Bureau
d'économie
théorique
et appliquée
(BETA)
UMR 7522

Documents de travail

« A note on second-order conditions in inequality constrained optimization »

Auteur

Gisèle Umbhauer

Document de Travail n° 2008 - **19**

Juillet 2008

Faculté des sciences économiques et de gestion
Pôle européen de gestion et d'économie (PEGE)
61 avenue de la Forêt Noire
F-67085 Strasbourg Cedex

Secétariat du BETA
Géraldine Manderscheidt
Tél. : (33) 03 90 24 20 69
Fax : (33) 03 90 24 20 70
manderscheidt@cournot.u–strasbg.fr
http://cournot2.u–strasbg.fr/beta

CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE

Nancy-Université
Université Nancy 2

UNIVERSITÉ LOUIS PASTEUR STRASBOURG

A note on second-order conditions in inequality constrained optimization

Gisèle Umbhauer[1][2]

Bureau d'Economie Théorique et Appliquée
Université Louis Pasteur, Strasbourg France

April 2008

Abstract

Let f, h_1………h_p be C^2 functions on $\mathbb{R}^n$. Consider the problem of optimizing f on the constraint set $\{x \in \mathbb{R}^n \ / \ h_j(x) \geq 0$ for j from 1 to p$\}$. We show that if $p \geq n$ and if the number of binding constraints is equal to n, the theorem of Kuhn-Tucker is a sufficient condition theorem regardless of the nature of the functions f and h.

Résumé

Soient f, h_1………h_p des fonctions C^2 définies sur $\mathbb{R}^n$. Considérons le problème qui consiste à optimiser f sur l'ensemble $\{x \in \mathbb{R}^n \ / \ h_j(x) \geq 0$ pour j de 1 à p$\}$. On suppose $p \geq n$ et on s'intéresse au cas où le nombre de contraintes serrées est égal à n. On montre dans ce cas que le théorème de Kuhn Tucker devient un théorème de conditions suffisantes, quelle que soit la nature des fonctions f et h.

JEL classification : C 61

Keywords: optimization; Kuhn Tucker; second-order conditions.

---

[1] Gisèle Umbhauer
e-mail address: umbhauer@cournot.u-strasbg.fr
postal address : BETA,Pôle Européen de Gestion et d'Economie (PEGE),
61 Avenue de la Forêt Noire
67085 Strasbourg Cedex France

[2] The author acknowledges Jacques Thépot for valuable suggestions and comments.

1. Introduction

Let f, h_1........h_p be C^2 functions on $\mathbb{R}^n$. Consider the problem of maximizing f on the constraint set $\{x\in \mathbb{R}^n \ / \ h_j(x) \geq 0$ for j from 1 to p$\}$. Consider a critical point x*, i.e. a point that meets the first-order conditions of the theorem of Kuhn Tucker.

Theorem 1: Theorem of Kuhn Tucker
Let E be the set of binding constraints at x*, and suppose that the rank at x* of the Jacobian matrix of binding constraints, rank $(Dh_E(x^*))$, is equal to |E|. Then, if x* is local maximum, there exists a vector $\lambda^*= (\lambda_1^*,\ldots,\lambda_p^*)$ such that:

$\lambda_j^* \geq 0$ for j from 1 to p (1)

$\lambda_j^* h_j(x^*)= 0$ for j from 1 to p (2)

$$Df(x^*) + \sum_{j=1}^{p} \lambda_j^* Dh_j(x^*) = 0 \qquad (3)$$

To check if x* is a local maximum, one can turn to the second-order conditions:
Consider the bordered Hessian H(x*):

$$H(x^*) = \begin{pmatrix} L''_{x_1x_1}(x^*) & . & L''_{x_nx_1}(x^*) & h_1'{}_{x_1}(x^*) & . & h_{|E|}'{}_{x_1}(x^*) \\ . & . & . & . & . & . \\ L''_{x_1x_n}(x^*) & . & L''_{x_nx_n}(x^*) & h_1'{}_{x_n}(x^*) & . & h_{|E|}'{}_{x_n}(x^*) \\ h_1'{}_{x_1}(x^*) & . & h_1'{}_{x_n}(x^*) & 0 & . & 0 \\ . & . & . & 0 & . & 0 \\ h_{|E|}'{}_{x_1}(x^*) & . & h_{|E|}'{}_{x_n}(x^*) & 0 & . & 0 \end{pmatrix}$$

Theorem 2: Second-order conditions (see for example [1])
For |E| <n, if the last n-|E| leading principal minors of H(x*) alternate in sign, the sign of the determinant of the largest matrix being the same as the sign of $(-1)^n$, then x* is a local maximum.

This theorem does not, at least not explicitly, treat the case |E|=n. As a matter of fact, theorem 2 checks the concavity of the objective function in the neighbourhood of x*, where the objective function is a function of n-|E| variables, the remaining |E| variables being implicit functions of the n-|E| other ones (by use of the implicit function theorem and the fact that |E| constraints are binding). Yet there are many (at least economic) non convex problems such that the number of binding constraints is exactly the number of variables. That is why we propose the extension of the second-order conditions to the case |E|=n.

2. Kuhn Tucker: a sufficient condition theorem in non convex problems

Kuhn-Tucker: a sufficient conditions theorem if |E|=n
When the number of binding constraints is equal to the number of variables, then a point x* that meets the conditions of the theorem of Kuhn Tucker is a local maximum, regardless of the nature of the functions f and h_j, j from 1 to p.

Proof:

It is straightforward that maximizing f on the constraint set $\{x \in \mathbb{R}^n \,/\, h_j(x) \geq 0$ for j from 1 to p$\}$ -(problem 1)- is equivalent to solving problem 2, which includes an additional variable y , $y \in \mathbb{R}$.

Problem 1
Max f(x)
$x \in \mathbb{R}^n$
u.c. $h_j(x) \geq 0$ j from 1 to p

Problem 2
Max $f(x) - y^2$
$x \in \mathbb{R}^n$ $y \in \mathbb{R}$
u.c. $h_j(x) \geq 0$ j from 1 to p

If x* is a local maximum of problem 1, (x*, y*=0) is a local maximum of problem 2, and vice versa. Moreover, if x* in problem 1 meets the conditions of the theorem of Kuhn-Tucker, then (x*, 0) meets the conditions of the theorem of Kuhn and Tucker in problem 2 and vice versa (the Kuhn Tucker multipliers, the binding constraints, the Jacobian matrix of binding constraints being the same in both problems for x* and (x*,0)).
Let us suppose that |E|=n, and, w.l.o.g., that the n binding constraints are the constraints h_j, j from 1 to n. Given that the number of variables in problem 2, (n+1), is higher than the number of binding constraints (n), we can apply the second-order conditions theorem to problem 2.
Given that $\mathcal{L}(x,y) = L(x) - y^2$, where $\mathcal{L}(x,y)$ is the Kuhn Tucker Lagrangian function of problem 2 and L(x) is the Kuhn Tucker Lagrangian function of problem 1, we get:

$$H(x^*,0) = \begin{pmatrix} L''_{x_1x_1}(x^*) & . & L''_{x_nx_1}(x^*) & 0 & h_1'{}_{x_1}(x^*) & . & h_n'{}_{x_1}(x^*) \\ . & . & . & 0 & . & . & . \\ L''_{x_1x_n}(x^*) & . & L''_{x_nx_n}(x^*) & 0 & h_1'{}_{x_n}(x^*) & . & h_n'{}_{x_n}(x^*) \\ 0 & . & 0 & -2 & 0 & . & 0 \\ h_1'{}_{x_1}(x^*) & . & h_1'{}_{x_n}(x^*) & 0 & 0 & . & 0 \\ . & . & . & 0 & 0 & . & 0 \\ h_n'{}_{x_1}(x^*) & . & h_n'{}_{x_n}(x^*) & 0 & 0 & . & 0 \end{pmatrix}$$

The only principal minor to calculate is detH(x*,0).
One gets:

$$\det H(x^*,0) = -2\begin{vmatrix} L''_{x_1x_1}(x^*) & . & L''_{x_nx_1}(x^*) & h_1'_{x_1}(x^*) & . & h_n'_{x_1}(x^*) \\ . & . & . & . & . & . \\ L''_{x_1x_n}(x^*) & . & L''_{x_nx_n}(x^*) & h_1'_{x_n}(x^*) & . & h_n'_{x_n}(x^*) \\ h_1'_{x_1}(x^*) & . & h_1'_{x_n}(x^*) & 0 & . & 0 \\ . & . & . & 0 & . & 0 \\ h_n'_{x_1}(x^*) & . & h_n'_{x_n}(x^*) & 0 & . & 0 \end{vmatrix}$$

Given that rank $(Dh_E(x^*)) = n$, each vector $(L''_{x_1x_j}(x^*);...,L''_{x_nx_j}(x^*))$, j from 1 to n, is a linear combination of the n vectors $(h_j'_{x_1}(x^*),...,h_j'_{x_n}(x^*))$, j from 1 to n. It follows that:

$$\det H(x^*,0) = -2\begin{vmatrix} 0 & . & 0 & h_1'_{x_1}(x^*) & . & h_n'_{x_1}(x^*) \\ . & . & . & . & . & . \\ 0 & . & 0 & h_1'_{x_n}(x^*) & . & h_n'_{x_n}(x^*) \\ h_1'_{x_1}(x^*) & . & h_1'_{x_n}(x^*) & 0 & . & 0 \\ . & . & . & 0 & . & 0 \\ h_n'_{x_1}(x^*) & . & h_n'_{x_n}(x^*) & 0 & . & 0 \end{vmatrix}$$

$$= (-2).\ (-1)^n \begin{vmatrix} h_1'_{x_1}(x^*) & . & h_1'_{x_n}(x^*) \\ . & . & . \\ h_n'_{x_1}(x^*) & . & h_n'_{x_n}(x^*) \end{vmatrix}^2$$

Given that rank$(Dh_E(x^*))=n$, it follows that:

$$(-1)^{n+1} \det H(x^*,0) = 2\begin{vmatrix} h_1'_{x_1}(x^*) & . & h_1'_{x_n}(x^*) \\ . & . & . \\ h_n'_{x_1}(x^*) & . & h_n'_{x_n}(x^*) \end{vmatrix}^2 > 0$$

So $(x^*,0)$ is a local maximum of problem 2, and x^* is a local maximum of problem 1, regardless of the nature of the functions f and h_i, j from 1 to p (provided they are C^2). Therefore, for $|E|= n$, each point satisfying the Kuhn Tucker conditions is a local maximum.

3. Illustration and minimization problems

It is interesting to illustrate by a simple example why, if the number of binding constraints is lower than n, they are additional second-order conditions for a critical point x* satisfying the Kuhn Tucker's conditions, whereas they are no additional conditions when $|E|=n$.

Example:

Max $(x_1-1)^3-x_2$

x_1x_2

u.c. $x_2 \geq 0$

$1 - x_1/2 - x_2 \geq 0$

It is immediate that the objective function f(x) is neither concave nor convex on $\mathbb{R}^2$, and that it is strictly convex (and therefore not concave) on the admissible set.

The Kuhn Tucker conditions lead to two critical points.

The first point is A=$\begin{pmatrix}1\\0\end{pmatrix}$. Only the first constraint is binding; so one has to check if the second order conditions hold, i.e. if $(-1)^2$détH(A) is strictly positive.

$$(-1)^2 \det H(A) = \begin{vmatrix} 6(x_1^*-1) & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{vmatrix} = 0$$ (because x*=1). Hence $(-1)^2$ detH(A) is not strictly positive. In fact, A is not a local maximum, because the objective function is strictly convex at the right of $x_1^*=1$. This strict convexity implies that it is possible to find a vector u that points into the admissible set (the triangle (B,C,D)) such that the value of f(x) grows when x moves into this direction (see figure 1).

The second critical point is B=$\begin{pmatrix}2\\0\end{pmatrix}$; both constraints are binding at B. Hence B is a local maximum because the number of binding constraints is equal to the number of variables. There is no additional condition to check: B is a local maximum despite the strict convexity of the objective function for x_1 higher than 1. One easily observes in figure 1 that the only fact that both constraints are binding at B ensures that the moves that increase the value of the objective function (i.e. the moves toward the dotted hyperplane) are not admissible.

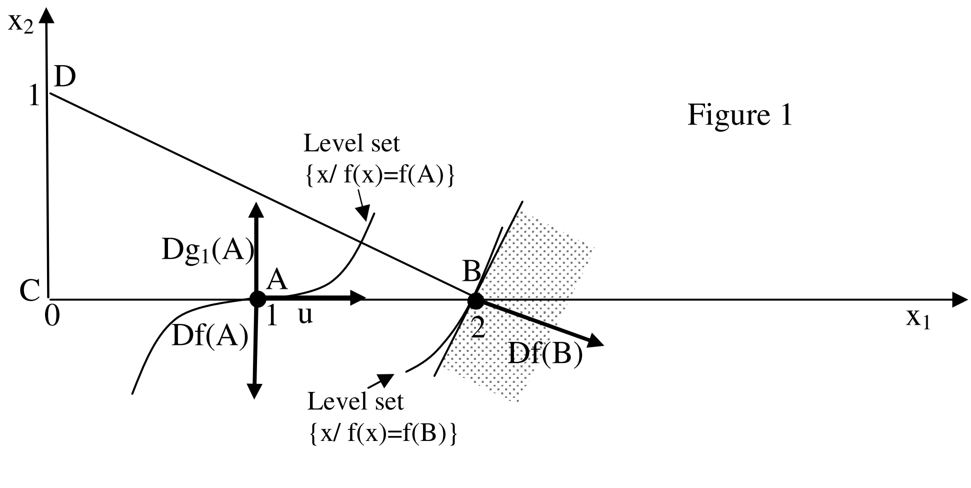

Figure 1

The result established in section 3 also holds for minimization problems.

Corollary: Let f, h_1.........h_p be C^2 functions on $\mathbb{R}^n$. Consider the problem of minimizing f on the constraint set $\{x \in \mathbb{R}^n \ / \ h_j(x) \geq 0$ for j from 1 to p$\}$. Consider a point x*, such that |E|=n, rank $(Dh_E(x^*))$=n, and there exists a vector $\lambda^* = (\lambda_1^*,\ldots,\lambda_p^*)$ such that:

$\lambda_j^* \geq 0$ for j from 1 to p (1)

$\lambda_j^* h_j(x^*) = 0$ for j from 1 to p (2)

and $Df(x^*) - \sum_{j=1}^{p} \lambda_j^* Dh_j(x^*) = 0$ (3)'.

Then x* is a local minimum.

The proof is the same as in section 2, by replacing (-2) by 2, so that:

$$(-1)^n \det H(x^*,0) = 2\begin{vmatrix} h_1'_{x_1}(x^*) & . & h_1'_{x_n}(x^*) \\ . & . & . \\ h_n'_{x_1}(x^*) & . & h_n'_{x_n}(x^*) \end{vmatrix}^2 > 0.$$ Hence x* is a local minimum.

The obtained results are useful for economists who often deal with problems such that the number of binding constraints is equal to the number of variables; it is interesting to know that, in that case, regardless of the nature of the functions f and h_j, j from 1 to p, (provided they are C^2), a point which satisfies the conditions of the theorem of Kuhn Tucker is a local optimum.

Acknowedgements

The author acknowledges Jacques Thépot for valuable suggestions and comments.

References

[1] C. P. Simon and L. Blume, Mathematics for economists, W.W. Norton & Company, New York, London 1994.