最適化 復習

Kosuke Toda

目次

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
  - $B(x, \varepsilon) = \{\boldsymbol{y} \in \mathbb{R}^n \mid \|\boldsymbol{y} - \boldsymbol{x}\| < \varepsilon\}$
  - ただし，$\|\cdot\|$ はユークリッドノルムであり，$\|\boldsymbol{x}\| = (\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x})^{1/2}$ である．
- $X \subseteq \mathbb{R}^n$ が開集合 (open set) である
  - $\forall \boldsymbol{x} \in X$, $\exists \varepsilon > 0$; $B(\boldsymbol{x}, \varepsilon) \subseteq X$
- $X \subseteq \mathbb{R}^n$ が閉集合 (closed set) である
  - X の補集合 X^{c} が開集合である．
- $\boldsymbol{x}$ が $X \subseteq \mathbb{R}^n$ の内点 (interior point) である
  - $\boldsymbol{x} \in X \subseteq \mathbb{R}^n$ に対して，$\exists \varepsilon > 0$; $B(\boldsymbol{x}, \varepsilon) \subseteq X$ が成立する．
- X の内部 (interior)，$\mathrm{Int}(X)$
  - $X \subseteq \mathbb{R}^n$ の内点の集合．
- $\boldsymbol{x}$ が $X \subseteq \mathbb{R}^n$ の触点 (contact point) である
  - $\boldsymbol{x} \in X \subseteq \mathbb{R}^n$ に対して，$\forall \varepsilon > 0$; $B(\boldsymbol{x}, \varepsilon) \cap X \neq \emptyset$ が成立する．
- X の閉包 (closure)，$\mathrm{Cl}(X)$
  - $X \subseteq \mathbb{R}^n$ の触点の集合．
  - X を含む最小の閉集合．
- X の境界 (boundary)，$\mathrm{Bd}(X)$
  - $\mathrm{Bd}(X) = \mathrm{Cl}(X) \setminus \mathrm{Int}(X)$.
- $\boldsymbol{x}$ は X の集積点である
  - 任意の $\varepsilon > 0$ に対して，$B(\boldsymbol{x}, \varepsilon) \cap X$ が $\boldsymbol{x}$ 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 $\{\boldsymbol{x}^i\}$, $i = 1, 2, \ldots$ を考える．$\forall\varepsilon > 0$, $\exists I_\varepsilon$; $\|\boldsymbol{x}^i - \boldsymbol{x}\| < \varepsilon$, $i \geq I_\varepsilon$ となる点 $\boldsymbol{x}$ が存在するとき，$\boldsymbol{x}$ は点列 $\{\boldsymbol{x}^i\}$ の極限 (limit) といい，点列 $\{\boldsymbol{x}^i\}$ は $\boldsymbol{x}$ に収束する (converge) という．
- 点列 $\{\boldsymbol{x}^i\}$ の集積点
  - $\{\boldsymbol{x}^i\}$ の部分点列 $\{\boldsymbol{x}^{l_i}\}$ が点 $\boldsymbol{x}$ に収束するとき，点 $\boldsymbol{x}$ を点列 $\{\boldsymbol{x}^i\}$ の集積点という．

有界閉集合 $X \subseteq \mathbb{R}^n$ について，$\{\boldsymbol{x}^i\} \subseteq X$ 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 $\boldsymbol{f} : X \to \mathbb{R}^m$ を考える ($X \subseteq R^n$).

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

  が成立するとき，$\boldsymbol{f}$ は点 $\boldsymbol{x}^0$ で連続 (continuous) であるという．
  - 任意の $\boldsymbol{x} \in X$ で連続となるとき，関数 $\boldsymbol{f}$ は X で連続という．
- 実数値関数
  - 値域が実数集合の関数 $f : X \to \mathbb{R}$.
- 実数値関数のクラス
  - $f : X \to \mathbb{R}(X \subseteq \mathbb{R}^n$ 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，$\partial f(\boldsymbol{x})/\partial x_i$, $i = 1, 2, \ldots, n$ 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，$\partial^2 f(\boldsymbol{x})/\partial x_i \partial x_j$, $i, j = 1, 2, \ldots, n$ 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 $\nabla f(\boldsymbol{x})$

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

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

$$\nabla^2 f(\boldsymbol{x}) = \begin{pmatrix} \dfrac{\partial^2 f(\boldsymbol{x})}{\partial {x_1}^2} & \dfrac{\partial^2 f(\boldsymbol{x})}{\partial x_1 \partial x_2} & \cdots & \dfrac{\partial^2 f(\boldsymbol{x})}{\partial x_1 \partial x_n} \\ \dfrac{\partial^2 f(\boldsymbol{x})}{\partial x_2 \partial x_1} & \dfrac{\partial^2 f(\boldsymbol{x})}{\partial {x_2}^2} & \cdots & \dfrac{\partial^2 f(\boldsymbol{x})}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial^2 f(\boldsymbol{x})}{\partial x_n \partial x_1} & \dfrac{\partial^2 f(\boldsymbol{x})}{\partial x_n \partial x_2} & \cdots & \dfrac{\partial^2 f(\boldsymbol{x})}{\partial {x_n}^2} \end{pmatrix}$$

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

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

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

を満たす $\theta \in (0,1)$ が存在する．

(注) 関数 $f : I \to \mathbb{R}\,(I \subseteq \mathbb{R})$, $f \in C^1$, $x_1, x_2 \in I$ 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　$f : X \to \mathbb{R}\,(X \subseteq \mathbb{R}^n)$, $f \in C^2$, $\boldsymbol{x}^1, \boldsymbol{x}^2 \in X$ に対して，

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

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

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

が成立する．

(注) 関数 $f : I \to \mathbb{R}\,(I \subseteq \mathbb{R})$, $f \in C^1$, $x_1, x_2 \in I$ 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，

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

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　$h_i : \mathbb{R}^n \to \mathbb{R}$, $\boldsymbol{x}^0 = (x_1^0, x_1^0, \ldots, x_n^0)^{\mathrm{T}} \in \mathbb{R}^n$ の近傍で，$h_i \in C^p\,(p \geq 1)$

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

が成立するとする．このとき，ヤコビ行列 (Jacobian matrix)

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

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

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

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

(注) 2 変数関数を考える．関数 $f : \Omega \to \mathbb{R}\,(\Omega \subseteq \mathbb{R}^2)$, $f \in C^p$, $p \geq 1$, $(x_1, x_2) \in \Omega$ 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

$$u(x) \coloneqq f(x, \phi(x)) = 0 \;\; (\forall x \in I)$$

が成立する．

2.3 凸集合

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

補題 2.1 $X_1, X_2, \ldots \subseteq \mathbb{R}^n$ を凸集合とすると，$\bigcap_{i=1,2,\ldots} X_i$ 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，$\boldsymbol{y} \notin X$ とする．このとき，

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

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

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

$$\boldsymbol{a}^{\mathrm{T}} \boldsymbol{x} \geq \boldsymbol{a}^{\mathrm{T}} \boldsymbol{y}, \; \forall \boldsymbol{x} \in X$$

なる $\boldsymbol{a} \neq \boldsymbol{0}$ 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，

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

なる $\boldsymbol{a} \neq \boldsymbol{0}$, 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 $\boldsymbol{x}^1, \boldsymbol{x}^2$ を結ぶ線分上の任意の内分点は，$x_\lambda = \lambda \boldsymbol{x}^1 + (1 - \lambda)\boldsymbol{x}^2$, $\lambda \in [0, 1]$ 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 $f : X \to \mathbb{R}$ を考える．任意の $\boldsymbol{x}^1, \boldsymbol{x}^2 \in X$，任意の $\lambda \in [0,1]$ に対して，

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

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 $\boldsymbol{x}^1, \boldsymbol{x}^2 \in X$ ($\boldsymbol{x}^1 \neq \boldsymbol{x}^2$)，任意の $\lambda \in (0,1)$ に対して，

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

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 $\boldsymbol{x}^1, \boldsymbol{x}^2 \in X$，任意の $\lambda \in [0,1]$ に対して，

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

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 $\boldsymbol{x}^1, \boldsymbol{x}^2 \in X$ ($\boldsymbol{x}^1 \neq \boldsymbol{x}^2$)，任意の $\lambda \in (0,1)$ に対して，

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

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 $f : X \to \mathbb{R}$ 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，

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

が凸集合となることである．$\mathrm{epi} f$ 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 $f : X \to \mathbb{R}$ 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，

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

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 $f : X \to \mathbb{R}$ 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，

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

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，

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

が成立することである．

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

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

と定められる．f が微分可能なとき，$f'(\boldsymbol{x}; \boldsymbol{d}) = \nabla f(\boldsymbol{x})^{\mathrm{T}} \boldsymbol{d}$ 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 $f : X \to \mathbb{R}$ 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 $H(\boldsymbol{x})$ が任意の $\boldsymbol{x} \in X$ に対して半正定 (positive semi-definite)，すなわち，

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

   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 $H(\boldsymbol{x})$ が任意の $\boldsymbol{x} \in X$ に対して正定 (positive definite)，すなわち，

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

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

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

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，$f(\boldsymbol{x} + r\boldsymbol{d}) = f(\boldsymbol{x}) + r\nabla f(\boldsymbol{x})^{\mathrm{T}} \boldsymbol{d} + o(r\|\boldsymbol{d}\|)$ である．さらに，$\lim_{r \to +0} o(r\|\boldsymbol{d}\|)/r = 0$ 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 $f : X \to \mathbb{R}$ が与えられたとき，f を最小にする解 $\boldsymbol{x} \in X \cap X$ 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．

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

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 $g_i : \mathbb{R}^n \to \mathbb{R}$, $i = 1, 2, \dots, m$ による不等式および等式制約を用いて，

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

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)

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

実行可能解 $\boldsymbol{x} \in S \cap X$ 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 $\boldsymbol{x}^*$ は，$\forall \boldsymbol{x} \in S \cap X,\ f(\boldsymbol{x}^*) \leq f(\boldsymbol{x})$ を満たす $\boldsymbol{x}^* \in S \cap X$ 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，

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

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 $\boldsymbol{x}^*$ が存在するならば $f^* = f(\boldsymbol{x}^*)$ 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

非線形計画問題とアルゴリズム

この章では，非線形計画問題の最適性条件および最適化問題を解くアルゴリズムを述べる．

4.1　非線形計画問題の最適性条件

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．

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

4.1.1　制約なし問題の最適性条件

最適化問題として，(4.1) を考える．

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

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，$\bar{\boldsymbol{x}} \in \mathbb{R}^n$ が問題 (4.1) の局所最適解であるための必要条件は，

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

となることである．

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

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

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

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

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，$\bar{\boldsymbol{x}}$ が問題 (4.1) の局所最適解であるための十分条件は，

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

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

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

が成立することである．

4.1.2　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．

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

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

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

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，$\boldsymbol{u} \in \mathbb{R}^m$, $\boldsymbol{g}(\boldsymbol{x}) = (g_1(\boldsymbol{x}), g_2(\boldsymbol{x}), \ldots, g_m(\boldsymbol{x}))^{\mathrm{T}}$ である．

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

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

ただし，$\bar{\boldsymbol{x}}$ 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，$\boldsymbol{g}(\boldsymbol{x}) = \mathbf{0}$ を満たす解 $\boldsymbol{x}$ を考えるとき，$\bar{\boldsymbol{x}}$ が近傍内で $F^k(\boldsymbol{x})$ 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．

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

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

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

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

を定義する．$\bar{B}(\bar{\boldsymbol{x}}, \varepsilon)$ 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 $\boldsymbol{x}^k$ による点列 $\{\boldsymbol{x}^k\}$ が $\bar{\boldsymbol{x}}$ に収束することを示す．

任意の k に対し，$\boldsymbol{x}^k$ の最適性より，

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

より，$\lim_{k\to\infty} \|\boldsymbol{g}(\boldsymbol{x}^k)\| = 0$ でなければならない[*1]．従って，$\{\boldsymbol{x}^k\}$ の任意の極限点 $\bar{\bar{\boldsymbol{x}}}$ は，$\boldsymbol{g}(\bar{\bar{\boldsymbol{x}}}) = 0$ 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，

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

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

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

つまり，点列 $\{\boldsymbol{x}^k\}$ は $\bar{\boldsymbol{x}}$ に収束する．点 $\bar{\boldsymbol{x}}$ 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，$\boldsymbol{x}^k$ も内点である．したがって，$\boldsymbol{x}^k$ は制約なしの最適化問題

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

の局所最適解でもある．

■等式制約問題の最適性十分条件

*1 これを満たさなければ，$k \to \infty$ で発散する．

第 5 章

線形計画問題とアルゴリズム

第 6 章

証明

この章には，補題や定理の証明を記載する．

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．$\forall \boldsymbol{x}^1, \boldsymbol{x}^2 \in X_1 \cap X_2$ に対し，$\boldsymbol{x}^1$ と $\boldsymbol{x}^2$ を結ぶ線分は，$\boldsymbol{x}^1 \in X_1, \boldsymbol{x}^2 \in X_1$ 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，$\boldsymbol{x}^1$ と $\boldsymbol{x}^2$ 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．　□

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，

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

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．$\boldsymbol{u} \in \mathbb{R}^n$ の L への射影を $\boldsymbol{v}$ とする．このとき，$\forall \boldsymbol{w} \in L$ に対して，

$$(\boldsymbol{u} - \boldsymbol{v})^{\mathrm{T}} (\boldsymbol{w} - \boldsymbol{v}) \leq 0 \tag{6.1}$$

が成立する．

証明 2（定理 6.1 の証明）

証明 3（補題 6.1 の証明）　$\boldsymbol{u}$ の L への射影 $\boldsymbol{v}$ と任意の点 $\boldsymbol{w} \in L$ を結ぶ線分上の点 $\boldsymbol{x}_\lambda = (1-\lambda)\boldsymbol{v} + \lambda \boldsymbol{w}$, $0 < \lambda < 1$ を考える．L は凸より，$\boldsymbol{x}_\lambda$ も L に含まれる．$\boldsymbol{v}$ の定義より，

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

[*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 $\boldsymbol{u} \in X$ に対し，$\|\boldsymbol{u}\| = \sqrt{(\boldsymbol{u}, \boldsymbol{u})}$ 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，

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

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，$\lambda \to 0$ とすると，(6.1) 式を得る． □

証明 4（補題 2.2 の証明）　$\boldsymbol{y} \in \mathbb{R}^n$ の X への射影を $\bar{\boldsymbol{y}}$ とする．定理 6.1 より，$\bar{\boldsymbol{y}}$ は $\boldsymbol{y}$ に対して一意に存在する．$\boldsymbol{y} \notin X$ より，$\boldsymbol{y} \neq \bar{\boldsymbol{y}}$ である．よって，

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

である．また，補題 6.1 より，

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

が成立する．(6.3) 式，(6.4) 式より，

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

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

証明 5（補題 2.3 の証明）　$\boldsymbol{y}$ は X の内点でないので，$\boldsymbol{y}$ の近傍 $B_k = \{\boldsymbol{x} \in \mathbb{R}^n \mid \|\boldsymbol{x}-\boldsymbol{y}\| < \varepsilon_k\}$ はある点 $\boldsymbol{x}^k \notin \mathrm{Cl}(X)$ を含む．$k \to \infty$ に伴い，$\varepsilon_k \to 0$ とすると，点列 $\{\boldsymbol{x}^k\}$ は $\mathrm{Cl}(X)$ の外部から $\boldsymbol{y}$ に収束する．$\boldsymbol{x}^k \notin \mathrm{Bd}(X)$ から，補題 2.2 より

$$(\boldsymbol{a}^k)^{\mathrm{T}}\boldsymbol{x} > (\boldsymbol{a}^k)^{\mathrm{T}}\boldsymbol{x}^k,\ \forall \boldsymbol{x} \in X$$

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

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

である．$\|\boldsymbol{a}\| = 1$ より，$\boldsymbol{a} \neq \boldsymbol{0}$ である． □

証明 6（補題 2.4 の証明）　集合 X, Y から集合 $S = \{\boldsymbol{z} \in \mathbb{R}^n \mid \boldsymbol{z} = \boldsymbol{x}-\boldsymbol{y},\ \boldsymbol{x} \in X, \boldsymbol{y} \in Y\}$ を定義すると，S も凸集合である[*2]．$X \cap Y = \emptyset$ より，$\boldsymbol{0}$ ベクトルは $\boldsymbol{0} \notin S$ である．補題 2.2，補題 2.3 より，

$$\boldsymbol{a}^{\mathrm{T}}\boldsymbol{z} \geq \boldsymbol{0},\ \forall \boldsymbol{z} \in S$$

を満たす $\boldsymbol{a} \neq \boldsymbol{0}$ が存在する．よって，S の定義より，

$$\boldsymbol{a}^{\mathrm{T}}\boldsymbol{x} \geq \boldsymbol{a}^{\mathrm{T}}\boldsymbol{y},\ \forall \boldsymbol{x} \in X, \forall \boldsymbol{y} \in Y$$

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

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

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

と書けるため，X と Y の凸性より，やはり S に属するから．

証明 7（定理 2.1 の証明）

証明 8（定理 2.2 の証明）

証明 9（定理 2.3 の証明）

証明 10（定理 2.4 の証明）

証明 11（定理 4.1 の証明）　$\boldsymbol{d} \in \mathbb{R}^n$, $\varepsilon > 0$ を考える．ε が十分小さければ，局所最適性より，

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

が成立する．Taylor の定理より，

$$\begin{aligned} f(\bar{\boldsymbol{x}} + \varepsilon \boldsymbol{d}) &= f(\bar{\boldsymbol{x}}) + \varepsilon \nabla f(\bar{\boldsymbol{x}})^{\mathrm{T}} \boldsymbol{d} + o(\varepsilon \|\boldsymbol{d}\|) \geq f(\bar{\boldsymbol{x}}) \\ f(\bar{\boldsymbol{x}} - \varepsilon \boldsymbol{d}) &= f(\bar{\boldsymbol{x}}) - \varepsilon \nabla f(\bar{\boldsymbol{x}})^{\mathrm{T}} \boldsymbol{d} + o(\varepsilon \|\boldsymbol{d}\|) \geq f(\bar{\boldsymbol{x}}) \end{aligned}$$

なので，

$$|\nabla f(\bar{\boldsymbol{x}}) \boldsymbol{d}| = \frac{o(\varepsilon \|\boldsymbol{d}\|)}{\varepsilon} \to 0\ (\varepsilon \to 0)$$

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

証明 12（定理 4.2 の証明）　定理 4.1 より，$\nabla f(\bar{\boldsymbol{x}}) = 0$．任意の $\boldsymbol{d} \in \mathbb{R}^n$, $\varepsilon > 0$ に対して，

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

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

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

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

証明 13（定理 4.3 の証明）　$f(\bar{\boldsymbol{x}} + \varepsilon \boldsymbol{d})$ に対して，Taylor の定理より，

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

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

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

$H(\bar{\boldsymbol{x}})$ が正定値であり，$\varepsilon \to 0$ 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，$f(\bar{\boldsymbol{x}} + \varepsilon \boldsymbol{d}) \geq f(\bar{\boldsymbol{x}})$ が成立する．つまり，$\bar{\boldsymbol{x}}$ は局所最適解である． □

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.