Least squares:

1) Given the (x,y) data

x	-1	2	3	5	6
y	3	1	-2	0	4

Find the least squares quadratic approximation $y = a + bx + cx^2$ as follows:

a) Write down the system of linear equations $y_i = a + bx_i + cx_i^2$, $i = 1,2,..,5$ you would like a,b,c to exactly satisfy.

In vector form, the equations to be satisfied are

$$\begin{bmatrix} 3 \\ 1 \\ -2 \\ 0 \\ 4 \end{bmatrix} = a\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} + b\begin{bmatrix} -1 \\ 2 \\ 3 \\ 5 \\ 6 \end{bmatrix} + c\begin{bmatrix} (-1)^2 \\ (2)^2 \\ (3)^2 \\ (5)^2 \\ (6)^2 \end{bmatrix}$$

b) Write the normal equations giving the $a,b,c,$ that come closest to solving the system in the least squares sense.

The system above asks that $\begin{bmatrix} 3 \\ 1 \\ -2 \\ 0 \\ 4 \end{bmatrix} = \begin{bmatrix} 1 & -1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \\ 1 & 5 & 25 \\ 1 & 6 & 36 \end{bmatrix}\begin{bmatrix} a \\ b \\ c \end{bmatrix}$

This is overdetermined, so we provide values (a,b,c) that minimize the sum of the squares of the errors. The equation has the form $\bar{b} = A\bar{x}$ so the least squares solution is obtained from the normal equations $A^T\bar{b} = (A^TA)\bar{x}$ which, numerically, becomes

$$\begin{bmatrix} 1 & -1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \\ 1 & 5 & 25 \\ 1 & 6 & 36 \end{bmatrix}^T\begin{bmatrix} 3 \\ 1 \\ -2 \\ 0 \\ 4 \end{bmatrix} = \begin{bmatrix} 6 \\ 17 \\ 133 \end{bmatrix} = \begin{bmatrix} 5 & 15 & 75 \\ 15 & 75 & 375 \\ 75 & 375 & 2019 \end{bmatrix}\bar{x}$$

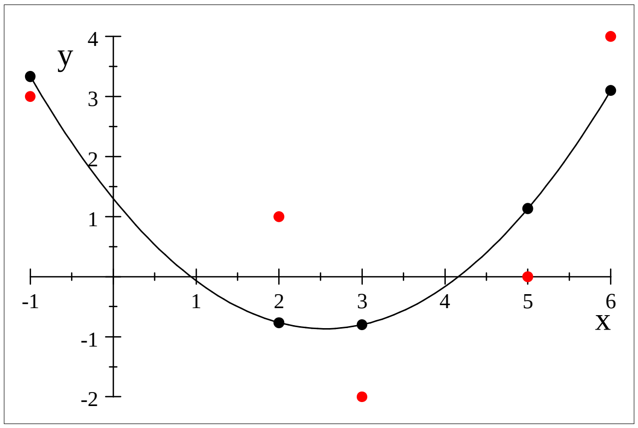
We solve for $\bar{x} = (A^TA)^{-1}A^T\bar{b} = \begin{bmatrix} 5 & 15 & 75 \\ 15 & 75 & 375 \\ 75 & 375 & 2019 \end{bmatrix}^{-1}\begin{bmatrix} 6 \\ 17 \\ 133 \end{bmatrix} = \begin{bmatrix} \frac{13}{10} \\ -\frac{17}{10} \\ \frac{1}{3} \end{bmatrix} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$

c) Find the least squares quadratic approximation to the data and plot it and the data (use whatever software you want for this part)

The quadratic approximation to the data is

$$\begin{bmatrix} 3 \\ 1 \\ -2 \\ 0 \\ 4 \end{bmatrix} = \bar{y} \cong \begin{bmatrix} 1 & -1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \\ 1 & 5 & 25 \\ 1 & 6 & 36 \end{bmatrix} \begin{bmatrix} \frac{13}{10} \\ -\frac{17}{10} \\ \frac{1}{3} \end{bmatrix} = \begin{bmatrix} \frac{10}{3} \\ -\frac{23}{30} \\ -\frac{4}{5} \\ \frac{17}{15} \\ \frac{31}{10} \end{bmatrix}$$

The quadratic itself is, of course, $\frac{13}{10} + \left(-\frac{17}{10}\right)x + \frac{1}{3}x^2$. Below, the original data is in red, the quadratic data is black, as is the graph of the quadratic.



d) From a linear algebra view point, what vector $\bar{b}$ is being projected onto what subspace V? How would you describe the projection $\bar{p}$ as it relates to the data of the problem?

The vector $\bar{b}$ is the data $\bar{y}$, being projected onto the subspace spanned by the powers x^0, x^1, x^2 evaluated at the points in x. The projection $\bar{p}$ is the quadratic data that comes closest to the y data at the points in x.

2) The least-squares approximation of data (x_i, y_i) with a constant function $f(x) = a$ is obtained when we set $a = y_{av}$, the average value of the y_i. Prove it. (What are the normal equations for a ?)

The equation to be "solved" is $\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} a$

The normal equations are $\begin{bmatrix} 1 & 1 & \vdots & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & 1 & \vdots & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} a$

If there are n datapoints, we have $\sum y_i = na$, $a = \frac{1}{n}\sum y_i = y_{av}$

3) In polynomial least squares curve fitting we approximate $\{(x_i, y_i)\}$ data with a function $f(x) = a_0 + a_1 x + \ldots + a_n x^n$. so as to minimize the sum of the squares of the errors $e_i = y_i - f(x_i)$. Show that when we do that the errors sum to zero: $\Sigma e_i = 0$. (Hint: The error vector is orthogonal to which vectors? For one of those vectors $\Sigma e_i = 0$ represents the orthogonality condition.)

The error vector must be orthogonal to the vectors in V, which in this case are the powers of x evaluated at the datapoints $\{x_i\}$. The first vector in V (the first column of the matrix A) is the vector of all $1's$ (as in problem 1 above). So the error is orthogonal to this vector, and taking the dot product of a vectors with all $1's$ gives you just the sum of the components, so $\sum e_i = 0$.

4.4 5,6,7,10,11,15,16,18,19,20,21,22,23,24

Supplementary probs:

If Q has orthonormal columns, the projection of $\bar{b}$ on the column space of Q is given by $\bar{p} = QQ^T\bar{b}$. a) How does this follow from the normal equations? Show this result by directly confirming that $\bar{b} - \bar{p}$ is orthogonal to $C(Q)$.

The normal equations are $Q^T\bar{b} = Q^TQ\bar{x} = I\bar{x} = \bar{x}$ because Q has orthonormal columns. Then $\bar{p} = Q\bar{x} = QQ^T\bar{b}$ gives the projection. Working directly with $\bar{p} = QQ^T\bar{b}$, we have $\bar{b} - \bar{p} = \bar{b} - QQ^T\bar{b}$ and the calculation $Q^T(\bar{b} - \bar{p}) = Q^T(\bar{b} - QQ^T\bar{b}) = Q^T\bar{b} - (Q^TQ)Q^T\bar{b} = \bar{0}$

b) If Q is an orthogonal matrix (square, with orthonormal columns) we know (how?) that $QQ^T = I$ so that $QQ^T\bar{x} = \bar{x}$ for all $\bar{x}$. If Q has orthonormal columns but is not square, in what sense might we say that $QQ^T \cong I$ is true? Hint: We don't have $QQ^T\bar{x} = \bar{x}$ but in what sense do we have $QQ^T\bar{x} \cong \bar{x}$?

Since Q is square, and $Q^TQ = I$ due to orthogonality, it must be the case that $Q^T = Q^{-1}$ so that $QQ^T = I$ as well (so the rows of Q are also orthonormal). From part a), we can recognize $QQ^T\bar{x}$ as the projection of $\bar{x}$ onto the column space of Q, and that is as close to $\bar{x}$ as we can get by using the columns of Q. So in that sense, we have $QQ^T\bar{x} \cong \bar{x}$ even when Q is not square.