On Finding the Top of the Hill,
or It's All One Tiny Ball.

By Kevin Hasker
ECON 225 Art for Economists (or Math for Economists)
December 19, 2015

I bet you didn't believe me when I said that Math was a minimalist art, and that therefore everything we would learn would fit into one tiny, tidy, little ball. Oh though of little faith. Really all we've been doing has been doing has been directed towards one objective: finding the top of the hill.

Finding the top of the hill is pathetically easy in one dimension, it's the combination of two insights.

1. The Ant Rule: At the top of the hill the ground is always flat ($f'(x) = 0$).

2. The slope of the hill is decreasing in every direction ($\frac{d}{dx} f'(x) = f''(x) \leq 0$).

But, when you think about it, most "hills" are actually made up of a bunch of tiny hill-ets, do you get to say you climbed Mount Everest when you get to the first point which satisfies these conditions? Boy, you are annoying, so what? I am a mathematician, what do I have to care about how *many* peaks there are? Oh, you want to be sure you have the *global* maximum? I don't care about that, I've found beautiful *local* conditions that say that at least in the area this point is a maximum—I'll leave it to you silly people to find the global maximum.[1]

Summary 1 *If $f'(x) = 0 \geq f''(x)$ then $f(x)$ is a **local maximum**, and x a **local maximizer**.*

We also care about the maximizer, to be specific consider the following extension. We have control variable, a, how is $x(a)$ going to change? One method is to solve for the explicit function $x(a)$. Yea, right, like that's going to happen very often. Of course it generally will for any problem we give you—because we will search for a function where it's true—but in reality? No. To be precise we will be maximizing $f(x, a)$ over x, and we will have the implicit function:

$$\frac{\partial f}{\partial x}(x^*(a), a) = 0$$

and we'll want to know when $x^*(a)$ exists. The answer to this question is to use the first fundamental theorem of calculus.

Theorem 2 (First Fundamental Theorem of Calculus) *If we know $f(x)$ for one x, and we know $f'(x)$ then:*

$$f(y) = f(x) + \int_x^y f'(z)\, dz$$

---

[1] Try making some assumptions that I will then proceed to laugh at and give counter-examples to.

So we use this as:

Theorem 3 (Implicit Function Theorem, 15.1) *If $\frac{\partial f}{\partial x}(x, a)$ is a C^1 function on a ball about (x_0, a_0) then if $\frac{\partial^2 f}{\partial x^2} \neq 0$ there is a function such that $x(a_0) = a_0$ and*

$$x(a) = x(a_0) + \int_{a_0}^{a} \frac{\partial f / \partial a}{\partial^2 f / \partial x^2} dz \ .$$

If we can find out that for all x, $\frac{\partial^2 f}{\partial x^2} \neq 0$, then we know that there is a global implicit function. OK, so we need $\frac{\partial^2 f}{\partial x^2} \leq 0$ for x to be a maximum, and we need $\frac{\partial^2 f}{\partial x^2} \neq 0$ in order to have an implicit function... Combined that would mean if $\frac{\partial^2 f}{\partial x^2} < 0$ then we have a global implicit function, right?[2] Isn't it interesting that this is equivalent to having a strictly concave function for f? Couldn't this be why we usually assume (strict) concavity? What is concavity and strict concavity?

Definition 4 *A function of n variables, f is*

1. ***Concave*** *if for all $x \in \mathbb{R}^n$, $y \in \mathbb{R}^n$, $x \neq y$ and $0 \leq \lambda \leq 1$*

$$f(\lambda x + (1 - \lambda) y) \geq \lambda f(x) + (1 - \lambda) f(y)$$

2. ***Strictly concave*** *if for all $x \in \mathbb{R}^n$, $y \in \mathbb{R}^n$, $x \neq y$ and $0 < \lambda < 1$*

$$f(\lambda x + (1 - \lambda) y) > \lambda f(x) + (1 - \lambda) f(y)$$

---

[2] You may notice a difference between what I wrote and what the book wrote. The reason they don't write it this was is that complicated derivative has to be evaluated at $(x(a), a)$, which is very hard. I wrote it this way so you could see the equivalence to the First Fundamental Theorem of Calculus.

In order to understand these definitions it's useful to look at a graph:
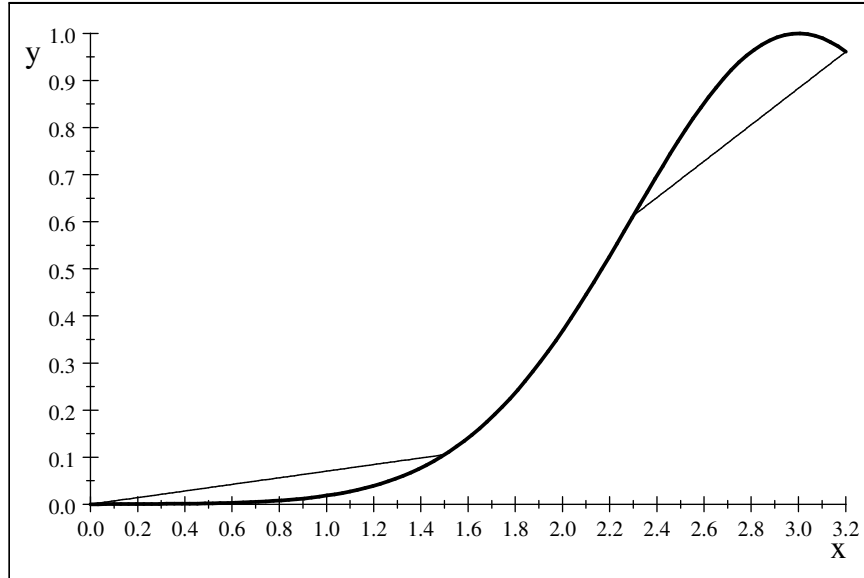
This is the function $\exp\left(-(x-3)^2\right)$ for $x \in [0, 3.2]$. It is concave when $x \geq 2.29$ and convex before that. So a line between $\exp\left(-(2.3-3)^2\right)$ and $\exp\left(-(3.2-3)^2\right)$ is always below the function because it is concave in that region and a line between $\exp\left(-(0-3)^2\right)$ and $\exp\left(-(1.5-3)^2\right)$ is always above the function because it's not concave.

Example 5 *Profit maximization over one variable.*
Let's ignore the fact that profit can't be below zero and that labor has to be demanded in a positive amount. Then the profit function is:

$$\pi(L) = pf(L) - wL \tag{1}$$

And using the ant-rule we know that:

$$\begin{aligned} pf'(L) - w &= 0 \text{ (first order condition)} \\ pf''(L) &\leq 0 \text{ (second order condition)} \end{aligned} \tag{2}$$

at any maximum. So, what can we find out about the input demand, $L(p, w)$? For this function we have the implicit function:

$$pf'(L(p, w)) - w = 0 \tag{3}$$

so we can easily use the chain rule:

$$\frac{\partial}{\partial p}\left(pf'\left(L\left(p,w\right)\right)-w\right) = \frac{\partial}{\partial p}\left(0\right) \qquad (4)$$

$$f'\left(L\left(p,w\right)\right)+pf''\left(L\left(p,w\right)\right)\frac{\partial L}{\partial p} = 0$$

$$\frac{\partial L}{\partial p} = -\frac{f'\left(L\left(p,w\right)\right)}{pf''\left(L\left(p,w\right)\right)}\ .$$

likewise for w:

$$\frac{\partial}{\partial w}\left(pf'\left(L\left(p,w\right)\right)-w\right) = \frac{\partial}{\partial w}\left(0\right) \qquad (5)$$

$$pf''\left(L\left(p,w\right)\right)\frac{\partial L}{\partial w}-1 = 0$$

$$\frac{\partial L}{\partial w} = \frac{1}{pf''\left(L\left(p,w\right)\right)}$$

which tells us that as long as $pf''\left(L\left(p,w\right)\right)\neq 0$ these derivatives exist. Combined with the second order condition $(pf''\left(L\left(p,w\right)\right)\leq 0)$ we realize that as long as $f''\left(L\left(p,w\right)\right)<0$ then these always exist, and we can also see that $\frac{\partial L}{\partial p}>0$ and $\frac{\partial L}{\partial w}<0$. Of course it's exact value will depend on the function, but that's pretty exciting because it's so general. We know that either the implicit function doesn't exist $(f''\left(L\left(p,w\right)\right)=0)$ or that input demand is increasing in output price and decreasing in input price.

1   Unconstrained Maximization with n Variables.

It's actually fairly simple to convert these two criteria for local maxima into a criterion for n variables, but unfortunately both of these conditions are now in terms of matrices. It's fairly obvious that the ground must be flat in every direction that we need to go, or $\partial f/\partial x_i = 0$ for $i = (1,2,3,...,n)$, equivalently the gradient of f is equal to zero, denoted $\bigtriangledown f = 0$. This brings us to one of the three important technical definitions for this test.

Definition 6 *The **gradient** of a function of n variables (f) is the first derivatives of f with regards to each of the variables written as a vector or column matrix, and is denoted $\bigtriangledown f$.*

$$\bigtriangledown f = \left[\frac{\partial f}{\partial x_i}\right]_{i=1...n} = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{bmatrix}$$

But what are the second order conditions? To understand this we have to remember the Taylor Series.

Definition 7 *The Taylor Series says that:*

$$f\left(x+v\right) = f\left(x\right) + v^T \bigtriangledown f\left(x\right) + \frac{1}{2}v^T D^2 f\left(x\right) v + \varepsilon \left(||v||^3\right)$$

where $D^2 f$ is the Hessian of f, and $\varepsilon\left(||v||^3\right)$ is a function which goes to zero much faster than the rest of the terms as $||v||^3 \to 0$.

This brings us to the second important technical term:

Definition 8 *The **Hessian** of a function of n variables (f) is the $n \times n$ matrix of second derivatives, and is denoted $D^2 f$.*

$$D^2 f = \left[\frac{\partial^2 f}{\partial x_i \partial x_j}\right]_{i=1...n}^{j=1...n} = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1 \partial x_1} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_n} \end{bmatrix}$$

Now if x is a local maximum we know two things. First for every $v \neq 0$ $f\left(x+v\right) \leq f\left(x\right)$, second $\bigtriangledown f\left(x\right) = 0$ (it is implicit that when I write a vector is equal to zero that the "zero" is vector of the same size—in this case both are n.) So we know that:

$$f\left(x+v\right) = f\left(x\right) + \frac{1}{2}v^T D^2 f\left(x\right) v + \varepsilon \left(||v||^3\right) \leq f\left(x\right) \tag{6}$$

for every $v \neq 0$, or:

$$\frac{1}{2}v^T D^2 f\left(x\right) v + \varepsilon \left(||v||^3\right) \leq f\left(x\right) - f\left(x\right) \tag{7}$$

and as $||v||^3 \to 0$ this means that:

$$\frac{1}{2}v^T D^2 f\left(x\right) v \leq 0 \ , \tag{8}$$

or that $D^2 f\left(x\right)$ is *negative semidefinite*. Oh no, what does this mean? Well the "semi" is a problem, just like before, so let's ignore it and focus on the negative definite, which means $\frac{1}{2}v^T D^2 f\left(x\right) v < 0$. (Just like in one dimension, we rule out the $= 0$ possibility. It's important that that's all we know, but it is such a headache in both cases.) From chapter 16 we learn that:

Definition 9 *For an $n \times n$ symmetric matrix, A, a **leading principal submatrix** (A_k) for $k = (1, 2, 3, ..., n)$ is the matrix created by dropping the last $n - k$ rows and columns.*

Definition 10 *A **leading principal minor** is $\det\left(A_k\right)$ for $k = (1, 2, 3, ..., n)$.*

Theorem 11 *Let A be an $n \times n$ symmetric matrix:*

1. *A is negative definite if and only if the sign of $\det(A_k)$ is the same as $(-1)^k$ and not equal to zero.*

2. *A is negative definite if and only if $-A$ is positive definite. $-A$ is positive definite if $\det(-A_k)$ is strictly positive.*

Summary 12 *If $f(x)$ is a local maximum then $\bigtriangledown f = 0$ and it is sufficient that $D^2 f(x)$ is negative definite.*
This means $\frac{\partial f}{\partial x_i} = 0$ and $\det\left(-D^2 f_i\right) > 0$ for $i = (1, 2, 3, ..., n)$

And I promise you that you don't really need to worry about the necessary conditions for f to be a maximum. I might ask this for an abstract question as a way to separate the star students from the crowd, but it's just too complex in general. If we assume $D^2 f(x)$ is strictly negative definite then we have n first order conditions ($\frac{\partial f}{\partial x_i} = 0$) and n second order conditions ($\det\left(-D^2 f_i\right) > 0$). Unfortunately the second order conditions are now about determinants, and that's a hideous formula, but that's the only way multiple variable optimization is more complicated than one variable optimization.

1.1 The Implicit Function Theorem

So... now we need to work up to the implicit function theorem in n variable problems. It's not actually that hard, we have the system of equations now:

$$\bigtriangledown f(x, a) = 0 \tag{9}$$

and we take the total differential of this system of equations with regards to (x, a) and get:

$$D^2 f(x, a)\, dx + \frac{\partial}{\partial a} \bigtriangledown f(x, a)\, da = 0 \tag{10}$$

$$\frac{dx}{da} = -\left[D^2 f(x, a)\right]^{-1} \frac{\partial}{\partial a} \bigtriangledown f(x, a) \tag{11}$$

Wasn't that simple? OK, so you didn't understand that at all. I don't blame you. Let's take the complete differential of the first order condition:

$$\frac{\partial f(x, a)}{\partial x_1} = 0 \tag{12}$$

$$\frac{\partial^2 f(x, a)}{\partial x_1 \partial x_1} dx_1 + \frac{\partial^2 f(x, a)}{\partial x_1 \partial x_2} dx_2 + ... + \frac{\partial^2 f(x, a)}{\partial x_1 \partial x_n} dx_n + \frac{\partial^2 f(x, a)}{\partial x_1 \partial a} da = 0$$

Now it should be clear that we have to do this with regards to each first order condition, simultaneously, or we would be goofing up. The terms multiplied by

$$dx = [dx_i]_{i=1...n} = \begin{bmatrix} dx_1 \\ dx_2 \\ \vdots \\ dx_n \end{bmatrix} \tag{13}$$

6

will then be $D^2 f(x, a)$. The terms multiplied by da will be:

$$\frac{\partial}{\partial a} \triangledown f(x, a) = \left[\frac{\partial^2 f(x, a)}{\partial x_i \partial a}\right]_{i=1\ldots n} = \begin{bmatrix} \frac{\partial^2 f(x,a)}{\partial x_1 \partial a} \\ \frac{\partial^2 f(x,a)}{\partial x_2 \partial a} \\ \vdots \\ \frac{\partial^2 f(x,a)}{\partial x_n \partial a} \end{bmatrix}. \tag{14}$$

Given these steps, as long as $D^2 f(x, a)^{-1}$ exists (or, in short, $\det\left(D^2 f(x, a)\right) \neq 0$) then

$$\frac{dx}{da} = \left[\frac{dx_i}{da}\right]_{i=1\ldots n} = \begin{bmatrix} \frac{dx_1}{da} \\ \frac{dx_2}{da} \\ \vdots \\ \frac{dx_n}{da} \end{bmatrix} \tag{15}$$

is $\frac{dx}{da} = -\left[D^2 f(x, a)\right]^{-1} \frac{\partial}{\partial a} \triangledown f(x, a)$, as stated in Equation 11. This is Theorem 15.7 from the text.

But how do we solve for $\frac{dx_i}{da}$? To do that I recommend using Crammer's Rule. I'm not saying that you can't do it by inverting the matrix if you so fancy, I'm just saying this is the standard methodology.

Theorem 13 (Cramer's Rule, 9.4) *Let A be $n \times n$ non-singular matrix ($\det(A) \neq 0$) and B_i be the matrix A with the i'th column replaced by the n element vector b. Then the solution to $Ax = b$ is*

$$x_i = \frac{\det(B_i)}{\det(A)} .$$

Here $b = -\frac{\partial}{\partial a} \triangledown f(x, a)$, $A = D^2 f(x, a)$ and $x_i = \frac{dx_i}{da}$. I hope in this setting you can understand why this is such a common tool, specifically the fact that x is a maximum means that we know at least the sign of $\det\left(D^2 f(x, a)\right)$. Please remember that it depends on how many variables you are maximizing over, if its 2 then this determinant is positive, if its three it's negative, and so on. But in essence this makes things a lot simpler. It's also not that unusual for us to care about some of the terms, like for example $\frac{\partial L}{\partial w}$.

Example 14 *Two input profit maximization.*

Again we're going to ignore that profits, labor, and capital all need to be positive (or zero) and just look at the simple case of maximizing profits over two inputs, labor and capital:

$$\pi(L, K) = pf(L, K) - wL - rK \tag{16}$$

the first order conditions are:

$$p\frac{\partial f(L, K)}{\partial L} - w = 0 \tag{17}$$

$$p\frac{\partial f(L, K)}{\partial K} - r = 0$$

Now notice that $\frac{\partial f(L,K)}{\partial L} = \frac{w}{p} > 0$ and $\frac{\partial f(L,K)}{\partial K} = \frac{r}{p} > 0$ in this situation, our Hessian is:

$$H = \begin{bmatrix} p\frac{\partial^2 f}{\partial L^2} & p\frac{\partial^2 f}{\partial L \partial K} \\ p\frac{\partial^2 f}{\partial L \partial K} & p\frac{\partial^2 f}{\partial K^2} \end{bmatrix} \tag{18}$$

and our second order conditions are that $p\frac{\partial^2 f}{\partial L^2} \leq 0$ and $\det \begin{bmatrix} p\frac{\partial^2 f}{\partial L^2} & p\frac{\partial^2 f}{\partial L \partial K} \\ p\frac{\partial^2 f}{\partial L \partial K} & p\frac{\partial^2 f}{\partial K^2} \end{bmatrix} \geq$

0 or $p^2 \left(\frac{\partial^2 f}{\partial L^2} \frac{\partial^2 f}{\partial K^2} - \left(\frac{\partial^2 f}{\partial L \partial K} \right)^2 \right) \geq 0$. (We also have, with a little analysis, that

$p\frac{\partial^2 f}{\partial K^2} \leq 0$.) Now we want to find out all the partial effects so we take the differential of this system of equations:

$$\begin{bmatrix} p\frac{\partial^2 f}{\partial L^2} & p\frac{\partial^2 f}{\partial L \partial K} \\ p\frac{\partial^2 f}{\partial L \partial K} & p\frac{\partial^2 f}{\partial K^2} \end{bmatrix} \begin{bmatrix} dL \\ dK \end{bmatrix} + \begin{bmatrix} \frac{\partial f}{\partial L} \\ \frac{\partial f}{\partial K} \end{bmatrix} dp - \begin{bmatrix} 1 \\ 0 \end{bmatrix} dw - \begin{bmatrix} 0 \\ 1 \end{bmatrix} dr = 0 \quad (19)$$

In order to find out the partial effects we have to first have $\det H = p^2 \left(\frac{\partial^2 f}{\partial L^2} \frac{\partial^2 f}{\partial K^2} - \left(\frac{\partial^2 f}{\partial L \partial K} \right)^2 \right) >$

0 and then we can see that:

$$\frac{\partial L}{\partial w} = \frac{1}{\det H} \det \begin{bmatrix} 1 & p\frac{\partial^2 f}{\partial L \partial K} \\ 0 & p\frac{\partial^2 f}{\partial K^2} \end{bmatrix} = \frac{p\frac{\partial^2 f}{\partial K^2}}{p^2 \left(\frac{\partial^2 f}{\partial L^2} \frac{\partial^2 f}{\partial K^2} - \left(\frac{\partial^2 f}{\partial L \partial K} \right)^2 \right)} \leq 0 \ (20)$$

$$\frac{\partial K}{\partial w} = \frac{1}{\det H} \det \begin{bmatrix} p\frac{\partial^2 f}{\partial L^2} & 1 \\ p\frac{\partial^2 f}{\partial L \partial K} & 0 \end{bmatrix} = -\frac{p\frac{\partial^2 f}{\partial L \partial K}}{p^2 \left(\frac{\partial^2 f}{\partial L^2} \frac{\partial^2 f}{\partial K^2} - \left(\frac{\partial^2 f}{\partial L \partial K} \right)^2 \right)}$$

who's sign depends on $\frac{\partial^2 f}{\partial L \partial K}$, like it sensibly should. We can further see that:

$$\frac{\partial K}{\partial r} = -\frac{p\frac{\partial^2 f}{\partial L^2}}{p^2 \left(\frac{\partial^2 f}{\partial L^2} \frac{\partial^2 f}{\partial K^2} - \left(\frac{\partial^2 f}{\partial L \partial K} \right)^2 \right)}, \ \frac{\partial L}{\partial r} = \frac{p\frac{\partial^2 f}{\partial K \partial L}}{p^2 \left(\frac{\partial^2 f}{\partial L^2} \frac{\partial^2 f}{\partial K^2} - \left(\frac{\partial^2 f}{\partial L \partial K} \right)^2 \right)} \leq 0$$

$$\tag{21}$$

by the same analysis, and finally that:

$$\frac{\partial L}{\partial p} = \frac{1}{\det H} \det \begin{bmatrix} -\frac{\partial f}{\partial L} & p\frac{\partial^2 f}{\partial L \partial K} \\ -\frac{\partial f}{\partial K} & p\frac{\partial^2 f}{\partial K^2} \end{bmatrix} = \frac{-p}{\det H} \left(\frac{\partial f}{\partial L} \frac{\partial^2 f}{\partial K^2} - \frac{\partial f}{\partial K} \frac{\partial^2 f}{\partial L \partial K} \right) \quad (22)$$

interestingly enough, this will not necessarily have a sure sign. It will be positive if $\frac{\partial^2 f}{\partial L \partial K} \leq 0$—or the two inputs are substitutes—but we can't be sure if they are compliments.

Example 15 *Profit Maximization with Cobb-Douglass production function.*
If $f(L,K) = L^\alpha K^\beta = Q$ then our first order conditions become:

$$p\alpha \frac{Q}{L} - w = 0$$
$$p\beta \frac{Q}{K} - r = 0$$

and we realize right away that we have to have $\alpha > 0$ and $\beta > 0$ if we want $K^* > 0$ and $L^* > 0$. To find the second order conditions it is best to write $\frac{\partial f}{\partial L} = \alpha L^{\alpha-1} K^\beta$ and $\frac{\partial f}{\partial K} = \beta L^\alpha K^{\beta-1}$. Notice these are the same as the ones I wrote above, it's just easier to take the second derivatives. When we do this we find that:

$$H = \begin{bmatrix} p\frac{\partial^2 f}{\partial L^2} & p\frac{\partial^2 f}{\partial L \partial K} \\ p\frac{\partial^2 f}{\partial L \partial K} & p\frac{\partial^2 f}{\partial K^2} \end{bmatrix} = \begin{bmatrix} p\alpha(\alpha-1)\frac{Q}{L^2} & p\alpha\beta\frac{Q}{LK} \\ p\alpha\beta\frac{Q}{LK} & p\beta(\beta-1)\frac{Q}{K^2} \end{bmatrix}$$

and in order for this to be negative definite we need $p\alpha(\alpha-1)\frac{Q}{L^2} < 0$, or $\alpha < 1$, and

$$\det \begin{bmatrix} p\alpha(\alpha-1)\frac{Q}{L^2} & p\alpha\beta\frac{Q}{LK} \\ p\alpha\beta\frac{Q}{LK} & p\beta(\beta-1)\frac{Q}{K^2} \end{bmatrix} = \frac{p^2 Q^2}{K^2 L^2}\alpha\beta(1-\alpha-\beta) > 0$$

which means that we have to have $(1-\alpha-\beta) > 0$. Now clearly we could figure out the explicit function of L and K in terms of (p,w,r)–but who wants to? It's much more fun to find the partial derivatives using the implicit function theorem.

$$\frac{\partial L}{\partial w} = \frac{1}{\det H}\det \begin{bmatrix} 1 & p\alpha\beta\frac{Q}{LK} \\ 0 & p\beta(\beta-1)\frac{Q}{K^2} \end{bmatrix} = \frac{p\beta(\beta-1)\frac{Q}{K^2}}{\frac{p^2 Q^2}{K^2 L^2}\alpha\beta(1-\alpha-\beta)} = -\frac{L^2}{Qp\alpha}\frac{1-\beta}{1-\alpha-\beta} < 0$$

$$\frac{\partial K}{\partial w} = \frac{1}{\det H}\det \begin{bmatrix} p\alpha(\alpha-1)\frac{Q}{L^2} & 1 \\ p\alpha\beta\frac{Q}{LK} & 0 \end{bmatrix} = \frac{-\frac{1}{KL}Qp\alpha\beta}{\frac{p^2 Q^2}{K^2 L^2}\alpha\beta(1-\alpha-\beta)} = -\frac{LK}{Qp}\frac{1}{1-\alpha-\beta} < 0$$

In this case, as you would expect, when you increase the sales price both of the inputs increase.

$$\frac{\partial L}{\partial p} = \frac{1}{\det H}\det \begin{bmatrix} -\alpha\frac{Q}{L} & p\alpha\beta\frac{Q}{LK} \\ -\beta\frac{Q}{K} & p\beta(\beta-1)\frac{Q}{K^2} \end{bmatrix} = \frac{\frac{1}{K^2 L}Q^2 p\alpha\beta}{\frac{p^2 Q^2}{K^2 L^2}\alpha\beta(1-\alpha-\beta)} = \frac{L}{p}\frac{1}{(1-\alpha-\beta)} > 0$$

$$\frac{\partial K}{\partial w} = \frac{1}{\det H}\det \begin{bmatrix} p\alpha(\alpha-1)\frac{Q}{L^2} & -\alpha\frac{Q}{L} \\ p\alpha\beta\frac{Q}{LK} & -\beta\frac{Q}{K} \end{bmatrix} = \frac{\frac{1}{KL^2}Q^2 p\alpha\beta}{\frac{p^2 Q^2}{K^2 L^2}\alpha\beta(1-\alpha-\beta)} = \frac{K}{p}\frac{1}{(1-\alpha-\beta)} > 0$$

OK, not only has this helped you see how these things might work in practice but you've realized the constraints we usually place on α and β are necessary. In order to have the results we want we must have $\alpha \in (0,1)$, $\beta \in (0,1)$ and $\alpha + \beta < 1$. It's a pain in the neck, but we need them all.

2　Constrained Maximization with n Variables.

Wouldn't be nice if we could stop now? But you notice that every time I did profit analysis I said we should ignore some constraints. Specifically these are $\pi(L^*, K^*) \geq 0$, $L^* \geq 0$, and $K^* \geq 0$. Constrained maximization is by far the more common, and unfortunately life gets a bit more complicated. I will only

talk about the case where we have k constraints of the form $g_j(x) \leq b_j$ and $x_i \geq 0$ ($j = (1, 2, 3, ..., k)$, $i = (1, 2, 3, ..., n)$). In this case we have the function:

$$KT(\lambda, x) = f(x) - \sum_{j=1}^{k} \lambda_j (g_j(x) - b_j) \ . \tag{23}$$

The book shows that $\lambda_j \geq 0$.[3] This means that we are maximizing this over the case $\lambda_j \geq 0$ and $x_i \geq 0$. This gives rise to the complimentary slackness conditions:

$$\frac{\partial KT}{\partial x_i} \leq 0, \ x_i \geq 0, \ x_i \frac{\partial KT}{\partial x_i} = 0 \tag{24}$$

$$\frac{\partial KT}{\partial \lambda_j} \geq 0, \ \lambda_j \geq 0, \ \lambda_j \frac{\partial KT}{\partial \lambda_j} = 0 \tag{25}$$

The first order conditions for the constraints (equations 25) are fairly obvious. $\frac{\partial KT}{\partial \lambda_j} = -(g_j(x) - b_j)$, we must always have $g_j(x) - b_j \leq 0$, so $\frac{\partial KT}{\partial \lambda_j} \geq 0$. Furthermore:

1. If $g_j(x) < b_j$ then the constraint is not binding and $\lambda_j = 0$, so $\frac{\partial KT}{\partial \lambda_j} > 0$, $\lambda_j \frac{\partial KT}{\partial \lambda_j} = 0$.

2. If $g_j(x) = b_j$ then $\lambda_j > 0$ but $\frac{\partial KT}{\partial \lambda_j} = 0$, so again we have $\lambda_j \frac{\partial KT}{\partial \lambda_j} = 0$.

---

[3] To see this you have to refer to Theorem 14.2, consider maximizing: $h_j(t) = g_j(x + tv)$ over v, then the v that maximizes this expression is $\bigtriangledown g_j(x)$. This means that $\bigtriangledown g_j(x)$ is pointing out of the set in which $g_j(x) \leq b_j$. If $\bigtriangledown f$ does not point in the same direction as $\bigtriangledown g_j(x)$ then it is pointing into the area where $g_j(x) < b_j$, and $\lambda_j = 0$. If it is pointing in the same direction then $\lambda_j > 0$.

It is essentially the same with regards to the x_i's consider a one dimensional problem with $x_i \geq b$. To be concrete look at $\exp\left(-(x-3)^2\right)$ again.
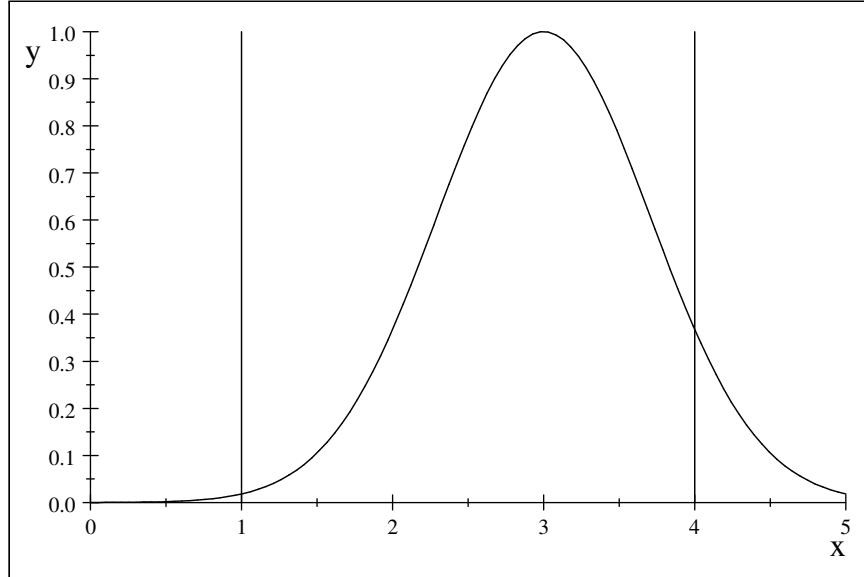
This function has the first derivative: $f'(x) = -e^{-(x-3)^2}(2x-6)$, and while it is not concave it does have a unique maximum when $f'(x) = 0$ or $x = 3$. If $b = 1$ (the left hand line) it's obvious that the maximum is achieved at the unconstrained maximum, $x = 3$. so in this case we have $x^* > b$ and $f'(x^*) = 0$. On the other hand if $b = 4$ (the right hand line) then we have $x^* = 4$ and $f'(4) = -e^{-(4-3)^2}(2(4)-6) = -2e^{-1} = -0.735\,76 < 0$. So in both cases we have $(x^* - b)f'(x^*) = 0$, $f'(x^*) \leq 0$, $x^* - b \geq 0$. In Section 18.6 they explain it based on the Lagrangian when $x_i \geq 0$ is included as an additional constraint.

So, that's not that much more difficult, but what about the second order conditions? Well first of all we drop variables where either $\frac{\partial KT}{\partial x_i} < 0$ or $\frac{\partial KT}{\partial \lambda_j} > 0$. These are variables where a small change in the objective is not going to change the fact that $x_i = 0$ or $\lambda_j = 0$. Now we face an additional constraint, which is that we have to make sure our constraints are satisfied. To do this we need to make sure that (for the binding constraints, let this be $1....k_0$) the Jacobian:

$$Dg = \begin{bmatrix} \frac{\partial g_1}{\partial x_1} & \frac{\partial g_1}{\partial x_2} & \cdots & \frac{\partial g_1}{\partial x_n} \\ \frac{\partial g_2}{\partial x_1} & \frac{\partial g_2}{\partial x_2} & \cdots & \frac{\partial g_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial g_{k_0}}{\partial x_1} & \frac{\partial g_{k_0}}{\partial x_2} & \cdots & \frac{\partial g_{k_0}}{\partial x_n} \end{bmatrix} \tag{26}$$

is such that:

$$Dgv = 0 \tag{27}$$

11

The reason for this goes back to the Taylor series again. Notice that we can also write the Taylor series as:

$$f\left(x+v\right) = f\left(x\right) + v^T \bigtriangledown f\left(x\right) + \varepsilon\left(\|v\|^2\right) \tag{28}$$

and again, for small $\|v\|$ the second term is trivial. Thus when we change v by a small amount we want to make sure that it doesn't change the value of $[g_j]_{j=1...n}$.

But what is the Jacobian? This is the third important technical definition.

Definition 16 *The **Jacobian** of a function g is the transpose of the gradient, or the matrix of first derivatives written as a row matrix:*

$$Dg = \left[\frac{\partial g}{\partial x_i}\right]^{i=1...n} = \left[\begin{array}{cccc} \frac{\partial g}{\partial x_1} & \frac{\partial g}{\partial x_2} & \cdots & \frac{\partial g}{\partial x_n} \end{array}\right]. \tag{29}$$

It is understand that the difference between a row matrix and a column matrix (vector) is that the former is a constraint, while the latter is a vector of variables. For a sequence of functions, $(g_j)_{j=1...k_0}$ the Jacobian is the matrix in equation 26.

We now have our result, which is:

Theorem 17 *(19.6) Assume all constraints are binding and that all variables are strictly greater than zero. If (λ^*, x^*) is critical point of $KT\left(\lambda, x\right)$ then it is a maximum if for all $v \neq 0$ such that $Dgv = 0$, then $v^T D_x^2 KT v < 0$. Where $D_x^2 KT = \left[\frac{\partial^2 KT}{\partial x_i \partial x_j}\right]_{i=1...n}^{j=1...n}$ or that the last $n - 2m$ leading principal minors of:*

$$D^2 KT = \left[\begin{array}{cc} 0 & Dg \\ Dg^T & D_x^2 KT \end{array}\right] \tag{30}$$

have the correct sign. These signs can be found in Theorem 16.4.

If only k_0 constraints are binding then there must be at least $n_0 \geq k_0$ variables with binding first order conditions, and the results of the theorem hold if $n_0 > k_0$. If $n_0 = k_0$ then there is only one possible solution, and it must be the maximum.

But, folks, this is ridiculous, because I promise you that you will only have to check second order conditions for functions of two variables with one binding first order condition. In that case:

Theorem 18 *Assume (μ^*, x^*, y^*) are a critical point of:*

$$L\left(\mu, x, y\right) = f\left(x, y\right) - \mu\left(h\left(x, y\right) - c\right) \tag{31}$$

then it is a maximum if

$$\det \left[\begin{array}{ccc} 0 & -\frac{\partial h}{\partial x} & -\frac{\partial h}{\partial y} \\ -\frac{\partial h}{\partial x} & \frac{\partial^2 f}{\partial x^2} - \mu\frac{\partial^2 h}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} - \mu\frac{\partial^2 h}{\partial x \partial y} \\ -\frac{\partial h}{\partial y} & \frac{\partial^2 f}{\partial x \partial y} - \mu\frac{\partial^2 h}{\partial x \partial y} & \frac{\partial^2 f}{\partial x^2} - \mu\frac{\partial^2 h}{\partial y^2} \end{array}\right] > 0 \tag{32}$$

if the determinant is strictly negative then it is a minimum.

Example 19 *Cost minimization:*

Consider the classic problem of minimizing cost $(wL + rK)$ such that output is above some given level $(f(L, K) \geq Q)$. Then the objective function we want to minimize is:

$$KT(\lambda, L, K) = wL + rK - \lambda(f(L, K) - Q) \tag{33}$$

where we have the usual constraints $\lambda \geq 0$, $L \geq 0$, $K \geq 0$. The first order conditions are:

$$\frac{\partial KT}{\partial \lambda} = -(f(L, K) - Q) \geq 0, \ \lambda \geq 0, \ \lambda\frac{\partial KT}{\partial \lambda} = 0 \tag{34}$$

$$\frac{\partial KT}{\partial L} = w - \lambda\frac{\partial f}{\partial L} \leq 0, \ L \geq 0, \ L\frac{\partial KT}{\partial L} = 0$$

$$\frac{\partial KT}{\partial K} = r - \lambda\frac{\partial f}{\partial K} \leq 0, \ K \geq 0, \ K\frac{\partial KT}{\partial K} = 0$$

assuming that $\lambda > 0$, $L > 0$ and $K > 0$ we have:

$$\det\left(D^2 KT\right) = \det \begin{bmatrix} 0 & -\frac{\partial f}{\partial L} & -\frac{\partial f}{\partial K} \\ -\frac{\partial f}{\partial L} & -\lambda\frac{\partial^2 f}{\partial L^2} & -\lambda\frac{\partial^2 f}{\partial L \partial K} \\ -\frac{\partial f}{\partial K} & -\lambda\frac{\partial^2 f}{\partial L \partial K} & -\lambda\frac{\partial^2 f}{\partial K^2} \end{bmatrix} < 0 \tag{35}$$

which is actually:

$$\det\left(D^2 KT\right) = \lambda\left(\left(\frac{\partial f}{\partial L}\right)^2 \frac{\partial^2 f}{\partial K^2} - 2\frac{\partial^2 f}{\partial L \partial K}\frac{\partial f}{\partial L}\frac{\partial f}{\partial K} + \left(\frac{\partial f}{\partial K}\right)^2 \frac{\partial^2 f}{\partial L^2}\right) < 0 \tag{36}$$

I don't expect you to understand that, but later on I'll want to refer back to it.

Now let's find the implicit functions. We now have three variables, one of which we don't care at all about (λ) but we still have to keep track of it. The total differential of our first order conditions is:

$$\begin{bmatrix} 0 & -\frac{\partial f}{\partial L} & -\frac{\partial f}{\partial K} \\ -\frac{\partial f}{\partial L} & -\lambda\frac{\partial^2 f}{\partial L^2} & -\lambda\frac{\partial^2 f}{\partial L \partial K} \\ -\frac{\partial f}{\partial K} & -\lambda\frac{\partial^2 f}{\partial L \partial K} & -\lambda\frac{\partial^2 f}{\partial K^2} \end{bmatrix}\begin{bmatrix} d\lambda \\ dL \\ dK \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}dQ + \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}dw + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}dr$$

$$\tag{37}$$

And using Crammer's Rule we see that:

$$\frac{\partial L}{\partial w} = \frac{1}{\det\left(D^2 KT\right)}\det\begin{bmatrix} 0 & 0 & -\frac{\partial f}{\partial K} \\ -\frac{\partial f}{\partial L} & -1 & -\lambda\frac{\partial^2 f}{\partial L \partial K} \\ -\frac{\partial f}{\partial K} & 0 & -\lambda\frac{\partial^2 f}{\partial K^2} \end{bmatrix} = \frac{1}{\det\left(D^2 KT\right)}\left(-\frac{\partial f}{\partial K}\right)\det\begin{bmatrix} -\frac{\partial f}{\partial L} & -1 \\ -\frac{\partial f}{\partial K} & 0 \end{bmatrix} \tag{38}$$

$$= -\frac{1}{-\det\left(D^2 KT\right)}\left(\frac{\partial f}{\partial K}\right)^2 < 0$$

$$\frac{\partial K}{\partial w} = \frac{1}{\det\left(D^2 KT\right)}\det\begin{bmatrix} 0 & -\frac{\partial f}{\partial L} & 0 \\ -\frac{\partial f}{\partial L} & -\lambda\frac{\partial^2 f}{\partial L^2} & -1 \\ -\frac{\partial f}{\partial K} & -\lambda\frac{\partial^2 f}{\partial L \partial K} & 0 \end{bmatrix} = \frac{1}{\det\left(D^2 KT\right)}\left(\frac{\partial f}{\partial L}\right)\det\begin{bmatrix} -\frac{\partial f}{\partial L} & -1 \\ -\frac{\partial f}{\partial K} & 0 \end{bmatrix}$$

$$= \frac{1}{-\det\left(D^2 KT\right)}\left(\frac{\partial f}{\partial L}\right)\left(\frac{\partial f}{\partial K}\right) > 0$$

13

and I'm not going to bother finding $\frac{\partial L}{\partial r}$ and $\frac{\partial K}{\partial r}$, it should be fairly obvious.

$$\frac{\partial L}{\partial Q} = \frac{1}{\det\left(D^2 KT\right)} \det \begin{bmatrix} 0 & -1 & -\frac{\partial f}{\partial K} \\ -\frac{\partial f}{\partial L} & 0 & -\lambda\frac{\partial^2 f}{\partial L \partial K} \\ -\frac{\partial f}{\partial K} & 0 & -\lambda\frac{\partial^2 f}{\partial K^2} \end{bmatrix} \tag{39}$$

$$= \frac{1}{\det\left(D^2 KT\right)} \left(\det \begin{bmatrix} -\frac{\partial f}{\partial L} & -\lambda\frac{\partial^2 f}{\partial L \partial K} \\ -\frac{\partial f}{\partial K} & -\lambda\frac{\partial^2 f}{\partial K^2} \end{bmatrix} - \frac{\partial f}{\partial K} \det \begin{bmatrix} -\frac{\partial f}{\partial L} & 0 \\ -\frac{\partial f}{\partial K} & 0 \end{bmatrix} \right)$$

$$= \frac{1}{\det\left(D^2 KT\right)} \lambda \left(\frac{\partial f}{\partial L}\frac{\partial^2 f}{\partial K^2} - \frac{\partial f}{\partial K}\frac{\partial^2 f}{\partial L \partial K} \right)$$

notice the similarity to $\frac{\partial L}{\partial p}$ in equation 22, again it is not sure, if you increase output you may not increase your demand for labor though you will increase either the demand for labor or capital.

Example 20 *Cost Minimization with Cobb-Douglass. Then the first order conditions become:*

$$\frac{\partial KT}{\partial \lambda} = -\left(L^\alpha K^\beta - Q\right) \geq 0,\ \lambda \geq 0,\ \lambda\frac{\partial KT}{\partial \lambda} = 0 \tag{40}$$

$$\frac{\partial KT}{\partial L} = w - \lambda\alpha\frac{Q}{L} \leq 0,\ L \geq 0,\ L\frac{\partial KT}{\partial L} = 0$$

$$\frac{\partial KT}{\partial K} = r - \lambda\beta\frac{Q}{K} \leq 0,\ K \geq 0,\ K\frac{\partial KT}{\partial K} = 0$$

and first of all it's obvious that $\lambda = 0$ can not be a solution, because then $\frac{\partial KT}{\partial L} = w > 0$. So if $\lambda > 0$ and $Q > 0$ can either L or K equaling zero? This is going to get a little technical, because I haven't assumed that $\alpha > 0$. First, if $\alpha < 0$ then $\frac{\partial KT}{\partial L} = w - \lambda\alpha\frac{Q}{L} > 0$, so we must have $\alpha > 0$. Likewise we must have $\beta > 0$. Given this if either $L = 0$ or $K = 0$ then $L^\alpha K^\beta = 0$ and $L^\alpha K^\beta < Q$. Therefore we conclude that $\lambda > 0$, $L > 0$, $K > 0$. Our first order conditions are:

$$-\left(L^\alpha K^\beta - Q\right) = 0 \tag{41}$$

$$w - \lambda\alpha\frac{Q}{L} = 0$$

$$r - \lambda\beta\frac{Q}{K} = 0$$

Now what are our second order conditions? To find them it is better to write $\alpha\frac{Q}{L} = \alpha L^{\alpha-1}K^\beta$ and $\beta\frac{Q}{K} = \beta L^\alpha K^{\beta-1}$. Then our second order conditions are:

$$D^2 KT = \begin{bmatrix} 0 & -\alpha\frac{Q}{L} & -\beta\frac{Q}{K} \\ -\alpha\frac{Q}{L} & -\lambda\alpha\left(\alpha-1\right)\frac{Q}{L^2} & -\lambda\alpha\beta\frac{Q}{LK} \\ -\beta\frac{Q}{K} & -\lambda\alpha\beta\frac{Q}{LK} & -\lambda\beta\left(\beta-1\right)\frac{Q}{K^2} \end{bmatrix} \tag{42}$$

And

$$\det \begin{bmatrix} 0 & -\alpha\frac{Q}{L} & -\beta\frac{Q}{K} \\ -\alpha\frac{Q}{L} & -\lambda\alpha(\alpha-1)\frac{Q}{L^2} & -\lambda\alpha\beta\frac{Q}{LK} \\ -\beta\frac{Q}{K} & -\lambda\alpha\beta\frac{Q}{LK} & -\lambda\beta(\beta-1)\frac{Q}{K^2} \end{bmatrix} = -\frac{Q^3\lambda}{K^2L^2}\alpha\beta(\alpha+\beta) < 0 \tag{43}$$

this will always be true as long as $\alpha > 0$ and $\beta > 0$, so we never have to worry about it. Now we could solve directly for $L(w,r,Q)$ and $K(w,r,Q)$ but what would be the fun in that? I would rather figure out the derivatives using the implicit function theorem:

$$\frac{\partial L}{\partial w} = \frac{1}{\det D^2 KT}\det \begin{bmatrix} 0 & 0 & -\beta\frac{Q}{K} \\ -\alpha\frac{Q}{L} & -1 & -\lambda\alpha\beta\frac{Q}{LK} \\ -\beta\frac{Q}{K} & 0 & -\lambda\beta(\beta-1)\frac{Q}{K^2} \end{bmatrix} = \frac{\frac{1}{K^2}Q^2\beta^2}{-\frac{Q^3\lambda}{K^2L^2}\alpha\beta(\alpha+\beta)} = -\frac{L^2}{Q\lambda}\frac{\beta}{\alpha(\alpha+\beta)} \tag{44}$$

$$\frac{\partial K}{\partial w} = \frac{1}{\det D^2 KT}\det \begin{bmatrix} 0 & -\alpha\frac{Q}{L} & 0 \\ -\alpha\frac{Q}{L} & -\lambda\alpha(\alpha-1)\frac{Q}{L^2} & -1 \\ -\beta\frac{Q}{K} & -\lambda\alpha\beta\frac{Q}{LK} & 0 \end{bmatrix} = \frac{-\frac{1}{KL}Q^2\alpha\beta}{-\frac{Q^3\lambda}{K^2L^2}\alpha\beta(\alpha+\beta)} = \frac{LK}{Q\lambda}\frac{1}{(\alpha+\beta)} \ .$$

Wasn't that fun? And in this case both inputs increase when the output does:

$$\frac{\partial L}{\partial Q} = \frac{1}{\det D^2 KT}\det \begin{bmatrix} 0 & -1 & -\beta\frac{Q}{K} \\ -\alpha\frac{Q}{L} & 0 & -\lambda\alpha\beta\frac{Q}{LK} \\ -\beta\frac{Q}{K} & 0 & -\lambda\beta(\beta-1)\frac{Q}{K^2} \end{bmatrix} = \frac{-\frac{1}{K^2L}Q^2\alpha\beta\lambda}{-\frac{Q^3\lambda}{K^2L^2}\alpha\beta(\alpha+\beta)} = \frac{L}{Q}\frac{1}{(\alpha+\beta)} \tag{45}$$

$$\frac{\partial K}{\partial Q} = \frac{1}{\det D^2 KT}\det \begin{bmatrix} 0 & -\alpha\frac{Q}{L} & -1 \\ -\alpha\frac{Q}{L} & -\lambda\alpha(\alpha-1)\frac{Q}{L^2} & 0 \\ -\beta\frac{Q}{K} & -\lambda\alpha\beta\frac{Q}{LK} & 0 \end{bmatrix} = \frac{-\frac{1}{KL^2}Q^2\alpha\beta\lambda}{-\frac{Q^3\lambda}{K^2L^2}\alpha\beta(\alpha+\beta)} = \frac{K}{Q}\frac{1}{(\alpha+\beta)} \ .$$

Notice that, in contrast to profit maximization, with cost minimization our only constraints are that $\alpha > 0$ and $\beta > 0$.

Example 21 *Utility Maximization, both in general and for the Cobb-Douglass.*

Why do I hate consumer theory? I can explain it to you by doing classic utility maximization problem:

$$KT(\lambda, x, y) = u(x, y) - \lambda(px + qy - I) \tag{46}$$

assume an interior solution and a binding constraint.

$$\frac{\partial KT}{\partial \lambda} = -(px + qy - I) \geq 0, \ \lambda \geq 0, \ \lambda\frac{\partial KT}{\partial \lambda} = 0 \tag{47}$$

$$\frac{\partial KT}{\partial x} = \frac{\partial u}{\partial x} - \lambda p \leq 0, \ x \geq 0, \ x\frac{\partial KT}{\partial x} = 0$$

$$\frac{\partial KT}{\partial y} = \frac{\partial u}{\partial y} - \lambda q \leq 0, \ y \geq 0, \ y\frac{\partial KT}{\partial y} = 0$$

And my Hessian is:

$$D^2 KT = \begin{bmatrix} 0 & -p & -q \\ -p & \frac{\partial^2 u}{\partial x^2} & \frac{\partial^2 u}{\partial x\partial y} \\ -q & \frac{\partial^2 u}{\partial x\partial y} & \frac{\partial^2 u}{\partial y^2} \end{bmatrix} \tag{48}$$

15

Now this is pretty mysterious, since it includes prices. It is better to analyze it by recognizing that from the first order conditions $p = \frac{1}{\lambda}\frac{\partial u}{\partial x}$ and $q = \frac{1}{\lambda}\frac{\partial u}{\partial y}$. Then

$$D^2 KT = \begin{bmatrix} 0 & -\frac{1}{\lambda}\frac{\partial u}{\partial x} & -\frac{1}{\lambda}\frac{\partial u}{\partial y} \\ -\frac{1}{\lambda}\frac{\partial u}{\partial x} & \frac{\partial^2 u}{\partial x^2} & \frac{\partial^2 u}{\partial x \partial y} \\ -\frac{1}{\lambda}\frac{\partial u}{\partial y} & \frac{\partial^2 u}{\partial x \partial y} & \frac{\partial^2 u}{\partial y^2} \end{bmatrix} \qquad (49)$$

This should look pretty familiar, except that $\begin{bmatrix} \frac{\partial^2 u}{\partial x^2} & \frac{\partial^2 u}{\partial x \partial y} \\ \frac{\partial^2 u}{\partial x \partial y} & \frac{\partial^2 u}{\partial y^2} \end{bmatrix}$ *is not multiplied by* -1 *it's pretty much the same matrix as for cost minimization. Indeed the condition for a maximum is the same:*

$$\det D^2 KT = -\frac{1}{\lambda^2}\left[\left(\frac{\partial u}{\partial y}\right)^2\frac{\partial^2 u}{\partial x^2} - 2\frac{\partial u}{\partial x}\frac{\partial u}{\partial y}\frac{\partial^2 u}{\partial x \partial y} + \left(\frac{\partial u}{\partial x}\right)^2\frac{\partial^2 u}{\partial y^2}\right] > 0 \quad (50)$$

$$\left(\frac{\partial u}{\partial y}\right)^2\frac{\partial^2 u}{\partial x^2} - 2\frac{\partial u}{\partial x}\frac{\partial u}{\partial y}\frac{\partial^2 u}{\partial x \partial y} + \left(\frac{\partial u}{\partial x}\right)^2\frac{\partial^2 u}{\partial y^2} < 0$$

now let's do the same comparative statics exercise. Let's look at the impact of p on the demand for x:

$$D^2 KT \begin{bmatrix} d\mu \\ dx \\ dy \end{bmatrix} = \begin{bmatrix} x \\ \lambda \\ 0 \end{bmatrix} dp \qquad (51)$$

$$\frac{\partial x}{\partial p} = \frac{1}{\det D^2 KT}\det\begin{bmatrix} 0 & x & -\frac{1}{\lambda}\frac{\partial u}{\partial y} \\ -\frac{1}{\lambda}\frac{\partial u}{\partial x} & \lambda & \frac{\partial^2 u}{\partial x \partial y} \\ -\frac{1}{\lambda}\frac{\partial u}{\partial y} & 0 & \frac{\partial^2 u}{\partial y^2} \end{bmatrix} \qquad (52)$$

$$\det\begin{bmatrix} 0 & x & -\frac{1}{\lambda}\frac{\partial u}{\partial y} \\ -\frac{1}{\lambda}\frac{\partial u}{\partial x} & \lambda & \frac{\partial^2 u}{\partial x \partial y} \\ -\frac{1}{\lambda}\frac{\partial u}{\partial y} & 0 & \frac{\partial^2 u}{\partial y^2} \end{bmatrix} = 0\det\begin{bmatrix} \lambda & \frac{\partial^2 u}{\partial x \partial y} \\ 0 & \frac{\partial^2 u}{\partial y^2} \end{bmatrix} - x\det\begin{bmatrix} -\frac{1}{\lambda}\frac{\partial u}{\partial x} & \frac{\partial^2 u}{\partial x \partial y} \\ -\frac{1}{\lambda}\frac{\partial u}{\partial y} & \frac{\partial^2 u}{\partial y^2} \end{bmatrix} - \frac{1}{\lambda}\frac{\partial u}{\partial y}\det\begin{bmatrix} -\frac{1}{\lambda}\frac{\partial u}{\partial x} & \lambda \\ -\frac{1}{\lambda}\frac{\partial u}{\partial y} & 0 \end{bmatrix} \quad (53)$$

$$= -x\left(-\frac{1}{\lambda}\frac{\partial u}{\partial x}\frac{\partial^2 u}{\partial y^2} + \frac{1}{\lambda}\frac{\partial u}{\partial y}\frac{\partial^2 u}{\partial x \partial y}\right) - \frac{1}{\lambda}\frac{\partial u}{\partial y}\left(\frac{\partial u}{\partial y}\right)$$

$$= \frac{1}{\lambda}\left[-\left(\frac{\partial u}{\partial y}\right)^2 + \left(\frac{\partial u}{\partial x}\frac{\partial^2 u}{\partial y^2} - \frac{\partial u}{\partial y}\frac{\partial^2 u}{\partial x \partial y}\right)x\right]$$

*and you might be surprised to find out that we can't determine this sign. The reason is because of the **income effect**, $\left(\frac{\partial u}{\partial x}\frac{\partial^2 u}{\partial y^2} - \frac{\partial u}{\partial y}\frac{\partial^2 u}{\partial x \partial y}\right)x$ but that's a topic for ECON 203. And why I hate utility maximization.*

Now let's look at things when we have the beautiful Cobb-Douglass function,

$u = x^\alpha y^\beta$. *Our first order conditions are:*

$$
\begin{aligned}
-(px + qy - I) &= 0 & (54) \\
\alpha \frac{u}{x} - \lambda p &= 0 \\
\beta \frac{u}{y} - \lambda q &= 0
\end{aligned}
$$

and while it's a little harder to prove we still need $\alpha > 0$ and $\beta > 0$, our Hessian is:

$$
D^2 KT = \begin{bmatrix} 0 & -\frac{1}{\lambda}\alpha\frac{u}{x} & -\frac{1}{\lambda}\beta\frac{u}{y} \\ -\frac{1}{\lambda}\alpha\frac{u}{x} & \alpha(\alpha-1)\frac{u}{x^2} & \alpha\beta\frac{u}{xy} \\ -\frac{1}{\lambda}\beta\frac{u}{y} & \alpha\beta\frac{u}{xy} & \beta(\beta-1)\frac{u}{y^2} \end{bmatrix} \tag{55}
$$

$$
\det D^2 KT = \det \begin{bmatrix} 0 & -\frac{1}{\lambda}\alpha\frac{u}{x} & -\frac{1}{\lambda}\beta\frac{u}{y} \\ -\frac{1}{\lambda}\alpha\frac{u}{x} & \alpha(\alpha-1)\frac{u}{x^2} & \alpha\beta\frac{u}{xy} \\ -\frac{1}{\lambda}\beta\frac{u}{y} & \alpha\beta\frac{u}{xy} & \beta(\beta-1)\frac{u}{y^2} \end{bmatrix} = \frac{u^3}{x^2 y^2}\alpha\frac{\beta}{\lambda^2}(\alpha + \beta) > 0
$$

and just like for cost minimization, all we need is that $\alpha > 0$ and $\beta > 0$. What about the partial derivatives?

$$
\frac{\partial x}{\partial p} = \frac{1}{\det D^2 KT} \det \begin{bmatrix} 0 & x & -\frac{1}{\lambda}\beta\frac{u}{y} \\ -\frac{1}{\lambda}\alpha\frac{u}{x} & \lambda & \alpha\beta\frac{u}{xy} \\ -\frac{1}{\lambda}\beta\frac{u}{y} & 0 & \beta(\beta-1)\frac{u}{y^2} \end{bmatrix} = \frac{-\frac{u^2}{y^2}\frac{\beta}{\lambda}(\alpha+\beta)}{\frac{u^3}{x^2 y^2}\alpha\frac{\beta}{\lambda^2}(\alpha+\beta)} = -\frac{x^2\lambda}{u}\frac{1}{\alpha} \tag{56}
$$

$$
\frac{\partial y}{\partial p} = \frac{1}{\det D^2 KT} \det \begin{bmatrix} 0 & -\frac{1}{\lambda}\alpha\frac{u}{x} & x \\ -\frac{1}{\lambda}\alpha\frac{u}{x} & \alpha(\alpha-1)\frac{u}{x^2} & \lambda \\ -\frac{1}{\lambda}\beta\frac{u}{y} & \alpha\beta\frac{u}{xy} & 0 \end{bmatrix} = \frac{0}{\frac{u^3}{x^2 y^2}\alpha\frac{\beta}{\lambda^2}(\alpha+\beta)} = 0
$$

Wow, isn't that weird? Of course if I had solved for the explicit demands: $x = \frac{\alpha}{\alpha+\beta}\frac{I}{p}$, $y = \frac{\beta}{\alpha+\beta}\frac{I}{q}$, it would have been a little easier to derive this, but my method is more fun. (And will work when you can't find those functions–an important plus.) Just for the fun of it:

$$
\frac{\partial x}{\partial I} = \frac{1}{\det D^2 KT} \det \begin{bmatrix} 0 & -1 & -\frac{1}{\lambda}\beta\frac{u}{y} \\ -\frac{1}{\lambda}\alpha\frac{u}{x} & 0 & \alpha\beta\frac{u}{xy} \\ -\frac{1}{\lambda}\beta\frac{u}{y} & 0 & \beta(\beta-1)\frac{u}{y^2} \end{bmatrix} = \frac{\frac{u^2}{xy^2}\alpha\frac{\beta}{\lambda}}{\frac{u^3}{x^2 y^2}\alpha\frac{\beta}{\lambda^2}(\alpha+\beta)} = \frac{x\lambda}{u}\frac{1}{\alpha+\beta} \tag{57}
$$

$$
\frac{\partial y}{\partial I} = \frac{1}{\det D^2 KT} \det \begin{bmatrix} 0 & -\frac{1}{\lambda}\alpha\frac{u}{x} & -1 \\ -\frac{1}{\lambda}\alpha\frac{u}{x} & \alpha(\alpha-1)\frac{u}{x^2} & 0 \\ -\frac{1}{\lambda}\beta\frac{u}{y} & \alpha\beta\frac{u}{xy} & 0 \end{bmatrix} = \frac{\frac{u^2}{x^2 y}\alpha\frac{\beta}{\lambda}}{\frac{u^3}{x^2 y^2}\alpha\frac{\beta}{\lambda^2}(\alpha+\beta)} = \frac{\lambda y}{u}\frac{1}{\alpha+\beta} > 0
$$

Example 22 *How to deal with multiple inequality constraints: Barter with a Transaction Cost.*

It is very common to have multiple inequality constraints. I've given you a question like this on a quiz and very well might on the final. And the methodology is not that difficult, though I haven't covered it yet in class.

17

Consider a model of trading with a transaction cost. It should be obvious that there is a cost for bring goods to and from the market, let us call this τ and assume that it is per-unit.

Now a barter model has you start with an initial endowment (x_0, y_0) and then your income is just the market prices times this endowment: $I = \tilde{p}x_0 + \tilde{q}y_0$.

But now let's combine the two, what happens? Now whether you are buying or selling the good affects the price you face.

Say that you are selling $y_s > 0$ units of good y to buy some good x.

You are going to get q in the market for each unit of y_s, but you are going to have to pay the transportation cost of getting it to the market, so your income at the market will be $(q - \tau)\, y_s$.

You will use this to buy extra units of x, let's call this x_b, and for each of these units you will first pay a price p then pay τ to get it home.

Thus for $x_b \geq 0$ and $y_s \geq 0$ your budget constraint is:

$$(p + \tau)\, x_b \leq (q - \tau)\, y_s \tag{58}$$

Now

$$y_s = y_0 - y^* \tag{59}$$

where y^* is the desired final consumption of y.

$$x_b = x^* - x_0 \tag{60}$$

so we can rewrite this budget constraint as:

$$
\begin{aligned}
(p + \tau)\,(x^* - x_0) &\leq& (q - \tau)\,(y_0 - y^*) \\
(p + \tau)\,x^* + (q - \tau)\,y^* &\leq& (p + \tau)\,x_0 + (q - \tau)\,y_0
\end{aligned}
\tag{61}
$$

Which is a much more natural form, and allows us to write

$$I_x = (p + \tau)\,x_0 + (q - \tau)\,y_0 \tag{62}$$

where the x is to make sure that we remember that we are buying x and selling y.

Reversing the logic, consider a case where we sell x and buy y then:

$$
\begin{aligned}
(q + \tau)\,y_b &\leq& (p - \tau)\,x_s \\
(q + \tau)\,(y^* - y_0) &\leq& (p - \tau)\,(x_0 - x^*) \\
(p - \tau)\,x^* + (q + \tau)\,y^* &\leq& (p - \tau)\,x_0 + (q + \tau)\,y_0
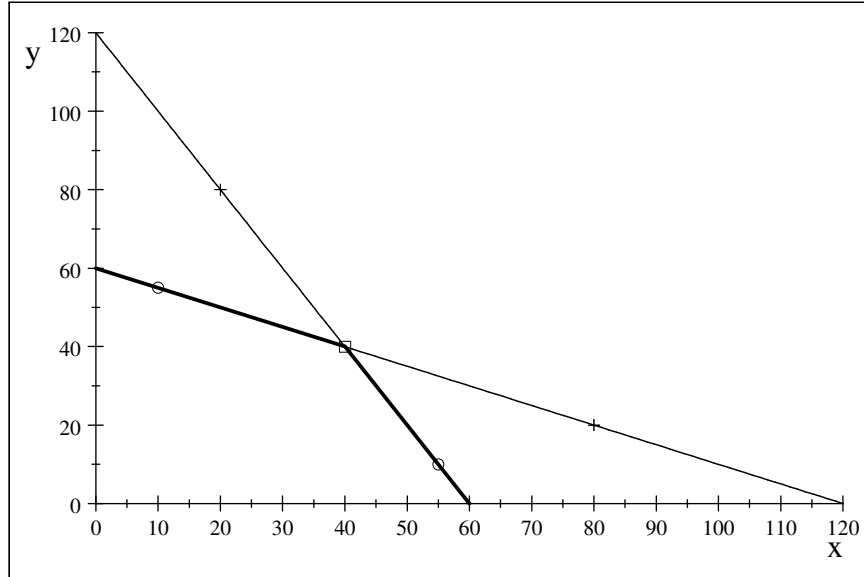\end{aligned}
\tag{63}
$$

so $I_y = (p - \tau)\,x_0 + (q + \tau)\,y_0$. Thus our Kuhn-Tucker function is:

$$KT(\lambda_x, \lambda_y, x, y) = U(x, y) - \lambda_x\,((p + \tau)\,x + (q - \tau)\,y - I_x) - \lambda_y\,((p - \tau)\,x + (q + \tau)\,y - I_y) \tag{64}$$

Now let's consider a specific example, and think about what the solution could be. Say that $(x_0, y_0) = (40, 40)$, and $(p, q, \tau) = (3, 3, 1)$ then the two constraints are:

$$
\begin{aligned}
4x + 2y &\leq& 240 \\
2x + 4y &\leq& 240
\end{aligned}
\tag{65}
$$

18

Let's graph these constraints and consider what we might find.

The light lines are the two constraints, the boundary of the true feasible points is the heavy dark line. Now, assuming utility is strictly monotonic/more is better, obviously we are going to be on one of the budget constraints. But which one? Well one way we could approach it is to solve the problem as if we had only one of the budget constraints. If we found the solution was like one of the circles, then that would be fine. On the other hand if we found it was like one of the crosses it would not work out. For example, at the lower cross we are buying x and selling y, but the prices we are considering assume we will sell x and buy y. So that doesn't work. So, how would we know that we have actually found the optimal point? It would be something like the following:

1. *Assume your budget constraint is $(p - \tau) x^* + (q + \tau) y^* \leq (p - \tau) x_0 + (q + \tau) y_0$, let (x^*, y^*) be the optimum. If $(p + \tau) x^* + (q - \tau) y^* \leq (p + \tau) x_0 + (q - \tau) y_0$ then you were right, and this is the optimum. If $(p + \tau) x^* + (q - \tau) y^* > (p + \tau) x_0 + (q - \tau) y_0$ continue.*

2. *Assume your budget constraint is $(p + \tau) x^* + (q - \tau) y^* \leq (p + \tau) x_0 + (q - \tau) y_0$, let (x^*, y^*) be the optimum. If $(p - \tau) x^* + (q + \tau) y^* \leq (p - \tau) x_0 + (q + \tau) y_0$ then you were right, and this is the optimum. If $(p - \tau) x^* + (q + \tau) y^* > (p - \tau) x_0 + (q + \tau) y_0$ continue.*

3. *Your optimum is (x_0, y_0)*

And this is general advice for how to proceed on any such problem. First consider case A, check to see if you are satisfy the other (or all of the other) constraint(s). Then consider case B, and so on. Only consider multiple constraints once every check of one constraint at a time has failed.

19

Let me work through this specific example when $u(x, y) = x^3 y$. I will want to consider different levels for τ, so let me just treat it as a variable. First notice that $(3 - \tau) 40 + (3 + \tau) 40 = 240$ for any level of τ, so my initial income will always be 240. The Kuhn-Tucker function is:

$$KT(\lambda_1, \lambda_2, x, y) = x^3 y - \lambda_x ((3 + \tau) x + (3 - \tau) y - 240) - \lambda_y ((3 - \tau) x + (3 + \tau) y - 240) \tag{66}$$

with associated first order conditions:

$$\frac{\partial KT}{\partial x} = 3x^2 y - \lambda_x (3 + \tau) - \lambda_y (3 - \tau) \leq 0, \ x \geq 0, \ x\frac{\partial KT}{\partial x} = 0 \tag{67}$$

$$\frac{\partial KT}{\partial y} = x^3 - \lambda_x (3 - \tau) - \lambda_y (3 + \tau) \leq 0, \ y \geq 0, \ y\frac{\partial KT}{\partial y} = 0$$

$$\frac{\partial KT}{\partial \lambda_x} = -((3 + \tau) x + (3 - \tau) y - 240) \geq 0, \ \lambda_x \geq 0, \ \lambda_x\frac{\partial KT}{\partial \lambda_x} = 0$$

$$\frac{\partial KT}{\partial \lambda_y} = -((3 - \tau) x + (3 + \tau) y - 240) \geq 0, \ \lambda_y \geq 0, \ \lambda_y\frac{\partial KT}{\partial \lambda_y} = 0$$

First of all, if $\lambda_1 = \lambda_2 = 0$ then this would imply $x = y = 0$ and $f(x, y) = 0$, but this obviously can't be true because (for example) if I contain my endowment my utility is $u = (40)^3 40 = 2,560,000$, so I know that either $\lambda_1 > 0$ or $\lambda_2 > 0$ or both. But considering the numbers involved I don't want to "compare maxima," so I will use the insight above. So if $\lambda_x > 0 = \lambda_y$ then:

$$\frac{\partial KT}{\partial x} = 3x^2 y - \lambda_x (3 + \tau) = 0 \tag{68}$$

$$\lambda_x = \frac{3x^2 y}{(3 + \tau)}$$

$$\lambda_x = \frac{x^3}{(3 - \tau)}$$

$$(3 + \tau) x = 3(3 - \tau) y$$

$$[3(3 - \tau) y] + (3 - \tau) y = 240$$

$$y = \frac{1}{4}\frac{240}{(3 - \tau)} = \frac{60}{3 - \tau}$$

$$x = \frac{3}{4}\frac{240}{(3 + \tau)} = \frac{180}{\tau + 3}$$

Now does it satisfy the other constraint? Do we have:

$$(3 - \tau) x + (3 + \tau) y \leq 240 \tag{69}$$

$$(3 - \tau) \left(\frac{180}{\tau + 3}\right) + (3 + \tau) \left(\frac{60}{3 - \tau}\right) \leq 240$$

If $\tau = 1$

$$(3 - \tau) \left(\frac{180}{\tau + 3}\right) + (3 + \tau) \left(\frac{60}{3 - \tau}\right) = (3 - (1)) \left(\frac{180}{(1) + 3}\right) + (3 + (1)) \left(\frac{60}{3 - (1)}\right) = 210 \tag{70}$$

so yes, it works. If $\tau = 2$ then:

$$(3 - \tau)\left(\frac{180}{\tau + 3}\right) + (3 + \tau)\left(\frac{60}{3 - \tau}\right) = (3 - (2))\left(\frac{180}{(2) + 3}\right) + (3 + (2))\left(\frac{60}{3 - (2)}\right) = 336 > 240 \tag{71}$$

So I only have to do the other case when $\tau = 2$. I worked it out for general τ, so I'm not going to bother setting $\tau = 2$ until I check it.

So in this case $\lambda_y > 0 = \lambda_x$ and our first order conditions become:

$$\lambda_y = \frac{3x^2 y}{(3 - \tau)} \tag{72}$$

$$\lambda_y = \frac{x^3}{(3 + \tau)}$$

$$3(3 + \tau)y = (3 - \tau)x$$

$$y = \frac{1}{4}\frac{240}{3 + \tau} = \frac{60}{\tau + 3} \tag{73}$$

$$x = \frac{3}{4}\frac{240}{3 - \tau} = \frac{180}{3 - \tau}$$

Does it satisfy the other budget constraint?

$$(3 + \tau)\left(\frac{180}{3 - \tau}\right) + (3 - \tau)\left(\frac{60}{\tau + 3}\right) \le 240 \tag{74}$$

If $\tau = 2$ then:

$$(3 + \tau)\left(\frac{180}{3 - \tau}\right) + (3 - \tau)\left(\frac{60}{\tau + 3}\right) = (3 + (2))\left(\frac{180}{3 - (2)}\right) + (3 - (2))\left(\frac{60}{(2) + 3}\right) = 912 \tag{75}$$

Whoa. What about when $\tau = 1$?

$$(3 + \tau)\left(\frac{180}{3 - \tau}\right) + (3 - \tau)\left(\frac{60}{\tau + 3}\right) = (3 + (1))\left(\frac{180}{3 - (1)}\right) + (3 - (1))\left(\frac{60}{(1) + 3}\right) = 390 \tag{76}$$

OK, so it fails both time. When $\tau = 1$ I have now confirmed and doubly confirmed that the optimum is $(x^*, y^*) = (45, 30)$. When $\tau = 2$ neither the optimization assuming I would sell x or sell y worked, so $(x^*, y^*) = (40, 40)$.

Notice my choosing the case where you buy x was not really by random chance. I looked at the utility function and said "man, this guy really likes x," so I went with that option first. Let me point out that since both of my constraints are linear I could have just considered the problem:

$$\max x^3 y - \lambda(px + qy - I) \tag{77}$$

$$3x^2y - \lambda p \;=\; 0 \tag{78}$$
$$x^3 - \lambda q \;=\; 0$$
$$\frac{1}{q}x^3 \;=\; \frac{3}{p}x^2y$$
$$y \;=\; \frac{1}{3}\frac{p}{q}x$$

$$px + q\left(\frac{1}{3}\frac{p}{q}x\right) \;=\; I \tag{79}$$
$$x \;=\; \frac{3}{4}\frac{I}{p}$$
$$y \;=\; \frac{1}{4}\frac{I}{q}$$

then plugged in our values of (p, q, I) with the different constraints. It would have taken less calculation, though I would have had to work with abstract variables. This should work any time your constraints are from the same family. (Both quadratic, both linear, things like that.)