# Econ 508-A

## The Envelope Theorem

Carmen Astorne-Figari
*Washington University in St. Louis*

August 16, 2010

# INTRODUCTION

Remember problems written in parametric form:

$$\max\{f(\boldsymbol{x}, \boldsymbol{\theta}) \mid \boldsymbol{x} \in X(\boldsymbol{\theta})\}$$

We want to know how the value of the objective function changes when parameters change.

Example: when solving the consumer's utility maximization problem, we would like to know how utility changes in response to a change in prices and/or income.

Other examples of applications of the envelope theorem:

- Cost function: cost minimization with parameters factor prices and output.
- Profit function: profit maximizaton with parameters input/output prices.

# STATEMENT

Given:

$$\begin{aligned} \max_{\boldsymbol{x}} \quad & f(\boldsymbol{x}, \boldsymbol{\theta}) \\ \text{s.t.} \quad & g_1(\boldsymbol{x}, \boldsymbol{\theta}) \leq 0 \\ & \vdots \\ & g_K(\boldsymbol{x}, \boldsymbol{\theta}) \leq 0 \end{aligned}$$

THEOREM: **Envelope Theorem**: Let $U \subseteq \mathbb{R}^L$ be open, let $\phi : U \rightarrow \mathbb{R}^N$ be differentiable, and suppose that $\phi(\boldsymbol{\theta})$ is a solution to the parametrized MAX problem for every $\boldsymbol{\theta} \in U$.
Let $f^v : U \rightarrow \mathbb{R}$ be defined by $f^v(\boldsymbol{\theta}) = f(\phi(\boldsymbol{\theta}), \boldsymbol{\theta})$. Fix $\boldsymbol{\theta}^* \in U$ and let $\phi(\boldsymbol{\theta}^*) = \boldsymbol{x}^*$. Let $E$ be the set of indices $k$ such that $g_k(\boldsymbol{x}^*, \boldsymbol{\theta}^*) = 0$. Then

$$Df^v(\boldsymbol{\theta}^*) = D_{\boldsymbol{\theta}} f(\boldsymbol{x}^*, \boldsymbol{\theta}^*) - \sum_{k \in E} \lambda_k D_{\boldsymbol{\theta}} g_k(\boldsymbol{x}^*, \boldsymbol{\theta}^*)$$

where the $\lambda_k$ are the KT multipliers. This holds for every $\boldsymbol{\theta} \in U$.

# Proof

(For the max problem only)

Define $s : U \to \mathbb{R}^{N+L}$ as $s(\boldsymbol{\theta}) = (\phi(\boldsymbol{\theta}), \boldsymbol{\theta})$. Then $f^v(\boldsymbol{\theta}) = f(s(\boldsymbol{\theta}))$.

By the chain rule,

$$
\begin{aligned}
D f^v(\boldsymbol{\theta}^*) &= Df(\boldsymbol{x}^*, \boldsymbol{\theta}^*) \quad Ds(\boldsymbol{\theta}^*) \\
&= \begin{bmatrix} D_{\boldsymbol{x}} f(\boldsymbol{x}^*, \boldsymbol{\theta}^*) & D_{\boldsymbol{\theta}} f(\boldsymbol{x}^*, \boldsymbol{\theta}^*) \end{bmatrix} \begin{bmatrix} D\phi(\boldsymbol{\theta}^*) \\ I \end{bmatrix}
\end{aligned}
$$

$$
D f^v(\boldsymbol{\theta}^*) = D_{\boldsymbol{x}} f(\boldsymbol{x}^*, \boldsymbol{\theta}^*) D\phi(\boldsymbol{\theta}^*) + D_{\boldsymbol{\theta}} f(\boldsymbol{x}^*, \boldsymbol{\theta}^*)
$$

By KT (1), $D_{\boldsymbol{x}} f(\boldsymbol{x}^*, \boldsymbol{\theta}^*) = \sum_{k \in E} \lambda_k D_{\boldsymbol{x}} g_k(\boldsymbol{x}^*, \boldsymbol{\theta}^*)$

Substituting, $D f^v(\boldsymbol{\theta}^*) = \sum_{k \in E} \lambda_k D_{\boldsymbol{x}} g_k(\boldsymbol{x}^*, \boldsymbol{\theta}^*) D\phi(\boldsymbol{\theta}^*) + D_{\boldsymbol{\theta}} f(\boldsymbol{x}^*, \boldsymbol{\theta}^*)$

# PROOF (2)

For any $k$, let $h_k(\boldsymbol{\theta}) = g_k(\phi(\boldsymbol{\theta}), \boldsymbol{\theta})$. For $k \in E$, $h_k(\boldsymbol{\theta}^*) = 0$ while for all other $\boldsymbol{\theta} \in U$, $h_k(\boldsymbol{\theta}) \leq 0$. Therefore, $\boldsymbol{\theta}^*$ maximizes $h_k$ on $U$.

Hence, $Dh_k(\boldsymbol{\theta}^*) = 0$ .

By the chain rule, $Dh_k(\boldsymbol{\theta}^*) = D_{\boldsymbol{x}} g_k(\boldsymbol{x}^*, \boldsymbol{\theta}^*) D\phi(\boldsymbol{\theta}^*) + D_{\boldsymbol{\theta}} g_k(\boldsymbol{x}^*, \boldsymbol{\theta}^*)$

On the other hand, for any $k \notin E$, $\lambda_k = 0$, by KT(2).

Substituting $D_{\boldsymbol{x}} g_k(\boldsymbol{x}^*, \boldsymbol{\theta}^*) D\phi(\boldsymbol{\theta}^*) = -D_{\boldsymbol{\theta}} g_k(\boldsymbol{x}^*, \boldsymbol{\theta}^*)$ into the expression for $D f^v(\boldsymbol{\theta}^*)$ yields the result. ■

# INTUITION

Solve the maximization problem and get the solution in terms of the parameters: $\boldsymbol{x}^* = \phi(\boldsymbol{\theta})$.

If we take this value and plug it into the objective function, we get $f^v = f(\phi(\boldsymbol{\theta}), \boldsymbol{\theta})$, the value function.

What happens with the value function when the parameters $\boldsymbol{\theta}$ change? we're looking for $Df^v(\boldsymbol{\theta})$.

What happens if $\boldsymbol{\theta}$ changes from $\boldsymbol{\theta}^*$ to $\hat{\boldsymbol{\theta}}$?

(i) The objective function changes from $f(\boldsymbol{x}, \boldsymbol{\theta}^*)$ to $f(\boldsymbol{x}, \hat{\boldsymbol{\theta}})$.

(ii) The maximizer changes from $\boldsymbol{x}^* = \arg\max f(\boldsymbol{x}, \boldsymbol{\theta}^*) = \phi(\boldsymbol{\theta}^*)$ to $\hat{\boldsymbol{x}} = \arg\max f(\boldsymbol{x}, \hat{\boldsymbol{\theta}}) = \phi(\hat{\boldsymbol{\theta}})$.

(iii) The value function, from $f^v(\boldsymbol{\theta}^*) = f(\phi(\boldsymbol{\theta}^*), \boldsymbol{\theta}^*)$ to $f^v(\hat{\boldsymbol{\theta}}) = f(\phi(\hat{\boldsymbol{\theta}}), \hat{\boldsymbol{\theta}})$.

# INTERIOR MAXIMUM

$\max f(\boldsymbol{x}, \boldsymbol{\theta})$ [picture; $A, B, C, D, E$]

We want: $f(\phi(\hat{\boldsymbol{\theta}}), \hat{\boldsymbol{\theta}}) - f(\phi(\boldsymbol{\theta}^*), \boldsymbol{\theta}^*) = CD$

By the Chain Rule, $D_{\boldsymbol{\theta}} f^v(\boldsymbol{\theta}^*) = \underbrace{D_{\boldsymbol{x}} f(\boldsymbol{x}^*, \boldsymbol{\theta}^*) D\phi(\boldsymbol{\theta}^*)}_{indirect\ effect} + \underbrace{D_{\boldsymbol{\theta}} f(\boldsymbol{x}^*, \boldsymbol{\theta}^*)}_{direct\ effect}$

Since $\boldsymbol{x}^*$ is optimal when $\boldsymbol{\theta} = \boldsymbol{\theta}^*$, by KT(1), $D_{\boldsymbol{x}} f(\boldsymbol{x}^*, \boldsymbol{\theta}^*) = 0$ (no indirect effect)

(Graphically, $f$ is flat around $\boldsymbol{x}^*$, so $DE \approx 0$.)

Hence, $D_{\boldsymbol{\theta}} f^v(\boldsymbol{\theta}^*) = D_{\boldsymbol{\theta}} f(\boldsymbol{x}^*, \boldsymbol{\theta}^*)$

(Graphically, CD can be approximated by BA)

# EXAMPLE

$$f(x,\theta) = -(x-2\theta)^2 + \theta^2$$

So $\phi(\theta) = 2\theta$

Then $\quad f^v(\theta) = f(\phi(\theta),\theta) = \theta^2$

By direct calculation, $\quad Df^v(\theta^*) = 2\theta^*$

By Envelope theorem, $\quad D_\theta f(x^*,\theta^*) = 4(x^* - 2\theta^*) + 2\theta^*$

Substituting for $x^*$, $\quad D_\theta f(x^*,\theta^*) = 2\theta^*$

# CONSTRAINED MAXIMUM

$$\begin{aligned} \max \quad & f(\boldsymbol{x}, \boldsymbol{\theta}^*) \\ \text{s.t.} \quad & g_1(\boldsymbol{x}, \boldsymbol{\theta}^*) \leq 0 \\ & \vdots \\ & g_k(\boldsymbol{x}, \boldsymbol{\theta}^*) \leq 0 \end{aligned}$$

[picture]

$$D f^v(\boldsymbol{\theta}^*) = D_{\boldsymbol{\theta}} f(\phi(\boldsymbol{\theta}^*), \boldsymbol{\theta}^*) - \sum_{k \in E} \lambda_k D_{\boldsymbol{\theta}} g_k(\phi(\boldsymbol{\theta}^*), \boldsymbol{\theta}^*)$$

The effect of the change of the parameters on the constraints is captured by $D_{\boldsymbol{\theta}} g_k(\phi(\boldsymbol{\theta}^*), \boldsymbol{\theta}^*)$.

Suppose it is positive, so that $g_k$ increases.

Then constraint $k$ becomes harder to satisfy. The feasible set shrinks, so $f$ may decrease.

By how much? Depends on $\lambda_k$.

# CONSTRAINED MAXIMUM (2)

Suppose that parameter $\theta_k$ affects only constraint $k$ additively:
$g_k(\boldsymbol{x}) - \theta_k \leq 0$

So increasing $\theta_k$ weakens the constraint.

By envelope theorem, at $\theta_k = 0$ for all $k$,

if $k \in E$ then $D_{\theta_k} f^v(0) = -\lambda_k(-1) = \lambda_k$.

if $k \notin E$ then KT2 implies that $\lambda_k = 0$,

while the envelope theorem implies that $D_{\theta_k} f^v(0) = \lambda_k$.

That is, $\lambda_k$ is the value of increasing constraint $k$.

# EXAMPLE

$$f(x, \theta) = \ln(x + 1)$$

$$x^* = \phi(\theta) = \theta$$

$$f^v(\theta) = \ln(\phi(\theta) + 1) = \ln(\theta + 1)$$

Direct calculation: $Df^v(\theta^*) = \frac{1}{1+\theta^*}$

Envelope theorem: $D_\theta f(x^*, \theta^*) = 0$, so $Df^v(\theta^*) = \lambda$

By KT(1), $\frac{1}{1+x} = \lambda(1)$

So $Df^v(\theta^*) = \frac{1}{1+x^*} = \frac{1}{1+\theta^*}$