

## 4 制約付き最適化問題

### 4.1 制約付き最適化問題

関数を最小化する際に変数に制約の付いた問題を扱う。

例 4.1. 縦横の辺の長さの和が 4 となる長方形の中で、面積が最大になるのはどのような長方形か?

この問題は次のように定式化できる。縦横の辺の長さをそれぞれ  $x$  と  $y$  とすると、面積は  $xy$  になる。辺の長さの合計が 4 となる長方形を考えるので、問題は

$$\begin{aligned} & \text{最大化 } f(x, y) := xy \\ & \text{制約 } x + y = 4, x \geq 0, y \geq 0 \end{aligned}$$

となる。ここで  $x + y = 4$  より、

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

となる。

点が  $A : (1, 3) \rightarrow B : (2, 2) \rightarrow C : (3, 1)$  と動くとき、関数  $f(x, y) = xy$  の値は、

$(x, y)$	A		B		C
$f(x, y)$	3	↗	4	↘	3

3) 線分上の増減表

と変化する。よって  $x + y = 4$  を満たす点で、B に近いものの中では、B で  $f(x, y)$  の関数値が一番大きくなる。

このような点を局所最大解と呼ぶ。さらに、B はすべての点に対しても最大解にもなっている。このような点を大域最大解と呼ぶ。したがって題意を満たす一辺が 2 の正方形のとき面積が最大になる。以下にきちんとした定義を紹介しよう。

今までと同様に最小化問題で主に説明するが最大化問題も同様である。また変数を一般に  $n$  個として扱う。

集合  $C$  を数ベクトル空間  $\mathbb{R}^n$  の部分集合とする.  $x = (x_1, x_2, \dots, x_n)$  のように書く. 以下のような問題を制約付き最小化問題と呼ぶ.

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

最小化問題 (P) において, 上記のような集合  $C$  を 実行可能領域,  $C$  の点を 実行可能解 と呼ぶ.

定義 4.1.  $\bar{x} \in C$  が  $\bar{x}$  に充分近いすべての  $x \in C$  に対して

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

のとき  $f(\bar{x})$  を (P) の局所最小値,  $\bar{x}$  を局所最小解と呼ぶ.

定義 4.2.  $\bar{x} \in C$  がすべての  $x \in C$  に対して

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

のとき  $f(\bar{x})$  を (P) の大域最小値,  $\bar{x}$  を大域最小解と呼ぶ.

## 4.2 等式制約が一つの場合

### 4.2.1 線形関数を円周上で最適化

まず円周上で線形関数 (一次関数) を最小化または最大化する問題を考える;

$$\begin{aligned} & \text{最小化又は最大化 } f(x, y) := x + y \\ & \text{制約 } x^2 + y^2 = 1. \end{aligned}$$

図 4.1 を見てみよう. 直線  $l_2$  は  $f(x, y) = x + y = -1$  を満たす直線である. このように関数の値が一定になるような集合を 等高線 という. 直線  $l_1, l_3$  もある値に対する  $f(x, y)$  の等高線である<sup>4)</sup>.  $f(x, y) = x + y$  より, 等高線上での  $f(x, y)$  の関数値は  $l_3, l_2, l_1$  の順で大きくなる. 以下で等高線と局所最適解の関係を見てみよう.

さて, 等高線  $l_2$  と円との交点の近くを詳しく見てみる. 拡大図 4.2 のように点 A, B, C を定める. 直線  $l_2$  は  $f(x, y)$  の等高線なので, 点が  $A \rightarrow B \rightarrow C$  と動くとき関数  $f(x, y)$  の値は,

$l_2$  と円の交点の近く

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

<sup>4)</sup> 一次関数の等高線は直線になる. 一般には曲線になる.

<sup>5)</sup> 曲線上の増減表

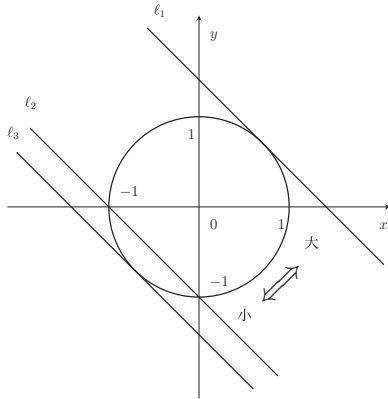


図 4.1  $x^2 + y^2 = 1$  の上で  $x + y$  を最小化又は最大化

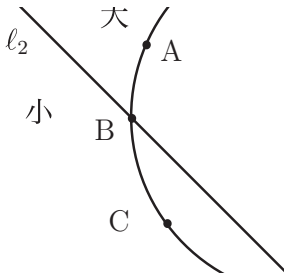


図 4.2  $l_2$  と円の交点の周りで拡大

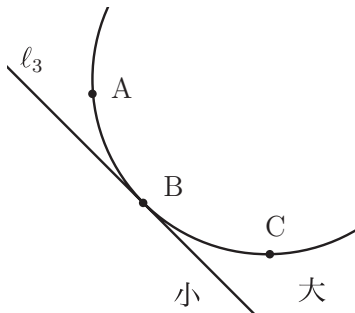


図 4.3  $l_3$  と円の接点の周りで拡大

等高線と実行可能領域  
交差している点は、局  
最適解にならないこと  
わかるだろう。

と変化するのを読み取れるだろう。よって、 $l_2$  と円の交点は局所最適解には  
ならない<sup>6)</sup>。一方、 $l_3$  と円の接点付近では (図 4.3 を見よ) ，

$l_3$  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。よって、 $l_3$  と円の接点  $(-1/\sqrt{2}, -1/\sqrt{2})$  の近くにある円周上

7) この場合特に大域最小解にもなっている。

8) 大域最大解でもある。

の点の中では、 $(-1/\sqrt{2}, -1/\sqrt{2})$  が  $f(x, y)$  の値を一番小さくする。したがって、 $(-1/\sqrt{2}, -1/\sqrt{2})$  が局所最小解である<sup>7)</sup>。同様に、等高線  $l_1$  と円の接点は局所最大解になっていることがわかるだろうか<sup>8)</sup>。

#### 4.2.2 一般の関数を円周上で最適化

一般の関数を円周上で最適化する問題を考えよう。目的関数が一次関数でなくても等高線を見れば局所最適解が分かる。一般の関数では等高線は曲線になる。目的関数の等高線が図 4.4 のようになった時、A, B, C, どの点が

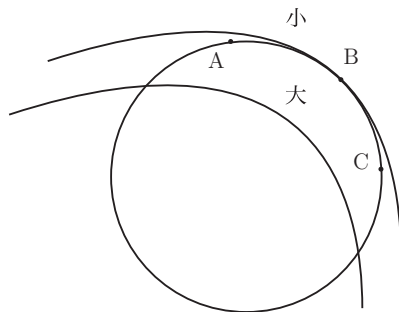


図 4.4 円（実行可能領域）と目的関数の等高線 1

局所最小解または局所最大解であるかわかるだろうか？ 点を  $A \rightarrow B \rightarrow C$  と動かすと、目的関数の値は、大  $\searrow$  小  $\nearrow$  大と変化するのを読み取れる。したがって、点 B が局所最小解になることがわかる。

それでは、目的関数の等高線が図 4.5 のようになった場合はどうだろうか？

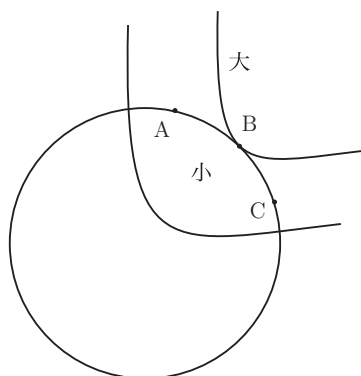


図 4.5 円（実行可能領域）と目的関数の等高線 2

この場合は、点を  $A \rightarrow B \rightarrow C$  と動かすと、目的関数の値は、小  $\nearrow$  大  $\searrow$

小と変化するので、点 B は局所最大解になる。

上記のように、ある点が局所最適解ならば、そこで等高線と実行可能領域が接することがわかる。また、目的関数の勾配ベクトルは等高線に直交することをふまえると、以下の定理を得る。

定理 4.1 (ラグランジュの乗数法). 最適化問題

最小化, 又は最大化  $f(x)$

$$\text{制約 } g(x) = 0$$

を考える.  $\bar{x}$  を局所最適解とする.  $\nabla g(\bar{x}) \neq (0, \dots, 0)$  ならば, ある数  $\lambda$  が存在して,

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

$$g(\bar{x}) = 0$$

が成り立つ.

解説. まず,  $g(x) = 0$  を満たす  $x$  の集合も  $g(x)$  の等高線の一種なので, 実行可能領域も  $\nabla g(x)$  と直交していることに注意する. いま, 上記の考察から  $\bar{x}$  が局所最適解のとき,  $f(x)$  の等高線と実行可能領域は  $\bar{x}$  で接している. このとき  $\nabla f(\bar{x})$  と  $\nabla g(\bar{x})$  は平行になっている. それを式で書くと  $\nabla f(\bar{x}) = \lambda \nabla g(\bar{x})$  となる.

証明. 2変数の最小化問題の場合を証明する.  $\bar{x} = (\bar{x}_1, \bar{x}_2)$  を局所最小解とする.  $g(\bar{x}) = 0$  は制約を満たすことから成り立つ. いま, 関数  $u(t) = (x_1(t), x_2(t))$  で,  $t = 0$  の近くで微分可能で,  $u(0) = \bar{x}$ ,  $g(u(t)) = 0$  を満たすものが存在するとしよう. すると定理の主張は簡単に示せる. 実際,  $\bar{x}$  は局所最小解で,  $u(t)$  は制約を満たしているので,  $f(u(t)) \geq f(\bar{x}) = f(u(0))$  が成り立つ. よって, 制約なしの最適化問題

$$\text{最小化 } f(u(t)), t \in \mathbb{R}$$

は  $t = 0$  で局所最小値を取る. したがって, 合成関数  $f \circ g$  の  $t = 0$  における微分  $\frac{d}{dt} f(u(t))|_{t=0} = \nabla f(\bar{x})u'(0) = 0$  が成り立つ. また,  $t = 0$  の近くで  $g(u(t)) = 0$  なので, 微分も  $\frac{d}{dt} g(u(t))|_{t=0} = \nabla g(\bar{x})u'(0) = 0$  を満たす. したがって,  $\nabla f(\bar{x}), \nabla g(\bar{x})$  2次元ベクトルで一つのベクトル  $u'(0)$  に直交しているので, ある定数  $\lambda$  が存在して,  $\nabla f(\bar{x}) = \lambda \nabla g(\bar{x})$  が成り立つ.

さて, 上記のような  $u(t)$  の存在を示そう. これには陰関数定理を用いる.  $\nabla g(\bar{x}) \neq 0$  より, 特に  $g_{x_2}(\bar{x}) \neq 0$  とする. すると, 陰関数定理より関数  $\phi$  で,  $\bar{x}_1$  の近くで微分可能で,  $g(x_1, \phi(x_1)) = 0$  ( $x_1$  は  $\bar{x}_1$  に近い任意の数),  $\phi(\bar{x}_1) = \bar{x}_2$  を満たすものが存在する. よって,  $t = 0$  の近くで,  $u(t) = (\bar{x}_1 + t, \phi(\bar{x}_1 + t))$

と定義すると上記を満たす関数  $u$  を得る. □

補足. 3 変数以上の場合には, 上の議論を工夫して  $\langle \nabla g(\bar{x}), d \rangle = 0$  を満たす任意のベクトル  $d$  に対して,  $u'(0) = d$  となるような  $u(t)$  を見つける必要がある. すると,  $\nabla g(\bar{x})$  に直交するすべてのベクトルに  $\nabla f(\bar{x})$  が直交することが言えて,  $\nabla f(\bar{x}) = \lambda \nabla g(\bar{x})$  が成り立つ.

補足. 証明を見れば分かるように最大化問題に対しても同様のことが言える.

定理の式を満たすような点  $\bar{x}$  を 実行可能領域に関する停留点と呼ぶ. 通常の停留点と同じように, 実行可能領域に関する停留点であっても最適解にならないものは存在する. 省略しても誤解がない場合は, 単に停留点と呼ぶ. 証

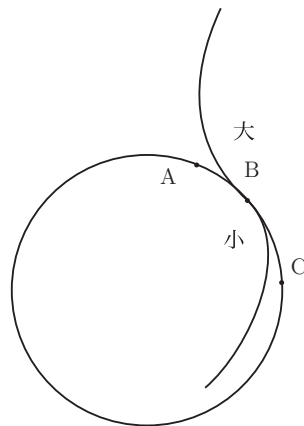


図 4.6 局所最適解にならない停留点

明はしないが (実数の性質を使う) 最適解の存在定理を一つ挙げておく.

<sup>9)</sup> 大きさに限りがあって, 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.

定理 4.2. 目的関数が連続で, 実行可能領域が有界閉集合<sup>9)</sup> ならば, 最適化問題は最小解と最大解を持つ.

制約付き最適化問題で特に等式制約を持つ問題は, 実行可能領域が有界閉集合になる場合が多い. 最適解の存在があらかじめ分かっていると議論が楽になり, 1 階微分の情報 (ラグランジュの乗数法) のみで, 最適解を探せることも多い.

### 4.2.3 等式制約付き最小化問題の解き方

この節では具体的に例題を解いてみよう. 以下の例題は最適解を持つ.

例 4.2.

$$\begin{aligned} \text{最小化 } f(x, y) &:= x - y \\ \text{制約 } g(x) &:= 2x^2 + 3y^2 - 1 = 0 \end{aligned}$$

ラグランジュの乗数法 (定理 4.1) より, 局所最適解は実行可能領域に関する停留点になる<sup>10)</sup>. 定義より, 停留点は

$$\begin{cases} \nabla f(x, y) = \lambda \nabla g(x, y), \\ g(x, y) = 0 \end{cases}$$

<sup>10)</sup>  $g(x, y) = 0$  を満たす点には  $\nabla g(x, y)$  を零ベクトルにする点はないので, 定理 4.1 を適用可能

を満たす  $(x, y)$  である. いま, 目的関数と制約式の勾配ベクトルは, それぞれ  $\nabla f(x, y) = (1, -1)$ ,  $\nabla g(x, y) = (4x, 6y)$  となるので,

$$\begin{cases} 1 = \lambda 4x \\ -1 = \lambda 6y \\ 2x^2 + 3y^2 - 1 = 0 \end{cases}$$

を満たす  $(x, y), \lambda$  を見つければ, その  $(x, y)$  が停留点になる ( $\lambda$  は存在すれば良く, 必ずしも求める必要はない).

第一式より  $4x \neq 0$  となるので,  $\lambda = 1/(4x)$  として第二式に代入すると,  $y = -2x/3$  を得る. これを第三式に代入すると,  $x = \pm\sqrt{3/10}$  が求まる. このとき  $y = \mp\sqrt{2/15}$  となる. よって停留点は  $(x, y) = (\pm\sqrt{3/10}, \mp\sqrt{2/15})$  である.

よって,  $(x, y) = (\pm\sqrt{3/10}, \mp\sqrt{2/15})$  での目的関数  $f$  の値を比べると, 最小解が  $(-\sqrt{3/10}, \sqrt{2/15})$ , 最小値が  $f(-\sqrt{3/10}, \sqrt{2/15}) = -\sqrt{3/10} - \sqrt{2/15}$  である.

例 4.3 (2 次形式の最小値).  $A = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$  を実対称行列として以下の最小化問題を考える.

$$\begin{aligned} (P) \text{ 最小化 } f(x, y) &:= \begin{bmatrix} x & y \end{bmatrix} A \begin{bmatrix} x \\ y \end{bmatrix} = ax^2 + 2bxy + cy^2 \\ \text{制約 } g(x, y) &= x^2 + y^2 - 1 \end{aligned}$$

すると, 最小値は  $A$  の最小固有値, 最小解はその固有ベクトルになる.

ラグランジュの乗数法 (定理 4.1) より, 最適解は実行可能領域に関する停留点になる<sup>11)</sup>. 停留点は

$$\begin{cases} \nabla f(x, y) = \lambda \nabla g(x, y), \\ g(x, y) = 0 \end{cases}$$

<sup>11)</sup>  $g(x, y) = 0$  を満たす点で  $\nabla g(x, y)$  は零ベクトルにならない.

を  $(x, y), \lambda$  について解けばよい. いま  $\nabla f(x, y) = (2ax + 2by, 2bx + 2cy)$ ,

$\nabla g(x, y) = (2x, 2y)$  より, 第一式は

$$\begin{bmatrix} 2ax + 2by \\ 2bx + 2cy \end{bmatrix} = \lambda \begin{bmatrix} 2x \\ 2y \end{bmatrix}$$

なので, 停留点の式は

$$\begin{cases} \begin{bmatrix} a & b \\ b & c \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \lambda \begin{bmatrix} x \\ y \end{bmatrix}, \\ x^2 + y^2 - 1 = 0 \end{cases}$$

となる. これは  $\lambda$  が  $A$  の固有値で,  $(x, y)$  がその固有ベクトルあり, かつ大きさが 1 であるということを示す. よって, 停留点は  $A$  の固有ベクトルと一致する<sup>12)</sup>.

<sup>12)</sup>  $2 \times 2$  行列なので固有ベクトルは 2 つある. ただしその 2 つが等しい場合もある.

ここで,  $\lambda, (\bar{x}, \bar{y})$ , をそれぞれ  $A$  の固有値と固有ベクトルとすると,

$$f(x, y) = [x \ y] A \begin{bmatrix} x \\ y \end{bmatrix} = [x \ y] \left( \lambda \begin{bmatrix} x \\ y \end{bmatrix} \right) = \lambda(x^2 + y^2) = \lambda$$

となる. よって, 停留点での  $f$  の値は固有値<sup>13)</sup> と一致する. したがって, 最小化問題の最小値は  $A$  の固有値の中で小さい方であり, 最小解は  $A$  の最小固有値に対する固有ベクトルである.<sup>14)</sup>

<sup>13)</sup> 線形代数の定理より, 実対象行列の固有値は実数となる

<sup>14)</sup> 同様に, 最大値は  $A$  の固有値の中で大きい方であり, 最大解は  $A$  の最大固有値に対する固有ベクトルである.

#### 4.2.4 等式制約が複数の場合

制約は複数ある場合もある.

**定理 4.3.** 最小化問題

最小化  $f(x)$

$$\text{制約 } g_1(x) = 0, g_2(x) = 0, \dots, g_m(x) = 0$$

を考える.  $\bar{x}$  を局所最小解とする.  $\nabla g_1(\bar{x}), \dots, \nabla g_m(\bar{x})$  が一次独立ならば, ある数  $\lambda_1, \dots, \lambda_m$  が存在して,

$$\nabla f(\bar{x}) = \lambda_1 \nabla g_1(\bar{x}) + \dots + \lambda_m \nabla g_m(\bar{x})$$

$$g_1(\bar{x}) = 0, \dots, g_m(\bar{x}) = 0$$

が成り立つ.

証明には次のような陰関数定理を用いる. 簡単のため変数の数が 3 つ, 等式の数 が 2 つの場合を記述する.

**定理 4.4.**  $g_1, g_2$  を 3 変数関数,  $g_1(\bar{x}) = 0, g_2(\bar{x}) = 0$  とする.



$$\text{rank} \begin{bmatrix} \nabla g_1(\bar{x})^T \\ \nabla g_2(\bar{x})^T \end{bmatrix} = 2$$

ならば, ある関数  $\phi_1, \phi_2$  で,

$$\begin{bmatrix} g_1(x_1, \phi_1(x), \phi_2(x)) \\ g_2(x_1, \phi_1(x), \phi_2(x)) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

を満たすものが存在する.

解説.  $g_1, g_2$  が線形だとすると,  $g_1(\bar{x}) = 0, g_2(\bar{x}) = 0$  は, ある  $2 \times 3$  行列を用いて,

$$\begin{bmatrix} g_1(\bar{x}) \\ g_2(\bar{x}) \end{bmatrix} = A\bar{x} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

と書ける. また,  $\text{rank} \begin{bmatrix} \nabla g_1(\bar{x})^T \\ \nabla g_2(\bar{x})^T \end{bmatrix} = \text{rank} A$  である.  $\text{rank} A = 2$  の場合, 方程式の解の自由度が 1 になるので 1 つのパラメータを用いて解を表せる. この定理が言っているのは,  $g_1, g_2$  が非線形の場合にも, 微分の階数が 2 であれば, 方程式の解は 1 つのパラメータと非線形関数を用いて表せる, ということである.  $\square$

問題 4.1. 1. 実行可能領域に関する停留点とその点での目的関数の値を求め, 最小解を探せ.

(a)

$$\begin{aligned} \text{最小化 } f(x, y) &:= 2x - 3y \\ \text{制約 } g(x, y) &:= 2x^2 + y^2 - 1 = 0 \end{aligned}$$

(b)

$$\begin{aligned} \text{最小化 } f(x, y, z) &:= x - y + 2z \\ \text{制約 } g(x, y, z) &:= x^2 + y^2 + 2z^2 - 1 = 0 \end{aligned}$$

(c)

$$\begin{aligned} \text{最小化 } f(x, y) &:= x^3 + y^3 \\ \text{制約 } g(x, y) &:= x^2 + y^2 - 1 = 0 \end{aligned}$$

2.  $x, y$  が原点を中心とする半径 1 の円周上の点であるとき,  $2x^2 - 6xy + 2y^2$  の最小値と最大値を求めよ.