5374 Fall '22

Numerical Linear Algebra Lecture 6

Warmup

3 vectors x, y, z

Dot product/inner product/scalar product x, y vectors in $\mathbb{R}^n$

$(x = (x_1, x_2, \ldots, x_n)$
$y = (y_1, y_2, \ldots, y_n))$

$$(x, y) \quad \text{or} \quad x \cdot y$$

often denotes the number

$$x_1 y_1 + \cdots + x_n y_n$$

This is often thought of in terms of matrix products, because we could think of x and y as $n \times 1$ matrices

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \qquad y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

Thought of as matrices, we can multiply

$$\underbrace{y^t x}_{(1\times 1 \text{ matrix})} \quad \text{and} \quad \underbrace{y x^t}_{(n \times n \text{ matrix})}$$

$$y^t \cdot x = (y_1 \cdots y_n)\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

$$= x_1 y_1 + \cdots + x_n y_n$$

$$x \otimes y := y x^t = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} (x_1 \cdots x_n)$$

$$= \begin{pmatrix} x_1 y_1 & x_2 y_1 & \cdots & x_n y_1 \\ x_1 y_2 & x_2 y_2 & & \vdots \\ \vdots & \vdots & & \vdots \\ x_1 y_n & x_2 y_n & & x_n y_n \end{pmatrix}$$

Note: if $z = \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix}$

$$x \otimes y = (y x^t) z = y\,(x^t z)$$

Given two vector spaces V_1 and V_2, we denote by $V_1 \otimes V_2$ the space of all expressions obtained by pairing all vectors in V_1 and V_2:

fix $v_1 \in V_1$, $v_2 \in V_2$

$$v_1 \otimes v_2 \in V_1 \otimes V_2$$

and then adding all finite linear combination of these, i.e.

if $v_1, v_2, \dots, v_k \in V_1$
$w_1, w_2, \dots, w_k \in V_2$
and $\alpha_1, \dots, \alpha_k \in \mathbb{R}$

Then $\alpha_1 v_1 \otimes w_1 + \dots + \alpha_k v_k \otimes w_k \in V_1 \otimes V_2$

With the rule that

$$(\alpha_1 v_1 + \alpha_2 v_2) \otimes w = \alpha_1 v_1 \otimes w + \alpha_2 v_2 \otimes w$$

$$v \otimes (\alpha_1 w_1 + \alpha_2 w_2) = \alpha_1 v \otimes w_1 + \alpha_2 v \otimes w_2 .$$

)

The condition number of a matrix, continued

If A is a $n \times n$ matrix, we defined

$$\mathrm{cond}(A) = \begin{cases} \|A\| \, \|A^{-1}\| & \text{if } A \text{ is invertible} \\ +\infty & \text{if } A \text{ is not invertible} \end{cases}$$

Here, the norm refers to the operator norm

$$\|A\| := \sup_{x \neq 0} \frac{\|Ax\|_2}{\|x\|_2}$$

$$= \sup_{\|x\|=1} \|Ax\|_2$$

Let's see lots of examples / special properties of cond(A)

EX 1 Let A be an invertible matrix and $\lambda \in \mathbb{R}$ ($\lambda \neq 0$).

Then $(\lambda A)^{-1} = \lambda^{-1} A^{-1}$, so

$$\|\lambda A\| = |\lambda| \, \|A\|$$
$$\|(\lambda A)^{-1}\| = |\lambda|^{-1} \|A^{-1}\|$$

So, $\text{cond}(\lambda A) = \cancel{|\lambda|} \|A\| \cancel{|\lambda|^{-1}} \|A^{-1}\|$
$= \|A\| \, \|A^{-1}\| = \text{cond}(A)$.

EX 2 If A preserves the Euclidean norm, then $\text{cond}(A) = 1$

(* ie. for all x, $\|Ax\|_2 = \|x\|_2$)

In this case A^{-1} also must preserve norms, since

$$\|x\|_2 \overset{\text{since } AA^{-1}=I}{=} \|A(A^{-1}x)\|_2 = \|A^{-1}x\|_2$$

Since A and A^{-1} preserve the Euclidean norm, their operator norms must be 1:

$$\|A\| = \sup_{x \neq 0} \frac{\|Ax\|_2}{\|x\|_2} = \sup_{x \neq 0} 1 = 1$$

$\|A^{-1}\| = 1$ (for the same reason)

$\Rightarrow \operatorname{cond}(A) = 1 \cdot 1 = 1$

EX 3

Let $A = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$

with $|\lambda_1| \geq |\lambda_2| > 0$

(Exercise: Show if $A = \begin{pmatrix} \alpha_1 & 0 \\ 0 & \alpha_2 \end{pmatrix}$, then $\|A\| = \max\{|\alpha_1|, |\alpha_2|\}$)

Then $$A^{-1} = \begin{pmatrix} \frac{1}{\lambda_1} & 0 \\ 0 & \frac{1}{\lambda_2} \end{pmatrix}$$

(From the previous exercise $\|A^{-1}\| = \frac{1}{|\lambda_2|}$)

Then in this case

$$\text{cond}(A) = \frac{|\lambda_1|}{|\lambda_2|}$$

In particular

$$\text{cond}\left(\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\right) = 1$$

$$\text{cond}\left(\begin{pmatrix} 3 & 0 \\ 0 & 10 \end{pmatrix}\right) = \frac{10}{3}$$

$$\text{cond}\left(\begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{100} \end{pmatrix}\right) = 100$$

EX 4 If

$$A = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \ddots & & \\ \vdots & & & \\ 0 & \cdots & & \lambda_n \end{pmatrix}, \quad \text{and } \lambda_k \neq 0 \ \forall k$$

Then $$\|A\| = \max_{1\le k\le n} |\lambda_k|$$

$$\|A^{-1}\| = \max_{1\le k\le n} |\lambda_k|^{-1} = \frac{1}{\min_{1\le k\le n} |\lambda_k|}$$

so $$\operatorname{cond}(A) = \frac{\max_{1\le k\le n} |\lambda_k|}{\min_{1\le k\le n} |\lambda_k|}$$

In particular, for these matrices
$$\operatorname{cond}(A) \ge 1$$

EX5 Let A be $n\times n$ and Q be a matrix that preserves Euclidean norm. Then

$$\operatorname{cond}(A) = \operatorname{cond}(AQ) = \operatorname{cond}(QA)$$

Why?

The operator norm has the following property:

given A, B $\quad \|AB\| \le \|A\|\,\|B\|$

(Proof: exercise! Use definition of $\|A\| = \sup_{x\ne 0} \frac{\|Ax\|_2}{\|x\|_2}$)

This property means several things, first, if A is invertible, then

$$I = A \cdot A^{-1}$$

$$\Rightarrow \quad 1 = \|I\| = \|A \cdot A^{-1}\| \leq \|A\| \cdot \|A^{-1}\|$$

so $\quad \text{cond}(A) \geq 1.$

But also, we see that

$$\|AQ\| \leq \|A\| \, \|Q\| = \|A\|$$
$$\|(AQ)^{-1}\| = \|Q^{-1}A^{-1}\| \leq \|Q^{-1}\| \, \|A^{-1}\| = \|A^{-1}\|$$

$$\Rightarrow \quad \|AQ\| \, \|(AQ)^{-1}\| \leq \|A\| \cdot \|A^{-1}\|$$

This shows that

$$\text{cond}(AQ) \leq \text{cond}(A)$$

But now, we can get the reverse inequality via a trick

$$\text{cond}(A) = \text{cond}(AQQ^{-1})$$
$$\leq \text{cond}(AQ)$$
(since Q^{-1} preserves norm)

$$\Rightarrow \quad \operatorname{cond}(A) = \operatorname{cond}(AQ)$$

The same argument shows that

$$\operatorname{cond}(QA) = \operatorname{cond}(A)$$

EX6 Combining the two observation from example 5, we see that given any two matrices Q_1 and Q_2 which preserve norm then

$$\operatorname{cond}(Q_1 A Q_2) = \operatorname{cond}(A).$$

EX 7 In linear algebra we learn that if A is a symmetric $n \times n$ matrix (i.e. $a_{ij} = a_{ji}$) Then there exists a matrix Q which preserves norm and a diagonal matrix D

$$D = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \ddots & & \\ \vdots & & & \\ 0 & & & \lambda_n \end{pmatrix}$$

where $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of A, and such that

$$A = QDQ^{-1}$$

Then

$$\text{cond}(A) = \text{cond}(QDQ^{-1})$$
$$= \text{cond}(D)$$

In condensed, if A is a symmetric matrix, and $\lambda_1, \dots, \lambda_n$ its eigenvalues, then

$$\text{cond}(A) = \frac{\max |\lambda_k|}{\min |\lambda_k|}$$

Remark : As we will see/review soon, a $n \times n$ matrix Q preserves norms if and only if

$$QQ^t = I$$

in other words, if

$$Q^{-1} = Q^t$$

Such matrices are more commonly called orthogonal matrices.

Example Given $\theta \in [0, 2\pi)$, let

$$A_\theta = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$$

Check that

$$A_{\theta_1} A_{\theta_2} = A_{\theta_1+\theta_2}$$

and check that

$$A_\theta^t = A_{-\theta} = A_\theta^{-1}$$

Example The following 4x4 matrix is also orthogonal

$$\left(\begin{array}{cc|cc} \cos(\theta_1) & -\sin(\theta_1) & 0 & 0 \\ \sin(\theta_1) & \cos(\theta_1) & 0 & 0 \\ \hline 0 & 0 & \cos(\theta_2) & -\sin(\theta_2) \\ 0 & 0 & \sin(\theta_2) & \cos(\theta_2) \end{array}\right)$$