# 数理計画法 第11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


## 復習：ヘッセ行列

関数 f のヘッセ行列

$$\mathrm{H}\,f(\boldsymbol{x}) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1 \partial x_1} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_n} \end{pmatrix}$$

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

$$f(\boldsymbol{x}+\boldsymbol{d}) = f(\boldsymbol{x}) + \nabla f(\boldsymbol{x})^T \boldsymbol{d} + \frac{1}{2}\boldsymbol{d}^T \mathrm{H} f(\boldsymbol{x}) \boldsymbol{d} + o(\|\boldsymbol{d}\|^2)$$

関数 f の(x における)
2次のテイラー展開

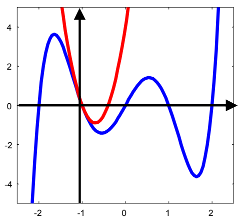
関数 f を多項式で表現したときの
2次項までの情報を取り出した式

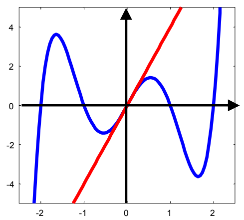
## 復習：ヘッセ行列とテイラー展開(続き)

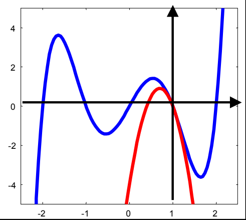
例2：$f(x) = (x-2)(x-1)x(x+1)(x+2) = x^5 - 5x^3 + 4x$

$\nabla f(x) = 5x^4 - 15x^2 + 4$　　　$\mathrm{H}\,f(x) = 20x^3 - 30x$

| | x = -1 のとき | x = 0 のとき | x = 1 のとき |
| --- | --- | --- | --- |
| テイラー展開 | $0 - 6d + 5d^2$ | $0 + 4d + 0d^2$ | $0 - 6d - 5d^2$ |







# 復習: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
⇔ 任意のベクトル y に対して $y^T A y \geq 0$

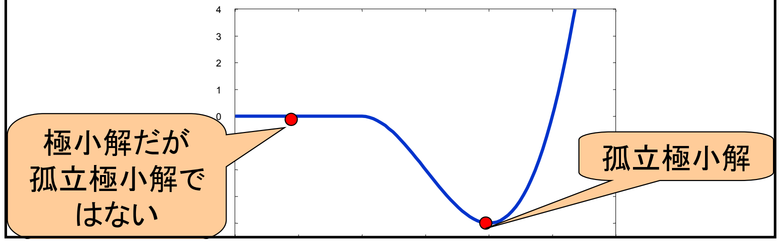
定義:正方行列 A は正定値
⇔ 任意の非零ベクトル y に対して $y^T A y > 0$

# 復習: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 ⇒ Hf(x*) は半正定値

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




# 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

定理:x*: 停留点, Hf(x*) : 正定値 ⇒ x*: 孤立極小解

# 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!

一般の関数 f は狭義2次凸関数とは限らない

⇒ f を2次のテイラー展開により近似

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

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

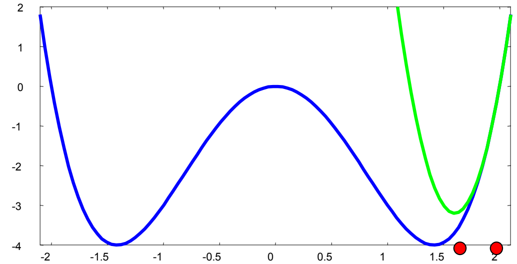
入力：関数 $f$ とその勾配ベクトル $\nabla f$, ヘッセ行列 $Hf$
初期点 $x^0$

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 $f(x) = x^4 - 4x^2$

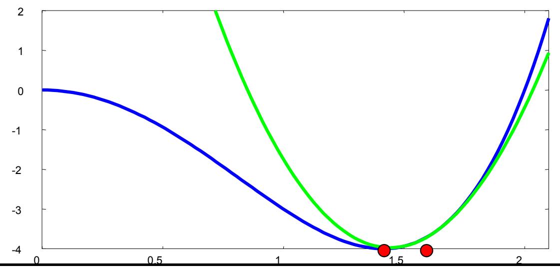
初期点 $x = 2$ においてfのテイラー近似を求める
⇒ $f(2+d) \fallingdotseq 0 + 16d + (40/2)d^2$
$d = -2/5$ のとき最小
⇒ 次の点は $x = 2 - 2/5 = 8/5$




## 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]

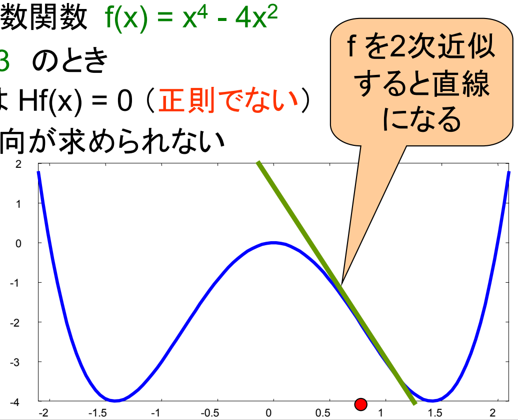
■ ヘッセ行列が正則でないと破綻

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]

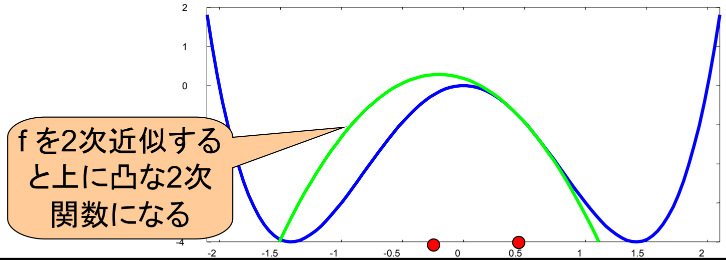
● ヘッセ行列が正定値でない場合には

目的関数値が増加する可能性あり

初期点 $x = 1/2$ のとき

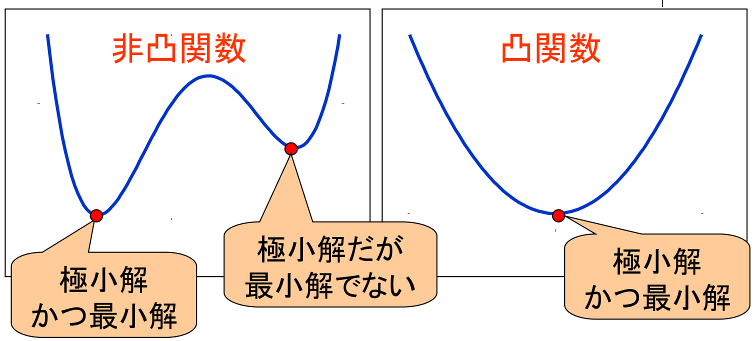
⇒ ヘッセ行列は $Hf(x) = -5$ (正定値でない)

⇒ ニュートン方向に進むと関数値が増加する




## 凸関数 [p.93]

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 [p.94]

定義:関数 f は凸関数

⇔ 任意のベクトル x, y

および任意の 0≦t≦1 に対して

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

注:教科書の定義と異なります

定義:関数 f は狭義凸関数

⇔ 任意の異なるベクトル x, y

および任意の 0<t<1 に対して

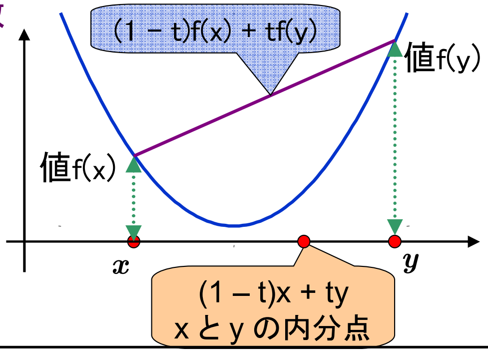
$$(1 - t)\, f(x) + t\, f(y) > f((1 - t)\, x + t\, y)$$

## 凸関数の定義（続き） [p.94]

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

狭義凸関数 ⇔ $(1-t)f(x) + t f(y) > f((1-t)x + t y)$

狭義凸関数の例




## 凸関数の定義（続き） [p.94]

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

狭義凸関数 ⇔ $(1-t)f(x) + t f(y) > f((1-t)x + t y)$

狭義凸関数の例

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）

$V$ : 正定値行列 ⇒ 狭義凸関数

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

## 凸関数の定義（続き） [p.94]

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

狭義凸関数 ⇔ $(1-t)f(x) + t f(y) > f((1-t)x + t y)$

狭義凸関数の例

2次関数 $f(x) = a x^2$ $(a > 0)$ は狭義凸関数

（証明）任意の異なる x, y と $0 < t < 1$ に対して、

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

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

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

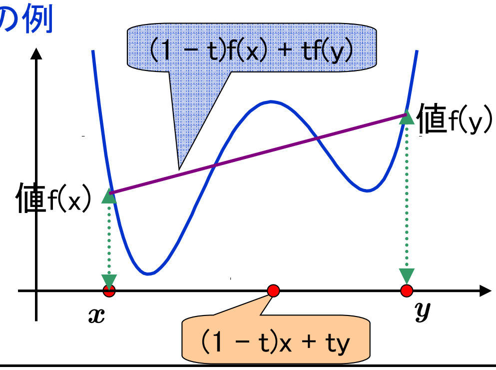
$= (t-t^2)\, a\,(x-y)^2$

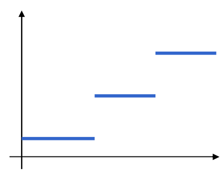
$> 0$ （$0 < t < 1$, $x \neq y$ より）

## 凸関数の定義（続き） [p.94]

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

非凸関数の例



## 凸関数の最適解の性質 [p.101]

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）

## レポート問題（今回で最後）

（1）関数 $f(x,y) = x^2 - xy + 2y^2 - 2x + e^{x+y}$
ついて次の問題を解きなさい.
(a)勾配ベクトルとヘッセ行列を計算せよ.
(b)原点(0,0)が極小解か否か, 最適性条件を用いて判定せよ.
(c)関数 f から最後の項 $e^{x+y}$ を削除したときの f に対し, すべての停留点を求めよ. さらに, 極小点, 極大点, 鞍点のいずれであるか, 2次の最適性条件を用いて判定せよ.

（2）関数 $f(x) = x^3 + 6x^2$ に対して
（a）初期点を x = 2 としてニュートン法を適用せよ。
（b）初期点を 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。

（3）関数 $f(x) = |x|$ は凸関数である. これを証明せよ. また, この関数は狭義凸ではない. 理由を説明せよ.