Lecture 26 - Polynomial Approximations & Least Squares

Note Title 4/26/2008

I. Approximations in $C([a,b])$ via polynomials.

Recall that if $\mathcal{U} = \text{Span}(\{\bar{u}_1, \dots, \bar{u}_n\}) \leq V$, then we can easily compute
orthogonal projections:

$$\text{Proj}_{\mathcal{U}} \bar{v} = \text{Proj}_{\bar{u}_1} \bar{v} + \dots + \text{Proj}_{\bar{u}_n} \bar{v}$$

$$= \frac{\langle \bar{u}_1, \bar{v} \rangle}{\langle \bar{u}_1, \bar{u}_1 \rangle} \bar{u}_1 + \dots + \frac{\langle \bar{u}_n, \bar{v} \rangle}{\langle \bar{u}_n, \bar{u}_n \rangle} \bar{u}_n$$

On the other hand, if $U = \text{Span}(\bar{v}_1, \dots, \bar{v}_n)$, then Gram-Schmidt lets us get an orthogonal set w/ same span.

Thus to find the approximation to $f \in C([a,b])$ by poly of deg $\leq n$, have 2 steps:

① Turn $\{1, x, \dots x^n\}$ into an orthogonal set: $\{p_0, p_1, \dots\}$

② Use above formula to find the projection.

<u>Ex</u>: In $C([0,1])$, let's apply ① to $\{1, x, x^2\}$: Gram-Schmidt

$p_0 = 1$ $\quad\quad \langle p_0, p_0 \rangle = \int_0^1 1 \, dx = 1$

$$\langle p_0, x \rangle = \int_0^1 1 \cdot x \, dx = 1/2$$

$$\langle p_0, x^2 \rangle = \int_0^1 1 \cdot x^2 \, dx = 1/3$$

$$p_1 = x - \text{Proj}_{1} x = x - \frac{\langle p_0, x \rangle}{\langle p_0, p_0 \rangle} p_0 = x - 1/2$$

$$\langle p_1, p_1 \rangle = \int_0^1 (x - 1/2)^2 \, dx = 1/12$$

$$\langle p_1, x^2 \rangle = \int_0^1 x^3 - 1/2 \cdot x^2 \, dx = 1/12$$

$$p_2 = x^2 - \text{Proj}_{p_0} x^2 - \text{Proj}_{p_1} x^2 = x^2 - \frac{\langle p_0, x^2 \rangle}{\langle p_0, p_0 \rangle} p_0 - \frac{\langle p_1, x^2 \rangle}{\langle p_1, p_1 \rangle} p_1$$

$$= x^2 - 1/3 - (x - 1/2) = x^2 - x + 1/6$$

So our orthogonal basis for P_2 is $\{1, x - 1/2, x^2 - x + 1/6\}$

A few things to note:
- p_i has degree i
- $\{1, x - 1/2\}$ is an orthogonal basis for P_1

These hold in general for P_n.

Let's use this to find the approx to e^x in P_1 over $(0,1)$:

$$\text{Proj}_{P_1} e^x = \frac{\langle p_0, e^x \rangle}{\langle p_0, p_0 \rangle} p_0 + \frac{\langle p_1, e^x \rangle}{\langle p_1, p_1 \rangle} p_1$$

$$\int_0^1 (x-\tfrac{1}{2})e^x\,dx = \left((x-\tfrac{1}{2})e^x - e^x\right)\Big|_0^1$$

u	dv	
$x-\frac{1}{2}$	e^x	
1	e^x	$+$
0	e^x	$-$

$$= (x-\tfrac{3}{2})e^x\Big|_0^1$$

$$= -\tfrac{e}{2}+\tfrac{3}{2}$$

$$\Rightarrow \text{Proj} = (e-1)\cdot 1 + (\tfrac{3}{2}-\tfrac{e}{2})(x-\tfrac{1}{2})$$

This is different from the power series. It is the linear poly closest to e^x using this notion of distance.

This told us how to approximate a function by poly. Now switch to data.

II. Least Squares

Set-up: more equations than unknowns

$\longleftrightarrow \quad A\bar{x}=\bar{b}, \qquad A\ m\times n, \quad m>n.$

We want to find the closest solution.

Fact If the columns of A are lin ind ($\Leftrightarrow$ rank $A=n$), then $A^t\cdot A$ is nonsingular.

Def - If A is $m\times n$, and $\text{rank}(A)=n$, then the pseudoinverse to A is

$$\text{pinv}(A) = (A^t\cdot A)^{-1}\cdot A^t$$

- The least squares solution to $A\bar{x}=\bar{b}$ is

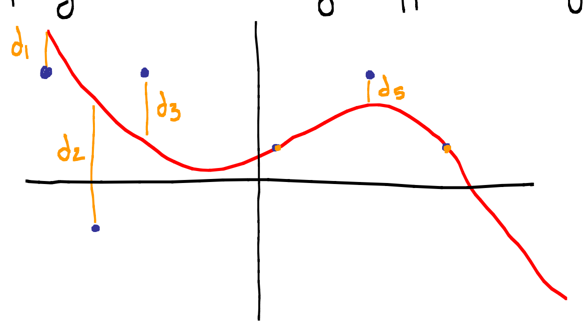
$$\bar{x}^* = \text{pinv}(A)\cdot\bar{b}$$

When A is $n\times n$, $\text{pinv}(A) = (A^t\cdot A)^{-1}\cdot A^t = A^{-1}\cdot A^{t^{-1}}\cdot A^t = A^{-1}$ (For an $n\times n$ matrix: $\text{rank}(A)=n$ $\Leftrightarrow$ A invertible).

If $A\bar{x}=b$ has a solution, $\bar{x}^*$ is the solution. When there is no solution, $\bar{x}^*$ is the point closest to being a solution.

Application: Curve fitting

Goal: find a poly of some deg approximating a set of data points.



$d_1^2+\dots+d_6^2$ gives a "distance" from the curve to our data set

If $\{(x_1,y_1),\dots,(x_n,y_n)\}$ is our data set, then the least squares curve is the poly p of deg at most k s.t.

$$(y_1 - p(x_1))^2 + \dots + (y_n - p(x_n))^2 \quad \text{is minimized.}$$

How does this relate to the above? The data points give a system.

Let $p(x) = a_0 + a_1 x + \dots + a_k x^k$. Our ideal curve passes through all points:

$$\begin{aligned} p(x_1) &= a_0 + a_1 x_1 + \dots + a_k x_1^k = y_1 \\ &\vdots \\ p(x_n) &= a_0 + a_1 x_n + \dots + a_k x_n^k = y_n \end{aligned} \quad \longleftrightarrow \quad \underbrace{\begin{bmatrix} 1 & x_1 & \dots & x_1^k \\ \vdots & \vdots & & \vdots \\ 1 & x_n & \dots & x_n^k \end{bmatrix}}_{A} \cdot \underbrace{\begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_k \end{bmatrix}}_{\bar{x}} = \underbrace{\begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}}_{\bar{b}}$$

If the x_i are all distinct, then the columns are linearly independent. Therefore the best curve is given by $\bar{x}^* = \operatorname{pinv}(A) \cdot \bar{b}$.

Ex

$$\begin{aligned} x + y &= 1 \\ y &= 2 \\ x + y &= 3 \end{aligned} \quad \longleftrightarrow \quad \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

Cols are lin ind $\Rightarrow$ pinv exists:

$$(A^t \cdot A) = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 2 \\ 2 & 3 \end{bmatrix}$$

$$\Rightarrow (A^t \cdot A)^{-1} = \frac{1}{2} \begin{bmatrix} 3 & -2 \\ -2 & 2 \end{bmatrix}$$

$$\Rightarrow \operatorname{pinv}(A) = (A^t \cdot A)^{-1} \cdot A^t = \frac{1}{2} \begin{bmatrix} 3 & -2 \\ -2 & 2 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 & -2 & 1 \\ 0 & 2 & 0 \end{bmatrix}$$

$$\Rightarrow \bar{x}^* = \frac{1}{2} \begin{bmatrix} 1 & -2 & 1 \\ 0 & 2 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 0 \\ 4 \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \end{bmatrix}$$