数理計画法　第12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　准教授)
shioura@dais.is.tohoku.ac.jp

期末試験について

- 日時：2月3日（木）午後1時より
- 受験資格者：以下の条件を満たす学生
  - 中間試験に合格している
  - ネットワーク最適化と非線形計画に関するレポートを一回以上提出
- 教科書等の持込は不可
- 座席はこちらで指定
- 試験内容：ネットワーク最適化，非線形計画の範囲（次回の内容まで）

  （詳しくはWeb上の過去問を参考にしてください）

成績評価の方法および基準

- 中間試験(50点)
- 期末試験(50点）
- 演習レポートの提出状況(最大20点程度)

  により評価
- 60点以上で合格
- 出席点は全く考慮しない
- レポート未提出者は試験の受験は不可

単位が不可となる条件

- 線形計画に関するレポートを一度も提出しない
- 中間試験の得点が25点未満
  - （レポートの提出状況が悪い場合）30点未満
- ネットワーク計画，非線形計画に関するレポートを一度も提出しない
- 期末試験の得点が25点未満
  - （レポートの提出状況が悪い場合）30点未満

以上の条件のいずれかに
該当する場合は単位が不可

復習：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

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

復習：ヘッセ行列とテイラー展開

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

関数 f(x) の x=a における二次のテイラー展開

$$f(x) = \underbrace{f(a) + \nabla f(a)^T(x-a) + \frac{1}{2}(x-a)^T \mathrm{H} f(a)(x-a)}_{\text{この部分が二次関数, f(x) を近似}} + \psi(x-a)$$

d = x - a とおくと（x の代わりに d が変数になる）

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

復習：極小解、極大解、鞍点 [p.99]

停留点 x* の分類

極小解(local minimum solution):

x* の付近だけに注目したとき、x* は最小

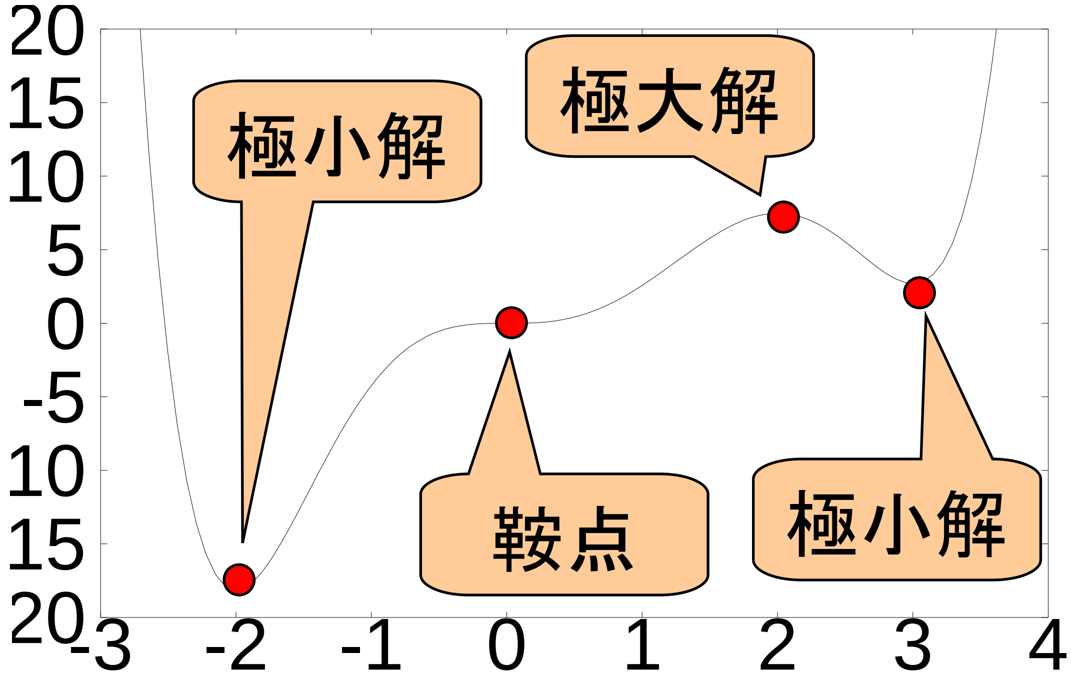
> あるδ＞0 が存在して、$\|x - x^*\| \leqq \delta$ を満たすすべての x に対して $f(x) \geqq f(x^*)$

極大解(local maximum solution):

x* の付近だけに注目したとき、x* は最大

鞍点(saddle point):

極小点でも極大点でもない停留点



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：

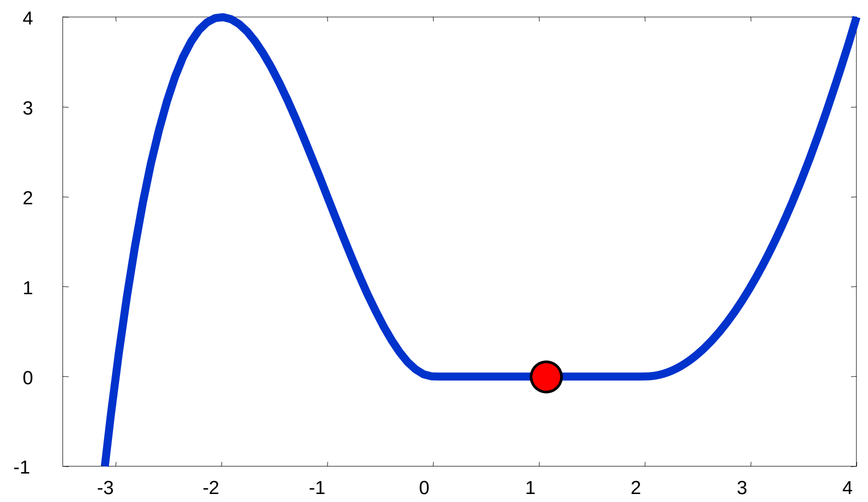
$x^* = 1$ は極小解

$0 \leqq x \leqq 2$ の範囲で $f(x) = 0$

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

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

半正定値



2次の最適性条件（十分条件） [p.100]

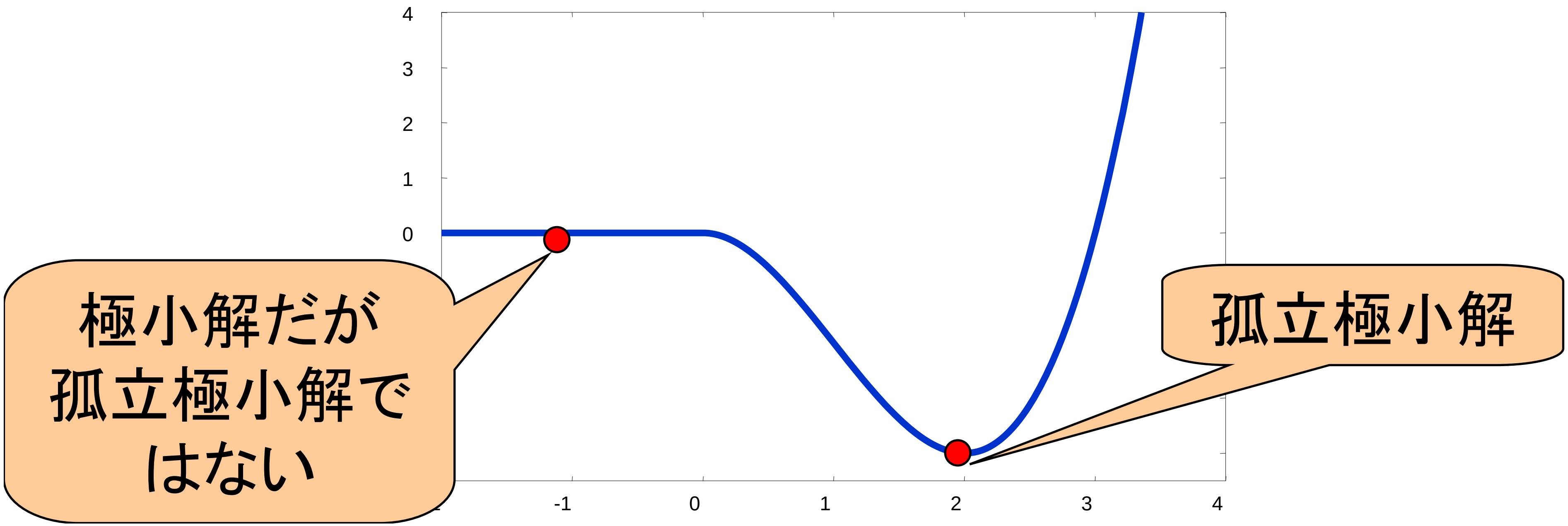
定理（2次の十分条件）:

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

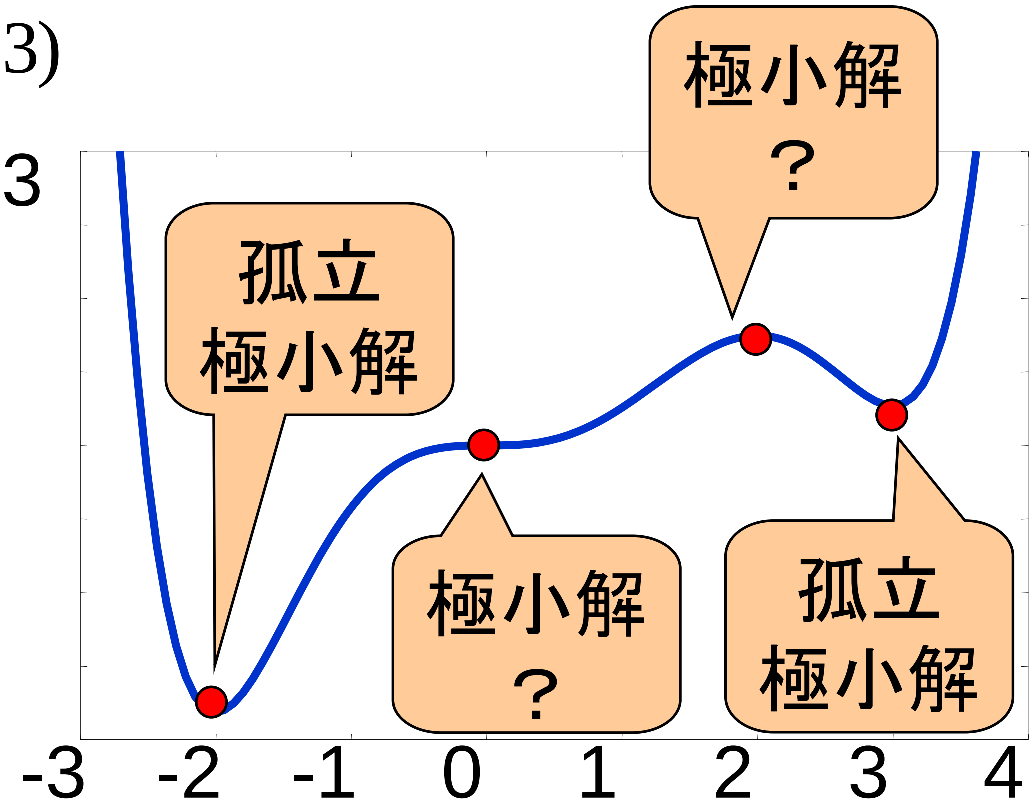
$$\mathrm{H}\, f(x) = 5x^4 - 12x^3 - 12x^2 + 24x$$

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

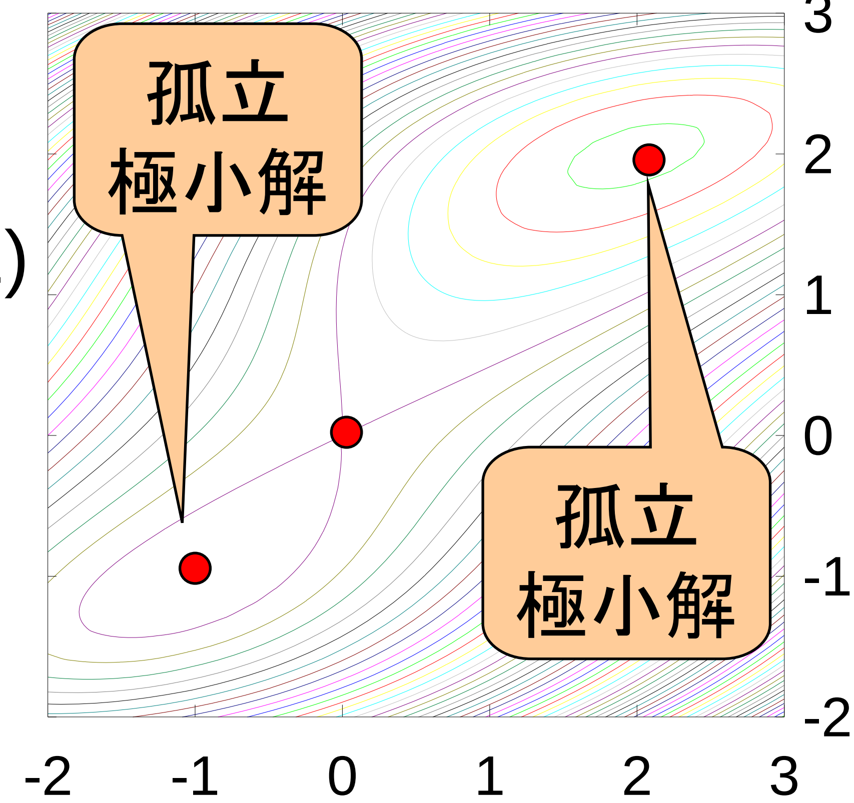
例2（教科書の例3．4）: $f(\boldsymbol{x}) = x_1^2 - 2x_1x_2 + \dfrac{1}{4}x_2^4 - \dfrac{1}{3}x_2^3$

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

⇒ 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
  $\Leftrightarrow$ 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 [p.105]

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 ⇔ 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)

⇒ 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 [p.105]

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 f は狭義2次凸とは限らない

⇒ f の代わりに,2次のテイラー近似 を使う

$$f(x+d) \simeq f(x) + \nabla f(x)^T(d) + \frac{1}{2}(d)^T \mathrm{H} f(x)(d)$$

ヘッセ行列 Hf(x) が正定値のとき

最適解は $d = -\mathrm{H}f(x)^{-1}\nabla f(x)$

ニュートン方向

⇒ x + d 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 [p.106]

現在の点 x を繰り返しニュートン方向へ移動、最適解に近づける

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

ニュートン法の例 [p.106]

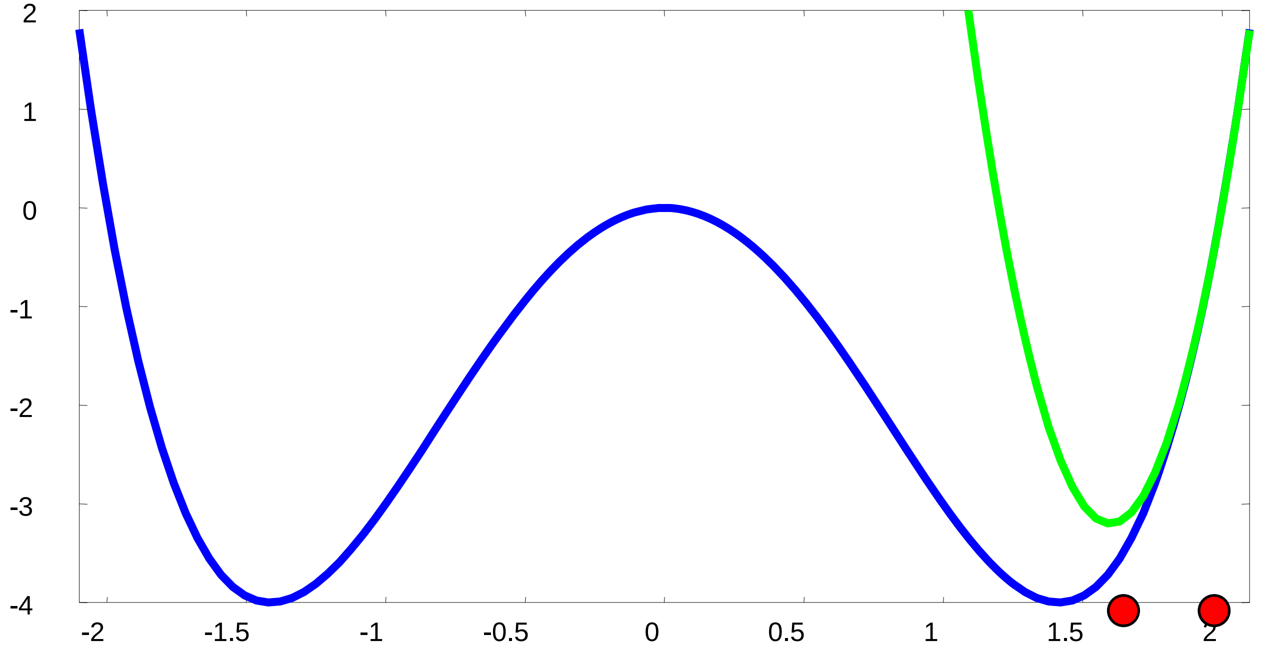
例1：一変数関数 $f(x) = x^4 - 4x^2$

初期点 $x = 2$ において f のテイラー近似を求める

⇒ $f(2+d) \fallingdotseq 0 + 16d + (40/2)d^2$

$d = -2/5$ のとき最小

⇒ 次の点は $x = 2 - 2/5 = 8/5$



ニュートン法の例 [p.106]

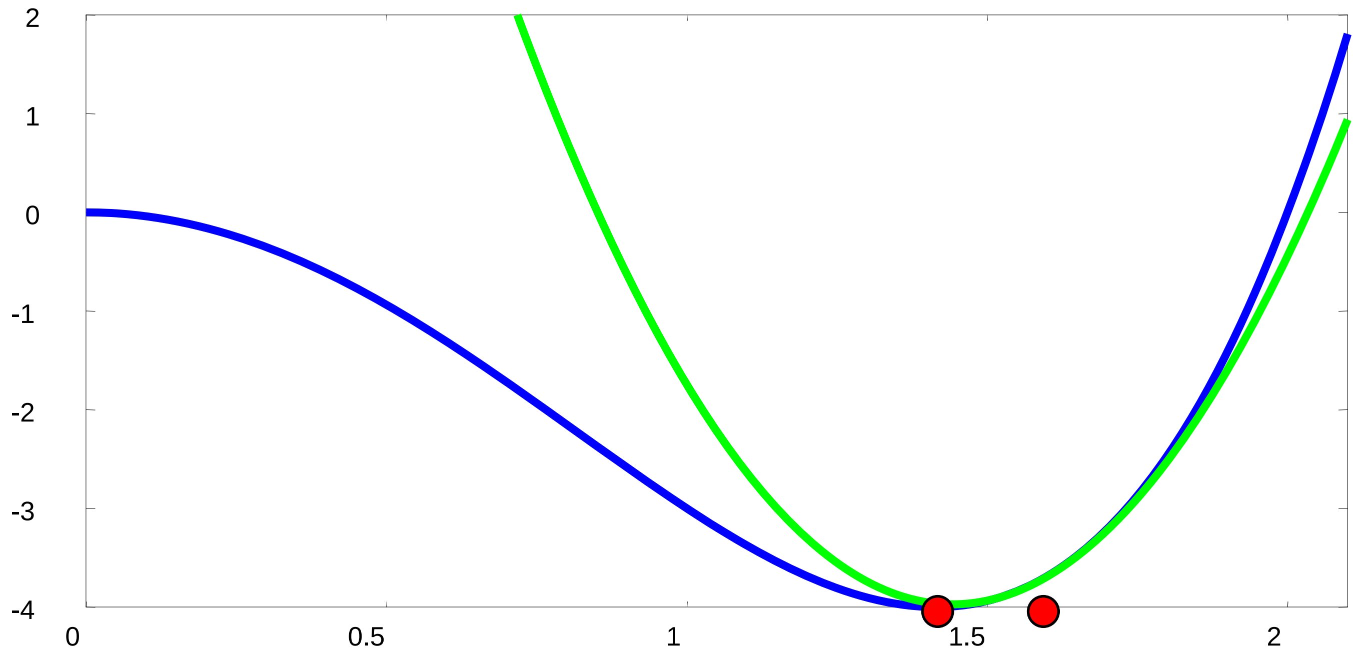
例1(続き) : 一変数関数 $f(x) = x^4 - 4x^2$

点 $x = 8/5$ において f のテイラー近似を求める

⇒ $f(8/5+d) \fallingdotseq -3.69 + 3.58d + 11.36d^2$

$d = -0.11$ のとき最小

⇒ 次の点は $x = 1.6 – 0.11 = 1.49$



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と破綻
- 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

ニュートン法の問題点 [p.107]

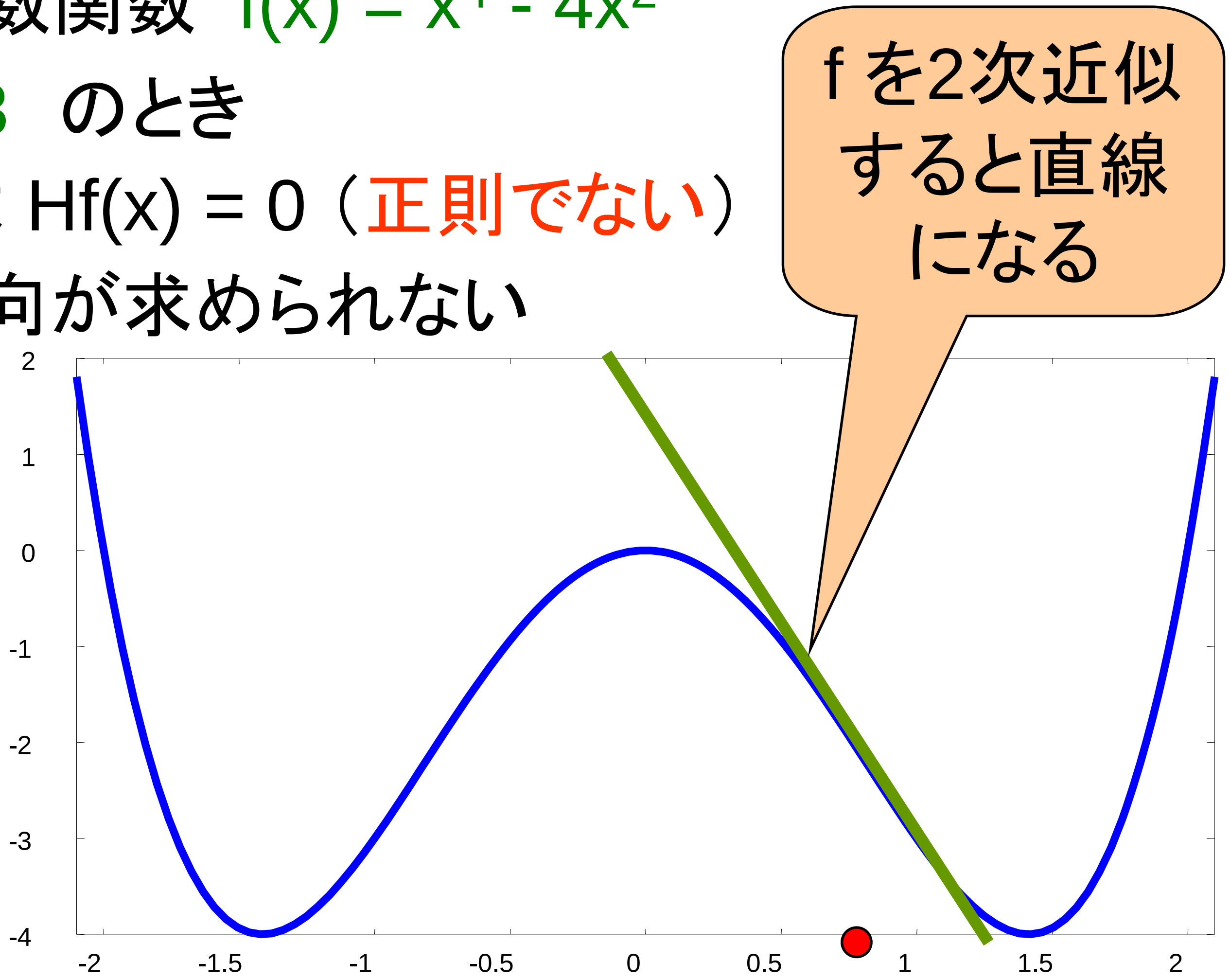
■ ヘッセ行列が正則でないと破綻

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



ニュートン法の問題点 [p.107]

- 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