3 condition numbers

Sunday, September 13, 2020 6:52 PM

Def. 8.9 For $A = (a_{ij}) \in \mathbb{C}^{n \times n}$, the norm $\|A\|_2$ is often called the **spectral norm**.

A few technical inequalities:

Prop. 8.8/8.11 Let $\|\ \|$ be any matrix norm, and let $B \in \mathbb{C}^{n\times n}$ s.t. $\|B\| < 1$.

(1) If $\|\ \|$ is a subordinate matrix norm, then $I+B$ is invertible and

$$\|(I+B)^{-1}\| \le \frac{1}{1-\|B\|}.$$

(2) If a matrix of the form $I+B$ is singular, then $\|B\| \ge 1$ for every matrix norm.

proof. (1) If $(I+B)u = 0$, then $Bu = -u \Rightarrow \|u\| = \|Bu\|$.

But $\|Bu\| \le \|B\|\,\|u\|$ for every subordinate norm.

Since $\|B\| < 1$, if $u \ne 0$, then $\|Bu\| < \|u\|$, a contradiction.

$\Rightarrow u = 0$ if $(I+B)u = 0$, so $I+B$ is invertible.

Then $I = (I+B)(I+B)^{-1} = (I+B)^{-1} + B(I+B)^{-1}$.

$$\Rightarrow (I+B)^{-1} = I - B(I+B)^{-1}.$$

$$\Rightarrow \|(I+B)^{-1}\| \le 1 + \|B\|\,\|(I+B)^{-1}\|$$

$$\Rightarrow \|(I+B)^{-1}\| \le \frac{1}{1-\|B\|}.$$

(2) If $I+B$ is singular, then -1 is an eigenvalue of B, and recall $\rho(B) \le \|B\|$ $\Rightarrow$ $1 \le \rho(B) \le \|B\|$. $\square$

spectral radius

We will need the following to deal with convergence of sequences of matrix powers

Prop. 8.9/8.12 $\forall A \in \mathbb{C}^{n\times n}$, $\forall \varepsilon > 0$, $\exists$ a subordinate matrix norm $\|\ \|$ s.t. $\|A\| \le \rho(A) + \varepsilon$.

proof. Recall the Schur decomposition, stating that $A = UTU^{-1}$ where U is a unitary matrix and T is upper triangular

proof. Recall the Schur decomposition, stating that
$A = UTU^{-1}$, where U is a unitary matrix and T is upper triangular.

Say
$$T = \begin{pmatrix} \lambda_1 & t_{12} & t_{13} & \cdots & t_{1n} \\ & \lambda_2 & \ddots & & \vdots \\ & & \ddots & \ddots & \vdots \\ & & & \ddots & t_{n-1,n} \\ 0 & & & & \lambda_n \end{pmatrix},$$
where $\lambda_1, \dots, \lambda_n$ are eigenvalues.

For every $\delta \neq 0$, define
$$D_\delta = \mathrm{diag}(1, \delta, \delta^2, \dots, \delta^{n-1}).$$

Then
$$(UD_\delta)^{-1} A (UD_\delta) = D_\delta^{-1} T D_\delta = \begin{pmatrix} \lambda_1 & \delta t_{12} & \delta^2 t_{13} & \cdots & \delta^{n-1} t_{1n} \\ & \lambda_2 & \ddots & & \vdots \\ & & \ddots & \ddots & \vdots \\ & & & \ddots & \delta t_{n-1,n} \\ & & & & \lambda_n \end{pmatrix}$$

We now construct $\| \ \| : \mathbb{C}^{n \times n} \to \mathbb{R}$ by defining
$$\|B\| = \|(UD_\delta)^{-1} B (UD_\delta)\|_\infty,$$
which is the matrix norm subordinate to the vector norm
$$v \mapsto \|(UD_\delta)^{-1} v\|_\infty.$$

$\forall \varepsilon > 0$, pick δ s.t. $\displaystyle\sum_{j=i+1}^{n} |\delta^{j-i} t_{ij}| \le \varepsilon, \qquad 1 \le i \le n-1$

$\Rightarrow \|A\| \le \rho(A) + \varepsilon$. (because $\| \ \|_\infty$ is the maximum over ℓ^1-norm of rows) $\square$

Aside: equality is generally not possible (e.g. $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$).

Condition numbers of matrices

Def 8.10 For any subordinate matrix norm $\| \ \|$, for any invertible A,
$$\mathrm{cond}(A) = \|A\| \, \|A^{-1}\|$$
is called the condition number of A relative to $\| \ \|$.

Condition numbers measure the sensitivity of the linear system $Ax = b$ to variations in A and b.

Condition numbers measure the sensitivity of the [illegible]
to variations in A and b.

Prop. 8.10/8.13 Let A be an invertible matrix and let

$$Ax = b$$
$$A(x+\delta x) = b + \delta b.$$

If $b \neq 0$, then

$$\frac{\|\delta x\|}{\|x\|} \leq \operatorname{cond}(A) \cdot \frac{\|\delta b\|}{\|b\|}.$$

Further, for a given A, $\exists\, b \neq 0$ and $\delta b \neq 0$ for which equality holds.

proof. $\delta x = A^{-1}\delta b$.

$$\Rightarrow \|\delta x\| \leq \|A^{-1}\| \, \|\delta b\|$$

$$\Rightarrow \|b\| \leq \|A\| \, \|x\|. \Rightarrow \|x\| \geq \frac{\|b\|}{\|A\|} \Rightarrow \frac{1}{\|x\|} \leq \frac{\|A\|}{\|b\|}$$

$$\Rightarrow \frac{\|\delta x\|}{\|x\|} \leq \frac{\|A^{-1}\| \, \|\delta b\|}{\frac{\|A\|}{\|b\|}} = \left(\|A\| \, \|A^{-1}\|\right) \cdot \frac{\|\delta b\|}{\|b\|}.$$

$$= \operatorname{cond}(A) \cdot \frac{\|\delta b\|}{\|b\|}.$$

Now just need to exhibit $b \neq 0$ and $\delta b \neq 0$ s.t. we get equality.

Because $\|\ \|$ is a subordinate matrix norm, $\exists\, x \neq 0$ and $\delta b \neq 0$ s.t.

$$\|A^{-1}\delta b\| = \|A^{-1}\| \, \|\delta b\| \quad \text{and} \quad \|Ax\| = \|A\| \, \|x\|.$$

(since subordinate matrix norms can be defined as supremums over closed unit balls).

Then $$\frac{\|\delta x\|}{\|x\|} = \frac{\|A\| \, \|\delta x\|}{\|A\| \, \|x\|} = \frac{\|A\| \, \|A^{-1}\delta b\|}{\|Ax\|} = \frac{\|A\| \, \|A^{-1}\| \, \|\delta b\|}{\|b\|}.$$ $\blacksquare$

Above, we have perturbed b. We can similarly perturb A.

Prop. 8.11/8.14 Let A be an invertible matrix, and let

$$Ax = b$$
$$(A+\delta A)(x+\delta x) = b.$$

[illegible] $\|\delta A\|$ [illegible] the

$(A+\Delta A)(x+\Delta x) = b.$

If $b \neq 0$, then $\dfrac{\|\Delta x\|}{\|x+\Delta x\|} \leq \text{cond}(A) \cdot \dfrac{\|\Delta A\|}{\|A\|}$ holds and is the best possible.

Basically, the condition number tells us how sensitive the system $Ax=b$ is to perturbations under $\| \ \|$.

Def. 8.11 Given a complex $n \times n$ matrix A, a triple (U, V, Σ) such that $A = V \Sigma U^T$, where U and V are $n \times n$ unitary matrices and $\Sigma = \text{diag}(\sigma_1, \ldots, \sigma_n)$ is a diagonal matrix of singular values $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n \geq 0$ is called a singular value decomposition (SVD) of A. If A is real, then U and V are orthogonal matrices.

Preview: Chapter 20: every real or complex rectangular matrix has an SVD. (Nice geometric interpretation that we will cover after adjoints and spectral theorems).

Note: $A = V\Sigma U^* \Rightarrow A^*A = U\Sigma^2 U^*$ and $AA^* = V\Sigma^2 V^*$

$\Rightarrow \sigma_1^2, \ldots, \sigma_n^2$ are eigenvalues of both A^*A and AA^*.

$\Rightarrow$ cols of U are eigenvectors for A^*A.

cols of V are eigenvectors for A^*A.

$$\Rightarrow \sqrt{\rho(A^*A)} = \sqrt{\rho(AA^*)} = \sigma_1.$$

Corollary 8.3/8.15 $\|A\|_2 = \sigma_1$ (spectral norm is equal to largest sing. val.)

$$= \|(\sigma_1, \ldots, \sigma_n)\|_\infty.$$

Corollary 8.4/8.16 $\|A\|_F = \sqrt{tr(A^*A)} = \sqrt{\sigma_1^2 + \cdots + \sigma_n^2} = \|(\sigma_1, \ldots, \sigma_n)\|_2.$

Prop. 8.12/8.17 For any invertible $A \in \mathbb{C}^{n\times n}$,

(because $1 = \|I\| \leq \|A\|\|A^{-1}\| = \text{cond}(A)$)

(1) $\text{cond}(A) \geq 1$

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

(1) $\text{cond}(A) \geq 1$

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

$\text{cond}(\alpha A) = \text{cond}(A) \quad \forall \alpha \in \mathbb{C} - \{0\}$.

(2) Let $\text{cond}_2(A)$ be the condition number of A w.r.t. the spectral norm. Then
$$\text{cond}_2(A) = \frac{\sigma_2}{\sigma_n}.$$

(3) If A is normal (i.e. $AA^* = A^*A$), then
$$\text{cond}_2(A) = \frac{|\lambda_1|}{|\lambda_n|},$$
where $\lambda_1, \dots, \lambda_n$ are the eigenvalues, sorted so that $|\lambda_1| \geq |\lambda_2| \geq \dots \geq |\lambda_n|$.

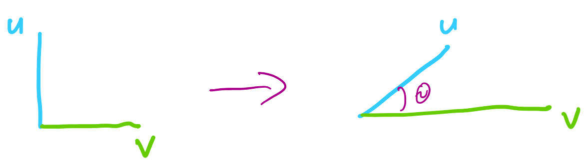
(4) If A is a unitary or orthogonal matrix, then $\text{cond}_2(A) = 1$.

(5) $\text{cond}_2(A) = \text{cond}_2(UA) = \text{cond}_2(AV) \quad \forall$ unitary U, V.

We very often care about unitary/orthogonal matrices since they are extremely well-conditioned for the spectral norm (indeed, they are "length-preserving").

<u>Geometric</u> <u>interpretation</u>

$\text{cond}_2(A) = \cot\left(\frac{\Theta(A)}{2}\right)$, where $\Theta(A)$ is the smallest angle between Au and Av, where u, v are orthonormal vectors.



<u>Caution</u>: It's tempting to think of condition number cond_2 as related to eigenvalues. However, this is only true for normal matrices.

Consider $A = \begin{pmatrix} 1 & 2 & & 0 \\ & 1 & 2 & \\ & & \ddots & 2 \\ 0 & & & 1 \end{pmatrix}$, with $\text{cond}_2(A) \geq 2^{n-1}$.

Or a Hilbert matrix $H^{(n)}_{ij} = \left(\frac{1}{i+j-1}\right)$, $\text{cond}_2(H^{(n)}) = O\left(\frac{(1+\sqrt{2})^{4n}}{\sqrt{n}}\right)$.