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 $\lim_{t \to 0} o(t)/t = 0$ 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 $\hat{\boldsymbol{x}}^0 = (x^0_{m+1}, \dots, x^0_n)^{\mathrm{T}} \in \mathbb{R}^{n-m}$ の近傍 $U = B(\hat{\boldsymbol{x}}^0, \varepsilon)$ が存在し，$\hat{\boldsymbol{x}} \in U$ に対して，

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

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

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

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

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

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

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

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．

- $\boldsymbol{x}^*$ 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　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．

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

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，

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

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，

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

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，

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

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

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

が成立することである．

ここで，次の例題を解く．

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

勾配ベクトルは，$\nabla f(\boldsymbol{x}) = (4x_1 + x_2 - 5, x_1 + 2x_2 - 3)^{\mathrm{T}}$ 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，

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

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．

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

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．

[*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．

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

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，$\boldsymbol{g}(\boldsymbol{x}) = \boldsymbol{0}$ 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[*2]．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．

[*2] これを満たさなければ，$k \to \infty$ 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 $g_i(\boldsymbol{x}),\ i=1,2,\dots,m$ が互いに線形独立であるとき，点 $\boldsymbol{x}$ 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，$\bar{\boldsymbol{x}}$ が正則であるとき，$\bar{\boldsymbol{x}} \in \mathbb{R}^n$ が問題 (4.2) の局所最適解であるための必要条件は，

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

を満たすラグランジュ乗数ベクトル $\boldsymbol{u} = (u_1, u_2, \dots, u_m)^{\mathrm{T}}$ 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，$\bar{\boldsymbol{x}} \in \mathbb{R}^n$ が問題 (4.2) の局所最適解であるための必要条件は，

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

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

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

である．

ここで，Lagrange 関数 (6.2) を用いると，等式制約条件 $\boldsymbol{g}(\boldsymbol{x}) = 0$ および条件 (4.9)，(4.10) は次のように表される．

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

■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 $g(\boldsymbol{x}) = 0$ の場合を考える．条件 (4.9) は，$\nabla f(\boldsymbol{x}) + u \nabla g(\boldsymbol{x}) = 0$ 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 $\nabla f(\boldsymbol{x})$ と $\nabla g(\boldsymbol{x})$ が同方向か，あるいは逆方向であることを述べている．$g(\boldsymbol{x}) = 0$ の範囲内で微小量移動することは，$\nabla g(\boldsymbol{x})$ の直交方向に進むことに相当する．この方向は $\nabla f(\boldsymbol{x})$ 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 $\nabla g_i(\boldsymbol{x}),\ i=1,2,\dots,m$ が張る空間内にベクトル $\nabla f(\boldsymbol{x})$ 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 $\boldsymbol{u}$ 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　$f, g_i \in C^2$, $i = 1, 2, \ldots, m$ であるとき，$\bar{\boldsymbol{x}} \in \mathbb{R}^n$ が問題 (4.2) の局所最適解であるための十分条件は，

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

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

ここで，次の例題を解く．

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

Lagrange 関数を $L(\boldsymbol{x}, u)$ とすると，

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

最適性の必要条件より，

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

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，

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

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 $\boldsymbol{y}$ に対し，$\boldsymbol{y}^{\mathrm{T}}(\nabla_{\boldsymbol{xx}})\boldsymbol{y} \geq 0$ 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 $\nabla_{\boldsymbol{xx}}$ を求めると，

$$\nabla_{\boldsymbol{xx}} = \begin{pmatrix} -1 & -1 \\ -1 & -1 \end{pmatrix}$$

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，$(\nabla g(\bar{\boldsymbol{x}}))^{\mathrm{T}} \boldsymbol{y} = \boldsymbol{0}$ を満たすベクトル y の集合は，パラメータ t を用いて，

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

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，$(\nabla g(\bar{\boldsymbol{x}}))^{\mathrm{T}}\boldsymbol{y} = \boldsymbol{0}$ を満たすベクトル y の集合は，パラメータ t を用いて，

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

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，

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

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．

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

実行可能解 $\boldsymbol{x}$ は，各不等式条件 $g_i(\boldsymbol{x}) = 0$ 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，$g_i(\boldsymbol{x}) = 0$ となるものを，$\boldsymbol{x}$ 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，

$$A(\boldsymbol{x}) = \{i \mid g_i(\boldsymbol{x}) = 0,\ i = 1, 2, \ldots, l\} \tag{4.13}$$

と記す．実行可能解 $\boldsymbol{x}$ において，等式で成立する条件の勾配ベクトル $\nabla g_i(\boldsymbol{x}), i \in A(\boldsymbol{x})$ および $\nabla g_i(\boldsymbol{x}), i = l+1, \ldots, m$ が線形独立であるとき，$\boldsymbol{x}$ 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 (6.2) 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

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

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

を満たすラグランジュ乗数ベクトル $\boldsymbol{u} \in \mathbb{R}^m$ 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，$f, g_i \in C^2,\ i = 1, 2, \ldots, m$ であるとき，$\bar{\boldsymbol{x}} \in \mathbb{R}^n$ 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，

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

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

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

 である．

ここで，(4.14) の 3 番目の条件は，

$$u_i g_i(\bar{\boldsymbol{x}}) = 0,\ i = 1, 2, \ldots, l \tag{4.17}$$

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　$f, g_i \in C^2,\ i = 1, 2, \ldots, m$ であるとき，$\bar{\boldsymbol{x}} \in \mathbb{R}^n$ 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 $u_i,\ i = l+1, \ldots, m$ に対する符号制約はない．

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

*6 KKT 条件は，$\bar{\boldsymbol{x}}$ から実行可能領域のどの方向を選んでも，$-\nabla f(\bar{\boldsymbol{x}})$ 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)．

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

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

ここで，次の例題を解く．

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

目的関数および制約式の勾配ベクトルはそれぞれ，

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

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，

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

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 $\boldsymbol{v}_1 = (1, \sqrt{2} - 1)^{\mathrm{T}}, \boldsymbol{v}_2 = (1, -\sqrt{2} - 1)^{\mathrm{T}}$ 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 $g_1(\boldsymbol{x}), g_2(\boldsymbol{x})$ と考えられる．よって，

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

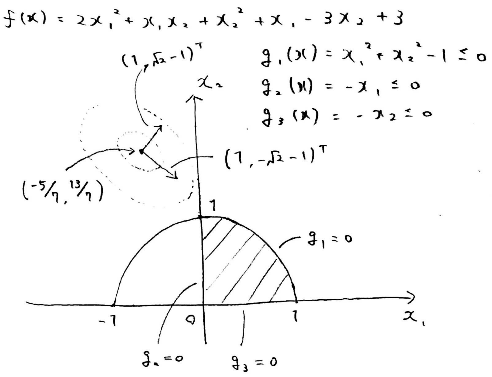

図 **4.1:** 実行可能領域と f の等高線

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，

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

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．$\bar{\boldsymbol{x}} = (0, 1)^{\mathrm{T}}$ における目的関数および有効制約の勾配ベクトルは，

$$\begin{aligned} \nabla f(\boldsymbol{x}) &= (2, -1)^{\mathrm{T}} \\ \nabla g_1(\boldsymbol{x}) &= (0, 2)^{\mathrm{T}} \\ \nabla g_2(\boldsymbol{x}) &= (-1, 0)^{\mathrm{T}} \end{aligned}$$

である．図 4.2 にこのベクトルを $\bar{\boldsymbol{x}}$ を始点に描いたものを示す．$\nabla g_i(\bar{\boldsymbol{x}})$ は，$\bar{\boldsymbol{x}}$ における $g_i(\boldsymbol{x})$ の増加方向を示す．これは，有効制約 $g_i(\boldsymbol{x}) = 0$ の $\bar{\boldsymbol{x}}$ における接線と直交する．また，$\nabla f(\bar{\boldsymbol{x}})$ は，$\bar{\boldsymbol{x}}$ における $f(\boldsymbol{x})$ の増加方向を示す．$f(\boldsymbol{x})$ を最小化するためには，逆方向 $-\nabla f(\boldsymbol{x})$ に進むのが望ましい．$-\nabla f(\boldsymbol{x})$ 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．

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

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

を定義する．

KKT 条件は，ベクトル $-\nabla f(\bar{\boldsymbol{x}})$ 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，$\bar{\boldsymbol{x}}$ から実行可能領域のどの方向を選んでも $-\nabla f(\bar{\boldsymbol{x}})$ 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，$\bar{\boldsymbol{x}}$ の局所最適性の必要条件となっている．

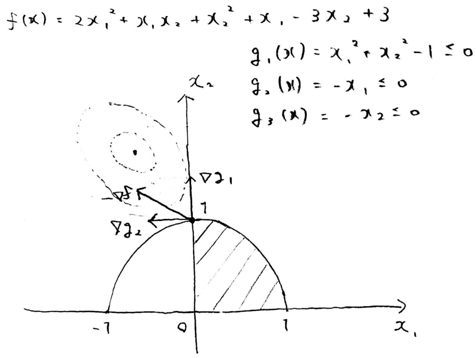

図 **4.2:** KKT 条件の幾何学的説明

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

$$\begin{array}{ll} \text{minimize} & f(\boldsymbol{x}) \\ \text{subject to} & \boldsymbol{x} \in S \cap X \end{array}$$

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

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

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

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

を満たす $\boldsymbol{x}$ の領域は凸集合である[*7]．

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

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

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

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．

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

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．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 $V = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_p\}$ を用いて，次のように表すことができる．

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

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

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

と表される．両辺に左から $\boldsymbol{c}^{\mathrm{T}}$ を掛けると，

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

となる．$\sum_{i=1}^{p} k_i = 1,\ k_i \geq 0,\ i = 1, 2, \ldots, p$ より，

$$\min_{i=1,2,\ldots,p} \boldsymbol{c}^{\mathrm{T}}\boldsymbol{v}_i \leq \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} \leq \max_{i=1,2,\ldots,p} \boldsymbol{c}^{\mathrm{T}}\boldsymbol{v}_i$$

となる．$i = 1, 2, \ldots, p$ について，$\boldsymbol{v}_i \in X$ より，実行可能領域 X での $\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}$ 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，

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

である．ただし，$\boldsymbol{c} = (c_1, c_2, \ldots, c_n)^{\mathrm{T}}$, $\boldsymbol{b} = (b_1, b_2, \ldots, b_m)^{\mathrm{T}}$ 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．$\boldsymbol{x} = (x_1, x_2, \ldots, x_n)^{\mathrm{T}}$ 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，

$$\text{maximize} \ \ \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}$$

の場合，$\boldsymbol{d} = -\boldsymbol{c}$ とすると，

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

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，

$$\boldsymbol{a}_i^{\mathrm{T}}\boldsymbol{x} \leq b_i$$

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

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

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，

$$\boldsymbol{a}_j^{\mathrm{T}}\boldsymbol{x} \geq b_j$$

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

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

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．

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

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

$$B\begin{pmatrix} x_1^B \\ x_2^B \\ \vdots \\ x_m^B \end{pmatrix} = \boldsymbol{b} \tag{5.4}$$

を解くことにより得られる一意的なベクトル $B^{-1}\boldsymbol{b}$ 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，$B^{-1}\boldsymbol{b} \geq \boldsymbol{0}$ 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

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

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，$A = (B, N), \boldsymbol{x} = (\boldsymbol{x}_B^{\mathrm{T}}, \boldsymbol{x}_N^{\mathrm{T}})^{\mathrm{T}}, \boldsymbol{c} = (\boldsymbol{c}_B^{\mathrm{T}}, \boldsymbol{c}_N^{\mathrm{T}})^{\mathrm{T}}$ と表すと，式 (5.5) は，

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

[*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．$\boldsymbol{x}_B$ は基底変数ベクトル，$\boldsymbol{x}_N$ は非基底変数ベクトル，$\boldsymbol{x}_B$ の各成分 x_j^B は基底変数，$\boldsymbol{x}_N$ 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 $\boldsymbol{x}_B$ を消去すると，(5.6) の問題は，

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

と等価になる．(5.7) の一部

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

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

$$\boldsymbol{x} = \begin{pmatrix} \boldsymbol{x}_B \\ \boldsymbol{x}_N \end{pmatrix} = \begin{pmatrix} B^{-1}\boldsymbol{b} \\ \boldsymbol{0} \end{pmatrix} \tag{5.9}$$

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

(5.8) 式を $\bar{\boldsymbol{c}}, \bar{N}, \bar{\boldsymbol{b}}$ を用いて表すと，

$$\begin{aligned} & z = \boldsymbol{c}_B^{\mathrm{T}} \bar{\boldsymbol{b}} + \bar{\boldsymbol{c}}^{\mathrm{T}}\boldsymbol{x}_N \\ & \boldsymbol{x}_B + \bar{N}\boldsymbol{x}_N = \bar{\boldsymbol{b}} \end{aligned} \tag{5.10}$$

となる．(5.7) 式において，

$$\bar{\boldsymbol{b}} = B^{-1}\boldsymbol{b} \geq \boldsymbol{0} \ \ \text{(実行可能性)} \tag{5.11}$$

が成立すれば，(5.9) 式の基底解は実行可能解となる．(5.11) 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，

$$\bar{\boldsymbol{c}}^{\mathrm{T}} = \boldsymbol{c}_N^{\mathrm{T}} - \boldsymbol{c}_B^{\mathrm{T}} B^{-1}N \geq \boldsymbol{0}^{\mathrm{T}} \ \ \text{(最適性)} \tag{5.12}$$

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，

$$\boldsymbol{x}_B = \bar{\boldsymbol{b}} - \begin{pmatrix} \bar{a}_{1r} \\ \bar{a}_{2r} \\ \vdots \\ \bar{a}_{mr} \end{pmatrix} x_r^N \tag{5.13}$$

[*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．

を満たせば，制約条件の最初の等式条件 $\boldsymbol{x}_B + \bar{N}\boldsymbol{x}_N = \bar{\boldsymbol{b}}$ を満足する．実際，$\boldsymbol{x}_N = \left(0, \cdots, x_r^N, \cdots, 0\right)^{\mathrm{T}}$ のとき，

$$\bar{N}\boldsymbol{x}_N = \begin{pmatrix} \bar{a}_{11} & \cdots & \bar{a}_{1r} & \cdots & \bar{a}_{1n-m} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ \bar{a}_{m1} & \cdots & \bar{a}_{mr} & \cdots & \bar{a}_{mn-m} \end{pmatrix} \begin{pmatrix} 0 \\ \vdots \\ x_r^N \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} \bar{a}_{1r} \\ \bar{a}_{2r} \\ \vdots \\ \bar{a}_{mr} \end{pmatrix} x_r^N$$

を満たす．さらに，非負条件を満たすには，$\boldsymbol{x}_B \geq \boldsymbol{0}$ より，

$$\bar{\boldsymbol{b}} - \begin{pmatrix} \bar{a}_{1r} \\ \bar{a}_{2r} \\ \vdots \\ \bar{a}_{mr} \end{pmatrix} x_r^N \geq \boldsymbol{0} \tag{5.14}$$

が成立すればよい．もし，

$$\begin{pmatrix} \bar{a}_{1r} \\ \bar{a}_{2r} \\ \vdots \\ \bar{a}_{mr} \end{pmatrix} \leq \boldsymbol{0} \tag{5.15}$$

が成立すれば，(5.14) 式より，実行可能性を損なうことなく x_r^N をいくらでも正に大きくすることができる．このとき，目的関数値をいくらでも小さくすることができるので，この解は非有解 (unbounded solution) である．すなわち，最適解は存在しない．

一方，(5.15) 式が成立しない場合，$\bar{a}_{ir}$ のいくつかは正となる．(5.14) 式を満たすには，$\bar{a}_{ir} > 0$ なるすべての $i \in \{1, 2, \ldots, m\}$ に対して，

$$\bar{b}_i - \bar{a}_{ir} x_r^N \geq 0 \tag{5.16}$$

を満たせばよい．すなわち，

$$x_r^N \leq \min_{i:\, \bar{a}_{ir} > 0} \frac{\bar{b}_i}{\bar{a}_{ir}} \tag{5.17}$$

を満たせばよい．いま，$\bar{c_r} < 0$ を考えているので，x_r^N が大きいほど目的関数値が小さくなる．よって，

$$x_r^N = \min_{i:\, \bar{a}_{ir} > 0} \frac{\bar{b}_i}{\bar{a}_{ir}} \tag{5.18}$$

とおくと，現在の実行可能基底解よりも目的関数値が小さい実行可能解が得られる．右辺の最小値を与える i を s とおく，つまり，(5.18) 式を満たす i を $i = s$ とすると，$\boldsymbol{x}_B$ の s 番目の成分 x_s^B の値は，(5.13) 式より，

$$x_s^B = \bar{b}_s - \bar{a}_{sr} x_r^N = \bar{b}_s - \bar{a}_{sr} \min_{i:\, \bar{a}_{ir} > 0} \frac{\bar{b}_i}{\bar{a}_{ir}} = \bar{b}_s - \bar{a}_{sr} \frac{\bar{b}_s}{\bar{a}_{sr}} = 0 \tag{5.19}$$

となる．すなわち，基底変数 x_s^B を非基底変数 x_r^N で入れ替えた実行可能基底解に対応する．したがって，この入れ替えを行った基底形式を求め，同様の議論を続けていけば，最適解が見つかるか，もしくは非有界であるかがわかる．この基底の入れ替えを行った後の基底形式を (5.7) 式から求める計算法を与えるのがピボット演算である．

5.4.4 ピボット演算

(5.7) 式に対応する基底形式 (5.8) は，$\bar{z} = \boldsymbol{c}_B^{\mathrm{T}} B^{-1} \boldsymbol{b}$ とすると，次のように書き換えられる．

$$z = \bar{z} + \begin{pmatrix} \bar{c}_1 & \cdots & \bar{c}_{n-m} \end{pmatrix} \begin{pmatrix} x_1^N \\ \vdots \\ x_{n-m}^N \end{pmatrix}$$

$$\begin{pmatrix} x_1^B \\ \vdots \\ x_m^B \end{pmatrix} + \begin{pmatrix} \bar{a}_{11} & \cdots & \bar{a}_{1n-m} \\ \vdots & & \vdots \\ \bar{a}_{m1} & \cdots & \bar{a}_{mn-m} \end{pmatrix} \begin{pmatrix} x_1^N \\ \vdots \\ x_{n-m}^N \end{pmatrix} = \begin{pmatrix} \bar{b}_1 \\ \vdots \\ \bar{b}_m \end{pmatrix}$$

つまり，

$$\begin{array}{llllll} x_1^B & & + \bar{a}_{11} x_1^N & + \bar{a}_{12} x_2^N & + \cdots + \bar{a}_{1n-m} x_{n-m}^N & = \bar{b}_1 \\ & x_2^B & + \bar{a}_{21} x_1^N & + \bar{a}_{22} x_2^N & + \cdots + \bar{a}_{2n-m} x_{n-m}^N & = \bar{b}_2 \\ & \ddots & \cdots & \cdots & \cdots & \\ & x_m^B & + \bar{a}_{m1} x_1^N & + \bar{a}_{m2} x_2^N & + \cdots + \bar{a}_{mn-m} x_{n-m}^N & = \bar{b}_m \\ & -z & + \bar{c}_1 x_1^N & + \bar{c}_2 x_2^N & + \cdots + \bar{c}_{n-m} x_{n-m}^N & = -\bar{z} \end{array} \tag{5.20}$$

のように書き換えられる．ピボット演算は次の手順で行われる．ただし，s は (5.18) 式を満たす i であり，r は基底形式 (5.8) の係数ベクトル $\bar{\boldsymbol{c}} = \boldsymbol{c}_N^{\mathrm{T}} - \boldsymbol{c}_B^{\mathrm{T}} B^{-1} N$ のうち，負の値を取る要素 $\bar{c_r}$ の添字である．

ピボット操作

手順 1　s 行を $\bar{a}_{sr}$ で割る．

手順 2　目的関数の行も含め，i 行 j 列の元の係数 d_{ij}^{old}，新しい係数を d_{ij}^{new} とすると，$i \neq s$ に対して，

$$d_{ij}^{new} = d_{ij}^{old} - d_{sj}^{old} \frac{\bar{a}_{ir}}{\bar{a}_{sr}}$$

と更新する．

つまり，ピボット演算は s 行 r 列の値を 1 に，r 列の他の行の値が 0 になるように等価変換することに相当する．

5.5 シンプレックス法

線形計画問題を解くアルゴリズムであるシンプレックス法 (simplex method) の基本形を与える．このアルゴリズムは，制約条件が定める凸多面体 X の頂点 (実行可能基底解) の 1 つから出発し，それに隣接する頂点の中でより目的関数値を小さくするものへ移動するという操作を反復して，最終的に最適解である頂点の 1 つに到達するか，あるいは発散を結論するアルゴリズムである．シンプレックス法は単体法とも呼ばれる．以下にシンプレックス法の手順を示す．

アルゴリズム SIMPLEX

Input　標準形の問題 (5.5)．ただし，$\boldsymbol{b} \geq \boldsymbol{0}$ を仮定し，(5.7) に等価変換する．

Output　最適解 $\bar{\boldsymbol{x}}$，あるいは「発散」の結論．

1. 初期実行可能基底解 $\boldsymbol{x}^0$ を求める．$k \leftarrow 0$.
2. $\bar{\boldsymbol{c}}^{\mathrm{T}} \coloneqq \boldsymbol{c}_N^{\mathrm{T}} - \boldsymbol{c}_B^{\mathrm{T}} B^{-1} N$ によって得られる $\bar{\boldsymbol{c}}$ の要素 $\bar{c}_j$ を用いて，
$$\min_{1 \leq j \leq n-m} \bar{c}_j = \bar{c}_r$$
となる r を 1 つ選ぶ．$\bar{c}_r \geq 0$ であれば，$\bar{\boldsymbol{x}} \leftarrow \boldsymbol{x}^k$ として計算終了．
3. $\bar{N} \coloneqq B^{-1}N(= (\bar{a}_{ij}))$ に対して，第 r 列のすべての成分 $\bar{a}_{ir}$ が $\bar{a}_{ir} \leq 0$ であれば，最小値が有界でない (問題が発散する) ことを結論して計算終了．
4. $\bar{a}_{ir}$ に正のものがあれば，
$$\min_{i:\, \bar{a}_{ir} > 0} \frac{\bar{b}_i}{\bar{a}_{ir}} = \frac{\bar{b}_s}{\bar{a}_{sr}}$$
となる s を 1 つ選ぶ．
5. $\bar{a}_{sr}$ に対するピボット操作により x_s^B の代わりに x_r^N を基底変数とする新しい実行可能基底解 $\boldsymbol{x}^{k+1}$ を求める．$k \leftarrow k+1$ として 2. に戻る．

通常，シンプレックスタブローを用いて問題 (5.5) を解く．最適化問題 (5.3) を例に説明する．

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

表 5.2 に問題 (5.3) のシンプレックスタブローを示す．

最適化問題 (5.3) において，決定変数ベクトルを $\boldsymbol{x} = (x_1, x_2, x_3, x_4, x_5)^{\mathrm{T}} \geq \boldsymbol{0}$ とする．さらに，

$$A = \begin{pmatrix} 2.5 & 5 & 1 & 0 & 0 \\ 5 & 6 & 0 & 1 & 0 \\ 3 & 2 & 0 & 0 & 1 \end{pmatrix},\ \boldsymbol{b} = \begin{pmatrix} 350 \\ 450 \\ 240 \end{pmatrix},\ \boldsymbol{c} = \begin{pmatrix} -4 \\ -5 \end{pmatrix}$$

とする．

サイクル 0 では，基底変数ベクトルを $\boldsymbol{x}_B = (x_3, x_4, x_5)^{\mathrm{T}}$，非基底変数ベクトルを $\boldsymbol{x}_N = (x_1, x_2)^{\mathrm{T}}$ としている．ただし，(5.6) における B 行列，N 行列はそれぞれ，

$$B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},\ N = \begin{pmatrix} 2.5 & 5 \\ 5 & 6 \\ 3 & 2 \end{pmatrix}$$

である．また，初期実行可能基底解は，$\boldsymbol{x}^0 = (x_1, x_2, x_3, x_4, x_5)^{\mathrm{T}} = (0, 0, 350, 450, 240)^{\mathrm{T}}$ である．$k \leftarrow 0$ とする．$\bar{\boldsymbol{c}}^{\mathrm{T}} = (-4, -5)$ であり，

$$\min_{1 \leq j \leq 2} \bar{c}_j = -5 = \bar{c}_2$$

表 **5.2:** シンプレックスタブロー

サイクル	基底	x_1	x_2	x_3	x_4	x_5	右辺
0	x_3	2.5	[5]	1			350
	x_4	5	6		1		450
	x_5	3	2			1	240
	$-z$	-4	-5				0
1	x_2	0.5	1	0.2			70
	x_4	[2]		-1.2	1		30
	x_5	2		-0.4		1	100
	$-z$	-1.5		1			350
2	x_2		1	-0.1	-0.25		62.5
	x_1	1		-0.6	0.5		15
	x_5			0.8	-0.5	1	70
	$-z$			0.1	0.75		372.5

より，$r = 2$．$\bar{N} = B^{-1}N(= N)$ の第 2 列の成分 $\bar{a}_{i2}$ のうち，

$$\min_{i:\, \bar{a}_{i2}>0} \frac{\bar{b}_i}{\bar{a}_{i2}} = 70 = \frac{\bar{b}_1}{\bar{a}_{12}}$$

より，$s = 1$．$\bar{a}_{12}$ に対するピボット操作をし，非基底変数 x_2 を基底変数とする．実行可能基底解は，$\boldsymbol{x}^1 = (x_1, x_2, x_3, x_4, x_5)^{\mathrm{T}} = (0, 70, 0, 30, 100)^{\mathrm{T}}$ である．$k \leftarrow 1$ とする．サイクル 1,2 についても同様である．サイクル 2 では $\bar{c}_1 = 0.1 > 0$ より，最適解は $\bar{\boldsymbol{x}} = \boldsymbol{x}^2 = (\bar{x}_1, \bar{x}_2, \bar{x}_3, \bar{x}_4, \bar{x}_5)^{\mathrm{T}} = (15, 62.5, 0, 0, 70)^{\mathrm{T}}$ である．また，最適値は -372.5 である．

5.6　2 段階法

上述のシンプレックス法では，実行可能基底解が得られてから最適解を得ることができる．しかし，初期実行可能基底解の求め方については述べていない．さらに，問題によっては A 行列がランク落ちしている場合や実行可能解が存在しない場合もあり，問題が与えられても直ちにシンプレックス法を適用できるとは限らない．

これらの問題点を克服する方法として，罰金法 (penalty method) と 2 段階法 (two phase method) がある．以下では 2 段階法について説明する．以後，線形計画問題 (5.5) の制約条件の右辺値 $b_i,\ i = 1, 2, \ldots, m$ がすべて非負であると仮定する．そうでなければ，両辺を (-1) 倍すればよい．

2 段階法では，線形計画問題 (5.5) の各制約条件の左辺に非負の人為変数 μ_i を加える．まず，実行可能性をテストする問題を解いた後に，続けて求められた実行可能基底解を用いて線形計画問題 (5.5) をシンプレックス法で解く 2 段階の解法である．

線形計画問題 (5.5) の実行可能性をテストする問題は，

$$
\begin{aligned}
&\text{minimize} \;\; w = \sum_{i=1}^{m} \mu_i \\
&\text{subject to} \;\; A\boldsymbol{x} + \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_m \end{pmatrix} = \boldsymbol{b} \\
&\qquad\qquad \boldsymbol{x} \geq \boldsymbol{0},\; \mu_i \geq 0,\; i = 1, 2, \ldots, m
\end{aligned}
\tag{5.21}
$$

と書ける．線形計画問題 (5.5) に実行可能解が存在すれば，この問題 (5.21) の最適値が 0 となる．逆に，問題 (5.21) の最適値が 0 となれば，線形計画問題 (5.5) に実行可能基底解が存在する．線形計画問題 (5.21) において，初期実行可能解を

$$
(x_1, x_2, \ldots, x_n, \mu_1, \mu_2, \ldots, \mu_m)^{\mathrm{T}} = (0, 0, \ldots, 0, b_1, b_2, \ldots, b_m)^{\mathrm{T}} \tag{5.22}
$$

としてシンプレックス法を適用する．問題 (5.21) の最適値が 0 であれば，$\mu_i = 0,\, i = 1, 2, \ldots, m$ であるので，得られた最適解は元の線形計画問題の実行可能解である．

このとき，

(a) $\mu_i = 0,\, i = 1, 2, \ldots, m$ がすべて非基底変数.
(b) μ_i のいずれかは基底変数であるがその値が 0 となり退化している.

の 2 通りが考えられる．

(a) の場合，変数 $\mu_i,\, i = 1, 2, \ldots, m$ の関与する部分をすべて取り去れば，得られた最適解は元の問題 (5.5) の実行可能基底解となる．つまり，基底形式が得られ，シンプレックス法により元の線形計画問題 (5.5) の最適解を計算することができる．

(b) の場合，求められた最適基底解のうち，μ_i が基底に入り，その変数値が 0 となっている場合は，問題 (5.21) の得られた最適解の基底形式の目的関数の係数が正となっている非基底変数を 0 で固定しなければならない[*3]．非基底変数を 0 で固定するには，これらの変数に関わる部分を以後削除すればよい．

ここで，$\hat{\boldsymbol{y}} = (\boldsymbol{x}^{\mathrm{T}}, \mu_1, \mu_2, \ldots, \mu_m)^{\mathrm{T}}$, $\hat{\boldsymbol{c}} = (\boldsymbol{c}^{\mathrm{T}}, 0, \ldots, 0)^{\mathrm{T}}$, $\hat{\boldsymbol{d}} = (\boldsymbol{0}^{\mathrm{T}}, 1, \ldots, 1)^{\mathrm{T}}$, $\hat{A} = (A, I)$(I は単位行列) と定めると，2 段階法の手順は次のように統一的に記述される．

[*3] この非基底変数が正の値を取れば，問題 (5.21) の目的関数値が正となり，問題 (5.5) の実行可能解ではなくなるから．

アルゴリズム 2PHASE_SIMPLEX

Input　標準形の問題 (5.5)．ただし，$b_i,\ i = 1, 2, \ldots, m$ をがすべて非負であることを仮定する (そうでなければ両辺を -1 倍する)．

Output　最適解 $\bar{\boldsymbol{x}}$，あるいは「実行不能」または「発散」の結論．

1. $\hat{\boldsymbol{y}}^0 = (\boldsymbol{0}^{\mathrm{T}}, \boldsymbol{b}^{\mathrm{T}})^{\mathrm{T}}$ 初期実行可能基底解として，シンプレックス法により，次の線形計画問題の最適解を求める．

$$\begin{aligned} \text{minimize} \quad & w = \hat{\boldsymbol{d}}^{\mathrm{T}} \hat{\boldsymbol{y}} \\ \text{subject to} \quad & \hat{A} \hat{\boldsymbol{y}} = \boldsymbol{b} \\ & \hat{\boldsymbol{y}} \geq \boldsymbol{0} \end{aligned} \tag{5.23}$$

2. 問題 (5.23) の最適値が正の値であれば，元の線形計画問題に実行可能解は存在しない．実行不能であることを結論して計算終了．
3. 問題 (5.23) の最適値が 0 であれば，得られた最適基底解に対応する基底形式

$$\begin{aligned} & w = 0 + (\hat{\boldsymbol{d}}_N^{\mathrm{T}} - \hat{\boldsymbol{d}}_B^{\mathrm{T}} \bar{\hat{N}}) \hat{\boldsymbol{y}}_N \\ & \hat{\boldsymbol{y}}_B + \bar{\hat{N}} \hat{\boldsymbol{y}}_N = \bar{\boldsymbol{b}} \end{aligned} \tag{5.24}$$

を考え，n 次元ベクトル $\bar{\hat{\boldsymbol{d}}}_N = (\bar{\hat{d}}_1, \bar{\hat{d}}_2, \ldots, \bar{\hat{d}}_n)^{\mathrm{T}} = (\hat{\boldsymbol{d}}_N^{\mathrm{T}} - \hat{\boldsymbol{d}}_B^{\mathrm{T}} \bar{\hat{N}})^{\mathrm{T}}$ を定める．$\bar{\hat{d}}_j > 0$ であれば，対応する変数 $\hat{y}_j^N$ を以後無視する．また，$\hat{y}_j,\ j = n, n+1, n+2, \ldots, n+m$ の中で非基底変数であるものも以後無視する．

4. 目的関数の係数ベクトルを $\hat{\boldsymbol{c}}$ に変更した基底形式

$$\begin{aligned} & z = 0 + (\hat{\boldsymbol{c}}_N^{\mathrm{T}} - \hat{\boldsymbol{c}}_B^{\mathrm{T}} \bar{\hat{N}}) \hat{\boldsymbol{y}}_N \\ & \hat{\boldsymbol{y}}_B + \bar{\hat{N}} \hat{\boldsymbol{y}}_N = \bar{\boldsymbol{b}} \end{aligned} \tag{5.25}$$

を求め，無視するすべての変数の列を削除して，SIMPLEX を実行する．

5. SIMPLEX の 2. で終了すれば，元の問題 (5.5) は非有界となる (発散する) ことを結論して計算終了．そうでなければ，最適解が得られる．このとき，無視された変数は非基底変数として扱われ，$\hat{\boldsymbol{y}}$ の最初の n 個の変数の値で構成されるベクトルが元の問題 (5.5) の最適解となる．

ここで，次の例題を解く．

$$\begin{aligned} \text{minimize} \quad & z = 9x_1 + 5x_2 + 8x_3 \\ \text{subject to} \quad & 2.5x_1 + 3x_2 + 5x_3 - x_4 = 200 \\ & 2.5x_1 + 2x_2 + 3x_3 - x_5 = 160 \\ & 3x_1 + x_2 + 2x_3 - x_6 = 120 \\ & x_1, x_2, x_3, x_4, x_5, x_6 \geq 0 \end{aligned} \tag{5.26}$$

問題 (5.26) の実行可能性をテストする問題 (第 1 段階の問題) は，

$$
\begin{aligned}
\text{minimize} \quad & w = \mu_1 + \mu_2 + \mu_3 \\
\text{subject to} \quad & 2.5x_1 + 3x_2 + 5x_3 - x_4 + \mu_1 = 200 \\
& 2.5x_1 + 2x_2 + 3x_3 - x_5 + \mu_2 = 160 \\
& 3x_1 + x_2 + 2x_3 - x_6 + \mu_3 = 120 \\
& x_1, x_2, x_3, x_4, x_5, x_6, \mu_1, \mu_2, \mu_3 \geq 0
\end{aligned}
\tag{5.27}
$$

である．第 1 段階のシンプレックスタブローを表 5.3 に示す．

表 5.3: シンプレックスタブロー (第 1 段階)

サイクル	基底	x_1	x_2	x_3	x_4	x_5	x_6	右辺
0	μ_1	2.5	3	[5]	−1			200
	μ_2	2.5	2	3		−1		160
	μ_3	3	1	2			−1	120
	$-w$	−8	−6	−10	1	1	1	−480
1	x_3	0.5	0.6	1	−0.2			40
	μ_2	1	0.2		0.6	−1		40
	μ_3	[2]	−0.2		0.4		−1	40
	$-w$	−3			−1	1	1	−80
2	x_3		0.65	1	−0.3		0.25	30
	μ_2		0.3		0.4	−1	[0.5]	20
	x_1	1	−0.1		0.2		−0.5	20
	$-w$		−0.3		−0.4	1	−0.5	−20
3	x_3		0.5	1	−0.5	0.5		20
	x_6		0.6		0.8	−2	1	40
	x_1	1	0.2		0.6	−1		40
	$-w$							0

第 1 段階の問題 (5.27) の最適値が 0 より，元の問題 (5.26) は実行可能であり，問題 (5.27) の最適解 $(x_1, x_2, x_3, x_4, x_5, x_6)^{\mathrm{T}} = (40, 0, 20, 0, 0, 40)^{\mathrm{T}}$ を初期実行可能基底解として解くことができる．基底変数は x_1, x_3, x_6 であり，目的関数を非基底変数のみを含む式に書き直すと，

$$
z = 520 - 0.8x_2 - 1.4x_4 + 5x_5 \tag{5.28}
$$

表 **5.4:** シンプレックスタブロー (第 2 段階)

サイクル	基底	x_1	x_2	x_3	x_4	x_5	x_6	右辺
	x_3		0.5	1	-0.5	0.5		20
$0'$	x_6		0.6		[0.8]	-2	1	40
	x_1	1	0.2		0.6	-1		40
	$-z$		-0.8		-1.4	5		-520
	x_3		0.875	1		-0.75	0.625	45
$1'$	x_4		0.75		1	-2.5	1.25	50
	x_1	1	-0.25			-0.5	-0.75	10
	$-z$		0.25			0.15	1.75	-450

である[*4]．

第 2 段階のタブローを表 5.4 に示す．タブローより，最適解は $(\bar{x}_1, \bar{x}_2, \bar{x}_3, \bar{x}_4, \bar{x}_5, \bar{x}_6)^{\mathrm{T}} = (10, 0, 45, 0, 50, 0)^{\mathrm{T}}$ であり，最適値は 450 である．

5.7 シンプレックス法の収束性

シンプレックス法は，退化が起こらないという仮定の下では，有限回の反復で最適解に到達することが示されており，多くの場合最適解を見つけることができる．しかし，この仮定が成立しない場合には，有限回の反復で最適解が求められるとは限らない．シンプレックス法のサイクルを繰り返すだけで最適解が求まらないような現象を巡回 (cycling) という[*5]．巡回が起こる場合，次の Bland の巡回対策を用いて巡回から抜け出し，最適解を求めることができる．

[*4] シンプレックスタブローより，

$$x_1 + 0.2x_2 + 0.6x_4 - x_5 = 40 \quad \therefore x_1 = 40 - 0.2x_2 - 0.6x_4 + x_5$$
$$0.5x_2 + x_3 - 0.5x_4 + 0.5x_5 = 20 \quad \therefore x_3 = 20 - 0.5x_2 + 0.5x_4 - 0.5x_5$$

なので，

$$\begin{aligned} z &= 9x_1 + 5x_2 + 8x_3 \\ &= 9(40 - 0.2x_2 - 0.6x_4 + x_5) + 5x_2 + 8(20 - 0.5x_2 + 0.5x_4 - 0.5x_5) \\ &= 520 - 0.8x_2 - 1.4x_4 + 5x_5 \end{aligned}$$

[*5] 目的関数値が変化しない．

Blandの巡回対策を含めたアルゴリズム SIMPLEX

Input　標準形の問題 (5.5)．ただし，$\boldsymbol{b} \geq \boldsymbol{0}$ を仮定し，(5.7) に等価変換する．

Output　最適解 $\bar{\boldsymbol{x}}$，あるいは「発散」の結論．

1　初期実行可能基底解 $\boldsymbol{x}^0$ を求める．$k \leftarrow 0$.

2B　$\bar{\boldsymbol{c}}^{\mathrm{T}} := \boldsymbol{c}_N^{\mathrm{T}} - \boldsymbol{c}_B^{\mathrm{T}} B^{-1} N$ によって得られる $\bar{\boldsymbol{c}}$ の要素 $\bar{c}_j$ を用いて，

$$\min\{j \mid \bar{c}_j\} = r$$

となる r を求める．

つまり，$\bar{c}_j < 0$ であるような j が 2 個以上存在するとき，添字 j の値が最小のものを r とする．

すべての $\bar{c}_j$ が $\bar{c}_j \geq 0$ であれば，$\bar{\boldsymbol{x}} \leftarrow \boldsymbol{x}^k$ として計算終了．

3　$\bar{N} := B^{-1}N(= (\bar{a}_{ij}))$ に対して，第 r 列のすべての成分 $\bar{a}_{ir}$ が $\bar{a}_{ir} \leq 0$ であれば，最小値が有界でない (問題が発散する) ことを結論して計算終了．

4B　$\bar{a}_{ir}$ に正のものがあれば，

$$\min_{i:\, \bar{a}_{ir} > 0} \frac{\bar{b}_i}{\bar{a}_{ir}} = \frac{\bar{b}_s}{\bar{a}_{sr}}$$

となる s を求める．

ただし，最小値を与える i が 2 つ以上存在するとき，添字 i の値が最小のものを s とする．

5　$\bar{a}_{sr}$ に対するピボット操作により x_s^B の代わりに x_r^N を基底変数とする新しい実行可能基底解 $\boldsymbol{x}^{k+1}$ を求める．$k \leftarrow k+1$ として 2B に戻る．

Blandの巡回対策を用いると巡回から抜け出すことができる．

定理 5.2　手順 2B で r を決定し，手順 4B で s を決定するとき，SIMPLEX は有限回の反復で終了する．

通常のシンプレックス法での列の選び方は，目的関数値の改善割合が大きい変数を基底に入れ，値を取らせることでより早く最適解に到達するだろうというヒューリスティックに基づいていた．このヒューリスティックは有効だが，巡回に陥り最適解が求まらないことがある．一方，Blandの巡回対策では，巡回することはないと保証されているが，効率的ではない．現実には，通常のシンプレックス法で開始し，目的関数が一定回数変化しなければ，Blandの巡回対策に切り替え，目的関数が変化すればもとの方法に戻すといった方法がとられることが多い．

5.8　線形計画問題の双対性

標準形の線形計画問題

$$\begin{array}{rl} \text{minimize} & \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} \\ \text{subject to} & A\boldsymbol{x} = \boldsymbol{b} \\ & \boldsymbol{x} \geq \boldsymbol{0} \end{array} \tag{5.29}$$

に対して，線形計画問題

$$\begin{aligned} &\text{maximize} \quad \boldsymbol{b}^{\mathrm{T}}\boldsymbol{y} \\ &\text{subject to} \quad A^{\mathrm{T}}\boldsymbol{y} \leq \boldsymbol{c} \end{aligned} \tag{5.30}$$

は双対問題である．一般に，元の問題 (主問題) を (P)，双対問題を (D) と記す．ここで，双対問題の変数ベクトル $\boldsymbol{y}$ は双対変数ベクトル (dual variable vector) と呼ばれる．線形計画問題の主問題 (5.29) と双対問題 (5.30) には次の弱双対定理が成立する．

定理 5.3（弱双対定理） $\boldsymbol{x}, \boldsymbol{y}$ をそれぞれ主問題 (5.29)，双対問題 (5.30) の実行可能解とすると，次の関係が成立する[*6]．

$$\boldsymbol{b}^{\mathrm{T}}\boldsymbol{y} \leq \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} \tag{5.31}$$

問題 (5.30) において，$\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}$ は上界値であり，問題 (5.29) において，$\boldsymbol{b}^{\mathrm{T}}\boldsymbol{y}$ は下界値である．つまり，主問題および双対問題に実行可能解が存在することがわかっているとき，主問題の実行可能解 $\boldsymbol{x}$ を見つければ，双対問題の最適値が $\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}$ 以下であることがわかり，双対問題の実行可能解 $\boldsymbol{y}$ を見つければ，主問題の最適値が $\boldsymbol{b}^{\mathrm{T}}\boldsymbol{y}$ 以上であることがわかる．

系 1 主問題 (5.29)，双対問題 (5.30) の実行可能解 $\boldsymbol{x}, \boldsymbol{y}$ について，

$$\boldsymbol{b}^{\mathrm{T}}\boldsymbol{y} = \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} \tag{5.32}$$

が成立する時，$\boldsymbol{x}, \boldsymbol{y}$ はそれぞれ主問題，双対問題の最適解となる．

定理 5.4（強双対定理 (Strong Duality)）

1. 主問題 (5.29) および双対問題 (5.30) のいずれにも実行可能解が存在すれば，いずれも最適解が存在し (非有界でない)，最適値は一致する．
2. 主問題が非有界であれば，双対問題に実行可能解が存在しない．逆に，双対問題が非有界であれば，主問題に実行可能解が存在しない．

証明

1. について．定理 5.3 より，(5.29)，(5.30) のいずれにも実行可能解が存在すれば，それぞれ下限値，上限値がある．よって，いずれの問題も非有界になることはなく，最適解は存在する．

主問題 (5.29) の最適解が存在するとき，最適基底解 $\hat{\boldsymbol{x}}$ も存在する．対応する基底行列を B とすると，基底形式は

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

であり，この基底形式に対応する基底解が最適解である．最適性より，

$$\boldsymbol{c}_N^{\mathrm{T}} - \boldsymbol{c}_B^{\mathrm{T}} B^{-1} N \geq \boldsymbol{0}$$

---

*6

$$\boldsymbol{b}^{\mathrm{T}}\boldsymbol{y} = \boldsymbol{y}^{\mathrm{T}}\boldsymbol{b} = \boldsymbol{y}^{\mathrm{T}}A\boldsymbol{x} \leq \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}$$

より，定理が得られる．

である．これと，

$$\boldsymbol{c}_B^{\mathrm{T}} - \boldsymbol{c}_B^{\mathrm{T}} = \boldsymbol{c}_B^{\mathrm{T}} - \boldsymbol{c}_B^{\mathrm{T}} B^{-1} B = \boldsymbol{0}$$

を合わせて，

$$\begin{aligned} (\boldsymbol{c}_B^{\mathrm{T}}\ \boldsymbol{c}_N^{\mathrm{T}}) - \boldsymbol{c}_B^{\mathrm{T}} B^{-1}(B\ N) &\geq \boldsymbol{0} \\ \therefore\ \boldsymbol{c}^{\mathrm{T}} - \boldsymbol{c}_B^{\mathrm{T}} B^{-1} A &\geq \boldsymbol{0} \end{aligned}$$

を得る．ここで，$\boldsymbol{\pi}^{\mathrm{T}} \coloneqq \boldsymbol{c}_B^{\mathrm{T}} B^{-1}$ とすると，$\boldsymbol{\pi}^{\mathrm{T}} A \leq \boldsymbol{c}^{\mathrm{T}}$ より，

$$A^{\mathrm{T}} \boldsymbol{\pi} \leq \boldsymbol{c}$$

となり，$\boldsymbol{\pi}$ は (5.30) の実行可能解である．目的関数値は，

$$\boldsymbol{b}^{\mathrm{T}} \boldsymbol{\pi} = \boldsymbol{\pi}^{\mathrm{T}} \boldsymbol{b} = \boldsymbol{c}_B^{\mathrm{T}} B^{-1} \boldsymbol{b} = \boldsymbol{c}^{\mathrm{T}} \hat{\boldsymbol{x}}$$

より，(5.29) の最適値と一致する．

弱双対定理より，(5.30) の任意の実行可能解 $\boldsymbol{y}$ に対して

$$\boldsymbol{b}^{\mathrm{T}} \boldsymbol{y} \leq \boldsymbol{c}^{\mathrm{T}} \hat{\boldsymbol{x}} = \boldsymbol{b}^{\mathrm{T}} \boldsymbol{\pi}$$

である．よって，$\boldsymbol{\pi}$ は (5.30) の最適解となり，$\boldsymbol{b}^{\mathrm{T}} \boldsymbol{\pi}$ は (5.30) の最適値となる．つまり，主問題 (5.29) と双対問題 (5.30) の最適値は一致する． □

2. について．(5.29) が非有界であるとき，(5.30) に実行可能解 $\boldsymbol{y}$ が存在すると仮定する．弱双対定理より，(5.29) の任意の実行可能解 $\boldsymbol{x}$ に対して，

$$\boldsymbol{c}^{\mathrm{T}} \boldsymbol{x} \geq \boldsymbol{b}^{\mathrm{T}} \boldsymbol{y}$$

が成立する．これは (5.29) が非有界であることに矛盾する．後半も同様に証明される． □

定理 5.4 1. の証明より，主問題の最適基底解 $\hat{\boldsymbol{x}}$ が見つかれば，双対問題の最適解の 1 つは主問題のシンプレックス乗数を用いて $\boldsymbol{\pi} = (B^{-1})^{\mathrm{T}} \boldsymbol{c}_B$ と与えられる．特に，主問題 (5.29) の最適基底解が退化していない場合は，制約条件の右辺値 $\boldsymbol{b}$ がどのように微小変化しても実行可能性が損なわれず*7，最適値は

$$\boldsymbol{c}^{\mathrm{T}} \hat{\boldsymbol{x}} = \boldsymbol{b}^{\mathrm{T}} \boldsymbol{\pi}$$

と書くことができる．つまり，$\boldsymbol{\pi}$ の第 i 成分 π_i は，i 番目の右辺値 b_i が微小変化するときの目的関数の変化量を示しており，b_i に対する感度 (sensitivity) と考えることができる．ただし，(5.29) の最適基底解が退化している場合には，複数の最適基底形式が存在し，それぞれにシンプレックス乗数が存在するので，ある 1 つのシンプレックス乗数のみからこのような情報が得られるとは限らない．

以上の定理より，主問題と双対問題の最適解の存在，非有界，実行不能の組み合わせの可否を表 5.5 にまとめる．

*7 最適性条件にも影響が出ず，最適性も損なわれない．

表 **5.5:** 主問題と双対問題の関係

主問題 / 双対問題	最適解存在	非有界	実行不能
最適解存在	◯	×	×
非有界	×	×	◯
実行不能	×	◯	◯

問題 (5.29) と (5.30) の非対称な組み合わせ以外に，次の互いに対称な 2 つの問題 (5.33), (5.34) も双対な関係にあることが知られている．

$$\begin{aligned} \text{minimize} \quad & \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} \\ \text{subject to} \quad & A\boldsymbol{x} \geq \boldsymbol{b} \\ & \boldsymbol{x} \geq \boldsymbol{0} \end{aligned} \tag{5.33}$$

$$\begin{aligned} \text{maximize} \quad & \boldsymbol{b}^{\mathrm{T}}\boldsymbol{y} \\ \text{subject to} \quad & A^{\mathrm{T}}\boldsymbol{y} \leq \boldsymbol{c} \\ & \boldsymbol{y} \geq \boldsymbol{0} \end{aligned} \tag{5.34}$$

双対定理より，主問題を解くことと双対問題を解くことが等価であることがわかる．

5.9　双対シンプレックス法

シンプレックス法では，実行可能基底解を最初に求める必要があった．初期実行可能基底解が簡単に与えられない場合は，2 段階法を用いて実行可能基底解を求めるなどの工夫が必要だった．しかし，容易に求められる基底解が双対問題の実行可能解ならば，双対問題を解くことにより，与えられた問題をシンプレックス法で解くことができる．この考えに基づいた手法が双対シンプレックス法である．

定理 5.4 1. の証明で示されていることと同様に，ある基底解

$$\begin{aligned} & z = \boldsymbol{c}_B^{\mathrm{T}}B^{-1}\boldsymbol{b} + (\boldsymbol{c}_N - \boldsymbol{c}_B^{\mathrm{T}}B^{-1}N)\boldsymbol{x}_N \\ & \boldsymbol{x}_B + B^{-1}N\boldsymbol{x}_N = B^{-1}\boldsymbol{b} \end{aligned} \tag{5.35}$$

において，

$$\boldsymbol{c}_N^{\mathrm{T}} - \boldsymbol{c}_B^{\mathrm{T}}B^{-1}N \geq \boldsymbol{0} \tag{5.36}$$

が成立していれば，

$$\boldsymbol{c}^{\mathrm{T}} - \boldsymbol{c}_B^{\mathrm{T}}B^{-1}A \geq \boldsymbol{0} \tag{5.37}$$

が成立し，$\boldsymbol{\pi} = (B^{-1})^{\mathrm{T}}\boldsymbol{c}_B$ は双対問題の実行可能解となる．よって，(5.36) 式が成立する基底解から双対問題を解くことと等価な計算を行うことで最適解を求めることができる．

(5.36) 式が成立する主問題の基底解が元の問題で実行可能である，つまり，$\bar{\boldsymbol{b}} = B^{-1}\boldsymbol{b}$ が成立していれば，その基底解は双対問題における最適解である．よって，$\bar{\boldsymbol{b}} \ngeq \boldsymbol{0}$ という基底解から目的関数を改善するようにピ

ボット演算を実行していくことにより，双対問題の最適解を求めることになる．これに関して，以下の定理が得られる．

定理 5.5 $\bar{\boldsymbol{b}}$ の第 r 成分 $\bar{b}_r$ が負，$\bar{b}_r < 0$ とし，$J = \{j \,|\, \bar{a}_{rj} < 0\} \neq \emptyset$ であるとする．このとき，

$$\max_{j \in J} \frac{\bar{c}_j}{\bar{a}_{rj}} = \frac{\bar{c}_s}{\bar{a}_{rs}} = \Delta,\ s \in J \tag{5.38}$$

とすると，$\bar{a}_{rs}$ を軸とするピボット演算を施せば，双対問題の実行可能性を失うことなく，双対問題の目的関数値を $\bar{z}$ から $\bar{b}_r\Delta$ だけ大きくする (改善する) ことができる[*8]．

定理 5.6 $\bar{b}_r < 0$ に対して，$J = \{j \,|\, \bar{a}_{rj} < 0\} = \emptyset$，すなわち，$\bar{a}_{rj} \geq 0,\ j = 1, 2, \ldots, n - m$ のとき，元の問題に実行可能解は存在しない．

証明

基底形式の第 r 番目の制約条件を考えると，

$$x_r + \sum_{j=1}^{n-m} \bar{a}_{rj} x_j = \bar{b}_r$$

となる．$x_r \geq 0,\ x_j \geq 0$ という非負条件より，左辺は非負の値しか取りえない．一方，右辺は $\bar{b}_r < 0$ より，明らかに非負条件を満たす元の問題の実行可能解は存在しない． □

以上より，(5.36) 式が成立する基底解が得られているとき，双対シンプレックス法のアルゴリズムは以下で表される．

アルゴリズム DUAL_SIMPLEX

Input 標準形の問題 (5.29)．

Output 最適解，あるいは「実行不能」の結論．

1. $\min_{\bar{b}_i < 0} \bar{b}_i = \bar{b}_r$ なる r を求める．これが $\bar{b}_r$ であれば，現在の基底解が最適解となり，計算終了．
2. $\bar{a}_{rj} \geq 0,\ j = 1, 2, \ldots, n - m$ であれば，実行可能解が存在しない．実行不能であることを結論し，計算終了．
3. $$\max_{\bar{a}_{rj} < 0} \frac{\bar{c}_j}{\bar{a}_{rj}} = \frac{\bar{c}_s}{\bar{a}_{rs}}$$
   なる s を求める．
4. x_r を基底から出し，x_s を基底に入れる $\bar{a}_{rs}$ を軸とするピボット演算を行い，新しい基底形式を求め，1. に戻る．

[*8] 主問題の制約条件の右辺値が負となっている第 r 行について，その係数が負となっている要素 $\bar{a}_{rj}$ を見る．目的関数の係数ベクトルの要素 $\bar{c}_j$ について，$\bar{a}_{rj}$ で割った値の絶対値が最小のもの (双対問題の実行可能性より $\bar{c}_j \geq 0$ なので，これは値が最大のものと等価である) の添字 s を見る．この $\bar{a}_{rs}$ を軸としてピボット演算をすればよい．

例として，次の線形計画問題を双対シンプレックス法で解く．

$$
\begin{aligned}
\text{minimize} \quad & 350x_1 + 450x_2 + 240x_3 \\
\text{subject to} \quad & 2.5x_1 + 5x_2 + 3x_3 - x_4 = 4 \\
& 5x_1 + 6x_2 + 2x_3 - x_5 = 5 \\
& x_1, x_2, x_3, x_4, x_5 \geq 0
\end{aligned}
\tag{5.39}
$$

この問題を双対シンプレックス法で解いたときのシンプレックスタブローを表 5.6 に示す．

表 5.6: シンプレックスタブロー (双対)

サイクル	基底	x_1	x_2	x_3	x_4	x_5	右辺
	x_4	-2.5	-5	-3	1		-4
0	x_5	$[-5]$	-6	-2		1	-5
	$-z$	350	450	240			0
	x_4		$[-2]$	-2	1		$-3/2$
1	x_1	1	$6/5$	$2/5$		$-1/5$	1
	$-z$		30	100		70	-350
	x_2		1	1	$-1/2$	$1/4$	$3/4$
2	x_1	1		$-4/5$	$3/5$	$-1/2$	$1/10$
	$-z$			70	15	62.5	-372.5

タブローより，最適解は $(\bar{x}_1, \bar{x}_2, \bar{x}_3, \bar{x}_4, \bar{x}_5)^{\mathrm{T}} = (1/10, 3/4, 0, 0, 0)^{\mathrm{T}}$ であり，最適値は 372.5 である．

5.10 再最適化

線形計画問題 (標準形) は，パラメータ $A, \boldsymbol{b}, \boldsymbol{c}$ が与えられれば定められる．しかし，現実の問題ではこれらを正確に見積もることが困難である場合も多い．このような場合，パラメータの微小変化に対して最適解や最適値がどのように変化するかを見る感度分析や，パラメータを変化した後の最適解を見る再最適化が有用である．ここでは，いくつかの場合の再最適化について述べる．

5.10.1 c が変化したとき

目的関数の係数ベクトルが $\boldsymbol{c}$ から $\boldsymbol{c}'$ に変化した場合，最適タブローの $-z$ の行のみが影響を受ける．基底形式より，$-z$ の行の非基底変数の列は

$$
\boldsymbol{c}_N^{\mathrm{T}} - \boldsymbol{c}_B^{\mathrm{T}} B^{-1} N \tag{5.40}
$$

の値をとり，右辺値の列は $-\boldsymbol{c}_B^{\mathrm{T}} B^{-1} \boldsymbol{b}$ を取る．

目的関数の係数ベクトルが $\boldsymbol{c}$ から $\boldsymbol{c}'$ に変化したとき，

$$
\boldsymbol{c}_N'^{\mathrm{T}} - \boldsymbol{c}_B'^{\mathrm{T}} B^{-1} N \leq \boldsymbol{0} \tag{5.41}
$$

であれば，最適性は保存されるので，最適解は変化しない．最適値は $\boldsymbol{c}_B'^{\mathrm{T}} B^{-1}\boldsymbol{b}$ に変化する．

(5.41) が成立しなければ，$-z$ の行のいずれかの成分は負となっている．右辺値 $B^{-1}\boldsymbol{b}$ は変化しておらず，非負のままであるので，$-z$ の行の非基底変数列を $\boldsymbol{c}_N'^{\mathrm{T}} - \boldsymbol{c}_B'^{\mathrm{T}} B^{-1}N$ で置き換えて，シンプレックス法を適用して $\boldsymbol{c}'$ に変化した後の最適解を求めることができる．

5.10.2 $\boldsymbol{b}$ が変化したとき

制約条件の右辺値ベクトル $\boldsymbol{b}$ が $\boldsymbol{b}'$ に変化した場合は，最適タブローの右辺値 (定数) の列のみが変化する．右辺値の列は $B^{-1}\boldsymbol{b}$ の値をとり，$\boldsymbol{b}'$ への変化後のこれらの値に注意すればよい．

$\boldsymbol{b}'$ に変化したとき，

$$B^{-1}\boldsymbol{b}' \geq \boldsymbol{0} \tag{5.42}$$

であれば実行可能性が保存され，最適基底は変化しない．ただし，最適解は $\boldsymbol{x}_B = B^{-1}\boldsymbol{b}'$, $\boldsymbol{x}_N = \boldsymbol{0}$ に変化する．

(5.42) が成立しなければ，$-z$ の行のいずれの成分も非負のまま，すなわち，双対問題の実行可能性を保っているので，右辺値の列で負になっている行を選び，双対シンプレックス法を適用すれば $\boldsymbol{b}'$ に変化したときの最適解が求められる．

最後に，次の例題でこれらの変化について確認する．

$$\begin{aligned}
\text{minimize} \quad & -2.5x_1 - 5x_2 - 3.4x_3 \\
\text{subject to} \quad & 2x_1 + 10x_2 + 4x_3 + x_4 = 425 \\
& 6x_1 + 5x_2 + 8x_3 + x_5 = 400 \\
& 7x_1 + 10x_2 + 8x_3 + x_6 = 600 \\
& x_1, x_2, x_3, x_4, x_5, x_6 \geq 0
\end{aligned}$$

この問題のシンプレックスタブローは表 5.7 で表され，最適解は $(x_1, x_2, x_3, x_4, x_5, x_6)^{\mathrm{T}} = (30, 34, 6.25, 0, 0, 0)^{\mathrm{T}}$，最適値は -266.25 となる．

まず，目的関数の係数ベクトルが $\boldsymbol{c} = (-2.5, -5, -3.4, 0, 0, 0)^{\mathrm{T}}$ から $\boldsymbol{c}' = (-3, -5, -4, 0, 0, 0)^{\mathrm{T}}$ に変化したとき，最適タブローの $-z$ の行の非基底変数 x_4, x_5, x_6 の列の値は，

$$B^{-1}N = \begin{pmatrix} 0.08 & -0.12 & 0.08 \\ 0.25 & 0.5 & -0.5 \\ -0.4 & -0.4 & 0.6 \end{pmatrix}$$

であるので，

$$\begin{pmatrix} 0 & 0 & 0 \end{pmatrix} - \begin{pmatrix} -5 & -4 & -3 \end{pmatrix} \begin{pmatrix} 0.08 & -0.12 & 0.08 \\ 0.25 & 0.5 & -0.5 \\ -0.4 & -0.4 & 0.6 \end{pmatrix} = \begin{pmatrix} 0.2 & 0.2 & 0.2 \end{pmatrix} \leq \boldsymbol{0}$$

となるので，解 $(x_1, x_2, x_3, x_4, x_5, x_6)^{\mathrm{T}} = (30, 34, 6.25, 0, 0, 0)^{\mathrm{T}}$ は最適解のままである．このとき，最適値のままである．このとき，最適値は，

$$z = -3 \times 30 - 5 \times 34 - 4 \times 6.25 = -285$$

表 **5.7:** シンプレックスタブロー

サイクル	基底	x_1	x_2	x_3	x_4	x_5	x_6	右辺
0	x_4	2	[10]	4	1			425
	x_5	6	5	8		1		400
	x_6	7	10	8			1	600
	$-z$	-2.5	-5	-3.4				0
1	x_2	0.2	1	0.4	0.1			42.5
	x_5	5		6	-0.5	1		187.5
	x_6	[5]		4	-1		1	175
	$-z$	-1.5		-1.4	0.5			212.5
2	x_2		1	0.24	0.14		-0.04	35.5
	x_5			[2]	0.5	1	-1	12.5
	x_1	1		0.8	-0.2		0.2	35
	$-z$			-0.2	0.2		0.3	265
3	x_2		1		0.08	-0.12	0.08	34
	x_3			1	0.25	0.5	-0.5	6.25
	x_1	1			-0.4	-0.4	0.6	30
	$-z$				0.25	0.1	0.2	266.25

となる．

さらに，この問題の目的関数の係数ベクトルが $\boldsymbol{c}'' = (-2,-5,-4,0,0,0)^{\mathrm{T}}$ に変化したとき，最適タブローの $-z$ の行の x_4, x_5, x_6 の列の値は，

$$\begin{pmatrix} 0 & 0 & 0 \end{pmatrix} - \begin{pmatrix} -5 & -4 & -2 \end{pmatrix} \begin{pmatrix} 0.08 & -0.12 & 0.08 \\ 0.25 & 0.5 & -0.5 \\ -0.4 & -0.4 & 0.6 \end{pmatrix} = \begin{pmatrix} 0.6 & 0.6 & -0.4 \end{pmatrix} \not\leq \mathbf{0}$$

となるので，解 $(x_1,x_2,x_3,x_4,x_5,x_6)^{\mathrm{T}} = (30,34,6.25,0,0,0)^{\mathrm{T}}$ は最適解でなくなる．

そこで，$-z$ の行を書き換え，シンプレックス法を適用すると，次のタブローのようになり，新しい目的係数ベクトル $\boldsymbol{c}'$ に対する最適解 $(x_1,x_2,x_3,x_4,x_5,x_6)^{\mathrm{T}} = (0,30,31.25,0,0,50)^{\mathrm{T}}$，最適値 $z=-275$ が求められる．

次に，右辺値ベクトルが $\boldsymbol{b} = (425,400,600)^{\mathrm{T}}$ から $\boldsymbol{b}' = (450,400,600)^{\mathrm{T}}$ に変化した場合と，$\boldsymbol{b}'' = (350,420,600)^{\mathrm{T}}$ に変化した場合を考える．

$\boldsymbol{b}'$ に変化した場合は，

$$B^{-1}\boldsymbol{b}' = \begin{pmatrix} 0.08 & -0.12 & 0.08 \\ 0.25 & 0.5 & -0.5 \\ -0.4 & -0.4 & 0.6 \end{pmatrix} \begin{pmatrix} 450 \\ 400 \\ 600 \end{pmatrix} = \begin{pmatrix} 36 \\ 12.5 \\ 20 \end{pmatrix} \geq \mathbf{0}$$

となり，実行可能性を損なわず，最適基底は変化しない．最適解は，$(x_1,x_2,x_3,x_4,x_5,x_6) =$

表 **5.8:** シンプレックスタブロー

サイクル	基底	x_1	x_2	x_3	x_4	x_5	x_6	右辺
	x_2		1		0.08	-0.12	0.08	34
I	x_3			1	0.25	0.5	-0.5	6.25
	x_1	1			-0.4	-0.4	[0.6]	30
	$-z$				0.6	0.6	-0.4	255
	x_2	$-2/15$	1		$2/15$	$-1/15$		30
II	x_3	$5/6$		1	$1/12$	$1/6$		31.25
	x_6	$5/3$			$-2/3$	$-2/3$	1	50
	$-z$	$2/3$			$1/3$	$1/3$		275

$(20, 36, 12.5, 0, 0, 0)^{\mathrm{T}}$，最適値は，

$$z = -2.5 \times 20 - 5 \times 36 - 3.4 \times 12.5 = -272.5$$

となる．

一方，$\boldsymbol{b}''$ に変化した場合は，

$$B^{-1}\boldsymbol{b}' = \begin{pmatrix} 0.08 & -0.12 & 0.08 \\ 0.25 & 0.5 & -0.5 \\ -0.4 & -0.4 & 0.6 \end{pmatrix} \begin{pmatrix} 350 \\ 420 \\ 600 \end{pmatrix} = \begin{pmatrix} 25.6 \\ -2.5 \\ 52 \end{pmatrix} \geq \mathbf{0}$$

となり，実行可能性を失う．そこで，双対シンプレックス法により，$\boldsymbol{b}''$ に変化した場合の最適解を求めると，次のタブローが得られ，$(x_1, x_2, x_3, x_4, x_5, x_6) = (49, 25.2, 0, 0, 0, 5)^{\mathrm{T}}$，最適値は，$-248.5$ となる．

表 **5.9:** シンプレックスタブロー

サイクル	基底	x_1	x_2	x_3	x_4	x_5	x_6	右辺
	x_2		1		0.08	-0.12	0.08	25.6
I	x_3			1	0.25	0.5	$[-0.5]$	-2.5
	x_1	1			-0.4	-0.4	0.6	52
	$-z$				0.25	0.1	0.2	249.5
	x_2		1	0.16	0.12	-0.04		25.2
II	x_6			-2	-0.5	-1	1	5
	x_1	1		1.2	-0.1	0.2		49
	$-z$			0.4	0.35	0.3		248.5

第 6 章

最適化問題の双対性

本章では，非線形計画問題および凸計画問題における双対問題の性質と役割について述べる．

6.1　ラグランジュ双対問題

非線形計画問題において，すべての制約条件が不等式で与えられている場合を考える．つまり，

$$
\begin{aligned}
\text{minimize} \quad & f(\boldsymbol{x}) \\
\text{subject to} \quad & g_i(\boldsymbol{x}) \leq 0,\ i = 1, 2, \ldots, m \\
& \boldsymbol{x} \in X
\end{aligned}
\tag{6.1}
$$

を考える．集合 $X \subseteq \mathbb{R}^n$ は任意だが，以下の議論では $X = \mathbb{R}^n$ あるいは $X = \mathbb{R}^n_+ \ (= \{\boldsymbol{x} \in \mathbb{R}^n \mid \boldsymbol{x} \geq \boldsymbol{0}\})$ を想定する．この問題の Lagrange 関数は，ラグランジュ乗数ベクトル $\boldsymbol{u}$ を用いて

$$
L(\boldsymbol{x}, \boldsymbol{u}) = f(\boldsymbol{x}) + \boldsymbol{u}^{\mathrm{T}} g(\boldsymbol{x}),\ \boldsymbol{x} \in X, \boldsymbol{u} \in \mathbb{R}^m_+ \tag{6.2}
$$

と定義される．

Lagrange 関数と元の最適化問題の関係を調べるため，$x \in X$ を固定し，$\boldsymbol{u}$ に関する最大化問題

$$
\begin{aligned}
\text{maximize} \quad & L(\boldsymbol{x}, \boldsymbol{u}) \\
\text{subject to} \quad & \boldsymbol{u} \in \mathbb{R}^m_+
\end{aligned}
\tag{6.3}
$$

を導入し，この最大値 (∞ の場合もある)$\sup_{\boldsymbol{u} \geq \boldsymbol{0}} L(\boldsymbol{x}, \boldsymbol{u})$ を $F(\boldsymbol{x})$ と表す．

1. $\exists i\ s.t.\ g_i(\boldsymbol{x}) > 0$ のとき
   添字 i に対応する u_i を大きくすれば問題 (6.3) の目的関数をいくらでも大きくすることができる．
2. $g(\boldsymbol{x}) \leq \boldsymbol{0}$ のとき
   $\boldsymbol{u}^{\mathrm{T}} g(\boldsymbol{x}) = 0$(例えば $\boldsymbol{u} = \boldsymbol{0}$) のとき，目的関数は最大値をとり，その値は $f(\boldsymbol{x})$ である．

よって，

$$
F(\boldsymbol{x}) = \begin{cases} f(\boldsymbol{x}), & g(\boldsymbol{x}) \leq \boldsymbol{0} \\ \infty, & \text{otherwise} \end{cases}
$$

となる．したがって，元の問題 (6.1) の最適値を f^* とすると，

$$f^* = \inf_{\boldsymbol{x} \in X} F(\boldsymbol{x}) = \inf_{\boldsymbol{x} \in X} \sup_{\boldsymbol{u} \geq \boldsymbol{0}} L(\boldsymbol{x}, \boldsymbol{u})$$

を得る．ただし，$g(\boldsymbol{x}) \leq \boldsymbol{0}$ を満たす実行可能解 $\boldsymbol{x}$ がなければこの値は ∞ である．場合によっては $-\infty$ に発散することもある．

補題 6.1（inf-sup と sup-inf） Lagrange 関数 $L(\boldsymbol{x}, \boldsymbol{u})$ は次の不等式を満たす[*1]．

$$\inf_{\boldsymbol{x} \in X} \sup_{\boldsymbol{u} \geq \boldsymbol{0}} L(\boldsymbol{x}, \boldsymbol{u}) \geq \sup_{\boldsymbol{u} \geq \boldsymbol{0}} \inf_{\boldsymbol{x} \in X} L(\boldsymbol{x}, \boldsymbol{u})$$

■双対問題と弱双対定理 Lagrange 関数 $L(\boldsymbol{x}, \boldsymbol{u})$ において，$\boldsymbol{u} \in \mathbb{R}^m_+$ を固定したのち，$\boldsymbol{x}$ に関する次の最小化問題を考える．

$$\begin{aligned} \text{minimize} \quad & L(\boldsymbol{x}, \boldsymbol{u}) \\ \text{subject to} \quad & \boldsymbol{x} \in X \end{aligned} \tag{6.4}$$

問題 (6.4) の最適値を $q(\boldsymbol{u})$ と表すと，

$$q(\boldsymbol{u}) = \inf_{\boldsymbol{x} \in X} L(\boldsymbol{x}, \boldsymbol{u}) \tag{6.5}$$

である．$\boldsymbol{u}$ の値によって $q(\boldsymbol{u}) = -\infty$ も取りうる．補題 6.1 より，

$$\inf_{\boldsymbol{x} \in X} F(\boldsymbol{x}) = \inf_{\boldsymbol{x} \in X} \sup_{\boldsymbol{u} \geq \boldsymbol{0}} L(\boldsymbol{x}, \boldsymbol{u}) \geq \sup_{\boldsymbol{u} \geq \boldsymbol{0}} \inf_{\boldsymbol{x} \in X} L(\boldsymbol{x}, \boldsymbol{u}) = \sup_{\boldsymbol{u} \geq \boldsymbol{0}} q(\boldsymbol{u}) \tag{6.6}$$

を得る．この最後の問題を元の最適化問題 (6.1) のラグランジュ双対問題 (Lagrangian dual problem) と呼ぶ．つまり，次の問題である．

$$\begin{aligned} \text{maximize} \quad & q(\boldsymbol{u}) \\ \text{subject to} \quad & \boldsymbol{u} \in \mathbb{R}^m_+ \end{aligned} \tag{6.7}$$

定理 6.1（弱双対定理 (weak duality)） 主問題 (6.1) の最適値を f^*，双対問題 (6.7) の最適値を q^* とすると，次の関係が成立する[*2]．

$$f^* \geq q^*$$

次の最適化問題を例に考える．

$$\begin{aligned} \text{minimize} \quad & f(\boldsymbol{x}) = 2x_1^2 + x_1 x_2 + x_2^2 - 5x_1 - 3x_2 + 4 \\ \text{subject to} \quad & g_1(\boldsymbol{x}) = x_1 + x_2 - 1 \leq 0 \\ & \boldsymbol{x} \in \mathbb{R}^2 \end{aligned}$$

[*1] 任意の $\tilde{\boldsymbol{x}} \in \mathbb{R}^n$ に対して

$$\sup_{\boldsymbol{u} \geq \boldsymbol{0}} L(\tilde{\boldsymbol{x}}, \boldsymbol{u}) \geq \sup_{\boldsymbol{u} \geq \boldsymbol{0}} \inf_{\boldsymbol{x} \in X} L(\boldsymbol{x}, \boldsymbol{u})$$

は明らか．左辺の $\inf_{\tilde{\boldsymbol{x}}}$ をとっても不等号は成立することからこの補題は成立する．

[*2] $f^* = \inf_{\boldsymbol{x} \in X} F(\boldsymbol{x})$ および $q^* = \sup_{\boldsymbol{u} \geq \boldsymbol{0}} q(\boldsymbol{u})$ より．

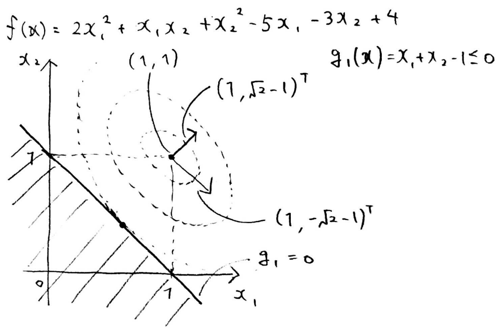

図 **6.1:** 実行可能領域と f の等高線

目的関数および制約式の勾配ベクトルは，

$$
\begin{aligned}
\nabla f(\boldsymbol{x}) &= (4x_1 + x_2 - 5, x_1 + 2x_2 - 3)^{\mathrm{T}} \\
\nabla g_1(\boldsymbol{x}) &= (1, 1)^{\mathrm{T}}
\end{aligned}
$$

である．よって，f の停留点は，$(1, 1)$ である．また，f のヘッセ行列は，

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

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 $\boldsymbol{v}_1 = (1, \sqrt{2} - 1)^{\mathrm{T}}, \boldsymbol{v}_2 = (1, -\sqrt{2} - 1)^{\mathrm{T}}$ である．この問題の実行可能領域および f の等高線を図 6.1 に示す．

$g_1(\boldsymbol{x})$ は有効制約であると考えられる．よって，KKT 条件は

$$
\begin{aligned}
&\begin{pmatrix} 4x_1 + x_2 - 5 \\ x_1 + 2x_2 - 3 \end{pmatrix} + u \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \\
&x_1 + x_2 - 1 = 0
\end{aligned}
$$

より，

$$
(\bar{x}_1, \bar{x}_2, u) = \left(\frac{3}{4}, \frac{1}{4}, \frac{7}{4}\right)
$$

を得る．容易にわかるようにこれは最適解であり，最適値は

$$
f^* = \frac{7}{8}
$$

である．

次に，この問題の双対問題を考える．Lagrange 関数は，

$$
L(\boldsymbol{x}, u) = 2x_1^2 + x_1 x_2 + x_2^2 + (u - 5)x_1 + (u - 3)x_2 + 4 - u
$$

である．u を固定して $q(u)$ を求めるため，停留点条件

$$
\nabla_{\boldsymbol{x}} L(\boldsymbol{x}, u) = \begin{pmatrix} 4x_1 + x_2 + u - 5 \\ x_1 + 2x_2 + u - 3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}
$$

を解くと，

$$x_1 = 1 - \frac{3}{7}u,\ x_2 = 1 - \frac{1}{7}u$$

を得る．この結果を $L(\boldsymbol{x}, u)$ に代入して整理すると，

$$q(u) = -\frac{2}{7}u^2 + u = -\frac{2}{7}\left(u - \frac{7}{4}\right)^2 + \frac{7}{8}$$

となる．したがって，双対問題は

$$\begin{aligned} &\text{maximize} \quad -\frac{2}{7}u^2 + u \\ &\text{subject to} \quad u \in \mathbb{R}_+ \end{aligned}$$

と書ける．この問題の最適解は $u = 7/4$ であり，最適値は

$$q^* = \frac{7}{8}$$

である．以上をまとめ，$f^* = q^*$ の成立が判明した． □

■鞍点と最適性　Lagrange 関数 $L(\boldsymbol{x}, \boldsymbol{u})$ に対し，次の条件を満たす $\boldsymbol{x}^* \in X$, $\boldsymbol{u}^* \in \mathbb{R}^m_+$ が存在するならば，$(\boldsymbol{x}^*, \boldsymbol{u}^*)$ を L の鞍点 (saddle point) という．

$$L(\boldsymbol{x}, \boldsymbol{u}^*) \geq L(\boldsymbol{x}^*, \boldsymbol{u}^*) \geq L(\boldsymbol{x}^*, \boldsymbol{u}),\ \forall \boldsymbol{x} \in X,\ \forall \boldsymbol{u} \in \mathbb{R}^m_+ \tag{6.8}$$

系 2　主問題 (6.1) と双対問題 (6.7) に対し，以下の条件はそれぞれ $\boldsymbol{x}^* \in X$ が主問題の，$\boldsymbol{u}^*$ が双対問題の最適解であるための十分条件である．

1. $\boldsymbol{x}^*$ は主問題の実行可能解，$\boldsymbol{u}^*$ は双対問題の実行可能解であり，さらに $f(\boldsymbol{x}^*) = q(\boldsymbol{u}^*)$ を満たす．
2. $(\boldsymbol{x}^*, \boldsymbol{u}^*)$ は Lagrange 関数 $L(\boldsymbol{x}, \boldsymbol{u})$ の鞍点である．

■等式制約がある場合の双対問題

$$\begin{aligned} \text{minimize} \quad & f(\boldsymbol{x}) \\ \text{subject to} \quad & g_i(\boldsymbol{x}) \leq 0,\ i = 1, 2, \ldots, l \\ & g_i(\boldsymbol{x}) = 0,\ i = l+1, \ldots, m \\ & \boldsymbol{x} \in X \end{aligned} \tag{6.9}$$

等式条件は，

$$g_i(\boldsymbol{x}) \leq 0,\ -g_i(\boldsymbol{x}) \leq 0,\ i = l+1, \ldots, m$$

と不等式の対で表し，対応するラグランジュ乗数をそれぞれ u_i^+, u_i^- と記すと，L この問題の Lagrange 関数は

$$L(\boldsymbol{x}, \boldsymbol{u}) = f(\boldsymbol{x}) + \sum_{i=1}^{l} u_i g_i(\boldsymbol{x}) + \sum_{i=l+1}^{m} (u_i^+ - u_i^-) g_i(x) \tag{6.10}$$

である．ただし，$u_i \geq 0, i = 1, 2, \ldots, l,\ u_i^+, u_i^- \geq 0, i = l+1, \ldots, m$ であるが，

$$u_i = u_i^+ - u_i^-,\ i = l+1, \ldots, m \tag{6.11}$$

とおくと，u_i は正負どちらの値もとることができて符号制約はなくなる．この u_i を用いると，LL この問題の Lagrange 関数は

$$L(\boldsymbol{x}, \boldsymbol{u}) = f(\boldsymbol{x}) + \sum_{i=1}^{m} u_i g_i(\boldsymbol{x})$$

ただし，

$$\begin{aligned} & u_i \geq 0,\ i = 1, 2, \ldots, l \\ & u_i\ :\ \text{符号制限なし},\ i = l+1, \ldots, m \end{aligned} \tag{6.12}$$

となる．つまり，等式条件に対するラグランジュ乗数の符号制限がなくなる点だけが異なる．

この $L(\boldsymbol{x}, \boldsymbol{u})$ から双対問題に至るまでは，不等式制約のみの議論をほぼそのまま適用することができる．得られる双対問題は，

$$\begin{aligned} \text{maximize} \quad & q(\boldsymbol{u}) \\ \text{subject to} \quad & \boldsymbol{u} \in \mathbb{R}^m \\ & u_i \geq 0,\ i = 1, 2, \ldots, l \end{aligned} \tag{6.13}$$

である．弱双対定理 6.1 もそのまま成立する．

6.2 凸計画問題の双対定理

主問題として

$$\begin{aligned} \text{minimize} \quad & f(\boldsymbol{x}) \\ \text{subject to} \quad & g_i(\boldsymbol{x}) \leq 0,\ i = 1, 2, \ldots, m \\ & \boldsymbol{x} \in \mathbb{R}^n \end{aligned} \tag{6.14}$$

を考える．ただし，$f(\boldsymbol{x})$ および $g_i(\boldsymbol{x}), i = 1, 2, \ldots, m$ はすべて凸関数である．さらに，実行可能領域は内点 $\boldsymbol{x}' \in \mathbb{R}^n$ を持つ (Slater の制約想定).

$$g_i(\boldsymbol{x}') < 0,\ i = 1, 2, \ldots, m \tag{6.15}$$

このとき，次の定理が成立する．

定理 6.2（凸計画問題の双対定理） 凸計画問題 (6.14) が Slater の制約想定 (6.15) を満たすと仮定し，最適値を f^*，その双対問題 (6.7) の最適値を q^* と記す．このとき，$f^* = q^*$ が成立する．

第 7 章

証明

この章には，補題や定理の証明 (一部) を記載する．

7.1　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．$\forall \boldsymbol{x}^1, \boldsymbol{x}^2 \in X_1 \cap X_2$ 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 7.1，補題 7.1 を用いる．

定理 7.1（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 7.1　定理 7.1 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，

$$(\boldsymbol{u} - \boldsymbol{v})^{\mathrm{T}}(\boldsymbol{w} - \boldsymbol{v}) \leq 0 \tag{7.1}$$

が成立する．

証明 2（補題 7.1 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，

$$\|\boldsymbol{v} - \boldsymbol{u}\|^2 \leq \|((1-\lambda)\boldsymbol{v} + \lambda\boldsymbol{w}) - \boldsymbol{u}\|^2 \tag{7.2}$$

[*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．

(7.2) 式を整理すると，

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

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

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

よって，$\lambda \to 0$ とすると，(7.1) 式を得る． □

証明 3（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 7.1 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，

$$\|\bar{\boldsymbol{y}}-\boldsymbol{y}\|^2 = (\bar{\boldsymbol{y}}-\boldsymbol{y})^{\mathrm{T}}(\bar{\boldsymbol{y}}-\boldsymbol{y}) > 0 \tag{7.3}$$

である．また，補題 7.1 より，

$$(\boldsymbol{y}-\bar{\boldsymbol{y}})^{\mathrm{T}}(\boldsymbol{x}-\bar{\boldsymbol{y}}) \le 0,\ x \in X \tag{7.4}$$

が成立する．(7.3) 式，(7.4) 式より，

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

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 4（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，

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

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

証明 5（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，

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

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

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

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．

7.2　4 章の証明

証明 6（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 7（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 8（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，

$$\frac{1}{2} \boldsymbol{d}^{\mathrm{T}} \nabla^2 f(\bar{\boldsymbol{x}}) \boldsymbol{d} + \frac{o(\varepsilon^2 \|\boldsymbol{d}\|^2)}{\varepsilon^2} \tag{7.5}$$

$H(\bar{\boldsymbol{x}})$ 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，(7.5) 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．　□

証明 9（定理 4.4 の証明）　(1 次の必要条件)

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

$$\begin{aligned} F^k(\boldsymbol{x}) &= f(\boldsymbol{x}) + \frac{k}{2} \|g(\boldsymbol{x})\|^2 + \frac{\alpha}{2} \|\boldsymbol{x} - \bar{\boldsymbol{x}}\|^2 \\ g(\boldsymbol{x}) &= (g_1(\boldsymbol{x}), \ldots, g_m(\boldsymbol{x}))^{\mathrm{T}} \end{aligned}$$

より，

$$\nabla F^k(\boldsymbol{x}^k) = \nabla f(\boldsymbol{x}^k) + k \nabla g(\boldsymbol{x}^k)^{\mathrm{T}} g(\boldsymbol{x}^k) + \alpha(\boldsymbol{x} - \bar{\boldsymbol{x}}) = 0 \tag{7.6}$$

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

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

が成立する．よって，m 次正方行列 $\nabla g(\boldsymbol{x}^k)\nabla g(\boldsymbol{x}^k)^{\mathrm{T}}$ も正則である．(7.6) の両辺に左から $(\nabla g(\boldsymbol{x}^k)\nabla g(\boldsymbol{x}^k)^{\mathrm{T}})\nabla g(\boldsymbol{x}^k)$ を乗じて整理すると，

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

を得る．$k \to \infty$ とすると，$\{\boldsymbol{x}^k\}$ は $\bar{\boldsymbol{x}}$ に収束するので，点列 $\{kg(\boldsymbol{x}^k)\}$ は，m 次元ベクトル

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

に収束する．(7.6) において，$k \to \infty$ とすると，

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

を得る． □

(2 次の必要条件)

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

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

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

$$\boldsymbol{y}^k = \boldsymbol{y} - \nabla g(\boldsymbol{x}^k)^{\mathrm{T}}(\nabla g(\boldsymbol{x}^k)\nabla g(\boldsymbol{x}^k)^{\mathrm{T}})^{-1}\nabla g(\boldsymbol{x}^k)\boldsymbol{y} \tag{7.7}$$

とする．$\boldsymbol{y}^k$ は，

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

を満たす[*3]．$k \to \infty$ とすると，$\boldsymbol{x}^k \to \bar{\boldsymbol{x}}$ であり，$\boldsymbol{y}$ の定義より，$\nabla g(\bar{\boldsymbol{x}}) = \boldsymbol{0}$ である．よって，(7.7) の第二項は $\boldsymbol{0}$ に収束し，$\boldsymbol{y}^k \to \boldsymbol{y}$ を得る．

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

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

を得る．$k \to \infty$ とすると，

$$\boldsymbol{y}^k \to \boldsymbol{y},\ \boldsymbol{x}^k \to \bar{\boldsymbol{x}},\ kg_i(\boldsymbol{x}^k) \to u_i$$

となる．よって，

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

となる．$\alpha > 0$ は任意より，$\alpha \to 0$ として，(4.10) を得る． □

[*3] $\boldsymbol{y}^k$ は y の空間 $V(\boldsymbol{x}^k)$ への射影点である．

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

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

である．これは，$\boldsymbol{x}'$ 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 $\bar{\boldsymbol{x}}, \bar{\bar{\boldsymbol{x}}}$ が存在することを仮定する．このとき，$f(\bar{\boldsymbol{x}}) = f(\bar{\bar{\boldsymbol{x}}})$ である．$\bar{\boldsymbol{x}}$ と $\bar{\bar{\boldsymbol{x}}}$ を結ぶ線分上の点 $\boldsymbol{x}_\alpha = (1-\alpha)\bar{\boldsymbol{x}} + \alpha\bar{\bar{\boldsymbol{x}}}, 0 < \alpha < 1$ を考えると，実行可能領域の凸性より，$\boldsymbol{x}_\alpha$ も実行可能解．f の狭義凸性より，

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

これは，$\bar{\boldsymbol{x}}$ 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.