最適化数学　講義ノート 1　(担当: 関口 良行)

1　最適化するとは？

以下, 問題を取り上げるときは, 最小化問題を扱うが最大化問題も同様なことが言える.

J を 数ベクトル空間 $\mathbb{R}^n$ 上の関数, C をその部分集合とする.

最小化問題とは以下のような問題 (P) を指す:

$$\begin{array}{lll} (P) & \text{最小化} & J(x) \\ & \text{制約} & x \in C \end{array}$$

ここで, 「$x \in C$」とは「x が C に含まれる」ことを意味する. また, 最小化する $J(x)$ を**目的関数**と呼ぶ.

例 1.

$$\begin{array}{ll} \text{最小化} & J(x) = x^2 + ax + b \\ \text{制約} & x \in \mathbb{R} \end{array}$$

最小値を取る点を見つけたいので, 式変形をすると

$$J(x) = x^2 + ax + b = \left(x + \frac{a}{2}\right)^2 - \frac{a^2}{4} + b \geq -\frac{a^2}{4} + b$$

となるので, 最小値は $-a^2/4 + b$ で最小解は $x = -a/2$ と分かる.

ここで, 最小解というものを改めて定義しておく.

定義. $\bar{x} \in C$ が任意の $x \in C$ に対して

$$J(x) \geq J(\bar{x})$$

のとき $J(\bar{x})$ を (P) の**大域最小値**, $\bar{x}$ を**大域最小解**と呼ぶ.

<u>大域最小解を見つけられれば最も良いが, それは一般的にに難しい.</u> そこで, より弱い解を探すことを目標とする.

定義. $\bar{x} \in C$ が $\bar{x}$ に**充分近い**任意の $x \in C$ に対して

$$J(x) \geq J(\bar{x})$$

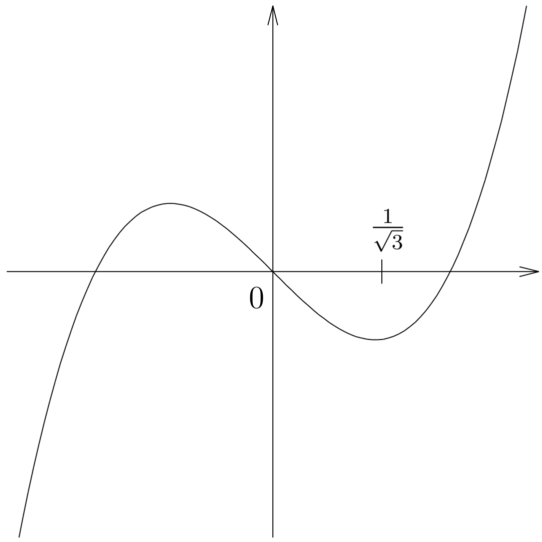
のとき $J(\bar{x})$ を (P) の**局所最小値**, $\bar{x}$ を**局所最小解**と呼ぶ.

例 2. 例 1 の場合は $x = -a/2$ は大域最小解となる.

一方, 次の最小化問題

$$\begin{aligned} \text{最小化} \quad & J(x) = x^3 - x \\ \text{制約} \quad & x \in \mathbb{R} \end{aligned}$$

については <u>大域最小解は存在しない</u>が, 右のグラフから分かるように $x = 1/\sqrt{3}$ に充分近い x ではグラフは下に凸になっているので, そこは局所最小解になっている.



2 制約無しの最適化問題

一般に最小化問題は制約があるものが多いが, まずは制約の無い最小化問題のから始める:

$$\text{最小化} \quad J(x)$$

制約無しの最小化問題の解析はこれからの理論の基礎になる.

また制約の無い最小化でも重要な応用例は多くある. 例えば, 観測データを一次関数で近似する最小二乗法は制約無しの最小化問題である.

2.1 最適性から分かること

簡単のため 2 変数の最小化問題を説明する.

$$\text{最小化} \quad J(x, y)$$

記号を少し用意する. $\|(x, y)\| = \sqrt{x^2 + y^2}$ とおく. また偏微分 J_x, J_y を用いて,

$$\nabla J(x, y) = (J_x(x, y), J_y(x, y))$$

とおき, このベクトルを**勾配ベクトル**と呼ぶ.

定理 1. 一次の最適性必要条件 $(\bar{x}, \bar{y})$ が局所最小解ならば,

$$\nabla J(\bar{x}, \bar{y}) = (0, 0)$$

が成り立つ.

証明. 局所最小解の定義より, $(\bar{x},\bar{y})$ に充分近い (x,y) に対して

$$J(x,y) \geq J(\bar{x},\bar{y})$$

が成り立つ. $\|(u,v)\| = 1$ となるような任意の (u,v) に対して,

$$h(t) = J((\bar{x},\bar{y}) + t(u,v))$$

とおく. 充分小さい $t > 0$ に対して, $(\bar{x},\bar{y}) + t(u,v)$ は $(\bar{x},\bar{y})$ に近くなるので,

$$h(t) \geq h(0)$$

が成り立つ. ここで, h を 0 でテーラー展開すると,

$$h(t) = h(0) + h'(0)t + o(t)$$

となるので, 上の不等式より $h'(0)t + o(t) \geq 0$ となり, 両辺を t で割ると

$$h'(0) + \frac{o(t)}{t} \geq 0$$

を得る. そこで, t を 0 に近づけると $o(t)/t$ も 0 に近づくので, $h'(0) \geq 0$ となる. 必要ならば t をさらに小さくすれば $h(-t) \geq h(0)$ も成り立つので, $h'(0) \leq 0$ も同様に得る. よって $h'(0) = 0$ が成り立つ.

今, $h'(t) = \frac{d}{dt}J(\bar{x}+tu, \bar{y}+tv) = J_x(\bar{x}+tu, \bar{y}+tv)u + J_x(\bar{x}+tu, \bar{y}+tv)v$ より,

$$h'(0) = J_x(\bar{x},\bar{y})u + J_y(\bar{x},\bar{y})v = 0$$

を得る. この等式は任意の $\|(u,v)\| = 1$ を満たす (u,v) に対して成り立つので, $\nabla J(\bar{x},\bar{y}) = (0,0)$ が成り立つ. □

定義. $\nabla J(\bar{x},\bar{y}) = (0,0)$ の時, $(\bar{x},\bar{y})$ を J の**停留点**と呼ぶ.

練習問題 1. 次の関数の停留点を求めよ.

$$(1)\ f(x,y) = x^2 - xy + 2y^2 - x - 2y \quad (2)\ f(x,y) = x^3 + y^3 - 3xy$$

2.2 停留点が局所最適解か?

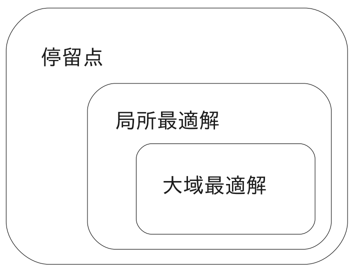
停留点は局所最適解とは限らない.

例 3.

$$\text{最小化 } J(x,y) = x^2 - y^2$$

では $\nabla J(0,0) = (0,0)$ となるが, 局所最適解ではない.

しかし, 局所最適解ならば常に停留点になるので最適解の候補となる点を得ることができる.



もちろん最適解が存在しない場合もある.

$(\bar{x}, \bar{y})$ を停留点とする．このとき, $(\bar{x}, \bar{y})$ を少し動かして (x', y') にしたとき,

(1). すべての動かし方で, $J(x', y') \geq J(\bar{x}, \bar{y})$ ならば, 局所最小解である

(2). すべての動かし方で, $J(x', y') \leq J(\bar{x}, \bar{y})$ ならば, 局所最大解である

(3). 動かし方により, $J(x', y')$ が $J(\bar{x}, \bar{y})$ よりも大きくなったり, 小さくなったりする場合はどちらでもない.

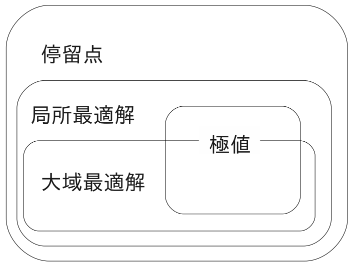
例 3 の $J(x, y) = x^2 - y^2$ の場合, 停留点 $(0, 0)$ から x だけ動かすと, J の値が大きくなり, y だけ動かすと J の値が小さくなるので, (3) の場合に当てはまる.

停留点の中から局所最適解を探す際に, 地道に $J(x', y')$ の値を調べなければならないこともあるが, 2 階の導関数を調べることで, 局所最適解を判定できたり, その候補をさらに絞り込むことができる.

注意

微分積分でつぎのようなことを学ぶ.

$\bar{x}$ に充分近い $x \neq \bar{x}$ に対して, $J(x) > J(\bar{x})$ が成り立つとき, J は $\bar{x}$ で**極小値**をとると言う．また, $J(x) < J(\bar{x})$ が成り立つとき, J は $\bar{x}$ で**極大値**をとると言う．二つを総称して, **極値**と呼ぶ.



極小値は不等号が “>” となるので, 局所最小値でさらに $\bar{x}$ 以外で等号の成り立たないものが, 極小値になる．大域最小値でも同様である.

局所最小値と極小値は別物であるが, それらを求める手法は結果的にあまり差が無いので, 講義では局所最小値を扱う.

例 4 (極小値にならない最小解)**.** $J(x, y) = 4x^2 + 4xy + y^2$ を最小化する問題を考える．$\nabla J(x, y) = (8x + 4y, 4x + 2y)$ なので, 停留点は $2x + y = 0$ を満たす．これより, 停留点はパラメータ t を用いて $x = t$ とすると, $(t, -2t)$ と書ける．この点上で J の値は $J(t, -2t) = 4t^2 - 8t^2 + 4t^2 = 0$ となる．よって, どんなに $(0, 0)$ の近くを考えても, そこに $J(0, 0)$ と同じ値をとる点があるので極小値にはならない．一方で, $J(x, y) = (2x + y)^2$ なので,

$$J(x, y) \geq 0 = J(t, -2t)$$

が成り立つ．よって, $(t, -2t)$ は大域最小解になる.