1. Comparative Statics

1.1. The Implicit Function Theorem. A useful feature of the theory of nonlinear programming is that via the First Order Necessary Conditions and Complementary Slackness, etc, we can often compute explicitly solutions to optimization problems in $n$ variables with $m$ constraints by solving the system of equations

$$\frac{\partial f(x^*)}{\partial x_i} + \sum_{j=1}^{m} \lambda_j^* \cdot \frac{\partial g_j(x^*)}{\partial x_i} \leq 0$$

$$\left( \frac{\partial f(x^*)}{\partial x_i} + \sum_{j=1}^{m} \lambda_j^* \cdot \frac{\partial g_j(x^*)}{\partial x_i} \right) x_i^* = 0$$

for all $i = 1, 2, \ldots, n$ and

$$\frac{\partial L(x^*, \lambda^*)}{\partial \lambda_j^*} = g_j(x^*) \geq 0$$

$$\lambda_j^* \cdot g_j(x^*) = 0$$

for all $j = 1, 2, \ldots, m$.

There are a variety of tools we can use to ensure that a solution exists. For example, the Weierstrass Theorem tells us that a continuous function achieves its maximum and minimum over a nonempty compact set. Thus, if we know that $f$ is $C^1$ and that the set $C = \{ x : G(x) \geq 0 \}$ is compact, we know that a solution exists. Alternatively, if we could express the First Order Necessary Conditions as equalities, we could apply a fixed point theorem or theorems from algebra to ensure a solution. (See later.)

Frequently, we want to learn even more. A typical economic problem is interested in learning how the solution to a constrained optimization problem changes with respect to changes in exogenous parameters of the problem. For example:

- How does a consumer's demand change with a change in price or income?
- How does the supply of output change with a change in the price of an input?
- How does the price of a product change with changes in the characteristics of a population?

Date: Summer 2013.

Notes compiled on August 7, 2013
Consider the parametrized problem

$$\max_{x \geq 0} f(x; q)$$

s.t. $G(x; q) \geq 0$

where $q$ is some $k$ dimensional vector of exogenous real numbers. Call a solution to this $x(q)$ and the value the solution attains, $V(q) = f(x(q); q)$.

(Note that $x(q)$ may not be unique but $V(q)$ is still well-defined, Why?) We are interested in knowing how $V$ and $x$ change with $q$, the exogenous parameters of the problem. (Note: $q$ is called 'exogenous' because these variables are not choice variables. They are treated as fixed and given by the decision-maker. The fact that the decision maker treats them as fixed and given does not mean that they never change. In fact, it is exactly the effects of changes in these variables that we are interested in.

**Remark.** The parametrized problem above really describes a family of optimization problems - each different value of the vector, $q$, yields a different member of the family - that is, a different optimization problem.

The first order necessary conditions from the Kuhn-Tucker Theorem indicates that it will often be useful to be able to solve more generally systems of equations of the form

$$T : \mathbb{R}^{k+p} \rightarrow \mathbb{R}^k; T(z, q) = 0;$$

In order to generate equations

$$z : \mathbb{R}^p \rightarrow \mathbb{R}^k; z(q)$$

The problem of finding an equation for a level set (that is, find $z(q)$ such that $T(z(q), q) = 0$) is an example of this type of problem. These are called implicit functions because rather than defining $z$ explicitly we are defining it implicitly via $T$.

Notice that it is not always possible to find such functions: Consider the simple system of a single equation,

$$q^2 + z^2 - 1 = 0$$

In this case we can solve explicitly for $z$. However we get $z = \pm \sqrt{1 - q^2}$

This is not a function in general because for any given $q$ there are more than one $z$ that solves the equation.

The role of the *implicit function theorem* is to tell us when it is possible to discover an implicit function from a system of equations.

**Theorem 1.** *(Implicit Function Theorem (15.7 SB)).* Let $T : \mathbb{R}^{k+p} \rightarrow \mathbb{R}^k$ be $C^1$ Suppose that $T(z^*, q^*) = 0$. If the $k \times k$ matrix formed by collecting the $k$ gradient vectors (w.r.t. $z$), $\nabla_z T(z^*, q^*)$, is invertible (or equivalently has full rank or is non-singular), then there exists $k C^1$ functions each mapping $\mathbb{R}^p \rightarrow \mathbb{R}^k$ such that

$$z_1(q^*) = z_1^*, z_2(q^*) = z_2, \ldots, z_k(q^*) = z_k^*$$

and

$$T(z(q); q) = 0 \text{ for all } q \in B_\epsilon(q^*), \text{ for some } \epsilon > 0$$
With the first-order effects of \( q \) on \( z \) being
\[
D_q z(q^*) = - [D_z T(z^*; q^*)]^{-1} D_q T(z^*; q^*)
\]

Remark. If inverting the \( D_z T(z^*; q^*) \) matrix is difficult or you only need to calculate a small number of effects, the above equation can be simplified to find \( dz_i/dq_j \)
\[
D_z T(z^*; q^*) \cdot D_q T(z^*; q^*) = - D_q T(z^*; q^*)
\]

And then we can use Cramer’s Rule (if \( Ax = b \) then \( x_i = |A|^{-1} A_i \) where \( A_i \) is \( A \) but with the \( i \)th column replaced with \( b \)) to get \( dz_i/dq_j \), the \( i \)th element of \( D_q z(q^*) \).

As an example, consider the classical utility maximization problem:
\[
\max_{x \in \mathbb{R}^n} U(x) \\
\text{s.t. } p \cdot x \leq I
\]

Where \( U \) is strictly quasi-concave and
\[
\nabla U(x) > 0 \\
\frac{\partial U(0)}{\partial x_i} = \infty
\]

We know that a solution to this problem satisfies \( x_i > 0, I - p \cdot x = 0 \) (why?) and satisfies the first order conditions
\[
\begin{align*}
\frac{\partial U(x)}{\partial x_1} - \lambda p_1 &= 0 \\
\frac{\partial U(x)}{\partial x_2} - \lambda p_2 &= 0 \\
&\vdots \\
\frac{\partial U(x)}{\partial x_n} - \lambda p_n &= 0 \\
I - px &= 0
\end{align*}
\]

This system of equations can be thought of as a mapping from \( \mathbb{R}^{2n+2} \to \mathbb{R}^{n+1} \) (that is, the mapping from the space in which the vector \((x, \lambda, p, I)\) live to the space of one half that dimensionality because we have \( n + 1 \) equations above.) Thus, to apply the Implicit Function Theorem, set \( z = (x, \lambda), q = (p, I) \). In other words we can create a function \( T : \mathbb{R}^{2n+2} \to \mathbb{R}^{n+1} \) given by
\[
T(x, \lambda, p, I) = \begin{pmatrix} \nabla U(x) - \lambda p = 0 \\
I - p \cdot x = 0 \end{pmatrix}
\]

If this created function \( T \) is \( C^1 \) and if the \((n+1) \times (n+1)\) matrix of derivatives of this function (with respect to \((x, \lambda)\) is invertible (draw what this matrix looks like), then by the Implicit Function Theorem, we know that there exist \( n + 1 \) \( C^1 \) functions,
\[
x_1(p, I), x_2(p, I), \ldots, x_n(p, I), \lambda(p, I)
\]
such that
\[
T(x_1(p, I), x_2(p, I), \ldots, x_n(p, I), \lambda(p, I), p, I) = 0
\]

for all \( p, I \) in a neighborhood of a given price income vector, \((p, I)\).
Because the $x$s satisfy the first order conditions of utility maximization (why) they are candidates for consumer demand. That is, we can determine the existence of continuously differentiable (Marshallian) consumer demand functions.

1.2. **Theorem of the Maximum.** There are more general versions of this theorem, but we can show this version using the implicit function theorem.

Consider the family of lagrangian problems

$$V(q) = \max_{x} f(x; q) \quad (CP_L)$$

**s.t.** $G(x; q) = 0$

This can be generalized to the KT in fairly obvious ways by restricting attention only to the constraints that are binding at a given solution. Define the function

$$T(x, \lambda; q) = \begin{bmatrix} \nabla f(x; q) + \lambda \cdot \nabla G(x; q) \\ G(x; q) \end{bmatrix}$$

The First Order Necessary Conditions at an optimum are represented by $T(x^*; \lambda^*; q) = 0$. We want to know about defining the solutions to the problem, $x^*(q), \lambda^*(q)$ as functions of $q$. The Implicit Function Theorem already tells us when we can. If the $(n+m+p) \times (n+m+p)$ matrix constructed by taking the derivative of $T$ with respect to $x$ and $\lambda$ is invertible, then we can find $C^1$ functions, $x^*(q), \lambda^*(q)$ such that $T(x^*(q), \lambda^*(q); q) = 0$ for $q$ in a neighborhood of $q^*$. That is, we need the matrix

$$\nabla T(x, \lambda; q) = \begin{bmatrix} \nabla^2 f(x; q) + \lambda \cdot \nabla^2 G(x; q) \\ \nabla G(x; q) \end{bmatrix}$$

to have full rank.

**Theorem 2.** (Theorem of the Maximum). Suppose $(CP_L)$ satisfies the conditions of the Implicit Function Theorem at $(x^*(q^*); q^*)$. If $f$ is $C^1$ at $(x^*(q^*), q^*)$, then $V(q)$ is $C^1$ at $q^*$.

This result extends fairly straightforwardly to Kuhn-Tucker problems.

1.3. **The Envelope Theorem.** If we can apply the Implicit Function Theorem to our First Order Necessary Conditions then we know that an $x(q)$ which solves the First Order Necessary Conditions exists and is $C^1$ and therefore $V$ is $C^1$. Now we are interested in learning how these variables change with parameters of the problem.

For this discussion, I am going to assume the problem is:

$$v(q) = \max_{x} f(x; q) \quad (CP_{KT})$$

**s.t.** $G(x; q) \geq 0$

$G : \mathbb{R}^n \rightarrow \mathbb{R}^m$

The parameter $q$ is a $p$-dimensional vector of exogenous variables. I assume that, at the solution, the First Order Necessary Conditions hold with equality and can ignore the Complementary Slackness Conditions. The Envelope Theorem holds more generally for Kuhn-Tucker problems, but the exposition is clearer in this case. Suppose that the problem is well behaved so we have that at a particular value, $q^*$, the solution, $x(q^*), \lambda(q^*)$ are $C^1$ and $V(q^*) = f(x(q^*); q^*)$ is $C^1$ in $q$. The envelope theorem tells us how $V$ changes.
Theorem 3. (Envelope Theorem (SB 19.5)). Suppose \((CP KT)\) satisfies the conditions of the Implicit Function Theorem at \(x^*(q^*), q^*\). If \(f\) is \(C^1\) at \(x^*(q^*), q^*\),
\[
\frac{df(x(q^*); q^*)}{dq} = \frac{\partial L(x(q^*), \lambda(q^*); q^*)}{\partial q}
\]

Remark. To determine how the value function changes with the parameter \(q\), all we need to do is see how the objective function and the constraint functions change with \(q\) directly. We do not need to include the impact of changes in the optimization variables, \(x\) and \(\lambda\). Crudely speaking, because we have already optimized \(L(x, \lambda, q)\) with respect to \(x\) and \(\lambda\), the first order conditions are satisfied so when we differentiate with respect to these variables, there is no first order effect on \(L\). The proof formalizes this logic.

Proof. (For the case where all constraints bind and \(x^* > 0\) Define
\[
\nabla_x f(x; q) = \frac{\partial f(x; q)}{\partial x}
\]
\[
\nabla_q f(x; q) = \frac{\partial f(x; q)}{\partial q}
\]
\[
\nabla_q x(q) = \frac{\partial x(q)}{\partial q}
\]
These are \(n \times 1\), \(p \times 1\) vectors and \(p \times n\) matrix respectively.

The First Order Necessary Conditions imply that
\[
(1.1) \quad \nabla_x f(x(q); q) = -\nabla_x G(x(q); q) \cdot \lambda
\]
\[
(1.2) \quad G(x(q); q) = 0
\]
for all \(q\). Note that the dimensionality of Equation \(1.1\) is \(n \times 1\) on the left side and \((n \times m)(m \times 1) = n \times 1\) on the right side. Totally differentiate Equation \(1.2\) with respect to \(q\). Since this equality holds for all \(q\), the total derivative stays zero for all \(q\). Thus,
\[
\nabla_q x(q) \cdot \nabla_x G(x(q); q) + \nabla_q G(x(q); q) = 0
\]
\[
\nabla_q x(q) \cdot \nabla_x G(x(q); q) = -\nabla_q G(x(q); q)
\]
Post multiply each side of the last equation by the \(m \times 1\) vector, \(\lambda\). This yields
\[
(1.3) \quad \nabla_q x(q) \cdot \nabla_x G(x(q); q) \cdot \lambda = -\nabla_q G(x(q); q) \cdot \lambda
\]
The left side is the product of a \(p \times n\) matrix, a \(n \times m\) matrix and a \(m \times 1\) vector respectively. The right side is the product of a \(p \times m\) matrix and a \(m \times 1\) vector. The first Equation (Equation \(1.1\)) from the First Order Necessary Conditions for \(f\) can be premultiplied by \(\nabla_q x(q)\) to get
\[
\nabla_q x(q) \cdot \nabla_x f(x(q); q) = -\nabla_q x(q) \cdot \nabla_x G(x(q); q) \cdot \lambda
\]
Now use Equation \(1.3\) above to get
\[
\nabla_q x(q) \cdot \nabla_x f(x(q); q) = \nabla_q G(x(q); q) \cdot \lambda
\]
By Definition, \(V(q) = f(x(q); q)\) so,
\[
\frac{dV(q)}{dq} = \nabla_q G(x(q); q) \cdot \lambda + \nabla_q f(x(q); q) \\
= \nabla_q \mathcal{L}(x, \lambda; q)|_{x=x(q), \lambda=\lambda(q)}
\]

Note that in words, the second equality is that the partial derivative of the Lagrangian, with respect to \( q \) evaluated at \( x = x(q), \lambda = \lambda(q) \), that is, at the optimal \( x \) and \( \lambda \). □

This result is useful because it allows us to see how the value function changes simply by looking at how \( q \) affects \( f \) and \( g \) directly. That is, we do not need to see how \( q \) affects \( f \) and \( g \) indirectly via changes in \( x \).

**Example.** Consider a general problem of the form,

\[
\max_x f(x) \\
\text{s.t. } q - g(x) \geq 0
\]

Thus, the parameter \( q \) as it gets bigger, makes it easier to satisfy the constraint. What would we gain from a small increase in \( q \) and thus a slight relaxation of the constraint? The Lagrangian is

\[
\mathcal{L}(x, \lambda, w) = f(x) + \lambda(q - g(x))
\]

The partial derivative of the Lagrangian with respect to \( q \) is \( \lambda \). Thus,

\[
\frac{\partial V(q)}{\partial q} = \lambda
\]

A small increase in \( q \) increases the value by approximately \( \lambda \). Thus the lagrange multiplier on the constraint is often called a shadow price. It describes the price of relaxing the constraint. Notice what happens if the constraint does not bind, \( \lambda = 0 \), and nothing happens to the constraint.

**Exercise.** Use this to show that in the consumer maximization problem, \( \lambda \) is the marginal utility of income.

**Example.** Consider the cost minimization problem,

\[
C(y, w) = \min_x w \cdot x \\
\text{s.t. } f(x) - y \geq 0
\]

The Lagrangian is

\[
\mathcal{L}(x, \lambda, w) = -w \cdot x + \lambda(f(x) - y)
\]

Denote the optimal solution to this problem to be \( x(y, w) \in \mathbb{R}^n \). Note that because of the transformation to a maximization problem in the Lagrangian, we must have

\[
\frac{\partial C(y, w)}{\partial w_i} = -\frac{\partial \mathcal{L}(x, \lambda, y, w)}{\partial w_i} = x_i(y, w)
\]

The partial derivative of the cost function with respect to \( w_i \) is \( x_i \) which is demand for factor \( i \). This result is known as Shephard’s lemma. Similarly, note that, by Young’s Theorem,

\[
\frac{\partial^2 C(y, w)}{\partial w_i \partial w_j} = \frac{\partial}{\partial w_j} x_i(y, w) = \frac{\partial}{\partial w_i} x_j(y, w) = \frac{\partial^2 C(y, w)}{\partial w_j \partial w_i}
\]
the change in demand for factor $i$ with respect to a small change in the price of factor $j$ equals the change in demand for factor $j$ in response to a small change in the price of factor $i$.

1.4. Correspondences and Fixed Point Theorems. Correspondences: In many parametrized optimization problems, it is natural to consider a generalization of the idea of a function.

**Definition 4.** (Correspondence). A correspondence is a transformation that maps a vector space into collections of subsets in another vector space.

For example, a correspondence $F : \mathbb{R}^n \rightarrow \mathbb{R}$ takes any $n$ dimensional vector and gives as its output a subset of $\mathbb{R}$. We focus on real-valued correspondences but, of course, there are many.

Motivation: There are a variety of ways that correspondences arise in economics. For example, the correspondence $g : \mathbb{R}^{n+1} \rightarrow \mathbb{R}^n$ given by $g(p, I) = \{x : p \cdot x \leq I\}$ gives the set of points in $\mathbb{R}^n$ for every vector of prices and income which are feasible for a consumer attempting to buy goods to maximize her utility. Also, the correspondence,

$$x(p, I) = \{x \in \mathbb{R}^n : x = \arg \max_{x} U(x), \text{ s.t. } p \cdot x \leq I\}$$

may yield a set of optimal consumption points rather than a single bundle if the consumers indifference curves are not strictly convex.

**Definition 5.** (Bounded Correspondence). A correspondence $F$ is bounded if for all $x$, and for all $y \in F(x)$, the size of $y$ is bounded. That is, $\|y\| \leq M$ for some number $M$.

In what follows, I restrict attention to correspondences that are bounded.

**Definition 6.** (Convex-Valued Correspondence). A correspondence $F$ is convex-valued if for all $x$, $F(x)$ is a convex set.

(Note, in a trivial sense, a function is a correspondence, and all functions are obviously convex-valued correspondences.)

**Definition 7.** (Upper Hemi-Continuous Correspondence). A correspondence $F$ is upper hemi-continuous at a point $x$, if for all sequences $\{x_n\}$ that converge to $x$ and all sequences $\{y_n\}$ such that $y_n \in F(x_n)$ that converge to $y$, then $y \in F(x)$.

Note: A bounded correspondences is upper hemi-continuous for all $x$ iff its graph (the set of points $\{(x, F(x))\}$) is a closed set.

**Definition 8.** (Lower Hemi-Continuous Correspondence). A correspondence $F$ is lower hemi-continuous at a point $x$, if for all sequences $\{x_n\}$ that converge to $x$ and for all $y \in F(x)$, there exists a sequence $\{y_n\}$ such that $y_n \in F(x_n)$ and the sequence converges to $y$.

Note: If a correspondence has a graph that is an open set then it is lower hemi-continuous for all $x$.

Note: The correspondence graphed in Figure [1.1] is upper hemi-continuous but not lower hemi-continuous.
1.4.1. Fixed Point Theorems.

**Definition 9.** (Fixed Point). A fixed point of a function $f : \mathbb{R}^n \to \mathbb{R}^n$ is a point $x$, such that $x = f(x)$. A fixed point of a correspondence, $F : \mathbb{R}^n \to \mathbb{R}^n$ is a point $x$, such that $x \in F(x)$.

Motivation: (1) Fixed points are used in many situations in economics. The problem of solving a set of equations can be described as a fixed point problem. (Suppose you are asked to find $x$ to solve $f(x) = 0$, where $f : \mathbb{R}^n \to \mathbb{R}^n$. Then define a new function, $g : \mathbb{R}^n \to \mathbb{R}^n$ such that $g(x) = x + f(x)$. If $x^*$ is a fixed point of $g$, then $x^*$ solves the system of equations because $x^* = g(x^*) = x^* + f(x^*)$. But this implies $f(x^*) = 0$, as desired.) Thus, fixed point theorems can be used to determine the existence of a solution to the Kuhn-Tucker first order conditions.

Motivation: (2) Fixed points are also crucial in the proofs of existence of general equilibrium in economics and in the existence of equilibria in games.

Intuition: If $f : \mathbb{R} \to \mathbb{R}$ then, a fixed point of $f$ is any point where the graph of $f$ crosses the 45 degree line.

A function can have many fixed points, a unique fixed point or no fixed points at all. The possibility of the latter makes us ask what conditions do we need to have on functions to be sure that it possesses a fixed point? This is the role of fixed point theorems.

**Theorem 10.** *(Brouwer's Fixed Point Theorem).* Suppose $f : \mathbb{R}^n \to \mathbb{R}^n$ and for some convex, compact set $C \subset \mathbb{R}^n$, $f$ maps $C$ into itself. (That is, if $x \in C$, then $f(x) \in C$.) If $f$ is continuous, then $f$, possesses a fixed point.

Note: There are a collection of conditions we typically need to check to apply Brouwer. 1) Continuity. 2) Convexity of $C$. 3) Compactness of $C$. 4) The fact that
f maps $C$ into itself. If we are looking for functions that do not have fixed points, then we should expect to have to violate at least one of these conditions.

**Theorem 11. (Kakutani’s Fixed Point Theorem).** Suppose $F : \mathbb{R}^n \to \mathbb{R}^n$ is convex-valued correspondence and for some convex, compact set $C \subset \mathbb{R}^n$, $F$ maps $C$ into itself. (That is, if $x \in C$, $F(x) \subset C$). If $F$ is upper hemicontinuous, then $F$, possesses a fixed point.

The previous fixed point theorems apply only in $\mathbb{R}^n$ and also do not tell us whether a mapping has a unique fixed point. The next fixed point theorem gives us conditions to ensure uniqueness and it also allows us to generalize to a broader class of mappings. This latter generalization will be useful in dynamic programming.

**Definition 12. (Contraction Mapping).** Suppose $f : \mathbb{R}^n \to \mathbb{R}^n$ such that $\|f(x) - f(y)\| \leq \theta \|x - y\|$ for $\theta < 1$ and for all $x, y$. Then $f$ is a contraction mapping

**Definition 13. (Contraction Mapping (Generalized)).** Let $C[a,b]$ be the set of all continuous functions $f : [a,b] \to \mathbb{R}$ with the “supnorm” metric

$$\|f\| = \max_{x \in [a,b]} \|f(x)\|$$

Suppose $T : C \to C$ (that is, $T$ takes a continuous function, does something to it and gives back a new, possibly different continuous function.) If, for all $f, g$ in $C$,

$$\|Tf - Tg\| \leq \theta \|f - g\|$$

for $\theta < 1$, then $T$ is a contraction mapping.

**Theorem 14. (Contraction Mapping Theorem).** If $f$ (respectively $T$) is a contraction mapping, it possesses a unique fixed point, $x^*$

**Remark.** The contraction mapping theorem is more general than stated here. (See de la Fuente, Section 7 and especially Theorem 7.16.) It applies to any contraction mapping from a complete metric space to itself.