5.4: #30) Fit a linear function of the form \( f(t) = c_0 + c_1 t \) to the data points \((0, 0), (0, 1), (1, 1)\), using least squares. Sketch your solution.

**Solution:** Two points determine a line, so it is unlikely that any line actually passes through all three points. Regardless, the conditions that our function goes through each of these points can be expressed as

\[
\begin{align*}
    c_0 + 0c_1 &= 0, \\
    c_0 + 0c_1 &= 1, \\
    c_0 + c_1 &= 1
\end{align*}
\]

We can express this as a single matrix equation

\[
\begin{pmatrix}
1 & 0 \\
1 & 0 \\
1 & 1
\end{pmatrix}
\begin{pmatrix}
c_0 \\
c_1
\end{pmatrix}
=
\begin{pmatrix}
0 \\
1 \\
1
\end{pmatrix}
\]

We are asked to solve this using least squares. The normal equation is

\[
\begin{pmatrix}
3 & 1 \\
1 & 1
\end{pmatrix}
\begin{pmatrix}
c_0 \\
c_1
\end{pmatrix}
=
\begin{pmatrix}
2 \\
1
\end{pmatrix}
\]

This gives a solution of \( c_0 = c_1 = \frac{1}{2} \). Therefore, our least squares fit gives \( f(t) = \frac{1}{2} + \frac{1}{2} t \). A picture of all of this can be found below:
5.4: #32) Fit a quadratic polynomial to the points (0,27), (1,0), (2,0), (3,0). Sketch the solution.

Solution: Three points determine a parabola, so it is unlikely that any parabola actually passes through all four of our points. Regardless, let us write out the conditions that our parabola passes through each point. Let the parabola be of the form $f(t) = at^2 + bt + c$. Then,

\[ 0a + 0b + c = 27, \quad a + b + c = 0, \quad 4a + 2b + c = 0, \quad 9a + 3b + c = 0 \]

Expressed as a matrix equation, we have

\[
\begin{pmatrix}
0 & 0 & 1 \\
1 & 1 & 1 \\
4 & 2 & 1 \\
9 & 3 & 1 \\
\end{pmatrix}
\begin{pmatrix}
a \\
b \\
c \\
\end{pmatrix}
=
\begin{pmatrix}
27 \\
0 \\
0 \\
0 \\
\end{pmatrix}
\]

The normal equation is

\[
\begin{pmatrix}
98 & 36 & 14 \\
36 & 14 & 6 \\
14 & 6 & 4 \\
\end{pmatrix}
\begin{pmatrix}
a \\
b \\
c \\
\end{pmatrix}
=
\begin{pmatrix}
0 \\
0 \\
27 \\
\end{pmatrix}
\]

Solving (via row reduction), we find that

\[ a = \frac{27}{4}, \quad b = -\frac{567}{20}, \quad c = \frac{513}{20} \quad \Rightarrow \quad f(t) = \frac{27}{4}t^2 - \frac{567}{20}t + \frac{513}{20} \]

\[ f(t) = \frac{27}{4}t^2 - \frac{567}{20}t + \frac{513}{20} \]

\[ y \]

\[ x \]

\[ -1 \quad 1 \quad 2 \quad 3 \quad 4 \]

\[ -10 \quad 10 \quad 20 \quad 30 \]
5.4: #36) Fit a function of the form \( f(t) = a + b \sin \left( \frac{2\pi}{366} t \right) + c \cos \left( \frac{2\pi}{366} t \right) \) to the points (32,10), (77,12), (121,14), (152,15). Find the maximum value of this function. Is it close to \( 15 + \frac{13}{60} + \frac{39}{3600} \approx 15.2275 \) ?

Solution: The conditions that our function passes through these points are

\[
\begin{align*}
1.5221b + .8529c &= 10, \\
1.9692b + .2464c &= 12, \\
1.8745b - .4851c &= 14, \\
1.5074b - .8617c &= 15
\end{align*}
\]

As a matrix equation, we have

\[
\begin{pmatrix}
1 & .5221 & .8529 \\
1 & .9692 & .2464 \\
1 & .8745 & -.4851 \\
1 & .5074 & -.8617
\end{pmatrix}
\begin{pmatrix}
a \\
b \\
c
\end{pmatrix}
=
\begin{pmatrix}
10 \\
12 \\
14 \\
15
\end{pmatrix}
\]

The normal equation is (please don’t compute this by hand)

\[
\begin{pmatrix}
4 & 2.87 & -.248 \\
2.87 & 2.23 & -.177 \\
-.248 & -.177 & 1.766
\end{pmatrix}
\begin{pmatrix}
a \\
b \\
c
\end{pmatrix}
=
\begin{pmatrix}
51 \\
36.71 \\
-8.23
\end{pmatrix}
\]

Solving yields

\[
a = 12.261, \quad b = .431, \quad c = -2.900 \implies f(t) = 12.261 + .431 \sin \left( \frac{2\pi}{366} t \right) - 2.9 \cos \left( \frac{2\pi}{366} t \right)
\]

The max of this guy was found by computer (for example wolframalpha.com). We get \( t = 174.404 \), and a max of \( \frac{15.1922}{3600} \approx 15.2275 \). We now have an estimate of the longest day of the year in Rome. Also, a picture for the curious:

Actually a very good fit!

---

1 Not a repeating decimal, surprisingly!
5.4: #40) Fit a function of the form $D(a) = ka^n$ to the points (.387,.241), (1,1), (5.203,11.86), (19.19,84.04), (39.53,248.6). Comment on Kepler’s laws if you happen to know them. Also, explain why $k$ should be close to 1.

Solution: Things aren’t quite as straightforward this time. The condition our function pass through the point (.387,.241) is

$$k \cdot (.386)^n = .241$$

This isn’t a nice linear equation like we had before, and so the least-squares method cannot be applied. However, we can take a hint from problem 39. For a moment, pretend that there is some function $D(a)$ that passes through all five data points. Then, $\ln(D(a))$ would be a function that passes through the same points, but with the natural log taken of their $y$ coordinates.

In other words, $\ln(D(a)) = \ln(ka^n) = \ln(k) + n \ln(a)$ would pass through the points (.387,-1.42), (1,0), (5.203,2.47), (19.19,4.43), (39.53,5.52). However, notice what the condition that $\ln(D(a))$ passes through the point (.387,-1.42) looks like:

$$\ln(k) + n \ln(.387) = -1.42$$

That is a linear equation in the variables $n$ and $\ln(k)$. So, we can apply least squares to fit the function $\ln(D(a))$, and then recover a good fit for $D(a)$ from this.\(^2\)

If we ask the computer for the answer, we get

$$\ln(k) = .0003, \quad n = 1.500, \quad \Longrightarrow \quad \ln(D(a)) = .0003 + 1.5 \ln(a) \quad \Longrightarrow \quad D(a) = 1.0003 \cdot a^{1.5}$$

I will not comment on Kepler’s laws here because it’s not that relevant to the math. Our function does fit them very nicely, however.

On the other hand, the reason $k$ should be 1 is because one of our data points is Earth. Because we have normalized everything to Earth standards, our function must obey $D(1) = 1$, which tells us that $k = 1$.

\(^2\)The resulting $D(a)$ will NOT be a least squares fit to the data. If you unwrap what we are doing, we are not minimizing the sum of the squares of the vertical distances from our function to the points. We are minimizing the sum of the squares of the differences between the logs of their $y$ coordinates. You may then wonder why we are using this method to find $D(a)$. It’s because it is actually possible to do the computation this way.