Giorgio Giorgi
Cesare Zuccotti

Some results and remarks on the envelope theorem

Quaderno di ricerca numero XII
Anno accademico 2008–2009

Copyright © MMIX
ARACNE EDITRICE S.r.l.

www.aracneeditrice.it
info@aracneeditrice.it

00173 Roma
via Raffaele Garofalo, 133 A/B
06 93781065
telefax 06 72678427

ISBN 978-88–548–2596-3

I diritti di traduzione, di memorizzazione elettronica,
di riproduzione e di adattamento anche parziale,
con qualsiasi mezzo, sono riservati per tutti i Paesi.

I edizione: luglio 2009

Finito di stampare nel mese di luglio del 2009
dalla tipografia « Braille Gamma S.r.l. » di Santa Rufina di Cittaducale (RI)
per conto della « Aracne editrice S.r.l. » di Roma
Printed in Italy

INDICE

1.	Introduction	p. 1
2.	The envelope theorem under twice differentiability assumptions	p. 4
3.	The envelope theorem without second order assumptions	p. 12
4.	The concave envelope theorem	p. 17

Abstract

We consider the usual envelope theorem as it appears in many books on microeconomics or mathematical economics. We point out the conditions under which the optimal value function is differentiable. We treat the unconstrained case and the general constrained case, under, respectively, second-order assumptions, no second-order assumptions, and concavity / convexity / linearity assumptions.

1. Introduction*

The *envelope theorem* is a basic result dealing with some optimization problems in microeconomic theory. It may be used, e.g., to prove Hotelling's lemma, Shepard's lemma and Roy's identity (see, e.g., Beavis and Dobbs (1990), Chiang and Wainwright (2005), Samuelson (1947), Silberberg (1971), Silberberg and Suen (2001), Takayama (1977, 1985)). This theorem deals with parametrized optimization problems and is usually stated without the suitable assumptions which guarantee the differentiability of the so-called *optimal value function* or simply assuming the differentiability of the same: see Bertocchi, Stefani and Zambruno (1992), Novshek (1993), Simon and Blume (1994), Varian (1978). This theorem has his own history which can be read in Silberberg (1999) and in Silberberg and Suen (2001).

* Giorgio Giorgi, Professore Ordinario di Matematica generale e di Matematica per l'economia e la finanza (secs-S/06). Università degli Studi di Pavia, Dipartimento di Ricerche Aziendali "Riccardo Argenziano", Via S. Felice, 7 - 27100 Pavia; tel. 0382986238 – e-mail: ggiorgi@eco.unipv.it

Cesare Zuccotti, Ricercatore di Teoria delle decisioni e di Applicazioni elettroniche per la gestione e l'ottimizzazione (secs-S/06). Università degli Studi di Pavia, Dipartimento di Ricerche Aziendali "Riccardo Argenziano", Via S. Felice, 7 - 27100 Pavia; tel. 0382986257 – e-mail: czuccotti@eco.unipv.it

We recall that, especially in the mathematical economics literature, the optimal value function (or marginal function or perturbation function or also indirect function) is defined by $f^*(a) = f(x^*(a), a)$, where x^* is the optimum value of a problem $\min_x$ or $\max_x f(x, a)$, $x \in \mathbb{R}^n$, and $a \in \mathbb{R}^k$ is a vector of parameters. So, $f^*(a)$ is always a function as there can only be one minimal or maximal value of $f(x, a)$ for each choice of $a \in \mathbb{R}^k$ (we suppose that the problem has a solution, otherwise we have to substitute "min" with "inf" and "max" with "sup"). Note, however, that $x^*(a)$ need not be a (vector) function, since the optimal choice of control variables need not be unique for a given $a \in \mathbb{R}^k$.

For example, W. Novshek (1993) gives the following two versions of the envelope theorem.

1.1 Without constraints

Let $\mathbb{X} \subseteq \mathbb{R}^n$ and $\mathbb{A} \subseteq \mathbb{R}^k$; assume that $f : \mathbb{X} \times \mathbb{A} \to \mathbb{R}$ is a continuously differentiable function, and for parameters $a \in \mathbb{R}^k$, $x^*(a) \in \mathbb{R}^n$ is the unique optimizer of $f(x, a)$ and $x^*(a)$ is continuously differentiable. Define the optimal value function

$$f^*(a) = f\left(x^*(a), a\right),$$

then

$$\frac{\partial f^*(a)}{\partial a_i} = \frac{\partial f}{\partial a_i}\left(x^*(a), a\right), \qquad \text{for } i = 1, 2, ..., k. \tag{1}$$

1.2 With equality constraints

Assume that $f : \mathbb{X} \times \mathbb{A} \to \mathbb{R}$ and

$$g_j : \mathbb{X} \times \mathbb{A} \to \mathbb{R}, \quad j = 1, 2, ..., m, (m < n),$$

are continuously differentiable functions and that for the parameters vector $a \in \mathbb{R}^k$ the vectors $x^*(a) \in \mathbb{R}^n$ and $\lambda^*(a) \in \mathbb{R}^m$ are such that $x^*(a)$ is the unique optimizer of $f(x, a)$ subject to $g_j(x, a) = 0$, $j = 1, 2, ..., m$, and such that the Lagrangian first-order conditions hold with multipliers $\lambda_j^*(a)$, $j = 1, 2, ..., m$. Assume that $x^*(a)$, $\lambda^*(a)$

are continuously differentiable; define $f^*(a) = f(x^*(a), a)$. Then, for $i = 1, 2, ..., k$,

$$\begin{aligned} \frac{\partial f^*(a)}{\partial a_i} &= \frac{\partial f}{\partial a_i}\Big(x^*(a), a\Big) + \sum_{j=1}^{m} \lambda_j^*(a) \cdot \frac{\partial g_j}{\partial a_i}\Big(x^*(a), a\Big) = \\ &= \frac{\partial L}{\partial a_i}\Big(x^*(a), \lambda^*(a), a\Big), \end{aligned}$$

where $L = f + \lambda g$ is the Lagrangian function.

For other similar statements the reader may see Afriat (1971), Beavis and Dobbs (1990), Carter (2001), Brinkhuis and Tikhomirov (2005), Sydsæter, Strom and Berck (1999), Takayama (1977, 1985), Varian (1978).

These results could benefit from some classical works on sensitivity and stability for mathematical programming problems, such as, e.g., the book of A. V. Fiacco (1983). See also Giorgi (1998, 2000) for similar remarks on these questions. In the present paper we shall recall some basic results on the envelope theorem, under suitable second-order differentiability assumptions, then we shall obtain the differentiability of the value function $f^*(a)$ without assuming second-order differentiability and finally we shall examine the "concave-convex-linear case", i.e. a concave mathematical programming problem with convex and linear constraints.

First we note that, with reference to the unconstrained case, if $f(x, a)$ is continuous on $\mathbb{X} \times \mathbb{A}$ and $\mathbb{X}$ is compact, then $f^*(a)$ is continuous on $\mathbb{A}$. If $x^*(a)$ is unique, for each $a \in \mathbb{A}$, then $x^*(a)$ is a continuous function of a. Moreover, if $f(x, a)$ is continuous in $\mathbb{X} \times \mathbb{A}$, $\mathbb{X}$ is compact, $x^*(a)$ is unique for $a = a^*$ and $\partial f / \partial a_i$, $i = 1, 2, ..., k$, exist and are continuous in a neighborhood of (x^*, a^*), then $f^*(a)$ is differentiable at a^* and we get (1), i.e.

$$\frac{\partial f^*(a^*)}{\partial a_i} = \frac{\partial f(x^*, a^*)}{\partial a_i}, \qquad \text{for } i = 1, 2, ..., k.$$

This is a corollary of a more general result of Milgrom and Segal (2002). See also Carter (2001) for a direct proof. The proofs of the results of Novshek (1993) (but also of Simon and Blume (1994), Takayama (1985), Castagnoli and Cigola (2006/2007), Bertocchi, Stefani and Zambruno (1992), Guerraggio and Salsa (1997), Mas-Colell, Whinston and Green (1995), Varian (1978)) simply rely on the classical chain rule on the differentiation of composite functions, as $f^*(\bullet)$ is *assumed* (directly or indirectly) to be differentiable. We have to note that Takayama (1977, 1985) remarks that f^* is not necessarily differentiable and that this fact was recognized by Samuelson (1947). As for what concerns the constrained case, especially in the presence of inequality constraints, the economical literature is even more inaccurate; see Giorgi (1998, 2000, 2006).

We have several deep papers on the existence of various *directional derivatives* of the optimal value function (see, e.g., Fiacco and Hutzler (1982), Gauvin and Dubeau (1982), Gauvin and Tolle (1977), Gauvin (1980, 1989), Gauvin and Janin (1990), Gollan (1984), Rockafellar (1984)) but here we shall be interested in the *differentiability* of the optimal value function, in accordance with the statement of the envelope theorem.

2. The envelope theorem under twice differentiability assumptions

We consider the following nonlinear programming problem

$$
(\mathsf{P}) \qquad
\begin{array}{ll}
 & \max\limits_{x \in \mathbb{R}^n} \text{imize} f(x) \\
\text{subject to} & g_i(x) \leq 0, \quad i = 1, 2, ..., m, \\
 & h_j(x) = 0, \quad j = 1, 2, ..., p, \\
\text{with} & f, \{g_i\}, \{h_j\} : \mathbb{R}^n \to \mathbb{R}.
\end{array}
$$

Moreover, it is assumed in the present section, that f, $\{g_i\}$ and $\{h_j\}$ are twice continuously differentiable around $x^* \in \mathbb{K}$ (feasible set of (P)). The Lagrangian associated with (P) is defined as:

$$
L(x, u, w) = f(x) - \sum_{i=1}^{m} u_i g_i(x) - \sum_{j=1}^{p} w_j h_j(x).
$$

We recall three basic constraint qualifications for (P):

a) The Mangasarian-Fromovitz Constraint Qualification (MFCQ)

holds at x^* if

i) the vectors $\bigtriangledown_x h_j(x^*)$, $j = 1, 2, .., p$, are linearly independent;

ii) there exists $z \in \mathbb{R}^n$ such that

$$\begin{aligned} &\bigtriangledown_x g_i(x^*)z < 0, \quad i \in \mathbb{I}(x^*) = \{i \mid g_i(x^*) = 0\}, \\ &\bigtriangledown_x h_j(x^*)z = 0, \quad j = 1, 2, ..., p. \end{aligned}$$

b) The Strict Mangasarian-Fromovitz Constraint Qualification (SMFCQ)

holds at x^* if for its associated multipliers u^*, w^* one has

i) $\bigtriangledown_x g_i(x^*)$, $i \in \mathbb{I}_+(u^*)$, $\bigtriangledown_x h_j(x^*)$, $j = 1, 2, ..., p$, are linearly independent;

ii) there exists $z \in \mathbb{R}^n$ such that

$$\begin{aligned} &\bigtriangledown_x g_i(x^*)z < 0, \quad i \in \mathbb{I}(x^*)\backslash\mathbb{I}_+(u^*), \\ &\bigtriangledown_x g_i(x^*)z = 0, \quad i \in \mathbb{I}_+(u^*), \\ &\bigtriangledown_x h_j(x^*)z = 0, \quad j = 1, 2, ..., p, \end{aligned}$$

where $\mathbb{I}_+(u^*) = \{i \in \mathbb{I}(x^*) : u_i^* > 0\}$.

Kyparisis (1985) has proved that (SMFCQ) is a necessary and sufficient conditions to have unique multipliers u^*, w^* for (P).

c) The Linear Independence condition (LI)

holds at x^* if the vectors

$$\begin{aligned} &\bigtriangledown_x g_i(x^*), \quad i \in \mathbb{I}(x^*), \\ &\bigtriangledown_x h_j(x^*), \quad j = 1, 2, ..., p, \end{aligned}$$

are linearly independent.

We recall the standard second-order sufficient conditions for a strict local maximum in problem (P). See, e.g., Fiacco and McCormick (1968), Fiacco (1983), McCormick (1983) and the paper of Giorgi and Zuccotti (2008) for extensions and remarks.

Theorem 2.1

Suppose that the Kuhn-Tucker conditions hold at x^* for (P) with some multiplier vectors u^* and w^* i.e.

$$\begin{aligned}
&\nabla_x L(x^*, u^*, w^*) = 0, \\
&u_i^* g_i(x^*) = 0, \quad i = 1, 2, ..., m, \\
&g_i(x^*) \leq 0, \quad i = 1, 2, ..., m, \\
&h_j(x^*) = 0, \quad j = 1, 2, ..., p, \\
&u_i^* \geq 0, \quad i = 1, 2, ..., m.
\end{aligned}$$

Suppose that the following additional second-order sufficient conditions (SOSC) hold at x^* with (u^*, w^*) :

$$z^T \nabla_x^2 L(x^*, u^*, w^*) z < 0, \quad \text{for all } z \neq 0,$$

such that

$$\begin{aligned}
&\nabla_x g_i(x^*) z \leq 0, \quad \text{for all } i \in \mathbb{I}(x^*), \\
&\nabla_x g_i(x^*) z = 0, \quad \text{for all } i \text{ such that } u_i^* > 0, \\
&\nabla_x h_j(x^*) z = 0, \quad j = 1, 2, ..., p.
\end{aligned}$$

Then x^* is a strict local maximum of (P).

We consider now the following general parametric nonlinear programming problem

$$(\mathsf{P}(a)) \qquad \begin{aligned}
&\max_{x \in \mathbb{R}^n} \text{imize} f(x, a) \\
\text{subject to} \quad &g_i(x, a) \leq 0, \quad i = 1, 2, ..., m, \\
&h_j(x, a) = 0, \quad j = 1, 2, ..., p,
\end{aligned}$$

where $a \in \mathbb{R}^k$ is the perturbation parameter,

$$f, \ g_i, \ h_j : \mathbb{R}^n \times \mathbb{R}^k \to \mathbb{R},$$

and f, $\{g_i\}$ and $\{h_j\}$ are $\mathbb{C}^2$ in (x, a) in some neighborhood of (x^*, a^*). The Lagrangian function associated to $(\mathsf{P}(a))$ is

$$L(x, u, w, a) = f(x, a) - \sum_{i=1}^{m} u_i g_i(x, a) - \sum_{j=1}^{p} h_j(x, a).$$

The following basic result was proved by Fiacco (1983). See also Giorgi (2000), Shapiro (1985).

Theorem 2.2

Suppose that the second-order sufficient conditions for a strict local maximum of $(\mathsf{P}(a^*))$ hold at x^* with associated Lagrange multiplier vectors u^* and w^*, that the (LI) condition holds at x^* for $(\mathsf{P}(a^*))$, and that the Strict Complementary Slackness condition (SCS) holds at x^* with respect to u^* for $(\mathsf{P}(a^*))$, i.e.

$$u_i^* > 0 \text{ when } g_i(x^*, a^*) = 0, \quad i = 1, 2, ..., m.$$

Then:

a) x^* is an isolated local maximum of $(\mathsf{P}(a^*))$ and the associated multiplier vectors u^* and w^* are unique.

b) For a in a neighborhood of a^*, there exists a unique (once) continuously differentiable vector function

$$y(a) = [x(a), u(a), w(a)]^T$$

satisfying the second-order sufficient conditions for a local maximum of $(\mathsf{P}(a))$ such that $y(a^*) = (x^*, u^*, w^*)^T$ and, hence $x(a)$ is a locally unique local maximum of $(\mathsf{P}(a))$, with associated unique multiplier vectors $u(a)$ and $w(a)$.

c) The (LI) condition and the (SCS) condition hold at $x(a)$ for a near a^*.

Fiacco (1983) also shows how to calculate the gradient $\nabla_a y(a)$.

We recall the attention on the (SCS) assumption, usually skipped in the various versions of sensitivity theorems or envelope theorems appearing in the mathematical economics literature.

See Carter (2001), Castagnoli and Peccati (1979), Intriligator (1971) and see Giorgi (1998, 2000, 2006) for similar considerations on these versions.

In differential stability results for $(\mathsf{P}(a))$ (see, e.g., the surveys in Fiacco (1983), Fiacco and Hutzler (1982), Gauvin and Janin (1990)) one usually employs the following standard definition of the optimal value function f^* :

$$f^*(a) = \begin{cases} \sup\limits_x \{f(x,a) \mid x \in \mathbb{K}(a) \neq \varnothing\}, \\ -\infty, \quad \text{if } \mathbb{K}(a) = \varnothing, \end{cases}$$

where

$$\mathbb{K}(a) = \left\{ \begin{array}{ll} x \mid & g_i(x,a) \leq 0, \quad i = 1,2,..,m, \\ & h_j(x,a) = 0, \quad j = 1,2,...,p \end{array} \right\},$$

and usually assumes some kind of compactness of the set $\mathbb{K}(a)$. However, when we deal with results having a local character, as in the present section, usually a "local" optimal value function f_l^* is considered, function defined as

$$f_l^*(a) = f\left(x(a), a\right),$$

where $x(a)$ is an isolated local maximum of $(\mathsf{P}(a))$. The following result is proved in Fiacco (1983).