

# 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 . . . . .	6
第 3 章	証明	7
参考文献		9

# 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.

- $\varepsilon$ -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,

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

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,

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

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\*<sup>1</sup>.

(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  $\phi$  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\*<sup>2</sup>.

補題 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 (分離超平面の存在)  $X \subseteq \mathbb{R}^n$  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 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 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、

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

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

## 2.4 凸関数

\*<sup>1</sup> 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.

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

## 第 3 章

# 証明

この付録には、補題や定理の証明を記載する。

**証明 1 (補題 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 1 は、(\*) を繰り返し用いることで示される。  $\square$

補題 2 の証明には、定理 1、補題 5 を用いる。

**定理 1 (射影定理)**  $X$  を Hilbert 空間<sup>\*1</sup>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 5** 定理 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、

$$(\mathbf{u} - \mathbf{v})^T (\mathbf{w} - \mathbf{v}) \leq 0 \quad (3.1)$$

が成立する。

**証明 2 (定理 1 の証明)**

**証明 3 (補題 5 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、 $\mathbf{x}_\lambda$  も  $L$  に含まれる。 $\mathbf{v}$  の定義より、

$$\|\mathbf{v} - \mathbf{u}\|^2 \leq \|((1 - \lambda)\mathbf{v} + \lambda\mathbf{w}) - \mathbf{u}\|^2 \quad (3.2)$$

---

<sup>\*1</sup> 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。



(3.2) 式を整理すると,

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

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

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

よって,  $\lambda \rightarrow 0$  とすると, (3.1) 式を得る. □

証明 4 (補題 2 の証明)  $\boldsymbol{y}$  の  $X$  への射影を  $\bar{\boldsymbol{y}}$  とする.  $\boldsymbol{y} \notin X$  より,  $\boldsymbol{y} \neq \bar{\boldsymbol{y}}$  である. よって,

$$\|\bar{\boldsymbol{y}} - \boldsymbol{y}\|^2 = (\bar{\boldsymbol{y}} - \boldsymbol{y})^T(\bar{\boldsymbol{y}} - \boldsymbol{y}) > 0 \quad (3.3)$$

である. また, 補題 5 より,

$$(\boldsymbol{y} - \bar{\boldsymbol{y}})^T(\boldsymbol{x} - \bar{\boldsymbol{y}}) \leq 0, \boldsymbol{x} \in X \quad (3.4)$$

が成立する. (3.3) 式, (3.4) 式より,

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

を得る. ここで,  $\boldsymbol{a} = \bar{\boldsymbol{y}} - \boldsymbol{y}$ ,  $b = (\bar{\boldsymbol{y}} - \boldsymbol{y})^T \bar{\boldsymbol{y}}$  と置くと補題 2 が示される. □

証明 5 (補題 3 の証明)

証明 6 (補題 4 の証明)

## 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.