Chapter 13

Nonlinear Multivariable Models[1]

Most problems in the real world involve many variables. So far, you have encountered two types of models that have multiple independent variables: linear models and multiplicative models. These are definitely the most commonly used multivariable models since they are easier to interpret and can cover a variety of situations. But they do not cover all the possibilities. Probably the next most commonly used model is a quadratic multivariable model. This is the generalization of a parabola. This chapter will introduce you to this model in several ways. First, you will learn how to create such models using regression and interaction terms. Then you will learn how to graph and visualize some of these models. This approach to graphing quadratics can then be used to graph other types of nonlinear models.

- *As a result of this chapter, students will learn*
  - √ How to interpret certain quadratic models of two variables
  - √ The different shapes that the graph of a function of two variables can assume
  - √ How to simplify models with more than two variables when there are surrogate relationships
  - √ The difference between substitute and complementary commodities
- *As a result of this chapter, students will be able to*
  - √ Create a contour plot of a function of two variables
  - √ Create a 3D surface plot of a function of two variables
  - √ Use the discrimminant to determine the shape of a quadratic model

13.1 Models with Numerical Interaction Terms

In a previous chapter, we discussed building models using interaction terms. However, we only dealt with two of the three types of interaction terms: the interaction of two categorical variables and the interaction of categorical variable with a numerical variable. In this section, we will talk about what happens when you allow two numerical variables to interact, and what happens when you interact a variable with itself.

The second case is actually slightly easier to understand. Interacting a variable with itself produces a new variable in which each observation is the square of one of the observations of the base variable. Thus, a model built from a variable interacted with itself is a nonlinear model, specifically a square or quadratic model. This gives us another way to think about creating simple nonlinear models. Consider the data shown in the graph below, which has indication of being a parabola. The independent variable is Units (of electricity) and the dependent variable is Cost.

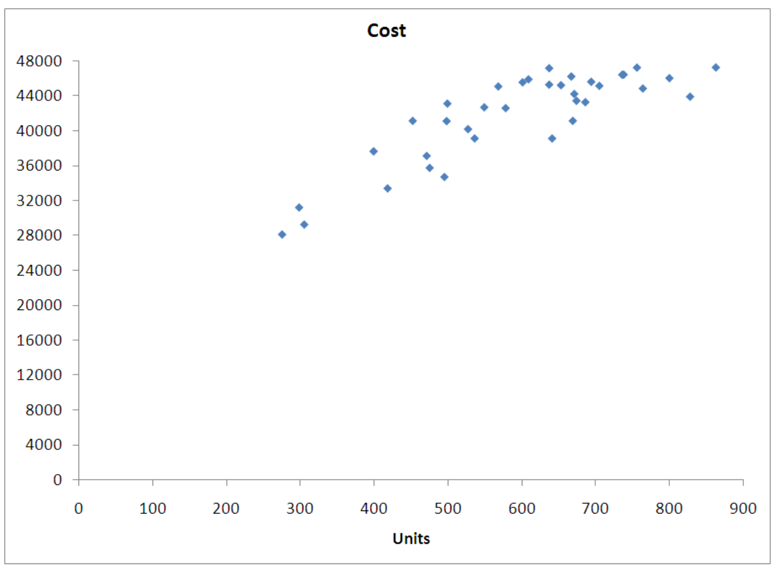

Figure 13.1: Electricity cost versus units used illustrating a nonlinear (possibly parabolic) relationship.

We can easily produce a quadratic model, and we find it has the equation

$$\text{Cost} = 5792.80 + 98.35 \cdot \text{Units - } 0.06\text{*Units} \cdot \text{Units.}$$

This model is clearly a parabola. It opens downward (as the graph shows) since the coefficient of the variable "Units · Units" is negative. (Of course, we don't expect there to be a discount for using too much electricity, so a quadratic model is perhaps not the most appropriate here, but you get the picture.)

The other situation - interacting two different numerical variables - is much harder to visualize, since we are dealing with at least three dimensions (one for each of the base variables plus one for the dependent variable). In the next section, you will work on interpreting such models and getting some sort of picture of what they might look like. For now, though, we concentrate on generating models of these two types, which are both quadratic models.

13.1.1 Definitions and Formulas

Interaction variable The product of two variables that constitutes a new variable and that captures, if it proves to be significant, the combined effect of the two original variables. Interaction terms can be created from any two variables. Most commonly, though, they are created from interacting either two categorical variables, or a categorical variable and a numerical variable (see chapter 10 for a discussion of such models).

Base Variable These are the original "uninteracted" variables from which the interaction terms were created.]

Quadratic model Any model made up of a combination of terms of the following forms: Constants, Constant · Variable, Constant · Variable2, Constant · Var1 · Var2.

Term A term is any object added to other objects in a mathematical expression. For example, in the function shown below, there are three terms: $3x$, 2 and $5xy$.

$$f(x, y) = 3x + 2 + 5xy$$

Factor In a mathematical expression, a factor is one quantity (a variable or constant) that is multiplied with other quantities to make a term. For example, in the function above, the factors of the term $5xy$ are 5, x, and y. The factors of the term $3x$ are 3 and x. The term "2" has only one factor, itself.

Factoring Mathematical/algebraic process of breaking terms into factored form so that several terms with similar factors can be grouped together. Often, this reveals hidden details of the model and can aid interpretation.

Self Interaction An interaction term created by multiplying or interacting a base variable with itself.

Joint Interaction An interaction term created by multiplying or interacting two different base variables.

13.1.2 Worked Examples

Example 13.1. Models built with one variable and self-interaction
Consider data on the Federal minimum wage, shown in "C13 MinWage.xls". This data shows the minimum wage (in dollars) at the end of each calendar year since 1950. Suppose we would like to build a model for this data in order to make projections about future labor costs for running a small company. Thus, we seek to explain the minimum wage, using the year as the independent variable.

One of the first things to note is that the years start in 1950 (when the minimum wage was established). This means that we are looking at large values for the independent variable, especially compared to the values of the minimum wage. It is helpful in situations like this to shift the independent variable to start at zero. StatPro can easily transform (under "StatPro/Data Utilities/Transform") the Year data into a new variable "Yr" representing the number of years since 1950. (This means that "Yr = 25" is the year 1950+25 = 1975.) One can also simply enter the formula "=A2 - 1950" in cell C2 and copy this down the column. Graphing the minimum wage versus the year since 1950 produces a graph like the following.

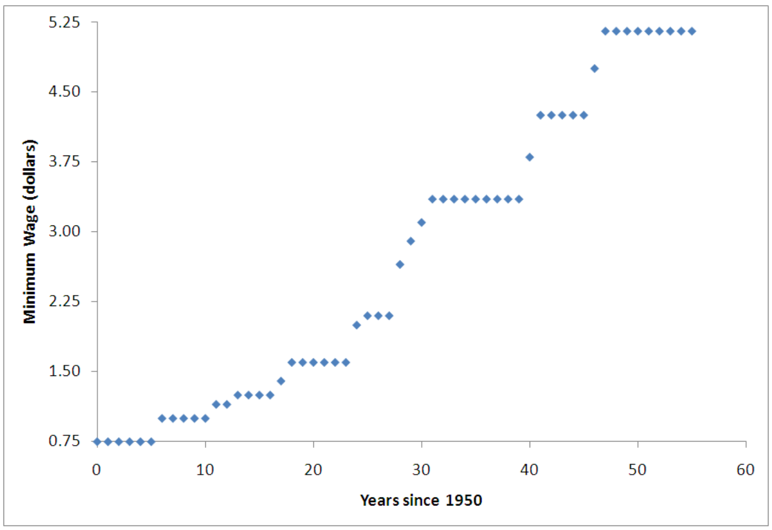

Figure 13.2: U.S. minimum wage versus years since 1950.

This graph clearly looks like part of a parabola, in spite of the high linear correlation. This means that it would be appropriate to introduce the interaction variable "Yr · Yr" and perform a multiple regression to build the model. The results of this are shown below.

The model equation is

$$\text{Minimum Wage} = 0.5196 + 0.0476 \cdot \text{Yr} + 0.0009 \cdot \text{Yr} \cdot \text{Yr}$$

We also see that the model has a coefficient of determination slightly worse than the linear model. This is due to the exact features of the graph; in particular, there are many years where the minimum wage does not change at all. The length of time the minimum wage stays constant seems to increase with time (since 1950) which stretches the graph out and makes the model slightly worse. A quadratic model, however, is clearly appropriate as can be determined from looking at the diagnostic graphs.

Results of simple regression for Price

Summary measures

Multiple R	0.9874
R-Square	0.9750
Adj R-Square	0.9740
StErr of Est	0.2539

ANOVA table

Source	df	SS	MS	F	p-value
Explained	2	133.0347	66.5174	1031.6093	0.0000
Unexplained	53	3.4174	0.0645		

Regression coefficients

	Coefficient	Std Err	t-value	p-value	Lower limit	Upper limit
Constant	0.5196	0.0983	5.2874	0.0000	0.3225	0.7167
Yr	0.0476	0.0083	5.7618	0.0000	0.0310	0.0642
Yr*Yr	0.0009	0.0001	5.8760	0.0000	0.0006	0.0011

One thing that is not apparent from this model, however, is what it means. Using a method called "completing the square" we can rewrite the model as

$$\text{Minimum Wage} + 0.1098 = 0.0009(\text{Yr} + 26.4444)^2$$

What this version of the model shows us is that the Minimum Wage plus about \$0.11 is modeled well by a scaled horizontally shifted power function! We can use the techniques of the last chapter to make sense of this power function: for every 1% increase in the number of years since 1950, the minimum wage should increase about 2% above its present level. In 2006, which is 56 years after 1950, a 1% increase in the year would be 0.01*56 = 0.56 years = 6.72 months. The minimum wage predicted by the model in 2006 is about \$6.01. The interpretation of the model is that we would expect the minimum wage to increase 2% (about \$0.12) to \$6.13 roughly six to seven months into the year 2006.

Example 13.2. Modeling with two interacting variables
Consider the data shown in file "C13 Production.xls". These data show the total number of hours (label "MachHrs") the manufacturing machinery at your plant ran each month. Also shown are the number of different production runs ("ProdRuns") each month and the overhead costs ("Overhead") incurred each month. In a previous chapter, we built the linear model shown below to explain these data.

$$\text{Overhead} = 3996.68 + 43.5364 \cdot \text{MachHrs} + 883.6179 \cdot \text{ProdRuns}$$

The model had a coefficient of determination of 0.8664 and a standard error of estimate of \$4,108.99, which was excellent compared to the standard deviation in overhead of \$10,916.81. In fact, it seemed the only problem with the model was the p-values for the constant term.

This was 0.5492, well above our 0.05 threshold for a "good" coefficient. So the question is can we improve on this without significantly complicating the model?

If we create all the possible interaction terms in the independent variables (these are MachHrs · MachHrs, ProdRuns · ProdRuns, and MachHrs · ProdRuns), we could create a full regression model and then reduce it by eliminating those variables with high p-values. Unfortunately, this produces a model with all p-values well above 0.05, leaving us no idea which to eliminate first. We need a better approach. Rather than begin with all the variables and eliminate, we will use stepwise regression to build the model up, one variable at a time. The result of this stepwise regression is the model below.

$$\text{Overhead} = 35{,}778.20 + 0.6240 \cdot \text{MachHrs} \cdot \text{ProdRuns} + 21.2566 \cdot \text{MachHrs}$$

This model has a coefficient of determination of 0.8628 and standard error of \$4,163.77, comparable to the linear model. However, the p-values for this model, including the constant term, are zero to four decimal places! Thus, the model more accurately shows the influential variables. But is this model too complex for interpretation?

One technique you may have encountered in previous mathematics classes is called factoring. Notice that the last two terms in the model both contain the same factor, MachHrs. Let's write the model in a different order without changing the model and then group the terms with similar factors together using parentheses, drawing that common factor out.

$$\text{Overhead} = 35{,}778.20 + (0.6240 \cdot \text{ProdRuns} + 21.2566) \cdot \text{MachHrs}$$

Now we notice that the model looks sort of linear. It's like the variable is MachHrs, the y-intercept is \$35,778.20 and the "slope" is 0.6240 · ProdRuns + 21.2566. Notice that since this is not a constant slope, we cannot truly call it such, but it can be interpreted this way: For each production run during the month, the cost of running the machinery for one hour increases by \$0.6240 from its base cost of \$21.26 per hour. So even though the model is quadratic and has an interaction term, it is still simple enough to interpret.

Example 13.3. Modeling with many interacting variables

In this example, we return to the commuter rail system introduced in an earlier chapter. If you recall, Ms. Carrie Allover needed a model to predict the number of weekly riders (in thousands of people) on her rail system based on the variables Price Per Ride, Income (representing average disposable income in the community), Parking Rate (for parking downtown instead of taking the rail system) and Population (in thousands of people). Previously, we developed a multilinear model for these data:

$$\begin{aligned}\text{Weekly Riders} \;=\; & -173.1971 - 139.3649 \cdot \text{Price per Ride} + 0.7763 \cdot \text{Population} \\ & -0.0309 \cdot \text{Income} + 131.0352 \cdot \text{Parking Rate}\end{aligned}$$

This model fit the data reasonably well, but we might ask whether we can do better, since the p-value for the constant term was so high (0.4389). Let's try a quadratic model. First, we create the interaction variables. There are four independent variables, so that gives us four variables representing self-interaction (Income · Income, Park · Park, Pop · Pop, Price

· Price) and $4 \cdot 3/2 = 6$ interaction terms created from two different variables. You can see the complete list of variables in "C13 Rail System.xls".

Clearly the full quadratic regression model will be complicated. Fortunately, many of the p-values in the full model are well above 0.05. Rather than build our model by eliminated variables one at a time, though, let's retrace our steps and perform a stepwise regression. We'll submit "Weekly Riders" as the response variable and we will submit all of the variables (the four base variables, the four square terms and the six interaction terms) as possible explanatory variables. StatPro will then build the model up from nothing adding in only the relevant variables rather than having us work from the full model and eliminate variables. The result is much simpler than we might have expected.

$$\begin{aligned}\text{Weekly Riders} &= 596.491 + 0.0002 \cdot \text{Pop} \cdot \text{Pop} - 0.0864 \cdot \text{Price} \cdot \text{Pop} \\ &\quad +36.0244 \cdot \text{Park} \cdot \text{Park} - 0.0229 \cdot \text{Income}\end{aligned}$$

This model has a coefficient of determination of 0.9342 and standard error of 23.0119, which are not very different from the linear model we started with, but we gain one significant advantage: all the p-values are significant.

Still, our model has four independent variables involved. This makes it extremely difficult to interpret. One way to do so would be to rewrite the model slightly by factoring the terms involving Population.

$$\begin{aligned}\text{Weekly Riders} &= 596.491 + \text{Pop} \cdot (0.0002 \cdot \text{Pop} - 0.0864 \cdot \text{Price}) \\ &\quad +36.0244 \cdot \text{Park} \cdot \text{Park} - 0.0229 \cdot \text{Income}\end{aligned}$$

This leaves us with a model indicating that:

- For each \$1 increase in disposable income, we expect 0.0229 thousand (about 23) fewer riders each week.
- Population has a generally positive effect on ridership, but its effect is mitigated by the price per ride; for each \$1 increase in ticket price, we expect the effect of population to be decreased by 0.0864 thousand riders per thousand people in the population.

Obviously, this model is complicated. Interpreting it is still difficult. However, we can reduce this model to a quadratic model of two variables by taking advantage of some of the natural correlations in the data. Looking at the correlations (table 13.2) shows us that there are strong linear relationships between Income and Parking Rates and between Price per Ride and Parking Rates. These relationships are shown in table 13.3 below.

	Weekly Riders	Price per Ride	Population	Income	Parking Rate
Weekly Riders	1.000				
Price per Ride	-0.804	1.000			
Population	0.933	-0.728	1.000		
Income	-0.810	0.961	-0.751	1.000	
Parking Rate	-0.698	0.958	-0.645	0.970	1.000

Model	Correlation	R^2	S_e
Income = 2046.8727 + 3191.5617 · Park	0.970	0.9408	505.1306
Price = -0.0929 + 0.5672 · Park	0.958	0.9176	0.1072

In the equation above, we substitute these relationships (replace Income with 2046.8727 + 3191.5617 · Park and replace Price with -0.0929 + 0.5672 · Park) and eliminate those two variables (which are surrogate variables for Parking Rate, apparently). The reduced model looks like

$$\begin{aligned}\text{Weekly Riders} &= 596.491 + 0.0002 \cdot \text{Pop} \cdot \text{Pop} - 0.0864 \cdot (-0.0929 + 0.5672 \cdot \text{Park}) \cdot \text{Pop} \\ &\quad +36.0244 \cdot \text{Park} \cdot \text{Park} - 0.0229(2046.8727 + 3191.5617 \cdot \text{Park}).\end{aligned}$$

Simplified, this model becomes

$$\begin{aligned}\text{Weekly Riders} &= 549.618 + 0.0002 \cdot \text{Pop} \cdot \text{Pop} + 0.00799 \cdot \text{Pop} - 0.0490 \cdot \text{Park} \cdot \text{Pop} \\ &\quad +36.0244 \cdot \text{Park} \cdot \text{Park} - 73.0868 \cdot \text{Park}.\end{aligned}$$

This two-variable quadratic model is simpler in many ways than the original nonlinear model. However, we will leave interpretation of this model to the next section, when we learn how to picture this model as a surface in three-dimensions.

13.1.3 Exploration 13A: Revenue and Demand Functions

File "C13 Exploration A.xls" contains weekly sales and revenue information for two different companies. The first worksheet, labeled "Company 1" shows the quantities of two complementary commodities that are sold by this company. These items are X and Y. The second sheet contains data on two substitute commodities sold by "Company 2".

1. Formulate a quadratic regression model for Company 1's revenue as a function of the quantity of each item that is produced and sold.

2. Formulate a quadratic regression model for Company 2's revenue as a function of the quantity of each item that is produced and sold.

   You should now have two revenue functions that look something like this:

   $$R(q_1, q_2) = Aq_1^2 + Bq_2^2 + Cq_1q_2 + Dq_1 + Eq_2 + F$$

   Where the capital letters are constants and variables q_1 and q_2 represent the quantity of goods of each type.

3. Explain why, in the revenue formula above, you would expect F, the constant term, to be zero. Do your regression models match this prediction?

   We are going to use these revenue functions to determine the demand functions for the products in each case. Recall that the demand function gives the unit price that the market will pay for something, given the supply (in this case the quantities q_1 and q_2) of the item(s) being sold. To find the demand functions, we need to write the revenue function in the form

   $$R(q_1, q_2) = q_1p_1 + q_2p_2$$

   In this formula, the p_1 and p_2 are the unit prices. We will assume that these are both linear functions of the two quantities.

4. What does it mean in the last sentence when it says that p_1 and p_2 are a linear function of the quantities? Give a sample function that could represent p_1 or p_2.

5. Try to find the demand functions for each situation. You can do this by (a) factoring the regression models you have formulated above and (b) assuming that the term with the coefficient C in the revenue formula is split equally between the two demand functions.

6. Use your demand functions to fill in the tables below, showing the estimated prices customers would pay at each company for different supplies of the two goods.

Company 1			
q_1	q_2	p_1	p_2
1000	1000		
1100	1000		
1000	1100		

Company 2			
q_1	q_2	p_1	p_2
2000	2500		
2100	2500		
2000	2600		

7. Based on your demand functions (you should now have four: two for each scenario) and your data in the tables above what do you think are meant by the terms "complementary commodities" and "substitute commodities"?

13.1.4 How To Guide

Interacting a variable with itself

StatPro will let you create a new interaction variable from two different variables; however, it will not allow you to interact a variable with itself. This is easy to do manually with a small formula, though. First, you create a new column for the interaction variable. To remind yourself what it represents, you might call it "Variable2" where "variable" is replaced with the name of the base variable. Second, you enter a formula in the first cell of the new column that computes the product of the variable with itself. Finally, you copy this formula down the column.

In the spreadsheet shown below (the original file without the interaction terms is "C13 Rail System.xls") you see that we have added a column for the new variable "Park2" which will be the interaction of the variable ParkingRate with itself (column G). Next, in cell G4 we entered =F4*F4 and then copied this down column G by double-clicking on the fill handle for cell G4.

	A	B	C	D	E	F	G
3	Year	Weekly_Riders	Price_per_Ride	Population	Income	Parking_Rate	Park2
4	1966	1,190	$0.15	1,800	$2,900	$0.50	0.25
5	1967	1,135	$0.15	1,790	$3,100	$0.50	0.25
6	1968	1,161	$0.15	1,780	$3,200	$0.60	0.36
7	1969	1,165	$0.25	1,778	$3,250	$0.60	0.36
8	1970	1,171	$0.25	1,750	$3,275	$0.60	0.36
9	1971	1,136	$0.25	1,740	$3,290	$0.70	0.49

Figure 13.3: Numerical-on-numerical interaction terms in the rail system data.

Pitfalls of numerical-numerical interaction variables

When creating models with numerical-numerical interaction variables, you are much more likely to encounter an error. The most common is the "multiple colinearity" error. In well-designed and collected data sets with several numerical variables, it is quite likely that there are many hidden relationships among the different variables. For example, if you attempt to predict car maintenance costs based on the two variables of age and mileage, we expect that older cars also have more mileage. Thus, the two independent variables are not truly independent.

Using mixed cell references to compute a table of function values

To compute a table of values for a function of two variables, say $f(x, y)$, it is very efficient to use mixed cell references. These are cell references with either the column or the row fixed (by placing a $ in front of it) but not both of them (that would be an absolute cell reference).

For example, to make a graph of $z = 4 + 2x - 3y + x^2 + 2y^2 - 5xy$, we first set up the spreadsheet with values of x across the first row (skipping the first cell) and values of y down the first column (skipping the first cell). So if we wanted to graph this function for values of x from -5 to 5 and values of y from -10 to 20, we might set the spreadsheet up as shown (see "C13 HowTo.xls" for the data and the sample graph; the labels for the "X Values" and "Y Values" were done using the "merge and center" feature). Note that the x values run left to right, and the y values run top to bottom. Also note that we have left the first cell blank; this is important in helping Excel "guess correctly" when it formats your graph (See "Making a 3D surface plot" in the next section).

D4 | =4+2*D$2-3*$B4+D2^2+2*B4^2-5*D$2*$B4

	A	B	C	D	E	F	G	H	I	J	K	L	M
1			X Values										
2			-5	-4	-3	-2	-1	0	1	2	3	4	5
3	Y Values	-10	-1	42	87	134	183	234	287	342	399	458	519
4		-8	-29	4	39	76	115	156	199	244	291	340	391
5		-6	-41	-18	7	34	63	94	127	162	199	238	279
6		-4	-37	-24	-9	8	27	48	71	96	123	152	183
7		-2	-17	-14	-9	-2	7	18	31	46	63	82	103
8		0	19	12	7	4	3	4	7	12	19	28	39
9		2	71	54	39	26	15	6	-1	-6	-9	-10	-9
10		4	139	112	87	64	43	24	7	-8	-21	-32	-41
11		6	223	186	151	118	87	58	31	6	-17	-38	-57
12		8	323	276	231	188	147	108	71	36	3	-28	-57
13		10	439	382	327	274	223	174	127	82	39	-2	-41
14		12	571	504	439	376	315	256	199	144	91	40	-9
15		14	719	642	567	494	423	354	287	222	159	98	39
16		16	883	796	711	628	547	468	391	316	243	172	103
17		18	1063	966	871	778	687	598	511	426	343	262	183
18		20	1259	1152	1047	944	843	744	647	552	459	368	279

Figure 13.4: Table of values for a function of two variables.

The next step is to enter the formula, paying careful attention to the types of cell references we need. Every time we refer to the x value in our formula, we need to fix the row (like C$2 in the formula highlighted above), since the x values are always in row 2, but will be in different columns depending on the specific x value we want to use. The y values are always in column B, so when we refer to a cell with a y value in it, we need to freeze the column (like $B3 in the formula shown above). The full formula (in cell C3) is shown below to make it easier to read. Of course, we could have set up cells with parameters for the function instead of typing the 4, 2, -3, etc. This would have allowed us to easily see how changing these numbers changes the shape of the graph, but it would have made the formula harder to read for this example. In general, you should always use parameters, rather than "hard coding" the numbers into the formula.

```
B2 = 4+2*C$2-3*$B3+C$22̂+2*$B32̂-5*C$2*$B3
```

Once these references are correct, we copy the formula to all the other cells in the table, as shown. The result shows us the value of z for a given value of x (the column) and y (the row). Thus, when $x = -2$ and $y = 10$, we find that $z = 274$. This procedure will work with any type of function of two variables, linear or nonlinear.

13.2 Interpreting Quadratic Models in Several Variables

When dealing with multivariable models, there are, literally, an infinite number of ways to explore them, depending on what kind of graph you want, which part of the model you want to graph, whether you would prefer looking at the data in a table of numbers, or a host of other possible choices. It helps to have some basic skills and options for visualizing functions with two independent variables. As we'll see, graphing them requires three dimensions, one for each independent variable and one for the dependent variable. Thus, if you want to graph a model with more than two independent variables, you need some mighty special paper!

Obviously, one way to gain an understanding of how the function behaves is to make a table of data. You've seen such tables before for functions of several variables, you just didn't realize it. One very common example relates to the weather. You've heard of wind chill probably. This is a measure of how cold the air feels, based not only on the actual temperature, but also on the wind speed. To use such a table (like the one below) you simply locate the intersection of the wind speed (down the left column) and air temperature (across the top row) to find the wind chill. Such a process defines a function of two variables. If we let W stand for the wind chill, S for wind speed and T for air temperature, then we could write

$$W = W(S, T)$$

to represent the relationship; this emphasizes that W is a function of S and T. For example, $W(25, 10) = -29$ indicating that a 25 mph wind on a 10 degree day makes the air feel like it is actually 29 degrees below zero!

Wind Speed (mph)	Ambient Air Temperature (degrees Fahrenheit)																
	35	30	25	20	15	10	5	0	-5	-10	-15	-20	-25	-30	-35	-40	-45
5	33	27	21	16	12	7	1	-6	-11	-15	-20	-26	-31	-35	-41	-47	-54
10	21	16	9	2	-2	-9	-15	-22	-27	-31	-38	-45	-52	-58	-64	-70	-77
15	16	11	1	-6	-11	-18	-25	-33	-40	-45	-51	-60	-65	-70	-78	-85	-90
20	12	3	-4	-9	-17	-24	-32	-40	-46	-52	-60	-68	-76	-81	-88	-96	-103
25	7	0	-7	-15	-22	-29	-37	-45	-52	-58	-67	-75	-83	-89	-96	-104	-112
30	5	-2	-11	-18	-26	-33	-41	-49	-56	-63	-70	-78	-87	-94	-101	-109	-117
35	3	-4	-13	-20	-27	-35	-43	-52	-60	-67	-72	-83	-90	-98	-105	-113	-123
40	1	-4	-15	-22	-29	-36	-45	-54	-62	-69	-76	-87	-94	-101	-107	-116	-128

But, making tables of the data from a function is only one way to study its behavior. And, the table of numbers may be difficult to read and interpret. In addition, the spacing of the values in the table may hide some important features. For example, the wind chill table makes it appear that no matter what, if the wind speed increases, the air feels colder (wind chill is lower). But what if between 20 and 25 mph, it actually gets a little warmer for some reason? Our table would not show this.

So, another common tool for studying such functions is to create 3D surface plots of them. If we copy the table above into Excel and create such a plot, we get a figure like the

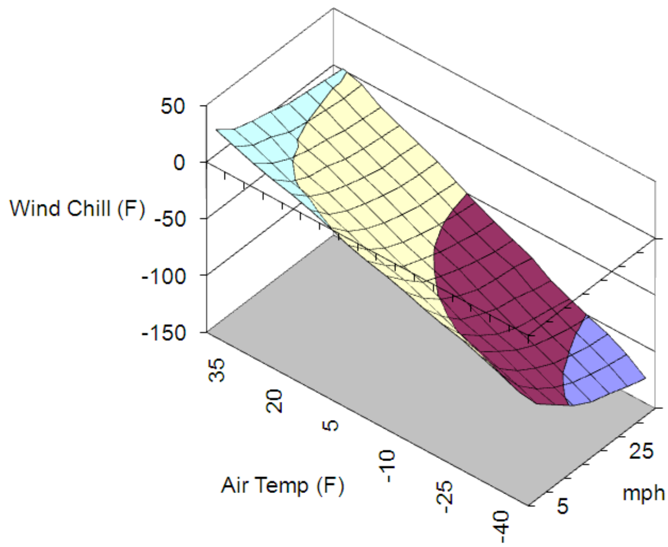

Figure 13.5: 3D plot of wind chill versus air temperature and wind speed.

one below. We can adjust the perspective of the graph, but otherwise, it has many of the same features as all the scatterplots we've used before.

In this section, we will use this graphical tool to help us understand the different types of quadratic models that we may get from applying the techniques of the previous section. In general, we will be dealing with models of the form

$$f(x_1, x_2) = E + A_1x_1 + A_2x_2 + B_1x_1^2 + B_2x_2^2 + Cx_1x_2$$

And will want to know what different shapes the graphs of such functions may take. Fortunately, there are only a few possibilities, and we will learn some ways of quickly classifying any function as being one of these types (either a bowl-shaped surface, a hill-shaped surface, or a saddle-shaped surface)

While it may seem restrictive to study such as specific class of functions, it turns out that there are several good reasons for it. The first is that it arises easily in modeling, as the techniques of the last section showed. The second is that if we zoom on the surface of any random function of two variables, on a small enough scale it looks like a quadratic. Thus, studying these objects gives us a lot of tools for understanding more complex objects.

13.2.1 Definitions and Formulas

Dimensions For each variable (independent or dependent) in a model, you need one dimension in order to create a graph of the model. Thus, a model like $y = f(x)$ needs two dimensions, one for y and one for x. A model like the general quadratic below needs three dimensions for its graph.

Surface Plot A graphic representation of a function of one variable (two dimensions) is a scatterplot. Creating a similar type of graph for a function of two variables requires three dimensions. Each point has three coordinates, and the height of the point above the xy-plane is the value of the function. When the points are connected together, they form a surface in three dimensions.

General Quadratic Model The general quadratic model we will use in this text is

$$f(x_1, x_2) = E + A_1x_1 + A_2x_2 + B_1x_1^2 + B_2x_2^2 + Cx_1x_2$$

In this, we assume that at least one of the B coefficients is non-zero. Other texts may refer to the model in slightly different terms, but the important things to note are that (1) this is a polynomial model (in two variables) and (2) the degree of each term (sum of the powers of each variable) is either 0, 1 or 2. For example, the terms with a B coefficient all have one variable raised to the second power and the other raised to the zeroth power, so they are degree 2. The cross term (the term with the C coefficient that involves both independent variables) has both variables raised to the first power, so its degree is $1 + 1 = 2$ as well.

Discriminant There are several mathematical objects that go by the name "discriminant". Each is used to discriminate between several alternatives. In this case, we are referring to a quantity that can be derived from the formula for the general quadratic that helps decide whether the graph will look like a bowl, a hill or a saddle. Using the symbols above, the discriminant is the quantity

$$D = 4B_1B_2 - C^2$$

The shape of the graph (as we will see in the examples), depends on this quantity in the following ways:

1. If $D > 0$ and $B_1 > 0$, then the graph will look like a bowl.
2. If $D > 0$ and $B_1 < 0$, then the graph will look like a hill.
3. If $D < 0$, then the graph will look like a saddle.
4. If $D = 0$, then the discriminant is not helpful.

There are two other possible shapes for the graph, which occur if the coefficients in front of all instances of one variable are zero. In that case, the graph looks like either a trough (if the remaining B coefficient is positive) or a speed bump (if the coefficient is negative).

Depending on your viewpoint and the exact values of your graph, you may not be able to see it has a particular shape, though (see example 5 (page 394)).

13.2.2 Worked Examples

Example 13.4. Looking at a multi-linear function
Recall the model from the previous section that represented our best, linear efforts to model the overhead based on the machine hours and production runs:

$$\text{Overhead} = 3996.68 + 43.5364 \cdot \text{MachHrs} + 883.6179 \cdot \text{ProdRuns}$$

File "C13 Production2.xls" shows a table of values for this function, calculated over a domain similar to that present in the data. Below is a 3D surface plot of these data, showing the linear structure.

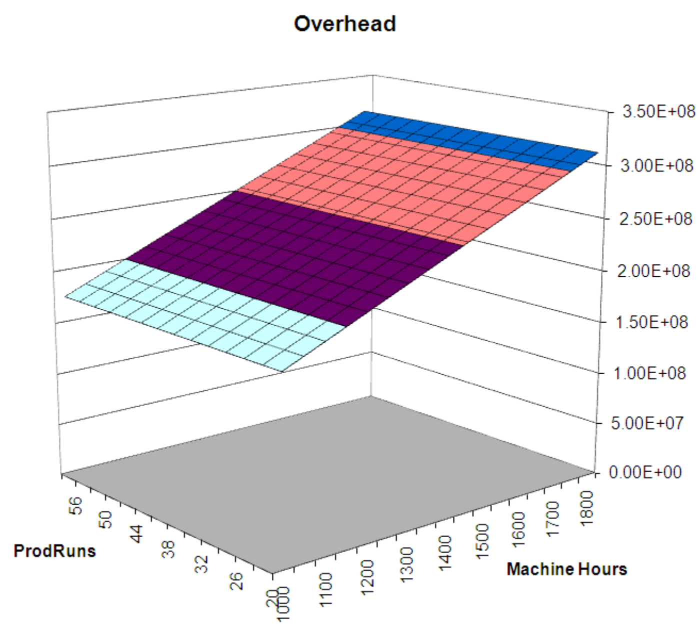

Figure 13.6: Linear two-variable model of overhead versus Production Runs and Machine Hours.

Notice that this graph appears to be a flat plane, like a piece of paper tilted at an angle. Any linear function of two variables has such a graph.

Example 13.5. Looking at a quadratic function of two variables
Here is one possible graph for a quadratic function of two variables. This is based on the quadratic model of the overhead costs found in "C13 Production2.xls" in the worksheet labeled "Example 13B2". It uses the model shown below.

$$\text{Overhead} = 35{,}778.20 + 0.6240 \cdot \text{MachHrs} \cdot \text{ProdRuns} + 21.2566 \cdot \text{MachHrs}$$

Notice that the formula in cell C5 uses mixed cell references (see the "How To Guide" for details) in order to calculate the overhead from a given number of machine hours (in column B) and a given number of production runs (row 5).

```
C5 = 35778.2 + 0.624*$B5*C$4 + 21.2566*$B5
```

The graph of this model is shown below.

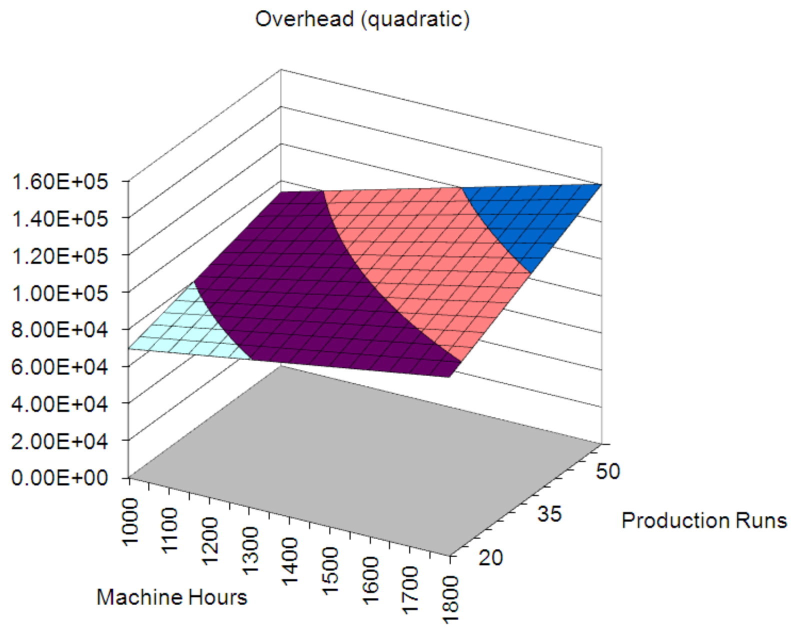

Figure 13.7: Quadratic two-variable model of overhead versus Production Runs and Machine Hours.

Notice that this graph also appears, at first glance, to be linear - like a plane. However, the contour lines on the surface between the different colored regions are curved, indicating that this is truly a nonlinear model. The reason it doesn't look quadratic is because of the particular set of values of MachHrs and ProdRuns we have used to graph the function. When we graph it over a larger region, we can clearly see the warped "saddle" shape of the surface become apparent. Of course, we could never have negative values of machine hours or production runs in a given month, so the actual data will never show this. Thus, we see that even when the data may be best represented by a nonlinear model, it may not be clear from the graph.

Also note that in the notation given in the "Definitions and Formulas" for the discriminant, we have $B_1 = B_2 = 0$ and $C = 0.6240$. This means that the discriminant, D, is -0.62402, which is less than zero, confirming that we should see a saddle in the graph.

For the sake of completeness, we view the graph of overhead from above (graphed on the region with all independent variables positive). Such a graph is called a contour plot and shows curves (called contours) that separate regions based on their coordinate in the third dimension. Notice that all of the contours are curved, another indication that the underlying graph is nonlinear. In fact, it can be shown that these curves are hyperbolas, a type of object closely related to parabolas.

Example 13.6. Another quadratic surface

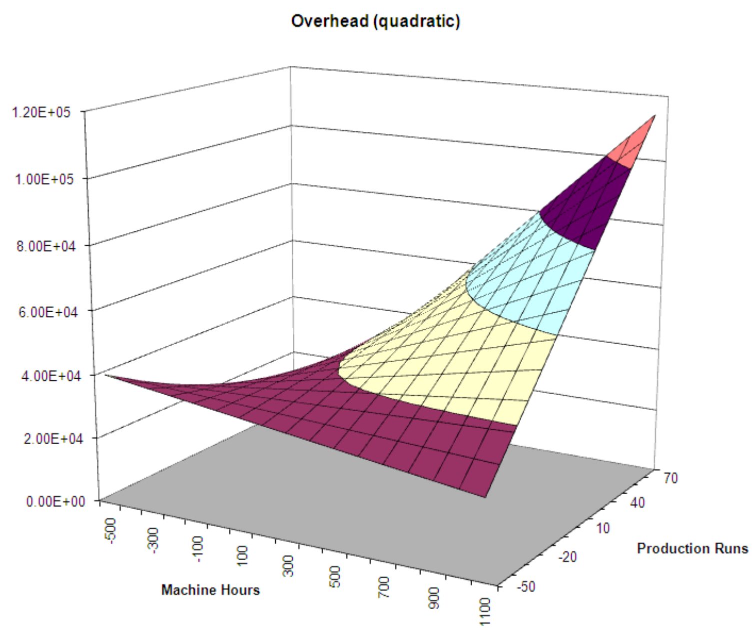

Figure 13.8: Quadratic two-variable model of overhead versus Production Runs and Machine Hours. Note that this is graphed over a different domain than in figure 13.7 emphasizing the nonlinear nature of the graph.

Let's look at a graph of the surface representing the quadratic Weekly Riders model from example 3 (page 384). This model, after reducing it to two variables, became

$$\begin{aligned}\text{Weekly Riders} \;&=\; 549.618 + 0.0002 \cdot \text{Pop} * \text{Pop} + 0.00799 \cdot \text{Pop} - 0.0490 \cdot \text{Park} \cdot \text{Pop} \\ &\quad +36.0244 \cdot \text{Park} \cdot \text{Park} - 73.0868 \cdot \text{Park}\end{aligned}$$

When graphed over the region with Parking Rates from \$0.50 to \$2.50 and Population between 1,000 thousand people and 2,000 thousand people, we appear to see a linear model. But a calculation of D gives $D = 0.0264$ which is positive. Since the coefficients of the squared terms are both positive, this seems to indicate that we should see a bowl-shaped surface. How are we to reconcile the calculation with the graph?

This is always part of the problem in graphing and interpreting nonlinear models, especially those of several variables: such functions tend to have large domains, and tend to look very different at different locations in the domain. To emphasize this, we look at the graph on a slightly expanded domain where the shape is more evident.

Example 13.7. Multiplicative models
As a final example, we will look at a graph of one of the other multivariable, nonlinear models we have encountered, the multiplicative model. The model below is a Cobb-Douglas production model. P represents the total production of the economy, L represents the units of labor available and K represents the units of capital invested. We met such models in the

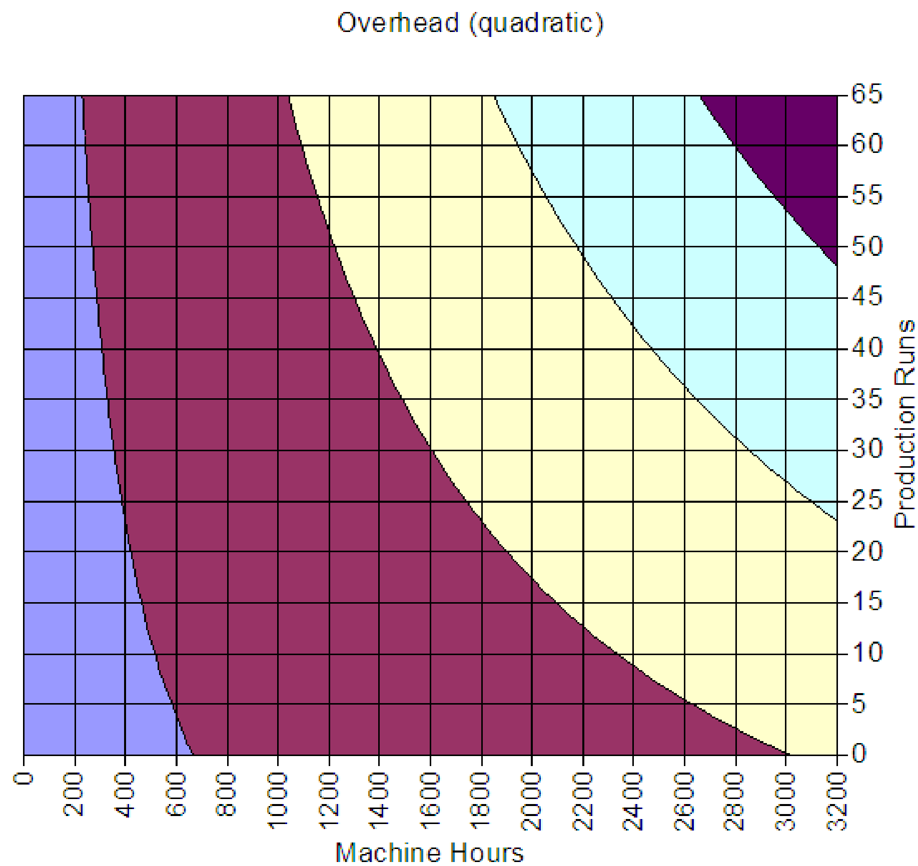

Figure 13.9: Contour view of the quadratic model of overhead. Note that the contours (or level curves) are not straight lines, as in a linear model, but are curved.

last chapter and applied parameter analysis to their interpretation. But what do they look like?

$$P = 0.939037L^{0.7689}K^{0.2471}$$

As you can see from the graph below, when we plot the production for reasonable values of the labor and capital (both positive) the contours look like those of a saddle-shaped surface, but the graph does not look like a saddle. The graph shows that if either of the inputs is zero (capital or labor) the production is zero. It also shows that if you increase either input (or both) you continue to get more output.

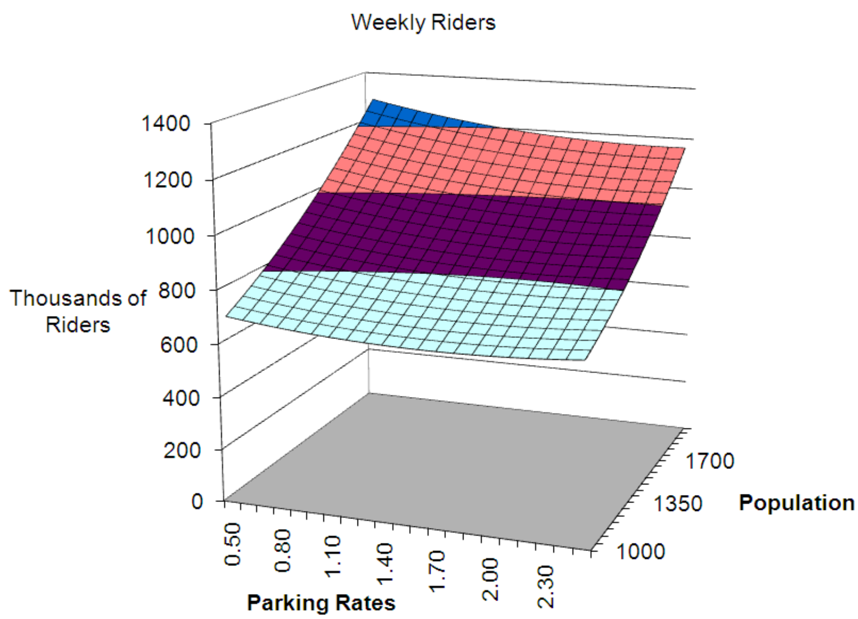

Figure 13.10: Quadratic model of weekly riders versus population and parking rates.

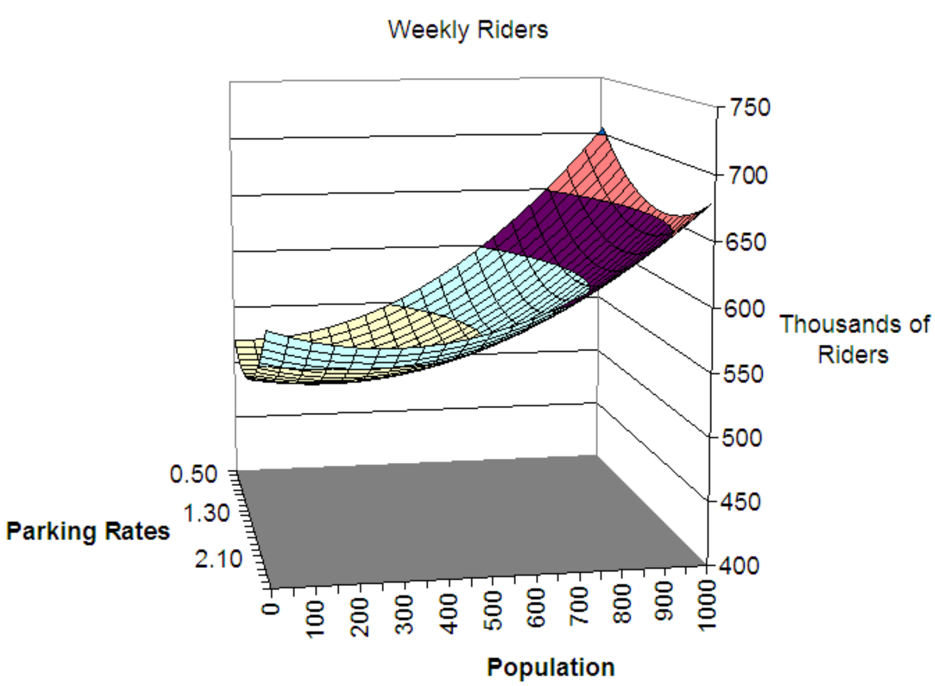

Figure 13.11: Different view of the graph in 13.10 showing the bowl-shape.

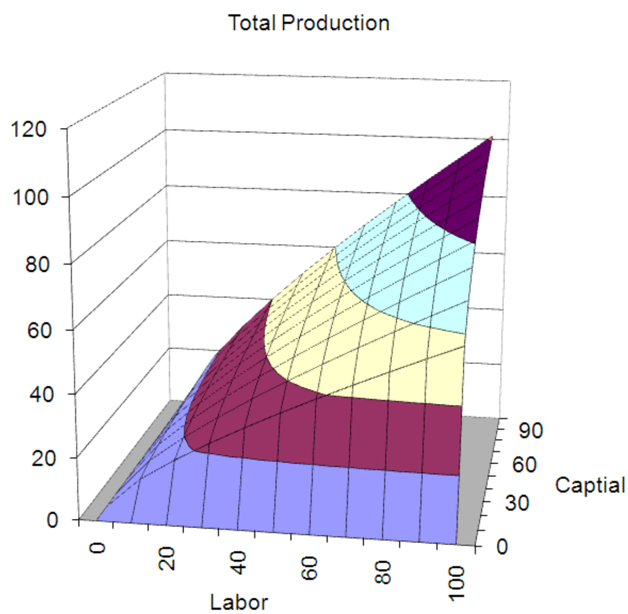

Figure 13.12: 3D plot of a Cobb-Douglas model, illustrating the nonlinear nature of the model.

13.2.3 Exploration 13B: Exploring Quadratic Models

In this exploration, you will get a chance to connect the different shapes of the quadratic graphs to the values of the coefficients and see some realistic examples where these different shaped graphs might occur. Consider the revenue generated from selling two different products. Since revenue is the quantity sold (q_1 will be the quantity of item 1 sold; likewise for item 2) times the unit price of the item (p_1 will be the unit price of item 1) we can reasonably assume that the revenue function looks something like this:

$$R(q_1, q_2) = q_1p_1 + q_2p_2$$

Depending on the particular goods, we might have the prices of each item related to the quantity of both items sold. Two common situations in which this occurs are when the items are either substitute commodities, which means that people buy one or the other, but not both, or when they items are complementary commodities, where people who buy one item tend to buy the other. For example, a car company might sell one model of SUV and one model of sedan; most people buy one or the other. Thus, sedans and SUVs tend to be substitute commodities. On the other hand, since all cars need tires, we expect increased car sales to result in increased tire sales; cars and tires are complementary commodities.

We could get these relationships for the prices from the demand functions for the two items. For now, we'll assume that the demands are linear in the prices so that:

$$p_1 = c_1 + a_1q_1 + b_1q_2 \text{ and } p_2 = c_2 + a_2q_1 + b_2q_2$$

In these expressions, the coefficients a, b, and c are all constants. The exact values of these constants depend on the relationship between the two commodities being sold.

Open the file "C13 Revenue Exploration.xls" to explore how these coefficients influence the shape of the graph and the decisions that you might make in order to achieve the best possible revenue. When you open the file, depending on your computer's security settings, you may need to click on the "Enable Macros" button in order to make the exploration active. If all is working properly, you should have two slider bars in the upper right corner and moving these around should change the shape of the graph; if it doesn't see the "How To Guide" below for details on adjusting the computer's security settings.

It is important to note that there are, potentially, six constants in the expression that you could change. We have rigged the exploration file, though, so that you can control just two of these with the slider bars, and the other four will change in a particular way. This makes it easier for you to see what is happening on the graph and allows you to focus your attention on the important features. The coefficients that you can change with the sliders are in cells C3 and D4: these represent the quantities a_1 and b_1 in the expressions above for the demand. You will also notice that the discriminant is calculated for you, in cell G1, to help you make some sense of what you are seeing.

Part A. First, move the sliders around to get a feel for how they interact and produce different shapes of the surface. Then concentrate on specific values of the coefficients that produce the different shapes. Finally, for one example of each shape, explain what the values of the coefficients mean in terms of the relationship between the two goods under investigation.

Interpretation of the Graphs

Now, focus on one of your graphs. The method we will use to interpret the graph is referred to as the "method of sections". The idea is that we fix the value of one of the independent variables; for example, we could let $q_2 = 500$. Now we imagine moving across the surface of the graph, always keeping q_2 fixed, but letting the other variable, q_1 increase. The interpretation follows by thinking about what happens to the dependent variable as the free variable increases at a fixed value of the other variable (the "sectioning variable"). For example, if you push the two sliders all the way to the right, so that cells J1 and J2 show the value of 1000, you have a graph that looks like a hill. Now, imagine setting $q_2 = 500$ and exploring the surface along this path by letting q_1 increase from 0 to 300. You might describe this exploration in the following way:

Along the section $q_2 = 500$, the total revenue seems to be increasing until the point where q_1 is about 200. Up the that point, the revenue is increasing, but at a decreasing rate (the hill is concave down). After $q_1 = 200$, the revenue begins to decrease as q_1 increases.

Similar statements can be made along any section (fixed value of one of the variables). This is very much like our interpretations of multivariable models that we have used before. The main differences are that (1) this is a graphical method and (2) we are referring to this as "sectioning in q_2" rather than "controlling for q_2" as we did in the algebraic versions.

Part B. Now, for each of the graphs you focused on in part A, describe several sections of the graph. Be sure to section the graph in both of the variables. You may want to change the viewing angle for the 3D graphs to help you visualize the surface better for some sectionings (See the How To guide for this).

13.2.4 How To Guide

Making a 3D surface plot

Once you have a table of values for a function of two variables, whether generated from actual data or from a formula (see above), it is relatively easy to create a surface plot. First highlight all of the table (in file "C13 HowTo.xls", we would highlight B2:M18). Notice that the first cell, the empty one, is included in this. Then from the insert ribbon, select "Other charts" and choose the first of the surface chart types.

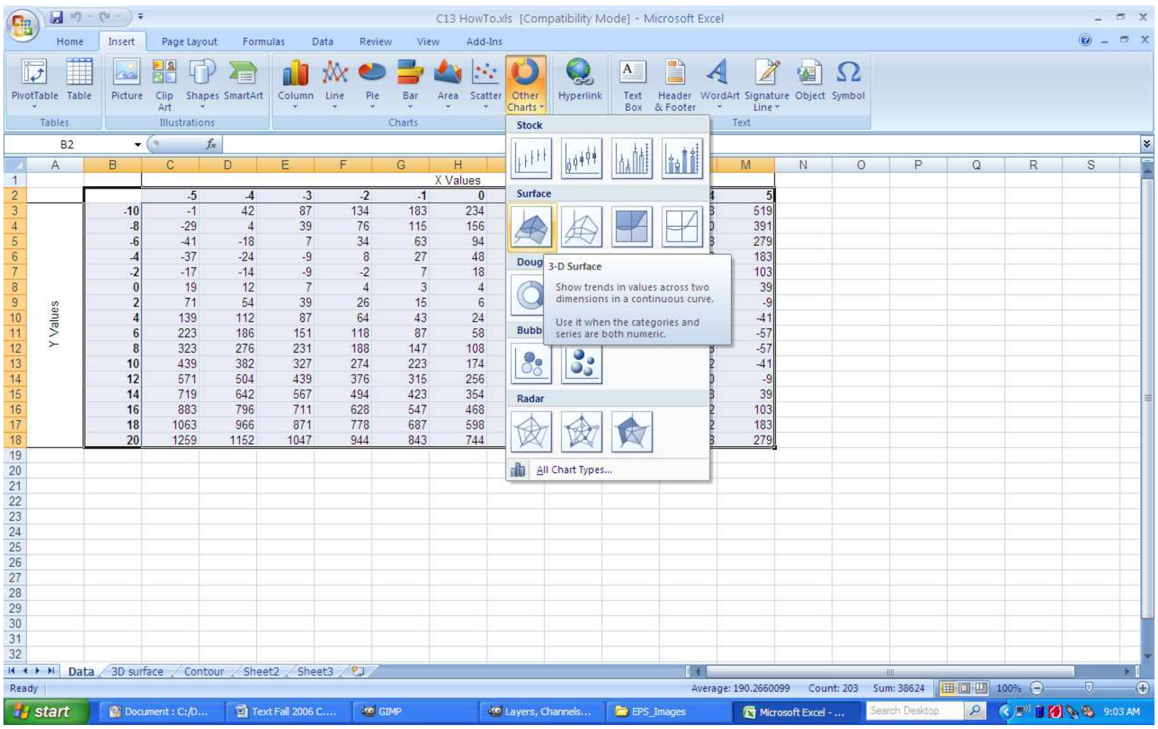

Figure 13.13: Inserting a 3D surface graph in Excel.

Adjusting security settings for macros in Excel

If you open a file with macros (for example, the sliders in exploration 13B) you will see one of three things happen. Either:

1. You will be asked if you want the macros to be active (medium security),
2. The macros will be automatically active and you will see nothing (low security), or
3. The macros will automatically be disabled (high security).

Macros are simply collections of instructions (a small program, basically) that have been connected to make them easier to run together, rather than having to repeat all of the commands each time you want to reproduce that set of actions. Macros are a common way to distribute computer viruses, so many recommend that you think carefully about enabling them all the time; we prefer to use medium security, so that we are asked before macros are enabled in a particular file.

To adjust your security settings, open the Excel Options menu by clicking on the Office button in the upper left corner of the Excel window and clicking on "Excel options" at the bottom. In the dialog box, click on the Trust Center (see figure 13.14) and click the "Trust center settings" button in the lower right corner. Select the level of security you wish to have from the screen in figure 13.15, then click "OK" and "OK" again to apply these settings.

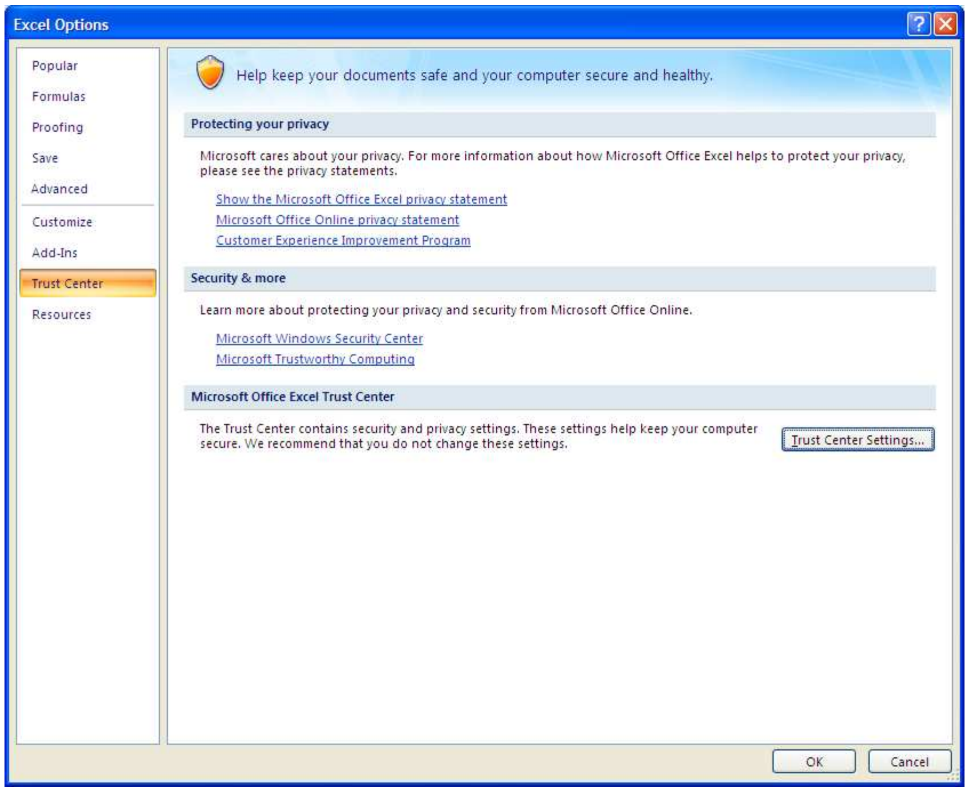

Figure 13.14: The Excel options menu, showing the trust center active.

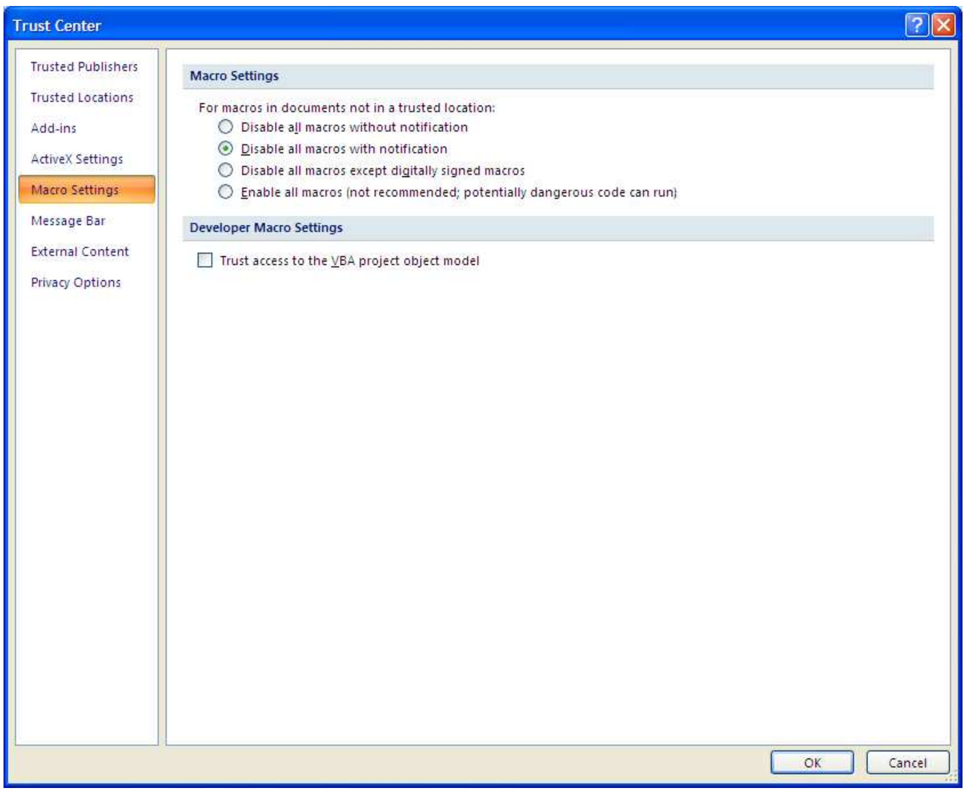

Figure 13.15: Adjusting the macro settings in the trust center.

13.3 Homework

13.3.1 Mechanics and Techniques Problems

13.1. Answer each of the following questions, given the function of two variables: $f(x, y) = 8xy - 3x^2 + 2y^2$.

1. Find the value of the function when $x = 2$ and $y = 1$.
2. Determine a value of y so that when $x = 10$, the function is equal to 124. You may use algebra, Goal Seek or some other method to find the answer, but explain your solution method.
3. Create a graph of the function of one variable $g(x)$ where $g(x) = f(x, 3)$.

13.2. Using the discrimminant identify the shape of the 3D surface plot of each function below. Describe the shape as being either: a bowl, a hill, a saddle, or impossible to tell.

1. $f(x, y) = 2x^2 - 3xy + y^2 + 4x - 5$
2. $g(x, y) = 3x^2 - 2xy + y^2 + 4y - 5$
3. $h(x, y) = -3x^2 + 2xy - y^2 + 4y - 5x + 1$
4. $k(x, y) = -0.3x^2 + 0.2xy - 0.1y^2 + 4y - 5x + 1$

13.3. Get Bent, Inc. sells assembled and unassembled recumbent bicycles. The estimated quantities demanded each year for the assembled and unassembled bikes are x and y units when the corresponding unit prices (in dollars) are

$$\begin{aligned} p &= 2000 - \frac{1}{5}x - \frac{1}{10}y \\ q &= 1600 - \frac{1}{10}x - \frac{1}{4}y \end{aligned}$$

1. Find the annual total revenue function, $R(x, y)$.
2. Find the approximate domain of the revenue function. That is, find the set of values of x and y such that the unit prices are all positive.
3. Create a 3D surface plot of the revenue function for all points (x, y) in the domain.
4. Create a 3D contour plot of the revenue function for all points (x, y) in the domain.

13.4. The revenue function below was developed as a model for the revenue data "Shaken and Stirred" collected regarding its sales of gin (x) and vodka (y). The sales quantities of each are measured in liters. The company would like to know if the revenue function supports the notion that their products are complementary commodities.

$$R(x, y) =$$

1. Factor the expression to put it into the form below. Assume that the mixed term (the xy term) splits equally into the two demand functions.
2. From your factored revenue function, identify the demand functions for gin (x) and vodka (y) sold by Shaken and Stirred.
3. Analyze your demand functions and explain whether the products are complementary commodities or substitute commodities.

13.5. The contour diagram below shows the total revenue from selling two different products.

1. Give at least four sets of production pairs (q_1, q_2) such that the revenue is positive.
2. Give at least four sets of production pairs (q_1, q_2) such that the revenue is greater than 200,000.

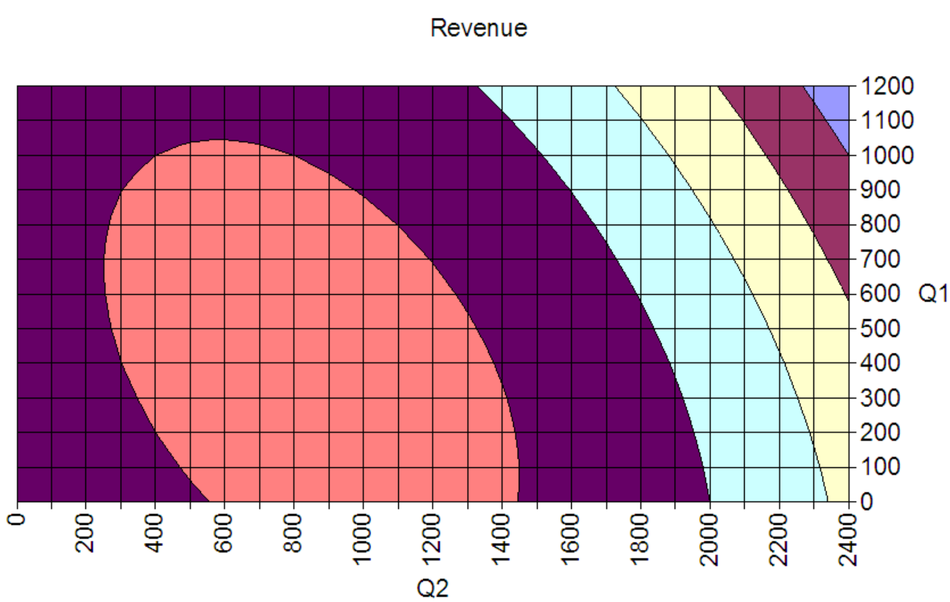

Figure 13.16: Revenue versus quantity of two products being sold, problem 5.

13.3.2 Application and Reasoning Problems

13.6. The graphs below show contour plots of the demand function for one product out of a pair of products sold by the same company. In each graph, the demand function plots the unit price when x and y units of the two products are demanded. Which company is selling two complementary commodities? Which is selling two substitute commodities? Explain your answer.

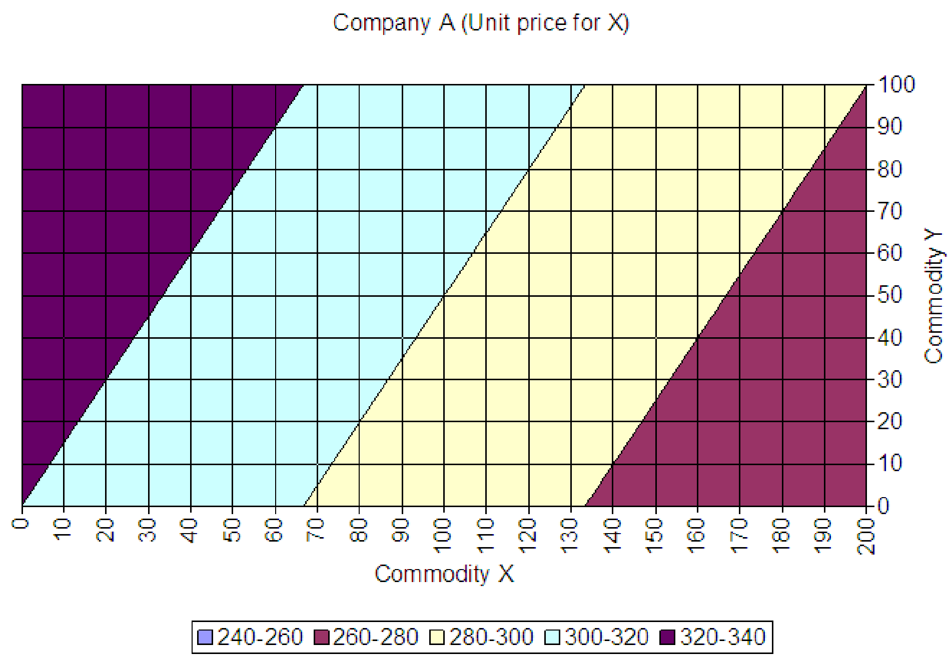

Figure 13.17: Contour plot of demand function for Company A in problem 6.

13.7. Metro Area Trucking has been gathering data regarding a different approach to predicting maintenance costs of its trucking fleet. There is a considerable growing body of research suggesting than uneven tire tread wear is related to maintenance costs for a variety of reasons including worn front end parts, worn or weak suspension, and even the vibrations of a roughly running engine. The surface of the roadway has been shown to affect uneven tire wear, which might relate to maintenance costs even apart from tire wear, and uneven tire wear is a direct contributor to high gasoline costs. Metro has developed an index for measuring uneven tire wear. Every three months the treads of the four tires of a van are each measured in three places by a digital gauge to the nearest 64th of an inch. The standard deviation of the three measurements taken on each tire is calculated and then scaled from 1 to 100 in whole numbers for easy reading. This is called the tire's tread index. The more uneven a tire is, the larger its standard deviation, and the higher its tread index. The largest index measured from the four tires on the van is recorded. The idea is that this index, which is a measure of the driving conditions to which the truck is subjected, interacted with the

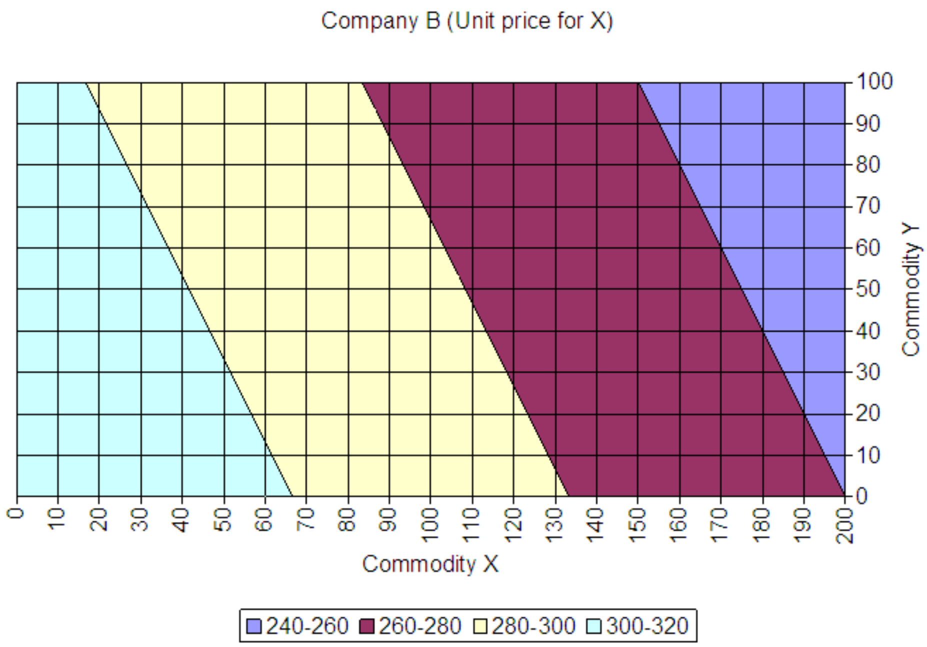

Figure 13.18: Contour plot of demand function for Company B in problem 6.

number of miles the truck is driven, might very well be a good predictor of maintenance cost.

1. From the data in C13 Truck Data.xls, build a model with interaction terms (self and joint)
2. Discuss the goodness of fit of your model
3. Interpret the model.

13.8. Consider the following model to explain the number of tickets sold each week in a large metro public transportation system:

$$\begin{aligned}\text{Riders} &= 1486.7960 + 0.0681 \cdot \text{Income} - 29.3 \cdot \text{TicketPrice} - 2.3324 \cdot \text{GasPrice} \cdot \text{GasPrice} \\ &\quad + 1.4625 \cdot \text{TicketPrice} \cdot \text{Income} + 13.8049 \cdot \text{TicketPrice} \cdot \text{TicketPrice}\end{aligned}$$

In this model, the variable "Income" represents average weekly disposable income for a family of four in the greater metropolitan area (in dollars), "TicketPrice" represents the price of a ticket on the transit system (in dollars), and "Gas Price" is the median price for a gallon of regular unleaded gas (in dollars).

But the model, with three variables, is too complicated for explaining to the city council at the upcoming meeting. You have noticed that, within the time span that this model was based upon, you found that

$$\text{Income} = 260.00 - 3.1 \cdot \text{TicketPrice}$$

Use this information to find a simpler way to express the model and interpret the simplified model both algebraically and graphically.

13.9. For a fixed amount of principal, A (in dollars), the monthly payment (\$) for a loan of t years at a fixed APR of r is given by the formula below.

$$P = f(A; r, t) = \frac{Ar}{12\left[1 - \left(1 + \frac{r}{12}\right)^{-12t}\right]}$$

1. Create a 3D surface plot for the monthly payment of such an amortized loan for a reasonable domain of t and r. Use $A =$ \$100,000 as the principal for the loan.
2. Using your graph, what happens to the monthly payments as the interest rate r increases, but the term of the loan (t) stays fixed? Does it depend on the value of t, or is the effect independent of t? Explain.
3. Using your graph, what happens to the monthly payments as the term of the loan t increases, but the interest rate (r) stays fixed? Does it depend on the value of r, or is the effect independent of r? Explain.

13.10. Home mortgage rates are designed so that the amount of principal and amount of interest in each payment varies over the life of the loan, but the monthly payment remains fixed. For a loan of A dollars and a term of t years, the total amount of principal paid by the end of month i of the loan is given by the formula below.

$$B = f(A, t; r, i) = A\left[\frac{\left(1 + \frac{r}{12}\right)^{i} - 1}{\left(1 + \frac{r}{12}\right)^{12t} - 1}\right]$$

1. Suppose you borrow \$100,000 for a home on a 30-year loan at 6.25% APR. How much will you have left to pay after 1 year (12 months)? After 5 years (60 months)? After 15 years (180 months)?
2. Suppose you borrow \$125,000 for a home on a 30-year loan. Create a 3D plot showing the amount of principal remaining after month i at an interest rate of r. Use values of r between 2% and 10%, in intervals of 0.25%. Make sure your graph covers the entire period of the loan.
3. From your graph, what can you infer about the amount of principal in each monthly payment when you are at the beginning of the load repayment? At the end?

13.3.3 Memo Problem

To:	Analysis Staff
From:	Project Management Director
Date:	May 29, 2008
Re:	Revenue Projections at Dream Grills

One of our smaller clients, Dream Grills, sells its one product, the Dream Grill 5000, in two forms: assembled and unassembled. Based on economics theories about substitute commodities, they have been making projections and analyses for their business plan based on the following models of their revenue.

$$\begin{aligned} R(Q_A, Q_U) &= Q_A P_A + Q_U P_U \\ P_A &= 462 - 0.1Q_U - 0.35Q_A \\ P_U &= 372 - 0.20Q_A - 0.16Q_U \end{aligned}$$

In these models, the P and Q refer to the price and the quantity of the two items; the subscripts A and the U refer to the "assembled" and "unassembled" versions of the product. Thus, the quantity P_A is the price of the assembled grills, based on the quantities of each version that are sold.

The company has collected revenue and quantity sales data for the last 50 weeks. Formulate a regression model for the revenue and compare the two models, yours and theirs, using graphical and analytical tools you feel are appropriate to illustrate the differences.

Attachments: C13 Revenue.xls