

## 3 最適化問題

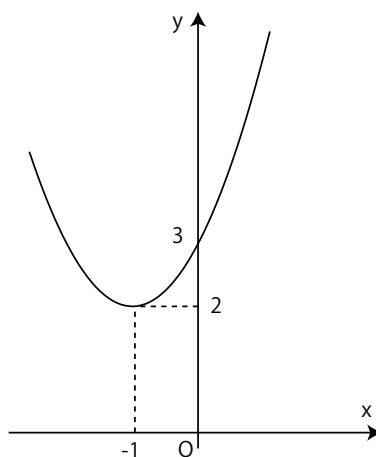
### 3.1 最適解

次の関数の最小値を探す.

$$\text{最小化 } f(x) = x^2 + 2x + 3$$

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

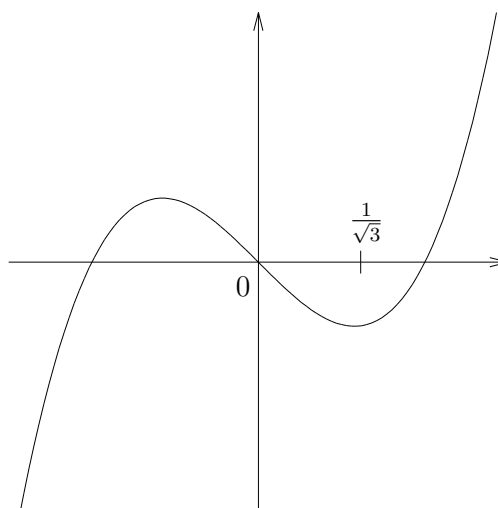
より, 最小値は 2, 最小解は  $x = -1$  である.  
この解は関数の形を考えればわかるように, すべての実数  $x$  に対して, 最小値を与える解になっていることがわかる. このような解を **大域最小解**, 値を **大域最小値** と呼ぶ.



一方, 次の最小化問題

$$\text{最小化 } f(x) = x^3 - x$$

については 大域最小解は存在しないが, 右のグラフから分かるように  $x = 1/\sqrt{3}$  に充分近いところではグラフは下に窪んでいる. 言い換えると, そこは局所的に最小になっているので,  $x = 1/\sqrt{3}$  を **局所最小解**, 値を **局所最小値** と呼ぶ.



一般の最小化問題に対して次の言葉を用意する.

$f$  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:

$$(P) \quad \begin{array}{l} \text{最小化} \quad f(x) \\ \text{制約} \quad g(x) \in C \end{array}$$

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  $f(x)$  を 目的関数 と呼ぶ。

ここで、最小解というものを改めて定義しておく。

**定義 3.1.**  $\bar{x} \in C$  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。

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、より見つけ易い解を探すことを目標とする。

**定義 3.2.**  $\bar{x} \in C$  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。

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、『最適化問題』、『局所最適値』、『局所最適解』などと呼ぶ。

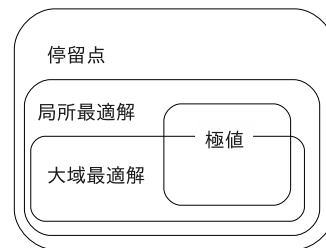
**注意**

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、 $f(x) > f(\bar{x})$  が成り立つとき、 $f$  は  $\bar{x}$  で極小値をとると言う。また、 $f(x) < f(\bar{x})$  が成り立つとき、 $f$  は  $\bar{x}$  で極大値をとると言う。二つを総称して、極値と呼ぶ。

局所最小値は等号も入れた『 $f(x) \geq f(\bar{x})$ 』で定義されるので、広義の極小値と同じである。同様に局所最適値は広義の極値と同じである。

この本では主に局所最小値を扱う。極小値と局所最小値が異なる例を以下に挙げる。



例 3.1 (極小値にならない最小値).  $f(x, y) = 4x^2 + 4xy + y^2$  を最小化する問題を考える.  $f(x, y) = (2x + y)^2$  より, すべての  $(x, y)$  に対して  $f(x, y) \geq 0$  となる. また  $f(x, y) = (2x + y)^2 = 0 \Leftrightarrow 2x + y = 0$  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,  $(x, y) = (t, -2t)$  という点の上では,  $f$  の値は  $f(t, -2t) = (2t - 2t)^2 = 0$  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.

## 3.2 制約無し最適化問題

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から勉強しよう:

$$\text{最小化 } f(x)$$

制約無し最小化問題の解析はこれからの理論の基礎になる.

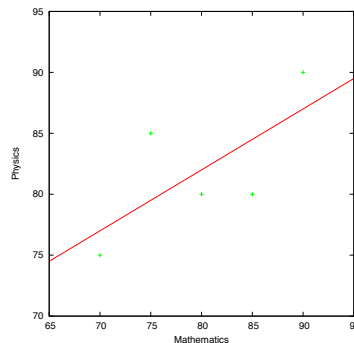
また制約の無い最小化でも重要な応用例は多くある.

例 3.2 (最小二乗法).

あるクラスの生徒 5 人の数学と物理の点数が以下ようになった.

数学	70	90	80	85	75
物理	75	90	80	80	85

数学と物理の点数がどのような関係を持っているか調べるために, これら点のなるべく近くを通る直線を求める. 直線を  $y = ax + b$  とおと, 点  $(60, 75)$  とこの直線との誤差は  $75 - (60a + b)$  と書ける (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制約無し最小化問題になる.



$$\begin{aligned} \text{最小化 } f(a, b) = & \{75 - (70a + b)\}^2 + \{90 - (90a + b)\}^2 + \{80 - (80a + b)\}^2 \\ & + \{80 - (85a + b)\}^2 + \{85 - (75a + b)\}^2 \end{aligned}$$

最小解を求めると,  $(a, b) = (1/2, 42)$  となるので, 最小二乗法で誤差を最小にする直線は  $y = \frac{1}{2}x + 42$  になる.

### 3.2.1 一次最適性必要条件

一変数関数  $f$  の場合は  $\bar{x}$  が局所最小値のとき,  $f'(\bar{x}) = 0$  が成り立つ. 多

変数関数の場合は以下が成り立つ.

定理 3.1 (一次の最適性必要条件).  $\bar{x}$  が局所最適解ならば,

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

が成り立つ.

定義 3.3.  $\nabla f(\bar{x}) = 0$  の時,  $\bar{x}$  を  $f$  の停留点と呼ぶ.

### 3.2.1.1 1 次最適性必要条件の直感的意味

2 変数関数の場合を説明する. テーラー展開より,

$$f((\bar{x}, \bar{y}) + (u, v)) = f(\bar{x}, \bar{y}) + \nabla f(\bar{x}, \bar{y})(u, v) + o(\|(u, v)\|)$$

となる. ここで,  $\nabla f(\bar{x}, \bar{y})(u, v) = f_x u + f_y v$  であった. 充分小さな数  $\delta > 0$  を使って, 特に  $(u, v) = (-\delta f_x, -\delta f_y)$  とおくと,

$$f((\bar{x}, \bar{y}) + \delta(u, v)) \approx f(\bar{x}, \bar{y}) - \delta(f_x^2 + f_y^2)$$

となる. (ここで  $\approx$  は”左辺は右辺に近い” ということを表す)

これより, もし,  $\nabla f(\bar{x}, \bar{y}) \neq 0$  なら右辺第二項は負になり, 右辺の値は  $f(\bar{x}, \bar{y})$  より小さい. よって, 点  $(\bar{x}, \bar{y})$  を  $(-\delta f_x, -\delta f_y)$  方向に少し動かせば,  $f$  の値が  $f(\bar{x}, \bar{y})$  より小さい点が見つかりそうである.

従って,  $(\bar{x}, \bar{y})$  が局所最適解になるには,  $\nabla f(\bar{x}, \bar{y}) = (0, 0)$  でないと駄目そうだと予想ができる. また,  $\nabla f(\bar{x}, \bar{y}) = (0, 0)$  のとき,  $(\bar{x}, \bar{y})$  の周りでは  $f(x, y)$  のグラフは”谷底”, “山の頂点” あるいは”平ら” になっていそう (下図参照).

### 3.2.1.2 定理 3.1 の証明

2 変数の場合に, 定理 3.1 の証明を挙げる.

$(x, y)$  のノルムを  $\|(x, y)\| = \sqrt{x^2 + y^2}$  とおく.  $(\bar{x}, \bar{y})$  を局所最小解とする. 定義より,  $(\bar{x}, \bar{y})$  に充分近い  $(x, y)$  に対して

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

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

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

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,

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

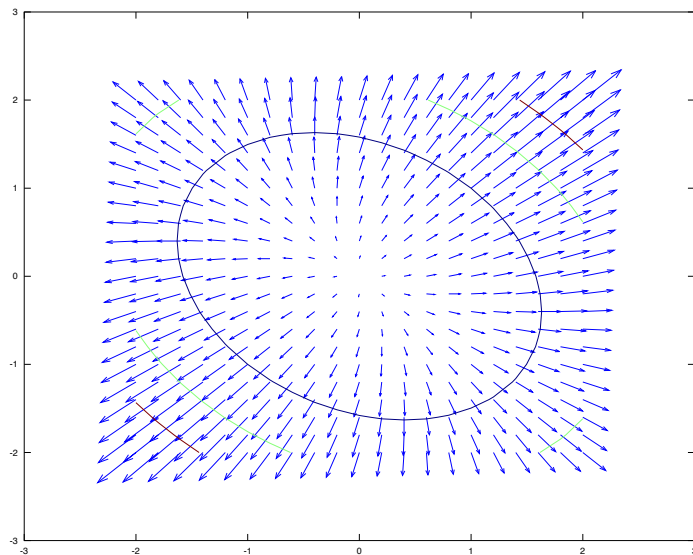


図 3.1  $f = 2x^2 + xy + 2y^2$  の等高線と勾配ベクトル

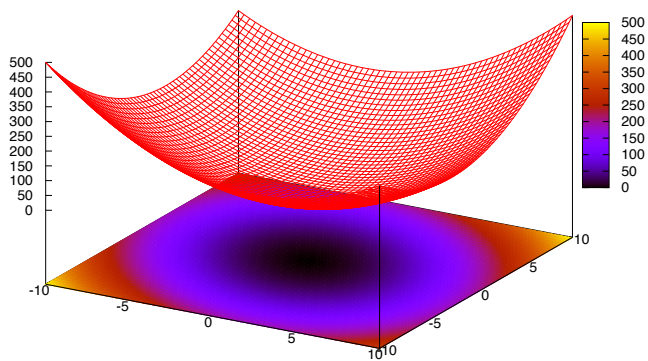


図 3.2  $f = 2x^2 + xy + 2y^2$  のグラフ

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、 $h'(t) = \frac{d}{dt} f(\bar{x} + tu, \bar{y} + tv) = f_x(\bar{x} + tu, \bar{y} + tv)u + f_y(\bar{x} + tu, \bar{y} + tv)v$  より、

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

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、 $\nabla f(\bar{x}, \bar{y}) = (0, 0)$  が成り立つ。

問題 3.1. 次の関数の停留点を求めよ。

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

### 3.2.1.3 停留点が局所最適か?

停留点は局所最適解とは限らない。

例 3.3.

$$\text{最小化 } f(x, y) = x^2 - y^2$$

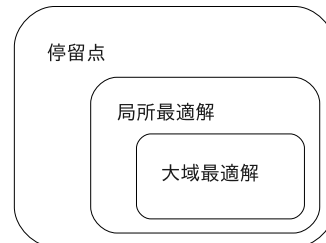
では  $\nabla f(0, 0) = (0, 0)$  となるが、局所最適解ではない。

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。

$(\bar{x}, \bar{y})$  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、 $f(x', y') \geq f(\bar{x}, \bar{y})$  ならば、局所最小解である
2. すべての動かし方で、 $f(x', y') \leq f(\bar{x}, \bar{y})$  ならば、局所最大解である
3. 動かし方により、 $f(x', y')$  が  $f(\bar{x}, \bar{y})$  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。

図 3.2.1.3 の  $f(x, y) = x^2 - y^2$  の場合、停留点  $(0, 0)$  から  $x$  だけ動かすと、 $f$  の値が大きくなり、 $y$  だけ動かすと  $f$  の値が小さくなるので、(3) の場合に当



もちろん最適解が存在しない場合もある。

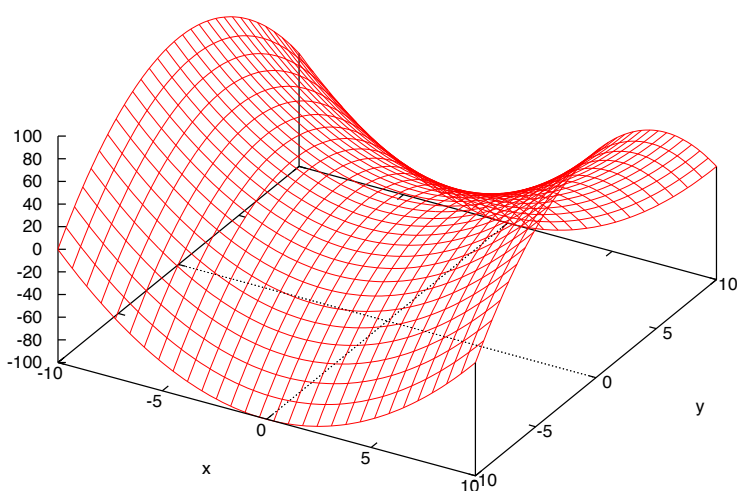


図 3.3  $x^2 - y^2$  のグラフ

てはまる。

停留点の中から局所最適解を探す際に、地道に  $f(x', y')$  の値を調べなければならぬこともあるが、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。

### 3.2.2 二次の最適性条件

我々の目標は局所最適解を求めることにある。しかし、1 階微分の情報を使って求まるのは停留点だけで、そのなかには局所最適解になっていないものもある。(例  $f(x) = x^3$  で、0 は停留点だが局所最適ではない) このままでは情報が足りないので、2 階微分の情報を使う。  $n$  変数関数  $f$  を最小化する問題を考える。

局所最適解を求めたい

→ 1 階微分を使って

停留点を求める

まだわからない

→ 2 階微分を使う

例 3.4.  $f(x, y) = x^3 - 3xy + y^3$  の局所最小解を探してみよう。下図を見ると、 $(x, y) = (1, 1)$  付近でグラフが下に窪んでいるのがわかる。そしてこの窪みの一番底が局所最小値になる。これを数学的な言い方で整理してみよう。まず、『局所的に下に窪んでいる』というのは『局所的に凸関数である』と言い換えられる。次に『窪みの一番底』というのは『勾配ベクトルが零ベクトル

ル』と言い換えられる. 実際には以下の定理が成り立つ.

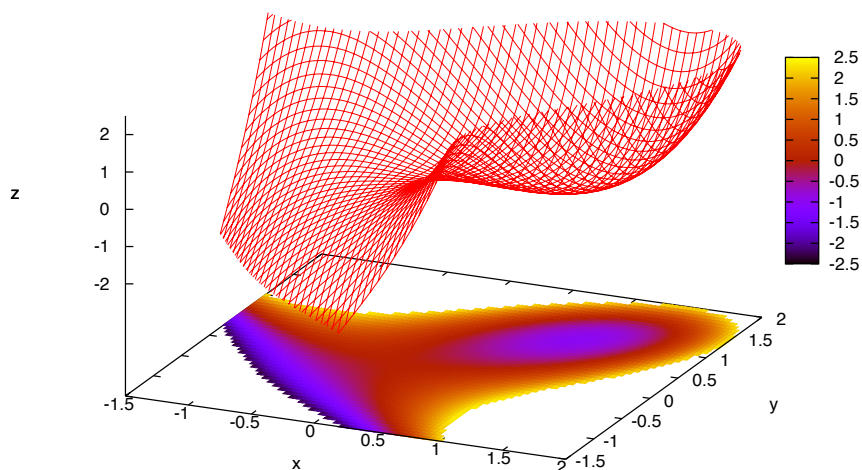


図 3.4  $z = x^3 - 3xy + y^3$  のグラフ

定理 3.2 (2 次の最適性条件).

(十分性)  $\nabla f(\bar{x}) = 0$  かつ  $\nabla^2 f(\bar{x})$  が正定値  $\Rightarrow \bar{x}$  は局所最小解

(必要性)  $\bar{x}$  が局所最小解である  $\Rightarrow \nabla f(\bar{x}) = 0$  かつ  $\nabla^2 f(\bar{x})$  が半正定値  
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, 負定値に置き換えたものが成り立つ.

(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.

解説. 半正定値行列の定義を見たときに非常に都合の良い条件に見えるかもしれないが, 実は関数が局所最小値をとっていれば, 自然に現れる性質である.

なので, 十分性にある正定値行列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のだ.

定理の証明は後ろの節で行う. □

### 3.2.2.1 2 次の最適性条件の幾何的イメージ

凸関数の概念を使うと, 2 次の最適性条件に幾何的なイメージをつけられ



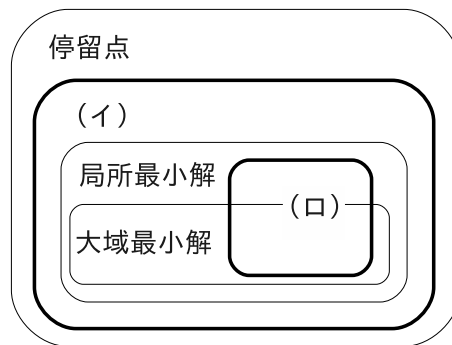
る。すべての  $(x, y)$  に対して、 $\nabla^2 f(x, y)$  が正定値であると  $f$  は凸関数になるのであった。

いま、 $\nabla^2 f(\bar{x}, \bar{y})$  が正定値で、 $(x, y)$  が  $(\bar{x}, \bar{y})$  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  $f_{xx}(\bar{x}, \bar{y}), f_{xy}(\bar{x}, \bar{y}), f_{yy}(\bar{x}, \bar{y})$  に近い。したがって、

$$\nabla^2 f(\bar{x}, \bar{y}) \text{ が正定値} \Rightarrow \nabla^2 f(x, y) \text{ が正定値} \Rightarrow f \text{ が「}(\bar{x}, \bar{y}) \text{ の近くで凸」}$$

従って、 $\nabla^2 f(\bar{x}, \bar{y})$  が正定値であり、さらに点  $(\bar{x}, \bar{y})$  が停留点ならば、その点は関数  $f$  の「局所的に凸」な部分の底にあるということを表している。

最適性と 2 次条件の関係



- (イ)  $\nabla f(\bar{x}, \bar{y}) = (0, 0)$  かつ  $\nabla^2 f(\bar{x}, \bar{y})$  が半正定値
- (ロ)  $\nabla f(\bar{x}, \bar{y}) = (0, 0)$  かつ  $\nabla^2 f(\bar{x}, \bar{y})$  が正定値

図の包含関係が示すように、 $\bar{x}$  が条件 (ロ) を満たさなくても最適解になることがある。実際、 $f(x, y) = x^4 + y^4$  の  $(0, 0)$  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。

### 3.2.3 最適性と凸性

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  $(\bar{x}, \bar{y})$  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適解を簡単に見つけられる。

**定理 3.3.**  $f: \mathbb{R}^n \rightarrow \mathbb{R}$  を凸関数とする。すると  $f$  の停留点は大域最小解になる。

*Proof.* 停留点を  $\bar{x}$  とする。  $f$  は凸なので、任意の  $y \in \mathbb{R}^n$  に対して、 $f(y) \geq f(\bar{x}) + \nabla f(\bar{x})(y - \bar{x})$  が成り立つ。いま  $\nabla f(\bar{x}) = 0$  なので、 $f(y) \geq f(\bar{x})$

が成り立つ. これは  $\bar{x}$  が  $f$  の大域最小解であることを意味する.

□

### 3.2.4 局所最適解の求め方

それでは 2 次の最適性条件を用いて例題を解いてみよう.

例 3.5.

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

の局所最適値を求めよ.

まず, 局所最適解であれば停留点になっているので, 方程式

$$\nabla f(x, y) = (2x + y - 9, x + 2y - 9) = (0, 0)$$

を解く. すると停留点は  $(x, y) = (3, 3)$  だと分かる. この点に対して 2 次の最適性条件を調べてみよう.

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. 固有値を求めると, 1, 3 となるので, ヘッセ行列は正定値である. よって,  $(3, 3)$  は局所最小解であり, 局所最小値は 0 となる. グラフの概形は下図を参照.

例 3.6.  $f(x, y) = x^3 - 3xy + y^3$

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

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

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

をえる. この解は  $(x, y) = (0, 0), (1, 1)$  になり, これが停留点になる.

次に, 2 次の最適性条件を調べる.

$(0, 0)$  ではヘッセ行列の行列式  $|\nabla^2 f(0, 0)|$  は  $|\nabla^2 f(0, 0)| < 0$  を満たすので, ヘッセ行列は不定値. よってこの点では局所最適ではない.

$(1, 1)$  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.

グラフの概形は図 3.4 を参照.

例 3.7.

$$\text{最小化 } f(x, y) = x^4 + y^4 - (x + y)^2$$

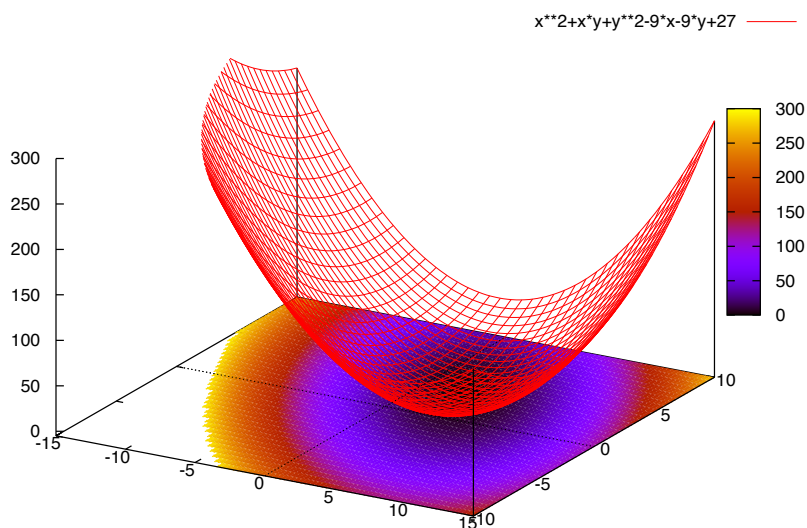


図 3.5  $x^2 + xy + y^2 - 9x - 9y + 27$  のグラフ.  $(x, y) = (3, 3)$  が停留点

の局所最適解を探し、そこでの値を求める.

まず、局所最適解であれば停留点になっているので、方程式

$$\nabla f(x, y) = (4x^3 - 2(x + y), 4y^3 - 2(x + y)) = (0, 0)$$

を解く. すると停留点は  $(x, y) = (0, 0), (1, 1), (-1, -1)$  だと分かる. この点に対して 2 次の最適性条件を調べてみよう.

ヘッセ行列を計算すると

$$\nabla^2 f(x, y) = \begin{bmatrix} 12x^2 - 2 & -2 \\ -2 & 12y^2 - 2 \end{bmatrix}$$

となる.

(i)点  $(1, 1), (-1, -1)$  で 2 次の条件を調べる.

関数  $f$  のヘッセ行列は

$$\nabla^2 f(1, 1) = \nabla^2 f(-1, -1) = \begin{bmatrix} 10 & -2 \\ -2 & 10 \end{bmatrix}$$

となる. 命題 2.2 を適用すると、この行列の 1 行 1 列の要素は 10 なので正. さらに行列式は  $10^2 - (-2)^2 > 0$  なので、ヘッセ行列は  $(1, 1), (-1, -1)$  で正定値になる. よって、2 次の最適性条件より、この二つの点は局所最小

解になり, その点での値は,  $f(1,1) = f(-1,-1) = -2$

補足.  $\nabla^2 f(1,1)$  の固有値は  $\begin{vmatrix} 10-\lambda & -2 \\ -2 & 10-\lambda \end{vmatrix} = 0$  を満たす  $\lambda$  であり,  $\lambda = 8, 12$  となる. よって固有値はすべて正であることが確かめられる.

(ii)点  $(0,0)$  で 2 次の条件を調べる.

関数  $f$  のヘッセ行列は,

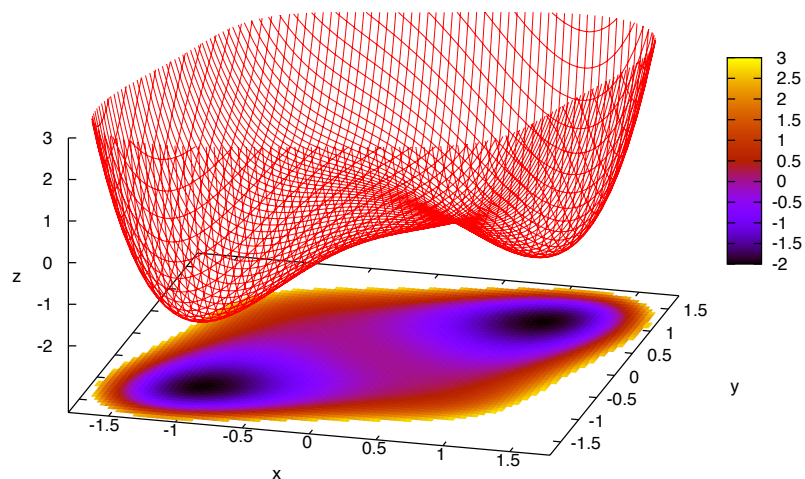
$$\nabla^2 f(0,0) = \begin{bmatrix} -2 & -2 \\ -2 & -2 \end{bmatrix}$$

となり, その行列式は 0 になり命題 2.2 では最適性の判定ができない. 固有値を計算してみると 0, 4 となる. よってヘッセ行列は点  $(0,0)$  で半正定値になる. この点では 2 次の最適性必要条件が成り立っているが, まだ局所最適解がそうでないか判断がつかない.

そこで地道に調べてみる.  $y = 0$  として  $x$  軸上のみで  $f$  の値の変化を調べる.  $f(x,0) = x^4 - x^2 = x^2(x^2 - 1)$  より充分小さい  $x$  で  $f$  は負の値をとる. 一方,  $y = -x$  という関係を満たす点での  $f$  の値の変化を調べると,  $f(x,-x) = x^4$  となり原点以外の点で正である. よって局所最小でも最大でもない.

まとめると,  $f$  は  $(1,1), (-1,-1)$  で局所最小値  $-2$  をとる. なお  $(0,0)$  は停留点だが局所最適解ではない.  $\square$

さて, いま解析した  $f(x,y) = x^4 + y^4 - (x+y)^2$  の概形は下の図のようになる.  $x, y$  どちらに関しても, どんどん大きな数を代入すれば  $f(x,y)$  の値がどんどん大きくなるのがわかる. また  $f(x,0)$  と同様に  $f(0,y)$  を調べると,  $y$  が 0 に近いとき,  $f$  の値が負になるのがわかる. 一方, 上で求めたように,  $f$  の停留点は 3 つしかなく,  $(1,1), (-1,-1)$  では  $f$  は  $-2$  まで窪んでいることもわかっている.



グラフを見ると確かに  $(1, 1), (-1, -1)$  は関数の局所的に凸な部分の底にあることがわかる。

問題 3.2. 以下の関数の局所最適解と局所最適値を求めよ。

(1)  $f(x, y) = x^3 - 3xy + y^3$    (2)  $f(x, y) = 3x^2 + 2xy + y^2$    (3)  $f(x, y) = x^3 + y^3 - 9xy + 1$

### 3.3 付録

#### 3.3.1 定理 3.2 の証明

証明. (必要性)  $(\bar{x}, \bar{y})$  を局所最小解とする. 任意の  $(u, v) \in R^2$  に対して,  $h(t) = f((\bar{x}, \bar{y}) + (u, v))$  とおく. すると充分小さな数  $t > 0$  に対して,  $h(t) \geq h(0)$  が成り立つ. いま, 0 に関するテーラー展開より,

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

となる. さらに, 1 次の最適性条件より,  $h'(0) = \nabla f(\bar{x}, \bar{y}) \cdot (u, v) = 0$  なので,

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

を得る. ここで, 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,  $h''(0) \geq 0$  が成り立つ. いま,  $h''(0) = \begin{bmatrix} u & v \end{bmatrix} \nabla^2 f(\bar{x}, \bar{y}) \begin{bmatrix} u \\ v \end{bmatrix}$  であり,

$h''(0) \geq 0$  は任意の  $(u, v)$  に対して成り立つので,  $\nabla^2 f(\bar{x}, \bar{y})$  は半正定値になる.

(十分性)  $(\bar{x}, \bar{y})$  が停留点で,  $\nabla^2 f(\bar{x}, \bar{y})$  を正定値とし, その最小固有値を  $\lambda$  とする. ここで, 定理 2.3 より,  $\lambda > 0$  となる. 2 変数に対するテーラー展開より, 小さな数  $\epsilon$  で  $0 < \epsilon < \lambda/2$  を満たすものに対しても,  $(0, 0)$  に充分近い  $(u, v)$  ならば,

$$\left| f((\bar{x}, \bar{y}) + (u, v)) - \left\{ f(\bar{x}, \bar{y}) + \nabla f(\bar{x}, \bar{y}) \cdot (u, v) + \frac{1}{2} \begin{bmatrix} u & v \end{bmatrix} \nabla^2 f(\bar{x}, \bar{y}) \begin{bmatrix} u \\ v \end{bmatrix} \right\} \right| < \epsilon \|(u, v)\|^2$$

が成り立つ (左辺は絶対値). さらに  $(\bar{x}, \bar{y})$  が停留点であることから, 特に

$$f((\bar{x}, \bar{y}) + (u, v)) > f(\bar{x}, \bar{y}) + \frac{1}{2} \begin{bmatrix} u & v \end{bmatrix} \nabla^2 f(\bar{x}, \bar{y}) \begin{bmatrix} u \\ v \end{bmatrix} - \epsilon \|(u, v)\|^2$$

を得る. ここで,  $t = \|(u, v)\|$  とおくと,  $\|t^{-1}(u, v)\| = 1$  なので, 定理 ?? を適用すると,

$$\begin{bmatrix} u & v \end{bmatrix} \nabla^2 f(\bar{x}, \bar{y}) \begin{bmatrix} u \\ v \end{bmatrix} \geq \lambda \|(u, v)\|^2$$

が成り立つことがわかる. よって上式と  $\epsilon$  の選び方より,

$$f((\bar{x}, \bar{y}) + (u, v)) > f(\bar{x}, \bar{y}) + \left( \frac{1}{2} \lambda - \epsilon \right) \|(u, v)\|^2 > f(\bar{x}, \bar{y})$$

を得る. いま  $(u, v)$  は  $(0, 0)$  に充分近い任意の点なので, これは  $(\bar{x}, \bar{y})$  が局所最小解であることを表す.

(否定)  $(\bar{x}, \bar{y})$  が局所最適解か局所最大解ならば, ヘッセ行列  $\nabla^2 f(\bar{x}, \bar{y})$  は半正定値か半負定値になるので, ヘッセ行列が不定値の場合は, そのどちらにもなっていない (定理 3.2 の図も参照).  $\square$