LARSON—MATH 511—CLASSROOM WORKSHEET 17
Least Squares

Sage/CoCalc

1. (a) Start the Chrome browser.
   (b) Go to `http://cocalc.com`
   (c) Login (likely using **your VCU email address**).
   (d) You should see an existing Project for our class. Click on that.
   (e) Click "New", then "Sage Worksheet", then call it **c17**.

Least Squares

If you have two-dimensional *data*, a collection of points (x_i, y_i), it is rarely the case that there is a line that passes through all of the points. One thing we might do is find a line that minimizes the sum of the squares of the difference between points on the line and the y-coordinates of the points. This is equivalent to finding a projection of the y-coordinates of the given points to the closest linear combination of some collection of vectors.

Here's a simple example. Consider the points $\{(0,0),(1,2),(2,1)\}$. We want a line $y = Cx + D$ that passes through these points exactly if there is one and otherwise is this *least squares* line.

Substituting in the points, we want:

$$C * 0 + D = 0,$$

$$C * 1 + D = 2,$$

$$C * 2 + D = 1.$$

No line passes through all of these points. Let's write a matrix equation of the form: $A\hat{x} = \hat{B}$, where the numbers in A and $\hat{B}$ are given, and we'll solve for $\hat{x}$.

The above equations can be written as:

$$\begin{bmatrix} 0 & 1 \\ 1 & 1 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} C \\ D \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix}$$

So we'll let $A = \begin{bmatrix} 0 & 1 \\ 1 & 1 \\ 2 & 1 \end{bmatrix}$, $\hat{x} = \begin{bmatrix} C \\ D \end{bmatrix}$, and $\hat{B} = \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix}$.

2. First we will define matrix A. Evaluate `A=matrix(RDF,3,2,[0,1,1,1,2,1])`. Evaluate `A` to check.

3. We'll also need A^T. Evaluate `At=A.transpose()`. Evaluate `At` to check.

4. We have seen the useful matrix A^TA before. We can find the product of matrices using the regular `*` multiplication operator. Evaluate `AtA=At*A`. Evaluate `AtA` to check.

5. We know AtA is invertible—why? We need its inverse. Evaluate `AtAinv=AtA.inverse()`. Evaluate `AtAinv` to check.

6. Now check that AtA and $AtAinv$ actually are inverses . Evaluate `AtA*AtAinv`. What do you get?

7. We found that, solving for $\hat{x}$, we get $\hat{x} = (A^TA)^{-1}A^T\hat{B}$. So now we need $\hat{B}$. Evaluate `B=matrix(RDF,3,1,[0,2,1])`. Evaluate `B` to check.

8. To find $\hat{x}$, evaluate `x=AtAinv*At*B`. Evaluate `x` to check.

9. What are C and D? Draw the line $y = Cx + D$, as well as the given data points.

Real Data

10. Open your CoCalc project Handouts folder, click on "least_squares.sage". We'll need this file and height-data.csv". You should probably move or copy then to your root/home directory.

11. We will run the code here step-by-step in your c17 worksheet.

Getting your classwork recorded

When you are done, before you leave class...

1. Click the "Make pdf" (Adobe symbol) icon and make a pdf of this worksheet. (If CoCalc hangs, click the printer icon, then "Open", then print or make a pdf using your browser).
2. Send me an email with an informative header like "Math 511—c17 worksheet attached" (so that it will be properly recorded).
3. Remember to attach today's classroom worksheet!