Computation and Computers in Geology

Simultaneous equations; fitting lines and polynomials to data

In the previous Excel exercises we investigated relative versus absolute addresses, graphed various functions, and worked with three array functions, **TRANSPOSE**(array), **MMULT**(array1,array2), and **MINVERSE**(matrix)). Now we can investigate some further uses of matrix algebra.

For the first application we will take a simple approach to the classic problem of trying to learn about the density of the mantle and core; we end up with two equations in two unknowns. In general, if we have an equal number of independent equations and unknowns then we can solve for the unknowns.

In our example we know the moment of inertia (Ie, $.3306*MeRe^2$) of Earth and its' radius (6,371 km) from astronomical observations, the mass (Me, $5.97*10^{24}$ kg) from Newton's work, the depth of the core-mantle boundary (2,900 km) from seismology. Given these values, how might we estimate the density of the core and mantle? We will use these abbreviations:

Mm = mantle mass	Pm = mantle density	Rm = mantle radius	Vm = mantle volume
Mc = core mass	Pc = core density	Rc = core radius	Vc = core volume
Im = mantle moment of inertia		Ic = core moment of inertia	

First, the mass of Earth is equal to the sum of the mass of the mantle and the mass of the core; the mass of each is the product of their volume and density:

$$\begin{aligned} Me &= Mm + Mc \\ &= PmVm + PcVc \end{aligned}$$

Similarly, the moment of inertia of Earth is equal to the sum of the moments of the mantle and the core; the moment of a uniform sphere (our simplifying assumption is that the density of the mantle and core are constant) is $2/5*Mass*radius^2$ of that sphere:

$$\begin{aligned} Ie &= Im + Ic \text{ ; recall moment for a uniform sphere } = 0.4Ms*Rs^2 \\ &= 0.4*Pm*(VeRe^2-VcRc^2) + 0.4*McRc^2 \\ &= 0.4*Pm*(VeRe^2-VcRc^2) + 0.4*PcVcRc^2 \end{aligned}$$

Now we have two equations in two unknowns (Pc, Pm) which, in matrix notation, yields:

$$\begin{bmatrix} 0.4*(Ve\mathrm{Re}^2 - VcRc^2) & 0.4*VcRc^2 \\ Vm & Vc \end{bmatrix} * \begin{bmatrix} Pm \\ Pc \end{bmatrix} = \begin{bmatrix} Ie \\ Me \end{bmatrix} \quad \text{or, } \mathbf{K*M = D}$$

To **solve for M**, and thus to determine Pm and Pc, find matrix of model parameters:

$$\mathbf{M = K^{-1} D}$$

Exercise 1:

1. Show that the result for my mantle and core problem is (Pm, Pc) = (4,151.5 kg/m^3, 12,560.8 kg/m^3). Make sure you get everything in the **MKS** system.

For the next application of matrix manipulation, consider the linear equation:

$$y_i = m*x_i + b$$

where m is the slope and b the intercept of the equation. For a simple example, graph the data for y=3*x + 1:

x	y
0	1
1	4
2	7
3	10
4	13

For this sort of equation we could write out a series of independent equations, one for each value of x that we are interested in:

$$y_1 = m*x_1 + b$$
$$y2 = m*x_2 + b$$
$$y3 = m*x_3 + b$$
$$\dots$$
$$y_n = m*x_n + b$$

We can also write this relation as a system of matrices:

$$\begin{bmatrix} y1 \\ y2 \\ y3 \\ \dots \\ yn \end{bmatrix} = \begin{bmatrix} x1 & 1 \\ x2 & 1 \\ x3 & 1 \\ \dots & \dots \\ xn & 1 \end{bmatrix} * \begin{bmatrix} m \\ b \end{bmatrix}$$

You can make up an example and use Excel's matrix multiplication function (MMULT(array1,array2)) to check out matrix multiplication. The standard matrix shorthand for writing such a matrix equation as the above is:

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

So why would we bother? You can consider the Y values in the matrices above as the observations resulting from some experiment or field measurement at locations X. If you have reason to believe the underlying equation governing the system you are investigating, say seismic arrival times as a function of distance, then the slope (m) and intercept (b) are simply the model parameters you wish to discover. That is, given the observations Y, you want to find the slope and intercept. Solving our system of equations, really n equations in 2 unknowns (m and b), is a classic least squares problem. We want to find the line (y = m*x + b) that best describes the system. If this was a course in statistics (it isn't) we would now spend a lot of time discussing "best". What we will do is find a classic solution that allows us to further develop our spreadsheet skills.

Given:

$$\mathbf{Y} = \mathbf{XM}$$

multiply both sides of the equation by the transpose of X:

$$\mathbf{X^t Y} = \mathbf{X^t XM}.$$

This yields a square matrix, just what we need to use the MINVERSE function. So, now find the inverse of $\mathbf{X^t X}$, which we will call $[\mathbf{X^t X}]^{-1}$.

Now multiply both sides of the equation by $[\mathbf{X^t X}]^{-1}$:

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

Note that the right side, $[\mathbf{X^t X}]^{-1}\,[\mathbf{X^t X}]\mathbf{M}$, has a matrix times its inverse which produces the identity matrix much like the equation 1/n * n produces the value one. Thus the equation yields:

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

and we have solved for the model parameters, in this case slope and intercept, of our linear equation.

Exercise 2:

2. Given:

$$Xi, Yi = \begin{bmatrix} 0 & .9 \\ 1 & 4.1 \\ 2 & 6.9 \\ 3 & 9.9 \\ 4 & 13.1 \end{bmatrix} \quad \text{thus, } X = \begin{bmatrix} 0 & 1 \\ 1 & 1 \\ 2 & 1 \\ 3 & 1 \\ 4 & 1 \end{bmatrix} \text{ and } \quad Y = \begin{bmatrix} .9 \\ 4.1 \\ 6.9 \\ 9.9 \\ 13.1 \end{bmatrix}$$

a. Find $\mathbf{X^t}$

After you enter the data as above, highlight (left mouse button down, drag) a block of the spreadsheet that has two rows and five columns. Next highlight or type in the array range you wish to transpose. Finally, **enter the array formula in Excel with CNTRL-SHIFT-ENTER (simultaneously).** That's an Excel peculiarity to remember (or find it in Help like I always have to do…).

b. Find $\mathbf{X^t X}$, and $[\mathbf{X^t X}]^{-1}$. Use MMULT(array1, array2) - **CNTRL-SHIFT-ENTER** and MINVERSE(array) - **CNTRL-SHIFT-ENTER**.

c. Show that $[\mathbf{X^t X}]^{-1}\,[\mathbf{X^t X}] = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

d. Find (m, b) and graph the data along with your new, best-fit, line; neat, clean, professional..

Optional – More Fun for Your Money!!

Extending the above to Parabolas, Polynomials and the like:

Above we saw that for a set of linear equations:

$y(x) = m*x + b$

we could write those equations in matrix form:

[Y] = [X][M] or, **Y=XM**

and use matrix algebra to get the solution for slope and intercept (m, b):

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

This works because the underlying equation {y = m*x +b} is linear in the parameters (m, b) of the model:

dy/dm = x; dy/db = 0

The same is true for a parabola $\{y(x) = a*x^2 + b*x + c\}$ or any polynomial of higher order:

$dy/da = x^2$; dy/db = x; dy/dc = 0.

Thus we can use the same technique as above to find the least-squares, best-fit parabola if we have data whose underlying function is parabolic. Consider these parabolic equations:

$$Y_1 = a*x_1^2 + b*x_1 + c$$
$$Y_2 = a*x_2^2 + b*x_2 + c$$
$$Y_3 = a*x_3^2 + b*x_3 + c$$
$$\ldots$$
$$Y_n = a*x_n^2 + b*x_n + c$$

Which can be written in matrix form as:

$$\begin{bmatrix} Y1 \\ Y2 \\ Y3 \\ \ldots \\ Y4 \end{bmatrix} = \begin{bmatrix} x_1^2 & x_1 & 1 \\ x_2^2 & x_2 & 1 \\ x_3^2 & x_3 & 1 \\ \ldots & \ldots & \ldots \\ x_n^2 & x_n & 1 \end{bmatrix} * \begin{bmatrix} a & b & c \end{bmatrix}$$

or

Y = XM

Whose solution for the model parameters is

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

Thus we now have the least-squares, best-fit a, b and c for our system of equations.

Problem 3:

Given these observations of temperature with depth:

Depth (km)	**Temperature (^{o}C)**
0	0
.1	3
.5	15
1	29
2	56
5	134
7	181
10	248

a. Use the matrix algebra approach to determine the best-fit parameters for a linear equation and a parabolic equation. This is the same as the linear approach except that your "X" matrix now has three columns (X^2 X 1).

b. Extrapolate both functions on your graph to predict temperature at 15 kilometers and 25 kilometers – consider the difference.

c. Make a nice graph which demonstrates which of the two functions best fits the data.