UNIT 7 MAXIMA AND MINIMA

Structure Page No.

7.1 Introduction 51
   Objectives 51
7.2 Local Maxima and Local Minima 52
7.3 Lagrange Multiplier 59
7.4 Summary 62
7.5 Hints/Solutions 63
7.6 Appendix 65

7.1 INTRODUCTION

In this unit, we discuss maxima and minima for real-valued functions of vector variables. You are already familiar with this concept from your undergraduate Real Analysis and Calculus courses. There you have seen that these concepts are studied locally also and they are called local maxima or local minima; together they are called local extrema. The extension of these concepts to the vector variable case helps to solve many real-life problems arising in economics, finance and other fields.

In Sec. 7.2, we shall discuss a necessary condition for a function to have local extrema in terms its partial derivatives. Then we shall discuss a sufficient condition for the existence of local extrema by using Taylor’s theorem for real valued functions of vector variables.

One of the main applications of the concept of maxima and minima is to solve optimization problems arising in economics such as expenditure minimization problem, profit maximization problem, utility maximization problem. Most of these problems are concerned with maximizing and minimizing real-valued n-variable function called objective function and there are some constraints also attached with the problem which are again represented as a functional relationship. Such problems can be solved by a method called Lagrange Multiplier method. In Sec. 7.3, we discuss this method. We shall briefly explain the utility of this method by giving a practical problem in optimization. The understanding of the method requires some techniques in linear algebra such as quadratic forms and the related matrix theory. You are advised to look into any standard book on Linear algebra that are available at your programme study centre or the Block 4 of IGNOU course on Linear Algebra with the Code MTE-02 titled Inner Products and Quadrics which is also available at your programme centre.

Objectives

After studying this unit, you should be able to

- define critical points, stationary points, saddle points, local maxima and local minima;
- state a necessary condition for functions to have local extrema and apply it;
- state and prove the theorem known as “second derivative test” which gives a sufficient condition for finding local maxima and minima;
Calculus in \( \mathbb{R}^n \)

- use Hessian for classifying local maxima and local minima; and
- apply Lagrange’s multiplier method for finding the stationary points when the variables are subject to some constraints.

### 7.2 LOCAL MAXIMA AND LOCAL MINIMA

This section deals with the concept of maxima and minima (or extrema) for real-valued functions of vector variables. You are already familiar with this concept for the one-variable case. In the case of one-variable you might have studied that there are functions which do not have an extrema at a point with respect to the whole domain whereas the functions have extrema at that point locally. These points are called local extrema. In this section we shall take up the study of local extrema for functions from \( \mathbb{R}^n \) to \( \mathbb{R} \).

**Definition 1:** Let \( f : E \subseteq \mathbb{R}^n \to \mathbb{R} \) be a function. A point \( a \in E \) is called a maximum point w.r.t. \( E \), if \( f(x) \leq f(a) \quad \forall x \in E \). A point \( a \in E \) is called a minimum point w.r.t. \( E \) if \( f(x) \geq f(a) \quad \forall x \in E \).

If a point \( a \in E \) is either maximum or minimum point w.r.t. \( E \), then that point is called an extreme point or point of extrema.

Now we define local extrema.

**Definition 2:** Let \( f : E \subseteq \mathbb{R}^n \to \mathbb{R} \) be a function where \( E \) is an open subset of \( \mathbb{R}^n \). A point \( a \in E \) is said to be a local maximum for \( f \) if there exists a neighbourhood \( E_a \) of \( a \) such that \( f(x) \leq f(a) \) for all \( x \in E_a \).

A local minima is similarly defined.

**Example 1:** Let us consider the function given by

\[
f(x, y) = (x + 1)^2 + (y - 3)^2 - 1.
\]

We first note that \( f(-1, 3) = -1 \).

Also \( f(x, y) \geq f(-1, 3) \) for all \( (x, y) \in \mathbb{R}^2 \).

This show that the function has a minimum at \( (-1, 3) \) and the minimum value is \( f(-1, 3) = -1 \). (See Fig. 1)

---

Fig. 1

***
Example 2: Let us consider the function

\[ f(x, y) = \frac{1}{2} - \sin(x^2 + y^2) \]

Here \( f(0, 0) = \frac{1}{2} \). Let us consider the neighbourhood \( U \) of \((0, 0)\) given by

\[ U = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 < \frac{\pi}{6} \} \]

Then for any \((x, y) \in U\), we have

\[ \sin(x^2 + y^2) > 0 \]

and therefore

\[ f(x, y) = \frac{1}{2} - \sin(x^2 + y^2) < \frac{1}{2} = f(0, 0). \]

Thus \( f(x, y) \leq f(0, 0) \) for all \((x, y) \in U\) in the disc. Note that \( f(x, y) \) can be greater than \( \frac{1}{2} \) for \((x, y) \notin U\). Hence \( f \) has a local minimum at \((0, 0)\).

***

Example 3: Let us consider another function given by

\[ f(x, y) = 1 + \sqrt{x^2 + y^2} \]

If we look at the graph of the function given below, then we can see that \( f \) has a minimum at \((0, 0)\).

If you closely look at the above examples, then you can notice that in the case of Examples 1 and 2 we have

\[ \frac{\partial f}{\partial x}(x_0, y_0) = 0 = \frac{\partial f}{\partial y}(x_0, y_0) \]

where \((x_0, y_0)\) is a point of extrema. Whereas we notice that in the case of Example 3, this is not the case as function does not have any first order partial derivatives at \((0, 0)\).

Now we state a result which shows that if all the first order partial derivatives of \( f : E \subseteq \mathbb{R}^n \rightarrow \mathbb{R} \) exists at a point \( a \in E \) where \( E \) is an open set, then they necessarily vanish at the points of extrema.

**Theorem 1:** Let \( f : E \subseteq \mathbb{R}^n \rightarrow \mathbb{R} \) be a function where \( E \) is an open subset of \( \mathbb{R}^n \). Suppose that all the first order partial derivatives of the function \( f \) exists at a point \( a \in E \). Then a necessary condition for the function to have a local extremum at the point \( a \) is that \( \frac{\partial f}{\partial x_i}(a) = 0 \) for \( i = 1, \ldots, n \).
Calculus in $\mathbb{R}^n$

**Proof:** Suppose that $f$ has a local extrema at the point $a = (a_1, a_2, \ldots, a_n)$.

Let us consider the real-valued function $\phi$ defined by

$$\phi(t) = f(t, a_2, \ldots, a_n).$$

Since $a$ is an extreme point of $f$, we get that $a_1$ is extreme for $\phi$. Then from the one-variable case you know that

$$\phi'(a_1) = \frac{\partial f}{\partial x_1}(a_1, a_2, \ldots, a_n) = 0$$

In this way we can show that $\frac{\partial f}{\partial x_i}(a_1, \ldots, a_n) = 0$ for each $i = 1, \ldots, n$.

Hence the result. $\square$

Now we make the following definition:

**Definition 3:** Let $f : E \subset \mathbb{R}^n \to \mathbb{R}$ be a function. A point $a \in E$ is called a critical point of $f$ if either

i) the partial derivatives of $f$ do not exist at $a$, or

ii) $\frac{\partial f}{\partial x_i}(a) = 0$ for $i = 1, \ldots, n$.

The points for which the condition (ii) is satisfied are called **stationary points**.

You may recall that all stationary points of a function need not be its point of local extrema. Such points are called saddle points. Note that a point $a \in E$ is called a saddle point if every neighbourhood $E_a$ of $a$ contains points $x \in E$ such that $f(x) > a$ and other points $y \in E_a$ such that $y < f(a)$.

Let us consider an example.

**Example 4:** Let us consider the function $f : \mathbb{R}^2 \to \mathbb{R}$ given by

$$f(x, y) = (y - x^2)(y - 2x^2).$$

Here we have $\frac{\partial f}{\partial x}(0, 0) = \frac{\partial f}{\partial y}(0, 0) = 0$. Thus, $(0, 0)$ is a stationary point. Now the graph of the function $f$ given below shows that $(0, 0)$ is not a point of local extrema. Note that the function $f$ assumes both positive and negative values in every neighbourhood of $(0, 0)$. Therefore $(0, 0)$ is a saddle point for the function $f$.  

![Fig. 3](image)

---

54
Next we discuss a sufficient condition in terms of second order partial derivatives to check whether a point is an extremum point.

**Theorem 2: (Second-derivative test for extrema):** Let \( f : E \to \mathbb{R} \) be a function defined on an open set \( E \subset \mathbb{R}^n \). Assume that the second-order partial derivatives \( D_{ij}f \) exist in an \( n \)-ball \( B(a) \) and are continuous at \( a \in \mathbb{R}^n \), where \( a \) is a stationary point of \( f \). Let

\[
Q(x) = \frac{1}{2} f''(a; x) = \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} D_{ij}f(a)x_i x_j
\]

where \( x = (x_1, \ldots, x_n) \). Then

- If \( Q(x) > 0 \) for all \( x \neq 0 \), \( f \) has a relative minimum at \( a \).
- If \( Q(x) < 0 \) for all \( x \neq 0 \), \( f \) has a relative maximum at \( a \).
- If \( Q(x) \) takes both positive and negative values, then \( f \) has a saddle point at \( a \).

**Proof:** We first apply Taylor's theorem to the function \( f \). Taking \( m = 2 \) and \( b = a + x \) in Taylor's theorem (see Sec. 5.4, Unit 5), we get that there exists a \( z \) which lies on the line segment joining \( a \) and \( a + x \) such that

\[
f(a + x) - f(a) = f'(a)x + \frac{1}{2} f''(z, x)
\]

and

\[
f''(z, x) = \sum_{i=1}^{n} \sum_{j=1}^{n} D_{ij}f(z)x_i x_j
\]

Since \( a \) is a stationary point, we have \( f'(a) = 0 \). Therefore Equation (2) becomes

\[
f(a + x) - f(a) = \frac{1}{2} f''(z, x)
\]

Therefore as \( a + x \) ranges over \( B(a) \), the algebraic sign of \( f(a + x) - f(a) \) is determined by that of \( f''(z; x) \). We can write Equation (2) in the form

\[
f(a + x) - f(a) = \frac{1}{2} f''(a, x) + \|x\|^2 E(x)
\]

where

\[
\|x\|^2 E(x) = \frac{1}{2} f''(z, x) - \frac{1}{2} f''(a; x)
\]

Substituting for \( f''(z, x) \) and \( f''(a; x) \), we get that

\[
\|x\|^2 |E(x)| \leq \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} |D_{ij}f(z) - D_{ij}f(a)| \|x\|^2
\]

Since the second order partial derivatives of \( f \) are continuous at \( a \) we get that \( E(x) \to 0 \) as \( x \to 0 \).

Now we rewrite Equation (3) in the form

\[
f(a + x) - f(a) = Q(x) + \|x\|^2 E(x)
\]
where \( Q(x) \) is as given in Equation (1).

The function \( Q \) is continuous at each point \( x \) in \( \mathbb{R}^n \). Let \( S = \{ x : \|x\| = 1 \} \) denotes the boundary of the \( n \)-ball \( B(0; 1) \). (Recall that we defined the norm function \( \| \cdot \| \) in Unit 5). If \( Q(x) > 0 \) for all \( x \neq 0 \), then \( Q(t) \) is positive on \( S \). Since \( S \) is compact, \( Q \) has a minimum on \( S \). Let us call it \( m \). Then \( m > 0 \).

Now \( Q(cx) = c^2Q(x) \) for every real number \( c \). Taking \( c = 1/\|x\| \) where \( x \neq 0 \) we see that \( cx \in S \) and hence \( c^2Q(x) \geq m \), so \( Q(x) \geq m\|x\|^2 \). Using this in (A0) we find

\[
f(a + x) - f(a) = Q(x) + \|x\|^2E(x) \geq m\|x\|^2 + \|x\|^2E(x)
\]

Since \( E(x) \to 0 \) as \( x \to 0 \), there is a positive number \( r \) such that \( |E(x)| < \frac{1}{2}m \) whenever \( 0 < \|x\| < r \). For such \( x \) we have \( 0 \leq \|x\|^2|E(x)| < \frac{1}{2}m\|x\|^2 \), so

\[
f(a + x) - f(a) > m\|x\|^2 - \frac{1}{2}m\|x\|^2 = \frac{1}{2}m\|x\|^2 > 0.
\]

Therefore \( f \) has a relative minimum at \( a \), which proves (a).

To prove (b) we use a similar argument, or apply part (a) to the function \(-f\).

Finally, we prove (c). For each \( \lambda > 0 \) we have, from (A0).

\[
f(a + \lambda x) - f(a) = \lambda Q(\lambda x) + \lambda^2\|x\|^2E(\lambda x) = \lambda^2\{Q(x) + \|x\|^2E(\lambda x)\}
\]

Suppose \( Q(x) \neq 0 \) for some \( x \). Since \( E(y) \to 0 \) as \( y \to 0 \), there is a positive \( r \) such that

\[
\|x\|^2E(\lambda x) < \frac{1}{2}|Q(x)| \quad \text{if} \quad 0 < \lambda < r.
\]

Therefore, for each such \( \lambda \) the quantity \( \lambda^2\{Q(x) + \|x\|^2E(\lambda x)\} \) has the same sign as \( Q(x) \). Therefore, if \( 0 < \lambda < r \), the difference \( f(a + \lambda x) - f(a) \) has the same sign as \( Q(x) \). Hence, if \( Q(x) \) takes both positive and negative values, it follows that \( f \) has a saddle point at \( a \).

\[\square\]

Note: A real-valued function \( Q \) defined on \( \mathbb{R}^n \) by an equation of the type

\[
Q(x) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}x_i x_j,
\]

where \( x = (x_1, \ldots, x_n) \) and the \( a_{ij} \) are real is called a quadratic form. The form is called symmetric if \( a_{ij} = a_{ji} \) for all \( i \) and \( j \) and is called positive definite if \( x \neq 0 \) implies \( Q(x) > 0 \), and negative definite if \( x \neq 0 \) implies \( Q(x) < 0 \).

You might be already familiar with quadratic forms from your undergraduate Linear algebra course. (You can refer to IGNOU course MTE-02, Block 4)

In general, it is not easy to determine whether a quadratic form is positive or negative definite. One criterion, involving determinants, can be described as follows. Let \( \Delta = \text{determinant of the matrix } [a_{ij}] \) and let \( \Delta_k \) denote the determinant of the \( k \times k \) matrix obtained by deleting the last \( (n - k) \) rows and columns of the matrix \([a_{ij}]\).
Also, put $\Delta_0 = 1$. From the theory of quadratic forms it is known that a necessary and sufficient condition for a symmetric form to be positive definite is that the $n + 1$ numbers $\Delta_0, \Delta_1, \ldots, \Delta_n$ be positive. The form is negative definite if and only if, the same $n + 1$ numbers are alternately positive and negative. The quadratic form which appears in Equation (1) is symmetric because the mixed partials $D_{ij} f(a)$ and $D_{ji} f(a)$ are equal. Therefore, under the conditions of Theorem 2, we see that $f$ has a local minimum at $a$ if the $(n + 1)$ numbers $\Delta_0, \Delta_1, \ldots, \Delta_n$ of the corresponding Jacobian matrix for $f$ are all positive, and a local maximum if these numbers are alternately positive and negative.

We have the following result:

**Theorem 3:** If $f : E \subset \mathbb{R}^n \to \mathbb{R}$, $E$ open in $\mathbb{R}^n$, has continuous first and second-order partial derivatives at $a$ where $a$ is a critical point of $f$, and $Hf$ is the Hessian of $f$ at $a$ (refer Unit 5 where we have defined the Hessian of a function) and $\Delta_k$ denote the determinant of $k \times k$ matrix obtained by deleting the last $(n - k)$ rows and column of the matrix. Then the following hold:

a) if $\Delta_{2k} < 0$ for some $k$ then $a$ is a saddle point of $f$,

b) if $\Delta_n \neq 0$ then

   (b1) $f$ has a local minimum at $a$ if and only if $\Delta_k > 0$ for all $k$,

   (b2) $f$ has a local maximum at $a$ if and only if $(-1)^k \Delta_k > 0$ for all $k$,

   c) if $\Delta_n = 0$ we call it a degenerate case and the test cannot be applied.

The case $n = 2$ can be handled directly and gives the following criterion.

**Theorem 4:** Let $f$ be a real-valued function with continuous second-order partial derivatives at a stationary point $a$ in $\mathbb{R}^2$. Let

$$ A = D_{1,1} f(a), \quad B = D_{1,2} f(a), \quad C = D_{2,2} f(a). $$

and let

$$ \Delta = \det \begin{bmatrix} A & B \\ B & C \end{bmatrix} = AC - B^2. $$

Then we have:

a) If $\Delta > 0$ and $A > 0$, $f$ has a local minimum at $a$.

b) If $\Delta > 0$ and $A < 0$, $f$ has a local maximum at $a$.

c) If $\Delta < 0$, $f$ has a saddle point at $a$. 

57
Calculus in $\mathbb{R}^n$

Note: You may recall that the conditions given in the theorem above resembles that of the one-variable case.

If $\Delta = 0$, then the test fails. Let us consider some examples.

**Example 5:** Let us check the following function $f : \mathbb{R}^3 \rightarrow \mathbb{R}$ for local extrema.

Let $f(x, y, z) = x^2y^2 + z^2 + 2x - 4y + z$.

We first note that

$$\nabla f(x, y, z) = (2xy^2 + 2, 2x^2y - 4, 2z + 1).$$

If $a$ is a critical point of $f$, then $a$ should satisfy the

$$2xy^2 + 2 = 0$$
$$2x^2y - 4 = 0$$
$$2z + 1 = 0$$

We solve these equations. From the third equation we get that $z = -1/2$. From the first two equations we see that $x$ and $y$ are non-zero. Hence $xy^2 = -1$ and $x^2y = 2$ imply $xy^2/x^2y = -1/2 = y/x$ and $x = -2y$. We have $-2y \cdot y^2 = -1$, i.e. $y^3 = 1/2$ and $y = 2^{-1/3}$. From $x = -2y$ we obtain $x = -2^{2/3}$ and conclude that $(-2^{2/3}, 2^{-1/3}, -1/2)$ is the only critical point of $f$.

A simple calculation shows

$$J_f = \begin{bmatrix} 2y^2 & 4xy & 0 \\ 4xy & 2x^2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

and

$$J_f(-2^{2/3}, 2^{-1/3}, 1/2) = \begin{bmatrix} 2^{1/3} & -4.2^{1/3} & 0 \\ -4.2^{1/3} & 2.2^{4/3} & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

Note $\Delta_1 = 2^{1/3} > 0$ and

$$\Delta_2 = \begin{vmatrix} 2^{1/3} & -4.2^{1/3} \\ -4.2^{1/3} & 2.2^{4/3} \end{vmatrix} = 2.2^{5/3} - 16.2^{2/3}$$

$$= 4.2^{2/3} - 16.2^{2/3} < 0$$

Therefore by Theorem 3, the critical point $(-2^{2/3}, 2^{-1/3}, -1/2)$ is a saddle point of $f$.

***

Your can try some exercises now.

---

E1) Find the critical points of $f(x, y, z) = x^2y + y^2z + z^2 - 2x$ and check whether they are extreme points.

In the next section we shall discuss the problem of finding maxima or minima subject to certain constraints.
Here we start with a practical situation.

Suppose a person consumes \( n \) commodities in nonnegative quantities. Then her utility from consuming \( x_i \geq 0 \) units of commodity \( i \) (\( i = 1, \ldots, n \)) is given by \( u(x_1, \ldots, x_n) \), where \( u : \mathbb{R}^n_+ \to \mathbb{R} \). Suppose she has an income of \( I \geq 0 \), and faces the price vector \( p = (p_1, \ldots, p_n) \), where \( p_i \geq 0 \) denotes the unit price of the \( i \)-th commodity. Her budget set (i.e., the set of feasible or affordable consumption bundles, given her income \( I \) and the prices \( p \)) is denoted \( B(p, I) \), and is given by

\[
B(p, I) = \{ x \in \mathbb{R}^n_+ | \sum_{i=1}^{n} p_i x_i \leq I \}
\]

Then her objective is to maximize the level of her utility over the set of affordable commodity bundles, i.e., to solve:

Maximize \( u(x) \) subject to \( x \in B(p, I) \).

There are many situations like this where the values of a given function \( f : \mathbb{R}^n \to \mathbb{R} \) are to be maximized or minimized over a given set \( E \subseteq \mathbb{R}^n \). Here we shall discuss a method for solving such problems that is developed by the mathematician Lagrange.

Let us consider another problem.

Suppose that \( f(x, y, z) \) represents the temperature at the point \((x, y, z)\) in space and we ask for the maximum or minimum value of the temperature on a certain surface. If the equation of the surface is given explicitly in the form \( z = h(x, y) \), then in the expression for \( f(x, y, z) \) we can replace \( z \) by \( h(x, y) \) to obtain the temperature on the surface as a function of \( x \) and \( y \) alone, say \( F(x, y) = f(x, y, h(x, y)) \). The problem is then reduced to finding the extreme value of \( F \).

However, in practice, certain difficulties arise. The equation of the surface might be given in an implicit form, say \( g(x, y, z) = 0 \) and it may be impossible, in practice, to solve this equation explicitly for \( z \) in terms of \( x \) and \( y \), or even for \( x \) or \( y \) in terms of the remaining variables. The problem might be further complicated by asking for the extreme values of the temperature at those points which lie on a given curve in space. Such a curve can be the intersection of two surfaces, say \( g_1(x, y, z) = 0 \) and \( g_2(x, y, z) = 0 \). If we could solve these two equations simultaneously, say for \( x \) and \( y \) in terms of \( z \), then we could introduce these expressions into \( f \) and obtain a new function of \( z \) alone, whose extrema we would then seek. In general, however, this procedure cannot be carried out and a more practicable method need to be sought. An elegant and useful method for attacking such problems was developed by Lagrange. The validity of the method is established by the Implicit Function Theorem which we described in Unit 6.

Lagrange's method provides a necessary condition for a point to be an extreme point which we shall explain now.

Let \( f : E \subseteq \mathbb{R}^n \to \mathbb{R}, E \subseteq \mathbb{R}^n \) an open set, be a function whose extreme values are sought when the variables are restricted by a certain number of side conditions, say \( g_1(x_1, \ldots, x_n) = 0, \ldots, g_m(x_1, \ldots, x_n) = 0 \).
We first form the linear combination

\[ L(x_1, \ldots, x_n) = f(x_1, \ldots, x_n) - \lambda_1 g_1(x_1, \ldots, x_n) - \cdots - \lambda_m g_m(x_1, \ldots, x_n), \] (6)

where \( \lambda_1, \ldots, \lambda_m \) are \( m \) constants. We then differentiate \( \phi \) with respect to each coordinate and consider the following system of \( n + m \) equations:

\[ \frac{\partial L}{\partial x_i} = 0, \quad i = 1, 2, \ldots, n, \] (7)

\[ g_k(x_1, \ldots, x_n) = 0, \quad k = 1, 2, \ldots, m. \] (8)

Lagrange, by his method, proved that if the point \((x_1, x_2, \ldots, x_n)\) is a point of extrema for \( f \), then it will also satisfy this system of \((n + m)\) equations. In practice we solve the “\( n + m \)” unknowns \( \lambda_1, \lambda_2, \ldots, \lambda_m \) and \( x_1, x_2, \ldots, x_n \). The point \((x_1, x_2, \ldots, x_n)\) so obtained is a stationary point. According to the Lagrange’s theorem this point can then be tested for maximum or minimum point by the already known methods.

The numbers \( \lambda_1, \lambda_2, \ldots, \lambda_m \) are introduced only to help to solve the system for \( x_1, x_2, \ldots, x_n \), and they are called Lagrange’s multipliers. One multiplier is introduced for each side condition. The function \( L \) in Equation (6) is called the Lagrangian function. Equations (7) and (8) are called Lagrangian Equations.

Now we state the Lagrange’s theorem, the proof of which involves implicit function theorem. We omit the proof here.

**Theorem 5:** Let \( f : E \subset \mathbb{R}^n \to \mathbb{R} \), \( E \) an open set in \( \mathbb{R}^n \), be such that the partial derivatives of \( f \) exists and are continuous on \( E \). Let \( g_1, \ldots, g_m \) be \( m \) real-valued functions defined on \( E \) such that partial derivatives of \( g_i \) exists and are continuous on \( E \) for \( i = 1, \ldots, m \). Let us assume that \( m < n \). Let \( X_0 \) be that subset of \( E \) on which each \( g_i \) vanishes for \( i = 1, \ldots, m \), that is,

\[ X_0 = \{ x \in E, g_i(x) = 0 \text{ for } i = 1, \ldots, m \}. \]

Assume that \( x_0 \in X_0 \) and assume that there exists a ball \( B(x_0) \) in \( \mathbb{R}^n \) such that \( f(x) \leq f(x_0) \) for all \( x \in X_0 \cap B(x_0) \) or such that \( f(x) \geq f(x_0) \) for all \( x \in X_0 \cap B(x_0) \). Assume also that the \( m \)-rowed determinant \( \det[Dg_i(x_0)] \neq 0 \).

Then there exist \( m \) real numbers \( \lambda_1, \ldots, \lambda_m \) such that they satisfy following \( n \) equations:

\[ \frac{\partial f}{\partial x_i}(x_0) - \sum_{k=1}^{m} \lambda_k \frac{\partial g_k}{\partial x_i}(x_0) = 0 \quad (i = 1, 2, \ldots, n). \] (9)

So the solution of an extremum problem by Lagrange’s method involves the following step:

**Step 1:** Form the Lagrangian function given in Equation (6)

**Step 2:** Form the Lagrangian equations given in Equations (7) and (8). The solution thus obtained is a stationary point.

**Step 3:** Check the stationary point for extrema by the methods already discussed in Sec. 7.2.
Here we state a sufficient condition for checking extrema when we have a single constraint. In this case the Equation 6 reduces to

$$L(x_1, \ldots, x_n) = f(x_1, \ldots, x_n) - \lambda g(x_1, \ldots, x_n).$$

(10)

To check that the stationary point obtained by Lagrange method is local maximum or local minimum, we need to compute the value of \( n - 1 \) principal minors of the following determinant

$$
\Delta_{n+1} =
\begin{vmatrix}
0 & \frac{\partial g}{\partial x_1} & \frac{\partial g}{\partial x_2} & \cdots & \frac{\partial g}{\partial x_n} \\
\frac{\partial g}{\partial x_1} & \frac{\partial^2 f}{\partial x_1^2} - \lambda \frac{\partial^2 g}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} - \lambda \frac{\partial^2 g}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} - \lambda \frac{\partial^2 g}{\partial x_1 \partial x_n} \\
\frac{\partial g}{\partial x_2} & \frac{\partial^2 f}{\partial x_2 \partial x_1} - \lambda \frac{\partial^2 g}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} - \lambda \frac{\partial^2 g}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n} - \lambda \frac{\partial^2 g}{\partial x_2 \partial x_n} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
\frac{\partial g}{\partial x_n} & \frac{\partial^2 f}{\partial x_n \partial x_1} - \lambda \frac{\partial^2 g}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} - \lambda \frac{\partial^2 g}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2} - \lambda \frac{\partial^2 g}{\partial x_n^2}
\end{vmatrix}
$$

If the signs of minors \( \Delta_3, \Delta_4, \Delta_5 \) are alternatively positive and negative, then extreme point is a local maximum. But if sign of all minors \( \Delta_3, \Delta_4, \Delta_5 \) are negative, the extreme point is a local minimum.

Let us see an example.

**Example 6:** Suppose we want to find the extreme values of the function

$$Z = 2x_1^2 + x_2^2 + 3x_3^2 + 10x_1 + 8x_2 + 6x_3 - 100$$

subject to the constraint

$$x_1 + x_2 + x_3 = 20, \quad x_1, x_2, x_3 \geq 0$$

**Solution:** Here \( n = 3 \) and \( m = 1 \). Let \( g(x_1, x_2, x_3) = x_1 + x_2 + x_3 - 20 \). Lagrangian function can be formulated as:

$$L(x, \lambda) = 2x_1^2 + x_2^2 + 3x_3^2 + 10x_1 + 8x_2 + 6x_3 - 100 - \lambda(x_1 + x_2 + x_3 - 20)$$

To obtain the stationary points, we solve the following system of equations.

$$\frac{\partial L}{\partial x_1} = 4x_1 + 10 - \lambda = 0; \quad \frac{\partial L}{\partial x_2} = 2x_2 + 8 - \lambda = 0$$

$$\frac{\partial L}{\partial x_3} = 6x_3 + 6 - \lambda = 0; \quad g(x_1, x_2, x_3) = x_1 + x_2 + x_3 - 20 = 0$$

Putting the value of \( x_1, x_2 \) and \( x_3 \) in the last equation \( g(x_1, x_2, x_3) = 0 \), and solving for \( \lambda \), we get \( \lambda = 30 \). Substituting the value of \( \lambda \) in the other three equations, we get the stationary point: \( (x_1, x_2, x_3) = 5, 11, 4 \).

To prove the sufficient condition whether the stationary point gives maximum or minimum value of the function we evaluate 2 principal minors as illustrated.
Calculus in \( \mathbb{R}^n \) in Sec. 7.2.

\[
\Delta_3 = \begin{vmatrix}
0 & \frac{\partial g}{\partial x_1} & \frac{\partial g}{\partial x_2} \\
\frac{\partial g}{\partial x_1} & \frac{\partial^2 f}{\partial x_1^2} - \lambda \frac{\partial^2 g}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} - \lambda \frac{\partial^2 g}{\partial x_1 \partial x_2} \\
\frac{\partial g}{\partial x_2} & \frac{\partial^2 f}{\partial x_2 \partial x_1} - \lambda \frac{\partial^2 g}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} - \lambda \frac{\partial^2 g}{\partial x_2^2}
\end{vmatrix} = \begin{vmatrix}
0 & 1 & 1 \\
1 & 4 & 0 \\
1 & 0 & 2
\end{vmatrix} = -6_{(5,11,4)}
\]

\[
\Delta_4 = \begin{vmatrix}
0 & 1 & 1 & 1 \\
1 & 4 & 0 & 0 \\
1 & 0 & 2 & 0 \\
1 & 0 & 0 & 6
\end{vmatrix} = 48
\]

Since sign of \( \Delta_3 \) and \( \Delta_4 \) are alternative, the stationary point: 
\((x_1, x_2, x_3) = (5, 11, 4)\) is a local maximum. At this point the value of the function is, \( Z = 281. \)

***

In the appendix we have given an illustrative example where we have explained how Lagrange Multiplier method can be used for modelling problem in economics.

You can try this exercise now.

E2) Find and clarify the extreme values of the following functions subject to the constraints given along side.

i) \( f(x_1, x_2, x_3) = x_1^2 + x_2^2 + x_3^2 \) subject to the constraint \( 4x_1 + x_2^2 + 2x_3 = 14, x_1, x_2, x_3 \geq 0. \)

ii) \( f(x_1, x_2) = 4x_1 + 6x_2 - 2x_1^2 - 2x_1x_2 - 2x_2^2 \) subject to the constraint \( x_1 + 2x_2 = 2, x_1, x_2 \geq 0. \)

With this we come to an end of this unit and to this block.

### 7.4 SUMMARY

In this unit, we have covered the following points for real-valued functions of vector-variable:

1. We have defined
   i) critical points and stationary points
   Let \( f : E \subset \mathbb{R}^n \rightarrow \mathbb{R} \), \( E \) an open set in \( \mathbb{R}^n \), be a function. A point \( a \in E \) is called a critical point of \( f \) if either
   i) the partial derivatives of \( f \) do not exist at \( a \), or
   ii) \( \frac{\partial f}{\partial x_i} (a) = 0 \) for all \( i \) such that \( 1 \leq i \leq n. \)

   The points for which the condition (ii) is satisfied are called stationary points.
ii) saddle point:
A point \( a \in E \) is called a **saddle point** if every neighbourhood \( E_a \) of a contains points \( x \in E \) such that \( f(x) > a \) and other points \( y \in E_a \) such that \( y < f(a) \).

iii) local maxima and local minima
Let \( f : E \subseteq \mathbb{R}^n \rightarrow \mathbb{R} \) be a function where \( E \) is an open subset of \( \mathbb{R}^n \). A point \( a \in E \) is said to be a local maximum for \( f \) if there exists a neighbourhood \( E_a \) of \( a \) such that \( f(x) \leq f(a) \) for all \( x \in E \).
A local minima is similarly defined.

2. We have established that a necessary condition for a function to have local extrema is \( \nabla f = 0 \) provide \( \nabla f \) exists.

3. We have derived a test called second derivative test for finding local extrema.

(Second-derivative test for extrema). Let \( f : E \rightarrow \mathbb{R} \) be a function defined on an open set \( E \subseteq \mathbb{R}^n \). Assume that the second-order partial derivatives \( D_{ij}f \) exist in an \( n \)-ball \( B(a) \) and are continuous at \( a \in \mathbb{R}^n \), where \( a \) is a stationary point of \( f \). Let

\[
Q(x) = \frac{1}{2} f''(a; x) = \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} D_{ij}f(a)x_ix_j
\]

where \( x = (x_1, \ldots, x_n) \). Then

a) If \( Q(x) > 0 \) for all \( x \neq 0 \), \( f \) has a relative minimum at \( a \).

b) If \( Q(x) < 0 \) for all \( x \neq 0 \), \( f \) has a relative maximum at \( a \).

c) If \( Q(x) \) takes both positive and negative values, then \( f \) has a saddle point at \( a \).

4. We have explained a method for classifying local maxima and local minima using the Hessian.

5. We have explained Lagrange's Multiplier Method.

### 7.5 HINTS/SOLUTIONS

E1) \( \nabla f(x, y, z) = (2xy - 2, x^2 + 2yz, y^2 + 2z) \)

and the critical points satisfy the equations

\[
2xy - 2 = 0, \quad x^2 + 2yz = 0 \quad \text{and} \quad y^2 + 2z = 0.
\]

Substituting \( z = -y^2/2 \) into the second equation implies \( y^3 = x^2 \). Hence, the first equation shows \( y^{5/2} = 1 \) and we have \( y = 1 \) and \( z = -1/2 \). From \( xy = -1 \) we get \( x = 1 \) and \( (1, 1, -1/2) \) is the only critical point of \( f \). We have

\[
H_{f(x,y,z)} = \begin{pmatrix}
2y & 2x & 0 \\
2x & 2z & 2y \\
0 & 2y & 2
\end{pmatrix}
\]

and

\[
H_{f(1,1,-1/2)} = \begin{pmatrix}
2 & 2 & 0 \\
2 & -1 & 2 \\
0 & 2 & 2
\end{pmatrix}
\]
Since $\det(2) > 0$ and

$$\det \begin{pmatrix} 2 & 2 \\ 2 & -1 \end{pmatrix} = -2 - 4 < 0$$

the point $(1, 1, -1/2)$ is a saddle point of $f$.

E2)  

i) **Hint:** The extreme point is $\left( \frac{81}{100}, \frac{7}{20}, \frac{7}{25} \right)$. It is a minimum and the minimum value is $\frac{857}{100}$.

ii) **Hint:** The extreme point is $\left( \frac{1}{3}, \frac{5}{6} \right)$. It is a maximum point and the value is 4.166.
A mathematical model for utility maximization problem in economics

[Source: Lagrange Multiplier Problems in Economics
John V. Baxley and John C. Moorhouse, Wake Forest University,
Winston-Salem, NC 27109. American Mathematical Monthly,
August-September 1984, Volume 91, Number 7, pp.404-412.]

Assume an individual obtains utility (i.e., satisfaction) from the consumption of two goods $x$ and $y$ which are purchased in the marketplace in quantities $X$ and $Y$, respectively, and from a certain quantity $L$ of leisure time $l$. The theory of consumer choice can then be characterized mathematically as a constrained optimization problem in which the individual chooses, in an optimal way, the amount of leisure time $l$ (and implicitly money income) and the quantities of $x$ and $y$ consumed. We shall assume that the individual spends all his income on the purchases of such goods. If this offends readers who are avid savers, they may assume that one of the consumption goods represents savings. To keep the problem manageable, however, we proceed on the assumption of a two-good world. Let $P_x$ and $P_y$ be the price per unit of $x$ and $y$, respectively. Let $T$ represent the total time available and $L$ the amount of leisure time chosen by the individual per period (e.g., per week, month, or year). Work time can then be defined as $T - L$. If the market-determined wage rate is $w$, the individual’s income is then $w(T - L)$. Since income equals consumption expenditures, the budget constraint becomes:

$$w(T_L) - P_x X - P_y Y = 0.$$  \hfill (11)

We assume that the utility $U$ obtained from a particular choice of $X$, $Y$, and $L$ is given by a utility function (unique to the individual)

$$U = u(X, Y, L).$$  \hfill (12)

Thus the consumer’s problem is to choose $X$, $Y$, $L$ to maximize $U$ subject to the constraint (11). Furthermore it is reasonable to assume that

$$u_x > 0, \quad u_y > 0, \quad u_L > 0,$$  \hfill (13)

$$u_{xx} < 0, \quad u_{yy} < 0, \quad u_{ll} < 0,$$  \hfill (14)

$$u_{xl} > 0, \quad u_{yl} > 0,$$  \hfill (15)

either $u_{xy} > 0$ or $u_{xy} < 0$ or $u_{xy} = 0$.  \hfill (16)

where subscripts denote, as usual, partial derivatives. The first order derivatives in Equation (12) are called marginal utilities. The inequalities in Equation (12) state that utility increases with higher level of consumption of $X$ and $Y$ and with more leisure time per period. The inequalities in Equation (13) state the “law of diminishing marginal utility.” That is, while utility increases with consumption of $x$, $y$, and $l$, it increases at a diminishing rate. The inequalities in Equation (14) state that the satisfaction one derives from consuming larger amounts of $x$ (or $y$) is enhanced by the availability of more leisure time. It takes time to “enjoy things”. The sign of $u_{xy}$ in Equation (15) depends on whether $x$ and $y$ are substitutes, complements, or unrelated. If the goods are substitutes, $u_{xy} < 0$; if complements, $u_{xy} > 0$; and if unrelated, $u_{xy} = 0$. 65
Applying the method of Lagrange, we introduce the multiplier \( \lambda \) and form the Lagrangian

\[
\nu(X, Y, L\lambda) = u(X, Y, L) + \lambda(w(T - L) - P_xX - P_yY).
\]  

(17)

Assuming the consumer maximizes utility, the optimal quantities \( X^*, Y^*, L^* \) and the multiplier \( \lambda^* \) necessarily satisfy the first-order conditions:

\[
\nu_\lambda = w(T - L) - P_xX - P_yY = 0, 
\]

(18)

\[
\nu_X = u_X - \lambda P_x = 0, 
\]

(19)

\[
\nu_Y = u_Y - \lambda P_y = 0, 
\]

(20)

\[
\nu_L = u_L - \lambda w = 0. 
\]

(21)

Observe that because of Equation (12) each of the equations imply that the optimal value \( \lambda^* > 0 \).

In the model, the individual chooses \( X, Y, \) and \( L \). The prices \( P_x, P_y \) and the wage rate \( w \) are given by market conditions beyond the individual’s influence or control. \( T \) is given by nature.