MATH 54 S215/S205: GROUP ASSIGNMENT FROM MONDAY, MARCH 5

This group assignment was intended to solidify what we've been doing lately. Many of you also saw the power of matrix math on a calculator for the first time. (It's worth learning. For example, the TI-83 and above can do matrix algebra including transposes and inverses, and solve several common problems. Your friend or GSI can show you how. Aside: if you have a TI-89 and you're going to take a physics or chemistry class during your lifetime, the calculator can handle units and physical constants for you—it's a tremendous time-saver!)

> Goal: to approximate $f(x) = e^x$ on the interval $[0, 1]$ by a quadratic polynomial. (Team work and use of calculators strongly encouraged.)

1

Let's start by trying to find a best-fit curve for this sample data:

x	$f(x)$
0.2	1.22
0.4	1.49
0.6	1.82
0.8	2.23

How can finding an approximation $f(x) = a_0 + a_1 x + a_2 x^2 + \text{error}$ be set up as a linear system of equations: $A \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = b$?

For each sample point (x, y), we want $(1)a_0 + (x)a_1 + (x^2)a_2 \approx y$. We get one equation like this for each sample point, yielding the (inconsistent) linear system

$$\underbrace{\begin{bmatrix} 1 & .2 & .04 \\ 1 & .4 & .16 \\ 1 & .6 & .36 \\ 1 & .8 & .64 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix}}_{x} = \underbrace{\begin{bmatrix} 1.22 \\ 1.49 \\ 1.82 \\ 2.23 \end{bmatrix}}_{b}$$

2

Solve the system via the normal equations ($\bar{x} = (A^T A)^{-1} A^T b$).

Retracted when I realized this hadn't been discussed in lecture: Also, use Gram-Schmidt to find the least-squares solution. Which method, if any, is better?

Your answer should be $(a_0, a_1, a_2) = (1.025, .805, .875)$.

You should ask me if you want to see the Gram-Schmidt method. In fact, Hill's section 4.5 is a good treatment of the method, but this is extra material beyond what's assigned for math 54.

3

If time allows, have a teammate figure out how to use more sample points: $x = 0.1, \ldots, 0.9$.

Your teammate would calculate e^x for each of those x values, and then do exactly the same thing as in number 1. To 3 decimal places, the answer should be $(a_0, a_1, a_2) = (1.017, .846, .836)$.

4

A more sophisticated approach is to use *orthogonal projection.* Convince yourself that P_2 is a subspace of $C[0, 1]$, with basis $\{1, x, x^2\}$. (The dot-product is the usual one for $C[0, 1]$: $f \cdot g = \int_0^1 f(x)g(x)dx$.)

Can you use the "normal equation" method to project f orthogonally onto P_2? (This method is a way to project a vector onto $CS(A)$: the projection of x onto $CS(A)$ is Px, where $P = A(A^TA)^{-1}A^T$.)

Can you use an orthonormal basis found via Gram-Schmidt to project f onto P_2?

Use the method that works.

To use the normal equation method, you would have to find a 3-column matrix A whose column space represents the subspace P_2 of $C[0, 1]$. So, in order to write down this matrix A, you would need a basis of $C[0, 1]$. But that's an *infinite-dimensional* vector space, so this method will not work!

However, you can find an orthonormal basis for P_2 via the Gram-Schmidt process. Start with the basis consisting of $v_1 = 1$, $v_2 = x$, $v_3 = x^2$. I illustrate two methods, which are (for all practical purposes) the same. Method I is often easier to compute by hand, but you may prefer computing these things by calculator. (I do.)

Method I: orthogonalize first, normalize later.

$$\begin{aligned}
u_1 &\leftarrow v_1 && = 1\\
u_2 &\leftarrow v_2 - \frac{v_2 \cdot u_1}{u_1 \cdot u_1} u_1 && = x - \frac{\int_0^1 (x)(1)dx}{\int_0^1 (1)(1)dx}(1)\\
& && = x - 1/2\\
u_3 &\leftarrow v_3 - \frac{v_3 \cdot u_1}{u_1 \cdot u_1} u_1 - \frac{v_3 \cdot u_2}{u_2 \cdot u_2} u_2 && = x^2 - \frac{\int_0^1 (x^2)(1)dx}{\int_0^1 (1)^2 dx}(1) - \frac{\int_0^1 (x^2)(x-1/2)dx}{\int_0^1 (x-1/2)^2 dx}(1)\\
& && = x^2 - 1/3 - \frac{1/12}{1/12}(x - 1/2) = x^2 - x + 1/6
\end{aligned}$$

$$\begin{aligned}
e_1 &\leftarrow u_1/\|u_1\| && = \frac{1}{\left[\int_0^1 (1)^2 dx\right]^{1/2}} && = 1\\
e_2 &\leftarrow u_2/\|u_2\| && = \frac{x - 1/2}{\left[\int_0^1 (x-1/2)^2 dx\right]^{1/2}} && = \sqrt{3}(2x - 1)\\
e_1 &\leftarrow u_1/\|u_1\| && = \frac{x^2 - x + 1/6}{\left[\int_0^1 (x^2 - x + 1/6)^2 dx\right]^{1/2}} && = \sqrt{5}(6x^2 - 6x + 1)
\end{aligned}$$

Method II: normalize first.

$$
\begin{aligned}
u_1 \leftarrow & v_1 && = 1 \\
e_1 \leftarrow & u_1/\|u_1\| && = \frac{1}{\left[\int_0^1 (1)^2 dx\right]^{1/2}} \\
& && = 1 \\
u_2 \leftarrow & v_2 - (v_2 \cdot e_1)e_1 && = x - \left[\int_0^1 (x)(1)dx\right](1) \\
& && = x - 1/2 \\
e_2 \leftarrow & u_2/\|u_2\| && = \frac{x-1/2}{\left[\int_0^1 (x-1/2)^2 dx\right]^{1/2}} \\
& && = \sqrt{3}(2x-1) \\
u_3 \leftarrow & v_3 - (v_3 \cdot e_1)e_1 - (v_3 \cdot e_2)e_2 && = x^2 - \left[\int_0^1 (x^2)(1)dx\right](1) \\
& && \quad - \left[\int_0^1 (x^2)(\sqrt{3}(2x-1))dx\right]\sqrt{3}(2x-1) \\
& && = x^2 - 1/3 - 1/6 \cdot 3^1(2x-1) = x^2 - x + 1/6 \\
e_3 \leftarrow & u_3/\|u_3\| && = \frac{x^2 - x + 1/6}{\left[\int_0^1 (x^2 - x + 1/6)^2 dx\right]^{1/2}} \\
& && = \sqrt{5}(6x^2 - 6x + 1)
\end{aligned}
$$

Either way, you get an orthonormal basis for P_2: $\{e_1, e_2, e_3\}$ with $e_1(x) = 1$, $e_2(x) = \sqrt{3}(2x-1)$, and $e_3(x) = \sqrt{5}(6x^2 - 6x + 1)$.

The e_j-component of f is just $f \cdot e_j$—that's what's nice about orthonormal bases. So the component of f lying in the subspace P_2 is

$$
\begin{aligned}
\sum_{j=1}^{3}(f \cdot e_j)e_j(x) = & \left[\int_0^1 (1)e^x dx\right](1) \\
& + \left[\int_0^1 \sqrt{3}(2x-1)e^x dx\right]\sqrt{3}(2x-1) \\
& + \left[\int_0^1 \sqrt{5}(6x^2-6x+1)e^x dx\right]\sqrt{5}(6x^2-6x+1)
\end{aligned}
$$

$$= (e-1)(1) + 3(3-e)(2x-1) + 5(7e-19)(6x^2-6x+1) \approx 1.013 + .851x + .839x^2.$$

5

> So, you could approximate $f(x)$ with any of the answers in problems 2–4. You could also use a second-order Taylor series. Which would you expect to best meet the goal of approximating $f(x)$ on the interval $[0, 1]$?

What it means to be a "good" approximation is open to interpretation, and might depend on how the approximation is being used. That said, you'd have good reason to expect that taking a larger (uniform) sample would improve the answer; in other words, #3 should give you a better answer than #2.

Compare this to #4: in #3, the least-squares answer minimizes $\epsilon(.1)^2 + \cdots + \epsilon(.9)^2$, where $\epsilon(x)$ is the error. So, in a sense, the approximation in #3 was found by controlling the error at nine points. But in #4 we're concerned with the entire function rather than a few of its values. In fact, by choosing the orthogonal projection of f onto P_2, we are minimizing the "length" of the error as measured by the dot-product $\epsilon \cdot \epsilon = \int_0^1 \epsilon(x)^2 dx$. This explains why the process is so valuable: in doing #4, you have effectively gone from controlling the error at nine points to controlling the error at *all points at once*, even though the computational steps of #4 are faster than those of #3!

So what about Taylor series? Aren't they supposed to be the best polynomial approximations possible? Well, yes and no—if you take a Taylor series centered at $x = x_0$, you get the best possible approximation to $f(x)$ *for x close to x_0*. Finding the best possible approximation to $f(x)$ on the whole interval $[0, 1]$ is a different problem. (For the record, the Taylor series centered at $x = 1/2$ is $1.65 + 1.65(x - .5) + .824(x - .5)^2 = 1.03 + .824x + .824x^2$.)