# Numerical Analysis

Math 370 Fall 2004 MWF 2:30 - 3:25pm
 Fowler North 5

*Worksheet 20*

**SUMMARY** Discrete Least Squares Approximation
**READING** Recktenwald, Sec 9.1, pp 455–468

**Linear Algebra Approach To Discrete Least Squares Aproximation**
What is the system of equations we are trying to solve when finding a line of best fit $y = ax+b$ for $m$ data points $(x_k, y_k)$?

$$\begin{aligned} ax_1 + b &= y_1 \\ ax_2 + b &= y_2 \\ \vdots &= \vdots \\ ax_m + b &= y_m \end{aligned}$$

1. What are the known parameters here and what are the unknown variables ?

2. Write the system as a linear system $A\vec{c} = \vec{y}$

3. What are the dimensions of the matrix $A$ and the vectors $\vec{c}$ and $\vec{y}$?

We define a vector called $\vec{r}$ known as the residual vector and try to minimize the size (norm) of this vector since an exact solution to the matrix equation above is extremely unlikely.

$$||\vec{r}||_2^2 = \vec{r}^T\vec{r} = (\vec{y} - A\vec{c})^T(\vec{y} - A\vec{c})$$

4. If we consider this norm to be $\rho = \rho(\vec{c})$ we can think of our problem as finding the value of $\vec{c}$ which mimizes the size of $\rho(\vec{c})$. (Sound familiar?). Let's simplify the function first, and then minimize. Make sure you understand each step.

$$\begin{aligned} \rho(c) &= y^Ty - y^TAc - (Ac)^Ty - (Ac)^TAc \\ &= y^Ty - y^TAc - c^TA^Ty + c^TA^TAc \\ &= y^Ty - 2y^TAc + c^TA^TAc \end{aligned}$$

Thus, the general form of the Normal Equations can be written:

**Least Squares Fit Using Other Functions**

There is no reason why one has to pick the linear function $P(x) = ax + b$ to fit the data to. That is, one does not have to do the fit to a linear function (or even transform the data so that a linear relationship can be established between input and output variables).
Suppose

$$y = P(x) = c_1 f_1(x) + c_2 f_2(x) + \ldots + c_k f_k(x) = \sum_{i=1}^{k} c_i f_i(x)$$

Thus the best fit problem will become one of finding the $k$ $c_i$ values which cause the least square error to be minimized.
Note: If $k$, the number of basis functions, is equal to the number of data points, $m$, then one is really doing **interpolation** instead of polynomial best fit.

**Exercise**

1. Write down the expression $E(\vec{c}) = E(c_1, c_2, \ldots, c_n)$ which must be minimised:

**Polynomial Of Best Fit**

To simplify things, let's assume that the basis functions are monomials and we have $m$ data points

$$P(x) = \sum_{k=0}^{n} a_k x^k, \qquad \text{where } n + 1 < m$$

2. Why does $n + 1$ have to be less than $m$?

EXAMPLE

Let's find the normal equations whose solution is the $a_k$ values. (HINT: write it in Matrix Form.)

The above represents the $n + 1$ **normal equations** in the $n + 1$ unknowns $a_0, a_1, \ldots a_n$

**Using `polyfit` in MATLAB**

Let's try and find a second degree polynomial which fits the data $(0, 1)$, $(0.25, 1.2840)$, $(0.5, 1.6487)$, $(0.75, 2.1170)$ and $(1.00, 2.7183)$

3. What will the dimensions of the matrix equation to be solved look like? (i.e. what is $n$ and what is $m$?)

4. Find the coeffecients of the polynomial using `polyfit`
(Write down what commands you use)

5. Plot the data and the polynomial of best fit on the same graph.
(Write down what commands you use)