ECON 6170 Section 9

TA: Patrick Ferguson

November 1, 2024

Implicit Function Theorem

Theorem 6 (Implicit function theorem)**.** *Suppose $f : X \times Y \subseteq \mathbb{R}^d \times \mathbb{R}^m \to \mathbb{R}^m$ is C^1 and $X \times Y$ is open.*[1] *Let $(x_0, y_0) \in X \times Y$ be a point at which*

(i) $f(x_0, y_0) = 0$;

(ii) $D_y f(x_0, y_0)$ *is invertible.*

Then:

(i) There exists $B_{\varepsilon_x}(x_0) \subseteq X$ and $B_{\varepsilon_y}(y_0) \subseteq Y$ such that for all $x \in B_{\varepsilon_x}(x_0)$ there exists a unique $y \in B_{\varepsilon_y}(y_0)$ such that $f(x, y) = 0$.

(ii) So there exists a function $g : B_{\varepsilon_x}(x_0) \to B_{\varepsilon_y}(y_0)$ that satisfies

(a) $g(x_0) = y_0$;

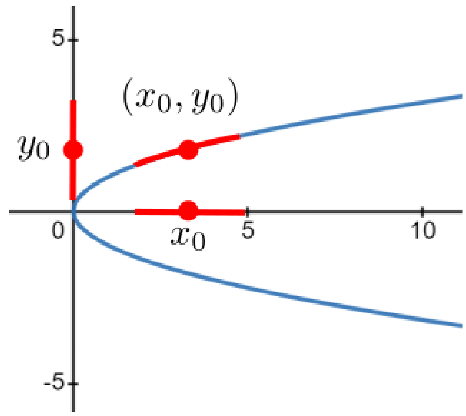
(b) $f(x, g(x)) = 0$ *for all* $x \in B_{\varepsilon_x}(x_0)$;

(c) g *is* C^1*, with derivative*

$$Dg(x) = -(D_y f(x, g(x)))^{-1} D_x f(x, g(x))$$

Example 1. Suppose $f : \mathbb{R}^2 \to \mathbb{R}$ is given by $f(x, y) := x - y^2$. The level set $\{(x, y) \mid f(x, y) = 0\}$ is shown below. The graph of the implicit function $g : B_{\varepsilon_x}(x_0) \to B_{\varepsilon_y}(y_0)$ is given by the red subset of the level set. Note that g maps *into* but not necessarily *onto* $B_{\varepsilon_y}(y_0)$, that is, there may be some $y \in B_{\varepsilon_y}(y_0)$ that are not in the range $g[B_{\varepsilon_x}(x_0)]$.

[1] To be explicit, we mean $X \subseteq \mathbb{R}^d$ and $Y \subseteq \mathbb{R}^m$.



Note also that if we took $x_0 = y_0 = 0$, then $D_y f(0,0) = \frac{\partial f(0,0)}{\partial y} = -2 \cdot 0 = 0$, violating hypothesis (ii). Indeed, for any $\varepsilon_x > 0$, any function defined on $B_{\varepsilon_x}(0)$ must violate $f(x, g(x)) = 0$ for $x < 0$. Similarly, at the graphed $(x_0, y_0) \gg 0$, we must choose ε_x small enough that it excludes $x < 0$.

Exercise 19. Prove the inverse function theorem: Suppose $f : X \subseteq \mathbb{R}^d \to Y \subseteq \mathbb{R}^d$ is C^1, $x_0 \in \operatorname{int} X$, and define $y_0 := f(x_0)$. If

(i) $Df(x_0)$ is invertible.

Then:

(i) There exists $B_{\varepsilon_x}(x_0) \subseteq X$ and $B_{\varepsilon_y}(y_0) \subseteq Y$ such that for all $y \in B_{\varepsilon_y}(y_0)$ there exists a unique $x \in B_{\varepsilon_x}(x_0)$ such that $f(x) = y$.

(ii) So there exists a function $g : B_{\varepsilon_x}(x_0) \to B_{\varepsilon_y}(y_0)$ that satisfies

(a) $(f \circ g)(y) = y$ for all $y \in B_{\varepsilon_y}(y_0)$;

(b) g is C^1, with derivative

$$Dg(y) = (Df(g(y)))^{-1}$$

Write

$$F(x, y) := y - f(x)$$

Then F is C^1, $F(x_0, y_0) = 0$, and $D_x F(x_0, y_0) = -Df(x_0)$ is invertible. It is WLOG to assume that $X \times Y$ is open.[2]

It follows that we can apply the implicit function theorem to obtain:

(i) There exists $B_{\varepsilon_y}(y_0) \subseteq Y$ and $B_{\varepsilon_x}(x_0) \subseteq X$ such that for all $y \in B_{\varepsilon_y}(y_0)$ there exists a unique $x \in B_{\varepsilon_x}(x_0)$ such that $F(x, y) = 0$. That is, $y = f(x)$.

(ii) So there exists a function $g : B_{\varepsilon_y}(y_0) \to B_{\varepsilon_x}(x_0)$ that satisfies

(a) $g(y_0) = x_0$;

(b) $F(g(y), y) = 0$ for all $y \in B_{\varepsilon_y}(y_0)$; that is, $y = (f \circ g)(y)$;

[2]Because $x_0 \in \operatorname{int} X$, Y can be extended to $\mathbb{R}^d$, and the Cartesian product of open sets is open.

(c) g is C^1, with derivative

$$\begin{aligned} Dg(y) &= -(D_x F(g(y), y))^{-1} D_y F(g(y), y) \\ &= -(D_x(y - f(g(y))))^{-1} D_y(y - f(g(y))) \\ &= -(-Df(g(y)))^{-1} I \\ &= (Df(g(y)))^{-1} \end{aligned}$$

Static Optimisation

Theorem 3 (Necessity)**.** *Let $f : \mathbb{R}^d \to \mathbb{R}$, $h_k : \mathbb{R}^d \to \mathbb{R}$, and $g_j : \mathbb{R}^d \to \mathbb{R}$ be C^1 for each $k \in \{1, \ldots, K\}$ and each $j \in \{1, \ldots, J\}$. Suppose x^* is a local maximum of f on the constraint set*

$$\Gamma := \left\{ x \in \mathbb{R}^d \mid h_k(x) = 0 \ \text{for } k = 1, \ldots, K \text{ and } g_j(x) \geq 0 \ \text{for } j = 1, \ldots, J \right\}$$

Let $E \subseteq \{1, \ldots, J\}$ denote the set of binding constraints at x^* and let $g_E := (g_j)_{j \in E}$. Suppose that

$$\operatorname{rank} \left(D \begin{bmatrix} h(x^*) \\ g_E(x^*) \end{bmatrix} \right) = K + |E| . \tag{9}$$

Then, there exists $\mu^* \in \mathbb{R}^K$ and $\lambda^* \in \mathbb{R}^J$ such that

$$\lambda_j^* \geq 0 \ \text{ for all } j \in \{1, \ldots, J\}, \tag{10}$$

$$\lambda_j^* g_j(x^*) = 0 \ \text{ for all } j \in \{1, \ldots, J\}, \tag{11}$$

$$\nabla f(x^*) + \sum_{k=1}^{K} \mu_k^* \nabla h_k(x^*) + \sum_{j=1}^{J} \lambda_j^* \nabla g_j(x^*) = 0^\mathsf{T}. \tag{12}$$

Section Exercise 1. Show that Theorems 1 and 2 are special cases of Theorem 3.

Theorem 1 is the case of only (K) equality constraints. The constraint qualification (9) then becomes

$$\operatorname{rank} Dh(x^*) = K$$

and the conclusion omits (10) and (11), and changes (12) to

$$\nabla f(x^*) + \sum_{k=1}^{K} \mu_k^* \nabla h_k(x^*) = 0^\mathsf{T}$$

Theorem 2 is the case of only (J) inequality constraints. The constraint qualification then becomes

$$\operatorname{rank} Dg_E(x^*) = |E|$$

The nonnegativity constraints (10) and complementary slackness conditions (11) are unchanged. The FOC becomes

$$\nabla f(x^*) + \sum_{j=1}^{J} \lambda_j^* \nabla g_j(x^*) = 0^\mathsf{T}$$

Section Exercise 2 (From MT3 2023 Q3)**.** Let $f : \mathbb{R}^d \to \mathbb{R}$ and $g_j : \mathbb{R}^d \to \mathbb{R}$ for $j = 1, \ldots, J$ all be C^1. Consider the following problem:

$$\max_{x \in X} f(x) \text{ st } g_j(x) \geq 0 \text{ for } j = 1, \ldots, J$$

Suppose x^* is a local maximum that satisfies the constraints.

(i) Suppose $g_1(\cdot) = g_j(\cdot)$ for $j = 2, \ldots, J$. Can the constraint qualification be satisfied? If not, what can we do?

Not in general. It follows from the question that $Dg_1(\cdot) = Dg_j(\cdot)$ for $j = 2, \ldots, J$. Moreover either all the constraints bind or none binds. It follows that if $J \geq 2$ and the constraints bind, $\operatorname{rank} Dg_E(x^*) = 1 < J = |E|$. If we remove all but the first constraint, the optimisation problem is unchanged, but the constraint qualification can be satisfied.

(ii) Suppose $g_1(\cdot) = -g_2(\cdot)$. Can the constraint qualification be satisfied? If not, what can we do?

No. We know that $Dg_1(\cdot) = -Dg_2(\cdot)$. Moreover, $g_1(x^*) \geq 0$ and $-g_1(x^*) \geq 0$ imply that $g_1(x^*) = 0$, so both constraints bind. It follows that $\operatorname{rank} Dg_E(x^*) \leq |E| - 1 < |E|$. We can resolve this by replacing the two inequality constraints with one equality constraint, $g_1(x) = 0$, and using the Theorem of Lagrange.

Section Exercise 3.

(i) Specialise Theorem 3 to the unconstrained case.

Proposition 2. *Let $f : \mathbb{R}^d \to \mathbb{R}$ be C^1. Suppose x^* is a local maximum of f on $\mathbb{R}^d$. Then,*

$$\nabla f(x^*) = 0^\mathsf{T}.$$

(ii) Let $X \subseteq \mathbb{R}^d$ be open and define $f|_X : X \to \mathbb{R}$ by $f|_X(x) = x$ for all $x \in X$. Show that $x^* \in X$ is a local maximum of $f|_X$ on X iff it is a local maximum of f on $\mathbb{R}^d$.

If X is open and $x^* \in X$, then there exists a sufficiently small $\varepsilon > 0$ such that $B := B_\varepsilon(x^*) \subseteq X$. Then x^* is a local maximum of $f|_X$ on X $\iff$ $f(x^*) \geq f(x)$ for all $x \in B_\delta(x^*) \cap X$ $\iff$ $f(x^*) \geq f(x)$ for all $x \in B_\delta(x^*)$ with $\delta < \varepsilon$ $\iff$ x^* is a local maximum of f on $\mathbb{R}^d$.

(iii) Show that it suffices in Proposition 1, that f be continuously differentiable at x^* (as opposed to everywhere in $\mathbb{R}^d$).

Note that the solution to part (ii) implies that the behaviour of f outside of $B_\delta(x^*)$ is irrelevant to whether x^* is a local maximum of f. But δ is an arbitrary positive real number. Suppose it were necessary that f be C^1 at $x \neq x^*$. Then choose $\delta < \|x - x^*\|$ to obtain a contradiction.

Exercise 1. Consider the equality-constrained problem from class notes:

$$\max_{x \in \mathbb{R}^d} f(x) \text{ st } h(x) = 0 \tag{1}$$

where $f : \mathbb{R}^d \to \mathbb{R}$ and $h_k : \mathbb{R}^d \to \mathbb{R}$, $k = 1, \ldots, K$ are all C^1. Define $\mathcal{L} : \mathbb{R}^d \times \mathbb{R}^K \to \mathbb{R}$ by

$$\mathcal{L}(x, \mu) := f(x) + \sum_{k=1}^{K} \mu_k h_k(x)$$

Let

$$S := \{(x,\mu) \mid \nabla \mathcal{L}(x,\mu) = 0\}$$

and define S_X as the projection of S onto the first d components of S, i.e.,

$$S_X := \{x \mid \text{ there exists } \mu \text{ such that } (x,\mu) \in S\}$$

Now consider the following problem:

$$\max_{x \in S_X} f(x) \tag{2}$$

(i) Show that if problem (1) attains a global maximum at $x^* \in \mathbb{R}^d$ and the constraint qualification holds at x^*, then a x° that solves (2) is also a global maximum of (1).

If problem (1) attains a global maximum at x^* and the constraint qualification holds, then by the Theorem of Lagrange, there exists $\mu^* \in \mathbb{R}^K$ such that

$$\nabla f(x^*) + \sum_{k=1}^{K} \mu_k^* \cdot \nabla h_k(x^*) = 0$$

But the left-hand side is just $\nabla_x \mathcal{L}(x^*,\mu^*)$. Moreover, the constraints imply $\nabla_\mu \mathcal{L}(x^*,\mu^*) = h(x^*) = 0$. Taken together, we have

$$\nabla \mathcal{L}(x^*,\mu^*) = 0$$

So $(x^*,\mu^*) \in S$ and $x^* \in S_X$. It follows that $f(x^\circ) \geq f(x^*)$. Moreover, $x^\circ \in S_X$ implies that there exists μ° such that $\nabla \mathcal{L}(x^\circ,\mu^\circ) = 0$. But $\nabla_\mu \mathcal{L}(x^\circ,\mu^\circ) = 0$ implies that x° satisfies the constraints. Therefore, x° is also a global maximiser for problem (1).

(ii) Show that (2) is equivalent to

$$\max_{(x,\mu) \in \mathbb{R}^d \times \mathbb{R}^K} \mathcal{L}(x,\mu) \tag{3}$$

if the latter has a solution.[3]

Let (x',μ') solve (3). Then Proposition 1 on unconstrained optimisation implies that $(x',\mu') \in S$, so $x' \in S_X$. Moreover, by definition of (x',μ'),

$$\mathcal{L}(x',\mu') = f(x') + \sum \mu_k' h_k(x') \geq f(x^\circ) + \sum \mu_k^\circ h_k(x^\circ) = \mathcal{L}(x^\circ,\mu^\circ)$$

where x° maximises (2). But $x', x^\circ \in S_X$ implies $h(x') = h(x^\circ) = 0$.It follows that

$$f(x') \geq f(x^\circ)$$

so x' is also a solution to (2). Conversely, $x' \in S_X$ and the definition of x° imply

$$f(x^\circ) \geq f(x')$$

Moreover, we know that $h(x^\circ) = h(x') = 0$ so

$$\mathcal{L}(x^\circ,\mu^\circ) = f(x^\circ) + \sum \mu_k^\circ h_k(x^\circ) \geq f(x') + \sum \mu_k' h_k(x') = \mathcal{L}(x',\mu')$$

for any μ°. It follows that (x°,μ°) solves (3).

[3] The text in red has been added—the result does not go through in its absence. Thank you to Wanxi for pointing this out.

Exercise 2. Consider the problem

$$\max_{x,y} f(x,y) \text{ st } h(x,y) = 0$$

where $f(x,y) := -y$ and $h(x,y) := y^3 - x^2$. Show that the unique solution to the problem is at 0; that the constraint qualification is violated at 0; and that there does not exist $\mu \in \mathbb{R}$ satisfying

$$\nabla f(x^*) + \sum_{k=1}^{K} \mu_k^* \nabla h_k(x^*) = 0$$

The equation $y^3 - x^2 = 0$ is equivalent to $y^3 = x^2$. In particular, this implies that $y \geq 0$. Maximising $-y$ is equivalent to minimising y, which is achieved by choosing $y = 0$. The constraint then implies that the optimal $x = y^{3/2} = 0$.

$$Dh(0,0) = \begin{bmatrix} \frac{\partial h(0,0)}{\partial x} & \frac{\partial h(0,0)}{\partial y} \end{bmatrix} = \begin{bmatrix} -2 \cdot 0 & 3 \cdot 0^2 \end{bmatrix} = \begin{bmatrix} 0 & 0 \end{bmatrix}$$

and the constraint qualification is that rank $Dh(x,y) = 1$. The rank of a matrix is the maximal number of its rows (or columns) that can comprise a linearly independent set. Here, we have one row, which is a zero vector, and the set $\{0\}$ is not linearly independent. Therefore rank $Dh(x,y) = 0$, violating the constraint qualification.

Note also that for any $\mu \in \mathbb{R}$,

$$\nabla f(0,0) + \mu \nabla h(0,0) = \begin{bmatrix} 0 & -1 \end{bmatrix} + \mu \begin{bmatrix} 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & -1 \end{bmatrix} \neq \begin{bmatrix} 0 & 0 \end{bmatrix}$$