A 4-steps elementary proof of existence of Lagrange multipliers

Gabriel Haeser * Daiana O. Santos †

Abstract We present a simplified proof of Lagrange's theorem using only elementary properties of sets and sequences.

Consider the problem of minimizing a real valued function $f(x)$, with $x \in \mathbb{R}^n$, subject to m equality constraints $h_1(x) = 0, \dots, h_m(x) = 0$ where all functions are continuously differentiable. Let $\overline{x}$ be a local solution with gradients $\nabla h_1(\overline{x}), \dots, \nabla h_m(\overline{x})$ linearly independent. We will show that there exist so-called Lagrange multipliers $\lambda_1, \dots, \lambda_m \in \mathbb{R}$ with $\nabla f(\overline{x}) + \sum_{i=1}^m \lambda_i \nabla h_i(\overline{x}) = 0$.

Step 1: Build the sequence

Let $\delta > 0$ be such that $f(\overline{x}) \leq f(x)$ for all x with $h(x) = 0$ and $\|x - \overline{x}\| \leq \delta$, where $h := (h_1, \dots, h_m)$ and $\|\cdot\|$ is the Euclidean norm. Consider the bounded sequence $\{x^k\}_{k\in\mathbb{N}}$ where each x^k, $k \in \mathbb{N}$, minimizes

$$\phi_k(x) := f(x) + \|x - \overline{x}\|^2 + k\|h(x)\|^2$$

over the compact set $\{x \in \mathbb{R}^n : \|x - \overline{x}\| \leq \delta\}$.

Step 2: Prove convergence

Take an arbitrary infinite set $K \subseteq \mathbb{N}$ such that $\{x^k\}_{k\in K}$ converges and let us show that its limit (say, x^*) is $\overline{x}$. Since

$$f(x^k) + \|x^k - \overline{x}\|^2 \leq \phi_k(x^k) \leq \phi_k(\overline{x}) = f(\overline{x}), \quad (1)$$

one must have by the definition of ϕ_k that $h(x^*) = 0$, since otherwise $\phi_k(x^k)$ would be unbounded. But $\|x^* - \overline{x}\| \leq \delta$ and, from the definition of δ, it follows that $f(\overline{x}) \leq f(x^*)$. Taking the limit for $k \in K$ in (1) we arrive at $f(x^*) + \|x^* - \overline{x}\|^2 \leq f(\overline{x}) \leq f(x^*)$, which implies that $x^* = \overline{x}$. This, together with the boundedness of $\{x^k\}_{k\in\mathbb{N}}$, proves that $\{x^k\}_{k\in\mathbb{N}}$ converges to $\overline{x}$.

*University of São Paulo. E-mail: ghaeser@ime.usp.br

†Federal University of São Paulo.
E-mail: daiana.santos@unifesp.br

Step 3: Take derivatives

For $k \in \mathbb{N}$ large enough one must have $\|x^k - \overline{x}\| < \delta$. Thus, x^k must be an unconstrained local minimizer of ϕ_k, implying $\nabla \phi_k(x^k) = 0$, that is,

$$\nabla f(x^k) + 2(x^k - \overline{x}) + \sum_{i=1}^m \lambda_i^k \nabla h_i(x^k) = 0, \quad (2)$$

with $\lambda_i^k := 2kh_i(x^k), i = 1, \dots, m$.

Step 4: Bound and take limit

It must be that $|\lambda_i^k| \leq M, i = 1, \dots, m$, for some $M \in \mathbb{R}$, since otherwise we could divide (2) by $\max\{|\lambda_i^k|, i = 1, \dots, m\}$ and take the limit in a suitable subsequence to contradict the linear independence assumption. The proof is completed using the continuity of the gradients in (2) by taking a subsequence such that each λ_i^k converges.

This proof is an adaptation of the ideas from [Bel69, Ber99, AHM11] and we note that the linear independence assumption may be replaced by the constant rank of the gradients nearby $\overline{x}$ with a simple additional step rewriting the sum $\sum_{i=1}^m \lambda_i^k \nabla h_i(x^k)$ using a constant index set corresponding to a basis of the space generated by $\nabla h_i(x^k), i = 1, \dots, m$. Inequality constraints $g_1(x) \leq 0, \dots, g_p(x) \leq 0$ may be treated similarly by adding the term $k\sum_{i=1}^p \max\{0, g_i(x)\}^2$ in the definition of $\phi_k(x)$.

References

[AHM11] R. Andreani, G. Haeser, and J.M. Martínez, *On sequential optimality conditions for smooth constrained optimization*, Optimization **60** (2011), no. 5, 627–641.

[Bel69] E.J. Beltrami, *A constructive proof of the Kuhn-Tucker multiplier rule*, J. Math. Anal. Appl. **26** (1969), 297–306.

[Ber99] D. P. Bertsekas, *Nonlinear programming*, Athena Scientific, 1999.