Global solution of nonlinear programming problem with equality and inequality constraints Via Differential equation Approach

2 authors:

Radhi Ali Zaboon
Mustansiriyiah University

Khalid A. A. Utub
University of Basrah

Global solution of nonlinear programming problem with equality and inequality constraints Via Differential equation Approach

By

Radhi Ali Zboon[1] , Hadi J. Mustafa[2] and Khalid Abdul-Zahra[3]

Abstract

The numerical solution of nonlinear programming problem with equality and inequality constraints based on the theoretical aspects has been developed using a suitable nonlinear autonomous differential system. Nonlinear autonomous differential system is introduced as the base of the theory instead of the usual approaches, and the relation between the critical points and local optima of the original optimization problem has been proved and presented. Asymptotic stability of the critical points (optimum) optima has also been proved and discussed. A numerical algorithm which has capable of finding (local) global optima depending on the nature of trajectory and critical point as well as their behavior has been presented. The numerical algorithm has been illustrated for different example to show its efficiency.

[1] Mathematics and computer Applications Department, College of Science, AL-Nahrain University, Baghdad – Iraq.
[2] Mathematics Department, College of Education, AL-Mustansiria University, Baghdad – Iraq.
[3]Mathematics Department, College of Education, AL-Mustansiria University, Baghdad – Iraq.

Introduction

A method for **s**olving equality constrained nonlinear programming problems by using differential equation approach whose critical point are solution to the optimization problem, had been discussed in [1]. This work was extended to include a muletiobjective nonlinear problem in [2].

A sensitivity analysis depending at the solution of [1] had also been discussed in [3]. Recent developments on parallel implementation of genetic algorithms, simulated annealing, tabu search, variable neighborhood search, and greedy randomized adaptive search procedures are discussed for obtain a global solution of problem [4].

In the present work, a differential equation approach to nonlinear programming problem with equality and inequality constrained has been developed and some numerical illustration and algorithm has also been presented. This method guarantee to find a local optimal solution as well as global one of nonlinear programming problem with equality and inequality constrained via a critical (asymptotic) point of a suitable nonlinear autonomous differential system.

1.1 Problem Formulation:

Consider the nonlinear optimization problem with equality and inequality constraints.

$$\left.\begin{array}{lll} & \text{Min} \quad f(x) & \\ \text{subject to} & & \\ & h_i(x) = 0 & i = 1,\dots,m \\ & h_i(x) \le 0 & i = m+1,\dots,p \\ & x \in R^n & \end{array}\right\} \tag{1.1}$$

$p \le n$,

$f, h \in C^2$

And for $x \in R^n$, let

$$J(x) = \{i : 1 \le i \le m \;\text{ or }\; h_i(x) = 0,\;\; i = m+1,\dots,p\}$$

We will know that the necessary condition for the problem (1.1) is that if the regular point x* (see 1.8.1) is a solution of the problem (1.1) then there exists Lagrange multipliers $\lambda(x^*)$ such that (x*, λ*) is a solution of the following system with equality and inequality constraints [5].

$$\left.\begin{array}{ll}\nabla f(x^*) + \sum_{i=1}^{p} \lambda_i^* \nabla h_i(x^*) = 0 & \\ \lambda_i^* h_i(x^*) = 0 & i = m+1,\ldots,p \\ h_i(x^*) = 0 & i = 1,\ldots,m \\ h_i(x^*) \le 0 & i = m+1,\ldots,p \\ \lambda_i^* \ge 0 & i = m+1,\ldots,p \end{array}\right\} \quad (1.2)$$

Where λ_i^* is the Lagrange multiplier and the symbol ∇ stands for the gradient operator, let w=p-m+n if we introduce the slack variables $y_{m+1},\ldots,y_p$, $F: R^w \to R$, $G: R^w \to R$, where R^W stands for $R^n \times R^{p-m}$ such that.

F(x, y) =f(x)

$$G(x,y) = \begin{bmatrix} g_i(x,y) = h_i(x) & i = 1,\ldots,m \\ g_i(x,y) = h_i(x) + \frac{1}{2} y^2 & i = m+1,\ldots,p \end{bmatrix} \quad (1.3)$$

The symbol (x, y) in this paper is standing for $(x_1,\ldots,x_n,y_{m+1},\ldots,y_p)$.
So, one can consider the following equality constrained optimization problem.

$$\left.\begin{array}{l} \min\limits_{(x,y)\in R^w} \quad F(x,y) \\ \text{subject to} \\ G(x,y) = 0 \\ (x,y) \in R^w \\ F,G \in C^2 \end{array}\right\} \quad (1.4)$$

Remarks 1.1.1:

At any rate most workers in optimization theory quickly reject squared slack variable approach (see (1.3)) with one or more of the following standard criticisms.

1. Squared slack variable increases the dimension of the problem [5].
2. Squared slack variable is less stable than non squared slack variable [5].
3. Squared slack variable leads to (asymptotic) singularities and in particular singular Hessians matrix [5].

Definition 1.1 [6] :

A point x^* satisfying the constraint h(x)=0 is side to be a regular point of the constraint if the gradient vector $\nabla h_1(x^*),\ldots,\nabla h_m(x^*)$ are linearly independent .

Definition 1.2 [7] :

Let $f \in C^1$ the transpose of derivative of f at x_0 is called the gradient of f at x_0 and denoted by $\nabla f(x_0)$.

Proposition 1.1.1:

Suppose $x \in R^n$, then

1. $x \in R^n$ solves problem (1.1) $\Leftrightarrow$ $(x, y) \in R^w$ solves problem (1.4) [5].
2. $x \in R^n$ is a regular point of (1.1) $\Leftrightarrow$ $(x, y) \in R^w$ is a regular point of (2.4) [5].
3. $x \in R^n$ satisfies the necessary condition of optimality of (1.1) $\Leftrightarrow$ $(x, y) \in R^w$ satisfies the necessary condition of optimality of (1.4) [5].

Now, let

$$z = \begin{bmatrix} z_1 \equiv x_1 \\ \cdot \\ \cdot \\ \cdot \\ z_n \equiv x_n \\ \hdashline z_{n+1} \equiv y_{n+1} \\ \cdot \\ \cdot \\ \cdot \\ z_{p-m+n} \equiv y_p \end{bmatrix}^T , \quad \dot{z} = \begin{bmatrix} \dot{z}_1 = \dfrac{dx_1}{dt} \\ \cdot \\ \cdot \\ \cdot \\ \dot{z}_n = \dfrac{dx_n}{dt} \\ \hdashline \dot{z}_{n+1} = \dfrac{dy_{m+1}}{dt} \\ \cdot \\ \cdot \\ \cdot \\ \dot{z}_{p-m+n} = \dfrac{dy_p}{dt} \end{bmatrix}^T \tag{1.5}$$

When T stands for transpose operator.
From above notation we can rewrite (1.4) by the form

$$\left.\begin{array}{l} \quad \underset{Z \in R^w}{\text{Min}} \;\; F(z) \\ \text{subject to} \\ \quad G(z) = 0 \\ \quad z \in R^w \\ \quad F, G \in C^2 \end{array}\right\} \tag{1.6}$$

By using (1.3) and (1.5) one can write,

$$a_{ij}(z) = \frac{\partial g_i(z)}{\partial z_j} \qquad i = 1, \ldots, p \;\; ; \quad j = 1, \ldots, w \tag{1.7}$$

and let A(z) be a matrix whose coefficients are $a_{ij}(z)$ and as follows:

$$A(z)=\left[\begin{array}{ccc:ccc} \frac{\partial g_1(z)}{\partial z_1} & \cdots & \frac{\partial g_1(z)}{\partial z_n} & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial g_m(z)}{\partial z_1} & \cdots & \frac{\partial g_m(z)}{\partial z_n} & 0 & \cdots & 0 \\ \hdashline \frac{\partial g_{m+1}(z)}{\partial z_1} & \cdots & \frac{\partial g_{m+1}(z)}{\partial z_n} & \frac{\partial g_{m+1}(z)}{\partial z_{n+1}} & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial g_p(z)}{\partial z_1} & \cdots & \frac{\partial g_p(z)}{\partial z_n} & 0 & \cdots & \frac{\partial g_p(z)}{\partial z_w} \end{array}\right]_{p\times w} \tag{1.8}$$

One can write the matrix A(z) as the following block matrix.

$$A(z)=\left[\begin{array}{c:c} A_{11}(z)_{m\times n} & 0_{m\times p\text{-}m} \\ \hdashline A_{21}(z)_{p-m\times n} & A_{22}(z)_{p-m\times p-m} \end{array}\right]_{p\times w} \tag{1.9}$$

Where

$$A_{11}(z)=\begin{bmatrix} \frac{\partial g_1(z)}{\partial z_1} & \cdots & \frac{\partial g_1(z)}{\partial z_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial g_m(z)}{\partial z_1} & \cdots & \frac{\partial g_m(z)}{\partial z_n} \end{bmatrix}_{p\text{-}m\times n}$$

$$A_{21}(z)=\begin{bmatrix} \frac{\partial g_{m+1}(z)}{\partial z_1} & \cdots & \frac{\partial g_{m+1}(z)}{\partial z_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial g_p(z)}{\partial z_1} & \cdots & \frac{\partial g_p(z)}{\partial z_n} \end{bmatrix}_{m\times n}$$

$$A_{22}(z)=\begin{bmatrix} \frac{\partial g_{m+1}(z)}{\partial z_{n+1}} & . \; . \; . & 0 \\ . & . & . \\ . & . & . \\ . & . & . \\ 0 & . \; . \; . & \frac{\partial g_p(z)}{\partial z_w} \end{bmatrix}_{p\text{-}m\times p\text{-}m}$$

And $0_{m\times p\text{-}m}$ stands for $m\times p-m$ zero matrix.

Let

$$B(z)=\left[\begin{array}{c:c} B_{11}(z)_{n\times n} & 0_{n\times p\text{-}m} \\ \hdashline 0_{p\text{-}m\times n} & B_{22}(z)_{p-m\times p-m} \end{array}\right]_{w\times w} \tag{1.10}$$

be arbitrary (some conditions will impose on this matrix later on) matrix.

To solve the problem (1.6) we shall introduce the following autonomous differential system, and assume that every z is a function of time t implicitly.Define the following differential system:

$$\text{I.}\quad \left[\begin{array}{c:c} B_{11}(z) & 0 \\ \hdashline 0 & B_{22}(z) \end{array}\right]_{w\times w} \left[\begin{array}{c} \dot{z}_1 \\ . \\ . \\ . \\ \dot{z}_n \\ \hdashline \dot{z}_{n+1} \\ . \\ . \\ . \\ \dot{z}_w \end{array}\right]_{w\times 1} = -\left[\begin{array}{c:c} A_{11}(z) & 0 \\ \hdashline A_{21}(z) & A_{22}(z) \end{array}\right]^T_{p\times w} \left[\begin{array}{c} \lambda_1(z) \\ . \\ . \\ . \\ \lambda_m(z) \\ \hdashline \lambda_{m+1}(z) \\ . \\ . \\ . \\ \lambda_p(z) \end{array}\right]_{p\times 1} - \left[\begin{array}{c} \nabla_{Z_1}F(z) \\ . \\ . \\ . \\ \nabla_{Z_n}F(z) \\ \hdashline \nabla_{Z_{n+1}}F(z) \\ . \\ . \\ . \\ \nabla_{Z_w}F(z) \end{array}\right]_{w\times 1} \tag{1.11}$$

$$\text{II.}\quad \left[\begin{array}{c} \frac{dg_1(z)}{dt} \\ . \\ . \\ . \\ \frac{dg_m(z)}{dt} \\ \hdashline \frac{dg_{m+1}(z)}{dt} \\ . \\ . \\ . \\ \frac{dg_p(z)}{dt} \end{array}\right]_{p\times 1} = -\left[\begin{array}{c} g_1(z) \\ . \\ . \\ . \\ g_m(z) \\ \hdashline g_{m+1}(z) \\ . \\ . \\ . \\ g_p(z) \end{array}\right]_{p\times 1} \tag{1.12}$$

Note that equation (1.12) can be written by using (1.7) and the total derivative definition to get the following:

$$\begin{bmatrix} \frac{\partial g_1(z)}{\partial z_1} & \cdots & \frac{\partial g_1(z)}{\partial z_n} & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ \frac{\partial g_m(z)}{\partial z_1} & \cdots & \frac{\partial g_m(z)}{\partial z_n} & 0 & \cdots & 0 \\ \frac{\partial g_{m+1}(z)}{\partial z_1} & \cdots & \frac{\partial g_{m+1}(z)}{\partial z_n} & \frac{\partial g_{m+1}(z)}{\partial z_{n+1}} & \cdots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ \frac{\partial g_p(z)}{\partial z_1} & \cdots & \frac{\partial g_p(z)}{\partial z_n} & 0 & \cdots & \frac{\partial g_p(z)}{\partial z_w} \end{bmatrix}_{p\times w} \begin{bmatrix} \dot{z}_1 \\ \cdot \\ \cdot \\ \cdot \\ \dot{z}_n \\ \dot{z}_{n+1} \\ \cdot \\ \cdot \\ \cdot \\ \dot{z}_w \end{bmatrix}_{w\times 1} = -\begin{bmatrix} g_1(z) \\ \cdot \\ \cdot \\ \cdot \\ g_m(z) \\ g_{m+1}(z) \\ \cdot \\ \cdot \\ \cdot \\ g_p(z) \end{bmatrix}_{p\times 1} \tag{1.13}$$

Where $\nabla F(z)$ is $w\times 1$ gradient vector see, and $\lambda(z)$ is a $p\times 1$ vector, and the $w\times w$ matrix $B(z)$ is arbitrary symmetric matrix, for arbitrary matrices $B_{11}(z)$, and $B_{22}(z)$ that will be selected later on chapter three.

Let $$\dot{z}_x = \begin{bmatrix} \dot{z}_1 \\ \cdot \\ \cdot \\ \cdot \\ \dot{z}_n \end{bmatrix}_{n\times 1} \text{ and } \dot{z}_y = \begin{bmatrix} \dot{z}_{n+1} \\ \cdot \\ \cdot \\ \cdot \\ \dot{z}_w \end{bmatrix}_{p-m\times 1} \tag{1.14}$$

$$\lambda_x(z) = \begin{bmatrix} \lambda_1(z) \\ \cdot \\ \cdot \\ \cdot \\ \lambda_m(z) \end{bmatrix}_{m\times 1} \text{ and } \lambda_y(z) = \begin{bmatrix} \lambda_{m+1}(z) \\ \cdot \\ \cdot \\ \cdot \\ \lambda_p(z) \end{bmatrix}_{p-m\times 1} \tag{1.15}$$

$$\nabla_x F(z) = \begin{bmatrix} \nabla_{z_1} F(z) \\ \cdot \\ \cdot \\ \cdot \\ \nabla_{z_n} F(z) \end{bmatrix}_{n\times 1} \text{ and } \nabla_y F(z) = \begin{bmatrix} \nabla_{z_{n+1}} F(z) \\ \cdot \\ \cdot \\ \cdot \\ \nabla_{z_w} F(z) \end{bmatrix}_{p-m\times 1} \tag{1.16}$$

$$g_a(z) = \begin{bmatrix} g_1(z) \\ \cdot \\ \cdot \\ \cdot \\ g_m(z) \end{bmatrix}_{m\times 1} \text{ and } g_b(z) = \begin{bmatrix} g_{m+1}(z) \\ \cdot \\ \cdot \\ \cdot \\ g_p(z) \end{bmatrix}_{p-m\times 1} \tag{1.17}$$

Thus from the system (1.11), (1.13), and from notations defined in (1.14)-(1.17) we have

$$B_{11}(z)\dot{z}_x = -A_{11}(z)^T \lambda_x(z) + A_{21}(z)^T \lambda_y(z) - \nabla_x F(z) \tag{1.18}$$

$$B_{22}(z)\dot{z}_y = -A_{22}(z)^T \lambda_y(z) - \nabla_y F(z) \tag{1.19}$$

$$A_{11}(z)\dot{z}_x = -g_a(z) \tag{1.20}$$

$$A_{21}(z)\dot{z}_x + A_{22}(z)\dot{z}_y = -g_b(z) \tag{1.21}$$

To solve the system (1.11) and (1.13) for $\dot{z}$ and $\lambda(z)$ uniquely, we want to show that the matrix

$$D(Z) = \left[\begin{array}{cc:cc} B_{11}(z) & 0 & A_{11}(z)^T & A_{21}(z)^T \\ 0 & B_{22}(z) & 0^T & A_{22}(z)^T \\ \hdashline A_{11}(z) & 0 & 0 & 0 \\ A_{21}(z) & A_{22}(z) & 0 & 0 \end{array}\right]_{2p-m+n\times 2p-m+n} \tag{1.22}$$

is non singular.

Then we need the following lemma.

Sometime in this work for simplicity we omit the variable z even when the matrix depends on it, for example A(z) will be written as A only.

Lemma 1.1.1:

The matrix D(z) in (1.22) is non singular, if the matrices B_{11}, B_{22} and $AB^{-1}A^T$ are nonsingular.

Proof:

At the beginning we shall prove that the matrix B in (1.10) is non singular.

Now let the inverse of B exists, i.e.

$$B^{-1}B = BB^{-1} = I \tag{1.23}$$

Let B^{-1} by the form.

$$B^{-1} = \left[\begin{array}{c:c} (b_{11})_{n\times n} & (b_{12})_{n\times p-m} \\ \hdashline (b_{21})_{p-m\times n} & (b_{22})_{p-m\times p-m} \end{array}\right]_{w\times w}$$

From (1.23) we get.

$$\left[\begin{array}{c:c} (B_{11})_{n\times n} & 0_{n\times p-m} \\ \hdashline 0_{p-m\times n} & (B_{22})_{p-m\times p-m} \end{array}\right]_{w\times w} \left[\begin{array}{c:c} (b_{11})_{n\times n} & (b_{12})_{n\times p-m} \\ \hdashline (b_{21})_{p-m\times n} & (b_{22})_{p-m\times p-m} \end{array}\right]_{w\times w} = \left[\begin{array}{c:c} I_{n\times n} & 0_{n\times p-m} \\ \hdashline 0_{p-m\times n} & I_{p-m\times p-m} \end{array}\right]_{w\times w}$$

$$\left[\begin{array}{c:c} (B_{11}b_{11})_{n\times n} & (B_{11}b_{12})_{n\times p-m} \\ \hdashline (B_{22}b_{21})_{p-m\times n} & (B_{22}b_{22})_{p-m\times p-m} \end{array}\right]_{w\times w} = \left[\begin{array}{c:c} I_{n\times n} & 0_{n\times p-m} \\ \hdashline 0_{p-m\times n} & I_{p-m\times p-m} \end{array}\right]_{w\times w}$$

From above, we have.

$B_{11}b_{11} = I$ (where I is of dimension n×n) (1.24)

$B_{11}b_{12} = 0$ (where 0 is of dimension n×p-m) (1.25)

$B_{22}b_{21} = 0$ (where 0 is of dimension p-m×n) (1.26)

$B_{22}b_{22} = I$ (where I is of dimension p-m×p-m) (1.27)

By using the invertibility of B_{11} on (1.24), we have

$$b_{11} = B_{11}^{-1} \tag{1.28}$$

By using the invertibility of B_{22} on (1.27), we have

$$b_{22} = B_{22}^{-1} \tag{1.29}$$

From (1.28) and (1.29) we have.

$$b_{12} = 0 \quad \text{and} \quad b_{21} = 0$$

So the matrix B^{-1} becomes

$$B^{-1} = \begin{bmatrix} (B_{11}^{-1})_{n\times n} & 0_{n\times p\text{-}m} \\ 0_{p\text{-}m\times n} & (B_{22}^{-1})_{p-m\times p-m} \end{bmatrix}_{w\times w}$$

Now let the inverse of D exists, i.e.

$$D^{-1}D = DD^{-1} = I \tag{1.30}$$

Let D^{-1} by the form

$$D = \begin{bmatrix} (d_{11})_{w\times w} & (d_{12})_{p\times w} \\ (d_{21})_{p\times w} & (d_{22})_{p\times p} \end{bmatrix}_{2p-m+n\times 2p-m+n}$$

From (1.30) we get.

$$\begin{bmatrix} (B)_{w\times w} & (A^T)_{p\times w} \\ (A)_{p\times w} & 0_{p\times p} \end{bmatrix} \begin{bmatrix} (d_{11})_{w\times w} & (d_{12})_{w\times p} \\ (d_{21})_{p\times w} & (d_{22})_{p\times p} \end{bmatrix} = \begin{bmatrix} I_{w\times w} & 0_{w\times p} \\ 0_{p\times w} & I_{p\times p} \end{bmatrix}$$

$$\begin{bmatrix} (Bd_{11} + A^T d_{21})_{w\times w} & (Bd_{12} + A^T d_{22})_{w\times p} \\ (A\,d_{11})_{p\times w} & (Ad_{12})_{p\times p} \end{bmatrix} = \begin{bmatrix} I_{w\times w} & 0_{w\times p} \\ 0_{p\times w} & I_{p\times p} \end{bmatrix}$$

From above we have, the following equation:

$Bd_{11} + A^T d_{21} = I$ (where I is of dimension w×w) (1.31)

$Bd_{12} + A^T d_{22} = 0$ (where I is of dimension w×p) (1.32)

$A\,d_{11} = 0$ (where I is of dimension p×w) (1.33)

$Ad_{12} = I$ (where I is of dimension p×p) (1.34)

From (1.31) one gets:

$$d_{11} = B^{-1} - B^{-1}A^T d_{21} \tag{1.35}$$

From (1.33) and (1.35) we get.

$$AB^{-1} - AB^{-1}A^T d_{21} = 0$$

to have a unique solution, we must assume that at last $AB^{-1}A^T$ is invertible, and

$$d_{21} = \left(AB^{-1}A^T\right)^{-1} AB^{-1} \tag{1.36}$$

From (1.35) and (1.36) we get.

$$d_{11} = B^{-1} - B^{-1}A^T\left(AB^{-1}A^T\right)^{-1} AB^{-1}$$

$$d_{11} = \left(I - B^{-1}A^T\left(AB^{-1}A^T\right)^{-1} A\right)B^{-1} \tag{1.37}$$

From (1.32) we have.

$$Bd_{12} + A^T d_{22} = 0 \tag{1.38}$$

$$d_{12} = -B^{-1}A^T d_{22} \tag{1.39}$$

From (1.34) and (1.39) one can obtain:

$$-AB^{-1}A^T d_{22} = I$$

$$d_{22} = -\left(AB^{-1}A^T\right)^{-1} \tag{1.40}$$

and also, from (1.39) and (1.40) one can get:

$$d_{12} = B^{-1}A^T\left(AB^{-1}A^T\right)^{-1} \tag{1.41}$$

From (1.28), (1.29), (1.36), (1.37), (1.40), (1.41), we have; that

$$D^{-1} = \left[\begin{array}{c:c} \left(I - B^{-1}A^T EA\right)B^{-1} & B^{-1}A^T E \\ \hdashline EAB^{-1} & -E \end{array}\right]$$

Where

$$E = \left(AB^{-1}A^T\right)^{-1} \tag{1.42}$$

Remarks 1.1.2:

Case 1:

If m=n, it means the matrix A is square matrix, and if its invertible then the term $\left(AB^{-1}A^T\right)^{-1}$ becomes

$$\left(AB^{-1}A^T\right)^{-1} = \left(A^T\right)^{-1} BA^{-1}$$

Case 2:
If $m \neq n$, then A is rectangular matrix, and if the term $AB^{-1}A^{T}$ is invertible, then there exists a unique matrix E.

$$E = \left(AB^{-1}A^{T}\right)^{-1}$$

Note that:

$$\begin{bmatrix} (A_{11})_{m\times n} & 0_{m\times p-m} \\ (A_{21})_{p-m\times n} & (A_{22})_{p-m\times p-m} \end{bmatrix} \begin{bmatrix} (B_{11})^{-1}_{n\times n} & 0_{m\times p-m} \\ 0 & (B_{22})^{-1}_{p-m\times p-m} \end{bmatrix} \begin{bmatrix} (A_{11})_{m\times n} & 0_{m\times p-m} \\ (A_{21})_{p-m\times n} & (A_{22})_{p-m\times p-m} \end{bmatrix}^{T}_{p\times w}$$

$$= \begin{bmatrix} \left(A_{11}B^{-1}_{11}A^{T}_{11}\right)_{m\times m} & \left(A_{11}B^{-1}_{11}A^{T}_{21}\right)_{m\times p-m} \\ \left(A_{21}B^{-1}_{11}A^{T}_{11}\right)_{p\text{-}m\times m} & \left(A_{21}B^{-1}_{11}A^{T}_{21} + A_{22}B^{-1}_{22}A^{T}_{22}\right)_{p\text{-}m\times p\text{-}m} \end{bmatrix}_{p\times p}$$

If B^{-1}_{11} is symmetric we assume the above matrix by the form.

$$= \begin{bmatrix} (T_{11})_{m\times m} & (T_{12})_{m\times p\text{-}m} \\ (T^{T}_{12})_{p\text{-}m\times m} & (T_{22})_{p\text{-}m\times p\text{-}m} \end{bmatrix}_{p\times p}$$

Now if T_{11} and T_{22} are symmetric matrices and as discussed earlier in (lemma (1.1.1)) one can find after simple calculation the inverse of the above matrix to be.

$$E = \begin{bmatrix} (e_{11})_{m\times m} & (e_{12})_{m\times p-m} \\ (e_{21})_{p-m\times m} & (e_{22})_{p-m\times p-m} \end{bmatrix}_{p\times p}$$

Where

$$e_{11} = T^{-1}_{11} + T^{-1}_{11}T_{12}(T_{22} - T^{T}_{12}T^{-1}_{11}T_{12})^{-1}T^{T}_{12}T^{-1}_{11}, \qquad e_{12} = -T^{-1}_{11}T_{12}(T_{22} - T^{T}_{12}T^{-1}_{11}T_{12})^{-1}$$

$$e_{21} = -(T_{22} - T^{T}_{12}T^{-1}_{11}T_{12})^{-1}T^{T}_{12}T^{-1}_{11}, \qquad e_{22} = (T_{22} - T^{T}_{12}T^{-1}_{11}T_{12})^{-1}$$

Now we can solve (1.11) and (1.13) for $\dot{z}$ and λ uniquely. From (1.11) we have.

$$\dot{z} = -B^{-1}A^{T}\lambda - B^{-1}\nabla F \tag{1.43}$$

By substituting (1.42) in (1.13), we have

$$A\left[-B^{-1}A^{T}\lambda - B^{-1}\nabla F\right] = -G$$

$$-AB^{-1}A^{T}\lambda - AB^{-1}\nabla F = -G$$

$$\lambda = -EAB^{-1}\nabla F + EG \tag{1.44a}$$

$$\begin{bmatrix} \lambda_x \\ \lambda_y \end{bmatrix}_{p\times 1} = -\begin{bmatrix} e_{11} & e_{12} \\ e_{21} & e_{22} \end{bmatrix}_{p\times p} \begin{bmatrix} A_{11} & 0 \\ A_{21} & A_{22} \end{bmatrix}_{p\times w} \begin{bmatrix} B_{11}^{-1} & \\ & B_{22}^{-1} \end{bmatrix}_{w\times w} \begin{bmatrix} \nabla_x F \\ \nabla_y F \end{bmatrix}_{w\times 1} + \begin{bmatrix} e_{11} & e_{12} \\ e_{21} & e_{22} \end{bmatrix}_{p\times p} \begin{bmatrix} g_a \\ g_b \end{bmatrix}_{p\times 1} \tag{1.44b}$$

From (1.42) and (1.43a) we get.

$$\dot{z} = -B^{-1}A^{T}[-EAB^{-1}\nabla F + EG] - B^{-1}\nabla F$$

$$\dot{z} = B^{-1}A^{T}EAB^{-1}\nabla F - B^{-1}A^{T}EG - B^{-1}\nabla F$$
$$\dot{z} = -\left(I - B^{-1}A^{T}EA\right)B^{-1}\nabla F - B^{-1}A^{T}EG$$
$$\dot{z} \equiv \varphi = -PB^{-1}\nabla F - \tilde{P}G \tag{1.45a}$$

and by the matrix form.

$$\begin{bmatrix} \dot{z}_x \\ \dot{z}_y \end{bmatrix} \equiv \begin{bmatrix} \varphi_1 \\ \varphi_2 \end{bmatrix} = -\begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix}\begin{bmatrix} B_{11}^{-1} & 0 \\ 0 & B_{22}^{-1} \end{bmatrix}\begin{bmatrix} \nabla_x F \\ \nabla_y F \end{bmatrix} - \begin{bmatrix} \tilde{p}_{11} & \tilde{p}_{12} \\ \tilde{p}_{21} & \tilde{p}_{22} \end{bmatrix}\begin{bmatrix} g_a \\ g_b \end{bmatrix} \tag{1.45b}$$

Where

$$\begin{bmatrix} (p_{11})_{n\times n} & (p_{12})_{n\times p\text{-}m} \\ (p_{21})_{p\text{-}m\times n} & (p_{22})_{p\text{-}m\times p\text{-}m} \end{bmatrix} = \begin{bmatrix} I_{n\times n} & 0_{n\times p\text{-}m} \\ 0_{p\text{-}m\times n} & I_{p\text{-}m\times p\text{-}m} \end{bmatrix} - \begin{bmatrix} (q_{11})_{n\times n} & (q_{12})_{n\times p\text{-}m} \\ (q_{21})_{p\text{-}m\times n} & (q_{22})_{p\text{-}m\times p\text{-}m} \end{bmatrix} \tag{1.46}$$

where

$$\begin{aligned} \begin{bmatrix} (q_{11})_{n\times n} & (q_{12})_{n\times p\text{-}m} \\ (q_{21})_{p\text{-}m\times n} & (q_{22})_{p\text{-}m\times p\text{-}m} \end{bmatrix} &= \begin{bmatrix} (B_{11}^{-1})_{n\times n} & 0_{n\times p\text{-}m} \\ 0_{p-m\times n} & (B_{22}^{-1})_{p-m\times p-m} \end{bmatrix}\begin{bmatrix} (A_{11})_{m\times n} & 0_{m\times p\text{-}m} \\ (A_{21})_{p-m\times n} & (A_{22})_{p-m\times p-m} \end{bmatrix}^{T} \\ &\quad \begin{bmatrix} (e_{11})_{m\times m} & (e_{12})_{m\times p-m} \\ (e_{21})_{p-m\times m} & (e_{22})_{p-m\times p-m} \end{bmatrix}\begin{bmatrix} (A_{11})_{m\times n} & 0_{m\times p\text{-}m} \\ (A_{21})_{p-m\times n} & (A_{22})_{p-m\times p-m} \end{bmatrix} \end{aligned} \tag{1.47}$$

$$\begin{bmatrix} (\tilde{p}_{11})_{n\times m} & (\tilde{p}_{12})_{m\times p\text{-}m} \\ (\tilde{p}_{21})_{p\text{-}m\times n} & (\tilde{p}_{22})_{p\text{-}m\times p\text{-}m} \end{bmatrix} = \begin{bmatrix} (B_{11}^{-1})_{n\times n} & 0_{n\times p\text{-}m} \\ 0_{p\text{-}m\times n} & (B_{22}^{-1})_{p-m\times p-m} \end{bmatrix}\begin{bmatrix} (A_{11})_{m\times n} & 0_{m\times p\text{-}m} \\ (A_{21})_{p-m\times n} & (A_{22})_{p-m\times p-m} \end{bmatrix}^{T}\begin{bmatrix} (e_{11})_{m\times m} & (e_{12})_{m\times p-m} \\ (e_{21})_{p-m\times m} & (e_{22})_{p-m\times p-m} \end{bmatrix} \tag{1.48}$$

<u>Remarks 1.1.3:</u>

One can calculate φ_1 and φ_2 as:

From (1.18) and (1.19)

$$\dot{z}_x \equiv \varphi_1 = -B_{11}^{-1}A_{11}^{T}\lambda_x - B_{11}^{-1}A_{22}^{T}\lambda_y - B_{11}^{-1}\nabla_x F \tag{1.49}$$
$$\dot{z}_y \equiv \varphi_2 = -B_{22}^{-1}A_{22}^{T}\lambda_y - B_{22}^{-1}\nabla_y F \tag{1.50}$$

From (1.43b) we have.

$$\lambda_x = -(e_{11}A_{11} + e_{12}A_{21})B_{11}^{-1}\nabla_x F - e_{12}A_{22}B_{22}^{-1}\nabla_y F + e_{11}g_a + e_{12}g_b \tag{1.51}$$
$$\lambda_y = -(e_{21}A_{11} + e_{22}A_{21})B_{11}^{-1}\nabla_x F - e_{22}A_{22}B_{22}^{-1}\nabla_y F + e_{21}g_a + e_{22}g_b \tag{1.52}$$

From (1.48) and (1.50), we have.

$$\begin{aligned} \dot{z}_x \equiv \varphi_1 = &\left(B_{11}^{-1}A_{11}^{T}(e_{11}A_{11} + e_{12}A_{21}) + B_{11}^{-1}A_{21}^{T}(e_{21}A_{11} + e_{22}A_{21}) - I\right) \\ &B_{11}^{-1}\nabla_x F - (B_{11}^{-1}A_{11}^{T}e_{11} + B_{11}^{-1}A_{22}^{T}e_{21})g_a + (B_{11}^{-1}A_{11}^{-1}e_{12}A_{22} \\ &+ B_{11}^{-1}A_{22}^{T}e_{22}A_{22})B_{22}^{-1}\nabla_y F - (B_{11}^{-1}A_{11}^{T}e_{12} + B_{11}^{-1}A_{22}^{T}e_{22})g_b \end{aligned} \tag{1.53}$$

$$\dot{z}_y \equiv \varphi_2 = B_{22}^{-1}A_{22}^{T}(e_{21}A_{11} - e_{22}A_{21})B_{11}^{-1}\nabla_x F + (B_{22}^{-1}A_{22}^{T}e_{22}A_{22} - I)B_{22}^{-1}\nabla_y F - B_{22}^{-1}A_{22}^{T}(e_{21}g_a + e_{22}g_b) \tag{1.54}$$

Because that φ of (1.44a) does not contain t explicitly then the system (1.44a) is autonomous, and we can prove the following useful identities.

1. $$\left(\begin{bmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{bmatrix}_{w\times w}\right)^2 = \begin{bmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{bmatrix}_{w\times w} \tag{1.55}$$

Where q_{11} is n×n matrix, q_{12} is n×p-m matrix, q_{21} is p-m×n matrix, q_{22} is p-m×p-m matrix,

2. $$\left(\begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix}_{w\times w}\right)^2 = \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix}_{w\times w} \tag{1.56}$$

Where p_{11} is n×n matrix, p_{12} is n×p-m matrix, p_{21} is p-m×n matrix, p_{22} is p-m×p-m matrix,

3. $$\begin{bmatrix} A_{11} & 0 \\ A_{21} & A_{22} \end{bmatrix}_{p\times w} \begin{bmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{bmatrix}_{w\times w} = \begin{bmatrix} A_{11} & 0 \\ A_{21} & A_{22} \end{bmatrix}_{p\times w} \tag{1.57}$$

4. $$\begin{bmatrix} A_{11} & 0 \\ A_{21} & A_{22} \end{bmatrix}_{p\times w} \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix}_{w\times w} = 0 \tag{1.58}$$

5. $$\begin{bmatrix} A_{11} & 0 \\ A_{21} & A_{22} \end{bmatrix}_{p\times w} \begin{bmatrix} \tilde{p}_{11} & \tilde{p}_{12} \\ \tilde{p}_{21} & \tilde{p}_{22} \end{bmatrix}_{w\times p} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}_{p\times p} \tag{1.59}$$

Where $\tilde{p}_{11}$ is n×m matrix, $\tilde{p}_{12}$ is m×p-m matrix, $\tilde{p}_{21}$ is p-m×m matrix, $\tilde{p}_{22}$ is p-m×p-m matrix,

6. $$\begin{bmatrix} B_{11} & 0 \\ 0 & B_{22} \end{bmatrix}_{w\times w} \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix}_{w\times w} = \left(\begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix}_{w\times w}\right)^T \begin{bmatrix} B_{11} & 0 \\ 0 & B_{22} \end{bmatrix}_{w\times w} \tag{1.60}$$

Proof:

1.

$$\left(\begin{bmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{bmatrix}\right)^2 = \begin{bmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{bmatrix}\begin{bmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{bmatrix}$$

$$= \begin{bmatrix} B_{11}^{-1} & 0 \\ 0 & B_{22}^{-1} \end{bmatrix}\begin{bmatrix} A_{11} & 0 \\ A_{21} & A_{22} \end{bmatrix}^T \begin{bmatrix} e_{11} & e_{12} \\ e_{21} & e_{22} \end{bmatrix}\begin{bmatrix} A_{11} & 0 \\ A_{21} & A_{22} \end{bmatrix}\begin{bmatrix} B_{11}^{-1} & 0 \\ 0 & B_{22}^{-1} \end{bmatrix}\begin{bmatrix} A_{11} & 0 \\ A_{21} & A_{22} \end{bmatrix}^T \begin{bmatrix} e_{11} & e_{12} \\ e_{21} & e_{22} \end{bmatrix}\begin{bmatrix} A_{11} & 0 \\ A_{21} & A_{22} \end{bmatrix}$$

$$=\left[\begin{array}{c:c} B_{11}^{-1} & 0 \\ \hdashline 0 & B_{22}^{-1} \end{array}\right]\left[\begin{array}{c:c} A_{11} & 0 \\ \hdashline A_{21} & A_{22} \end{array}\right]^{T}\left[\begin{array}{c:c} e_{11} & e_{12} \\ \hdashline e_{21} & e_{22} \end{array}\right]\left[\begin{array}{c:c} A_{11} & 0 \\ \hdashline A_{21} & A_{22} \end{array}\right] \text{ by using (1.47)}$$

$$=\left[\begin{array}{c:c} q_{11} & q_{12} \\ \hdashline q_{21} & q_{22} \end{array}\right]$$

For simplicity we shall prove the following notation by the form.

2.

From (1.46) we get.

$P^2 = (I-q)^2 = I - 2q + q^2$, where p, I and q are w×w matrices see (1.46), (1.47)

$P^2 = I - 2q + q$

$= I - q$, by using (1.46) we have

$= p$

3.

$Aq = AB^{-1}A^{T}EA$, see (1.47)

$= A$, see (1.42)

4.

$AP = A(I - q)$, see (1.46)

$= A(I - B^{-1}A^{T}EA)$, see (1.47)

$= A - AB^{-1}A^{T}EA)$, see (1.42)

$= A - A$

$=0$

5.

$A\tilde{P} = AB^{-1}A^{T}\left(AB^{-1}A^{T}\right)^{-1}$, see (1.42) and (1.48)

$=I$

6.

$BP = B\cdot(I - q)$, see (1.46)

$= B\cdot(I - B^{-1}A^{T}EA)$, see (1.47)

$= B - A^{T}EA$ (1.61)

$P^{T}B = (I - q)^{T}B$, see (1.46)

$$= B - q^T B$$
$$= B - (B^{-1}A^T\left(AB^{-1}A^T\right)^{-1}A)^T B$$, see (1.42) and (1.47)
$$= B - (A^T\left(\left(AB^{-1}A^T\right)^{-1}\right)^T AB^{-1})B$$
$$= B - A^T EA \qquad (1.62)$$

From (1.61) and (1.62) we have.

$B\,P = P^T B$

1.2. The Tangent Plane of the Problem:

Let $M(\hat{z}) = [M_1(\hat{z}) \;\; M_2(\hat{z})]^T$ be the tangent plane of the $G(\hat{z})$, at $(\hat{z})$ $(G(\hat{z}))$ such that,

$$\begin{bmatrix} M_1(\hat{z}) \\ \cdots \\ M_2(\hat{z}) \end{bmatrix}_{w\times1} = \left\{ \begin{bmatrix} K_1 \\ \cdots \\ K_2 \end{bmatrix}_{p\times1} : \begin{bmatrix} A_{11}^T(\hat{z})_{\;n\times m} & A_{21}^T(\hat{z})_{\;n\times p-m} \\ 0 & A_{22}^T(\hat{z})_{\;p\text{-}m\times p-m} \end{bmatrix}_{w\times p} \begin{bmatrix} K_1 \\ \cdots \\ K_2 \end{bmatrix}_{p\times1} = 0 \right\} \qquad (1.63)$$

The matrix $P(\hat{z})$ is projection operator which projects any vector in R^w into $M(\hat{z})$.

It's clear that the term $-\tilde{P}(z)\,G(z)$ in (1.45b) generates the solution.

By multiplying both sides of (1.45b) by A(z) we get:

$$\begin{bmatrix} A_{11} & 0 \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} \dot{z}_x \\ \dot{z}_y \end{bmatrix} = -\begin{bmatrix} A_{11} & 0 \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix} \begin{bmatrix} B_{11}^{-1} & \\ & B_{22}^{-1} \end{bmatrix} \begin{bmatrix} \nabla_x F \\ \nabla_y F \end{bmatrix} - \begin{bmatrix} A_{11} & 0 \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} \tilde{p}_{11} & \tilde{p}_{12} \\ \tilde{p}_{21} & \tilde{p}_{22} \end{bmatrix} \begin{bmatrix} g_a \\ g_b \end{bmatrix}$$

From (1.58) and (1.59) we have.

$$\begin{bmatrix} A_{11} & 0 \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} \dot{z}_x \\ \dot{z}_y \end{bmatrix} = -\begin{bmatrix} g_a \\ g_b \end{bmatrix}$$

$$\begin{bmatrix} A_{11}\,\dot{z}_x \\ A_{21}\,\dot{z}_x + A_{22}\,\dot{z}_y \end{bmatrix} = -\begin{bmatrix} g_a \\ g_b \end{bmatrix}$$

From above and (1.12), (1.13) we have.

$$\begin{bmatrix} dg_a/dt \\ dg_b/dt \end{bmatrix} = -\begin{bmatrix} g_a \\ g_b \end{bmatrix} \qquad (1.64)$$

From (1.64) we have.

$$\begin{bmatrix} dg_1/dt \\ . \\ . \\ . \\ dg_m/dt \end{bmatrix} = -\begin{bmatrix} 1 & 0 & . & . & 0 \\ 0 & . & & & . \\ . & & . & & . \\ . & & & . & 0 \\ 0 & . & . & 0 & 1 \end{bmatrix} \begin{bmatrix} g_1 \\ . \\ . \\ . \\ g_m \end{bmatrix} \tag{1.65}$$

Then

$$\frac{dg_i(z(t))}{dt} = -g_i(z(t)) \qquad i = 1,\ldots,m \tag{1.66}$$

Let the solution of (1.66) by the form

$$g_i(z(t)) = g_i(z(0))e^{-t} \qquad i = 1,\ldots,m \tag{1.67}$$

Since any solution must satisfy the system of differential equation (1.66).

$$\frac{dg_i(z(t))}{dt} = \frac{d}{dt}(g_i(z(0))e^{-t}), \qquad i = 1,\ldots,m$$

$$\frac{dg_i(z(t))}{dt} = -g_i(z(0))e^{-t} \qquad , \qquad i = 1,\ldots,m$$

$$\frac{dg_i(z(t))}{dt} = -g_i(z(t)) \qquad , \qquad i = 1,\ldots,m$$

From (1.66) we have.

$$\begin{bmatrix} g_1(z(t)) \\ . \\ . \\ . \\ g_m(z(t)) \end{bmatrix} = \begin{bmatrix} g_1(z(0)) \\ . \\ . \\ . \\ g_m(z(0)) \end{bmatrix} e^{-t} \tag{1.68}$$

By the same way from (1.64) we have.

$$\begin{bmatrix} g_{m+1}(z(t)) \\ . \\ . \\ . \\ g_p(z(t)) \end{bmatrix} = \begin{bmatrix} g_{m+1}(z(0)) \\ . \\ . \\ . \\ g_p(z(0)) \end{bmatrix} e^{-t} \tag{1.69}$$

From (1.68) and (1.69) we have.

$$\begin{bmatrix} g_1(z(t)) \\ \cdot \\ \cdot \\ \cdot \\ g_m(z(t)) \\ g_{m+1}(z(t)) \\ \cdot \\ \cdot \\ \cdot \\ g_p(z(t)) \end{bmatrix} = \begin{bmatrix} g_1(z(0)) \\ \cdot \\ \cdot \\ \cdot \\ g_m(z(0)) \\ g_{m+1}(z(0)) \\ \cdot \\ \cdot \\ \cdot \\ g_p(z(0)) \end{bmatrix} e^{-t} \tag{1.70}$$

To solve the optimization problem (1.6) by using the differential equation system (1.11) and (1.13), we have established first the following notation.

$$y = [0 \ , \ . \ . \ . , \ 0 \ \ y_{m+1} \ \ , . \ . \ . \ , \ y_p]^T \tag{1.71}$$

$$Y = \begin{bmatrix} 0 & & & & & & & \\ & \cdot & & & & 0 & & \\ & & \cdot & & & & & \\ & & & \cdot & & & & \\ & & & & 0 & & & \\ & & & & & y_{m+1} & & \\ & & & & & & \cdot & \\ & 0 & & & & & \cdot & \\ & & & & & & & \cdot \\ & & & & & & & y_p \end{bmatrix}, \ \Lambda = \begin{bmatrix} 1 & & & & & & & \\ & \cdot & & & & 0 & & \\ & & \cdot & & & & & \\ & & & \cdot & & & & \\ & & & & 1 & & & \\ & & & & & \lambda_{m+1} & & \\ & & & & & & \cdot & \\ & 0 & & & & & \cdot & \\ & & & & & & & \cdot \\ & & & & & & & \lambda_p \end{bmatrix} \tag{1.72}$$

$$\ell(x, \lambda) = f + \lambda^T \ \ h \tag{1.73}$$

$$L(z, \lambda) = F + \lambda^T \ \ G \tag{1.74}$$

Where f, λ, h are defined in (1.1), (1.2) and F, λ, G are defined in (1.5). We will also partition the vector $u \subset R^p$ and the $p \times p$ diagonal matrix U (U=diag (u)) into a part corresponding to the equality constrained and a part corresponding to the inequality constrained. These parts will be subscripted with E (Equality) and I (Inequality) respectively, i.e.

$$U_E = \begin{bmatrix} u_1 & & & & \\ & \cdot & & & \\ & & \cdot & & \\ & & & \cdot & \\ & & & & u_m \end{bmatrix}, \quad U_I = \begin{bmatrix} u_{m+1} & & & & \\ & \cdot & & & \\ & & \cdot & & \\ & & & \cdot & \\ & & & & u_p \end{bmatrix} \tag{1.75}$$

$$u_I = \begin{bmatrix} u_{1`} \\ . \\ . \\ . \\ u_m \end{bmatrix}, \quad u_E = \begin{bmatrix} u_{m+1`} \\ . \\ . \\ . \\ u_p \end{bmatrix} \tag{1.76}$$

From (1.75) we get.

$$\nabla_z L(z,\lambda) = \left[\nabla_z (F + \lambda^T G)\right]^T$$
$$= \left[\nabla_x f + \nabla_x \lambda^T G, \nabla_y \lambda^T G\right]^T$$

Since from (1.73) we get

$$\nabla_x \ell(z,\lambda) = \nabla_x f + \nabla_x \lambda^T \ h \tag{1.77}$$

From (1.75)

$$\nabla_y \lambda^T = \Lambda_I \tag{1.78}$$

From (1.75) and (1.76)

$$\nabla_y G = y_I \tag{1.79}$$

Then from (1.77), (1.78) and (1.79) we have.

$$\nabla_z L(z,\lambda) = \begin{bmatrix} \nabla_x \ell(z,\lambda) \\ \cdots\cdots\cdots \\ \Lambda_I y_I \end{bmatrix}_{w\times 1} \tag{1.80}$$

From (1.74) we get.

$$\nabla_{(z,\lambda)} L(z,\lambda) = \left[\nabla_{(z,\lambda)} (F + \lambda^T G)\right]^T$$
$$= \left[\nabla_x f + \nabla_x \lambda^T G, \nabla_y \lambda^T G, G\right]^T$$

$$\nabla_{(z,\lambda)} L(z,\lambda) = \begin{bmatrix} \nabla_x \ell(z,\lambda) \\ \cdots\cdots\cdots \\ \Lambda_I y_I \\ \cdots\cdots\cdots \\ G \end{bmatrix}_{2p-m+n\times 1} \tag{1.81}$$

$$\nabla_z^2 L(z,\lambda) = \nabla_z \begin{bmatrix} \nabla_x \ell(z,\lambda) \\ \text{-----------} \\ \Lambda_I y_I \end{bmatrix} = \begin{bmatrix} \nabla_z (\nabla_x \ell(z,\lambda)) \\ \text{-----------} \\ \nabla_z (\Lambda_I y_I) \end{bmatrix} \tag{1.82}$$

$$\nabla_z (\nabla_x \ell(z,\lambda)) = \left[\nabla_x^2 \ell(z,\lambda), \nabla_y \nabla_x \ell(z,\lambda)\right]^T$$

By using (1.77)

$$= \left[\nabla_x^2 \ell(z,\lambda), 0\right]^T \tag{1.83}$$

$$\nabla_z(\Lambda_I y_I) = [\nabla_x(\Lambda_I y_I), \nabla_y(\Lambda_I y_I)]^T$$
$$= [0, \Lambda_I]^T \tag{1.84}$$

Then from (1.82), (1.83) and (1.84) we get.

$$\nabla_z^2 L(z,\lambda) = \begin{bmatrix} \nabla_x^2 \ell(z,\lambda) & \vdots & 0 \\ \cdots & & \cdots \\ 0 & \vdots & \Lambda_I \end{bmatrix}_{w\times w} \tag{2.85}$$

Where Λ_I is in (1.78), from (1.82) we have that.

$$\nabla_{(z,\lambda)}^2 L(z,\lambda) = \nabla_{(z,\lambda)} \begin{bmatrix} \nabla_x \ell(z,\lambda) \\ \cdots \\ \Lambda_I y_I \\ \cdots \\ G \end{bmatrix} = \begin{bmatrix} \nabla_{(z,\lambda)}(\nabla_x \ell(z,\lambda)) \\ \cdots \\ \nabla_{(z,\lambda)}(\Lambda_I y_I) \\ \cdots \\ \nabla_{(z,\lambda)} G \end{bmatrix} \tag{1.86}$$

$$\nabla_{(z,\lambda)}(\nabla_x \ell(z,\lambda)) = [\nabla_x^2 \ell(z,\lambda), \nabla_y \nabla_x \ell(z,\lambda), \nabla_\lambda \nabla_x \ell(z,\lambda), \nabla_\lambda \nabla_x \ell(z,\lambda)]^T$$
$$= [\nabla_x^2 \ell(z,\lambda), 0, \nabla g_a, \nabla g_b]^T \tag{1.87}$$

$$\nabla_{(z,\lambda)}(\Lambda_I y_I) = [\nabla_x(\Lambda_I y_I), \nabla_y(\Lambda_I y_I), \nabla_\lambda(\Lambda_I y_I), \nabla_\lambda(\Lambda_I y_I)]^T$$
$$= [0, \Lambda_I, 0, Y_I)]^T \tag{1.88}$$

$$\nabla_{(z,\lambda)} G(z) = [\nabla_x G, \nabla_y G, \nabla_\lambda G, \nabla_\lambda G]^T$$
$$= \begin{bmatrix} \nabla g_a & \vdots & 0 & \vdots & 0 & \vdots & 0 \\ \nabla g_b & \vdots & Y_I & \vdots & 0 & \vdots & 0 \end{bmatrix}^T \tag{1.89}$$

Then from (1.86), (1.87), (1.88) and (1.89) we get.

$$\nabla_{(z,\lambda)}^2 L(z,\lambda) = \left[\begin{array}{cc|cc} \nabla_x^2 \ell(z,\lambda) & 0 & \nabla g_a & \nabla g_b \\ 0 & \Lambda_I & 0 & Y_I \\ \hline \nabla g_a^T & 0 & 0 & 0 \\ \nabla g_b^T & Y_I & 0 & 0 \end{array}\right]_{2p-m+n\times 2p-m+n} \tag{1.90}$$

Remark (1.2.1):

By using the lemma (1.1.1) and the fact $\nabla_x^2 \ell(x,\lambda)$ and Λ_I are non-singular then the Hessian matrix given by (1.90) is not singular at the solution as long as we have strict complementarity. Strict complementarity means that at the solution not both λ_i and y_i are zero [1]. This assumption is both standard and mild.

1.3 Asymptotic Properties of the Fundamental Equation:

In this section we have driver asymptotic properties of the system of differential equation (1.45b), when $t \to \infty$. The relation between constrained optimal points of the original optimization problem (1.6) and asymptotically stable critical points of the system of differential equation (1.45b) will be clarified. For the stability theory of ordinary differential equation, see, [8], [9] and [10].

We have denoted the solution of the system (1.45b) which passes through Z = ζ, at t=0 by $\Pi(\zeta, t)$, and the whole trajectory by $C(\zeta)$. i.e.

$$C(\zeta) = \{\Pi(\zeta, t) \quad : t \in T(\zeta)\} \tag{1.91}$$

Where $T(\zeta) = (t_a(\zeta), t_b(\zeta))$ is the maximum interval of the existence solution. Let S be a subset of R^w. The set S is called an invariant set if $C(\zeta) \subset S$ for any $\zeta \in$ S. We have denoted a set of points which satisfy the following equality constraint by S. i.e.

$$S = \{z: G(z) = 0, z \in R^w\} \tag{1.92}$$

It's clear that S is invariant set [1].

We can easily prove from (1.70) that, if $\Pi(\zeta, t)$ has a positive limit set $\Gamma^+(\zeta)$,

$$\lim_{t\to\infty}\begin{bmatrix} g_1(z(t)) \\ . \\ . \\ . \\ g_m(z(t)) \\ g_{m+1}(z(t)) \\ . \\ . \\ . \\ g_p(z(t)) \end{bmatrix} = \lim_{t\to\infty}\begin{bmatrix} g_1(z(0)) \\ . \\ . \\ . \\ g_m(z(0)) \\ g_{m+1}(z(t)) \\ . \\ . \\ . \\ g_p(z(0)) \end{bmatrix} e^{-t} = 0$$

Then $\Gamma^+(\zeta) \subset S$ (1.93)

The following important theorem, has been proved.

<u>**Theorem 1.3.1:**</u>

If D (z^*) in (1.21) is non singular matrix, a necessary and sufficient condition that (z^*) is a constrained stationary point is that (z^*) is a critical point of the system (1.11), (1.13).

<u>*Proof*</u>

If (z^*) is a constrained stationary point, then it satisfies the necessary condition (1.2) of the problem (1.6) there exists a vector $\lambda(z^*) \in R^p$ such that.

$$\begin{bmatrix} A_{11} & 0 \\ A_{21} & A_{22} \end{bmatrix}^T \begin{bmatrix} \lambda_1 \\ . \\ . \\ . \\ \lambda_m \\ \lambda_m \\ . \\ . \\ . \\ \lambda_p \end{bmatrix} + \begin{bmatrix} \nabla_{x_1} F \\ . \\ . \\ . \\ \nabla_{x_n} F \\ \nabla_{y_{m+1}} F \\ . \\ . \\ . \\ \nabla_{y_{p-m}} F \end{bmatrix} = 0, \begin{bmatrix} g_1 \\ . \\ . \\ . \\ g_m \\ g_{m+1} \\ . \\ . \\ . \\ g_p \end{bmatrix} = 0 \tag{1.94}$$

Along the point z=z*

And hence by using (1.94) and the invertible of B we have $\dot{z} \equiv 0$ at (z*) and $\lambda = \lambda^*$, Since D(z*) of (1.22) is nonsingular, then the system (1.11) and (1.13) has the unique solution on $\dot{z}$ and λ and by using the result above and (1.45a) we have. $\varphi(z^*) = 0$.

Conversely,

If (z*) is a critical point of the system (1.11) and (1.13), then $\varphi(z^*) = 0$

By using (1.45a), we have $\dot{z} \equiv 0$.Since D(z*) of (1.22) is nonsingular, then the system (1.11) and (1.13) has the unique solution on $\dot{z}$ and λ and by using the result above and the system (1.11) and (1.13), we have.

$$\begin{bmatrix} A_{11} & 0 \\ A_{21} & A_{22} \end{bmatrix}^T \begin{bmatrix} \lambda_1 \\ . \\ . \\ . \\ \lambda_m \\ \lambda_m \\ . \\ . \\ . \\ \lambda_p \end{bmatrix} + \begin{bmatrix} \nabla_{x_1} F \\ . \\ . \\ . \\ \nabla_{x_n} F \\ \nabla_{y_{m+1}} F \\ . \\ . \\ . \\ \nabla_{y_{p-m}} F \end{bmatrix} = 0, \begin{bmatrix} g_1 \\ . \\ . \\ . \\ g_m \\ g_{m+1} \\ . \\ . \\ . \\ g_p \end{bmatrix} = 0$$

Along the point z=z*

Then (z*) satisfies the necessary condition (1.2) of the problem (1.6), and thus (z*) is a constrained stationary point.

From the above theorem, it's clear that λ is the Lagrange multiplier of the original problem (1.6). The next corollary can be proved from the theorem and the second-order sufficient condition for local optimal solution.

Theorem 1.3.2:

Suppose there is a point z^* and a $\lambda(z^*) \in R^p$ satisfying (1.94), suppose also that the matrix

$$H = \nabla^2 F(z^*) + \lambda(z^*)^T \nabla^2 G(z^*) \tag{1.95}$$

(see (equation 1.6) and (theorem 1.3.1) for details)
is positive definite on the tangent plane M(z) of (1.63), that is, for $K \in M(z),\ K \neq 0$ there holds $K^T H(z^*) K > 0$. Then z^* is a strict local minimum of F(z) subject to G(z)=0.

Proof:

See [11]

Corollary 1.3.1:
If at the critical point z^* the matrix

$$H = \nabla^2 F(z^*) + \lambda(z^*)^T \nabla^2 G(z^*)$$

is positive definite on the tangent plane $M(z^*)$ (see (section 1.2)) of the constraint surface, that is $K^T HK > 0$ for any $K \in M(z^*), K \neq 0$, then z^* is a strict local minimal point.

Proof:

Let z^* be a critical point of the system (1.45b), by (theorem (1.3.1)), then $\varphi(z^*) = 0$

Since $D(z^*)$ is non singular at z^*, and from theorem (1.3.1) z^* is constrained stationary point, as well as there exists $\lambda(z^*)$ satisfying (1.94).
If (1.96) is positive definite on the tangent plane $M(z^*)$ of (1.63) of the constraint surface , for any $K \in M(z^*), K \neq 0$,
by theorem (1.3.2). Then z^* is a strict local minimal point.

Theorem 1.3.3:

Let (z^*) be a strict local minimal point of (1.6), if there exists a neighborhood E1 of (z^*), for any $(z) \in E1 \cap S$, the matrix B (see (1.10)) is positive definite on the tangent plane M(z) (see (section 1.2)) of (1.63), then (z^*) is asymptotically stable on S.

Proof:

Since S is invariant (see (section 1.3)), it's sufficient to consider trajectories on S.
Then for a solution $z(t) \in E1 \cap S$,

$$\frac{dF}{dt} = -\nabla F^T \dot{z}$$

$$\frac{dF}{dt} = -\nabla F^T P\, B^{-1} \nabla F, \text{ from (1.56),}$$

$$= -\nabla F^T PP\, B^{-1} \nabla F$$

And from (1.60)

$$= -\nabla F^T B^{-1} P^T B P B^{-1} \nabla F$$
$$= -\dot{z}^T B \dot{z}$$

Where $\nabla F, P, B$ in (1.45a),
We have from the assumptions.

$$\frac{dF(z)}{dt} < 0 \quad , (z) \neq (z^*) \tag{1.96}$$

$$\frac{dF(z)}{dt} = 0 \quad , \text{at } (z^*) \tag{1.97}$$

From the other hand the strict minimality of z^*, there exists a neighborhood E2 of z^* such that for any $z \in E2 \cap S$, that.

$$F(z) > F(z^*) \quad , \quad z \neq z^* \tag{1.98}$$

From (1.96) and (1.97) the function
$V(z) = F(z) - F(z^*)$

Satisfies the following condition,

$V(z) > 0 \qquad , z \neq z^*$

$V(z) = 0 \qquad , \text{at } z^*$

$$\frac{dV(z)}{dt} < 0 \qquad , z \neq z^*$$

$$\frac{dV(z)}{dt} = 0 \qquad , \text{at } z^*$$

For $z(t) \in E1 \cap E2 \cap S$

Thus the function V(z) is Lyapunov function, and z^* is asymptotically stable on S, of the system (1.45b).

<u>Corollary 1.3.2:</u>

Let z^* be a strict local maximal point of (1.6), if there exists a neighborhood E1 of z^*, for any $z \in E1 \cap S$, the matrix B of (1.10) is negative definite on the tangent plane M(z) of (1.63), then z^* is asymptotically stable on S.

<u>Proof</u>

It's trivial, (see (theorem 1.3.3)).

<u>Remark (1.3.1):</u>

The above theorem is concerned with the asymptotic stability on the set S, next we shall consider stability on the space R^w.

Lemma 1.3.1:

A local minimal (maximal) point z^* which satisfies the assumption of Theorem 1.3.3 (Corollary 1.3.2) is stable.

Proof:

Suppose that z^* is not stable. Then for any neighborhoods E1 and E2 of z^*, there exists a point $z \in E2$ such that,

$C^+(z) \not\subset \overline{E1}$, ($\overline{E1}$ is the closure of E1)

$C^+(z) = \{\Pi(z,t) \ : 0 \le t \in T(z)\}$

Where T(z) is the maximum interval of the existence of the solution.

Since we can choose a sequence $\{z^k\} \subset E2$, converging to z^* such that $C^+(z^k) \not\subset \overline{E1}$ and clearly there exists $t^k>0$.

Where

$\Pi(z^k,t) \in E1 \qquad , 0 \le t \le t^k$

$J^k \equiv \Pi(z^k,t^k) \in \partial E1$, ($\partial E1$ is the boundary of $\overline{E1}$, since $\partial E1$ is compact).

$\{J^k\}$ must have a convergent subsequence to say J.

For convergence of notation we assume that the sequence $\{J^k\}$ is itself convergent to J.

From $G(z^*)=0$, we have

$G(z^k) \to 0$, where $z^k \to z^*$

From (1.70) we have

$G(J^k)= G(z^k)\ e^{-t^k}$

And from the fact that $G(z^k) \to 0,\ \ t^k \to 0$,and by the continuity $G(J^k) \to G(J)$

We can conclude that

G (J) =0, i.e. $J \in S$.

From the relation

$z^k(t^k)=\Pi(J^k,-t^k)$ and the continuity of $\Pi(z,t)$ with respect to z and t we have.

$\Pi(J,-t) \in E2 \cap S$, for some t>0.

This contradicts the asymptotic stability of z^* on S (theorem 1.3.3), since E2 is arbitrary. Thus for any neighborhood E1 of z^*, there exists a neighborhood E2 of z^* such that $C^+(z) \subset E1$ for any $z \in E2$.

This is stability of z^*.

Clearly, stability alone is not sufficient to solve the optimization problem (1.6), the next theorem is of the utmost importance in this section.

Theorem 1.3.4:

A local minimal (maximal) point z^* which satisfies the assumption of Theorem 1.3.3 (Corollary 1.3.2) is asymptotically stable.

Proof:

From the stability of z^* for any neighborhood E1 of z^* there exists a neighborhood E2 of z^* such that $C^+(z) \subset E1$. For any $z \in$ E2. Therefore the positive limit set $\Gamma^+(z)$ satisfies $\Gamma^+(z) \subset \overline{E1}$ [14].

From (1.93) we have.

$\Gamma^{+}(z) \subset \overline{E1_s}$

Where $E1_s \equiv S \bigcap E1$,

By the asymptotic stability of z^* on S, we can choose E1 such that.

$\Gamma^{+}(J) = (z^*)$

For any $J \in \overline{E1_s}$, because E1 is arbitrary, If we take

$J \in \Gamma^{+}(z) \subset \overline{E1_s}$

Then from $\Gamma^{+}(z) = \Gamma^{+}(J)$

We have

$\Gamma^{+}(z) = z^*$.

Thus z^* is asymptotically stable.

It's easily seen from the above theorem that if the differential equation satisfies the assumption of Theorem 1.3.3 (Corollary 1.3.2) any trajectory starting from a point within some (sufficiently small) neighborhood of the local (minimal / maximal) point z converges to z^* as $t \to \infty$.

1.4 Useful Comments:

The following comments for numerical solution of optimization problem (1.6) are based on theoretical above.

Step 1: Trajectory continuation:

1. From the discussion given in (section (1.3)), it's clear that we can use the system of differential equation (1.11) and (1.13) to solve the constrained optimization problem (1.6). However in that original form (1.6), the region of the convergence is restricted to some neighborhood of the critical point.

2. To enlarge the region of the convergence, the most obvious way to do this is the use of the Marquardt-Levenberg [12] for making the matrix B(z) of (1.10) positive definite on tangent plane M(z) (see (section 1.2)), and one can use the technique of the trajectory continuation[1].

3. The trajectory continuation can be simplified as follows.

We consider the nature of the trajectories on the feasible set

$S = \{(z) \ : \ G(z) = 0 \quad , z \in R^{p}\}$

Let ζ be a point on S, and $D(\zeta)$ of (1.22) be nonsingular. Then we can define a unique solution $\Pi(\zeta, t)$ passing through ζ at t=0, the nature of the trajectory $C(\zeta)$ (see section (1.3)) is classified according to the maximal interval of existence of the solution $T(\zeta) = (t_a(\zeta), t_b(\zeta))$, and notice that :

a) If $t_a(\zeta)(t_b(\zeta))$ is finite the trajectory $C(\zeta)$ (see section (1.3)) starts from (ends at) the point (ζ_0), where $D(\zeta_0)$ is singular.

b) If $t_a(\zeta)(t_b(\zeta))$ is not bounded $C(\zeta)$ (see section (1.3)) will tend to infinity, a critical point or cyclic trajectory. Other wise $C(\zeta)$ it's self is a critical point or cyclic trajectory. Cyclic trajectory is possible only when there exists region U in which B(z) (see (1.10)) is indefinite on the tangent plane $M(z)$, $z \in U$. That does not mean a region with an indefinite B(z) always results in cyclic trajectories [1].

For trajectory which ends up at the singular point of the matrix D(z) (see (1.22)) continuation of the trajectory is possible by using Branin's idea [1] of sign reversal of the equation .

We shall now formulate the method of trajectory continuation. We shall introduce the signature into the system

$$\dot{z} \equiv \varphi_\mu(z) = -\mu P(z)B(z)^{-1}\nabla F(z) \tag{1.99}$$

Where $\mu = \mp 1$

The system (1.99) is identical to the system (1.45a) with B(z) replaced by μB(z).Trajectory with signatures of the opposite signs passing through the same point differs only in the direction of the trajectory which is defined with respect to the subsidiary variable t. This can be easily seen from the relation

$$\varphi_+(z) = -\varphi_-(z) \tag{1.100}$$

Where

$\varphi_\pm(z)$ denotes $\varphi_\mu(z)$

With positive / negative signature

Now continuation may be possible between the trajectories with opposite signatures which end and start at the point (ζ_0) where $|D(\zeta_0)| = 0$ (determinant of the matrix $D(\zeta_0)$) respectively. This rule is also used for the continuation of a cross a critical point.

Note that, it may not be always possible to continue trajectories. But suppose that these continuations are possible. Then we can continue to integrate the trajectory and find local constrained extrema one after another.

In general, it will not be possible to locate all of the extrema, but there are cases in which it becomes possible.

For example, let the feasible set S of (1.92) be bounded and connected. Then it's possible for the continued trajectory to come into regions of convergence of all the extrema and to pass through them.

We can generalize these procedures to the case of trajectory from infeasible point, in this case the equation becomes.

$$\dot{z} \equiv \varphi_\mu(z) = -\mu P(z)B(z)^{-1}\nabla F(z) - \tilde{P}(z)G(z) \tag{1.101}$$

Where the signature is included only in first term of $\varphi_\mu(z)$ to preserve the solution,

$$\begin{aligned}&\left[g_1(z(t)) \quad ,\dots, \quad g_m(z(t)) \quad g_{m+1}(z(t)) \quad ,\dots, \quad g_p(z(t))\right]^T \\ &\qquad = \left[g_1(z(0)) \quad ,\dots, \quad g_m(z(0)) \quad g_{m+1}(z(0)) \quad ,\dots, \quad g_p(z(0))\right]^T e^{-t}\end{aligned}$$

Step 2: Selection of Arbitrary matrix B(z)*:*

The advantage of this method is the depending of the arbitrary selection of the matrix B(z).where

$$B(z) = \begin{bmatrix} B_{11}(z)_{n\times n} & 0_{n\times p-m} \\ 0_{p-m\times n} & B_{22}(z)_{p-m\times p-m} \end{bmatrix}_{w\times w} \tag{1.102}$$

The selection can be done so that the following conditions are satisfied.

1. The matrix B (z) of (1.10) should be positive definite matrix to overcome the problem of singularity in (1.11), (1.22) and to ensure it's invertibility ($B(z)^{-1}$ exists).

2. B(z) should be selected in such a way that the linearized system

$$\dot{z} \equiv \varphi_\mu(z) = -\mu P(z)B(z)^{-1}\nabla F(z) - \tilde{P}(z)G(z)$$

is stable, to ensure the solution goes to its critical point as time increases without tend where μ is the switching point.

Note that, if B(z) is constant matrix and the system (1.11) and (1.13) is time invariant one can select it by using the Eigenvalues of the system (1.11) and (1.13) (future work).

Step 3: Explicit form of the matrix B(z):

In this step we have suggested examples of the matrix B(z) to give a definite algorithm for solving the optimization problem (1.6)

1. The most simple form of B(z)will be

$$B=I \tag{1.103}$$

The system (1.101) becomes

$$\varphi_\mu(z) = -\mu p(z)\nabla F(z)^T - \tilde{p}(z)G(z) \tag{1.104}$$

and the equations (1.42),(1.46)-(1.48) become:

$$p=\left[\begin{array}{c:c}(p_{11})_{n\times n} & (p_{12})_{n\times p\text{-}m}\\ \hdashline (p_{21})_{p\text{-}m\times n} & (p_{22})_{p\text{-}m\times p\text{-}m}\end{array}\right]=\left[\begin{array}{c:c}I_{n\times n} & 0_{n\times p\text{-}m}\\ \hdashline 0_{p\text{-}m\times n} & I_{p\text{-}m\times p\text{-}m}\end{array}\right]-\left[\begin{array}{c:c}(q_{11})_{n\times n} & (q_{12})_{n\times p\text{-}m}\\ \hdashline (q_{21})_{p\text{-}m\times n} & (q_{22})_{p\text{-}m\times p\text{-}m}\end{array}\right] \tag{1.105}$$

$$q=\left[\begin{array}{c:c}(q_{11})_{n\times n} & (q_{12})_{n\times p\text{-}m}\\ \hdashline (q_{21})_{p\text{-}m\times n} & (q_{22})_{p\text{-}m\times p\text{-}m}\end{array}\right]=\left[\begin{array}{c:c}(B_{11}^{-1})_{n\times n} & 0_{n\times p\text{-}m}\\ \hdashline 0_{p\text{-}m\times n} & (B_{22}^{-1})_{p-m\times p-m}\end{array}\right]\left[\begin{array}{c:c}(A_{11})_{m\times n} & 0_{m\times p\text{-}m}\\ \hdashline (A_{21})_{p-m\times n} & (A_{22})_{p-m\times p-m}\end{array}\right]^{T}$$
$$\left[\begin{array}{c:c}(e_{11})_{m\times m} & (e_{12})_{m\times p-m}\\ \hdashline (e_{21})_{p-m\times m} & (e_{22})_{p-m\times p-m}\end{array}\right]\left[\begin{array}{c:c}(A_{11})_{m\times n} & 0_{m\times p\text{-}m}\\ \hdashline (A_{21})_{p-m\times n} & (A_{22})_{p-m\times p-m}\end{array}\right] \tag{1.106}$$

$$\tilde{p}=\left[\begin{array}{c:c}(\tilde{p}_{11})_{n\times m} & (\tilde{p}_{12})_{m\times p\text{-}m}\\ \hdashline (\tilde{p}_{21})_{p\text{-}m\times n} & (\tilde{p}_{22})_{p\text{-}m\times p\text{-}m}\end{array}\right]=\left[\begin{array}{c:c}(B_{11}^{-1})_{n\times n} & 0_{n\times p\text{-}m}\\ \hdashline 0_{p\text{-}m\times n} & (B_{22}^{-1})_{p-m\times p-m}\end{array}\right]\left[\begin{array}{c:c}(A_{11})_{m\times n} & 0_{m\times p\text{-}m}\\ \hdashline (A_{21})_{p-m\times n} & (A_{22})_{p-m\times p-m}\end{array}\right]^{T}$$
$$\left[\begin{array}{c:c}(e_{11})_{m\times m} & (e_{12})_{m\times p-m}\\ \hdashline (e_{21})_{p-m\times m} & (e_{22})_{p-m\times p-m}\end{array}\right] \tag{1.107}$$

$$E=\left[\begin{array}{c:c}(e_{11})_{m\times m} & (e_{12})_{m\times p-m}\\ \hdashline (e_{21})_{p-m\times m} & (e_{22})_{p-m\times p-m}\end{array}\right]=\left(\left[\begin{array}{c:c}(A_{11})_{m\times n} & 0_{m\times p\text{-}m}\\ \hdashline (A_{21})_{p-m\times n} & (A_{22})_{p-m\times p-m}\end{array}\right]\left[\begin{array}{c:c}(A_{11})_{m\times n} & 0_{m\times p\text{-}m}\\ \hdashline (A_{21})_{p-m\times n} & (A_{22})_{p-m\times p-m}\end{array}\right]^{T}\right)^{-1} \tag{1.118}$$

Tanabe [1] proposed the similar equation

$$\dot{z}=-p(z)\nabla F(z)^{T} \tag{1.109}$$

Which is direct generalization of Rosen's projection method [13] to a differential form, the equation (1.109) is identical to (1.104) on the feasible set S except for existence of the signature. Though a local minimal point is asymptotically stable for (3.17) with positive signature, it is true only on the set S for (1.109). Moreover, even if we have calculated the trajectory starting from a point on S, the existence of the calibration term $-\tilde{p}(z)G(z)$ in (1.104) is very desirable from a numerical viewpoint. From the discussion given in Section (2.3), we conclude that a local minimal (maximal) point is asymptotically stable under the positive (negative) signature equation, and unstable under the negative (positive) signature equation. From a computational point of view, however, the system (1.104)-(1.107) has a serious defect, because trajectories described by the system (1.104)-(1.107) may become complicated near the critical point, say, z^* .This is easily seen from the equation near z^*.

$$\frac{dz}{dt}=\frac{\partial\phi_\mu(z^*)}{\partial z}(z-z^*) \tag{1.110}$$

$$\varphi_\mu(z^*)=-\mu p(z^*)\nabla F(z^*)^T-\tilde{p}(z^*)G(z^*)$$

$$\frac{\partial\varphi_\mu(z^*)}{\partial z^*}=\frac{\partial}{\partial z^*}\left[-\mu p(z^*)\nabla F(z^*)^T-\tilde{p}(z^*)G(z^*)\right]$$
$$=-\mu\left[\frac{\partial}{\partial z^*}p(z^*)\nabla F(z^*)^T+\frac{\partial}{\partial z^*}\mu\tilde{p}(z^*)G(z^*)\right]$$

$$= -\mu\left[\left(\frac{\partial}{\partial z^*}p(z^*)\right)\nabla F(z^*)^T + p(z^*)\nabla^2 F(z^*) + \left(\frac{\partial}{\partial z^*}\mu\tilde{p}(z^*)\right)G(z^*) + \mu\tilde{p}(z^*)\nabla G(z^*)\right]$$

By using (1.105), (1.107) and some calculation we get,

$$\frac{\partial\varphi_\mu(z^*)}{\partial z^*} = -\mu\{\nabla^2 F(z^*) + \lambda(z^*)^T\nabla^2 G(z^*) - A(z^*)(A(z^*)A(z^*)^T)^{-1}A(z^*)$$
$$\left[\nabla^2 F(z^*)^T + \lambda(z^*)^T\nabla^2 G(z^*) - \mu I\right]\}$$

2. The second form we have proposed as.

$$B(z) = \nabla^2 F(z) + \lambda_0(z)^T\nabla^2 G(z)$$
$$\lambda_0(z) = -(A(z)A(z)^T)^{-1}A(z)\nabla F(z)^T$$

Where $\lambda_0(z)$ is the first-order estimate of the Lagrange multiplier λ^*, that is
$\|\lambda_0(z) - \lambda(z^*)\| \sim \|z - z^*\|$

If the local minimal/maximal point satisfies the second-order sufficient condition, then the assumption of theorem 1.3.3 (Corollary 1.3.2) is satisfied by the continuity of B(z). Therefore, a local minimal/maximal point is asymptotically stable under the positive signature equation, and unstable under the negative signature equation. The difficulty encountered in the equation with B=I is now removed. This is easily seen from the relation

$$\frac{\partial\varphi_+(z^*)}{\partial z} = -B(z)^{-1}\left[\nabla^2 F(z) + \lambda(z)^T\nabla^2 G(z) - A(z)^T(A(z)B(z)^{-1}A(z)^T)^{-1}\right.$$
$$\left.A(z)B(z)^{-1}(\nabla^2 F(z) + \lambda(z)^T\nabla^2 G(z) - B(z))\right]_{Z=Z^*}$$
$$= -I$$

The solution near z^* becomes (μ=+1)
$$z(t) \approx z^* + \left(z(0) - z^*\right)e^{-t}$$

We use this form of B(z) in our work.

Step 4:

Calculate the matrix $B = \left[\begin{array}{c|c} B_{11}(z)_{n\times n} & 0_{n\times p-m} \\ \hline 0_{p-m\times n} & B_{22}(z)_{p-m\times p-m}\end{array}\right]_{w\times w}$ in such a way, (see (step 3)).

Step 5:

Calculate $\varphi_\mu(z) = -\mu P(z)B(z)^{-1}\nabla F(z) - \tilde{P}(z)G(z)$, (see (step 1))

Step 6: **(Method of Integration)**

Some numerical methods of integration can be used for system of differential equation to integrate the trajectory numerically if the inequalities

$\|\varphi_{+}(z^{k})\| \le \varepsilon_1$, $\|G(z^{k})\| \le \varepsilon_2$, are satisfies for a suitable small positive real number ε_1 and ε_2.

And if the above inequalities are satisfied, then an approximate feasible critical point of the system (1.103) is obtained.

<u>Remark 3.2.1:</u>

In our work forth-order Runge-kutta method for system of differential equation is used to integrate the trajectory numerically.

Step 7:(**Trajectory continuation)**

If the determinant of D(z) in (1.22) changes it's sign the signature μ is changed.When a critical point is reached, the trajectory is projected through the point along the direction $-P(z)B(z)^{-1}\nabla F(z)$ and integration is continued with an opposite signature μ=-μ

<u>Remark 3.2.2:</u>

A MATLAB software has been used to obtain the numerical solution by using the above comments.

3.3 Illustration:

Case 1

Min $\quad f(x) = x_1^3 + x_2^3 + x_3^3$

Subject to $\quad h(x) = \dfrac{1}{4}x_1^2 + \dfrac{1}{2}x_2^2 + \dfrac{1}{2}x_3^2 - 1$

Initial point: $\quad x^0 = [1.0,2.0,3.0]^T$

Solution

It's clear that the above problem is nonlinear and it's not convex, then there exists infinite solution to this problem. To solve this example we use the second formula of the matrix B(z) (see (step 3)). We have solved this problem by using the above comments, and different step sizes of forth-order Runge-Kutta method, then we have got all possible critical points and hence the solution of the optimization problem is as discussed in (section 1.3).The numerical solution for different step-size of Runge-Kutta method is shown in the following table, where the symbol

H1 stands for $\left\|\varphi(z^k)\right\|$, (see critical point).

H2 stands for $h(x_1,x_1,x_3)$, (feasible region).

H3 stands for $f(x_1,x_1,x_3)$, (minimum value).

Note that some solutions are repeated (periodic solution), and in this table are omitted

Table (1)

x_1	x_2	x_3	H1	H2	H3
6.67E-01	1.33E+00	-3.27E-07	5.12E-07	5.29E-07	2.67E+00
-4.85E-01	-9.70E-01	-9.70E-01	7.96E-07	1.08E-06	-1.94E+00
5.64E-07	-1.00E+00	-1.00E+00	7.17E-07	3.79E-07	-2.00E+00
-1.31E-07	-1.41E+00	5.59E-08	5.20E-07	7.07E-07	-2.83E+00
6.67E-01	-2.76E-07	1.33E+00	6.70E-07	2.91E-07	2.67E+00
2.00E+00	5.28E-08	-6.27E-07	7.18E-07	3.44E-07	8.00E+00
-6.67E-01	3.66E-07	-1.33E+00	8.88E-07	3.20E-07	-2.67E+00
-4.75E-07	1.41E+00	-1.22E-08	4.89E-07	1.64E-07	2.83E+00
-2.00E+00	-1.30E-07	-9.32E-08	4.23E-07	3.92E-07	-8.00E+00
-6.24E-09	1.00E+00	1.00E+00	4.49E-07	6.06E-07	2.00E+00
-1.22E-07	5.07E-07	1.41E+00	5.40E-07	1.96E-07	2.83E+00
-6.67E-01	-1.33E+00	9.54E-08	6.13E-07	1.09E-07	-2.67E+00
-2.65E-07	-5.30E-07	-1.41E+00	8.24E-07	8.10E-07	-2.83E+00
9.76E-07	-1.00E+00	-1.00E+00	9.77E-07	5.25E-08	-2.00E+00
4.85E-01	9.70E-01	9.70E-01	9.89E-07	1.10E-07	1.94E+00

Table (1) presents all possible different solutions via all step-size
(1, 0.5, 0.25, 0.125, 0.0625, 0.03125) have obtained.

Case 2

Min $f(x) = (x_1 - 2)^2 + (x_1 - 2x_2)^2$

Subject to $x_1^2 - x_2 \le 0$

Initial point $[0,1]^T$

Solution

It's clear that the above problem is nonlinear and its convex, then there exist unique solution to this problem. To solve this example we use the second formula of the matrix B(z) (see (step 3)).

We solve the above problem by using the comments (step 5-step 7) and different step sizes of forth-order Runge-Kutta method, then we get the critical point and hence the solution of the optimization problem is as discussed in (section 1.3).The numerical solution for different step-size of Runge-Kutta method is shown in the following table, where the symbol

h stands for the step size.

H1 stands for $\|\varphi(z^k)\|$, (see critical point).

H2 stands for $h(x_1,x_1,x_3)$, (feasible region).

H3 stands for $f(x_1,x_1,x_3)$, (minimum value).

Table (2)

h	t	x_1	x_2	H1	H2	H3
1	16	0.94658	0.89413	9.58 E-07	-5.65 E-07	1.9462
0.5	31	0.94558	0.89413	1.45 E-07	3.61 E-07	1.9462
0.25	62	0.94558	0.89413	6.87 E-07	-8.29 E-07	1.9462
0.125	127	0.94558	0.89413	6.21 E-07	-7.19 E-07	1.9462
0.0625	253	0.94558	0.89413	7.65 E-07	2.90 E-07	1.9462
0.03125	505	0.94558	0.89413	1.92 E-07	-1.19 E-07	1.9462

Table (2) presents the solution via all possible step-size
(1, 0.5, 0.25, 0.125, 0.0625, 0.03125)

Remark 3.3.1:

The local solution has global on the convex nonlinear optimization problem.

3.6 The Advantage:

The main advantage of this method is to find all possible local solutions and a global solution which is very important in application and makes this method a good technique to be followed even when there exists the difficulties.

3.7 Future Study:

The following subjects will be taken into accounts in a future work, if possible

1. Sensitivity analysis of this method.
2. The optimum selection of the matrix B(z).
3. Generalizing this method to optimal control problem.

REFERENCE

1. Hiroshi, Y., “A Differential Equation Approach to Nonlinear Programming”, Mathematical Programming, Vol. 18, p.p.155-168, (1980).
2. Ali, F. M., “Technique For Solving Multobjective Nonlinear Programming Using Differential Equation Approach”, J. inform. Sci.,Vol. 18, p.p. 351-357, (1997).
3. Fatma M. Ali, ” sensitivity analysis for parametric vector optimization problem using differential equation approach”, Ijmms 25:9, pp.621-628, (2001).
4. Mauricio G.C.Resende, Panos M. Pardalos, And Sandra Duni Eksiogla. “Parallel metaheuristics for combinatorial optimization “, ttt://www.ima.umn.edu./index.html.
5. R .A .Tapia, “On the role of slack variables in Quasi-Newton method for constrained optimization ”, North-Holland puplishing, Pp.235-246,(1980).
6. David G. Luenberger, “ introduction to linear and nonlinear programming “, Addison-Wesley publishing Company ,(1973).
7. McCormick, G. P., “Nonlinear programming “, John Willy and Sons, (1983).
8. Nohal, J. A., and Brauer, F., “Qualitative Theory of Ordinary Differential Equation”, W. A. Benjamin, (1968).
9. Jane, C., “Differential Equation Introduction and Qualitative Theory”, Marcel Dekker,(1980).
10. Arrowsmith, D. K., and Place, C. M., “Ordinary Differential Equations”, Chapman and Hall, (1982).
11. Fiacco, A. V., and McCormick, G. P., “Nonlinear Programming Sequential Unconstrained Minimization Techniques “,John Willy and Sons, (1968).
12. Marquardt, D., “An Algorithmfor Least-Squares Estimation of Nonlinear Parameters “, SIAM Journal of Applied Mathematics, Vol. 11, p.p. 431-441,(1963).
13. Roseen, J. B., “The Gradient Projection Method for Nonlinear Programming. Part ii. Nonlinear Constraints”, SIAM Journal Vol 9, p.p. 514-432, (1961).

الخلاصة

لقد طورت طريقة عددية (معتمدين على الجانب التحليلي للمسالة) لحل مسائل الامثلية غير الخطية المقيدة (قيد الدراسة)، مستنده على إيجاد نظام من المعادلات التفاضلية غير الخطية وغير متغيرة مع الزمن (autonomous).

لقد درست تحليليا العلاقة بين النقاط الحرجة للمعادلات التفاضلية غير الخطية والحلول المحلية (local optimum) لمسالة الامثلية غير الخطية المقيدة (قيد الدراسة).

السلوك المحاذي واستقرارية النقاط الحرجة لنظام المعادلات التفاضلية غير الخطية تم دراسته ومناقشته.

أسلوبية عددية لإيجاد الحلول المثلى (local and global optimum) للمسالة الأصلية تم تطويرها بالاعتماد على سلوك وطبيعة حل المعادلات التفاضلية غير الخطية، وتم تطبيقها لإيجاد حلول لبعض المسائل المختارة لدراسة كفاءة الطريقة.