Linear Models

When we say that y is a **linear function** of x, we mean that the graph of the function is a line, so we can use the slope-intercept form of the equation of a line to write a formula for the function as

$$y = f(x) = mx + b$$

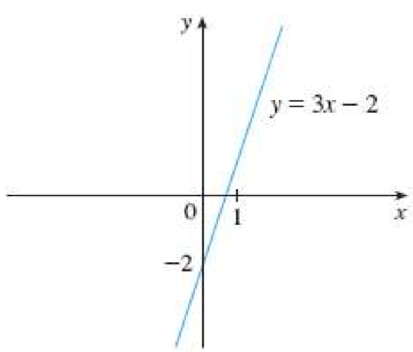
where m is the slope of the line and b is the y-intercept.

Note

The coordinate geometry of lines is reviewed in Appendix B.

A characteristic feature of linear functions is that they grow at a constant rate. For instance, Figure 2 shows a graph of the linear function $f(x) = 3x - 2$ and a table of sample values. Notice that whenever x increases by 0.1, the value of $f(x)$ increases by 0.3. So $f(x)$ increases three times as fast as x. Thus the slope of the graph $y = 3x - 2$, namely 3, can be interpreted as the rate of change of y with respect to x.

Figure 2



x	$f(x) = 3x - 2$
1.0	1.0
1.1	1.3
1.2	1.6
1.3	1.9
1.4	2.2
1.5	2.5

Example 1

(a) As dry air moves upward, it expands and cools. If the ground temperature is $20°\mathrm{C}$ and the temperature at a height of 1 km is $10°\mathrm{C}$, express the temperature T (in $°\mathrm{C}$) as a function of the height h (in kilometers), assuming that a linear model is appropriate.

(b) Draw the graph of the function in part (a). What does the slope represent?

(c) What is the temperature at a height of 2.5 km?

Solution

(a) Because we are assuming that T is a linear function of h, we can write

$$T = mh + b$$

We are given that $T = 20$ when $h = 0$, so

$$20 = m \cdot 0 + b = b$$

In other words, the y-intercept is $b = 20$.

We are also given that $T = 10$ when $h = 1$, so

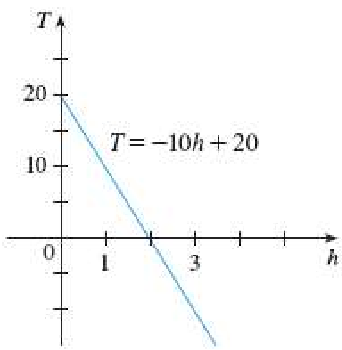
$$10 = m \cdot 1 + 20$$

The slope of the line is therefore $m = 10 - 20 = -10$ and the required linear function is

$$T = -10h + 20$$

(b) The graph is sketched in Figure 3. The slope is $m = -10°\mathrm{C/km}$ and this represents the rate of change of temperature with respect to height.

Figure 3



(c) At a height of $h = 2.5$ km, the temperature is

$$T = -10\,(2.5) + 20 = -5°\mathrm{C}$$

If there is no physical law or principle to help us formulate a model, we construct an **empirical model**, which is based entirely on collected data. We seek a curve that “fits" the data in the sense that it captures the basic trend of the data points.

Example 2

Table 1 lists the average carbon dioxide level in the atmosphere, measured in parts per million at Mauna Loa Observatory from 1980 to 2012. Use the data in Table 1 to find a model for the carbon dioxide level.

Table 1

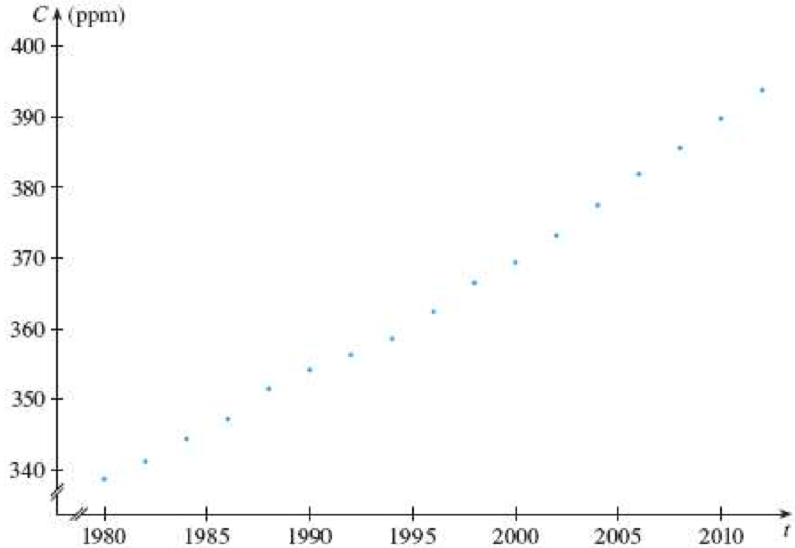
Year	**CO_2 level (in ppm)**
1980	338.7

Year	CO_2 level (in ppm)
1982	341.2
1984	344.4
1986	347.2
1988	351.5
1990	354.2
1992	356.3
1994	358.6
1996	362.4
1998	366.5
2000	369.4
2002	373.2
2004	377.5
2006	381.9
2008	385.6
2010	389.9
2012	393.8

Solution We use the data in Table 1 to make the scatter plot in Figure 4, where t represents time (in years) and C represents the CO_2 level (in parts per million, ppm).

Figure 4

Scatter plot for the average CO_2 level



Notice that the data points appear to lie close to a straight line, so it's natural to choose a linear model in this case. But there are many possible lines that approximate these data points, so which one should we use? One possibility is the line that passes through the first and last data points. The slope of this line is

$$\frac{393.8 - 338.7}{2012 - 1980} = \frac{55.1}{32} = 1.721875 \approx 1.722$$

We write its equation as

$$C - 338.7 = 1.722\,(t - 1980)$$

or

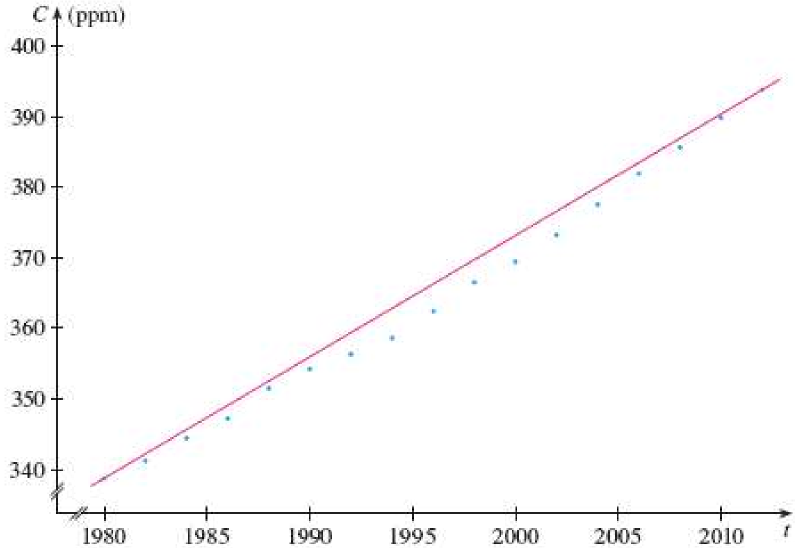
1

$$C = 1.722t - 3070.86$$

Equation 1 gives one possible linear model for the carbon dioxide level; it is graphed in Figure 5.

Figure 5

Linear model through first and last data points



Notice that our model gives values higher than most of the actual CO_2 levels. A better linear model is obtained by a procedure from statistics called *linear regression*. If we use a graphing calculator, we enter the data from Table 1 into the data editor and choose the linear regression command. (With Maple we use the fit[leastsquare] command in the stats package; with Mathematica we use the Fit command.) The machine gives the slope and y-intercept of the regression line as

$$m = 1.71262 \qquad b = -3054.14$$

So our least squares model for the CO_2 level is

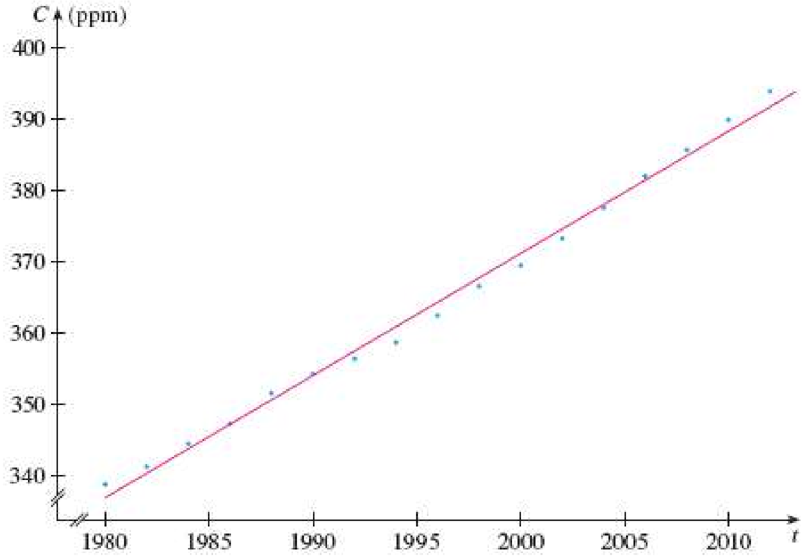
2

$$C = 1.71262t - 3054.14$$

In Figure 6 we graph the regression line as well as the data points. Comparing with Figure 5, we see that it gives a better fit than our previous linear model.

Figure 6

The regression line



Note

A computer or graphing calculator finds the regression line by the method of **least squares**, which is to minimize the sum of the squares of the vertical distances between the data points and the line. The details are explained in Section 14.7.

Example 3

Use the linear model given by Equation 2 to estimate the average CO_2 level for 1987 and to predict the level for the year 2020. According to this model, when will the CO_2 level exceed 420 parts per million?

Solution Using Equation 2 with $t = 1987$, we estimate that the average CO_2 level in 1987 was

$$C(1987) = (1.71262)(1987) - 3054.14 \approx 348.84$$

This is an example of *interpolation* because we have estimated a value *between* observed values. (In fact, the Mauna Loa Observatory reported that the average CO_2 level in 1987 was 348.93 ppm, so our estimate is quite accurate.)

With $t = 2020$, we get

$$C(2020) = (1.71262)(2020) - 3054.14 \approx 405.35$$

So we predict that the average CO_2 level in the year 2020 will be 405.4 ppm. This is an example of *extrapolation* because we have predicted a value *outside* the time frame of observations. Consequently, we are far less certain about the accuracy of our prediction.

Using Equation 2, we see that the CO_2 level exceeds 420 ppm when

$$1.71262t - 3054.14 > 420$$

Solving this inequality, we get

$$t > \frac{3474.14}{1.71262} \approx 2028.55$$

We therefore predict that the CO_2 level will exceed 420 ppm by the year 2029. This prediction is risky because it involves a time quite remote from our observations. In fact, we see from Figure 6 that the trend has been for CO_2 levels to increase rather more rapidly in recent years, so the level might exceed 420 ppm well before 2029.

Chapter 1: Functions and Models Linear Models
Book Title: Calculus: Early Transcendentals
Printed By: Troy Jeffery (tradozprime@gmail.com)