数理計画法　第10回

非線形最適化

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

非線形計画問題(Nonlinear Programming Problem)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　$x\,y \geqq 1$

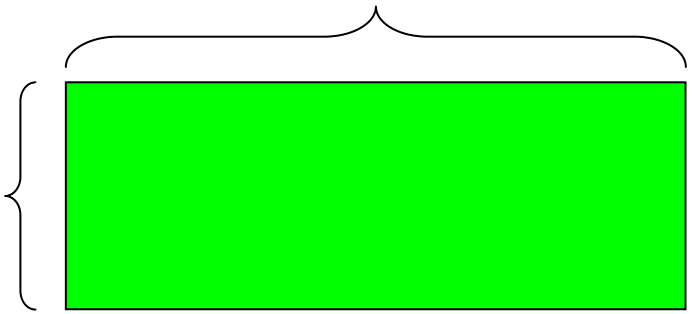
$x, y \geqq 0$

非線形の不等式

→ 非線形計画問題

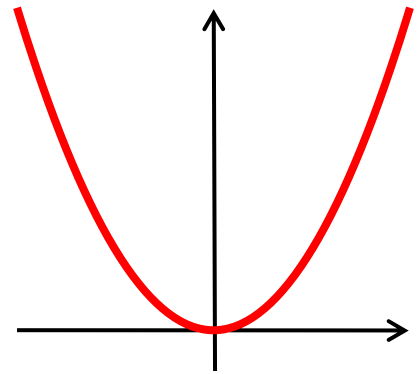
y：横の長さ

x：縦の長さ



注意:線形計画問題は非線形計画問題の特殊ケース

非線形関数の例(その1) [p.87]

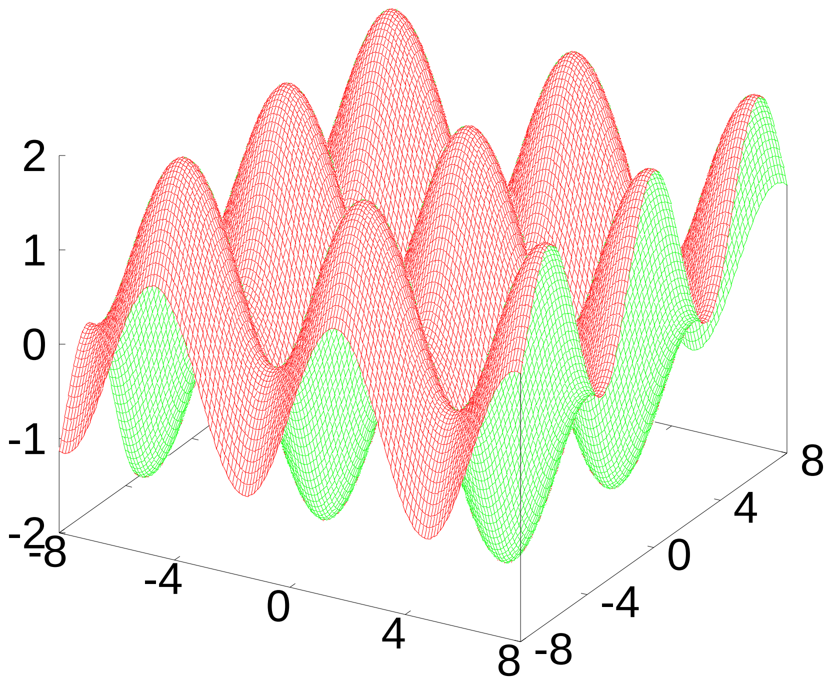


$$f_1(x) = x^2$$

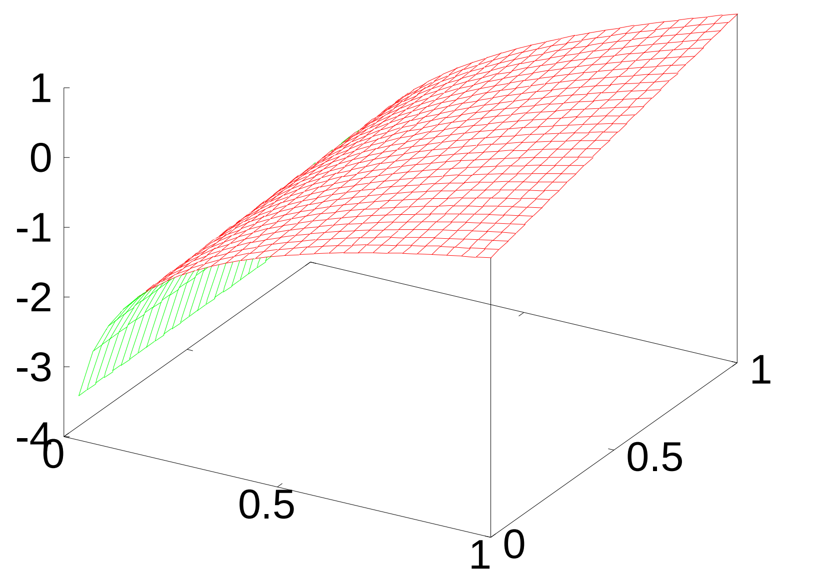
非線形関数(nonlinear function)
--- 線形でない関数

微分可能(differentiable)な非線形関数の例

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



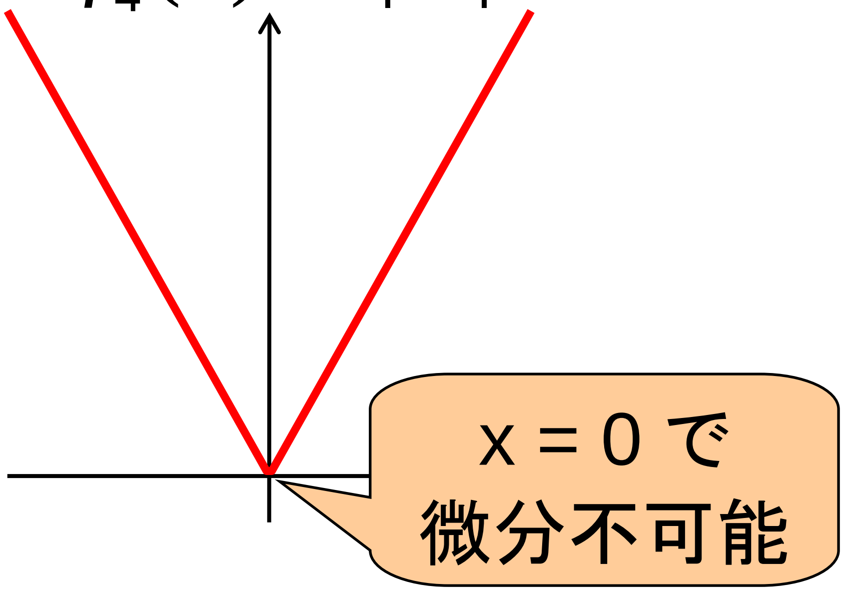
$$f_3(x_1, x_2) = x_1 x_2 + \log x_1$$



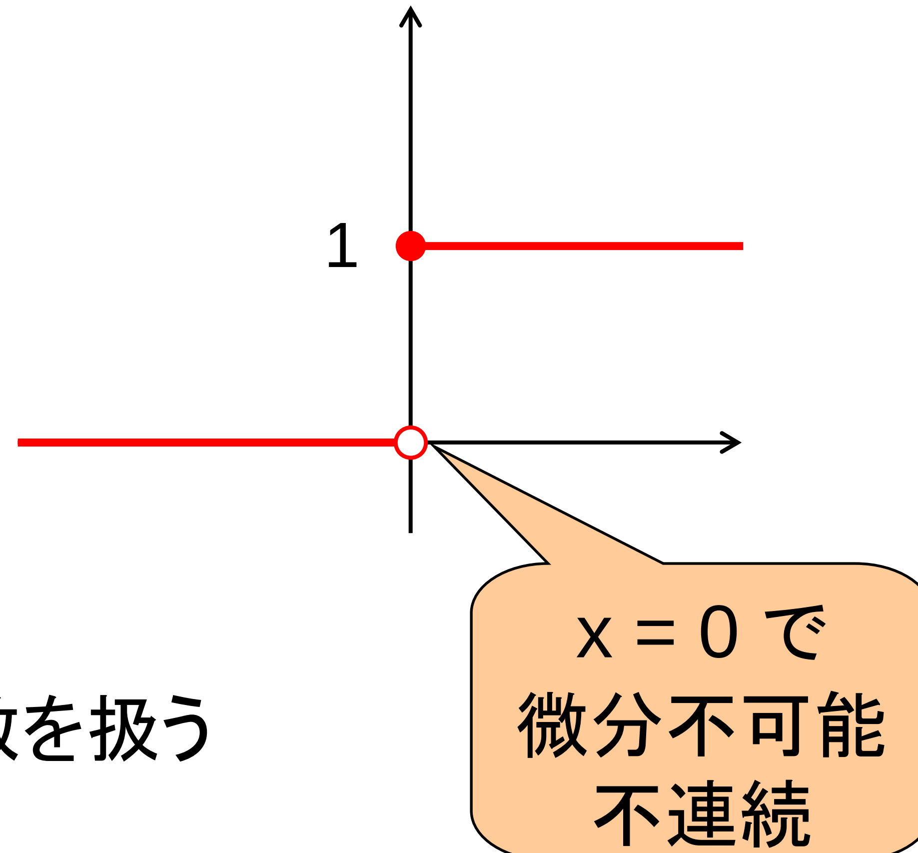
非線形関数の例（その2） [p.88]

微分不可能(nondifferentiable)
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:
主に何回でも微分可能な関数を扱う

非線形計画問題の分類 [p.6,97,131]

制約なし最適化問題
(unconstrained optimization problem)

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
(constrained optimization problem)

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

問題：　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

非線形計画問題の分類 [p.6,97,131]

制約つき問題と制約なし問題の関係

- 制約つき問題は制約なし問題に変形できる

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

最小化　f(x) + h(x)　　条件　なし

$$h(x)=\begin{cases} 0 & (g_i(\boldsymbol{x}) \le 0, i=1,2,\ldots,m) \\ M & (\text{その他}) \end{cases}$$

Mは十分に大きい正数

（注意）この制約なし問題を直接解くことは実用上難しい

- 制約なし問題を繰り返し解くことにより、制約つき問題を解くことが出来る

→p.146 ペナルティ関数法、バリア関数法

勾配ベクトル [p.89]

関数 f の勾配ベクトル(gradient vector)

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

一変数関数の場合は $\nabla f(x) = f'(x)$

例:

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

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

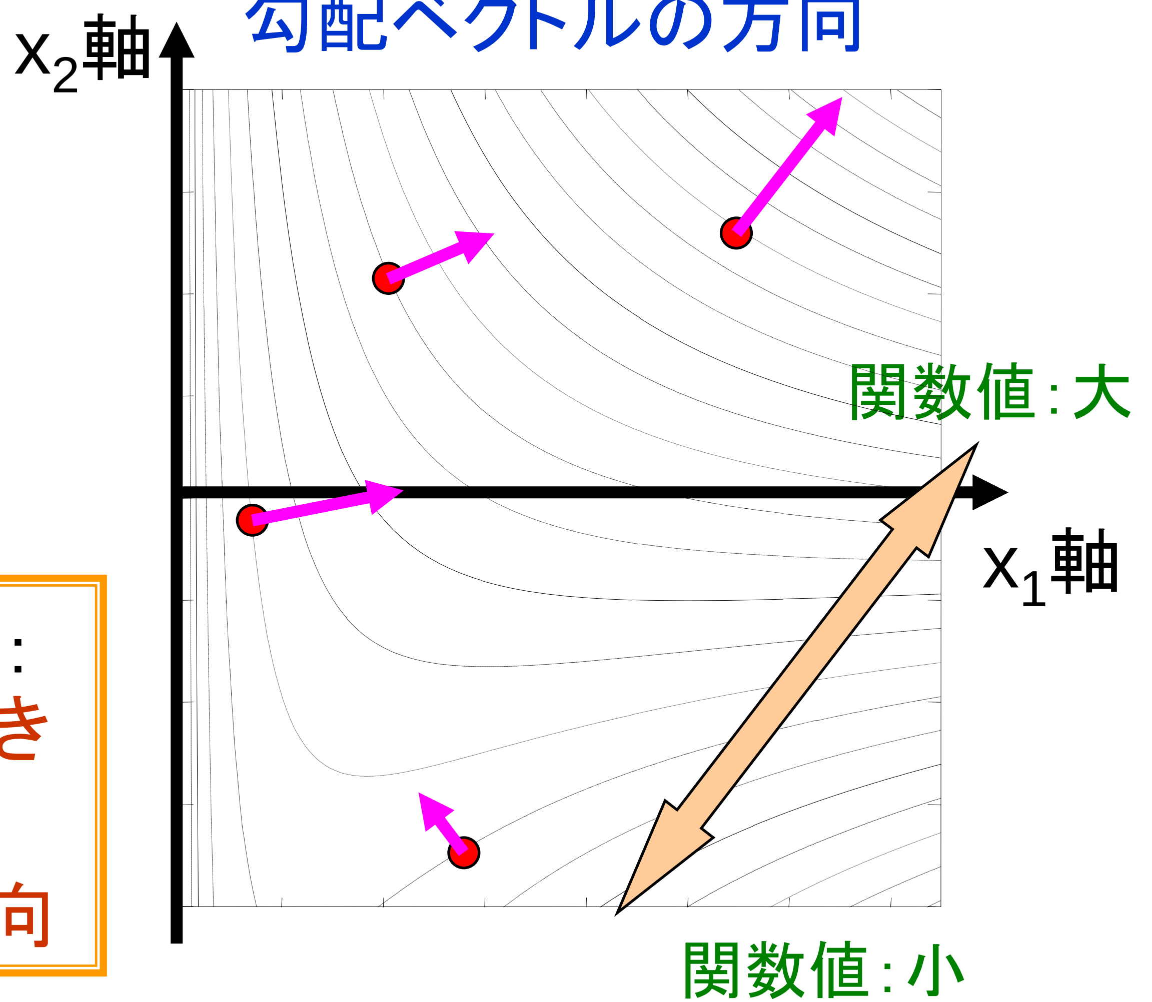
$$\Rightarrow \nabla f_2(\boldsymbol{x}) = \begin{pmatrix} \cos x_1 \\ -\sin x_2 \end{pmatrix}$$

勾配ベクトル（続き）[p.89]

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

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

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[p.89]

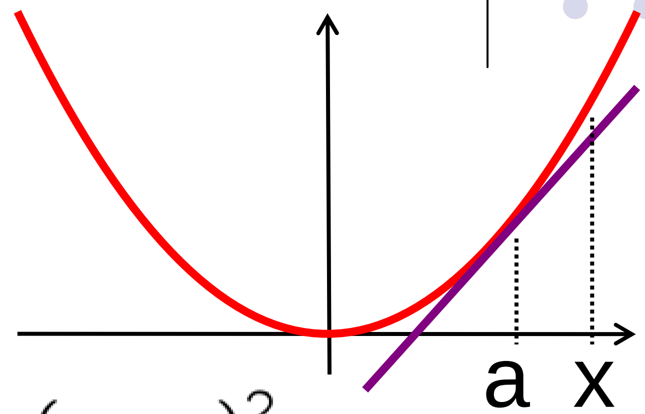
任意の関数 f は，任意のベクトル a を使って，
次の形に表現できる

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

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

関数 f(x) の x=a における
一次のテイラー展開
(Taylor expansion)

一次のテイラー展開



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

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

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

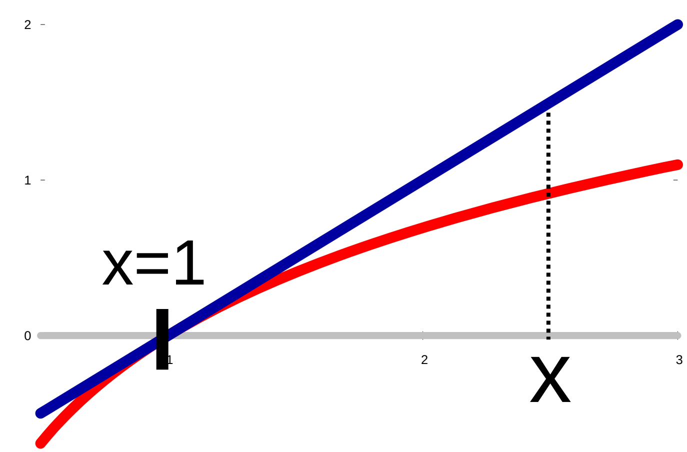
例2：$f_2(x) = \log x \qquad f_2'(x) = 1/x$

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

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



一次のテイラー近似[p.89]

関数 f(x) の x=a における
一次のテイラー展開

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

$\varphi(x-a)$ の項は他の項に比べて十分小さい値と仮定
➔無視する

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

関数 f(x) の x=a における
一次のテイラー近似
(Taylor approximation)

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

一次のテイラー近似

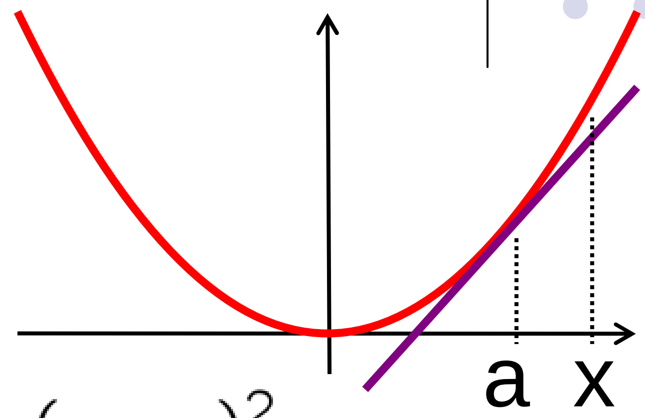
例1：$f_1(x) = x^2 \qquad f_1'(x) = 2x,$

f_1 の x=a における一次のテイラー展開

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

ここで

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



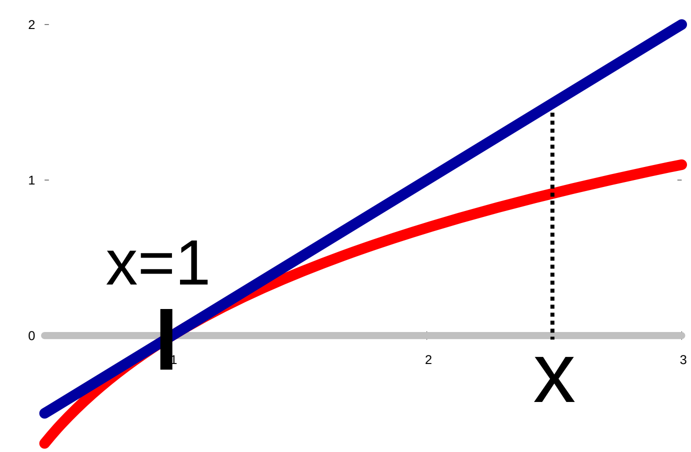
例2：$f_2(x) = \log x \qquad f_2'(x) = 1/x$

f_2 の x=1 における一次のテイラー展開

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

ここで

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



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

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

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

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

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　$f(\boldsymbol{x} + \delta \nabla f(\boldsymbol{x})) > f(\boldsymbol{x})$

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(optimality condition):

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(stationary point)　⇔　$\nabla f(x) = 0$

> 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　⇒　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):

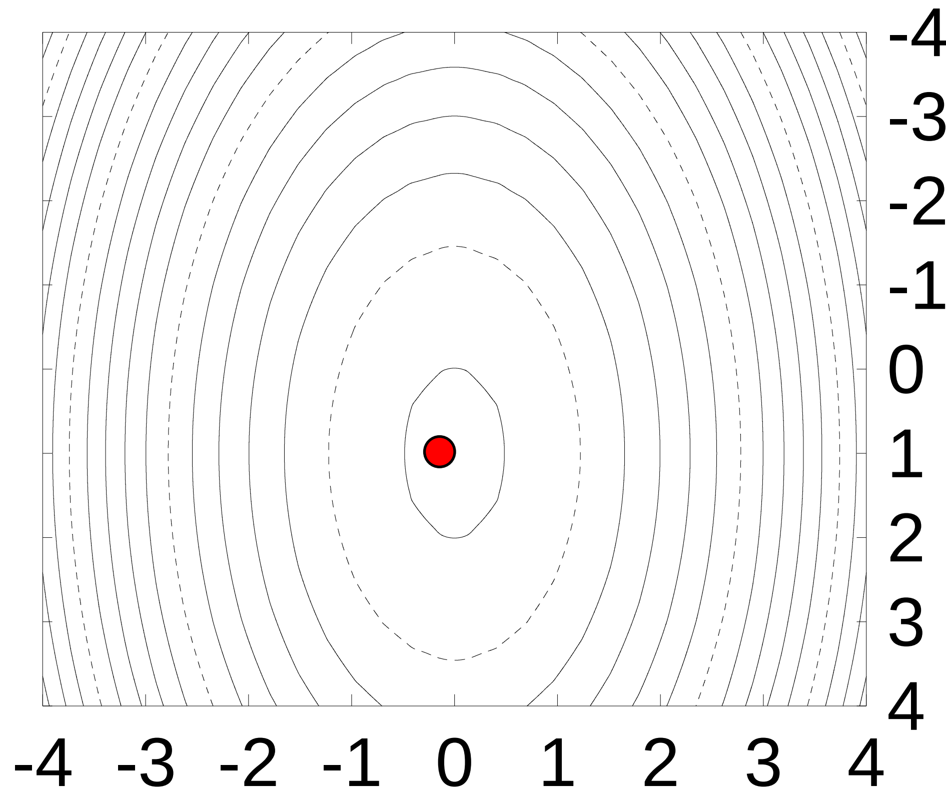
x*: 制約なし問題の最適解 ⇒ ∇f(x*) = 0

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 [p.97]

※「x*は停留点 ⇒ x*は最適解」は必ずしも成り立たない

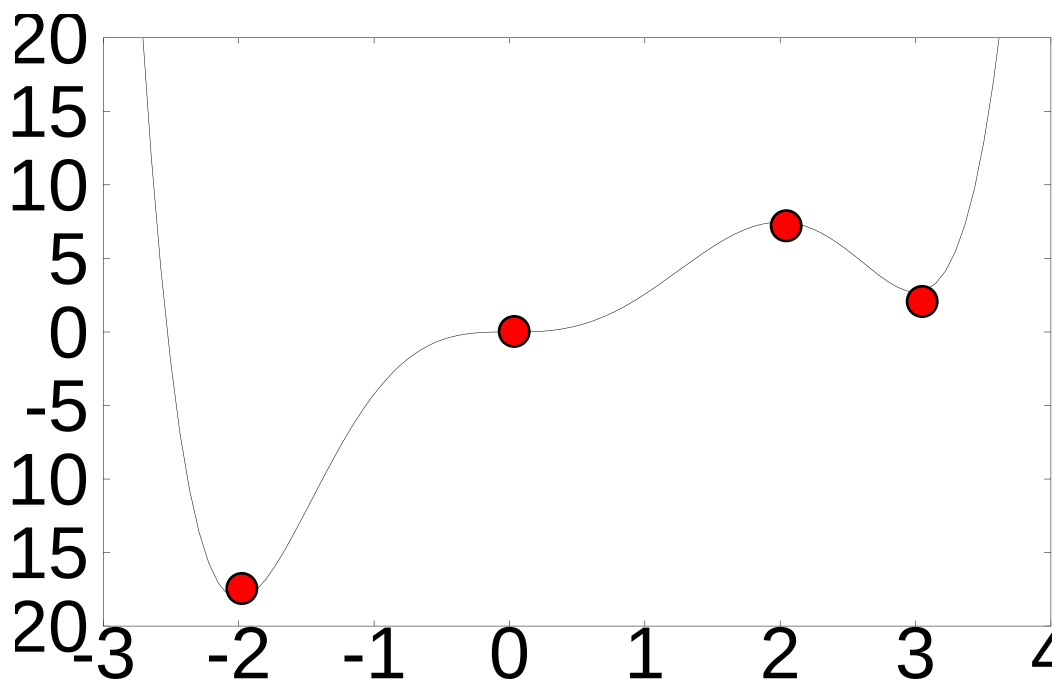
例：

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

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

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 [p.99]

停留点 x* の分類

極小解(local minimum solution):

x* の付近だけに注目したとき、x* は最小

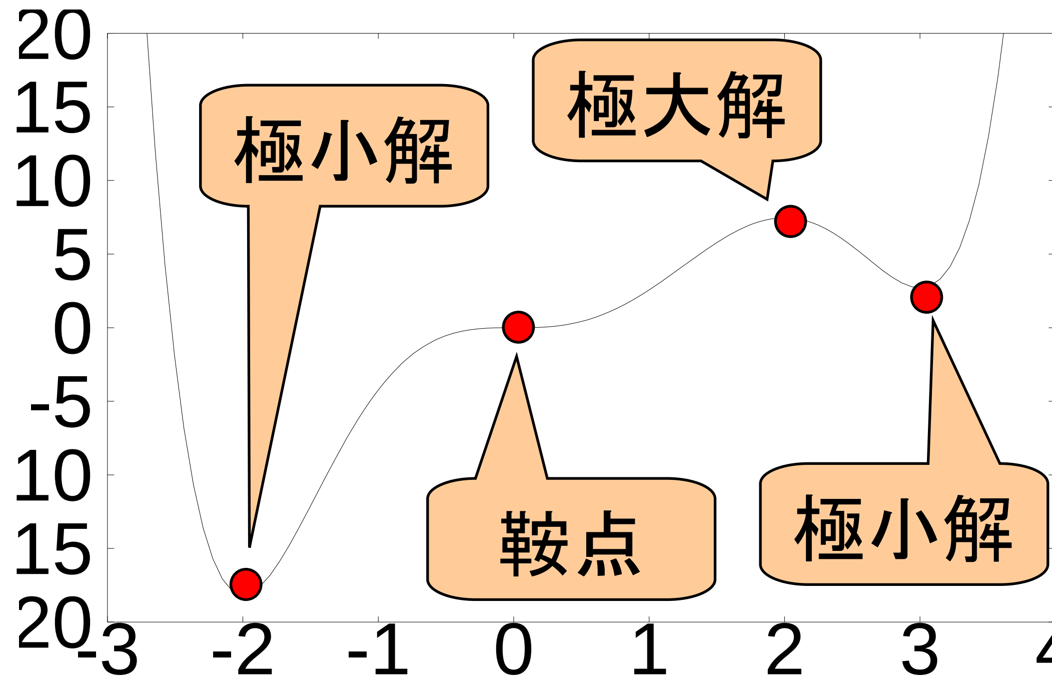
> ある δ >0 が存在して、||x − x*|| ≦ δ を満たす
> すべての x に対して f(x) ≧ f(x*)

極大解(local maximum solution):

x* の付近だけに注目したとき、x* は最大

鞍点(saddle point):

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 (Steepest Descent Method)[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 $f(x-\alpha\nabla f(x))$ 条件 $\alpha>0$

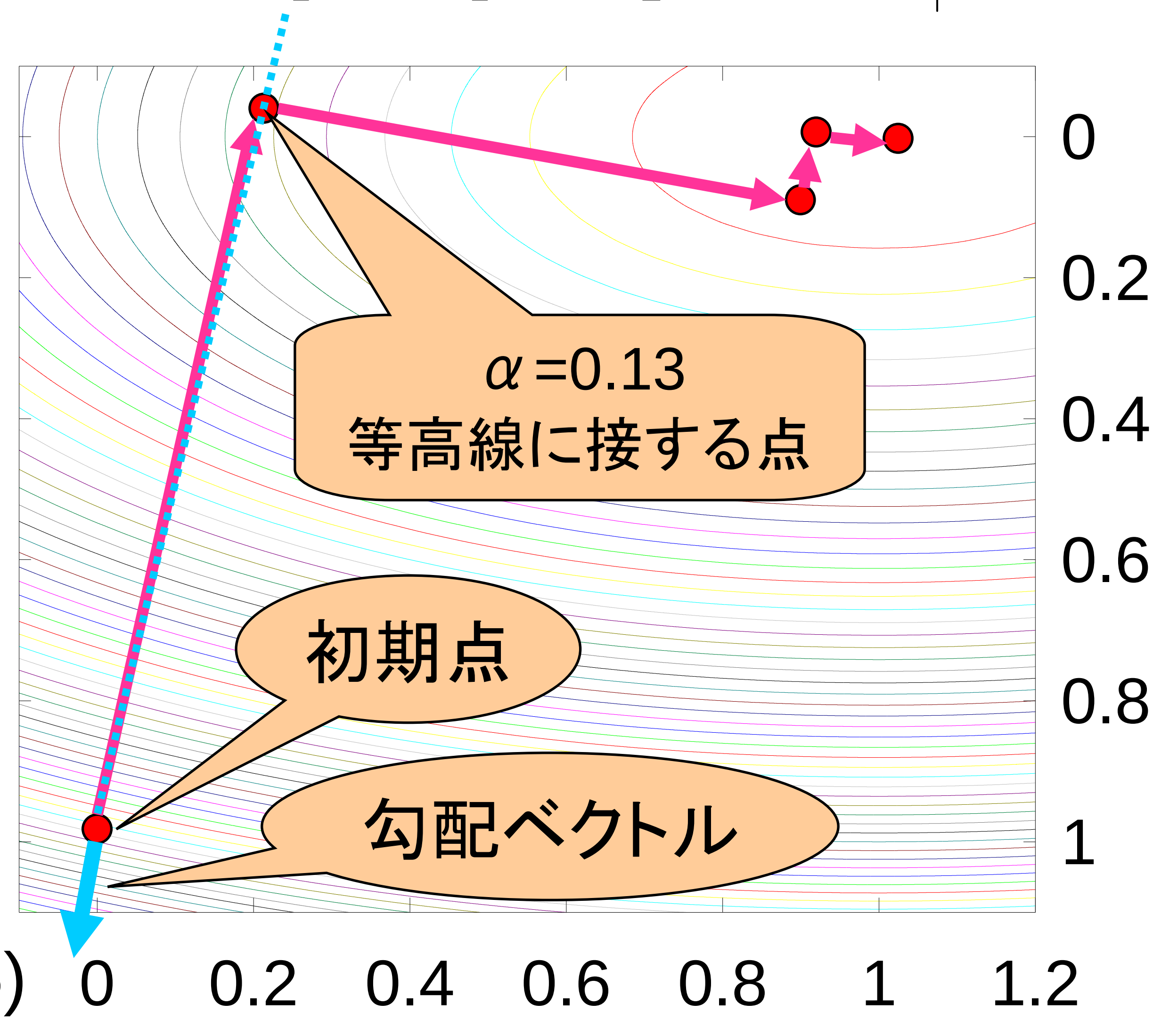
直線探索(line search)と呼ばれる

最急降下法の実行例 [p.103]

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

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

- $(x_1,x_2) = (0,1)$ 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 [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

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: 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展開を求めなさい．

問題3:次の2つの非線形計画問題

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

「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．