Caltech Division of the Humanities and Social Sciences

Classical Envelope Theorem

KC Border

Revised October, 2006
v. 2015.11.11::14.20

The graph of a real-valued function f on $[0,1]$ is a curve in the plane, and we shall also refer to the curve itself as f. Given a one-dimensional parametrized family $f_\alpha\colon [0,1] \to \boldsymbol{R}$ of such curves, where α runs over some interval, the curve $h\colon [0,1] \to \boldsymbol{R}$ is the **envelope** of the family if each point on the curve h is tangent to the graph of one of the curves f_α and each curve f_α is tangent to h. (See, e.g., Apostol [1, p. 342] for this definition.) That is, for each α, there is some t and also for each t, there is some α, satisfying

$$f_\alpha(t) = h(t) \text{ and } f'_\alpha(t) = h'(t).$$

If the correspondence between curves and points on the envelope is one-to-one, then we may regard h as a function of α.

Consider now an unconstrained parametrized maximization problem. Let $x^*(p)$ be the value of the control variable x that maximizes $f(x,p)$, where p is our parameter of interest. For fixed x, the function

$$\varphi_x(p) = f(x,p)$$

defines a curve (or more generally a surface). The value function $V(p)$ satisfies

$$V(p) = f\bigl(x^*(p), p\bigr) = \max_x \varphi_x(p).$$

Under appropriate conditions, the graph of the value function V will be the envelope of the curves (surfaces) φ_x. "Envelope theorems" in maximization theory are concerned with the tangency conditions this entails.

To get a picture of this result, imagine a plot of the graph of f. It is the surface $z = f(x,p)$ in (x,p,z)-space. Orient the graph so that the x-axis is perpendicular to the page and the p-axis runs horizontally across the page, and the z-axis is vertical. The high points of the surface (minus perspective effects) determine the graph of the value function V. Here is an example:

1 Example Let

$$f(x,p) = p - (x-p)^2 + 1, \quad 0 \leqslant x, p \leqslant 2.$$

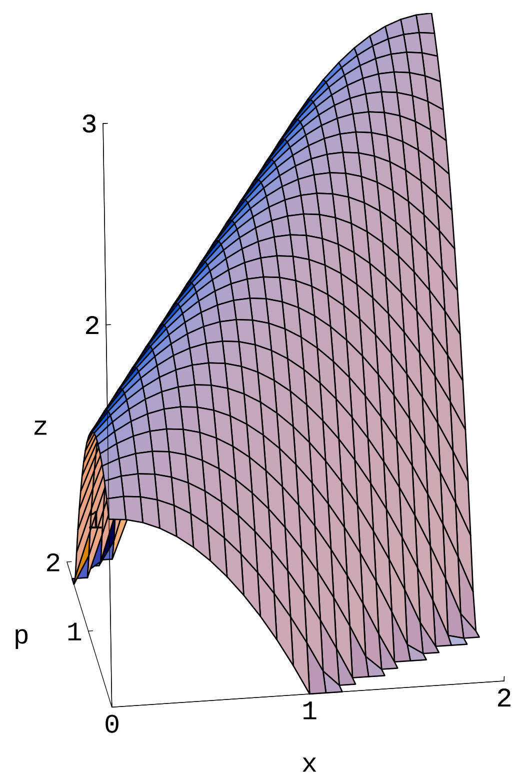

Figure 1. Graph of $f(x, p) = p - (x - p)^2 + 1$.

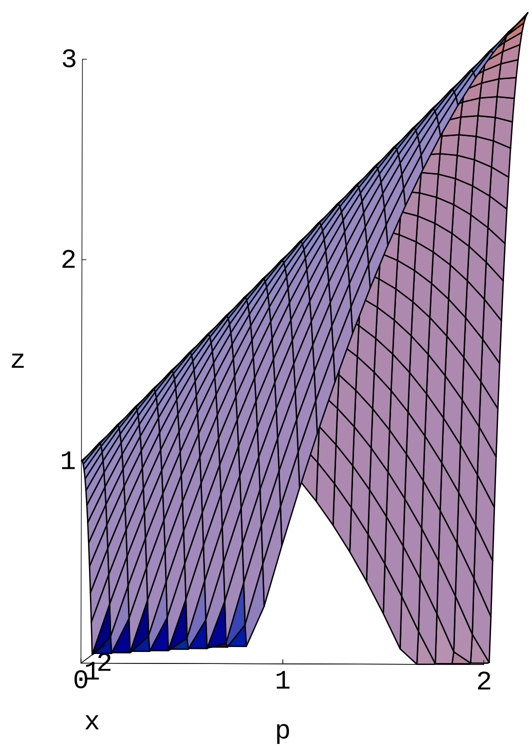

Figure 2. Graph of $f(x,p) = p - (x-p)^2 + 1$ viewed from the side.

See Figure 1. Then given p, the maximizing x is given by $x^*(p) = p$, and $V(p) = p + 1$. The side-view of this graph in Figure 2 shows that the high points do indeed lie on the line $z = 1 + p$. For each x, the function φ_x is given by

$$\varphi_x(p) = p - (x - p)^2 + 1.$$

The graphs of these functions and of V are shown for selected values of x in Figure 3. Note that the graph of V is the envelope of the family of graphs of the functions φ_x. Consequently the slope of V is the slope of the φ_x to which it is tangent, that is,

$$V'(p) = \frac{\partial f}{\partial p}\Big|_{x=x^*(p)=p} = 1 + 2(x - p)\Big|_{x=p} = 1.$$

This last observation is one version of the Envelope Theorem. □

1 An envelope theorem for unconstrained maximization

The following theorem is proven in my on-line notes on maximization.

2 Envelope theorem for unconstrained maximization *Let X be a metric space and P an open subset of R^n. Let $w\colon X \times P \to \boldsymbol{R}$ and assume $\frac{\partial w}{\partial p}$ exists and is continuous in $X \times P$. For each $p \in P$, let $x^*(p)$ maximize $w(x, p)$ over X. Set*

$$V(p) = w(x^*(p), p).$$

Assume that $x^*\colon P \to X$ is a continuous function. Then V is continuously differentiable and

$$DV(p) = \frac{\partial w}{\partial p}(x^*(p), p).$$

2 Constrained Maxima

3 Theorem *Let $X \subset \boldsymbol{R}^n$ and $P \subset \boldsymbol{R}^\ell$ be open, and assume that the functions $f, g_1, \ldots, g_m\colon X \times P \to \boldsymbol{R}$ are C^1. For each $p \in P$, let $x^*(p)$ be an interior constrained local maximizer of $f(x, p)$ subject to $g(x, p) = 0$. Define the Lagrangean*

$$L(x, \lambda; p) = f(x, p) + \sum_{i=1}^{m} \lambda_i g_i(x, p),$$

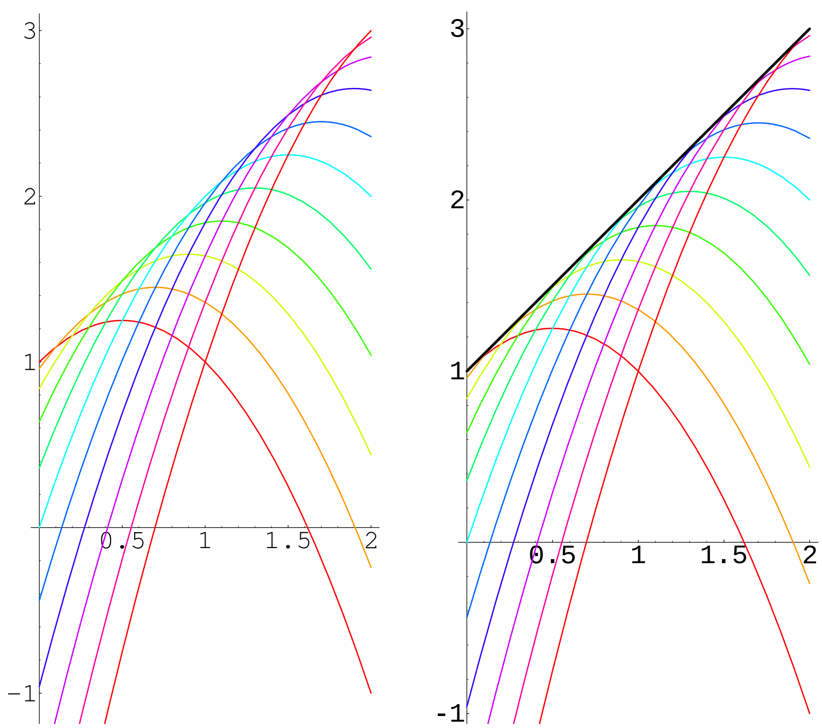

Figure 3. Graph of $V(p) = p + 1$ as the envelope of the family of curves $\{\varphi_x(p)\colon x = 0.0, 0.2, \dots, 2.0\}$, where $\varphi_x(p) = p - (x - p)^2 + 1 = f(x, p)$.

and assume that the conclusion of the Lagrange Multiplier Theorem holds for each p, that is, there exist real numbers $\lambda_1^*(p), \ldots, \lambda_m^*(p)$, such that the first order conditions

$$\frac{\partial L(x^*(p), \lambda^*(p), p)}{\partial x} = \nabla_x f(x^*(p), p) + \sum_{i=1}^{m} \lambda_i^*(p) \nabla_x g_i(x^*(p), p) = 0$$

are satisfied. Assume that $x^* \colon P \to X$ and $\lambda^* \colon P \to \boldsymbol{R}^{\mathrm{m}}$ are C^1. Set

$$V(p) = f\left(x^*(p), p\right).$$

Then V is C^1 and

$$\frac{\partial V(p)}{\partial p_j} = \frac{\partial L(x^*(p), \lambda^*(p), p)}{\partial p_j} = \frac{\partial f(x^*, p)}{\partial p_j} + \sum_{i=1}^{m} \lambda_i^* \frac{\partial g_i(x^*, p)}{\partial p_j}.$$

Proof: Clearly V is C^1 as the composition of C^1 functions. Since x^* satisfies the constraints, we have

$$V(p) = f(x^*(p), p) = f(x^*(p), p) + \sum_{i=1}^{m} \lambda_i^*(p) g_i(x^*(p), p).$$

Therefore by the chain rule,

$$\begin{aligned}
\frac{\partial V(p)}{\partial p_j} &= \left(\sum_{k=1}^{n} \frac{\partial f(x^*, p)}{\partial x_k} \frac{\partial x_k^*}{\partial p_j} \right) + \frac{\partial f(x^*, p)}{\partial p_j} \\
&\quad + \sum_{i=1}^{m} \left\{ \frac{\partial \lambda_i^*}{\partial p_j} g_i(x^*, p) + \lambda^* \left[\left(\sum_{k=1}^{n} \frac{\partial g_i(x^*, p)}{\partial x_k} \frac{\partial x_k^*}{\partial p_j} \right) + \frac{\partial g_i(x^*, p)}{\partial p_j} \right] \right\} \\
&= \frac{\partial f(x^*, p)}{\partial p_j} + \sum_{i=1}^{m} \lambda_i^* \frac{\partial g_i(x^*, p)}{\partial p_j} \\
&\quad + \sum_{i=1}^{m} \frac{\partial \lambda_i^*}{\partial p_j} g_i(x^*, p) \qquad (1) \\
&\quad + \sum_{k=1}^{n} \left(\frac{\partial f(x^*, p)}{\partial x_k} + \sum_{i=1}^{m} \lambda^* \frac{\partial g_i(x^*, p)}{\partial x_k} \right) \frac{\partial x_k^*}{\partial p_j}. \qquad (2)
\end{aligned}$$

The theorem now follows from the fact that both terms (1) and (2) are zero. Term (1) is zero since x^* satisfies the constraints, and term (2) is zero, since the first order conditions imply that each $\frac{\partial f(x^*,p)}{\partial x_k} + \sum_{i=1}^{m} \lambda^* \frac{\partial g_i(x^*,p)}{\partial x_k} = 0$. ∎

2.1 Another Envelope Theorem

The previous theorem assume only that the conclusion of Lagrange Multiplier Theorem held. This version requires the assumptions of the Lagrange Multiplier Theorem to hold, but dispenses with the assumption that the multipliers are a C^1 function of the parameters. At the moment, there is an uncomfortable gap in the proof, so label it a conjecture.

4 Conjecture *Let $X \subset \boldsymbol{R}^{\mathrm{n}}$ and $P \subset \boldsymbol{R}^{\ell}$ be open, and assume that the functions $f, g_1, \dots, g_m \colon X \times P \to \boldsymbol{R}$ are C^1. For each $p \in P$, let $x^*(p)$ be an interior constrained local maximizer of $f(x,p)$ subject to $g(x,p) = 0$. Assume that for each p, the gradients (with respect to x) ${g_i}'_x$ are linearly independent at $(x^*(p), p)$. Assume that $x^* \colon P \to X$ is C^1.*

Define the Lagrangean

$$L(x, \lambda; p) = f(x,p) + \sum_{i=1}^{m} \lambda_i g_i(x,p).$$

Then for each p there exist real numbers $\lambda_1^*(p), \dots, \lambda_m^*(p)$, such that the first order conditions Notation!!!!

$$\frac{\partial L(x^*(p), \lambda^*(p), p)}{\partial x} = f'_x(x^*(p), p) + \sum_{i=1}^{m} \lambda_i^*(p) {g_i}'_x(x^*(p), p) = 0$$

are satisfied. Set

$$V(p) = f\left(x^*(p), p\right).$$

Then V is C^1 and

$$\frac{\partial V(p)}{\partial p_j} = \frac{\partial L(x^*(p), \lambda^*(p), p)}{\partial p_j} = \frac{\partial f(x^*, p)}{\partial p_j} + \sum_{i=1}^{m} \lambda_i^*(p) \frac{\partial g_i(x^*, p)}{\partial p_j}.$$

The main idea of this proof appears in many places, e.g., Silberberg [4, 5], Clarke et. al. [2], and Diamond and McFadden [3] (who attribute it to Gorman).

Proof: As in the proof of Theorem 3, the function V is clearly C^1. Now observe that we can embed our maximization in the family of problems

$$\text{maximize } f(x,p) \text{ subject to } g(x,p) - \alpha = 0 \qquad (\mathrm{P}(\alpha))$$

where α ranges over a neighborhood 0 in $\boldsymbol{R}^{\mathrm{m}}$. The first thing we have to show is that for each α, there is some (x,p) satisfying $g(x,p) + \alpha = 0$. We have already assumed that for each p there is some x_p satisfying $g(x_p, p) = 0$. Indeed $x_p = x^*(p)$ works. Now consider the function

$$h_p(x, \alpha) = g(x, p) - \alpha.$$

By hypothesis $h_p(x_p, 0) = 0$. The Jacobian of h with respect to x is just the Jacobian of g, which is of full rank by our linear independence hypothesis. Therefore by the Implicit Function Theorem, there is a neighborhood U of 0 in $\boldsymbol{R}^{\mathrm{m}}$ such that $\alpha \in U$ implies the existence of some $\hat{x}_p(\alpha)$ such that $h_p(\hat{x}_p(\alpha), \alpha) = 0$. Thus each problem $\mathrm{P}(\alpha)$ is feasible.

One gap in the proof is to show that in fact each $\mathrm{P}(\alpha)$ has an optimal solution. Assume for now that this is so, and let $x^*(p, \alpha) = \hat{x}_p(\alpha)$ be the optimum. Another gap is to show that x^* is a differentiable function of both p and α. (This is another application of the Implicit Function Theorem, but I'll leave it out for now.) Modify the definition of V so that

$$V(p, \alpha) = f\big(x^*(p, \alpha), p\big).$$

Now for any x and p, if we set $\alpha = g(x, p)$, then x satisfies $g(x, p) + \alpha = 0$. In particular, the value $f(x, p)$ is less than or equal to the optimal value $V\big(p, g(x, p)\big)$. In other words,

$$h(x, p) = V\big(p, g(x, p)\big) - f(x, p) \geqslant 0,$$

and is equal to zero for $x = x^*\big(p, g(x, p)\big)$. Thus minima of h occur whenever $x = x^*(p, 0)$. The first order conditions for this minimum are that

$$\begin{aligned} \frac{\partial h}{\partial x_j} &= 0 \qquad j = 1, \ldots, n, \\ \frac{\partial h}{\partial p_i} &= 0 \qquad i = 1, \ldots, m. \end{aligned}$$

The first group of first order conditions imply

$$\frac{\partial h}{\partial x_j} = \sum_{k=1}^{m} \frac{\partial V}{\partial \alpha_i} \frac{\partial g_k}{\partial x_j} - \frac{\partial f}{\partial x_j} = 0,$$

which tells us that

$$\lambda_i^* = -\frac{\partial V}{\partial \alpha_i}$$

are the desired Lagrange multipliers. The second group of first order conditions imply

$$\frac{\partial h}{\partial p_i} = \frac{\partial V}{\partial p_i} + \sum_{k=1}^{m} \frac{\partial V}{\partial \alpha_i} \frac{\partial g_k}{\partial p_i} - \frac{\partial f}{\partial p_i} = 0,$$

or using the Lagrange multipliers defined above

$$\frac{\partial V}{\partial p_i} = \frac{\partial f}{\partial p_i} + \sum_{k=1}^{m} \lambda^* \frac{\partial g_k}{\partial p_i},$$

where of course the partials are evaluated at the optimizers. ∎

References

[1] T. M. Apostol. 1967. *Calculus*, 2d. ed., volume 1. Waltham, Massachusetts: Blaisdell.

[2] F. H. Clarke, Y. S. Ledyaev, R. J. Stern, and P. R. Wolenski. 1998. *Nonsmooth analysis and control theory*. Number 178 in Graduate Texts in Mathematics. New York: Springer–Verlag.

[3] P. A. Diamond and D. L. McFadden. 1974. Some uses of the expenditure function in public finance. *Journal of Public Economics* 3:3–21.

[4] E. Silberberg. 1971. The Le Chatelier principle as a corollary to a generalized envelope theorem. *Journal of Economic Theory* 3(2):146–155.
DOI: 10.1016/0022-0531(71)90012-3

[5] ——— . 1974. A revision of comparative statics methodology in economics, or how to do comparative statics on the back of an envelope. *Journal of Economic Theory* 7(2):159–172.
DOI: 10.1016/0022-0531(74)90104-5