10.2　2次形式と最小（最大）化問題

[問題 10-01] $\boldsymbol{A} = (a_{ij})$ を 2×2 定数行列とする．関数 $g : \mathbb{R}^2 \to \mathbb{R}$ を

$$g(\boldsymbol{x}) = \boldsymbol{x}^T \boldsymbol{A} \boldsymbol{x} \quad \text{for} \quad \boldsymbol{x} = (x_1, x_2)^T \in \mathbb{R}^2$$

と定義する．このとき，

(a) $g(\boldsymbol{x})$ を x_1, x_2 で表せ．

(b) 任意の $\boldsymbol{x} \in \mathbb{R}^2$ に対して，

$$g(\boldsymbol{x}) = \boldsymbol{x}^T \left(\frac{\boldsymbol{A}}{2} + \frac{\boldsymbol{A}^T}{2} \right) \boldsymbol{x}$$

が成立することを示せ．行列 $\boldsymbol{Q} = \left(\frac{\boldsymbol{A}}{2} + \frac{\boldsymbol{A}^T}{2} \right)$ は対称行列になる．

ここでは，$\boldsymbol{Q} = (q_{ij})$ を $n \times n$ 定数行列，$\boldsymbol{x} \in \mathbb{R}^n$ に対して，2次形式

$$g(\boldsymbol{x}) = \boldsymbol{x}^T \boldsymbol{Q} \boldsymbol{x} = \sum_{i=1}^{n} \sum_{j=1}^{n} q_{ij} x_i x_j$$

を扱う．2次形式を考えるときには，定数行列 $\boldsymbol{Q}$ が対称行列である．すなわち，$\boldsymbol{Q}^T = \boldsymbol{Q}$ であることを仮定しても一般性を失わない（[問題 10-01] を参照）．

例：

n 次元空間 $\mathbb{R}^n$ で関数の最小化を考える．$f : \mathbb{R}^n \to \mathbb{R}$ を2階連続微分可能な関数とする．すなわち，

$$\frac{\partial^2 f(\boldsymbol{x})}{\partial x_i \partial x_j} \tag{5}$$

が存在して，(5) は変数 $\boldsymbol{x}$ に関して連続（関数）であるとする．点（ベクトル）$\boldsymbol{x}^*$ で Taylor 展開をすると

$$f(\boldsymbol{x}) = f(\boldsymbol{x}^*) + \boldsymbol{\nabla} f(\boldsymbol{x}^*)^T (\boldsymbol{x} - \boldsymbol{x}^*) + \frac{1}{2} (\boldsymbol{x} - \boldsymbol{x}^*)^T \boldsymbol{H} f(\boldsymbol{x}^*) (\boldsymbol{x} - \boldsymbol{x}^*) + o(\|\boldsymbol{x} - \boldsymbol{x}^*\|^2) \tag{6}$$

として表現できる．ただし，

$$\boldsymbol{\nabla} f(\boldsymbol{x}^*) = \left. \begin{pmatrix} \frac{\partial f(\boldsymbol{x})}{\partial x_1} \\ \frac{\partial f(\boldsymbol{x})}{\partial x_2} \\ \vdots \\ \frac{\partial f(\boldsymbol{x})}{\partial x_n} \end{pmatrix} \right|_{\boldsymbol{x} = \boldsymbol{x}^*}, \qquad \boldsymbol{H} f(\boldsymbol{x}^*) = \left. \left(\frac{\partial f(\boldsymbol{x})}{\partial x_i \partial x_j} \right) \right|_{\boldsymbol{x} = \boldsymbol{x}^*}.$$

である．ここで，

$$\frac{1}{2} (\boldsymbol{x} - \boldsymbol{x}^*)^T \boldsymbol{H} f(\boldsymbol{x}^*) (\boldsymbol{x} - \boldsymbol{x}^*)$$

は2次形式になっている．さらに，$\boldsymbol{x}^*$ が関数 f の極小点あるいは極大点であると仮定すると，

$$\boldsymbol{\nabla} f(\boldsymbol{x}^*) = \boldsymbol{0}$$

が成り立つ．この場合，近似式 (6) は

$$f(\boldsymbol{x}) = f(\boldsymbol{x}^*) + \frac{1}{2} (\boldsymbol{x} - \boldsymbol{x}^*)^T \boldsymbol{H} f(\boldsymbol{x}^*) (\boldsymbol{x} - \boldsymbol{x}^*) + o(\|\boldsymbol{x} - \boldsymbol{x}^*\|^2)$$

となり，誤差を無視した右辺を

$$g(\boldsymbol{x}) := f(\boldsymbol{x}^*) + \frac{1}{2}(\boldsymbol{x}-\boldsymbol{x}^*)^T \boldsymbol{H}f(\boldsymbol{x}^*)(\boldsymbol{x}-\boldsymbol{x}^*)$$

と定義する．このとき，

$$\boldsymbol{x}^* \text{ が } g(\boldsymbol{x}) \text{ の最小点} \quad \Leftrightarrow \quad \boldsymbol{H}f(\boldsymbol{x}^*) \text{ が非負定値（半正定値），すなわち}$$
$$\boldsymbol{u}^T \boldsymbol{H}f(\boldsymbol{x}^*)\boldsymbol{u} \geq 0, \quad \forall \boldsymbol{u} \in \mathbb{R}^n$$

[問題 10-02] 2 階連続微分可能な関数 $f : \mathbb{R}^n \to \mathbb{R}$ に対して，点 $\boldsymbol{x}^* \in \mathbb{R}^n$ での 2 次近似を $g : \mathbb{R}^n \to \mathbb{R}$ とする．すなわち，

$$g(\boldsymbol{x}) = f(\boldsymbol{x}^*) + \boldsymbol{\nabla}f(\boldsymbol{x}^*)^T(\boldsymbol{x}-\boldsymbol{x}^*) + \frac{1}{2}(\boldsymbol{x}-\boldsymbol{x}^*)^T \boldsymbol{H}f(\boldsymbol{x}^*)(\boldsymbol{x}-\boldsymbol{x}^*)$$

で定義する．このとき，

$$\boldsymbol{x}^* \text{ が } g(\boldsymbol{x}) \text{ の最小点} \Leftrightarrow \boldsymbol{\nabla}f(\boldsymbol{x}^*) = \boldsymbol{0} \text{ かつ } \boldsymbol{H}f(\boldsymbol{x}^*) \text{ が非負定値（半正定値）}$$

となることを証明せよ．このことより，

$$\boldsymbol{x}^* \text{ が } f(\boldsymbol{x}) \text{ の極小点} \Rightarrow \boldsymbol{\nabla}f(\boldsymbol{x}^*) = \boldsymbol{0} \text{ かつ } \boldsymbol{H}f(\boldsymbol{x}^*) \text{ が非負定値（半正定値）}$$

が導かれる．

定義１０．１：$\boldsymbol{Q}$ を $n \times n$ 次元の行列とする．

$\boldsymbol{Q}$ は非負定値（半正定値）	$\Leftrightarrow$	$\boldsymbol{u}^T\boldsymbol{Q}\boldsymbol{u} \geq 0,$	$\forall \boldsymbol{u} \in \mathbb{R}^n$
$\boldsymbol{Q}$ は正定値	$\Leftrightarrow$	$\boldsymbol{u}^T\boldsymbol{Q}\boldsymbol{u} > 0,$	$\forall \boldsymbol{u} \neq 0$
$\boldsymbol{Q}$ は非正定値（半負定値）	$\Leftrightarrow$	$\boldsymbol{u}^T\boldsymbol{Q}\boldsymbol{u} \leq 0,$	$\forall \boldsymbol{u} \in \mathbb{R}^n$
$\boldsymbol{Q}$ は負定値	$\Leftrightarrow$	$\boldsymbol{u}^T\boldsymbol{Q}\boldsymbol{u} < 0,$	$\forall \boldsymbol{u} \neq 0$

例１：

$\boldsymbol{Q} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ を 2 次元の単位行列とする．

$$\boldsymbol{x}^T\boldsymbol{Q}\boldsymbol{x} = (x_1\ x_2)\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = x_1^2 + x_2^2 > 0, \qquad \forall \boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \neq \boldsymbol{0}.$$

ゆえに，正定値行列である．

例２：

$\boldsymbol{Q} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$ を 2 次元の対角行列とする．

$$\boldsymbol{x}^T\boldsymbol{Q}\boldsymbol{x} = (x_1\ x_2)\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \lambda_1 x_1^2 + \lambda_2 x_2^2.$$

ゆえに，

$$\lambda_1 \geq 0 \text{ かつ } \lambda_2 \geq 0 \quad \Rightarrow \quad \boldsymbol{Q} \text{ は非負定値行列である．}$$
$$\lambda_1 > 0 \text{ かつ } \lambda_2 > 0 \quad \Rightarrow \quad \boldsymbol{Q} \text{ は正定値行列である．}$$

不等式が逆の場合はそれぞれ非正定値，負定値となる．λ_1, λ_2 は $\boldsymbol{Q}$ の固有値になっている．$n \times n$ の対角行列の場合は

$$\begin{aligned} \boldsymbol{Q} \text{ は非負定値行列（半正定値行列）} &\Leftrightarrow \text{ すべての固有値が非負} \\ \boldsymbol{Q} \text{ は正定値行列} &\Leftrightarrow \text{ すべての固有値が正} \\ \boldsymbol{Q} \text{ は非正定値行列（半負定値行列）} &\Leftrightarrow \text{ すべての固有値が非正} \\ \boldsymbol{Q} \text{ は負定値行列} &\Leftrightarrow \text{ すべての固有値が負} \end{aligned}$$

このことは，一般の対称行列でも成立する．これを,『対角化』を用いて後程証明する．

例３：

$\boldsymbol{Q} = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$ をすべての要素が正の行列とする．

$$\boldsymbol{x}^T \boldsymbol{Q} \boldsymbol{x} = (x_1 \ x_2) \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = x_1^2 + 4x_1 x_2 + x_2^2.$$

$x_1 = x_2 = 1$ とすると，$\boldsymbol{x}^T \boldsymbol{Q} \boldsymbol{x} = 6 > 0$．$x_1 = 1,\ x_2 = -1$ とすると，$\boldsymbol{x}^T \boldsymbol{Q} \boldsymbol{x} = -2 < 0$ となる．したがって，この行列は非負定値でも非正定値でもない．

例４：

$$(x_1 - x_2)^2 = x_1^2 - 2x_1 x_2 + x_2^2 \geq 0, \qquad \forall \boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

であるから，この２次形式は非負定値である．この行列表現は,

$$(x_1 \ x_2) \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

となる．$\boldsymbol{Q} = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$ には負の要素があるが，$\boldsymbol{Q}$ は非負定値行列である．

例３，例４で分かるとおり，要素がすべて非負（正）の行列と非負定値（正定値）行列は異なる．

[問題 10-03] $n \times n$ 対称行列 $\boldsymbol{Q}$ に対して，変数ベクトル $\boldsymbol{x} \in \mathbb{R}^n$ に関する２次形式を $g(\boldsymbol{x}) = \boldsymbol{x}^T \boldsymbol{Q} \boldsymbol{x}$ で定義する．以下の４つの命題について考える．

(a) $\boldsymbol{Q}$ は正定値行列.

(b) $g(\boldsymbol{x})$ は $\bar{\boldsymbol{x}} = \boldsymbol{0}$ で最小値をとる．

(c) $g(\boldsymbol{x})$ は $\bar{\boldsymbol{x}} = \boldsymbol{0}$ で極小値をとる．すなわち，ある $\varepsilon > 0$ が存在し，$\|\boldsymbol{x}\| \leq \varepsilon$ なる任意の $\boldsymbol{x} \in \mathbb{R}^n$ に対して，$g(\bar{\boldsymbol{x}}) \leq g(\boldsymbol{x})$ が成り立つ.

(d) $\boldsymbol{Q}$ は非負定値行列.

このとき,

$$(a) \Rightarrow (b) \Leftrightarrow (c) \Leftrightarrow (d)$$

が成立することを証明せよ.

[問題 10-04] $f : \mathbb{R} \to \mathbb{R}$ を 2 階連続微分可能な関数としたとき，以下について説明せよ．

(a) $\boldsymbol{x}^*$ が f の極小点であるための必要条件．

(b) $\boldsymbol{x}^*$ が f の極小点であるための十分条件．

(c) $\boldsymbol{x}^*$ が f の極大点であるための必要条件．

(d) $\boldsymbol{x}^*$ が f の極大点であるための十分条件．

一般に，2 階連続微分可能な関数 $f : \mathbb{R}^n \to \mathbb{R}$ に対して，

$$
\begin{aligned}
&\boldsymbol{x}^* \text{ が } f \text{ の極小点} \quad \Rightarrow \quad \nabla f(\boldsymbol{x}^*) = \boldsymbol{0},\ \boldsymbol{H} f(\boldsymbol{x}^*) \text{ が非負定値（半正定値）} \\
&\boldsymbol{x}^* \text{ が } f \text{ の極小点} \quad \Leftarrow \quad \nabla f(\boldsymbol{x}^*) = \boldsymbol{0},\ \boldsymbol{H} f(\boldsymbol{x}^*) \text{ が正定値} \\
&\boldsymbol{x}^* \text{ が } f \text{ の極大点} \quad \Rightarrow \quad \nabla f(\boldsymbol{x}^*) = \boldsymbol{0},\ \boldsymbol{H} f(\boldsymbol{x}^*) \text{ が非正定値（半負定値）} \\
&\boldsymbol{x}^* \text{ が } f \text{ の極大点} \quad \Leftarrow \quad \nabla f(\boldsymbol{x}^*) = \boldsymbol{0},\ \boldsymbol{H} f(\boldsymbol{x}^*) \text{ が負定値}
\end{aligned}
$$

10.3 固有値

定義１０．２：$\boldsymbol{A}$ を n 次正方行列とする．

λ は $\boldsymbol{A}$ の 固有値 で，$\boldsymbol{p}$ は λ に対する固有ベクトル である．

$$\Updownarrow$$

$$\boldsymbol{A}\boldsymbol{p} = \lambda \boldsymbol{p}, \quad \boldsymbol{p} \neq \boldsymbol{0} \tag{7}$$

固有値 λ が正の実数のときには，線形写像 $\boldsymbol{A}$ によって同じ方向に移されるベクトルが固有値 λ に対応する固有ベクトル $\boldsymbol{p}$ である（図 21 参照）．固有値が負の実数の場合には，その固有値に対応する固有ベクトルは線形写像 $\boldsymbol{A}$ によって逆の方向に移される．

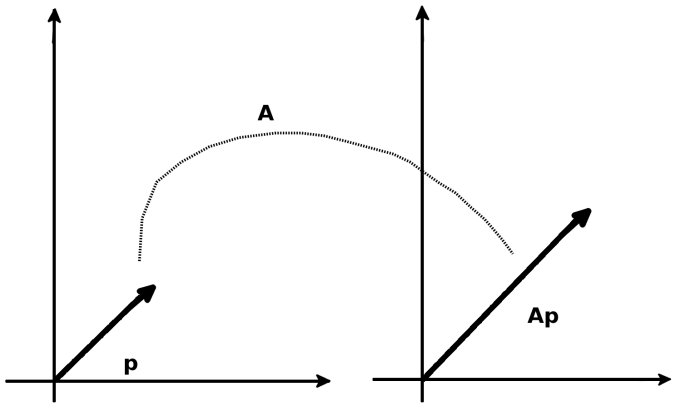

図 21: 固有ベクトル $\boldsymbol{p}$ を線形写像 $\boldsymbol{A}$ によって移した場合．

一般には，λ は複素数，$\boldsymbol{p}$ は複素ベクトルになる．固有値，固有ベクトルの条件 (7) は

$$(\boldsymbol{A}-\lambda\boldsymbol{I})\boldsymbol{p}=\boldsymbol{0}, \qquad \boldsymbol{p}\neq\boldsymbol{0}$$

と書ける．ただし，$\boldsymbol{I}$ は $n\times n$ 単位行列である．したがって，$\boldsymbol{A}-\lambda\boldsymbol{I}$ は正則でない．すなわち，

$$\det(\boldsymbol{A}-\lambda\boldsymbol{I})=0 \qquad \text{（固有方程式）}$$

これは一般的に n 次の代数方程式となり，根の重複度も含め n 個の根をもつ．したがって，固有値は（重複度も含めると）n 個になる．

[問題 10-05]

(a) $\boldsymbol{A}=\begin{pmatrix} 3 & 1 \\ -4 & 4 \end{pmatrix}$ の固有方程式を導き，固有値を計算せよ．

(b) $\boldsymbol{A}=\begin{pmatrix} 2 & -1 \\ 4 & 1 \end{pmatrix}$ の固有方程式を導き，固有値と固有ベクトルを計算せよ．

(c) $\boldsymbol{A}=\begin{pmatrix} 3 & 4 \\ 4 & 2 \end{pmatrix}$ の固有方程式を導き，固有値と固有ベクトルを計算せよ．

定理１０．３：$\lambda_1,\lambda_2,\ldots,\lambda_n$ を $n\times n$ 実対称行列の固有値とする．

(a) $\lambda_1,\lambda_2,\ldots,\lambda_n$ は実数であり，対応する固有ベクトルも実数ベクトルとしてとれる．

(b) 固有値 λ_i に対応する固有ベクトル $\boldsymbol{p}_i$ をお互いに直交し，かつ，その長さを 1 にとることができる．すなわち，

$$(\boldsymbol{p}_i)^T\boldsymbol{p}_j=0 \quad (i\neq j), \qquad \|\boldsymbol{p}_i\|=1. \quad (i=1,2,\ldots,n)$$

したがって，$\boldsymbol{p}_1,\boldsymbol{p}_2,\ldots,\boldsymbol{p}_n$ は線形独立になり，n 次元空間 $\mathbb{R}^n$ の基底をなす．

証明（略）：

(a) λ を $\boldsymbol{A}$ の固有値とする．すなわち，

$$\boldsymbol{A}\boldsymbol{x}=\lambda\boldsymbol{x}, \quad \boldsymbol{x}\neq\boldsymbol{0}$$

を仮定する．複素共役を $\bar{}$ で表すとする．$\boldsymbol{A}\boldsymbol{x}=\lambda\boldsymbol{x}$ の複素共役をとると，$\boldsymbol{A}\bar{\boldsymbol{x}}=\bar{\lambda}\bar{\boldsymbol{x}}$ となる．したがって，

$$\lambda\bar{\boldsymbol{x}}^T\boldsymbol{x}=\bar{\boldsymbol{x}}^T(\lambda\boldsymbol{x})=\bar{\boldsymbol{x}}^T\boldsymbol{A}\boldsymbol{x}=\boldsymbol{x}^T\boldsymbol{A}\bar{\boldsymbol{x}}=\boldsymbol{x}^T(\bar{\lambda}\bar{\boldsymbol{x}})=\bar{\lambda}\boldsymbol{x}^T\bar{\boldsymbol{x}}=\bar{\lambda}\bar{\boldsymbol{x}}^T\boldsymbol{x}$$

となる．ここで，$\bar{\boldsymbol{x}}^T\boldsymbol{x}$ はゼロでない実数であるので $\lambda=\bar{\lambda}$ を得る．

さらに，$\boldsymbol{x}=\boldsymbol{a}+i\boldsymbol{b}$，$\boldsymbol{a},\boldsymbol{b}\in\mathbb{R}^n$ とすると，

$$\boldsymbol{A}\boldsymbol{x}=\boldsymbol{A}(\boldsymbol{a}+i\boldsymbol{b})=\lambda(\boldsymbol{a}+i\boldsymbol{b})=\lambda\boldsymbol{x}$$

となるので，$\boldsymbol{A}\boldsymbol{a}=\lambda\boldsymbol{a}$ となる実数の固有ベクトル $\boldsymbol{a}\in\mathbb{R}^n$ が存在することになる．