}} \right)^{-1} \cdot \underbrace{\underline{V} \cdot \Sigma^{*T} \cdot \underline{U}^{*T}}_{\underline{A}^{*T}}$$

$$= \left(\underbrace{\underline{A}^{*T} \cdot \underline{A}}_{\text{invertible}} \right)^{-1} \cdot \underline{A}^{*T}$$

- left pseudoinverse:

$$\underline{A}^{+} \cdot \underline{A} = \left(\underline{A}^{*T} \cdot \underline{A} \right)^{-1} \cdot \underline{A}^{*T} \cdot \underline{A} = E$$

Right Pseudoinverse

- full rank $M \times N$ matrix $\underline{A}$:

  $\mathrm{rank}(\underline{A}) = M \leq N$

- pseudoinverse:

$$\underline{A}^+ = \underline{V} \cdot \Sigma^+ \cdot \underline{U}^{*T}$$

$$= \underline{V} \cdot \Sigma^{*T} \cdot \underbrace{(\Sigma \cdot \Sigma^{*T})^{-1}}_{\begin{pmatrix} 1/\sigma_1^2 & & 0 \\ & \ddots & \\ 0 & & 1/\sigma_M^2 \end{pmatrix}} \cdot \underline{U}^{*T}$$

$$= \underline{V} \cdot \Sigma^{*T} \cdot \underline{U}^{*T} \cdot \underline{U} \cdot (\Sigma \cdot \underline{V}^{*T} \cdot \underline{V} \cdot \Sigma^{*T})^{-1} \cdot \underline{U}^{*T}$$

$$= \underbrace{\underline{V} \cdot \Sigma^{*T} \cdot \underline{U}^{*T}}_{\underline{A}^{*T}} \cdot (\underbrace{\underline{U} \cdot \Sigma \cdot \underline{V}^{*T}}_{\underline{A}} \cdot \underbrace{\underline{V} \cdot \Sigma^{*T} \cdot \underline{U}^{*T}}_{\underline{A}^{*T}})^{-1}$$

$$= \underline{A}^{*T} \cdot (\underbrace{\underline{A} \cdot \underline{A}^{*T}}_{\text{invertible}})^{-1}$$

- right pseudoinverse:

$$\underline{A} \cdot \underline{A}^+ = \underline{A} \cdot \underline{A}^{*T} \cdot (\underline{A} \cdot \underline{A}^{*T})^{-1} = E$$

Total Least Squares (TLS)

- Find matrices $\underline{\Delta}_A$ and $\underline{\Delta}_B$ of minimum square Frobenius norm $\|(\underline{\Delta}_A, \underline{\Delta}_B)\|_F^2$ such that

$$(\underline{A} + \underline{\Delta}_A) \cdot \underline{X} = (\underline{B} + \underline{\Delta}_B).$$

- reformulation of the constraint:

$$\underbrace{(\underbrace{\underline{A} + \underline{\Delta}_A}_{M \times N}, \underbrace{\underline{B} + \underline{\Delta}_B}_{M \times L})}_{\substack{\text{must have a} \\ \text{nullspace of} \\ \text{dimension } L}} \cdot \begin{pmatrix} \overset{N \times L}{\underline{X}} \\ \underset{L \times L}{-E} \end{pmatrix} = \underbrace{0}_{M \times L}$$

$\Rightarrow \operatorname{rank}(\underline{A} + \underline{\Delta}_A, \underline{B} + \underline{\Delta}_B) = N$

$\Rightarrow (\underline{A} + \underline{\Delta}_A, \underline{B} + \underline{\Delta}_B)$ is a rank N approximation of $(\underline{A}, \underline{B})$

• Using the singular value decomposition

$$\left(\underline{A}+\underline{\Delta}_A ,\ \underline{B}+\underline{\Delta}_B\right) = \underbrace{\left(\overset{\overset{\text{first } N \text{ columns}}{\uparrow}}{\underline{U}_1} ,\ \underline{U}_2\right)}_{\underline{U},\ n\times n} \cdot \underbrace{\begin{pmatrix} \overset{\overset{N\times N}{\uparrow}}{\Sigma_1} & 0 \\ 0 & 0 \end{pmatrix}}_{n\times N+L} \cdot \underbrace{\begin{pmatrix} \overset{\overset{\text{first } N \text{ rows}}{\uparrow}}{\underline{V}_1^{*T}} \\ \underline{V}_2^{*T} \end{pmatrix}}_{\underline{V},\ N+L\times N+L}$$

one obtains

$$\left(\underline{A}+\underline{\Delta}_A ,\ \underline{B}+\underline{\Delta}_B\right) = \underline{U}_1 \cdot \Sigma_1 \cdot \underline{V}_1^{*T} .$$

$$\Rightarrow \underset{\substack{\downarrow \\ \text{full rank}}}{\underline{U}_1} \cdot \Sigma_1 \cdot \underline{V}_1^{*T} \cdot \begin{pmatrix} \underline{X} \\ -E \end{pmatrix} = 0$$

$$\underline{V}_1^{*T} \cdot \begin{pmatrix} \underline{X} \\ -E \end{pmatrix} = 0$$

$\Rightarrow$ The columns of $\begin{pmatrix} \underline{X} \\ -E \end{pmatrix}$ must be orthogonal to the columns of $\underline{V}_1$, i.e., they must be linear combinations of the orthogonal complement $\underline{V}_2$:

$$\begin{pmatrix} \underline{X} \\ -E \end{pmatrix} = \underline{V}_2 \cdot \underline{G}$$

- Using

$$\underline{V} = \begin{pmatrix} \overset{N\times N}{\underline{V}_{11}} & \overset{N\times L}{\underline{V}_{12}} \\ \underset{L\times N}{\underline{V}_{21}} & \underset{L\times L}{\underline{V}_{22}} \end{pmatrix}$$

one obtains

$$\underline{V}_{22} \cdot \underline{C} = -E$$

$$\underline{C} = -\underline{V}_{22}^{-1}$$

and finally

$$\underline{X} = \underline{V}_{12} \cdot \underline{C}$$

$$= -\underline{V}_{12} \cdot \underline{V}_{22}^{-1} .$$

- $\underline{V}$ can also be computed by the eigendecomposition of

$$\begin{pmatrix} \underline{A}^{*T} \\ \underline{B}^{*T} \end{pmatrix} \cdot (\underline{A}, \underline{B}) = \underline{V} \cdot \Sigma^T \cdot \Sigma \cdot \underline{V}^{*T}$$

(instead of the singular value decomposition of $(\underline{A}, \underline{B})$).