数学解析第１　第４回講義ノート

3　極値問題

まず初めに，通常の極値問題を考えよう．2 変数関数に対する極値問題の扱いは 1 年次の数学 3A で習って来ている．それを n 変数関数に対して拡張することは難しくないが，ここでは線形代数の知識を用いてもう少し見通しの良い証明を与えることにしよう．

定理 3.1　Ω を $\mathbf{R}^n$ の領域，$\boldsymbol{a} \in \Omega$，および $f \in C^1(\Omega)$ とする．このとき，f が $\boldsymbol{a}$ で極値をとれば，$\nabla f(\boldsymbol{a}) = \mathbf{0}$ が成り立つ．

証明　各 $j \in \{1, 2, \ldots, n\}$ を固定するごとに，x_j の 1 変数関数

$$\varphi_j(x_j) := f(a_1, \ldots, a_{j-1}, x_j, a_{j+1}, \ldots, a_n)$$

は C^1 級であり，$x_j = a_j$ において極値をとる．したがって，

$$\varphi_j'(a_j) = \frac{\partial f}{\partial x_j}(\boldsymbol{a}) = 0 \qquad (1 \leq j \leq n)$$

となり，$\nabla f(\boldsymbol{a}) = \mathbf{0}$ が従う．（証明終）

1 変数関数のときもそうであったように，n 変数関数のときも $\nabla f(\boldsymbol{a}) = \mathbf{0}$ であっても，f が $\boldsymbol{a}$ で極値をとるとは限らないことに注意しよう．

定義 3.1　$\nabla f(\boldsymbol{a}) = \mathbf{0}$ を満たす点 $\boldsymbol{a}$ を，f の停留点とよぶ．

この用語を用いれば，定理 3.1 は極値点は常に停留点であることを主張している．しかし，その逆は必ずしも正しくない．つまり，停留点であっても極値点でない場合があることに注意しよう．いずれにしても，極値点を見つけるときは，まずその候補として停留点を見つければよい．次に，1 変数関数のとき極値の判定に 2 階の導関数が用いられたように，n 変数関数のときにも 2 階の偏導関数が極値の判定に用いられる．そこで，次の定義をしておこう．

定義 3.2　$f \in C^2(\Omega)$ に対して，$n \times n$ 実対称行列

$$\begin{aligned} H_f(\boldsymbol{x}) &:= \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(\boldsymbol{x}) \right)_{1 \leq i,j \leq n} \\ &= \begin{pmatrix} \dfrac{\partial^2 f}{\partial x_1^2}(\boldsymbol{x}) & \dfrac{\partial^2 f}{\partial x_1 \partial x_2}(\boldsymbol{x}) & \cdots & \dfrac{\partial^2 f}{\partial x_1 \partial x_n}(\boldsymbol{x}) \\ \dfrac{\partial^2 f}{\partial x_1 \partial x_2}(\boldsymbol{x}) & \dfrac{\partial^2 f}{\partial x_2^2}(\boldsymbol{x}) & \cdots & \dfrac{\partial^2 f}{\partial x_2 \partial x_n}(\boldsymbol{x}) \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial^2 f}{\partial x_1 \partial x_n}(\boldsymbol{x}) & \dfrac{\partial^2 f}{\partial x_2 \partial x_n}(\boldsymbol{x}) & \cdots & \dfrac{\partial^2 f}{\partial x_n^2}(\boldsymbol{x}) \end{pmatrix} \end{aligned}$$

を f の Hesse 行列とよぶ．

定理 3.2（極値の判定） Ω を $\mathbf{R}^n$ の領域，$\boldsymbol{a} \in \Omega$，$f \in C^2(\Omega)$，および $\nabla f(\boldsymbol{a}) = \mathbf{0}$ とする．このとき，

(1) $H_f(\boldsymbol{a})$ の固有値が全て正ならば，f は $\boldsymbol{a}$ で極小となる．
(2) $H_f(\boldsymbol{a})$ の固有値が全て負ならば，f は $\boldsymbol{a}$ で極大となる．
(3) $H_f(\boldsymbol{a})$ が正と負の固有値をもてば，$\boldsymbol{a}$ は f の鞍点である．

この定理を証明する前に，線形代数の定理を一つ思い出しておこう．

補題 3.1 (1) 実対称行列は直交行列で対角化可能である．すなわち，任意の $n \times n$ 実対称行列 H に対して，ある $n \times n$ 直交行列 T が存在し，

$$T^{-1}HT = \begin{pmatrix} \lambda_1 & & & \text{\huge 0} \\ & \lambda_2 & & \\ & & \ddots & \\ \text{\huge 0} & & & \lambda_n \end{pmatrix}$$

が成り立つ．ここで，$\lambda_1, \lambda_2, \dots, \lambda_n$ は H の固有値である．

(2) T を $n \times n$ 直交行列とすると，$T^{-1} = T^t$（T の転置行列）および

$$\|T\boldsymbol{x}\| = \|\boldsymbol{x}\| \qquad (\forall \boldsymbol{x} \in \mathbf{R}^n)$$

が成り立つ.（このとき，$\boldsymbol{y} = T\boldsymbol{x}$ により，直交座標系 $\boldsymbol{x}$ から他の直交座標系 $\boldsymbol{y}$ への変換が与えられる．逆に，そのような変換を与える行列 T は必ず直交行列になる.）

定理 3.2 の証明 f を $\boldsymbol{a}$ のまわりで Taylor 展開しよう．$f \in C^2(\Omega)$ であることから，

$$\begin{aligned} f(\boldsymbol{x}) &= \sum_{|\alpha| \le 2} \frac{f^{(\alpha)}(\boldsymbol{a})}{\alpha!}(\boldsymbol{x} - \boldsymbol{a})^\alpha + R(\boldsymbol{x}) \\ &= f(\boldsymbol{a}) + \sum_{i=1}^n \frac{\partial f}{\partial x_i}(\boldsymbol{a})(x_i - a_i) + \frac{1}{2}\sum_{i=1}^n \sum_{j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j}(\boldsymbol{a})(x_i - a_i)(x_j - a_j) + R(\boldsymbol{x}) \\ &= f(\boldsymbol{a}) + (\boldsymbol{x} - \boldsymbol{a}) \cdot \nabla f(\boldsymbol{a}) + \frac{1}{2}(\boldsymbol{x} - \boldsymbol{a}) \cdot \big(H_f(\boldsymbol{a})(\boldsymbol{x} - \boldsymbol{a})\big) + R(\boldsymbol{x}) \end{aligned}$$

が成り立つ．ここで，$R(\boldsymbol{x})$ は剰余項であり，

$$\lim_{\boldsymbol{x} \to \boldsymbol{a}} \frac{R(\boldsymbol{x})}{\|\boldsymbol{x} - \boldsymbol{a}\|^2} = 0 \tag{3.1}$$

を満たす．また，上式における行列とベクトルの積の計算では，ベクトルは全て縦ベクトルとみなして計算している．ここで，仮定より $\nabla f(\boldsymbol{a}) = \mathbf{0}$ であるから，

$$f(\boldsymbol{x}) - f(\boldsymbol{a}) = \frac{1}{2}(\boldsymbol{x} - \boldsymbol{a}) \cdot \big(H_f(\boldsymbol{a})(\boldsymbol{x} - \boldsymbol{a})\big) + R(\boldsymbol{x})$$

が得られる．次に，$\boldsymbol{x}$ が $\boldsymbol{a}$ の近傍を動くとき，上式の右辺の符号変化を調べよう．補題 3.1 より，対称行列 $H_f(\boldsymbol{a})$ は，ある直交行列 T により対角化される．すなわち，

$$T^{-1}H_f(\boldsymbol{a})T = \begin{pmatrix} \lambda_1 & & \text{\huge 0} \\ & \ddots & \\ \text{\huge 0} & & \lambda_n \end{pmatrix} =: \Lambda$$

が成り立つ．ここで，$\lambda_1, \dots, \lambda_n$ は $H_f(\boldsymbol{a})$ の固有値である．この直交行列 T を用いて，$\boldsymbol{y} := T^{-1}(\boldsymbol{x}-\boldsymbol{a})$ により独立変数を $\boldsymbol{x}$ から $\boldsymbol{y}$ へ変換しよう．これは，$\boldsymbol{a}$ を原点とする新しい直交座標系を用いていることに対応する．このとき，

$$\begin{aligned} f(\boldsymbol{x}) - f(\boldsymbol{a}) &= \frac{1}{2}(T\boldsymbol{y}) \cdot (H_f(\boldsymbol{a})T\boldsymbol{y}) + R(\boldsymbol{x}) \\ &= \frac{1}{2}\boldsymbol{y} \cdot (T^t H_f(\boldsymbol{a})T\boldsymbol{y}) + R(\boldsymbol{x}) \\ &= \frac{1}{2}\boldsymbol{y} \cdot (\Lambda\boldsymbol{y}) + R(\boldsymbol{x}) \\ &= \frac{1}{2}(\lambda_1 y_1^2 + \cdots \lambda_n y_n^2) + R(\boldsymbol{x}) \end{aligned}$$

上の計算では，Euclid 内積と転置行列の関係 $(A\boldsymbol{x}) \cdot \boldsymbol{y} = \boldsymbol{x} \cdot (A^t\boldsymbol{y})$ $(\boldsymbol{x}, \boldsymbol{y} \in \mathbf{R}^n)$ および $T^t = T^{-1}$ という事実を用いた．

(1) の場合：$\lambda_j > 0$ $(1 \le j \le n)$ であるから，

$$\lambda_0 := \min\{\lambda_1, \dots, \lambda_n\}$$

とおくと，$\lambda_j \ge \lambda_0 > 0$ および $\|\boldsymbol{y}\| = \|T\boldsymbol{y}\| = \|\boldsymbol{x}-\boldsymbol{a}\|$ に注意すれば，

$$\begin{aligned} f(\boldsymbol{x}) - f(\boldsymbol{a}) &\ge \frac{1}{2}\lambda_0(y_1^2 + \cdots + y_n^2) + R(\boldsymbol{x}) \\ &= \frac{1}{2}\lambda_0\|\boldsymbol{x}-\boldsymbol{a}\|^2 + R(\boldsymbol{x}) \\ &= \left(\frac{1}{2}\lambda_0 + \frac{R(\boldsymbol{x})}{\|\boldsymbol{x}-\boldsymbol{a}\|^2}\right)\|\boldsymbol{x}-\boldsymbol{a}\|^2 \end{aligned}$$

ここで，剰余項 $R(\boldsymbol{x})$ は (3.1) を満たすことから，十分小さな正定数 $\delta > 0$ が存在し，$0 < \|\boldsymbol{x}-\boldsymbol{a}\| < \delta$ を満たす任意の $\boldsymbol{x}$ に対して

$$\frac{|R(\boldsymbol{x})|}{\|\boldsymbol{x}-\boldsymbol{a}\|^2} \le \frac{1}{4}\lambda_0$$

が成り立つ．したがって，

$$f(\boldsymbol{x}) - f(\boldsymbol{a}) \ge \frac{1}{4}\lambda_0\|\boldsymbol{x}-\boldsymbol{a}\|^2 > 0 \quad \text{for} \quad 0 < \|\boldsymbol{x}-\boldsymbol{a}\| < \delta$$

となり，f は $\boldsymbol{a}$ で極小値をとることが分かる．

(2) の場合：$\lambda_j < 0$ $(1 \le j \le n)$ であるから，

$$\lambda_0 := \max\{\lambda_1, \dots, \lambda_n\}$$

とおくと，$\lambda_j \le \lambda_0 < 0$ より，

$$\begin{aligned} f(\boldsymbol{x}) - f(\boldsymbol{a}) &\le \frac{1}{2}\lambda_0(y_1^2 + \cdots + y_n^2) + R(\boldsymbol{x}) \\ &= \left(\frac{1}{2}\lambda_0 + \frac{R(\boldsymbol{x})}{\|\boldsymbol{x}-\boldsymbol{a}\|^2}\right)\|\boldsymbol{x}-\boldsymbol{a}\|^2 \end{aligned}$$

となる．したがって，(1) の場合と同様にして，十分小さな正定数 $\delta > 0$ が存在し，

$$f(\boldsymbol{x}) - f(\boldsymbol{a}) \le \frac{1}{4}\lambda_0\|\boldsymbol{x}-\boldsymbol{a}\|^2 < 0 \quad \text{for} \quad 0 < \|\boldsymbol{x}-\boldsymbol{a}\| < \delta$$

が成り立つ．ゆえに，f は $\boldsymbol{a}$ で極大値をとることが分かる．

(3) の場合：簡単のために，$\lambda_1 > 0 > \lambda_2$ としよう．このとき，まず y_1 軸に沿って f の値の変化を調べてみよう．すなわち，$\boldsymbol{y} = (y_1, 0, \ldots, 0)$ 上の f の値を見てみよう．このとき，$\boldsymbol{x} = \boldsymbol{a} + (T\boldsymbol{e}_1)y_1$ および $\boldsymbol{e}_1 = (1, 0, \ldots, 0)^t$ であり，

$$f(\boldsymbol{x}) - f(\boldsymbol{a}) = \frac{1}{2}\lambda_1 y_1^2 + R(\boldsymbol{x}), \qquad \lim_{y_1 \to 0} \frac{R(\boldsymbol{x})}{y_1^2} = 0$$

が成り立つ．これより，y_1 軸に沿って f の値の変化を調べると，f は $y_1 = 0$ （すなわち，$\boldsymbol{x} = \boldsymbol{a}$）において極小になることが分かる．同様にして，$y_2$ 軸に沿って f の値の変化を調べると，すなわち，$\boldsymbol{y} = (0, y_2, 0, \ldots, 0)$ 上の f の値を見てみると，f は $y_2 = 0$（すなわち，$\boldsymbol{x} = \boldsymbol{a}$）において極大になることが分かる．ゆえに，この場合 $\boldsymbol{a}$ は f の鞍点である．（証明終）

例 3.1　$f(x, y) = x^3 - y^3 - 3x + 12y$ の極大値と極小値を求めてみよう．まず，f の停留点を求めていく．

$$\nabla f(x, y) = (3x^2 - 3, -3y^2 + 12)$$

より，$\nabla f(x, y) = \mathbf{0}$ を解くと

$$(x, y) = (\pm 1, \pm 2) \qquad \text{（複合任意）}$$

となる．これらが f の停留点である．次に，定理 3.2 を用いて，これらの点が極値点であるかどうかを判定する．

$$H_f(x, y) = \begin{pmatrix} 6x & 0 \\ 0 & -6y \end{pmatrix}$$

より，f の Hesse 行列 $H_f(x, y)$ の固有値は $6x, -6y$ である．したがって，f は $(x, y) = (1, -2)$ で極小，$(x, y) = (-1, 2)$ で極大となり，$(x, y) = \pm(1, 2)$ は f の鞍点である．

条件付き極値問題　実際の応用では，上のような何かの関数の極大値や極小値を求めるような問題ばかりではなく，ある制約条件下で関数の極大値や極小値を求める問題も，しばしば現れる．そのような制約条件は，適当な関数 $g_1, \ldots, g_m$ を用いて

$$g_1(\boldsymbol{x}) = \cdots = g_m(\boldsymbol{x}) = 0$$

と書ける場合が多い．以下では，そのような制約条件下での $f(\boldsymbol{x})$ の極大値・極小値を求めていこう．そのような問題のことを，条件付き極値問題という．この場合には，もはや定理 3.1 は役に立たないが，その代わりに次の定理が有用である．

定理 3.3（Lagrange の未定乗数法）　Ω を $\mathbf{R}^n$ の領域，f および $\boldsymbol{g} = (g_1, \ldots, g_m)$ を Ω で定義された C^1 級関数とし，$S := \{\boldsymbol{x} \in \Omega \,|\, \boldsymbol{g}(\boldsymbol{x}) = \mathbf{0}\}$ とおく．このとき，

(1)　f は $\boldsymbol{a} \in S$ において S 上極値をとり

(2)　$\operatorname{rank}(D\boldsymbol{g}(\boldsymbol{a})) = m \ (\leq n)$

ならば，ある $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_m) \in \mathbf{R}^m$ が存在し

$$Df(\boldsymbol{a}) + \boldsymbol{\lambda} D\boldsymbol{g}(\boldsymbol{a}) = \mathbf{0}$$

が成り立つ．この最後の式は，次式と同値である．

$$\big(\nabla(f + \lambda_1 g_1 + \cdots + \lambda_m g_m)\big)(\boldsymbol{a}) = \mathbf{0}$$