HOKKAIDO UNIVERSITY

| | |
|---|---|
| Title | A Short Derivation of the Kuhn-Tucker Conditions |
| Author(s) | Tanaka, Yoshihiro |
| Citation | Discussion Paper, Series A, 234, 1-6 |
| Issue Date | 2011-02 |
| Doc URL | https://hdl.handle.net/2115/45177 |
| Type | departmental bulletin paper |
| File Information | DPA234_new.pdf |




Hokkaido University Collection of Scholarly and Academic Papers : HUSCAP

Discussion Paper Series $A$, No. 2011-234

# A Short Derivation of the Kuhn-Tucker Conditions

by

Yoshihiro TANAKA
*Associate Professor*

Graduate School of Economics and
Business Administration
Hokkaido University
Kita 9, Nishi 7, Kita-ku
Sapporo 060-0809, Japan

E-mail: tanaka@econ.hokudai.ac.jp

February 2011

# A short derivation of the Kuhn-Tucker conditions

Yoshihiro Tanaka*

Graduate School of Economics and Business Administration, Hokkaido University,
Kita 9 Nishi 7, Kita-ku, Sapporo 060-0809, Japan

Abstract

The Kuhn-Tucker conditions have been used to derive many significant results in economics. However, thus far, their derivation has been a little bit troublesome. The author directly derives the Kuhn-Tucker conditions by applying a corollary of Farkas's lemma under the Mangasarian-Fromovitz constraint qualification.

*JEL classification*: C60, C61

*Keywords*: Farkas's lemma, Kuhn-Tucker conditions, the method of Lagrange multipliers, economics

*E-mail: tanaka@econ.hokudai.ac.jp

# 1 Introduction

The Kuhn-Tucker conditions have been used to derive many significant results in economics, particularly in decision problems that occur in static situations, for instance, to show the existence of an equilibrium for a competitive economy [6], to carry out the first-order approach to principal-agent problems [8], and to examine the need for land reform [3].

The Kuhn-Tucker conditions for the optimization problem with inequality and equality constraints have a comprehensive form that incorporates the method of Lagrange multipliers (introduced by Lagrange in 1788) in a natural way; therefore, the simple derivation of the Kuhn-Tucker conditions would shed light on the problem's true nature.

In this paper, the Kuhn-Tucker conditions under the Mangasarian-Fromovitz constraint qualification are derived directly by applying a corollary of Farkas's lemma without resorting to the Fritz John conditions.

# 2 Farkas's lemma

First, I focus on Farkas's lemma (or the Farkas-Minkowski theorem) with regard to the linear inequality system.

**Proposition 1** *(Farkas's lemma)* *For $A \in \mathbb{R}^{m\times n}$ and $c \in \mathbb{R}^n$, either*

*(a)* $\exists x \in \mathbb{R}^n, \quad c^T x > 0$ *and* $Ax \geq 0$

*or*

*(b)* $\exists u \geq 0, \quad c + A^T u = 0,$

*but never both.* ■

I extend the above result to the linear system including equalities.

**Corollary 1** *For $A \in \mathbb{R}^{m'\times n}$, $B \in \mathbb{R}^{\ell\times n}$, and $c \in \mathbb{R}^n$, either*

*(a)* $\exists x \in \mathbb{R}^n, \quad c^T x > 0$ *and* $Ax \geq 0,\ Bx = 0$

*or*

*(b)* $\exists u \geq 0,\ \exists v \in \mathbb{R}^\ell, \quad c + A^T u + B^T v = 0,$

*but never both.*

*Proof.* By substituting $A := \begin{bmatrix} A \\ B \\ -B \end{bmatrix}$, $u := \begin{bmatrix} u \\ \xi \\ \zeta \end{bmatrix}$, $\xi \geq 0,\ \zeta \geq 0,\ \xi,\ \zeta \in \mathbb{R}^\ell$ in Proposition 1, we obtain the result by putting $v := \xi - \zeta$. ■

# 3 The Kuhn-Tucker conditions

The problem to be addressed is as follows:

$$\begin{array}{lll} (P) \quad \text{maximize} & f(x) & \\ \text{subject to} & g_i(x) \geq 0, & i = 1, ..., m, \\ & h_j(x) = 0, & j = 1, ..., \ell, \end{array}$$

where $f : \mathbb{R}^n \to \mathbb{R}$, $g_i : \mathbb{R}^n \to \mathbb{R},\ i = 1, ..., m$, and $h_j : \mathbb{R}^n \to \mathbb{R},\ i = 1, ..., \ell$ are continuously differentiable functions.

Here, we should pay attention to the fact that the problem (P) naturally includes the optimization problem with equalities considered by Lagrange in the $18^{th}$ century.

I postulate the following Mangasarian-Fromovitz constraint qualification (MF) [5] in association with (P). I define $I(\bar{x}) \equiv \{i \mid g_i(\bar{x}) = 0, i = 1, ..., m\}$.

(MF) For $\bar{x} \in \mathbb{R}^n$, $\nabla h_j(\bar{x}),\ j = 1, ..., \ell$ are linearly independent, and there exists an $\exists d \in \mathbb{R}^n$ s.t. $\nabla g_i(\bar{x})^T d > 0,\ i \in I(\bar{x})$ and $\nabla h_j(\bar{x})^T d = 0,\ j = 1, ..., \ell$.

*Remarks*

- The linearly independent constraint qualification, which is usually assumed in practice, implies (MF) (see [7]).
- (MF) is equal to the Cottle constraint qualification without the presence of equality constraints, and if the problem (P) is a concave program without

equality constraints, the Slater constraint qualification implies the Cottle constraint qualification [1].

Next, I establish the main result.

**Theorem 1** *Suppose that $\bar{x} \in \mathbb{R}^n$ is a local solution for (P), and that the constraint qualification (MF) holds at $\bar{x}$. Then, for $\exists\ \bar{\lambda}_i \geq 0,\ i = 1, ..., m$ and $\exists\ \bar{\mu}_1, ..., \bar{\mu}_\ell \in \mathbb{R}$,*

$$\begin{aligned} \nabla f(\bar{x}) + \sum_{i=1}^{m} \bar{\lambda}_i \nabla g_i(\bar{x}) + \sum_{j=1}^{\ell} \bar{\mu}_j \nabla h_j(\bar{x}) = 0, \\ \bar{\lambda}_i g_i(\bar{x}) = 0, \quad \bar{\lambda}_i \geq 0, \quad i = 1, ..., m. \end{aligned} \tag{1}$$

*Proof.* At a local solution $\bar{x} \in \mathbb{R}^n$, if we choose $x^k$ in the feasible region such that $x^k - \bar{x} = t_k s + o(t_k)$,

$$\begin{aligned} \frac{g_i(x^k) - g_i(\bar{x})}{t_k} &= \frac{1}{t_k}[t_k \nabla g_i(\bar{x})^T s + o(t_k)] \\ &= \nabla g_i(\bar{x})^T s + \frac{o(t_k)}{t_k} \\ &\geq 0 \quad \text{as } t_k \downarrow 0, \quad i \in I(\bar{x}), \\ \frac{h_j(x^k) - h_j(\bar{x})}{t_k} &= \frac{1}{t_k}[t_k \nabla h_j(\bar{x})^T s + o(t_k)] \\ &= \nabla h_j(\bar{x})^T s + \frac{o(t_k)}{t_k} \\ &= 0 \quad \text{as } t_k \downarrow 0, \quad j = 1, ..., \ell, \end{aligned}$$

which shows that $s \in \mathbb{R}^n$ satisfies $\nabla g_i(\bar{x})^T s \geq 0$, $i \in I(\bar{x})$ and $\nabla h_j(\bar{x})^T s = 0$, $j = 1, ..., \ell$.
Then, for a local solution $\bar{x}$, it follows that

$$\nabla f(\bar{x})^T s > 0 \tag{2}$$

does *not* hold, since, if so, $\frac{f(x^k) - f(\bar{x})}{t_k} = \frac{1}{t_k}[t_k \nabla f(\bar{x})^T s + o(t_k)] > 0$ as $t_k \downarrow 0$ for $x^k \to \bar{x}$, which contradicts the local optimality of $f$ at $\bar{x}$.
Note that (MF) guarantees the existence of such $0 \neq s = \exists d \in \mathbb{R}^n$ from the implicit function theorem [[4], Appendix D with $h(x_I, x_{II}) = 0$ and $(x_I, x_{II}) \in \Lambda \subset \mathbb{R}^{n-\ell} \times \mathbb{R}^\ell$]

if $|I(\bar{x})|+\ell \geq 1$; otherwise (2) does not hold for $0 \neq \forall s \in \mathbb{R}^n$. By applying Corollary 1 to (2) and (MF), even if the active constraints are empty, we obtain

$$\nabla f(\bar{x}) + \sum_{i \in I(\bar{x})} \bar{\lambda}_i \nabla g_i(\bar{x}) + \sum_{j=1}^{\ell} \bar{\mu}_j \nabla h_j(\bar{x}) = 0,$$
$$\bar{\lambda}_i g_i(\bar{x}) = 0, \quad \bar{\lambda}_i \geq 0, \quad i \in I(\bar{x}),$$

or, equivalently,

$$\nabla f(\bar{x}) + \sum_{i=1}^{m} \bar{\lambda}_i \nabla g_i(\bar{x}) + \sum_{j=1}^{\ell} \bar{\mu}_j \nabla h_j(\bar{x}) = 0,$$
$$\bar{\lambda}_i g_i(\bar{x}) = 0, \quad \bar{\lambda}_i \geq 0, \quad i = 1, ..., m$$

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

It is important to note that when applying Corollary 1 to (MF), $\nabla g_i(\bar{x})^T d > 0$, $i \in I(\bar{x})$ implies that $\nabla g_i(\bar{x})$, $i \in I(\bar{x})$ are in a convex cone (or pointed), unless $I(\bar{x}) \neq \phi$. Therefore, the Lagrange multipliers $\bar{\lambda}$ and $\bar{\mu}$ are bounded if and only if (MF) is satisfied at $\bar{x}$ (see [2]).

# 4 Concluding remarks

In this paper, the Kuhn-Tucker conditions under the Mangasarian-Fromovitz constraint qualification were derived directly by applying a corollary of Farkas's lemma without resorting to the Fritz John conditions.

Considerable effort has been devoted to the generalization of Farkas's lemma. However, what seems to be lacking is a discrete version of Farkas's lemma under a mild condition; such a version would be theoretically meaningful and would be help solve the discrete optimization problems that emerge in the economics studying indivisible goods.

# References


[1] Bazarra, M.S., Goode, J.J., Shetty, C.M., 1972. "Constraint qualifications revisited," *Management Science* 18, 567–573.

[2] Gauvin, J., 1977. "A necessary and sufficient regularity condition to have bound multipliers in nonconvex programming," *Mathematical Programming* 12, 136–138.

[3] Grossman, H.I., 1994. "Production, appropriation, and land reform," *The American Economic Review* 84, 705–712.

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

[5] Mangasarian, O.L., Fromovitz, S., 1967. "The Fritz John necessary optimality conditions in the presence of equality and inequality constraints," *Journal of Mathematical Analysis and Applications* 17, 37–47.

[6] Negishi, T., 1960. "The existence of an equilibrium for a competitive economy," *Metroeconomica* 12, 92–97.

[7] Nocedal, J., Wright, S.J., 2006. *Numerical Optimization – 2nd Ed.*, Springer.

[8] Rogerson, W.P., 1985. "The first-order approach to principal-agent problems," *Econometrica* 53, 1357–1367.