From last time, the idea behind the Gerschgorin theorem is that if we have a matrix *A*, with

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

Make a set of disks in the complex plane

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

The eigenvalues must lie in $\bigcup D_i$

Example:

$$A = \begin{pmatrix} 100 & 1 \\ 2 & 99 \end{pmatrix} \qquad \begin{array}{l} |\lambda - 100| \leq 1 \\ |\lambda - 99| \leq 2 \end{array}$$

Taking the transpose of A doesn't change anything

Example:

$$\begin{pmatrix} 100 & 1 \\ -3 & 2 \end{pmatrix} \qquad \begin{array}{l} |\lambda - 100| \leq 1 \\ |\lambda - 2| \leq 3 \end{array}$$

$$\begin{pmatrix} 1 & 10^6 \\ 0 & 1 \end{pmatrix}$$ the eigenvalues are $\lambda_1 = \lambda_2 = 1$
but the theorem just says

$|1-\lambda| \leq 10^6$, which isn't useful

So, let's replace *A* with *DAD*$^{-1}$

where D is diagonal, e.g.

$$D = \begin{pmatrix} \varepsilon & 0 \\ 0 & 1 \end{pmatrix}$$

This gives DAD^{-1} = $\begin{pmatrix} 1 & 10^6\varepsilon \\ 0 & 1 \end{pmatrix}$

So the 10^6 term can be made arbitrarily small

Upcoming homework: sequence of diagonal transformations on a bidiagonal matrix

There exist fancier methods such as "ovals of cassini"

Singular Value Decomposition

Wednesday, October 05, 2005
11:13 AM

A is an m $\times$ n with m≥n matrix; its SVD is

$A = U\Sigma V^T$

1) $U^TU = I_m$ (both orthogonal)
2) $V^TV = I_n$

3) $\Sigma = \begin{pmatrix} \sigma_1 & & 0 \\ & \ddots & \\ 0 & & \sigma_n \end{pmatrix}_{m\times n}$ (almost diagonal)

σ_i are "singular values"

Say $\sigma_1, \dots, \sigma_r > 0$, and $\sigma_{r+1}, \dots, \sigma_n = 0$

This is different from eigenvalues, for

$A = \begin{pmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{pmatrix}$ has n zero eigenvalues

but $\sigma_i = 1$ for $i = 1, \dots, n-1$, $\sigma_n = 0$

"Rank" is not very well-defined; what is the rank of

$\begin{pmatrix} 1 & 0 \\ 0 & 10^{-6} \end{pmatrix}$?

It depends on the application -- maybe machine precision makes $10^{-6} \approx 1$

We wrote

$$A = U\Sigma V^T \Leftrightarrow AV = U\Sigma$$

Partition the vectors: $V = [v_1,\dots,v_n]$

$$A v_i = \sigma_i u_i \qquad i = 1,\dots,n$$

$$A v_i = \vec{0} \qquad i = r+1,\dots,n$$

This tells us that the singular values form a basis?

If we redefine

$$\tilde{U} = [\tilde{u}_1,\dots,u_r] \qquad \tilde{V} = [v_1,\dots,v_r]$$

then

$$A = \tilde{U}\tilde{\Sigma}\tilde{V}^T$$ with $\tilde{U}: m\times r$, $\tilde{\Sigma}: r\times r$, $\tilde{V}: n\times r$

Suppose

$$\underset{m\times m}{A} = U\Sigma V^T$$

$\sigma_m > 0$ (meaning all strictly positive)

$\sigma_m > 0$ (meaning all strictly positive)

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

Easy to compute

Recall the 2-norm of a matrix

$$\|A\|_2^2 = \max_{\vec{x}\neq 0} \frac{x^TA^TAx}{x^Tx} = \lambda_{max}(A^TA)$$

Now,

$$A^TA = V\Sigma^TU^TU\Sigma V^T = \underset{\text{orthog}}{V}\,\underset{\text{diag}}{\Sigma^T}\,\underset{\text{d}}{\Sigma}\,\underset{\text{o}}{V^T}$$

Therefore,

$$\lambda_{max}(A^TA) = \sigma_{max}^2(A^TA)$$

$$\Rightarrow \|A\|_2 = \sigma_{max}(A)$$

Also, note that

$$\|\vec{x}\|_2 = \|Qx\|_2, \quad Q^TQ = I$$

Not hard to show that

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

Example:

Π a permutation matrix

$$\|\Pi A\|_2 = \|A\|_2 = \sigma_1$$

Which isn't true for eigenvalues; we would require $\Pi A \Pi^T$ for that

Example

A: $m \times n$, $\boldsymbol{b}$: $m \times 1$

Want $\min \|b - Ax\|_2$

$= \min \|b - U\Sigma V^T x\|_2$

Since that whole expression is a vector, we can multiply it by orthogonal transformation

$= \min \|U^T b - \Sigma V^T x\|_2$

Say $U^T b = c$; $V^T x = y$

$= \min\limits_{y} \|c - \Sigma y\|_2$

$$= \min \left[\sum_{i=1}^{r} (c_i - \sigma_i y_i)^2 + \sum_{i=r+1}^{m} c_i^2 \right]^{1/2}$$

The only thing we can vary is the y_i's

Set $\hat{y}_i = c_i / \sigma_i$ for $i = 1, 2, \ldots, r$

The $\hat{y}_i$ for $i = r+1, \ldots, n$ (m?) don't matter

$$\min \|\vec{c} - \Sigma \vec{y}\|_2 = \left(c_{r+1}^2 + \cdots + c^2_{\underset{n?}{m}} \right)^{1/2}$$

Remember

$$V^T x = y, \quad \text{so} \quad x = Vy$$

We want to choose $y_{r+1}, \ldots, y_n$ so that

$\|\vec{x}\|_2$ is minimized

Choose $y_{r+1} = \ldots = y_n = 0$

We have $\vec{x} = V\vec{y} = V\Sigma^+ c$

Where

$$\Sigma^+ = \begin{pmatrix} \sigma_1^{-1} & & \\ & \ddots & \\ & & \sigma_r^{-1} \\ & & \end{pmatrix} \quad n \times m$$

this gives $\vec{x} = A^T b$ where

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

Call A^+ the "pseudoinverse"

Pseudoinverse

Wednesday, October 05, 2005
11:42 AM

Say A is $m \times n$; want X satisfying

1) $(AX)^T = AX$
2) $(XA)^T = XA$
3) $AXA = A$
4) $XAX = X$

Such an X is called a "pseudo-inverse"

If such an X exists, it is unique; in fact, it always exists

Moore-Penrose

If we take the singular value decomposition,

$A = U\Sigma V^T, \quad A^{-1} = V\Sigma^{-1}U^T$

Cool theorem

Wednesday, October 05, 2005
11:47 AM

Let $\mathcal{X} = \{x \mid \|b - Ax\|_2 = \min\}$

$\tilde{x} \in \mathcal{X}$ such that

$\|\tilde{x}\|_2 = \min$

$\hat{x} = A^+ \vec{b}$

Projection Matrices

Wednesday, October 05, 2005
11:53 AM

Let $\vec{b} = \vec{b}_1 + \vec{b}_2$; write

$$b_1 = AA^+b, \quad b_2 = (I - AA^+)b$$

What's AA^+? It's

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

$$= U \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} U^T =: P$$

Observe that $P^2 = P$ and
$I - AA^+ = I - P = (I - P)^2 =: P^\perp$
By this property, $(AA^+)^2 = AA^+AA^+$

if we write $A^+ = X$ then the above is
$(AXA)X = AX = (AX)^T$

So, to get a "projection matrix" we really only require 2 of the pseudoinverse axioms to be satisfied

Another way to solve the minimal least squares problem is to make a projection matrix

$$A\vec{x} = Pb = b_1$$

$r = b - A\bar{x}$ is called the "residual vector"

$$= b - AA^{+}b = (I - AA^{+})b = P^{\perp}b$$

Quadratic Forms

Wednesday, October 05, 2005
12:00 PM

$A = A^T = U\Lambda U^T$, $\lambda_1 \geq \ldots \geq \lambda_n \geq 0$

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

$$=: \frac{\bar{y}^T \Lambda \bar{y}}{y^T y} = \frac{\sum_{i=1}^{n} \lambda_i y_i^2}{\sum_{i=1}^{n} y_i^2} \leq \lambda_1$$

Note that $\lambda_n \leq$

In HW, will consider

$$? \leq \frac{x^T A x}{x^T B x} \leq ?$$

Bilinear Forms

Wednesday, October 05, 2005
12:03 PM

A: m × n

look at $\dfrac{u^T A v}{\|u\|_2 \|v\|_2} = \dfrac{|\tilde{u}^T U \Sigma V^T v|}{\|U^T \tilde{u}\|_2 \cdot \|V^T \tilde{v}\|_2}$

$$=: \frac{|x^T \Sigma y|}{\|\tilde{x}\|_2 \|y\|_2} \leq \frac{\sigma_1 \|\tilde{x}\|_2 \|y\|_2}{\|x\|_2 \|y\|_2} = \sigma_1$$

by using Cauchy-Schwarz

$$\max_{u,v} \frac{|u^T A v|}{\|u\|_2 \|v\|_2} = \sigma_1$$

Proof of SVD

Wednesday, October 05, 2005
12:06 PM

Begin with a matrix *A*, look at

$$\tilde{A} = \begin{pmatrix} 0 & A \\ A^T & 0 \end{pmatrix}$$

Since $\tilde{A}$ is symmetric, it can be decomposed

$$\tilde{A} = Z \Lambda Z^T$$

$$\tilde{A}^2 = \begin{pmatrix} 0 & A \\ A^T & 0 \end{pmatrix}\begin{pmatrix} 0 & A \\ A^T & 0 \end{pmatrix} = \begin{pmatrix} AA^T & 0 \\ 0 & A^TA \end{pmatrix}$$

The eigenvalues $\lambda(\tilde{A}^2) = \sigma^2(A)$ are equal to the singular values squared

Important useful theorem:

The eigs of *AB* are the eigs of *BA*

e.g. $a = \boldsymbol{u}\boldsymbol{u}^T$
$\boldsymbol{u}^T\boldsymbol{u}$ is only a number, both have only one eigenvalue

Anyway, take

$$\begin{pmatrix} 0 & A \end{pmatrix} \vec{z} = \sigma \vec{z}$$

$\begin{pmatrix} \cdot & \cdot \\ A & \emptyset \end{pmatrix}$

$\vec{z} = \begin{pmatrix} \vec{x} \\ \vec{y} \end{pmatrix}$

$$\begin{pmatrix} \emptyset & A \\ A^T & \emptyset \end{pmatrix} \begin{pmatrix} \vec{x} \\ -\vec{y} \end{pmatrix} = \begin{pmatrix} -Ay \\ A^T\vec{x} \end{pmatrix} = \begin{pmatrix} -\sigma\vec{x} \\ \sigma y \end{pmatrix} = -\sigma \begin{pmatrix} \vec{x} \\ -y \end{pmatrix}$$

$=: -\sigma \mathfrak{z}$

This tells us $\vec{z}^T \mathfrak{z} = \emptyset$

$$\left.\begin{array}{l} \vec{x}^T\vec{x} - \vec{y}^T\vec{y} = \emptyset \\ \vec{x}^T\vec{x} + y^T y = 1 \end{array}\right. \Rightarrow \begin{array}{l} x^T x = 1/2 \\ y^T y = 1/2 \end{array}$$