数理手法（数理最適化）第13回
非線形計画
ニュートン法

塩浦昭義
東京工業大学 経営工学系 准教授

shioura.a.aa@m.titech.ac.jp

http://www.me.titech.ac.jp/~shioura/shioura/teaching/TUmp17/index.html

期末試験について

- 日時：1月17日（水）13:05〜14:35
- 場所： 工2号館 212講義室（授業の部屋）
- 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8回〜第13回の講義で教えたところ
  - ネットワーク最適化
  - 非線形計画
- 50点満点，20点以下は不合格
- 中間と合わせて51点以上は合格，50点以下は単位不可

2次の最適性条件(必要条件)証明

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**

証明: $A=Hf(x^*)$ とおく.
背理法: A は半正定値でないと仮定
➔ $||y||=1$ なるベクトル y が存在して, $y^TAy<0$
以下に示すように,
ある $\varepsilon>0$ に対して $f(x^*+\varepsilon' y)<f(x^*)$ $(0<\forall\varepsilon'<\varepsilon)$ となり, 矛盾.
$x = x^*$ での2次のテイラー展開と $\nabla f(x^*)=0$ を使うと,

$$f(x^*+\varepsilon y)=f(x^*)+\nabla f(x^*)^T(\varepsilon y)+\frac{1}{2}(\varepsilon y)^T A(\varepsilon y)+\psi(\varepsilon y)$$

$$=f(x^*)+\varepsilon^2\left(\frac{1}{2}y^TAy+\frac{\psi(\varepsilon y)}{\varepsilon^2}\right)$$

テイラー展開の性質より, ある $\varepsilon>0$ が存在して, $0<\forall\varepsilon'<\varepsilon$ に対して

$$\frac{1}{2}y^TAy+\frac{\psi(\varepsilon' y)}{(\varepsilon')^2}<0 \quad \therefore f(x^*+\varepsilon' y)<f(x^*)$$

2次の最適性条件(十分条件)証明

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

証明の概略:

x = x* での2次のテイラー展開 $\tilde{f}$ を考えると,

$\tilde{f}$ は凸関数, x* が最小解 ∴ x* は$\tilde{f}$ の極小解

x* のある近傍において, $\tilde{f}$ と f は十分に近い ∴ x* は f の極小解

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

凸関数の特徴付け(その2)

定理: f: 凸関数, 微分可能(ヘッセ行列が定義可能)
$\Leftrightarrow$ 任意のベクトル x に対して
ヘッセ行列 $Hf(x)$ が半正定値

証明は略

一変数凸関数の場合:
関数 f は凸関数$\Leftrightarrow$任意のx に対して二階微分$f''(x) \geq 0$

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

2次関数$f(x)=\frac{1}{2}x^TVx+cx+c_0$の係数行列 V が

正定値行列のとき，最小解（最適解）は簡単に求められる！

- $\nabla f(x)=Vx+c$ ➔ 停留点は $x^*=-V^{-1}c$ のみ
- ヘッセ行列 ＝ V, 正定値行列 ➔ 停留点は最小解

⇩

2次の十分条件より x^* は最小解

※ 正定値行列は正則行列（逆行列をもつ）
　半正定値行列は正則とは限らない

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：

V が正定値の2次関数に対して最適解は簡単に求められる！

ただし，一般の関数は2次とは限らない

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

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

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

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

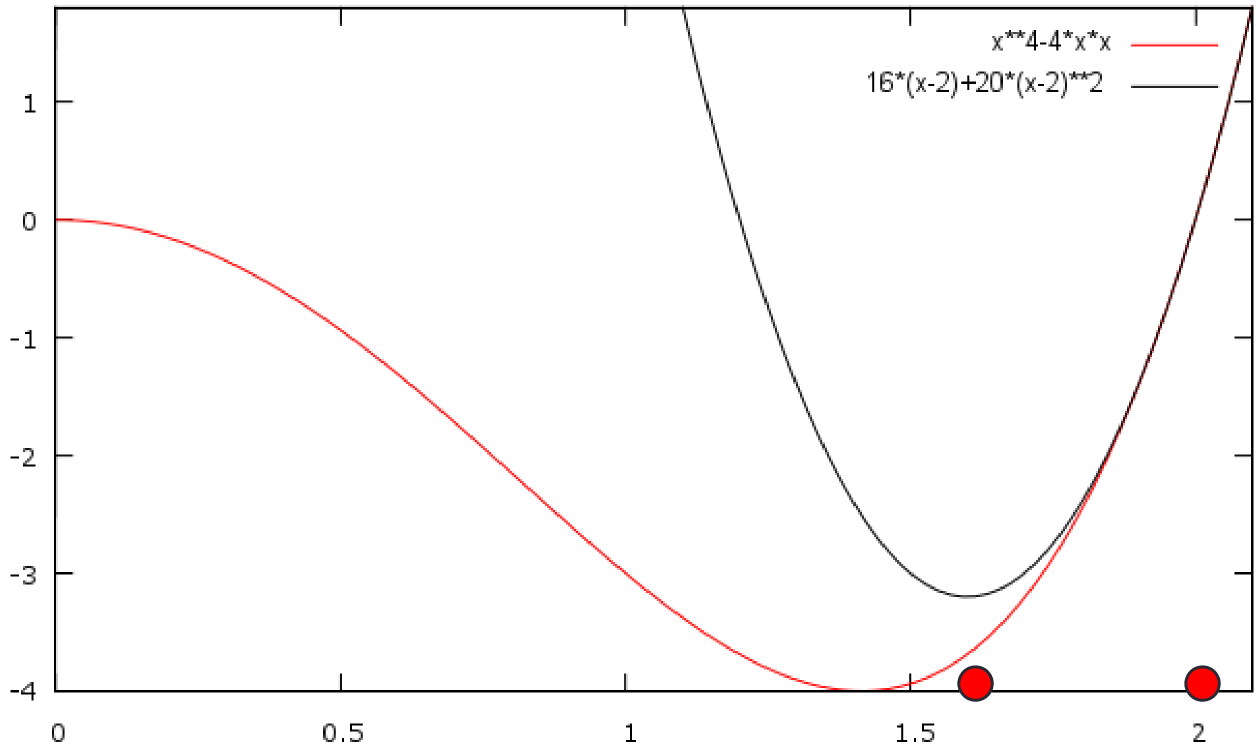
ニュートン法のアルゴリズム

現在の点 x から $x - Hf(x)^{-1}\nabla f(x)$ への移動を繰り返す
($-Hf(x)^{-1}\nabla f(x)$ を，xにおけるニュートン方向と呼ぶ)

入力：関数 f，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

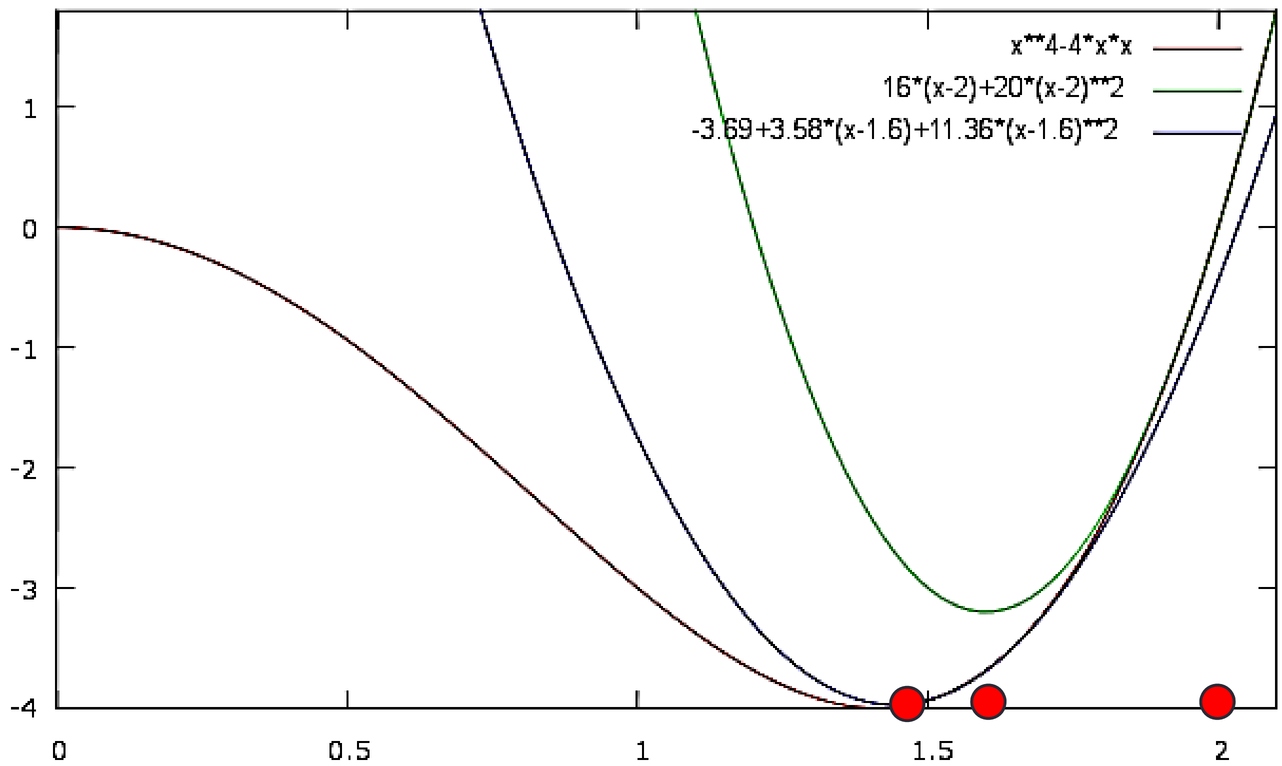
ニュートン法の実行例その1

- 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
- $x^{(1)} := 1.6$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
- $x^{(2)} \coloneqq 1.49$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)

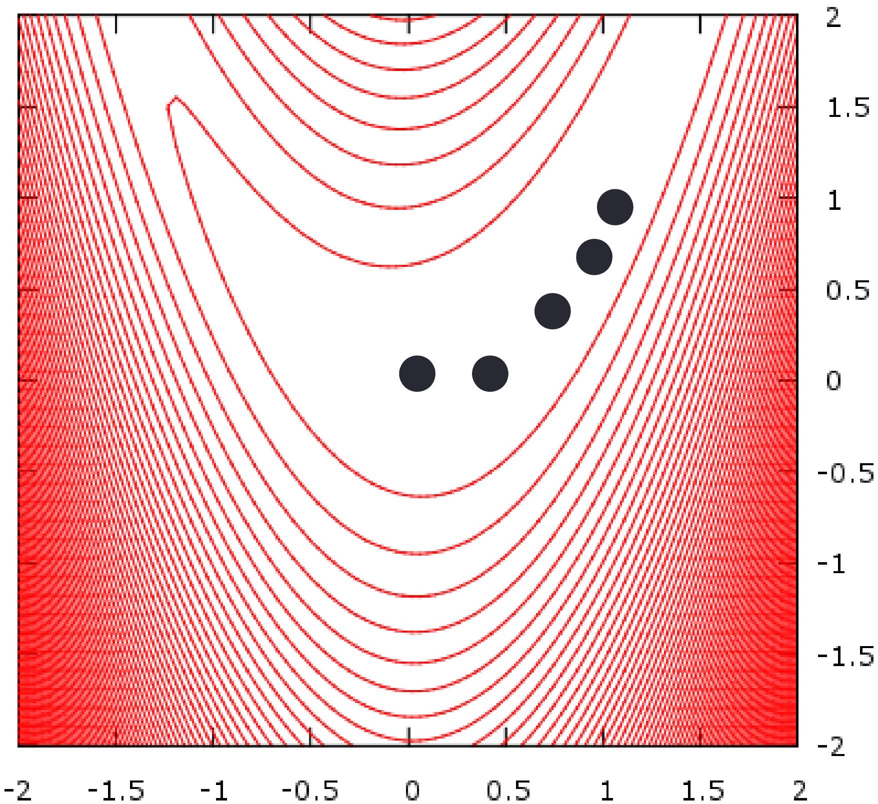


表 4.2　関数 (4.19) に対するニュートン法の計算結果

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

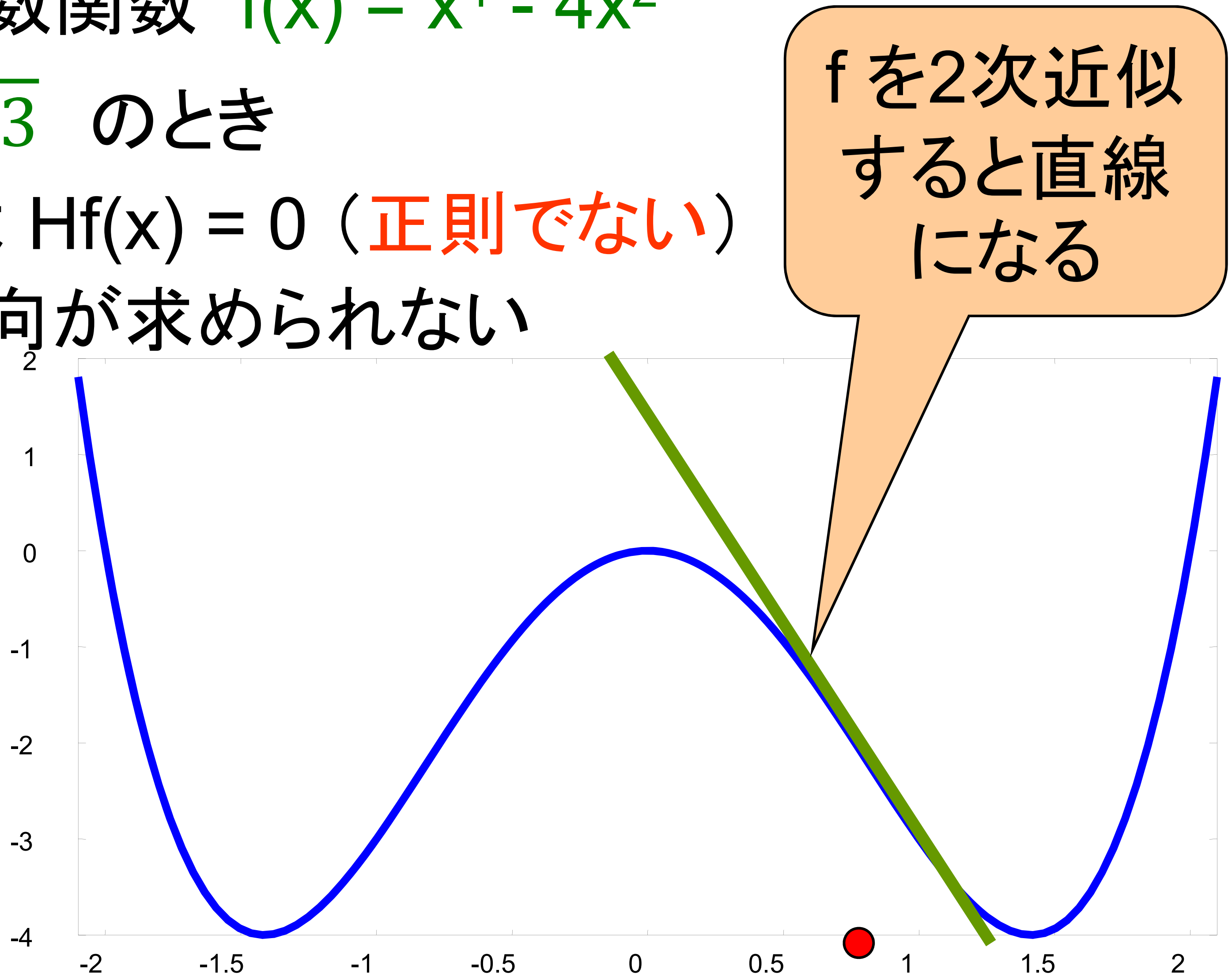
■ ヘッセ行列が正則でないと破綻

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



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