EC9A0: Pre-sessional Advanced Mathematics Course

Constrained Optimisation: Inequality Constraints

Pablo F. Beker
Department of Economics
University of Warwick

September 2023

Lecture Outline

1. Constrained Optimisation with Inequality Constraints
   - Introduction
   - Counter-Example
   - Kühn-Tucker Theorem
   - Sufficient Conditions
   - Example
   - Quasi-Concave Problems

Introduction

- Suppose $f : \mathbb{R} \to \mathbb{R}$ is differentiable, $(a, b) \in \mathbb{R}$ and $a < b$.
- We would like to solve the problem:

$$\max f(x) : x \geq a \text{ and } x \leq b. \tag{1}$$

- If $x^* \in (a, b)$ solves (1), x^* is a local maximizer of f and $f'(x^*) = 0$.
- If $x^* = b$ solves (1), $f'(x^*) \geq 0$.
- If $x^* = a$ solves (1), $f'(x^*) \leq 0$.
- Thus, if x^* solves the problem, there exist $\lambda_a^*, \lambda_b^* \in \mathbb{R}_+$ such that:

$$\begin{aligned} f'(x^*) - \lambda_b^* + \lambda_a^* &= 0, \\ \lambda_a^*(x^* - a) &= 0, \\ \lambda_b^*(b - x^*) &= 0. \end{aligned}$$

- It is customary to define a function $\mathcal{L} : \mathbb{R}^3 \to \mathbb{R}$ by

$$\mathcal{L}(x, \lambda_a, \lambda_b) = f(x) + \lambda_b(b - x) + \lambda_a(x - a),$$

called *the Lagrangean*, and with which the FOC can be re-written as

$$\frac{\partial \mathcal{L}}{\partial x}(x^*, \lambda_a^*, \lambda_b^*) = 0.$$

Introduction

- We will show how this Lagrangean method works and explain when it fails.
- Suppose $D \subseteq \mathbb{R}^K$, K finite, is open.
- $f : D \to \mathbb{R}$
- $g : D \to \mathbb{R}^J$ and $b \in \mathbb{R}^J$, with $J \leq K$.
- We would like to solve:

$$\max_{x \in D} f(x) \text{ s.t. } g(x) - b \geq 0. \tag{2}$$

- The "usual" method says that one should try to find $(x^*, \lambda^*) \in D \times \mathbb{R}^J_+$ such that $D_x \mathcal{L}(x^*, \lambda^*) = 0$, $g(x^*) - b \geq 0$ and $\lambda^* \cdot g(x^*) = 0$.
- It is as if there were a theorem that states:

  If $x^* \in D$ locally solves Problem (2), then there exists $\lambda^* \in \mathbb{R}^J_+$ such that $D_x \mathcal{L}(x^*, \lambda^*) = 0$, $g(x^*) - b \geq 0$ and $\lambda^* \cdot (g(x^*) - b) = 0$.
- Although this statement recognizes the local character and states only necessary conditions, it neglects the constraint qualification.

Counter-Example

- Consider the problem
$$\max_{(x,y)\in\mathbb{R}^2} x \text{ s.t. } 0 \leq y \leq (1-x)^3. \tag{3}$$
- The Lagrangean of this problem can be written as
$$\mathcal{L}(x,y,\lambda_1,\lambda_2) = x + \lambda_1((1-x)^3 - y) + \lambda_2 y.$$
- Although $(1,0)$ solves (3), there is no (λ_1,λ_2) s.t. $(1,0,\lambda_1,\lambda_2)$ solves:
  1. $1 - 3\lambda_1^*(1-x^*)^2 = 0$
  2. $-\lambda_1^* + \lambda_2^* = 0$;
  3. $\lambda_1^* \geq 0$ and $\lambda_2^* \geq 0$;
  4. $(1-x^*)^3 - y^* \geq 0$ and $y^* \geq 0$; and
  5. $\lambda_1^*((1-x^*)^3 - y^*) = 0$ and $\lambda_2^* y^* = 0$.
- If the FOC were to hold even without the constraint qualification, the system of equations would have to have a solution.

Kühn-Tucker Theorem

Theorem (Kühn - Tucker)

Let $f : D \to \mathbb{R}$ and $g : D \to \mathbb{R}^J$ are both $\mathbb{C}^1$. Suppose that $x^* \in D$ is a local maximiser of f on the constraint set and $g_i(x^*) = b_i$ for $i = 1, ..., I \leq J$. Suppose that $\text{rank}(D\tilde{g}(x^*)) = I$ for $\tilde{g} : D \to \mathbb{R}^I$ defined by $\tilde{g}(x) = (g_j(x))_{j=1}^I$. Then, there exists $\lambda^* \in \mathbb{R}^J$ such that

1. $\frac{\partial \mathcal{L}}{\partial x_k}(x^*, \lambda^*) = 0$, *for all* $k = 1, ..., K$,
2. $\lambda_j^* \cdot (g_j(x^*) - b_j) = 0$ *for all* $j = 1, ..., J$,
3. $\lambda_j^* \geq 0$ *for all* $j = 1, ..., J$, *and*
4. $g_j(x^*) - b_j \geq 0$ *for all* $j = 1, ..., J$.

- With inequality constraints, the sign of λ does matter.
- It is crucial to notice that the process does not amount to maximizing $\mathcal{L}$.
  - In general, $\mathcal{L}$ does not have a maximum;
  - One finds a saddle point of $\mathcal{L}$.

Sufficient Conditions

Theorem

Suppose $f : D \to \mathbb{R} \in$ *and* $g : D \to \mathbb{R}^J$ *are both* $\mathbb{C}^2$. *Suppose there exists* $(x^*, \lambda^*) \in \mathbb{R}^K \times \mathbb{R}^J$ *such that:*

1. $\frac{\partial \mathcal{L}}{\partial x_k}(x^*, \lambda^*) = 0$, *for all* $k = 1, ..., K$,
2. $\lambda_j^* \cdot (g_j(x^*) - b_j) = 0$ *for all* $j = 1, ..., J$,
3. $\lambda_j^* \geq 0$ *for all* $j = 1, ..., J$, *and*
4. $g_j(x^*) - b_j \geq 0$ *for all* $j = 1, ..., J$.
5. $\Delta^\top D^2_{x,x}\mathcal{L}(x^*, \lambda^*)\Delta < 0$ *for all* $\Delta \in \left\{\mathbb{R}^J \setminus \{0\} : \Delta \cdot Dg(x^*) = 0\right\}$.

Example

- Suppose $f(x,y,z) = xyz$,

$$g(x,y,z) = \begin{bmatrix} -(x+y+z) \\ x \\ y \\ z \end{bmatrix}, \qquad b = \begin{bmatrix} -1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

- Then,

$$Dg(x,y,z) = \begin{bmatrix} -1 & -1 & -1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

- A solution exists because the objective function is continuous and the constraint set is nonempty and compact.
- Since at most 3 constraints can be binding at the same time, the CQ holds.
- Let's form the Kühn -Tucker Lagrangean function:

$$\mathcal{L}(x,y,z,\lambda) = xyz + \lambda(1-x-y-z) + \lambda_x x + \lambda_y y + \lambda_z z$$

Example (cont.)

- The FONC are,

$$
\begin{array}{llll}
(1) & \frac{\partial \mathcal{L}(\cdot)}{\partial x} = yz - \lambda + \lambda_x = 0 & (8) & \lambda \geq 0 \\
(2) & \frac{\partial \mathcal{L}(\cdot)}{\partial y} = xz - \lambda + \lambda_y = 0 & (9) & \lambda_x \geq 0 \\
(3) & \frac{\partial \mathcal{L}(\cdot)}{\partial z} = xy - \lambda + \lambda_z = 0 & (10) & \lambda_y \geq 0 \\
(4) & \lambda(1 - x - y - z) = 0 & (11) & \lambda_z \geq 0 \\
(5) & \lambda_x x = 0 & (12) & x + y + z = 1 \\
(6) & \lambda_y y = 0 & (13) & x \geq 0 \\
(7) & \lambda_z z = 0 & (14) & y \geq 0
\end{array}
$$

$$(15) \quad z \geq 0$$

- Since the global maximiser exists and the only points that solve the FONC are (x, y, z) such that $xyz = 0$ and $(x, y, z) = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$, it follows that the latter is the global maximiser.

Quasi-Concave Problems

Theorem

Let $f : D \to \mathbb{R}$ and $g : D \to \mathbb{R}^J$. Suppose f is $\mathbb{C}^1$. Assume there exists $(x^*, \lambda^*) \in \mathbb{R}^K \times \mathbb{R}^J$ such that:

1. $\frac{\partial \mathcal{L}}{\partial x_k}(x^*, \lambda^*) = 0$, *for all* $k = 1, ..., K$,
2. $\lambda_j^* \cdot (g_j(x^*) - b_j) = 0$ *for all* $j = 1, ..., J$,
3. $\lambda_j^* \geq 0$ *for all* $j = 1, ..., J$,
4. $g_j(x^*) - b_j \geq 0$ *for all* $j = 1, ..., J$,
5. f *is quasi-concave with* $\nabla f(x^*) \neq 0$, *and*
6. $g_j(x)$ *is quasi-concave for all* $j = 1, \ldots, J$.

Then x^* is a global maximiser in problem (2)

Proof

Proof: Suppose x^* is not a global maximizer.

1. Then, $f(x) > f(x^*)$ for some $x \in \mathbb{R}^K$ s.t. $g_j(x) \geq b_j$ for every j.
2. Since f is quasi-concave with $\nabla f(x^*) \neq 0$, then $\nabla f(x^*)(x - x^*) > 0$.
3. Since $g_j(\cdot)$ is quasi-concave, $\nabla g_j(x^*)(x - x^*) \geq 0$ for all $j = 1, \dots, J$.
4. Hence, $\sum_{j=1}^{J} \lambda_j \nabla g_j(x^*)(x - x^*) \geq 0$ as $\lambda_j \geq 0$.
5. But by the first K-T condition,

$$\sum_{j=1}^{J} \lambda_j \nabla g_j(x^*)(x - x^*) = -\nabla f(x^*)(x - x^*) < 0$$

a contradiction.