

数理手法

(数理最適化) 第11回

非線形計画その1

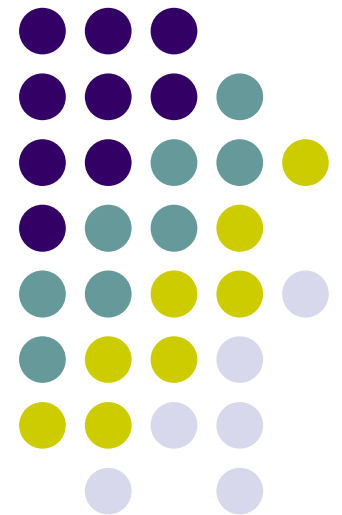
一次の最適性条件と最急降下法

塩浦昭義

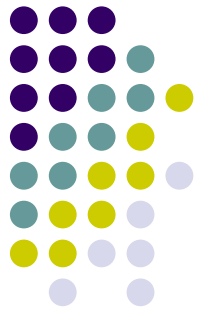
東京工業大学 経営工学系 准教授

shioura.a.aa@m.titech.ac.jp

<http://www.me.titech.ac.jp/~shioura/shioura/teaching/TUmp17/index.html>



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

$$\begin{array}{l} \text{最小化} \quad 2x + 2y \\ \text{条件} \quad xy \geq 1 \\ \quad \quad x, y \geq 0 \end{array}$$

非線形の
制約条件

例2: 線形制約つき関数最大化問題

$$\begin{array}{l} \text{最大化} \quad -3x^2 + 5y^3 + 2xy^2 - 4y \\ \text{条件} \quad 3x + 5y \leq 12 \\ \quad \quad 7x + 2y \leq 8 \\ \quad \quad x, y \geq 0 \end{array}$$

非線形の
目的関数

制約なし問題 (unconstrained problem)

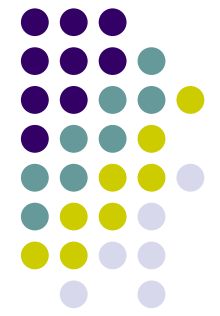
$$\text{最小化 } f(x) \quad \text{条件} \quad \text{なし}$$

制約つき問題 (constrained problem)

$$\text{最小化 } f(x) \quad \text{条件} \quad g_i(x) \leq 0 \quad (i = 1, 2, \dots, m)$$

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)

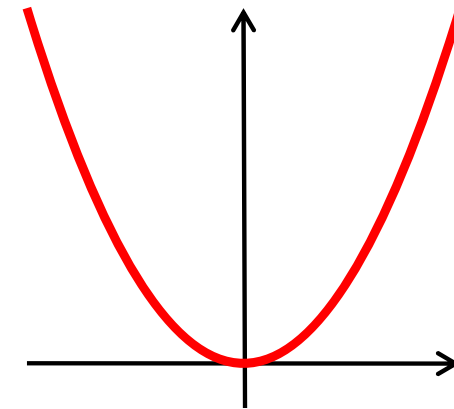
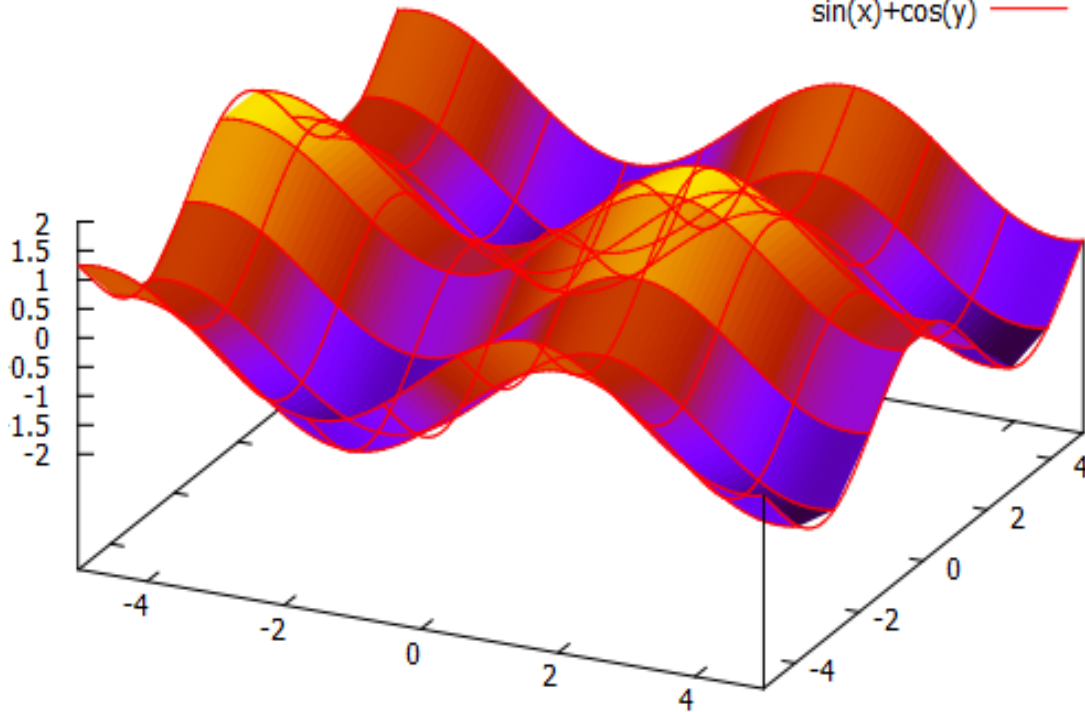


非線形関数 --- 線形でない関数

微分可能な非線形関数の例

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

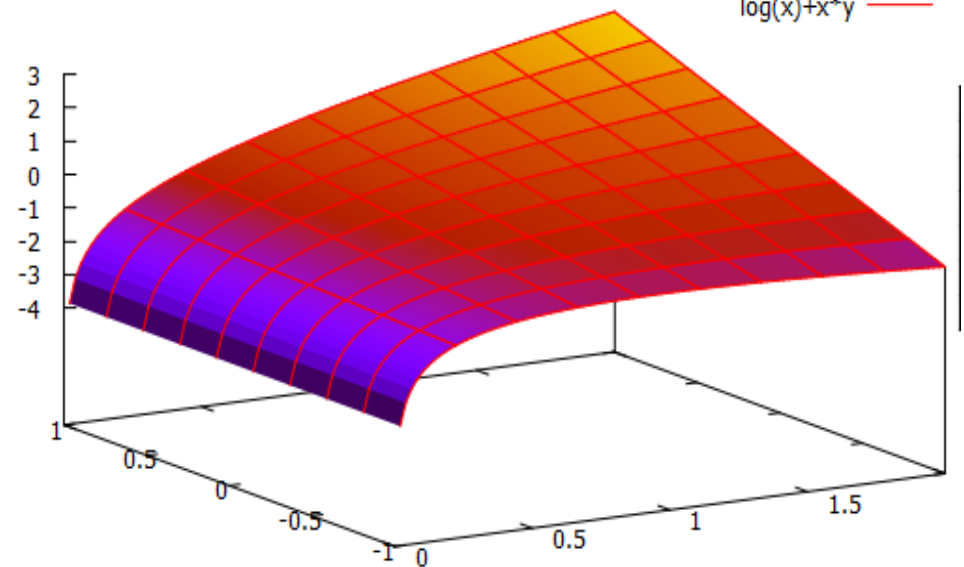
sin(x)+cos(y) —



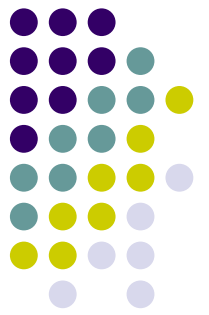
$$f_1(x) = x^2$$

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

log(x)+x*y —

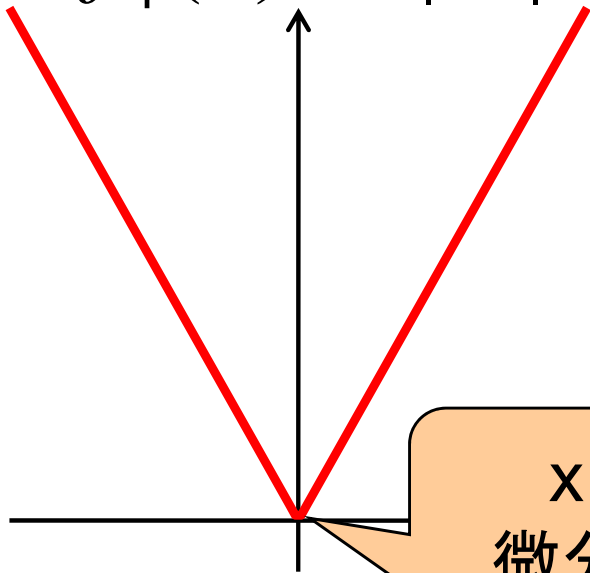


非線形関数の例(その2)



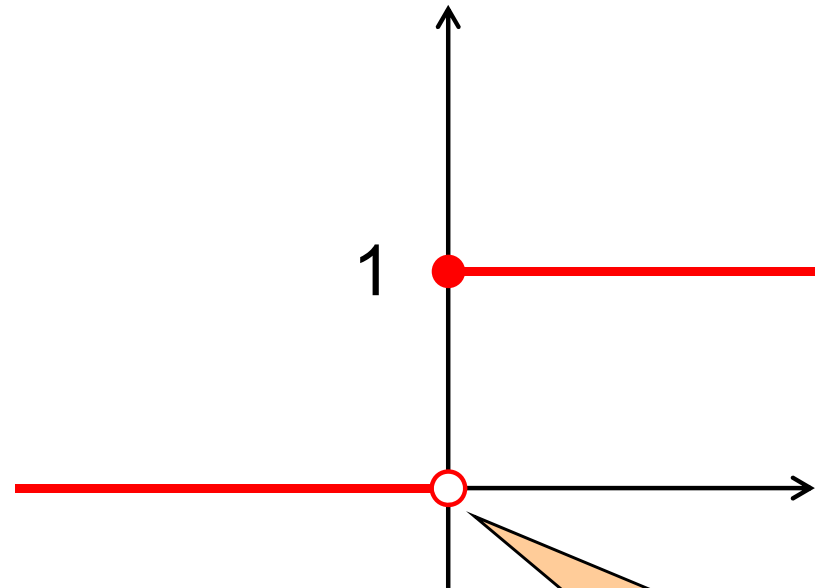
微分不可能な非線形関数の例

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



x = 0 で
微分不可能

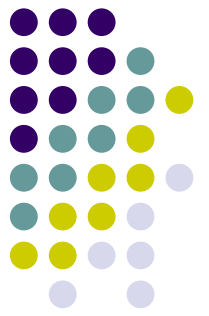
$$f_5(x) = \begin{cases} 1 & (x \geq 0 \text{ のとき}) \\ 0 & (x < 0 \text{ のとき}) \end{cases}$$



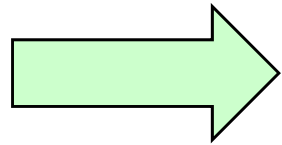
x = 0 で
微分不可能
不連続

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

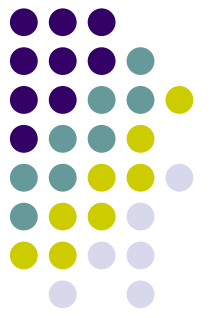


制約なし問題 最小化 $f(x) + h(x)$ 条件 なし

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

- 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)

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

一変数関数の場合 $\nabla f(x) = f'(x)$

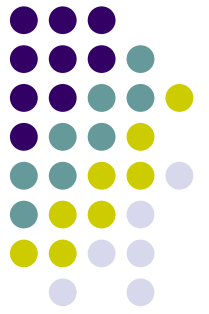
例:

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

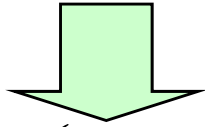
$$f_2(x_1, x_2) = \sin x_1 + \cos x_2$$

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

勾配ベクトル(続き)

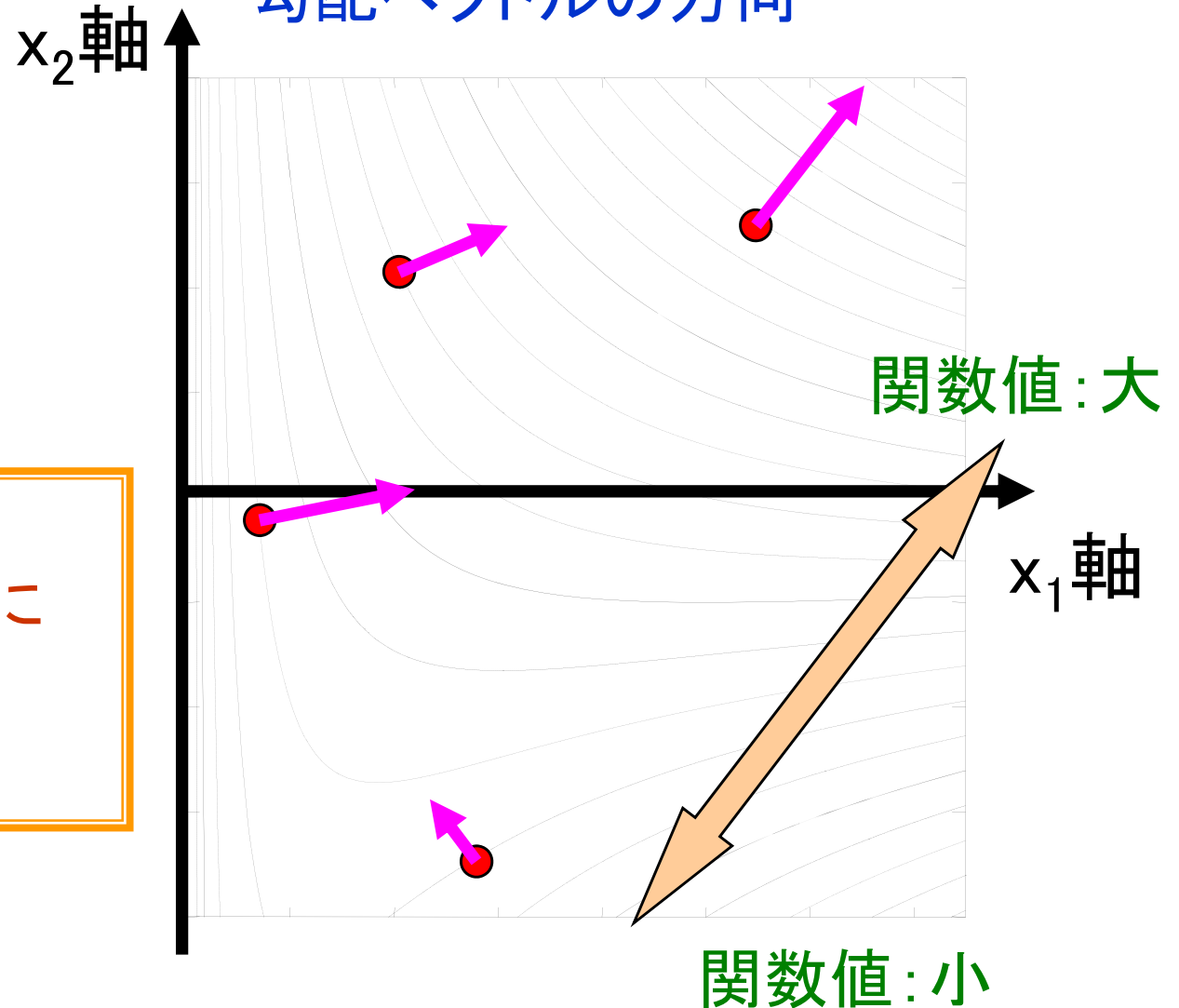


$$f_3(x_1, x_2) = x_1 x_2 + \log x_1$$



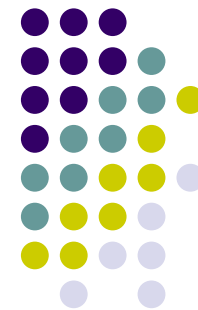
$$\nabla f_3(\mathbf{x}) = \begin{pmatrix} x_2 + 1/x_1 \\ x_1 \end{pmatrix}$$

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

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

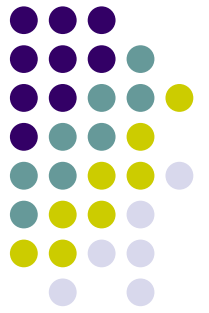
$$f(x) = f(a) + \nabla f(a)^T (x - a) + \varphi(x - a)$$

関数 $f(x)$ の $x = a$ における
一次のテイラー展開

$\varphi(d)$ は $d = (d_1, d_2, \dots, d_n)$ に関する
2次以上の項からなる n 変数多項式
(定数項, 1次の項は存在しない)

→ 任意のベクトル d に対して

$$\lim_{\varepsilon \rightarrow 0} \frac{\varphi(\varepsilon d)}{\varepsilon} = 0$$



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

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

$x \simeq a$ 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) \rightarrow 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: $f_1(x) = x^2$ $f'_1(x) = 2x$,

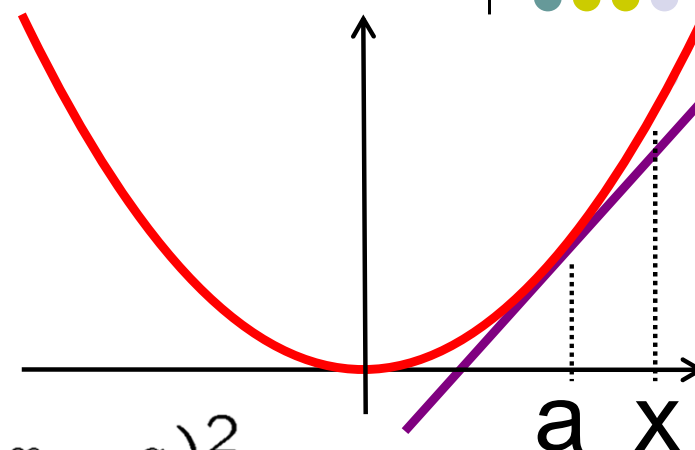
x=aでのテイラー展開

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

ここで

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

テイラー近似 $\tilde{f}_1(x) = a^2 + 2a(x - a)$



例2: $f_2(x) = \log x$ $f'_2(x) = 1/x$

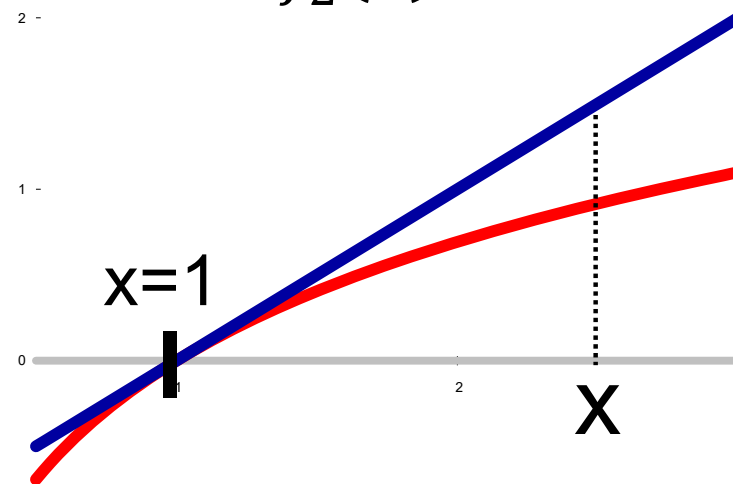
x=1でのテイラー展開

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

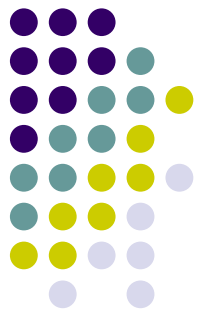
ここで

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

テイラー近似 $\tilde{f}_2(x) = x - 1$



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

命題: 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, $f(y - \delta \nabla f(y)) < f(y)$ 成立

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

一次のテイラー展開において $x = y + d$, $a = y$ とおくと,

$$\begin{aligned} f(y + d) &= f(y) + \nabla f(y)^T d + \varphi(d) \\ &= f(y) - \delta \|\nabla f(y)\|^2 + \varphi(-\delta \nabla f(y)) \\ &= f(y) - \delta \left(\|\nabla f(y)\|^2 - \frac{\varphi(-\delta \nabla f(y))}{\delta} \right) \quad (1) \end{aligned}$$

任意のベクトル d に対して $\lim_{\varepsilon \rightarrow 0} \frac{\varphi(\varepsilon d)}{\varepsilon} = 0$ なので,

$$\text{十分小さい } \delta > 0 \text{ に対し, } \frac{\varphi(-\delta \nabla f(y))}{\delta} < \|\nabla f(y)\|^2 \quad (2)$$

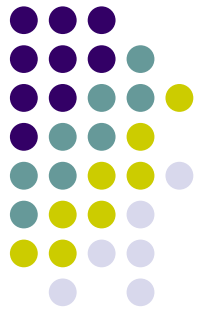
式 (1), (2) より $f(y - \delta \nabla f(y)) = f(y + d) < f(y)$

勾配ベクトルの性質

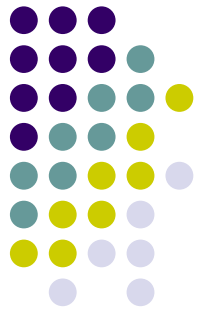
勾配ベクトルの方向に進むと関数値が増える

命題: 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, $f(y + \delta \nabla f(y)) > f(y)$ 成立

証明は省略(直前の命題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 $\Leftrightarrow \nabla f(x) = 0$

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 $\rightarrow x^*$ は停留点

証明: $\nabla f(x^*) \neq 0$ と仮定

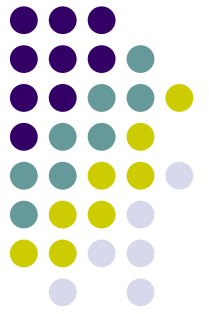
勾配ベクトルの性質より、十分小さい $\delta > 0$ に対して

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

x^* が最適解であることに矛盾

$$\therefore \nabla f(x^*) = 0$$

最適性条件



定理(制約なし問題の最適性条件):

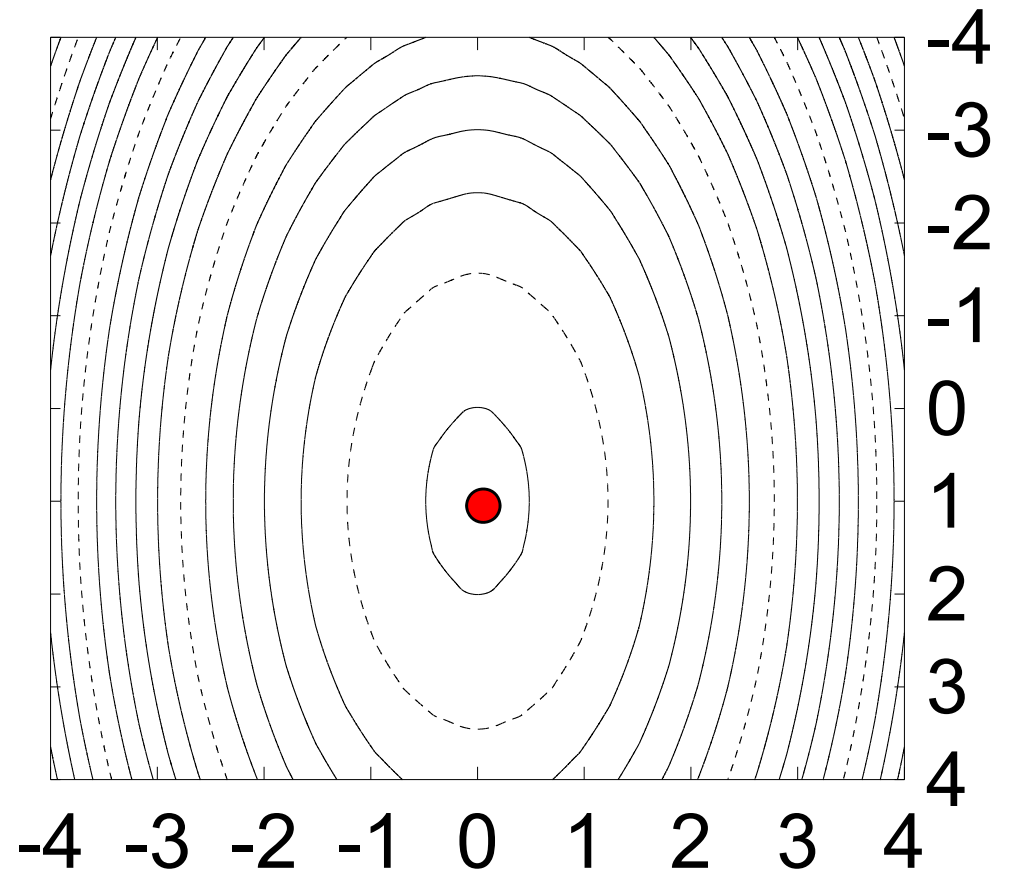
$$x^*: \text{制約なし問題の最適解} \Rightarrow \nabla f(x^*) = 0$$

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

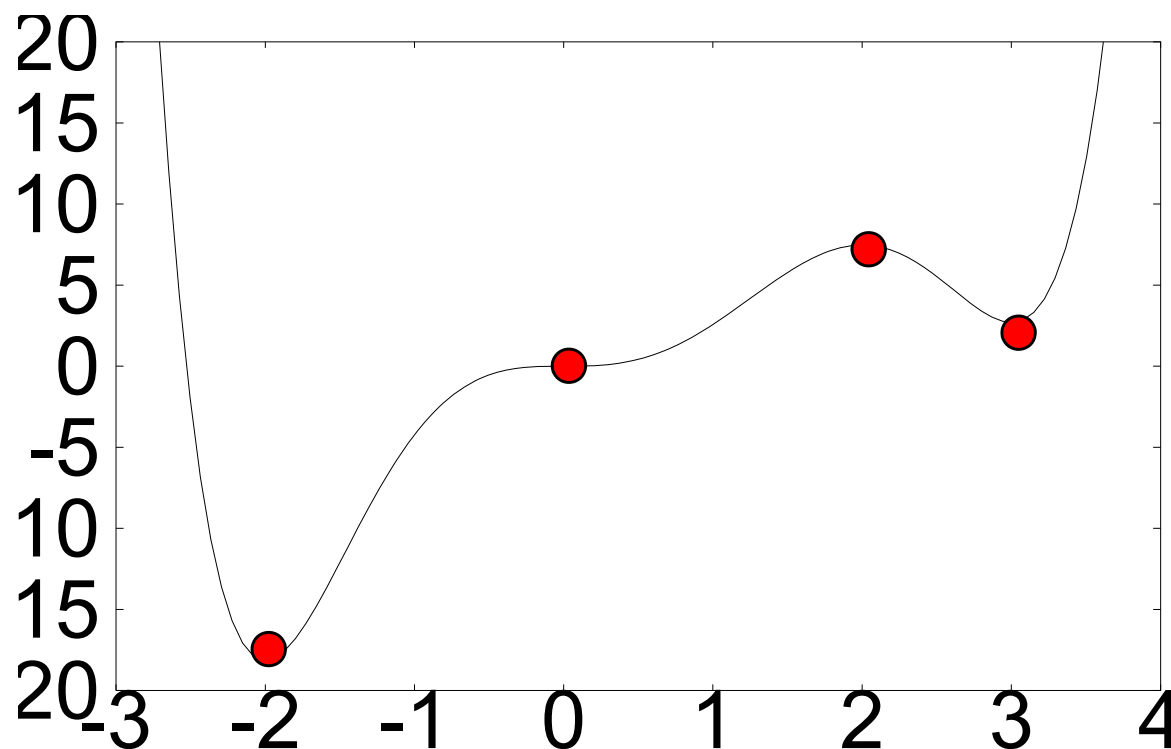


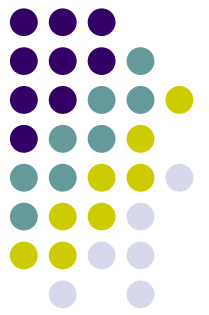
※「 x^* は停留点 $\Rightarrow x^*$ は最適解」は必ずしも成り立たない

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

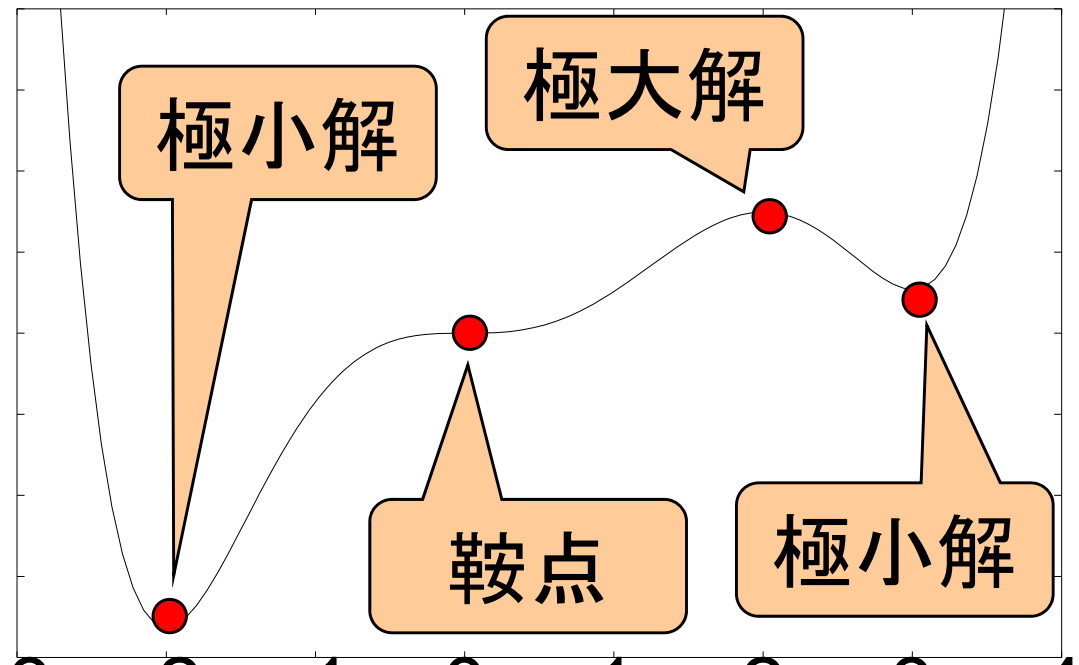
停留点 x^* の分類

極小解: x^* の付近だけに注目したとき、 x^* は最小

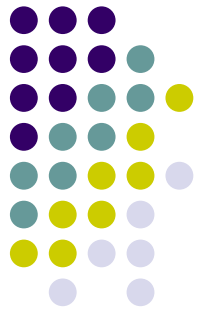
ある $\delta > 0$ が存在して、 $\|x - x^*\| \leq \delta$ を満たすすべての x に対して $f(x) \geq f(x^*)$

極大解: x^* の付近だけに注目したとき、 x^* は最大

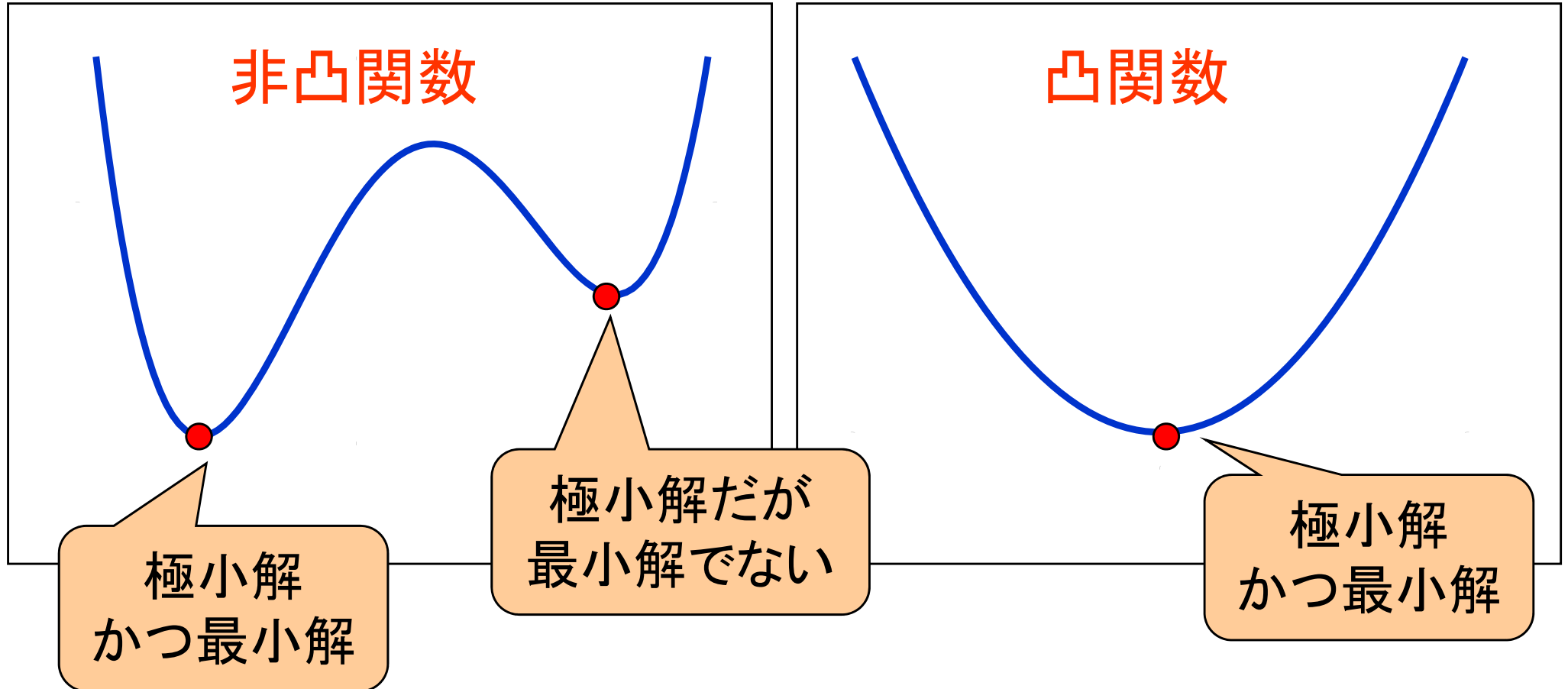
鞍点: 極小点でも極大点でもない停留点



凸関数



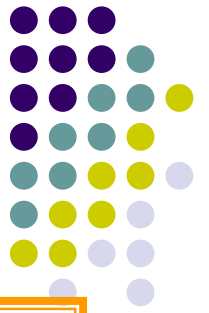
最小化しやすい関数の形は？



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

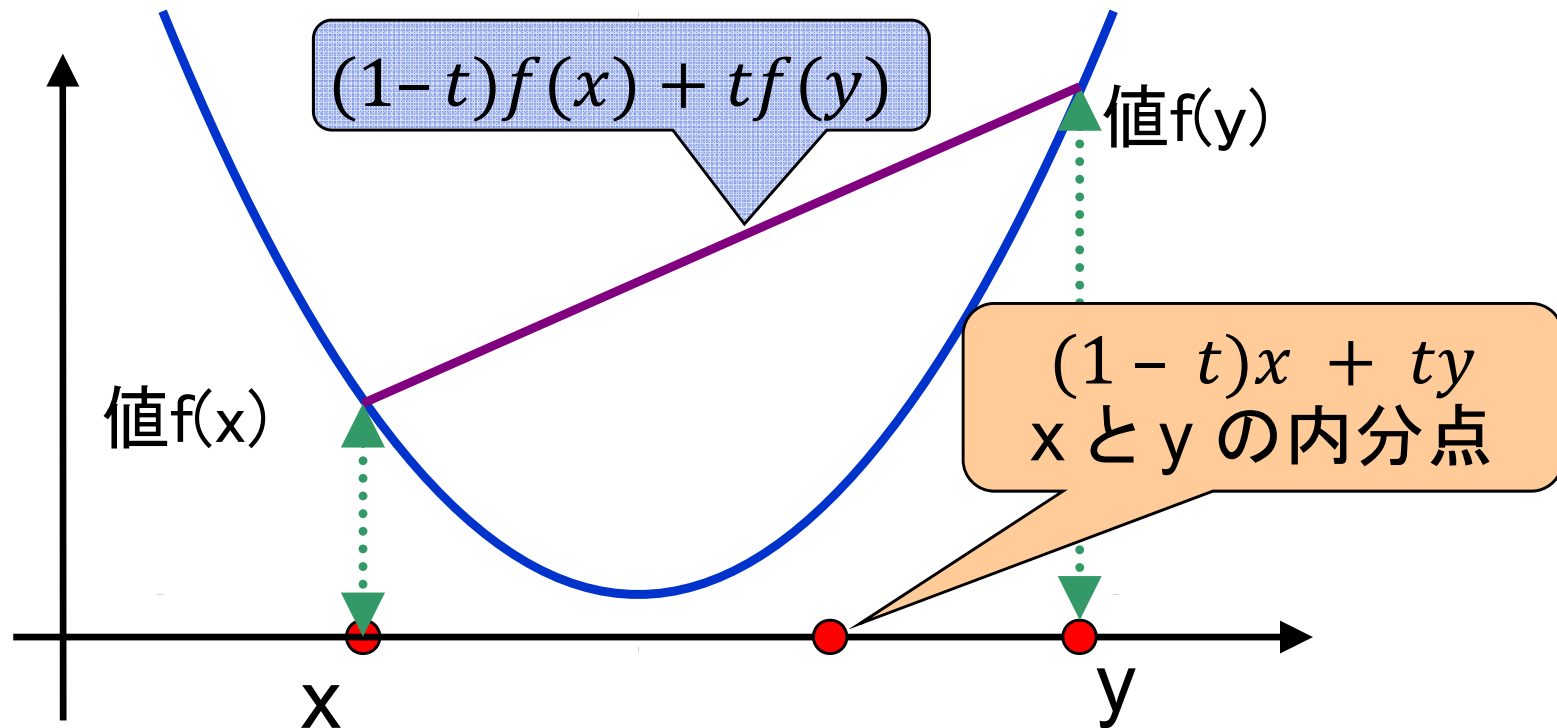
凸関数の定義

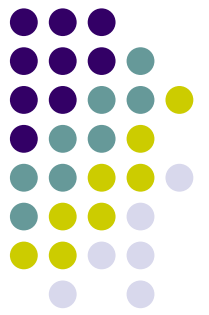


定義: 関数 f は**凸関数**

⇔ 任意の異なるベクトル x, y および任意の $0 < t < 1$ に対し

$$(1-t)f(x) + tf(y) \geq f((1-t)x + ty)$$

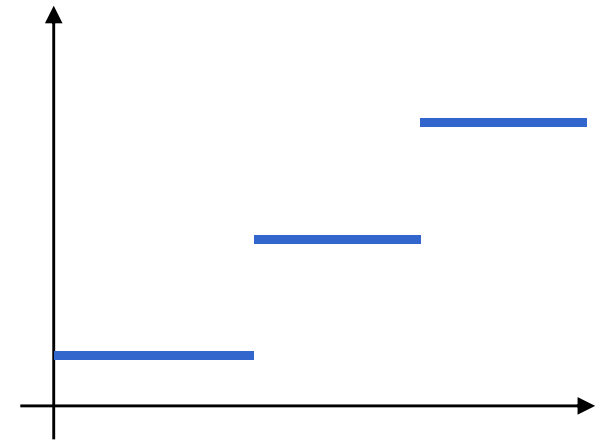
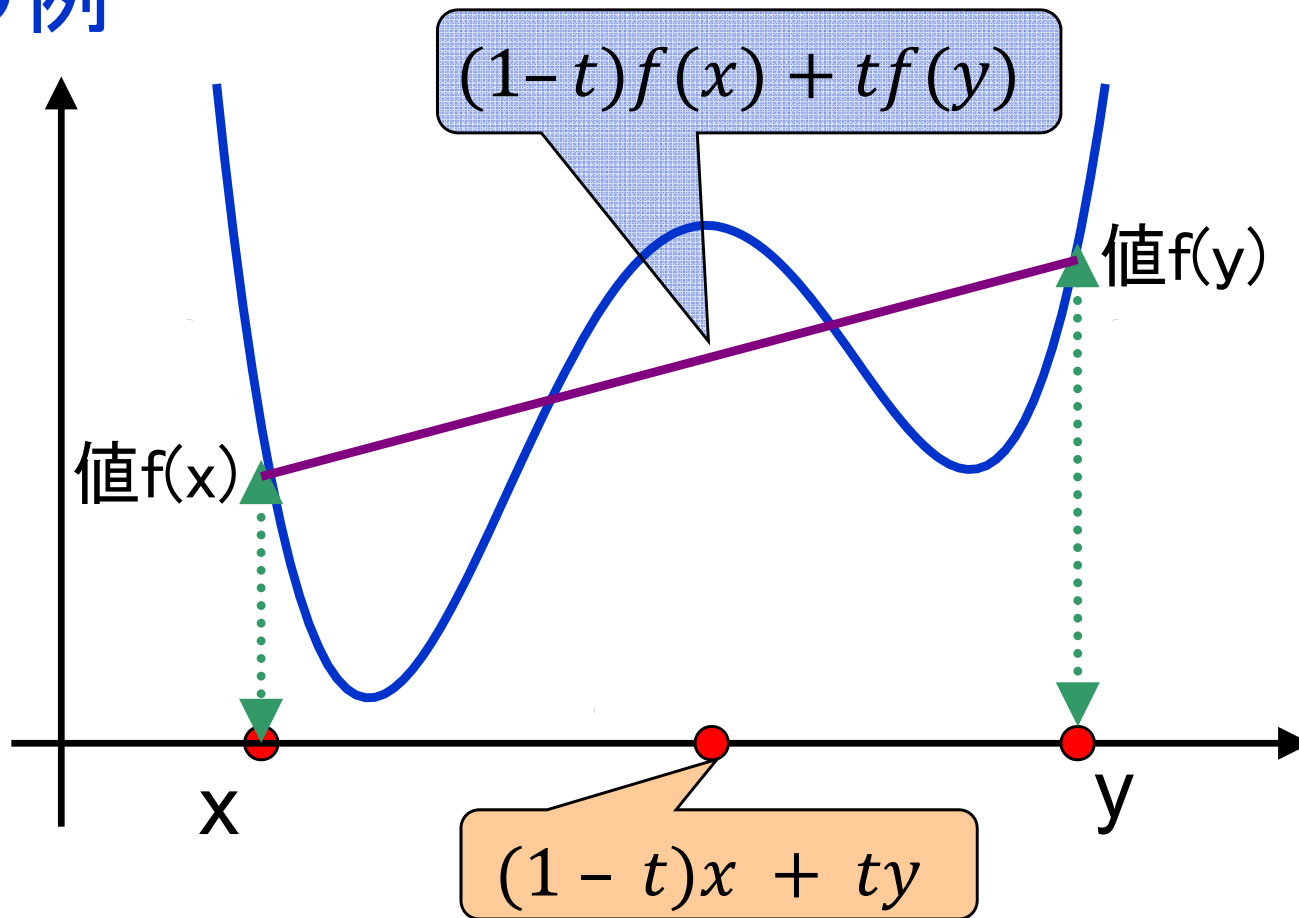




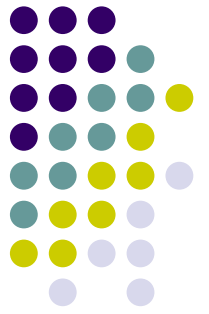
凸関数の定義(続き)

$$\text{凸関数} \Leftrightarrow (1-t)f(x) + tf(y) \geq f((1-t)x + ty)$$

非凸関数の例



2次の凸関数



$$\text{凸関数} \Leftrightarrow (1-t)f(x) + tf(y) \geq f((1-t)x + ty)$$

- 1変数の2次関数 $f(x) = ax^2 + bx + c$ ($a > 0$) は凸関数

(証明) 任意の異なる x, y と $0 < t < 1$ に対して、

$$(1-t)f(x) + tf(y) - f((1-t)x + ty)$$

$$= (1-t)ax^2 + t ay^2 - a[(1-t)x + ty]^2$$

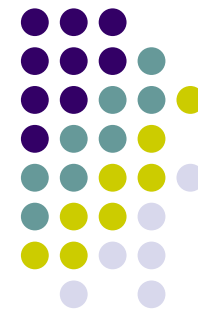
$$= (1-t)ax^2 + t ay^2 - a(1-t)^2 x^2 - a t^2 y^2 - 2a(1-t)t x y$$

$$= (t-t^2)ax^2 + (t-t^2)ay^2 - 2a(t-t^2)xy$$

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

$$> 0 \quad (0 < t < 1, x \neq y \text{より})$$

2次の凸関数(続き)



$$\text{凸関数} \Leftrightarrow (1-t)f(x) + tf(y) \geq f((1-t)x + ty)$$

より一般に,

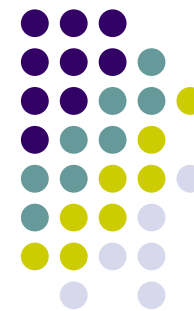
$$\text{2次関数 } f(\mathbf{x}) = \frac{1}{2} \mathbf{x}^T V \mathbf{x} + \mathbf{c}^T \mathbf{x} + c_0$$

(V : $n \times n$ 行列, \mathbf{c} : n 次元ベクトル, c_0 : 定数)

は V が半正定値行列 \rightarrow 凸関数

例:

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



1次関数は凸関数

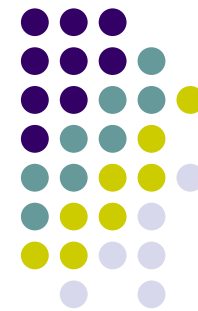
$$\text{凸関数} \Leftrightarrow (1-t)f(x) + tf(y) \geq f((1-t)x + ty)$$

- (n変数の)1次関数 $f(x) = \sum_{i=1}^n a_i x_i + c$ は凸関数
とくに $(1-t)f(x) + tf(y) = f((1-t)x + ty)$ をみたら

注意:

応用では, 非自明な凸関数がしばしば現れる

凸関数の特徴付け

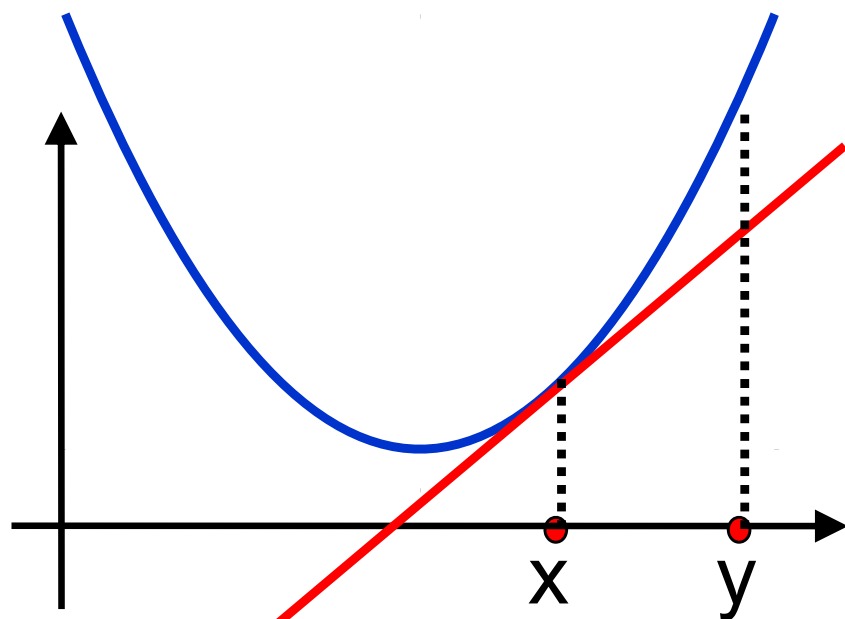


定理: f : 凸関数, 微分可能 (勾配ベクトルが定義可能)

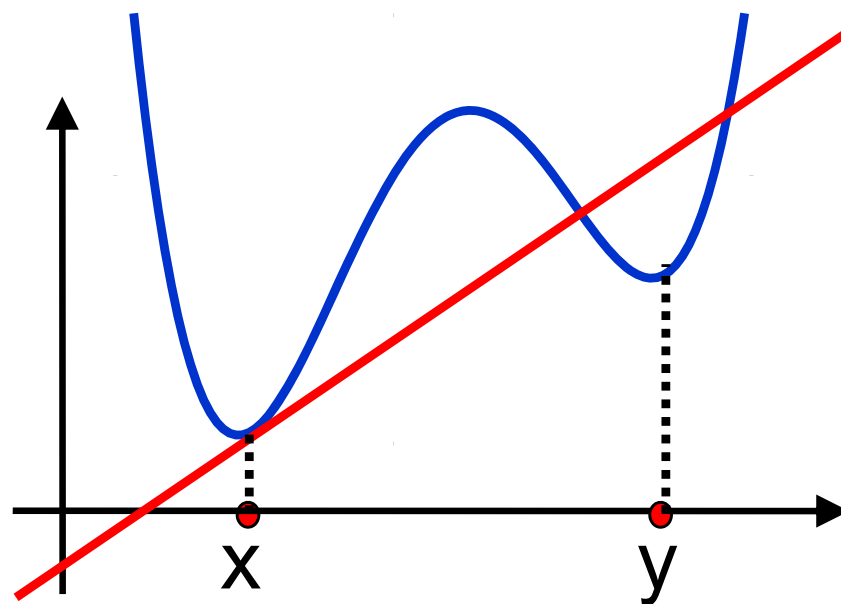
↔ 任意のベクトル x, y に対して次の不等式が成立

$$f(y) \geq f(x) + \nabla f(x)(y - x)$$

証明は略

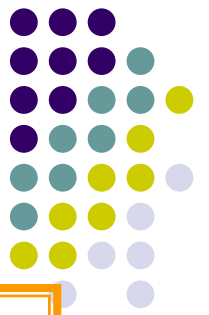


一変数凸関数の場合: x における
接線 $y = f(x) + \nabla f(x)(y - x)$
より $f(y)$ は上にある



一変数非凸関数の場合は
成り立たない

凸関数の最適解の必要条件



定理: f : 凸関数, 微分可能 (勾配ベクトルが定義可能) ならば
 x^* : f の停留点 ($\nabla f(x^*)=0$) $\iff x^*$ は制約なし問題の最適解

証明: 「 \leftarrow 」はすでに証明したので、「 \rightarrow 」を示す.
 f は凸関数なので, 任意の x, y に対して次が成り立つ

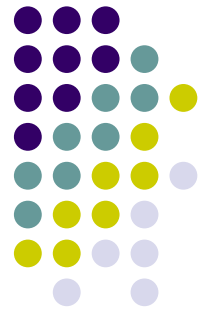
$$f(y) \geq f(x) + \nabla f(x)(y - x)$$

$x = x^*$ を代入すると, $\nabla f(x^*)=0$ なので

$$f(y) \geq f(x^*) + \nabla f(x^*)(y - x^*) = f(x^*)$$

ベクトル y は任意なので, x^* は最適解

凸関数の最適解の必要条件



定理: f : 凸関数, x^* : f の極小解
 $\leftrightarrow x^*$ は制約なし問題の最適解

証明: 「 \leftarrow 」は自明なので、「 \rightarrow 」を示す.

極小解の定義より, ある $\varepsilon > 0$ が存在して、

任意の x に対し $\|x - x^*\| < \varepsilon$ ならば $f(x) \geq f(x^*)$

$f(y) < f(x^*)$ なる y が存在すると仮定

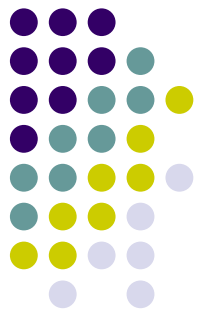
f は凸関数

$\Rightarrow 0 < t < 1$ なる任意の t に対して

$$f((1-t)y + tx^*) \leq (1-t)f(y) + tf(x^*) < f(x^*)$$

t を1に近づけると

$(1-t)y + tx^*$ と x^* の距離 $< \varepsilon$ (矛盾)



制約なし問題の解法1: 最急降下法

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

⇒ 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



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

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

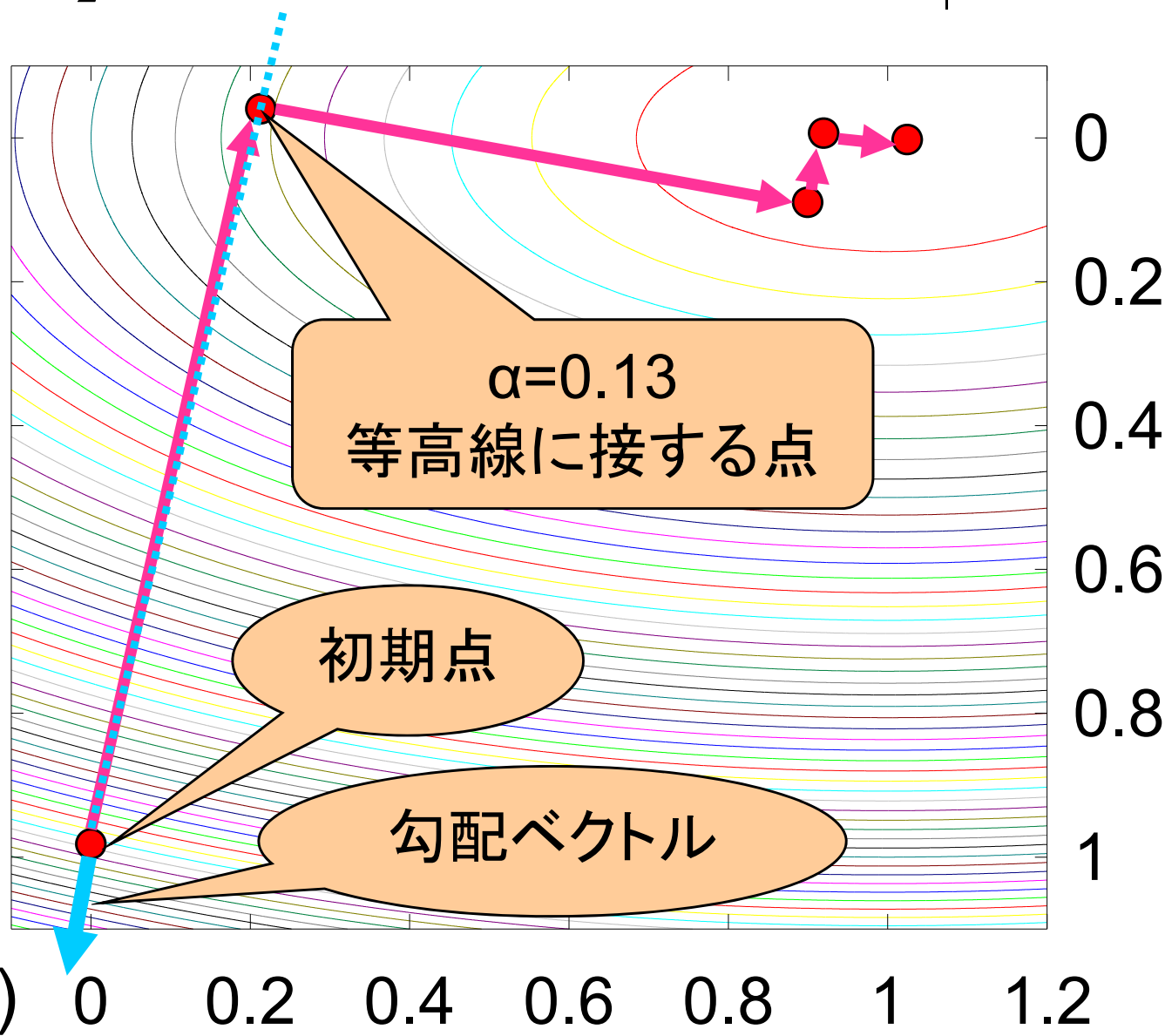
• $(x_1, x_2) = (0, 1)$ から
スタート

• $\nabla f(0, 1) = (-2, 8)$

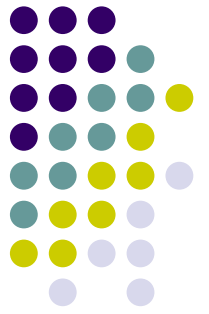
• $f(0 + 2\alpha, 1 - 8\alpha)$
を最小にするのは
 $\alpha = 0.13$

• 次の点は

$(x_1, x_2) = (0.26, -0.05)$



最急降下法のアルゴリズム



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

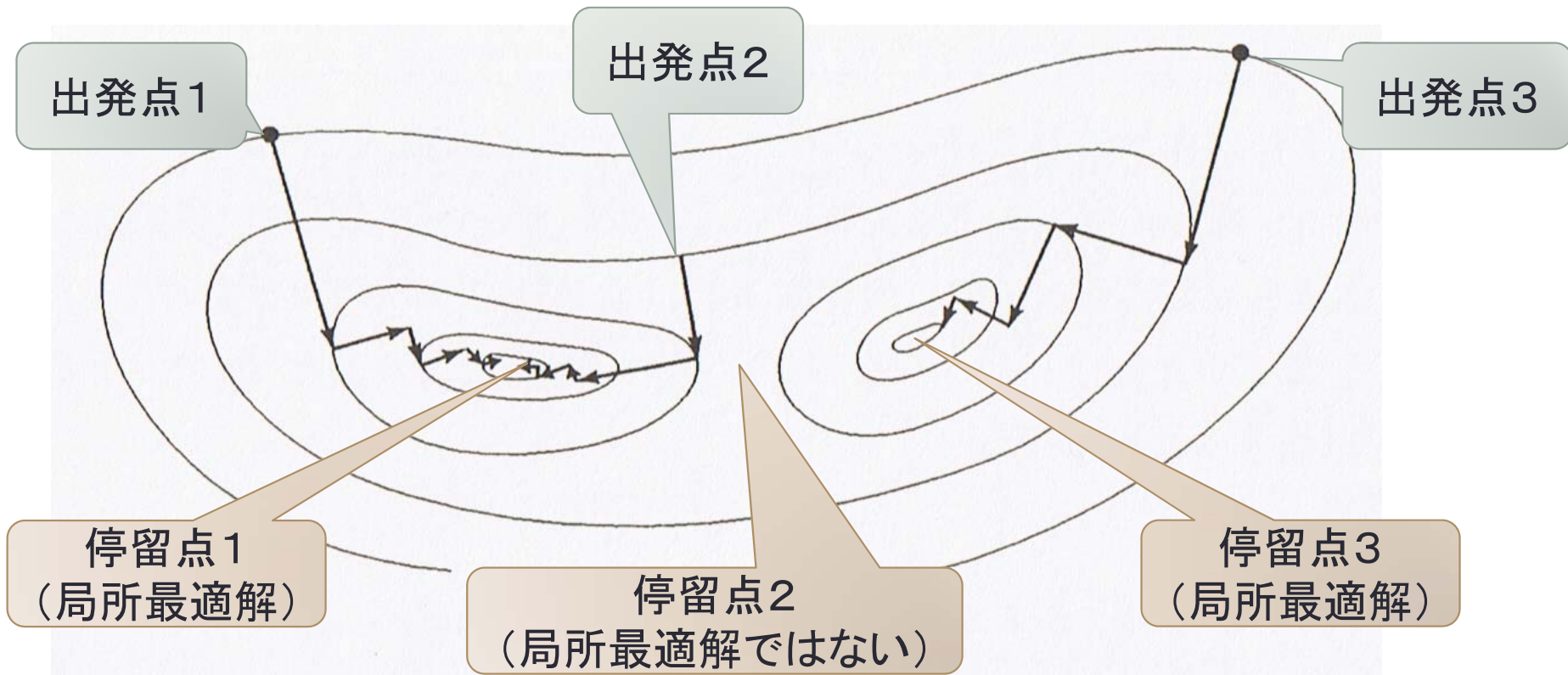
最小化 $f(x^k - \alpha \nabla f(x^k))$ 条件 $\alpha > 0$

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
 - 凸関数に対しては、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

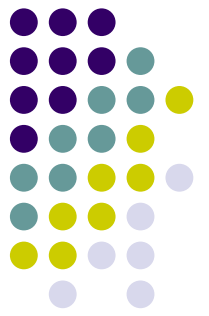
- 最適解に十分近いことをどうやって判定する？

(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

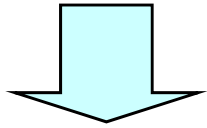
(方法2) 最適解の近くでは x^k があまり変化しない

→ $\|x^{k+1} - x^k\|$ の値が十分小さくなったら終了



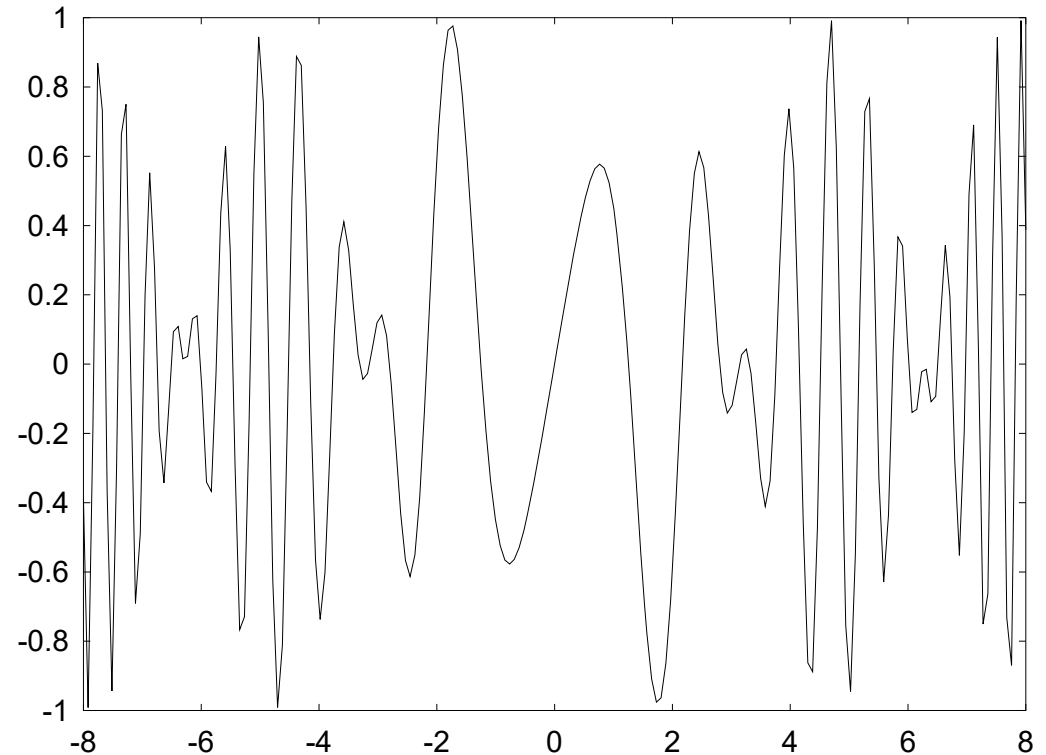
最適解の判定 (つづき)

- 非線形計画問題では
近似的最適解すら求めることが困難なことが多い



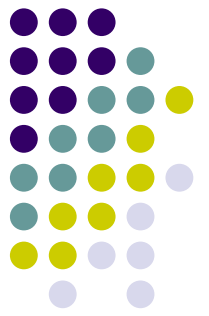
極小解または停留点を
求めることで我慢する

- 極小解は良い解であることが多い
- 凸関数では
極小解 \Leftrightarrow 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: 以下の関数 $f_1(x_1, x_2)$, $f_2(x_1, x_2)$, $f_3(x_1, x_2)$ に対して、 $(x_1, x_2) = (a_1, a_2)$ における一次のテイラー近似を求めなさい。

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

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

問題2: 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。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

問題3: (i) 1次関数 $f(x) = \sum_{i=1}^n a_i x_i + c$ ($x \in \mathbb{R}^n$) および
(ii) 1変数の絶対値関数 $f(x) = |x|$ ($x \in \mathbb{R}$)
が凸関数であることを証明せよ。

