Welcome to Ma221! Lecture 7, Sep 8

SVD: $A^{n\times n} = \underset{\text{orthog}}{U}\ \overset{\text{diag}}{\Sigma}\ \underset{\text{orthog}}{V^T}$

"thin SVD" $A^{m\times n}$, $m>n$

$$A = \begin{bmatrix} \ \\ \ \\ \ \end{bmatrix} = \begin{array}{c} \\ m \end{array}\overset{\ \ n \quad m-n}{\begin{bmatrix} U_1 \mid U_2 \end{bmatrix}} \begin{array}{c} n \\ m-n \end{array}\begin{bmatrix} \Sigma \\ 0 \end{bmatrix} \ n\overset{n}{\begin{bmatrix} V^T \end{bmatrix}}$$

$$= \begin{bmatrix} U_1 \end{bmatrix} \begin{bmatrix} \Sigma \end{bmatrix} \begin{bmatrix} V^T \end{bmatrix}$$

Fact 1: $A^{n\times n}$ nonsing.

can solve $Ax=b$ in $O(n^2)$ ops, given SVD

$$A^{-1}b = (U\Sigma V^T)^{-1} b$$
$$= V(\Sigma^{-1}(U^T b))$$

Gaussian Elim cheaper

SVD gives "free" error bound

Fact 2: $m>n$ solve $\underset{x}{\operatorname{argmin}} \|Ax-b\|_2$

$x = V\Sigma^{-1}U^T b$ if A full rank

generalizes "A^{-1}" to rectangular matrices using "thin SVD" proof: $A = \hat{U}\hat{\Sigma}V^T$ $\quad \hat{U} = \begin{array}{c}\left[\begin{array}{c|c} U & U' \end{array}\right]^{m\times m} \\ \begin{array}{cc} n & m-n \end{array}\end{array}$

$$\hat{\Sigma} = \begin{array}{c} n \\ \left[\begin{array}{c} \Sigma \\ \hline 0 \end{array}\right] \end{array} \begin{array}{l} n \\ m-n \end{array}$$

$$\|Ax-b\|_2^2 = \|\hat{U}\hat{\Sigma}V^Tx - b\|_2^2$$

$$= \|\hat{U}^T(\quad " \quad)\|_2^2$$

$$= \|\hat{\Sigma}V^Tx - \hat{U}^Tb\|_2^2$$

$$= \left\|\begin{bmatrix}\Sigma \\ 0\end{bmatrix}V^Tx - \begin{bmatrix}U^T \\ U'^T\end{bmatrix}b\right\|_2^2$$

$$= \left\|\begin{bmatrix}\Sigma V^Tx - U^Tb \\ 0 \quad -U'^Tb\end{bmatrix}\right\|_2^2$$

$$= \|\Sigma V^Tx - U^Tb\|_2^2 + \|U'^Tb\|_2^2$$

$$= 0 + \|U'^Tb\|_2^2$$

if $x = V\Sigma^{-1}U^Tb$ QED

Def: $A^{m\times n} = U_{m\times n}\Sigma V^T_{n\times n}$ full rank

$A^+ = V\Sigma^{-1}U^T$ is

Moore-Penrose pseudo inverse

pinv(A) in Matlab

All extends (LS and pinv) to

underdetermined case ($m<n$)

rank deficient

regularized LS: $\operatorname{argmin}_x \|Ax-b\|_2^2 + \lambda^2\|x\|_2^2$

Q3.13 has more properties of A^+

Fact 3: $A = A^T$ real, with

evals $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$

orthogonal evecs $V = [v(1), \ldots, v(n)]$

$A = V \Lambda V^T \quad (AV = V\Lambda) \quad Av(i) = \lambda_i v(i)$

$= \text{SVD}$ if all $\lambda_i \geq 0$

otherwise: $A = (VD)(D\Lambda)V^T = \text{SVD}$

$D = \mathrm{diag}(\mathrm{sign}(\lambda_i))$

Fact 4: using SVD:

$\square = A^T A = (U\Sigma V^T)^T (U \Sigma V^T)$ only need thin SVD

$= V \Sigma^T \underbrace{U^T U}_{I} \Sigma V^T$

$= V \Sigma^T \Sigma V^T = \square \diagdown \square$

$=$ eigen decomp of $A^T A$

Fact 5: $A A^T = (U\Sigma V^T)(U\Sigma V^T)^T$

$= U \Sigma \underbrace{V^T V}_{I} \Sigma^T U^T$

$= U \Sigma \Sigma^T U^T = \underbrace{\begin{bmatrix} U & U' \end{bmatrix}}_{\text{orthog}} \begin{bmatrix} \diagdown & 0 \\ 0 & 0 \end{bmatrix} \underbrace{\begin{bmatrix} U^T \\ U'^T \end{bmatrix}}_{\text{orthog}}$

Fact 6: $H = \begin{bmatrix} 0 & A^T \\ A & 0 \end{bmatrix} = H^T$ $m+n \times m+n$

H has evals $\pm \sigma_i$ of A

evecs $\frac{1}{\sqrt{2}} \begin{bmatrix} v(i) \\ \pm u(i) \end{bmatrix}$

$\Rightarrow$ algs for sym eigenproblem closely related to eigs for SVD

Fact 7: $\|A\|_2 = \|U\Sigma V^T\|_2 = \|\Sigma\|_2 = \sigma_1$

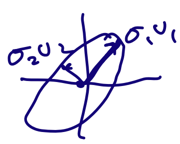
$\|A^{-1}\|_2 = \|V\Sigma^{-1}U^T\|_2 = \|\Sigma^{-1}\|_2 = \frac{1}{\sigma_n}$

$\sigma_1 \ge \cdots \ge \sigma_n > 0$

Def: $\kappa(A) = \frac{\sigma_1}{\sigma_n}$ = condition number of A

Fact 8: Let S be the unit sphere in $\mathbb{R}^n$

then $A \cdot S$ is an ellipsoid centered at 0



with principal axes u_i, length σ_i

proof: $s = [s_1, \dots, s_n]$ $\|s\|_2 = 1$

$$As = U\Sigma \underbrace{V^T s}_{\text{unit}} = U\Sigma\hat{s} = \sum_i u_i(\sigma_i \hat{s}_i)$$

$$u_1 = \begin{bmatrix}1\\0\end{bmatrix} \quad u_2 = \begin{bmatrix}0\\1\end{bmatrix} \quad As = \begin{bmatrix}\sigma_1\hat{s}_1\\ \sigma_2\hat{s}_2\end{bmatrix} = \begin{bmatrix}x\\y\end{bmatrix}$$

$$\begin{bmatrix}0 & 100\\ 0 & 0\end{bmatrix}$$

$$\left(\frac{x}{\sigma_1}\right)^2 + \left(\frac{y}{\sigma_2}\right)^2 = \hat{s}_1^2 + \hat{s}_2^2 = 1$$

Fact 9: Suppose

$\sigma_1 \ge \sigma_2 \ge \cdots \ge \sigma_r > 0 = \sigma_{r+1} = \cdots = \sigma_n$

$\Rightarrow$ rank$(A) = r$

null space = span$(v_{r+1}, v_{r+2}, \dots, v_n)$

range space = span$(u_1, \dots, u_r)$

$$Ax = U\Sigma V^T x = \sum_{i=1}^{r} u_i \sigma_i \underline{(v_i^T x)}$$

Fact 10: Matrix A_k of rank k closest to A in 2-norm is

$$A_k = \sum_{i=1}^{k} u_i \sigma_i v_i^T = U \Sigma_k V^T$$

$$\Sigma_k = \begin{bmatrix} \sigma_1 & & & & & \\ & \ddots & & & & \\ & & \sigma_k & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{bmatrix}$$

$$\|A_k - A\|_2 = \sigma_{k+1}$$

In particular, closest non-full rank matrix to A is at distance $\sigma_n = \sigma_{min}$

recall $\|A^{-1}\|_2 = \frac{1}{\sigma_{min}}$

proof: Easy: A_k has rank k ✓

and right distance from A:

$$\|A - A_k\|_2 = \| \underbrace{\sum_{i=k+1}^{n} u_i \sigma_i v_i^T}_{\text{SVD of } A - A_k} \| = \sigma_{k+1}$$

Suppose B has rank k,

need to show $\|A - B\|_2 \geq \sigma_{k+1}$

nullspace of B has dimension $n-k$

space spanned by $\{v_1, \ldots, v_{k+1}\}$ has dim $k+1$

$$\underbrace{\text{nullspace}(B)}_{\dim = n-k} \cap \underbrace{\text{span}\{v_1, \ldots, v_{k+1}\}}_{\dim = k+1} \ni h \neq 0 \qquad n-k + k+1 = n+1$$

these two spaces intersect, $\|h\|_2 = 1$

$$\|A - B\|_2 \geq \|(A-B)h\|_2 = \|Ah\|_2 \quad \text{since } Bh = 0$$

$$= \| U \Sigma \underbrace{V^T h}_{x} \|_2 \qquad x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{k+1} \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

$$= \|\Sigma x\|_2 = \left\| \begin{bmatrix} \sigma_1 x_1 \\ \sigma_2 x_2 \\ \vdots \\ \sigma_{k+1} x_{k+1} \\ 0 \\ \vdots \\ 0 \end{bmatrix} \right\|_2 \geq \sigma_{k+1} \|x\|_2 = \sigma_{k+1}$$

QED

Start using SVD, norms to analyze condition number for A^{-1} and for solving $Ax = b$.
if A (and b) change "a little" how much can A^{-1} (and $A^{-1}b$) change?

If $|x| < 1$, $\frac{1}{1-x} = 1 + x + x^2 + \cdots$

Generalize to matrices:

Lemma: If $\|X\| < 1$ any operator norm then $I - X$ nonsingular:

$$(I-X)^{-1} = \sum_{i=0}^{\infty} X^i \qquad \|(I-X)^{-1}\| \leq \frac{1}{1-\|X\|}$$