B1

ERASMUS UNIVERSITEIT ROTTERDAM

Nederlandsche Economische Hogeschool

ECONOMETRIC INSTITUTE (DEPT. MATH.)

Report 7320 / M ,0

NOTE ON KUHN-TUCKER THEORY AND STABILITY
(second order theory)

Michiel Hazewinkel

Oct. 16, 1973

Preliminary

NOTE ON KUHN-TUCKER THEORY AND STABILITY

(second order theory)

Michiel Hazewinkel

Contents

1. INTRODUCTION, NOTATIONS AND CONVENTIONS

This note is more or less a complement to [1], in which the "first order theory" was dealt with. For motivation we refer to the introduction of [1].

In this note we consider C^2-perturbations of programming problems

$$\text{(MP)} \qquad \begin{array}{l} \text{maximize } f(X),\ x \in \mathbb{R}^n \\ \text{subject to } g_i(x) \leq 0 \quad i = 1,\dots,m \end{array}$$

where f and the g_i are supposed to be twice continuously differential functions.

A problem

$$\text{(MP')} \qquad \begin{array}{l} \text{Maximize } f'(x),\quad x \in \mathbb{R}^n \\ \text{subject to } g_i'(x) \leq 0 \quad i = 1,\dots,m \end{array}$$

is an ε - C^2-perturbation of (MP) if $|f(x) - f'(x)| < \varepsilon$

$|g_i(x) - g_i'(x)| < \varepsilon$ for all $i = 1,\dots,m$, $x \in \mathbb{R}^n$,

$|\frac{\partial g_i}{\partial x_j}(x) - \frac{\partial g_i'}{\partial x_j}(x)| < \varepsilon$, $|\frac{\partial f}{\partial x_j}(x) - \frac{\partial f'}{\partial x_j}(x)| < \varepsilon$

for all $x \in \mathbb{R}^n$; $i = 1,\dots,m$; $j = 1,\dots,n$ and $|\frac{\partial^2 g_i}{\partial x_j \partial x_\ell}(x) - \frac{\partial^2 g_i}{\partial x_j \partial x_\ell}(x)| < \varepsilon$

$|\frac{\partial^2 f}{\partial x_j \partial x_\ell}(x) - \frac{\partial^2 f'}{\partial x_j \partial x_\ell}(x)| < 0$ for all $x \in \mathbb{R}^n$; $i = 1,\dots,m$; $j = 1,\dots,n$;

$l = 1,\dots,n$.

As in [1] is the aim of this note to prove theorems comparing the solutions of (MP') to those of (MP) in certain nice cases and to show that most problems are nice.

The proper setting for this are again mathematical programming problems (D,f) where D is a compact C^2-differential set, and ε-C^2-perturbations of these. The definitions of these things are completely analogous to the definitions given in [1] in the C^1-case and will therefore not be given explicitly here.

In a way C^2-perturbations are the natural ones to consider. This is reflected in two facts. Firstly one need not worry any more about solutions which are not on the boundary of the feasible region as in [1], and secondly one does not need to consider generalized perturbations (cf. [1]).

We use the same notations and conventions as in [1].

2. STABILITY OF NICE PROBLEMS

We consider the mathematical programming problem

$$\text{(MP)} \qquad \begin{array}{l} \text{maximize } f(x),\ x \in \mathbb{R}^n \\ \text{subject to } g_i(x) \leq 0,\ i = 1,2,\ \ldots,m \end{array}$$

where f and the g_i are functions of class C^2; i.e. they are twice continuously differentiable.

(2.1) Definition.

The problem (MP) is called nice if

1^o) (MP) has only finitely many local solutions.
Let $\hat{x}$ be any local solution of (MP) and suppose that

$$g_{i_1}(\hat{x}) = .. = g_{i_k}(\hat{x}) = 0,\ g_l(\hat{x}) < 0 \text{ for } l \in \{1,2,\ldots,m\} \setminus \{i_1,\ldots,i_k\}.$$

Then we further require

2^o) The vectors $\nabla g_{i_j}(\hat{x})$, $j = 1,\ldots,k$ are linearly independant.

3^o) $\nabla f(\hat{x})$ is the relative interior of the cone spanned by the $\nabla g_{i_j}(\hat{x})$, $j = 1,\ldots,k$

4^o) If $<\mu\ ,\ \nabla g_{i_j}(\hat{x})> = 0$ for $j = 1,\ldots,k$ and $\mu \neq 0$ then

$$\sum_{i,j} \frac{\partial^2 H}{\partial x_i \partial x_j} (\hat{x},\hat{\lambda})\mu_i\mu_j < 0$$

where $H(x,\lambda) := f(x) - \sum_{\ell=1}^{m} \lambda_\ell \, g_\ell(x)$ and the $\hat{\lambda}_1,\dots, \hat{\lambda}_m$ are such that

$$\nabla f(\hat{x}) - \sum_{\ell=1}^{m} \hat{\lambda}_\ell \nabla g_\ell(\hat{x}) = 0,\ \hat{\lambda}_\ell g_\ell(\hat{x}) = 0,\ \hat{\lambda}_\ell \geq 0,\ \ell = 1,\dots,m$$

(Such $\hat{\lambda}_\ell$ exist by the necessary first order Kuhn-Tucker conditions in view of 2°) and they are unique by 2°), 3°); Note that $\hat{\lambda}_{i_j} > 0$ for $j = 1,\dots,k$ by 3°).

Condition 4°) is the sufficient second order condition of Kuhn-Tucker theory

(2.2) Definition.

Let (MP) be a programming problem, and $\hat{x}$ a solution of (MP). By renumbering the g_i if necessary we can assume that $g_1(\hat{x}) = .. = g_k(\hat{x}) = 0$, $g_{k+1}(\hat{x}) < 0,\dots, g_m(\hat{x}) < 0$. We now define the function $F : \mathbb{R}^{n+k} \to \mathbb{R}^{n+k}$ as

$$F\begin{pmatrix} x_1 \\ \vdots \\ x_n \\ \lambda_1 \\ \vdots \\ \lambda_k \end{pmatrix} = \begin{pmatrix} \frac{\partial f}{\partial x_1}(x) - \sum_{i=1}^{k} \lambda_i \frac{\partial g_i}{\partial x_1}(x) \\ \vdots \\ \frac{\partial f}{\partial x_n}(x) - \sum_{i=1}^{k} \lambda_i \frac{\partial g_i}{\partial x_n}(x) \\ g_1(x) \\ \vdots \\ g_k(x) \end{pmatrix}$$

(Note that F depends to some extend on $\hat{x}$).

(2.3) Proposition.

Let (MP) be a nice programming problem and $\hat{x}$ a local solution of (MP), let F be the corresponding function as defined in (2.2). Then the Jacobian matrix of F at $(\hat{x},\hat{\lambda})$ is nonsingular.

Proof. The Jacobian matrix $JF(\hat{x},\hat{\lambda})$ of F at $(\hat{x},\hat{\lambda})$ is equal to

$$\begin{pmatrix}
\frac{\partial^2 f}{\partial x_1^2}(\hat{x}) - \sum_{i=1}^{k}\hat{\lambda}_i \frac{\partial^2 g_i}{\partial x_1^2}(\hat{x}) & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_1}(\hat{x}) - \sum_{i=1}^{k}\hat{\lambda}_i \frac{\partial^2 g_i}{\partial x_n \partial x_1}(\hat{x}) & -\frac{\partial g_1}{\partial x_1}(\hat{x}) \cdots -\frac{\partial g_k}{\partial x_1}(\hat{x}) \\
\frac{\partial^2 f}{\partial x_1 \partial x_n}(\hat{x}) - \sum_{i=1}^{k}\hat{\lambda}_i \frac{\partial^2 g_i}{\partial x_1 \partial x_n}(\hat{x}) & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_n}(\hat{x}) - \sum_{i=1}^{k}\hat{\lambda}_i \frac{\partial^2 g_i}{\partial x_n \partial x_n}(\hat{x}) & -\frac{\partial g_1}{\partial x_n}(\hat{x}) \cdots -\frac{\partial g_k}{\partial x_n}(\hat{x}) \\
\frac{\partial g_1}{\partial x_1}(\hat{x}) & \cdots & \frac{\partial g_1}{\partial x_n}(\hat{x}) & 0 \cdots 0 \\
\vdots & & \vdots & \vdots \quad \vdots \\
\frac{\partial g_k}{\partial x_1}(\hat{x}) & \cdots & \frac{\partial g_k}{\partial x_n}(\hat{x}) & 0 \cdots 0
\end{pmatrix}$$

Suppose that e.g. the columns are linearly dependant. Then there are $(\mu_1,\ldots,\mu_n)$ and $(\nu_1,\ldots,\nu_k)$ such that

$$(2.3.1) \quad \sum_{\ell=1}^{n} \mu_\ell\left(\frac{\partial^2 f}{\partial x_\ell \partial x_j}(\hat{x}) - \sum_{i=1}^{k}\hat{\lambda}_i \frac{\partial^2 g_i}{\partial x_\ell \partial x_j}(\hat{x})\right) - \sum_{s=1}^{k} \nu_s \frac{\partial g_s}{\partial x_j}(\hat{x}) = 0 \qquad j = 1,\ldots,n$$

and

$$(2.3.2) \quad \sum_{\ell=1}^{n} \mu_\ell \frac{\partial g_s}{\partial x_\ell}(\hat{x}) = 0 \qquad s = 1,\ldots,k$$

Multiplying (2.3.1) with μ_j and summing over all j we obtain

$$(2.3.3) \quad \sum_{\ell,j=1}^{n}\left(\frac{\partial^2 f}{\partial x_\ell \partial x_j}(\hat{x}) - \sum_{i=1}^{k}\hat{\lambda}_i \frac{\partial^2 g_i}{\partial x_\ell \partial x_j}(\hat{x})\right)\mu_\ell\mu_j - \sum_{j=1}^{n}\sum_{s=1}^{k}\mu_j\nu_s \frac{\partial g_s}{\partial x_j}(\hat{x}) = 0$$

Which implies

$$(2.3.4) \quad \sum_{\ell,j=1}^{n} \frac{\partial^2 H}{\partial x_\ell \partial x_j}(\hat{x},\hat{\lambda})\mu_\ell\mu_j = 0$$

because of (2.3.2) and $\hat{\lambda}_{k+1} = \ldots = \hat{\lambda}_m = 0$ (Cf. (2.1)). But this contradicts condition 4°) of (2.1) in view of (2.3.2), because the vector μ is such that $\langle\mu, \nabla g_i(x)\rangle = 0$ for $i = 1,\ldots,k$. This proves the proposition.

(2.4) Stability theorem.

Let (MP) be a nice programming problem and $\hat{x}$ a local solution of (MP). Then there exist an $\varepsilon > 0$ and a $\delta > 0$ such that every δ-C^2-perturbation of (MP) has precisely one local solution $\hat{x}'$ within distance ε from $\hat{x}$.

Proof. Consider the function F associated to (MP) and $\hat{x}$. (Cf. (2.2)). A δ-C^2-perturbation of (MP) results in an δ-C^1-perturbation F' of F. Therefore using transversality and (2.3) we see that for sufficiently small ε and δ, there is precisely one solution $(\hat{x},\hat{\lambda}')$ of $F'(x,\lambda) = 0$. If δ is small enough we also have that the $\nabla g_i(\hat{x}_i')$, $i = 1,\ldots,k$ are independant and that $\nabla f'(\hat{x}')$ is in the relative interior of the cone spanned by the $\nabla \hat{g}_i(\hat{x}')$, $i = 1,\ldots,k$. Further (if δ is small enough), $g_j(\hat{x}') < 0$ for $j = k+1,\ldots,m$. Finally because we are dealing with a δ-C^2-perturbation the second order Kuhn-Tucker condition (2.1),4°) remains fulfilled for the perturbed problem, which proves that $\hat{x}'$ is a local solution. The solution $\hat{x}'$ is unique because for sufficiently small ε and δ for every local solution x within ε of x there must be a $\tilde{\lambda}$ such that $F(\tilde{x},\tilde{\lambda}) = 0$ (because the $\nabla g_i(\tilde{x})$ are independant for $i = 1,\ldots,k$).

(2.5) Remark.

The local solution x' of (MP') satisfies conditions 2°, 3°, 4° of (2.1)

(2.6) Remark.

Let (D,f) be a mathematical programming problem consisting of a differential set D (cf. [1] for definitions) and a function f defined on D. (We now of course require that the local defining functions of D are C^2 and also that f is C^2). We call such a programming problem nice if

1°) D is compact
2°) (D,f) has precisely one solution $\hat{x}$ for which the obvious analogues of 2°,3°,4° of (2.1) hold.

We now define δ-C^2-perturbations of (D,f) in the natural way (cf. [1] for C^1-perturbations).One then has the following.

(2.7) Theorem.

If (D,f) is a nice programming problem then there is a $\varepsilon > 0$ such that every ε-C^2-perturbation of (D,f) is nice.

3. MOST PROBLEMS ARE NICE

(3.1) Proposition.

Let (MP) be programming problem and $\hat{x}$ a local solution of (MP). Then for every $\varepsilon > 0$ there exists an ε-C^2-perturbation (MP) of (MP) such that

1^o $\hat{x}$ is a local solution of (MP')

2^o The conditions 2^o, 3^o, 4^o hold for the perturbed solved programming problem ((MP'), $\hat{x}$)

Proof. By means of the same arguments as used in [1], (5,1) one finds an $\frac{1}{2}\varepsilon$-C^2-perturbation (MP") of (MP) such that $\hat{x}$ is a unique local solution of (MP") for which conditions 2^o and 3^o of (2.1) hold. To get 4^o as well one adds a function $(\sum_{i,j}\delta\, a_{ij}(x_i-\hat{x}_i)(x_j-\hat{x}_j))\phi(x-\hat{x})$ to f for δ sufficiently small, where (a_{ij}) is a negative definite matrix and $\phi : R^n \to R^1$ is a C^2-function such that $0 \leq \phi(x) \leq 1$ for all x, $\phi(0) = 1$, $\phi(x) = 0$ if $||x|| \geq 1$.

q.e.d.

(3.2) We define $\mathfrak{MP}_2$ as the space of programming problems (D,f) where D is a compact C^2-differential set and f a C^2-function on D, topologized by taking as a basis for the open neighbourhoods of (D,f) the sets of all ε-C^2-perturbations of (D,f). (Note that we do not take generalized ε-C^2-perturbations of (D,f).

3.3) Theorem.

The set of nice programming problems in $\mathfrak{MP}_2$ is open and dense. (Nice is taken in the sense of (2.6).

Proof. Use (2,4), (2,7), (3.1) and the compactness of D.

REFERENCES

[1]. M. Hazewinkel, P. de Weerd (1973). On stability in mathematical programming (first order theory). Report 7306 of the Econometric Institute, Erasmus Univ. Rotterdam.