Lecture 8.1: Singular Value Decomposition

Optimization and Computational Linear Algebra for Data Science

Léo Miolane

PCA

- Data matrix $\boxed{A \in \mathbb{R}^{n \times m}}$
- "Covariance matrix" $S = A^\mathsf{T} A \in \mathbb{R}^{m \times m}$.
- S is symmetric positive semi-definite.
- **Spectral Theorem:** there exists an orthonormal basis $v_1, \dots, v_m$ of $\mathbb{R}^m$ such that the v_i's are eigenvectors of S associated with the eigenvalues $\lambda_1 \geq \cdots \geq \lambda_m \geq 0$.

Exercise: Show $\operatorname{rank}(S) = \operatorname{rank}(A) \overset{(\text{def})}{=} r$

Consequence: $\begin{cases} \lambda_1 \geq \cdots \geq \lambda_r > 0 \\ \lambda_{r+1} = \cdots = \lambda_m = 0 \end{cases}$

Singular values/vectors

square roots of eigenvalues of A^TA

For $i = 1, \ldots, m$:

- we define $\sigma_i = \sqrt{\lambda_i}$, called the i^{th} singular value of A.
- we call v_j the i^{th} right singular vector of A.

What about the left singular vectors ?

We define for $i = 1, \ldots, r$

$$\boxed{\underset{\in \mathbb{R}^n}{u_i} = \frac{1}{\sigma_i} \underset{\in \mathbb{R}^n}{(A v_i)}}$$

the i^{th} left singular vector of A.

If $r < n$: we are going to add vectors $u_{r+1} \ldots u_n$ in order to get an orthonormal basis $(u_1 \ldots u_n)$ of $\mathbb{R}^n$

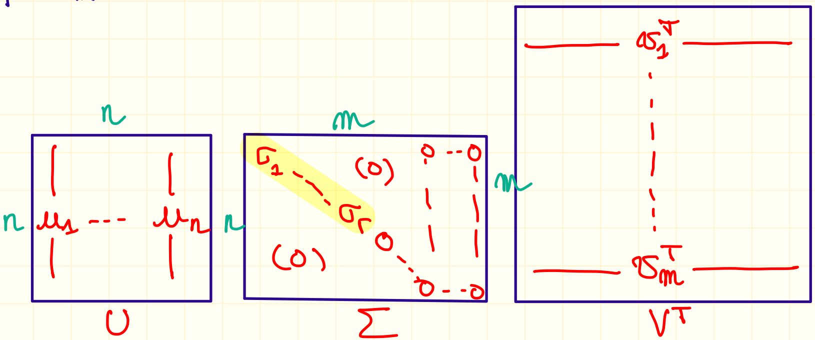
Singular Value decomposition

Theorem

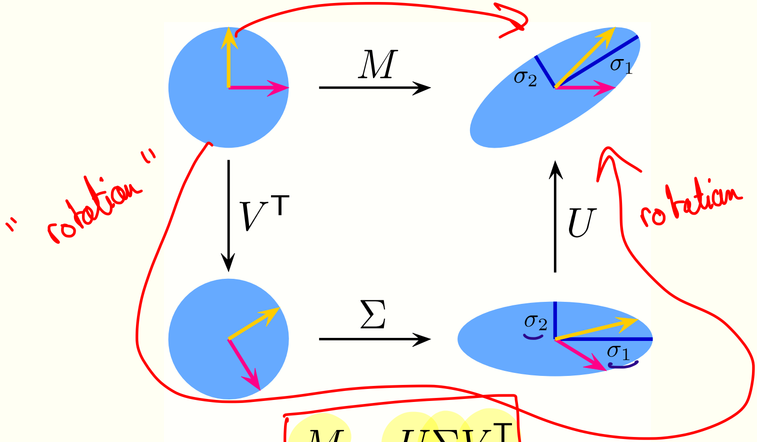
Let $A \in \mathbb{R}^{n\times m}$. Then there exists two orthogonal matrices $U \in \mathbb{R}^{n\times n}$ and $V \in \mathbb{R}^{m\times m}$ and a matrix $\Sigma \in \mathbb{R}^{n\times m}$ such that $\Sigma_{1,1} \geq \Sigma_{2,2} \geq \cdots \geq 0$ and $\boxed{\Sigma_{i,j} = 0}$ for $\boxed{i \neq j}$, that verify

$$\boxed{A = U\Sigma V^\mathsf{T}.}$$

A $n \times m$



Geometric interpretation of $U\Sigma V^\top$



$$\boxed{M = U\Sigma V^\top}$$

rotation

scaling

rotation

$$\begin{array}{cc} & m \\ n & \begin{pmatrix} 3 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \end{array}$$