数理手法（数理最適化）第12回
非線形計画
二次の最適性条件

塩浦昭義
東京工業大学 経営工学系 准教授

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

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

期末試験について

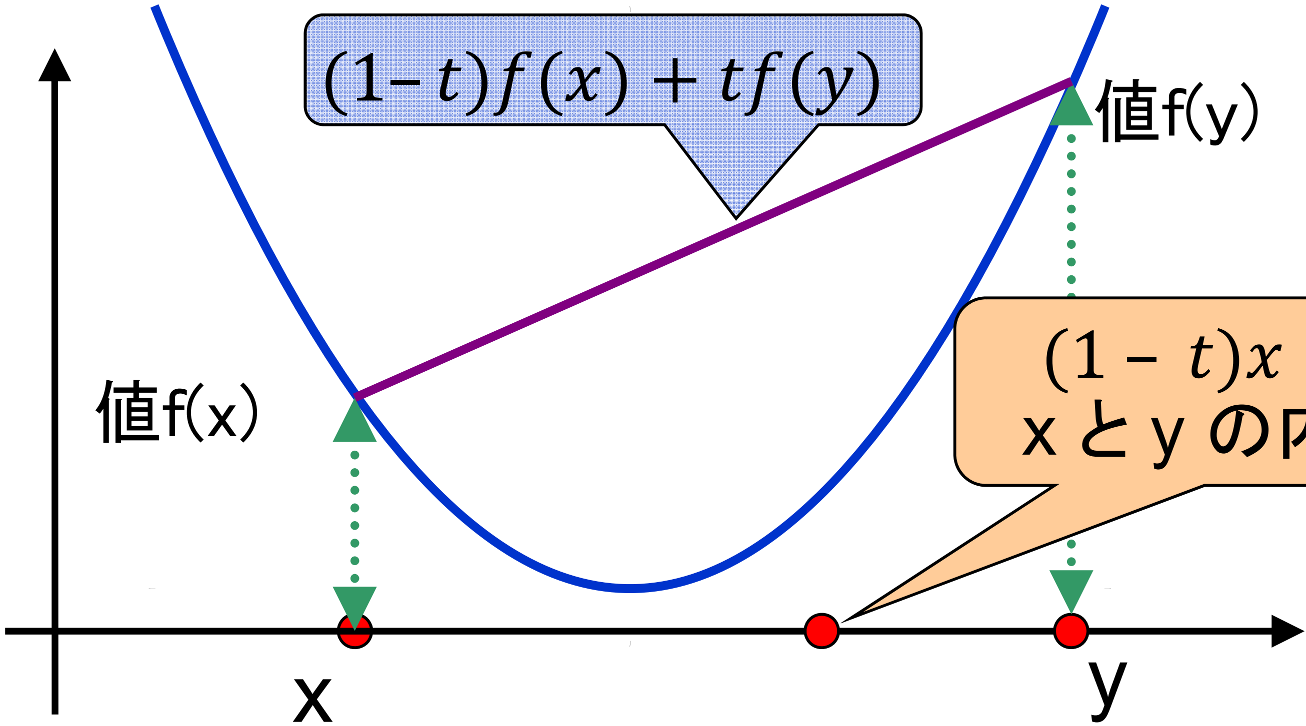
- 日時：1月17日（水）13:05～14:35
- 場所： 工2号館 212講義室（授業の部屋）
- 手書きのA4用紙一枚のみ持ち込み可（印刷やコピーは不可）
  - これも採点の対象，試験終了後に回収します
- 教科書，ノート等の持ち込みは不可
- 座席はこちらで指定
- 試験内容：第8回～第13回の講義で教えたところ
  - ネットワーク最適化
  - 非線形計画
- 50点満点，20点以下は不合格
- 中間と合わせて51点以上は合格，50点以下は単位不可

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

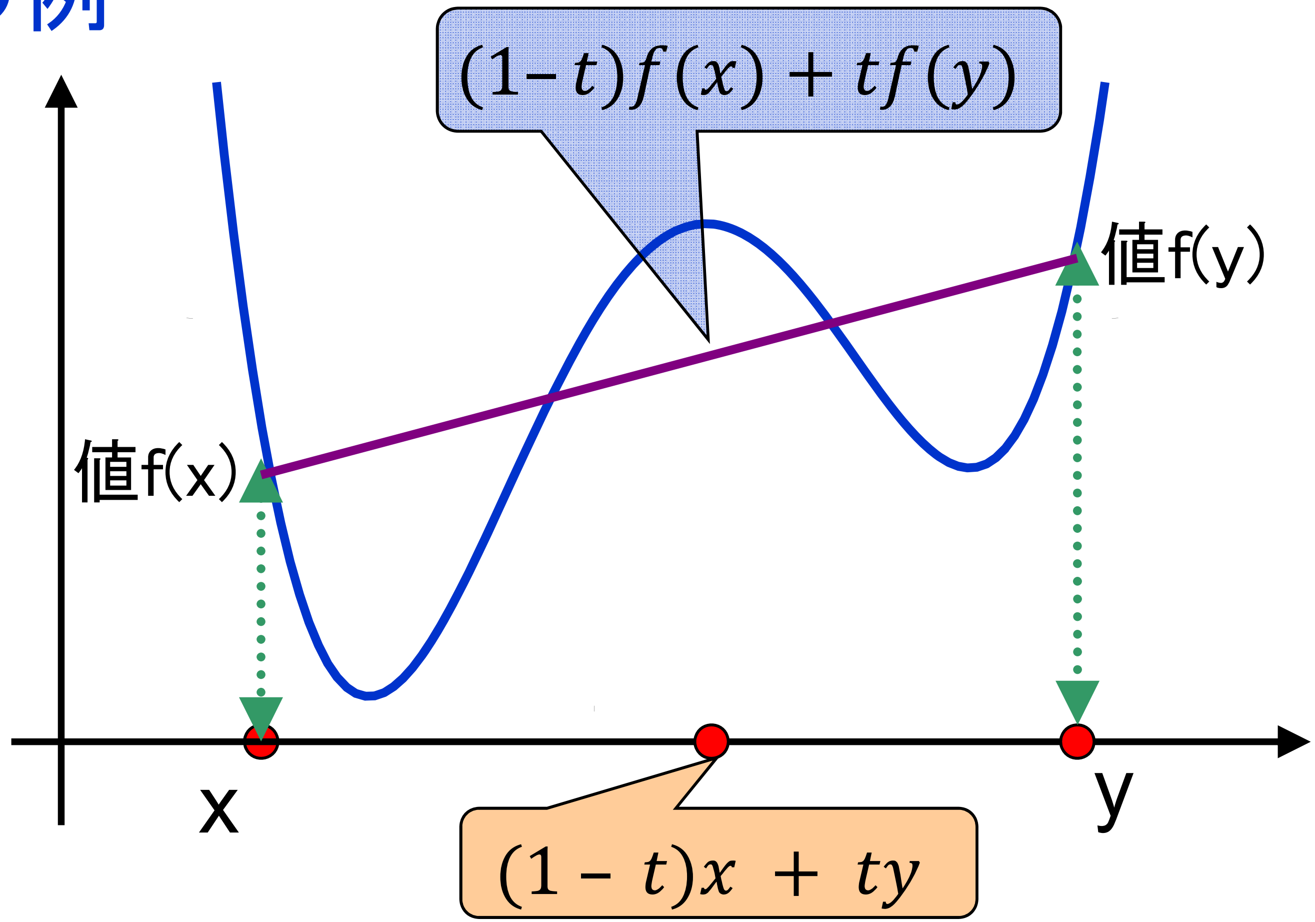
$$(1-t)f(x) + t\,f(y) \geq f((1-t)\,x + t\,y)$$

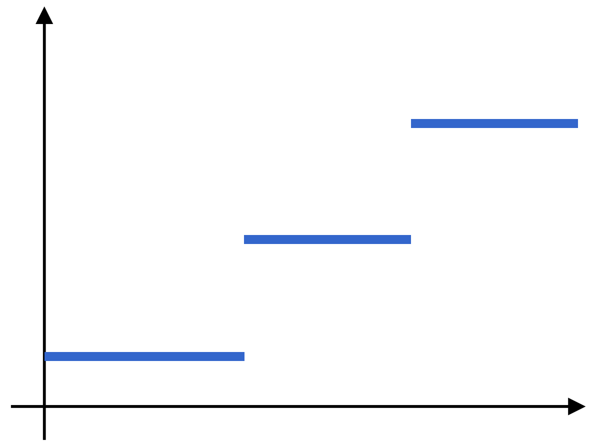


凸関数の定義(続き)

凸関数 ⇔ $(1-t)f(x)+t f(y) \geqq f((1-t)x+t y)$

非凸関数
の例



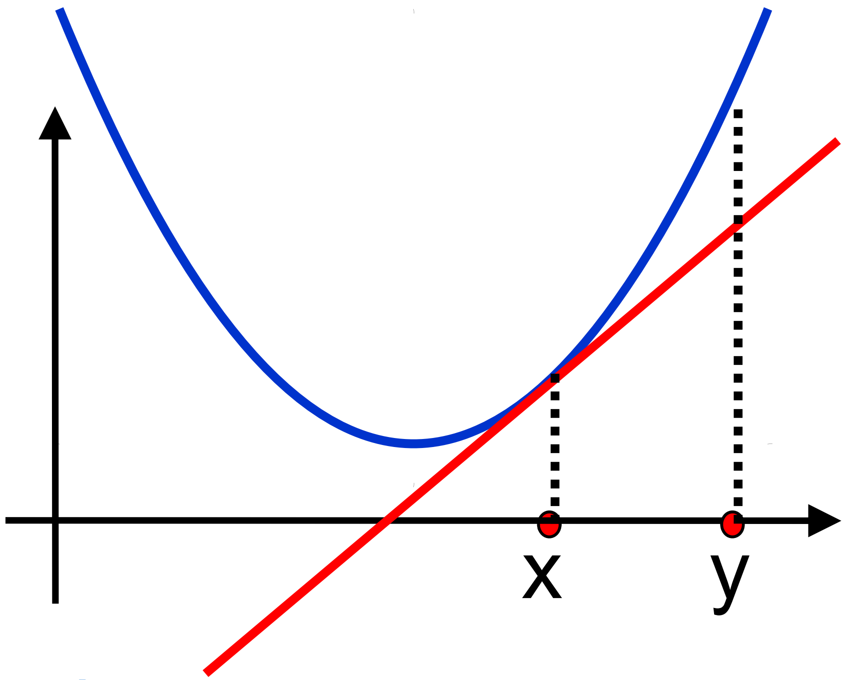


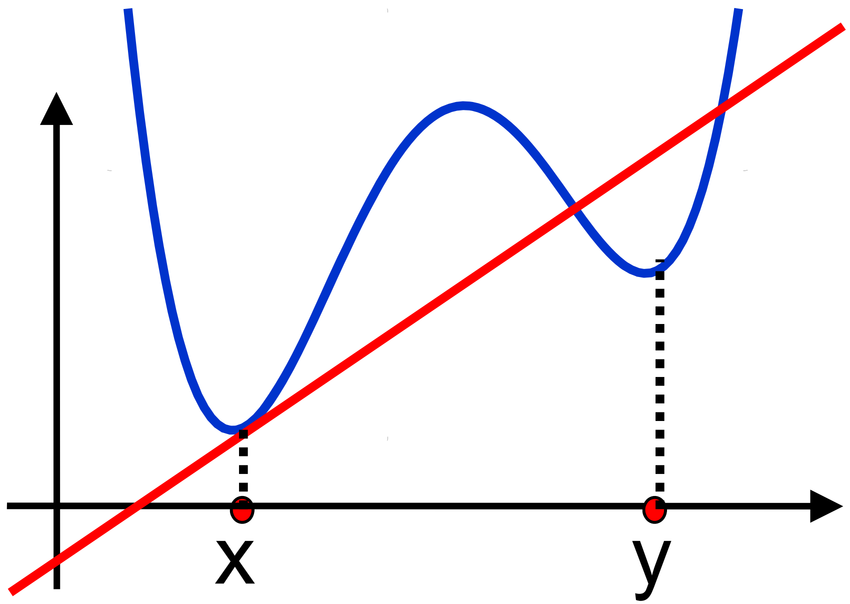
凸関数の特徴付け

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（∇f(x*)=0）←→ x*は制約なし問題の最適解

証明：「←」はすでに証明したので,「→」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，∇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
　　　　$\Leftarrow\Rightarrow$ 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 $f(x) \geqq f(x^*)$

$f(y) < f(x^*)$ 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
$f((1-t)\,y + t\,x^*) \leqq (1-t)\,f(y) + t\,f(x^*) < f(x^*)$
t を1に近づけると
　$(1-t)\,y + t\,x^*$ 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
⇒　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　$\alpha > 0$

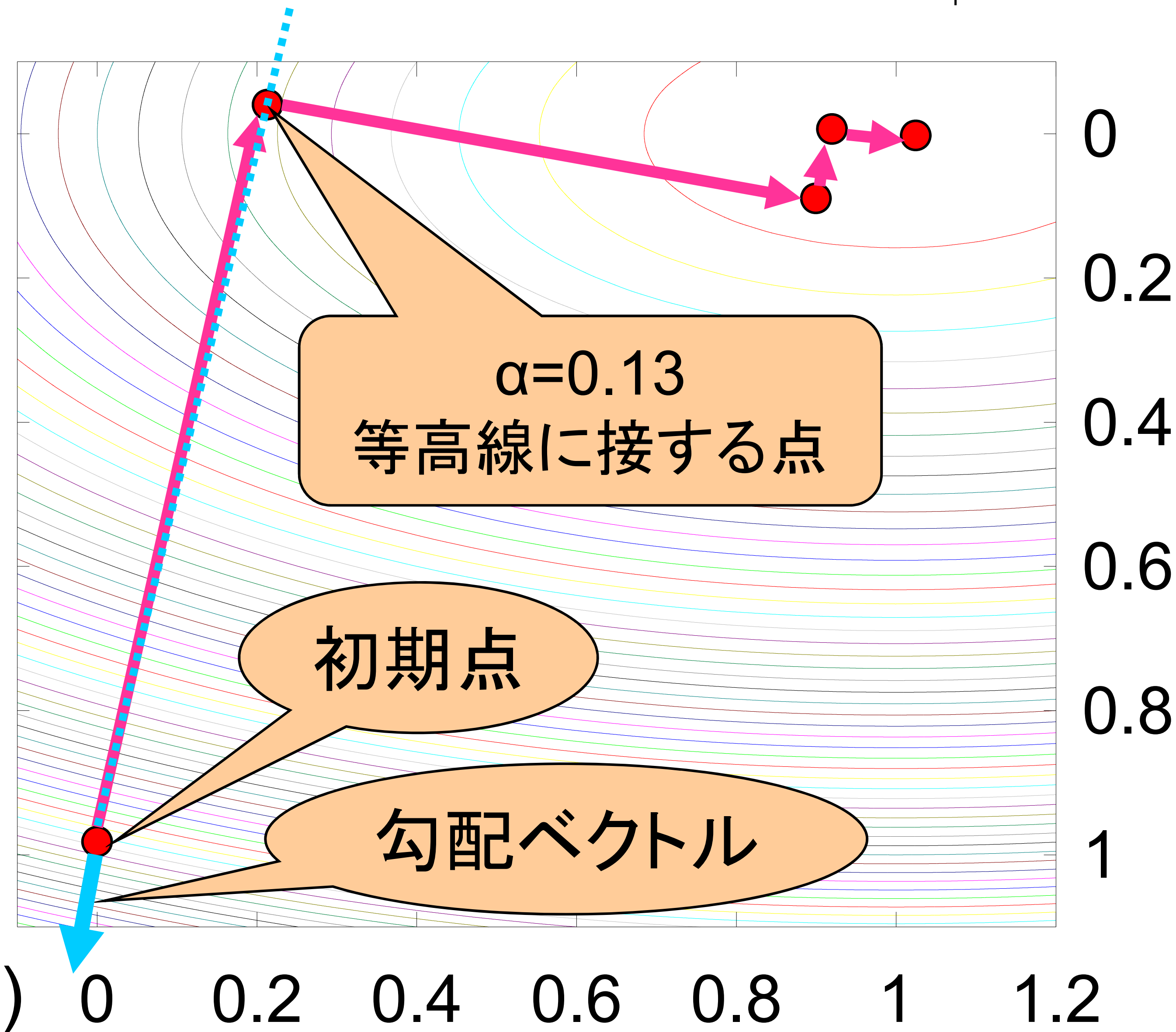
直線探索と呼ばれる

最急降下法の実行例

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

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

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

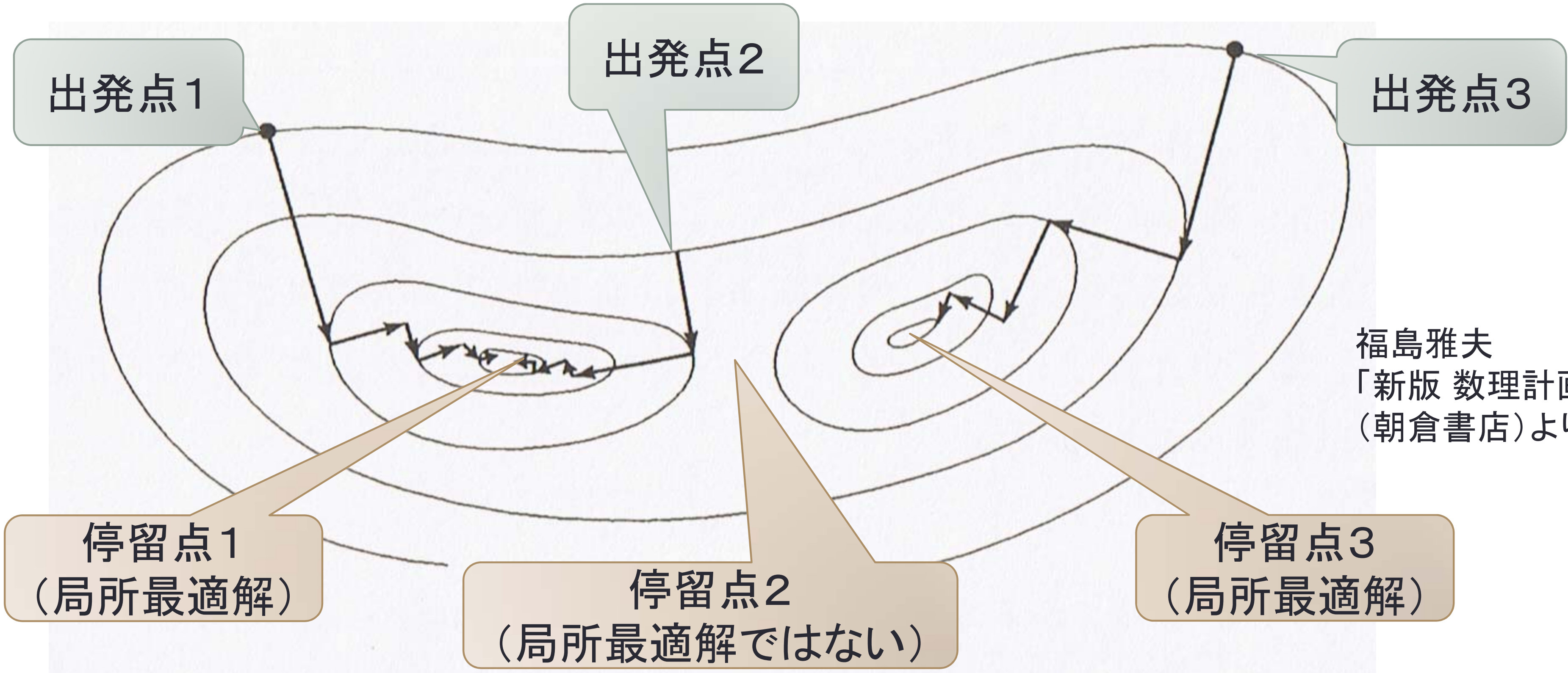
最小化　$f(x^k - \alpha \nabla f(x^k))$　条件　$\alpha > 0$

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



福島雅夫
「新版 数理計画入門」
（朝倉書店）より

- 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

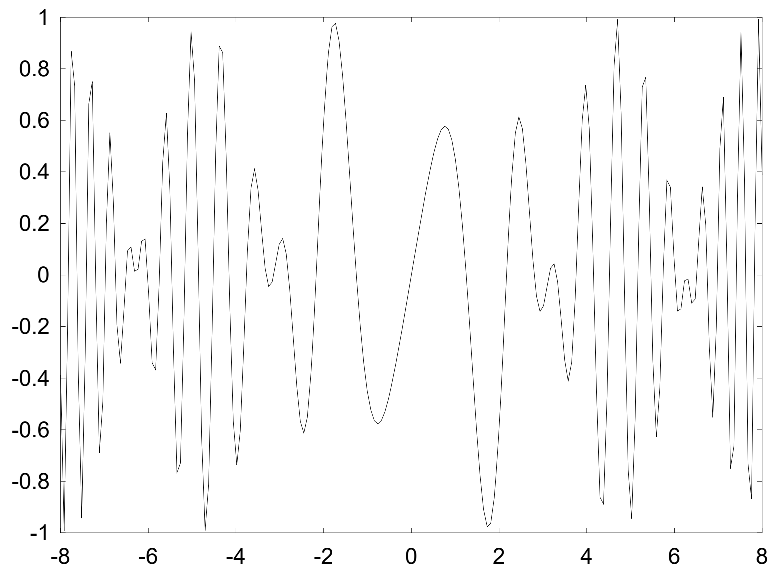
（方法2）最適解の近くでは x^k があまり変化しない

⇒ $\|x^{k+1} - x^k\|$ 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
求めることで妥協する



- 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

関数のヘッセ行列

- 定義：n変数関数 f のヘッセ行列 $\mathrm{H}f(x)$

　　　$\Leftrightarrow$ f の2次偏微分係数を要素とするｎｘｎ行列

$$\mathrm{H}f(x_1, x_2, \ldots, x_n) = \begin{bmatrix} \frac{\partial^2 f(x)}{\partial x_1^2} & \frac{\partial^2 f(x)}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f(x)}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f(x)}{\partial x_2 \partial x_1} & \frac{\partial^2 f(x)}{\partial x_2^2} & \cdots & \frac{\partial^2 f(x)}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f(x)}{\partial x_n \partial x_1} & \frac{\partial^2 f(x)}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f(x)}{\partial x_n^2} \end{bmatrix}$$

- f が1変数関数のときは，ヘッセ行列は2階微分と同じ

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

ヘッセ行列の例

例:

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

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

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

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

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

二次のテイラー展開

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

関数$f(x)$の$x=a$における
二次のテイラー展開

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

関数$\psi(d)$は $d = (d_1, d_2, \ldots, d_n)$ に関する
3次以上の項から構成される n 変数多項式関数
（定数項，一次の項，二次の項は含まれない）
➔任意のベクトル d に対して

$$\lim_{\varepsilon \to 0} \frac{\varphi(\varepsilon d)}{\varepsilon^2} = 0$$

二次のテイラー近似

関数$f(x)$の$x = a$における二次のテイラー展開

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

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

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

関数$f(x)$の$x = a$における
二次のテイラー近似

- 二次関数
  $\nabla \tilde{f}(a) = \nabla f(a), \mathrm{H}\tilde{f}(a) = \mathrm{H}f(a)$
- $x \simeq a$のとき $\tilde{f}(x) \simeq f(x)$,
  とくに $\tilde{f}(a) = f(a)$

二次のテイラー近似の例

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

$$
\begin{aligned}
f_1(x) &= a^2 + (2a)(x-a) + \frac{1}{2}\cdot 2(x-a)^2 + \psi(x-a) \\
&= x^2 + \psi(x-a)
\end{aligned}
$$

つまり, $\psi(x = a) = 0$であり,

f_1の二次のテイラー近似 $= f_1$そのもの

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

の二次のテイラー近似は f に一致

V: $n \times n$行列
c: n 次元ベクトル
c_0: 定数

二次のテイラー近似の例

例2：　$f_2(x) = \log x \;\Rightarrow f'(x) = \frac{1}{x}, f''(x) = -\frac{1}{x^2}$

f_2 の x=1 における二次のテイラー展開

$$f_2(x) = \log 1 + \frac{1}{1}(x-1) - \frac{1}{1^2}(x-1)^2 + \psi(x-a)$$

なお，$\psi(x-a) = \frac{(x-1)^3}{3} - \frac{(x-1)^4}{4} + \frac{(x-1)^5}{5} - \cdots$

二次のテイラー近似の例

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

$\nabla f(x) = 5x^4 - 15x^2 + 4$　　　　$\mathrm{H}\, f(x) = 20x^3 - 30x$

a = -1 のとき

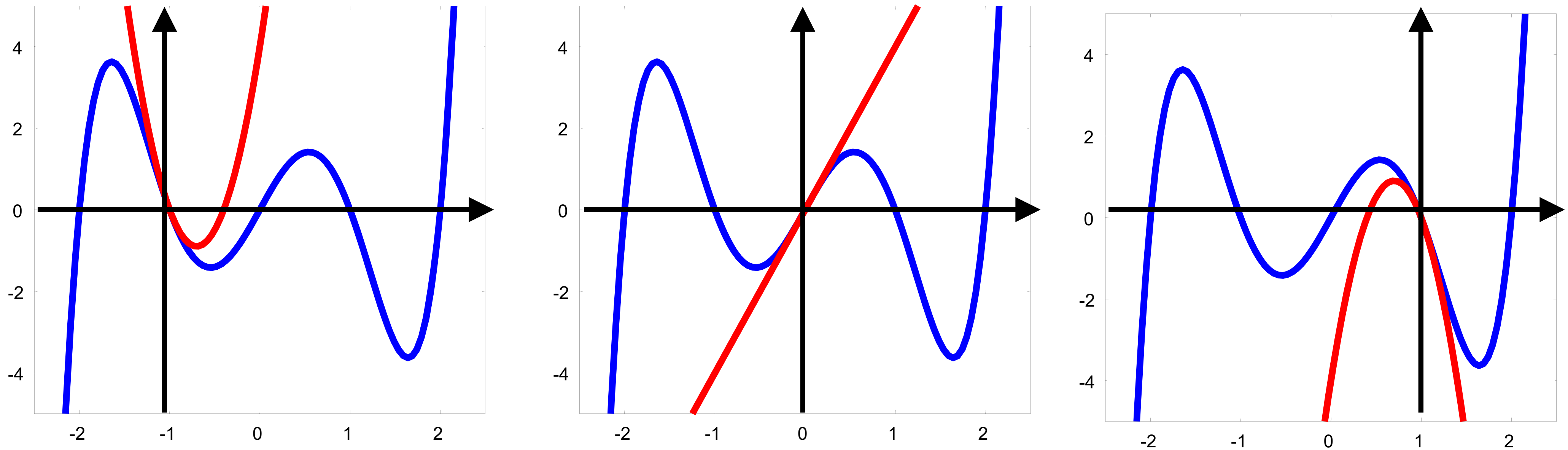
$0 - 6(x+1) + 5(x+1)^2$

a = 0 のとき

$0 + 4x + 0x^2$

a = 1 のとき

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



極小解，極大解の判定方法

- 一変数関数 f の場合
  - 極小，極大ならば傾き（一回微分） $f'(x)$ が 0
  - 極小，極大は二回微分 $f''(x)$ を使って判定
    - $f''(x) > 0$ ならば極小，$f''(x) < 0$ ならば極大
    - $f''(x) = 0$ の場合は不明
- 多変数関数の場合
  - 極小，極大ならば勾配ベクトル $f'(x)$ がゼロベクトル
  - 極小，極大はヘッセ行列 $\mathrm{H}f(x)$ を使って判定
    - どうやって判定？ --- 行列の（半）正定値性を使う

正定値（半正定値）・・・・行列が「正（非負）」

行列の正定値性、半正定値性

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 ⇔ $a_{11} > 0$

※ A が2×2対称行列のとき、

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

行列の正定値性、半正定値性

※Ａが2×2対称行列のとき、

Aは正定値 ⇔ $a_{11} > 0$, $a_{22} > 0$, $a_{11}a_{22} - a_{12}^2 > 0$

$a_{11} \neq 0$ のとき　$y^T A y = a_{11}\left(y_1 + \frac{a_{12}}{a_{11}} y_2\right)^2 + \left(a_{22} - \frac{a_{12}^2}{a_{11}}\right) y_2^2$　（☆）

であることを使うと証明できる

[➔の証明] $y^T A y$ に $y=(1,0)$, $(0,1)$, $(-a_{12}/a_{11}, 1)$ を代入すれば得られる

[🡐の証明] $y \neq 0$ より $\left(y_1 + \frac{a_{12}}{a_{11}} y_2\right) \neq 0$ または $y_2 \neq 0$

$\therefore \left(y_1 + \frac{a_{12}}{a_{11}} y_2\right)^2 > 0$ または $y_2^2 > 0$

上記の等式（☆）および仮定した不等式条件より，$y^T A y > 0$

2次の凸関数(再掲)

凸関数 ⇔ $(1-t)f(x) + t\,f(y) \geqq f((1-t)x + t\,y)$

より一般に,

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

（V: n ×n 行列, c: n次元ベクトル, c_0: 定数）

はＶが半正定値行列 ➔ 凸関数

例：

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

2次の最適性条件（必要条件）

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：

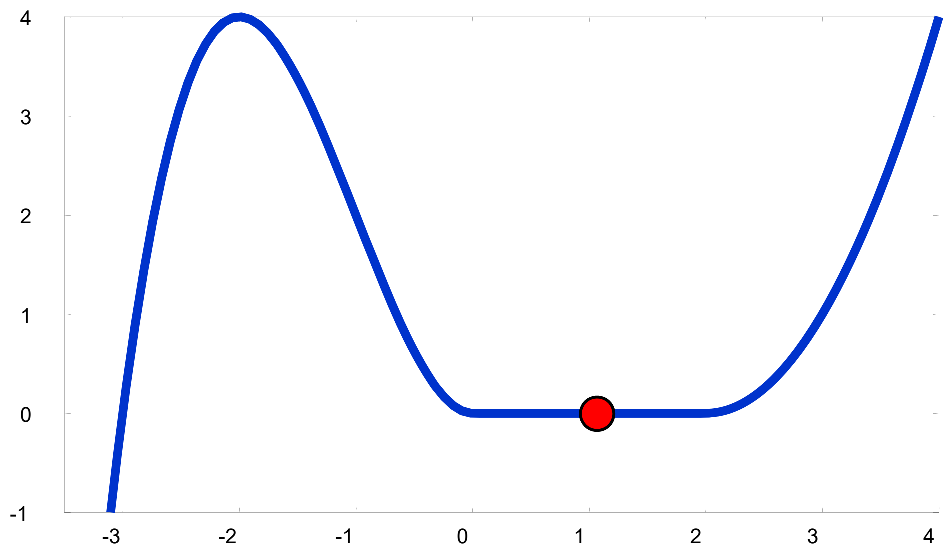
$x^* = 1$ は極小解

$0 \leqq x \leqq 2$ の範囲で $f(x) = 0$

⇒ $\nabla f(x^*) = f'(x^*) = 0$

$Hf(x^*) = f''(x^*) = 0$

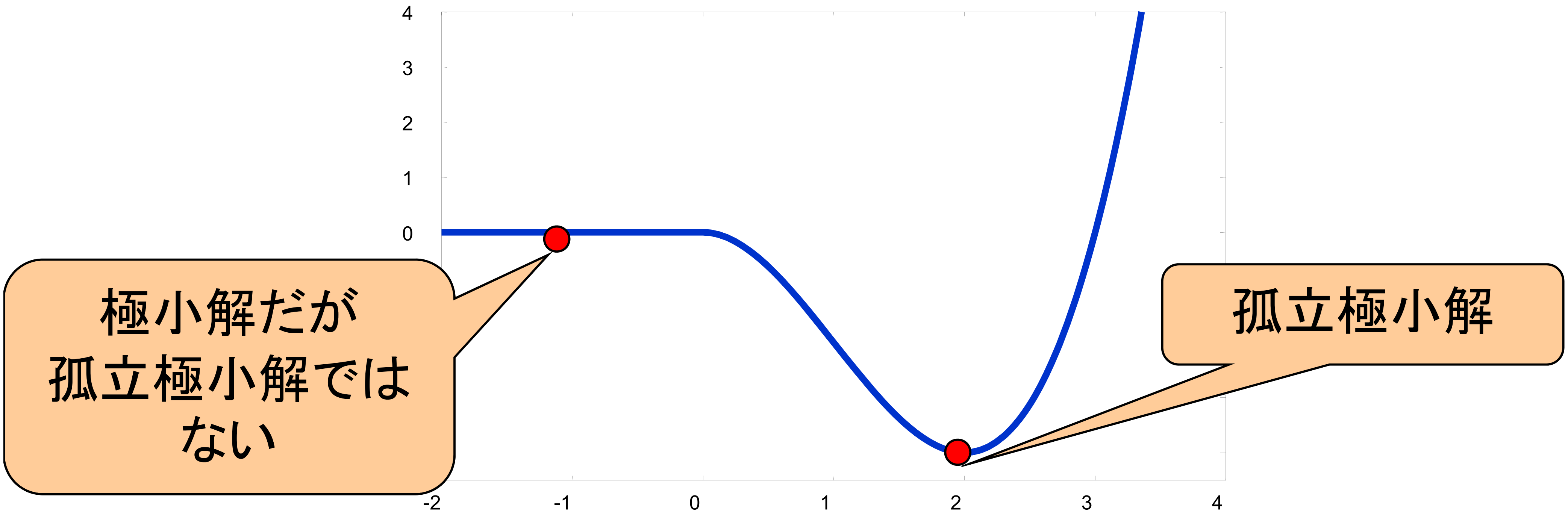
半正定値



2次の最適性条件（十分条件）

定理（2次の十分条件）：
x^* は停留点，　$Hf(x^*)$ は正定値
　⇒　x^*: 制約なし問題の（孤立）極小解

定義：x^* は孤立極小解
⇔　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)の例

定理:Hf(x*) は正定値 ⇒ (孤立)極小解

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

勾配を計算: $f'(x) = (x+2)x^2(x-2)(x-3)$

⇒ 停留点はx = -2, 0, 2, 3

2階微分を計算:

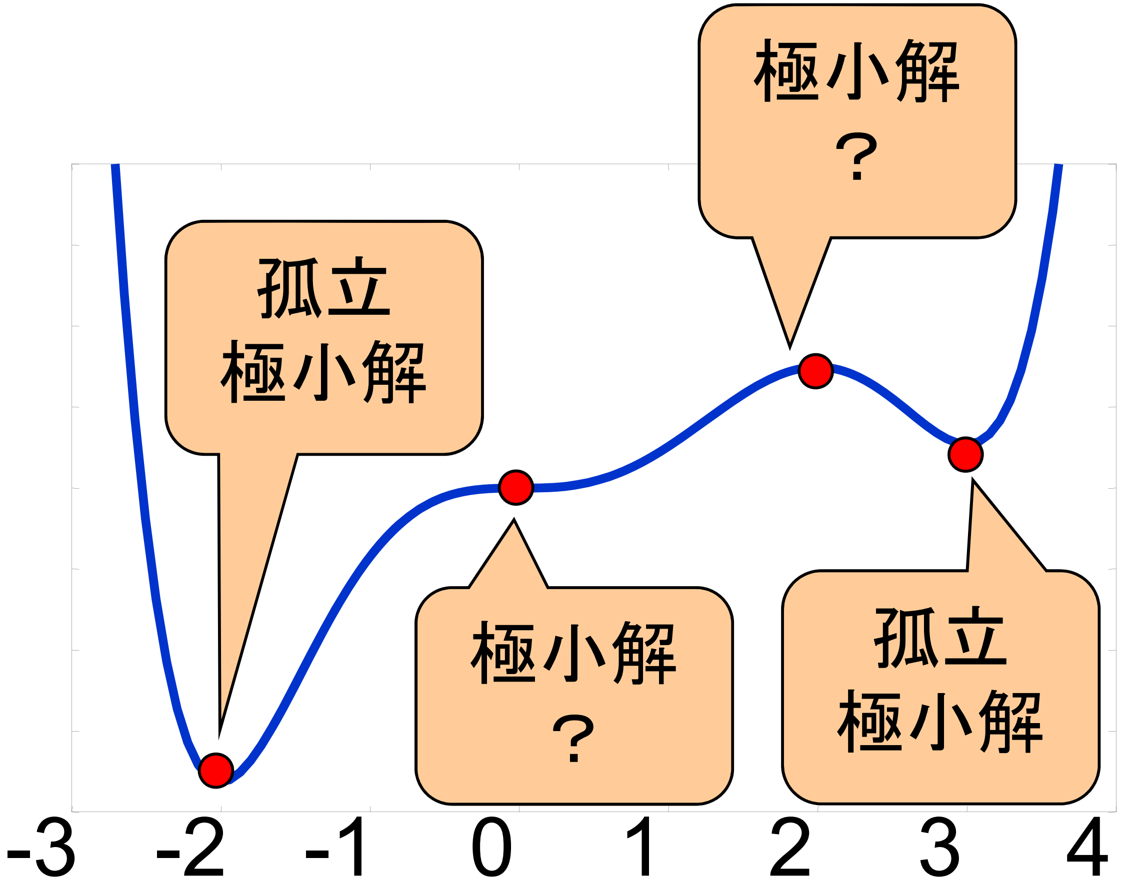
$$f''(x) = 5x^4 - 12x^3 - 12x^2 + 24x$$

⇒ 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）の例

> 定理： x* は停留点，Hf(x*) は正定値
> ⇒　x*:（孤立）極小解

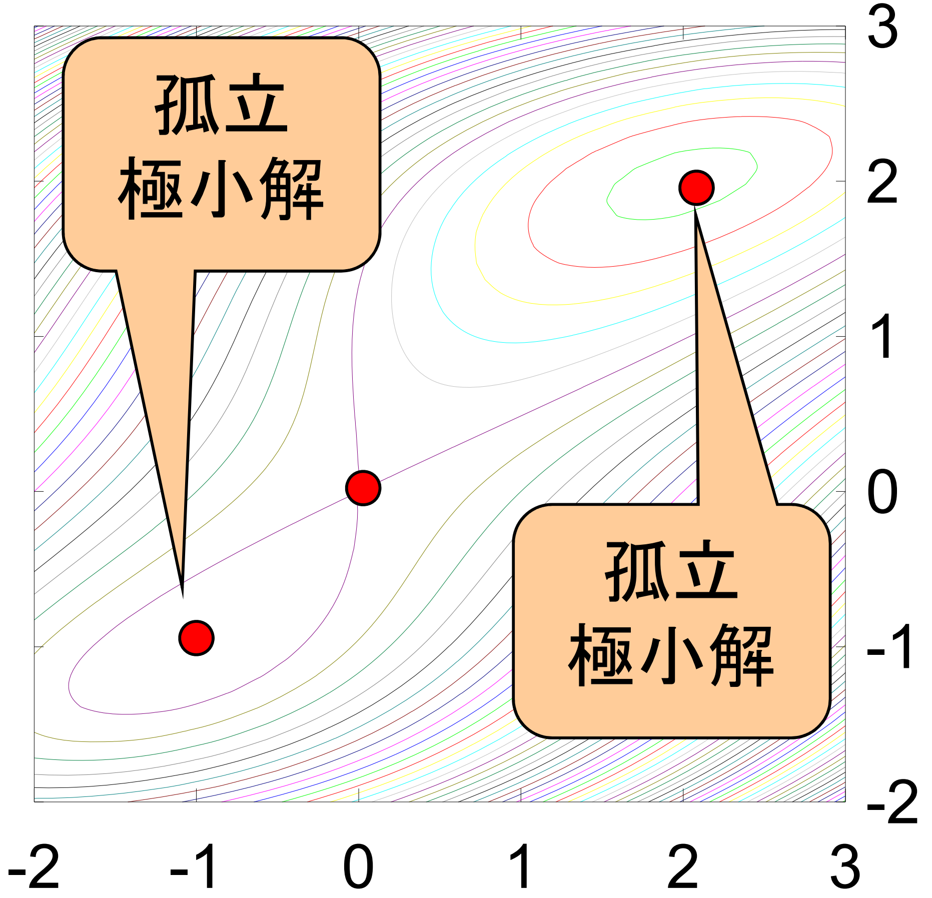
例2 $$f(\boldsymbol{x}) = x_1^2 - 2x_1x_2 + \frac{1}{4}x_2^4 - \frac{1}{3}x_2^3$$

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

⇒ 停留点は(0,0), (-1, -1), (2, 2)

$$\mathrm{H} f(\mathbf{x}) = \begin{pmatrix} 2 & -2 \\ -2 & 3x_2^2 - 2x_2 \end{pmatrix}$$

⇒ (-1, -1), (2, 2) は孤立極小解



2次の最適性条件の例

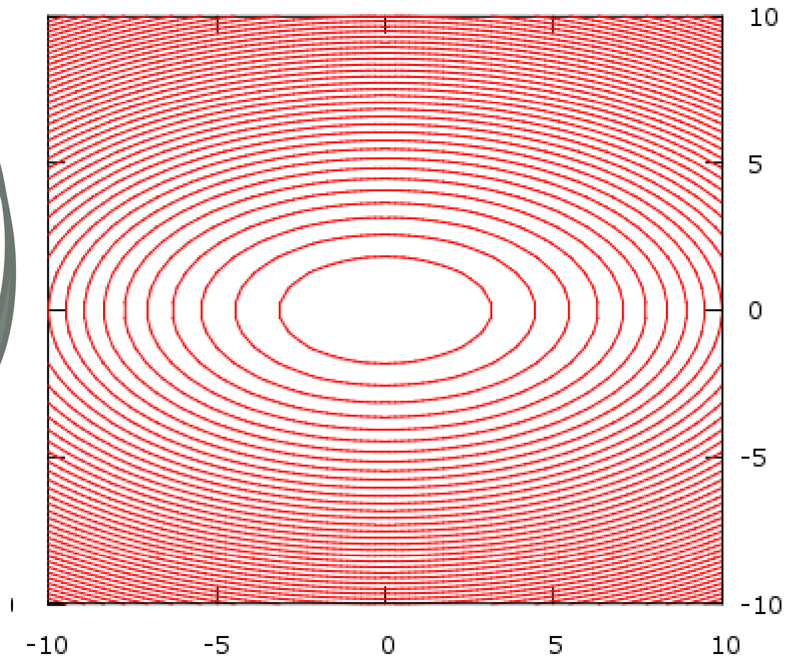
例3 : $f(x_1, x_2) = x_1^2 + 3x_2^2$

- $\nabla f(x_1, x_2) = \begin{bmatrix} 2x_1 \\ 6x_2 \end{bmatrix}, \mathrm{H}f(x_1, x_2) = \begin{bmatrix} 2 & 0 \\ 0 & 6 \end{bmatrix}$
- $\nabla f(x_1, x_2)$がゼロベクトルとなるのは (0,0) のみ🡐停留点
- $\nabla^2 f(0,0) = \begin{bmatrix} 2 & 0 \\ 0 & 6 \end{bmatrix}$ は正定値行列

  ➔(0, 0) は孤立極小解

任意の非ゼロベクトル(y_1, y_2)に対して

$$\begin{bmatrix} y_1 & y_2 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 6 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = 2y_1^2 + 6y_2^2 > 0$$



2次の最適性条件の例

例4 : $f(x_1, x_2) = x_1^2 + x_2^4$

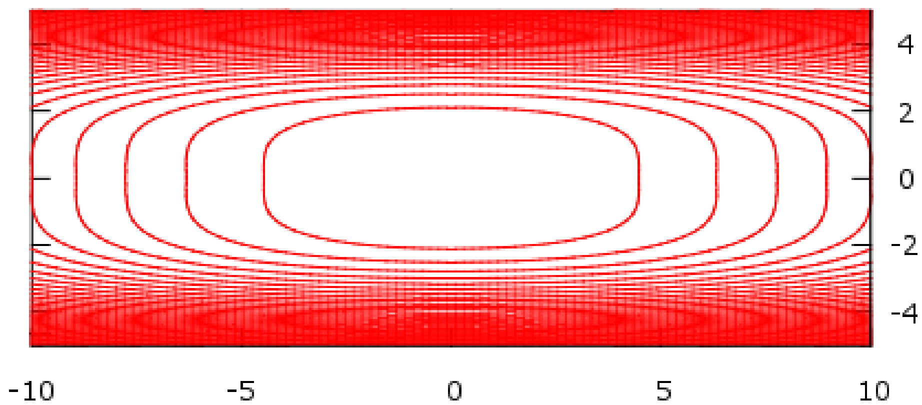
- $\nabla f(x_1, x_2) = \begin{bmatrix} 2x_1 \\ 4x_2^3 \end{bmatrix}, \nabla^2 f(x_1, x_2) = \begin{bmatrix} 2 & 0 \\ 0 & 12x_2^2 \end{bmatrix}$
- $\nabla f(x_1, x_2)$がゼロベクトルとなるのは (0,0) のみ（実は最適解）
- $\nabla^2 f(0,0) = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}$ は半正定値だが，正定値ではない

➔(0, 0) が極小解かどうかは，ヘッセ行列を使って
判定できない（実際には極小解）

任意のベクトル(y_1, y_2)に対して

$$\begin{bmatrix} y_1 & y_2 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = 2y_1^2 \geq 0$$

$y_1 = 0$のときは$y_2 \neq 0$でも値は0



2次の最適性条件（必要条件）証明

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**

証明: $A=Hf(x^*)$ とおく.
背理法: A は半正定値でないと仮定

➔ ある単位ベクトル y に対し，$y^TAy<0$

以下に示すように，
ある $\varepsilon>0$ に対して $f(x^*+\varepsilon' y)<f(x^*)$ $(0<\forall\varepsilon'<\varepsilon)$ となり，矛盾.
$x = x^*$ での2次のテイラー展開と $\nabla f(x^*)=0$ を使うと，

$$f(x^*+\varepsilon y)=f(x^*)+\nabla f(x^*)^T(\varepsilon y)+\frac{1}{2}(\varepsilon y)^T A(\varepsilon y)+\psi(\varepsilon y)$$

$$=f(x^*)+\varepsilon^2\left(\frac{1}{2}y^TAy+\frac{\psi(\varepsilon y)}{\varepsilon^2}\right)$$

テイラー展開の性質より，ある $\varepsilon>0$ が存在して，$0<\forall\varepsilon'<\varepsilon$ に対して

$$\frac{1}{2}y^TAy+\frac{\psi(\varepsilon' y)}{(\varepsilon')^2}<0 \quad \therefore\ f(x^*+\varepsilon' y)<f(x^*)$$

2次の最適性条件(十分条件)証明

定理(2次の十分条件):

x* は停留点, Hf(x*) は正定値 ⇒ x*: 制約なし問題の(孤立)極小解

証明の概略:

x = x* での2次のテイラー展開 $\tilde{f}$ を考えると,

$\tilde{f}$ は凸関数, x* が最小解 ∴ x* は $\tilde{f}$ の極小解

x* のある近傍において, $\tilde{f}$ と f は十分に近い ∴ x* は f の極小解

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

凸関数の特徴付け(その2)

定理: f: 凸関数, 微分可能(ヘッセ行列が定義可能)
⟺ 任意のベクトル x に対して
ヘッセ行列 $Hf(x)$ が半正定値

証明は略

一変数凸関数の場合:
関数 f は凸関数⟺任意のx に対して二階微分$f''(x) \geq 0$

演習問題

問 1：次の3つの関数に対し，$(x_1, x_2) = (1,2)$における二次のテイラー近似を求めなさい（ log の底はeとする）

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

$$f_3(x, y) = x^2 - xy + 2y^2 - 2x + e^{x+y}$$

問 2 ：関数$f(x, y) = \frac{1}{4}x^4 - \frac{1}{2}x^2 + \frac{1}{3}y^3 + 2y^2$ について考える

(a) 勾配ベクトルとヘッセ行列を計算せよ.

(b) すべての停留点（勾配ベクトルがゼロの点）を求めよ.

さらに，2次の最適性条件（十分条件）を用いて，極小解を求めよ.

問3：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 ⇔ $a_{11} \geqq 0$, $a_{22} \geqq 0$, $a_{11}a_{22} - a_{12}^2 \geqq 0$