Edward R. Morey

I. Minimization & Maximization

(without & with constraints)

1. Procedure for Maximizing (or Minimizing) a Function of One Variable

2. Maximizing or Minimizing Functions of Two Variables

3. Maximizing or Minimizing Functions of N Variables

4. Maximizing or Minimizing Functions of 2 Variables Subject to a Constraint

Method I Direct Substitution
Method II The Lagrangian Technique

5. Maximizing or Minimizing Functions of N Variables Subject to a Constraint

Method I Direct Substitution
Method II The Lagrangian Technique

*Thanks to Laura Langhoff for layout, typing and graphics.

I. Minimization & Maximization

Most of the allocational problems that we consider in economics can be formulated as constrained optimization problems

That is, maximizing an objective function subject to a constraint

for example, the consumer maximizes utility subject to a budget constraint

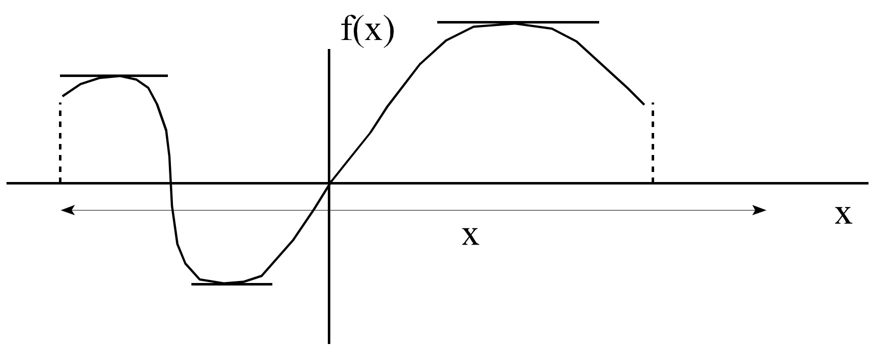
Many estimation problems in econometrics can formulated as constrained optimization problems

for example, OLS where one minimizes the sum of squared residuals subject to the constraint that the fitted line is linear and maximum likelihood estimation which assumes a particular density function

1. Procedure for Maximizing (or Minimizing) a Function of One Variable

Let's begin our consideration of optimization by outlining the procedure for maximizing or minimizing unconstrained functions of one variable.

Given a **differentiable** function f(x) defined over some domain of x, where x is a scalar find the maximum and the minimum value of f(x). [We will generally, but not always, assume differentiability.]



What does it mean to say a function is differentiable? Why do economists often assume differentiability? Do differentiable functions drive the behavior of economic agents?

We have to distinguish between necessary and sufficient conditions for a point to be the maximum (or minimum) of some function f(x)

necessary means the condition is required - necessary doesn't imply that the condition alone will guarantee the result

if A is necessary for B $\Leftrightarrow$ not A $\Rightarrow$ not B

sufficient means presence of the condition is enough to guarantee the result, however, sufficient conditions are not necessarily required to guarantee the result

if A is sufficient for B $\Leftrightarrow A \Rightarrow B$

Let x^o denote a specific value of x.

A **necessary** condition for a point $f(x^o)$ to be a max or min in the **interior** of the domain x is that

(i) $f'(x^0) = 0$

i.e. the instantaneous slope of the curve [f(x), x] must be zero at the point $[f(x^0), x^0]$. For a min, we want to be at the **bottom** of a valley where the slope will be zero, and for a max, we want to be on the top of a hill where the instantaneous slope will also be zero.

Would this be true if we dropped the assumption of differentiability?

An aside: What is a derivative?

Consider the function $y = f(x)$, where x is a scalar. Denote the derivative of $f(x)$ with respect to x as $\frac{dy}{dx} \equiv f'(x)$.

The slope of $f(x)$ at x^0, from x^0 to $(x^0 + t)$ is

$$\frac{f(x^0 + t) - f(x^0)}{t}$$

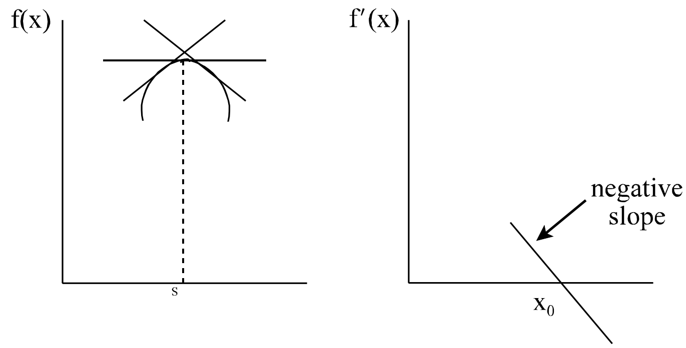
Note that the slope is undefined when $t = 0$; the definition of a derivative requires the limit concept, where $\lim\limits_{m \to 0} f(m)$ is the number $f(m)$ approaches as m approaches zero.

$$f'(x^0) = \lim_{t \to 0} \frac{f(x^0 + t) - f(x^0)}{t}$$

Sufficient conditions for a **local interior** max at $x = x^0$ are:

(i) $f'(x^0) = 0$ and

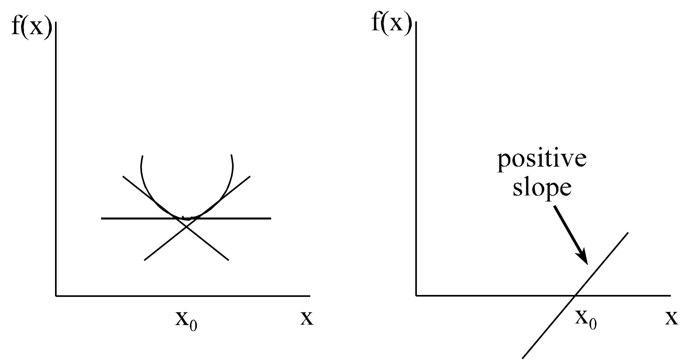
(ii) $f''(x^0) < 0$ which implies that the slope of the function is decreasing at x^0 (i.e. that we are at the top of a hill), provided that the function is twice differentiable in the neighborhood of x^0.



Sufficient conditions for a **local interior minimum** at $x = x^0$ are:

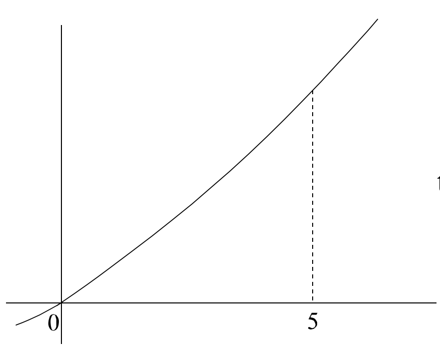
(i) $f'(x^0) = 0$ and

(ii) $f''(x^0) > 0$



> Why was I so careful about using the adjective "interior" when defining the necessary and sufficient conditions for a local max or min?

$f'(x^0) = 0$ is <u>not</u> necessary for a local min or max if we drop the adjective "*interior*." For example: find the maximum of $f(x) = ax^2$ for $0 \le x \le 5$



the max is at $x = 5$ but at $x = 5$ $f'(5) \neq 0$

Whenever we examine max or min we have to be careful that we understand all the adjectives. One should check for the presence or absence of adjectives such as *local*, *interior*, *global*, and *unique*.

Example: Given the total cost function $C(x) = a + bx + cx^2$ where output $x \geq 0$ and a, b, c are positive constants, find the $x > 0$ which minimizes the average cost defined as $AC(x) \equiv C(x)/x = ax^{-1} + b + cx$. Call this x. x^0.

Solution: Find the critical points, that is, points x^0 such that $f'(x^0) = 0$

Now $AC'(x) = -ax^{-2} + 0 + c = 0$

$\Rightarrow x^{-2} = c/a$

$\Rightarrow x^2 = a/c$

$\Rightarrow x^0 = (a/c)^{½} > 0$

the symbol "⇒" means implies

There is only one critical point which is positive, $x^0 = (a/c)^{½}$. Now lets check to see whether this point is a local min

evaluated at $x = x^0 \equiv (a/c)^{½}$

$AC''(x) = 2ax^{-3}$ is $AC''(x^0) = 2a(x^0)^{-3} = 2a((a/c)^{1/2})^{-3} = 2a^{-1/2}c^{3/2} > 0$ if a and c are positive

So, if a, b and c are positive, we have a local min which turns out to be a global min in this case, in which case we might denote the specific value of x that min AC(x), x*.

How do we know it is a global min?

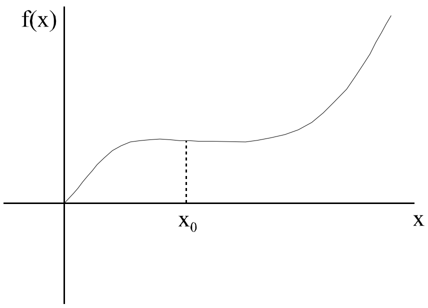
How, in general, do we find a **local** min of f(x)?

(i) Find all points x such that $f'(x) = 0$.
These are called critical points

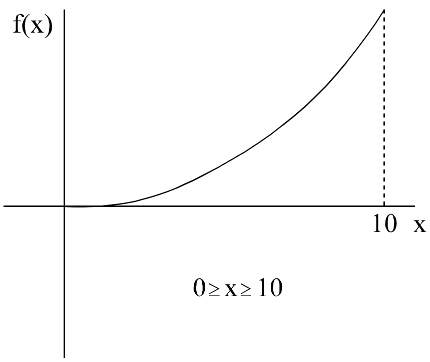
(ii) Those critical points x^0 for which $f''(x^0) > 0$ are all **local** min. Pick the x^0 which gives rise to the smallest local min. This point will generally be a **global** min but one must also check (iii).

(iii) Check the endpoints in the domain of x. Also check all the critical points x^0 for which $f''(x^0) = 0$. Any of these points could be the global min.

The following examples illustrate the need for (iii).

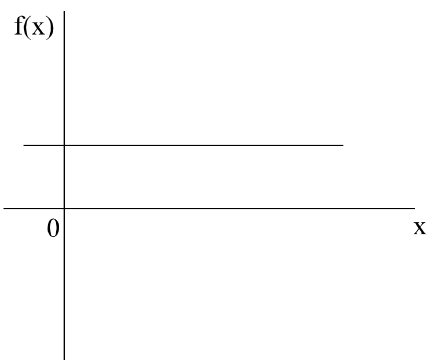


Normally if $f'(x^0) = 0$ and $f''(x^0) = 0$, x^0 is an inflection point, as in this example, but there are some functions for which $f'(x^0) = 0$, $f''(x^0) = 0$, and x^0 is a local min. (See below for an example.)

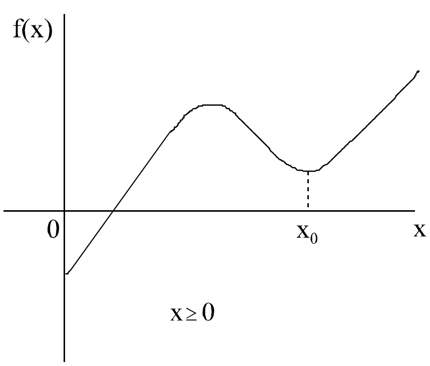


$0 \leq x \leq 10$

For this function there are no critical points in the domain of x but there is a local min which is also the global min.



For this function $f'(x) = 0$ and $f''(x) = 0$ for all x. Every value of x is a local and global max, every value of x is a local and global min. However, there is no unique max or min.



$x \geq 0$

For this function $x = x^0$ is a critical point and a local interior min but it is not the global min. The global min is at the end point $x = 0$.

In economics we **often** make assumptions that guarantee the min or max is an interior min or max. These assumptions will not always be "*realistic*" but they greatly limit the math tools we will need. If one admits the possibility of corner solutions, it is possible that at the max or min, x^0, $f'(x^0) \neq 0$. Checking corners is easy if the number of corners is small. When x is a scalar there is a maximum of two corners (e.g., if $-\infty \leq x \leq \infty$ there are no corners, if $x \geq a$ there is one corner, at a , and if $b \geq x \geq a$ there are two corners, one at a and one at b)

Up to this point we have assumed that x in f(x) is a scalar. Now we need to generalize and make x a vector.

$$x \equiv \begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ \cdot \\ x_N \end{bmatrix} \equiv (x_1,\ x_2,\ .\ .\ .\ x_N)'$$

but to keep things simple we'll often assume N = 2. i.e. $f(x) = f(x_1, x_2)$. How might one denote a specific value of x_1? x_1^0. However, unless it causes ambiguity, we will omit the superscript ^{0}.

2. Maximizing or Minimizing Functions of Two Variables

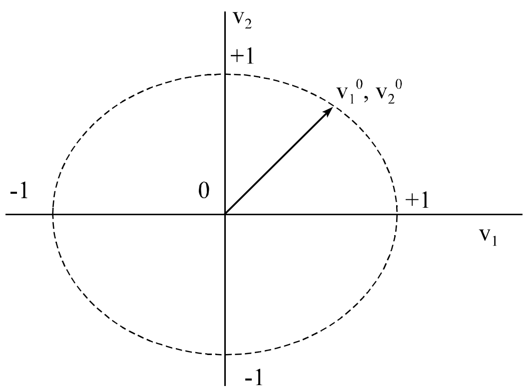
Directional Derivatives

Sydsaeter and Hammond page 541

To find the max or min of a multivariate function requires a new concept - directional derivatives.

Directional derivatives identify conditions for a local interior max (min) that are both necessary and sufficient.

Definition: A direction $v \equiv (v_1, v_2)$ is merely two numbers v_1 and v_2 such that $v_1^2 + v_2^2 = 1$.Note: the equation $v_1^2 + v_2^2 = 1$ is the equation for a circle of radius one. Geometrically, a direction can be represented as a point on a circle of radius one.



For example: $v_1 = 1$ $v_2 = 0$ denotes the direction east
$v_1 = 0$ $v_2 = 1$ denotes the direction north
$v_1 = .8$ $v_2 = .6$ denotes a north-easterly direction

Further note that the distance between two points in *x-y* space, (x, y) and $(x + tv_1, y + tv_2)$, is $\sqrt{[(x + tv_1) - x]^2 + [(y + tv_2) - y]^2}$.

Now that we know what a direction is and how to measure the distance between two points, we can define directional derivatives.

Definition: Given a function of two variables f(x, y), (where x and y are each scalars), defined over a region S, and a direction $v \equiv (v_1, v_2)$, the directional derivative of f(x, y) in the direction v evaluated at the point (x, y) is defined as

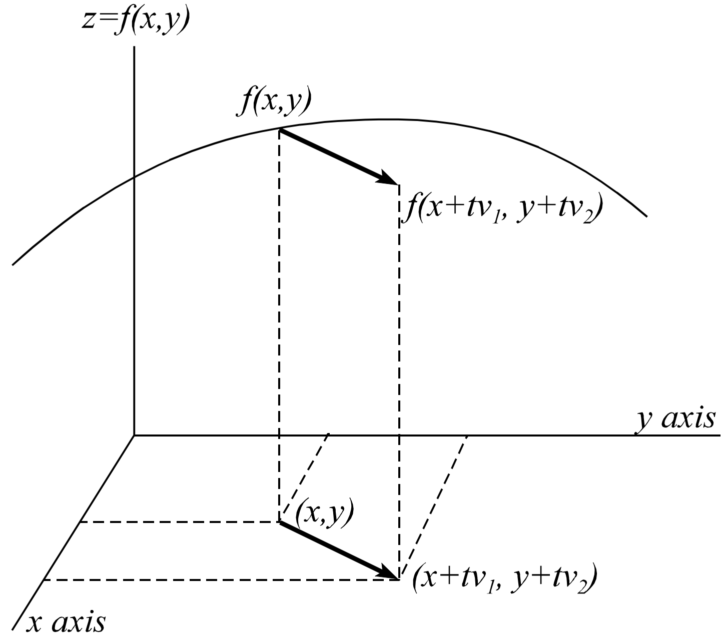
$$D_v f(x,y) = \lim_{t \to 0} \frac{f(x+tv_1, y+tv_2) - f(x,y)}{\sqrt{[x+tv_1) - x]^2 + [(y+tv_2) - y]^2}}$$

$$= \lim_{t \to 0} \frac{f(x+tv_1, y+tv_2) - f(x,y)}{\sqrt{t^2(v_1^2 + v_2^2)}}$$

$$= \lim_{t \to 0} \frac{f(x+tv_1, y+tv_2) - f(x,y)}{t}$$

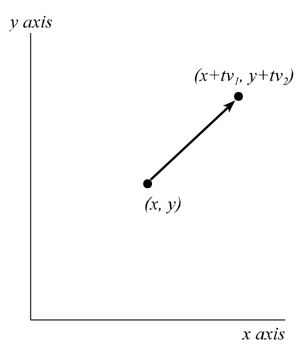
Note that the term $\frac{f(x+tv_1, y+tv_2) - f(x,y)}{t}$ is the slope of the function between the points (x, y) and $(x+tv_1, y+tv_2)$ where t denotes the distance between the two points and v denotes the direction one must go from (x, y) to get to $(x+tv_1, y+tv_2)$.

Graphically

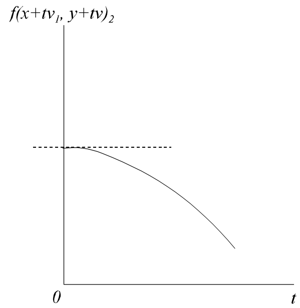


or looking at the function from some different perspectives

From above



OR in terms of t



Geometrically, finding the directional derivative D_v f(x, y) works as follows: start at a point (x, y) and draw a line pointing in the direction v. Any point on this line can be written as $(x+tv_1, y+tv_2)$. v denotes the direction the line takes from the point (x, y), t is the distance along the line from the point (x, y). Now look at the function of the single variable t defined by $g(t) = f(x+tv_1\ y+tv_2)$. The slope of the function

$$g'(t) = \frac{dg(t)}{dt}$$

$g'(t)$ is D_v f(x,y).

$D_v f(x,y)$ is the limit of the slope of f(x,y) in the direction (v_1, v_2) at the point (x, y).

Example: Let $f(x, y) = x + 2x^{1/2}y^{1/2} + y$ where $x \geq 0, y \geq 0$.

Calculate $D_v f(x, y)$ for the following directions v.

(i) the direction $v = (1, 0)$

$$D_v\ f(x,y) \equiv \lim_{t \to 0} \frac{f(x+tv_1, y+tv_2) - f(x,y)}{t}$$

$$= \lim_{t \to 0} \frac{f(x+t,y) - f(x,y)}{t}$$

$$= \lim_{t \to 0} \frac{(x+t) + 2(x+t)^{1/2}y^{1/2} + y - [x+2x^{1/2}y^{1/2}+y]}{t}$$

$$= \lim_{t \to 0} \frac{t+2y^{1/2}\{(x+t)^{1/2}-x^{1/2}\}}{t}$$

$$= 1 + \lim_{t \to 0} 2y^{1/2} \left\{ \frac{(x+t)^{1/2} - x^{1/2}}{t} \right\}$$

$$= 1 + g'(x)2y^{1/2} \quad \text{where } g(x) \equiv x^{1/2}$$

$$= 1 + \frac{1}{2} x^{-1/2}\, 2y^{1/2}$$

$$= 1 + x^{-1/2}\, y^{1/2}$$

(ii) the direction $v = (0, 1)$

$$D_v\ f(x,y) \equiv \lim_{t \to 0} \frac{f(x+tv_1, y+tv_2) - f(x,y)}{t}$$

$$= \lim_{t \to 0} \frac{f(x,y+t) - f(x,y)}{t}$$

$$= \lim_{t \to 0} \frac{x+2x^{1/2}(y+t)^{1/2} + (y+t) - [x+2x^{1/2}y^{1/2}+y]}{t}$$

$$= \lim_{t \to 0} \frac{2x^{1/2}\{(y+t)^{1/2}-y^{1/2}\}+t}{t}$$

$$= 1 + 2x^{1/2} \lim_{t \to 0} \left\{ \frac{(y+t)^{1/2} - y^{1/2}}{t} \right\}$$

$$= 1 + 2x^{1/2}\, g'(y) \quad \text{where } g(y) \equiv y^{1/2}$$

$$= 1 + 2x^{1/2} \frac{1}{2} y^{-1/2}$$

$$= 1 + x^{1/2}\, y^{-1/2}$$

Definition: The directional derivatives in the directions given by the coordinate axes are called the first order partial derivatives of f:

$$D_{(1,0)}\, f(x^0,y^0) \equiv \lim_{t\to 0} \frac{f(x^0+t,y^0) - f(x^0,y^0)}{t} \equiv \frac{\partial f(x^0,y^0)}{\partial x} \equiv f_1(x_0,y_0) \equiv f_x(x^0,y^0)$$

$$D_{(0,1)}\, f(x^0,y^0) \equiv \lim_{t\to 0} \frac{f(x^0,y^0+t) - f(x^0,y^0)}{t} \equiv \frac{\partial f(x^0,y^0)}{\partial y} \equiv f_2(x_0,y_0) \equiv f_y(x^0,y^0)$$

Note the three different notations for partial derivatives.

How to calculate first order partial derivatives

When partially differentiating a function of several variables with respect to one of these variables, just treat the other variables as constants and use the usual rules for differentiating functions of one variable.

Example: $f(x, y) = x + 2x^{1/2}\, y^{1/2} + y; \quad x \geq 0, y \geq 0.$

$f_x(x, y) = 1 + 2\ ½\ x^{-1/2}\, y^{1/2} + 0 = 1 + x^{-1/2}\, y^{½}$

$f_y(x, y) = 0 + 2x^{1/2}\ ½\ y^{-1/2} + 1 = x^{1/2}\, y^{-1/2} + 1$

Example: $f(x, y) = xy^2$

$f_x(x, y) = y^2 \qquad f_y(x, y) = 2xy$

Economic Example of partial derivatives and their economic interpretation: The two-input Cobb-Douglas Production Function

Suppose that output Y is produced by two inputs, $L \equiv$ number of units of labor used in the period of analysis, and $K \equiv$ number of units of capital services used in the period of analysis.

The technological relationship between the output and the inputs is summarized by means of a function $Y = f(L, K)$, $L \geq 0$, $K \geq 0$.

$f(L, K)$ is called the production function. It identifies **max** output as a function of the quantities of labor and capital used.

Often $f(L, K)$ is assumed to have the following functional form.

$Y = f(L, K) = aL^{\alpha}K^{1-\alpha}$ where $a > 0$ is a constant and $0 < \alpha < 1$ is another constant

This is a Cobb-Douglas production function with constant returns to scale.
The partial derivatives of a production function have fancy economic names:

Definition: the marginal product of labor is

$$f_L(L,K) \text{ or } \frac{\partial f(L,K)}{\partial L} = a\alpha L^{\alpha-1}K^{1-\alpha} \text{ in the Cobb Douglas case}$$

note that for this C-D, the marginal product of $L\downarrow$ as $L\uparrow$, holding K fixed ($\Rightarrow$ diminishing marginal product)

Definition: the marginal product of capital is

$$f_K(L,K) \text{ or } \frac{\partial f(L,K)}{\partial K} = (1-\alpha)aL^{\alpha}K^{-\alpha}.$$

(First Order) Necessary conditions for the Unconstrained Interior Maximum or Minimum of a Function of two Variables f(x, y)

Let $f(x, y)$ be a once continuously differentiable function defined over some set S. Then a necessary condition for a point (x^0, y^0) belonging to the interior of S to be a local minimum or maximum is that:

$$\text{(i)} \quad \frac{\partial f(x^0,y^0)}{\partial x} = 0 \quad \text{and (ii)} \quad \frac{\partial f(x^0,y^0)}{\partial y} = 0 \ .$$

For a max, we want to be at the top of a hill in all directions. This requires the slopes to be zero in the coordinate axis directions.

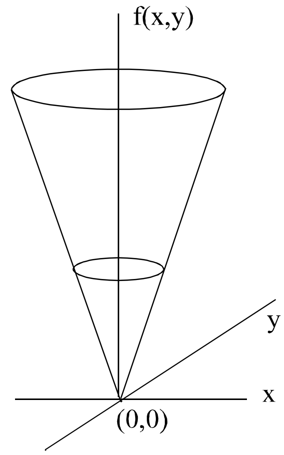
For a min, we want to be at the bottom of a hill in all directions. This requires the slopes to be zero in the coordinate axis directions.

In general, equations (i) and (ii) give us two equations in two unknowns, which we try to solve for x^0 and y^0.

Example: Minimize w.r.t. x and y the function $f(x, y) = x^2 + y^2$.

$$\frac{\partial f(x^0, y^0)}{\partial x} = 2x^0 = 0$$

$$\frac{\partial f(x^0, y^0)}{\partial y} = 2y^0 = 0$$



The solution is $(x^0, y^0) = (0, 0)$

The possible minimum is therefore

$$f(x^0, y^0) = 0^2 + 0^2 = 0$$

Note that this really looks more like the bottom half of a ball, than a cone

The first order necessary conditions for a maximum or minimum do not enable us to determine whether we have found a local minimum or maximum (or neither).

<u>(Second Order) Sufficient Conditions for a Local Interior Max</u>

A point (x^0, y^0) locally maximizes a twice continuously differentiable function f(x, y) if for every direction v, we have

(i) $D_v f(x^0, y^0) = 0$ (we are at the top of a hill in all directions)

(ii) $D_v [D_v f(x^0, y^0)] < 0$ (the slope of the directional derivative function Dv f(x,y) is decreasing in every direction around the point (x^0, y^0))

Can you prove that (i) and (ii) are not necessary?

<u>(Second Order) Sufficient Conditions for a Local Interior Min</u>

A point (x^0, y^0) locally minimizes a twice differentiable function f(x, y) if for every direction v, we have

(i) $D_v f(x^0, y^0) = 0$

(ii) $D_v [D_v f(x^0, y^0)] > 0$

Recall the example, $f(x,y) = x^2 + y^2$. Let's check whether $(x^0, y^0) \equiv (0, 0)$ satisfies the second order sufficient conditions for a local min. We need to calculate $D_v\, f(x^0, y^0)$ and $D_v\, [D_v\, f(x^0, y^0)]$ for every direction v.

$$D_v\ f(x,y) \equiv \lim_{t \to 0} \frac{f(x+tv_1, y+tv_2) - f(x,y)}{t}$$

$$\text{noting that } f(x,y) = x^2+y^2$$

$$= \lim_{t \to 0} \frac{(x+tv_1)^2 + (y+tv_2)^2 - x^2 - y^2}{t}$$

$$= \lim_{t \to 0} \frac{2xtv_1 + t^2v_1^2 + 2ytv_2 + t^2v_2t^2}{t}$$

dividing through by t and then take the limit as $t \to 0$

$$D_v\, f(x, y) = 2xv_1 + 2yv_2$$

evaluate the directional derivative at point $(x^0, y^0) = (0, 0)$

$$D_v\, f(x^0, y^0) \equiv D_v\, f(0, 0) = 0 + 0 = 0 \quad \text{for every } v_1, v_2 \text{ such that } v_1^2 + v_2^2 = 1$$

So at the point (0, 0) the directional derivative is zero **in every** direction.

Now let's derive $D_v\, [D_v\, f(x, y)]$ and then evaluate it at the point (0, 0) to see if it's strictly positive in every direction.

$$D_v\ [D_v\ f(x,y)] \equiv \lim_{t \to 0} \frac{D_v\ f(x+tv_1,\ y+tv_2) - D_v\ f(x,y)}{t}$$

where

$$D_v\, f(x, y) = 2xv_1 + 2yv_2$$

and

$$D_v\, f(x+tv_1, y+tv_2) = 2(x+tv_1)v_1 + 2(y+tv_2)v_2$$

making these substitutions

$$= \lim_{t \to 0} \frac{2(x+tv_1)v_1 \;+\; 2(y+tv_2)v_2 \;-\; [2xv_1+2yv_2]}{t}$$

$$= \lim_{t \to 0} \frac{2(tv_1^2 \;+\; tv_2^2)}{t}$$

$$= 2(v_1^2 \;+\; v_2^2)$$

$$= 2$$

$D_v\,[D_v\,f(0_0, 0_0)] > 0$ for every direction v

Therefore

$(x^0, y^0) \equiv (0, 0)$ satisfies the second order sufficient conditions for a local minimum

Problem

Suppose a firm produces a single output Y. The quantity produced in period t is Y(t). Y(t) is produced using capital services K(t) and labor L(t) where t = 1, 2, 3. The following data is available:

t	Y(t)	K(t)	L(t)
1	1	1	1
2	2	2	1
3	2	1	2

Let's assume that the firm's technology can be approximated with the linear production function

$$Y_t = f(K_t, L_t) = aL_t + bK_t + e_t$$

where t = 1, 2, 3; a and b are positive constants and e_t is the approximation error in year t

As an aside, does this functional form seem like a reasonable functional form for a production function?

We do not know the true values of a and b but we can estimate them given our data on inputs and output for the three years. One method of obtaining estimates for a and b is to choose the $\hat{a}$ and $\hat{b}$ which minimizes the sum of the squares of the approximation errors. This gives rise to the following minimization problem:

minimize --

w.r.t. a and b $\quad e_1^2 + e_2^2 + e_3^2 = \sum_{t=1}^{3} (Y_t - aL_t - bK_t)^2$

plugging in the data

$$f(a,b) \sum_{t=1}^{3} e_t^2 = 9 - 14a - 14b + 10ab + 6a^2 + 6b^2$$

so we want to find the point $(\hat{a}, \hat{b})$ that minimizes the function

$$f(a, b) = 9 - 14a - 14b + 10ab + 6a^2 + 6b^2$$

a) Write out the first order necessary conditions for a min of f(a,b).

b) Solve for $(\hat{a}, \hat{b})$. If $(\hat{a}, \hat{b})$ minimize f(a,b), the solution $(\hat{a}, \hat{b})$ is called the least square estimates of a and b.

c) Use directional derivatives to prove that the point you found is a local minimum. Hint - to prove this you will need to demonstrate that $20v_1v_2 + 12 > 0$ if $v_1^2 + v_2^2 = 1$. Maybe *Mathematica* can help you to demonstrate this last step. (continued)

d) For 6808, Write a *Mathematica* notebook to solve this problem. Hint: maybe use the command "FindMinimum" - - absolute.nb
3) Now generalize your notebook so that it will solve the problem for any three observations that you read or write in at the beginning of the notebook.
4) Now consider a different minimization problem. Instead of minimizing the sum or squared residuals, find the estimates of a and b that minimize the sum of the absolute values of the deviations. Can you solve this problem using derivatives? Write a *Mathematica* notebook to solve this problem for the original three observations. - - absolute.nb

Problem

Assume that in the population that

(1) $y_i = \beta x_i + \varepsilon_i$ where $i = 1,\ldots,N$

and each ε_i is an independent draw from the density function

(2) $f(\varepsilon) = e^{-\varepsilon} \exp[e^{-\varepsilon}]$

$f(\varepsilon)$ is a simple Extreme value density function with mode=0, mean = .57721 and var $= \sigma_\varepsilon^2 = \pi^2 / 6$.

Assume a random sample with one observation (i=1) where $y_1 = 10$ and $x_1 = 2$.

a) Intuitively come up with an estimate of β.

Alternatively, the maximum likelihood estimate of β, β_{ml}, is that β that maximizes the probability of obtaining this sample.

Note that given equations (1) and (2), the probability of observing a particular y given a particular x is

(3) $f(y|x) = e^{-(y-\beta x)} \exp[e^{-(y-\beta x)}]$

which is the density function for an extreme value distribution with a mode of βx and a mean of $\beta x + .57721$.

The maximum likelihood estimate of β is the β that maximizes

(4) $L(\beta) = e^{-(10-\beta 2)} \exp[e^{-(10-\beta 2)}] = e^{(\beta 2-10)} \exp[e^{-(\beta 2-10)}]$

Equation (4) is a likelihood function.

b) If in 6808, use *Mathematica* to graph this likelihood function and to find the maximum likelihood estimate of β. Does the estimate surprise you? Is it a biased estimate? Demonstrate that you will get the same estimate if you maximize the ln of the likelihood function rather than the likelihood function - - maxlikev.nb .
c) Now find that same estimate using a derivative.

3. Maximizing or Minimizing Functions of N Variables

Let $f(x) = f(x_1, x_2, x_3, .., x_N)$ be a once continuously differentiable function defined over some set S. Then a necessary (but not sufficient) condition for a point $(x_1^0, x_2^0, .., x_N^0)$ belonging to the interior of s to be a local min or max

(1) $$\frac{\partial f(x^0)}{\partial x_1} \equiv \frac{\partial f(x_1^0, x_2^0, .., x_N^0)}{\partial x_1} = 0$$

(2) $$\frac{\partial f(x^0)}{\partial x_2} = 0$$

•

•

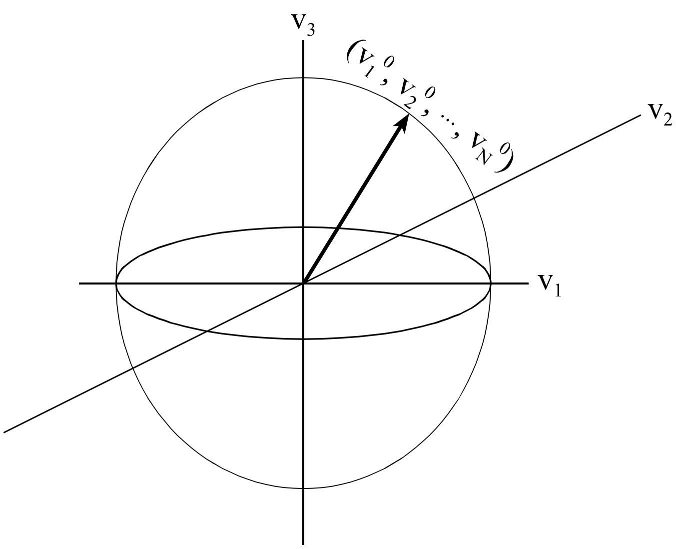
(N) $$\frac{\partial f(x^0)}{\partial x_N} = 0$$

i.e. we need the directional derivatives in the directions given by the coordinate axes to be zero

The second-order sufficient conditions for a local interior max or min can be developed in a manner analogous to the two variable case considered on pages 12 and 13. Just replace (x^0, y^0) by $(x_1^0, x_2^0, ..., x_N^0)$ and the same conditions can be used; however, the direction v is now $v \equiv (v_1, v_2, v_3, .., v_N)$ where $v_1^2 + v_2^2 + .. + v_N^2 = 1$.

When $n > 2$ the conditions are equivalent to the conditions for $n = 2$ but there are now n dimensions to check

Example: if n = 3, $v_1^2 + v_2^2 + v_3^3 = 1$ is the equation for a sphere of radius one



$v^0 \equiv (v_1^0, v_2^0, v_3^0)$ denotes a direction

v^0 is a point on the surface of this 3 dimensional sphere

Things are obviously even more difficult to visualize when n ≥ 4

The Sufficient conditions for a min ⇒ one must be at the bottom of a hill in every direction in n-dimensional space.

Problem Assume $f(x) = x_1^2 + x_2^2 + x_3^2 + .. + x_N^2$

a) Find a candidate for a local interior min of f(x)
b) Check whether your candidate $(x_1^0, x_2^0, .., x_N^0)$ is truly a local interior min

4. Maximizing or Minimizing functions of 2 Variables Subject to a Constraint

Economics is basically a problem of constrained optimization

resources are scarce and economic agents want to do as well as they can

society wants to max its welfare given its resource constraints and given the state of technical knowledge

consumers want to max their utility given their budget constraint and parametric prices

competitive firms attempt to max their profits given the state of technical knowledge and parametric input and output prices

for a given level of output and for a given state of technical knowledge, a firm attempts to minimize its costs

mathematically, the problem is one of maximizing or minimizing a function of N variables $f(x_1, x_2, .., x_N)$ w.r.t. $x_1, x_2, .., x_N$ belonging to some set S, where the x_i's are subject to one or more constraints.

The constraints are of the form

$$g(x_1, x_2, .., x_N) \geq 0 \text{ and } h(x_1, x_2, .., x_N) \geq 0$$

i.e. the x_i's must satisfy the above inequalities.

To keep things simple we will assume that there is only one constraint and it is an equality constraint.

$$g(x_1, x_2, .., x_N) = 0$$

How restrictive is it to assume $g(x_1, x_2, \ldots, x_n) = 0$ rather than ≥ 0?

Consider the 2 variable constrained minimization problem

$$\min_{\text{w.r.t. } x_1 \text{ and } x_2} f(x_1, x_2) \quad \text{subject to } g(x_1, x_2) = 0$$

It turns out that there are two methods that we can use to solve the above problem.

Method I: Direct Substitution

Use the constraint equation $g(x_1, x_2) = 0$ to solve for one of the variables in terms of the other; e.g. solve for x_1 as a function of x_2 and suppose this functional relation can be written as

$x_1 = h(x_2)$ NOTE: It is not always possible to solve $g(x_1, x_2) = 0$ for x_1.

Would direct substitution be applicable if the constraint was $g(x_1, x_2) \geq 0$? Note that the assumption that $g(x_1, x_2) \geq 0$ is the same constraint as $-g(x_1, x_2) \leq 0$

Now use this equation to eliminate x_1 from the objective function $f(x_1, x_2)$. We are left with the following unconstrained minimization problem

$$\min_{\substack{\text{w.r.t.} \\ x_2}} \; f(h(x_2), x_2) \equiv m(x_2)$$

and thus we can simply apply the tools developed for maximizing or minimizing an unconstrained function of one variable.

Example: $\min_{\text{w.r.t. } x_1, x_2} f(x_1, x_2)$ s.t. $g(x_1, x_2) = 0$

where $f(x_1, x_2) \equiv x_1^2 + x_2^2$ and $g(x_1, x_2) \equiv x_1 + x_2 - 1 = 0$

Solution: Use the equation $g(x_1, x_2) \equiv x_1 + x_2 - 1 = 0$ to solve for x_1 as a function of $x_2 : x_1 = 1 - x_2 \equiv h(x_2)$

Now substitute $h(x_2)$ into $f(x_1, x_2)$ and minimize w.r.t. x_2 :

$$
\begin{aligned}
f(h(x_2), x_2) \equiv m(x_2) &\equiv [h(x_2)]^2 + x_2^2 \\
&= (1 - x_2)^2 + x_2^2 \\
&= 1 - 2x_2 + x_2^2 \\
&= 2x_2^2 - 2x_2 + 1
\end{aligned}
$$

$$\frac{\partial m(x_2)}{\partial x_2} = 4x_2 - 2 + 0 \overset{\text{set}}{=} 0$$

or

$$4x_2 = 2$$

or

$$x_2^0 = 1/2 \quad \text{and} \quad x_1^0 = h(x_2^0) = 1 - x_2^0 = 1 - 1/2 = 1/2$$

So the point $(x_1^0, x_2^0) \equiv (½, ½)$ is a candidate for the constrained min.

Now check whether (½, ½) is in fact a local min of $f(x_1, x_2)$ s.t. $g(x_1, x_2) = 0$. To do this we need to check whether the function of one variable, $f[h(x_2), x_2] \equiv m(x_2)$ has a local min at $x_2^0 = ½$. How do we do this?

Check $\dfrac{d^2 m(x_2)}{dx_2^2} = \dfrac{d[4x_2 - 2]}{dx_2} = 4$

so $\dfrac{d^2 m(x_2)}{dx_2^2} = 4 > 0$ implies we have a local min. There is only one direction to check.

Since the equation $(4x_2 - 2 = 0)$ has only one solution $(x_2^0 = ½)$

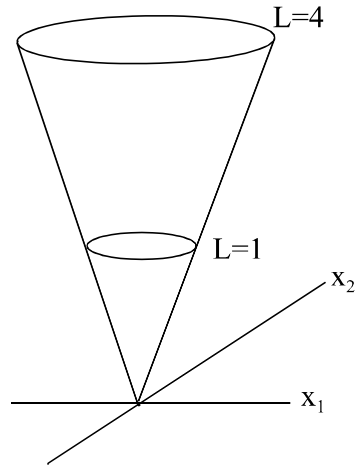
and since there is not a local min at a corner (no restrictions where placed on the domain of S)

we know that

$(x_1^0, x_2^0) \equiv (½, ½)$ is the global min of our constrained minimization problem and
$f(x_1^0, x_2^0) = (1/2)^2 + (1/2)^2 = 1/2$.

<u>Graphical Analysis of the Example</u>

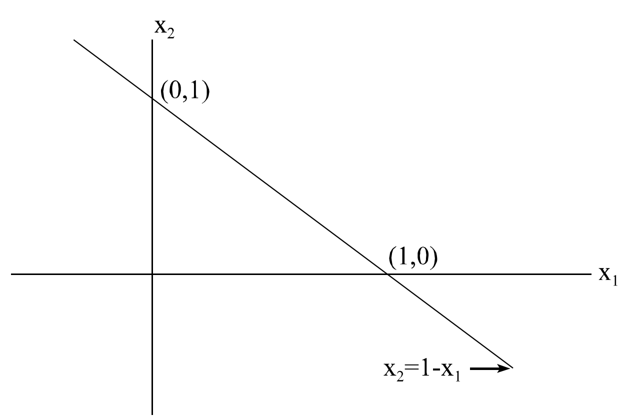
Graphically, the function $f(x_1, x_2) = x_1^2 + x_2^2$ is a 3-dimensional cone whose bottom is balancing on the x_1 x_2 plane at the point where $x_1 = 0$ and $x_2 = 0$



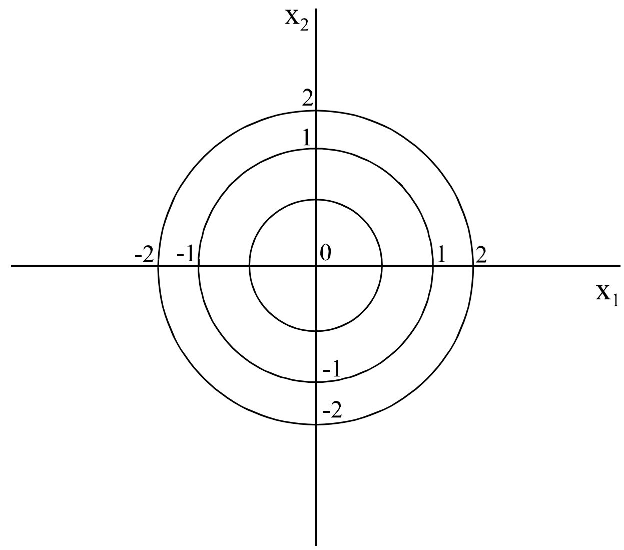
The constraint is the line defined by the equation $x_1 + x_2 - 1 = 0$.

I'm not going to try and draw it. One can visualize this plane as a straight line that sits on the x_1 x_2 plane.

Looking at the x_1 x_2 plane from above the constraint is the straight line.

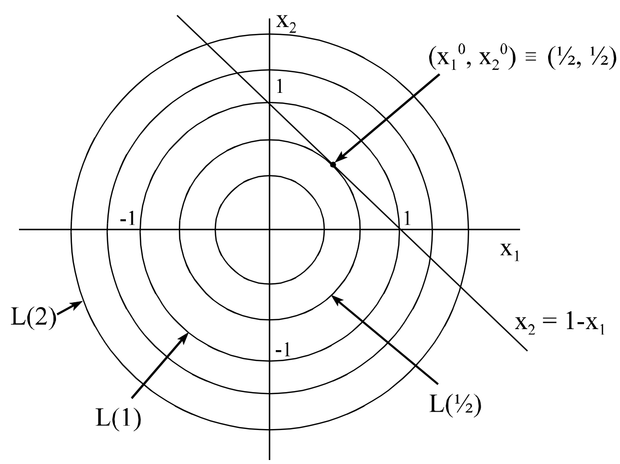


Looking at the cone from above (i.e. projecting the cone onto the 2-dimensional x_1 x_2 plane) one sees circles.



Our constrained minimization problem in 3 dimensions is one of finding the lowest point on the cone that also is a point on the constraint line.

Graphically, in terms of our projections onto the 2 dimensional x_1 x_2 plane, the constrained min problem is one of finding the point on the constraint line that lies on the smallest circle. i.e.



Notice that the solution is at the point of tangency between the constraint line and one of the circles.

These circles have a more technical name. All the points (x_1, x_2) for which $f(x_1, x_2) \leq k$ are the lower level set of $f(x_1, x_2) \leq k$

indicates it is the lower level set.

Graphically, in terms of our projection onto the 2 dimensional x_1 x_2 plane, $L_l(k)$ consists of all points that are in the circle of radius k. e.g. $L_l(1) \equiv \{(x_1, x_2): x_1^2 + x_2^2 \leq 1\}$. That is $L_l(1)$ is the set of x_1's and x_2's which lie within (including the boundary) a circle of radius 1, any such point has $f(x_1, x_2) \leq 1$. It can be visualized as the projection onto the x_1 x_2 plane of a two-dimensional cone cut parallel to the x_1x_2 plane, k units above that plane.

for future reference note that
$M(k) \equiv \{(x_1, x_1) : x_1^2 + x_2^2 = k\}$ is a subset of
$L_l(k) \equiv \{(x_1, x_2) : x_1^2 + x_2^2 \leq k\}$ and includes only the points on the boundary of the circle. Alternatively,
$N(k) \equiv \{(x_1, x_2) : x_1^2 + x_2^2 < k\}$ includes only the interior points of the circle. And $L_u(k)= \equiv \{(x_1, x_2): x_1^2 + x_2^2 \geq k\}$ is an *upper-level set*.

Our constrained min problem is to identify the point belonging to the set $J(1) \equiv \{(x_1, x_2) : x_2 + x_1 = 1\}$ that belongs to the lowest lower-level set $L_l(k)$. Again note that the optimum point $(x_1^0, x_2^0) \equiv (½, ½)$ is at the tangency between two sets J(1) and $L_l(½)$. i.e. just where the two sets touch.

Since the constraint line is tangent to M(½) at the optimal point, the slope of the constraint line must be equal to slope of the circle of radius ½ at this point. This is easily demonstrated using total
differentials, as we shall soon see.

If $u(x_1, x_2)$ is a utility function, what are

$$L_u(k) = \{(x_1, x_2):\ u(x_1, x_2) > k\}\ ?$$

and

$$M(k) = \{(x_1, x_2):\ u(x_1, x_2) = k\}\ ?$$

Now let's go on to method II.

Method II: The Lagrangian Method

for minimizing $f(x_1, x_2)$ s.t. $g(x_1, x_2) = 0$
w.r.t. x_1, x_2

Form the function $\mathcal{L}(x_1, x_2, \lambda) \equiv f(x_1, x_2) - \lambda\, g(x_1, x_2)$.

$\mathcal{L}$(☺) is called the Lagrangian function and λ is called the Lagrangian multiplier. Let's begin by asserting that the Lagrangian can be used to solve our constrained min problem. Then we will apply the Lagrangian technique to the previous example and show that it gives the same answer as the method of direct substitution.[1]

Finally, we will prove that the Lagrangian technique will, in general, identify the constrained min.

In order to obtain the first order necessary conditions for a constrained min (or max) partially differentiate the Lagrangian $\mathcal{L}(x_1, x_2, \lambda)$ w.r.t. x_1, x_2, and λ and set the resulting partial derivatives equal to zero:

$$\frac{\partial \mathcal{L}}{\partial x_1}(x_1, x_2, \lambda) = \frac{\partial f(x_1, x_2)}{\partial x_1} - \lambda \frac{\partial g(x_1, x_2)}{\partial x_1} \overset{\text{set}}{=} 0$$

$$\frac{\partial \mathcal{L}}{\partial x_2}(x_1, x_2, \lambda) = \frac{\partial f(x_1, x_2)}{\partial x_2} - \lambda \frac{\partial g(x_1, x_2)}{\partial x_2} \overset{\text{set}}{=} 0$$

$$\frac{\partial \mathcal{L}}{\partial \lambda}(x_1, x_2, \lambda) = -g(x_1, x_2) \overset{\text{set}}{=} 0$$

Solve the 3 equations for the 3 unknowns and call the solution x_1^0, x_2^0, λ^0
It turns out that for the previous example this x_1^0 and x_2^0 will be equal to the x_1^0, and x_2^0 that we would have obtained if we had used the method of direct substitution.

WOW!

[1] If it did not, we would know the Lagrangian technique is a technique that does not always give the correct answer.

Example: min $f(x_1, x_2)$ s.t. $g(x_1, x_2) = 0$
w.r.t.
x_1, x_2
where $f(x_1, x_2) = x_1^2 + x_2^2$ and $g(x_1, x_2) = x_1 + x_2 - 1 = 0$

Form the Lagrangian

$$\mathcal{L}(x_1, x_2, \lambda) \equiv f(x_1, x_2) - \lambda g(x_1, x_2)$$
$$= x_1^2 + x_2^2 - \lambda[x_1 + x_2 - 1]$$

Derive the first order necessary conditions by taking the partial derivatives and setting them equal to zero.

$$\frac{\partial \mathcal{L}}{\partial x_1}(x_1, x_2, \lambda) = 2x_1 - \lambda \overset{\text{set}}{=} 0$$

$$\frac{\partial \mathcal{L}}{\partial x_2}(x_1, x_2, \lambda) = 2x_2 - \lambda \overset{\text{set}}{=} 0$$

$$\frac{\partial \mathcal{L}}{\partial \lambda}(x_1, x_2, \lambda) = -[x_1 + x_2 - 1] \overset{\text{set}}{=} 0$$

If one solves these three equations for x_1, x_2 and x_3 one obtains $x_1^0 = ½$ $x_2^0 = ½$ $\lambda^0 = 1$.
We get the same answer we got with the method of direct substitution but with the additional piece of information that $\lambda^0 = 1$.

since $(x_1^0, x_2^0) \equiv (½, ½)$ was the solution before it must still be the solution.

So, the Lagrangian technique gave us the correct answer, at least for this example.

We don't know why the Lagrangian technique gave us the correct answer, and we don't know if will always work. Let's see if we can show that is will always work.

We want to first show that the first order conditions of the Lagrangian are necessary conditions for the solution to our constrained min problem.

<u>The proof will take the following form.</u>

We know from using method I that an interior solution to the constrained optimization problem

min $f(x_1, x_3)$ subject to $g(x_1, x_2) = 0$
w.r.t.
x_1 and x_2

is characterized by the following conditions:

a) at the solution point (x_1^0, x_2^0), $g(x_1^0, x_2^0) = 0$ (that is, the constraint is fulfilled)

b) the slope of the function defining the set $M(f(x_1^0, x_2^0)) \equiv \{(x_1, x_2): f(x_1, x_2) = f(x_1^0, x_2^0)\}$ evaluated at the point (x_1^0, x_2^0) equals the slope of the constraint function, $g(x_1, x_2) = 0$, evaluated at the point (x_1^0, x_2^0)

(that is, as we saw earlier with out geometric representation, the solution (x_1^0, x_2^0) is at the tangency of the level set of the objective function and the constraint set.)

Therefore, if the solution to the first order conditions from the Lagrangian is characterized by properties a) and b) we will have shown that the Lagrangian technique fulfills the necessary conditions for a min.

The logic of the proof proceeds as follows

1) we know that the solution to the constrained optimization problem is characterized by properties a) and b)

2) therefore, the solution to any other problem (e.g. the Lagrangian problem) will also be the solution to the constrained optimization problem if it is characterized by properties a) and b)

We therefore need to show that the solution to the Lagrangian problem has properties a) and b). To do this we first need to digress and introduce the concept of a total differential.

Definition: the total differential of the function $y = f(x) \equiv f(x_1, x_2, .., x_n)$ is

$$dy = \frac{\partial f(\cdot)}{\partial x_1} dx_1 + \frac{\partial f(\cdot)}{\partial x_2} dx_2 + \ldots + \frac{\partial f(\cdot)}{\partial x_n} dx_n$$

The total differential tells us how much the value of y changes when the value of one or more of the x_i's change. Note that if $dx_i = 0$ for all i, $i \neq j$, the equation for the total differential reduces to

$$dy = \frac{\partial f(\cdot)}{\partial x_j} dx_j$$

rearranging terms on obtains $\left.\frac{dy}{dx_j}\right|_{\substack{dx_i=0 \\ \forall i \neq j}} = \frac{\partial f(.)}{\partial x_j}$ This says that a partial derivative can be described at the ratio of two restricted total differentials.

Now let's use the concept of total differentials to show that the Lagrangian technique always works.

Define the solution to the Lagrangian problem as $(x_1^0, x_2^0, \lambda^0)$.

First we will use the concept of total differentials to derive the slope of the function defining the set

$M[f(x_1^0, x_1^0) \equiv \{(x_1, x_2) : f(x_1, x_2) = f(x_1^0, x_2^0)\}$

and then evaluate that slope at the point (x_1^0, x_2^0).

Then we will use the concept of total differentials to derive the slope of the constraint function evaluated at (x_1^0, x_2^0).

Finally we show that the Lagrangian technique guarantees that the slopes of the two functions are equal at this point, and that the constraint is fulfilled at this point.

First the set of points (x_1, x_2) that belong to the set $M\{f(x_1^0, x_2^0)]$ are also described by the function

$h(x_1, x_2) = f(x_1, x_2) - c = 0 \quad \text{where} \quad c \equiv f(x_1^0, x_2^0)$

We want the slope of this function $\left(\frac{dx_2}{dx_1}\right)$ evaluated at (x_1^0, x_2^0). Consider the total differential of h, dh, where $dh = f_1 dx_1 + f_2 dx_2$. Given that $h(x_1, x_2) = 0$ for those x's where the function is identified,
$dh = d0 = 0$, and $dh = f_1 dx_1 + f_2 dx_2 = 0$

Rearranging terms one obtains $\frac{dx_2}{dx_1} = -f_1 / f_2$ = the slope of the curve defined by the set $M[f(x_1^0, x_2^0)]$. Evaluating this slope at (x_1^0, x_2^0) one obtains

$$\left.\frac{dx_2}{dx_1}\right| = -\frac{\partial f(x_1^0, x_2^0)}{\partial x_1} \Bigg/ \frac{\partial f(x_1^0, x_2^0)}{\partial x_2} \equiv \text{ slope of the objective function eval at the critical}$$

pt

The constraint is $g(x_1, x_2) = 0$. By the same argument as above.

$dg = g_1\, dx_1 + g_2\, dx_2 = 0$

Rearranging terms one obtains

$$\frac{dx_2}{dx_1} = -g_1 / g_2 \equiv \text{ the slope of the constraint function } g(x_1, x_2) = 0$$

Evaluating this at (x_1^0, x_2^0) one obtains

$$\left.\frac{dx_2}{dx_1}\right| = -\frac{\partial g(x_1^0, x_2^0)}{\partial x_1} \Bigg/ \frac{\partial g(x_1^0, x_2^0)}{\partial x_2} \equiv \text{ slope of the constraint function eval at the critical}$$

pt

Therefore, the two functions h( ) and g( ) will be tangent at the point (x_1^0, x_2^0) if

$$\frac{\partial f(x_1^0,x_2^0)}{\partial x_1} \Bigg/ \frac{\partial f(x_1^0,x_2^0)}{\partial x_2} = \frac{\partial g(x_1^0,x_2^0)}{\partial x_1} \Bigg/ \frac{\partial g(x_1^0,x_2^0)}{\partial x_2}$$

this is property b) in mathematical form. Property a) in mathematical form is $g(x_1^0, x_2^0) = 0$. Together these are necessary conditions for the constrained min.

<u>Finally</u>

As defined earlier, the first order conditions for the Lagrangian problem are

1) $$\frac{\partial f(x_1^0,x_2^0)}{\partial x_1} - \lambda^0 \frac{\partial g(x_1^0,x_2^0)}{\partial x_1} = 0$$

2) $$\frac{\partial f(x_1^0,x_2^0)}{\partial x_2} - \lambda^0 \frac{\partial g(x_1^0,x_2^0)}{\partial x_2} = 0$$

3) $$-g(x_1^0,x_2^0) = 0$$

One can see that the solution to these equations (x_1^0, x_2^0, λ^0) guarantee that properties a) and b) are fulfilled. Equation 3) says that the constraint is fulfilled (property a). Assuming $\lambda^0 \neq 0$, one can rearrange equations 1) and 2) to obtain property b).

$$\frac{\partial f(x_1^0,x_2^0)}{\partial x_1} \Bigg/ \frac{\partial f(x_1^0,x_2^0)}{\partial x_2} = \frac{\partial g(x_1^0,x_2^0)}{\partial x_1} \Bigg/ \frac{\partial g(x_1^0,x_2^0)}{\partial x_2}$$

QED

Example: min $f(x_1, x_2)$ s.t. $g(x_1, x_2) = 0$
w.r.t.
x_1, x_2
where $f(x_1, x_2) \equiv x_1^2 + x_2^2$
and $g(x_1, x_2) \equiv x_1 + x_2 - 1$

We have already found that the optimal point is $(x_1^0, x_2^0) = (½, ½)$. Let us now check algebraically whether the two curves were tangent at this point.

the slope of the one curve is $$-\frac{\partial f(x_1^0,x_2^0)}{\partial x_1} \Bigg/ \frac{\partial f(x_1^0,x_2^0)}{\partial x_2} = -\frac{2x_1^0}{2x_2^0} = -2(1/2) \Big/ 2(1/2) = -1$$

the slope of the other curve is $$\frac{-\partial g(x_1^0,x_2^0)}{\partial x_1} \Bigg/ \frac{\partial g(x_1^0,x_2^0)}{\partial x_2} = -1 \Big/ 1 = -1$$

Thus, the slopes are equal as the diagram on page 25 indicates.

Constrained Maximization Using the Lagrangian Technique

Note that if we wish to solve the following constrained maximization problem:

$$\max_{x_1, x_2} f(x_1, x_2) \quad \text{s.t.} \quad g(x_1, x_2) = 0$$
w.r.t.
x_1, x_2

We do not have to redo the theory outlined above for a constrained min problem. Why?

The necessary conditions for a constrained min and max are the same because maximizing f(x) is the same as minimizing -f(x), and minimizing f(x) is the same as maximizing -f(x).

That is, the above maximization problem can be written as

$$\min \; - f(x_1, x_2) \quad \text{s.t.} \quad g(x_1, x_2) = 0$$
w.r.t.
x_1, x_2

Second Order Sufficient Conditions for the Solution to our constrained Min Problem

Up to this point we have only examined necessary conditions. Let's now try to get an intuitive feel for the sufficient conditions described in terms of directional derivatives of the objective function.

A point (x_1^0, x_2^0) locally minimizes a twice differentiable function $f(x_1, x_2)$ subject to the constraint $g(x_1, x_2) = 0$ if

(i) $D_v\, f(x_1^0, x_2^0) = 0$ if $g(x_1^0, x_1^0) = 0$
and
(ii) $D_v\, [D_v\, f(x_1^0, x_2^0)] > 0$ if $g(x_1^0, x_2^0) = 0$

These are the same conditions that we had for the unconstrained minimization problem (see page 13) except that we don't have to check in every direction. We only have to check whether $D_v\, f(x_1^0, x_2^0) = 0$ and $D_v\, [D_v\, f(x_1^0, x_2^0)] > 0$ in the directions consistent with the constraint set.

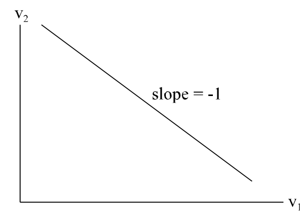
As in our previous example, where the constraint set is a straight line, there is only one direction to check. This can be visualized by examining the diagram on page 27. The point (x_1^0, x_2^0) is the constrained min if (x_1^0, x_2^0) is the point on the straight line $x_2 = 1 - x_1$ that results in the lowest value of y.

We know from page 17 that

$D_v f(x_1, x_2) = 2v_1x_1 + 2v_2x_2$ so $D_v f(½, ½) = v_1 + v_2 \neq 0$ unless $v_2 = -v_1$. That is, in most directions the directional derivative does not equal zero.

However, the directional derivative doesn't need to equal 0, except along the direction of the constraint. The constraint $g(x_1, x_2) = 0$ denotes a direction where

$g(x_1, x_2) = x_1 + x_2 - 1$ which implies $x_2 = 1 - x_1$



implies $\frac{dx_2}{dx_1} = -1$

i.e. $\frac{dv_2}{dv_1} = -1 \Rightarrow dv_2 = -dv_1$ along the constraint

so along the constraint

$v_2 = -v_1$ implies $D_v f(½, ½) = 0$ if $g(½, ½) = 0$

Therefore, at the point (x_1^0, x_2^0), $D_v f(x_1^0, x_2^0) = 0$ if $g(x_1^0, x_2^0) = 0$, and at the point (x_1^0, x_2^0), $D_v [D_v f(x_1^0, x_2^0)] = 2 > 0$ in all directions, so is positive in the direction of the constraints.

We are at the bottom of the hill on the constraint line - any movement away from the point x_1^0, x_2^0, along the constraint line, will increase the value of y.

To be sure we have a constrained minimum at (x_1^0, x_2^0), second-order sufficient conditions should be checked independent of whether the point (x_1^0, x_2^0) was identified using the method of direct substitution or the Lagrangian method.

At this point one might wonder why we bothered with the Lagrangian technique.

At first glance, the method of direct substitution appears more intuitive and easier to use, however

1) Sometimes it is difficult to solve the constraint $g(x_1, x_2) = 0$ for $x_2 = h(x_1)$ explicitly and thus we can not use Method I but we can use the Lagrangian technique.

and

2) The Lagrangian technique provides more information. The solution value of the Lagrangian multiplier (λ^0) tells us how much the value of the objective function, $f(x_1, x_2)$, evaluated at the point (x_1^0, x_2^0), will change if the constraint is relaxed slightly.

i.e. $$\lambda^0 = \frac{\partial f(x_1^0, x_2^0)}{\partial b}$$

where $g(x_1, x_2) = 0 \equiv b - m(x_1, x_2)$

λ^0 often has an interesting economic interpretation.

An example of point 2

Assume we have the Lagrangian function

$$\mathcal{L} = f(x_1, x_2) + \lambda[b - m(x_1, x_2)]$$

it can be shown that

$$\lambda^0 = \frac{\partial f(x_1^0, x_2^0)}{\partial b}$$

i.e. λ^0 tells us how sensitive the constrained min value of the objective function is to a change in the value of the constraint.

Often in economics the objective function has the dimension of value.

e.g. $f(x_1, x_2) = \text{cost}$

$\text{cost} = wL + rK$

and the problem is to min costs subject to a production function $y = g(K, L)$ and an output constraint $\bar{y}$.

The Lagrangian in this case is $\mathcal{L} = wL + rK + \lambda[\bar{y} - g(K, L)]$, therefore

$$\lambda^0 = \frac{\partial[wL^0 + rK^0]}{\partial \bar{y}}$$

it tells us how much min costs will rise if output is increased by a marginal amount from $\bar{y}$

This is fairly easy to show.

The 1st order conditions for the Lagrangian are

$$\frac{\partial \mathcal{L}}{\partial L} = \frac{\partial f}{\partial L} - \lambda \frac{\partial g}{\partial L} = 0, \qquad \frac{\partial \mathcal{L}}{\partial K} = \frac{\partial f}{\partial K} - \lambda \frac{\partial g}{\partial K} = 0, \quad \text{and} \quad \frac{\partial \mathcal{L}}{\partial \lambda} = \bar{y} - g(K,L) = 0$$

We have 3 equations and three unknowns.

Solve them to obtain

$$L^0 = L(\bar{y}, w, r),\ K^0 = K(\bar{y}, w, r), \text{ and } \lambda^0 = \lambda(\bar{y}, w, r)$$

Plug these back into the Lagrangian function to obtain

$$\mathcal{L}^0(\bar{y}, w, r) = w[L(\bar{y}, w, r)] + r[K(\bar{y}, w, r)] + \lambda(\bar{y}, w, r)\,[\bar{y} - g(L(\bar{y}, w, r), K(\bar{y}, w, r))]$$

We now have the Lagrangian, evaluated at L^0, K^0, λ^0, as a function of $\bar{y}$.
Differentiate it w.r.t. $\bar{y}$ to obtain

$$\frac{\partial \mathcal{L}^0(\bar{y},w,r)}{\partial \bar{y}} = w\frac{\partial L(\bar{y},w,r)}{\partial \bar{y}} + r\frac{\partial K(\bar{y},w,r)}{\partial \bar{y}} + \frac{\partial \lambda(\bar{y},w,r)}{\partial \bar{y}}\left[\bar{y} - g(L(\bar{y},w,r),K(\bar{y},w,r))\right]$$

$$+ \lambda(\bar{y},w,r)\left\{1 - \frac{\partial g(L(\bar{y},w,r),K(\bar{y},w,r))}{\partial L(\bar{y},w,r)}\frac{\partial L(\bar{y},w,r)}{\partial \bar{y}} \frac{-\partial g(L(\bar{y},w,r),K(\bar{y},w,r))}{\partial K(\bar{y},w,r)}\frac{\partial K(\bar{y},w,r)}{\partial \bar{y}}\right\}$$

noting that $L^0 = L(\bar{y}, w, r)$, $K^0 = K(\bar{y}, w, r)$ and $\lambda^0 = \lambda(\bar{y}, w, r)$ this can be written more simply

$$\frac{\partial \mathcal{L}^0(\bar{y},w,r)}{\partial \bar{y}} = w\frac{\partial L^0}{\partial \bar{y}} + r\frac{\partial K^0}{\partial \bar{y}} + \frac{\partial \lambda^0}{\partial \bar{y}}\left[\bar{y} - g(K^0,\ L^0)\right] +$$

$$\lambda^0\left\{1 - \frac{\partial g(k^0,\ L^0)}{\partial K^0}\frac{\partial K^0}{\partial \bar{y}} - \frac{\partial g(K^0,\ L^0)}{\partial L^0}\frac{\partial L^0}{\partial \bar{y}}\right\}$$

rearranging terms one obtains

$$\frac{\partial \mathcal{L}^0(\bar{y},w,r)}{\partial \bar{y}} = \frac{\partial L^0}{\partial \bar{y}}\left\{w - \lambda^0\frac{\partial g(K^0,L^0)}{\partial L^0}\right\} + \frac{\partial K^0}{\partial \bar{y}}\left\{r - \lambda^0\frac{\partial g(K^0,L^0)}{\partial k^0}\right\} + \frac{\partial \lambda^0}{\partial \bar{y}}\left\{\bar{y} - g(K^0,L^0)\right\}$$

but the first order conditions tell us everything in each set of brackets { } is zero so

$\dfrac{\partial \mathcal{L}^0(\bar{y},w,r)}{\partial \bar{y}} = \lambda^0$. We have, to this point, proven that λ^0 equals the derivative of the Lagrangian function w.r.t. y, evaluated at (L^0, K^0, and λ^0).

To finish the proof, we need to show that

$$\frac{\partial \mathcal{L}^0(\bar{y},w,r)}{\partial \bar{y}} = \frac{\partial\,[wL^0 + rK^0]}{\partial \bar{y}}$$

This is quite trivial. At the point (K^0, L^0, λ^0),

$$\bar{y} - g(K^0, L^0) = 0 \text{ so } \mathcal{L}^0(\bar{y}, w, r) \equiv wL^0 + rK^0 \equiv wL(\bar{y}, w, r) + rK(\bar{y}, w, r)$$

i.e. the two functions are identical so their derivatives w.r.t. to $\bar{y}$ must be identical. Therefore

$$\lambda^0 = \frac{\partial f(K^0,L^0)}{\partial \bar{y}} \equiv \frac{\partial\,[wL^0 + rK^0]}{\partial \bar{y}}$$

QED

Looking at the Lagrangian, one might be <u>tempted</u> to conclude that $\frac{\partial \mathcal{L}^0}{\partial \bar{y}} = \lambda^0$, because the partial of $\mathcal{L}$ w.r.t. $\bar{y}$ <u>appears</u> to be λ.
Why should this temptation be resisted?

Suppose a production manager is in charge of a plant which produces y by combining two inputs L and K. Suppose that the technology of the plant can be described with the production function $y = L^{1/2} K^{1/2}$.

Suppose further that the production manager can hire any number of units of L at the given wage rate $w > 0$ and she can rent any number of units of K at a given price $r > 0$.

The production manager's problem is to minimize the cost of producing a given output level $y > 0$ subject to the production function constraint

$$\min_{\substack{\text{w.r.t.} \\ L>0\ K>0}} \quad wL + rK \quad \text{s.t. } y = L^{1/2} K^{1/2}$$

a) Use the method of direct substitution to solve this problem. Hint: Replace the constraint $y = L^{1/2} K^{1/2}$ with the equivalent constraint $y^2 = LK$. The solution (L^0, K^0) are the conditional demand for L and K. These two demand functions will be functions of y, w, and r.

b) If $w = 2$, $r = 2$, and $y = 3$, calculate the resulting L^0 and K^0.

c) Go back to part a) and check that the second-order sufficient conditions for minimizing a function of one variable are satisfied.

5. Maximizing or Minimizing Functions of N Variables Subject to a Constraint

The basic constrained minimization problem is:

$$\min_{\text{w.r.t. } x_1, x_2, \ldots, x_N} \quad f(x_1, x_2, ..., x_N) \quad \text{s.t. } g(x_1, x_2, ..., x_N) = 0$$

There are two methods for solving this problem, just as in the two variable case.

Method I: Direct Substitution

Use the constraint equation $g(x_1, x_2, ... x_N)$ to solve for one variable is terms of the others; e.g.

$$x_1 = h(x_2, x_3, ..., x_N)$$

Now use this equation to eliminate x_1 from the objective function, and then solve the following unconstrained minimization problem w.r.t. the N-1 variables $x_2, x_3, ..., x_N$.

$$\min_{\text{w.r.t. } x_2, x_3, .., x_N} \quad f(h(x_2, x_3, ... x_N), x_2, ... x_N) \equiv b(x_2, x_3, ..., x_N)$$

the first order conditions are

$$\begin{aligned}
\frac{\partial b}{\partial x_2}(x_2^0, x_3^0, \ldots x_N^0) &\overset{\text{set}}{=} 0 \\
\frac{\partial b}{\partial x_3}(x_2^0, x_3^0, \ldots x_N^0) &\overset{\text{set}}{=} 0 \\
&\vdots \\
\frac{\partial b}{\partial x_N}(x_2^0, x_3^0, \ldots x_N^0) &\overset{\text{set}}{=} 0
\end{aligned}$$

there are $N-1$ equations and N-1 unknowns $x_2^0, ... x_N^0$, $x_1^0 = h(x_2^0, ... x_N^0)$

The point $(x_1^0, x_2^0, .., x_N^0)$ is a candidate as solution to our constrained minimization problem. It fulfills the necessary conditions for a constrained interior local min. However, one can not be sure that it is a constrained local min until the second-order conditions have been checked.

Method II: The Lagrangian Method

Define the Lagrangian

$$\mathcal{L}(x_1, x_2, ..., x_N, \lambda) \equiv f(x_1, x_2, ... x_N) - \lambda\, g(x_1, x_2, ..., x_N)$$

and partially differentiate $\mathcal{L}$ with respect to the x's and the λ. Set the resulting partials equal to zero.

$$\frac{\partial \mathcal{L}}{\partial x_1}(x_1^0, x_2^0, \ldots x_N^0) \overset{\text{set}}{=} 0$$

$$\frac{\partial \mathcal{L}}{\partial x_2}(x_1^0, x_2^0, \ldots x_N^0) \overset{\text{set}}{=} 0$$

$$\vdots$$

$$\frac{\partial \mathcal{L}}{\partial x_N}(x_1^0, x_2^0, \ldots x_N^0) \overset{\text{set}}{=} 0$$

$$\frac{\partial \mathcal{L}}{\partial \lambda}(x_1^0, x_2^0, \ldots x_N^0) \overset{\text{set}}{=} 0$$

there are N+1 equations and N+1 unknowns $x_1^0, x_2^0, \ldots x_N^0, \lambda^0$

<u>Geometrically, the constrained minimization problem</u>

$$\min_{\substack{\text{w.r.t.} \\ x_1 \ldots x_N}} f(x_1, x_2, \ldots, x_N) \quad \text{s.t.} \quad g(x_1, x_2, \ldots, x_N) = 0$$

<u>can be conceptualized in the following way.</u>

The constraint function $g(x_1, x_2, \ldots x_N) = 0$ identifies a set of points in N-dimensional space

$$Q(0) \equiv \{(x_1, x_2, \ldots, x_N) \text{ in } R_N : g(x_1, x_2, \ldots, x_N) = 0\}$$

The constraint is a stationary cloud floating in N-dimensional space. The objective function $y = f(x_1, x_2, \ldots, x_N)$ identifies a set of point in N+1 dimensional space.

The lower level set of $f(\cdot) = y$ is

$$\mathcal{L}_l(y) \equiv \{(x_1, x_2, \ldots x_N) \text{ in } R_N : f(x_1, x_2, \ldots x_N) \le y\}$$

Each lower level set is a set of points in N-dimensional space.

We are looking for that point $(x_1^0, x_2^0, \ldots x_N^0)$ that belongs to the N dimensional constraint set and the level set, $\mathcal{L}_l(y^0)$, such that there is no point that belongs to the constraint set and the set $\mathcal{L}_l(y^1)$ where $y^1 < y^0$. That is, we want that x vector consistent with the constraint that generates a y that is as low as any y generated by any x vector consistent with the constraint.

The point $(x_1^0, x_2^0, \ldots x_N^0)$, will, in general, be a point where the set Q(0) and the set $\mathcal{L}_l(y^0)$ touch in N dimensional space.

Often, but not always, the point $(x_1^0, x_2^0, \ldots, x_N^0)$ will be a point of tangency between the two sets. One can search for such a tangency point using either the method of direct substitution or the Lagrangian technique.

You are now prepared to search for constrained mins and maxs. However, given your equipment, the hunt will not always be successful. For example, you might miss a wild max or min lurking on the corner. Or you might capture what appears to be a wild min (max) only to later discover that it is a wild max (min). This could lead to danger (embarrassment) or both.

Be careful.

The following discussion of convex sets and curvature properties might help you to identify a max or min.

The End