数理計画法　第10回
第3章　非線形計画
3.2　制約なし最適化

担当：　塩浦昭義
(情報科学研究科　助教授)
shioura@dais.is.tohoku.ac.jp

復習：制約なし最適化問題，勾配ベクトル

入力：目的関数 f(x) のみ

最小化　f(x)　条件　なし（$x \in R^n$）

関数 f の勾配ベクトル

$$\nabla f(x) = \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{pmatrix}$$

一変数関数の場合は

$$\nabla f(x) = f'(x)$$

復習：一次のテイラー展開[p.89]

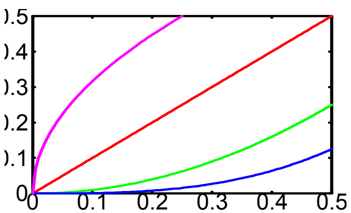
関数 f は勾配ベクトルを傾きとする線形関数により近似できる

$$f(\boldsymbol{x}+\boldsymbol{d}) = f(\boldsymbol{x}) + \nabla f(\boldsymbol{x})^T \boldsymbol{d} + o(\|\boldsymbol{d}\|)$$

$$\|\mathbf{d}\| = \sqrt{d_1^2 + d_2^2 + \cdots + d_n^2}$$

記号 $o(\alpha)$：$\lim_{\alpha \to 0} \frac{\varphi(\alpha)}{\alpha} = 0$ を満たす関数 $\varphi(\alpha)$ を表す

例えばα^2，α^3など，
αより速く0に
近づく関数を表す

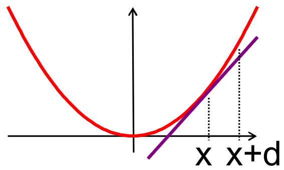


復習：一次のテイラー展開

例1：$f_1(x) = x^2$　　$f_1'(x) = 2x$，$f_1''(x) = 2$

f_1 の x における一次のテイラー展開

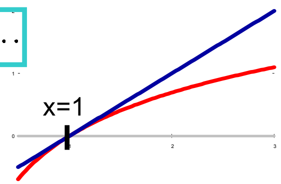
$$f_1(x+d) = (x+d)^2 = x^2 + (2x)d + \boxed{d^2}$$
$$= x^2 + (2x)d + o(|d|)$$



例2：$f_2(x) = \log x$　　$f_2'(x) = 1/x$，$f_2''(x) = -1/x^2$，…

f_2 の x=1 における一次のテイラー展開

$$f_2(1+d) = 0 + d \boxed{- \frac{d^2}{2} + \frac{d^3}{3} + \frac{d^4}{4} - \cdots}$$
$$= 0 + d + o(|d|)$$



復習：勾配ベクトルの性質 [p.89]

勾配ベクトルと逆の方向に進むと関数値が減る

性質：任意の x に対し、∇f ≠0 ならば
十分小さい $\delta>0$に対して　$f(\boldsymbol{x}-\delta\nabla f(\boldsymbol{x})) < f(\boldsymbol{x})$

証明: $\delta=\varepsilon/\|\nabla f(\boldsymbol{x})\|, \boldsymbol{d}=-\delta\nabla f(\boldsymbol{x})$ とおく（ε：正の数）

テイラー展開より
$$f(\boldsymbol{x}+\boldsymbol{d}) = f(\boldsymbol{x})+\nabla f(\boldsymbol{x})^T\boldsymbol{d}+o(\|\boldsymbol{d}\|) = f(\boldsymbol{x})-\varepsilon\|\nabla f(\boldsymbol{x})\|+o(\varepsilon)$$

正数εを十分小さくする $\Rightarrow$ $\dfrac{o(\varepsilon)}{\varepsilon}$ は0に近い値

$$-\varepsilon\|\nabla f(\boldsymbol{x})\|+o(\varepsilon)=\varepsilon\left(-\|\nabla f(\boldsymbol{x})\|+\frac{o(\varepsilon)}{\varepsilon}\right)<0$$

$$\therefore\ f(\boldsymbol{x}+\boldsymbol{d}) < f(\boldsymbol{x})$$

復習：最急降下法

入力：関数 f とその勾配ベクトル∇f
初期点 x^0

ステップ0: k =0 とする
ステップ1: x^k が最適解に十分近ければ終了
ステップ2：最急降下方向 $-\nabla f(x^k)$ を計算
ステップ3：直線探索問題
最小化　$f(x^k-\alpha\nabla f(x^k))$　条件　$\alpha>0$
を解き、解をα^kとする
ステップ4: $x^{k+1}=x^k-\alpha^k\nabla f(x^k)$ とおく
ステップ5: k=k+1として、ステップ1に戻る

ヘッセ行列 [p.90]

関数 f のヘッセ行列

$$\mathrm{H}\,f(\boldsymbol{x})=\begin{pmatrix} \dfrac{\partial^2 f}{\partial x_1\partial x_1} & \dfrac{\partial^2 f}{\partial x_1\partial x_2} & \cdots & \dfrac{\partial^2 f}{\partial x_1\partial x_n} \\ \dfrac{\partial^2 f}{\partial x_2\partial x_1} & \dfrac{\partial^2 f}{\partial x_2\partial x_2} & \cdots & \dfrac{\partial^2 f}{\partial x_2\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial^2 f}{\partial x_n\partial x_1} & \dfrac{\partial^2 f}{\partial x_n\partial x_2} & \cdots & \dfrac{\partial^2 f}{\partial x_n\partial x_n} \end{pmatrix}$$

一変数関数の場合は2階導関数に一致

$$\mathrm{H}\,f(x)=f''(x)$$

ヘッセ行列（続き） [p.89]

例：

$f_1(x)=x^2$　　$\nabla f_1(x)=f_1'(x)=2x$, $\mathrm{H}f_1(x)=f_1''(x)=2$

$$f_2(x_1,x_2)=\sin x_1+\cos x_2 \Rightarrow \nabla f_2(\boldsymbol{x})=\begin{pmatrix}\cos x_1\\ -\sin x_2\end{pmatrix}$$

$$\Rightarrow \mathrm{H}\,f_2(\boldsymbol{x})=\begin{pmatrix}-\sin x_1 & 0\\ 0 & -\cos x_2\end{pmatrix}$$

$$f_3(x_1,x_2)=x_1x_2+\log x_1$$

$$\Rightarrow \nabla f_3(\boldsymbol{x})=\begin{pmatrix}x_2+1/x_1\\ x_1\end{pmatrix} \Rightarrow \mathrm{H}\,f_3(\boldsymbol{x})=\begin{pmatrix}-1/x_1^2 & 1\\ 1 & 0\end{pmatrix}$$

ヘッセ行列とテイラー展開[p.90]

関数 f は勾配ベクトルとヘッセ行列により表現される2次関数により近似できる

$$f(\boldsymbol{x}+\boldsymbol{d}) = f(\boldsymbol{x}) + \nabla f(\boldsymbol{x})^T \boldsymbol{d} + \frac{1}{2}\boldsymbol{d}^T \mathrm{H} f(\boldsymbol{x})\boldsymbol{d} + o(\|\boldsymbol{d}\|^2)$$

関数 f の(x における)
2次のテイラー展開

関数 f を多項式で表現したときの
2次項までの情報を取り出した式

ヘッセ行列とテイラー展開(続き)[p.90]

$$f(\boldsymbol{x}+\boldsymbol{d}) = f(\boldsymbol{x}) + \nabla f(\boldsymbol{x})^T \boldsymbol{d} + \frac{1}{2}\boldsymbol{d}^T \mathrm{H} f(\boldsymbol{x})\boldsymbol{d} + o(\|\boldsymbol{d}\|^2)$$

例1: $f_1(x) = x^2$　　$\nabla f_1(x) = 2x$　　$\mathrm{H} f_1(x) = 2$

左辺 $= f_1(x+d) = (x+d)^2$

右辺 $= x^2 + (2x)d + d^2 = (x+d)^2$

∴2次関数 f_1 の2次のテイラー展開は f_1 に等しい

※一般に、2次関数 $f(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^T V \boldsymbol{x} + \boldsymbol{c}^T \boldsymbol{x} + c_0$

のテイラー展開は f に等しい

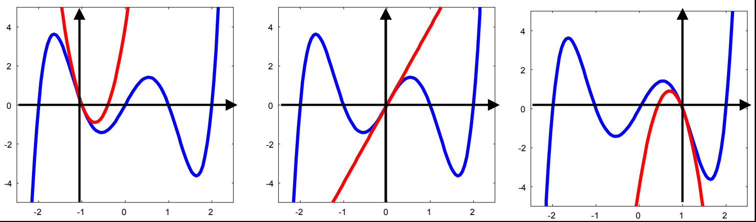
V: n×n行列
c: n 次元ベクトル
c_0: 定数

ヘッセ行列とテイラー展開(続き)[p.90]

例2: $f(x) = (x-2)(x-1)x(x+1)(x+2) = x^5 - 5x^3 + 4x$

$\nabla f(x) = 5x^4 - 15x^2 + 4$　　$\mathrm{H} f(x) = 20x^3 - 30x$

	x = -1 のとき	x = 0 のとき	x = 1 のとき
テイラー展開	$0 - 6d + 5d^2$	$0 + 4d + 0d^2$	$0 - 6d - 5d^2$



行列の正定値性、半正定値性[p.99]

正定値(半正定値)・・・・行列が「正(非負)」

定義:正方行列 A は半正定値
　⇔ 任意のベクトル y に対して $y^T A y \geqq 0$

定義:正方行列 A は正定値
　⇔ 任意の非零ベクトル y に対して $y^T A y > 0$

※ A が1×1行列のとき、
Aは半正定値 ⇔ $a_{11} \geqq 0$,　Aは正定値 ⇔ $a_{11} > 0$

行列の正定値性、半正定値性[p.99]

定義：正方行列 A は半正定値
⇔ 任意のベクトル y に対して $y^T A y \geqq 0$

定義：正方行列 A は正定値
⇔ 任意の非零ベクトル y に対して $y^T A y > 0$

※ A が2×2行列のとき、
Aは正定値 ⇔ $a_{11} > 0,\ a_{22} > 0,\ a_{11}a_{22} - a_{12}^2 > 0$
Aは半正定値 ⇔ $a_{11} \geqq 0,\ a_{22} \geqq 0,\ a_{11}a_{22} - a_{12}^2 \geqq 0$

$\begin{pmatrix} 2 & -2 \\ -2 & 5 \end{pmatrix}$正定値　$\begin{pmatrix} 2 & -2 \\ -2 & 2 \end{pmatrix}$半正定値　$\begin{pmatrix} 2 & -2 \\ -2 & 0 \end{pmatrix}$半正定値ではない

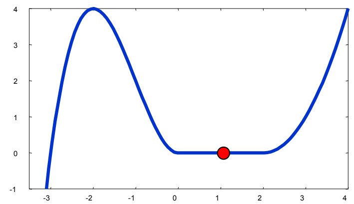
2次の最適性条件（必要条件）[p.99]

ヘッセ行列を用いた最適性条件

定理（2次の必要条件）：
x*: 制約なし問題の極小解
⇒ Hf(x*) は半正定値

例：
x^* = 1 は極小解
0≦x≦2 の範囲で f(x) = 0

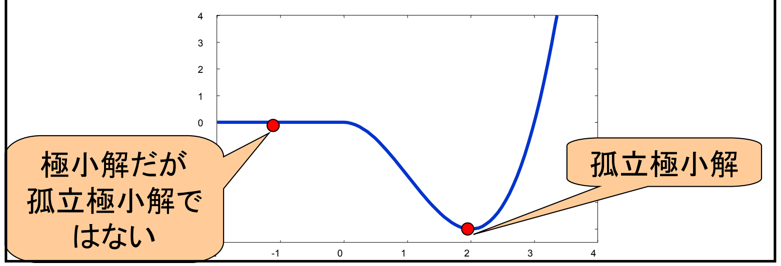
⇒ $\nabla f(x^*) = f'(x^*) = 0$
$\mathrm{H}f(x^*) = f''(x^*) = 0$
半正定値



2次の最適性条件（十分条件）[p.100]

定理（2次の十分条件）：
x* は停留点，Hf(x*) は正定値
⇒ x*: 制約なし問題の（孤立）極小解

定義：x* は孤立極小解
⇔ x* は極小、近傍内に同じ関数値をもつ点が存在しない



2次の最適性条件（十分条件）[p.100]

定理：Hf(x*) は正定値 ⇒ （孤立）極小解

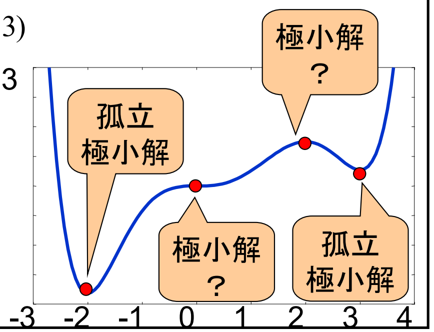
例1：$f(x) = \frac{1}{6}x^6 - \frac{3}{5}x^5 - x^4 + 4x^3$

$\nabla f(x) = (x+2)x^2(x-2)(x-3)$

⇒ 停留点はx = -2, 0, 2, 3

$\mathrm{H}\,f(x) = 5x^4 - 12x^3 - 12x^2 + 24x$

⇒ $\mathrm{H}f(-2) = 80 > 0$
$\mathrm{H}f(0) = 0$
$\mathrm{H}f(2) = -16 < 0$
$\mathrm{H}f(3) = 45 > 0$



2次の最適性条件（十分条件） [p.100]

定理: x* は停留点，Hf(x*) は正定値
　⇒　x*: （孤立）極小解

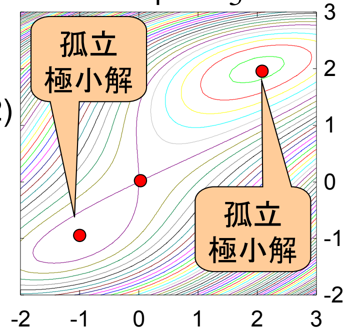
例2（教科書の例3．4）： $f(\boldsymbol{x}) = x_1^2 - 2x_1x_2 + \frac{1}{4}x_2^4 - \frac{1}{3}x_2^3$

$$\nabla f(\boldsymbol{x}) = \begin{pmatrix} 2x_1 - 2x_2 \\ x_2^3 - x_2^2 - 2x_1 \end{pmatrix}$$

⇒ 停留点は(0,0), (-1, -1), (2, 2)

$$\mathrm{H}f(\boldsymbol{x}) = \begin{pmatrix} 2 & -2 \\ -2 & 3x_2^2 - 2_2 \end{pmatrix}$$

⇒ (-1, -1), (2, 2) は
孤立極小解



極大解に関する性質

- $\boldsymbol{x}^*$ は関数 f の（孤立）極大解
　⇔　$\boldsymbol{x}^*$ は関数 $-f$ の（孤立）極小解
- $\boldsymbol{x}^*$ における関数 $-f$ のヘッセ行列は $-\mathrm{H}f(\boldsymbol{x})$

⇒　極大解であるための条件

定理:
x*: 制約なし問題の極大解　⇒　− Hf(x*) は半正定値

定理:
x* は停留点，− Hf(x*) は正定値
　⇒　x*: 制約なし問題の（孤立）極大解

2次の必要条件の証明 [p.99]

定理　x*: 極小解　⇒　Hf(x*) は半正定値

証明: ・Hf(x*) が半正定値でないと仮定

⇒ ある u （||u|| = 1 と仮定）に対し u^T Hf(x*)u ＜ 0

この値を -δ とおく

・x* は極小解　⇒　∇f(x*)＝0

・2次のテイラー展開の式に d = εu を代入（ε＞0）

$$f(\boldsymbol{x}^* + \varepsilon \boldsymbol{u}) = f(\boldsymbol{x}^*) + \varepsilon \nabla f(\boldsymbol{x}^*)^T \boldsymbol{u} + \frac{1}{2}\varepsilon^2 \boldsymbol{u}^T \mathrm{H}f(\boldsymbol{x}^*)\boldsymbol{u} + o(\varepsilon^2)$$

$$= f(\boldsymbol{x}^*) - \frac{1}{2}\varepsilon^2\delta + o(\varepsilon^2) = f(\boldsymbol{x}^*) - \varepsilon^2\left(\frac{1}{2}\delta - \frac{o(\varepsilon^2)}{\varepsilon^2}\right)$$

・ε を十分小さくする　⇒　δ/2 − o(ε²)/ε² > 0
　⇒ f(x* + εu) < f(x*)　　[矛盾]

レポート問題

問題１:下記の4つの関数のヘッセ行列を計算しなさい

$f_1(x_1, x_2) = x_1 + 2x_2$　　$f_2(x_1, x_2) = x_1^2 + x_2^2 - 1$

$f_3(x_1, x_2) = x_1 \log x_2 - x_2 \log x_1$

(ただし，$\boldsymbol{x}$はn次元ベクトル，$\boldsymbol{V}$は$n\times n$対称行列)

問題２:関数 f_2, f_3 に対し，x＝(1，2)におけるテイラー展開を求めなさい．

問題3：対称な2×2行列 A に対し、次の関係を証明せよ。
　Aは半正定値 ⇔ $a_{11} \geq 0$, $a_{22} \geq 0$, $a_{11}a_{22} - a_{12}^2 \geq 0$
　（ヒント：教科書の問題3.7の答えを参考にせよ）