

解析概論 I 第 12 回講義内容 (2004 年 7 月 16 日) 井上淳

1 年次の微積分学では時間的制約で講義されなかったであろう条件付き極値問題を考察する。この話題は是非とも理解して欲しい事柄でもある。

条件付き極値問題

定理 0.1 U を \mathbb{R}^n の開集合とし、 $f: U \rightarrow \mathbb{R}$ 、 $g: U \rightarrow \mathbb{R}^m$ は共に C^1 -級とする。いま

(a) $a \in S = \{x \in U \mid g(x) = 0\}$ において、 f が S 上の極値¹に達し、かつ

(b) $\text{rank } g'(a) = m$

であるとすれば、 $\lambda = (\lambda_1, \dots, \lambda_m) \in \mathbb{R}^m$ が存在して

$$f'(a) = \lambda g'(a)$$

が成立つ。

証明: $\text{rank } g'(a) = m$ より $n \geq m$ である。必要ならば座標 x_i の番号を入れ替えて

$$\frac{\partial(g_1, \dots, g_m)}{\partial(x_{n-m+1}, \dots, x_n)} \neq 0$$

と仮定してよい。このとき、陰関数定理より、 a の近傍で方程式 $g(x) = 0$ を x_{n-m+1}, \dots, x_n について解くことができる。今、 $x = {}^t(x_1, \dots, x_n) = {}^t(y, z)y = {}^t(x_1, \dots, x_{n-m})z = (x_{n-m+1}, \dots, x_n)$ と分解し、同様に $a = {}^t(b, c)b = {}^t(a_1, \dots, a_{n-m})c = {}^t(a_{n-m+1}, \dots, a_n)$ と書こう。このとき、 b, c の \mathbb{R}^{n-m} 、 \mathbb{R}^m における近傍 V 、 W と C^1 -級関数 $\varphi: V \rightarrow W$ が存在して、 $V \times W \subset U$ であり、 $x = {}^t(y, z) \in V \times W$ に対して

$$x \in S (g(z) = 0) \iff z = \varphi(y)$$

となる。更に、

$$\varphi'(y) = -\left(\frac{\partial g}{\partial z}\right)^{-1} \cdot \left(\frac{\partial g}{\partial y}\right)$$

で与えられる。関数を $F(y) = f(y, \varphi(y))$ と定めると、これは $n - m$ 次元の開集合 V 上で定義された C^1 -級関数で、一点 $b \in V$ で極値に達するから $F'(b) = 0$ となる。一方、

$$F'(y) = \frac{\partial f}{\partial y}(x) + \frac{\partial f}{\partial z}(x) \frac{\partial \varphi}{\partial y}(y) = \frac{\partial f}{\partial y}(x) - \frac{\partial f}{\partial z}(x) \left(\frac{\partial g}{\partial z}(x)\right)^{-1} \left(\frac{\partial g}{\partial y}(x)\right)$$

である。

$$\frac{\partial f}{\partial z}(a) \left(\frac{\partial g}{\partial z}(a)\right)^{-1} = \lambda \in \mathbb{R}^m$$

と定義すると、上の式は

$$\frac{\partial f}{\partial z}(a) = \lambda \frac{\partial g}{\partial z}(a)$$

¹ a のある開近傍 W が存在して、 $f(a)$ は $S \cap W$ 上の最大値または最小値となる

となる。また $F'(b) = 0$ を計算すれば

$$\frac{\partial f}{\partial y}(a) = \lambda \frac{\partial g}{\partial y}(a)$$

となる。故に

$$f'(a) = \left(\frac{\partial f}{\partial y}(a), \frac{\partial f}{\partial z}(a) \right) = \lambda \left(\frac{\partial g}{\partial y}(a), \frac{\partial g}{\partial z}(a) \right) = \lambda g'(a). \quad \square$$

系 0.1 (Lagrange 未定乗数法) U を \mathbb{R}^n の開集合とし、 $f : U \rightarrow \mathbb{R}$ 、 $g : U \rightarrow \mathbb{R}^m$ は共に C^1 -級とする。 $a \in S = \{x \in U \mid g(x) = 0\}$ が S 上での関数 f の極値点であるとき、次の (i), (ii) のいずれかが成立つ：

(i) $U \times \mathbb{R}^m$ 上の関数 $\Phi(x, \lambda) = f(x) - \lambda g(x)$ ($\lambda = (\lambda_1, \dots, \lambda_m) \in \mathbb{R}^m$) に対し、ある $\lambda^0 \in \mathbb{R}^m$ が存在して、次式が成立つ。

$$\Phi'(a, \lambda^0) = 0. \quad \text{i.e.} \quad \frac{\partial \Phi}{\partial x_j}(a, \lambda^0) = 0, \quad \frac{\partial \Phi}{\partial \lambda_k}(a, \lambda^0) = 0. \quad (1 \leq j \leq n, 1 \leq k \leq m).$$

或いは、

(ii) $\text{rank } g'(a) < m$.

証明：もし (ii) でないとすると、 $\text{rank } g'(a) = m$ である。故に上の定理より $f'(a) = \mu g'(a)$ となる定数 $\mu \in \mathbb{R}^m$ がある。一方、 $a \in S$ とすると $g(a) = 0$ だから $\lambda = \mu$ として

$$\Phi'(a, \lambda) = (f'(a) - \lambda g'(a), -g(a)) = 0$$

が、即ち (i) が成立する。 \square

例 1 [Hadamard の定理、評価式]: $x_j^2 + y_j^2 + z_j^2 = A_j^2$ ($j = 1, 2, 3$) という条件のもとで $\Delta = \det \begin{bmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{bmatrix}$

の最大値は $|A_1 A_2 A_3|$ に等しいこと、 $|\Delta| \leq (\sqrt{3}M)^3$ を示せ。但し、 $A_1 A_2 A_3 \neq 0$, $M = \max_j(|x_j|, |y_j|, |z_j|)$ とする。

解：

$$f(x_1, y_1, z_1, x_2, \dots, z_3) = \Delta - \frac{1}{2} \sum_{j=1}^3 \lambda_j (x_j^2 + y_j^2 + z_j^2 - A_j^2)$$

とおく。 \hat{x}_i を x_i での余因子すると、行列式の行に関する展開 $x_i \hat{x}_i + y_i \hat{y}_i + z_i \hat{z}_i = \Delta$ となる。故に、もしこの関数の極値があるならばその点で

$$f_{x_j} = \hat{x}_j - \lambda_j x_j = 0, \quad f_{y_j} = \hat{y}_j - \lambda_j y_j = 0, \quad f_{z_j} = \hat{z}_j - \lambda_j z_j = 0 \quad (1)$$

を満たす。もし $i \neq j$ ならば $\hat{x}_i x_j + \hat{y}_i y_j + \hat{z}_i z_j = 0$ となる。実際、 $i = 1, j = 2$

$$\det \begin{bmatrix} x_2 & y_2 & z_2 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{bmatrix} = \hat{x}_1 x_2 + \hat{y}_1 y_2 + \hat{z}_1 z_2 = 0, \quad \text{etc.}$$

故に、(1) より

$$\hat{x}_i x_j + \hat{y}_i y_j + \hat{z}_i z_j = \lambda_i (x_i x_j + y_i y_j + z_i z_j) = 0 \quad (i \neq j)$$

となる。これより

$$\Delta^2 = \begin{vmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{vmatrix} \begin{vmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{vmatrix} = \begin{vmatrix} a_1^2 & 0 & 0 \\ 0 & a_2^2 & 0 \\ 0 & 0 & a_3^2 \end{vmatrix} = a_1^2 a_2^2 a_3^2 \quad \therefore \Delta = \pm a_1 a_2 a_3$$

x_j, y_j, z_j の範囲は有界閉領域だから Δ は最大値 $|a_1 a_2 a_3|$ と最小値 $-|a_1 a_2 a_3|$ を取る。 $M = \max_j(|x_j|, |y_j|, |z_j|)$ だから、

$$a_j^2 \leq 3M^2, \quad |a_j| \leq \sqrt{3}M, \quad |\Delta| \leq (\sqrt{3}M)^3. \quad \square$$

註：この定理は $n \times n$ 行列に対してもそのまま拡張される。

例 2： $p_1 > 0, \dots, p_n > 0, p_1 + \dots + p_n = 1$ とする。 $H(p_1, \dots, p_n) = -\sum_{j=1}^n p_j \log p_j$ をエントロピー (entropy) と呼び、 H は $p_1 = \dots = p_n = 1/n$ で最大値をとる。

解：

$$F(p_1, \dots, p_n) = -\sum_{j=1}^n p_j \log p_j - \lambda(p_1 + \dots + p_n - 1)$$

と定めると、極値をとる点では

$$\frac{\partial F}{\partial p_j} = -\log p_j - 1 - \lambda = 0 \quad (j = 1, \dots, n) \because p_1 = \dots = p_n (= e^{-(1+\lambda)})$$

を満たす。一方、 (p_1, \dots, p_n) の範囲は有界、閉で、境界で 0、中で正だから、この点で最大値を持つ。 \square

練習問題 1: 定円に内接する三角形の中で面積最大のものが正三角形であることを、Lagrange の未定乗数法で証明せよ。(一般化して) 定円に内接する最大面積の n 角形を求めよ。

練習問題 2: $yz + zx + xy = 1$ から定まる関数 $z = z(x, y)$ の第 2 次導関数を求めよ ($z = \frac{1-xy}{x+y}$ と解いて計算したものと、陰関数のまま合成関数の微分則を用いて計算したものと計算量を比較せよ)。

練習問題 3: [変数変換] 関数 $u(x, y)$ の変数を $x + y = \xi, y = \xi\eta$ と変換して関数 $\tilde{u}(\xi, \eta) = u(\xi - \xi\eta, \xi\eta)$ とすると

$$(xu_x + yu_{xy} + u_x)(x, y) \Big|_{x+y=\xi, y=\xi\eta} = (\xi\tilde{u}_{\xi\xi} - \eta\tilde{u}_{\xi\eta} + \tilde{u}_{\xi})(\xi, \eta).$$

しばしば、上の等式は $xu_x + yu_{xy} + u_x = \xi\tilde{u}_{\xi\xi} - \eta\tilde{u}_{\xi\eta} + \tilde{u}_{\xi}$ と簡単に表記される。

練習問題 4: [微分方程式関連] $f(x, y)f_{xy}f(x, y) = f_x f(x, y)f_y f(x, y)$ が成立つとき、 $f(x, y) = \varphi(x)\psi(y)$ と表される。

=== おまけ ===

無限次元空間での条件付き極値問題 : 曲線 $x = x(t)$ に対して汎関数

$$J[x] = \int_a^b F(t, x(t), x'(t)) dt$$

を、許容される曲線群

$$\mathcal{X}_\ell = \{x(\cdot) \in C^1([a, b] : \mathbb{R}) \mid x(a) = A, x(b) = B, K[x] = \int_a^b G(t, x(t), x'(t)) dt = \ell\}$$

の中で考える。このとき、以下の定理が成立する。

定理 0.2 F と G を、 $F, G \in C^1([a, b] \times \mathbb{R} \times \mathbb{R} : \mathbb{R})$ であり、かつ x と x' を固定するとき、 $C^2[a, b]$ なるものとする。 $x = x_c(t)$ を $J[x]$ の極値と仮定するとき、 $x = x_c(t)$ が $K[x]$ の極値でないとすると、ある λ があって、 $x = x_c(t)$ は汎関数

$$\Phi(x, \dot{x}) = \int_a^b (F(t, x(t), \dot{x}(t)) + \lambda G(t, x(t), \dot{x}(t))) dt$$

の極値になっている。即ち、 $x = x_c(t)$ は

$$F_x - \frac{d}{dt}F_{\dot{x}} + \lambda \left(G_x - \frac{d}{dt}G_{\dot{x}} \right) = 0$$

を満たす。

少し雑に説明すると、もし $x_c(t)$ が汎関数 $\Phi(x, x')$ の極値になっているとするならば、その $x_c(t)$ を少し動かした時どうなっているべきだろうと考える。そこで、任意の $\varphi(t)$ と 0 の近くの τ を考え、 $x_c(t) + \tau\varphi(t)$ が許容されているとしよう。任意の $\varphi(t)$ を一つ止めて τ の関数

$$\Phi(x_c + \tau\varphi, \dot{x}_c + \tau\dot{\varphi}) = \int_a^b (F(t, x_c(t) + \tau\varphi(t), \dot{x}_c(t) + \tau\dot{\varphi}(t)) + \lambda G(t, x_c(t) + \tau\varphi(t), \dot{x}_c(t) + \tau\dot{\varphi}(t))) dt$$

を考える。“極値”だとすると

$$\left. \frac{d}{d\tau} \Phi(x_c + \tau\varphi, \dot{x}_c + \tau\dot{\varphi}) \right|_{\tau=0} = 0$$

となるべきだろう。この式を計算する。

$$\begin{aligned} \left. \frac{d}{d\tau} \Phi(x_c + \tau\varphi, \dot{x}_c + \tau\dot{\varphi}) \right|_{\tau=0} &= \int_a^b \left[F_x \varphi + F_{\dot{x}} \dot{\varphi} + \lambda \left(G_x \varphi + G_{\dot{x}} \dot{\varphi} \right) \right] dt \\ &= \int_a^b \left[F_x - \frac{d}{dt}F_{\dot{x}} + \lambda \left(G_x - \frac{d}{dt}G_{\dot{x}} \right) \right] \varphi(t) dt \end{aligned}$$

ここで $H(t) = F_x(t, x_c(t) + \tau\varphi(t), \dot{x}_c(t) + \tau\dot{\varphi}(t))$ とおくと部分積分より

$$\int_a^b H(t) \dot{\varphi}(t) dt = - \int_a^b \frac{d}{dt} H(t) \varphi(t) dt$$

となること、連続関数 $I(t)$ に対し十分沢山の φ で

$$\int_a^b I(t) \varphi(t) dt = 0 \implies I(t) = 0$$

となることを用いた。□

古典力学であらわれる Lagrangian density $L(t, \gamma, \dot{\gamma})$ ($\gamma = (\gamma_1, \dots, \gamma_n)$, $\dot{\gamma} = (\dot{\gamma}_1, \dots, \dot{\gamma}_n)$) に対して

$$\mathcal{L}(\gamma, \dot{\gamma}) = \int_0^T L(t, \gamma, \dot{\gamma}) dt$$

とするとき

$$\delta \mathcal{L}(\gamma, \dot{\gamma}) = 0 \quad \text{とは} \quad \frac{d}{dt} L_{\dot{\gamma}_j} = - \frac{\partial}{\partial \gamma_j} L_{\gamma_j}$$

としたのもので、上と同様の推論で導かれる。

==== おまけ終わり ====

==== 付け足し ====

練習問題 1: 定円に内接する三角形の中で面積最大のものが正三角形であることを、Lagrange の未定乗数法で証明せよ。(一般化して) 定円に内接する最大面積の n 角形を求めよ。

解答例: 定円の半径を r とし、内接三角形の 3 辺の中心角を $x, y, 2\pi - x - y$ とすると、三角形の面積 $S = S(x, y)$ は

$$S = \frac{r^2}{2} [\sin x + \sin y - \sin(x + y)]$$

と与えられる。もし頂角の一つが鈍角であれば、その対辺を中心を超えて移動し内接させれば面積は大きくなるから、面積最大になるものは直角または鋭角三角形である。故に

$$D = \{(x, y) \mid 0 < x \leq \pi, 0 < y \leq \pi, \pi \leq x + y < 2\pi\}$$

で S の最大値を調べれば良い。極値点では

$$S_x = \frac{r^2}{2} [\cos x - \cos(x+y)] = 0, \quad S_y = \frac{r^2}{2} [\cos y - \cos(x+y)] = 0$$

となるから、 $\cos x = \cos(x+y) = \cos y$ であり、これを満たす極値点 (x, y) は D 内では $x = y = 2\pi/3$ のみである。また、

$$\begin{aligned} \det \begin{pmatrix} S_{xx} & S_{xy} \\ S_{yx} & S_{yy} \end{pmatrix} &= \det \begin{pmatrix} -\sin x + \sin(x+y) & \sin(x+y) \\ \sin(x+y) & -\sin y + \sin(x+y) \end{pmatrix} \\ &= \sin x \sin y - (\sin x + \sin y) \sin(x+y) \quad (> 0 \text{ at } x = y = 2\pi/3), \end{aligned}$$

より、 $x = y = 2\pi/3$ で極大値をとる。一方 S は D 内で正、 D の境界 ∂D で 0 だから、内部に最大値をとる点が存在する。故に、それが点 $x = y = 2\pi/3$ であり、そこでの最大値は $3\sqrt{3}r^2/4$ である。

練習問題 2: $yz + zx + xy = 1$ から定まる関数 $z = z(x, y)$ の第 2 次導関数を求めよ ($z = \frac{1-xy}{x+y}$ と解いて計算したものと、陰関数のまま合成関数の微分則を用いて計算したものと計算量を比較せよ)。

解答例: この例の場合は、関数 z が簡単に x, y の陽な関数として $z = \frac{1-xy}{x+y}$ と簡単に表示できるので、それを偏微分して

$$z_x = \frac{-y(x+y) - (1-xy)}{(x+y)^2} = \frac{-1-y^2}{(x+y)^2}, \quad z_y = \frac{-x(x+y) - (1-xy)}{(x+y)^2} = \frac{-1-x^2}{(x+y)^2}.$$

陰関数のまま合成関数の微分則を用いて計算すると、

$$yz_x + z + xz_x + y = 0, \quad z + yz_y + xz_y + x = 0.$$

最初の式に x と y を掛けて、

$$x(x+y)z_x + xy + xz = 0, \quad y(x+y)z_x + y^2 + yz = 0 \implies (x+y)^2 z_x + xy + y^2 + xz + yz = 0$$

となる。これと $yz + zx + xy = 1$ より $z_x = \frac{-1-y^2}{(x+y)^2}$ が求められる。

=====

メモ: 中間試験の感想を見て、私の意図したことが必ずしも伝わっていないことを痛感した。そこで、何故、普通の Euclid 空間だけでなく距離空間と言う概念で特に Banach 空間で微分を考えたかの説明を再度した。また、極限概念の一つの拡張である上極限、下極限について、期末試験でも出題することを約束し、勉強してくれることを期待した。受講者は 40 名前後か、中間試験を受けた人数程度。