数理計画法
（数理最適化）第13回

非線形計画

ニュートン法，凸関数

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30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ィ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

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$f(x) = \frac{1}{2}x^T Vx + cx + c_0$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(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 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

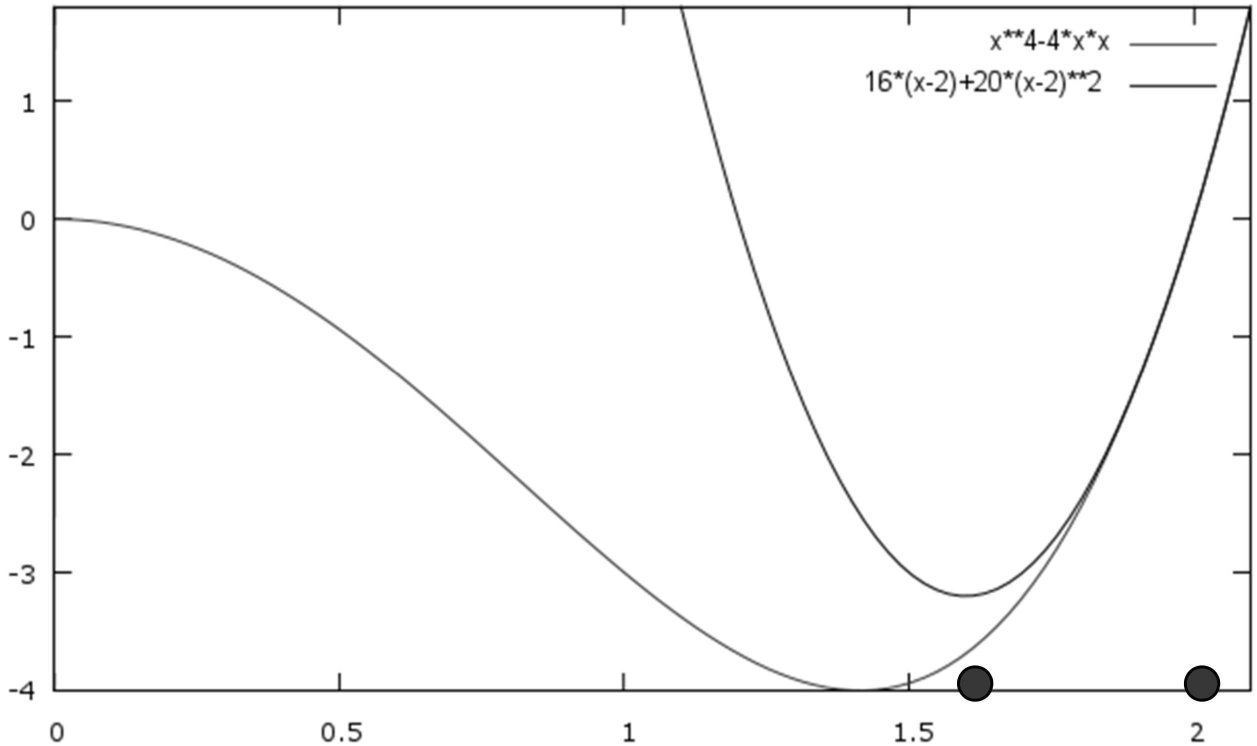
ステップ2：ニュートン方向 $-Hf(x^k)^{-1}\nabla f(x^k)$ を計算

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

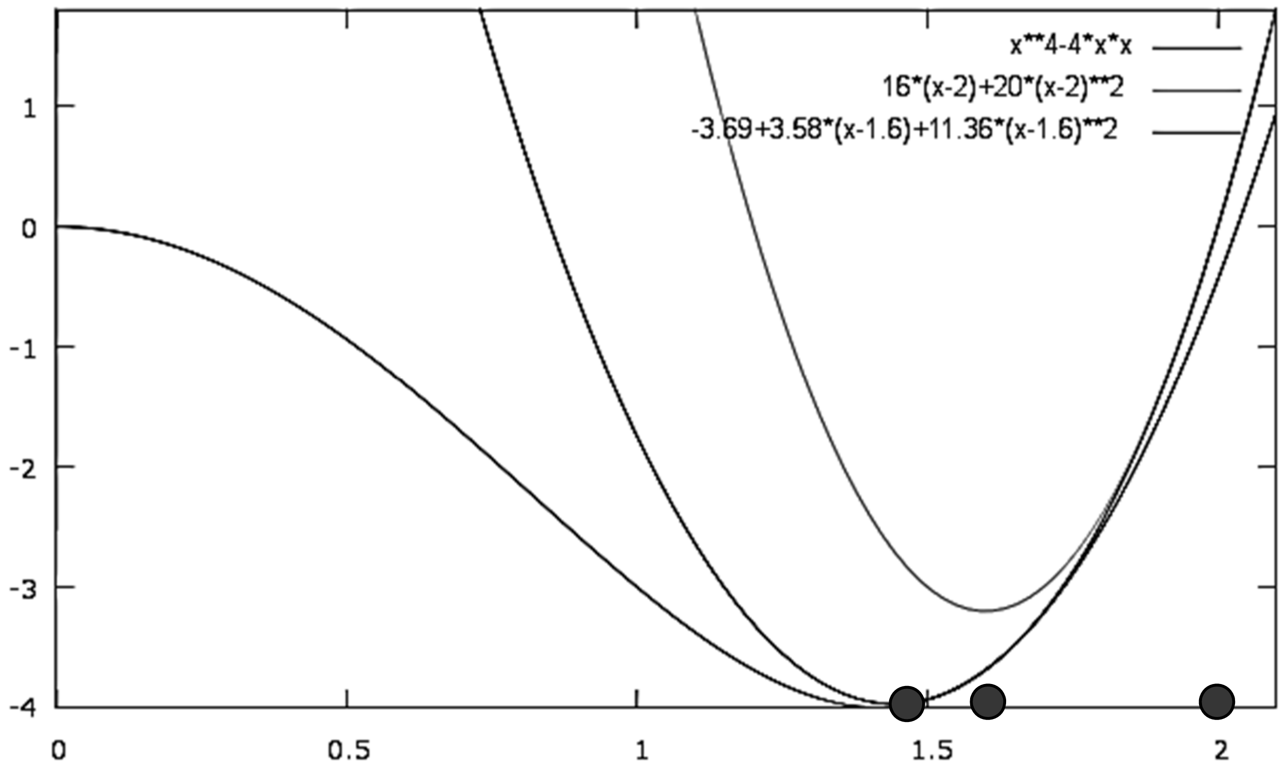
ニュートン法の実行例その1

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

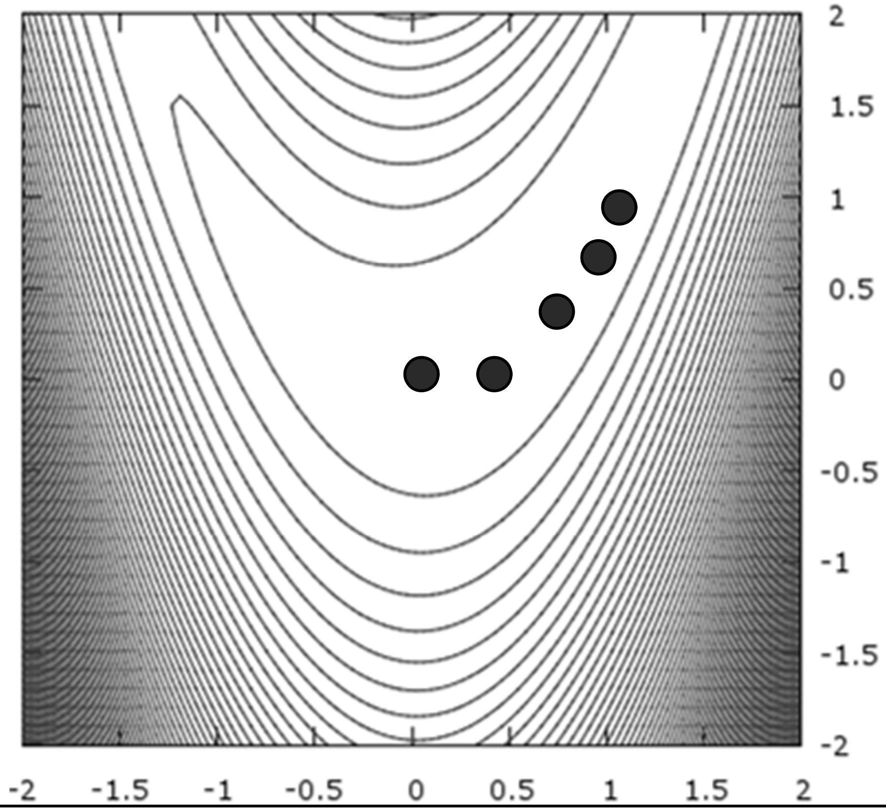


表 4.2　関数 (4.19) に対するニュートン法の計算結果

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

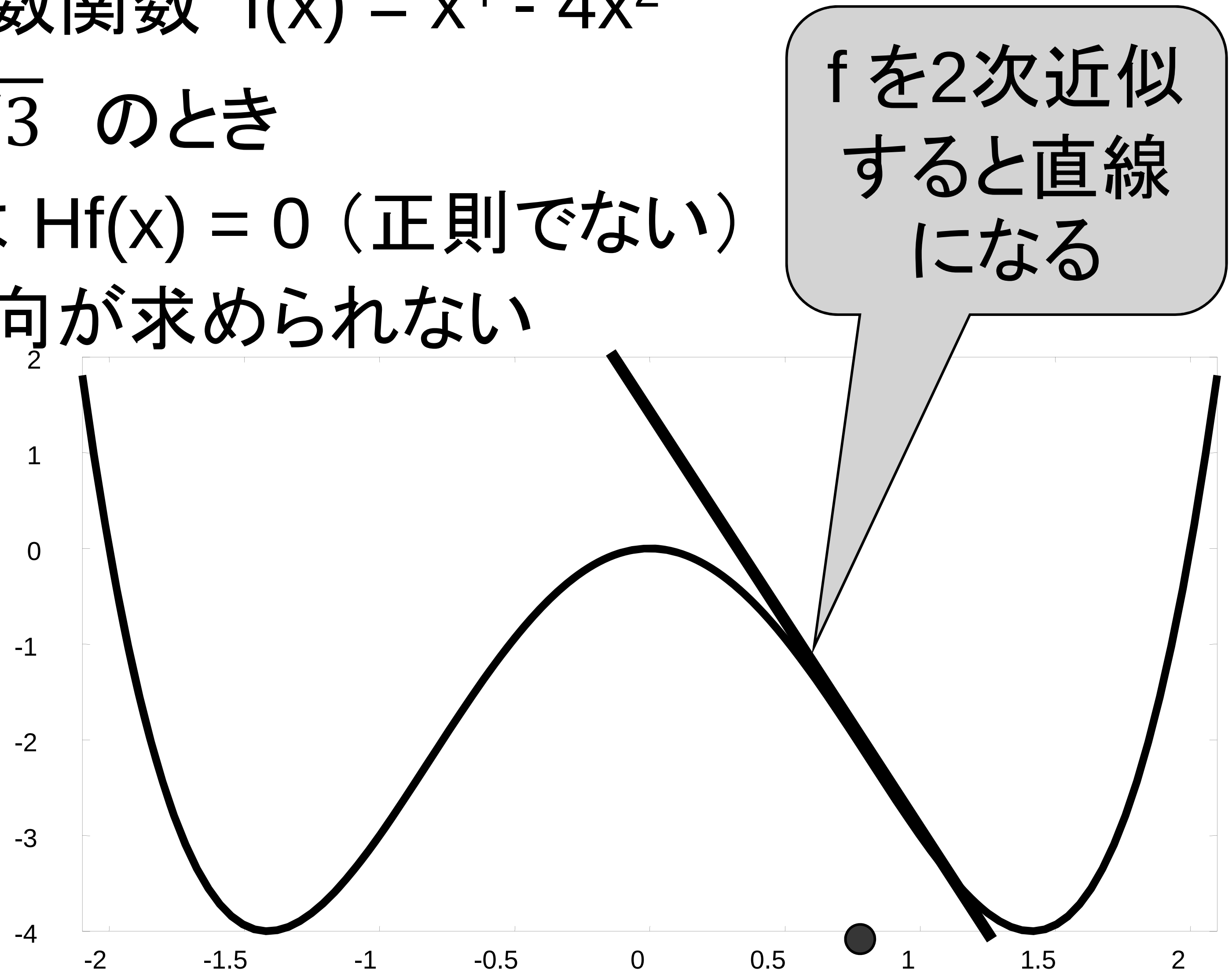
■ ヘッセ行列が正則でないと破綻

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



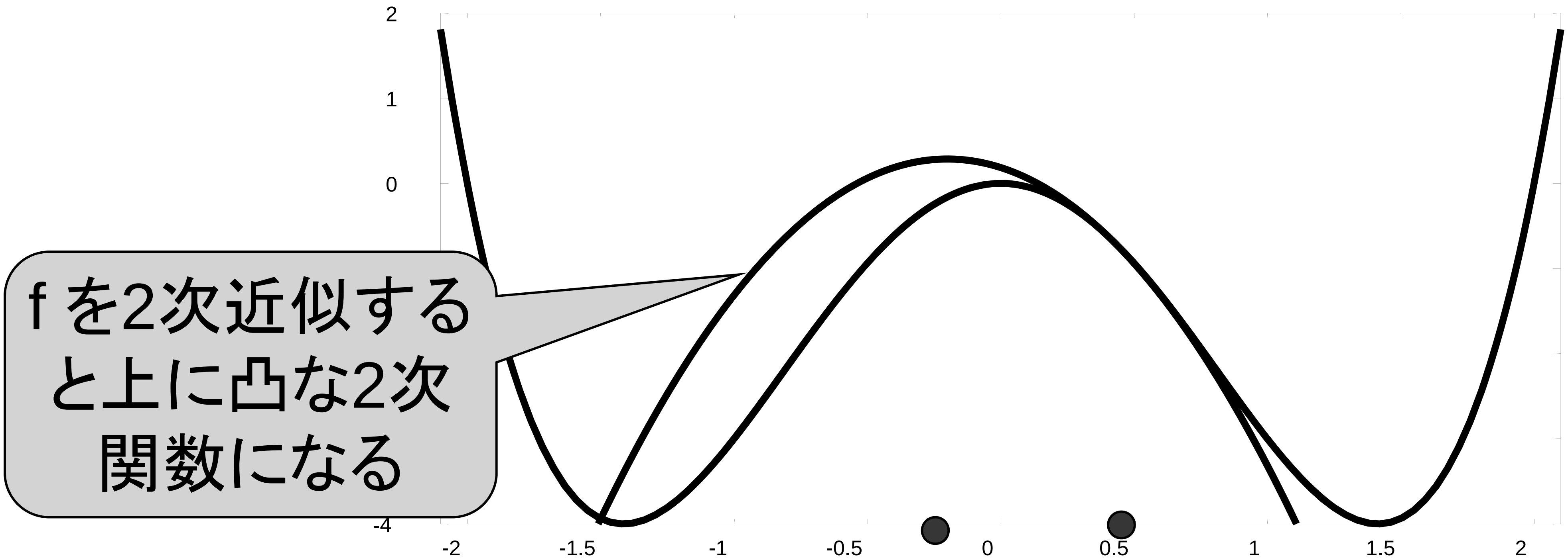
ニュートン法の問題点

- 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

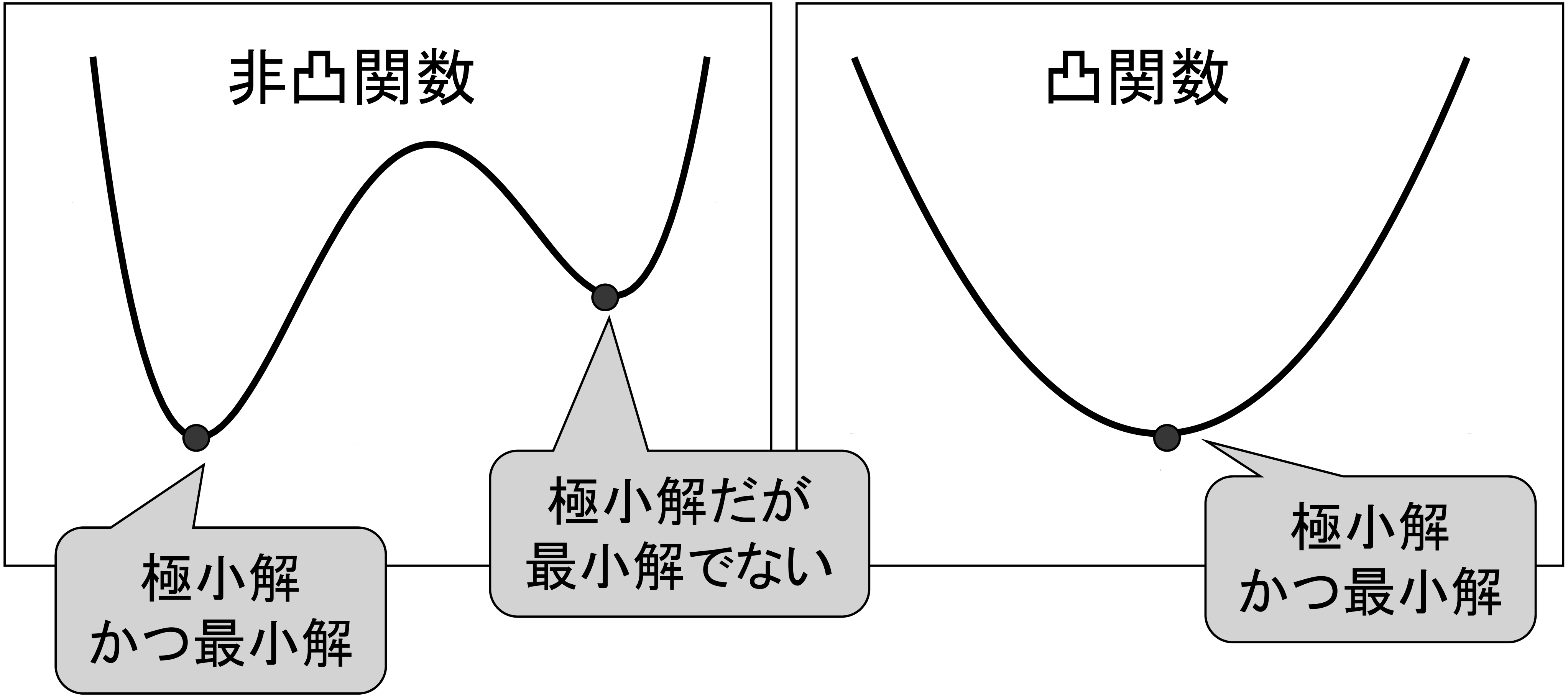
⇒　ヘッセ行列は Hf(x) = -5（正定値でない）

⇒　ニュートン方向に進むと関数値が増加する



凸関数

最小化しやすい関数の形は？



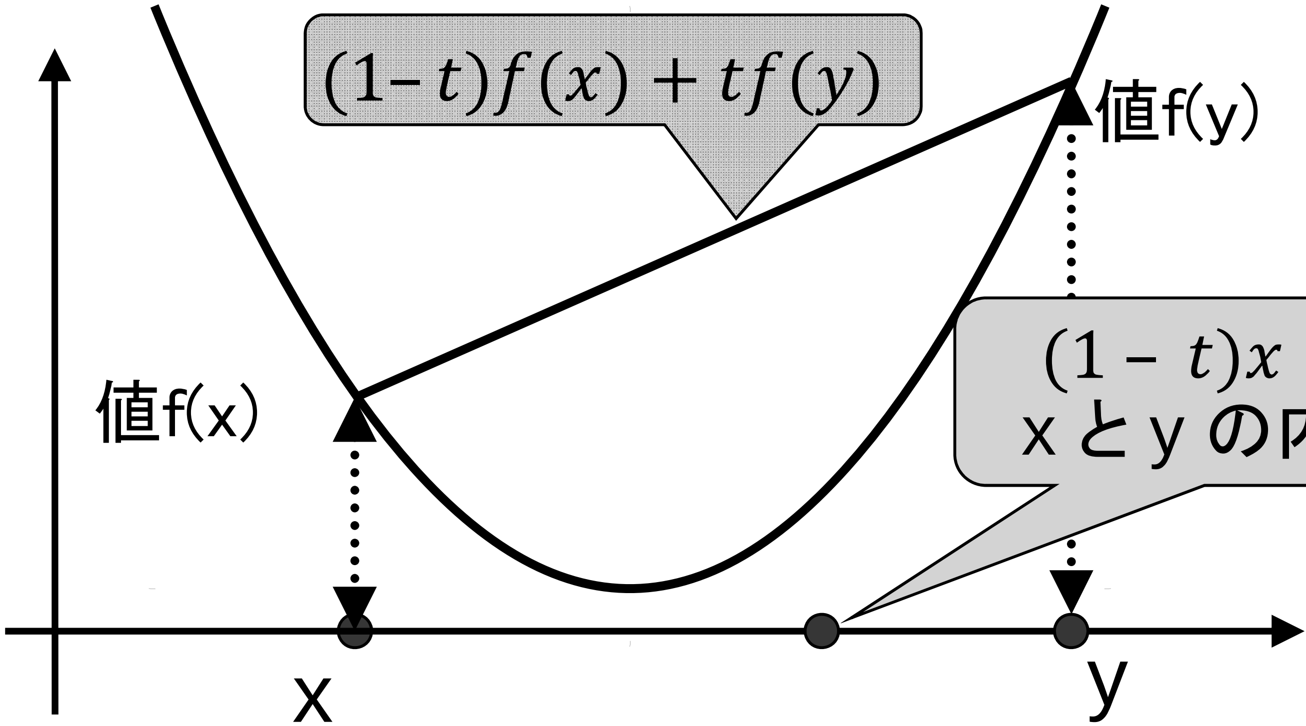
最小解でない極小解がある
➔最小化が難しい

極小解が一つ
➔最小化しやすい

凸関数の定義

定義：関数 f は凸関数
⇔　任意の異なるベクトル x, y および任意の $0 < t < 1$ に対し

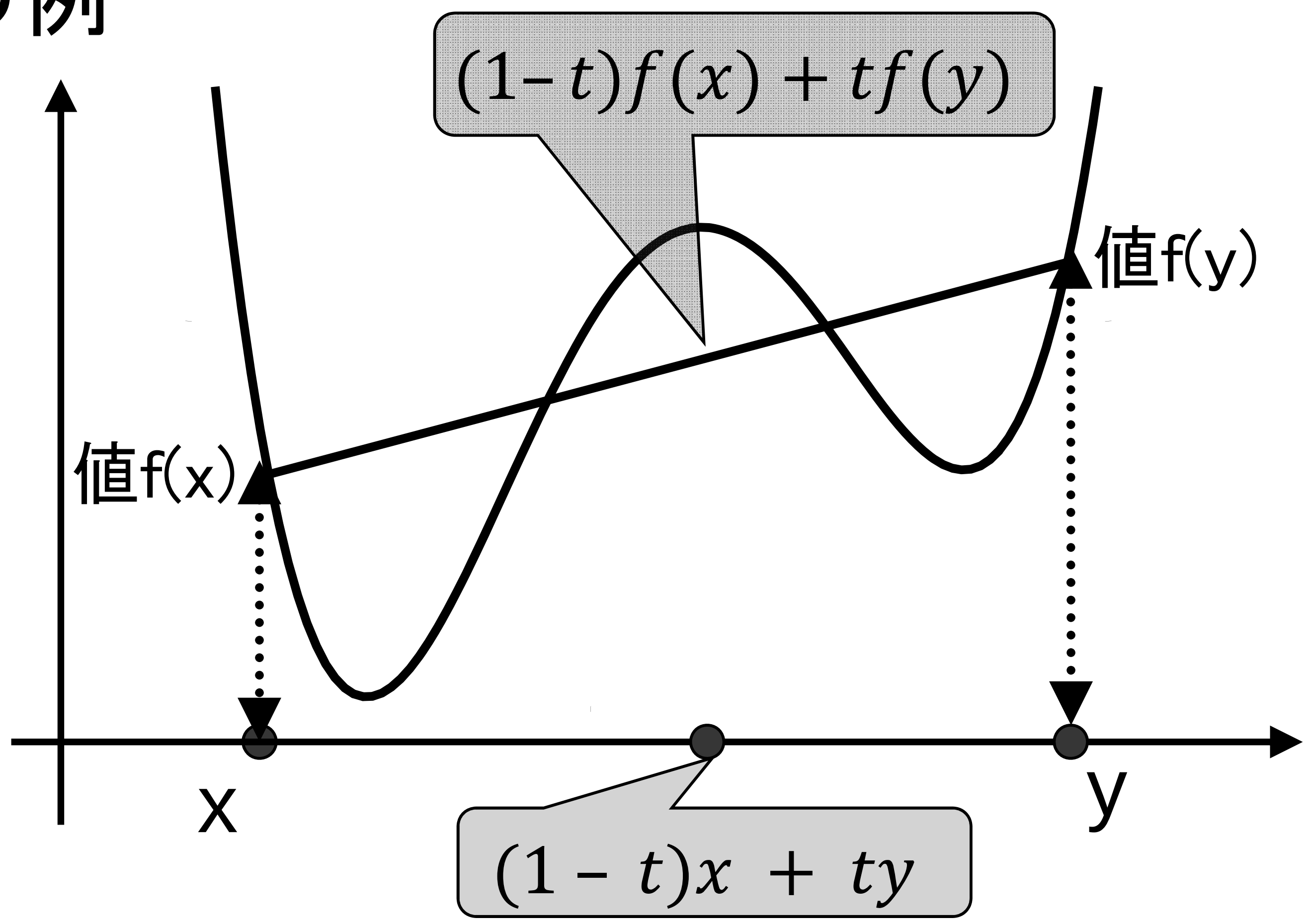
$$(1-t)f(x) + t\,f(y) \geq f((1-t)\,x + t\,y)$$

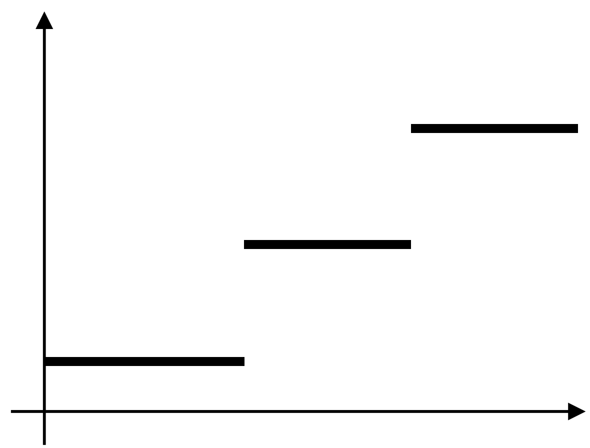


凸関数の定義(続き)

凸関数 ⇔ $(1-t)f(x)+tf(y) \geqq f((1-t)x+ty)$

非凸関数の例





2次の凸関数

凸関数 $\Leftrightarrow$ $(1-t)f(x) + t f(y) \geqq f((1-t)x + t y)$

2次関数 $f(x) = a x^2 + bx + c$ $(a > 0)$ は凸関数

（証明）任意の異なる x, y と $0 < t < 1$ に対して、

$$(1-\mathrm{t})f(x) + tf(y) - f((1-t)x + ty)$$

$$= (1-t)\,ax^2 + t\,a\,y^2 - a\,[(1-t)\,x + t\,y]^2$$

$$= (1-t)\,ax^2 + t\,a\,y^2 - a\,(1-t)^2 x^2 - a\,t^2 y^2 - 2a\,(1-t)\,t\,x\,y$$

$$= (t-t^2)\,ax^2 + (t-t^2)\,a\,y^2 - 2a\,(t-t^2)\,x\,y$$

$$= (t-t^2)\,a\,(x-y)^2$$

> 0　　$(0 < t < 1, x \neq y$より$)$

2次の凸関数（続き）

凸関数 ⇔ $(1-t)f(x) + t\,f(y) \geqq f((1-t)\,x + t\,y)$

より一般に,

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

（V: n ×n 行列, c: n次元ベクトル, c_0: 定数）

はＶが半正定値行列 ➔ 凸関数

例：

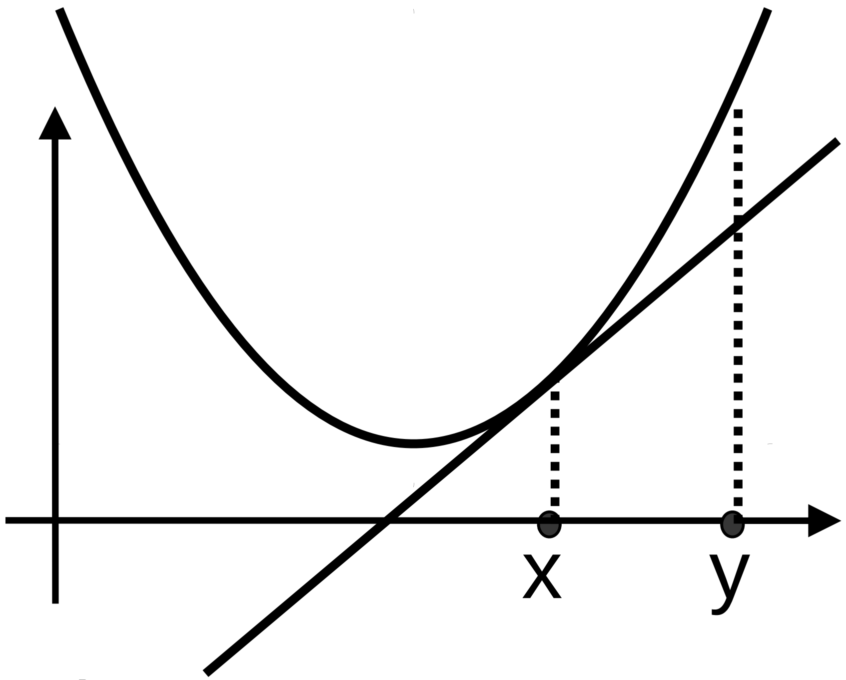
$$f(x_1, x_2) = \frac{1}{2}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}^T \begin{pmatrix} 2 & -2 \\ -2 & 5 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

凸関数の特徴付け(その1)

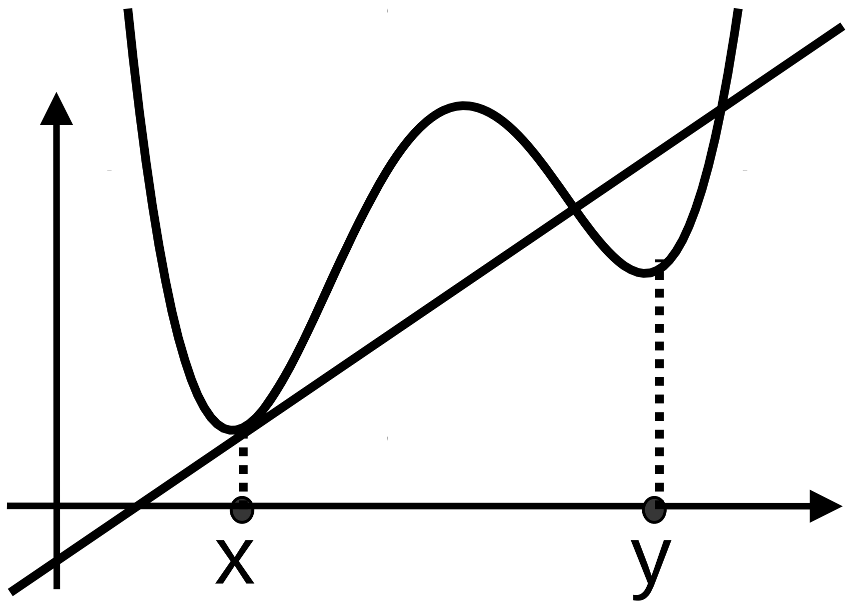
定理: f: 凸関数, 微分可能(勾配ベクトルが定義可能)
⇔ 任意のベクトル x, y に対して次の不等式が成立

$$f(y) \geq f(x) + \nabla f(x)(y-x)$$

証明は略



一変数凸関数の場合: x における接線 $y = f(x) + \nabla f(x)(y-x)$ より f(y) は上にある



一変数非凸関数の場合は成り立たない

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)
←→ 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←→任意のx に対して二階微分$f''(x) \geq 0$

凸関数の最適解の必要条件

定理: f: 凸関数, 微分可能（勾配ベクトルが定義可能）
x*: f の停留点 (∇f(x*)=0)
⇒ x*は制約なし問題の最適解

証明: f は凸関数なので, 任意のx, y に対して次が成り立つ

$$f(y) \geq f(x) + \nabla f(x)(y - x)$$

x = x* を代入すると, ∇f(x*)=0なので

$$f(y) \geq f(x^*) + \nabla f(x^*)(y - x^*) = f(x^*)$$

すなわち, 任意のベクトル y の関数値より,
x* の関数値は少ない（または等しい）
∴ x*は最適解

凸関数の最適解の必要条件

定理: f: 凸関数, x^*: f の極小解
⇒ x^*は制約なし問題の最適解

証明: x^* は極小解
⇒ ある $\varepsilon > 0$が存在して、
任意の x に対し $\|x - x^*\| < \varepsilon$ ならば $f(x) \geqq f(x^*)$

$f(y) < f(x^*)$ なる y が存在すると仮定
f は凸関数
⇒ $0 < t < 1$ なる任意の t に対して
$f((1-t)y + t x^*) \leqq (1-t) f(y) + t f(x^*) < f(x^*)$
t を1に近づけると
$(1-t)y + t x^*$ と x^* の距離 $< \varepsilon$（矛盾）