ECON 6170

Problem Set 9

Gabe Sekeres

November 6, 2024

Worked with Fenglin Ye on Exercise 19.

1 Exercises from Class Notes

Exercise 18. Let $F : \mathbb{R}^4 \to \mathbb{R}^2$. Suppose the conditions for the Implicit Function Theorem are satisfied at all points, and that $F(x_1^\star, x_2^\star, y_1^\star, y_2^\star) = 0$. Let $h = (h_1, h_2)$ denote the implicitly defined function of (x_1, x_2) for the relation $F(x_1, x_2, y_1, y_2) = (0, 0)$ near $(x_1^\star, x_2^\star, y_1^\star, y_2^\star)$. Give explicit formulae for $\frac{\partial h_i}{\partial x_j}$, for $i, j \in \{1, 2\}$.
Solution. We have from the Implicit Function Theorem that $h = (h_1, h_2)$ is differentiable at all (x_1, x_2) sufficiently close to $(x_1^\star, x_2^\star)$, and that

$$Dh(x) = -(D_yF(x, h(x)))^{-1} D_xF(x, h(x))$$

Recall that Dh, D_yF, and D_xF are all 2×2 matrices of partial derivatives. In expanded terms, we have that at some $x = (x_1, x_2)$,

$$\begin{bmatrix} \frac{\partial h_1}{\partial x_1}(x) & \frac{\partial h_1}{\partial x_2}(x) \\ \frac{\partial h_2}{\partial x_1}(x) & \frac{\partial h_2}{\partial x_2}(x) \end{bmatrix} = - \begin{bmatrix} \frac{\partial F_1}{\partial h_1(x)}(x) & \frac{\partial F_1}{\partial h_2(x)}(x) \\ \frac{\partial F_2}{\partial h_1(x)}(x) & \frac{\partial F_2}{\partial h_2(x)}(x) \end{bmatrix}^{-1} \begin{bmatrix} \frac{\partial F_1}{\partial x_1}(x) & \frac{\partial F_1}{\partial x_2}(x) \\ \frac{\partial F_2}{\partial x_1}(x) & \frac{\partial F_2}{\partial x_2}(x) \end{bmatrix}$$

Recall that the inverse of a 2×2 matrix has a closed form:

$$\begin{bmatrix} \frac{\partial F_1}{\partial h_1(x)}(x) & \frac{\partial F_1}{\partial h_2(x)}(x) \\ \frac{\partial F_2}{\partial h_1(x)}(x) & \frac{\partial F_2}{\partial h_2(x)}(x) \end{bmatrix}^{-1} = \frac{1}{\frac{\partial F_2}{\partial h_2(x)}(x)\frac{\partial F_1}{\partial h_1(x)}(x) - \frac{\partial F_1}{\partial h_2(x)}(x)\frac{\partial F_2}{\partial h_1(x)}(x)} \begin{bmatrix} \frac{\partial F_2}{\partial h_2(x)}(x) & -\frac{\partial F_1}{\partial h_2(x)}(x) \\ -\frac{\partial F_2}{\partial h_1(x)}(x) & \frac{\partial F_1}{\partial h_1(x)}(x) \end{bmatrix}$$

Taking the matrix inner product of $(D_yF(x, h(x)))^{-1}$ and $D_xF(x, h(x))$, we get

$$\begin{bmatrix} \frac{\partial F_2}{\partial h_2(x)}(x)\frac{\partial F_1}{\partial x_1}(x) - \frac{\partial F_1}{\partial h_2(x)}(x)\frac{\partial F_2}{\partial x_1}(x) & \frac{\partial F_2}{\partial h_2(x)}(x)\frac{\partial F_1}{\partial x_2}(x) - \frac{\partial F_1}{\partial h_2(x)}(x)\frac{\partial F_2}{\partial x_2}(x) \\ \frac{\partial F_1}{\partial h_1(x)}(x)\frac{\partial F_2}{\partial x_1}(x) - \frac{\partial F_2}{\partial h_1(x)}(x)\frac{\partial F_1}{\partial x_1}(x) & \frac{\partial F_1}{\partial h_1(x)}(x)\frac{\partial F_2}{\partial x_2}(x) - \frac{\partial F_2}{\partial h_1(x)}(x)\frac{\partial F_1}{\partial x_2}(x) \end{bmatrix}$$

Thus, we have that the respective partial derivatives are

$$\frac{\partial h_1}{\partial x_1}(x) = -\frac{\frac{\partial F_2}{\partial h_2(x)}(x)\frac{\partial F_1}{\partial x_1}(x) - \frac{\partial F_1}{\partial h_2(x)}(x)\frac{\partial F_2}{\partial x_1}(x)}{\frac{\partial F_2}{\partial h_2(x)}(x)\frac{\partial F_1}{\partial h_1(x)}(x) - \frac{\partial F_1}{\partial h_2(x)}(x)\frac{\partial F_2}{\partial h_1(x)}(x)}$$

$$\frac{\partial h_1}{\partial x_2}(x) = -\frac{\frac{\partial F_2}{\partial h_2(x)}(x)\frac{\partial F_1}{\partial x_2}(x) - \frac{\partial F_1}{\partial h_2(x)}(x)\frac{\partial F_2}{\partial x_2}(x)}{\frac{\partial F_2}{\partial h_2(x)}(x)\frac{\partial F_1}{\partial h_1(x)}(x) - \frac{\partial F_1}{\partial h_2(x)}(x)\frac{\partial F_2}{\partial h_1(x)}(x)}$$

$$\frac{\partial h_2}{\partial x_1}(x) = -\frac{\frac{\partial F_1}{\partial h_1(x)}(x)\frac{\partial F_2}{\partial x_1}(x) - \frac{\partial F_2}{\partial h_1(x)}(x)\frac{\partial F_1}{\partial x_1}(x)}{\frac{\partial F_2}{\partial h_2(x)}(x)\frac{\partial F_1}{\partial h_1(x)}(x) - \frac{\partial F_1}{\partial h_2(x)}(x)\frac{\partial F_2}{\partial h_1(x)}(x)}$$

$$\frac{\partial h_2}{\partial x_2}(x) = -\frac{\frac{\partial F_1}{\partial h_1(x)}(x)\frac{\partial F_2}{\partial x_2}(x) - \frac{\partial F_2}{\partial h_1(x)}(x)\frac{\partial F_1}{\partial x_2}(x)}{\frac{\partial F_2}{\partial h_2(x)}(x)\frac{\partial F_1}{\partial h_1(x)}(x) - \frac{\partial F_1}{\partial h_2(x)}(x)\frac{\partial F_2}{\partial h_1(x)}(x)}$$

I refuse to simplify these on principle (the principle is mostly laziness).

Exercise 19. Prove the Inverse Function Theorem:

Theorem 1. *Suppose $f : X \subseteq \mathbb{R}^d \to Y \subseteq \mathbb{R}^d$. Let $x_0 \in \text{int}(X)$ and define $y_0 = f(x_0)$. If $f \in C^1$ and $Df(x_0)$ is invertible, then there exists an open ball $B_{\varepsilon_X}(x_0) \subseteq X$ and an open ball $B_{\varepsilon_Y}(y_0) \subseteq Y$ such that for all $y \in B_{\varepsilon_Y}(y_0)$ there exists a unique $x \in B_{\varepsilon_X}(x_0)$ such that $f(x) = y$. Therefore, the equation $f(x) = y$ implicitly defines a function $g : B_{\varepsilon_Y}(y_0) \to B_{\varepsilon_X}(x_0)$ with the property*

$$f(g(y)) = y \;\forall\; y \in B_{\varepsilon_Y}(y_0)$$

Moreover, g is differentiable at any $y \in B_{\varepsilon_Y}(y_0)$ and

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

Proof. Define a function $F : \text{int}\, Y \times \text{int}\, X \to \mathbb{R}$ by $F(y, x) = y - f(x)$. Note that this function meets the conditions of the Implicit Function Theorem, as f is continuous, and the Cartesian product of open sets is open. Moreover, at (y_0, x_0), we have that $F(y_0, x_0) = y_0 - f(x_0) = 0$. Since $Df(x_0)$ is invertible, we have that by the Implicit Function Theorem, there exists an open ball $B_{\varepsilon_X}(x_0) \subseteq \text{int}\, X$ and an open ball $B_{\varepsilon_Y}(y_0) \subseteq \text{int}\, Y$ such that for each $x \in B_{\varepsilon_X}(x_0)$, there exists a unique $y \in B_{\varepsilon_Y}(y_0)$ such that $F(y, x) = 0 \implies f(x) = y$.

We have that the implicitly defined function implies that $y - f(g(y)) = 0$ for some $g : B_{\varepsilon_Y}(y_0) \to B_{\varepsilon_X}(x_0)$. From the fact that $y - f(f^{-1}(y)) = 0$, we get that $g(y) = f^{-1}(y)$. Note that $f(g(y)) = f(f^{-1}(y)) = y \;\forall\; y \in B_{\varepsilon_Y}(y_0)$.

Finally, from the Implicit Function Theorem, we have that

$$Dg(y) = -(D_x F(y, g(y)))^{-1} D_y F(y, g(y)) = -(Df(g(y)))^{-1} \cdot -1 = (Df(g(y)))^{-1}$$

$\square$

Exercise 3. Prove the following: Suppose f is $\mathbf{C}^2$ on X, where $\text{int}(X)$ is convex, and that f is concave. Fix $x^\star \in \text{int}(X)$. The following are equivalent:

(i) $\nabla f(x^\star) = 0$

(ii) f has a local maximum at $x^\star$

(iii) f has a global maximum at $x^\star$

Proof. (i) $\implies$ (ii): Since $\text{int}(X)$ is definitionally open, we can find some $\varepsilon > 0$ such that $B_\varepsilon(x^\star) \subseteq \text{int}(X)$. Consider some $y \in B_\varepsilon(x^\star)$, where $y \neq x$. From Proposition 14 in the notes, we have that f being concave implies that

$$\nabla f(x^\star)(y - x^\star) \geq f(y) - f(x^\star) \implies 0 \geq f(y) - f(x^\star) \implies f(x^\star) \geq f(y)$$

Thus, $x^\star$ is a local maximum of f.

(ii) $\implies$ (iii): Fix $\varepsilon > 0$ such that $x^\star \geq y \;\forall\; y \in B_\varepsilon(x^\star)$. Consider some $x' \in X$. Since $\text{int}(X)$ is convex, $\alpha x^\star + (1-\alpha)x' \in \text{int}(X)$ for all $\alpha \in (0, 1)$. Additionally, there exists α sufficiently close to 1 that $\alpha x^\star + (1 - \alpha)x' \in B_\varepsilon(x^\star)$. Since f is concave, it is also quasiconcave, and we have that

$$f(\alpha x^\star + (1 - \alpha)x') \geq \min\{f(x^\star), f(x')\}$$

Since $\alpha x^\star + (1-\alpha)x' \in B_\varepsilon(x^\star)$, we have that $f(x^\star) \geq f(\alpha x^\star + (1-\alpha)x')$, so it must be the case that $f(x^\star) \geq f(x')$, and thus $x^\star$ is a global maximum of f.

(iii) $\implies$ (i): Take some sequence $\{y_n\} \in \text{int}(X)$ such that $y_n \to x^\star$, and where $y_n \ll x^\star$ for all $n \in \mathbb{N}$. Take another sequence $\{z_n\}\,\text{int}(X)$ such that $z_n \to x^\star$, and where $z_n \gg x^\star$ for all $n \in \mathbb{N}$. Note that these sequences exist because $\text{int}(X)$ is open, meaning that there exists ε such that $B_\varepsilon(x^\star) \subseteq \text{int}(X)$. Since $x^\star$ is a global maximum, we have that $f(y_n) - f(x^\star) \leq 0 \; \forall \; n \in \mathbb{N}$, and that $f(z_n) - f(x^\star) \leq 0 \; \forall \; n \in \mathbb{N}$. Since f is concave, we have that

$$f(y_n) - f(x^\star) \leq \nabla f(x^\star)(y_n - x^\star) \; \forall \; n \in \mathbb{N}$$

and

$$f(z_n) - f(x^\star) \leq \nabla f(x^\star)(z_n - x^\star) \; \forall \; n \in \mathbb{N}$$

Taking the limit, we get that from the definition of gradients

$$\lim_{n\to\infty} \frac{f(y_n) - f(x^\star)}{y_n - x^\star} = \nabla f(x^\star) = \lim_{n\to\infty} \frac{f(z_n) - f(x^\star)}{z_n - x^\star}$$

where all the terms of the left limit are positive and all the terms of the right limit are negative, since $y_n - x^\star < 0$ and $z_n - x^\star > 0$. Thus, $\nabla f(x^\star) = 0$. □

2 Additional Exercises

Exercise 1. Consider the following problem:

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

(i) ***Proof.***

We have that $x^\star$ is a global maximum of $f(x)$ on the constraint set $h(x) = 0$. Since the constraint qualification holds, and since global maxima are also local maxima, we have that there exists $\mu^\star \in \mathbb{R}^K$ such that

$$\nabla f(x) + \sum_{k=1}^{K} \mu_k^\star \nabla h_k(x) = 0 \implies \nabla \mathcal{L}(x^\star, \mu^\star) = 0$$

Thus, the set S_X is nonempty, and $x^\star \in S_X$. Say that we have some $x^\circ \in S_X$ such that $x^\circ = \max_{x \in S_X} f(x)$. It must be the case that $x^\circ \geq x^\star$ because $x^\star$ is not necessarily the maximum. Further, since $\exists\; \mu^\circ$ s.t. $(x^\circ, \mu^\circ) \in S \implies \nabla \mathcal{L}(x^\circ, \mu^\circ) = 0 \implies h_k(x^\circ) = 0 \; \forall \; k \in \{1, \ldots, K\} \implies h(x) = 0$. Thus, x° is feasible in the primal problem, and since $x^\circ \geq x^\star$, x° is a global maximum. □

(ii) ***Proof.*** Take some $x^\star$ such that $f(x^\star) = \max_{x \in S_X} f(x)$. It must also be the case that (by (i)) $f(x^\star) = \max f(x)$ s.t. $h(x) = 0$, which means that $\mathcal{L}(x^\star, \mu) = f(x) + \sum_{k=1}^{K} \mu_k \cdot 0$ for any $\mu \in \mathbb{R}^K$. Thus, it follows that $f(x^\star) = \max_{x \in S_X} f(x) \leq \max_{(x,\mu) \in \mathbb{R}^d \times \mathbb{R}^K} \mathcal{L}(x, \mu)$. Next, take some $(x', \mu') \in \mathbb{R}^d \times \mathbb{R}^K$ such that $\mathcal{L}(x', \mu') = \max_{(x,\mu) \in \mathbb{R}^d \times \mathbb{R}^K} \mathcal{L}(x, \mu)$. Since (x', μ') is a maximizer, we must have that $\nabla \mathcal{L}(x', \mu') = 0$, which implies that $(x', \mu') \in S$. Thus, $x' \in S_X$, and $f(x') \leq \max_{x \in S_X} f(x)$. Thus, we have that $\max_{x \in S_X} f(x) \equiv \max_{(x,\mu) \in \mathbb{R}^d \times \mathbb{R}^K} \mathcal{L}(x, \mu)$. □

Exercise 2. Consider the following problem:

$$\max_{(x,y) \in \mathbb{R}^2} -y \qquad \text{s.t. } y^3 - x^2 = 0$$

Note that the constraint implies that at any optimal solution, we will have that $y^3 = x^2 \implies y \geq 0$, where $y > 0$ at any $y \neq 0$. Thus, the problem is maximized when $y = 0$, and the constraint is satisfied only when $x = 0$, implying that the unique solution is at $(0, 0)$.

Note that since $2 > 1$, the constraint qualification does not hold. We have that

$$\nabla\mathcal{L}(x^\star, y^\star, \mu) = \nabla f(0) + \mu \cdot \nabla h(0,0) = 0$$

which becomes

$$-1 + \mu(3 \cdot 0 - 2 \cdot 0) = 0 \Longrightarrow -1 = 0$$

Thus, this equation has no solutions.

Exercise 3. Let $f(x,y) = \frac{1}{3}x^3 - \frac{3}{2}y^2 + 2x$ and $g(x,y) = x - y$. We have that the constraint qualification is

$$\text{rank}(Dg(x)) = \text{rank}\begin{bmatrix} 1 & -1 \end{bmatrix} = 1 = K$$

Thus, the constraint qualification holds everywhere. We will solve

$$\max_{(x,y,\mu)\in\mathbb{R}^2\times\mathbb{R}} \frac{1}{3}x^3 - \frac{3}{2}y^2 + 2x + \mu(x-y)$$

We get the first order condition

$$\nabla\mathcal{L}(x,y,\mu) = \begin{bmatrix} x^2 + 2 + \mu & -3y - \mu & x - y \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \end{bmatrix}$$

These imply that $x = y$, and that $x^2 + 2 = 3y$. These imply that

$$(x^\star, y^\star, \mu^\star) \in \{(2,2,-6), (1,1,-3)\}$$

which have the attained values:

$$\mathcal{L}(2,2,-6) = \frac{2}{3} \qquad \text{and} \qquad \mathcal{L}(1,1,-3) = \frac{5}{6}$$

However, note that $(x,y) = (4,4)$ is attainable in the primal problem, and $f(4,4) = \frac{16}{3} > f(-2,-2), f(-1,-1)$. Thus, these solutions to (4) are not solutions to (1).

Exercise 4. Exercises 2 and 3 tell us that we have to be extremely careful trying to solve a primal problem using the Lagrangian. In Exercise 2, the constraint qualification was violated and there existed no solutions to the Lagrangian. In Exercise 3, the function we optimized over was convex, rather than concave, so the Lagrangian did not identify local maxima, instead identifying local minima. The function actually went to infinity as x, y increased, so it does not attain a maximum in the constraint set.