Chapter Equality and Inequality Constrained Optimization of Multivariable Functions

1- Introduction

In this section we look at problems of the following general form:

$$\min_{x\in\mathbb{R}^n} f(\mathrm{x})$$
$$s.t. \quad g_j(\mathrm{x}) = b_j$$
$$h_j(\mathrm{x}) \le c_j$$

We call the above problem, a Non-Linear programming Optimization Problem (NLP). In it, $f(\mathrm{x})$ is called the objective function, $g_i(x) = b_i$ are **equality constraints**, and $h_i(x) \le c_i$ are in**equality constraints**.

Note that any optimization problem can be written in the **above canonical form**. For example if we want to **maximize** a function $f(x)$, we can do this by **minimizing** $[-f(x)]$.

It turns out that it is easier not to solve (NLP) directly, but instead solve another, related problem (**Lagrange**'s or **Kuhn-Tucker's)** for x^* and then verify that x^* solves the original NLP as well. We will also be interested in whether we are obtaining all solutions to the NLP in this way, i.e., whether it is true that if x^* solves the NLP it solves the related problem as well. Thus we would like to see when the **Lagrange's** or **Kuhn-Tucker's** methods are both necessary and sufficient for obtaining solutions to the original NLP.

2- Nonlinear Programming Optimization with Equality Constraints

In this section we consider the optimization of continuous functions when subjected to equality constraints. This optimization problem is formulated as follows:

$$\min_{X\in R^n} f(\mathbf{x})$$
$$subject\ to:$$
$$g_j(\mathbf{x}) = 0, \quad j = 1, 2, 3, \dots, m$$

Where:

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$: is the design variable vector.

$f(\boldsymbol{X})$: represents the objective function.

$g_i(\boldsymbol{X}) = 0, \quad for\ i = 1,2, \dots, m$: represent the equality constraints, with the equality constraints functions are defined as: $g_i | R^n \to R$.

Here, $m \le n$, otherwise, if $m > n$, the optimization problem becomes then overdefined, and in general, there will be no solution to overdefined problem.

Generally, in engineering, constrained optimization problems are more common than unconstrained problems, but it is important to understand that algorithms for unconstrained optimization problems are the core of algorithms used for solving constrained problems.

3- Solving Equality Constrained Optimization problem

We now try to make some general statements about how to solve equality constrained optimization problem defined and stated earlier. In fact, to solve the multivariable optimization problem with equality constraints, there several methods are available. We shall discuss three fundamental methods:

- Direct substitution method,

- Constrained variation method,
- Lagrange multipliers method.

2.1 Solution by direct Substitution Method

For an optimization 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 the expressions are substituted in the original objective function, it will result a new objective function with only $(n-m)$ variables.

The new objective function will not be subjected to any constraint; hence, it will be treated as an unconstrained optimization problem, where its optimum can be found by using the unconstrained optimization techniques discussed previously.

The method of ***direct substitution***, although it seems (appears) to be simple in theory, is practically not convenient. The reason for this is that the constraint equations will be nonlinear for most of the practical problems, and often it becomes impossible to solve them and express any m variables in terms of the remaining $(n-m)$ variables.

However, the method of ***direct substitution*** might prove to be very simple and direct for solving simple optimization problems as it is shown in the following example.

Example 1:

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 $2x_1, 2\,x_2, 2\,x_3$,

The volume of the box is given by:

$$V = 2x_1 \times 2\,x_2 \times 2\,x_3 = 8x_1\,x_2\,x_3$$

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

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

This problem has three design variables x_1, x_2, x_3 and one equality constraint.

Hence the equality constraint can be used eliminate any one of the design variables from the objective function. If we choose to eliminate the design variable x_3, the constraint equation gives:

$$x_3{}^2 = 1 - x_1{}^2 - x_2{}^2 \Rightarrow x_3 = \sqrt{1 - x_1{}^2 - x_2{}^2}$$

We substitute in the objective function, we find:

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

Which can be maximized as an unconstrained optimization problem of two variables.

Step 1: Necessary Conditions

The necessary condition for finding the maximum point of the function $f(x_1, x_2)$ is to have **zero gradient**, that is:

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

Therefore;

$$\begin{cases} \frac{\partial}{\partial x_1} f(\boldsymbol{X}) = \frac{\partial}{\partial x_1}\left[8x_1\, x_2 \sqrt{1-{x_1}^2 - {x_2}^2}\right] = 8\, x_2 \left[\sqrt{1-{x_1}^2 - {x_2}^2} - \frac{{x_1}^2}{\sqrt{1-{x_1}^2 - {x_2}^2}}\right] = 0 \\ \frac{\partial}{\partial x_2} f(\boldsymbol{X}) = \frac{\partial}{\partial x_2}\left[8x_1\, x_2 \sqrt{1-{x_1}^2 - {x_2}^2}\right] = 8\, x_1 \left[\sqrt{1-{x_1}^2 - {x_2}^2} - \frac{{x_2}^2}{\sqrt{1-{x_1}^2 - {x_2}^2}}\right] = 0 \end{cases}$$

$$\begin{cases} 8\, x_2 \left[\frac{1-{x_1}^2 - {x_2}^2}{\sqrt{1-{x_1}^2 - {x_2}^2}} - \frac{{x_1}^2}{\sqrt{1-{x_1}^2 - {x_2}^2}}\right] = 0 \\ 8\, x_1 \left[\frac{1-{x_1}^2 - {x_2}^2}{\sqrt{1-{x_1}^2 - {x_2}^2}} - \frac{{x_2}^2}{\sqrt{1-{x_1}^2 - {x_2}^2}}\right] = 0 \end{cases} \Rightarrow \begin{cases} 8\, x_2 \left[\frac{1-{x_1}^2 - {x_2}^2 - {x_1}^2}{\sqrt{1-{x_1}^2 - {x_2}^2}}\right] = 0 \\ 8\, x_1 \left[\frac{1-{x_1}^2 - {x_2}^2 - {x_2}^2}{\sqrt{1-{x_1}^2 - {x_2}^2}}\right] = 0 \end{cases}$$

$$\Rightarrow \begin{cases} 8\, x_2 (1-{x_1}^2 - {x_2}^2 - {x_1}^2) = 0 \\ 8\, x_1 (1-{x_1}^2 - {x_2}^2 - {x_2}^2) = 0 \end{cases} \Rightarrow \begin{cases} x_2 = 0 \\ 1 - 2{x_1}^2 - {x_2}^2 = 0 \\ x_1 = 0 \\ 1 - {x_1}^2 - 2{x_2}^2 = 0 \end{cases} \Rightarrow \begin{cases} x_2 = \frac{1}{\sqrt{3}} \\ x_1 = \frac{1}{\sqrt{3}} \end{cases}$$

From which, $x_3 = \frac{1}{\sqrt{3}}$

Hence, the candidate solution to the problem is:

$$\begin{cases} x_1 = \frac{1}{\sqrt{3}} \\ x_2 = \frac{1}{\sqrt{3}} \\ x_3 = \frac{1}{\sqrt{3}} \end{cases}$$

To find whether the obtained solution corresponds to a minimum or maximum, we need to verify the sufficient condition; that is evaluating the Hessian matrix at the candidate solution point.

Step 2: Sufficient Condition

The Hessian matrix is defined as:

$$H(\boldsymbol{X}) = \frac{\partial^2}{\partial x_i \partial x_j} f(\boldsymbol{X}), \quad for \begin{cases} i = 1,2, \ldots, n \\ j = 1,2, \ldots, n \end{cases}$$

It results that:

$$H(\boldsymbol{X}^*) = \begin{bmatrix} \frac{\partial^2}{\partial x_1 \partial x_1} f(\boldsymbol{X}^*) & \frac{\partial^2}{\partial x_1 \partial x_2} f(\boldsymbol{X}^*) \\ \frac{\partial^2}{\partial x_2 \partial x_1} f(\boldsymbol{X}^*) & \frac{\partial^2}{\partial x_2 \partial x_2} f(\boldsymbol{X}^*) \end{bmatrix} = \begin{bmatrix} -\frac{32}{\sqrt{3}} & -\frac{16}{\sqrt{3}} \\ -\frac{16}{\sqrt{3}} & -\frac{32}{\sqrt{3}} \end{bmatrix}$$

To determine the nature of the Hessian matrix at the candidate solution, we use the determinant method as follows:

$$H_1 = |a_{11}| = -\frac{32}{\sqrt{3}} < 0, j = 1$$

$$H_2 = \begin{vmatrix} -\frac{32}{\sqrt{3}} & -\frac{16}{\sqrt{3}} \\ -\frac{16}{\sqrt{3}} & -\frac{32}{\sqrt{3}} \end{vmatrix} = \frac{(32)^2}{3} - \frac{(16)^2}{3} = \frac{(2 \times 16)^2 - (16)^2}{3} = \frac{(16)^2(4-1)}{3} = (16)^2 > 0, j = 2$$

Consequently, the Hessian matrix is ***negative definite*** at (x_1, x_2), which implies that the solution $({x_1}^*, {x_2}^*, {x_3}^*) = (\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}})$ is relative maximum.

2.2 Solving using Lagrange Multipliers Method

2.2.1 General optimization Problem with Equality Constraints

We now try to make some general statements about how to solve general constrained optimization problem for the function of n variables and with m equations representing equality constraints defined and stated earlier using Lagrange multipliers method. In so doing, we shall present the whole procedure of solving the optimization problem using Lagrange Multipliers.

The general equality constrained optimization problem is formulated as follows:

$$Minimize\ f(\boldsymbol{X})$$

$$subject\ to:$$

$$g_j(\boldsymbol{X}) = 0, \quad j = 1, 2, 3, \dots, m$$

Where:

$$\boldsymbol{X} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ \cdot \\ x_n \end{bmatrix}$$

- **Constructing the Lagrange Function L:**

  The Lagrange function (L) in this case is defined by introducing the Lagrange multiplier λ_j for each constraint function $g_j(\boldsymbol{X})$ as follows:

$$L(X,\lambda) = f(\boldsymbol{X}) + \sum_{j=1}^{m} \lambda_j g_j(\boldsymbol{X}) = L(x_1, x_2, \dots, x_n, \lambda_1, \lambda_2, \dots, \lambda_m) = f(\boldsymbol{X}) + \lambda_1 g_1(\boldsymbol{X}) + \cdots + \lambda_m g_m(\boldsymbol{X})$$

The function $L(X,\lambda)$ is called the ***Lagrangian function*** or simply Lagrangian, where, of course:

$$\begin{cases} X \in R^n \\ \lambda \in R^n \end{cases}$$

In using Lagrange multipliers denoted by λ_i, the original optimization problem will be replaced with the mathematically equivalent problem of minimizing:

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

- **Necessary Conditions for a General Optimization Problem**

By treating the Lagrange function (L) as a function of $(n+m)$ unknown variables; $x_1, x_2, \dots, x_n, \lambda_1, \lambda_2, \dots, \lambda_m$, the necessary conditions for the Extremum of L, which also corresponds to the Extremum of the original objective function $f(\boldsymbol{X})$ are given as follows:

$$\begin{cases} \dfrac{\partial}{\partial x_i} L = \dfrac{\partial}{\partial x_i} f(\boldsymbol{X}) + \displaystyle\sum_{j=1}^{m} \lambda_j \dfrac{\partial g_j(\boldsymbol{X})}{\partial x_i} = 0, & for: i = 1,2,3, \dots\dots, n \\ \dfrac{\partial}{\partial \lambda_j} L = g_j(\boldsymbol{X}) = 0, & for: j = 1,2,3, \dots\dots, m \end{cases} \quad (1)$$

The set of equations (1) represents $(n+m)$ equations of $(n+m)$ unknowns, $(x_1, x_2, \dots, x_n, \lambda_1, \lambda_2, \dots, \lambda_m)$ where the solution to these equations simultaneously gives the vectors:

$$\boldsymbol{X}^* = \begin{bmatrix} x_1{}^* \\ x_2{}^* \\ \vdots \\ . \\ x_n{}^* \end{bmatrix}, \quad \lambda^* = \begin{bmatrix} \lambda_1{}^* \\ \lambda_2{}^* \\ \vdots \\ . \\ \lambda_m{}^* \end{bmatrix}$$

The vector $\boldsymbol{X}^*$ corresponds to the solution of the optimization problem; that is the relative Extremum of the constrained function $f(\boldsymbol{X})$.

Regarding the necessary conditions for the existence of the solution, the optimal solution must satisfy:

$$\nabla L(X^*, \lambda^*) = 0$$

And, hence, instead of computing:

$$\boldsymbol{\nabla} = \left[\frac{\boldsymbol{\partial}}{\boldsymbol{\partial x_1}} \; \frac{\boldsymbol{\partial}}{\boldsymbol{\partial x_2}} \; \cdots \; \frac{\boldsymbol{\partial}}{\boldsymbol{\partial x_n}}\right]^{\boldsymbol{T}}$$

We compute:

$$\begin{bmatrix} \nabla_X \\ \\ \nabla_\lambda \end{bmatrix}$$

Such that:

$$\nabla_X = \left[\frac{\boldsymbol{\partial}}{\boldsymbol{\partial x_1}} \; \frac{\boldsymbol{\partial}}{\boldsymbol{\partial x_2}} \; \cdots \; \frac{\boldsymbol{\partial}}{\boldsymbol{\partial x_n}}\right]^{\boldsymbol{T}}$$

$$\nabla_\lambda = \left[\frac{\boldsymbol{\partial}}{\boldsymbol{\partial \lambda_1}} \; \frac{\boldsymbol{\partial}}{\boldsymbol{\partial \lambda_2}} \; \cdots \; \frac{\boldsymbol{\partial}}{\boldsymbol{\partial \lambda_m}}\right]^{\boldsymbol{T}}$$

Consequently, we see that a necessary condition for stationary point of the objective function $f(X)$ subject to our constraints is that $(X^*, \; \lambda^*)$ form a stationary point of the Lagrangian function.

Remark:

The vector λ^* **of the Lagrange multipliers** provides the Sensitivity information, as it will be discussed in the next subsection.

❖ Sufficient Conditions for a General Optimization Problem

Of course, to solve whether the stationary point X^* is a local minimizer, requires additional information (the nature of the Hessian matrix), which defines the sufficient condition.

Using Lagrange multipliers method, the sufficient condition for the equality constrained objective function $f(\boldsymbol{X})$ to have a relative (local) minimum at the point $\boldsymbol{X} = \boldsymbol{X}^*$ is given by the following theorem:

❖ **Theorem**: **Sufficient Condition**:

A sufficient condition for the equality constrained objective function $f(\boldsymbol{X})$ to have a relative (local) minimum at the point $\boldsymbol{X} = \boldsymbol{X}^*$ is that the **Quadratic**, **Q**, defined by:

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

Evaluated at $\boldsymbol{X} = \boldsymbol{X}^*$ must be ***Positive definite*** for all values of $d\boldsymbol{X}$ for which the constraints are satisfied.

❖ **Proof**:

.*

*

*

*

*

*

**

**

Important Notes:

1. If

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

Is negative for all choices of admissible variations dx_i, the stationary vector values $\boldsymbol{X}^*$ will be a constrained maximum of the objective function $f(\boldsymbol{X})$.

2. It has been shown that a necessary condition for the Quadratic matrix $\boldsymbol{Q}$, to be positive (negative) definite for all admissible variations $d\boldsymbol{X}$ is that each root of the polynomial z_i, defined by the following determinant equation, be positive (negative).

$$\begin{vmatrix} L_{11}-z & L_{12} & \dots & L_{1n} & g_{11} & g_{21} & \cdots & g_{m1} \\ L_{21} & L_{22}-z & \cdots & L_{21} & g_{12} & g_{22} & \cdots & g_{m2} \\ \vdots & \vdots & \dots & \vdots & \vdots & \vdots & \dots & \vdots \\ L_{n1} & L_{n2} & \dots & L_{nn}-z & g_{1n} & g_{2n} & \cdots & g_{mn} \\ g_{11} & g_{12} & \cdots & g_{1n} & 0 & 0 & \cdots & 0 \\ g_{21} & g_{22} & \cdots & g_{2n} & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ g_{m1} & g_{m2} & \cdots & g_{mn} & 0 & 0 & \cdots & 0 \end{vmatrix} = 0$$

Where:

$$L_{ij} = \frac{\boldsymbol{\partial^2 L}}{\boldsymbol{\partial} x_i x_j}(\boldsymbol{X}^*, \lambda^*)$$

$$g_{ij} = \frac{\boldsymbol{\partial} g_i}{\boldsymbol{\partial} x_j}(\boldsymbol{X}^*)$$

3. The previous determinant equation, on expansion, leads to an (n-m)th –order polynomial in the variable z. If some of the roots of this polynomial are positive while the others are negative, the Quadratic matrix is indefinite and hence, the stationary point $\boldsymbol{X}^*$ **is not an extreme point.**

The application of the necessary and sufficient conditions in the Lagrange multiplier method is illustrated with the help of the following example.

2.2.2 Interpretation of Lagrange Multipliers

To find and understand the physical meaning of Lagrange Multipliers, consider the following optimization problem involving only single equality constraint.

$$Minimize\ f(\boldsymbol{X})$$

$$subject\ to:$$

$$g(\boldsymbol{X}) = b, \quad or: g(\boldsymbol{X}) - b = 0$$

Where:

b: is a constant.

Using Lagrange multiplier method, the necessary conditions to be satisfied for the solution of the problem are:

$$\begin{cases} \left(\frac{\partial f}{\partial x_i} + \lambda \frac{\partial g}{\partial x_i}\right)\Big|_{(x_1^*, x_2^*)} = 0, \qquad i = 1,2,3, \ldots\ldots, n \\ g(\boldsymbol{X}) = 0 \end{cases}$$

2.2.3 Case study: Two Variable Optimization Problem with One Constraint Equation

The basic feature of **Lagrange multiplier method** for solving ***equality constrained*** optimization problem is first explained for the case of simple objective function of two variables with one constraint equation.

The extension of the method to a general problem of n variables with m constraint equations will be discussed later.

Consider the optimization problem defined and formulated as:

$$Minimize\ f(x_1, x_2)$$

$$subject\ to:$$

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

❖ **Necessary Condition**:

For this problem, the necessary condition for the existence of an extreme point at $\boldsymbol{X} = \boldsymbol{X}^*$ is found to be:

$$\nabla f(x_1, x_2) = 0$$

Using the ***constrained variation*** method, this necessary condition for this typical problem can be expressed as:

$$\left(\frac{\partial f}{\partial x_1} - \frac{\frac{\partial f}{\partial x_2}}{\frac{\partial g}{\partial x_2}} \frac{\partial g}{\partial x_1}\right)\Bigg|_{(x_1^*, x_2^*)} = 0$$

Now we define the quantity:

$$\lambda = -\left(\frac{\frac{\partial f}{\partial x_2}}{\frac{\partial g}{\partial x_2}}\right)\Bigg|_{(x_1^*, x_2^*)}$$

Which we call it **Lagrange Multiplier**,

Consequently, the necessary condition can be rewritten as:

$$\left(\frac{\partial f}{\partial x_1} + \lambda \frac{\partial g}{\partial x_1}\right)\Big|_{(x_1^*, x_2^*)} = 0 \qquad (1)$$

From the definition of **Lagrange multiplier**, we can write:

$$\lambda = -\left(\frac{\frac{\partial f}{\partial x_2}}{\frac{\partial g}{\partial x_2}}\right)\Bigg|_{(x_1^*, x_2^*)} \Rightarrow \lambda + \left(\frac{\frac{\partial f}{\partial x_2}}{\frac{\partial g}{\partial x_2}}\right)\Bigg|_{(x_1^*, x_2^*)} = 0 \Rightarrow \left(\frac{\partial f}{\partial x_2} + \lambda \frac{\partial g}{\partial x_2}\right)\Big|_{(x_1^*, x_2^*)} = 0 \qquad (2)$$

In addition, the constraint equation has to be satisfied at the extreme point, that is;

$$g(x_1, x_2)|_{(x_1^*, x_2^*)} = 0 \qquad (3)$$

Thus, equations (1) to (3) represent the ***necessary conditions*** for the point (x_1^*, x_2^*) to be an extreme point of the objective function of two variables $f(x_1, x_2)$.

Notice that, to define the Lagrange multiplier, it is required that:

$$\left(\frac{\partial g}{\partial x_2}\right)\Big|_{(x_1^*, x_2^*)} \neq 0$$

In fact, the derivation of the necessary conditions by the method of Lagrange multipliers requires that at least one of the partial derivatives of the constraint equation $g(x_1, x_2)$ be non zero at the extreme point.

- **Lagrange Function L**

The necessary conditions given by the equations (1)-(3), are more commonly generated by constructing a function L known as Lagrange Function, which is defined as:

$$L(x_1, x_2, \lambda) = f(x_1, x_2) + \lambda g(x_1, x_2) \qquad (4)$$

By treating the Lagrange Function L as a function of three variables, (x_1, x_2, λ), the necessary conditions for its Extremum are given as:

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

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

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

Example1:

Using **Lagrange multiplier method**, find the solution for the following equality constrained optimization problem formulated as:

$$Minimize\ f(x, y) = kx^{-1}y^{-2}$$

$$subjected\ to:$$

$$g(x, y) = x^2 + y^2 - a^2 = 0$$

Solution:

Step 1: **building the Lagrange Function**

In using Lagrange multiplier method, we start building Lagrange function $L(x, y, \lambda)$ as follows:

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

Step 2: necessary conditions for the existence of extreme points of the objective function

The necessary conditions for the function $f(x, y)$ to have extreme points using Lagrange multiplier are as follows:

$$\frac{\partial}{\partial x} L(x, y, \lambda) = 0 \Rightarrow \frac{\partial}{\partial x}(kx^{-1}y^{-2} + \lambda x^2 + \lambda y^2 - \lambda a^2) = 0 \Rightarrow -ky^{-2}x^{-2} + 2\lambda x = 0$$

$$\frac{\partial}{\partial y} L(x, y, \lambda) = 0 \Rightarrow \frac{\partial}{\partial y}(kx^{-1}y^{-2} + \lambda x^2 + \lambda y^2 - \lambda a^2) = 0 \Rightarrow -2ky^{-3}x^{-1} + 2\lambda y = 0$$

$$\frac{\partial}{\partial \lambda} L(x, y, \lambda) = 0 \Rightarrow \frac{\partial}{\partial \lambda}(kx^{-1}y^{-2} + \lambda x^2 + \lambda y^2 - \lambda a^2) = 0 \Rightarrow x^2 + y^2 - a^2 = 0$$

We end up with the following set of equation in terms of the design variables:

$$\begin{cases} -ky^{-2}x^{-2} + 2\lambda x = 0 \\ -2ky^{-3}x^{-1} + 2\lambda y = 0 \\ x^2 + y^2 - a^2 = 0 \end{cases} \Rightarrow \begin{cases} +2\lambda = \dfrac{ky^{-2}x^{-2}}{x} & (E_1) \\ +2\lambda = \dfrac{2ky^{-3}x^{-1}}{y} & (E_2) \\ x^2 + y^2 - a^2 = 0 & (E_3) \end{cases}$$

From equations ((E_1)) and (E_2), we obtain:

$$\frac{ky^{-2}x^{-2}}{x} = \frac{2ky^{-3}x^{-1}}{y} \Rightarrow y^{-2}x^{-3} = 2y^{-4}x^{-1} \Rightarrow x^{-2} = 2y^{-2} \Rightarrow \frac{1}{x^2} = 2\frac{1}{y^2} \Rightarrow x = \frac{1}{\sqrt{2}}y$$

$$\Rightarrow y = x\sqrt{2}$$

Substituting this in equation (E_3), we get:

$$x^2 + 2x^2 - a^2 = 0 \Rightarrow 3x^2 = a^2 \Rightarrow \begin{cases} x = \dfrac{a}{\sqrt{3}} \\ y = \dfrac{a\sqrt{2}}{\sqrt{3}} \end{cases}$$

Therefore, the candidate optimum point of the objective function is at: $(x, y) = (x^*, y^*) = \left(\frac{a}{\sqrt{3}}, \frac{a\sqrt{2}}{\sqrt{3}}\right)$.

To find the nature of the point for the objective function, we appeal for the Hessian matrix of the function as it is evaluated at this point.

Step 3: sufficient condition

The Hessian matrix of the given objective function is generally defined as:

$$H(\boldsymbol{X}) = \frac{\partial^2}{\partial x_i \partial x_j} f(\boldsymbol{X}), \quad for \begin{cases} i = 1,2, \dots, n \\ j = 1,2, \dots, n \end{cases}$$

For function of two variables, its Hessian becomes:

$$H(\boldsymbol{X}) = \begin{bmatrix} \dfrac{\partial^2}{\partial x \partial x} f(\boldsymbol{X}) & \dfrac{\partial^2}{\partial x \partial y} f(\boldsymbol{X}) \\ \dfrac{\partial^2}{\partial y \partial x} f(\boldsymbol{X}) & \dfrac{\partial^2}{\partial y \partial y} f(\boldsymbol{X}) \end{bmatrix} = \begin{bmatrix} 2kx^{-3}y^{-2} & 2kx^{-2}y^{-3} \\ 2kx^{-2}y^{-3} & 2kx^{-1}y^{-4} \end{bmatrix}$$

Therefore:

$$H(\boldsymbol{X}^*) = H\left(\frac{a}{\sqrt{3}}, \frac{a\sqrt{2}}{\sqrt{3}}\right) = \begin{bmatrix} 2k\left(\frac{a}{\sqrt{3}}\right)^{-3}\left(\frac{a\sqrt{2}}{\sqrt{3}}\right)^{-2} & 2k\left(\frac{a}{\sqrt{3}}\right)^{-2}\left(\frac{a\sqrt{2}}{\sqrt{3}}\right)^{-3} \\ 2k\left(\frac{a}{\sqrt{3}}\right)^{-2}\left(\frac{a\sqrt{2}}{\sqrt{3}}\right)^{-3} & 2k\left(\frac{a}{\sqrt{3}}\right)^{-1}\left(\frac{a\sqrt{2}}{\sqrt{3}}\right)^{-4} \end{bmatrix}$$

Obviously, the value of the Hessian matrix depends on the values of parameters 'a' and 'k', hence, when these are known,

- If $H(\boldsymbol{X}^*)$ is ***positive definite***, the candidate solution is relative minimum of the objective function.
- If $H(\boldsymbol{X}^*)$ is ***negative definite***, the candidate solution is relative maximum of the objective function.
- If $H(\boldsymbol{X}^*)$ is ***indefinite***, the candidate solution will be a saddle point of the objective function.

Example 2:

Consider the objective function defined by:

$$f(\boldsymbol{X}) = 2x_1 + x_2 + 10$$

Subjected to the equality constraint:

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

1. Find the maximum of the function using Lagrange multiplier method.
2. Also, find the effect of changing the right-hand side of the constraint on the optimum value of f.

Solution:

When using Lagrange multipliers method to solve the optimization problem, we follow the steps below:

Step 1: building the Lagrange Function (L):

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

Step 2: necessary conditions for the existence of Extremum

The necessary conditions are:

$$\begin{cases} \frac{\partial}{\partial x_1} L(x_1, x_2, \lambda) = 0 \Rightarrow \frac{\partial}{\partial x_1} \left(2x_1 + x_2 + 10 + \lambda(x_1 + 2x_2{}^2 - 3)\right) = 0 \Rightarrow 2 + \lambda = 0 \\ \frac{\partial}{\partial x_2} L(x_1, x_2, \lambda) = 0 \Rightarrow \frac{\partial}{\partial x_2} \left(2x_1 + x_2 + 10 + \lambda(x_1 + 2x_2{}^2 - 3)\right) = 0 \Rightarrow 1 + 4\lambda x_2 = 0 \\ \frac{\partial}{\partial \lambda} L(x_1, x_2, \lambda) = 0 \Rightarrow \frac{\partial}{\partial \lambda} \left(2x_1 + x_2 + 10 + \lambda(x_1 + 2x_2{}^2 - 3)\right) = 0 \Rightarrow x_1 + 2x_2{}^2 - 3 = 0 \end{cases} \quad (E_2)$$

From which, we get:

$$\begin{cases} \lambda = -2 \\ x_2 = \frac{-1}{4\lambda} = \frac{1}{8} = 0.125 \\ x_1 = -2\left(\frac{1}{8}\right)^2 + 3 = \frac{-1 + 64 * 12}{64 * 4} = 2.99 \end{cases}$$

Therefore, the point $\boldsymbol{X}^* = \begin{bmatrix} x_1{}^* \\ x_2{}^* \end{bmatrix} = \begin{bmatrix} 2.99 \\ 0.125 \end{bmatrix}$ is a candidate to be an extreme point for the function $f(\boldsymbol{X})$.

To determine the nature of this point (min, or max), we appeal to the nature of the Hessian matrix of the given function evaluated at this point.

The Hessian matrix is given as:

$$H(\boldsymbol{X}) = \begin{bmatrix} \frac{\partial^2}{\partial x_1 \partial x_1} f(\boldsymbol{X}) & \frac{\partial^2}{\partial x_1 \partial x_2} f(\boldsymbol{X}) \\ \frac{\partial^2}{\partial x_2 \partial x_1} f(\boldsymbol{X}) & \frac{\partial^2}{\partial x_2 \partial x_2} f(\boldsymbol{X}) \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Result:

We cannot use the Hessian matrix of the objective function as measure of a sufficient condition for a candidate point to be an Extremum (min or max) for a constrained optimization problem.

☑ **The Hessian matrix condition is solely used for unconstrained optimization problem.**

To continue solving this constrained optimization problem, we use the nature of the Quadratic matrix, Q. **Refer to pp.106/830**.

The application of the sufficient conditions mentioned previously yields:

$$\begin{vmatrix} L_{11} - z & L_{12} & g_{11} \\ L_{21} & L_{22} - z & g_{12} \\ g_{11} & g_{12} & 0 \end{vmatrix} = 0 \Rightarrow \begin{vmatrix} -z & 0 & -1 \\ 0 & -4\lambda - z & -4x_2 \\ -1 & -4x_2 & 0 \end{vmatrix} = 0$$

For :

$$\begin{cases} \lambda = -2 \\ x_2 = 0.125 \\ x_1 = 2.99 \end{cases}$$

It results that:

$$\begin{vmatrix} -z & 0 & -1 \\ 0 & 8 - z & -0.52 \\ -1 & -0.52 & 0 \end{vmatrix} = 0 \Rightarrow -z(0.52 * 0.52) - 0 + (-1)(8 - z) = 0$$

$$\Rightarrow -0.2704z + z - 8 = 0 \Rightarrow z = \frac{+8}{1 - 0.2704} = 10.965$$

Which reveals that the ***Quadratic Matrix*** **Q** is ***positive definite***, indicating that the solution point is **relative minimum**.

4- Multivariable Optimization with Inequality Constraints

We consider the optimization problem defined as:

$$Minimize\ f(\boldsymbol{X})$$
$$subject\ to:$$
$$g_j(\boldsymbol{X}) \le 0, \quad j = 1,2,3, \dots \dots, m$$

Rule:

As a rule for solving an optimization problem with an inequality constraints is by converting and transforming it to an equality constraints. This is done by adding a nonnegative slack variables, y_j^2, as:

$$g_j(\boldsymbol{X}) + y_j^2 = 0, \quad j = 1,2,3, \dots \dots, m$$

Where the values of the slack variables are unknown.

The optimization problem becomes after transformation reformulated as:

$$Minimize\ f(\boldsymbol{X})$$
$$subject\ to:$$
$$G_j(\boldsymbol{X}, \boldsymbol{Y}) = g_j(\boldsymbol{X}) + y_j^2 = 0, \quad j = 1,2,3, \dots \dots, m$$

Where:

$\boldsymbol{Y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ . \\ y_m \end{bmatrix} = [y_1\ \ y_2 \dots \dots .\, y_m]^T$ is the vector of the slack variables.

This problem can be conveniently solved using Lagrange multipliers. To do so, we are guided by the following steps.

Step 1: construction of Lagrange Function L:

We begin by constructing the Lagrange function as:

$$L(\boldsymbol{X}, \boldsymbol{Y}, \lambda) = f(\boldsymbol{X}) + \lambda_1 g_1(\boldsymbol{X}, \boldsymbol{Y}) + \dots + \lambda_m g_m(\boldsymbol{X}, \boldsymbol{Y}) = f(\boldsymbol{X}) + \sum_{j=1}^{m} \lambda_j G_j(\boldsymbol{X}, \boldsymbol{Y})$$

Where:

$\lambda = \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ . \\ \lambda_m \end{bmatrix} = [\lambda_1\ \ \lambda_2 \dots \dots .\, \lambda_m]^T$: is the vector of the Lagrange multipliers.

Step 2: Necessary Conditions

The step is to apply the necessary conditions in order to determine the stationary (candidate) points of the objective function.

Using Lagrange function, the necessary conditions are stated as follows:

$$\begin{cases} \dfrac{\partial}{\partial x_i} L(\boldsymbol{X}, \boldsymbol{Y}, \lambda) = \dfrac{\partial}{\partial x_i} f(\boldsymbol{X}) + \displaystyle\sum_{j=1}^{m} \lambda_j \dfrac{\partial g_j(\boldsymbol{X})}{\partial x_i} = 0, \quad i = 1,2,3, \dots \dots, n \\ \dfrac{\partial}{\partial \lambda_j} L(\boldsymbol{X}, \boldsymbol{Y}, \lambda) = G_j(\boldsymbol{X}, \boldsymbol{Y}) = g_j(\boldsymbol{X}) + y_j^2 = 0, \quad j = 1,2,3, \dots \dots, m \\ \dfrac{\partial}{\partial y_j} L(\boldsymbol{X}, \boldsymbol{Y}, \lambda) = 2\lambda_j y_j = 0, \quad j = 1,2,3, \dots \dots, m \end{cases} \quad (1)$$

It can be seen that the set of equations (1) represent (n + 2 m) equations in (n + 2m) unknown variables contained in the vectors $\boldsymbol{X}, \boldsymbol{Y}, \lambda$.

The solution of the equations (1) gives thus the optimum solution vector $\boldsymbol{X}^*$, and the corresponding the optimum Lagrange multipliers vector λ^*, the optimum slack variables vector $\boldsymbol{Y}^*$.

Important remarks

- The second equation of the necessary conditions ensures the constraints $g_j(\boldsymbol{X}) \leq 0, j = 1,2,3, \ldots, m$ are satisfied.
- The third equation of the necessary conditions implies two cases:
    - ✓ Either $\lambda_j = 0$, which means that the jth constraint is inactive (those constraints that are satisfied with an equality sign, $g_j = 0$, at the optimum point are called ***active constraints***, while those that are satisfied with strict inequality sign, $g_j < 0$, are termed ***inactive constraints***) and hence can be ignored.
    - ✓ Or: $y_j = 0$, which means that the jth constraint is active, $g_j = 0$ at the optimum point.

pp.112/830