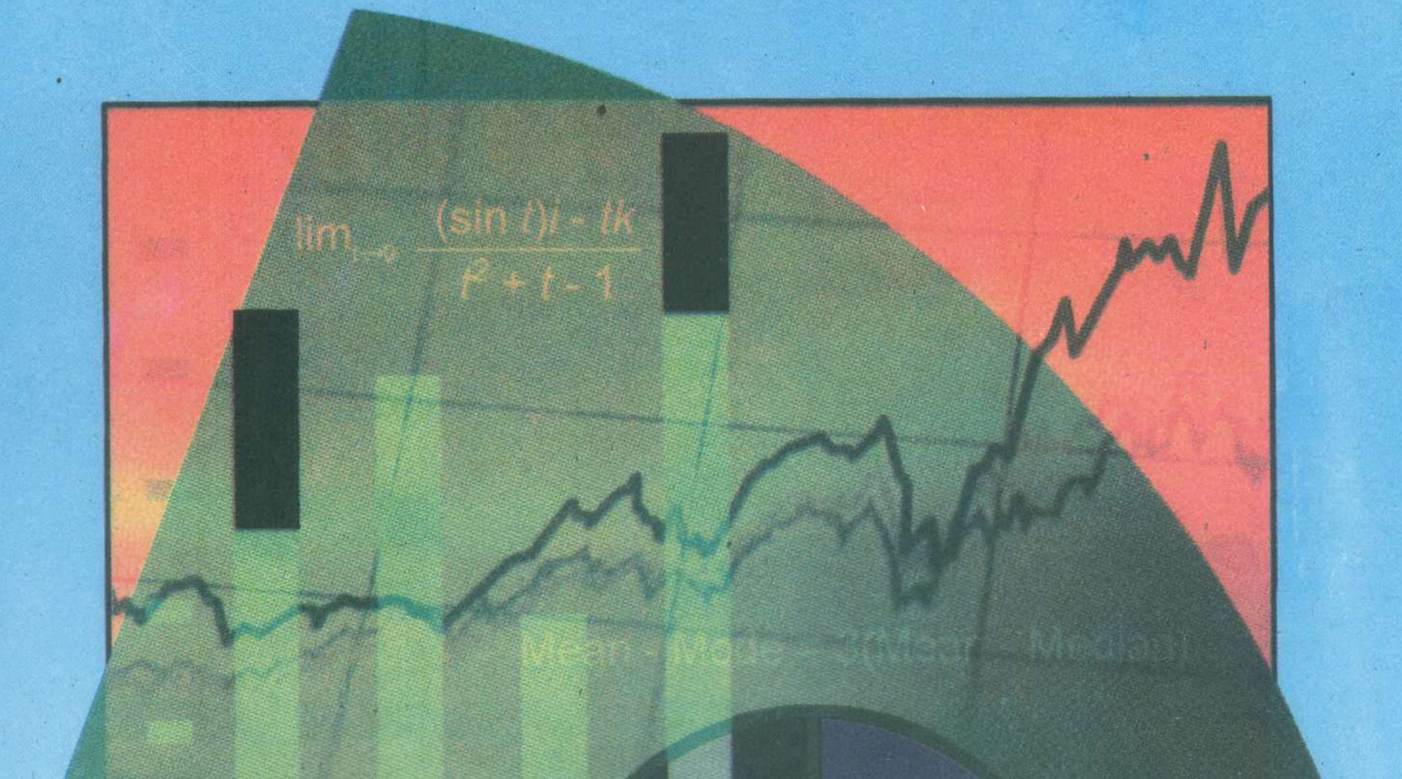
ignou
THE PEOPLE'S
UNIVERSITY

Indira Gandhi National Open University
School of Social Sciences

MEC-103
QUANTITATIVE METHODS



Extreme Values and Optimisation

2

"शिक्षा मानव को बन्धनों से मुक्त करती है और आज के युग में तो यह लोकतंत्रा की भावना का आधार भी है। जन्म तथा अन्य कारणों से उत्पन्न जाति एवं वर्गगत विषमताओं को दूर करते हुए मनुष्य को इन सबसे ऊपर उठाती है।"

- इन्दिरा गाँधी

"Education is a liberating force, and in our age it is also a democratising force, cutting across the barriers of caste and class, smoothing out inequalities imposed by birth and other circumstances."

- Indira Gandhi

ignou
THE PEOPLE'S
UNIVERSITY

Indira Gandhi
National Open University
School of Social Sciences

MEC-103
Quantitative Methods

Block

2

EXTREME VALUES AND OPTIMISATION

Expert Committee

Course Editor : **Prof. Gopinath Pradhan**

Programme Coordinators

Prof. Gopinath Pradhan, IGNOU, New Delhi

Prof. Kaustuva Barik, IGNOU, New Delhi

Course Coordinator

Prof. Gopinath Pradhan

Block Preparation Team

Unit	Writer
4	Sunando Basu
5 & 6	Rittwik Chatterjee

PRINT PRODUCTION TEAM

Mr. S. Burman
D.R.(P) MPDD-IGNOU,
Maidan Garhi, Delhi-110068

Mr. Tilak Raj
A.R.(P) MPDD-IGNOU,
Maidan Garhi Delhi-110068

Mr. Yashpal Sharma
S.O.(P) MPDD-IGNOU,
Maidan Garhi Delhi-110068

November 2018 (Reprint)

ISBN : 978-93-86375-13-1

Further information on the Indira Gandhi National Open University Courses may be obtained from the University's Office at Maidan Garhi, New Delhi-110 068 or visit University's http:// www.ignou.ac.in.

Printed and Published on behalf of the Indira Gandhi National Open University, New Delhi by Registrar, MPDD. IGNOU

Printed at: A-One Offset Printers, 5/34, Kirti Nagar Indl. Area, New Delhi-110015

BLOCK 2 EXTREME VALUES AND OPTIMISATION

This block covers the techniques used for finding optimal values for an objective function. For that purpose, it divides the themes into three units. Of these, the first two have many common areas of attaining maximum and minimum values of a function and would look repetitive. Therefore, it may be necessary to keep in mind the attempt of solving generalised optimisation problems from an initial coverage specific to maximum or minimum values. The last unit of the block dealing with constraint optimisation is somewhat technical and it would be better to get familiarise with linear programming approach while covering the theme.

Unit 4, the first theme of the block, explains the basic idea of arriving at a maximum or minimum point. It starts with single variable functions and covers multivariate cases. Unconstrained optimisation is discussed in **Unit 5,** which highlights the method of attaining the optimal values and explains in a more general framework the maximum and minimum positions covered in its preceding unit. The last theme of the block **Unit 6** is on constrained optimisation and deals with the techniques of finding solution to objective funcions when choice variables have constraints.

BLOCK 2 EXTREME VALUES AND OPTIMISATION

This block covers the techniques used for finding optimal values of multivariable functions. [illegible]

Unit 4, the first of the block, [illegible]

UNIT 4 MAXIMA AND MINIMA

Structure

4.0 OBJECTIVES

After going through this unit, you will be able to:

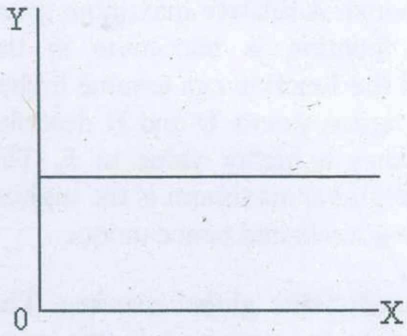
- understand the identification process of maximum and minimum points;
- prove the necessary conditions for maximum and minimum for functions;
- distinguish between a global and local maximum minimum; and
- characterise an extremum and distinguish it from a point of inflexion (or infection).

4.1 INTRODUCTION

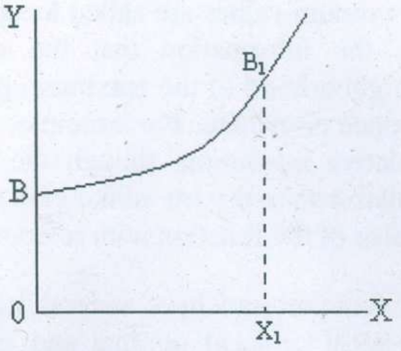
To begin with, we will introduce the notion of a stationary point. The points, at which first order derivatives are zero, are called stationary points. At these points, the function comes to a standstill momentarily. The values of the function at those points are called stationary values.

4.2 CONCEPT OF MAXIMA AND MINIMA

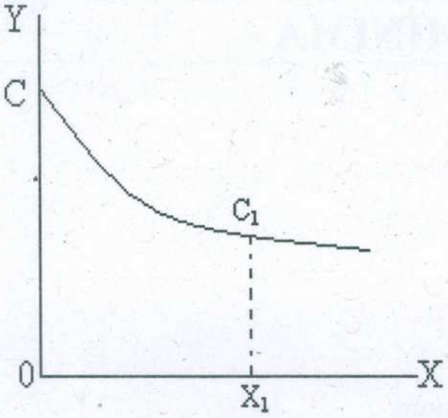
To start with, we assume the function $y = f(x)$ is differentiable (and hence continuous) throughout its range. The function may be represented by one of the following types of graphs:



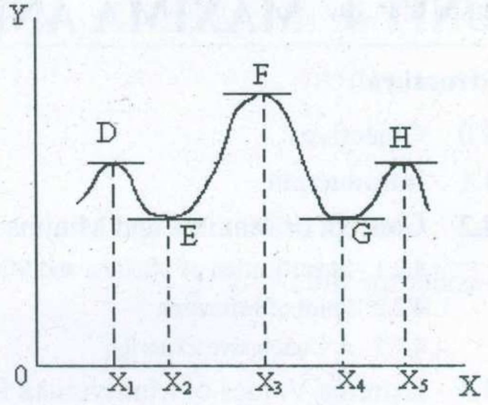
i) Constant



ii) Monotonically increasing



iii) Monotonically decreasing



iv) Fluctuating

In the first case, the function is the same for all values of x. We shall consider the values of y as both maximum and minimum value.

In the second case, the value of the function is monotonically increasing. Let us restrict the domain of x to the interval $0 \le x \le x_1$. Then in this closed interval, y can assume a maximum of B_1x_1 and a minimum of OB. These maximum and minimum values of the function are called absolute maximum (global) and absolute (global) minimum respectively. Such maximum and minimum values of a function are called extreme values or extrema. Accordingly in the third case, OC becomes the global maxima and X_1C_1 the global minima if we restrict in the interval of x to $0 \le x \le x_1$.

Let us consider the fourth case. Here y oscillates with changes in x. The function generates three peaks, D, F and H and two bottoms E and G. The value of the function at D is the highest in comparison to the values at other points in its the immediate neighborhood. Symbolically,

$f(x_1 - \in) < f(x_1)$ and $f(x_1 + \in) < f(x_1)$ as $\in \to 0$.

Thus, we can say that the value of the function at D is maxima, at least in some small interval $x_1 - \in < x < x + \in$. Similarly, in another small interval $x_3 - \in < x < x_3 + \in$, the value of the function is maximum at x = x_3 and in the neighborhood of F,

$f(x_3 - \in) < f(x_3)$ and $f(x_3 + \in) < f(x_3)$.

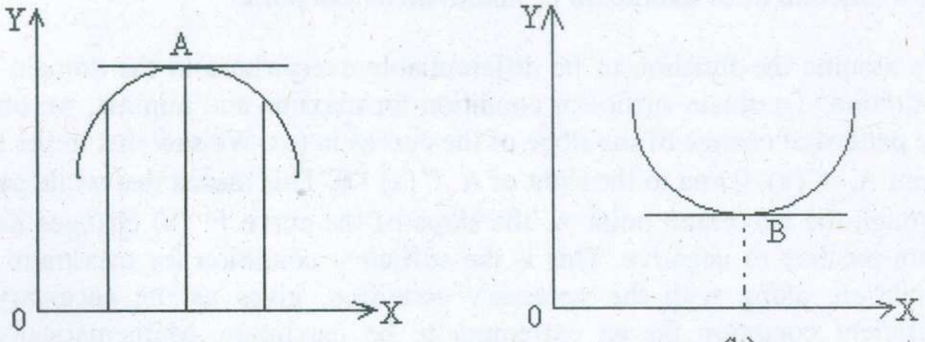
The function attains another maximum in the neighborhood of point F. These maximum values are called local (relative) maxima. A relative maximum gives us the information that the value of the function is maximum in the neighborhood of the maximum point only, and the function can assume higher values elsewhere. For example, in the fourth figure, points D and H describe relative maximum, though the function assumes a higher value at F. The relative maxima are unlike global maxima. The global maximum is the highest value of the function with reference to its entire domain and hence unique.

A function can have several local maxima and unique global maxima. The concept of local minima and global minima can be explained in the same manner. For a continuous function, the global maximum must always be greater than the global minimum.

The local extrema always appear as point interior to an interval, however small. These local extrema are thus strictly interior extrema. In the following discussion, we shall consider only these types of extrema.

4.2.1 Identification of Maxima and Minima

Consider the following figure:



(a) (b)

The extremum described by (a) is a maximum, while that of (b) is a minimum. In Figure (a) the curve y = f(x) takes a turn at point A. To the left of A, the curve is increasing i.e., $f'(x)>0$ while to the its right f(x) is decreasing i.e., $f'(x) <0$. Thus, at the turning point A, $f'(x)$ must be zero i.e., $f'(x_0) =0$. Similarly, at point B in Figure (b), $f'(x_0)=0$. Therefore, we can conclude that an extremum can take place only at the stationary points where $f'(x)=0$. The extreme values are always stationary values.

Let us construct ε - neighborhood of the point x_0 in Figure (a). As A is local maximum, we know that in the interval $x_0 - \varepsilon < x < x_0 + \varepsilon$, the value of f(x) at $x = x_0$ is the highest.

Thus, to the left of x_0,

$$x < x_0 \rightarrow f(x) < f(x_0) \rightarrow \frac{f(x)-f(x_0)}{x-x_0}$$

$$> 0 \rightarrow \lim_{x \to x_0^-} \frac{f(x)-f(x_0)}{x-x_0} > 0.$$

Let the value of this limit be $1 \geq 0$. Then

$$\lim_{x \to x_0^-} \frac{f(x)-f(x_0)}{x-x_0} > 1.$$

To the right of x_1,

$$x > x_0 \rightarrow f(x) < f(x_0) \rightarrow \frac{f(x)-f(x_0)}{x-x_0} < 0$$

$$\rightarrow \lim_{x \to x_0^+} \frac{f(x)-f(x_0)}{x-x_0} < 0 = -l_2(say) \text{ where, } l_2 \geq 0$$

Thus, we get the result that the right hand limit and left hand limit of x_0 assume two different values. Naturally, the limit does not exist. But the curve is smooth and continuous. So the, limit must exist at x_0 also. This can happen, if

$$\lim_{x \to x_0^-} \frac{f(x) - f(x_0)}{x - x_0} = \lim_{x \to x+} \frac{f(x) - f(x_0)}{x - x_0}$$

$l_1 = - l_2$ i.e. $l_1 + l_2 = 0$. i.e., if $l_1 = l_2 = 0$.

The value of the first derivative at that point is zero is the necessary condition for a function to be maximum or minimum at that point.

We assume the function to be differentiable everywhere in the domain of its definition. To obtain sufficient condition for maxima and minima, we observe the pattern of change of the slope of the curves in (a). We saw that to the left of point A, $f'(x)>0$ and to the right of A, $f'(x)<0$. This means that while passing through the maximum point A, the slope of the curve $f'(x)$ changes its sign from positive to negative. This is the sufficient condition for maximum. This condition, along with the necessary condition, gives us the necessary and sufficient condition for an extremum to be maximum. Mathematically, the sufficient condition means $f''(x_0)=0$ and above two definitions are not equivalent. But we shall assume $f''(x_0)$ assumes a non-zero value.

Let us now summarise the result we have obtained so far.

i) Necessary condition

a) for maxima: $f''(x)=0$

b) for minima: $f'(x)=0$

ii) Sufficient condition

a) for maxima: $f'(x)=0$

$f''(x)<0$

b) for minima: $f'(x)=0$

$f''(x)>0$.

Example: i) Find the maxima and minima for the following function:

$y = 3x^4 - 10x^3 + 6x^2 + 5$

Solution:

First order condition (F.O.C.):

$12x^3 - 30x^2 + 12x = 0$

or, $3x(4x - 2)(x - 2) = 0$.

Either, $x = 0$ or, $x = 2$ or $x = ½$.

Second, order condition (S.O.C.):

At $x = 0$, $f''(x) = 12>0$.

At $x = 2$, $f''(x) = - 9 < 0$.

Hence the function attains maximum at $x = \frac{1}{2}$ and minimum at x=0 and x = 2.

iii) Show that for the function $y = x + \frac{1}{x}$, maximum value is less than its minimum value.

$f'(x) = 1 - \frac{1}{x^2}$, $f''(x) = 1 - \frac{2}{x^3}$

$f'(x) = 0 \rightarrow x = 1$ or -1

$f''(1) = 2 > 0 \rightarrow f(x)$ has its minima at x = 1 and f (1) = 2

$f''(-1) = -2 < 0 \rightarrow f(x)$ has its minima at x = -1 and f (-1) = -2

Thus, the maximum value is less than its minimum value.

Check Your Progress 1

1) Show that f(x) = $x^3 - 6x^2 + 24x + 4$ has neither a maximum nor a minimum.

....

....

....

....

2) Examine whether $x^{1/x}$ possesses a maximum or a minimum value.

....

....

....

....

3) Find the maxima and minima of 1 + 2 Sin x + 3 Cos^2 x. $\left(0 \le x \le \frac{\pi}{2}\right)$.

....

....

....

....

4.2.2 Points of Inflextion

Now we will consider the cases for which the second derivative is zero. If the second derivative is zero at a point, the function may have a maximum, a minimum or a point of inflextion. The point of inflextion is defined as a point at which a curve changes its curvature.

At the point of inflextion, a curve does not change its direction of movement but only the nature of its slope. At each inflextion point, the first order derivative of the curve either reaches a maximum or a minimum. The sufficient condition for a point of inflextion is

$f''(x) = 0$ and $f'''(x) \neq 0$.

Thus, if a function has $f'(x)=0, f''(x)=0$ and $f'''(x)\neq 0$ at the point x, the point is said to be stationary and inflextional.

At this stage, we state an important result:

If a monotonic function is first convex and then concave, the point of inflextion generates a maximum $f'(x)$. Similarly, if the function is first concave and then convex, the point of inflextion generates a minimum $f'(x)$.

Example:

Find the point of inflextion for the function f(x) $=x^3+2$

Here, $f'(x)=3x^2, f''(x)=6x$ and $f'''(x)=6\neq 0$

When $f''(x)=6x=0$, x=0. Thus, the function has point of inflextion at the origin.

Check Your Progress 2

1) Find the point of inflextion for the function $f(x)=x^3+x^2+x+1$

....

....

....

....

....

4.2.3 A Conclusive Criterion

Let us expand the function f(x) around the point $x=a$.

$$f(x)-f(a)=f'(a)(x-a)+\frac{f''(a)}{2!}(x-a)^2+......+\frac{f^n(x^*)}{n!}(x-a)^n$$

where, $x^*=a+\theta(x-a), 0\leq\theta\leq 1$.

The above series is the Taylor series with Lagrange's form of the remainder.

Form this result we will consider various possible cases:

Case-I: $f'(a)\neq 0$

For simplicity, we assume $n=0$. By Taylor's expansion,

$f(x)-f(a)=f'(x^*)(x-a)$.

Since $f'(a)\neq 0, f'(x^*)\neq 0$. To the left of x*, $(x-a)<0$ and to it right, $(x-a)>0$. Thus to the left and right of $x=a$, $f(x)-f(a)$ has opposite signs. This implies the extrema of a function is characterized by $f'(x)=0$.

Case-II: $f'(a)=0, f''(a)\neq 0$

Suppose n = 1. From the Taylor series,

$$f(x)-f(a)=f'(a)(x-a)+\frac{f''(x^*)}{2!}(x-a)^2=\frac{1}{2}f''(x^*)(x-a)^2$$

Now, $(x-a)^2 > 0 \rightarrow$ sign of $f(x)-f(a)$ depends on the sign of $f''(x^*)$ which again depends on the sign of $f''(a)$.

Thus if i) $f''(a)>0$, $f(x)-f(a)>0$ and the curve is minimum at $x=a$

ii) $f''(a)<0$, $f(x)-f(a)<0$ and the curve is maximum at $x=a$

Case-III:

$f'(a)=0, f''(a)=0$

and

$f'''(a)\neq 0$

Assume n=2.

Here from the Taylor expansion,

$$f(x)-f(a)=f'(a)(x-a)+\frac{f''(a)}{2!}(x-a)^2+\frac{f'''(x^*)}{3!}(x-a)^3=$$
$$\frac{1}{6}f'''(x^*)(x-a)^3$$

Here again $f'''(x^*)\neq 0$ and depends on the sign of $f'''(a)$. $(x-a)^3>0$ to the right of a and <0 to the left of a. Thus, the whole right hand expression has opposite signs at the two opposite sides of a. Thus a is not an extrema rather a point of inflextion.

Case-IV:

$f'(a)=0, f''(a)=0, f'''(a)=0$

and

$f^4(a)\neq 0$

Here assume n=4. Again by Taylor expansion,

$$f(x)-f(a)=\frac{1}{24}f^4(x^*)(x-a)^4.$$

Since $(x-a)^4>0$, the sign of $f(x)-f(a)$ depends on the sign of $f^4(x^*)$ which in turn depends on the sign of $f^4(a)$.

We can conclude that

i) $f^4(a)>0$, $f(x)-f(a)>0$ and the curve is minimum at $x=a$

ii) $f^4(a)<0$, $f(x)-f(a)<0$ and the curve is maximum at $x=a$

Generalizing the above argument we may conclude,

when $f'(a)=0, f''(a)=0 .... f^{(n-1)}(a)=0$ & $f^n(a)\neq 0$ the point $x=a$ has

i) a maximum if n is even and $f^n(a)<0$
ii) a minimum if n is even and $f^n(a)>0$
iii) an inflextion if n is odd.

4.3 EXTREME VALUES OF MULTIVARIATE FUNCTIONS

Let $y = f(x)$ and the first order condition for optimization is $f'(x)=0$.

Now, $dy = f'(x)$.

Thus, F.O.C. for optimization requires $dy = 0$.

To find out the extrema of bivariate function, we proceed according to the same argument.

Consider the bivariate function $y = f(x_1, x_2)$.

$dy = f_1\,dx_1 + f_2\,dx_2$.

The first order condition for optimization is

$dy = f_1dx_1 + f_2\,dx_2 = 0$

As both dx_1 and dx_2 are not equal to zero, the above equation can hold only when $f_1 = f_2 = 0$.

Hence, F.O.C. for optimum of a bivariate function are

$f_1=0$ and $f_2=0$.

4.3.1 Sufficient Condition for Extreme Values

A function $y = f(x_1, x_2)$ is said to be at maximum at a point (a, b) if $f(a, b)>f(x_1, x_2)$ for all values of (x_1, x_2) in the domain $a-\delta < x_1 < a+\delta$ and $b-\delta < x_2 < b+\delta$. When $x_1 \neq a,\ x_2 \neq b$ and δ is a very small quantity, i.e.,

$f(a + h, b + k) - f(a, b) < 0$

where, $|h| < \delta$ and $|k| < \delta$.

This implies, the function $y = f(x_1, x_2)$ is said to have a maximum at the point (a, b) if any movement from the point in any direction gives us a lower value of y.

If $dy = 0$ at the point of maximum and $dy<0$ else where, we can conclude that the sufficient condition to have a maximum is that the change in dy should be diminishing i.e.,

$d(dy) < 0 \rightarrow d^2y < 0$.

From $y = f(x_1, x_2)$ we have

$dy = f_1dx_1 + f_2\,dx_2 = 0$

or, $d^2y = d(f_1dx_1 + f_2\,dx_2)$

i.e., $d^2y = d(f_1dx_1) + d(f_2\,dx_2)$

$= df_1(dx_1) + df_2(dx_2)$

$= (f_{11}\,dx_1 + f_{12}\,dx_2)\,dx_1 + (f_{21}\,dx_1 + f_{22}\,dx_2)\,dx_2$

By Young's theorem, $f_{12} = f_{21}$.

Thus, $d^2y < 0$ implies

$f_{11}\,(dx_1)^2 + 2f_{12}\,dx_1\,dx_2 + f_{22}\,(dx_2)^2<0$ --------------(i)

Note that f_{11}, f_{12} and f_{22} are derived on any particular point, say (a,b). Thus, their values can be assumed to be constant. However, dx_1 and dx_2 are variables.

Let, $u = dx_1$, $v = dx_2$, $a = f_{11}$, $b = f_{22}$, $h = f_{12}$.

The left hand expression of (i) can be written in a quadratic form as

$$q = au^2 + 2huv + bv^2$$

$$= a\left(u^2 + \frac{2h}{a}uv + \frac{h^2}{a^2}v^2\right) + \left(b - \frac{h^2}{a}\right)v^2$$

$$= a\left(u + \frac{h}{a}v\right)^2 + \frac{ab-h^2}{a}(v^2).$$

Since both $\left(u + \frac{h}{a}v\right)^2$ and v^2 are positive, sign of q depends entirely on the values of the coefficients a, b and h. The inequality $q<0$ holds if $a < 0$ and $ab - h^2 > 0$.

Again, since the product a b must be positive in order to get positive $(ab-h^2)$, both a and b must bear the same sign. Thus, if $a<0$, $b<0$ and $ab-h^2>0$, then $q<0$.

Alternatively, $d^2y<0$, if $f_{11} < 0$, $f_{22} < 0$ and $f_{11}\ f_{22} - f_{12}^2 > 0$. This is the sufficient condition for a maximum.

The sufficient condition for a minimum of a bivariate function is

$f_{11} > 0$, $f_{22} > 0$, $f_{11}\ f_{22} - f_{12}^2 > 0$ which can be derived in a similar way.

Let us now consider the case when $f_{11}\ f_{22} < (f_{12})^2$. A stationary point of a bivariate function, where $f_1 = 0 = f_2$ is called

i) A saddle point if

   $f_{11}\ f_{22} < f_{12}^2$ and f_{11} and f_{22} have different signs.

ii) An inflexion point if $f_{11}\ f_{22} < f_{12}^2$ and f_{11}, f_{22} have the same sign.

   If $f_{11}\ f_{22} = f_{12}^2$, the change becomes inconclusive.

Check Your Progress 3

1) Find the stationary values and test whether they are maximum or minimum for

   $z = 3x^2 + 6xy + 7y^2$

   ……………………………………………………………………………………

   ……………………………………………………………………………………

   ……………………………………………………………………………………

   ……………………………………………………………………………………

   ……………………………………………………………………………………

2) Show that the only stationary point of the function $z = 4x^2 - 2y^2 + 7xy$ is a saddle point.

   ……………………………………………………………………………………

....

....

....

....

....

3) If $y = x_1^3 + 2x_2^3 - 6x_1^2 + 9x_2^2 - 63x_1 - 60x_2 - 24$

find the optimum points and verify the nature for the function.

....

....

....

....

4) Examine $y = -x_1^3 + 9x_1 - 4x_2^2$ for relative extrema.

....

....

....

....

4.3.2 The Case of Extrema with More than Two Variables

First, consider the case of two variables and take the expression

$q = au^2 + 2huv + bv^2$.

$= au^2 + huv + huv + bv^2$.

This may be put in the following matrix form:

$$q = [v] \begin{bmatrix} a & h \\ h & b \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix}$$

Since we are not concerned with the value of x and v, matrix of order 2x2 i.e., $\begin{bmatrix} a & h \\ h & b \end{bmatrix}$ supplies us the required conditions for q to be positive or negative.

i) q will have positive sign

if $|a| > 0$ and $\begin{vmatrix} a & h \\ h & b \end{vmatrix} > 0$

ii) q will have negative sign

if $|a| < 0$ and $\begin{vmatrix} a & h \\ h & b \end{vmatrix} > 0$

Next, consider the position of $|a|$ and $\begin{vmatrix} a & h \\ h & b \end{vmatrix}$.

The determinant $|a|$ is the sub-determinant of H that consists of the first element of the principal diagonal and is called the first principal minor of H.

The determinant $\begin{vmatrix} a & h \\ h & b \end{vmatrix}$ can also be considered as sub-determinant of H. Since it involves the first and second elements on the principal diagonal of H, it is called the second principal minor of H. In the case of function of two variables, only two principal minors are available but there will be three if the function involves three variables.

4.3.3 Function of Three Variables and Extrema

Assume, $z = f(x, y, w)$.

Then $dz = f_x dx + f_y dy + f_w dw$.

$d^2z = f_{xx}dx^2 + f_{xy}dxdy + f_{xw}dxdw + f_{yy}(dy)^2 + f_{yw}dydw + f_{wx}dwdx + f_{wy}dwdy + f_{yx}dydx + f_{ww}(dw)^2$.

$$= \begin{bmatrix} dx & dy & dw \end{bmatrix} \begin{bmatrix} fxx & fxy & fxw \\ fyx & fyy & fyw \\ fwx & fwy & fww \end{bmatrix} \begin{bmatrix} dx \\ dy \\ dw \end{bmatrix}$$

The determinant of the matrix $\begin{bmatrix} fxx & fxy & fxw \\ fyx & fyy & fyw \\ fwx & fwy & fww \end{bmatrix}$

shall have three principal minors namely,

$D_1 = |fxx|$, $|D_2| = \begin{vmatrix} fxx & fxy \\ fyx & fyy \end{vmatrix}$,

and $|D_3| = \begin{vmatrix} fxx & fxy & fxw \\ fyx & fyy & fyw \\ fwx & fwy & fww \end{vmatrix}$

Our problem is to determine what restrictions should be placed upon D_1, D_2 and D_3, so that, for any value of dx and dy, the value of d^2z will turn out to be positive or negative.

For this, we once again put the quadratic form of three variables in a way that the three variables always appear as components of some square.

Let $q = ax^2 + by^2 + cw^2 + 2hxy + 2fyx + 2gwx$ be

$d^2z = f_{xx}(dx^2) + f_{yy}(dy^2) + f_{ww}(dw)^2 + 2f_{xy}dxdy + 2f_{yw}dydw + 2f_{xw}dxdw$.

We can rewrite q in the form

$$q = a\left(x + \frac{h}{a}y + \frac{q}{a}w\right)^2 + \frac{ab - h^2}{a}\left(y + \frac{af - gh}{ab - h^2}\right)^2 + \left(\frac{abc - af_2 - bg^2 - ch^2 + 2fgh}{ab - h^2}\right)w^2$$

So that

i) for q to be positive, the following must hold:

$a > 0$, $ab - h^2 > 0$, and

$abc - af^2 - bg^2 - ch^2 + 2fgh > 0.$

The above three conditions can be put as the three principal minors of the matrix

$$\begin{bmatrix} a & h & g \\ h & b & f \\ g & f & c \end{bmatrix}$$

$D_1>0$, $D_2>0$ and $D_3>0$.

Translating these elements in terms of partial derivatives, we have the following three conditions for d^2z to be positive for the given function to have minimum value.

$$D_1 = |fxx| > 0, \; D_2 = \begin{vmatrix} fxx & fxy \\ fyx & fyy \end{vmatrix} > 0,$$

$$D_3 = \begin{vmatrix} fxx & fxy & fxw \\ fyx & fyy & fyw \\ fwx & fwy & fww \end{vmatrix} > 0.$$

ii) for q to be negative, following conditions must hold:

$$D_1 = a < 0, \; D_2 = \begin{vmatrix} a & h \\ h & b \end{vmatrix} > 0,$$

$$D_3 = \begin{vmatrix} a & h & g \\ h & b & f \\ g & f & c \end{vmatrix} < 0.$$

or, in terms of partial derivatives for d^2z to be negative (or the given function to have maximum value), following conditions must hold good:

$$D_1 = |fxx| < 0, \; D_2 = \begin{vmatrix} fxx & fxy \\ fyx & fyy \end{vmatrix} > 0,$$

$$D_3 = \begin{vmatrix} fxx & fxy & fxw \\ fyx & fyy & fyw \\ fwx & fwy & fww \end{vmatrix} < 0.$$

i.e., the principal minors of the determinant of d^2z should be successively negative, positive and negative.

4.3.4 Function of n Variables and Extrema

When there're n in dependent variables, the objective function may be expressed as

$z = f(x_1, x_2, ... x_n)$.

The total differential will then be.

$dz = f_1\, dx_1 + \text{---} + f_n dx_n$

so that F.O.C. for minimum or maximum will be that all the n first order partial derivatives must vanish (since dz = 0) and therefore,

$f_1 = f_2 ---- = f_n = 0.$

For the second order condition

i) d^2z will have positive sign (i.e., function will have the minimum value) if all the n principal minors of the Hessian determinant:

$$|H| = \begin{vmatrix} f_{11} & -- & f_{1n} \\ f_{21} & f_{22} & f_{2n} \\ -- & -- & -- \\ f_{n1} & f_{n2} & f_{nn} \end{vmatrix} \text{ are positive.}$$

ii) d^2z will have negative sign (i.e., function will have maximum value, if n principal minors of H posses alternate signs - the first being negative.

Let us now summarize the extremum conditions in the tabulated form:

For $z = f(x_1, x_2, --, x_n)$ to be

Condition	Maximum	Minimum
F.O.C.	dz = 0 i.e., $f_1 = f_2 = .. = f_n = 0$	dz = 0 i.e., $f_1=f_2 = -- =f_n = 0$
S.O.C	$d^2z < 0$, i.e., $\|H_1\| < 0$, $\|H_2\| > 0$, $\|H_3\| > 0$, --	$d^2z>0$ i.e., $\|H_1\| > 0$ $\|H_2\|>0$, $\|H_3\| >0$, ---

Example:

i) Indicate whether the function

$Z = f(x_1, x_2, x_3) = x_1^2 - 3x_1 x_2 + 3x_2^2 + 4x_2 x_3 + 6x_3^2$ has maximum or minimum value.

F.O.C.

$f_1 = 2x_1 - 3x_2 = 0$

$f_2 = -3x_1 + 6x_2 + 4x_3 = 0.$

$f_3 = 4x_2 + 12x_3 = 0.$

Solving these equations, we get $x_1=0$, $x_2 = 0$ and $x_3 = 0$.

$$|H| = \begin{vmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{vmatrix} = \begin{vmatrix} 2 & -3 & 0 \\ -3 & 6 & 4 \\ 0 & 4 & 12 \end{vmatrix}$$

for $(x_1, x_2, x_3) = (0, 0, 0)$.

Where $f_{ij} = f_{x_i x_j}$ $\forall$ i, j = 1, 2, 4.

$|H_1| = 2 > 0$, $|H_2| = \begin{vmatrix} 2 & -3 \\ -3 & 6 \end{vmatrix} > 0$

and $|H_3| = \begin{vmatrix} 2 & -3 & 0 \\ -3 & 6 & 4 \\ 0 & 4 & 12 \end{vmatrix} > 0.$

i.e., all principal minors are positive. Hence, the given function has minimum value equal to zero at point (0, 0, 0).

ii) Determine whether the function $\upsilon = f(x, y, z)$

$= -x^3 + 3xz + 2y - y^2 - 3z^2$.

F.O.C: $f_1 = -3x^2 + 3z = 0$

$f_2 = 2 - 2y = 0$

$f_3 = 3x - 6z = 0$

Solving these equations we get $x = 0$, or ½; $y = 1$; $z = 0$ or ¼.

Thus, there are two stationary points (0, 1, 0) and (1/2, 1, 1/4) and we have to examine whether the given function has maximum or minimum value at those points.

$$\text{S.O.C: } |H| = \begin{vmatrix} f_{xx} & f_{xy} & f_{xz} \\ f_{yx} & f_{yy} & f_{yz} \\ f_{zx} & f_{zy} & f_{zz} \end{vmatrix} = \begin{vmatrix} -6x & 0 & 3 \\ 0 & -2 & 0 \\ 3 & 0 & -6 \end{vmatrix}$$

At (0, 1, 0)

$$|H| = \begin{vmatrix} 0 & 0 & 3 \\ 0 & -2 & 0 \\ 3 & 0 & -6 \end{vmatrix}$$

$f_{xx} = 0$ Thus, S.O.C. cannot be fulfilled.

At (1/2, 1, ¼),

$$|H| = \begin{vmatrix} -3 & 0 & 3 \\ 0 & -2 & 0 \\ 3 & 0 & -6 \end{vmatrix}$$

$|H_1| = -3 < 0$, $|H_2| = 6 > 0$, $|H_3| = -18 < 0$

which alternate in sign.

Hence at (1/2, 1, ¼),

the given function possesses maximum value which will be $\frac{17}{16}$.

Check Your Progress 4

1) Examine the following functions for maxima or minima.

a) $z = 2x^2 + xy + 4y^2 + xz + z^2 + 2$

b) $z = e^{2x} - e^y + e^{z2} - 2(x + e) + y$.

....

....

....

....

4.4 LET US SUM UP

In this unit, we have discussed the extrema of a function and the condition under which it attains (relative/ absolute) extrema. We started with single variable case for simplicity extended the argument for multivariable cases. The emphasis was on the technique of how to obtain extrema of a function, which is differentiable. We have ignored the pathological cases when the function is discontinuous or non-differentiable.

4.5 KEY WORDS

Stationary Points: The points, at which first order derivatives are zero, are called stationary points. The values of the function at those points are called stationary values.

Local (relative) Maxima and Global (absolute) Maxima: A function attains local maximum at any particular point in the domain of definition implies that the value of the function is maximum in the neighborhood of that point only, and the function can assume higher values elsewhere. This can be non-unique. The global maximum is the highest value of the function with reference to its entire domain and hence unique.

Local (relative) Minima and Global (absolute) Minima: A function attains local minimum at any particular point in the domain of definition implies that the value of the function is minimum in the neighborhood of that point only, and the function can assume lower values elsewhere. This can be non-unique. The global minimum is the lowest value of the function with reference to its entire domain and hence unique.

Points of Inflextion: The point of inflextion is defined as a point at which a curve changes its curvature. The sufficient condition for a point of inflextion is $f''(x) = 0 \; and \; f'''(x) \neq 0$.

Thus, if a function has $f'(x) = 0, f''(x) = 0 \; and \; f'''(x) \neq 0$ at the point x, the point is said to be stationary and inflextional.

4.6 SOME USEFUL BOOKS

Allen, R.G.D. (1938), *Mathematical Analysis for Economists*, St. Martin's Press, New York.

Chiang, A.C. (1974), *Fundamental Methods of Mathematical Economics*, 2nd edition, McGrow-Hill Book Company, New York.

Henderson, James M. and Quandt, Richard E. (1980), *Microeconomic Theory*, McGrow-Hill Book Company, New York.

4.7 ANSWER OR HINTS TO CHECK YOUR PROGRESS EXERCISES

Check Your Progress 1

1) $f'(x) = 3(x^2 - 4x + 8) = 3\{(x-2)^2 + 4\} > 0 \quad \forall x$. Thus, the function does not have a maximum or a minimum.

2) Examine whether $x^{1/x}$ possesses a maximum or a minimum value.

$$y = x^{1/x} \rightarrow \log\ y = \frac{1}{x}\log x$$

$$\therefore \frac{1}{y}.\frac{dy}{dx} = \frac{1}{x^2} - \frac{1}{x^2}\log x = \frac{1}{x^2}(1-\log x) \text{ ---(i)}$$

$$\therefore \text{ when } \frac{dy}{dx} = 0, 1-\log x = 0 \ \therefore x = e.$$

$$\text{Again, } -\frac{1}{y^2}.\left(\frac{dy}{dx}\right)^2 + \frac{1}{y}.\frac{d^2y}{dx^2}$$

$$= \frac{x^2\left(-1/x\right) - (1-\log x)2x}{x^4}$$

$$= \frac{-3+2\log x}{x^2}$$

$$\text{when x=e, } \frac{d^2y}{dx^2} = e^{1/e}.\frac{-3+2}{e^3} \quad = -e^{1/e}/e^3$$

which is negative.

$\therefore$ for x = e, the function is a maximum and the maximum value is $e^{1/e}$.

3) To find the maxima and minima of 1 + 2 Sin x + 3 Cos^2 x. $\left(0 \le x \le \frac{\pi}{2}\right)$ get

f(x) = 1 + 2 Sin x+ 3 Cos^2 x

f′(x) = 2 Cos x – 6 Cox x Sin x

f′(x) = 0 when 2 Cos x(1 – 3 Sin x) = 0.

i.e., when Cos x = 0 and when Sin $x = \frac{1}{3}$.

f′′(x) = - 2 Sin x – 6 (Cos^2x-Sin^2x)

when Cos x = 0, $x = \frac{\pi}{2}$ $\therefore$ Sin x =1

$\therefore$ f′′(x) = - 2 + 6 = 4 > 0.

For Cos x = 0, f(x) is a minimum and the minimum value is 3.

When Sin $x = \frac{1}{3}$

f′′(x) = - 2 Sin x – 6(1 – 2 Sin^2x)

$$f''(x)\Big|_{Sinx=\frac{1}{3}} = -\frac{2}{3} - 6\left(1-\frac{2}{9}\right) < 0$$

Therefore, for Sin $x = \frac{1}{3}$, f(x) is a maximum and the maximum value is $4\frac{1}{3}$.

Check Your Progress 2

1) $x = -1/3$ is the point of inflextion.

Check Your Progress 3

1) Solution:

Here, $fx = 6x + 6y$, $fy = 6x + 14y$.

$fxx = 6$, $fxy = 6$, $fyy = 14$.

F.O.C. requires $fx = fy = 0$ i.e.,

$6x + 6y = 0$ ....(i)

$6x + 14y = 0$ ....(ii)

Solving (i) and (ii) for x and y we get, $x = 0$, $y = 0$.

The given function reaches its minimum value at the stationary point and its minimum value is zero. This is because

$$fxx\Big|_{\substack{x=0\\y=0}} = 6 > 0, fyy\Big|_{\substack{x=0\\y=0}} = 14 > 0.$$

Also, $fxx.fyy - (fxy)^2\Big|_{\substack{x=0\\y=0}}$

$= 84 - 36 = 48 > 0$.

2) Solution:

Here $fx = 8x + 7y$, $fy = -4y + 7x$, $fxx = 8$, $fxy = 7$, $fyy = -4$

F.O.C. for having stationary values is $fx = fy = 0$

i.e., $8x + 7y = 0$

$-4y + 7x = 0$

Solution of these equations gives us $x=0$, $y = 0$. Thus, the given function has only one stationary point at (0, 0). Now fxx and fyy have opposite signs and $fxx.fyy - (fxy)^2 < 0$, the stationary point for the given function is a saddle point.

3) $x_1 = -3$, $x_2 = 2$ saddle point

$x_1 = 7$, $x_2 = -5$ saddle point

$x_1 = -3$, $x_2 = -5$ relative maximum

$x_1 = 7$, $x_2 = 2$ relative minimum

4) The necessary conditions for maximum or minimum are

$f_1 = -3x_1^2 + 0 = 0$ ...(i)

$f_2 = -8x_2 = 0$ ...(ii)

from (i) and (ii) we get $x_1 = \pm\sqrt{3}$, $x_2 = 0$.

Thus, we get two stationary points for this function, namely, $(\sqrt{3},0)$ and $(-\sqrt{3},0)$. At stationary point, $(\sqrt{3},0)$ the function has a maximum. At $(-\sqrt{3},0)$ the function has a saddle point.

Check Your Progress 4

1) a) No maxima or minima

b) Maximum at (0,0,0)

UNIT 5 UNCONSTRAINED OPTIMISATION

Structure

5.0 OBJECTIVES

After going through this unit, we should able to:

- solve simple unconstrained optimisation problems; and
- solve simple economic problems dealing with unconstrained optimisation.

5.1 INTRODUCTION

In this unit, our objective is to find and explain the necessary and sufficient conditions for any unconstrained optimisation problem. Optimisation in mathematical sense (and therefore for our purpose) means either maximisation or, minimisation of different mathematical functions. Optimisation is broadly distinguishable in two respects viz., Unconstrained Optimisation (which is our current focus) and Constrained Optimisation (which we discuss in the next unit). Unconstrained optimisation deals with those problems whose domain is not compressed by any constrained. We are going to use simple calculus to solve these kinds of optimisation problems. First, let us we find and discuss the necessary first order conditions for unconstrained optimisation before moving on to the sufficient condition or, the second order condition. In the next section, we will focus on some simple economic applications that deal with unconstrained optimisation.

5.2 UNCONSTRAINED OPTIMISATION

Under optimisation exercise we either maximise or, minimise some mathematical functions consists of either one variable or, several variables. To optimize a mathematical function, we need to find necessary first order conditions from which the optimum values of the variables can be found. But to find whether the value is maximum or, minimum we have to check the sufficient second order condition.

5.2.1 First Order Condition

Consider the function $y = f(x)$, which can be plotted as shown below. The function has the maximum point A and the minimum point B. Note that at these two points the value of y must be stationary. We can say that it is a necessary conditions for an extremum of y that $dy = 0$ instantaneously as x varies. This condition is known as the differential version of the first order condition for an extremum and we have already discussed it in the preceding unit. As we have said earlier, this condition is necessary but not sufficient for either a maximum or, a minimum. Therefore, it is clear that the necessary first order condition for both maxima and minima are the same.

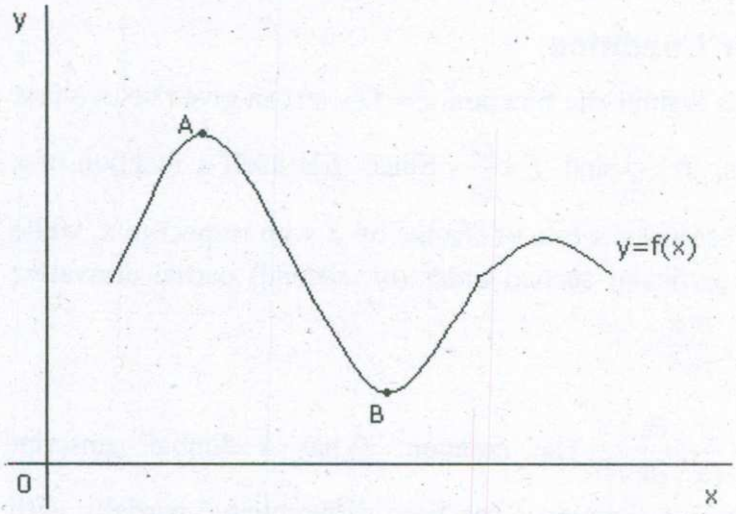


Fig 5.1: First order condition for optimum

We will see that the above condition is equivalent to the derivative version of the first order condition $\frac{dy}{dx} = 0$ or, $f^{'}(x) = 0$. We note that when there is no change in x (dx = 0), dy will automatically be zero. But this, of course, is not what the first order condition is all about. What the first order condition requires is that dy be zero as x is varied, that is, as arbitrary (positive or, negative, but not zero) infinitesimal changes of x occur. In such a context, with $dx \neq 0$, dy can be zero if and only if $f^{'}(x) = 0$. Thus, the derivative condition $f^{'}(x) = 0$ and the differential condition dy = 0 "for arbitrary nonzero values of dx" are indeed equivalent.

Now suppose $y = f(x_1, x_2, ..., x_n)$ then in this case, the first order condition for optimisation is given by the n-equation viz., $\frac{\partial y}{\partial x_1} = f^1(x) = 0$, $\frac{\partial y}{\partial x_2} = f^2(x) = 0$, ..., $\frac{\partial y}{\partial x_n} = f^n(x) = 0$. From these n equations, we can find $x_1^*, x_2^*, ..., x_n^*$.

Example: Find the optimal values of $y = x^3 - x$ in the domain $-\infty < x < +\infty$.

Solution: The necessary first order condition gives $\frac{dy}{dx} = 0$, which can be transformed to $3x^2 - 1 = 0$ - (I). From equation (I) we get at optimum either $x^* = +\frac{1}{\sqrt{3}}$ or, $x^* = -\frac{1}{\sqrt{3}}$. Putting these values of x, we get optimum y i.e., y^*.

Example: Find the optimal values of $z = y^2 - xy + x^2$ on the domain $-\infty < x < +\infty$ and $-\infty < y < +\infty$.

Solution: The necessary first order condition gives $\frac{\partial z}{\partial x}=0$ and $\frac{\partial z}{\partial y}=0$, from which we can get $-y+2x=0$ and $2y-x=0$. We solve this and get $x^*=0$ and $y^*=0$. Finally putting these optimum values into the original equation we get $z^*=0$.

5.2.2 Second Order Condition

For the time being, let us assume the function $z=f(x,y)$ can give rise two first order partial derivatives, $f_x = \frac{\partial z}{\partial x}$ and $f_y = \frac{\partial z}{\partial y}$. Since f_x is itself a function of x (as well as y), we can measure the rate of change of f_x with respect to x, while y remains fixed, by a particular second order (or, second) partial derivative denoted by either f_{xx} or, $\frac{\partial^2 z}{\partial x^2}$.

$f_{xx} \equiv \frac{\partial}{\partial x}(f_x)$ Or, $\frac{\partial^2 z}{\partial x^2} \equiv \frac{\partial}{\partial x}\left(\frac{\partial z}{\partial x}\right)$. The notation f_{xx} has a double subscript signifying that the primitive function f has been differentiated partially with respect to x twice, whereas the notation $\frac{\partial^2 z}{\partial x^2}$ resembles that of $\frac{d^2 z}{dx^2}$ except for the use of the partial symbol. In a perfectly analogues manner, we can use the second partial derivative $f_{yy} \equiv \frac{\partial}{\partial y}(f_y)$ or, $\frac{\partial^2 z}{\partial y^2} \equiv \frac{\partial}{\partial y}\left(\frac{\partial z}{\partial y}\right)$ to denote the rate of change of f_y with respect to y, while x is held constant.

Recall however, that f_x is also a function of x. Hence, there can be written two more partial derivatives: $f_{xy} \equiv \frac{\partial^2 z}{\partial x \partial y} \equiv \frac{\partial}{\partial x}\left(\frac{\partial z}{\partial y}\right)$ and $f_{yx} \equiv \frac{\partial^2 z}{\partial y \partial x} \equiv \frac{\partial}{\partial y}\left(\frac{\partial z}{\partial x}\right)$.

These are called cross (or, mixed) partial derivatives because each measures the rate of change of one first order partial derivative with respect to the "other" variable. We can also note that using total differential equation we get $d^2 z \equiv d(dz) \equiv \frac{\partial(\partial z)}{\partial x}dx + \frac{\partial(\partial z)}{\partial y}dy$.

Using the concept of $d^2 z$, we can state the second order sufficient condition for a maximum of $z=f(x,y)$ as follows: $d^2 z<0$ for arbitrary values of dx and dy, not both zero – (a). The rationale behind is very similar to that of the $d^2 z$ condition for the one variable case and it can be explained by means of a figure that depicts the bird's-eye view of a surface. Let point A on the surface – the point lying directly above the point (x_0, y_0) in the domain – satisfy the first order condition. Then the point A is a prospective candidate for a maximum. Whether it in fact qualifies depends on the surface configuration in the neighborhood of A. If an infinitesimal movement away from A in any direction along the surface invariable results in a decrease in z – that is, if dz < 0 for arbitrary values of dx and dy, not both zero – A is a pick of a dome. Given that dz = 0 at point A, however, the condition dz < 0 at point in the neighborhood of A amounts to the stipulation that dz is decreasing, that is, $d^2 z \equiv d(dz)<0$, for arbitrary values of dx and dy, not both zero. Thus (a) constitutes a sufficient condition for identifying a stationary value as a maximum of z. Analogous reasoning would show that a counterpart second order sufficient condition for identifying value as a minimum of z = f(x, y) is as follows: $d^2z > 0$ for arbitrary values of dx and dy, not both zero – (b).

The reason why (a) and (b) are sufficient conditions but not necessary conditions, is that it is again possible for d^2z to take a zero value at a maximum or, a minimum. For this reason, second order necessary conditions must be stated with weak inequalities as follows: For maximum of z: $d^2z \le 0$ for arbitrary values of dx and dy, not both zero. For minimum of z: $d^2z \ge 0$ for arbitrary values of dx and dy, not both zero – (c). In the following, however, we pay more attention to the second order sufficient conditions.

For operational convenience, second order differential conditions can be translated into equivalent conditions on second order derivatives. In the two-variable case, this would entail restrictions on the signs of the second partial derivatives f_{xx}, f_{yy} and f_{xy}. The actual translation would require knowledge of quadratic forms. But we may first introduce the main results here: For any values of dx and dy, not both zero, we have

$d^2z < 0$ if and only if $f_{xx} < 0$, $f_{yy} < 0$, and $f_{xx}f_{yy} > f_{xy}^2$

$d^2z > 0$ if and only if $f_{xx} > 0$, $f_{yy} > 0$, and $f_{xx}f_{yy} > f_{xy}^2$.

Note that the sign of d^2z hinges not only on f_{xx} and f_{yy}, which have to do with the surface configuration around point A, but also on the cross partial derivative f_{xy}. The role played by this latter partial derivative is to ensure that the surface in question will yield (two-dimensional) cross sections with the same type of configuration in all possible directions.

Table 5. 1: Conditions for Relative Extremum: z=f(x, y)

Condition	**Maximum**	**Minimum**
First order necessary condition	$f_x = f_y = 0$	$f_x = f_y = 0$
Second order sufficient condition	$f_{xx}, f_{yy} < 0$	$f_{xx}, f_{yy} > 0$
	and $f_{xx}f_{yy} > f_{xy}^2$	and $f_{xx}f_{yy} > f_{xy}^2$

The above result, together with the first order condition, enables us to construct Table 5.1. It should be understood that all the second partial derivatives therein are to be evaluated at the stationary point where $f_x = f_y = 0$. It should also be stressed that the second order sufficient condition is not necessary for an extremum. In particular, if a stationary value is characterised by $f_{xx}f_{yy} = f_{xy}^2$ in violation of that condition, that stationary value may nevertheless turn out to be an extremum. On the other hand, in the case of another type of violation, with a stationary point characterised by $f_{xx}f_{yy} < f_{xy}^2$, we can identify that point as a saddle point, because the sign of d^2z will in that case be indefinite (positive for same values of dx and dy, but negative for others).

For n-variable case we can generalise the above second order condition. The satisfaction of the first order condition earmarks certain values of z as the stationary values of the objective function. If at a stationary value of z we find that d^2z is positive definite, this will suffice to establish that value of z as a minimum. Analogously, the negative definiteness of d^2z is a sufficient condition for the stationary value to be maximum.

When there are n choice variables, the objective function may be expressed as $z = f(x_1, x_2, ..., x_n)$.

The total differential will then be $dz = f_1dx_1 + f_2dx_2 + ... + f_ndx_n$ so that the necessary condition for extremum (dz = 0 for arbitrary dx_i) means that all the n first order partial derivatives are required to be zero.

The second order differential d^2z will again be a quadratic form, which can be expressed as an $n \times n$ array. The coefficients of that array, properly arranged, will now give the (symmetric) Hessian

$$|H| = \begin{bmatrix} f_{11} & f_{12} & \cdots & f_{1n} \\ f_{21} & f_{22} & \cdots & f_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ f_{n1} & f_{n2} & \cdots & f_{nn} \end{bmatrix}$$

with principle minors $|H_1| = f_{11}$, $|H_2| = \begin{bmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{bmatrix}$,..., $|H_n| = |H|$. The second order sufficient condition for extremum is, as before, that all the n principal minors be positive (for a minimum in z) and that they duly alternate in sign (for a maximum in z), the first one being negative.

In summary, then – if we concentrate on the determinantal test – we have the criteria as listed in Table 5.2, which is valid for an objective function of any number of choice variables. As special cases, we can have n = 1 or, n = 2. When n = 1, the objective function is z = f(x), and the condition for maximization, $f_1 = 0$ and $|H_1| < 0$, reduce to $f'(x) = 0$ and $f''(x) < 0$, exactly as we learned above. Similarly, when n = 2, the objective function is $z = f(x_1, x_2)$, so that the first order condition for maximum is $f_1 = f_2 = 0$, whereas the second order sufficient condition becomes $f_{11} < 0$ and $\begin{bmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{bmatrix} = f_{11}f_{22} - f_{12}{}^2 > 0$, which is merely a restatement of the information given above.

Table 5. 2: Determinantal Test for Relative Optimum: $z = f(x_1, x_2, ..., x_n)$

Condition	Maximum	Minimum
First order necessary condition	$f_1 = f_2 = ... = f_n$	$f_1 = f_2 = ... = f_n$
Second order sufficient condition	$(-1)^i \mid H_i \mid > 0\ \forall$ i=1,...,n	$\mid H_i \mid > 0\ \forall$ i=1,...,n

Example: Find the extreme value(s) of $z = x^2 + xy + 2y^2 + 3$ and determine whether they are maximum or, minimum.

Solution: The relevant first order conditions are the following:

$f_x = 2x + y = 0$ and $f_y = x + 4y = 0$. Solving these two equations we get $x^* = 0$ and $y^* = 0$. Note that $f_{xx} = 2 > 0$ and $f_{yy} = 4 > 0$ and as $f_{xy} = 1$ we have $f_{xx}f_{yy} > f_{xy}{}^2$. So at $x^* = 0$ and $y^* = 0$ the function z is minimized.

Example: Find the extreme value(s) of $z = -x^2 + xy - y^2 + 2x + y$ and determine whether they are maximum or, minimum.

Solution: The relevant first order conditions are the following:

$f_x = -2x + y + 2 = 0$ and $f_y = x - 2y + 1 = 0$. Solving these two equations we get $x^* = 5/3$ and $y^* = 4/3$. Now note that $f_{xx} = -2 < 0$ and $f_{yy} = -2 < 0$ and as $f_{xy} = 1$ we have $f_{xx}f_{yy} > f_{xy}^2$. So at $x^* = 5/3$ and $y^* = 4/3$ the function z is maximised.

Check Your Progress 1

1) Find the extreme value(s) of $z = -x^2 + xy - y^2 + x + 5y$ and determine whether they are maximum or, minimum.

....

....

....

....

....

....

....

2) Find the extreme value(s) of $y = x^2 - 6$ and determine whether they are maximum or, minimum.

....

....

....

....

....

....

....

3) Find the extreme value(s) of $y = x^3 - 2x^2 + x - 6$ and determine whether they are maximum or, minimum.

....

....

....

....

………………………………………………………………………………

………………………………………………………………………………

………………………………………………………………………………

4) Find the extreme value(s) of $z = x^2 - 2x - y^2$ and determine whether they are maximum or, minimum.

………………………………………………………………………………

………………………………………………………………………………

………………………………………………………………………………

………………………………………………………………………………

………………………………………………………………………………

………………………………………………………………………………

………………………………………………………………………………

5.3 SOME ECONOMIC PROBLEMS

In this section, we will give some economic application concerning unconstrained optimisation. Our first example is concerning the multi-product firm who takes the prices as given i.e., she produces under competitive structure. The second example is also regarding multi-product firm but here the firm acts as a monopolist where she can choose how much to produce and at what price. In both the cases the firm's objective is to maximise her own profit where profit is defined as net revenue minus net cost. These objectives we take for simplicity i.e., one can think of other objective of the firm viz., revenue maximisation, but that is not going to be discussed in this unit. For simplicity again, we take cost function to be the same in both example.

Example 1: This example is regarding the problem of a multi-product firm. Let us first postulate a two-product firm under circumstances of pure competition. Since with pure competition the prices of both commodities must be taken as exogenous, these will be given by $p_1 = 5$ and $p_2 = 3$, respectively. Accordingly, the firm's revenue function will be $R = p_1q_1 + p_2q_2 = 5q_1 + 3q_2$, where q_i represents the output level of the i^{th} product per unit of time. The firm's cost function is assumed to be $C = 2q_1^2 + 2q_2^2 + q_1q_2$.

Note that $\dfrac{\partial C}{\partial q_1} = 4q_1 + q_2$ (the marginal cost of the first product) is a function not only of q_1 but also of q_2. Similarly, the marginal cost of the second product also depends, in part, on the output level of the first product. Thus, according to the assumed cost function, the two commodities are seen to be technically related in production.

The profit function of this hypothetical firm can now be written readily as $\pi = R - C = 5q_1 + 3q_2 - 2q_1^2 - 2q_2^2 - q_1q_2$, a function of two choice variables (q_1 and q_2) and two price parameters. It is our task to find the levels of q_1 and q_2, which in combination, will maximise π. For this purpose, we find the first order partial derivatives of the profit function:

$$\pi_1\left(\equiv \frac{\partial \pi}{\partial q_1}\right) = 5 - 4q_1 - q_2 \text{ and } \pi_2\left(\equiv \frac{\partial \pi}{\partial q_2}\right) = 3 - 4q_2 - q_1 \text{ - (a)}$$

Setting these both equal to zero, to satisfy the necessary condition for maximum, we get the two simultaneous equations $4q_1 + q_2 = 5$ and $q_1 + 4q_2 = 3$, which yield the unique solution $q_1^* = \frac{4p_1 - p_2}{15} = \frac{17}{15}$ and $q_2^* = \frac{4p_2 - p_1}{15} = \frac{7}{15}$.

Thus, the maximum profit that the firm will get is $\pi^* = 5.51$ (approx).

To be sure that this does represent a maximum profit, let us check the second order condition. The second partial derivatives, obtainable by partial differentiation of equation (a), give us the following Hessian:

$$|H| = \begin{bmatrix} \pi_{11} & \pi_{12} \\ \pi_{21} & \pi_{22} \end{bmatrix} = \begin{bmatrix} -4 & -1 \\ -1 & -4 \end{bmatrix}$$

Since $|H_1| = -4 < 0$ and $|H_2| = 15 > 0$, the Hessian matrix (or, d^2z) is negative definite, and the solution does maximise the profit. In fact, since the signs of the principal minors do not depends on where they are evaluated, d^2z is in this case everywhere negative definite. Thus, the objective function must be strictly concave, and the maximum profit found above is actually a unique absolute maximum.

Example 2: This example again regarding a multi-product firm but with the modification that in earlier case the firm face given price i.e., multi-product firm under competitive structure, in this example the multi-product firm is a monopoly in both the product i.e., the firm can decide how much quantity to be produced and at the same time she can choose any price for her product. Now by virtue of this new market-structure assumption, the revenue function must be modified to reflect the fact that the prices of the two products will now vary with their output levels (which are assumed to be identical with their sales levels, no inventory accumulation being contemplated in the model). The exact manner in which prices will vary with output levels is, of course, to be found in the demand functions for the firm's two products.

Suppose that the demand facing the monopolistic firm, are as follows:

$$q_1 = 40 - 2p_1 + p_2 \text{ and } q_2 = 15 + p_1 - p_2 \text{ - (a)}$$

These equations reveal that the two commodities are related in consumption; specifically, they are substitute goods, because an increase in the price of one will raise the demand for the other. As given, equation (a) expresses the quantities demanded q_1 and q_2 as functions of prices, but for our present purposes, it will be more convenient to have prices p_1 and p_2 expressed in terms of the sales volumes q_1 and q_2, that is, to have average revenue functions for the two products. Since equation (a) can be rewritten as

$-2p_1 + p_2 = q_1 - 40$ and $p_1 - p_2 = q_2 - 15$, we may (considering q_1 and q_2 as parameters) apply Cramer's rule to solve for p_1 and p_2 as follows:

$$p_1 = 55 - q_1 - q_2 \text{ and } p_2 = 70 - q_1 - 2q_2 \text{ - (b)}$$

These constitute the desired average-revenue functions, since $p_1 \equiv AR_1$ and $p_2 \equiv AR_2$. Consequently, the firm's total-revenue function can be written as:

$$R = p_1q_1 + p_2q_2$$
$$= (55 - q_1 - q_2)q_1 + (70 - q_1 - 2q_2)q_2$$
$$= 55q_1 + 70q_2 - 2q_1q_2 - q_1^2 - 2q_2^2$$

If we again assume the total cost function to be $C = q_1^2 + q_1q_2 + q_2^2$, then the profit function will be:

$$\pi = R - C$$
$$= 55q_1 + 70q_2 - 3q_1q_2 - 2q_1^2 - 3q_2^2 \text{ - (c)},$$

which is an objective function with two choice variables. Once the profit maximizing output levels q_1^* and q_2^* are found, however, the optimal prices p_1^* and p_2^* are easy enough to find from equation (b).

The objective function yields the following first and second partial derivatives:

$$\pi_1 = 55 - 3q_1 - 4q_2 \qquad \pi_2 = 70 - 3q_1 - 6q_2$$

$$\pi_{11} = -4 \qquad \pi_{12} = \pi_{21} = -3 \qquad \pi_{22} = -6$$

To satisfy the first order condition for a maximum of π, we must have $\pi_1 = \pi_2 = 0$; that is $4q_1 + 3q_2 = 55$ and $3q_1 + 6q_2 = 70$, thus the solution output levels (per unit of time) are $(q_1^*, q_2^*) = (8, 7\frac{2}{3})$.

Upon substitution of this result into equation (b) and (c), respectively, we find that $p_1^* = 39\frac{1}{3}$ $p_2^* = 46\frac{2}{3}$ and $\pi^* = 488\frac{1}{3}$ (per unit of time)

Inasmuch as the Hessian is $\begin{bmatrix} -4 & -3 \\ -3 & -6 \end{bmatrix}$, we have $|H_1| = -4 < 0$ and $|H_2| = 15 > 0$, so that the value of π^* does represent the maximum profit. Here, the signs of the principle minors are again independent of where they are evaluated. Thus, the Hessian matrix is everywhere negative definite, implying that the objective function is strictly concave and that it has a unique absolute maximum.

5.4 LET US SUM UP

In this unit we emphasise an important tool, which has enormous application in economic viz., unconstrained optimisation. Unconstrained optimisation problem deals with optimisation (i.e., either minimisation or, maximisation) problems where the domain of objective function is not limited by any constrained. There are many real life economic problems for which unconstrained optimisation technique is necessary. Any problem regarding profit maximisation, we use the technique of unconstrained optimisation. In this unit, we emphasise the technique, how to solve unconstrained optimisation problems. For that we first introduced the first order necessary conditions for optimisation. We saw that if there are n choice variables we must have n number of equality as first order condition to solve the optimum value of each variable. We also saw that the first order conditions are same for both maximisation and minimisation problem, and therefore we can't get whether by putting the optimum values of variables we get maximum or, minimum value of the objective function. To solve this kind of problem we therefore introduce second order sufficient condition for optimisation. We saw second order condition is different for maximisation and minimisation exercise. We also saw that for a given objective function and for given optimum values of choice variables, if the corresponding Hessian matrix is negative definite at optimum then the optimum values give the maximum value of the objective function. Analogously, if the Hessian matrix is positive definite at optimum then the optimum values give the minimum value of the objective function. Therefore, from the first order condition (supported by second order condition), we can determine the maximum or, the minimum value of the objective function. There is a notion of necessary second order condition which tells that the necessary second order condition for maximization is that, the Hessian matrix should be negative semi-definite and the necessary second order condition for minimisation is that, the Hessian matrix should be positive semi-definite.

Finally, we give some interesting economic applications that clear the techniques of unconstrained optimisation problem. For that our first example is a problem regarding multi-product firm under competitive structure. Then we give an example again concerning multi-product firm but now the firm acting as monopolist in all markets.

5.5 KEY WORDS

Positive Definite Matrix: A symmetric matrix A is positive definite if it is true that for every vector x, the real number which results from the multiplication x'Ax > 0 for all nonzero x. Such a matrix has real positive eigenvalues.

Positive Semi-definite Matrix: A symmetric matrix A is positive definite if it is true that for every vector x, the real number which results from the multiplication x'Ax ≥ 0 for all nonzero x.

Negative Definite Matrix: A symmetric matrix A is negative definite if it is true that for every vector x, the real number which results from the multiplication x'Ax ≤ 0 for all nonzero x. Such a matrix has real negative eigenvalues.

Negative Semi-definite Matrix: A symmetric matrix A is negative semi-definite if it is true that for every vector x, the real number which results from the multiplication x'Ax ≤ 0 for all nonzero x.

Stationary Point: In calculus, a **stationary point** is referred to a point on a graph where the tangent to the graph is parallel to the *x*-axis or, equivalently, where the derivative of the function equals zero.

Necessary and Sufficient Condition for Optimum: Given a function $z = f(x_1, x_2, ..., x_n)$, the conditions are obtained as the following:

Condition	Maximum	Minimum
First order necessary condition	$f_1 = f_2 = ... = f_n$	$f_1 = f_2 = ... = f_n$
Second order sufficient condition	$(-1)^i \mid H_i \mid > 0\ \forall$ i =1,...,n	$\mid H_i \mid > 0\ \forall$ i=1,...,n

MEC-103-2-6F-7K

5.6 SOME USEFUL BOOKS

Henderson, James M. and Quandt, Richard E. (1980), *Microeconomic Theory*, McGrow-Hill Book Company, New York.

Allen, R.G.D. (1938), *Mathematical Analysis for Economists*, St. Martin's Press, New York.

Varian, Hal (1992), *Microeconomic Analysis*, W.W. Norton & Company, Inc., Yew York.

5.7 ANSWERS OR HINTS TO CHECK YOUR PROGRESS EXERCISE

Check Your Progress 1

1) See the example given in Sub-section 5.2.2.

2) The objective function is $y = x^2 - 6$. Therefore, from the first order condition for optimisation, we get $2x = 0$. Solving this equation we have x = 0. The second order condition gives that $d^2y = 2 > 0$, which is not a function of x. Therefore, we have $d^2y = 2$ at x = 0. So at x = 0, we reach the minima of the objective function.

3) The objective function is $y = x^3 - 2x^2 + x - 6$. Therefore, we get the first order condition $3x^2 - 4x + 1 = 0$ or, either x = $\frac{1}{3}$ or, x = 1. Note that $d^2y = 6x - 4$. So, $d^2y > 0$ at x = 1 and $d^2y < 0$ at x = $\frac{1}{3}$. Therefore, at x = 1, we have the minima and at x = $\frac{1}{3}$.

4) The objective function is $z = x^2 - 2x - y^2$. The necessary first order condition yields $2x - 2 = 0$ and 2y = 0. From these two first order conditions we get x = 1 and y = 0. The corresponding Hessian matrix is $|H| = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$, so we get that $|H_1| = 2 > 0$ and $|H_2| = 4 > 0$. Therefore, we have the minimum value of z at x =1 and y = 0.

UNIT 6 CONSTRAINED OPTIMISATION

Structure

6.0 OBJECTIVES

After going through this unit, you will be able to:

- solve the basic optimisation problems with equality as well as inequality constraint using Lagrange method; and
- use Kuhn-Tucker method while solving optimisation problems with inequality constraint.

6.1 INTRODUCTION

When an economic project such as the production of a specified level of output is to be carried out, there are normally a number of alternative ways of accomplishing it. One (or more) of these alternatives will, however, be more desirable than others from the standpoint of a producer. Need, therefore, arises to choose the best alternative available without compromising the basic objective of the project. Essentially, the constraint optimisation helps arrive the decision.

The most common criterion of choice among alternatives in economics is the goal of maximising something (such as maximising a firm's profit, a consumer's utility, or the rate of growth of firm or of a country's economy) or, of minimising something (such as minimising the cost of producing a given output). Economically, we may categorize such maximisation and minimisation problems under the general heading of optimisation, meaning "the quest for the best". From a purely mathematical point of view, however, the terms maximum and minimum do not carry with them any connotation of optimality. Therefore, the collective term for maximum and minimum, as mathematical concepts, is the more matter-of-fact designation extremum, meaning an extreme value.

In formulating an optimisation problem, the first order of business is to delineate an objective function in which the dependent variable represents the object to maximisation or minimisation and in which the set of independent variables indicates the objects whose magnitudes the economic unit in question can pick and choose, with a view to optimising. We shall, therefore, refer to the independent variables as choice variables. They can also be called decision variables, or policy variables. The essence of the optimisation process is simply to find the set of values of the choice variables that will yield the desired extremum of the objective function.

6.2 EQUALITY CONSTRAINED OPTIMISATION

We have been given an objective function $f: R^n \to R$ to maximise, subject to k constraints. That is, there are k functions, $g^1: R^n \to R, g^2: R^n \to R, ..., g^k: R^n \to R$, and we wish to

Maximise f(x) over all $x \in R^n$ such that $g^1(x)=0,..., g^k(x)=0$.

More compactly, collect the constraint functions (looking them as component functions) into one function $g: R^n \to R$, where $g(x) = (g^1(x),...,g^k(x))$. Then what we want is to

Maximise f(x) over all $x \in R^n$ such that $g(x)_{1xk} = \theta_{1xk}$.

The theorem of Lagrange provides the necessary conditions for a local optimum x^*. By local we mean that the value $f(x^*)$ is a maximum or minimum compared to other values f(x) for all x contained in some open set U containing x^* such that x satisfies the k constraints. Thus, the problem considered is to provide necessary conditions for a maximum or a minimum of f(x) over all $x \in S$, where $S = U \cap \{x \in R^n \mid g(x) = \theta\}$, for some open set U.

The following example illustrates the principle of 'no arbitrage' underlying a maximum. A more general illustration, with more than one constraint, requires a little bit of the machinery of linear inequalities for the illustration of which is out of the scope of this unit. However, such inequalities go a long way in optimisation theories of asset pricing, Nash equilibrium etc.

Example: Suppose x^* solves Max $U(x_1, x_2)$ subject to (s.t.), $I - p_1x_1 - p_2x_2 = 0$ and suppose $x^* > \theta$. Then reallocating a small amount of income from one good to the other does not increase utility. Say, income dI > 0 is shifted from good 2 to good 1. So $dx_1 = (dI / p_1) > 0$ and $dx_2 = -(dI / p_2) < 0$. Note that this reallocation satisfies the budget constraint, since

$p_1(x_1 + dx_1) + p_2(x_2 + dx_2) = I$. The change in utility is $dU = U_1dx_1 + U_2dx_2$, or $dU = [(U_1 / p_1) - (U_2 / p_2)]dI \leq 0$, since the change in utility cannot be positive at a maximum. Therefore, $(U_1 / p_1) - (U_2 / p_2) \leq 0$ -------- (1).

Similarly, dI > 0 shifted from good 1 to good 2 does not increase utility, so that $[-(U_1 / p_1) + (U_2 / p_2)]dI \leq 0$, or $[-(U_1 / p_1) + (U_2 / p_2)] \leq 0$ ------- (2).

Equations (1) and (2) imply $(U_1(x^*) / p_1) = (U_2(x^*) / p_2) = \lambda^*$ ----- (3).

That is, the marginal utility of the last bit of income equals $(U_1(x^*)/p_1) = (U_2(x^*)/ p_2)$ at the optimum. Also, (3) implies $U_1(x^*)=\lambda^* p_1$, $U_2(x^*)=\lambda^* p_2$. Along with $p_1x^*_1+p_2x^*_2=I$, these are the first order conditions of the Lagrangean function

$$\text{Max } L(x, \lambda) = U(x_1, x_2) + \lambda[I - p_1x_1 - p_2x_2]$$

More generally, suppose $F: R^n \to R$ and $G: R^n \to R$, and suppose x^* solves Max F(x) s.t. $c - G(x) = 0$.

Contemplate a change dx in x^* that respects the constraint G(x) = c.

That is, $dG = G_1x_1+G_2x_2 = 0$.

Therefore, $G_1x_1 = -G_2x_2 = dc$ and $dx_1 = (dc/G_1)$, $dx_2 = (dc/G_2)$. If dc>0, then change dx implies $dx_1>0$, $dx_2<0$. F does not increase at the maximum, x^*. So,

$dF = F_1x_1 + F_2x_2 \leq 0$,

or, $[(F_1 / G_1) - (F_2 / G_2)]dc \leq 0$.

Therefore,

$(F_1(x^*) / G_1(x^*)) = (F_2(x^*) / G_2(x^*)) = \lambda^*$ ----------- (4).

Caveat: We have assumed that $G_1(x^*)$ and $G_2(x^*)$ are not both zero at x^*. This is called the constraint qualification. Again, note that equation (4) can be taken as the first order condition (FONC) of the problem

$$\text{Max } L(x, \lambda) = F(x) + \lambda [c - G(x)].$$

Let's go back to the utility example. At the optimum (x^*, λ^*), suppose you increase income by ΔI. Buying more x_1 implies utility increases by $(U_1(x^*)/p_1)\,\Delta I$, approximately. Buying more x_2 implies utility increases by $(U_2(x^*)/p_2)\,\Delta I$. At the optimum, $(U_1(x^*)/p_1) = (U_2(x^*)/p_2) = \lambda^*$. So in either case, utility increases by $\lambda^*\Delta I$, where λ^* gives the change in the objective (here, the objective is utility), at an optimum that results from relaxing the constraint a little bit.

The interpretation is the same in the more general case: If $G(x) = c$, and c is increased by Δc. Suppose x_1 alone is then increased,

so, $\Delta G = g_1 dx_1 = \Delta c$,

or, $dx_1 = (\Delta c/G_1)$.

That means, at x^*, F increases by $dF = F_1 dx_1 = (F_1(x^*)/G_1(x^*))\Delta c = \lambda^*\Delta c$. If instead x_2 is changed, F increases by $dF = F_2 dx_2 = (F_2(x^*)/G_2(x^*))\Delta c = \lambda^*\Delta c$.

Now we discuss the Lagrange theorem in detail. The set up is the following:

$f : R^n \to R$ is the objective function, $g^i : R^n \to R$, i = 1,…,n k, k<n are the constraint functions.

Let $g : R^n \to R^k$ be the function given by $g(x) = (g^1(x),\ldots\ldots, g^k(x))$. $Df(x) = (\partial f(x)/\partial x_1, \ldots, \partial f(x)/\partial x_n)$.

$$Dg(x) = \begin{pmatrix} (\partial g^1(x)/\partial x_1) & \cdots & (\partial g^1(x)/\partial x_n) \\ \vdots & \ddots & \vdots \\ (\partial g^k(x)/\partial x_1) & \cdots & (\partial g^k(x)/\partial x_n) \end{pmatrix} = \begin{pmatrix} Dg^1(x) \\ \vdots \\ Dg^k(x) \end{pmatrix}$$

Thus, Dg(x) is a $k \times n$ matrix.

The theorem below provides a necessary condition for a local maximum or local minimum. Note that x^* is a local maximum (respectively, minimum) of f on the constraint set $\{x \in R^n \mid g^i(x) = 0, i = 1, \ldots, k\}$ if $f(x^*) \geq f(x)$ (respectively for $\leq f(x)$) for all $x \in U$ for some open set U containing x^*, s.t. $g^i(x) = 0$, i = 1,…,k. Thus x^* is a maximum on the set $S = U \cap \{x \in R^n \mid g^i(x) = 0, i = 1, \ldots, k\}$.

Theorem of Lagrange: Let
$f : R^n \to R$ and $g^1 : R^n \to R, g^2 : R^n \to R, \ldots, g^k : R^n \to R$, k < n be once continuously differentiable functions. Suppose x^* is a maximum or a minimum of f on the set $S = U \cap \{x \in R^n \mid g^i(x) = 0, i = 1, \ldots, k\}$, for some open set $U \subseteq R^n$. Then there exists real numbers $\mu^*, \lambda_1^*, \ldots, \lambda_k^*$, not all zero, such that

$$\mu Df(x^*) + \sum_{i=1}^{k} \lambda_i^* Dg^i(x^*) = \theta_{1 \times n}.$$

Moreover, if rank $(Dg(x^*)) = k$, then we may put $\mu = 1$.

Notes:

1) The condition rank $(Dg(x^*)) = k$ is called the constraint qualification. The first part of the theorem says that at a local maximum or minimum, under the assumption of continuous differentiability of f and g^i, $i = 1,\ldots, k$, the vectors $Df(x^*)$, $Dg^1(x^*),\ldots, Dg^k(x^*)$ are linearly dependent. The constraint qualification (CQ) basically assumes that the vectors $Dg^1(x^*),\ldots, Dg^k(x^*)$ are linearly independent. In that case, $\mu Df(x^*)+\sum_{i=1}^{k}\lambda_i^* Dg^i(x^*)=\theta_{1\times n}$ implies that μ cannot equal to zero, for then $\sum_{i=1}^{k}\lambda_i^* Dg^i(x^*)=\theta_{1\times n}$, which along with linear independence implies $\lambda_i^* = 0$, $i = 1,\ldots, k$. This cannot be. So, if the CQ holds, then $\mu \neq 0$, and we can divide through by μ.

2) In most applications, the CQ holds. We usually check first whether it holds, and then proceed. Suppose it does hold. Note that $\mu Df(x^*)+\sum_{i=1}^{k}\lambda_i^* Dg^i(x^*)=\theta_{1\times n}$ subsumes the following n equations: $(\partial f(x^*)/\partial x_j)+\sum_{i=1}^{k}\lambda_i^*(\partial g^i(x^i)/\partial x_j)=0, j=1,\ldots,n$. Note also that this leads to the 'usual' procedure for finding equality constrained maximum or minimum, by setting up a Lagrangean function:

$$L(x,\lambda)=f(x)+\sum_{i=1}^{k}\lambda_i g^i(x),$$

and solving the FONC

$$(\partial L(x,\lambda)/\partial x_i)=(\partial f(x)/\partial x_i)+\sum_{i=1}^{k}\lambda_i(\partial g^i(x)/\partial x_i)=0, j=1,\ldots,n$$

$(\partial L(x,\lambda)/\partial \lambda_i)=g^i(x)=0, i=1,\ldots,k$ which is (n+k) equations in (n+k) variables $x_1,\ldots, x_n, \lambda_1,\ldots, \lambda_k$.

Second order condition: These conditions are characterised by definiteness or semi-definiteness of the Hessian of the Lagrangean function, which is the appropriate function to look at in this constrained optimisation problem. In addition, we don't have to check the appropriate inequality for the quadratic form for all x; only those x are relevant that satisfy the constraints. Second order conditions in general captures the change of the curvature of the objective function around the local maximum or minimum, i.e., how the graph curves as we move from x^* to a nearby x. In constrained optimisation, we cannot move from x^* to any arbitrary x nearby; the move must be to an x, which satisfies the constraints. That is, such a move must leave all $g^i(x)$ at 0. In other words, $dg^i(x) = Dg^i(x).dx = 0$, where dx is a vector $x - x^*$ that must be orthogonal to $Dg^i(x)$. Thus, it is suffices to evaluate the appropriate quadratic form at all vectors x that are orthogonal to all the gradients of the constraint functions.

Since $L(x,\lambda)=f(x)+\sum_{i=1}^{k}\lambda_i g^i(x)$,

$$D^2L(x,\lambda)_{n\times n}=D^2 f(x)_{n\times n}+\sum_{i=1}^{k}\lambda_i D^2 g^i(x)_{n\times n}.$$

where $D^2 f(x) = \begin{pmatrix} f_{11}(x) & \cdots & f_{1n}(x) \\ \vdots & \ddots & \vdots \\ f_{n1}(x) & \cdots & f_{nn}(x) \end{pmatrix}$ and $D^2 g^i(x) = \begin{pmatrix} g^i{}_{11}(x) & \cdots & g^i{}_{1n}(x) \\ \vdots & \ddots & \vdots \\ g^i{}_{n1}(x) & \cdots & g^i{}_{nn}(x) \end{pmatrix}$

So, $D^2 L(x,\lambda)_{n\times n} = \begin{pmatrix} f_{11}(x) + \sum_{i=1}^{k} \lambda_i g^i{}_{11}(x) & \ldots & f_{1n}(x) + \sum_{i=1}^{k} \lambda_i g^i{}_{1n}(x) \\ \vdots & \ddots & \vdots \\ f_{n1}(x) + \sum_{i=1}^{k} \lambda_i g^i{}_{n1}(x) & \cdots & f_{nn}(x) + \sum_{i=1}^{k} \lambda_i g^i{}_{nn}(x) \end{pmatrix}$ is the second order derivative of L with respect to the x variables. Note that $D^2L(x,\lambda)$ is symmetric, so we may work with this quadratic form.

At a given $x^* \in R^n$, $Dg(x^*)_{k\times n} = \begin{pmatrix} Dg^1(x^*) \\ \vdots \\ Dg^k(x^*) \end{pmatrix}$. So the set of all vectors x that are orthogonal to all the gradient vectors of the constraint functions at x^* is the Null Space of $Dg(x^*)$, $N(Dg(x^*)) = \{x \in R^n \mid Dg(x^*)x = \theta_{k\times 1}\}$.

Theorem of Second Order Optimum:

Suppose there exists $(x^*_{n\times 1}, \lambda^*_{k\times 1})$ such that Rank $(Dg(x^*))=k$

and $Df(x^*) + \sum_{i=1}^{k} Dg^i(x^*) = \theta$.

i) A necessary condition: If f has a local maximum (respectively, local minimum) on S at point x^*, then $x^T D^2 L(x^*,\lambda^*)x \le 0$, (respectively ≥ 0) for all $x \in N(Dg(x^*))$.

ii) A sufficient condition: If $x^T D^2 L(x^*,\lambda^*)x < 0$, (respectively for >0), for all $x \in N(Dg(x^*))$, $x \neq \theta$, then x^* is a strict local maximum (respectively for strict local minimum) of f on S.

Note the theorem means that for maximum the second order condition should be that the bordered Hessian matrix of the problem is negative semi-definite (necessary) or negative definite (sufficient) given the FONC and for minimum it should be positive semi-definite (necessary) or positive definite (sufficient given the FONC), where the bordered Hessian is the matrix of second derivatives of $L(x, \lambda)$ with respect to x and λ.

Check Your Progress 1

1) Let $U(x, y) = x^2y$ and $p_x=2$, $p_y=1$. $I=200$. Find out x^* and y^*, if $I=p_xx+p_yy$ holds.

....

....

....

....

....

2) Let $U(x, y) = x^2+y$ and $p_x=2$, $p_y=1$. $I=200$. Find out x^* and y^*, if $I=p_xx+p_yy$ holds.

....

....

....

....

....

....

....

6.3 INEQUALITY CONSTRAINED OPTIMISATION

The problem is to find the maximum or the minimum of $f: R^n \to R$ on the set $\{x \in R^n \mid g^i(x) \geq 0, i = 1, ..., k\}$, where $g^i: R^n \to R$ are the k constraint functions. At the optimum, the constraints are now allowed to be binding (or tight or effective), i.e., $g^i(x) = 0$, as before, or slack (or non-binding), i.e., $g^i(x) > 0$.

Example: Max $U(x_1, x_2)$ s.t. $x_1 \geq 0$, $x_2 \geq 0$, $I - p_1x_1 - p_2x_2$. If we do not know whether $x_i = 0$, for some I, at the utility maximum, or whether $x_i > 0$, then clearly we cannot use the theorem of Lagrange. Similarly, if there is a bliss point, then we do not know in advance whether there at the budget constrained optimum the budget constraint is binding or slack. Again, we cannot then use the theorem of Lagrange, to use which we need to be assured that the constraint is binding.

Note the general nature of a constraint of the form $g^i(x) \geq 0$. If we have a constraint $h(x) \leq 0$, this is equivalent to $-h(x) \geq 0$. And something like $h(x) \leq c$ is equivalent to $c - h(x) \geq 0$.

We use Kuhn-Tucker Theory to address optimisation problems with inequality constraints. The main result is a first order necessary condition that is slightly different from that of the theorem of Lagrange; the main difference is that the first order condition $g^i(x) = 0$, $i = 1,\ldots,k$ in the theorem of Lagrange are replaced by the conditions $\lambda_i\, g^i(x) = 0$, $i = 1,\ldots,k$ in Kuhn-Tucker theory.

Kuhn-Tucker Theory:

Recall that the problem is to find the maximum or the minimum of $f: R^n \to R$ on the set $\{x \in R^n \mid g^i(x) \geq 0, i = 1, ..., k\}$, where $g^i: R^n \to R$ are the k constraint functions. The main theorem deals with local maxima and minima, though.

Suppose that l of the k constraints bind at the optimum x^*. Denote the corresponding constraint functions as $(g^i)_{i \in \Phi}$, where Φ is the set of indexes of the binding constraints. Let $g_\Phi: R^n \to R^l$ be the function whose l components are the constraint functions of the binding constraints. That is, $g_\Phi(x) = (g^i(x))_{i \in \Phi}$.

$$Dg_\Phi(x) = \begin{pmatrix} Dg^i(x) \\ \vdots \\ Dg^m(x) \end{pmatrix}$$, where i,…, m are the indexes of the binding constraints.

So, $Dg_\Phi(x)$ is a $l \times n$ matrix.

We now state FONC for the problem. The theorem below is a consolidation of the Fritz-John and the Kuhn-Tucker theorems.

The Kuhn-Tucker (KT) Theorem: Let

$f: R^n \to R$, and $g^i: R^n \to R$, i = 1,…, k

be once continuously differentiable functions. Suppose x^* is a maximum of f on the set

$S = U \cap \{x \in R^n \mid g^i(x) \geq 0, i = 1,...,k\}$, for some open set $U \subseteq R^n$.

Then there exist real numbers μ, λ_1^*, …, $.\lambda_k^*$, not all zero such that $\mu Df(x^*) + \sum_{i=1}^{k} \lambda_i^* Dg^i(x^*) = \theta_{1\times n}$.

Moreover, if $g^i(x^*) > 0$ for some i, then $\lambda_i^* = 0$. If in addition,

$\text{rank}Dg_{\Phi}(x^*) = 1$, then we may take μ to be equal to 1. Furthermore, $\lambda_i^* \geq 0$, i = 1,…, k and $\lambda_i^* > 0$ for some I implies $g^i(x^*) = 0$.

Suppose the constraint qualification, $\text{Rank}Dg_{\Phi}(x^*) = 1$, is met at the optimum. Then the KT equations are the following (n+k) equations in the n+k variables $x_1,\ldots, x_n, \lambda_1,\ldots, \lambda_k$.

$\lambda_i g^i(x^*) = 0, i = 1,...,k, \lambda_i^* \geq 0, g^i(x^*) \geq 0$ with complementary slackness ----(1)

$Df(x^*) + \sum_{i=1}^{k} \lambda_i^* Dg^i(x^*) = \theta$ ------------ (2)

If x^* is a local minimum of f on S, then $-f(x^*)$ attains a local maximum value. Thus for minimisation, while equation (1) stays the same, equation (2) changes to $-Df(x^*) + \sum_{i=1}^{k} \lambda_i^* Dg^i(x^*) = \theta$ ------------ (2')

Equation (1) and (2) are known as the Kuhn-Tucker conditions. Note finally that since the conditions of KT theorem are not sufficient conditions for local optima, there may be points that satisfy equations (1) and (2) or (2') without being local optima. This method does not require checking for concavity of objective functions and constraints, and does not require checking any second order condition.

Next, we give one useful theorem that gives some link between local and global maxima, where the crucial assumption is that the objective function is concave.

Theorem: Suppose $S \subseteq R^n$ is a convex set, and $f: S \to R$ is concave. Then

1) Any local maximum of f is a global maximum of f.
2) The set of maximizers of f on S, $\arg\max\{f(x) \mid x \in S\}$ is either empty or convex.

Check Your Progress 2

1) Consider the following utility maximisation problem, in which the utility function of the consumer is given by – $U(x, y) = x^{0.5}y^{0.5}$. The consumer income is I = 100 and it is given to him that $p_x = 2$ and $p_y = 3$. Find the utility maximising x and y when the budget constraint of the consumer may or may not bind. (Apply the Kuhn-Tucker method).

……………………………………………………………………………

……………………………………………………………………………

……………………………………………………………………………

....

....

....

2) In this problem you have to maximise $U(x, y) = (x - 3)(y - 2)$ subject to the following constraints –

   i) $x \geq 0$.

   ii) $y \geq 0$.

   iii) $5x+3y \leq 50$.

   ....

   ....

   ....

   ....

   ....

   ....

6.4 SIMPLE ECONOMIC PROBLEMS

In this section, we present two simple utility maximisation problems.

Example 1: In this problem our objective is to

maximise $U(x_1, x_2) = x_1 + x_2$,

over the set $\{x = (x_1, x_2) \in R^2 \mid x_1 \geq 0, x_2 \geq 0, I - p_1x_1 - p_1x_2 = 0\}$,

where $I > 0$, $p_1 > 0$ and $p_2 > 0$ are given.

So there are three inequality constraints:

$g^1(x_1, x_2) = x_1 \geq 0$, $g^2(x_1, x_2) = x_2 \geq 0$ and $g^3(x_1, x_2) = I - p_1x_1 - p_2x_2 \geq 0$

At the maximum x^*, any combination of these three could bind; so there are eight possibilities. However, since U is strictly increasing, the budget constraint binds at the maximum, therefore $g^3(x^*) = 0$. Moreover, $g^1(x^*) = g^2(x^*) = 0$ is not possible since consuming 0 of both goods gives utility equal to 0, which is clearly not a maximum.

Therefore, we have to check just three possibilities out of the eight.

Case (1) $g^1(x^*) > 0$, $g^2(x^*) > 0$, $g^3(x^*) = 0$

Case (2) $g^1(x^*) = 0$, $g^2(x^*) > 0$, $g^3(x^*) = 0$

Case (3) $g^1(x^*) > 0$, $g^2(x^*) = 0$, $g^3(x^*) = 0$

Before using the KT conditions, we verify that –

i) A global max exists (here, because the utility function is continuous and the budget set is compact).

ii) The CQ holds at all three relevant combinations of binding constraints described above.

Indeed, for Case (1), $Dg_\Phi(x) = Dg^3(x) = (-p_1, -p_2)$, so $\text{Rank}[Dg^3(x)] = 1$, so CQ holds. For Case (2), $Dg_\Phi(x) = \begin{pmatrix} Dg^1(x) \\ Dg^2(x) \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -p_1 & -p_2 \end{pmatrix}$, so $\text{Rank}[Dg_\Phi(x)] = 2$. For Case (3), $Dg_\Phi(x) = \begin{pmatrix} Dg^2(x) \\ Dg^3(x) \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -p_1 & -p_2 \end{pmatrix}$, so $\text{Rank}[Dg_\Phi(x)] = 2$.

Thus for the maximum, x^*, there exists a λ^* such that (x^*, λ^*) will be a solution to the KT FONCs. Of course, there could be other (x, λ)'s that are solutions as well, but a simple comparison of $U(x)$ for all candidate solutions will isolate for us the maximum.

$L(x, \lambda) = x_1 + x_2 + \lambda_1 x_1 + \lambda_2 x_2 + \lambda_3(I - p_1 x_1 - p_2 x_2)$

The KT conditions are –

$\lambda_1(\partial L / \partial \lambda_1) = \lambda_1 x_1 = 0, \lambda_1 \geq 0, x_1 \geq 0$, with CS ... (1)

$\lambda_2(\partial L / \partial \lambda_2) = \lambda_2 x_2 = 0, \lambda_2 \geq 0, x_2 \geq 0$, with CS ... (2)

$\lambda_3(\partial L / \partial \lambda_3) = \lambda_3(I - p_1 x_1 - p_2 x_2) = 0, \lambda_3 \geq 0, (I - p_1 x_1 - p_2 x_2) \geq 0$, with CS ...(3)

$(\partial L / \partial x_1) = 1 + \lambda_1 - \lambda_3 p_1 = 0$ ...(4)

$(\partial L / \partial x_2) = 1 + \lambda_2 - \lambda_3 p_2 = 0$ ...(5)

Since we don't know which of these three cases select the constraints that bind at the maximum, we must try all three.

Case 1: Since $x_1 > 0$, $x_2 > 0$, (1) and (2) imply $\lambda_1 = \lambda_2 = 0$. Plugging these in equation (4) and (5), we have $1 = \lambda_3^* p_1 = \lambda_3^* p_2$. Since utility is strictly increasing, relaxing the budget constraint will increase utility. So the marginal utility of income, $\lambda_3^* > 0$. Thus $\lambda_3^* p_1 = \lambda_3^* p_2$ implies $p_1 = p_2$. From equation (4) it can also be seen that $\lambda_3^* = 1/p_1 > 0$.

Therefore, if at a local maximum both x_1 and x_2 are strictly positive, then it must be that their prices are equal. All (x_1, x_2) that solve equation (3) are solutions. The utility in any such case equals $x_1 + ((I - p_1 x_1)/p_2) = I/p$, where $p = p_1 = p_2$. Note that in this case, $(U_1/p_1) = (U_2/p_2) = 1/p$.

Case 2: $x_1 = 0$ implies, from equation (3), that $x_2 = (I/p_2)$. Since this is greater than 0, equation (2) implies $\lambda_2 = 0$. Hence from equation (5), $\lambda_3^* p_2 = 1$. Since $\lambda_1^* \geq 0$, equation (4) and (5) imply $\lambda_3^* p_1 = 1 + \lambda_1 \geq 1 = \lambda_3^* p_2$. Moreover, since $\lambda_3^* > 0$, this implies $p_1 \geq p_2$.

That is, if it is the case that at the maximum, $x_1^* = 0$, $x_2^* > 0$, then it must be that $p_1 \geq p_2$. Note that is this case,

$(U_2/p_2) = (1/p_2) \geq (U_1/p_1) = (1/p_1)$.

For completeness sake, equation (5) implies $\lambda_3^* = (1/p_2)$. So from equation (4), $\lambda_1^* = (p_1/p_2) - 1$. So the unique critical point of $L(x, \lambda)$ is

$(x^*, \lambda^*) = (x_1^*, x_2^*, \lambda_1^*, \lambda_2^*, \lambda_3^*) = (0, (I/p_2), (p_1/p_2) - 1, 0, (1/p_2))$.

Case 3: This case is similar, and we get that $x_2^* = 0$, $x_1^* > 0$ occurs only if $p_1 \leq p_2$. We have $(x^*, \lambda^*) = ((I/p_2), 0, 0, (p_2/p_1) - 1, 1/p_1)$.

We see that which cases apply depends upon the price ratio (p_2/p_1). If $(p_2/p_1) = 1$, then all these three cases are relevant, and all $(x_1, x_2) \in R_+^2$ such that the budget constraint binds are utility maxima. But if $(p_2/p_1) > 1$, then only Case (2) is applies, and the solution to the KT conditions in that case is the utility maximum. Similarly, if $(p_2/p_1) < 1$, only Case (3) applies.

Example 2: Max $U(x_1, x_2) = (x_1/(1+x_1))+(x_2/(1+x_1))$, subject to $x_1 \geq 0$, $x_2 \geq 0$, $p_1x_1+p_2x_2 \leq I$.

Check that the indifference curves are downward sloping, convex and that they cut the axes. This last condition is due to the additive form of the utility function, and may result in 0 consumption of one of the goods at the utility maximum. Exactly as in Example 1, we are assured that a global max exists, that the CQ is met at the optimum, and that there are only three relevant cases of binding constraints to check.

The Kuhn-Tucker conditions are:

$$\lambda_1(\partial L/\partial\lambda_1) = \lambda_1 x_1 = 0, \lambda_1 \geq 0, x_1 \geq 0, \text{ with CS} \qquad \ldots(1)$$

$$\lambda_2(\partial L/\partial\lambda_2) = \lambda_2 x_2 = 0, \lambda_2 \geq 0, x_2 \geq 0, \text{ with CS} \qquad \ldots(2)$$

$$\lambda_3(\partial L/\partial\lambda_3) = \lambda_3(I - p_1x_1 - p_1x_1) = 0, \lambda_1 \geq 0, (I - p_1x_1 - p_1x_1) \geq 0, \text{ with CS} \qquad \ldots(3)$$

$$(\partial L/\partial x_1) = (\frac{1}{(1+x_1)^2}) + \lambda_1 - \lambda_3 p_1 = 0 \qquad \ldots(4)$$

$$(\partial L/\partial x_2) = (\frac{1}{(1+x_2)^2}) + \lambda_2 - \lambda_3 p_2 = 0 \qquad \ldots(5)$$

Case 1: $x_1 > 0$, $x_2 > 0$ implies $\lambda_1 = \lambda_2 = 0$. Equation (4) implies $\lambda_3 > 0$, so that equation (4) and (5) give $\left(\frac{(1+x_1)}{(1+x_2)}\right) = \left(\frac{p_1}{p_2}\right)^{\frac{1}{2}}$.

Using equation (3), which gives $x_2 = \left(\frac{(I - p_1x_1)}{p_2}\right)$, above, we get $\left(\frac{(p_2 + I - p_1x_1)}{(p_2(1+x_1))}\right) = \left(\frac{p_1}{p_2}\right)^{\frac{1}{2}}$, so simple consumptions yield

$$x_1 = \frac{\left(I + p_2 - (p_1p_2)^{\frac{1}{2}}\right)}{\left(p_1 + (p_1p_2)^{\frac{1}{2}}\right)}$$

$$x_2 = \frac{\left(I + p_1 - (p_1p_2)^{\frac{1}{2}}\right)}{\left(p_2 + (p_1p_2)^{\frac{1}{2}}\right)}$$

$$\lambda_3 = \frac{1}{p_1(1+x_1)^2} = \frac{1}{p_2(1+x_2)^2}$$

$x_1 > 0$, $x_2 > 0$ implies $I > (p_1p_2)^{\frac{1}{2}} - p_1, I > (p_1p_2)^{\frac{1}{2}} - p_2$. If either of these fails, then we are not in the regime of Case (1).

Case 2: $x_1 = 0$ with equation (3) implies $x_2 = \frac{I}{p_2}$. Since this is positive, $\lambda_2 = 0$, so equation (5) implies $\lambda_3 = \frac{1}{\left(1+\left(\frac{I}{p_2}\right)^2\right)} p_2$.

$\lambda_1 = \lambda_3 p_1 - 1$ (from $x_1 = 0$ and equation (4)). $\lambda_1 = \frac{p_1p_2}{(p_2+I)^2} - 1$. For this to be greater than zero it is required that $\frac{p_1p_2}{(p_2+I)^2} \geq 1$, that is $I \leq (p_1p_2)^{\frac{1}{2}} - p_2$.

Utility equals $\frac{x_2}{(1+x_2)} = \frac{I}{(p_2+I)}$.

$$\left(x_1^*, x_2^*, \lambda_1^*, \lambda_2^*, \lambda_3^*\right) = \left(0, \frac{I}{p_2}, \frac{p_1p_2}{(p_2+I)^2} - 1, 0, \frac{p_2}{(p_2+I)^2}\right).$$

Case (3): By symmetry, the solution is

$$\left(x_1^*, x_2^*, \lambda_1^*, \lambda_2^*, \lambda_3^*\right) = \left(\frac{I}{p_1}, 0, 0, \frac{p_1p_2}{(p_1+I)^2} - 1, \frac{p_1}{(p_1+I)^2}\right)$$

and for this case to hold it is necessary that $\frac{p_1p_2}{(p_1+I)^2} \geq 1$, or $I \leq (p_1p_2)^{\frac{1}{2}} - p_1$.

To summarise, suppose $p_1 = p_2 = p$, then $(p_1p_2)^{1/2} - p_1 = (p_1p_2)^{1/2} - p_2 = 0$. So since $I > 0$, we are in the regime of Case I, and $x_1 = x_2 = (I/2p)$ at the maximum. Suppose on the other hand that $p_1 < p_2$ (the contrary can be worked out similarly). Then $p_2 > (p_1p_2)^{1/2} > p_1$. Thus earlier $I > (p_1p_2)^{1/2} - p_1$, in which case we use Case (1), or $I \leq (p_1p_2)^{1/2} - p_1$ in which case we use Case (3). Case (2), that in which a positive amount of good 2 and zero amount of good 1 is consumed at the maximum, does not apply.

6.5 LET US SUM UP

In this unit, we emphasise the basic theory of constrained optimisation, with its division into two main parts viz., constrained optimisation in case of equality constraints and constrained optimisation in case of inequality constraints. We consider two different techniques for solving these two different types of problems. In case of optimisation under equality constraints, we apply Lagrangean method to solve those problems, and in the later case, that is, in the case for optimisation under inequality constraints, we apply the Kuhn-Tucker theory. We saw that this theory is different in many respects. First of all, the first order conditions of Lagrangean theory and the Kuhn-Tucker theory are different. Secondly, in Kuhn-Tucker theory there is no need to check the second order condition, while in case of Lagrangean case one must check it to ensure whether the optimum value (that we get from first order condition) is maximum or minimum. Finally, in Kuhn-Tucker theory there is a new set of conditions called "complementary slackness condition", which does not exist in Lagrangean method. It can be noted that if objective function is concave and constraint set is convex, the local maxima is same as global maxima, which proves that equality of local and global maxima under certain conditions. For the second order condition of Lagrangean method one can notice that the necessary second order condition for maxima is that the corresponding bordered Hessian matrix should be negative semi-definite and the sufficient second order condition for maxima is that the same matrix should be negative definite. For the case of minima, the necessary second order condition is that the bordered Hessian matrix should be positive semi-definite and the sufficient second order condition is that the same matrix is positive definite. We then present two examples of optimisation under inequality constraints with increasing difficulty. We solve both the problems using Kuhn-Tucker method. For this method, it can be seen that for any optimisation problem if the number of variables is n and the number of constraint is k, then the first order condition for optimisation should have 2(n+k) inequality and (n+k) equality and the first order condition for maxima and minima are different in Kuhn-Tucker theory. Therefore, there is no need to check second order condition.

6.6 KEY WORDS

Theorem of Lagrange:

Let $f: R^n \to R$ and $g^1: R^n \to R, g^2: R^n \to R, \ldots, g^k: R^n \to R$, k < n be once continuously differentiable functions. Suppose x^* is a Max or a Min of f on the set $S = U \cap \{x \in R^n \mid g^i(x) = 0, i = 1, \ldots, k\}$, for some open set $U \subseteq R^n$. Then there exist real numbers $\mu^*, \lambda_1^*, \ldots, \lambda_k^*$, not all zero, such that

$$\mu Df(x^*) + \sum_{i=1}^{k} \lambda_i^* Dg^i(x^*) = \theta_{1\times n}$$

Moreover if rank($Dg(x^*)$) = k, then we may put $\mu = 1$.

Theorem of Second Order Optimum:

Suppose there exists $(x^*_{n\times 1}, \lambda^*_{k\times 1})$ such that

rank($Dg(x^*)$) = k and $Df(x^*) + \sum_{i=1}^{k} Dg^i(x^*) = \theta$.

i) A necessary condition: If f has a local maximum (respectively for local minimum) on S at point x^*, then $x^T D^2 L(x^*, \lambda^*) x \le 0$, (respectively for ≥ 0) for all $x \in N(Dg(x^*))$.

ii) A sufficient condition: If $x^T D^2 L(x^*, \lambda^*) x < 0$, (respectively for >0), for all $x \in N(Dg(x^*))$, $x \neq \theta$, then x^* is a strict local maximum (respectively for strict local minimum) of f on S.

The Kuhn-Tucker (KT) Theorem:

Let $f: R^n \to R$, and $g^i: R^n \to R$, i = 1,..., k be once continuously differentiable functions. Suppose x^* is a maximum of f on the set $S = U \cap \{x \in R^n \mid g^i(x) \ge 0, i = 1, \ldots, k\}$, for some open set $U \subseteq R^n$. Then there exist real numbers μ, λ_1^*, ..., λ_k^*, not all zero such that $\mu Df(x^*) + \sum_{i=1}^{k} \lambda_i^* Dg^i(x^*) = \theta_{1\times n}$.

Moreover, if $g^i(x^*) > 0$ for some i, then $\lambda_i^* = 0$. If in addition, Rank$Dg_{\Phi}(x^*) = 1$, then we may take μ to be equal to 1. Furthermore, $\lambda_i^* \ge 0$, i = 1,..., k, and $\lambda_i^* > 0$ for some I implies $g^i(x^*) = 0$.

6.7 SOME USEFUL BOOKS

Henderson, James M. and Quandt, Richard E. (1980), *Microeconomic Theory*, McGrow-Hill Book Company, New York.

Allen, R.G.D. (1938), *Mathematical Analysis for Economists*, St. Martin's Press, New York.

Varian, Hal (1992), *Microeconomic Analysis*, W.W. Norton & Company, Inc., New York.

6.8 ANSWERS OR HINTS TO CHECK YOUR PROGRESS EXERCISES

Check Your Progress 1

1) The Lagrange function of the problem: $L(x,\lambda) = x^2y + \lambda\,[200 - 2x - y]$

   The FONC for this problem are:

   $2xy = 2\lambda$ -(1), $x^2 = \lambda$ -(2) and $200 = 2x + y$ -(3). Solving these three equations, we get $x^* = (200/3)$, $y^* = (200/3)$ and $\lambda^* = (200/3)^2$.

   Note that CQ satisfied because $p_x > 0$ and $p_y > 0$.

   The second order condition of this problem is –

   $|\overline{H}| = \begin{pmatrix} 0 & -2 & -1 \\ -2 & 2(200/3) & 2(200/3) \\ -1 & 2(200/3) & 0 \end{pmatrix}$. One can check that this matrix is negative definite. Therefore, second order condition fulfilled.

2) The problem can be solved exactly the same way as done in the previous one.

Check Your Progress 2

1) See the method of solving optimisation problem under inequality constraint in Section 6.3. You should able to get $x^* = 25$ and $y^* = (50/3)$.

   Note that Lagrangean function is

   $L(x, y, \lambda) = x^{0.5}y^{0.5} + \lambda\,[100 - 2x - 3x]$

   The first order Kuhn-Tucker conditions are:

   $0.5x^{-0.5}y^{0.5} - 2\lambda \le 0,\ x \ge 0,\ x[0.5x^{-0.5}y^{0.5} - 2\lambda] = 0$

   $0.5x^{0.5}y^{-0.5} - 3\lambda \le 0,\ y \ge 0,\ y[0.5x^{0.5}y^{-0.5} - 3\lambda] = 0$

   $100 - 2x - 3x \le 0,\ \lambda \ge 0,\ \lambda[100 - 2x - 3x]$

2) See the method of solving optimisation problem under inequality constraint in Section 6.3. You should able to get $x^* = 6.1$ and $y^* = 6.5$.

   Note that the Lagrangean function is

   $L(x, y, \lambda) = (x - 3)(y - 2) + \lambda\,[50 - 5x - 3y]$

   The first order Kuhn-Tucker conditions are:

   $(y - 2) - 5\lambda \le 0,\ x \ge 0,\ x[(y - 2) - 5\lambda] = 0$

   $(x - 3) - 3\lambda \le 0,\ y \ge 0,\ y[(x - 3) - 3\lambda] = 0$

   $[50 - 5x - 3y] \ge 0,\ \lambda \ge 0,\ \lambda[50 - 5x - 3y] = 0$

NOTES

MPDD/IGNOU/P.O.7K/November 2018 (Reprint)

ISBN : 978-93-86375-13-1