数理計画法　第8回

第3章　非線形計画

3.1　非線形計画法の基礎
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：非線形計画問題の分類

制約なし最適化問題

入力：目的関数 f(x) のみ

最小化　f(x)　条件　なし　($x \in R^n$)

制約つき最適化問題

入力：目的関数 f(x)
　　および制約を表す関数 $g_i(x)$ (i = 1, 2, …, m)

最小化　f(x)　条件　$g_i(x) \leqq 0$ (i = 1, 2, …, m)

この講義では、制約なし問題を主に扱う

復習：勾配ベクトル

関数 f の勾配ベクトル

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

一変数関数の場合は $\nabla f(x) = f'(x)$

例：

$f_1(x) = x^2 \Rightarrow \nabla f_1(x) = 2x$

$f_2(x_1, x_2) = \sin x_1 + \cos x_2$

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

勾配ベクトルとテイラー展開[p.89]

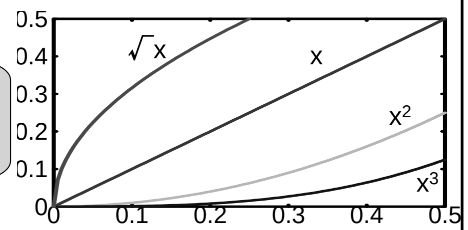
関数 f は勾配ベクトルを傾きとする線形関数により近似できる

(一次の) テイラー展開

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

記号 o(α)：α→0のときにg(α)/α→0となる関数g(α)を表す.

α², α³など，
αより速く0に
近づく関数を表す



勾配ベクトルとテイラー展開[p.89]

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

例：線形関数の場合 $f(x) = \boldsymbol{a}^T x + b$

$$\nabla f(x) = \boldsymbol{a}$$

$$f(x+\boldsymbol{d}) = \boldsymbol{a}^T(x+\boldsymbol{d}) + b$$
$$= f(x) + \nabla f(x+\boldsymbol{d})^T$$

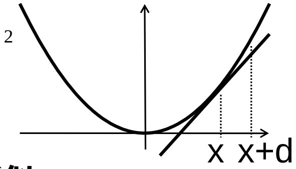
テイラー展開はもとの関数に一致

勾配ベクトルとテイラー展開[p.89]

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

例： $f_1(x) = x^2 \qquad \nabla f_1(x) = 2x$

$$f_1(x+d) = (x+d)^2 = x^2 + (2x)d + d^2$$
$$= x^2 + (2x)d + o(|d|)$$



関数 f を点 x において線形関数で近似

$(x+d)^2$ **と** $x^2 + 2xd$ **の誤差　ー　**$|d|$**が小さければ十分小さい**

勾配ベクトルの性質 [p.89]

勾配ベクトルの方向に進むと関数値が増える

性質：任意の x に対し、∇f ≠0 ならば
十分小さい δ＞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 ⇨ $\dfrac{o(\varepsilon)}{\varepsilon}$ **は0に近い値**

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

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

勾配ベクトルの性質（続き） [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 δ＞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 ⇨ $\dfrac{o(\varepsilon)}{\varepsilon}$ **は0に近い値**

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

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

制約なし問題の最適性条件 [p.97]

制約なし最適化問題：　**最小化**　f(x)

最適性条件：**ベクトル** x **が非線形計画問題の最適解であるための必要条件**

> **定理** (**制約なし最適化問題の最適性条件**)：
> x*: **制約なし問題の最適解** ⇒ **∇** f(x*) = **0**

証明：　**∇** f(x*)**≠0 と仮定**

勾配ベクトルの性質より、十分小さい δ＞0に対して

$$f(\boldsymbol{x}^* - \delta \nabla f(\boldsymbol{x}^*)) < f(\boldsymbol{x}^*)$$

x* **が最適解であることに矛盾**　　**∴ ∇** f(x*) **＝0**

制約なし問題の最適性条件 [p.97]

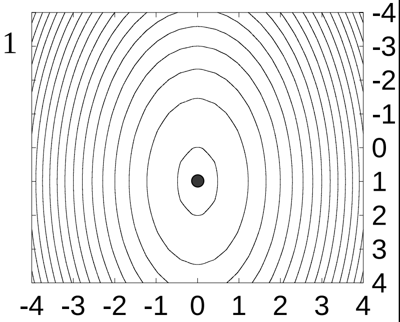
> **定理** (**制約なし最適化問題の最適性条件**)：
> x*: **制約なし問題の最適解** ⇒ **∇** f(x*) = **0**

例：
$$f(\boldsymbol{x}) = x_1^2 - 2x_1 + 4x_2^2$$
$$= (x_1 - 1)^2 + 4x_2^2 - 1$$

(x_1,x_2) = (1,0) **が最適解**

$$\nabla f(\boldsymbol{x}) = \begin{pmatrix} 2x_1 - 2 \\ 8x_2 \end{pmatrix}$$

$$\nabla f(1,0) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$



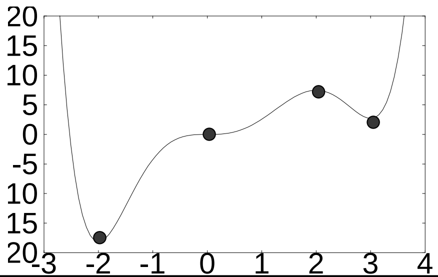
停留点 [p.98]

> **定義**：x **は停留点** ⇔ **∇** f(x) **＝0**

※「x***は停留点** ⇒ x***は最適解**」**は必ずしも成り立たない**

例：
$$f(x) = \frac{1}{6}x^6 - \frac{3}{5}x^5 - x^4 + 4x^3 \qquad \nabla f(x) = (x+2)x^2(x-2)(x-3)$$

停留点はx = -2, 0, 2, 3
最適解はx = -2 **のみ**



極小解、極大解、鞍点 [p.99]

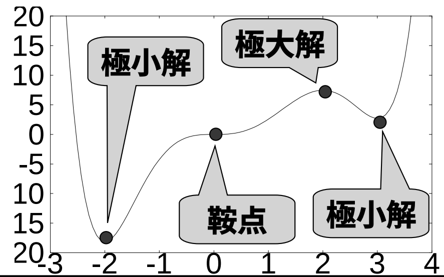
停留点 x* **の分類**

極小解: x* **の付近だけに注目したとき、**x* **は最小**

> **ある δ ＞**0 **が存在して、**||x – x*||**≦ δ を満たす**
> **すべての** x **に対して** f(x) **≧**f(x*)

極大解: x* **の付近だけに注目したとき、**x* **は最大**

鞍点：**極小点でも極大点でもない停留点**



制約なし問題の解法1：最急降下法
[p.102]

最急降下法のアイディア：
勾配ベクトルと逆の方向に進むと関数値が減る

現在の点 x を x－α∇f(x) により更新
⇒ 関数値 f(x) を減らしていく

ステップサイズ

ステップサイズの選び方：
次の一変数最適化問題を（近似的に）解く
最小化 f(x－α∇f(x)) 条件 α＞0

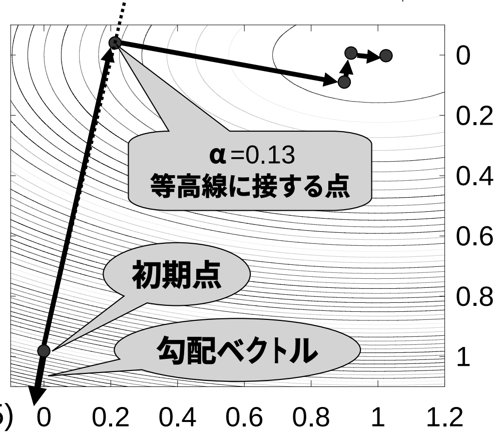
直線探索と呼ばれる

最急降下法の実行例 [p.103]

教科書の例3.2,3.8: $f(\boldsymbol{x}) = x_1^2 - 2x_1 + 4x_2^2$

$$\nabla f(\boldsymbol{x}) = \begin{pmatrix} 2x_1 - 2 \\ 8x_2 \end{pmatrix}$$

- $(x_1, x_2) = (0,1)$ からスタート
- ∇f(0,1) = (-2,8)
- f(0 + 2α,1 – 8α) を最小にするのは α=0.13
- 次の点は $(x_1, x_2) = (0.26, -0.05)$



最急降下法のアルゴリズム [p.102]

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 α＞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

最適解の判定 [p.104,105]

- 非線形計画問題では
  最適解を正確に求めることは困難
  ⇒ 最適解に十分近い解（近似最適解）を求める

例: $f(x) = x^4 - 4x^2$
この関数を最小にする x は 0, $\pm\sqrt{2}$
無理数をコンピュータで表現することは不可能

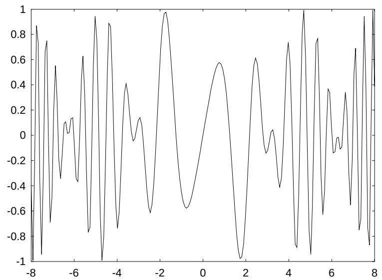
- 最適解に十分近いことをどうやって判定する？
  （方法1） 最適解 x* に対し ||∇f(x)|| = 0 が成り立つ
  ⇒ ||∇f(x)|| の値が十分小さくなったら終了
  （方法2） 最適解の近くでは x^k があまり変化しない
  ⇒ $||x^{k+1} - x^k||$ の値が十分小さくなったら終了

最適解の判定（つづき）[p.104,105]

•非線形計画問題では
近似最適解すら求めることが困難なことが多い

⇩

極小解または停留点を
求めることで我慢する



•極小解は良い解であること
が多い

•ある種の非線形関数では
極小解⇔最小解

定理：ある仮定の下では、
最急降下法の求める点列は停留点に収束する

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：

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

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

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

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

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

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

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

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

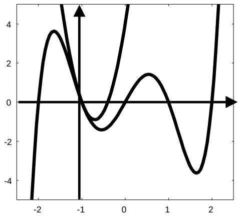
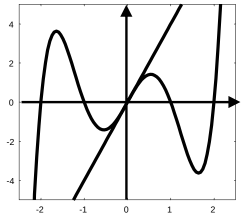
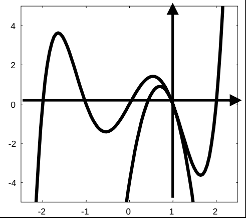
V: n×n行列
c: n 次元ベクトル
c_0: 定数

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

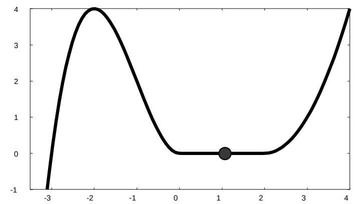
2次の最適性条件 (必要条件) [p.99]

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

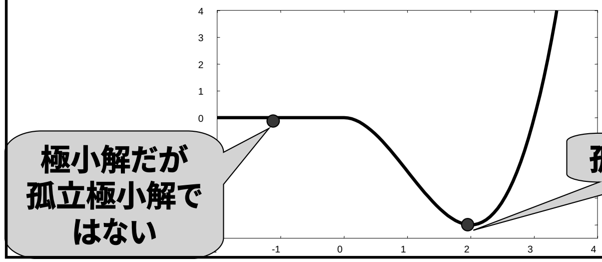
⇒ ∇f(x*) = f '(x*) = 0
Hf(x*) = f ''(x*) = 0
半正定値



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**

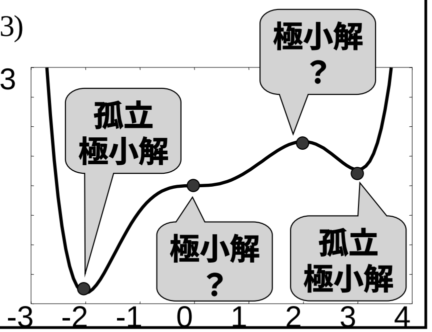
例1： $f(x)=\frac{1}{6}x^6-\frac{3}{5}x^5-x^4+4x^3$

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

⇒ 停留点はx = -2, 0, 2, 3

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

⇒ Hf(-2) = 80 > 0
Hf(0) = 0
Hf(2) = -16 < 0
Hf(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

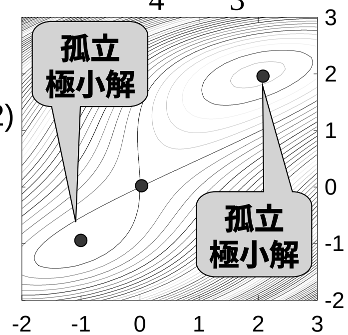
例2 (教科書の例3．4)： $f(\boldsymbol{x})=x_1^2-2x_1x_2+\frac{1}{4}x_2^4-\frac{1}{3}x_2^3$

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

⇒ 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

- x^* は関数 f の (孤立) 極大解
  ⇔ x^* は関数 – f の (孤立) 極小解
- x^* における関数 – f のヘッセ行列は – Hf(**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

レポート問題 (締め切り:1月12日)

(1) 教科書p.162, 問題3.1

(2) 関数 $f(x,y) = (x-2)^4 + (x-2y)^2$ に対して、初期点を (0, 3) として最急降下法を適用せよ。
資料に添付してある等高線の図を使って実行すればよい.
(数値を具体的に計算するのは一回目の反復だけでよい).