SM233 – Project A – Curve Fitting
(Due Friday 1/25)

Background. Up to an $n-1^{th}$-degree polynomial can be fitted to *n* data points. A polynomial of degree $n-1$ will produce a "perfect" fit. For example consider a set of four data points x_1, y_1 , x_2, y_2 , x_3, y_3 , x_4, y_4 . A cubic will satisfy the equations:

$$\begin{cases} y_1 = a_3 x_1^3 + a_2 x_1^2 + a_1 x_1^1 + a_0 \\ y_2 = a_3 x_2^3 + a_2 x_2^2 + a_1 x_2^1 + a_0 \\ y_3 = a_3 x_3^3 + a_2 x_3^2 + a_1 x_3^1 + a_0 \\ y_4 = a_3 x_4^3 + a_2 x_4^2 + a_1 x_4^1 + a_0 \end{cases},$$

where a_i are undetermined constants. This equation can be written in matrix form:

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} = \begin{bmatrix} x_1^3 & x_1^2 & x_1 & 1 \\ x_2^3 & x_2^2 & x_2 & 1 \\ x_3^3 & x_3^2 & x_3 & 1 \\ x_4^3 & x_4^2 & x_4 & 1 \end{bmatrix} \begin{bmatrix} a_3 \\ a_2 \\ a_1 \\ a_0 \end{bmatrix},$$

or in shorthand form: $Y = XA$. Using matrix operations we find the solution for A by left multiplying both sides of the equation by the inverse of the X, i.e.:

$$X^{-1}Y = X^{-1}XA = IA \Rightarrow A = X^{-1}Y,$$

where A contains the coefficients of a cubic that fits the four data points.

For a polynomial of degree $< n-1$ the fit will least squares approximation. Suppose we wish to fit the same four data points to a quadratic, i.e.

$$\begin{cases} y_1 = a_2 x_1^2 + a_1 x_1^1 + a_0 \\ y_2 = a_2 x_2^2 + a_1 x_2^1 + a_0 \\ y_3 = a_2 x_3^2 + a_1 x_3^1 + a_0 \\ y_4 = a_2 x_4^2 + a_1 x_4^1 + a_0 \end{cases} \Rightarrow\Rightarrow \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} = \begin{bmatrix} x_1^2 & x_1 & 1 \\ x_2^2 & x_2 & 1 \\ x_3^2 & x_3 & 1 \\ x_4^2 & x_4 & 1 \end{bmatrix} \begin{bmatrix} a_2 \\ a_1 \\ a_0 \end{bmatrix} \Rightarrow\Rightarrow Y = XA$$

In order to get a least squares fit for A we first left multiply both sides by X^T:

$$X^T Y = X^T XA$$

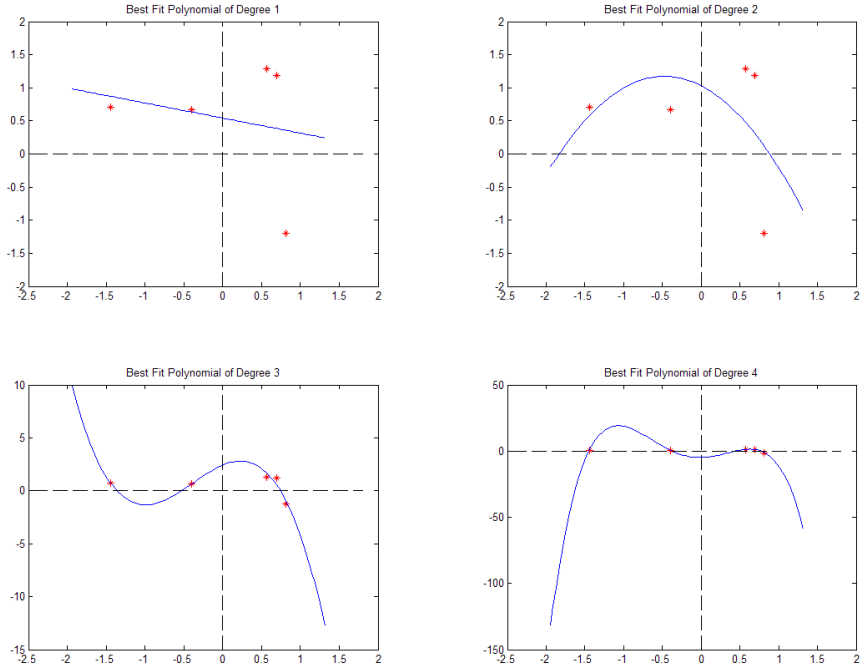
We than multiply both sides by the inverse of $X^T X$, i.e.

$$(X^T X)^{-1} X^T Y = (X^T X)^{-1}(X^T X)A = IA \Rightarrow A = (X^T X)^{-1} X^T Y$$

We now have solved for the coefficients in A such that the resulting parabola will be a least squares best fit for the data points.

Project.

- Create an *x*-vector and a *y*-vector consisting of 4 random variables each.
- Determine the coefficients for the line of best fit.
- Create a plot for -1<x<2 using these coefficients.
- Determine the coefficients for the parabola of best fit (use the same x,y-vectors).
- Plot the result on a separate plot for -1<x<2.
- Determine the coefficients for the cubic that fits the points.
- Plot the results.
- Consolidate outputs on a Microsoft Word document. Note that plots can be copied and pasted into MS Word.
- Emulate the sample plots below (note: you do not need to find a quartic)



Sources. For more information see:

http://mathworld.wolfram.com/LeastSquaresFittingPolynomial.html