Optimization Methods in Economics [1]

John Baxley
Department of Mathematics
Wake Forest University

June 20, 2015

[1] Notes (revised Spring 2015) to Accompany the textbook Introductory Mathematical Economics by D. W. Hands

Contents

Preface

This material is written for a half-semester course in optimization methods in economics. The central topic is comparative statics for economics problems with many variables. The ideal reader is approximately equally prepared in mathematics and economics. He or she will have studied mathematics through vector calculus and linear algebra and have completed intermediate courses in both microeconomics and macroeconomics.

It is intended that the text material be roughly half mathematics and half economics. Essentially all students in the course are engaged in the joint major at Wake Forest in mathematical economics, which is provided as a cooperative project of the Departments of Economics and Mathematics at Wake Forest. This effort began in the mid-seventies and has flourished, primarily because of a deep commitment on the part of members of the faculties of both departments. The contributions of these faculty members have been characterized by a respect for both disciplines and a commitment to appreciate and understand a dual point of view. Looking at the material simultaneously from the angles of a mathematician and an economist has been a fertile intellectual discipline.

An important part of any education should be becoming adept at learning from books. Students in mathematics complain, perhaps more than other students, about the difficulty of books. It is not really true that mathematicians purposefully make it difficult to learn from books. The fault, dear reader, lies with the subject. Mathematics is not a narrative subject. Mathematics lives on an intellectual terrain, in a person's mind. Words and symbols are put on paper attempting to describe that intellectual terrain. It is necessary that readers somehow translate these words and symbols into a vision in their own minds. Probably no two people "see" this vision exactly the same, and that is probably good. By seeing the material from different angles, different valuable insights are gained. So part of reading a book in mathematics is for the reader to create his or her own vision of the material and attempt to describe, using words and symbols, what that vision looks like.

So the material here is the result of my interaction with some of the material in Hands' book. It is the attempt to describe my version of the vision. Naturally, it seems clearer to me than the attempt made by Hands. Whether that is true for you remains to be seen. Nevertheless, it gives a second version of the material and covers exactly the material for Math 254. One could view the result as a set of "Cliff notes" for Hands' book.

The author cannot commit his version of these ideas to paper without expressing his great appreciation to Professor John Moorhouse of the Department of Economics at Wake Forest. For twenty-five years, it was my privilege to work with him in a jointly taught seminar in mathematical economics, hear him lecture on much of the material in this text, formulate and attack interesting problems with him, and learn to see the subject through his intellectual eyes. His influence is present on each page of this draft.

In particular, I have learned from Professor Moorhouse a very valuable pedagogical prin-

ciple for teaching this material. Coming to such a subject from the influence of most books in applied mathematics written by mathematicians, my natural inclination would have been to present the material in a pendulum fashion: a section on mathematical methods, then a section on applications of these methods to economics problems chosen to illustrate these methods, and repeat this cycle over and over. John Moorhouse taught me a better way: begin with the economics problem, pose it carefully, and then solve it by using the mathematical tools. Do this again with one or two more problems using the same tools, and finally explain the tools. This is like teaching basketball by first having the students watch a well-played game by capable players, then watch a couple more, and finally let the students try their hands, explaining some of the better moves. This way the applications are the objects of study, they are the nouns. The mathematical methods are the means, they are the adjectives.

So I have taken economics as the central subject with mathematics providing the methods: I seek to let economics carry the mathematics as a truck would carry its cargo. I am convinced that this is the best way to present the material. On the one hand, the students are predominantly economics students who like mathematics or feel somewhat talented in mathematics. They are not particularly interested in mathematics for its own sake, but they are attracted to the idea of becoming better in economics than their competition because they are skilled users of mathematical methods. On the other hand, mathematics is not a static subject: it is best learned not in isolation but in action, solving the problems it was created to solve. A student who learns a mathematical idea in tandem with an application has a stronger hold on the idea and is more likely to know when and how to use it. Students who learn mathematics in isolation from applications often remind one of a carpenter who has learned how to make all the tools but has never used one: he or she foolishly tries to use a screwdriver to drive a nail or a hammer to drive a screw.

The reader will form his or her own opinion as to the degree of my success in this effort. It has been clear to me throughout my academic career that a professor sees a text rather differently from a student. All of us who teach have had the experience of using a "beautiful" textbook which our students did not appreciate, or find that our students like a book with which we are unimpressed. It is a rare book indeed which is praised by both faculty and students. In this case, I hope the student reader will find these notes helpful.

Chapter 1

Elementary Comparative Statics

Max-min problems play a central role in every calculus course. Finding relative (local) maxima and minima using the derivative and applying the first or second derivative test is the name of the game in curve-sketching as well as the "applied" problems in the calculus books. The student who comes to economics from such calculus courses often feels betrayed. Slowly it becomes evident that economists do not spend their time finding maxima and minima. In fact, quite the opposite is true. Unlike the typical math problem where one "finds the maximum", the economist *assumes* that the economic agent (firm, consumer, etc.) is instinctively maximizing. The fundamental assumption is that somehow such economic agents have a built in computer or natural instinct which leads them to maximizing behavior. The central question for the economist is not: find the maximum, but: how will the agent adjust maximizing behavior if some variable which he cannot control undergoes a change. For example, how will the quantity of snack crackers sold in the marketplace change if the price of a related good like Coca-cola rises? This question assumes that consumers are maximizing their utility and as they face higher prices for Coke, they will make adjustments in their expenditures which may effect the amount of snack crackers which are sold.

One of the interesting sidelights of this state of affairs is that economists deeply wish that second order conditions were necessary rather than sufficient for an optimum. You will recall the second derivative test: if c is a critical point of a function $f(x)$ (i.e. $f'(c) = 0$), then $f''(c) < 0$ is a sufficient condition for c to be a maximum. However, this condition is not necessary, for $f(x) = 9 - x^4$ has a maximum at the critical point $c = 0$ but $f''(0) = 0$. If we are math students and are on the prowl for maxima, the second derivative test can be used to determine if c is a maximum, but if we are economists we want the thought to flow the other way: if we know that c gives a maximum, we would like to conclude that $f''(c) < 0$. Unfortunately, we can only conclude that $f''(c) \leq 0$. The possibility that $f''(c) = 0$ is often disastrous for economic analysis, because in the analysis this value occurs in a denominator and leads to division by zero. Wishful thinking has led many economists to argue that this disaster is somehow very unlikely and can be safely ignored; some economists have actually referred to the disaster as "pathological" in nature. The example $f(x) = 9 - x^4$ does not look very pathological!

As we proceed, you will find that we will often have to assume that second order conditions known only to be sufficient actually hold at the maximum. We don't really have a viable choice. Either we make this assumption and draw an interesting economics conclusion, or we don't make the assumption and no conclusion can be drawn! The only fringe benefit, attractive for students with no pride, of not making the assumption is that the course becomes the empty set and study time is minimized.

1.1 Static Equilibrium: A One-Good World

Hands, pp. 91-96. (pp. 113-119, 1st ed.)

We begin with a very simple example which you may have seen before. If so, that is fine. We choose it because we can illustrate two elementary methods on it and it does not require any deep thought. It is the simple supply and demand problem:

$$Q^D = a - bP, \tag{1.1}$$

$$Q^S = c + dP. \tag{1.2}$$

Here Q^D represents the quantity demanded of some good and Q^S represents the quantity supplied. Both depend on the price P of the good and these demand and supply "curves" are just straight lines. We assume that the three parameters a, b, d are positive; c may be positive or negative (why?). Thus the demand curve slopes down and the supply curve slopes up, as elementary economics books insist. Such a good is called a normal good. Our problem is not a "real world" problem. Demand and supply curves are not straight lines and our world has many more than one good; all these goods interact in complicated ways. Nonetheless, this simple idealized problem is instructive.

Let us first re-think the problem in the style of an introductory economics text. This style may be called diagrammatic analysis. Such diagrams (graphs) are drawn at the beginning of Chapter 3 in Hands' book. However, unlike Hands, I prefer to put Q on the vertical axis and P on the horizontal axis. If you draw such a graph of the demand and supply curves on the same axis, you get a unique intersection point (P^*, Q^*) of the supply and demand curves giving the *static equilibrium* values of the price P^* and the quantity Q^*; thus P^* is the "market-clearing" price of the good. The question of most interest to an economist is: "What happens to the static equilibrium price and quantity if there is a shift in the demand curve or supply curve?" After a shift upwards in the demand curve parallel to itself, it is clear that both the static equilibrium price and the static equilibrium quantity has increased. However, if the supply curve shifts upward parallel to itself, the static equilibrium price decreases while the static equilibrium quantity increases. A shift upward in the demand curve is represented by an increase in a; a shift upward in the supply curve is represented by an increase in c. You should draw diagrams which will help you make conclusions about shifts in the static equilibrium if the slope of the supply curve (or the demand curve) increases. (Do not be confused by the differences in conclusions caused by putting P in the vertical axis and Q on the horizontal axis as Hands does.)

The answers to the questions we have been entertaining are called comparative statics. The values of P^* and Q^* here are the numerical values we seek to understand; mathematicians might call them the dependent variables; economists call them the endogenous variables. The values of the four numbers a, b, c, d are viewed as numbers which affect the values of P^* and Q^*; mathematicians might call them the independent variables (or the parameters of the model); economists call them the exogenous variables. Although the actual quantitative value of a shift in P^* might be of interest, the focus is usually only on the *direction* of the shift. We have seen that an increase in a in our simple model causes both an increase in P^* and Q^*, so both have positive responses to an increase in a. Thus these comparative statics are positive. However P^* has a negative response to an increase in c; this comparative static is negative.

It is instructive now to reconsider the entire problem from the point of view of algebra (and a little calculus). Static equilibrium occurs precisely when quantity demanded agrees with quantity supplied, i.e. $Q^D = Q^S$. Thus we may equate these two to get

$$a - bP = c + dP,$$

and solve for P to get

$$P^* = \frac{a-c}{b+d}.$$

Thus $P^* > 0$ as long as $a > c$. This was the case in all the diagrams we have been drawing, and may be assumed true in any case of interest. (Does your economic intuition tell you why?) Once we know P^*, we may substitute it into either the demand curve or the supply curve to get

$$Q^* = \frac{ad+bc}{b+d}.$$

These formulas show clearly that Q^* will increase if either a or c increases, but the effect of an increase in b or d on Q^* is not obvious. Similarly, an increase in a clearly increases P^*, while an increase in c decreases P^*. The effect of an increase in b or d is clearly to decrease P^*. These effects can be made quantitative by computing partial derivatives:

$$\frac{\partial P^*}{\partial a} = \frac{1}{b+d} > 0, \quad \frac{\partial P^*}{\partial c} = -\frac{1}{b+d} < 0,$$

$$\frac{\partial P^*}{\partial b} = \frac{\partial P^*}{\partial d} = -\frac{a-c}{(b+d)^2} < 0,$$

and

$$\frac{\partial Q^*}{\partial a} = \frac{d}{b+d} > 0, \quad \frac{\partial Q^*}{\partial c} = \frac{b}{b+d} > 0,$$

$$\frac{\partial Q^*}{\partial b} = -\frac{d(a-c)}{(b+d)^2} < 0, \quad \frac{\partial Q^*}{\partial d} = \frac{b(a-c)}{(b+d)^2} > 0.$$

Of course, all these inequalities agree with the conclusions from diagrammatic analysis. (Otherwise, all rational efforts would be worthless!) But more comes from this little bit of calculus: we now have formulas for the quantitative effect of shifts in exogenous variables on endogenous variables. For example, the effect of a shift in a on P^* is more pronounced if $b + d$ is small and almost negligible if $b + d$ is large. This realization might be a little slower to obtain from diagrammatic analysis.

All this thought has been concerned with an impossibly simple one-good world. Yet we have achieved from very elementary thought processes results which confirm our intuition and widespread economic conviction. Let's stay in this radically oversimplified one-good world for a little longer, but let's make things a little more realistic by giving up the simple linear supply and demand functions. Let's assume that

$$Q^D = f(P, a), \tag{1.3}$$

$$Q^S = g(P, b), \tag{1.4}$$

where demand depends not only on price but also some exogenous variable a, which might represent taste factors, and supply depends not only on price but also on some exogenous variable b, which might represent certainly technological factors. How can we attack this problem? We can no longer solve for the equilibrium values P^* and Q^*. Before proceeding, we must take action to prevent notational confusion. It helps enormously if we use subscript notation for partial derivatives. So let's marshal all the reasonable economic information we might use:

$$f_P(P,a) = \frac{\partial f}{\partial P} < 0, \quad g_P(P,b) = \frac{\partial g}{\partial P} > 0.$$

If you don't believe these inequalities, you should not be a student in economics! We do not know the signs of the partial derivatives of f and g with respect to the exogenous variables, but let's assume

$$f_a(P,a) = \frac{\partial f}{\partial a} > 0, \quad g_b(P,b) = \frac{\partial g}{\partial b} > 0.$$

Can you "tell an economic story" which would justify these?

Now we must get serious about the problem. It helps to remember where our true interests lie: it is not the values of P^* and Q^* which beg for attention so much as the effect of shifts in a and b on these values. Thus we only care about partial derivatives. The unknown value of P^* must satisfy the static equilibrium requirement:

$$Q^D = Q^S \text{ or } Q^D - Q^S = f(P^*,a) - g(P^*,b) = 0.$$

If we want to know the effect of a shift in a on P^*, the obvious thing to do is the differentiate this equation with respect to a, reminding ourselves loudly to remember that $P^*$is a function of a:

$$f_P(P^*,a)\frac{\partial P^*}{\partial a} + f_a(P^*,a) - g_P(P^*,b)\frac{\partial P^*}{\partial a} = 0.$$

We have used the chain rule from vector calculus. It is a fact that mistakes with the chain rule are the most common cause of errors in working with comparative statics. No resolution will be more valuable to you than to resolve right now that you will do everything in your power to utterly master the chain rule. Be always alert when the chain rule is in use and focus all your concentration to be careful. So examine the equation above and be absolutely certain you are comfortable with the calculation.

Amazingly, obtaining the desired comparative static is now simple. We just solve this last equation for it:

$$\frac{\partial P^*}{\partial a} = \frac{-f_a(P^*,a)}{f_P(P^*,a) - g_P(P^*,b)} > 0.$$

Does this conclusion agree with your economic intuition? If the taste (i.e. desire) for the good increases, will the equilibrium price rise? You should not delay even a minute to practice this calculation by finding the comparative static showing the effect on P^* of an increase in b.

But we are not done. There are two other comparative statics: the effect on Q^* of increases in a or b. How will we find these? Well, once we know P^* (don't start yelling that we don't; just pretend that we do), we can get Q^* by plugging P^* into either the demand curve or the supply curve. After all they are equal here. If we plug into the supply curve, we get

$$Q^* = g(P^*,b),$$

and differentiating with respect to a we get

$$\frac{\partial Q^*}{\partial a} = g_P(P^*,b)\frac{\partial P^*}{\partial a} = \frac{-f_a(P^*,a)g_p(P^*,b)}{f_P(P^*,a) - g_P(P^*,b)},$$

where we used the already known comparative static for P^*. You should plug into the demand curve and repeat this work (a little more complicated) to see that you get the same answer. You should also find the fourth comparative static: for the effect on Q^* of an increase in b.

It is worth pausing for a moment to think about these last calculations purely mathematically. We have used basically implicit differentiation to turn the out-of-reach algebra problem (solve for P^*) into a simple linear equation for the comparative static. Notice that the solution of our problem required dividing; the nature of the denominator in such expressions is always significant. The denominator here is $f_P - g_P$. Reflect on this a moment. If we go back to our static equilibrium equation $Q^D - Q^S = f - g = 0$, you see that this denominator is the derivative of this expression with respect to the endogenous variable of interest. In this case, our fundamental economic understanding guarantees us that we have not divided by zero (always a foolish thing to do!). This division worry will come back to turn most of our dreams in this course into nightmares. The road ahead is not so smooth. Knowing that a denominator is not zero so catastrophe does not strike, and even knowing the sign of the denominator (after all we want to know the sign of our comparative static), will be a major worry for us.

1.2 Exercises

1. For the linear model (equations (1.1)-(1.2)), verify the partial derivatives for the comparative statics given in the text for Q^*. Do you see why the signs of these partials are not so obvious just from looking at the formula for Q^*.
2. Look again at the nonlinear model in equations (1.3)- (1.4).
   (a) Find the comparative static showing the response of P^* to an increase in b.
   (b) Now find the comparative static showing the response of Q^* to an increase in b.
3. Let's consider the effect of a tax on the static equilibrium model. Now the price paid by consumers for our one good is $P = c + t$, where c is the price the seller receives and t is the tax. Now we have $Q^D = f(P), \quad Q^S = g(c)$, where we only consider t as an exogenous variable.
   (a) Find the comparative statics results showing the dependence of P^*, c^*, and Q^* on an increase in the tax t.
   (b) Can you conclude a sign for these comparative statics. Do these results satisfy your intuition?

1.3 Static Equilibrium: A Two-Good World

We will be less talkative in this section, leaving it to you to editorialize on most aspects. We consider the static equilibrium problem in a world with two goods A and B. We express the demand and supply functions for each good as

$$\begin{aligned} Q_A^D &= f(P_A, P_B, a), \quad Q_A^S = g(P_A, b), \\ Q_B^D &= F(P_A, P_B, c), \quad Q_B^S = G(P_B, d), \end{aligned} \tag{1.5}$$

so that the demand for each good depends, as expected, on the prices of both goods, but the supply of a good depends only on its own price; demand and supply functions all depend on exogenous variables denoted by a, b, c, d. What knowledge do economics principles give us about the functions in (1.5)? Certainly,

$$f_{P_A} < 0, \quad g_{P_A} > 0, \quad F_{P_B} < 0, \quad G_{P_B} > 0.$$

The signs of certain other partials, e.g. f_{P_B}, depend on whether or not A and B are complements, substitutes, or unrelated. We would expect f_{P_B} to be positive if A and B are substitutes, negative if complements, zero if unrelated. Similar statements can be made about F_{P_A}. You can sign the derivatives with respect to the exogenous variables by telling appropriate economic fables.

Let's continue with the mathematics. Static equilibrium conditions are

$$Q_A^D - Q_A^S = f(P_A^*, P_B^*, a) - g(P_A^*, b) = 0,$$
$$Q_B^D - Q_B^S = F(P_A^*, P_B^*, c) - G(P_B^*, d) = 0,$$

Again, we are not able to solve for equilibrium prices and quantities unless we know much more specific information about the supply and demand functions, but we can use the method of implicit differentiation. For example, if we want to know comparative statics with respect to the exogenous variable a, we differentiate both equations with respect to a to get:

$$f_{P_A}\frac{\partial P_A^*}{\partial a} + f_{P_B}\frac{\partial P_B^*}{\partial a} + f_a - g_{P_A}\frac{\partial P_A^*}{\partial a} = 0,$$

and

$$F_{P_A}\frac{\partial P_A^*}{\partial a} + F_{P_B}\frac{\partial P_B^*}{\partial a} - G_{P_B}\frac{\partial P_B^*}{\partial a} = 0.$$

Digest the fact that a change in the value of a directly affects the demand for A and thus the price of A causing an indirect effect on the supply of A, but it also affects the price of B indirectly because a change in the price of A will have repercussions on the demand for good B and consequent effects on the price of B. So we don't do something foolish like believing that the left side of the equilibrium equation for good B is constant when a changes using the argument that no a appears explicitly in this equation.

Look carefully at these last two equations. Think of them as two equations for the two desired comparative statics $\dfrac{\partial P_A^*}{\partial a}$ and $\dfrac{\partial P_B^*}{\partial a}$. These are simple linear equations. Before solving them, let's inventory the different ways we know how to solve such systems.

1. We know a method learned in our first algebra course. It consists of either solving the first equation for either unknown in terms of the other and substituting the result into the second equation (I call this the substitution method) or multiplying each equation by an appropriate factor and subtracting the two resulting equations to eliminate one of the unknowns (I call this the addition- subtraction method). There is no reason why we should not use that method here; no method is simpler.

2. We can solve using determinants; this method is called Cramer's rule. You may have learned it in a school algebra class or again in linear algebra. Probably a majority of books written by economists would use this method.

3. We can rewrite the equation as a matrix equation involving a 2×2 matrix and solve by multiplying both sides of this matrix equation by the inverse of the matrix. Since the four entries in the coefficient matrix are partial derivatives rather than numbers, it

would not be wise to try to invert using Gaussian elimination on the augmented matrix so we would be better advised to invert using cofactors; this method should have been learned in linear algebra if not earlier.

You should solve the system by your method of choice (actually it would be smart to do it all three ways for review; the last two methods will be our method of choice when we deal with problems involving more equations later). The result is

$$\frac{\partial P_A^*}{\partial a} = \frac{-f_a(F_{P_B} - G_{P_B})}{\Delta}, \quad \frac{\partial P_B^*}{\partial a} = \frac{f_a F_{P_A}}{\Delta},$$

where

$$\Delta = (f_{P_A} - g_{P_A})(F_{P_B} - G_{P_B}) - f_{P_B} F_{P_A}.$$

It is easy to attach a sign to the numerators of these two comparative statics (reason it out!). The first numerator unambiguously has the same sign as f_a; the sign of the second numerator depends on whether or not A and B are substitutes, complements, or unrelated. The denominator presents no problem if A and B are unrelated, for then Δ is definitely positive. However, if A and B are complements or substitutes, Δ has the form of a positive number minus a positive number and it looks like a tug of war: who will win? Economists generally believe the first product always wins so that Δ is unambiguously positive in every case. The usual article of faith states: the own price effects *always* dominate the cross price effects. Have you been properly indoctrinated by your economics professors so that you accept this article of faith? If so, then you will have no difficulty signing these comparative statics. After doing so, you should not rest until you have compared the results with economic principles and either been comforted or distressed by the results. If you are comforted, then your enthusiasm for the article of faith will be strengthened!

You will probably be helped by studying another static equilibrium problem in economics which is mathematically similar to the one just analyzed. It can be found on pp. 96-101 (pp. 119-125, 1st ed.) in the textbook by Hands and also leads to two equilibrium equations so there are two endogenous variables. It is a macroeconomics problem which combines the goods market and the money market to form a Keynesian model of the economy.

1.4 Exercises

1. For the model in (1.5), find the comparative statics which relate the impact of an increase in a on the equilibrium quantities Q_A^*, Q_B^*. Can you sign these?

2. For the model in (1.5), find the comparative statics which reflect the impact on both equilibrium prices and both equilibrium quantities of an increase in b. Attempt to sign them.

1.5 Profit Maximization of a Firm

Hands, Sect. 3.2 and pp. 220-222. (pp. 271-272 1st ed.)

A firm engaged in the manufacture of some good employs capital and labor. Resources need to be divided between capital and labor so as to maximize profits. This decision is made depending on the relative costs of capital and labor as well as the effectiveness of each in the production function of the firm. We shall take the position here that the firm is purely competitive and its output does not effect prices, so the price of its product as well as the cost of capital and labor are given.

Let $f(K, L)$ be the production function of our firm and let p be the unit price of its product. Then $pf(K, L)$ is the revenue of the firm. The profit of the firm, excluding fixed costs, is then

$$\pi(K, L) = pf(K, L) - wL - \nu K,$$

where w is the wage rate of labor and ν is the unit cost of capital. Of course, the true profit of the firm would be obtained by subtracting fixed costs, but the maximization problem is the same.

In the short term, capital may be considered constant and the problem has only one endogenous variable: L. You can apply elementary one-variable calculus and differentiate the profit function with respect to L to find the first order condition that the optimal value L^* of labor must satisfy:

$$pf_L(K, L^*) - w = 0.$$

To find the impact on L^* of an increase in the exogenous variable w, we would use the now familiar step of implicit differentiation to get

$$pf_{LL}(K, L^*)\frac{\partial L^*}{\partial w} - 1 = 0.$$

We now arrive at the moment of truth which has been anticipated earlier. To solve for the desired comparative static, we must divide by $pf_{LL}(K, L^*)$. We may certainly assume that the unit price $p > 0$. But what about the second derivative $f_{LL}(K, L^*)$? Since we are assuming that the firm is maximizing profits, we definitely know that this second derivative is not positive, for then the second derivative test implies a minimum. A negative second derivative is a sufficient but not necessary condition for a maximum. It just might happen that the second derivative is zero at the maximum (as for the example $f(x) = 9 - x^4$). So in order to make progress, we must *assume* that the sufficient second order condition holds. Then we get

$$\frac{\partial L^*}{\partial w} = \frac{1}{pf_{LL}(K, L^*)}.$$

Note that the assumed second order condition not only allows division but guarantees that this comparative static is negative. Thus an increase in wage rate will lead to a decrease in the amount of labor employed by the firm.

You should meditate some more about this second order condition. Do you see in this simple case that the second order condition can be justified by purely economics principles? One would expect that the production function exhibits diminishing returns to labor. This fortuitous nonmathematical defense of the second order condition is not usually available in more realistic problems.

Let's turn our attention to the more interesting exploration of the long term optimization problem where K cannot be assumed constant and we have both K and L as endogenous variables. Notice that increasing the number of endogenous variables is an added burden on the model; the model is expected to explain more things. We are basically asking the model to let us tell it less while it tells us more.

As usual, let's gather together all the information we learn from economics principles about the production function $f(K, L)$. We expect

$$f_K > 0, \quad f_{KK} < 0, \quad f_L > 0, \quad f_{LL} < 0,$$

which state that production increases with more capital or labor, but that the production function exhibits diminishing returns to capital or to labor. What can we say about f_{KL}? From vector calculus we know that it is the same as f_{LK}, at least if f has continuous second partial derivatives (which we expect is true of any reasonable production function). But is this mixed second partial positive, negative, or zero? The answer is that any of these is conceivable. If this mixed partial is positive, then capital and labor are complementary inputs, if negative, they are substitutes, if zero, they are unrelated. Think about this until you are comfortable! We apply the necessary first order conditions from multivariate calculus. The optimal choice K^* and L^* of capital and labor must satisfy the equations:

$$\pi_K(K^*, L^*) = pf_K(K^*, L^*) - \nu = 0,$$
$$\pi_L(K^*, L^*) = pf_L(K^*, L^*) - w = 0.$$

If we knew a specific production function (you will have the chance to deal with such specific cases in the homework problems), we could now solve these two equations for the optimal values of K^* and L^*. Here we must be more devious. Recall the fundamental point that we are not so much interested in finding K^* and L^* as we are in answering the comparative static question: what will be the direction of shift in K^* of L^* if either the wage rate w or the unit cost of capital ν should experience a marginal increase? Thus what we really would like to explore are the comparative statics $\dfrac{\partial K^*}{\partial w}$ and $\dfrac{\partial K^*}{\partial \nu}$ or alternatively those regarding labor L^*. To do this we will resort to the now familiar method of implicit differentiation. For example, to obtain the comparative static results for the wage rate w, we differentiate both equations with respect to w, keeping in mind that K^* and L^* both depend on w:

$$pf_{KK}(K^*, L^*)\frac{\partial K^*}{\partial w} + pf_{KL}(K^*, L^*)\frac{\partial L^*}{\partial w} = 0,$$
$$pf_{LK}(K^*, L^*)\frac{\partial K^*}{\partial w} + pf_{LL}(K^*, L^*)\frac{\partial L^*}{\partial w} - 1 = 0.$$

As expected, these are linear equations to determine the two comparative statics with respect to w. We can add the "1" to both sides of the second equation and then solve by Cramer's rule (or one of the other two methods) to get

$$\frac{\partial K^*}{\partial w} = \frac{-f_{KL}(K^*, L^*)}{\Delta},$$

$$\frac{\partial L^*}{\partial w} = \frac{f_{KK}(K^*, L^*)}{\Delta},$$

where $\Delta = p(f_{KK}(K^*, L^*)f_{LL}(K^*, L^*) - (f_{KL}(K^*, L^*))^2)$.

We know that the numerator of the second comparative static above is negative (diminishing returns of capital). What about the denominator? Here we have arrived at a serious juncture. What would wishful thinking have us hope? Surely your economic expectation is that this comparative static turns out negative: if the cost of labor increases, the firm will employ less labor. So we would pray to the god of mathematical economics and ask that $\Delta > 0$. Will we get what we wish? Well, it's close. Recall the second derivative test from multivariable calculus. (This is the optimal time to dig out your multivariable calculus book and review the second derivative test for a max or a min for functions of two variables!) If we have a critical point (x^*, y^*) of a function $F(x, y)$ of two variables, then the sufficient condition for a maximum is that $D = F_{xx}F_{yy} - F_{xy}^2 > 0$ while $F_{xx} < 0$ at the critical point. Most books call D the discriminant. If $D > 0$ while $F_{xx} > 0$ at the critical point, then the critical point is a minimum. If $D < 0$ at the critical point, then the critical point is a saddle point. Notice that in our case

$$\pi_{KK} = pf_{KK}, \quad \pi_{LL} = pf_{LL}, \quad \pi_{KL} = pf_{KL}.$$

So Δ has the same sign as D since $D = p\Delta$. Since our firm is a profit maximizer by assumption, then clearly Δ cannot be negative (i.e., we are not at a saddle point!). Since $\pi_{KK} = pf_{KK} < 0$, then $\Delta > 0$ would be a sufficient condition for a maximum. If only this sufficient condition were necessary! But maybe $\Delta = 0$; this may happen. We proceed in the time honored tradition of economics by *assuming* that the sufficient condition holds at the maximum: $\Delta > 0$. (We cannot prove this mathematically; there is no obvious argument from economics principles.) Then we obtain the conclusion that $\dfrac{\partial L^*}{\partial w} < 0$. What about the other comparative static with respect to w? Well, if we continue to assume that the sufficient condition holds, the denominator $\Delta > 0$. The sign of the numerator depends on whether or not capital and labor are substitutes, complements, or unrelated in the production function. If you think capital and labor are usually substitutes, then the comparative static $\dfrac{\partial K^*}{\partial w} > 0$. This also agrees with economic intuition: if the unit cost of labor increases, and capital and labor are substitutes, we will adjust by substituting capital for labor: employ less labor and more capital.

For practice, you should explore the comparative statics with respect to ν and p. Assuming the sufficient conditions for a maximum hold, you should encounter no surprises.

1.6 Exercises

1. Fill in the mathematical details leading to the comparative statics in the previous section with respect to w.

2. Find the comparative statics with respect to ν and p for the problem of the previous section. Be sure to consider and comment on their signs.

3. Find the short-run perfectly competitive firm's maximizing level of labor L^* for the explicit production function $f(L) = 4\sqrt{L}$. What are the explicit comparative statics in this case?

4. Suppose that the production function of a perfectly competitive firm (in the long run) has the explicit form
$$f(K, L) = \ln(1 + L) + \ln(1 + K).$$
   (a) Find explicit formulas for the maximizing values K^* and L^*. These are called the input demand functions. They should be functions of all three exogenous variables w, ν, p.
   (b) Are the input demand functions homogeneous of any degree?
   (Note: A function h of several variables, say w, ν, p is called homogeneous of degree r if $h(\lambda w, \lambda\nu, \lambda p) = \lambda^r f(w, \nu, p)$ for all values of w, ν, p and any $\lambda > 0$. We shall see the significance of homogeneous functions in the next chapter.)
   (c) Find all six comparative statics for this explicit problem.
   (d) Find the maximum profit π^*.
   (e) Verify that for this profit function
$$\frac{\partial \pi^*}{\partial w} = -L^*, \qquad \frac{\partial \pi^*}{\partial \nu} = -K^*.$$

5. For the general perfectly competitive firm considered in the text, if K^*, L^* are the optimal values of capital and labor, then the value $y^* = f(K^*, L^*)$ is called the supply function and $\pi^* = py^* - wL^* - \nu K^*$ is called the profit function; K^*, L^* are called the input demand functions.
   (a) Show that, even in this general case,
$$\frac{\partial \pi^*}{\partial w} = -L^*, \qquad \frac{\partial \pi^*}{\partial \nu} = -K^*, \qquad \frac{\partial \pi^*}{\partial p} = y^*.$$
   This result is called Hotelling's lemma.
   (b) Show that
$$\frac{\partial L^*}{\partial \nu} = \frac{\partial K^*}{\partial w}.$$

Chapter 2

Comparative Statics in Many Variables

In this chapter, our purpose is to consider problems with many endogenous variables. For example, we shall consider a static equilibrium problem for n interrelated goods. In the previous chapter we considered a model for 2 goods. You will recall that the two good model required the consideration of 2×2 matrices and determinants. Since we will now be forced to entertain much larger matrices and determinants, it is clear that we must marshal some serious mathematical tools. It will not do to evaluate such large determinants by brute force. In order to make the transition as easy as possible we will begin by reconsideration of the problem (1.3), (1.4) involving only one good.

2.1 A Recapitulation

Hands, Sect. 6.3.

Recall that when we first considered the problem (1.3), (1.4), we took advantage of the static equilibrium requirement $Q^D = Q^S$ to reduce the problem to only one equation $f(P^*, a) - g(P^*, b) = 0$. This caused a certain uneven attitude regarding equilibrium price and quantity. Our one equation eliminated the equilibrium quantity. Of course there was no serious loss because we could deal successfully with comparative statics for Q^* after dealing with those for P^*. In our reconsideration of the problem, we shall refrain from eliminating the equilibrium quantity and keep both equations

$$Q^* - f(P^*, a) = 0,$$

$$Q^* - g(P^*, b) = 0.$$

We shall view these as two equations for the two endogenous variables P^*, Q^*. Note that if we define the vector function

$$F(P^*, Q^*, a, b) = \begin{bmatrix} Q^* - f(P^*, a) \\ Q^* - g(P^*, b) \end{bmatrix},$$

our equations can be described by the single vector equation

$$F(P^*, Q^*, a, b) = 0, \tag{2.1}$$

where this last zero is the two dimensional zero vector. This function F has a 4 dimensional domain and a 2 dimensional range. If we are able to solve (2.1) for P^*, Q^* in terms of a, b, we could write the solution in the form

$$\begin{bmatrix} P^* \\ Q^* \end{bmatrix} = H(a,b) = \begin{bmatrix} h_1(a,b) \\ h_2(a,b) \end{bmatrix}.$$

Note that H has a 2 dimensional domain and a 2 dimensional range.

You will recall from vector calculus the definition of the derivative of such functions as F and H. Each derivative is a matrix: the derivative matrix for F is a 2×4 matrix; the derivative matrix for H is a 2×2 matrix. In fact,

$$F'(P^*, Q^*, a, b) = \begin{bmatrix} -f_P(P^*, a) & 1 & -f_a(P^*, a) & 0 \\ \\ -g_P(P^*, b) & 1 & 0 & -g_b(P^*, b) \end{bmatrix}, \tag{2.2}$$

where the top row is the gradient vector of the first component of F (the four partial derivatives of this component) and the second row is the gradient vector of the second component of F (the four partial derivatives of this component). Similarly the derivative of H is

$$H'(a,b) = \begin{bmatrix} \frac{\partial h_1}{\partial a} & \frac{\partial h_1}{\partial b} \\ \frac{\partial h_2}{\partial a} & \frac{\partial h_2}{\partial b} \end{bmatrix}.$$

It does not matter in what order the variables are listed, but it is *essential* to observe one particular order consistently. In particular for F, it is desirable to always list the endogenous variables first, then the exogenous ones.

The left 2×2 submatrix of F' above which involve only partial derivatives with respect to the endogenous variables is denoted by $F_{(P,Q)}$ and the right 2×2 submatrix involving only partial derivatives with respect to the exogenous variables is denoted by $F_{(a,b)}$. This notation is consistent with our subscript notation for partial derivatives and reminds us that we have differentiated F with respect to some of the variables, pretending the other variables are constant.

Recall that all these derivative matrices are often referred to as Jacobian matrices. Thus a Jacobian matrix is just a matrix whose entries are certain first partial derivatives. Later, we shall use the term Hessian to refer to matrices whose entries are second partial derivatives.

With respect to our vector equation (2.1), there are two questions of importance, one theoretical, the other practical. The theoretical question concerns the existence of a solution for the endogenous variables. One naturally is worried that a tragedy might occur and there be no solutions. The reaction to such an event (this event happens rather often in real life) is to go back to the economics and figure out where the mistake was made because the only conclusion is that economics principles were somehow violated and the equations do not really describe economic reality. The practical question is how one finds H' because you will have noticed that the entries in H' are precisely the four comparative statics we want to find.

A famous theorem, the implicit function theorem, successfully deals simultaneously with these two questions. It is in two parts: the first part tells us conditions which guarantee that, at least theoretically, the vector equation can be solved; the second part gives a formula for H'. Here it is.

Theorem 1 *Suppose that F is a function with an $n + m$ dimensional domain and an m dimensional range. We write the function in the form $F(y, x)$, where x is an n-vector and*

y is an m-vector; y represents the m endogenous variables and x the n exogenous variables. Suppose that all partial derivatives which are entries in the derivative matrix F' are continuous in some neighborhood of the point (y^*, x^*) and $F(y^*, x^*) = 0$ (i.e. (y^*, x^*) is an equilibrium point). Suppose further that the $m \times m$ partial derivative matrix $F_y(y^*, x^*)$ is invertible. Then $F(y, x) = 0$ can be solved (theoretically) for $y = H(x)$ in terms of x for all points x in some neighborhood of x^*. (This is the theoretical part of the theorem.) Furthermore, each partial derivative in the $m \times n$ matrix $H'(x)$ exists and is continuous in this neighborhood of x^* and $H'(x)$ may be found with the formula

$$F_y(y, x)H'(x) = -F_x(y, x)$$

where it is required to substitute $H(x)$ for y. The matrix $H'(x)$ contains all the comparative statics for the problem. Note that F_y is the left $m \times m$ portion of $F'(y, x)$ and F_x is the right $m \times n$ portion of $F'(y, x)$. Also check that these sizes are compatible for the multiplication in the formula for $H'(x)$.

Of course, textbook statements of the above theorem are much shorter because all the editorial remarks are omitted. The word neighborhood in the above statement means exactly what you would expect: for example a neighborhood of (y^*, x^*) consists of this point and all points "close to" it. The size of the neighborhood is unimportant: think of it as some spherical "ball" centered at (y^*, x^*). The actual size of the positive radius is immaterial.

Let us now solve the problem (1.3), (1.4) using the implicit function theorem. The hypotheses for the theorem are satisfied by assumption: we assume that our functions have continuous derivatives (it seems unrealistic that such functions would behave badly) and we are assuming that the market is in static equilibrium right now, which means that the current price P^* and quantity Q^* are truly values which place the market in static equilibrium. Thus the equations have a solution $H(a, b)$ (the notation is $y^* = (P^*, Q^*)$ and $x^* = (a, b)$). We can then use the formula for $H'(x)$, which is actually a streamlined version of implicit differentiation. In our case, we have

$$F_y = F_{(P,Q)} = \begin{bmatrix} -f_P(P^*, a) & 1 \\ -g_P(P^*, b) & 1 \end{bmatrix}, \quad F_x = F_{(a,b)} = \begin{bmatrix} -f_a(P^*, a) & 0 \\ 0 & -g_b(P^*, b) \end{bmatrix}.$$

The 2×2 matrix $F_{(P,Q)}$ is easily inverted and the formula gives us

$$\begin{aligned} H'(a, b) &= -\frac{1}{\Delta} \begin{bmatrix} 1 & -1 \\ g_P(P^*, b) & -f_P(P^*, a) \end{bmatrix} \begin{bmatrix} -f_a(P^*, a) & 0 \\ 0 & -g_b(P^*, b) \end{bmatrix} \\ &= -\frac{1}{\Delta} \begin{bmatrix} -f_a(P^*, a) & g_b(P^*, b) \\ -g_P(P^*, b) f_a(P^*, a) & g_b(P^*, b) f_P(P^*, a) \end{bmatrix}, \end{aligned}$$

where $\Delta = g_P(P^*, b) - f_P(P^*, a)$. You can compare to see that these answers agree with those for the comparative statics found before.

Do not misunderstand. We are not claiming that what we just did is superior to what we did before. In fact, our earlier work on this problem was simpler. It is always simpler to plant an azalea with a shovel than with a bulldozer. Heavy equipment is unwieldy and unnecessary for a small task. The implicit function theorem is heavy equipment. So we use it for the heavy jobs which we will soon face. For a first exposure, we wanted to use it on a simple problem which we had already solved.

It may be helpful if we reiterate in outline form the steps in applying the implicit function theorem.

1. Write the equations to be solved in the form $F(y, x) = 0$ where y is a vector containing all the endogenous variables of the problems (these are the ones for which you solve) and x is a vector containing all the other (exogenous) variables. Check to be sure you have the same number of equations as you do endogenous variables.

2. Calculate the derivative matrix $F'(y, x)$. If there are m endogenous variables, the left $m \times m$ submatrix will be $F_y(y, x)$ and the remaining right submatrix will be $F_x(y, x)$.

3. Check to be sure that F_y is invertible. In most economics problems where the matrices are not small, this invertibility cannot be verified by brute force (say by computing the determinant) but must be implied by some economic assumptions together with mathematical tools. Then you know that the equations can be solved (at least theoretically) for the endogenous variables y in terms of the exogenous variables x.

4. Use the formula $F_y H'(x) = -F_x$ to find $H'(x)$. If the matrix F_y is small, its inverse can be found using determinants (cofactors). $H'(x)$ is the matrix which contains all the comparative statics.

5. If you are interested in only one comparative static, then you can use Cramer's rule. If there are zeros in the matrices, then only relatively few determinants may need computing. It would be a shame (and a disgrace) to spend lots of time computing unnecessary cofactors calculating the full inverse since surely no one enjoys such work.

2.2 Exercises

1. Reconsider the long run profit maximization of a competitive firm from Section 1.5 (i.e., the first order conditions in the middle of page 11). Apply the implicit function theorem to re-derive the comparative statics.

2. Consider again the static equilibrium model (1.5) for 2 goods in Section 1.3. Do not eliminate the equilibrium quantities, so you will have four equations involving the four endogenous variables for the two equilibrium prices and the two equilibrium quantities. Use the implicit function theorem to re-derive the comparative statics for good A relative to the exogenous variable a. Take advantage of the final advice of this section.

2.3 The n Good Equilibrium Model

Hands, Sect. 7.1.

The model we now discuss is a key problem for us to understand. It is an important and nontrivial economics problem, which makes serious use of some of the most important mathematical tools we want to master. In addition to the implicit function theorem, we will need an important matrix theorem, which we will introduce when needed. We prefer to dive right in on the economics problem and encounter the mathematics in the context of the application.

The problem is a general equilibrium model involving n goods. These goods, numbered 1 through n, have positive prices designated by p_i. The system is in static equilibrium if the demand Q_i^D of the ith good is equal to its supply Q_i^S. If demand for the ith good exceeds supply, then $z^{(i)} = Q_i^D - Q_i^S$ represents the excess demand. We interpret $z^{(i)} < 0$ to mean that supply of the ith good exceeds demand. We assume that the value of $z^{(i)}$ depends only on the current prices p_j of all n goods and one exogenous variable α. So we write

$$z^{(i)} = z^{(i)}(p, \alpha) = z^{(i)}(p_1, p_2, \cdots, p_n, \alpha).$$

The system will be in equilibrium precisely when the price vector p takes values which imply that each $z^{(i)}(p, \alpha) = 0$. We will assume that the excess demand for each good depends only on the relative values of the prices, not on the absolute level of prices. Thus, if all prices doubled suddenly without change in α, neither consumers or suppliers would see a reason to shift demand or supply from one of the goods to another. I suppose what this means is that if inflation is uniform, no changes in demand or supply will occur. Perhaps you can give a better justification of this assumption. Mathematically, we have

$$z^{(i)}(p, \alpha) = z^{(i)}(\lambda p, \alpha), \tag{2.3}$$

for each $i = 1, 2, \cdots, n$. In particular, taking $\lambda = \frac{1}{p_n}$, we can replace the functions $z^i(p, \alpha)$ with $z^i(r, \alpha)$, where $r_i = p_i/p_n$ is the ratio of the ith price to the nth price p_n. Of course, we then have $r_n = 1$. We thus convert the problem involving n endogenous prices to a problem involving $n-1$ endogenous ratios, since the ratio $r_n = 1$ is constant and no longer to be found. It is common in economics to think of these ratios as prices, expressed not in terms of currency values (like dollars or German marks, for example) but in terms of the cost of one particular good, usually called the "numeraire" good. I want to emphasize that these new endogenous variables are *ratios* of prices, rather than prices, so when we get comparative statics which show the direction of change of, say, r_2 as α increases, we do not really know whether the price of good 2 rises or falls, only the direction of change in this ratio. However, the conclusion we do get may be more meaningful than a comparative static for p_2 because if all prices were to double, we do not gain any real knowledge. Note that (2.3) is a mathematical statement that the excess demand functions are homogeneous of degree zero in the endogenous variables. You should review the definition of homogeneous which appeared in problem 4(b) of Exercises 1.6. If you insert a factor of $\lambda^0 = 1$ on the left side of (2.3), you are looking at the definition of degree 0 in the endogenous variables. Note that the exogenous variable α is not multiplied by λ.

So now we have changed our problem from one with n endogenous variables $p_1, p_2, \cdots, p_n$ to one with $n-1$ endogenous variables $r_1, r_2, \cdots, r_{n-1}$, since $r_n = 1$ is now a known constant. Mathematically, this situation is troubling, for we now have n equilibrium equations ($z^{(i)}(r) =$

0 for $i = 1, 2, \cdots, n$) and only $n-1$ prices $(r_1, r_2, \cdots, r_{n-1})$ to specify. Fortunately there is a way to eliminate an equation; it is called Walras' law and is an axiom of economics. Walras' law states

$$\sum_{i=1}^{n} p_i z^{(i)}(p, \alpha) = 0,$$

whether or not the system is in equilibrium; it is trivially true if the system is in equilibrium. Economic arguments may be advanced that Walras' law is plausible. Read pages 190-192 in Hands' book for one such argument. Mathematically, Walras' law allows us to ignore the equation $z^{(n)}(p, \alpha) = 0$; if the first $n-1$ markets are in equilibrium, the last one has no choice.

To make the problem interesting (and worthwhile), we assume that an increase in α indicates a shift of demand from good n to the other $n-1$ goods. Thus, we assume that

$$z_\alpha^{(n)}(p, \alpha) < 0, \quad z_\alpha^{(i)}(p, \alpha) \geq 0 \text{ for } i = 1, 2 \cdots, n-1.$$

We assume that the economy is in equilibrium for the current value $\alpha = \alpha^*$; i.e., that the equilibrium equations $z^{(i)}(r^*, \alpha^*) = 0$ are all satisfied for some ratio vector $r^* = (r_1^*, r_2^*, \cdots, r_{n-1}^*, 1)$.

Our goal is to explore the comparative statics for the price ratios, i.e. we want to determine the signs of $\dfrac{\partial r_i^*}{\partial \alpha}$ for each $i = 1, 2, \cdots, n-1$. In particular, we might wonder if all equilibrium ratios of the first $n-1$ goods increase when there is a shift away from demand for good n. If we can use the implicit function theorem, we know how to get the comparative statics: the column vector of these partial derivatives is given by the expression

$$-(F_y)^{-1} F_x,$$

where the Jacobian matrix F_y, is of size $n-1$ with the element in the ith row and jth column being

$$z_j^{(i)}(r^*),$$

($z_j^{(i)}$ denotes the partial derivative of $z^{(i)}$ with respect to r_j). So the ith row of F_y is obtained from the $n-1$ partial derivatives of the left side of the ith equilibrium equation with respect to the $n-1$ price ratios, and F_x is the column vector whose ith row is just the partial derivative of the left side of the ith equilibrium equation with respect to α. Note that the matrix F_y is precisely the matrix which needs to be invertible in order to apply the implicit function theorem. The jth entry in F_x is $z_\alpha^{(j)}$, so all entries in F_x are non-negative by assumption.

Since an increase in α shifts demand from good n to all the other goods, we might suspect that all the comparative statics are positive, or at least non-negative. Without further assumptions, it seems hopeless to know if F_y is invertible, or the signs of the entries in $(F_y)^{-1}$. Clearly, all the comparative statics will be non-negative if the elements of $-(F_y)^{-1}$ are all non-negative. Since $-(F_y)^{-1} = (-F_y)^{-1}$, we are interested in conditions that will guarantee that F_y is invertible and that the entries in $(-F_y)^{-1}$ are non-negative.

Now we need a powerful theorem from linear algebra. This theorem is too specialized and advanced to be a part of a first linear algebra course. So we interrupt our flow of thought to state this theorem.

Theorem 2 *Suppose B is an $n \times n$ matrix whose off-diagonal elements are all non-positive, and suppose that there is some n-vector v all of whose entries are nonnegative for which all entries in Bv are positive. Then B is invertible and every entry in B^{-1} is non-negative.*

Clearly this theorem is an answer to an economist's prayer. It delivers almost everything we desire, *if* we can verify the hypotheses.

We want to apply this theorem to the matrix $B = -F_y$. So first we examine the off-diagonal elements of F_y. They are just the partial derivatives of the excess demand functions $z^{(i)}$ with respect to all the prices p_j for $j \neq i$. What is the economic meaning of these partials being non-negative? Just this: if an increase in the price of one good causes the excess demand of all other goods to be non-decreasing, the goods are all substitutes for each other, or at least unrelated. Thus if we assume the goods are gross substitutes, or unrelated, then all the off-diagonal entries in F_y will be non-negative. Hence the off-diagonal entries in $B = -F_y$ will be non-positive and the first hypothesis of Theorem 2 will be satisfied.

How about the vector v that we need? Will we need to accept more assumptions about the system? Since we are assuming that $z^{(i)}(r, \alpha) = z^{(i)}(\lambda r, \alpha)$ for each i, we can differentiate both sides with respect to λ and then substitute $\lambda = 1$ to get (be careful with the chain rule!)

$$0 = \sum_{j=1}^{n} z_j^{(i)}(r, \alpha) r_j, \tag{2.4}$$

for each $i = 1, 2, \cdots, n$. Therefore, rearranging gives

$$\sum_{j=1}^{n-1} (-1) z_j^{(i)}(r, \alpha) r_j = z_n^{(i)}(r, \alpha),$$

for $i = 1, 2, \cdots, n$, since $r_n = 1$. In particular, we have

$$\sum_{j=1}^{n-1} (-1) z_j^{(i)}(r^*, \alpha^*) r_j^* = z_n^{(i)}(r^*, \alpha^*).$$

Our previous assumption only guarantees that the right side of this equation is non-negative and we need it to be positive in order to use our theorem. Thus, we need to strengthen our assumption so that $z_n^i(r^*) > 0$, i.e., good n and good i are substitutes for all $i = 1, 2, \cdots, n-1$. For $i = 1, 2, \cdots, n-1$, the left hand sides are just the elements in $(-F_y)w$, where w is the $n-1$ dimensional vector $w = (r_1^*, r_2^*, \cdots, r_{n-1}^*)$. Using this w as v, the second hypothesis of Theorem 2 is satisfied. (The technique we just used is very valuable and we will return to it later when we discuss Euler's theorem and homogeneous functions.)

Summarizing, the assumption that goods n and i are substitutes for each $i = 1, 2, \cdots, n-1$, and all other pairs are substitutes or unrelated, assures us that both hypotheses of Theorem 2 are satisfied and we may conclude that all elements of $(-F_y)^{-1}$ are non-negative. Hence all comparative statics $\dfrac{\partial r_i^*}{\partial \alpha}$ for each $i = 1, 2, \cdots, n-1$ are at least non-negative. We can say a little more. Since no column of $(-F_y)^{-1}$ can be all zeros (in such a case the determinant would be zero and so would have no inverse), then a little thought will convince you that at least one comparative static will be positive if at least one $z_\alpha^i > 0$. If we assume that $z_\alpha^i > 0$ for all $i = 1, 2, \cdots, n-1$, i.e. increasing α increases demand for all goods except good n, then every comparative static will be positive.

Regarding Theorem 2, if all off-diagonal elements in the matrix B are strictly negative, then every entry in B^{-1} will be stricktly positive. Hands only proves it for 2 by 2 matrices. Hands references several sources for Theorem 2 (the one I know is by Berman and Plemmons (yes our own Prof. Plemmons at WFU; that book, published in 1979 is now republished as a classic by the Soc. of Industrial and Applied Math.) This modified statement of Theorem 2 is in fact true for n by n matrices and can be found in Berman and Plemmons. Thus all of our comparative statics are positive if only one $z_\alpha^i > 0$.

2.4 Exercises

1. Use economics principles to argue that Walras' law is reasonable.
2. Use (2.4) and assumption about substitutes to show that $z_i^{(i)}(p) < 0$. Thus each excess demand function slopes downward with respect to its own price.
3. Differentiate Walras' law with respect to α and show that $z_\alpha^{(n)} < 0$ forces at least one $z_\alpha^{(i)}$ to be positive. Conclude that at least one comparative static is positive.

2.5 Competitive Firm with n Inputs

Hands, Sect 7.2, 7.3.

We now extend the discussion of Section 1.5 to the case of a competitive firm with $n > 2$ inputs. In Section 1.5, the two inputs were capital and labor. Here the inputs may represent different classes of labor, whose wage rates are different, different kinds of capital whose unit costs are different, as well as other kinds of inputs. Here we will need a second powerful theorem in linear algebra, as well as second order conditions for problems with many variables.

So the problem is for a perfectly competitive firm to maximize profits when the production function exploits n inputs. So the firm seeks to maximize profits

$$\Pi = pf(x) - \sum_{k=1}^{n} w_k x_k,$$

where the vector of inputs is $x = (x_1, x_2, \cdots, x_n)$ with corresponding cost vector $w = (w_1, w_2, \cdots, w_n)$. The number p is the unit market price of the firm's product. Since the firm is perfectly competitive it accepts the market price p and the unit price w_i of the ith input x_i. Of course, we assume from economics priciples that the production function $f(x)$ satisfies the conditions

$$f_i = \frac{\partial f}{\partial x_i} > 0 \quad \text{and} \quad f_{ii} = \frac{\partial^2 f}{\partial x_i^2} < 0$$

for each input x_i.

The first order necessary conditions for the optimum values $(x_1^*, x_2^*, \cdots, x_n^*)$ are

$$\Pi_i = pf_i(x) - w_i = 0, \tag{2.5}$$

for each input $i = 1, 2, \ldots, n$. This gives n equations which we would like to solve for the values $(x_1^*, x_2^*, \cdots, x_n^*)$ in terms of the w_i's and p, after which we would like to find the partial derivatives (comparative statics) of each x_i^* with respect to each w_j and p. The implicit function theorem enables us to skip the task of finding explicit formulas for the x_i^*'s. Assuming that the firm is able to maximize profits for the current values of w and p, the equations (1) have a solution for the current values of w, p. Assuming also that the Jacobian matrix corresponding to the equations (1), evaluated at these values, is invertible, then the hypotheses of the implicit function theorem are satisfied. Thus we are guaranteed, at least theoretically, that the solution x^* of (1) exists for other nearby values of w and p, and the comparative statics can be found by the (generic) formula

$$F_y H'(x) = -F_x.$$

Of course we have to translate these generic symbols. For our problem, we replace y by x (the variable we are solving for), we replace x by (w, p) (all other variables), F stands for all the left sides of (2.5), and $H'(x)$ is the matrix whose entries are the desired comparative statics.

Filling in the indicated matrices, we get

$$\begin{bmatrix} pf_{11}(x^*) & pf_{12}(x^*) & \cdots & pf_{1n}(x^*) \\ pf_{21}(x^*) & pf_{22}(x^*) & \cdots & pf_{2n}(x^*) \\ \vdots & \vdots & \vdots & \vdots \\ pf_{n1}(x^*) & pf_{n2}(x^*) & \cdots & pf_{nn}(x^*) \end{bmatrix} \begin{bmatrix} \frac{\partial x_1^*}{\partial w_1} & \frac{\partial x_1^*}{\partial w_2} & \cdots & \frac{\partial x_1^*}{\partial w_n} & \frac{\partial x_1^*}{\partial p} \\ \frac{\partial x_2^*}{\partial w_1} & \frac{\partial x_2^*}{\partial w_2} & \cdots & \frac{\partial x_2^*}{\partial w_n} & \frac{\partial x_2^*}{\partial p} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{\partial x_n^*}{\partial w_1} & \frac{\partial x_n^*}{\partial w_2} & \cdots & \frac{\partial x_n^*}{\partial w_n} & \frac{\partial x_n^*}{\partial p} \end{bmatrix} = -\begin{bmatrix} -1 & 0 & \cdots & 0 & f_1(x^*) \\ 0 & -1 & \cdots & 0 & f_2(x^*) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & -1 & f_n(x^*) \end{bmatrix} \tag{2.6}$$

The first matrix on the left side of this equation is called the Hessian matrix for Π evaluated at x^*. We denote this Hessian matrix by the symbol $H_\Pi(x^*)$. Economic intuition indicates that we would expect each of the comparative statics

$$\frac{\partial x_k^*}{\partial w_k}$$

to be negative, indicating that the amount of input k used will go down if the cost w_k of that input increases. This is called the "own price" effect. We would expect a "cross price" effect, given by

$$\frac{\partial x_k^*}{\partial w_j},$$

for $k \neq j$, to be more complicated; its sign would depend on the mix of complements and substitutes in the input vector. So there is no use working hard to analyze them. We will examine carefully the bottom entry

$$\frac{\partial x_n^*}{\partial w_n}$$

which is the "own price" effect of the nth input. Later we will also examine closely the comparative statics in the last column of the second matrix in (2.6).

To find

$$\frac{\partial x_n^*}{\partial w_n},$$

we need to solve the matrix equation

$$H_\Pi(x^*) \begin{bmatrix} \frac{\partial x_1^*}{\partial w_n} \\ \frac{\partial x_2^*}{\partial w_n} \\ \vdots \\ \frac{\partial x_n^*}{\partial w_n} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ +1 \end{bmatrix} \tag{2.7}$$

for the last entry in the unknown vector. Using Cramer's rule (solving by determinants), we see that the denominator is the determinant of $H_\Pi(x^*)$ and the numerator, after expansion by cofactors down the last column, is the determinant of the upper left $n-1 \times n-1$ "corner" of $H_\Pi(x^*)$. (A submatrix of given matrix obtained by crossing out corresponding rows and columns of the given matrix is called a *principal minor*; we will consider such principal minors in more detail later.) Can we attach a sign to this quotient of determinants without evaluating the (possibly) large determinants involved? Not always. But we are at a maximum, and there are second order conditions for a maximum for problems with many variables just as there are for problems with two variables. We will discuss these conditions in generality in Section 2.7.1. As always, second order conditions are sufficient for a maximum but not necessary, so such conditions are not guaranteed to hold at a maximum. Nevertheless, there is high probability that the second order conditions for a maximum are satisfied at the maximum, and the only way to make progress is to *assume* that these conditions hold. As we shall see in Section 2.7.1, these conditions tell us the numerator and denominator in our fraction of determinants will have opposite signs and the "own price" comparative static

$$\frac{\partial x_n^*}{\partial w_n}$$

will be negative as expected. We emphasize that we can draw this conclusion only if the second order conditions for a maximum are satisfied. In cases where the second order conditions fail to be true, we seem to be fenced away from this conclusion. Just how unfortunate this state of affairs seems varies from person to person. Most economists seem not very bothered.

Let us denote the last column of comparative statics (those which give the impact of a change in p on the optimal values of x^*) by the letter v. Then we have

$$H_\Pi(x^*)v = -\nabla f,$$

where, as in vector calculus, ∇f denotes the vector of first partial derivatives of f. In vector calculus, this vector is usually written as a "row" vector; here we write it as a column vector.

There is really no hope of assigning a definite sign to the individual comparative statics in v; in fact, some of them may be negative and some positive (but see problem 4 in Exercises 2.6 below for conclusions with addditional assumptions). However, we can use them to show a related interesting fact: the effect on the optimal output $y^* = f(x^*)$ of an increase in selling price p is positive, as our economic intuition would predict.

Using the information above, we calculate

$$\frac{\partial y^*}{\partial p} = \sum_{k=1}^{n} f_k(x^*)\frac{\partial x_k^*}{\partial p} = \nabla f(x^*) \cdot v = -H_\Pi(x^*)v \cdot v.$$

Now we come to our first use of a second powerful theorem in linear algebra. It is fairly complicated to state, so we postpone a full statement until Section 2.7. A consequence of this theorem is that the second order conditions for a maximum force the dot product on the right side of the last equation, ignoring the minus sign, to be negative. Hence the right side of this last equation is positive, implying that the left side is also positive, giving the expected sign of the comparative static

$$\frac{\partial y^*}{\partial p} > 0.$$

Note that again, to draw this conclusion, we need to know that the second order conditions for a maximum are satisfied; otherwise we are left in an uncomfortable position. Thus the second order conditions assume a position of great value for they are the key to both of the economic conclusions we have drawn.

We conclude this section with a few comments about the Hotelling lemma and a result called the envelope theorem. The first part of Hotelling's lemma says that

$$\frac{\partial \Pi^*}{\partial w_i} = -x_i^*$$

for each i. Mathematically, this means that if you choose $x = x^*(\alpha)$ to maximize a function of the form

$$z = g(x, \alpha)$$

with α as a parameter, then the derivative of the maximum value of the function $z^* = g(x^*(\alpha), \alpha)$ with respect to the parameter α is just the same as you would get if you did something which would normally be illegal and just differentiated the function $z = g(x, \alpha)$ with respect to α, without first maximizing, treating x as not depending on α (which is not true) and then substituted x^* for x! This fact is called the envelope theorem and is presented in Section 2.7.4.

The second part of Hotelling's lemma says that

$$\frac{\partial \Pi^*}{\partial p} = y^*.$$

You will have a chance in one of the homework problems to prove it; a simple way is to use the envelope theorem.

One final observation: the first order conditions (2.5) for a maximum are not changed if each w_i and p are replaced by λw_i and λp. So x_k^* is not changed if all prices are scaled the same. Thus the solutions of these first order conditions are not changed and so the functions $x_k^*(w, p)$ are not changed by such a scaling of prices. It follows that they are thus homogeneous of degree zero. This fact will be needed in problem 2(a) in Exercises 2.6. If a function is known to be homogeneous of some degree r, one usually uses this knowledge to apply Euler's theorem, discussed in Section 2.7.3 below. You should jump forward and read that section now because you will need to use Euler's theorem to do problem 2(b) of Exercises 2.6.

2.6 Exercises

1. Here is an explicit example to work out. Suppose the production function of a competitive firm is

$$y = f(x_1, x_2, x_3) = \sqrt{x_1} + \sqrt{x_2} + \sqrt{x_3}.$$

   (a) Using the first order conditions, find explicit formulas for each imput demand function $x_i^*(w, p)$.
   (b) Are the input demand functions homogeneous of any degree?
   (c) Find all four comparative statics for x_1^*.
   (d) Find the supply function y^*.
   (e) Is the supply function homogeneous of any degree?
   (f) Find the profit function Π^*.

2. (a) Show that the maximum profit function Π^* of a competitive firm is homogeneous of degree 1 and that the supply function y^* is homogeneous of degree 0.
   (b) If y^* is the supply function of a competitive firm, show that $\partial y^*/\partial w_i < 0$ for at least one value of i.

3. (a) Prove the second part of Hotelling's lemma:
$$\frac{\partial \Pi^*}{\partial p} = y^*.$$
   (b) Show that for a competitive firm,
$$\frac{\partial y^*}{\partial w_i} = -\frac{\partial x_i^*}{\partial p}.$$

4. Show that if a competitive firm satisfies the conditions that $f_{ij} \geq 0$ whenever $i \neq j$, then the comparative statics $\partial x_i^*/\partial w_j$ will all be nonpositive when $i \neq j$. (This says that if all inputs are normal, then inputs will never be substitutes.) Also show that $\partial x_i^*/\partial p > 0$ for all i.

5. Suppose that our firm is competitive in the product market but is a monopsonist in the input market (i.e., the unit cost w_i of each input is an increasing function of the quantity used: $w_i = w_i(x_i)$ and $dw_i/dx_i > 0$ for all i. Find the first order conditions for this firm and interpret them.

2.7 A Mathematical Interlude

In this section we need to go back and pick up some pieces. It is important to take the time to make sure we understand the first and second order conditions for an optimum and we will do that first. Next we will discuss more completely the important matrix results we have needed. Finally, we will devote some attention to homogeneous functions and the simple fact called Euler's theorem.

2.7.1 Unconstrained Optimization Problems

Hands (pp. 259-262)

We review some material from vector calculus, which you probably know well in the case of a function of two variables. In order to gain maximum clarity, we first discuss the case of functions of three variables.

FIRST ORDER NECESSARY CONDITIONS:

If $u(x, y, z)$ has an optimum (either a maximum or a minimum) at the point (x^*, y^*, z^*), then (x^*, y^*, z^*) must be a solution of the three simultaneous equations
$$u_x(x^*, y^*, z^*) = 0, \quad u_y(x^*, y^*, z^*) = 0, \quad u_z(x^*, y^*, z^*) = 0.$$
Any point satisfying these three equations is called a critical point for $u(x, y, z)$.

SECOND ORDER SUFFICIENT CONDITIONS:

Suppose (x^*, y^*, z^*) is a critical point of $u(x, y, z)$. Let H be the Hessian matrix

$$H = \begin{bmatrix} u_{xx} & u_{xy} & u_{xz} \\ u_{yx} & u_{yy} & u_{yz} \\ u_{zx} & u_{zy} & u_{zz} \end{bmatrix}$$

and let H_k be the upper left $k \times k$ submatrix of H. That is, $H_1 = u_{xx}$,

$$H_2 = \begin{bmatrix} u_{xx} & u_{xy} \\ u_{yx} & u_{yy} \end{bmatrix},$$

$H_3 = H$. Assume that $\det H \neq 0$ when evaluated at (x^*, y^*, z^*).

1. If $\det H_k > 0$ when evaluated at (x^*, y^*, z^*) for $k = 1, 2, 3$, then $u(x, y, z)$ has a relative minimum at (x^*, y^*, z^*).
2. If $\det H_1 = u_{xx} < 0$, $\det H_2 > 0$ and $\det H_3 < 0$ when evaluated at (x^*, y^*, z^*), then $u(x, y, z)$ has a relative maximum at (x^*, y^*, z^*).
3. If (1) and (2) both fail, then $u(x, y, z)$ has neither a maximum or a minimum at (x^*, y^*, z^*).

Note: If $\det H = 0$ at (x^*, y^*, z^*), then no conclusion can be drawn. In this case, $u(x, y, z)$ might have a relative maximum, a relative minimum, or neither at (x^*, y^*, z^*).

You need to think long and hard about the difference between conditions which are necessary and those which are only sufficient. In a typical calculus class, one starts with a function and wants to find maximum or minimum values. One first finds critical points (the points that satisfy the first order necessary conditions) and then tests each critical point to see if it is a max or a min or neither. Second order conditions are used for this test and if such conditions are met, then a conclusion may be drawn.

However, an economist is not usually searching for a maximum. Instead, an economist *assumes* he or she is at a maximum and wishes to draw conclusions about how such a maximum will shift if certain parameters in the problem should change. Thus an economist needs statements of the form: if I am presently at a maximum, what conclusions can I draw? So I need *necessary* conditions. Sufficient conditions may not be satisfied at a maximum. Thus, economists are in a curious position. They deeply wish sufficient conditions were actually necessary. Many important economic conclusions (certainly in this course) depend logically on *assuming* that sufficient conditions are actually satisfied. The result is that one is always a little nervous!

We pass to the case of a function of n variables $x_1, x_2, \cdots x_n$. We use the vector notation $x = (x_1, x_2, \cdots, x_n)$.

FIRST ORDER NECESSARY CONDITIONS:

If $u(x)$ has an optimum at the point x^*, then x^* must be a solution of the n simultaneous equations

$$u_{x_1}(x^*) = 0, \quad u_{x_2}(x^*) = 0, \quad \cdots, \quad u_{x_n}(x^*) = 0.$$

Any point satisfying these n equations is called a critical point for $u(x)$.

SECOND ORDER SUFFICIENT CONDITIONS:

Suppose x^* is a critical point of $u(x)$. Let H be the Hessian matrix

$$H = \begin{bmatrix} u_{x_1x_1} & u_{x_1x_2} & \cdots & u_{x_1x_n} \\ u_{x_2x_1} & u_{x_2x_2} & \cdots & u_{x2x_n} \\ \vdots & \vdots & & \vdots \\ u_{x_nx_1} & u_{x_nx_2} & \cdots & u_{x_nx_n} \end{bmatrix}$$

and let $H_k, k = 1, 2, \cdots, n$ be the upper left $k \times k$ submatrix of H. Assume that $\det H \neq 0$ when evaluated at x^*.

1. If $\det H_k > 0$ when evaluated at x^* for $k = 1, 2, \cdots, n$,then $u(x)$ has a relative minimum at x.
2. If $(-1)^k \det H_k > 0$ when evaluated at x^*, for $k = 1, 2, \cdots, n$, then $u(x)$ has a relative maximum at x^*. (This condition is just that $\det H_k$ alternates in sign, beginning with a negative sign, as k goes from 1 to n.)
3. If (1) and (2) both fail, then $u(x)$ has neither a maximum or a minimum at x^*.

Note: If $\det H = 0$ at x^*, then no conclusion can be drawn.

The matrices H_k above are called leading principal minors of the matrix H. One gets the submatrix H_k by deleting the last $n - k$ rows and the last $n - k$ columns of H. If you take a square matrix and delete some of the rows and the corresponding columns, what is left is called a principal minor. For example, if you delete the second, fifth, and ninth rows of a 12×12 matrix and also delete the second, fifth, and ninth columns, you will have a 9×9 principal minor. You should realize that the numbers on the diagonal of a principal minor were also on the diagonal of the original larger matrix. Thus the second order conditions above are conditions on the determinants of all the *leading* principal minors. There is a discrepancy between my statement of the second order conditions and the one in many textbooks. Many books (for example the one by Hands) place sign conditions on all principal minor determinants. For example, for a maximum the second order conditions state that all 1×1 principal minor determinants be negative, all 2×2 principal minor determinants be positive, all 3×3 principal minor determinants be negative, etc. A matrix satisfying all these conditions is called an NP matrix (more about them later).

This appears to be many more second order conditions than my discussion indicates. Your mind may rest at ease. There is a theorem in linear algebra which says that a matrix which satisfies the fewer conditions I gave earlier automatically satisfies all these additional conditions. What difference does it make? Well, if you were going to use the second order conditions in the proper way, to show that a critical point really is a maximum, then my description seems better: you have fewer conditions to verify. However, if you are an economist reversing the situation so that you want to use the second order conditions (which are not always true at a maximum, but hopefully are usually true), then the book's description places more second order conditions at your disposal, and these extra conditions might come in handy. We should emphasize again that second order conditions are sufficient but not necessary at an optimum. But if we are going to *assume* that they hold at an optimum, the assumption that all the additional conditions hold is no greater risk than assuming that the fewer I stated above hold there.

2.7.2 Matrix Theorems

Hands, Sect. 6.4, 6.6.

We now discuss in some detail the circle of ideas around the two important matrix theorems we have used. Suppose A is an $n \times n$ matrix. Recall from the last section that A is called an NP matrix if all 1×1 principal minor determinants are negative, all 2×2 principal minor determinants are positive, all 3×3 principal minor determinants are negative, and this alternation of sign continues up to the determinant of the entire matrix A. Note that you can think of the whole matrix as a principal minor of itself; there is only one principal minor that large. Do you see how many principal minors have size $n - 1$?

A matrix is called a P matrix if the determinant of every principal minor is positive. I suppose the P is for the word positive and the NP is for negative and positive, to remind you that the signs of the principal minor determinants alternate in sign as the size increases.

Here is some additional terminology. A square matrix A is called an M matrix if it is simultaneously a P matrix and all off-diagonal entries are nonpositive, i.e., $a_{ij} \leq 0$ if $i \neq j$. With this terminology, we can now state the first big matrix theorem, dealing with M matrices, of which Theorem 2 is only a part.

Theorem 3 *Suppose B is an $n \times n$ matrix and all the off-diagonal entries are nonpositive. The following conditions are then equivalent, i.e., if any one of them is true, so are the other two:*

(a) *There exists an n-vector x all of whose entries are nonnegative for which all entries in Bx are positive.*

(b) *B is a P matrix.*

(c) *B is invertible and every entry in B^{-1} is nonnegative.*

The part of this theorem which we stated earlier as Theorem 2 is that if (a) is true then (c) is true. As indicated at the end of Section 2.3, with the added knowledge that the off-diagonal entries in B are strictly negative, one can conclude that all entries in B^{-1} are strictly positive.

It is important to know that a square matrix A is a P matrix if and only if its negative $-A$ is an NP matrix. Think about this: if a matrix is negated, that just means that each row has been multiplied by negative one. Multiplying a row by negative one changes the sign of the determinant of a matrix. So if there are an even number of rows in a matrix and it is negated, its determinant does not change sign; if there are an odd number of rows, the sign of the determinant does change. Do you see why this is important? The theorem above only addresses the case of P matrices. If you suspect that a matrix is an NP matrix, you should negate it and try to apply the theorem above to the negated matrix.

Note that the conditions for a P matrix are the second order conditions for a minimum and the conditions for an NP matrix are the second order conditions for a maximum. So it should be apparent why our Theorem 3 should be so valuable in economics. It is a fact (see the book by Hands [4], p. 231, 2nd ed. (p. 283, 1st ed.)) that no area of mathematics has been influenced as much by economics as the theory of M matrices. Economists have provided the major questions and problems for this theory and some of the results have been obtained first by economists. The reason should be clear.

To discuss the second big theorem, we need to consider another property of matrices. Some more terminology: a square matrix A is called positive definite if $x^T Ax = x \cdot Ax > 0$ for every nonzero n-vector x; A is called negative definite if $x^T Ax = x \cdot Ax < 0$ for every nonzero n-vector x. These ideas have close connections with second order conditions because the proof of the second order conditions involves finding out if the relevant Hessian matrices are positive definite or negative definite. Actually, the following theorem must be known before the second order conditions can be stated.

Theorem 4 *A symmetric matrix A is positive definite if and only if A is a P matrix and a symmetric matrix A is negative definite if and only if A is an NP matrix.*

Notice that the Hessian matrices are all symmetric because of the theorem in vector calculus that the mixed partials are equal, i.e., $u_{x_i x_j} = u_{x_j x_i}$. Theorem 4 says that the conditions for a maximum can be described equally well by saying the second order conditions hold (the Hessian is an NP matrix) or by saying that the Hessian is negative definite, which was the fact we needed when analyzing the effect on output of a competitive firm by an increase in sales price.

2.7.3 Homogeneous Functions and Euler's Theorem

Hands, Sect. 2.2.

Suppose $x = (x_1, x_2, \cdots, x_n)$ is an n-vector and $f(x)$ is a real valued function of x. Thus for a given vector x, $f(x)$ is a real number. Then the function f is called *homogeneous of degree r* if $f(\lambda x) = \lambda^r f(x)$ for every n-vector x. Here is a simple example for the case $n = 3$:

$$f(x_1, x_2, x_3) = x_1^2 + 3x_2^2 - 2x_1x_3.$$

You can see that

$$f(\lambda x_1, \lambda x_2, \lambda x_3) = \lambda^2 x_1^2 + 3\lambda^2 x_2^2 - 2\lambda^2 x_1 x_3 = \lambda^2 f(x_1, x_2, x_3).$$

So this function is homogeneous of degree 2. The important result about homogeneous functions is called Euler's theorem. Here it is.

Theorem 5 *If $f(x)$ is a differentiable function of the n-vector x, then $f(x)$ is homogeneous of degree r if and only if*

$$rf(x) = \sum_{i=1}^{n} f_i(x)x_i = \nabla f(x) \cdot x,$$

for every n-vector x, where

$$f_i = \frac{\partial f}{\partial x_i}.$$

To prove one part of this theorem, you only need to differentiate both sides of the equation

$$f(\lambda x) = \lambda^r f(x)$$

with respect to λ to get

$$\sum_{i=1}^{n} f_i(\lambda x)x_i = r\lambda^{r-1} f(x),$$

and then substitute $\lambda = 1$. This is essentially the calculation we did to arrive at equation (2.4).

2.7.4 The Envelope Theorem

Here it is:

Theorem 6 *Suppose the function $g(x,\alpha)$ depends on a vector $x = (x_1, x_2, \cdots, x_n)$ of endogenous variables and some exogenous variable α. (There could, of course, be more than one exogenous variable.) Suppose that $x^*(\alpha)$ is a critical point of this function, i.e., all of the first partial derivatives of g with respect to the x_i's are zero when evaluated at $x * (\alpha)$. So $x^*(\alpha)$ could be a relative maximum, a relative minimum, or a saddle point of the function g. Note carefully that the critical point is a function of the exogenous variable α. Then the derivative of the function $h(\alpha) = g(x^*(\alpha), \alpha)$ is given simply by*

$$h'(\alpha) = \frac{\partial g}{\partial \alpha}(x^*(\alpha), \alpha),$$

i.e., the partial derivative of the original function $g(x,\alpha)$ with respect to α before finding the critical point $x^*(\alpha)$ and then evaluate the result at the critical point $x^*(\alpha)$.

The proof is simply achieved by doing the derivative $h'(\alpha)$ the proper way. Just write

$$z^* = g(x^*(\alpha), \alpha)$$

and differentiate using the chain rule from vector calculus to get:

$$\frac{dz^*}{d\alpha} = \sum_{k=1}^{n} g_k(x^*(\alpha), \alpha)\frac{\partial x_k^*}{\partial \alpha} + g_{n+1}(x^*(\alpha), \alpha) = g_{n+1}(x^*(\alpha), \alpha),$$

since all the terms in the sum are zero by the first order conditions.

The envelope theorem essentially allows you to skip the chain rule calculation by differentiating and then substituting the critical point, rather than first substituting the critical point and then differentiating.

2.8 Exercises

1. Do the second order conditions for long-run profit maximization hold for a competitive firm with production function $f(K,L) = L^{3/2}K^{3/2}$?

2. If a matrix A is 10×10, how many principal minors are there of order 9? How many of order 8?

3. Verify the following comments about Theorem 3.

   (a) Suppose (c) is true about an $n \times n$ matrix B. Show that (a) is also true for B, without the assumption that all the off-diagonal entries in B are nonpositive, by letting y be the n-vector consisting of all 1's and showing that $x = B^{-1}y$ satisfies the requirements of part (a).

   (b) Show that if x satisfies (a) and the off-diagonal entries in B are nonpositive, then in fact every entry in x is positive, by supposing that some entry in x is zero and examining the corresponding entry in $y = Bx$.

   (c) Suppose an $n \times n$ matrix A satisfies the condition that every off-diagonal entry is nonnegative. Apply Theorem 3 to the matrix $B = -A$ and describe what (a), (b),and (c) say about A.

Chapter 3

Optimization with Equality Constraints

Optimization problems in economics often involve constraints. A firm may be in the position of trying to minimize costs while fulfilling an order with a fixed output requirement. A government agency may be trying to maximize output with a fixed budget. Or consider the theory of consumer behavior. An individual consumer wants to maximize utility but, as we all know, does not usually live in conditions of unlimited available resources. Instead, the consumer divides available budget on the expenditure of a number of consumer goods. In an effort to understand the economics questions of interest and the mathematical techniques involved, we will focus attention almost exclusively on the utility maximization problem. We begin with a two good world, but later pass to an n good world. In this chapter we shall also meet the concept of dual problems in economics as we consider the problem dual to the utility maximization problem; this is the problem of the consumer who first decides on a specific level of utility to achieve and seeks to minimize costs in achieving this level of utility. In the process we will meet the famous Slutsky equations. The mathematics of this chapter is basically the method of Lagrange multipliers.

3.1 Utility Maximization in a Two Good World

Compare Hands, Sect. 8.2.

We begin with a very simple problem in the theory of consumer behavior. We imagine an individual who lives in an imaginary world with two consumer goods. The only decision facing this individual is how to divide her available budget resources between these two goods. We will call the two goods X and Y, and use x to denote the number of units of X, y the number of units of Y that she buys. Our heroine has a utility function unique to her, which we denote by $U(x, y)$, which measures the number of utils of utility which she receives from consuming x units of X and y units of Y. She has available a total of B dollars. Her problem is to

$$\text{Maximize } U(x, y)$$

subject to the constraint

$$p_1x + p_2y = B.$$

You see that she expects to spend her total budget B by dividing it into expenditures p_1x on X and p_2y on Y. Here p_1, p_2 are the unit prices of X and Y, respectively. Clearly, x and y are endogenous variables and p_1, p_2, B are exogenous variables. The usual economic assumptions are that

$$U_x > 0, \qquad U_y > 0$$

(more is better) and

$$U_{xx} < 0, \qquad U_{yy} < 0$$

(diminishing returns). The sign of $U_{xy} = U_{yx}$ depends on whether or not X and Y are substitutes, complements, or unrelated in our individuals utility function.

You will recognize this as a constrained maximization problem from vector calculus. One method of solving such a problem is to use Lagrange multipliers, a method you should recall from vector calculus. However, do not fret; we will review it carefully. There are basically two ways of thinking about this method. The historical way, associated with Lagrange himself, is to form a new function called the Lagrangian, involving an additional variable λ, called the Lagrange multiplier:

$$L(\lambda, x, y) \equiv U(x, y) + \lambda(B - p_1 x - p_2 y),$$

and apply to it the first order conditions for a critical point:

$$L_\lambda(\lambda, x, y) = B - p_1 x - p_2 y = 0,$$

$$L_x(\lambda, x, y) = U_x(x, y) - \lambda p_1 = 0,$$

$$L_y(\lambda, x, y) = U_y(x, y) - \lambda p_2 = 0.$$

(Because of things to come, you will find it less confusing if you always put the λ first among the variables and differentiate first with respect to λ so that the pattern is always the same.) Another way to arrive at these same three equations is to say that the optimum values of x and y have to satisfy the budget constraint and also the requirement that $\nabla U = \lambda \nabla g$, where the constraint has been written in the form $g(x, y) = 0$. The mathematical theory assures us that the desired solution of our maximization problem (if indeed there is one) will be found among the simultaneous solutions of this system of three equations. If we denote our optimal choice of x and y by x^* and y^*, then there will be an accompanying λ^* so that these three values satisfy the three equations above. In the absence of specific knowledge of our friend's utility function, we cannot solve these equations. But we are undaunted by that problem, for our economics perspective only asks the question of comparative statics. We know how to make progress on this problem. The implicit function theorem is by now an old friend. Denoting the left sides of our three first order conditions by the vector function $F(\lambda^*, x^*, y^*)$, we have

$$F(\lambda^*, x^*, y^*) = \begin{bmatrix} B - p_1 x^* - p_2 y^* \\ U_x(x^*, y^*) - \lambda^* p_1 \\ U_y(x^*, y^*) - \lambda^* p_2 \end{bmatrix}.$$

We know that our comparative static matrix can be found from the equation

$$F_y H' = -F_x,$$

where the generic variable $y = (\lambda^*, x^*, y^*)$ gives our endogenous variables and the generic variable $x = (p_1, p_2, B)$ gives our exogenous variables. Note that the "artificial" variable λ^*, called the Lagrange multiplier, is viewed as an endogenous variable. We calculate

$$F_y = \begin{bmatrix} 0 & -p_1 & -p_2 \\ -p_1 & U_{xx} & U_{xy} \\ -p_2 & U_{yx} & U_{yy} \end{bmatrix},$$

and

$$F_x = \begin{bmatrix} -x^* & -y^* & 1 \\ -\lambda^* & 0 & 0 \\ 0 & -\lambda^* & 0 \end{bmatrix}.$$

We know what to do. To find any desired comparative static, all we have to do is to solve using Cramer's rule. Let's start with one we think we can predict. For example, do you feel assured that your economics intuition tells you what will happen to the amount of good X our consumer will buy if the price p_1 of good X rises? Would you not expect that this comparative static will be negative? Let's check and see. Using Cramer's rule, the desired comparative static $\partial x^*/\partial p_1$ will be the quotient of two determinants. The denominator will be the $\det F_y$; the matrix F_y is called a bordered Hessian. The first row and first column are called borders and the 2×2 matrix which remains after deleting the borders you recognize as the Hessian of the utility function. The numerator is

$$\det \begin{bmatrix} 0 & x^* & -p_2 \\ -p_1 & \lambda^* & U_{xy} \\ -p_2 & 0 & U_{yy} \end{bmatrix} = -x^* \left(p_2 U_{xy} - p_1 U_{yy}\right) - \lambda^* p_2^2.$$

We now need to sign the numerator and the denominator. If you work out the denominator determinant, you arrive at ambiguous results. We shall see later that second order conditions for this problem indicate that the denominator determinant is positive. Thus, assuming the second order conditions hold (not guaranteed of course), we can conclude that the denominator is positive.

Fortunately, the first order conditions give us a sign for λ^*, since for example the second first order condition tells us that

$$\lambda^* = \frac{U_x}{p_1} > 0.$$

If $U_{xy} \geq 0$ (complements or unrelated), it is clear that the parenthetical expression will be positive, so in this case the numerator is negative and the comparative static is unambiguously negative, as expected. If $U_{xy} < 0$ (X and Y are substitutes), the sign of our parenthetical expression is ambiguous. Only by taking the position that "own" effects always dominate "cross" effects can we conclude that the comparative static is negative in the case of substitutes. In the homework, you are asked to show similarly that $\partial x^*/\partial B > 0$ if X and Y are complements or unrelated.

The results here are very reassuring, but in fact they are misleading. It is natural to expect that an n good world, with $n > 2$, would give similar conclusions. Unfortunately, this expectation is doomed. In the homework, you are asked to work out the "own price" effect in a 3 good world and face the fact that the sign of this comparative static is generally ambiguous. The ball is then in the economist's court. Some explanation must be given. We will see in a later section that the explanation has to do with income and substitution effects and involve some famous equations named for Slutsky.

3.2 Exercises

1. Suppose our individual has the specific utility function $U(x, y) = xy$.

   (a) Find explicit formulas for our individual's demand functions x^* and y^*, as well as λ^*.

(b) Find the comparative statics showing the effect on x^* of an increase in p_1 and of an increase in B.

(c) Are the goods X and Y complements, substitutes, or unrelated?

(d) Let $U^* = U(x^*, y^*)$ be the maximum utility achieved by our individual under the budget constraint. Find the derivative $\partial U^*/\partial B$ and show that it is the same as λ^*.

2. For the general two good problem considered in the previous section, find the comparative statics $\partial x^*/\partial p_2$ and $\partial x^*/\partial B$. Show that the expected sign for $\partial x^*/\partial B$ is obtained for complementary or unrelated goods. What about the sign for $\partial x^*/\partial p_2$?

3. For the general two good problem considered in the previous section, let $U^* = U(x^*, y^*)$ be the consumer's maximum utility.

   (a) Use the chain rule to find an expression for $\partial U^*/\partial B$ and then use the first order conditions to eliminate the partial derivatives of U.

   (b) Differentiate the budget constraint, which holds at the maximum, with respect to B and combine your result with part (a) to conclude that $\partial U^*/\partial B = \lambda^*$. Thus λ^*, which started out as an "artificial" variable has a meaning in economics: it is the marginal utility of income to the consumer.

   (c) Explain why $U^* = L(\lambda^*, x^*, y^*)$ and use the envelope theorem to differentiate U^* with respect to B and quickly get the result of part (b).

4. Consider utility maximization in a 3 good world:

$$\text{maximize} \qquad U(x, y, z)$$

$$\text{subject to} \qquad p_1 x + p_2 y + p_3 z = B.$$

   Find the comparative static $\partial z^*/\partial p_3$ and see what happens when you try to sign the numerator. The denominator may be signed from second order conditions we will discuss later, as can the coefficient of λ^* in the numerator. It is the other part of the numerator that presents the problem.

3.3 Choice Between Labor and Leisure

Compare Hands, Sect. 8.3.

Before passing to a consideration of an n good utility maximization problem, it is of interest to enrich our two good model by allowing our consumer to choose how much of her time is devoted to labor. Mathematically we must deal with 4×4 determinants; economically, we will be able to draw some interesting conclusions. The material in this section is taken from the paper Lagrange Multiplier Problems in Economics, which appeared in the American Mathematical Monthly, vol. 91, 1984, pp.404-412. A more extended discussion and further results may be found there.

In this model, the utility function $U(x, y, \ell)$ depends not only on the amounts of X and Y consumed, but also the amount ℓ of leisure time $\mathcal{L}$ available to the consumer. The consumer implicitly chooses the amount of work time by explicitly choosing the number of leisure hours. Letting T be the total time available (for example, $T = 168$ hours per week), the number of hours worked is $T - \ell$. If w (an exogenous variable) is the wage rate of our consumer, then $w(T - \ell)$ is the total income available to our consumer for expenditures. Thus our utility maximization problem becomes

$$\text{Maximize } U(x, y, \ell)$$

subject to the constraint

$$p_1 x + p_2 y = w(T - \ell).$$

The natural assumptions are now

$$U_x > 0, \qquad U_y > 0, \qquad U_\ell > 0,$$
$$U_{xx} < 0, \qquad U_{yy} < 0, \qquad U_{\ell\ell} < 0,$$
$$U_{x\ell} > 0, \qquad U_{y\ell} > 0.$$

The sign of U_{xy} depends, as before, on whether or not X and Y are substitutes, complements, or unrelated. We naturally assume leisure time $\mathcal{L}$ is complementary to X or Y (it takes time to enjoy one's purchases).

As before, we apply the method of Lagrange multipliers. We let

$$L(\lambda, x, y, \ell) = U(x, y, \ell) + \lambda(w(T - \ell) - p_1 x - p_2 y),$$

and then the first order conditions are

$$\begin{aligned} L_\lambda &= w(T - \ell) - p_1 x - p_2 y = 0, \\ L_x &= U_x - \lambda p_1 = 0, \\ L_y &= U_y - \lambda p_2 = 0, \\ L_\ell &= U_\ell - \lambda w = 0. \end{aligned}$$

The optimal values $\lambda^*, x^*, y^*, \ell^*$ are simultaneous solutions of these four equations. Note that each of the last three first order conditions imply that $\lambda^* > 0$. Since there are four endogenous variables and three exogenous variables (p_1, p_2, w), then there are twelve comparative statics we might find. In the now routine way, we apply the implicit function theorem to find these comparative statics. Using the generic notation, we compute

$$F_y = \begin{bmatrix} 0 & -p_1 & -p_2 & -w \\ -p_1 & U_{xx} & U_{xy} & U_{x\ell} \\ -p_2 & U_{yx} & U_{yy} & U_{y\ell} \\ -w & U_{\ell x} & U_{\ell y} & U_{\ell\ell} \end{bmatrix}.$$

To apply the implicit function theorem, we need to know that F_y is nonsingular. This will be true if the second order conditions we discuss later are satisfied. In fact, the second order conditions for this problem turn out to be that the determinant of F_y is negative and the determinant of the upper left 3×3 corner of F_y is positive at the optimal point $(\lambda^*, x^*, y^*, \ell^*)$. In particular, F_y is nonsingular.

Before finding any comparative statics, we discuss the meaning of the Lagrange multiplier λ^* for this problem. The optimal point $(\lambda^*, x^*, y^*, \ell^*)$ is the solution of the first order conditions and therefore is a function of the exogenous variables p_1, p_2, w. Letting $U^* = U(x^*, y^*, \ell^*)$ be the maximum utility for this consumer under the budget constraint, we express U^* in terms of the exogenous variables. To find the derivative of U^* with respect to w, we could differentiate with the chain rule but it is quicker to use the envelope theorem. Since $L^* = L(\lambda^*, x^*, y^*, \ell^*) = U^*$, we instead differentiate L with respect to w to get quickly

$$\frac{\partial L}{\partial w} = \lambda(T - \ell)$$

and then replace endogenous variables with their optimal values to get

$$\frac{\partial U^*}{\partial w} = \frac{\partial L^*}{\partial w} = \lambda^*(T - \ell^*).$$

Thus

$$\lambda^* = \frac{\partial U^*}{\partial w} \frac{1}{T - \ell^*}. \tag{3.1}$$

Thus the Lagrange multiplier may be interpreted as the marginal utility of wage rate per hour of work. A further refinement of this interpretation will appear in the exercise set.

We return to the question of computing the twelve comparative statics. Using the implicit function theorem, we have

$$F_y \begin{bmatrix} \frac{\partial \lambda^*}{\partial p_1} & \frac{\partial \lambda^*}{\partial p_2} & \frac{\partial \lambda^*}{\partial w} \\ \frac{\partial x^*}{\partial p_1} & \frac{\partial x^*}{\partial p_2} & \frac{\partial x^*}{\partial w} \\ \frac{\partial y^*}{\partial p_1} & \frac{\partial y^*}{\partial p_2} & \frac{\partial y^*}{\partial w} \\ \frac{\partial \ell^*}{\partial p_1} & \frac{\partial \ell^*}{\partial p_2} & \frac{\partial \ell^*}{\partial w} \end{bmatrix} = - \begin{bmatrix} -x^* & -y^* & T - \ell^* \\ -\lambda^* & 0 & 0 \\ 0 & -\lambda^* & 0 \\ 0 & 0 & -\lambda^* \end{bmatrix}. \tag{3.2}$$

We wish first to find and analyze the comparative static for the impact on x^* of an increase in its own price p_1. This comparative static is in row 2, column 1 of the comparative static matrix. Solving by Cramer's rule, the first column on the right of (3.2) is used and we get

$$\frac{\partial x^*}{\partial p_1} = \frac{x^* C_{12} + \lambda^* C_{22}}{\det F_y},$$

where C_{12} is the cofactor obtained by removing the first row and second column of F_y and C_{22} is the cofactor obtained by removing the second row and second column of F_y. Computing, we obtain

$$C_{22} = -p_2^2 U_{\ell\ell} - w^2 U_{yy} + 2wp_2 U_{\ell y}$$

and

$$C_{12} = p_1 \left[U_{yy} U_{\ell\ell} - U_{\ell y}^2 \right] - p_2 \left[U_{xy} U_{\ell\ell} - U_{\ell y} U_{x\ell} \right] + w \left[U_{xy} U_{y\ell} - U_{yy} U_{x\ell} \right].$$

Our assumptions on the utility function tell us that $C_{22} > 0$. If we assume that $U_{xy} \geq 0$ (X and Y are complements or unrelated), the second and third terms of C_{12} are unambiguously positive. If we accept the article of faith that $U_{\ell\ell}U_{yy}-U_{\ell y}^2 \geq 0$ as economists generally do, then $C_{12} > 0$. Since λ^* and x^* are positive, we use the second order condition to sign $\det F_y < 0$ and we conclude that this comparative static is unambiguously negative, in agreement with intuition. (You will find some interesting comments about the article of faith in the paper from which this discussion is taken.)

We now use Cramer's rule to find the comparative static showing the impact on ℓ^* of an increase in the wage rate w. We get

$$\frac{\partial \ell^*}{\partial w} = \frac{\lambda^* C_{44} - (T-\ell^*)C_{14}}{\det F_y}, \tag{3.3}$$

where

$$C_{44} = -\left(p_1 U_{xx} + p_2 U_{yy}\right) + 2p_1 p_2 U_{xy}$$

and

$$C_{14} = p_1\left(U_{yx}U_{\ell y} - U_{\ell x}U_{yy}\right) - p_2\left(U_{xx}U_{\ell y} - U_{\ell x}U_{xy}\right) + w\left(U_{xx}U_{yy} - U_{xy}^2\right).$$

Continuing to assume that $U_{xy} \geq 0$, the assumptions of diminishing returns and that leisure is a complement of each good imply that $C_{44} > 0$. Further, the first and second terms in C_{14} are positive. Clinging to the article of faith, we see that $C_{14} > 0$. Thus the sign of this comparative static is ambiguous.

However, we take a closer look at this result. The first term in the numerator on the right of (3.3) is positive; since the denominator is negative (the second order condition), we see that the effect of this term is downward. This term is called the *substitution effect* of a wage increase: when the wage rate increases, there is a strong incentive to work longer hours, thus choosing less leisure time. The second term in the numerator of (3.3) is negative, the effect of this term is thus upward. This term is called the *income effect*: a wage hike increases income and allows our consumer to work less hours, opting for more leisure time, without suffering a loss of income. In general, a wage increase gives rise to both an income effect and a substitution effect which have opposite influences on the individual's choice.

We can gain more insight into (3.3) if we substitute for λ^* from (3.1) and use a little algebra to get

$$\frac{\partial \ell^*}{\partial w} = \frac{C_{44}\frac{\partial U^*}{\partial w} - (T-\ell^*)^2 C_{14}}{(T-\ell^*)\det F_y}.$$

From this equation, you can see that if either the individual is working long hours ($T-\ell^*$ is large) or the marginal effect of an increase in wage rate on U^* is small, then the income effect dominates and leisure time will rise is the wage rate increases. In the opposite situation, the substitution effect dominates and leisure time will go down. In other words, relatively high income, whether because of a high wage rate and/or long hours of work, tends to lead a person to opt for more leisure time when w rises.

You should see from this problem that such mathematical models in economics are rich in interesting conclusions. Although we cannot devote further time to it, you will find other interpretations, particularly regarding elasticities, in the paper mentioned earlier.

3.4 Exercises

The questions below refer to the problem and notation of the previous section.

1. Find the comparative static giving the impact on x^* of an increase in p_2. Can you sign this comparative static?

2. Find the comparative static giving the impact on the maximum utility U^* of an increase in p_1. What is it's sign? (Hint: Use the envelope theorem.)

3. Use the following steps to gain a second interesting interpretation of the Lagrange multiplier λ^*.

   (a) Let $B = w(T - \ell^*)$. Since ℓ^* is a function of the exogenous variables p_1, p_2, w, this equation expresses B as a function of these exogenous variables. Differentiate both sides of this equation with respect to B to derive

   $$(T - \ell^*)(\epsilon_S + 1)\frac{\partial w}{\partial B} = 1,$$

   where

   $$\epsilon_S = \frac{w}{T - \ell^*}\frac{\partial (T - \ell^*)}{\partial w}$$

   is called the individual's elasticity of supply of labor with respect to wage rate.

   (b) Assuming $B = w(T - \ell^*(p_1, p_, w))$ can be solved for w in terms of p_1, p_2, B, you can think of w as a function of these three variables. Thus, you can view $U^* = U^*(p_1, p_2, w)$ as a function of p_1, p_2, B and differentiate both sides with respect to B and then use the result of part (a) and (3.1) to obtain

   $$\lambda^* = (\epsilon_S + 1)\frac{\partial U^*}{\partial B}.$$

   Interpret this equation.

3.5 Utility Maximization in an n Good World

Hands, Sect. 8.5.

The previous sections were designed primarily to prepare for this problem. Not only will we need second order conditions for a problem with many variables, but we will be introduced to the concept of dual problems in economics. We will try to understand in detail the relationship between utility maximization and cost minimization. We will begin with the usual utility maximization problem of which the problem in Section 3.1 is the special case of two goods. We will call it Problem I to distinguish it from the dual cost minimization problem considered later.

Problem I - Consumer Maximizes Utility

$$\text{Maximize} \quad U(x)$$

$$\text{subject to} \quad p \cdot x = \sum_{i=1}^{n} p_i x_i = \bar{B}.$$

Here $x = (x_1, x_2, \cdots, x_n)$ are the quantities of n goods $X_1, X_2, \cdots, X_n$ from which our individual consumer makes choices, U is the utility function for that individual, and p is the price vector. The unit price of X_i is p_i, and $\bar{B}$ is the total funds available to our consumer.

The Lagrangian is

$$L(\lambda, x) = U(x) + \lambda(\bar{B} - p \cdot x).$$

The first order necessary conditions which the optimal x^* and its corresponding λ^* must satisfy are

$$\bar{B} - p \cdot x^* = 0,$$

$$U_i(x^*) - \lambda^* p_i = 0, \qquad \text{for} \quad i = 1, 2, \cdots, n.$$

If the hypotheses of the implicit function theorem are satisfied, the optimal values (λ^*, x^*) are functions of the exogenous variables $(p, \bar{B})$. Thus we have

$$\begin{aligned} \lambda^* &= \lambda^*(p, \bar{B}), \\ x_k^* &= x_k^*(p, \bar{B}), \qquad \text{for} \quad k = 1, 2, \cdots, n, \\ U^* &= U(x^*) = U(x^*(p, \bar{B})). \end{aligned}$$

The functions $x_k^*(p, \bar{B})$ are called the consumer demand functions. To use the implicit function theorem, you need as usual to know that the matrix F_y for this problem is invertible. If you assume the second order conditions for this problem hold at the maximum (as always, these conditions are sufficient but not necessary), you will have free of charge that F_y is invertible and you will know the sign of its determinant (which tells you the sign of the denominators of all comparative statics). You will also know the signs of many other determinants but none of these will tell you the sign of the numerator of any comparative static. In case $n = 2$, we succeeded in Section 3.1 to sign these numerators (at least in the case the goods were complements or unrelated) for "own price" comparative statics because we could expand the numerator determinants. An exercise in Section 3.2 should have convinced you that when $n > 2$, things are not so simple. I hope you can imagine how the complexity increases as n gets larger. The difficulty with signing the numerators of the "own price" comparative statics is caused by the fact that the relevant columns in the matrix F_x for this problem have two nonzero entries. Thus when the numerator determinant is expanded down that column, one gets two terms in the numerator. Although second order conditions sign one of these terms, the other rejects all efforts to sign.

Thus, it is a fact that for $n > 2$, none of the comparative statics $\partial x_k^*/\partial p_j$, $\partial x_k^*/\partial \bar{B}$ can be signed.

We now pass to the dual problem of cost minimization. It is a great stroke of fortune that comparative statics in the dual problem are more cooperative.

Problem II - Consumer Minimizes Expenditures

$$\text{Minimize} \quad E = p \cdot x = \sum_{i=1}^{n} p_i x_i$$

$$\text{subject to} \quad U(x) = \bar{U}.$$

Now the Lagrangian is

$$L(\lambda, x) = p \cdot x + \lambda(\bar{U} - U(x)),$$

and the first order conditions satisfied by the optimal choices $(\hat{\lambda}, \hat{x})$ are

$$\bar{U} - U(\hat{x}) = 0,$$

$$p_i - \hat{\lambda} U_i(\hat{x}) = 0, \qquad \text{for} \quad i = 1, 2, \cdots, n.$$

If the hypotheses of the implicit function theorem are satisfied, the optimal values $(\hat{\lambda}, \hat{x})$ are again functions of the exogenous variables $(p, \bar{U})$. Thus we have

$$\begin{aligned}
\hat{\lambda} &= \hat{\lambda}(p, \bar{U}), \\
\hat{x}_k &= \hat{x}_k(p, \bar{U}), \qquad \text{for} \quad k = 1, 2, \cdots, n, \\
\hat{E} &= E(p, \bar{U}) = p \cdot \hat{x}.
\end{aligned}$$

The functions $\hat{x}_k(p, \bar{U})$ are called the compensated demand functions. To see how fortunate we are, let's calculate, using the generic notation of the implicit function theorem,

$$F_y = \begin{bmatrix} 0 & -U_1(\hat{x}) & -U_2(\hat{x}) & \cdots & -U_n(\hat{x}) \\ -U_1(\hat{x}) & -\hat{\lambda} U_{11}(\hat{x}) & -\hat{\lambda} U_{12}(\hat{x}) & \cdots & -\hat{\lambda} U_{1n}(\hat{x}) \\ -U_2(\hat{x}) & -\hat{\lambda} U_{21}(\hat{x}) & -\hat{\lambda} U_{22}(\hat{x}) & \cdots & -\hat{\lambda} U_{2n}(\hat{x}) \\ \vdots & \vdots & \vdots & & \vdots \\ -U_n(\hat{x}) & -\hat{\lambda} U_{n1}(\hat{x}) & -\hat{\lambda} U_{n2}(\hat{x}) & \cdots & -\hat{\lambda} U_{nn}(\hat{x}) \end{bmatrix}$$

and

$$F_x = \begin{bmatrix} 0 & 0 & \cdots & 0 & 1 \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix}.$$

You already see the good fortune: each column in F_x has only one nonzero entry. Thus when we solve by Cramer's rule and expand down one of these columns in the numerator, there will only be one nonzero term. We will see details momentarily. If you write out the comparative static matrix $H'(x)$, you will see that the bottom row contains the comparative statics for $\hat{x}_n$ and the next to last column has the particular comparative static $\partial \hat{x}_n / \partial p_n$; this gives the impact on $\hat{x}_n$ of an increase in its own price. Cramer's rule gives

$$\begin{aligned}
\frac{\partial \hat{x}_n}{\partial p_n} &= \frac{\begin{vmatrix} 0 & -U_1(\hat{x}) & -U_2(\hat{x}) & \cdots & -U_{n-1}(\hat{x}) & 0 \\ -U_1(\hat{x}) & -\hat{\lambda} U_{11}(\hat{x}) & -\hat{\lambda} U_{12}(\hat{x}) & \cdots & -\hat{\lambda} U_{1,n-1}(\hat{x}) & 0 \\ -U_2(\hat{x}) & -\hat{\lambda} U_{21}(\hat{x}) & -\hat{\lambda} U_{22}(\hat{x}) & \cdots & -\hat{\lambda} U_{2,n-1}(\hat{x}) & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ -U_n(\hat{x}) & -\hat{\lambda} U_{n1}(\hat{x}) & -\hat{\lambda} U_{n2}(\hat{x}) & \cdots & -\hat{\lambda} U_{n,n-1}(\hat{x}) & -1 \end{vmatrix}}{\det(F_y)} \\
&= \frac{-\det H_{B,n}}{\det H_{B,n+1}},
\end{aligned}$$

where $H_{B,n+1} = F_y$ and $H_{B,n}$ is the upper left $n \times n$ corner of F_y. This notation is from Section 3.7, where we will consider second order conditions for these large constrained optimization problems. The letter "H" is for Hessian, "B" is for border, and "n" is for the size, so $H_{B,n}$ is a bordered Hessian of size $n+1$. The second order conditions for a minimum in such a constrained problem in Section 3.7 will tell us that both of these determinants are negative, if $n \geq 3$. So for $n \geq 3$, our comparative static is negative, *if* we assume that the second order conditions hold at the minimum. For the case $n = 2$ the second order conditions do not sign the numerator, but you will verify in a homework problem that the numerator is easily calculated and its sign is obvious. Thus for the "consumer minimizes expenditures" problem, the second order conditions sign all the "own" price comparative statics. Unfortunately, this is not the case for the "cross" price comparative statics. Even though each of the columns in F_x have only one nonzero entry, the situation is this: when you use Cramer's rule to solve for any "cross" comparative static, the numerator determinant does not have the lone -1 on the diagonal. Thus when you expand down that column, you do not get one of the appropriate principal minors of the denominator and the second order conditions do not give signs for non-principal minors. In particular, $\partial \hat{x}_n/\partial \bar{U}$ cannot be signed in general, but you will verify in a homework problem that for a two good world with complementary goods, this comparative static is positive. The same is true for $\partial \hat{x}_1/\partial p_2$ in a two good world.

So for the "consumer minimizes expenditures" problem, it is a fact that the second order conditions sign all the "own" price comparative statics, but none of the others. But this is certainly an improvement over the situation for the "consumer maximizes utility" problem.

Now we pass to a discussion of the connection between Problems I and II. This connection is the reason for calling them dual problems.

Connection between Problems I and II:

At first, there is no discernible relationship between the comparative statics

$$\frac{\partial x_k^*}{\partial p_j} \quad \text{and} \quad \frac{\partial \hat{x}_k}{\partial p_j} \tag{3.4}$$

that we seek in Problems I and II, respectively. The functions x_k^* in Problem I are functions of the exogenous variables $(p, \bar{B})$ while the functions $\hat{x}_k$ in Problem II are functions of the exogenous variables $(p, \bar{U})$ and these two sets of functions are obtained by solving two different systems of first order conditions. There is certainly no reason to think the comparative statics in the two problems are equal. Our goal now is to discover a relationship between the two sets of comparative statics. This relationship was first dicovered by Slutsky and gives important insight into income and substitution effects, and in the case of the "own-price" price compartive static allows us to take advantage of our ability to sign the ones in Problem II.

Here is the connection. If we solve Problem II for a given utility level $\bar{U}$ and find $\hat{E} = E(p, \bar{U})$ (the minimum expenditure) and then solve Problem I for the budget level $\bar{B} = \hat{E}$, we will get the *same* optimal values:

$$\begin{aligned} x_k^*(p, \bar{B}) &= \hat{x}_k(p, \bar{U}), \qquad \text{for} \quad k = 1, 2, \cdots, n \\ U^* &= U(x^*) = \bar{U}. \end{aligned}$$

Thus the consumer will buy the same amount for each good, spend the same amount of money, and obtain the same level of utility in both scenarios.

You should think about this last paragraph until you are convinced of its truth.

The main step is to differentiate both sides of the equation $\hat{x}_k(p, \bar{U}) = x_k^*(p, \bar{B})$ with respect to p_j. The left side depends on p_j in only one variable; the right side depends on p_j directly via the jth variable and indirectly via the last variable $\bar{B} = \hat{E} = E(p, \bar{U})$ which depends on p_j. So, differentiating with the chain rule gives

$$\frac{\partial \hat{x}_k}{\partial p_j} = \frac{\partial x_k^*}{\partial p_j} + \frac{\partial x_k^*}{\partial \bar{B}} \frac{\partial \bar{B}}{\partial p_j}.$$

This is almost Slutsky's equation. We just need to make a simplification. Since $\bar{B} = \hat{E} = p \cdot \hat{x}$ and $\hat{E} = L(\hat{\lambda}, \hat{x})$ (the Lagrangian for Problem II) and $(\hat{\lambda}, \hat{x})$ is a critical point of this Lagrangian $L(\lambda, x)$, we can find $\partial \bar{B}/\partial p_j = \partial \hat{E}/\partial p_j$ from the envelope theorem. (Note: since we only know that a minimum for $E(x, p)$ is one of the critical points of the Lagrangian, not necessarily a maximum or a minimum of the Lagrangian, we need to know here that the envelope theorem works for any critical point, not just maxima and minima.) We have

$$L(\lambda, x) = p \cdot x + \lambda(\bar{U} - U(x)),$$

so

$$\frac{\partial L}{\partial p_j} = x_j,$$

and evaluating at $(\lambda, x) = (\hat{\lambda}, \hat{x})$, we get

$$\frac{\partial \bar{B}}{\partial p_j} = \frac{\partial \hat{E}}{\partial p_j} = \hat{x}_j.$$

Thus

$$\frac{\partial \hat{x}_k}{\partial p_j} = \frac{\partial x_k^*}{\partial p_j} + \hat{x}_j \frac{\partial x_k^*}{\partial \bar{B}},$$

and using $\hat{x}_j = x_j^*$, we solve this last equation to get

$$\frac{\partial x_k^*}{\partial p_j} = \frac{\partial \hat{x}_k}{\partial p_j} - x_j^* \frac{\partial x_k^*}{\partial \bar{B}}.$$

Now we have the main Slutsky equation. The first term on the right is called the substitution effect; the second is called the income effect. Note that the substitution effect measures the impact on the quantity of the kth good purchased when the price of the jth good increases in the "consumer minimizes expenditures" problem, while the income effect measures the impact on the quantity of the kth good purchased when the available income goes up in the "consumer maximizes utility" problem, scaled by the amount of the jth good purchased.

The second Slutsky equation is the "own-price" special case when $k = j$:

$$\frac{\partial x_k^*}{\partial p_k} = \frac{\partial \hat{x}_k}{\partial p_k} - x_k^* \frac{\partial x_k^*}{\partial \bar{B}}.$$

Since the first term is unambiguously negative, as we showed above, it follows that the comparative static $\partial x_k^*/\partial p_k$, which we were earlier unable to sign, is unambiguously negative, unless the kth good is an inferior good ($\partial x_k^*/\partial \bar{B}$ is negative) and the income effect outweighs the substitution effect. If this actually happens, the good is called a Giffin good.

How do we interpret the first Slutsky equation? If the kth good is a normal good, the income effect is always negative. Thus the other two partial derivatives are certainly not the same. The substitution effect could be negative or positive. If the kth good is normal, it is clear that the effect on x_k^* of an increase in the price p_j of any good is always more negative (or less positive) than the effect on $\hat{x}_k$. Think about this until it makes intuitive economic sense to you.

The results of this section are complex, properly representing the truth about the complicated world of consumer behavior. Try to digest the difference between the two basic problems we have discussed. Our conclusions in this section are among the most prized in the theory of consumer behavior.

3.6 Exercises

1. Suppose our consumer in a 3 good world has the utility function

$$U(x_1, x_2, x_3) = a_1 \ln(x_1) + a_2 \ln(x_2) + a_3 \ln(x_3)$$

where $a_1 + a_2 + a_3 = 1$ and each a_i is positive.

   (a) Solve the first order conditions for the consumer's demand functions, assuming the consumer maximizes utility.

   (b) Find all the comparative statics for each good with respect to each price and the available budget B.

   (c) Find all the comparative statics for λ^*, and interpret.

2. For the utility function of the previous problem, assume the consumer minimizes expenditures and solve the first order conditions for the compensated demand functions. Then find a few of the more interesting comparative statics.

3. Show that the first order conditions for Problem I imply that

$$\frac{U_i(x^*)}{U_j(x^*)} = \frac{p_i}{p_j}$$

for any two different goods X_i and X_j. Do you recognize the left side as the marginal rate of substitution (MRS) and do you see that this equation states an elementary fact in principles of economics?

4. For the "consumer minimizes expenditures" problem in a two good world, verify that

   (a) the comparative static $\partial \hat{x}_2 / \partial p_2$ is negative,

   (b) the comparative static $\partial \hat{x}_2 / \partial p_1$ is positive,

   (c) the comparative static $\partial \hat{x}_2 / \partial \bar{U}$ is positive.

5. For the "consumer maximizes utility" problem:

   (a) Show that $\partial U^* / \partial \bar{B} = \lambda^*$.

   (b) Suppose the consumer's utility function is homogeneous of degree r. Show that

$$\frac{\bar{B}}{U^*} \frac{\partial U^*}{\partial \bar{B}} = r.$$

   (c) Show that

$$\frac{\partial U^*}{\partial p_j} = -x_j^* \frac{\partial U^*}{\partial \bar{B}}.$$

3.7 Mathematical Interlude

Our main purpose in this section is to discuss the first and second order conditions for constrained optimization problems. A secondary purpose is to give a simple example which will illustrate the advantage of the Lagrange multiplier method, even when there seems a more elementary approach and will also show that one needs to think rather carefully about this method for solving constrained optimization problems.

3.7.1 The Lagrange Multiplier Method

Hands, Sect. 8.1.

Before discussing the general case, we examine the case of a function of three variables and one constraint.

The problem is to find the optimum (either maximum or minimum) values of $u(x, y, z)$ subject to the constraint $g(x, y, z) = 0$.

FIRST ORDER NECESSARY CONDITIONS:

If (a, b, c) is an optimum point for the constrained problem, then there is a number λ^* so that (λ^*, a, b, c) is a critical point for the unconstrained problem of optimizing the Lagrangian function

$$L(\lambda, x, y, z) = u(x, y, z) + \lambda g(x, y, z)$$

of the four variables (λ, x, y, z). If (λ^*, a, b, c) is a critical point for this unconstrained problem, then (a, b, c) is called a critical point for the corresponding constrained problem. The critical points are thus determined by the solutions of the four simultaneous equations

$$\begin{aligned} L_\lambda &= g(x, y, z) = 0, \\ L_x &= u_x + \lambda g_x = 0, \\ L_y &= u_y + \lambda g_y = 0, \\ L_z &= u_z + \lambda g_z = 0. \end{aligned}$$

Please understand that every optimum point for $u(x, y, z)$ is found among the critical points for $L(\lambda, x, y, z)$. However, an optimum point for $u(x, y, z)$ may correspond to a saddle point for $L(\lambda, x, y, z)$. This is why the second order conditions below differ from those for the unconstrained problem.

SECOND ORDER SUFFICIENT CONDITIONS:

Suppose (a, b, c) is a critical point for the constrained problem. Let H_B be the bordered Hessian

$$H_B = \begin{bmatrix} 0 & g_x & g_y & g_z \\ g_x & L_{xx} & L_{xy} & L_{xz} \\ g_y & L_{yx} & L_{yy} & L_{yz} \\ g_z & L_{zx} & L_{zy} & L_{zz} \end{bmatrix}.$$

Note that is the Hessian for $L(\lambda, x, y, z)$. As before, let $H_{B,k}$ be the upper left $k \times k$ submatrix of H_B. That is,

$$H_{B,3} = \begin{bmatrix} 0 & g_x & g_y \\ g_x & L_{xx} & L_{xy} \\ g_y & L_{yx} & L_{yy} \end{bmatrix}, \qquad H_{B,4} = H_B.$$

Here are the second order conditions:

1. If $\det H_{B,3} < 0$, $\det H_{B,4} < 0$ when evaluated at the critical point (λ^*, a, b, c), then (a, b, c) is a relative minimum for the constrained problem.

2. If $\det H_{B,3} > 0$, $\det H_{B,4} < 0$ when evaluated at the critical point (λ^*, a, b, c), then (a, b, c) is a relative maximum for the constrained problem.

3. If both the above fail, no conclusion can be drawn.

Now we describe the general case of functions $u(x)$ of n variables $x = (x_1, x_2, \cdots, x_n)$. We also allow several constraints, but require that the number m of constraints be less than the number n of variables. The problem now is to find optimum values of $u(x)$ subject to the m constraints $g_1(x) = 0, g_2(x) = 0, \cdots, g_m(x) = 0$.

FIRST ORDER NECESSARY CONDITIONS:

If $b = (b_1, b_2, \cdots, b_n)$ is an optimum point for the constrained problem, then there exists $\lambda^* = (\lambda_1^*, \lambda_2^*, \cdots, \lambda_m^*)$ so that (λ^*, b) is a critical point of the unconstrained problem of optimizing the Lagrangian function

$$L(\lambda, x) = u(x) + \sum_{k=1}^{m} \lambda_k g_k(x)$$

of the $n + m$ variables $(x_1, x_2, \cdots, x_n, \lambda_1, \lambda_2, \cdots, \lambda_m)$. If (λ^*, b) is a critical point for this unconstrained problem, then b is called a critical point for the corresponding constrained problem. The critical points are thus determined as the solutions of the $n + m$ simultaneous equations

$$g_i(x) = 0, \qquad i = 1, 2, \cdots, m,$$

$$\frac{\partial u}{\partial x_i} + \sum_{k=1}^{m} \lambda_k \frac{\partial g_k}{\partial x_i} = 0, \qquad i = 1, 2, \cdots, n,$$

in the $m + n$ "unknowns" $x_1, x_2, \cdots, x_n, \lambda_1, \lambda_2, \cdots, \lambda_m$.

SECOND ORDER SUFFICIENT CONDITIONS:

Suppose $b = (b_1, b_2, \cdots, b_n)$ is a critical point for the constrained problem. Let H_B be the bordered Hessian

$$H_B = \begin{bmatrix} 0 & B \\ B^T & H \end{bmatrix},$$

where 0 is $m \times m$, B is $m \times n$, and H is $n \times n$; H_B is the Hessian matrix for $L(\lambda, x)$ and H may be viewed as the Hessian matrix for the $L(\lambda, x)$ with λ held constant, pretending that only the x_i's are variable. Also

$$B = \begin{bmatrix} \frac{\partial g_1}{\partial x_1} & \frac{\partial g_1}{\partial x_2} & \cdots & \frac{\partial g_1}{\partial x_n} \\ \frac{\partial g_2}{\partial x_1} & \frac{\partial g_2}{\partial x_2} & \cdots & \frac{\partial g_2}{\partial x_n} \\ \vdots & \vdots & & \vdots \\ \frac{\partial g_m}{\partial x_1} & \frac{\partial g_m}{\partial x_2} & \cdots & \frac{\partial g_m}{\partial x_n} \end{bmatrix}.$$

The top m rows and the left m columns (containing B) are called the borders. Let $H_{B,k}$ be the upper left $k \times k$ submatrix of H_B. Here are the second order conditions, which are sign conditions on $n - m$ determinants.

1. If $(-1)^m \det H_{B,k} > 0$ when evaluated at the critical point (λ^*, b), for $k = 2m + 1, 2m + 2, \cdots, m + n$, then b is a relative minimum for the constrained problem. Thus, if the number of constraints is odd, the conditions are that $\det H_{B,k} < 0$ for all $k = 2m+1, 2m+2, \cdots, m + n$. If m is even, then $\det H_{B,k} > 0$ for all $k = 2m + 1, 2m + 2, \cdots, m + n$.

2. If $(-1)^{k-m} \det H_{B,k} > 0$ when evaluated at the critical point (λ^*, b), for $k = 2m+1, 2m+2, \cdots, m+n$, then b is a relative maximum for the constrained problem. Thus, the conditions are that $\det H_{B,k}$ alternate in sign, with the largest one, $\det H_{B,n+m} = \det H_B$ being positive if n is even, negative if n is odd.

3. If both the above fail, no conclusion can be drawn.

In some books, the bordered Hessians are written with the zero block in the lower right corner and the borders at the bottom and right. The conditions can be described equally well either way, but there are subtle differences which are hard to remember. If is wise to always set up the problem the same way (we choose as above) and not get confused.

Since the variables $x_1, x_2, \cdots, x_n$ could be arranged in any order it follows that the order of the rows and columns of the H block is a matter of choice. Thus, the second order conditions can be used to sign other determinants of carefully chosen principal minors which could be obtained by rearranging rows and columns which do not affect the zero block. But be very careful!

3.7.2 A Simple Enlightening Example

It seems to be widely believed, and is used in a variety of books to obtain certain conclusions, that if (a, b) provides a maximum (or a minimum) for the constrained problem, then (λ^*, a, b) provides a maximum (or a minimum) for the Lagrangian. For example, in Hands' book, this belief is used to prove the necessity of the Kuhn-Tucker conditions for a constrained optimization problem which has inequality constraints. The belief is not justified and we give below a counter-example. It happens in the example below that the solution to a constrained minimization problem provides a saddle point for the Lagrangian. This is not unusual; in fact such an event is probably more common than the alternative.

Here is the example. It is a very good example to use to illustrate the advantage of the Lagrange method. I will point out why as I do the example. Given the parabola $x^2 = 4y$ and a point $(0, b)$, with $b > 0$, on the y-axis, we wish to find the minimum distance from the given point to the parabola. Thus our problem is

$$\min f(x,y) \equiv x^2 + (y-b)^2$$

$$\text{subject to} \quad g(x,y) \equiv 4y - x^2 = 0.$$

If one chooses to use the simplest method (not the Lagrange method), one would solve the constraint for $x^2 = 4y$, and eliminate x in the objective function to get

$$h(y) \equiv f(x,y) = 4y + (y-b)^2.$$

Minimizing this function of one variable, one differentiates and gets the necessary condition $4 + 2(y-b) = 0$ from which $y = b-2$ and $x = \pm 2\sqrt{b-2}$. An immediate cause for alarm is that this critical point does not exist if $b < 2$. But for $0 < b < 2$ the geometry clearly shows that a solution to the minimization problem must exist. Clearly something is wrong. The problem is that this approach "loses" critical points. If one starts over, using the same idea, but eliminating y rather than x from the objective function, one finds the "lost" critical point $(0, 0)$. There is a simple explanation for what has happened. It is provided by the implicit function theorem, which says that when one solves an equation like $g(x, y) = 0$ for one of the variables, say x, in terms of the other variable y, it is important that the derivative $\dfrac{\partial g}{\partial x}$ is not zero. But that partial derivative in this example is $-2x$, which is zero at the missing critical point $(0,0)$.

One of the main advantages of the Lagrange method is that it does not require solving the constraint for one of the variables, and avoids the problem encountered in this example altogether.

Let's look at the solution using the Lagrange method. The necessary conditions become

$$\begin{aligned} L_\lambda &= 4y - x^2 = 0, \\ L_x &= 2x - 2\lambda x = 0, \\ L_y &= 2(y-b) + 4\lambda = 0. \end{aligned}$$

The second condition requires that either $x = 0$ or $\lambda = 1$. In the first case, we get $y = 0$ from the first condition and $\lambda = b/2$ from the third condition. In the second case, we get $y = b - 2$ from the third condition, and then $x = \pm 2\sqrt{b-2}$ from the first condition. So we have three critical points if $b > 2$ and only one (the origin $(0,0)$) if $b \leq 2$. An evaluation of the objective function at the critical points shows that the minimum occurs at both the non-zero critical points if $b > 2$; if $b \leq 2$, the minimum must occur at $(0,0)$. Thus the method works as expected, finding all critical points, and in particular, the minimum.

Here is the main point. For $b > 2$, the two points where the minimum occurs for the constrained minimization problem actually correspond to saddle points for the Lagrangian. We verify this statement for the point in the first quadrant. That is, we will show that the point $(\lambda^*, a, b) = (1, 2\sqrt{b-2}, b-2)$ is a saddle point for $L(\lambda, x, y)$. Our method is to freeze the value of y at $b-2$ and allow the other two variables to vary around the point (λ^*, a). The resulting function of two variables is

$$h(\lambda, x) = x^2 + 4 + \lambda(4b - 8 - x^2).$$

Applying two dimensional max-min methods to this problem, we compute the first partials and equate them to zero. The resulting equations produce, as expected $\lambda = \lambda^* = 1$, $x = a = 2\sqrt{b-2}$. From elementary vector calculus, we know that this point will be a saddle point if the second partials satisfy the conditions $h_{xx}h_{yy} - (h_{xy})^2 < 0$ at the critical point (λ^*, a). Computing this quantity gives the result $-4x_0^2 = -16(b-2) < 0$. Thus we have a saddle point for the two dimensional problem.

From the nature of a saddle point, there must be a curve in the plane $y = b-2$ in (λ, x, y) space, passing through (λ^*, a, b), along which $L(\lambda, x, y)$ attains a maximum at (λ^*, a, b). Since the Lagrangian has a minimum at this point moving along the curve in the $\lambda = \lambda^* = 1$ plane defined by $x^2 = 4y$, it is clear that the point (λ^*, a, b) must also be a saddle point for the Lagrangian in three dimensions.

Chapter 4

Optimization with Inequality Constraints

We will now consider constrained optimization problems with inequality constraints. For example, we might wish to allow our familiar consumer to maximize utility without spending all available funds. Perhaps somehow, the utility function reaches a maximum and further consumption of goods is actually of negative value, somewhat like eating more when one is already gorged. In the first section, we will describe two such economics problems. Then we will pass to the mathematical tools, and finally, in our last section, we will apply the mathematical machinery to gain some interesting conclusions about our two problems.

4.1 Two Problems

Our first problem is that of utility maximization in a two good world. However, this time we will allow our consumer the possibility of encountering a maximum without spending all available funds. We will also not allow our consumer to buy negative amounts of either good. In our previous encounter with this problem, we were silent on this point. We just implicitly assumed that a maximum would not occur with negative quantities. But if one good was vastly preferred to the other, and we don't explicitly rule out purchases of negative quantities (viewed I suppose as our consumer having the ability to convert units of the despised good into cash, which can then be used to get additional units of the desirable good), we have no way of preventing such things from occurring. So we restate our "consumer maximizes utility" problem in the form

$$\begin{array}{ll} \text{maximize} & U(x,y) \\ \text{subject to} & p_1x + p_2y \leq B, \quad x \geq 0, y \geq 0 \end{array}$$

We shall return to this problem after we have discussed first order necessary conditions for such a problem. These necessary conditions are called the Kuhn-Tucker conditions.

Our second problem is the theory of a firm operating as a monopoly. We assume that this monopolist is subject to some regulatory agency, whose purpose is to limit the firm's profits to a *fair* rate of return. We shall assume that the profit of our firm is limited to a certain fixed percentage of the capital employed by the firm. The idea is to disallow the firm from increasing profits without expanding by a proportionate amount. However, there is an undesirable side effect: such a regulated firm is led to distort its use of inputs to favor the use of capital. This outcome is called the A-J effect. Our ultimate goal is to see how this effect comes about.

The revenue of our firm can be expressed as $R = py$ where p is the unit sales price of our firm's product and y is the number of units produced per period. Since the firm is a monopolist, it is a price setter. If the price is set at a level p, then demand for the good will allow us to sell a certain number of units y; clearly as we increase p, y will decrease. Thus $y = y(p)$ is a function of p. If we imagine solving this equation for p in terms of y, we can just as well think of $p = p(y)$ so that p is a function of y. This means that we could view our firm as making a decision regarding how many units to manufacture, and then accepting the market clearing price. At any rate, the output $y = f(L, K)$ is a function of the amount of labor L and capital K employed in our production and thus our revenue can be regarded as $R = yp(y) = f(L, K)p(f(L, K))$, so R is a function of K and L. We shall write this function as $R = R(L, K)$.

Thus our problem is

$$\begin{array}{ll} \text{maximize} & \Pi(L, K) = R(L, K) - wL - \nu K \\ \text{subject to} & R(L, K) - wL - \nu K \leq sK, \quad L \geq 0, K \geq 0, \end{array}$$

where s represents the fraction (percentage) of K to which the firm's profit is limited. So $s > 0$.

You easily see the similarities between these two problems. They are both constrained maximization problems, in which the constraints consist of a collection of inequalities. We shall return to the second problem also after we have prepared the mathematical tools.

4.2 Kuhn-Tucker Conditions

Hands, Sect. 9.1, 9.2.

Here we explain briefly, and only partially, why the Kuhn-Tucker conditions are valid necessary conditions for an inequality constrained optimization problem.

To gain a visual understanding we begin with the simple case of two variables and one inequality constraint. As always, we use the word *optimize* to mean either *maximize* or *minimize*.

$$\begin{array}{ll} \text{optimize} & f(x_1, x_2) \\ \text{subject to} & g(x_1, x_2) \geq 0. \end{array} \tag{4.1}$$

In order to make the scenario specific, assume that the graph of $g(x_1, x_2) = 0$ is a simple closed curve (like a circle or an ellipse) and that the points inside the curve are the points where $g(x_1, x_2) > 0$ while the points outside the curve are those where $g(x_1, x_2) < 0$. Thus, we are seeking the point x^* inside or on the curve which gives the optimum value of $f(x_1, x_2)$. There are two possibiliteis: either the opitmum occurs inside the curve or on the curve. In the first case, the usual first order necessary conditions hold true:

$$\frac{\partial f}{\partial x_i} = 0, \qquad \text{for} \quad i = 1, 2.$$

Since the gradient vector $\nabla f(x^*)$ is just the vector with these two partial derivatives as components, then our necessary condition can be abbreviated as $\nabla f(x^*) = 0$).

In the second case, we say that the constraint is binding, and x^* satisfies $g(x) = 0$, the equality constraint from Chapter 3. Using the Lagrange multiplier λ, the necessary conditions there can be abbreviated as $\nabla f(x^*) + \lambda^* \nabla g(x^*) = 0$. This equation tells us that $\nabla f(x^*)$ points in either the same or the opposite directions as $\nabla g(x^*)$, depending on the sign of λ^*.

But we can say more about the second case. We can convince ourselves that in this case, it must be true that $\lambda^* \geq 0$. Recall that in vector calculus, we learned that $\nabla f(x^*)$ is perpendicular to the level set $f(x) = f(x^*)$ and points in the direction of maximum increase of $f(x)$, the "uphill" direction. Since $g(x) > 0$ inside the curve and $g(x) < 0$ outside the curve, then uphill for g must be toward the interior. If uphill for $f(x)$ pointed into the interior also, then there would be points inside the curve giving greater values for f than its value at x^* on the curve. So $\nabla f(x^*)$ could not point into the interior if its maximum is on the curve. Hence $\nabla f(x^*)$ must point in the opposite direction as $\nabla g(x^*)$, and necessarily $\lambda^* \geq 0$

Similar thought shows that if $f(x^*)$ is the minimum value of f for all points satisfying the *inequality* constraint $g(x) \geq 0$, then $\lambda^* \leq 0$.

Thus, for the binding case, there exists a number λ^* (nonnegative for a maximum, nonpositive for a minimum) so that

$$\frac{\partial f}{\partial x_i} + \lambda^* \frac{\partial g}{\partial x_i} = 0, \quad \text{for} \quad i = 1, 2.$$

This statement of necessary conditions for the binding case can be used in both cases by the simple expedient of letting $\lambda^* = 0$ in the nonbinding case. Thus the first order conditions take the form:

There exists a number λ^* (nonnegative for a maximum, nonpositive for a minimum) so that

$$\frac{\partial f}{\partial x_i} + \lambda^* \frac{\partial g}{\partial x_i} \quad = \quad 0, \quad \text{for} \quad i = 1, 2$$

$$\lambda^* g(x^*) \quad = \quad 0.$$

These are the Kuhn-Tucker conditions for problem (4.1) and may be used to determine the three variables (λ, x_1, x_2)

Next consider optimization problems with more variables and with several inequality constraints. We shall suppose in our discussion that the number m of constraints is less than the number n of variables. So let $x = (x_1, x_2, \cdots, x_n)$. The optimization problem now is

$$\begin{array}{lll} \text{optimize} & f(x) & \\ \text{subject to} & g_k(x) \geq 0, \quad \text{for} & k = 1, 2, \cdots, m. \end{array} \tag{4.2}$$

Let x^* be the point where the optimum occurs. It may be that $g_k(x^*) > 0$ for each k. Or it may happen that some of the constraints are binding.

(i) Suppose all the constraints are binding. Then it can be shown, as in the case where there is only one constraint, that there are numbers $\lambda_1^*, \lambda_2^*, \cdots, \lambda_m^*$ (all nonnegative at a maximum, nonpositive at a minimum) so that

$$\nabla f(x^*) + \sum_{i=1}^{m} \lambda_i^* \nabla g_i(x^*) = 0.$$

This is a vector equation with n components, so there are actually n equations here. These n equations together with the m constraints, which are equations in this case, are the $n+m$ first order necessary conditions, which determine the $n+m$ values of x^* and the λ^*'s. These are the same as the conditions in Chapter 3 for equality constraints, except in this case we know the signs of the λ^*'s.

(ii) Suppose at least one of the constraints is binding. Assume that p of the constraints are binding. We can assume that we have arranged the constraints in such an order that the first p constraints are binding. Hence x^* is at least a relative optimum for the problem

$$\begin{array}{lll} \text{optimize} & f(x) & \\ \text{subject to} & g_k(x) = 0, \quad \text{for} & k = 1, 2, \cdots, p. \end{array}$$

Case 1 results then assure us that there exist numbers $\lambda_1^*, \lambda_2^*, \cdots, \lambda_p^*$ (all nonnegative at a maximum, nonpositive at a minimum) for which

$$\nabla f(x^*) + \sum_{i=1}^{p} \lambda_i^* \nabla g_i(x^*) = 0.$$

These n equations together with the p binding constraints determine the $n+p$ values of x^* and the λ^*'s.

(iii) Suppose none of the constraints are binding. In this case, x^* is a (at least relative) optimum for the unconstrained problem. Thus we must have

$$\nabla f(x^*) = 0.$$

These n equations determine the n values of x^*.

Thus we have the needed necessary conditions. The only problem is that they are difficult to use since we do not generally know which case we are in. We get around this problem by the following clever trick. For any nonbinding constraints $g_k(x) \geq 0$ in Case 2 or Case 3, we can let $\lambda_k^* = 0$ and then in all three cases the necessary conditions read

$$\nabla f(x^*) + \sum_{i=1}^{m} \lambda_i^* \nabla g_i(x^*) = 0. \tag{4.3}$$

By adding the conditions

$$\lambda_i^* g_i(x^*) = 0, \qquad \text{for} \quad i = 1, 2, \cdots, m, \tag{4.4}$$

we have $n + m$ conditions to determine the $n + m$ values of x^* and the λ^*'s. Note that the conditions in (4.4) are necessary since they clearly hold for both binding and nonbinding constraints. Do not forget that all λ^*'s are nonnegative for a maximum and nonpositive for a minimum.

These are the Kuhn-Tucker first order necessary conditions for (4.2). They are true even if $m \geq n$, which is fortunate since we encounter this situation immediately.

In economics problems, the x_k often denote quantities of goods or inputs. Thus in optimization problems, it is natural to have constraints like $x_k \geq 0$ for all k. These are especially simple inequality constraints, so we reformulate the Kuhn-Tucker conditions for such a problem. In order not to be awkward, we focus on the case where our optimization problem is a maximization problem. (The alternative case of a minimization problem should be clear.)

$$\begin{array}{lll} \max & f(x) & \\ \text{subject to} & g_k(x) \geq 0, \quad \text{for} \quad k = 1, 2, \cdots, m, & x \in R_+^n. \end{array} \tag{4.5}$$

The notation $x \in R_+^n$ is just shorthand for the n inequalities $x_k \geq 0$ for $k = 1, 2, \cdots, n$. So there are $m + n$ inequality constraints in the above problem: more constraints than variables. From the above discussion, the first order Kuhn-Tucker conditions then require that there exist non-negative numbers $\lambda_1^*, \lambda_2^*, \cdots, \lambda_m^*$ and $\mu_1^*, \mu_2^*, \cdots, \mu_n^*$ for which

$$\nabla f(x^*) + \sum_{k=1}^{m} \lambda_k^* \nabla g_k(x^*) + \sum_{j=1}^{n} \mu_j^* e_j = 0,$$

and $\lambda_k^* g_k(x^*) = 0$, for $1 \leq k \leq m$, and $\mu_j^* x_j^* = 0$ for $1 \leq j \leq n$. We have simplified the last displayed equation using the fact that the gradient vector for $h_k(x) = x_k$ is just the kth standard basis vector e_k from linear algebra.

It is common to express these last conditions in a slightly different way. The last displayed equation implies that

$$\nabla f(x^*) + \sum_{k=1}^{m} \lambda_k^* \nabla g_k(x^*) \leq 0, \tag{4.6}$$

(i.e., each component of the vector on the left is non-positive). Further, if $\mu_j^* > 0$ for some j, then (since $\mu_j^* x_j^* = 0$) $x_j^* = 0$. On the other hand, if $\mu_j^* = 0$ for some j, then the jth component of (4.6) satisfies

$$\frac{\partial f}{\partial x_j}(x^*) + \sum_{k=1}^{m} \lambda_k^* \frac{\partial g_k}{\partial x_j}(x^*) = 0.$$

Thus we see that it is always the case that every component of (4.6) satisfies

$$\frac{\partial f}{\partial x_j}(x^*) + \sum_{k=1}^{m} \lambda_k^* \frac{\partial g_k}{\partial x_j}(x^*) \leq 0, \tag{4.7}$$

and

$$x_j^* \left[\frac{\partial f}{\partial x_j}(x^*) + \sum_{k=1}^{m} \lambda_k^* \frac{\partial g_k}{\partial x_j}(x^*) \right] = 0. \tag{4.8}$$

Also, we know that

$$\lambda_k^* \geq 0 \tag{4.9}$$

and

$$\lambda_k^* g_k(x^*) = 0. \tag{4.10}$$

Further, we know the constraints

$$g_k(x^*) \geq 0, \tag{4.11}$$

and

$$x_j^* \geq 0 \tag{4.12}$$

are satisfied. The conditions (4.7), (4.8), (4.12) holding for all $j = 1, 2, \cdots, n$ and (4.9), (4.10), (4.11) holding for all $k = 1, 2, \cdots, m$ are the usual Kuhn-Tucker first order necessary conditions for a maximum in (4.5). If (4.5) is changed to find the minimum, then the Kuhn-Tucker conditions are the same except that the inequalities in (4.7) and (4.9) are reversed.

4.3 Analysis of the Two Problems

Hands, Sect. 9.3.

We return to the two problems introduced in Section 4.1 and apply the Kuhn-Tucker conditions from Section 4.2 to see what conclusions can be drawn.

4.3.1 Utility Maximization

We restate the utility maximization problem from Section 4.1:

$$\begin{array}{ll} \text{maximize} & U(x, y) \\ \text{subject to} & B - p_1 x - p_2 y \geq 0, \quad x \geq 0, y \geq 0. \end{array}$$

Applying the Kuhn-Tucker conditions, the optimal values λ^*, x^*, y^* satisfy the requirements

$$\frac{\partial U}{\partial x} - \lambda^* p_1 \leq 0,$$

$$\frac{\partial U}{\partial y} - \lambda^* p_2 \leq 0,$$

$$x^* \left[\frac{\partial U}{\partial x} - \lambda^* p_1 \right] = 0,$$

$$y^* \left[\frac{\partial U}{\partial y} - \lambda^* p_2 \right] = 0,$$

$$B - p_1x^* - p_2y^* \geq 0,$$

$$\lambda^* [B - p_1x^* - p_2y^*] = 0,$$

$$x^* \geq 0, \quad y^* \geq 0, \quad \lambda^* \geq 0.$$

These conditions do not allow both x^* and y^* to be zero. We analyze the two remaining cases. First suppose that neither x^* or y^* is zero. Then the Kuhn-Tucker conditions imply that

$$\frac{\partial U}{\partial x} - \lambda^* p_1 = 0, \qquad \frac{\partial U}{\partial y} - \lambda^* p_2 = 0,$$

both of which imply that $\lambda^* > 0$ and therefore that $B - p_1x^* - p_2y^* = 0$, so we are back to equality constraints so the conclusions in Section 3.1 hold.

So now suppose that $x^* > 0$, $y^* = 0$. (The other case is handled similarly.) Now the Kuhn-Tucker conditions tell us that

$$\frac{\partial U}{\partial y} \leq \lambda^* p_2,$$

$$\frac{\partial U}{\partial x} = \lambda^* p_1,$$

so $\lambda^* > 0$. Hence also $p_1x^* = B$ so all available funds are being used. Our conditions imply that

$$\frac{U_x(x^*, y^*)}{U_y(x^*, y^*)} \geq \frac{p_1}{p_2}.$$

You will recognize the economics interpretation of the left side of this inequality; it is the marginal rate of substitution of X for Y. So in the case that utility maximization leads our consumer to buy no units of Y, then this marginal rate of substitution is greater than or equal to the price ratio p_1/p_2.

It was implicit in our treatment of the two good world in Section 3.1 that we were assuming that the consumer was buying positive amounts of both goods. Using the Kuhn-Tucker conditions, we have been able to extend our results to the case where one of the goods is ignored by the consumer, and this could not happen unless the marginal rate of substitution satisfies the above inequality. However, if this is our consumer's maximizing behavior, we are no longer able to make any comparative statics conclusions. We cannot apply the implicit function theorem to the inequalities which are now our first order conditions.

4.3.2 Rate of Return Regulation

We now consider the second problem from Section 4.1:

$$\begin{array}{ll} \text{maximize} & \Pi(L, K) = R(L, K) - wL - \nu K \\ \text{subject to} & (s + \nu)K + wL - R(L, K) \geq 0 \quad L \geq 0, K \geq 0, \end{array}$$

where $s > 0$.

The Kuhn-Tucker necessary conditions for this problem are

$$R_L - w + \lambda^*(w - R_L) \leq 0,$$

$$L^* [R_L - w + \lambda^*(w - R_L)] = 0,$$

$$R_K - \nu + \lambda^*(s + \nu - R_K) \leq 0,$$
$$K^* \left[R_K - \nu + \lambda^*(s + \nu - R_K)\right] = 0,$$
$$(s + \nu)K^* + wL^* - R(L^*, K^*) \geq 0,$$
$$\lambda^* \left[(s + \nu)K^* + wL^* - R(L^*, K^*)\right] = 0,$$
$$\lambda^* \geq 0, \quad L^* \geq 0, \quad K^* \geq 0.$$

Let us assume that both L^*, K^* are positive. (Our experience in the utility maximization problem indicates that our conclusions would be rather slim in other cases. It also seems likely in practice that this would be the only case to consider.) Now the first and third conditions above become equalities and we can remove the second and fourth conditions. The first condition, as an equality, can be rewritten as

$$(1 - \lambda^*)(R_L - w) = 0$$

and the third, as an equality, similarly becomes

$$(1 - \lambda^*)R_K = \nu - \lambda^*(s + \nu).$$

Since $s > 0$, this last equation forces $\lambda^* \neq 1$.

Thus

$$R_L = w, \qquad R_K = \frac{\nu - \lambda^*(s + \nu)}{1 - \lambda^*}.$$

Dividing we get

$$\frac{R_K}{R_L} = \frac{\nu}{w} - \frac{\lambda^* s}{w(1 - \lambda^*)}. \tag{4.13}$$

Recall (see Section 3.1) that $R(L, K) = p(f(L, K))f(L, K)$. It is a simple differentiation problem to find R_K and R_L and verify that

$$\frac{R_K}{R_L} = \frac{f_K}{f_L}.$$

Thus we can rewrite (4.13) in the form

$$\frac{f_K}{f_L} = \frac{\nu}{w} - \frac{\lambda^* s}{w(1 - \lambda^*)}. \tag{4.14}$$

The equation (4.14) will be key to our conclusions. We are interested in what effect, if any, the regulatory agent has on the behavior of our monopolistic firm. If $\lambda^* = 0$, then the constraint is not binding and the regulation has no effect on the firm's behavior. So let us consider the case that $\lambda^* > 0$. Our firm will produce a certain maximizing output $y^* = f(L^*, K^*)$. In a homework problem you will be asked to show that if our firm produced the output y^* in a cost minimizing way, then the equation $f_K/f_L = \nu/w$ would be true. Our equation (4.14) shows that if the firm changes its behavior because of the regulation, it will change in such a way that the ratio of the marginal product of capital to the marginal product of labor is no longer equal to the ratio of the input prices. If $\lambda^* < 1$ then the last term in (4.14) is positive so the effect of the regulation is that the ratio f_K/f_L is lower than the cost minimizing ratio would be.

Let us assume that $\lambda^* < 1$ for a moment. We want to understand the conclusion more clearly. The equation $f(L, K) = y^*$ can be solved for L in terms of K. Let's differentiate $f(L, K) = y^*$ implicitly with respect to K. We get $f_L dL/dK + f_K = 0$ and thus

$$-\frac{dL}{dK} = \frac{f_K}{f_L}.$$

Thus the derivative on the left is negative and the effect of the regulation is to decrease its absolute value. This can only happen to the slope on the curve $f(L, K) = y^*$ if one moves so that K becomes larger and L becomes smaller and therefore the capital-labor ratio becomes larger. (Sketch a graph and see this!) This is the so-called A-J effect, named for H. Averch and L. L. Johnson from their 1962 paper in the American Economic Review. It means that if a monopolistic firm is subject to a regulation of the kind considered here (profits are limited to a percentage of capital employed), the firm will adjust so that it produces its output in a more capital intensive way than the cost minimizing way.

The only task left is to investigate the inequality $\lambda^* < 1$. Since $\lambda^* > 0$, the regulatory constraint is binding. Since we are assuming that both K^* and L^* are positive, these constraints are not binding. Thus our problem is equivalent to the equality constrained problem

$$\begin{array}{ll} \text{maximize} & \Pi(L, K) = R(L, K) - wL - \nu K \\ \text{subject to} & (s + \nu)K + wL - R(L, K) = 0, \end{array}$$

where $s > 0$. The Lagrangian for this problem is

$$\mathcal{L}(\lambda, L, K) = R(L, K) - wL - \nu K + \lambda[(s + \nu)K + wL - R(L, K)].$$

The first order conditions are

$$\mathcal{L}_\lambda = (s + \nu)K^* + wL^* - R(L^*, K^*) = 0,$$

$$\mathcal{L}_L = R_L - w + \lambda^*(w - R_L) = 0,$$

$$\mathcal{L}_K = R_K - \nu + \lambda^*(s + \nu - R_K) = 0.$$

In particular, the third first order condition gives, as before, $(1 - \lambda^*)R_K = \nu - \lambda^*(s + \nu)$ and the second gives $(1 - \lambda^*)R_L = (1 - \lambda^*)w$. Thus we see again that $\lambda^* \neq 1$ (since $s > 0$) so $R_L = w$ as before. Let us see what the second order conditions from Chapter 3 tell us about the optimal value of λ^*. Here $n = 2$, $m = 1$ so we have a sign condition on only $n - m = 1$ determinant (the entire bordered Hessian H_B) and its determinant is positive since $n = 2$ is even. We quickly compute, using the fact that $R_L = w$ at the optimal point,

$$H_B = \begin{bmatrix} 0 & 0 & s + \nu - R_K \\ 0 & (1 - \lambda^*)R_{LL} & (1 - \lambda^*)R_{LK} \\ s + \nu - R_K & (1 - \lambda^*)R_{KL} & (1 - \lambda^*)R_{KK} \end{bmatrix}.$$

Expanding across the top row is easy and we get

$$\det(H_B) = -(1 - \lambda^*)(s + \nu - R_K)^2 R_{LL}.$$

If $R_{LL} < 0$ when evaluated at the optimum point (L^*, K^*), then we must have $\lambda^* < 1$ to satisfy the second order condition. If $R_{LL} > 0$, then $\lambda^* > 1$ is necessary to satisfy the second order condition.

We want to make a case that $R_{LL} < 0$ so that the A-J effect in fact holds. If you take the equation $R(L, K) = p(f(L, K))f(L, K)$, differentiate it twice with respect to L (you should check me on this!) *carefully* with the chain rule, you will get something of a mess and you will find that the natural assumptions on the production function $f(L, K)$ and the natural assumptions on the behavior of the function $p(y) = p(f(L, K))$ lead to an ambiguous sign on R_{LL}. So this approach is not helpful. So we turn to economics for some guidance. Remember this firm is a monopoly. Would you expect that $R_L > 0$ and $R_{LL} < 0$? Don't jump to a conclusion. Certainly, other things being equal, increasing labor should increase production, but since more production leads to a lower market price, net revenue could conceivably drop so R_L might be negative at certain levels of production. Similarly, at certain production levels it might happen that increasing labor could not only increase revenue, but even the marginal revenue might increase so $R_{LL} > 0$ at these production levels. But only the behavior at the optimum levels of L and K is relevant. At this level, you would expect revenue to still be increasing with L but that costs are increasing sufficiently rapidly to plateau our profits at the maximum. Also, it seems natural to assume that marginal revenue (R_L) is declining at this point, consistent with the idea that the marginal cost of labor (a constant) is overtaking the marginal product of labor.

Think about this for yourself. You know more economics than I do. Is the argument above wishful thinking because we have already decided we like the conclusion (the A-J effect) or do we find it believable independently of the A-J effect?

We have completed our last problem in this study of static equilibrium and optimization in economics, and we have arrived again at a situation that is now rather familiar. Here again we cannot draw a conclusion without pinning our hopes on second order conditions, which are sufficient for a maximum, but do not necessarily hold at a maximum. Here again mathematics has helped us to be more precise in our thinking and provided us a clear path between fundamental economic assumptions and the implications which they determine, but mathematics alone cannot give conclusions about economic behavior. Careful mathematical processes and clear economic understanding must combine forces in partnership to obtain conclusions.

What makes it all so intellectually rich and fascinating is that mathematics, expected to furnish the equipment to solve economics problems, ends up just as often raising questions about the fundamental economics assumptions, while economics problems, expected to provide an arena for mathematics to exhibit its power, ends up just as often raising questions about fundamental mathematical methods, usually because the need is to run mathematical logic in reverse of its natural order.

That note is as good as any on which to end this truly hybrid effort.

4.4 Exercises

1. Apply the method of Chapter 3 to the constrained optimization problem

$$\begin{aligned} &\text{minimize} \qquad C = wL + \nu K \\ &\text{subject to} \qquad f(L, K) = \bar{y} \end{aligned}$$

   to show that the firm will employ labor and capital to satisfy the relation $f_K/f_L = \nu/w$.

2. Differentiate $R(L, K) = p(f(L, K))f(L, K)$ twice with respect to L to confirm the claim

in the text that R_{LL} cannot be signed using the natural assumptions on the production function and the function $p(y)$.

3. Suppose our monopolistic firm works under the production function $y = \sqrt{LK}$ and the demand for its product is given by $p = 8 - y$. Solve this regulated firm's profit maximization firm with $\nu = w = s = 1$ to find explicit functions for L^* and K^*, and $y*$. Then solve the cost minimization problem for the cost minimizing values of L_0 and K_0, and finally show that $K^*/L^* > K_0/L_0$ so that this firm exhibits the A-J effect.

4. For the general monopoly rate of return problem considered in the text, find the comparative static showing the effect of an increase in s on K^*. What is the sign of this comparative static?

5. Suppose our monopolist maximizes revenue instead of profit. Re-do the problem in the text to discover if this monopolist would exhibit the A-J effect. Discuss your conclusions.

6. The consumption of goods by a consumer takes time as well as money. So it is interesting to consider the utility maximization problem with both a budget and a time constraint. So use the Kuhn-Tucker conditions to solve the problem

$$\text{maximize} \qquad U(x, y) = xy$$

$$\text{subject to} \qquad p_1 x + p_2 y \leq B, \quad t_1 x + t_2 y \leq T.$$

   Here t_1 is the time it takes to consume a unit of x and t_2 is the time it takes to consume a unit of y. Analyze the case that $x^* > 0$, $y^* > 0$. If you like, you may consider the explicit values $p_1 = 1$, $p_2 = 2$, $B = 40$, $t_1 = t_2 = 1$, $t = 24$. Discuss your conclusions. Would both constraints normally be binding?

7. The previous problem is much more interesting when there are three or more goods. Give it a try with three goods and see what happens.

Bibliography

[1] Averch, H. and L. L. Johnson, Behavior of the firm under regulatory constraint, American Economic Review 52 (1962), 1053-1069.

[2] Baxley, J. and J. Moorhouse, Lagrange Multiplier Problems in Economics, American Math Monthly 91 (1984), 404-412.

[3] Berman, A. and R. J. Plemmons, Non-negative Matrices in the Mathematical Sciences, SIAM, 1994.

[4] Hands, D. W., Introductory Mathmatical Economics, Second Edition, Oxford Univ. Press, 2004 (1st ed., Heath, 1991).

[5] Hess, Peter, Using Mathematics in Economic Analysis, Prentice-Hall, 2002.

[6] Kuhn, H. and A. Tucker, Nonlinear Programming, in Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, ed. J. Neyman, 481-492, Univ of California Press.

[7] Marsden, J. and Tromba, Vector Calculus, W. H. Freeman.

[8] Slutsky, E., On the theory of the budget of the consumer, in Readings in Price Theory, eds. G. J. Stigler and K. E. Boulding, 27-56, Chicago: Irwin.