I. Linear Algebra

I.8. Singluar Value Decomposition

Lecture based on

https://github.com/gwthomas/math4ml (Garrett Thomas, 2018)

Prof. Dr. Christoph Lippert

Digital Health & Machine Learning

Singular value decomposition (SVD) is a widely applicable tool in linear algebra.

Its strength stems partially from the fact that *every matrix* $\mathbf{A} \in \mathbb{R}^{m \times n}$ has an SVD (even non-square matrices)!

The decomposition goes as follows:

$$\mathbf{A} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^{\top}$$

where $\mathbf{U} \in \mathbb{R}^{m \times m}$ and $\mathbf{V} \in \mathbb{R}^{n \times n}$ are orthogonal matrices.

Further, $\boldsymbol{\Sigma} \in \mathbb{R}^{m \times n}$ is a diagonal matrix with the **singular values** of $\mathbf{A}$ (denoted σ_i) on its diagonal.

By convention, the singular values are given in non-increasing order, i.e.

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_{\min(m,n)} \geq 0$$

Only the first r singular values are nonzero, where r is the rank of $\mathbf{A}$.

Observe that the SVD factors provide eigendecompositions for $\mathbf{A}^\top\mathbf{A}$ and $\mathbf{A}\mathbf{A}^\top$:

$$\mathbf{A}^\top\mathbf{A} = (\mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^\top)^\top\mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^\top = \mathbf{V}\boldsymbol{\Sigma}^\top\mathbf{U}^\top\mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^\top = \mathbf{V}\boldsymbol{\Sigma}^\top\boldsymbol{\Sigma}\mathbf{V}^\top$$
$$\mathbf{A}\mathbf{A}^\top = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^\top(\mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^\top)^\top = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^\top\mathbf{V}\boldsymbol{\Sigma}^\top\mathbf{U}^\top = \mathbf{U}\boldsymbol{\Sigma}\boldsymbol{\Sigma}^\top\mathbf{U}^\top$$

It follows that the columns of $\mathbf{V}$ (the **right-singular vectors** of $\mathbf{A}$) are eigenvectors of $\mathbf{A}^\top\mathbf{A}$.

Similarly, the columns of $\mathbf{U}$ (the **left-singular vectors** of $\mathbf{A}$) are eigenvectors of $\mathbf{A}\mathbf{A}^\top$.

Note, that the matrices $\boldsymbol{\Sigma}^\top\boldsymbol{\Sigma}$ and $\boldsymbol{\Sigma}\boldsymbol{\Sigma}^\top$ are not necessarily the same size.

However, both are diagonal with the squared singular values σ_i^2 on the diagonal (plus possibly some zeros).

Thus the singular values of $\mathbf{A}$ are the square roots of the eigenvalues of $\mathbf{A}^\top\mathbf{A}$ (or equivalently, of $\mathbf{A}\mathbf{A}^\top$)[1].

[1] Recall that $\mathbf{A}^\top\mathbf{A}$ and $\mathbf{A}\mathbf{A}^\top$ are positive semi-definite, so their eigenvalues are nonnegative, and thus taking square roots is always well-defined.