数理計画法　第12回

非線形計画

二次の最適性条件とニュートン法

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

期末試験について

- 日時：1月24日（木）13:00～14:30
- 場所：総合研究棟110講義室（この部屋）
- 手書きのA4用紙一枚のみ持ち込み可（印刷やコピーは不可）
  - これも採点の対象，試験終了後に回収します
- 教科書，ノート等の持ち込みは不可
- 座席はこちらで指定
- 試験内容：第7回目以降の講義で教えたところ
  - ネットワーク計画，非線形計画
  - 中間試験でやったところは範囲外
- 50点満点，29点以下は原則として不合格
- インフルエンザやノロウィルス感染時は無理に来ないこと
  - 事前にメールの連絡の上，後日医師の診断書を持参すれば，再試験を認めます

最適解の判定

- 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正確に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 $\|\nabla f(x)\| = 0$ が成り立つ

⇒ $\|\nabla f(x)\|$ の値が十分小さくなったら終了

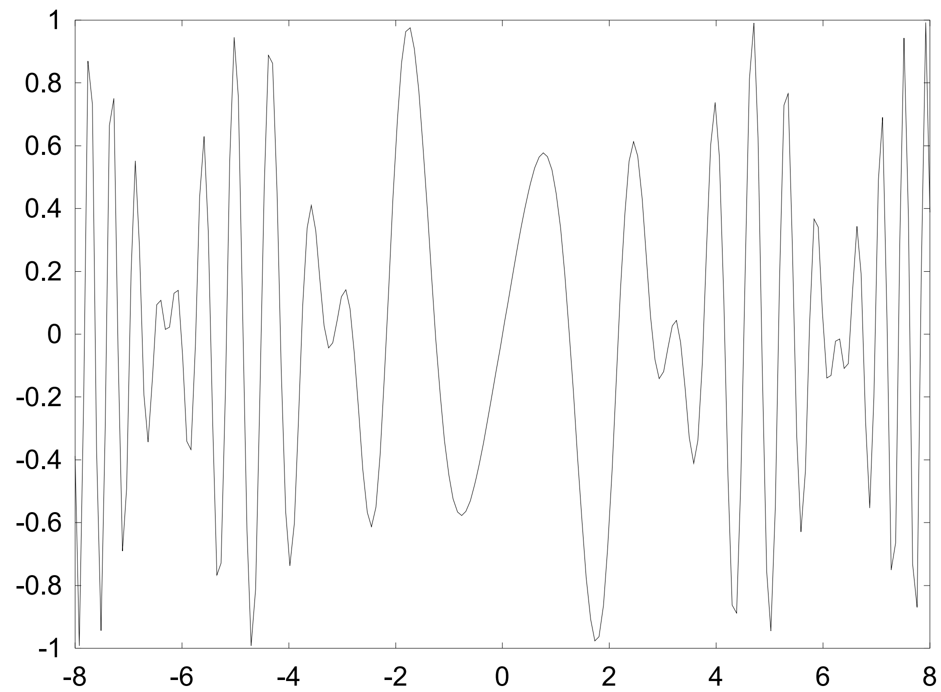
（方法2）最適解の近くでは x^k があまり変化しない

⇒ $\|x^{k+1} - x^k\|$ の値が十分小さくなったら終了

最適解の判定（つづき）

•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（凸関数）では

極小解⇔最小解

定理：ある仮定の下で，最急降下法の求める点列は
停留点に収束する

関数のヘッセ行列

- 定義：n変数関数 f のヘッセ行列 $\nabla^2 f(x)$

  ⇔f の2次偏微分係数を要素とするｎｘｎ 行列

$$\nabla^2 f(x_1, x_2, \dots, x_n) = \begin{bmatrix} \frac{\partial^2 f(x)}{\partial x_1^2} & \frac{\partial^2 f(x)}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f(x)}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f(x)}{\partial x_2 \partial x_1} & \frac{\partial^2 f(x)}{\partial x_2^2} & \cdots & \frac{\partial^2 f(x)}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f(x)}{\partial x_n \partial x_1} & \frac{\partial^2 f(x)}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f(x)}{\partial x_n^2} \end{bmatrix}$$

- f が1変数関数のときは，ヘッセ行列は2階微分と同じ

$$\nabla^2 f(x) = f''(x)$$

ヘッセ行列の例

例 :

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

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

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

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

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

二次のテイラー展開

任意の関数 f はベクトル $a \in \mathbb{R}^n$ を使って
次の形に表現できる

$$f(x) = f(a) + \nabla f(a)^T (x-a) + \frac{1}{2}(x-a)^T \mathrm{H} f(a)(x-a) + \psi(x-a)$$

関数$\psi(x-a) = \psi(x_1 - a_1, \ldots, x_n - a_n)$は
$x_1 - a_1, \ldots, x_n - a_n$ に関する3次以上の項から
構成される n 変数多項式関数
（定数項，一次の項，二次の項は含まれない）

関数$f(x)$の$x = a$における
二次のテイラー展開

二次のテイラー近似

関数$f(x)$の$x = a$における二次のテイラー展開

$$f(x) = f(a) + \nabla f(a)^T(x-a) + \frac{1}{2}(x-a)^T \mathrm{H}f(a)(x-a) + \psi(x-a)$$

$x \simeq a$ のとき, $\psi(x-a)$の値は他の項に比べて
十分小さい（0に近い） ➔無視できる

$$\tilde{f}(x) = f(a) + \nabla f(a)^T(x-a) + \frac{1}{2}(x-a)^T \mathrm{H}f(a)(x-a)$$

関数$f(x)$の$x = a$における
二次のテイラー近似

- 二次関数
  $\nabla \tilde{f}(a) = \nabla f(a), \mathrm{H}\tilde{f}(a) = \mathrm{H}f(a)$
- $x \simeq a$のとき $\tilde{f}(x) \simeq f(x)$,
  とくに $\tilde{f}(a) = f(a)$

二次のテイラー近似の例

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

$$
\begin{aligned}
f_1(x) &= a^2 + (2a)(x-a) + \frac{1}{2}\cdot 2(x-a)^2 + \psi(x-a) \\
&= x^2 + \psi(x-a)
\end{aligned}
$$

つまり, $\psi(x=a) = 0$であり,

f_1の二次のテイラー近似 $=$ f_1そのもの

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

の二次のテイラー近似は f に一致

V: $n \times n$行列
c: n 次元ベクトル
c_0: 定数

二次のテイラー近似の例

例２：　$f_2(x) = \log x \;\Rightarrow f'(x) = \frac{1}{x}, f''(x) = -\frac{1}{x^2}$

f_2 の x=1 における二次のテイラー展開

$$f_2(x) = \log 1 + \frac{1}{1}(x-1) - \frac{1}{1^2}(x-1)^2 + \psi(x-a)$$

なお，$\psi(x-a) = \frac{(x-1)^3}{3} - \frac{(x-1)^4}{4} + \frac{(x-1)^5}{5} - \cdots$

二次のテイラー近似の例

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

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

a = -1 のとき

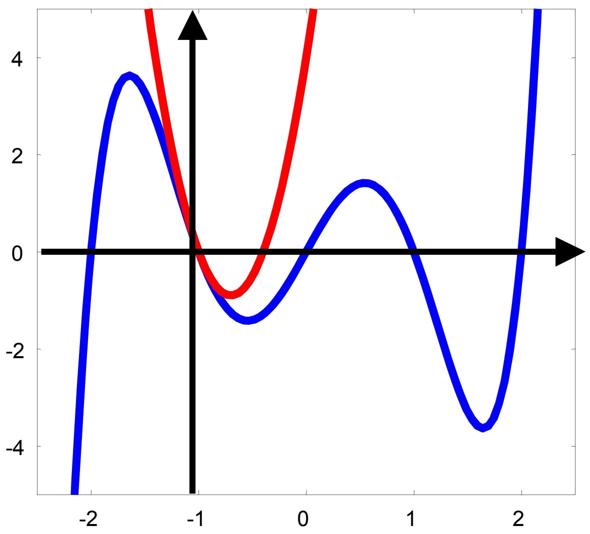
$0 - 6(x+1) + 5(x+1)^2$

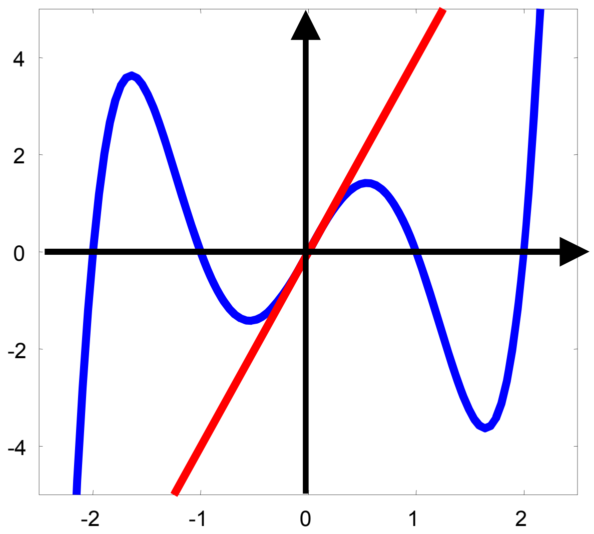
a = 0 のとき

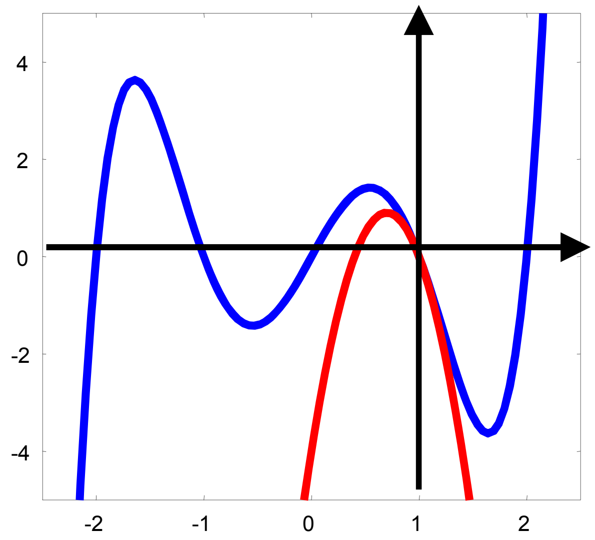
$0 + 4x + 0x^2$

a = 1 のとき

$0 - 6(x-1) - 5(x-1)^2$







行列の正定値性、半正定値性

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

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、

板書で証明

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

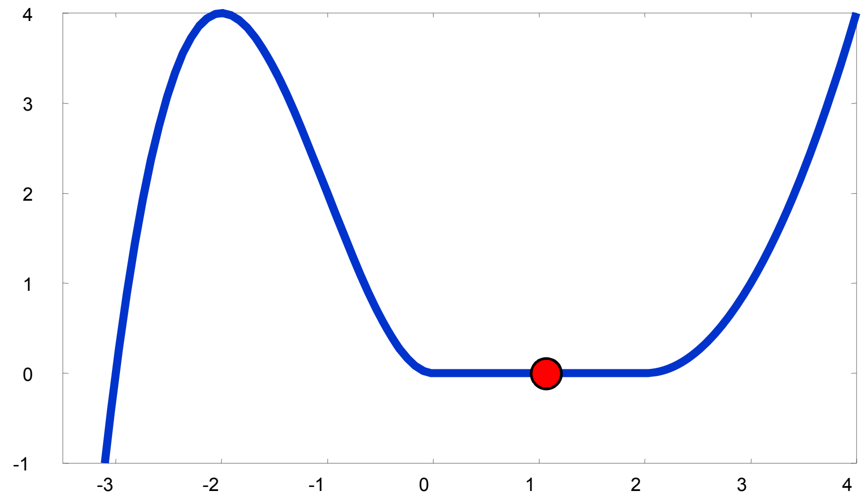
2次の最適性条件（必要条件）

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）

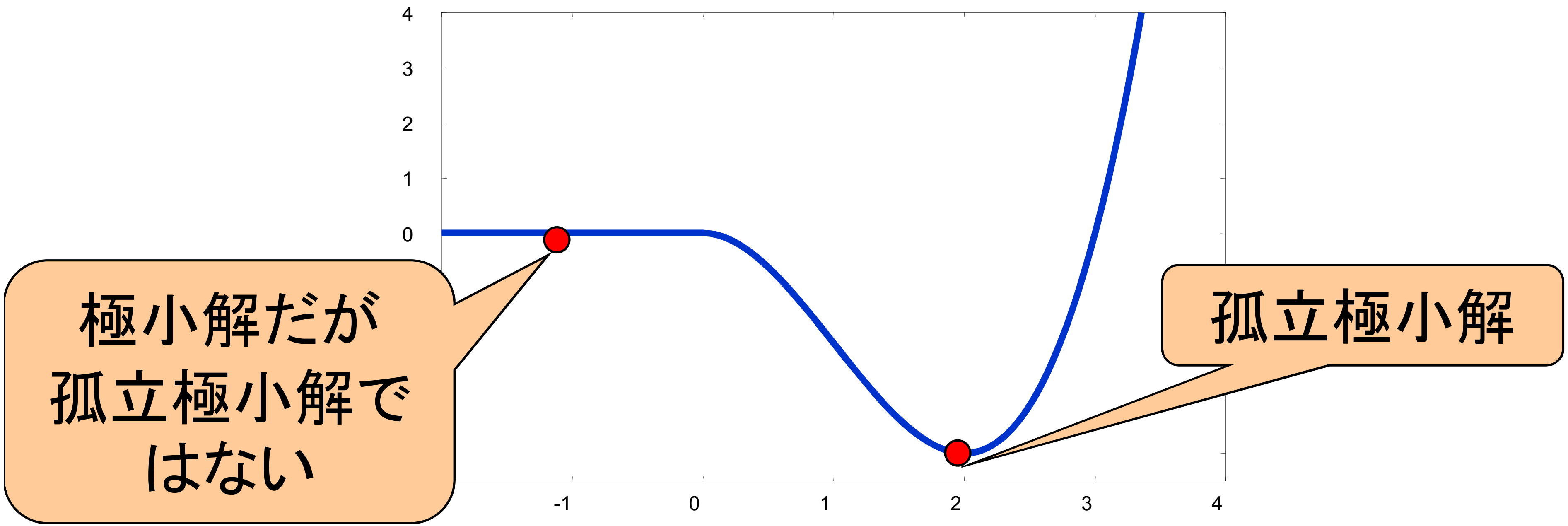
定理（2次の十分条件）：

x^* は停留点， $Hf(x^*)$ は正定値

$\Rightarrow$ x^*: 制約なし問題の（孤立）極小解

定義：x^* は孤立極小解

$\Leftrightarrow$ 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）の例

定理：Hf(x*) は正定値　⇒　（孤立）極小解

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

勾配を計算：$f'(x) = (x+2)x^2(x-2)(x-3)$

⇒ 停留点はx = -2, 0, 2, 3

2階微分を計算：

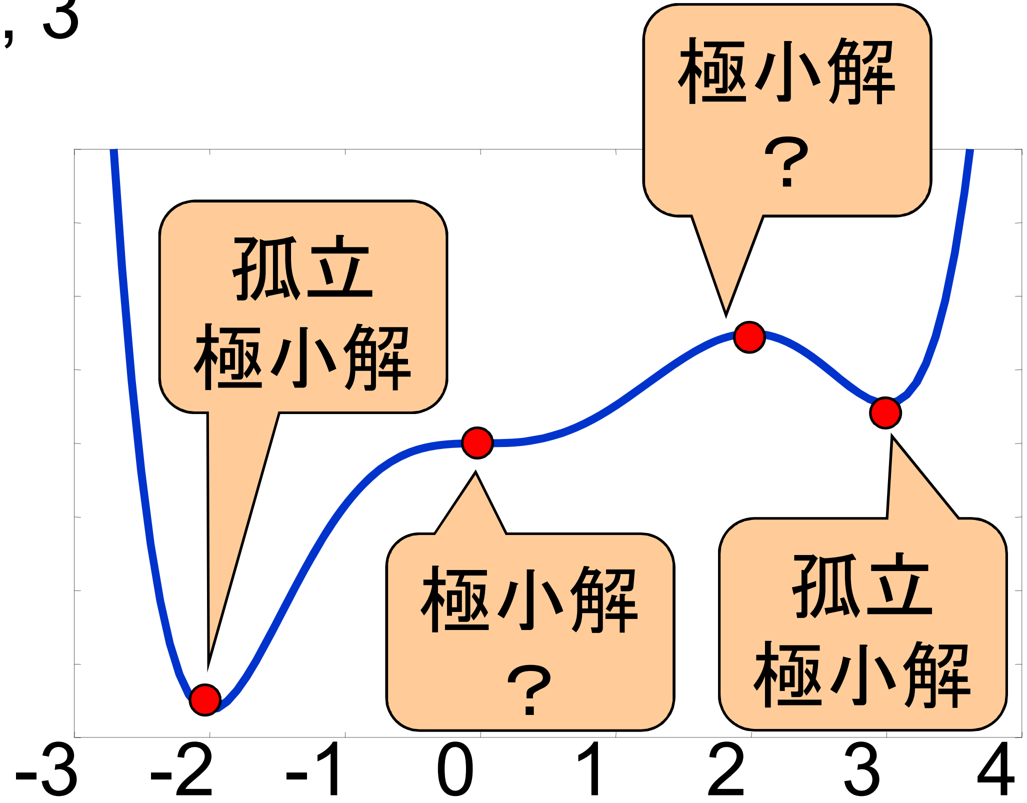
$$f''(x) = 5x^4 - 12x^3 - 12x^2 + 24x$$

⇒ 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)の例

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

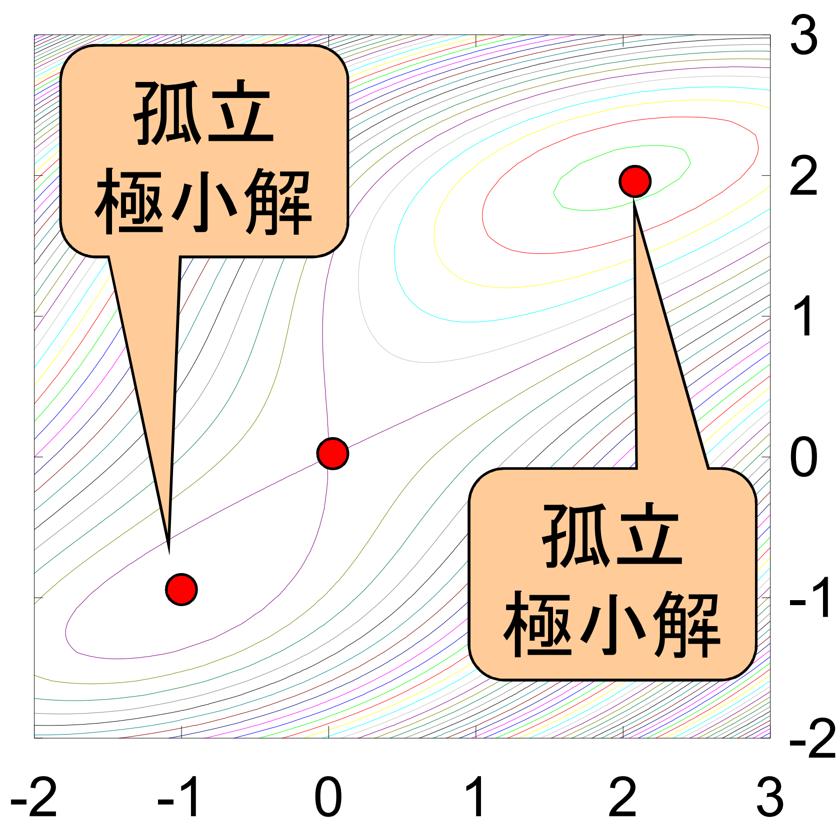
例2 $f(\boldsymbol{x}) = x_1^2 - 2x_1x_2 + \frac{1}{4}x_2^4 - \frac{1}{3}x_2^3$

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

停留点は(0,0), (-1, -1), (2, 2)

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

(-1, -1), (2, 2) は孤立極小解

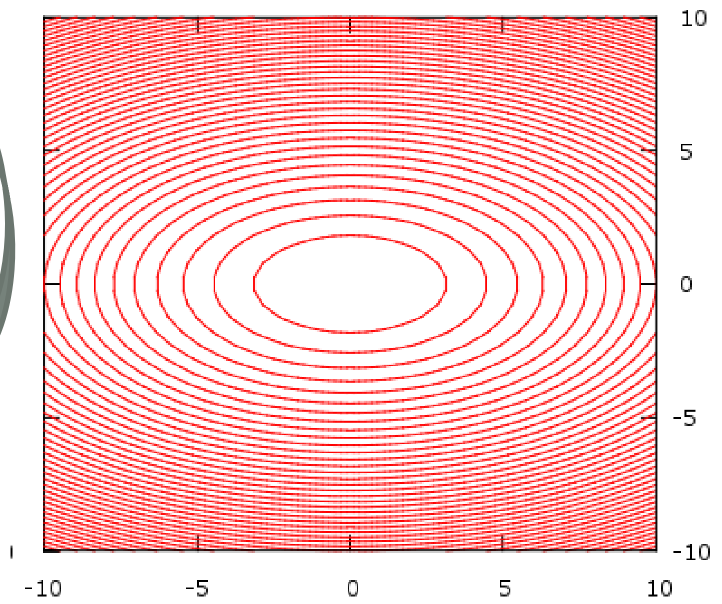


2次の最適性条件の例

例3：$f(x_1, x_2) = x_1^2 + 3x_2^2$

- $\nabla f(x_1, x_2) = \begin{bmatrix} 2x_1 \\ 6x_2 \end{bmatrix}, \mathrm{H}f(x_1, x_2) = \begin{bmatrix} 2 & 0 \\ 0 & 6 \end{bmatrix}$
- $\nabla f(x_1, x_2)$がゼロベクトルとなるのは (0,0) のみ🡐停留点
- $\nabla^2 f(0,0) = \begin{bmatrix} 2 & 0 \\ 0 & 6 \end{bmatrix}$ は正定値行列

🡒(0, 0) は孤立極小解



任意の非ゼロベクトル(y_1, y_2)に対して

$$\begin{bmatrix} y_1 & y_2 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 6 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = 2y_1^2 + 6y_2^2 > 0$$

2次の最適性条件の例

例4 : $f(x_1, x_2) = x_1^2 + x_2^4$

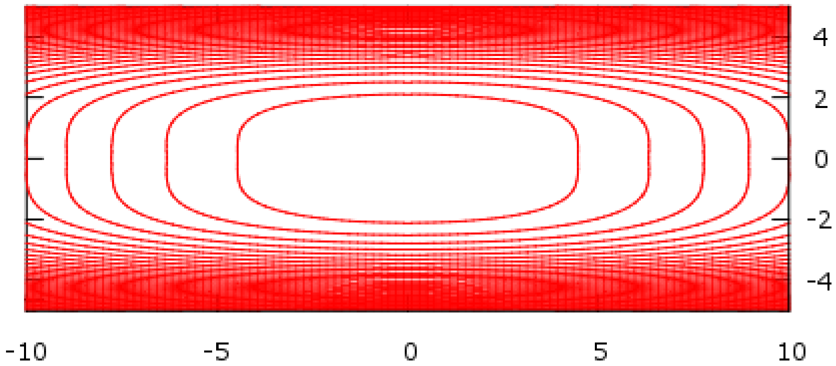
- $\nabla f(x_1, x_2) = \begin{bmatrix} 2x_1 \\ 4x_2^3 \end{bmatrix}, \nabla^2 f(x_1, x_2) = \begin{bmatrix} 2 & 0 \\ 0 & 12x_2^2 \end{bmatrix}$
- $\nabla f(x_1, x_2)$がゼロベクトルとなるのは (0,0) のみ（実は最適解）
- $\nabla^2 f(0,0) = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}$ は半正定値だが，正定値ではない

➔(0, 0) が極小解かどうかは，ヘッセ行列を使って判定できない（実際には極小解）

任意のベクトル(y_1, y_2)に対して

$$\begin{bmatrix} y_1 & y_2 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = 2y_1^2 \geq 0$$

$y_1 = 0$のときは$y_2 \neq 0$でも値は0



極大解に関する性質

- $\boldsymbol{x}^*$ は関数 f の(孤立)極大解
  $\Leftrightarrow$ $\boldsymbol{x}^*$ 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 $-\mathrm{H}f(\boldsymbol{x})$

$\Rightarrow$ 極大解であるための条件

定理:

x^*: 制約なし問題の極大解 $\Rightarrow$ $-\mathrm{H}f(x^*)$ は半正定値

定理:

x^* は停留点, $-\mathrm{H}f(x^*)$ は正定値

$\Rightarrow$ x^*: 制約なし問題の(孤立)極大解

制約なし問題の解法2:ニュートン法

定義:2次関数 $f(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^T V \boldsymbol{x} + \boldsymbol{c}^T \boldsymbol{x} + c_0$

は狭義2次凸関数 $\Leftrightarrow$ V は正定値行列

ニュートン法のアイディア:
狭義2次凸関数の最適解は簡単に求められる!

$$\nabla f(\boldsymbol{x}) = V\boldsymbol{x} + \boldsymbol{c} \qquad \mathrm{H}\, f(\boldsymbol{x}) = V$$

停留点は $x^* = -V^{-1}c$ のみ, ヘッセ行列は V (正定値)

$\Rightarrow$ 2次の十分条件より x^* は最適解

制約なし問題の解法2:ニュートン法

ニュートン法のアイディア:

狭義2次凸関数の最適解は簡単に求められる!

ただし,一般の関数は狭義2次凸とは限らない

⇒ 元の関数 f の代わりに,二次のテイラー近似 $\tilde{f}$ を使う

$$\tilde{f}(x) = f(a) + \nabla f(a)^T(x-a) + (x-a)^T \mathrm{H} f(a)(x-a)$$

- ヘッセ行列 Hf(x) が正定値のとき,
  $\tilde{f}$ の最適解は $x = a - Hf(x)^{-1}\nabla f(x)$
- $\tilde{f}$ は f の良い近似

⇒ $a - Hf(x)^{-1}\nabla f(x)$ は
f の最適解のより良い近似解と期待できる

ニュートン法のアルゴリズム

現在の点 x を $x - Hf(x)^{-1}\nabla f(x)$ へ移動させることを繰り返す
($-Hf(x)^{-1}\nabla f(x)$ を, xにおけるニュートン方向と呼ぶ)

入力:関数 f とその勾配ベクトル∇f, ヘッセ行列$\mathrm{H}f$
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

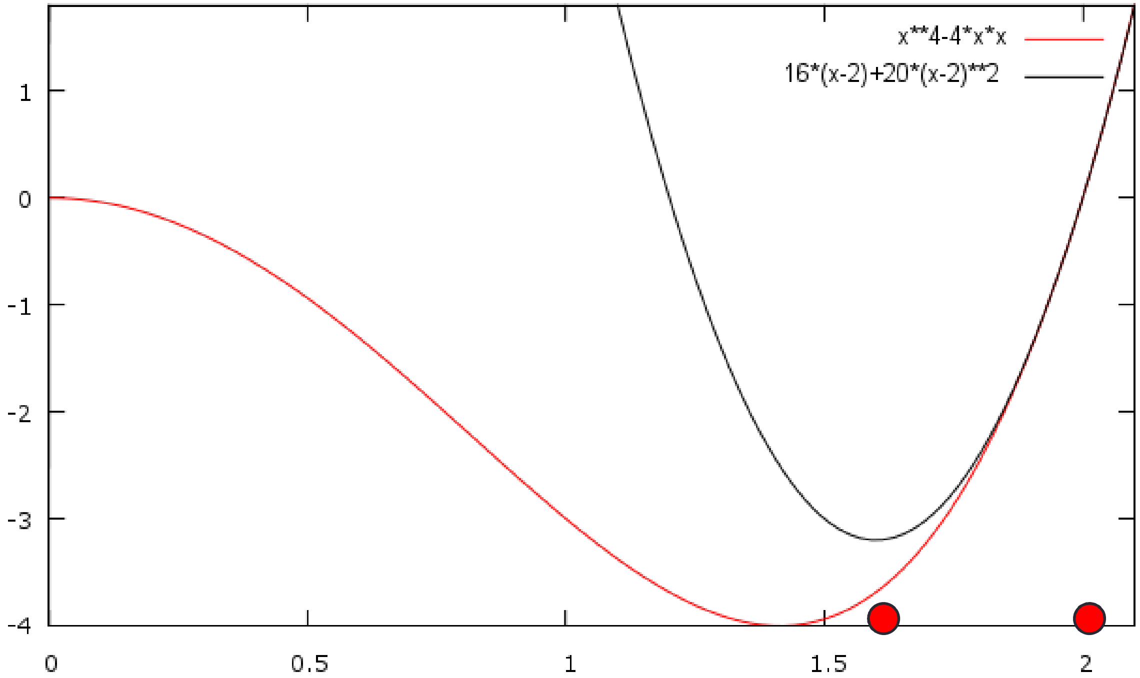
ステップ2:ニュートン方向 $-Hf(x^k)^{-1}\nabla f(x^k)$ を計算

ステップ3: $x^{k+1} = x^k - Hf(x^k)^{-1}\nabla f(x^k)$ とおく

ステップ4: $k = k + 1$として、ステップ1に戻る

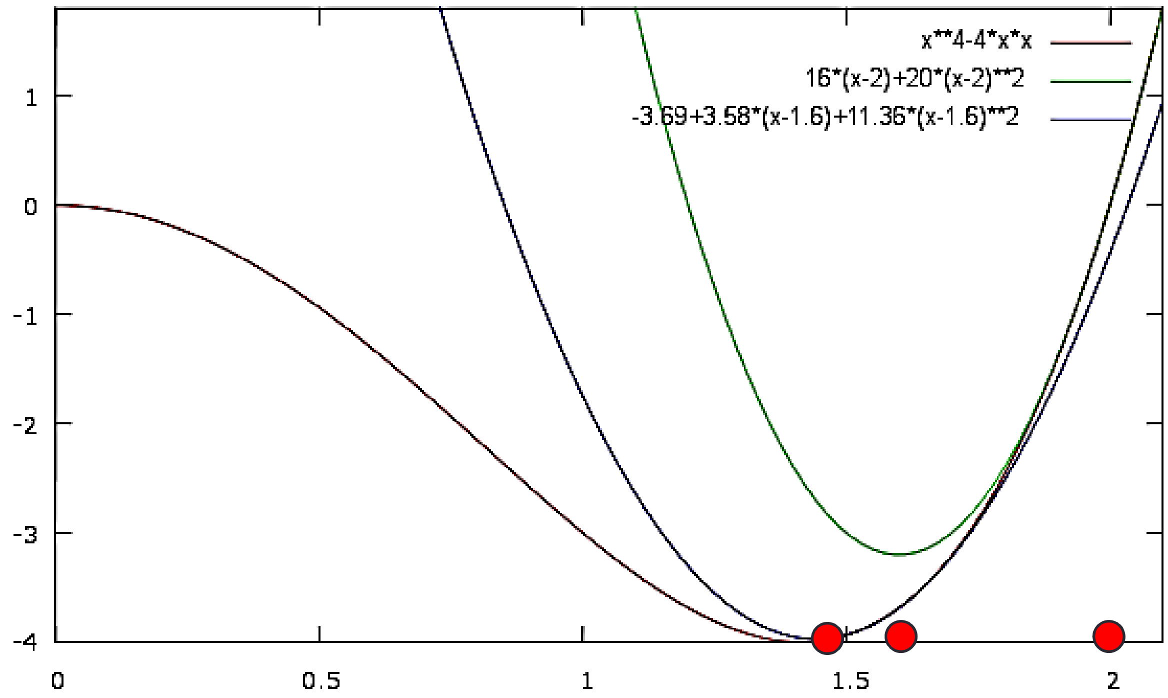
ニュートン法の実行例その1

- 一変数関数 $f(x) = x^4 - 4x^2$
- 初期点 $x^{(0)} = 2$
- テイラー近似は$\tilde{f}(x) = 16(x-2) + 20(x-2)^2$
- これが最小になるのは$x = 2 - 0.4 = 1.6$のとき
- $x^{(1)} \coloneqq 1.6$とおく



ニュートン法の実行例その1

- 一変数関数 $f(x) = x^4 - 4x^2$
- 点 $x^{(1)} = 1.6$
- テイラー近似は$\tilde{f}(x) = -3.69 + 3.58(x - 1.6) + 11.36(x - 1.6)^2$
- これが最小になるのは$x = 1.6 - 0.11 = 1.49$のとき
- $x^{(2)} := 1.49$とおく



ニュートン法の例2

- 関数 $f(x) = (x_1 - 1)^2 + 10(x_1^2 - x_2)^2$ に適用
  - 初期解(0,0), 最適解は(1,1)
  - 6回の反復で最適解に到達
    - 最急降下法では100回反復後でも(0.91, 0.82)

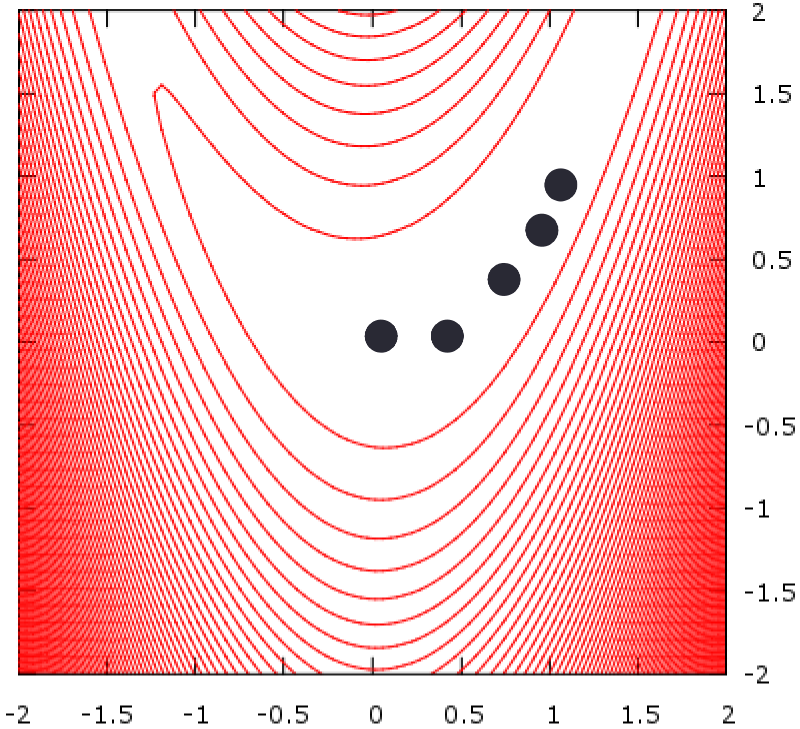


表 4.2　関数 (4.19) に対するニュートン法の計算結果

反復 k	$\boldsymbol{x}^{(k)}$	$f(\boldsymbol{x}^{(k)})$	$\lVert\nabla f(\boldsymbol{x}^{(k)})\rVert$
0	$(0.00000, 0.00000)$	0.10000×10^{1}	0.20000×10^{1}
1	$(0.32341, 0.00000)$	0.56717×10^{0}	0.20919×10^{1}
2	$(0.73455, 0.46247)$	0.12990×10^{0}	0.23209×10^{1}
3	$(0.91297, 0.85632)$	0.12775×10^{-1}	0.11054×10^{1}
4	$(1.00450, 1.01041)$	0.39429×10^{-4}	0.54177×10^{-1}
5	$(0.99997, 0.99995)$	0.16624×10^{-8}	0.46482×10^{-3}
6	$(1.00000, 1.00000)$	0.39340×10^{-17}	0.17062×10^{-7}

レポート問題（今回で最後）

問 1：関数 f_1, f_2 に対し，$(x_1, x_2) = (1,2)$における二次のテイラー展開を求めなさい．

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

$$f(x, y) = x^2 - xy + 2y^2 - 2x + e^{x+y}$$

問 2：関数$f(x, y) = \frac{1}{4}x^4 - \frac{1}{2}x^2 + \frac{1}{3}y^3 + 2y^2$ について考える

(a) 勾配ベクトルとヘッセ行列を計算せよ．

(b) すべての停留点（勾配ベクトルがゼロの点）を求めよ．

さらに，2次の最適性条件（十分条件）を用いて，極小解を求めよ．

問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 $\Leftrightarrow a_{11} \geqq 0,\ a_{22} \geqq 0,\ a_{11}a_{22} - a_{12}^2 \geqq 0$