

The Fredholm alternative and the Lagrange multiplier theorem

The Lagrange multiplier theorem (Theorem 5.2) is the central tool of optimization as practised by economists. The most straightforward derivation of the Lagrange multiplier theorem is given by applying the Fredholm alternative, as is done in the text (Section 5.3.1).[1]

Three different proofs of the Fredholm alternative are given in the text, deriving it in turn from the quotient theorem (Exercise 3.48), the Hahn-Banach theorem (Exercise 3.199) and the Farkas lemma (Exercise 3.237). Each of these involves some sophisticated reasoning. The exposition of optimization in chapter 5 assumes finite-dimensional spaces. In this case, a much simpler proof of the Fredholm alternative can be given, using only basic results in linear algebra. This is the purpose of this supplementary note.

Proposition 1 (Fredholm alternative) *Let $f_1, f_2, \ldots, f_m$ be linear functionals on $\Re^n$ and let*

$$S = \{\, \mathbf{x} \in X : f_j(\mathbf{x}) = 0,\ j = 1, 2, \ldots, m \,\}$$

Suppose that $f_0 \neq 0$ is another linear functional such that such that $f_0(\mathbf{x}) = \mathbf{0}$ for every $\mathbf{x} \in S$. Then, there exist numbers $\lambda_1, \lambda_2, \ldots, \lambda_m \in \Re$ such that

$$f_0(\mathbf{x}) = \sum_{i=1}^{m} \lambda_i f_i(\mathbf{x})$$

In other words, f_0 is a linear combination of $f_1, f_2, \ldots, f_m$.

Proof. Each f_j can be represented by a vector $\mathbf{a}_j \in \Re^n$ (Proposition 3.4), that is for each f_j there exists $\mathbf{a}_j \in \Re^n$ such that

$$f_j(\mathbf{x}) = \mathbf{a}_j^T \mathbf{x} \text{ for every } \mathbf{x} \in X$$

and S can be alternatively expressed as

$$S = \{\, \mathbf{x} \in X : \mathbf{a}_j^T \mathbf{x} = 0,\ j = 1, 2, \ldots, m \,\}$$

The orthogonal complement of S is

$$S^\perp = \{\, \mathbf{a} \in \Re^n : \mathbf{a}^T \mathbf{x} = 0 \text{ for every } \mathbf{x} \in S \,\}$$

By definition, each $\mathbf{a}_j \in S^\perp, j = 1, 2, \ldots, m$. $S^\perp$ is a subspace (Supplementary exercise 3.1).

Without loss of generality, assume that $\{\, \mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_l \,\}$ form a maximal linearly independent subset of $\{\, \mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m \,\}$. Then, S has dimension $n - l$ (Exercise 3.24) which implies that $S^\perp$ has dimension l (Supplementary exercise 3.4). That is, the l linearly independent elements $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_l$ span $S^\perp$.

Since $f_0 = \mathbf{a}_0^T \mathbf{x}$ for some $\mathbf{a}_0 \in \Re^n$ (Proposition 3.4), $f_0(\mathbf{x}) = \mathbf{a}_0^T \mathbf{x} = \mathbf{0}$ for every $\mathbf{x} \in S$. Therefore $\mathbf{a}_0 \in S^\perp = \text{lin}\ \{\, \mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_l \,\}$ and there exist numbers $\lambda_1, \lambda_2, \ldots, \lambda_l$ such that

$$\mathbf{a}_0 = \sum_{j=1}^{l} \lambda_j \mathbf{a}_j$$

[1] An alternative proof using the implicit function theorem is given in Section 5.3.3. This is the approach adopted by most textbooks, but it less intuitive and notationally burdensome.

which implies that

$$f_0(\mathbf{x}) = \mathbf{a}_0^T \mathbf{x} = \sum_{j=1}^{l} \lambda_j \mathbf{a}_j^T \mathbf{x} = \sum_{j=1}^{l} \lambda_i f_j(\mathbf{x})$$

Setting $\lambda_j = 0,\ j = l+1, \dots, m$, we conclude that

$$f_0(\mathbf{x}) = \sum_{j=1}^{m} \lambda_j f_j(\mathbf{x})$$

as required. □

The Lagrange mutliplier theorem is a straightforward application (p. 518-519). Provided that the gradients $\nabla g_j(\mathbf{x}^*)$ are linearly independent, a necessary condition for $\mathbf{x}^*$ to be a local optimum of

$$\max_{\mathbf{x} \in X} f(\mathbf{x})$$
$$\text{subject to } g_j(\mathbf{x}) = 0, j = 1, 2, \dots, m$$

is that $Df[\mathbf{x}^*](\mathbf{dx}) = 0$ for all perturbations $\mathbf{dx}$ in

$$S = \{\, \mathbf{dx} \in X : Dg_j(\mathbf{x}) = 0,\ j = 1, 2 \dots m \,\}$$

The derivatives Df and Dg_j are linear functionals. Setting $f_0 = Df[\mathbf{x}^*]$ and $f_j = Dg_j[\mathbf{x}^*],\ j = 1, 2, \dots, m$ and applying the Fredholm alternative, we conclude that there exist multipliers $\lambda_1, \lambda_2, \dots, \lambda_m$ such that

$$Df[\mathbf{x}^*] = \sum_{j=1}^{m} \lambda_j Dg_j[\mathbf{x}^*]$$

or alternatively

$$\nabla f(\mathbf{x}^*) = \sum_{j=1}^{m} \lambda_j \nabla g_j(\mathbf{x}^*)$$