Partial Derivatives, Total Derivatives, and the Envelope Theorem

The Envelope Theorem is a shortcut for taking the derivative of an optimized function with respect to a parameter. The key to understanding the envelope theorem is to deeply understand the difference between partial and total derivatives.

Partial Derivatives

Consider the function

$$f(x,z) = 2xz + z^2.$$

When taking partial derivatives with respect to one variable (say x), we treat the other variable (z) as a constant. Therefore,

$$\frac{\partial f}{\partial x} = 2z.$$

Partial derivatives are denoted by the ∂ symbol, or sometimes as f_x (partial derivative of $f(x,z)$ with respect to x).

Total Derivatives

When taking a total derivative, we must acknowledge that one variable can depend on another. So in our example, z can be a function of x. Therefore, if we change x, it will have both a *direct effect* on f and an *indirect effect* on f via z. Again, let

$$f(x,z) = 2xz + z^2.$$

However, suppose we also have a constraint that

$$z(x) = x^2.$$

This means that the total derivative is

$$\frac{df}{dx} = \underbrace{\frac{\partial f}{\partial x}}_{direct\ effect} + \underbrace{\frac{\partial f}{\partial z} \cdot \frac{dz}{dx}}_{indirect\ effect} = 2z + (2x + 2z) \cdot 2x$$

Introducing the Envelope Theorem via Example

Suppose we have the following utility function

$$u(x;a) = -x^2 + ax$$

where a is a parameter and x is a choice variable. We want to know $\frac{d}{da}u(x^*(a);a)$ where $u(x^*(a);a)$ is the maximized value of $u(x;a)$. We can solve this two ways:

1. First, solve for the optimal x^* by taking the first order condition and setting it equal to zero:

$$-2x^* + a = 0 \implies x^* = \frac{a}{2}.$$

If we plug $x^*(a)$ back in to $u(x;a)$ we get the maximum value of u given a:

$$\begin{aligned}
u(x^*(a);a) &= -(x^*(a))^2 + ax^*(a) \\
&= -\left(\frac{a}{2}\right)^2 + a\left(\frac{a}{2}\right) \\
&= \frac{a^2}{4}
\end{aligned}$$

Now we can just take $\frac{d}{da}u(x^*(a);a)$:

$$\frac{d}{da}u(x^*(a);a) = \frac{a}{2}.$$

2. Can we do this without first solving for x^* and plugging back in? The answer is yes, and it gets to the heart of the envelope theorem. Let's totally differentiate the maximized utility function with respect to a. The key here is to remember that x^* is a function of a. Therefore,

$$\frac{d}{da}u(x^*(a);a) = \frac{\partial u}{\partial x}(x^*(a);a) \cdot \frac{dx^*}{da} + \frac{\partial u}{\partial a}(x^*(a);a).$$

But we know something very important about $\frac{\partial u}{\partial x}(x^*(a);a)$. It's zero! This is because $x^*(a)$ is the optimal value of x. Therefore, we can rewrite

$$\underbrace{\frac{\partial u}{\partial x}(x^*(a);a)}_{=0} \cdot \frac{dx^*}{da} + \frac{\partial u}{\partial a}(x^*(a);a) = \frac{\partial u}{\partial a}(x^*(a);a).$$

We now just take

$$\frac{\partial u}{\partial a}(x^*(a);a) = x|_{x=x^*(a)} = \frac{a}{2}.$$

That's the same answer that we got before! And notice something interesting: the total derivative in this case is equal to the partial derivative:

$$\frac{d}{da}u(x^*(a);a) = \frac{\partial}{\partial a}u(x^*(a),a).$$

The Envelope Theorem

Now that we've worked through an example, here is the statement of the Envelope Theorem:

Theorem 1. *(Envelope Theorem for the unconstrained case). Let f(x;a) be a continuous function of $x \in R^n$ and a scalar a. For a given value of a, consider the maximization*

$$\max_x f(x;a)$$

and let the vector $x^*(a) = (x_1^*(a), x_2^*(a), ..., x_n^*(a))$ be the solution to this maximization. Further suppose that $x^*(a)$ is a continuous function. Then the total derivative of the maximized f is equal to the partial derivative of the maximized f. In other words,

$$\underbrace{\frac{d}{da}f(x^*(a);a)}_{\text{total derivative}} = \underbrace{\frac{\partial}{\partial a}f(x^*(a);a)}_{\text{partial derivative}}.$$

Proof. Taking the total derivative w.r.t. a yields:

$$\frac{d}{da}f(x^*(a);a) = \sum_{i=1}^{n} \frac{\partial}{\partial x_i} f(x^*(a);a) \cdot \frac{\partial}{\partial a} x_i^*(a) + \frac{\partial}{\partial a} f(x^*(a);a)$$

But what do we know about $\frac{\partial}{\partial x_i} f(x^*(a);a)$? Since $x^*(a)$ is optimal choice of x, then every partial derivative must be zero. In other words,

$$\frac{\partial}{\partial x_i} f(x^*(a);a) = 0 \quad \forall i = 1, 2, ...n.$$

Therefore,

$$\sum_{i=1}^{n} \underbrace{\frac{\partial}{\partial x_i} f(x^*(a);a)}_{=0} \cdot \frac{\partial}{\partial a} x_i^*(a) + \frac{\partial}{\partial a} f(x^*(a);a) = \frac{\partial}{\partial a} f(x^*(a);a).$$

Intuition

Remember that when you take a total derivative, you have both direct effects and indirect effects. In our example above, when we change a, that has a direct effect of $f(x^*(a);a)$. However, when we change a we also change $x^*(a)$ (because the agent will re-optimize), which in turn changes $f(x^*(a);a)$. These are the indirect effects. In other words, we can write

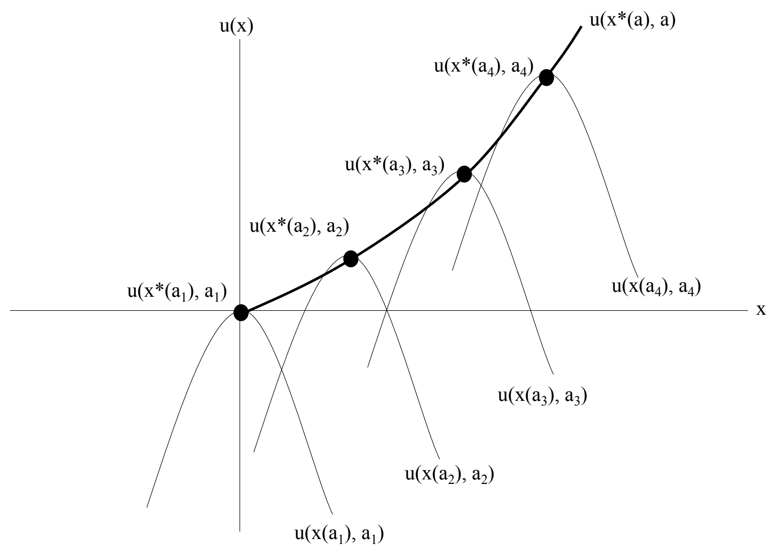
$$\frac{d}{da}f(x^*(a);a) = \underbrace{\sum_{i=1}^{n} \frac{\partial}{\partial x_i} f(x^*(a);a) \cdot \frac{\partial}{\partial a} x_i^*(a)}_{indirect\ effects} + \underbrace{\frac{\partial}{\partial a} f(x^*(a);a)}_{direct\ effect}.$$

However, think about these indirect effects. What is the slope of $f(x;a)$ when we are close to $x^*(a)$? It's zero in all n dimensions! So if we move a around, each $x_i^*(a)$ will change, but $f(x^*(a);a)$ won't change much. In other words, if we are at the optimal x, the indirect effects don't do much. In fact, they drop out altogether. Only the direct effect remains.

Visualization

Return to our example, where $u(x;a) = -x^2 + ax$. If we fix a value of $a = a_i$, this becomes a univariate function (a downward facing parabola). Figure 1 below plots $u(x;a)$ for four different values of a_i. The peak of each parabola is $u(x^*(a_i);a_i)$. The function $u(x^*(a);a)$ is traced out by the peaks of each $u(x^*(a_i);a_i)$ (the "upper envelope" - shown by the bold line on the graph).

Figure 1: The Envelope Theorem



Application to Consumer Theory

Let's apply the above abstract setup to a consumer's expenditure minimization problem. Let $f(\cdot)$ be the consumer's expenditure function, x be a vector of goods purchased, and a be goods prices that the consumer takes as given. In words, the envelope theorem tells us that–for a consumer who is already minimizing expenditure subject to a minimum utility constraint–the total effect of a single good's price increase on minimized expenditure is simply the direct effect (i.e. the price increase times the number of that good the consumer was previously purchasing). Even though the consumer may reallocate their consumption decision in response to the price change due to various income and substitution effects, the envelope theorem tells us that there is no (first-order) effect on minimized expenditure of these other changes!

You can also apply the constrained envelope theorem to a price change in the utility maximization problem, which leads to a result referred to as "Roy's Identity". For that problem, the only direct effect of a price change on the utility maximization problem is the effect on the budget constraint. The utility value of this tightening of the budget constraint is the value of Lagrange multiplier. And any corresponding changes in consumption do not have (first-order) effects on maximized utility!

The above discussion is a preview of "Shephard's Lemma" and "Roy's Identity", respectively. These apply the constrained envelope theorem to a price change in the expenditure minimization problem and the utility maximization problem. We will discuss these in more detail in the consumer theory section of the course.