Lecture 26. Properties of SVD.

SVD. Given $A \in \mathbb{R}^{m\times n}$, there are orthogonal matrices $U \in \mathbb{R}^{m\times m}$, $V \in \mathbb{R}^{n\times n}$ s.t.

$$AV = U\Sigma, \quad \Sigma = \left(\begin{array}{ccc|c} \sigma_1 & & & 0 \\ & \ddots & & \\ & & \sigma_r & \\ \hline & 0 & & 0 \end{array}\right) \in \mathbb{R}^{m\times n}$$

$\sigma_1 \geq \cdots \geq \sigma_r > 0.$

1. Let $m \geq n$. $\Sigma = \left(\begin{array}{c} \begin{matrix} \sigma_1 & & & & \\ & \ddots & & & \\ & & \sigma_r & & \\ & & & 0 & \\ & & & & \ddots \; 0 \end{matrix} \\ \hline 0 \end{array}\right)$ $\sigma_1, \dots, \sigma_r, \underbrace{0 \cdots 0}_{n-r}$

are the singular values of A.

Let $m \leq n$, $\Sigma = \left(\begin{array}{c|c} \underbrace{\begin{matrix} \sigma_1 & & & & \\ & \ddots & & & \\ & & \sigma_r & & \\ & & & 0 & \\ & & & & \ddots \; 0 \end{matrix}} & 0 \end{array}\right)$

$\sigma_1, \dots, \sigma_r, \underbrace{0 \cdots 0}_{m-r}$ are the singular values of A.

In general, $\sigma_1, \cdots, \sigma_r, \sigma_{r+1}, \cdots, \sigma_{\min\{m,n\}}$

are the singular values of A.

2. let the $V = [v_1, \cdots, v_n]$, $U = [u_1, \cdots, u_m]$

Compare the i-th columns of both sides of

$$(*) \quad AV = U\Sigma$$

we obtain $\quad Av_i = \sigma_i u_i$

Taking transpose of both sides of $(*)$, we

obtain

$$(**) \quad A^T U = V\Sigma^T.$$

Compare the i-th columns of both sides of $(**)$

we obtain $\quad A^T u_i = \sigma_i v_i$

(σ_i, u_i, v_i) is called the i-th, $i=1,\cdots,\min\{n,m\}$

singular value ~~triplet~~ of A.

3. SVD and spectral decomposition.

from $AV = U\Sigma$, $A = U\Sigma V^T$.

we obtain $A^TA = (U\Sigma V^T)^T (U\Sigma V^T)$

$$= V\Sigma^T U^T U \Sigma V^T$$

$$= V \Sigma^T \Sigma V^T = V \begin{pmatrix} \sigma_1^2 & & & & \\ & \ddots & & & \\ & & \sigma_r^2 & & \\ & & & 0 & \\ & & & & \ddots \\ & & & & & 0 \end{pmatrix} V^T$$

the spectral decomposition of A^TA.

i.e., V are the eigenvectors of A^TA,

A^TA eigenvalues are

$$\sigma_1^2 \cdots \sigma_r^2, \underbrace{0, \cdots, 0}_{n-r}$$

Similarly, $AA^T = U \begin{pmatrix} \sigma_1^2 & & & & & \\ & \ddots & & & & \\ & & \sigma_r^2 & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix} U^T$

U are the eigenvectors of AA^T, and its eigenvalues are

$$\sigma_1^2, \cdots, \sigma_r^2, \underbrace{0, \cdots, 0}_{m-r}$$

---

4. Four fundamental subspaces of $A \in \mathbb{R}^{m \times n}$

$r = r(A)$, $U = [u_1, \cdots, u_m]$, $V = [v_1, \cdots, v_n]$

$C(A) = \text{Span}\{u_1, \cdots, u_r\}$

$N(A) = \text{Span}\{v_{r+1}, \cdots, v_n\}$

$C(A^T) = \text{Span}\{v_1, \cdots, v_r\}$

$N(A^T) = \text{Span}\{u_{r+1}, \cdots, u_m\}$

5. linear system $Ax=b$ and SVD.

plug in $A = U\Sigma V^T$ into $Ax=b$.

$$U\Sigma V^T x = b \Rightarrow \Sigma V^T x = U^T b. \quad (*)$$

let $V^T x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$; $U^T b = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \in \mathbb{R}^m$

(*) implies
$$\left(\begin{array}{ccc|c} \sigma_1 & & & \\ & \ddots & & 0 \\ & & \sigma_r & \\ \hline & 0 & & 0 \end{array}\right) \left(\begin{array}{c} x_1 \\ \vdots \\ x_r \\ \hline x_{r+1} \\ \vdots \\ x_n \end{array}\right) = \left(\begin{array}{c} b_1 \\ \vdots \\ b_r \\ \hline b_{r+1} \\ \vdots \\ b_m \end{array}\right)$$

$Ax=b$ has a solution iff $b_{r+1} = \cdots = b_m = 0$.

Otherwise, let's look at least squares problem

$$\|b - Ax\|^2 = \|U^T(b - Ax)\|^2$$

$$= \left\| \left(\begin{array}{c} b_1 \\ \vdots \\ b_r \\ \hline b_{r+1} \\ \vdots \\ b_m \end{array}\right) - \left(\begin{array}{ccc|c} \sigma_1 & & & \\ & \ddots & & 0 \\ & & \sigma_r & \\ \hline & 0 & & 0 \end{array}\right) \left(\begin{array}{c} x_1 \\ \vdots \\ x_r \\ \hline x_{r+1} \\ \vdots \\ x_n \end{array}\right) \right\|^2$$

$$= \left\| \begin{pmatrix} b_1 \\ \vdots \\ b_r \end{pmatrix} - \begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_r \end{pmatrix} \right\|^2 + \left\| \begin{pmatrix} b_{r+1} \\ \vdots \\ b_m \end{pmatrix} \right\|^2$$

the above is minimized if. $x_i = \frac{b_i}{\sigma_i}$, $i = 1, \cdots, r$.

A general solution for the least squares problem

$$x = V \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = V \begin{pmatrix} \frac{b_1}{\sigma_1} \\ \vdots \\ \frac{b_r}{\sigma_r} \\ \hline x_{r+1} \\ \vdots \\ x_n \end{pmatrix}$$

$$= \left(\frac{b_1}{\sigma_1}\right) v_1 + \cdots + \left(\frac{b_r}{\sigma_r}\right) v_r + x_{r+1} v_{r+1} + \cdots + x_n v_n$$

where $x_{r+1}, \cdots, x_n$ arbitrary.

Or using subspace notation

$$x \in \underbrace{\left(\frac{b_1}{\sigma_1}\right) v_1 + \cdots + \left(\frac{b_r}{\sigma_r}\right) v_r}_{\text{a special solution}} + \underbrace{\mathrm{Span}\{v_{r+1}, \cdots v_n\}}_{N(A)}$$

6. for symmetric $A = A^T \in \mathbb{R}^{n\times n}$

for $x \neq 0$. $R(x) = \dfrac{x^T A x}{x^T x}$ is called a Rayleigh quotient.

In your homework VI, you showed.

$$\lambda_{min} \leq R(x) \leq \lambda_{max}$$

where λ_{min}, and λ_{max} are the smallest and largest eigenvalues of A, respectively.

Now consider A^TA. $x^T A^T A x = \|Ax\|^2$

and $$\frac{x^T A^T A x}{x^T x} = \left(\frac{\|Ax\|}{\|x\|}\right)^2$$

eigenvalues of A^TA are $\sigma_1^2, \cdots \sigma_r^2 \cdot 0, \cdots 0,$

Hence $$\sigma_{min} \leq \frac{\|Ax\|}{\|x\|} \leq \sigma_{max}$$

where σ_{max}, and σ_{min} are the smallest and largest singular values of A, respectively.

$$\sigma_{max}(A) = \max_{x \neq 0} \frac{\|Ax\|}{\|x\|}$$ is defined to be the spectral norm (or 2-norm) of A. $\sigma_{max}(A) = \|A\|_2$

There's another norm of $A \in \mathbb{R}^{m \times n}$

$$\|A\|_F^2 = \sum_{i,j} a_{ij}^2$$ (Sum of squares of all elements of A)

1) $\|A\|_F^2 = \text{trace}(A^T A) = \text{trace}(AA^T)$

$$= \sum_{i=1}^{\min(m,n)} \sigma_i^2 \geq \|A\|_2^2 = \sigma_{max}^2(A)$$

7. $A = \sigma_1 u_1 v_1^T + \cdots + \sigma_r u_r v_r^T$

decomposition as a sum of r rank-1 matrices

<u>proof.</u> $A = U\Sigma V^T$

$$= [\sigma_1 u_1, \cdots, \sigma_r u_r, 0 \cdots 0]\, V^T$$

$$= \sigma_1 u_1 v_1^T + \cdots + \sigma_r u_r v_r^T$$

8. for $k \leq r = \mathrm{rank}(A)$ let.

$$A_k = \sigma_1 u_1 v_1^T + \cdots + \sigma_k u_k v_k^T$$

Then 1) $\mathrm{rank}(A_k) = k$.

$$\|A - A_k\| = \min_{\mathrm{rank}(B) \leq k} \|A - B\|$$

The norm $\|\cdot\|$ can be $\|\cdot\|_2$, $\|\cdot\|_F$.

or ~~Spectr~~ Nuclear Norm $\|A\|_N = \sigma_1 + \cdots + \sigma_r$