Caltech Division of the Humanities and Social Sciences

Notes on the Implicit Function Theorem

KC Border
v. 2019.12.06::17.04

1 Implicit Function Theorems

The Implicit Function Theorem is a basic tool for analyzing extrema of differentiable functions.

Definition 1 *An equation of the form*

$$f(x, p) = y \tag{1}$$

implicitly defines x as a function of p *on a domain P if there is a function ξ on P for which $f(\xi(p), p) = y$ for all $p \in P$. It is traditional to assume that $y = 0$, but not essential.*

The use of zero in the above equation serves to simplify notation. The condition $f(x, p) = y$ is equivalent to $g(x, p) = 0$ where $g(x, p) = f(x, p) - y$, and this transformation of the problem is common in practice.

The implicit function theorem gives conditions under which it is possible to solve for x as a function of p in the neighborhood of a known solution $(\bar{x}, \bar{p})$. There are actually many implicit function theorems. If you make stronger assumptions, you can derive stronger conclusions. In each of the theorems that follows we are given a subset X of $\boldsymbol{R}^{\mathrm{n}}$, a metric space P (of parameters), a function f from $X \times P$ into $\boldsymbol{R}^{\mathrm{n}}$, and a point $(\bar{x}, \bar{p})$ in the interior of $X \times P$ such that $D_x f(\bar{x}, \bar{p})$ exists and is invertible. Each asserts the existence of neighborhoods U of $\bar{x}$ and W of $\bar{p}$ and a function $\xi\colon W \to U$ such that $f(\xi(p), p) = f(\bar{x}, \bar{p})$ for all $p \in W$. They differ in whether ξ is uniquely defined (in U) and how smooth it is. The following table serves as a guide to the theorems. For ease of reference, each theorem is stated as a standalone result.

Theorem	Hypotheses	Conclusion
All	f is continuous on $X \times P$ $D_x f(\bar{x}, \bar{p})$ is invertible	$f(\xi(p), p) = f(\bar{x}, \bar{p})$ for all p in W $\xi(\bar{p}) = \bar{x}$
2		ξ is continuous at $\bar{p}$
3	$D_x f$ is continuous on $X \times P$	ξ is unique in U ξ is continuous on W
4	$Df(\bar{x}, \bar{p})$ (wrt x, p) exists	ξ is differentiable at $\bar{p}$
5	Df (wrt x, p) exists on $X \times P$ $D_x f$ is continuous on $X \times P$	ξ is unique in U ξ is differentiable on W
1	f is C^k on $X \times P$	ξ is unique in U ξ is C^k on W

The first result is due to Halkin [13, Theorem B].

Theorem 2 (Implicit Function Theorem 0) *Let X be a subset of $\boldsymbol{R}^{\mathrm{n}}$, let P be a metric space, and let $f\colon X \times P \to \boldsymbol{R}^{\mathrm{n}}$ be continuous. Suppose the derivative $D_x f$ of f with respect to x exists at a point and that $D_x f(\bar{x}, \bar{p})$ is invertible. Let*

$$\bar{y} = f(\bar{x}, \bar{p}).$$

Then for any neighborhood U of $\bar{x}$, there is a neighborhood W of $\bar{p}$ and a function $\xi\colon W \to U$ such that:

a. $\xi(\bar{p}) = \bar{x}$.

b. $f(\xi(p), p) = \bar{y}$ *for all* $p \in W$.

c. ξ *is continuous at the point* $\bar{p}$.

However, it may be that ξ is neither continuous nor uniquely defined on any neighborhood of $\bar{p}$. There are two ways to strengthen the hypotheses and derive a stronger conclusion. One is to assume the derivative with respect to x exists and is continuous on $X \times P$. The other is to make P a subset of a Euclidean space and assume that f has a derivative with respect to (x, p) at the single point $(\bar{x}, \bar{p})$.

Taking the first approach allows us to conclude that the function ξ is uniquely defined and moreover continuous. The following result is Theorem 9.3 in Loomis and Sternberg [19, pp. 230–231].

Theorem 3 (Implicit Function Theorem 1a) *Let X be an open subset $\boldsymbol{R}^{\mathrm{n}}$, let P be a metric space, and let $f\colon X \times P \to \boldsymbol{R}^{\mathrm{n}}$ be continuous. Suppose* *the derivative $D_x f$ of f with respect to x exists at each point (x, p) and is continuous on $X \times P$. Assume that $D_x f(\bar{x}, \bar{p})$ is invertible. Let*

$$\bar{y} = f(\bar{x}, \bar{p}).$$

Then there are neighborhoods $U \subset X$ and $W \subset P$ of $\bar{x}$ and $\bar{p}$, and a function $\xi\colon W \to U$ such that:

a. *$f(\xi(p); p) = \bar{y}$ for all $p \in W$.*

b. *For each $p \in W$, $\xi(p)$ is the unique solution to (1) lying in U. In particular, then*

$$\xi(\bar{p}) = \bar{x}.$$

c. *ξ is continuous on W.*

The next result, also due to Halkin [13, Theorem E] takes the second approach. It concludes that ξ is differentiable at a single point. Related results may be found in Hurwicz and Richter [14, Theorem 1], Leach [16, 17], Nijenhuis [21], and Nikaidô [22, Theorem 5.6, p. 81].

Theorem 4 (Implicit Function Theorem 1b) *Let X be a subset of $\boldsymbol{R}^{\mathrm{n}}$, let P be an open subset of $\boldsymbol{R}^{\mathrm{m}}$, and let $f\colon X \times P \to \boldsymbol{R}^{\mathrm{n}}$ be continuous. Suppose the derivative Df of f with respect to (x, p) exists at $(\bar{x}, \bar{p})$. Write $Df(\bar{x}, \bar{p}) = (T, S)$, where $T\colon \boldsymbol{R}^{\mathrm{n}} \to \boldsymbol{R}^{\mathrm{n}}$ and $S\colon \boldsymbol{R}^{\mathrm{m}} \to \boldsymbol{R}^{\mathrm{m}}$, so that $Df(\bar{x}, \bar{p})(h, z) = Th + Sz$. Assume T is invertible. Let*

$$\bar{y} = f(\bar{x}, \bar{p}).$$

Then there is a neighborhood W of $\bar{p}$ and a function $\xi\colon W \to X$ satisfying

a. *$\xi(\bar{p}) = \bar{x}$.*

b. *$f(\xi(p), p) = \bar{y}$ for all $p \in W$.*

c. *ξ is differentiable (hence continuous) at $\bar{p}$, and*

$$D\xi(\bar{p}) = -T^{-1} \circ S.$$

The following result is Theorem 9.4 in Loomis and Sternberg [19, p. 231]. It strengthens the hypotheses of both Theorems 3 and 4. In return we get differentiability of ξ on W.

Theorem 5 (Semiclassical Implicit Function Theorem) *Let $X \times P$ be an open subset of $\boldsymbol{R}^{\mathrm{n}} \times \boldsymbol{R}^{\mathrm{m}}$, and let $f \colon X \times P \to \boldsymbol{R}^{\mathrm{n}}$ be differentiable. Suppose the derivative $D_x f$ of f with respect to x is continuous on $X \times P$. Assume that $D_x f(\bar{x}, \bar{p})$ is invertible. Let*

$$\bar{y} = f(\bar{x}, \bar{p}).$$

Then there are neighborhoods $U \subset X$ and $W \subset P$ of $\bar{x}$ and $\bar{p}$ on which equation (1) uniquely defines x as a function of p. That is, there is a function $\xi \colon W \to U$ such that:

a. *$f(\xi(p); p) = \bar{y}$ for all $p \in W$.*

b. *For each $p \in W$, $\xi(p)$ is the unique solution to (1) lying in U. In particular, then*
$$\xi(\bar{p}) = \bar{x}.$$

c. *ξ is differentiable on W, and*

$$\begin{bmatrix} \dfrac{\partial \xi_1}{\partial p_1} & \cdots & \dfrac{\partial \xi_1}{\partial p_m} \\ \vdots & & \vdots \\ \dfrac{\partial \xi_n}{\partial p_1} & \cdots & \dfrac{\partial \xi_n}{\partial p_m} \end{bmatrix} = - \begin{bmatrix} \dfrac{\partial f_1}{\partial x_1} & \cdots & \dfrac{\partial f_1}{\partial x_n} \\ \vdots & & \vdots \\ \dfrac{\partial f_n}{\partial x_1} & \cdots & \dfrac{\partial f_n}{\partial x_n} \end{bmatrix}^{-1} \begin{bmatrix} \dfrac{\partial f_1}{\partial p_1} & \cdots & \dfrac{\partial f_1}{\partial p_m} \\ \vdots & & \vdots \\ \dfrac{\partial f_n}{\partial p_1} & \cdots & \dfrac{\partial f_n}{\partial p_m} \end{bmatrix}.$$

The classical version maybe found, for instance, in Apostol [3, Theorem 7-6, p. 146], Rudin [24, Theorem 9.28, p. 224], or Spivak [26, Theorem 2-12, p. 41]. Some of these have the weaker statement that there is a unique function ξ within the class of continuous functions satisfying both $\xi(\bar{p}) = \bar{x}$ and $f(\xi(p); p) = 0$ for all p. Dieudonné [8, Theorem 10.2.3, p. 272] points out that the C^k case follows from the formula for $D\xi$ and the fact that the mapping from invertible linear transformations to their inverses, $A \mapsto A^{-1}$, is C^∞. (See Marsden [20, Lemma 2, p. 231].)

Classical Implicit Function Theorem *Let $X \times P$ be an open subset of $\boldsymbol{R}^{\mathrm{n}} \times \boldsymbol{R}^{\mathrm{m}}$, and let $f \colon X \times P \to \boldsymbol{R}^{\mathrm{n}}$ be C^k, for $k \geqslant 1$. Assume that $D_x f(\bar{x}, \bar{p})$ is invertible. Let*

$$\bar{y} = f(\bar{x}, \bar{p}).$$

Then there are neighborhoods $U \subset X$ and $W \subset P$ of $\bar{x}$ and $\bar{p}$ on which equation (1) uniquely defines x as a function of p. That is, there is a function $\xi \colon W \to U$ such that:

a. *$f(\xi(p); p) = \bar{y}$ for all $p \in W$.*

b. *For each $p \in W$, $\xi(p)$ is the unique solution to (1) lying in U. In particular, then*

$$\xi(\bar{p}) = \bar{x}.$$

c. *ξ is C^k on W, and*

$$\begin{bmatrix} \dfrac{\partial \xi_1}{\partial p_1} & \cdots & \dfrac{\partial \xi_1}{\partial p_m} \\ \vdots & & \vdots \\ \dfrac{\partial \xi_n}{\partial p_1} & \cdots & \dfrac{\partial \xi_n}{\partial p_m} \end{bmatrix} = - \begin{bmatrix} \dfrac{\partial f_1}{\partial x_1} & \cdots & \dfrac{\partial f_1}{\partial x_n} \\ \vdots & & \vdots \\ \dfrac{\partial f_n}{\partial x_1} & \cdots & \dfrac{\partial f_n}{\partial x_n} \end{bmatrix}^{-1} \begin{bmatrix} \dfrac{\partial f_1}{\partial p_1} & \cdots & \dfrac{\partial f_1}{\partial p_m} \\ \vdots & & \vdots \\ \dfrac{\partial f_n}{\partial p_1} & \cdots & \dfrac{\partial f_n}{\partial p_m} \end{bmatrix}.$$

As a bonus, let me throw in the following result, which is inspired by Apostol [4, Theorem 7.21].

Theorem 6 (Lipschitz Implicit Function Theorem) *Let P be a compact metric space and let $f\colon \boldsymbol{R} \times P \to \boldsymbol{R}$ be continuous and assume that there are real numbers $0 < m < M$ such that for each p*

$$m \leqslant \frac{f(x,p) - f(y,p)}{x - y} \leqslant M.$$

Then there is a unique function $\xi\colon P \to \boldsymbol{R}$ satisfying $f\big(\xi(p), p\big) = 0$. Moreover, ξ is continuous.

An interesting extension of this result to Banach spaces and functions with compact range may be found in Warga [28].

1.1 Proofs of Implicit Function Theorems

The proofs given here are based on fixed point arguments and are adapted from Halkin [12, 13], Rudin [24, pp. 220–227], Loomis and Sternberg [19, pp. 229–231], Marsden [20, pp. 230–237], and Dieudonné [8, pp. 265—273]. Another sort of proof, which is explicitly finite dimensional, of the classical case may be found in Apostol [3, p. 146] or Spivak [26, p. 41].

The first step is to show that for each p (at least in a neighborhood of $\bar{p}$) there is a zero of the function $f(x,p) - \bar{y}$, where $\bar{y} = f(\bar{x}, \bar{p})$. As is often the case, the problem of finding a zero of $f - \bar{y}$ is best converted to the problem of finding a fixed point of some other function. The obvious choice is to find a fixed point of $\pi_X - (f - \bar{y})$ (where $\pi_X(x,p) = x$), but the obvious choice is not clever enough in this case. Let

$$T = D_x f(\bar{x}, \bar{p}).$$

Define $\varphi\colon X \times P \to \boldsymbol{R}^{\mathrm{n}}$ by $\varphi = \pi_X - T^{-1}(f - \bar{y})$. That is,

$$\varphi(x,p) = x - T^{-1}\big(f(x,p) - \bar{y}\big). \tag{2}$$

Note that $\varphi(x,p) = x$ if and only if $T^{-1}\big(f(x,p) - \bar{y}\big) = 0$. But the invertibility of T^{-1} guarantees that this happens if and only if $f(x,p) = \bar{y}$. Thus the problem of finding a zero of $f(\cdot,p) - \bar{y}$ is equivalent to that of finding a fixed point of $\varphi(\cdot,p)$. Note also that

$$\varphi(\bar{x}, \bar{p}) = \bar{x}. \tag{3}$$

Observe that φ is continuous and also has a derivative $D_x\varphi$ with respect to x whenever f does. In fact,

$$D_x\varphi(x,p) = I - T^{-1}D_x f(x,p).$$

In particular, at $(\bar{x}, \bar{p})$, we get

$$D_x\varphi(\bar{x}, \bar{p}) = I - T^{-1}T = 0. \tag{4}$$

That is, $D_x\varphi(\bar{x}, \bar{p})$ is the zero transformation.

Recall that for a linear transformation A, its operator norm $\|A\|$ is defined by $\|A\| = \sup_{|x|\leqslant 1} |Ax|$, and satisfies $|Ax| \leqslant \|A\| \cdot |x|$ for all x. If A is invertible, then $\|A^{-1}\| > 0$.

Proof of Theorem 2: Let X, P, and $f\colon X \times P \to \boldsymbol{R}^{\mathrm{n}}$ be as in the hypotheses of Theorem 2.

In order to apply a fixed point argument, we must first find a subset of X that is mapped into itself. By the definition of differentiability and (4) we can choose $r > 0$ so that

$$\frac{\big|\varphi(x,\bar{p}) - \varphi(\bar{x},\bar{p})\big|}{|x - \bar{x}|} \leqslant \frac{1}{2} \qquad \text{for all } x \in \bar{B}_r(\bar{x}).$$

Noting that $\varphi(\bar{x}, \bar{p}) = \bar{x}$ and rearranging, it follows that

$$\big|\varphi(x,\bar{p}) - \bar{x}\big| \leqslant \frac{r}{2} \qquad \text{for all } x \in \bar{B}_r(\bar{x}).$$

For each p set $m(p) = \max_x \bigl|\varphi(x,p) - \varphi(x,\bar p)\bigr|$ as x runs over the compact set $\bar B_r(\bar x)$. Since φ is continuous (and $\bar B_r(\bar x)$ is a fixed set), the Maximum Theorem **??** implies that m is continuous. Since $m(\bar p) = 0$, there is some $\varepsilon > 0$ such that $|m(p)| < \frac{r}{2}$ for all $p \in B_\varepsilon(\bar p)$. That is,

$$\bigl|\varphi(x,p) - \varphi(x,\bar p)\bigr| < \frac{r}{2} \qquad \text{for all } x \in \bar B_r(\bar x),\ p \in B_\varepsilon(\bar p).$$

For each $p \in B_\varepsilon(\bar p)$, the function φ maps $\bar B_r(\bar x)$ into itself, for

$$\begin{array}{rcl} \bigl|\varphi(x,p) - \bar x\bigr| & \leqslant & \bigl|\varphi(x,p) - \varphi(x,\bar p)\bigr| + \bigl|\varphi(x,\bar p) - \bar x\bigr| \\ & < & \dfrac{r}{2} + \dfrac{r}{2} \\ & = & r. \end{array}$$

That is, $\varphi(x,p) \in \bar B_r(\bar x)$. Since φ is continuous and $\bar B_r(\bar x)$ is compact and convex, by the Brouwer Fixed Point Theorem (e.g., [5, Corollary 6.6, p. 29]), there is some $x \in \bar B_r(\bar x)$ satisfying $\varphi(x,p) = x$, or in other words $f(x,p) = 0$.

We have just proven parts (a) and (b) of Theorem 2. That is, for every neighborhood X of $\bar x$, there is a neighborhood $W = \bar B_{\varepsilon(r)}(\bar p)$ of $\bar p$ and a function ξ from $\bar B_{\varepsilon(r)}(\bar p)$ into $\bar B_r(\bar x) \subset X$ satisfying $\xi(\bar p) = \bar x$ and $f\bigl(\xi(p),p\bigr) = 0$ for all $p \in W$. (Halkin actually breaks this part out as Theorem A.)

We can use the above result to construct a ξ that is continuous at $\bar p$. Start with a given neighborhood U of $\bar x$. Construct a sequence of $r_1 > r_2 > \cdots > 0$ satisfying $\lim r_n = 0$ and for each n consider the neighborhood $U_n = U \cap B_{r_n}(\bar x)$. From the argument above there is a neighborhood W_n of $\bar p$ and a function ξ_n from W_n into $U_n \subset U$ satisfying $\xi_n(\bar p) = \bar x$ and $f\bigl(\xi_n(p),p\bigr) = 0$ for all $p \in W_n$. Without loss of generality we may assume $W_n \supset W_{n+1}$ (otherwise replace W_{n+1} with $W_n \cap W_{n+1}$), so set $W = W_1$. Define $\xi \colon W \to U$ by $\xi(p) = \xi_n(p)$ for $p \in W_n \setminus W_{n+1}$. Then ξ is continuous at $\bar p$, satisfies $\xi(\bar p) = \bar x$, and $f\bigl(\xi(p),p\bigr) = 0$ for all $p \in W$. ∎

Note that the above proof used in an essential way the compactness of $\bar B_r(\bar x)$, which relies on the finite dimensionality of $\boldsymbol{R}^{\mathrm{n}}$. The compactness was used first to show that $m(p)$ is finite, and second to apply the Brouwer fixed point theorem.

Theorem 3 adds to the hypotheses of Theorem 2. It assumes that $D_x f$ exists everywhere on $X \times P$ and is continuous. The conclusion is that there are *some* neighborhoods U of $\bar x$ and W of $\bar p$ and a continuous function $\xi \colon W \to U$ such that $\xi(p)$ is the unique solution to $f\bigl(\xi(p),p\bigr) = 0$ lying in U. It is the uniqueness of $\xi(p)$ that puts a restriction on U. If U is too large, say $U = X$, then the solution need not be unique. (On the other hand, it is easy to show, as does Dieudonné [8, pp. 270–271], there is at most one continuous ξ, provided U is connected.) The argument we use here, which resembles that of Loomis and Sternberg, duplicates some of the proof of Theorem 2, but we do not actually need to assume that the domain of f lies in the finite dimensional space $\boldsymbol{R}^{\mathrm{n}} \times \boldsymbol{R}^{\mathrm{m}}$, any Banach spaces will do, and the proof need not change. This means that we cannot use Brouwer's theorem, since closed balls are not compact in general Banach spaces. Instead, we will be able to use the Mean Value Theorem and the Contraction Mapping Theorem.

Proof of Theorem 3: Let X, P, and let $f\colon X \times P \to \boldsymbol{R}^{\mathrm{n}}$ obey the hypotheses of Theorem 3. Set $T = D_x f(\bar{x}, \bar{p})$, and recall that T is invertible.

Again we must find a suitable subset of X so that each $\varphi(\cdot, p)$ maps this set into itself. Now we use the hypothesis that $D_x f$ (and hence $D_x\varphi$) exists and is continuous on $X \times P$ to deduce that there is a neighborhood $\bar{B}_r(\bar{x}) \times W_1$ of $(\bar{x}, \bar{p})$ on which the operator norm $\|D_x\varphi\|$ is strictly less than $\frac{1}{2}$. Set $U = \bar{B}_r(\bar{x})$. Since $\varphi(\bar{x}, \bar{p}) = \bar{x}$ and since φ is continuous (as f is), we can now choose W so that $\bar{p} \in W$, $W \subset W_1$, and $p \in W$ implies

$$\left|\varphi(\bar{x}, p) - \bar{x}\right| < \frac{r}{2}.$$

We now show that for each $p \in W$, the mapping $x \mapsto \varphi(x, p)$ is a contraction that maps $\bar{B}_r(\bar{x})$ into itself. To see this, note that the Mean Value Theorem (or Taylor's Theorem) implies

$$\varphi(x, p) - \varphi(y, p) = D_x\varphi(z, p)(x - y),$$

for some z lying on the segment between x and y. If x and y lie in $\bar{B}_r(\bar{x})$, then z too must lie in $\bar{B}_r(\bar{x})$, so $\|D_x\varphi(z, p)\| < \frac{1}{2}$. It follows that

$$\left|\varphi(x, p) - \varphi(y, p)\right| < \frac{1}{2}\left|x - y\right| \qquad \text{for all } x, y \in \bar{B}_r(\bar{x}),\ p \in B_\varepsilon(\bar{p}), \tag{5}$$

so $\varphi(\cdot, p)$ is a contraction on $\bar{B}_r(\bar{x})$ with contraction constant $\frac{1}{2}$.

To see that $\bar{B}_r(\bar{x})$ is mapped into itself, let (x, p) belong to $\bar{B}_r(\bar{x}) \times W$ and observe that

$$\begin{aligned} \left|\varphi(x, p) - \bar{x}\right| &\leqslant \left|\varphi(x, p) - \varphi(\bar{x}, p)\right| + \left|\varphi(\bar{x}, p) - \bar{x}\right| \\ &< \frac{1}{2}\left|x - \bar{x}\right| + \frac{r}{2} \\ &< r. \end{aligned}$$

Thus $\varphi(x, p) \in \bar{B}_r(\bar{x})$.

Since $\bar{B}_r(\bar{x})$ is a closed subset of the complete metric space $\boldsymbol{R}^{\mathrm{n}}$, it is complete itself, so the Contraction Mapping Theorem guarantees that there is a unique fixed point of $\varphi(\cdot, p)$ in $\bar{B}_r(\bar{x})$. In other words, for each $p \in W$ there is a unique point $\xi(p)$ lying in $U = \bar{B}_r(\bar{x})$ satisfying $f(\xi(p), p) = 0$.

It remains to show that ξ must be continuous on W. This follows from a general result on parametric contraction mappings, presented as Lemma 7 below, which also appears in [19, Corollary 4, p. 230]. ∎

Note that the above proof nowhere uses the finite dimensionality of $\boldsymbol{R}^{\mathrm{m}}$, so the theorem actually applies to a general Banach space.

Proof of Theorem 4: For this theorem, in addition to the hypotheses of Theorem 2, we need P to be a subset of a Euclidean space (or more generally a Banach space), so that it makes sense to partially differentiate with respect to p. Now assume f is differentiable with respect to (x, p) at the point $(\bar{x}, \bar{p})$.

There is a neighborhood W of $\bar{p}$ and a function $\xi \colon W \to X$ satisfying the conclusions of Theorem 2. It turns out that under the added hypotheses, such a function ξ is differentiable at $\bar{p}$.

We start by showing that ξ is locally Lipschitz continuous at $\bar{p}$. First set

$$\Delta(x, p) = f(x, p) - f(\bar{x}, \bar{p}) - T(x - \bar{x}) - S(p - \bar{p}).$$

Since Df exists at $(\bar{x}, \bar{p})$, there exists $r > 0$ such that $B_r(\bar{x}) \times B_r(\bar{p}) \subset X \times W$ and if $|x - \bar{x}| < r$ and $|p - \bar{p}| < r$, then

$$\frac{|\Delta(x, p)|}{|x - \bar{x}| + |p - \bar{p}|} < \frac{1}{2\,\|T^{-1}\|},$$

which in turn implies

$$\left|T^{-1}\Delta(x, p)\right| < \frac{1}{2}|x - \bar{x}| + \frac{1}{2}|p - \bar{p}|.$$

Since ξ is continuous at $\bar{p}$ and $\xi(\bar{p}) = \bar{x}$, there is some $r \geqslant \delta > 0$ such that $|p - \bar{p}| < \delta$ implies $|\xi(p) - \bar{x}| < r$. Thus

$$\left|T^{-1}\Delta(\xi(p), p)\right| < \frac{1}{2}\left|\xi(p) - \bar{x}\right| + \frac{1}{2}|p - \bar{p}| \qquad \text{for all } p \in B_\delta(\bar{p}). \tag{6}$$

But $f(\xi(p), p) - f(\bar{x}, \bar{p}) = 0$ implies

$$\left|T^{-1}\Delta(\xi(p), p)\right| = \left|(\xi(p) - \bar{x}) + T^{-1}S(p - \bar{p})\right|. \tag{7}$$

Therefore, from the facts that $|a+b| < c$ implies $|a| < |b| + c$, and $\xi(\bar p) = \bar x$, equations (6) and (7) imply

$$\big|\xi(p) - \xi(\bar p)\big| < \big|T^{-1}S(p-\bar p)\big| + \frac{1}{2}\big|\xi(p) - \xi(\bar p)\big| + \frac{1}{2}|p-\bar p| \quad \text{for all } p \in B_\delta(\bar p)$$

or,

$$\big|\xi(p) - \xi(\bar p)\big| < (2\|T^{-1}S\| + 1)|p - \bar p| \quad \text{for all } p \in B_\delta(\bar p).$$

That is, ξ satisfies a local Lipschitz condition at $\bar p$. For future use set $M = 2\|T^{-1}S\| + 1$.

Now we are in a position to prove that $-T^{-1}S$ is the differential of ξ at $\bar p$. Let $\varepsilon > 0$ be given. Choose $0 < r < \delta$ so that $|x - \bar x| < r$ and $|p - \bar p| < r$ implies

$$\frac{\big|\Delta(x,p)\big|}{|x-\bar x| + |p - \bar p|} < \frac{\varepsilon}{(M+1)\,\|T^{-1}\|},$$

so

$$\big|(\xi(p) - \xi(\bar p)) + T^{-1}S(p-\bar p)\big| = \big|T^{-1}\Delta\big(\xi(p),p\big)\big| < \frac{\varepsilon}{(M+1)}\big(|\xi(p)-\xi(\bar p)| + |p-\bar p|\big) \leqslant \varepsilon|p-\bar p|,$$

for $|p - \bar p| < r$, which shows that indeed $-T^{-1}S$ is the differential of ξ at $\bar p$. ∎

Proof of Theorem 5: ************ ∎

Proof of Theorem 1: ************ ∎

Proof of Theorem 6: Let f satisfy the hypotheses of the theorem. Let $C(P)$ denote the set of continuous real functions on P. Then $C(P)$ is complete under the uniform norm metric, $\|f - g\| = \sup_p |f(p) - g(p)|$ [2, Lemma 3.97, p. 124]. For each p define the function $\psi_p \colon \boldsymbol{R} \to \boldsymbol{R}$ by

$$\psi_p(x) = x - \frac{1}{M} f(x,p).$$

Note that $\psi_p(x) = x$ if and only if $f(x,p) = 0$. If ψ_p has a unique fixed point $\xi(p)$, then we shall have shown that there is a unique function ξ satisfying $f\big(\xi(p),p\big) = 0$. It suffices to show that ψ_p is a contraction.

To see this, write

$$\begin{aligned}
\psi_p(x) - \psi_p(y) &= x - y - \frac{f(x,p) - f(y,p)}{M} \\
&= \left(1 - \frac{1}{M}\frac{f(x,p) - f(y,p)}{x - y}\right)(x - y).
\end{aligned}$$

By hypothesis

$$0 < m \leqslant \frac{f(x,p) - f(y,p)}{x - y} \leqslant M,$$

so

$$|\psi_p(x) - \psi_p(y)| \leqslant (1 - \frac{m}{M})|x - y|.$$

This shows that ψ_p is a contraction with constant $1 - \frac{m}{M} < 1$.

To see that ξ is actually continuous, define the function $\psi\colon C(P) \to C(P)$ via

$$\psi g(p) = g(p) - \frac{1}{M} f\big(g(p), p\big).$$

(Since f is continuous, ψg is continuous whenever g is continuous.) The pointwise argument above is independent of p, so it also shows that $|\psi g(p) - \psi h(p)| \leqslant (1 - \frac{m}{M})|g(p) - h(p)|$ for any functions g and h. Thus

$$\|\psi g - \psi h\| \leqslant (1 - \frac{m}{M})\|g - h\|.$$

In other words ψ is a contraction on $C(P)$, so it has a unique fixed point $\bar{g}$ in $C(P)$, so $\bar{g}$ is continuous. But $\bar{g}$ also satisfies $f\big(\bar{g}(p), p\big)$, but since $\xi(p)$ is unique we have $\xi = \bar{g}$ is continuous. ∎

Lemma 7 (Continuity of fixed points) *Let $\varphi\colon X \times P \to X$ be continuous in p for each x, where X is a complete metric space under the metric d and P is a metrizable space. Suppose that φ is a uniform contraction in x. That is, there is some $0 \leqslant \alpha < 1$ such that*

$$d\big(\varphi(x,p) - \varphi(y,p)\big) \leqslant \alpha d(x,y)$$

for all x and y in X and all p in P. Then the mapping $\xi\colon P \to X$ from p to the unique fixed point of $\varphi(\cdot, p)$, defined by $\varphi\big(\xi(p), p\big) = \xi(p)$, is continuous.

Proof: Fix a point p in P and let $\varepsilon > 0$ be given. Let ρ be a compatible metric on P and using the continuity of $\varphi(x, \cdot)$ on P, choose $\delta > 0$ so that $\rho(p,q) < \delta$ implies that

$$d\big(\varphi(\xi(p), p), \varphi(\xi(p), q)\big) < (1 - \alpha)\varepsilon.$$

So if $\rho(p,q) < \delta$, then

$$\begin{aligned} d\big(\xi(p), \xi(q)\big) &= d\big(\varphi(\xi(p), p), \varphi(\xi(q), q)\big) \\ &\leqslant d\big(\varphi(\xi(p), p), \varphi(\xi(p), q)\big) + d\big(\varphi(\xi(p), q), \varphi(\xi(q), q)\big) \\ &< (1 - \alpha)\varepsilon + \alpha d\big(\xi(p), \xi(q)\big) \end{aligned}$$

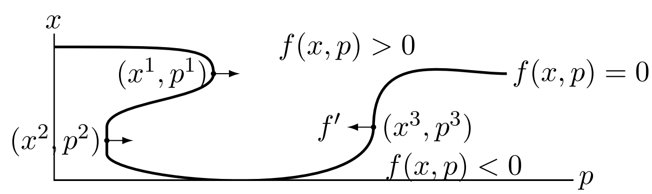

Figure 1. Looking for implicit functions.

so

$$(1-\alpha)d\big(\xi(p),\xi(q)\big) < (1-\alpha)\varepsilon$$

or

$$d\big(\xi(p),\xi(q)\big) < \varepsilon,$$

which proves that ξ is continuous at p. ∎

1.2 Examples

Figure 1 illustrates the Implicit Function Theorem for the special case $n = m = 1$, which is the only one I can draw. The figure is drawn sideways since we are looking for x as a function of p. In this case, the requirement that the differential with respect to x be invertible reduces to $\frac{\partial f}{\partial x} \neq 0$. That is, in the diagram the gradient of f may not be horizontal. In the figure, you can see that the points, (x^1,p^1), (x^2,p^2), and (x^3,p^3), the differentials $D_x f$ are zero. At (x^1,p^1) and (x^2,p^2) there is no way to define x as a continuous function of p locally. (Note however, that if we allowed a discontinuous function, we could define x as a function of p in a neighborhood of p^1 or p^2, but not uniquely.) At the point (x^3,p^3), we can uniquely define x as a function of p near p^3, but this function is not differentiable.

Another example of the failure of the conclusion of the Classical Implicit Function Theorem is provided by the function from Example **??**.

Example 8 (Differential not invertible) Define $f\colon \boldsymbol{R}\times\boldsymbol{R}\to\boldsymbol{R}$ by

$$f(x,p) = -(x-p^2)(x-2p^2).$$

Consider the function implicitly defined by $f(x,p) = 0$. The function f is zero along the parabolas $x = p^2$ and $x = 2p^2$, and in particular $f(0,0) = 0$. See Figure 2 on page 20. The hypothesis of the Implicit Function Theorem is not satisfied since $\frac{\partial f(0,0)}{\partial x} = 0$. The conclusion also fails. The problem here is not that a smooth implicit function through $(x,p) = (0,0)$ fails to exist.

The problem is that it is not unique. There are four distinct continuously differentiable implicitly defined functions. □

Example 9 (Lack of continuous differentiability) Consider again the function $h(x) = x + 2x^2 \sin \frac{1}{x^2}$ from Example **??**. Recall that h is differentiable everywhere, but not continuously differentiable at zero. Furthermore, $h(0) = 0$, $h'(0) = 1$, but h is not monotone on any neighborhood of zero. Now consider the function $f(x,p) = h(x) - p$. It satisfies $f(0,0) = 0$ and $\frac{\partial f(0,0)}{\partial x} \neq 0$, but it there is no unique implicitly defined function on any neighborhood, nor is there any continuous implicitly defined function.

To see this, note that $f(x,p) = 0$ if and only if $h(x) = p$. So a unique implicitly defined function exists only if h is invertible on some neighborhood of zero. But this is not so, for given any $\varepsilon > 0$, there is some $0 < p < \frac{\varepsilon}{2}$ for which there are $0 < x < x' < \varepsilon$ satisfying $h(x) = h(x') = p$. It is also easy to see that no continuous function satisfies $h(\xi(p)) = p$ either. □

If X is more than one-dimensional there are subtler ways in which $D_x f$ may fail to be continuous. The next example is taken from Dieudonné [8, Problem 10.2.2, p. 273].

Example 10 Define $f \colon \boldsymbol{R}^2 \to \boldsymbol{R}^2$ by

$$f_1(x,y) = x$$

and

$$f_2(x,y) = \begin{cases} y - x^2 & 0 \leqslant x^2 \leqslant y \\ \dfrac{y^2 - x^2 y}{x^2} & 0 \leqslant y < x^2 \\ -f_2(x,-y) & y < 0. \end{cases}$$

Dieudonné claims that f is everywhere differentiable on $\boldsymbol{R}^2$, and $Df(0,0)$ is the identity mapping, but Df is not continuous at the origin. I'll let you ponder that. Work out the details.

Furthermore in every neighborhood of the origin there are distinct points (x,y) and (x',y') with $f(x,y) = f(x',y')$. To find such a pair, pick a (small) $x' = x > 0$ and set $y = x^2$ and $y' = -x^2$. Then $f_1(x,y) = f_1(x',y') = x$, and $f_2(x,y) = f_2(x',y') = 0$.

This implies f has no local inverse, so the equation $f(x,y) - p = f(0,0) = 0$ does not uniquely define (x,y) as a function of $p = (p_1,p_2)$ near the origin. □

1.3 Implicit vs. inverse function theorems

In this section we discuss the relationship between the existence of a unique implicitly defined function and the existence of an inverse function. These results are quite standard and may be found, for instance, in Marsden [20, p. 234].

First we show how the implicit function theorem can be used to prove an inverse function theorem. Suppose $X \subset \boldsymbol{R}^{\mathrm{n}}$ and $g \colon X \to \boldsymbol{R}^{\mathrm{n}}$. Let P be a neighborhood of $\bar{p} = g(\bar{x})$. Consider $f \colon X \times P \to \boldsymbol{R}^{\mathrm{n}}$ defined by

$$f(x,p) = g(x) - p.$$

Then $f(x,p) = 0$ if and only if $p = g(x)$. Thus if there is a unique implicitly defined function $\xi \colon P \to X$ implicitly defined by $f\bigl(\xi(p),p\bigr) = 0$, it follows that g is invertible and $\xi = g^{-1}$. Now compare the Jacobian matrix of f with respect to x and observe that it is just the Jacobian matrix of g. Thus each of the implicit function theorems has a corresponding inverse function theorem.

We could also proceed in the other direction, as is usually the case in textbooks. Let $X \times P$ be a subset of $\boldsymbol{R}^{\mathrm{n}} \times \boldsymbol{R}^{\mathrm{m}}$, and let $f \colon X \times P \to \boldsymbol{R}^{\mathrm{n}}$, and suppose $f(\bar{x},\bar{p}) = 0$. Define a function $g \colon X \times P \to \boldsymbol{R}^{\mathrm{n}} \times P$ by

$$g(x,p) = \bigl(f(x,p),p\bigr).$$

Suppose g is invertible, that is, it is one-to-one and onto. Define $\xi \colon P \to X$ by

$$\xi(p) = \pi_x\bigl(g^{-1}(0,p)\bigr).$$

Then ξ is the unique function implicitly defined by

$$\bigl(f\bigl(\xi(p),p\bigr),p\bigr) = 0$$

for all $p \in P$. Now let's compare hypotheses. The standard Inverse Function Theorem, e.g. [20, Theorem 7.1.1, p. 206], says that if g is continuously differentiable and has a nonsingular Jacobian matrix at some point, then there is a neighborhood of the point where g is invertible. The Jacobian matrix for $g(x,p) = \big(f(x,p),p\big)$ above is

$$\begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} & \frac{\partial f_1}{\partial p_1} & \cdots & \frac{\partial f_1}{\partial p_m} \\ \vdots & & \vdots & \vdots & & \vdots \\ \frac{\partial f_n}{\partial x_1} & \cdots & \frac{\partial f_n}{\partial x_n} & \frac{\partial f_n}{\partial p_1} & \cdots & \frac{\partial f_n}{\partial p_m} \\ \\ 0 & \cdots & 0 & 1 & & 0 \\ \vdots & & \vdots & & \ddots & \\ 0 & \cdots & 0 & 0 & & 1 \end{bmatrix}.$$

Since this is block diagonal, it is easy to see that this Jacobian matrix is nonsingular at $(\bar{x},\bar{p})$ if and only if the derivative $D_x f(\bar{x},\bar{p})$, is invertible.

1.4 Global inversion

The inverse function theorems proven above are local results. Even if the Jacobian matrix of a function never vanishes, it may be that the function does not have an inverse everywhere. The following example is well known, see e.g., [20, Example 7.1.2, p. 208].

Add a section on the Gale–Nikaidô Theorem.

Example 11 (A function without a global inverse) Define $f\colon \boldsymbol{R}^2 \to \boldsymbol{R}^2$ via

$$f(x,y) = (e^x \cos y, e^x \sin y).$$

Then the Jacobian matrix is

$$\begin{pmatrix} e^x \cos y & -e^x \sin y \\ e^x \sin y & e^x \cos y \end{pmatrix}$$

which has determinant $e^{2x}(\cos^2 y + \sin^2 y) = e^{2x} > 0$ everywhere. Nonetheless, f is not invertible since $f(x,y) = f(x, y+2\pi)$ for every x and y. □

2 Applications of the Implicit Function Theorem

2.1 A fundamental lemma

A **curve** in $\boldsymbol{R}^{\mathrm{n}}$ is simply a function from an interval of $\boldsymbol{R}$ into $\boldsymbol{R}^{\mathrm{n}}$, usually assumed to be continuous.

Fundamental Lemma on Curves *Let U be an open set in $\boldsymbol{R}^{\mathrm{n}}$ and let $g\colon U \to \boldsymbol{R}^{\mathrm{m}}$. Let $x^* \in U$ satisfy $g(x^*) = 0$, and suppose g is differentiable at x^*. Assume that ${g_1}'(x^*), \dots, {g_m}'(x^*)$ are linearly independent. Let $v \in \boldsymbol{R}^{\mathrm{n}}$ satisfy*

$$
{g_i}'(x^*) \cdot v = 0, \quad i = 1, \dots, m.
$$

Then there exists $\delta > 0$ and a curve $\hat{x}\colon (-\delta, \delta) \to U$ satisfying:

1. *$\hat{x}(0) = x^*$.*
2. *$g\big(\hat{x}(\alpha)\big) = 0$ for all $\alpha \in (-\delta, \delta)$.*
3. *$\hat{x}$ is differentiable at 0. Moreover, if g is C^k on U, then $\hat{x}$ is C^k on $(-\delta, \delta)$.*
4. *$\hat{x}'(0) = v$.*

Proof: Since the ${g_i}'(x^*)$s are linearly independent, $n \geqslant m$, and without loss of generality, we may assume the coordinates are numbered so that the $m \times m$ matrix

$$
\begin{bmatrix}
\frac{\partial g_1}{\partial x_1} & \cdots & \frac{\partial g_1}{\partial x_m} \\
\vdots & & \vdots \\
\frac{\partial g_m}{\partial x_1} & \cdots & \frac{\partial g_m}{\partial x_m}
\end{bmatrix}
$$

is invertible at x^*.

Fix v satisfying ${g_i}'(x^*) \cdot v = 0$ for all $i = 1, \dots, m$. Rearranging terms we have

$$
\sum_{j=1}^{m} \frac{\partial g_i(x^*)}{\partial x_j} \cdot v_j = - \sum_{j=m+1}^{n} \frac{\partial g_i(x^*)}{\partial x_j} \cdot v_j \quad i = 1, \dots, m,
$$

or in matrix terms

$$
\begin{bmatrix}
\frac{\partial g_1}{\partial x_1} & \cdots & \frac{\partial g_1}{\partial x_m} \\
\vdots & & \vdots \\
\frac{\partial g_m}{\partial x_1} & \cdots & \frac{\partial g_m}{\partial x_m}
\end{bmatrix}
\begin{bmatrix} v_1 \\ \vdots \\ v_m \end{bmatrix}
= -
\begin{bmatrix}
\frac{\partial g_1}{\partial x_{m+1}} & \cdots & \frac{\partial g_1}{\partial x_n} \\
\vdots & & \vdots \\
\frac{\partial g_m}{\partial x_{m+1}} & \cdots & \frac{\partial g_m}{\partial x_n}
\end{bmatrix}
\begin{bmatrix} v_{m+1} \\ \vdots \\ v_n \end{bmatrix},
$$

so

$$
\begin{bmatrix} v_1 \\ \vdots \\ v_m \end{bmatrix}
= -
\begin{bmatrix}
\frac{\partial g_1}{\partial x_1} & \cdots & \frac{\partial g_1}{\partial x_m} \\
\vdots & & \vdots \\
\frac{\partial g_m}{\partial x_1} & \cdots & \frac{\partial g_m}{\partial x_m}
\end{bmatrix}^{-1}
\begin{bmatrix}
\frac{\partial g_1}{\partial x_{m+1}} & \cdots & \frac{\partial g_1}{\partial x_n} \\
\vdots & & \vdots \\
\frac{\partial g_m}{\partial x_{m+1}} & \cdots & \frac{\partial g_m}{\partial x_n}
\end{bmatrix}
\begin{bmatrix} v_{m+1} \\ \vdots \\ v_n \end{bmatrix}.
$$

Observe that these conditions completely characterize v. That is, for any $y \in \boldsymbol{R}^{\mathrm{n}}$,

$$\left(g_i{}'(x^*)\cdot y = 0,\ i = 1,\dots,m, \text{ and } y_j = v_j,\ j = m{+}1,\dots,n\right) \quad\Longrightarrow\quad y = v. \tag{8}$$

Define the C^∞ function $f\colon \boldsymbol{R}^{\mathrm{m}} \times \boldsymbol{R} \to \boldsymbol{R}^{\mathrm{n}}$ by

$$f(z,\alpha) = (z_1,\dots,z_m, x^*_{m+1} + \alpha v_{m+1},\dots,x^*_n + \alpha v_n).$$

Set $z^* = (x^*_1,\dots,x^*_m)$ and note that $f(z^*,0) = x^*$. Since x^* is an interior point of U, there is a neighborhood W of z^* and an interval $(-\eta,\eta)$ in $\boldsymbol{R}$ so that for every $z \in W$ and $\alpha \in (-\eta,\eta)$, the point $f(z,\alpha)$ belongs to $U \subset \boldsymbol{R}^{\mathrm{n}}$. Finally, define $h\colon W \times (-\eta,\eta) \to \boldsymbol{R}^{\mathrm{m}}$ by

$$h(z,\alpha) = g\big(f(z,\alpha)\big) = g(z_1,\dots,z_m, x^*_{m+1} + \alpha v_{m+1},\dots,x^*_n + \alpha v_n).$$

Observe that h possesses the same degree of differentiability as g, since h is the composition of g with the C^∞ function f.

Then for $j = 1,\dots,m$, we have $= \frac{\partial h_i}{\partial z_j}(z,0) = \frac{\partial g_i}{\partial x_j}(x)$, where $z = (x_1,\dots,x_m)$. Therefore the m-vectors

$$\begin{bmatrix} \dfrac{\partial h_i(z^*;0)}{\partial z_1} \\ \vdots \\ \dfrac{\partial h_i(z^*;0)}{\partial z_m} \end{bmatrix}, \quad i = 1,\dots,m$$

are linearly independent.

But $h(z^*,0) = 0$, so by the Implicit Function Theorem 4 there is an interval $(-\delta,\delta) \subset (-\eta,\eta)$ about 0, a neighborhood $V \subset W$ of z^* and a function $\zeta\colon (-\delta,\delta) \to V$ such that

$$\zeta(0) = z^*,$$

$$h\big(\zeta(\alpha),\alpha\big) = 0, \quad \text{for all } \alpha \in (-\delta,\delta),$$

and ζ is differentiable at 0. Moreover, by Implicit Function Theorem 1, if g is C^k on U, then h is C^k, so ζ is C^k on $(-\delta,\delta)$.

Define the curve $\hat{x}\colon (-\delta,\delta) \to U$ by

$$\hat{x}(\alpha) = f\big(\zeta(\alpha),\alpha\big) = \big(\zeta_1(\alpha),\dots,\zeta_m(\alpha), x^*_{m+1} + \alpha v_{m+1},\dots,x^*_n + \alpha v_n\big). \tag{9}$$

Then $\hat{x}(0) = x^*$,

$$g\big(\hat{x}(\alpha)\big) = 0 \text{ for all } \alpha \in (-\delta,\delta),$$

and $\hat{x}$ is differentiable at 0, and if g is C^k, then $\hat{x}$ is C^k. So by the Chain Rule,

$$g_i{}'(x^*) \cdot \hat{x}'(0) = 0, \quad i = 1, \ldots, m.$$

Now by construction (9), $\hat{x}'_j(0) = v_j$, for $j = m+1, \ldots, n$. Thus (8) implies $\hat{x}'(0) = v$. ∎

2.2 A note on comparative statics

"Comparative statics" analysis tells us how equilibrium values of endogenous variables $x_1, \ldots, x_n$ (the things we want to solve for) change as a function of the exogenous parameters $p_1, \ldots, p_m$. (As such it is hardly unique to economics.) Typically we can write the equilibrium conditions of our model as the zero of a system of equations in the endogenous variables and the exogenous parameters:

$$\begin{array}{rcl} F_1(x_1, \ldots, x_n; p_1, \ldots, p_m) & = & 0 \\ & \vdots & \\ F_n(x_1, \ldots, x_n; p_1, \ldots, p_m) & = & 0 \end{array} \tag{10}$$

This implicitly defines x as a function of p, which we will explicitly denote $x = \xi(p)$, or

$$(x_1, \ldots, x_n) = (\xi_1(p_1, \ldots, p_m), \ldots, \xi_n(p_1, \ldots, p_m)).$$

This explicit function, if it exists, satisfies the implicit definition

$$F\big(\xi(p); p\big) = 0 \tag{11}$$

for at least a rectangle of values of p. The Implicit Function Theorem tells that such an explicit function exists whenever it is possible to solve for all its partial derivatives.

Setting $G(p) = F\big(\xi(p); p\big)$, and differentiating G_i with respect to p_j, yields, by equation (11),

$$\sum_k \frac{\partial F_i}{\partial x_k} \frac{\partial \xi_k}{\partial p_j} + \frac{\partial F_i}{\partial p_j} = 0 \tag{12}$$

for each $i = 1, \ldots, n$, $j = 1, \ldots, m$. In matrix terms we have

$$\begin{bmatrix} \frac{\partial F_1}{\partial x_1} & \cdots & \frac{\partial F_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial F_n}{\partial x_1} & \cdots & \frac{\partial F_n}{\partial x_n} \end{bmatrix} \begin{bmatrix} \frac{\partial \xi_1}{\partial p_1} & \cdots & \frac{\partial \xi_1}{\partial p_m} \\ \vdots & & \vdots \\ \frac{\partial \xi_n}{\partial p_1} & \cdots & \frac{\partial \xi_n}{\partial p_m} \end{bmatrix} + \begin{bmatrix} \frac{\partial F_1}{\partial p_1} & \cdots & \frac{\partial F_1}{\partial p_m} \\ \vdots & & \vdots \\ \frac{\partial F_n}{\partial p_1} & \cdots & \frac{\partial F_n}{\partial p_m} \end{bmatrix} = 0. \tag{13}$$

Provided $\begin{bmatrix} \frac{\partial F_1}{\partial x_1} & \cdots & \frac{\partial F_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial F_n}{\partial x_1} & \cdots & \frac{\partial F_n}{\partial x_n} \end{bmatrix}$ has an inverse (the hypothesis of the Implicit Function Theorem) we can solve this:

$$\begin{bmatrix} \frac{\partial \xi_1}{\partial p_1} & \cdots & \frac{\partial \xi_1}{\partial p_m} \\ \vdots & & \vdots \\ \frac{\partial \xi_n}{\partial p_1} & \cdots & \frac{\partial \xi_n}{\partial p_m} \end{bmatrix} = - \begin{bmatrix} \frac{\partial F_1}{\partial x_1} & \cdots & \frac{\partial F_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial F_n}{\partial x_1} & \cdots & \frac{\partial F_n}{\partial x_n} \end{bmatrix}^{-1} \begin{bmatrix} \frac{\partial F_1}{\partial p_1} & \cdots & \frac{\partial F_1}{\partial p_m} \\ \vdots & & \vdots \\ \frac{\partial F_n}{\partial p_1} & \cdots & \frac{\partial F_n}{\partial p_m} \end{bmatrix} \tag{14}$$

The old-fashioned derivation (see, e.g., Samuelson [25, pp. 10–14]) of this same result runs like this: "Totally differentiate" the ith row of equation (10) to get

$$\sum_k \frac{\partial F_i}{\partial x_k}\, dx_k + \sum_\ell \frac{\partial F_i}{\partial p_\ell}\, dp_\ell = 0 \tag{15}$$

for all i. Now set all dp_ℓ's equal to zero except p_j, and divide by dp_j to get

$$\sum_k \frac{\partial F_i}{\partial x_k} \frac{dx_k}{dp_j} + \frac{\partial F_i}{\partial p_j} = 0 \tag{16}$$

for all i and j, which is equivalent to equation (12). For further information on total differentials and how to manipulate them, see [3, Chapter 6].

Using Cramer's Rule (e.g. [4, pp. 93–94]), we see then that

$$\frac{dx_i}{dp_j} = \frac{\partial \xi_i}{\partial p_j} = - \frac{\begin{vmatrix} \frac{\partial F_1}{\partial x_1} & \cdots & \frac{\partial F_1}{\partial x_{i-1}} & \frac{\partial F_1}{\partial p_j} & \frac{\partial F_1}{\partial x_{i+1}} & \cdots & \frac{\partial F_1}{\partial x_n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ \frac{\partial F_n}{\partial x_1} & \cdots & \frac{\partial F_n}{\partial x_{i-1}} & \frac{\partial F_n}{\partial p_j} & \frac{\partial F_n}{\partial x_{i+1}} & \cdots & \frac{\partial F_n}{\partial x_n} \end{vmatrix}}{\begin{vmatrix} \frac{\partial F_1}{\partial x_1} & \cdots & \frac{\partial F_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial F_n}{\partial x_1} & \cdots & \frac{\partial F_n}{\partial x_n} \end{vmatrix}}. \tag{17}$$

Or, letting Δ denote the determinant of $\begin{bmatrix} \frac{\partial F_1}{\partial x_1} & \cdots & \frac{\partial F_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial F_n}{\partial x_1} & \cdots & \frac{\partial F_n}{\partial x_n} \end{bmatrix}$, and letting $\Delta_{i,j}$

denote the determinant of the matrix formed by deleting its i-th row and j-th column, we have

$$\frac{\partial \xi_i}{\partial p_j} = -\sum_{k=1}^{n} (-1)^{i+k} \frac{\partial F_k}{\partial p_j} \frac{\Delta_{k,i}}{\Delta}. \tag{18}$$

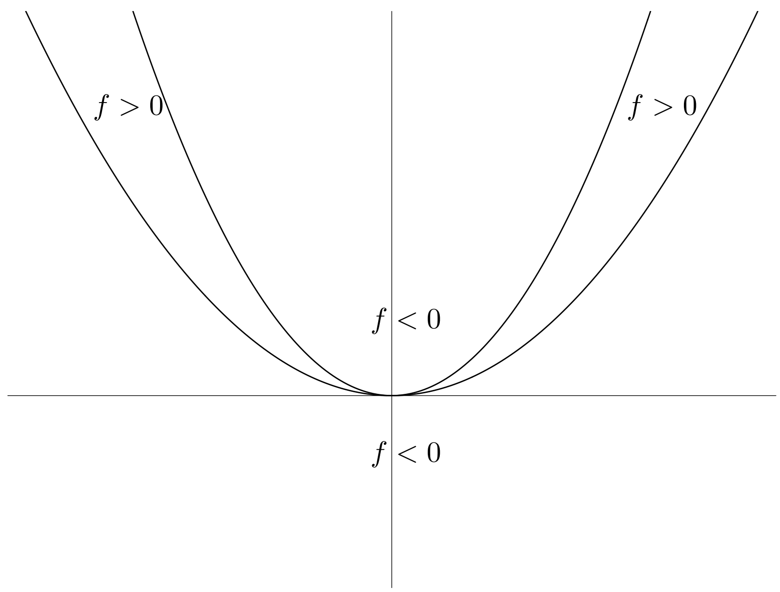

Figure 2. $f(x, y) = -(y - x^2)(y - 2x^2)$.

References

[1] S. N. Afriat. 1971. Theory of maxima and the method of Lagrange. *SIAM Journal of Applied Mathematics* 20:343–357. http://www.jstor.org/stable/2099955

[2] C. D. Aliprantis and K. C. Border. 2006. *Infinite dimensional analysis: A hitchhiker's guide*, 3d. ed. Berlin: Springer–Verlag.

[3] T. M. Apostol. 1957. *Mathematical analysis: A modern approach to advanced calculus*. Addison-Wesley series in mathematics. Reading, Massachusetts: Addison Wesley.

[4] ——— . 1969. *Calculus*, 2d. ed., volume 2. Waltham, Massachusetts: Blaisdell.

[5] K. C. Border. 1985. *Fixed point theorems with applications to economics and game theory*. New York: Cambridge University Press.

[6] J.-P. Crouzeix. 1981. Continuity and differentiability properties of quasiconvex functions on $\mathbb{R}^n$. In S. Schaible and W. Ziemba, eds., *Generalized Concavity in Optimization and Economics*, pages 109–130. NY: Academic Press.

[7] J.-P. Crouzeix and J. A. Ferland. 1982. Criteria for quasi-convexity and pseudo-convexity: Relationships and comparisons. *Mathematical Programming* 23:193–205. DOI: 10.1007/BF01583788

[8] J. Dieudonné. 1969. *Foundations of modern analysis*. Number 10-I in Pure and Applied Mathematics. New York: Academic Press. Volume 1 of Treatise on Analysis.

[9] E. E. Diewert, M. Avriel, and I. Zang. 1981. Nine kinds of quasiconcavity and concavity. *Journal of Economic Theory* 25(3):397–420. DOI: 10.1016/0022-0531(81)90039-9

[10] D. Gale and H. Nikaidô. 1965. Jacobian matrix and global univalence of mappings. *Mathematische Annalen* 159(2):81–93. DOI: 10.1007/BF01360282

[11] W. Ginsberg. 1973. Concavity and quasiconcavity in economics. *Journal of Economic Theory* 6(6):596–605. DOI: 10.1016/0022-0531(73)90080-X

[12] H. Halkin. 1972. Necessary conditions for optimal control problems with infinite horizons. Discussion Paper 7210, CORE.

[13] ——— . 1974. Implicit functions and optimization problems without continuous differentiability of the data. *SIAM Journal on Control* 12(2):229–236. DOI: 10.1137/0312017

[14] L. Hurwicz and M. K. Richter. 2003. Implicit functions and diffeomorphisms without C^1. *Advances in Mathematical Economics* 5:65–96.

[15] F. John. 1948. Extremum problems with inequalities as subsidiary conditions. In K. O. Friedrichs, O. E. Neugebauer, and J. J. Stoker, eds., *Studies and Essays: Courant Anniversary Volume*, pages 187–204. New York: Interscience.

[16] E. B. Leach. 1961. A note on inverse function theorems. *Proceedings of the American Mathematical Society* 12:694–697. http://www.jstor.org/stable/2034858

[17] ——— . 1963. On a related function theorem. *Proceedings of the American Mathematical Society* 14:687–689. http://www.jstor.org/stable/2034971

[18] A. Leroux. 1984. Other determinantal conditions for concavity and quasi-concavity. *Journal of Mathematical Economics* 13(1):43–49. DOI: 10.1016/0304-4068(84)90023-5

[19] L. H. Loomis and S. Sternberg. 1968. *Advanced calculus*. Reading, Massachusetts: Addison–Wesley.

[20] J. E. Marsden. 1974. *Elementary classical analysis*. San Francisco: W. H. Freeman and Company.

[21] A. Nijenhuis. 1974. Strong derivatives and inverse mappings. *American Mathematical Monthly* 81(9):969–980.

[22] H. Nikaidô. 1968. *Convex structures and economic theory*. Mathematics in Science and Engineering. New York: Academic Press.

[23] K. Otani. 1983. A characterization of quasi-concave functions. *Journal of Economic Theory* 31(1):194–196. DOI: 10.1016/0022-0531(83)90030-3

[24] W. Rudin. 1976. *Principles of mathematical analysis*, 3d. ed. International Series in Pure and Applied Mathematics. New York: McGraw Hill.

[25] P. A. Samuelson. 1965. *Foundations of economic analysis*. New York: Athenaeum. Reprint of the 1947 edition published by Harvard University Press.

[26] M. Spivak. 1965. *Calculus on manifolds*. Mathematics Monograph Series. New York: Benjamin.

[27] H. Uzawa. 1958. The Kuhn–Tucker conditions in concave programming. In K. J. Arrow, L. Hurwicz, and H. Uzawa, eds., *Studies in Linear and Non-linear Programming*, number 2 in Stanford Mathematical Studies in the Social Sciences, chapter 3, pages 32–37. Stanford, California: Stanford University Press.

[28] J. Warga. 1978. An implicit function theorem without differentiability. *Proceedings of the American Mathematical Society* 69(1):65–69. http://www.jstor.org/stable/2043190.pdf