More on Norms

Monday, October 03, 2005
11:00 AM

Norms are important in two ways:

1) Measuring accuracy of approximations
2) Convergence proofs

Not proven, but useful: norms are continuous, i.e.
$\|\boldsymbol{x}\| \to 0$ $\boldsymbol{x} \Rightarrow 0$ componentwise

(3 norm axioms)

Matrix Norms

Monday, October 03, 2005
11:14 AM

1) $f(\boldsymbol{A}) > 0 \Leftrightarrow \boldsymbol{A} \neq 0$
   $= 0 \Leftrightarrow \boldsymbol{A} = 0$
2) $f(\boldsymbol{A}+\boldsymbol{B}) \leq f(\boldsymbol{A}) + f(\boldsymbol{B})$
3) $f(\alpha\boldsymbol{A}) = |\alpha| f(\boldsymbol{A})$
4) $f(\boldsymbol{AB}) \leq f(\boldsymbol{A}) f(\boldsymbol{B})$ (new)

Example: the Frobenius norm:

$\|\mathbf{A}\|_F = (\sum_{i=1,\ldots,n} \sum_{j=1,\ldots,n} |a_{ij}|^2)^{1/2}$
Not hard to show that this satisfies the four above properties

Note that (4) says that

$\|A^2\| \leq \|A\|^2$, $\|A^n\| \leq \|A\|^n$, so if $\|A\| < 1$,
then $A^n \to \emptyset$

Natural Norm

Monday, October 03, 2005
11:20 AM

$$\|A\| = \max_{\tilde{x} \neq 0} \|Ax\| / \|\tilde{x}\|$$

We say the vector norm "induces" the matrix norm

1) if $A = 0$, then $\|A\tilde{x}\| = 0$

if $A \neq 0$, say $a_{ij} \neq 0$, and consider e_j

Say $\vec{y}$ is the vector that maximizes

$$2) \ \|\alpha A\| = \max_{x} \frac{\|\alpha A\vec{x}\|}{\|\vec{x}\|} = \frac{\|\alpha(A\vec{y})\|}{\|\vec{y}\|} = |\alpha| \frac{\|A\vec{y}\|}{\|\vec{y}\|}$$

Pick arbitrary $\vec{z} \neq 0$, note that ← Aside

$$\|A\vec{z}\| = \left\| A \|\vec{z}\| \cdot \frac{\vec{z}}{\|\vec{z}\|} \right\| = \|\vec{z}\| \cdot \left\| A \frac{\vec{z}}{\|\vec{z}\|} \right\| \leq \|\vec{z}\| \cdot \|A\|$$

3) Say $(A+B)$ has its max at $\vec{y}_0$,

$$\frac{\|(A+B)\vec{y}_0\|}{\|y_0\|} \leq \frac{\|Ay_0\|}{\|y_0\|} + \frac{\|By_0\|}{\|y_0\|} \leq \|A\| + \|B\|$$

what about (4)?

assume that $\max \dfrac{\|AB\vec{x}\|}{\|\vec{x}\|}$ is attained at

some vector $\vec{y}$

Note that $B\vec{y}$ is a vector, so

$$\frac{\|A(B\vec{x})\|}{\|\vec{x}\|} \leq \|A\| \cdot \frac{\|B\vec{y}\|}{\|\vec{y}\|} \leq \|A\| \cdot \|B\|$$

What is $\|A\|_\infty$?

$$\max \frac{\|A\vec{x}\|_\infty}{\|\vec{x}\|_\infty} = \max_{\|y\|_\infty = 1} \|A\vec{y}\|_\infty$$

$$= \max_{\|y\|_\infty = 1} \max_{1 \le i \le m} \left(\sum_{j=1}^{n} a_{ij} y_j \right)$$

$$\le \max_{\|y\|_\infty = 1} \max_{1 \le i \le m}$$

$$\max_{1 \le i \le m} \sum_{j=1}^{n} |a_{ij}|$$

so $\|A\|_\infty \le \max_{1 \le i \le m} \sum_{j=1}^{n} |a_{ij}|$

$$= \sum_{j=1}^{n} |a_{Ij}|$$ for some I

$y_j = 1$ if $a_{Ij} \ge 0$

-1 if $a_{Ij} < 0$

$$\|A\vec{y}\| = \sum_{j=1}^{n} |a_{Ij}|$$

So $\|A\|_\infty = \max \sum |a_{ij}|$

So $\|A\|_\infty = \max\limits_{1 \le i \le m} \sum\limits_{j=1}^{n} |a_{ij}|$

So if a matrix looks like

$$\begin{pmatrix} 1/2 & -1/4 \\ -1/8 & 2 \end{pmatrix} \quad \text{then } \|A\|_\infty = 2\tfrac{1}{8}$$

$$\begin{pmatrix} 0 & 1/2 & & \\ 1/4 & \ddots & \ddots & \\ & \ddots & \ddots & 1/2 \\ & & 1/4 & 0 \end{pmatrix} \quad \text{then} \quad \|A\|_\infty = 3/4 \quad \text{hence}$$

$$A^n \to 0$$

A more delicate case:

$$\begin{pmatrix} 0 & 1/2 & 0 \\ 1/2 & 0 & 1/2 \\ 0 & 1/2 & 0 \end{pmatrix} \to 0$$

by eigenvalue analysis

The 2-norm

Monday, October 03, 2005
11:42 AM

Recall that

$$\|\vec{x}\|_2 = \left(\sum |x_i|^2 \right)^{1/2} = \left(\vec{x}^T \vec{x} \right)^{1/2}$$

$$\|A\|_2 = \max_{\vec{x} \neq 0} \frac{\|Ax\|_2}{\|x\|_2} = \max_{\vec{x} \neq 0} \left(\frac{x^T A^T A x}{x^T x} \right)^{1/2}$$

How large can this get? Now, $A^T A$ is symmetric and hence can be diagonalized. Further, it is positive semidefinite, meaning that the quadratic form is nonnegative. Write

$A^T A = U \Sigma U^T$ where $U^T U = I$ and

$$\Sigma = \begin{pmatrix} \sigma_1^2 & & 0 \\ & \ddots & \\ 0 & & \sigma_n^2 \end{pmatrix}$$ are the ordered eigenvalues

for each $x^T A^T A x$, write $\dfrac{x^T U \Sigma U^T x}{x^T U U^T x}$

$$= \frac{w^T \Sigma w}{w^T w}$$ where $w = U^T x$

$$= \frac{\sum_{i=1}^{n} \sigma_i^2 w_i^2}{\sum_{i=1}^{n} w_i^2} \leq \sigma_1^2 \frac{\sum w_i^2}{\sum w_i^2} \leq \sigma^{1^2}$$

So,

$$\frac{x^T A^T A x}{x^T x} \leq \sigma_1^2$$

$$\|A\|_2^2 \leq \sigma_1^2$$

so we have an upper bound on $\|A\|_2$; does

it ever achieve this? Recall

$$A^T A = U \Sigma U^T$$

$$A^T A U = U \Sigma \qquad U = (\vec{u}_1, \dots, \vec{u}_n)$$

$$A^T A u_i = \sigma_i^2 u_i \Rightarrow A^T A u_1 = \sigma_1^2 u_1$$

$$\frac{x^T A^T A x}{x^T x} = \frac{x^T U \Sigma U^T x}{x^T x}$$

Here, if we choose $\vec{x} = \vec{e}_1$ then the above quantity is σ_1^2

So, $\|A\|_2 = \sigma_1$

So, $\|A\|_2 = \sigma_1$

Where $\sigma_1 = [\lambda_{max} \text{ of } A^T A]^{1/2}$

2-norm is also sometimes called the "spectral norm"

Norms and eigenvalues

Monday, October 03, 2005
12:00 PM

Remember that every matrix has eigenvectors:

$$A\vec{x}_i = \lambda_i \vec{x}_i \qquad i = 1, \dots, k$$

where k may be less than n

We can make the following claim:

$$\|A\vec{x}_i\| = \|\lambda_i \vec{x}_i\| \leq |\lambda_i| \cdot \|\vec{x}_i\|$$

Aside: $A = \begin{pmatrix} \rho & 1 \\ 0 & \rho \end{pmatrix}$, for example, has only one eigenvector

so

$$|\lambda_i| = \frac{\|A\vec{x}_i\|}{\|\vec{x}_i\|} \leq \|A\|$$

$$\|A\| \geq \max_{1 \leq i \leq n} |\lambda_i| = \rho(A)$$

Call $\rho(A)$ the "spectral radius"

So with any norm, we get a bound on the spectral radius

Example:

$$A: \begin{pmatrix} 1/2 & 1/4 \\ 3/4 & 10 \end{pmatrix} \qquad \|A\|_\infty : 10\,3/4$$

so $\rho(A) \leq 10\tfrac{3}{4}$

One can prove that there exists a norm where

$$\|A\| \leq \rho(A) + \varepsilon$$

for a particular A

Consider the matrix

$$A = \begin{pmatrix} 2 & -1 & & 0 \\ -1 & \ddots & \ddots & \\ & \ddots & \ddots & 1 \\ 0 & & -1 & 2 \end{pmatrix}$$

$$\|A\|_\infty = 4 \qquad \rho(A) = 2 + 2\cos\frac{\pi}{n+1} < 4$$

Monday, October 03, 2005
12:10 PM

All norms are equivalent

$$c_2 \|A\|_\alpha \le \|A\|_\beta \le c_1 \|A\|_\alpha$$

$$\frac{1}{\sqrt{n}} \|A\|_\infty \le \|A\|_2 \le \sqrt{n} \|A\|_\infty$$

Gerschgorin Theorem

Monday, October 03, 2005
12:11 PM

We have $A\vec{x} = \lambda\vec{x}$ for k eigenvectors

$$\sum_{j=1}^{n} a_{ij} x_i = \lambda x_i$$

$$(a_{ii} - \lambda) x_i = -\sum_{j \neq i} a_{ij} x_j$$

$$|a_{ii} - \lambda| \cdot |x_i| \leq \sum_{j \neq i} |a_{ij}| \cdot |x_j| \quad i = 1, \dots, n$$

$$|x_i| \leq |x_I|$$

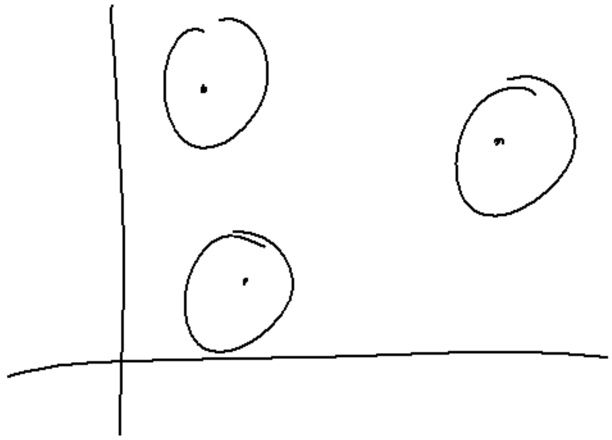
$$|a_{II} - \lambda| = \sum_{I \neq i} |a_{Ij}| \cdot \frac{|x_j|}{|x_I|} \leq \sum_{j \neq I} |a_{Ij}|$$

We don't know for which i this happens, so define

$$r_i = \sum_{j \neq i} |a_{ij}|$$

$$|\lambda - a_{ii}| \leq r_i$$

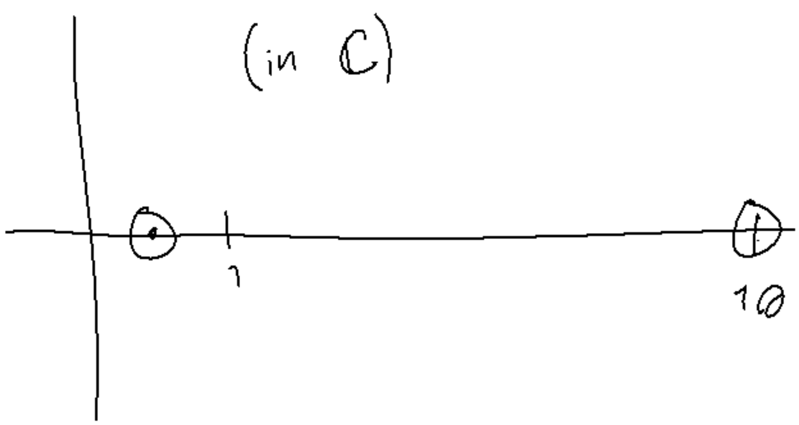
This gives a set of disks in the complex plane:



Take their union, and the eigenvalues must lie in this union

Example:

$$\begin{pmatrix} 1/2 & 1/4 \\ 1/8 & 10 \end{pmatrix} \quad \text{gives} \quad \begin{aligned} |\lambda - 1/2| &\le 1/4 \\ |\lambda - 10| &\le 1/8 \end{aligned}$$



$$\begin{pmatrix} 0 & 1/2 & & \\ 1/2 & \ddots & \ddots & 0 \\ & \ddots & \ddots & 1/2 \\ 0 & & 1/2 & 0 \end{pmatrix}$$

$$\begin{pmatrix} 0 & \tfrac{1}{2} & & \\ \tfrac{1}{2} & 0 & \ddots & \\ & \ddots & \ddots & \tfrac{1}{2} \\ & & \tfrac{1}{2} & 0 \end{pmatrix}$$

Symmetric, so all real eigenvalues; further, they all lie between -1 and 1

$$\begin{pmatrix} 2 & -1 & & \\ -1 & \ddots & \ddots & \\ & \ddots & \ddots & -1 \\ & & -1 & 2 \end{pmatrix}$$

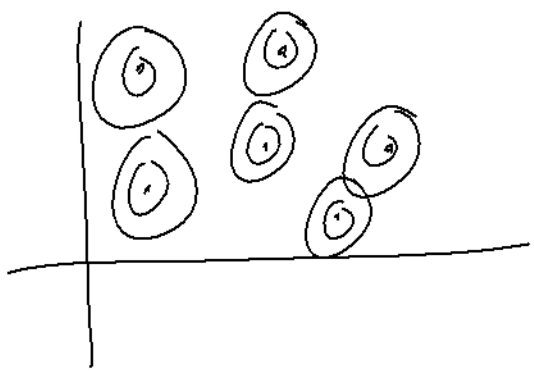
this says $|2-\lambda| \le 2$

So the eigenvalues lie between 0 and 4

How about the matrix

$$A(\varepsilon) = (D + \varepsilon R)$$

$$R = \begin{pmatrix} 0 & & a_{ij} \\ & \ddots & \\ a_{ij} & & 0 \end{pmatrix} \qquad D = \begin{pmatrix} a_{11} & & 0 \\ & \ddots & \\ 0 & & a_{nn} \end{pmatrix}$$



When $\varepsilon = 0$, the eigenvalues are the a_{ii}; the general theorem says that if we have disjoint domains, the eigenvalues lie in each of them