V-2-6. 最大・最小

V-2-6-1. 最大・最小から二次形式の最大最小へ

二次形式に正の実数が与えられるとベクトル$\boldsymbol{x}$ の矢印の先端の軌跡を描くことが出来ます。このことは、ベクトル$\boldsymbol{x}$の長さが次の式によって制約されるということを意味します。

$$\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x} = d$$

従って、二次形式の形によっては、ベクトルの長さが極値を持つことになります。たとえば、行列$\boldsymbol{A}$が正定置であれば、軌跡は多次元の超楕円になります。したがって、超楕円の最も長い軸に沿って、中心から延ばしたベクトルの長さが最大で、最も短い超楕円の軸にそって伸ばしたベクトルの長さが長さの最小値になります。行列が非正定置の時には、形状が複雑になり、閉じた形にならないので、他の制約がないと最大・最小を決めることはできません。たとえば

$$\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x} = \boldsymbol{1}$$

のときに

$$\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x} = \boldsymbol{X}^T\boldsymbol{\Lambda}\boldsymbol{X}$$

と変形して、

$$\boldsymbol{\Lambda} = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_p \end{pmatrix}$$

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p > 0$$

$$\frac{{X_1}^2}{\left(\frac{1}{\sqrt{\lambda_1}}\right)^2} + \frac{{X_2}^2}{\left(\frac{1}{\sqrt{\lambda_2}}\right)^2} + \cdots + \frac{{X_p}^2}{\left(\frac{1}{\sqrt{\lambda_p}}\right)^2} = 1$$

のように対角化できます。図 59 は 2 次元平面の楕円の対角化を示しています。

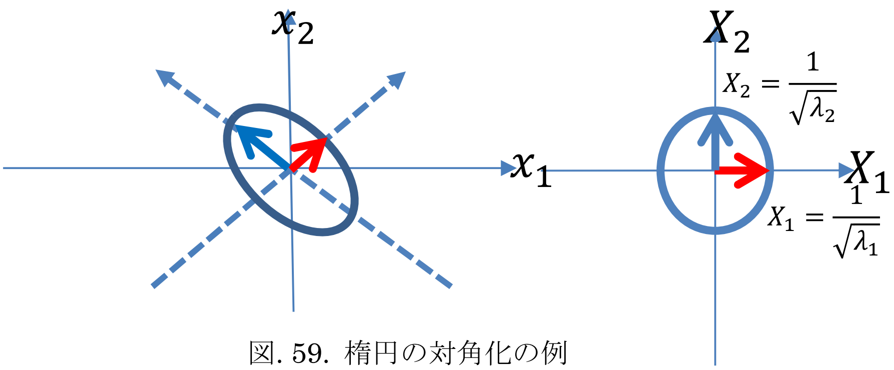

図. 59. 楕円の対角化の例

$$\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x} = \boldsymbol{1}$$

を

$$\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x} = \boldsymbol{X}^T\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}\boldsymbol{X}$$

のように対角化していますが、二次形式にもどすと

$$\frac{{X_1}^2}{\left(\frac{1}{\sqrt{\lambda_1}}\right)^2}+\frac{{X_2}^2}{\left(\frac{1}{\sqrt{\lambda_2}}\right)^2}=1$$

図 59 からもベクトルの絶対値の最大値が $\frac{1}{\sqrt{\lambda_2}}$で、最小値が $\frac{1}{\sqrt{\lambda_1}}$であることは明らかです。

このことから行列$\boldsymbol{A}$が正定置であるとき、次の式の最大値・最小値が固有値の最大値・最小値になることは明らかです。

$$\lambda_{smallest} \le \frac{\boldsymbol{x^T A x}}{\boldsymbol{x^T x}} \le \lambda_{largest}$$

この説明は極めて感覚的な説明で、単純すぎて、納得できる人とできない人とがいるでしょう。もう少し説明を加えます。 $\frac{\boldsymbol{x^T A x}}{\boldsymbol{x^T x}}$ という式の意味は $\boldsymbol{x^T A x}$ と $\boldsymbol{x^T x}$ の比です。分子の$\boldsymbol{x^T A x}$は超楕円の大きさで、$\boldsymbol{x^T x}$ はベクトル xの長さの平方です。私たちが議論しているのは、この比の最大値・最小値とそれが得られる条件です。ここで、行列$\boldsymbol{A}$ は正定値の対称行列です。

$\boldsymbol{A}$ を対象化します。

$$\boldsymbol{P}^{-1}\boldsymbol{AP}=\boldsymbol{\Lambda}$$

$$\boldsymbol{\Lambda}=\begin{pmatrix}\lambda_1 & 0 & \cdots & 0\\ 0 & \lambda_2 & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & \lambda_p\end{pmatrix}$$

$$\lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_p > 0$$

$$\boldsymbol{A}^{\frac{1}{2}}=\boldsymbol{P\Lambda}^{\frac{1}{2}}\boldsymbol{P}^{-1}$$

$\boldsymbol{A}$ は対称行列

$$\boldsymbol{P}^{-1}=\boldsymbol{P}^T$$

$$\boldsymbol{B}^{\frac{1}{2}}=\boldsymbol{P\Lambda}^{\frac{1}{2}}\boldsymbol{P}^T$$

ここで、以下の様にベクトル$\boldsymbol{X}$を定めます。

$$\boldsymbol{X}=\boldsymbol{P}^T\boldsymbol{x}$$

$$\boldsymbol{x}=\boldsymbol{PX}$$

$$\frac{\boldsymbol{x}^T\boldsymbol{Bx}}{\boldsymbol{x}^T\boldsymbol{x}}=\frac{\boldsymbol{x}^T\boldsymbol{Bx}}{(\boldsymbol{PX})^T\boldsymbol{PX}}=\frac{\boldsymbol{x}^T\boldsymbol{B}^{\frac{1}{2}}\boldsymbol{B}^{\frac{1}{2}}\boldsymbol{x}}{\boldsymbol{X}^T\boldsymbol{P}^T\boldsymbol{PX}}=\frac{\boldsymbol{x}^T\boldsymbol{B}^{\frac{1}{2}}\boldsymbol{B}^{\frac{1}{2}}\boldsymbol{x}}{\boldsymbol{X}^T\boldsymbol{X}}=\frac{\boldsymbol{x}^T\boldsymbol{P\Lambda}^{\frac{1}{2}}\boldsymbol{P}^T\boldsymbol{P\Lambda}^{\frac{1}{2}}\boldsymbol{P}^T\boldsymbol{x}}{\boldsymbol{X}^T\boldsymbol{X}}=\frac{\boldsymbol{x}^T\boldsymbol{P\Lambda}^{\frac{1}{2}}\boldsymbol{\Lambda}^{\frac{1}{2}}\boldsymbol{P}^T\boldsymbol{x}}{\boldsymbol{X}^T\boldsymbol{X}}$$

$$=\frac{\boldsymbol{x}^T\boldsymbol{P\Lambda P}^T\boldsymbol{x}}{\boldsymbol{X}^T\boldsymbol{X}}=\frac{\boldsymbol{X}^T\boldsymbol{\Lambda X}}{\boldsymbol{X}^T\boldsymbol{X}}$$

$$\frac{\boldsymbol{X}^T\boldsymbol{\Lambda X}}{\boldsymbol{X}^T\boldsymbol{X}}=\frac{\sum_{\boldsymbol{i=1}}^{\boldsymbol{p}}\lambda_i{X_i}^2}{\sum_{\boldsymbol{i=1}}^{\boldsymbol{p}}{X_i}^2}$$

$\lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_p > 0$

$$\frac{\sum_{i=1}^{p} \lambda_i {y_i}^2}{\sum_{i=1}^{p} {y_i}^2} \le \lambda_1 \frac{\sum_{i=1}^{p} {y_i}^2}{\sum_{i=1}^{p} {y_i}^2} = \lambda_1$$

スペクトル分解します。

$$\boldsymbol{A} = \lambda_1 \boldsymbol{e_1}\boldsymbol{e_1}^T + \lambda_2 \boldsymbol{e_2}\boldsymbol{e_2}^T + \cdots + \lambda_p \boldsymbol{e_p}\boldsymbol{e_p}^T$$

$$\boldsymbol{x}^T\boldsymbol{Ax} = \lambda_1 \boldsymbol{x}^T\boldsymbol{e_1}\boldsymbol{e_1}^T\boldsymbol{x} + \lambda_2 \boldsymbol{x}^T\boldsymbol{e_2}\boldsymbol{e_2}^T\boldsymbol{x} + \cdots + \lambda_p \boldsymbol{x}^T\boldsymbol{e_p}\boldsymbol{e_p}^T\boldsymbol{x}$$

$\boldsymbol{x}$に$\boldsymbol{e_1}$を代入します。

$$\boldsymbol{e}_1{}^T\boldsymbol{Ae} = \lambda_1 \boldsymbol{e}_1{}^T\boldsymbol{e_1}\boldsymbol{e_1}^T\boldsymbol{e}_1 + \lambda_2 \boldsymbol{e}_1{}^T\boldsymbol{e_2}\boldsymbol{e_2}^T\boldsymbol{e}_1 + \cdots + \lambda_p \boldsymbol{e}_1{}^T\boldsymbol{e_p}\boldsymbol{e_p}^T\boldsymbol{e}_1$$

$\boldsymbol{e}_i$ と $\boldsymbol{e}_j$ は直交していますから

$$\boldsymbol{e}_i{}^T\boldsymbol{e}_j = 0 \quad i \ne j$$

$$\boldsymbol{e}_i{}^T\boldsymbol{e}_i = 1$$

$$\boldsymbol{e}_1{}^T\boldsymbol{Ae} = \lambda_1 \boldsymbol{e}_1{}^T\boldsymbol{e_1}\boldsymbol{e_1}^T\boldsymbol{e}_1 = \lambda_1$$

結論

行列$\boldsymbol{A}$ が正定値の対称行列であるとき

$$\max_{x \ne 0} \frac{\boldsymbol{x}^T\boldsymbol{Ax}}{\boldsymbol{x}^T\boldsymbol{x}} = \lambda_1$$

同様に

$$\min_{x \ne 0} \frac{\boldsymbol{x}^T\boldsymbol{Ax}}{\boldsymbol{x}^T\boldsymbol{x}} = \lambda_p$$

$$\boldsymbol{\Lambda} = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_p \end{pmatrix}$$

$$\lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_p > 0$$

式 63

V-2-6-2. コーシー・シュワルツの不等式

今のところ私たちは、上記の結論を何に使うのか知りません。この項の後半でその応用について触れますが、その前に、不等式の意味や解釈を考えるために、 その他の不等式について解説します。

コーシー・シュワルツの不等式は、最適化などによく使われる不等式ですが、しばしば変形した形で登場します。最も一般的な形は以下の形です。

$$\left(a_1b_1 + a_2b_2 + \cdots + a_pb_p\right)^2 \le \left({a_1}^2 + {a_2}^2 + \cdots + {a_p}^2\right)\left({b_1}^2 + {b_2}^2 + \cdots + {b_p}^2\right)$$

式 64

この不等式の、最も簡単で洗練された証明は、この式をベクトルに変形することによって得られます。

次のベクトル$\boldsymbol{a}$ 、$\boldsymbol{b}$があるとします。

$$\boldsymbol{a} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_p \end{pmatrix}, \quad \boldsymbol{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_p \end{pmatrix}$$

２つのベクトルの内積は次の通りです。

$$\boldsymbol{a} \cdot \boldsymbol{b} = |\boldsymbol{a}||\boldsymbol{b}| \cos\theta$$

$$-\frac{\pi}{2} \leq \theta \leq \frac{\pi}{2}$$

$$0 \leq \cos\theta \leq 1$$

次に

$$\boldsymbol{a} \cdot \boldsymbol{b} \leq |\boldsymbol{a}||\boldsymbol{b}| \qquad \text{i}$$

$$\boldsymbol{a} \cdot \boldsymbol{b} = \boldsymbol{a}^{\mathrm{T}}\boldsymbol{b} = (a_1 \quad a_2 \quad \cdots \quad a_p)\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_p \end{pmatrix} = a_1b_1 + a_2b_2 + \cdots + a_pb_p$$

$$|\boldsymbol{a}| = \sqrt{{a_1}^2 + {a_2}^2 + \cdots + {a_p}^2}, \qquad |\boldsymbol{b}| = \sqrt{{b_1}^2 + {b_2}^2 + \cdots + {b_p}^2}$$

式ｉにこれらを代入すると

$$a_1b_1 + a_2b_2 + \cdots + a_pb_p \leq \sqrt{{a_1}^2 + {a_2}^2 + \cdots + {a_p}^2}\sqrt{{b_1}^2 + {b_2}^2 + \cdots + {b_p}^2}$$

２乗して

$$\left(a_1b_1 + a_2b_2 + \cdots + a_pb_p\right)^2 \leq \left({a_1}^2 + {a_2}^2 + \cdots + {a_p}^2\right)\left({b_1}^2 + {b_2}^2 + \cdots + {b_p}^2\right)$$

大変洗練された証明ですが、よく考えると、内積はベクトルの長さの積よりも小さいということを言っているにすぎません。

コーシーシュワルツの不等式の変形の例(1)

$$\boldsymbol{a} = \left(\boldsymbol{B}^{\frac{1}{2}}\boldsymbol{\alpha}\right)$$

$$\boldsymbol{b} = \left(\boldsymbol{B}^{-\frac{1}{2}}\boldsymbol{\beta}\right)$$

$\boldsymbol{B}$ は対称行列

$\boldsymbol{a}$ と $\boldsymbol{b}$ をコーシー・シュワルツの不等式に入れます。

$$(\boldsymbol{a}^T\boldsymbol{a})(\boldsymbol{b}^T\boldsymbol{b}) \geq (\boldsymbol{a}^T\boldsymbol{b})^2$$

$$\text{左辺} = \left(\boldsymbol{B}^{\frac{1}{2}}\boldsymbol{\alpha}\right)^T\left(\boldsymbol{B}^{\frac{1}{2}}\boldsymbol{\alpha}\right)\left(\boldsymbol{B}^{-\frac{1}{2}}\boldsymbol{\beta}\right)^T\left(\boldsymbol{B}^{-\frac{1}{2}}\boldsymbol{\beta}\right)$$

$$= \left(\boldsymbol{\alpha}^T \boldsymbol{B}^{\frac{1}{2}^T} \boldsymbol{B}^{\frac{1}{2}} \boldsymbol{\alpha}\right)\left(\boldsymbol{\beta}' \boldsymbol{B}^{-\frac{1}{2}^{\boldsymbol{T}}} \boldsymbol{B}^{-\frac{1}{2}} \boldsymbol{\beta}\right)$$

$$= (\boldsymbol{\alpha}^T \boldsymbol{B\alpha})(\boldsymbol{\beta}^T \boldsymbol{B}^{-\mathbf{1}} \boldsymbol{\beta})$$

$$\text{右辺} = \left(\boldsymbol{B}^{\frac{1}{2}} \boldsymbol{\alpha}\right)^T \left(\boldsymbol{B}^{-\frac{1}{2}} \boldsymbol{\beta}\right)^T \left(\boldsymbol{B}^{\frac{1}{2}} \boldsymbol{\alpha}\right)\left(\boldsymbol{B}^{-\frac{1}{2}} \boldsymbol{\beta}\right)$$

$$= \left(\boldsymbol{\alpha}^T \boldsymbol{B}^{\frac{1}{2}^T} \boldsymbol{B}^{-\frac{1}{2}} \boldsymbol{\beta}\right)\left(\boldsymbol{\alpha}^T \boldsymbol{B}^{\frac{1}{2}^T} \boldsymbol{B}^{-\frac{1}{2}} \boldsymbol{\beta}\right)$$

$$= (\boldsymbol{\alpha}^T \boldsymbol{I\beta})(\boldsymbol{\alpha}^T \boldsymbol{I\beta})$$

$$= (\boldsymbol{\alpha}^T \boldsymbol{\beta})^2$$

$$(\boldsymbol{\alpha}^T \boldsymbol{B\alpha})(\boldsymbol{\beta}^T \boldsymbol{B}^{-\mathbf{1}} \boldsymbol{\beta}) \geq (\boldsymbol{\alpha}^{\boldsymbol{T}} \boldsymbol{\beta})^2$$

また

$$\boldsymbol{B}^{\frac{1}{2}} \boldsymbol{\alpha} = c\boldsymbol{B}^{-\frac{1}{2}} \boldsymbol{\beta}$$

のとき

$$(\boldsymbol{\alpha}^T \boldsymbol{B\alpha})(\boldsymbol{\beta}^T \boldsymbol{B}^{-\mathbf{1}} \boldsymbol{\beta}) = (\boldsymbol{\alpha}^{\boldsymbol{T}} \boldsymbol{\beta})^2$$

式 65

この不等式で $\boldsymbol{\alpha}^T \boldsymbol{B\alpha}$ がスカラーでBが正定値であるとき、不等号の向きを変えずに$(\boldsymbol{\alpha}^T \boldsymbol{B\alpha})$を右辺に移項できるので、

$$\frac{(\boldsymbol{\alpha}^T \boldsymbol{\beta})^2}{\boldsymbol{\alpha}^T \boldsymbol{B\alpha}} \leq \boldsymbol{\beta}^T \boldsymbol{B}^{-\mathbf{1}} \boldsymbol{\beta}$$

さらに、左辺を$\boldsymbol{\alpha}$、右辺を$\boldsymbol{\beta}$の関数だと考えると

$$\mathbf{F}(\boldsymbol{\alpha}) = \frac{(\boldsymbol{\alpha}^T \boldsymbol{\beta})^2}{\boldsymbol{\alpha}^T \boldsymbol{B\alpha}}$$

なので、

$$\boldsymbol{B}^{\frac{1}{2}} \boldsymbol{\alpha} = c\boldsymbol{B}^{-\frac{1}{2}} \boldsymbol{\beta}$$
$$\boldsymbol{\alpha} = c\boldsymbol{B}^{-1} \boldsymbol{\beta}$$

の時

$$\max_{\boldsymbol{\alpha} \neq 0} \mathbf{F}(\boldsymbol{\alpha}) = \boldsymbol{\beta}^T \boldsymbol{B}^{-\mathbf{1}} \boldsymbol{\beta}$$

もう少しはっきりさせるために$\boldsymbol{\alpha}$ を $\boldsymbol{x}$と書くと、$\boldsymbol{B}$が正定値の対称行列ならば、

$$\boldsymbol{x} = c\boldsymbol{B}^{-1} \boldsymbol{\beta}$$

のとき

$$\max_{x \neq 0} \frac{(\boldsymbol{x}' \boldsymbol{\beta})^2}{\boldsymbol{x}' \boldsymbol{Bx}} = \boldsymbol{\beta}' \boldsymbol{B}^{-\mathbf{1}} \boldsymbol{\beta}$$

コーシー・シュワルツの不等式の変形例(2)

$$\boldsymbol{a} = \left(\boldsymbol{F}^{\frac{1}{2}} \boldsymbol{\alpha}\right)$$

$$\boldsymbol{b} = \left(\boldsymbol{G}^{\frac{1}{2}}\boldsymbol{\alpha}\right)$$

F、G は対称行列

$\boldsymbol{a}$ を $\boldsymbol{b}$ コーシー・シュワルツの不等式に入れます。

$$(\boldsymbol{a}^T\boldsymbol{a})(\boldsymbol{b}^T\boldsymbol{b}) \geq (\boldsymbol{a}^T\boldsymbol{b})^2$$

$$左辺 = \left(\boldsymbol{F}^{\frac{1}{2}}\boldsymbol{\alpha}\right)^T \left(\boldsymbol{F}^{\frac{1}{2}}\boldsymbol{\alpha}\right)\left(\boldsymbol{G}^{-\frac{1}{2}}\boldsymbol{\alpha}\right)^T \left(\boldsymbol{G}^{-\frac{1}{2}}\boldsymbol{\alpha}\right)$$

$$= \left(\boldsymbol{\alpha}^T\boldsymbol{F}^{\frac{1}{2}^T}\boldsymbol{F}^{\frac{1}{2}}\boldsymbol{\alpha}\right)\left(\boldsymbol{\alpha}^T\boldsymbol{G}^{\frac{1}{2}^{\boldsymbol{T}}}\boldsymbol{G}^{\frac{1}{2}}\boldsymbol{\alpha}\right)$$

$$= (\boldsymbol{\alpha}^T\boldsymbol{F}\boldsymbol{\alpha})(\boldsymbol{\alpha}^T\boldsymbol{G}\boldsymbol{\alpha})$$

$$右辺 = \left(\boldsymbol{G}^{\frac{1}{2}}\boldsymbol{\alpha}\right)^T \left(\boldsymbol{F}^{\frac{1}{2}}\boldsymbol{\alpha}\right)$$

$$= \boldsymbol{\alpha}^T\boldsymbol{G}^{\frac{1}{2}^T}\boldsymbol{F}^{\frac{1}{2}}\boldsymbol{\alpha}$$

元の不等式に戻します

$$(\boldsymbol{\alpha}^T\boldsymbol{F}\boldsymbol{\alpha})(\boldsymbol{\alpha}^T\boldsymbol{G}\boldsymbol{\alpha}) \geq \boldsymbol{\alpha}^T\boldsymbol{G}^{\frac{1}{2}^T}\boldsymbol{F}^{\frac{1}{2}}\boldsymbol{\alpha}$$

$(\boldsymbol{\alpha}^T\boldsymbol{G}\boldsymbol{\alpha})^2 > \boldsymbol{0}$ だから、両辺を$(\boldsymbol{\alpha}^T\boldsymbol{G}\boldsymbol{\alpha})^2$で割って

$$\frac{(\boldsymbol{\alpha}^T\boldsymbol{F}\boldsymbol{\alpha})}{(\boldsymbol{\alpha}^T\boldsymbol{G}\boldsymbol{\alpha})} \geq \frac{\boldsymbol{\alpha}^T\boldsymbol{g}^{\frac{1}{2}^T}\boldsymbol{F}^{\frac{1}{2}}\boldsymbol{\alpha}}{(\boldsymbol{\alpha}^T\boldsymbol{G}\boldsymbol{\alpha})^2}$$

等号が成立つ条件は

$$\boldsymbol{a} = c\boldsymbol{b}$$

$$（\ \boldsymbol{a} = \left(\boldsymbol{F}^{\frac{1}{2}}\boldsymbol{\alpha}\right),\quad \boldsymbol{b} = \left(\boldsymbol{G}^{\frac{1}{2}}\boldsymbol{\alpha}\right)）$$

$$\boldsymbol{F}^{\frac{1}{2}}\boldsymbol{\alpha} = c\boldsymbol{G}^{\frac{1}{2}}\boldsymbol{\alpha}$$

したがって

$$\boldsymbol{F}^{\frac{1}{2}}\boldsymbol{\alpha} = c\boldsymbol{G}^{\frac{1}{2}}\boldsymbol{\alpha}$$

の時

$$\frac{(\boldsymbol{\alpha}^T\boldsymbol{F}\boldsymbol{\alpha})}{(\boldsymbol{\alpha}^T\boldsymbol{G}\boldsymbol{\alpha})} \text{は最低値} \frac{(\boldsymbol{\alpha}^T\boldsymbol{F}\boldsymbol{\alpha})}{(\boldsymbol{\alpha}^T\boldsymbol{G}\boldsymbol{\alpha})} = \frac{\boldsymbol{\alpha}^T\boldsymbol{g}^{\frac{1}{2}^T}\boldsymbol{F}^{\frac{1}{2}}\boldsymbol{\alpha}}{(\boldsymbol{\alpha}^T\boldsymbol{G}\boldsymbol{\alpha})^2}$$

$$\frac{(\boldsymbol{\alpha}^T\boldsymbol{G}\boldsymbol{\alpha})}{(\boldsymbol{\alpha}^T\boldsymbol{F}\boldsymbol{\alpha})} \text{は最大値}$$

$$\frac{(\boldsymbol{\alpha}^T\boldsymbol{G}\boldsymbol{\alpha})}{(\boldsymbol{\alpha}^T\boldsymbol{F}\boldsymbol{\alpha})} = \frac{(\boldsymbol{\alpha}^T\boldsymbol{G}\boldsymbol{\alpha})^2}{\boldsymbol{\alpha}^T\boldsymbol{g}^{\frac{1}{2}^T}\boldsymbol{F}^{\frac{1}{2}}\boldsymbol{\alpha}}$$

となります。

V-2-6-3. ラグランジュの未定乗数法

上述のコーシー・シュワルツの不等式の変形例２つは、ラグランジェの未定乗数法の特定の場合だとも考えられます。ラグランジュの未定乗数法は、制約条件が与えられた時の最大最小問題の解法です。図 60 はラグランジュの未定乗数法の理論的な説明図です。

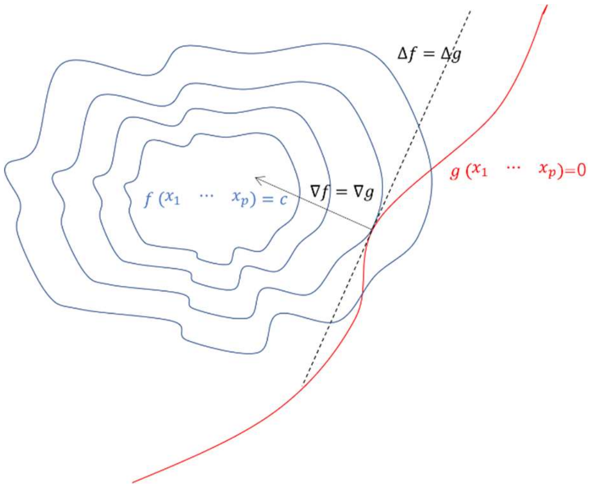

図 60.最小値を求める関数と制約条件の関係

赤い線は制約条件の軌跡です。たとえば$g(x_1 \quad \cdots \quad x_p) = 0$のような制約です。青い線は最低値を求める関数で$f(x_1 \quad \cdots \quad x_p) = c$のような関数で、変数$\boldsymbol{x}$によって変動します。$c$ の増加に伴って、青い線で囲まれた部分が拡大するとします。ある特定の値をcがとるとき青い線と赤い線が接します。それ以上拡大すると赤い線と青い線は交差します。赤い線と青い線が接点あるいは交点を持つ時、関数$f(x_1 \quad \cdots \quad x_p)$は制約条件$g(x_1 \quad \cdots \quad x_p) = 0$を満たしています。$f(x_1 \quad \cdots \quad x_p)$が制約条件$g(x_1 \quad \cdots \quad x_p) = 0.$の内側に囲い込まれていれば、青い線や赤い線がどんなに複雑な形をしていても、最初に接点を持った時が、制約条件下での c の最小値であることは明らかです。そして、最後に接点を持つときが最大値になるはずです。赤いラインと青いラインが接点を持つ時、二つの線は接線、接する平面を共有し($\Delta f = \lambda \Delta g$)、法線ベクトルを共有するはずです($\nabla f = \lambda \nabla g$)。法線ベクトルは、超平面の傾斜を表すものですから、勾配ベクトルと呼ばれることもあります。

$f(x_1 \quad \cdots \quad x_p)$の勾配はベクトルの偏微分によって求めることが出来ます。

$$\nabla f = \frac{\partial f(\boldsymbol{x})}{\partial \boldsymbol{x}} = \begin{pmatrix} \dfrac{\partial f(x_1 \quad \cdots \quad x_p)}{\partial x_1} \\ \vdots \\ \dfrac{\partial f(x_1 \quad \cdots \quad x_p)}{\partial x_p} \end{pmatrix}$$

$$\lambda \nabla g = \lambda \frac{\partial g(\boldsymbol{x})}{\partial \boldsymbol{x}} = \lambda \begin{pmatrix} \dfrac{\partial g(x_1 \quad \cdots \quad x_p)}{\partial x_1} \\ \vdots \\ \dfrac{\partial g(x_1 \quad \cdots \quad x_p)}{\partial x_p} \end{pmatrix}$$

法線ベクトルの共有は次の式であらわせますから

$$\nabla f = \lambda \nabla g$$

$$\nabla f - \lambda \nabla g = 0$$

$$\frac{\partial f(\boldsymbol{x})}{\partial \boldsymbol{x}} - \lambda \frac{\partial g(\boldsymbol{x})}{\partial \boldsymbol{x}} = 0$$

$$\frac{\partial \big(f(\boldsymbol{x}) - \lambda g(\boldsymbol{x})\big)}{\partial \boldsymbol{x}} = 0$$

ここでラグランジュの未定乗数法の関数を以下の式で表します。

$$L(\boldsymbol{x}, \lambda) = f(\boldsymbol{x}) - \lambda g(\boldsymbol{x})$$

λをラグランジュの未定乗数といいます。

ラグランジュの未定乗数法は、次のように表せます。.

$$\frac{\partial L(\boldsymbol{x}, \lambda)}{\partial(\boldsymbol{x}, \lambda)} = \begin{pmatrix} \frac{\partial L(\boldsymbol{x}, \lambda)}{\partial x_1} \\ \vdots \\ \frac{\partial L(\boldsymbol{x}, \lambda)}{\partial x_p} \\ \frac{\partial L(\boldsymbol{x}, \lambda)}{\partial \lambda} \end{pmatrix}$$

ここで

$$L(\boldsymbol{x}, \lambda) = f(\boldsymbol{x}) - \lambda g(\boldsymbol{x})$$

$$\boldsymbol{x} = \begin{pmatrix} \boldsymbol{x_1} \\ \vdots \\ \boldsymbol{x_p} \end{pmatrix}$$

です。

式 66

計算例

1. $x_1{}^2 + x_2{}^2 + x_3{}^2 = 1$ という制約条件で$x_1 + x_2 + x_3$の極値を求めます

   極値を求める関数 $f(x_1 \quad x_2 \quad x_3)$=$x_1 + x_2 + x_3$

   制約条件 $g(x_1 \quad x_2 \quad x_3)$=$x_1{}^2 + x_2{}^2 + x_3{}^2 - 1 = 0$

$$\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

ラグランジェ関数

$$L(x_1 \quad x_2 \quad x_3 \quad \lambda) = x_1 + x_2 + x_3 - \lambda(x_1{}^2 + x_2{}^2 + x_3{}^2 - 1)$$

$$\frac{\partial L(x_1 \quad x_2 \quad x_3 \quad \lambda)}{\partial \boldsymbol{x}}$$

$$\frac{\partial L(x_1 \quad x_2 \quad x_3 \quad \lambda)}{\partial x_1} = 1 - 2\lambda x_1 = 0 \qquad \text{i}$$

$$\frac{\partial L(x_1 \quad x_2 \quad x_3 \quad \lambda)}{\partial x_2} = 1 - 2\lambda x_2 = 0 \qquad \text{ii}$$

$$\frac{\partial L(x_1 \quad x_2 \quad x_3 \quad \lambda)}{\partial x_3} = 1 - 2\lambda x_3 = 0 \qquad \text{iii}$$

$$\frac{\partial L(x_1 \quad x_2 \quad x_3 \quad \lambda)}{\partial \lambda} = -({x_1}^2 + {x_2}^2 + {x_3}^2 - 1) = 0 \qquad \text{iv}$$

I、ii、iii から

$$x_1 = \frac{1}{2\lambda} \qquad \text{i}'$$

$$x_2 = \frac{1}{2\lambda} \qquad \text{ii}'$$

$$x_3 = \frac{1}{2\lambda} \qquad \text{iii}'$$

i’、 ii’、iii’ を iv に代入

$$\left(\frac{1}{2\lambda}\right)^2 + \left(\frac{1}{2\lambda}\right)^2 + \left(\frac{1}{2\lambda}\right)^2 - 1 = 0$$

$$\frac{3}{4\lambda^2} = 1$$

$$\lambda^2 = \frac{3}{4}$$

$$\lambda = \pm\frac{\sqrt{3}}{2}$$

ラグランジュの未定乗数の値を i’、 ii’、iii’に代入

$$x_1 = \pm\frac{2}{2\sqrt{3}} = \pm\frac{1}{\sqrt{3}} = \pm\frac{\sqrt{3}}{3}$$

$$x_2 = \pm\frac{2}{2\sqrt{3}} = \pm\frac{1}{\sqrt{3}} = \pm\frac{\sqrt{3}}{3}$$

$$x_3 = \pm\frac{2}{2\sqrt{3}} = \pm\frac{1}{\sqrt{3}} = \pm\frac{\sqrt{3}}{3}$$

λ	$-\frac{\sqrt{3}}{2}$	$\frac{\sqrt{3}}{2}$
x_1	$-\frac{\sqrt{3}}{3}$	$\frac{\sqrt{3}}{3}$
x_2	$-\frac{\sqrt{3}}{3}$	$\frac{\sqrt{3}}{3}$

$$x_3 \quad -\frac{\sqrt{3}}{3} \quad \frac{\sqrt{3}}{3}$$

$$x_1+x_2+x_3 \quad -\sqrt{3} \quad \sqrt{3}$$

最大値 $\sqrt{3}$、最小値$-\sqrt{3}$

2. ${x_1}^2+{x_2}^2=1$と${x_2}^2+{x_3}^2=1$という制約条件で$x_1+x_2+x_3$の極値を求めます。
   極値を求める関数 $f(x_1 \quad x_2 \quad x_3)$=$x_1+x_2+x_3$
   制約条件 $g_1(x_1 \quad x_2 \quad x_3)$=${x_1}^2+{x_2}^2-1=0$
   $g_2(x_1 \quad x_2 \quad x_3)$=${x_2}^2+{x_3}^2-1=0$

$$\boldsymbol{x}=\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

$$L(x_1 \quad x_2 \quad x_3 \quad \lambda_1 \quad \lambda_2)= x_1+x_2+x_3-\lambda_1({x_1}^2+{x_2}^2-1)-\lambda_2({x_2}^2+{x_3}^2-1)$$

$$\frac{\partial L(x_1 \quad x_2 \quad x_3 \quad \lambda_1 \quad \lambda_2)}{\partial \boldsymbol{x}}$$

$$\frac{\partial L(x_1 \quad x_2 \quad x_3 \quad \lambda_1 \quad \lambda_2)}{\partial x_1}=1-2\lambda_1 x_1=0 \qquad \mathrm{i}$$

$$\frac{\partial L(x_1 \quad x_2 \quad x_3 \quad \lambda_1 \quad \lambda_2)}{\partial x_2}=1-2(\lambda_1+\lambda_2)x_2=0 \qquad \mathrm{ii}$$

$$\frac{\partial L(x_1 \quad x_2 \quad x_3 \quad \lambda_1 \quad \lambda_2)}{\partial x_3}=1-2\lambda_2 x_3=0 \qquad \mathrm{iii}$$

$$\frac{\partial L(x_1 \quad x_2 \quad x_3 \quad \lambda_1 \quad \lambda_2)}{\partial \lambda_1}={x_1}^2+{x_2}^2-1=0 \qquad \mathrm{iv}$$

$$\frac{\partial L(x_1 \quad x_2 \quad x_3 \quad \lambda_1 \quad \lambda_2)}{\partial \lambda_1}={x_2}^2+{x_3}^2-1=0 \qquad \mathrm{v}$$

I、ii、iii から

$$x_1=\frac{1}{2\lambda_1} \qquad \mathrm{i}'$$

$$x_2=\frac{1}{2(\lambda_1+\lambda_2)} \qquad \mathrm{ii}'$$

$$x_3=\frac{1}{2\lambda_2} \qquad \mathrm{iii}'$$

$$g_1(x_1 \quad x_2 \quad x_3)-g_2(x_1 \quad x_2 \quad x_3)=({x_1}^2+{x_2}^2-1)-({x_2}^2+{x_3}^2-1)={x_1}^2-{x_3}^2=0$$

$${x_1}^2-{x_3}^2=(x_1+x_3)(x_1-x_3)=0$$

$$x_1=x_3 \text{ または } x_1=-x_3$$

$x_1=-x_3$のとき、

$$\lambda_1=-\lambda_2$$

$$\lambda_1+\lambda_2=0.$$

$(\lambda_1+\lambda_2)$ は ii′の分母だから、$x_1=-x_3$は採用されない。したがって、

$$x_1=x_3,\quad \lambda_1=-\lambda_2=\lambda$$

これを i′、ii’に入れる。

$$x_1=\frac{1}{2\lambda_1}=\frac{1}{2\lambda}\qquad \text{i}''$$

$$x_2=\frac{1}{2(\lambda_1+\lambda_2)}=\frac{1}{4\lambda}\qquad \text{ii}''$$

i′′、ii” を iv に代入

$$x_1{}^2+x_2{}^2-1=\frac{1}{4\lambda^2}+\frac{1}{16\lambda^2}-1=\frac{5}{16\lambda^2}-1=0\qquad \text{iv}$$

$$\lambda^2=\frac{5}{16}$$

$$\lambda=\pm\frac{\sqrt{5}}{4}$$

これを i’、 ii’ 、iii’に代入

$$x_1=\pm\frac{1}{2\frac{\sqrt{5}}{4}}=\pm\frac{2}{\sqrt{5}}$$

$$x_2=\pm\frac{1}{4\frac{\sqrt{5}}{4}}=\pm\frac{1}{\sqrt{5}}$$

$$x_3=\pm\frac{1}{2\frac{\sqrt{5}}{4}}=\pm\frac{2}{\sqrt{5}}$$

λ	$-\frac{\sqrt{5}}{4}$	$\frac{\sqrt{5}}{4}$
x_1	$-\frac{2}{\sqrt{5}}$	$\frac{2}{\sqrt{5}}$
x_2	$-\frac{1}{\sqrt{5}}$	$\frac{1}{\sqrt{5}}$
x_3	$-\frac{2}{\sqrt{5}}$	$\frac{2}{\sqrt{5}}$
$x_1+x_2+x_3$	$-\sqrt{5}$	$\sqrt{5}$

最大値 $\sqrt{5}$ 、*最小値* $-\sqrt{5}$