



## Multiple Regression

If  $y$  is a function of more than one independent variable, the matrix equations that express the relationships among the variables are expanded to accommodate the additional data. This is called **multiple regression**.

Suppose you measure a quantity  $y$  for several values of  $x_1$  and  $x_2$ . Enter these into MATLAB at the command line, as follows:

```
x1 = [.2 .5 .6 .8 1.0 1.1]';
x2 = [.1 .3 .4 .9 1.1 1.4]';
y  = [.17 .26 .28 .23 .27 .24]';
```

A model of this data is of the form

$$y = a_0 + a_1x_1 + a_2x_2$$

**Multiple regression** solves for unknown coefficients  $a_0$ ,  $a_1$ , and  $a_2$  by minimizing the sum of the squares of the deviations of the data from the model (*least-squares fit*).

Construct and solve the set of simultaneous equations by forming the Vandermonde matrix,  $X$ , and solving for the parameters by using the backslash operator:

```
X = [ones(size(x1)) x1 x2];
a = X\y

a =
    0.1018
    0.4844
   -0.2847
```

The least-squares fit model of the data is

$$y = 0.1018 + 0.4844x_1 - 0.2847x_2$$

To validate the model, find the maximum of the absolute value of the deviation of the data from the model:

```
Y = X*a;
MaxErr = max(abs(Y - y))

MaxErr =
    0.0038
```

This value is sufficiently small when compared to the data values and indicates a good fit.