Optimization Techniques

CHAPTER

2

Carly Fiorina, chairperson of Hewlett-Packard, purchased the Compaq Computer Company in 2002. Why? Because she and her colleagues felt the acquisition would enhance the performance of Hewlett-Packard. As we learned in Chapter 1, managerial economics is concerned with the ways in which managers make decisions in order to maximize the effectiveness or performance of the organization they manage. To understand how this can be done, you must understand the basic optimization techniques taken up in this chapter.

To begin, we describe the nature of marginal analysis. While simple in concept, marginal analysis is a powerful tool that illuminates many central aspects of decision making regarding resource allocation. Economists think at the margins. It is intuitive that one would undertake a project if the additional (marginal) benefit one received from undertaking the project exceeded the additional (marginal) cost incurred by undertaking the project. Virtually all the rules of optimal behavior of firms and individuals that we study are driven by this concept.

Next, we examine the basic elements of differential calculus, including the rules of differentiation and the use of a derivative to maximize a function (such as profit) or minimize a function (such as cost). Differentiation tells us what changes will occur in one variable (the dependent variable) when a small (marginal) change is made in another variable (the independent variable). Therefore, the marginal analysis first discussed can be implemented by the use of differentiation.

Finally, we take up constrained optimization and include an optional section on Lagrangian multipliers. While we want to maximize the profits of our firm (or minimize the costs of production), such maximization or minimization is often subject to constraints (such as producing a certain amount to adhere to a contract or utilizing a certain amount of labor in a union agreement).

Since Lagrangian multipliers require more mathematical sophistication than the rest of this chapter, the section in which they are discussed can be skipped without loss of continuity.

Functional Relationships

To understand the optimization techniques described in this chapter, you must know how economic relationships are expressed. Frequently, the relationship between two or more economic variables can be represented by a table or graph. For example, Table 2.1 shows the relationship between the price charged by the Cherry Corporation and the number of units of output the company sells per day. Figure 2.1 represents the same relationship using a graph.

While tables and graphs are extremely helpful and are used often in this book, another way of expressing economic relationships is through equations. How can the relationship between the number of units sold and the price in Table 2.1 (and Figure 2.1) be expressed in the form of an equation? One way is to use the following functional notation:

$$Q = f(P) \tag{2.1}$$

where Q is the number of units sold and P is price. This equation should be read as: "The number of units sold is a function of price," which means that

TABLE 2.1

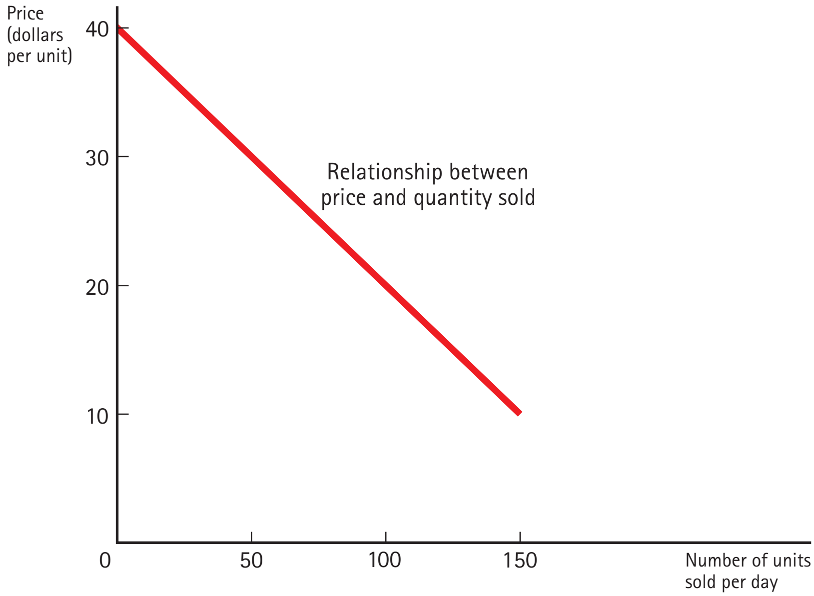
Relationship between Price and Quantity Sold, Cherry Corporation

Price per unit	Number of units sold per day
\$10	150
\$20	100
\$30	50
\$40	0

FIGURE 2.1

Relationship between Price and Quantity Sold, Cherry Corporation

This graph presents the data in Table 2.1.



the number of units sold *depends* on price. In other words, the number of units sold is the *dependent* variable, and price is the *independent* variable.

While equation (2.1) is useful, it does not tell us how the number of units sold depends on price. A more specific representation of this relationship is

$$Q = 200 - 5P \tag{2.2}$$

Comparing this equation with the data in Table 2.1 (and Figure 2.1), you can verify that these data conform to this equation. For example, if the price equals \$10, the number of units sold should be $200 - 5(10) = 150$, according to the equation. This is exactly what Table 2.1 (and Figure 2.1) shows. Regardless of what price you choose, the number of units sold is the same, no matter whether you consult Table 2.1, Figure 2.1, or equation (2.2).

Marginal Analysis

Whether economic relationships are expressed as tables, graphs, or equations, marginal analysis has enabled managers to use these relationships more

TABLE 2.2

Relationship between Output and Profit, Roland Corporation

(1) Number of units of output per day	(2) Total profit	(3) Marginal profit	(4) Average profit
0	0		—
1	100	100	100
2	250	150	125
3	600	350	200
4	1,000	400	250
5	1,350	350	270
6	1,500	150	250
7	1,550	50	221.4
8	1,500	−50	187.5
9	1,400	−100	155.5
10	1,200	−200	120

effectively. The **marginal value** of a dependent variable is defined as the change in this dependent variable associated with a one-unit *change* in a particular independent variable. As an illustration, consider Table 2.2, which indicates in columns 1 and 2 the total profit of the Roland Corporation if the number of units produced equals various amounts. In this case, total profit is the dependent variable and output is the independent variable. Therefore, the marginal value of profit, called **marginal profit**, is the *change* in total profit associated with a one-unit *change* in output.

Column 3 of Table 2.2 shows the value of marginal profit. If output increases from zero to one unit, column 2 shows that total profit increases by $100 (from $0 to $100). Therefore, marginal profit in column 3 equals $100 if output is between zero and one unit. If output increases from one to two units, total profit increases by $150 (from $100 to $250). Therefore, marginal profit in column 3 equals $150 if output is between one and two units.

The central point to bear in mind about a marginal relationship of this sort is that the dependent variable—in this case, total profit—is maximized when its marginal value shifts from positive to negative. To see this, consider Table 2.2. So long as marginal profit is positive, the Roland Corporation can raise its total profit by increasing output. For example, if output is between five and six units, marginal profit is positive ($150); therefore, the firm's total profit goes up (by $150) if output is increased from five to six units. But, when marginal profit shifts from positive to negative, total profit falls, not goes up, with any further increase in output. In Table 2.2, this point is reached when the firm produces seven units of output. If output is increased beyond this point (to eight units), marginal profit shifts from positive to negative—and total profit goes down (by $50). So, as stated previously, the dependent variable—in this case, total profit—is maximized when its marginal value shifts from positive to negative.

Since managers are interested in determining how to maximize profit (or other performance measures), this is a very useful result. It emphasizes the importance of looking at marginal values, and the hazards that may arise if average values are used instead. In Table 2.2, **average profit**—that is, total profit divided by output—is shown in column 4. It may seem eminently reasonable to choose the output whose average profit is highest. Countless managers have done so. But this is not the correct decision if one wants to maximize profit. Instead, as stressed in the previous paragraph, one should choose the output where marginal profit shifts from positive to negative.

To prove this, one need only find the output in Table 2.2 at which average profit is highest. Based on a comparison of the figures in column 4, this output is five units; and according to column 2, total profit at this output equals $1,350. In the paragraph before last, we found that the output where marginal profit shifts from positive to negative is seven units; and according to column 2, total profit at this point equals $1,550. Clearly, total profit is $200 higher if output is seven rather than five units. Hence, if the managers of this firm were to choose the output at which average profit is highest, they would sacrifice $200 per day in profits.

It is important to understand the relationship between average and marginal values: Because the marginal value represents the change in the total, the average value must increase if the marginal value is greater than the average value; by the same token, the average value must decrease if the marginal value is less than the average value. Table 2.2 illustrates these propositions. For the first to fifth units of output, marginal profit is greater than average profit. Since the extra profit from each additional unit is greater than the average, the average is pulled up as more is produced. For the sixth to tenth units of output, marginal profit is less than average profit. Since the extra profit from each additional unit is less than the average, the average is pulled down as more is produced.

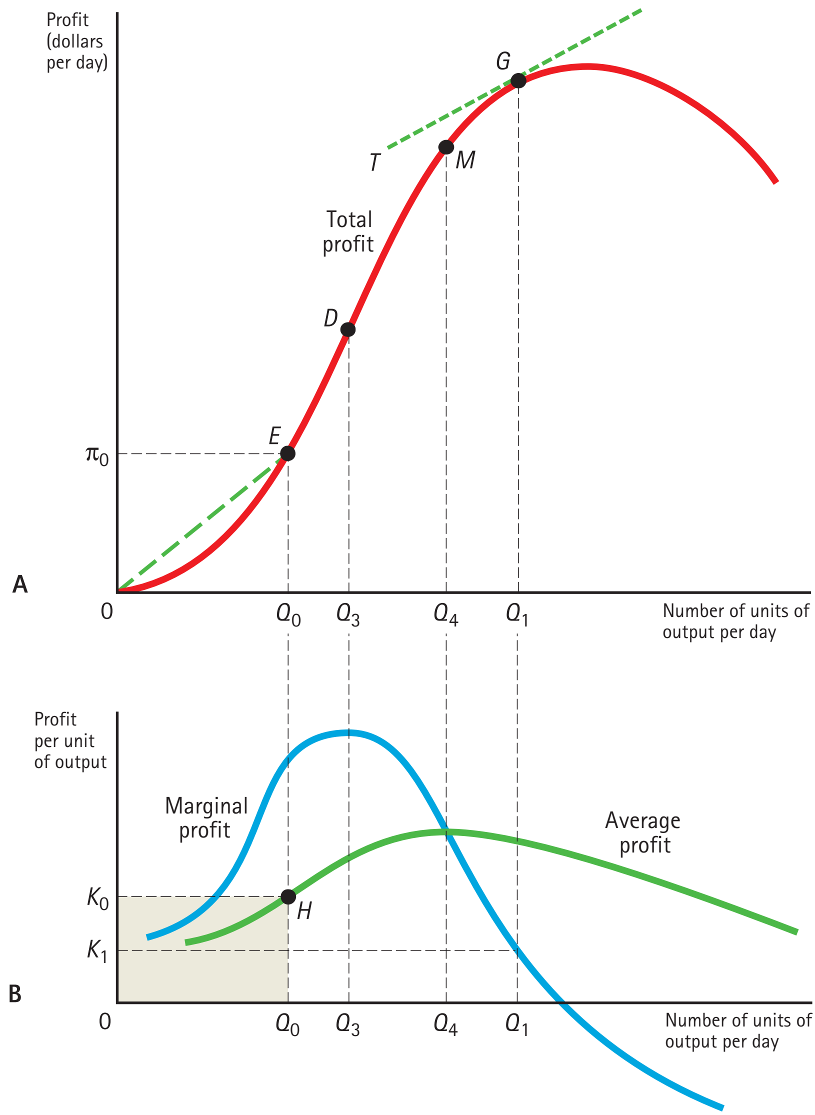
Relationships between Total, Marginal, and Average Values

To explore further the relationships between total, marginal, and average values, consider Figure 2.2, which shows the relationships between total, average, and marginal profit, on the one hand, and output, on the other hand, for the

FIGURE 2.2

Total Profit, Average Profit, and Marginal Profit, Roland Corporation

The average and marginal profit curves in panel B can be derived geometrically from the total profit curve in panel A.



Roland Corporation. The relationship between output and profit is exactly the same as in Table 2.2, but rather than using particular numbers to designate output or profit, we use symbols such as Q_0 and Q_1 for output levels and π_0 for a profit level. The results are of general validity, not true for only a particular set of numerical values.

At the outset, note that Figure 2.2 contains two panels. The upper panel (panel A) shows the relationship between total profit and output, while the lower panel (panel B) shows the relationship between average profit and marginal profit, on the one hand, and output, on the other. The horizontal scale of panel A is the same as that of panel B, the result being that a given output, like Q_0, is the same distance from the origin (along the horizontal axis) in panel A as in panel B.

In practice, one seldom is presented with data concerning both (1) the relationship between total profit and output and (2) the relationship between average profit and output, because it is relatively simple to derive the latter relationship from the former. How can this be done? Take any output, say Q_0. *At this output, average profit equals the slope of the straight line from the origin to point E, the point on the total profit curve corresponding to output* Q_0. To see that this is the case, note that average profit at this output level equals π_0/Q_0, where π_0 is the level of total profit if output is Q_0. Because the slope of any straight line equals the vertical distance between two points on the line divided by the horizontal distance between them, the slope of the line from the origin to point E equals π_0/Q_0.[1] Thus, the slope of line 0*E* equals average profit at this output. (In other words, K_0 in panel B of Figure 2.2 is equal to the slope of line 0*E*.) To determine the relationship between average profit and output from the relationship between total profit and output, we repeat this procedure for each level of output, not Q_0 alone. The resulting average profit curve is shown in panel B.

Turning to the relationship between marginal profit and output (in panel B), it is relatively simple to derive this relationship too from the relationship between total profit and output (in panel A). Take any output, say Q_1. *At this output, marginal profit equals the slope of the tangent to the total profit curve (in panel A) at the point where output is* Q_1. In other words, marginal profit equals the slope of line *T* in Figure 2.2, which is tangent to the total profit curve at point *G*. As a first step toward seeing why this is true, consider Figure 2.3, which provides a magnified picture of the total profit curve in the neighborhood of point *G*.

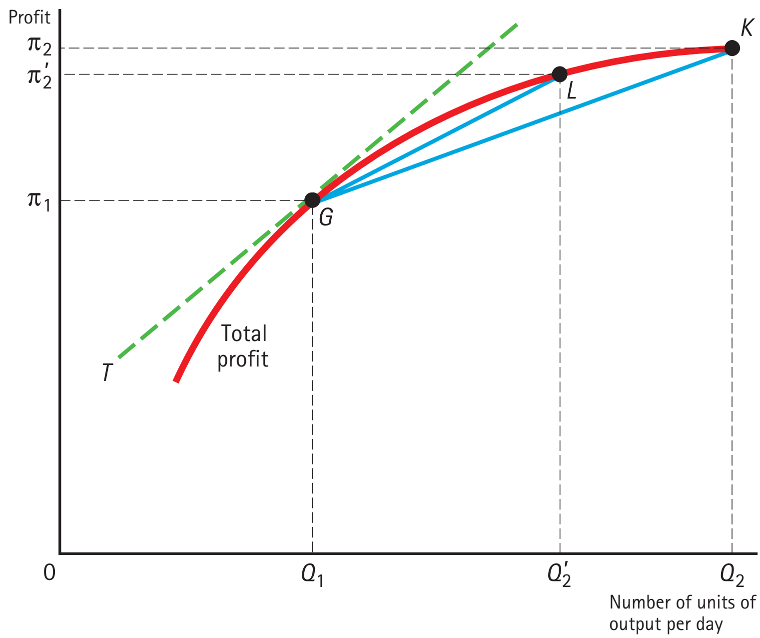
Recall that marginal profit is defined as the extra profit resulting from a very small increase (specifically, a one-unit increase) in output. If output

[1]The vertical distance between the origin and point E equals π_0, and the horizontal distance between these two points equals Q_0. Therefore, the vertical distance divided by the horizontal distance equals π_0/Q_0.

FIGURE 2.3

Marginal Profit Equals the Slope of the Tangent to the Total Profit Curve

As the distance between Q_1 and Q_2 becomes extremely small, the slope of line *T* becomes a very good estimate of $(\pi_2 - \pi_1)/(Q_2 - Q_1)$.



increases from Q_1 to Q_2, total profit increases from π_1 to π_2, as shown in Figure 2.3. Therefore, the extra profit per unit of output is $(\pi_2 - \pi_1)/(Q_2 - Q_1)$, which is the slope of the *GK* line. But this increase in output is rather large. Suppose that we decrease Q_2 so that it is closer to Q_1. In particular, let the new value of Q_2 be Q_2'. If output increases from Q_1 to Q_2', the extra profit per unit of output equals $(\pi_2' - \pi_1)/(Q_2' - Q_1)$, which is the slope of the *GL* line. If we further decrease Q_2 until the distance between Q_1 and Q_2 is extremely small, the slope of the tangent (line *T*) at point *G* becomes a very good estimate of $(\pi_2 - \pi_1)/(Q_2 - Q_1)$. In the limit, for changes in output in a very small neighborhood around Q_1, the slope of the tangent is marginal profit. (This slope equals K_1 in panel B of Figure 2.2.) To determine the relationship between marginal profit and output from the relationship between total profit and output, we repeat this procedure for each level of output, not Q_1 alone. The resulting marginal profit curve is shown in panel B of Figure 2.2.

Sometimes one is given an average profit curve like that in panel B but not the total profit curve. To derive the latter curve from the former, note that total profit equals average profit times output. Hence, if output equals Q_0, total profit equals K_0 times Q_0. In other words, π_0 in panel A equals the area of rectangle $0K_0HQ_0$ in panel B. To derive the relationship between total profit and output from the relationship between average profit and output, we repeat this procedure for each level of output. That is, we find the area of the appropriate rectangle of this sort corresponding to each output, not Q_0 alone. The resulting total profit curve is shown in panel A.

Finally, two further points should be made concerning the total, average, and marginal profit curves in Figure 2.2. First, you should be able to tell by a glance at panel A that marginal profit increases as output rises from zero to Q_3 and that it decreases as output rises further. Why is this so obvious from panel A? Because the slope of the total profit curve increases as one moves from the origin to point *D*. In other words, lines drawn tangent to the total profit curve become steeper as one moves from the origin to point *D*. Since marginal profit equals the slope of this tangent, it must increase as output rises from zero to Q_3. To the right of point *D*, the slope of the total profit curve decreases as output increases. That is, lines drawn tangent to the total profit curve become less steep as one moves to the right of point *D*. Consequently, since marginal profit equals the slope of this tangent, it too must decrease when output rises beyond Q_3.

Second, panel B of Figure 2.2 confirms the following proposition: *The average profit curve must be rising if it is below the marginal profit curve, and it must be falling if it is above the marginal profit curve.* At output levels below Q_4, the average profit curve is below the marginal profit curve; therefore, the average profit curve is rising because the higher marginal profits are pulling up the average profits. At output levels above Q_4, the average profit curve is above the marginal profit curve; therefore, the average profit curve is falling because the lower marginal profits are pulling down the average profits. At Q_4, the straight line drawn from the origin to point *M* is just tangent to the total cost curve. Therefore, the average profit and marginal profit are equal at output Q_4.

The Concept of a Derivative

In the case of the Roland Corporation, we used Table 2.2 (which shows the relationship between the firm's output and profit) to find the profit-maximizing output level. Frequently, a table of this sort is too cumbersome or inaccurate to be useful for this purpose. Instead, an equation is used to represent the relationship between the variable we are trying to maximize (in this case, profit) and the variable or variables under the control of the decision maker (in this case, output). Given an equation of this sort, the powerful concepts and techniques of differential calculus can be employed to find optimal solutions to the decision maker's problem.

In previous sections, we defined the *marginal value* as the change in the dependent variable resulting from a one-unit change in an independent variable. If Y is the dependent variable and X is the independent variable,

$$Y = f(X) \tag{2.3}$$

according to the notation in equation (2.1). Using Δ (called *delta*) to denote change, a change in the independent variable can be expressed as ΔX, and a change in the dependent variable can be expressed as ΔY. Thus, the marginal value of Y can be estimated by

$$\frac{\text{Change in } Y}{\text{Change in } X} = \frac{\Delta Y}{\Delta X} \tag{2.4}$$

For example, if a two-unit increase in X results in a one-unit increase in Y, $\Delta X = 2$ and $\Delta Y = 1$; this means that the marginal value of Y is about $^1/_2$. That is, the dependent variable Y increases by about $^1/_2$ if the independent variable X increases by 1.[2]

Unless the relationship between Y and X can be represented as a straight line (as in Figure 2.4), the value of $\Delta Y/\Delta X$ is not constant. For example, consider the relationship between Y and X in Figure 2.5. If a movement occurs from point G to point H, a relatively small change in X (from X_1 to X_2) is associated with a big change in Y (from Y_1 to Y_2). Therefore, between points G and H, the value of $\Delta Y/\Delta X$, which equals $(Y_2 - Y_1)/(X_2 - X_1)$, is relatively large. On the other hand, if a movement occurs from point K to point L, a relatively large change in X (from X_3 to X_4) is associated with a small change in Y (from Y_3 to Y_4). Consequently, between points K and L, the value of $\Delta Y/\Delta X$, which equals $(Y_4 - Y_3)/(X_4 - X_3)$, is relatively small.

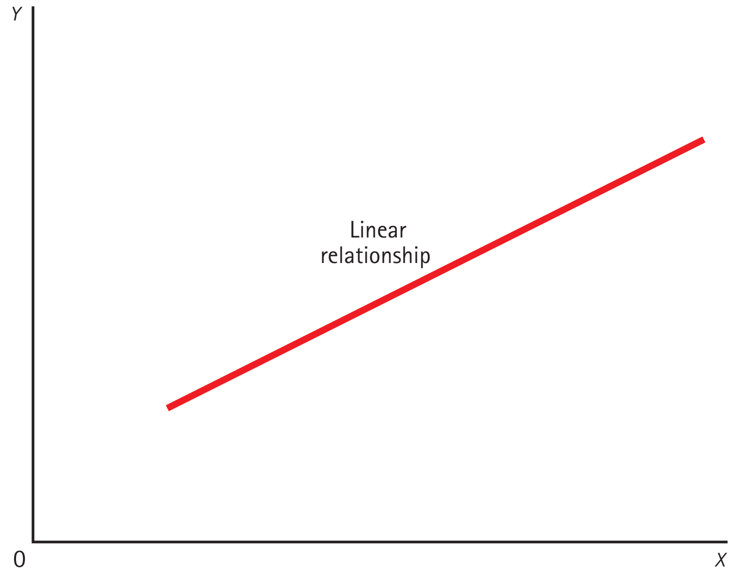
The value of $\Delta Y/\Delta X$ is related to the steepness or flatness of the curve in Figure 2.5. Between points G and H, the curve is relatively *steep*; this means that a *small* change in X results in a *large* change in Y. Consequently, $\Delta Y/\Delta X$ is relatively large. Between points K and L, the curve is relatively *flat*; this means that a *large* change in X results in a *small* change in Y. Consequently, $\Delta Y/\Delta X$ is relatively small.

[2]Why do we say that Y increases by about 1/2, rather than by exactly 1/2? Because Y may not be linearly related to X. More is said on this subject in the next paragraph of the text.

FIGURE 2.4

Linear Relationship between *Y* and *X*

The relationship between *Y* and *X* can be represented as a straight line.



The derivative of Y with respect to X is defined as the limit of $\Delta Y/\Delta X$ as ΔX approaches zero. Since the derivative of *Y* with respect to *X* is denoted by dY/dX, this definition can be restated as

$$\frac{dY}{dX} = \lim_{\Delta X \to 0} \frac{\Delta Y}{\Delta X} \tag{2.5}$$

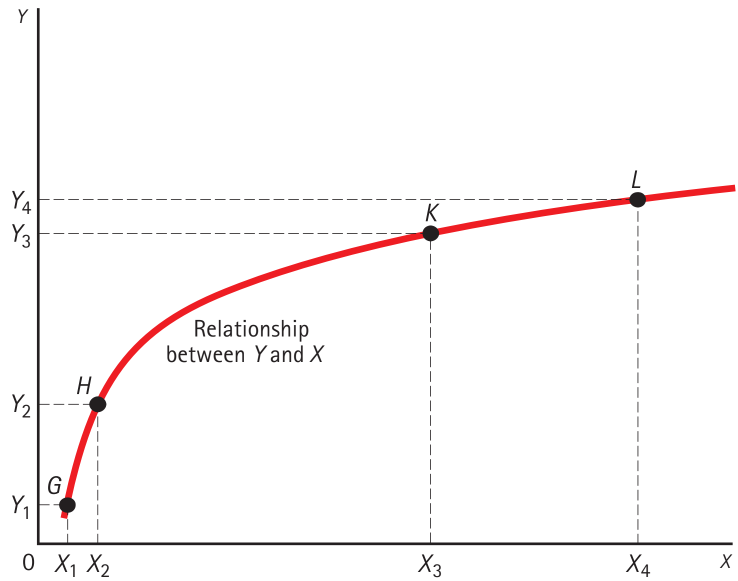
which is read "The derivative of *Y* with respect to *X* equals the limit of the ratio $\Delta Y/\Delta X$ as ΔX approaches zero." To understand what is meant by a limit, consider the function $(X - 2)$. What is the limit of this function as *X* approaches 2? Clearly, as *X* gets closer and closer to 2, $(X - 2)$ gets closer and closer to zero. What is the limit of this function as *X* approaches zero? Clearly, as *X* gets closer and closer to zero, $(X - 2)$ gets closer and closer to -2.

Graphically, the derivative of *Y* with respect to *X* equals the *slope* of the curve showing *Y* (on the vertical axis) as a function of *X* (on the horizontal axis). To see this, suppose we want to find the value of the derivative of *Y* with respect to *X* when *X* equals X_5 in Figure 2.6. A rough measure is the

FIGURE 2.5

How the Value of $\Delta Y/\Delta X$ Varies Depending on the Steepness or Flatness of the Relationship between Y and X

Between points G and H, since the curve is steep, $\Delta Y/\Delta X$ is large. Between points K and L, since the curve is flat, $\Delta Y/\Delta X$ is small.



value of $\Delta Y/\Delta X$ when a movement is made from point A to point C; this measure equals

$$(Y_7 - Y_5)/(X_7 - X_5)$$

which is the slope of the AC line. A better measure is the value of $\Delta Y/\Delta X$ when a movement is made from point A to point B; this measure equals

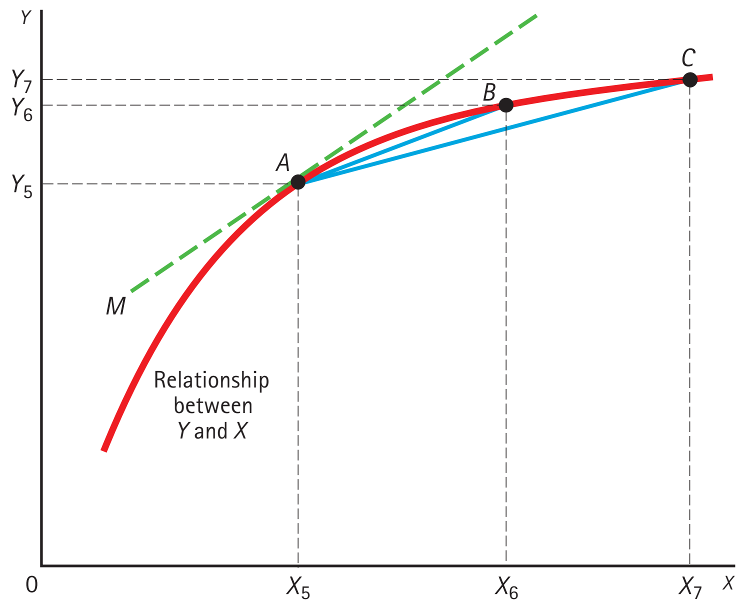
$$(Y_6 - Y_5)/(X_6 - X_5)$$

which is the slope of the AB line. Why is the latter measure better than the former? Because the distance between points A and B is less than the distance between points A and C, what we want is the value of $\Delta Y/\Delta X$ when ΔX is as small as possible. Clearly, *in the limit, as ΔX approaches zero, the ratio $\Delta Y/\Delta X$ is equal to the slope of the line M, which is drawn tangent to the curve at point A.*

FIGURE 2.6

Derivative as the Slope of the Curve

When X equals X_5, the derivative of Y with respect to X equals the slope of line M, the tangent to the curve at point A.



How to Find a Derivative

Managers want to know how to optimize the performance of their organizations. If Y is some measure of organizational performance and X is a variable under a particular manager's control, he or she would like to know the value of X that maximizes Y. To find out, it is very useful, as we shall see in the next section, to know the derivative of Y with respect to X. In this section, we learn how to find this derivative.

Derivatives of Constants

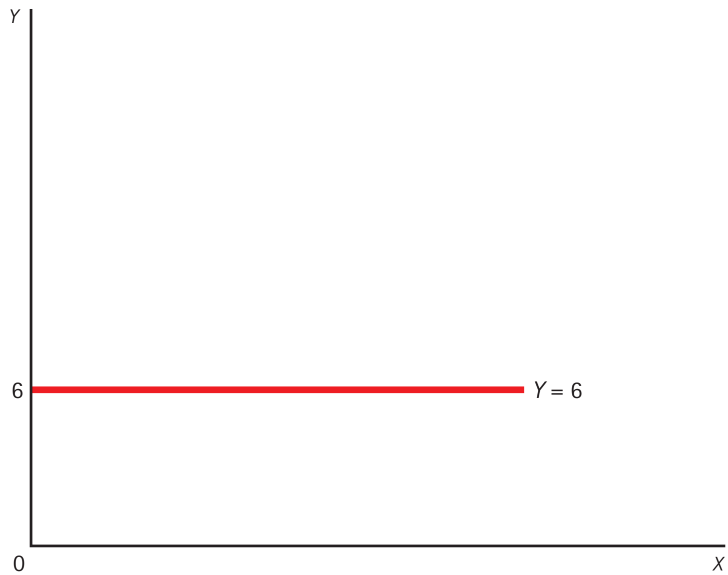
If the dependent variable Y is a constant, its derivative with respect to X is always zero. That is, if $Y = a$ (where a is a constant),

$$\frac{dY}{dX} = 0 \tag{2.6}$$

FIGURE 2.7

Case in Which $Y = 6$

In this case, dY/dX equals zero, since the slope of this horizontal line equals zero.



EXAMPLE Suppose that $Y = 6$, as shown in Figure 2.7. Since the value of Y does not change as X varies, dY/dX must be equal to zero. To see how this can also be shown geometrically, recall from the previous section that dY/dX equals the slope of the curve showing Y as a function of X. As is evident from Figure 2.7, this slope equals zero, which means that dY/dX must equal zero.

Derivatives of Power Functions

A power function can be expressed as

$$Y = aX^b$$

where a and b are constants. If the relationship between X and Y is of this kind, the derivative of Y with respect to X equals b times a multiplied by X raised to the $(b - 1)$ power:

$$\frac{dY}{dX} = baX^{b-1} \tag{2.7}$$

EXAMPLE Suppose that

$$Y = 3X$$

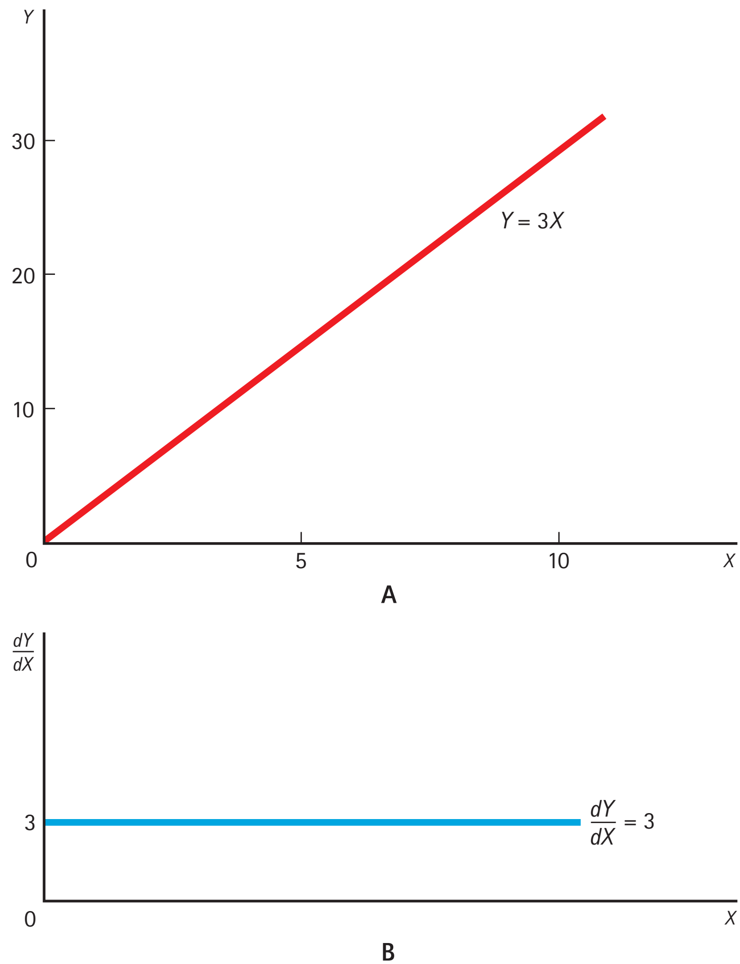
which is graphed in panel A of Figure 2.8. Applying equation (2.7), we find that

$$\frac{dY}{dX} = 1 \cdot 3 \cdot X^0 = 3$$

FIGURE 2.8

Case in Which $Y = 3X$

In this case, dY/dX equals 3, since the slope of the line in panel A equals 3.



since $a = 3$ and $b = 1$. Therefore, the value of dY/dX graphed in panel B of Figure 2.8 is 3, regardless of the value of X. This makes sense, since the slope of the line in panel A is 3, regardless of the value of X. Recall once again from the previous section that dY/dX equals the slope of the curve showing Y as a function of X. In this case (as in Figure 2.7), this "curve" is a straight line.

EXAMPLE Suppose that

$$Y = 2X^2$$

which is graphed in panel A of Figure 2.9. Applying equation (2.7), we find that

$$\frac{dY}{dX} = 2 \cdot 2 \cdot X^1 = 4X$$

since $a = 2$ and $b = 2$. Therefore, the value of dY/dX, which is graphed in panel B of Figure 2.9, is proportional to X. As would be expected, dY/dX is negative when the slope of the curve in panel A is negative and positive when this slope is positive. Why? Because, as we stressed repeatedly, dY/dX equals this slope.

Derivatives of Sums and Differences

Suppose that U and W are two variables, each of which depends on X. That is,

$$U = g(X) \text{ and } W = h(X)$$

The functional relationship between U and X is denoted by g and that between W and X is denoted by h. Suppose further that

$$Y = U + W$$

In other words, Y is the sum of U and W. If so, the derivative of Y with respect to X is equal to the sum of the derivatives of the individual terms:

$$\frac{dY}{dX} = \frac{dU}{dX} + \frac{dW}{dX} \tag{2.8}$$

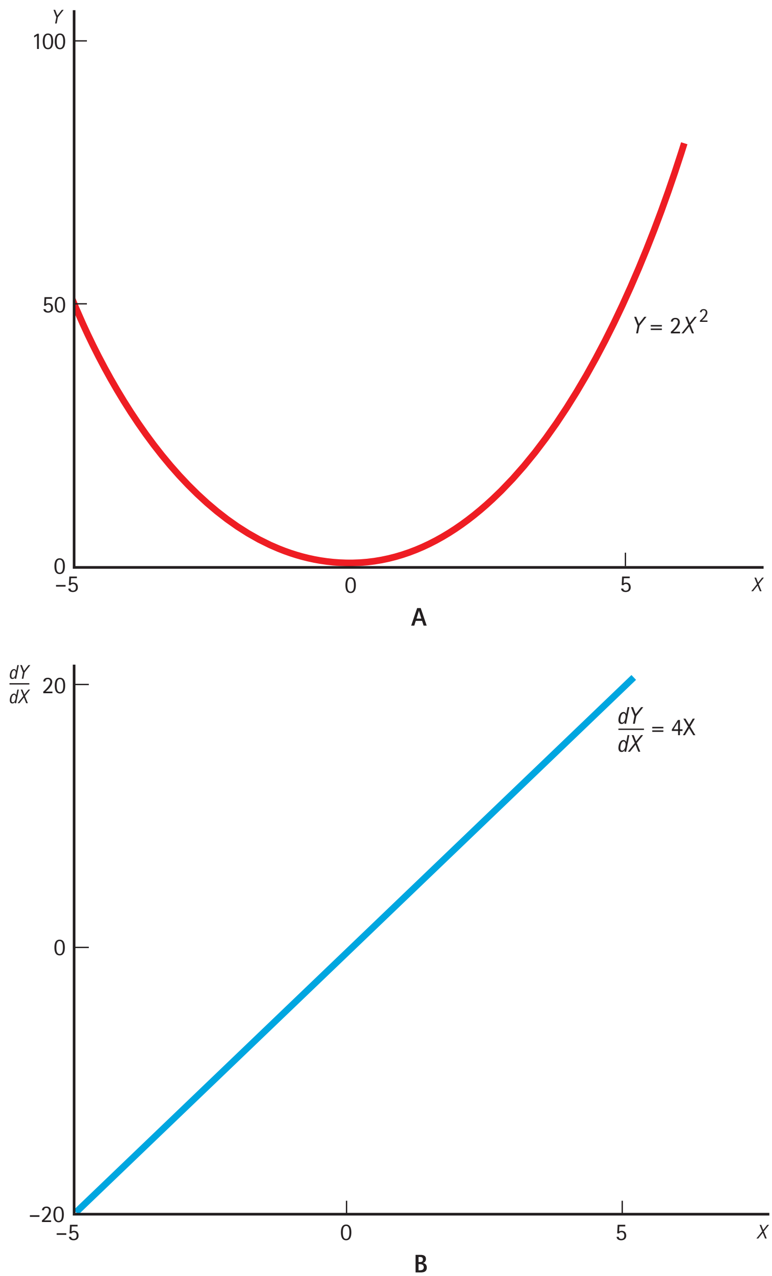
On the other hand, if

$$Y = U - W$$

FIGURE 2.9

Case in Which $Y = 2X^2$

In this case, $dY/dX = 4X$, since the slope of the curve in panel A equals $4X$.



CONCEPTS IN CONTEXT

The Allocation of the TANG Brand Advertising Budget

In the 1960's, US astronauts drank TANG on their trips in space. Forty years later, Kraft Foods sells TANG in more than 80 countries. Eighty five percent of TANG's sales come from developing markets, where Kraft customizes the beverage to the specific taste preferences and nutritional needs of each market.

Managers and analysts use differential calculus to help solve all sorts of problems. Consider the work that Young and Rubicam, the prominent advertising agency, did for one of its Kraft foods (General Foods at the time) accounts, TANG beverages. TANG is the trademark for an instant breakfast drink with an orange flavor. Young and Rubicam did a study to estimate the effects of advertising expenditures on the sales of TANG and found that the relationships between advertising expenditures and sales in two districts were*

$$S_1 = 10 + 5A_1 - 1.5A_1^2$$

and

$$S_2 = 12 + 4A_2 - 0.5A_2^2$$

where S_1 is the sales of TANG (in millions of dollars per year) in the first district, S_2 is its sales in the second district, A_1 is the advertising expenditure on TANG in the first district, and A_2 is the advertising expenditure in the second district.

Young and Rubicam wanted to determine the amount of additional sales that an extra dollar of advertising would generate in each district. To answer this question, the derivative of sales with respect to advertising expenditure was calculated for each district:

$$\frac{dS_1}{dA_1} = 5 - 3A_1$$

and

$$\frac{dS_2}{dA_2} = 4 - A_2$$

Thus in each district, the effect on sales of an extra dollar of advertising expense, depended on the amount spent on advertising. Supposing that $0.5 million was being spent on advertising in the first district and that $1 million was being spent on advertising in the second district:

$$\frac{dS_1}{dA_1} = 5 - 3(0.5) = 3.5$$

and

$$\frac{dS_2}{dA_2} = 4 - 1 = 3$$

Consequently, an extra dollar of advertising generated an extra $3.50 of sales in the first district and an extra $3.00 of sales in the second district.

On the basis of these findings, Young and Rubicam made a number of recommendations to General Foods (recall that TANG is now part of Kraft Foods) concerning the regional allocation of the TANG advertising budget. In particular, it recommended that, if General Foods wanted to boost the total sales of TANG, more should be spent on advertising in the first district and less should be spent on it in the second district. This would not mean an increase in General Foods's total advertising budget, since the extra advertising expenditure in the first district would be offset by the reduced advertising expenditure in the second district.

How did Young and Rubicam come to this conclusion? The fact that an extra dollar of advertising would result in a greater addition to sales in the first district than in the second indicated that a reallocation of the advertising budget was called for. To see this, consider what would happen if a dollar extra were spent on advertising in the first district and a dollar less spent in the second. The result, as indicated previously, would be an extra \$3.50 of sales in the first district and a \$3.00 reduction in sales in the second. The overall effect would be a \$3.50 − \$3.00 = \$0.50 increase in total sales. Thus, if General Foods wanted to increase the sales of TANG beverages, a reallocation of the advertising budget in favor of the first district was to be recommended.†

*Although these equations are of the form derived by Young and Rubicam, the numerical coefficients are hypothetical. For present purposes, this is of no consequence. Our purpose here is to describe the general features of this case and how differential calculus played a role, not the specific numbers. The methods that can be used to estimate such coefficients are described in Chapter 5. Also, note that sales in each district are assumed to depend on the level of advertising in this district only.

†F. DeBruicker, J. Quelch, and S. Ward, *Cases in Consumer Behavior* (2d ed; Englewood Cliffs, NJ: Prentice-Hall, 1986). This case has been simplified in various respects for pedagogical reasons.

the derivative of Y with respect to X is equal to the difference between the derivatives of the individual terms:

$$\frac{dY}{dX} = \frac{dU}{dX} - \frac{dW}{dX} \tag{2.9}$$

EXAMPLE Consider the case in which $U = g(X) = 3X^3$ and $W = h(X) = 4X^2$. If $Y = U + W = 3X^3 + 4X^2$,

$$\frac{dY}{dX} = 9X^2 + 8X \tag{2.10}$$

To see why, recall from equation (2.8) that

$$\frac{dY}{dX} = \frac{dU}{dX} + \frac{dW}{dX} \tag{2.11}$$

Applying equation (2.7),

$$\frac{dU}{dX} = 9X^2 \text{ and } \frac{dW}{dX} = 8X$$

Substituting these values of the derivatives into equation (2.11), equation (2.10) follows.

EXAMPLE Suppose that $Y = U - W$, where $U = 8X^2$ and $W = 9X$. Then

$$\frac{dY}{dX} = 16X - 9$$

since, according to equation (2.9),

$$\frac{dY}{dX} = \frac{dU}{dX} - \frac{dW}{dX}$$

and, applying equation (2.7),

$$\frac{dU}{dX} = 16X \text{ and } \frac{dW}{dX} = 9$$

Derivatives of Products

The derivative of the product of two terms is equal to the sum of the first term multiplied by the derivative of the second plus the second term times the derivative of the first. Consequently, if $Y = UW$,

$$\frac{dY}{dX} = U\frac{dW}{dX} + W\frac{dU}{dX} \tag{2.12}$$

EXAMPLE If $Y = 6X(3 - X^2)$, we can let $U = 6X$ and $W = 3 - X^2$; then

$$\begin{aligned}\frac{dY}{dX} &= 6X\frac{dW}{dX} + (3 - X^2)\frac{dU}{dX} \\ &= 6X(-2X) + (3 - X^2)(6) \\ &= -12X^2 + 18 - 6X^2 \\ &= 18 - 18X^2\end{aligned}$$

The first term, $6X$, is multiplied by the derivative of the second term, $-2X$, and the result is added to the second term, $3 - X^2$, times the derivative of the first, 6. As indicated, the result is $18 - 18X^2$.

Derivatives of Quotients

If $Y = U/W$, the derivative of Y with respect to X equals

$$\frac{dY}{dX} = \frac{W(dU/dX) - U(dW/dX)}{W^2}. \tag{2.13}$$

In other words, the derivative of the quotient of two terms equals the denominator times the derivative of the numerator minus the numerator times the derivative of the denominator—all divided by the square of the denominator.

EXAMPLE Consider the problem of finding the derivative of the expression

$$Y = \frac{5X^3}{3 - 4X}$$

If we let $U = 5X^3$ and $W = 3 - 4X$,

$$\frac{dU}{dX} = 15X^2 \text{ and } \frac{dW}{dX} = -4$$

Consequently, applying equation (2.13),

$$\begin{aligned} \frac{dY}{dX} &= \frac{(3 - 4X)(15X^2) - 5X^3(-4)}{(3 - 4X)^2} \\ &= \frac{45X^2 - 60X^3 + 20X^3}{(3 - 4X)^2} \\ &= \frac{45X^2 - 40X^3}{(3 - 4X)^2} \end{aligned}$$

Derivatives of a Function of a Function (Chain Rule)[3]

Sometimes a variable depends on another variable, which in turn depends on a third variable. For example, suppose that $Y = f(W)$ and $W = g(X)$. Under these circumstances, the derivative of Y with respect to X equals

$$\frac{dY}{dX} = \left(\frac{dY}{dW}\right)\left(\frac{dW}{dX}\right) \tag{2.14}$$

In other words, to find this derivative, we find the derivative of Y with respect to W and multiply it by the derivative of W with respect to X.

EXAMPLE Suppose that $Y = 4W + W^3$ and $W = 3X^2$. To find dY/dX we begin by finding dY/dW and dW/dX:

$$\begin{aligned} \frac{dY}{dW} &= 4 + 3W^2 \\ &= 4 + 3(3X^2)^2 \\ &= 4 + 27X^4 \\ \frac{dW}{dX} &= 6X \end{aligned}$$

[3]This section can be skipped without loss of continuity.

Then, to find dY/dX, we multiply dY/dW and dW/dX:

$$\frac{dY}{dX} = (4 + 27X^4)(6X)$$
$$= 24X + 162X^5$$

Using Derivatives to Solve Maximization and Minimization Problems

Having determined how to find the derivative of Y with respect to X, we now take up the way to determine the value of X that maximizes or minimizes Y. *The central point to recognize is that a maximum or minimum point can occur only if the slope of the curve showing Y on the vertical axis and X on the horizontal axis equals zero.* To see this, suppose that Y equals the profit of the Monroe Company and X is its output level. If the relationship between Y and X is as shown by the curve in panel A of Figure 2.10, the maximum value of Y occurs when $X = 10$, and at this value of X the slope of the curve equals zero.

Since the derivative of Y with respect to X equals the slope of this curve, it follows that Y can be a maximum or minimum only if this derivative equals zero. To see that Y really is maximized when this derivative equals zero, note that the relationship between Y and X in Figure 2.10 is

$$Y = -50 + 100X - 5X^2 \tag{2.15}$$

which means that

$$\frac{dY}{dX} = 100 - 10X \tag{2.16}$$

Therefore, if this derivative equals zero,

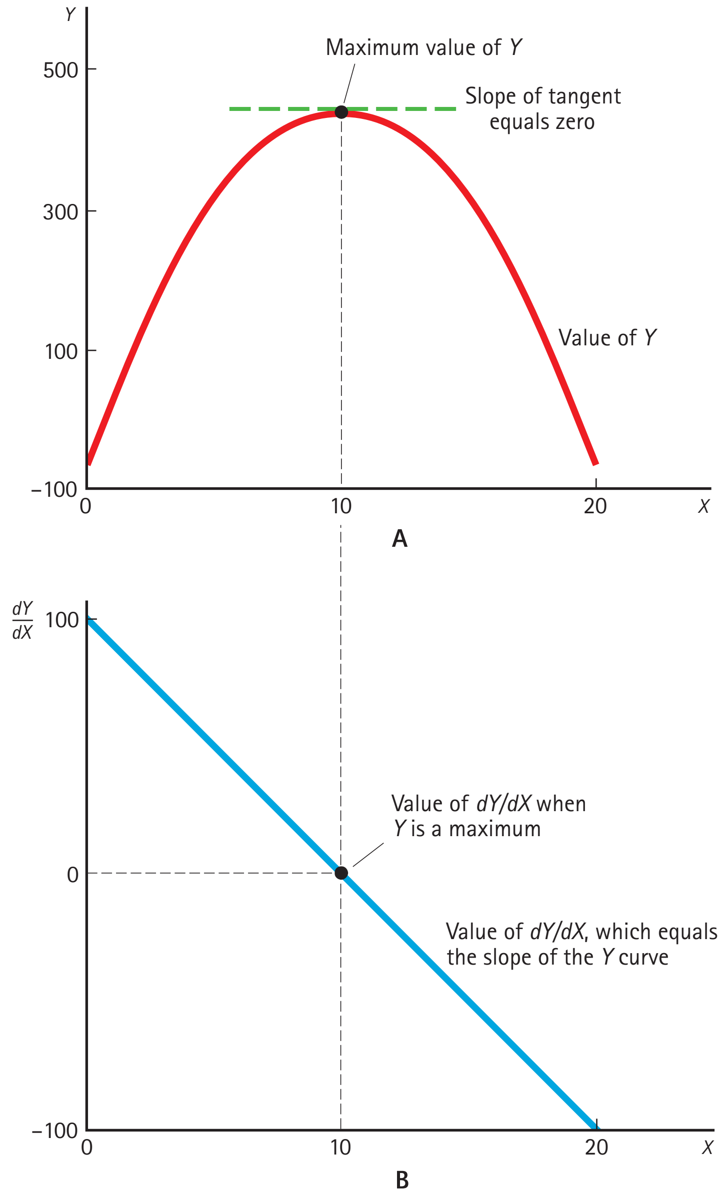
$$100 - 10X = 0$$
$$X = 10$$

This is the value of X where Y is maximized, as we saw in the previous paragraph. The key point here is that, *to find the value of X that maximizes or minimizes Y, we must find the value of X where this derivative equals zero.* Panel B of Figure 2.10 shows graphically that this derivative equals zero when Y is maximized.

FIGURE 2.10

Value of the Derivative When Y Is a Maximum

When Y is a maximum (at $X = 10$), dY/dX equals zero.



However, on the basis of only the fact that this derivative is zero, one cannot distinguish between a point on the curve where Y is maximized and a point where Y is minimized. For example, in Figure 2.11, this derivative is zero both when $X = 5$ and when $X = 15$. In the one case (when $X = 15$), Y is a maximum; in the other case (when $X = 5$), Y is a minimum. To distinguish between a maximum and a minimum, one must find the *second derivative of Y with respect to X, which is denoted* d^2Y/dX^2 *and is the derivative of* dY/dX. For example, in Figure 2.10, the second derivative of Y with respect to X is the derivative of the function in equation (2.16); therefore, it equals -10.

The second derivative measures the slope of the curve showing the relationship between dY/dX (the first derivative) and X. Just as the first derivative (that is, dY/dX) measures the slope of the Y curve in panel A of Figure 2.11, so the second derivative (that is, d^2Y/dX^2) measures the slope of the dY/dX curve in panel B of Figure 2.11. In other words, just as the first derivative measures the slope of the total profit curve, the second derivative measures the slope of the marginal profit curve. The reason why the second derivative is so important is that it is always *negative* at a point of *maximization* and always *positive* at a point of *minimization*. Therefore, *to distinguish between maximization and minimization points, all we have to do is determine whether the second derivative at each point is positive or negative.*

To understand why the second derivative is always negative at a maximization point and always positive at a minimization point, consider Figure 2.11. When the second derivative is negative, this means that the slope of the dY/dX curve in panel B is negative. Because dY/dX equals the slope of the Y curve in panel A, this in turn means that the slope of the Y curve decreases as X increases. At a maximum point, such as when $X = 15$, this must be the case. On the other hand, when the second derivative is positive, this means that the slope of the dY/dX curve in panel B is positive, which is another way of saying that the slope of the Y curve in panel A increases as X increases. At a minimum point, such as when $X = 5$, this must be the case.

EXAMPLE To illustrate how one can use derivatives to solve maximization and minimization problems, suppose that the relationship between profit and output at the Kantor Corporation is

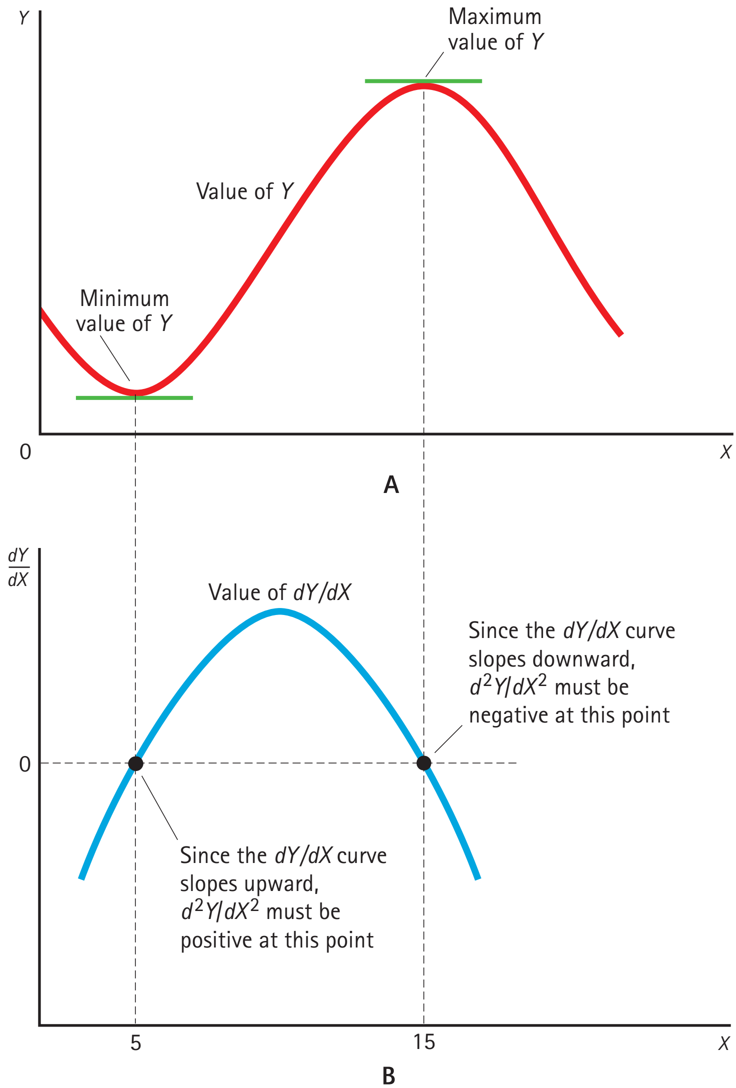
$$Y = -1 + 9X - 6X^2 + X^3$$

where Y equals annual profit (in millions of dollars) and X equals annual output (in millions of units). This equation is valid only for values of X that equal 3 or less; capacity limitations prevent the firm from producing more than 3 million units per year. To find the values of output that maximize or

FIGURE 2.11

Using the Second Derivative to Distinguish Maxima from Minima

At maxima (such as $X = 15$), d^2Y/dX^2 is negative; at minima (such as $X = 5$), d^2Y/dX^2 is positive.



minimize profit, we find the derivative of Y with respect to X and set it equal to zero:

$$\frac{dY}{dX} = 9 - 12X + 3X^2 = 0 \tag{2.17}$$

Solving this equation for X, we find that two values of X—1 and 3—result in this derivative being zero.[4]

To determine whether each of these two output levels maximizes or minimizes profit, we find the value of the second derivative at these two values of X. Taking the derivative of dY/dX, which is shown in equation (2.17) to equal $9 + 12X + 3X^2$, we find that

$$\frac{d^2Y}{dX^2} = -12 + 6X$$

If $X = 1$

$$\frac{d^2Y}{dX^2} = -12 + 6(1) = -6$$

Since the second derivative is negative, profit is a maximum (at 3) when output equals 1 million units. If $X = 3$,

$$\frac{d^2Y}{dX^2} = -12 + 6(3) = 6$$

Since the second derivative is positive, profit is a minimum (at -1) when output equals 3 million units.

[4]If an equation is of the general quadratic form, $Y = aX^2 + bX + c$, the values of X where Y is 0 are

$$X = \frac{-b \pm (b^2 - 4ac)^{0.5}}{2a}$$

In the equation in the text, $a = 3$, $b = -12$, and $c = 9$. Hence,

$$X = \frac{12 \pm (144 - 108)^{0.5}}{6} = 2 \pm 1$$

Therefore, $Y = 0$ when X equals 1 or 3.

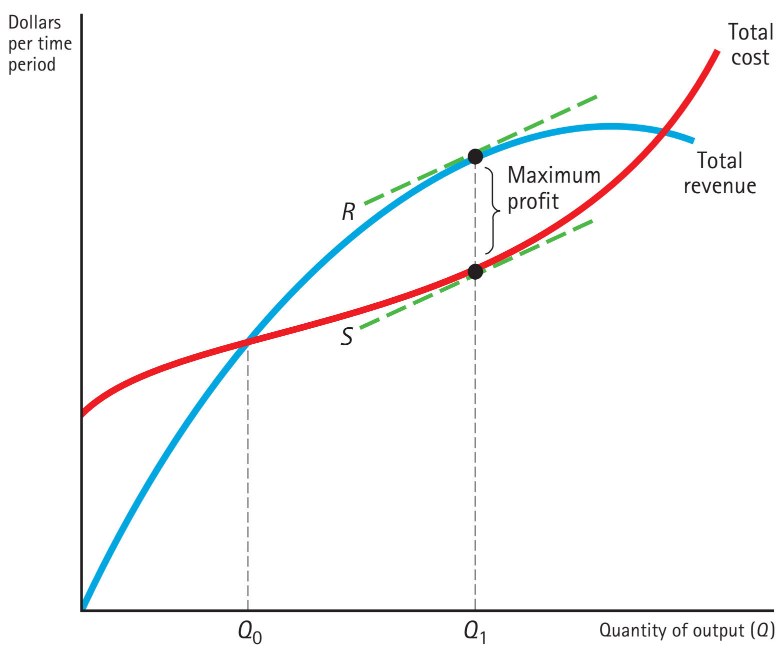
Marginal Cost Equals Marginal Revenue and the Calculus of Optimization

Once you know how elementary calculus can be used to solve optimization problems, it is easy to see that the fundamental rule for profit maximization—set marginal cost equal to marginal revenue—has its basis in the calculus of optimization. Figure 2.12 shows a firm's total cost and total revenue functions. Since total profit equals total revenue minus total cost, it is equal to the vertical distance between the total revenue and total cost curves at any level of output. This distance is maximized at output Q_1, where the slopes of the total revenue and total cost curves are equal. Since the slope of the total revenue curve is marginal revenue and the slope of the total cost curve is marginal cost, this means that profit is maximized when marginal cost equals marginal revenue.

FIGURE 2.12

Marginal-Revenue-Equals-Marginal-Cost Rule for Profit

At the profit-maximizing output of Q_1, marginal revenue (equal to the slope of line *R*) equals marginal cost (the slope of line *S*).



ANALYZING MANAGERIAL DECISIONS

The Optimal Size of a Nursing Home

The nursing home industry takes in about $70 billion a year and is growing rapidly because of the aging of the U.S. population. According to Peter Sidoti of Nat West Securities Corporation, "Nursing homes are the only area in health care where there is a shortage."* A study by Niccie McKay of Trinity University has estimated that the average cost per patient-day of a nursing home (owned by a chain of for-profit homes) is

$$Y = A - 0.16X + 0.00137X^2$$

where X is the nursing home's number of patient-days per year (in thousands) and A is a number that depends on the region in which the nursing home is located (and other such factors) but not on X.

(a) On the basis of the results of this study, how big must a nursing home be (in terms of patient-days) to minimize the cost per patient-day? (b) Show that your result minimizes, rather than maximizes, the cost per patient-day. (c) Is the number of patient-days a good measure of a nursing home's output? Why or why not?

SOLUTION (a) To find the value of X that minimizes Y, we set the derivative of Y with respect to X equal to zero:

$$= -0.16 + 0.00274X = 0$$

Therefore, $X = 0.16/0.00274 = 58.4$ thousands of days.

(b) Since

$$\frac{d^2Y}{dX^2} = 0.00274$$

d^2Y/dX^2 is positive. So, Y must be a minimum, not a maximum, at the point where $dY/dX = 0$.

(c) It is a crude measure, since some patients require far more complex and intensive care than others.†

New York Times, February 27, 1994, p. F5.

†For further discussion, see N. McKay, "The Effect of Chain Ownership on Nursing Home Costs," *Health Services Research*, April 1991.

Inspection of Figure 2.12 shows that Q_1 must be the profit-maximizing output. Outputs below Q_0 result in losses (since total cost exceeds total revenue) and obviously do not maximize profit. As output increases beyond Q_0, total revenue rises more rapidly than total cost, so profit must be going up. So long as the slope of the total revenue curve (which equals marginal revenue) exceeds the slope of the total cost curve (which equals marginal cost), profit will continue to rise as output increases. But, when these slopes become equal (which means that marginal revenue equals marginal cost), profit no longer will rise but will be at a maximum. Since these slopes become equal at an output of Q_1, this must be the profit-maximizing output. After output Q_1, profit decreases (because marginal cost exceeds marginal revenue).

Using calculus, one can readily understand why firms maximize profit by setting marginal cost equal to marginal revenue. The first thing to note is that

$$\pi = \text{TR} - \text{TC},$$

where π equals total profit, TR equals total revenue, and TC equals total cost. Taking the derivative of π with respect to Q (output), we find that

$$\frac{d\pi}{dQ} = \frac{d\text{TR}}{dQ} - \frac{d\text{TC}}{dQ}$$

For π to be a maximum, this derivative must be zero, so it must be true that

$$\frac{d\text{TR}}{dQ} = \frac{d\text{TC}}{dQ} \tag{2.18}$$

And since marginal revenue is defined as $d\text{TR}/dQ$ and marginal cost is defined as $d\text{TC}/dQ$, marginal revenue must equal marginal cost.[5]

Partial Differentiation and the Maximizations of Multivariable Functions

Up to this point, we have been concerned solely with situations in which a variable depends on only one other variable. Although such situations exist, in many cases, a variable depends on a number (often a large number) of other

[5]Two points should be noted. (1) For profit to be maximized, $d^2\pi/dQ^2$ must be negative. (2) The analysis in this section (as well as in earlier sections) results in the determination of a *local* maximum. Sometimes a local maximum is not a global maximum. For example, under some circumstances, the profit-maximizing (or loss-minimizing) output is zero, as we shall see in Chapter 10.

CONCEPTS IN CONTEXT

An Alleged Blunder in the Stealth Bomber's Design

Managerial economics is of great use in the aerospace industry, but this does not mean that errors will not occur sometimes. The B-2 "Stealth" bomber cost billions of dollars to develop. According to Joseph Foa, an emeritus professor of engineering at George Washington University, its design is fundamentally flawed because two aerodynamicists made a mistake: They mistook a minimum point for a maximum point.

The B-2 is basically a jet-powered "flying wing" aircraft. In a secret study for the Air Force, the two aerodynamicists, William Sears and Irving Ashkenas (then at the Northrop Corporation), used mathematical formulas to determine how an aircraft's volume should be proportioned between wing and fuselage to maximize its range. Taking the derivative of range with respect to volume, they found that this derivative equaled zero when the total volume was almost all in the wing. Hence, they concluded that a "flying wing" design would maximize range.

But in a subsequent analysis, Foa showed that the second derivative was positive, not negative, under these circumstances. Hence, the "flying wing" design *minimized* the range; it did not maximize it. In Foa's words, "The flying wing was the aerodynamically worst possible choice of configuration."

This is a very interesting example of how important it is to look at the second derivative to make sure that you do not confuse a maximization point with a minimization point. While the backers of the B-2 bomber claim that it is a good plane despite this error, no one denies that the error is an embarrassment.*

*This discussion is based on W. Biddle, "Skeleton Alleged in the Stealth Bomber's Closet," *Science*, May 12, 1989.

variables, not just one. For example, the Merrimack Company produces two goods, and its profit depends on the amount that it produces of each good. That is,

$$\pi = f(Q_1, Q_2) \tag{2.19}$$

where π is the firm's profit, Q_1 is its output of the first good, and Q_2 is its output of the second good.

To find the value of each of the independent variables (Q_1 and Q_2 in this case) that maximizes the dependent variable (π in this case), we need to know the marginal effect of each independent variable on the dependent variable, *holding constant the effect of all other independent variables*. For example, in this case, we need to know the marginal effect of Q_1 on π when Q_2 is held constant, and we need to know the marginal effect of Q_2 on π when Q_1 is held constant. To get this information, we obtain the partial derivative of π with respect to Q_1 and the partial derivative of π with respect to Q_2.

To obtain the partial derivative of π with respect to Q_1, denoted $\partial\pi/\partial Q_1$, one applies the rules for finding a derivative (on pages 47–60) to equation (2.19), but treats Q_2 as a constant. Similarly, to obtain the partial derivative of π with respect to Q_2, denoted $\partial\pi/\partial Q_2$, one applies these rules to equation (2.19), but treats Q_1 as a constant.

EXAMPLE Suppose that the relationship between the Merrimack Company's profit (in thousands of dollars) and its output of each good is

$$\pi = -20 + 113.75Q_1 + 80Q_2 - 10Q_1^2 - 10Q_2^2 - 5Q_1Q_2 \tag{2.20}$$

To find the partial derivative of π with respect to Q_1, we treat Q_2 as a constant and find that

$$\frac{\partial\pi}{\partial Q_1} = 113.75 - 20Q_1 - 5Q_2$$

To find the partial derivative of π with respect to Q_2, we treat Q_1 as a constant and find that

$$\frac{\partial\pi}{\partial Q_2} = 80 - 20Q_2 - 5Q_1$$

Once we have derived the partial derivatives, it is relatively simple to determine the values of the independent variables that maximize the dependent variable. All we have to do is set *all the partial derivatives equal to zero*. In the case of the Merrimack Company,

$$\frac{\partial\pi}{\partial Q_1} = 113.75 - 20Q_1 - 5Q_2 = 0 \tag{2.21}$$

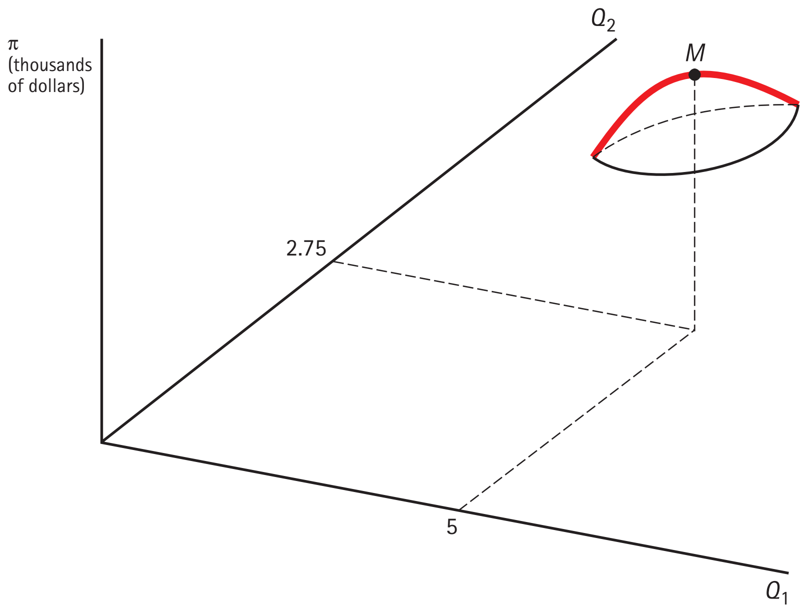
$$\frac{\partial\pi}{\partial Q_2} = 80 - 20Q_2 - 5Q_1 = 0 \tag{2.22}$$

Equations (2.21) and (2.22) are two equations in two unknowns. Solving them simultaneously, we find that profit is maximized when $Q_1 = 5.0$ and $Q_2 = 2.75$. In other words, to maximize profit, the firm should produce 5.0 units of the

FIGURE 2.13

Relationship between π, Q_1, and Q_2

At M, the point where π is a maximum, the surface representing this relationship is flat; its slope with regard to either Q_1 or Q_2 is zero.



first good and 2.75 units of the second good per period of time. If it does this, its profit will equal \$374.375 thousand per period of time.[6]

To see why all the partial derivatives should be set equal to zero, consider Figure 2.13, which shows the relationship in equation (2.20) between π, Q_1, and Q_2, in the range where π is close to its maximum value. As you can see, this relationship is represented by a three-dimensional surface. The maximum value of π is at point M, where this surface is level. A plane tangent to this surface at point M is parallel to the Q_1Q_2 plane; in other words, its slopes with respect to either Q_1 or Q_2 must be zero. Since the partial derivatives in equations (2.21) and (2.22) equal these slopes, they too must equal zero at the maximum point, M.[7]

[6]Inserting 5.0 for Q_1 and 2.75 for Q_2 in equation (2.20), we find that

$$\pi = -20 + 113.75(5) + 80(2.75) - 10(5)^2 - 10(2.75)^2 - 5(5)(2.75) = 374.375$$

[7]The second-order conditions for distinguishing maxima from minima can be found in any calculus book. For present purposes, a discussion of these conditions is not essential. Also note that the techniques presented in this section result in a local maximum, not necessarily a global maximum (recall footnote 5).

ANALYZING MANAGERIAL DECISIONS

The Effects of Advertising on the Sales of TANG

On page 56, we encountered Young and Rubicam's study to estimate the effects of advertising expenditures on the sales of TANG, an instant breakfast drink marketed by General Foods (currently marketed by Kraft Foods). Specifically, the agency found that the relationship between advertising expenditures and sales in two districts were

$$S_1 = 10 + 5A_1 - 1.5A_1^2$$

and

$$S_2 = 12 + 4A_2 - 0.5A_2^2$$

where S_1 is TANG's sales (in millions of dollars per year) in the first district, S_2 is its sales in the second district, A_1 is the advertising expenditure (in millions of dollars per year) on TANG in the first district, and A_2 is the advertising expenditure in the second district.

(a) If General Foods wants to maximize TANG's sales in the first district, how much should it spend on advertising? (b) If General Foods wants to maximize TANG's sales in the second district, how much should it spend on advertising? (c) Show that your answers to parts a and b maximize, rather than minimize, sales. (d) Would you recommend that General Foods attempt to maximize TANG's sales? Why or why not?

SOLUTION (a) To find the value of A_1 that maximizes S_1, we set the derivative of S_1 with respect to A_1 equal to zero:

$$\frac{dS_1}{dA_1} = 5 - 3A_1 = 0$$

Here, $A_1 = 5/3$ million dollars. (b) To find the value of A_2 that maximizes S_2, we set the derivative of S_2 with respect to A_2 equal to zero:

$$\frac{dS_2}{dA_2} = 4 - A_2 = 0$$

Here, $A_2 = 4$ million dollars. (c) Since $d^2S_1/dA_1^2 = -3$, S_1 must be a maximum at the point where $dS_1/dA_1 = 0$. Since $d^2S_2/dA_2^2 = -1$, S_2 must be a maximum at the point where $dS_2/dA_2 = 0$. If S_1 and S_2 were minimized, not maximized, the second derivatives would be positive, not negative. (d) No. As stressed in Chapter 1, firms generally are assumed to be interested in maximizing profit, not sales. In general, a firm is unlikely to increase its sales if it means a decrease in its profits. But, in some cases, a firm may do this because, although profits may fall in the short run, they may increase in the long run. For example, a firm may make some sales at a loss to gain customers who eventually will enhance the firm's profits.*

*See F.S. DeBruicker, J.A. Quelch, and S. Ward, *Cases in Consumer Behavior* (2nd ed.; Englewood Cliffs, NJ: Prentice-Hall, 1986).

Constrained Optimization

As we learned in Chapter 1, managers of firms and other organizations generally face constraints that limit the options available to them. A production manager may want to minimize his or her firm's costs but may not be permitted to produce less than is required to meet the firm's contracts with its customers. The top managers of a firm may want to maximize profits, but in the short run, they may be unable to change its product or augment its plant and equipment.

Constrained optimization problems of this sort can be solved in a number of ways. In relatively simple cases in which there is only one constraint, one can use this constraint to express one of the decision variables—that is, one of the variables the decision maker can choose—as a function of the other decision variables. Then, one can apply the techniques for unconstrained optimization described in the previous sections. In effect, what one does is convert the problem to one of unconstrained maximization or minimization.

To illustrate, suppose that the Kloster Company produces two products and that its total cost equals

$$TC = 4Q_1^2 + 5Q_2^2 - Q_1Q_2 \qquad (2.23)$$

where Q_1 equals its output per hour of the first product and Q_2 equals its output per hour of the second product. Because of commitments to customers, the number produced of both products combined cannot be less than 30 per hour. Kloster's president wants to know what output levels of the two products minimize the firm's costs, given that the output of the first product plus the output of the second product equals 30 per hour.

This constrained optimization problem can be expressed as follows:

$$\begin{aligned} &\text{Minimize} && TC = 4Q_1^2 + 5Q_2^2 - Q_1Q_2 \\ &\text{subject to} && Q_1 + Q_2 = 30 \end{aligned}$$

Of course, the constraint is that $(Q_1 + Q_2)$ must equal 30. Solving this constraint for Q_1, we have

$$Q_1 = 30 - Q_2$$

Substituting $(30 - Q_2)$ for Q_1 in equation (2.23), it follows that

$$\begin{aligned} TC &= 4(30 - Q_2)^2 + 5Q_2^2 - (30 - Q_2)Q_2 \\ &= 4(900 - 60Q_2 + Q_2^2) + 5Q_2^2 - 30Q_2 + Q_2^2 \\ TC &= 3600 - 270Q_2 + 10Q_2^2 \qquad (2.24) \end{aligned}$$

The methods of unconstrained optimization just described can be used to find the value of Q_2 that minimizes TC. As indicated in earlier sections, we must obtain the derivative of TC with respect to Q_2 and set it equal to zero:

$$\frac{d\text{TC}}{dQ_2} = -270 + 20Q_2 = 0$$
$$20Q_2 = 270$$
$$Q_2 = 13.5$$

To make sure that this is a minimum, not a maximum, we obtain the second derivative, which is

$$\frac{d^2\text{TC}}{dQ_2^2} = 20$$

Since this is positive, we have found a minimum.

To find the value of Q_1 that minimizes total cost, recall that the constraint requires that

$$Q_1 + Q_2 = 30$$

which means that

$$Q_1 = 30 - Q_2$$

Since we know that the optimal value of Q_2 is 13.5, it follows that the optimal value of Q_1 must be

$$Q_1 = 30 - 13.5 = 16.5$$

Summing up, if the Kloster Company wants to minimize total cost subject to the constraint that the sum of the output levels of its two products be 30, it should produce 16.5 units of the first product and 13.5 units of the second product per hour.[8] (In other words, it should produce 33 units of the first product and 27 units of the second product every 2 hours.)

[8]Substituting 16.5 for Q_1 and 13.5 for Q_2 in equation (2.23), the firm's total cost will equal

$$\begin{aligned}\text{TC} &= 4(16.5)^2 + 5(13.5)^2 - (16.5)(13.5)\\ &= 4(272.25) + 5(182.25) - 222.75\\ &= 1089 + 911.25 - 222.75\\ &= 1777.5, \quad \text{or} \quad \$1{,}777.50\end{aligned}$$

Lagrangian Multipliers[9]

If the technique described in the previous section is not feasible because the constraints are too numerous or complex, the method of Lagrangian multipliers can be used. This method of solving constrained optimization problems involves the construction of an equation—the so-called Lagrangian function—that combines the function to be minimized or maximized and the constraints. This equation is constructed so that two things are true: (1) When this equation is maximized (or minimized), the original function we want to maximize (or minimize) is in fact maximized (or minimized). (2) All the constraints are satisfied.

To illustrate how one creates a Lagrangian function, consider once again the problem faced by the Kloster Company. As indicated in the previous section, this firm wants to minimize $TC = 4Q_1^2 + 5Q_2^2 - Q_1Q_2$, subject to the constraint that $Q_1 + Q_2 = 30$. The first step in constructing the Lagrangian function for this firm's problem is to restate the constraint so that an expression is formed that is equal to zero:

$$30 - Q_1 - Q_2 = 0 \tag{2.25}$$

Then, if we multiply this form of the constraint by an unknown factor, designated λ (*lambda*), and add the result to the function we want to minimize (in equation (2.23)), we get the Lagrangian function, which is

$$L_{TC} = 4Q_1^2 + 5Q_2^2 - Q_1Q_2 + \lambda(30 - Q_1 - Q_2) \tag{2.26}$$

For reasons specified in the next paragraph, we can be sure that, if we find the unconstrained maximum (or minimum) of the Lagrangian function, the solution will be exactly the same as the solution of the original constrained maximization (or minimization) problem. In other words, to solve the constrained optimization problem, all we have to do is optimize the Lagrangian function. For example, in the case of the Kloster Company, we must find the values of Q_1, Q_2, and λ that minimize L_{TC} in equation (2.26). To do this, we must find the partial derivative of L_{TC} with respect to each of these three variables—Q_1, Q_2, and λ:

$$\frac{\partial L_{TC}}{\partial Q_1} = 8Q_1 - Q_2 - \lambda$$

$$\frac{\partial L_{TC}}{\partial Q_2} = -Q_1 + 10Q_2 - \lambda$$

$$\frac{\partial L_{TC}}{\partial \lambda} = -Q_1 - Q_2 + 30$$

[9]This section can be skipped without loss of continuity.

As indicated in the section before last, we must set all three of these partial derivatives equal to zero in order to minimize L_{TC}:

$$8Q_1 - Q_2 - \lambda = 0 \qquad (2.27)$$
$$-Q_1 + 10Q_2 - \lambda = 0 \qquad (2.28)$$
$$-Q_1 - Q_2 + 30 = 0 \qquad (2.29)$$

It is important to note that the partial derivative of the Lagrangian function with regard to λ (that is, $\partial L_{TC}/\partial\lambda$), when it is set equal to zero (in equation (2.29)), is the constraint in our original optimization problem (recall equation (2.25)). This, of course, is always true because of the way the Lagrangian function is constructed. So, if this derivative is zero, we can be sure that this original constraint is satisfied. And, if this constraint is satisfied, the last term on the right of the Lagrangian function is zero, so the Lagrangian function boils down to the original function that we wanted to maximize (or minimize). Consequently, by maximizing (or minimizing) the Lagrangian function, we solve the original constrained optimization problem.

Returning to the Kloster Company, equations (2.27), (2.28), and (2.29) are three simultaneous equations with three unknowns—Q_1, Q_2, and λ. If we solve this system of equations for Q_1 and Q_2, we get the optimal values of Q_1 and Q_2. Subtracting equation (2.28) from equation (2.27), we find that

$$9Q_1 - 11Q_2 = 0 \qquad (2.30)$$

Multiplying equation (2.29) by 9 and adding the result to equation (2.30), we can solve for Q_2:

$$\begin{array}{r} -9Q_1 - 9Q_2 + 270 = 0 \\ 9Q_1 - 11Q_2 = 0 \\ \hline -20Q_2 + 270 = 0 \\ Q_2 = 270/20 = 13.5 \end{array}$$

Therefore, the optimal value of Q_2 is 13.5. Substituting 13.5 for Q_2 in equation (2.29), we find that the optimal value of Q_1 is 16.5.

The answer we get is precisely the same as in the previous section: The optimal value of Q_1 is 16.5, and the optimal value of Q_2 is 13.5. In other words, the Kloster Company should produce 16.5 units of the first product and 13.5 units of the second product per hour. But, the method of Lagrangian multipliers described in this section is more powerful than that described in the previous section for at least two reasons: (1) It can handle more than a single constraint, and (2) the value of λ provides interesting and useful information to the decision maker.

Specifically λ, called the Lagrangian multiplier, measures the change in the variable to be maximized or minimized (TC in this case) if the constraint is relaxed by one unit. For example, if the Kloster Company wants to minimize total cost subject to the constraint that the total output of both products is 31 rather than 30, the value of λ indicates by how much the minimum value of TC will increase. What is the value of λ? According to equation (2.27),

$$8Q_1 - Q_2 - \lambda = 0$$

Since $Q_1 = 16.5$ and $Q_2 = 13.5$,

$$\lambda = 8(16.5) - 13.5 = 118.5$$

Consequently, if the constraint is relaxed so that total output is 31 rather than 30, the total cost will go up by $118.50.

For many managerial decisions, information of this sort is of great value. Suppose that a customer offers the Kloster Company $115 for one of its products, but to make this product, Kloster would have to stretch its total output to 31 per hour. On the basis of the findings of the previous paragraph, Kloster would be foolish to accept this offer, since this extra product would raise its costs by $118.50, which is $3.50 more than the amount the customer offers to pay for it.

Comparing Incremental Costs with Incremental Revenues

Before concluding this chapter, we must point out that many business decisions can and should be made by comparing incremental costs with incremental revenues. Typically, a manager must choose between two (or more) courses of action, and what is relevant is the difference in costs between the two courses of action, as well as the difference in revenues between them. For example, if the managers of a machinery company are considering whether to add a new product line, they should compare the incremental cost of adding the new product line (that is, the extra cost resulting from its addition) with the incremental revenue (that is, the extra revenue resulting from its addition). If the incremental revenue exceeds the incremental cost, the new product line will add to the firm's profits.

Note that *incremental* cost is not the same as *marginal* cost. Whereas marginal cost is the extra cost from a very small (one-unit) increase in output, **incremental cost** is the extra cost from an output increase that may be very substantial. Similarly, **incremental revenue**, unlike marginal revenue, is the

CONSULTANT'S CORNER

Planning to Meet Peak Engineering Requirements*

A leading computer manufacturer, after analyzing the history of its product development projects, found regular patterns of laborpower buildup and phaseout in its projects. Specifically, the number of engineers required to carry out such a project t months after the start of the project could be approximated reasonably well by

$$Y = at - bt^2, \qquad \text{for } 0 \leq t \leq a/b$$

where Y is the number of engineers required t months after the start of the project, and a and b are numbers that vary from project to project (and that depend on the kind of product being developed).

The computer manufacturer wanted to use these results to estimate when the number of engineers required to carry out a particular product development project would hit its peak and how great this peak requirement would be. Estimates of this sort would help the firm's managers plan the allocation and utilization of the firm's engineering staff and alert them to situations in which its staff might have to be expanded or supplemented. The project for which the firm's managers wanted these estimates was to begin immediately. On the basis of previous experience with projects of this type, the firm's managers estimated that a would be about 18 and b would be about 1 for this project.

If you were a consultant to this firm, how would you make the estimates wanted by the firm's managers?

Source: This section is based on an actual case, although the equations and the situation have been simplified somewhat for pedagogical purposes.

extra revenue from an output increase that may be very substantial. For example, suppose that you want to see whether a firm's profits will increase if it doubles its output. If the incremental cost of such an output increase is \$5 million and the incremental revenue is \$6 million, the firm will increase its profits by \$1 million if it doubles its output. Marginal cost and marginal revenue cannot tell you this, because they refer to only a very small increase in output, not to a doubling of it.

While it may seem very easy to compare incremental costs with incremental revenues, there in fact are many pitfalls. One of the most common errors is the failure to recognize the irrelevance of sunk costs. Costs incurred in the past often are irrelevant in making today's decisions. Suppose you are going to make a trip and you want to determine whether it will be cheaper to drive your car or to go by plane. What costs should be included if you drive your car? Since the only incremental costs incurred will be the gas and oil (and a certain amount of wear and tear on tires, engine, and so on), they are the only costs that should be included. Costs incurred in the past, such as the original price of the car, and costs that will be the same regardless of whether you make the trip by car or plane, such as your auto insurance, should not be included. On the other hand, if you are thinking about buying a car to make this and many other trips, these costs should be included.[10]

To illustrate the proper reasoning, consider the actual case of an airline that has deliberately run extra flights that do no more than return a little more than their out-of-pocket costs. Assume that this airline is faced with the decision of whether to run an extra flight between city *A* and city *B*. Assume that the fully allocated costs—the out-of-pocket costs plus a certain percent of overhead, depreciation, insurance, and other such costs—are $5,500 for the flight. Assume that the out-of-pocket costs—the actual sum that this airline has to disburse to run the flight—are $3,000 and the expected revenue from the flight is $4,100. In such a case, this airline will run the flight, which is the correct decision, since the flight will add $1,100 to profit. The incremental revenue from the flight is $4,100, and the incremental cost is $3,000. Overhead, depreciation, and insurance would be the same whether the flight is run or not. Therefore, fully allocated costs are misleading here; the relevant concept of costs is out-of-pocket, not fully allocated, costs.

Errors of other kinds can also mar firms' estimates of incremental costs. For example, a firm may refuse to produce and sell some items because it is already working near capacity and the incremental cost of producing them is judged to be very high. In fact, however, the incremental cost may not be so high because the firm may be able to produce these items during the slack season (when there is plenty of excess capacity), since the potential customers may be willing to accept delivery then.

Also, incremental revenue frequently is misjudged. Take the case of a firm that is considering the introduction of a new product. The firm's managers may estimate the incremental revenue from the new product without taking proper account of the effects of the new product's sales on the sales of the firm's existing products. Whereas they may think that the new product will not cut into the sales of existing products, it may in fact do so, with the result that their estimate of incremental revenue may be too high.

Summary

1. Functional relationships can be represented by tables, graphs, or equations. The marginal value of a dependent variable is defined as the change in this

[10]This example is worked out in more detail in the paper by E. Grant and W. Ireson in E. Mansfield, ed., *Managerial Economics and Operations Research: Techniques, Applications, and Cases,* 5th ed.: New York: W. W. Norton, 1983).

variable associated with a one-unit change in a particular independent variable. The dependent variable achieves a maximum when its marginal value shifts from positive to negative.

2. The derivative of Y with respect to X, denoted dY/dX, is the limit of the ratio $\Delta Y/\Delta X$ as ΔX approaches zero. Geometrically, it is the slope of the curve showing Y (on the vertical axis) as a function of X (on the horizontal axis). We have provided rules that enable us to find the value of this derivative.
3. To find the value of X that maximizes or minimizes Y, we determine the value of X where dY/dX equals zero. To tell whether this is a maximum or a minimum, we find the second derivative of Y with respect to X, denoted d^2Y/dX^2, which is the derivative of dY/dX. If this second derivative is negative, we have found a maximum; if it is positive, we have found a minimum.
4. A dependent variable often depends on a number of independent variables, not just one. To find the value of each of the independent variables that maximizes the dependent variable, we determine the partial derivative of Y with respect to each of the independent variables, denoted $\partial Y/\partial X$ and set it equal to zero. To obtain the partial derivative of Y with respect to X, we apply the ordinary rules for finding a derivative; however, all independent variables other than X are treated as constants.
5. Managers of firms and other organizations generally face constraints that limit the options available to them. In relatively simple cases in which there is only one constraint, we can use this constraint to express one of the decision variables as a function of the other decision variables, and we can apply the techniques for unconstrained optimization.
6. In more complex cases, constrained optimization problems can be solved by the method of Lagrangian multipliers. The Lagrangian function combines the function to be maximized or minimized and the constraints. To solve the constrained optimization problem, we optimize the Lagrangian function.
7. Many business decisions can and should be made by comparing incremental costs with incremental revenues. Typically, a manager must choose between two (or more) courses of action, and what is relevant is the difference between the costs of the two courses of action, as well as the difference between their revenues.

Problems

1. One very important question facing hospitals is this: How big must a hospital be (in terms of patient-days of care) to minimize the cost per patient-