حامعة جنوب الوادى

كلية التربية بالغردقة

الفرقة الرابعة عام رياضيات (Math)

المادة : (Pure 13) جزء (Operations Research)

إستاذ المادة / د. اسماعيل جاد امين

الفصل الدراسي الأول

First Course in Operations Research

Content:

- 1 Problem formulation of operation research
- 2 Modeling Life and scientific problems
- One-dimensional minimization methods
- 3 Linear programming problem
- 4 Geometry of linear programming problems
- Classical optimization techniques
- Multi-variable optimization with no constrains
- Multi-variable optimization with equality constraints
- Unconstrained multi-variable optimization methods
    - Direct methods
    - Indirect methods

Recommended references

1. Wenyu Sun and Ya-Xiang Yuan, Optimization theory and methods: nonlinear programming, 2006 Springer Science & Business Media, LLC.
2. Hamdy Taha, Operations research: An introduction (Eight Edition), 2006, Prentice Hall.
3. Frederick S. Hillier and Gerald J. Lieberman, Introduction to operations research, Seventh edition, 2000, McGraw-Hill.

Chapter 1

Introduction

Operation research, for short OR, is the act of obtaining the best result under given circumstances. Thus, we may have several solutions for a certain problem and our aim is to find the best solution among those solutions which leads to the presentation of the optimization problem.

1 – Problem Formulation

1. Define the quantity to be maximize or minimize.
   This quantity is called objective function.
2. Define the constraints
   Those are the restriction under which we have to solve our problem.
3. Define the non-negative constraints
   We have to be sure that all the variables are of non-negative type. If this is not the case, then we have to modify them as we will see later on in our study.

Examples

Example1:

The Haty shop makes its sandwiches from a combination beef and goat meat. The beef contains 80% meat and 20% fat, and it costs 24 pounds per kilo. The goat meat contains 68% meat and 32% fat, and it costs 18 pounds per kilo. What is the amount of meat from each type must be used in each kilo of meat if it wants to minimize its costs and keep the ratio of fat so that no more than 25%?

Solution 1:

Let $x_1$weight of beef meat and $x_2$weight of goat meat

Objective function is

minimize $z = 24x_1+18x_2$

The constrains

(1) Rate of fat

$$0.20x_1+0.32x_2 \leq 0.25$$

(2) Per kilo

$$x_1+x_2=1$$

Non-negative condition

$$x_1 \geq 0, x_2 \geq 0$$

Thus, the final formula for the linear programming problem is

Minimize $z=24x_1+18x_2$

Subject to

$$0.20x_1+0.32x_2 \leq 0.25$$

$$x_1+x_2=1$$

$$x_1 \geq 0. x_2 \geq 0$$

Example2:

A factory wants in the production of 2 models. The first one needs 3 units of wood; and 3 units of iron; 5 units of aluminum, models II needs a single unit of wood; 8 units of iron; 4 units of aluminum. If you know that the maximum available of wood is 53 units, Steel 127 and 100 for aluminum. Form the mathematical model in the following cases

أ- إذا علم ان السلعة الاولى تعطى ربحاً قدره الوحدة والثانية تعطي وحدتان

ب- إذا علم ان السلعة الأولى تعطي ربحا قدره وحدتان والثانية تعطي وحدة واحدة.

A - If the first model is given a profit of unit and the second 2 units.

B – If the first model gives a profit of two units and the second gives a single unit.

Solution 2:

Let the factory produce x unit of 1^{st} one and y from the 2^{nd}.

Objective function

(a) $Max\ Z = x + 2y$ (b) $Max\ Z = 2x + y$

and the constraints are

For wood;

$$3x + y \leq 53$$

For iron;

$$3x + 8y \leq 127$$

For Aluminum;

$$5x + 4y \leq 100$$

Non-negative condition

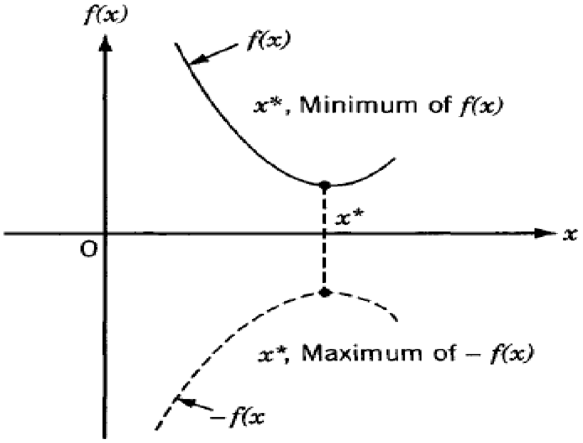
$$x \geq 0\ ,\quad y \geq 0$$

دوجل

Nonlinear Programming I: One-Dimensional Minimization Methods

1. Introduction

It can be seen from the blow figure that if a point $x*$ corresponds to the minimum value of a function $f(x)$, the same point also corresponds to the maximum value of the negative of the function, - $f(x)$.



Optimization can be taken to mean minimization since the maximum of a function can be found by seeking the minimum of the negative of the same function.

2. Statement of an optimization problem

An optimization or a mathematical programming problem can be stated as follows.

$$\text{Find } \mathbf{X} = \begin{Bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{Bmatrix} \text{ which minimizes } f(\mathbf{X})$$

subject to the constraints

$$g_j(\mathbf{X}) \leq 0, \quad j = 1,2,\ldots,m$$
$$l_j(\mathbf{X}) = 0, \quad j = 1,2,\ldots,p \tag{1.1}$$

Where, X is an n-dimensional vector called the design vector, $f(X)$ is termed the objective function, and $g_j(X)$ and $l_j(X)$ are known as inequality and equality constraints, respectively.

- The problem stated in Eq. (1.1) is called a constrained optimization problem.

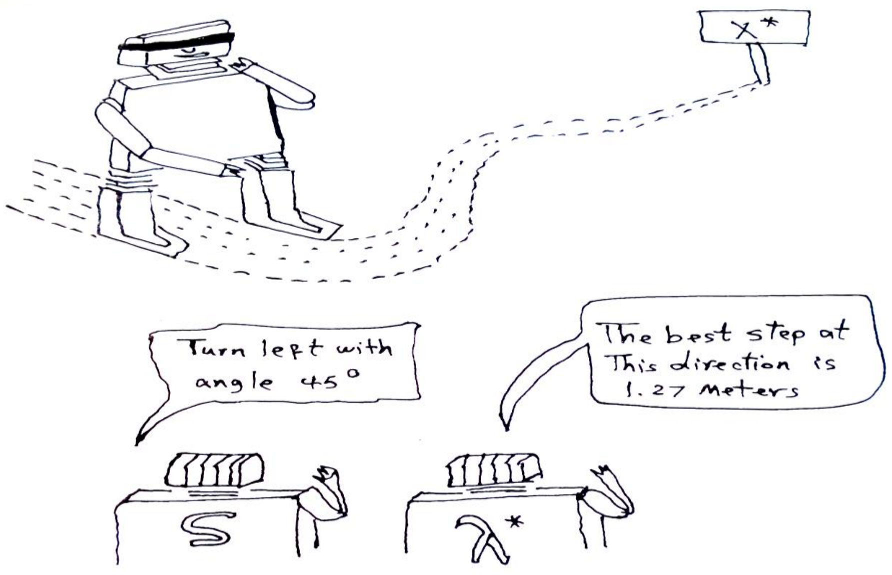
<u>The algorithm that treats a nonlinear programming problem.</u>

1. Start with an initial trial point $\mathbf{X}_1$.
2. Find a suitable direction $\mathbf{S}_i$ ($i = 1$ to start with) which points in the general direction of the optimum.
3. Find an appropriate step length λ_i^* for movement along the direction $\mathbf{S}_i$.
4. Obtain the new approximation $\mathbf{X}_{i+1}$ as

$$\mathbf{X}_{i+1} = \mathbf{X}_i + \lambda_i^* \mathbf{S}_i$$

5. Test whether $\mathbf{X}_{i+1}$ is optimum. If $\mathbf{X}_{i+1}$ is optimum, stop the procedure. Otherwise, set a new $i = i + 1$ and repeat step (2) onward.

- From this algorithm, we conclude that finding a minimum of single variable objective function is an important step (step3) in solving unconstrained multivariable optimization problem. So we start with studying unconstrained single optimization problem



Theorem 1: Necessary Condition

If a function f(x) is defined in the interval *a* <*x* <*b* and have a relative minimum at $x = x^*$, where *a* < x^* <*b,* and if the derivative *df(x)/dx* = $f'(x^*)$ exists as a finite number at $x = x^*$, then $f'(x^*) = 0$.

Theorem 2: Sufficient Condition:

Let $f'(x^*) = f''(x^*) = \ldots = f^{(n-1)}(x^*) = 0$, but $f^{(n)}(x^*) \neq 0$. Then $f(x^*)$ is

(i) A minimum value of $f(x^*)$ if $f^{(n)}(x^*) > 0$ and *n* is even;

(ii) A maximum value of $f(x^*)$ if $f^{(n)}(x^*) < 0$ and *n* is even;

(iii) Neither a maximum nor a minimum if *n* is odd.

<u>Example:</u>

Use theorems 1 and 2 to find the optimum values of

$$f(x) = 12x^5 - 45x^4 + 40x^3 + 5$$

Answer:

$$f'(x) = 60x^4 - 3 * 60x^3 + 60 * 2 * x^2$$
$$= 60x^2(x^2 - 3x + 2)$$
$$= 60x^2(x-1)(x-2) = 0$$

The extreme points are $x = 0, x = 1$ and $x = 2$

$x = 0$	$x = 1$	$x = 2$
$f''(x)= 240x^3 - 540x^2 + 240x$ $f''(0) = 0$ We evaluate the nex derivative $f'''(x)= 3 * 240x^2 - 2 * 540x + 240$ $f'''(0)= +240$, Order of derivative i odd. So this point is neithe maximum nor minimun	$f''(1) = -60$ this point is relative maximum $f_{Max} = 12(1) - 45(1) + 40(1) + 5$ $=12$	$f''(2) = 240$ this point is relative minimum $f_{Min} = -11$

Excercises 3:

Find the maxima and minima, if any, of the functions

(a) $f(x) = \frac{x^4}{(x-1)(x-3)^3}$

(b) $f(x) = 4x^3 - 18x^2 + 27x - 7$

(c) $f(x) = 10x^6 - 48x^5 + 15x^4 + 200x^3 - 120x^2 - 480x + 100$

Answer:

A unimodal function is one that has only one peak in a given interval	T	F
A unimodal function is one that has several peaks in a given interval	T	F
In the **Interval halving method**, the function value at the middle point of the interval will be available in the stage except the first stage	T	F
The interval of uncertainty of the **Interval halving method** remaining at the end of n experiments) $n \geq 3$ and odd) is given by $L_n = \left(\frac{1}{2}\right)^{\frac{(n-1)}{2}} L_0$	T	F
the best value for the eliminating part of the interval in Fibonacci method assuming we conduct a large number of iterations is 0.38٢	T	F

Unconstrained single optimization problem

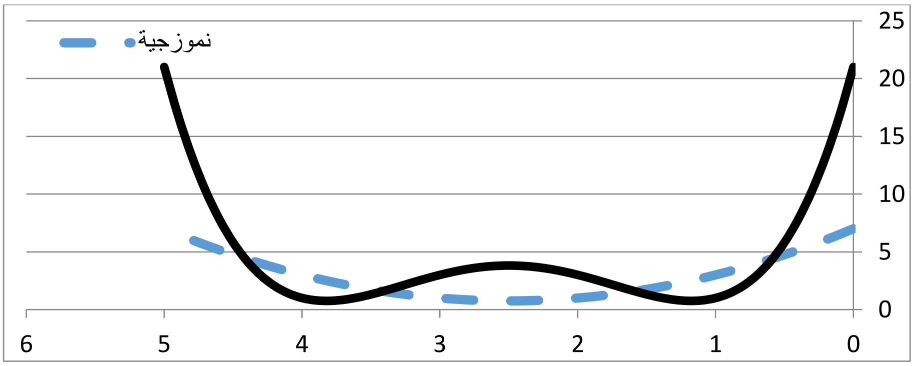
Analytical methods (differential calculus methods)	Numerical methods		
	Elimination methods	Interpolation methods	
		Requiring no derivatives	Requiring derivatives
	Unrestricted search	(quadratic)	Cubic
	Exhaustive search		Direct root
	Dichotomous search		Newton
	Fibonacci method		Quasi-Newton
	Golden section method		Secant

Unimodal function

A unimodal function is one that has only one peak in a given interval

A *unimodal function* is one that has only one peak (maximum) or valley (minimum) in a given interval. Thus a function of one variable is said to be *unimodal* if, given that two values of the variable are on the same side of the optimum, the one nearer the optimum gives the better functional value (i.e., the smaller value in the case of a minimization problem). This can be stated mathematically as follows:

A function $f(x)$ is unimodal if (i) $x_1 < x_2 < x^*$ implies that $f(x_2) < f(x_1)$, and (ii) $x_2 > x_1 > x^*$ implies that $f(x_1) < f(x_2)$, where x^* is the minimum point.



Elimination methods

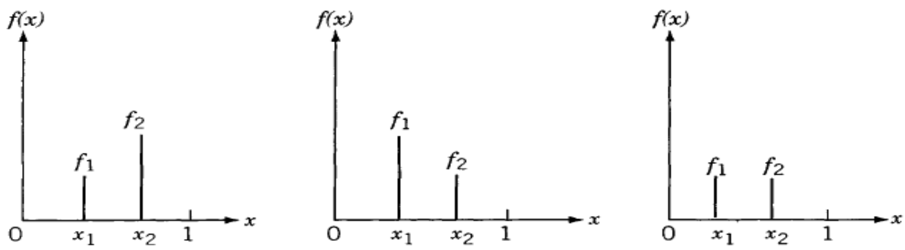
Unrestricted search

In the most practical problems, the optimum solution is known to lie within restricted ranges of the design variables. In some cases this range is not known, and hence the reach has to be made with no restrictions on the values of the variables.

Search with fixed step size

The most elementary approach for such a problem is to use a fixed step size and move from an initial guess points in a favorable direction (positive or negative). The step size used must be small in the relation to the final accuracy desired. Although this method is very simple to implement, it is not efficient in many cases. This method is described in the following steps:

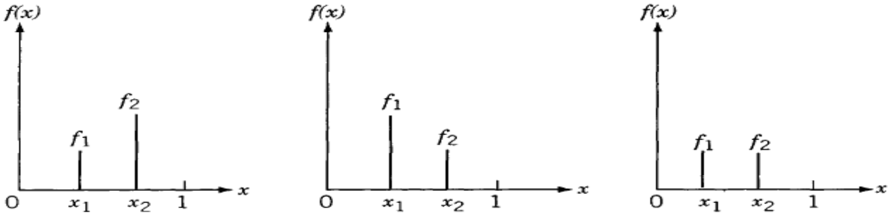
1. Start with an initial guess point, say, x_1.
2. Find $f_1 = f(x_1)$.
3. Assuming a step size s, find $x_2 = x_1 + s$.
4. Find $f_2 = f(x_2)$.
5. If $f_2 < f_1$, and if the problem is one of minimization, the assumption of unimodality indicates that the desired minimum cannot lie at $x < x_1$.



(a) (b) (c)

Hence the search can be continued further along points $x_3, x_4, \ldots$ using the unimodality assumption while testing each pair of experiments. This procedure is continued until a point, $x_i = x_1 + (i - 1)s$, shows an increase in the function value.

6. The search is terminated at x_i, and either x_{i-1} or x_i can be taken as the optimum point.
7. Originally, if $f_2 > f_1$, the search should be carried in the reverse direction at points $x_{-2}, x_{-3}, \ldots$, where $x_{-j} = x_1 - (j - 1)s$.
8. If $f_2 = f_1$, the desired minimum lies in between x_1 and x_2, and the minimum point can be taken as either x_1 or x_2.
9. If it happens that both f_2 and f_{-2} are greater than f_1, it implies that the desired minimum will lie in the double interval $x_{-2} < x < x_2$.



(a) (b) (c)

Example:

Use unrestricted search with Fixed Step Size to Find the maximum of

$$f(x) = \begin{cases} \frac{1}{2}x, & x \le 2 \\ 3 - x, & x > 2 \end{cases}$$

by starting from $x_1 = 0$ **with an initial step size of 0.4.**

Solution: this problem corresponds to Find the minimum of

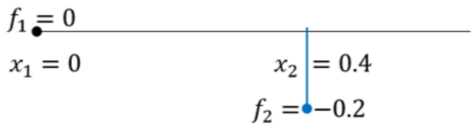
$$f(x) = \begin{cases} -0.5x; & x \le 2 \\ x - 3; & x > 2 \end{cases}$$

$x_1 = 0$, $\quad f(x_1) = f(0) = 0 \; S = 0.4$

$$x_2 = x_1 + S = 0.4$$

$$f(x) = \begin{cases} -\frac{1}{2}x, & x \le 2 \\ 3 - x, & x > 2 \end{cases}$$

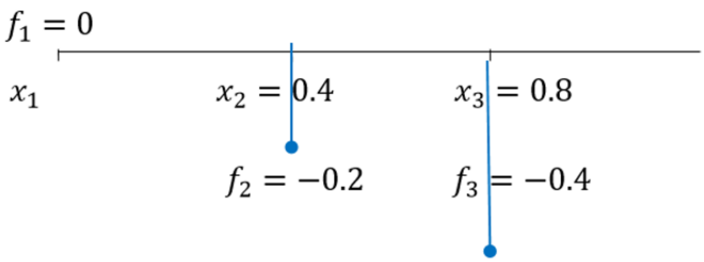
$$f(x_2) = f(0.4) = -\frac{1}{2}(0.4) = -0.2$$



$$x_3 = x_2 + S = 0.4 + 0.4 = 0.8$$

$$f(x) = \begin{cases} -\frac{1}{2}x, & x \le 2 \\ 3 - x, & x > 2 \end{cases}$$

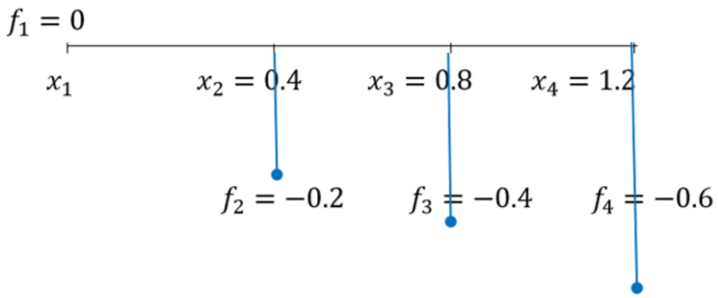
$$f(x_3) = f(0.8) = -\frac{1}{2}(0.8) = -0.4$$



$$x_4 = x_3 + S = 0.8 + 0.4 = 1.2$$

$$f(x) = \begin{cases} -\frac{1}{2}x, & x \le 2 \\ 3 - x, & x > 2 \end{cases}$$

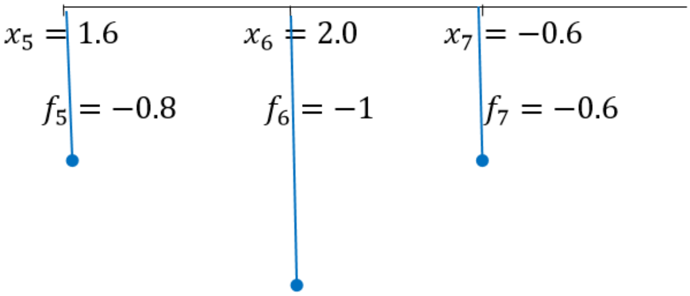
$$f(x_4) = f(1.2) = -\frac{1}{2}(1.2) = -0.6$$



$$x_5 = x_4 + S = 1.2 + 0.4 = 1.6 \qquad f(x_5) = f(1.6) = -\frac{1}{2}(1.6) = -0.8$$

$$x_6 = x_5 + S = 1.6 + 0.4 = 2.0 \qquad f(x_6) = f(2.0) = -\frac{1}{2}(2.0) = -1$$

$$x_7 = x_6 + S = 2.0 + 0.4 = 2.4 \qquad f(x_7) = f(2.4) = 2.4 - 3 = -0.6$$



$\therefore \; x_6 = 2.0$ *is the minimum point and* $f(2.0) = -1$

Fibonacci method

As stated earlier, the *Fibonacci method* can be used to find the minimum of a function of one variable even if the function is not continuous. This method, like many other elimination methods, has the following limitations:

1. The initial interval of uncertainty, in which the optimum lies, has to be known.
2. The function being optimized has to be unimodal in the initial interval of uncertainty.
3. The exact optimum cannot be located in this method. Only an interval known as the *final interval of uncertainty* will be known. The final interval of uncertainty can be made as small as desired by using more computations.
4. The number of function evaluations to be used in the search or the resolution required has to be specified beforehand.

This method makes use of the sequence of Fibonacci numbers, $\{F_n\}$, for placing the experiments. These numbers are defined as

$$F_0 = F_1 = 1$$

$$F_n = F_{n-1} + F_{n-2}, \qquad n = 2,3,4,\ldots$$

which yield the sequence 1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, ...

Procedure. Let L_0 be the initial interval of uncertainty defined by $a \leq x \leq b$ and n be the total number of experiments to be conducted. Define

$$L_2^* = \frac{F_{n-2}}{F_n} L_0$$

and place the first two experiments at points x_1 and x_2, which are located at a distance of L_2^* from each end of L_0.[†] This gives[‡]

$$x_1 = a + L_2^* = a + \frac{F_{n-2}}{F_n} L_0$$

$$x_2 = b - L_2^* = b - \frac{F_{n-2}}{F_n} L_0 = a + \frac{F_{n-1}}{F_n} L_0$$

Discard part of the interval by using the unimodality assumption.

Example:

Use Fibonacci method to find the maximum of

$$f(x) = \begin{cases} \frac{1}{2}x, & x \le 2 \\ 3 - x, & x > 2 \end{cases}$$

by starting from [0,3] **with n=6**

Solution:

This problem corresponds to

Find the minimum of

$$f(x) = \begin{cases} -0.5x; & x \le 2 \\ x - 3; & x > 2 \end{cases}$$

the sequence of Fibonacci numbers, is

$$f_0 = f_1 = 1, \quad 1,1,2,3,5,8,13,21$$

$n = 6$ $\quad f_0 = f_1 = 1, \quad 1,1,2,3,5,8,13,21$

$$L^* = \frac{f_{n-2}}{f_n} L_o = \frac{f_4}{f_6}(3-0) = \frac{5}{13}(3) = 1.15$$

$$x_1 = a + L^* = 0 + 1.15 = 1.15$$

$$x_2 = b - L^* = 3 - 1.15 = 1.85$$

$a = 0 \qquad x_1 = 1.15 \qquad x_2 = 1.85 \qquad b = 3$

$f_a = 0 \qquad f_1 = -0.57 \qquad f_2 = -0.925 \qquad f_b = 0$

Discard part of the interval by using the unimodality assumption.

$$[a = 1.15, b = 3]$$

$a = 0 \qquad x_1 = 1.15 \qquad x_2 = 1.85 \qquad b = 3$

$f_a = 0 \qquad f_1 = -0.57 \qquad f_2 = -0.925 \qquad f_b = 0$

$f_0 = f_1 = 1, \quad 1,1,2,3,5,8,13,21 \qquad n = 5$

$$L^* = \frac{f_{n-2}}{f_n} L_o = \frac{3}{8}(3 - 1.15) = \frac{3}{8}(1.85) = 0.694$$

$$x_1 = a + L^* = 1.15 + 0.694 = 1.84,$$

$$x_2 = b - L^* = 3 - 0.694 = 2.31$$

$a = 1.15 \qquad x_1 = 1.84 \qquad x_2 = 2.31 \qquad b = 3$

$f_a = -058 \qquad f_1 = -0.92 \qquad f_2 = -0.69 \qquad f_b = 0$

Discard part of the interval by using the unimodality assumption.

$$[a = 1.15, b = x_2 = 2.31]$$

$f_0 = f_1 = 1,\ \ 1,1,2,3,5,8,13,21 / \quad n = 4$

$$L^* = \frac{f_{n-2}}{f_n} L_o = \frac{2}{5}(2.31 - 1.15) = \frac{2}{5}(1.16) = 0.464$$

$x_1 = a + L^* = 1.15 + 0.464 = 1.614\,,$

$x_2 = b - L^* = 2.31 - 0.464 = 1.846$

$a = 1.15 \qquad x_1 = 1.614 \qquad x_2 = 1.846 \qquad b = 2.31$

$f_a = -0.57 \quad f_1 = -0.807 \qquad f_2 = -0.923 \quad f_b = -0.69$

Discard part of the interval by using the unimodality assumption.

$$[a = x_1 = 1.614, b = 2.31]$$

$n = 3$

$$L^* = \frac{f_{n-2}}{f_n} L_o = \frac{1}{3}(2.31 - 1.614) = \frac{1}{3}(0.696) = 0.232$$

$x_1 = a + L^* = 1.614 + 0.232 = 1.846\,,$

$x_2 = b - L^* = 2.31 - 0.232 = 2.078$

$a = 1.614 \qquad x_1 = 1.846 \qquad x_2 = 2.078 \qquad b = 2.31$

$f_a = -0.807 \quad f_1 = -0.923 \quad f_2 = -0.922 \quad f_b = -0.69$

n	3	4	5	6	7

$a = 1.614$ $\quad x_1 = 1.846$ $\quad x_2 = 2.078$ $\quad b = 2.31$

$f_a = -0.807$ $\quad f_1 = -0.923$ $\quad f_2 = -0.922$ $\quad f_b = -0.69$

The Last interval is [1.846, 2.31]

Thus the minimum must located at the middle

Hence the minimum is x*=2.078

Golden Section Method

The *golden section method* is same as the Fibonacci method except that in the Fibonacci method the total number of experiments to be conducted has to be specified before beginning the calculation, whereas this is not required in the golden section method. In the Fibonacci method, the location of the first two experiments is determined by the total number of experiments, *n*. In the golden section method we start with the assumption that we are going to conduct a large number of experiments. Of course, the total number of experiments can be decided during the computation.

Example:

Deduce the best value for the eliminating part of the interval in Fibonacci method assuming we conduct a large number of iterations.

$f_0 = f_1 = 1, \quad 1,1,2,3,5,8,13,21$

$$L^* = \frac{f_{n-2}}{f_n} L_o$$

$\frac{f_{n-2}}{f_n}$	$\frac{f_1}{f_3}=\frac{1}{3}$ $=0.33$	$\frac{f_2}{f_4}=\frac{2}{5}$ $=0.4$	$\frac{3}{8}=0.37$	$\frac{5}{13}$ $=0.382$	$\frac{8}{21}$ $=0.382$

$$\lim_{n\to\infty}\frac{f_{n-٢}}{f_n}=٠.٣٨٢$$

The algorithm of method of Golden Section Method

Procedure. The procedure is same as the Fibonacci method except that the location of the first two experiments is defined by

$L^* = 0.382\mathrm{L_o}$

1. Let L_o be the initial interval : $L_o = [a, b]$

2. Define $L^* = 0.382\mathrm{L_o}$

3. Put points of test to be $x_1 = a + L^*, x_2 = b - L^*$

4. Eliminate the non-desired part of the interval depending on the unimodality property

5. define the new interval $L_o = [a, b]$, repeat steps 2-5 until a desired accuracy is obtained.

In step 5, we can use one of the following accuracy formula:

$$|f(x_1) - f(x_2)| \leq \varepsilon$$

Or

$$|L_o| \leq \varepsilon$$

Where ε is small chosen value (such as 0.1).

Exercise:

Use Golden Section method to find the maximum of

$$f(x) = \begin{cases} \frac{1}{2}x, & x \leq 2 \\ 3 - x, & x > 2 \end{cases}$$

By starting from [0,3] with n=6

2

Chapter 4: Linear Programming Problem

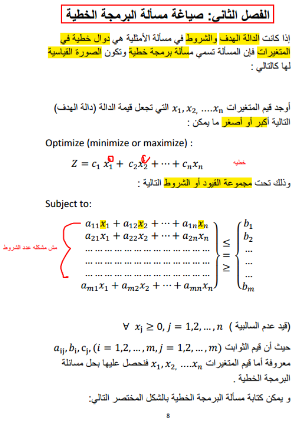
Standard form of a linear programming problem

The general linear programming problem can be stated in the following standard form:

1. *Scalar form*

$$\text{Minimize } f(x_1,x_2,\ldots,x_n) = c_1x_1 + c_2x_2 + \cdots + c_nx_n$$

subject to the constraints


الفصل الثانى: صياغة مسألة البرمجة الخطية

إذا كانت الدالة الهدف والشروط في مسألة الأمثلية هي دوال خطية في المتغيرات فإن المسألة تسمي مسألة برمجة خطية وتكون الصورة القياسية لها كالتالي :

أوجد قيم المتغيرات $x_1, x_2, \ldots x_n$ التي تجعل قيمة الدالة (دالة الهدف) التالية أكبر أو أصغر ما يمكن :

Optimize (minimize or maximize) :

$$Z = c_1 x_1 + c_2 x_2 + \cdots + c_n x_n$$

خطيه

وذلك تحت مجموعة القيود أو الشروط التالية :

Subject to:

مش مشكله عدد الشروط

$$\left.\begin{matrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \cdots \\ \cdots \\ \cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{matrix}\right\} \begin{matrix} \le \\ = \\ \ge \end{matrix} \left\{\begin{matrix} b_1 \\ b_2 \\ \cdots \\ \cdots \\ \cdots \\ b_m \end{matrix}\right.$$

$\forall\ x_j \ge 0, j = 1,2,\ldots,n$ (قيد عدم السالبية)

حيث أن قيم الثوابت $a_{ij}, b_i, c_j, (i = 1,2,\ldots,m, j = 1,2,\ldots,m)$ معروفة أما قيم المتغيرات $x_1, x_2, \ldots x_n$ فنحصل عليها بحل مسائلة البرمجة الخطية .

و يمكن كتابة مسألة البرمجة الخطية بالشكل المختصر التالي:

8

$$a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1$$
$$a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2$$
$$\vdots$$
$$a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m$$
$$x_1 \ge 0$$
$$x_2 \ge 0$$
$$\vdots$$
$$x_n \ge 0$$

where c_j, b_j, and a_{ij} ($i = 1,2,\ldots,m$; $j = 1,2,\ldots,n$) are known constants, and x_j are the decision variables.

2. Matrix form

$$\text{Minimize } f(\mathbf{X}) = \mathbf{c}^T\mathbf{X}$$

subject to the constraints

$$\mathbf{aX} = \mathbf{b}$$

$$\mathbf{X} \geq \mathbf{0}$$

where

$$\mathbf{X} = \begin{Bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{Bmatrix}, \quad \mathbf{b} = \begin{Bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{Bmatrix}, \quad \mathbf{c} = \begin{Bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{Bmatrix},$$

$$\mathbf{a} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

أوجد قيمة $X = \begin{Bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{Bmatrix}$

التي تحقق القيمة الصغري للدالة

$$F(X) = \sum_{i=1}^{n} c_i \; x_i$$

مع تحقق الشروط

$$g(X) = \sum_{j=1}^{n} a_{ij} \; x_j \geq b_i \qquad j = 1,2,\ldots,m$$

$$x_i \geq 0 \qquad i = 1,2,\ldots,n$$

حيث c_i, a_{ij}, b_i جميعها ثوابت.

The characteristics of a linear programming problem, stated in the standard form, are:

1. The objective function is of the minimization type.
2. All the constraints are of the equality type.
3. All the decision variables are nonnegative.

It is now shown that any linear programming problem can be expressed in the standard form by using the following transformations.

How to write any LPP as standard form???

The maximization of a function $f(x_1, x_2, \ldots, x_n)$ is equivalent to the minimization of the negative of the same function. For example, the objective function

$$\text{minimize } f = c_1x_1 + c_2x_2 + \cdots + c_nx_n$$

is equivalent to

$$\text{maximize } f' = -f = -c_1x_1 - c_2x_2 - \cdots - c_nx_n$$

Consequently, the objective function can be stated in the minimization form in any linear programming problem.

2. In most engineering optimization problems, the decision variables represent some physical dimensions, and hence the variables x_j will be nonnegative. However, a variable may be unrestricted in sign in some problems. In such cases, an unrestricted variable (which can take a positive, negative, or zero value) can be written as the difference of two nonnegative variables. Thus if x_j is unrestricted in sign, it can be written as

$$\boxed{x_j = x_j' - x_j'', \text{ where } x_j' \geq 0 \text{ and } x_j'' \geq 0}$$

Slack variable

A non-negative variable which must be added to an inequality constraint of the form $\leq$ to be in an equality form

If a constraint appears in the form of a ''less than or equal to'' type of inequality as

$$a_{k_1}x_1 + a_{k_2}x_2 + \cdots + a_{kn}x_n \leq b_k$$

it can be converted into the equality form by adding a nonnegative slack variable x_{n+1} as follows:

$$a_{k_1}x_1 + a_{k_2}x_2 + \cdots + a_{kn}x_n + \boxed{x_{n+1}} = b_k$$

surplus variable

A non-negative variable which must be subtracted from an inequality constraint of the form $\geq$ to be in an equality form

if the constraint is in the form of a ''greater than or equal to'' type of inequality as

$$a_{k1}x_1 + a_{k2}x_2 + \cdots + a_{kn}x_n \geq b_k$$

it can be converted into the equality form by subtracting a variable as

$$a_{k1}x_1 + a_{k2}x_2 + \cdots + a_{kn}x_n - x_{n+1} = b_k$$

where x_{n+1} is a nonnegative variable known as a *surplus variable*.

Exercises

Detect which of the following Mathematical statements is true and which is false.

1	The Matrix form of the standard form of linear programming problem is Minimize $f(X) = c^T X$, where c is unknown constant.	☐	☐
2	The decision variables in the standard form of linear programming problem must be positive or zero	☐	☐
3	The maximization of a function is equivalent to the minimization of the negative of the same function	☐	☐
4	An unrestricted variable can be written as the difference of two nonnegative variables to agree the standard form of linear programming problem	☐	☐
5	A non negative variable which must be added to an inequality constraint of the form $\leq$ to be in an equality form is called slack variable	☐	☐
6	A non negative variable which must be subtracted from an inequality constraint of the form $\geq$ to be in an equality form is called surplus variable	☐	☐
7	A non negative variable which must be added to an inequality constraint of the form $\leq$ to be in an equality form is called surplus variable	☐	☐
8	A non negative variable which must be subtracted from an inequality constraint of the form $\geq$ to be in an equality form is called slack variable	☐	☐

Classify of LP

Geometry of linear programming problems

The following general geometrical characteristics can be noted from the graphical solution.

1. The feasible region is a convex polygon.
2. The optimum value occurs at an extreme point or vertex of the feasible region.

Example:
Find the set of points that satisfies the following set of inequalities:

$$4x + 5y \leq 33, x + 4y \geq 11, 2x - 3y \geq -11$$

Answer:

We consider the line

$$4x + 5y = 33$$

$$x = 0 \ \rightarrow 5y = 33 \ \rightarrow y = \frac{33}{5} = 6\frac{3}{5} = 6.6$$

$$4x + 5y = 33$$

$$y = 0 \ \rightarrow 4x = 33 \ \rightarrow x = \frac{33}{4} = 8\frac{1}{4} = 8.25$$

$$\left(0\,, \frac{33}{5}\right), \left(\frac{33}{4}, 0\right)$$

(0,0) satisfies $4x + 5y \leq 33$ then this inequality is satisfied by the set of point down and left the line passes through $\left(0\,, \frac{33}{5}\right), \left(\frac{33}{4}, 0\right)$

The line $x + 4y = 11$

$$x = 0 \ \rightarrow 4y = 11 \ \rightarrow y = \frac{11}{4} = 2\frac{3}{4}$$

$$y = 0 \ \rightarrow x = 11$$

The point to that satisfies the inequality are over and on the line

The line $\quad 2x - 3y = -11$

$$x = 0 \rightarrow -3y = -11 \rightarrow y = \frac{11}{3} = 3\frac{2}{3} = 3.67$$

$$y = 0 \rightarrow 2x = -11 \rightarrow x = \frac{-11}{2} = -5\frac{1}{2} = -5.5$$

(0,0) satisfies $2x - 3y \geq -11$ then this inequality is satisfied by the set of point down the line

Solving $4x + 5y = 33$, $x + 4y = 11$,

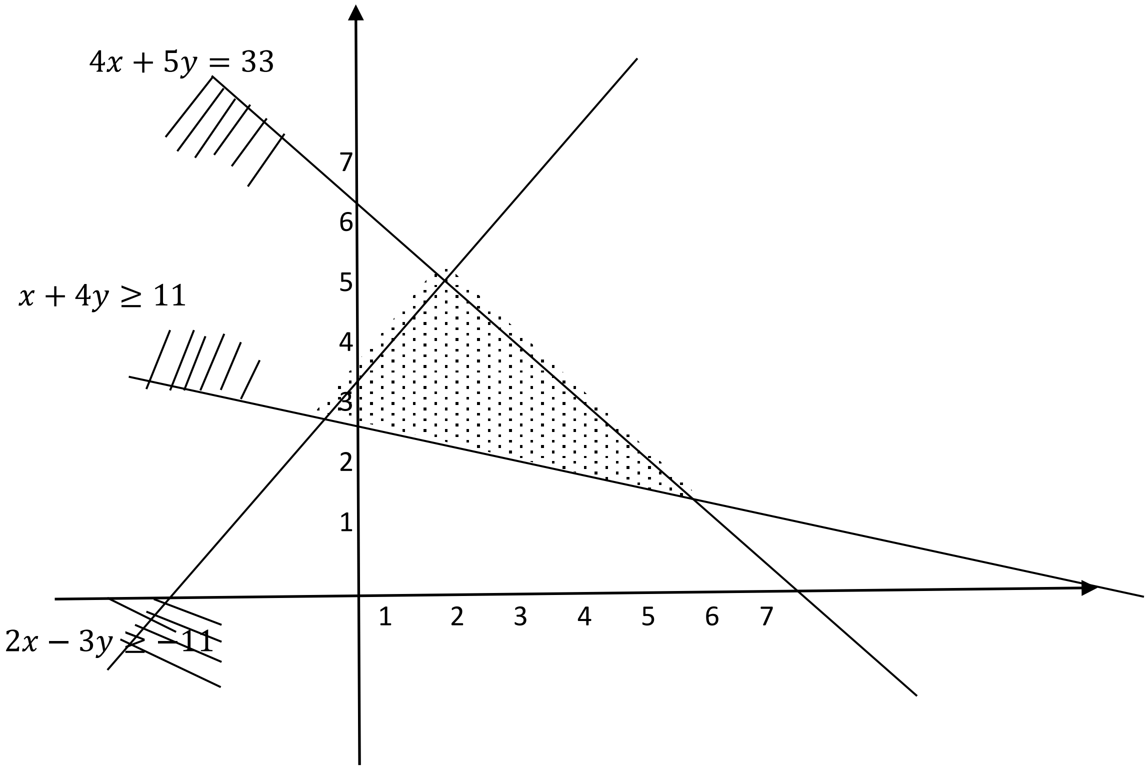
we obtain P$(7,1)$

Solving $4x + 5y = 33$, $2x - 3y = -11$,

we obtain Q$(2,5)$

Solving $x + 4y = 11$, $2x - 3y = -11$

We obtain R$(-1,3)$



The set of points that satisfy the three inequalities are those inside and at the triangle described at the figure ABC.

The optimum value occurs at an extreme point or vertex of the feasible region.

Exercises

Select the correct word

(1) The inequality $4x + 5y \leq 33$ is satisfied by the set of points down and left the line passes through

□	$\left(\mathbf{0},\frac{\mathbf{33}}{\mathbf{5}}\right),\left(\frac{\mathbf{33}}{\mathbf{4}},\mathbf{0}\right)$	□	$\left(\frac{\mathbf{33}}{\mathbf{5}},\mathbf{0}\right),\left(\frac{\mathbf{33}}{\mathbf{4}},\mathbf{0}\right)$	□	$\left(\frac{\mathbf{33}}{\mathbf{5}},\mathbf{0}\right),\left(\mathbf{0},\frac{\mathbf{33}}{\mathbf{4}}\right)$	□	$(\mathbf{1},\mathbf{0}),(\mathbf{0},\mathbf{1})$

(2)The inequality $2x - 3y \geq -11$ is satisfied by the set of point down and left the line passes through

□	$\left(\frac{11}{3},\mathbf{0}\right),\left(\frac{-11}{2},\mathbf{0}\right)$	□	$\left(\mathbf{0},\frac{11}{3}\right),\left(\frac{-11}{2},\mathbf{0}\right)$	□	$\left(\frac{11}{3},\mathbf{0}\right),\left(\mathbf{0},\frac{-11}{2}\right)$	□	$(\mathbf{9},\mathbf{0}),(\mathbf{0},\mathbf{6})$

Example

Find the solution of the following LP problem graphically:

Maximize$f(x, y) = 3x + y + 2$,

Subject to $2x + y + 9 \geq 0$, $3y - x + 6 \geq 0$, $x + 2y \leq 3$, $y \leq x + 3$

Answer:

$$2x + y + 9 = 0$$

$$2x + y = -9 \rightarrow x = 0, y = -9$$

$$y = 0, x = -4.5$$

(0,0) Satisfies it, so the proposed area is up right the line

----

$$3y - x + 6 \geq 0$$

$$3y - x = -6 \rightarrow x = 0, y = -2$$

$$y = 0, x = 6$$

(0,0) Satisfies it, so the proposed area is up Left the line

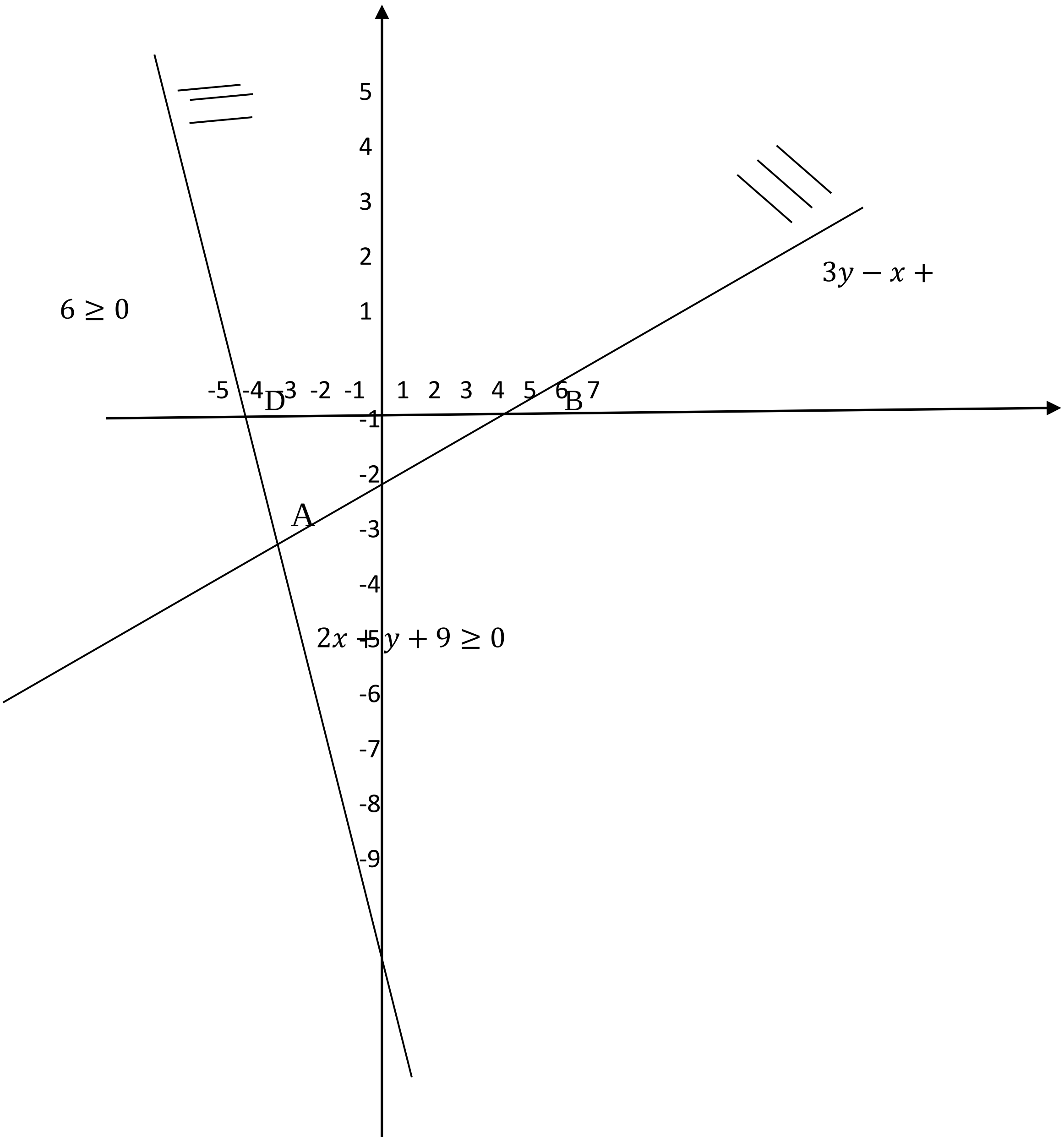
The intersection of

$2x + y = -9,\ 3y - x = -6$

is obtained by solving these two eqs. To obtain

$$x = -3, y = -3$$

$$f(at\ A) = 3(-3) + (-3) + 2 = -10$$



$x + 2y \leq 3$

=============

$x + 2y = 3 \Rightarrow x = 0, y = 1.5$

$y = 0, x = 3$

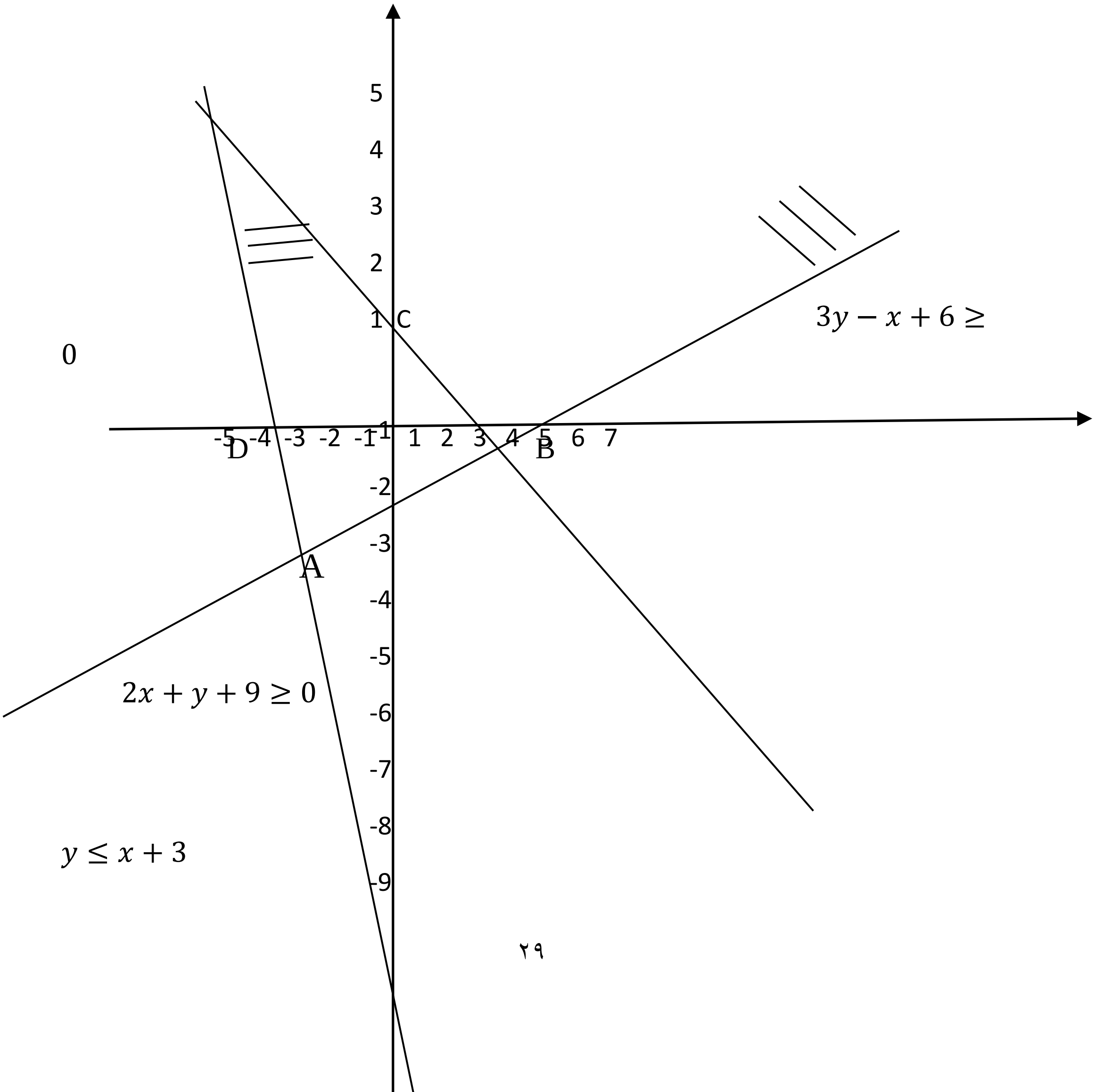
(0,0) satisfies it, so the proposed area is Down Left the line

===========

The intersection of $3y - x = -6, x + 2y = 3$

is obtained by solving these two eqs. To obtain $x = 4.5, y = -0.6$

$f(at\ B) = 3(4.5) + (-0.6) + 2 = 14$



$y - x = 3 \Rightarrow x = 0, y = 3$

$$y = 0, x = -3$$

(0,0) satisfies it, so the proposed area is Down Wright the line

===========

The intersection of $y - x = 3$, $x + 2y = 3$

Is obtained by solving these two eqs. To obtain $x = -1, y = 2$

$$f(at\ C) = 3(-1) + (2) + 2 = 1$$

===========

The intersection of $y + x = 3$, $2x + y = -9$

Is obtained by solving these two eqs. To obtain $x = -4, y = -12$

$$f(x, y) = 3x + y + 2$$

Thus,

$$f_A = -10 \text{ at } A(-3, -3)$$

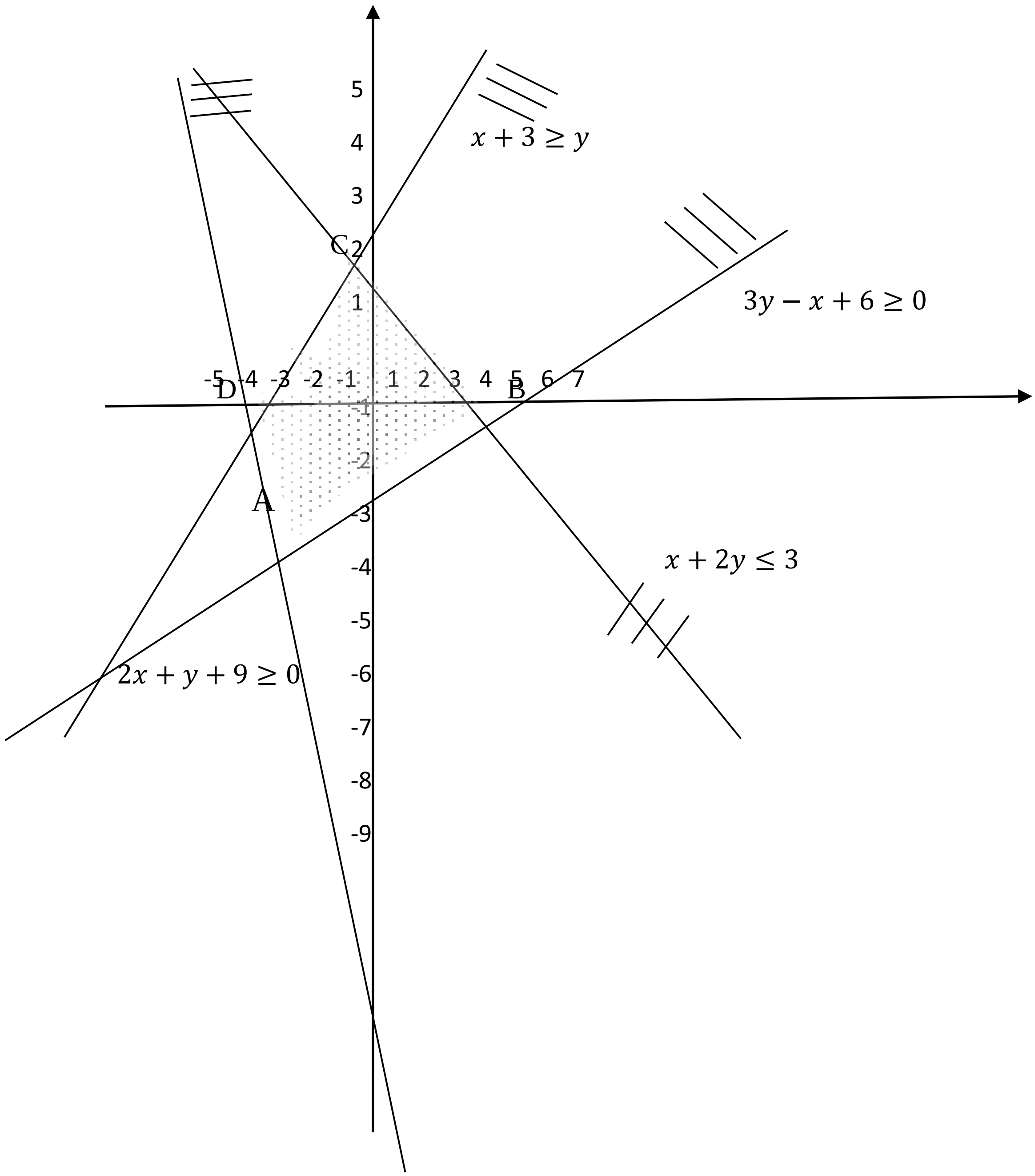
$$f_C = 1 \text{ at } \quad C(-1,2)$$

$$f_B = 14 \text{ at } \quad B(4.2, -0.6)$$

$$f_D = -11 \text{ at } D(-4, -1)$$

Hence the Maximum value is $f_B = 14$ at $B(4.2, -0.6)$

And the Minimum value is $f_D = -11$ at $D(-4, -1)$


$x + 3 \geq y$
$3y - x + 6 \geq 0$
$x + 2y \leq 3$
$2x + y + 9 \geq 0$
A
B
C
D

Exercises

Consider the following graph that represent four inequalities constraints of linear programming problem with the objective function $f(x, y) = 3x + y + 2$. Answer the following:

(1) The Point of intersection A is

□	$(4.2, -0.6)$	□	$(-3, -3)$	□	$(-1,2)$	□	$(\mathbf{1}, \mathbf{0}), (\mathbf{0}, \mathbf{1})$

(2) The Point of intersection B is

□	$(4.2, -0.6)$	□	$(-3, -3)$	□	$(-1,2)$	□	$(\mathbf{1}, \mathbf{0}), (\mathbf{0}, \mathbf{1})$

(3) The Point of intersection C is

□	$(-3, -3)$	□	$(4.2, -0.6)$	□	$(-1,2)$	□	$(-\mathbf{1}, \mathbf{0}), (\mathbf{0}, -\mathbf{1})$

(4) The Point of intersection D is

□	$(-4.2,0.6)$	□	$(-3, -3)$	□	$(4.2, -0.6)$	□	$(-4, -1)$

(5) The Maximum value occurs at

□	$B(4.2, -0.6)$	□	$D(-4, -1)$	□	$C(-1,2)$	□	$(\mathbf{1}, \mathbf{0}), (\mathbf{0}, \mathbf{1})$

(9) the Minimum value occurs at

□	$B(4.2, -0.6)$	□		□		□	$(-\mathbf{1}, \mathbf{0}), (\mathbf{0}, -\mathbf{1})$

Example:

A manufacturing firm produces two machine parts using lathes, milling machines, and grinding machines. The different machining times required for each part, the machining times available on different machines, and the profit on each machine part are given in the following table.

Type of Machine	Machining Time Required (min) Machine Part I	Machine Part II	Maximum Time Available per Week (min)
Lathes	10	5	2500
Milling machines	4	10	2000
Grinding machines	1	1.5	450
Profit per unit	\$50	\$100	

Determine the number of parts I and II to be manufactured per week to maximize the profit.

Solution

Let the number of machine parts I and II manufactured per week be denoted by x and y, respectively.

The constraints due to the maximum time limitations on the various machines are given by

$$10x + 5y \leq 2500 \qquad (E_1)$$

$$4x + 10y \leq 2000 \qquad (E_2)$$

$$x + 1.5y \leq 450 \qquad (E_3)$$

Since the variables x and y cannot take negative values, we have

$$x \geq 0, \quad y \geq 0 \qquad (E_4)$$

The total profit is given by

$$f(x,y) = 50x + 100y \qquad (E_5)$$

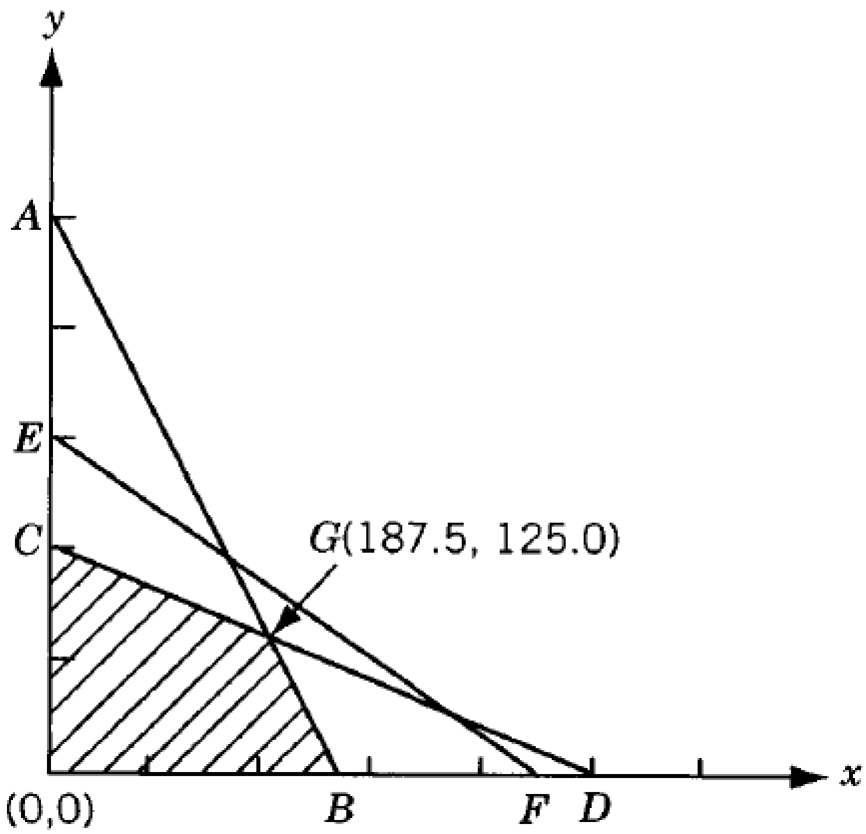
Exercise:
Find the solution of the following LP problem graphically:

Maximize
$$f(x,y) = 50x + 100y$$
subject to
$$10x + 5y \leq 2500$$
$$4x + 10y \leq 2000$$
$$x + 1.5y \leq 450$$

$$x \geq 0, y \geq 0$$

Thus the problem is to determine the nonnegative values of x and y that satisfy the constraints stated in Eqs. (E_1) to (E_3) and maximize the objective function given by Eq. (E_5). The inequalities (E_1) to (E_4) can be plotted in the xy plane and the feasible region identified as shown in Fig. 3.3. Our objective is to find

at least one point out of the infinite points in the shaded region of Fig. which maximizes the profit function (E_5).



Definitions and Theorems

Definitions

1. Point in n-Dimensional Space

$$(x_1, x_2, \ldots, x_n).$$

2. ***Line Segment in n-Dimensions (L)*** If the coordinates of two points A and B are given by $x_j^{(1)}$ and $x_j^{(2)}$ $(j = 1,2,\ldots,n)$, the line segment (L) joining these points is the collection of points $\mathbf{X}$ (λ) whose coordinates are given by $x_j = \lambda x_j^{(1)} + (1 - \lambda)x_j^{(2)}$, $j = 1,2,\ldots,n$, with $0 \leq \lambda \leq 1$.

Thus

$$L = \{\mathbf{X} | \mathbf{X} = \lambda\mathbf{X}^{(1)} + (1 - \lambda)\mathbf{X}^{(2)}\}$$

In one dimension, for example, it is easy to see that the definition is in accordance with our experience (Fig.):

$$x^{(2)} - x(\lambda) = \lambda[x^{(2)} - x^{(1)}], \qquad 0 \le \lambda \le 1$$

Figure Line segment.

whence

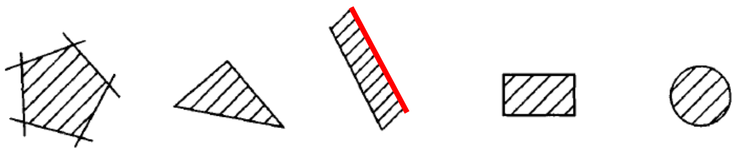
$$x(\lambda) = \lambda x^{(1)} + (1 - \lambda) x^{(2)}, \qquad 0 \le \lambda \le 1$$

Convex Set

4. Convex Set A convex set is a collection of points such that if $\mathbf{X}^{(1)}$ and $\mathbf{X}^{(2)}$ are any two points in the collection, the line segment joining them is also in the collection. A convex set, S, can be defined mathematically as follows:

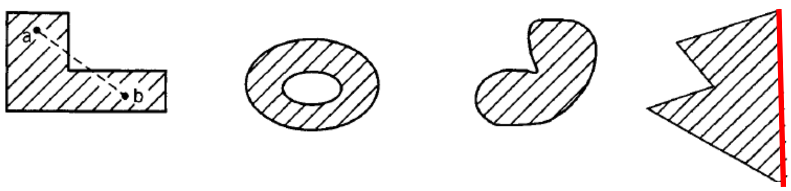
If $X^{(1)}, X^{(2)} \in S$, implies

$$X = \lambda X^{(1)} + (1 - \lambda) X^{(2)} \in S, \quad 0 \le \lambda \le 1.$$



Convex Set

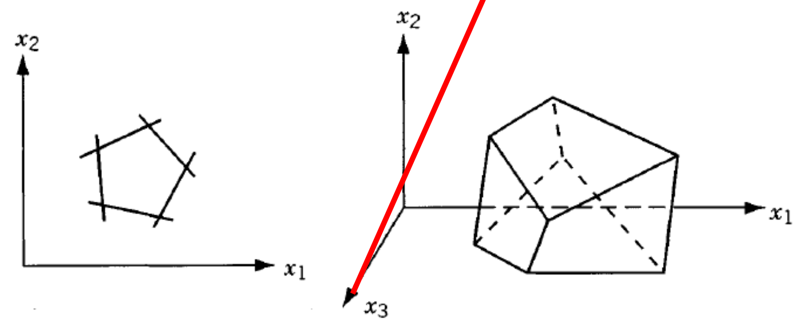
A set containing only one point is always considered to be convex



Non Convex Set

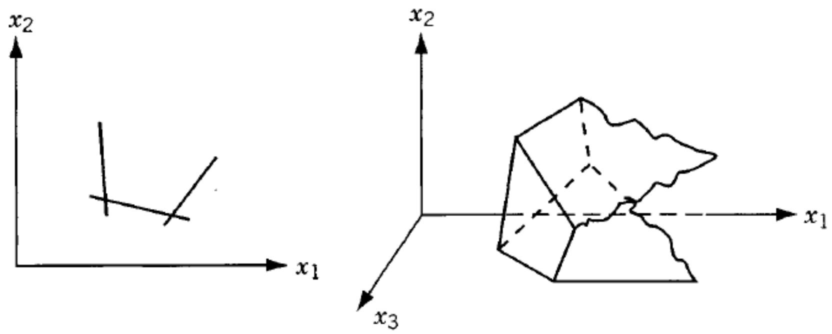
Convex Polyhedron and Convex Polytope

A convex polyhedron is a set of points common to one or more half-spaces. A convex polyhedron that is bounded is called a convex polytope.



(a) (b)

Figure *a* and *b* represent convex polytopes in two and three dimensions,



(c) (d)

and Fig. c and *d* denote convex polyhedra in two and three dimensions.

It can be seen that a convex polygon, shown in Fig. a and c, can be considered as the intersection of one or more half-planes.

Vertex or Extreme Point

This is a point in the convex set that **does not lie** on a line segment joining two other points of the set. For example,

every point on the circumference of a circle and each **corner** point of a polygon can
be called a vertex or extreme point.

7. ***Feasible Solution*** In a linear programming problem, any solution that satisfies the constraints

$$\mathbf{aX} = \mathbf{b}$$
$$\mathbf{X} \geq \mathbf{0}$$

is called a *feasible solution.*

8. Basic Solution A basic solution is one in which *n* - *m* variables are set equal to zero. A basic solution can be obtained by setting *n* - *m* variables to zero and solving the constraint Eqs. (3.2) simultaneously.

9. ***Basis*** The collection of variables not set equal to zero to obtain the basic solution is called the basis.

10. Basic Feasible Solution This is a basic solution that satisfies the non negativity conditions of the problem

$$\mathbf{aX} = \mathbf{b} \qquad \mathbf{X} \geq \mathbf{0}$$

11. Non degenerate Basic Feasible Solution This is a basic feasible solution that has got exactly *m* positive x_i.

12. Optimal Solution A feasible solution that optimizes the objective function is called an optimal solution.

13. Optimal Basic Solution This is a basic feasible solution for which the
objective function is optimal.

Theorems

The basic theorems of linear programming can now be stated and proved.

Theorem 1 The intersection of any number of convex sets is also convex.

Proof: Let the given convex sets be represented as R_i $(i = 1,2,\ldots,K)$ and their intersection as R, so that[‡]

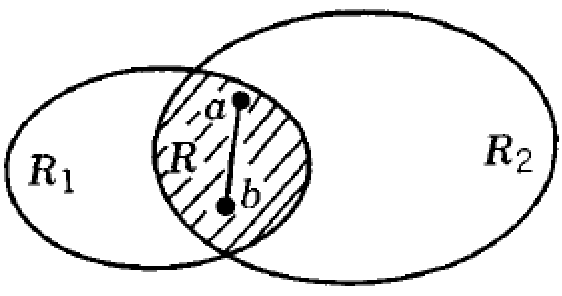
$$R = \bigcap_{i=1}^{K} R_i$$

If the points $\mathbf{X}^{(1)}, \mathbf{X}^{(2)} \in R$, then from the definition of intersection,

$$\mathbf{X} = \lambda \mathbf{X}^{(1)} + (1 - \lambda)\, \mathbf{X}^{(2)} \in R_i \qquad (i = 1,2,\ldots,K)$$
$$0 \le \lambda \le 1$$

Thus

$$\mathbf{X} \in R = \bigcap_{i=1}^{K} R_i$$



and the theorem is proved.

Theorem 2 The ~~feasible region~~ of a linear programming problem is convex.

Proof: The feasible region S of a standard linear programming problem is defined as

$$S = \{\mathbf{X} | \mathbf{aX} = \mathbf{b}, \mathbf{X} \geq 0\}$$

Let the points $\mathbf{X}_1$ and $\mathbf{X}_2$ belong to the feasible set S so that

$$\mathbf{aX}_1 = \mathbf{b}, \quad \mathbf{X}_1 \geq 0$$

$$\mathbf{aX}_2 = \mathbf{b}, \quad \mathbf{X}_2 \geq 0$$

Multiply the 1st eq by λ and the second by $1-\lambda$, and adding them, we obtain:

$$\mathbf{a}[\lambda\mathbf{X}_1 + (1-\lambda)\mathbf{X}_2] = \lambda\mathbf{b} + (1-\lambda)\mathbf{b} = \mathbf{b}$$

that is,

$$\mathbf{aX}_\lambda = \mathbf{b}$$

where

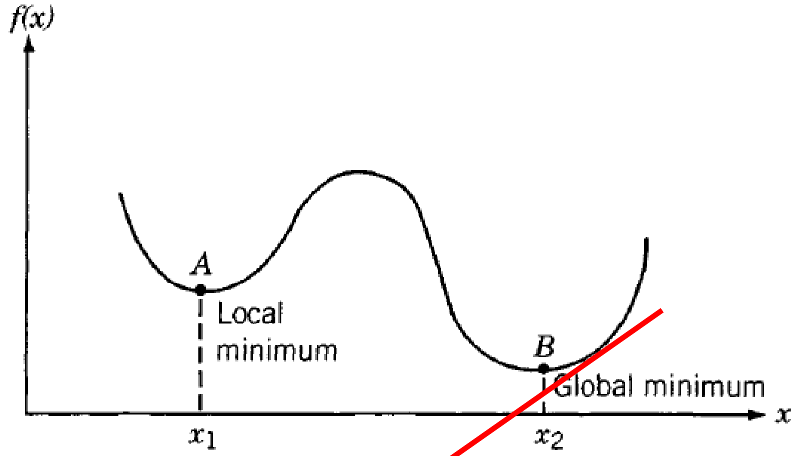
$$\mathbf{X}_\lambda = \lambda\mathbf{X}_1 + (1-\lambda)\mathbf{X}_2$$

Thus the point $\mathbf{X}_\lambda$ satisfies the constraints and if

$$0 \leq \lambda \leq 1, \quad \mathbf{X}_\lambda \geq 0$$

Hence the theorem is proved.

Theorem 3 Any local minimum solution is global for a linear programming problem.



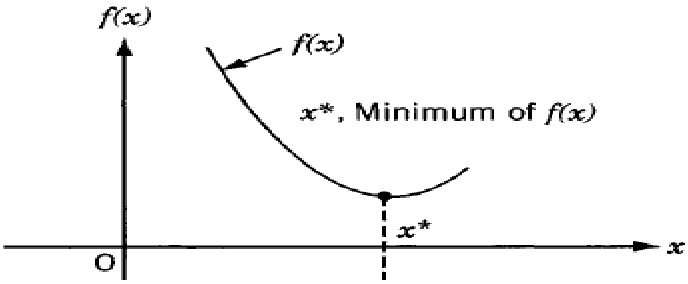
Theorem 4: Every basic feasible solution is an extreme point of the convex set of feasible solutions.

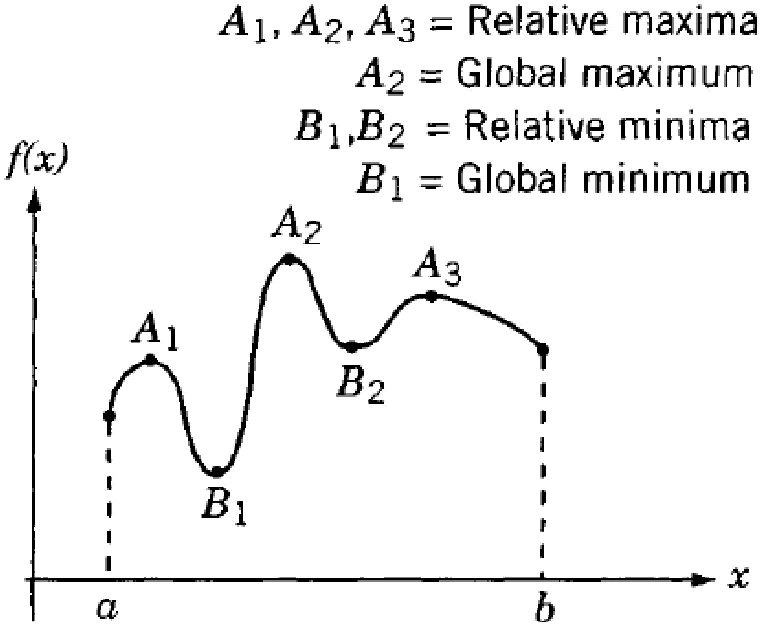
Theorem 5 Let S be a closed convex polyhedron. Then the minimum of a linear function over S is attained at an extreme point of S.

Classical Optimization Techniques

Single-variable optimization

A function of one variable $f(x)$ is said to have a relative or local minimum at $x = x^*$ if $f(x^*) < f(x^* + h)$ for all sufficiently small positive and negative values of h.





Theorem 1: Necessary Condition

If a function $f(x)$ is defined in the interval $a < x < b$ and have a relative minimum at $x = x^*$, where $a < x^* < b$, and if the derivative *df(x)/dx* = $f'(x^*)$ exists as a finite number at $x = x^*$, then $f'(x^*)$ = 0.

Theorem 2: Sufficient Condition:

Let $f'(x^*) = f''(x^*) = \ldots = f^{(n-1)}(x^*) = 0$, but $f^{(n)}(x^*) \neq 0$. Then $f(x^*)$ is

(i) a minimum value of $f(x^*)$ if $f^{(n)}(x^*) > 0$ and *n* is even;

(ii) a maximum value of $f(x^*)$ if $f^{(n)}(x^*) < 0$ and *n* is even;

(iii) neither a maximum nor a minimum if *n* is odd.

(b) Use theorems in(a) to find the optimum values of

$$f(x) = 12x^5 - 45x^4 + 40x^3 + 5$$

Answer:

$$f'(x) = 60x^4 - 3 * 60x^3 + 60 * 2 * x^2$$
$$= 60x^2(x^2 - 3x + 2)$$
$$= 60x^2(x-1)(x-2) = 0$$

The extreme points are

$x = 0, x = 1$ and $x = 2$

$x = 0$	$x = 1$	$x = 2$
$f''(x) = 240x^3 - 540x^2 + 240x$ $f''(0) = 0$ We evaluate the next derivative $f'''(x) = 3 * 240x^2 - 2 * 540x + 240$ $f'''(0) = +240$, order of derivative is odd. So this point is neither maximum nor minimum	$f''(1) = -60$ this point is relative maximum $f_{Max} = 12(1) - 45\ (1) + 40(1) + 5$ =12	$f''(2) = 240$ this point is relative minimum $f_{Min} = -11$

<u>Excercises:</u>

(1)Find the maxima and minima, if any, of the functions

$$f(x) = \frac{x^4}{(x-1)(x-3)^3}$$

$$f(x) = 4x^3 - 18x^2 + 27x - 7$$

$$f(x) = 10x^6 - 48x^5 + 15x^4 + 200x^3 - 120x^2 - 480x + 100$$

[2] Detect which of the following Mathematical statements is true and which is false. Write the false one(s) in the correct case.

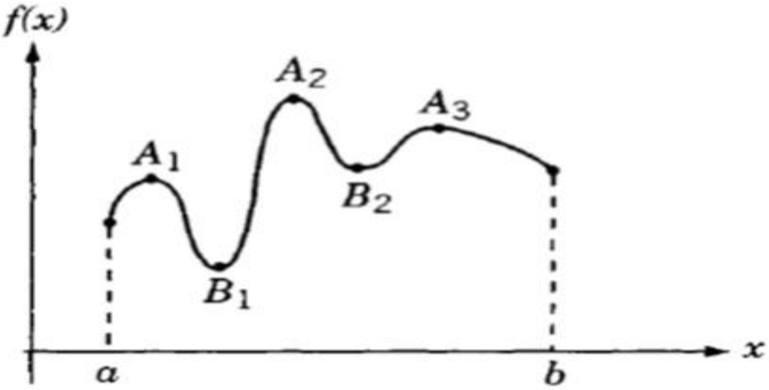


Figure I

In Figure I,

1	A_1 is relative minimum	□	□
2	A_2 is Global Maximum	□	□
	A_3 is relative Maximum	□	□
	B_1 is Global minimum	□	□
	B_2 is Global minimum	□	□
	The necessary condition for a function $f(x)$ to have a relative minimum at $x = x^*$, is $f'(x^*) = 0$.	□	□
	The sufficient condition for a function $f(x)$ to have a relative minimum at $x = x^*$ depends on the order (even- or odd) of the first non zero derivative of $f(x)$.	□	□

[3] Select the correct word

(1) A function of one variable $f(x)$ is said to have a relative or local minimum at $x = x^*$ if $f(x^*) \ldots f(x^* + h)$ for all sufficiently small positive and negative values of h.

□	$<$	□	$\geq$	□	$\leq$	□	Else

Consider using the necessary and sufficient condition to find the optimum values of
$f(x) = 12x^5 - 45x^4 + 40x^3 + 5$. Answer the following questions:

$$f'(x) = ax^2(x-b)(x-c)$$

(1) $a = \cdots$

□	40	□	80	□	60	□	1

(1) $b = \cdots$

□	40	□	1	□	60	□	10

(1) $c = \cdots$

□	40	□	1	□	60	□	2

(1)The extreme point $x = \cdots$ is neither maximum nor minimum

□	0	□	1	□	60	□	2

The extreme point $x = \cdots$ is relative maximum

□	40	□	1	□	60	□	10

The extreme point $x = \cdots$ is relative minimum

□	40	□	1	□	60	□	2

Multivariable optimization with no constraints

Definition: ***r_th*** *Differential of* f: If all partial derivatives of the function f through order $r \geq 1$ exist and are continuous at a point X*, the polynomial

$$d^r f(\mathbf{X}^*) = \underbrace{\sum_{i=1}^{n} \sum_{j=1}^{n} \cdots \sum_{k=1}^{n}}_{r \text{ summations}} h_i h_j \cdots h_k \frac{\partial^r f(\mathbf{X}^*)}{\partial x_i \, \partial x_j \cdots \partial x_k}$$

is called the r^th^ differential of f at X*.

For example :

when $r = 1$ and $n = 3$, we have

$$df(X^*) = \sum_{i=1}^{3} h_i \frac{\partial f}{\partial x_i} = h_1 \frac{\partial f}{\partial x_1} + h_2 \frac{\partial f}{\partial x_2} + h_3 \frac{\partial f}{\partial x_3}$$

Which corresponds $df = \frac{\partial f}{\partial x_1} dx_1 + \frac{\partial f}{\partial x_2} dx_2 + \frac{\partial f}{\partial x_3} dx_3$

When $r = 2$ and $n = 3$, we have

$$d^2 f(\mathbf{X}^*) = d^2 f(x_1^*, x_2^*, x_3^*) = \sum_{i=1}^{3} \sum_{j=1}^{3} h_i h_j \frac{\partial^2 f(\mathbf{X}^*)}{\partial x_i \, \partial x_j}$$

$$= h_1^2 \frac{\partial^2 f}{\partial x_1^2}(\mathbf{X}^*) + h_2^2 \frac{\partial^2 f}{\partial x_2^2}(\mathbf{X}^*) + h_3^2 \frac{\partial^2 f}{\partial x_3^2}(\mathbf{X}^*)$$

$$+ 2h_1 h_2 \frac{\partial^2 f}{\partial x_1 \, \partial x_2}(\mathbf{X}^*) + 2h_2 h_3 \frac{\partial^2 f}{\partial x_2 \, \partial x_3}(\mathbf{X}^*) + 2h_1 h_3 \frac{\partial^2 f}{\partial x_1 \, \partial x_3}(\mathbf{X}^*)$$

The Taylor's series expansion of a function f (X) near a point X* is given by

$$f(\mathbf{X}) = f(\mathbf{X}^*) + df(\mathbf{X}^*) + \frac{1}{2!} d^2f(\mathbf{X}^*) + \frac{1}{3!} d^3f(\mathbf{X}^*)$$

$$+ \cdots + \frac{1}{N!} d^Nf(\mathbf{X}^*) + R_N(\mathbf{X}^*, \mathbf{h})$$

Example 3 : Find the second-order Taylor's series approximation of the function

$$f(x_1, x_2, x_3) = x_2^2 x_3 + x_1 e^{x_3}$$

near the point

$$\mathbf{X}^* = \begin{Bmatrix} 1 \\ 0 \\ -2 \end{Bmatrix}.$$

SOLUTION The second-order Taylor's series approximation of the function f about point $\mathbf{X}^*$ is given by

$$f(\mathbf{X}) = f\begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix} + df\begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix} + \frac{1}{2!} d^2f\begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix}$$

where

$$f\begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix} = e^{-2}$$

$$df\begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix} = h_1 \frac{\partial f}{\partial x_1}\begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix} + h_2 \frac{\partial f}{\partial x_2}\begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix} + h_3 \frac{\partial f}{\partial x_3}\begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix}$$

$$= [h_1 e^{x_3} + h_2(2x_2x_3) + h_3 x_2^2 + h_3 x_1 e^{x_3}]\begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix} = h_1 e^{-2} + h_3 e^{-2}$$

$$d^2f\begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix} = \sum_{i=1}^{3} \sum_{j=1}^{3} h_i h_j \frac{\partial^2 f}{\partial x_i \, \partial x_j}\begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix} = \left(h_1^2 \frac{\partial^2 f}{\partial x_1^2} + h_2^2 \frac{\partial^2 f}{\partial x_2^2} + h_3^2 \frac{\partial^2 f}{\partial x_3^2} \right.$$

$$\left. + 2h_1h_2 \frac{\partial^2 f}{\partial x_1 \, \partial x_2} + 2h_2h_3 \frac{\partial^2 f}{\partial x_2 \, \partial x_3} + 2h_1h_3 \frac{\partial^2 f}{\partial x_1 \, \partial x_3} \right)\begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix}$$

$$= [h_1^2(0) + h_2^2(2x_3) + h_3^2(x_1e^{x_3}) + 2h_1h_2(0) + 2h_2h_3(2x_2)$$

$$+ 2h_1h_3(e^{x_3})]\begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix} = -4h_2^2 + e^{-2}h_3^2 + 2h_1h_3e^{-2}$$

Thus the Taylor's series approximation is given by

$$f(\mathbf{X}) \simeq e^{-2} + e^{-2}(h_1 + h_3) + \frac{1}{2!}(-4h_2^2 + e^{-2}h_3^2 + 2h_1h_3e^{-2})$$

Theorem : Necessary Condition If $f(\mathbf{X})$ has an extreme point (maximum or minimum) at $\mathbf{X} = \mathbf{X}^*$ and if the first partial derivatives of $f(\mathbf{X})$ exist at $\mathbf{X}^*$, then

$$\frac{\partial f}{\partial x_1}(\mathbf{X}^*) = \frac{\partial f}{\partial x_2}(\mathbf{X}^*) = \cdots = \frac{\partial f}{\partial x_n}(\mathbf{X}^*) = 0$$

Proof: The proof given for Theorem can easily be extended to prove the present theorem. However, we present a different approach to prove this theorem. Suppose that one of the first partial derivatives, say the kth one, does not vanish at $\mathbf{X}^*$. Then, by Taylor's theorem,

X=X*+h

$$f(\mathbf{X}^* + \mathbf{h}) = f(\mathbf{X}^*) + \sum_{i=1}^{n} h_i \frac{\partial f}{\partial x_i}(\mathbf{X}^*) + \frac{1}{2!} d^2 f(\mathbf{X}^* + \theta\mathbf{h}),$$

that is,

$$f(\mathbf{X}^* + \mathbf{h}) - f(\mathbf{X}^*) = h_k \frac{\partial f}{\partial x_k}(\mathbf{X}^*) + \frac{1}{2!} d^2 f(\mathbf{X}^* + \theta\mathbf{h}), \qquad 0 < \theta < 1$$

Since $d^2f(\mathbf{X}^* + \theta\mathbf{h})$ is of order h_i^2, the terms of order $\mathbf{h}$ will dominate the higher-order terms for small $\mathbf{h}$. Thus the sign of $f(\mathbf{X}^* + \mathbf{h}) - f(\mathbf{X}^*)$ is decided by the sign of $h_k\, \partial f(\mathbf{X}^*)/\partial x_k$. Suppose that $\partial f(\mathbf{X}^*)/\partial x_k > 0$. Then the sign of $f(\mathbf{X}^* + \mathbf{h}) - f(\mathbf{X}^*)$ will be positive for $h_k > 0$ and negative for $h_k < 0$. This means that $\mathbf{X}^*$ cannot be an extreme point. The same conclusion can be obtained even if we assume that $\partial f(\mathbf{X}^*)/\partial x_k < 0$. Since this conclusion is in contradiction with the original statement that $\mathbf{X}^*$ is an extreme point, we may say that $\partial f/\partial x_k = 0$ at $\mathbf{X} = \mathbf{X}^*$. Hence the theorem is proved.

Theorem : Sufficient Condition A sufficient condition for a stationary point $\mathbf{X}^*$ to be an extreme point is that the matrix of second partial derivatives (Hessian matrix) of $f(\mathbf{X})$ evaluated at $\mathbf{X}^*$ is (i) positive definite when $\mathbf{X}^*$ is a relative minimum point, and (ii) negative definite when $\mathbf{X}^*$ is a relative maximum point.

Proof: From Taylor's theorem we can write

$$f(\mathbf{X}^* + \mathbf{h}) = f(\mathbf{X}^*) + \sum_{i=1}^{n} h_i \frac{\partial f}{\partial x_i}(\mathbf{X}^*) + \frac{1}{2!} \sum_{i=1}^{n} \sum_{j=1}^{n} h_i h_j \left.\frac{\partial^2 f}{\partial x_i\, \partial x_j}\right|_{\mathbf{X}=\mathbf{X}^*+\theta\mathbf{h}},$$

$$0 < \theta < 1$$

Since $\mathbf{X}^*$ is a stationary point, the necessary conditions give (Theorem

$$\frac{\partial f}{\partial x_i} = 0, \qquad i = 1,2,\ldots,n$$

Thus Eq. ( ) reduces to

$$f(\mathbf{X}^* + \mathbf{h}) - f(\mathbf{X}^*) = \frac{1}{2!} \sum_{i=1}^{n} \sum_{j=1}^{n} h_i h_j \left.\frac{\partial^2 f}{\partial x_i\, \partial x_j}\right|_{\mathbf{X}=\mathbf{X}^*+\theta\mathbf{h}}, \qquad 0 < \theta < 1$$

Therefore, the sign of

$$f(\mathbf{X}^* + \mathbf{h}) - f(\mathbf{X}^*)$$

will be same as that of

$$\sum_{i=1}^{n} \sum_{j=1}^{n} h_i h_j \left.\frac{\partial^2 f}{\partial x_i\, \partial x_j}\right|_{\mathbf{X}=\mathbf{X}^*+\theta\mathbf{h}}$$

$$Q = \sum_{i=1}^{n} \sum_{j=1}^{n} h_i h_j \left.\frac{\partial^2 f}{\partial x_i\, \partial x_j}\right|_{\mathbf{X}=\mathbf{X}^*}$$

is positive. This quantity Q is a quadratic form and can be written in matrix form as

$$Q = \mathbf{h}^T \mathbf{J} \mathbf{h}|_{\mathbf{X}=\mathbf{X}^*}$$

where

$$\mathbf{J}|_{\mathbf{X}=\mathbf{X}^*} = \left[\left.\frac{\partial^2 f}{\partial x_i\, \partial x_j}\right|_{\mathbf{X}=\mathbf{X}^*} \right]$$

$$\begin{bmatrix} \frac{\partial^2 f}{\partial x_1 \partial x_1} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \frac{\partial^2 f}{\partial x_1 \partial x_3} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2 \partial x_2} & \frac{\partial^2 f}{\partial x_2 \partial x_3} \\ \frac{\partial^2 f}{\partial x_3 \partial x_1} & \frac{\partial^2 f}{\partial x_3 \partial x_2} & \frac{\partial^2 f}{\partial x_3 \partial x_3} \end{bmatrix}$$

is the matrix of second partial derivatives and is called the *Hessian matrix* of $f(\mathbf{X})$.

Definition:

A matrix A will be positive definite if all its eigenvalues are positive;
that is, all the values of λ that satisfy the determinantal equation

$$|A - \lambda I| = 0$$

should be positive. Similarly, the matrix [A] will be negative definite if its eigenvalues are negative.

Another test that can be used to find the positive definiteness of a matrix A of order n involves evaluation of the determinants

$$A = |a_{11}|,$$

$$A_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix},$$

$$A_3 = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{32} \end{vmatrix}, \ldots,$$

$$A_n = \begin{vmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n} \\ \vdots & & & & \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nn} \end{vmatrix}$$

The matrix **A** will be positive definite if and only if all the values A_1, A_2, A_3, . . . , A_n are positive. The matrix **A** will be negative definite if and only if the sign of A_j is $(-1)^j$ for $j = 1,2,\ldots,n$. If some of the A_j are positive and the remaining A_j are zero, the matrix **A** will be positive semidefinite.

A matrix A will be positive definite if and only if all its determinants are positive;
A matrix A will be negative definite if and only if all its determinant A_k satisfies: $(-1)^k, k = 1,2, ..$
A matrix A will be semi-definite if some of its determinant are positive, and the remaining are zeros

Saddle Point

In the case of a function of two variables, $f(x,y)$, the Hessian matrix may be neither positive nor negative definite at a point (x^*,y^*) at which

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

In such a case, the point (x*,y*) is called a *saddle point.*

The characteristic of a saddle point is that it corresponds to a relative minimum or maximum of f(x,*y)* with respect to one variable , say, *x* (the other variable being fixed at *y* = *y**) and a relative maximum or minimum of *f(x,y)* with respect to the second variable *y* (the other variable being fixed at x*).

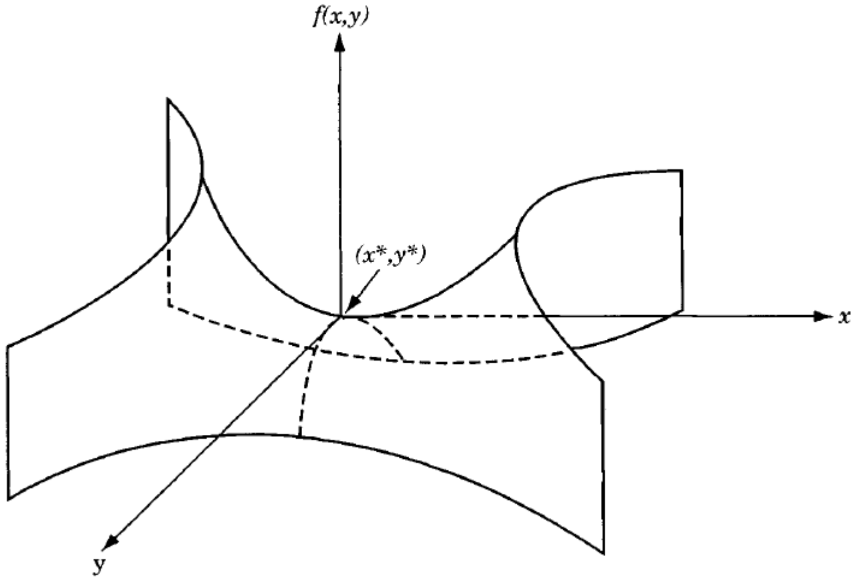


Figure 2.5 Saddle point of the function $f(x,y) = x^2 - y^2$.

1	A matrix A will be positive definite if all its eigenvalues are positive;	□	□
2	A matrix A will be positive definite if and only if all its determinants are positive;	□	□
3	A matrix A will be negative definite if and only if all its determinant A_k satisfies: $(-1)^k$	□	□
4	A matrix A will be semidefinite definite if some of its determinant are positive, and the remaining are zeros	□	□
5	A saddle point is corresponds to a relative minimum of f(x,*y)* with respect to one variable and a relative maximum with respect to the second variable	□	□

[Q3] Find the extreme points of the function

$$f(x_1, x_2) = x_1^3 + x_2^3 + 2x_1^2 + 4x_2^2 + 6$$

The necessary condition is

$\frac{\partial f}{\partial x_1} = 3x_1^2 + 4x_1 = x_1(3x_1 + 4) = 0$ $\frac{\partial f}{\partial x_2} = 3x_2^2 + 8x_2 = x_2(3x_2 + 8) = 0$	So $x_1(3x_1 + 4) = 0$ $x_2(3x_2 + 8) = 0$

Consider finding the extreme points of the function $f(x_1, x_2) = x_1^3 + x_2^3 + 2x_1^2 + 4x_2^2 + 6$. Answer the following

(1) The necessary condition yields

□	$x_1(3x_1 + 4) = 0,$ $x_2(3x_2 + 8) = 0$	□	$x_1(3x_1 - 4) = 0,$ $x_2(3x_2 + 8) = 0$
□	$x_1(3x_1 + 4) = 0,$ $x_2(3x_2 - 8) = 0$	□	$x_1(x_2 - 4) = 0,$ $x_2(x_1 - 8) = 0$

Consider finding the extreme points of the function $f(x_1, x_2) = x_1^3 + x_2^3 + 2x_1^2 + 4x_2^2 + 6$. Answer the following

(2) The solutions of the necessary condition equations are

□	$x_1 = 0, \frac{4}{3}, \ x_2 = 0, \frac{-8}{3}$	□	$x_1 = 0, \frac{-4}{3}, \ x_2 = 0, \frac{-8}{3}$
□	$x_1 = 0, \frac{-4}{3}, \ x_2 = 0, \frac{8}{3}$	□	Else

Consider finding the extreme points of the function $f(x_1, x_2) = x_1^3 + x_2^3 + 2x_1^2 + 4x_2^2 + 6$. Answer the following

(3) The necessary condition equations are satisfied at the points

☐	$(0,0), \left(\frac{-8}{3}, 0\right), \left(0, \frac{-4}{3}\right), \left(\frac{-4}{3}, \frac{-8}{3}\right)$	☐	$(0,0), \left(\frac{8}{3}, 0\right), \left(0, \frac{4}{3}\right), \left(\frac{4}{3}, \frac{8}{3}\right)$
☐	$(0,0), \left(0, \frac{-8}{3}\right), \left(\frac{-4}{3}, 0\right), \left(\frac{-4}{3}, \frac{-8}{3}\right)$	☐	$(0,1), \left(\frac{8}{3}, 1\right), \left(1, \frac{4}{3}\right), \left(\frac{4}{3}, \frac{-8}{3}\right)$

Consider finding the extreme points of the function $f(x_1, x_2) = x_1^3 + x_2^3 + 2x_1^2 + 4x_2^2 + 6$. Answer the following

(3) The following point satisfies the necessary condition

☐	$\left(0, \frac{-4}{3}\right)$	☐	$\left(\frac{4}{3}, \frac{8}{3}\right)$
☐	$\left(\frac{-8}{3}, 0\right)$	☐	$\left(\frac{-4}{3}, \frac{-8}{3}\right)$

To find the nature of these extreme points, we have to use the sufficiency conditions. The second-order partial derivatives of f are given by

$\frac{\partial f}{\partial x_1} = 3x_1^2 + 4x_1$ $\frac{\partial f}{\partial x_2} = 3x_2^2 + 8x_2$	$\frac{\partial^2 f}{\partial {x_1}^2} = 6x_1 + 4, \frac{\partial^2 f}{\partial {x_2}^2} = 6x_2 + 8, \frac{\partial^2 f}{\partial x_1 \partial x_2} = 0$
$\begin{bmatrix} \frac{\partial^2 f}{\partial x_1 \partial x_1} & \frac{\partial^2 f}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2 \partial x_2} \end{bmatrix}$	The Hessian matrix of f is given by $\begin{bmatrix} 6x_1 + 4 & 0 \\ 0 & 6x_2 + 8 \end{bmatrix}$

If $J_1 = |6x_1 + 4|$ and $J_2 = \begin{vmatrix} 6x_1 + 4 & 0 \\ 0 & 6x_2 + 8 \end{vmatrix}$, the values of J_1 and J_2 and the nature of the extreme point are as given below.

$$f(x_1, x_2) = x_1^3 + x_2^3 + 2x_1^2 + 4x_2^2 + 6$$

Point $\mathbf{X}$	Value of J_1	Value of J_2	Nature of $\mathbf{J}$	Nature of $\mathbf{X}$	$f(\mathbf{X})$
$(0,0)$	$+4$	$+32$	Positive definite	Relative minimum	6
$(0,-\frac{8}{3})$	$+4$	-32	Indefinite	Saddle point	418/27
$(-\frac{4}{3},0)$	-4	-32	Indefinite	Saddle point	194/27
$(-\frac{4}{3},-\frac{8}{3})$	-4	$+32$	Negative definite	Relative maximum	50/3

(4)To find the nature of these extreme points, we use the sufficiency conditions. The second-order partial derivatives of f are given by

□	$\frac{\partial^2 f}{\partial x_1^2}=0, \frac{\partial^2 f}{\partial x_2^2}=6x_2+8, \frac{\partial^2 f}{\partial x_1 \partial x_2}=6x_1+4$	□	$\frac{\partial^2 f}{\partial x_1^2}=6x_1+4, \frac{\partial^2 f}{\partial x_2^2}=0, \frac{\partial^2 f}{\partial x_1 \partial x_2}=6x_2+8$
□	$\frac{\partial^2 f}{\partial x_1^2}=6x_1+4, \frac{\partial^2 f}{\partial x_2^2}=6x_2+8, \frac{\partial^2 f}{\partial x_1 \partial x_2}=0$	□	Else

(5) The Hessian matrix of f is given by....

□	$\begin{bmatrix} 0 & 6x_1+4 \\ 0 & 6x_2+8 \end{bmatrix}$	□	$\begin{bmatrix} 6x_1+4 & 0 \\ 0 & 6x_2+8 \end{bmatrix}$
□	$\begin{bmatrix} 6x_1+4 & 0 \\ 6x_2+8 & 0 \end{bmatrix}$	□	$\begin{bmatrix} 6x_1+4 & 3x_1 \\ 2x_2 & 6x_2+8 \end{bmatrix}$

(6) The nature of the extreme point (0,0) is

□	Relative minimum	□	Saddle point	□	Relative maximum	□	indefinite points

(7) The nature of the extreme point $\left(0,\frac{-8}{3}\right)$ is

□	Relative minimum	□	Saddle point	□	Relative maximum	□	indefinite points

(8) The nature of the extreme point $\left(\frac{-4}{3},0\right)$ is

□	Relative minimum	□	Saddle point	□	Relative maximum	□	indefinite points

(9) The nature of the extreme point $\left(\frac{-4}{3},\frac{-8}{3}\right)$ is

□	Relative minimum	□	Saddle point	□	Relative maximum	□	indefinite points

(10) The Relative maximum of the function is

□	194/27	□	6	□	50/3	□	$\mathbf{\frac{418}{27}}$

[1] Answer whether each of the following quadratic forms is positive definite, negative definite, or neither.

(a) $f = x_1^2 - x_2^2$

(b) $f = 4x_1x_2$

(c) $f = x_1^2 + 2x_2^2$

(d) $f = -x_1^2 + 4x_1x_2 + 4x_2^2$

(e) $f = -x_1^2 + 4x_1x_2 - 9x_2^2 + 2x_1x_3 + 8x_2x_3 - 4x_3^2$

(2) Match the following equations and their characteristics.

(a) $f = 4x_1 - 3x_2 + 2$	Relative maximum at (1, 2)
(b) $f = (2x_1 - 2)^2 + (x_1 - 2)^2$	Saddle point at origin
(c) $f = -(x_1 - 1)^2 - (x_2 - 2)^2$	No minimum
(d) $f = x_1x_2$	Inflection point at origin
(e) $f = x^3$	Relative minimum at (1, 2)

(4) Determine whether each of the following matrices is positive definite, negative definite, or indefinite by finding its eigenvalues.

$$[A] = \begin{bmatrix} 3 & 1 & -1 \\ 1 & 3 & -1 \\ -1 & -1 & 5 \end{bmatrix}$$

$$[B] = \begin{bmatrix} 4 & 2 & -4 \\ 2 & 4 & -2 \\ -4 & -2 & 4 \end{bmatrix}$$

$$[C] = \begin{bmatrix} -1 & -1 & -1 \\ -1 & -2 & -2 \\ -1 & -2 & -3 \end{bmatrix}$$

$$[A] = \begin{bmatrix} -14 & 3 & 0 \\ 3 & -1 & 4 \\ 0 & 4 & 2 \end{bmatrix}$$

(5) Determine whether each of the following matrices is positive definite, negative definite, or indefinite by evaluating the signs of its submatrices.

$$[A] = \begin{bmatrix} 3 & 1 & -1 \\ 1 & 3 & -1 \\ -1 & -1 & 5 \end{bmatrix}$$

$$[B] = \begin{bmatrix} 4 & 2 & -4 \\ 2 & 4 & -2 \\ -4 & -2 & 4 \end{bmatrix}$$

$$[C] = \begin{bmatrix} -1 & -1 & -1 \\ -1 & -2 & -2 \\ -1 & -2 & -3 \end{bmatrix}$$

$$[A] = \begin{bmatrix} 4 & -3 & 0 \\ -3 & 0 & 4 \\ 0 & 4 & 2 \end{bmatrix}$$

(6) Express the function

$$f(x_1, x_2, x_3) = -x_1^2 - x_2^2 + 2x_1x_2 - x_3^2 + 6x_1x_3 + 4x_1 - 5x_3 + 2$$

in matrix form as

$$f(X) = \frac{1}{2} X^T [A]X + B^T X + C$$

and determine whether the matrix [A] is positive definite, negative definite, or indefinite.

(7) The profit per acre of a farm is given by

$$20x_1 + 26x_2 + 4x_1x_2 - 4x_1^2 - 3x_2^2$$

where x_1 and x_2 denote, respectively, the labor cost and the fertilizer cost. Find the values of x_1 and x_2 to maximize the profit.

where $X1$ and $X2$ denote, respectively, the labor cost and the fertilizer
cost. Find the values of $X1$ and $X2$ to maximize the profit.

Multivariable Optimization With Equality Constraints

In this section we consider the optimization of continuous functions subjected to equality constraints:

$$\text{Minimize } f = f(\mathbf{X})$$

$$\text{subject to}$$

$$g_j(\mathbf{X}) = 0, \qquad j = 1,2,\ldots,m$$

Where

$$\mathbf{X} = \begin{Bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{Bmatrix}$$

Here *m* is less than or equal to *n;* otherwise (if *m* > *n),* the problem becomes over defined and, in general, there will be no solution. There are several methods available for the solution of this problem. The methods of direct substitution, constrained variation, and Lagrange multipliers are discussed in the following sections.

(*) In the equality constraints optimization problem, the number of constraints must be ….. the number of variable.

□	<	□	≥	□	≤	□	Else

Solution by Direct Substitution

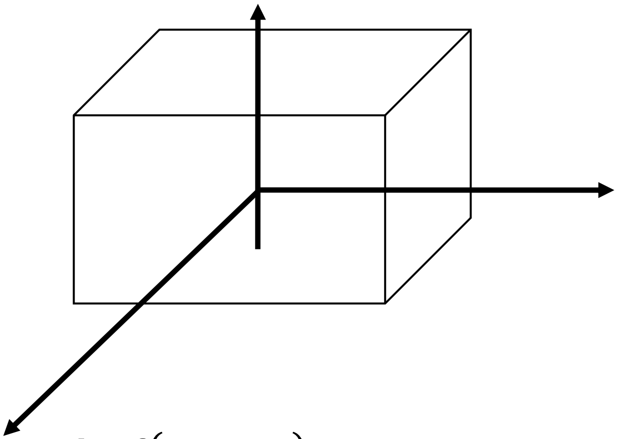
For a problem with *n* variables and *m* equality constraints, it is theoretically possible to solve simultaneously the *m* equality constraints and express any set of *m* variables in terms of the remaining *n* - *m* variables. When these expressions are substituted into the original objective function, there results a new objective function involving only *n* - *m* variables. The new objective function is not subjected to any constraint, and hence its optimum can be found by using the unconstrained optimization techniques discussed in Section 2.3.

[Q4] Find the dimensions of a box of largest volume that can be inscribed in a sphere of unit radius.

SOLUTION Let the origin of the Cartesian coordinate system x_1, x_2, x_3 be at the center of the sphere and the sides of the box be $2x_1$, $2x_2$, and $2x_3$. The

أوجد أبعاد صندوق بحيث يكون له أكبر حجم يمكن احتواؤه في كرة نصف قطرها الوحدة.

$\leftarrow \quad 2x_1 \quad \rightarrow P(x_1, x_2, x_3)$



volume of the box is given by $f(x_{1,}x_2, x_3) = 8x_1x_2x_3$

Since the corners of the box lie on the surface of the sphere of unit radius, x_1, x_2, and x_3 have to satisfy the constraint

$$x_1^2 + x_2^2 + x_3^2 = 1$$

This problem has three design variables and one equality constraint. Hence the equality constraint can be used to eliminate any one of the design variables from the objective function. If we choose to eliminate x_3, Eq. (E_2) gives

$$x_3 = (1 - x_1^2 - x_2^2)^{1/2} \qquad (E_3)$$

Thus the objective function becomes

$$f(x_1, x_2) = 8x_1x_2(1 - x_1^2 - x_2^2)^{1/2} \qquad (E_4)$$

$$f(x_{1,}x_2) = 8x_1x_2\sqrt{1 - x_1^2 - x_2^2}$$

which can be maximized as an unconstrained function in two variables.

The necessary conditions for the maximum of f give

$$\frac{\partial f}{\partial x_1} = 8x_2 \left[(1 - x_1^2 - x_2^2)^{1/2} - \frac{x_1^2}{(1 - x_1^2 - x_2^2)^{1/2}} \right] = 0 \qquad (E_5)$$

$$\frac{\partial f}{\partial x_2} = 8x_1 \left[(1 - x_1^2 - x_2^2)^{1/2} - \frac{x_2^2}{(1 - x_1^2 - x_2^2)^{1/2}} \right] = 0 \qquad (E_6)$$

Equations (E_5) and (E_6) can be simplified to obtain

$$1 - 2x_1^2 - x_2^2 = 0$$

$$1 - x_1^2 - 2x_2^2 = 0$$

from which it follows that $x_1^* = x_2^* = 1/\sqrt{3}$ and hence $x_3^* = 1/\sqrt{3}$. This solution gives the maximum volume of the box as

$$f_{\max} = \frac{8}{3\sqrt{3}}$$

For the sufficient condition, it is clear that the Hessian matrix is negative definite. Hence the point X_1 is maximum for the given function.

1	In the equality constraints optimization problem, the number of constraints must be less than or equal to the number of variable.	☐	☐
2	If the number of constraints is greater than the number of variable in the equality constraints optimization problem, the problem becomes over defined	☐	☐
3	Max. $f(x_1, x_2, x_3) = 8x_1x_2x_3$ subject to $x_1^2 + x_2^2 + x_3^2 = 1$ is equivalent to Max. $f(x_1, x_2) = 8x_1x_2\sqrt{1 - x_1^2 - x_2^2}$	☐	☐
4	The extreme points for the problem Max. $f(x_1, x_2, x_3) = 8x_1x_2x_3$ subject to $x_1^2 + x_2^2 + x_3^2 = 1$ is $\left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)$.	☐	☐
5	The maximum value for the problem Max. $f(x_1, x_2, x_3) = 8x_1x_2x_3$ subject to $x_1^2 + x_2^2 + x_3^2 = 1$ is $\frac{8}{3\sqrt{3}}$.	☐	☐

[Q4]

Minimize $f = 9 - 8x_1 - 6x_2 - 4x_3 + 2x_1^2 + 2x_2^2 + x_3^2 + 2x_1x_2 + 2x_1x_3$ subject to $x_1 + x_2 + 2x_3 = 3$ by direct substitution,

[Q5] Consider the problem Minimize $f(X) = \frac{x_1^2+x_2^2+x_3^2}{2}$

Subject to $g_1(X) = x_1 - x_2 = 0$, $g_2(X) = x_1 + x_2 + x_3 = 1$

By Direct substitution

[Q6] Find the value of x, y, and z that maximize the function $f(x, y, z) = \frac{6xyz}{x+2y+2z}$

When x, y, and z are restricted by the relation xyz = 16.

Solution by the Method of Lagrange Multipliers

The basic features of the Lagrange multiplier method is given initially for a simple problem of two variables with one constraint. The extension of the method to a general problem of *n* variables with *m* constraints is given later.

Problem with Two Variables and One Constraint.
Consider the problem:
Minimize $f(x_1, x_2)$ subject to $g(x_1, x_2) = 0$

1	Lagrange multiplier for the problem [Minimize $f(x_1, x_2)$ subject to $g(x_1, x_2) = 0$], evaluated at the extreme point is $\lambda = -\frac{\frac{\partial f}{\partial x_2}}{\frac{\partial g}{\partial x_2}}$	☐	☐

The necessary conditions generated by constructing a function L, known as the Lagrange function, as

$$L(x_1,x_2,\lambda) = f(x_1,x_2) + \lambda g(x_1,x_2)$$

By treating L as a function of the three variables x_1, x_2, and λ, the necessary conditions for its extremum are given by

$$\frac{\partial L}{\partial x_1}(x_1,x_2,\lambda) = \frac{\partial f}{\partial x_1}(x_1,x_2) + \lambda \frac{\partial g}{\partial x_1}(x_1,x_2) = 0$$

$$\frac{\partial L}{\partial x_2}(x_1,x_2,\lambda) = \frac{\partial f}{\partial x_2}(x_1,x_2) + \lambda \frac{\partial g}{\partial x_2}(x_1,x_2) = 0$$

$$\frac{\partial L}{\partial \lambda}(x_1,x_2,\lambda) = g(x_1,x_2) = 0$$

2	One of the necessary conditions for extremum of Lagrange multiplier solution for the problem [Minimize $f(x_1, x_2)$ subject to $g(x_1, x_2) = 0$], evaluated at the extreme point is $g(x_1, x_2) = 0$	☐	☐

[Q9] Find the solution of Minimize $f = k/xy^2$ Subject *to*
$x^2 + y^2 = a^2$ using the necessary condition of Lagrange multiplier method

SOLUTION

The Lagrange function is

$$L(x,y,\lambda) = f(x,y) + \lambda g(x,y) = kx^{-1}y^{-2} + \lambda(x^2 + y^2 - a^2)$$

The necessary conditions for the minimum of *f(x, y)* give

$$\frac{\partial L}{\partial x} = -kx^{-2}y^{-2} + 2x\lambda = 0 \qquad (E_1)$$

$$\frac{\partial L}{\partial y} = -2kx^{-1}y^{-3} + 2y\lambda = 0 \qquad (E_2)$$

$$\frac{\partial L}{\partial \lambda} = x^2 + y^2 - a^2 = 0 \qquad (E_3)$$

Equations (E1) and (E2) yield

$$2\lambda = \frac{k}{x^3y^2} = \frac{2k}{xy^4}$$

$$\frac{1}{x^2} = \frac{2}{y^2}$$

from which the relation $x^* = (1/\sqrt{2})\ y^*$ can be obtained. This relation, along with Eq. (E_3), gives the optimum solution as

$$x^* = \frac{a}{\sqrt{3}} \quad \text{and} \quad y^* = \sqrt{2}\,\frac{a}{\sqrt{3}}$$

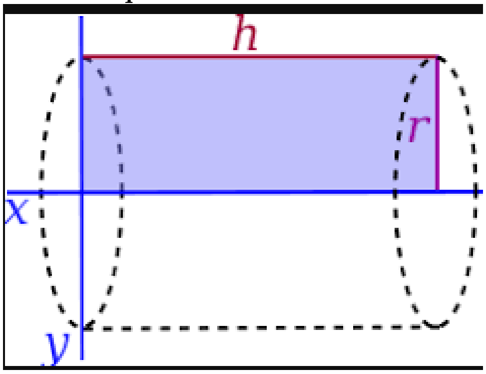
Sufficiency Conditions for a General Problem

Theorem : Sufficient Condition A sufficient condition for $f(\mathbf{X})$ to have a relative minimum at $\mathbf{X}^*$ is that the quadratic, Q, defined by

$$Q = \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 L}{\partial x_i\, \partial x_j}\, dx_i\, dx_j$$

evaluated at $\mathbf{X} = \mathbf{X}^*$ must be positive definite for all values of $d\mathbf{X}$ for which the constraints are satisfied.

[Q11]Find the dimensions of a cylindrical tin (with top and bottom) made up of sheet metal to maximize its volume such that the total surface area is equal to $A0 = 24\,\pi$.



Let the radius of the tin is $r = x_1$ and the length is $h = x_2$. respectively, the problem can be stated as:

$$\text{Maximize } f(x_1, x_2) = \pi x_1^2 x_2$$

subject to

$$2\pi x_1^2 + 2\pi x_1 x_2 = A_0 = 24\pi$$

The Lagrange function is

$$L(X, \lambda) = f(X) + \sum_{j=1}^{m} \lambda_j g_j(X)$$

$$L(x_1, x_2, \lambda) = \pi x_1^2 x_2 + \lambda(2\pi x_1^2 + 2\pi x_1 x_2 - A_0)$$

and the necessary conditions for the maximum of f give

$$\frac{\partial L}{\partial x_1} = 2\pi x_1 x_2 + 4\pi\lambda x_1 + 2\pi\lambda x_2 = 0 \qquad (E_1)$$

$$\frac{\partial L}{\partial x_2} = \pi x_1^2 + 2\pi\lambda x_1 = 0 \qquad (E_2)$$

$$\frac{\partial L}{\partial \lambda} = 2\pi x_1^2 + 2\pi x_1 x_2 - A_0 = 0 \qquad (E_3)$$

Equations (E_1) and (E_2) lead to

$$\lambda = -\frac{x_1 x_2}{2x_1 + x_2} = -\frac{1}{2} x_1$$

that is,

$$x_1 = \tfrac{1}{2} x_2 \tag{E_4}$$

and Eqs. (E_3) and (E_4) give the desired solution as

$$x_1^* = \left(\frac{A_0}{6\pi}\right)^{1/2}, \quad x_2^* = \left(\frac{2A_0}{3\pi}\right)^{1/2}, \quad \text{and} \quad \lambda^* = -\left(\frac{A_0}{24\pi}\right)^{1/2}$$

This gives the maximum value of f as

$$f^* = \left(\frac{A_0^3}{54\pi}\right)^{1/2}$$

If $A_0 = 24\pi$, the optimum solution becomes

$$x_1^* = 2, \quad x_2^* = 4, \quad \lambda^* = -1, \quad \text{and} \quad f^* = 16\pi$$

To see that this solution really corresponds to the maximum of f, we apply the sufficiency condition of Eq. (2.44). In this case

$$L_{11} = \left.\frac{\partial^2 L}{\partial x_1^2}\right|_{(\mathbf{X}^*, \lambda^*)} = 2\pi x_2^* + 4\pi\lambda^* = 4\pi$$

$$L_{12} = \left.\frac{\partial^2 L}{\partial x_1\, \partial x_2}\right|_{(\mathbf{X}^*, \lambda^*)} = L_{21} = 2\pi x_1^* + 2\pi\lambda^* = 2\pi$$

$$L_{22} = \left.\frac{\partial^2 L}{\partial x_2^2}\right|_{(\mathbf{X}^*, \lambda^*)} = 0$$

Now since

$$\frac{\partial L}{\partial \lambda} = 2\pi x_1^2 + 2\pi x_1 x_2 - A_0 = 0$$

$$\frac{\partial^2 L}{\partial x_1 \partial \lambda}\bigg|_{(\mathbf{X}^*,\lambda^*)} = 4\pi x_1^* + 2\pi x_2^* = 16\pi$$

And since

$$\frac{\partial L}{\partial \lambda} = 2\pi x_1^2 + 2\pi x_1 x_2 - A_0 = 0$$

$$\frac{\partial^2 L}{\partial x_2 \partial \lambda}\bigg|_{(\mathbf{X}^*,\lambda^*)} = 2\pi x_1^* = 4\pi$$

And since

$$\frac{\partial L}{\partial \lambda} = 2\pi x_1^2 + 2\pi x_1 x_2 - A_0 = 0$$

$$\frac{\partial^2 L}{\partial \lambda \partial \lambda} = 0$$

$$\begin{bmatrix} \frac{\partial^2 L}{\partial x_1 \partial x_1} & \frac{\partial^2 L}{\partial x_1 \partial x_2} & \frac{\partial^2 L}{\partial x_1 \partial \lambda} \\ \frac{\partial^2 L}{\partial x_2 \partial x_1} & \frac{\partial^2 L}{\partial x_2 \partial x_2} & \frac{\partial^2 L}{\partial x_2 \partial \lambda} \\ \frac{\partial^2 L}{\partial \lambda \partial x_1} & \frac{\partial^2 L}{\partial \lambda \partial x_2} & \frac{\partial^2 L}{\partial \lambda \partial \lambda} \end{bmatrix}$$

For test the positiveness of Hessian matrix:

$$H = \begin{bmatrix} 4\pi & 2\pi & 16\pi \\ 2\pi & 0 & 4\pi \\ 16\pi & 4\pi & 0 \end{bmatrix}$$

$$|H - \lambda I| = 0$$

$$\begin{bmatrix} 4\pi - \lambda & 2\pi & 16\pi \\ 2\pi & 0 - \lambda & 4\pi \\ 16\pi & 4\pi & 0 - \lambda \end{bmatrix} = 0$$

$$272\pi^2\lambda + 192\pi^3 = 0$$

$$\lambda = -\frac{12}{17}\pi$$

Since the value of λ is negative, the point $(x_1{}^*, x_2{}^*)$ corresponds to the maximum of f.

==========================

[Q12] Find the maximum of the function

$f(\mathbf{X}) = 2x_1 + x_2 + 10$ subject to

$g(\mathbf{X}) = x_1 + 2x_2^2 = 3$

using the Lagrange multiplier method.

SOLUTION The Lagrange function is given by

$$L(\mathbf{X},\lambda) = 2x_1 + x_2 + 10 + \lambda(3 - x_1 - 2x_2^2) \qquad (E_1)$$

The necessary conditions for the solution of the problem are

$$\frac{\partial L}{\partial x_1} = 2 - \lambda = 0$$

$$\frac{\partial L}{\partial x_2} = 1 - 4\lambda x_2 = 0 \qquad (E_2)$$

$$\frac{\partial L}{\partial \lambda} = 3 - x_1 - 2x_2^2 = 0$$

The solution of Eqs. (E_2) is

$$\mathbf{X}^* = \begin{Bmatrix} x_1^* \\ x_2^* \end{Bmatrix} = \begin{Bmatrix} 2.97 \\ 0.13 \end{Bmatrix}$$

$$\lambda^* = 2.0 \qquad (E_3)$$

Sufficient condition is Homework

[Q5]

Find the admissible and constrained variations at the point $\mathbf{X} = \begin{Bmatrix} 0 \\ 4 \end{Bmatrix}$ for the following problem:

$$\text{Minimize } f = x_1^2 + (x_2 - 1)^2$$

subject to

$$-2x_1^2 + x_2 = 4$$

$$\text{Minimize } f = 9 - 8x_1 - 6x_2 - 4x_3 + 2x_1^2 + 2x_2^2 + x_3^2 + 2x_1x_2 + 2x_1x_3$$

subject to

$$x_1 + x_2 + 2x_3 = 3$$

by **(a)** direct substitution, **(b)** Lagrange multiplier method.

[Q6] Consider the problem

$$\text{Minimize } f(X) = \frac{x_1^2+x_2^2+x_3^2}{2}$$

Subject to

$$g_1(X) = x_1 - x_2 = 0$$
$$g_2(X) = x_1 + x_2 + x_3 = 1$$

By Lagrange multipliers method.

==

[Q7] (b) Minimize $f(X) = \frac{x_1^2+x_2^2+x_3^2}{2}$ (1)

Subject to $g_1(X) = x_1 - x_2 = 0,$ (2)

$g_2(X) = x_1 + x_2 + x_3 = 1$ (3)

By Lagrange multipliers method.

[Q8] find the value of x, y, and z that maximize the function

$$f(x, y, z) = \frac{6xyz}{x + 2y + 2z}$$

When x, y, and z are restricted by the relation xyz = 16.

Unconstrained Multivariable Optimization Techniques

This chapter deals with the various methods of solving the unconstrained min·imization problem:

$$\text{Find } \mathbf{X} = \begin{Bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{Bmatrix} \text{ which minimizes } f(\mathbf{X})$$

As discussed in Chapter 1, a point X* will be a relative minimum of f (X) if the necessary conditions

$$\frac{\partial f}{\partial x_i}(\mathbf{X} = \mathbf{X}^*) = 0, \qquad i = 1,2,\ldots,n \qquad (\ \ 2)$$

are satisfied. The point $\mathbf{X}^*$ is guaranteed to be a relative minimum if the Hessian matrix is positive definite, that is,

$$\mathbf{J}_{\mathbf{X}^*} = [J]_{\mathbf{X}^*} = \left[\frac{\partial^2 f}{\partial x_i\, \partial x_j}(\mathbf{X}^*)\right] = \text{positive definite} \qquad (\ \ .3)$$

Equations (2) and (3) can be used to identify the optimum point during numerical computations. However, if the function is not differentiate, Eqs. (2) and (3) cannot be applied to identify the optimum point.

Classification of Unconstrained Minimization Methods

Several methods are available for solving an unconstrained minimization problem. These methods can be classified into two broad categories as direct search methods and descent methods as indicated in Table 1.

TABLE 6.1 Unconstrained Minimization Methods

Direct Search Methods	Descent Methods
Random search method	Steepest descent (Cauchy) method
Grid search method	Fletcher–Reeves method
Univariate method	Newton's method
Pattern search methods	Marquardt method
Powell's method	Quasi-Newton methods
Hooke–Jeeves method	Davidon–Fletcher–Powell method
Rosenbrock's method	Broyden–Fletcher–Goldfarb–Shanno method
Simplex method	

Direct Search Methods: Do not require the derivatives of the function.
Descent Methods: Require the derivatives of the function.

	Direct Search Methods don’t require the derivatives of the function.	☐	☐
	Descent Search require the derivatives of the function.	☐	☐

1.2 General Approach

[Q1]Draw the flowchart of general iterative scheme of unconstrained multivariable optimization

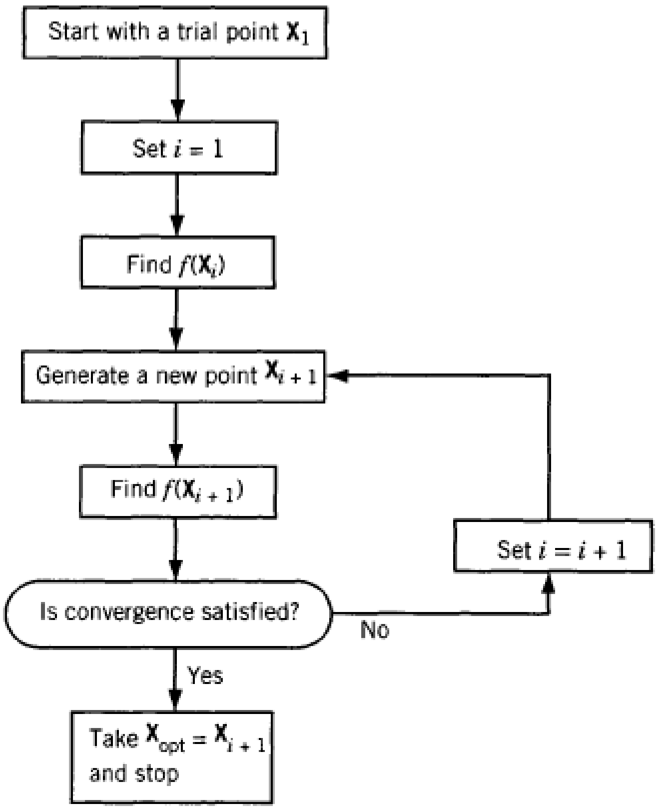


Figure 6.3 General iterative scheme of optimization.

3 Rate of Convergence

Different iterative optimization methods have different rates of convergence.
In general, an optimization method is said to have convergence of order *p* if

$$\frac{\|X_{i+1} - X^*\|}{\|X_i - X^*\|^p} \leq K, K \geq 0, p \geq 1 \qquad (4)$$

where $\mathbf{X}_i$ and $\mathbf{X}_{i+1}$ denote the points obtained at the end of iterations i and $i + 1$, respectively, $\mathbf{X}^*$ represents the optimum point, and $\|\mathbf{X}\|$ denotes the length or norm of the vector $\mathbf{X}$:

$$\|\mathbf{X}\| = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$$

If p = 1 and 0 ≤ k ≤ 1, the method is said to be linearly convergent (corresponds to slow convergence), If p = 2, the method is said to be quadratically convergent (corresponds to fast convergence). An optimization method is said to have superlinear convergence (corresponds to fast convergence) if

$$\lim_{k \to \infty} \frac{\|X_{k+1} - X^*\|}{\|X_k - X^*\|} \to 0 \qquad (6)$$

The definitions of rates of convergence given in Eqs. (.4) and (.6) are applicable to single-variable as well as multivariable optimization problems. In the case of single-variable problems, the vector, $\mathbf{X}_i$, for example, degenerates to a scalar, x_i.

An iterative optimization method satisfies $\lim_{k \to \infty} \frac{\|X_{k+1} - X^*\|}{\|X_k - X^*\|} \to 0$, is *said to be ..... convergence*

□	quadratically	□	superlinear	□	linearly	□	super

	An iterative optimization method *is said to be quadratically convergent if* $\frac{\|X_{i+1} - X^*\|}{\|X_i - X^*\|^p} \leq K, K \geq 0,$ $p = 2$	□	□
	Different iterative optimization methods have the same rates of convergence.	□	□
	Some of the methods for solving constrained minimization problems require the use of unconstrained minimization techniques.	□	□
	The study of unconstrained minimization techniques provides the basic understanding necessary for the study of constrained minimization methods.	□	□
	All the unconstrained minimization methods are iterative in nature.	□	□

Indirect search (descent) methods

Gradient of a function

The gradient of a function is an n-component vector given by

$$\underset{n \times 1}{\nabla f} = \left[\frac{\partial f}{\partial x_1} \frac{\partial f}{\partial x_2} \frac{\partial f}{\partial x_3} \cdots \frac{\partial f}{\partial x_n}\right]^T$$

The gradient has a very important property. If we move along the gradient direction from any point in n-dimensional space, the function value increases at the fastest rate. Hence the gradient direction is called the *direction of steepest ascent.* Unfortunately, the direction of steepest ascent is a local property and not a global one. This is illustrated in Fig. 6.15, where the gradient

Since the gradient vector represents the direction of steepest ascent, the negative of the gradient vector denotes the direction of steepest descent. Thus any method that makes use of the gradient vector can be expected to give the minimum point faster than one that does not make use of the gradient vector. All the descent methods make use of the gradient vector, either directly or indirectly, in finding the search directions. Before considering the descent methods of minimization, we prove that the gradient vector represents the direction of steepest ascent.

[Q1] Prove that the gradient vector represents the direction of steepest ascent.

Theorem 6.3 The gradient vector represents the direction of steepest ascent.

Proof: Consider an arbitrary point **X** in the *n*-dimensional space. Let *f* denote the value of the objective function at the point **X**. Consider a neighboring point $\mathbf{X} + d\mathbf{X}$ with

$$d\mathbf{X} = \begin{Bmatrix} dx_1 \\ dx_2 \\ \vdots \\ dx_n \end{Bmatrix}$$

where $dx_1, dx_2, \ldots, dx_n$ represent the components of the vector $d\mathbf{X}$. The magnitude of the vector $d\mathbf{X}$, ds, is given by

$$d\mathbf{X}^T d\mathbf{X} = (ds)^2 = \sum_{i=1}^{n} (dx_i)^2$$

If $f + df$ denotes the value of the objective function at $\mathbf{X} + d\mathbf{X}$, the change in f, df, associated with $d\mathbf{X}$ can be expressed as

$$df = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} dx_i = \nabla f^T d\mathbf{X} \qquad (1)$$

If $\mathbf{u}$ denotes the unit vector along the direction $d\mathbf{X}$ and ds the length of $d\mathbf{X}$, we can write

$$d\mathbf{X} = \mathbf{u}\, ds$$

The rate of change of the function with respect to the step length ds is given by Eq. (1) as

$$\frac{df}{ds} = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} \frac{dx_i}{ds} = \nabla f^T \frac{d\mathbf{X}}{ds} = \nabla f^T \mathbf{u} \qquad (2)$$

The value of df/ds will be different for different directions and we are interested in finding the particular step $d\mathbf{X}$ along which the value of df/ds will be maximum. This will give the direction of steepest ascent.[†] By using the definition of the dot product, Eq. (2) can be rewritten as

$$\frac{df}{ds} = \|\nabla f\| \, \|\mathbf{u}\| \cos\theta$$

where $\|\nabla f\|$ and $\|\mathbf{u}\|$ denote the lengths of the vectors ∇f and $\mathbf{u}$, respectively, and θ indicates the angle between the vectors ∇f and $\mathbf{u}$. It can be seen that df/ds will be maximum when $\theta = 0°$ and minimum when $\theta = 180°$. This indicates that the function value increases at a maximum rate in the direction of the gradient (i.e., when $\mathbf{u}$ is along ∇f).

[Q2] Prove that the maximum rate of change of f at any point X is equal to the magnitude of the gradient vector at the same point. Then show what we can do if the Evaluation of the Gradient poses certain problem

Theorem 4 The maximum rate of change of f at any point X is equal to the magnitude of the gradient vector at the same point.

Proof: The rate of change of the function f with respect to the step length s along a direction $\mathbf{u}$ is given by Eq. (6.62). Since df/ds is maximum when $\theta = 0°$ and $\mathbf{u}$ is a unit vector, Eq. (6.62) gives

$$\left(\frac{df}{ds}\right)\bigg|_{\max} = \|\nabla f\|$$

which proves the theorem.

<u>Evaluation of the Gradient</u>

[Q3]"The evaluation of the gradient poses certain problems". Discuss this sentence.

The evaluation of the gradient requires the computation of the partial derivatives $\partial f/\partial x_i$, $i = 1,2,\ldots,$ft. There are three situations where the evaluation of the gradient poses certain problems:

1. The function is differentiable at all the points, but the calculation of the components of the gradient, $\partial f/\partial x_i$, is either impractical or impossible.
2. The expressions for the partial derivatives $\partial f/\partial x_i$ can be derived, but they require large computational time for evaluation.
3. The gradient ∇f is not defined at all the points.

In the first case, we can use the forward finite-difference formula

$$\frac{\partial f}{\partial x_i}\bigg|_{\mathbf{X}_m} \simeq \frac{f(\mathbf{X}_m + \Delta x_i \mathbf{u}_i) - f(\mathbf{X}_m)}{\Delta x_i}, \qquad i = 1,2,\ldots,n \qquad (6.63)$$

Steepest descent (Cauchy) method

The use of the negative of the gradient vector as a direction for minimization was first made by Cauchy in 1847. In this method we start from an initial trial point X_1 and iteratively move along the steepest descent directions until the optimum point is found. The steepest descent method can be summarized by the following steps:

[Q4](a)Summarize the steps of steepest descent method for Multivariable Unconstrained Minimization problem.

1. Start with an arbitrary initial point $\mathbf{X}_1$. Set the iteration number as $i = 1$.
2. Find the search direction $\mathbf{S}_i$ as

$$\mathbf{S}_i = -\nabla f_i = -\nabla f(\mathbf{X}_i)$$

3. Determine the optimal step length λ_i^* in the direction $\mathbf{S}_i$ and set

$$\mathbf{X}_{i+1} = \mathbf{X}_i + \lambda_i^* \mathbf{S}_i = \mathbf{X}_i - \lambda_i^* \nabla f_i$$

4. Test the new point, $\mathbf{X}_{i+1}$, for optimality. If $\mathbf{X}_{i+1}$ is optimum, stop the process. Otherwise, go to step 5.
5. Set the new iteration number $i = i + 1$ and go to step 2.

The method of steepest descent may appear to be the *best unconstrained minimization* technique since each one-dimensional search starts in the "best" direction. However, owing to the fact that the steepest descent direction is a local property, the method is not really effective in most problems.

[Q4](b)Use steepest descent method to Minimize the following Multivariable Unconstrained Minimization problem starting from X= {0 0)^T

$$f(x_1, x_2) = x_1 - x_2 + 2x_1^2 + 2x_1 x_2 + x_2^2$$

SOLUTION

Iteration 1

The gradient of f is given by

$$\nabla f = \begin{Bmatrix} \partial f/\partial x_1 \\ \partial f/\partial x_2 \end{Bmatrix} = \begin{Bmatrix} 1 + 4x_1 + 2x_2 \\ -1 + 2x_1 + 2x_2 \end{Bmatrix}$$

$$\nabla f_1 = \nabla f(\mathbf{X}_1) = \begin{Bmatrix} 1 \\ -1 \end{Bmatrix}$$

Therefore,

$$\mathbf{S}_1 = -\nabla f_1 = \begin{Bmatrix} -1 \\ 1 \end{Bmatrix}$$

To find $\mathbf{X}_2$, we need to find the optimal step length λ_1^*. For this, we minimize $f(\mathbf{X}_1 + \lambda_1\mathbf{S}_1) = f(-\lambda_1, \lambda_1) = \lambda_1^2 - 2\lambda_1$ with respect to λ_1. Since $df/d\lambda_1 = 0$ at $\lambda_1^* = 1$, we obtain

$$\mathbf{X}_2 = \mathbf{X}_1 + \lambda_1^*\mathbf{S}_1 = \begin{Bmatrix} 0 \\ 0 \end{Bmatrix} + 1\begin{Bmatrix} -1 \\ 1 \end{Bmatrix} = \begin{Bmatrix} -1 \\ 1 \end{Bmatrix}$$

As $\nabla f_2 = \nabla f(\mathbf{X}_2) = \begin{Bmatrix} -1 \\ -1 \end{Bmatrix} \neq \begin{Bmatrix} 0 \\ 0 \end{Bmatrix}$, $\mathbf{X}_2$ is not optimum.

Iteration 2

$$\mathbf{S}_2 = -\nabla f_2 = \begin{Bmatrix} 1 \\ 1 \end{Bmatrix}$$

To minimize

$$\begin{aligned} f(\mathbf{X}_2 + \lambda_2\mathbf{S}_2) &= f(-1 + \lambda_2, 1 + \lambda_2) \\ &= 5\lambda_2^2 - 2\lambda_2 - 1 \end{aligned}$$

we set $df/d\lambda_2 = 0$. This gives $\lambda_2^* = \frac{1}{5}$, and hence

$$\mathbf{X}_3 = \mathbf{X}_2 + \lambda_2^*\mathbf{S}_2 = \begin{Bmatrix} -1 \\ 1 \end{Bmatrix} + \frac{1}{5}\begin{Bmatrix} 1 \\ 1 \end{Bmatrix} = \begin{Bmatrix} -0.8 \\ 1.2 \end{Bmatrix}$$

Since the components of the gradient at $\mathbf{X}_3$, $\nabla f_3 = \begin{Bmatrix} 0.2 \\ -0.2 \end{Bmatrix}$, are not zero, we proceed to the next iteration.

Iteration 3

$$\mathbf{S}_3 = -\nabla f_3 = \begin{Bmatrix} -0.2 \\ 0.2 \end{Bmatrix}$$

As

$$f(\mathbf{X}_3 + \lambda_3\mathbf{S}_3) = f(-0.8 - 0.2\lambda_3,\ 1.2 + 0.2\lambda_3)$$
$$= 0.04\lambda_3^2 - 0.08\lambda_3 - 1.20, \qquad \frac{df}{d\lambda_3} = 0 \text{ at } \lambda_3^* = 1.0$$

Therefore,

$$\mathbf{X}_4 = \mathbf{X}_3 + \lambda_3^*\mathbf{S}_3 = \begin{Bmatrix} -0.8 \\ 1.2 \end{Bmatrix} + 1.0 \begin{Bmatrix} -0.2 \\ 0.2 \end{Bmatrix} = \begin{Bmatrix} -1.0 \\ 1.4 \end{Bmatrix}$$

The gradient at $\mathbf{X}_4$ is given by

$$\nabla f_4 = \begin{Bmatrix} -0.20 \\ -0.20 \end{Bmatrix}$$

Since $\nabla f_4 \neq \begin{Bmatrix} 0 \\ 0 \end{Bmatrix}$, $\mathbf{X}_4$ is not optimum and hence we have to proceed to the next iteration. This process has to be continued until the optimum point, $\mathbf{X}^* = \begin{Bmatrix} -1.0 \\ 1.5 \end{Bmatrix}$, is found.

Convergence Criteria.

The following criteria can be used to terminate the iterative process:

1. When the change in function value in two consecutive iterations is small:

$$\left|\frac{f(X_{i+1}) - f(X_i)}{f(X_i)}\right| \leq \varepsilon$$

2. When the partial derivatives (components of the gradient) of/are small:

$$\left|\frac{\partial f}{\partial x_i}\right| \leq \varepsilon_2, \qquad i = 1,2,\ldots,n$$

3. When the change in the design vector in two consecutive iterations is small:

$$|\mathbf{X}_{i+1} - \mathbf{X}_i| \leq \varepsilon_3$$

Conjugate Gradient (Fletcher-Reeves) Method

The convergence characteristics of the steepest descent method can be improved greatly by modifying it into a conjugate gradient method (which can be considered as a conjugate directions method involving the use of the gradient of the function).

We saw that any minimization method that makes use of the conjugate directions is quadratically convergent. T

his property of quadratic convergence is very useful because it ensures that the method will minimize a quadratic function in *n* steps or less. Since any general function can be approximated reasonably well by a quadratic near the optimum point, any quadratically convergent method is expected to find the optimum point in a finite number of iterations.

We have seen that Powell's conjugate direction method requires *n* single variable minimizations per iteration and sets up a new conjugate direction at the end of each iteration. Thus it requires, in general, *n2* single-variable minimizations to find the minimum of a quadratic function. On the other hand, if we can evaluate the gradients of the objective function, we can set up a new conjugate direction after every one-dimensional minimization, and hence we can achieve faster convergence. The construction of conjugate directions and development of the Fletcher-Reeves method are discussed in this section.

Development of the Fletcher-Reeves Method

[Q5] Develop the Fletcher-Reeves Method

Consider the development of an algorithm by modifying the steepest descent method applied to a quadratic function $f(\mathbf{X}) = \frac{1}{2}\mathbf{X}^T\mathbf{A}\mathbf{X} + \mathbf{B}^T\mathbf{X} + C$ by imposing the condition that the successive directions be mutually conjugate. Let $\mathbf{X}_1$ be the starting point for the minimization and let the first search direction be the steepest descent direction:

$$\mathbf{S}_1 = -\nabla f_1 = -\mathbf{A}\mathbf{X}_1 - \mathbf{B}$$

$$\mathbf{X}_2 = \mathbf{X}_1 + \lambda_1^*\mathbf{S}_1$$

or

$$\mathbf{S}_1 = \frac{\mathbf{X}_2 - \mathbf{X}_1}{\lambda_1^*}$$

where λ_1^* is the minimizing step length in the direction $\mathbf{S}_1$, so that

$$\mathbf{S}_1^T \nabla f|_{\mathbf{X}_2} = 0 \qquad 1$$

Equation (1) can be expanded as

$$\mathbf{S}_1^T[\mathbf{A}(\mathbf{X}_1 + \lambda_1^*\mathbf{S}_1) + \mathbf{B}] = 0$$

from which the value of λ_1^* can be found as

$$\lambda_1^* = \frac{-\mathbf{S}_1^T(\mathbf{A}\mathbf{X}_1 + \mathbf{B})}{\mathbf{S}_1^T \mathbf{A}\mathbf{S}_1} = -\frac{\mathbf{S}_1^T \nabla f_1}{\mathbf{S}_1^T \mathbf{A}\mathbf{S}_1}$$

Now express the second search direction as a linear combination of $\mathbf{S}_1$ and $-\nabla f_2$:

$$\mathbf{S}_2 = -\nabla f_2 + \beta_2 \mathbf{S}_1$$

where β_2 is to be chosen so as to make $\mathbf{S}_1$ and $\mathbf{S}_2$ conjugate. This requires that

$$\mathbf{S}_1^T \mathbf{A}\mathbf{S}_2 = 0$$

Substituting Eq. ( ) into Eq. ( ) leads to

$$\mathbf{S}_1^T\mathbf{A}\,(-\nabla f_2 + \beta_2\mathbf{S}_1) = 0$$

Equations ( ) and ( ) yield

$$-\frac{(\mathbf{X}_2 - \mathbf{X}_1)^T}{\lambda_1^*}\,\mathbf{A}\,(\nabla f_2 - \beta_2\mathbf{S}_1) = 0$$

The difference of the gradients $(\nabla f_2 - \nabla f_1)$ can be expressed as

$$(\nabla f_2 - \nabla f_1) = (\mathbf{A}\mathbf{X}_2 + \mathbf{B}) - (\mathbf{A}\mathbf{X}_1 + \mathbf{B}) = \mathbf{A}(\mathbf{X}_2 - \mathbf{X}_1)$$

With the help of Eq. ( ), Eq. ( ) can be written as

$$(\nabla f_2 - \nabla f_1)^T(\nabla f_2 - \beta_2\mathbf{S}_1) = 0$$

where the symmetricity of the matrix **A** has been used. Equation (.) can be expanded as

$$\nabla f_2^T \nabla f_2 - \nabla f_1^T \nabla f_2 - \beta_2 \nabla f_2^T \mathbf{S}_1 + \beta_2 \nabla f_1^T \mathbf{S}_1 = 0 \qquad (\quad)$$

Since $\nabla f_1^T \nabla f_2 = -\mathbf{S}_1^T \nabla f_2 = 0$ from Eq. ( ), Eq. ( ) gives

$$\beta_2 = -\frac{\nabla f_2^T \nabla f_2}{\nabla f_1^T \mathbf{S}_1} = \frac{\nabla f_2^T \nabla f_2}{\nabla f_1^T \nabla f_1} \qquad (\quad)$$

Next we consider the third search direction as a linear combination of $\mathbf{S}_1$, $\mathbf{S}_2$, and $-\nabla f_3$ as

$$\mathbf{S}_3 = -\nabla f_3 + \beta_3 \mathbf{S}_2 + \delta_3 \mathbf{S}_1$$

where the values of β_3 and δ_3 can be found by making $\mathbf{S}_3$ conjugate to $\mathbf{S}_1$ and $\mathbf{S}_2$. By using the condition $\mathbf{S}_1^T \mathbf{A} \mathbf{S}_3 = 0$, the value of δ_3 can be found to be zero When the condition $\mathbf{S}_2^T \mathbf{A} \mathbf{S}_3 = 0$ is used, the value of β_3 can be obtained as

$$\beta_3 = \frac{\nabla f_3^T \nabla f_3}{\nabla f_2^T \nabla f_2}$$

so that Eq. ( ) becomes

$$\mathbf{S}_3 = -\nabla f_3 + \beta_3 \mathbf{S}_2$$

where β_3 is given by Eq. ( ). In fact, Eq. ( ) can be generalized as

$$\mathbf{S}_i = -\nabla f_i + \beta_i \mathbf{S}_{i-1} \qquad (\quad)$$

where

$$\beta_i = \frac{\nabla f_i^T \nabla f_i}{\nabla f_{i-1}^T \nabla f_{i-1}} \qquad (\quad)$$

Equations ( ) and ( ) define the search directions used in the Fletcher–Reeves method

	If we move along the gradient direction from any point in n-dimensional space, the function value increases at the fastest rate	□	□
	The gradient vector represents the direction of steepest ascent.	□	□
	The maximum rate of change of f at any point X is equal to the magnitude of the gradient vector at	□	□

	the same point		
	The convergence criteria of steepest descent method that can be used to terminate the iterative process, when the change in function value in two consecutive iterations is small is $\left\|\frac{f(X_{i+1})-f(X_i)}{f(X_i)}\right\| \leq \varepsilon$. .	□	□
	All the unconstrained minimization methods are iterative in nature.	□	□