数理計画法 第13回

塩浦昭義
情報科学研究科 准教授

shioura@dais.is.tohoku.ac.jp
http://www.dais.is.tohoku.ac.jp/~shioura/teaching

期末試験について

- 日時：2月2日（木）午後1時～2時30分（日時確定）
- 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
  - ネットワーク計画と非線形計画に関するレポートを1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：ネットワーク計画，非線形計画の範囲（今回の内容まで）
  - （詳しくはWeb上の過去問を参考にしてください）

※ 期末試験の出来が悪く，かつ救済を希望する場合は
2月3日（金）17:00までに問い合わせること
（救済できない可能性もあり）

関数のヘッセ行列

- 非線形計画では，多変数関数の微分が重要な役割を果たす
- 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

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

- f が2変数以上の関数の時は，幾何学的な情報を含む（次のスライド）
- f が1変数関数のときは，ヘッセ行列は2階微分と同じ

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

2次の最適性条件

- 非線形計画問題に対する最適性条件
～実行可能解が（局所的）最適解であるための
十分条件または必要条件

- 定理[最適解であるための2次の必要条件]:
ベクトル x^* が関数 f に関する制約なし問題の局所的最適解
➔ $\nabla^2 f(x^*)$は半正定値行列

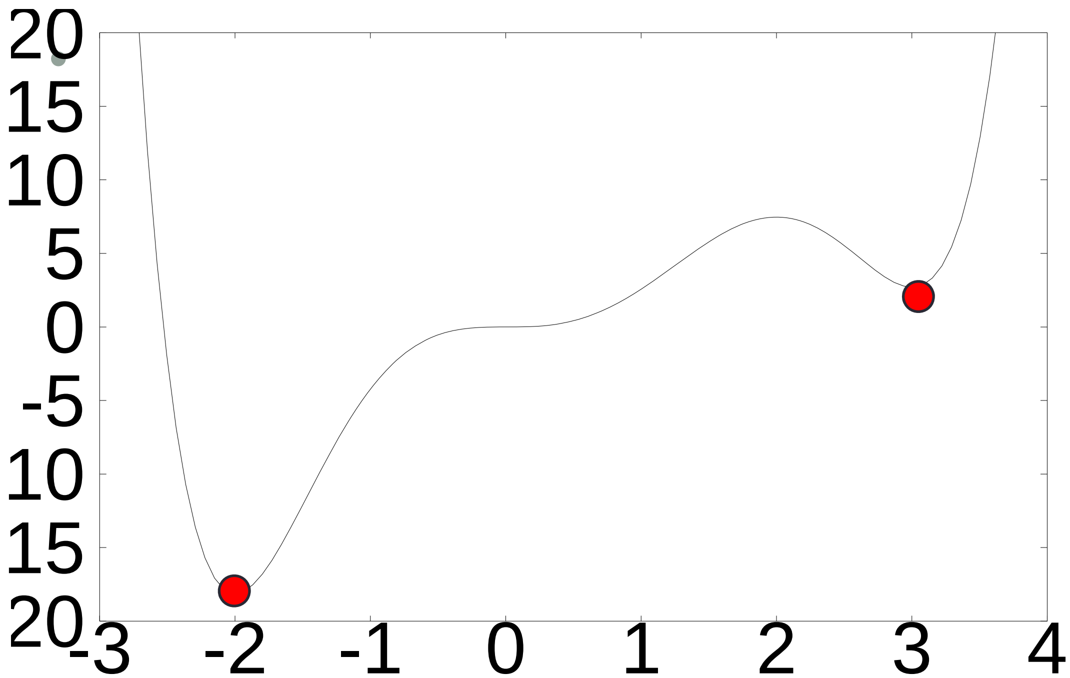
- 定義 : $n \times n$行列Aは半正定値
⇔ 任意のn次元ベクトルxに対し, $x^T A x \geq 0$

f が一変数関数の場合,
$\nabla^2 f(x^*)$は半正定値行列⇔$f''(x^*) \geq 0$が成り立つ
「極小点での傾きの増加率は非負」

2次の最適性条件：例1

- 目的関数：$f(x) = \frac{1}{6}x^6 - \frac{3}{5}x^5 - x^4 + 4x^3$ → 最小化
- 2階微分を計算：$f''(x) = 5x^4 - 12x^3 - 12x^2 + 24x$
- 局所的最適解は$x = -2, +3$

$$f''(-2) = 80 \geq 0, f''(3) = 45 \geq 0$$



2次の最適性条件

- 定理[最適解であるための2次の十分条件]:

ベクトル x^* が$\nabla f(x) = 0$を満たし，かつ$\nabla^2 f(x^*)$が正定値行列

➔x^* は関数 f に関する制約なし問題の局所的最適解

- 定義 : $n \times n$行列Aは正定値

  $\Leftrightarrow$ 任意の非ゼロ n 次元ベクトルxに対し，$x^T A x > 0$

f が一変数関数の場合，

$\nabla^2 f(x^*)$が正定値行列⇔$f''(x^*) > 0$が成り立つ

行列の正定値性，半正定値性

- 定義：$n \times n$行列Aは半正定値
  $\Leftrightarrow$ 任意のn次元ベクトルxに対し，$x^T Ax \geq 0$
- 定義：$n \times n$行列Aは正定値
  $\Leftrightarrow$ 任意の非ゼロ n 次元ベクトルxに対し，$x^T Ax > 0$
- 2×2対称行列$A = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$ $(a, b, c$は実数$)$の場合，

Aは半正定値$\Leftrightarrow a \geq 0, b \geq 0, ac - b^2 \geq 0$

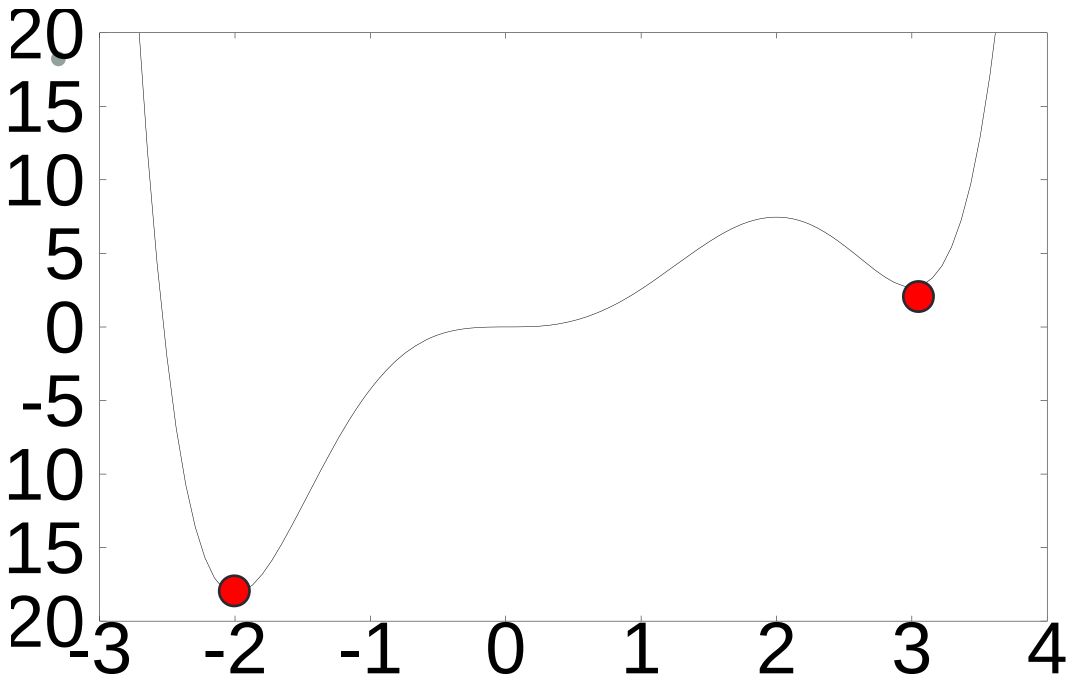
Aは正定値$\Leftrightarrow a > 0, b > 0, ac - b^2 > 0$

$\begin{bmatrix} 2 & 0 \\ 0 & 6 \end{bmatrix}$ は正定値，$\begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}$ は半正定値だが，正定値ではない

$\begin{bmatrix} 2 & -1 \\ -1 & 6 \end{bmatrix}$ は正定値，$\begin{bmatrix} 2 & 4 \\ 4 & 6 \end{bmatrix}$ は半正定値ではない

2次の最適性条件:例1(続き)

- 目的関数:$f(x) = \frac{1}{6}x^6 - \frac{3}{5}x^5 - x^4 + 4x^3$ → 最小化
- 2階微分を計算:$f''(x) = 5x^4 - 12x^3 - 12x^2 + 24x$
- $x \ = \ -2, +3$ は$\nabla f(x) = 0$を満たす
- $f''(-2) = 80 > 0, f''(3) = 45 > 0$➔ともに局所的最適解

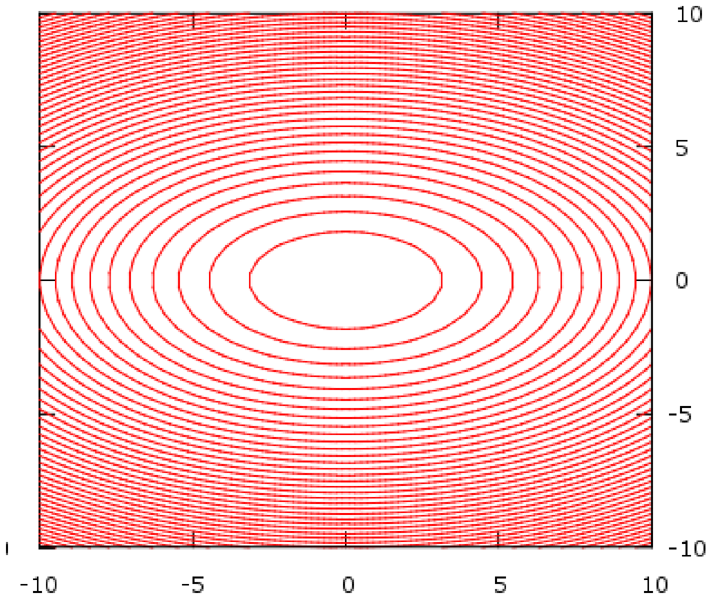


2次の最適性条件:例2

- 目的関数: $f(x_1, x_2) = x_1^2 + 3x_2^2$ → 最小化
- $\nabla f(x_1, x_2) = \begin{bmatrix} 2x_1 \\ 6x_2 \end{bmatrix}, \nabla^2 f(x_1, x_2) = \begin{bmatrix} 2 & 0 \\ 0 & 6 \end{bmatrix}$
- $\nabla f(x_1, x_2)$がゼロベクトルとなるのは (0,0) のみ
- $\nabla^2 f(0,0) = \begin{bmatrix} 2 & 0 \\ 0 & 6 \end{bmatrix}$ は正定値行列➔(0, 0) は局所的最適解

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：例3

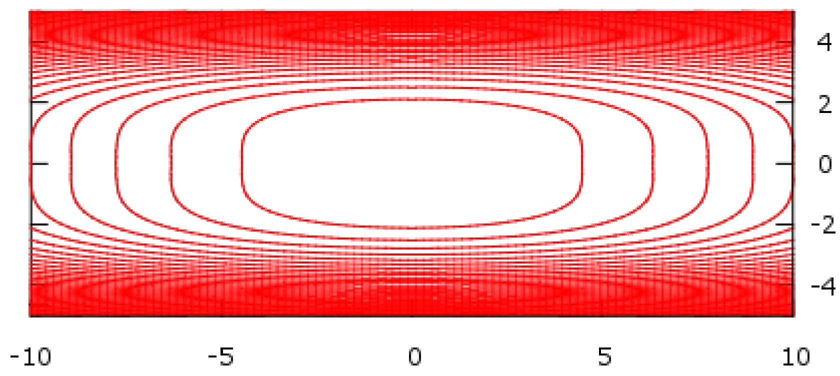
- 目的関数：$f(x_1, x_2) = x_1^2 + x_2^4$ → 最小化
- $\nabla f(x_1, x_2) = \begin{bmatrix} 2x_1 \\ 4x_2^3 \end{bmatrix}, \nabla^2 f(x_1, x_2) = \begin{bmatrix} 2 & 0 \\ 0 & 12x_2^2 \end{bmatrix}$
- $\nabla f(x_1, x_2)$がゼロベクトルとなるのは (0,0) のみ（実は最適解）
- $\nabla^2 f(0,0) = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}$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

➔(0, 0) が局所的最適解かどうかは，ヘッセ行列を使って判定できない（実際には局所的最適解）

任意のベクトル(y_1, y_2)に対して

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

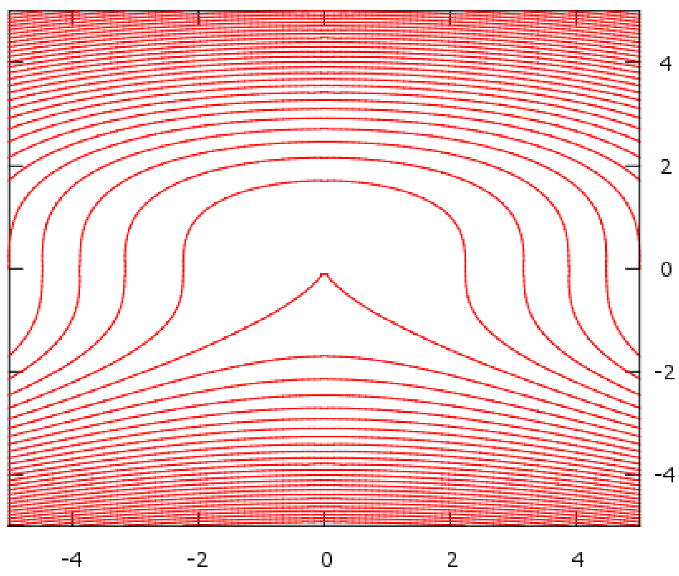
$y_1 = 0$のときは$y_2 \neq 0$でも値は0



2次の最適性条件：例4

- 目的関数：$f(x_1, x_2) = x_1^2 + x_2^3$ → 最小化
- $\nabla f(x_1, x_2) = \begin{bmatrix} 2x_1 \\ 3x_2^2 \end{bmatrix}, \nabla^2 f(x_1, x_2) = \begin{bmatrix} 2 & 0 \\ 0 & 6x_2 \end{bmatrix}$
- $\nabla f(x_1, x_2)$がゼロベクトルとなるのは (0,0) のみ
- $\nabla^2 f(0,0) = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}$ は半正定値だが，正定値ではない

➔(0, 0) が局所的最適解かどうかは，
ヘッセ行列を使って判定できない
（実際には局所的最適解ではない
任意の$\varepsilon > 0$に対して，
$$f(0, -\varepsilon) = -\varepsilon^3 < 0）$$



2次関数とその凸性

- n 変数の2次関数 f は,
  実数 c, n次元ベクトル$b = (b_1, b_2, \dots, b_n)$, $n \times n$対称行列$A = \left(a_{ij}\right)$
  を使って次の形に書ける

$$
\begin{aligned}
f(x) &= c + b^T x + \frac{1}{2} x^T A x \\
&= c + \sum_{i=1}^{n} b_i x_i + \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j
\end{aligned}
$$

- 性質 : 2次関数は凸関数←→行列Aが半正定値行列

- 2次関数が凸のとき, 最小点は$\nabla f(x) = b + Ax = 0$の解
  (Aが正定値行列ならば, 解は $x = A^{-1} b$)

テイラー展開を使った近似

- n変数関数 f, n次元ベクトル $\bar{x}$
- 定義:関数 f の $\bar{x}$ における2次のテイラー展開 $\tilde{f}$

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

  - 2次のテイラー展開は2次関数
    - ヘッセ行列が半正定値行列ならば凸2次関数
    - ヘッセ行列が正定値行列ならば, $\tilde{f}$の最小点は

$$x' = \bar{x} - \nabla^2 f(\bar{x})^{-1} \nabla f(\bar{x})$$

ニュートン法：考え方

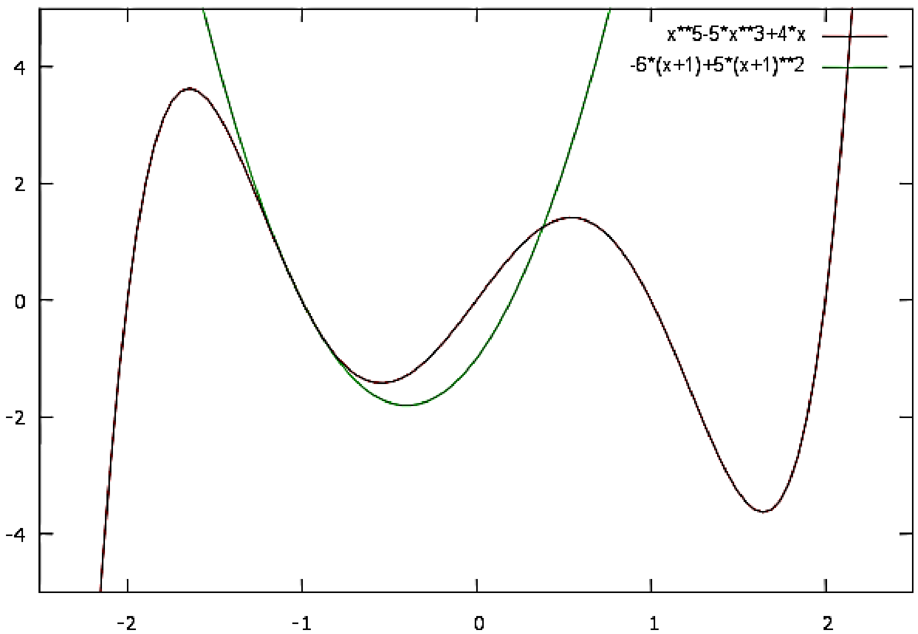
- 関数 f は，2次のテイラー展開 $\tilde{f}$によって近似的に表現できる

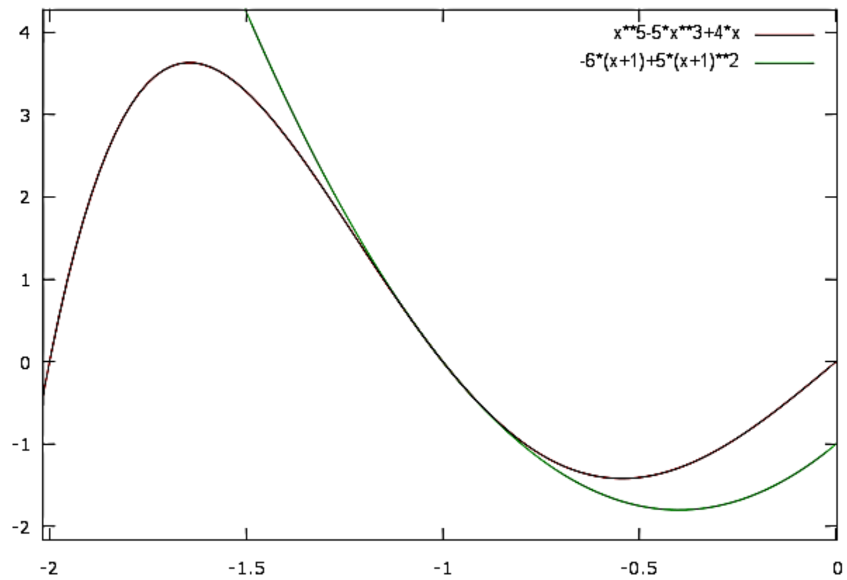
$$\tilde{f}(x) = f(\bar{x}) + \nabla f(\bar{x})^T(x-\bar{x}) + \frac{1}{2}(x-\bar{x})^T \nabla^2 f(\bar{x})(x-\bar{x})$$

➔ （ヘッセ行列が正定値のとき） $\tilde{f}$の最小点は$x' = \bar{x} - \nabla^2 f(\bar{x})^{-1} \nabla f(\bar{x})$

x'は fの大域的(もしくは局所的)最適解に近い（ことが期待される）

- この考えに基づいて x を繰り返し更新➔ニュートン法





ニュートン法のアルゴリズム

仮定：ヘッセ行列$\nabla^2 f(\bar{x})$は常に正定値行列

ステップ0：出発点$x^{(0)}$を選択，$k \coloneqq 0$とおく

ステップ1：$\nabla f\left(x^{(k)}\right) \simeq 0$ならば終了

そうでなければ，$d^{(k)} = -\nabla^2 f\left(x^{(k)}\right)^{-1} \nabla f\left(x^{(k)}\right)$とおく

ステップ2：$x^{(k+1)} \coloneqq x^{(k)} + d^{(k)}$とおく.

$k \coloneqq k + 1$とおいて、ステップ1に戻る

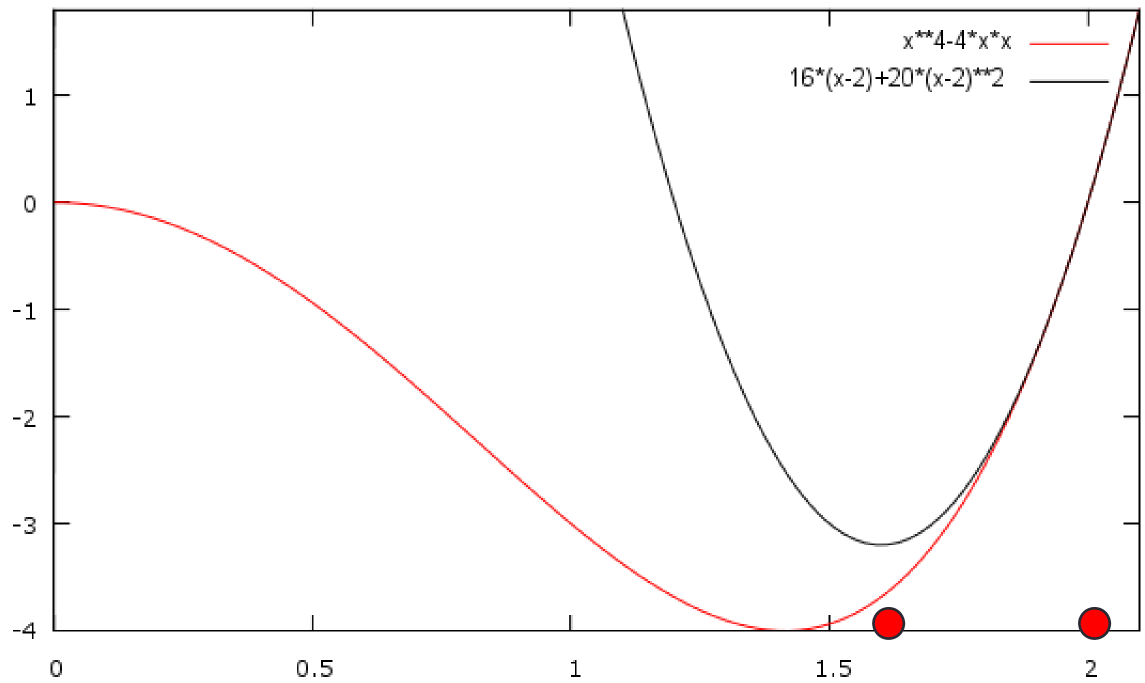
関数 f の $\bar{x}$ における2次のテイラー展開 $\tilde{f}$

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

の最小点

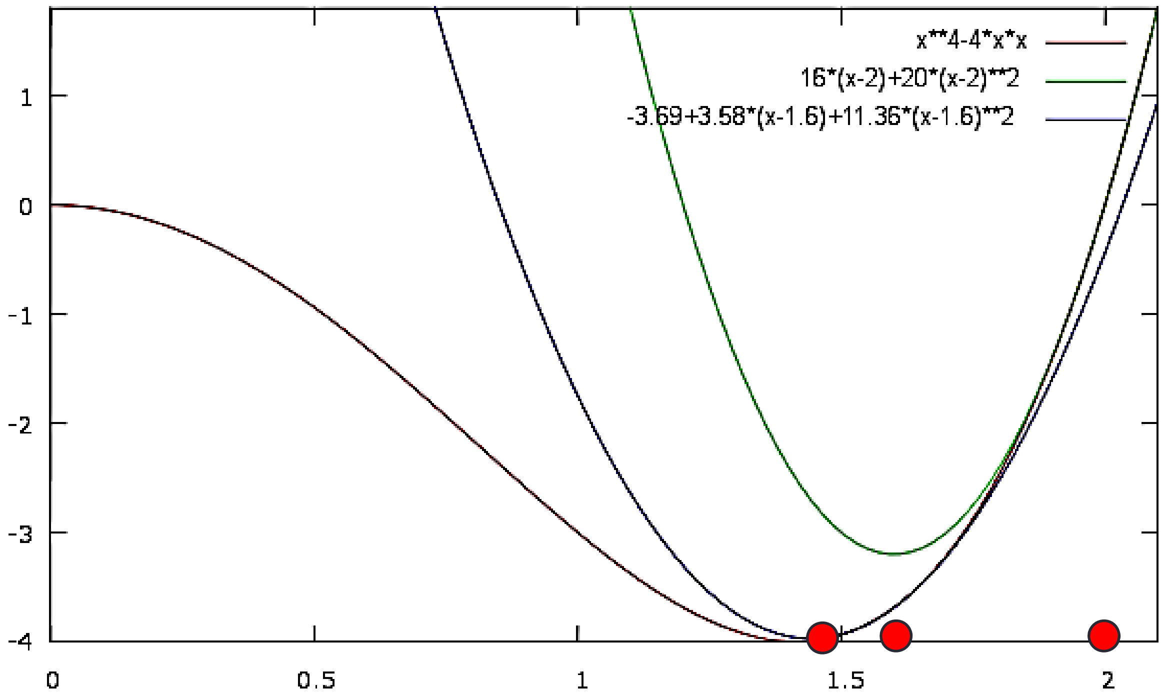
ニュートン法の実行例その1

- 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



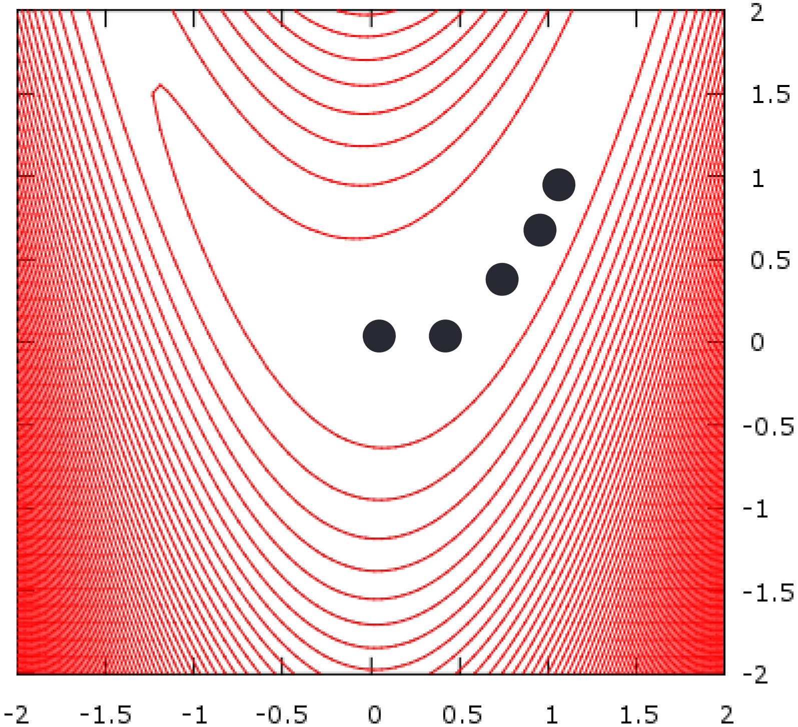
ニュートン法の実行例その1

- 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
  - 初期階(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)



p. 118表4. 2を参照

ニュートン法の性質

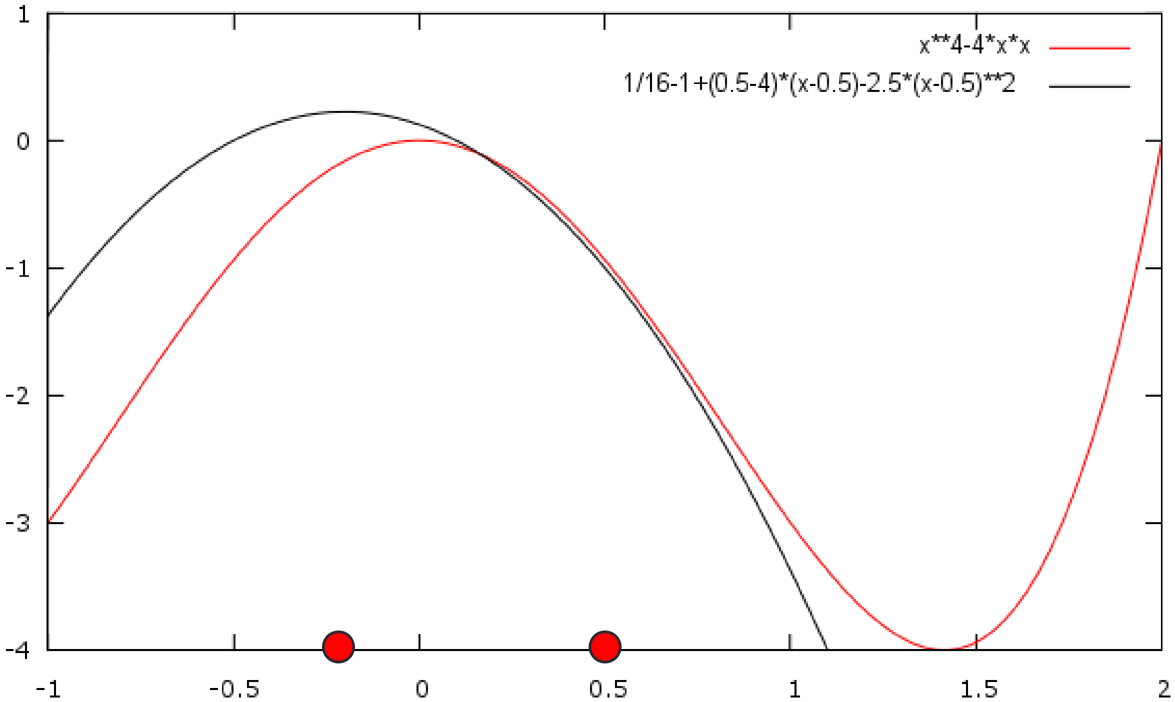
- 二次収束をする➔非常に速く収束する
  - ある定数$0 < \beta < 1$とある整数k_0に対して，
    $k \geq k_0$ ならば
    $$||x^{(k+1)} - x^*|| \leq \beta ||x^{(k)} - x^*||^2$$
    が成立（x^*は極限の点，収束先）
- 「ヘッセ行列$\nabla^2 f(\bar{x})$は常に正定値行列」という仮定が必要
  - この仮定が常に成り立つのは，特別な凸関数のみ
  - 一般の関数の場合，この仮定は局所的にしか保証されない
    ➔ニュートン法は局所最適解に近いところでのみ利用可能

ニュートン法の失敗例

- 一変数関数 $f(x) = x^4 - 4x^2$
- $f'(x) = 4x^3 - 8x, f''(x) = 12x^2 - 8$
- ヘッセ行列が正定値⇔$f''(x) > 0$ ⇔$x^2 > \frac{2}{3}$
- $x = 0.5$のとき,
- $f''(x) = -5 < 0$

➔テイラー展開は上に凸の二次関数

➔ニュートン法で次の点に進むと関数値が増える

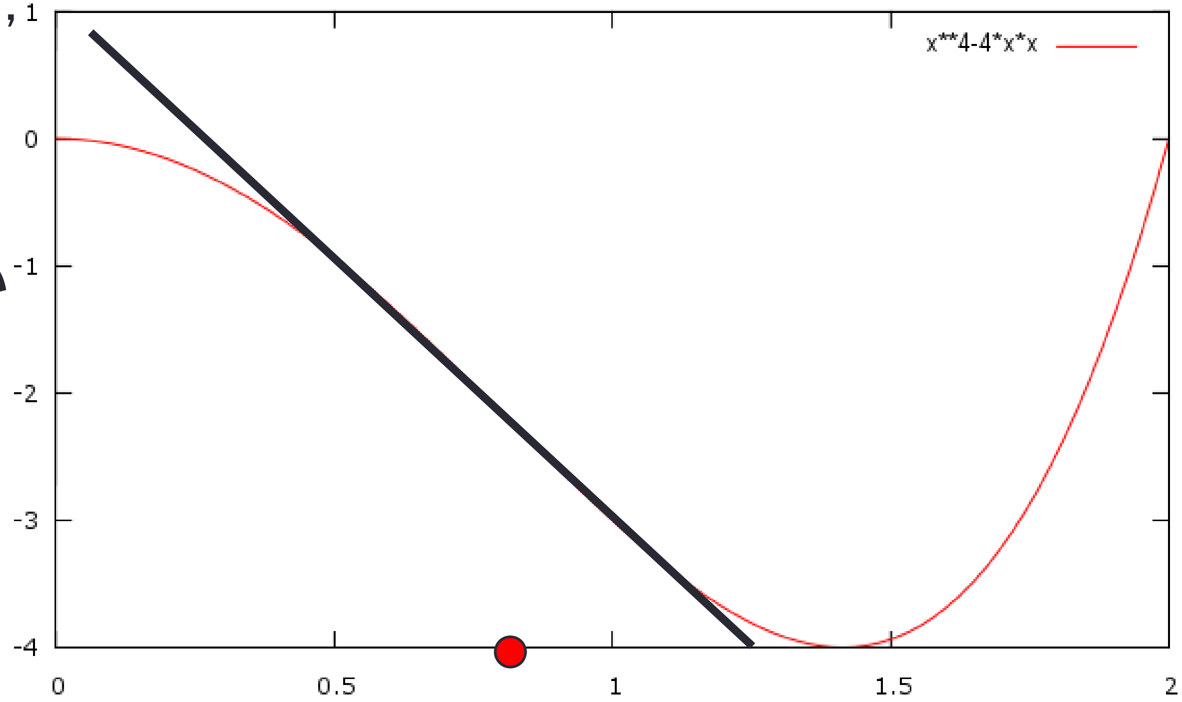


ニュートン法の失敗例

- 一変数関数 $f(x) = x^4 - 4x^2$
- $f'(x) = 4x^3 - 8x, f''(x) = 12x^2 - 8$
- ヘッセ行列が正定値➡➡$f''(x) > 0$ ➡➡$x^2 > \frac{2}{3}$
- $x = \sqrt{2/3} \cong 0.81$のとき,
- $f''(x) = 0$

➔テイラー展開は直線

➔次の点が求められない



演習問題(提出の必要なし)

問題1:関数 $f(x) = x^3 + 6x^2$ に対して

(a)出発点を $x = 2$ としてニュートン法を適用せよ。

(b)出発点を $x = 1$ としてニュートン法を適用せよ。

それぞれ、反復は2回行うこと。

問題2:関数 $f(x,y) = \frac{1}{2}(x+y)^2 + (y-1)^2$ に対して,

出発点を $(x,y) = (1,1)$ としてニュートン法を適用せよ.

反復は2回行うこと.

ヒント:$\alpha > 1$ のとき, 行列 $\begin{bmatrix} 1 & 1 \\ 1 & \alpha \end{bmatrix}$ の逆行列は $\begin{bmatrix} \frac{\alpha}{\alpha-1} & \frac{-1}{\alpha-1} \\ \frac{-1}{\alpha-1} & \frac{1}{\alpha-1} \end{bmatrix}$