An EffectiveTechnique to Optimization

Daw Phyu Win Nwe ,Faculty of Computing, University of Information Technology, Yangon, Myanmar

Email: phyuwinnwe.pyay@gmail.com

Daw Ohnmar Kyu ,Department of Engineering Mathematics, Technological University (Hinthata),Hinthata, Myanmar

Email: ohmarkyu2019@gmail.com

Daw Aye Aye Thinn,Faculty of Computing, University of Information Technology, Yangon, Myanmar

Email: thinnmyatnoe76@gmail.com

Abstract

The purpose of this research is to find optimal solution of non-affine C^1 function by means of constructing respective Lagrange multiplier function together with Fritz John and Karush-Kuhn- Tucker conditions. These conditions are relating to problems P and Q. Problem P is the minimization of convex functional value with the constrained set of convex inequalities. Problem Q is the minimization of convex functional value with the constrained set of convex inequalities, linear inequalities and convex equalities. Lagrange multiplier function method is well-known but its manipulation is rather complicated. An effective way to handle Lagrange function is presented, in this paper. It is more convenient and more available than the Simplex method due to G.Dantzig. Most of the functions in this paper is non-affine C^1-functions. Moreover, some illustrative examples are also discussed where necessary.

Keywords

Convex Function, Concave Function, Linear Programming, Convex Programming, Karush-Kuhn-Tucker conditions, Lagrange Function and Fritz-John conditions.

1.Introduction

In the literature of optimization of functional value, there are two mains parts such as linear and non-linear optimization. Linear optimization has been developed as linear programming, since the middle years of 20th century. Meanwhile, non-linear optimization has been developed as convex programming quadratic programming and fractional programming since nearly at the end of 20th century. Many researchers concerning these phenomena have been trying in order to get new theories and new techniques which are available to social and environmental development of human societies. These research papers have been done by I.N.Gass, R.T Rockefeller, Mclindon and their followers. This research paper is one of the fruitful results of those researchers, especially for the research paper is organized as the technical development of programming methods. In the simplex method, pivot element is needed in any column of matrix in our consideration and the remaining elements must be

vanished by row operations. The KKT method is not difficult as simplex method and KKT method gives us optimal solution by method of reducing impossible cases.

2. Non-affine and C^1 Function.

2.1 Non-affine Function

2.1.1 Definition [4]

Let $f: \mathbb{R}^n \longmapsto \mathbb{R}$.

$x \longmapsto f(x)$ is an affine function if there exists nonzero constant c:

$$f(x) = L(x) + c$$

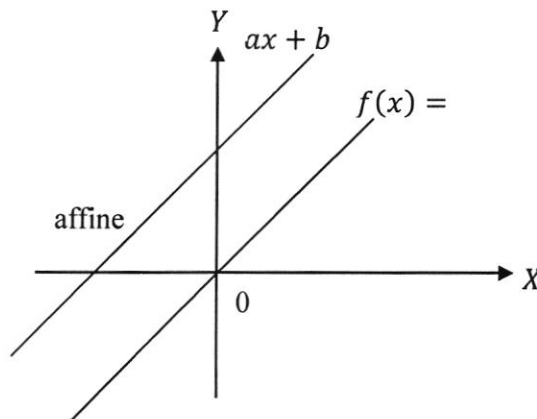where $L: \mathbb{R}^n \longmapsto \mathbb{R}$ is linear.

The function which is not an affine function is called a non-affine function.

Example

Every linear function is non-affine.

Every concave function is non-affine.

Every convex function is non-affine.

2.1.2 Definition [3]

Let $f: \mathbb{R}^n \longmapsto \mathbb{R}$

$$x \longmapsto f(x)$$

We say $f \in C^1$ if and f' are continuous.

Hence, $C^1(\mathbb{R}) = \{f | f: \mathbb{R} \to \mathbb{R}, f \ and \ f' \text{are continuous}\}$ and

$C^\infty(\mathbb{R}) = \{f | f: \mathbb{R} \to \mathbb{R}, f, f', f'', \dots \text{ are continuous}\}$.

Example

$|\bullet|: \mathbb{R} \longmapsto \mathbb{R}$

$$x \longmapsto |x| = \begin{cases} x & , x \geq 0 \\ -x & , x < 0 \end{cases}$$

$$f'(x) = \begin{cases} 1 & ,x > 0 \\ -1 & ,x < 0 \end{cases}$$

$f'(0)$ does not exist. So $f \notin C^1(\mathbb{R})$,

but $f \in C^1(R\backslash\{0\})$.

3. Fritz-John and Karush-Kuhn-Tucker Conditions [3]

In the first part, we study Fritz-John conditions and Karush-Kuhn-Tucker conditions with their proofs for problems P and Q.

3.1   Fritz-John Optimal Conditions for Problems P and Q [3]

3.1.1   Lagrange Functions [3]

Definition

Let $f_i: \mathbb{R}^n \longmapsto \mathbb{R}$ and $h_j: \mathbb{R}^n \longmapsto \mathbb{R}$ be non-affine and C^1- function. $(i = 0,1, ..., q)(j = 1, ..., r)$. The problem P is $min f_0(x): f_i(x) \le 0$ and the problem Q is $min f_0(\text{x}): f_i(x) \le 0, (i = 1, ..., q),...$

Then **Lagrange function for problem P is**

$L: \mathbb{R}^{m+n+q+1} \longmapsto \mathbb{R}$

$$(x, \lambda) \longmapsto L(x, \lambda) = \lambda_0 f_0(x) + \lambda_1 f_1(x) + \cdots + \lambda_q f_q(x) + \sum_{k=1}^{m} v_k(a_k^T x - b)$$

Definition [3]

Let $f: \mathbb{R}^n \longmapsto \mathbb{R}$

Then **Lagrange function for problem Q is**

$L: \mathbb{R}^{n+m+q+r+1} \longmapsto \mathbb{R}$

$(x, \lambda, \mu, \mu^*) \longmapsto L(x, \mu, \mu^*)$

$L(x, \lambda, \mu, \mu^*) = f_0(x) + \sum_{k=1}^{m} \mu_k(a_k^T x - b) + \sum_{i=1}^{q} \mu_i f_i(\text{x}) + \sum_{j=1}^{r} \mu_j^* h_j(x)$

Consider $\min_{x \in c} f(x)$

where $f: \mathbb{R}^n \to \mathbb{R}$

$x \mapsto f(x) \in C^1(R^n)$ and

Q: $\min_{x \in D} f(x)$

where , $C = \mathbb{R}^n \cap \{x | f_i(x) \le 0, a_k^T x - b \le 0\}$, $1 \le i \le q$, $1 \le k \le m$.

$D = C \cap \{x \mid h_j(x) = 0, 1 \le j \le r\}$.

3.1.2   Definition [3][2]

We say that constraint set D as above admits Lagrange multipliers λ_i^* and μ_j^* at a point $x^* \in D$ if for every $f_0 \in C^1(\mathbb{R})$ for which x^* is a local minimizer of a problem, there exist vector $\lambda^* = (\lambda_1^*, ..., \lambda_m^*)$ and $\mu^* = (\mu_1^*, ..., \mu_r^*)$ that satisfy the following conditions.

$\frac{d}{dx}$(Lagrange   function)= $[\nabla f_0(x^*) + \sum_{k=1}^m v_k(a_k^T x - b) + \sum_{j=1}^r \lambda_i^* \nabla h_j(x^*) + \sum_{i=1}^q \mu_j^* \nabla f(x^*)]^T y \geq$ 0,(1)

$\forall y \in T(x^*)$ such that N(x)=T(x^*)

$$\mu_j^* \geq 0, j = 1,2,\dots,r, \qquad (2)$$

$$\mu_j^* = 0, \forall_j \in A(x^*), \qquad (3)$$

where $A(x^*) = \{i | f_i(x^*) = 0\}$, the tangent cone is

$T(x^*) =$

$\left\{ y \in \mathbb{R}^n \middle| \text{either } y = 0 \text{ or a sequence } \{x_k\} \subset \mathbb{R}^n : x_k \neq x \text{ for all } k \text{ and } x_k \rightarrow x, \frac{x_k - x}{\|x_k - x\|} \rightarrow \frac{y}{\|y\|} \right\}$

$N(x) = \{z \in \mathbb{R}'' | \exists \{x_k\} \subset \mathbb{R}^n : x_k \rightarrow x, z_k \rightarrow z, z_k \in T(x_k)^* \}.$

A pair (λ^*, μ^*)atisfying (1), (2) and (3) is called a Lagrange multiplier vector corresponding to f and x^*.

These conditions are called **Fritz-John(FJ) conditions.**

3.1.3   Proposition [3]

Let x^*be a local minimum of the problem (1), (2) and (3). Then there exist $\mu_0^*, \lambda_1^*, \dots, \lambda_m^*; \mu_1^*, \dots, \mu_r^*$ satisfying the following conditions:-

(i) $-(\mu_0^* \nabla f(x^*) + \sum_{i=1}^m \lambda_i^* \nabla h_i(x^*) + \sum_{j=1}^r \mu_j^* \nabla g_j(x^*)) \in N(x^*)$

(ii) $\mu_j^* \geq 0$ for all $j = 0,1,\dots,r.$

(iii) $\mu_0^*, \lambda_1^*, \dots, \lambda_m^*; \mu_1^*, \dots, \mu_r^*$ are not equal to 0.

(iv) If the index set $I \cup J$ is non-empty where

$$I = \{i | \lambda_i^* \neq 0\}, \ J = \{j | j \neq 0, \mu_j^* > 0\},$$

There exists a sequence $\{x^k\} \subset \mathbb{R}^n$ that converges to x^* and is such that for all k,

$$f(x^k) < f(x^*),$$

$$\lambda_i^* h_i(x^k) > 0, \forall i \in I \in, \mu_j^* g_j(x^k) > 0, \quad \forall j \in J \qquad (4)$$

$$|h_i(x^k)| = 0\big(w(x^k)\big), \forall i \notin I, g_j^+(x^k) = 0\big(w(x^k)\big), \forall i \notin J \qquad (5)$$

where $g^+(x) = \max\{0, g_j(x)\}$ and $w(x) = \min\{\min_{i \in j} |h_i(x)|, \min_{i \in j} g_j^+(x)\}$

Proof See [2].

3.2   Karush-Kuhn-Tucker (KKT) Conditions [3]

3.2.1 Definition [3]

Let A be an $m \times n$ matrix with rows $a_k^T, 1 \leq k \leq m, b \in \mathbb{R}^m$ and $f_i: \mathbb{R}^n \mapsto \mathbb{R}, 0 \leq i \leq q$ be non-affine continuously differentiable functions.

We consider the **problem P** is

$$min\{f_0(x)|x \in C\}$$

where

$$C = \{x \in \mathbb{R}^n | a_k^T x \le b_k, 1 \le k \le m, f_i(x) \le 0, 1 \le i \le q\}$$

and the **problem Q** is

$$min\{f_0(x)|x \in D\}$$

where

$$D = C \cap \{x \in \mathbb{R}^n | h_j(x) \le 0, 1 \le j \le r\},$$

$h_j: \mathbb{R}^n \longmapsto \mathbb{R}, 1 \le j \le r$, are non-affine and continuously differentiable.

We denote R_+^l by the non-negative orthant of R^l.

The *FJ* necessary conditions for the problem P are given by the following:-

If x_p is a local minimizer of problem P; then there exist vectors $\lambda \ne 0; \lambda \in \mathbb{R}_+^{q+1}$ and $v \in \mathbb{R}_+^m$ satisfying the condition (FJP)

$$\sum_{i=0}^q \lambda_i \nabla f_i(x_p) + \sum_{k=1}^m v_k a_k = 0,$$

$$\lambda_i f_i(x_p) = 0, 1 \le i \le q \text{ and } v_k(a_k^T x_p - b_k) = 0, 1 \le k \le m.$$

These conditions are called KKT conditions.

For optimization Q, the *FJ* conditions are as follows:-

If x_Q is a local minimizer of problem Q, then there exist vectors

$(\lambda, v) \in \mathbb{R}_+^{q+1+m}, \mu \in \mathbb{R}^r$ with $(\lambda, \mu) \ne 0$ satisfying the conditions (FJQ)

$$\sum_{i=0}^q \lambda_i \nabla f_i(x_Q) + \sum_{j=1}^r \mu_i \nabla h_i(x_Q) + \sum_{k=1}^m v_k(a_k^T x_k - b_k) = 0,$$

$$\lambda_i f_i(x_Q) = 0, 1 \le i \le q \text{ and } v_k(a_k^T x_Q - b_k) = 0, 1 \le k \le m.$$

If λ_0 given in conditions *FJP* and *FJQ* can be chosen positive, then the resulting conditions are called the **Karush-Kuhn-Tucker (KKT) conditions for problem P and Q**, respectively.

A sufficient condition for λ_0 to be positive is given by a so-called first-order constraint qualification.

3.2.2 Definition[3]

For *FJ* and KKT conditions for problems P and Q let $\delta > 0$ and $\bar{x} \in \mathbb{R}^n, N(\bar{x}, \delta)$ denote a δ-neighbourhood of $\bar{x}$ given by

$$N(\bar{x}, \delta) = \{x \in \mathbb{R}^n | \ ||x - \bar{x}|| \le \delta\}.$$

A vector x_p **is called a local minimizer of P** (respectively Q) if $x_p \in C$ (respectively, $x_p \in D$) and there $\delta > 0$ such that $f_0(x_p) \le f_0(x)$ for every $x \in C \cap N(x_p, \delta)$ $\left(\text{respectively}, x \in D \cap N(x_p, \delta)\right)$.

We introduce the index sets $I(x) = \{i|f_i(x) = 0, 1 \le i \le q\}$,

$K(x) = \{k|a_k^T x_Q - b_k, 0, 1 \le k \le m\}$ and

$B(x) = $ the matrix consisting rows $a_k^T, k \in K(x)$.

We say that the **Mangasarian-Fromovitz(MF) constraint qualification** for P holds at a point x if there exists some d_0 satisfying

$B(x)d_0 \leq 0$ and $\max_{i \in I(x)}\{\nabla f_i(x_p)^T d_0\} < 0.$

MF1. $\nabla h_j(x)$, $1 \leq j \leq r$ are linearly indepent.

MF2. $\lin\{\nabla h_j(x), 1 \leq j \leq r\} \cap \lin\{a_k, k \in K(x)\} = \{0\}$

MF3. There exists some d_0 satisfying

$$B(x)d_0 \leq 0, \nabla h_j(x)^T d_0 = 0, 0 \leq j \leq r, \quad \text{and} \quad \max_{i \in I(x)}\{\nabla f_i(x)^T d_0\} < 0.$$

3.2.3 Proposition [3]

Let x_p be a local minimizer of problem (P). Then FJ conditions for P holds.

3.2.4 Proposition [2]

Let x_p be a local minimizer of problem (P). Then KKT conditions for P holds.

Their proofs will be mentioned later after some lemmas are described.

3.2.5 Lemma [3]

If x_p if a local minimizer of problem (P), then $\max \{\nabla f_i(x_p)^T d : i \in I(x_p) \cup \{0\}\} \geq 0$ for every d such that $B(x_p)d \leq 0$.

3.2.6 Lemma [3]

Let $\Delta_s \subseteq \mathbb{R}_+^s$ be the unit simplex. If B is a p × n matrix and $c_i \in \mathbb{R}^n, 1 \leq i \leq s$, some given vectors, then following conditions are equivalent;

1. For every d $\in \mathbb{R}^n$ satisfying $Bd \leq 0$ it holds that $\max_{1 \leq i \leq s} c_i^T d \geq 0$.

2. There exists some $\lambda \in \Delta_s$ and $\mu \in \mathbb{R}_+^p$ satisfying $\sum_{i=1}^s \lambda_i c_i + B^T \mu = 0$.

3.2.7   Proof of FJ conditions for problem (P)[3]

It is followed by combining Lemma 3.2.5 and 3.2.6.

3.2.8   Proof of KKT conditions for Problem (P)

Proof See [3].

3.2.9   Lemma [3]

For any sequence $\varepsilon_l \downarrow 0$ follows that $\lim_{l \uparrow \infty} x_Q(\varepsilon_l) = x_Q$.

Proof : See[3].

3.2.10   Proposition [3]

Let x_Q be a local minimizer of problem Q. Then FJ conditions for Q holds.

3.2.11   Proposition [3]

Let x_Q be a local minimizer of problem Q. Then KKT consider for Q holds. Their proofs will be separately shown in the following;

3.2.12   Proof of FJ conditions for problem (Q)[3]

3.2.13   Proof of K.K.T conditions for problem Q[3]

4. Some Practical Problems

In the second part, we compute some practical problems by means of KKT conditions.

4.1 Application of Lagrange Multiplier Functions

4.2 Application to Linear Fractional Programming Problem

Consider $\max\left\{\frac{3x+2}{2x+1}\middle| x \geq 0\right\}$.

It is equivalent to the following linear programming problem.

$\max\left\{\frac{3y+2t}{2y+t}\middle| y \geq 0\right\}$ by letting $x = \frac{y}{t}$.

At first we will solve the problem

$\min\{-3y - 2t | 2y + t - 1 \leq 0, 1 - 2y - t \leq 0, -y \leq 0\}$.

Here, we have respective Lagrange multiplier function is

$L(y, t, \lambda_1, \lambda_2, \lambda_3) = (-3y - 2t) + \lambda_1(2y + t - 1) + \lambda_2(1 - 2y - t) + \lambda_3(-y)$

By KKT method,

$-3 + 2\lambda_1 - 2\lambda_2 - \lambda_3 = 0$ ..... (i)

$-2 + \lambda_1 - \lambda_2 \quad\quad = 0$ ..... (ii)

$\lambda_1 = 0$ or $2y + t - 1 = 0$ ..... (iii)

$\lambda_2 = 0$ or $1 - 2y - t = 0$ ..... (iv)

$\lambda_3 = 0$ or $\quad\quad y = 0$ ..... (v)

There are 8 cases for λ_1, λ_2 and λ_3.

By considering all the cases, we finally obtain optimal solution is $x = \frac{y}{t} = 0$ and optimal value is 2.

4.3 Application to finding amount of chemicals in Pharmaceutical Industry

In a pharmaceutical industry, amount of chemical used in drugs is important. Effectiveness of a drug depends on its ingredients with chemicals. The effectiveness of amounts x and y of chemical c_1 and c_2 is measured by a function $E(x, y)$. For example c_1 and c_2 be chemicals with costs a dollars per kilogram for c_1 and b dollars per kilogram for c_2 respectively. Then $E(x, y) = ax + by = c$. Suppose that $E(x, y) = xy + 2x$. Then it may give $2x + y = 30$. We will find maximum value of λ_2. Now, the respective Lagrange function is

$L(x, y, \lambda) = (-xy - 2x) + \lambda_1(2x + y - 30) + \lambda_2(30 - 2x - y) + \lambda_3(-x) + \lambda_4(-y)$

By KKT method, $-y - 2 + 2\lambda_1 - 2\lambda_2 - \lambda_3 = 0$

$\quad\quad\quad -x + \lambda_1 - \lambda_2 - \lambda_4 = 0$

$\quad\quad\quad \lambda_1 = 0$ or $2x + y = 30$

$\quad\quad\quad \lambda_2 = 0$ or $2x + y = 30$

$\quad\quad\quad \lambda_3 = 0$ or $\quad x = 0 \quad$ and

$\quad\quad\quad \lambda_4 = 0$ or $\quad y = 0$

Then there are 16 cases to be computed. By considering all cases, impossible cases are rejected. Then we obtain optimal solution by the 8th case: $\lambda_1 > 0, \lambda_2 = \lambda_3 = \lambda_4 = 0, x = 8$ and $y = 14$ give optimal value. So, minimum of $E(x, y)$ is -128 dollar and hence maximum of $E(x, y)$ is 128 dollar. The pharmaceutical industry managers buy 8 kilogram of chemical c_1 and 14 kilogram of chemical c_2. So, maximum effectiveness is 128 dollar.

4. Conclusions

We have presented an efficient algorithm in order to compute optimization problems in social mathematics by using respective Lagrange function some practical problems have been solved by KKT algorithm. KKT algorithm will help us in easy computing optimization problems instead of simplex method. I think this paper is very efficient to Engineering Mathematics. Moreover, KKT method can be applied to the problem how to build darns and reservoirs with specific measurements that minimizing condition of flood and sanding a certain without of water where farmers and cultivators needed. As well as, KKT method is applicable for solving problem where we erect radio telescope tower to minimize interference of magnetic field of the planet in our galaxies.

References

[1] A. Antoniou and Wu-Sheng Lu, Pratical Optimization Algorithm and Engineering Applications, Springer Verlag, New York,2007.

[2] D. P. Bertsekas, A Nedic and A. E. Ozdaglar, *Convex Analysis and Optimization*, Massachusetts Institute of Technology, Massachusetts, 2003.

[3] S.I. Birbil, J. B. G., Frenk and G.J.Still, An elementary Proof of the Fritz-John and Karush-Kuhn-Tucker Conditions in nonlinear programing ,European Journal of Operational Research, p.479-484, Vol. No.180,2006.

[4] R. T. Rockafeller. *Convex Analysis*. Princeton University Press, 1972.

[5]George B.Thomas, Calculus, 12th Edition, Early Transcendentals, Addison-Weseley, Boston, USA, 2006.