Chapter 14

Optimization and Analysis of Models[1]

This chapter is designed to help you take you knowledge of building models to the next level - applying them to solve problems involving questions about optimization. In general, optimization is the process of trying to make something as efficient as possible, or as large as possible or as cheap as possible. It's the study of minimizing or maximizing a quantity, like profit, as a function of some other quantity, like production. In order to optimize a quantity, though, we need a few things. The first is a skill you already have - the ability to create a model equation that represents how the quantity to be optimized varies as a function of some other quantity. For example, we might produce a model equation describing how the profits of a company depend on the number of items they produce, since the more you produce: (a) the more you can sell, generating more revenue but (b) the more it costs, in labor and materials. The other tool that you need is a knowledge of marginal analysis, which measures how a change in the independent variable will cause a change in the dependent variable in a model. We will focus our study on the marginal analysis and optimization of polynomial models, although this is only the tip of the iceberg.

- *As a result of this chapter, students will learn*
  - √ What marginal analysis is
  - √ How to interpret the results of marginal analysis
  - √ What the derivative of a power function is
  - √ What the derivative of a polynomial function is
- *As a result of this chapter, students will be able to*
  - √ Compute the derivative of a power function
  - √ Compute the derivative of a polynomial
  - √ Maximize or minimize a polynomial, using both algebra and Excel

14.1 Calculus with Powers and Polynomials

We have spent some time discussing the basic families of functions. These functions can be used to model the behavior of various real-world business situations. For example, suppose we have data based on the total cost of paying back a loan (for a fixed principal and fixed payback period). We can use this data to develop a function, call it $C(r)$, which represents this cost as a function of different interest rates on the loan. Suppose interest rates are increasing. How will this affect the cost of paying back the loan?

This question really centers on how the function C changes as the interest rate r increases. To answer this question, we will turn to our knowledge of families of functions. In particular, we will use what we know about the parameter A in the general formula for a linear function, $y = A + Bx$.

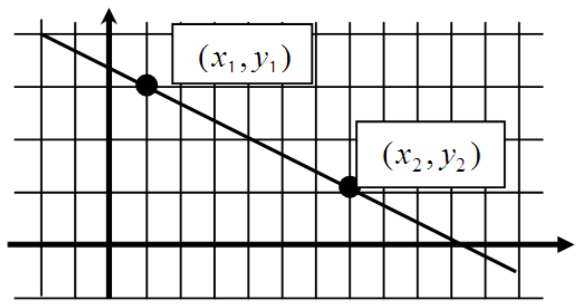


Figure 14.1: Slope between two points.

Look at the graph of the linear function shown in figure 14.1. Also shown on the graph are two points. These points are labeled with the coordinates (x_1, y_1) and (x_2, y_2). What is the total change in the linear function between the two points?

Between these points, there is a change of $y_2 - y_1$. This is just the vertical separation between the two points. Now, how quickly is the function changing at the first point? This is not a question of total change, but of the rate of change of the function. Another way of asking this question is "If I make a small change in x from x_1 to x_2, how much will the function change?" To answer this question, we look at the slope of the line. As you may recall, the slope of a line can be calculated from the formula

$$\text{slope} = A = \frac{y_2 - y_1}{x_2 - x_1}.$$

For the function above, we see that the two points have coordinates $(1, 3)$ and $(7, 1)$. Thus, the slope of the line is $(1 - 3)/(7 - 1) = -2/6 = -1/3$. The negative tells us that the function (in this case a straight line) is decreasing. This means that, as we move from left to right, the value y of the function gets smaller. There are several nice things about straight lines that we can see from this example. First, unlike nonlinear functions, the slope of a straight line is exactly the same at every single value of x. This means that the slope of the function at the first point is $-1/3$ and slope at the second point is also $-1/3$ and the slope at $x = 249$ is also $-1/3$. Second, it is easy to calculate the slope of a straight line. We simply look at the change in the values of the function (the y values) and divide this by the change in the x values between the two points. This will not hold for any other family of functions.

To find the slope of a nonlinear function, we take advantage of a property of smooth functions. As illustrated in the graphs in figure 14.2, if we have the graph of a nonlinear function, and we zoom in on the graph, it begins to look linear.

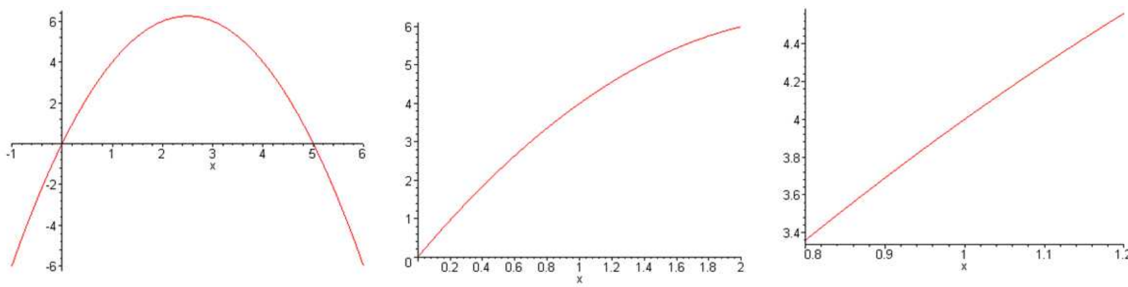


Figure 14.2: Series of graphs showing how the function changes as we zoom in on $x = 1$.

For some functions, we need to zoom in more, and for others we zoom in less to see this linear-like appearance. In order to calculate the slope, we will use this feature, called local linearity, to determine the slope of a functions at any point. Specifically, if we pick two points on the function, and draw a line between them, we will call the slope of this line the average rate of change of the function. If we call these two points $(x_1, f(x_1))$ and $(x_2, f(x_2))$, then the average rate of change between the points is

$$\text{average rate of change} = \frac{f(x_2) - f(x_1)}{x_2 - x_1}.$$

Notice that the graph in figure 14.3 shows how the average rate of change can be quite different from the actual rate of change (called the instantaneous rate of change or derivative).

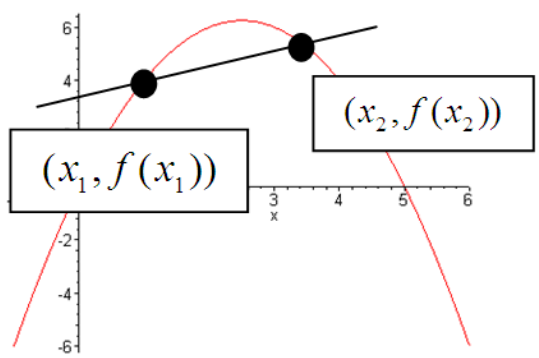


Figure 14.3: Average slope between two points.

However, if we move the second point closer to the first, we can get a more accurate approximation to the instantaneous rate of change of the function near the first point. If the two points are close enough, the average rate of change will be a very good approximation to the instantaneous rate of change. This fact will help us in many cases where we only have data, instead of an actual function.

14.1.1 Definitions and Formulas

Quotient A quotient is simply the result of dividing one quantity by another quantity.

Average slope The average slope between two points on a function is what you get when you start with a function (f), evaluate it at two points (say x_1 and x_2) and then take the difference of these values, $f(x_2) - f(x_1)$ and divide it by the distance between the two x-values $(x_2 - x_1)$. Thus,

$$\text{average slope} = \frac{f(x_2) - f(x_1)}{x_2 - x_1}.$$

Note that the order is important! If you start with x_2 first in the numerator, you must also start with x_2 in the denominator. The graph below shows the basic idea and illustrates why it's called average slope and not the actual slope. The dashed line between the two points represents the average slope of the function (the curved line) between those two points. In between the two points, though, notice that there are places where the curve has a more negative slope than the average slope and places where the slope is even positive!

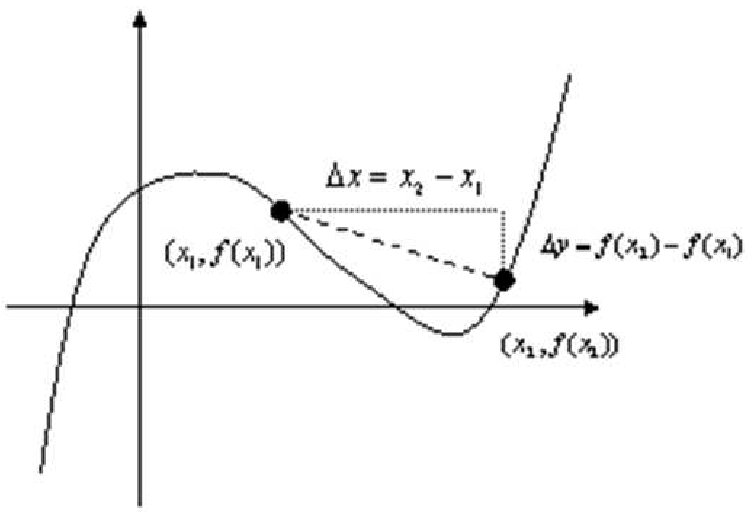


Figure 14.4: Average slope between two points.

Difference quotient The difference quotient is another way of writing the average slope. Instead of looking at the average slope between x_1 and x_2, we look at the average slope between x_1 and $x_1 + h$, where we think of h as a small number. So $x_1 + h$ is another way of writing x_2. This form of x_2 allows us to focus on how the function changes at

x_1. Using $x_1 + h$ in place of x_2 changes the denominator of the average slope formula. Instead of $x_2 - x_1$, we have $(x_1 + h) - x1 = h$. So, the average slope formula takes on a new name and a new look:

$$\text{Difference quotient} = \frac{f(x_1 + h) - f(x_1)}{h}$$

Consider the line passing through the point $(x_1, f(x_1))$ and having the same slope as the difference quotient, with a fixed value of h, say 1. If we look at this line for smaller and smaller values of h (say 0.1, 0.01, 0.001, etc.) we see that the line eventually becomes "parallel" with the function at the point $(x_1, f(x_1))$. This visual process of watching the line become parallel can be carried out mathematically through a limit.

Marginal Analysis This is a financial/business term for the process of finding the instantaneous rate of change of a function at a point. Essentially, this is a difference quotient, and it is useful for answering the question "If my independent variable increases by 1 unit, how much will my dependent variable increase (or decrease)?" Another way to think of this is:

How much bang do I get for each additional buck that I spend?

Marginal Cost Basically, when the word "marginal" is followed by a term like "cost", it means that you are looking at the instantaneous rate of change of the cost function, which is just its derivative.

Marginal Profit Instantaneous rate of change of the profit function.

Marginal Revenue Instantaneous rate of change of the revenue function.

Derivative function The derivative function is a function derived from the slopes of another function. Basically, at each point $(x, f(x))$ the function has a slope, usually denoted by $f'(x)$. If we collect all these slopes into a new function, so that plugging in a value of the independent variable, x, results in the slope of f at that point, then we have the derivative function. The derivative of a function at a point is also denoted by the notation $\frac{\partial f}{\partial x}$ which indicates that we are interested in the slope of f in the x-direction. Thus, a positive number for the derivative means that as x increases (we always move to the right) the value of f is increasing. Likewise, a negative value of the derivative indicates that the function is decreasing at that point. Officially, the derivative of a function at a point is computed by taking the difference quotient and letting h go to zero. This is noted mathematically by the "limit of the difference quotient":

$$f'(x) = \lim_{h \to 0} \left[\frac{f(x + h) - f(x)}{h} \right]$$

Second derivative Since the first derivative of a function is (usually) a function itself, we can take its derivative. We refer to the derivative of the derivative of a function as the second derivative. It is denoted by f'' or $\frac{\partial^2 f}{\partial x^2}$. Since the derivative tells how fast the function is changing, the second derivative tells us how fast the first derivative is changing. Thus, it measures the rate of change of the slope, which is called concavity. In a graph, concavity is easy to see: it refers to the direction and steepness of the way the function bends. If it bends up (looks like a cup) then the concavity is positive. If it bends down (looks like a frown) then the concavity is negative. If the function is almost flat, then the concavity is close to zero.

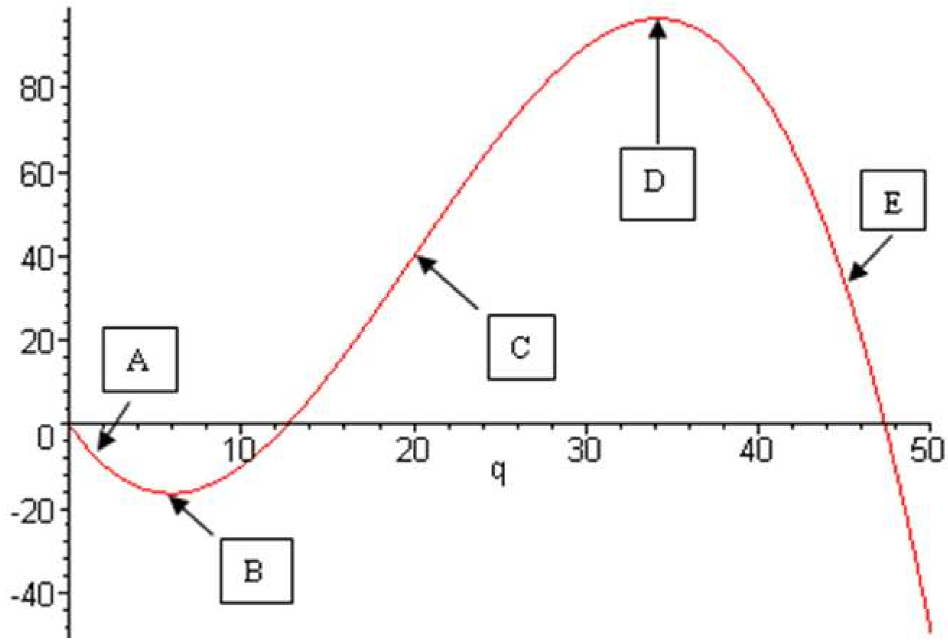


Figure 14.5: Graph and explanation showing the connections between f, f', and f''.

In this graph, there are five points marked A - E. The function and its derivatives are described at each of these points below.

A. Here the function is negative, the slope is negative (it is a decreasing graph) and the second derivative (concavity) is zero, since the graph is basically flat. Thus, $f(A) < 0$, $f'(A) < 0$, $f''(A) = 0$.

B. Here we have the function negative (it is below the x-axis, the line $y = 0$). The slope is zero, since the graph is horizontal at this point. The concavity is positive since the graph is curving upward. Thus, $f(B) < 0$, $f'(B) = 0$, $f''(A) > 0$.

C. Here the function is positive, the slope is positive (it is an increasing graph) and the second derivative (concavity) is zero, since the graph is basically flat. Thus, $f(C) > 0$, $f'(C) > 0$, $f''(C) = 0$.

D. Here we have the function positive (above the x-axis, $y = 0$), the slope is zero, since the graph is horizontal at this point, and the concavity is negative since the graph is curving downward. Thus, $f(D) > 0$, $f'(D) = 0$, $f''(D) < 0$.

E. Here the function is positive, the slope is negative (it is a decreasing graph) and the second derivative (concavity) is zero, since the graph is basically flat. Thus, $f(E) > 0$, $f'(E) < 0$, $f''(E) = 0$. In addition, since the function much steeper at point E than at point A, we know that the slope at E is more negative. Thus, we can also say that $f'(E) < f'(A)$.

14.1.2 Worked Examples

Example 14.1. Marginal Analysis with Difference Quotients
In example 1 (page 318) we developed a model for the cost of electricity as a function of the number of units of electricity produced. Later in that chapter we used parameter analysis to explore how the function behaved. This analysis was all in terms of percent changes, which is somewhat limiting. In this example, we are going to use marginal analysis through the difference quotient to interpret how much each unit of electricity affects the total cost of producing the electricity. (In later examples, we will refine this process using a shortcut method called the derivative.) The cost model we will use is the square root model given by

$$\text{Cost} = 6{,}772.56 + 1{,}448.74 \cdot \text{Sqrt(Units)}.$$

Suppose that we are currently producing 500 units of electricity. How much would it cost to produce one more unit of electricity? We can put this into an Excel spreadsheet to compute it fairly easily. The results are shown below, and were obtained by setting up a formula for the difference quotient of the function, with a variable for h so that we can let h get very small. This lets us see what the instantaneous rate of change of the cost function is.

A	6772.56				
B	1448.74				
X	500				
H	X+H	F(X)	F(X+H)	DF=F(X+H)-F(X)	DF/H
10	510	39167.37	39489.72	322.3443693	32.23444
1	501	39167.37	39199.75	32.37862999	32.37863
0.1	500.1	39167.37	39170.61	3.239319164	32.39319
0.01	500.01	39167.37	39167.7	0.323946492	32.39465
0.001	500.001	39167.37	39167.4	0.032394795	32.3948

From this, it seems that when current production is at 500 units, each additional unit of electricity will cost approximately \$32.39. In contrast, if are currently producing 1,000 units of electricity, the marginal cost is about \$22.91 per unit.

A	6772.56				
B	1448.74				
X	1000				
H	X+H	F(X)	F(X+H)	DF=F(X+H)-F(X)	DF/H
10	1010	52585.74	52814.24	228.4960877	22.84961
1	1001	52585.74	52608.64	22.9008669	22.90087
0.1	1000.1	52585.74	52588.03	2.290601805	22.90602
0.01	1000.01	52585.74	52585.97	0.229065334	22.90653
0.001	1000.001	52585.74	52585.76	0.022906585	22.90658

Example 14.2. Finding the Derivative of a Power Function
While it is possible to use basic algebra and the definition of the derivative (as a limit of the difference quotient) for marginal analysis this process can be tedious and will be difficult for some of the basic functions. Instead, we're going to use Excel to experiment to find a shortcut for the derivative of a power function. We begin with the power function $f(x) = x^2$. Here's an outline of what we'll do:

1. Set up a spread sheet that has places to enter the parameters of the function (A and B).

2. We also add in a place for us to enter a value for h, the number we need in the difference quotient.

3. We create columns for x and $f(x)$. Excel will compute the values for the column under $f(x)$ from the values listed under the x column.

4. Next we add columns to compute $x + h$ and $f(x + h)$.

5. We add a column to compute the difference quotient from the data we've already set up.

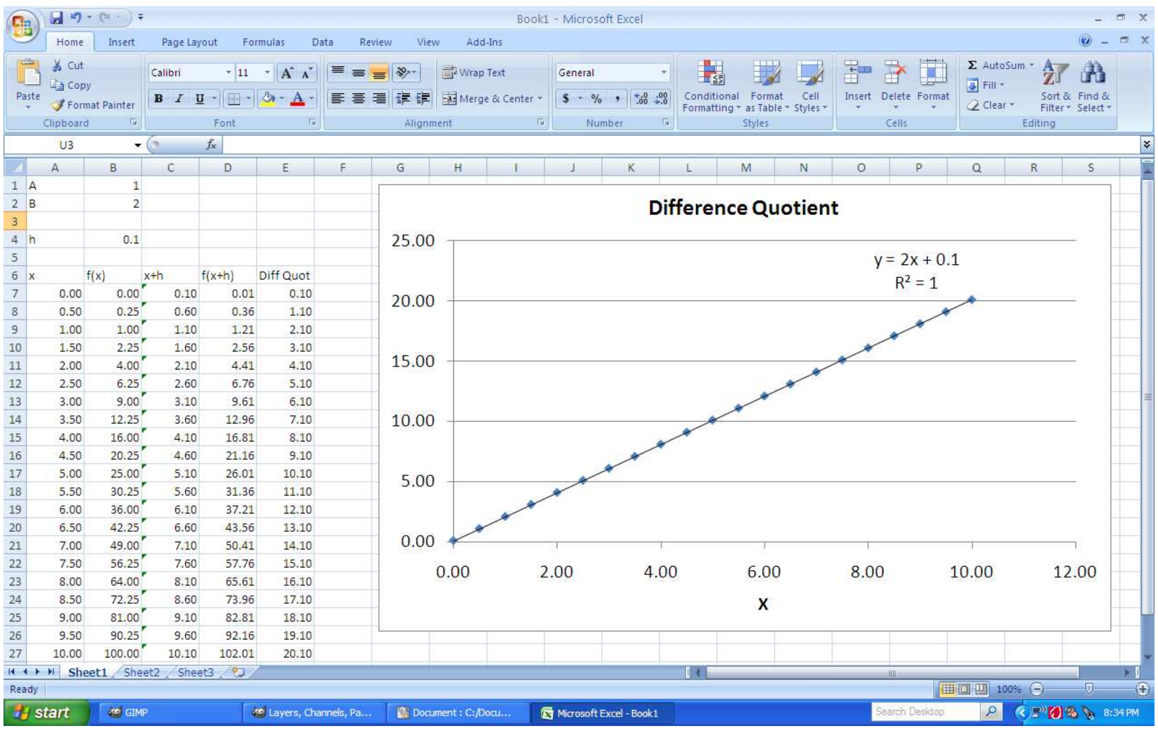
6. Since we have lots of x values (running down the table,) we now have a bunch of points of the form (x, difference quotient of f at x). If we make a scatterplot of these points, we can fit a trendline to these data and determine the equation of the difference quotient in the process. Thus, we are close to experimentally determining an equation for the derivative function.

7. Up till now we have kept h fixed. We can then simulate the limit of the difference quotient by making h a smaller and smaller number, until we think we see what the "real" equation would be with h equal to zero. (Note that we can't actually set h to be zero, since we would be dividing by zero, which gives an error!)

The screen shots below will show you what our spread sheet looks like at the end of this process. To go through this procedure, open the file "C14 SquareDerivative.xls". Starting with the power function

$$f(x) = Ax^B = x^2.$$

and setting h initially to 0.1, and listing x values from 0 to 10 in steps of 0.5, we get the table of data (using steps 1-6 above) shown in figure 14.6.



A	1			
B	2			
h	0.1			
x	f(x)	x+h	f(x+h)	Diff Quot
0.00	0.00	0.10	0.01	0.10
0.50	0.25	0.60	0.36	1.10
1.00	1.00	1.10	1.21	2.10
1.50	2.25	1.60	2.56	3.10
2.00	4.00	2.10	4.41	4.10
2.50	6.25	2.60	6.76	5.10
3.00	9.00	3.10	9.61	6.10
3.50	12.25	3.60	12.96	7.10
4.00	16.00	4.10	16.81	8.10
4.50	20.25	4.60	21.16	9.10
5.00	25.00	5.10	26.01	10.10
5.50	30.25	5.60	31.36	11.10
6.00	36.00	6.10	37.21	12.10
6.50	42.25	6.60	43.56	13.10
7.00	49.00	7.10	50.41	14.10
7.50	56.25	7.60	57.76	15.10
8.00	64.00	8.10	65.61	16.10
8.50	72.25	8.60	73.96	17.10
9.00	81.00	9.10	82.81	18.10
9.50	90.25	9.60	92.16	19.10
10.00	100.00	10.10	102.01	20.10

Figure 14.6: Difference quotient worksheet.

Now, what kind of trendline does the difference quotient make? It looks like a straight line, so let's add a linear trendline to the graph. Excel produces a fairly accurate equation:

$$y = 2x + 0.1 \text{ with } R^2 = 1.$$

But, we have a (mathematically speaking) pretty large value of h. Let's vary h and collect the results of the trendline into a table like the one below. Notice that as h gets smaller, the y-intercept of the trendline decreases. Since the derivative is the limit as h goes to zero of this difference quotient, we can reasonably conjecture that as h shrinks down to zero, so does the y-intercept, leading us to the following simple rule:

The derivative of the function $y = x^2$ is the function $y = 2x$.

h	y	R^2
0.1	$y = 2x + 0.1$	1
0.01	$y = 2x + 0.01$	1
0.001	$y = 2x + 0.001$	1
0.0001	$y = 2x + 0.0001$	1
0.00001	$y = 2x + 0.00001$	1
0.000001	$y = 2x + 0.000001$	1
0.0000001	$y = 2x + 0.0000001$	1

However, this only gives us the derivative of one single power function. What about all the other ones? How can we determine their derivatives without going through this fairly lengthy process every time? We've actually almost got the answer, since our spreadsheet is set up to allow us to change the parameters in the power function and find rules for those as well. This is what the exploration in this section is all about - finding the rules for ALL of the power functions. It turns out to be relatively simple.

Example 14.3. Marginal Analysis with Derivatives
Suppose we know that our costs for producing q thousand goods are $C(q) = q^2$, where C is measured in millions of dollars. If we are currently producing 10,000 goods, how will our costs increase if we add an additional 1,000 goods to the production?

For this situation, we are currently producing $q = 10$ thousand items and want to know what happens to the cost if we produce $q = 11$ thousand items. This is an increase of 1 (in our units of q) so it is a question about the marginal cost. Since the marginal cost is really just the derivative (slope) of the cost function, we can use the last example to help us out. In that example, we used Excel, difference quotients, and regression to learn that the derivative of x^2 is $2x$. Thus, the derivative of the cost function, denoted C', is $C'(q) = 2q$ and the marginal cost of producing 10,000 goods is $C'(10) = 2(10) = 20$. The units of the derivative are (units of function/units of independent variable) so the complete answer is:

> If the cost of producing q goods is $C(q) = q^2$, where C is measured in millions of dollars and q is measured in thousands of goods, then the marginal cost of producing 10,000 goods is 20 million dollars per thousand goods.

This means that if we want to increase production to 11,000 goods, we can expect an increase in the costs of about 20 million dollars. If we wanted to produce 12,000 goods the cost would increase by approximately (\$20 million per thousand goods) $\cdot$ (2 thousand goods) = 40 million dollars (since 12,000 is 2,000, or 2 thousands, greater than 10,000.) If, on the other hand, we decrease the production to 9,500 goods, then the cost will change by about (\$20 million per thousand goods) $\cdot$ (-0.5 thousand goods) = -10 million dollars. (The negative sign simply means that the cost decreases if we decrease production.)

14.1.3 Exploration 14A: Finding the Derivative of a General Power Function

Using the file "C14 PowerDerivative.xls" try to determine the shortcut derivative rules for general power functions. Phrase each of your rules as a sentence like the one in italics in example 2 (page 422). The tables below may help you to organize your results in order to make sense of them. For each, you should probably use h = 1E-6 or smaller.

A. Start with changing A and see what that does. Complete the following table to help you record your observations and make conjectures about the general form of the derivative of the function $f(x) = Ax^2$.

A	B	$F(x)$	$F'(x)$

Your sentence describing the shortcut rule:

B. Now set $A = 1$ and see if you can find the derivative rule for $f(x) = x^B$. Start with integer powers of B to find the pattern, then test your pattern for non-integer values of B. You may need to delete the row containing $x = 0.0$ from the data table in order to use the appropriate trendline.

A	B	$F(x)$	$F'(x)$

Your sentence describing the shortcut rule:

C. Finally, try to combine your rules above to find the general shortcut rule for the derivative of the function $f(x) = Ax^B$.

A	B	$F(x)$	$F'(x)$

Your sentence describing the shortcut rule:

D. For the ultimate challenge, try to find out what the derivative rule for polynomials is. Start with a simple one, like $f(x) = x^3 + x^2 + 1$ and see if you can figure out what happens. (Hint: Polynomials are really just sums of power functions with non-negative integer powers.)

14.1.4 How To Guide

Typesetting and formatting equations in MS Word

Unknown to most people using Microsoft Word, there is a way to type set complicated mathematical formulas. To activate the Equation Editor in Word, first select the point where you want to insert the equation. Then activate the "Insert" ribbon, and choose "Equation". You can then choose to insert a pre-set equation from the drop-down list that appears, or you can insert a new equation from scratch by selecting "insert new equation" at the bottom of the list.

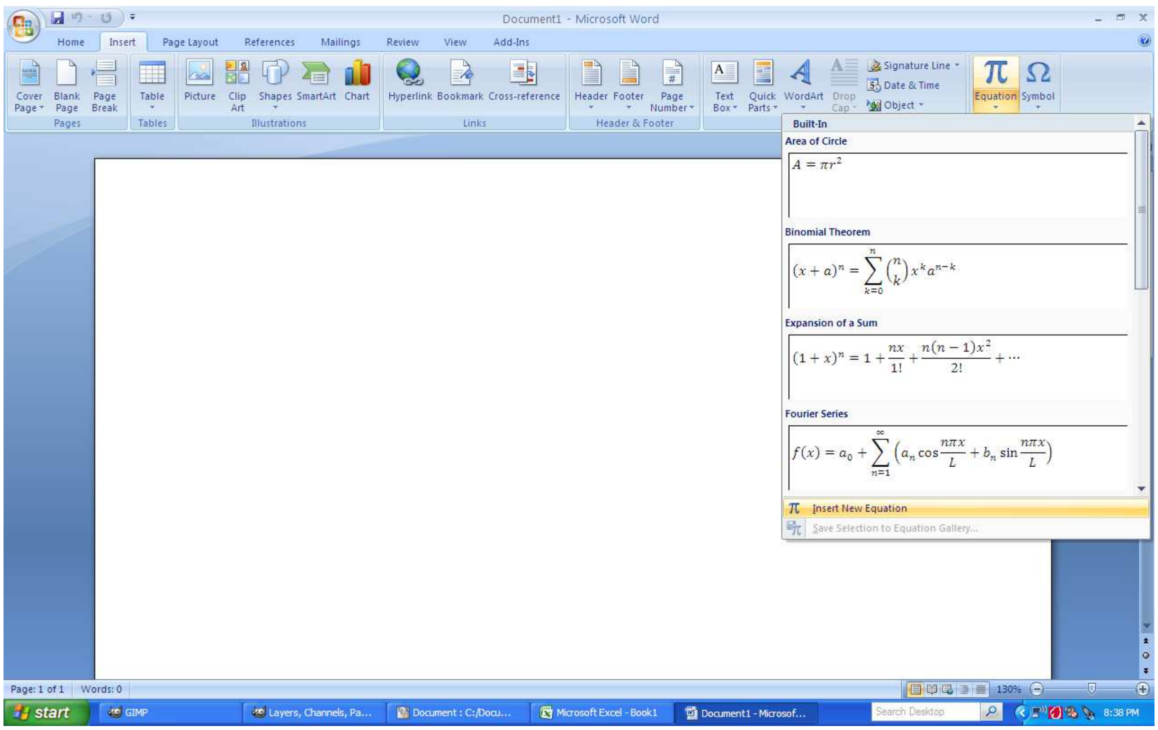


Figure 14.7: Inserting a new equation from the insert ribbon.

The screen will appear almost normal, except for two things - you will now see an area marked "type equation here" and the equation ribbon is activated (see figure 14.8). You can start typing the equation with the keyboard and then use the buttons on the ribbon to access more symbols and tools. Some of the more common symbols are are directly accessible. The remaining symbols are grouped into categories such as fraction, script, radical, integral, and bracket. For example, to enter a formula to show x^2, we first set up a placeholder for the symbols using the script category. From the script menu, select the button with the large empty box in the center and a small empty box in the upper right corner - the one that looks like a superscript. You can now enter the "x" into the large box (the base) and the exponent 2 into the small box.

To make a more complicated formula, say $x^2 + \frac{1}{x}$, start as before, then, after you type the superscript of 2, hit the right arrow button on the keyboard - this will move the cursor

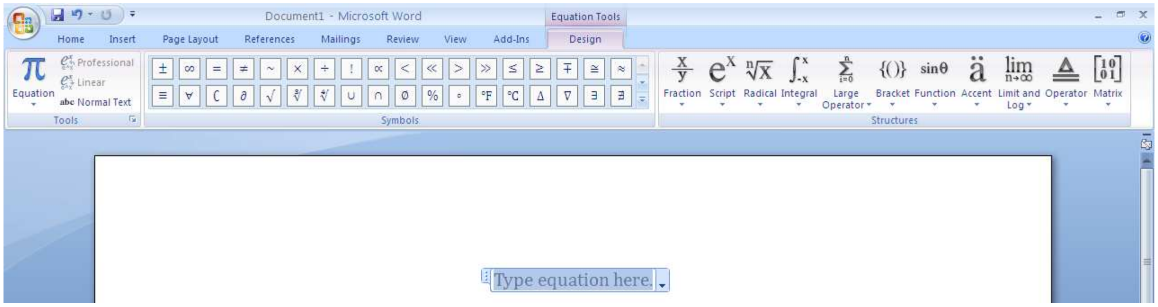


Figure 14.8: The equation ribbon.

out of the exponent so that you can continue typing in the main part of the formula. Type a "+" then add a fraction with the tools in the second group on the bottom row. The first fraction (with two empty boxes, one above and one below the line) is best. Click on each box and type in the symbols you need (a "1" on top and an "x" on bottom). When you are done editing the formula, click the mouse outside the boxed in region of the equation. Once inserted, you can copy, paste, and modify the equation. (To modify it, just double-click on it.)

14.2 Extreme Calculus!

Now that we've learned a little about marginal analysis, we can apply this knowledge to help answer questions that are really important. For example, suppose we would like to minimize the cost of producing our product, working on the theory that this will save us money. How would we go about this process of optimizing the cost?

First of all, we need to know what causes the cost of production to vary. Typically, the simplest quantity that determines production cost is, you guessed it, the number of items that we produce. After all, each one of them uses a certain amount of materials that aren't free; each one of them requires labor; production probably involves machines which use electricity and so forth. So, we could start by getting together data that shows the total cost each month (or week or whatever) along with the total cost of production that month (or week or whatever). We can then use our model-building skills to determine an equation that represents the cost of production as a function of the number of items produced.

Now, how can this help us find the amount of production that will result in the lowest overall cost? We actually have several tools available. We could create a table of values from the function and look for the lowest cost. That could be difficult, though, since our table will only show some of the possible values: it may be that we skip over the best spot if we're not very careful. We could also graph the function, but then scale is an issue; we may have to keep redrawing the graph on larger and larger scales to see where this minimum occurs. The most commonly used approach, though, is based on marginal analysis.

Think about it this way. We could imagine "walking" along the function in the direction of increasing production. As we do this, the slope along which we climb is determined by the rate of change of the function - marginal analysis. If the marginal cost is negative, we are going downhill; this means that by increasing the production we can decrease the costs a little. If the slope is very large and negative, then we are far from the minimum cost. As we get closer to the minimum of the cost, this slope will level out. In fact, if we go too far, we could wind up increasing the costs - like climbing out of a hole. That means that we need to go back in order to decrease the cost.

This idea of walking along the function is a little hard to implement on a computer. It's much easier to think about what the function must look like near the minimum cost. We know that on one side of the minimum, the slope is negative, because we are decreasing the cost as we increase production. On the other side of the minimum (we've gone too far!) the slope is positive. Now the slope is the marginal cost. This is a number associated with each value of production. If it is negative on one side of the minimum, and positive on the other side of the minimum, then we can conclude (assuming a mathematical property called continuity) that at the minimum, the slope (marginal cost) is exactly zero. This basic idea can be used to solve any optimization problem - simply set the marginal whatever to zero and solve the resulting equation.

14.2.1 Definitions and Formulas

Critical point Any point on the graph of function f where the derivative is zero is a critical point. Thus, we can find all the critical points by solving the equation $f'(x) = 0$. Often, this will be a nonlinear equation and will require some algebra to solve.

Extrema An extrema is some "extreme" point on a function: either a maximum or a minimum.

Local Maximum A local maximum is a point on the graph of a function that is higher than all the points that are close by it. Thus, the point looks like the top of a hill. Point D in the graph at the end of the definitions fromthe last section is an example of a local maximum. See the graph below.

Local Minimum A local minimum is a point on the graph of a function that is lower than all the points that are close by it. Thus, the point looks like the bottom of a valley. Point B in the graph at the end of the definitions from the last section is an example of a local minimum. See the graph below.

Global Maximum A global maximum is the highest point on a function anywhere not just when compared to points near it. Most functions have lots of hills and valleys; only the highest peak in the "mountain range of the function" would be the global maximum. See the graph below.

Global Minimum A global minimum is the lowest point on a function anywhere not just when compared to points near it. Most functions have lots of hills and valleys; only the lowest valley in the "mountain range of the function" would be the global maximum. See the graph below.

Optimization This is the process of finding and classifying all the extrema for a function and then using this to solve some problem. For example, we may have a function that describes our profits from manufacturing a quantity q of a good. Optimization would help us answer the question "How many of this good should we make in order to get the highest profit?"

Second Derivative Test Solving the equation $f'(x) = 0$ only finds extreme points. You then need to classify the points as maxima or minima (the plurals of maximum and minimum, respectively). One way to do this is by graphing the function. The other way is by evaluating the second derivative of the function at the critical point. If the second derivative is negative, you have a maximum (the graph is concave down, as at Point D in figure 14.5, pge 420.) If the second derivative is positive, you have a minimum (the graph is concave down, as at Point B in 14.5.) If the second derivative is zero, then you don't have a maximum or a minimum, necessarily.

14.2.2 Worked Examples

Example 14.4. Using optimization to sketch polynomials
This example assumes that you have learned (from the last section) the following derivative rule:

> The Sum Rule for Derivatives: The derivative of the sum of two functions is the sum of the derivatives of the two functions. In other words, $(f + g)' = f' + g'$.

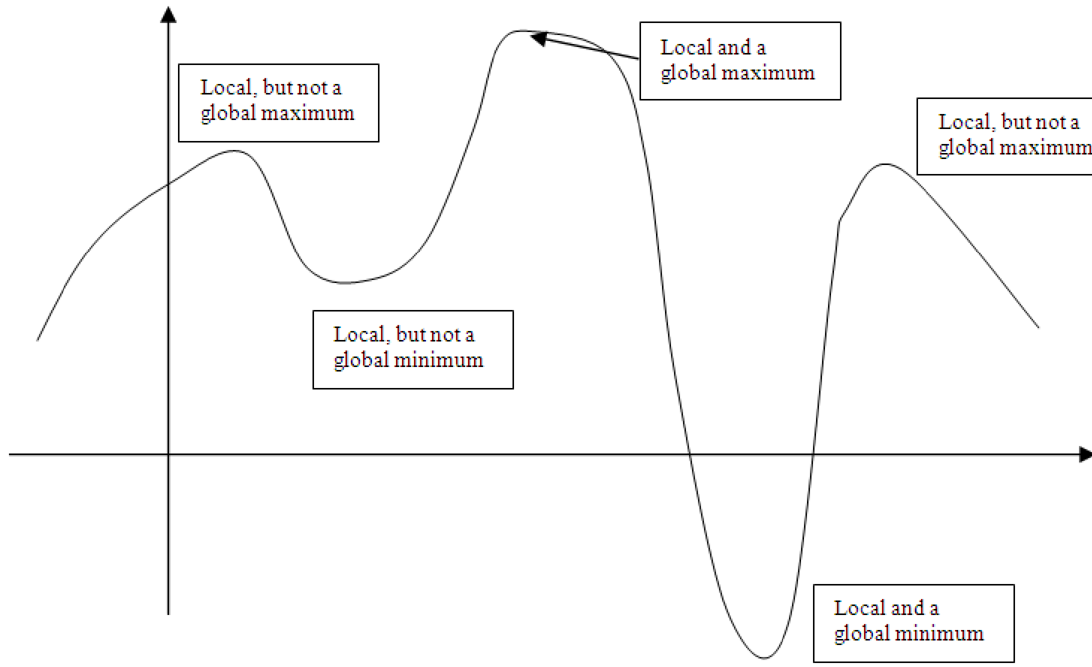


Figure 14.9: Example of a function with several local maxima and minima.

Since a polynomial is just a sum of power functions, we can use this rule to determine the derivative of a polynomial: It's just the sum of the derivatives of the individual power functions that make up the polynomial. Thus, the derivative of the polynomial $f(x) = 3x^4 - 5x^3 + 2x - 7$ is just $f'(x) = 12x^3 - 15x^2 + 2$. (The derivative of $3x^4$ is $12x^3$. The derivative of $-5x^3$ is $-15x^2$. The derivative of $2x$ is $2(1)x^{1-1} = 2x^0 = 2$, and the derivative of a constant is zero.) We can use this to learn about the properties of polynomials and what they look like.

For example, suppose we have a general fifth degree polynomial. Thus, the function can be written (generally) as $g(x) = a_5x^5 + a_4x^4 + a_3x^3 + a_2x^2 + a_1x + a_0$, where the a's represent constants. What would the derivative of this function be? Well, we apply the power rule to each term and get: $g'(x) = 5a_5x^4 + 4a_4x^3 + 3a_3x^2 + 2a_2x + a_1$. This is a fourth degree polynomial, as expected. How can this help us visualize the graph of g?

For starters, notice that if we try to find all the critical points of g we will have to solve the equation $g'(x) = 0$. This is a fourth degree polynomial equation and can have, at most, four solutions. Thus, there are at most four critical points in the graph of g. If we were to locate these critical points, we could begin to sketch the graph. Let's take the specific polynomial $h(x) = 6x^5 + 15x^4 - 130x^3 - 210x^2 + 720x + 300$. It's derivative is $h'(x) = 30x^4 + 60x^3 - 390x^2 - 420x + 720$. To find the critical points, we set this derivative equal to zero and solve the equation. Since this equation can be factored as

$$0 = 30(x^4 + 2x^3 - 13x^2 - 14x + 24) = 30(x - 1)(x + 2)(x - 3)(x + 4)$$

we see that the derivative is zero at the points where $x = 1, -2, 3, -4$. There are four critical points. By plugging them into the function, we can find the y-coordinate of these points, and then graph them. Finally, we notice that since the leading term is a fifth degree power function with a positive coefficient, the function is increasing to the right. Since it is an odd-degree polynomial, it must do the opposite on the left, so it decreases to the left. In the end, we can sketch the graph quite accurately.

Example 14.5. Maximizing Profits with Derivatives
Suppose that the cost of producing q goods is

$$C(q) = 0.01q^3 - 0.6q^2 + 13q$$

and we sell these goods for \$7 apiece. How many of our product should we make (and sell) in order to maximize our profit?

The revenue function will be $R(q) = 7q$. This comes from the fact that revenue is simply the number of products sold times the selling price per product. The profit function (remember: profit = revenue - cost) will then be

$$P(q) = 7q - 0.01q^3 + 0.6q^2 - 13q.$$

The marginal profit is given by the derivative of the profit (which can be computed using the rules we have developed so far). We find that

$$P'(q) = 7 - 0.03q^2 + 1.2q - 13 = -0.03q^2 + 1.2q - 6.$$

We set this function equal to zero and do some algebra (anmely, the quadratic formula) to find that when $q = 5.86$ and $q = 34.14$ the profit function has a critical point. We can also get these results by entering the following data in EXCEL.

	A	B
1	q	1
2	P'(q)	=-0.03*B1ˆ2+1.2*B1-6

We then use the goal seek procedure with the following information. "Set Value" to B2, "To value" 0, and "By changing" B1. Note that this will only locate the first extreme point, $q = 5.86$. To be sure you do not miss the other points, it is good to first graph the function and visually locate some values that are close the extreme points. Then enter one of these values in cell B1. From the graph of $P(q)$ we find that there is an extreme point near $q = 30$. If we enter 30 in cell B1 and then repeat the goal seek procedure described above (with B2, 0, and B1 in the "Goal Seek" dialog box) the computer will locate the other extreme point.

Now, which of these two points is a maximum and which is a minimum? To answer this, we'll apply the second derivative test. This is simple; we just find the second derivative of the profit function and evaluate its sign at each critical point. The second derivative of the profit function is just the derivative of the first derivative. So, we find

$$P''(q) = -0.06q + 1.2.$$

We then compute easily that $P''(5.86) = 0.8484$, which is positive. This indicates that the point $(q = 5.86, P = -16.57)$ is a local minimum - not a good place to be! At the other point we find that $P''(34.14) = -0.8484$, so the point $(q = 34.14, P = 96.57)$ is a local maximum for the profit function. That's where we want our production and sales!

This tells us that if we sell our product at \$7 each and incur a cost given by the function above then we can achieve a maximum profit of \$96.57 dollars by making and selling 34.14 units of our product.

Example 14.6. Minimizing Average Cost
Suppose that we have a fixed cost of \$2000 each month. This cost includes electricity, rent, and equipment. In addition, if it costs us \$12 per good manufactured (including materials and labor), we have a total monthly cost of

$$C(q) = 12q + 2000.$$

Suppose that instead of minimizing the total cost, we now we want to minimize the average cost function. The average cost function, $\bar{C}(q)$, is basically the cost function divided by the quantity produced (i.e., average cost = total cost of making q goods divided by q.) Thus, the average cost function for this scenario is

$$\bar{C}(q) = 12 + \frac{2000}{q}.$$

This is not a polynomial (the $1/q$ term is really $q^(- 1)$, which not a positive integer power) but we can use the sum rule and product rule to get its derivative:

$$\bar{C}'(q) = 0 + 2000\frac{d}{dq}\left(q^{-1}\right) = 2000(-1)q^{-2} = \frac{-2000}{q^2}.$$

Now, this function is not like our other examples: there is no minimum! We cannot solve the equation $\bar{C}'(q) = 0$ because no value of q will solve this. However, we notice that as q increases, the derivative of the average cost decreases (the derivative is always negative.) This means that making more of our product will always reduce the average cost.

If, instead, we had a slightly more realistic cost function (taking secondary costs into effect) like

$$C(q) = 0.05q^2 + 12q + 2000,$$

then we can optimize the function. Following the same steps as before, we get the average cost function as

$$\bar{C}(q) = 0.05q + 12 + \frac{2000}{q}$$

and we find its derivative as

$$\bar{C}'(q) = 0.05 - \frac{2000}{q^2}.$$

Setting this derivative to zero, we get

$$0 = 0.05 - \frac{2000}{q^2} \to 0.05 = \frac{2000}{q^2} \to q^2 = \frac{2000}{0.05} \to q = \sqrt{\frac{2000}{0.05}} = 200.$$

Thus, to minimize the average cost of producing the goods, we should make 200 goods. This is especially useful, since the cost function itself is only minimized for a negative number of goods! (Try it. You should get the derivative of the cost function as $C'(q) = 0.1q + 12$ which is minimized at $q = -12/0.1 = -120$ goods.)

14.2.3 Exploration 14B: Simple Regression Formulas

We have made extensive use of simple regression so far in this book. But how does simple regression work? How does the computer know how to compute the slope and y-intercept of the line that will minimize the total squared error in our approximation to the data? Wait a minute. That phrase "minimize the squared error" sounds important. It sounds like we can use calculus to find the answer.

First, let's do this with an example. Consider the following data points. We want to find the best fit (least-squares) regression line for these data.

x	1	2	3	4	5
y	5	7	8	9.5	12

What we want to do is to minimize the total squared error between the data and the regression line. If the line has the regression equation $y = A + Bx$, fill in the rows of the Excel table (file "C14 Regression.xls") with the appropriate calculation for each data point. (For now, just guess a value of the slope and y-intercept. Place these values as parameters on the spreadsheet.) Now, add up all the squared errors to get the total error, $E(A, B)$. This is a function of two variables, and we could treat it with calculus directly, but we'll simplify everything slightly by noting that the regression line always passes through the point $(\bar{x}, \bar{y})$ which means that $\bar{y} = A + B\bar{x}$. Rearranging this, we get $A = \bar{y} - B\bar{x}$. Now, let's put all this into Excel. You should have a sheet that looks a lot like the one below.

Now that we have the formulas entered, we can minimize the error function using the solver routine in Excel. Click on the cell containing the error value. Then click on "Tools/ Solver" and enter the values as shown in the screen shot below. It should very quickly find the value of the slope (B) that minimizes the total squared error. Now run simple regression on the data (Y = response, X = explanatory) to see what the regression routine gives as the best values for the parameters.

If we do this in general, using calculus and algebra, we can find some interesting facts. The total error function will look like (remember, we have eliminated the A variable with the relationship above)

$$E(B) = \sum_{i=1}^{n} (y_i - (A + Bx_i))^2 = \sum_{i=1}^{n} (y_i - A - Bx_i)^2 = \sum_{i=1}^{n} (y_i - Bx_i - (\bar{y} - B\bar{x}))^2$$

We can rearrange this last expression to be a little friendlier:

$$E(B) = \sum_{i=1}^{n} [(y_i - \bar{y}) - B(x_i - \bar{x})]^2 = \sum_{i=1}^{n} \left[(y_i - \bar{y})^2 - 2B(x_i - \bar{x})(y_i - \bar{y}) + B^2(x_i - \bar{x})^2\right]$$

This can be rearranged a little to get an expression that really looks like a second degree polynomial in B (with ugly coefficients - but they're just numbers!)

$$E(B) = B^2 \sum_{i=1}^{n}(x_i - \bar{x})^2 - 2B \sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}) + \sum_{i=1}^{n}(y_i - \bar{y})^2$$

The derivative of this is just

$$E'(B) = 2B \sum (x_i - \bar{x})^2 - 2 \sum (x_i - \bar{x})(y_i - \bar{y})$$

Setting this right hand side of this last expression equal to zero and solving for the parameter B we see that the error is minimized when

$$B = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2}.$$

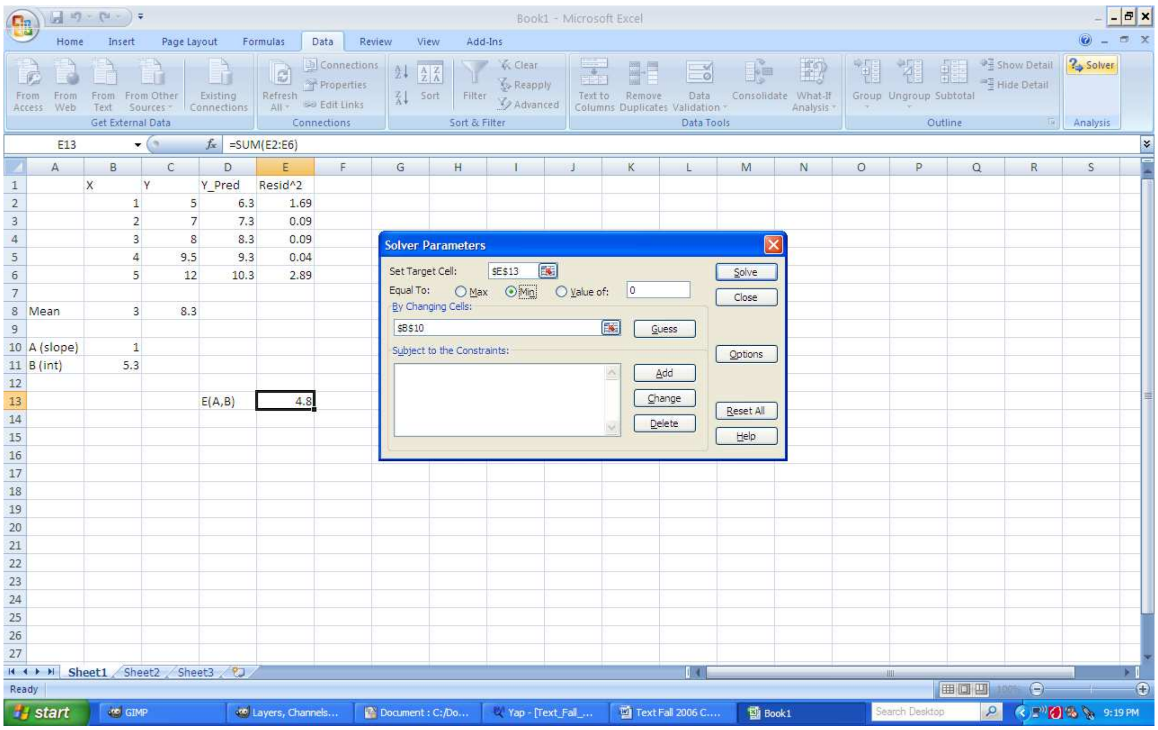


Figure 14.10: Screen shot for minimizing the total squared error.

14.2.4 How To Guide

Introduction to using SOLVER to minimize and maximize a function.

Excel has a very powerful equation solving tool built into it. This routine has limitations, and it certainly won't work for solving equations that don't have numerical values for the parameters, but it is a powerful tool for solving specific problems.

To use the solver, you need to have two things set up on your spreadsheet:

1. A cell that calculates something (the target cell)
2. Other cells (virtually any number of them) whose values are used in the calculation of the target cell (the parameter cells)

Using the solver is easy, once it's set up. Select the target cell. Then activate the solver routine with "Tools/ Solver". In the dialog box, make sure the "Set Target Cell" refers to the correct target cell. Then, click on the option for what solution you want: either maximum, minimum, or exact value - like goal seek. Finally, click in the space next to "By changing cells" and then highlight the parameter cells on the worksheet (use the control key to select multiple, non-adjacent parameter cells). Finally, hit the SOLVE button and let Excel compute.

Since the process is numerical, there are some errors that may occur. First, Excel may not find a solution. This can happen for a variety of reasons, but most often it's related to having the stating values of the parameter cells too far from the solution, so try changing the values of the parameter cells and starting over. You might also get problems if your target cell involves calculations with logarithms, since the process may need to try a variety of values for each parameter and this may lead to computing the log of a negative number, which is impossible.

Figure 14.11: Setting the parameters for the solver routine.

14.3 Homework

14.3.1 Mechanics and Techniques Problems

14.1. The function $\bar{c}(q) = 0.1q + 3 + \frac{2}{q}$ represents the average cost for producing q of a product. (Assuming that $q > 0$.) Find the minimum average cost and the number of goods that should be produced in order to achieve this minimum.

14.2. The function $\bar{c}(q) = \frac{10{,}484.69}{q} - 2.250 + 0.000328q$ gives the average cost for producing q goods.

1. Find a formula for the total cost of producing q goods by multiplying the average cost function by the number of goods produced.
2. Find the minimum total cost and the number of goods that should be produced in order to achieve this minimum total cost.

14.3. Given the points $(1, 12), (2, 7), (3, 5)$ and $(4, 6)$, assume that a linear function fits these points. Assume that the linear function passes through the point $(\bar{x}, \bar{y})$ so that the y-intercept, A, is given by $A = \bar{y} - B\bar{x}$ where B is the slope of the least-squares regression line.

1. Write down the exact error function, $E(B)$, as a function of slope for the total squared error between the data points and the regression line.
2. Minimize your total squared error function to find the slope of the least-squares regression line. Show all steps and explain all work.

14.3.2 Application and Reasoning Problems

14.4. We are given the following information regarding a product:

$$\begin{aligned} \text{Demand function:} \quad p &= 400 - 2q \\ \text{Average Cost:} \quad \bar{c} &= 0.2q + 4 + \frac{400}{q} \end{aligned}$$

1. The demand function gives the price people are willing to pay for the product, based on its availability (measured by q, the production). Use this to find the revenue function for this product.
2. Find the total cost function.

3. Find the profit function (profit = revenue - cost).

4. Use your profit function to determine the maximum profit and the number of goods to produce in order to achieve this maximum profit.

5. Based on your optimization of the profit, what is the price (look at the demand function!) at which the maximum profit occurs.

6. Suppose that the government imposes a tax of \$22 per unit on the product. What is the new maximum profit, number of goods needed to achieve maximum, and price?

14.5. Consider the profit graphs of each of the two companies shown in figure 14.12 from two different perspectives: the managers of the company (who want to keep their jobs) and the shareholders in the company (who want to make more money). For each graph, consider everything: the value of the function plotted on the graph, the rate of change of that function and the concavity (rate of change of the rate of change). Answer the following questions:

1. What might the managers say about this profit scenario in order to justify that they have been doing a good job leading the company and should keep their jobs?

2. What might the shareholders say to challenge the way the managers have run the company?

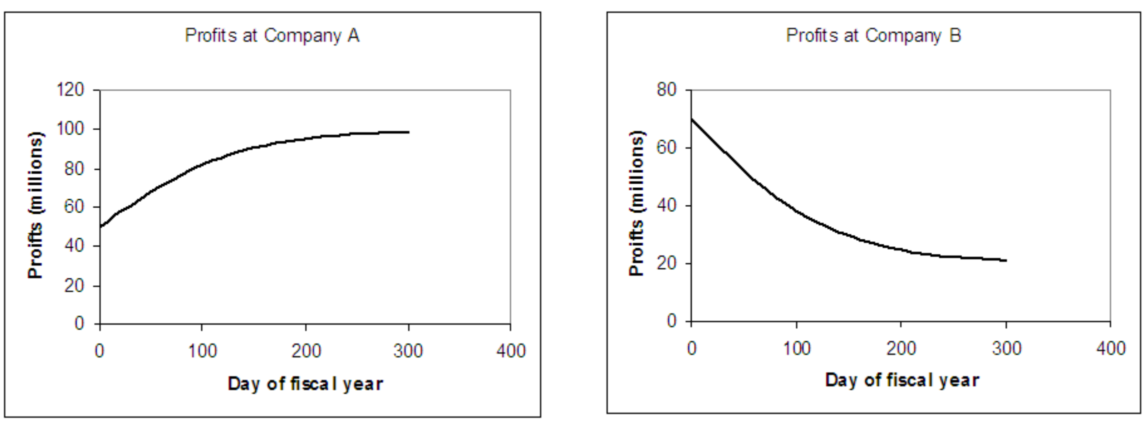


Figure 14.12: Profits over a year at two companies. Which is doing better?

14.6. Re-examine the situation in problem 5, only this time, imagine that the graphs given show the rate of change in the profits (millions of dollars per day) rather than the profits themselves.

14.7. Data file C14 MacroSoft Profits.xls contains data on weekly profits over each of the past 52 weeks. The profits are in thousands of dollars. Also shown are the corresponding number of units of software sold each week. At yesterday's board meeting, the operations manager claimed that the data shows profits are increasing as we produce more units of software. This means that the company can produce as much software as they want and continue to make profits. The CEO never believes news this good. Analyze the data, build some models for the profits, and analyze this claim.

14.3.3 Memo Problem

To: Analysis Staff
From: Project Director
Date: May 29, 2008
Re: Profit analysis for MacroSoft Software Company

A small, but up-and-coming software firm called MacroSoft has contacted us concerning a new software package they have developed. The CEO of the company, Bob Doors, has asked us to analyze three different production scenarios and to report on the findings. For each of the scenarios, he wants us to assume that the average cost of producing q million copies of the software is given by the function (with $q > 0$): $\bar{C}(q) = 0.01q^2 - 0.6q + 10$.

The units of this average cost function are in millions of dollars per million copies. Further, he expects that users will pay \$9.95 per copy of the software. Each of the three scenarios is described below. Mr. Doors has asked that the report contain both analytical calculations and spreadsheet calculations to verify these.

- Scenario A. In this production scenario, the company needs to know how many copies to produce (and then sell) in order to minimize the average cost for producing each copy of the software.
- Scenario B. In this production scenario, the company needs to know the total number of copies that it should produce (and sell) in order to minimize the total cost for producing the entire quantity of software.
- Scenario C. In this scenario, cost is no object. The company is interested in maximizing the profit earned from manufacturing and selling the software, no matter how many copies it takes to do it and regardless of the costs involved.

Your final report should include advice for manufacturing under each scenario and an overall comparison of each scenario, including: average cost, total cost, revenue, and profits. These should be in a nice table, and should be clearly explained for Mr. Doors - I know him, and he doesn't read anything that isn't fully explained and absolutely clear. Further, he would like a final recommendation on which of the scenarios his company should follow at the present. Again, keep in mind that this is a start-up company with limited production capacity.

Attachment: No attachment - you should create your own file to analyze this problem.