Applied Microeconomics: Firm and Household

Lecture 2: Math Review

Jason Kerwin

Department of Applied Economics
University of Minnesota

September 6-8, 2016

Outline

- Calculus of One Variable
- Calculus of Several Variables

Functions

Definition: Given a set X called the *domain* and another set Y, a **function** f from X to Y, (written $f : X \to Y$) is a rule for assigning a unique element of Y to each element of X.

1. Note that every element of X is assigned to exactly one element of Y, whereas an element of Y may be assigned to any number of elements of X.
2. The **range** of the function is the set of elements of Y that are assigned to at least one element of X.
3. Suppose the variables x and y represent the elements of X and Y, respectively. Then, the function can be represented as
    - $y = f(x)$, where x is the independent variable, y is the dependent variable.

Functions

Definition: A function $f(x)$ is **continuous** at point c iff $\lim_{x\to c} f(x) = f(c)$. A function is called a continuous function if it is continuous at every value in its domain.

Definition: A function $f(x)$ is **differentiable** if its derivative $f'(x) = \left(\frac{df(x)}{dx_1}, ..., \frac{df(x)}{dx_n}\right)$ exists at each $x_0 \in X$.

Note: A differentiable function is always continuous, but a continuous function may or may not be differentiable.

Slopes and Elasticity

Derivatives,

1. The derivative of f, $dy/dx = f'(x)$ measures how fast the function is changing as x changes.
   - Geometrically, $f'(x)$ is the slope of $f(x)$ with respect to x
2. The second derivative of f, $d^2y/dx^2 = f''(x)$ measures how fast the first derivative is changing.

<u>Definition:</u> The percentage change in dependent variable for a given percentage change in independent variable is called the **elasticity** of the curve. Formally,

- $\epsilon = \frac{dy}{dx}\frac{x}{y} = f'(x)\frac{x}{y}$

Optimization

A firm's profit maximization problem (for a given production technology $f(x)$) can be formalized as:

- $\max\limits_{y,x} \pi(x, y) = py - wx, \quad s.t.\ y \leqslant f(x).$
  - The firm chooses its output y and input level x. Prices (p and w) are given.
  - The firm cannot choose output level that is beyond the feasible technology. Can it choose less?

Optimization

A firm's profit maximization problem (for a given production technology $f(x)$) can be formalized as:

- $\max\limits_{y,x} \pi(x,y) = py - wx, \quad s.t. \; y \leqslant f(x).$
  - The firm chooses its output y and input level x. Prices (p and w) are given.
  - The firm cannot choose output level that is beyond the feasible technology. Can it choose less?

Note that, because firm can choose any level of $y \leqslant f(x)$, so long as $p > 0$ a profit maximizing firm will always choose $y = f(x)$. Why?

Optimization

A firm's profit maximization problem (for a given production technology $f(x)$) can be formalized as:

- $\max_{y,x} \pi(x,y) = py - wx, \quad s.t.\; y \leqslant f(x).$
  - The firm chooses its output y and input level x. Prices (p and w) are given.
  - The firm cannot choose output level that is beyond the feasible technology. Can it choose less?

Note that, because firm can choose any level of $y \leqslant f(x)$, so long as $p > 0$ a profit maximizing firm will always choose $y = f(x)$. Why?

By substituting the (binding) constraint into the objective function we can transform the problem into a simpler form: an **unconstrained** optimization problem.

- $\max_{x} \pi(x) = pf(x) - wx$

General Rules for an Unconstrained Optimization

Let $f(x;\alpha)$ denote an objective function (e.g., profit, cost, production, or a utility function), where x is the vector of choice variables and α is the vector of model parameters.

The necessary and sufficient conditions for a **maximum** are:

- $\frac{\partial f(x^*;\alpha)}{\partial x} = f'(x^*;\alpha) = 0$
- $\frac{\partial^2 f(x^*;\alpha)}{\partial x^2} = f''(x^*;\alpha)$ is negative semi-definite.
- Solution: $x^* = x(\alpha)$

The necessary and sufficient conditions for a **minimum** are:

- $\frac{\partial f(x^*;\alpha)}{\partial x} = f'(x^*;\alpha) = 0$
- $\frac{\partial^2 f(x^*;\alpha)}{\partial x^2} = f''(x^*;\alpha)$ is positive semi-definite.
- Solution: $x^* = x(\alpha)$

Definition: The **Hessian matrix** is the square matrix of the second order partial derivatives of a function.

$$H(x) = \begin{bmatrix} \frac{\partial^2 f(x)}{\partial x_1^2} & \frac{\partial^2 f(x)}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f(x)}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f(x)}{\partial x_2 \partial x_1} & \frac{\partial^2 f(x)}{\partial x_2^2} & \cdots & \frac{\partial^2 f(x)}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f(x)}{\partial x_n \partial x_1} & \frac{\partial^2 f(x)}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f(x)}{\partial x_n^2} \end{bmatrix} = \begin{bmatrix} f_{11}(x) & f_{12}(x) & \cdots & f_{1n}(x) \\ f_{21}(x) & f_{22}(x) & \cdots & f_{2n}(x) \\ \vdots & \vdots & \ddots & \vdots \\ f_{n1}(x) & f_{n2}(x) & \cdots & f_{nn}(x) \end{bmatrix}$$

Definition: A function $f(x)$ is **twice differentiable** if its Hessian is defined at every point in its domain.

Young's Theorem: if $f(x)$ is twice differentiable the value of $f''(x)$ is invariant to the order of the differentiation: $f_{12} = f_{21}$.

- This implies that the Hessian is symmetric across the diagonal.

Mean value theorem, MVT

The **mean value theorem** helps us understand why $f'(x) = 0$ is a necessary condition for an optimum of $f(x)$.

The theorem states that if $f(x)$ is differentiable on the interval $a \leqslant x \leqslant b$ then there exists an x_0 such that

- $f'(x_0) = \frac{f(b)-f(a)}{b-a}$

It states that, at some point within the interval the slope of the cord is equal to the slope of the curve. The theorem can be rewritten as:

- $f(b) = f(a) + (b-a)f'(x_0)$

Mean value theorem, MVT

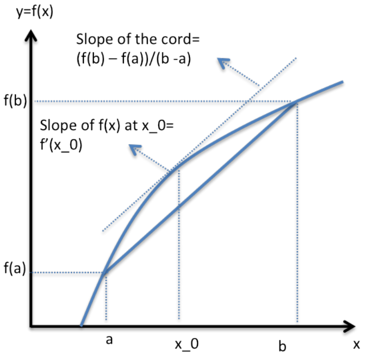
The **mean value theorem** helps us understand why $f'(x) = 0$ is a necessary condition for an optimum of $f(x)$.

The theorem states that if $f(x)$ is differentiable on the interval $a \leqslant x \leqslant b$ then there exists an x_0 such that

- $f'(x_0) = \frac{f(b)-f(a)}{b-a}$

It states that, at some point within the interval the slope of the cord is equal to the slope of the curve. The theorem can be rewritten as:

- $f(b) = f(a) + (b-a)f'(x_0)$



An Application of the MVT

Suppose that $x^* > 0$ maximizes a differentiable function $f(x)$. Also, let $x \geqslant 0$ denote any arbitrary point within the domain of $f(x)$. By the MVT there exists an x_0 such that

- $f(x^*) = f(x) + (x^* - x)f'(x_0)$
- $f(x^*) - f(x) = (x^* - x)f'(x_0)$

An Application of the MVT

Suppose that $x^* > 0$ maximizes a differentiable function $f(x)$. Also, let $x \geqslant 0$ denote any arbitrary point within the domain of $f(x)$. By the MVT there exists an x_0 such that

- $f(x^*) = f(x) + (x^* - x)f'(x_0)$
- $f(x^*) - f(x) = (x^* - x)f'(x_0)$

Because $x^* > 0$ maximizes $f(x)$ it must be true that

- $f(x^*) \geqslant f(x), \forall x$

An Application of the MVT

Suppose that $x^* > 0$ maximizes a differentiable function $f(x)$. Also, let $x \geqslant 0$ denote any arbitrary point within the domain of $f(x)$. By the MVT there exists an x_0 such that

- $f(x^*) = f(x) + (x^* - x)f'(x_0)$
- $f(x^*) - f(x) = (x^* - x)f'(x_0)$

Because $x^* > 0$ maximizes $f(x)$ it must be true that

- $f(x^*) \geqslant f(x), \forall x$
- $f(x^*) - f(x) \geqslant 0, \forall x$. Therefore,

An Application of the MVT

Suppose that $x^* > 0$ maximizes a differentiable function $f(x)$. Also, let $x \geqslant 0$ denote any arbitrary point within the domain of $f(x)$. By the MVT there exists an x_0 such that

- $f(x^*) = f(x) + (x^* - x)f'(x_0)$
- $f(x^*) - f(x) = (x^* - x)f'(x_0)$

Because $x^* > 0$ maximizes $f(x)$ it must be true that

- $f(x^*) \geqslant f(x), \forall x$
- $f(x^*) - f(x) \geqslant 0, \forall x$. Therefore,
- $(x^* - x)f'(x_0) \geqslant 0$

An Application of the MVT

Suppose that $x^* > 0$ maximizes a differentiable function $f(x)$. Also, let $x \geqslant 0$ denote any arbitrary point within the domain of $f(x)$. By the MVT there exists an x_0 such that

- $f(x^*) = f(x) + (x^* - x)f'(x_0)$
- $f(x^*) - f(x) = (x^* - x)f'(x_0)$

Because $x^* > 0$ maximizes $f(x)$ it must be true that

- $f(x^*) \geqslant f(x), \forall x$
- $f(x^*) - f(x) \geqslant 0, \forall x$. Therefore,
- $(x^* - x)f'(x_0) \geqslant 0$

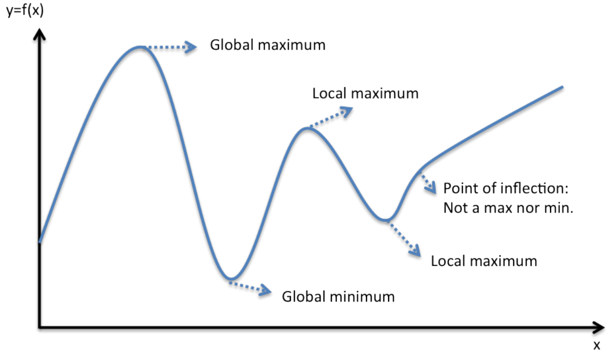
Implying that

- $f'(x_0) \leqslant 0$ if $x^* \leqslant x$ (slope is rising as we approach x^* from below)
- $f'(x_0) \geqslant 0$ if $x^* \geqslant x$ (slope is falling if we move beyond x^*)

If we choose x in a close neighbourhood of x^*, (i.e., $x \cong x^*$), then $f'(x_0) \cong f'(x^*) = 0$. In words, the slope of the function must be zero at the maximum, (also true for minimum). *This is the necessary condition*.

Local vs. Global optimum

The FOC helps us to find potential optima. However, we need to do further analysis to determine **i)** whether each optimum is a maximum or a minimum, and **ii)** whether it is a local or a global optimum. To this end, we discuss concavity and convexity.



Concavity & Convexity: Formal definitions

Definition: A function $f(X)$ is *concave* (*convex*) if for all $x_1, x_2 \in X$ and $\lambda \in [0, 1]$

- $f(\lambda x_1 + (1-\lambda)x_2) \underset{(\leqslant)}{\geqslant} \lambda f(x_1) + (1-\lambda)f(x_2)$

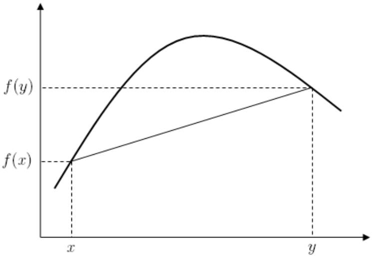
Definition: A function $f(X)$ is *strictly concave* (*strictly convex*) if strict inequality holds when $\lambda \in (0, 1)$

- $f(\lambda x_1 + (1-\lambda)x_2) \underset{(<)}{>} \lambda f(x_1) + (1-\lambda)f(x_2)$

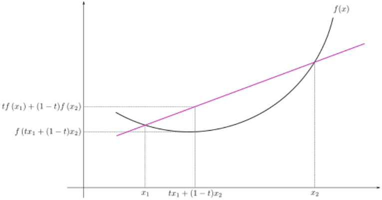
If f is concave (convex) and continuously differentiable the graph of f lies everywhere on or below (above) a tangent plane.

Concavity-Convexity: graphical representation

Concave function



Convex function



The Hessian and curvature conditions

We can determine the curvature of a function using its Hessian:

Definition: A continuously differentiable function $f(X)$ is concave (convex) iff its Hessian is negative (positive) semi-definite at each point in X.

- $f(X)$ is strictly concave (convex) if its Hessian is negative (positive) definite.

The Hessian and curvature conditions

We can determine the curvature of a function using its Hessian:

Definition: A continuously differentiable function $f(X)$ is concave (convex) iff its Hessian is negative (positive) semi-definite at each point in X.

- $f(X)$ is strictly concave (convex) if its Hessian is negative (positive) definite.

Definition: A matrix is negative semidefinite if its principal minors alternate signs starting from a nonpositive.

- Similarly, it is positive semidefinite if its all principal minors are non-negative.

The Hessian and curvature conditions

We can determine the curvature of a function using its Hessian:

Definition: A continuously differentiable function $f(X)$ is concave (convex) iff its Hessian is negative (positive) semi-definite at each point in X.

- $f(X)$ is strictly concave (convex) if its Hessian is negative (positive) definite.

Definition: A matrix is negative semidefinite if its principal minors alternate signs starting from a nonpositive.

- Similarly, it is positive semidefinite if its all principal minors are non-negative.

Formally, suppose $H(x) = \begin{bmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{bmatrix}$ is the Hessian of $f(x_1, x_2)$.

- H is positive semidefinite if $f_{11} \geqslant 0$ and $f_{11}f_{22} - f_{12}f_{21} \geqslant 0$
- H is negative semidefinite if $f_{11} \leqslant 0$ and $f_{11}f_{22} - f_{12}f_{21} \geqslant 0$

Profit Maximization

Previously we set up the following profit maximization problem:

- $\max_{x} \pi(x) = pf(x) - wx$

Denote the solution to this optimization problem (the **decision rule**) as

- $x^*(p, w)$
- $y^*(p, w) = f(x^*(p, w))$

where $x^*(p, w)$ are the profit maximizing firm's unconditional input demand functions, and $y^*(p, w)$ is its supply function. It's important to note that the optimal quantities are functions of model parameters. Next, we seek to discuss the following:

- Derive $y^*(p, w)$ and $x^*(p, w)$
- Investigate the properties of $y^*(p, w)$ and $x^*(p, w)$
- Analyze how a profit maximizing firm's decision rules change under changing market conditions (comparative statics).

Profit Maximization: 2 input 1 output

Without loss of generality, let's focus on the 2 input case, $i = 1, 2$.

- $\max_{x_1, x_2} \pi = pf(x_1, x_2) - w_1 x_1 - w_2 x_2$

The FOCs for profit maximization are:

1. $\frac{\partial \pi}{\partial x_1} = \pi_1 = pf_1(x_1, x_2) - w_1 = 0$
    - $pf_1(x_1, x_2) = w_1$
2. $\frac{\partial \pi}{\partial x_2} = \pi_2 = pf_2(x_1, x_2) - w_2 = 0$
    - $pf_2(x_1, x_2) = w_2$

Profit Maximization: 2 input 1 output

1. $pf_1(x_1, x_2) = w_1$
2. $pf_2(x_1, x_2) = w_2$

Interpretation: The FOCs say that, for an interior solution, a profit maximizing firm sets the marginal contribution of each factor to revenues, pf_i, (the value of marginal product factor) equal to the marginal cost of that factor, w_i.

Implication:

- Profit maximization takes place at points only where marginal products f_i are positive (regardless of p and w).

Profit Maximization: 2 input 1 output

1. $pf_1(x_1, x_2) = w_1$
2. $pf_2(x_1, x_2) = w_2$

Interpretation: The FOCs say that, for an interior solution, a profit maximizing firm sets the marginal contribution of each factor to revenues, pf_i, (the value of marginal product factor) equal to the marginal cost of that factor, w_i.

Implication:

- Profit maximization takes place at points only where marginal products f_i are positive (regardless of p and w).

Also, note that by dividing (1) and (2) we obtain

- $\frac{f_1(x_1,x_2)}{f_2(x_1,x_2)} = \frac{w_1}{w_2}$

The above equality states that at the optimum a profit maximizing firm sets the marginal rate of technical substitution (MRTS = ratio of marginal products) equal to the input price ratio.

Profit Maximization: 2 input 1 output

The SOC for a maximum is that the Hessian has to be *negative semi-definite.* The second partials are:

- $\pi_{11} = pf_{11}$
- $\pi_{22} = pf_{22}$
- $\pi_{12} = pf_{12}$

Hence, the Hessian is:

- $H(x) = \begin{bmatrix} pf_{11} & pf_{12} \\ pf_{12} & pf_{22} \end{bmatrix}$

- H is negative semidefinite if $f_{11} \leqslant 0$ and $f_{11}f_{22} - f_{12}^2 \geqslant 0$
- From the SOC, it must be true that $f_{22} \leqslant 0$
- Note that if the production function is *concave*, the FOC is also sufficient for an optimum.

Implications of the SOC for Profit Maximization

1. $f_{11} \leqslant 0$ and $f_{22} \leqslant 0$
2. $f_{11} f_{22} - f_{12}^2 \geqslant 0$

The conditions in 1 state that a profit maximizing firm employs inputs at a level where technology exhibits *diminishing marginal returns* in each input.

Q: Why can't operating at a level with increasing marginal returns be profitable?

Implications of the SOC for Profit Maximization

1. $f_{11} \leqslant 0$ and $f_{22} \leqslant 0$
2. $f_{11}f_{22} - f_{12}^2 \geqslant 0$

The conditions in 1 state that a profit maximizing firm employs inputs at a level where technology exhibits *diminishing marginal returns* in each input.

Q: Why can't operating at a level with increasing marginal returns be profitable?

If hiring the first unit of input was profitable, without diminishing marginal returns the firm would hire that input without bound.

Condition 2 arises from the fact that factors can be dependent (hiring more of one factor affects the productivity of the other). The condition requires that the decrease in the own marginal productivity of an input from an additional unit cannot be overcompensated from productivity increases coming from cross effects.

Constrained Optimization

So far we have focused on unconstrained optimization

- This is easier - always make your problem into the unconstrained version if it can be readily done
- But sometimes we need to live with the constraint(s)
- In this case we set up a **Lagrangian**

The Lagrangian

Suppose $u = u(x, y)$ and consumers face a budget constraint, $p_x x + p_y y = I$

- First we set up the budget constraint as an equation equal to zero:
  - $g(x, y) = I - p_x x - p_y y = 0$
- Then we build the Lagrangian:
  - $L(x, y, \lambda) = u(x, y) + \lambda(I - p_x x - p_y y)$
- λ is known as the **Lagrange multiplier** for the budget constraint

The Lagrangian: FOCs

The FOCs for an optimum are found by differentiating the Lagrangian w.r.t. its three arguments and setting them equal to zero:

- $\frac{\partial L}{\partial x}|_{x=x^*,y=y^*} = 0$
- $\frac{\partial L}{\partial y}|_{x=x^*,y=y^*} = 0$
- $\frac{\partial L}{\partial \lambda}|_{x=x^*,y=y^*} = 0$

We can then solve for the optimal values, x^* and y^*.

Note that the last FOC just returns the budget constraint.

The Lagrangian: SOCs

The SOCs for a constrained optimum are summarized by the **bordered Hessian**:

$$H_b = \begin{vmatrix} L_{\lambda\lambda} & L_{\lambda x} & L_{\lambda y} \\ L_{x\lambda} & L_{xx} & L_{xy} \\ L_{y\lambda} & L_{yx} & L_{yy} \end{vmatrix}$$

This can be simplified because i) L is a C2 function so the second derivatives are all symmetric and ii) L_λ is just the budget constraint $g(x, y)$:

$$H_b = \begin{vmatrix} 0 & g_x & g_y \\ g_x & L_{xx} & L_{xy} \\ g_y & L_{xy} & L_{yy} \end{vmatrix}$$

The Lagrangian: SOSC

The SOSC for a local maximum is that the bordered Hessian must be negative definite. The SOSC for a local minimum is that the bordered Hessian must be positive definite.

- A matrix is **positive definite** if all its leading principal minors $|H_{bk}|$, except the first, are positive.
- A matrix is **negative definite** if all its leading principal minors $|H_{bk}|$, except the first, alternate in sign, with $|H_{b2}|$ being negative, $|H_{b3}|$ being positive, etc.
- $|\cdot|$ here is the determinant operator The kth leading **principal**

  minor of a matrix is the determinants of the submatrix formed by selecting its first k rows and columns

Cramer's Rule

We often set up optimization problems with several FOCs, and want to solve for the choice variables.

Cramer's Rule often greatly simplifies this process. Suppose we have a set of *n* equations with *n* unknowns:

$$\begin{aligned}
a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\
a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\
\vdots \qquad\qquad &\vdots \quad \vdots \\
a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n &= b_n
\end{aligned}$$

Cramer's Rule

We can write this in matrix form:

$$Ax = b$$

Cramer's Rule says that the solution for each x_i is given by

$$x_i = \frac{|A_i|}{|A|}$$

where A_i is the matrix formed by replacing column i of A with the vector b.

The implicit function theorem (IFT)

Sometimes solving the FOCs for the choice variables analytically is hard or impossible. But in many cases, what we care about is not the specific values of x^*, y^*, etc., but rather how they change as we shift the parameters - the *comparative statics*. We can often find these using the **implicit function theorem**, which is a very powerful result. For a single-variable optimization problem,

$$\text{If } G(x, \alpha) = 0 \text{ then } \frac{dx}{d\alpha} = -\frac{\partial G/\partial \alpha}{\partial G \partial x}.$$

Thus we can use the partial derivatives of the FOC to find the comparative statics for x, *without having to solve for* x.

The IFT with multiple variables

The IFT can be generalized to a case where there are multiple equations that have to be equal to zero. Suppose that

- $G_1(x, y, \beta, \alpha) = 0$
- $G_2(x, y, \beta, \alpha) = 0$

then

$$\frac{\partial x}{\partial \alpha} = -\frac{\begin{vmatrix} \frac{\partial G_1}{\partial \alpha} & \frac{\partial G_1}{\partial y} \\ \frac{\partial G_2}{\partial \alpha} & \frac{\partial G_2}{\partial y} \end{vmatrix}}{\begin{vmatrix} \frac{\partial G_1}{\partial x} & \frac{\partial G_1}{\partial y} \\ \frac{\partial G_2}{\partial x} & \frac{\partial G_2}{\partial y} \end{vmatrix}}$$

and likewise for $\frac{\partial y}{\partial \alpha}$.

The Envelope Theorem

The **envelope theorem** is a tool for simplifying the analysis of optimized functions.

In words, it says that if we are looking at changes in the maximized value of a function, we can ignore the direct effect of changes in the choice variable(s), because at an optimum the FOC must hold so the relevant partial derivative is equal to zero.

For some reason, confusing “explanations” of this theorem abound. I will be as clear as I can, and also see the supplementary PDF about this on Moodle.

Deriving the Envelope Theorem

Suppose we choose x to maximize the function $g(x, \alpha)$. The optimal choice of x^* is, in general, a function of α: $x^* = x^*(\alpha)$.

We want to know how the maximized value $M(\alpha) = g(x^*(\alpha), \alpha)$ changes when α changes. By the chain rule,

$$\frac{dM(\alpha)}{d\alpha} = \frac{\partial g(x^*(\alpha), \alpha)}{\partial x^*} \frac{\partial x^*(\alpha)}{\partial \alpha} + \frac{\partial g(x^*(\alpha), \alpha)}{\partial \alpha}$$

The FOC tells us we can ignore the first term.

This is the envelope theorem! Thus

$$\frac{dM(\alpha)}{d\alpha} = \frac{\partial g(x^*(\alpha), \alpha)}{\partial \alpha}$$

An example using the Envelope Theorem

A firm maximizes a one-input profit function, $\pi(L, p, w) = pf(L) - wL$. Its optimal value of L is a function of output prices p and wages w, $L^* = L^*(p, w)$. Our goal is to examine how profits vary with the price of output p.

$$\frac{d\pi(L, p, w)}{dp} = f(L(p, w)) + p\frac{\partial L^*(p, w)}{\partial p} - w\frac{\partial L^*(p, w)}{\partial p}$$

By the envelope theorem, terms two and three are zero because they are multiplied by the FOC. So we get a result known as Hotelling's Lemma:

$$\frac{d\pi(L, p, w)}{dp} = f(L(p, w))$$