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 \((h, \varphi)\)-generalized directional derivative and \((h, \varphi)\)-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,

Key Words: Generalized Algebraic Operations, Generalized Gradient, Farkas Lemma, Karush-Kuhn-Tucker Conditions, Lagrange Multipliers.

2010 Mathematics Subject Classification: Primary 49K35

Received: 27.05.2016

Revised: 20.06.2016

Accepted: 29.06.2016
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 debated by Preda [11]-[14], in his works, generalizing different types of coevesities to obtain necessary conditions for optimization problems.

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

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

The main aim of this paper is to follow the line of [16] to give some optimality 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 applying 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.


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

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

and the \( \varphi \)-scalar multiplication of \( \alpha \in R \) and \( \lambda \in R \) by

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

3. The \((h, \varphi)\)-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 \( x^i \in R^n, i = 1, 2, \ldots, m \),

\[ \oplus_{i=1}^m x^i = x^1 \oplus x^2 \oplus \ldots \oplus x^m \]

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

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

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:

\[ \oplus_{i=1}^m x^i = h^{-1}(\sum_{i=1}^m h(x^i)) \]

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

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

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

\[ \varphi(\lambda [\cdot ] \alpha) = \lambda \varphi(\alpha) \]
\[ h(\lambda \otimes x) = \lambda h(x) \]

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

**Definition 2.1** [17] A real valued function \( f : \mathbb{R}^n \rightarrow \mathbb{R} \) is said to be \((h, \varphi)\)-Lipschitz at \( x \in \mathbb{R}^n \) if there exists to positive constants \( \epsilon, k \) such that

\[ |f(z) - f(y)|_{h, \varphi} \leq k \| z - y \|_{(h, \varphi)}, \forall z, y \in B_{\epsilon, (h, \varphi)}(x). \]

\( f \) is said to be \((h, \varphi)\)-locally Lipschitz on \( \mathbb{R}^n \) if \( f \) is \((h, \varphi)\)-Lipschitz at every \( x \in \mathbb{R}^n \).

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

\[ f^*(x, v) = \lim_{y \to x, t \to 0} \sup_t \frac{1}{t} \left( f(y + t \otimes v) - f(y) \right), \]

\[ \partial^* f(x) = \{ \xi^* \in \mathbb{R}^n ; f^*(x, v) \geq (\xi^* T v)_{(h, \varphi)}, \forall v \in \mathbb{R}^n \}. \]

The relation between \((h, \varphi)\)-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 \( \mathbb{R} \). Then \( f^*(x, v) = \varphi^{-1}(f^o(h(x), h(v))) \), where \( f^o \) is Clarke directional derivative.

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

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

**Definition 2.2**[1] Let \( f \) be a real-valued function defined on \( \mathbb{R}^n \). The function \( f \) is said to be \((h, \varphi)\)-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 \((h, \varphi)\)-differentiable on \( \mathbb{R}^n \), if is \((h, \varphi)\)-differentiable in any \( x \in \mathbb{R}^n \).

**Remark 2.1**[4] Every differentiable function at \( x \) is \((h, \varphi)\)-differentiable at \( x \), where \( h(t) = t, t \in \mathbb{R}^n \) and \( \varphi(\alpha) = \alpha, \alpha \in \mathbb{R} \).
We consider the following optimization problem:

\((P)\max f(x)\)

\[\begin{align*}
\text{s.t. } & g_i(x) \geq 0, \ i \in I \\
& h_j(x) = 0, \ j \in J
\end{align*}\]

where, \(f : \mathbb{R}^n \to \mathbb{R}\), \(g_i : \mathbb{R}^n \to \mathbb{R}, i \in I\), \(h_j : \mathbb{R}^n \to \mathbb{R}, j \in J\) are continuously and \((h, \varphi)\)-differentiable functions, and \(|I| = m, |J| = l, m, l \in \mathbb{Z}_+\). If there are no 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:

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

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

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

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

**Lemma 2.1** For \(A \in \mathbb{R}^{n \times m}, B \in \mathbb{R}^{n \times 1}, \) and \(c \in \mathbb{R}^n\), either

(a) \(\exists y \geq 0, \exists z \in \mathbb{R}^l, c + Ay + Bz = 0\)

or

(b) \(\exists x \in \mathbb{R}^n, c^T x > 0 \text{ and } A^T x \geq 0, B^T x = 0\)

but never both.

In the rest of the paper, we further assume \(h : \mathbb{R}^n \to \mathbb{R}^n\) is a continuous one-to-one and onto function. Similarly, suppose \(\varphi : \mathbb{R} \to \mathbb{R}\) 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 $\bar{x} \in \mathbb{R}^n$ 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 $\bar{\mu}_j \in \mathbb{R}, j \in J$, such that

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

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

Proof: Let $\bar{x} \in \mathbb{R}^n$, $x^k$ 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 $\varphi$ is strictly monotone increasing on $\mathbb{R}$. How $g_i$ is $(h, \varphi)$-differentiable in $\bar{x}$, so $\hat{g}_i(t) = \varphi g_i h^{-1}(t)$ is differentiable in $t = h(\bar{x})$ we get

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

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

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

By division with $t_k$, we have

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

and together with monotony of $\varphi$ we obtain

$$\frac{\varphi g_i(x^k) \ominus \varphi g_i(\bar{x})}{t_k} = \varphi((\nabla^* g_i(\bar{x}), s)_{h,\varphi}) + \frac{o(t_k)}{t_k} \geq 0,$$

when $t_k \to 0, i \in I(\bar{x})$.

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

Similarly, we obtain

$$\frac{\varphi h_j(x^k) \ominus \varphi h_j(\bar{x})}{t_k} = \varphi((\nabla^* h_j(\bar{x}), s)_{h,\varphi}) + \frac{o(t_k)}{t_k} \geq 0,$$

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
\[
(\nabla^* f(\bar{x}), s)_{h,\varphi} > 0 \tag{2}
\]
does not hold, since, if,
\[
\varphi(f(x^k) - f(\bar{x}))/t_k = \varphi((\nabla^* f(\bar{x}), s)_{h,\varphi}) + \alpha(t_k)/t_k \geq 0,
\]
as \(t_k \to 0\) for \(x^k \to \bar{x}\)
then the local optimality of \(f\) at \(\bar{x}\) is contradicted.

Therefore, using the fact that \(\varphi\) 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 \(h: \mathbb{R}^n \to \mathbb{R}^n\) 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

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

or, equivalently,

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

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 \( d \in \mathbb{R}^n \) from the implicit function theorem\([6]\), if \( |I(\bar{x})| + l \geq 1 \); otherwise \((2)\) does not hold for \( s \neq 0, s \in \mathbb{R}^n \).

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

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

So,

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

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

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

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

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

From calibration 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 Farkas lemma, without involving the Fritz John conditions, or without introducing the tangent cones. The boundedness of Lagrange multipliers was also shown.

References


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