Numerical Analysis – Homework 5

James Diffenderfer

February 28, 2017

Exercise 1. Find the value of α which minimizes

$$F(\alpha) = \int_{-1}^{1} \left(x^2 - (x+\alpha)\right)^2 \, dx.$$

Be sure to show that your solution is, in fact, the global minimum by using an additional test.

Proof.

$$\begin{aligned}
F(\alpha) &= \int_{-1}^{1} \left(x^2 - (x+\alpha)\right)^2 \, dx \\
&= \int_{-1}^{1} \left(x^4 - 2x^3 + x^2(1-2\alpha) + 2x\alpha + \alpha^2\right) \, dx \\
&= \int_{-1}^{1} \left(x^4 + x^2(1-2\alpha) + \alpha^2\right) \, dx + \int_{-1}^{1} \left(-2x^3 + 2x\alpha\right) \, dx \\
&= 2\int_{0}^{1} \left(x^4 + x^2(1-2\alpha) + \alpha^2\right) \, dx \\
&= \frac{16}{15} - \frac{4}{3}\alpha + 2\alpha^2.
\end{aligned}$$

Solving $F'(\alpha) = 4\alpha - \frac{4}{3} = 0$ yields $\alpha = \frac{1}{3}$. Since $F'(x) < 0$ for $x \in \left(-\infty, \frac{1}{3}\right)$ and $F'(x) > 0$ for $x \in \left(\frac{1}{3}, \infty\right)$ we conclude that $\alpha = \frac{1}{3}$ is the value at which $F(\alpha)$ attains a global minimum. ☐

Exercise 2. For each n, let $x_0^{(n)}, x_1^{(n)}, \ldots, x_n^{(n)}$ be the zeros of the n^{th} order Chebychev polynomial T_n on $[-1,1]$ and $Q_n(x)$ be the degree n interpolating polynomial to e^x on $[-1,1]$ with the nodes $x_0^{(n)}, x_1^{(n)}, \ldots, x_n^{(n)}$. Show that

$$\lim_{n\to\infty} \|Q_n - e^x\|_\infty = 0.$$

Proof. Recall that there exists a $\eta \in [-1,1]$ such that

$$Q_n(x) - e^x = \frac{e^\eta}{(n+1)!}(x-x_0)(x-x_1)\cdots(x-x_n) = \frac{e^\eta}{(n+1)!}T_n(x),$$

for $x \in [-1,1]$. Hence,

$$\|Q_n(x) - e^x\|_\infty = \max_{x\in[-1,1]} \frac{|e^x T_n(x)|}{(n+1)!}.$$

Since $\max_{x\in[-1,1]} e^x = e$ and $\max_{x\in[-1,1]} |T_n(x)| = 1$ we have that

$$\|Q_n(x) - e^x\|_\infty \le \frac{e}{(n+1)!},$$

for all $x \in [-1,1]$. Thus,

$$\lim_{n\to\infty} \|Q_n(x) - e^x\|_\infty \le \lim_{n\to\infty} \frac{e}{(n+1)!} = 0,$$

for all $x \in [-1,1]$, which concludes the proof. ☐

Exercise 3. Find the second order (i.e., quadratic) least squares approximation to the function e^x on the interval $[1,3]$ with respect to the weight $\alpha(x) \equiv 1$ using the fact that the first three orthonormal polynomials on $[-1,1]$ with respect to the weight $\alpha(x) \equiv 1$ are

$$\phi_0(x) = \sqrt{1/2}, \quad \phi_1(x) = \sqrt{3/2}\, x, \quad \text{and} \quad \phi_2(x) = \sqrt{5/8}\, (3x^2 - 1).$$

Proof. For $t \in [-1,1]$ define $x = t + 2$. Using

$$\phi_0(t) = \sqrt{1/2}, \quad \phi_1(t) = \sqrt{3/2}\, t, \quad \text{and} \quad \phi_2(t) = \sqrt{5/8}\, (3t^2 - 1)$$

and by defining $f(t) = e^{t+2}$ we have

$$p_2(t) = \langle f(t), \phi_0(t)\rangle \phi_0(t) + \langle f(t), \phi_1(t)\rangle \phi_1(t) + \langle f(t), \phi_2(t)\rangle \phi_2(t).$$

Since

$$\langle f(t), \phi_0\rangle = e^2 \sqrt{\frac{1}{2}} \int_{-1}^{1} e^t \, dt \approx 12.28050385,$$

$$\langle f(t), \phi_1\rangle = e^2 \sqrt{\frac{3}{2}} \int_{-1}^{1} te^t \, dt =\approx 6.6584034568,$$

and

$$\langle f(t), \phi_2\rangle = e^2 \sqrt{\frac{5}{8}} \int_{-1}^{1} (3t^2 - 1)e^t \, dt \approx 1.6721557017,$$

we have that

$$\begin{aligned} p_2(t) &= 12.28050385\sqrt{\frac{1}{2}} + 6.6584034568\sqrt{\frac{3}{2}}\, t + 1.6721557017\sqrt{\frac{5}{8}}\, (3t^2 - 1) \\ &= 3.96587t^2 + 8.15485t + 7.36167. \end{aligned}$$

Substituting $t = x - 2$ we find that

$$p_2(x) = 3.96587(x-2)^2 + 8.15485(x-2) + 7.36167$$

which can be simplified to

$$p_2(x) = 3.96587x^2 - 7.70863x + 6.91545.$$

□

Exercise 4. Programming Exercise:

(a) Write a program that computes the linear least squares approximation for a data set $(x_1, y_1), \ldots, (x_m, y_m)$.

(b) Generate 20 data points for $k = 1, \ldots, 20$ with $x_k = k/2$ and $y_k = 5x_k + 2 + \varepsilon_k$, where ε_k is a random number uniformly distributed in $[-2, 2]$.

(c) Use your program to find the best linear least squares fit to your data $(x_1, y_1), \ldots, (x_{20}, y_{20})$.

(d) How close is your computed answer to the expected one? Why is the estimate of the slope so much better than that for the intercept?

Proof. The output of the function for one of the randomly generated data sets was

```
y-int:  1.7809335  and  slope:  4.9832468
```

which had the plot corresponding to Figure 1.

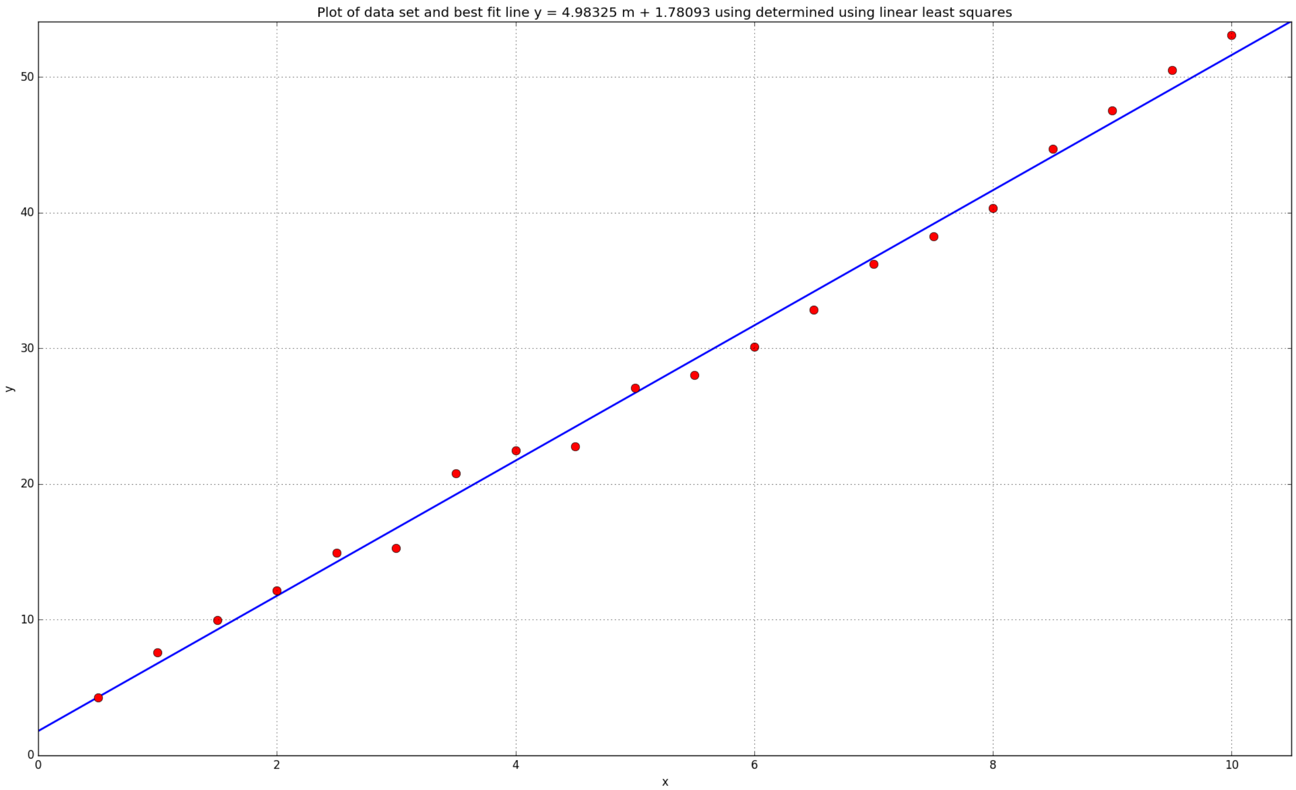

Figure 1: Plot of twenty data points and line of best fit determing using linear least squares.

Suppose now that we use the same set up for the data points and we let n be the number of data points used. Fix n and let $\varepsilon = \sum_{k=1}^{n} \varepsilon_k$ and $\delta = \sum_{k=1}^{n} x_k \varepsilon_k$. Since $\sum_{k=1}^{n} x_k = \frac{1}{4}n^2 + \frac{1}{4}n$, $\sum_{k=1}^{n} x_k^2 = \frac{1}{12}n^3 + \frac{1}{8}n^2 + \frac{1}{24}n$, $\sum_{k=1}^{n} y_k = \frac{5}{4}n^2 + \frac{13}{4}n + \varepsilon$ and $\sum_{k=1}^{n} = x_k y_k = \frac{5}{12}n^3 + \frac{9}{8}n^2 + \frac{17}{24}n + \delta$ we have that

$$\begin{aligned} a_1 &= \frac{\left(\sum_{k=1}^{n} x_k\right)\left(\sum_{k=1}^{n} y_k\right) - n\sum_{k=1}^{n} x_k y_k}{\left(\sum_{k=1}^{n} x_k\right)^2 - n\sum_{k=1}^{n} x_k^2} \\ &= \frac{\left(\frac{1}{4}n^2 + \frac{1}{4}n\right)\left(\frac{5}{4}n^2 + \frac{13}{4}n + \varepsilon\right) - n\left(\frac{5}{12}n^3 + \frac{9}{8}n^2 + \frac{17}{24}n + \delta\right)}{\left(\frac{1}{4}n^2 + \frac{1}{4}n\right)^2 - n\left(\frac{1}{12}n^3 + \frac{1}{8}n^2 + \frac{1}{24}n\right)} \\ &= \frac{-\frac{5}{48}n^4 + \frac{1}{4}n^2\varepsilon + \frac{5}{48}n^2 - \delta n + \frac{1}{4}\varepsilon n}{\frac{1}{48}n^2 - \frac{1}{48}n^4} \\ &= \frac{5n^3 - (12\varepsilon + 5)n - 12\varepsilon + 48\delta}{n^3 - n} \end{aligned}$$

and

$$\begin{aligned} a_0 &= \frac{\left(\sum_{k=1}^{n} y_k\right) - a_1\sum_{k=1}^{n} x_k}{n} \\ &= \frac{\left(\frac{5}{4}n^2 + \frac{13}{4}n + \varepsilon\right) - \frac{5n^3 - (12\varepsilon+5)n - 12\varepsilon + 48\delta}{n^3 - n}\left(\frac{1}{4}n^2 + \frac{1}{4}n\right)}{n} \\ &= \frac{2n^2 + 2(2\varepsilon - 1)n + 2\varepsilon - 12\delta}{n^2 - n} \end{aligned}$$

However, by increasing the number of data points we observe that

Number of Data Points	y-intercept	Slope
20	1.7809335	4.9832468
200	1.9579963	5.0024152
2000	2.0229506	4.9999168
20000	1.9870547	5.0000045
200000	2.0062279	4.9999999
2000000	1.9992016	5

□

Exercise 5. Derive the central three point formula

$$f'(x) = \frac{f(x+h) - f(x-h)}{2h} - \frac{f^{(3)}(\eta)h^2}{6}$$

for some $\eta \in [x-h, x+h]$ using a pair of Taylor polynomials with remainders in the same fashion as the derivation of the three point formula for $f''(x)$ done in class.

Proof. Suppose $f \in C^3(\mathbb{R})$. Fix $x \in \mathbb{R}$ and $h > 0$. By Taylor's Theorem, there exists a $\eta_1 \in [x, x+h]$ such that

$$f(x+h) = f(x) + f'(x)h + \frac{1}{2}f''(x)h^2 + \frac{1}{6}f'''(\eta_1)h^3. \tag{1}$$

Similarly, by Taylor's theorem there exists a $\eta_2 \in [x-h, x]$ such that

$$f(x-h) = f(x) - f'(x)h + \frac{1}{2}f''(x)h^2 - \frac{1}{6}f'''(\eta_1)h^3. \tag{2}$$

Taking the difference of (1) and (2) we find that

$$f(x+h) - f(x-h) = 2f'(x)h + \frac{f'''(\eta_1) + f'''(\eta_2)}{6}h^3.$$

Hence, since $h > 0$ we have that

$$f'(x) = \frac{f(x+h) - f(x-h)}{2h} + \frac{f'''(\eta_1) + f'''(\eta_2)}{12}h^2. \tag{3}$$

Since $f \in C^3(\mathbb{R})$ and $(f'''(\eta_1) + f'''(\eta_2))/2$ is between $f'''(\eta_1)$ and $f'''(\eta_2)$, by the Intermediate Value Theorem there exists a $\eta \in (\eta_2, \eta_1) \subseteq [x-h, x+h]$ such that

$$f'''(\eta) = \frac{f'''(\eta_1) + f'''(\eta_2)}{2}. \tag{4}$$

Substituting (4) into (3) yields

$$f'(x) = \frac{f(x+h) - f(x-h)}{2h} + \frac{f'''(\eta)}{6}h^2, \tag{5}$$

for some $\eta \in [x-h, x+h]$, the desired result. □