The Envelope Theorem

CSCI 1440/2440

2025-02-05

We prove the celebrated envelope theorem. Then, by way of this theorem, we derive the symmetric equilibrium in first-price auctions and Myerson's payment characterization for DSIC auctions. When originally drafted, these notes followed the presentation in Quint;[1] by now, there are likely deviations.

[1] Dan Quint. Some beautiful theorems with beautiful proofs. University of Wisconsin–Madison, 2014

1 *Envelope Theorem*

Consider the following parameterized optimization problem:

$$V(\theta) = \max_{a \in A} f(a; \theta)$$

We write $f(\cdot; \theta)$ to indicate that f is "parameterized" by some $\theta \in \Theta$.[2] Via this parameterization, this seeming optimization over the set A is in fact an optimization over a strategy space of functions $s : \Theta \to A$.

[2] Such parameterizations are also sometimes denoted with subscripts instead of semicolons: i.e., $f_\theta(\cdot)$.

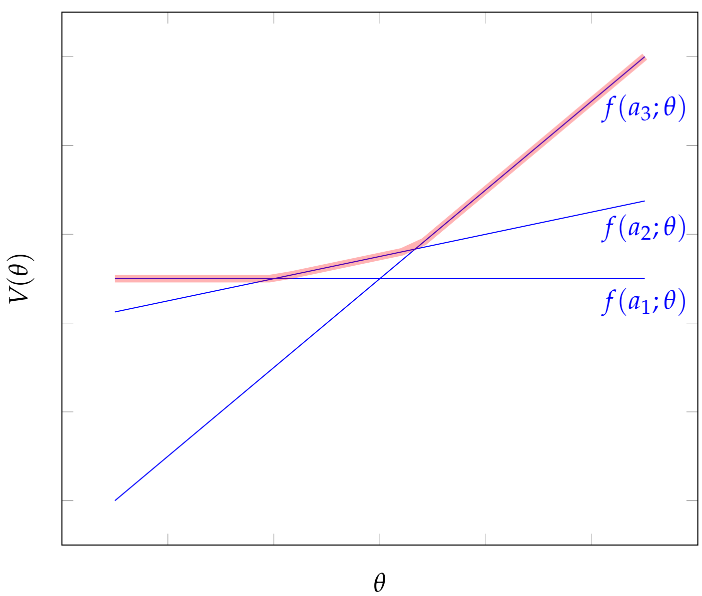

Figure 1: An illustration of an upper envelope (in red).

We prove the simplest version of the envelope theorem, where only the so-called **value function** V depends on the parameter θ; the constraint (or **choice**) set A does not.

Theorem 1.1. *Assume $A^*(\theta)$ is nonempty for all $\theta \in \Theta$, with $s^*(\theta)$ an element of $A^*(\theta)$, so that $V(\theta) = f(s^*(\theta); \theta)$. If f is continuously differentiable with respect to both a (the decision variable) and θ (the parameter), and $s^*(\theta)$ is continuously differentiable with respect to θ,[3] then*

[3] This assumption can be dropped. Indeed, the upper envelope in the example plotted in Figure 1 is not smooth; it has sharp corners. Thus, the assumption does not hold.

$$V'(\theta) = \frac{\mathrm{d}V(\theta)}{\mathrm{d}\theta} = \frac{\mathrm{d}\max_{a\in A} f(a;\theta)}{\mathrm{d}\theta} = \frac{\partial f(s^*(\theta);\theta)}{\partial\theta}$$

The envelope theorem isolates the effect of changes in the parameter θ on V as dependent only on the direct effect of changes in the parameter θ on f, regardless of the indirect effect of such changes on the optimizer $s^*(\theta)$. In particular, while the chain rule would suggest that we need to consider this indirect effect, the key takeaway of the envelope theorem is that we do not! Not only does the envelope theorem preclude the need for unnecessary computations, it skirts the potential difficulty of optimizing $s^*(\theta)$ when it is not differentiable, as in Figure 1.

Proof Sketch. Recall the chain rule from multivariable calculus: If $z(t) = f(x,y)$, where $x = x(t)$ and $y = y(t)$, and everything is continuously differentiable (i.e., x and y with respect to t, and f with respect to x and y), then

$$\frac{\mathrm{d}z}{\mathrm{d}t} = \frac{\partial f}{\partial x}\frac{\mathrm{d}x}{\mathrm{d}t} + \frac{\partial f}{\partial y}\frac{\mathrm{d}y}{\mathrm{d}t}$$

N.B. The notation $\mathrm{d}x/\mathrm{d}t$ is a **total derivative**; it describes how the function x varies with respect to *all* its variables, in this case, t. The notation $\partial f/\partial x$, in contrast, is a **partial derivative**; it describes how the function f varies with respect to the variable x only, holding all other variables constant.

By assumption, the function f is continuously differentiable with respect to both a and θ, and $s^*(\theta)$ is continuously differentiable with respect to θ. Therefore, we can apply the chain rule to the parameterized value function $V(\theta) = f(s^*(\theta);\theta)$, which yields:

$$\frac{\mathrm{d}V}{\mathrm{d}\theta} = \frac{\partial f(s^*(\theta);\theta)}{\partial a}\frac{\mathrm{d}s^*(\theta)}{\mathrm{d}\theta} + \frac{\partial f(s^*(\theta);\theta)}{\partial\theta}$$

The first summand in this expression can be understood as follows:

$$\frac{\partial f(s^*(\theta);\theta)}{\partial a} = \left.\frac{\partial f(a;\theta)}{\partial a}\right|_{a=s^*(\theta)}$$

But the partial derivative of f with respect to a evaluated at an optimum, namely $s^*(\theta)$, is necessarily 0, by the first-order optimality conditions. Hence, only the third term survives:

$$\frac{\mathrm{d}V}{\mathrm{d}\theta} = \frac{\partial f(s^*(\theta);\theta)}{\partial\theta} = \left.\frac{\partial f(a;\theta)}{\partial\theta}\right|_{a=s^*(\theta)}$$

□

Example 1.2. The following functions are shown in Figure 1:

$$f(a_1;\theta) = 1$$
$$f(a_2;\theta) = 1/4\,\theta + 3/2$$
$$f(a_3;\theta) = \theta + 1$$

Next observe the following:

$$s^*(\theta) = \begin{cases} a_1 & \text{if } \theta < -2 \\ a_2 & \text{if } -2 < \theta < 2/3 \\ a_3 & \text{if } \theta > 2/3 \end{cases}$$

Thus,

$$V(\theta) = f(s^*(\theta); \theta) = \begin{cases} f(a_1; \theta) & \text{if } \theta < -2 \\ f(a_2; \theta) & \text{if } -2 < \theta < 2/3 \\ f(a_3; \theta) & \text{if } \theta > 2/3 \end{cases}$$

Therefore, by the envelope theorem:

$$V'(\theta) = \begin{cases} 0 & \text{if } \theta < -2 \\ 1/4 & \text{if } -2 < \theta < 2/3 \\ 1 & \text{if } \theta > 2/3 \end{cases}$$

In particular, the derivative of the value function with respect to θ can be computed without differentiating through the optimizer $s^*(\theta)$.

2 *Key Observation*

Recall that an auction is defined by two rules, an allocation rule and a payment rule. We consider a single-parameter auction for one good in which each bidder's i's valuation/value for the good is described by one number, $v_i \in T_i$, based on which she chooses her bid $b_i \in B_i$.

Fixing all other agents' strategies $\mathbf{s}_{-i} : T_{-i} \to B_{-i}$, we abbreviate bidder i's payment when she bids b_i by $p_i(b_i) \doteq p_i(b_i, \mathbf{s}_{-i}(\cdot))$. Her quasilinear utility if she is allocated the good is then $v_i - p_i(b_i)$. Furthermore, since her allocation probability is $x_i(b_i)$, her expected utility is $x_i(b_i)(v_i - p_i(b_i))$.

In this setting, the envelope theorem yields an interesting insight about the optimal expected utility function, namely that *its derivative is the (expected) allocation function*.

Theorem 2.1. *Given a bidder i with value v_i and quasilinear utility function $u_i(b_i) = x_i(b_i)(v_i - p_i(b_i))$, so that her* optimal *utility function is given by*

$$U(v_i) = \max_{b_i \in B_i} x_i(b_i)(v_i - p_i(b_i)) \ .$$

The derivative of her optimal utility function with respect to her value is her allocation function: i.e.,

$$\frac{\mathrm{d}U(v_i)}{\mathrm{d}v_i} = x_i(b_i^*) \ ,$$

where $b_i^* = s_i^*(v_i)$ denotes a utility-maximizing bid.

Proof. Let $f(b_i; v_i) = x_i(b_i)(v_i - p_i(b_i))$, and observe the following:

$$\frac{\partial f(b_i; v_i)}{\partial v_i} = \frac{\partial x_i(b_i)(v_i - p_i(b_i))}{\partial v_i} = x_i(b_i)$$

Therefore, by the envelope theorem,

$$\frac{\mathrm{d}U(v_i)}{\mathrm{d}v_i} = \frac{\partial f(s_i^*(v_i); v_i)}{\partial v_i} = \frac{\partial f(b_i; v_i)}{\partial v_i}\Big|_{b_i = s_i^*(v_i)} = x_i(b_i)\Big|_{b_i = s_i^*(v_i)} = x_i(b_i^*)$$

□

3 *Equilibrium Derivations via the Envelope Theorem*

Next, let's use the envelope theorem to analyze a symmetric first-price auction, meaning one in which the bidders' values are drawn i.i.d. from a bounded distribution F on $[\underline{v}, \overline{v}]$, for some $\underline{v} \leq \overline{v} \in \mathbb{R}$.

Assume a symmetric equilibrium $s^*(v)$ that is non-decreasing in v, so that a bidder with the highest value wins. At such an equilibrium, if the allocation probability is $x(v)$ and the winner pays her bid $s^*(v)$, each bidder's expected utility at equilibrium is given by $U^*(v) = x(v)(v - s^*(v))$. The probability that a bidder i with value v wins is the probability that $v \geq v_j$, for all $j \neq i$: i.e., $F^{n-1}(v)$. Each bidder's expected utility at equilibrium is thus:

$$U^*(v) = F^{n-1}(v)(v - s^*(v)) \quad . \tag{1}$$

By Theorem 2.1, ${\mathrm{d}U^*(v)}/{\mathrm{d}v} = x(s^*(v))$. Moreover, by the monotonicity assumption (i.e., $s^*(v)$ is non-decreasing in v), $x(s^*(v)) = F^{n-1}(v)$. Therefore, by the fundamental theorem of calculus,

$$U^*(v) = \int_{\underline{v}}^{v} F^{n-1}(t)\, \mathrm{d}t \quad , \tag{2}$$

as $U^*(\underline{v}) = 0$. Setting these two expressions for $U^*(v)$ (Equations 1 and 2) equal to one another yields

$$F^{n-1}(v)(v - s^*(v)) = \int_{\underline{v}}^{v} F^{n-1}(t)\, \mathrm{d}t \quad , \tag{3}$$

from which it follows that

$$s^*(v) = v - \frac{\int_{\underline{v}}^{v} F^{n-1}(t)\, \mathrm{d}t}{F^{n-1}(v)} \quad . \tag{4}$$

Finally, since[4]

$$vF^{n-1}(v) - \int_{\underline{v}}^{v} F^{n-1}(t)\, \mathrm{d}t = \int_{\underline{v}}^{v} t\, \mathrm{d}F^{n-1} \quad , \tag{5}$$

it follows that

$$s^*(v) = \frac{\int_{\underline{v}}^{v} t\, \mathrm{d}F^{n-1}}{F^{n-1}(v)} \quad . \tag{6}$$

[4] See Math'l Aside at the start of Lecture 5 on Myerson's optimal auction design.

In other words, at equilibrium in a symmetric first-price auction, bidders shade their bids in such a way that the result is the expected bid of the bidder with the second-highest value, conditioned on their value being highest.

Remark 3.1. This derivation establishes *necessary* conditions for s^* to be a symmetric equilibrium of a symmetric first-price auction: i.e., if s^* is such an equilibrium, then it must take the form of Equation 4 (or equivalently, Equation 6).

Last week, we proved this same result in the special case of $F = U[0,1]$. We can easily recover last week's result from this week's as follows. First, since $\underline{v} = 0$ and $F^{n-1}(t) = t^{n-1}$,

$$\int_0^v F^{n-1}(t)dt = \int_0^v t^{n-1}dt = \frac{1}{n}t^n\Big|_0^{\overline{v}} = \frac{v^n}{n} \ .$$

Second, plugging this calculation into Equation 4, and again using the fact that $F^{n-1}(v) = v^{n-1}$, yields:

$$s^*(v) = v - \frac{v^n}{nv^{n-1}} = v - \frac{v}{n} = v\left(1 - \frac{1}{n}\right) = \left(\frac{n-1}{n}\right)v \ .$$

4 Myerson's Payment Formula via the Envelope Theorem

Finally, again using the envelope theorem, we prove (one direction of) Myerson's payment characterization theorem—that the DSIC assumption implies Myerson's payment formula.

Proof. Reverting back to our original notation, we denote bidder i's optimal utility (or value function) by V_i. By Theorem 2.1, $V_i'(v_i) = x_i(s_i^*(v_i))$. Moreover, by the DSIC assumption, bidder i's expected utility is maximized at v_i: i.e., $s_i^*(v_i) = v_i$. Therefore, $V_i'(v_i) = x_i(v_i)$.

More specifically, $V_i'(v_i, \mathbf{v}_{-i}) = x_i(v_i, \mathbf{v}_{-i})$. But then, by **??** (which invokes the fundamental theorem of calculus),

$$V_i(v_i, \mathbf{v}_{-i}) - V_i(\underline{v}_i, \mathbf{v}_{-i}) = \int_{\underline{v}_i}^{v_i} V_i'(z, \mathbf{v}_{-i})\,\mathrm{d}z = \int_{\underline{v}_i}^{v_i} x_i(z, \mathbf{v}_{-i})\,\mathrm{d}z \ .$$

Next, letting $p_i(v_i, \mathbf{v}_{-i})$ denote bidder i's *expected* payment, we can also express bidder i's expected utility $V_i(v_i, \mathbf{v}_{-i})$ as $v_i x_i(v_i, \mathbf{v}_{-i}) - p_i(v_i, \mathbf{v}_{-i})$.

It now follows that

$$v_i x_i(v_i, \mathbf{v}_{-i}) - p_i(v_i, \mathbf{v}_{-i}) - \left(\underline{v}_i x_i(\underline{v}_i, \mathbf{v}_{-i}) - p_i(\underline{v}_i, \mathbf{v}_{-i})\right) = \int_{\underline{v}_i}^{v_i} x_i(z, \mathbf{v}_{-i})\,\mathrm{d}z \ .$$

In other words,

$$p_i(v_i, \mathbf{v}_{-i}) = v_i x_i(v_i, \mathbf{v}_{-i}) - \int_{\underline{v}_i}^{v_i} x_i(v_i, \mathbf{v}_{-i})\,\mathrm{d}z + p_i(\underline{v}_i, \mathbf{v}_{-i}) - \underline{v}_i x_i(\underline{v}_i, \mathbf{v}_{-i}) \ .$$

□

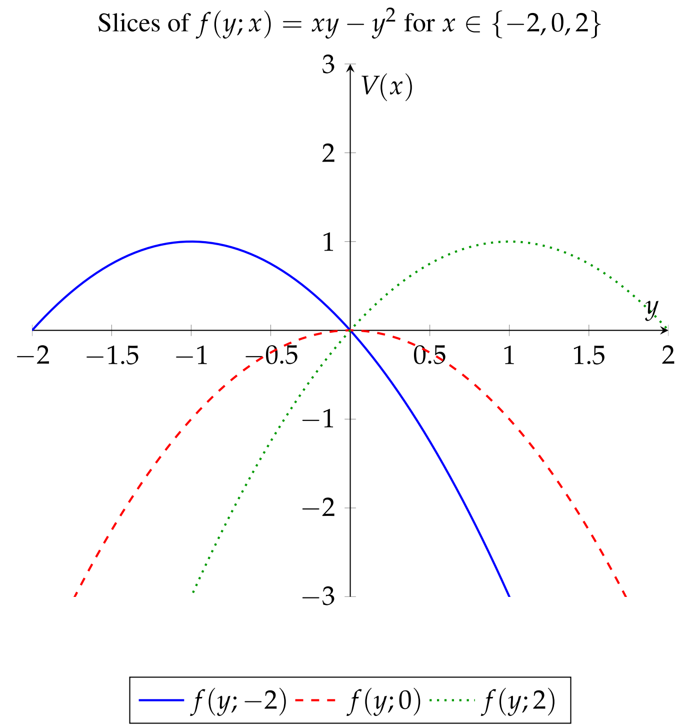

Figure 2: Another example of the envelope theorem.

A Another Example of the Envelope Theorem

First off, let's calculate $y^*(x)$, because when we use envelope theorem, we assume this value is known *a priori*. Holding x constant, we compute the partial derivative of $f(y;x)$ with respect to y and set it equal to zero:

$$\frac{\partial(xy - y^2)}{\partial y} = x - 2y = 0 \ ,$$

so that $y^*(x) = \frac{1}{2}x$.

Now, let's proceed to calculate the derivative of the value function with respect to the parameter the long way, using the chain rule:

$$\frac{\mathrm{d}V(x)}{\mathrm{d}x} = \frac{\partial f(y^*(x);x)}{\partial y}\frac{\mathrm{d}y^*(x)}{\mathrm{d}x} + \frac{\partial f(y^*(x);x)}{\partial x}$$

There are three terms to compute:

1.
$$\frac{\partial f(y;x)}{\partial y} = \frac{\partial(xy - y^2)}{\partial y} = x - 2y$$

Therefore,

$$\frac{\partial f(y;x)}{\partial x}\Big|_{y=y^*(x)} = x - 2y^*(x) = x - 2\left(\frac{1}{2}x\right) = x - x = 0$$

N.B. This conclusion follows from the first-order optimality conditions, which is precisely the insight that gives rise to the envelope theorem!

2.
$$\frac{\mathrm{d}y^*(x)}{\mathrm{d}x} = \frac{\mathrm{d}\frac{1}{2}(x)}{\mathrm{d}x} = \frac{1}{2}$$

3.
$$\frac{\partial f(y;x)}{\partial x} = \frac{\mathrm{d}(xy - y^2)}{\partial x} = y$$

Therefore,
$$\frac{\partial f(y;x)}{\partial x}\Big|_{y=y^*(x)} = y^*(x) = \frac{1}{2}x$$

Putting it all together:
$$\frac{\mathrm{d}V(x)}{\mathrm{d}x} = 0\left(\frac{1}{2}\right) + \frac{1}{2}x = \frac{1}{2}x$$

The envelope theorem uses the fact that the first term in this computation is necessarily 0 to conclude:

$$\frac{\mathrm{d}V(x)}{\mathrm{d}x} = \frac{\partial f(y^*(x);x)}{\partial x} = \frac{\partial f(y(x);x)}{\partial x}\Big|_{y=y^*(x)}$$

In other words, the envelope theorem renders it unnecessary to carry out steps 1 and 2 in this derivation. Step 3 alone suffices!

References

[1] Dan Quint. Some beautiful theorems with beautiful proofs. University of Wisconsin–Madison, 2014.