Economics 101A
Section Notes
GSI: David Albouy

Notes on Calculus and Optimization

1 Basic Calculus

1.1 Definition of a Derivative

Let $f(x)$ be some function of x, then the **derivative** of f, if it exists, is given by the following limit

$$\frac{df(x)}{dx} = \lim_{h\to 0}\frac{f(x+h)-f(x)}{h} \qquad \text{(Definition of Derivative)}$$

although often this definition is hard to apply directly. It is common to write $f'(x)$, or $\frac{df}{dx}$ to be shorter, or if $y = f(x)$ then $\frac{dy}{dx}$ for the derivative of y with respect to x.

1.2 Calculus Rules

Here are some handy formulas which can be derived using the definitions, where $f(x)$ and $g(x)$ are functions of x and k is a constant which does not depend on x.

$$\frac{d}{dx}[kf(x)] = k\frac{df(x)}{dx} \qquad \text{(Constant Rule)}$$

$$\frac{d}{dx}[f(x)+g(x)] = \frac{df(x)}{dx}+\frac{df(x)}{dy} \qquad \text{(Sum Rule)}$$

$$\frac{d}{dx}[f(x)g(x)] = \frac{df(x)}{dx}g(x)+f(x)\frac{dg(x)}{dx} \qquad \text{(Product Rule)}$$

$$\frac{d}{dx}\left[\frac{f(x)}{g(x)}\right] = \frac{\frac{df(x)}{dx}g(x)-\frac{dg(x)}{dx}f(x)}{[g(x)]^2} \qquad \text{(Quotient Rule)}$$

$$\frac{d}{dx}f[g(x)] = \frac{f[g(x)]}{dg}\frac{dg(x)}{dx} \qquad \text{(Chain Rule)}$$

For specific forms of f the following rules are useful in economics

$$\frac{d}{dx}x^k = kx^{k-1} \qquad \text{(Power Rule)}$$

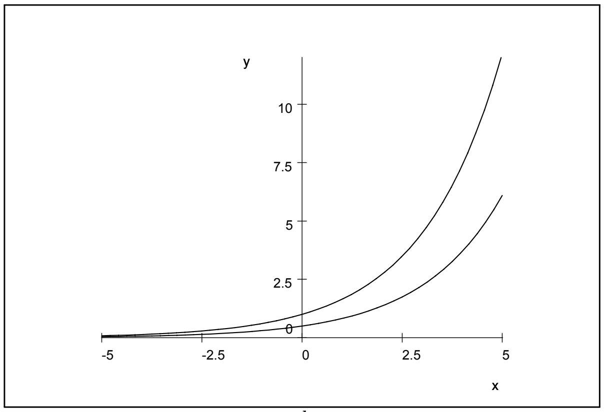
$$\frac{d}{dx}e^x = e^x \qquad \text{(Exponent Rule)}$$

$$\frac{d}{dx}\ln x = \frac{1}{x} \qquad \text{(Logarithm Rule)}$$

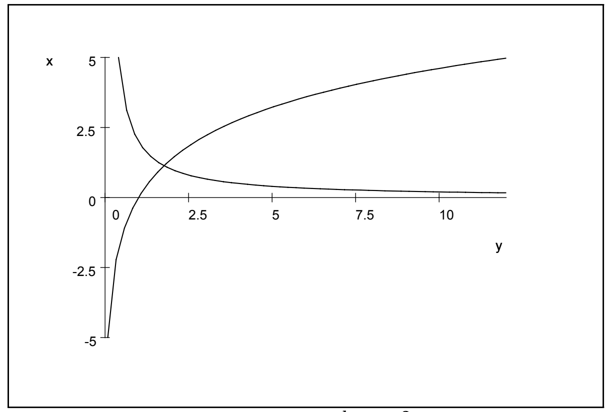
Finally assuming that we can invert $y = f(x)$ by solving for x in terms of y so that $x = f^{-1}(y)$ then the following rule applies

$$\frac{df^{-1}(y)}{dy} = \frac{1}{\frac{df(f^{-1}(y))}{dx}} \qquad \text{(Inverse Rule)}$$

Example 1 *Let $y = f(x) = e^{x/2}$, then using the exponent rule and the chain rule, where $g(x) = x/2$, we get $\frac{df(x)}{dx} = \frac{d}{dx}\left(e^{x/2}\right) = \frac{d}{d(x/2)}\left(e^{x/2}\right)\cdot\frac{d}{dx}\left(\frac{x}{2}\right) = e^{x/2}\left(\frac{1}{2}\right) = \frac{1}{2}e^{x/2}$. The inverse function of f can be gotten by solving for x in terms of y. $y = e^{x/2} \Rightarrow \ln y = \ln\left(e^{x/2}\right) = \frac{x}{2}\ln(e) = \frac{x}{2} \Rightarrow x = 2\ln y$ using the facts from precalculus that $\ln e^u = u\ln e$ and that $\ln e = 1$. The derivative of $f^{-1}(y)$ can be solved directly as $\frac{d}{dy}(2\ln y) = 2\frac{d}{dy}(\ln y) = 2\left(\frac{1}{y}\right) = \frac{2}{y}$ or indirectly through the Inverse Rule as $\frac{df^{-1}(y)}{dy} = \frac{1}{\frac{1}{2}e^{x/2}} = \frac{1}{\frac{1}{2}e^{2\ln y/2}} = \frac{2}{e^{\ln y}} = \frac{2}{y}$ using the fact that $e^{\ln y} = y$.*



$y = \exp(x/2) \quad \frac{dy}{dx} = \frac{1}{2}\exp(x/2)$

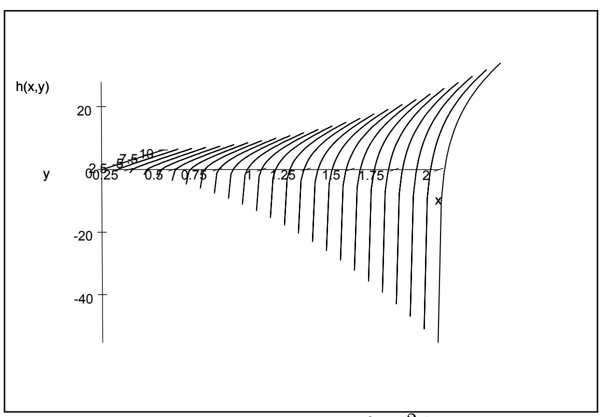


$x = 2\ln y \quad \frac{dx}{dy} = \frac{2}{y}$

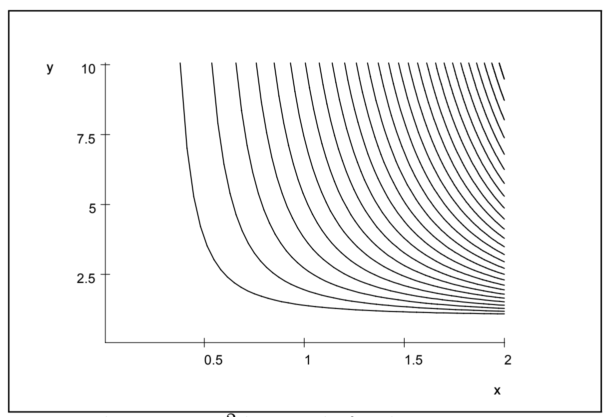
1.3 Partial Differentiation

Taking derivatives of functions in several variables is very straightforward (at least with the functions used in economics). Say we have a function in two variables, x and y, which we can write as $h(x, y)$. Partial differentiation, which uses a "∂" ("del") instead of a "d", can be done in either in x or in y, by writing either "$\frac{\partial}{\partial x}$" or "$\frac{\partial}{\partial y}$" in front of $h(x, y)$, respectively, and then taking the derivative with respect to the variable chosen as we would in the single variable case, treating the other variable *as a constant*.

Example 2 *Let* $h(x, y) = 3x^2 \ln y$*, shown below in the graphs for ranges* $0 \leq x \leq 2$ *and* $0 \leq y \leq 10$



3-D representation of $3x^2 \ln y$



Level curves $3x^2 \ln y = k$ *for* $k = 1, 2, ..., 27$

Note the second plot of the level curves are like what you would find in a topographic contour map. Applying the partial differentiation rules we get the partial derivatives

$$\frac{\partial}{\partial x}\left(3x^2 \ln y\right) = 3\ln y\left(\frac{\partial}{\partial x}x^2\right) = 3\ln y\,(2x) = 6x\ln y$$

$$\frac{\partial}{\partial y}\left(3x^2 \ln y\right) = 3x^2\left(\frac{\partial}{\partial y}\ln y\right) = 3x^2\left(\frac{1}{y}\right) = \frac{3x^2}{y}$$

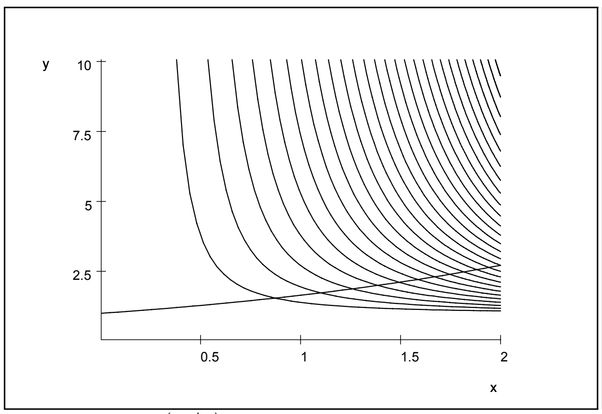
1.4 Total Differentiation

Say that we know already that y is a function of x, say $y = f(x)$ and we want to see how a function in both variables, say $h(x, y)$ responds when we let *both* x change *and* y *change according to how it should according to* f. The we could rewrite our expression as a function in a single variable $H(x) = h(x, f(x))$ and take the **total derivative** with respect to x to get

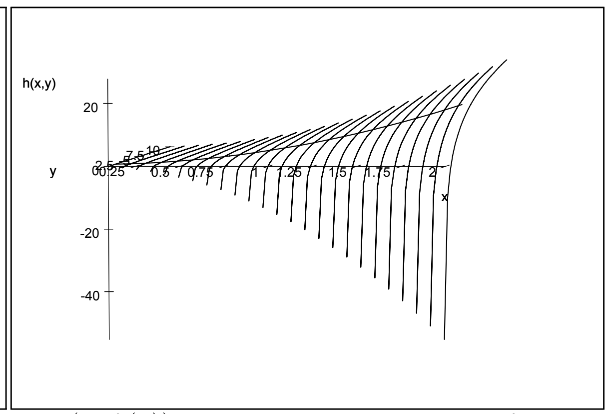
$$\frac{dH(x)}{dx} = \frac{\partial h(x, f(x))}{\partial x} + \frac{\partial h(x, f(x))}{\partial y} \frac{df(x)}{dx} \qquad \text{(Total Derivative)}$$

where the second term makes use of the chain rule.

Example 3 *Let* $y = f(x) = e^{x/2}$ *and again* $h(x, y) = 3x^2 \ln y$. *Then* $\frac{df(x)}{dx} = \frac{1}{2}e^{x/2}$. *And so by the above formulas we have* $\frac{dH(x)}{dx} = 6x \ln y + \frac{3x^2}{y}\frac{1}{2}e^{x/2}$. *Notice that this expression is in both* x *and* y. *It can be simplified to be just a function of* x *by substituting in* $y = e^{x/2}$ *so* $\frac{dH(x)}{dx} = 6x \ln e^{x/2} + \frac{3x^2}{e^{x/2}}\frac{1}{2}e^{x/2} = 3x^2 + \frac{3x^2}{2} = \frac{9}{2}x^2$. *Of course had we substituted in directly for* y *from the beginning we would have gotten* $H(x) = 3x^2 \ln e^{x/2} = \frac{3}{2}x^3$ *and differentiating directly* $\frac{dH(x)}{dx} = \frac{9}{2}x^2$, *but this is not as general a treatment.*



$\exp(x/2)$ *crossing level curves*



$h(x, f(x))$ *is given by curve on the surface*

1.5 Implicit Differentiation

Imagine that it is known that y depends on x, but that we do not have an "explicit" formula for y in terms of x, but that we do know that together x and y must satisfy some equation

$$h(x, y) = k \qquad \text{(Equation in 2 variables)}$$

where $h(x, y)$ is a function and k is just a constant (any equation in x and y can be written like this). Sometimes it is difficult to solve "explicitly" for y in terms of x, i.e. to find the function f such that $y = f(x)$. Nevertheless by the **implicit function theorem** we can still find the derivative of f by differentiating "implicitly" with respect to x, treating y as a function of x, so we can write

$$H(x) = h(x, f(x)) = k$$

Using what we learned about total differentiation we can differentiate totally with respect to x to get

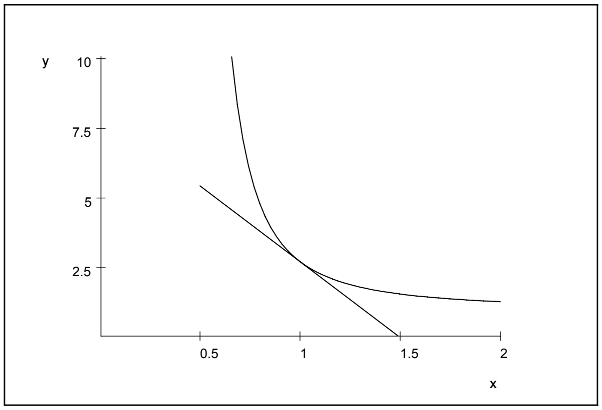
$$\frac{dH(x)}{dx} = \frac{\partial h(x, f(x))}{\partial x} + \frac{\partial h(x, f(x))}{\partial y} \frac{df(x)}{dx} = 0$$

and then solving for $\frac{df(x)}{dx}$ we get

$$\frac{df(x)}{dx} = -\frac{\frac{\partial h(x, f(x))}{\partial x}}{\frac{\partial h(x, f(x))}{\partial y}} \qquad \text{(Implicit Derivative)}$$

Note that this trick does not work if $\frac{\partial h(x, f(x))}{\partial y} = 0$ since the derivative is undefined.

Example 4 *Assume that our equation is* $h(x,y) = 3x^2 \ln y = 3$. *Then using the derivative we took above we have* $\frac{dH(x)}{dx} = 6x \ln y + \frac{3x^2}{y}\frac{df(x)}{dx} = 0$. *Solving for* $\frac{df(x)}{dx}$ *we get* $\frac{df(x)}{dx} = -(6x\ln y) / \left(\frac{3x^2}{y}\right)$. *This expression is in both* x *and* y*, but we could evaluate it meaningfully for any value of* x *and* y *satisfying* $3x^2 \ln y = 3$ *for instance* $x = 1$ *and* $y = e$ *which means that* $\frac{df(1)}{dx} = -(6 \cdot 1) / \left(\frac{3}{e}\right) = -2e$. *Of course this derivative may be different for different values of* x.



Implicit derivative at $x = 1, y = e$ *is* $-2e$

2 Optimization

2.1 Increasing and Decreasing Functions

By looking at the definition of a derivative we can say "loosely speaking" that for h small $h \approx 0$ and positive $h > 0$ that

$$\frac{df(x)}{dx} \approx \frac{f(x+h) - f(x)}{h}$$

or that

$$\frac{df(x)}{dx} h = f(x+h) - f(x)$$

Now we say that a function f is **increasing** at x if we have $f(x+h) > f(x)$ or $f(x+h) - f(x) > 0$. However since $h > 0$ this implies that $\frac{df(x)}{dx} > 0$. Equivalently a function f is **decreasing** at x if we have $f(x+h) < f(x)$ which then implies that $\frac{df(x)}{dx} < 0$. Therefore the sign of the derivative tells us whether or not a function is increasing (the derivative is positive) or decreasing (the derivative is negative).

2.2 Unconstrained Optimization

2.2.1 One Variable

Consider first the case with one variables x where $x \in R$, i.e. x can take on any real value. We wish to maximize the **objective function** $f(x)$ and there are no constraints. Thus we solve

$$\max_x f(x)$$

Assuming $f(x)$ has a maximum (for example f does not go to ∞)[1] and that it is differentiable everywhere, then any maximum (x^*) must solve the following **first order necessary conditions (FOC)**

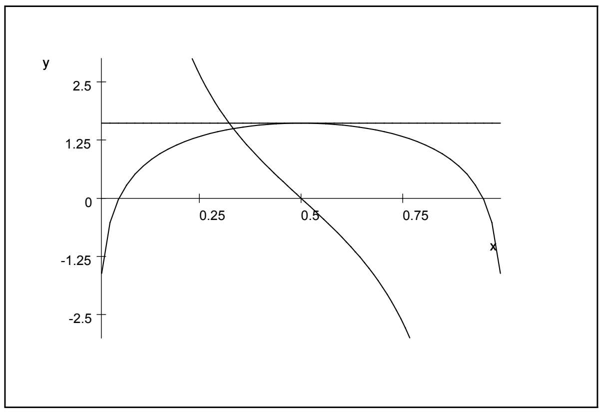
$$\frac{df(x^*)}{dx} = 0 \tag{FOC}$$

[1] **Weierstrauss's Theorem** say shat $f(x,y)$ will be guaranteed to have a maximum if f is continuous and x and y are defined over a closed and bounded domain. See an analysis book for more.

This is a single equation in a single unknown and so it should be solvable for x^*. and y^*.

Why does this make sense? Well say $\frac{df(x^*)}{dx} \neq 0$ then that means that either $\frac{df(x^*)}{dx} > 0$ or $\frac{df(x^*)}{dx} < 0$, i.e. f is either increasing or decreasing. If f is increasing however we could increase the value of f by increasing x a "little bit" to $x^* + h$ in which case x^* never gave us our maximum. Similarly if f is decreasing we could reduce x a little bit and increase the value of x. Either case is impossible at a maximum and therefore we must have that $\frac{df(x^*)}{dx} = 0$. Remeber that a *necessary* condition is different from a *sufficient* condition in that we could have $\frac{df}{dx}(x) = 0$ and not be at a maximum, but at a minimum or a "saddle point."[2] Also we have to be careful to check points where f is not differentiable. Third, there may be several points that satisfy the condition $\frac{df(x)}{dx} = 0$ Typically we will try and avoid these situations and deal with points where x^* is uniquely determined by (FOC), but not always.

Example 5 $f(x) = \ln x + \ln(1-x) + 3$. *The FOC implies* $\frac{df(x^*)}{dx^*} = \frac{1}{x^*} - \frac{1}{1-x^*} = 0 \Rightarrow x^* = 1 - x^* \Rightarrow x^* = 1/2$. *The maximum value of* f *achieves is then* $f(1/2) = 2\ln(1/2) + 3 \cong 1.614$



Graphic Representation of FOC

2.2.2 Two Variables

Now consider the case with two variables x and y, where both $x, y \in R$. Here we wish to solve

$$\max_{x,y} f(x,y)$$

With similar conditions we can expect the maximum (x^*, y^*) must solve the following first order necessary conditions (FOC)

$$\frac{\partial f(x^*, y^*)}{\partial x} = 0 \tag{FOC1}$$

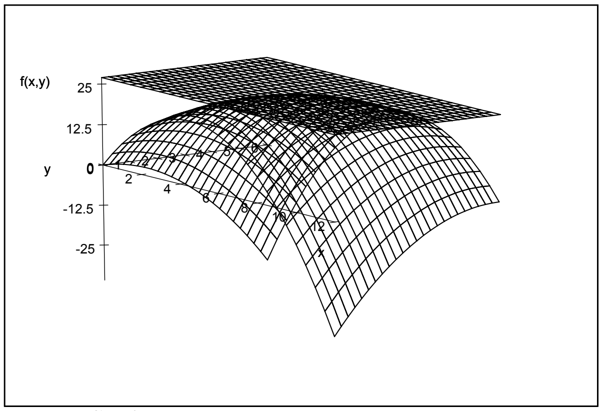
$$\frac{\partial f(x^*, y^*)}{\partial y} = 0 \tag{FOC2}$$

This is a system of two equations in two unknowns, and thus should usually be solvable for both x^* and y^*. The intuition for these conditions is similar to that above.

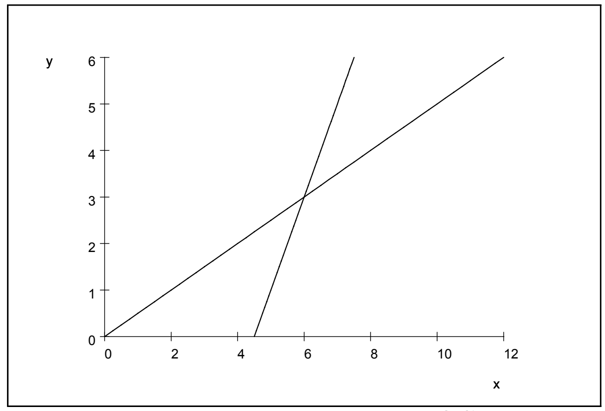
Example 6 *Let* $f(x,y) = -x^2 + 9x + xy - y^2$. *The FOC are* $\frac{\partial f(x^*)}{\partial x} = -2x^* + y^* + 9 = 0$, $\frac{\partial f(x^*)}{\partial y} = -2y^* + x^* = 0$. *The second equation implies* $x^* = 2y^*$ *which can be substituted into the first equation to yield* $-4y^* + y^* + 9 = 0 \Rightarrow y^* = 3$ *and* $x^* = 6$. *The maximum value attained is* $f(6,3) = -6^2 + 54 + 18 - 3^2 =$

[2] The **second order necessary condition (SOC)** states that the derivate of f should be decreasing, which implies that the second derivative be negative $\frac{d^2 f}{dx^2}(x^*) = \frac{d}{dx}\left[\frac{df}{dx}(x^*)\right] \leq 0$. This implies that f will be increasing, then flat (at the maximum) and then decreasing afterwards. If f were first decreasing, then flat, and then increasing, that would imply that x^* was at a minimum.

27. *Note that we can be sure that this is a unique and global maximum because f is everywhere concave as the graph suggests.*



Surface tangent to horizontal plane.



Equations determined by FOC

2.3 Optimization with Equality Constraints

Say there is a **constraint** on x and y which we can write as $g(x,y)=k$. There are two ways of solving the ensuing maximization problem. The first via **substitution**, the second via a **Lagrangean.**

2.3.1 Substitution

Pursuing the first route we use the constraint equation to solve for one variable in terms of the other, i.e. we find another function $y=h(x)$, such that $g(x,h(x))=k$. This can then be substituted into the objective function to get a maximization in a single variable

$$\max_x f(x,h(x))$$

which taking the total derivative yields the FOC

$$\frac{\partial f(x^*,h(x^*))}{\partial x}+\frac{\partial f(x^*,h(x^*))}{\partial y}\frac{dh(x^*)}{dx}=0 \tag{Substitution FOC}$$

We could use this one equation to solve for x^*, and then find $y^*=h(x^*)$.

Differentiating the budget constraint implicitly we can solve for the slope of $h(x)$

$$\frac{\partial g(x,h(x))}{\partial x}+\frac{\partial g(x,h(x))}{\partial y}\frac{dh(x)}{dx}=0\Rightarrow\frac{dh(x)}{dx}=-\frac{\frac{\partial g(x,h(x))}{\partial x}}{\frac{\partial g(x,h(x))}{\partial y}}$$

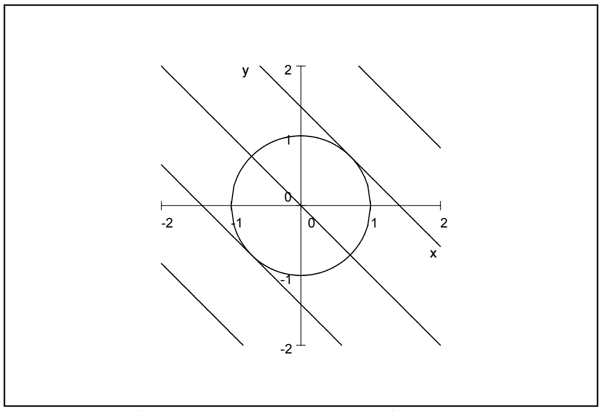
Plugging this into the FOC and rearranging means that at the optimum

$$\frac{\frac{\partial f(x^*,h(x^*))}{\partial x}}{\frac{\partial f(x^*,h(x^*))}{\partial y}}=\frac{\frac{\partial g(x^*,h(x^*))}{\partial x}}{\frac{\partial g(x^*,h(x^*))}{\partial y}} \tag{Tangency}$$

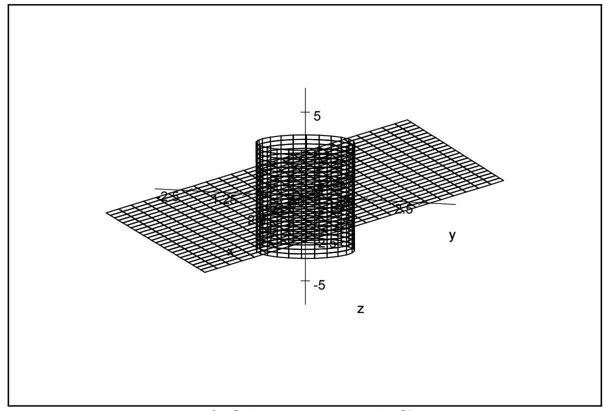
which is a general tangency condition, which we see many versions of throughout economics. It says that the constraint g and the level curve of f at $f^*=f(x^*,y^*)$ should be **tangent** at the maximum, which means the curves meet and have the same slope.. A **level curve** of f at z is just all the points x,y such that $f(x,y)=z$ (like with indifference curves)

Example 7 *Let $f(x,y)=x+y$ subject to $g(x,y)=x^2+y^2=1=c$. Then $h(x)=y=\sqrt{1-x^2}$, and so $f\left(x,\sqrt{1-x^2}\right)=x+\sqrt{1-x^2}$. The FOC is $1-\frac{x}{\sqrt{1-x^2}}=0\Rightarrow\sqrt{1-x^2}=x\Rightarrow 1-x^2=x^2\Rightarrow x^2=\frac{1}{2}\Rightarrow$*

$x^* = \frac{1}{\sqrt{2}}$ *and* $y^* = \frac{1}{\sqrt{2}}$ *achieving* $f\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right) = \sqrt{2}$ *Note that there is a second tangency at the minimum* $x^* = -\frac{1}{\sqrt{2}}$ *and* $y^* = -\frac{1}{\sqrt{2}}$.



Constraint and Level Curves



3-D Plot of Objective and Constraint

2.3.2 Lagrangean

An equivalent way of solving this problem which at first may seem more difficult, but which in fact can be very useful (and sometimes easier) is to maximize a **Lagrangean**

$$\mathfrak{L}(x, y, \lambda) = f(x, y) + \lambda [c - g(x, y)] \qquad \text{(Lagrangean)}$$

where λ is a **Lagrange multiplier**, which we maximize just as we would an unconstrained problem, taking into account the extra variable λ. The FOC are then

$$\frac{\mathfrak{L}(x^*, y^*, \lambda^*)}{\partial x} = \frac{\partial f}{\partial x}(x^*, y^*) - \lambda^* \frac{\partial g}{\partial x}(x^*, y^*) = 0 \qquad \text{(Lagrange FOC1)}$$

$$\frac{\mathfrak{L}(x^*, y^*, \lambda^*)}{\partial y} = \frac{\partial f}{\partial y}(x^*, y^*) - \lambda^* \frac{\partial g}{\partial y}(x^*, y^*) = 0 \qquad \text{(Lagrange FOC 2)}$$

$$\frac{\mathfrak{L}(x^*, y^*, \lambda^*)}{\partial \lambda} = c - g(x^*, y^*) = 0 \qquad \text{(Lagrange FOC3)}$$

This is a system of three equations with three unknowns, and therefore should be relatively straightforward to solve. Notice that third FOC is just the constraint restated. Solve (Lagrange FOC1) and (Lagrange FOC2) for λ and you will get

$$\lambda^* = \frac{\frac{\partial f}{\partial x}(x^*, y^*)}{\frac{\partial g}{\partial x}(x^*, y^*)} = \frac{\frac{\partial f}{\partial y}(x^*, y^*)}{\frac{\partial g}{\partial y}(x^*, y^*)} \qquad \text{(Lambda)}$$

The second equation eliminates λ for the purposes of problem solving, which combined with the constraint (Lagrange FOC3) gives a 2 equation system in two unknowns (x^*, y^*). λ^* can then be solved for by plugging back into (Lambda). Notice that the second part of (Lambda) can be rearranged to produce the same result as (Tangency), making the equivalence of the two approaches obvious.

2.3.3 Interpreting the Lagrange Multiplier

So why all the fuss about the Lagrange multiplier? One reason why is that the Lagrange multiplier has an interesting interpretation. If we were to increase c by one unit how much higher would f go? The answer is λ^*: it tells us the value of relaxing the constraint by one unit (e.g. one dollar).

More formally, take c to be a parameter (like prices or income) that is fixed over the optimization. Then really all of our solutions x^*, y^*, λ^* are dependent on c. We can then write $x^*(c), y^*(c), \lambda^*(c)$, as the parameter c changes, these solutions will change. Now plug $x^*(c), y^*(c)$ into f and we will get a new function $F(c) = f(x^*(c), y^*(c))$ which gives us the maximum value of f we can get for a given c. Now take the total derivative of F with respect to c:

$$\begin{aligned}
\frac{dF(c)}{dc} &= \frac{\partial f(x^*(c), y^*(c))}{\partial x}\frac{dx^*(c)}{dc} + \frac{\partial f(x^*(c), y^*(c))}{\partial y}\frac{dy^*(c)}{dc} \\
&= \lambda^*(c)\frac{\partial g(x^*(c), y^*(c))}{\partial x}\frac{dx^*(c)}{dc} + \lambda^*(c)\frac{\partial g(x^*(c), y^*(c))}{\partial y}\frac{dy^*(c)}{dc} \\
&= \lambda^*(c)\left[\frac{\partial g(x^*(c), y^*(c))}{\partial x}\frac{dx^*(c)}{dc} + \frac{\partial g(x^*(c), y^*(c))}{\partial y}\frac{dy^*(c)}{dc}\right]
\end{aligned}$$

where the second equality comes from using (Lagrange FOC1) and (Lagrange FOC2). The constraint implies that $g(x^*(c), y^*(c)) = c$. Differentiating the constraint totally with respect to c gives us

$$\frac{\partial g(x^*(c), y^*(c))}{\partial x}\frac{dx^*(c)}{dc} + \frac{\partial g(x^*(c), y^*(c))}{\partial y}\frac{dy^*(c)}{dc} = 1$$

so the entire term in brackets above equals one leaving us with

$$\frac{dF(c)}{dc} = \lambda^*(c) \qquad \text{(Lambda Interpretation)}$$

which is what we wanted to show: an increase in one unit of c increases F (or f) by λ^*.

Example 8 *Let $f(x, y) = x + y$ subject to $g(x, y) = x^2 + y^2 = 1 = c$. Then*

$$\mathfrak{L}(x, y, \lambda) = x + y + \lambda\left(1 - x^2 - y^2\right)$$

with the FOC

$$\begin{aligned}
\frac{\partial \mathfrak{L}}{\partial x} &= 1 - \lambda 2x = 0 \\
\frac{\partial \mathfrak{L}}{\partial x} &= 1 - \lambda 2y = 0 \\
\frac{\partial \mathfrak{L}}{\partial x} &= 1 - x^2 - y^2 = 0
\end{aligned}$$

Solving for λ we have $\lambda = \frac{1}{2x} = \frac{1}{2y} \Rightarrow x = y$ Putting this into the third equation we have $1 - x^2 - x^2 = 0 \Rightarrow 2x^2 = 1 \Rightarrow x^2 = \frac{1}{2} \Rightarrow x^* = \frac{1}{\sqrt{2}} = y^*$. Also $\lambda^* = \frac{1}{2\frac{1}{\sqrt{2}}} = \frac{1}{\sqrt{2}}$.

2.4 Optimization with Inequality Constraints

In economics it is much more common to start with inequality constraints of the form $g(x, y) \leq c$. The constraint is said to be **binding** if at the optimum $g(x^*, y^*) = c$, and it is said to be **slack** if at the optimum $g(x^*, y^*) < c$, clearly it must be one or the other. In this case it is not clear whether or not we can use the substitution method, since that would be invalid if the constraint is slack. The Lagrangean used above is still useful however except that the FOC's above are amended to yield the following **"Kuhn-Tucker" first order conditions**

$$\frac{\partial f}{\partial x}(x^*, y^*) - \lambda^*\frac{\partial g}{\partial x}(x^*, y^*) = 0 \qquad \text{(KT1)}$$

$$\frac{\partial f}{\partial y}(x^*, y^*) - \lambda^*\frac{\partial g}{\partial y}(x^*, y^*) = 0 \qquad \text{(KT2)}$$

$$c - g(x^*, y^*) \geq 0, \ \lambda^* \geq 0, \ \lambda^*\left[c - g(x^*, y^*)\right] = 0 \qquad \text{(KT3, KT4, KT5)}$$

The first two conditions are identical to the Lagrangean FOC's. The third condition now has three parts. The first is just a restatement of the constraint. The second part says that λ^* is always non-negative.

Mathematically this this ensures a correct answer: if you get an (x, y) combination that implies a negative λ, then this is not a solution. Note that if we had written $\mathfrak{L}(x, y, \lambda) = f(x, y) - \lambda[c - g(x, y)]$ by mistake, the correct λ would be non-positive. This would also occur if the constraint is written as $c - g(x, y) \leq 0$ instead of $c - g(x, y) \geq 0$. Although mathematically the choice of sign is a formality, economically we want the Lagrange multiplier to have a positive sign because of its economic interpretation as a shadow price.

The third part (the **"complementary slackness"** condition) says that either λ^* or $c - g(x^*, y^*)$ must be zero. If $\lambda^* = 0$, then we can forget about the constraint and the Kuhn-Tucker conditions turn into the unconstrained FOC considered earlier. If $\lambda^* > 0$ then the constraint must be binding then the problem turns into the standard Lagrangean considered above. Thus the Kuhn-Tucker conditions provide a neat mathematical way of integrating the unconstrained and constrained cases into a single formulation. Unfortunately you typically have to check both cases, unless you know in advance if it's one or the other, which often happens.

Note that the interpretation of λ^*, still stands. If $\lambda^* = 0$, this means that the constraint is slack $g(x, y) < c$. Changing c by a "tiny bit" won't make the constraint binding and so the constraint will still be slack, and therefore $F(c)$ won't change at all.

Example 9 *A **nonnegativity constraint** is just a constraint of the form $x \geq 0$. Consider the one variable case where we wish to maximize $f(x)$ subject to $x \geq 0$. Then the Lagrangean will be $\mathfrak{L}(x, \lambda) = f(x) + \lambda x$. The first order Kuhn-Tucker conditions are then*

$$f'(x^*) + \lambda = 0, \;\; x^* \geq 0, \; \lambda^* \geq 0, \; x^* \lambda^* = 0$$

If the nonnegativity constraint does not bind then $f'(x^*) = 0$. If it does bind, then $f'(x^*) = -\lambda^* < 0$ and $x^* = 0$. Thus we can rephrase the FOC's without mentioning λ^* as

$$f'(x^*) \leq 0 \text{ with } "=" \text{ if } x^* > 0$$

As a specific case consider maximizing the function $f(x) = -2x$ subject to the non-negativity constraint $x \geq 0$. In this case the FOC is $-2 + \lambda^* = 0$. So $\lambda^* = 2$ and therefore as $x^* \lambda^* = 0$, then $x^* = 0$.

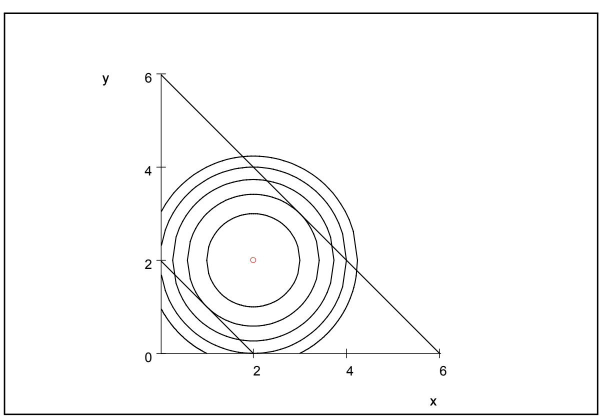
Example 10 *Let $f(x, y) = -(x-2)^2 - (y-2)^2 + 5$ and $g(x, y) = x + y \leq c$, where either $c = 2$ or $c = 6$. The Lagrangean is*

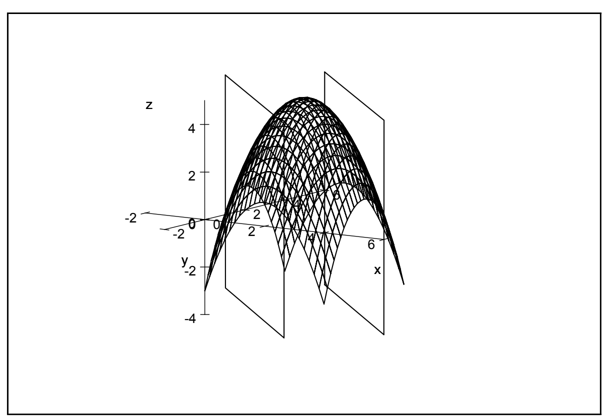
$$\mathfrak{L}(x, y, \lambda) = -(x-2)^2 - (y-2)^2 + 5 + \lambda(c - x - y)$$

and the FOC are

$$\begin{gathered} -2(x-2) - \lambda = 0 \\ -2(y-2) - \lambda = 0 \\ c - x - y \geq 0, \; \lambda \geq 0, \; \lambda(c - x - y) = 0 \end{gathered}$$

The first two conditions imply $x = y = 2 - \lambda/2$, and so $x + y = 4 - \lambda$. Plugging this into the third condition gives $c \geq 4 - \lambda$ or $\lambda \geq 4 - c$. If $c = 2$ this means that $\lambda \geq 2 > 0$ and so the fifth condition implies that $2 - x - y = 0$, i.e. the constraint binds, and so $x = y = 1$. If $c = 6$, then the third condition gives $\lambda \geq -2$. If $\lambda \neq 0$, then the fifth condition implies $x = y = 3 = 2 - \lambda/2$, by the first condition. But then $\lambda = -2$ which conflicts with $\lambda \geq 0$. Therefore it must be the case that $\lambda = 0$, and the unconstrained optimum $x = y = 2$ is achieved.





2.5 More than One Constraint (Optional)

When there is more than one constraints say $g(x,y) \leq c$ and $h(x,y) \leq d$ we can just write a longer Lagrangean with two multipliers λ and μ

$$\mathfrak{L}(x,y,\lambda,\mu) = f(x,y) + \lambda[c - g(x,y)] + \mu[d - h(x,y)]$$

The FOC's are derived in the same way by taking the derivatives with respect to each variable (including the multipliers) and setting each equal to zero. If the constraints are not known to be binding then conditions like KT3-KT5 need to be written for each constraint (see the example below). While this may get tedious it should eventually boil down to a system of 4 equations in 4 unknowns. Moreover, if both constraints bind and are independent, then (x^*, y^*) will simply be whatever satisfies both constraints $g(x^*, y^*) = c$, and $h(x^*, y^*) = d$, a system of two equations in two unknowns and the first two FOC equations are only useful for determining the multipliers.

More control variables like x and y only implies more derivatives. More constraints simply imply adding more Lagrange multipliers with or without the Kuhn-Tucker subtleties. There are other issues related to optimization that are in fact quite interesting and worth learning more about, but we won't go over them here.[3]

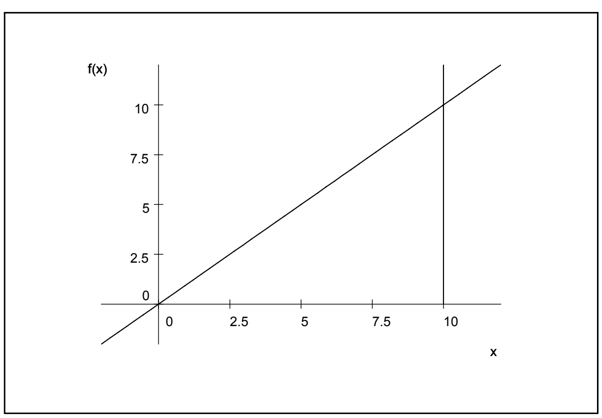
Example 11 ***"Bang-Bang" Solution*** *Consider the very straightforward problem of maximizing* $f(x) = ax + b$ *a simple linear function where* $a \gtreqless 0$ *is a constant of indeterminate sign. Without a constraint there is no solution to this problem (infinity isn't technically considered a "solution"). However consider the constraints* $\times \geq 0$ *and* $x \leq 10$. *The Lagrangean will then be*

$$\mathfrak{L}(x,\lambda,\mu) = ax + b + \lambda x + \mu(10 - x)$$

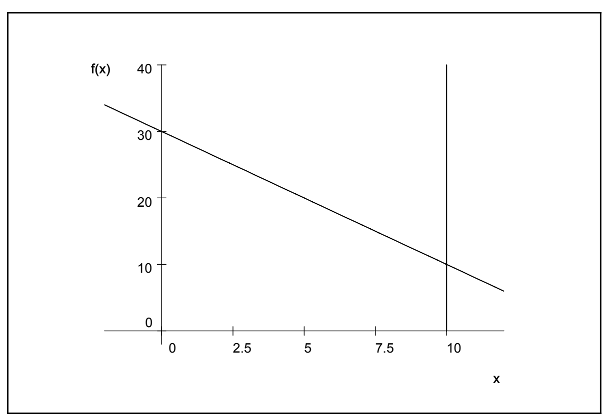
with FOC

$$a + \lambda^* - \mu^* = 0$$
$$x^* \geq 0,\ \lambda^* \geq 0,\ x^*\lambda^* = 0$$
$$x^* \leq 10,\ \mu^* \geq 0,\ (10 - x)^* \mu^* = 0$$

Realize that both constraints cannot be binding at the same time so either $\lambda^* = 0$ *or* $\mu^* = 0$. *Also the FOC implies* $a = \mu^* - \lambda^*$ *There are then 3 cases worth considering here (1)* $a > 0$ *then it must be that* $a = \mu^* > 0$ *and* $\lambda^* = 0$. *Otherwise* $a = -\lambda^* > 0$ *contradicts* $\lambda^* \geq 0$. *The complementary slackness* μ *constraint must be binding and so* $x^* = 10$ *(2)* $a < 0$ *where by similar reasoning* $a = \lambda^*$ *and* $x^* = \mu^* = 0$ *and (3)* $a = 0$ *in which case* $\mu^* = \lambda^* = 0$ *and* x^* *can take any value between 0 and 100 as they all yield* $ax^* = 0x^* = 0$. *This kind of a solution is known as a "bang-bang" solution - either you "bang"* x *at the top or at the bottom. You go for the maximum or minimum possible value of* x.



$a = 1,\ b = 0, x^* = 10$



$a = -2, b = 30,\ x^* = 0$

[3] A typical multivariate calculus book should have a chapter on this subject for equality constraints. Simon and Blume (1994) offers a full treatment and is a good read and reference with many economic applications.