

第3章 最適化問題

3.1 最適化問題とは？

最適化問題とは関数を最小化，又は最大化する問題である．まず，以下の具体例を紹介しよう．

例 9. 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

$$\text{最大化 } f(x, y) := xy$$

$$\text{制約 } x + y = 4, x \geq 0, y \geq 0$$

となる．このような問題を（制約付き）最適化問題と呼ぶ．

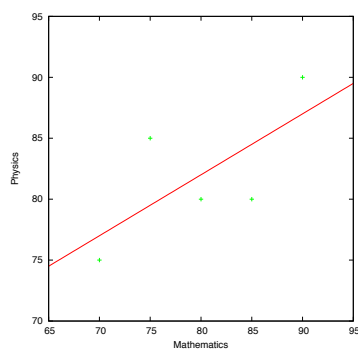
例 10. 平面に 4 点 $(1, 3)$, $(2, 5)$, $(3, 5)$, $(4, 9)$ が与えられたとき，これらの点の最も近くを通る直線は？

実験データの整理

この問題は実験データの整理に使われる．例えば，ばねにおもりを吊るした時の伸びが以下ようになったとしよう．

| | | | | | |
|----------------|-----|-----|-----|-----|-----|
| 重さ (g) | 20 | 40 | 60 | 80 | 100 |
| ばねの伸び (cm) | 1.9 | 3.8 | 6.1 | 8.1 | 9.9 |

これらの点を平面上に描くと，フックの法則より，一直線上にあるはずである．しかし，実際には計測誤差があり，これらの点は直線上にない．上の問題を解くと，この誤差のあるデータから最も確からしい直線を求めることができる．



上記の問題は以下の最適化問題を用いて解くことができる．直線は一般に

$$y = ax + b$$

と書け、点 $(1, 3)$ と直線との誤差は

$$3 - (1 \cdot a + b)$$

となる（丁度直線上にあれば誤差は 0）．この誤差の二乗の和を最小にする直線を求める方法を最小二乗法と呼ぶ．これは以下のような最小化問題になる．

$$\begin{aligned} \text{最小化 } f(a, b) = & \{3 - (a + b)\}^2 + \{5 - (2a + b)\}^2 \\ & + \{5 - (3a + b)\}^2 + \{9 - (4a + b)\}^2 \end{aligned}$$

制約 なし

このような問題を 制約無し最適化問題 と呼ぶ．

上記の問題では、長方形と直線を求める問題が 2 変数関数の最小値を求める問題になった．このように、数理的な問題は最適化問題として定式化できる場合が多い．

3.1.1 制約無しの最適化問題

一般に最小化問題は、例 9 のように、 x を選ぶ範囲になんらかの制約があるものが多いが、まずは制約の無い最小化問題から勉強しよう：

$$\begin{aligned} \text{最小化 } & f(x) \\ \text{制約 } & \text{無し} \end{aligned} \tag{3.1}$$

ここで、最小化する関数 $f(x)$ を 目的関数 と呼ぶ．関数を最大化する問題を最大化問題と呼び、最小化問題と最大化問題をまとめて、最適化問題 と呼ぶ．

最適解の定義

最適化問題の解である最適解にはいくつかの種類がある．制約無しの最小化問題に対しては、次のように定義される．

定義. \bar{x} がすべての $x \in \mathbb{R}^n$ に対して

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

のとき $f(\bar{x})$ を (3.1) の大域最小値、 \bar{x} を大域最小解と呼ぶ．

大域最小解を見つけられれば最も良いが、それは一般的に難しい．そこで、より見つけ易い以下の解を探ることがある．

定義. \bar{x} に充分近いすべての $x \in \mathbb{R}^n$ に対して

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

のとき $f(\bar{x})$ を (3.1) の局所最小値、 \bar{x} を局所最小解と呼ぶ．

最大化問題の場合は不等式が逆向きになり、それぞれ『最小』を『最大』に置き換える．最小化と最大化をひとまとめにして、局所最適値、局所最適解 などと呼ぶ．

例 11.

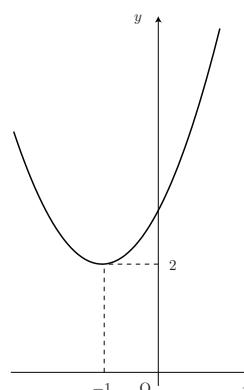
$$\text{最小化 } f(x) = x^2 + 2x + 3$$

$$f(x) = x^2 + 2x + 3 = (x + 1)^2 + 2$$

より、最小値は 2，最小解は $x = -1$ である。
ここで、

$$f(x) \geq f(-1) = 2 \text{ (全ての } x \text{)}$$

となる。つまり、この最小解はすべての実数 x に対して、最小値を与える解になっていることがわかる。このような最小解は大域最小解、値は大域最小値になる。



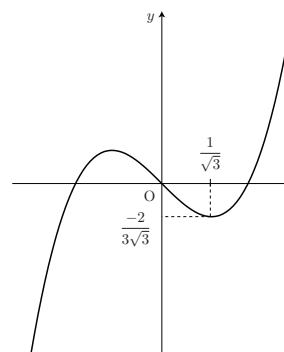
例 12. 一方、最小化問題

$$\text{最小化 } f(x) = x^3 - x$$

には、図より大域最小解は存在しない。しかし、

$$f(x) \geq f\left(\frac{1}{\sqrt{3}}\right) = \frac{-2}{3\sqrt{3}} \text{ (} \frac{1}{\sqrt{3}} \text{ に十分近い } x \text{)}$$

となる。このように、局所的に最小になっている点が局所最小解で、値が局所最小値である。



極値問題

微分積分で扱う 極値問題 も最適化問題である。ただし、最適解の中でも特別なものを求めることを目的としている。

定義.

- $x \neq \bar{x}$ を満たし、 \bar{x} に充分近い x に対して、 $f(x) > f(\bar{x})$ が成り立つとき、 f は \bar{x} で 極小値 をとると言う。
- $x \neq \bar{x}$ を満たし、 \bar{x} に充分近い x に対して、 $f(x) < f(\bar{x})$ が成り立つとき、 f は \bar{x} で 極大値 をとると言う。

極小値と極大値をひとまとめにして、極値 と呼ぶ。

局所最小値は等号も含んだ「 $f(x) \geq f(\bar{x})$ 」で定義されるので、極小値は局所最小値になる。注意して欲しいのは、極値を取る点は局所的に一意的最適解になっているということである。包含関係は図 3.1 のようになる。

局所最小値を持つが、極小値を持たない例をあげる。

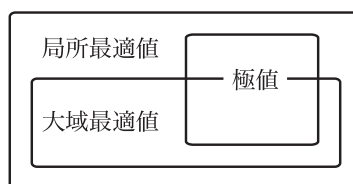


図 3.1: 最適値と極値の関係図

例 13 (極小値にならない最小値).

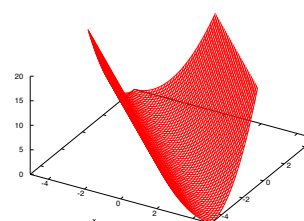
$$\text{最小化 } f(x, y) = x^2 + 2xy + y^2$$

$f(x, y) = (x + y)^2$ より, すべての (x, y) に対して $f(x, y) \geq 0$ となる. また

$$f(x, y) = (x + y)^2 = 0 \Leftrightarrow x + y = 0$$

である. 例えば, $f(0, 0) = 0$ より, 0 は f の大域最小値, $(0, 0)$ は f の大域最小解になる.

一方, $(x, y) = (t, -t)$ という点の上では, f の値は $f(t, -t) = (t - t)^2 = 0$ となる. したがって, どんなに $(0, 0)$ の近くを考えても, そこに $f(0, 0)$ と同じ値をとる点があるので, f は $(0, 0)$ で極小値をとらない (最小解が一意でない).

図 3.2: $x^2 + 2xy + y^2$ のグラフ

3.2 最適性条件

最適解を求めるために, まず最適解と関係する性質と, それらと関数のグラフとの関係について学ぶ. 最大化問題は目的関数に (-1) をかけると最小化問題になるので, この本では主に最小化問題についてのみ解説する.

3.2.1 1 次最適性条件

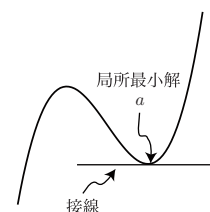
1 変数関数の場合を復習する. 点 \bar{x} が局所最小解となるのは, 右図のような関係が典型例である. このとき, 図より \bar{x} における f の接線の傾きは 0 である. よって, $f'(\bar{x}) = 0$ が成り立つ.

同様のことが多変数関数でも成り立つ.

定理 6 (一次最適性条件). 点 \bar{x} が局所最適解ならば,

$$\nabla f(\bar{x}) = \mathbf{0} \quad (\text{零ベクトル})$$

が成り立つ.



この定理の結論にある条件は非常に重要なものなので、以下のような名前をつける。
定義. 点 a が,

$$\nabla f(a) = \mathbf{0}$$

を満たすとき, a を f の停留点と呼ぶ.

解説. 次の具体例でこの定理の意味を考えよう. 最小化問題

$$\text{最小化 } f(x, y) = x^3 - 3xy + y^3$$

の局所最小解は $(x, y) = (1, 1)$ である. 目的関数の勾配ベクトルは

$$\nabla f(x, y) = \begin{bmatrix} 3x^2 - 3y \\ 3y^2 - 3x \end{bmatrix}$$

である. いま, 目的関数 f のグラフは図 3.3 のようになる.

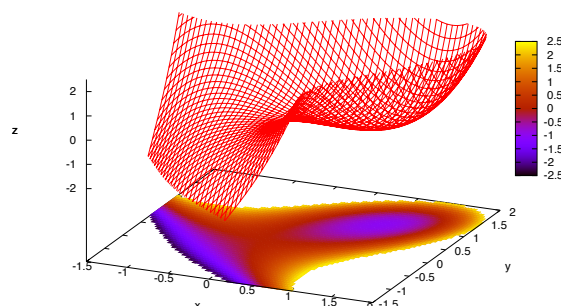


図 3.3: $z = x^3 - 3xy + y^3$ のグラフ

ここで, 点 $(x, y, z) = (1, 1, f(1, 1))$ はグラフのどの位置にあるだろうか? 正解をいうと, 点 $(1, 1, f(1, 1))$ は, 座標の大体の位置からグラフの下に窪んだ部分にあることがわかり. さらに局所最小解であるので, その窪みの一番底に位置しているのである. 2変数関数の場合は, このような形と位置の関係が局所最小解の典型的な例になる.

ここで, 窪みの一番底に接する平面を描くと, 水平な平面になるだろう. いま, $u = (x, y)$ とすると, 点 $(1, 1)$ における f の接平面は公式より

$$z = f(1, 1) + f_x(1, 1)(x - 1) + f_y(1, 1)(y - 1)$$

となる. 一方, 点 $(1, 1, f(1, 1))$ を通り, 水平な平面の式は

$$z = f(1, 1) \quad (\text{定数})$$

であり, この二つの平面が等しいので,

$$\nabla f(1, 1) = \begin{bmatrix} f_x(1, 1) \\ f_y(1, 1) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

が導かれる.

練習問題 1. 次の関数の停留点を求めよ.

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

定理 6 の証明

2 変数の場合を示す. ベクトル (x, y) の大きさを $\|(x, y)\| = \sqrt{x^2 + y^2}$ とおく. (a, b) を局所最小解とする. 定義より, (a, b) に充分近い (x, y) に対して

$$f(x, y) \geq f(a, b)$$

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

$$h(t) = f((a, b) + t(x, y))$$

とおく. 充分小さい $t > 0$ に対して, $(a, b) + t(x, y)$ は (a, b) に近くなるので,

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

が成り立つ. よって $t = 0$ は $h(t)$ の局所最小値なので, $h'(0) = 0$ が成り立つ.

いま

$$h'(t) = \frac{d}{dt} \{f(a + tx, b + ty)\} = f_x(a + tx, b + ty)x + f_y(a + tx, b + ty)y$$

より,

$$h'(0) = f_x(a, b)x + f_y(a, b)y = 0$$

を得る. (x, y) の選び方より, この等式は任意の $\|(x, y)\| = 1$ を満たす (u, v) に対して成り立つ. したがって, それぞれ $(x, y) = (1, 0)$ と $(x, y) = (0, 1)$ を代入することにより, $\nabla f(a, b) = \mathbf{0}$ が成り立つ.

3.2.2 停留点と局所最適解

局所最適解ならば停留点となるが, 停留点であっても, 局所最適解とは限らない.

例 14.

$$\text{最小化 } f(x, y) = x^2 - y^2$$

では $\nabla f(0, 0) = \mathbf{0}$ となるが, 局所最小解ではない. 実際, $f(0, 0)$ と, 点 $(0, 0)$ に近い点 (x, y) での f の値を比べても, $f(0, 0)$ は最も小さい値にはなっていない.

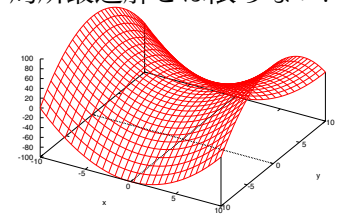


図 3.4: $x^2 - y^2$ のグラフ

一方で, 局所最適解が存在すれば, 定理 6 よりそれは常に停留点になる. よってこのとき, 停留点をすべて見つければ最適解は必ずその中にある. このことから, 停留点を見つけることは重要なのである.

3.2.3 2 次の最適性条件

3.2. 最適性条件

1 階微分の情報を用いて、停留点が求められたが、そのなかには局所最適解になっていないものもある。それでは、停留点の中から、どのように局所最適解を見つけたら良いだろうか？

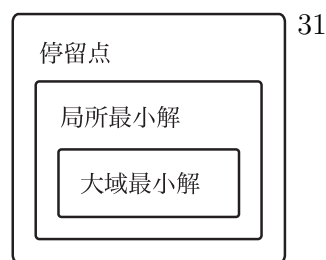


図 3.5: 最適解と停留点の関係図

グラフからの推測

定理 6 の解説の例をもう一度挙げる。

$$\text{最小化 } f(x, y) = x^3 - 3xy + y^3$$

f の停留点は $(0, 0)$ と $(1, 1)$ である。コンピュータでグラフを描くと図 3.6 のようになった。いま、点 $(0, 0)$ はグラフが捻れた位置、点 $(1, 1)$ は下に窪んだ位置にある。グラフがわかっているので、 $(0, 0)$ は局所最小解ではなく、 $(1, 1)$ が局所最小解であることがわかる。

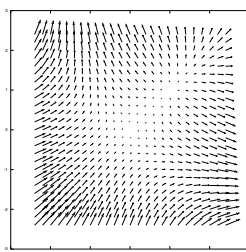
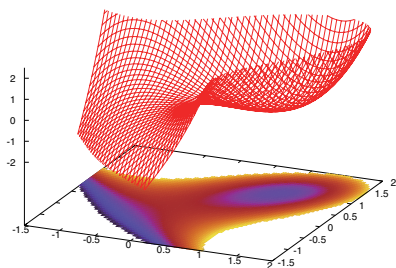


図 3.6: $z = x^3 - 3xy + y^3$ のグラフ
図 3.7: $x^3 - 3xy + y^3$ の勾配ベクトル

ヘッセ行列を用いた判定法

それでは、グラフがわからないときに、目的関数 f の式から停留点が局所最小解かどうかを知ることはできないだろうか？

停留点 \bar{x} を局所最小解とする。すると、グラフ全体の形がわからなくても、少なくとも点 \bar{x} ではグラフは局所的に下に窪んでいることになる。これを別の言葉で言い換えてみよう。すると、『局所的に下に窪んでいる』というのは、点 \bar{x} の近くで、関数 f は『局所的に凸関数である』と言い換えて良いだろう。これを数学的に表現し整理すると、以下の定理を得る。

定理 7 (2 次の最適性条件).

- (1). (必要性) \bar{x} が局所最小解である $\Rightarrow \nabla f(\bar{x}) = \mathbf{0}$ かつ $\nabla^2 f(\bar{x})$ が半正定値
- (2). (十分性) $\nabla f(\bar{x}) = \mathbf{0}$ かつ $\nabla^2 f(\bar{x})$ が正定値
 $\Rightarrow \bar{x}$ は局所最小解である。さらにそこで極小値を取る。

(3). (否定) $\nabla^2 f(\bar{x})$ が不定値のとき, \bar{x} は局所最適解ではない.

局所最大解についても, それぞれ対応する箇所を半負定値, 負定値, 極大値に置き換えたものが成り立つ.

解説. 定理直前の解説より, (1) が推測できる. (2), (3) はこの解説のあとに説明し, 定理の証明は後の節で行う. 凸関数や半正定値の定義を見たときに, 非常に都合の良い性質に見えたかもしれないが, 実は関数が局所最小値をもてば, 自然に現れる性質である. また, 十分性にある正定値の条件も非常に強い条件に見えるが, あながちそうとは言えない. \square

2 次の最適性条件の幾何的イメージ

凸関数の概念を用いて, 2 次の最適性条件の (2) ついて考察してみよう.

まず復習すると, すべての (x, y) に対して, $\nabla^2 f(x, y)$ が正定値であると f は狭義凸関数になるのであった. いま, $\nabla^2 f(a, b)$ が正定値で, (x, y) が (a, b) に十分近い点であると仮定する. すると $f_{xx}(x, y), f_{xy}(x, y), f_{yy}(x, y)$ の値も $f_{xx}(a, b), f_{xy}(a, b), f_{yy}(a, b)$ に近い. したがって,

$$\begin{aligned} \nabla^2 f(a, b) \text{ が正定値} &\Rightarrow \nabla^2 f(x, y) \text{ が正定値} \\ &\Rightarrow f \text{ が } \text{「}(a, b) \text{ の近くで狭義凸関数」} \end{aligned}$$

従って, $\nabla^2 f(a, b)$ が正定値であり, さらに点 (a, b) が停留点ならば, その点は関数 f の「局所的に狭義凸」な部分の底にあるということを表している. また狭義凸であることから, 窪みの底はちゃんと突き出ている, 点 (a, b) と同じ値をとる点が近くにはない. よって, そこで極小値を取る.

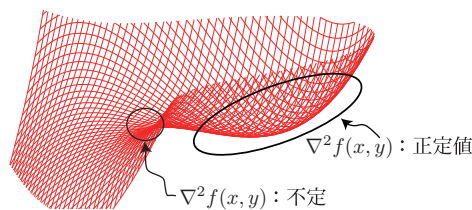


図 3.8: グラフとヘッセ行列の正値性との関係

次に条件 (3) について考えよう. ヘッセ行列 $\nabla^2 f(a, b)$ が不定のときは, 行列式 $|\nabla^2 f(x, y)| < 0$ が (a, b) に近い (x, y) で成り立つ. よって, f が (a, b) の近くで凸関数にも凹関数にならず, グラフが捻れているということを表している.

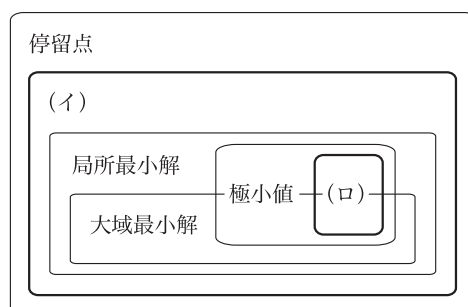


図 3.9: 最適性と 2 次の条件の関係

(イ) $\nabla f(a, b) = \mathbf{0}$ かつ $\nabla^2 f(a, b)$ が半正定値

(ロ) $\nabla f(a, b) = \mathbf{0}$ かつ $\nabla^2 f(a, b)$ が正定値

2 次の最適性十分条件で分からない最適解

図 3.9 の包含関係が示すように, (a, b) が条件 (ロ) を満たさなくても最小解になることがある.

例 15.

$$\text{最小化 } f(x, y) = x^4 + y^4$$

を考える. いま,

$$\nabla f(x, y) = \begin{bmatrix} 4x^3 \\ 4y^3 \end{bmatrix}, \quad \nabla^2 f(x, y) = \begin{bmatrix} 12x^2 & 0 \\ 0 & 12y^2 \end{bmatrix}$$

となり, 停留点は $(0, 0)$ である. $(0, 0)$ でのヘッセ行列はゼロ行列になるので条件 (ロ) は満たさない. しかし, 任意の (x, y) で $f(x, y) \geq 0$ なので, $(0, 0)$ は大域最小解になっている. さらに $(x, y) \neq (0, 0)$ のとき, $f(x, y) > 0$ なので, $(0, 0)$ は極小値である.

3.3 局所最適解の求め方

それでは 1 次と 2 次の最適性条件を用いて例題を解いてみよう.

例 16.

$$\text{最小化 } f(x, y) = x^2 + xy + y^2 - 9x - 9y + 27$$

の局所最適値を求めよ.

この例では, 定理と用語の関連性を明確にするため冗長な解答を書く. 実際に解答を書く場合は他の例を参考にして欲しい.

まず1次の最適性条件(定理6)より, 局所最適解が存在すれば, それらは全て停留点になる. よって, 停留点を求める. 勾配ベクトルは

$$\nabla f(x, y) = \begin{bmatrix} 2x + y - 9 \\ x + 2y - 9 \end{bmatrix}$$

なので, 停留点は連立方程式

$$\begin{cases} 2x + y - 9 = 0 \\ x + 2y - 9 = 0 \end{cases}$$

の解となる. これを解くと, $(x, y) = (3, 3)$ を得る.

この点に対して2次の最適性条件(定理7)を調べる. ヘッセ行列を計算すると

$$\nabla^2 f(x, y) = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$

となるので, $(x, y) = (0, 0)$ のとき,

$$\nabla^2 f(0, 0) = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$

である. この行列の正値性を調べる. 正値性の調べ方は以下の二通りある.

(解法1) 行列 $\nabla^2 f(0, 0)$ の固有値を求めると 1, 3 となるので, 定理4より $\nabla^2 f(0, 0)$ は正定値である.

(解法2) $f_{xx}(0, 0) > 0$ かつ $|\nabla^2 f(0, 0)| > 0$ なので, 定理5より, 行列 $\nabla^2 f(0, 0)$ は正定値である.

よって, 定理??より $(3, 3)$ は局所最小解であり, 局所最小値は $f(3, 3) = 0$ となる. グラフの概形は下図を参照して欲しい.

例 17.

$$f(x, y) = x^3 - 3xy + y^3$$

の局所最適値を求めよ.

勾配ベクトルとヘッセ行列を求めると,

$$\nabla f(x, y) = \begin{bmatrix} 3x^2 - 3y \\ 3y^2 - 3x \end{bmatrix}, \quad \nabla^2 f(x, y) = \begin{bmatrix} 6x & -3 \\ -3 & 6y \end{bmatrix}$$

となる. まず停留点を求める. $\nabla f(x, y) = \mathbf{0}$ を書き下すと, 連立方程式

$$\begin{cases} 3x^2 - 3y = 0 \\ 3y^2 - 3x = 0 \end{cases}$$

を得る. これを解くと停留点は $(x, y) = (0, 0), (1, 1)$ となる.

次にヘッセ行列の正値性を調べる (2変数の場合は, 計算が楽なことから行列式を使う方法を用いる).

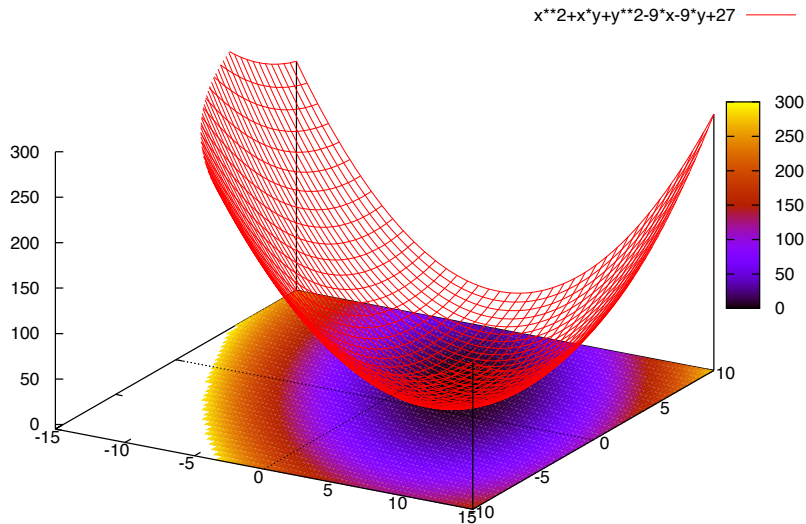


図 3.10: $x^2 + xy + y^2 - 9x - 9y + 27$ のグラフ. $(x, y) = (3, 3)$ が停留点

- $(0, 0)$ ではヘッセ行列の行列式は $|\nabla^2 f(0, 0)| < 0$ を満たすので, ヘッセ行列は不定値. よってこの点は局所最適解ではない.
- $(1, 1)$ では $f_{xx}(1, 1) = 6 > 0$ かつ $|\nabla^2 f(1, 1)| = 27 > 0$ なので, ヘッセ行列は正定値. よってこの点は局所最小解で, 局所最小値は -1 である.

これを表に書くと

| (x, y) | $(0, 0)$ | $(1, 1)$ |
|----------------------|----------|----------|
| $ \nabla^2 f(x, y) $ | - | + |
| $f_{xx}(x, y)$ | | + |
| | 不定 | 小 |
| $f(x, y)$ | | -1 |

と書ける. グラフの概形は図 3.3 を参照.

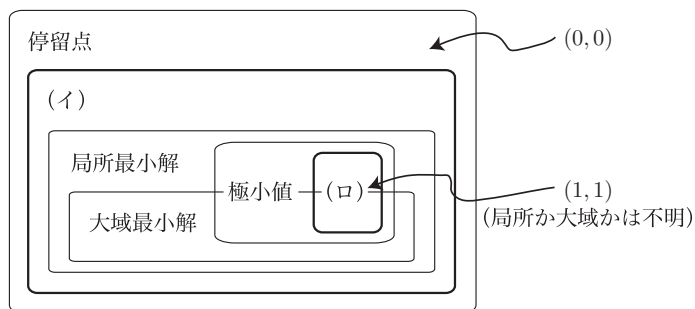


図 3.11: $x^3 - 3xy + y^3$ の停留点 $(0, 0)$, $(1, 1)$ の分類

(イ) $\nabla f(a, b) = \mathbf{0}$ かつ $\nabla^2 f(a, b)$ が半正定値

(ロ) $\nabla f(a, b) = \mathbf{0}$ かつ $\nabla^2 f(a, b)$ が正定値

例 18.

$$f(x, y) = x^3 + y^3 - 6x - 6y$$

の局所最適値を求めよ.

勾配ベクトルとヘッセ行列は

$$\nabla f(x, y) = \begin{bmatrix} 3x^2 - 6 \\ 3y^2 - 6 \end{bmatrix}, \quad \nabla^2 f(x, y) = \begin{bmatrix} 6x & 0 \\ 0 & 6x \end{bmatrix}$$

となる.

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

を解くと、停留点は $(x, y) = (\pm\sqrt{2}, \pm\sqrt{2})$ (順不同) となる. 符号と最適性の関係を調べると,

| (x, y) | $(\sqrt{2}, \sqrt{2})$ | $(-\sqrt{2}, -\sqrt{2})$ | $(\pm\sqrt{2}, \mp\sqrt{2})$ |
|----------------------|------------------------|--------------------------|------------------------------|
| $ \nabla^2 f(x, y) $ | + | + | - |
| $f_{xx}(x, y)$ | + | - | |
| | 小 | 大 | 不定 |
| $f(x, y)$ | $-8\sqrt{2}$ | $8\sqrt{2}$ | |

となる. よって, $(x, y) = (\sqrt{2}, \sqrt{2})$ のとき, 局所最小値 $-8\sqrt{2}$, $(x, y) = (-\sqrt{2}, -\sqrt{2})$ のとき, 局所最大値 $8\sqrt{2}$ をとる.

局所最適解は全て求まっているか?

さて, 例の解法で局所最適解が全て求まっているか確認してみよう. 特に例 18 を取り上げる. 例 18 の関数グラフは複雑な形であり, 図 3.12 はグラフの一部しか表していないので, 局所最適解に漏れがあるのではないか心配である.

まず, 停留点が 4 つ求まった. すると, 1 次の最適性条件より局所最適解は存在すれば全て, この 4 つの中に必ず含まれている. さらに, この中の $(\sqrt{2}, \sqrt{2})$, $(-\sqrt{2}, -\sqrt{2})$ がそれぞれ局所最小, 局所最大と判定され, $(\pm\sqrt{2}, \mp\sqrt{2})$ はヘッセ行列が不定であることから局所最適解ではないと判定される. したがって, $(\sqrt{2}, \sqrt{2})$, $(-\sqrt{2}, -\sqrt{2})$ は局所最適解であり, それら以外に局所最適解は存在しないと言えるのである.

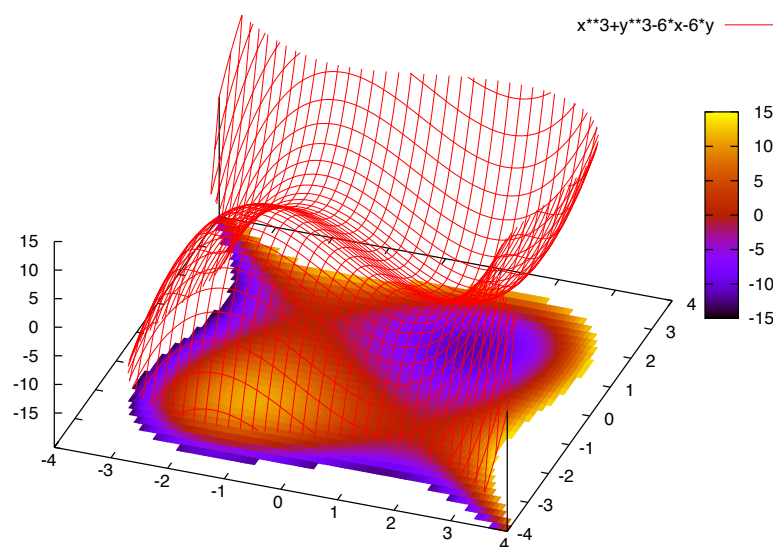


図 3.12: $x^3 + y^3 - 6x - 6y$ のグラフ

局所最適値と極値

3.1.1 節の終りで説明したように、極値は局所最適値の中でも特別なものである。しかし、上述の例で求めた解はすべて極値になっている（図 ?? を参照）。これは、例で取り上げた問題はすべての局所最適値が極値となっている問題だからである。計算のし易さから、この本では敢えてそのような問題のみを扱う（実際にはほとんどの本でこの方針をとっている）。

3.4 凸性と最適解

2 階微分の情報を使って 2 次の最適性条件を調べても、ある点 \bar{x} が局所最適解であることしかわからない。それでは大域最適解を見つけるにはどうすれば良いだろうか？実は関数が凸であるときは大域最小解を簡単に見つけられる。

定理 8. n 変数関数 f を凸関数とする。すると f の停留点は大域最小解になる。

証明. 停留点を $a \in \mathbb{R}^n$ とする。 f は凸なので、任意の $x \in \mathbb{R}^n$ に対して、

$$f(x) \geq f(a) + \nabla f(a)(x - a)$$

が成り立つ。いま $\nabla f(a) = \mathbf{0}$ なので、 $f(x) \geq f(a)$ が成り立つ。これは a が f の大域最小解であることを意味する。 \square

凸関数では停留点と大域最小解が一致してしまう。これが最適化問題における凸関数の最も重要な点である。

停留点 = 局所最小解 = 大域最小解

図 3.13: 凸関数の最小解と停留点の関係

練習問題 2. 以下の関数の局所最適解と局所最適値を求めよ.

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

$$(3) f(x, y) = x^3 + y^3 - 9xy + 1$$

3.5 定理 7 の証明

証明. (必要性) (a, b) を局所最小解とする. 任意の $(x, y) \in \mathbb{R}^2$ に対して, $h(t) = f((a, b) + t(x, y))$ とおく. すると

$$h(t) \geq h(0) \quad (\text{十分小さな } t) \quad (3.2)$$

が成り立つ. いま, $t = 0$ に関するテーラー展開をすると,

$$h(t) = h(0) + h'(0)t + h''(0)t^2 + o(t^2) \quad (3.3)$$

を得る. ここで, 1 次の最適性条件より

$$h'(0) = \nabla f(a, b) \cdot \begin{bmatrix} x \\ y \end{bmatrix} = 0$$

なので, (3.2) と (3.3) より

$$h''(0)t^2 + o(t^2) \geq 0$$

を得る. さらに, 両辺を t^2 で割り $t \rightarrow 0$ とすると, $o(t^2)/t^2 \rightarrow 0$ となるので,

$$h''(0) = [x \ y] \nabla^2 f(a, b) \begin{bmatrix} x \\ y \end{bmatrix} \geq 0 \quad (3.4)$$

が成り立つ. いま, $h(t)$ の定義に用いた (x, y) は任意であるので, 不等式 3.4 は任意の (x, y) について成り立つことになる. したがって $\nabla^2 f(a, b)$ は半正定値である.

(十分性) (a, b) を停留点, $\nabla^2 f(a, b)$ を正定値とし, その最小固有値を λ とする. ここで, 定理 ?? より, $\lambda > 0$ となる. いま, 2 変数関数に対するテーラー展開より,

$$f((a, b) + (x, y)) = f(a, b) + \nabla f(a, b) \cdot \begin{bmatrix} x \\ y \end{bmatrix} + \frac{1}{2} [x \ y] \nabla^2 f(a, b) \begin{bmatrix} x \\ y \end{bmatrix} + o(\|(x, y)\|^2)$$

が成り立つ。ここで,

$$\lim_{\delta \rightarrow 0^+} \frac{o(\delta^2)}{\delta^2} = 0$$

より, 小さな数 ε で $0 < \varepsilon < \lambda/2$ を満たすものに対して, $(0, 0)$ に充分近い (x, y) ならば ($\delta = \|(x, y)\|$ として),

$$|o(\|(x, y)\|^2)| \leq \varepsilon \|(x, y)\|^2$$

が成り立つ (左辺は絶対値)。よって,

$$\left| f((a, b) + (x, y)) - \left\{ f(a, b) + \nabla f(a, b) \cdot \begin{bmatrix} x \\ y \end{bmatrix} + \frac{1}{2} [x \ y] \nabla^2 f(a, b) \begin{bmatrix} x \\ y \end{bmatrix} \right\} \right| < \varepsilon \|(x, y)\|^2$$

を得る。さらに (a, b) が停留点であることから,

$$f((a, b) + (x, y)) > f(a, b) + \frac{1}{2} [x \ y] \nabla^2 f(a, b) \begin{bmatrix} x \\ y \end{bmatrix} - \varepsilon \|(x, y)\|^2$$

となる。ここで,

$$u = \frac{1}{\|(x, y)\|} \begin{bmatrix} x \\ y \end{bmatrix}$$

とおくと, $\|u\| = 1$ なので, 定理 ?? より

$${}^t u \nabla^2 f(a, b) u \geq \lambda,$$

$$[x \ y] \nabla^2 f(a, b) \begin{bmatrix} x \\ y \end{bmatrix} \geq \lambda \|(x, y)\|^2$$

が成り立つことがわかる。よって上式と ε の選び方より,

$$f((a, b) + (x, y)) > f(a, b) + \left(\frac{1}{2} \lambda - \varepsilon \right) \|(x, y)\|^2 > f(a, b)$$

を得る。いま (x, y) は $(0, 0)$ に充分近い任意の点なので, これは (a, b) が局所最小解であることを表す。

(否定) (a, b) が局所最適解ならば, ヘッセ行列 $\nabla^2 f(a, b)$ は半正定値か半負定値になる。対偶をとると, ヘッセ行列が不定値の場合は, 局所最適解ではない (定理 7 の図も参照)。 \square

まとめ

| |
|---------------|
| 最適化 $f(x, y)$ |
| 制約 なし |

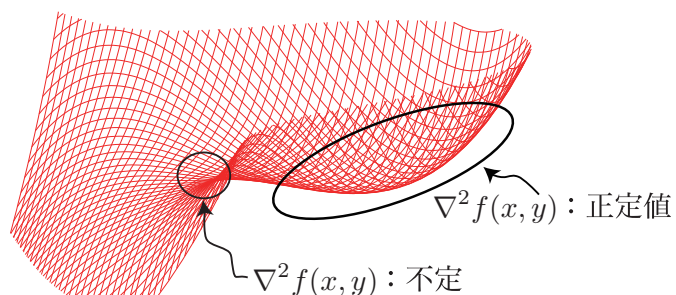
Step 1 停留点を求める

連立方程式 $\begin{cases} f_x(x, y) = 0 \\ f_y(x, y) = 0 \end{cases}$ の解 $(x, y) = (a, b)$ を求める

Step 2 $\nabla^2 f(a, b)$ の正値性を調べる

$\nabla^2 f(a, b)$: 正定値, 又は負定値 $\implies (a, b)$: 局所最適解

$\nabla^2 f(a, b)$: 不定 $\implies (a, b)$: 最適解でない



停留点 $\nabla f(a, b) = \mathbf{0}$

(イ)

局所最小解

極小値

(ロ)

大域最小解

(イ) $\nabla f(a, b) = \mathbf{0}$ かつ $\nabla^2 f(a, b)$ が半正定値

(ロ) $\nabla f(a, b) = \mathbf{0}$ かつ $\nabla^2 f(a, b)$ が正定値