21 Feb 2022 Finishing SVD

For any matrix $A \in \mathbb{R}^{m \times n}$ we found there are vectors $v_1, v_2, \ldots, v_n$ such that:

(1) The matrix $B = A^T A$ is a symmetric $n \times n$ matrix with eigenvectors $v_1, \ldots, v_n$ sorted by decreasing eigenvalue.

(2) For $1 \le k \le n$ the column space of the matrix $V_k = \begin{bmatrix} | & | & \cdots & | \\ v_1 & v_2 & \cdots & v_k \\ | & | & \cdots & | \end{bmatrix}$ is the best k-dimensional linear subspace of $\mathbb{R}^n$ approximating the rows of A.

If $\hat{A}$ is a matrix whose rows belong to a k-dimensional linear subspace of $\mathbb{R}^n$,
$\Longleftrightarrow$ $\text{rank}(\hat{A}) = k$.

In particular, if $\hat{a}_i^T$ denotes projection of a_i^T onto the best-fit k-dimensional subspace, and let

$$\hat{A}_k = \begin{bmatrix} — \hat{a}_1^T — \\ — \hat{a}_2^T — \\ \vdots \\ — \hat{a}_m^T — \end{bmatrix}$$

Then the $\sum_{i=1}^{m} \|a_i - \hat{a}_i\|^2$ objective can be written as

$$\|A - \hat{A}_k\|_F^2$$

So the $\hat{A}_k$ matrix defined above, whose rows belong to span of $v_1^T, \ldots, v_k^T$, is the rank-k matrix that minimizes $\|A - \hat{A}_k\|_F^2$.

The coefficients that express $\hat{a}_i$ as a linear combo of $v_1, \ldots, v_k$.

Suppose $\forall i \quad \hat{a}_i^T = \sum \tilde{u}_{ij} v_j^T$

Then $\hat{A}_k = \tilde{U} V_k^T$ $\qquad \tilde{U} = \begin{bmatrix} & \tilde{u}_{ij} & \end{bmatrix}$

Write $\tilde{U} = U_k D_k$

diagonal matrix with $(D_k)_{ii} = \|u_i\|_2$.

columns have $\|u\|_2 = 1$

(i.e. rescale each column of $\tilde{U}$ to have length 1.)

$$\hat{A}_k = U_k D_k V_k^T.$$

Let $k = n$. $\hat{A}_n = A$, $\qquad A = U_n D_n V_n^T$

$$= \boxed{UDV^T}$$

Columns of U: left singular vectors

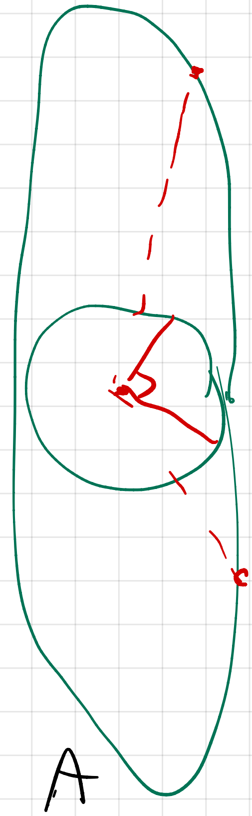
Diag entries of D: singular values

$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n \geq 0$

singular value decomp. of A

$\sigma_1^2, \ldots, \sigma_n^2$ are eigvals of A^TA

Columns of V: right singular vectors.



The equation $A = UDV^T$

means

$$a_{ij} = \sum_k u_{ik} \underbrace{d_{kk}}_{\sigma_k} v_{jk}$$

$$= \sum_k \sigma_k u_{ik} v_{jk}$$

Summarize as

$$A = \sum_{k=1}^{\min\{m,n\}} \sigma_k u_k v_k^T$$

We derived that best rank-r approx to A in Frobenius norm is

$$\hat{A}_r = \sum_{k=1}^{r} \sigma_k u_k v_k^T$$

Best rank-r approx to A^T is $\sum_{k=1}^{r} \sigma_k v_k u_k^T$

To find eigenvectors of $B = A^TA$, do "power iteration": start with random vector v, compute $B^k v$ for some large k, it will be nearly parallel to v_1.

If $v = c_1 v_1 + c_2 v_2 + \cdots + c_n v_n$

$$B^k v = c_1 (B^k v_1) + \cdots + c_n (B^k v_n)$$

$$= c_1 \sigma_1^{2k} v_1 + c_2 \sigma_2^{2k} v_2 + \cdots + c_n \sigma_n^{2k} v_n .$$