LEAST SQUARES DATA FITTING

Experiments generally have *error* or *uncertainty* in measuring their outcome. *Error* can be human error, but it is more usually due to inherent limitations in the equipment being used to make measurements. *Uncertainty* can be due to lack of precise definition or of human variation in what is being measured. (For example, how do you measure how much you like something?)

We often want to represent the experimental data by some functional expression. Interpolation is often unsatisfactory because it fails to account for the error or uncertainty in the data.

We may know from theory that the data is taken from a particular form of function (e.g. a quadratic polynomial), or we may choose to use some particular type of formula to represent the data.

EXAMPLE

Consider the following data.

Table 1: Empirical data

x_i	y_i	x_i	y_i
1.0	−1.945	3.2	0.764
1.2	−1.253	3.4	0.532
1.4	−1.140	3.6	1.073
1.6	−1.087	3.8	1.286
1.8	−0.760	4.0	1.502
2.0	−0.682	4.2	1.582
2.2	−0.424	4.4	1.993
2.4	−0.012	4.6	2.473
2.6	−0.190	4.8	2.503
2.8	0.452	5.0	2.322
3.0	0.337		

From the following Figure 1 it appears to be approximately linear.

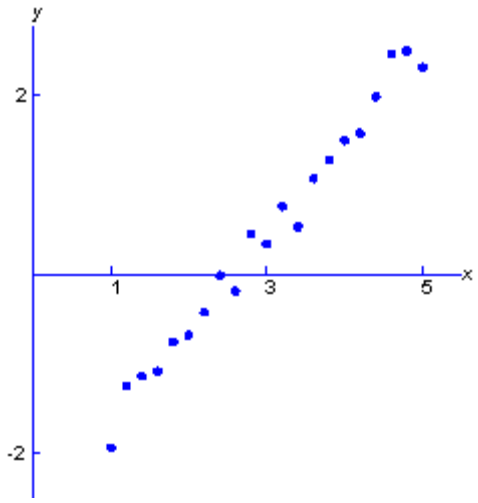

Figure 1: The plot of empirical data

An experiment seeks to obtain an unknown functional relationship

$$y = f(x) \tag{1}$$

involving two related variables x and y. We choose varying values of x, say, $x_1, x_2, \ldots, x_n$. Then we measure a corresponding set of values for y. Let the actual measurements be denoted by $y_1, \ldots, y_n$, and let

$$\epsilon_i = f(x_i) - y_i$$

denote the unknown measurement errors. We want to use the points $(x_1, y_1), \ldots, (x_n, y_n)$ to determine the analytic relationship (1) as accurately as possible.

Often we suspect that the unknown function $f(x)$ lies within some known class of functions, for example, polynomials. Then we want to choose the member of that class of functions that will best approximate the unknown function $f(x)$, taking into account the experimental errors $\{\epsilon_i\}$.

As an example of such a situation, consider the data in Table 1 and the plot of it in Figure 1. From this plot, it is reasonable to expect $f(x)$ to be close to a linear polynomial,

$$f(x) = mx + b \tag{2}$$

Assuming this to be the true form, the problem of determining $f(x)$ is now reduced to that of determining the constants m and b.

We can choose to determine m and b in a number of ways. We list three such ways.

1. Choose m and b so as to minimize the quantity

$$\frac{1}{n}\sum_{i=1}^{n}|f(x_i)-y_i|$$

which can be considered an average approximation error.

2. Choose m and b so as to minimize the quantity

$$\sqrt{\frac{1}{n}\sum_{i=1}^{n}[f(x_i)-y_i]^2}$$

which can also be considered an average approximation error. It is called the *root mean square error* in the approximation of the data $\{(x_i, y_i)\}$ by the function $f(x)$.

3. Choose m and b so as to minimize the quantity

$$\max_{1\le i\le n}|f(x_i)-y_i|$$

which is the maximum error of approximation.

All of these can be used, but #**2** is the favorite, and we now comment on why. To do so we need to understand more about the nature of the unknown errors $\{\epsilon_i\}$.

Standard assumption: Each error ϵ_i is a random variable chosen from a *normal probability distribution*. Intuitively, such errors satisfy the following.
(1) If the experiment is repeated many times for the same $x = x_i$, then the associated unknown errors ϵ_i in the empirical values y_i will be likely to have an average of zero.
(2) For this same experimental case with $x = x_i$, as the size of ϵ_i increases, the likelihood of its occurring will decrease rapidly.

This is the *normal error assumption*. We also assume that the individual errors ϵ_i, $1 \leq i \leq n$, are all random variables from the same normal probability distribution function, meaning that the size of ϵ_i is unrelated to the size of x_i or y_i.

Assume $f(x)$ is in a known class of functions, call it C. An example is the assumption that $f(x)$ is linear for the data in Table 1.

Then among all functions $\widehat{f}(x)$ in C, it can be shown that the function $\widehat{f}^*$ that is most likely to equal f will also minimize the expression

$$E = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left[\widehat{f}(x_i) - y_i\right]^2} \tag{3}$$

among all functions $\widehat{f}$ in C.

This is called the *root-mean-square error* in the approximation of the data $\{y_i\}$ by $\widehat{f}(x)$. The function $\widehat{f}^*(x)$ that minimizes E relative to all $\widehat{f}$ in C is called the *least squares approximation* to the data $\{(x_i, y_i)\}$.

EXAMPLE

Return to the data in Table 1, pictured in Figure 1. The least squares approximation is given by

$$\widehat{f}^*(x) = 1.06338x - 2.74605 \qquad (4)$$

It is illustrated graphically in Figure 2.

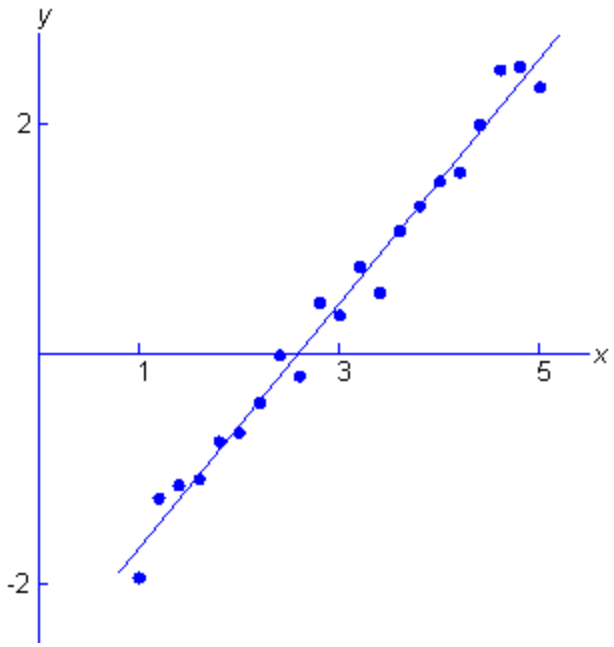

Figure 2: The linear least squares fit $\widehat{f}^*(x)$

CALCULATING THE LEAST SQUARES APPROXIMATION

How did we calculate $\widehat{f}^*(x)$? We want to minimize

$$E \equiv \sqrt{\frac{1}{n}\sum_{i=1}^{n}[f(x_i) - y_i]^2}$$

when considering all possible functions $f(x) = mx + b$. Note that minimizing E is equivalent to minimizing the sum, although the minimum values will be different. Thus we seek to minimize

$$G(b, m) = \sum_{i=1}^{n}[mx_i + b - y_i]^2 \qquad (5)$$

as b and m are allowed to vary arbitrarily.

The choices of b and m that minimize $G(b, m)$ will satisfy

$$\frac{\partial G(b, m)}{\partial b} = 0, \qquad \frac{\partial G(b, m)}{\partial m} = 0 \qquad (6)$$

Use

$$\frac{\partial G}{\partial b} = \sum_{i=1}^{n} 2\left[mx_i + b - y_i\right]$$

$$\frac{\partial G}{\partial m} = \sum_{i=1}^{n} 2\left[mx_i + b - y_i\right] x_i = \sum_{i=1}^{n} 2\left[mx_i^2 + bx_i - x_i y_i\right]$$

This leads to the linear system

$$\begin{aligned} nb + \left(\sum_{i=1}^{n} x_i\right) m &= \sum_{i=1}^{n} y_i \\ \left(\sum_{i=1}^{n} x_i\right) b + \left(\sum_{i=1}^{n} x_i^2\right) m &= \sum_{i=1}^{n} x_i y_i \end{aligned} \tag{7}$$

This is uniquely solvable if the determinant is nonzero,

$$n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2 \neq 0 \tag{8}$$

This is true unless

$$x_1 = x_2 = \cdots = x_n = \text{constant}$$

and this is false for our case.

For our example in Table 1,

$$\sum_{i=1}^{n} x_i = 63.0 \qquad \sum_{i=1}^{n} x_i^2 = 219.8$$

$$\sum_{i=1}^{n} y_i = 9.326 \qquad \sum_{i=1}^{n} x_i y_i = 60.7302$$

Using this in (7), the linear system becomes

$$\begin{aligned} 21b + 63.0m &= 9.326 \\ 63.0b + 219.8m &= 60.7302 \end{aligned}$$

The solution is

$$b \doteq -2.74605 \qquad m \doteq 1.06338$$

$$\widehat{f}^*(x) = 1.06338x - 2.74605$$

The root-mean-square-error in $\widehat{f}^*(x)$ is

$$E \doteq 0.171$$

Recall the graph of $\widehat{f}^*(x)$ is given in Figure 2.

GENERALIZATION

To represent the data $\{(x_i, y_i) \mid 1 \leq i \leq n\}$, let

$$\widehat{f}(x) = a_1\varphi_1(x) + a_2\varphi_2(x) + \cdots + a_m\varphi_m(x) \quad (9)$$

$a_1, a_2, \ldots, a_m$ arbitrary numbers, $\varphi_1(x), \ldots, \varphi_m(x)$ given functions.

If $\widehat{f}(x)$ is to be a quadratic polynomial, write

$$\widehat{f}(x) = a_1 + a_2x + a_3x^2 \quad (10)$$

$$\varphi_1(x) \equiv 1, \qquad \varphi_2(x) = x, \qquad \varphi_3(x) = x^2$$

Under the normal error assumption, the function $\widehat{f}(x)$ is to be chosen to minimize the root-mean-square error

$$E = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left[\widehat{f}(x_i) - y_i\right]^2}$$

Consider $m = 3$. Then

$$\widehat{f}(x) = a_1\varphi_1(x) + a_2\varphi_2(x) + a_3\varphi_3(x)$$

Choose a_1, a_2, a_3 to minimize

$$G(a_1, a_2, a_3) = \sum_{j=1}^{n} [a_1\varphi_1(x_j)+a_2\varphi_2(x_j)+a_3\varphi_3(x_j)-y_j]^2$$

At the minimizing point (a_1, a_2, a_3),

$$\frac{\partial G}{\partial a_1} = 0, \qquad \frac{\partial G}{\partial a_2} = 0, \qquad \frac{\partial G}{\partial a_3} = 0$$

This leads to the three equations. For $i = 1, 2, 3$,

$$0 = \frac{\partial G}{\partial a_i} = \sum_{j=1}^{n} 2[a_1\varphi_1(x_j)+a_2\varphi_2(x_j)+a_3\varphi_3(x_j)-y_j]\varphi_i(x_j)$$

$$\left[\sum_{j=1}^{n} \varphi_1(x_j)\varphi_i(x_j)\right] a_1 + \left[\sum_{j=1}^{n} \varphi_2(x_j)\varphi_i(x_j)\right] a_2 + \left[\sum_{j=1}^{n} \varphi_3(x_j)\varphi_i(x_j)\right] a_3 = \sum_{j=1}^{n} y_j\varphi_i(x_j), \quad (11)$$

Apply this to the quadratic formula

$$\widehat{f}(x) = a_1 + a_2 x + a_3 x^2$$

$$\varphi_1(x) \equiv 1, \qquad \varphi_2(x) = x, \qquad \varphi_3(x) = x^2$$

Then the three equations are

$$\begin{aligned}
n a_1 + \left[\sum_{j=1}^{n} x_j\right] a_2 + \left[\sum_{j=1}^{n} x_j^2\right] a_3 &= \sum_{j=1}^{n} y_j \\
\left[\sum_{j=1}^{n} x_j\right] a_1 + \left[\sum_{j=1}^{n} x_j^2\right] a_2 + \left[\sum_{j=1}^{n} x_j^3\right] a_3 &= \sum_{j=1}^{n} y_j x_j \\
\left[\sum_{j=1}^{n} x_j^2\right] a_1 + \left[\sum_{j=1}^{n} x_j^3\right] a_2 + \left[\sum_{j=1}^{n} x_j^4\right] a_3 &= \sum_{j=1}^{n} y_j x_j^2
\end{aligned} \tag{12}$$

This can be shown a nonsingular system due to the assumption that the points $\{x_j\}$ are distinct.

Generalization. Let

$$\widehat{f}(x) = a_1\varphi_1(x) + a_2\varphi_2(x) + \cdots + a_m\varphi_m(x)$$

The root-mean-square error E in (3) is minimized with the coefficients $a_1, \ldots, a_m$ satisfying

$$\sum_{k=1}^{m} a_k \left[\sum_{j=1}^{n} \varphi_k(x_j)\varphi_i(x_j)\right] = \sum_{j=1}^{n} y_j\varphi_i(x_j) \qquad (13)$$

for $i = 1, \ldots, m$. For the special case of a polynomial of degree $(m-1)$,

$$\widehat{f}(x) = a_1 + a_2x + a_3x^2 + \cdots + a_mx^{m-1}$$

write

$$\varphi_1(x) = 1, \quad \varphi_2(x) = x, \quad \varphi_3(x) = x^2, \quad \ldots, \varphi_m(x) = x^{m-1} \qquad (14)$$

System (13) becomes

$$\sum_{k=1}^{m} a_k \left[\sum_{j=1}^{n} x_j^{i+k-2}\right] = \sum_{j=1}^{n} y_j x_j^{i-1}, \qquad i = 1, 2, \ldots, m \qquad (15)$$

When $m = 3$, this yields the system (12) obtained earlier.

ILL-CONDITIONING

This system (15) is nonsingular (for $m < n$). Unfortunately it is increasingly ill-conditioned as the degree $m-1$ increases.

The condition number for the matrix of coefficients can be very large for fairly small values of m, say, $m = 4$.

For this reason, it is seldom advisable to use

$$\varphi_1(x) = 1, \quad \varphi_2(x) = x, \quad \varphi_3(x) = x^2, \\ \ldots, \varphi_m(x) = x^{m-1}$$

to do a least squares polynomial fit, except for degree ≤ 2.

To do a least squares fit to data $\{(x_i, y_i) \mid 1 \leq i \leq n\}$ with a higher degree polynomial $\widehat{f}(x)$, write

$$\widehat{f}(x) = a_1\varphi_1(x) + \cdots + a_m\varphi_m(x)$$

with $\varphi_1(x), \ldots, \varphi_m(x)$ so chosen that the matrix of coefficients in

$$\sum_{k=1}^{m} a_k \left[\sum_{j=1}^{n} \varphi_k(x_j)\varphi_i(x_j) \right] = \sum_{j=1}^{n} y_j\varphi_i(x_j)$$

is not ill-conditioned.

There are optimal choices of these functions $\varphi_j(x)$, with $\deg(\varphi_j) = j - 1$ and with the coefficient matrix becoming diagonal.

IMPROVED BASIS FUNCTIONS

A nonoptimal but still satisfactory choice in general can be based on the Chebyshev polynomials $\{T_k(x)\}$ of Section 5.5, and a somewhat better choice is the Legendre polynomials of Section 5.7.

Suppose that the nodes $\{x_i\}$ are chosen from an interval $[\alpha, \beta]$. Introduce modified Chebyshev polynomials

$$\varphi_k(x) = T_{k-1}\left(\frac{2x - \alpha - \beta}{\beta - \alpha}\right), \qquad \alpha \le x \le \beta, \qquad k \ge 1 \tag{16}$$

Then degree $(\varphi_k) = k - 1$; and any polynomial $\widehat{f}(x)$ of degree $(m - 1)$ can be written as a combination of $\varphi_1(x), \ldots, \varphi_m(x)$.

EXAMPLE

Consider the following data.

Table 2: Data for a cubic least squares fit

x_i	y_i	x_i	y_i
0.00	0.486	0.55	1.102
0.05	0.866	0.60	1.099
0.10	0.944	0.65	1.017
0.15	1.144	0.70	1.111
0.20	1.103	0.75	1.117
0.25	1.202	0.80	1.152
0.30	1.166	0.85	1.265
0.35	1.191	0.90	1.380
0.40	1.124	0.95	1.575
0.45	1.095	1.00	1.857
0.50	1.122		

From the following Figure 3 it appears to be approximately cubic. We begin by using

$$\widehat{f}(x) = a_1 + a_2 x + a_3 x^2 + a_4 x^3 \qquad (17)$$

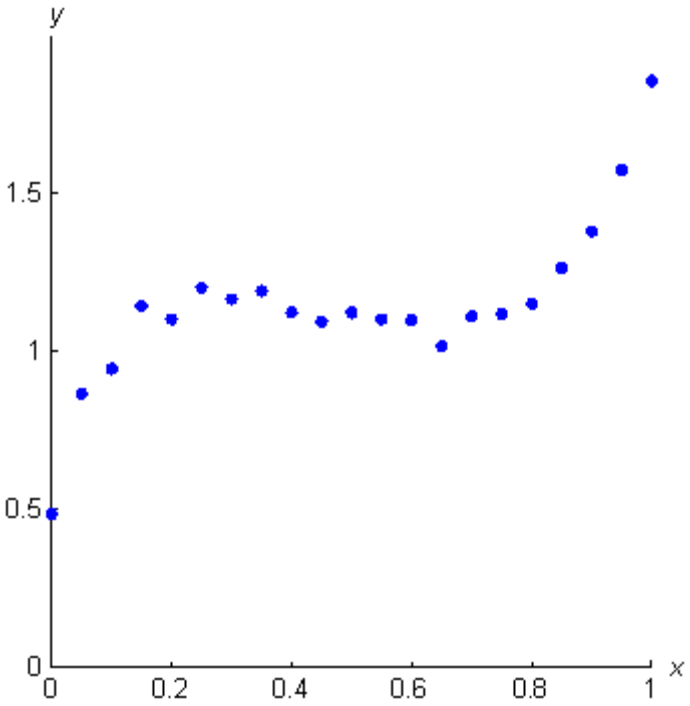

Figure 3: The plot of data of Table 2

The resulting linear system (15), denoted here by $La = b$, is given by

$$L = \begin{bmatrix} 21 & 10.5 & 7.175 & 5.5125 \\ 10.5 & 7.175 & 5.5125 & 4.51666 \\ 7.175 & 5.5125 & 4.51666 & 3.85416 \\ 5.5125 & 4.51666 & 3.85416 & 3.38212 \end{bmatrix}$$

$$a = [a_1, a_2, a_3, a_4]^T$$

$$b = [24.1180, 13.2345, 9.46836, 7.55944]^T$$

The solution is

$$a = [0.5747,\ 4.7259,\ -11.1282,\ 7.6687]^T$$

The condition number is

$$\text{cond}(L) = \|L\|\,\|L^{-1}\| \doteq 22000 \tag{18}$$

This is very large; it may be difficult to obtain an accurate answer for $La = b$.

To verify this, perturb b above by adding to it the perturbation

$$[0.01,\ -0.01,\ 0.01,\ -0.01]^T$$

This will change b in its second place to the right of the decimal point, within the range of possible perturbations due to errors in the data. The solution of the new perturbed system is

$$a = [0.7408,\ 2.6825,\ -6.1538,\ 4.4550]^T$$

This is very different from the earlier result for a.

The main point here is that use of

$$\widehat{f}(x) = a_1 + a_2 x + a_3 x^2 + a_4 x^3$$

leads to a rather ill-conditioned system of linear equations for determining $\{a_1, a_2, a_3, a_4\}$.

A better basis.

Use the modified Chebyshev functions of (16) on $[\alpha, \beta] = [0, 1]$:

$$f(x) = a_1\varphi_1(x) + a_2\varphi_2(x) + a_3\varphi_3(x) + a_4\varphi_4(x)$$

$$\begin{aligned}
\varphi_1(x) &= T_0(2x-1) \equiv 1 \\
\varphi_2(x) &= T_1(2x-1) = 2x - 1 \\
\varphi_3(x) &= T_2(2x-1) = 8x^2 - 8x + 1 \\
\varphi_4(x) &= T_3(2x-1) = 32x^3 - 48x^2 + 18x - 1
\end{aligned}$$

The values $\{a_1, a_2, a_3, a_4\}$ are completely different than in the representation (17).

The linear system (13) is again denoted by $La = b$:

$$L = \begin{bmatrix} 21 & 0 & -5.6 & 0 \\ 0 & 7.7 & 0 & -2.8336 \\ -5.6 & 0 & 10.4664 & 0 \\ 0 & -2.8336 & 0 & 11.01056 \end{bmatrix}$$

$$b = [24.118,\ 2.351,\ -6.01108,\ 1.523576]^T$$

The solution is

$$a = [1.160969,\ 0.393514,\ 0.046850,\ 0.239646]^T$$

The linear system is very stable with respect to the type of perturbation made in b with the earlier approach to the cubic least squares fit, using (17).

This is implied by the small condition number of L.

$$\text{cond}(L) = \|L\|\,\|L^{-1}\| \doteq (26.6)(0.1804) \doteq 4.8$$

Relatively small perturbations in b will lead to relatively small changes in the solution a.

The graph of $\widehat{f}(x)$ is shown in Figure 4.

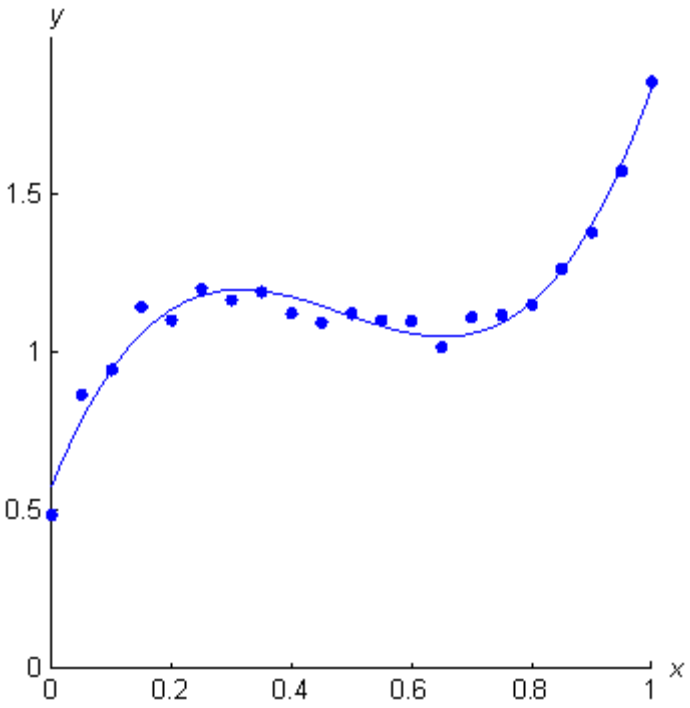

Figure 4: The cubic least squares fit for Table 2

To give some idea of the accuracy of $\widehat{f}(x)$ in approximating the data in Table 2, we easily compute the root-mean-square error from (3) to be

$$E \doteq 0.0421$$

a fairly small value when compared with the function values of $\widehat{f}(x)$.