International Journal of Mathematics and Statistics Invention (IJMSI)
E-ISSN: 2321 – 4767 P-ISSN: 2321 - 4759
www.ijmsi.org Volume 4 Issue 4 || April. 2016 || PP-28-33

Optimality conditions for Equality Constrained Optimization Problems

EFOR, T. E
Department of Mathematics and Computer Science Ebonyi state University, Abakaliki. Nigeria

ABSTRACT : *This paper focuses on the equality constrained optimization problems. It considered the Lagrange multiplier method for solving equality constrained problem. The Lagrange First and second order optimality conditions for equality problems are discussed. Also the paper uses the fact that the definiteness of an $n \times n$ symmetric matrix can be characterized in terms of the sub matrix to apply the Hessian Borded Matrix method to check for optimality condition for equality constrained with many constraints. Examples are used to illustrate each of the method.*

KEYWORDS: *Bordered Hessian matrix, Constraint qualification, Equality constrained, Local minimizer, Lagrange multiplier.*

I. Introduction

Optimization is a mathematical procedure for determining optimal allocation of scarce resources. It has found practical applications in almost all sectors of life, for instance in economics, engineering etc. This can also be defined as the selection of the best element from available resources and achieve maximum output.

Generally, optimization problems are divided into two namely: unconstrained and constrained optimization problems. Unconstrained problems do not arise often in practical applications. Sundaram in his paper [1] considered such problems. Because optimality conditions for constrained problems become a logical extension of the conditions for unconstrained problems. Shuonan in [2] applied the constrained optimization methods on a simple toy problem .

Constrained optimization is approached somewhat differently from unconstrained optimization because the goal is not to find the global optima. However, constrained optimization methods use unconstrained optimization as a sub-step. In [3] ,the author used Newton's Method to solve for minimizers of an equality constraint problems. T his paper will only consider Lagrangian multiplier method and Hessian Borded matrix method to solve for optima of the equality constrained optimization problem.

For the purpose of understanding, I hereby state the standard form of unconstrained optimization problems as follows

$$Minimize f(x)$$

Subject to $x \in D$

where D is some open subset of $\mathbb{R}^n$

Here, to study this type of optimization problem, the study of the theory is under assumptions of differentiability, the principal objective here is to identify necessary conditions that the derivative of f must satisfy at an optimum. Unconstrained problem does not literally mean the absence of constraints, but it refers to a situation in which we can move (at least) a small distance away from the optimum point in any direction without leaving the feasible set.

II. Constrained Optimization Problems

Now, we turn our attention to constrained optimization problems as our interest

In constrained optimization problem, the objective function is subject to certain constraints. For instance, in industries, most constraints are encountered such as transportation, technology, time, labour, etc

The standard model for the constrained optimization problem is posed as follows;

$$Minimize\ f(x)$$
$$Subject\ to\ x \in K$$

where K is some special set in $\mathbb{R}^n$ that can assume any of the following forms in practice;

(i) $K_1 = P \cap \{x \in \mathbb{R}^n : g_i(x) = 0\}, i = 1, \ldots, m.$

(ii) $K_2 = P \cap \{x \in \mathbb{R}^n : h_j(x) \leq 0\}, j = 1, \ldots, k.$

(iii) $K_3 = P \cap \{x \in \mathbb{R}^n : g_i(x) = 0, h_j(x) \leq 0\}$

Where P is open subset of $\mathbb{R}^n$ and $g: \mathbb{R}^n \to \mathbb{R}^m, h: \mathbb{R}^n \to \mathbb{R}^k$,

Basically, there are essentially two types of constrained problems, namely equality and inequality constraints. But for the purpose of this paper, we concentrate on equality constrained problems.

2.1. Equality Constrained Problems:

The model is posed as

$$\begin{aligned} &minimize\ f(x) \\ &Subject\ to\ K_1 \end{aligned}$$

2.2. Inequality constrained problems

The model is posed as;

$$\begin{aligned} &Minimize\ f(x) \\ &Subject\ to\ K_2 \end{aligned}$$

We also have the general mixed constrained problem which is given by

$$\begin{aligned} &Minimize\ f(x) \\ &Subject\ to\ K_3 \end{aligned}$$

2.3. Notations

We give some notations and definitions that we will need throughout this paper .

(1).The gradient of a real-valued function f is given by

$$\nabla f(x) = \left(\frac{\partial}{\partial x_1} f(x), \dots, \frac{\partial}{\partial x_n} f(x)\right) = \frac{\partial}{\partial x_i} f(x), i = 1, \dots, n$$

(2).The Hessian matrix of f denoted by

$$Hf(x) = \left[\frac{\partial^2}{\partial x_i \partial x_j} f(x)\right]_{i,j=1,\dots,n} = \nabla^2 f(x)$$

(3).The Jacobian matrix of a vector – valued function $f(x) = (f_1(x), \dots f_m(x))^T$ can also be written in the form

$\nabla f(x) = (\nabla(f_1(x), \dots f_m(x))^T$,where, $\nabla f(x))^T$, means the transpose of $\nabla f(x)$.

The fundamental tool for deriving optimality condition in optimization algorithms is the so called lagrangain function

$\mathcal{L}(x,\lambda) = f(x) + \sum_{i=1}^m \lambda_i g_i(x), for\ all\ x \epsilon \mathbb{R}^n, \lambda = (\lambda_1, \lambda_2, \dots, \lambda_k) \epsilon \mathbb{R}^n$.. The purpose of $\mathcal{L}(x,\lambda)$ is to link the objective function $f(x)$ and the constraints set, $g_i(x) = 0, i = 1, \dots, k$. The variable λ_i are called the lagrangian multipliers associated to the minimize x^*

III. Optimality condition for equality constrained problems.

3.1 OPTIMALITY CONDITION FOR EQUALITY CONSTRAINED PROBLEMS.

Here, we are interested in the problem (P), that is

$$\begin{aligned} &Minimize\ f(x) \\ &subject\ to\ x \epsilon K \neq \emptyset \end{aligned}$$

Where $K = P \cap \{x \in \mathbb{R}^n : g_i(x) = 0\}, P \subset \mathbb{R}^n\ is\ open.$

In this section we are interested in the conditions in which f and the constraint sets will satisfy at an optimum.

3.I.1 The Theorem of Lagrange.

In [4], Huijuan presented the rule for the lagrange multipliers and introduced its application to the field of power systems economic operation. The Theorem of Lagrange provides a powerful characterization of optima of equality-constrained optimization problems in terms of the behavior of the objective function and the constraint functions g at this point

Theorem 3.1 (The Theorem of Lagrange)

Let $f: \mathbb{R}^n \to \mathbb{R}\ and\ g_i: \mathbb{R}^n \to \mathbb{R}^m$ be C^1function, $i = 1, \dots, m.$ Suppose x *is a local maximum or minimum of $f$on the set

$K = P \cap \{x \mid g_i(x) = 0, i = 1, \dots m\},$

Where $P \subset \mathbb{R}^n$ is open. Suppose also that $\nabla g(x^*) = m$
then there exist a vector $\lambda^* = (\lambda_1^*, \lambda_2^*, \ldots \lambda_m^*) \in \mathbb{R}^m$ such that,
$\nabla f(x^*) + \sum_{i=1}^{m} \lambda^* \nabla g_i(x^*) = 0$
where λ_i^* are called the Lagrangian multiplier associated with the local optimum x^*, $\ell(\nabla g(x^*))$ is the rank of the gradient vector $g_i(x)$ that is, the rank of $\nabla g(x^*) = \frac{\partial g_i(x^*)}{\delta x_j}, j = 1, \ldots, n, i = 1, \ldots, m.$
We defined the Lagrangian function $L: K \times \mathbb{R}^n \to \mathbb{R}, \forall x \in \mathbb{R}^n$ as follows:

$$\mathcal{L}(x, \lambda) = f(x) + \sum_{i=1}^{m} \lambda_i g_i(x), \lambda \in \mathbb{R}^m$$

The function L is called the Lagrangian associated with the problem (P). With this function, we get the Lagrangian equations as follows:

$$\frac{\partial L}{\partial x_j}(x, \lambda) = 0, j = 1, \ldots, n,$$

$$\frac{\partial L}{\partial \lambda_i}(x, \lambda) = 0, i = 1, \ldots, m$$

Definition 3.1: Let $x^* \in K$. We say that x^* is a regular point of K if the vectors $\nabla g_i(x^*), \ldots, \nabla g_m(x^*)$ of $\mathbb{R}^m$ are linear independent.
In matrix form, x^* is a regular point of K if the Jacobian matrix of g at x^*,
denoted by $J(g, x^*) = \left[\frac{\partial g_i(x^*)}{\partial x_j}\right]_{j=1,\ldots,n, i=1,\ldots,m}$
has rank , m (i.e, full rank) ,where $g = (g_1, g_2, \cdots g_m)$

Remark 1: If a pair (x^*, λ^*) satisfies the twin conditions that $g(x^*) = 0$ and
$\nabla f(x^*) + \sum_{i=1}^{m} \lambda_i \nabla g_i(x^*) = 0$
we will say that (x^*, λ^*) meets the first order conditions of the Theorem of Lagrange, or (x^*, λ^*) meets the first order necessary conditions in equality – constrained optimization problems.
The Theorem of lagrange only provides necessary conditions for local optima at x^*, and that, only for those local optima x^* which also meet the condition that $\ell(\nabla g(x^*)) = m$. These conditions are not asserted to be sufficient: that is, the Theorem does not mean that if there exist (x^*, λ^*) such that $g(x^*) = 0$,
and $\nabla f(x^*) + \sum_{i=1}^{m} \lambda_i^* \nabla g_i(x^*) = 0$
then x^* must either be a local maximum or a local minimum, even if x also meets the rank condition $\ell(\nabla g(x^*)) = m$

Example 3. Let f and g be functions on $\mathbb{R}^2$defined by $f(x_1, x_2) = x_1^2 + x_2^2$ *and* $g(x_1, x_2) = x_1 - x_2.$
Consider the equality. Constrained optimization problem of minimizing $f(x_1, x_2)$ over the set
$K = \{(x_1, x_2) \in \mathbb{R}^2 | : g(x_1, x_2) = 0\}$
Let (x_1^*, x_2^*)be the point $(0,0)$ and let$\lambda^* = 0$.
Then $g(x_1^*, x_2^*) = 0$
Therefore, (x_1^*, x_2^*) is a feasible point.
Moreover,$\nabla g(x_1, x_2) = (1, -1)$for any $(x_1, x_2) \in \mathbb{R}^2$.
Here $\ell(g\nabla(x_1^*, x_2^*)) = m = 1$,, that is full rank.
$\nabla f(x_1, x_2) = (2x_1, 2x_2),$
Substituting in

$$\nabla f(x^*) + \sum_{i=1}^{m} \lambda_i^* \nabla g_i(x^*)$$

We have
$(0,0) + (1,1) = (0,0)$
This implies that it satisfies the condition of the theorem of Lagrange, but it is not sufficient.

3.2 Constraint qualification: The condition in the theorem of Lagrange that the rank of $\nabla f(x^*)$ be equal to the number of constraints is called the constraint qualification under equality constraints. It ensures that $\nabla g(x^*)$ contains an invertible $m \times m$ sub-matrix, which may be used to define the vector λ^*.

However , if the constraint qualification is violated, then the conclusions of the theorem may also fail. That is, if $x^*$is a local optimum at which $\ell\left(\nabla g(x^*)\right) < m$, then there need not exist a vector λ^* such that

The set of all critical point of L contains the set of all local maxima and minima of f on K at which the constraint qualification is met

$\nabla f(x^*) + \sum_{i=1}^{m} \lambda_i^* \nabla g_i(x^*)$

A particular implications of this property is the following proposition

Proposition 3.4: Suppose the following two conditions hold:

(1) A global minimizer x^*exists to the given problem

(2)The constraint qualification is met at x^*. Then there exists λ^* such that (x^*, λ^*) is a critical point of L .

Example 3.4: Consider the problem

Minimize $f(x) = x_1^2 + 3x_2^2$

$subject \quad x_1^2 + x_2^2 = 1$

$g(x) = x_1^2 + x_2^2 - 1$ and the constraint equation reduces to $g(x) = 0$ and $K = P \cap \{x \in \mathbb{R}^2 | g(x) = 0\}$.

Let $P = \mathbb{R}^2$, then $K = \{x \in \mathbb{R}^2 : x_1^2 + 3x_2^2 = 1\}$

Applying Lagrangian multiplier method we construct the Lagrangian function,

$L(x, \lambda) = f(x) + \lambda_1 g_1(x), \forall\ (x, \lambda) \in K \times \mathbb{R}$

The vector $\lambda = \lambda_1 \in \mathbb{R}$ is called the vector of Lagrangian Multipliers,

$L(x, \lambda) = x_1^2 + 3x_2^2 + \lambda(x_1^2 + x_2^2 - 1)$

Now we first show that the two conditions of proposition 3.4 (namely; existence of global minimizer, and the constraint qualification) are met, so the critical points of the Lagrangian function L will, contain the set of global minima, f is a continuous function on K and K is compact thus by the existence theorem, there exists a minimizer of f in K.

We check for the constraint qualification. The derivative of the constraint function g at any point $x_1, x_2 \in \mathbb{R}$ is given by $(x_1, x_2) \in \mathbb{R}^2$

Given by $\nabla g(x_1, x_2) = ((2x_1, 2x_2),$

and $\lambda \nabla g(x) = 0 \Rightarrow 2\lambda x_1 = 0$ or $2\lambda x_2 = 0$

Since x_1 and x_2 cannot be zero simultaneously on K otherwise, $x_1^2 + x_2^2 = 1$would fail. It implies that we must have $\ell(\nabla g(x_1, x_2)) = 1$ at $(x_1, x_2) \in K$.

Therefore, the constraint qualification holds everywhere on K. Since, $\ell(\nabla g(x_1, x_2)) = m = 1,$,

then there exist $\lambda \in \mathbb{R}$ such that for x^* a minimizer, we have, $\nabla f(x^*) + \lambda \nabla g(x^*) = 0 \left(g(x^*)\right) = 0$

The Lagrangian equation reduces to:

$$\frac{\partial L}{\partial x_1}(x_1, x_2, \lambda) = 2x_1 + 2\lambda x_2 = 0 \qquad (1)$$

$$\frac{\partial L}{\partial x_2}(x_1, x_2, \lambda) = 6x_2 + 2\lambda x_2 = 0 \qquad (2)$$

$$\frac{\partial L}{\partial x_i}(x_1, x_2, \lambda) = x_1^2 + x_2^2 = 1 \qquad (3)$$

$2x_1(1 + \lambda) = 0$

$2x_2(3 + \lambda) = 0$

$x_1^2 + x_2^2 = 1$

The critical points of L are the solutions $(x_1^*, x_2^*, \lambda^*) \in \mathbb{R}^3$

Solving equations above we obtain candidates for minimizer as

$$(x_1^*, x_2^*, \lambda^*) = \begin{cases} (0,1,-3) \\ (0,-1,-3) \\ (1,0,-1) \\ (-1,0,-1) \end{cases}$$

Evaluating f at these four points

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

$f(0,1) = f(0,-1) = 3$

$f(1,0) = f(-1,0) = 1$

Hence the points $(0,1)$ and $(-1,0)$ are the global minimizer of f on K while the points $(0,1)$ and $(0-1)$ are the global maximers of f on K .

3.3 LAGRANGE SECOND ORDER OPTIMALITY CONDITIONS FOR EQUALITY CONSTRAINED PROBLEM.

Now, we look into a situations where the candidates for minimizers are so many that it involves too many computations. To avoid such computations, we give the Lagrange second order conditions for equality constrained problem. Note that second order conditions for equality constrained problem pertain only to local maxima and minima just like the case of unconstrained problems, this is because of the fact that differentiability is only a local property.

Theorem 3.5: Let $f: K \subset \mathbb{R}^n \to \mathbb{R}$ and $g: \mathbb{R}^n \to \mathbb{R}^k$ be C^2 functions.
Suppose there exists point $x^* \in K$ and $\lambda \in \mathbb{R}^k$ such that $\ell\,(\nabla g(x^*)) = k$
and $\nabla f(x^*) + \sum_{i=1}^{m} \lambda_i^* \nabla g_i(x^*) = 0$.
Define $Z(x^*) = \{z \in \mathbb{R}^n | \nabla g(x^*)\, z = 0\}$
and let $\nabla^2 L^*((x^* \lambda^*) = \nabla^2 f(x^*) + \sum_{i=1}^{k} \lambda_i^* \nabla^2 g_i(x^*)$
denote the $n \times n$ symmetric matrix of $L((x^* \lambda^*)$

(1) If f has a local minimum on K at x^* then $z^t \nabla^2 L^* z \geq 0$ for all $z \in Z(x^*)$.
(2) If f has a local maximum on f at x^* then $z^t \nabla^2 L^* z \leq 0$ for all $z \in Z(x^*)$

Applying this theorem to example 3.4,
We have that,

$$\nabla^2 L^*(x, \lambda) = \begin{pmatrix} 2+2\lambda & 0 \\ 0 & 6+2\lambda \end{pmatrix}$$

Substituting λ^* into $\nabla^2 L^*(x, \lambda)$
For $\lambda = -1$, $\nabla^2 L^*(x^*, -1) = \begin{bmatrix} 0 & 0 \\ 0 & 4 \end{bmatrix}$
For $\lambda = -3, \nabla^2 L^*(x, -3) = \begin{bmatrix} -4 & 0 \\ 0 & 0 \end{bmatrix}$
To obtain
$Z(x^*) = \{z \in \mathbb{R}^n | \nabla g(x^*)\, z = 0\}$
$=\{(z_1, z_2) \in \mathbb{R}^2 | 2x_1 z_1 + 2x_2 z_2\} = 0$
Let $z_2 = kx_1, k \in \mathbb{R} \Rightarrow z_1 = -kx_2 \Rightarrow Z(x^*) = \{k(-x_2, x_1) | k \in \mathbb{R}\}$
We see from the example that points $(0,1)$; $\lambda = -3$ and $(0,-1)$; $\lambda = -3$ gives $z^t \nabla^2 L^*(x^*, \lambda^*) z < 0$ and hence they are local maximizers while points $(1,0)$; and $(,-1,0)$; $\lambda = -1$ gives $z^t \nabla^2 L^*(x^*, \lambda^*)\, z > 0$ hence are local minimizers of f on K.

Observe in theorem 3.5 that the definiteness of a symmetric $n \times n$ matrix say Q can be completely characterized in terms of the sub-matrices of Q. Because of this observation we turn to related problem: the characterization of the definiteness of A on the set $\{z \neq 0 | pz = 0\}$ where P is a $k \times n$ matrix of rank k, because of this characterization we state alternative way to check for the conditions of Theorem 3.5. This theorem uses what is called **bordered Hessian matrix denoted by H^B.**
The condition is stated as follows:
Define;

$$H^B = \begin{bmatrix} O & P \\ P^T & Q \end{bmatrix}_{(k+n)\times(k+n)}$$

Where $P = \begin{bmatrix} \nabla g_1(x) \\ \vdots \\ \nabla g_k \end{bmatrix}_{k \times n}$ and $Q = \left[\frac{\partial^2 L(x, \lambda)}{\partial x_i \partial x_j}\right]_{nxn} \forall\, i, j.$

The matrix H^B is called the Bodered Hessian matrix.
Once the critical points (x^*, λ^*) are computed for the Lagrangean function $L(x^*, \lambda^*)$ and Borded Hessian matrix H^B evaluated at (x^*, λ^*), then x^* is

(1) A maximum point if, starting with the principal major of order $2k+1$,, the last $n-k$, principal minor of H^B form an alternating sign pattern with $(-1)^{k+1}$
(2) A minimum point if, starting with principal minor determinant of order $2k+1$, the last $n-k$ principal minor determinants of H^B have the sign of $(-1)^k$

Example 3.5 consider the problem
Minimize $f(x) = x_1^2 + x_2^2 + x_3^2$

Subject to
$x_1 + x_2 + 3x_3 = 2$
$5x_1 + 2x_2 + x_3 = 5$
We construct the lagrangain function
$L(x,\lambda) = x_1^2 + x_2^2 + x_3^2 + \lambda_1(x_1 + x_2 + 3x_3 - 2) + \lambda_2(5x_1 + 2x_2 + x_3 - 5)$
The critical points will be obtained thus;

$\frac{\partial L}{\partial x_1} = 2x_1 + \lambda_1 + 5\lambda_2 = 0$ (4)

$\frac{\partial L}{\partial x_2} = 2x_2 + \lambda_1 + 2\lambda_2 = 0$ (5)

$\frac{\partial L}{\partial x_3} = 2x_3 + 3\lambda_1 + \lambda_2 = 0$ (6)

$\frac{\partial L}{\partial \lambda_1} = x_1 + x_2 + 3x_3 = 2$ (7)

$\frac{\partial L}{\partial \lambda_2} = 5x_1 + 2x_2 + x_3 = -5$ (8)

Solving these equation , we have
$x^* = (x_1, x_2, x_3) = (0.8043, 0.3478, 0.2826)$
$\lambda^* = (\lambda_1, \lambda_2,) = (0.0870, 0.3043)$
We construct the H^B

$$H^B = \begin{bmatrix} O & P \\ P^T & Q \end{bmatrix}$$

$$P = \begin{bmatrix} \nabla g_1(x) \\ \nabla g_2(x) \end{bmatrix} = \begin{bmatrix} \frac{\partial g_1(x)}{\partial x_1} & \frac{\partial g_1(x)}{\partial x_2} & \frac{\partial g_1(x)}{\partial x_3} \\ \frac{\partial g_2(x)}{\partial x_1} & \frac{\partial g_2(x)}{\partial x_2} & \frac{\partial g_2(x)}{\partial x_3} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 3 \\ 5 & 2 & 1 \end{bmatrix}_{2\times 3}$$

$$P^T = \begin{bmatrix} 1 & 5 \\ 1 & 2 \\ 3 & 1 \end{bmatrix}_{3\times 2}$$

$$Q = \frac{\partial^2 L(x,\lambda)}{\partial x_i \partial x_j} = \begin{bmatrix} \frac{\partial^2 L}{\partial x_1^2} & \frac{\partial^2 L}{\partial x_1 \partial x_2} & \frac{\partial^2 L}{\partial x_1 \partial x_3} \\ \frac{\partial^2 L}{\partial x_2 \partial x_1} & \frac{\partial^2 L}{\partial x_2^2} & \frac{\partial^2 L}{\partial x_2 \partial x_3} \\ \frac{\partial^2 L}{\partial x_3 \partial x_1} & \frac{\partial^2 L}{\partial x_3 \partial x_2} & \frac{\partial^2 L}{\partial x_3^2} \end{bmatrix} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

$$H^B = \begin{bmatrix} 0 & 0 & 1 & 1 & 3 \\ 1 & 1 & 5 & 2 & 1 \\ 1 & 5 & 2 & 0 & 0 \\ 1 & 2 & 0 & 2 & 0 \\ 3 & 1 & 0 & 0 & 2 \end{bmatrix}$$

Since $n = 3$ and $k = 2, n - k = 1$ and we need to check the determinant of H^B only.
We must have the sign of $(-1)^k$ for the critical point x^* to be a minimum.
Determinant of $H^B = (-1)^2 460 > 0$. Therefore x^* is a minimum point since $H^B > 0$

IV. CONCLUSION

Herein, an attempt has being made in discussing the Lagrange first and second optimality conditions. Methods for checking optimality conditions for equality constrained problems has being examined and this paper have showed that for complexity computation, Hessian Borded Matrix method becomes a better approach for checking optimal point.

RFERENCES

[1] Rangarajan, K. Sundaram, A first course in optimization theory, Cambridge, United States; University Press, (1998).
[2] Shuonan Dong Methods for Constrained Optimization, 18.086 Project 2 Spring, 2006
[3] Jean P.Vert, Nonlinear optimization (c), Insead, Spring, 2006
[4] Huijuan Li, Lagrange Multipliers and their Applications *University of Tennessee, Knoxville,TN 37921 USA*, 2008