MODULE 7 MAXIMA AND MINIMA OF FUNCTIONS OF SEVERAL VARIABLES, STATIONARY POINT, LAGRANGE'S METHOD OF MULTIPLIERS

UNIT 1 MAXIMISATION AND MINIMISATION OF FUNCTIONS OF SEVERAL VARIABLES

CONTENTS

1.0 INTRODUCTION

Def. Stationary (or critical) point. For a function $y = f(x)$ of a single variable, a stationary (or critical) point is a point at which $dy/dx = 0$; for a function $u = f(x_1, x_2, ... ,x_n)$ of n variables it is a point at which

1. $$\frac{\partial u}{\partial x_1} = 0 \qquad \frac{\partial u}{\partial x_2} = 0 \ldots \quad \frac{\partial u}{\partial x_n} = 0$$

In the case of a function $y = f(x)$ of a single variable, a stationary point corresponds to a point on the curve at which the tangent to the curve is horizontal. In the case of a function $y = f(x, y)$ of two variables a stationary point corresponds to a point on the surface at which the tangent plane to the surface is horizontal.

In the case of a function $y = f(x)$ of a single variable, a stationary point can be any of the following three: a maximum point, a minimum point or an inflection point. For a function $y = f(x, y)$ of two variables, a stationary point can be a maximum point, a minimum point or a saddle point. For a function of n variables it can be a maximum point, a minimum point or a point that is analogous to an inflection or saddle point.

2.0 OBJECTIVE

At the end of this unit, you should be able to:

- recognise problems on maximum and minimum functions of several variables;
- know the necessary condition for a maxima or minima function of several variable;
- know the Sufficient condition for a maxima or minima function of several variable;
- identify the maxima and minima of functions subject to constraints;
- know the method of finding maxima and minima of functions subject to constraints;
- identify the different types of examples of maxima and minima functions of several variables; and
- solve problems on maxima and minima functions of several variables.

3.0 MAIN CONTENT

Maxima and minima of functions of several variables

A function f(x, y) of two independent variables has a **maximum** at a point (x_0, y_0) if $f(x_0, y_0) \geq f(x, y)$ for all points (x, y) in the neighborhood of (x_0, y_0). Such a function has a **minimum** at a point (x_0, y_0) if $f(x_0, y_0) \leq f(x, y)$ for all points (x, y) in the neighborhood of (x_0, y_0).

A function $f(x_1, x_2, ... , x_n)$ of *n* independent variables has a **maximum** at a point $(x_1', x_2', ... ,x_n')$ if $f(x_1', x_2', ... , x_n') \leq f(x_1, x_2, ... , x_n)$ at all points in the neighborhood of $(x_1', x_2', ... , x_n')$.Such a function has a minimum at a point $(x_1', x_2', ... ,x_n')$ if $f(x_1', x_2', ... , x_n')f(x_1, x_2, ...,x_n)$ at all points in the neighborhood of $(x_1', x_2', ... , x_n')$.

Necessary condition for a maxima or minima: A necessary condition for a function f(x, y) of two variables to have a maxima or minima at point (x_0, y_0) is that

$$\frac{\partial f}{\partial x} = 0 , \qquad \frac{\partial f}{\partial y} = 0$$

at the point (i.e. that the point be a **stationary point).**

In the case of a function $f(x_1, x_2, ... ,x_n)$ of n variables, the condition for the function to have a maximum or minimum at point $(x_1', x_2', ... , x_n')$ is that

$$\frac{\partial f}{\partial x_1} = 0 \qquad \frac{\partial f}{\partial x_2} = 0,\ldots, \ \frac{\partial f}{\partial x_n} = 0$$

at that point (i.e. that the point be a **stationary point).**

To find the maximum or minimum points of a function you first locate the stationary points using (1) above. After locating the stationary points you then examine each stationary point to determine if it is a maximum or minimum. To determine if a point is a maximum or minimum you may consider values of the function in the neighborhood of the point as well as the values of its first and second partial derivatives. You also may be able to establish what it is by arguments of one kind or other. The following theorem may be useful in establishing maximums and minimums for the case of functions of two variables.

Sufficient condition for a maximum or minimum of a function

z = f(x, y). Let z = f(x, y) have continuous first and second partial derivatives in the neighborhood of point (x_0, y_0). If at the point (x_0, y_0)

$$\frac{\partial f}{\partial x} = 0 \, , \qquad \frac{\partial f}{\partial y} = 0$$

and

$$\Delta = \left(\frac{\partial^2 f}{\partial x \partial y}\right)^2 - \frac{\partial^2 f}{\partial x^2}\frac{\partial^2 f}{\partial y^2} < 0$$

then there is a maximum at (x_0, y_0) if

$$\frac{\partial^2 f}{\partial x^2} < 0$$

and a minimum if

$$\frac{\partial^2 f}{\partial x^2} > 0$$

If $\Delta > 0$, point (x_0, y_0) is a saddle point (neither maximum nor minimum). If $\Delta = 0$, the nature of point (x_0, y_0) is undecided. More investigation is necessary.

Example. Find the maxima and minima of function $z = x^2 + xy + y^2 - y$.

Solution..

$$\frac{\partial z}{\partial x} = 2x + y, \quad \frac{\partial z}{\partial y} = x + 2y - 1$$

$$\frac{\partial^2 z}{\partial x^2} = 2, \quad \frac{\partial^2 z}{\partial x \partial y} = 1, \quad \frac{\partial^2 z}{\partial y^2} = 2$$

2x + y = 0

x + 2y = 1

x = -1/3, y = 2/3

This is the stationary point. At this point Δ > 0 and

$$\frac{\partial^2 z}{\partial x^2} > 0$$

and the point is a minimum. The minimum value of the function is - 1/3.

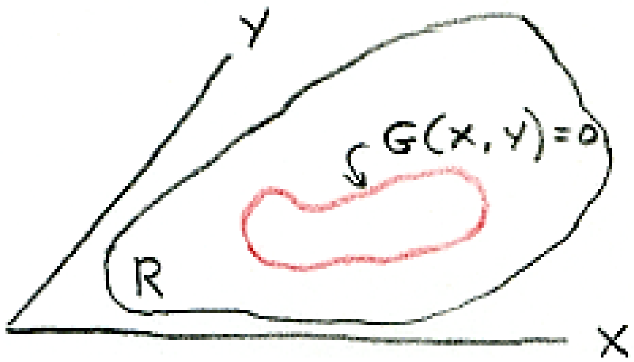


Fig. 1

Maxima and minima of functions subject to constraints. Let us set ourselves the following problem: Let F(x, y) and G(x, y) be functions defined over some region R of the x-y plane. Find the points at which the function F(x, y) has maximums subject to the side condition.

G(x, y) = 0. Basically you are asking the question: At what points on the solution set of G(x, y) = 0 does F(x, y) have maximums? The solution set of G(x, y) = 0 corresponds to some curve in the plane. See Figure 1. The solution set (i.e. locus) of G(x, y) = 0 is shown in red. Figure 2 shows the situation in three dimensions where function z = F(x, y) is shown rising up above the x-y plane along the curve G(x, y) = 0. The problem is to find the maximums of z = F(x, y) along the curve

G(x, y) = 0.

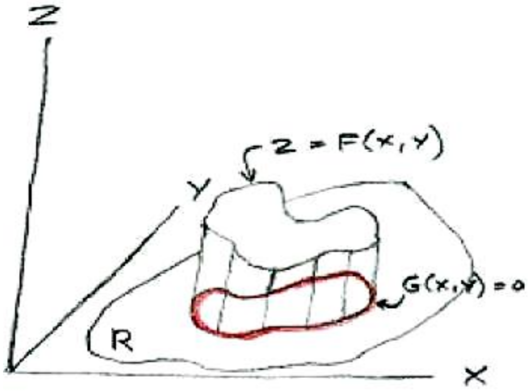


Fig. 2

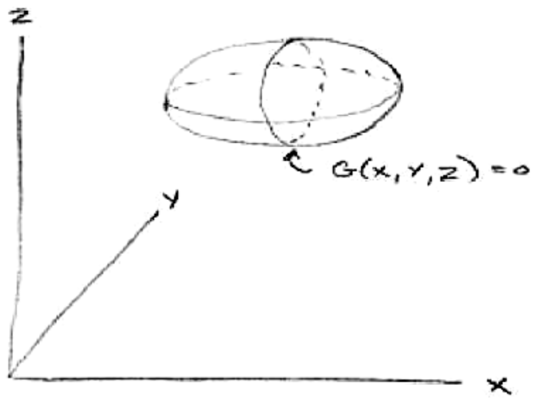


Fig. 3

Let us now consider the same problem in three variables. Let F(x, y, z) and G(x, y, z) be functions defined over some region R of space. Find the points at which the function F(x, y, z) has maximums subject to the side condition G(x, y, z) = 0. Basically we are asking the question: At what points on the solution set of G(x, y, z) = 0 does F(x, y, z) have maximums? G(x, y, z) = 0 represents some surface in space. In Figure 3, G(x, y, z) = 0 is depicted as a spheroid in space. The problem then is to find the maximums of the function F(x, y, z) as evaluated on this spheroidal surface.

Let us now consider another problem. Suppose instead of one side condition we have two. Let F(x, y, z), G(x, y, z) and H(x, y, z) be functions defined over some region R of space. Find the points at which the function F(x, y, z) has maximums subject to the side conditions

G(x, y, z) = 0
H(x, y, z) = 0.

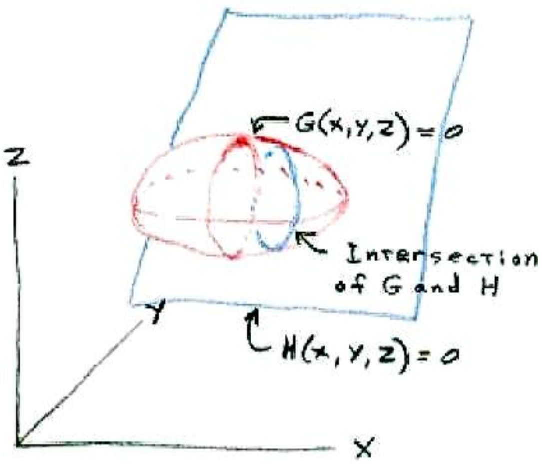


Fig. 4

Here we wish to find the maximum values of F(x, y, z) on that set of points that satisfy both equations 2) and 3). Thus if D represents the solution set of G(x, y, z) = 0 and E represents the solution set of H(x, y, z) = 0 we wish to find the maximum points of F(x, y, z) asevaluated on set F = D ∩E (i.e. the intersection of sets D and E). In Fig. 4 G(x, y, z) = 0 is depicted as an ellipsoid and H(x, y, z) = 0 as a plane. The intersection of the ellipsoid and the plane is the set F on which F(x, y, z) is to be evaluated.

The above can be generalized to functions of *n* variables
$F(x_1, x_2, ... ,x_n)$, $G(x_1, x_2, ... , x_n)$, etc. and *m* side conditions.

Methods for finding maxima and minima of functions subject to constraints

1. **Method of direct elimination.** Suppose you wish to find the maxima or minima of a function F(x, y) with the constraint Φ(x, y) = 0. Suppose you are so lucky that Φ(x, y) = 0 can be solved explicitly for y, giving y = g(x). You can then substitute g(x) for y in F(x, y) and then find the maximums and minimums of F(x, g(x)) by standard methods. In some cases, it may be possible to do this kind of thing. You express some of the variables in the equations of constraint in terms of other variables and then substitute into the function whose extrema are sought, and find the extrema by standard methods.

2. **Method of implicit functions.** Suppose you wish to find the maxima or minima of a function u = F(x, y, z) with the constraint Φ(x, y, z) = 0. You note that Φ(x, y, z) = 0 defines z implicitly as a function of x and y i.e. z = f(x, y). You thus seek the extrema of the quantity

   u = F(x, y, f(x, y)).

The necessary condition for a stationary point, as given by (1) above, becomes

4) $\frac{\partial u}{\partial x} = F_1 + F_3 \frac{\partial z}{\partial x} = 0 \qquad \frac{\partial u}{\partial y} = F_2 + F_3 \frac{\partial z}{\partial y} = 0$

(where F1 represents the partial of F with respect to x, etc.) Taking partials of Φ with respect to x and y it follows that

5) $\Phi_1 + \Phi_3 \frac{\partial z}{\partial x} = 0 \qquad \Phi_2 + \Phi_3 \frac{\partial z}{\partial y} = 0.$

(since the partial derivative of a function that is constant is zero).
From the pair of equations consisting of the first equation in 4) and 5) you can eliminate$\partial z/\partial x$giving

6) $F_1\phi_3 - F_3\phi_1 = 0$

From the pair of equations consisting of the second equation in 4) and 5) you can eliminate $\partial z/\partial y$ giving

7) $F_2\phi_3 - F_3\phi_2 = 0$

Equations 6) and 7) can be written in determinant form as

8) $\begin{vmatrix} F_1 & F_3 \\ \Phi_1 & \Phi_3 \end{vmatrix} = 0 \qquad \begin{vmatrix} F_2 & F_3 \\ \Phi_2 & \Phi_3 \end{vmatrix} = 0$

Equations 8) combined with the equation Φ(x, y, z) = 0 give you three equations which you can solve simultaneously for x, y, z to obtain the stationary points of function F(x, y, z). The maxima and minima will be among the stationary points.

This same method can be used for functions of an arbitrary number of variables and an arbitrary number of side conditions (smaller than the number of variables).

Extrema for a function of four variables with two auxiliary equations: Suppose you wish to find the maxima or minima of a function

u = F(x, y, z, t)

with the side conditions

9) Φ(x, y, z, t) = 0 ψ(x, y, z, t) = 0.

Equations 9) define variables z and t implicitly as functions of x and y i.e.

10) $z = f_1(x,y) \quad t = f_2(x, y)$.

We thus seek the extrema of the quantity

$$u = F(x, y, f_1(x, y), f_2(x, y)) .$$

The necessary condition for a stationary point, as given by 1) above, becomes

11) $$\frac{\partial u}{\partial x} = F_1 + F_3\frac{\partial z}{\partial x} + F_4\frac{\partial t}{\partial x} = 0 \quad \frac{\partial u}{\partial y} = F_2 + F_3\frac{\partial z}{\partial y} + F_4\frac{\partial t}{\partial x} = 0$$

Taking partials of Φ with respect to x and y it follows that

12) $$\Phi_1 + \Phi_3\frac{\partial z}{\partial x} + \Phi_4\frac{\partial t}{\partial x} = 0 \quad \Phi_2 + \Phi_3\frac{\partial f}{\partial y} + \Phi_4\frac{\partial t}{\partial y} = 0.$$

Taking partials of ψ with respect to x and y it follows that

13) $$\psi_1 + \psi_3\frac{\partial z}{\partial x} + \psi_4\frac{\partial t}{\partial x} = 0 \quad \psi_2 + \psi_3\frac{\partial f}{\partial y} + \psi_4\frac{\partial t}{\partial y} = 0.$$

From 12) and 13) we can derive the conditions

14) $$\begin{vmatrix} F_1 & F_3 & F_4 \\ \Phi_1 & \Phi_3 & \Phi_4 \\ \psi_1 & \psi_3 & \psi_4 \end{vmatrix} = 0 \quad \begin{vmatrix} F_2 & F_3 & F_4 \\ \Phi_2 & \Phi_3 & \Phi_4 \\ \psi_2 & \psi_3 & \psi_4 \end{vmatrix} = 0$$

Equations 14) combined with the auxiliary equations Φ(x, y, z, t) = 0 and ψ(x, y, z, t) = 0 give you four equations which you can solve simultaneously for x, y, z, t to obtain the stationary points of function F(x, y, z, t). The maxima and minima will be among the stationary points.

Extrema for a function of n variables with p auxiliary equations.
The p equations corresponding to equation 14) above for the case of a function of n variables

$u = F(x_1, x_2, \ldots .x_n)$

and p auxiliary equations (i.e. side conditions)

$\Phi(x_1, x_2, \ldots ,x_n) = 0$
$\Psi(x_1, x_2, \ldots , x_n) = 0$
....
$\Omega(x_1, x_2, \ldots ,x_n) = 0$

are

15) $$\begin{vmatrix} F_k & F_{n-p+1} \cdots & F_n \\ \Phi_k & \Phi_{n-p+1} & \Phi_n \\ \cdots & \cdots & \cdots \\ \Omega_k & \Omega_{n-p+1} & \Omega_n \end{vmatrix} = 0 \quad k = 1,2,\ldots,\text{p}$$

These p equations along with the p auxiliary equations

$\Phi(x_1, x_2, \ldots, x_n) = 0$
$\Psi(x_1, x_2, \ldots, x_n) = 0$
....
$\Omega(x_1, x_2, \ldots, x_n) = 0$

can be solved simultaneously for the n variables $x_1, x_2, \ldots .x_n$ to obtain the stationary points of $F(x_1, x_2, \ldots .x_n)$. The maxima and minima will be among the stationary points.

Geometrical interpretation for extrema of functionF(x, y, z) with a constraint: We shall now present a theorem that gives a geometrical interpretation for the case of extrema values of functions of type F(x, y, z) with a constraint.

Theorem 1: Suppose the functions F(x, y, z) and 0(x, y, z) have continuous first partial derivatives throughout a certain region R of space. Let the equation 0(x, y, z) = 0 define a surface S, every point of which is in the interior of R, and suppose that the three partial derivatives ϕ_1, ϕ_2, ϕ_3 are never simultaneously zero at a point of S. Then a necessary condition for the values of F(x, y, z) on S to attain an extreme value (either relative or absolute) at a point of S is that F_1, F_2, F_3 be proportional to ϕ_1, ϕ_2, ϕ_3 at that point. If C is the value of F at the point, and if the constant of proportionality is not zero, the geometric meaning of the proportionality is that the surface S and the surface F(x, y, z) = C are tangent at the point in question.

Rationale behind theorem: From 8) above, a necessary condition for F(x, y, z) to attain a maxima or minima (i.e. a condition for a stationary point) at a point P is that

$F_1\phi_3 - F_3\phi_1 = 0 \qquad F_2\phi_3 - F_3\phi_2 = 0$

or

16) $$\frac{F_1}{\Phi_1} = \frac{F_3}{\Phi_3}\frac{F_2}{\Phi_2} = \frac{F_3}{\Phi_3}\,.$$

Thus at a stationary point the partial derivatives F_1, F_2, $F_3$and Φ_1, Φ_2, Φ_3 are proportional. Now the partial derivatives F_1, F_2, $F_3$and 01, 02, 03 represent the gradients of the functions F and Φ; and the gradient, at any point P, of a scalar point function y(x, y, z) is a vector that is normal to that level surface of ψ(x, y, z) that passes through point P. If C is the value of F at the stationary point P, then the vector (F_1, F_2, F_3) at point P is normal to the surface F(x, y, z) = C at P. Similarly, the vector (Φ_1, Φ_2, Φ_3) at point P is normal to the surface Φ(x, y, z) = 0 at P. Since the partial derivatives F_1, F_2, $F_3$and Φ_1, Φ_2, $\Phi_3$are proportional, the normals to the two surfaces point in the same direction at P and the surfaces must be tangent at point P.

Example. Consider the maximum and minimum values of $F(x, y, z) = x^2 + y^2 + z^2$ on the surface of the ellipsoid

$$G(x,y,z) = \frac{x^2}{64}+\frac{y^2}{36}+\frac{z^2}{25} = 1.$$

Since F(x, y, z) is the square of the distance from (x, y, z) to the origin, it is clear that we are looking for the points at maximum and minimum distances from the center of the ellipsoid. The maximum occurs at the ends of the longest principal axis, namely at (±8, 0, 0). The minimum occurs at the ends of the shortest principal axis, namely at (0, 0, ±5). Consider the maximum point (8, 0, 0). The value of F at this point is 64, and the surface F(x, y, z) = 64 is a sphere. The sphere and the ellipsoid are tangent at (8, 0, 0) as asserted by the theorem. In this case the ratios G_1:G_2:G_3 and F_1:F_2:F_3 at (8, 0, 0) are 1/4 : 0 : 0 and 16 : 0 : 0 respectively.

This example brings out the fact that the tangency of the surfaces (or the proportionality of the two sets of ratios), is a necessary but not a sufficient condition for a maximum or minimum value of F, for you note that the condition of proportionality exists at the points (0, ±6, 0), which are the ends of the principal axis of intermediate length. But the value of F in neither a maximum nor a minimum at this point.

Case of extrema of function F(x, y) with a constraint: A similar geometrical interpretation can be given to the problem of extremal values for F(x, y) subject to the constraint Φ(x, y) = 0. Here you have a curve defined by the constraint, and a one-parameter family of curves F(x, y) = C. At a point of extremal value of F the curve F(x, y) = C through the point will be tangent to the curve defined by the constraint.

Lagrange's Method of Multipiers: Let F(x, y, z) and Φ(x, y, z) be functions defined over some region R of space. Find the points at which the function F(x, y, z) has maximums and minimums subject to the side condition Φ(x, y, z) = 0. Lagrange's method for solving this problem consists of forming a third function G(x, y, z) given by

17) $G(x, y, z) = F(x, y, z) + \lambda\Phi(x, y, z)$,

where λ is a constant (i.e. a parameter) to which you will later assign a value, and then finding the maxima and minima of the function G(x, y, z). A reader might quickly ask, "Of what interest are the maxima and minima of the function G(x, y, z)? How does this help us solve the problem of finding the maxima and minima of F(x, y, z)?" The answer is that examination of 17) shows that for those points corresponding to the solution set of Φ(x, y, z) = 0 the function G(x, y, z) is equal to the function F(x, y, z) since at those points equation 17) becomes

$G(x, y, z) = F(x, y, z) + \lambda\cdot 0$

Thus, for the points on the surface Φ(x, y, z) = 0, functions F and G are equal so the maxima and minima of G are also the maxima and minima of F. The procedure for finding the maxima and minima of G(x, y, z) is as follows: You regard G(x, y, z) as a function of three independent variables and write down the necessary conditions for a stationary point using 1) above:

18) $F_1 + \lambda\Phi_1 = 0 \qquad F_2 + \lambda\Phi_2 = 0 \qquad F_3 + \lambda\Phi_3 = 0$

We then solve these three equations along with the equation of constraint Φ(x, y, z) = 0 to find the values of the four quantities x, y, z, λ. More than one point can be found in this way and this will give you the locations of the stationary points. The maxima and minima will be among the stationary points thus found.

Let us now observe something. If equations 18) are to hold simultaneously, then it follows from the third of them that λ must have the value $\lambda = -\frac{F_3}{\Phi_3}$.

If you substitute this value of λ into the first two equations of 18) we obtain

$F_1\Phi_3 - F_3\Phi_1 = 0 \qquad F_2\Phi_3 - F_3\Phi_2 = 0$

or

19) $\begin{vmatrix} F_1 & F_3 \\ \Phi_1 & \Phi_3 \end{vmatrix} = 0 \qquad \begin{vmatrix} F_2 & F_3 \\ \Phi_2 & \Phi_3 \end{vmatrix} = 0$

You note that the two equations of 19) are identically the same conditions as 8) above for the previous method. Thus using equations 19) along with the equation of constraint Φ(x, y, z) = 0 is exactly the same procedure as the previous method in which you used equations 8) and the same constraint.

One of the great advantages of Lagrange's method over the method of implicit functions or the method of direct elimination is that it enables you to avoid making a choice of independent variables. This is sometimes very important; it permits the retention of symmetry in a problem where the variables enter symmetrically at the outset.

Lagrange's method can be used with functions of any number of variables and any number of constraints (smaller than the number of variables). In general, given a function $F(x_1, x_2, \ldots, x_n)$ of n variables and h side conditions $\Phi_1 = 0, \Phi_2 = 0, \ldots,$ $\Phi_h = 0$, for which this function may have a maximum or minimum, equate to zero the partial derivatives of the auxiliary function $F + \lambda_1\Phi_1 + \lambda_2\Phi_2 + \ldots + \lambda_h\Phi_h$ with respect to $x_1, x_2, \ldots, x_n$, regarding $\lambda_1, \lambda_2, \ldots, \lambda_h$ as constants, and solve these n equations simultaneously with the given h side conditions, treating the λ's as unknowns to be eliminated.

The parameter λ in Lagrange's method is called Lagrange's multiplier.

Further examples

Example 1

Let us find the critical points of

$$z = f(x, y) = \exp\left(-\frac{1}{3}x^3 + x - y^2\right)$$

The partial derivatives are

$$f_x(x, y) = (-x^2 + 1)\exp\left(-\frac{1}{3}x^3 + x - y^2\right)$$

$$f_y(x, y) = -2y\exp\left(-\frac{1}{3}x^3 + x - y^2\right)$$

$f_x = 0$ if $1-x^2=0$ or the exponential term is 0. $f_y=0$ if $-2y=0$ or the exponential term is 0. The exponential term is not 0 except in the degenerate case. Hence you require $1 - x^2=0$ and $-2y=0$, implying $x=1$ or $x=-1$ and $y=0$. There are two critical points (-1,0) and (1,0)

The Second Derivative Test for Functions of Two Variables

How can we determine if the critical points found above are relative maxima or minima? You apply a second derivative test for functions of two variables.

Let (x_c,y_c) be a critical point and define

$$D(x_c, y_c) = f_{xx}(x_c, y_c)f_{yy}(x_c, y_c) - \left[f_{xy}(x_c, y_c)\right]^2$$

You have the following cases:

- If D>0 and f_{xx} (x_c , y_c) < 0, then f(x,y) has a relative maximum at
- x_c,y_c .
- If D>0 and f_{xx} (x_c ,y_c) <0, then f(x,y) has a relative minimum at x_c,y_c .
- If D<0, then f(x,y) has a saddle point at x_c,y_c
- If D=0, the second derivative test is inconclusive.

An example of a saddle point is shown in the example below.

Example: Continued

For the example above, we have

$$f_{xx}(x,y) = (-2x + (1-x^2)^2)exp\left(-\frac{1}{3}x^3 + x - y^2\right),$$

$$f_{yy}(x,y) = (-2 + 4y^2)exp\left(-\frac{1}{3}x^3 + x - y^2\right),$$

$$f_{xy}(x,y) = -2y(1-x^2)exp\left(-\frac{1}{3}x^3 + x - y^2\right),$$

For x=1 and y=0, we have D(1,0)=4exp(4/3)> 0 with
f_{xx}(1,0)=-2exp(2/3)<0. Hence, (1,0) is a relative maximum.
For x=-1 and y=0, we have D(-1,0)=-4exp(-4/3)<0. Hence, (-1,0) is a saddle point.

Example 2: Maxima and Minima in a Disk

Another example of a bounded region is the disk of radius 2 centered at the origin. You proceed as in the previous example, determining in the 3 classes above. (1,0) and (-1,0) lie in the interior of the disk.

The boundary of the disk is the circle $x^2+ y^2=4$. To find extreme points on the disk you parameterize the circle. A natural parameterization is
x=2cos(t) and y=2sin(t) for $0 \leq t \leq 2\pi$. You substitute these expressions into z=f(x,y) and obtain

$$z = f(x,y) = f(\cos(t), \sin(t)) = esp\left(-\frac{8}{3}\cos^3 t + 2\cos t - 4\sin^2 t\right) = g(t)$$

On the circle, the original functions of 2 variables is reduced to a function of 1 variable. You can determine the extrema on the circle using techniques from calculus of one variable.

In this problem there are not any corners. Hence, you determine the global max and min by considering points in the interior of the disk and on the circle. An alternative method for finding the maximum and minimum on the circle is the method of Lagrange multipliers.

4.0 CONCLUSION

You have been introduced to maximum and minimum functions of several variables, necessary condition for a maxima or minima function of several variables, problems on maximum and minimum functions of several variables etc.

5.0 SUMMARY

A summary of maximum and minimum functions of several variables are as follows:

A function f(x, y) of two independent variables has a **maximum** at a point (x_0, y_0) if $f(x_0, y_0) \geq f(x, y)$ for all points (x, y) in the neighborhood of (x_0, y_0). Such a function has a **minimum** at a point (x_0, y_0) if $f(x_0, y_0) \leq f(x, y)$ for all points (x, y) in the neighborhood of (x_0, y_0).

Solve the following problem, Find the maxima and minima of function
$z = x^2 + xy + y^2 - y$.

Solution

$$\frac{\partial z}{\partial x} = 2x + y, \quad \frac{\partial z}{\partial y} = x + 2y - 1 \quad \frac{\partial^2 z}{\partial x^2} = 2, \quad \frac{\partial^2 z}{\partial x \partial y} = 1, \quad \frac{\partial^2 z}{\partial y^2} = 2 .$$

$2x + y = 0$, $x + 2y = 1$

$x = -1/3$, $y = 2/3$

This is the stationary point. At this point $\Delta > 0$ and

$$\frac{\partial^2 z}{\partial x^2} > 0$$

and the point is a minimum. The minimum value of the function is - 1/3.

6.0 TUTOR-MARKED ASSIGNMENT

1. Determine the critical points and locate any relative minimum, maxima and saddle points offunctions f defined by

   $F(x,y) = 2x^2 - 2xy + 2y^4 - 6x$

2. Determine the critical points and locate any relative minimum, maxima and saddle points of functions f defined by

   $F(x, y) = 2x^4 - 4xy + y^3 + 4$

3. Determine the critical points and locate any relative minimum, maxima and saddle points of functions f defined by
   $F(x,y)= x^4 - y^4 + xy$.

Determine the critical points of the functions below and find out whether each point corresponds to a relative minimum, maximum and saddle point, or no conclusion can be made

4. $F(x,y)= x^2 + 3y^2 - 2xy - 8x$

5. $F(x,y)=x^3 + 12x + y^3 + 3y^2 - 9y$

7.0 REFERENCES/FURTHER READING

Taylor. Advanced Calculus

Osgood.Advanced Calculus.

James and James.Mathematics Dictionary.

Mathematics, Its Content, Methods and Meaning.Vol.

UNIT 2 LAGRANGE MULTIPLIERS

CONTENTS

1.0 INTRODUCTION

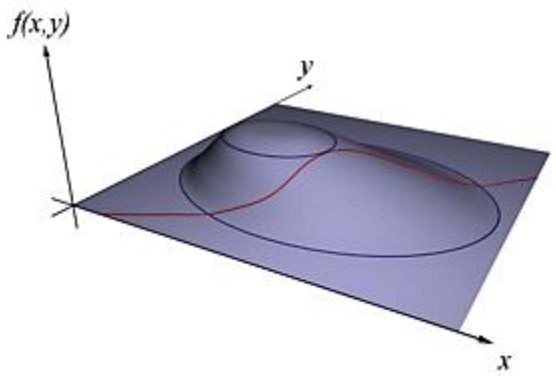


Figure 1: Find x and y to maximize f(x,y) subject to a constraint (shown in red) g(x,y) = c.

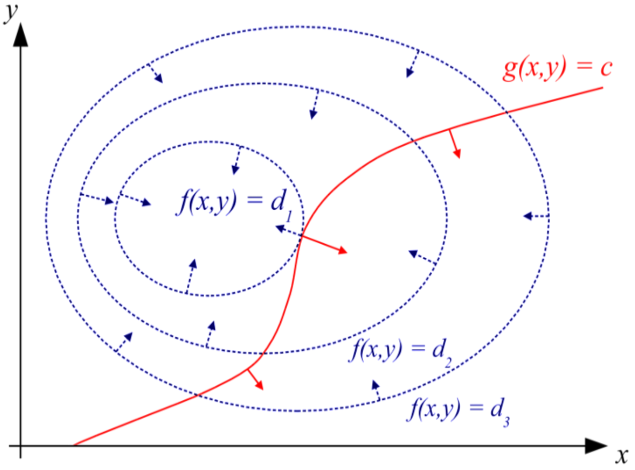


Figure 2: Contour map of Figure 1. The red line shows the constraint g(x,y) = c. The blue lines are contours of f(x,y). The point where the red line tangentially touches a blue contour is your solution.

In mathematical optimization, the method of Lagrange multipliers (named after Joseph Louis Lagrange) provides a strategy for finding the maxima and minima o f a function subject to constraints.

For instance (see Figure 1), consider the optimization problem

maximize $f(x, y)$
subject to $g(x, y) = c.$

We introduce a new variable (λ) called a Lagrange's multiplier, and study the Lagrange's function defined by

$$\Lambda(x, y, \lambda) = f(x, y) + \lambda \cdot \Big(g(x, y) - c\Big),$$

where the λ term may be either added or subtracted. If f(x,y) is a maximum for the original constrained problem, then there exists λ such that (x,y,λ) is a stationary point for the Lagrange's function (stationary points are those points where the partial derivatives of Λ are zero). However, not all stationary points yield a solution of the original problem. Thus, the method of Lagrange's multipliers yield a necessary condition for optimality in cons trained problems

2.0 OBJECTIVES

After studying this unit, you should be able to:

- identify problem which could be solve by languages multiplier;
- explain single and multiple constraints;
- explain the interpretation of language's multiplier; and
- solve problems with the use of language's multiplier.

3.0 MAIN CONTENT

One of the most common problems in calculus is that of finding maxima or minima (in general, "extrema") of a function, but it is often difficult to find a closed form for the function being extremized. Such difficulties often arise when one wishes to maximize or minimize a function subject to fixed outside conditions or constraints. The method of Lagrange's multipliers is a powerful tool for solving this class of problems without the need to explicitly solve the conditions and use them to eliminate extra variables.

Consider the two-dimensional problem introduced above:

maximize $f(x, y)$

subject to $g(x, y) = c.$

We can visualize contours of *f* given by

$$f(x,y) = d$$

for various values of d, and the contour of g given by g(x,y) = c.

Suppose you walk along the contour line with g = c. In general the contur lines of f and g may be distinct, so following the contour line for g = c one could intersect with or cross the contour lines of f. This is equivalent to saying that while moving along the contour line for g = c the value of f can vary. Only when the contour line for g = c meet contour lines of f tangentially, do you not increase or decrease the value of f — that is, when the contour lines touch but do not cross

The contour lines of f and g touch when the tangent vectors of the contour lines are parallel. Since the gradient of a function is perpendicular to the contour lines, this is the same as saying that the gradients of f and g are parallel. Thus you want points *(x,y)* where *g*(*x*,*y*) = *c* and

$$\nabla_{x,y} f = -\lambda \nabla_{x,y} g$$

where

$$\nabla_{x,y} f = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \right)$$

and

$$\nabla_{x,y} g = \left(\frac{\partial g}{\partial x}, \frac{\partial g}{\partial y} \right)$$

are the respective gradients. The constant λ is required because although the two gradient vectors are parallel, the magnitudes of the gradient vectors are generally not equal.

To incorporate these conditions into one equation, you introduce an auxiliary function

$$\Lambda(x,y,\lambda) = f(x,y) + \lambda \cdot \Big(g(x,y) - c \Big),$$

and solve

$$\nabla_{x,y,\lambda}\Lambda(x,y,\lambda)=0.$$

This is the method of Lagrange's multipliers. Note that $\nabla_\lambda\Lambda(x,y,\lambda)=0$ implies $g(x,y)=c$

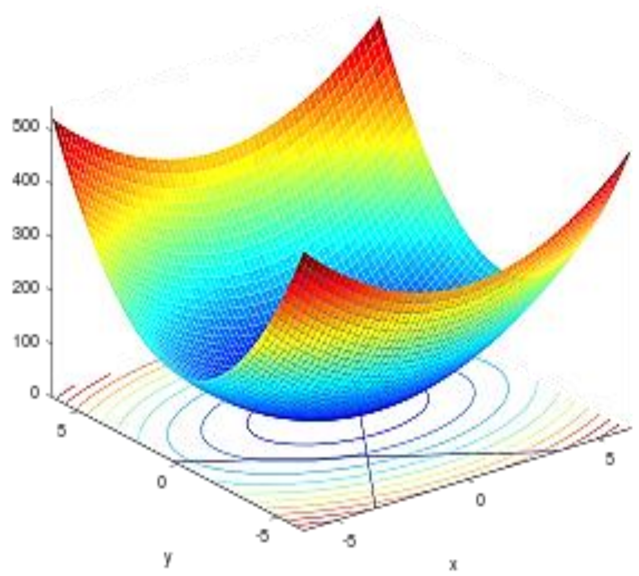
Not necessarily extrema

The constrained extrema of f are critical points of the Lagrangian Λ, but they are not local extrema of Λ (see Example 2 below).

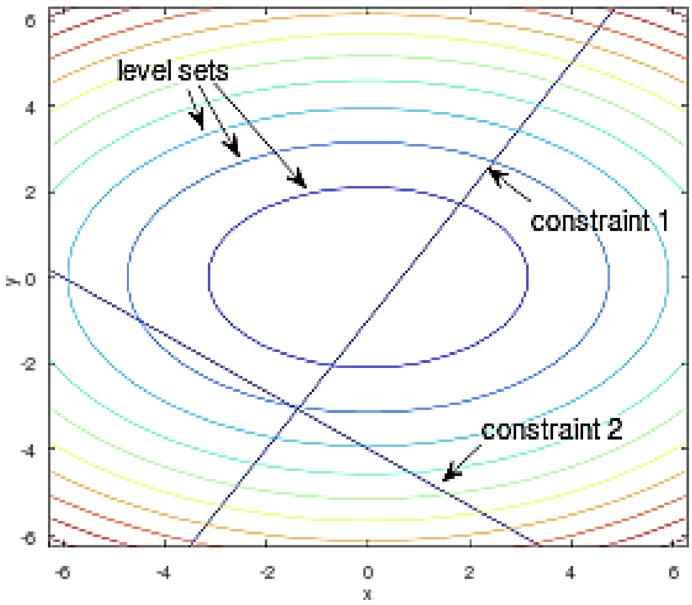
One may reformulate the Lagrangian as a Hamiltonian, in which case the solutions are local minima for the Hamiltonian. This is done in optimal control theory, in the form of Pontryagin's minimum principle.

The fact that solutions of the Lagrangian are not necessarily extrema also poses difficulties for numerical optimization. This can be addressed by computing the magnitude of the gradient, as the zeros of the magnitude are necessarily local minima, as illustrated in the numerical optimization example.

Handling multiple constraints



A paraboloid, some of its level sets (aka contour lines) and 2 line constraints.



Zooming in on the levels sets and constraints, you see that the two constraint lines intersect to form a "joint" constraint that is a point.

Since there is only one point to analyze, the corresponding point on the paraboloid is automatically a minimum and maximum. Yet the simplified reasoning presented in sections above seems to fail because the level set definitely appears to "cross" the point and at the same time its gradient is not parallel to the gradients of either constraint. This shows you must refine our explanation of the method to handle the kinds of constraints that are formed when you have more than one constraint acting at once.

The method of *Lagrange's multipliers* can also accommodate multiple constraints. To see how this is done, you need to reexamine the problem in a slightly different manner because the concept of "crossing" discussed above becomes rapidly unclear when you consider the types of constraints that are created when you have more than one constraint acting together.

As an example, consider a paraboloid with a constraint that is a single point (as might be created if we had 2 line constraints that intersect). The level set (i.e., contour line) clearly appears to "cross" that point and its gradient is clearly not parallel to the gradients of either of the two line constraints. Yet, it is obviously a maximum and a minimum because there is only one point on the paraboloid that meets the constraint.

While this example seems a bit odd, it is easy to understand and is representative of the sort of "effective" constraint that appears quite often when you deal with multiple constraints intersecting. Thus, you take a slightly different approach below to explain and derive the Lagrange's Multipliers method with any number of constants.

Throughout this section, the independent variables will be denoted by $x_1, x_2, \ldots, x_N$ as a group, you will denote them as $p = (x_1\ x_2, \ldots, x_n)$. Also, the function being analyzed will be denoted by and the constraints will be represented by the equations $g_1\ (p) = 0,\ g_2\ (p) = 0,\ \ldots, g_M\ (p) = 0.$

The basic idea remains essentially the same: if we consider only the points that satisfy the constraints (i.e. are *in* the constraints), then a point (*p*, *f* (*p*))is a stationary point (i.e. a point in a "flat" region) of *f* if and only if the constraints at that point do not allow movement in a direction where *f* changes value.

Once you have located the stationary points, you need to do further tests to see if you have found a minimum, a maximum or just a stationary point that is neither.

You start by considering the level set of *f*at (*p*, *f* (*p*)). The set of vectors $\{u_L\}$ the directions in which you can move and still remain in the same level set are the directions where the value of *f* does not change (i.e. the change equals zero). Thus, for every vector *v* in $\{v_L\}$, the following relation must hold:

$$\Delta f = \frac{df}{dx_1}v_{x1} + \frac{df}{dx_2}v_{x2} + \cdots + \frac{df}{dx_N}v_{xN} = 0$$

where the notation v_{xk} above means the *xK*-component of the vector *v*. The equation above can be rewritten in a more compact geometric form that helps our intuition:

$$\underbrace{\begin{bmatrix} \frac{df}{dx_1} \\ \frac{df}{dx_2} \\ \vdots \\ \frac{df}{dx_N} \end{bmatrix}}_{\nabla f} \cdot \underbrace{\begin{bmatrix} v_{x1} \\ v_{x2} \\ \vdots \\ v_{xN} \end{bmatrix}}_{v} = 0 \quad \Rightarrow \quad \nabla f \cdot v = 0$$

This makes it clear that if you are at *p*, then *all* directions from this point that do *not* change the value of *f must be perpendicular* to $\nabla f\,(p)$ the gradient of *f* at *p*).

Now let us consider the effect of the constraints. Each constraint limits the directions that you can move from a particular point and still satisfy the constraint. You can use the same procedure, to look for the set of vectors $\{v_C\}$ containing the directions in which you can move and still satisfy the constraint. As above, for every vector *v* in $\{v_C\}$, the following relation must hold:

$$\Delta g = \frac{dg}{dx_1}v_{x1} + \frac{dg}{dx_2}v_{x2} + \cdots + \frac{dg}{dx_N}v_{xN} = 0 \quad \Rightarrow \quad \nabla g \cdot v = 0$$

From this, you see that at point p, all directions from this point that will still satisfy thisconstraint must be perpendicular to $\nabla g(p)$.

Now you are ready to refine our idea further and complete the method: a point on f is a constrained stationary point if and only if the direction that changes f violates at least one of the constraints. (You can see that this is true because if a direction that changes f did not violate any constraints, then there would a "legal" point nearby with a higher or lower value for f and the current point would then not be a stationary point.)

Single constraint revisited

For a single constraint, you use the statement above to say that at stationary points the direction that changes *f* is in the same direction that violates the constraint. To determine if two vectors are in the same direction, you note that if two vectors start from the same point and are "in the same direction", then one vector can always "reach" the other by changing its length and/or flipping to point the opposite way along the same direction line. In this way, you can succinctly state that two vectors point in the same direction if and only if one of them can be multiplied by some real number such that they become equal to the other. So, for your purposes, you require that:

$$\nabla f(p) = \lambda \nabla g(p) \qquad \Rightarrow \qquad \nabla f(p) - \lambda \nabla g(p) = \mathbf{0}$$

If yu now add another simultaneous equation to guarantee that you only perform this test when you are at a point that satisfies the constraint, you end up with 2 simultaneous equations that when solved, identify all constrained stationary points:

$$\begin{cases} g(p) = 0 & \text{means point satisfies constraint} \\ \nabla f(p) - \lambda \nabla g(p) = 0 & \text{means point is a stationary point} \end{cases}$$

Note that the above is a succinct way of writing the equations. Fully expanded, there are N + 1 simultaneous equations that need to be solved for the N + 1 variables which are λ and $x_1, x_2, \ldots, x_N$: :

$$\begin{aligned} g(x_1, x_2, \ldots, x_N) &= 0 \\ \frac{df}{dx_1}(x_1, x_2, \ldots, x_N) - \lambda \frac{dg}{dx_1}(x_1, x_2, \ldots, x_N) &= 0 \\ \frac{df}{dx_2}(x_1, x_2, \ldots, x_N) - \lambda \frac{dg}{dx_2}(x_1, x_2, \ldots, x_N) &= 0 \\ &\vdots \\ \frac{df}{dx_N}(x_1, x_2, \ldots x_N) - \lambda \frac{dg}{dx_N}(x_1, x_2, \ldots, x_N) &= 0 \end{aligned}$$

Multiple constraints

For more than one constraint, the same reasoning applies. If there is more than one active together, each constraint contributes a direction that will violate it. Together, these "violation directions" form a "violation space", where infinitesimal movement in any direction within the space will violate one or more constraints. Thus, to satisfy multiple constraints you can state (using this new terminology) that at the stationary points, the direction that changes *f* is in the "violation space" created by the constraints acting jointly.

The *violation space* created by the constraints consists of all points that can be reached by adding any combination of scaled and/or flipped versions of the individual violation direction vectors. In other words, all the points that are "reachable" when you use the individual violation directions as the basis of the space. Thus, you can succinctly state that v is in the space defined by $b_1, b_2, \ldots, b_M$ if and only if there exists a set of "multipliers" $\lambda_1, \lambda_2, \ldots, \lambda_M$ such that:

$$\sum_{k=1}^{M} \lambda_k b_k = v$$

which for our purposes, translates to stating that the direction that changes f at p is in the"violation space" defined by the constraints $g_1, g_2, \ldots, g_M$if and only if

$$\sum_{k=1}^{M} \lambda_k \nabla g_k(p) = \nabla f(p) \quad \Rightarrow \quad \nabla f(p) - \sum_{k=1}^{M} \lambda_k \nabla g_k(p) = 0$$

As before, you now add simultaneous equation to guarantee that you only perform this test when you are at a point that satisfies every constraint, you end up with simultaneous equations that when solved, identify all constrained stationary points:

$$\begin{aligned} g_1(p) &= 0 \\ g_2(p) &= 0 \quad \text{these mean the point satisfies all constraints} \\ &\vdots \\ g_M(p) &= 0 \end{aligned}$$

The method is complete now (from the standpoint of solving the

$$\nabla f(p) - \sum_{k=1}^{M} \lambda_k \nabla g_k(p) = 0 \quad \text{this means the point is a stationary point}$$

problem of finding stationary points) but as mathematicians delight in doing, these equations can be further condensed into an even more elegant and succinct form. Lagrange must have cleverly noticed that the equations above look like partial derivatives of some larger scalar function L that takes all the $x_1, x_2, \ldots, x_N$ and all the $\lambda_1, \lambda_2, \ldots, \lambda_M$ as inputs. Next,he might then have noticed that setting every equation equal to zero is exactly what one would have to do to solve for the *unconstrained* stationary points of that larger function. Finally, he showed that a larger function L with partial derivatives that are exactly the ones you require can be constructed very simply as below:

$$L(x_1, x_2, \ldots, x_N, \lambda_1, \lambda_2, \ldots, \lambda_M)$$
$$= f(x_1, x_2, \ldots, x_N) - \sum_{k=1}^{M} \lambda_k g_k (x_1, x_2, \ldots, x_N)$$

Solving the equation above for its unconstrained stationary points generates exactly the same stationary points as solving or the constrained stationary points of f under the constraints $g_1, g_2, \ldots, g_M$

In Lagrange's honor, the function above is called a *Lagrangian*, the scalars $\lambda_1, \lambda_2, \ldots, \lambda_M$ are called *Lagrange Multipliers* and this optimization method itself is called *The Method of Lagrange's Multipliers.*

The method of Lagrange's multipliers is generalized by the Karush–Kuhn Tucker conditions, which can also take into account inequality constraints of the form $h(x) \leq c$.

Interpretation of the Lagrange's multipliers

Often the Lagrange's multipliers have an interpretation as some quantity of interest. To see why this might be the case, observe that:

$$\frac{\partial L}{\partial g_k} = \lambda_k.$$

So, λ_k is the rate of change f the quantity being optimized as a function of the constraint variable. As examples, in Lagrangian mechanics the equations of motion are derived by finding stationary points of the action, the time integral of the difference between kinetic and potential energy. Thus, the force on a particle due to a scalar potential, $F = -\nabla V$, can be interpreted as a Lagrange's multiplier determining the change in action (transfer of potential to kinetic energy) following a variation in the particle's constrained trajectory. In economics, the optimal profit to a player is calculated subject to a constrained space f actions, where a Lagrange's multiplier is the increase in the value of the objective function due to the relaxation of a given constraint (e.g. through an increase in income or bribery or other means) – the marginal cost of a constraint, called the shadow price.

In control theory this is formulated instead as costate equations.

Examples

Example 1

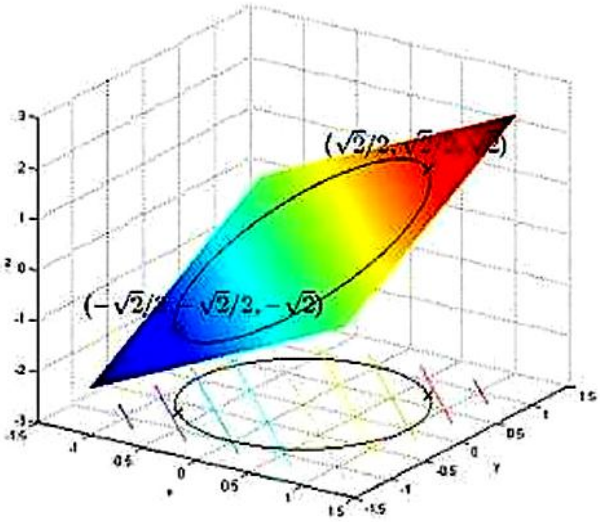


Fig. 3: Illustration of the constrained optimization problem

Suppose you wishes to maximize $f(x,y) = x + y$ subject to the constraint $x^2 + y^2 = 1$. The feasible set is the unit circle, and the level sets of f are diagonal lines (with slope-1), so you can see graphically that the maximum occurs at $(\sqrt{2}/2, \sqrt{2}/2)$, and the minimum occurs $(-\sqrt{2}/2, -\sqrt{2}/2)$,

Formally, set $g(x,y) - c = x^2 + y^2 - 1$, and

$$\Lambda(x,y,\lambda) = f(x,y) + \lambda(g\,x,y) - c) = x + y + \lambda(x^2 + y^2 - 1)$$

Set the derivative $d\Lambda = 0$, which yields the system of equations:

$$\begin{aligned} \frac{\partial \Lambda}{\partial x} &= 1 + 2\lambda x &&= 0, \qquad \text{(i)} \\ \frac{\partial \Lambda}{\partial y} &= 1 + 2\lambda y &&= 0, \qquad \text{(ii)} \\ \frac{\partial \Lambda}{\partial \lambda} &= x^2 + y^2 - 1 &&= 0, \qquad \text{(iii)} \end{aligned}$$

As always, the $\partial\lambda$ equation ((iii) here) is the original constraint.

Combining the first two equations yields $x = y$ (explicitly, $\lambda \neq 0$, otherwise (i) yields 1 = 0, so you have $x = -1/(2\lambda) = y$.

Substituting into (iii) yields $2x^2 = 1$, so $x = y = \pm\sqrt{2/2}$2and

$\lambda = \mp\sqrt{2/2}$2, showing the stationary points are($\sqrt{2/2}$2, $\sqrt{2/2}$2) and

(-$\sqrt{2/2}$2, $-\sqrt{2/2}$2) . Evaluating the objective function f on these yield $f(\sqrt{2/2}2, \sqrt{2/2}2) = \sqrt{2}$ and $f(-\sqrt{2/2}2, -\sqrt{2/2}2) = -\sqrt{2}$, thus the maximum is √2, which is attained at ($\sqrt{2/2}$2, $\sqrt{2/2}$2), and the minimum is - $\sqrt{2}$, which is attained at ($\sqrt{2/2}$2, $-\sqrt{2/2}$2).

Example 2

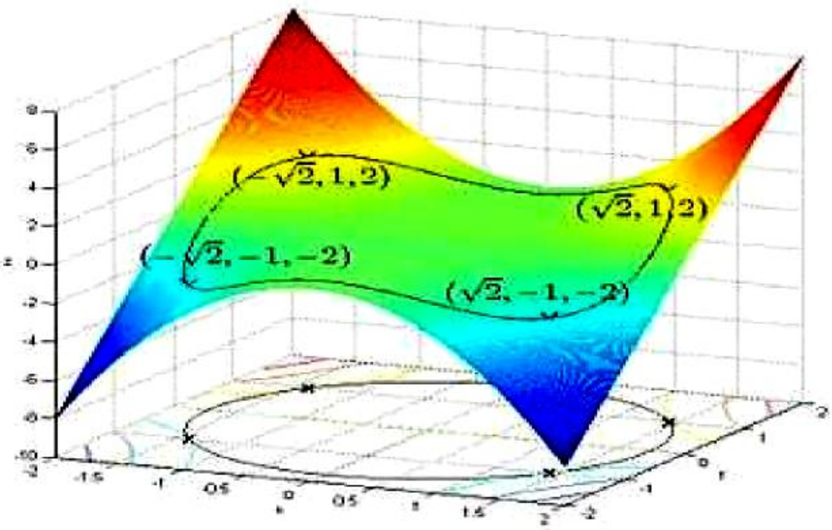


Fig. 4. Illustration of the constrained optimization problem

Suppose you wants to find the maximum values of

$F(x,y) = x^2y$

with the condition that the x and y coordinates lie on the circle around the origin with radius √3, that is, subject to the constraint

$g(x,y) = x^2 + y^2 = 3.$

As there is just a single constraint, you will use only one multiplier, say λ.

The constraint g(x, y)-3 is identically zero on the circle of radius √3. So any multiple of g(x, y)-3 may be added to f(x, y) leaving f(x, y) unchanged in the region of interest (above the circle where our original constraint is satisfied). Let

$$\Lambda(x,y,\lambda) = f(x,y) + \lambda(g(x,y) - 3) = x^2y + \lambda(x^2 + y^2 - 3).$$

The critical values of Λ occur where its gradient is zero. The partial derivatives are

$$\frac{\partial\Lambda}{\partial x} = 2xy + 2\lambda x \quad = 0, \qquad \text{(i)}$$
$$\frac{\partial\Lambda}{\partial y} = x^2 + 2\lambda y \quad = 0, \qquad \text{(ii)}$$
$$\frac{\partial\Lambda}{\partial \lambda} = x^2 + y^2 - 3 \quad = 0. \qquad \text{(iii)}$$

Equation (iii) is just the original constraint. Equation (i) implies x = 0 or $\lambda = -y$. In the first case, if x = 0 then you must have $y = \pm\sqrt{3}$ by (iii) and then by (ii) $\lambda = 0$. In the second case, if $\lambda = -y$ and substituting into equation (ii) you have that,

$x^2 - 2y^2 = 0$.

Then $x^2 = 2y^2$. Substituting into equation (iii) and solving for y gives this value of y:

$y = \pm 1$.

Thus there are six critical points:

$$(\sqrt{2},1); \quad (-\sqrt{2},1); \quad (\sqrt{2},-1); \quad (-\sqrt{2},-1); \quad (0,\sqrt{3}); \quad (0,-\sqrt{3}).$$

Evaluating the objective at these points, you find

$$f(\pm\sqrt{2},1) = 2; \quad f(\pm\sqrt{2},-1) = -2; \quad f(0,\pm\sqrt{3}) = 0.$$

Therefore, the objective function attains the global maximum (subject to the constraints) at $(\pm\sqrt{2},1)$ and the global minimum at $(\pm\sqrt{2}, -1)$. The point $(0,\sqrt{3})$ is a local minimum and $(0,\sqrt{3})$ is a local maximum, as may be determined by consideration of the Hessian matrix of Λ.

Note that while $(\sqrt{2}, 1, -1)$ is a critical point of Λ, it is not a local extremum. You have $\Lambda(\sqrt{2}+\epsilon, 1, -1+\delta) = 2 + \delta(\epsilon^2 + (2\sqrt{2})\epsilon)$ Givenany neighborhood of $(\sqrt{2}, 1, -1)$, you can choose a small positive ∈ and a small δ of either sign to get Λ values both greater and less than 2.

Example: Entropy

Suppose you wish to find the discrete probability distribution on the point $\{x_1, x_2, \ldots, x_n\}$ with maximal information entropy. This is the same as saying that you wish to find the least biased probability distribution on the points $\{x_1, x_3, \ldots, x_n\}$. In other words, you wish to maximize the Shannon entropy equation:

$$f(p_1, p_2, \ldots, p_n) = -\sum_{j=1}^{n} p_j \log_2 p_j.$$

For this to be a probability distribution the sum of the probabilities p_i at each point x_i must equal 1, so our constraint is $g(\vec{p}) - 1$:

$$g(p_1, p_2, \ldots, p_n) = \sum_{j=1}^{n} p_j.$$

$$\vec{p}^*$$

You use Lagrange multipliers to find the point of maximum entropy, $\vec{p}$, across all discrete probability distributions $\vec{p}$ on $\{x_1, x_2, \ldots, x_n\}$. You require that:

$$\frac{\partial}{\partial \vec{p}}(f + \lambda(g-1))\bigg|_{\vec{p}=\vec{p}^*} = 0,$$

which gives a system of n equations, $k = 1,\ldots,n$, such that:

$$\frac{\partial}{\partial p_k}\left(-\sum_{j=1}^{n} p_j \log_2 p_j + \lambda\left(\sum_{j=1}^{n} p_j - 1\right)\right)\bigg|_{p_k=p_k^*} = 0.$$

Carrying out the differentiation of these n equations, you get

$$-\left(\frac{1}{\ln 2} + \log_2 p_k^*\right) + \lambda = 0.$$

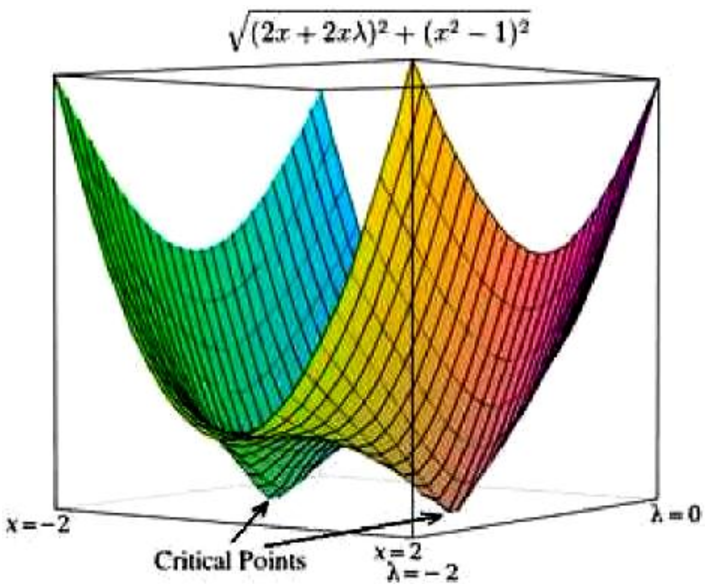
This shows that all p_k^* are equal (because they depend on λ only). By using the constraint $\Sigma_j pj = 1$, you find

$$p_k^* = \frac{1}{n}.$$

Hence, the uniform distribution is the distribution with the greatest entropy, among distributions on n points.

Example: Numerical optimization

Lagrange’s multipliers cause the critical points to occur at saddle points.



The magnitude of the gradient can be used to force the critical point to occur at local minima.

With Lagrange’s multipliers, the critical points occur at saddle points, rather than at local maxima (or minima). Unfortunately, many numerical optimization techniques, such as hill climbing, gradient descent, some of the quasi-Newton methods, among others, are designed to find local maxima (or minima) and not saddle points. For this reason, you must either modify the formulation to ensure that it's a minimization problem (for example, by extremizing the square of the gradient of the Lagrangian as below), or else use an optimization technique that finds stationary points (such as Newton's method without an extremum seeking line search) and not necessarily extrema.

As a simple example, consider the problem of finding the value of x that minimizes $f(x) = x^2$, constrained such that $x^2 = 1$. (This problem is somewhat pathological because there are only two values that satisfy this constraint, but it is useful for illustration purposes because the corresponding unconstrained function can be visualized in three dimensions.)

Using Lagrange’s multipliers, this problem can be converted into an unconstrained optimization problem:

$\Lambda(x,\lambda) = x^2 + \lambda(x^2 - 1)$

The two critical points occur at saddle points where $x = 1$ and $x = -1$.

In order to solve this problem with a numerical optimization technique, you must first transform this problem such that the critical points occur at local minima. This is done by computing the magnitude of the gradient of the unconstrained optimization problem.

First, you compute the partial derivative of the unconstrained problem with respect to each variable:

$$\frac{\partial \Lambda}{\partial x} = 2x + 2x\lambda$$

$$\frac{\partial \Lambda}{\partial \lambda} = x^2 - 1$$

If the target function is not easily differentiable, the differential with respect to each variable can be approximated as

$$\frac{\partial \Lambda}{\partial x} \approx \frac{\Lambda(x+\epsilon, \lambda) - \Lambda(x, \lambda)}{\epsilon}$$

$$\frac{\partial \Lambda}{\partial \lambda} \approx \frac{\Lambda(x, \lambda+\epsilon) - \Lambda(x, \lambda)}{\epsilon}$$

where $\in$ is a small value.

Next, you compute the magnitude of the gradient, which is the square root of the sum of the squares of the partial derivatives:

$$h(x,\lambda) = \sqrt{(2x+2x\lambda)^2 + (x^2-1)^2} \approx \sqrt{\left(\frac{\Lambda(x+\epsilon,\lambda) - \Lambda(x,\lambda)}{\epsilon}\right)^2 + \left(\frac{\Lambda(x,\lambda+\epsilon) - \Lambda(x,\lambda)}{\epsilon}\right)^2}$$

(Since magnitude is always non-negative, optimizing over the squared-magnitude is equivalent to optimizing over the magnitude. Thus, the "square root" may be omitted from these equations with no expected difference in the results of optimization.)

The critical points of h occur at x = 1 and x = − 1, just as in Λ. Unlike the critical points in Λ, however, the critical points in h occur at local minima, so numerical optimization techniques can be used to find them.

4.0 CONCLUSION

In this unit, you have studied how to identify problem which could be solve by langrage's multiplier. You studied single and multiple constraints. You have studied the interpretation of Lagrange's multiplier. You could solve problems with the use of Lagrange's multiplier.

5.0 SUMMARY

In this unit, you have:

- identified problems which could be solved by Lagrange’s multipliers
- known single and multiple constraints
- known the interpretation of Lagrange’s multiplier
- solved problems with the use of Langrage’s multiplier

Problems

Problem 1. Let $h(x, y) = x^2 + .3y^2 + .4y + 1$ be our objective function. (Note that the coefficients are decimals 0 .3 and 0 .4 and not 3 and 4.) Let and the ellipse *g(x, y)* = 1 be your constraint. Find the maximum and the minimum values of *h(x, y)* subject to *g(x, y)* = 1 following the steps below.

(a) Plot the 3d graph of the function *h(x, y)*, the ellipse *g(x, y)* = 1 in the xy-plane and the curve on the graph z = *h(x, y)*, corresponding to the values of *h(x, y)*, in one coordinate system. Use a parametric representation of the ellipse that you should know from last semester. How many solutions you will expect the Lagrangian system of equations to have, explain your reasoning.

(b) Define the Lagrangian function for the optimization problem and set up the corresponding system of equations.

(c) Find solutions to the system using the solve command. Check that you didn't obtain any extraneous solutions. Is the number of solutions what you expected?

(d) Using results of (c), find the minimum and the maximum values of *h(x, y)* subject to the constraint *g(x, y)* = 1.

6.0 TUTOR-MARKED ASSIGNMENT

1. Find the maximum and minimum of *f(x, y)* = 5x – 3y subject to the constraint $x^2 + y^2 = 136$

2. Find the maximum and minimum values of *f(x, y, z)* = xyz subject to the constraint $x + y + z = 1$. Assume that $x, y, z, \geq 0$

3. Find the maximum and minimum values of $f(x,y) = 4x^2 + 10y^2$ on the disk $x^2 + y^2 \leq 4$

4. Find the maximum and minimum of *f(x, y, z)* = 4y – 2z subject to the constraints $2x – y – z = 2$ and $x^2 + y^2 = 1$

7.0 REFERENCES/FURTHER READING

Bertsekas, Dimitri P. (1999). Nonlinear Programming (Second ed.). Cambridge, MA.: Athena Scientific. ISBN 1-886529-00-0.

Grundlehren der MathematischenWissenschaften [Fundamental Principles of Mathematical Sciences]. 306. Berlin: Springer-Verlag. pp. 136–193 (and Bibliographical comments on pp. 334–335).ISBN 3-540-56852-2.

Hiriart-Urruty, Jean-Baptiste; Lemaréchal, Claude (1993)."XII Abstract duality for practitioners". Convex analysis and minimization algorithms, Volume II: Advanced theory and bundle methods.

Lasdon, Leon S. (1970). Optimization theory for large systems.Macmillan series in operations research. New York: The Macmillan Company. pp. xi+523. MR337317.

Lasdon, Leon S. (2002). Optimization theory for large systems (reprint of the 1970 Macmillan ed.). Mineola, New York: Dover Publications, Inc..pp. xiii+523. MR1888251.

Lemaréchal, Claude (2001). "Lagrangian relaxation".In Michael Jünger and Denis Naddef. Computational combinatorial optimization: Papers from the Spring School held in SchloßDagstuhl, May 15–19, 2000. Lecture Notes in Computer Science. 2241. Berlin: Springer-Verlag. pp. 112–156. doi:10.1007/3-540-45586-8_4. ISBN 3-540-42877-1.MRdoi:[http://dx.doi.org/10.1007%2F3-540-45586-8_4 10.1007/3-540-45586-8_4 1900016.[[Digital object identifier|doi]]:[http://dx.doi.org/10.1007%2F3-540-45586-8_4 10.1007/3-540-45586-8_4]].

Vapnyarskii, I.B. (2001), "Lagrange multipliers", in Hazewinkel, Michiel, Encyclopedia of Mathematics, Springer, ISBN 978-1556080104, http://www.encyclopediaofmath.org/index.php?title=Lagrange_multipliers.

UNIT 3 APPLICATIONS OF LANGRANGES MULTIPLIER

CONTENTS

1.0 INTRODUCTION

Optimization problems, which seek to minimize or maximize a real function, play an important role in the real world. It can be classified into unconstrained optimization problems and constrained optimization problems. Many practical uses in science, engineering, economics, or even in our everyday life can be formulated as constrained optimization problems, such as the minimization of the energy of a particle in physics; how to maximize the profit of the investments in economics. In unconstrained problems, the stationary points theory gives the necessary condition to find the extreme points of the objective function f (x_1; ¢ ¢ ¢ ;x_n). The stationary points are the points where the gradient *rf* is zero, that is each of the partial derivatives is zero. All the variables in f (x_1; ¢ ¢ ¢ ;x_n) are independent, so they can be arbitrarily set to seek the extreme of f. However when it comes to the constrained optimization problems, the arbitration of the variables does not exist. The constrained optimization problems can be formulated into the standard form.

2.0 OBJECTIVES

At the end of this unit, you should be able to:

- apply the lagranges multiplier on a pringle surface;
- apply Lagrange's multiplier on Economics;
- apply Lagrange's multiplier on control theory; and
- solve problems with the application of Lagrange's multiplier.

3.0 MAIN CONTENT

There are many cool applications for the Lagrange's multiplier method. For example, you will show how to find the extrema on the world famous Pringle surface. The Pringle surface can be given by the equation

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

Let us bound this surface by the unit circle, giving us a very happy pringle. :) In this case, the boundary would be

$$G(x,y) = x^2 + y^2 - 1$$

The first step is to find the extrema on an unbounded f.

Economics

Constrained optimization plays a central role in economics. For example, the choice problem for a consumer is represented as one of maximizing a utility function subject to a budget constraint. The Lagrange's multiplier has an economic interpretation as the shadow price associated with the constraint, in this example the marginal utility of income.

Control theory

In optimal control theory, the Lagrange's multipliers are interpreted as costate variables, and Lagrange's multipliers are reformulated as the minimization of the Hamiltonian, in Pontryagin's minimum principle.

Example 1 Find the dimensions of the box with largest volume if the total surface area is 64 cm^2.

You first need to identify the function that you're going to optimize as well as the constraint. Let's set the length of the box to be *x*, the width of the box to be y and the height of the box to be *z*. Let's also note that because you're dealing with the dimensions of a box it is safe to assume that *x*, *y*, and z are all positive quantities.

You want to find the largest volume and so the function that you want to optimize is given by,

$f(x,y,z) = xyz$ Next you know that the surface area of the box must be a constant 64. So this is the constraint. The surface area of a box is simply the sum of the areas of each of the sides so the constraint is given by,
$2xy + 2xz + 2yz = 64 \quad \Rightarrow xy + xz + yz = 32$

Note that you divided the constraint by 2 to simplify the equation a little. Also, you get the function $g(x, y, z)$ from this.

$$g(x,y,z) = xy + xz + yz$$

Here are the four equations that you need to solve.

(1) $yz = \lambda(y + z)$ $\qquad (f_x = \lambda g_x)$

(2) $xz = \lambda(x+z)$ $\left(f_y = \lambda g_y\right)$
$xy + xz + yz = 32$ $(g(x,y,z) = 32)$

(3) $xy = \lambda(x+y)$ $(f_z = \lambda g_z)$

(4) There are many ways to solve this system. You'll solve it in the following way. Let's multiply equation (1) by x, equation (2) by y and equation (3) by z. This gives,

$xyz = \lambda x(y+z)$

(5) $xyz = \lambda y(x+z)xyz = \lambda y(x+z)$

(6) $xyz = \lambda z(x+y)$

(7) Now notice that you can set equations (5) and (6) equal. Doing this gives,

$$\lambda x(y+z)xyz = \lambda y(x+z)$$

$\lambda(xy + xz) - \lambda(yx + yz) = 0$

$\lambda(xz - yz) = 0 \quad \Rightarrow \quad \lambda = 0 \quad \text{or} xz = yz$

This gave two possibilities. The first, $\lambda = 0\lambda = 0\lambda = 0$ is not possible since if this was the case equation (1) would reduce to

$yz = 0 \quad \Rightarrow \quad y = 0 \, \text{or} z = 0$

Since you are talking about the dimensions of a box neither of these are possible so you can discount $\lambda = 0$. This leaves the second possibility.

$xz = yz$

Since you know that$z \neq 0$(again since you are talking about the dimensions of a box) you can cancel the z from both sides. This gives,
$x = y$ (8)

Next, let's set equations (6) and (7) equal. Doing this gives,

$$\lambda y(x+z) = \lambda z(x+y)$$

$$\lambda(yx + yz - zx - zy) = 0$$

$$\lambda(yx - zx) = 0 \qquad \Rightarrow \qquad \lambda = 0 \text{ or } yx = zx$$

As already discussed you know that $\lambda = 0$ won't work and so this leaves,

$$yx = zx$$

You can also say that $z \neq 0$ since you are dealing with the dimensions of a box so you must have,

(9) $$z = y$$

Plugging equations (8) and (9) into equation (4) you get,

$$y^2 + y^2 + y^2 = 3y^2 = 32 \qquad y = \pm\sqrt{\frac{32}{3}} = \pm 3.266$$

However, you know that y must be positive since you are talking about the dimensions of a box. Therefore the only solution that makes physical sense here is

$$x = y = z = 3.266$$

So, it looks like you've got a cube here.

You should be a little careful here. Since you've only got one solution you might be tempted to assume that these are the dimensions that will give the largest volume. The method of Lagrange's Multipliers will give a set of points that will either maximize or minimize a given function subject to the constraint, provided there actually are minimums or maximums.

The function itself $f(x, y, z) = xyz$ will clearly have neither minimums nor maximums unless you put some restrictions on the variables. The only real restriction that you've got is that all the variables must be positive. This, of course, instantly means that the function does have a minimum, zero.

The function will not have a maximum if all the variables are allowed to increase without bound. That however, can't happen because of the constraint,

$$xy = xz = yz = 32$$

Here you've got the sum of three positive numbers (because x, y, and z are positive) and the sum must equal 32. So, if one of the variables gets very large, say x, then because each of the products must be less than 32 both y and z must be very small to make sure the first two terms are less than 32. So, there is no way for all the variables to increase without bound and so it should make some sense that the function, $f(x, y, z) = xyz$, will have a maximum.

This isn't a rigorous proof that the function will have a maximum, but it should help to visualize that in fact it should have a maximum and so you can say that you will get a maximum volume if the dimensions are:$x = y = z = 3.266$

Notice that you never actually found values for λ in the above example. This is fairlystandard for these kinds of problems. The value of λ isn't really important to determining if the point is a maximum or a minimum so often you will not bother with finding a value for it. On occasion you will need its value to help solve.

Example 2
Find the maximum and minimum of $f(x, y, z) = 5x - 3y$ subject to the constraint

$x^2 + y^2 = 136$

Solution

This one is going to be a little easier than the previous one since it only has two variables. Also, note that it's clear from the constraint that region of possible solutions lies on a disk of radius $\sqrt{136}$ which is a closed and bounded region and hence by the Extreme Value Theorem you know that a minimum and maximum value must exist.

Here is the system that you need to solve.

$$5 = 2\lambda x$$
$$-3 = 2\lambda y$$
$$x^2 + y^2 = 136$$

Notice that, as with the last example, you can't have $\lambda = 0$ since that would not satisfy thefirst two equations. So, since you know that $\lambda \neq 0$, you can solve the first two equationsfor x and y respectively.

This gives,

$$x = \frac{5}{2\lambda} \qquad y = \frac{3}{2\lambda}$$

Plugging these into the constraint gives,

$$\frac{25}{4\lambda^2} + \frac{9}{4\lambda^2} = \frac{17}{2\lambda^2} = 136$$

We can solve this forλ

$$\lambda^2 = \frac{1}{16} \qquad \Rightarrow \qquad \lambda = \pm\frac{1}{4}$$

Now, that you know λ you can find the points that will be potential maximums and/orminimums.

If $\lambda = \pm\frac{1}{4}$ you get, $x = -10, \quad y = 6$

and if

$x = 10, \qquad y = -6$

To determine if you have maximums or minimums you just need to plug these into the function. Also recall from the discussion at the start of this solution that you know these will be the minimum and maximums because the Extreme Value Theorem tells you that minimums and maximums will exist for this problem.

Here are the minimum and maximum values of the function.

$f(-10,6) = -68$ Minimum at $(-10,6)$

$f(10,-6) = 68$ Maximum at $(10,-6)$

Example 3

- Set up equations for the volume and the cost of building the silo.
- Using the Lagrange's multiplier method, find the cheapest way to build the silo.
- Do these dimensions seem reasonable? Why?

Next, you will look at the cost of building a silo of volume 1000 cubic meters. The curved surface on top of the silo costs \$3 per square meter to build, while the walls cost \$1 per square meter.

Of course, if all situations where this simple, there would be no need for the Lagrange's multiplier method, since there are other methods for solving 2 variable functions that are much nicer. However, with a greater number of variables, the Lagrange's multiplier method is much more fun.

For the next example, imagine you are working at the State Fair (since you're so desperate for money that you can't even buy a bagel anymore). You find yourself at the snowcone booth, and your boss, upon hearing that you are good at math, offers you a bonus if you can design the most efficient snowcone. You assume the snowcone will be modelled by a half-ellipsoid perched upon a cone.

Your boss only wants to use 84 square centimeters of paper per cone, and wants to have it hold the maximum amount of snow. This can be represented in 3 variables: **r** (the radius of the cone), **h** (the height of the cone), and s (the height of the half-ellipsoid). In order to keep the snow from tumbling off the cone, **s** cannot be greater than 1.5*r. You have provided hints for the equations if you need them.

4.0 CONCLUSION

In this unit, you should be able to apply the Lagrange's multiplier on a pringle surface, apply Lagrange's multiplier on Economics, apply Lagrange's multiplier on control theory and solve problems with the application of lagrange multiplier

5.0 SUMMARY

The Lagrange's multipliers method is a very sufficient tool for the nonlinear optimization problems which are capable of dealing with both equality constrained and inequality constrained nonlinear optimization problems. Many computational programming methods, such as the barrier and interior point method, penalizing and augmented Lagrange method, The Lagrange's multipliers method and its extended methods are widely applied in science, engineering, economics and our everyday life.

6.0 TUTOR-MARKED ASSIGNMENT

1. Find the dimensions of the box with largest volume if the total surface area is 64 cm^2.

2. Consider two curves on the xy-plane: $y = e^x$ and $y = -(x-2)^2$. Find two points (x,y),(X,Y) on each of the two curves, respectively, whose distance apart is as small as possible. Use the method of Lagrange's multipliers. Make a graph that illustrates your solution.

3. Find the maximum and minimum values of $f(x,y,z) = xyz$ subject to the constraint $x + y + z = 1$ Assume that $x, y, z \geq 0$

4. Find the maximum and minimum values of $f(x,y) = 4x^2 + 10y^2$ on the disk $x^2 + y^2 \leq 4$

5. Find the maximum and minimum of $f(x,y,z) = 4y - 2z$ subject to the constraints $2x - y - z = 2$ and $x^2 + y^2 = 1$

7.0 REFERENCES/FURTHER READING

G. B. Arfken and H. J. Weber, Mathematical Methods for Physicists (Elsevier academic, 2005), 5th ed.

N. Schofield, Mathematical Methods in Economics and Social Choice (Springer, 2003), 1st ed.

D. P. Bertsekas, Constrained Optimization and Lagrange Multiplier Methods (Academic Press, 1982), 1st ed.

R. Courant, Differential and Integral Calculus (Interscience Publishers, 1937), 1st ed.

D. P. Bertsekas, Nonlinear Programming (Athena Scientific, 1999), 2nd ed.

M. Crow, Computational Methods for Electric Power Systems (CRC, 2003), 1st ed.