

最適化 復習

Kosuke Toda

目次

| | | |
|-------|-------------------------|----|
| 第 1 章 | まえがき | 2 |
| 第 2 章 | 数学的準備 | 3 |
| 2.1 | 諸定義 | 3 |
| 2.2 | 関数 | 4 |
| 2.3 | 凸集合 | 6 |
| 2.4 | 凸関数 | 7 |
| 第 3 章 | 最適化問題の定式化 | 9 |
| 3.1 | 最適化問題 | 9 |
| 3.2 | 用語 | 9 |
| 第 4 章 | 非線形計画問題 | 11 |
| 4.1 | 制約なし問題の最適性条件 | 11 |
| 4.2 | 等式制約問題の最適性条件 | 12 |
| 4.3 | 不等式制約問題の最適性条件 | 16 |
| 4.4 | 凸計画問題 | 20 |
| 第 5 章 | 線形計画問題 | 21 |
| 5.1 | 線形計画問題の例 | 21 |
| 5.2 | 線形計画問題の最適解 | 22 |
| 5.3 | 標準形の線形計画問題 | 22 |
| 5.4 | 基底解と最適解 | 23 |
| 第 6 章 | 証明 | 26 |
| 6.1 | 2 章の証明 | 26 |
| 6.2 | 4 章の証明 | 28 |
| | 参考文献 | 31 |

第 1 章

まえがき

様々な問題に対して，最も効率的になるように意思決定をする手法としてオペレーションズ・リサーチ (OR) というものがある [1]．これは，現実の問題を数理モデルに置換し，問題を解決する．最適化理論 (optimization theory) は，OR の基礎理論の 1 つであり，理論だけでなく，現実の問題を解くための手法を提供してきた [2]．本資料では，特に連続最適化に焦点を当てる．

第 2 章

数学的準備

2.1 諸定義

n 次元実数空間 \mathbb{R}^n を考える.

- ε -近傍
 - $B(x, \varepsilon) = \{\mathbf{y} \in \mathbb{R}^n \mid \|\mathbf{y} - \mathbf{x}\| < \varepsilon\}$
 - ただし, $\|\cdot\|$ はユークリッドノルムであり, $\|\mathbf{x}\| = (\mathbf{x}^T \mathbf{x})^{1/2}$ である.
- $X \subseteq \mathbb{R}^n$ が開集合 (open set) である
 - $\forall \mathbf{x} \in X, \exists \varepsilon > 0; B(\mathbf{x}, \varepsilon) \subseteq X$
- $X \subseteq \mathbb{R}^n$ が閉集合 (closed set) である
 - X の補集合 X^c が開集合である.
- \mathbf{x} が $X \subseteq \mathbb{R}^n$ の内点 (interior point) である
 - $\mathbf{x} \in X \subseteq \mathbb{R}^n$ に対して, $\exists \varepsilon > 0; B(\mathbf{x}, \varepsilon) \subseteq X$ が成立する.
- X の内部 (interior), $\text{Int}(X)$
 - $X \subseteq \mathbb{R}^n$ の内点の集合.
- \mathbf{x} が $X \subseteq \mathbb{R}^n$ の触点 (contact point) である
 - $\mathbf{x} \in X \subseteq \mathbb{R}^n$ に対して, $\forall \varepsilon > 0; B(\mathbf{x}, \varepsilon) \cap X \neq \emptyset$ が成立する.
- X の閉包 (closure), $\text{Cl}(X)$
 - $X \subseteq \mathbb{R}^n$ の触点の集合.
 - X を含む最小の閉集合.
- X の境界 (boundary), $\text{Bd}(X)$
 - $\text{Bd}(X) = \text{Cl}(X) \setminus \text{Int}(X)$.
- \mathbf{x} は X の集積点である
 - 任意の $\varepsilon > 0$ に対して, $B(\mathbf{x}, \varepsilon) \cap X$ が \mathbf{x} と異なる要素を含む.
- 孤立点 (isolated point)
 - X の集積点でない X の触点

- X が有界である (bounded)
 - $\exists \varepsilon > 0; X \subseteq B(\mathbf{0}, \varepsilon)$.
- 収束
 - 点列 $\{\mathbf{x}^i\}, i = 1, 2, \dots$ を考える. $\forall \varepsilon > 0, \exists I_\varepsilon; \|\mathbf{x}^i - \mathbf{x}\| < \varepsilon, i \geq I_\varepsilon$ となる点 \mathbf{x} が存在するとき, \mathbf{x} は点列 $\{\mathbf{x}^i\}$ の極限 (limit) といい, 点列 $\{\mathbf{x}^i\}$ は \mathbf{x} に収束する (converge) という.
- 点列 $\{\mathbf{x}^i\}$ の集積点
 - $\{\mathbf{x}^i\}$ の部分点列 $\{\mathbf{x}^{i_k}\}$ が点 \mathbf{x} に収束するとき, 点 \mathbf{x} を点列 $\{\mathbf{x}^i\}$ の集積点という.

有界閉集合 $X \subseteq \mathbb{R}^n$ について, $\{\mathbf{x}^i\} \subseteq X$ なる無限点列は少なくとも 1 つの集積点をもつ.

2.2 関数

- 連続性
 - 関数 $f: X \rightarrow \mathbb{R}^m$ を考える ($X \subseteq \mathbb{R}^n$).

$$\forall \varepsilon > 0, \exists \delta > 0; \|\mathbf{x} - \mathbf{x}^0\| < \delta \Rightarrow \|f(\mathbf{x}) - f(\mathbf{x}^0)\| < \varepsilon$$

が成立するとき, f は点 \mathbf{x}^0 で連続 (continuous) であるという.

- 任意の $\mathbf{x} \in X$ で連続となると, 関数 f は X で連続という.

- 実数値関数

- 値域が実数集合の関数 $f: X \rightarrow \mathbb{R}$.

- 実数値関数のクラス

- $f: X \rightarrow \mathbb{R} (X \subseteq \mathbb{R}^n \text{ は開集合})$ を考える.
- f が連続であるとき, X 上で C^0 級と呼ばれ, $f \in C^0$ と記す.
- $f \in C^0$ で, $\partial f(\mathbf{x})/\partial x_i, i = 1, 2, \dots, n$ が存在し, 連続であれば, f は X 上で C^1 級と呼ばれ, $f \in C^1$ と表す.
- $f \in C^1$ で, $\partial^2 f(\mathbf{x})/\partial x_i \partial x_j, i, j = 1, 2, \dots, n$ が存在し, 連続であれば, f は X 上で C^2 級と呼ばれ, $f \in C^2$ と表す.
- 以後 C^k 級も同様に定義される.

- 勾配ベクトル $\nabla f(\mathbf{x})$

$$\nabla f(\mathbf{x}) = \left(\frac{\partial f(\mathbf{x})}{\partial x_1}, \frac{\partial f(\mathbf{x})}{\partial x_2}, \dots, \frac{\partial f(\mathbf{x})}{\partial x_n} \right)$$

- ヘッセ行列 $H(\mathbf{x}) = \nabla^2 f(\mathbf{x})$

$$\nabla^2 f(\mathbf{x}) = \begin{pmatrix} \frac{\partial^2 f(\mathbf{x})}{\partial x_1^2} & \frac{\partial^2 f(\mathbf{x})}{\partial x_1 \partial x_2} & \dots & \frac{\partial^2 f(\mathbf{x})}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f(\mathbf{x})}{\partial x_2 \partial x_1} & \frac{\partial^2 f(\mathbf{x})}{\partial x_2^2} & \dots & \frac{\partial^2 f(\mathbf{x})}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f(\mathbf{x})}{\partial x_n \partial x_1} & \frac{\partial^2 f(\mathbf{x})}{\partial x_n \partial x_2} & \dots & \frac{\partial^2 f(\mathbf{x})}{\partial x_n^2} \end{pmatrix}$$

■Weierstrass の定理 有界閉集合 $X \subseteq \mathbb{R}^n$ 上の連続な実数値関数 $f(\mathbf{x})$ は X 内の点で最大値, 最小値をとる.

■平均値の定理 $f : X \rightarrow \mathbb{R} (X \subseteq \mathbb{R}^n), f \in C^1, \mathbf{x}^1, \mathbf{x}^2 \in X$ に対して,

$$f(\mathbf{x}^1) = f(\mathbf{x}^2) + \nabla f(\theta \mathbf{x}^1 + (1 - \theta) \mathbf{x}^2)^T (\mathbf{x}^1 - \mathbf{x}^2)$$

を満たす $\theta \in (0, 1)$ が存在する.

(注) 関数 $f : I \rightarrow \mathbb{R} (I \subseteq \mathbb{R}), f \in C^1, x_1, x_2 \in I$ に対しては,

$$f(x_1) = f(x_2) + f'(\theta x_1 + (1 - \theta)x_2)(x_1 - x_2)$$

を満たす $\theta \in (0, 1)$ が存在する. これは, 点 $x_1, x_2 \in I$ を結ぶ線分内に, $f(x_1), f(x_2)$ を結ぶ直線と同じ傾きとなる接線を持つ点 $c = \theta x_1 + (1 - \theta)x_2, \theta \in (0, 1)$ が存在することに相当する.

■Taylor の定理 $f : X \rightarrow \mathbb{R} (X \subseteq \mathbb{R}^n), f \in C^2, \mathbf{x}^1, \mathbf{x}^2 \in X$ に対して,

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

が成立する. ただし, o は $\lim_{t \rightarrow 0} o(t)/t = 0$ なる関数である. 同様に, $f \in C^1$ に対して,

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

が成立する.

(注) 関数 $f : I \rightarrow \mathbb{R} (I \subseteq \mathbb{R}), f \in C^1, x_1, x_2 \in I$ に対しては,

$$f(x_1) = f(x_2) + f'(x_2)(x_1 - x_2) + \frac{f''(x_2)}{2} (x_1 - x_2)^2 + o((x_1 - x_2)^2)$$

$$f(x_1) = f(x_2) + f'(x_2)(x_1 - x_2) + o(x_1 - x_2)$$

が成立する. o を用いずに表すと,

$$f(x_1) = f(x_2) + f'(x_2)(x_1 - x_2) + \frac{f''(x_2)}{2} (x_1 - x_2)^2 + \cdots + \frac{f^{(n-1)}(x_2)}{(n-1)!} (x_1 - x_2)^{n-1} + R_n$$

$$R_n := \frac{f^{(n)}(\theta x_1 + (1 - \theta)x_2)}{n!} (x_1 - x_2)^n$$

を満たす $\theta \in (0, 1)$ が存在する. $n = 1$ のときは平均値の定理. つまり, Taylor の定理は平均値の定理の一般化と考えることができる.

■陰関数の定理 $h_i : \mathbb{R}^n \rightarrow \mathbb{R}, \mathbf{x}^0 = (x_1^0, x_1^0, \dots, x_n^0)^T \in \mathbb{R}^n$ の近傍で, $h_i \in C^p (p \geq 1)$

$$h_i(\mathbf{x}^0) = 0, i = 1, 2, \dots, m$$

が成立するとする. このとき, ヤコビ行列 (Jacobian matrix)

$$J(\mathbf{x}^0) = \begin{pmatrix} \frac{\partial h_1(\mathbf{x}^0)}{\partial x_1} & \cdots & \frac{\partial h_1(\mathbf{x}^0)}{\partial x_m} \\ \vdots & \vdots & \vdots \\ \frac{\partial h_m(\mathbf{x}^0)}{\partial x_1} & \cdots & \frac{\partial h_m(\mathbf{x}^0)}{\partial x_m} \end{pmatrix}$$

が正則ならば、ある $\varepsilon > 0$ に対して $\hat{\mathbf{x}}^0 = (x_{m+1}^0, \dots, x_n^0)^T \in \mathbb{R}^{n-m}$ の近傍 $U = B(\hat{\mathbf{x}}^0, \varepsilon)$ が存在し、 $\hat{\mathbf{x}} \in U$ に対して、

1. $\phi_i \in C^p, i = 1, 2, \dots, m$
2. $x_i = \phi(\hat{\mathbf{x}}), i = 1, 2, \dots, m$
3. $h_i(\phi_1(\hat{\mathbf{x}}), \phi_2(\hat{\mathbf{x}}), \dots, \phi_m(\hat{\mathbf{x}}), \hat{\mathbf{x}}) = 0, i = 1, 2, \dots, m$

となる陰関数 (implicit function) $\phi_i, i = 1, 2, \dots, m$ が存在する*¹.

(注) 2 変数関数を考える. 関数 $f: \Omega \rightarrow \mathbb{R} (\Omega \subseteq \mathbb{R}^2), f \in C^p, p \geq 1, (x_1, x_2) \in \Omega$ とする. $f(x_1, x_2) = 0$ のとき, $x = x_1$ を含む開区間 I および I 上で定義された C^p 級の関数 ϕ がただ 1 つ存在し, $\phi(x_1) = x_2$ および

$$u(x) := f(x, \phi(x)) = 0 \quad (\forall x \in I)$$

が成立する.

2.3 凸集合

$X \subseteq \mathbb{R}^n$ とする. $\forall \mathbf{x}^1, \mathbf{x}^2 \in X, \forall \lambda \in [0, 1]$ に対して, $\lambda \mathbf{x}^1 + (1 - \lambda) \mathbf{x}^2 \in X$ が成立するとき, X は凸集合 (convex set) である*².

補題 2.1 $X_1, X_2, \dots \subseteq \mathbb{R}^n$ を凸集合とすると, $\bigcap_{i=1,2,\dots} X_i$ も凸集合となる.

補題 2.2 $X \subseteq \mathbb{R}^n$ を空でない閉凸集合, $\mathbf{y} \notin X$ とする. このとき,

$$\mathbf{a}^T \mathbf{x} \geq b > \mathbf{a}^T \mathbf{y}, \forall \mathbf{x} \in X$$

となる分離超平面 (separating hyperplane) $\{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{a}^T \mathbf{x} = b\}$ ($\mathbf{a} \neq \mathbf{0}$) が存在する.

補題 2.3 $X \subseteq \mathbb{R}^n$ を空でない凸集合, $\mathbf{y} \in \text{Bd}(X)$ とする. このとき,

$$\mathbf{a}^T \mathbf{x} \geq \mathbf{a}^T \mathbf{y}, \forall \mathbf{x} \in X$$

なる $\mathbf{a} \neq \mathbf{0}$ が存在する.

補題 2.4 $X, Y \subseteq \mathbb{R}^n$ を空でない凸集合, $X \cap Y = \emptyset$ とする. このとき,

$$\begin{aligned} \mathbf{a}^T \mathbf{x} &\geq b, \forall \mathbf{x} \in X, \\ \mathbf{a}^T \mathbf{y} &\leq b, \forall \mathbf{y} \in Y \end{aligned} \tag{2.1}$$

なる $\mathbf{a} \neq \mathbf{0}, b$ が存在する.

*¹ 解析学の教科書では、陰関数が唯一存在すると記されている.

*² 2 点 $\mathbf{x}^1, \mathbf{x}^2$ を結ぶ線分上の任意の内分点は, $x_\lambda = \lambda \mathbf{x}^1 + (1 - \lambda) \mathbf{x}^2, \lambda \in [0, 1]$ と書ける.

2.4 凸関数

$X \subseteq \mathbb{R}^n$ を空でない凸集合とする．関数 $f: X \rightarrow \mathbb{R}$ を考える．任意の $\mathbf{x}^1, \mathbf{x}^2 \in X$ ，任意の $\lambda \in [0, 1]$ に対して，

$$f(\lambda \mathbf{x}^1 + (1 - \lambda) \mathbf{x}^2) \leq \lambda f(\mathbf{x}^1) + (1 - \lambda) f(\mathbf{x}^2)$$

が成立するとき， f は凸関数 (convex function) であるという．また，任意の $\mathbf{x}^1, \mathbf{x}^2 \in X$ ($\mathbf{x}^1 \neq \mathbf{x}^2$)，任意の $\lambda \in (0, 1)$ に対して，

$$f(\lambda \mathbf{x}^1 + (1 - \lambda) \mathbf{x}^2) < \lambda f(\mathbf{x}^1) + (1 - \lambda) f(\mathbf{x}^2)$$

が成立するとき， f は強意の凸関数 (strictly convex function) であるという．さらに，任意の $\mathbf{x}^1, \mathbf{x}^2 \in X$ ，任意の $\lambda \in [0, 1]$ に対して，

$$f(\lambda \mathbf{x}^1 + (1 - \lambda) \mathbf{x}^2) \leq \max(f(\mathbf{x}^1), f(\mathbf{x}^2))$$

が成立するとき， f は準凸関数 (quasi-convex function) であるという．また，任意の $\mathbf{x}^1, \mathbf{x}^2 \in X$ ($\mathbf{x}^1 \neq \mathbf{x}^2$)，任意の $\lambda \in (0, 1)$ に対して，

$$f(\lambda \mathbf{x}^1 + (1 - \lambda) \mathbf{x}^2) < \max(f(\mathbf{x}^1), f(\mathbf{x}^2))$$

が成立するとき， f は強意の準凸関数 (strictly quasi-convex function) であるという． $-f$ が凸関数，強意の凸関数，準凸関数，強意の準凸関数であるとき，それぞれ， f は凹関数 (concave function)，強意の凹関数，準凹関数，強意の準凹関数であるという．

定理 2.1 $X \subseteq \mathbb{R}^n$ を空でない凸集合とする．関数 $f: X \rightarrow \mathbb{R}$ が凸関数であるための必要十分条件は， \mathbb{R}^{n+1} の部分集合，

$$\text{epi} f = \{(\mathbf{x}, r) \mid \mathbf{x} \in X, r \geq f(\mathbf{x}), r \in \mathbb{R}\}$$

が凸集合となることである． $\text{epi} f$ を関数 f のエピグラフ (epigraph) という．

定理 2.2 $X \subseteq \mathbb{R}^n$ を空でない凸集合とする．関数 $f: X \rightarrow \mathbb{R}$ が準凸関数であるための必要十分条件は， \mathbb{R}^n の部分集合，

$$L(r) = \{\mathbf{x} \in \mathbb{R}^n \mid f(\mathbf{x}) \leq r\}$$

が任意の $r \in \mathbb{R}$ に対して凸集合となることである． $L(r)$ を関数 f のレベル集合 (level set) という．

定理 2.3 $X \subseteq \mathbb{R}^n$ を空でない開凸集合とする．関数 $f: X \rightarrow \mathbb{R}$ が微分可能とする． f が凸関数であるための必要十分条件は，

$$f(\mathbf{x}^2) \geq f(\mathbf{x}^1) + \nabla f(\mathbf{x}^1)^T (\mathbf{x}^2 - \mathbf{x}^1), \forall \mathbf{x}^1, \mathbf{x}^2 \in X$$

が成立することである．また， f が強意の凸関数であるための必要十分条件は，

$$f(\mathbf{x}^2) > f(\mathbf{x}^1) + \nabla f(\mathbf{x}^1)^T (\mathbf{x}^2 - \mathbf{x}^1), \forall \mathbf{x}^1, \mathbf{x}^2 \in X \text{ s.t. } \mathbf{x}^1 \neq \mathbf{x}^2$$

が成立することである．

(参考) $\mathbf{x} \in X$ ($X \subseteq \mathbb{R}^n$ は開集合), $\mathbf{d} \in \mathbb{R}^n$ とする. $f: X \rightarrow \mathbb{R}$ に対して, 方向微分係数 (directional derivative) は,

$$f'(\mathbf{x}; \mathbf{d}) = \lim_{r \rightarrow +0} \frac{f(\mathbf{x} + r\mathbf{d}) - f(\mathbf{x})}{r}$$

と定められる. f が微分可能なとき, $f'(\mathbf{x}; \mathbf{d}) = \nabla f(\mathbf{x})^T \mathbf{d}$ が成立する*³.

定理 2.4 $X \subseteq \mathbb{R}^n$ を空でない開凸集合とする. 関数 $f: X \rightarrow \mathbb{R}$ が 2 階微分可能とする.

1. f が凸関数であるための必要十分条件は, f のヘッセ行列 $H(\mathbf{x})$ が任意の $\mathbf{x} \in X$ に対して半正定 (positive semi-definite), すなわち,

$$\mathbf{y}^T H(\mathbf{x}) \mathbf{y} \geq 0, \forall \mathbf{y} \in \mathbb{R}^n$$

が成立することである.

2. f が強意の凸関数であるための十分条件は, f のヘッセ行列 $H(\mathbf{x})$ が任意の $\mathbf{x} \in X$ に対して正定 (positive definite), すなわち,

$$\mathbf{y}^T H(\mathbf{x}) \mathbf{y} > 0, \forall \mathbf{y} \in \mathbb{R}^n, \mathbf{y} \neq \mathbf{0}$$

が成立することである*⁴.

(凸関数の連続性) $X \subseteq \mathbb{R}^n$ を空でない凸集合とする. 関数 $f: X \rightarrow \mathbb{R}$ が凸関数であれば, f は X の任意の内点で連続である.

Taylor の公式 $f: X \rightarrow \mathbb{R}$ ($X \subseteq \mathbb{R}^n$) が $\bar{\mathbf{x}} \in X$ の近傍で 2 階微分可能なとき,

$$f(\mathbf{x}) = f(\bar{\mathbf{x}}) + \nabla f(\bar{\mathbf{x}})^T (\mathbf{x} - \bar{\mathbf{x}}) + \frac{1}{2} (\mathbf{x} - \bar{\mathbf{x}})^T H((1 - \lambda)\bar{\mathbf{x}} + \lambda\mathbf{x}) (\mathbf{x} - \bar{\mathbf{x}})$$

なる $\lambda \in (0, 1)$ が存在する.

*³ f が微分可能のとき, Taylor の定理より, $f(\mathbf{x} + r\mathbf{d}) = f(\mathbf{x}) + r\nabla f(\mathbf{x})^T \mathbf{d} + o(r\|\mathbf{d}\|)$ である. さらに, $\lim_{r \rightarrow +0} o(r\|\mathbf{d}\|)/r = 0$ より上式を得る.

*⁴ この逆は一般に成立しない.

第 3 章

最適化問題の定式化

本章では、一般の最適化問題について定式化を行い、用語の定義を行う。

3.1 最適化問題

最適化問題を口語的に定義すると、「与えられた条件の下で何らかの関数を最小化、もしくは最大化する問題」である。最適化問題は以下のように定義される。ここでは、最小化問題を最適化問題とする。

基礎となる空間 X , $S \subseteq \mathbb{R}^n$, X 上で定義された関数 $f: X \rightarrow \mathbb{R}$ が与えられたとき, f を最小にする解 $\mathbf{x} \in S \cap X$ を求める問題を最適化問題 (optimization problem), あるいは計画問題 (programming problem) という。つまり、最適化問題は、(3.1) 式で定義される問題である。

$$\begin{aligned} & \text{minimize : } f(\mathbf{x}) \\ & \text{subject to : } \mathbf{x} \in S \cap X \end{aligned} \tag{3.1}$$

例えば X として, \mathbb{R}^n や離散集合などが考えられる。連続最適化問題においては通常, S は関数 $g_i: \mathbb{R}^n \rightarrow \mathbb{R}$, $i = 1, 2, \dots, m$ による不等式および等式制約を用いて,

$$S = \{\mathbf{x} \in \mathbb{R}^n \mid g_i(\mathbf{x}) \leq 0, i = 1, 2, \dots, l, g_i(\mathbf{x}) = 0, i = l + 1, \dots, m\} \tag{3.2}$$

と表現される。

3.2 用語

$S \cap X$ 実行可能領域 (feasible region)

- $S \cap X \neq \emptyset$ ならば, 実行可能 (feasible)
- $S \cap X = \emptyset$ ならば, 実行不能 (infeasible)

$\mathbf{x} \in S \cap X$ 実行可能解 (feasible solution)

実行可能解 $\mathbf{x} \in S \cap X$ のうち, 目的関数値を最小にする解を最適解 (optimal solution) という。つまり, 最適化問題 3.1 の最適解 \mathbf{x}^* は, $\forall \mathbf{x} \in S \cap X, f(\mathbf{x}^*) \leq f(\mathbf{x})$ を満たす $\mathbf{x}^* \in S \cap X$ である。

最適化問題 (3.1) において、最適解は常に存在するわけではない。実行不能であるとき、あるいはいくらでも目的関数値を小さくできる場合などでは、最適解は存在しない*1。

最適化問題は、基礎となる空間 X 、目的関数 f 、制約式 (3.2) における g_i により次のように分類される。

- 非線形計画問題 (nonlinear programming problem)
 - $X = \mathbb{R}^n$
 - f や S を定義する g_i に制限を置かない。
- 線形計画問題 (linear programming problem)
 - f, g_i がすべて線形 (1 次関数)。
- 整数計画問題 (integer programming problem)
 - $X = \mathbb{Z}^n$
 - f, g_i が線形。
 - すべての変数が整数変数の全整数計画問題 (all-integer programming problem)、整数変数と実数変数が混在する混合整数計画問題 (mixed-integer programming problem) に分類される。
- 組合せ最適化問題 (combinatorial optimization problem)
 - 離散集合 X に対する最適化問題。
 - グラフ理論など。

最適化問題 (3.1) は、

$$f^* = \inf_{x \in S \cap X} f(x)$$

を求める問題と解釈できる。 f^* を最適値 (optimal value) という。最適値 f^* は常に定義され、最小化問題については、実行不能ならば $f^* = \infty$ 、発散するならば $f^* = -\infty$ であり、最適解 x^* が存在するならば $f^* = f(x^*)$ である。

*1 実行不能であるとき、この問題を非有界 (unbounded) であるといい、いくらでも目的関数値を小さくできる場合、この問題は発散する (divergent) という。

第 4 章

非線形計画問題

本章では、非線形計画問題の最適性条件について述べる。

非線形計画問題においては、最適性の定義を次のように緩めることが多い。

- \mathbf{x}^* は局所最小解 (locally minimal solution) である。
 - $\mathbf{x}^* \in S$ に対し、ある近傍 B が存在し、任意の $\mathbf{x} \in S \cap B$ に対して $f(\mathbf{x}^*) \leq f(\mathbf{x})$ が成立する。
- \mathbf{x}^* は局所最大解 (locally maximum solution) である。
 - $\mathbf{x}^* \in S$ に対し、ある近傍 B が存在し、任意の $\mathbf{x} \in S \cap B$ に対して $f(\mathbf{x}^*) \geq f(\mathbf{x})$ が成立する。

ここでは、最小化問題を基本としているので、局所最小解を局所最適解 (locally optimal solution) ともいう。また、本来の最適解を大域最適解 (globally optimal solution) という。まず、局所最適解の必要条件、十分条件について述べ、局所最適解が大域最適解になる場合について述べる。

4.1 制約なし問題の最適性条件

本節では最適化問題として、(4.1) を考える。

$$\begin{aligned} & \text{minimize : } f(\mathbf{x}) \\ & \text{subject to : } \mathbf{x} \in \mathbb{R}^n \end{aligned} \tag{4.1}$$

ただし、 f は任意の非線形関数であり、適当な微分可能性を仮定する。

問題 (4.1) の最適性の必要条件として、次の 2 つの定理を示す。

定理 4.1 $f \in C^1$ のとき、 $\bar{\mathbf{x}} \in \mathbb{R}^n$ が問題 (4.1) の局所最適解であるための必要条件は、

$$\nabla f(\bar{\mathbf{x}}) = \mathbf{0}$$

となることである。

定理 4.2 $f \in C^2$ のとき、 $\bar{\mathbf{x}}$ が問題 (4.1) の局所最適解であるための必要条件は、

$$\nabla f(\bar{\mathbf{x}}) = \mathbf{0}$$

かつ、 $\bar{\mathbf{x}}$ におけるヘッセ行列 $H(\bar{\mathbf{x}})$ が半正定値、すなわち、

$$\mathbf{d}^T H(\bar{\mathbf{x}}) \mathbf{d} \geq 0, \forall \mathbf{d} \in \mathbb{R}^n$$

が成立することである。

さらに、局所最適解の十分条件は、定理 4.3 で与えられる。

定理 4.3 $f \in C^2$ のとき、 $\bar{\mathbf{x}}$ が問題 (4.1) の局所最適解であるための十分条件は、

$$\nabla f(\bar{\mathbf{x}}) = \mathbf{0}$$

かつヘッセ行列 $H(\bar{\mathbf{x}})$ が正定値、すなわち、

$$\mathbf{d}^T H(\bar{\mathbf{x}}) \mathbf{d} > 0, \forall \mathbf{d} \in \mathbb{R}^n, \mathbf{d} \neq \mathbf{0}$$

が成立することである。

ここで、次の例題を解く。

$$\begin{aligned} &\text{minimize : } f(\mathbf{x}) = 2x_1^2 + x_1x_2 + x_2^2 - 5x_1 - 3x_2 + 4 \\ &\text{subject to : } \mathbf{x} \in \mathbb{R}^2 \end{aligned}$$

勾配ベクトルは、 $\nabla f(\mathbf{x}) = (4x_1 + x_2 - 5, x_1 + 2x_2 - 3)^T$ であるので、 f の停留点は、 $x_1 = 1, x_2 = 1$ である。また、 f のヘッセ行列は、

$$H(\mathbf{x}) = \begin{pmatrix} 4 & 1 \\ 1 & 2 \end{pmatrix}$$

であり、固有値 $\lambda_1 = 3 + \sqrt{2}, \lambda_2 = 3 - \sqrt{2}$ がともに正より、正定値である。よって、定理 4.3 よりこの停留点は局所最適解である*1。□

4.2 等式制約問題の最適性条件

本節では最適化問題として、(4.2) を考える。

$$\begin{aligned} &\text{minimize : } f(\mathbf{x}) \\ &\text{subject to : } g_i(\mathbf{x}) = 0, i = 1, 2, \dots, m \\ &\mathbf{x} \in \mathbb{R}^n \end{aligned} \tag{4.2}$$

等式制約問題 (4.2) における主要な結果は、目的関数 $f(\mathbf{x})$ の代わりに Lagrange 関数 (Lagrangian function)

$$L(\mathbf{x}, \mathbf{u}) = f(\mathbf{x}) + \mathbf{u}^T \mathbf{g}(\mathbf{x}) \tag{4.3}$$

を用いると、前節の定理 4.1, 定理 4.2 を自然に拡張できるという点にある。ただし、 $\mathbf{u} \in \mathbb{R}^m, \mathbf{g}(\mathbf{x}) = (g_1(\mathbf{x}), g_2(\mathbf{x}), \dots, g_m(\mathbf{x}))^T$ である。

*1 f の等高線は、停留点を中心とした、2 軸が固有ベクトルの楕円となる。

■等式条件のペナルティ関数 問題 (4.2) とパラメータ $k = 1, 2, \dots$ から次の関数を定義する.

$$F^k(\mathbf{x}) = f(\mathbf{x}) + \frac{k}{2}\|\mathbf{g}(\mathbf{x})\|^2 + \frac{\alpha}{2}\|\mathbf{x} - \bar{\mathbf{x}}\|^2 \quad (4.4)$$

ただし, $\bar{\mathbf{x}}$ は, 問題 (4.2) の局所最適解である. また, $\alpha > 0$ とする. (4.4) 式において, 第 2 項は, 等式制約の違反に対するペナルティと解釈することができる. 第 3 項は, $\mathbf{g}(\mathbf{x}) = \mathbf{0}$ を満たす解 \mathbf{x} を考えるとき, $\bar{\mathbf{x}}$ が近傍内で $F^k(\mathbf{x})$ の唯一の局所最適解となるように導入されたものである.

$\bar{\mathbf{x}}$ は局所最適解であるので, 適当な $\varepsilon > 0$ を選べば, $\bar{\mathbf{x}}$ の近傍の閉包

$$\bar{B}(\bar{\mathbf{x}}, \varepsilon) = \{\mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x} - \bar{\mathbf{x}}\| \leq \varepsilon\} \quad (4.5)$$

において, そこに属す任意の実行可能解 \mathbf{x} に対して, $f(\bar{\mathbf{x}}) \leq f(\mathbf{x})$ である最適化問題

$$\begin{aligned} &\text{minimize : } F^k(\mathbf{x}) \\ &\text{subject to : } \mathbf{x} \in \bar{B}(\bar{\mathbf{x}}, \varepsilon) \end{aligned} \quad (4.6)$$

を定義する. $\bar{B}(\bar{\mathbf{x}}, \varepsilon)$ は有界閉集合より, Weierstrass の定理から最小解が存在する. 最小解 \mathbf{x}^k による点列 $\{\mathbf{x}^k\}$ が $\bar{\mathbf{x}}$ に収束することを示す.

任意の k に対し, \mathbf{x}^k の最適性より,

$$\begin{aligned} F^k(\mathbf{x}^k) &= f(\mathbf{x}^k) + \frac{k}{2}\|\mathbf{g}(\mathbf{x}^k)\|^2 + \frac{\alpha}{2}\|\mathbf{x}^k - \bar{\mathbf{x}}\|^2 \\ &\leq F^k(\bar{\mathbf{x}}) = f(\bar{\mathbf{x}}) \end{aligned} \quad (4.7)$$

より, $\lim_{k \rightarrow \infty} \|\mathbf{g}(\mathbf{x}^k)\| = 0$ でなければならない*2. 従って, $\{\mathbf{x}^k\}$ の任意の極限点 $\bar{\bar{\mathbf{x}}}$ は, $\mathbf{g}(\bar{\bar{\mathbf{x}}}) = \mathbf{0}$ を満たすので, 最適化問題 (4.2) の実行可能解である. (4.7) 式から,

$$f(\bar{\bar{\mathbf{x}}}) + \frac{\alpha}{2}\|\bar{\bar{\mathbf{x}}} - \bar{\mathbf{x}}\|^2 \leq f(\bar{\mathbf{x}})$$

を得る. また, $\bar{\mathbf{x}}$ が, 元の問題 (4.2) の実行可能解であることから, $f(\bar{\mathbf{x}}) \leq f(\bar{\bar{\mathbf{x}}})$. これと上式を合わせて,

$$\|\bar{\bar{\mathbf{x}}} - \bar{\mathbf{x}}\|^2 = 0 \quad \therefore \bar{\bar{\mathbf{x}}} = \bar{\mathbf{x}}$$

つまり, 点列 $\{\mathbf{x}^k\}$ は $\bar{\mathbf{x}}$ に収束する. 点 $\bar{\mathbf{x}}$ は問題 (4.6) の実行可能領域の内点である. また, k が十分大きければ, \mathbf{x}^k も内点である. したがって, \mathbf{x}^k は制約なしの最適化問題

$$\begin{aligned} &\text{minimize : } F^k(\mathbf{x}) \\ &\text{subject to : } \mathbf{x} \in \mathbb{R}^n \end{aligned} \quad (4.8)$$

の局所最適解でもある.

*2 これを満たさなければ, $k \rightarrow \infty$ で発散する.

■等式制約問題の最適性必要条件 問題 (4.2) の m 個の制約関数の勾配ベクトル $g_i(\mathbf{x})$, $i = 1, 2, \dots, m$ が互いに線形独立であるとき, 点 \mathbf{x} は正則 (regular) であるという. 問題 (4.2) の最適性の必要条件は, 定理 4.4 で与えられる.

定理 4.4 1. 1 次の最適性必要条件

$f, g_i \in C^1$, $i = 1, 2, \dots, m$ であり, $\bar{\mathbf{x}}$ が正則であるとき, $\bar{\mathbf{x}} \in \mathbb{R}^n$ が問題 (4.2) の局所最適解であるための必要条件は,

$$\nabla f(\bar{\mathbf{x}}) + \sum_{i=1}^m u_i \nabla g_i(\bar{\mathbf{x}}) = \mathbf{0} \quad (4.9)$$

を満たすラグランジュ乗数ベクトル $\mathbf{u} = (u_1, u_2, \dots, u_m)^T$ が存在することである*3.

2. 2 次の最適性必要条件

1 の仮定に加え, $f, g_i \in C^2$, $i = 1, 2, \dots, m$ であるとき, $\bar{\mathbf{x}} \in \mathbb{R}^n$ が問題 (4.2) の局所最適解であるための必要条件は,

$$\begin{aligned} \nabla f(\bar{\mathbf{x}}) + \sum_{i=1}^m u_i \nabla g_i(\bar{\mathbf{x}}) &= \mathbf{0} \\ \mathbf{y}^T \left(\nabla^2 f(\bar{\mathbf{x}}) + \sum_{i=1}^m u_i \nabla^2 g_i(\bar{\mathbf{x}}) \right) \mathbf{y} &\geq 0, \quad \forall \mathbf{y} \in V(\bar{\mathbf{x}}) \end{aligned} \quad (4.10)$$

を満たすラグランジュ乗数ベクトル $\mathbf{u} = (u_1, u_2, \dots, u_m)^T$ が存在することである. ただし,

$$V(\bar{\mathbf{x}}) = \{\mathbf{y} \in \mathbb{R}^n \mid \nabla g(\bar{\mathbf{x}})\mathbf{y} = \mathbf{0}\} \quad (4.11)$$

である.

ここで, Lagrange 関数 (4.3) を用いると, 等式制約条件 $\mathbf{g}(\mathbf{x}) = \mathbf{0}$ および条件 (4.9), (4.10) は次のように表される.

$$\begin{aligned} \nabla_{\mathbf{u}} L(\bar{\mathbf{x}}, \mathbf{u}) &= \mathbf{0} \\ \nabla_{\mathbf{x}} L(\bar{\mathbf{x}}, \mathbf{u}) &= \mathbf{0} \\ \mathbf{y}^T \nabla_{\mathbf{xx}}^2 L(\bar{\mathbf{x}}, \mathbf{u}) \mathbf{y} &\geq 0, \quad \forall \mathbf{y} \in V(\bar{\mathbf{x}}) \end{aligned}$$

■最適性条件の幾何学的解釈 条件 (4.9) の幾何学的な意味を考察する. 簡単のため, 制約条件が 1 つの場合, すなわち, 制約条件が $g(\mathbf{x}) = 0$ の場合を考える. 条件 (4.9) は, $\nabla f(\mathbf{x}) + u \nabla g(\mathbf{x}) = \mathbf{0}$ と書くことができる. u の符号をどちらにとっても良いことから, この条件は, 2 つのベクトル $\nabla f(\mathbf{x})$ と $\nabla g(\mathbf{x})$ が同方向か, あるいは逆方向であることを述べている. $g(\mathbf{x}) = 0$ の範囲内で微小量移動することは, $\nabla g(\mathbf{x})$ の直交方向に進むことに相当する. この方向は $\nabla f(\mathbf{x})$ とも直行しているので, 目的関数値は変化せず, 最適性の必要条件を与える. 複数の制約条件を持つ一般の場合, 条件 (4.9) は, ベクトル $\nabla g_i(\mathbf{x})$, $i = 1, 2, \dots, m$ が張る空間内にベクトル $\nabla f(\mathbf{x})$ が位置していることを要求している. これも上と同様の解釈が可能である.

*3 ラグランジュ乗数ベクトル \mathbf{u} は, 等式制約条件の右辺の変化が目的関数値にどのような影響を持つかを示すベクトルである.

■等式制約問題の最適性十分条件 制約なし問題の最適性の十分条件 (定理 4.3) を等式制約の場合へ拡張する。十分条件は、定理 4.5 で与えられる。

定理 4.5 $f, g_i \in C^2, i = 1, 2, \dots, m$ であるとき, $\bar{\mathbf{x}} \in \mathbb{R}^n$ が問題 (4.2) の局所最適解であるための十分条件は,

$$\begin{aligned}\nabla_{\mathbf{u}} L(\bar{\mathbf{x}}, \mathbf{u}) &= \mathbf{0} \\ \nabla_{\mathbf{x}} L(\bar{\mathbf{x}}, \mathbf{u}) &= \mathbf{0} \\ \mathbf{y}^T \nabla_{\mathbf{x}\mathbf{x}}^2 L(\bar{\mathbf{x}}, \mathbf{u}) \mathbf{y} &> 0, \forall \mathbf{y} (\neq \mathbf{0}) \in V(\bar{\mathbf{x}})\end{aligned}$$

を満たすラグランジュ乗数ベクトル $\mathbf{u} = (u_1, u_2, \dots, u_m)^T$ が存在することである。

ここで、次の例題を解く。

$$\begin{aligned}\text{minimize : } f(\mathbf{x}) &= -x_1 x_2 \\ \text{subject to : } g(\mathbf{x}) &= x_1^2 + x_2^2 - 1 = 0 \\ \mathbf{x} &\in \mathbb{R}^2\end{aligned}$$

Lagrange 関数を $L(\mathbf{x}, u)$ とすると,

$$L(\mathbf{x}, u) = f(\mathbf{x}) + u g(\mathbf{x}) = -x_1 x_2 + u(x_1^2 + x_2^2 - 1)$$

最適性の必要条件より,

$$\begin{aligned}\nabla_u L(\mathbf{x}, u) &= x_1^2 + x_2^2 - 1 = 0 \\ \nabla_{\mathbf{x}} L(\mathbf{x}, u) &= \begin{pmatrix} -x_2 \\ -x_1 \end{pmatrix} + u \begin{pmatrix} 2x_1 \\ 2x_2 \end{pmatrix} = \mathbf{0}\end{aligned}$$

この連立方程式を解くと,

$$(\bar{x}_1, \bar{x}_2, u) = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, \frac{1}{2} \right), \left(-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}, \frac{1}{2} \right), \left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}, -\frac{1}{2} \right), \left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, -\frac{1}{2} \right)$$

が得られる。前半の 2 つの解について,

$$\begin{aligned}\nabla_{\mathbf{x}\mathbf{x}} L(\bar{\mathbf{x}}, u) &= \nabla^2 f(\bar{\mathbf{x}}) + u \nabla^2 g(\bar{\mathbf{x}}) \\ &= \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} + \frac{1}{2} \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}\end{aligned}$$

から、固有値 λ_1, λ_2 は、 $\lambda_1 = 2, \lambda_2 = 0$ 。よって、この行列は半正定値であり、任意の \mathbf{y} に対し、 $\mathbf{y}^T (\nabla_{\mathbf{x}\mathbf{x}}) \mathbf{y} \geq 0$ を満たす。よって、この 2 解は最適性の必要条件を満たす。

次に、後半の 2 解についても同様に $\nabla_{\mathbf{x}\mathbf{x}}$ を求めると,

$$\nabla_{\mathbf{x}\mathbf{x}} = \begin{pmatrix} -1 & -1 \\ -1 & -1 \end{pmatrix}$$

から、固有値 λ_1, λ_2 は、 $\lambda_1 = 0, \lambda_2 = -2$ 。よって、この行列は半負定値である。2 解に対し、 $(\nabla g(\bar{\mathbf{x}}))^T \mathbf{y} = \mathbf{0}$ を満たすベクトル \mathbf{y} の集合は、パラメータ t を用いて,

$$V(\bar{\mathbf{x}}) = \{\mathbf{y} = (t, t)^T \mid t \in \mathbb{R}\}$$

と書ける。これを用いると、

$$\begin{pmatrix} t & t \end{pmatrix} \begin{pmatrix} -1 & -1 \\ -1 & -1 \end{pmatrix} \begin{pmatrix} t \\ t \end{pmatrix} = -4t^2$$

である。これは、 $t \neq 0$ で負の値を取り、最適性の必要条件を満たさない。よって、この2解は局所最適解ではない。

必要条件を満たす2解に対して、 $(\nabla g(\bar{\mathbf{x}}))^T \mathbf{y} = \mathbf{0}$ を満たすベクトル \mathbf{y} の集合は、パラメータ t を用いて、

$$V(\bar{\mathbf{x}}) = \{\mathbf{y} = (t, -t)^T \mid t \in \mathbb{R}\}$$

である。よって、

$$\begin{pmatrix} t & -t \end{pmatrix} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} t \\ -t \end{pmatrix} = 4t^2 \geq 0$$

であり、 $t \neq 0$ では正の値を取る。よって、これは最適性の十分条件を満たす。

以上より、この問題の局所最適解は、

$$\bar{\mathbf{x}} = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right)^T, \left(-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}} \right)^T$$

であり、その目的関数値は $-1/2$ である。 □

4.3 不等式制約問題の最適性条件

本節では最適化問題として、(4.12) を考える。

$$\begin{aligned} & \text{minimize : } f(\mathbf{x}) \\ & \text{subject to : } g_i(\mathbf{x}) \leq 0, \quad i = 1, 2, \dots, l \\ & \quad \quad \quad g_i(\mathbf{x}) = 0, \quad i = l+1, \dots, m \\ & \quad \quad \quad \mathbf{x} \in \mathbb{R}^n \end{aligned} \tag{4.12}$$

実行可能解 \mathbf{x} は、各不等式条件 $g_i(\mathbf{x}) = 0$ を等式で満たす場合と真の不等式で満たす場合がある。不等式条件において、 $g_i(\mathbf{x}) = 0$ となるものを、 \mathbf{x} において有効である (active) という。有効制約条件の添字集合を、

$$A(\mathbf{x}) = \{i \mid g_i(\mathbf{x}) = 0, \quad i = 1, 2, \dots, l\} \tag{4.13}$$

と記す。実行可能解 \mathbf{x} において、等式で成立する条件の勾配ベクトル $\nabla g_i(\mathbf{x})$, $i \in A(\mathbf{x})$ および $\nabla g_i(\mathbf{x})$, $i = l+1, \dots, m$ が線形独立であるとき、 \mathbf{x} は正則 (regular) であるという。

■不等式制約問題の最適性条件 Lagrange 関数 (4.3) に基づく等式制約問題の最適性条件 (定理 4.4) の一般化を行う。問題 (4.12) の最適性の必要条件は、定理 4.6 で与えられる。

定理 4.6 1. 1 次の最適性必要条件

$f, g_i \in C^1, i = 1, 2, \dots, m$ であり, $\bar{\mathbf{x}}$ が正則であるとき, $\bar{\mathbf{x}} \in \mathbb{R}^n$ が問題 (4.12) の局所最適解であるための必要条件は,

$$\begin{aligned}\nabla_{\mathbf{x}} L(\bar{\mathbf{x}}, \mathbf{u}) &= \nabla f(\bar{\mathbf{x}}) + \sum_{i=1}^m u_i \nabla g_i(\bar{\mathbf{x}}) = \mathbf{0} \\ u_i &\geq 0, \quad i = 1, 2, \dots, l \\ u_i &= 0, \quad i \in \{1, 2, \dots, l\} - A(\bar{\mathbf{x}})\end{aligned}\tag{4.14}$$

を満たすラグランジュ乗数ベクトル $\mathbf{u} \in \mathbb{R}^m$ が存在することである*4.

2. 2 次の最適性必要条件

1 の仮定に加え, $f, g_i \in C^2, i = 1, 2, \dots, m$ であるとき, $\bar{\mathbf{x}} \in \mathbb{R}^n$ が問題 (4.12) の局所最適解であるための必要条件は, (4.14) に加え,

$$\mathbf{y}^T \left(\nabla^2 f(\bar{\mathbf{x}}) + \sum_{i=1}^m u_i \nabla^2 g_i(\bar{\mathbf{x}}) \right) \mathbf{y} \geq 0, \quad \forall \mathbf{y} \in V'(\bar{\mathbf{x}})\tag{4.15}$$

を満たすラグランジュ乗数ベクトル $\mathbf{u} = (u_1, u_2, \dots, u_m)^T$ が存在することである. ただし,

$$V'(\bar{\mathbf{x}}) = \{ \mathbf{y} \in \mathbb{R}^n \mid \nabla g_i(\bar{\mathbf{x}})^T \mathbf{y} = 0, i \in A(\bar{\mathbf{x}}) \cup \{l+1, \dots, m\} \}\tag{4.16}$$

である.

ここで, (4.14) の 3 番目の条件は,

$$u_i g_i(\bar{\mathbf{x}}) = 0, \quad i = 1, 2, \dots, l\tag{4.17}$$

と書くことができる*5. この条件は相補性条件 (complementarity condition) と呼ばれる. また, 定理 4.6 の式 (4.14) および (4.15) は KKT 条件と呼ばれている*6.

■不等式制約問題の最適性十分条件 定理 4.5 にならって, 不等式制約問題 (4.12) に対する 2 次の最適性十分条件を与える.

定理 4.7 $f, g_i \in C^2, i = 1, 2, \dots, m$ であるとき, $\bar{\mathbf{x}} \in \mathbb{R}^n$ が問題 (4.12) の局所最適解であるための十分条件は,

*4 不等式制約のうち, 有効でないものに対する u_i は 0. 等式制約の $u_i, i = l+1, \dots, m$ に対する符号制約はない.

*5 $g_i(\bar{\mathbf{x}}) < 0$ ならば $u_i = 0$, $g_i(\bar{\mathbf{x}}) = 0$ ならば $u_i \geq 0$.

*6 KKT 条件は, $\bar{\mathbf{x}}$ から実行可能領域のどの方向を選んでも, $-\nabla f(\bar{\mathbf{x}})$ と鋭角をなすことができず, 目的関数値を減少させることができないことを示している (後述).

$$\begin{aligned}
g_i(\bar{\mathbf{x}}) &\leq 0, \quad i = 1, 2, \dots, l \\
g_i(\bar{\mathbf{x}}) &= 0, \quad i = l + 1, \dots, m \\
\nabla f(\bar{\mathbf{x}}) + \sum_{i=1}^m u_i \nabla g_i(\bar{\mathbf{x}}) &= \mathbf{0} \\
u_i &> 0, \quad i \in A(\bar{\mathbf{x}}) \\
u_i &= 0, \quad i \in \{1, 2, \dots, l\} - A(\bar{\mathbf{x}}) \\
\mathbf{y}^T \left(\nabla^2 f(\bar{\mathbf{x}}) + \sum_{i=1}^m u_i \nabla^2 g_i(\bar{\mathbf{x}}) \right) \mathbf{y} &> 0, \quad \forall \mathbf{y} (\neq \mathbf{0}) \in V'(\bar{\mathbf{x}})
\end{aligned} \tag{4.18}$$

を満たすラグランジュ乗数ベクトル $\mathbf{u} = (u_1, u_2, \dots, u_m)^T$ が存在することである。ただし、 $A(\bar{\mathbf{x}})$ の定義は (4.13) 式、 $V'(\bar{\mathbf{x}})$ の定義は (4.16) 式にある。

ここで、次の例題を解く。

$$\begin{aligned}
&\text{minimize : } f(\mathbf{x}) = 2x_1^2 + x_1x_2 + x_2^2 + x_1 - 3x_2 + 3 \\
&\text{subject to : } g_1(\mathbf{x}) = x_1^2 + x_2^2 - 1 \leq 0 \\
&\quad \quad \quad g_2(\mathbf{x}) = -x_1 \leq 0 \\
&\quad \quad \quad g_3(\mathbf{x}) = -x_2 \leq 0 \\
&\quad \quad \quad \mathbf{x} \in \mathbb{R}^2
\end{aligned}$$

目的関数および制約式の勾配ベクトルはそれぞれ、

$$\begin{aligned}
\nabla f(\mathbf{x}) &= (4x_1 + x_2 + 1, x_1 + 2x_2 - 3)^T \\
\nabla g_1(\mathbf{x}) &= (2x_1, 2x_2)^T \\
\nabla g_2(\mathbf{x}) &= (-1, 0)^T \\
\nabla g_3(\mathbf{x}) &= (0, -1)^T
\end{aligned}$$

よって、 f の停留点は、 $(-5/7, 13/7)$ である。また、 f のヘッセ行列は、

$$\nabla^2 f(\mathbf{x}) = \begin{pmatrix} 4 & 1 \\ 1 & 2 \end{pmatrix}$$

から、固有値は $\lambda_1 = 3 + \sqrt{2}, \lambda_2 = 3 - \sqrt{2}$ であり、固有値 λ_1, λ_2 に属する固有ベクトルはそれぞれ $\mathbf{v}_1 = (1, \sqrt{2} - 1)^T, \mathbf{v}_2 = (1, -\sqrt{2} - 1)^T$ である。この問題の実行可能領域および f の等高線を図 4.1 に示す。

有効制約は $g_1(\mathbf{x}), g_2(\mathbf{x})$ と考えられる。よって、

$$\begin{aligned}
&\begin{pmatrix} 4x_1 + x_2 + 1 \\ x_1 + 2x_2 - 3 \end{pmatrix} + u_1 \begin{pmatrix} 2x_1 \\ 2x_2 \end{pmatrix} + u_2 \begin{pmatrix} -1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \\
&x_1^2 + x_2^2 - 1 = 0 \\
&-x_1 = 0 \\
&u_3 = 0
\end{aligned}$$

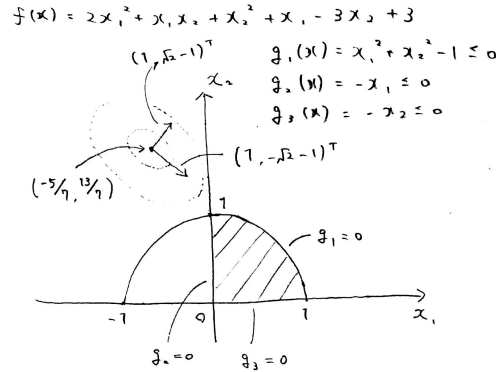


図 4.1: 実行可能領域と f の等高線

より，実行可能解で 1 次の最適性必要条件を満たすものは，

$$(\bar{x}_1, \bar{x}_2, u_1, u_2) = \left(0, 1, \frac{1}{2}, 2\right)$$

また，

$$\begin{pmatrix} 4 & 1 \\ 1 & 2 \end{pmatrix} + \frac{1}{2} \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} + 2 \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 5 & 1 \\ 1 & 3 \end{pmatrix}$$

は正定より，2 次の必要条件も満たす．さらに，2 次の十分条件を満たすことより，この解は局所最適解である．

以上より，この問題の局所最適解は，

$$\bar{x} = (0, 1)^T$$

であり，その目的関数値は 1 である． □

KKT 条件の幾何学的解釈をこの問題を例に示す． $\bar{x} = (0, 1)^T$ における目的関数および有効制約の勾配ベクトルは，

$$\nabla f(\bar{x}) = (2, -1)^T$$

$$\nabla g_1(\bar{x}) = (0, 2)^T$$

$$\nabla g_2(\bar{x}) = (-1, 0)^T$$

である．図 4.2 にこのベクトルを \bar{x} を始点に描いたものを示す． $\nabla g_i(\bar{x})$ は， \bar{x} における $g_i(x)$ の増加方向を示す．これは，有効制約 $g_i(x) = 0$ の \bar{x} における接線と直交する．また， $\nabla f(\bar{x})$ は， \bar{x} における $f(x)$ の増加方向を示す． $f(x)$ を最小化するためには，逆方向 $-\nabla f(\bar{x})$ に進むのが望ましい． $-\nabla f(\bar{x})$ と鋭角をなす方向へ進むことができれば，目的関数値は減少する．

$\nabla g_i(\bar{x})$ を，非負の重み $u_i \geq 0$ によって加えたベクトルの領域

$$G(\bar{x}) = \{u_1 \nabla g_1(\bar{x}) + u_2 \nabla g_2(\bar{x}) \mid u_1 \geq 0, u_2 \geq 0\}$$

を定義する．

KKT 条件は，ベクトル $-\nabla f(\bar{x})$ がこの領域内に入ることを要求している．つまり，この条件は， \bar{x} から実行可能領域のどの方向を選んでも $-\nabla f(\bar{x})$ と鋭角をなすことができず，目的関数値を減少させることができないことを表す．つまり， \bar{x} の局所最適性の必要条件となっている．

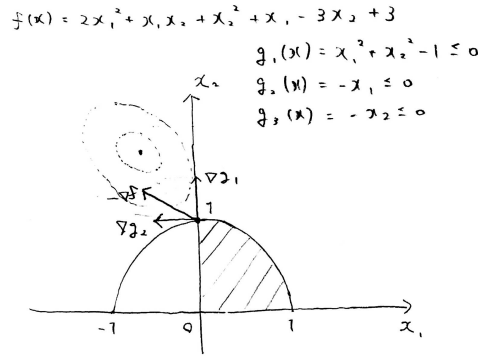


図 4.2: KKT 条件の幾何学的説明

4.4 凸計画問題

最適化問題 (3.1) において，実行可能領域 $S \cap X$ が凸集合であり，目的関数 f が凸関数であるものを凸計画問題という．凸計画問題に対して，次の定理が成立する．

定理 4.8 最適化問題

$$\begin{aligned} & \text{minimize} : f(\mathbf{x}) \\ & \text{subject to} : \mathbf{x} \in S \cap X \end{aligned}$$

が凸計画問題であれば，任意の局所最適解は大域最適解である．さらに，目的関数 f が狭義凸関数のとき，大域最適解が存在すれば唯一である．

線形計画問題は，実行可能領域が凸多面体であり，目的関数が凸関数であるため，凸計画問題である．非線形計画問題について，実行可能領域が

$$S = \{\mathbf{x} \in \mathbb{R}^n \mid g_i(\mathbf{x}) \leq 0, i = 1, 2, \dots, m\} \quad (4.19)$$

である場合を考える．このとき， $g_i(\mathbf{x})$ が凸関数であれば，

$$g_i(\mathbf{x}) \leq 0$$

を満たす \mathbf{x} の領域は凸集合である*7．

等式制約 $g_i(\mathbf{x}) = 0$ は， $g_i(\mathbf{x}) \leq 0$ と $-g_i(\mathbf{x}) \leq 0$ を合わせたものである． $g_i(\mathbf{x})$ と $-g_i(\mathbf{x})$ がともに凸関数であるのは，1 次式 $g_i(\mathbf{x}) = \mathbf{a}^T \mathbf{x} - b$ に限る．

*7 この条件を満たす任意の 2 点 $\mathbf{x}^1, \mathbf{x}^2$ を結ぶ線分上の点 $\mathbf{x}_\alpha = (1 - \alpha)\mathbf{x}^1 + \alpha\mathbf{x}^2$ は， g_i の凸性より，

$$g_i(\mathbf{x}_\alpha) = g_i((1 - \alpha)\mathbf{x}^1 + \alpha\mathbf{x}^2) \leq (1 - \alpha)g_i(\mathbf{x}^1) + \alpha g_i(\mathbf{x}^2) \leq 0$$

を満たす．補題 2.1 より，(4.19) は凸集合である．

□

第 5 章

線形計画問題

本章では，線形計画問題の定式化，最適性，およびアルゴリズムについて述べる．

5.1 線形計画問題の例

線形計画問題の例として，しばしば生産計画問題が挙げられる．例えば，次のような問題である．

生産計画問題 A 社では，2 種類の製品 P_1, P_2 を同じ原料から生産している． P_1 1 トン生産するのに要する原料は 2.5 トン，電力は 5kwh，労力は 3 人時， P_2 1 トン生産するのに要する原料は 5 トン，電力は 6kwh，労力は 2 人時である．1 日の原料，電力，労力の使用可能量は，それぞれ 350 トン，450kwh，240 人時である． P_1, P_2 1 トン当たりの粗利益は，それぞれ，4 万円，5 万円である．このとき，総粗利益を最大にするには， P_1, P_2 を何トンずつ生産すればよいか．

P_1, P_2 の 1 日の生産量 (トン) をそれぞれ x_1, x_2 とする．この問題は表 5.1 のように表される．

表 5.1: 生産計画問題

| | P_1 | P_2 | |
|-----|----------|--------|-----|
| 原料 | $2.5x_1$ | $5x_2$ | 350 |
| 電力 | $5x_1$ | $6x_2$ | 450 |
| 労力 | $3x_1$ | $2x_2$ | 240 |
| 粗利益 | $4x_1$ | $5x_2$ | |

以上より，この生産計画問題は次の線形計画問題となる．

$$\begin{aligned} & \text{maximize} && 4x_1 + 5x_2 \\ & \text{subject to} && 2.5x_1 + 5x_2 \leq 350 \\ & && 5x_1 + 6x_2 \leq 450 \\ & && 3x_1 + 2x_2 \leq 240 \\ & && x_1, x_2 \geq 0 \end{aligned} \tag{5.1}$$

5.2 線形計画問題の最適解

4.4 節で述べたように、線形計画問題の実行可能領域は凸多面体であり、目的関数も線形 (凸関数) であるので、線形計画問題は凸計画問題である。よって、任意の局所最適解は大域最適解である。

線形計画問題の最適解は、実行可能領域の多面体の頂点に存在する。この事実を確認する。一般に多面体 X は、すべての頂点からなる集合 $V = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_p\}$ を用いて、次のように表すことができる。

$$X = \left\{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{x} = \sum_{i=1}^p k_i \mathbf{v}_i, \sum_{i=1}^p k_i = 1, k_i \geq 0, i = 1, 2, \dots, p \right\}$$

また、線形計画問題の目的関数は、 $\mathbf{c}^T \mathbf{x}$ と表される。実行可能解 $\mathbf{x} \in X$ を考えると、 $\sum_{i=1}^p k_i = 1, k_i \geq 0, i = 1, 2, \dots, p$ を満たす $(k_1, k_2, \dots, k_p)^T$ が存在して、

$$\mathbf{x} = \sum_{i=1}^p k_i \mathbf{v}_i$$

と表される。両辺に左から \mathbf{c}^T を掛けると、

$$\mathbf{c}^T \mathbf{x} = \mathbf{c}^T \left(\sum_{i=1}^p k_i \mathbf{v}_i \right) = \sum_{i=1}^p k_i \mathbf{c}^T \mathbf{v}_i$$

となる。 $\sum_{i=1}^p k_i = 1, k_i \geq 0, i = 1, 2, \dots, p$ より、

$$\min_{i=1,2,\dots,p} \mathbf{c}^T \mathbf{v}_i \leq \mathbf{c}^T \mathbf{x} \leq \max_{i=1,2,\dots,p} \mathbf{c}^T \mathbf{v}_i$$

となる。 $i = 1, 2, \dots, p$ について、 $\mathbf{v}_i \in X$ より、実行可能領域 X での $\mathbf{c}^T \mathbf{x}$ の最大値および最小値を与える解の 1 つは、頂点集合 V 内にある。

5.3 標準形の線形計画問題

5.3.1 線形計画問題の標準形

変数が非負で等式制約条件しかもたない最小化問題を標準形 (standard form) の線形計画問題という。すなわち、線形計画問題の標準形は、

$$\begin{aligned} & \text{minimize } \mathbf{c}^T \mathbf{x} \\ & \text{subject to } A\mathbf{x} = \mathbf{b} \\ & \mathbf{x} \geq \mathbf{0} \end{aligned} \tag{5.2}$$

である。ただし、 $\mathbf{c} = (c_1, c_2, \dots, c_n)^T$, $\mathbf{b} = (b_1, b_2, \dots, b_m)^T$ はそれぞれ n 次元、 m 次元の実数定ベクトルである。 A は $m \times n$ ($m \leq n$) の実数定行列である。 $\mathbf{x} = (x_1, x_2, \dots, x_n)^T$ は n 次元の実数変数ベクトルであり、決定変数ベクトルと呼ばれる。

5.3.2 標準形への変形

最大化問題の場合 線形計画問題が最大化問題の場合、つまり、

$$\text{maximize } \mathbf{c}^T \mathbf{x}$$

の場合、 $\mathbf{d} = -\mathbf{c}$ とすると、

$$\text{maximize } \mathbf{c}^T \mathbf{x} \Leftrightarrow \text{minimize } \mathbf{d}^T \mathbf{x}$$

と最小化問題に変形できる。

不等式条件 \leq の場合 制約条件の 1 つが、

$$\mathbf{a}_i^T \mathbf{x} \leq b_i$$

である場合、スラック変数 (slack variable) s ($s \geq 0$) を導入して、

$$\mathbf{a}_i^T \mathbf{x} \leq b_i \Leftrightarrow \mathbf{a}_i^T \mathbf{x} + s = b_i, s \geq 0$$

と等式条件に変形される。

不等号条件 \geq の場合 制約条件の 1 つが、

$$\mathbf{a}_j^T \mathbf{x} \geq b_j$$

である場合、サープラス変数 (surplus variable) t ($t \geq 0$) を導入して、

$$\mathbf{a}_j^T \mathbf{x} \geq b_j \Leftrightarrow \mathbf{a}_j^T \mathbf{x} - t = b_j, t \geq 0$$

と等式条件に変形される。

自由変数 (非負条件の無い変数) が存在する場合 x_k に対して、 $x_k \geq 0$ という条件が存在せず、負の値も取る場合は、 x_k を次の条件を満たす 2 つの変数 x_k^+ と x_k^- に分解して取り扱われる。

$$x_k = x_k^+ - x_k^-, x_k^+ \geq 0, x_k^- \geq 0$$

例えば、最適化問題 (5.1) を標準形の線形計画問題に変形すると、(5.3) になる。

$$\begin{aligned} & \text{minimize } -4x_1 - 5x_2 \\ & \text{subject to } 2.5x_1 + 5x_2 + x_3 = 350 \\ & \quad 5x_1 + 6x_2 + x_4 = 450 \\ & \quad 3x_1 + 2x_2 + x_5 = 240 \\ & \quad x_1, x_2, x_3, x_4, x_5 \geq 0 \end{aligned} \tag{5.3}$$

5.4 基底解と最適解

5.4.1 用語

基底行列 基底行列 (basic matrix) B とは、制約条件の係数行列 A の m 個 (制約条件数) の列ベクトルから作られる $m \times m$ の正則行列のことをいう ($|B| \neq 0$)。

基底解 線形計画問題の基底解 (basic solution) とは, ある基底行列 B で選ばれなかった $(n - m)$ 個の列に対応する変数 (非基底変数 (nonbasic variable)) を 0 とおき, 選ばれた m 個の変数 (基底変数 (basic variable)) に対する正則な連立方程式

$$B \begin{pmatrix} x_1^B \\ x_2^B \\ \vdots \\ x_m^B \end{pmatrix} = \mathbf{b} \quad (5.4)$$

を解くことにより得られる一意的なベクトル $B^{-1}\mathbf{b}$ のことをいう.

実行可能基底解 実行可能解であり, 基底解である解を実行可能基底解 (basic feasible solution) という. 基底行列を B とすると, $B^{-1}\mathbf{b} \geq \mathbf{0}$ のときに基底行列 B に対応する基底解は実行可能基底解となる.

退化 基底解では, m 個の基底変数のみが正の値を取りうるが, この m 個の基底変数の内少なくとも 1 つが 0 となる基底解は, 退化しているという.

5.4.2 線形計画問題の基本定理

標準形の線形計画問題 (5.2) において, 次の定理が成立する.

定理 5.1 (線形計画問題の基本定理) 標準形の線形計画問題 (5.2) が与えられたとき, 次の 1, 2 が成立する.

1. 実行可能解が存在するならば, 必ず実行可能基底解が存在する.
2. 最適解が存在するならば, 実行可能基底解の中に最適解が存在する.

定理 5.1 より, 線形計画問題の最適解を求めるには, 基底解の中で, 実行可能かつ最適なものを見つければよい*1.

5.4.3 基底形式

標準形の線形計画問題

$$\begin{aligned} &\text{minimize } z = \mathbf{c}^T \mathbf{x} \\ &\text{subject to } A\mathbf{x} = \mathbf{b} \\ &\mathbf{x} \geq \mathbf{0} \end{aligned} \quad (5.5)$$

において, $m \times n$ 行列 A から m 個の列を選び基底行列 B とする. ただし, B は正則行列とする. 残りの列からなる行列を N とする. 列の順番と変数の順番を適当に変えることにより, $A = (B, N)$, $\mathbf{x} = (\mathbf{x}_B^T, \mathbf{x}_N^T)^T$, $\mathbf{c} = (\mathbf{c}_B^T, \mathbf{c}_N^T)^T$ と表すと, 式 (5.5) は,

$$\begin{aligned} &\text{minimize } z = \mathbf{c}_B^T \mathbf{x}_B + \mathbf{c}_N^T \mathbf{x}_N \\ &\text{subject to } B\mathbf{x}_B + N\mathbf{x}_N = \mathbf{b} \\ &\mathbf{x}_B, \mathbf{x}_N \geq \mathbf{0} \end{aligned} \quad (5.6)$$

*1 実行可能領域の端点を辿っていけば最適解に到達する. 後述のシンプレックス法はこの考えに基づいている.

と表される。 \mathbf{x}_B は基底変数ベクトル、 \mathbf{x}_N は非基底変数ベクトル、 \mathbf{x}_B の各成分 x_j^B は基底変数、 \mathbf{x}_N の各成分 x_j^N は非基底変数と呼ばれる。 B は正則より、 目的関数から \mathbf{x}_B を消去すると、 (5.6) の問題は、

$$\begin{aligned} & \text{minimize } z = \mathbf{c}_B^T B^{-1} \mathbf{b} + (\mathbf{c}_N^T - \mathbf{c}_B^T B^{-1} N) \mathbf{x}_N \\ & \text{subject to } \mathbf{x}_B + B^{-1} N \mathbf{x}_N = B^{-1} \mathbf{b} \\ & \mathbf{x}_B, \mathbf{x}_N \geq \mathbf{0} \end{aligned} \quad (5.7)$$

と等価になる。 (5.7) の一部

$$\begin{aligned} z &= \mathbf{c}_B^T B^{-1} \mathbf{b} + (\mathbf{c}_N^T - \mathbf{c}_B^T B^{-1} N) \mathbf{x}_N \\ \mathbf{x}_B + B^{-1} N \mathbf{x}_N &= B^{-1} \mathbf{b} \end{aligned} \quad (5.8)$$

は、基底形式 (basic form) あるいは正準形 (cannoncial form) と呼ばれる。 このとき、基底行列は B であり、対応する基底解は

$$\mathbf{x} = \begin{pmatrix} \mathbf{x}_B \\ \mathbf{x}_N \end{pmatrix} = \begin{pmatrix} B^{-1} \mathbf{b} \\ \mathbf{0} \end{pmatrix} \quad (5.9)$$

となる。 以後、 $\bar{\mathbf{c}}^T = \mathbf{c}_N^T - \mathbf{c}_B^T B^{-1} N$, $\bar{N} = B^{-1} N$, $\bar{\mathbf{b}} = B^{-1} \mathbf{b}$ と定める。 また、 \mathbf{c}_N^T , \bar{N} , $\bar{\mathbf{b}}$ の各成分をそれぞれ、 \bar{c}_j , \bar{a}_{ij} , \bar{b}_i と表す。

(5.7) 式において、

$$\bar{\mathbf{b}} = B^{-1} \mathbf{b} \geq \mathbf{0} \quad (\text{実行可能性}) \quad (5.10)$$

が成立すれば、 (5.9) 式の基底解は実行可能解となる。 (5.10) が満たされると仮定する。

基底形式 (5.8) のある非基底変数 x_r^N の係数 \bar{c}_r が負であれば、この非基底変数に正の値を取らせることにより、目的関数が改善できる*2。 逆に、すべての非基底変数 x_j^N の係数 \bar{c}_j が正、すなわち、

$$\bar{\mathbf{c}}^T = \mathbf{c}_N^T - \mathbf{c}_B^T B^{-1} N \geq \mathbf{0}^T \quad (\text{最適性}) \quad (5.11)$$

が成立すれば、この実行可能基底解は最適解の1つである。

次に、 \bar{c}_r が負であるとき、実行可能性を維持する、つまり、 x_r^N に正の値を取らせることが可能であるかどうかを考える。 ここで、 x_r^N 以外の非基底変数 x_j^N の値を0に固定する。 このとき、

$$\mathbf{x}_B = \bar{\mathbf{b}} - \begin{pmatrix} \bar{a}_{1r} \\ \bar{a}_{2r} \\ \vdots \\ \bar{a}_{mr} \end{pmatrix} x_r^N \quad (5.12)$$

を満たせば、制約条件の最初の等式条件 $\mathbf{x}_B + B^{-1} N \mathbf{x}_N = B^{-1} \mathbf{b}$ を満足する*3。

*2 $\bar{c}_r < 0$ ならば、 $x_r^N > 0$ を取ることで、実行可能性を維持したまま (5.8) 式の z の値を小さくすることができる。

*3 このとき、

第 6 章

証明

この章には、補題や定理の証明 (一部) を記載する。

6.1 2 章の証明

証明 1 (補題 2.1 の証明) 項が 2 つのとき、つまり、 $X_1, X_2 \subseteq \mathbb{R}^n$ が凸集合ならば、 $X_1 \cap X_2$ も凸集合である (*) を証明する。 $\forall \mathbf{x}^1, \mathbf{x}^2 \in X_1 \cap X_2$ に対し、 \mathbf{x}^1 と \mathbf{x}^2 を結ぶ線分は、 $\mathbf{x}^1 \in X_1, \mathbf{x}^2 \in X_1$ より X_1 に含まれる ($\because X_1$ は凸集合)。同様に X_2 にも含まれる。よって、 \mathbf{x}^1 と \mathbf{x}^2 を結ぶ線分は、 $X_1 \cap X_2$ に含まれる。つまり、 $X_1 \cap X_2$ は凸集合である。補題 2.1 は、(*) を繰り返し用いることで示される。 \square

補題 2.2 の証明には、定理 6.1、補題 6.1 を用いる。

定理 6.1 (射影定理) X を Hilbert 空間^{*1}とし、 $L \subset X$ を閉部分空間とする。このとき、

$$\mathbf{u} \in X \text{ (given)} \Rightarrow \exists! \mathbf{v} \in L \text{ s.t. } (\mathbf{u} - \mathbf{v}, \mathbf{w}) = 0, \forall \mathbf{w} \in L$$

が成立する (射影が一意に存在する)。

補題 6.1 定理 6.1 の X が $X = \mathbb{R}^n$ であり、 $L \subseteq \mathbb{R}^n$ が空でない閉凸集合とする。 $\mathbf{u} \in \mathbb{R}^n$ の L への射影を \mathbf{v} とする。このとき、 $\forall \mathbf{w} \in L$ に対して、

$$(\mathbf{u} - \mathbf{v})^T (\mathbf{w} - \mathbf{v}) \leq 0 \quad (6.1)$$

が成立する。

証明 2 (補題 6.1 の証明) \mathbf{u} の L への射影 \mathbf{v} と任意の点 $\mathbf{w} \in L$ を結ぶ線分上の点 $\mathbf{x}_\lambda = (1-\lambda)\mathbf{v} + \lambda\mathbf{w}$, $0 < \lambda < 1$ を考える。 L は凸より、 \mathbf{x}_λ も L に含まれる。 \mathbf{v} の定義より、

$$\|\mathbf{v} - \mathbf{u}\|^2 \leq \|((1-\lambda)\mathbf{v} + \lambda\mathbf{w}) - \mathbf{u}\|^2 \quad (6.2)$$

^{*1} 距離 (ノルム) を持つ集合をノルム空間という。内積を持つ線形空間を内積空間という。ノルム空間 X 内の任意のコーシー列が収束するとき、 X は完備であるといい、完備性を持つノルム空間 X を Banach 空間という。また、内積空間 X 上の点 $\mathbf{u} \in X$ に対し、 $\|\mathbf{u}\| = \sqrt{(\mathbf{u}, \mathbf{u})}$ を内積から誘導されるノルムと呼ぶ。内積から誘導されるノルム空間 X が Banach 空間であるとき、 X を Hilbert 空間という。実数空間 \mathbb{R}^n は完備性を持つ。

(6.2) 式を整理すると,

$$\begin{aligned} (v - u, v - u) &\leq \|(v - u) + \lambda(w - v)\|^2 \\ &\leq (v - u, v - u) + 2\lambda(v - u)(w - v) + \lambda^2(w - v, w - v) \end{aligned}$$

いま, $\lambda \neq 0, \lambda > 0$ であり, $L \subseteq \mathbb{R}^n$ であるから,

$$2(u - v)(w - v) \leq \lambda \|w - v\|^2$$

よって, $\lambda \rightarrow 0$ とすると, (6.1) 式を得る. \square

証明 3 (補題 2.2 の証明) $y \in \mathbb{R}^n$ の X への射影を \bar{y} とする. 定理 6.1 より, \bar{y} は y に対して一意に存在する. $y \notin X$ より, $y \neq \bar{y}$ である. よって,

$$\|\bar{y} - y\|^2 = (\bar{y} - y)^T(\bar{y} - y) > 0 \quad (6.3)$$

である. また, 補題 6.1 より,

$$(y - \bar{y})^T(x - \bar{y}) \leq 0, x \in X \quad (6.4)$$

が成立する. (6.3) 式, (6.4) 式より,

$$(\bar{y} - y)^T x \geq (\bar{y} - y)^T \bar{y} > (\bar{y} - y)^T y, x \in X$$

を得る. ここで, $a = \bar{y} - y, b = (\bar{y} - y)^T \bar{y}$ と置くと, \bar{y} の一意性より補題 2.2 が示される. \square

証明 4 (補題 2.3 の証明) y は X の内点でないので, y の近傍 $B_k = \{x \in \mathbb{R}^n \mid \|x - y\| < \varepsilon_k\}$ はある点 $x^k \notin \text{Cl}(X)$ を含む. $k \rightarrow \infty$ に伴い, $\varepsilon_k \rightarrow 0$ とすると, 点列 $\{x^k\}$ は $\text{Cl}(X)$ の外部から y に収束する. $x^k \notin \text{Bd}(X)$ から, 補題 2.2 より

$$(a^k)^T x > (a^k)^T x^k, \forall x \in X$$

なる $a^k \in \mathbb{R}^n$ が存在する. $a^k \neq 0$ より, $\|a^k\| = 1$ を仮定する. 点列 $\{a^k\}$ の極限点を a とすると,

$$a^T x \geq a^T y, \forall x \in X$$

である. $\|a\| = 1$ より, $a \neq 0$ である. \square

証明 5 (補題 2.4 の証明) 集合 X, Y から集合 $S = \{z \in \mathbb{R}^n \mid z = x - y, x \in X, y \in Y\}$ を定義すると, S も凸集合である^{*2}. $X \cap Y = \emptyset$ より, 0 ベクトルは $0 \notin S$ である. 補題 2.2, 補題 2.3 より,

$$a^T z \geq 0, \forall z \in S$$

を満たす $a \neq 0$ が存在する. よって, S の定義より,

$$a^T x \geq a^T y, \forall x \in X, \forall y \in Y$$

を得る. よって, $b = \inf_{x \in X} a^T x$ とおくと, (2.1) 式を得る. \square

^{*2} 2 点 $z^1 = x^1 - y^1, z^2 = x^2 - y^2 \in S, x^1, x^2 \in X, y^1, y^2 \in Y$ を結ぶ線分上の点の考えると,

$$\begin{aligned} (1 - \lambda)z^1 + \lambda z^2 &= (1 - \lambda)(x^1 - y^1) + \lambda(x^2 - y^2) \\ &= ((1 - \lambda)x^1 + \lambda x^2) - ((1 - \lambda)y^1 + \lambda y^2) \end{aligned}$$

と書けるため, X と Y の凸性より, やはり S に属するから.

6.2 4章の証明

証明 6 (定理 4.1 の証明) $\mathbf{d} \in \mathbb{R}^n, \varepsilon > 0$ を考える. ε が十分小さければ, 局所最適性より,

$$f(\bar{\mathbf{x}} + \varepsilon \mathbf{d}) \geq f(\bar{\mathbf{x}}), f(\bar{\mathbf{x}} - \varepsilon \mathbf{d}) \geq f(\bar{\mathbf{x}})$$

が成立する. Taylor の定理より,

$$f(\bar{\mathbf{x}} + \varepsilon \mathbf{d}) = f(\bar{\mathbf{x}}) + \varepsilon \nabla f(\bar{\mathbf{x}})^T \mathbf{d} + o(\varepsilon \|\mathbf{d}\|) \geq f(\bar{\mathbf{x}})$$

$$f(\bar{\mathbf{x}} - \varepsilon \mathbf{d}) = f(\bar{\mathbf{x}}) - \varepsilon \nabla f(\bar{\mathbf{x}})^T \mathbf{d} + o(\varepsilon \|\mathbf{d}\|) \geq f(\bar{\mathbf{x}})$$

なので,

$$|\nabla f(\bar{\mathbf{x}}) \mathbf{d}| = \frac{o(\varepsilon \|\mathbf{d}\|)}{\varepsilon} \rightarrow 0 \quad (\varepsilon \rightarrow 0)$$

より, $\nabla f(\bar{\mathbf{x}}) = 0$ □

証明 7 (定理 4.2 の証明) 定理 4.1 より, $\nabla f(\bar{\mathbf{x}}) = 0$. 任意の $\mathbf{d} \in \mathbb{R}^n, \varepsilon > 0$ に対して,

$$f(\bar{\mathbf{x}} + \varepsilon \mathbf{d}) \geq f(\bar{\mathbf{x}}) \quad (\because \bar{\mathbf{x}} \text{ が局所最適解})$$

Taylor の定理と $\nabla f(\bar{\mathbf{x}}) = 0$ より,

$$\frac{1}{2} \mathbf{d}^T \nabla^2 f(\bar{\mathbf{x}}) \mathbf{d} + \frac{o(\varepsilon^2 \|\mathbf{d}\|^2)}{\varepsilon^2} \geq 0$$

が成立する. よって, $\varepsilon \rightarrow 0$ として $\mathbf{d}^T H(\bar{\mathbf{x}}) \mathbf{d} \geq 0$ を得る. □

証明 8 (定理 4.3 の証明) $f(\bar{\mathbf{x}} + \varepsilon \mathbf{d})$ に対して, Taylor の定理より,

$$f(\bar{\mathbf{x}} + \varepsilon \mathbf{d}) - f(\bar{\mathbf{x}}) = \varepsilon \nabla f(\bar{\mathbf{x}})^T \mathbf{d} + \frac{\varepsilon^2}{2} \mathbf{d}^T \nabla^2 f(\bar{\mathbf{x}}) \mathbf{d} + o(\varepsilon^2 \|\mathbf{d}\|^2)$$

$\nabla f(\bar{\mathbf{x}}) = 0$ を代入し, 両辺を ε^2 で割ると, 右辺は,

$$\frac{1}{2} \mathbf{d}^T \nabla^2 f(\bar{\mathbf{x}}) \mathbf{d} + \frac{o(\varepsilon^2 \|\mathbf{d}\|^2)}{\varepsilon^2} \tag{6.5}$$

$H(\bar{\mathbf{x}})$ が正定値であり, $\varepsilon \rightarrow 0$ であれば, 第二項は 0 に収束する. よって, ε が十分小さく, $\varepsilon \neq 0$ であれば, (6.5) 式は正となる. $\varepsilon = 0$ を含めて, $f(\bar{\mathbf{x}} + \varepsilon \mathbf{d}) \geq f(\bar{\mathbf{x}})$ が成立する. つまり, $\bar{\mathbf{x}}$ は局所最適解である. □

証明 9 (定理 4.4 の証明) (1 次の必要条件)

ペナルティ関数 $F^k(\mathbf{x})$ に関する議論により, (4.8) の局所最適解 \mathbf{x}^k に定理 4.1 を適用できる.

$$F^k(\mathbf{x}) = f(\mathbf{x}) + \frac{k}{2} \|g(\mathbf{x})\|^2 + \frac{\alpha}{2} \|\mathbf{x} - \bar{\mathbf{x}}\|^2$$

$$g(\mathbf{x}) = (g_1(\mathbf{x}), \dots, g_m(\mathbf{x}))^T$$

より,

$$\nabla F^k(\mathbf{x}^k) = \nabla f(\mathbf{x}^k) + k \nabla g(\mathbf{x}^k)^T g(\mathbf{x}^k) + \alpha(\mathbf{x} - \bar{\mathbf{x}}) = 0 \tag{6.6}$$

を得る。 $\bar{\mathbf{x}}$ は正則であるから、 k が十分大きければ \mathbf{x}^k も正則。 よって、

$$\text{rank}(\nabla g(\mathbf{x}^k)^T) = m, \text{rank}(\nabla g(\mathbf{x}^k)) = m$$

が成立する。 よって、 m 次正方行列 $\nabla g(\mathbf{x}^k)\nabla g(\mathbf{x}^k)^T$ も正則である。 (6.6) の両辺に左から $(\nabla g(\mathbf{x}^k)\nabla g(\mathbf{x}^k)^T)^{-1}\nabla g(\mathbf{x}^k)$ を乗じて整理すると、

$$kg(\mathbf{x}^k) = -(\nabla g(\mathbf{x}^k)\nabla g(\mathbf{x}^k)^T)^{-1}\nabla g(\bar{\mathbf{x}})(\nabla f(\mathbf{x}^k) + \alpha(\mathbf{x} - \bar{\mathbf{x}}))$$

を得る。 $k \rightarrow \infty$ とすると、 $\{\mathbf{x}^k\}$ は $\bar{\mathbf{x}}$ に収束するので、 点列 $\{kg(\mathbf{x}^k)\}$ は、 m 次元ベクトル

$$\mathbf{u} = -(\nabla g(\mathbf{x}^k)\nabla g(\mathbf{x}^k)^T)^{-1}\nabla g(\bar{\mathbf{x}})\nabla f(\bar{\mathbf{x}})$$

に収束する。 (6.6) において、 $k \rightarrow \infty$ とすると、

$$\nabla f(\bar{\mathbf{x}}) + \nabla g(\bar{\mathbf{x}})^T \mathbf{u} = \mathbf{0}$$

を得る。 □

(2 次の必要条件)

(4.8) の局所最適解 \mathbf{x}^k に定理 4.2 を適用できる。

$$\nabla^2 F^k(\mathbf{x}^k) = \nabla^2 f(\mathbf{x}^k) + k\nabla g(\mathbf{x}^k)^T \nabla g(\mathbf{x}^k) + k \sum_{i=1}^m g_i(\mathbf{x}^k) \nabla^2 g_i(\mathbf{x}^k) + \alpha I$$

は半正定値である。 ただし、 I は単位行列を示す。 $\mathbf{y} \in V(\bar{\mathbf{x}})$ を任意に選び、 これに基づき、

$$\mathbf{y}^k = \mathbf{y} - \nabla g(\mathbf{x}^k)^T (\nabla g(\mathbf{x}^k)\nabla g(\mathbf{x}^k)^T)^{-1} \nabla g(\mathbf{x}^k) \mathbf{y} \quad (6.7)$$

とする。 \mathbf{y}^k は、

$$\nabla g(\mathbf{x}^k) \mathbf{y}^k = \mathbf{0}$$

を満たす^{*3}。 $k \rightarrow \infty$ とすると、 $\mathbf{x}^k \rightarrow \bar{\mathbf{x}}$ であり、 \mathbf{y} の定義より、 $\nabla g(\bar{\mathbf{x}}) = \mathbf{0}$ である。 よって、 (6.7) の第二項は $\mathbf{0}$ に収束し、 $\mathbf{y}^k \rightarrow \mathbf{y}$ を得る。

次に、 $\nabla^2 F^k(\mathbf{x}^k)$ の半正定値性と $\nabla g(\mathbf{x}^k) \mathbf{y}^k = \mathbf{0}$ を用いて、

$$(\mathbf{y}^k)^T \nabla^2 F^k(\mathbf{x}^k) \mathbf{y}^k = (\mathbf{y}^k)^T \left(\nabla^2 f(\mathbf{x}^k) + k \sum_{i=1}^m g_i(\mathbf{x}^k) \nabla^2 g_i(\mathbf{x}^k) \right) \mathbf{y}^k + \alpha \|\mathbf{y}^k\|^2 \geq 0$$

を得る。 $k \rightarrow \infty$ とすると、

$$\mathbf{y}^k \rightarrow \mathbf{y}, \mathbf{x}^k \rightarrow \bar{\mathbf{x}}, kg_i(\mathbf{x}^k) \rightarrow u_i$$

となる。 よって、

$$(\mathbf{y})^T \left(\nabla^2 f(\bar{\mathbf{x}}) + \sum_{i=1}^m u_i \nabla^2 g_i(\bar{\mathbf{x}}) \right) \mathbf{y} + \alpha \|\mathbf{y}\|^2 \geq 0, \forall \mathbf{y} \in V(\bar{\mathbf{x}})$$

となる。 $\alpha > 0$ は任意より、 $\alpha \rightarrow 0$ として、 (4.10) を得る。 □

^{*3} \mathbf{y}^k は \mathbf{y} の空間 $V(\mathbf{x}^k)$ への射影点である。

証明 10 (定理 4.8 の証明) 局所最適解 \mathbf{x}' が大域最適解ではなく、それとは異なる大域最適解 $\bar{\mathbf{x}}$ が存在したとする。このとき、 $f(\bar{\mathbf{x}}) < f(\mathbf{x}')$ を満たす。 \mathbf{x}' の任意の近傍 B と \mathbf{x}' と $\bar{\mathbf{x}}$ を結ぶ線分上の点 $\mathbf{x}_\alpha = (1 - \alpha)\mathbf{x}' + \alpha\bar{\mathbf{x}}, 0 \leq \alpha \leq 1$ を考えると、実行可能領域の凸性より、 \mathbf{x}_α も実行可能解である。十分小さな $\alpha > 0$ では、 $\mathbf{x}_\alpha \in B$ が成立する。 f の凸性より、

$$f(\mathbf{x}_\alpha) \leq (1 - \alpha)f(\mathbf{x}') + \alpha f(\bar{\mathbf{x}}) < f(\mathbf{x}')$$

である。これは、 \mathbf{x}' が局所最適解であることに矛盾する。

f が狭義凸関数であるとし、異なる大域最適解 $\bar{\mathbf{x}}, \bar{\bar{\mathbf{x}}}$ が存在することを仮定する。このとき、 $f(\bar{\mathbf{x}}) = f(\bar{\bar{\mathbf{x}}})$ である。 $\bar{\mathbf{x}}$ と $\bar{\bar{\mathbf{x}}}$ を結ぶ線分上の点 $\mathbf{x}_\alpha = (1 - \alpha)\bar{\mathbf{x}} + \alpha\bar{\bar{\mathbf{x}}}, 0 < \alpha < 1$ を考えると、実行可能領域の凸性より、 \mathbf{x}_α も実行可能解。 f の狭義凸性より、

$$f(\mathbf{x}_\alpha) < (1 - \alpha)f(\bar{\mathbf{x}}) + \alpha f(\bar{\bar{\mathbf{x}}}) = f(\bar{\mathbf{x}})$$

これは、 $\bar{\mathbf{x}}$ が大域最適解であることに矛盾する。 □

参考文献

- [1] オペレーションズ・リサーチとは. 公益社団法人日本オペレーションズ・リサーチ学会.
<https://www.orsj.or.jp/whatisor/whatisor.html>. (参照 2020/3/23)
- [2] 茨木俊秀. (2011). 最適化の数学. 共立出版株式会社.