Least Squares Homework: Matlab

For this homework assignment, we'll be finding model equations for data sets. First, download three data sets from our class website:

- `HanfordData1.mat` This data is from an article in *Journal of Environmental Health*, May-June 1965, Volume 27, Number 6, pgs 883-897, by Robert Fadeley. The idea here is that there has been radioactive waste seeping into the Columbia River, and 8 Oregon counties and the city of Portland have been exposed to contamination. In the data file, there are three vectors of data:
  - `Index` is a vector with index values. These values try to measure an exposure risk to radiation by assuming that the exposure is directly proportional to river frontage and inversely proportional both to the distance from Hanford and the average depth away from the river.
  - `Deaths` is the number of cancer deaths per 100,000 residents.
  - `Names` is a listing of the county names (and Portland) in order.
- `Body1.mat` is a listing of some body measurements of 507 people. It includes:
  - `KneeGirth` is a vector of knee girth measurements (in cm)
  - `Height` is a vector of heights (in cm)
  - `menidx` and `womenidx` are the indices of men and women. For example, to look at heights associated only with women, in Matlab type: `Height(womenidx)`
- `mammals.mat` This data has two vectors, w is the body weight of some mammals, and r is the heart rate. The first entry of each represents a shrew, and the last entry is for an elephant. We will assume that the relationship is:

$$r = Aw^k$$

  where A, k are the unknowns. First put this into a linear relationship, then find A, k.

For each of the first two data sets, use linear algebra to find the best fitting line to the data (you can check with the Data Fitting tools). Turn in a printout of the command window (or script file) showing your computations, a plot of the original data (as discrete circles) and the line. Be sure you have the slope and intercept written down on the printout.

For the last data set, also turn in either a printout from the command window or the script file. It should be clear from the computations that you are first linearizing the data. Turn in two plots: The first plot will show the linearized data together with the best fitting line, The second plot will show the *original* data with the best fitting *curve*.