EC9A0: Pre-sessional Advanced Mathematics Course

Constrained Optimisation: Equality Constraints

Pablo F. Beker
Department of Economics
University of Warwick

September 2020

Lecture Outline

Introduction

- Suppose $D \subseteq \mathbb{R}^K$, K finite, is open.
- $f : D \to \mathbb{R}$
- $g : D \to \mathbb{R}^J$, with $J \leq K$.
- We would like to solve:

$$\max_{x \in D} f(x) \text{ s.t. } g(x) = 0, \tag{1}$$

- In the previous notation, one wants to find

$$\max_{x \in D'} f(x)$$

  where $D' = \{x \in D | g(x) = 0\}$.
- We will analyse when the Lagrangean method can be used.
- We will derive necessary and sufficient conditions for a constrained global maximum.

Pseudo-Theorem

- The method that is usually applied consists of the following steps:
  1. Defining the Lagrangean function $\mathcal{L} : D \times \mathbb{R}^J \to \mathbb{R}$, by
  $$\mathcal{L}(x, \lambda) = f(x) + \sum_{j=1}^{J} \lambda_j g_j(x);$$
  2. Finding $(x^*, \lambda^*) \in D \times \mathbb{R}^J$ such that $D\mathcal{L}(x^*, \lambda^*) = 0$.
- That is, a recipe is applied as though there is a "Theorem" that states:

Let $f : D \to \mathbb{R}$ and $g : D \to \mathbb{R}^J$ be differentiable. Then $x^* \in D$ solves Problem (1) if and only if there exists $\lambda^* \in \mathbb{R}^J$ such that (x^*, λ^*) solves:

$$\frac{\partial f(x^*)}{\partial x_i} + \sum_{j=1}^{J} \lambda_j^* \frac{\partial g(x^*)}{\partial x_i} = 0, \text{ for all } i = 1, ..., K.$$

Countexample

- $f(x_1, x_2) = x_1 x_2$ and $g(x_1, x_2) = (1 - x_1 - x_2)^3$.

$$x^* \text{ solves } \max_{x \in \mathbb{R}^2} f(x) \; s.t. \; g(x) = 0 \Leftrightarrow x^* \text{ solves } \max_{x \in \mathbb{R}^2_+} f(x) \; s.t. \; g(x) = 0.$$

- The second problem has a solution by Weierstrass Theorem.
- The unique maximiser is $(x_1^*, x_2^*) = (\frac{1}{2}, \frac{1}{2})$.
- According to the "Theorem" there is λ^* such that $(x_1^*, x_2^*, \lambda^*)$ solves:

$$\begin{aligned}
(a) \qquad & \frac{\partial \mathcal{L}}{\partial x_1} = 0 \quad \Leftrightarrow \quad x_2 - 3\lambda(1 - x_1 - x_2)^2 = 0 \\
(b) \qquad & \frac{\partial \mathcal{L}}{\partial x_2} = 0 \quad \Leftrightarrow \quad x_1 - 3\lambda(1 - x_1 - x_2)^2 = 0 \\
(c) \qquad & \frac{\partial \mathcal{L}}{\partial \lambda} = 0 \quad \Leftrightarrow \quad (1 - x_1 - x_2)^3 = 0
\end{aligned}$$

- A solution to this system of equations does not exist.
- Equation (c) implies that at any solution it must be the case that $x_1^* + x_2^* = 1$.
- (a) and (b) imply that both x_1^* and x_2^* are zero, a contradiction.

Intuitive Argument

- Suppose $D = \mathbb{R}^2$ and $J = 1$, Given $f : \mathbb{R}^2 \to \mathbb{R}$ and $g : \mathbb{R}^2 \to \mathbb{R}$.
- We want to solve

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

- Suppose:
  - A1 There is $h : \mathbb{R} \to \mathbb{R}$ such that $g(x, y) = 0$ if and only if $y = h(x)$.
  - A2 The function h is differentiable.
- A "crude" method would be to study the unconstrained problem

$$\max_{x\in\mathbb{R}} F(x), \qquad \text{(P}^*\text{)}$$

where $F : \mathbb{R} \to \mathbb{R}$ is defined by $F(x) = f(x, h(x))$.

Intuitive argument

- $g(x, h(x)) = 0 \Rightarrow g'_x(x, h(x)) + g'_y(x, h(x))h'(x) = 0,$
- $h'(x) = -\frac{g'_x(x,h(x))}{g'_y(x,h(x))}.$
- Apply FONC to (P*): x^* solves $\max_{x \in \mathbb{R}} F(x)$ only if $F'(x^*) = 0$.

$$\begin{aligned} f'_x(x^*, h(x^*)) + f'_y(x^*, h(x^*))h'(x^*) &= 0, \\ &\Updownarrow \\ f'_x(x^*, h(x^*)) - f'_y(x^*, h(x^*))\frac{g'_x(x^*,h(x^*))}{g'_y(x^*,h(x^*))} &= 0. \end{aligned}$$

- Define $y^* = h(x^*)$ and $\lambda^* = -\frac{\partial_y f(x^*,y^*)}{\partial_y g(x^*,y^*)} \in \mathbb{R}$,
- Then, we get that (x^*, y^*, λ^*) solves

$$f'_x(x^*, y^*) + \lambda^* g'_x(x^*, y^*) = 0,$$
$$f'_y(x^*, y^*) + \lambda^* g'_y(x^*, y^*) = 0.$$

Intuitive argument

- The "crude" method has shown that:

  Let $f : D \to \mathbb{R}$ and $g : D \to \mathbb{R}^J$ be differentiable and (A1)-(A2) hold. If $x^* \in D$ is a local maximiser in (1), there exists $\lambda^* \in \mathbb{R}^J$ such that (x^*, λ^*) solves:

$$\frac{\partial f(x^*)}{\partial x_i} + \sum_{j=1}^{J} \lambda_j^* \frac{\partial g(x^*)}{\partial x_i} = 0, \text{ for all } i = 1, ..., K.$$

- Under what conditions (A1) and (A2) hold?
- Under what conditions h exists and is differentiable?

Implicit Function Theorem

- We assumed h exists and
- We assumed $g'_y(x^*, y^*) \neq 0$. Of course, $g_x(x^*, y^*) \neq 0$ would be enough.
- What we actually require is $Dg(x^*, y^*)$ has rank 1, its maximum possible.
- Is this a general result, or does it only work in our simplified case?

Theorem The Implicit Function Theorem

Let $D \subseteq \mathbb{R}^K$ and let $g : D \to \mathbb{R}^J \in \mathbb{C}^1$, with $J < K$. If $y^* \in \mathbb{R}^J$ and $(x^*, y^*) \in D$ is such that $\text{rank}(D_y g(x^*, y^*)) = J$, then there exist $\varepsilon, \delta > 0$ and $h : B_\varepsilon(x^*) \to B_\delta(y^*) \in \mathbb{C}^1$ such that:

1. $\forall x \in B_\varepsilon(x^*)$, $(x, h(x)) \in D$;
2. $\forall x \in B_\varepsilon(x^*)$, $g(x, y) = g(x^*, y^*)$ *for* $y \in B_\delta(y^*)$ *iff* $y = h(x)$;
3. $\forall x \in B_\varepsilon(x^*)$, $Dh(x) = -D_y g(x, h(x))^{-1} D_x g(x, h(x))$.

First Order Necessary Conditions

Theorem

Let $f : D \to \mathbb{R}$ and $g : D \to \mathbb{R}^J$ be $\mathbb{C}^1$. If $x^* \in D$ is a local maximiser in (1) and $\text{rank}(Dg(x^*)) = J$, there exists $\lambda^* \in \mathbb{R}^J$ such that (x^*, λ^*) solves:

$$\begin{aligned} \frac{\partial f(x^*)}{\partial x_i} + \sum_{j=1}^{J} \lambda_j^* \frac{\partial g(x^*)}{\partial x_i} &= 0, \text{ for all } i = 1, ..., K. \\ g_j(x^*) &= 0 \text{ for all } j = 1, ..., J. \end{aligned}$$

Second Order Necessary Conditions

- The SONC for problem (P^*) is that $F''(x^*) \leq 0$. Note that:

$$F''(x) = f'_{xx}(x, h(x)) + [f'_{xy}(x, h(x)) + f'_{yx}(x, h(x))]h'(x) + f'_{yy}(x, h(x))h'(x)^2 \\ + f'_y(x, h(x))h''(x)),$$

$$h''(x) = -\frac{\partial}{\partial x}\left(\frac{g_x(x,h(x))}{g_y(x,h(x))}\right) = -\frac{1}{g_y(x,h(x))}\begin{bmatrix} 1 & h'(x) \end{bmatrix} D^2 g(x, h(x)) \begin{bmatrix} 1 \\ h'(x) \end{bmatrix}$$

- Substituting h'' and writing in matrix form, $F'' \leq 0$ becomes

$$[1\ h'(x)]D^2 f(x, h(x))\begin{bmatrix} 1 \\ h'(x) \end{bmatrix} - \frac{f'_y(x,h(x))}{g'_y(x,h(x))}[1\ h'(x)]D^2 g(x, h(x))\begin{bmatrix} 1 \\ h'(x) \end{bmatrix} \leq 0$$

$$\Leftrightarrow \begin{pmatrix} 1 & h'(x^*) \end{pmatrix} D^2_{(x,y)}\mathcal{L}(x^*, y^*, \lambda^*) \begin{pmatrix} 1 \\ h'(x^*) \end{pmatrix} \leq 0.$$

Second Order Necessary Conditions

- This condition is satisfied if $D^2_{(x,y)}\mathcal{L}(x^*, y^*, \lambda^*)$ is negative semi-definite.
- Notice that

$$\begin{pmatrix} 1 & h'(x^*) \end{pmatrix} \cdot Dg(x^*, y^*) = 0,$$

so it suffices that we guarantee that for every $\Delta \in \mathbb{R}^2 \setminus \{0\}$ such that $\Delta \cdot Dg(x^*) = 0$ we have that $\Delta^\top D^2_{(x,y)}\mathcal{L}(x^*, y^*, \lambda^*)\Delta \leq 0$.

- So, in summary, we have argued that:

Let $f : D \to \mathbb{R}$ *and* $g : D \to \mathbb{R}^J$ *be* $\mathbb{C}^1$*. If* $x^* \in D$ *is a local maximiser in* (1) *and rank*$(Dg(x^*)) = J$*, then* $\Delta^\top D^2_{xx}\mathcal{L}(x^*, \lambda^*)\Delta \leq 0$ *for all* $\Delta \in \mathbb{R}^2 \setminus \{0\}$ *such that* $\Delta \cdot Dg(x^*) = 0$*.*

First and Second Order Necessary Conditions

Theorem Lagrange - FONC and SONC

Let $f : D \to \mathbb{R}$ and $g : D \to \mathbb{R}^J$ be $\mathbb{C}^2$ with $J \leq K$. If $x^* \in D$ is a local maximiser in (1) and $rank(Dg(x^*)) = J$, then there exists $\lambda^* \in \mathbb{R}^J$ such that

1. $D_{(x,\lambda)}\mathcal{L}(x^*, \lambda^*) = 0.$
2. $\Delta^\top D^2_{xx}\mathcal{L}(x^*, \lambda^*)\Delta \leq 0$ *for all* $\Delta \in \mathbb{R}^J \setminus \{0\}$ *satisfying* $\Delta \cdot Dg(x^*) = 0;$

Necessary Conditions are not Sufficient

- The existence of $(x^*, \lambda^*) \in \mathbb{R}^K \times \mathbb{R}^J$ such that

$$\begin{aligned} \frac{\partial f(x^*)}{\partial x_i} + \sum_{j=1}^{J} \lambda_j^* \frac{\partial g_j(x^*)}{\partial x_i} &= 0, \text{ for all } i = 1, ..., K. \\ g_j(x^*) &= 0 \text{ for all } j = 1, ..., J. \end{aligned}$$

might not be sufficient for x^* to be a local maximiser of Problem 1.

Counterexample

- $f(x_1, x_2) = -(\frac{1}{2} - x_1)^3$ and $g(x_1, x_2) = 1 - x_1 - x_2$.
- $(x_1^*, x_2^*, \lambda^*) = (\frac{1}{2}, \frac{1}{2}, 0)$ satisfies the constraint qualification, it solves

$$
\begin{array}{llll}
(a) & \frac{\partial \mathcal{L}}{\partial x_1} = 0 & \Longleftrightarrow & 3\left(\frac{1}{2} - x_1\right)^2 - \lambda = 0 \\
(b) & \frac{\partial \mathcal{L}}{\partial x_2} = 0 & \Longleftrightarrow & -\lambda = 0 \\
(c) & \frac{\partial \mathcal{L}}{\partial \lambda} = 0 & \Longleftrightarrow & 1 - x_1 - x_2 = 0
\end{array}
$$

  and satisfies the (necessary) second order condition since

$$
\begin{aligned}
\frac{\partial \mathcal{L}}{\partial x_i, x_i}(x_1^*, x_2^*, \lambda^*) &= 0, \textit{for } \mathrm{i} = 1, 2 \\
\frac{\partial \mathcal{L}}{\partial x_i, x_j}(x_1^*, x_2^*, \lambda^*) &= 0, \textit{for } \mathrm{i} \neq j.
\end{aligned}
$$

- However, $(x_1^*, x_2^*) = (\frac{1}{2}, \frac{1}{2})$ is not a local maximiser since $f(\frac{1}{2}, \frac{1}{2}) = 0$ but $(\frac{1}{2} + \varepsilon, \frac{1}{2} - \varepsilon)$ is also in the constrained set and $f(\frac{1}{2} + \varepsilon, \frac{1}{2} - \varepsilon) > 0$ for any $\varepsilon > 0$.

First and Second Order Sufficient Conditions

Theorem Lagrange - FOSC and SOSC

Let $f : D \to \mathbb{R}$ *and* $g : D \to \mathbb{R}^J$ *be* $\mathbb{C}^2$*, with* $J \leq K$*. If* $(x^*, \lambda^*) \in \mathbb{R}^K \times \mathbb{R}^J$ *satisfy:*

1. $D_{(x,\lambda)}\mathcal{L}(x^*, \lambda^*) = 0$ *and*
2. $\Delta^\top D^2_{xx}\mathcal{L}(x^*, \lambda^*)\Delta < 0$ *for all* $\Delta \in \left\{\mathbb{R}^J \setminus \{0\} : \Delta \cdot Dg(x^*) = 0\right\}$.

Then, x^* *is a local maximiser in Problem* (1)*.*