

非線形計画

第2章において、線形計画問題が最適解をもつ場合は、実行可能集合の端点、つまり境界で最小値もしくは最大値をとることを学んだ。一方、1変数関数の場合は、高等学校で学んだように、2次関数ならば頂点と境界での関数値を比較した。また、一般の関数の場合は微分して極大・極小を求め、極値と境界での関数値を比較した。それでは、多変数関数を目的関数とするような非線形計画問題ではどのように考えればよいだろうか。もし、制約条件がないのであれば、極大・極小を求めることができれば最適解・最適値の候補となる。制約条件がある場合には、境界で最適値をとる解を特徴づける必要がある。

そこで、まず、2変数関数の場合について考えてみる。線形数学では、行列の応用として偏微分を使った最適性の必要条件と十分条件が説明されている(例えば、文献[14]参照)。ここでは、各点における微分係数にあたる勾配ベクトルの値とその変化を表すヘッセ行列の定値性が重要となっている。しかし、これは制約条件のない非線形計画問題に対するものである。制約条件が与えられているときは、その実行可能集合を表す制約関数についての特徴づけも必要となってくる。そして、それらの特徴づけは一般の多変数関数の場合でも同様である。

本章では、多変数の偏微分の基礎事項および制約条件のない非線形計画問題に対する最適性条件の概要を述べ、不等式制約のある非線形計画問題に対する最適性の必要条件について説明する。その際に、ラグランジュ乗数も紹介する。また、等式制約のある非線形計画問題に対する最適性条件も本質は同じで、最適性の1次の必要条件はラグランジュ未定乗数法として知られている。なお、本書で解説する不等式制約のある問題に対する最適性条件は1次までとし、2次の最適性条件については[5], [11], [16], [18]などを参照されたい。

4.1 多変数関数の微分

一辺の長さが x である正方形の面積を y とすると、 $y = x^2$ という関係式が成立する。このとき、 y は x の関数 (function) であるといい、 $y = f(x)$ で表す。中学校や高等学校では、1 次関数 $f(x) = ax + b$ や 2 次関数 $g(x) = ax^2 + bx + c$ を学習している。

次に、正方形を長方形の場合にして考えてみる。例えば、各辺の長さが x と y である長方形の面積を z とすると、 $z = xy$ という関係式が成り立ち、 z は x と y の関数であることがわかる。このような関数を **2 変数関数** といい、一般に $z = f(x, y)$ と表す。同様に、 z が $x_1, x_2, \dots, x_n$ の関数になっているとき **多変数関数** といい、 $z = f(x_1, x_2, \dots, x_n)$ と表す。例えば、2 変数関数としての 1 次関数や 2 次関数は、

$$f(x, y) = c_1x + c_2y + d, \quad g(x, y) = a_1x^2 + 2bxy + a_2y^2 + c_1x + c_2y + d$$

となるが、ベクトルの内積や行列の積を用いて次のようにも表せる。

$$f(x, y) = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}^T \begin{pmatrix} x \\ y \end{pmatrix} + d,$$

$$g(x, y) = \begin{pmatrix} x \\ y \end{pmatrix}^T \begin{pmatrix} a_1 & b \\ b & a_2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}^T \begin{pmatrix} x \\ y \end{pmatrix} + d$$

さらに、 $\mathbf{c} = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}$, $\mathbf{x} = \begin{pmatrix} x \\ y \end{pmatrix}$, $A = \begin{pmatrix} a_1 & b \\ b & a_2 \end{pmatrix}$ というように、ベクトルと行列の記号を使って、次のように表すこともできる。

$$f(\mathbf{x}) = \mathbf{c}^T \mathbf{x} + d, \quad g(\mathbf{x}) = \mathbf{x}^T A \mathbf{x} + \mathbf{c}^T \mathbf{x} + d$$

ここで、 A は対称行列 ($A^T = A$) となっているが、 $a_{12} + a_{21} = 2b$ となるような

行列 $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ を選んでも、 $g(\mathbf{x}) = \mathbf{x}^T A \mathbf{x} + \mathbf{c}^T \mathbf{x} + d$ と表せる。

1 次形式と 2 次形式

対称行列 A で表された $\mathbf{x} = (x, y)^T$ に関する 2 次式 $\mathbf{x}^T A \mathbf{x} + \mathbf{c}^T \mathbf{x} + d$ において、1 次の項 $\mathbf{c}^T \mathbf{x}$ を **1 次形式** (linear form) といったように、2 次の項 $\mathbf{x}^T A \mathbf{x}$ を (x と y に関する) **2 次形式** (quadratic form) という。

なお, $3x^2 - 2xy + 3y^2$ のような展開された式も 2 次形式という.

一般に, 対称行列とは限らない n 次の実行列 Q を使用して, ($\mathbb{R}^n$ で定義された実数値) **多変数 2 次関数** f を表すと次のようになる.

$$f(\mathbf{x}) = \frac{1}{2} \mathbf{x}^T Q \mathbf{x} + \mathbf{c}^T \mathbf{x} + d \quad (4.1)$$

ここで, Q の第 i 行, 第 j 列の成分を a_{ij} とし, 実ベクトル $\mathbf{c}$, $\mathbf{x}$ の第 i 成分をそれぞれ c_i , x_i と表すと次のようになる.

$$f(\mathbf{x}) = \frac{1}{2} \sum_{i=1}^n x_i \left(\sum_{j=1}^n a_{ij} x_j \right) + \sum_{i=1}^n c_i x_i + d,$$

あるいは

$$f(\mathbf{x}) = \frac{1}{2} \sum_{i,j=1}^n x_i a_{ij} x_j + \sum_{i=1}^n c_i x_i + d$$

2 変数関数 $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ において, 点 (x, y) が限りなく点 (a, b) に近づくとき, どのような近づき方でも, $f(x, y)$ の値が一定の実数 α に限りなく近づくならば,

$$\left[(x, y) \rightarrow (a, b) \text{ のとき } f(x, y) \rightarrow \alpha \right] \quad \text{または} \quad \lim_{(x,y) \rightarrow (a,b)} f(x, y) = \alpha$$

と書いて, 「点 (x, y) が点 (a, b) に近づいたときの $f(x, y)$ の**極限 (極限值)** は α である」という. 1 変数のときと同じように, 「点 (x, y) が点 (a, b) に近づいたときの $f(x, y)$ の極限が $f(a, b)$ である」とき, つまり,

$$\lim_{(x,y) \rightarrow (a,b)} f(x, y) = f(a, b)$$

となるとき, 関数 $f(x, y)$ は点 (a, b) において**連続である**という (点 $(x, y) = (a, b)$ で連続ともいう). また, すべての点で連続のとき, f は $\mathbb{R}^2$ **上で連続である**といい, f を**連続関数**という. なお, $\mathbb{R}^n$ 上で定義される多変数実数値関数 $f: \mathbb{R}^n \rightarrow \mathbb{R}$ に対しても 2 変数関数の場合と同様に連続性が定義できる.

次に, 2 変数関数の微分について考えてみよう. 点 $(x, y) = (a, b)$ の近くで定義される連続関数 $z = f(x, y)$ を考える.

偏微分可能性・偏微分係数

(1) 関数 f が点 $(x, y) = (a, b)$ で x に関して**偏微分可能**であるとは,

$$\lim_{h \rightarrow 0} \frac{f(a+h, b) - f(a, b)}{h}$$

が有限値として存在するときをいう。この極限値を

$$f_x(a, b) \quad \text{または} \quad \frac{\partial}{\partial x} f(a, b)$$

と表し、 $(x, y) = (a, b)$ における (x) に関する f の**偏微分係数**という。

(2) 関数 f が点 $(x, y) = (a, b)$ で y に関して**偏微分可能**であるとは、

$$\lim_{k \rightarrow 0} \frac{f(a, b+k) - f(a, b)}{k}$$

が有限値として存在するときをいう。この極限値を

$$f_y(a, b) \quad \text{または} \quad \frac{\partial}{\partial y} f(a, b)$$

と表し、 $(x, y) = (a, b)$ における (y) に関する f の**偏微分係数**という。

(3) 偏微分係数を求めることを**偏微分する**という。もし、すべての点 (x, y) で f が偏微分可能ならば、この (x, y) に対して $f_x(x, y)$ や $f_y(x, y)$ を対応させるとそれぞれ (x, y) の関数となるので、 f_x を f の x に関する**偏導関数**、 f_y を f の y に関する**偏導関数**という。つまり、次のように書くときは、右辺の極限値が有限で存在することを意味している。

$$f_x(x, y) = \lim_{h \rightarrow 0} \frac{f(x+h, y) - f(x, y)}{h},$$

$$f_y(x, y) = \lim_{k \rightarrow 0} \frac{f(x, y+k) - f(x, y)}{k}$$

ここで、「偏微分する」とは、特定の文字を変数と考え、それ以外の文字を定数(変数の係数)とみなして微分することをさす。例えば、 x について偏微分するとは、 x 以外をただの定数とみなして x の式と考えて微分することである。このときは、1 変数の基本的な公式がすべて成り立つ。つまり、 $(f(x) \pm g(x))' = f'(x) \pm g'(x)$ (複号同順) のように項別に微分を行ってかまわないし、1 変数で成立する公式は使ってもよい。

◆**演習 4.1.** 次の関数を偏微分せよ。つまり、 x に関する偏微分係数と y に関する偏微分係数を求めよ。

$$(1) f(x, y) = xy$$

$$(2) f(x, y) = x^2 + y^2$$

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

$$(4) f(x, y) = 5x^2 - 6xy + 5y^2 - 10x + 6y$$

2 階偏微分可能性・2 次の偏導関数

- (1) 関数 f が点 $(x, y) = (a, b)$ で **2 階偏微分可能**であるとは, f の偏導関数 f_x と f_y が再び偏微分可能であるときをいう. つまり,

$$(f_x)_x(a, b), \quad (f_x)_y(a, b), \quad (f_y)_x(a, b), \quad (f_y)_y(a, b)$$

が存在する場合である. そこで, これらを点 (a, b) における f の **2 階偏微分係数**といい,

$$f_{xx}(a, b), \quad f_{xy}(a, b), \quad f_{yx}(a, b), \quad f_{yy}(a, b)$$

と表すことにする. また, f_x と f_y がすべての点 (x, y) で偏微分可能ならば, $f_{xx}, f_{xy}, f_{yx}, f_{yy}$ を f の **2 次偏導関数** (2 階偏導関数) という.

- (2) 2 次偏導関数 $f_{xx}, f_{xy}, f_{yx}, f_{yy}$ は次のようにも表す.

$$\begin{aligned} f_{xx}(x, y) &= \frac{\partial^2}{\partial x^2} f(x, y), & f_{xy}(x, y) &= \frac{\partial^2}{\partial y \partial x} f(x, y), \\ f_{yx}(x, y) &= \frac{\partial^2}{\partial x \partial y} f(x, y), & f_{yy}(x, y) &= \frac{\partial^2}{\partial y^2} f(x, y) \end{aligned}$$

◆**演習 4.2.** 次の関数の 2 次偏導関数をすべて求めよ.

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

ここまで 2 変数関数についての偏微分について述べてきたが, $\mathbb{R}^n$ 上で定義される多変数実数値関数 $f: \mathbb{R}^n \rightarrow \mathbb{R}$ に対しても 2 変数関数の場合と同様に偏微分や 2 階偏微分が定義できる. それらを 1 次形式や 2 次形式の形で表現することによって, 関数の増減を調べることができる.

勾配ベクトルとヘッセ行列

多変数関数 $z = f(x_1, x_2, \dots, x_n)$ に対して, $\mathbf{x} = (x_1, x_2, \dots, x_n)$ を用いて $z = f(\mathbf{x})$ と表す. このとき, 関数 f が 2 階偏微分可能であるならば, f の**勾配ベクトル** (gradient vector) と**ヘッセ行列** (Hessian matrix) がそれぞれ次のように定義できる.

$$\nabla f(\mathbf{x}) = \begin{pmatrix} \frac{\partial f}{\partial x_1}(\mathbf{x}) \\ \vdots \\ \frac{\partial f}{\partial x_n}(\mathbf{x}) \end{pmatrix}, \quad \nabla^2 f(\mathbf{x}) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(\mathbf{x}) & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(\mathbf{x}) \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(\mathbf{x}) & \cdots & \frac{\partial^2 f}{\partial x_n^2}(\mathbf{x}) \end{pmatrix}$$

特に, f が 2 階連続的微分可能ならばヘッセ行列は (実) 対称行列となる.

例えば, 1 次関数 $f(x_1, \dots, x_n) = c_1x_1 + \dots + c_nx_n + d$ については,

$$f_{x_1}(x_1, \dots, x_n) = c_1, \quad \dots, \quad f_{x_n}(x_1, x_2, \dots, x_n) = c_n$$

であるので, $\mathbf{c} = (c_1, \dots, c_n)^T$, $\mathbf{x} = (x_1, \dots, x_n)^T$ と表せば, $\nabla f(\mathbf{x}) = \mathbf{c}$ となる. これは第1章(19ページ)ですでにみていて, 1 次関数のグラフは傾斜が一定である平面であることを意味している.

◆演習 4.3. 次の関数の勾配ベクトルとヘッセ行列を求めよ.

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

また, 1 変数の微分可能な関数の微分係数がそのグラフの接線の傾きを表していた. それでは, 2 変数や多変数の関数のグラフの接線の傾きに対応するものは何だろうか.

全微分可能 (微分可能)

f を $\mathbb{R}^n$ 上の実数値関数とする. f が $\mathbf{x}_0 \in \mathbb{R}^n$ で**全微分可能**(単に, 「微分可能」ともいう)であるとは, あるベクトル $\mathbf{a} \in \mathbb{R}^n$ が存在して, $\mathbf{x}_0$ の近くのすべての点 $\mathbf{x}$ に対して以下の式が成立するときをいう.¹⁾

$$f(\mathbf{x}) = f(\mathbf{x}_0) + \mathbf{a}^T(\mathbf{x} - \mathbf{x}_0) + o(\|\mathbf{x} - \mathbf{x}_0\|), \quad \lim_{t \rightarrow 0} \frac{o(t)}{t} = 0 \quad (4.2)$$

この $\mathbf{a}$ を f の点 $\mathbf{x}_0$ における**微分**といい, $f'(\mathbf{x}_0)$ または $Df(\mathbf{x}_0)$ と表す. f が $\mathbb{R}^n$ のすべての点で全微分可能であるとき, f は $\mathbb{R}^n$ で**全微分可能**(微分可能)であるという.

定理 4.1

f を $\mathbb{R}^n$ 上の実数値関数とすると, 次が成立する.

- (1) f が $\mathbb{R}^n$ で全微分可能ならば連続である.
- (2) f が $\mathbb{R}^n$ で全微分可能ならば, $\mathbf{x} = (x_1, \dots, x_n)^T$ の各 x_k ($k = 1, \dots, n$) について f は偏微分可能であって,

$$f'(\mathbf{x}) = \begin{pmatrix} \frac{\partial f}{\partial x_1}(\mathbf{x}) \\ \vdots \\ \frac{\partial f}{\partial x_n}(\mathbf{x}) \end{pmatrix}$$

1) ノルム $\|\cdot\|$ の定義については, 18 ページをみよ.

が成立する．すなわち， $f'(x) = \nabla f(x)$ となる．

注意 4.1. 関数 $f: \mathbb{R}^n \rightarrow \mathbb{R}$ が偏微分可能で，そのすべての偏導関数が連続であれば（このとき， f が連続的偏微分可能という）， f は全微分可能であることが知られている．このとき，(4.2) は次のように表せる．

$$f(x) = f(x_0) + \nabla f(x_0)^T (x - x_0) + o(\|x - x_0\|) \quad \text{ただし, } \lim_{t \rightarrow 0} \frac{o(t)}{t} = 0$$

よって， $d \in \mathbb{R}^n$ に対して $x = x_0 + \lambda d$ を考えると， λ が十分小さくなると x は x_0 に近づくため， $f(x) - f(x_0)$ の正負は $\nabla f(x)^T d$ の正負に左右されることがわかる．

4.1 節の問題

[1] (4.1) で与えられた n 変数の 2 次関数の勾配ベクトルとヘッセ行列を，行列 Q ，ベクトル c ， x を使って表せ．途中の式変形には，それぞれの成分を使用して変形の仕方がわかるように示せ．

4.2 非線形計画問題の最適性条件

まず，「制約条件がない最適化問題」，つまり，無制約最適化問題に対する最適性条件を考えてみる．目的関数や制約関数の変数 x の定義域が $\mathbb{R}^n$ の場合を考え，それぞれ実数値関数とする．つまり，目的関数を $f: \mathbb{R}^n \rightarrow \mathbb{R}$ とする．

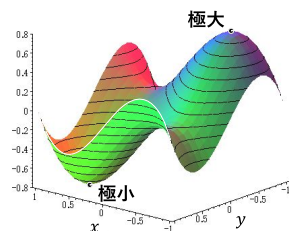


図 4.1 $f(x, y)$ の極大・極小

関数の極小と極大

多変数関数 $y = f(x)$ が点 $x = a_0$ の近く²⁾のすべての点 x に対して，

- (1) $f(x) \geq f(a_0)$ をみたす（関数のグラフが盆地になっている）とき，関数 $f(x)$ は点 a_0 で**極小値** (local minimum) をとる，または**極小**になる (minimal) という．
- (2) $f(x) \leq f(a_0)$ をみたす（関数のグラフが丘になっている）とき，関数 $f(x)$ は点 a_0 で**極大値** (local maximum) をとる，または**極大**になる (maximal) という．

極小値と極大値を総称して**極値** (extremum, local optimum) という．また，極値をとるときの点 a_0 を**局所的最適解** (local optimal solution) という．

関数の最小と最大

多変数関数 $y = f(\mathbf{x})$ が点 $\mathbf{x} = \mathbf{a}_0$ の近くだけでなく、すべての点 $\mathbf{x}$ で

- (1) $f(\mathbf{x}) \geq f(\mathbf{a}_0)$ をみたすとき、関数 $f(\mathbf{x})$ は点 $\mathbf{a}_0$ で**最小値** (global minimum) をとる、または**最小になる** (minimum) といい、
- (2) $f(\mathbf{x}) \leq f(\mathbf{a}_0)$ をみたすとき、関数 $f(\mathbf{x})$ は点 $\mathbf{a}_0$ で**最大値** (global maximum) をとる、または**最大になる** (maximum) という。

最小値と最大値を総称して**最適値** (optimal value, global optimum) という。また、最適値をとるときの点 $\mathbf{a}_0$ を**大域的最適解** (global optimal solution) という。

ここで、微分可能性の定義の (4.2) より、容易に次の定理が証明できる。

定理 4.2

$f: \mathbb{R}^n \rightarrow \mathbb{R}$ が $\mathbf{x}_0 \in \mathbb{R}^n$ において全微分可能ならば、 $\nabla f(\mathbf{x}_0)^T \mathbf{d} < 0$ となる $\mathbf{d} \in \mathbb{R}^n$ に対して、ある $\delta > 0$ が存在して、次が成立する。

$$0 < \lambda < \delta \text{ となるすべての } \lambda \text{ に対して } f(\mathbf{x}_0 + \lambda \mathbf{d}) < f(\mathbf{x}_0) \quad (4.3)$$

証明. f が全微分可能なので、注意 4.1 より、(4.2) は

$$\frac{f(\mathbf{x}_0 + \lambda \mathbf{d}) - f(\mathbf{x}_0)}{\lambda} = \nabla f(\mathbf{x}_0)^T \mathbf{d} + \frac{o(\lambda \|\mathbf{d}\|)}{\lambda \|\mathbf{d}\|} \|\mathbf{d}\|, \quad \text{ただし } \lim_{t \rightarrow 0} \frac{o(t)}{t} = 0$$

となる。ここで、 $\nabla f(\mathbf{x}_0)^T \mathbf{d} < 0$ が仮定されているので、 λ を十分小さくすると、右辺は負となる。よって、左辺も負であり、 $f(\mathbf{x}_0 + \lambda \mathbf{d}) < f(\mathbf{x}_0)$ が成立する。□

(4.3) をみたすような $\mathbf{d}$ を f の点 $\mathbf{x}_0$ における**減少方向ベクトル**という。

定理 4.2 を利用して以下の最適性条件が導ける。

定理 4.3: 最適性の 1 次の必要条件

多変数関数 $y = f(\mathbf{x})$ は、点 $\mathbf{a}_0$ の近くで連続的偏微分可能とする。 f が点 $\mathbf{a}_0$ で極値をとるならば、 $\nabla f(\mathbf{a}_0) = \mathbf{0}$ が成り立つ。これを**最適性の 1 次の必要条件** (first order necessary condition) という。

2) “近く”とは、正確には、点 $\mathbf{a}_0$ のある開近傍 $N(\mathbf{a}_0)$ が存在して、その要素 $\mathbf{x} \in N(\mathbf{a}_0)$ について考えるということである。

証明. f が点 $\mathbf{a}_0$ で極小値をとる場合のみ示す。 $\nabla f(\mathbf{a}_0) \neq \mathbf{0}$ と仮定すると、 $\mathbf{d} = -\nabla f(\mathbf{a}_0)$ とすることで、 $\nabla f(\mathbf{a}_0)^T \mathbf{d} < 0$ が成り立つ。 よって、 定理 4.2 により、 $\mathbf{a}_0$ のいくらでも近くに $f(\mathbf{a}_0 + \lambda \mathbf{d}) < f(\mathbf{a}_0)$ となる点 $\mathbf{a}_0 + \lambda \mathbf{d}$ が存在することになり、 仮定に矛盾する。 $\square$

次に、 式 (4.2) にならって、 関数 f の $\mathbf{x}_0$ における 2 次のテーラー展開を考え、 ヘッセ行列を使って f の 2 階微分可能性を定義する。

$$f(\mathbf{x}) = f(\mathbf{x}_0) + \nabla f(\mathbf{x}_0)^T (\mathbf{x} - \mathbf{x}_0) + \frac{1}{2} (\mathbf{x} - \mathbf{x}_0)^T \nabla^2 f(\mathbf{x}_0) (\mathbf{x} - \mathbf{x}_0) + o(\|\mathbf{x} - \mathbf{x}_0\|^2) \quad (4.4)$$

$$\text{ただし, } \lim_{t \rightarrow 0} \frac{o(t)}{t} = 0$$

このとき、 ヘッセ行列は対称行列となるので、 その 2 次形式について考えてみる。

2 次形式の符号 (正値・負値・半正値・半負値)

$\mathbf{x} = (x_1, \dots, x_n)^T$ についての 2 次形式 $\mathbf{x}^T A \mathbf{x}$ または行列 A は、 零ベクトルでないすべての $\mathbf{x}$ に対して、

- (1) $\mathbf{x}^T A \mathbf{x} > 0$ のとき、 **正値** (positive definite) (または**正定値**)
 - (2) $\mathbf{x}^T A \mathbf{x} < 0$ のとき、 **負値** (negative definite) (または**負定値**)
 - (3) $\mathbf{x}^T A \mathbf{x} \geq 0$ のとき、 **半正値** (positive semidefinite) (または**半正定値**)
 - (4) $\mathbf{x}^T A \mathbf{x} \leq 0$ のとき、 **半負値** (negative semidefinite) (または**半負定値**)
- であるという。

もちろん、 正値ならば半正値でもあるし、 負値ならば半負値でもある。 特に、 2 次の正方行列の場合には以下のような特徴づけができる。

定理 4.4

2 次の正方行列 $A = \begin{pmatrix} a & c \\ c & b \end{pmatrix}$ に対して、 その 2 次形式の符号について次が成立する。

- (1) A が正値であるための必要十分条件は、 $a > 0$ かつ $ab - c^2 > 0$ である。
- (2) A が負値であるための必要十分条件は、 $a < 0$ かつ $ab - c^2 > 0$ である。
- (3) A が半正値であるならば、 $a \geq 0$ かつ $ab - c^2 \geq 0$ である。

(4) A が半負値であるならば, $a \leq 0$ かつ $ab - c^2 \geq 0$ である.

証明. 任意の $x = \begin{pmatrix} x \\ y \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ に対して, 行列 A の2次形式の符号を考えてみる.

$$\begin{aligned} x^T A x &= (x \ y) \begin{pmatrix} a & c \\ c & b \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = (x \ y) \begin{pmatrix} ax + cy \\ cx + by \end{pmatrix} \\ &= ax^2 + 2cxy + by^2 \end{aligned} \quad (4.5)$$

$$= \begin{cases} a \left(x + \frac{c}{a}y\right)^2 + \left(b - \frac{c^2}{a}\right)y^2 & (a \neq 0) \\ b \left(y + \frac{c}{b}x\right)^2 - \frac{c^2}{b}x^2 & (a = 0 \text{ かつ } b \neq 0) \\ 2cxy & (a = 0 \text{ かつ } b = 0) \end{cases} \quad (4.6)$$

となるので, 同時に0とならない, どのような x と y についても2次形式 (4.5) が正値となるためには, $a > 0$ かつ $b - \frac{c^2}{a} > 0$, つまり, $ab - c^2 > 0$ でなければならないことがすぐにわかる. 同様に負値となるためには, $a < 0$ かつ $b - \frac{c^2}{a} < 0$, つまり, $ab - c^2 > 0$ でなければならない. さらに, これらの逆が成立することも (4.6) から容易にわかる.

また, 2次形式 (4.5) が半正値であるとする, $a = 0$ のときは (4.6) から $b \geq 0$ かつ $c = 0$ でなければならず, $ab - c^2 = 0$ となる. また, $a > 0$ のときは2次形式 (4.5) を x の2次式とみて, その判別式 D について $D/4 = c^2 - ab \leq 0$ となることから $ab - c^2 \geq 0$ が得られる. (4.5) が半負値の場合も同様である. $\square$

注意 4.2. 上記の命題の (3), (4) の対偶を考えると次が成立する.

- $a < 0$ または $ab - c^2 < 0$ ならば A は半正値でない.
- $a > 0$ または $ab - c^2 < 0$ ならば A は半負値でない.

したがって, $ab - c^2 < 0$ ならば A は半正値でも半負値でもない.

ここで, 2変数関数の勾配ベクトルとヘッセ行列は次のようであった.

$$\nabla f(x, y) = \begin{pmatrix} f_x(x, y) \\ f_y(x, y) \end{pmatrix}, \quad \nabla^2 f(x, y) = \begin{pmatrix} f_{xx}(x, y) & f_{xy}(x, y) \\ f_{yx}(x, y) & f_{yy}(x, y) \end{pmatrix}$$

定理 4.4 と注意 4.2 より, 2変数関数のヘッセ行列の定値性について偏微分係数を計算することで確認することが可能となる.

系 4.1: ヘッセ行列の定値性

2変数関数 $z = f(x, y)$ に対して, 点 $(x, y) = (a, b)$ におけるヘッセ行列の定値性について次が成立する.

(1) $\nabla^2 f(a, b)$ が正値であるための必要十分条件は,

$$f_{xx}(a, b) > 0 \text{ かつ } f_{xx}(a, b)f_{yy}(a, b) - f_{xy}(a, b)f_{yx}(a, b) > 0 \text{ である.}$$

(2) $\nabla^2 f(a, b)$ が負値であるための必要十分条件は,

$$f_{xx}(a, b) < 0 \text{ かつ } f_{xx}(a, b)f_{yy}(a, b) - f_{xy}(a, b)f_{yx}(a, b) > 0 \text{ である.}$$

(3) $\nabla^2 f(a, b)$ が半正値ならば

$$f_{xx}(a, b) \geq 0 \text{ かつ } f_{xx}(a, b)f_{yy}(a, b) - f_{xy}(a, b)f_{yx}(a, b) \geq 0 \text{ である.}$$

(4) $\nabla^2 f(a, b)$ が半負値ならば

$$f_{xx}(a, b) \leq 0 \text{ かつ } f_{xx}(a, b)f_{yy}(a, b) - f_{xy}(a, b)f_{yx}(a, b) \geq 0 \text{ である.}$$

(5) $f_{xx}(a, b)f_{yy}(a, b) - f_{xy}(a, b)f_{yx}(a, b) < 0$ ならば, $\nabla^2 f(a, b)$ は半正値でも半負値でもない.

さて, 多変数関数 $f: \mathbb{R}^n \rightarrow \mathbb{R}$ に対して, f が 2 階連続的偏微分可能であれば, (4.4) から 2 次の最適性条件を導ける.

定理 4.5: 最適性の 2 次の必要条件

多変数関数 $y = f(x)$ は, 点 a_0 の近くで 2 階連続的偏微分可能とする.

(1) f が点 a_0 で極小ならば, $\nabla f(a_0) = \mathbf{0}$ かつ $\nabla^2 f(x)$ が半正値となる.

(2) f が点 a_0 で極大ならば, $\nabla f(a_0) = \mathbf{0}$ かつ $\nabla^2 f(x)$ が半負値となる.

これらを**最適性の 2 次の必要条件** (second order necessary condition) という.

定理 4.6: 最適性の 2 次の十分条件

多変数関数 $y = f(x)$ は, 点 a_0 の近くで 2 階連続的偏微分可能とする.

(1) $\nabla f(a_0) = \mathbf{0}$ かつ $\nabla^2 f(x)$ が正値ならば, f は点 a_0 で極小値をとる.

(2) $\nabla f(a_0) = \mathbf{0}$ かつ $\nabla^2 f(x)$ が負値ならば, f は点 a_0 で極大値をとる.

これらを**最適性の 2 次の十分条件** (second order sufficient condition) という.

これらの最適性条件を利用して、多変数関数の極値の判定ができる。

○例 4.1. 2変数実数値関数 $f(x, y) = x^2 + y^2$ は常に非負の値をとり、 $x = y = 0$ のときのみ最小値 0 となる。よって、極小値も 0 である。このとき、それぞれの偏微分係数の値を計算すると

$$\nabla f(x, y) = \begin{pmatrix} f_x(x, y) \\ f_y(x, y) \end{pmatrix} = \begin{pmatrix} 2x \\ 2y \end{pmatrix}, \quad \nabla f(0, 0) = \begin{pmatrix} f_x(0, 0) \\ f_y(0, 0) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

$$\nabla^2 f(x, y) = \begin{pmatrix} f_{xx}(x, y) & f_{xy}(x, y) \\ f_{yx}(x, y) & f_{yy}(x, y) \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}, \quad \nabla^2 f(0, 0) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

となり、 $f_{xx}(0, 0) = 2 > 0$ かつ $f_{xx}(0, 0)f_{yy}(0, 0) - f_{xy}(0, 0)f_{yx}(0, 0) = 4 > 0$ が成立している。したがって、 $\nabla f(0, 0) = \mathbf{0}$ で $\nabla^2 f(0, 0)$ が正値なので、最適性の 2 次の十分条件より、『関数 f は点 $(x, y) = (0, 0)$ で極小値をとる』ことがわかるが、これは最初に述べたことに合致している。□

○例 4.2. 2変数実数値関数 $f(x, y) = xy$ は x と y が同じ符号の場合は常に正の値をとるが、異なる符号の場合は常に負の値をとる。よって、関数 f について、 $(x, y) = (0, 0)$ のいくらでも近いところに関数の値が負の場合と正の場合が起こりうることになり、点 $(x, y) = (0, 0)$ では f は極値をとらないことがわかる。このとき、関数 f の勾配ベクトルは

$$\nabla f(x, y) = \begin{pmatrix} f_x(x, y) \\ f_y(x, y) \end{pmatrix} = \begin{pmatrix} y \\ x \end{pmatrix}$$

となるので、 $\nabla f(x, y) = \mathbf{0}$ を解くと極値の候補は $(x, y) = (0, 0)$ となる。また、 $(x, y) = (0, 0)$ におけるヘッセ行列は次のように計算される。

$$\nabla^2 f(x, y) = \begin{pmatrix} f_{xx}(x, y) & f_{xy}(x, y) \\ f_{yx}(x, y) & f_{yy}(x, y) \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \nabla^2 f(0, 0) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

したがって、ヘッセ行列に関する最適性の 2 次の必要条件について考えてみると、 $f_{xx}(0, 0)f_{yy}(0, 0) - f_{xy}(0, 0)f_{yx}(0, 0) = -1 < 0$ となるので、系 4.1 の (5) により、 $\nabla^2 f(0, 0)$ は半正値でも半負値でもない。したがって、最適性の 2 次の必要条件の対偶により『関数 f が点 $(x, y) = (0, 0)$ において極値をとらない』ことがわかるが、これは最初に述べたことに合致している。□

さて次に、「制約条件のある場合の非線形計画問題」を考えてみよう。

$$(\text{NLP})^3 \quad \begin{cases} \text{Min} & f(\mathbf{x}) \\ \text{s.t.} & g_1(\mathbf{x}) \leq 0, g_2(\mathbf{x}) \leq 0, \dots, g_m(\mathbf{x}) \leq 0 \end{cases} \quad (4.7)$$

ここで, $f: \mathbb{R}^n \rightarrow \mathbb{R}$, $g_i: \mathbb{R}^n \rightarrow \mathbb{R}$ ($i = 1, \dots, m$) は 2 階連続的偏微分可能な実数値関数とする. また,

$$I(\mathbf{x}_0) = \{i \mid g_i(\mathbf{x}_0) = 0, 1 \leq i \leq m\}$$

を点 $\mathbf{x}_0$ において**アクティブ**な制約条件式の番号の集合とする. つまり, アクティブな制約⁴⁾とは, $\mathbf{x}_0$ が制約集合の境界となっている場合 ($\mathbf{x}_0$ において, その制約式が等号で成立している) を表していて, $\mathbf{x}_0$ から移動する方向によっては, 制約条件をみたさなくなることの意味している. このとき, 定理 4.2 の (4.3) を利用して次の結果が成立する.

定理 4.7

$f: \mathbb{R}^n \rightarrow \mathbb{R}$ と $g_i: \mathbb{R}^n \rightarrow \mathbb{R}$ ($i \in I(\mathbf{x}_0)$) が $\mathbf{x}_0$ で連続的偏微分可能とし, また, 残りの g_j ($j \notin I(\mathbf{x}_0)$) が $\mathbf{x}_0$ で連続であるとする. このとき, $\mathbf{x}_0$ が問題 (NLP) の局所的最適解ならば,

$$\nabla f(\mathbf{x}_0)^T \mathbf{d} < 0, \quad \nabla g_i(\mathbf{x}_0)^T \mathbf{d} < 0 \quad (i \in I(\mathbf{x}_0)) \quad (4.8)$$

が同時に成り立つベクトル $\mathbf{d} \in \mathbb{R}^n$ は存在しない.

よって, 定理 4.7 の結果にゴルダンの二者択一定理 (定理 3.6) を利用すると, 次の 1 次の最適性の必要条件が得られる.

定理 4.8: フリッツ-ジョンの条件

$f: \mathbb{R}^n \rightarrow \mathbb{R}$ と $g_i: \mathbb{R}^n \rightarrow \mathbb{R}$ ($i \in I(\mathbf{x}_0)$) が $\mathbf{x}_0$ で連続的偏微分可能とし, また, 残りの g_j ($j \notin I(\mathbf{x}_0)$) が $\mathbf{x}_0$ で連続であるとする. このとき, $\mathbf{x}_0$ が問題 (NLP) の局所的最適解ならば, すべてが 0 ではない実数 u_0, u_i ($i \in I(\mathbf{x}_0)$) が存在して, 以下の条件が成立する.

$$u_0 \nabla f(\mathbf{x}_0) + \sum_{i \in I(\mathbf{x}_0)} u_i \nabla g_i(\mathbf{x}_0) = \mathbf{0}, \quad (4.9)$$

$$u_0 \geq 0, u_i \geq 0 \quad (i \in I(\mathbf{x}_0)) \quad (4.10)$$

これを問題 (NLP) に対する**フリッツ-ジョン** (Fritz John) の**最適性必要条件**という.

3) (NLP) は “**N**on**L**inear programming **P**roblem” の頭文字である.

4) 有効制約ともいう.

証明. $\nabla f(\mathbf{x}_0)$ を1列目におき, 2列目から $i \in I(\mathbf{x}_0)$ について順番に $\nabla g_i(\mathbf{x}_0)$ を並べた行列 A を考えると, 定理 4.7 の結果から $A^T \mathbf{d} < 0$ をみたす $\mathbf{d} \in \mathbb{R}^n$ が存在しないことになる. ここで, ゴルダンの二者択一定理を利用すれば, 定理の結果が得られる. $\square$

注意 4.3. $\mathbf{x}_0$ でアクティブでない制約式 g_j ($j \notin I(\mathbf{x}_0)$) に対応する係数を $u_j = 0$ とすれば, $u_j \nabla g_j(\mathbf{x}_0) = \mathbf{0}$ とすることができる. そこで, $\mathbf{u} = (u_1, \dots, m)^T$ とおくと, フリッツ・ジョンの最適性必要条件は以下のようにも表すことができる.

$$u_0 \nabla f(\mathbf{x}_0) + \sum_{i=1}^m u_i \nabla g_i(\mathbf{x}_0) = \mathbf{0}, \quad (4.11)$$

$$(u_0, \mathbf{u}) \neq \mathbf{0}, (u_0, \mathbf{u}) \geq \mathbf{0},$$

$$u_j g_j(\mathbf{x}_0) = 0 \quad (j = 1, \dots, m) \quad (4.12)$$

この (4.12) を**相補性条件** (complementarity condition) という.

さらに条件を追加すると, 次の定理も成立する.

定理 4.9: カルシュ・クーン・タッカー条件

上記の定理と同じ仮定のもとで, $\mathbf{x}_0$ が問題 (NLP) の局所的最適解ならば, $\{\nabla g_i(\mathbf{x}_0) | i \in I(\mathbf{x}_0)\}$ が1次独立であるとき, 実数 u_i ($i \in I(\mathbf{x}_0)$) が存在して, 以下の条件が成立する.

$$\nabla f(\mathbf{x}_0) + \sum_{i \in I(\mathbf{x}_0)} u_i \nabla g_i(\mathbf{x}_0) = \mathbf{0}, \quad (4.13)$$

$$u_i \geq 0 \quad (i \in I(\mathbf{x}_0)) \quad (4.14)$$

これを問題 (NLP) に対する**カルシュ・クーン・タッカー** (Karush-Kuhn-Tucker) の**最適性必要条件**という.

証明. 定理 4.8 の結果に1次独立制約想定を利用する. $\square$

注意 4.4. 上記の「 $\{\nabla g_i(\mathbf{x}_0) | i \in I(\mathbf{x}_0)\}$ が1次独立」を**1次独立制約想定** (constraint qualification) という. 本書では, これを単に「 $I(\mathbf{x}_0)$ が1次独立制約想定をみたす」ということにする. また, カルシュ・クーン・タッカーの最適性必要条件における目的関数と制約関数の勾配ベクトルの関係式 (4.13) と (4.14) は

$$-\nabla f(\mathbf{x}_0) = \sum_{i \in I(\mathbf{x}_0)} u_i \nabla g_i(\mathbf{x}_0), \quad u_i \geq 0 \quad (i \in I(\mathbf{x}_0)) \quad (4.15)$$

と変形できるので, $\mathbf{x}_0$ での目的関数の勾配ベクトルの (-1) 倍が制約関数の勾配ベクトルの非負1次結合で表されていることを示している. よって, 第2章の線形計画問題の実行可能集合の解釈で考えたように, (2.12) と同じ形で表すことができる.

$$-\nabla f(\mathbf{x}_0) \in \text{cc}\{\nabla g_i(\mathbf{x}_0) \mid i \in I(\mathbf{x}_0)\} \quad (4.16)$$

注意 4.5. (4.9), (4.11), (4.13) において, 関数の係数として使用されている u_i を **ラグランジュ乗数** (Lagrangian multiplier) といい, それらを成分とするベクトルを **ラグランジュ乗数ベクトル** という. (4.13) と (4.14) の場合も相補性条件 (4.12) を仮定することで次のように表せる.

$$\nabla f(\mathbf{x}_0) + \sum_{i=1}^m u_i \nabla g_i(\mathbf{x}_0) = \mathbf{0}, \quad (4.17)$$

$$u_j \geq 0, \quad u_j g_j(\mathbf{x}_0) = 0 \quad (j = 1, \dots, m) \quad (4.18)$$

このとき, $\mathbf{g}(\mathbf{x}) = (g_1(\mathbf{x}), \dots, g_m(\mathbf{x}))^T$ とすると, (4.17) の左辺は $\mathbf{x}$ と $\mathbf{u} = (u_1, \dots, u_m)^T$ の関数

$$L(\mathbf{x}, \mathbf{u}) = f(\mathbf{x}) + \mathbf{u}^T \mathbf{g}(\mathbf{x})$$

の勾配ベクトルとなっていることがわかる. この $L: \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}$ を **ラグランジュ関数** (Lagrangian function) という. 一般に, 等式制約をもつ最適化問題 (非線形計画問題) の解を求めるためにラグランジュ関数を考えて (4.17) をみたすラグランジュ乗数ベクトル $\mathbf{u}$ を求めることによって解 $\mathbf{x}_0$ を求める方法を **ラグランジュ未定乗数法** という. そのときは, $\mathbf{u} \geq \mathbf{0}$ である必要はない.

ここまで, 不等式制約のある非線形計画問題に対する最適性の必要条件を考えてきたが, 不等式制約と等式制約が混在した問題に対する最適性条件についてはもう少し議論が必要である. また, 制約のある問題に対する最適性の十分条件については, 実行可能集合やヘッセ行列に関する特徴づけが必要となる. そのためには, いろいろな数学的な準備が必要となるので, 等式制約が含まれた非線形計画問題に対する最適性条件や 2 次の最適性条件, および **制約想定** については [1], [5], [11] などを参照されたい.

○例 4.3. 非線形計画問題

$$(\text{NLP}) \quad \begin{cases} \text{Min} & f(\mathbf{x}) \\ \text{s.t.} & g_1(\mathbf{x}) \leq 0 \\ & g_2(\mathbf{x}) \leq 0 \\ & g_3(\mathbf{x}) \leq 0 \end{cases}$$

の制約関数の等高線が図 4.2 のようになっていて, 実行可能領域を S として, 2 点 $\mathbf{x}_0$ と $\mathbf{x}_1$ を選んでみる. このとき, どちらも実行可能解であり, それぞれのアクティブな制約式については, $I(\mathbf{x}_0) = \{1, 2\}$, $I(\mathbf{x}_1) = \{2, 3\}$ となっている. また, 目的関数の等高線が図 4.3 のようになっているとき, $\mathbf{x}_0$ は (NLP) の局所最適解であるが, $\mathbf{x}_1$ は局所最適解でない. 実際, $I(\mathbf{x}_0)$ と $I(\mathbf{x}_1)$ のどちら

も 1 次独立制約想定をみたしている．しかし，微分可能な場合は各点の勾配ベクトルがその点を通る等高線の接線に垂直になっていることに注意すると， x_0 では (4.13) と (4.14) をみたしていて，(4.16) が成立しているが， x_1 では成立していないことが容易にわかる． □

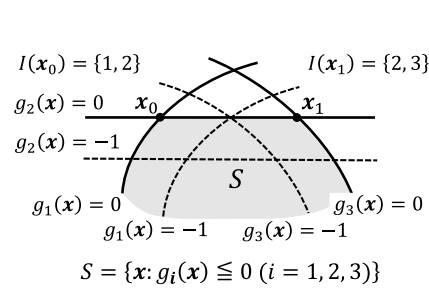


図 4.2 (NLP) の実行可能領域

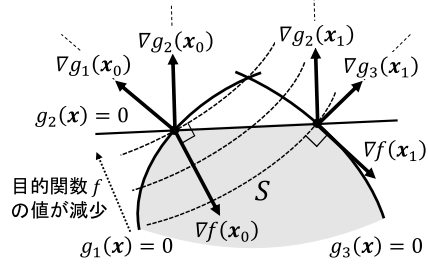


図 4.3 目的関数とアクティブな制約関数の勾配ベクトル

○例 4.4. (4.16) の状況を確認するために，具体的な非線形計画問題を考えよう．

(NLP) $\left\{ \begin{array}{l} \text{Min} \\ \text{s.t.} \end{array} \right.$

$\begin{cases} x^2 - y \\ x^2 - 2x + y - 1 \leq 0 \\ x + y - 1 \leq 0 \\ -y \leq 0 \end{cases}$

(4.19)

問題 (NLP) の目的関数を $f(x,y) = x^2 - y$ とおき，制約関数を

$$g_1(x,y) = x^2 - 2x + y - 1,$$
$$g_2(x,y) = x + y - 1, \quad g_3(x,y) = -y$$

とすると，(NLP) は (4.7) の形となっている．また，それぞれの関数は $\mathbb{R}^2$ で定義された実数値関数で微分可能である．それらの勾配ベクトルを求めてみると

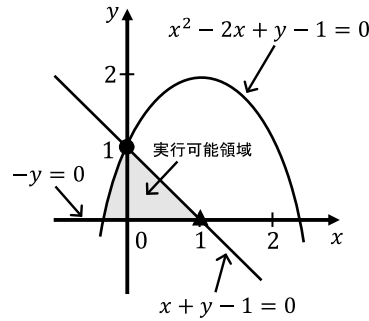


図 4.4 例 4.4 における (4.19) の実行可能領域

$$\nabla f(x,y) = \begin{pmatrix} 2x \\ -1 \end{pmatrix},$$

$$\nabla g_1(x,y) = \begin{pmatrix} 2x-2 \\ 1 \end{pmatrix}, \quad \nabla g_2(x,y) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \nabla g_3(x,y) = \begin{pmatrix} 0 \\ -1 \end{pmatrix}$$

となる．また， xy 平面にそれぞれの等高線を描き，実行可能領域を図示すると図 4.4 のようになる．

ここで, $\mathbf{x}_0 = (x, y) = (0, 1)$ と $\mathbf{x}_1 = (x, y) = (1, 0)$ において $I(\mathbf{x}_0) = \{1, 2\}$, $I(\mathbf{x}_1) = \{2, 3\}$ となり, $I(\mathbf{x}_0)$ と $I(\mathbf{x}_1)$ のどちらも 1 次独立制約想定をみたしている. また,

$$\nabla f(0, 1) = \begin{pmatrix} 0 \\ -1 \end{pmatrix}, \quad \nabla g_1(0, 1) = \begin{pmatrix} -2 \\ 1 \end{pmatrix}, \quad \nabla g_2(0, 1) = \begin{pmatrix} 1 \\ 1 \end{pmatrix},$$

$$\nabla f(1, 0) = \begin{pmatrix} 2 \\ -1 \end{pmatrix}, \quad \nabla g_2(1, 0) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \nabla g_3(1, 0) = \begin{pmatrix} 0 \\ -1 \end{pmatrix}$$

となるので, $\mathbf{x}_0 = (0, 1)$ では, $u_1 = \frac{1}{3}$, $u_2 = \frac{2}{3}$ で (4.13) と (4.14) をみたしている. 一方, $\mathbf{x}_1 = (1, 0)$ においては, (4.13) を u_1, u_2 について解くと, $u_1 = -2$, $u_2 = -3$ となり, (4.14) をみたしていない. 実際, $\mathbf{x}_0 = (0, 1)$ は最適解となり, $\mathbf{x}_1 = (1, 0)$ は最適解にならない. $\square$

4.2 節の問題

[1] 次の非線形計画問題の最適解と最適値を求めよ.

$$(\text{NLP}) \quad \begin{cases} \text{Min} & (x-1)^2 + (y-2)^2 \\ \text{s.t.} & 2x^2 - y \leq 0 \\ & -x^2 + y - 1 \leq 0 \\ & x \leq 0 \end{cases}$$

章末問題

[1] 2 変数実数値関数 $f(x, y) = x^3 - 3xy + y^3$ の極値を調べよ. ただし, 最適性条件のどれを利用して候補を絞り込んでいったかがわかるように述べよ.

[2] 2 変数実数値関数 $f(x, y) = 5x^2 - 6xy + 5y^2 - 10x + 6y$ の極値を調べよ. ただし, 最適性条件のどれを利用して候補を絞り込んでいったかがわかるように述べよ.

[3] 次の非線形計画問題について, 以下の各問いに答えよ.

$$(\text{NLP}) \quad \begin{cases} \text{Min} & (x+1)^2 + (y+2)^2 \\ \text{s.t.} & -x^2 + y \leq 0 \\ & 2x^2 - y \leq 1 \\ & x \geq 0 \end{cases}$$

(1) 問題 (NLP) の実行可能領域を xy 平面に描き, 最適解と最適値を求めよ.

(2) (1) で求めた最適解において, 1 次独立制約想定とカルーシュ・クーン・タッカーの最適性必要条件が成立していることを示せ.

(3) 最適解でない実行可能解を 1 つ取り上げ, カルーシュ・クーン・タッカーの最適性必要条件が成り立っていないことを示せ.