Matrix Theory, Math6304
Lecture Notes from November 13, 2012

taken by Tasadduk Chowdhury

Last Time:

1. More on least squares.

2. Abbreviated SVD.

3. Minimizing norm among solutions of normal equations

$$A^*Ax = A^*b$$
$$\hat{x} = W_r\Sigma_r^{-1}V_r^*b.$$

From last time,

$$\varphi(x) = \|b - Ax\|^2 = \|V_r^*b - \Sigma_r W_r^* x\|^2 + \|V'^*b\|^2 \geq \|V'^*b\|^2.$$

To see that we can achieve equality by minimizing φ, recall minimizers, observe $A^*Ax = A^*b$, in terms of abbreviated SVD,

$$W_r\Sigma_r V_r^* V_r \Sigma_r W_r^* x = W_r \Sigma_r V_r^* b.$$

So we have $W_r\Sigma_r^2 W_r^* x = W_r\Sigma_r V_r^* b$. Using that $W_r^*W_r = I$, we get

$$\Sigma_r^2 W_r^* x = \Sigma_r V_r^* b. \tag{1}$$

So $\|V_r^*b - \Sigma_r W_r^* x\| = 0$ if and only if x solves $A^*Ax = A^*b$.

Next, compute the norm of x which solves (1),

$$\begin{aligned}\|x\|^2 &= \left\| \begin{bmatrix} W_r^* \\ W'^* \end{bmatrix} x \right\|^2 \\ &= \|W_r^*x\|^2 + \|W'^*x\|^2 \\ &= \|\Sigma_r^{-1}V_r^*b\|^2 + \|W'^*x\|^2 \geq \|\Sigma_r^{-1}V_r^*b\|^2,\end{aligned}$$

and equality holds iff $W'^*x = 0$. If this is the case, then

$$\begin{aligned}W^*x = \begin{bmatrix} W_r^* \\ W'^* \end{bmatrix} x &= \begin{bmatrix} \Sigma_r^{-1}V_r^*b \\ 0 \end{bmatrix} \\ x &= \begin{bmatrix} W_r & W' \end{bmatrix} \begin{bmatrix} \Sigma_r^{-1}V_r^*b \\ 0 \end{bmatrix} \\ &= W_r\Sigma_r^{-1}V_r^*b.\end{aligned}$$

We see that the minimizer $\hat{x}$ for the norm among all solutions to the normal equations is characterized by

$$\hat{x} = W_r \Sigma_r^{-1} V_r^* b.$$

Remark

- We conclude there is a linear map $b \mapsto \hat{x} \equiv A^\dagger b$, which gives this unique minimum norm least-squares solution.
- If A is invertible, then $A^\dagger = W\Sigma^{-1}V^*$, or equivalently $A^\dagger = A^{-1}$. For this reason, $A^\dagger$ is called the *pseudo-inverse* of A.

5 Matrix Norms and Spectral Radius

5.1 From inner product to Euclidean norm

Recall that inner product $\langle \cdot, \cdot \rangle$ on $\mathbb{C}^n$ induces Euclidean norm $\|\cdot\|$ by $\|x\| = \sqrt{\langle x, x\rangle}$, for $x \in \mathbb{C}^n$. The defining properties of a norm are positive-definiteness, homogeneity, and triangle inequality.

We can get a norm on matrices from an inner product.

5.1.1 Definition. The *Hilbert-Schmidt inner product* or *Frobenius inner product* for $A, B \in M_n$ is

$$\langle A, B\rangle = \sum_{i,j=1}^{n} a_{i,j}\overline{b_{i,j}} = \mathrm{trace}[AB^*],$$

with $A = [a_{i,j}]_{i,j=1}^n, B = [b_{i,j}]_{i,j=1}^n$. This inner product induces the *Frobenius norm*,

$$\|A\| = \sqrt{\mathrm{trace}[AA^*]} = \sqrt{\sum_{i,j=1}^{n} |a_{i,j}|^2} = \sqrt{\mathrm{trace}[A^*A]}.$$

5.1.2 Remark. We observe that if $\sigma_1, \ldots, \sigma_n$ are the singular values of A, then

$$\|A\| = \sqrt{\sum_{j=1}^{n} \sigma_j^2}.$$

So Frobenius norm can be interpreted as Euclidean norm of singular values.

The Frobenius norm of a matrix product has a convenient estimate. To prepare this, we recall the Cauchy-Schwarz inequality.

5.1.3 Lemma (Cauchy-Schwarz). *Take $\mathbb{C}^n$ with the usual inner product, then for $x, y \in \mathbb{C}^n$,*

$$|\langle x, y\rangle| \le \|x\|\|y\|.$$

Proof. Without loss of generality, assume $x \neq 0$. We have that the orthogonal projection Qy of y onto $\mathbb{C}_x$ satisfies

$$\|y\|^2 = \|Qy\|^2 + \|y - Qy\|^2,$$

so $\|Qy\|^2 \leq \|y\|^2$. Since

$$\|Qy\|^2 = \left\| \frac{x}{\|x\|} \left\langle \frac{x}{\|x\|}, y \right\rangle \right\| = \frac{|\langle x, y\rangle|^2}{\|x\|^2},$$

we have

$$|\langle x, y\rangle|^2 \leq \|x\|^2 \|y\|^2.$$

□

5.1.4 Proposition. *Let $A, B \in M_n$ and $\|\cdot\|_F$ the Frobenius norm, then*

$$\|AB\|_F \leq \|A\|_F \|B\|_F.$$

Proof.

$$\begin{aligned}\|AB\|_F^2 &= \operatorname{trace}[ABB^*A^*]\\ &= \operatorname{trace}[(A^*A)(BB^*)]\\ &\leq (\operatorname{trace}[(A^*A)^2])^{\frac{1}{2}} (\operatorname{trace}[(BB^*)^2])^{\frac{1}{2}} \quad \text{(Cauchy-Schwarz)}\\ &= \|A^*A\|_F \|BB^*\|_F.\end{aligned}$$

If $\sigma_1, \sigma_2, \ldots, \sigma_n$ are the singular values of A,

$$\begin{aligned}\|A^*A\|_F &= \left(\sum_{j=1}^n (\sigma_j^2)^2 \right)^{\frac{1}{2}}\\ &\leq \left(\left(\sum_{j=1}^n \sigma_j^2 \right)^2 \right)^{\frac{1}{2}}\\ &= \sum_{j=1}^n \sigma_j^2 = \operatorname{trace}[A^*A] = \|A\|_F^2.\end{aligned}$$

Using the fact that singular values of B^* are identical to those of B and repeating this for for B gives

$$\|BB^*\|_F \leq \|B\|_F^2.$$

Putting these two estimates together gives

$$\|AB\|_F^2 \leq \|A\|_F^2 \|B\|_F^2.$$

Taking square root on both sides gives the claimed inequality. □

This property of the Frobenius norm under matrix multiplication is called *sub-multiplicativity*.

5.1.5 Definition. A function $|||\cdot|||: M_n \to \mathbb{R}$ is called a *matrix norm* if for all $A, B \in M_n, \lambda \in \mathbb{C}$,

i) $|||A||| \geq 0$ and $|||A||| = 0$ if and only if $A = 0$.

ii) $|||\lambda A||| = |\lambda| |||A|||$.

iii) $|||A + B||| \leq |||A||| + |||B|||$.

iv) $|||AB||| \leq |||A||| |||B|||$.

The preceding proposition shows that the Frobenius norm is a matrix norm. We have other choices for matrix norms.

5.1.6 Theorem. *If $\|\cdot\|$ is a norm on $\mathbb{C}^n$, then it induces a matrix norm $|||\cdot|||$ on M_n by*

$$|||A||| = \max_{\|x\| \leq 1} \|Ax\|.$$

Proof. i) We have $|||A||| \geq 0$. If $|||A||| = 0$, then $\max_{\|x\|=1} \|Ax\| = 0$. So $Ax = 0$ for all $x, \|x\| = 1$. Thus, $A = 0$

ii) $|||\lambda A||| = \max_{\|x\| \leq 1} \|\lambda Ax\| = |\lambda| |||A|||$.

iii)

$$\begin{aligned} \max_{\|x\| \leq 1} \|(A + B)x\| &\leq \max_{\|x\| \leq 1} \|Ax\| + \max_{\|x\| \leq 1} \|Bx\| \\ &= |||A||| + |||B|||. \end{aligned}$$

iv)

$$\max_{\substack{\|x\| \leq 1 \\ x \neq 0}} \|ABx\| \leq \max_{\substack{\|x\| \leq 1 \\ x \neq 0}} |||A||| \|Bx\| = |||A||| |||B|||.$$

□