§6.4 2次形式 ～ 最大・最小値問題の基礎

2次形式の係数行列 $\overset{n\times n 行列}{A}=\{a_{ij}\}$.

2次形式： $x\in\mathbb{R}^n$ に対し、その2次同次式 $\displaystyle\sum_{i,j=1}^{n} a_{ij}x_ix_j = x^TAx =: A\{x\}$ のコト.

($^\forall a_{ij}\in\mathbb{R}$)

注： $\underset{\substack{\shortparallel \\ x^TAx}}{(x^TAx)^T} = x^TA^Tx$ より、 $x^TAx = x^T\underbrace{\left(\frac{A+A^T}{2}\right)}_{実対称行列}x$ であることから、

スカラーなので.

係数行列 A は実対称行列であることを前提としてしまっても問題ない. 以降, A は実対称とする.

例. $A\{x\} = x_1^2+2x_1x_2+3x_2^2$ は,

$$= \underset{\substack{\uparrow \\ a_{11}=1}}{x_1^2} + \underset{\substack{\uparrow \\ a_{12}=1}}{x_1x_2} + \underset{\substack{\uparrow \\ a_{21}=1}}{x_2x_1} + \underset{\substack{\uparrow \\ a_{22}=3}}{3x_2^2} = (x_1\ x_2)\underbrace{\begin{pmatrix}1 & 1\\ 1 & 3\end{pmatrix}}_{これがA}\begin{pmatrix}x_1\\ x_2\end{pmatrix}$$

を意味するので となっている.

2次形式の標準形 ～ 2種類あるよ.

標準形その1

λ_i は A の固有値

#6-3-5 より、実対称行列 A は直交行列 P で対角化可能，つまり、 $A = P\begin{pmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{pmatrix}P^T$.

これを2次形式 $A\{x\} = x^TAx$ に代入すると、

$$= x^T P\begin{pmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{pmatrix}\underbrace{P^Tx}_{\overset{\shortparallel}{y}\text{として}} = y^T\begin{pmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{pmatrix}y = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2 \quad \sim(*1)$$

となる.

この、Aの固有値を用いた形(*1)を $A\{x\}$ の標準形と呼ぶこともある.

標準形その2

(*1)の y_i を、$\sqrt{|\lambda_i|}\,y_i =: z_i$ と変換すると、

$$\overset{A\{x\}}{\overset{\shortparallel}{(*1)}} = \underbrace{z_1^2 + z_2^2 + \cdots + z_p^2}_{\lambda_i>0\text{ に対応}} \underbrace{- z_{p+1}^2 - z_{p+2}^2 - \cdots - z_{p+q}^2}_{\lambda_i<0\text{ に対応.}} \sim(*2)$$

この(*2)を $A\{x\}$ の標準形と呼ぶこともある.

def. この (p,q) を $A\{x\}$ の符号数と言う.

$$\left\{\begin{array}{l}\text{この項数 } p = \#\{\text{正固有値}\} \\ \text{この項数 } q = \#\{\text{負固有値}\}\end{array}\right.$$

Th. (Sylvester の 慣性則). $A\{\boldsymbol{x}\}$ の符号数は $\overset{n\times n}{A}$ によって一意に定まる. ← ほぼ自明だが、標準形に至る過程中の直交行列 P が一意でないので念の為.

A>0 と書いたりする.

def. $A\{\boldsymbol{x}\}$ が正定値 $\overset{\text{def.}}{\Leftrightarrow}$ $A\{\boldsymbol{x}\}>0$ for $\forall \boldsymbol{x} \neq \boldsymbol{0}$. (負定値の場合は符号が逆).

($\Leftrightarrow$ $\underline{\forall\text{固有値 of } A > 0}$) ← 標準形を考えれば自明.

→ こうである行列を 正定値な行列と言う.

def. $A\{\boldsymbol{x}\}$ が半正定値 $\overset{\text{def}}{\Leftrightarrow}$ $A\{\boldsymbol{x}\} \geq 0$ for $\forall \boldsymbol{x} \neq \boldsymbol{0}$ (半負定値の場合は以下同文)

($\Leftrightarrow$ $\forall$固有値 of $A \geq 0$)

(ちょっとオマケ)

- $A\{\boldsymbol{x}\}$ が正定値 $\Leftrightarrow$ $\boldsymbol{a}^T A \boldsymbol{b}$ for $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^n$ を $\boldsymbol{a}$ と $\boldsymbol{b}$ の内積として扱える.
- 〃 $\Leftrightarrow$ A の $\underline{\forall\text{主小行列式}} > 0$.

→ $\det\begin{pmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & & \\ a_{k1} & \cdots & a_{kk} \end{pmatrix}$, $k \leq n$ のコト.

#6-4-4.

Th. 二次多項式の最大・最小判定 〜 2次形式の重要な応用！

← 何故か
この教科書にはのってない…

実変数 $x_1, x_2, \cdots, x_n$ の二次多項式 $F(x_1, \cdots, x_n)$ を実対称行列 A を用いて

$F(x_1, \cdots, x_n) = \boldsymbol{x}^T A \boldsymbol{x} - 2\boldsymbol{x}^T \boldsymbol{b} + c$ と表わした時、

A が正定値ならば F は $\boldsymbol{x}_* := A^{-1}\boldsymbol{b}$ にて最小になる。（負定値の場合も同様）

↑ 応用上、最小値問題は最終的に二次多項式で近似することが多いので、これはとても重要。

← $\boldsymbol{b} = A\boldsymbol{x}_*$ を使っている

(証) $F(\boldsymbol{x}) - F(\boldsymbol{x}_*) = \boldsymbol{x}^T A \boldsymbol{x} - 2\boldsymbol{x}^T A \boldsymbol{x}_* - \{\boldsymbol{x}_*^T A \boldsymbol{x}_* - 2\boldsymbol{x}_*^T A \boldsymbol{x}_*\}$

$= \boldsymbol{x}^T A \boldsymbol{x} - \boldsymbol{x}^T A \boldsymbol{x}_* - \boldsymbol{x}_*^T A \boldsymbol{x} + \boldsymbol{x}_*^T A \boldsymbol{x}_* = (\boldsymbol{x} - \boldsymbol{x}_*)^T A (\boldsymbol{x} - \boldsymbol{x}_*)$

$\underbrace{> 0}_{A\text{の正定値性から}}$ when $\boldsymbol{x} \neq \boldsymbol{x}_*$ となるので、$F(\boldsymbol{x}_*)$ が最小値。

例. $F(x,y) = x^2+2xy+2y^2-6y$ の最小値は？

$$= \boldsymbol{x}^T \underbrace{\begin{pmatrix}1&1\\1&2\end{pmatrix}}_{A}\boldsymbol{x} - 2\boldsymbol{x}^T \underbrace{\begin{pmatrix}0\\3\end{pmatrix}}_{\mathbb{b}\ \text{として}}$$

と書けるので（ $\boldsymbol{x}=\begin{pmatrix}x\\y\end{pmatrix}$ として ）これを調べる.

$A=\begin{pmatrix}1&1\\1&2\end{pmatrix}$ の固有値を求めると、$0=\varphi_A(\lambda)=\begin{vmatrix}1-\lambda&1\\1&2-\lambda\end{vmatrix} = \lambda^2-3\lambda+2-1 = \lambda^2-3\lambda+1$ より

$\lambda=\frac{1}{2}\left(3\pm\sqrt{9-4}\right)=\frac{1}{2}\left(3\pm\sqrt{5}\right)$ で、この2つの固有値が両方とも正. つまり、Aは正定値.

よって $\boldsymbol{x}_* = A^{-1}\mathbb{b}$ で F は最小値.

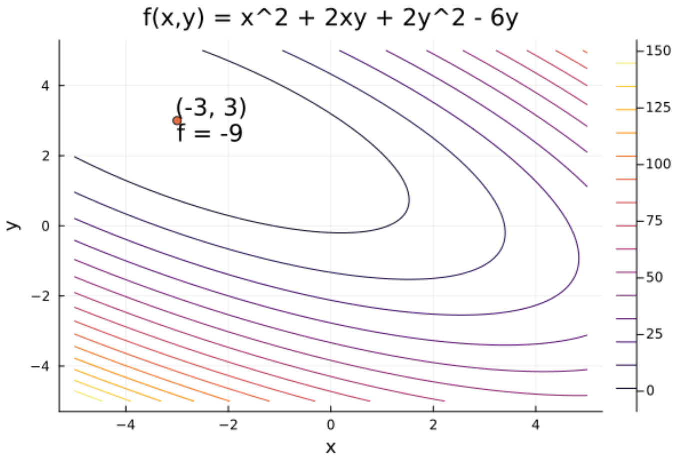
具体的には

$$\underbrace{\left(\underbrace{\begin{matrix}1&1\\1&2\end{matrix}}_{A}\;\vdots\;\underbrace{\begin{matrix}0\\3\end{matrix}}_{\mathbb{b}}\right)}_{\text{拡大係数行列}} \to \begin{pmatrix}1&1&0\\0&1&3\end{pmatrix} \to \left(\begin{array}{cc:c}1&0&-3\\0&1&3\end{array}\right)$$

より $\boldsymbol{x}_*=\begin{pmatrix}-3\\3\end{pmatrix}$ で、

その時の $F = 9-\cancel{2\,9}+\cancel{2\,9}-18=-9$ // ← これが F の最小値.

#6-4-6



← 前ページの $f(x,y)$ の大きさを (x,y)平面上で等高線プロットしたもの.

確かに $\mathbf{x} = \begin{pmatrix}-3\\3\end{pmatrix}$ で $f(x,y)$ が最小になっていることが見てわかる.

(おまけ)

Aは正定値とする.

$F(x) = x^T A x - 2x^T \boldsymbol{b} + c$ が $x_* := A^{-1}\boldsymbol{b}$ で最小になる時の $F(x_*)$ の値は？

→ $x_* = A^{-1}\boldsymbol{b} \Leftrightarrow Ax_* = \boldsymbol{b}$ なことを用いて、

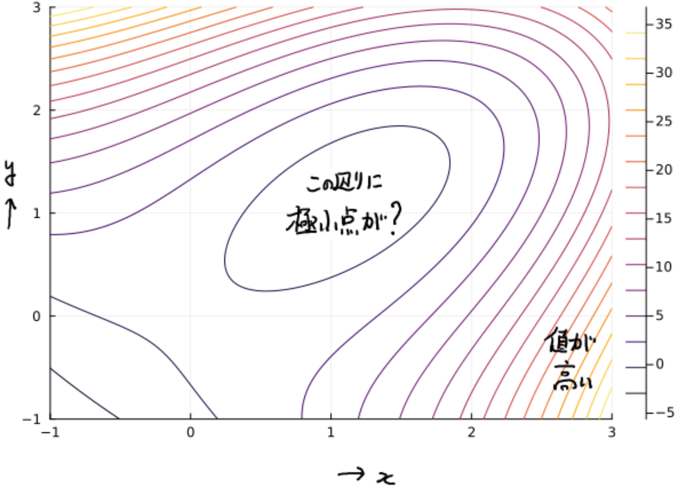
$$F(x_*) = x_*^T A x_* - 2x_*^T \boldsymbol{b} + c = x_*^T \boldsymbol{b} - 2x_*^T \boldsymbol{b} + c = \underline{-x_*^T \boldsymbol{b} + c} \ //$$

cの値

例1. #6-4-5 の F と x_* だと、$F(x_*) = -\begin{pmatrix}-3\\3\end{pmatrix}^T \cdot \begin{pmatrix}0\\3\end{pmatrix} + 0 = -9$ で正しいことが確認できる.

#6-4-8

f(x, y) の大まかな等高線図



応用問題

具体的な$\overbrace{\text{数字}}^{\text{の}\, x,\, y}$に対して

$$\nabla f(x,y) = \begin{pmatrix} \frac{\partial}{\partial x} f(x,y) \\ \frac{\partial}{\partial y} 〃 \end{pmatrix} \text{や}$$

$$H_f(x,y) = \begin{pmatrix} \frac{\partial^2}{\partial x^2} f & \frac{\partial^2}{\partial x \partial y} 〃 \\ \frac{\partial^2}{\partial x \partial y} 〃 & \frac{\partial^2}{\partial y^2} 〃 \end{pmatrix}$$

が数字で得られる時、左図のようなケースで極小点を近似で良いので求められるだろうか？

解法例

最初はどうにかして、決める。

1. 極小点の近似を $\mathbf{x}_k := (x_k, y_k)$ とする。

2. $\mathbf{x}_k$ 付近で f を Taylor展開すると、

$$f(\mathbf{x}) \simeq f(\mathbf{x}_k) + \underbrace{\nabla f(\mathbf{x}_k)}_{\text{分かる}}{}^T(\mathbf{x}-\mathbf{x}_k) + \frac{1}{2}(\mathbf{x}-\mathbf{x}_k)^T \underbrace{H_f(\mathbf{x}_k)}_{\text{分かる}}(\mathbf{x}-\mathbf{x}_k)$$

なので、

二次形式のギロンにより

$$\mathbf{x}-\mathbf{x}_k = -H_f(\mathbf{x}_k)^{-1}\nabla f(\mathbf{x}_k)$$

、即ち、

$$\mathbf{x} = \mathbf{x}_k - H_f(\mathbf{x}_k)^{-1}\nabla f(\mathbf{x}_k)$$

が、$\mathbf{x}_k$ よりもより良い極小点の近似！と思える。

よって、

$$\mathbf{x}_{k+1} := \mathbf{x}_k - H_f(\mathbf{x}_k)^{-1}\nabla f(\mathbf{x}_k)$$

として、近似点を改善する。　以下、くり返せば良い。

例題. 前ページ(#6-4-8)の話を、実際に試してみよう.

(函数形を陽に知っているのはインチキだが、あくまで練習台として) $f(x, y) = x^3 + y^3 - \underbrace{3.6}\,xy$ に対して.

扱いにくいので、$3.6 = 3a$, $a = 1.2$ として
おく.

まず、近似点 x_k の改善 step は $x_{k+1} = x_k - H_f(x_k)^{-1} \nabla f(x_k)$ であるので.

必要なのは ∇f の値と H_f の値.

各々計算すると、

$$\nabla f = \begin{pmatrix} 3x^2 - 3ay \\ 3y^2 - 3ax \end{pmatrix} = 3\begin{pmatrix} x^2 - ay \\ y^2 - ax \end{pmatrix}, \quad H_f = 3\begin{pmatrix} 2x & -a \\ -a & 2y \end{pmatrix}$$

なので、$\mathbb{b} := H_f^{-1} \nabla f$ とすると.
($H_f \mathbb{b} = \nabla f$ を解けば $\mathbb{b}$)

$$H_f^{-1} = \frac{1}{3} \frac{1}{4xy - a^2} \begin{pmatrix} 2y & a \\ a & 2x \end{pmatrix}$$

より、

$$\mathbb{b} = \frac{1}{4xy - a^2} \begin{pmatrix} 2y & a \\ a & 2x \end{pmatrix} \begin{pmatrix} x^2 - ay \\ y^2 - ax \end{pmatrix}.$$

よって.

$$x_{k+1} = x_k - \frac{1}{4xy - a^2} \begin{pmatrix} 2y & a \\ a & 2x \end{pmatrix} \begin{pmatrix} x^2 - ay \\ y^2 - ax \end{pmatrix}$$

として x_k を更新することになる.

よって、例えば極小点近似値として $x_1=(1,1)$ と（勘かなにかで）決めたとすると、

$$x_2 = x_1 - \frac{1}{4-a^2}\begin{pmatrix}2 & a\\ a & 2\end{pmatrix}\begin{pmatrix}1-a\\ 1-a\end{pmatrix} = x_1 - \frac{1-a}{4-a^2}\begin{pmatrix}2+a\\ 2+a\end{pmatrix} = x_1 - \frac{(1-a)\cancel{(2+a)}}{\underset{2-a}{\cancel{4-a^2}}}\begin{pmatrix}1\\ 1\end{pmatrix}$$

$$= \frac{(\overset{1}{\cancel{2}}-\cancel{a})-(\cancel{1}-\cancel{a})}{2-a}\begin{pmatrix}1\\ 1\end{pmatrix} = \frac{1}{2-a}\begin{pmatrix}1\\ 1\end{pmatrix} = \frac{1}{0.8}\begin{pmatrix}1\\ 1\end{pmatrix} = \boxed{\frac{5}{4}}\begin{pmatrix}1\\ 1\end{pmatrix},$$

$$x_3 = x_2 - \frac{1}{4\cdot(\frac{5}{4})^2-a^2}\begin{pmatrix}2\cdot\frac{5}{4} & a\\ a & 2\cdot\frac{5}{4}\end{pmatrix}\begin{pmatrix}(\frac{5}{4})^2-a(\frac{5}{4})\\ 〃\end{pmatrix} = x_2 - \frac{(\frac{5}{4})(\frac{5}{4}-a)}{\underset{\frac{5}{2}-a}{\cancel{(\frac{5}{2})^2-a^2}}}\cdot\cancel{(\frac{5}{2}+a)}\begin{pmatrix}1\\ 1\end{pmatrix}$$

$$= \left(\frac{5}{4}\right)\cdot\frac{(\overset{\frac{5}{4}}{\cancel{\frac{5}{2}}}-\cancel{a})-(\cancel{\frac{5}{4}}-\cancel{a})}{\frac{5}{2}-a}\cdot\begin{pmatrix}1\\ 1\end{pmatrix} = \left(\frac{5}{4}\right)^2\cdot\frac{1}{\frac{5}{2}-\frac{6}{5}}\begin{pmatrix}1\\ 1\end{pmatrix} = \left(\frac{5}{4}\right)^2\cdot\frac{10}{13}\begin{pmatrix}1\\ 1\end{pmatrix} = \boxed{\frac{125}{104}}\begin{pmatrix}1\\ 1\end{pmatrix},$$

$$x_4 = x_3 - \frac{1}{4\cdot(\frac{125}{104})^2-a^2}\begin{pmatrix}2\cdot\frac{125}{104} & a\\ a & 2\cdot\frac{125}{104}\end{pmatrix}\begin{pmatrix}(\frac{125}{104})^2-a(〃)\\ 〃\end{pmatrix} = x_3 - \left(\frac{125}{104}\right)\cdot\frac{(\frac{125}{104}-a)}{\underset{\frac{125}{52}-a}{\cancel{(\frac{125}{52})^2-a^2}}}\cdot\cancel{(\frac{125}{52}+a)}\begin{pmatrix}1\\ 1\end{pmatrix}$$

$$= \frac{125}{104}\cdot\frac{(\frac{125}{\underset{104}{\cancel{52}}}-\cancel{a})-(\cancel{\frac{125}{104}}-\cancel{a})}{\frac{125}{52}-a}\begin{pmatrix}1\\ 1\end{pmatrix} = \left(\frac{125}{104}\right)^2\frac{1}{\frac{125}{52}-\frac{6}{5}}\begin{pmatrix}1\\ 1\end{pmatrix} = \boxed{\left(\frac{5^7}{2^4\cdot 13\cdot 313}\right)}\begin{pmatrix}1\\ 1\end{pmatrix}$$

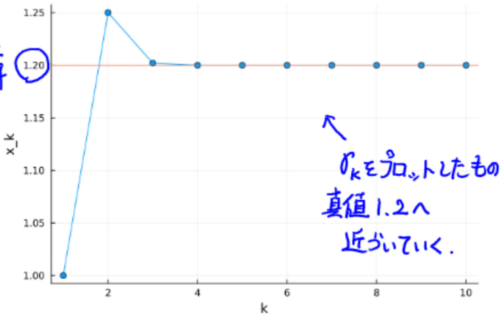
$\bigcirc$ 部分を見ると、結局、 $\bigcirc_{New} = \bigcirc_{old}^2 \Big/ \left(2\bigcirc_{old} - 6/5\right)$ となっていることが分かる。つまり、（→次ページ）

$x_1 = 1 \cdot v$, ただし $v := \begin{pmatrix}1\\1\end{pmatrix}$, に対し、

$x_2 = \boxed{\frac{5}{4}}\, v$, $x_3 = \boxed{\frac{125}{104}}\, v$, $x_4 = \boxed{\cdots}$, $x_k = \gamma_k v$ として、γ_k をプロットしてみると、

$\frac{5}{4} = 1.25$, $\frac{125}{104} = 1.2019\ldots$, $\cdots = 1.200003\ldots$

$\underset{\sim\sim}{1.2}$ に近づいていく。



正解 $\boxed{1.20}$

γ_k をプロットしたもの。
真値 1.2 へ近づいていく。

自分で別途に計算して確かめておこう。

実は、$x_* = \begin{pmatrix}1.2\\1.2\end{pmatrix}$ が真の極小点で、

x_k はそこへ向かって近づいているのだ。

つまり、近似の改善がうまくいった、のだ。