T. Mitra, Fall 2014

Intermediate Mathematical Economics (Econ 6170)

Fall 2014

Notes on the Lagrange Multiplier Theorem

# 1 Introduction

In these notes, we state and prove a general version of the Lagrange Multiplier Theorem, with multiple equality constraints.

It is useful to keep in mind that the theorem provides Lagrange's *necessary* conditions that must hold at a point of constrained local maximum (provided a constraint qualification holds). If one finds a pair $(\hat{x}, \hat{\lambda})$ where $\hat{x}$ is in the constraint set, and $\hat{\lambda}$ the vector of Lagrange multipliers associated with the equality constraints, such that the pair satisfies the Lagrange conditions, this does not ensure in general that $\hat{x}$ is a point of constrained global or local maximum.

However, Lagrange's theorem, when combined with Weierstrass theorem on the existence of a constrained maximum, can be a powerful method for solving a class of constrained optimization problems. Because neither theorem uses convex structures, this method can be very useful in solving optimization problems in economics in the presence of non-convexities. The method has also been used in the mathematics literature to provide proofs of various analytical inequalities.

# 2 Statement of Result

Here is a statement of the Lagrange theorem.

**Theorem 1** *Let $X$ be an open subset of $\mathbb{R}^n$. Let $f, g^1, ..., g^m$ be continuously differentiable real valued functions on $X$. Suppose $\bar{x} \in X$ is a point of local maximum of $f$, subject to the constraints $g^j(x) = 0$ for $j \in \{1, ..., m\}$. That is, there is $\delta' > 0$ such that for all $x \in B(\bar{x}, \delta') \cap C$, where $C \equiv \{x \in A : g^j(x) = 0$ for all $j \in \{1, ..., m\}\}$, we have $f(x) \leq f(\bar{x})$.*

*Assume that the following constraint qualification holds: the $m \times n$ matrix $(D_i g^j(\bar{x}))$ has rank equal to $m$. Then, there exists $\lambda \in \mathbb{R}^m$ such that:*

$$D_i f(\bar{x}) = \lambda_1 D_i g^1(\bar{x}) + \cdots + \lambda_m D_i g^m(\bar{x}) \text{ for all } i \in \{1, ..., n\} \tag{1}$$

# 3 Proof of the Result

We break up the proof into several steps.

**Step 1:**

Choose $0 < \delta'' < \delta'$, such that $B(\bar{x}, \delta'')$ is contained in $X$; this can be done since $X$ is open in $\mathbb{R}^n$. Then for $0 < \delta < \delta''$, the closed ball $\bar{B}(\bar{x}, \delta) \equiv \{x \in \mathbb{R}^n : ||x - \bar{x}|| \leq \delta\}$ is contained in $X$. Further, since $\bar{B}(\bar{x}, \delta) \subset B(\bar{x}, \delta')$, we know that for all $x \in \bar{B}(\bar{x}, \delta) \cap C$, we have $f(x) \leq f(\bar{x})$, where:

$$C \equiv \{x \in A : g^j(x) = 0 \text{ for all } j \in \{1, ..., m\}\} \tag{2}$$

**Step 2:**

We claim that for each $s \in \mathbb{N}$, there is $N \in \mathbb{N}$ (with $N$ depending on $s$) such that for all $x \in X$, satisfying $||x - \bar{x}|| = (\delta/s)$, we have:

$$f(x) - f(\bar{x}) - ||x - \bar{x}||^2 - N \sum_{j=1}^{m} [g^j(x)]^2 < 0 \tag{3}$$

If the claim is not true, then there is some $s \in \mathbb{N}$, such that for every choice of $N \in \mathbb{N}$, condition (3) is violated for some $x(N) \in X$, satisfying $||x(N) - \bar{x}|| = (\delta/s)$.That is, there is some $s \in \mathbb{N}$, such that for every choice of $N \in \mathbb{N}$, there is a corresponding $x(N) \in X$, with $||x(N) - \bar{x}|| = (\delta/s)$, such that:

$$f(x(N)) - f(\bar{x}) - ||x(N) - \bar{x}||^2 \geq N \sum_{j=1}^{m} [g^j(x(N))]^2 \tag{4}$$

This can be rewritten as:

$$\frac{f(x(N)) - f(\bar{x}) - ||x(N) - \bar{x}||^2}{N} \geq \sum_{j=1}^{m} [g^j(x(N))]^2 \tag{5}$$

The sequence $\{x(N)\}_{N=1}^{\infty}$ is in $\bar{B}(\bar{x}, \delta)$, a closed and bounded set, and therefore has a convergent subsequence, $\{x(N_k)\}_{k=1}^{\infty}$converging to $\hat{x}$. Then $\hat{x} \in \mathbb{R}^n$, and $||\hat{x} - \bar{x}|| = (\delta/s)$, by continuity of the norm on $\mathbb{R}^n$. Thus, $\hat{x} \in \bar{B}(\bar{x}, \delta) \subset X$. Since $f, g^1, ..., g^m$ are continuous functions on $X$, we can also infer that $f(x(N_k)) \to f(\hat{x})$ and $[g^j(x(N_k))] \to [g^j(\hat{x})]$ as $N_k \to \infty$.Now, we can use (5) and let $N_k \to \infty$ to infer that:

$$0 = \lim_{N_k \to \infty} \frac{f(x(N_k)) - f(\bar{x}) - ||x(N_k) - \bar{x}||^2}{N_k} \geq \lim_{N_k \to \infty} \sum_{j=1}^{m} [g^j(x(N_k))]^2 = \sum_{j=1}^{m} [g^j(\hat{x})]^2 \tag{6}$$

since weak inequalities are preserved in the limit. This implies that:

$$\sum_{j=1}^{m} [g^j(\hat{x})]^2 = 0 \tag{7}$$

and so $\hat{x} \in C$. Also $\hat{x} \in \bar{B}(\bar{x}, \delta) \subset B(\bar{x}, \delta')$, and so $\hat{x} \in C \cap B(\bar{x}, \delta')$. Consequently, we must have:

$$f(\hat{x}) \leq f(\bar{x}) \tag{8}$$

Using (4), we also have:

$$f(x(N_k)) - f(\bar{x}) - ||x(N_k) - \bar{x}||^2 \geq 0 \tag{9}$$

and so letting $N_k \to \infty$ in (9), and using the continuity of $f$ on $X$, we get:

$$f(\hat{x}) - f(\bar{x}) - ||\hat{x} - \bar{x}||^2 \geq 0 \tag{10}$$

Using (8) and (10), we must have $||\hat{x} - \bar{x}||^2 = 0$. But, this is a contradiction to our earlier finding that $||\hat{x} - \bar{x}|| = (\delta/s) > 0$. This contradiction establishes our claim.

We have established that for each $s \in \mathbb{N}$, there is $N \in \mathbb{N}$ (with $N$ depending on $s$) such that for all $x \in X$, satisfying $||x - \bar{x}|| = (\delta/s)$, (3) holds. Examining (3), we then see that if we replace $N$ in the left hand side of (3) by $N' > N$, the inequality in (3) will still hold. Thus, we can always choose $N \geq s$ such that for all $x \in X$, satisfying $||x - \bar{x}|| = (\delta/s)$, (3) holds. To be even more specific we can choose the *smallest* $N \geq s$ such that for all $x \in X$, satisfying $||x - \bar{x}|| = (\delta/s)$, (3) holds. We will assume this about the choice of $N$ in what follows; note that this makes $N$ a well-defined function of $s$, which we can denote by $h : \mathbb{N} \to \mathbb{N}$.

**Step 3:**

In this step, we fix an arbitrary $s \in \mathbb{N}$, and pick a corresponding $N \in \mathbb{N}$ such that $N = h(s)$. By Step 2, with this choice of $N$, for all $x \in X$ satisfying $||x - \bar{x}|| = (\delta/s)$, we have:

$$f(x) - f(\bar{x}) - ||x - \bar{x}||^2 - N \sum_{j=1}^{m} [g^j(x)]^2 < 0 \tag{11}$$

With $s$ fixed (and therefore $N = h(s)$ also fixed), define:

$$F(x) = f(x) - f(\bar{x}) - ||x - \bar{x}||^2 - N \sum_{j=1}^{m} [g^j(x)]^2 \text{ for all } x \in X \tag{12}$$

Clearly, $F$ is a well-defined function from $X$ to $\mathbb{R}$. Since $\bar{B}(\bar{x}, \delta/s)$ is a non-empty, closed and bounded subset of $X$, and $F$ a continuous function on $\bar{B}(\bar{x}, \delta/s)$, there is $x' \in \bar{B}(\bar{x}, \delta/s)$ such that $F(x) \leq F(x')$ for all $x \in \bar{B}(\bar{x}, \delta/s)$, by an application of Weierstrass Theorem. Clearly, $\bar{x} \in \bar{B}(\bar{x}, \delta/s)$ and $F(\bar{x}) = 0$ by (12), so $F(x') \geq 0$. This implies by using (11) that $||x' - \bar{x}|| \neq (\delta/s)$. Thus, $x'$ must belong to the *open ball* $B(\bar{x}, \delta/s)$, which is an open set in $\mathbb{R}^n$. To summarize, the function $F$ attains a maximum at $x'$ among all points $x \in B(\bar{x}, \delta/n)$. We can therfore use the first-order conditions of unconstrained optimization to infer that $D_i F(x') = 0$ for all $i \in \{1, ..., n\}$. These conditions can be explicitly written as:

$$D_i f(x') - 2(x'_i - \bar{x}_i) = 2N[g^1(x') D_i g^1(x') + \cdots + g^m(x') D_i g^m(x')] \text{ for all } i \in \{1, ..., n\} \tag{13}$$

Define:

$$K = [1 + \sum_{j=1}^{m} \{2N g^j(x')\}^2]^{1/2}, \ \alpha = (1/K), \ \beta_j = \{2N g^j(x')\}/K \text{ for } j \in \{1, ..., m\} \tag{14}$$

Dividing through by $K$ in (13), we get:

$$\alpha D_i f(x') - (2/K)(x'_i - \bar{x}_i) = [\beta_1 D_i g^1(x') + \cdots + \beta_m D_i g^m(x')] \text{ for all } i \in \{1, ..., n\} \tag{15}$$

Note that by (14), we have $(\alpha, \beta) \in \mathbb{R}^{m+1}$, and:

$$||(\alpha, \beta)|| = 1 \tag{16}$$

**Step 4:**

We will now let $s \in \mathbb{N}$ vary, and specifically let $s \to \infty$. For this purpose, we need to introduce some notation. Given any $s \in \mathbb{N}$, we have $N = h(s)$. We will therefore have the function $F$, defined in (12), dependent on $s$; and we can indicate this by writing:

$$F(x, s) = f(x) - f(\bar{x}) - ||x - \bar{x}||^2 - h(s) \sum_{j=1}^{m} [g^j(x)]^2 \text{ for all } x \in X \tag{17}$$

in place of (12). This function will attain a maximum on the set $\bar{B}(\bar{x}, \delta/s)$, and we will indicate a point in $\bar{B}(\bar{x}, \delta/s)$ at which the maximum is attained by $x'(s)$.

Then, as in Step 3, we will obtain:

$$D_i f(x'(s)) - 2(x'_i(s) - \bar{x}_i) = 2h(s)[g^1(x'(s)) D_i g^1(x'(s)) + \cdots + g^m(x'(s)) D_i g^m(x'(s))] \text{ for all } i \in \{1, ..., n\} \tag{18}$$

in place of (13). We define:

$$K(s) = [1+\sum_{j=1}^{m}\{2h(s)g^j(x'(s))\}^2]^{1/2},\ \alpha(s) = (1/K(s)),\ \beta_j(s) = \{2h(s)g^j(x'(s))\}/K(s)\ \text{ for } j \in \{1,...,m\} \tag{19}$$

analogous to (14). Dividing through by $K(s)$ in (13), we get:

$$\alpha(s)D_if(x'(s)) - (2/K(s))(x'_i(s) - \bar{x}_i) = [\beta_1(s)D_ig^1(x'(s)) + \cdots + \beta_m(s)D_ig^m(x'(s))]\ \text{ for all } i \in \{1,...,n\} \tag{20}$$

analogous to (15). Finally, note that by (19), we have $(\alpha(s), \beta(s)) \in \mathbb{R}^{m+1}$, and:

$$||(\alpha(s), \beta(s))|| = 1 \tag{21}$$

analogous to (16).

We now let $s \to \infty$. Since $h(s) \geq s$, we have $h(s) \to \infty$. Since $x'(s) \in B(\bar{x}, \delta/s)$, we have $x'(s) \to \bar{x}$ as $s \to \infty$. Because of (21), $(\alpha(s), \beta(s))$ is in a bounded set, and has a convergent subsequence $(\alpha(s_k), \beta(s_k))$ converging to $(\alpha, \beta) \in \mathbb{R}^{m+1}$. Using (21), we must have:

$$||(\alpha, \beta)|| = 1 \tag{22}$$

Now, letting $s_k \to \infty$, and noting that $(\alpha(s_k), \beta(s_k))$ converges to $(\alpha, \beta)$, while $x'(s_k) \to \bar{x}$, continuity of $D_if$ and $D_ig^1, ..., D_ig^m$ on $X$ imply, using (20), that:

$$\alpha D_if(\bar{x}) = [\beta_1 D_ig^1(\bar{x}) + \cdots + \beta_m D_ig^m(\bar{x})]\ \text{ for all } i \in \{1,...,n\} \tag{23}$$

noting that $0 \leq (2/K(s)) \leq 2$ for all $s \in \mathbb{N}$.

**Step 5:**

It will be noted that (23) can be derived without the use of the Constraint Qualification. We use the Constraint Qualification at this stage to ensure that $\alpha \neq 0$. For if $\alpha = 0$, then by (22), we must have $||\beta|| = 1$, and so $\beta \neq 0$. Then (23) implies that the $m \times n$ matrix $(D_ig^j(\bar{x}))$ has rank less than $m$. This violates the Constraint Qualification. Thus, $\alpha \neq 0$, and defining $\lambda \in \mathbb{R}^m$ by $\lambda = \beta/\alpha$, and dividing through in (23) by $\alpha \neq 0$, we obtain (1). This concludes the proof.

**Remark:**

Since the Constraint Qualification ensures that $\alpha \neq 0$, it must ensure that $K(s_k)$ remains bounded above as $s_k \to \infty$. This, in turn, means that the CQ ensures that as $s_k$ goes to infinity, $\sum_{j=1}^{m}[g^j(x'(s_k))]^2$ goes to 0 at least as fast. That is, as $s_k \to \infty$, $x'(s_k)$ must converge to $\bar{x}$ fast enough to make this happen. Although this information is inessential for the proof of Steps 4 and 5, it is an implication of the approach used in those steps.

**Reference:**

The above approach to the Theorem is based entirely on the paper by McShane (1973). Our setting is deliberately simpler than the one considered in McShane's paper (which includes inequality contraints as well as equality constraints). We have supplied a bit more of the details of the steps of the proof.

An alternative approach to the theorem that is often used makes explicit use of the Implicit Function Theorem.

E.J. McShane (1973), The Lagrange Multiplier Rule, *American Mathematical Monthly*, volume 80, No. 8, 922-925.