

数理計画法 第12回

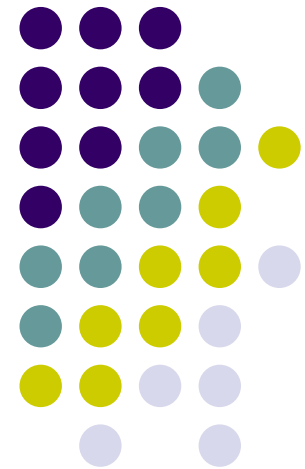
第3章 非線形計画

3.2 制約なし最適化

担当： 塩浦昭義

(情報科学研究科 助教授)

shioura@dais.is.tohoku.ac.jp



復習：非線形計画問題とは？



目的関数や制約式が**必ずしも線形でない**
数理計画問題

例：長方形の外周最小化問題

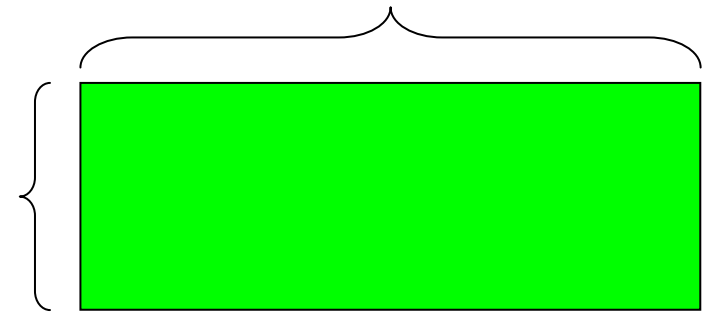
最小化 $2x + 2y$

条件 $xy \geq 1$

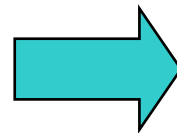
$x, y \geq 0$

x: 縦の長さ

y: 横の長さ



非線形の
不等式



非線形計画問題

復習：非線形計画問題の分類



制約なし最適化問題

入力： 目的関数 $f(x)$

問題： 最小化 $f(x)$ 条件 なし

制約つき最適化問題

入力： 目的関数 $f(x)$, 制約を表す関数 $g_i(x)$ ($i = 1, 2, \dots, m$)

問題： 最小化 $f(x)$ 条件 $g_i(x) \leq 0$ ($i = 1, 2, \dots, m$)

この講義では、**制約なし問題**を主に扱う

復習: 勾配ベクトル



関数 f の勾配ベクトル

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

一変数関数の場合は $\nabla f(x) = f'(x)$

例:

$$f_1(x) = x^2 \longrightarrow \nabla f_1(x) = 2x$$

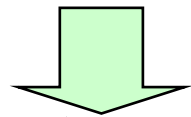
$$f_2(x_1, x_2) = \sin x_1 + \cos x_2$$

$$\longrightarrow \nabla f_2(\mathbf{x}) = \begin{pmatrix} \cos x_1 \\ -\sin x_2 \end{pmatrix}$$

復習: 勾配ベクトル

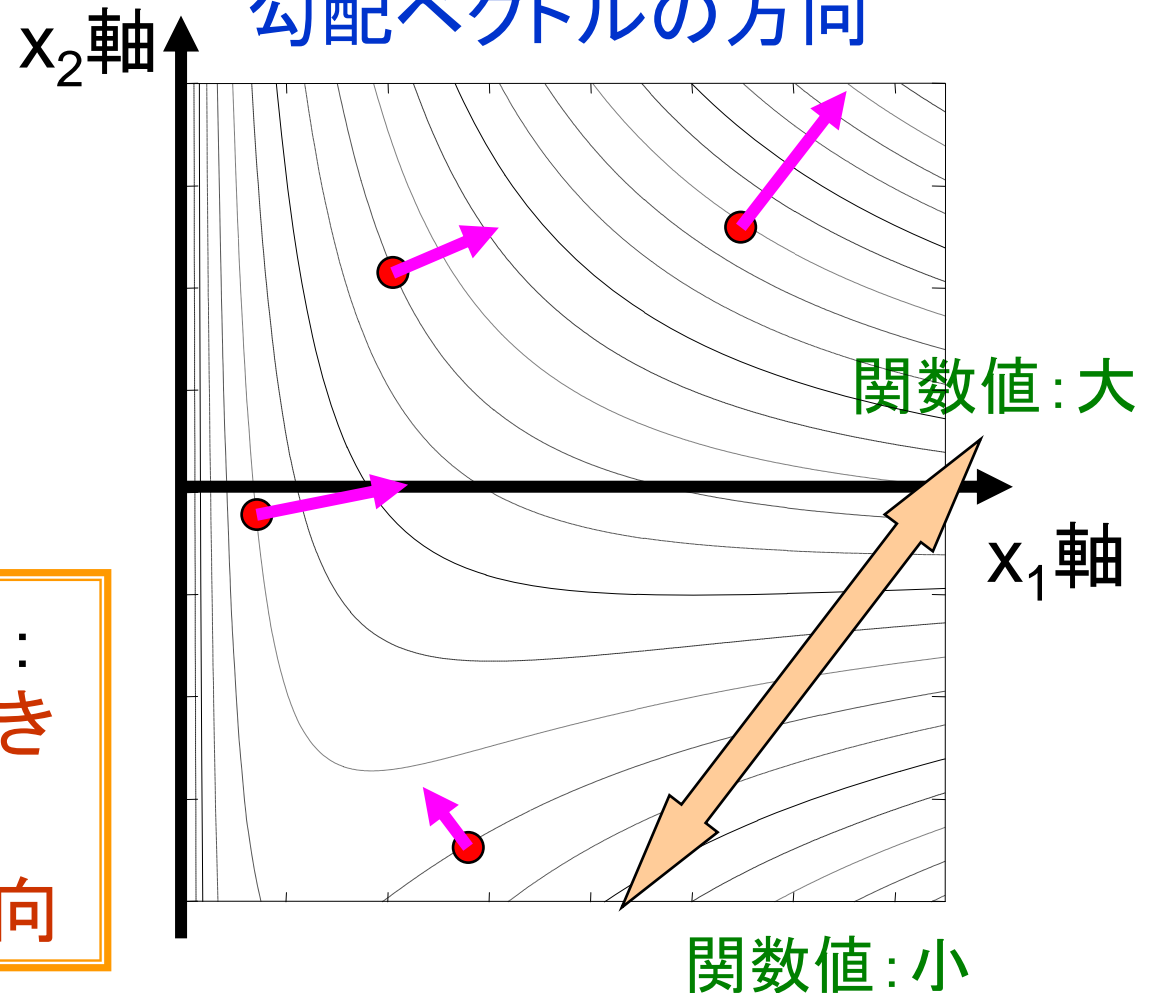


$$f_3(x_1, x_2) = x_1 x_2 + \log x_1$$



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

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[p.89]



任意の関数 f は次の形に表現できる (a は定数ベクトル)

関数 $f(x)$ の $x=a$ における
一次のテイラー展開

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

関数 $\varphi(x - a) = \varphi(x_1 - a_1, x_2 - a_2, \dots, x_n - a_n)$ は $x_1 - a_1, x_2 - a_2, \dots, x_n - a_n$ 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)

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

関数 $f(x)$ の $x=a$ における
一次のテイラー近似

- 線形関数
- $x = a$ のとき値は $f(a)$
- 傾きは勾配ベクトル $\nabla f(a)$
- $x=a$ の近くで関数 f を近似

一次のテイラー展開

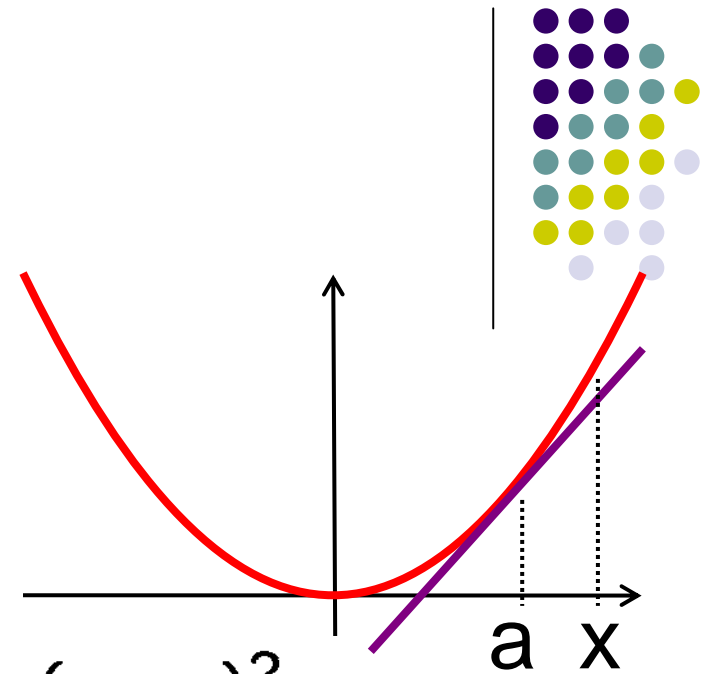
例1: $f_1(x) = x^2$ $f_1'(x) = 2x$,

f_1 の $x=a$ における一次のテイラー展開

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

ここで

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



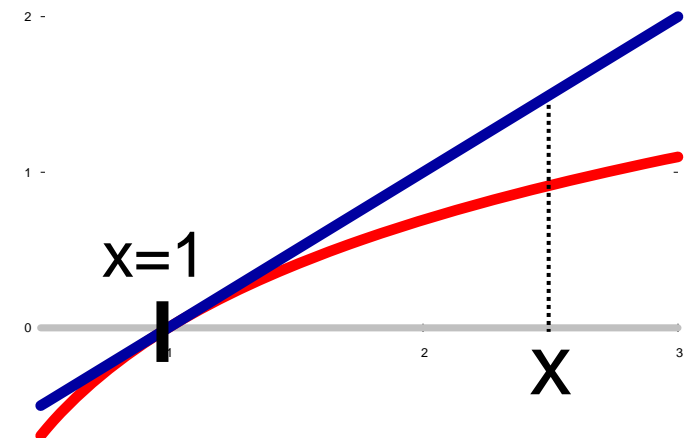
例2: $f_2(x) = \log x$ $f_2'(x) = 1/x$

f_2 の $x=1$ における一次のテイラー展開

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

ここで

$$\begin{aligned} \varphi(x - 1) \\ = -\frac{(x-1)^2}{2} + \frac{(x-1)^3}{3} - \frac{(x-1)^4}{4} + \dots \end{aligned}$$



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 $f(y - \delta \nabla f(y)) < f(y)$

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 $f(y - \delta \nabla f(y)) < f(y)$

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

$\therefore \delta$ 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 [p.89]



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 x に対し、 $\nabla f \neq 0$ ならば

十分小さい $\delta > 0$ に対して $f(x + \delta \nabla f(x)) > f(x)$

証明は省略(直前の性質と同様に証明できる)

最適性条件 [p.97]



制約なし最適化問題: 最小化 $f(x)$

最適性条件: ベクトル x が非線形計画問題の最適解であるための**必要条件**

x は**停留点** $\Leftrightarrow \nabla f(x) = 0$

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 $\Rightarrow x^*$ 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

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

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

最適性条件 [p.97]



定理(制約なし最適化問題の最適性条件):

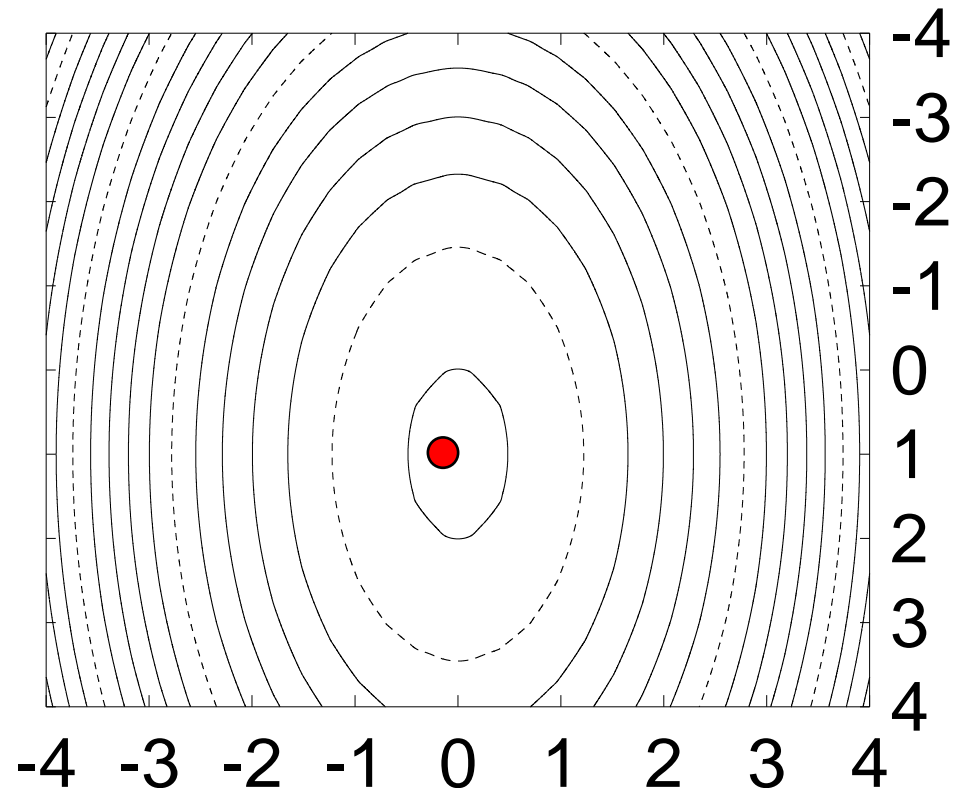
$$x^*: \text{制約なし問題の最適解} \Rightarrow \nabla f(x^*) = 0$$

例: $f(\mathbf{x}) = x_1^2 - 2x_1 + 4x_2^2$
 $= (x_1 - 1)^2 + 4x_2^2 - 1$

$(x_1, x_2) = (1, 0)$ が最適解

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

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



最適性条件 [p.97]

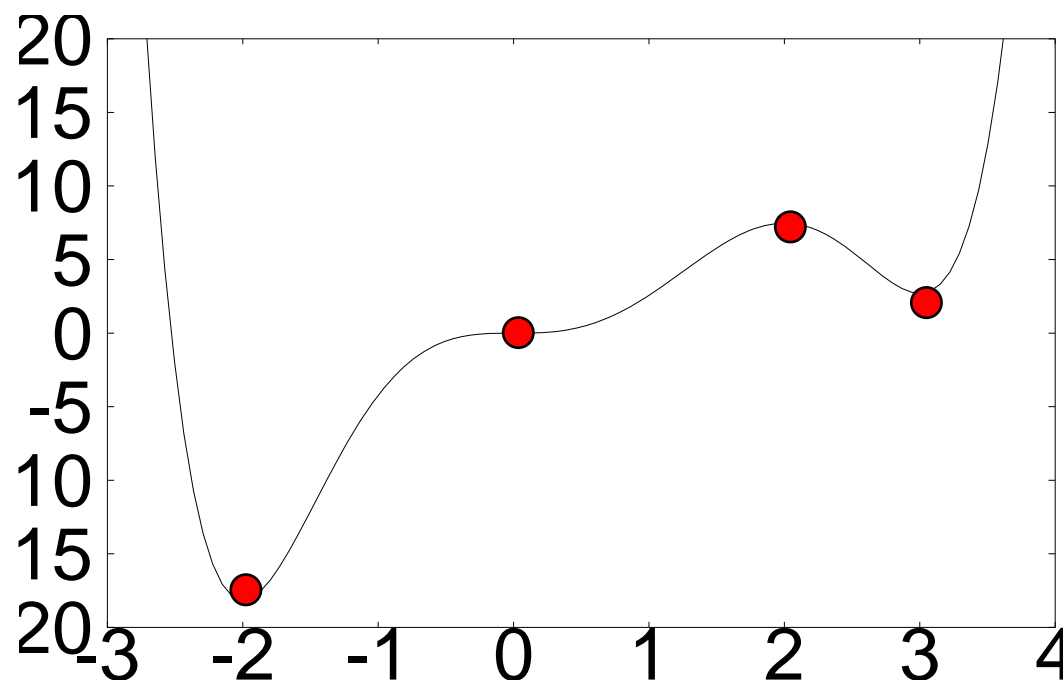


※「 x^* は停留点 $\Rightarrow x^*$ は最適解」は必ずしも
成り立たない

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 [p.99]



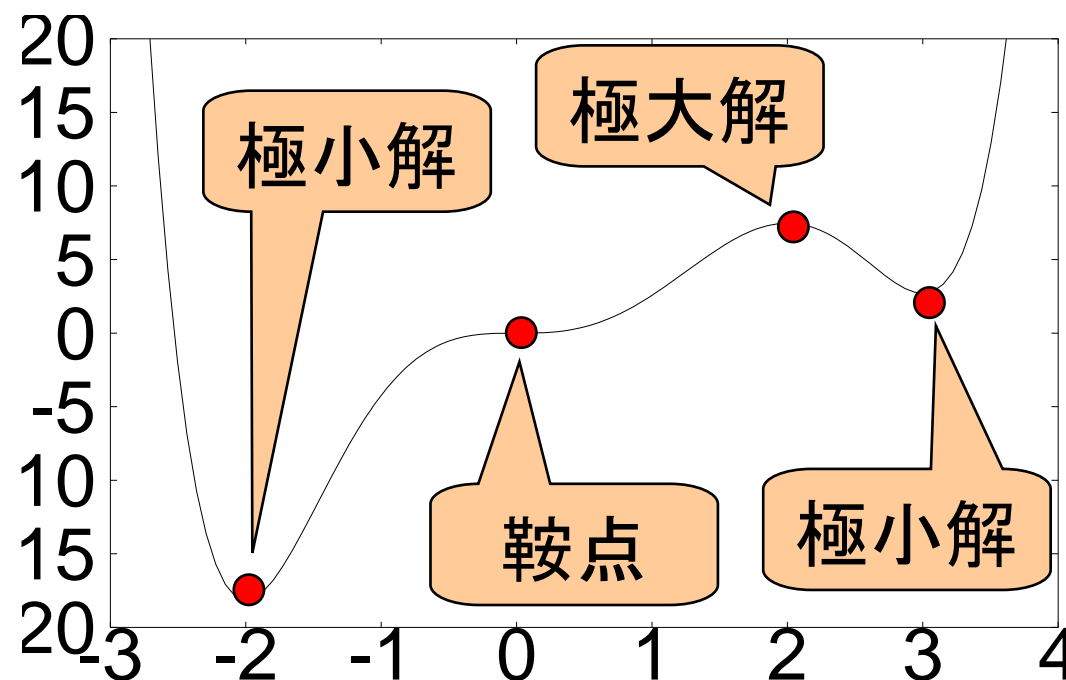
停留点 x^* の分類

極小解: x^* の付近だけに注目したとき、 x^* は最小

ある $\delta > 0$ が存在して、 $\|x - x^*\| \leq \delta$ を満たすすべての x に対して $f(x) \geq f(x^*)$

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

[p.102]



最急降下法のアイデア:

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

\Rightarrow 関数値 $f(x)$ を減らしていく

ステップサイズ

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 [p.103]



教科書の例3.2,3.8: $f(\mathbf{x}) = x_1^2 - 2x_1 + 4x_2^2$

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

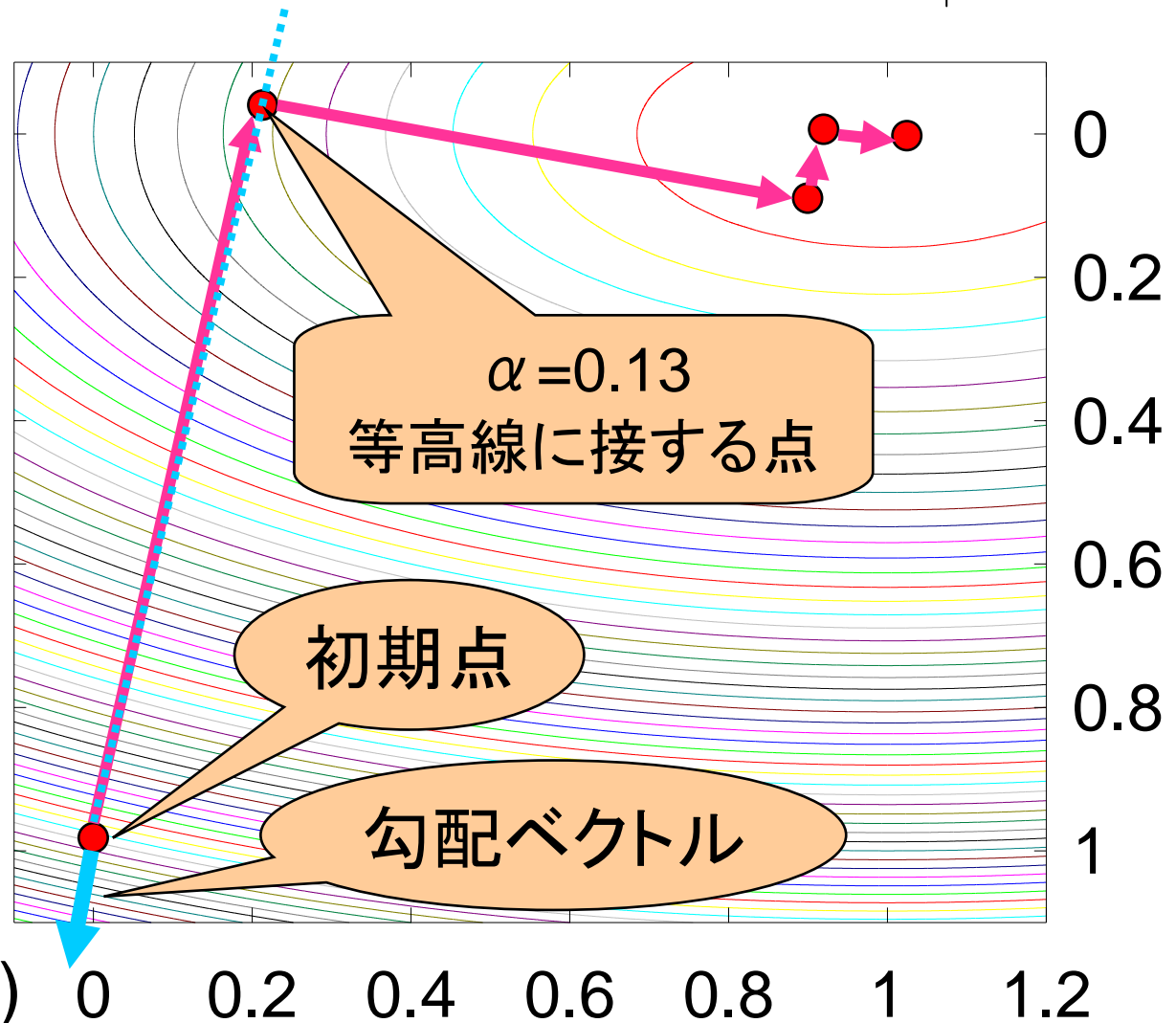
• $(x_1, x_2) = (0, 1)$ からスタート

• $\nabla f(0, 1) = (-2, 8)$

• $f(0 + 2\alpha, 1 - 8\alpha)$ 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 [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

ヘッセ行列 [p.90]



関数 f のヘッセ行列

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

一変数関数の場合は2階導関数に一致

$$H f(x) = f''(x)$$

ヘッセ行列 (続き) [p.89]



例:

$$f_1(x) = x^2 \quad \nabla f_1(x) = f_1'(x) = 2x, \quad \text{H}f_1(x) = f_1''(x) = 2$$

$$f_2(x_1, x_2) = \sin x_1 + \cos x_2 \quad \longrightarrow \quad \nabla f_2(\mathbf{x}) = \begin{pmatrix} \cos x_1 \\ -\sin x_2 \end{pmatrix}$$
$$\quad \longrightarrow \quad \text{H}f_2(\mathbf{x}) = \begin{pmatrix} -\sin x_1 & 0 \\ 0 & -\cos x_2 \end{pmatrix}$$

$$f_3(x_1, x_2) = x_1 x_2 + \log x_1$$
$$\longrightarrow \quad \nabla f_3(\mathbf{x}) = \begin{pmatrix} x_2 + 1/x_1 \\ x_1 \end{pmatrix} \longrightarrow \quad \text{H}f_3(\mathbf{x}) = \begin{pmatrix} -1/x_1^2 & 1 \\ 1 & 0 \end{pmatrix}$$

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)

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

この部分が二次関数, $f(x)$ を近似

関数 $\psi(x-a) = \varphi(x_1 - a_1, x_2 - a_2, \dots, x_n - a_n)$ は
 $x_1 - a_1, x_2 - a_2, \dots, x_n - a_n$ 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: $f_1(x) = x^2$ $\nabla f_1(x) = 2x$ $H f_1(x) = 2$

$$\begin{aligned} f(x) &= a^2 + (2a)(x-a) + \frac{1}{2} \cdot 2(x-a)^2 + \psi(x-a) \\ &= x^2 + \psi(x-a) \end{aligned}$$

$$\therefore \psi(x-a) = 0$$

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

の2次のテイラー展開において、
 $\psi(x-a) = 0$

V: $n \times n$ 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: $f(x) = (x-2)(x-1)x(x+1)(x+2) = x^5 - 5x^3 + 4x$

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

$$H f(x) = 20x^3 - 30x$$

$a = -1$ のとき

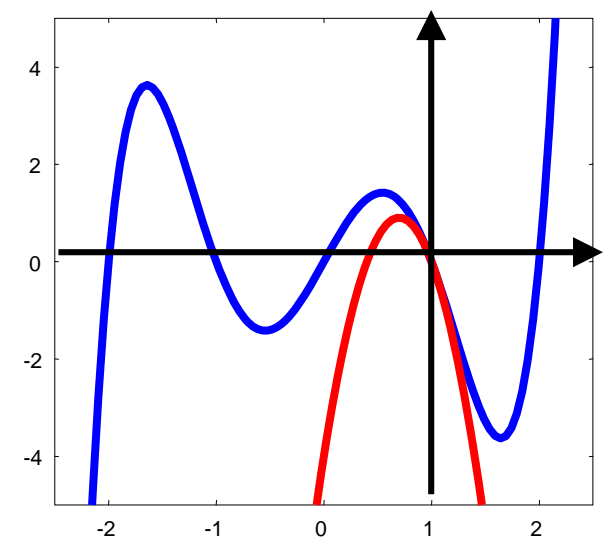
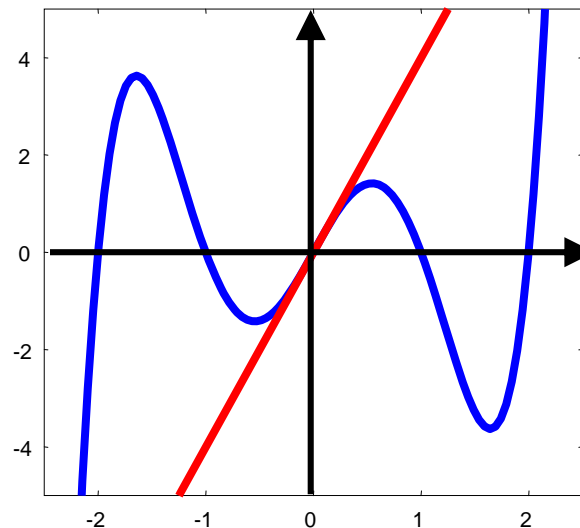
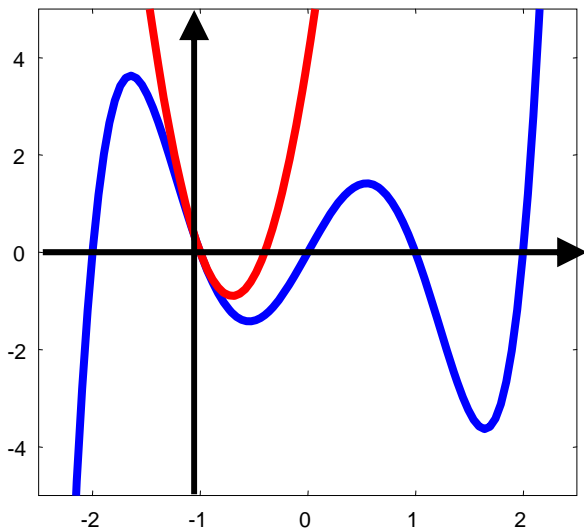
$$0 - 6(x+1) + 5(x+1)^2$$

$a = 0$ のとき

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

$a = 1$ のとき

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



演習問題



問題 1: 下記の4つの関数の勾配ベクトルとヘッセ行列を計算しなさい

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

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

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

レポート問題



問題2: 次の2つの非線形計画問題

「最大化 $f_1(x_1, x_2)$ 条件 $f_2(x_1, x_2) \leq 5$ 」

「最小化 $f_2(x_1, x_2)$ 条件 $f_1(x_1, x_2) = 3$ 」

を(手計算で)解きなさい. また, 問題および最適解を図で表しなさい.

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。(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