Lecture 17 : Rapidly Decreasing Singular Values

$X = n \times n$ real matrix

singular values $\sigma_1(X) \geq \sigma_2(X) \geq \cdots \geq \sigma_n(X) \geq 0$

Fact: If $\sigma_{k+1}(X) = 0$, $\sigma_k(X) > 0$:

1. $\text{rank}(X) = k$
2. $X = u_1 v_1^T + \cdots + u_k v_k^T$

   $+ \cdots +$
3. $\dim(\text{col. space}) = \dim(\text{row space}) = k$

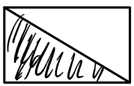
Def: X = low rank, if $2kn < n^2 \Rightarrow k < n/2$

Often, we demand $k \ll n/2$

What do low rank matrices look like?

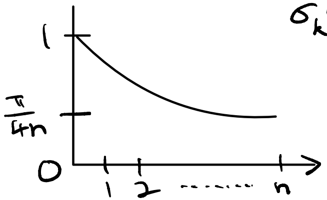
rank-1 $\Rightarrow$ matrix highly aligned

Triangular flag



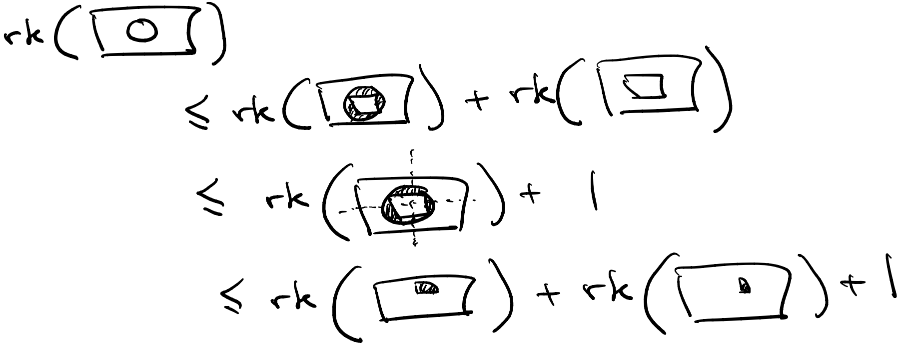
$$= \begin{bmatrix} 1 & & 0 \\ \vdots & \ddots & \\ 1 & \cdots & 1 \end{bmatrix} = X, \quad X^{-1} = \begin{bmatrix} 1 & & \\ -1 & \ddots & \\ & \ddots & \ddots \\ & & -1 \quad 1 \end{bmatrix}, \quad (X^T X)^{-1} = \begin{bmatrix} 1 & -1 & & \\ -1 & 2 & \ddots & \\ & \ddots & \ddots & -1 \\ & & -1 & 2 \end{bmatrix}$$

$$\Rightarrow \sigma_k(X) = \left[2 \sin\left(\frac{\pi(2k-1)}{2(2n+1)} \right) \right] \qquad \sigma_1(X) \approx \frac{2n}{\pi}, \quad \sigma_n(X) \approx \frac{1}{2}$$



$\Rightarrow$ Hence, triangular numbers are bad for low rank.

Japan Flag



$$\leq \lceil r(1-\sqrt{2}/2)\rceil + \lceil r(1-\sqrt{2}/2)\rceil + 1 = \frac{1}{2}r + 1$$

$\Rightarrow$ low rank matrix

Numerical rank

For $0 < \varepsilon < 1$ (ε = tolerance)

$\text{rank}_\varepsilon(x) = k$ $\quad \sigma_{k+1}(x) \leq \varepsilon\, \sigma_1(x)$

$\quad \sigma_k(x) > \varepsilon\, \sigma_1(x)$

$\text{rank}_0(x) = \text{rank}(x)$

when $\varepsilon = 0$

Eckart-Young

$$\sigma_{k+1}(x) = \|x - x_k\|_2$$

x_k = best rank-k

Matrices of numerical low rank

- All low rank matrices

$$H_{jk} = \frac{1}{j+k-1}, \quad V = \begin{bmatrix} 1 & x_1 & \cdots & x_1^{n-1} \\ \vdots & \vdots & & \vdots \\ 1 & x_n & & x_n^{n-1} \end{bmatrix} \quad x_i\text{'s real}$$

(Hilbert) Vandermonde

often, we want V^{-1}, but difficult because V has low numerical rank.

Why so many low rank matrices in the world?

Reason 1: The world is smooth (Reade, 1983)

Ex. $p(x,y) = 1 + x + xy$

$$x_{jk} = p(j,k) = 1 + j + jk$$

$$x = \begin{bmatrix} 1 & \cdots & 1 \\ \vdots & & \vdots \\ 1 & \cdots & 1 \end{bmatrix} + \begin{bmatrix} 1 & \cdots & 1 \\ 2 & \cdots & 2 \\ \vdots & & \vdots \\ n & \cdots & n \end{bmatrix} + \begin{bmatrix} 1 & 2 & \cdots & n \\ 2 & 4 & \cdots & 2n \\ \vdots & \vdots & \ddots & \vdots \\ n & 2n & \cdots & n^2 \end{bmatrix}$$

All rank-1 matrices

$\text{rank}(x) \leq 3$

In general, $p(x,y) = \sum_{s=0}^{m-1} \sum_{t=0}^{m-1} a_{st} x^s y^t$

$$x_{jk} = p(j,k) \quad , \text{rank}(x) \le m^2$$

If $x_{jk} = \frac{1}{j+k-1}$, $f(x,y) = \frac{1}{x+y-1}$ $\Rightarrow$ Smooth function like this can be well approximated by polynomials, therefore have low numerical rank.

Find p $\quad |f(x,y) - \underbrace{p(x,y)}_{\text{approx. of } f}| \le \varepsilon/n \, \|x\|_2$

$$Y_{jk} = p(j,k) = \text{finite rank}$$

$$\|x - Y\|_2 \le \varepsilon \|x\|_2$$

$\text{rank}(H_{1000}) = 1000$, $\text{rank}_\varepsilon(H_{1000}) = 28$ $\quad \varepsilon = 10^{-15}$

But, Reade's argument gets $\text{rank}_\varepsilon(H_{1000}) \le 719$

Another Reason : The world is Sylvester, matrices satisfy $AX - XB = C$ for some A, B, C

Thm : If X satisfies $AX - XB = C$ (and A = normal, B = normal), then

$$\sigma_{rk+1}(X) \le \underbrace{Z_k(E,F)}_{\text{Zolotarev no.}} \sigma_1(X)$$

$$\text{rank}(C) = r$$

E = set eigs of A

F = set eigs of B