

最適化数学 第5回

[今回の項目]

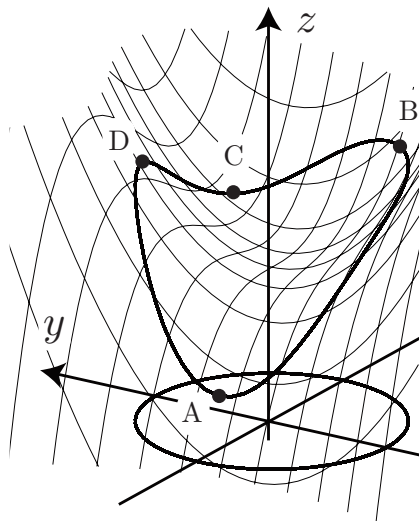
- ① 制約つき最適化問題
- ② 最適解の種類
- ③ 曲線上の増減表
- ④ ラグランジュ乗数法
- ⑤ 例題

制約つき最適化問題

2変数関数を
円周上で最小化する問題を考える：

$$\begin{aligned} & \text{最小化又は最大化 } x^3 - xy + y^2 + 2 \\ & \text{制約 } x^2 + y^2 - 1 = 0 \end{aligned}$$

図より、最適解は点 A である。
それでは、このような図がないとき、
計算によって最適解を見つけるにはどうしたらよいだろうか？

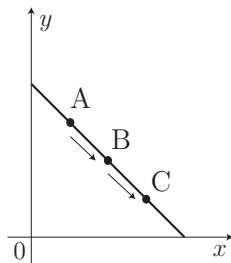


周長が一定のとき，面積最小の長方形は？

最大化 $f(x, y) := xy$

制約 $x + y = 4, x \geq 0, y \geq 0$

(x, y)	A		B		C
$f(x, y)$	3	↗	4	↘	3



ここで $x + y = 4$ より，

$$xy = x(4 - x) = -x^2 + 4x = -(x - 2)^2 + 4$$

と変形できる．点が $A: (1, 3) \rightarrow B: (2, 2) \rightarrow C: (3, 1)$ と動くとき， $f(x, y) = xy$ の値は，表のように変化する．

よって $x + y = 4$ を満たし，B に近い点の中で， $f(x, y)$ の値が一番大きくなるのは B においてである．したがって，一辺が 2 の正方形のとき面積が最大になる．

制約つき最適化問題の定義

集合 C を数ベクトル空間 \mathbb{R}^n の部分集合として,

$$\text{最小化 } f(x)$$

$$\text{制約 } x \in C$$

を **制約つき最小化問題** と呼ぶ. ここで, 集合 C を **実行可能領域**, C の点を **実行可能解** と呼ぶ. ただし, 「 $x \in C$ 」とは「 x が C に含まれる」ことを意味する.

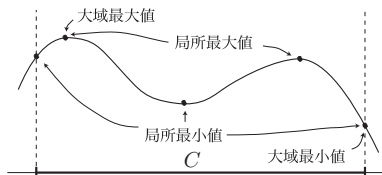
C は, 例えば上述の制約

$$x + y = 4, x \geq 0, y \geq 0$$

のように式を用いて表される. このような式を **制約式** と呼ぶ. 当然, 制約つき最適化問題においても最小化する関数 $f(x)$ は **目的関数** と呼ばれる.

最適解の種類

$$\begin{aligned} \text{(P) 最小化} \quad & f(x) \\ \text{制約} \quad & x \in C \end{aligned}$$



[定義]

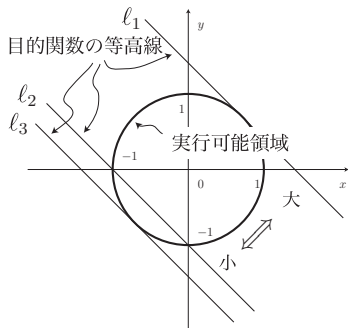
$\bar{x} \in C$ が \bar{x} に 十分近い C 上のすべての x に対して $f(x) \geq f(\bar{x})$ のとき, \bar{x} を 局所最小解, $f(\bar{x})$ を (P) の 局所最小値 と呼ぶ.

[定義]

$\bar{x} \in C$ が C 上のすべての x に対して $f(x) \leq f(\bar{x})$ のとき, \bar{x} を 大域最小解, $f(\bar{x})$ を (P) の 大域最小値 と呼ぶ.

最小解と最大解を合わせて最適解と呼ぶ. 最適値も同様である.

1 次関数を円周上で最適化



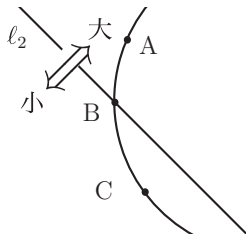
始めに、円周上で 1 次関数を最小化または最大化する問題を考える;

$$\begin{aligned} \text{最小化または最大化} \quad & x + y \\ \text{制 約} \quad & x^2 + y^2 = 1. \end{aligned}$$

直線 l_2 は、 $f(x, y) := x + y = -1$ を満たす直線であり、 f の等高線になっている。また、直線 l_1, l_3 もある値に対する f の等高線である。ここで、等高線上での $f(x, y)$ の値は l_3, l_2, l_1 の順で大きくなる。以下で、 f の等高線と 局所最適解 の関係について勉強しよう。

円と l_2 の位置関係

まず, f の等高線 l_2 と円との交点の近くを詳しく調べる.



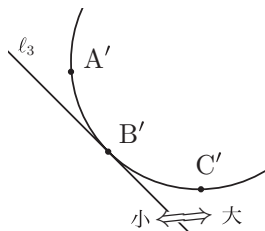
図で点が $A \rightarrow B \rightarrow C$ と動くとき, 関数 f の値は

l_2 と円の交点の近く

(x, y)	A		B		C
$f(x, y)$	大	\searrow	小	\searrow	もっと小

と変化するのが読み取れるだろう. よって, 等高線 l_2 と円の交点
は局所最適解にはならない.

円と l_3 の位置関係



一方，上図より，等高線 l_3 と円の接点付近では

l_3 と円の接点の近く

(x, y)	A'		B'		C'
$f(x, y)$	大	\searrow	小	\nearrow	大

と変化する．ここで， l_3 は円の接線で，接点は $(-1/\sqrt{2}, -1/\sqrt{2})$ である．表より接点に近い円周上の点の中では，接点 $(-1/\sqrt{2}, -1/\sqrt{2})$ で $f(x, y)$ の値が一番小さくなる．したがって， $(-1/\sqrt{2}, -1/\sqrt{2})$ が局所最小解である．

一般の関数の等高線

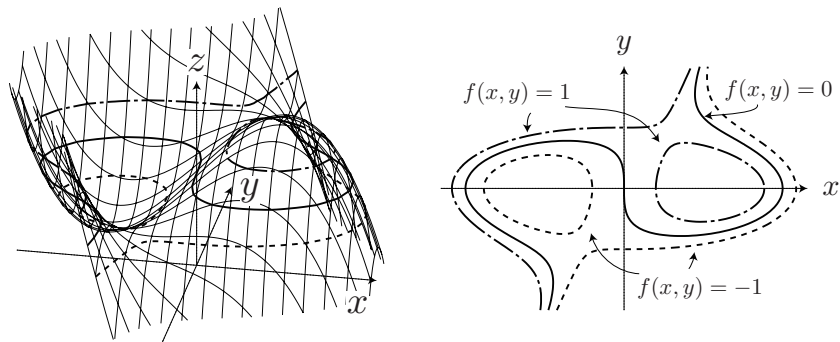
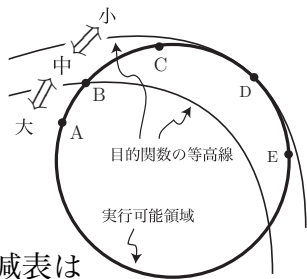


Figure : $f(x, y) = -x^3 - 3xy^2 + y^3 + 3x$ のグラフと等高線

一般の関数を円周上で最適化

目的関数の等高線が図のようになっていると仮定する。このとき、A, B, C, D, E のどの点が局所最小解または局所最大解だろうか？



図で点を $A \rightarrow B \rightarrow C$ と動かすと、増減表は

(x, y)	A		B		C
$f(x, y)$	大	\searrow	中	\searrow	小

となる。したがって、点 B は局所最適解ではない。一方、点を $C \rightarrow D \rightarrow E$ と動かすと、増減表は

(x, y)	C		D		E
$f(x, y)$	小	\searrow	もっと小	\nearrow	小

となるので、点 D は局所最小解になる。

一般の関数を一般の曲線上で最適化

上記の考察より,

局所最適解で目的関数の等高線と実行可能領域は接する
ということがわかる. これより以下の定理を得る.

[定理] (ラグランジュ乗数法)

$$\begin{array}{ll} \text{最小化または最大化} & f(x) \\ \text{制 約} & g(x) = 0 \end{array}$$

に対して, \bar{x} を局所最適解とする. $\nabla g(\bar{x}) \neq \mathbf{0}$ ならば, ある数 λ が存在して,

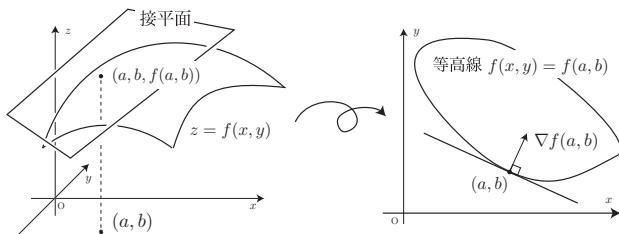
$$\nabla f(\bar{x}) = \lambda \nabla g(\bar{x}), \quad g(\bar{x}) = 0$$

が成り立つ.

勾配ベクトルと等高線は直交している

曲面 $z = f(x, y)$ の (a, b) における接平面：

$$z = f(a, b) + f_x(a, b)(x - a) + f_y(a, b)(y - b)$$



点 $(a, b, f(a, b))$ を通り xy 平面に平行な平面 $z = f(a, b)$ で、グラフと接平面を切った時の断面は、右図になる。グラフの断面は等高線、接平面の断面は接線になる。さらに、この接線は接平面と平面 $z = f(a, b)$ との交線なので、

$$\text{[等高線の接線の式]} \quad f_x(a, b)(x - a) + f_y(a, b)(y - b) = 0.$$

よって、勾配ベクトル $\nabla f(a, b) = \begin{bmatrix} f_x(a, b) \\ f_y(a, b) \end{bmatrix}$ と等高線は直交している。

勾配ベクトルの向きは、関数の増える方向

さらに、接平面の式

$$z = f(a, b) + f_x(a, b)(x - a) + f_y(a, b)(y - b)$$

で、点 (a, b) を $\nabla f(a, b)$ 方向へ移動した点

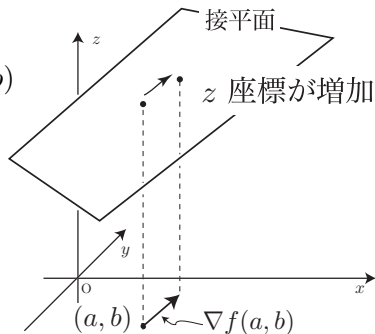
$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} a \\ b \end{bmatrix} + \nabla f(a, b) = \begin{bmatrix} a + f_x(a, b) \\ b + f_y(a, b) \end{bmatrix}$$

における z 座標を調べると

$$z = f(a, b) + f_x(a, b)^2 + f_y(a, b)^2 > f(a, b)$$

となるので、 $\nabla f(a, b)$ は接平面の z 座標が増加する方向を向いている。よって、

勾配ベクトルは関数値が増える方向を向いている



[定理] (ラグランジュ乗数法)

$$\begin{array}{ll} \text{最小化または最大化} & f(x, y) \\ \text{制 約} & g(x, y) = 0 \end{array}$$

に対して, $\bar{x} = (a, b)$ を局所最適解とする. $\nabla g(\bar{x}) \neq \mathbf{0}$ ならば, ある数 λ が存在して, 以下が成り立つ:

$$\nabla f(\bar{x}) = \lambda \nabla g(\bar{x}), \quad g(\bar{x}) = 0.$$

解説.

制約は $g(x, y) = 0$ なので, 実行可能領域は g の等高線. よって,

$$\nabla f(\bar{x}) \perp \{f \text{ の等高線} \} \text{「接する」} \{ \text{実行可能領域} \} \perp \nabla g(\bar{x})$$

よって, $\nabla f(\bar{x})$ と $\nabla g(\bar{x})$ は平行. □

制約付き停留点

[定義]

ある実数 λ に対して,
$$\begin{cases} \nabla f(\bar{x}) = \lambda \nabla g(\bar{x}) \\ g(\bar{x}) = 0 \end{cases}$$

を満たす点 \bar{x} を **制約つき停留点** と呼ぶ. ただし, 省略しても誤解がない場合は, 単に停留点と呼ぶ.

実行可能領域に関する停留点

局所最適解

大域最適解

最適解の存在定理

証明はしないが、最適解の存在定理を一つ挙げておく。

[定理]

目的関数が連続で、実行可能領域が有界閉集合ならば、最適化問題は
大域最小解と大域最大解を持つ。

有界閉集合とは、大きさに限りがあつて、中も端もつまっている
集合。例えば、 $[0, 1]$ は有界閉集合だが、 $(0, 1)$ 、 $[0, \infty)$ は有界閉集
合ではない。特に等式制約を持つ最適化問題は、実行可能領域が
有界閉集合になる場合が多い。

例題

最小化 $f(x, y) := x - y$

制 約 $g(x, y) := 2x^2 + 3y^2 - 1 = 0$

練習問題

- (1) 最小化 $x - y$
制約 $2x^2 + 3y^2 = 1$
- (2) 最小化 $3x^2 + 2y^2 + 4z^2 + 4xy + 4xz$
制約 $x + y + z = 1$