数理計画法（数理最適化）第11回

非線形計画その1

一次の最適性条件と最急降下法

担当：　塩浦昭義
(情報科学研究科　准教授)
shioura@dais.is.tohoku.ac.jp

非線形計画問題とは？

目的関数や制約条件が必ずしも線形でない数理最適化問題

例1：長方形の外周最小化問題

最小化 $2x + 2y$

条件 $xy \geqq 1$

$x, y \geqq 0$

非線形の制約条件

例2：線形制約つき関数最大化問題

最大化 $-3x^2 + 5y^3 + 2xy^2 - 4y$

条件 $3x + 5y \leq 12$

$7x + 2y \leq 8$

$x, y \geqq 0$

非線形の目的関数

制約なし問題 (unconstrained problem)

最小化 $f(x)$ 条件 なし

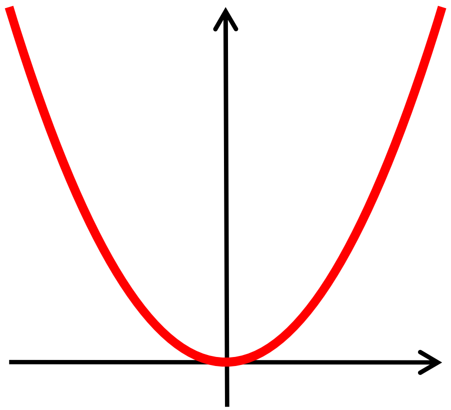
制約つき問題 (constrained problem)

最小化 $f(x)$ 条件 $g_i(x) \leqq 0 \ (i = 1, 2, \ldots, m)$

この講義では、制約なし問題を主に扱う

非線形関数の例(その1)

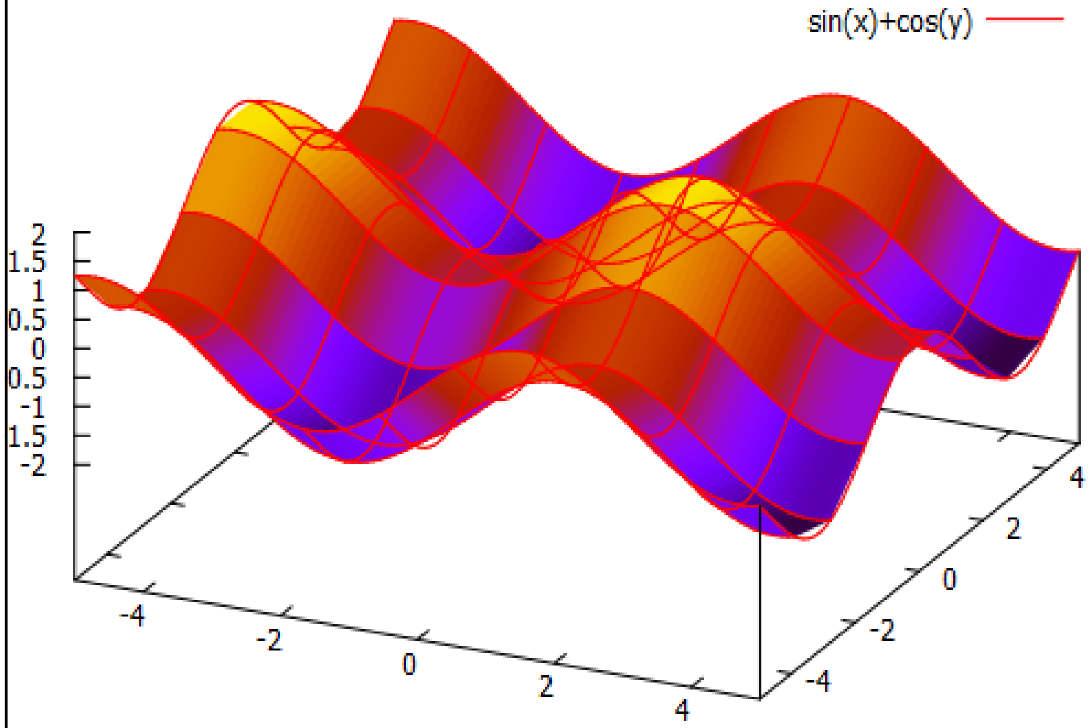
非線形関数 --- 線形でない関数



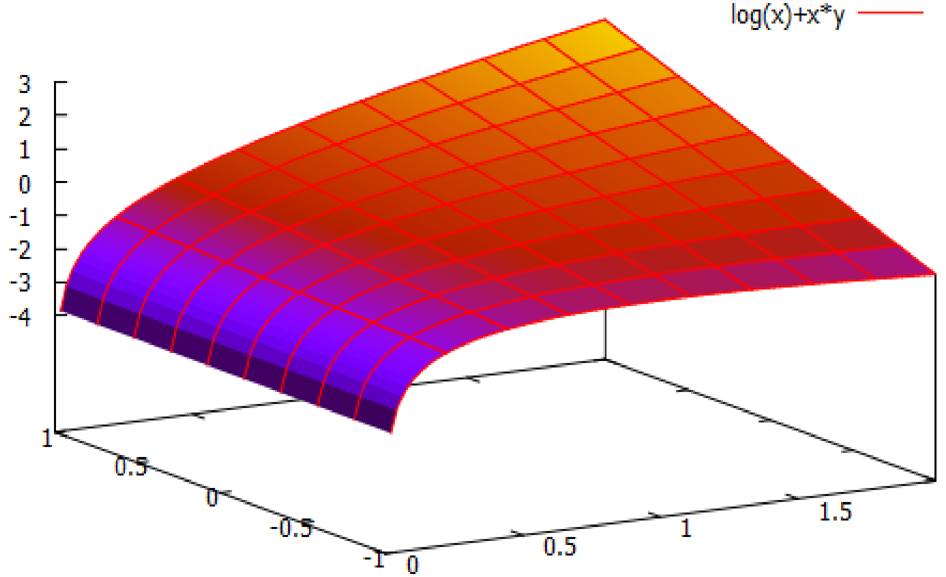
$$f_1(x) = x^2$$

微分可能な非線形関数の例

$$f_2(x_1, x_2) = \sin x_1 + \cos x_2$$



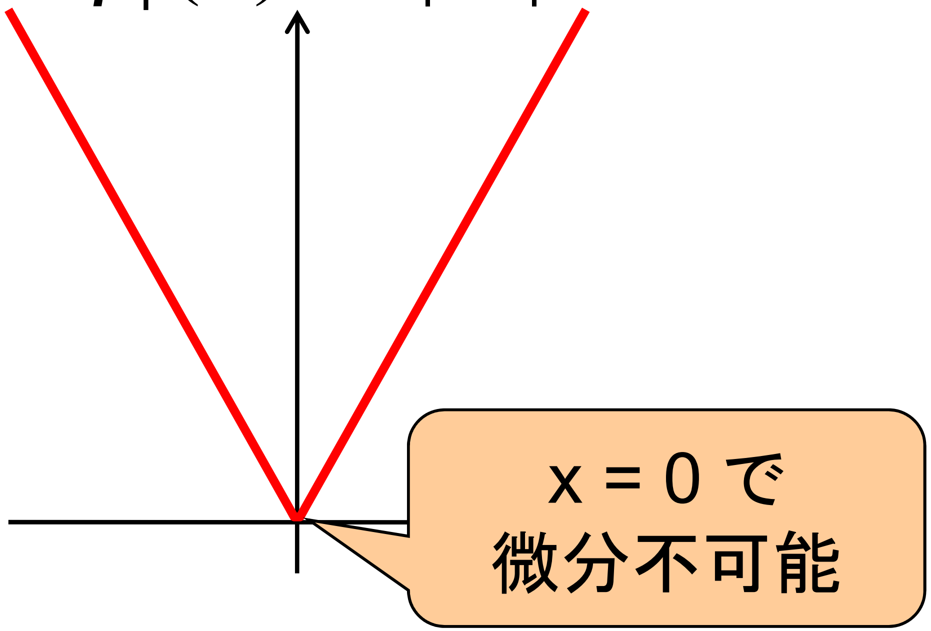
$$f_3(x_1, x_2) = x_1 x_2 + \log x_1$$



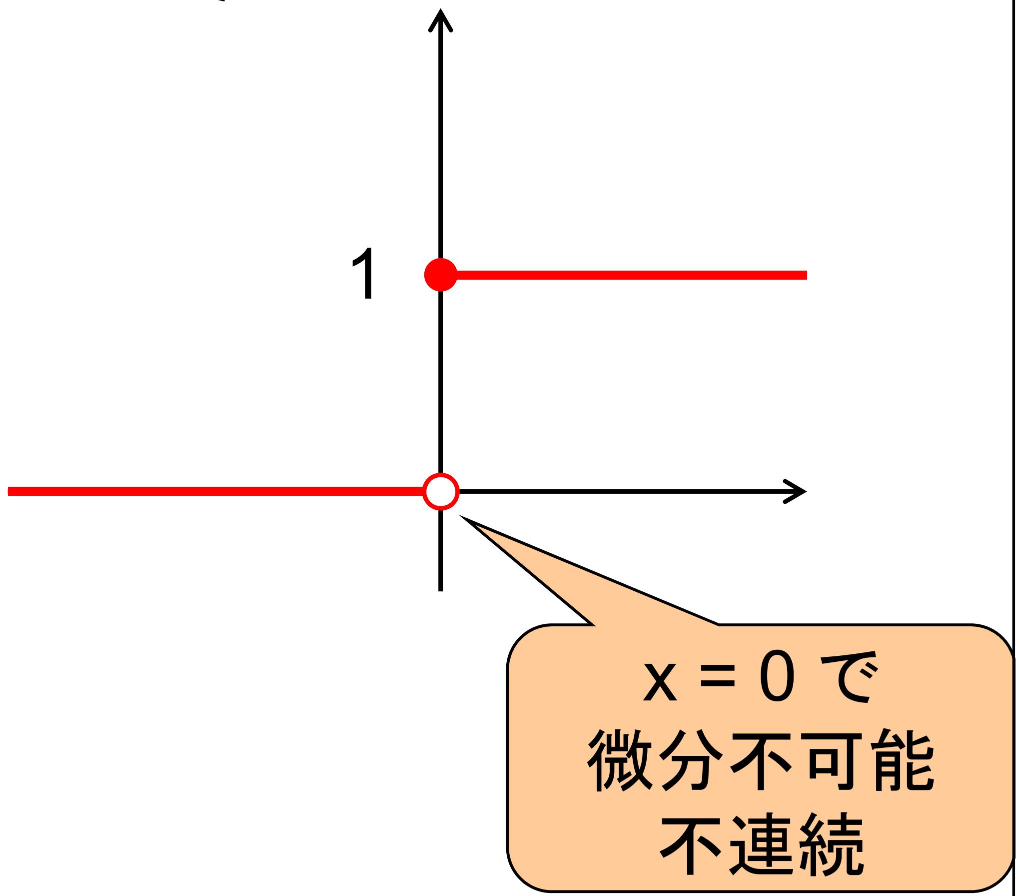
非線形関数の例(その2)

微分不可能な非線形関数の例

$$f_4(x) = |x|$$



$$f_5(x) = \begin{cases} 1 & (x \geq 0 のとき) \\ 0 & (x < 0 のとき) \end{cases}$$



この授業:
主に2回微分可能な関数を扱う

制約つき問題の制約なし問題への帰着

制約つき問題 $\boxed{\text{最小化}\ f(x)\quad \text{条件}\ g_i(x) \leqq 0\ (i = 1, 2, \ldots, m)}$

➔ 制約なし問題 $\boxed{\text{最小化}\ f(x) + h(x)\quad \text{条件}\ \text{なし}}$

$$h(x) = \begin{cases} 0 & (x\text{が元の制約を満たすとき}) \\ M(\text{十分大きい数}) & (\text{それ以外のとき}) \end{cases}$$

- 関数hは微分不可能
  ➔この制約なし問題を直接解くことは実用上難しい
- 関数hを滑らかな関数で近似➔解きやすい制約なし問題
  これを繰り返し解いて，制約つき問題の（近似）最適解を求める
  ➔ペナルティ関数法，バリア関数法

勾配ベクトル

関数 f の勾配ベクトル(gradient vector)

$$\nabla f(\boldsymbol{x}) = \begin{pmatrix} \dfrac{\partial f}{\partial x_1} \\ \vdots \\ \vdots \\ \dfrac{\partial f}{\partial x_n} \end{pmatrix}$$

一変数関数の場合　$\nabla f(x) = f'(x)$

例:

$$f_1(x) = x^2 \quad \Rightarrow \quad \nabla f_1(x) = 2x$$

$$f_2(x_1, x_2) = \sin x_1 + \cos x_2$$

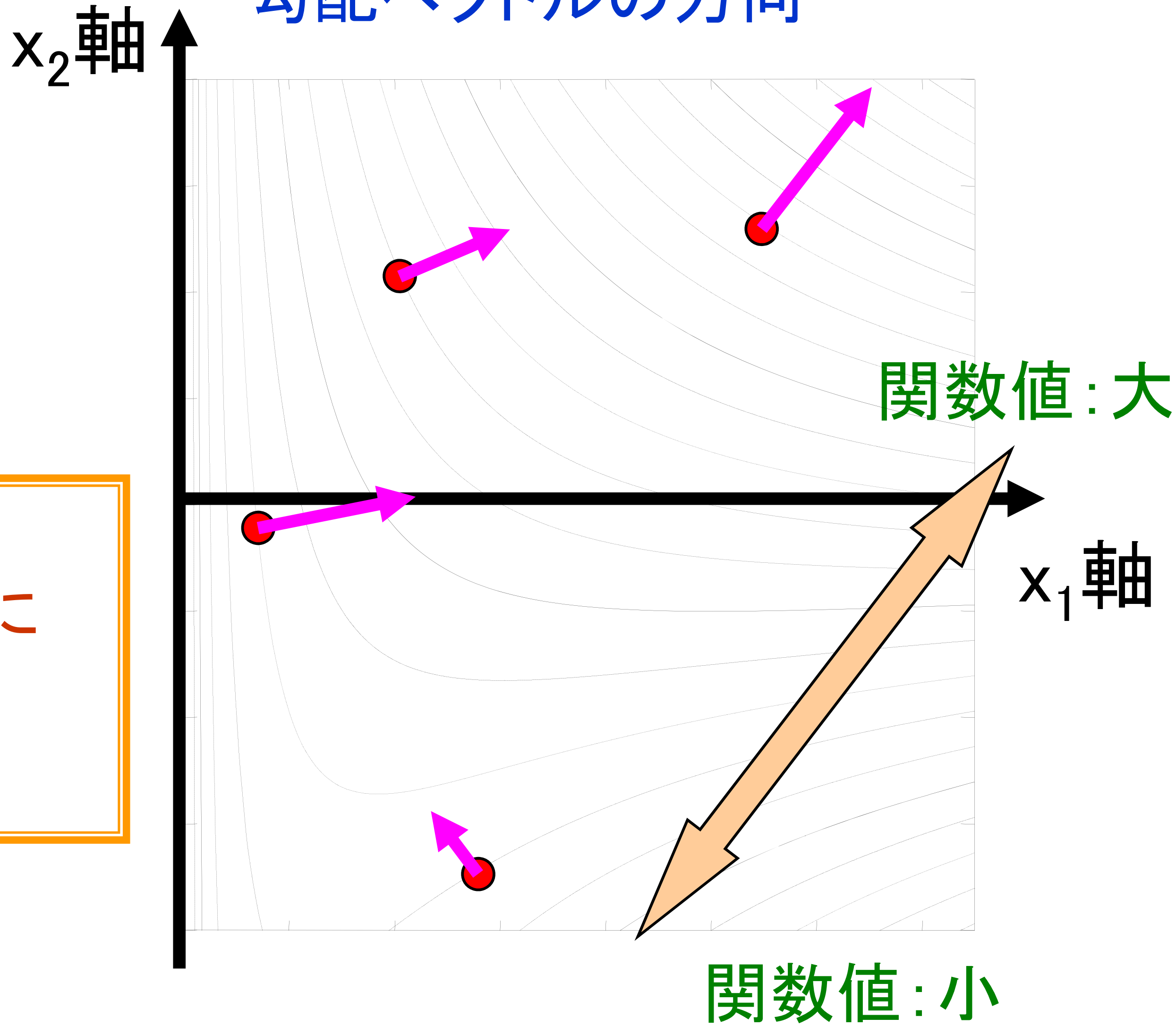
$$\Rightarrow \quad \nabla f_2(\boldsymbol{x}) = \begin{pmatrix} \cos x_1 \\ -\sin x_2 \end{pmatrix}$$

勾配ベクトル(続き)

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

$$\nabla f_3(\boldsymbol{x}) = \begin{pmatrix} x_2 + 1/x_1 \\ x_1 \end{pmatrix}$$

関数f_3の等高線と
勾配ベクトルの方向



勾配ベクトルのイメージ:

- 関数という山を登るときに
  最も急な方向
- 関数値が増加する方向

一次のテイラー展開

任意の関数 f はベクトル $a \in \mathbb{R}^n$ を使って
次の形に表現できる

$$f(x) = f(a) + \nabla f(a)^T (x - a) + \varphi(x - a)$$

関数$f(x)$の$x = a$における
一次のテイラー展開

関数$\varphi(x-a) = \varphi(x_1 - a_1, x_2 - a_2, \ldots, x_n - a_n)$ は
$x_1 - a_1, x_2 - a_2, \ldots, x_n - a_n$ に関する
2次以上の項からなるn変数多項式関数
(定数項, 一次の項は全く含まれない)

$x - a$ が十分小さいとき,
$\varphi(x - a)$は十分小さい

一次のテイラー近似

関数$f(x)$の$x = a$における
一次のテイラー展開

$$f(x) = \boxed{f(a) + \nabla f(a)^T (x - a)} + \varphi(x - a)$$

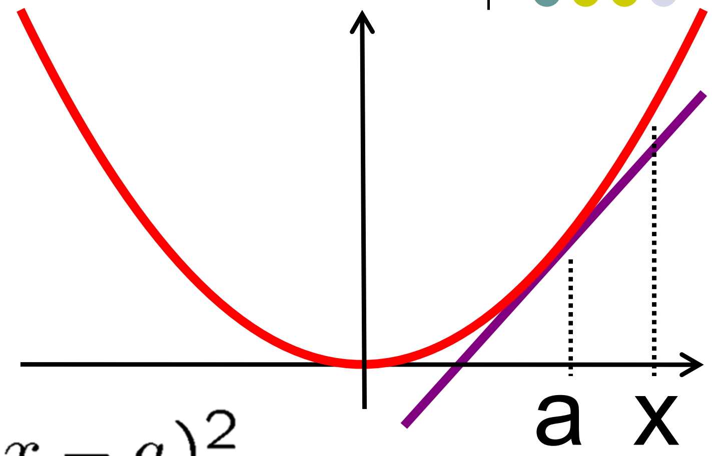
$x \simeq a$ のとき，$\varphi(x - a)$の値は他の項に比べて
十分小さい（0に近い） ➔無視できる

$$\tilde{f}(x) = f(a) + \nabla f(a)^T (x - a)$$

関数$f(x)$の$x = a$における
一次のテイラー近似

- 線形関数，傾き$= \nabla f(a)$
- $x \simeq a$のとき $\tilde{f}(x) \simeq f(x)$，
  とくに $\tilde{f}(a) = f(a)$

テイラー展開とテイラー近似の例



例1： $f_1(x) = x^2 \qquad f_1'(x) = 2x,$

テイラー展開

$$f(x) = a^2 + 2a(x-a) + \varphi(x-a)$$

ここで

$$\varphi(x-a) = f(x) - \{a^2 + 2a(x-a)\} = (x-a)^2$$

テイラー近似 $\widetilde{f_1}(x) = a^2 + 2a(x-a)$

例2： $f_2(x) = \log x \qquad f_2'(x) = 1/x$

テイラー展開

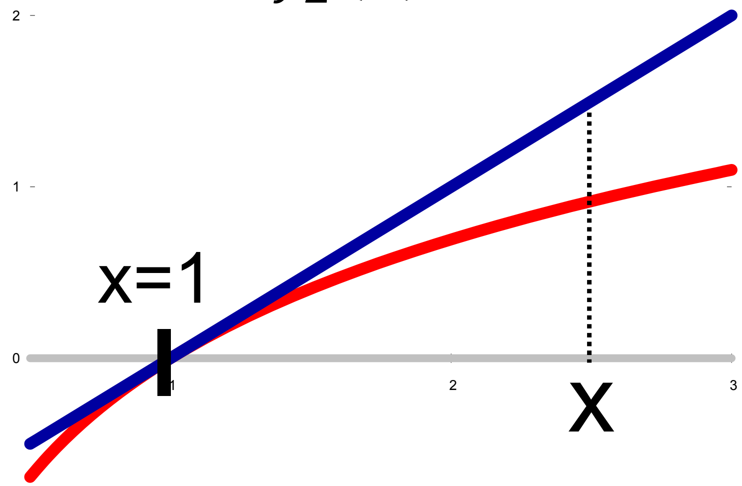
$$f_2(x) = 0 + \frac{1}{1}(x-1) + \varphi(x-1)$$

ここで

$$\varphi(x-1)$$

$$= -\frac{(x-1)^2}{2} + \frac{(x-1)^3}{3} - \frac{(x-1)^4}{4} + \cdots$$

テイラー近似 $\widetilde{f_2}(x) = x - 1$



勾配ベクトルの性質

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$y \in \mathbb{R}^n$に対し, $\nabla f(y) \neq 0$ならば,
十分小さい$\delta > 0$に対し, $f\big(y - \delta \nabla f(y)\big) < f(y)$成立

証明: $d = -\delta \nabla f(y)$ とおく（δ: 正の実数）

一次のテイラー展開において x = y + d, a = y とおくと,

$$
\begin{aligned}
f(y+d) &= f(y) + \nabla f(y)^T d + \varphi(d) \\
&= f(y) - \delta \|\nabla f(y)\|^2 + \varphi(-\delta \nabla f(y))
\end{aligned}
$$

φは2次以上の項からなる多項式関数

$\Rightarrow$ $\varphi(-\delta \nabla f(y))$ は δ に関する2次以上の項からなる一変数関数

勾配ベクトルの性質

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$y \in \mathbb{R}^n$に対し, $\nabla f(y) \neq 0$ならば,
十分小さい$\delta > 0$に対し, $f\big(y - \delta \nabla f(y)\big) < f(y)$成立

証明の続き:

$\varphi(-\delta \nabla f(y))$ は δ に関する2次以上の項からなる一変数関数

$\Rightarrow \varphi(-\delta \nabla f(y))/\delta$ は δ に関する1次以上の項からなる

$\therefore$ δ を十分小さくすると, $\varphi(-\delta \nabla f(y))/\delta$ は0 に近づく

$$\therefore\ -\delta ||\nabla f(y)||^2 + \varphi(-\delta \nabla f(y))$$
$$= -\delta \{||\nabla f(y)||^2 - \varphi(-\delta \nabla f(y))/\delta\} < 0$$
$$\therefore f(y - \delta \nabla f(y)) < f(y)$$

勾配ベクトルの性質

勾配ベクトルの方向に進むと関数値が増える

性質：任意の$y \in \mathbb{R}^n$に対し，$\nabla f(y) \neq 0$ならば，
十分小さい$\delta > 0$に対し，$f\big(y + \delta \nabla f(y)\big) > f(y)$成立

証明は省略（直前の性質と同様に証明できる）

最適性条件

非線形計画問題の最適性条件:
ベクトル x が最適解であるための必要条件(または十分条件)

定義: x は停留点 ⇔ $\nabla f(x)=0$

定理(制約なし問題の最適性条件):
　　x^*: 制約なし問題の最適解 ⇒ x^*は停留点

証明: 　$\nabla f(x^*)\neq 0$ と仮定

勾配ベクトルの性質より、十分小さい $\delta>0$に対して

$$f(\boldsymbol{x}^*-\delta\nabla f(\boldsymbol{x}^*)) < f(\boldsymbol{x}^*)$$

x^* が最適解であることに矛盾 　$\therefore \nabla f(x^*)=0$

最適性条件

定理(制約なし問題の最適性条件):

x*: 制約なし問題の最適解 ⇒ ∇f(x*) = 0

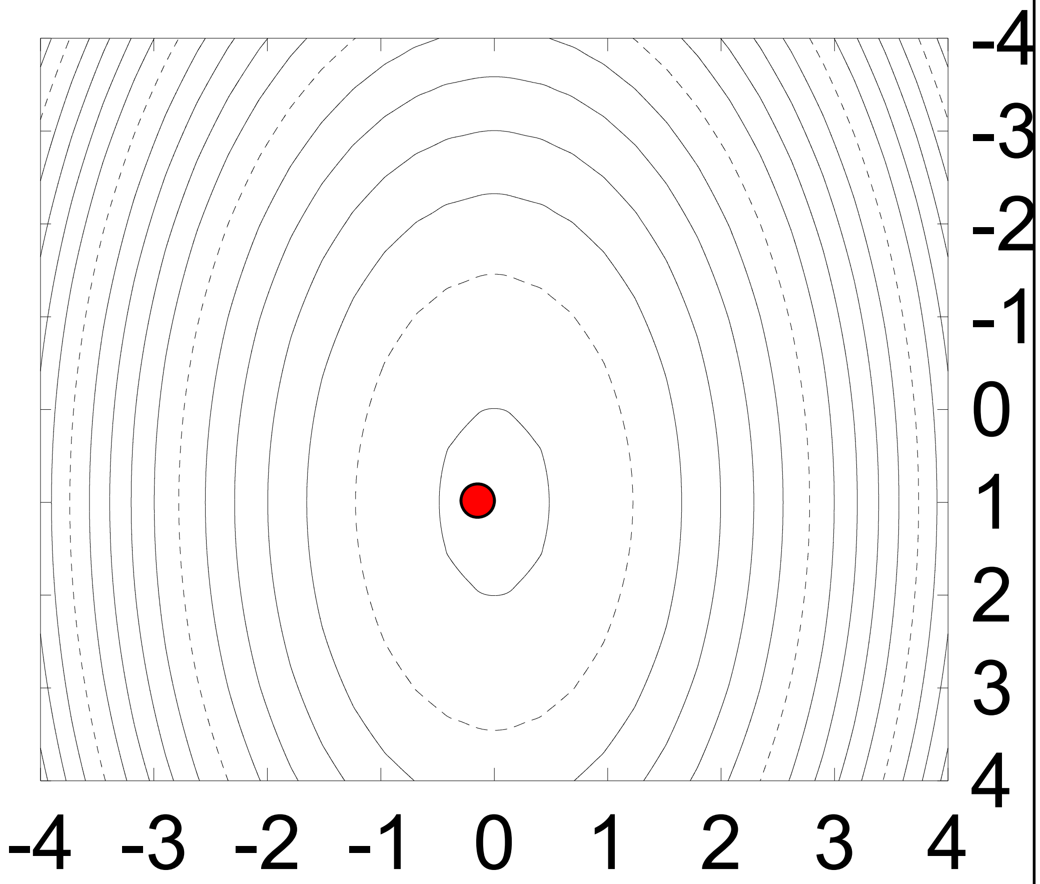
例: $f(\boldsymbol{x}) = x_1^2 - 2x_1 + 4x_2^2$

$$= (x_1 - 1)^2 + 4x_2^2 - 1$$

$(x_1,x_2) = (1,0)$ が最適解

$$\nabla f(\boldsymbol{x}) = \begin{pmatrix} 2x_1 - 2 \\ 8x_2 \end{pmatrix}$$

$$\nabla f(1,0) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$



最適性条件

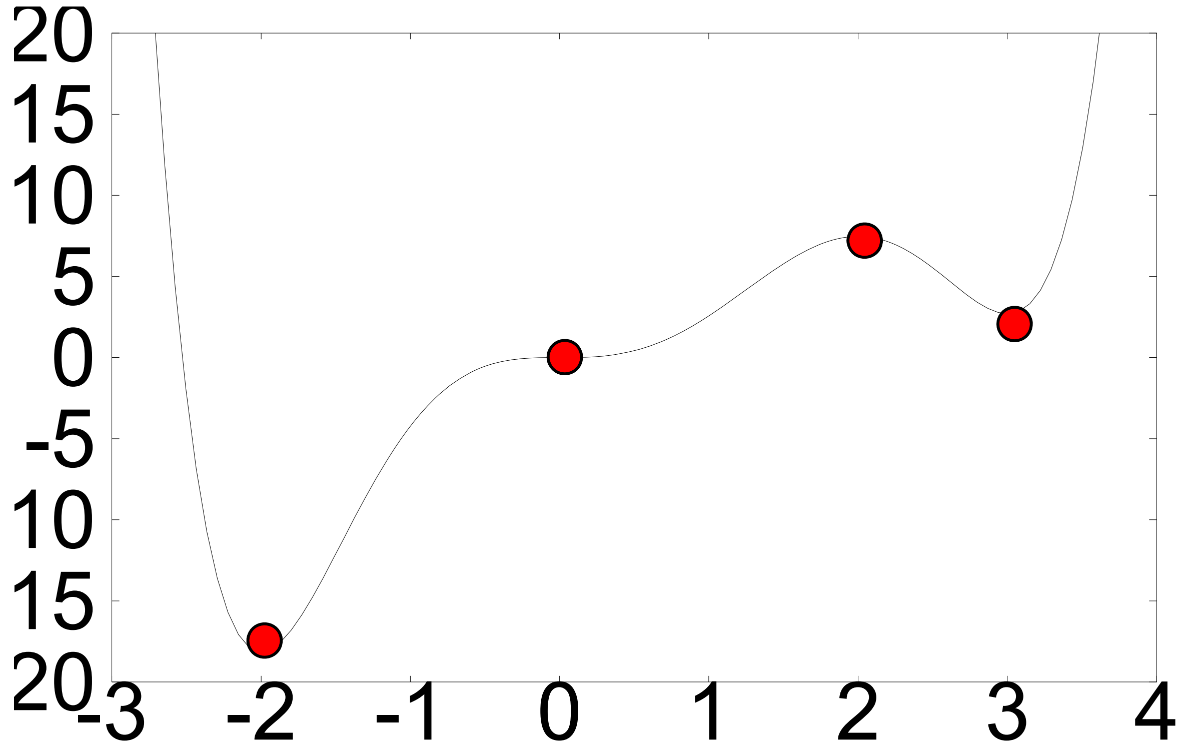
※「x*は停留点 ⇒ x*は最適解」は必ずしも成り立たない

例：

$$f(x) = \frac{1}{6}x^6 - \frac{3}{5}x^5 - x^4 + 4x^3$$

$$\nabla f(x) = (x+2)x^2(x-2)(x-3)$$

停留点はx = -2, 0, 2, 3
最適解はx = -2 のみ



極小解, 極大解, 鞍点

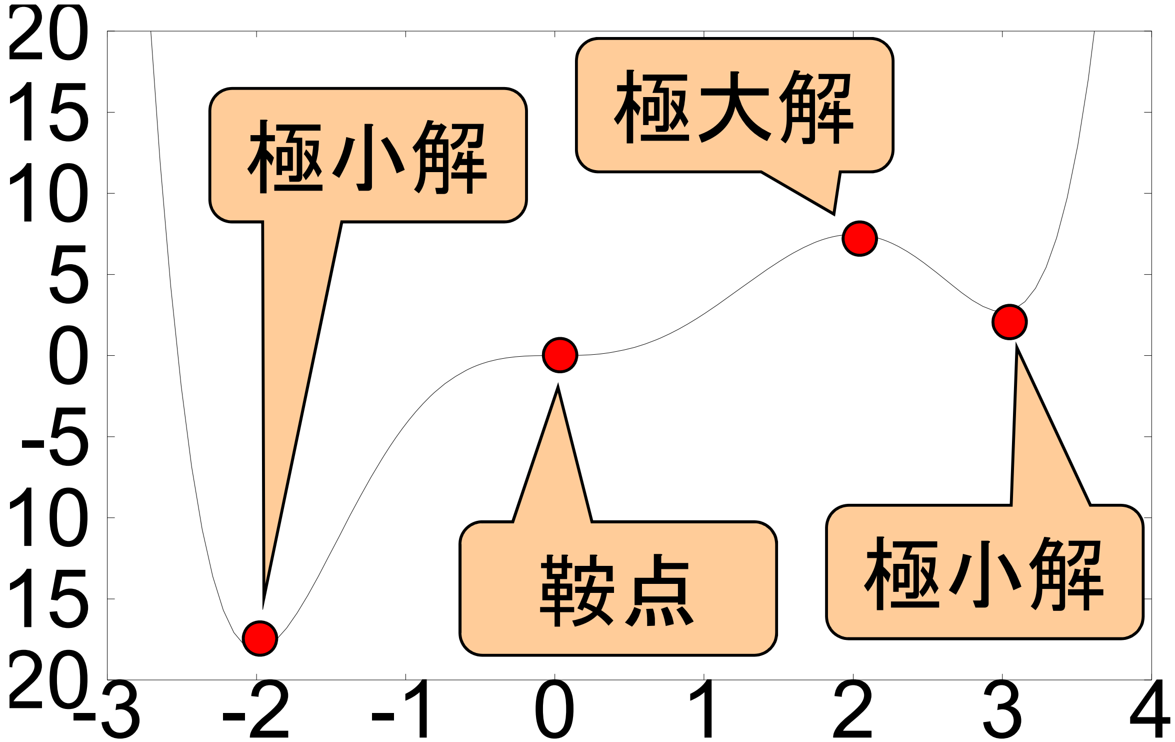
停留点 x* の分類

極小解: x* の付近だけに注目したとき、x* は最小

> あるδ＞0 が存在して, ||x − x*|| ≦ δを満たすすべての x に対して f(x) ≧ f(x*)

極大解 : x* の付近だけに注目したとき, x* は最大

鞍点 : 極小点でも極大点でもない停留点



制約なし問題の解法1：最急降下法

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 $x - \alpha \nabla f(x)$ 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　$f(x - \alpha \nabla f(x))$　条件　$\alpha > 0$

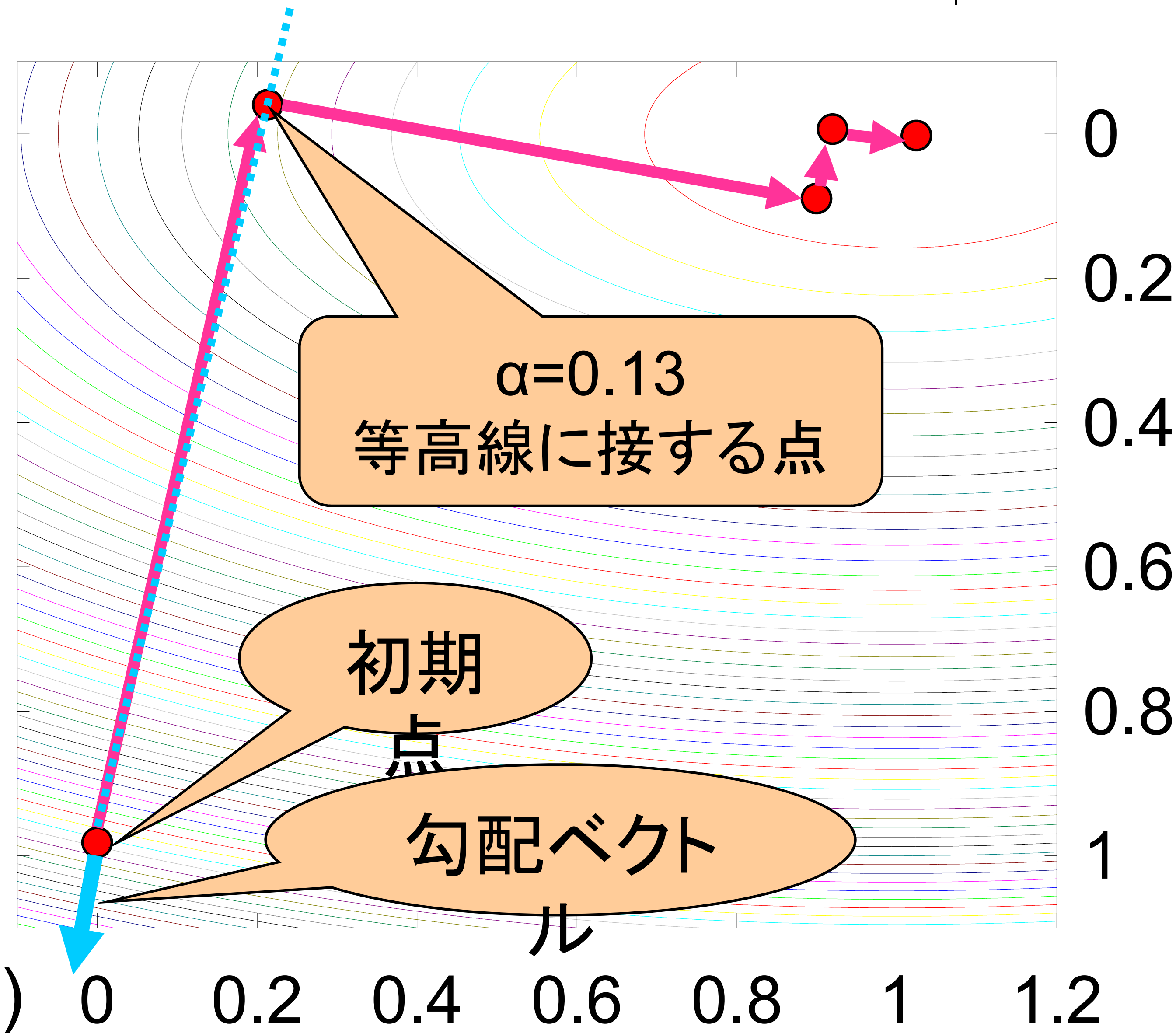
直線探索と呼ばれる

最急降下法の実行例

例: $f(\boldsymbol{x}) = x_1^2 - 2x_1 + 4x_2^2$

$$\nabla f(\boldsymbol{x}) = \begin{pmatrix} 2x_1 - 2 \\ 8x_2 \end{pmatrix}$$

- $(x_1, x_2) = (0,1)$ からスタート
- $\nabla f(0,1) = (-2,8)$
- $f(0 + 2\alpha, 1 - 8\alpha)$ を最小にするのは $\alpha=0.13$
- 次の点は $(x_1, x_2) = (0.26, -0.05)$



最急降下法のアルゴリズム

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

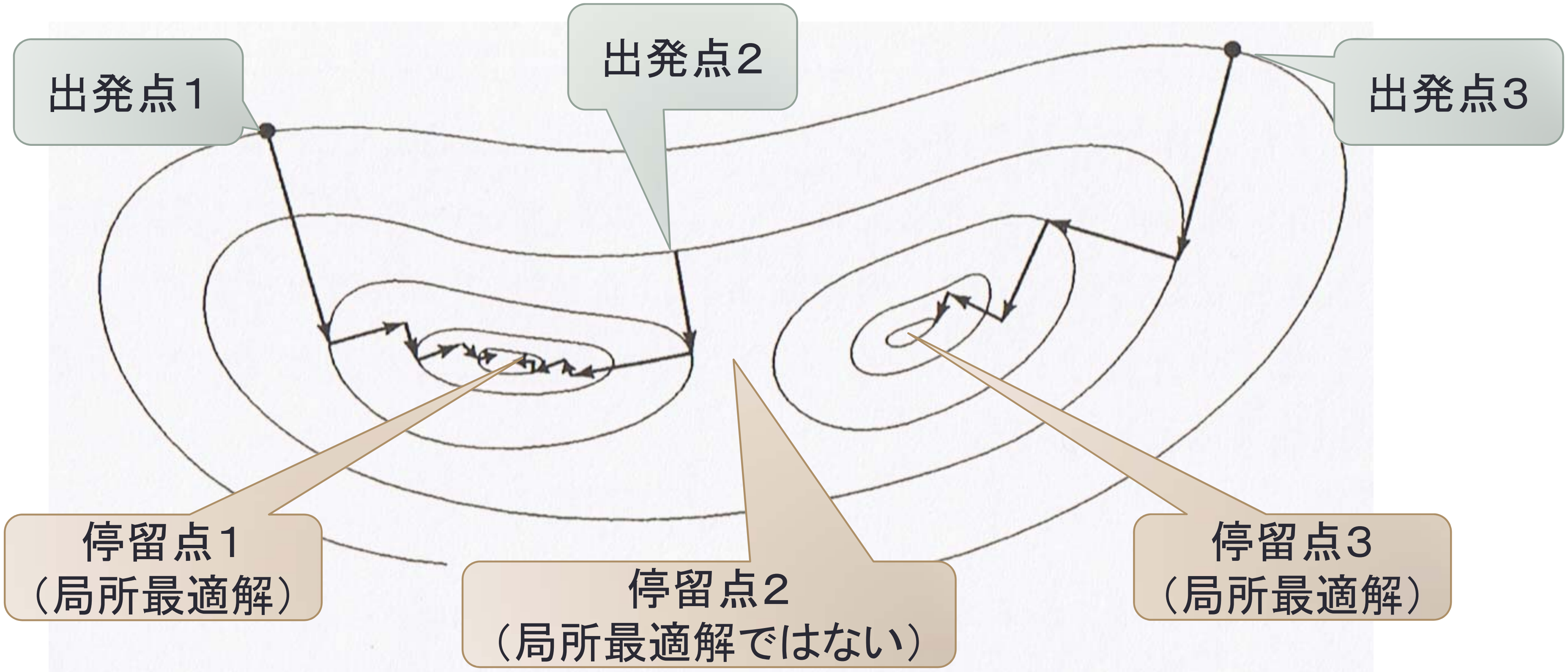
最小化　$f(x^k - \alpha \nabla f(x^k))$　条件　$\alpha > 0$

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

最急降下法の実行例その2



- 最急降下法は，必ず停留点（$\nabla f(x) = 0$となる点）に収束（大域的収束性）
  - 出発点の選び方次第では，局所的最適解に収束
  - 凸関数の場合，必ず大域的最適解に収束

最適解の判定

- 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正確に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

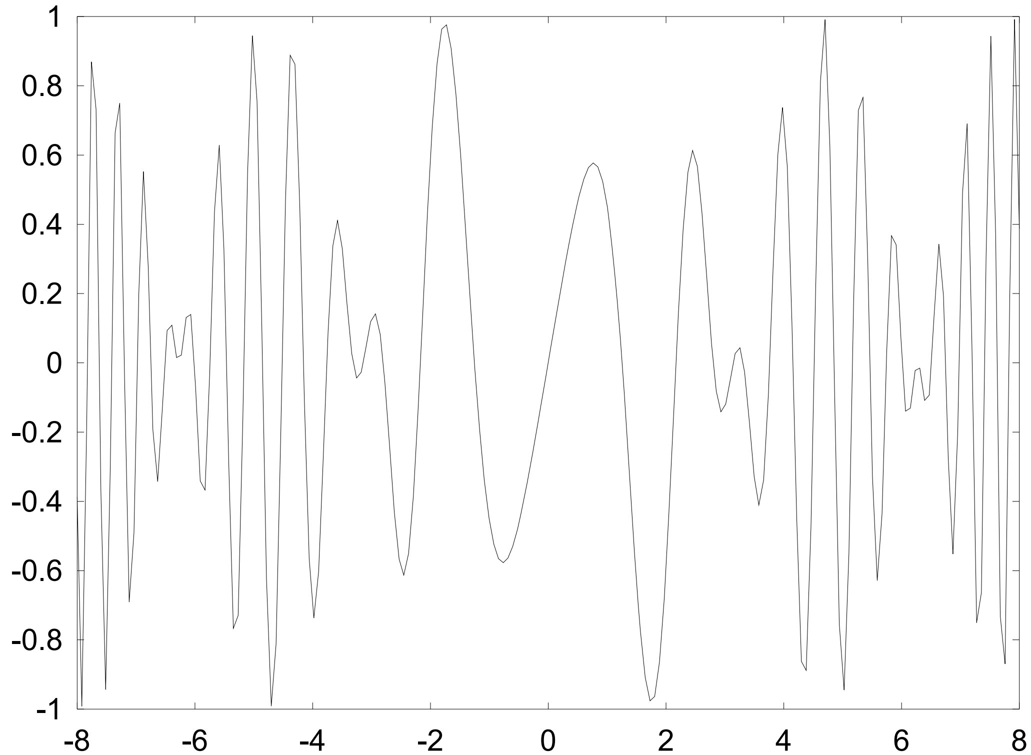
最適解の判定（つづき）

•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(凸関数)
では
極小解⇔最小解



定理:ある仮定の下で，最急降下法の求める点列は
停留点に収束する

演習問題

問題１:下記の4つの関数の勾配ベクトルを計算しなさい

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

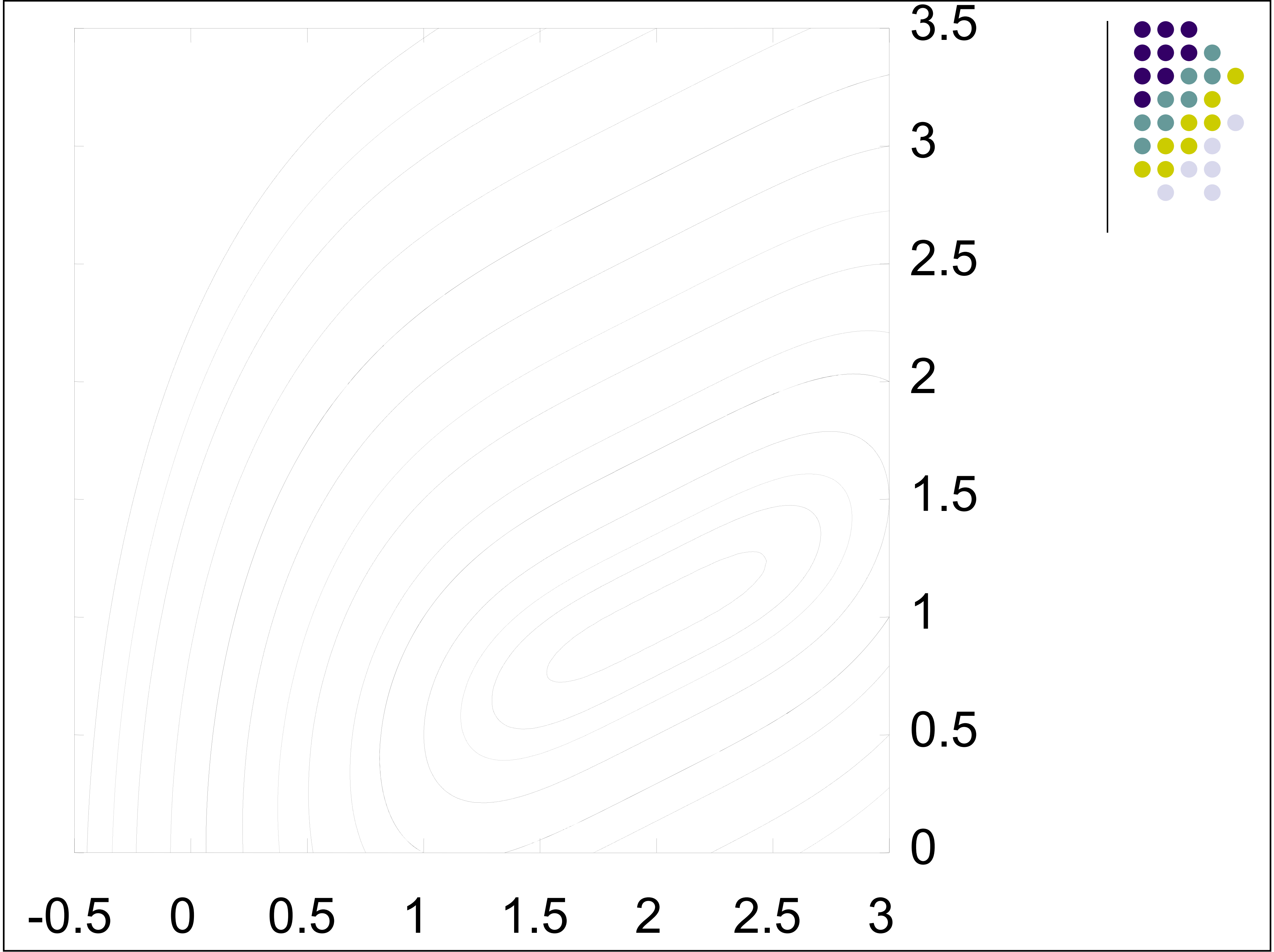
$$f_3(x_1,\ x_2)\ =\ x_1\ \log x_2 - x_2 \log x_1$$

$$f_4(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^T\boldsymbol{V}\boldsymbol{x} \quad \text{(ただし，}\boldsymbol{x}\text{は}n\text{次元ベクトル，}\boldsymbol{V}\text{は}n\times n\text{対称行列)}$$

問題2: 問題1の3つの関数 $f_1(x_1, x_2)$, $f_2(x_1, x_2)$, $f_3(x_1, x_2)$ に対して，(a_1,a_2) における一次のテイラー近似を求めなさい.

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.（具体的な数値は計算しなくてもよい）

ポイント: 点の動きを表す折れ線の角度は必ず90度



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