Computation and Computers in Geology

Matrix operations and fitting functions

In the last set of exercises we saw that for a system of observations that were linearly related to their collection in space or time such as:

$y_1 = m*x_1 + b$
$y2 = m*x2 + b$
…
$y_n = m*x_n + b.$

We could write a matrix equation:

$$[Y]=[X][M] \text{ or, simply } \mathbf{Y} = \mathbf{XM}$$

and then solve for the model parameters (M; slope and intercept) with:

$$[\mathbf{X^t}*\mathbf{X}]^{-1}*\mathbf{X}^t*\mathbf{Y} =\mathbf{M}.$$

In the case above we have more observations than unknowns so the system is said to be over determined; this yields the least squares solution. It was also a good way to introduce Excel's matrix capabilities. However, it turns out the Excel also has built in statistical functions including one for least squares, best-fit lines.

Excel has a function, **LINEST()** which calculates the least squares, best-fit line and a set of associated statistics. **LINEST()** calculates the statistics for a line by using the "least squares" method to calculate a straight line that best fits your data, and returns an array that describes the line. Because this function returns an array of values, it must be entered as an array formula (CNTRL-SHIFT-ENTER). Excel has a couple pages of details on **LINEST()** in its' help file – you should look at the help file when you are working with real data or if you have problems with my explanations.

Example: To experiment with **LINEST()** we will first generate some noisy "data" from a linear equation with known model parameters:

- In the upper left hand corner of your spreadsheet enter values for slope and intercept (say 2.5, 2) into cells A1 and A2 respectively.

- In column A starting at A4, generate a series of integers from 1 through 50 (check out "*filling cells, series*" in Excel's help to find an easy way to do this).

- In B4, adjacent to the first integer in column A, enter the formula
  =$A$1*A4+$A$2+RAND()*20-10
  and copy it down to the end of the values in column A. This formula uses absolute references to incorporate the slope and intercept. The formula also uses Excel's **RAND** function to add a little noise, distributed between –10 and 10, to the generated values.

- Now, highlight D4 and E4, then type in the formula =**LINEST**(
  Highlight the Y values (column B)
  put in a comma
  Highlight the X values (column A)

Hit CNTRL-SHIFT-ENTER to enter the matrix equation.
Excel will return the slope and intercept in D4 and E4.

Exercises:

1. Graph the noisy data from above (use markers without lines in a XY plot for the data) and plot the least squares line on the same graph. Read the help file in **LINEST()** and experiment with recovering the rest of the statistical information it will supply.

2. Adapt the equation so that **RAND()** is in a column of its own next to the x values. Now generate three parallel columns by(substituting **RAND**()*10-5; **RAND**()*20-10; **RAND**()*40-20 in the equation above. Compare the recovered model parameters (m, b) for each case. Make independent graphs of the three functions so that you can visually inspect the noise added to the data. Here's a figure to show you how scatter increases with the changes in RAND().

3. **Background for problem #3:**

   Radiometric dating relies on radioactive decay which is a statistical property. The probability that any radioactive parent atom will decay per second is called the decay rate. If the decay rate is λ then in a time interval dt the probability that a given nucleus will decay isλdt. If there are P parent nuclei then Pλdt will decay in time dt. Thus the change in the number of parent nuclei in time dt is:

   $$dP = -P\lambda dt \qquad \text{or}$$

   $$\frac{dP}{P} = -\lambda dt \qquad \text{or, integrating both sides}$$

   $$\int_{P_0}^{P} \frac{dP}{P} = -\lambda \int_0^t dt$$

   which yields the common expressions for radioactive decay:

   $$P = P_0 \exp(-\lambda t) \qquad \text{and} \qquad P_0 = P\exp(\lambda t)$$

   While the parent atoms decrease, the amount of daughter atom increases; D is the difference between P and P_0:

   $$D = P_0 - P$$
   $$D = P\exp(\lambda t) - P$$
   $$D = P(\exp(\lambda t) - 1)$$

   Unfortunately there is usually some unknown initial amount of daughter isotope in the environment so that the measured amount of daughter isotope is the sum of that produced by decay with the initial concentration, D_0:

   $$D = D_0 + P(\exp(\lambda t) - 1)$$

We dismiss the need to know the initial amount of daughter product by using a third isotope to normalize the equations. For example ^{87}Rubidium decays to ^{87}Strontium; non-radiogenic ^{86}Sr has about the same initial abundance as the ^{87}Sr so we use that for normalization. Using ^{87}Rb as P, and ^{87}Sr for D in our last equation then dividing through by ^{86}Sr gives:

$$\frac{^{87}Sr}{^{86}Sr} = \left(\frac{^{87}Sr}{^{86}Sr}\right)_0 + \left(\frac{^{87}Rb}{^{86}Sr}\right) * \left(\exp(\lambda t) - 1\right).$$

This linear equation yields the isochron for the Rubidium-Strontium geochronologic system; λ=0.14201E^{-10}/year in this system. Thus we measure isotopic abundances, plot them in ^{87}Sr/^{86}Sr – ^{87}Rb/^{86}Sr space and the slope of the isochron yields the age of the rock; the intercept yields the initial ratio of ^{87}Sr/^{86}Sr.

OK – Here's problem #3:

Plot the following data and use **LINEST**() to find the age of the rock and the initial ^{87}Sr/^{86}Sr ratio. You'll have to use the slope and {exp(λt)-1} to solve for the age, t:

^{87}Rb/^{86}Sr	^{87}Sr/^{86}Sr
0.45	0.728
0.7	0.737
0.9	0.748
0.95	0.748
1.15	0.759
1.6	0.779
1.65	0.782
2.25	0.808

4. The Sm-Nd system is much like the Rb-Sr system:

$$\frac{^{143}Nd}{^{144}Nd} = \left(\frac{^{143}Nd}{^{144}Nd}\right)_0 + \left(\frac{^{147}Sm}{^{144}Nd}\right) * \left(\exp(\lambda t) - 1\right).$$

However, for this system, the decay constant (λ) is 6.54*10^{-12}. Given the following data determine the age and initial ratio:

^{143}Nd/^{144}Nd	^{147}Sm/^{144}Nd
0.5105	0.12
0.5122	0.18
0.5141	0.24
0.5153	0.28
0.516	0.3

Where do you think this rock is from?