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

期末試験について

- 日時：２月５日（木）午後１時より
- 受験資格者：ネットワーク計画，非線形計画に関するレポートを一回以上提出した学生のみ
- 教科書等の持込は不可
- 座席は指定
- 試験内容：ネットワーク計画，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）

復習：勾配ベクトルに関する性質

制約なし最適化問題：　最小化　$f(x)$

定理（制約なし最適化問題の最適性条件）：

x^*: 制約なし問題の最適解

⇒　x^*は停留点 $\nabla f(x)=0$

勾配ベクトルと逆の方向に進むと関数値が減る

性質：任意のベクトル$y \in \mathbf{R}^n$ に対し，$\nabla f(y) \neq 0$ ならば十分小さい $\delta > 0$ に対して $f(y - \delta \nabla f(y)) < f(y)$

制約なし問題の解法1:最急降下法 [p.102]

最急降下法のアイディア:
勾配ベクトルと逆の方向に進むと関数値が減る

現在の点 x を x－α ∇f(x) により更新

ステップサイズ

⇒ 関数値 f(x) を減らしていく

ステップサイズの選び方:
次の一変数最適化問題を(近似的に)解く

最小化 f(x－α ∇f(x)) 条件 α >0

直線探索と呼ばれる

最急降下法のアルゴリズム [p.102]

入力：関数 f とその勾配ベクトル ∇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：最急降下方向 $-\nabla f(x^k)$ を計算

ステップ3：直線探索問題

最小化　$f(x^k - \alpha \nabla f(x^k))$　条件　$\alpha > 0$

を解き、解を α^k とする

ステップ4： $x^{k+1} = x^k - \alpha^k \nabla f(x^k)$ とおく

ステップ5： $k=k+1$ として、ステップ1に戻る

最適解の判定 [p.104,105]

- 非線形計画問題では

  最適解を正確に求めることは困難

⇒ 最適解に十分近い解（近似最適解）を求める

> 例: $f(x) = x^4 - 4x^2$
>
> この関数を最小にする x は 0, $\pm\sqrt{2}$
>
> 無理数をコンピュータで表現することは不可能

- 最適解に十分近いことをどうやって判定する？

（方法1）最適解 x^* に対し $\|\nabla f(x)\| = 0$ が成り立つ

⇒ $\|\nabla f(x)\|$ の値が十分小さくなったら終了

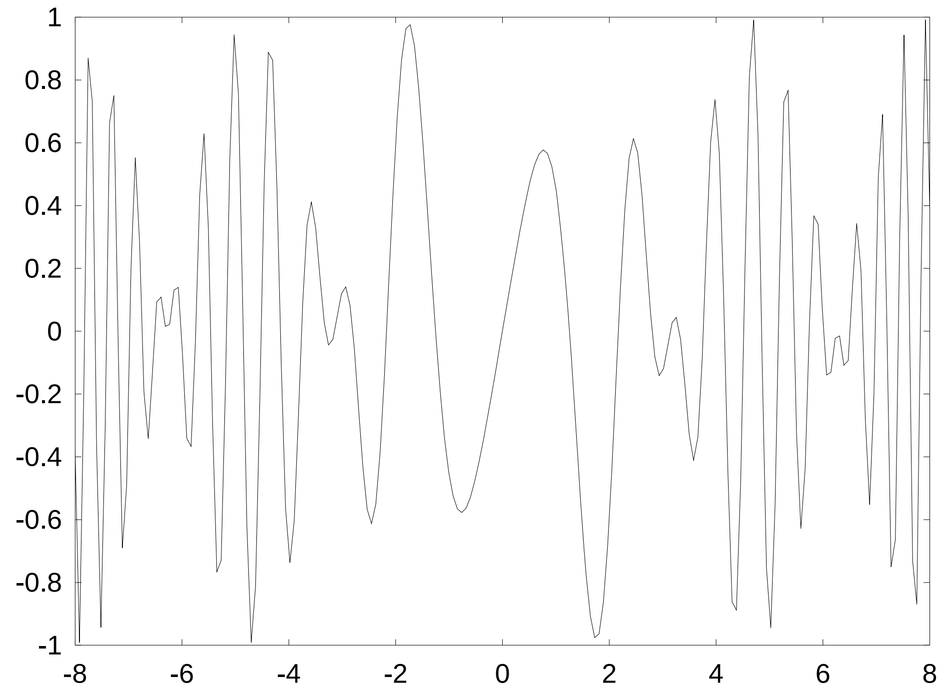
（方法2）最適解の近くでは x^k があまり変化しない

⇒ $\|x^{k+1} - x^k\|$ の値が十分小さくなったら終了

最適解の判定（つづき）[p.104,105]

- 非線形計画問題では
  近似最適解すら求めることが困難なことが多い

極小解または停留点を
求めることで我慢する



- 極小解は良い解であること
  が多い
- ある種の非線形関数では
  極小解⇔最小解

定理：ある仮定の下では、
　　最急降下法の求める点列は停留点に収束する

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

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

ヘッセ行列（続き）[p.89]

例：

$$f_1(x) = x^2 \qquad \nabla f_1(x) = f_1'(x) = 2x,\ \mathrm{H} f_1(x) = f_1''(x) = 2$$

$$f_2(x_1,\ x_2) = \sin x_1 + \cos x_2 \quad \Rightarrow \quad \nabla f_2(\boldsymbol{x}) = \begin{pmatrix} \cos x_1 \\ -\sin x_2 \end{pmatrix}$$

$$\Rightarrow \quad \mathrm{H} f_2(\boldsymbol{x}) = \begin{pmatrix} -\sin x_1 & 0 \\ 0 & -\cos x_2 \end{pmatrix}$$

$$f_3(x_1,\ x_2) = x_1 x_2 + \log x_1$$

$$\Rightarrow \quad \nabla f_3(\boldsymbol{x}) = \begin{pmatrix} x_2 + 1/x_1 \\ x_1 \end{pmatrix} \quad \Rightarrow \quad \mathrm{H} f_3(\boldsymbol{x}) = \begin{pmatrix} -1/x_1^2 & 1 \\ 1 & 0 \end{pmatrix}$$

ヘッセ行列とテイラー展開[p.90]

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（a は定数ベクトル）

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

この部分が二次関数, f(x) を近似

関数 $\psi(x-a) = \varphi(x_1 - a_1, x_2 - a_2, \ldots, x_n - a_n)$ は $x_1 - a_1, x_2 - a_2, \ldots, x_n - a_n$ に関する３次以上の項からなる n 変数多項式関数

(定数項，一次，二次の項は全く含まれない)

ヘッセ行列とテイラー展開（続き）[p.90]

例1：　$f_1(x) = x^2$　　$\nabla f_1(x) = 2x$　　$\mathrm{H}\, f_1(x) = 2$

$$f(x) = a^2 + (2a)(x-a) + \frac{1}{2}\cdot 2(x-a)^2 + \psi(x-a)$$
$$= x^2 + \psi(x-a)$$
$$\therefore\ \psi(x-a) = 0$$

※一般に、2次関数　$f(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^T V \boldsymbol{x} + \boldsymbol{c}^T\boldsymbol{x} + c_0$

の2次のテイラー展開において，
ψ(x-a) = 0

V: n×n行列
c: n 次元ベクトル
c_0: 定数

ヘッセ行列とテイラー展開（続き）[p.90]

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

$\nabla f(x) = 5x^4 - 15x^2 + 4$

$\mathrm{H} f(x) = 20x^3 - 30x$

a = -1 のとき

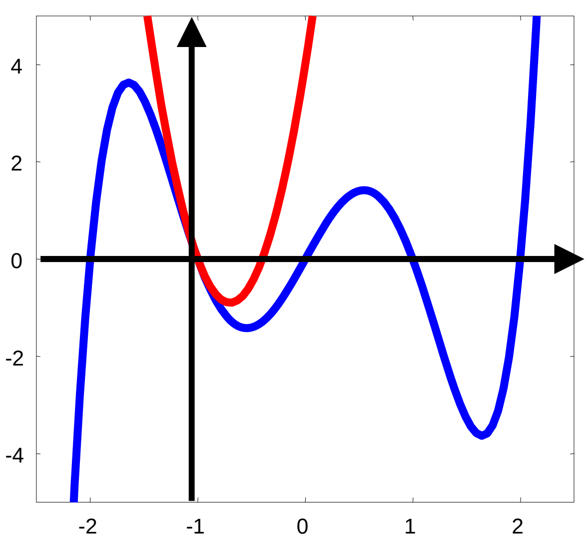
$$0 - 6(x+1) + 5(x+1)^2$$

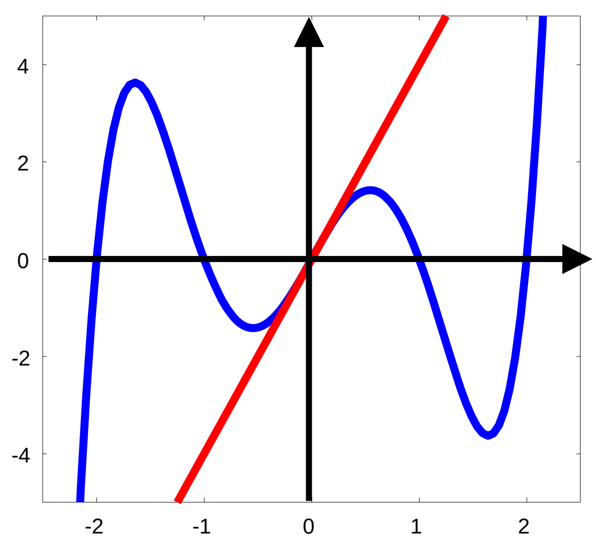
a = 0 のとき

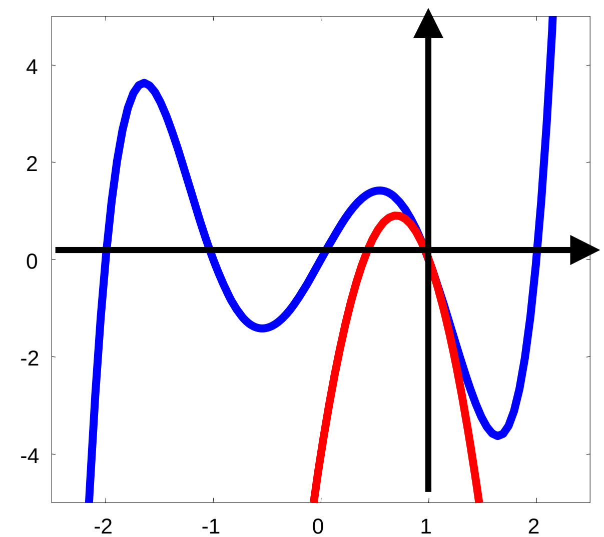
$$0 + 4x + 0x^2$$

a = 1 のとき

$$0 - 6(x-1) - 5(x-1)^2$$







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

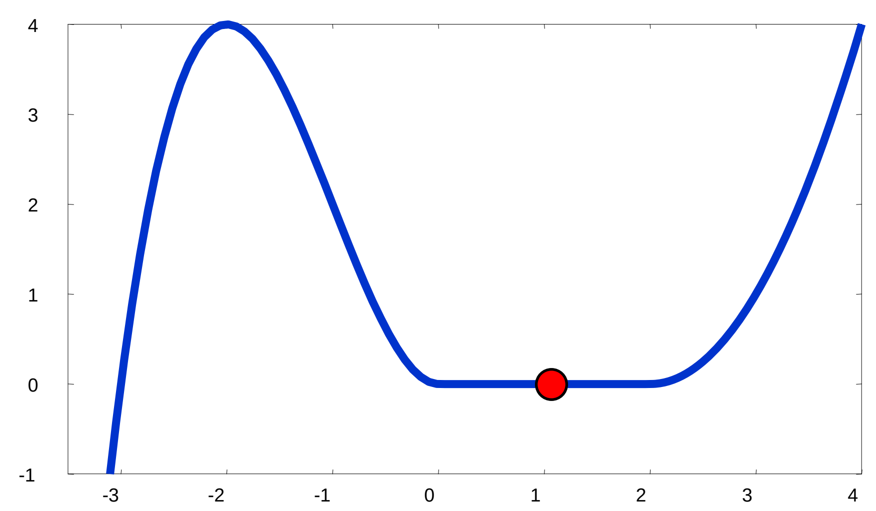
2次の最適性条件（必要条件）[p.99]

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 $f(x) = 0$

⇒ $\nabla f(x^*) = f'(x^*) = 0$
$Hf(x^*) = f''(x^*) = 0$
半正定値



２次の最適性条件（十分条件） [p.100]

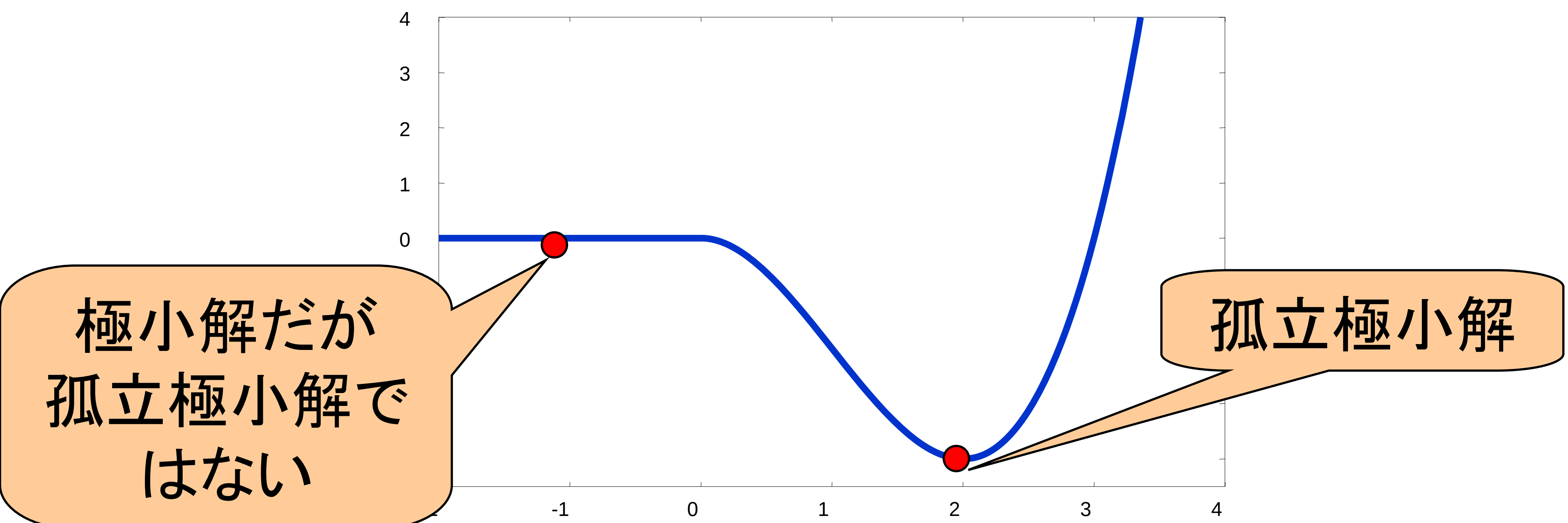
定理（２次の十分条件）：

x^* は停留点，　$Hf(x^*)$ は正定値

　$\Rightarrow$　x^*: 制約なし問題の（孤立）極小解

定義：x^* は孤立極小解

$\Leftrightarrow$　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

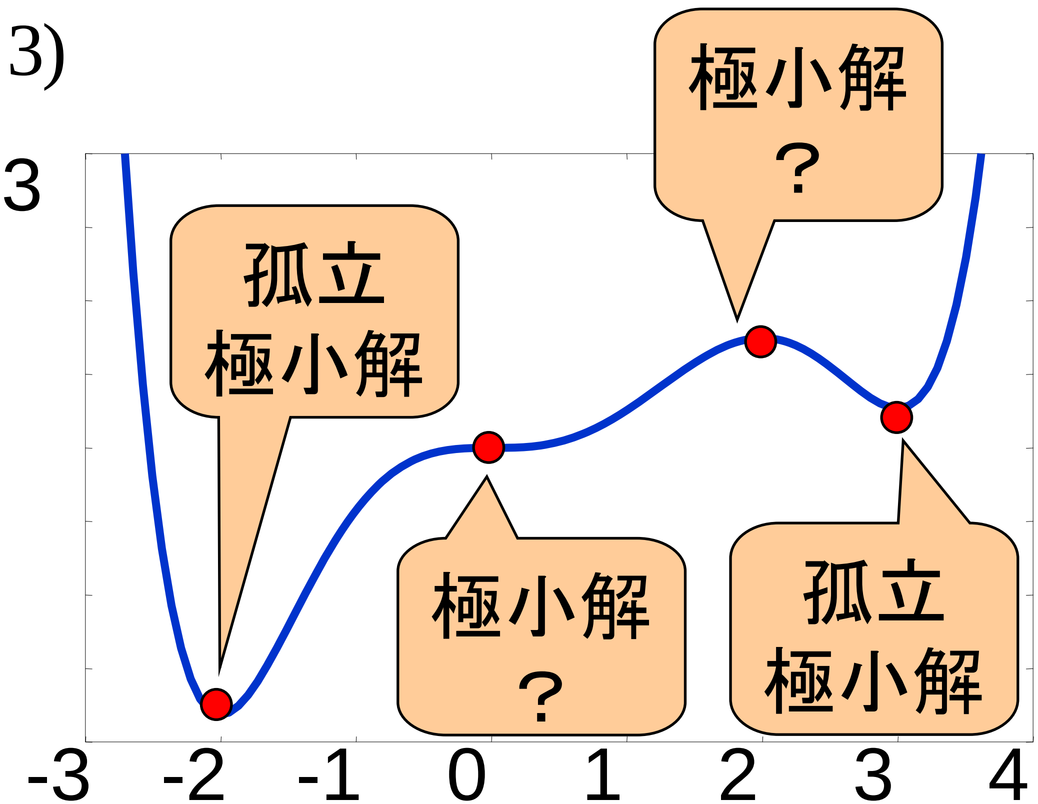
$$\mathrm{H}f(x) = 5x^4 - 12x^3 - 12x^2 + 24x$$

⇒ 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

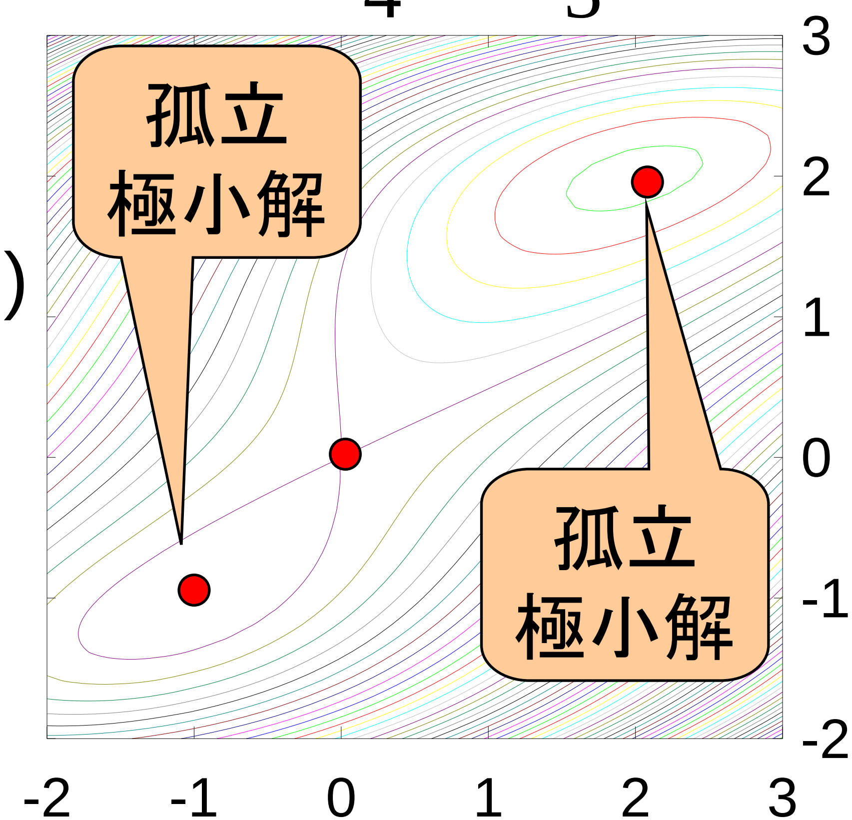
例2（教科書の例3．4）： $f(\boldsymbol{x}) = x_1^2 - 2x_1x_2 + \frac{1}{4}x_2^4 - \frac{1}{3}x_2^3$

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

⇒ 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

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 $d = -Hf(x)^{-1} \nabla f(x)$

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

レポート問題（今回で最後です）

問題 1：関数 f_1, f_2 に対し，x＝（1，2）における2次のテイラー展開を求めなさい．

$$f_1(x_1,\ x_2) = x_1 \log x_2 - x_2 \log x_1 \qquad f_2(x_1,\ x_2) = x_1^2 + x_2^2 - 1$$

問題2:

関数 $f(x_1, x_2) = x_1^3 + 2x_1x_2^2 - 4x_2^2 + x_2^3$ について考える．

(1) 関数 f の勾配ベクトル $\nabla f(x_1, x_2)$ を計算せよ．

(2) 任意のベクトル (a_1, a_2) に対し，$\nabla f(a_1, a_2) = (d_1, d_2)$ とする．このとき，$(d_1, d_2) \neq (0, 0)$ ならば，十分小さい $\delta > 0$ に対して $f(a_1 - \delta d_1, a_2 - \delta d_2) < f(a_1, a_2)$ が成り立つことを証明せよ．

（ヒント：勾配ベクトルの性質の証明を参考にしてください）

レポート問題（4点満点）

問題3： 関数 $f(x,y) = x^2 - xy + 2y^2 - 2x + e^{x+y}$

に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.

問題5： 対称な2×2行列 A に対し、次の関係を証明せよ。

Aは半正定値 ⇔ $a_{11} \geqq 0$, $a_{22} \geqq 0$, $a_{11}a_{22} - a_{12}{}^2 \geqq 0$

（ヒント：教科書の問題3.7の答えを参考にせよ）

先週の演習問題の答え

問題3： 関数 $f(x,y) = (x-2)^4 + (x-2y)^2$ に対して、初期点を(0, 3) として最急降下法を適用せよ。資料に添付してある等高線の図を使って実行すること.（数値はおおまかに計算すればよい）

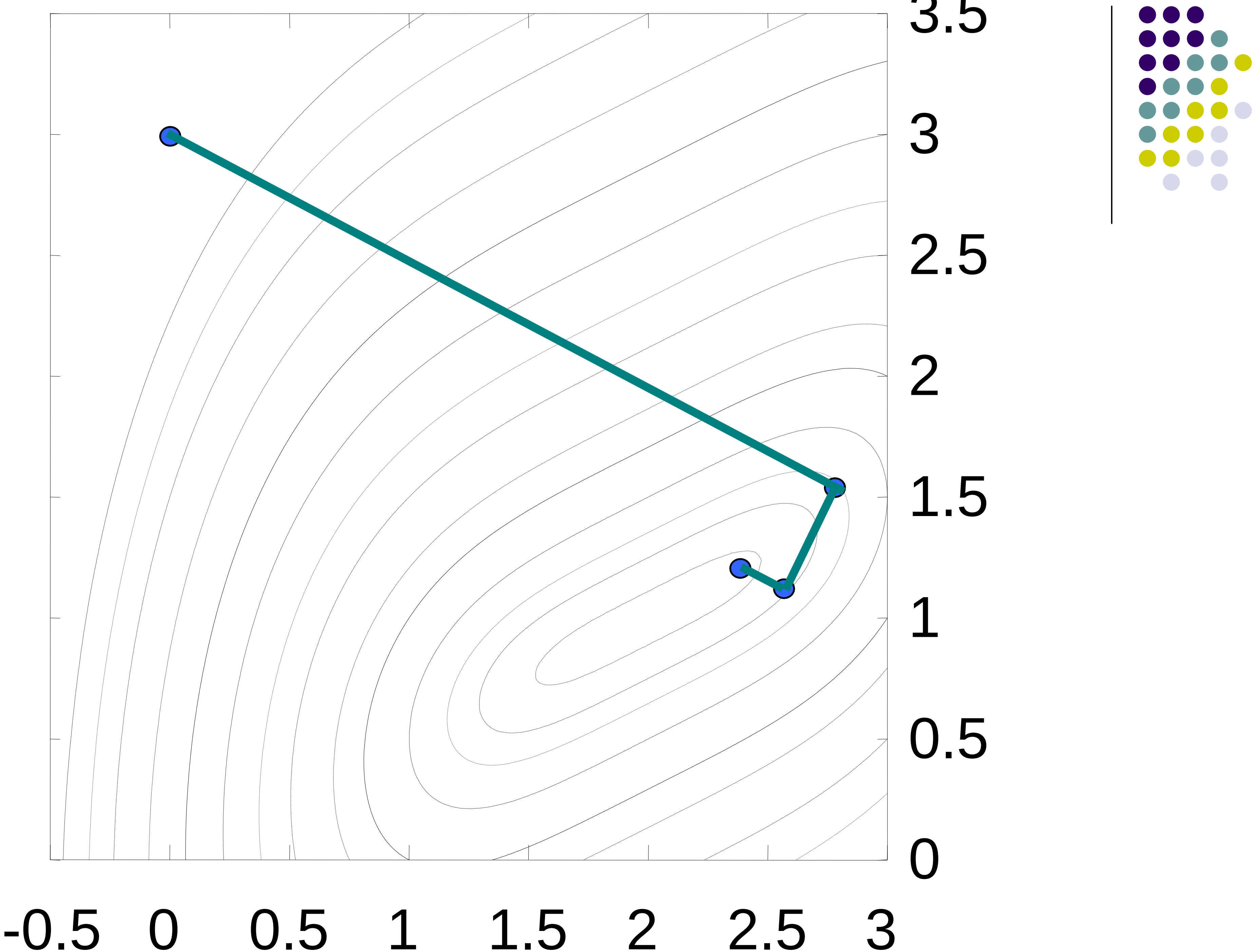
ポイント: 点の動きを表す折れ線の角度は必ず90度

点の動きは次の通り

(0.00, 3.00)➔(2.70, 1.51)

➔(2.52, 1.20)➔(2.43, 1.25)

➔…


3.5
3
2.5
2
1.5
1
0.5
0
-0.5
0
0.5
1
1.5
2
2.5
3