ラグランジュ乗数法—陰関数定理の応用例

1　等式制約付き最適化問題

変数 $x \in \mathbf{R}^n$ に関する等式制約の最適化問題

$$\mathrm{Minimize} f(x) \quad \mathrm{s.t.} \quad g_i(x) = 0 \qquad (i = 1, \ldots, m)$$

に対するラグランジュ乗数法を説明する ($m \leq n$)．記号の定義：

$$\nabla f(x) = \left(\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n} \right), \qquad \nabla g_i(x) = \left(\frac{\partial g_i}{\partial x_1}, \ldots, \frac{\partial g_i}{\partial x_n} \right).$$

定理（最適性の必要条件）　点 x^* が

$$g_i(x^*) = 0 \qquad (i = 1, \ldots, m) \tag{1}$$

を満たすとする．点 x^* が最適解ならば，$\boxed{\text{適当な条件}}$（後述）の下で

$$\nabla f(x^*) = \sum_{i=1}^{m} \lambda_i^* \nabla g_i(x^*) \tag{2}$$

を満たす $\lambda^* \in \mathbf{R}^m$ が存在する．■

条件式 (1), (2) を合わせると，$m+n$ 個の変数 $\lambda_1^*, \ldots, \lambda_m^*$, $x_1^*, \ldots, x_n^*$ に関する $m+n$ 個の方程式となり，これを解けば最適解（の候補）が求められる．これを**ラグランジュ乗数法**あるいは**ラグランジュ未定乗数法**という．なお，**ラグランジュ関数** $L(x, \lambda)$ を

$$L(x, \lambda) = f(x) - \sum_{i=1}^{m} \lambda_i g_i(x) \tag{3}$$

($x \in \mathbf{R}^n$, $\lambda \in \mathbf{R}^m$) と定義すると，上の二つの条件は次の形に表現できる:

$$\text{条件 (2)} \iff \frac{\partial L}{\partial x_j}(x^*, \lambda^*) = 0 \quad (j = 1, \ldots, n),$$
$$\text{条件 (1)} \iff \frac{\partial L}{\partial \lambda_i}(x^*, \lambda^*) = 0 \quad (i = 1, \ldots, m).$$

2　証明

第 1 段：線形近似による大筋の理解

方向を表すベクトル $d \in \mathbf{R}^n$ と実数 $t \in \mathbf{R}$ を使ってベクトル $x^* + td$ を考える．内積を $\cdot$ で表すとき，テイラー展開より

$$f(x^* + td) = f(x^*) + t\ (\nabla f(x^*) \cdot d) + O(t^2), \tag{4}$$
$$g_i(x^* + td) = g_i(x^*) + t\ (\nabla g_i(x^*) \cdot d) + O(t^2) \tag{5}$$

となるので，t の 1 次の項まで考えるとき，x^* が最適解ならば

$$\nabla g_i(x^*) \cdot d = 0 \quad (i = 1, \ldots, m) \text{ を満たす任意の } d \text{ に対して } \nabla f(x^*) \cdot d = 0 \tag{6}$$

が成り立つ．$\nabla g_i(x^*)$ を第 i 行ベクトルとする $m \times n$ 行列を G とすると，式 (6) の条件は「$Gd = \mathbf{0} \Rightarrow \nabla f(x^*) \cdot d = 0$」と同値であり，さらに，これは

$$(\nabla f(x^*))^\top \in (\mathrm{Ker}\, G)^\perp \tag{7}$$

と言い換えることができる（$^\perp$ は直交補空間を表す）．ここで

$$\textbf{線形代数における基本公式}: \qquad (\mathrm{Ker}\, G)^\perp = \mathrm{Im}\, G^\top$$

により，条件 (7) は

$$\nabla f(x^*) \text{ が } \nabla g_1(x^*), \ldots, \nabla g_m(x^*) \text{ の 1 次結合でかける} \tag{8}$$

と同値である．この 1 次結合の係数を $\lambda_1^*, \ldots, \lambda_m^*$ とすれば (2) が成り立つ．

第 2 段：陰関数定理を利用した厳密な証明

定理中の $\boxed{\text{適当な条件}}$ とは：

(i) 点 x^* の近傍において関数 $f(x)$, $g_i(x)$ $(i = 1, \ldots, m)$ は連続微分可能 (C^1),
(ii) $\{\nabla g_1(x^*), \ldots, \nabla g_m(x^*)\}$ は 1 次独立.

$\nabla g_i(x^*)$ を第 i 行ベクトルとする $m \times n$ 行列を G とする．G の適当な m 本の列ベクトルを選べば線形独立である．一般性を失うことなく，最後の m 列（$n - m + 1, \ldots, n$）が線形独立と仮定し，$x = (y, z)$ （ただし $y \in \mathbf{R}^{n-m}$, $z \in \mathbf{R}^m$）と分けて，$x^* = (y^*, z^*)$, $f(x) = f(y, z)$, $g_i(x) = g_i(y, z)$ と考える．

行列 G も $G = [N \ B]$ と分ける．ここで $N = \left(\frac{\partial g_i}{\partial y_k} \mid i = 1, \ldots, m; k = 1, \ldots, n - m\right)$ は $m \times (m - n)$ 行列，$B = \left(\frac{\partial g_i}{\partial z_j} \mid i, j = 1, \ldots, m\right)$ は $m \times m$ 行列であり，B は正則である．したがって，**陰関数定理**により，(y^*, z^*) の近傍において $z = \varphi(y)$ と解ける．ここで φ は連続微分可能で，

$$\left(\frac{\partial \varphi_i}{\partial y_k}\right) = -B^{-1} N$$

である．

元の制約付き最適化問題は，変数 z を消去した関数 $F(y) = f(y, \varphi(y))$ の制約なしの最小化問題と（局所的には）等価であり，後者の最適性の必要条件は $\nabla_y F(y^*) = \mathbf{0}$ である．一方，(y^*, z^*) において

$$\nabla_y F = \nabla_y f - \nabla_z f \cdot B^{-1} N$$

であるから，$\lambda^* = \nabla_z f \cdot B^{-1}$ （m 次元の横ベクトル）とおくと，

$$\nabla_y f = \lambda^* N, \qquad \nabla_z f = \lambda^* B$$

となる．$\nabla f = (\nabla_y f, \nabla_z f)$, $G = [N \ B]$ だから，上の式はまとめて $\nabla f = \lambda^* G = \sum_{i=1}^m \lambda_i^* \nabla g_i(x^*)$ となる．　　以上 (2012-01-10)