Econ 205 - Slides from Lecture 10

Joel Sobel

September 2, 2010

Example

Find the tangent plane to $\{\mathbf{x} \mid x_1 x_2 - x_3^2 = 6\} \subset \mathbb{R}^3$ at $\widehat{\mathbf{x}} = (2, 5, 2)$.
If you let $f(\mathbf{x}) = x_1 x_2 - x_3^2$, then this is a level set of f for value 6.

$$\begin{aligned}\nabla f(\mathbf{x}) &= (\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \frac{\partial f}{\partial x_3}) \\ &= (x_2, x_1, -2x_3)\end{aligned}$$

$$\begin{aligned}\nabla f(\widehat{\mathbf{x}}) &= \nabla f(\mathbf{x}) \mid_{\mathbf{x}=(2,5,2)} \\ &= (5, 2, -4)\end{aligned}$$

Tangent Plane:

$$\begin{aligned}\{\widehat{\mathbf{x}} + \mathbf{y} \mid \mathbf{y} \cdot \nabla f(\widehat{\mathbf{x}}) = 0\} &= \{(2, 5, 2) + (y_1, y_2, y_3) \mid 5y_1 + 2y_2 - 4y_3 = 0\} \\ &= \{\mathbf{x} \mid 5x_1 - 10 + 2x_2 - 10 - 4x_3 + 8 = 0\} \\ &= \{5x_1 + 2x_2 - 4x_3 = 12\}\end{aligned}$$

Suppose $f(x, y, z) = 3x^2 + 2xy - z^2$.

- $\nabla f(x, y, z) = (6x + 2y, 2x, -2z)$.
- $f(2, 1, 3) = 7$.
- The level set of f when $f(x, y, z) = 7$ is $\{(x, y, z) : f(x, y, z) = 7\}$.
- This set is a (two dimensional) surface in $\mathbb{R}^3$: It can be written $F(x, y, z) = 0$ (for $F(x, y, z) = f(x, y, z) - 7$).
- The equation of the tangent to the level set of f is a (two-dimensional) hyperplane in $\mathbb{R}^3$.
- At the point $(2, 1, 3)$, the hyperplane has normal equal to $\nabla f(2, 1, 3) = (12, 4, -6)$.
- Hence the equation of the hyperplane to the level set at $(2, 1, 3)$ is equal to:

$$(12, 4, -6) \cdot (x - 2, y - 1, z - 3) = 0$$

or

$$12x + 4y - 6z = 10.$$

The graph of f is a three-dimensional subset of $\mathbb{R}^4$

$$\{(x, y, z, w) : w = f(x, y, z)\}.$$

A point on this surface is $(2, 1, 3, 7) = (x, y, z, w)$.
The tangent hyperplane at this point can be written:

$$w - 7 = \nabla f(2, 1, 3) \cdot (x - 2, y - 1, z - 3) = 12x + 4y - 6z - 10$$

or

$$12x + 4y - 6z - w = 3.$$

Homogeneous Functions

Definition

The function $F : \mathbb{R}^n \longrightarrow \mathbb{R}$ is homogeneous of degree k if $F(\lambda x) = \lambda^k F(x)$ for all λ.

Homogeneity of degree one is weaker than linearity:
All linear functions are homogeneous of degree one, but not conversely.
For example, $f(x, y) = \sqrt{xy}$ is homogeneous of degree one but not linear.

Theorem (Euler's Theorem)

If $F : \mathbb{R}^n \longrightarrow \mathbb{R}$ be a differential at x and homogeneous of degree k, then $\nabla F(x) \cdot x = kF(x)$.

Proof.

Fix x. Consider the function $H(\lambda) = F(\lambda x)$. This is a composite function, $H(\lambda) = F \circ G(\lambda)$, where $G : \mathbb{R} \longrightarrow \mathbb{R}^n$, such that $G(\lambda) = \lambda x$. By the chain rule, $DH(\lambda) = DF(G(\lambda))DG(\lambda)$. If we evaluate this when $\lambda = 1$ we have

$$DH(1) = \nabla F(x) \cdot x. \tag{1}$$

On the other hand, we know from homogeneity that $H(\lambda) = \lambda^k x$. Differentiating the right hand side of this equation yields $DH(\lambda) = k\lambda^{k-1}F(\lambda x)$ and evaluating when $\lambda = 1$ yields

$$DH(1) = kF(x). \tag{2}$$

Combining equations (1) and (2) yields the theorem. □

Comments

1. I A cost function depends on the wages you pay to workers. If all of the wages double, then the cost doubles. This is homogeneity of degree one.
2. A consumer's demand behavior is homogeneous of degree zero.
   Demand is a function $\phi(p, w)$ that gives the consumer's utility maximizing feasible demand given prices p and wealth w.
   The demand is the best affordable consumption for the consumer.
   The consumptions x that are affordable satisfy $p \cdot x \leq w$ (and possibly another constraint like non-negativity).
   If p and w are multiplied by the same factor, λ, then the budget constraint remains unchanged.
   Hence the demand function is homogeneous of degree zero.

Euler's Theorem provides a nice decomposition of a function F.
Suppose that F describes the profit produced by a team of n agents, when agent i contributes effort x_i.
How such the team divide the profit it generates?
If F is linear, the answer is easy: If $F(x) = p \cdot x$, then just give agent i $p_i x_i$.
Here you give each agent a constant "per unit" payment equal to the marginal contribution of her effort.
When F is non-linear, it is harder to figure out the contribution of each agent.
The theorem states that if you pay each agent her marginal contribution ($D_{e_i} f(x)$) per unit, then you distribute the surplus fully if F is homogeneous of degree one.

Higher-Order Derivatives

If $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ is differentiable, then its partial derivatives, $D_{e_i} f(x) = D_i f(x)$ can also be viewed as functions from $\mathbb{R}^n$ to $\mathbb{R}$. You can imagine taking a derivative with respect to one variable, then other then the first again, and so on, creating sequences of the form

$$D_{i_k} \cdots D_{i_2} D_{i_1} f(x).$$

Provided that all of the derivatives are continuous in a neighborhood of x, the order in which you take partial derivatives does not matter. We denote an kth derivative $D^k_{i_1,\ldots,i_n} f(x)$, where k is the total number of derivatives and i_j is the number of partial derivatives with respect to the jth argument (so each i_j is a non-negative integer and $\sum_{j=1}^n i_j = x$. Except for the statement of Taylor's Formula, we will have little interest in third or higher derivatives.

Second derivatives are important

A real-valued function of n variables will have n^2 second derivatives, which we sometimes think of as terms in a square matrix:

$$\underset{n\times n}{D^2 f(\mathbf{x})} = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(\mathbf{x}) & \dots & \frac{\partial^2 f}{\partial x_n \partial x_1}(\mathbf{x}) \\ \vdots & & \vdots \\ \frac{\partial^2 f}{\partial x_1 \partial x_n}(\mathbf{x}) & \dots & \frac{\partial^2 f}{\partial x_n^2}(\mathbf{x}) \end{pmatrix}$$

We say $f \in C^k$ if the kth order partials exist and are continuous.

Taylor Approximations

- It looks scary, but it really is a one-variable theorem.
- If you know about a function at the point $a \in \mathbb{R}^n$ and you want to approximate the function at x.
- Consider the function $F(t) = f(xt + a(1-t))$. $F : \mathbb{R} \longrightarrow \mathbb{R}$ and so you can use one-variable calculus.
- $F(1) = x$ and $F(0) = a$.
- If you want to know about the function f at x using information about the function f at a, you really just want to know about the one-variable function F at $t = 1$ knowing something about F at $t = 0$.
- Multivariable version of Taylor's Theorem: apply the one variable version of the theorem to F.
- The chain rule describes the derivatives of F (in terms of f) and there are a lot of these derivatives.

First Order Approximation

Consider $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ such that f is differentiable. At $\mathbf{a} \in \mathbb{R}^n$ the 1st degree Taylor Polynomial of f is

$$P_1(\mathbf{x}) \equiv f(\mathbf{a}) + \nabla f(\mathbf{a}) \cdot (\mathbf{x} - \mathbf{a})$$

The first-order approximation should be familiar. Notice that you have n "derivatives" (the partials).
If we write $f(x) = P_1(x, a) + E_2(x, a)$ for the first-order approximation with error of f at x around the point a, then we have

$$\lim_{\mathbf{x} \to \mathbf{a}} \frac{|E_2(\mathbf{x}, \mathbf{a})|}{||\mathbf{x} - \mathbf{a}||} = \lim_{\mathbf{x} \to \mathbf{a}} \frac{|f(\mathbf{x}) - f(\mathbf{a}) - Df(\mathbf{a}) \cdot (\mathbf{x} - \mathbf{a})|}{||\mathbf{x} - \mathbf{a}||}$$
$$= 0$$

Thus as before, as $\mathbf{x} \to \mathbf{a}$, E_2 converges to 0 faster than $\mathbf{x}$ to $\mathbf{a}$.

Second Order Approximation

If $f \in C^2$, the 2nd degree Taylor approximation is

$$f(\mathbf{x}) = \underbrace{f(\mathbf{a}) + \underset{1\times n}{\nabla f(\mathbf{a})}\underset{n\times 1}{(\mathbf{x} - \mathbf{a})} + \frac{1}{2}\underset{1\times n}{(\mathbf{x} - \mathbf{a})'}\underset{n\times n}{D^2 f(\mathbf{a})}\underset{n\times 1}{(\mathbf{x} - \mathbf{a})}}_{P_2(\mathbf{x},a)} + E_3(\mathbf{x}, \mathbf{a})$$

where

$$\frac{1}{2}\underbrace{\underset{1\times n}{(\mathbf{x} - \mathbf{a})'}\underset{n\times n}{D^2 f(\mathbf{a})}\underset{n\times 1}{(\mathbf{x} - \mathbf{a})}}_{1\times 1} = \frac{1}{2}\mathbf{W}$$

$$= \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}(x_i - a_i)\frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{a})(x_j - a_j)$$

where

$$W = (x_1 - a_1, \ldots, x_n - a_n) \cdot \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(\mathbf{a}) & \ldots & \frac{\partial^2 f}{\partial x_n \partial x_1}(\mathbf{a}) \\ \vdots & & \vdots \\ \frac{\partial^2 f}{\partial x_1 \partial x_n}(\mathbf{a}) & \ldots & \frac{\partial^2 f}{\partial x_n^2}(\mathbf{a}) \end{pmatrix} \cdot \begin{pmatrix} x_1 - a_1 \\ \vdots \\ x_n - a_n \end{pmatrix}$$

Notice that the error term is a quadratic form.

Notation

Define $D_h^k f$ to be a kth derivative:

$$D_h^k f = \sum_{j_1+\cdots+j_n=k} \binom{k}{j_1 \cdots j_n} h_1^{j_1} \cdots h_n^{j_n} D_1^{j_1} \cdots D_n^{j_n} f,$$

where the summation is taken over all n-tuples of $j_1, \ldots, j_n$ of non-negative integers that sum to k and the symbol

$$\binom{k}{j_1 \cdots j_n} = \frac{k!}{j_1! \cdots j_n!}.$$

General Form of Taylor's Theorem

Theorem (Taylor's Theorem)

If f is a real-valued function in C^{k+1} defined on an open set containing the line segment connecting a to x, then there exists a point η on the segment such that

$$f(x) = P_k(x, a) + E_k(x, a)$$

where $P_k(x, a)$ is a kth order Taylor's Approximation:

$$P_k(x, a) = \sum_{r=0}^{k} \frac{D^r_{x-a}(a)}{r!}$$

and $E_k(x, a)$ is the error term:

$$E_k(x, a) = \frac{D^{k+1}_{x-a}(\eta)}{(k+1)!}.$$

Coda

Moveover, the error term satisfies:

$$\lim_{\mathbf{h}\to\mathbf{0}} \frac{E_k(\mathbf{a}+\mathbf{h},\mathbf{a})}{||\mathbf{h}||^k} = 0$$

Inverse and Implicit Functions: Introduction

Theme:
Whatever you know about linear functions is true *locally* about differentiable functions.

Inverse Functions: Review

1. One variable, linear case: $f(x) = ax$. Invertible if and only if $a \neq 0$.
2. General:

Definition

We say the function $f\colon X \longrightarrow Y$ is one-to-one if

$$f(x) = f(x') \quad \Longrightarrow \quad x = x'$$

Recall $f^{-1}(S) = \{x \in X \mid f(x) \in S\}$ for $S \subset Y$.
Now let's consider $f^{-1}(y)$ for $y \in Y$. Is f^{-1} a function?
If f is one-to-one, then $f^{-1}\colon f(X) \longrightarrow X$ is a function.

3. f is generally not invertible as the inverse is not one-to-one. But in the neighborhood (circle around a point x_0), it may be strictly increasing so it is one-to-one locally and therefore locally invertible.

Theorem

If $f : \mathbb{R} \longrightarrow \mathbb{R}$ *is* C^1 *and*

$$f'(x_0) \neq 0,$$

then $\exists\, \varepsilon > 0$ *such that* f *is strictly monotone on the open interval* $(x_0 - \varepsilon, x_0 + \varepsilon)$.

Locally

Definition

The function $f : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ is locally invertible at x_0 if there is a $\delta > 0$ and a function $g : B_\delta(f(x_0)) \longrightarrow \mathbb{R}^n$ such that $f \circ g(y) \equiv y$ for $y \in B_\delta(f(x_0))$ and $g \circ f(x) \equiv x$ for $x \in B_\delta(x_0)$.

So f is locally invertible at x_0, then we can define g on $(x_0 - \varepsilon, x_0 + \varepsilon)$ such that

$$g(f(x)) = x$$

In the one-variable case, a linear function has (constant) derivative. When derivative is non zero and not equal to zero, the function is invertible (globally).
Differentiable functions with derivative not equal to zero at a point are invertible locally.
For one variable functions, if the derivative is always non zero and continuous, then the inverse can be defined on the entire range of the function.

Higher Dimensions

- When is $f : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ invertible?
- Linear functions can be represented as multiplication by a square matrix.
- Invertibility of the function is equivalent to inverting the matrix.
- A linear function is invertible (globally) if its matrix representation is invertible.

Inverse Function Theorem

Theorem

If $f : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ is differentiable at x_0 and $Df(x_0)$ is invertible, then f is locally invertible at x_0. Moreover, the inverse function, g is differentiable at $f(x_0)$ and $Dg(f(x_0)) = (Df(x_0))^{-1}$.

The theorem asserts that if the linear approximation of a function is invertible, then the function is invertible locally.
Unlike the one variable case, the assumption that Df is globally invertible does not imply the existence of a global inverse.

One Variable Again

$$g'(y^0) = \frac{1}{f'(x^0)}$$

so the formula for the derivative of the inverse generalizes the one-variable formula.

More General Problem

Given $G : \mathbb{R}^{m+n} \longrightarrow \mathbb{R}^n$. Let $x \in \mathbb{R}^m$ and $y \in \mathbb{R}^n$.
Can we solve the system of equations: $G(x, y) = 0$?
n equations in $n + m$ variables.
Search for a solution to the equation that gives x as a function of y.
The problem of finding an inverse is really a special case where $n = m$ and $G(x, y) = f(x) - y$.

Who Cares?

1. Equations characterize an economic equilibrium (market clearing price; first-order condition)
2. y variables are parameters.
3. "Solve" a model for a fixed value of the parameters.
4. What happens when parameters change?
5. The implicit function theorem says (under a certain condition), if you can solve the system at a give y_0, then you can solve the system in a neighborhood of y_0. Furthermore, it gives you expressions for the derivatives of the solution function.

Why Call it Implicit Function Theorem?

- If you could write down the system of equations and solve them to get an explicit representation of the solution function, great.
- ... explicit solution and a formula for solution for derivatives.
- Life is not always so easy.
- IFT assumes existence of solution and describes "sensitivity" properties even when explicit formula is not available.

Examples

$$f : \mathbb{R}^2 \longrightarrow \mathbb{R}$$

Suppose $f(x, z) = 0$ is an *identity* relating x and z.
How does z depends on x?

$$f(x, z) = x^3 - z = 0$$

Explicit solution possible.

$$x^2 z - z^2 + \sin x \log z + \cos x = 0,$$

then there is no explicit formula for z in terms of x.

Lower-order case

If solving $f(x, z) = 0$ gives us a function

$$g\colon (x_0 - \varepsilon, x_0 + \varepsilon) \longrightarrow (z_0 - \varepsilon, z_0 + \varepsilon)$$

such that

$$f(x, g(x)) = 0,$$

$\forall\, x \in (x_0 - \varepsilon, x_0 + \varepsilon)$
then we can define

$$h\colon (x_0 - \varepsilon, x_0 + \varepsilon) \longrightarrow \mathbb{R}$$

by

$$h(x) = f(x, g(x))$$

$h(x) = 0$ for an interval, then $h'(x) = 0$ for any x in the interval.

Also . . .

$$h'(x) = D_1 f(x, g(x)) + D_2 f(x, g(x)) Dg(x)$$

Since this expression is zero, it yields a formula for $Dg(x)$ provided that $D_2 f(x, g(x)) \neq 0$.

Explicit Calculation of Implicit Function Theorem Formula

If f and g are differentiable, we calculate

$$y = \begin{pmatrix} x \\ z \end{pmatrix}, \quad G(x) = \begin{pmatrix} x \\ g(x) \end{pmatrix}$$

$$h(x) = [f \circ G](x)$$

$$\begin{aligned} h'(x) &= Df(x_0, z_0) DG(x^0) \\ &= \left(\frac{\partial f}{\partial x}(x_0, z_0), \ \frac{\partial f}{\partial z}(x_0, z_0) \right) \cdot \begin{pmatrix} 1 \\ g'(x^0) \end{pmatrix} \\ &= \frac{\partial f}{\partial x}(x_0, z_0) + \frac{\partial f}{\partial z}(x_0, z_0) g'(x_0) \\ &= 0 \end{aligned}$$

This gives us

$$g'(x_0) = -\frac{\frac{\partial f}{\partial x}(x_0, z_0)}{\frac{\partial f}{\partial z}(x_0, z_0)}$$

More Generally

$$f\colon \mathbb{R} \times \mathbb{R}^m \longrightarrow \mathbb{R}^m$$

$$f(x, \mathbf{z}) = \begin{pmatrix} f_1(x, \mathbf{z}) \\ f_2(x, \mathbf{z}) \\ \vdots \\ f_m(x, \mathbf{z}) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \mathbf{0}$$

where $x \in \mathbb{R}$, and $\mathbf{z} \in \mathbb{R}^m$.
And we have

$$g\colon \mathbb{R} \longrightarrow \mathbb{R}^m$$

$$\mathbf{z} = g(x)$$

IFT

Theorem

Suppose

$$f: \mathbb{R} \times \mathbb{R}^m \longrightarrow \mathbb{R}^m$$

is C^1 *and write* $F(x, \mathbf{z})$ *where* $x \in \mathbb{R}$ *and* $\mathbf{z} \in \mathbb{R}^m$. *Assume*

$$|D_z f(x_0, \mathbf{z}_0)| = \begin{vmatrix} \frac{\partial f_1}{\partial z_1} & \cdots & \frac{\partial f_1}{\partial z_m} \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial z_1} & \cdots & \frac{\partial f_m}{\partial z_m} \end{vmatrix} \neq 0$$

and

$$f(x_0, \mathbf{z}_0) = \mathbf{0}.$$

There exists a neighborhood of (x_0, z_0) *and a function* $g: \mathbb{R} \longrightarrow \mathbb{R}^m$ *defined on the neighborhood of* x_0, *such that* $\mathbf{z} = g(x)$ *uniquely solves* $f(x, \mathbf{z}) = \mathbf{0}$ *on this neighborhood.*

This didn't fit

Furthermore the derivatives of g are given by

$$\underset{m\times 1}{Dg(x_0)} = -[\underset{m\times m}{D_z f(x_0, \mathbf{z}_0)}]^{-1} \underset{m\times 1}{D_x f(x_0, \mathbf{z}_0)}$$

Comments

1. This is hard to prove.
2. A standard proof uses a technique (contraction mappings) that you'll see in macro.
3. Hard part: Existence of the function g that gives z in terms of x.
4. Computing the derivatives of g is a simple application of the chain rule.

The Easy Part

$$f(x, g(x)) = 0$$

And we define

$$H(x) \equiv f(x, g(x))$$

And thus

$$\begin{aligned} D_x H(x) &= D_x f(x_0, \mathbf{z}_0) + D_z f(x_0, \mathbf{z}_0) D_x g(x_0) \\ &= \mathbf{0} \end{aligned}$$

$$\implies D_z f(x_0, \mathbf{z}_0) D_x g(x_0) = -D_x f(x_0, \mathbf{z}_0)$$

Multiply both sides by the inverse:

$$\underbrace{[D_z f(x_0, \mathbf{z}_0)]^{-1} \cdot [D_z f(x_0, \mathbf{z}_0)]}_{=\mathbf{I}_m} \cdot D_x g(x_0) = -[D_z f(x_0, \mathbf{z}_0)]^{-1} D_x F(x_0, \mathbf{z}_0)$$

$$\implies D_x g(x_0) = -[D_z f(x_0, \mathbf{z}_0)]^{-1} D_x F(x_0, \mathbf{z}_0)$$

Reminders

1. The implicit function theorem thus gives you a guarantee that you can (locally) solve a system of equations in terms of parameters.
2. Theorem is a local version of a result about linear systems.
3. Above, only one one parameter.
4. In general, parameters $x \in \mathbb{R}^n$ rather than $x \in \mathbb{R}$.

IFT - the real thing

Theorem

Suppose

$$f : \mathbb{R}^n \times \mathbb{R}^m \longrightarrow \mathbb{R}^m$$

is C^1 and write $F(\mathbf{x}, \mathbf{z})$ where $\mathbf{x} \in \mathbb{R}^n$ and $\mathbf{z} \in \mathbb{R}^m$.
Assume

$$|D_z f(\mathbf{x}_0, \mathbf{z}_0)| = \begin{vmatrix} \frac{\partial f_1}{\partial z_1} & \cdots & \frac{\partial f_1}{\partial z_m} \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial z_1} & \cdots & \frac{\partial f_m}{\partial z_m} \end{vmatrix} \neq 0$$

and

$$f(\mathbf{x}_0, \mathbf{z}_0) = \mathbf{0}.$$

There exists a neighborhood of $(\mathbf{x}_0, \mathbf{z}_0)$ and a function $g : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ defined on the neighborhood of $\mathbf{x}_0$, such that $\mathbf{z} = g(\mathbf{x})$ uniquely solves $f(\mathbf{x}, \mathbf{z}) = \mathbf{0}$ on this neighborhood.

Still Doesn't Fit

Furthermore the derivatives of g are given by implicit differentiation (use chain rule)

$$\underset{m\times n}{Dg(\mathbf{x}_0)} = -[\underset{m\times m}{D_z f(\mathbf{x}_0, \mathbf{z}_0)}]^{-1} \underset{m\times n}{D_x f(\mathbf{x}_0, \mathbf{z}_0)}$$

More Comments

1. Verifying the formula is just the chain rule.
2. Keep track of the dimensions of the various matrices.
3. Keep in mind the intuitive idea that you usually need exactly m variables to solve m equations is helpful.
4. This means that if the domain has n extra dimensions, typically you will have n parameters – the solution function will go from $\mathbb{R}^n$ into $\mathbb{R}^m$.
5. The implicit function theorem proves that a system of equations has a solution **if** you already know that a solution exists at a point.
6. Repeat: Theorem says: **If** you can solve the system once, then you can solve it locally.
7. Theorem does not guarantee existence of a solution.
8. In this respect linear case is special.
9. The theorem provides an explicit formula for the derivatives of the implicit function. Don't memorize it. Compute the derivatives of the implicit function by "implicitly differentiating" the system of equations.

Example

A monopolist produces a single output to be sold in a single market. The cost to produce q units is $C(q) = q + .5q^2$ dollars and the monopolist can sell q units for the price of $P(q) = 4 - \frac{q^5}{6}$ dollars per unit. The monopolist must pay a tax of one dollar per unit sold.

1. Show that the output $q^* = 1$ that maximizes profit (revenue minus tax payments minus production cost).
2. How does the monopolist's output change when the tax rate changes by a small amount?

The monopolist picks q to maximize:

$$q(4 - \frac{q^5}{6}) - tq - q - .5q^2.$$

The first-order condition is

$$q^5 + q - 3 + t = 0$$

and the second derivative of the objective function is $-5q^4 - 1 < 0$. Conclude: at most one solution to this equation; solution must be a (global) maximum.

Details

- Plug in $q = 1$ to see that this value does satisfy the first-order condition when $t = 1$.
- How the solution $q(t)$ to:

$$q^5 + q - 3 + t = 0$$

varies as a function of t when t is close to one?

- We know that $q(1) = 1$ satisfies the equation.
- LHS of the equation is increasing in q, so IFT holds.
- Differentiation yields:

$$q'(t) = -\frac{1}{5q^4 + 1}. \tag{3}$$

In particular, $q'(1) = -\frac{1}{6}$.

- To obtain the equation for $q'(t)$ you could use the general formula or differentiate the identity:

$$q(t)^5 + q(t) - 3 + t \equiv 0$$

  with respect to one variable (t) to obtain

$$5q(t)q'(t) + q'(t) + 1 = 0,$$

  and solve for $q'(t)$.
- Equation is linear in q'.

Implicit Differentiation

- This technique of "implicit differentiation" is fully general.
- In the example you have $n = m = 1$ so there is just one equation and one derivative to find.
- In general, you will have an identity in n variables and m equations.
- If you differentiate each of the equations with respect to a fixed parameter, you will get m linear equations for the derivatives of the m implicit functions with respect to that variable.
- The system will have a solution if the invertibility condition in the theorem is true.