Singular Value Decomposition (applications)

1) Determining the rank of a matrix

Suppose $\boldsymbol{A}$ is a $m \times n$ rectangular matrix where $m > n$:

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

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

Rank of a matrix

For general rectangular matrix $\boldsymbol{A}$ with dimensions $m \times n$, the reduced SVD is:

$$\boldsymbol{A} = \boldsymbol{U_R \Sigma_R V_R}^{\boldsymbol{T}}$$

Rank of a matrix

- The rank of **A** equals the number of non-zero singular values which is the same as the number of non-zero diagonal elements in $\boldsymbol{\Sigma}$.
- Rounding errors may lead to small but non-zero singular values in a rank deficient matrix, hence the rank of a matrix determined by the number of non-zero singular values is sometimes called "effective rank".
- The right-singular vectors (columns of $\boldsymbol{V}$) corresponding to vanishing singular values span the null space of **A**.
- The left-singular vectors (columns of $\boldsymbol{U}$) corresponding to the non-zero singular values of **A** span the range of **A**.

2) Pseudo-inverse

- **Problem:** if **A** is rank-deficient, **Σ** is not be invertible
- **How to fix it:** Define the Pseudo Inverse
- **Pseudo-Inverse of a diagonal matrix**:

$$(\boldsymbol{\Sigma}^{+})_i = \begin{cases} \frac{1}{\sigma_i}, & \text{if } \sigma_i \neq 0 \\ 0, & \text{if } \sigma_i = 0 \end{cases}$$

- **Pseudo-Inverse of a matrix $\boldsymbol{A}$**:

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

3) Matrix norms

The Euclidean norm of an orthogonal matrix is equal to 1

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

The Euclidean norm of a matrix is given by the largest singular value

$$\|\boldsymbol{A}\|_2 = \max_{\|\boldsymbol{x}\|_2=1} \|\boldsymbol{Ax}\|_2 = \max_{\|\boldsymbol{x}\|_2=1} \left\|\boldsymbol{U}\,\boldsymbol{\Sigma}\,\boldsymbol{V}^T\boldsymbol{x}\right\|_2 = \max_{\|\boldsymbol{x}\|_2=1} \left\|\boldsymbol{\Sigma}\,\boldsymbol{V}^T\boldsymbol{x}\right\|_2$$

$$= \max_{\left\|\boldsymbol{V}^T\boldsymbol{x}\right\|_2=1} \left\|\boldsymbol{\Sigma}\,\boldsymbol{V}^T\boldsymbol{x}\right\|_2 = \max_{\|\boldsymbol{y}\|_2=1} \|\boldsymbol{\Sigma}\,\boldsymbol{y}\|_2$$

Where we used the fact that $\|\boldsymbol{U}\|_2 = 1$, $\|\boldsymbol{V}\|_2 = 1$. Since $\boldsymbol{\Sigma}$ is diagonal we get:

$$\|\boldsymbol{A}\|_2 = \max(\sigma_i) = \sigma_{max}$$

σ_{max} is the largest singular value

4) Norm for the inverse of a matrix

The Euclidean norm of the inverse of a square-matrix is given by:

Assume here $\boldsymbol{A}$ is full rank, so that $\boldsymbol{A}^{-1}$ exists

$$\|\boldsymbol{A}^{-1}\|_2 = \max_{\|\boldsymbol{x}\|_2=1} \left\|(\boldsymbol{U}\,\boldsymbol{\Sigma}\,\boldsymbol{V}^{\boldsymbol{T}})^{-1}\boldsymbol{x}\right\|_2$$

$$\|\boldsymbol{A}^{-1}\|_2 = \max_{\|\boldsymbol{x}\|_2=1} \left\|\boldsymbol{V}\,\boldsymbol{\Sigma}^{-1}\boldsymbol{U}^{\boldsymbol{T}}\boldsymbol{x}\right\|_2$$

Since $\|\boldsymbol{U}\|_2 = 1$, $\|\boldsymbol{V}\|_2 = 1$ and $\boldsymbol{\Sigma}$ is diagonal then

$$\|\boldsymbol{A}^{-1}\|_2 = \frac{1}{\sigma_{min}}$$

σ_{min} is the smallest singular value

5) Norm of the pseudo-inverse matrix

The norm of the pseudo-inverse of a $m \times n$ matrix is:

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

$$\|\boldsymbol{A}^{+}\|_2 = \frac{1}{\sigma_r}$$

where σ_r is the smallest **non-zero** singular value. This is valid for any matrix, regardless of the shape or rank.

Note that for a full rank square matrix, $\|\boldsymbol{A}^{+}\|_2$ is the same as $\|\boldsymbol{A}^{-1}\|_2$.

Zero matrix: If $\boldsymbol{A}$ is a zero matrix, then $\boldsymbol{A}^{+}$ is also the zero matrix, and $\|\boldsymbol{A}^{+}\|_2 = 0$

6) Condition number of a matrix

The condition number of a matrix is given by

$$cond_2(\boldsymbol{A}) = \|\boldsymbol{A}\|_2 \, \|\boldsymbol{A}^+\|_2$$

If the matrix is full rank: $rank(\boldsymbol{A}) = min(m, n)$

$$cond_2(\boldsymbol{A}) = \frac{\sigma_{max}}{\sigma_{min}}$$

where σ_{max} is the largest singular value and σ_{min} is the smallest singular value

If the matrix is rank deficient: $rank(\boldsymbol{A}) < min(m, n)$

$$cond_2(\boldsymbol{A}) = \infty$$

7) Low-Rank Approximation

We will again use the SVD to write the matrix A as a sum of outer products (of left and right singular vectors) – here for $m > n$ without loss of generality:

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

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

$$= \sigma_1 \boldsymbol{u}_1 \mathbf{v}_1^T + \sigma_2 \boldsymbol{u}_2 \mathbf{v}_2^T + \cdots + \sigma_n \boldsymbol{u}_n \mathbf{v}_n^T$$

7) Low-Rank Approximation (cont.)

The best **rank-$\boldsymbol{k}$** approximation for a $m \times n$ matrix $\boldsymbol{A}$, (where $k \leq min(m, n)$) is the one that minimizes the following problem:

$$\min_{A_k} \|A - A_k\|$$
$$\text{such that} \quad \text{rank}(A_k) \leq k.$$

When using the induced 2-norm, the best **rank-$\boldsymbol{k}$** approximation is given by:

$$\boldsymbol{A}_k = \sigma_1 \boldsymbol{u}_1 \mathbf{v}_1^T + \sigma_2 \boldsymbol{u}_2 \mathbf{v}_2^T + \cdots + \sigma_k \boldsymbol{u}_k \mathbf{v}_k^T$$

$$\sigma_1 \geq \sigma_2 \geq \sigma_3 \ldots \geq 0$$

Note that $rank(\boldsymbol{A}) = n$ and $rank(\boldsymbol{A}_k) = k$ and the norm of the difference between the matrix and its approximation is

Example: Image compression

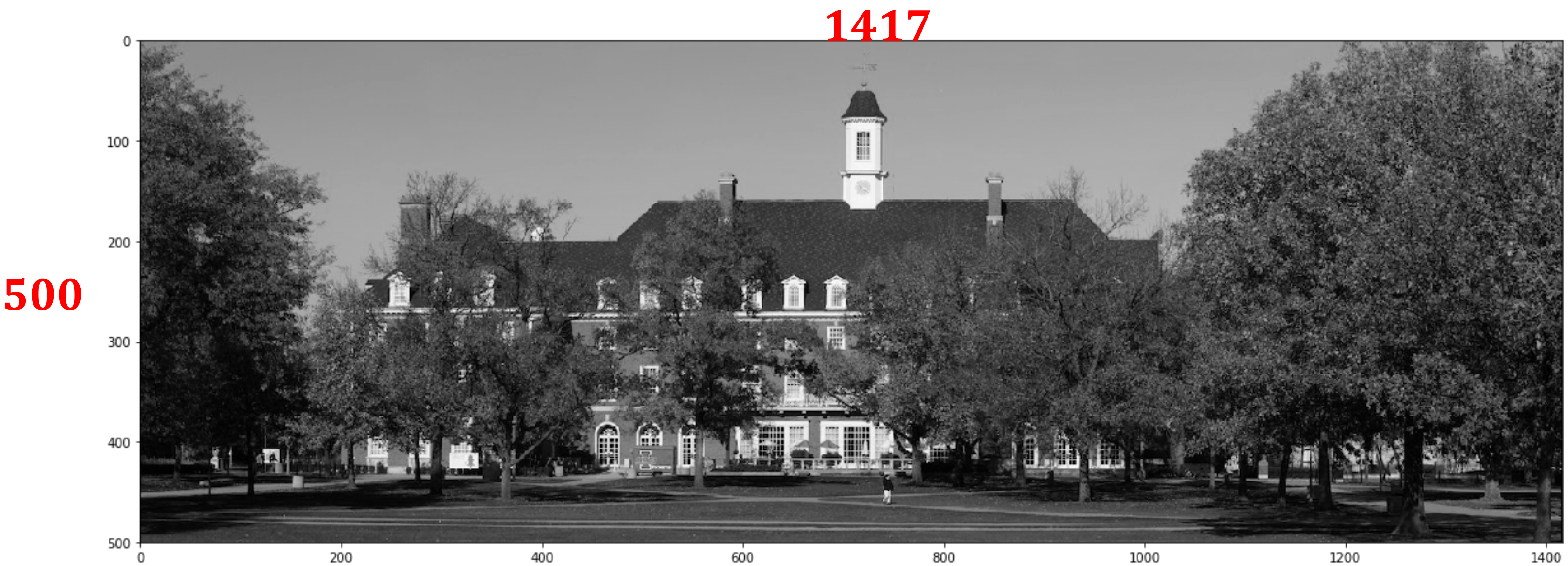

Image using rank-50 approximation

8) Using SVD to solve square system of linear equations

If $\boldsymbol{A}$ is a $n \times n$ square matrix and we want to solve $\boldsymbol{A}\,\boldsymbol{x} = \boldsymbol{b}$, we can use the SVD for $\boldsymbol{A}$ such that