A good definition of rank:

$M \in \mathbb{R}^{n \times d}$ suppose $M = \sum_{\ell=1}^{k} \boxed{x_\ell \underline{y_\ell^T}}$ $x_\ell \in \mathbb{R}^n$, $y_\ell \in \mathbb{R}^d$

$\Rightarrow \operatorname{rank}(M) \le k$

For example SVD: $M = UDV^T$ & D contains only k non-zero ~~non-singular~~ values $\operatorname{rank}(M) = k$

$$M = \sum_{\ell=1}^{k} D_{\ell\ell}\, u_\ell \underline{v_\ell^T}$$

What is SVD optimizing?

Let $M \in \mathbb{R}^{n \times d}$. Approximate M with $\hat{M}$ s.t. $\operatorname{rank}(\hat{M}) = k$.

$$\min_{\hat{M} :\, \underline{\operatorname{rank}(\hat{M}) = k}} \sum_{i,j} \underline{\left(M_{ij} - \hat{M}_{ij} \right)^2}$$

$\boxed{\text{solution}}$ $\hat{M} = U_{1:k}\, \underline{D_{1:k}}\, V_{1:k}^T$ where $M = UDV^T$

$U_{1:k}$ contains top k singular vectors in

$D_{1:k}$ top k in D

$V_{1:k}$ top k in V

This gives an "embedding"

Define $\hat{U} = U_{1:k}$ $\hat{D} = D_{1:k}$ $\hat{V} = V_{1:k}$.

$\hat{u}_i \in \mathbb{R}^k$ is k^{th} row of $\hat{U}$

$\hat{v}_i \in \mathbb{R}^k$ " " " " $\hat{V}$.

$$\approx \hat{M}_{ij} = [\hat{U} \hat{D} \hat{V}^T]_{ij} = u_i D v_j^T = \sum_{\ell=1} u_{i\ell} v_{j\ell} d_{\ell\ell}$$

$$= \langle \boxed{\hat{u}_i \hat{D}}, \boxed{\hat{v}_j} \rangle$$

- $\hat{u}_i \hat{D} \in \mathbb{R}^k$ is an embedding of the ith row.

- $\hat{v}_j \in \mathbb{R}^k$ is an embedding of the jth column.

J