EC9A0: Pre-sessional Advanced Mathematics Course

Constrained Optimisation II: Inequality Constraints

By Pablo F. Beker[1]

1 Introduction

Suppose that $f : \mathbb{R} \to \mathbb{R}$ is differentiable, and suppose that for $a, b \in \mathbb{R}$, $a < b$, we want to find $x^* \in [a, b]$ such that $f(x) \leq f(x^*)$ at all $x \in [a, b]$. That is, we want to solve the problem

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

If $x^* \in (a, b)$ solves the problem, then x^* is a local maximizer of f and $f'(x^*) = 0$. If, alternatively, $x^* = b$ solves the problem, then it must be that $f'(x^*) \geq 0$. Finally, if $x^* = a$ solves the problem, it follows that $f'(x^*) \leq 0$.

It is then straightforward that if x^* solves the problem, then there exist $\lambda_a^*, \lambda_b^* \in \mathbb{R}_+$ such that $f(x^*) - \lambda_b^* + \lambda_a^* = 0$, $\lambda_a^*(x^* - a) = 0$ and $\lambda_b^*(b - x^*) = 0$.[2] It is customary to define a function

$$\mathcal{L} : \mathbb{R}^3 \to \mathbb{R}; \mathcal{L}(x, \lambda_a, \lambda_b) = f(x) + \lambda_b(b - x) + \lambda_a(x - a),$$

which is called *the Lagrangean*, and with which the first condition can be re-written as

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

In this section we show how these Lagrangean methods work, and emphasize when they fail.

2 Inequality constraints

As before, let $f : D \to \mathbb{R} \in \mathbf{C}^1$, $g : D \to \mathbb{R}^J \in \mathbf{C}^1$ and $b \in \mathbb{R}^J$. Now suppose that we have to solve the problem

$$\max_{x \in D} f(x) : g(x) - b \geq 0. \tag{1}$$

Again, 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 (1), then there exists $\lambda^* \in \mathbb{R}_+^J$ such that $D_x\mathcal{L}(x^*, \lambda^*) = 0$, $g(x^*) - b \geq 0$ and $\lambda^* \cdot g(x^*) = 0$.*

Now, even though in this statement we are recognizing the local character and (only) the necessity of the result, we still have to worry about constraint qualification. To see that this is the case, consider the following example:

[1] Based on notes by Andrés Carvajal

[2] The second and third condition simply express that (i) if $x^* \in (a, b)$, then $\lambda_a^* = 0$ and $\lambda_b^* = 0$; (ii) if $x^* = b$, then $\lambda_a^* = 0$ and $\lambda_b^* \geq 0$; and (iii) if $x^* = a$, then $\lambda_a^* \geq 0$ and $\lambda_b^* = 0$.

EXAMPLE 1. *Consider the problem*

$$\max_{(x,y)\in\mathbb{R}^2} -((x-3)^2+y^2) : 0 \le y \le -(x-1)^3.$$

The Lagrangean of this problem can be written as

$$\mathcal{L}(x,y,\lambda_1,\lambda_2) = -(x-3)^2 - y^2 + \lambda_1(-(x-1)^3 - y) + \lambda_2 y.$$

Notice that, although $(1,0)$ *solves the problem, there is no* (λ_1,λ_2) *such that* $(1,0,\lambda_1,\lambda_2)$ *solves the following system:*
(i) $-2(x^*-3) - 3\lambda_1^*(x^*-1)^2 = 0$ *and* $-2y^* - \lambda_1^* + \lambda_2^* = 0$*;*
(ii) $\lambda_1^* \ge 0$ *and* $\lambda_2^* \ge 0$*;*
(iii) $-(x^*-1)^3 - y^* \ge 0$ *and* $y^* \ge 0$*; and*
(iv) $\lambda_1^*(-(x^*-1)^3 - y^*) = 0$ *and* $\lambda_2^* y^* = 0$*.*

If the first order conditions were to hold even without the constraint qualification, the system of equations in the previous example would necessarily have to have a solution. The point of the example is just that the theorem requires the constraint qualification condition:

THEOREM 1 (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* $i = 1,...,I \le J$*. 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. $\lambda_j^* \ge 0$ *for all* $j = 1,...,J,$ *and*
4. $g_j(x^*) - b_j \ge 0$ *for all* $j = 1,...,J.$

Proof: See proof of Theorem 18.4 (page 480) in Simon and Blume's book.

Importantly, notice that with inequality constraint the sign of λ does matter: this is because of the geometry of the theorem: a local maximizer is attained when the feasible directions, as determined by the gradients of the binding constraints is exactly opposite to the desired direction, as determined by the gradient of the objective function. Obviously, locally only the binding constraints matter, which explains why the constraint qualification looks more complicated here than with equality constraints. Finally, it is crucial to notice that the process does not amount to maximizing $\mathcal{L}$: in general, $\mathcal{L}$ does not have a maximum; what one finds is a saddle point of $\mathcal{L}$.

The following theorem states the second order (sufficient) conditions:

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

Proof: See Section 30.5 (page 841) in Simon and Blume's book.

EXAMPLE 2. *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 Constraint Qualification holds. Let's form the Kühn -Tucker Lagrangean function:

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

The FONC are,

$$\begin{array}{llllll}
(1) & \frac{\partial \mathcal{L}(\cdot)}{\partial x} = yz - \lambda + \lambda_x = 0 & (8) & \lambda \geq 0 & (15) & z \geq 0 \\
(2) & \frac{\partial \mathcal{L}(\cdot)}{\partial y} = xz - \lambda + \lambda_y = 0 & (9) & \lambda_x \geq 0 & & \\
(3) & \frac{\partial \mathcal{L}(\cdot)}{\partial z} = xy - \lambda + \lambda_z = 0 & (10) & \lambda_y \geq 0 & & \\
(4) & \lambda(1-x-y-z) = 0 & (11) & \lambda_z \geq 0 & & \\
(5) & \lambda_x x = 0 & (12) & x+y+z = 0 & & \\
(6) & \lambda_y y = 0 & (13) & x \geq 0 & & \\
(7) & \lambda_z z = 0 & (14) & y \geq 0 & &
\end{array}$$

Case 1: $\lambda = 0$

$(1) \Rightarrow yz + \lambda_x = 0$. *Then, (9), (14) and (15)* $\Rightarrow yz = 0$ *and* $\lambda_x = 0$
$(2) \Rightarrow xz + \lambda_y = 0$. *Then, (10), (13) and (15)* $\Rightarrow xz = 0$ *and* $\lambda_y = 0$
$(3) \Rightarrow xy + \lambda_z = 0$. *Then, (11), (13) and (14)* $\Rightarrow xy = 0$ *and* $\lambda_z = 0$
Hence, $f(x,y,z) = f(0,0,0) = 0$.

Case 2: $\lambda > 0$

$(4) \Rightarrow x+y+z = 1$. *Suppose that* $x = 0$. *Then,* $(2)+(3) \Rightarrow \lambda_y = \lambda_z = \lambda > 0$. *Thus,* $(6)+(7) \Rightarrow y = 0, z = 0$. *Hence,* $x+y+z = 0 < 1$, *a contradiction. It follows that* $x > 0$. *An analogous reasoning shows that* $y > 0$ *and* $z > 0$ *and, therefore, (5), (6) and (7)* $\Rightarrow \lambda_x = \lambda_y = \lambda_z = 0$. *Hence, (1), (2) and (3)* $\Rightarrow yz = xz = xy$ *and* $x+y+z = 1$. *It follows that* $x = y = z = \frac{1}{3}$ *and* $\lambda = \frac{1}{9}$ *We conclude that* $f(x,y,z) = f(\frac{1}{3},\frac{1}{3},\frac{1}{3}) = \frac{1}{27} > 0$.

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

2.1 Quasi-concave Problems

As before, it must be noticed that there is a gap between necessity and sufficiency, and that the theorem only gives local solutions. For the former problem, there is no solution. For the latter, one can study concavity of the objective function and convexity of the feasible set.

Theorem 3. *Let $f : D \to \mathbb{R} \in$ 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. $\lambda_j g_j(x)$ *is quasi-concave.*
7. x^* *satisfies the constraint qualification.*

Then x^* is a global maximiser in problem (1)

Proof: See proof of Theorem M.K.3 (page 961) and discussion in page 962 of Mas-Colell et al's book.

Remark: Under the conditions of the Theorem, the constraint qualification holds if the constraint set has non-empty interior, see footnote 25 in page 962 of Mas-Colell et al.