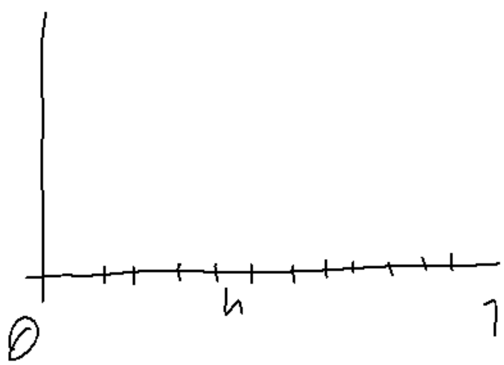
Consider the 2nd-order ODE given by

$$-y'' + \sigma y = f \qquad y(0) = \alpha, \; y(1) = \beta$$



Write

$$-y''(x_i) \approx \frac{-y(x_{i-1}) - 2y(x_i) - y(x_{i+1})}{h^2}$$

$$y'(x_i) \approx \frac{y(x_{i+1}) - y(x_{i-1})}{2h}$$

or

$$y'(x_i) \approx \frac{y(x_{i+1}) - y(x_i)}{}$$

$1 \qquad\qquad h$

we had

$$-y'' + \sigma y' = f$$

$$\frac{y_{i-1} + 2y_i - y_{i+1}}{h^2} + \sigma\left(\frac{y_{i+1} - y_{i-1}}{2h}\right) = f_i$$

for $i = 1, \dots, N$

This is the difference approximation to the solution

$$-\left(1 + \frac{\sigma h}{2}\right) y_{i+1} + 2y_i - \left(1 - \frac{\sigma h}{2}\right) y_{i-1} = h^2 f_i$$

with $f_i = f(x_i)$

Just by rearranging. Now we have a system of linear equations

This gives us a tridiagonal matrix

$$\begin{pmatrix} a & b & & \\ c & \ddots & \ddots & \\ & \ddots & \ddots & b \\ & & c & a \end{pmatrix} \begin{pmatrix} \\ y \\ \\ \end{pmatrix} = g$$

Ty =g

(constant diagonal is called" Toeplitz")

We had $-y'' + \sigma y' = f$; note that when $\sigma = 0$, the tridiagonal matrix is symmetric

What can be said about

$|y(x_i) - y_i|$?

How good is our a?

In the future we will talk about matrices of the form

$$\begin{pmatrix} A & C & & 0 \\ B & A & \ddots & \\ & \ddots & \ddots & C \\ 0 & & B & A \end{pmatrix}$$

Block Toeplitz Matrix

With all matrices p x p

Data Fitting

Suppose we have a data set

$$\{x_i, y_i\}_{i=1}^{n}$$

And we want to approximate the data by a polynomial

$$p_\ell(x) = a_0 + a_1 x + \cdots + a_\ell x^\ell$$

Do the following:

$$\sum_{i=1}^{n} \left(y_i - p_\ell(x_i)\right)^2 = \min$$

A classic least-squares problem

Construct a matrix X like the following;

$$X = \begin{pmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^\ell \\ 1 & x_2 & x_2^2 & \cdots & x_2^\ell \\ & \vdots & & & \\ 1 & x_n & x_n^2 & \cdots & x_n^\ell \end{pmatrix}$$

It has dimensions n x (l+1)

Also have a vector

$$\tilde{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

In matrix terms, we are looking at

$$(\tilde{y} - Xa)^T (y - Xa) =: \varphi(a)$$

Wednesday, September 28, 2005
11:17 AM

$$\varphi(a) = (\bar{y} - X\bar{a})^T(\bar{y} - Xa)$$

$$= y^T y - y^T X a - a^T X^T y$$

$$+ a^T X^T X a$$

We must have $y^T X a = a^T X^T y$

So the above is equivalent to

$$y^T y - 2a^T x^T y + a^t x^T x a$$

look at the gradient

$$d\varphi/da = -2X^T y + 2X^T X a$$

set $\text{grad}\, \varphi = 0$, we get

$$X^T X a = X^T y$$

Called the "normal equations"

Consider a new vector x^r given by

$$x^r = \begin{pmatrix} x_1^r \\ | \\ x^r \end{pmatrix}$$

not sure

$$\begin{pmatrix} \vdots \\ x_n^r \end{pmatrix}$$

not sure

Now, note that

$$(X^T X)_{i+1,j+1} = (x^{i+1})^T x^{j+1}$$

$$= \sum_{k=1}^{n} x_k^{i+j} =: \mu_{i+j}$$

In other words, the (i+1,j+1) element of $X^T X$ only depends on those i+1, j+1 elements this gives

$$X^TX = \begin{pmatrix} \mu_0 & \mu_1 & \mu_2 & \cdots & \mu_\ell \\ \mu_1 & & & & \\ \mu_2 & & & & \\ \vdots & & & & \end{pmatrix}$$

this is called a "Hankel" matrix

It remains to show that $\varphi(\hat{a})$ is the minimum, but we won't do it now

In general, the above method described is actually awful for data fitting
--e.g. do we really want to fit a single polynomial over the whole interval? That's why ore might use splines

Scaling/centering

Wednesday, September 28, 2005
11:29 AM

If the x_i's vary greatly, we might want to write

$$p_\ell(x) = a_0 + a_1(x-\mu) + a_2(x-\mu)^2 + \cdots + a_\ell(x-\mu)^\ell$$

Maybe write this instead as

$$p_\ell(x) = b_0 q_0(x) + \cdots + b_\ell q_\ell(x)$$

$q_k(x)$: polynomial of degree k

i.e. choose a different basis of polynomials (orthogonal)

Not sure

$$\sum_{i=1}^{k} q_r(x) q_s(x) = 0, \quad r \neq s$$

We'll have matrices of the form

$$X = \begin{pmatrix} & & \\ & & \\ & & \end{pmatrix}$$

$$Q = \begin{pmatrix} q_0(x_1) & q_1(x_1) & \cdots & q_\ell(x_1) \\ \vdots & & & \vdots \\ q_0(x_n) & q_1(x_n) & \cdots & q_\ell(x_n) \end{pmatrix}$$

Now, instead of forming

$$X^T X a = X^T y$$

we'll have

$$Q^T Q b = Q^T y$$

What's true about Q^TQ? we have

$$Q^TQ = \begin{pmatrix} \ddots & 0 \\ 0 & \ddots \end{pmatrix}$$

by orthogonality

Wednesday, September 28, 2005
11:35 AM

Another aspect of this problem (a dynamic situation) say

$$X_l = \left(\hat{x}^0, \tilde{x}^1, \ldots, \tilde{x}^l \right)$$

$$X_{l+1} = \left(X_l \ \vdots\ \tilde{x}^{l+1} \right)$$

$$X_{l+1}^T X_{l+1} \tilde{a} = X_{l+1}^T y$$

Now we have to recompute all the $\tilde{a}$'s
--but if we choose orthogonal Q, then:

$q_{l+1}(x)$

$$Q_{l+1} = \left(q_0(x), q_1(x), \ldots, q_{l+1}(x) \right)$$

therefore, when we solve for the new system of equations,

$$Q_{l+1}^T Q_{l+1} b_{l+1} = Q_{l+1}^T y$$

$$Q_{l+1}' Q_{l+1} b_{l+1} = Q_{l+1}' y$$

So we don't have to change the old coefficients

Often, we have constraints, like

$$\tilde{a} \geq \tilde{0}$$

$$C^T \tilde{a} = \tilde{0}$$

or even constraining length:

$$a^T a = \alpha^2$$

Anyway, we wanted to minimize

$$\varphi(a) = \sum_{i=1}^{n} \left(y_i - p_k(x_i)\right)^2$$

or maybe

$$\varphi(a) = \sum_{i=1}^{n} \left|y_i - p_k(x_i)\right|$$

or

$$\min \max \left| y_i - p_k(x_i)\right|$$

Start with a vector $\tilde{x}$, then find a function satisfying

$$f(\tilde{x}) \geq 0$$
$$f(\tilde{x}) = 0 \Leftrightarrow \tilde{x} = \tilde{0}$$
$$f(\alpha\tilde{x}) = \alpha f(\tilde{x})$$
$$f(\tilde{x} + \tilde{y}) \leq f(\tilde{x}) + f(\tilde{y})$$

Such a function is called a *norm*

Some examples:

$$nu \rightarrow \quad \gamma(\tilde{x}) = \sum_{i=1}^{n} |x_i|$$

Is this a norm? Yes

The p^{th} norm, $\|\tilde{x}\|_p$, is given by

$$\left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p}$$

We have

$$\|\tilde{x}\|_p = \left(\sum_{i=1}^{k} |x_i|^p\right)^{1/p} \leq \left(n \max |x_i|^p\right)^{1/p}$$
$$\rightarrow \max |x_i|$$

On the other hand,

$$\|\tilde{x}\|_p \geq \left(\max |x_i|^p\right)^{1/p} = \max |x_i|$$

So $\|\tilde{x}\|_\infty = \max |x_i|$

In this class we do 1,2, ∞; the ∞
norm is sometimes called the Chebyshev norm

One can also do a weighted norm, in which entries
are weighted componentwise:

$$\left(\sum_{i=1}^{n} w_i |x_i|^p \right)^{1/p}$$

The "energy norm"

$$\left(\tilde{x}^T A \tilde{x} \right)^{1/2}$$

$$\|x+y+z\| \le \|x+y\| + \|z\| \le \|x\| + \|y\| + \|z\|$$

Etc...

Differences and Norms

Wednesday, September 28, 2005
12:04 PM

Note that

$$\|\tilde{x}\| = \|(x-y)+y\| \leq \|x-y\| + \|y\|$$

so $\|x-y\| \geq \|x\| - \|y\|$

and

$$\|y-x\| \geq \|y\| - \|x\|$$

so

$$\|x-y\| \geq \big|\, \|x\| - \|y\| \,\big|$$

Equivalence of Norms

Wednesday, September 28, 2005
12:06 PM

Say we have two norms

$$\|x\|_\alpha, \ \|x\|_\beta$$

Suppose

$$\|\tilde{x}\|_\alpha \leq \|\tilde{x}\|_\beta \leq c_2 \|\tilde{x}\|_\alpha$$

for all $\tilde{x}$

For example, the 2-norm and ∞-norm

$$\left(\sum_{i=1}^{n} |x_i|^2\right)^{1/2} \leq n^{1/2} \max |x_i| = n^{1/2} \|\tilde{x}\|_\infty$$

The Cauchy-Schwarz inequality

$$|x^T y| \leq \|\tilde{x}\|_{?} \cdot \|\tilde{y}\|_2$$

We talk about absolute and relative error:

$$\|x - y\| \quad \text{vs} \quad \frac{\|x - y\|}{\|x\|}$$

Also pointwise

$x_i - y_i = z_i$ then take $\|\tilde{z}\|,$

$$\frac{x_i - y_i}{x_i} = z_i \,,\quad \text{then take} \quad \|\tilde{z}\|_1$$

Next time: matrix norms