

数理計画法 (数理最適化) 第12回

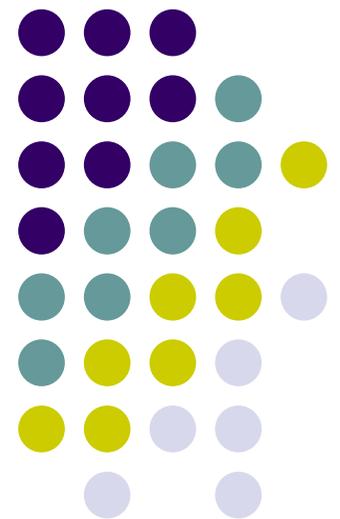
非線形計画

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29日(木) 13:00~14:30
- **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
 - ネットワーク最適化, 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イルス感染時は無理に来ないこと**
 - 事前にメール連絡の上, 後日医師の診断書を持参すれば, 再試験を実施します

関数のヘッセ行列

- 定義: n 変数関数 f のヘッセ行列 $Hf(x)$

$\leftrightarrow f$ の2次偏微分係数を要素とする $n \times n$ 行列

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

- f が1変数関数のときは, ヘッセ行列は2階微分と同じ
 $Hf(x) = f''(x)$

ヘッセ行列の例



例:

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

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

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

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

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



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

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

関数 $\psi(x - a) = \psi(x_1 - a_1, \dots, x_n - a_n)$ は
 $x_1 - a_1, \dots, x_n - a_n$ 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**

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

$x \simeq a$ 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

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

関数 $f(x)$ の $x = a$ における
二次のテイラー近似

- 二次関数
 $\nabla \tilde{f}(a) = \nabla f(a), \text{H}\tilde{f}(a) = \text{H}f(a)$
- $x \simeq a$ 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: $f_1(x) = x^2$ $\nabla f_1(x) = 2x$ $H f_1(x) = 2$

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

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 $f(\mathbf{x}) = \frac{1}{2} \mathbf{x}^T V \mathbf{x} + \mathbf{c}^T \mathbf{x} + c_0$

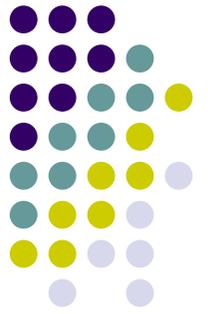
の二次のテイラー近似は f に一致

V : $n \times n$ 行列

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

二次のテイラー近似の例



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

$$\nabla f(x) = 5x^4 - 15x^2 + 4$$

$$H f(x) = 20x^3 - 30x$$

$a = -1$ のとき

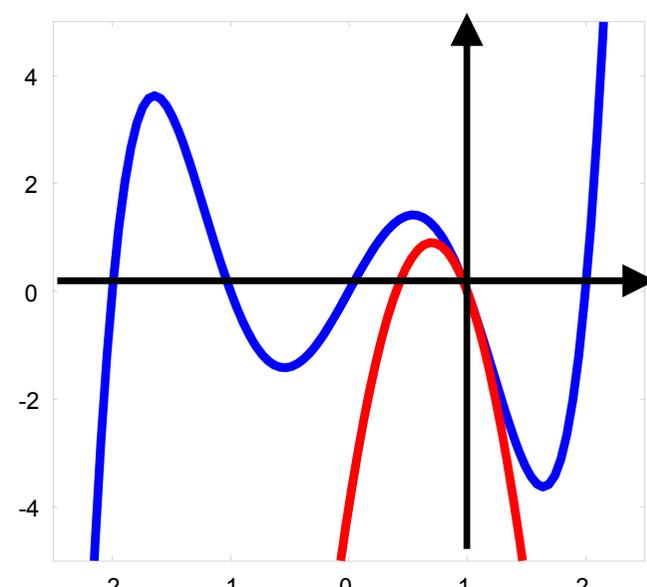
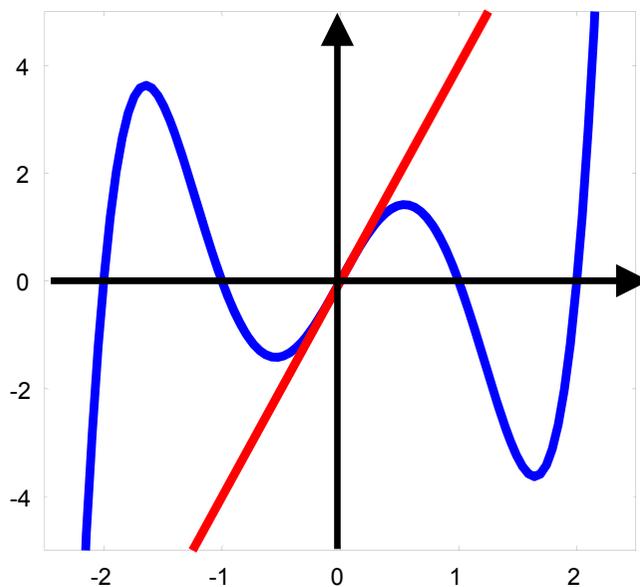
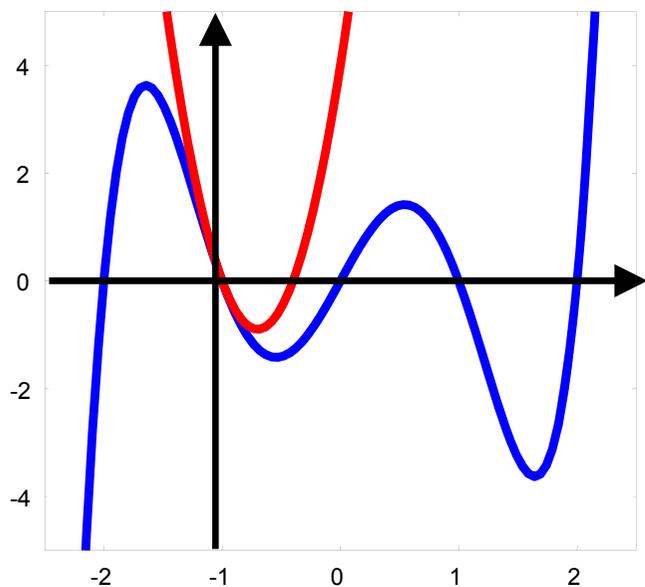
$$0 - 6(x+1) + 5(x+1)^2$$

$a = 0$ のとき

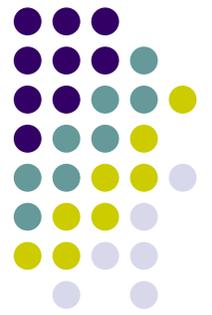
$$0 + 4x + 0x^2$$

$a = 1$ のとき

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



極小解, 極大解の判定方法



- 一変数関数 f の場合
 - 極小, 極大ならば傾き(一回微分) $f'(x)$ が 0
 - 極小, 極大は二回微分 $f''(x)$ を使って判定
 - $f''(x) > 0$ ならば極小, $f''(x) < 0$ ならば極大
 - $f''(x) = 0$ の場合は不明
- 多変数関数の場合
 - 極小, 極大ならば勾配ベクトル $f'(x)$ がゼロベクトル
 - 極小, 極大はヘッセ行列 $Hf(x)$ を使って判定
 - どうやって判定? --- 行列の(半)正定値性を使う

正定値(半正定値)・・・行列が「正(非負)」

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

$$\Leftrightarrow \text{任意のベクトル } y \text{ に対して } y^T A y \geq 0$$

定義: 正方行列 A は正定値

$$\Leftrightarrow \text{任意の非零ベクトル } y \text{ に対して } y^T A y > 0$$

※ A が 1×1 行列のとき、

$$A \text{ は半正定値} \Leftrightarrow a_{11} \geq 0, \quad A \text{ は正定値} \Leftrightarrow a_{11} > 0$$

行列の正定値性、半正定値性



定義: 正方行列 A は半正定値

$$\Leftrightarrow \text{任意のベクトル } y \text{ に対して } y^T A y \geq 0$$

定義: 正方行列 A は正定値

$$\Leftrightarrow \text{任意の非零ベクトル } y \text{ に対して } y^T A y > 0$$

板書で証明

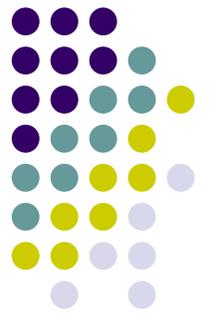
※ A が 2×2 行列のとき、

$$A \text{ は正定値} \Leftrightarrow a_{11} > 0, a_{22} > 0, a_{11}a_{22} - a_{12}^2 > 0$$

$$A \text{ は半正定値} \Leftrightarrow a_{11} \geq 0, a_{22} \geq 0, a_{11}a_{22} - a_{12}^2 \geq 0$$

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

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 \Rightarrow $Hf(x^*)$ は半正定値

例:

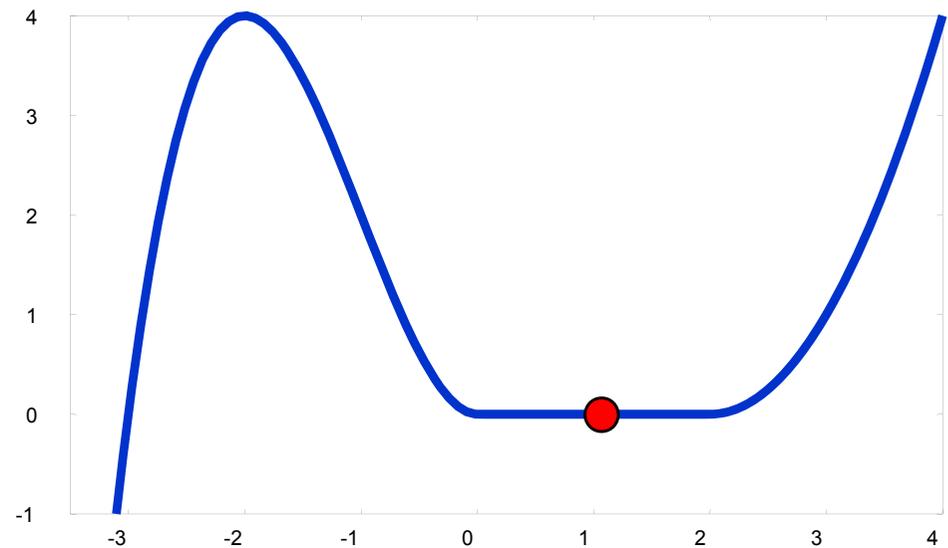
$x^* = 1$ は極小解

$0 \leq x \leq 2$ の範囲で $f(x) = 0$

$$\Rightarrow \nabla f(x^*) = f'(x^*) = 0$$

$$Hf(x^*) = f''(x^*) = 0$$

半正定値



2次の最適性条件(十分条件)



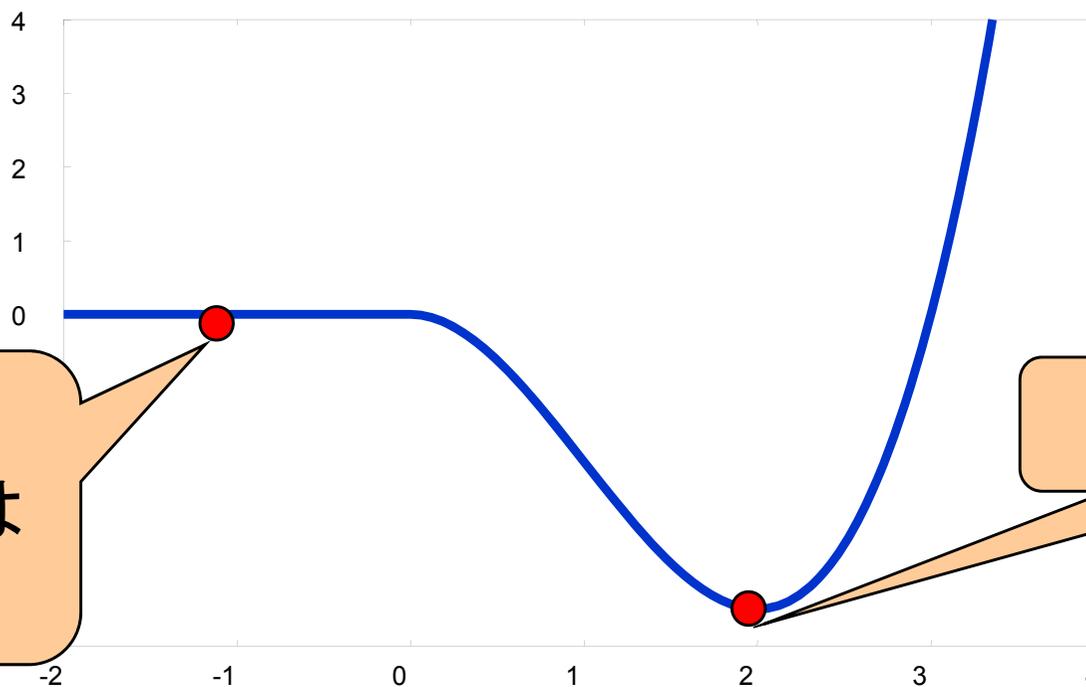
定理(2次の十分条件):

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



極小解だが
孤立極小解では
ない

孤立極小解

2次の最適性条件(十分条件)の例



定理: $Hf(x^*)$ は正定値 \Rightarrow (孤立)極小解

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

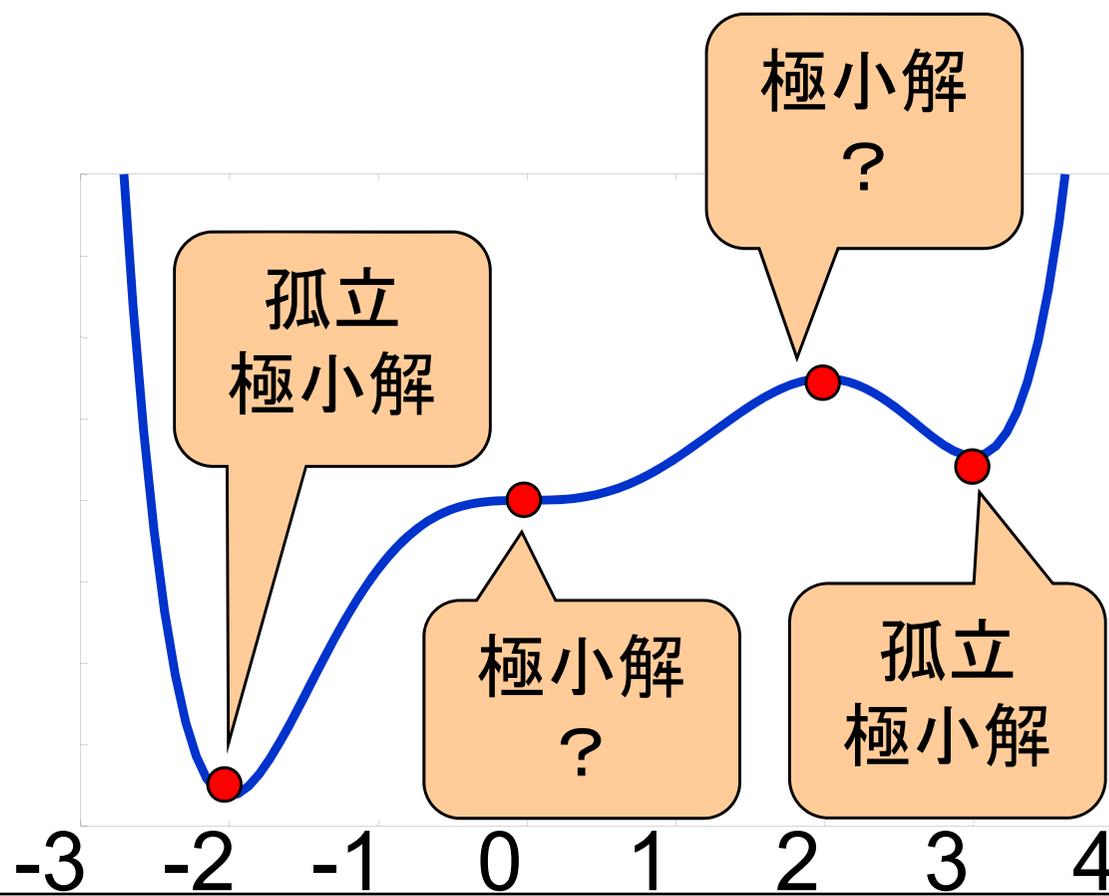
勾配を計算: $f'(x) = (x+2)x^2(x-2)(x-3)$

→ 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
 $\Rightarrow x^*$: (孤立)極小解

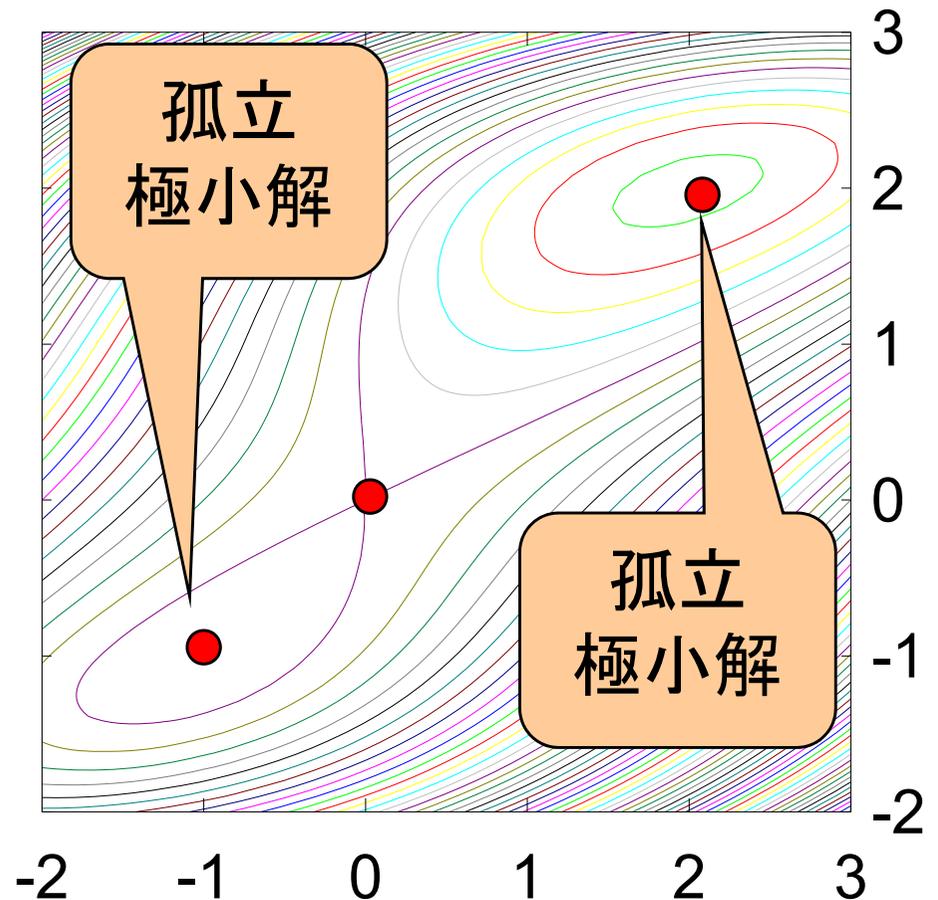
例2 $f(\mathbf{x}) = x_1^2 - 2x_1x_2 + \frac{1}{4}x_2^4 - \frac{1}{3}x_2^3$

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

➡ 停留点は $(0,0)$, $(-1, -1)$, $(2, 2)$

$$Hf(\mathbf{x}) = \begin{pmatrix} 2 & -2 \\ -2 & 3x_2^2 - 2x_2 \end{pmatrix}$$

➡ $(-1, -1)$, $(2, 2)$ は孤立極小解



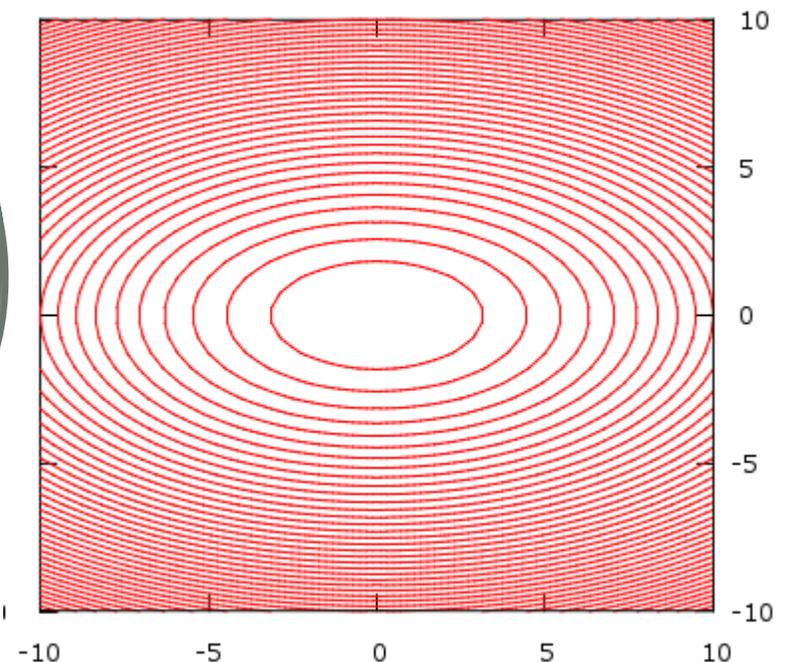
2次の最適性条件の例

例3: $f(x_1, x_2) = x_1^2 + 3x_2^2$

- $\nabla f(x_1, x_2) = \begin{bmatrix} 2x_1 \\ 6x_2 \end{bmatrix}$, $Hf(x_1, x_2) = \begin{bmatrix} 2 & 0 \\ 0 & 6 \end{bmatrix}$
- $\nabla f(x_1, x_2)$ 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

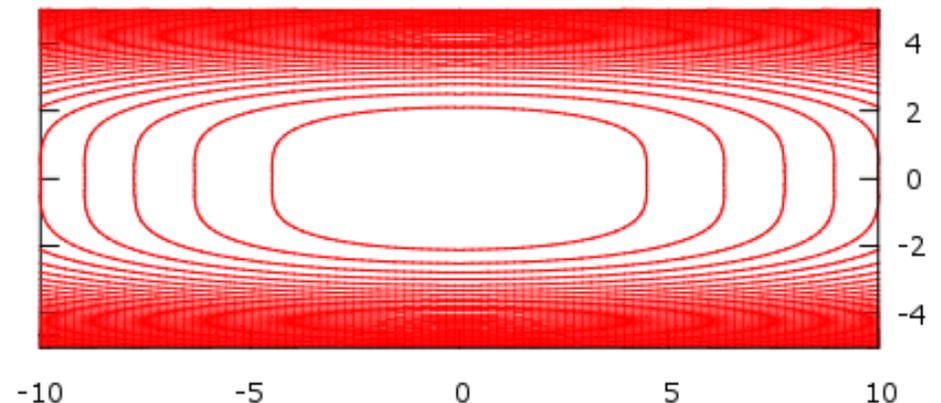
例4: $f(x_1, x_2) = x_1^2 + x_2^4$

- $\nabla f(x_1, x_2) = \begin{bmatrix} 2x_1 \\ 4x_2^3 \end{bmatrix}$, $\nabla^2 f(x_1, x_2) = \begin{bmatrix} 2 & 0 \\ 0 & 12x_2^2 \end{bmatrix}$
- $\nabla f(x_1, x_2)$ 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

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

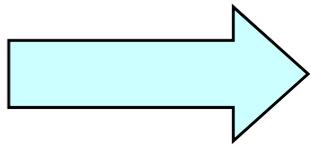
$y_1 = 0$ のときは $y_2 \neq 0$ でも値は0



極大解に関する性質



- 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 $\Rightarrow -Hf(x^*)$ は半正定値

定理:

x^* は停留点, $-Hf(x^*)$ 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



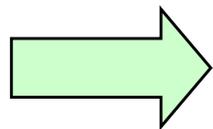
定義: 2次関数 $f(\mathbf{x}) = \frac{1}{2} \mathbf{x}^T V \mathbf{x} + \mathbf{c}^T \mathbf{x} + c_0$
は狭義2次凸関数 $\Leftrightarrow V$ は正定値行列

ニュートン法のアイデア:

狭義2次凸関数の最適解は簡単に求められる!

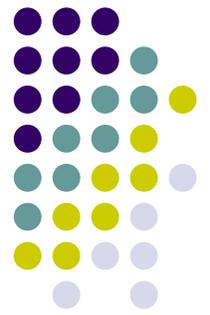
$$\nabla f(\mathbf{x}) = V \mathbf{x} + \mathbf{c} \quad H f(\mathbf{x}) = V$$

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



2次の十分条件より \mathbf{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

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

- 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 Hf
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

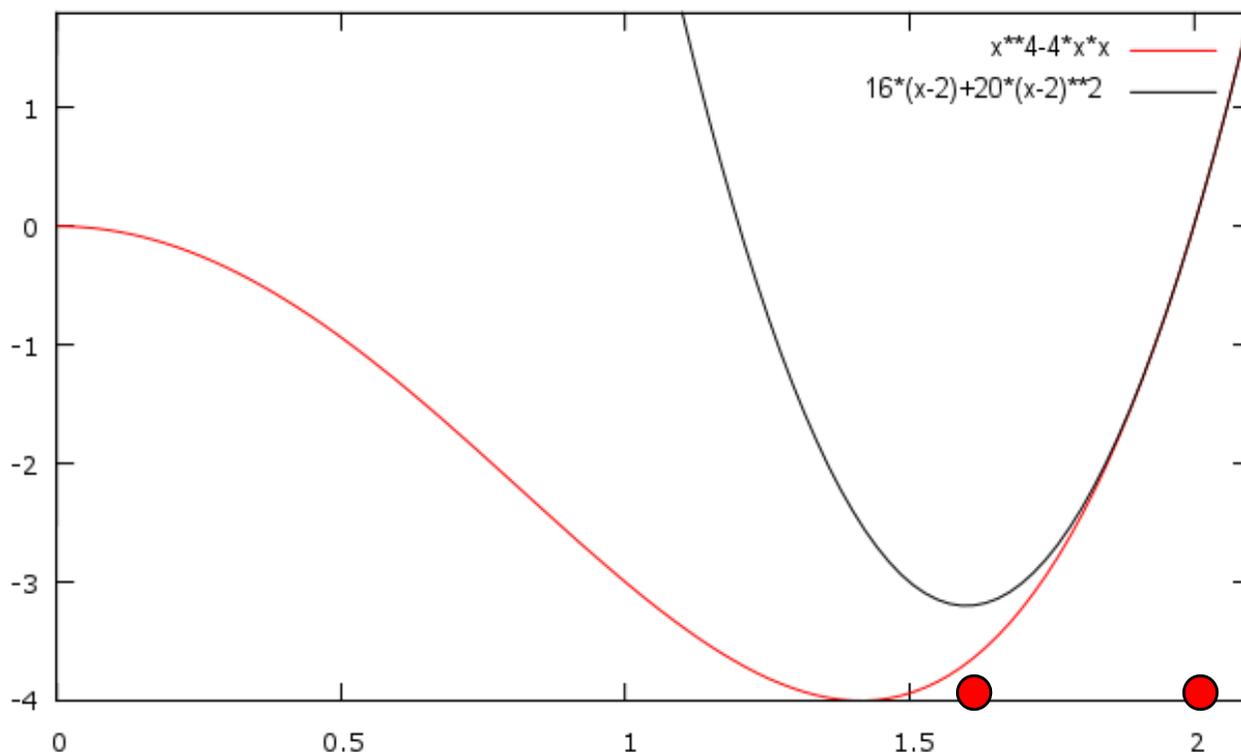
ステップ3: $x^{k+1} = x^k - Hf(x^k)^{-1}\nabla f(x^k)$ とおく

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

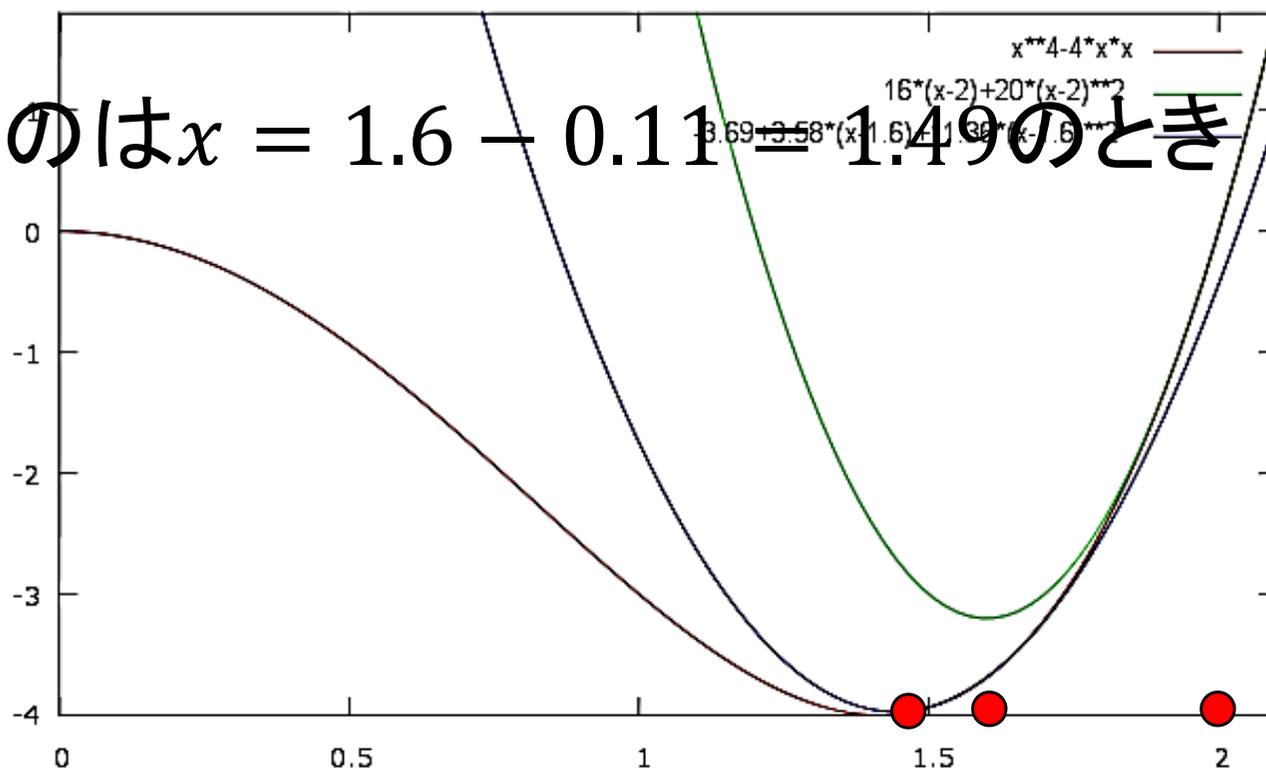
- 一変数関数 $f(x) = x^4 - 4x^2$
- 初期点 $x^{(0)} = 2$
- テイラー近似は $\tilde{f}(x) = 16(x - 2) + 20(x - 2)^2$
- これが最小にな
- $x^{(1)} := 1.6$ とおく





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



ニュートン法の特徴 [p.107]



長所:

- 最急降下法より**反復回数が少ない**
 - 狭義2次凸関数に対しては**一反復**で終了
- 直線探索が不要

短所:

- **ヘッセ行列の逆行列の計算が必要**
 - **ヘッセ行列の計算**ができないと破綻
 - **ヘッセ行列が正則**でないで破綻
- **ヘッセ行列が正定値でない場合には**
目的関数値が増加する可能性あり



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)

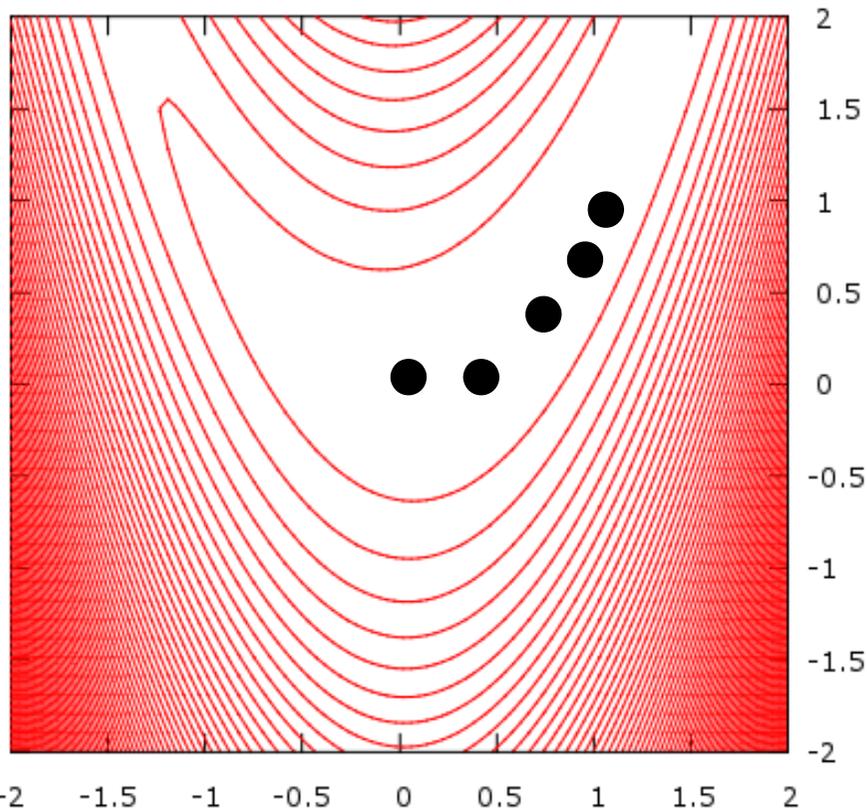


表 4.2 関数 (4.19) に対するニュートン法の計算結果

反復 k	$x^{(k)}$	$f(x^{(k)})$	$\ \nabla f(x^{(k)})\ $
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}

福島雅夫
「新版 数理計画入門」
(朝倉書店)より

ニュートン法の問題点



■ ヘッセ行列が**正則**でないと破綻

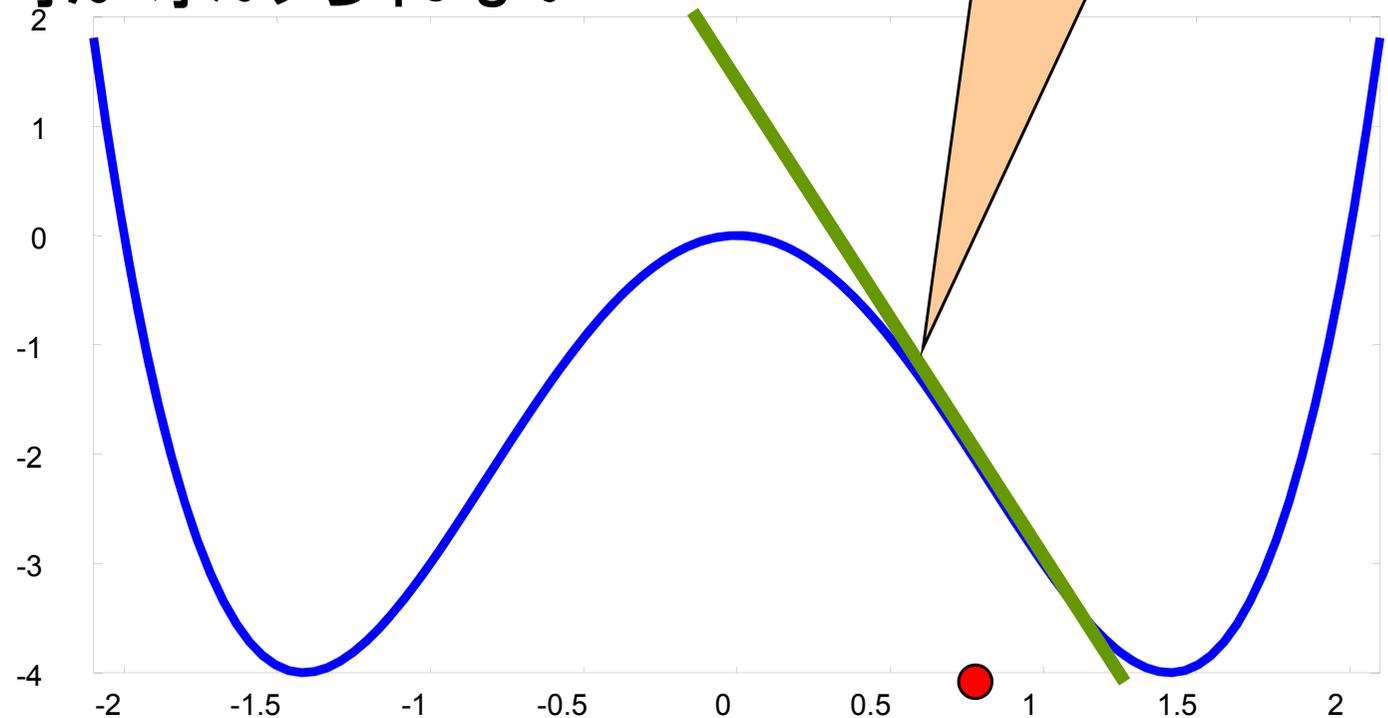
例1(続き): 一変数関数 $f(x) = x^4 - 4x^2$

初期点 $x = \sqrt{2/3}$ のとき

⇒ ヘッセ行列は $Hf(x) = 0$ (**正則でない**)

⇒ ニュートン方向が求められない

f を2次近似
すると直線
になる





ニュートン法の問題点

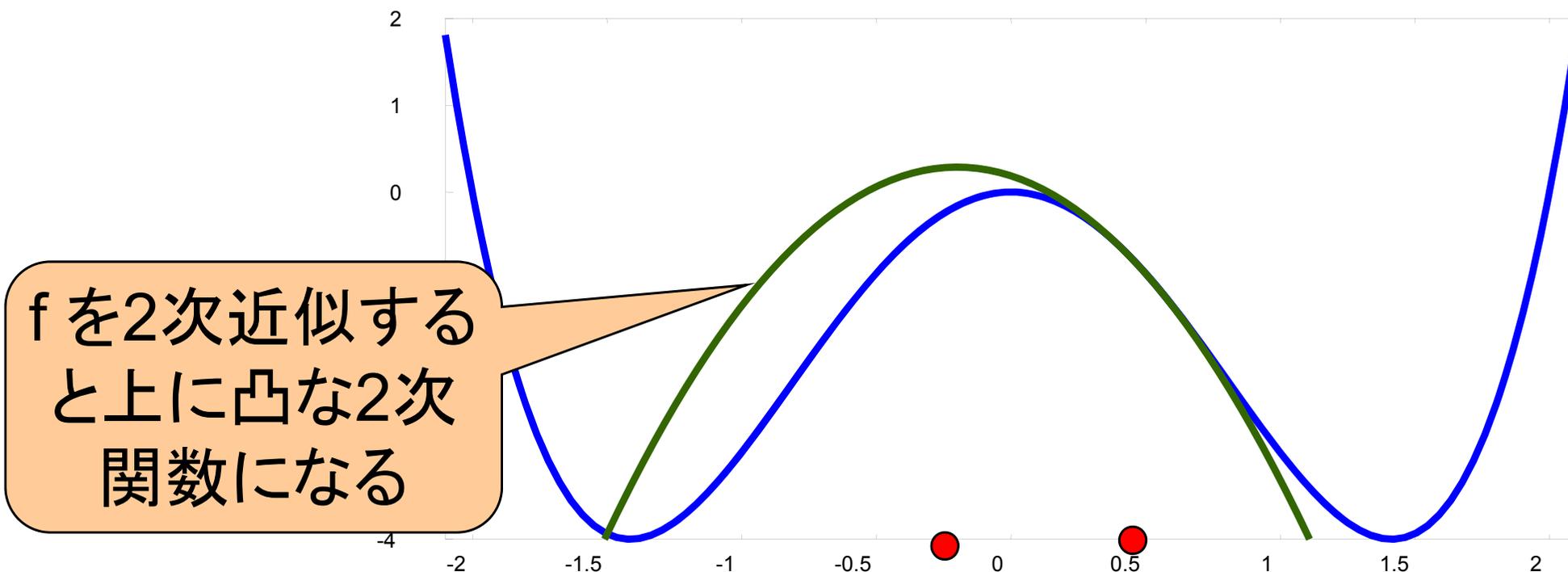
- ヘッセ行列が正定値でない場合には

目的関数値が増加する可能性あり

初期点 $x = 1/2$ のとき

⇒ ヘッセ行列は $Hf(x) = -5$ (正定値でない)

⇒ ニュートン方向に進むと関数値が増加する



レポート問題(今回で最後)



問1: 次の3つの関数に対し、 $(x_1, x_2) = (1, 2)$ における二次のテイラー近似を求めなさい(\log の底はeとする)

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

$$f_3(x, y) = x^2 - xy + 2y^2 - 2x + e^{x+y}$$

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。

$$A \text{ は半正定値} \Leftrightarrow a_{11} \geq 0, a_{22} \geq 0, a_{11}a_{22} - a_{12}^2 \geq 0$$