The Avriel-Ben-Tal algebraic operations approach for a short version proof of the Karush-Kuhn-Tucker optimality conditions

Veronica CORNACIU and Ileana IOANA

Abstract

In this paper, by using (h, φ)-generalized directional derivative and (h, φ)-generalized gradient, the authors directly derives the Karush-Kuhn-Tucker conditions by applying a corollary of Farkas lemma under the Mangasarian-Fromovitz constraint qualification.Furthermore, the boundedness of Lagrange multipliers is showed.

1 Introduction

In mathematical programming involving differentiable functions, the Karush-Kuhn-Tucker conditions provide necessary conditions for an optimum, under certain qualifications on the constraints.

Many authors studied optimality conditions for vector optimization problems involving constraints that are defined by single-valued mappings and obtained optimality conditions in terms of Lagrange-Kuhn-Tucker multipliers.

The Karush-Kuhn-Tucker conditions have been also used to derive many significant results, for example, in economics, in decision problems that occur in static situations, to show the existence of an equilibrium for a competitive economy, to carry out the first-order approach to principal-agent problems,

and to examine the need for land reform [5],[9],[15]. The Karush-Kuhn-Tucker
conditions and/or the method of Lagrange multipliers appear also in standard
microeconomics textbooks. Mas-Colell, Whinston and Green [8], gives some
Karush-Kuhn-Tucker optimality conditions for an optimization problem with
both inequality and equality constraints. One of the most used method to
determine the Karush-Kuhn-Tucker conditions for the optimization problem
with inequality and equality constraints incorporates the method of Lagrange
multipliers (introduced by Lagrange in 1788); therefore, the simple derivation
of the Karush-Kuhn-Tucker conditions would shed light on the problems true
nature.

The problem of optimization is intensely debatet by Preda [11]-[14], in his
works, generalizing different types of covexities to obtain necessary conditions
for optimization problems.

The study of the optimality conditions also appear in [18], where are intro-
duced five kinds of cones, which are used to establish the constraints qualifica-
tions, under which the generalized Karush-Kuhn-Tucker necessary conditions
are developed for a class of generalized differentiable single-objective and mul-
tiojective programming problems also.

The Karush-Kuhn-Tucker conditions under the Mangasarian-Fromovitz con-
straint qualification are also obtained in [16].

The main aim of this paper is to follow the line of [16] to give some opti-
mality necessary conditions for optimization problems but in the the frame of
the pseudo-Avriel-Ben-Tal algebraic operations.

The paper is organized as follows. Section 2 contains preliminaries and
related results that will be used to obtain the main result of the paper.

In Section 3, the Karush-Kuhn-Tucker optimality conditions under the
Mangasarian-Fromovitz constraint qualification are derived directly by apply-
ing a corollary of Farkas lemma without resorting to the Fritz John conditions.
Also the boundedness of Lagrange multipliers is shown.

Throughout the paper, we denote by R the set of real numbers and
denote by R^n the collection of n-dimensional real vectors.
Ben-Tal [3] introduced certain generalized operations of addition and multi-
plication.

1. Let h be an n vector-valued continuous function defined on R^n, with the
   inverse function h^{-1}. Define the h-vector addition of $x \in R^n$ and $y \in R^n$
   as

$$x \oplus y = h^{-1}(h(x) + h(y)),$$

   and the h-scalar multiplication of $x \in R^n$ and $\lambda \in R$ as

$$\lambda \otimes x = h^{-1}(\lambda h(x)).$$

2. Let φ be a real-valued continuous functions defined on R, with the inverse function φ^{-1}. Then φ-addition of two numbers, $\alpha \in R$ and $\beta \in R$, is given by

$$\alpha[+]\beta = \varphi^{-1}(\varphi(\alpha) + \varphi(\beta)),$$

and the φ-scalar multiplication of $\alpha \in R$ and $\lambda \in R$ by

$$\lambda[\cdot]\alpha = \varphi^{-1}(\lambda\varphi(\alpha)).$$

3. The (h, φ)-inner product of vectors $x, y \in R^n$ is defined as

$$(x^T y)_{h,\varphi} = \varphi^{-1}(h(x)^T h(y)).$$

We put, for $x^i \in R^n, i = 1, 2, \ldots, m$,

$$\overset{m}{\underset{i=1}{\oplus}} x^i = x^1 \oplus x^2 \oplus \ldots \oplus x^m$$

and for $\alpha_i \in R, i = 1, 2, \ldots, m$,

$$[\textstyle\sum_{i=1}^m] \alpha_i = \alpha_1[+]\alpha_2[+]\ldots[+]\alpha_m$$

For $x, y \in R^n$ and $\alpha, \beta \in R$ the operations of subtraction are

$$x\Theta y = x \oplus ((-1) \otimes y), \alpha[-]\beta = \alpha[+]((-1)[\cdot]\beta).$$

Using Ben-Tal generalized algebraic operation, it is easy to obtain the following relations:

$$\overset{m}{\underset{i=1}{\oplus}} x^i = h^{-1}\left(\textstyle\sum_{i=1}^m h(x^i)\right)$$

$$[\textstyle\sum_{i=1}^m] \alpha_i = \varphi^{-1}\left(\textstyle\sum_{i=1}^m \varphi(\alpha_i)\right)$$

$$x\Theta y = h^{-1}(h(x) - h(y))$$

$$\alpha[-]\beta = \varphi^{-1}(\varphi(\alpha) - \varphi(\beta))$$

$$\varphi(\lambda[\cdot]\alpha) = \lambda\varphi(\alpha)$$

$$h(\lambda \otimes x) = \lambda h(x)$$

Denote: $\hat{f}(t) = \varphi(f(h^{-1}(t)))$. For simplicity, write $\hat{f}(t) = \varphi f h^{-1}(t)$.

Definition 2.1 [17] A real valued function $f : R^n \to R$ is said to be (h, φ) - Lipschitz at $x \in R^n$ if there exists to positive constants ϵ, k such that $|f(z)[-]f(y)|_{h,\varphi} \le k[\cdot] \parallel z \Theta y \parallel_{(h,\varphi)}, \forall z, y \in B_{\epsilon,(h,\varphi)}(x)$. f is said to be (h, φ) - locally Lipschitz on R^n if f is (h, φ) - Lipschitz at every $x \in R^n$.

Let f be a Lipschitz and real-valued function defined on R^n. For all $x, v \in R^n$, the (h, φ) - generalized directional derivative of f with respect to direction v and the (h, φ) - generalized gradient of f at x, denoted by $f^*(x, v)$ and $\partial^* f(x)$, respectively, are defined as follows [4].

$$f^*(x, v) = \lim_{y \to x, t \to 0} sup \frac{1}{t}[\cdot](f(y \oplus t \otimes v))[-]f(y)),$$

$$\partial^* f^*(x) = \left\{ \xi^* \in R^n; f^*(x, v) \ge (\xi^{*T} v)_{(h,\varphi)}, \forall v \in R^n \right\},$$

The relation between (h, φ)-generalized directional derivative and Clarke directional derivative can be given by the following theorem.

Theorem 2.1 [4]. Let f be a real valued function, $\varphi(t)$ be strictly increasing and continuous on R. Then $f^*(x, v) = \varphi^{-1}(\hat{f}^\circ(h(x), h(v)))$, where f° is Clarke directional derivative.

Theorem 2.2 [4]. Let f be a real valued function, $\varphi(t)$ be strictly increasing and continuous on R. Then

$$\partial^*(x) = h^{-1}(\partial \hat{f}(h(x))) = \left\{ h^{-1}(\xi); \xi \in \partial(\hat{f}(t)|_{t=h(x)}) \right\}.$$

Definition 2.2[1] Let f be a real-valued function defined on R^n. The function f is said to be (h, φ)-differentiable at x, if $\hat{f}(t)$ is differentiable at $t = h(x)$. Denote $\nabla^* f(x) = h^{-1}(\nabla f(t)|_{t=h(x)})$.

In addition, f is (h, φ)-differentiable on R^n, if is (h, φ)-differentiable in any $x \in R^n$.

Remark 2.1[4] Every differentiable function at x is (h, φ)-differentiable at x, where $h(t) = t, t \in R^n$ and $\varphi(\alpha) = \alpha, \alpha \in R$.

We consider the following optimization problem:
$(P) \max f(x)$
    s.t. $g_i(x) \geq 0, i \in I$
        $h_j(x) = 0, j \in J$

where, $f : R^n \to R$, $g_i : R^n \to R, i \in I, h_i : R^n \to R, j \in J$ are continuously and (h, φ)-differentiable functions, and $|I| = m, |J| = l(m, l \in Z_+)$. If there are not inequality (equality) constraints, we assume that $m = 0(l = 0)$.

We define $I(\bar{x}) \equiv \{i | g_i(\bar{x}) = 0, i \in I\}$.

Now we give the Mangasarian-Fromovitz constraint qualification $(MF)_{h,\varphi}$ (as a generalization of the Mangasarian and Fromovitz constraint [6]) in association with (P) and with our operators:

$(MF)_{h,\varphi}$: For $\bar{x} \in R^n, \nabla^* h_j(\bar{x}), j \in J$, are linearly independent, and there exist $d \in R^n$ s.t.

$$(\nabla^* g_i(\bar{x}), d)_{h,\varphi} > 0, i \in I(\bar{x}) \text{ and } (\nabla^* h_j(\bar{x}), d)_{h,\varphi} = 0, j \in J.$$

Remark: The linearly independent constraint qualification, which is usually assumed in practice, implies $(MF)_{h,\varphi}$.

The following lemma is a generalization of the Corollar 2 to Theorem 2.4.5 from [2] :

Lemma 2.1 For $A \in R^{n \times m}, B \in R^{n \times 1}$, and $c \in R^n$, either

$$(a) \exists y \geq 0, \exists z \in R^1, c + Ay + Bz = 0$$
or
$$(b) \exists x \in R^n, c^T x > 0 \text{ and } A^T x \geq 0, B^T x = 0$$
but never both.

In the rest of the paper, we further assume $h : R^n \to R^n$ is a continuous one-to-one and onto function. Similarly, suppose $\varphi : R \to R$ is a continuous one-to-one strictly monotone and onto function.

2   Result

We now establish the main result, which includes the boundedness of Lagrange multipliers under $(MF)_{h,\varphi}$.

Theorem 3.1.: If $\bar{x} \in R^n$ is a local solution for (P), and the constraint qualification$(MF)_{h,\varphi}$ holds at $\bar{x}$, then, there exist $\bar{\lambda}_i \geq 0, i \in I$ and there exist $\bar{\mu}_j \in R, j \in J$, such that

$$\nabla^* f(\bar{x}) \oplus \bigoplus_{i \in I} \bar{\lambda}_i \otimes \nabla^* g_i(\bar{x}) \oplus \bigoplus_{j \in J} \bar{\mu}_j \otimes \nabla^* h_j(\bar{x}) = 0, \qquad (1)$$

$$\bar{\lambda} \otimes g_i(\bar{x}) = 0, \bar{\lambda}_i \geq 0, i \in I,$$

and furthermore $\bar{\lambda}, \bar{\mu}$ are bounded.

Proof: Let $\bar{x} \in R^n$, x^k in the feasible region such that $x^k \Theta \bar{x} = t_k \otimes s \oplus o(t_k)$. Without loss of generality, we suppose φ is strictly monotone increasing on R. How g_i is (h, φ)-differentiable in $\bar{x}$, so $\hat{g}_i(t) = \varphi g_i h^{-1}(t)$ is differentiable in $\bar{t} = h(\bar{x})$ we get

$$\hat{g}_i(\bar{t} + t_k h(s)) = \hat{g}_i(\bar{t}) + t_k h(s)^T \nabla \hat{g}_i(\bar{t}) + o(t_k), \text{ where } o(t_k) \to 0,$$

when $t_k \to 0$
so, we obtain

$$\varphi g_i h^{-1}(h(\bar{x}) + t_k h(s)) = \varphi g_i h^{-1}(h(\bar{x})) + t_k h(s)^T h(\nabla^* g_i(\bar{x})) + o(t_k).$$

Therefore

$$\varphi g_i(\bar{x} \oplus t_k \otimes s) = \varphi g_i(\bar{x}) + t_k \varphi((\nabla^* g_i(\bar{x}), s)_{h,\varphi}) + o(t_k).$$

So

$$\varphi g_i(\bar{x} \oplus t_k \otimes s) - \varphi g_i(\bar{x}) = t_k \varphi((\nabla^* g_I(\bar{x}), s)_{h,\varphi}) + o(t_k).$$

By division with t_k, we have

$$\frac{\varphi g_i(\bar{x} \oplus t_k \otimes s) - \varphi g_i(\bar{x})}{t_k} = \varphi((\nabla^* g_i(\bar{x}), s)_{h,\varphi}) + \frac{o(t_k)}{t_k}$$

and together with monotony of φ we obtain

$$\frac{\varphi g_i(x^k) - \varphi g_i(\bar{x})}{t_k} = \varphi((\nabla^* g_i(\bar{x}), s)_{h,\varphi}) + \frac{o(t_k)}{t_k} \geq 0, \text{ when } t_k \to 0, i \in I(\bar{x}).$$

which show that $s \in R^n$ satisfies $(\nabla^* g_i(\bar{x}), s)_{h,\varphi} \geq 0, i \in I(\bar{x})$.

Similarly, we obtain

$$\frac{\varphi h_j(x^k) - \varphi h_j(\bar{x})}{t_k} = \varphi((\nabla^* h_j(\bar{x}), s)_{h,\varphi}) + \frac{o(t_k)}{t_k} \geq 0, \text{ when } t_k \to 0, j \in J.$$

which show us that $(\nabla^* h_j(\bar{x}), s)_{h,\varphi} = 0, i \in J$.

Then, for a local solution $\bar{x}$, it follows that

$$(\nabla^* f(\bar{x}), s)_{h,\varphi} > 0 \qquad (2)$$

does not hold, since, if,

$$\frac{\varphi f(x^k) - \varphi f(\bar{x})}{t_k} = \varphi((\nabla^* f(\bar{x}), s)_{h,\varphi}) + \frac{o(t_k)}{t_k} \geq 0, \text{ as } t_k \to 0 \text{ for } x^k \to \bar{x}$$

then the local optimality of f at $\bar{x}$ is contradicted.

Therefore, using the fact that φ,is strictly monotone increasing, the following system:

$$h(s)^T h(\nabla^* f(\bar{x})) > 0$$

$$h(s)^T h(\nabla^* g_i(\bar{x})) \geq 0 \text{ for } i \in I(\bar{x})$$

$$h(s)^T h(\nabla^* h_j(\bar{x})) = 0 \text{ for } j \in J$$

is inconsistent.

So, because $h : R^n \to R^n$ is a one-to-one and onto function, there does not exist a z satisfying

$$z^T h(\nabla^* f(\bar{x})) > 0$$

$$z^T h(\nabla^* g_i(\bar{x})) \geq 0 \text{ for } i \in I(\bar{x})$$

$$z^T h(\nabla^* h_j(\bar{x})) = 0 \text{ for } j \in J$$

We take $c = h(\nabla^* f(\bar{x}))^T$, $A = (h(\nabla^* g_i(\bar{x})))_{i \in I(\bar{x})}^T$, $B = (h(\nabla^* h_j(\bar{x})))_{j \in J}^T$

It follows from the above discussion that the system

$$c^T X > 0, A^T X \geq 0, B^T X = 0$$

is inconsistent.

By Lemma 2.1, there exist $\bar{\lambda}_i \in R_+^{|I(\bar{x})|}, \bar{\mu} \in R^{|J|}$ such that

$$h(\nabla^* f(\bar{x})) + \sum_{i \in I(\bar{x})} \bar{\lambda}_i h(\nabla^* g_i(\bar{x})) + \sum_{j \in J} \bar{\mu}_j h(\nabla^* h_j(\bar{x})) = 0$$

Applying h^{-1} , we obtain

$$h^{-1}(h(\nabla^* f(\bar{x})) + \sum_{i \in I(\bar{x})} \bar{\lambda}_i h(\nabla^* g_i(\bar{x})) + \sum_{j \in J} \bar{\mu}_j h(\nabla^* h_j(\bar{x}))) = h^{-1}(0)$$

which leads to

$$\nabla^* f(\bar{x}) \oplus \oplus_{i \in I(\bar{x})} \bar{\lambda}_i \otimes \nabla^* g_i(\bar{x}) \oplus \oplus_{j \in J} \bar{\mu}_j \otimes \nabla^* h_j(\bar{x}) = 0, \quad (3)$$

$$\bar{\lambda}_i \otimes g_i(\bar{x}) = 0, \bar{\lambda} \geq 0, i \in I(\bar{x})$$

or, equivalently,

$$\nabla^* f(\bar{x}) \oplus \oplus_{i \in I} \bar{\lambda}_i \otimes \nabla^* g_i(\bar{x}) \oplus \oplus_{j \in J} \bar{\mu}_j \otimes \nabla^* h_j(\bar{x}) = 0,$$

$$\bar{\lambda}_i \otimes g_i(\bar{x}) = 0, \bar{\lambda}_i \geq 0, i \in I$$

for $\bar{\lambda}_i = 0, i \in \ I - I(\bar{x})$.

$(MF)_{(h,\varphi)}$ guarantees the existence of such s, $s \neq 0$, so then exists $\ d \in R^n$ from the implicit function theorem[6], if $|I(\bar{x})| + l \geq 1$; otherwise (2) does not hold for $s \neq 0, s \in R^n$.

From (3) we obtain that exists an $\ d \in R^n$ which satisfy $(MF)_{h,\varphi}$ and

$$(\nabla^* f(\bar{x}) \oplus \oplus_{i \in I(\bar{x})} \bar{\lambda}_i \otimes \nabla^* g_i(\bar{x}), d)_{h,\varphi} = 0,$$

So,

$$(\nabla^* f(\bar{x}), d)_{h,\varphi} [+] [\sum_{i \in I(\bar{x})}] \bar{\lambda}_i [\cdot] (\nabla^* g_i(\bar{x}), d)_{h,\varphi} = 0$$

So if, $|I(\bar{x})| \neq 0$ and if function φ it is only strictly monotone increasing on R, we have:

$$-\frac{\varphi((\nabla^* f(\bar{x}), d)_{h,\varphi})}{|I(\bar{x})| \cdot \varphi(\min(\nabla^* g_i(\bar{x}), d)_{h,\varphi})} \geq \bar{\lambda}_i \geq 0,$$

and if $|I(\bar{x})| = 0$, then $\bar{\lambda}$ vanishes.

Therefore, (3) reduced to (a bounded vector) $\oplus \oplus_{j \in J} \bar{\mu}_j \otimes \nabla^* h_j(\bar{x}) = 0$.

From calification constraints $(MF)_{h,\varphi}$, we have that $\nabla^* h_j(\bar{x})$, $j \in J$ are linearly independent, and so $\bar{\mu}$ is determined to a single bounded vector.

3   Conclusions

In this paper, using the algebraic operations introduced by Ben-Tal, the Karush-Kuhn-Tucker optimality conditions under the Mangasarian-Fromovitz constraint qualification were derived directly by applying a corollary of Farkass lemma, without involving the Fritz John conditions, or without introducing the tangent cones. The boundedness of Lagrange multipliers was also shown.

References

[1] Avriel, M. , Nonlinear Programming: Analysis and Methods, Prentice Hall, Englewood Cliffs, NJ, 1976.

[2] Bazaraa, M.S., Shetty, C.M., Nonlinear Programming: Theory and Algorithms. 3nd Edition, Wiley, New York, 2001.

[3] Ben-Tal A., On generalized means and generalized convex functions, J.Optim.Theory Appl., 21:1-13, 1977.

[4] Dehui,Y, Altannar,C, Xiaoling, L, Pardalos, P,M, Generalized convexities and generalized gradients based on algebraic operations, J. Math. Anal. Appl., 321(2):675-690, ,2006.

[5] Grossman, H.I. Production, Appropriation, and Land Reform. American Economic Review, 84, 705-712, 1994.

[6] Mangasarian, O.L., Fromovitz, S., The Fritz John Necessary Optimality Conditions in the Presence of Equality and Inequality Constraints, Journal of Mathematical,1967.

[7] Mangasarian, O.L., Nonlinear Programming, SIAM, Philadelphia, 1994.

[8] Mas-Colell, A., Whinston, M.D., Green, J.R., Microeconomic Theory, Oxford University Press, New York. Analysis and Applications, 17, 37-47, 1995.

[9] Negishi, T., Welfare Economics and Existence of an Equilibrium for a Competitive Economy, Metroeconomica, 12, 92-97, 1960.

[10] Nocedal, J., Wright, S.J., Numerical Optimization. 2nd Edition, Springer, Berlin, 2006.

[11] Preda, V., On Efficiency and Duality for Multi-objective Programs, Journal of Mathematical Analysis and Applications, 166(8), 365-377, 1992.

[12] Preda, V., Some optimality conditions for multiobjective programming problems with set functions, Revue Roumaine de Mathmatiques Pures et Appliques 39 (3), 233-248, 1994.

[13] Preda, V., Duality for multiobjective fractional programming problems involving n-set functions, Analysis and Topology, 569-583, 1998.

[14] Preda, V., Stanciu, D.E., New suficient conditions for B-preinvex and some extinsion, Proceedings of The Roumanian Academy, 12(3),197-202, 2011.

[15] Rogerson, W.P., The First-Order Approach to Principal-Agent Problems, Econometrica, 53, 1357-1367, 1985.

[16] Tanaka, Y, A short derivation of the Kuhn-Tucker conditions, Open Journal of Optimization, 4(2), 47-50,2015.

[17] Zang, Q.X. On sufficiency and duality of solutions for nonsmooth (h, φ)-semi-infinite programming, Acta Math.Appl. Sinica 24, 129138, 2001.

[18] Xu, Y., Liu, S., Kuhn-Tucker Necessary Conditions for (h, φ)-multiobjective optimization problems, Journal of Systems Science and Complexity, 2004.

Veronica CORNACIU,
PhD Student, Doctoral School of Mathematics,
University of Bucharest
Academiei nr.14, Bucharest, 010014, Romania
Assistant Lecturer, *Titu Maiorescu* University
Faculty of Computer Science
Vacaresti nr.187, Bucharest,004051,Romania
E-mail: veronica_zanfir@yahoo.com

Ileana IOANA,
PhD Student, Doctoral School of Mathematics,
University of Bucharest,Academiei nr.14,
Bucharest,010014,Romania