多変数の微分積分学1 第12回

桂田 祐史

2013年7月8日

目 次

この授業用の WWW ページは
`http://www.math.meiji.ac.jp/~mk/lecture/tahensuu1-2013/`

今日は §8.4 からである。

8 極値問題

8.1 まずは問題から

K を $\mathbf{R}^2$ 内の3点 $(0,0)$, $(1,0)$, $(0,1)$ を頂点とする三角形とする: $K=\{(x,y)\in\mathbf{R}^2\ x\ge 0,\ y\ge 0,\ x+y\le 1\}$. $f\colon K\to\mathbf{R}$ を

$$f(x,y)=3x^2+2y^2+2xy-2x-2y+1$$

で定めるとき、f の最大値、f の最小値を求めよ。

とりあえず微分してみましょう。

$$f'(x,y)=\begin{pmatrix}\dfrac{\partial f}{\partial x} & \dfrac{\partial f}{\partial y}\end{pmatrix}=(6x+2y-2 \;\; 4y+2x-2)\,.$$

これから

$$f'(x,y)=0 \quad \Leftrightarrow \quad (x,y)=\left(\frac{1}{5},\frac{2}{5}\right).$$

さて、ここからどうしたら良いか？1変数関数の場合のような増減表は書けない。そもそも f' の符号を調べるというのが、どう多変数関数に拡張したら良いか分からない ($f'(x,y)$ はベクトルなので)。

しかし、すぐ後で証明するように

> f が定義域の内点 a で極大 (or 極小) ならば $f'(a)=0$

は多変数関数でも成立する。

また

> $f'(a)=0$, a の十分近くで $f''>0 \Longrightarrow f$ は a で極小

は多変数関数への拡張が出来る (今回、定理を紹介する)。

8.2 内点 a で極値を取れば $f'(a)=0$

まず極値を定義しないと話が始まりません。

定義 8.1 (極大, 極小) $A\subset\mathbf{R}^n$, $f\colon A\to\mathbf{R}$, $a\in A$ とするとき、f が a で極大とは、

$$\exists\varepsilon>0\quad\text{s.t.}\quad f(a)=\max_{x\in A\cap B(a;\varepsilon)}f(x)=\max_{\substack{x\in A\\ \|x-a\|<\varepsilon}}f(x)$$

が成り立つことをいう。また f が a で狭義の極大とは、

$$\exists\varepsilon>0\quad\text{s.t.}\quad(\forall x\in A:0<\|x-a\|<\varepsilon)\quad f(a)>f(x)$$

が成り立つことをいう。「極小」「狭義の極小」も同様である。

注意 8.2 内点でしか極値を考えないという立場もある。後の条件付き極値問題とのからみで上のような定義を採用した。上の定理は、どちらの流儀でも考えても良い (開集合に属する点は、すべて内点なので)。 ■

定理 8.3 (微分可能な関数が内点で極値を取れば微分は 0) Ω が $\mathbf{R}^n$ の開集合、$f\colon\Omega\to\mathbf{R}$ は全微分可能で、$a\in\Omega$, f は a で極大 (or 極小) ならば、$f'(a)=0$ (これは $\nabla f(a)=0$ とも書ける).

証明 Ω が開集合であるから、$\exists r>0$ s.t. $B(a;r)\subset\Omega$. 各 $i\in\{1,\ldots,n\}$ に対して、

$$\varphi_i\colon(a_i-r,a_i+r)\ni x_i\mapsto f\left(a_1,\ldots,a_{i-1},x_i,a_{i+1},\ldots,a_n\right)\in\mathbf{R}$$

を考えると、これは $x_i = a_i$ で極大値を取る。ゆえに

$$0 = \varphi_i'(a_i) = \frac{\partial f}{\partial x_i}(a) = 0.$$

これが任意の i について成り立つから、

$$f'(a) = \left(\frac{\partial f}{\partial x_1}(a) \cdots \frac{\partial f}{\partial x_n}(a)\right) = (0 \cdots 0) = 0. \blacksquare$$

幾何学的考察　この定理を図形的に考えてみよう。一般に関数 f のグラフ $z = f(\vec{x})$ の $\vec{x} = \vec{a}$ における接超平面は、

$$z = f'(\vec{a})(\vec{x} - \vec{a}) + f(\vec{a})$$

であった。$n = 2$ の場合、

$$z = f'(a, b)\begin{pmatrix} x - a \\ y - b \end{pmatrix} + f(a, b)$$

となるが、$f'(a, b) = 0$ であれば、

$$z = f(a, b).$$

これは xy 平面に水平な平面である。 $\blacksquare$

上の定理の逆は成り立たない。すなわち $f'(a) = 0$ であっても、f が a で極値を取らないということがありうる。これは 1 変数関数でもそうである。(反例: $f(x) = x^3$, $a = 0$ とすると、$f'(a) = 0$ であるが、f は a で極大でも極小でもない。)

8.3　Hesse 行列の符号による極値の判定定理

Taylor の定理を $k = 2$ で用いる。f が C^2 級ならば、十分小さい $\forall h \neq 0$ に対して、

$$\begin{aligned} f(a + h) &= f(a) + \left(d^2 f\right)_a (h) + \frac{1}{2!}\left(d^2 f\right)_{a+\theta h}(h) \\ &= f(a) + f'(a)h + \frac{1}{2}\sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(a + \theta h) h_i h_j. \end{aligned}$$

$f'(a) = 0$ とすると、$f'(a)h = 0$ なので

$$f(a + h) = f(a) + \frac{1}{2}\sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(a + \theta h) h_i h_j.$$

$\|h\|$ が小さいとき、右辺第 2 項は「大体」 h の 2 次式なので (かなり乱暴な議論だけれど)、符号が一定になる場合がある。

$$\exists \varepsilon > 0 \quad (\forall h : 0 < \|h\| < \varepsilon) \quad \sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(a + \theta h) h_i h_j > 0 \quad \Longrightarrow \quad f \text{ は } a \text{ で極小},$$

$$\exists \varepsilon > 0 \quad (\forall h : 0 < \|h\| < \varepsilon) \quad \sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(a + \theta h) h_i h_j < 0 \quad \Longrightarrow \quad f \text{ は } a \text{ で極大}.$$

もちろん以下では厳密な議論を行なう。

定義 8.4 (Hesse 行列) C^2 級の関数 f に対して、

$$H(x) := \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(x) \right)$$

とおき、これを f の x における **Hesse 行列**と呼ぶ。

Hesse 行列は実対称行列である。これを使うと、上の式は

$$f(a+h) = f(a) + f'(a)h + \frac{1}{2}\left(H(a+\theta h)h, h\right) = f(a) + \frac{1}{2}\left(H(a+\theta h)h, h\right)$$

と書ける。ポイントは

$$\left(d^2 f\right)_x (h) = (H(x)h, h),$$

すなわち 2 次の微分が Hesse 行列を係数とする 2 次形式になることである。

次の定理がこの節のメインである。

定理 8.5 (Hesse 行列の符号による極値の判定) Ω が $\mathbf{R}^n$ の開集合、$f\colon \Omega \to \mathbf{R}$ が C^2 級, $a \in \Omega$, $f'(a) = 0$, $H(a) := f$ の a における Hesse 行列 とするとき、次の (i), (ii), (iii) が成り立つ。

(i) $H(a)$ が正値 $\Longrightarrow$ f は a で狭義の極小となる。

(ii) $H(a)$ が負値 $\Longrightarrow$ f は a で狭義の極大となる。

(iii) $H(a)$ が不定符号 $\Longrightarrow$ f は a で極値を取らない。

注意 8.6 正値でも負値でも不定符号でもない場合がある (次項で述べる)。そういう場合は、もっと詳しく調べないと判定できない。 ■

8.4 実対称行列の正値性、負値性

定義 8.7 $A = (a_{ij})$ を n 次実対称行列とする。

(i) A が正値 $\overset{\text{def.}}{\Leftrightarrow}$ A の固有値がすべて正。

(ii) A が負値 $\overset{\text{def.}}{\Leftrightarrow}$ A の固有値がすべて負。

(iii) A が不定符号 $\overset{\text{def.}}{\Leftrightarrow}$ A の固有値に正のもの、負のものがある。

例 8.8 $A = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$ のとき、固有値は $2, 3$ で、すべて正であるから、A は正値。

$A = \begin{pmatrix} -1 & 0 \\ 0 & -2 \end{pmatrix}$ のとき、固有値は $-1, -2$ で、すべて負であるから、A は負値。

$A = \begin{pmatrix} 5 & 0 \\ 0 & -2 \end{pmatrix}$ のとき、固有値は $5, -2$ で、正のもの、負のもの両方あるので、A は不定符号。

$A = \begin{pmatrix} 3 & 0 \\ 0 & 0 \end{pmatrix}$ のとき、固有値は $3, 0$ で、すべて正でもないし、すべて負でもないし、不定符号でもない (負の固有値がない) ので、A は正値でも、負値でも、不定符号でもない。■

8.5　2次形式の符号＝係数行列の符号

定理 8.9 (線形代数の復習: 実対称行列の符号と 2 次形式の符号) $A = (a_{ij})$ が n 次実対称行列とするとき、次の (i), (ii), (iii) が成り立つ。

(i) A が正値 $\Leftrightarrow \forall h \in \mathbf{R}^n \setminus \{0\} \quad (Ah, h) > 0.$

(ii) A が負値 $\Leftrightarrow \forall h \in \mathbf{R}^n \setminus \{0\} \quad (Ah, h) < 0.$

(iii) A が不定符号 $\Longrightarrow \exists h, h' \in \mathbf{R}^n$ s.t. $(Ah, h) > 0, (Ah', h') < 0.$

(証明のあらすじ) 適当な実直交行列 U が存在して、

$$U^T A U = \begin{pmatrix} \lambda_1 & & \text{\Large 0} \\ & \ddots & \\ \text{\Large 0} & & \lambda_n \end{pmatrix}.$$

このとき $x = Uy$ とおくと ($y = U^T x$ とおくと)、

$$(Ax, x) = (AUy, Uy) = \left(U^T AUy, y\right) = \left(\begin{pmatrix} \lambda_1 & & \text{\Large 0} \\ & \ddots & \\ \text{\Large 0} & & \lambda_n \end{pmatrix} y, y\right) = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2.$$

後は $x = 0 \Leftrightarrow y = 0$ に注意して符号を考えれば良い。■

8.6　実対称行列の符号の判定 (オーソドックス版)

与えられた実対称行列が正値か、負値か、不定符号か、あるいはそのいずれでもないか、判定する必要が生じることがある。

「そんなのは線形代数の話だから、微積分の授業ではカットする」と言いたいのをぐっとこらえて。

実対称行列 A が正値であるとは、定義によれば A の固有値がすべて 0 であることだから、A の特性方程式 $\det(\lambda I - A) = 0$ の解がすべて正であることと同値である。残念ながら特性方程式は解くのが大変な場合が多く、3 次方程式の場合で既に難しい。

定理 8.10 (首座小行列式の符号による正値性、負値性の判定) n 次実対称行列 $A = (a_{ij})$ に対して、A_k を A の k 次首座行列とするとき、以下の (1), (2) が成り立つ。

(1) A が正値 $\Longleftrightarrow \forall k \in \{1, \cdots, n\} \quad \det A_k > 0.$

(2) A が負値 $\Longleftrightarrow \forall k \in \{1, \cdots, n\} \quad (-1)^k \det A_k > 0.$ ($\det A_1$, $\det A_2$, $\ldots$, $\det A_n$ の符号が交互に負, 正, 負, 正,$\ldots$)

補題 8.11 対称行列の逆行列は対称行列である:

$$A^T = A, \quad \det A \neq 0 \quad \Rightarrow \quad (A^{-1})^T = A^{-1}.$$

証明 A が対称かつ正則であるとする。$AA^{-1} = A^{-1}A = I$ の転置を取ると、$(A^{-1})^T A^T = A^T (A^{-1})^T = I^T$. A が対称であることから、$(A^{-1})^T A = A(A^{-1})^T = I$. これから、$(A^{-1})^T = A^{-1}$. ■

定理 8.10 の証明 以下の証明は齋藤 [1] pp. 156–157 を書き直したものである。

(必要性) A は正値であると仮定する。$k \in \{1, \cdots, n\}$, $x \in \mathbf{R}^k$ とするとき、$x' := (x, \overbrace{0, \cdots, 0}^{n-k\text{ 個}}) \in \mathbf{R}^n$ とおくと、$(A_k x, x) = (Ax', x')$. $x \neq 0$ とすると、$x' \neq 0$ であるから、A の正値性の仮定から

$$(A_k x, x) = (Ax', x') > 0.$$

ゆえに A_k は正値であり、その固有値 $\lambda_j^{(k)}$ $(j = 1, \cdots, k)$ はすべて正であるから、

$$\det A_k = \prod_{j=1}^{k} \lambda_j^{(k)} > 0.$$

(十分性) 「任意の n 次実対称行列 A に対して、$\det A_k > 0$ $(k = 1, \cdots, n)$ ならば A は正値である」を、n に関する数学的帰納法によって証明する。$n = 1$ のとき、$A = a_{11}$ について $\det A_1 = a_{11}$ であるから、$\det A_1 > 0$ であれば A は正値である。$n - 1$ のときに成り立つと仮定する。A は n 次実対称行列で、$\det A_k > 0$ $(k = 1, \cdots, n)$ を満たすと仮定すると、A_{n-1} は正値である。

$$A = \begin{pmatrix} A_{n-1} & \boldsymbol{b} \\ \boldsymbol{b}^T & c \end{pmatrix}$$

とブロック分けする。

$$P := \begin{pmatrix} I_{n-1} & A_{n-1}^{-1}\boldsymbol{b} \\ \mathbf{0}^T & 1 \end{pmatrix}, \quad B := \begin{pmatrix} A_{n-1} & \mathbf{0} \\ \mathbf{0}^T & c - \boldsymbol{b}^T A_{n-1}^{-1}\boldsymbol{b} \end{pmatrix} \quad (I_{n-1} \text{ は } n-1 \text{ 次単位行列})$$

とおくと、

$$P^T B P = A. \tag{1}$$

(これは良くある変換のようだが、分かりやすい説明はないものか…)

実際

$$\begin{aligned}
P^T B P &= \begin{pmatrix} I_{n-1} & \mathbf{0} \\ (A_{n-1}^{-1}\boldsymbol{b})^T & 1 \end{pmatrix} \begin{pmatrix} A_{n-1} & \mathbf{0} \\ \mathbf{0}^T & c - \boldsymbol{b}^T A_{n-1}^{-1}\boldsymbol{b} \end{pmatrix} \begin{pmatrix} I_{n-1} & A_{n-1}^{-1}\boldsymbol{b} \\ \mathbf{0}^T & 1 \end{pmatrix} \\
&= \begin{pmatrix} I_{n-1} & \mathbf{0} \\ \boldsymbol{b}^T A_{n-1}^{-1} & 1 \end{pmatrix} \begin{pmatrix} A_{n-1} & \mathbf{0} \\ \mathbf{0}^T & c - \boldsymbol{b}^T A_{n-1}^{-1}\boldsymbol{b} \end{pmatrix} \begin{pmatrix} I_{n-1} & A_{n-1}^{-1}\boldsymbol{b} \\ \mathbf{0}^T & 1 \end{pmatrix} \\
&= \begin{pmatrix} A_{n-1} & \mathbf{0} \\ \boldsymbol{b}^T & c - \boldsymbol{b}^T A_{n-1}^{-1}\boldsymbol{b} \end{pmatrix} \begin{pmatrix} I_{n-1} & A_{n-1}^{-1}\boldsymbol{b} \\ \mathbf{0}^T & 1 \end{pmatrix} \\
&= \begin{pmatrix} A_{n-1} & \boldsymbol{b} \\ \boldsymbol{b}^T & \boldsymbol{b}^T A_{n-1}^{-1}\boldsymbol{b} + c - \boldsymbol{b}^T A_{n-1}^{-1}\boldsymbol{b} \end{pmatrix} \\
&= \begin{pmatrix} A_{n-1} & \boldsymbol{b} \\ \boldsymbol{b}^T & c \end{pmatrix} = A.
\end{aligned}$$

A が正値であることを示すには、B が正値であることを示せば十分である。
$\det P = 1$ に注意すると、(1) から

$$\det A = \det A_{n-1}\left(c - \boldsymbol{b}^T A_{n-1}^{-1}\boldsymbol{b}\right).$$

$\det A > 0$, $\det A_{n-1} > 0$ より $c - \boldsymbol{b}^T A_{n-1}^{-1}\boldsymbol{b} > 0$.
$x = (x_1, \cdots, x_n)^T \in \mathbf{R}^n$ に対して、$x' := (x_1, \cdots, x_{n-1})^T$ とおくと、

$$(Bx, x) = (A_{n-1}x', x') + (c - \boldsymbol{b}^T A_{n-1}^{-1}\boldsymbol{b})x_n^2.$$

$x \neq 0$ より $x' \neq 0$ または $x_n \neq 0$ であることに注意すると、$(Bx, x) > 0$. ゆえに B は正値である。■

- 与えられた行列が対角行列ならば、対角成分が固有値なので、判定は簡単である。

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

  は対角行列で、固有値は 1, 2, 3 でいずれも正なので A は正値。

- 与えられた行列が 2 行 2 列ならば、特性方程式を解いて固有値を求めるのもアリ。

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$$

  の特定多項式は $\det(\lambda I - A) = \lambda^2 - (2+1)\lambda + (2\cdot 1 - 1\cdot 1) = \lambda^2 - 3\lambda + 1$. 固有値は $\lambda = \dfrac{3 \pm \sqrt{3^2 - 4\cdot 1\cdot 1}}{2} = \dfrac{3 \pm \sqrt{5}}{2} > 0$ であるから、いずれも正である。ゆえに A は正値である。

- 与えられた行列の首座小行列式 $\det A_k$ の符号を調べる。すべて正ならば A は正値、$\det A_k$ $(k = 1, 2, \ldots, n)$ が順に $- + - + \cdots$ であれば A は負値。例えば

$$A = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}$$

  の首座小行列式は

$$\begin{aligned} &\det A_1 = 2 > 0, \quad \det A_2 = 2\cdot 2 - 1\cdot 1 = 3 > 0, \\ &\det A_3 = 2\cdot 2\cdot 2 + 1\cdot 1\cdot 1 + 1\cdot 1\cdot 1 - 1\cdot 2\cdot 1 - 2\cdot 1\cdot 1 - 2\cdot 1\cdot 1 \\ &\qquad = 8 + 1 + 1 - 2 - 2 - 2 = 4 > 0 \end{aligned}$$

  であるから正値である。

- A が 2 次実対称行列で $\det A < 0$ ならば不定符号であり、$\det A = 0$ ならば正値でも負値でも不定符号でもない。
  (A の固有値を λ_1, λ_2 とするとき、$\lambda_1\lambda_2 = \det A$ であるから。)

問 11　次の行列が正値であるか、負値であるか、不定符号であるか、そのどれでもないか、判定せよ。

(1) $\begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}$　(2) $\begin{pmatrix} -2 & 0 \\ 0 & -1 \end{pmatrix}$　(3) $\begin{pmatrix} 2 & 0 \\ 0 & -1 \end{pmatrix}$　(4) $\begin{pmatrix} 4 & 3 \\ 3 & 1 \end{pmatrix}$

(5) $\begin{pmatrix} 4 & 3 \\ 3 & 4 \end{pmatrix}$　(6) $\begin{pmatrix} -4 & -3 \\ -3 & -4 \end{pmatrix}$　(7) $\begin{pmatrix} 4 & 2 \\ 2 & 1 \end{pmatrix}$　(8) $\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$

(9) $\begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{pmatrix}$　(10) $\begin{pmatrix} -2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -3 \end{pmatrix}$　(11) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 0 \end{pmatrix}$　(12) $\begin{pmatrix} 4 & 3 & 0 \\ 3 & 4 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

(13) $\begin{pmatrix} -4 & 3 & 0 \\ 3 & -4 & 0 \\ 0 & 0 & -2 \end{pmatrix}$　(14) $\begin{pmatrix} 4 & 1 & 0 \\ 1 & 4 & 1 \\ 0 & 1 & 4 \end{pmatrix}$　(15) $\begin{pmatrix} 1 & 2 & 3 \\ 2 & 2 & 0 \\ 3 & 0 & 3 \end{pmatrix}$　(固有値計算しづらい)

8.7　Hesse 行列の符号による極値の判定定理の証明

定理 8.12 (Hesse 行列の符号による極値の判定定理) Ω が $\mathbf{R}^n$ の開集合、$f\colon \Omega \to \mathbf{R}$ が C^2 級, $a \in \Omega$, $f'(a) = 0$, $H(a) := f$ の a における Hesse 行列 とするとき、次の (i), (ii), (iii) が成り立つ。

(i) $H(a)$ が正値 $\Longrightarrow$ f は a で狭義の極小となる。

(ii) $H(a)$ が負値 $\Longrightarrow$ f は a で狭義の極大となる。

(iii) $H(a)$ が不定符号 $\Longrightarrow$ f は a で極値を取らない。

証明　Ω が開集合であるから、$\exists\varepsilon > 0$ s.t. $B(a;\varepsilon) \subset \Omega$. Taylor の定理から、

$$\begin{aligned}(\forall h \in \mathbf{R}^2 : \|h\| < \varepsilon)(\exists\theta \in (0,1)) \quad f(a+h) &= f(a) + f'(a)h + \frac{1}{2}\left(H(a+\theta h)h, h\right) \\ &= f(a) + \frac{1}{2}\left(H(a+\theta h)h, h\right).\end{aligned}$$

(i) $H(a)$ が正値ならば、$\forall k \in \{1, \cdots, n\}$ $\det H_k(a) > 0$. f が C^2 級で、$H(x)$ は連続だから、$\exists\varepsilon' \in (0,\varepsilon)$, $(\forall h \in \mathbf{R}^n\colon \|h\| < \varepsilon')(\forall\theta \in (0,1))$ $\det H_k(a+\theta h) > 0$. このとき $H(a+\theta h)$ は正値である。ゆえに $h \neq 0$ ならば $(H(a+\theta h)h, h) > 0$ であるから、$f(a+h) > f(a)$. ゆえに f は a で狭義の極小である。

(ii) (i) と同様に証明できるので省略する。

(iii) $H(a)$ が不定符号とする。$\exists\lambda, \mu$: $H(a)$ の固有値で、$\lambda > 0$ かつ $\mu < 0$. $\exists u, v \in \mathbf{R}^n \setminus \{0\}$ s.t. $H(a)u = \lambda u$, $H(a)v = \mu v$. $0 < \|u\| < \varepsilon$, $0 < \|v\| < \varepsilon$ として良い (固有ベクトルは実数倍しても固有ベクトルであるから)。

$$g(t) := f(a+tu), \quad h(t) := f(a+tv)$$

とおくと、

$$\begin{aligned} g'(t) &= f'(a+tu)u, \quad g''(t) = \left(H(a+tu)u, u\right), \\ h'(t) &= f'(a+tv)v, \quad h''(t) = \left(H(a+tv)v, v\right), \\ g'(0) &= f'(a)u = 0, \quad g''(0) = (H(a)u, u) = \lambda(u,u) > 0, \\ h'(0) &= f'(a)v = 0, \quad h''(0) = (H(a)v, v) = \mu(v,v) < 0 \end{aligned}$$

である。$g(t)$ ($|t| < 1$) は $t = 0$ で狭義の極小、$h(t)$ ($|t| < 1$) は $t = 0$ で狭義の極大である。ゆえに f は a で極値を取らない (変化する方向によって極小となったり極大となったりする峠点 (鞍点, saddle point) であるから)。■

8.8 典型例

例題 8.1 $f(x, y) = x^3 + y^3 - 3xy$ の極値を求めよ。

解 まず導関数を計算しよう:

$$\nabla f(x) = \begin{pmatrix} f_x \\ f_y \end{pmatrix} = \begin{pmatrix} 3x^2 - y \\ 3y^2 - x \end{pmatrix},$$

$$f \text{ の Hesse 行列 } H = \begin{pmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{pmatrix} = \begin{pmatrix} 6x & -3 \\ -3 & 6y \end{pmatrix}.$$

f の停留点を求める。

$$\begin{aligned} \nabla f(x, y) = 0 \quad &\Leftrightarrow \quad \begin{cases} 3x^2 - 3y = 0 \\ 3y^2 - 3x = 0 \end{cases} \Leftrightarrow \begin{cases} x^2 - y = 0 \\ y^2 - x = 0 \end{cases} \\ &\Leftrightarrow \quad \begin{cases} y = x^2 \\ x^4 - x = 0 \end{cases} \\ &\Leftrightarrow \quad x = y = 0 \quad \text{or} \quad x = y = 1. \end{aligned}$$

(1) 点 $\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ においては

$$H = \begin{pmatrix} 0 & -3 \\ -3 & 0 \end{pmatrix}, \quad \therefore \det H = -9 < 0.$$

H は不定符号であるから、この点は極値点ではない。

(2) 点 $\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ においては

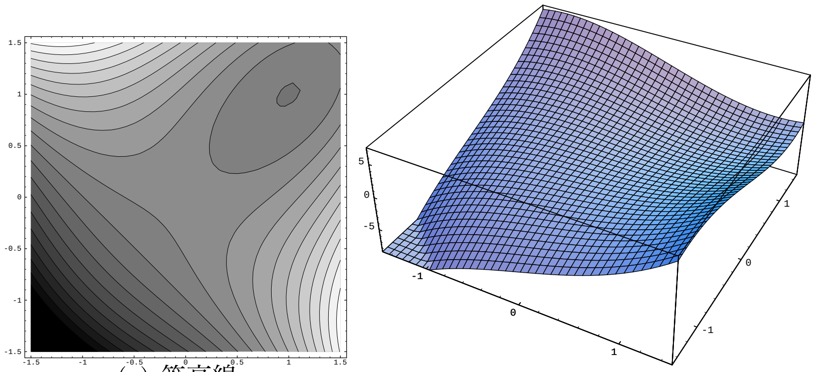
$$H = \begin{pmatrix} 6 & -3 \\ -3 & 6 \end{pmatrix}, \quad \therefore \det D_1 = 6 > 0, \quad \det D_2 = \det H = 6 \cdot 6 - (-3)(-3) = 27 > 0.$$

H は正値であるから、この点は極小点である。値は

$$f(1, 1) = 1^3 + 1^3 - 3 \cdot 1 \cdot 1 = -1.$$

以上をまとめると、$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ のとき極小値 -1. ■

図 1 は $[-1.5, 1.5] \times [-1.5, 1.5]$ で関数の様子を調べたものである。



(a) 等高線　　(b) 鳥瞰図(ちょうかんず)

図 1: $f(x,y) = x^3 + y^3 - 3xy$.

8.9 実対称行列の符号の判定 — 少し真剣にアルゴリズムの追求

8.9.1 固有値計算作戦

固有多項式 $\varphi(\lambda) := \det(\lambda I - A)$ を計算して、その根を求める。$n = 2$ のときは 2 次方程式の解の公式で計算可能である。実際、$A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$ のとき、

$$\det(\lambda I - A) = \lambda^2 - t\lambda + D, \quad t := a + b, \quad D := ad - bc$$

であるから、

$$\lambda = \frac{t \pm \sqrt{t^2 - 4D}}{2}.$$

しかし $n \geq 3$ になると困難が生じる。$n = 3$ であっても、いわゆる不還元の場合には、解が虚数の 3 乗根を含む形で表されることになる (もちろん、微積分を使えば何とか処理できるが)。

8.9.2 首座小行列式作戦

$k = 1, 2, \cdots$ に対して、$\det A_k$ (A_k は A の k 次首座小行列) を計算して符号を調べる。

(i) すべて正である ($\forall k \in \{1, \cdots, n\}$ $\det A_k > 0$) ことは、A が正値であるための必要十分条件である。

(ii) 負から始まり、負と正が交互に現れる ($\forall k \in \{1, \cdots, n\}$ $(-1)^k \det A_k > 0$) ことは、A が負値であるための必要十分条件である。

(iii) 上の (i), (ii) のいずれでもない場合、$\det A$ を計算する。もし $\det A \neq 0$ であるならば、A は不定符号である。$\det A = 0$ のときは、一般には面倒だが、

  (a) $n = 2$ の場合は、正値、負値、不定符号のいずれでもないと結論できる ($\det A < 0$ $\iff$ A は不定符号)。

  (b) また $n = 3$ の場合は、固有多項式が容易に因数分解可能で、符号の判定は容易である。結論だけ書いておくと

$$A = \begin{pmatrix} a & b & c \\ b & d & e \\ c & e & f \end{pmatrix}$$

の固有多項式は

$$\det(\lambda I - A) = \lambda^3 - (a+d+f)\lambda^2 + (-b^2 - c^2 - e - 2 + ad + af + df)\lambda - \det A$$

であるから、$\det A = 0$ を満たすとき、

$$\lambda_1\lambda_2 = -b^2 - c^2 - e - 2 + ad + af + df$$

が負ならば不定符号、そうでないならば正値でも負値でも不定符号でもない。

(c) $n \geq 4$ の場合は研究課題であろう。

8.9.3 Gauss の消去法作戦

A の対角線から下を掃き出す。A が正値であれば、対角線に非零要素は現れず、行交換なしに最後まで上三角化が進められて、対角線に正数が並ぶはずである。一方、A が負値であれば、対角線に非零要素は現れず、行交換なしに最後まで上三角化が進められて、対角線に負数が並ぶはずである。そのいずれでもない場合、必要ならば行交換を施して計算を進めて A の行列式を計算する (行交換を全部で r 回した場合、最終的には対角成分の積 $\times(-1)^r = \det A$ である)。$\det A \neq 0$ ならば、A は 0 を固有値に持たず、正値でも負値でもないので、A は実は不定符号であることが分かる。$\det A = 0$ の場合は少々難しいが、シフトしてみるなどして (つまり A の代りに、$A + \sigma I$ (σ は適当に選ぶ実数) を調べる)、「何とかなる」場合が多いであろう。

$A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 2 & 0 \\ 3 & 0 & 3 \end{pmatrix}$ に対する Gauss の消去法は、行交換なしに

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 2 & 0 \\ 3 & 0 & 3 \end{pmatrix} \to \begin{pmatrix} 1 & 2 & 3 \\ 0 & -2 & -6 \\ 0 & -6 & -3 \end{pmatrix} \to \begin{pmatrix} 1 & 2 & 3 \\ 0 & -2 & -6 \\ 0 & 0 & 15 \end{pmatrix}$$

となるので、いわゆる符号は $(2,1)$ (正の固有値は 2 個、負の固有値が 1 個) で、不定符号である。

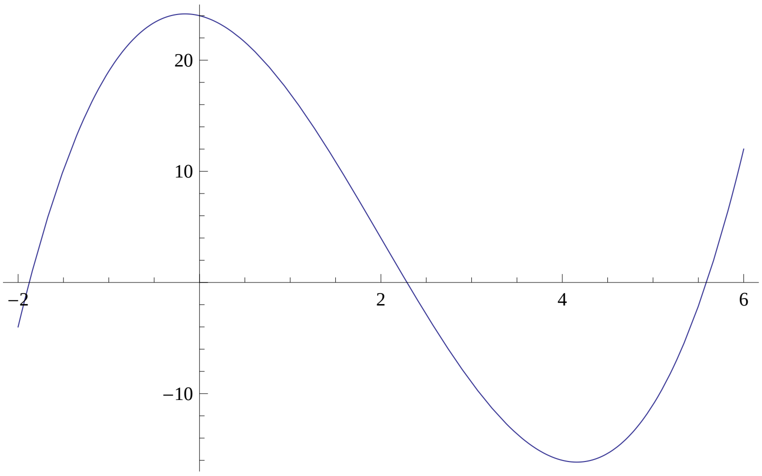
$p(\lambda) := \det(\lambda I - A)$ の根は

$$\begin{aligned}
\lambda = &\frac{1}{3}\left(6 + \frac{7\ 6^{2/3}}{\sqrt[3]{-9 + i\sqrt{1977}}} + \sqrt[3]{6i\left(\sqrt{1977} + 9i\right)}\right), \\
&2 - \frac{\left(1 + i\sqrt{3}\right)\sqrt[3]{-9 + i\sqrt{1977}}}{6^{2/3}} + \frac{-7 + 7i\sqrt{3}}{\sqrt[3]{6i\left(\sqrt{1977} + 9i\right)}}, \\
&2 + \frac{i\left(\sqrt{3} + i\right)\sqrt[3]{-9 + i\sqrt{1977}}}{6^{2/3}} + \frac{-7 - 7i\sqrt{3}}{\sqrt[3]{6i\left(\sqrt{1977} + 9i\right)}}.
\end{aligned}$$

分かりづらいけれど、これは (もちろん) いずれも実数で

$$\lambda \fallingdotseq 5.580664\cdots, 2.2874\cdots, -1.877074\cdots.$$

p のグラフは次のようになる。



8.9.4 処世術

Gauss の消去法で行列の符号が計算できることは、知らない人も結構いると思われる。ペーパーテストで Gauss の消去法で符号を調べるのは少し危険があるかもしれない。

8.10 多変数関数の最大最小問題やりかけの問題

極値問題の最初に、

(2) $$f(x,y) := 3x^2 + 2y^2 + 2xy - 2x - 2y + 1$$

の、

(3) $$K := \{(x,y) \in \mathbf{R}^2; x \geq 0,\ y \geq 0,\ x + y \leq 1\}$$

における最大値、最小値を求めよ、という問題を提示した。

f の $\mathbf{R}^2$ での極値を求めてみよう。$\nabla f(x,y) = 0 \iff (x,y) = \left(\frac{1}{5}, \frac{2}{5}\right)$. この点で Hesse 行列は正値であるので、$f$ は狭義の極小になることが分かる。極小値 $f\left(\frac{1}{5}, \frac{2}{5}\right) = \frac{2}{5}$.

定理 8.13 (Weierstrass の最大値定理) コンパクト集合 K で定義された連続関数 $f\colon K \to \mathbf{R}$ は、必ず最大値と最小値を持つ。

系 8.14 K が $\mathbf{R}^n$ の空でない有界閉集合とするとき、連続関数 $f\colon K \to \mathbf{R}$ は必ず最大値と最小値を持つ。

(3) の K は明らかに閉集合である (多項式関数 (それは連続!) と $\geq, \leq$ で定義されている)。「有界」については定義を思い出そう。

定義 8.15 (有界集合) $A \subset \mathbf{R}^n$ とする。A が有界であるとは、

$$\exists R \in \mathbf{R} \quad \text{s.t.} \quad \forall x \in A \quad \|x\| \leq R$$

が成り立つことをいう。

A が有界であるとは、要するに、原点を中心とする十分大きな半径 R の閉球 $\overline{B}(0;R)$ に A が含まれることである。

(3) の K は、

$$K \subset \overline{B}(0;1)$$

を満たすので、有界性の条件が $R=1$ として成り立つ。ゆえに (3) の K は有界である。

従って Weierstrass の最大値定理によって、(2) で定義された f は K で最大値、最小値を持つ。

K を内部

$$K^\circ = \{(x,y) \in \mathbf{R}^2; x>0,\ y>0,\ x+y<1\}$$

と境界

$$\begin{aligned}
&K^b = \overline{K} \setminus K^\circ = E_1 \cup E_2 \cup E_2 = (\text{三角形の周}),\\
&E_1 := \{(x,0); x \in [0,1]\}, \quad E_2 := \{(0,y); y \in [0,1]\}, \quad E_2 := \{(x,1-x); x \in [0,1]\}
\end{aligned}$$

に分けて最大値、最小値を考える。

$$\max_{K^b} f = \max\left\{\max_{E_1} f, \max_{E_2} f, \max_{E_3} f\right\}$$

であるが、

$$\begin{aligned}
\max_{E_1} f &= \max_{x\in[0,1]} f(x,0) = 2 \quad (= f(1,0)),\\
\max_{E_2} f &= \max_{y\in[0,1]} f(0,y) = 1 \quad (= f(0,0) = f(0,1)),\\
\max_{E_3} f &= \max_{x\in[0,1]} f(x,1-x) = 2 \quad (= f(1,0))
\end{aligned}$$

であるから、

$$\max_{K^b} f = 2 \quad (= f(1,0)).$$

一方、

$$\min_{K^b} f = \min\{\min_{E_1} f, \min_{E_2} f, \min_{E_3} f\}$$

において

$$\begin{aligned}
\min_{E_1} f &= \min_{x\in[0,1]} f(x,0) = \frac{2}{3} \quad (= f\left(\frac{1}{3}, 0\right)),\\
\min_{E_2} f &= \min_{y\in[0,1]} f(0,y) = \frac{1}{2} \quad (= f\left(0, \frac{1}{2}\right),\\
\min_{E_3} f &= \min_{x\in[0,1]} f(x,1-x) = \frac{2}{3} \quad (= f\left(\frac{1}{3}, \frac{2}{3}\right))
\end{aligned}$$

であるから、

$$\min_{K^b} f = \frac{1}{2} \quad \left(f(\left(0, \frac{1}{2}\right))\right).$$

f が $a \in K$ で最大になったとする ($f(a) = \max_K f$)。$a \notin K^\circ$ である。実際、$a \in K^\circ$ であれば、$\nabla f(a) = 0$, f は a で極大となるはずであるが、K° で $\nabla f(x) = 0$ となる点は $x = \left(\frac{1}{5}, \frac{2}{5}\right)$ のみで、そこで f は狭義の極小となるので、極大とはなりえない。ゆえに $a \in K^b$. ゆえに

$$\max_K f = \max_{K^b} f = 2 \quad (= f(1,0)).$$

一方、$f\left(\dfrac{1}{5},\dfrac{2}{5}\right)=\dfrac{2}{5}<\dfrac{1}{2}=\max_{K^b} f$ であるから、f は K^b で最小値を取らず、K° で最小値を取る。それは極小値であるから、$f\left(\dfrac{1}{5},\dfrac{2}{5}\right)$ でなければならない。

まとめると、f は $\left(\dfrac{1}{5},\dfrac{2}{5}\right)$ で最小値 $\dfrac{2}{5}$ を取り、$(1,0)$ で最大値 2 を取る。■

A　実対称行列の符号の判定 (Gaussの消去法作戦)

連立1次方程式の解法に、「Gauss の消去法」と呼ばれるものがある。数学の授業では習わないかも知れないが、コンピューター関係の授業ではしばしば紹介される (B 節を見よ)。この方法の前半部分は行列の変形を行うが、それを対称行列の符号の判定に利用できる。

定理 A.1 (Gaussの消去法による正値性、負値性の判定法) n 次実対称行列 $A=(a_{ij})$ に対して、以下の (1), (2) が成り立つ。

(1) A が正値 $\iff$ A が行交換なしの Gauss の消去法で上三角行列に変形され、その対角成分がすべて正となる。

(2) A が負値 $\iff$ A が行交換なしの Gauss の消去法で上三角行列に変形され、その対角成分がすべて負となる。

証明　(この定理の証明だけが欲しい人には、余分なことがたくさん書いてあって分かり辛いので、勧められないが、桂田 [3] にはとりあえず書いてある。要点だけ説明しておくと、

$$\text{Gauss の消去法で現れる } k \text{ 番目のピボット} = \text{上三角行列の対角成分} = \frac{\det A_k}{\det A_{k-1}}$$

(ただし $\det A_0 = 1$ とする) であることが証明できる。) ■

実は、Gauss の消去法は、平方完成と関係がある。実対称行列 A を係数とする2次形式を、(係数) $\times$ (多項式)^2 の和の形に変形したとき、出て来た係数は Gauss の消去法で得られた上三角行列の対角成分となっている。証明はしないが、例を見てもらおう。

例 A.2

$$A=\begin{pmatrix}1&0&2\\0&1&2\\2&2&-1\end{pmatrix}\to\begin{pmatrix}1&0&2\\0&1&2\\0&2&-5\end{pmatrix}\to\begin{pmatrix}1&0&2\\0&1&2\\0&0&-9\end{pmatrix}.$$

正値でも負値でもない。また行列式は 0 でない (-9 に等しい)。ゆえに不定符号である (実際固有値は $\lambda=1,1,-1$ である)。A を係数とする2次形式を平方完成してみよう。

$$\begin{aligned}x^2+y^2-z^2+4xz+4yz &= (x+2z)^2-4z^2+y^2-z^2+4yz\\&=(x+2z)^2+y^2+4yz-5z^2\\&=(x+2z)^2+(y+2z)^2-9z^2.\end{aligned}$$

係数の 1, 1, -9 に注目。

もう一つ、

$$A=\begin{pmatrix}2&1&1\\1&2&1\\1&1&2\end{pmatrix}\to\begin{pmatrix}2&1&1\\0&\frac{3}{2}&\frac{1}{2}\\0&\frac{1}{2}&\frac{3}{2}\end{pmatrix}\to\begin{pmatrix}2&1&1\\0&\frac{3}{2}&\frac{1}{2}\\0&0&\frac{4}{3}.\end{pmatrix}$$

$\det A_1 = 2 > 0$, $\det A_2 = 2^2 - 1^2 = 3 > 0$, $\det A_3 = \det A = 4 > 0$ であるから、これは正値である (固有値は 4, 1, 1)。2 次形式の平方完成は

$$\begin{aligned}
2x^2+2y^2+2z^2+2xy+2yz+2zx &= 2(x^2+(y+z)x)+2y^2+2z^2+2yz \\
&= 2\left(x+\frac{y+z}{2}\right)^2 - \frac{(y+z)^2}{2} + 2y^2+2z^2-2yz \\
&= 2\left(x+\frac{y+z}{2}\right)^2 + \frac{3y^2}{2} + yz + \frac{3z^2}{2} \\
&= 2\left(x+\frac{y+z}{2}\right)^2 + \frac{3}{2}\left(y^2+\frac{2}{3}yz\right) + \frac{3z^2}{2} \\
&= 2\left(x+\frac{y+z}{2}\right)^2 + \frac{3}{2}\left(y+\frac{z}{3}\right)^2 - \frac{z^2}{6} + \frac{3z^2}{2} \\
&= 2\left(x+\frac{y+z}{2}\right)^2 + \frac{3}{2}\left(y+\frac{z}{3}\right)^2 + \frac{4}{3}z^2.
\end{aligned}$$

係数 2, 3/2, 4/3 が、やはり上の Gauss の消去法の結果の対角成分として現れている。 ■

B　ガウス (Gauss) の消去法のアルゴリズム

連立 1 次方程式の解法として、線形代数の教科書には **クラーメル (Cramer) の公式**や**掃き出し法** (Jordan（ヨルダン）の消去法ともいう) が説明されていることが多いが、ガウスの消去法は、掃き出し法を改良したものである[1]。

例として次の方程式を取りあげて説明しよう。

$$(4)\qquad \begin{cases} 2x_1 + 3x_2 - x_3 = 5 \\ 4x_1 + 4x_2 - 3x_3 = 3 \\ -2x_1 + 3x_2 - x_3 = 1. \end{cases}$$

掃き出し法では係数行列と右辺のベクトルを並べた行列を作り、それに

1. ある行に 0 でない定数をかける。
2. 2 つの行を入れ換える。
3. ある行に別の行の定数倍を加える。

のような操作 — **行に関する基本変形**と呼ぶ — をほどこして、連立方程式の係数行列に相当する部分を単位行列にするのであった。

$$\begin{pmatrix} 2 & 3 & -1 & 5 \\ 4 & 4 & -3 & 3 \\ -2 & 3 & -1 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & \frac{3}{2} & -\frac{1}{2} & \frac{5}{2} \\ 4 & 4 & -3 & 3 \\ -2 & 3 & -1 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & \frac{3}{2} & -\frac{1}{2} & \frac{5}{2} \\ 0 & -2 & -1 & -7 \\ 0 & 6 & -2 & 6 \end{pmatrix} \to$$

$$\to \begin{pmatrix} 1 & \frac{3}{2} & -\frac{1}{2} & \frac{5}{2} \\ 0 & 1 & \frac{1}{2} & \frac{7}{2} \\ 0 & 6 & -2 & 6 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & -\frac{5}{4} & -\frac{11}{4} \\ 0 & 1 & \frac{1}{2} & \frac{7}{2} \\ 0 & 0 & -\frac{5}{2} & -\frac{15}{2} \end{pmatrix} \to \begin{pmatrix} 1 & 0 & -\frac{5}{4} & -\frac{11}{4} \\ 0 & 1 & \frac{1}{2} & \frac{7}{2} \\ 0 & 0 & 1 & 3 \end{pmatrix} \to$$

[1] 見方によっては、ガウスの消去法は中学校で習う加減法 (初めて習う解法！) そのものであり、大学の線形代数で習う解法は、実用性では退化していると言えなくもない (？)。

$$\to \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 3 \end{pmatrix}, \quad \text{ゆえに} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}.$$

ガウスの消去法も、前半の段階はこの方法に似ていて、同様の変形を用いて掃き出しを行なうのだが、以下のように対角線の下側だけを 0 にする。

$$\begin{pmatrix} 2 & 3 & -1 & 5 \\ 4 & 4 & -3 & 3 \\ -2 & 3 & -1 & 1 \end{pmatrix} \to \begin{pmatrix} 2 & 3 & -1 & 5 \\ 0 & -2 & -1 & -7 \\ 0 & 6 & -2 & 6 \end{pmatrix} \to \begin{pmatrix} 2 & 3 & -1 & 5 \\ 0 & -2 & -1 & -7 \\ 0 & 0 & -5 & -15 \end{pmatrix}.$$

最後の行列は

$$2x_1 + 3x_2 - x_3 = 5, \quad -2x_2 - x_3 = -7, \quad -5x_3 = -15$$

ということを表しているので、後の方から順に

$$x_3 = \frac{-15}{-5} = 3, \quad x_2 = \frac{-7 + x_3}{-2} = 2, \quad x_1 = \frac{5 - 3x_2 + x_3}{2} = \frac{5 - 3 \times 2 + 3}{2} = 1$$

と解くことが出来る。前半の対角線の下側を 0 にする掃き出しの操作を**前進消去** (forward elimination)、後半の代入により解の値を求める操作を**後退代入** (backward substitution) と呼ぶ。

参考文献

[1] 齋藤正彦(さいとうまさひこ)：線型代数入門, 東京大学出版会 (1966).

[2] 桂田祐史：多変数の微分積分学 1 講義ノート, `http://www.math.meiji.ac.jp/~mk/lecture/tahensuu1-2013/tahensuu1-2011.pdf`.

[3] 桂田祐史：発展系の数値解析の続き, `http://www.math.meiji.ac.jp/~mk/labo/text/heat-fdm-0-add.pdf` (1997 年〜).