数理計画法 第11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

非線形計画問題

- 非線形計画問題は，下記のように表される問題
  - 目的関数：　$f(x_1, x_2, \ldots, x_n)$ → 最小（または最大）
  - 制約条件：　$x \in S$
    - $f(x_1, x_2, \ldots, x_n)$は変数$x_1, \ldots, x_n$に関する非線形関数（目的関数）
    - S はベクトル$(x_1, x_2, \ldots, x_n)$の集合（実行可能集合）
      - Sは線形または非線形な等式・不等式で与えられることが多い

- 制約条件のない問題（制約なし問題）を考えることも多い
  - 目的関数：　$f(x_1, x_2, \ldots, x_n)$ → 最小（または最大）

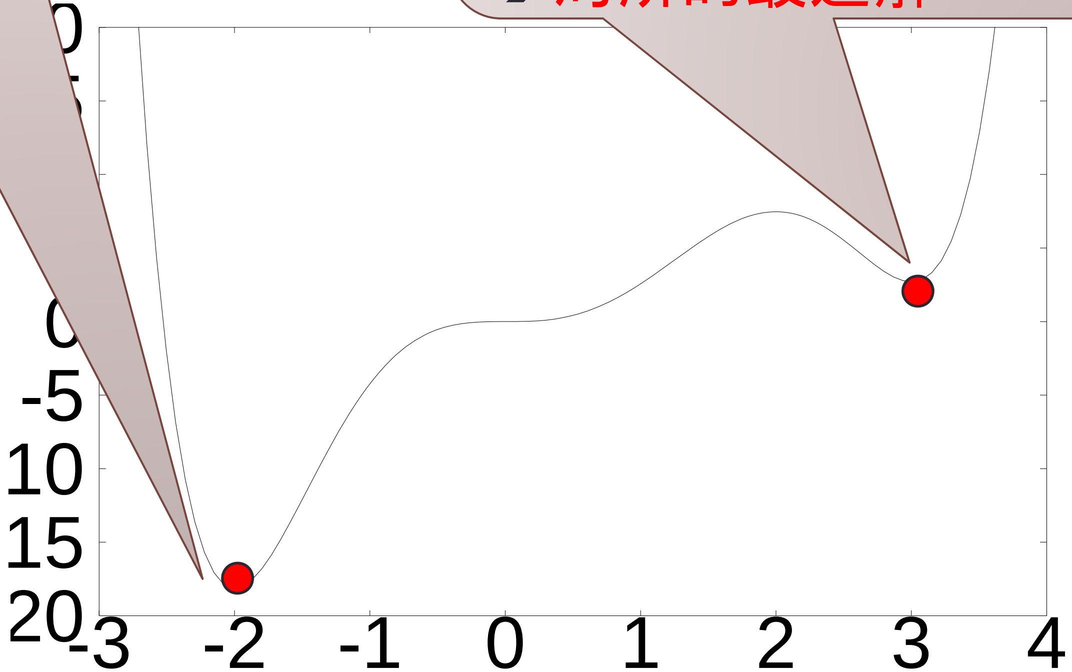
- この授業では，主に制約なし問題を扱う

制約なし問題の例

- 目的関数 : $f(x) = \frac{1}{6}x^6 - \frac{3}{5}x^5 - x^4 + 4x^3$ → 最小化

全ての実行可能解の中でf(x)が最小
➔大域的最適解

実行可能解 x=-3 の周り（近傍）では, いずれの実行可能解の目的関数値も f(-3) 以上
➔局所的最適解



大域的最適解と局所的最適解

- 定義：実行可能解 x は大域的最適解

⇔全ての実行可能解の中で目的関数値が最小

  - 普通の意味での最適解と同じ意味

- 定義：実行可能解 x は局所的最適解（局所的最小点）

⇔ x の周り（近傍）の全て実行可能解の

目的関数値は f(x) 以上

[厳密な定義] ある$\varepsilon > 0$が存在して，$\|y - x\|_2 \leq \varepsilon$を満たす任意の$y$に対して$f(x) \leq f(y)$が成り立つ

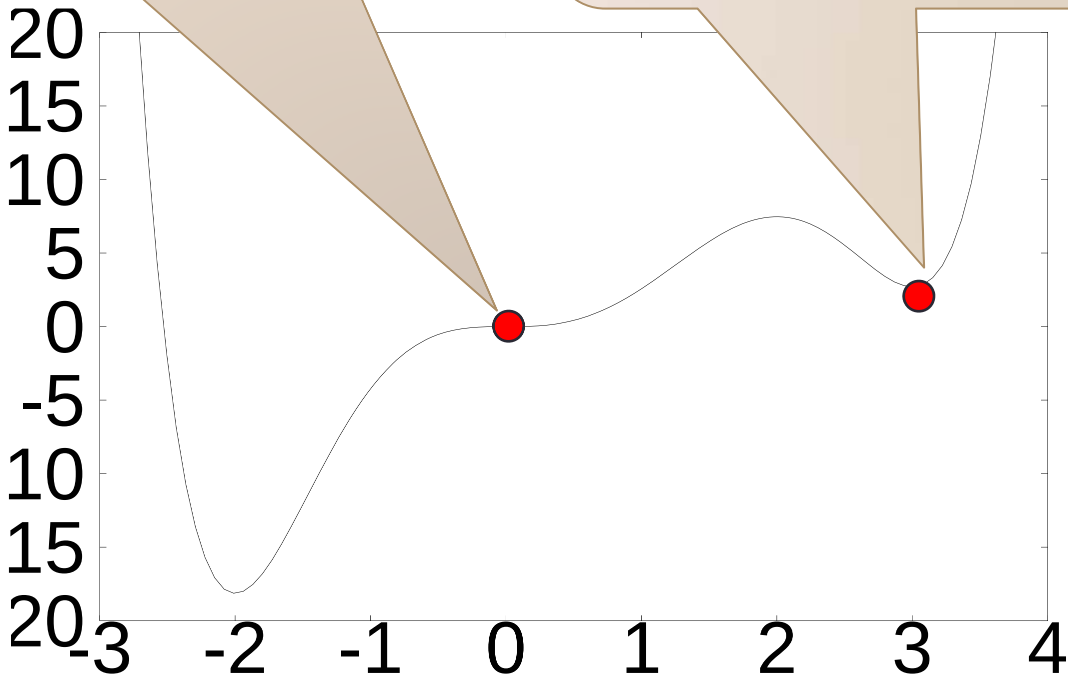
- 大域的最適解は局所的最適解のひとつである
- 局所的最適解は多数存在することもある

制約なし問題の例（再掲）

- 目的関数：$f(x) = \frac{1}{6}x^6 - \frac{3}{5}x^5 - x^4 + 4x^3$ → 最小化

任意の$\varepsilon > 0$に対して
$f(0) > f(0-\varepsilon)$成立
➔x=0 は局所的最適解ではない

$\varepsilon = 0.5$とおくと,
$|y-3| \leq \varepsilon$を満たす全ての
yに対して$f(3) \leq f(y)$成立
➔x=3 は局所的最適解

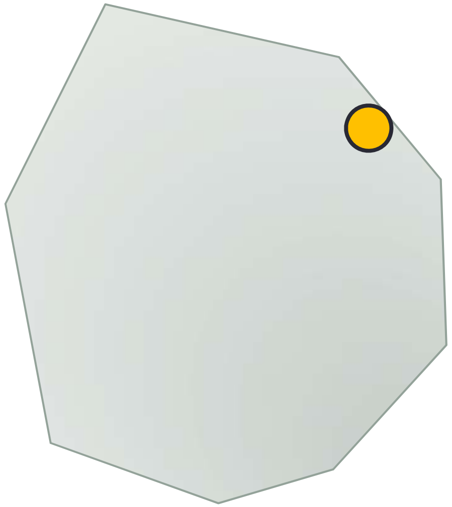


非線形計画問題の解きやすさ，難しさ（その1）

- 2次元の集合 S の中で，一番下にある点を求めたい
  ➔ボールを転がして求めよう

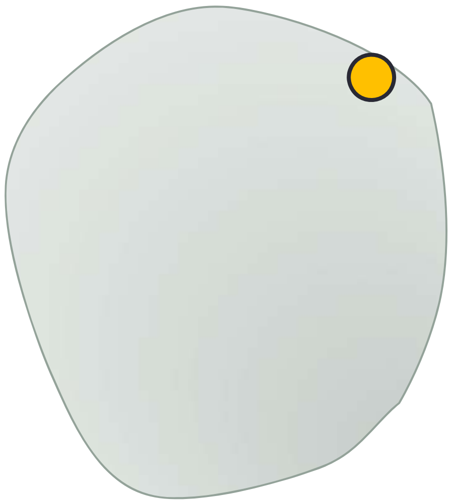
例1：Sが多面体

- ボールは一番下の点で必ず止まる
- 角で止まる



例2：Sがへこみのない集合

- ボールは一番下の点で必ず止まる

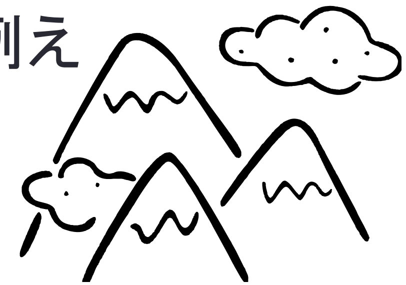


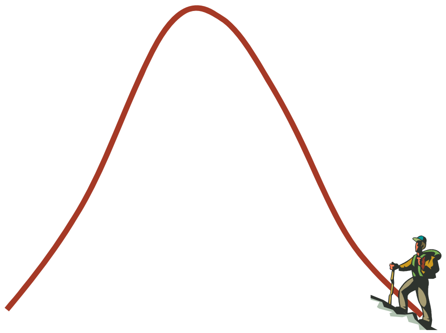
例3：Sがへこみのある集合

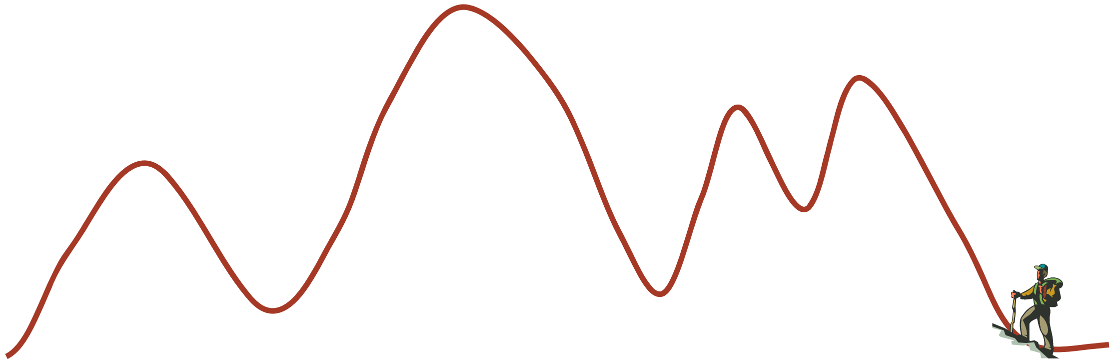
- ボールは一番下の点で必ず止まるとは限らない

非線形計画問題の解きやすさ，難しさ（その2）

- 非線形関数の制約なし最大化を考える➔山登りの例え
  - 最適解（最大解）を求める
    ＝ 山脈の最も高いところを見つける
  - 局所最適解 ＝ 山の頂（いただき）







- 山の頂がただ一つ
  ➔どんどん上に登ると，必ず最も高いところに辿り着く
- 山の頂が複数ある
  ➔どんどん上に登っても，最も高いところに辿り着く保証はない

非線形計画問題の解きやすさ，難しさ（その3）

- 非線形計画問題
  - 目的関数： $f(x_1, x_2, \ldots, x_n)$ → 最小
  - 制約条件：　$x \in S$

- 集合 S にへこみがない➔解きやすい
  へこみがある➔難しい

- 関数 f に局所的最適解がただ一つ➔解きやすい
  局所最適解がたくさん存在➔難しい

凸集合と凸関数（その1）

- Sが凸集合➔へこみがない（穴もない）
- f が凸関数➔局所的最適解はただ一つ

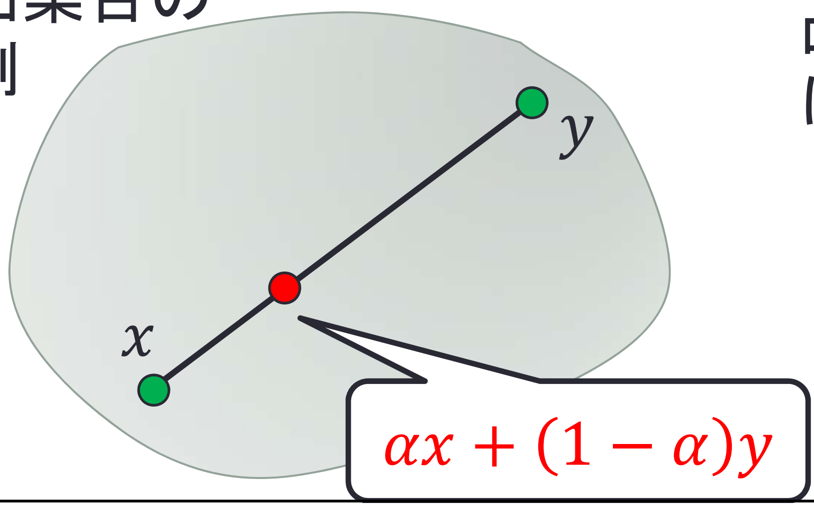
非線形計画問題は解きやすい

- 定義：n次元ベクトルの集合 S は凸集合

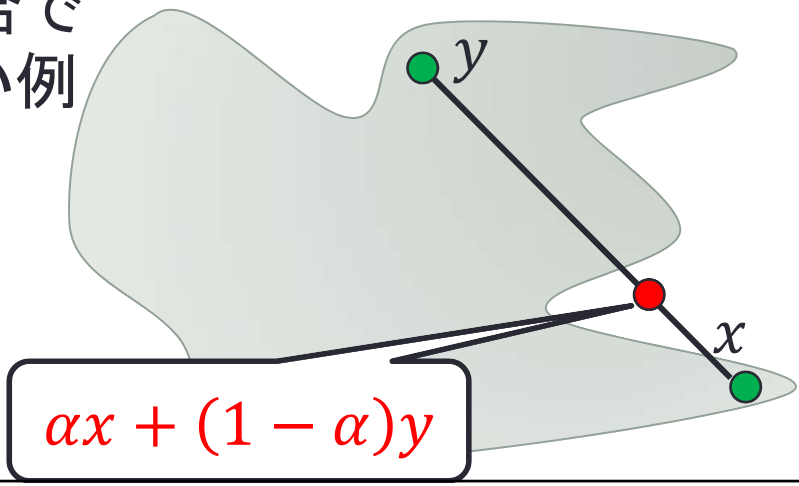
  ⇔任意の $x, y \in S$ と α $(0 \le \alpha \le 1)$ に対して

  $\alpha x + (1-\alpha)y \in S$ が成り立つ

凸集合の例



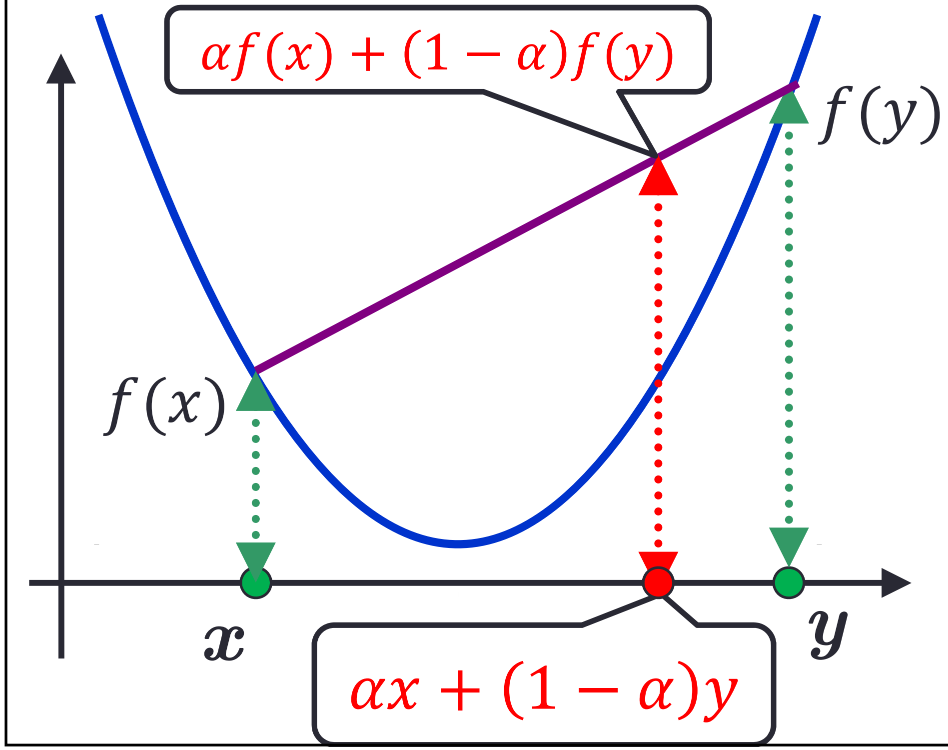
凸集合ではない例



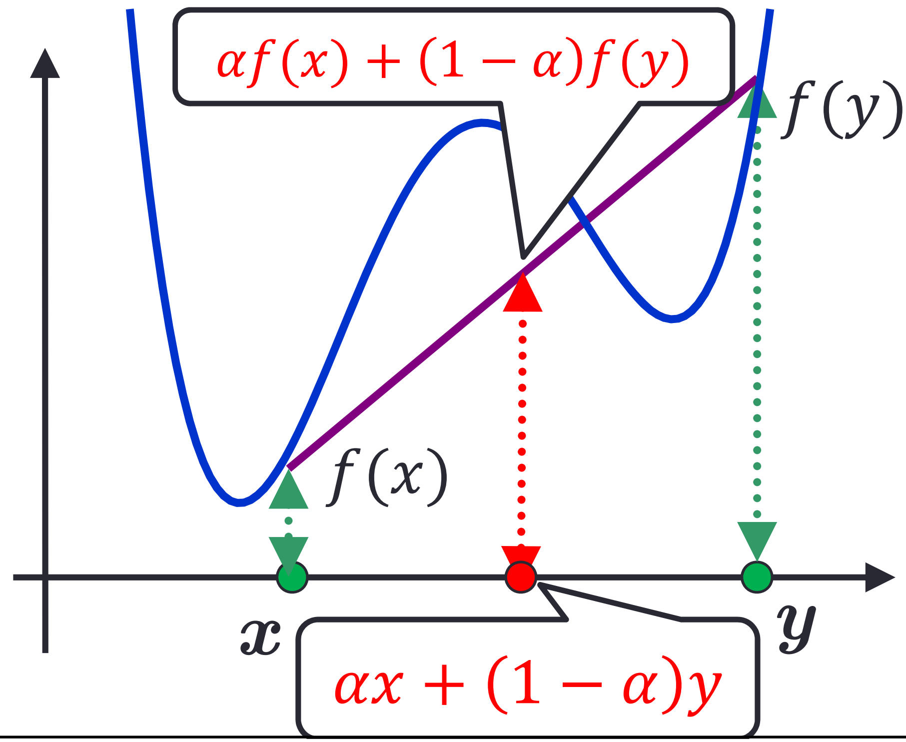
凸集合と凸関数(その2)

- 定義: 変数$x_1, \dots, x_n$に関する関数$f(x_1, x_2, \dots, x_n)$は凸関数
  ⟺任意のベクトル x, yとα $(0 \le \alpha \le 1)$に対して
  $f(\alpha x + (1-\alpha)y) \le \alpha f(x) + (1-\alpha)f(y)$が成り立つ

凸関数の例



凸関数ではない例



凸関数と局所的最適解の関係

- 定理：f が凸関数ならば，（f に関する制約なし問題の）
  任意の局所的最適解は大域的最適解である

証明：背理法により証明する．

- x^*は局所的最適解，大域的最適解ではないと仮定

➔$f(y^*) < f(x^*)$を満たすベクトルy^*が存在（$y^* \neq x^*$である）

- x^*は局所的最適解

➔ある$\varepsilon > 0$が存在して，$\|y - x^*\|_2 \leq \varepsilon$を満たす

任意の y に対して $f(x^*) \leq f(y)$成立

fは凸関数➔$0 \leq \alpha \leq 1$を満たす任意のαに対して

$f(\alpha x^* + (1-\alpha)y^*) \leq \alpha f(x^*) + (1-\alpha)f(y^*)$成立

とくに，$0 < \alpha < 1$ならば $\alpha f(x^*) + (1-\alpha)f(y^*) < f(x^*)$

しかし，αを1に十分近い値とすると，$\|(\alpha x^* + (1-\alpha)y^*) - x^*\|_2 \leq \varepsilon$ 成立

➔ $f(x^*) \leq f(\alpha x^* + (1-\alpha)y^*)$成立（矛盾）

関数の勾配

- 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 の勾配（ベクトル）$\nabla f(x)$

  ⇔各変数で f を偏微分したものを並べたベクトル

$$\nabla f(x_1, x_2, \dots, x_n) = \begin{bmatrix} \dfrac{\partial f(x)}{\partial x_1} \\ \dfrac{\partial f(x)}{\partial x_2} \\ \vdots \\ \dfrac{\partial f(x)}{\partial x_n} \end{bmatrix}$$

- 勾配ベクトルの方向
  - その点において関数値の増加率（関数の傾き）が最大となる方向
  - 等高線とは垂直の関係

関数の勾配：例

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

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

$a = (0,0)$のとき

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

$b = (2,0)$のとき

$$\nabla f(b) = \begin{bmatrix} +10 \\ +6 \end{bmatrix}$$

$c = (3,1)$のとき

$$\nabla f(c) = \begin{bmatrix} +14 \\ -2 \end{bmatrix}$$

教科書106ページの
図4．3を参照のこと

制約なし問題の最適性条件

- 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

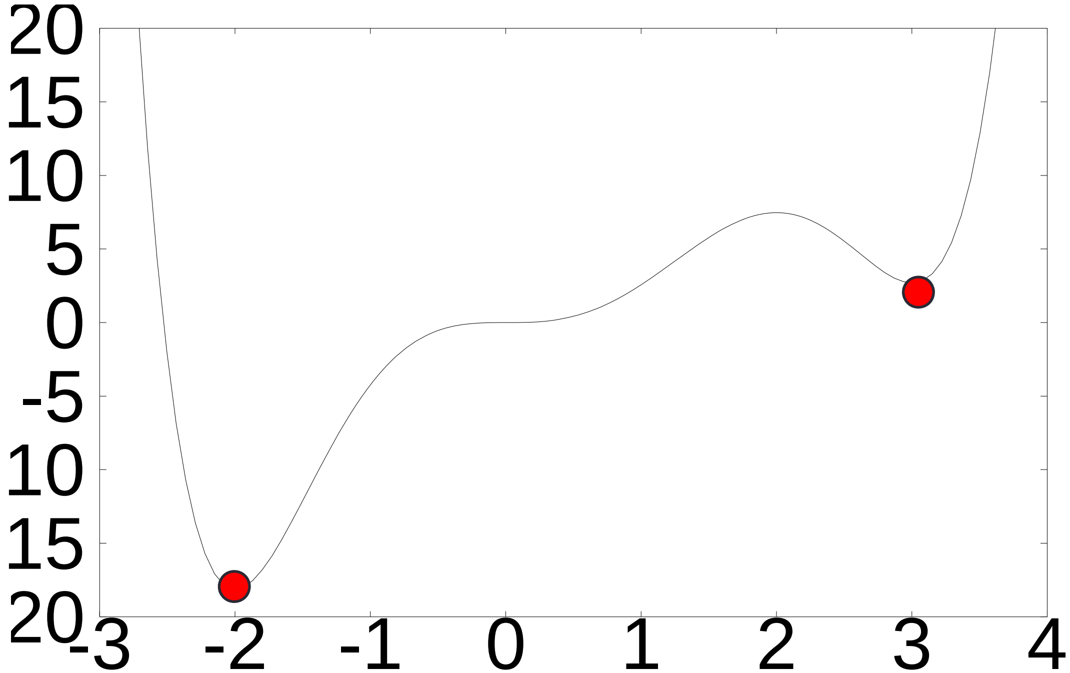
- 定理[最適解であるための必要条件]:
ベクトル x^* が関数 f に関する制約なし問題の局所的最適解
➔ $\nabla f(x^*) = 0$が成り立つ

とくに，f が一変数関数の場合，$f'(x^*) = 0$が成り立つ

「極小点での傾きは0」

制約なし問題の最適性条件：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
- 勾配を計算：$f'(x) = (x+2)x^2(x-2)(x-3)$
- 局所的最適解はx = −2, +3, $f'(-2)$=$f'(+3)$=0



注意：x= 0, +2 に対しても $f'(x)$=0 が成り立つが，x は局所的最適解ではない

制約なし問題の最適性条件：例2

- $f(x_1, x_2) = 5x_1^2 - 6x_1x_2 + 5x_2^2 - 10x_1 + 6x_2$
  の局所的最適解を求めたい（存在しない場合もありうる）
- 「局所的最適解➔$\nabla f(x) = 0$」なので，次のやり方で計算できる
  1. $\nabla f(x) = 0$の解を全て求める.
  2. 各々の解が局所的最適解か否かをチェックする.

$\nabla f(x_1, x_2) = \begin{bmatrix} 10x_1 - 6x_2 - 10 \\ -6x_1 + 10x_2 + 6 \end{bmatrix}$ ➔ 解は(1,0)のみ

さらに，任意の(y_1, y_2)に対して$f(1,0) \leq f(y_1, y_2)$ が証明できる

➔(1,0) は局所的最適解（とくに大域的最適解）

$$f(x_1, x_2) = 5\left(x_1 - \frac{3x_2+5}{5}\right)^2 + \frac{16}{5}x_2^2 - 5$$

制約なし問題の最適性条件：証明

「局所的最適解➔ $\nabla f(x^*) = 0$」を証明する

証明で重要な性質：ノルム$\|d\|$が十分に小さい場合，

$$f(x+d) \simeq f(x) + \nabla f(x)^T d \text{ が成立}$$

証明（概略）：背理法を使う

$\nabla f(x^*) \neq 0$のとき，

$d = -\delta \nabla f(x^*)$ (δは十分に小さい正の数)とおく

➔ $f(x^* + d) \simeq f(x^*) + \nabla f(x^*)^T d$

$$= f(x^*) - \delta \|\nabla f(x^*)\|^2 < f(x^*)$$

x^* が局所的最適解であることに矛盾

制約なし問題の最適性条件：補足

- 「$\nabla f(x^*) = 0$」は勾配（一次微分）を用いた必要条件

➔局所的最適解であるための「一次の必要条件」と呼ばれる

- 「局所的最適解➔ $\nabla f(x^*) = 0$」の逆は必ずしも成り立たない
  - $\nabla f(x) = 0$を満たす x は停留点と呼ばれる
  - f が凸関数のとき，逆は成り立つ

- 定理：凸関数fに対し，ベクトルx^*が$\nabla f(x^*) = 0$を満たす

  ➔ x^*は大域的最適解

証明（概略）：f は凸関数➔次の性質が成り立つ

「任意のx, yに対して$f(x) - f(y) \geq \nabla f(y)^T(x - y)$」

$y = x^*$を代入すると$f(x) - f(x^*) \geq 0$ ➔ x^* は大域的最適解

レポート問題

問1：関数$f(x) = |x|$ ($|x|$は値xの絶対値)が凸関数であることを，凸関数の定義に基づいて証明せよ．

問2：次の関数が凸関数でないことを，凸関数の定義に基づいて証明せよ．

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

問3：下記の関数の勾配ベクトルを計算せよ．

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

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

問4：下記の関数の停留点，および局所的最適解を全て求めよ．

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