第1章　制約無し最適化

1.1　復習：2変数関数の極値

1.1.1　判別式による方法

関数 $f(x,y)$ の極値を求めるには，連立方程式

$$\begin{cases} f_x(x,y)=0 \\ f_y(x,y)=0 \end{cases}$$

を解いて停留点 $(x,y)=(a,b)$ を求め，さらに

$$\Delta(x,y)=f_{xx}(x,y)f_{yy}(x,y)-(f_{xy}(x,y))^2$$

とおくとき，

定理 1.1.1 (2変数関数の極大・極小).
i) $\Delta(a,b)>0$ かつ $f_{xx}(a,b)>0$ ならば $f(x,y)$ は (a,b) で極小.
ii) $\Delta(a,b)>0$ かつ $f_{xx}(a,b)<0$ ならば $f(x,y)$ は (a,b) で極大.
iii) $\Delta(a,b)<0$ ならば $f(x,y)$ は (a,b) で極値をとらない.

という定理を用いた.

例題

1. $f(x,y)=x^3+xy^2+x^2+y^2-8x$ の極値を求めよ.
2. a,b について場合分けして，$f(x,y)=x^2+axy+by^2$ の極値を求めよ.

1.1.2　2変数関数のテイラー展開と2次形式

上の極値の判別法を一般化したいが，いきなり n 変数でやるとややこしいので，まず2変数の場合を整理する.

$(x,y)=(a+h,b+k)$ として $f(x,y)$ を (a,b) を中心にテーラー展開すると

$$f(a+h,b+k) \;=\; f(a,b)+hf_x(a,b)+kf_y(a,b)+\cdots+\frac{1}{m!}\Big(h\frac{\partial}{\partial_x}+k\frac{\partial}{\partial_y}\Big)^m f(a,b)+R_{m+1}$$

ここで R_{m+1} はここで R_{m+1} は適当な実数 $0<\theta<1$ によって

$$R_{m+1}=\frac{1}{(m+1)!}\Big(h\frac{\partial}{\partial_x}+k\frac{\partial}{\partial_y}\Big)^{m+1} f(a+\theta h,b+\theta k)$$

と書かれる．(a,b) の近傍で R_{m+1} は $|(h,k)|^m$ よりも十分小さい．

特に 2 次まで展開は

$$\begin{aligned} f(a+h,b+k) \;&=\; f(a,b)+hf_x(a,b)+kf_y(a,b) \\ &\quad +\frac{1}{2!}\Big(h^2f_{xx}(a,b)+2hkf_{xy}(a,b)+k^2f_{yy}(a,b)\Big)+o(|(h,k)|^2) \end{aligned}$$

となる．ここで 2 次の項は

$$\frac{1}{2}\Big(h^2f_{xx}+2hkf_{xy}+k^2f_{yy}\Big) \;=\; \frac{1}{2}\Big(h,k\Big)\begin{pmatrix} f_{xx}(a,b) & f_{xy}(a,b) \\ f_{xy}(a,b) & f_{yy}(a,b)\end{pmatrix}\begin{pmatrix} h \\ k\end{pmatrix}$$

と書ける．行列

$$H(x,y) \;=\; \begin{pmatrix} f_{xx}(x,y) & f_{xy}(x,y) \\ f_{xy}(x,y) & f_{yy}(x,y)\end{pmatrix}$$

をヘッセ行列という．

定理 1.1.2 (2 次形式の符号)**.**
$h^2f_{xx}(a,b)+2hkf_{xy}(a,b)+k^2f_{yy}(a,b)$ に対して次が成立する．

$(h,k)\neq(0,0)$ において常に正	$\Longleftrightarrow$	$H(a,b)$ の固有値は全て正（正定値）
$(h,k)\neq(0,0)$ において常に負	$\Longleftrightarrow$	$H(a,b)$ の固有値は全て負（負定値）
$(h,k)\neq(0,0)$ において正にも負にもなる	$\Longleftrightarrow$	$H(a,b)$ の固有値は正のもの と負のものある（不定符号）

さらに次が成立する．

定理 1.1.3 (対称行列の符号の判定)**.**

$H(a,b)$ が正定値	$\Longleftrightarrow$	$\Delta(a,b)>0$ かつ $f_{xx}(a,b)>0$
$H(a,b)$ が負定値	$\Longleftrightarrow$	$\Delta(a,b)>0$ かつ $f_{xx}(a,b)<0$
$H(a,b)$ が不定符号	$\Longleftrightarrow$	$\Delta(a,b)<0$

停留点の近傍において，$f(x,y)$ を $f(a,b)$ と 2 次の項までで近似することにより，この 2 つの定理から，前節の定理 1.1.1 が得られる．

1.2　多変数の極値問題

1.2.1　多変数のテイラー展開

$\mathbf{x} = (x_1, x_2, \cdots, x_n)$, $\mathbf{a} = (a_1, a_2, \cdots, a_n)$, $\mathbf{h} = (h_1, h_2, \cdots, h_n)$, $\dfrac{\partial}{\partial x_i} = \partial_i$ と書くとき，

$$\begin{aligned} f(\mathbf{a}+\mathbf{h}) &= f(\mathbf{a}) + \Big(h_1\partial_1 + \cdots + h_n\partial_n\Big) f(\mathbf{a}) + \frac{1}{2!}\Big(h_1\partial_1 + \cdots + h_n\partial_n\Big)^2 f(\mathbf{a}) \\ &\quad + \cdots + \frac{1}{m!}\Big(h_1\partial_1 + \cdots + h_n\partial_n\Big)^m f(\mathbf{a}) + R_{m+1} \end{aligned}$$

ここで R_{m+1} は適当な実数 $0 < \theta < 1$ によって

$$R_{m+1} = \frac{1}{(m+1)!}\Big(h_1\partial_1 + \cdots + h_n\partial_n\Big)^{m+1} f(\mathbf{a}+\theta\mathbf{h})$$

と書かれる．特に R_{m+1} は h_i たちの斉次 $m+1$ 次式なので

$$R_{m+1} = o(|\mathbf{h}|^m)$$

である（$\displaystyle\lim_{|\mathbf{h}|\to 0} \frac{R_{m+1}}{|\mathbf{d}|^m} = 0$ ということ）．

2 次までとると

$$f(\mathbf{a}+\mathbf{h}) = f(\mathbf{a}) + \mathbf{h}\cdot\nabla f(\mathbf{a}) + \frac{1}{2}\mathbf{h}^\top (Hf(\mathbf{a}))\mathbf{h} + o(|\mathbf{h}|^2)$$

となる．ここで ∇f は

$$\nabla f(\mathbf{x}) = \begin{pmatrix} f_{x_1}(\mathbf{x}) \\ f_{x_2}(\mathbf{x}) \\ \vdots \\ f_{x_n}(\mathbf{x}) \end{pmatrix}$$

（$f(\mathbf{x})$ の $\mathbf{x}$ における勾配ベクトル gradient）であり，Hf は

$$Hf(\mathbf{x}) = \begin{pmatrix} f_{x_1x_1}(\mathbf{x}) & f_{x_1x_2}(\mathbf{x}) & \cdots & f_{x_1x_n}(\mathbf{x}) \\ f_{x_1x_2}(\mathbf{x}) & f_{x_2x_2}(\mathbf{x}) & \cdots & f_{x_2x_n}(\mathbf{x}) \\ \vdots & \vdots & \ddots & \vdots \\ f_{x_1x_n}(\mathbf{x}) & f_{x_2x_n}(\mathbf{x}) & \cdots & f_{x_nx_n}(\mathbf{x}) \end{pmatrix}$$

（$f(\mathbf{x})$ の $\mathbf{x}$ におけるヘッセ行列）である．

1.2.2　2 次形式

一般に 2 次形式 (斉次 2 次多項式)$g(\mathbf{h}) = \sum_{1\le i\le j\le n} a_{ij}h_ih_j$ は

$$g(\mathbf{d}) = (h_1, h_2, \cdots, h_n)\begin{pmatrix} a_{11} & \frac{a_{12}}{2} & \cdots & \frac{a_{1n}}{2} \\ \frac{a_{12}}{2} & a_{22} & \cdots & \frac{a_{2n}}{2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{a_{1n}}{2} & \frac{a_{2n}}{2} & \cdots & a_{nn} \end{pmatrix}\begin{pmatrix} h_1 \\ h_2 \\ \vdots \\ h_n \end{pmatrix} = \mathbf{h}^\top A\mathbf{h}$$

と書ける．テイラー展開の 2 次の項は (簡単のため $f_{x_i} = f_i$ などと書くと)

$$\begin{aligned}\frac{1}{2!}(h_1\partial_1 + \cdots h_n\partial_n)^2 f(\mathbf{a}) &= \frac{1}{2}\Big(f_{11}h_1^2 + f_{22}h_2^2 + \cdots + f_{nn}h_n^2 \\ &\quad +2f_{12}h_1h_2 + \cdots + 2f_{n-1,n}h_{n-1}h_n\Big)\end{aligned}$$

なので確かに $\dfrac{1}{2}\mathbf{h}^\top(Hf(\mathbf{a}))\mathbf{h}$ となっている．

定理 1.2.1 (2 次形式の符号)**.** 2 次形式 $g(\mathbf{h}) = \mathbf{h}^\top A\mathbf{h}$ の $\mathbf{h} \neq 0$ における値に関して次が成立する．

常に 0 以上	$\Longleftrightarrow$	A の固有値は全て 0 以上 (半正定値)
常に正	$\Longleftrightarrow$	A の固有値は全て正 (正定値)
常に 0 以下	$\Longleftrightarrow$	A の固有値は全て 0 以下 (半正定値)
常に負	$\Longleftrightarrow$	A の固有値は全て負 (正定値)
正にも負にもなる	$\Longleftrightarrow$	A の固有値は正のものと負のものがある (不定値)

定理 1.2.2 (正定値の判定条件)**.**

2 次形式 $g(\mathbf{h}) = \mathbf{h}^\top A\mathbf{h}$ が正定値

$\Longleftrightarrow$　A の固有値は全て正

$\Longleftrightarrow$　A の主座小行列式 $|A_1|, |A_2|, \cdots, |A_n|$ は全て正

$\Longleftrightarrow$　$h'_i = \sum_{j=1}^{n} b_{ij}h_j$ という形の変換で

$g(\mathbf{x}) = h_1'^2 + h_2'^2 + \cdots + h_n'^2$ と書ける *(ラグランジュの方法)*.

1.2.3　極値問題

最小化	$f(\mathbf{x})$
制約条件	なし $(\mathbf{x} \in \mathbb{R}^n)$

を制約なし最適化問題という．

定理 1.2.3 (多変数関数の極大・極小).
1. $\mathbf{a}$ が $f(\mathbf{x})$ の極値ならば $\nabla f(\mathbf{a}) = \mathbf{0}$.
2. $\nabla f(\mathbf{a}) = \mathbf{0}$（停留点）のとき,

(1) $Hf(\mathbf{a})$ が正定値 $\Longrightarrow$ $f(\mathbf{a})$ は極小値.

(2) $Hf(\mathbf{a})$ が負定値 $\Longrightarrow$ $f(\mathbf{a})$ は極大値.

(3) $Hf(\mathbf{a})$ が不定符号 $\Longrightarrow$ $f(\mathbf{a})$ は極値でない.

が成立する.

例題

1. $f(x, y, z) = 2xy + 2yz + 2zx$ のヘッセ行列の固有値を求めよ.

2. $f(x, y, z) = 2x^3 + 6xy^2 + 3x^2 - 3y^2 - z^2$ の極値を求めよ.

3. n 点 (x_i, y_i) との距離の 2 乗の和が最小になる点 (a, b) を求めよ.

4. (最小 2 乗法により回帰直線を求める) 直線 $y = ax + b$ と n 点 (x_i, y_i) に関して, $x = x_i$ における y 座標の差の 2 乗の和
$$\sum_{i=1}^{n}(y_i - (ax_i + b))^2$$
が最小になるときの a, b を求めよ