数理計画法 第12回

塩浦昭義
情報科学研究科 准教授

shioura@dais.is.tohoku.ac.jp
http://www.dais.is.tohoku.ac.jp/~shioura/teaching

期末試験について

- 日時：2月2日（木）午後1時～2時30分（日時が確定しました）
- 受験資格者：以下の条件を満たす学生
  - 中間試験に合格している
  - ネットワーク計画と非線形計画に関するレポートを1回以上提出
- 教科書等の持込は不可
- 座席はこちらで指定
- 試験内容：ネットワーク計画，非線形計画の範囲（次回の内容まで）
  - （詳しくはWeb上の過去問を参考にしてください）

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：

- 勾配ベクトルの逆方向$d = -\nabla f(x)$は，目的関数値の減少率が最も大きい方向
- 勾配ベクトルと逆の方向に（少し）進むと関数値が減る

現在の点 x を $x + \alpha d$ $(d = -\nabla f(x))$ により繰り返し更新，
関数値 $f(x)$ を減らす

ステップ幅

ステップ幅の選び方：

次の1変数最小化問題を（近似的に）解く

目的関数： $f(x + \alpha d) \to$ 最小化

制約条件： $\alpha \geq 0$

直線探索と呼ばれる

最急降下法のアルゴリズム

ステップ0：出発点$x^{(0)}$を選択，$k \coloneqq 0$とおく

ステップ1：$\nabla f\left(x^{(k)}\right) = 0$ならば終了$\left(x^{(k)}\right.$は最適解$\left.\right)$.

そうでなければ，$d^{(k)} = -\nabla f\left(x^{(k)}\right)$とおく.

ステップ2：直線探索問題

目的関数：$f\left(x^{(k)} + \alpha d^{(k)}\right) \to$ 最小化，制約条件：$\alpha \geq 0$

を近似的に解き、解（ステップ幅）を$\alpha^{(k)}$とする.

$x^{(k+1)} \coloneqq x^{(k)} + \alpha^{(k)} d^{(k)}$とおく.

$k \coloneqq k + 1$とおいて、ステップ1に戻る

※現実には，$\nabla f\left(x^{(k)}\right) = 0$ が満たされることは期待できない

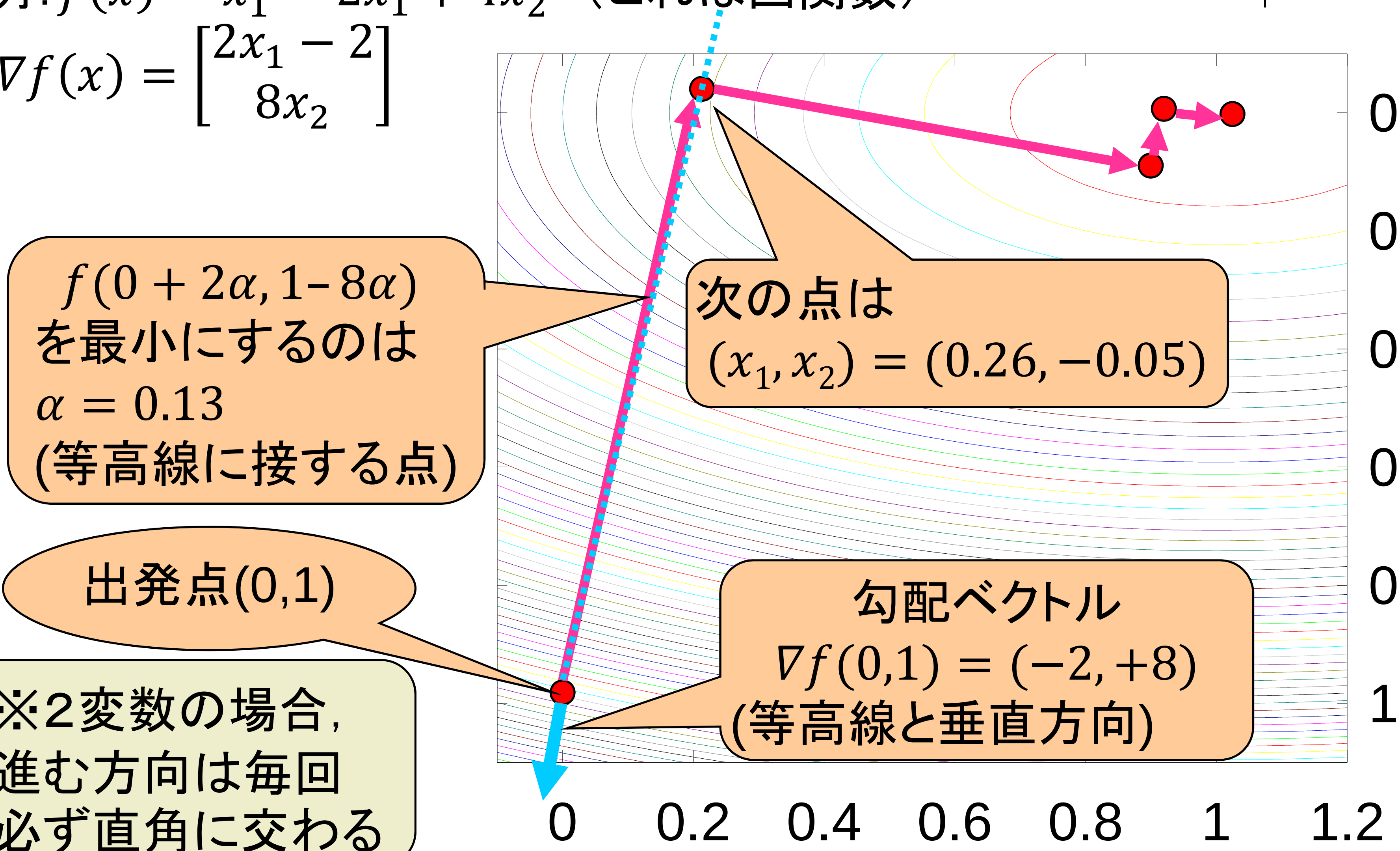
➔緩和した条件「勾配ベクトルのノルム（長さ）が十分に小さい」

$||\nabla f\left(x^{(k)}\right)|| < \varepsilon$ (εは十分小さい正数)に置き換える

最急降下法の実行例その1

例 : $f(x) = x_1^2 - 2x_1 + 4x_2^2$ （これは凸関数）

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



※2変数の場合，
進む方向は毎回
必ず直角に交わる

最急降下法の実行例その2

教科書113ページの図4．4参照

- 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

最急降下法の性質

- 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
- ある種の凸関数（狭義凸2次関数）に対しては，一次収束をする
  - 一次収束：「毎回の反復で，現在の点と極限との距離が，一定の割合で減少すること」

    ある定数$0 < \beta < 1$とある整数k_0に対して，

    $k \geq k_0$ ならば

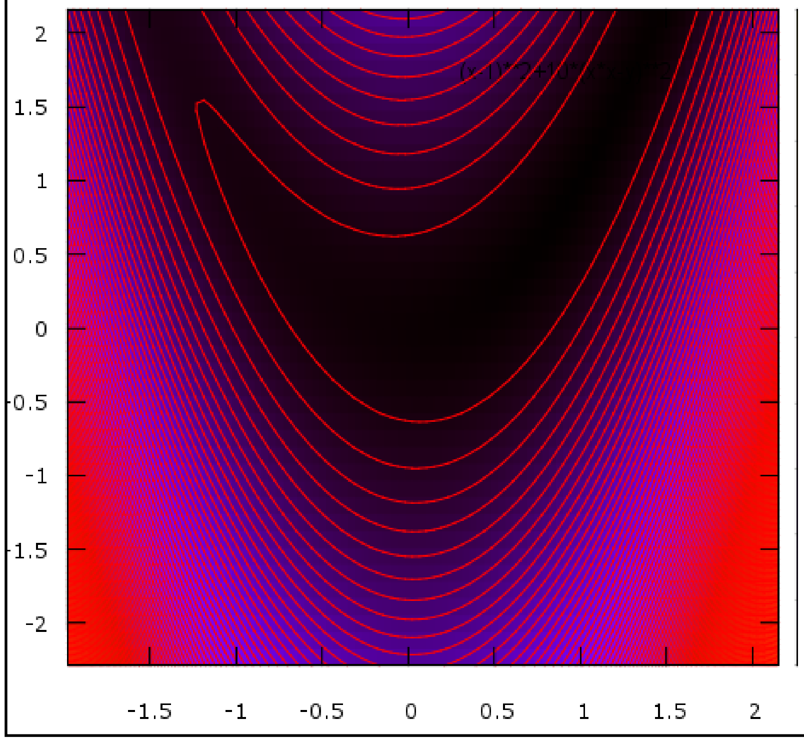
    $$||x^{(k+1)} - x^*|| \leq \beta ||x^{(k)} - x^*||$$

    が成立（x^*は極限の点，収束先）
  - 一次収束する場合でも，収束のスピードは遅いこともある

一次収束の例1

- 関数 $f(x) = (x_1 - 1)^2 + 10(x_1^2 - x_2)^2$ に適用
  - 最適解は(1,1)
  - β ＝ 約0.99なので，

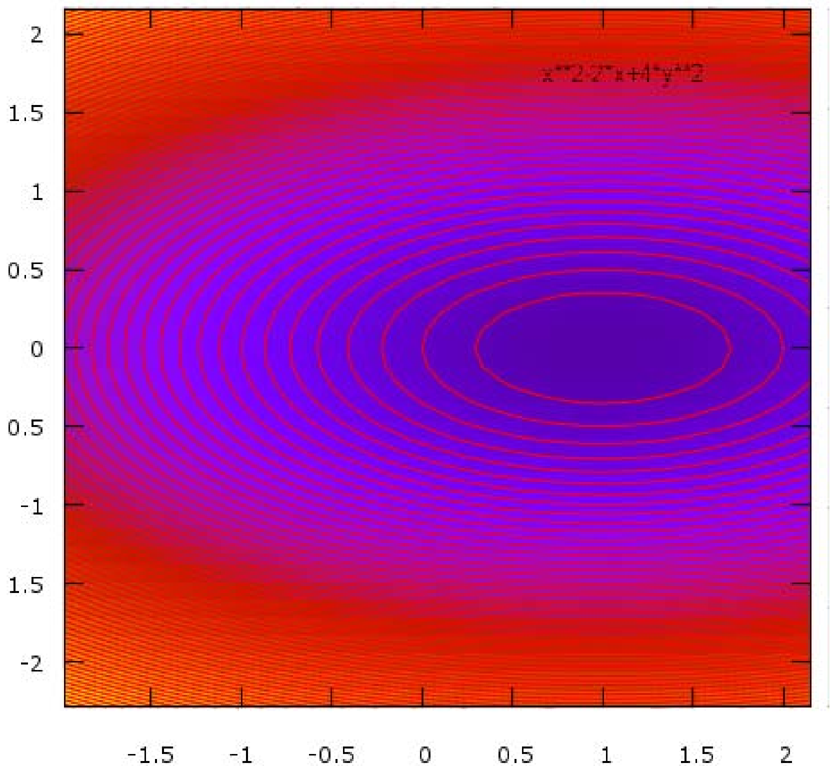
収束のスピードは遅い



教科書115ページの
表4．1参照

一次収束の例2

- 狭義凸2次関数 $f(x) = x_1^2 - 2x_1 + 4x_2^2$ に適用
- 最適解は(1, 0)
- $\beta =$ 約0.523 なので収束のスピードが速い



田村，村松:「最適化法」，共立出版，2002，p103
の表を参照

関数のヘッセ行列

- 非線形計画では，多変数関数の微分が重要な役割を果たす
- 定義：n変数関数 f の**ヘッセ行列** $\nabla^2 f(x)$

　　←→f の2次偏微分係数を要素とするｎｘｎ 行列

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

- ｆが2変数以上の関数の時は，幾何学的な情報を含む（次のスライド）
- ｆが1変数関数のときは，ヘッセ行列は2階微分と同じ

$$\nabla^2 f(x) = f''(x)$$

関数のヘッセ行列

- 非線形計画では，多変数関数の微分が重要な役割を果たす
- 定義：n変数関数 f のヘッセ行列 $\nabla^2 f(x)$

　　　←→f の2次偏微分係数を要素とするｎｘｎ 行列

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

- ｆ が2変数以上の関数の時は，幾何学的な情報を含む（次のスライド）
- ｆ が1変数関数のときは，ヘッセ行列は2階微分と同じ

$$\nabla^2 f(x) = f''(x)$$

関数のヘッセ行列：例

- $f(x_1, x_2) = 5x_1^2 - 6x_1x_2 + 5x_2^2 - 10x_1 + 6x_2$

$$\nabla f(x_1, x_2) = \begin{bmatrix} 10x_1 - 6x_2 - 10 \\ -6x_1 + 10x_2 + 6 \end{bmatrix}$$

$$\nabla^2 f(x_1, x_2) = \begin{bmatrix} 10 & -6 \\ -6 & 10 \end{bmatrix}$$

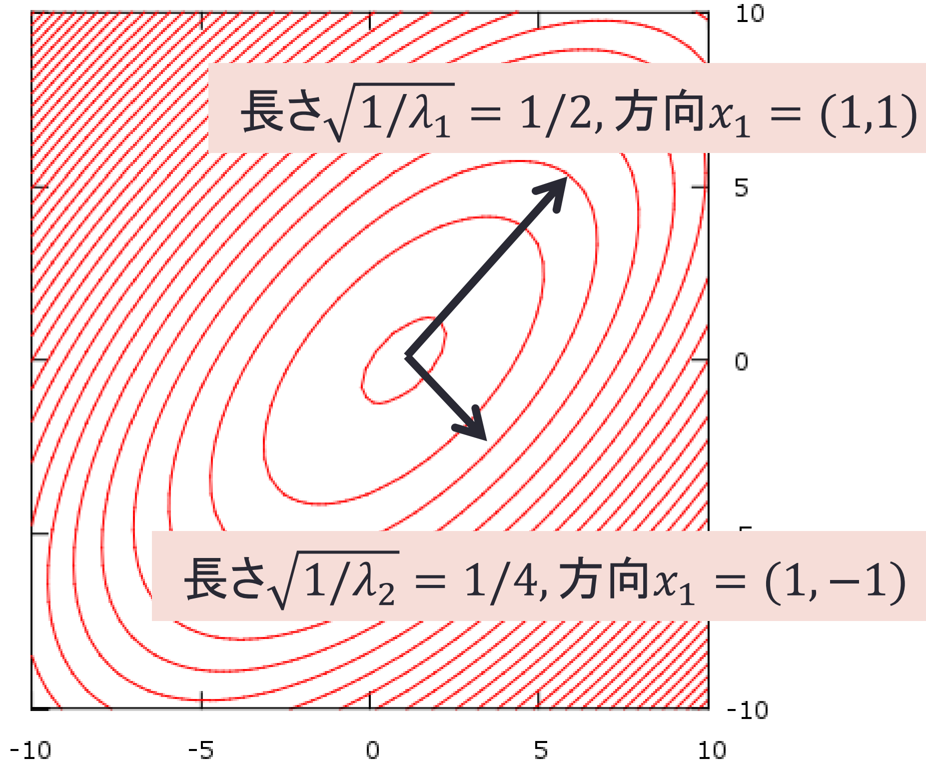
凸二次関数なので，等高線は楕円

ヘッセ行列の固有ベクトルと固有値

$\lambda_1 = 4, x_1 = (1,1)$

$\lambda_2 = 16, x_1 = (1,-1)$

これらの情報が楕円の形状を表す



2次関数とその凸性

- n 変数の2次関数 f は，
  実数 c，n次元ベクトル$b = (b_1, b_2, \dots, b_n)$，$n \times n$対称行列$A = (a_{ij})$
  を使って次の形に書ける

$$f(x) = c + b^T x + \frac{1}{2} x^T A x$$

$$= c + \sum_{i=1}^{n} b_i x_i + \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j$$

$\begin{bmatrix} 2 & 0 \\ 0 & 6 \end{bmatrix}, \begin{bmatrix} 2 & -1 \\ -1 & 4 \end{bmatrix}$
は半正定値

- 性質：2次関数は凸関数←→行列Aが半正定値行列

- 定義：$n \times n$行列Aは半正定値
  $\Leftrightarrow$ 任意のn次元ベクトルxに対し，$x^T A x \geq 0$

$\begin{bmatrix} 2 & 0 \\ 0 & -1 \end{bmatrix}, \begin{bmatrix} 2 & 2 \\ 2 & 1 \end{bmatrix}$は
半正定値ではない

テイラー展開を使った近似

- n変数関数 f，n次元ベクトル $\bar{x}$
- 定義：関数 f の $\bar{x}$ における2次のテイラー展開 $\tilde{f}$

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

  - 2次のテイラー展開は2次関数
    - ヘッセ行列が半正定値行列ならば凸2次関数
  - $\bar{x}$のまわりで$\tilde{f}$はfを近似している

    （xが$\bar{x}$に近いとき，$\tilde{f}(x)$と$f(x)$の誤差は小さい）

テイラー展開を使った近似の例

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

$$\nabla f(x) = 5x^4 - 15x^2 + 4 \qquad \nabla^2 f(x) = 20x^3 - 30x$$

a = -1 のとき

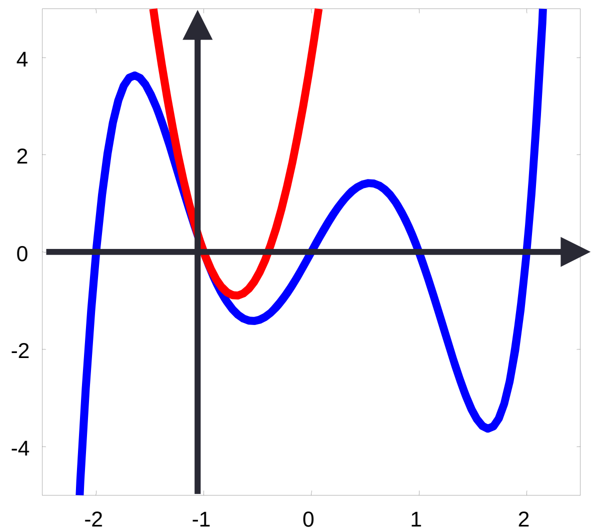
$$0 - 6(x+1) + 5(x+1)^2$$

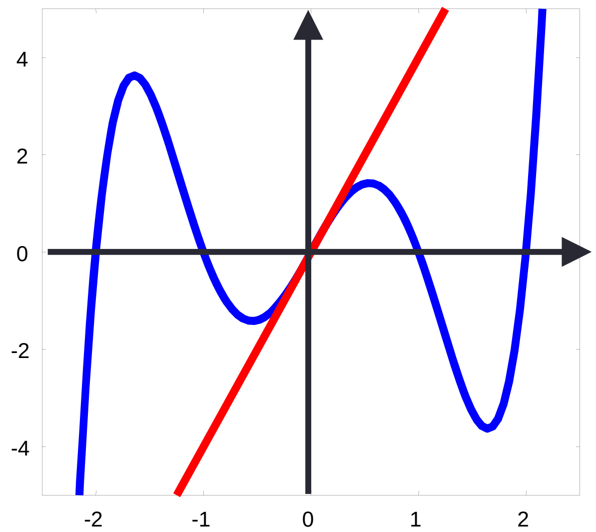
a = 0 のとき

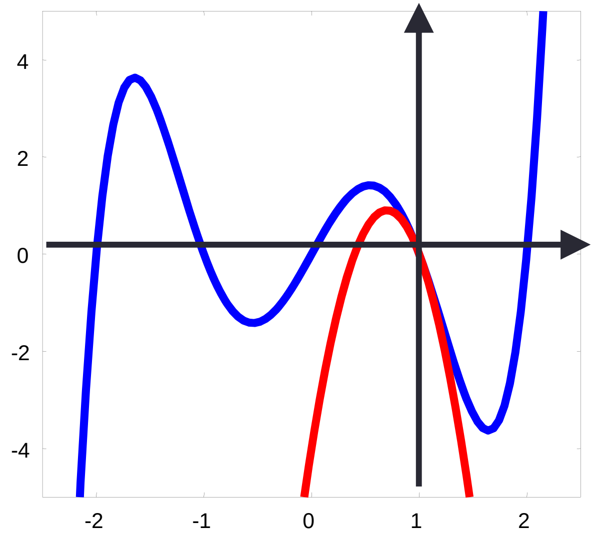
$$0 + 4x + 0x^2$$

a = 1 のとき

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







2次の最適性条件

- 非線形計画問題に対する**最適性条件**
～実行可能解が(局所的)最適解であるための
十分条件または必要条件

- **定理[最適解であるための2次の必要条件]:**
ベクトル x^* が関数 f に関する制約なし問題の局所的最適解
➔ $\nabla^2 f(x^*)$**は半正定値行列**

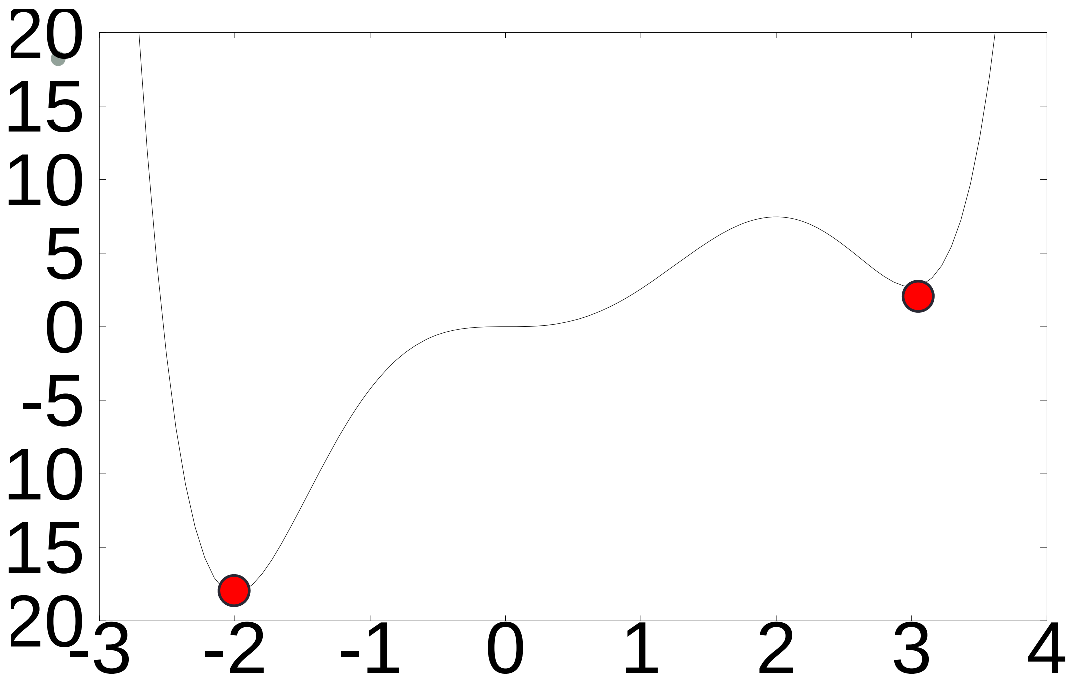
- **定義**: $n \times n$行列Aは**半正定値**
⇔ 任意のn次元ベクトルxに対し, $x^T A x \geq 0$

f が一変数関数の場合,
$\nabla^2 f(x^*)$は半正定値行列⇔$f''(x^*) \geq 0$が成り立つ
「極小点での傾きの増加率は非負」「下に凸」

2次の最適性条件：例1

- 目的関数：$f(x) = \frac{1}{6}x^6 - \frac{3}{5}x^5 - x^4 + 4x^3$ → 最小化
- 2階微分を計算：$f''(x) = 5x^4 - 12x^3 - 12x^2 + 24x$
- 局所的最適解は$x = -2, +3$

$$f''(-2) = 80 \geq 0, f''(3) = 45 \geq 0$$



2次の最適性条件

- 定理[最適解であるための2次の十分条件]:

ベクトル x^* が$\nabla f(x) = 0$を満たし，かつ$\nabla^2 f(x^*)$が正定値行列

➔x^* は関数 f に関する制約なし問題の局所的最適解

- 定義 : $n \times n$行列Aは正定値

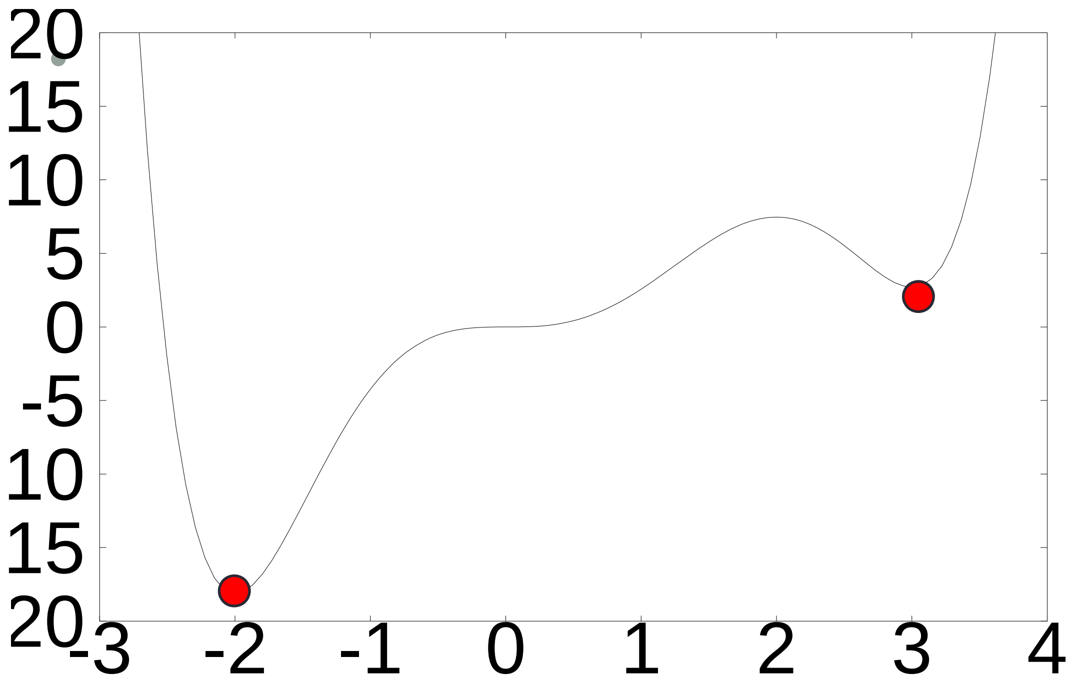
  $\Leftrightarrow$ 任意の非ゼロ n 次元ベクトルxに対し，$x^T Ax > 0$

f が一変数関数の場合，

$\nabla^2 f(x^*)$が正定値行列←→$f''(x^*) > 0$が成り立つ

2次の最適性条件：例1（続き）

- 目的関数：$f(x) = \frac{1}{6}x^6 - \frac{3}{5}x^5 - x^4 + 4x^3$ → 最小化
- 2階微分を計算：$f''(x) = 5x^4 - 12x^3 - 12x^2 + 24x$
- $x\ =\ -2, +3$ は$\nabla f(x) = 0$を満たす
- $f''(-2) = 80 > 0, f''(3) = 45 > 0$➔ともに局所的最適解

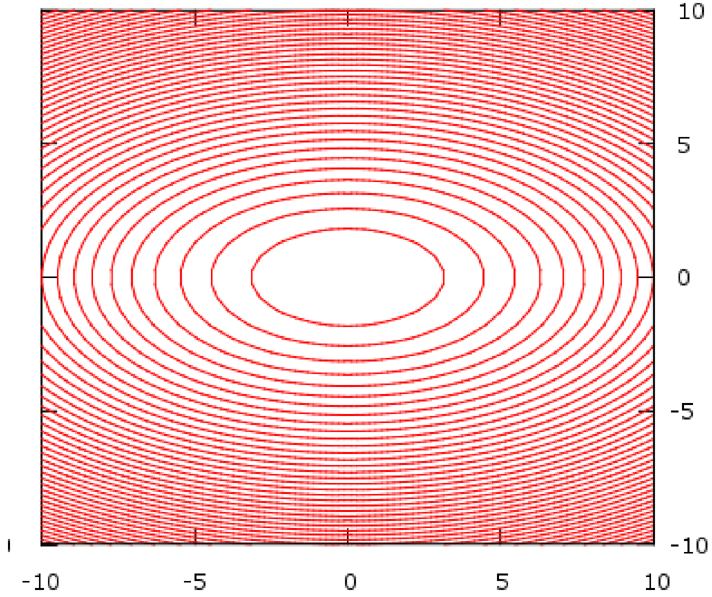


2次の最適性条件：例2

- 目的関数：$f(x_1, x_2) = x_1^2 + 3x_2^2$ → 最小化
- $\nabla f(x_1, x_2) = \begin{bmatrix} 2x_1 \\ 6x_2 \end{bmatrix}, \nabla^2 f(x_1, x_2) = \begin{bmatrix} 2 & 0 \\ 0 & 6 \end{bmatrix}$
- $\nabla f(x_1, x_2)$がゼロベクトルとなるのは (0,0) のみ
- $\nabla^2 f(0,0) = \begin{bmatrix} 2 & 0 \\ 0 & 6 \end{bmatrix}$ は正定値行列➔(0, 0) は局所的最適解

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:例3

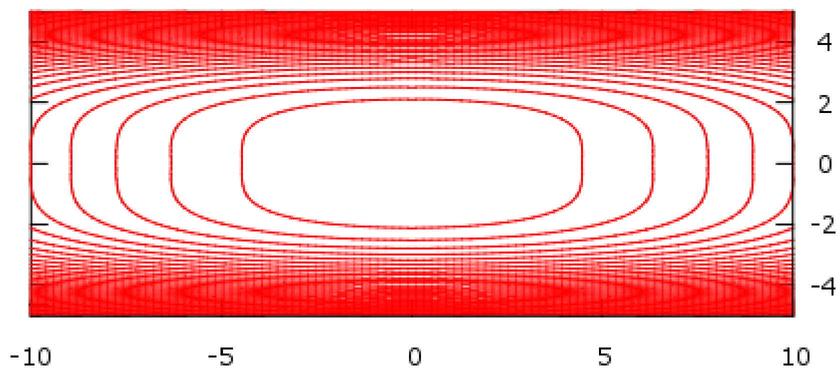
- 目的関数: $f(x_1, x_2) = x_1^2 + x_2^4$ → 最小化
- $\nabla f(x_1, x_2) = \begin{bmatrix} 2x_1 \\ 4x_2^3 \end{bmatrix}, \nabla^2 f(x_1, x_2) = \begin{bmatrix} 2 & 0 \\ 0 & 12x_2^2 \end{bmatrix}$
- $\nabla f(x_1, x_2)$がゼロベクトルとなるのは (0,0) のみ(実は最適解)
- $\nabla^2 f(0,0) = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}$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

➔(0, 0) が局所的最適解かどうかは,ヘッセ行列を使って判定できない(実際には局所的最適解)

任意のベクトル(y_1, y_2)に対して

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

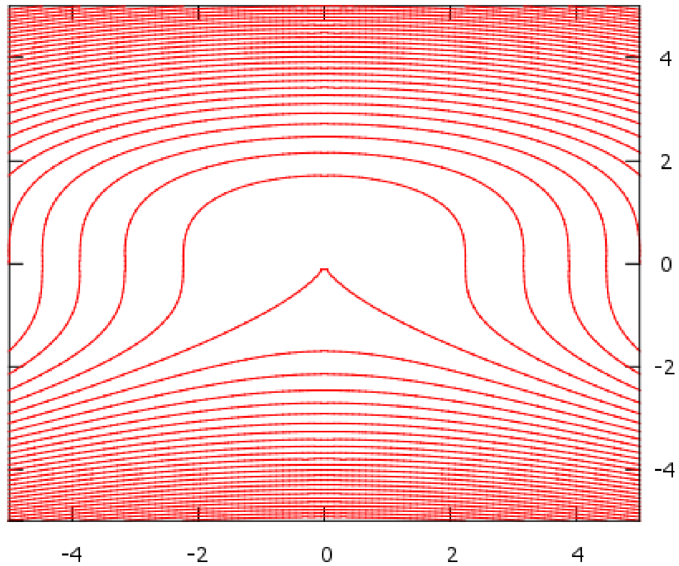
$y_1 = 0$のときは$y_2 \neq 0$でも値は0



2次の最適性条件：例4

- 目的関数：$f(x_1, x_2) = x_1^2 + x_2^3$ → 最小化
- $\nabla f(x_1, x_2) = \begin{bmatrix} 2x_1 \\ 3x_2^2 \end{bmatrix}, \nabla^2 f(x_1, x_2) = \begin{bmatrix} 2 & 0 \\ 0 & 6x_2 \end{bmatrix}$
- $\nabla f(x_1, x_2)$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) が局所的最適解かどうかは，
ヘッセ行列を使って判定できない
（実際には局所的最適解ではない
任意の$\varepsilon > 0$に対して，
$$f(0, -\varepsilon) = -\varepsilon^3 < 0)$$



レポート問題（締切：1月26日，今回で最後）

問1：等高線で表された2つの関数に対して，最急降下法を実行して得られる点列を求めよ．計算は全て，添付の図を使って行うこと．出発点は左上に図示された点とし，3回～6回程度の反復を行うこと．（注意：勾配ベクトルの方向は等高線と必ず垂直．等高線のないところは自分で補完してください）

問2：関数 f, g に対し，$x = (1,2)$ における2次のテイラー展開を計算せよ．

$$f(x_1, x_2) = x_1 \log x_2 - x_2 \log x_1 \qquad g(x_1, x_2) = x_1^2 + x_2^2 - 1$$

問3：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，局所的最適解を求めよ．