最適化

酒匂貴市

平成 20 年 1 月 3 日

目 次

第 I 部
基礎事項

1　最適化問題の基礎事項

1.1　最適化問題と用語

定義 1.1 X を $\boldsymbol{R^n}$ などの適当な集合とし、その部分集合を S、f を X を定義域とする実数値関数とする。ある $x^* \in S$ が、$\forall x \in S$ について $f(x^*) \leq f(x)$ を満たすとき、$f(x^*)$ を最小値、x^* を最小解という。また、ある $x^* \in S$ が、$\forall x \in S$ について $f(x^*) \geq f(x)$ を満たすとき、$f(x^*)$ を最大値、x^* を最大解という。このとき S, f について最小解もしくは最大解を求める問題を**最適化問題**といい

$$\left|\begin{array}{ll} \textit{minimize} & f(x) \\ \textit{s.t.} & x \in S \end{array}\right. \qquad \text{[P1]}$$

$$\left|\begin{array}{ll} \textit{maximize} & f(x) \\ \textit{s.t.} & x \in S \end{array}\right.$$

などと表す[1]。最適化問題に対して、求めたい最小解もしくは最大解を**最適解・大局的最適解**という。このとき、関数 f を**目的関数**、集合 S を**許容領域・実行可能領域・許容集合**といい、S の元を**許容解・実行可能解**という。また、許容集合を定める条件を**制約**といい、これが式で表されるとき**制約式**とも呼ぶ。

最大化問題は目的関数を -1 倍することにより、最小化問題となるので、以後一般的な視点では、**最小化問題のみを扱うものとする**。

定義 1.2 （最適化問題の解のパターン）[P1] において、許容解が存在するとき、最適化問題は**実行可能**であるといい、許容解が存在しないとき、つまり $S = \phi$ であるとき、最適化問題は**実行不可能**であるという。また、許容集合の目的関数による像 $f(S)$ が下に有界であれば、最適化問題は**有界**であるといい、有界でなければ**非有界**であるという。有界であれば、実数の順序完備性（付録命題 *A.1*）から、$f(S)$ には下限 $\inf\limits_{x \in S} f(x)$ が存在する。この下限を**最適値**と呼ぶ。最適値が存在しても、最適値を実現する許容解が存在するとは限らない。最適値を実現する許容解が存在すれば、それは最適解である。

最適化問題のパターンのまとめ

実行不可能		
実行可能	非有界	
実行可能	有界	最適解存在せず
実行可能	有界	最適解が存在

最適解の存在については、次の定理がある。

定理 1.1 $f(S)$ が有界閉集合・コンパクトであるときは、$f(S)$ に最小値が存在し、したがって、最適解が存在する。また、$f(S)$ が下に有界でかつ閉集合であれば、やはり最適解が存在する。さらに、許容集合 S がコンパクトで、f が連続関数ならば、最適解が存在する。

(proof)
定理 C.9 定理 C.10 定理 C.11 より示される。 証明終

[1] s.t. という略語は subject to と such that という二つの意味で用いられるが、最適化問題に関しては subject to の方が意味としてより適切であろう。

1.1.1 無限大の扱い

実際には、関数は $\pm\infty$ をとると考えることも多い。凸解析の分野では、積極的にこの概念を取り入れている。次の定義を置く。

定義 1.3 適当な集合 X を定義域とする関数 $f : X \to \boldsymbol{R^n} \cup \{\pm\infty\}$ について

$$\mathrm{dom}f \equiv \{x \in X; -\infty < f(x) < \infty\}$$

を**実効定義域・実効領域**[2]という。

しかし、一般の最適化問題では、$\pm\infty$ をとらない関数 ($\mathrm{dom}f^c = \phi$) と考えたり、定義域ではないとしてあらかじめ許容解から除いておく (許容集合 $S \subset \mathrm{dom}f$ となる) ことが多い。また、以後、$\mathrm{dom}f = \phi$ であるような極めて特殊な関数は扱わないとする。これは実用的にはまったく問題ない仮定である。

1.2 最適化問題への一般的なアプローチ

1.2.1 最適化問題の分類

一般的には、最適解に到達することは容易ではない。まず、実際の問題に即して、目的関数や制約に制限を加え、クラスに分類し、それぞれに固有の方法を用いることが考えられる。分類されたクラスの例としては

1. 線形計画問題
   $\cdots$ 目的関数が線形関数（一次関数）で、制約が線形関数の等式及び不等式で表されるもの。

2. 非線形計画問題
   $\cdots$ 目的関数が $\boldsymbol{R^n}$ を定義域とする一般の関数で、制約も同様の関数の等式及び不等式で表されるもの。

3. 凸計画問題
   $\cdots$ 目的関数が $\boldsymbol{R^n}$ を定義域とする凸関数で、制約が不等式に対しては凸関数、等式に対しては線形関数の形で表されるもの。

4. 錐線形計画問題
   $\cdots$ 詳細後述

5. 半正定値計画問題
   $\cdots$ 詳細後述

6. 二次錐計画問題
   $\cdots$ 詳細後述

7. 制約なし計画問題
   $\cdots$ 制約がなく、許容集合が目的関数の定義域と一致する ($S = X$) もの。

8. 離散計画問題
   $\cdots$ 離散的な構造を持つ問題の総称。組み合わせ最適化問題とも言う。

9. 整数計画問題
   $\cdots$ 線形計画問題に、整数性の制約を加えたもの。

のようなものがある。これらの間では

非線形計画問題 $\supset$ 凸計画問題 $\supset$ 錐線形計画問題 $\supset$ 線形計画問題・半正定値計画問題・二次錐計画問題

という包含関係が存在しており、また、離散計画問題ではその定義は明確なものではない。

[2]effective domain

1.2.2　最適性条件

また、最適解が満たすべき必要条件により絞り込むことが考えられる。

定義 1.4 最適解であることの必要条件を**最適性条件**という。

対偶を考えることにより、最適性条件を満たさない許容解は最適解ではないといえる。最適性条件の一つとして、局所的最適解がある。

定義 1.5 [P1] で、X を距離空間とし、$x^* \in X$ について、そのある近傍 $V_\epsilon(x^*)$ が存在して

$$\left| \begin{array}{ll} minimize & f(x) \\ s.t. & x \in S \cap V_\epsilon(x^*) \end{array} \right.$$

の最適解が x^* であるとき、x^* を**局所（的）最適解**という。

いわゆる、極値・停留点は、局所的最適解となるものである。大局的最適解は、局所的最適解であるが、逆は一般には成立しない。ただし、前に挙げた凸計画問題では、局所的最適解が大局的最適解であることが保証されている。

1.2.3　アルゴリズム

最適化問題は、実用を念頭に置いた問題であり、解を得ることが最たる目的である。そのために、アルゴリズムによって十分実用的な近似解を得ることは、有力な方法である。

1.2.4　問題の書き換え

制約を別の**同値な制約**に置き換え、問題を書き換えることもよくなされる。この同値な制約で置き換えた場合には、書き換えられた問題は元の問題と最適化問題としてはまったく同じである。また、最適解が満たすべき必要条件を制約として付加することもある。この場合には、許容集合が小さくなるが、最適解は変化せず、また、許容性も変化しないので、実用的には、最適解に焦点があることが多いこともあり、有用な書き換えである。

第 II 部
非線形計画問題

2　非線形計画問題の概要

2.1　非線形計画問題の表現

非線形計画問題とは、すでに述べたように、目的関数が $\boldsymbol{R^n}$ を定義域とする一般の関数で、制約も同様の関数の等式及び不等式で表される最適化問題である。より記号的に表現すれば

$$\left|\begin{array}{lll} \text{minimize} & f(\boldsymbol{x}) & (f : \boldsymbol{R^n} \to \boldsymbol{R}) \\ \text{s.t.} & g_i(\boldsymbol{x}) \geq 0 & (i = 1, \cdots, l) \\ & g_i(\boldsymbol{x}) = 0 & (i = l+1, \cdots, m) \\ & (g_i : \boldsymbol{R^n} \to \boldsymbol{R}) & \end{array}\right. \qquad [\text{P2}]$$

となる。要するに、式の形で表される最適化問題 のことである。

ところで、等式は、不等式 2 個を以て同値に書き換えられる、すなわち $X = Y \leftrightarrow X \leq Y, X \geq Y$ である。こうして等式制約を同値変形すると、すべて不等式制約となって、形式上扱いやすく、この形

$$\left|\begin{array}{lll} \text{minimize} & f(\boldsymbol{x}) & \\ \text{s.t.} & g_i(\boldsymbol{x}) \geq 0 & (i = 1, \cdots, m) \end{array}\right. \qquad [\text{P3}]$$

で非線形計画問題を記述することも多い。この形式を**不等式形**という。また、不等式制約をまとめて $\boldsymbol{g}(\boldsymbol{x}) \geq \boldsymbol{0}$ とも記述する。また

$$\left|\begin{array}{lll} \text{minimize} & f(\boldsymbol{x}) & \\ \text{s.t.} & g_i(\boldsymbol{x}) \geq 0 & (i = 1, \cdots, m) \\ & \boldsymbol{x} \geq \boldsymbol{0} & \end{array}\right. \qquad [\text{P4}]$$

の形を、**不等式標準形**という。不等式形にした任意の非線形最適化問題に対して、各変数を

$$\begin{cases} x_i = x_i^+ - x_i^- \\ x_i^+ \geq 0 \\ x_i^- \geq 0 \end{cases}$$

によって二個の非負制約変数によって表すことによって、実現できる。さらに

$$\left|\begin{array}{lll} \text{minimize} & f(\boldsymbol{x}) & \\ \text{s.t.} & g_i(\boldsymbol{x}) = 0 & (i = 1, \cdots, m) \\ & \boldsymbol{x} \geq \boldsymbol{0} & \end{array}\right. \qquad [\text{P5}]$$

の形を、**等式標準形**という。不等式標準形から、各不等式に対しては、新たな変数**（スラック変数）**x_i^s を導入して

$$\begin{cases} g_i(\boldsymbol{x}) - x_i^s = 0 \\ x_i^s \geq 0 \end{cases}$$

と書き換えることにより、任意の非線形計画問題を等式標準形で記述することができる。以上で述べたことを、定理の形でまとめておく。

定理 2.1 任意の非線形計画問題は不等式形・不等式標準形・等式標準形に書き換えられる。

2.2　共通基礎事項

定義 2.1　（ラグランジュ関数）非線形計画問題 [P3] において、目的関数 f の定義域の部分集合を V とし、$\mathbf{0} \leq \boldsymbol{\lambda} = (\lambda_1, \cdots, \lambda_m) \in \boldsymbol{R^m}$ とするとき

$$\begin{aligned} L(\boldsymbol{x}, \boldsymbol{\lambda}) &\equiv f(\boldsymbol{x}) - \lambda_1 g_1(\boldsymbol{x}) - \cdots - \lambda_m g_m(\boldsymbol{x}) \\ &= f(\boldsymbol{x}) - \boldsymbol{\lambda}^T \boldsymbol{g}(\boldsymbol{x}) \end{aligned}$$

である関数 $L_V(\boldsymbol{x}, \boldsymbol{\lambda}) : V \times \boldsymbol{R^m} \to \boldsymbol{R}$ を**ラグランジュ関数**といい、ベクトル $\boldsymbol{\lambda}$ を**（一般化）ラグランジュ乗数（ラグランジュの未定乗数）**という。$V = \boldsymbol{R^n}$ の場合は、単純に $L(\boldsymbol{x}, \boldsymbol{\lambda})$ と表す。

2.2.1　多変数解析

証明はせずに、列挙にとどめる。

定理 2.2　（多変数の平均値の定理）$f : \boldsymbol{R^n} \to \boldsymbol{R}$ が $[\boldsymbol{x}, \boldsymbol{x}'] \subset \boldsymbol{R^n}$ で微分可能ならば

$$0 \leq \exists\theta \leq 1, \quad f(\boldsymbol{x}') - f(\boldsymbol{x}) = (\boldsymbol{x}' - \boldsymbol{x})^T \nabla f(\theta\boldsymbol{x}' + (1-\theta)\boldsymbol{x})$$

となる。

定義 2.2 二回微分可能な関数 $f : \boldsymbol{R^n} \to \boldsymbol{R}$ に対して

$$\nabla^2 f \equiv \begin{pmatrix} \frac{\partial^2 f}{\partial x_1 \partial x_1} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_n} \end{pmatrix}$$

を**ヘッセ行列**という。

定理 2.3　（2次の多変数テイラー展開）$f : \boldsymbol{R^n} \to \boldsymbol{R}$ が $[\boldsymbol{x}, \boldsymbol{x}'] \subset \boldsymbol{R^n}$ で2回微分可能ならば

$$0 \leq \exists\theta \leq 1, \quad f(\boldsymbol{x}') - f(\boldsymbol{x}) = (\boldsymbol{x}' - \boldsymbol{x})^T \nabla f(\boldsymbol{x}) + \frac{1}{2}(\boldsymbol{x}' - \boldsymbol{x})^T \nabla^2 f(\theta\boldsymbol{x}' + (1-\theta)\boldsymbol{x})(\boldsymbol{x}' - \boldsymbol{x})$$

となる。

2.2.2　凸集合

定義 2.3 線形空間の部分空間 F と $\forall \boldsymbol{x}, \boldsymbol{y} \in F, 0 \leq \forall t \leq 1$ について

$$t\boldsymbol{x} + (1-t)\boldsymbol{y} \in F$$

が成り立つとき、F を**凸集合**という。

補題 2.4 A, B が凸集合ならば、$A \cap B$ も凸集合である。

(proof)
$\forall \boldsymbol{x}, \boldsymbol{y} \in A \cap B, 0 \leq \forall t \leq 1$ について

$$t\boldsymbol{x} + (1-t)\boldsymbol{y} \in A \text{ かつ } t\boldsymbol{x} + (1-t)\boldsymbol{y} \in B$$

であるから、つまり

$$t\boldsymbol{x} + (1-t)\boldsymbol{y} \in A \cap B$$

である。 $\boxed{\text{証明終}}$

定理 2.5 （分離定理）閉凸集合 $F \subset \boldsymbol{R^n}$ と $z \notin F$ なる点 $\boldsymbol{z}$ について

$$\forall x \in F, \quad \boldsymbol{a}^T\boldsymbol{z} < \Theta < \boldsymbol{a}^T\boldsymbol{x} \quad (\boldsymbol{a}^T\boldsymbol{z} - \Theta < 0 < \boldsymbol{a}^T\boldsymbol{x} - \Theta)$$

なる $\boldsymbol{a} \in \boldsymbol{R^n}, \Theta \in \boldsymbol{R}$ が存在する[3]。

(proof)

$F = \phi$ である場合は、自明である。$F \neq \phi$ であるとき、適当な $\boldsymbol{f} \in F$ をとり、$L \equiv \|\boldsymbol{f} - \boldsymbol{z}\|$ として

$$G = F \cap \overline{V}_L(\boldsymbol{z})$$

とおく。$\overline{V}_L(\boldsymbol{z})$ は明らかに有界なので、G も有界である。また、補題 B.13 より $\overline{V}_L(\boldsymbol{z})$ は閉集合であり、F も閉集合なので、定理 B.1 より G も閉集合である。すなわち、G は有界閉集合であるので、定理 C.6 より G はコンパクトである。

$$f_z(\boldsymbol{x}) \equiv \|\boldsymbol{x} - \boldsymbol{z}\|$$

とおくと、定理 B.18 より距離関数は連続なので、定理 C.11 より $f_z(G)$ には最小値が存在する。このとき、f_z の最小値を与える $\boldsymbol{g} \in G$ が少なくともひとつは存在する。また、$F \backslash G$ の元から z への距離は $\|\boldsymbol{f} - \boldsymbol{z}\| \geq \|\boldsymbol{g} - \boldsymbol{z}\|$ よりも大きいので、$\boldsymbol{g}$ は F の f_z を最小にする元でもある。これを用いて

$$\boldsymbol{a} \equiv \boldsymbol{g} - \boldsymbol{z}$$

とおく。

このとき $\forall \boldsymbol{x} \in F, 0 \leq \forall t < 1$ について、$f_z(t\boldsymbol{g} + (1-t)\boldsymbol{x}) \geq f_z(\boldsymbol{g})$ であるから

$$\begin{aligned}
\|t\boldsymbol{g} + (1-t)\boldsymbol{x} - \boldsymbol{z}\|^2 &\geq \|\boldsymbol{g} - \boldsymbol{z}\|^2 = \|\boldsymbol{a}\|^2 \\
\|(1-t)(\boldsymbol{x} - \boldsymbol{g}) + \boldsymbol{a}\|^2 &\geq \|\boldsymbol{a}\|^2 \\
(1-t)^2\|\boldsymbol{x} - \boldsymbol{g}\|^2 + 2(1-t)\boldsymbol{a}^T(\boldsymbol{x} - \boldsymbol{g}) + \|\boldsymbol{a}\|^2 &\geq \|\boldsymbol{a}\|^2 \\
(1-t)\|\boldsymbol{x} - \boldsymbol{g}\|^2 + 2\boldsymbol{a}^T(\boldsymbol{x} - \boldsymbol{g}) &\geq 0 \quad \because \text{ここでは} t < 1 \\
2\boldsymbol{a}^T(\boldsymbol{x} - \boldsymbol{g}) + 2\|\boldsymbol{a}\|^2 &\geq 2\|\boldsymbol{a}\|^2 - (1-t)\|\boldsymbol{x} - \boldsymbol{g}\|^2
\end{aligned}$$

であるが、$\|\boldsymbol{a}\| > 0$ であり、かつ $1 - t > 0$ をいくらでも小さくできるので、$2\|\boldsymbol{a}\|^2 - (1-t)\|\boldsymbol{x} - \boldsymbol{g}\|^2 > 0$ となるよう $0 < \exists t < 1$ をとれる。このとき

$$\begin{aligned}
2\boldsymbol{a}^T(\boldsymbol{x} - \boldsymbol{g}) + 2\|\boldsymbol{a}\|^2 \geq 2\|\boldsymbol{a}\|^2 - (1-t)\|\boldsymbol{x} - \boldsymbol{g}\|^2 &> 0 \\
2\boldsymbol{a}^T(\boldsymbol{x} - \boldsymbol{g}) + 2\boldsymbol{a}^T(\boldsymbol{g} - \boldsymbol{z}) &> 0 \\
\boldsymbol{a}^T(\boldsymbol{x} - \boldsymbol{z}) &> 0 \\
\boldsymbol{a}^T\boldsymbol{x} &> \boldsymbol{a}^T\boldsymbol{z}
\end{aligned}$$

である。このとき、例えば $\Theta \equiv \dfrac{\boldsymbol{a}^T\boldsymbol{x} + \boldsymbol{a}^T\boldsymbol{z}}{2}$ とすれば

$$\forall x \in F, \quad \boldsymbol{a}^T\boldsymbol{z} < \Theta < \boldsymbol{a}^T\boldsymbol{x}$$

である。 証明終

2.2.3 錐

定義 2.4 線形空間の部分空間 K について

$$\forall \boldsymbol{x} \in K, \forall \lambda > 0 \quad \lambda\boldsymbol{x} \in K$$

が成り立つとき、K は**錐**であるという。凸集合である錐を凸錐、さらに閉集合である錐を閉凸錐という。

[3]超平面 $\{\boldsymbol{x} \in \boldsymbol{R^n}; \boldsymbol{a}^T\boldsymbol{x} - \Theta = 0\}$ によって、F と z が分離されるということを表している。

定義 2.5 計量線形空間 V の錐 K に対し

$$K^* \equiv \{\boldsymbol{v} \in V; \forall \boldsymbol{u} \in K, (\boldsymbol{u}, \boldsymbol{v}) \geq 0\}$$

を、K の**双対錐**という。さらに $K = K^*$ となるならば、K は**自己双対錐**であるという。

(例)

例えば、線形空間は錐であり、特に線形部分空間も錐である。線形部分空間については、容易にその双対錐が直交補空間であることがわかる。また、非負制約 $\{\boldsymbol{x} \in \boldsymbol{R^n}; \boldsymbol{x} \geq \boldsymbol{0}\}$ は閉凸錐である。さらに、非負制約は自己双対錐でもある。(次の補題参照)

補題 2.6 $K \equiv \{\boldsymbol{x} \in \boldsymbol{R^n}; \boldsymbol{x} \geq \boldsymbol{0}\}$(非負制約)とすると $K^* = K$ である。

(proof)

$\boldsymbol{x} \in K^*$ ならば、$\forall \boldsymbol{u} \in K \leftrightarrow \boldsymbol{u} \geq \boldsymbol{0}$ に対して $\boldsymbol{u}^T \boldsymbol{x} \geq 0$ である。よって、特に $\boldsymbol{u} = \boldsymbol{e}_i = (0, \cdots, 1_{(i\ 要素)}, \cdots, 0)^T$ とすれば

$$\boldsymbol{u}^T \boldsymbol{x} = x_i \geq 0$$

が必要である。逆に $\boldsymbol{x} \geq \boldsymbol{0}$ ならば、$\forall \boldsymbol{u} \in K \leftrightarrow \boldsymbol{u} \geq \boldsymbol{0}$ に対して $\boldsymbol{u}^T \boldsymbol{x} \geq 0$ である。よって

$$K^* = \{\boldsymbol{x} \in \boldsymbol{R^n}; \boldsymbol{x} \geq \boldsymbol{0}\} = K$$

である。 $\boxed{\text{証明終}}$

定理 2.7 (錐の分離定理)閉凸錐 $K \subset \boldsymbol{R^n}$ と $z \notin K$ なる点 $\boldsymbol{z}$ について

$$\forall \boldsymbol{x} \in K, \quad \boldsymbol{a}^T \boldsymbol{z} < 0 \leq \boldsymbol{a}^T \boldsymbol{x}$$

なる $\boldsymbol{a} \in \boldsymbol{R^n}$ が存在する。

(proof)

$K \subset \boldsymbol{R^n}$ は閉凸集合なので、定理 2.5 より

$$\forall x \in F, \quad \boldsymbol{a}^T \boldsymbol{z} < \Theta < \boldsymbol{a}^T \boldsymbol{x}$$

なる $\boldsymbol{a} \in \boldsymbol{R^n}, \Theta \in \boldsymbol{R}$ が存在する。$\forall \lambda > 0$ に対して $\lambda \boldsymbol{a} \in F$ なので

$$\boldsymbol{a}^T \boldsymbol{z} < \Theta < \lambda(\boldsymbol{a}^T \boldsymbol{x})$$

である。$\boldsymbol{a}^T \boldsymbol{x} < 0$ と仮定すると、$\lambda > 0$ は任意なので、十分大きく取れば

$$\lambda(\boldsymbol{a}^T \boldsymbol{x}) < \Theta$$

とできて矛盾する。よって $0 \leq \boldsymbol{a}^T \boldsymbol{x}$ である。また、$\Theta > 0$ と仮定すると、$\lambda > 0$ は任意なので、十分小さく取れば

$$\Theta > \lambda(\boldsymbol{a}^T \boldsymbol{x})$$

とできて矛盾する。よって $\Theta \leq 0$ であるから、以上をまとめると

$$\boldsymbol{a}^T \boldsymbol{z} < \Theta \leq 0 \leq \boldsymbol{a}^T \boldsymbol{x}$$

となる。 $\boxed{\text{証明終}}$

定理 2.8 錐 K について $K \subset (K^*)^*$

(proof)
$\forall \boldsymbol{x} \in K$ について、$\forall \boldsymbol{u} \in K^*$ に対して $(\boldsymbol{x}, \boldsymbol{u}) \geq 0$ であり

$$(K^*)^* \equiv \{\boldsymbol{v} \in V; \forall \boldsymbol{u} \in K^*, (\boldsymbol{u}, \boldsymbol{v}) \geq 0\}$$

であるから、$\boldsymbol{x} \in (K^*)^*$ である。よって示された。 証明終

定理 2.9 閉凸錐 $K \subset \boldsymbol{R^n}$ について $(K^*)^* = K$

(proof)
$\forall \boldsymbol{y} \notin K$ について、錐の分離定理 2.7 より

$$\forall \boldsymbol{x} \in K, \quad \boldsymbol{a}^T \boldsymbol{y} < 0 \leq \boldsymbol{a}^T \boldsymbol{x}$$

となる $\boldsymbol{a} \in \boldsymbol{R^n}$ が存在する。このとき $\boldsymbol{a} \in K^*$ であり、したがって、$\boldsymbol{a}^T \boldsymbol{y} < 0$ より $\boldsymbol{y} \notin (K^*)^*$ である。よって、対偶を考えることにより $(K^*)^* \supset K$ であり、上の定理より $(K^*)^* \subset K$ なので、あわせて示される。
証明終

補題 2.10 錐 $K_1 \subset K_2$ について $K_1^* \supset K_2^*$

(proof)
$\boldsymbol{x} \in K_2^*$ ならば

$$\begin{aligned} \forall \boldsymbol{u} \in K_2, \quad (\boldsymbol{x}, \boldsymbol{u}) \geq 0 \\ \rightarrow \forall \boldsymbol{u} \in K_1 \subset K_2, \quad (\boldsymbol{x}, \boldsymbol{u}) \geq 0 \end{aligned}$$

つまり、$\boldsymbol{u} \in K_1^*$ である。よって示された。 証明終

2.2.4 Farkas の補題と閉凸多面錐

定義 2.6 V を n 次元計量線形空間とし、錐 $K \subset V$、$\boldsymbol{a}_1, \cdots, \boldsymbol{a}_m \in V$ とするとき

$$C[\boldsymbol{a}_1, \cdots, \boldsymbol{a}_m; K] \equiv \left\{ \begin{pmatrix} (\boldsymbol{a}_1, \boldsymbol{x}) \\ \vdots \\ (\boldsymbol{a}_m, \boldsymbol{x}) \end{pmatrix} \in \boldsymbol{R^m}; \boldsymbol{x} \in K \right\}$$

を、ベクトル $\boldsymbol{a}_1, \cdots, \boldsymbol{a}_m$ に対する錐 K の**像**という。像は明らかに錐である。同様に、K が凸錐であれば、その像も凸錐である。

定義 2.7 $m \times n$ 行列 $A = (\boldsymbol{a}_1, \cdots, \boldsymbol{a}_m)^T$ について、K を閉凸錐である非負制約とした像

$$C[A] \equiv C[\boldsymbol{a}_1, \cdots, \boldsymbol{a}_m; K(= \text{非負制約})] = \{A\boldsymbol{x}; \boldsymbol{x} \geq \boldsymbol{0}\}$$

を**有限生成錐**という。非負制約が凸錐なので、有限生成錐も凸錐である。

定義 2.8 $m \times n$ 行列 $A = (\boldsymbol{a}_1, \cdots, \boldsymbol{a}_m)^T$ について

$$K(A) \equiv \{\boldsymbol{x} \in \boldsymbol{R^n}; A\boldsymbol{x} \leq \boldsymbol{0}\}$$

を**凸多面錐**という。行列の線形性より、凸多面錐は明らかに凸錐である。

定理 2.11 有限生成錐は閉凸錐である。

(proof)

すでに凸錐であることは分かっているので、閉集合であることを示せばよい。まず、ベクトル $\boldsymbol{a}^T = (a_1, \cdots, a_n)$ に対する凸多面錐 $C[\boldsymbol{a}^T]$ について考える。

(i)$\boldsymbol{a} = \boldsymbol{0}$ のとき。$C[\boldsymbol{a}^T] = \{0\} = [0, 0]$ なので、補題 C.8 より閉区間は閉集合なので、$C[\boldsymbol{a}^T]$ は閉集合であり、$C[\boldsymbol{a}^T]^c$ は開集合である。

(ii)$a_i = 0$ なる成分が $\boldsymbol{a} \neq \boldsymbol{0}$ にあるとき。i 成分は $C[\boldsymbol{a}^T]$ の範囲に影響しないので、$a_i = 0$ なる成分を除いた、全成分が 0 でないベクトル $\boldsymbol{a}'^T$ をとると $C[\boldsymbol{a}^T] = C[\boldsymbol{a}'^T]$ であるので、$C[\boldsymbol{a}'^T]$ で考えればよく、無視してよい。

(iii)$\|\boldsymbol{a}\| \neq 0$ で符号の異なる成分が存在するとき。$a_i > 0, a_j < 0$ とする。このとき $\{a_i x_i; x_i \geq 0\} = [0, \infty), \{a_j x_j; x_j \geq 0\} = (-\infty, 0]$ であるから $C[\boldsymbol{a}^T] \supset \boldsymbol{R}$ つまり $C[\boldsymbol{a}^T] = \boldsymbol{R}$ であり、$C[\boldsymbol{a}^T]^c = \phi$ は開集合である。

(iv)$\boldsymbol{a} > \boldsymbol{0} \quad or \quad \boldsymbol{a} < \boldsymbol{0}$ であるとき。$C[\boldsymbol{a}^T]^c = \{\boldsymbol{a}^T\boldsymbol{x}; \boldsymbol{x} \not\geq \boldsymbol{0}\} = \{\boldsymbol{a}^T\boldsymbol{x}; \exists x_j < 0\}$ である。$\forall y \in C[\boldsymbol{a}^T]^c$ を考えると $y = \boldsymbol{a}^T\boldsymbol{x}, x_j < 0$ である。このとき $0 < \forall c < |x_j| = -x_j$ をとると $x_j + c < 0, x_j - c < 0$ であるから

$$y \pm a_j c = \boldsymbol{a}^T(\boldsymbol{x} \pm c\boldsymbol{e}_j) \in C[\boldsymbol{a}^T]^c$$

である。よって

$$0 < \frac{\epsilon}{|a_j|} < |x_j|$$

なる $\epsilon > 0$ をとることができて

$$V_\epsilon(y) = (y - \epsilon, y + \epsilon) \subset C[\boldsymbol{a}^T]^c$$

となる。つまり $C[\boldsymbol{a}^T]^c$ は開集合である。

(i)(ii)(iii)(iv) より、$C[\boldsymbol{a}^T]^c = \phi$ は $\forall \boldsymbol{a} \in \boldsymbol{R^n}$ に対して開集合である。さて、$m \times n$ 行列 $A = (\boldsymbol{a}_1, \cdots, \boldsymbol{a}_m)^T$ について

$$\begin{aligned} C[A]^c &= \{A\boldsymbol{x}; \boldsymbol{x} \not\geq \boldsymbol{0}\} = \{\boldsymbol{a}_1^T\boldsymbol{x}; \boldsymbol{x} \not\geq \boldsymbol{0}\} \times \cdots \times \{\boldsymbol{a}_m^T\boldsymbol{x}; \boldsymbol{x} \not\geq \boldsymbol{0}\} \\ &= C[\boldsymbol{a}_1]^c \times C[\boldsymbol{a}_m]^c \end{aligned}$$

である[4]。よって定理 C.1 より、$C[A]^c$ は開集合であり、$C[A]$ は閉集合である。 証明終

定理 2.12 凸多面錐は閉凸錐である。

(proof)

すでに凸錐であることは分かっているので、閉集合であることを示せばよい。まず、ベクトル $\boldsymbol{a}^T = (a_1, \cdots, a_n)$ に対する凸多面錐 $K[\boldsymbol{a}^T]$ について考える。$\boldsymbol{a} = \boldsymbol{0}$ である場合は $K[\boldsymbol{a}^T] = \boldsymbol{R^n}$ であり、閉集合である。

$\boldsymbol{a} \neq \boldsymbol{0}$ である場合。$\forall \boldsymbol{y} \in K[\boldsymbol{a}^T]^c$ を考えると $\boldsymbol{a}^T\boldsymbol{y} > 0$ となる。$\epsilon > 0$ と $\forall \boldsymbol{b} \in V_\epsilon(\boldsymbol{y})$ について

$$\boldsymbol{b} = \boldsymbol{y} + \boldsymbol{e} \quad (\|\boldsymbol{e}\| < \epsilon)$$

[4]各成分に共通の $\boldsymbol{x}$ に対する条件が課されている集合だからこそ成立したのであり、一般には直積の補集合は補集合の直積とは一致しない。

である。これについて

$$\begin{aligned}\boldsymbol{a}^T\boldsymbol{e} &= \sum_{i=1}^{n} a_i e_i \\ &\geq -\sum_{i=1}^{n}|a_i e_i| = -\sum_{i=1}^{n}|a_i||e_i| \\ &> -\epsilon\left(\sum_{i=1}^{n}|a_i|\right) \quad \because |e_i| < \epsilon \\ \therefore \boldsymbol{a}^T\boldsymbol{b} = \boldsymbol{a}^T\boldsymbol{y} + \boldsymbol{a}^T\boldsymbol{e} &> \boldsymbol{a}^T\boldsymbol{y} - \epsilon\left(\sum_{i=1}^{n}|a_i|\right)\end{aligned}$$

であるが、$\boldsymbol{a} \neq \boldsymbol{0}$ であれば $\sum_{i=1}^{n}|a_i| > 0$ であり、また $\boldsymbol{a}^T\boldsymbol{y} > 0$ なので

$$\frac{\boldsymbol{a}^T\boldsymbol{y}}{\left(\sum_{i=1}^{n}|a_i|\right)} \geq \epsilon > 0$$

なる ϵ をとることができ、このとき

$$\boldsymbol{a}^T\boldsymbol{b} > \boldsymbol{a}^T\boldsymbol{y} - \epsilon\left(\sum_{i=1}^{n}|a_i|\right) \geq 0$$

より、$\boldsymbol{b} \in K[\boldsymbol{a}^T]^c$ である。つまり、$K[\boldsymbol{a}^T]^c$ は開集合であり $K[\boldsymbol{a}^T]$ は閉集合である。

ここで、$A = (\boldsymbol{a}_1, \cdots, \boldsymbol{a}_m)^T$ として

$$K(A) = \{\boldsymbol{x} \in \boldsymbol{R^n}; \boldsymbol{a}_1^T\boldsymbol{x} \leq 0, \cdots, \boldsymbol{a}_m^T\boldsymbol{x} \leq 0\} = K(\boldsymbol{a}_1^T) \cap \cdots \cap K(\boldsymbol{a}_m^T)$$

であり、各 $K(\boldsymbol{a}_1^T), \cdots, K(\boldsymbol{a}_m^T)$ は閉集合なので、その共通部分である $K(A)$ も閉集合である。 証明終

定理 2.13 （錐に対する *Farkas* の補題）V を n 次元計量線形空間とし、凸錐 $K \subset V$、$\boldsymbol{a}_1, \cdots, \boldsymbol{a}_m \in V$、$\boldsymbol{b} \in \boldsymbol{R^m}$ とする。もし像 $C[\boldsymbol{a}_1, \cdots, \boldsymbol{a}_m; K]$ が閉集合ならば

1. $\boldsymbol{b} \in C[\boldsymbol{a}_1, \cdots, \boldsymbol{a}_n; K]$
2. $\boldsymbol{b}^T\boldsymbol{y} < 0, \sum_{i=1}^{m} y_i\boldsymbol{a}_i \in K^*$ なる $\boldsymbol{y} \in \boldsymbol{R^m}$ が存在する

のどちらか一方のみが必ず成り立つ。

(proof)
まず、両方成り立つと仮定する。1. が成り立てば、$(\boldsymbol{a}_i, \boldsymbol{x}) = \boldsymbol{b}_i (i = 1, \cdots, m)$ となる $\boldsymbol{x} \in K$ が存在する。このとき

$$\begin{aligned} 0 &> \boldsymbol{b}^T\boldsymbol{y} \\ &= \sum_{i=1}^{m} y_i\boldsymbol{b}_i \\ &= \sum_{i=1}^{m} y_i(\boldsymbol{a}_i, \boldsymbol{x}) \\ &= \left(\sum_{i=1}^{m} y_i\boldsymbol{a}_i, \boldsymbol{x}\right) \geq 0 \quad \because \sum_{i=1}^{m} y_i \in K^*, \boldsymbol{x} \in K \end{aligned}$$

となって矛盾するので、両方成立することはない。

1. が成り立たないとすると、K が凸錐であることより $C[\boldsymbol{a}_1, \cdots, \boldsymbol{a}_n; K] \subset \boldsymbol{R^m}$ は凸錐であり、また、条件より閉集合なので、閉凸錐である。よって、錐の分離定理 2.7 より

$$\forall \boldsymbol{u} \in C[\boldsymbol{a}_1, \cdots, \boldsymbol{a}_n; K], \quad \boldsymbol{b}^T \boldsymbol{y} = \boldsymbol{y}^T \boldsymbol{b} < 0 \leq \boldsymbol{y}^T \boldsymbol{u}$$

なる $\boldsymbol{y} \in \boldsymbol{R^m}$ が存在する。このとき、$\forall \boldsymbol{x} \in K$ に対して

$$\boldsymbol{u} \equiv \begin{pmatrix} (\boldsymbol{a}_1, \boldsymbol{x}) \\ \vdots \\ (\boldsymbol{a}_m, \boldsymbol{x}) \end{pmatrix} \in C[\boldsymbol{a}_1, \cdots, \boldsymbol{a}_n; K]$$

であるので

$$\begin{aligned} 0 &\leq \boldsymbol{y}^T \boldsymbol{u} \\ &= \sum_{i=1}^{m} y_i (\boldsymbol{a}_i, \boldsymbol{x}) \\ &= \left(\sum_{i=1}^{m} y_i \boldsymbol{a}_i, \boldsymbol{x} \right) \end{aligned}$$

である。$\boldsymbol{x} \in K$ は任意だったので

$$\sum_{i=1}^{m} y_i \boldsymbol{a}_i \in K^*$$

であり、上とあわせて 2. が成立する。

以上より示された。 $\boxed{\text{証明終}}$

定理 2.14 （*Farkas* の補題）$m \times n$ 行列 $A, \boldsymbol{b} \in \boldsymbol{R^m}$ に対して

1. $A\boldsymbol{x} = \boldsymbol{b}, \boldsymbol{x} \geq \boldsymbol{0}$ なる $\boldsymbol{x} \in \boldsymbol{R^n}$ が存在する

2. $\boldsymbol{b}^T \boldsymbol{y} < 0, A^T \boldsymbol{y} \geq \boldsymbol{0}$ なる $\boldsymbol{y} \in \boldsymbol{R^m}$が存在する

のどちらか一方のみが必ず成り立つ。

(proof)

$A^T = (\boldsymbol{a}_1, \cdots, \boldsymbol{a}_m), \boldsymbol{a}_i \in \boldsymbol{R^n}$ とおき、$K = \{\boldsymbol{x} \in \boldsymbol{R^n}; \boldsymbol{x} \geq \boldsymbol{0}\}$（非負制約）とすると、これは凸錐である。このとき、補題 2.6 より $K^* = K$ であり、また

$$C[\boldsymbol{a}_1, \cdots, \boldsymbol{a}_m; K] = \{A\boldsymbol{x}; \boldsymbol{x} \geq \boldsymbol{0}\} = C[A]$$

であり、これは定理 2.11 より閉集合である。さらに、$\boldsymbol{y} \in \boldsymbol{R^m}$ に対して

$$\sum_{i=1}^{m} y_i \boldsymbol{a}_i = A^T \boldsymbol{y}$$

である。よって、直前の定理より

1. $A\boldsymbol{x} = \boldsymbol{b}, \boldsymbol{x} \geq \boldsymbol{0}$ なる $\boldsymbol{x} \in \boldsymbol{R^n}$ が存在する
2. $\boldsymbol{b}^T \boldsymbol{y} < 0, A^T \boldsymbol{y} \geq \boldsymbol{0}$ なる $\boldsymbol{y} \in \boldsymbol{R^m}$が存在する

のどちらか一方のみが必ず成り立つ。 証明終

Farkas の補題には、言い換えがいくつかある。

定理 2.15 （*Farkas* の補題 *2*） $A\boldsymbol{x}=\boldsymbol{b},\boldsymbol{x}\geq\boldsymbol{0}$ なる $\boldsymbol{x}\in\boldsymbol{R^n}$ が存在することと

$$A^T\boldsymbol{y}\geq\boldsymbol{0}\rightarrow\boldsymbol{b}^T\boldsymbol{y}\geq 0$$

が成立することは同値である。

定理 2.16 （*Farkas* の補題 *3*）

$$C[A]=K^*(-A^T)$$
$$C^*[A]=K(-A^T)$$

(proof)

$$\begin{aligned}\boldsymbol{b}\in C[A] &\leftrightarrow & A^T\boldsymbol{y}\geq\boldsymbol{0} \text{ ならば } \boldsymbol{b}^T\boldsymbol{y}\geq 0 \quad \because Farkas \text{ の補題 3}\\ &\leftrightarrow & -A^T\boldsymbol{y}\leq\boldsymbol{0} \text{ ならば } \boldsymbol{b}^T\boldsymbol{y}\geq 0\\ &\leftrightarrow & \boldsymbol{y}\in K(-A^T) \text{ ならば } \boldsymbol{b}^T\boldsymbol{y}\geq 0\\ &\leftrightarrow & \boldsymbol{b}\in K^*(-A^T)\end{aligned}$$

より $C[A]=K^*(-A^T)$ である。このとき

$$\begin{aligned}C^*[A]&=(K^*)^*(-A^T)\\ &=K(-A^T) \quad \because \text{定理 2.12 より } K(-A^T) \text{ が閉凸錐であることから定理 2.9}\end{aligned}$$

であり、示される。 証明終

二者択一型の、少し異なる定理も存在する。

定理 2.17 $m\times n$ 行列 A に対して

1. $A\boldsymbol{x}=\boldsymbol{0},\boldsymbol{x}\geq\boldsymbol{0},\boldsymbol{x}\neq\boldsymbol{0}$ なる $\boldsymbol{x}\in\boldsymbol{R^n}$ が存在する

2. $A^T\boldsymbol{y}<\boldsymbol{0}$ なる $\boldsymbol{y}\in\boldsymbol{R^m}$ が存在する

のどちらか一方のみが必ず成り立つ。

(proof)

両方とも成り立つとする。$\boldsymbol{x}\geq\boldsymbol{0},\boldsymbol{x}\neq\boldsymbol{0}$ と $\boldsymbol{b}<\boldsymbol{0}$ に対しては

$$\boldsymbol{x}^T\boldsymbol{b}=\sum_{i=1}^{n}x_ib_i<0$$

が成立する。よって

$$\boldsymbol{x}^T(A^T\boldsymbol{y})<0$$

が成り立たなければならない。ところが、この $\boldsymbol{x}$ について $A\boldsymbol{x}=\boldsymbol{0}$ がなりたつので

$$\boldsymbol{x}^T(A^T\boldsymbol{y})=(A\boldsymbol{x})^T\boldsymbol{y}=0$$

となり、矛盾する。したがって、両方とも成り立つことはない。

ここで、2. が成り立たないとする。このとき、$\forall \boldsymbol{y} \in \boldsymbol{R^n}$ に対して $A^T \boldsymbol{y} \geq \boldsymbol{0}$ が成り立つ。よって、$A^T(-\boldsymbol{y}) = -A^T \boldsymbol{y} \geq \boldsymbol{0}$ であるから

$$\forall \boldsymbol{y} \in \boldsymbol{R^n}, \quad A^T \boldsymbol{y} = \boldsymbol{0}$$

である。このとき $A^T = (\boldsymbol{a}_1, \cdots, \boldsymbol{a}_m)$ とすると

$$A^T \begin{pmatrix} 0 \\ \vdots \\ 1(i \text{ 成分}) \\ \vdots \\ 0 \end{pmatrix} = \boldsymbol{a}_i = \boldsymbol{0}$$

である。つまり $A = \hat{0}$ である。よって、このとき $\boldsymbol{x} \geq \boldsymbol{0}, \boldsymbol{x} \neq \boldsymbol{0}$ なるどんな項をとっても $A\boldsymbol{x} = 0$ である。よって示された。 $\boxed{\text{証明終}}$

3　ラグランジュ関数の鞍点及び双対理論

3.1　ラグランジュ関数の鞍点

定義 3.1 不等式形式の非線形計画問題 [P3] において、$V \subset \boldsymbol{R^n}$ なる集合 V をとる。このとき、V を定義域とするラグランジュ関数 $L_V(\boldsymbol{x}, \boldsymbol{\lambda}) = f(\boldsymbol{x}) - \boldsymbol{\lambda}^T \boldsymbol{g}(\boldsymbol{x})$ について、$\exists \boldsymbol{x}^* \in V, \exists \boldsymbol{\lambda}^* \geq \boldsymbol{0}$ が存在して $\forall \boldsymbol{x} \in V, \forall \boldsymbol{\lambda} \geq \boldsymbol{0}$ に対して

$$L_V(\boldsymbol{x}^*, \boldsymbol{\lambda}) \leq L_V(\boldsymbol{x}^*, \boldsymbol{\lambda}^*) \leq L_V(\boldsymbol{x}, \boldsymbol{\lambda}^*) \tag{1}$$

が成立するならば、$(\boldsymbol{x}^*, \boldsymbol{\lambda}^*)$ を $L_V(\boldsymbol{x}, \boldsymbol{\lambda})$ の**鞍点**[5]という。

定理 3.1 不等式形式の非線形計画問題 [P3] において、$\boldsymbol{x}^* \in V, \boldsymbol{\lambda}^* \geq \boldsymbol{0}$ であるとき、$(\boldsymbol{x}^*, \boldsymbol{\lambda}^*)$ が $L_V(\boldsymbol{x}, \boldsymbol{\lambda})$ の鞍点であることと、以下の *3* 条件が全て成立することは同値である。

$$L_V(\boldsymbol{x}^*, \boldsymbol{\lambda}^*) \leq L_V(\boldsymbol{x}, \boldsymbol{\lambda}^*) \tag{2}$$

$$\boldsymbol{\lambda}^{*T} \boldsymbol{g}(\boldsymbol{x}^*) = \boldsymbol{0} \tag{3}$$

$$\boldsymbol{g}(\boldsymbol{x}^*) \geq \boldsymbol{0} \quad （\boldsymbol{x}^* \text{が許容解}） \tag{4}$$

(proof)

必要性から。(2) は定義の右側そのものである。また、$\forall \boldsymbol{\lambda} \geq \boldsymbol{0}$ について

$$\begin{aligned} L_V(\boldsymbol{x}^*, \boldsymbol{\lambda}) &\leq L_V(\boldsymbol{x}^*, \boldsymbol{\lambda}^*) \\ f(\boldsymbol{x}^*) - \boldsymbol{\lambda}^T \boldsymbol{g}(\boldsymbol{x}^*) &\leq f(\boldsymbol{x}^*) - \boldsymbol{\lambda}^{*T} \boldsymbol{g}(\boldsymbol{x}^*) \\ \boldsymbol{0} &\leq (\boldsymbol{\lambda} - \boldsymbol{\lambda}^*)^T \boldsymbol{g}(\boldsymbol{x}^*) \end{aligned} \tag{5}$$

であるから、$\boldsymbol{\lambda} > \boldsymbol{\lambda}^*$ に対しては

$$\boldsymbol{g}(\boldsymbol{x}^*) \geq \boldsymbol{0}$$

が必要である。また、(5) において $\boldsymbol{\lambda} = \boldsymbol{0}$ とすれば

$$\boldsymbol{0} \leq -\boldsymbol{\lambda}^{*T} \boldsymbol{g}(\boldsymbol{x}^*)$$

であり、一方 $\boldsymbol{g}(\boldsymbol{x}^*) \geq \boldsymbol{0}, \boldsymbol{\lambda}^* \geq \boldsymbol{0}$ より

$$\boldsymbol{0} \leq \boldsymbol{\lambda}^{*T} \boldsymbol{g}(\boldsymbol{x}^*)$$

なので、あわせて

$$\boldsymbol{0} = \boldsymbol{\lambda}^{*T} \boldsymbol{g}(\boldsymbol{x}^*)$$

である。

次に十分性について。$\forall \boldsymbol{\lambda} \geq \boldsymbol{0}$ に対し

$$\begin{aligned} L_V(\boldsymbol{x}^*, \boldsymbol{\lambda}^*) &= f(\boldsymbol{x}^*) - \boldsymbol{\lambda}^{*T} \boldsymbol{g}(\boldsymbol{x}^*) \\ &= f(\boldsymbol{x}^*) - 0 \quad \because \text{式 (3)} \\ &\leq f(\boldsymbol{x}^*) - \boldsymbol{\lambda}^T \boldsymbol{g}(\boldsymbol{x}^*) \quad \because \text{式 (4) 及び} \boldsymbol{\lambda} \geq \boldsymbol{0} \\ &= L_V(\boldsymbol{x}^*, \boldsymbol{\lambda}) \end{aligned}$$

[5] 鞍点は集合 V の選び方に依存することには注意が必要である。

であり、これと式 (2) より

$$L_V(\boldsymbol{x}^*, \boldsymbol{\lambda}) \leq L_V(\boldsymbol{x}^*, \boldsymbol{\lambda}^*) \leq L_V(\boldsymbol{x}, \boldsymbol{\lambda}^*)$$

となる。 証明終

定理 3.2 不等式形式の非線形計画問題 [P3] において、許容集合を S として、$S \subset V, \boldsymbol{x}^* \in V, \boldsymbol{\lambda}^* \geq \boldsymbol{0}$ であるとき、$(\boldsymbol{x}^*, \boldsymbol{\lambda}^*)$ が $L_V(\boldsymbol{x}, \boldsymbol{\lambda})$ の鞍点ならば、$\boldsymbol{x}^*$ は [P3] の**最適解**であり、$L_V(\boldsymbol{x}^*, \boldsymbol{\lambda}^*)$ は最適値である。

(proof)
$\forall \boldsymbol{x} \in S \subset V$ に対し

$$\begin{aligned}
f(\boldsymbol{x}) &\geq f(\boldsymbol{x}) - \boldsymbol{\lambda}^{*T}\boldsymbol{g}(\boldsymbol{x}) \quad \because \boldsymbol{\lambda}^{*T} \geq \boldsymbol{0}, \boldsymbol{g}(\boldsymbol{x}) \geq \boldsymbol{0} (\leftarrow \boldsymbol{x} \in S) \\
&\geq L_V(\boldsymbol{x}, \boldsymbol{\lambda}^*) \\
&\geq L_V(\boldsymbol{x}^*, \boldsymbol{\lambda}^*) \quad \because \text{式 (2)} \\
&= f(\boldsymbol{x}^*) - \boldsymbol{\lambda}^{*T}\boldsymbol{g}(\boldsymbol{x}^*) \\
&= f(\boldsymbol{x}^*) \quad \because \text{式 (3)}
\end{aligned}$$

であり、$\boldsymbol{x}^*$ は最適解である。また、このとき既に上に示している通り、$f(\boldsymbol{x}^*) = L_V(\boldsymbol{x}^*, \boldsymbol{\lambda}^*)$ であり、$L_V(\boldsymbol{x}^*, \boldsymbol{\lambda}^*)$ は最適値である。 証明終

3.2 双対問題と鞍点

定義 3.2 不等式形式の非線形計画問題 [P3] において

$$LP_V(\boldsymbol{x}) \equiv \sup_{\boldsymbol{\lambda} \geq \boldsymbol{0}} L_V(\boldsymbol{x}, \boldsymbol{\lambda}) \tag{6}$$

$$LD_V(\boldsymbol{x}) \equiv \inf_{\boldsymbol{x} \in V} L_V(\boldsymbol{x}, \boldsymbol{\lambda}) \tag{7}$$

を、それぞれ $L_V(\boldsymbol{x}, \boldsymbol{\lambda})$ の**主関数・双対関数**という。

定理 3.3 不等式形式の非線形計画問題 [P3] は、S を許容集合として、$S \subset V$ となるよう V を選べば、主関数を用いて

$$\left| \begin{array}{ll} minimize & LP_V(\boldsymbol{x}) \\ s.t. & \boldsymbol{x} \in S(\subset V) \end{array} \right. \qquad \text{[P6]}$$

と書き換えられる。

(proof)
$\forall \boldsymbol{x} \in S$ について、$\boldsymbol{\lambda} \geq \boldsymbol{0}, \boldsymbol{g}(\boldsymbol{x}) \geq \boldsymbol{0}$ より $L_V(\boldsymbol{x}, \boldsymbol{\lambda}) = f(\boldsymbol{x}) - \boldsymbol{\lambda}^T \boldsymbol{g}(\boldsymbol{x})$ を最大化する $\boldsymbol{\lambda}$ は $\boldsymbol{\lambda} = \boldsymbol{0}$ である。よってこのとき

$$LP_V(\boldsymbol{x}) = \sup_{\boldsymbol{\lambda} \geq \boldsymbol{0}} L_V(\boldsymbol{x}, \boldsymbol{\lambda}) = f(\boldsymbol{x})$$

であり、定理が示される。 証明終

補題 3.4 $\forall \boldsymbol{x} \in V, \forall \boldsymbol{\lambda} \geq \boldsymbol{0}$ について

$$LD_V(\boldsymbol{\lambda}) \leq L_V(\boldsymbol{x}, \boldsymbol{\lambda}) \leq LP_V(\boldsymbol{x})$$

(proof)
上限・下限の定義より明らか。 証明終

上の補題は、常に主関数が双対関数以上であることを表しているが、ラグランジュ関数の鞍点においては、これが一致する。それを示すのが次の定理である。

定理 3.5 不等式形式の非線形計画問題 [P3] において、$\boldsymbol{x}^* \in V, \boldsymbol{\lambda}^* \geq \boldsymbol{0}$ であるとき、$(\boldsymbol{x}^*, \boldsymbol{\lambda}^*)$ が $L_V(\boldsymbol{x}, \boldsymbol{\lambda})$ の鞍点ならば

$$\begin{aligned}\sup_{\boldsymbol{\lambda}\geq\boldsymbol{0}} LD_V(\boldsymbol{\lambda}) &= L_V(\boldsymbol{x}^*, \boldsymbol{\lambda}^*) = \inf_{\boldsymbol{x}\in V} LP_V(\boldsymbol{x}) \\ &= LD_V(\boldsymbol{\lambda}^*) = LP_V(\boldsymbol{x}^*)\end{aligned}$$

が成立する。

(proof)
補題 3.4 より $\boldsymbol{x} \in V, \boldsymbol{\lambda} \geq \boldsymbol{0}$ は任意なので

$$\sup_{\boldsymbol{\lambda}\geq\boldsymbol{0}} LD_V(\boldsymbol{\lambda}) \leq \inf_{\boldsymbol{x}\in V} LP_V(\boldsymbol{x})$$

である。また

$$\begin{aligned}\sup_{\boldsymbol{\lambda}\geq\boldsymbol{0}} LD_V(\boldsymbol{\lambda}) &\geq LD_V(\boldsymbol{\lambda}^*) = \inf_{\boldsymbol{x}\in V} L_V(\boldsymbol{x}, \boldsymbol{\lambda}^*) \quad \because \text{上限} \\ &\geq L_V(\boldsymbol{x}^*, \boldsymbol{\lambda}^*) \quad \because \text{鞍点の定義} \\ &\geq \sup_{\boldsymbol{\lambda}\geq\boldsymbol{0}} L_V(\boldsymbol{x}^*, \boldsymbol{\lambda}) = LP_V(\boldsymbol{x}^*) \quad \because \text{鞍点の定義} \\ &\geq \inf_{\boldsymbol{x}\in V} LP_V(\boldsymbol{x}) \quad \because \text{下限}\end{aligned}$$

なので、あわせて途中の式はすべて等号で結ばれ、示される。 証明終

ところで、不等式形式の非線形計画問題 [P3] は主関数 $LP_V(\boldsymbol{x})$ の最小化問題 [P6] に書き換えられ、また、許容集合を S とするとき、$S \subset V$ ならば定理 3.2 より [P3] の最適解すなわち [P6] の最適解が鞍点の $\boldsymbol{x}^*$ に等しく、最適値[6] は $L_V(\boldsymbol{x}^*, \boldsymbol{\lambda}^*)$ に等しい。つまり、上の定理 3.5 の等式の値は、$S \subset V$ の場合には、[P3] の最適値に等しいということである。よって、上の左側の等式は、最適値と $\sup_{\boldsymbol{\lambda}\geq\boldsymbol{0}} LD_V(\boldsymbol{\lambda})$ が等しいということをあらわしており、$S \subset V$ でラグランジュ関数に鞍点が存在する場合には、$\sup_{\boldsymbol{\lambda}\geq\boldsymbol{0}} LD_V(\boldsymbol{\lambda})$ を考えることにより、最適値が得られるということである。このことについては、最適解について調べる別ルートの可能性として、検討してみる価値があるといえよう。そこで次の「双対問題」が導入される。

定義 3.3 不等式形式の非線形計画問題 [P3] に対し、$\sup_{\boldsymbol{\lambda}\geq\boldsymbol{0}} LD_V(\boldsymbol{\lambda})$ を求めることに相当する問題

$$\left|\begin{array}{ll} maximize & LD_V(\boldsymbol{\lambda}) \\ s.t. & \boldsymbol{\lambda} \geq \boldsymbol{0} \end{array}\right. \qquad [\mathrm{P7}]$$

を**双対問題**という。これに対し、元の問題を**主問題**という。双対関数 $LD_V(\boldsymbol{\lambda}) = -\infty$ となる場合は、考えても意味が無いので、これが下に有界になるよう $LD_V(\boldsymbol{\lambda}) > -\infty$ という制約を加えてよい。

双対問題を考えるには、双対関数を計算する必要があり、これ自身も最適化問題である。

[6] $\inf_{\boldsymbol{x}\in V} LP_V(\boldsymbol{x})$ は、V が直接的には許容集合でないので、それだけでこれが [P3] の最適値であるとはいえない。

定義 3.4 不等式形式の非線形計画問題 [P3] について、$\boldsymbol{\lambda} \geq \mathbf{0}$ に対して、双対関数 $LD_V(\boldsymbol{\lambda})$ を求めることに相当する問題

$$\left| \begin{array}{ll} minimize & L_V(\boldsymbol{x}, \boldsymbol{\lambda}) \\ s.t. & \boldsymbol{x} \in V \end{array} \right. \qquad \text{[P8]}$$

を**ラグランジュ問題**という。$V = \boldsymbol{R^n}$ と選べる場合には、無制約最適化問題となる。

以上に導入した事柄を用いると、定理 3.5 を次のように発展させて述べることができる。

定理 3.6 不等式形式の非線形計画問題 [P3] において、許容集合を S とし、$\boldsymbol{x}^* \in V, \boldsymbol{\lambda}^* \geq \mathbf{0}, S \subset V$ であるとき、$(\boldsymbol{x}^*, \boldsymbol{\lambda}^*)$ が $L_V(\boldsymbol{x}, \boldsymbol{\lambda})$ の鞍点であれば、$\boldsymbol{x}^*, \boldsymbol{\lambda}^*$ がそれぞれ主問題・双対問題の最適解であり、最適値に関して

$$LD_V(\boldsymbol{\lambda}^*) = L_V(\boldsymbol{x}^*, \boldsymbol{\lambda}^*) = LP_V(\boldsymbol{x}^*) = f(\boldsymbol{x}^*)$$

が成立する。

(proof)
$S \subset V$ なので、主問題 [P3] は主関数 $LP_V(\boldsymbol{x})$ の最小化問題 [P6] に書き換えられ、また定理 3.2 より [P3] の最適解すなわち [P6] の最適解が鞍点の $\boldsymbol{x}^*$ に等しく、最適値は $L_V(\boldsymbol{x}^*, \boldsymbol{\lambda}^*)$ に等しい。

双対問題 [P7] については、定理 3.5 から

$$\sup_{\boldsymbol{\lambda} \geq \mathbf{0}} LD_V(\boldsymbol{\lambda}) = LD_V(\boldsymbol{\lambda}^*) = L_V(\boldsymbol{x}^*, \boldsymbol{\lambda}^*)$$

なので、双対問題の最適値 $\sup_{\boldsymbol{\lambda} \geq \mathbf{0}} LD_V(\boldsymbol{\lambda}) = L_V(\boldsymbol{x}^*, \boldsymbol{\lambda}^*)$ が $LD_V(\boldsymbol{\lambda}^*)$ によって実現されているということになり、$\boldsymbol{\lambda}^*$ は双対問題の最適解である。 $\boxed{\text{証明終}}$

定理 3.7 不等式形式の非線形計画問題 [P3] において、許容集合を S とし、$\boldsymbol{x}^* \in V, \boldsymbol{\lambda}^* \geq \mathbf{0}, S \subset V$ であるとき、$(\boldsymbol{x}^*, \boldsymbol{\lambda}^*)$ が $L_V(\boldsymbol{x}, \boldsymbol{\lambda})$ の鞍点であることと

$$LD_V(\boldsymbol{\lambda}^*) = f(\boldsymbol{x}^*) \neq \pm\infty \quad \text{かつ} \quad \boldsymbol{x}^* \in S$$

であることは同値である。ただし、目的関数 f については S で、双対関数 LD_V については $\boldsymbol{\lambda} \geq \mathbf{0}$ で、有限の値をとる点があるとする。

(proof)
必要性に関しては、定理 3.6 と、f 及び LD_V が有限の値をとり得ることより

$$LD_V(\boldsymbol{\lambda}^*) = f(\boldsymbol{x}^*) \neq \pm\infty$$

であり、また、定理 3.1 の式 (4) より $\boldsymbol{x}^* \in S$ である。

十分性に関して。条件より任意の $\boldsymbol{x} \in V$ について

$$\begin{aligned} L_V(\boldsymbol{x}, \boldsymbol{\lambda}^*) &\geq LD_V(\boldsymbol{\lambda}^*) = f(\boldsymbol{x}) \\ &\geq f(\boldsymbol{x}^*) - (\boldsymbol{\lambda}^*)^T \boldsymbol{g}(\boldsymbol{x}^*) \quad \because \boldsymbol{x}^* \in S, \boldsymbol{\lambda}^* \geq \mathbf{0} \\ &= L_V(\boldsymbol{x}^*, \boldsymbol{\lambda}^*) \end{aligned}$$

であり、$\boldsymbol{x}=\boldsymbol{x}^*$ とすれば、途中の式もすべて等号で結ばれるので

$$LD_V(\boldsymbol{\lambda}^*)=L_V(\boldsymbol{x}^*,\boldsymbol{\lambda}^*) \quad (\boldsymbol{\lambda}^*)^T\boldsymbol{g}(\boldsymbol{x}^*)=0$$

となる。よって定理 3.1 より $(\boldsymbol{x}^*,\boldsymbol{\lambda}^*)$ は $L_V(\boldsymbol{x},\boldsymbol{\lambda})$ の鞍点である。 証明終

系 3.8 許容解 $\bar{\boldsymbol{x}}\in S,\bar{\boldsymbol{\lambda}}\geq \mathbf{0}$ について、

$$LD_V(\bar{\boldsymbol{\lambda}})=f(\bar{\boldsymbol{x}})\neq\pm\infty$$

となるならば、$\bar{\boldsymbol{x}},\bar{\boldsymbol{\lambda}}$ はそれぞれ主問題・双対問題の最適解である。

3.3 弱双対定理及び双対性

定理 3.9 （弱双対定理）不等式形式の非線形計画問題 [P3] について、許容集合を S とするとき、$S\subset V$ であれば、主問題の最適値を P^*、双対問題の最適値を D^* とおくと

$$D^*=\sup_{\boldsymbol{\lambda}\geq\mathbf{0}}LD_V(\boldsymbol{\lambda})\leq\inf_{\boldsymbol{x}\in V}LP_V(\boldsymbol{x})\leq P^* \tag{8}$$

となる。

(proof)

$$\begin{aligned}
P^* &= \inf_{\boldsymbol{x}\in S}LP_V(\boldsymbol{x}) && \because S\subset V \text{ より [P3] と [P6] は同値}\\
&\geq \inf_{\boldsymbol{x}\in V}LP_V(\boldsymbol{x}) && \because S\subset V\\
&\geq \sup_{\boldsymbol{\lambda}\geq\mathbf{0}}LD_V(\boldsymbol{\lambda}) && \because \text{補題 3.4}\\
&= D^*
\end{aligned}$$

証明終

定理 3.10 （弱双対定理 *2*）不等式形式の非線形計画問題 [P3] について、許容集合を S とするとき、$S\subset V$ であれば、任意の主問題・双対問題の許容解 $\boldsymbol{x}^*,\boldsymbol{\lambda}^*$ に対して

$$LD_V(\boldsymbol{\lambda})\leq f(\boldsymbol{x})$$

となる。

定義 3.5 不等式形式の非線形計画問題 [P3] について、主問題の最適解を $\boldsymbol{x}^*$、双対問題の最適解を $\boldsymbol{\lambda}^*$ とおくと弱双対定理 *3.9* より最適値の差には

$$f(\boldsymbol{x}^*)-LD_V(\boldsymbol{\lambda}^*)\geq 0$$

という関係が成り立つが、この差が 0 であるとき、すなわち主問題と双対問題の最適値が一致する場合には、問題 [P3] は**双対性**が成立するという。逆に一致しない場合は、**双対ギャップ・双対性の隙間**があるという。

弱双対定理の証明から分かるように、双対ギャップは、V と許容集合 S が同じでないことと、主関数・双対関数の差、すなわち、上限と下限をとる順序の差との二つの要因から生じる。主問題の最適値を得ずに、双対性の成立を、任意の問題に対して判定するのは困難であり、あらかじめ双対性が成り立つ問題のクラスを特定しておくことが有効である。凸計画問題は、その代表格である。

3.4 弱双対定理の応用

定理 3.11 双対問題が非有界ならば、主問題に有限の目的関数値を与える許容解は存在しない。また、主問題が非有界ならば、双対問題に有限の双対関数値を与える許容解は存在しない。

目的関数は、許容解に対しては有限である場合も多い。もしくは、許容集合を目的関数値が有限である範囲にあらかじめ絞っておく考え方もある。その場合には、上の定理は「**双対問題が非有界ならば、主問題は実行不能である。**」ということになる。

定理 3.12 $\boldsymbol{x}, \boldsymbol{\lambda}$ をそれぞれ主問題・双対問題の許容解、P^*, D^* をそれぞれ主問題・双対問題の最適値とすると

$$LD_V(\boldsymbol{\lambda}) \leq D^* \leq P^* \leq f(\boldsymbol{x})$$

である。

この定理により、最適値を厳密な不等式によって評価できる。

4 局所最適性と Karush-Kuhn-Tucker 条件

不等式形の非線形計画 [P3] に対して、目的関数及び制約式に数回の微分可能性を仮定して、局所最適性を論じる。

4.1 接錐

定義 4.1 計量線形空間 X において、$S \subset X, \boldsymbol{x}^* \in S$ が与えられたとき $\boldsymbol{x}_n \to \boldsymbol{x}^*$ なる S の無限点列 $\{\boldsymbol{x}_n\}(\boldsymbol{x}_n \neq \boldsymbol{x}^*)$ が存在して

$$\frac{\boldsymbol{t}}{\|\boldsymbol{t}\|} = \lim_{n\to\infty} \frac{\boldsymbol{x}_n - \boldsymbol{x}^*}{\|\boldsymbol{x}_n - \boldsymbol{x}^*\|}$$

となるような $\boldsymbol{t}$ を、$\boldsymbol{x}^*$ における S の**接ベクトル**と呼ぶ。また、接ベクトル全体と $\boldsymbol{0}$ をあわせた集合は、明らかに錐である。これを**接錐**と呼び $T_S(\boldsymbol{x}^*)$ と表す。

定理 4.1 不等式形の非線形計画 [P3] について、目的関数を $f : \boldsymbol{R^n} \to \boldsymbol{R}$、許容集合を S とする。このとき $\boldsymbol{x}^* \in S$ が極小解であり、f が $\boldsymbol{x}^*$ で微分可能ならば

$$\forall \boldsymbol{t} \in T_S(\boldsymbol{x}^*), \quad \boldsymbol{t}^T \nabla f(\boldsymbol{x}^*) \geq \boldsymbol{0}$$

である。これは

$$\nabla f(\boldsymbol{x}^*) \in T_S^*(\boldsymbol{x}^*)$$

と書き換えられる[7]。

(proof)

$\boldsymbol{t} \in T_S(\boldsymbol{x}^*)$ に対して、$\boldsymbol{x}_n \to \boldsymbol{x}^*$ なる S の無限点列 $\{\boldsymbol{x}_n\}(\boldsymbol{x}_n \neq \boldsymbol{x}^*)$ が存在して

$$\frac{\boldsymbol{t}}{\|\boldsymbol{t}\|} = \lim_{n\to\infty} \frac{\boldsymbol{x}_n - \boldsymbol{x}^*}{\|\boldsymbol{x}_n - \boldsymbol{x}^*\|}$$

となる。$\boldsymbol{x}^*$ が $S \cap V$ の極小解なので、十分大きな k に対して、適当な $0 \leq \theta \leq 1$ が存在して

$$\begin{aligned} 0 &\leq \frac{f(\boldsymbol{x}_k) - f(\boldsymbol{x}^*)}{\|f(\boldsymbol{x}_k) - f(\boldsymbol{x}^*)\|} \\ &= \left(\frac{\boldsymbol{x}_k - \boldsymbol{x}^*}{\|f(\boldsymbol{x}_k) - f(\boldsymbol{x}^*)\|} \right)^T \nabla f(\theta \boldsymbol{x}_k + (1-\theta)\boldsymbol{x}^*) \quad \because \text{平均値の定理 2.2} \end{aligned}$$

であるから、$k \to \infty$ とすると

$$0 \leq \left(\frac{\boldsymbol{t}}{\|\boldsymbol{t}\|} \right)^T \nabla f(\boldsymbol{x}^*)$$

が成り立つ。 証明終

4.2 有効な制約・線形化錐

制約式について、以下のことが言える。

定理 4.2 $g : \boldsymbol{R^n} \to \boldsymbol{R}$ が $\boldsymbol{x}$ で連続であるとき、$g(\boldsymbol{x}) > 0$ ならば、任意の $\boldsymbol{e} \in \boldsymbol{R^n}$ に対して、十分小さな $\delta > 0$ をとれば

$$g(\boldsymbol{x} + \delta \boldsymbol{e}) \geq 0$$

とできる。

[7] $T_S^*(\boldsymbol{x}^*)$ は $T_S(\boldsymbol{x}^*)$ の双対錐。

(proof)

$\|\boldsymbol{e}\| = 0$ なら明らか。$\|\boldsymbol{e}\| \neq 0$ のとき。定理 B.16 において、$\epsilon = g(\boldsymbol{x}) > 0$ に対して

$$g(V_{\frac{\delta}{\|\epsilon\|}}(a)) \subset V_\epsilon(g(a))$$

となる δ が存在する。よって示された。 証明終

これは、微小な変化という観点では、$g(\boldsymbol{x}) > 0$ ならば $\boldsymbol{x}$ において g という制約式は制約として機能していない、ということをあらわしている。そこで、次の定義が置かれる。

定義 4.2 許容点 $\boldsymbol{x}^*$ において、$g_i(\boldsymbol{x}^*) = 0$ なる制約式を $\boldsymbol{x}^*$ において**有効な制約**という。また、$g_i(\boldsymbol{x}^*) > 0$ なる制約式を $\boldsymbol{x}^*$ において**無効な制約**という。さらに、$I(\boldsymbol{x}^*)$ を有効な制約を与える添字集合

$$I(\boldsymbol{x}^*) \equiv \{i; g_i(\boldsymbol{x}^*) = 0\}$$

と定義する。

定義 4.3 許容集合を S、$\boldsymbol{x}^* \in S$ とする。$I(\boldsymbol{x}^*) = \{i(1), \cdots, i(M)\}$ として $A = (\nabla g_{i(1)}(\boldsymbol{x}^*), \cdots, \nabla g_{i(M)}(\boldsymbol{x}^*))^T$ とおくとき

$$L_S(\boldsymbol{x}^*) \equiv K(-A) = \{\boldsymbol{u} \in \boldsymbol{R^n}; \forall i \in I(\boldsymbol{x}^*), \boldsymbol{u}^T \nabla g_i(\boldsymbol{x}^*) \geq 0\}$$

を S の許容点 $\boldsymbol{x}^*$ における**線形化錐**という。定理 *2.12* より線形化錐は閉凸錐である。

補題 4.3 許容点 $\boldsymbol{x}^*$ に対して $T_S(\boldsymbol{x}^*) \subset L_S(\boldsymbol{x}^*)$

(proof)

$g_i, \forall i \in I(\boldsymbol{x}^*)$ について $g_i(\boldsymbol{x}^*) = 0$ は許容集合における最小解のひとつなので、定理 4.1 より、$\forall \boldsymbol{t} \in T_S(\boldsymbol{x}^*)$ に対し

$$0 \leq \left(\frac{\boldsymbol{t}}{\|\boldsymbol{t}\|}\right)^T \nabla g(\boldsymbol{x}^*) \quad \therefore 0 \leq \boldsymbol{t}^T \nabla g(\boldsymbol{x}^*)$$

であり $\boldsymbol{t} \in L_S(\boldsymbol{x}^*)$ となる。よって示された。 証明終

4.3 Karush-Kuhn-Tucker 条件

定義 4.4 不等式形の非線形計画 [P3] について、$\boldsymbol{x} \in \boldsymbol{R^n}$ に対し

$$\nabla f(\boldsymbol{x}) - \sum_{i=1}^{m} \lambda_i \nabla g_i(\boldsymbol{x}) = \boldsymbol{0} \tag{9}$$

$$\lambda_i g_i(\boldsymbol{x}) = 0 \quad (i = 1, \cdots, m) \tag{10}$$

$$\boldsymbol{g}(\boldsymbol{x}) \geq \boldsymbol{0} \tag{11}$$

$$\boldsymbol{\lambda} \geq \boldsymbol{0} \tag{12}$$

を満たすベクトル $\boldsymbol{\lambda} \in \boldsymbol{R^m}$ が存在することを ***Karush-Kuhn-Tucker* 条件・*KKT* 条件**[8] といい、これを満たす点 $\boldsymbol{x}$ を ***KKT* 点**という。また、このとき $\boldsymbol{\lambda}$ を**一般化ラグランジュ乗数**という。さらに、式 (10) は、有効な制約のみが条件において意味を成すことをあらわしており、**相補性条件**という。なお、KKT 条件はラグランジュ関数 $L(\boldsymbol{x}, \boldsymbol{\lambda})$ を用いて

$$\nabla_{\boldsymbol{x}} L(\boldsymbol{x}, \lambda) = \boldsymbol{0} \tag{13}$$

$$\boldsymbol{\lambda}^T \nabla_{\boldsymbol{\lambda}} L(\boldsymbol{x}, \lambda) = \boldsymbol{0} \tag{14}$$

$$\nabla_{\boldsymbol{\lambda}} L(\boldsymbol{x}, \lambda) \leq \boldsymbol{0} \tag{15}$$

$$\boldsymbol{\lambda} \geq \boldsymbol{0} \tag{16}$$

[8] Kunh-Tucker 条件という流儀もある（特に経済関係ではそうよばれる）。これは Karush の修士論文が 1939 年と、線形計画法が提案される前で、Kuhn と Tucker が非線形最適化問題の最適性条件として発表する 11 年前でもあり、当時、注目を浴びなかったことに由来している。

ともあらわせる。

定義 4.5 不等式形の非線形計画 [P3] において、条件 Q の成立のもと、許容解 $\boldsymbol{x}$ について

$$KKT\text{ 点でない} \to \text{極小解でない}$$

となるならば、条件 Q を ***1* 次の制約想定**と呼び、$\boldsymbol{x}$ において *1* 次の制約想定が満たされるという。

制約想定については、後に詳しく見るとして、この定義のもと、当然ながら次の定理が成り立つ。

定理 4.4 不等式形の非線形計画 [P3] について、極小解 $\boldsymbol{x}^*$ において *1* 次の制約想定が満たされるならば、KKT 条件が成立する。

4.4 1 次の制約想定

次の定理の条件こそが 1 次の制約想定の基本である。狭義にはこれを 1 次の制約想定という。

定理 4.5 不等式形の非線形計画 [P3] において、S を許容集合、$\boldsymbol{x}^*$ を許容解とするとき

$$\text{条件} : T_S(\boldsymbol{x}^*) = L_S(\boldsymbol{x}^*)$$

は *1* 次の制約想定である。

(proof)

$\boldsymbol{x}^*$ が極小解であるとき、KKT 条件を満たすことを示す。まず、許容解であることから

$$\boldsymbol{g}(\boldsymbol{x}) \geq \boldsymbol{0}$$

である。$I(\boldsymbol{x}^*) = \{i(1), \cdots, i(M)\}$ として $A = (\nabla g_{i(1)}(\boldsymbol{x}^*), \cdots, \nabla g_{i(M)}(\boldsymbol{x}^*))^T$ とする。このとき

$$\begin{aligned}
\nabla f(\boldsymbol{x}^*) \in T_S^*(\boldsymbol{x}^*) &= L_S^*(\boldsymbol{x}^*) \quad \because \text{定理 4.1 及び条件} \\
&= K^*(-A) \\
&= C[A^T] \quad \because Farkas\text{ の補題 3(定理 2.16)} \\
&= \{A^T\boldsymbol{x}; \boldsymbol{x} \geq \boldsymbol{0}\}
\end{aligned}$$

である。よって $\lambda_i^* \geq 0 (i \in I(\boldsymbol{x}^*))$ によって

$$\nabla f(\boldsymbol{x}^*) = \sum_{k=1}^{M} \lambda_{i(k)}^* g_{i(k)}(\boldsymbol{x}^*) = \sum_{i \in I(\boldsymbol{x}^*)} \lambda_i^* g_i(\boldsymbol{x}^*)$$

となる。ここで、$i \notin I(\boldsymbol{x}^*)$ に対して $\lambda_i^* = 0$ とすれば

$$\begin{aligned}
\nabla f(\boldsymbol{x}^*) - \sum_{i=1}^{m} \lambda_i^* \nabla g_i(\boldsymbol{x}^*) &= \boldsymbol{0} \\
\lambda_i^* g_i(\boldsymbol{x}^*) &= 0 \quad (i = 1, \cdots, m) \\
\boldsymbol{\lambda}^* &\geq \boldsymbol{0}
\end{aligned}$$

であり、KKT 条件を満たす。 $\boxed{\text{証明終}}$

制約想定は十分条件ということもあり、いろいろなものがあるが、議論の都合上、まず、Cottle の制約想定を挙げる。

定理 4.6 （*Cottle* の制約想定）下の条件が成り立つとき $T_S(\boldsymbol{x}^*) = L_S(\boldsymbol{x}^*)$ であり

$$\forall i \in I(\boldsymbol{x}^*) \text{ に対して} \begin{cases} \boldsymbol{u}^T \nabla g_i(\boldsymbol{x}^*) > 0 & g_i \neq \text{一次式} \\ \boldsymbol{u}^T \nabla g_i(\boldsymbol{x}^*) \geq 0 & g_i = \text{一次式} \end{cases}$$

$$\text{となる } \boldsymbol{u} \in \boldsymbol{R^M} \text{が存在する}$$

したがって、この条件は *1* 次の制約想定である。

(proof)

$\forall \boldsymbol{y} \in L_S(\boldsymbol{x}^*)$ について、$\boldsymbol{y} = \boldsymbol{0}$ ならば明らかに $\boldsymbol{y} \in T_S(\boldsymbol{x}^*)$ である。$\boldsymbol{y} \neq \boldsymbol{0}$ のとき、$0 < \exists w < 1$ をとり

$$\boldsymbol{x}_k \equiv \boldsymbol{x}^* + w^k \boldsymbol{y} + w^{2k} \boldsymbol{u}$$

と定義すると、$\boldsymbol{x}_k \to \boldsymbol{x}^*$ であり

$$\boldsymbol{x}_k - \boldsymbol{x}^* = w^k(\boldsymbol{y} + w^k \boldsymbol{u})$$

となる。$\boldsymbol{y} \neq \boldsymbol{0}$ なので $\boldsymbol{x}_k \neq \boldsymbol{x}^*$ である。ここで $i \notin I(\boldsymbol{x}^*)$ に対しては $g_i(\boldsymbol{x}^*) > 0$ なので定理 4.2 より、十分大きな k に対して

$$g_i(\boldsymbol{x}_k) = g_i(\boldsymbol{x}^* + w^k(\boldsymbol{y} + w^k \boldsymbol{u})) \geq 0 \quad (i \notin I(\boldsymbol{x}^*))$$

である。有効な制約 $i \in I(\boldsymbol{x}^*)$ に対しては、平均値の定理 2.2 より

$$g_i(\boldsymbol{x}_k) - g_i(\boldsymbol{x}^*) = w^k(\boldsymbol{y} + w^k \boldsymbol{u})^T \nabla g_i(\boldsymbol{x}^* + \theta w^k(\boldsymbol{y} + w^k \boldsymbol{u})) \quad 0 \leq \theta \leq 1$$

である。$\boldsymbol{y}^T \nabla g_i(\boldsymbol{x}^*) \geq 0$ であり、g_i が一次式でないならば $\boldsymbol{u}^T \nabla g_i(\boldsymbol{x}^*) > 0$ であるから

$$(\boldsymbol{y} + w^k \boldsymbol{u})^T \nabla g_i(\boldsymbol{x}^*) > 0$$

となる。よって定理 4.2 より十分大きな k に対して

$$(\boldsymbol{y} + w^k \boldsymbol{u})^T \nabla g_i(\boldsymbol{x}^* + \theta w^k(\boldsymbol{y} + w^k \boldsymbol{u})) \geq 0$$

である。g_i が一次式ならば $\nabla g_i(\boldsymbol{x}) = const$ なので、やはり

$$\begin{aligned} (\boldsymbol{y} + w^k \boldsymbol{u})^T \nabla g_i(\boldsymbol{x}^* + \theta w^k(\boldsymbol{y} + w^k \boldsymbol{u})) &= (\boldsymbol{y} + w^k \boldsymbol{u})^T \nabla g_i(\boldsymbol{x}^*) \\ &= \boldsymbol{y}^T \nabla g_i(\boldsymbol{x}^*) + w^k \boldsymbol{u}^T \nabla g_i(\boldsymbol{x}^*) \\ &\geq 0 \quad \because \text{一次式に対する条件および } \boldsymbol{y}^T \nabla g_i(\boldsymbol{x}^*) \geq 0 \end{aligned}$$

であり、一次式であろうとなかろうと

$$g_i(\boldsymbol{x}_k) = g_i(\boldsymbol{x}^*) + w^k(\boldsymbol{y} + w^k \boldsymbol{u})^T \nabla g_i(\boldsymbol{x}^* + \theta w^k(\boldsymbol{y} + w^k \boldsymbol{u})) \geq 0 \quad (i \in I(\boldsymbol{x}^*))$$

となる。以上より、十分大きな k に対して $\{x_k\}$ は許容集合 S の x^* に収束する $\boldsymbol{x}_k \neq \boldsymbol{x}^*$ なる無限点列である。よって、接錐の定義より

$$\lim_{k \to \infty} \frac{\boldsymbol{x}_k - \boldsymbol{x}^*}{\|\boldsymbol{x}_k - \boldsymbol{x}^*\|} = \lim_{k \to \infty} \frac{w^k(\boldsymbol{y} + w^k \boldsymbol{u})}{\|w^k(\boldsymbol{y} + w^k \boldsymbol{u})\|} = \frac{\boldsymbol{y}}{\|\boldsymbol{y}\|} \in T_S(\boldsymbol{x}^*) \quad \therefore \boldsymbol{y} \in T_S(\boldsymbol{x}^*)$$

であり、つまり $L_S(\boldsymbol{x}^*) \subset T_S(\boldsymbol{x}^*)$ である。また、定理 4.3 より $T_S(\boldsymbol{x}^*) \subset L_S(\boldsymbol{x}^*)$ であるから、あわせて

$$L_S(\boldsymbol{x}^*) = T_S(\boldsymbol{x}^*)$$

である。 $\boxed{\text{証明終}}$

系 4.7 下の条件が成り立つとき $T_S(\boldsymbol{x}^*) = L_S(\boldsymbol{x}^*)$ である。

$$\forall i \in I(\boldsymbol{x}^*) \text{ に対して } \boldsymbol{u}^T \nabla g_i(\boldsymbol{x}^*) > 0 \text{ となる } \boldsymbol{u} \in \boldsymbol{R^M} \text{が存在する}$$

したがって、この条件は *1* 次の制約想定である。

4.4.1 John の条件

KKT 条件より弱くなるが、制約想定を必要としない極小解の必要条件、すなわち、最適性条件を導くこともできる。

補題 4.8 定理 *4.7* の制約想定が成立しない許容解 $\boldsymbol{x}^*$ について

$$\sum_{i=1}^{m} \lambda_i \nabla g_i(\boldsymbol{x}) = \boldsymbol{0} \tag{17}$$

$$\lambda_i g_i(\boldsymbol{x}) = 0 \quad (i = 1, \cdots, m) \tag{18}$$

$$\boldsymbol{g}(\boldsymbol{x}) \geq \boldsymbol{0} \tag{19}$$

$$\boldsymbol{\lambda} \geq \boldsymbol{0} \tag{20}$$

$$\boldsymbol{\lambda} \neq \boldsymbol{0} \tag{21}$$

を満たすベクトル $\boldsymbol{\lambda} \in \boldsymbol{R^m}$ が存在する。

(proof)

$I(\boldsymbol{x}^*) = \{i(1), \cdots, i(M)\}$ として $A = (\nabla g_{i(1)}(\boldsymbol{x}^*), \cdots, \nabla g_{i(M)}(\boldsymbol{x}^*))^T$ とおくとき、定理 4.7 の制約想定が成立しないということは

$$-A\boldsymbol{u} < 0 \text{ なる } \boldsymbol{u} \in \boldsymbol{R^n} \text{が存在しない}$$

ということであり、このとき定理 2.17 より

$$A^T \boldsymbol{\lambda}' = \boldsymbol{0}, \boldsymbol{\lambda}' \geq \boldsymbol{0}, \boldsymbol{\lambda}' \neq \boldsymbol{0}$$

なる $\boldsymbol{\lambda}' \in \boldsymbol{R^M}$ が存在する。ここで $i = 1, \cdots, m$ に対して

$$\lambda_i = \begin{cases} \lambda'_j : i(j) = i & i \in I(\boldsymbol{x}^*) \text{ のとき} \\ 0 & i \notin I(\boldsymbol{x}^*) \text{ のとき} \end{cases}$$

とすれば、明らかに

$$\lambda_i g_i(\boldsymbol{x}) = 0 \quad (i = 1, \cdots, m)$$
$$\boldsymbol{\lambda} \geq \boldsymbol{0}$$
$$\boldsymbol{\lambda} \neq \boldsymbol{0}$$

であり、また $A^T \boldsymbol{\lambda}' = 0$ より

$$\begin{aligned} A^T \boldsymbol{\lambda}' &= (\nabla g_{i(1)}(\boldsymbol{x}^*), \cdots, \nabla g_{i(M)}(\boldsymbol{x}^*))\boldsymbol{\lambda}' \\ &= \sum_{j=1}^{M} \lambda'_j \nabla g_{i(j)}(\boldsymbol{x}^*) \\ &= \sum_{i=1}^{m} \lambda_i \nabla g_i(\boldsymbol{x}^*) = 0 \end{aligned}$$

である。また許容解なので $\boldsymbol{g}(\boldsymbol{x}) \geq \boldsymbol{0}$ は当然成立する。よって示された。 $\boxed{\text{証明終}}$

これと、KKT 条件を組み合わせることによって、必要条件を導ける。

定理 4.9 （*John* の条件）不等式形の非線形計画 [P3] について、極小解 $\boldsymbol{x}^*$ に対し

$$\lambda_0 \nabla f(\boldsymbol{x}) - \sum_{i=1}^{m} \lambda_i \nabla g_i(\boldsymbol{x}) = \boldsymbol{0} \tag{22}$$

$$\lambda_i g_i(\boldsymbol{x}) = 0 \quad (i = 1, \cdots, m) \tag{23}$$

$$\boldsymbol{g}(\boldsymbol{x}) \geq \boldsymbol{0} \tag{24}$$

$$\boldsymbol{\lambda} \geq \boldsymbol{0} \tag{25}$$

$$\boldsymbol{\lambda} \neq \boldsymbol{0} \tag{26}$$

を満たすベクトル $\boldsymbol{\lambda} \in \boldsymbol{R}^{m+1}$ が存在する。

$\lambda_0 \neq 0$ の場合が KKT 条件に相当する。これは、式 (22) を λ_0 で割ることによって導ける。$\lambda_0 = 0$ の場合が KKT 条件を満たさない極小解の満たす条件である。この条件には目的関数の情報が入っておらず、やや特殊な条件といえるだろう。なお、John の条件を導くのには定理 4.7 の制約想定を用いたが、KKT 条件を満たさない極小解がそれを満たすとは限らない。

4.4.2 正規条件

定義 4.6 許容解 $\boldsymbol{x}^*$ に対して $\{\nabla g_i(\boldsymbol{x}^*)\}_{i \in I(\boldsymbol{x}^*)}$ が線形独立であるとき、$\boldsymbol{x}^*$ において**正規条件**が満たされるという。

定理 4.10 正規条件は *1* 次の制約想定である。

(proof)

正規条件のもと、極小解が KKT 条件を満たすことを示せばよい。John の条件 4.9 より極小解 $\boldsymbol{x}^*$ について

$$\lambda_0 \nabla f(\boldsymbol{x}) - \sum_{i=1}^{m} \lambda_i \nabla g_i(\boldsymbol{x}) = \boldsymbol{0} \tag{27}$$

$$\lambda_i g_i(\boldsymbol{x}) = 0 \quad (i = 1, \cdots, m) \tag{28}$$

$$\boldsymbol{g}(\boldsymbol{x}) \geq \boldsymbol{0} \tag{29}$$

$$\boldsymbol{\lambda} \geq \boldsymbol{0} \tag{30}$$

$$\boldsymbol{\lambda} \neq \boldsymbol{0} \tag{31}$$

を満たすベクトル $\boldsymbol{\lambda} \in \boldsymbol{R}^{m+1}$ が存在する。ここで、$\lambda_0 = 0$ と仮定すると

$$\sum_{i=1}^{m} \lambda_i \nabla g_i(\boldsymbol{x}) = \sum_{i \in I(\boldsymbol{x}^*)} \lambda_i \nabla g_i(\boldsymbol{x}) = \boldsymbol{0}$$

であるが、この線形関係について $\{\nabla g_i(\boldsymbol{x}^*)\}_{i \in I(\boldsymbol{x}^*)}$ が線形独立なので

$$\lambda_i = 0 \quad (i \in I(\boldsymbol{x}^*))$$

である。よって、これと、仮定及び式 (28) より

$$\boldsymbol{\lambda} = \boldsymbol{0}$$

ということになり、John の条件の $\boldsymbol{\lambda} \neq \boldsymbol{0}$ に矛盾する。よって、$\lambda_0 \neq 0$ であり、式 (27) を λ_0 で割ることにより、KKT 条件をみたすことが導かれる。 $\boxed{\text{証明終}}$

4.5 無制約最適化問題

無制約の最適化問題に対しては、定理 4.7 の制約想定が常に成立するので、常に極小解は KKT 条件を満たす。KKT 条件も、制約条件がないので一般化ラグランジュ乗数を導入する必要もない。よって、次の定理のようになる。

定理 4.11 無制約の最適化問題

$$\left|\ \textit{minimize} \quad f(\boldsymbol{x})\right. \qquad \text{[P9]}$$

に対しては、極小解について常に KKT 条件が成立し、その KKT 条件は

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

である。

4.6 等式制約・ラグランジュの未定乗数法

等式制約のみによる問題

$$\left|\ \begin{array}{ll} \text{minimize} & f(\boldsymbol{x}) \\ \text{s.t.} & \boldsymbol{g}(\boldsymbol{x}) = 0 \end{array}\right. \qquad \text{[P10]}$$

における KKT 条件を考える。これは、不等式形 [P3] に変形することができる。つまり

$$\left|\ \begin{array}{ll} \text{minimize} & f(\boldsymbol{x}) \\ \text{s.t.} & \boldsymbol{g}(\boldsymbol{x}) \geq 0 \\ & -\boldsymbol{g}(\boldsymbol{x}) \geq 0 \end{array}\right. \qquad \text{[P11]}$$

と記述できる。これに対し、KKT 条件を考えて整理すればよい。

定理 4.12 等式制約のみの非線形最適化問題 [P10] の KKT 条件は

$$\nabla f(\boldsymbol{x}) - \sum_{i=1}^{m} \lambda_i \nabla g_i(\boldsymbol{x}) = \mathbf{0} \tag{32}$$

$$\boldsymbol{g}(\boldsymbol{x}) = \mathbf{0} \tag{33}$$

である。

(proof)

不等式形で記述した問題 [P11] について KKT 条件は

$$\begin{aligned} \nabla f(\boldsymbol{x}) - \sum_{i=1}^{m} \lambda_i^+ \nabla g_i(\boldsymbol{x}) - \sum_{i=1}^{m} \lambda_i^- \nabla\{-g_i(\boldsymbol{x})\} &= \mathbf{0} \\ \lambda_i^{\pm} g_i(\boldsymbol{x}) &= 0 \quad (i = 1, \cdots, m) \\ \boldsymbol{g}(\boldsymbol{x}) &\geq \mathbf{0} \\ -\boldsymbol{g}(\boldsymbol{x}) &\geq \mathbf{0} \\ \boldsymbol{\lambda}^{\pm} &\geq \mathbf{0} \end{aligned}$$

を満たすベクトル $\boldsymbol{\lambda}^+, \boldsymbol{\lambda}^- \in \boldsymbol{R^m}$ が存在する、となる。これらを整理すると

$$\begin{aligned}\nabla f(\boldsymbol{x}) - \sum_{i=1}^{m}(\lambda_i^+ - \lambda_i^-)\nabla g_i(\boldsymbol{x}) &= \mathbf{0} \\ \lambda_i^{\pm} g_i(\boldsymbol{x}) &= 0 \quad (i = 1, \cdots, m) \\ \boldsymbol{g}(\boldsymbol{x}) &= \mathbf{0} \\ \boldsymbol{\lambda}^{\pm} &\geq \mathbf{0}\end{aligned}$$

$\lambda_i^{\pm} g_i(\boldsymbol{x}) = 0 \quad (i = 1, \cdots, m)$ は $\boldsymbol{g}(\boldsymbol{x}) = 0$ より明らかなので、冗長である[9]。$(\lambda_i^+ - \lambda_i^-), \lambda_i^{\pm} \geq 0$ は結局任意の実数を表現でき、変数を二つに分けて記述する意味がないので $\boldsymbol{\lambda} = \boldsymbol{\lambda}^+ - \boldsymbol{\lambda}^-$ とすると結局

$$\begin{aligned}\nabla f(\boldsymbol{x}) - \sum_{i=1}^{m}\lambda_i \nabla g_i(\boldsymbol{x}) &= \mathbf{0} \\ \boldsymbol{g}(\boldsymbol{x}) &= \mathbf{0}\end{aligned}$$

となる。 $\boxed{\text{証明終}}$

　等式制約の最適化問題に対して、この KKT 条件を用いて極値を得るのが、いわゆる、**ラグランジュの未定乗数法**である。等式制約と不等式制約が混在している問題についても、容易に類推が可能であろう。ちなみにこの条件をラグランジュ関数を用いて表すと

$$\nabla_{\boldsymbol{x},\boldsymbol{\lambda}} L(\boldsymbol{x}, \boldsymbol{\lambda}) = 0$$

となり、制約想定のもと、等式制約の最適化問題は、ラグランジュ関数の無制約最適化問題に書き換えられる。

4.7　2 次の条件

定理 4.13 （*2* 次の必要条件）*1* 次の制約想定のもと、$\boldsymbol{x}^*$ が極小解であり、目的関数と制約式がそれぞれその近傍で *2* 回微分可能であるとする。$\boldsymbol{x}^*$ は *KKT* 条件を満たすので、その一般化ラグランジュ乗数を $\boldsymbol{\lambda}(\boldsymbol{x}^*)$ としたとき

$$\begin{aligned}&I^+(\boldsymbol{x}^*) \equiv \{i; \lambda(\boldsymbol{x}^*)_i > 0\} \\ &W(\boldsymbol{x}^*) \equiv \{\boldsymbol{x}; i \in I^+(\boldsymbol{x}^*) \text{ に対して } g_i(\boldsymbol{x}) = 0, i \notin I^+(\boldsymbol{x}^*) \text{ に対して } g_i(\boldsymbol{x}) \geq 0\}\end{aligned}$$

と定義する。このとき、$\nabla_{\boldsymbol{x}}^2 L$ をラグランジュ関数のヘッセ行列として

$$\forall \boldsymbol{t} \in T_W(\boldsymbol{x}^*), \quad \boldsymbol{t}^T \nabla_{\boldsymbol{x}}^2 L(\boldsymbol{x}^*, \boldsymbol{\lambda}(\boldsymbol{x}^*))\boldsymbol{t} \geq 0$$

である。

(proof)

$\boldsymbol{t} \in T_{W(\boldsymbol{x}^*)}$ に対して、$\boldsymbol{x}_k \to \boldsymbol{x}^*$ なる $W(\boldsymbol{x}^*)$ の無限点列 $\{\boldsymbol{x}_k\}(\boldsymbol{x}_k \neq \boldsymbol{x}^*)$ が存在して

$$\lim_{k\to\infty} \frac{\boldsymbol{x}_k - \boldsymbol{x}^*}{\|\boldsymbol{x}_k - \boldsymbol{x}^*\|} = \frac{\boldsymbol{t}}{\|\boldsymbol{t}\|}$$

[9] 等式制約ではすべてが有効な制約である事に対応している

となる。また、$\boldsymbol{x}^*$ が極小解なので、十分大きな k に対して、$0 \leq \exists\theta \leq 1$ が存在して

$$
\begin{aligned}
0 &\leq f(\boldsymbol{x}_k) - f(\boldsymbol{x}^*) \\
&= L(\boldsymbol{x}_k, \boldsymbol{\lambda}(\boldsymbol{x}^*)) - L(\boldsymbol{x}^*, \boldsymbol{\lambda}(\boldsymbol{x}^*)) + \boldsymbol{\lambda}(\boldsymbol{x}^*)^T \boldsymbol{g}(\boldsymbol{x}_k) + \boldsymbol{\lambda}(\boldsymbol{x}^*)^T \boldsymbol{g}(\boldsymbol{x}^*) \\
&= L(\boldsymbol{x}_k, \boldsymbol{\lambda}(\boldsymbol{x}^*)) - L(\boldsymbol{x}^*, \boldsymbol{\lambda}(\boldsymbol{x}^*)) \quad \because \lambda_i > 0 \text{ ならば } g_i(\boldsymbol{x}_k) = 0 \text{ および } KKT \text{ 条件の相補性条件} \\
&= (\boldsymbol{x}_k - \boldsymbol{x}^*)^T \nabla_{\boldsymbol{x}} L(\boldsymbol{x}^*, \boldsymbol{\lambda}(\boldsymbol{x}^*)) + \frac{1}{2}(\boldsymbol{x}_k - \boldsymbol{x}^*)^T \nabla_{\boldsymbol{x}}^2 L(\theta\boldsymbol{x}_k + (1-\theta)\boldsymbol{x}^*, \boldsymbol{\lambda}(\boldsymbol{x}^*))(\boldsymbol{x}_k - \boldsymbol{x}^*) \quad \because \text{テイラー展開 (定理 2.3)} \\
&= \frac{1}{2}(\boldsymbol{x}_k - \boldsymbol{x}^*)^T \nabla_{\boldsymbol{x}}^2 L(\theta\boldsymbol{x}_k + (1-\theta)\boldsymbol{x}^*, \boldsymbol{\lambda}(\boldsymbol{x}^*))(\boldsymbol{x}_k - \boldsymbol{x}^*) \quad \because KKT \text{ 条件} \\
0 &\leq \left(\frac{\boldsymbol{x}_k - \boldsymbol{x}^*}{\|\boldsymbol{x}_k - \boldsymbol{x}^*\|}\right)^T \nabla_{\boldsymbol{x}}^2 L(\theta\boldsymbol{x}_k + (1-\theta)\boldsymbol{x}^*, \boldsymbol{\lambda}(\boldsymbol{x}^*)) \left(\frac{\boldsymbol{x}_k - \boldsymbol{x}^*}{\|\boldsymbol{x}_k - \boldsymbol{x}^*\|}\right)
\end{aligned}
$$

であるから、$k \to \infty$ とすると

$$0 \leq \boldsymbol{t}^T \nabla_{\boldsymbol{x}}^2 L(\boldsymbol{x}^*, \boldsymbol{\lambda}(\boldsymbol{x}^*))\boldsymbol{t}$$

である。 $\boxed{\text{証明終}}$

定義 4.7 $W(\boldsymbol{x}^*)$ に対応する線形化錐は、$g_i(\boldsymbol{x}) = 0 \leftrightarrow g_i(\boldsymbol{x}) \geq 0, g_i(\boldsymbol{x}) \leq 0$ と見ることにより

$$L_W(\boldsymbol{x}^*) \equiv \{\boldsymbol{u} \in \boldsymbol{R}^{\boldsymbol{n}}; i \in I^+(\boldsymbol{x}^*) \text{ に対して } \boldsymbol{u}^T \nabla g_i(\boldsymbol{x}^*) = 0, i \in I(\boldsymbol{x}^*) \text{ に対して } \boldsymbol{u}^T \nabla g_i(\boldsymbol{x}^*) \geq 0\}$$

となる。条件 Q のもと

$$L_W(\boldsymbol{x}^*) = T_W(\boldsymbol{x}^*)$$

が成立するとき、条件 Q を ***2* 次の制約想定**と呼ぶ。この成立のもと、定理 *4.13* は、より検証しやすい形となる。正規条件は、*2* 次の制約想定である。(証明略)

定理 4.14 (*2* 次の十分条件)KKT 点 $\boldsymbol{x}^*$ の周りで目的関数 f および制約式 g_i が二回微分可能であるとき、$\forall \boldsymbol{u} \in L_W(\boldsymbol{x}^*), \boldsymbol{u} \neq \boldsymbol{0}$ に対して

$$\boldsymbol{u}^T \nabla_{\boldsymbol{x}}^2 L(\boldsymbol{x}^*, \boldsymbol{\lambda}(\boldsymbol{x}^*))\boldsymbol{u} > 0$$

となるならば、$\boldsymbol{x}^*$ は孤立極小解である。

(proof)
証明略 $\boxed{\text{証明終}}$

第 III 部
線形計画問題

5　線形計画問題の概要

5.1　線形計画問題の表現

線形計画問題とは、目的関数が $\boldsymbol{R}^{\boldsymbol{n}}$ を定義域とする線形（1 次）関数で、制約も同様の線形関数の等式及び不等式で表される最適化問題である。より記号的に表現すれば

$$\left|\begin{array}{ll} \text{minimize} & \boldsymbol{c}^T\boldsymbol{x} \\ \text{s.t.} & \boldsymbol{a}_i^T\boldsymbol{x} \geq 0 \quad (i = 1, \cdots, l) \\ & \boldsymbol{a}_i^T\boldsymbol{x} = 0 \quad (i = l+1, \cdots, m) \end{array}\right. \qquad \text{[P12]}$$

となる。

線形計画問題も、非線形計画問題の一種なので、同様にして定理 2.1 にあるように、不等式形・不等式標準形・等式標準形に書き換えられる。以下に、線形計画問題における各表現について記す。また、制約式については、多くの式をまとめて行列の形で表現することが多い。

不等式形

$$\left|\begin{array}{ll} \text{minimize} & \boldsymbol{c}^T\boldsymbol{x} \\ \text{s.t.} & A\boldsymbol{x} \geq \boldsymbol{b} \end{array}\right. \qquad \text{[P13]}$$

不等式標準形

$$\left|\begin{array}{ll} \text{minimize} & \boldsymbol{c}^T\boldsymbol{x} \\ \text{s.t.} & A\boldsymbol{x} \geq \boldsymbol{b} \\ & \boldsymbol{x} \geq \boldsymbol{0} \end{array}\right. \qquad \text{[P14]}$$

等式標準形

$$\left|\begin{array}{ll} \text{minimize} & \boldsymbol{c}^T\boldsymbol{x} \\ \text{s.t.} & A\boldsymbol{x} = \boldsymbol{b} \\ & \boldsymbol{x} \geq \boldsymbol{0} \end{array}\right. \qquad \text{[P15]}$$

5.2　解のパターン

定理 5.1 線形計画問題の値域は閉集合である。

(proof)

不等式形 [P13] で考える。このとき、許容集合 $S = \{\boldsymbol{x}; A\boldsymbol{x} \geq \boldsymbol{b}\} = \{\boldsymbol{x}; \boldsymbol{a}_i^T\boldsymbol{x} \geq \boldsymbol{b}_i(i = 1, \cdots, m)\}$ であり、値域 $f(S) = \{\boldsymbol{c}^T\boldsymbol{x}; A\boldsymbol{x} \geq \boldsymbol{b}\}$ である。実行不能ならば、$f(S) = \phi$ であり、明らかに閉集合である。実行可能であるとき。$\boldsymbol{c} = \boldsymbol{0}$ ならば、$f(S) = \{0\}$ で閉集合である。$\boldsymbol{c} \neq \boldsymbol{0}$ ならば

$$f(\boldsymbol{R}^{\boldsymbol{n}}) = f(S \cup S^c) = f(S) \cup f(S^c) = \boldsymbol{R}$$

なので $f(S)^c = f(S^c) = \{\boldsymbol{x}; \exists j, \boldsymbol{a}_j^T\boldsymbol{x} < \boldsymbol{b}_j\}$ である。このとき、$\forall\boldsymbol{x} \in f(S)^c$ に対して、十分小さな $\epsilon > 0$ をとれば

$$\boldsymbol{a}_j^T\boldsymbol{x} < \boldsymbol{a}_j^T\boldsymbol{x} + \epsilon|\boldsymbol{a}_j^T\boldsymbol{x}| < \boldsymbol{b}_j$$

とできるので

$$\{\boldsymbol{x} + \alpha\boldsymbol{x}; |\alpha| < \epsilon\} \subset S^c$$

である。よって

$$f(S)^c = f(S^c) \supset f(\{\boldsymbol{x}+\alpha\boldsymbol{x}; |\alpha| < \epsilon\}) = \{\boldsymbol{c}^T\boldsymbol{x}+\alpha\boldsymbol{c}^T\boldsymbol{x}; |\alpha| < \epsilon\} = (\boldsymbol{c}^T\boldsymbol{x}-\epsilon\boldsymbol{c}^T\boldsymbol{x}, \boldsymbol{c}^T\boldsymbol{x}-\epsilon\boldsymbol{c}^T\boldsymbol{x}) = V_{\epsilon\boldsymbol{c}^T\boldsymbol{x}}(\boldsymbol{c}^T\boldsymbol{x})$$

であり、$\boldsymbol{x} \in f(S)^c$ は任意だったので、$f(S)^c$ は開集合である。よって、$f(S)$ は閉集合である。 証明終

定理 5.2 実行可能で有界な線形計画問題は、最適解を持つ。

(proof)
直前の定理より、線形計画問題の値域 $f(S) \subset \boldsymbol{R}$ は閉集合である。また、条件より $f(S)$ は下に有界であり、また、実行可能なので $f(S) \neq \phi$ である。よって、定理 C.10 より $f(S)$ は最小値を持つ。$f(S)$ は像なので、対応する S の元が存在し、これが最適解である。 証明終

上の定理は、「実行可能・有界・最適解なし」というパターンが、線形計画問題では存在しないことを表している。よって、線形計画問題の解のパターンは、以下のようになる。

線形計画問題のパターンのまとめ

実行不可能		
実行可能	非有界	
実行可能	有界	最適解が存在

5.3 線形計画問題における KKT 条件

定理 5.3 線形計画問題は、実行可能であれば制約想定が満たされる。すなわち、極小解は KKT 条件を満たす。

(proof)
[P12] において、任意の許容解 $\boldsymbol{x}^*$ に対して

$$(\boldsymbol{x}^*)^T \nabla g_i(\boldsymbol{x}^*) = \boldsymbol{a}_i^T \boldsymbol{x}^* = g_i(\boldsymbol{x}^*) \geq 0$$

であるから、Cottle の制約想定（定理 4.6）が常に満たされる。 証明終

定理 5.4 （線形計画問題における KKT 条件）等式標準形の線形計画問題 [P15] に対する KKT 条件は

$$\boldsymbol{c}^T\boldsymbol{x} = \boldsymbol{b}^T\boldsymbol{y} \tag{34}$$

$$A^T\boldsymbol{y} \leq \boldsymbol{c} \tag{35}$$

$$A\boldsymbol{x} = \boldsymbol{b} \tag{36}$$

$$\boldsymbol{x} \geq \boldsymbol{0} \tag{37}$$

を満たす $\boldsymbol{y} \in \boldsymbol{R^n}$ が存在することと、同値である。

(proof)
KKT 条件をそのまま書けば、$\boldsymbol{y}, \boldsymbol{\lambda} \in \boldsymbol{R^n}$ が存在して

$$\begin{aligned} \boldsymbol{c} - A^T\boldsymbol{y} - \boldsymbol{\lambda} &= \boldsymbol{0} \\ \boldsymbol{\lambda}^T\boldsymbol{x} &= 0 \\ A\boldsymbol{x} = \boldsymbol{b}, \quad \boldsymbol{x} &\geq \boldsymbol{0} \\ \boldsymbol{\lambda} &\geq \boldsymbol{0} \end{aligned}$$

である。まず必要性から。すでに許容性 $A\boldsymbol{x}=\boldsymbol{b}, \boldsymbol{x}\geq\boldsymbol{0}$ は示されている。

$$\boldsymbol{\lambda}=\boldsymbol{c}-A^T\boldsymbol{y}$$

であるので、残りの条件から

$$\begin{aligned}\boldsymbol{\lambda}=\boldsymbol{c}-A^T\boldsymbol{y}\geq\boldsymbol{0} &\quad \therefore A^T\boldsymbol{y}\leq\boldsymbol{c}\\ \boldsymbol{\lambda}^T\boldsymbol{x}=0 &\\ (\boldsymbol{c}-A^T\boldsymbol{y})^T\boldsymbol{x}=\boldsymbol{c}^T\boldsymbol{x}-\boldsymbol{y}^TA\boldsymbol{x}=0 &\\ \boldsymbol{c}^T\boldsymbol{x}-\boldsymbol{y}^T\boldsymbol{b}=0 &\quad \therefore \boldsymbol{c}^T\boldsymbol{x}=\boldsymbol{b}^T\boldsymbol{y}\end{aligned}$$

である。

　十分性に関しては $\boldsymbol{\lambda}\equiv\boldsymbol{c}-A^T\boldsymbol{y}$ とおくことにより

$$\begin{aligned}\boldsymbol{\lambda}\geq\boldsymbol{0}\\ \boldsymbol{c}-A^T\boldsymbol{y}-\boldsymbol{\lambda}=\boldsymbol{0}\end{aligned}$$

であり

$$\begin{aligned}\boldsymbol{\lambda}^T\boldsymbol{x}&=(\boldsymbol{c}-A^T\boldsymbol{y})^T\boldsymbol{x}\\ &=\boldsymbol{c}^T\boldsymbol{x}-\boldsymbol{y}^TA\boldsymbol{x}\\ &=\boldsymbol{c}^T\boldsymbol{x}-\boldsymbol{y}^T\boldsymbol{b}=\boldsymbol{c}^T\boldsymbol{x}-\boldsymbol{b}^T\boldsymbol{y}\\ &=0\end{aligned}$$

となる。 $\boxed{\text{証明終}}$

定理 5.5 （線形計画問題における *KKT* 条件 *2*）不等式標準形の線形計画問題 [P14] に対する *KKT* 条件は

$$\boldsymbol{c}^T\boldsymbol{x}=\boldsymbol{b}^T\boldsymbol{y} \tag{38}$$
$$A^T\boldsymbol{y}\leq\boldsymbol{c} \tag{39}$$
$$\boldsymbol{y}\geq\boldsymbol{0} \tag{40}$$
$$A\boldsymbol{x}\geq\boldsymbol{b} \tag{41}$$
$$\boldsymbol{x}\geq\boldsymbol{0} \tag{42}$$

を満たす $\boldsymbol{y}\in\boldsymbol{R^n}$ が存在することと、同値である。

(proof)
　不等式標準形 [P14] に対しては、スラック変数 $\boldsymbol{s}$ を導入することで、等式標準形

$$\left|\begin{array}{ll}\text{minimize} & \begin{pmatrix}\boldsymbol{c}^T & \boldsymbol{0}^T\end{pmatrix}\begin{pmatrix}\boldsymbol{x}\\ \boldsymbol{s}\end{pmatrix}\\ \text{s.t.} & \begin{pmatrix}A & -E\end{pmatrix}\begin{pmatrix}\boldsymbol{x}\\ \boldsymbol{s}\end{pmatrix}=\boldsymbol{b}\\ & \begin{pmatrix}\boldsymbol{x}\\ \boldsymbol{s}\end{pmatrix}\geq\boldsymbol{0}\end{array}\right.$$

に変形でき、これに定理 5.4 を適用して、スラック変数を整理することにより得られる。 $\boxed{\text{証明終}}$

5.4 線形計画のトピックス

5.4.1 l_∞ ノルム最小化問題

定義 5.1 $\boldsymbol{v} \in \boldsymbol{R^n}$ に対して

$$\|\boldsymbol{v}\|_\infty \equiv \max_i |(\boldsymbol{v})_i|$$

を、**l_∞ ノルム・*sup* ノルム・最大値ノルム**といい、行列 A、$\boldsymbol{b} \in \boldsymbol{R^n}$ について

$$\left|\begin{array}{ll} minimize & \|\boldsymbol{x}\|_\infty \\ s.t. & \boldsymbol{x} \in S \subset \boldsymbol{R^n} \end{array}\right. \qquad \text{[P16]}$$

を **l_∞ ノルム最小化問題**という。

定理 5.6 l_∞ ノルム最小化問題 [P16] は

$$\left|\begin{array}{ll} minimize & t \\ s.t. & -t\boldsymbol{1} \le \boldsymbol{x} \le t\boldsymbol{1} \\ & \boldsymbol{x} \in S \subset \boldsymbol{R^n} \end{array}\right. \qquad \text{[P17]}$$

に書き換えられる。ただし、$\boldsymbol{1} \in \boldsymbol{R^n}$ は全成分が 1 のベクトルである。

(proof)

問題 [P16] の最適解・最適値を $\boldsymbol{x}_A \in S, t_A$ とし、問題 [P17] の最適解・最適値を $\boldsymbol{x}_B \in S, t_B$ とする。このとき

$$\begin{aligned} t_A &= \max_i |(\boldsymbol{x}_A)_i| \\ &\ge |(\boldsymbol{x}_A)_i| \\ \therefore -t_A\boldsymbol{1} &\le \boldsymbol{x}_A \le t_A\boldsymbol{1} \end{aligned}$$

であるから、$\boldsymbol{x}_A, t_A$ は [P17] の許容解であり

$$t_A \ge t_B$$

となる。

また、逆に [P17] の最適解・最適値を $\boldsymbol{x}_B, t_B$ について、$i = 1, \cdots, n$ に対して $t_B \ge 0$ であり

$$|(\boldsymbol{x}_B)_i| \le t_B$$

が成立しているが、すべての $i = 1, \cdots, n$ に対して等号が成立しないとすると

$$|(\boldsymbol{x}_B)_i| < t' < t_B \quad (i = 1, \cdots, n)$$

となる t' がとれて、t_B が最適値であることに矛盾する。よって、$\exists j$ について

$$\begin{aligned} |(\boldsymbol{x}_B)_j| &= t_B \\ &\ge |(\boldsymbol{x}_B)_i| \quad (i = 1, \cdots, n) \end{aligned}$$

であるから

$$t_B = \max_i |(\boldsymbol{x}_B)_i| = \|\boldsymbol{x}_B\|_\infty$$

であり、$\boldsymbol{x}_B$ は [P16] の許容解ということになり

$$t_A \le t_B = \|\boldsymbol{x}_B\|_\infty$$

となる。よって、あわせて

$$\|\boldsymbol{x}_A\|_\infty = t_A = t_B = \|\boldsymbol{x}_B\|_\infty$$

であり、確かに [P17] の最適解 $\boldsymbol{x}_B$ が [P16] の最適値を与える許容解、すなわち、最適解となっている。 $\boxed{\text{証明終}}$

よって、S が線形制約であれば、線形計画問題に書き換えられるということである。

5.4.2 l_1 ノルム最小化問題

定義 5.2 $\boldsymbol{v} \in \boldsymbol{R^n}$ に対して

$$\|\boldsymbol{v}\|_p \equiv \left(\sum_{i=1}^{n} |(\boldsymbol{v})_i|^p\right)^{\frac{1}{p}}$$

を、l_p **ノルム**といい、行列 A、$\boldsymbol{b} \in \boldsymbol{R^n}$ について

$$\left|\begin{array}{ll} minimize & \|\boldsymbol{x}\|_1 \\ s.t. & \boldsymbol{x} \in S \subset \boldsymbol{R^n} \end{array}\right. \qquad [P18]$$

を l_1 **ノルム最小化問題**という。

定理 5.7 l_1 ノルム最小化問題 [P18] は

$$\left|\begin{array}{ll} minimize & \boldsymbol{1}^T\boldsymbol{t} \\ s.t. & -\boldsymbol{t} \le \boldsymbol{x} \le \boldsymbol{t} \\ & \boldsymbol{t} \ge \boldsymbol{0} \\ & \boldsymbol{x} \in S \subset \boldsymbol{R^n} \end{array}\right. \qquad [P19]$$

に書き換えられる。ただし、$\boldsymbol{1} \in \boldsymbol{R^n}$ は全成分が 1 のベクトルである。

(proof)

[P18] の最適解を $\boldsymbol{x}_A \in S$、[P19] の最適解を $\boldsymbol{x}_B \in S, \boldsymbol{t}_B$ とする。このとき、$\boldsymbol{t}_A \equiv (|(\boldsymbol{x}_A)_1|, \cdots, |(\boldsymbol{x}_A)_n|)^T$ とおくと

$$\begin{aligned} -\boldsymbol{t}_A \le \boldsymbol{x}_A \le \boldsymbol{t}_A \\ \boldsymbol{t}_A \ge \boldsymbol{0} \\ \boldsymbol{x}_A \in S \end{aligned}$$

を満たすので、$\boldsymbol{x}_A, \boldsymbol{t}_A$ は [P19] の許容解であり、したがって

$$\boldsymbol{1}^T\boldsymbol{t}_A = \sum_{i=1}^{n} |(\boldsymbol{x}_A)_i| = \|\boldsymbol{x}_A\|_1 \ge \boldsymbol{1}^T\boldsymbol{t}_B$$

である。

逆に $\boldsymbol{x}_B, \boldsymbol{t}_B$ について、$-\boldsymbol{t}_B \le \boldsymbol{x}_B \le \boldsymbol{t}_B, \boldsymbol{t} \ge \boldsymbol{0}$ より

$$-(\boldsymbol{t}_B)_i \le (\boldsymbol{x}_B)_i \le (\boldsymbol{t}_B)_i \quad \leftrightarrow \quad |(\boldsymbol{x}_B)_i| \le (\boldsymbol{t}_B)_i$$

であるが、$\boldsymbol{t}_B$ は最適値であり、目的関数の最小化には各要素が小さければ小さいほどよいので

$$|(\boldsymbol{x}_B)_i| = (\boldsymbol{t}_B)_i$$

である。このとき

$$\begin{aligned}\mathbf{1}^T \boldsymbol{t}_B &= \sum_{i=1}^{n} (\boldsymbol{t}_B)_i = \sum_{i=1}^{n} |(\boldsymbol{x}_B)_i| \\ &= \|\boldsymbol{x}_B\|_1 \\ &\geq \inf_{\boldsymbol{x}\in S} \|\boldsymbol{x}\|_1 = \|\boldsymbol{x}_A\|_1\end{aligned}$$

から、上とあわせて

$$\|\boldsymbol{x}_A\|_1 = \mathbf{1}^T \boldsymbol{t}_B = \|\boldsymbol{x}_B\|_1$$

となり、確かに [P19] の最適解 $\boldsymbol{x}_B$ が、[P18] の最適値を与える許容解、すなわち最適解となっている。 証明終

これも、やはり S が線形制約ならば線形計画問題である。

6 線形計画問題の双対問題

6.1 双対問題の導出

等式標準形の線形計画問題 [P15]

$$\left|\begin{array}{ll} \text{minimize} & \boldsymbol{c}^T\boldsymbol{x} \\ \text{s.t.} & A\boldsymbol{x} = \boldsymbol{b} \\ & \boldsymbol{x} \geq \boldsymbol{0} \end{array}\right.$$

について考える。ラグランジュ関数は、$\boldsymbol{x}$ に関する定義域を $V = \boldsymbol{R^n}$ ととることができて

$$\begin{aligned} L(\boldsymbol{x},\boldsymbol{y},\boldsymbol{\lambda}) &= \boldsymbol{c}^T\boldsymbol{x} - \boldsymbol{y}^T(A\boldsymbol{x}-\boldsymbol{b}) - \boldsymbol{\lambda}^T\boldsymbol{x} \\ &= (\boldsymbol{c} - A^T\boldsymbol{y} - \boldsymbol{\lambda})^T\boldsymbol{x} + \boldsymbol{b}^T\boldsymbol{y} \quad (\boldsymbol{\lambda} \geq \boldsymbol{0}) \end{aligned} \tag{43}$$

である。このとき、双対関数を得るためのラグランジュ問題 [P8] は

$$\left|\begin{array}{ll} \text{minimize} & L_V(\boldsymbol{x},\boldsymbol{y},\boldsymbol{\lambda}) = (\boldsymbol{c} - A^T\boldsymbol{y} - \boldsymbol{\lambda})^T\boldsymbol{x} + \boldsymbol{b}^T\boldsymbol{y} \\ \text{s.t.} & \boldsymbol{x} \in \boldsymbol{R^n} \quad \boldsymbol{y},\boldsymbol{\lambda}\ \text{は定数とみなす} \end{array}\right. \tag{P20}$$

という、無制約の線形計画問題となる。最適解の必要条件は、定理 4.11 より

$$\nabla_{\boldsymbol{x}} L_V(\boldsymbol{x},\boldsymbol{\lambda}) = \boldsymbol{c} - A^T\boldsymbol{y} - \boldsymbol{\lambda} = 0$$

である。逆に $\boldsymbol{c} - A^T\boldsymbol{y} - \boldsymbol{\lambda} \neq 0$ ならば最適解は存在しないということになる。つまり、双対関数は

$$LD(\boldsymbol{x},\boldsymbol{\lambda}) = \begin{cases} \boldsymbol{b}^T\boldsymbol{y} & \boldsymbol{c} - A^T\boldsymbol{y} - \boldsymbol{\lambda} = 0 \\ -\infty & \boldsymbol{c} - A^T\boldsymbol{y} - \boldsymbol{\lambda} \neq 0 \end{cases}$$

である。双対問題は最大化問題であり $LD_V(\boldsymbol{\lambda}) = -\infty$ の場合を考える意味がないので、$LD_V(\boldsymbol{\lambda}) > -\infty$ という制約を付加できる。このもと、[P15] の双対問題 [P7] は

$$\left|\begin{array}{ll} \text{maximize} & \boldsymbol{b}^T\boldsymbol{y} \\ \text{s.t.} & \boldsymbol{\lambda} \geq \boldsymbol{0} \\ & \boldsymbol{c} - A^T\boldsymbol{y} - \boldsymbol{\lambda} = 0 \end{array}\right. \tag{P21}$$

となる。$\boldsymbol{\lambda}$ は目的関数に現れないので、これを明示しない形での制約

$$\boldsymbol{c} - A^T\boldsymbol{y} - \boldsymbol{\lambda} = 0, \boldsymbol{\lambda} \geq \boldsymbol{0} \quad \leftrightarrow \quad A^T\boldsymbol{y} \leq \boldsymbol{c}$$

に書き換えることにより、双対問題は

等式標準形の双対問題

$$\left|\begin{array}{ll} \text{maximize} & \boldsymbol{b}^T\boldsymbol{y} \\ \text{s.t.} & A^T\boldsymbol{y} \leq \boldsymbol{c} \end{array}\right. \tag{P22}$$

と記述できる。

定理 6.1 [P15] の双対問題の双対問題は主問題 [P15] に一致する。

(proof)

双対問題 [P22] を等式標準形に変形し、その双対問題を考える。同値な問題 [P21] において、$\boldsymbol{y} = \boldsymbol{y}^+ - \boldsymbol{y}^-, \boldsymbol{y}^\pm \geq \boldsymbol{0}$ とすることによって、[P22] は

$$
\left|\begin{array}{ll}
\text{minimize} & -\begin{pmatrix} \boldsymbol{b}^T & -\boldsymbol{b}^T & \boldsymbol{0}^T \end{pmatrix} \begin{pmatrix} \boldsymbol{y}^+ \\ \boldsymbol{y}^- \\ \boldsymbol{\lambda} \end{pmatrix} \\
\text{s.t.} & \begin{pmatrix} A^T & -A^T & E^T \end{pmatrix} \begin{pmatrix} \boldsymbol{y}^+ \\ \boldsymbol{y}^- \\ \boldsymbol{\lambda} \end{pmatrix} = \boldsymbol{c} \\
& \begin{pmatrix} \boldsymbol{y}^+ \\ \boldsymbol{y}^- \\ \boldsymbol{\lambda} \end{pmatrix} \geq \boldsymbol{0}
\end{array}\right.
$$

という、等式標準形に変形できる。この双対問題は

$$
\left|\begin{array}{ll}
\text{maximize} & \boldsymbol{c}^T \boldsymbol{x}' \\
\text{s.t.} & \begin{pmatrix} A \\ -A \\ E \end{pmatrix} \boldsymbol{x}' \leq -\begin{pmatrix} \boldsymbol{b} \\ -\boldsymbol{b} \\ \boldsymbol{0} \end{pmatrix}
\end{array}\right.
$$

であるが、$\boldsymbol{x} = -\boldsymbol{x}'$ として、整理すると

$$
\left|\begin{array}{ll}
\text{maximize} & -\boldsymbol{c}^T \boldsymbol{x} \\
\text{s.t.} & A\boldsymbol{x} \geq \boldsymbol{b} \\
& A\boldsymbol{x} \leq \boldsymbol{b} \\
& \boldsymbol{x} \geq \boldsymbol{0}
\end{array}\right.
$$

であり、これは主問題と同値である。 $\boxed{\text{証明終}}$

この定理から、不等式形 [P13]

$$
\left|\begin{array}{ll}
\text{minimize} & \boldsymbol{c}^T \boldsymbol{x} \\
\text{s.t.} & A\boldsymbol{x} \geq \boldsymbol{b}
\end{array}\right.
$$

の双対問題も得られる。つまり、不等式形の問題は

$$
\left|\begin{array}{ll}
\text{maximize} & \boldsymbol{c}^T (-\boldsymbol{x}) \\
\text{s.t.} & A(-\boldsymbol{x}) \leq -\boldsymbol{b}
\end{array}\right.
$$

と同値であり、これは等式標準形の双対問題なので、その双対問題は等式標準形の形

不等式形の双対問題

$$
\left|\begin{array}{ll}
\text{maximize} & \boldsymbol{b}^T \boldsymbol{y} \\
\text{s.t.} & A^T \boldsymbol{y} = \boldsymbol{c} \\
& \boldsymbol{y} \geq \boldsymbol{0}
\end{array}\right. \qquad \text{[P23]}
$$

で与えられる。

不等式標準形 [P14]

$$
\left|\begin{array}{ll}
\text{minimize} & \boldsymbol{c}^T \boldsymbol{x} \\
\text{s.t.} & A\boldsymbol{x} \geq \boldsymbol{b} \\
& \boldsymbol{x} \geq \boldsymbol{0}
\end{array}\right.
$$

に対しては、スラック変数 $\boldsymbol{s}$ を導入することで、等式標準形

$$
\left|\begin{array}{ll}
\text{minimize} & \begin{pmatrix} \boldsymbol{c}^T & \boldsymbol{0}^T \end{pmatrix} \begin{pmatrix} \boldsymbol{x} \\ \boldsymbol{s} \end{pmatrix} \\
\text{s.t.} & \begin{pmatrix} A & -E \end{pmatrix} \begin{pmatrix} \boldsymbol{x} \\ \boldsymbol{s} \end{pmatrix} = \boldsymbol{b} \\
& \begin{pmatrix} \boldsymbol{x} \\ \boldsymbol{s} \end{pmatrix} \geq \boldsymbol{0}
\end{array}\right.
$$

に変形でき、この双対をとることにより

不等式標準形の双対問題

$$\begin{array}{ll} \text{maximize} & \boldsymbol{b}^T\boldsymbol{y} \\ \text{s.t.} & A^T\boldsymbol{y} \leq \boldsymbol{c} \\ & \boldsymbol{y} \geq \boldsymbol{0} \end{array} \qquad \text{[P24]}$$

となる。この双対問題が 主問題に一致 することは、ほぼ自明であろう。

線形計画問題における双対問題についても、当然ながら、非線形の双対問題に関する命題が成立する。

6.2 弱双対定理

線形計画問題においても、弱双対定理 3.9 及び 3.10 が成立している。また、線形計画では、特に次のことが言える。

定理 6.2 線形計画問題において、主問題または双対問題が非有界ならば、他方は実行不可能である。

(proof)
定理 3.11 より、片方が非有界ならば、他方の許容解は無限の目的関数値しかとらないが、線形計画問題においては、$\boldsymbol{R}^{\boldsymbol{n}}$ の許容解に対しては、目的関数値も $\boldsymbol{R}$ の元であり、無限の目的関数値を取らない。よって、許容解が存在しない、つまり、実行不可能だということになる。 証明終

6.3 強双対定理

定理 6.3 （強双対定理）線形計画問題において、主問題または双対問題が最適解を持てば、他方も最適解を持ち、それぞれの最適値は一致する。

(proof)
等式標準形 [P15] で考える。$\boldsymbol{x}^*$ が主問題の最適解ならば、定理 5.3 から KKT 条件を満たすので、定理 5.4 より

$$\begin{aligned} \boldsymbol{c}^T\boldsymbol{x}^* &= \boldsymbol{b}^T\boldsymbol{y} \\ A^T\boldsymbol{y} &\leq \boldsymbol{c} \\ A\boldsymbol{x}^* &= \boldsymbol{b} \\ \boldsymbol{x}^* &\geq \boldsymbol{0} \end{aligned}$$

を満たす $\boldsymbol{y} \in \boldsymbol{R}^{\boldsymbol{n}}$ が存在する。このとき $\boldsymbol{y}$ は双対問題 [P22] の許容解であり、したがって、$\boldsymbol{c}^T\boldsymbol{x}^* = \boldsymbol{b}^T\boldsymbol{y}$ から系 3.8 より双対問題の最適解である。よって、主問題が最適解を持つ場合について示された。双対問題が最適解を持つ場合についても、双対問題の双対問題が主問題なので、示されている。 証明終

(proof)
（Farkas の補題を直接用いる別証明）不等式標準形 [P14] で考える。主問題の最適解 $\boldsymbol{x}^*$ に対して

$$\begin{aligned} \boldsymbol{b}^T\boldsymbol{y} - v = \boldsymbol{c}^T\boldsymbol{x}^*, &\quad v \geq 0 \\ A^T\boldsymbol{y} + \boldsymbol{s} = \boldsymbol{c}, &\quad \boldsymbol{s} \geq \boldsymbol{0} \\ \boldsymbol{y} \leq \boldsymbol{0} & \end{aligned}$$

なる $\boldsymbol{y}$ が存在するとする。このとき $\boldsymbol{y}$ は双対問題の許容解であり、さらに弱双対定理及び系 3.8 より $\boldsymbol{y}$ は双対問題の最適解である。わざわざ $\boldsymbol{c}^T\boldsymbol{x}^* = \boldsymbol{b}^T\boldsymbol{y}$ を不等式で表して、スラック変数 v を導入したのは、Farkas の補題を用いたとき制約式が増えるからである。ここで、上の条件を満たす $\boldsymbol{y}$ が存在しないとする。このとき Farkas の補題 2.14 より

$$\begin{aligned} \boldsymbol{c}^T\boldsymbol{u} &< -w(\boldsymbol{c}^T\boldsymbol{x}^*) \\ A\boldsymbol{u} &\geq -w\boldsymbol{b} \\ \boldsymbol{u} &\geq \boldsymbol{0} \\ w &\geq 0 \end{aligned}$$

となる $\boldsymbol{u} \in \boldsymbol{R^n}, w \in \boldsymbol{R}$ が存在する。$w = 0$ とすると

$$\begin{aligned} \boldsymbol{c}^T\boldsymbol{u} &< 0 \\ A\boldsymbol{u} &\geq \boldsymbol{0} \\ \boldsymbol{u} &\geq \boldsymbol{0} \end{aligned}$$

より

$$\begin{aligned} \boldsymbol{c}^T(\boldsymbol{x}^* + \boldsymbol{u}) &< \boldsymbol{c}^T\boldsymbol{x}^* \\ A(\boldsymbol{x}^* + \boldsymbol{u}) &\geq \boldsymbol{b} \\ (\boldsymbol{x}^* + \boldsymbol{u}) &\geq \boldsymbol{0} \end{aligned}$$

であるから、$\boldsymbol{x}^*$ が最適解であることに矛盾する。したがって、$w < 0$ である。このとき

$$\begin{aligned} \boldsymbol{c}^T\frac{\boldsymbol{u}}{-w} &< \boldsymbol{c}^T\boldsymbol{x}^* \\ A\frac{\boldsymbol{u}}{-w} &\geq \boldsymbol{b} \\ \frac{\boldsymbol{u}}{-w} &\geq \boldsymbol{0} \end{aligned}$$

であるが、これはやはり $\boldsymbol{x}^*$ が最適解であることに矛盾する。よって、条件を満たす $\boldsymbol{y}$ が存在し、主問題に最適解が存在すれば、双対問題にも最適解が存在し、その最適解は一致する。 $\boxed{\text{証明終}}$

定理 6.4 線形計画問題において、KKT 条件は、最適解であることの必要十分条件である。

(proof)
　等式標準形 [P15] で考える。必要条件であることは定理 5.3 によって示されている。逆に KKT 条件を満たすとき、定理 5.4 より

$$\begin{aligned} \boldsymbol{c}^T\boldsymbol{x} &= \boldsymbol{b}^T\boldsymbol{y} \\ A^T\boldsymbol{y} &\leq \boldsymbol{c} \\ A\boldsymbol{x} &= \boldsymbol{b} \\ \boldsymbol{x} &\geq \boldsymbol{0} \end{aligned}$$

を満たす $\boldsymbol{y} \in \boldsymbol{R^n}$ が存在する。このとき $\boldsymbol{x}$ は主問題の許容解であり、$\boldsymbol{y}$ は双対問題 [P22] の許容解である。したがって、$\boldsymbol{c}^T\boldsymbol{x} = \boldsymbol{b}^T\boldsymbol{y}$ から系 3.8 より $\boldsymbol{x}, \boldsymbol{y}$ はそれぞれ主問題・許容解の最適解であり、確かに十分条件でもある。 $\boxed{\text{証明終}}$

　KKT 条件は、「**主許容性＋双対許容性＋最適値の一致**」となっている。任意の線形計画問題は等式標準形に変形できるので、等式標準形以外の形でも、これは変わらない。

定理 6.5 （相補性定理）不等式標準形の線形計画問題 [P15] において、主問題・双対問題の許容解 $\boldsymbol{x}, \boldsymbol{y}$ がいずれも最適解である必要十分条件は

$$\boldsymbol{x}^T(A^T\boldsymbol{y} - \boldsymbol{c}) = 0 \tag{44}$$

$$\boldsymbol{y}^T(A\boldsymbol{x} - \boldsymbol{b}) = 0 \tag{45}$$

が成立することである。

(proof)

定理 6.4 より KKT 条件が最適解である必要十分条件であり、その条件は定理 5.5 にある通りだが、すでに許容性は成立しているので

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

が成立すればよい。これについては

$$\begin{aligned} \boldsymbol{c}^T\boldsymbol{x} &\geq (A^T\boldsymbol{y})^T\boldsymbol{x} \\ &= \boldsymbol{y}^T A\boldsymbol{x} \\ &\geq \boldsymbol{y}^T\boldsymbol{b} = \boldsymbol{b}^T\boldsymbol{y} \end{aligned}$$

であるが、定理の条件が成り立つことは、上の変形において不等号が等号として成立することであり、$\boldsymbol{c}^T\boldsymbol{x} = \boldsymbol{b}^T\boldsymbol{y}$ が成り立つことと同値である。よって示された。 $\boxed{\text{証明終}}$

定理 6.5 の条件は**相補性**と呼ばれる。等式標準形に対しては、以下のようになる。

定理 6.6 （相補性定理 *2*）等式標準形の線形計画問題 [P14] において、主問題・双対問題の許容解 $\boldsymbol{x}, \boldsymbol{y}$ がいずれも最適解である必要十分条件は

$$\boldsymbol{x}^T(A^T\boldsymbol{y} - \boldsymbol{c}) = 0 \tag{46}$$

が成立することである。

6.4 解のパターン

定理 6.2 及び強双対定理 6.3 と、主問題・双対問題いずれも実行不可能な問題が実際に構成できること（容易）により、主問題と双対問題の解のパターンは、下の表のように限られる。◯ が可能性の存在を示し、× がありえないことを示す。

	双対実行不可能	双対非有界	双対最適解が存在
主実行不可能	◯	◯	×
主非有界	◯	×	×
主最適解が存在	×	×	◯

上の表と強双対定理 6.3 より、次のことも成立する。

定理 6.7 （強双対定理 *2*）線形計画問題において、主問題と双対問題の両方とも実行可能ならば、主最適解・双対最適解が存在し、最適値は一致する。

6.5 双対問題の例

6.5.1 ゼロ和 2 人ゲーム

プレイヤー p と q がおり、プレイヤー p が戦略 i、プレイヤー q が戦略 j をとった場合の利得を $a_{ij}(\{A\}_{ij} = a_{ij})$ とし、プレイヤー p はその最小化、プレイヤー q はその最大化を目論んでいるとする。混合戦略を考え、プレイヤー p が戦略 i をとる確率を $p_i(\{\boldsymbol{p}\}_i = p_i)$、プレイヤー q が戦略 j をとる確率を $q_j(\{\boldsymbol{q}\}_j = q_j)$ とする。このとき、利得の期待値は

$$f(\boldsymbol{p}, \boldsymbol{q}) = \boldsymbol{p}^T A \boldsymbol{q}$$

となる。また $a_{ij} \geq 0$ と仮定する。これは、全利得に同じ定数を足してもゲームは本質的に変わらない（利得の期待値も定数が付加されるだけ）ことを鑑みると、一般性を失わない。

プレイヤー p が、最悪でも得られる利得を最適化（最小化）するという $\min - \max$ 戦略を採用しているとする。これを考えるには、まず混合戦略 $\boldsymbol{p}$ に対する、最悪な利得の期待値を計算する必要がある。$A = (\boldsymbol{a}_1, \cdots, \boldsymbol{a}_m)$ とおいて、最適化問題として記述すると以下のようになる。

$$\left|\begin{array}{ll} \text{maximize} & \boldsymbol{p}^T A \boldsymbol{q} = \displaystyle\sum_{j=1}^{m} (\boldsymbol{a}_j^T \boldsymbol{p}) q_j \\ \text{s.t.} & \boldsymbol{1}^T \boldsymbol{q} \left(= \displaystyle\sum_{j=1}^{m} q_j \right) = 1 \\ & \boldsymbol{q} \geq \boldsymbol{0} \\ & \boldsymbol{p} \text{ は定数とみなす} \end{array}\right.$$

明らかにこの最適解は $\boldsymbol{a}_j^T \boldsymbol{p}$ が最大となる戦略 j に全確率 1 を配分する混合戦略であり、そのときの最適値は

$$\max_j (\boldsymbol{a}_j^T \boldsymbol{p}) = \|A^T \boldsymbol{p}\|_\infty$$

である。よって、プレイヤー p が $\min - \max$ 戦略を得るための最適化問題 は

$$\left|\begin{array}{ll} \text{minimize} & \|A^T \boldsymbol{p}\|_\infty \\ \text{s.t.} & \boldsymbol{1}^T \boldsymbol{p} = 1 \\ & \boldsymbol{p} \geq \boldsymbol{0} \end{array}\right. \qquad \text{[P25]}$$

と記述できる。これは、l_∞ ノルム最小化問題なので、定理 5.6 より

$$\left|\begin{array}{ll} \text{minimize} & t \\ \text{s.t.} & -t\boldsymbol{1} \leq A^T \boldsymbol{p} \leq t\boldsymbol{1} \\ & \boldsymbol{1}^T \boldsymbol{p} = 1 \\ & \boldsymbol{p} \geq \boldsymbol{0} \end{array}\right.$$

という線形計画問題に書き換えられる。この**双対問題**をとり、整理すると

$$\left|\begin{array}{ll} \text{maximize} & \min_i (A\boldsymbol{q} - A\boldsymbol{q}')_i \\ \text{s.t.} & \boldsymbol{1}^T \boldsymbol{q} + \boldsymbol{1}^T \boldsymbol{q}' = 1 \\ & \boldsymbol{q}, \boldsymbol{q}' \geq \boldsymbol{0} \end{array}\right.$$

という形になる。ところが、仮定より $A\boldsymbol{q}' \geq \boldsymbol{0}$ なので

$$A\boldsymbol{q} - A\boldsymbol{q}' \leq A\boldsymbol{q} - A\boldsymbol{0}$$

であり、この問題は最大化問題なので $\boldsymbol{q}' = \boldsymbol{0}$ という制約を付加しても、最適解が等しい問題が得られる。そのときの問題は

$$\left|\begin{array}{ll} \text{maximize} & \min_i (A\boldsymbol{q})_i \\ \text{s.t.} & \boldsymbol{1}^T\boldsymbol{q} = 1 \\ & \boldsymbol{q} \geq \boldsymbol{0} \end{array}\right. \qquad \text{[P26]}$$

となる。これは、前と同様の考え方により、プレイヤー q の $\max - \min$ 戦略を得るための問題である。そして、双対問題に制約が付加された問題であるから、許容集合は小さくなっているが、最適解が変化しないので、強双対定理は変わらず成立する。この場合には、いずれの問題にも許容解は必ず存在するので、強双対定理 6.7 の内容は、「**プレイヤー** p **の** $\min - \max$ **戦略・プレイヤー** q **の** $\max - \min$ **戦略のいずれも最適解を持ち、最適値は一致する**。」ということになる。このときの最適解である混合戦略は、**最適混合戦略**といわれる。

第 IV 部
付録

A 上限・下限及び実数の連続性

A.1 実数の上限・下限

定義 A.1 実数の上に部分集合 X が与えられたとき、実数 a が以下の条件

1. $b \in X$ なら $b \leq a$
2. $c < a$ なら $c < x$ なる $x \in X$ が存在する

を満たせば、a を X の**上限**であるといい、$a = \sup X$ と表す。

定義 A.2 実数の上に部分集合 X が与えられたとき、実数 a が以下の条件

1. $b \in X$ なら $b \geq a$
2. $a < c$ なら $x < c$ なる $x \in X$ が存在する

を満たせば、a を X の**下限**であるといい、$a = \inf X$ と表す。

定理 A.1 上限及び下限は一意的である。

(proof)
容易。証明略。 $\boxed{\text{証明終}}$

命題 A.1 実数に関しては、上に有界な部分集合は必ず上限を持ち、下に有界な部分集合は必ず下限を持つ。このことを、実数は**順序完備**であるという。これは、実数の公理のひとつである。[10]

あるいは、有界でない場合は $\pm\infty$ であると定義することにより、常に上限下限の値を考えることもできる。また、実数の部分集合 X に対し、上限 $\sup X$ 及び下限 $\inf X$ は、必ずしも X には属さないことには注意が必要である。それに対し、最大値・最小値は存在するとは限らない。上限 $\sup X$ 及び下限 $\inf X$ が X に属せば、それらはそれぞれ最大値・最小値である。上限及び下限は一意的なので、最大値・最小値が存在すれば、それらはまた上限・下限に等しい。

A.2 数列の上限・下限

数列に対しても、同じ値を取る項を同一視し、数列を実数の集合とみなすことで、上限・下限を定義できる。

定理 A.2 実数列 $\{a_n\}, \{b_n\}$ について すべての n について $a_n \leq b_n$ なら

$$\sup_n a_n \leq \sup_n b_n \quad \inf_n a_n \leq \inf_n b_n$$

[10] 上限下限は、より一般には、順序集合に対して定義できる。また、順序完備な順序体を実数体といい、その元を実数という。

(proof)
上限について。$\sup_n a_n > \sup_n b_n$ と仮定すると、上限の定義より $\sup_n b_n < a_m$ となる数列 $\{a_n\}$ の項 a_m が存在する。このとき、条件より $\sup_n b_n < a_m \leq b_m$ となることになり、上限 $\sup_n b_n$ の定義に矛盾する。したがって $\sup_n a_n \leq \sup_n b_n$ である。下限についても同様。 $\boxed{\text{証明終}}$

定理 A.3 上に有界な単調増加数列 $\{a_n\}$ は収束し $\displaystyle\lim_{n\to\infty} a_n = \sup_n a_n$

(proof)
まず、上に有界ならば、$\sup_n a_n$ が存在する (順序完備性)。

上限の定義よりすべての n について $a_n \leq \sup_n a_n$ である。また、任意の正の実数 ε について、当然 $\sup_n a_n - \varepsilon < \sup_n a_n$ なので $\sup_n a_n - \varepsilon < a_m(\varepsilon)$ となる項 $a_m(\varepsilon)$ が存在する。このとき、数列 $\{a_n\}$ は単調増加なので、$n \geq m(\varepsilon)$ であれば

$$\sup_n a_n - \varepsilon < a_m(\varepsilon) \leq a_n \leq \sup_n a_n < \sup_n a_n + \varepsilon$$

つまり $|a_n - \sup_n a_n| < \varepsilon$ となる。ε は任意だったので

$$\lim_{n\to\infty} a_n = \sup_n a_n$$

である。 $\boxed{\text{証明終}}$

定理 A.4 下に有界な単調減少数列 $\{a_n\}$ は収束し $\displaystyle\lim_{n\to\infty} a_n = \inf_n a_n$

(proof)
上と同様。 $\boxed{\text{証明終}}$

A.2.1 数列の上極限・下極限

上に有界な実数列 $\{a_n\}$ に対し、その始めの m 項を除いた数列 $\{a_{m+1}, a_{m+2}, \cdots\}$ の下限を以て新しい数列 $\{b_m\}$ を定義する、すなわち

$$b_m \equiv \inf_{n>m} a_n$$

とする。m が大きくなれば、下限を取る対象の数列は小さくなっていくので、数列 $\{b_m\}$ は単調増加である。また、$\{a_n\}$ は上に有界なので、$\{b_m\}$ も上に有界である。したがって、定理 A.3(単調有界) より、$\{b_m\}$ は収束する。ここで、以下の定義を置く。

定義 A.3 実数列 $\{a_n\}$ に対し、その**下極限**を

$$\underline{\lim}\, a_n \equiv \begin{cases} \displaystyle\lim_{m\to\infty} \inf_{n>m} a_n \quad (= \sup_m \inf_{n>m} a_n) & \{a_n\} \text{ が下に有界であるとき} \\ -\infty & \{a_n\} \text{ が下に有界でないとき} \end{cases}$$

と定義する。上の議論より、下極限は必ず存在する。

同様にして、上極限を定義する。

定義 A.4 実数列 $\{a_n\}$ に対し、その**上極限**を

$$\overline{\lim}\, a_n \equiv \begin{cases} \lim_{m\to\infty} \sup_{n>m} a_n \quad (= \inf_m \sup_{n>m} a_n) & \{a_n\} \text{ が上に有界であるとき} \\ \infty & \{a_n\} \text{ が上に有界でないとき} \end{cases}$$

と定義する。定理 *A.4* から上極限は必ず存在する。

定理 A.5 $\underline{\lim}\, a_n \leq \overline{\lim}\, a_n$

(proof)
定義から明らか。 $\boxed{\text{証明終}}$

定理 A.6 $\underline{\lim}\, a_n = \overline{\lim}\, a_n$ であることと、$\{a_n\}$ が確定する、つまり、収束するか $\pm\infty$ に発散することは同値であり、このとき $\lim_{n\to\infty} a_n = \underline{\lim}\, a_n = \overline{\lim}\, a_n$ である。

(proof)
必要性から。
(i)a_n が上下に有界であるとき。$\underline{\lim}\, a_n = \overline{\lim}\, a_n = a$ とおく。上限・下限の定義より $n > m$ なら

$$\inf_{n>m} a_n \leq a_n \leq \sup_{n>m} a_n$$

である。$m \to \infty$ の極限をとれば $n \to \infty$ であり、挟撃により

$$\lim_{n\to\infty} a_n = a$$

となる。
(ii)a_n が上に有界でなく、下に有界なとき。まず、$\overline{\lim}\, a_n = \infty = \underline{\lim}\, a_n$ である。下限の定義より $n > m$ なら $\inf_{n>m} a_n \leq a_n$ である。したがって、このとき $m \to \infty$ の極限をとれば $n \to \infty$ でもあり

$$\infty = \overline{\lim}\, a_n \leq \lim_{n\to\infty} a_n$$

つまり

$$\lim_{n\to\infty} a_n = \infty$$

となる。
(iii)a_n が下に有界でなく、上に有界なとき。まず、$\underline{\lim}\, a_n = -\infty = \underline{\lim}\, a_n$ である。上限の定義より $n > m$ なら $a_n \leq \sup_{n>m} a_n$ である。したがって、このとき $m \to \infty$ の極限をとれば $n \to \infty$ でもあり

$$\lim_{n\to\infty} a_n \leq \overline{\lim}\, a_n = -\infty$$

つまり

$$\lim_{n\to\infty} a_n = -\infty$$

となる。
(iiii) 上下いずれにも有界でなければ、上極限・下極限は一致しないので、その場合は考察の必要がない。

次は、十分性について。
(i) $\lim_{n\to\infty} a_n = a$ と収束するとき。任意の正の実数 ε に対して、$n \geq n_0(\varepsilon)$ のとき

$$a - \varepsilon < a_n < a + \varepsilon$$

となる。したがって、$m > n_0(\varepsilon)$ のとき

$$a - \varepsilon \leq \sup_{n>m} a_n \leq \inf_{n>m} a_n \leq a + \varepsilon$$

である。[11] これは

$$\lim_{m\to\infty} \sup_{n>m} a_n = \lim_{m\to\infty} \inf_{n>m} a_n = a$$

をあらわしている。つまり

$$\underline{\lim}\, a_n = \overline{\lim}\, a_n = a$$

である。

(ii) $\lim\limits_{n\to\infty} a_n = \infty$ と発散するとき。まず、上に有界ではないので、上極限の定義より

$$\overline{\lim}\, a_n = \infty$$

である。また、n を十分大きくとれば $a_n = \infty$ (どの実数よりも大きくできる) なので m が同じくらい十分大きければ $\inf_{n>m} a_n = \infty$ である。したがって

$$\underline{\lim}\, a_n = \lim_{m\to\infty} \inf_{n>m} a_n = \infty = \overline{\lim}\, a_n$$

である。

(iii) $\lim\limits_{n\to\infty} a_n = -\infty$ と発散するとき。まず、下に有界ではないので、上極限の定義より

$$\underline{\lim}\, a_n = -\infty$$

である。また、n を十分大きくとれば $a_n = -\infty$ (どの実数よりも小さくできる) なので m が同じくらい十分大きければ $\sup_{n>m} a_n = -\infty$ である。したがって

$$\overline{\lim}\, a_n = \lim_{m\to\infty} \inf_{n>m} a_n = -\infty = \underline{\lim}\, a_n$$

である。 証明終

定理 A.7 任意の n で $a_n \leq b_n$ なら $\overline{\lim}\, a_n \leq \overline{\lim}\, b_n \quad \underline{\lim}\, a_n \leq \underline{\lim}\, b_n$

(proof)

定理 A.2 より

$$\sup_{n>m} a_n \leq \sup_{n>m} b_n \qquad \inf_{n>m} a_n \leq \inf_{n>m} b_n$$

したがって、$m \to \infty$ とすれば

$$\overline{\lim}\, a_n \leq \overline{\lim}\, b_n \qquad \underline{\lim}\, a_n \leq \underline{\lim}\, b_n$$

となる。 証明終

A.3 実数の連続性

順序完備性 (命題 A.1) と同値な命題群[12] は**完備性・実数の連続性**と呼ばれ、実数の重要な性質となっている。定理 A.3 や定理 A.4（有界単調収束）も、その一つである。ここでは、順序完備性を公理として考え、そこから他の完備性を諸定理として導く。逆は、他書を参照されたい。

[11] 上限・下限は数列に属する値とは限らないので等号が付加されているが、本質的ではない。
[12] より一般には、順序体に関する命題群である。

定理 A.8 実数体 $\boldsymbol{R}$ はコーシー完備[13] である。

(proof)
任意のコーシー列 $\{a_n\}$ について、任意の $\varepsilon > 0$ に対してある自然数 L が存在して

$$\forall n, m \geq L \quad \to \quad |a_n - a_m| < \varepsilon \tag{47}$$

が成立する。

$n \geq L$ であるとき

$$a_L - \varepsilon < a_n < a_L + \varepsilon$$

であるから、任意の自然数 n について

$$\min(a_1, \cdots, a_{L-1}, a_L - \varepsilon) \leq a_n \leq \max(a_1, \cdots, a_{L-1}, a_L + \varepsilon)$$

である。つまり、a_n は有界である。よって、定義より、上極限・下極限は有限である。

また、式 (47) より、$L < L_1$ なる L_1 について

$$\begin{aligned}
-\varepsilon &< a_n - a_m < \varepsilon \\
-\varepsilon &\leq \sup_{n > L_1} (a_n - a_m) \\
&= \sup_{n > L_1} a_n - a_m \leq \varepsilon \quad \because \text{定理 } A.2 \\
-\varepsilon &\leq a_m - \sup_{n > L_1} a_n \leq \varepsilon \\
-\varepsilon &\leq \inf_{m > L_1} a_m - \sup_{n > L_1} a_n \leq \varepsilon \quad \because \text{定理 } A.2
\end{aligned}$$

であり、$L_1 \to \infty$ とすることにより

$$\left|\underline{\lim}\, a_n - \overline{\lim}\, a_n\right| \leq \varepsilon$$

であるが、$\left|\underline{\lim}\, a_n - \overline{\lim}\, a_n\right|$ は有限かつ定数で、$\varepsilon > 0$ は任意なので、$0 < \left|\underline{\lim}\, a_n - \overline{\lim}\, a_n\right|$ と仮定すると、$\varepsilon' < \left|\underline{\lim}\, a_n - \overline{\lim}\, a_n\right|$ なる ε' がとれて矛盾する。よって

$$\underline{\lim}\, a_n = \overline{\lim}\, a_n (= \text{有限})$$

である。これより定理 A.6 から $\{a_n\}$ は収束する。 証明終

定理 A.9 （ボルツァーノ・ワイエルシュトラスの性質）任意の有界無限実数列から収束部分列をとることができる。

(proof)
任意の有界無限実数列 $\{a_n\}$ から、実際に収束部分列 $\{a_{\varphi(n)}\}$ を構成する。

有界なので $\forall n$ について $a_n \in [-b, b]$ が成立する。ここで、$\varphi(0) = 0, I_0 = [-b, b]$ と置く。そして、以下の手順を繰り返す。

> 区間 I_{k-1} を二等分した時、少なくとも片方には無限の $\{a_n\}$ の項が属しているので、それを I_k と置く。
>
> $$S_k = \{\text{ 自然数 } n; a_n \in I_k, n > \varphi(k-1)\}$$

[13] さらにアルキメデスの公理を満たすことが、順序完備との必要十分条件である。

を考えるとき、$S_k = \phi$ とすると $a_n \in I_k$ なる項は $n \leq \varphi(k-1)$ となることになり、I_k に無限の $\{a_n\}$ の項が属していることに矛盾する。よって、$S_k \neq \phi$ であり、また S_k の元は 1 以上なので、S_k には最小値が存在する。その S_k の最小値を $\varphi(k)$ とする。

こうして構成した $\varphi(n), I_n$ については

$$\varphi(0) < \varphi(1) < \varphi(2) < \cdots$$

$$a_{\varphi(n)} \in I_n \supset I_{n+1} \supset \cdots$$

が成立する。よって $\forall n, m \geq L$ について

$$a_{\varphi(n)}, a_{\varphi(m)} \in I_L$$

である。ところで、区間 I_L の幅は $\left(\frac{1}{2}\right)^L (2b)$ なので

$$|a_{\varphi(n)} - a_{\varphi(m)}| \leq \left(\frac{1}{2}\right)^L (2b)$$

となる。よって、任意の $\epsilon > 0$ に対して $\left(\frac{1}{2}\right)^L (2b) < \epsilon$ となるよう L をとれば

$$|a_{\varphi(n)} - a_{\varphi(m)}| \leq \epsilon$$

となるのだから、$\{a_{\varphi(n)}\}$ はコーシー列である。よって、定理 A.8 より、$\{a_{\varphi(n)}\}$ は収束する。 証明終

B　位相・距離空間

B.1　位相空間と開集合・閉集合

定義 B.1 集合 X に、以下の性質を持つ部分集合族 $\mathfrak{O}$(**開集合族**) が与えられたとき、$(X, \mathfrak{O})$ を**位相空間**といい、$\mathfrak{O}$ に属す集合を**開集合**という。

1. $O_\gamma \in \mathfrak{O}(\gamma \in \Gamma)$ のとき　$\displaystyle\bigcup_{\gamma\in\Gamma} O_\gamma \in \mathfrak{O}$ [14]
2. $O_1, O_2 \in \mathfrak{O}(\gamma \in \Gamma)$ のとき　$O_1 \cap O_2 \in \mathfrak{O}$ [15]
3. $X, \phi \in \mathfrak{O}$

定義 B.2 位相空間 $(X, \mathfrak{O})$ において、X の部分集合 F について $F^c \in \mathfrak{O}$ であるとき、F を**閉集合**という。

定理 B.1 位相空間 $(X, \mathfrak{O})$ において、閉集合について以下が成り立つ。

1. $\gamma \in \Gamma$ なる F_γ がすべて閉集合のとき　$\displaystyle\bigcap_{\gamma\in\Gamma} O_\gamma$ も閉集合である。
2. F_1, F_2 が閉集合のとき　$O_1 \cup O_2$ も閉集合である。
3. X, ϕ は閉集合である。

(proof)
閉集合の補集合が開集合であることから、開集合の定義に適用すればよい。詳細略。 証明終

B.1.1　位相空間の直積

定理 B.2 位相空間 $(X, \mathfrak{O}_X)$,$(Y, \mathfrak{O}_Y)$ について

$$\mathfrak{O}_X \times \mathfrak{O}_Y \equiv \{A \times B | A \in O_X, B \in O_Y\}$$

と定義すると、$\mathfrak{O}_X \times \mathfrak{O}_Y$ は開集合族の条件を満たし、$(X \times Y, \mathfrak{O}_X \times \mathfrak{O}_Y)$ は位相空間となる。

(proof)
証明略。 証明終

定義 B.3 位相空間 $(X, \mathfrak{O}_X)$,$(Y, \mathfrak{O}_Y)$ について、上の定理より位相空間となる $(X \times Y, \mathfrak{O}_X \times \mathfrak{O}_Y)$ を、**積位相空間**または単純に直積と言う。

[14]「開集合の和集合は開集合になる」ことを要請している。
[15]「二つの開集合の共通部分は開集合になる」ことを要請している。

B.2　コンパクト

定義 B.4 位相空間 $(X, \mathfrak{O})$ と X の部分集合 Y において、開集合の族 $\{O_\gamma\}_{\gamma\in\Gamma}$ が

$$Y \subset \bigcup_{\gamma\in\Gamma} O_\gamma$$

をみたすとき、$\{O_\gamma\}_{\gamma\in\Gamma}$ を Y の**開被覆**であるという。[16]特に、有限個の開集合の族 $\{O_n\}$ によって

$$Y \subset \bigcup_{n=1}^{N} O_n$$

と表されるとき、$\{O_n\}$ は**有限開被覆**であるという。

定義 B.5 位相空間 $(X, \mathfrak{O})$ と X の部分集合 Y において、$\{O_\gamma\}_{\gamma\in\Gamma}$ が Y の開被覆ならば、つまり

$$Y \subset \bigcup_{\gamma\in\Gamma} O_\gamma$$

ならば、$\{O_\gamma\}_{\gamma\in\Gamma}$ からとった有限個の部分集合族 $O_1, O_2, \cdots, O_s$ が既に Y の有限開被覆である、つまり

$$Y \subset \bigcup_{n=1}^{s} O_n$$

であるとき、Y は**コンパクト**であるという。

定理 B.3 コンパクトな位相空間の部分閉集合はコンパクトである。

(proof)
X をコンパクトな位相空間とし、その部分閉集合を F とおく。$\{O_\gamma\}_{\gamma\in\Gamma}$ を F の開被覆とする。

$$X = F \cup F^c \subset \bigcup_{\gamma\in\Gamma} O_\gamma \cup F^c$$

であり F^c は開集合なので、$\{O_\gamma\}_{\gamma\in\Gamma} \cup F^c$ は X の開被覆である。X はコンパクトなので、この中から X の有限部分開被覆 $\{O_1, \cdots, O_s, F^c\}$ をとれる。このとき

$$X \subset \bigcup_{n=1}^{s} O_s \cup F^c$$

となる。これより、X の要素で $\bigcup_{n=1}^{s} O_s$ に含まれない要素は F^c に含まれるということがいえる。この対偶をとれば、F の要素は $\bigcup_{n=1}^{s} O_s$ に含まれるということになり、

$$F \subset \bigcup_{n=1}^{s} O_s$$

であるから、F は開被覆の有限部分開被覆によって覆われることになる。つまり、F はコンパクトである。
証明終

定理 B.4 位相空間 $(X, \mathfrak{O}_X)$,$(Y, \mathfrak{O}_Y)$ がコンパクトである時、積位相空間 $(X \times Y, \mathfrak{O}_X \times \mathfrak{O}_Y)$ もコンパクトである。

[16]包含関係ではなく"="で定義する流儀もあるが、こちらの方が使いやすい。

(proof)

$X\times Y$ の開被覆 $\{O_\gamma\}_{\gamma\in\Gamma}$ が与えられた時、そのうちの有限個によって

$$X\times Y\subset\bigcup_{n=1}^{N}O_n$$

となればよいのだが、そのままでは扱いにくいので[17]、開被覆の部分開集合も含めることにより

$$\{O'_\gamma\}_{\gamma\in\Gamma'}=\{O'\in\mathfrak{O}_X\times\mathfrak{O}_Y|O'\subset O_\gamma(\gamma\in\Gamma)\}$$

と拡張した集合族 $\{O'_\gamma\}_{\gamma\in\Gamma'}$ を考える。これは、明らかにもとの開被覆 $\{O_\gamma\}_{\gamma\in\Gamma}$ を含むので、$\{O'_\gamma\}_{\gamma\in\Gamma'}$ も $X\times Y$ の開被覆である。すなわち

$$X\times Y\subset\bigcup_{\gamma\in\Gamma'}O'_\gamma$$

となる。このとき $O'_\gamma=O'^X_\gamma\times O'^Y_\gamma(O'^X_\gamma\in X,O'^Y_\gamma\in Y)$ と表すことにすると、積位相空間の開集合族 $\mathfrak{O}_X\times\mathfrak{O}_Y$ の定義より O'^X_γ,O'^Y_γ はそれぞれ X,Y の開集合であり

$$X\subset\bigcup_{\gamma\in\Gamma'}O'^X_\gamma\quad Y\subset\bigcup_{\gamma\in\Gamma'}O'^Y_\gamma$$

となる[18]。このとき Y は開被覆 $\{O'^Y_\gamma\}_{\gamma\in\Gamma'}$ によって覆われているので、Y がコンパクトであることより、その中から Y の有限部分開被覆 $\{O'^Y_1,\cdots,O'^Y_s\}$ をとれる。

さて、有限開被覆の各開集合に対し、$O'_\gamma=O'^X_\gamma\times O'^Y_\gamma$ の関係から得られる O'^X_n 全体の成す集合

$$\mathbb{X}\equiv\{A|A\times O'^Y_n(n=1,\cdots,s)=O'_\gamma(\gamma\in\Gamma')\}$$

を考える。ここで「Y の有限部分開被覆 $\{O'^Y_1,\cdots,O'^Y_s\}$ をどのようにとっても、$\mathbb{X}$ に属するどの集合にも含まれない X の元 x_0 がある」と仮定する。とはいえ、$\{O'_\gamma\}_{\gamma\in\Gamma'}=\{O'^X_\gamma\times O'^Y_\gamma\}_{\gamma\in\Gamma'}$ は $X\times Y$ の開被覆なので、$x_0\in O'^X_{\gamma0}(\gamma_0\in\Gamma')$ となる X の開集合 $O'^X_{\gamma0}$ が存在する。このとき、対応する $O'^Y_{\gamma0}$ を Y の有限部分開被覆に加えれば、加えられた Y の有限部分開被覆 $\{O'^Y_1,\cdots,O'^Y_s,O'^Y_{\gamma0}\}$ に対する $\mathbb{X}$ については、$x_0\in O'^X_{\gamma0}\in\mathbb{X}$ となり、仮定に矛盾する。したがって、Y の有限部分開被覆 $\{O'^Y_1,\cdots,O'^Y_s\}$ をうまくとれば、X の任意の元は $\mathbb{X}$ に属する集合のどれかには含まれる。つまり

$$X\subset\bigcup_{A\in\mathbb{X}}A$$

となるということであり、$\mathbb{X}$ は X の開被覆である。

X はコンパクトなので、$\mathbb{X}$ の有限部分被覆 $\{O'^X_1,\cdots,O'^X_t\}$ によって

$$X\subset\bigcup_{n=1}^{t}O'^X_n$$

と覆われる。また、$\{O'^Y_1,\cdots,O'^Y_s\}$ は Y の有限開被覆であったので

$$Y\subset\bigcup_{m=1}^{s}O'^Y_m$$

と覆われる。以上より

$$X\times Y\subset\bigcup_{n=1}^{t}O'^X_n\times\bigcup_{m=1}^{s}O'^Y_m=\bigcup_{n=1}^{t}\bigcup_{m=1}^{s}(O'^X_n\times O'^Y_m)$$

[17]Y の有限開被覆が得られても、$\{O_\gamma\}_{\gamma\in\Gamma}$ で対応する X の集合族が被覆になるとは限らない。
[18]逆は成立しないのが、直積を扱う上での注意である。これは平面を考えればよくわかる。

である。$\mathbb{X}$ の定義より

$$O_n'^X \times O_m'^Y (n=1,\cdots,t \quad m=1,\cdots,s) \quad = \quad O'_\gamma(\gamma \in \Gamma') \quad \subset \quad O_\gamma(\gamma \in \Gamma)$$

である。このとき、上の包含関係で $O_n'^X \times O_m'^Y$ に対応する O_γ を O_{nm} と表すことにすれば、$O_n'^X \times O_m'^Y \subset O_{nm}$ なので

$$X \times Y \subset \bigcup_{n=1}^{t} \bigcup_{m=1}^{s} (O_n'^X \times O_m'^Y) \subset \bigcup_{n=1}^{t} \bigcup_{m=1}^{s} O_{nm}$$

である。よって、もとの開被覆 $\{O_\gamma\}_{\gamma \in \Gamma}$ の有限部分開被覆 $\{O_{11}, O_{12}, \cdots, O_{1s}, \cdots, O_{t1}, \cdots, O_{ts}\}$ によって $X \times Y$ は覆われたので、積位相空間 $(X \times Y, \mathfrak{O}_X \times \mathfrak{O}_Y)$ はコンパクトである。 証明終

B.3 距離空間

定義 B.6 集合 X に

$$関数 \; d : X \times X \to \boldsymbol{R}$$

が与えられ、任意の $x, y, z \in X$ について以下の性質を満たすとき、(X, d) を**距離空間**といい、d を**距離関数**、$d(x, y)$ を x, y の距離という。

1. $d(x,y) \geq 0 \quad d(x,y) = 0$ なら $x = y$
2. $d(x,y) = d(y,x)$
3. $d(x,y) \leq d(x,z) + d(z,y)$ 　三角不等式

定義 B.7 距離空間 (X, d) と $x \in X$ について、$V_\epsilon(x) \equiv \{p | d(x,p) < \epsilon\}$ と定義し、これを x の ϵ **近傍**もしくは**開球**という。

定義 B.8 距離空間 (X, d) と $x \in X$ について、$\overline{V}_\epsilon(x) \equiv \{p | d(x,p) \leq \epsilon\}$ と定義し、これを x の**閉球**という。

補題 B.5 距離空間 (X, d) と、その部分空間 Y について、$y \in Y, x \in Y^c$ ならば $d(x, y) > 0$ となる。

(proof)

$d(x, y) = 0$ ならば距離空間の定義から $x = y$ となり、$x = y \in Y$ となって矛盾する。 証明終

B.3.1 距離空間における収束

距離空間においては、距離関数による実数との対応づけによって、収束の概念を定義できる。

定義 B.9 距離空間 (X, d) において、X の点列[19] $\{x_n\}$ が

$$\lim_{n \to \infty} d(x_n, a) = 0$$

をみたすとき、点列 $\{x_n\}$ は a に収束するといい

$$\lim_{n \to \infty} x_n = a$$

といったように、数列の収束に準じて表す。

[19] 点とは、一般に集合の元のことを表しており、点列とは、単に要素の列のことである。

数列の極限と同様に、点列の要素がすべてある集合に含まれていても、その点列の極限がその集合に含まれているとは限らない。また、三角不等式によって極限の一意性が保証される。(次の定理を参照)

定理 B.6 距離空間 (X,d) において、X の点列 $\{x_n\}$ が、$n\to\infty$ のとき $x_n\to a$,$x_n\to b$ ならば $a=b$ である。

(proof)
$d(a,b)>0$ と仮定する。三角不等式から $d(a,b)\le d(a,x_n)+d(b,x_n)$ であるが、$n\to\infty$ のとき右辺は 0 に収束するので、$d(a,x_n)+d(b,x_n)<d(a,b)$ とすることができ、矛盾する。したがって、$d(a,b)=0$ であり、距離空間の定義から $a=b$ となる。 証明終

補題 B.7 距離空間 (X,d) において、X の点列 $\{x_n\}$ が x に収束することと、任意の $\varepsilon>0$ に対し、十分大きな n をとれば $x_n\in V_\varepsilon(x)$ となることは同値である。

(proof)
簡単な言い換え、証明略。 証明終

定義 B.10 距離空間 (X,d) と X の点列 $\{x_n\}$ について、任意の $\varepsilon>0$ に対してある自然数 L が存在して

$$\forall n,m\ge L \quad\to\quad d(x_n,x_m)<\epsilon$$

となるとき、点列 $\{x_n\}$ を**コーシー列**という。また、X の任意のコーシー列が収束することを**コーシー完備**であるといい、このとき、X を**完備距離空間**[20] であるという。

B.3.2 距離空間の開集合・閉集合

定義 B.11 距離空間 (X,d) において、X の部分集合 O が $\forall x\in O$ について、各 x に対し適当な $\epsilon>0$ をとると $V_\epsilon(x)\subset O$ となるとき、O を**距離空間の意味で開集合**であるという。また、X,ϕ は距離空間の意味の開集合であると定義する。

定義 B.12 距離空間 (X,d) において、X の部分集合 F について、F^c が距離空間の意味で開集合であるとき、F は**閉集合**であるという。

定理 B.8 $\gamma\in\Gamma$ なる O_γ がすべて距離空間の意味で開集合であるとき $\displaystyle\bigcup_{\gamma\in\Gamma}O_\gamma$ も距離空間の意味で開集合である。

(proof)
$\forall x\in\displaystyle\bigcup_{\gamma\in\Gamma}O_\gamma$ について、$c\in\Gamma$ のどれかを選べば $x\in O_c$ である。したがって、適当な $\epsilon>0$ をとれば $V_\epsilon(x)\subset O_c\subset\displaystyle\bigcup_{\gamma\in\Gamma}O_\gamma$ よって、$\displaystyle\bigcup_{\gamma\in\Gamma}O_\gamma$ は距離空間の意味で開集合である。 証明終

定理 B.9 O_1,O_2 が距離空間の意味で開集合のとき $O_1\cup O_2$ も距離空間の意味で開集合である。

[20] 全順序集合には順序完備の概念があるが、これとコーシー完備は同値ではない。実数 (順序体) は距離空間であり、かつ、全順序集合であるが、その完備性とは、通常、順序完備のほうを指しており、コーシー完備かつアルキメデス的であることと同値である。

(proof)
$O_1 \cap O_2 = \phi$ のときは、定義よりこれは距離空間の意味で開集合である。$O_1 \cap O_2 \neq \phi$ のとき、$\forall x \in O_1 \cap O_2$ について $x \in O_1$ かつ $x \in O_2$ である、O_1, O_2 が距離空間の意味で開集合なので、適当な $\epsilon > 0$ をとれば $V_\epsilon(x) \subset O_1$ かつ $V_\epsilon(x) \subset O_2$ よって $V_\epsilon(x) \subset O_1 \cap O_2$ となり、$O_1 \cap O_2$ は開集合である。 証明終

以上の定理と、距離空間の意味での開集合の定義より、次のことが言える。

定理 B.10 距離空間 (X, d) において、距離空間の意味での開集合からなる集合族 $\mathfrak{O}$ をとると、$(X, \mathfrak{O})$ は位相空間になり、距離空間の意味の開集合は位相空間の意味でも開集合となる。

したがって、通常、距離空間においては、距離空間の開集合・閉集合のことを「距離空間の」という序詞を付けずに呼んでも問題ないわけである。次の定理は、距離空間の閉集合の、最も本質的な性質である。

定理 B.11 距離空間 (X, d) において、X の部分集合 F が閉集合であるとき、F の点列 $\{x_n\}$ が x に収束するなら、$x \in F$ である。[21]

(proof)
$x \notin F$ と仮定すると、$x \in F^c$ であるが、F^c は開集合なので適当な正数 ϵ を選べば

$$V_\epsilon(x) \subset F^c$$

となる。点列 $\{x_n\}$ は x に収束しているので、補題 B.7 より十分大きな n に対して

$$x_n \in V_\epsilon(x) \subset F^c$$

となることになるが、これは $\{x_n\}$ が F の点列であることに矛盾する。したがって $x \in F$ である。 証明終

補題 B.12 開球は開集合である。

(proof)
開球を $V_r(a)$ とする。$\forall b \in V_r(a)$ について、$d(a, b) < r$ である。ここで、$r - d(a, b) > 0$ なので、$V_{r-d(a,b)}(b)$ という b の $r - d(a, b)$ 近傍 を考えることができる。$\forall x \in V_{r-d(a,b)}(b)$ について

$$\begin{aligned} d(b, x) &< r - d(a, b) \\ d(b, x) + d(a, b) &< r \\ d(a, x) &< r \quad \because \text{三角不等式} \end{aligned}$$

であるから、$x \in V_r(a)$ である。つまり、任意の $b \in V_r(a)$ に対し、正数 $r - d(a, b)$ をとることで $V_{r-d(a,b)}(b) \subset V_r(a)$ となるので、開球 $V_r(a)$ は開集合である。 証明終

補題 B.13 閉球は閉集合である。

(proof)
閉球を $\overline{V}_r(a)$ とする。$\forall b \in \overline{V}_r(a)^c$ について、$d(a, b) \geq r$ である。ここで、$d(a, b) - r > 0$ なので、

[21] 選択公理を用いれば、逆も成り立つ。

$V_{d(a,b)-r}(b)$ という b の $d(a,b)-r$ 近傍 を考えることができる。$\forall x \in V_{d(a,b)-r}(b)$ について

$$\begin{aligned} d(b,x) &< d(a,b)-r \\ r+d(b,x) &< d(a,b) \\ r+d(b,x) &< d(a,x)+d(b,x) \quad \because \text{三角不等式} \\ r &< d(a,x) \end{aligned}$$

であるから、$x \in \overline{V}_r(a)^c$ である。つまり、任意の $b \in \overline{V}_r(a)^c$ に対し、正数 $d(a,b)-r$ をとることで $V_{d(a,b)-r} \subset \overline{V}_r(a)^c$ となるので、$\overline{V}_r(a)^c$ は開集合である。よって、閉球 $\overline{V}_r(a)$ は閉集合である。 証明終

B.4 近傍

定義 B.13 X を位相空間、$a \in X$ とする。このとき、X の部分集合 A について

$$a \in B \subset A$$

となる開集合 B が存在するとき、A は点 a の**近傍**であるという。

定理 B.14 位相空間 X とその部分集合 A について、A が開集合であることと、A の任意の点について $N \subset A$ なる近傍 N が存在することは、同値である。

(proof)
A が開集合であれば、A の任意の点について $N=A$ が $N \subset A$ なる近傍である。逆に、A の任意の点について $N \subset A$ なる近傍 N が存在するとき。そのような、$a \in A$ の近傍を $N(a)$ とすると、$\forall a \in A$ について、近傍の定義より

$$a \in O(a), O \subset N(a)$$

なる開集合 $O(a)$ が存在する。これについて $a \in O(a) \subset \bigcup_{a' \in A} O(a')$ であるから、つまり

$$A \subset \bigcup_{a' \in A} O(a')$$

である。また、$O(a) \subset N(a) \subset A$ であるから

$$\bigcup_{a' \in A} O(a') \subset A$$

である。よって

$$A = \bigcup_{a' \in A} O(a')$$

となる。$O(a')$ はそれぞれ開集合なので、その和集合である A も開集合である。 証明終

補題 B.15 距離空間では、そのある点 x について、ϵ 近傍 $V_\epsilon(x)$ は、x の近傍である。また、任意の x の近傍 N について

$$\exists \epsilon > 0, V_\epsilon(x) \subset N$$

となる。

(proof)
容易。詳細略。 証明終

B.5 連続写像

定義 B.14 X, Y を位相空間、f を $f : X \to Y$ なる写像、a を X の点とする。$f(a) \in Y$ の任意の近傍 N_Y に対して

$$f(N_X) \subset N_Y$$

となる $a \in X$ の近傍 N_X が存在するとき、f は**点 a で連続**であるという。また、$\forall x \in X$ について f が点 x で連続であるならば、f は**連続**であるという。

位相空間での近傍には、全空間も含まれており、近いという感覚はあまりない。連続の定義では、任意の近傍を対象とすることによって、一般の位相空間では表現しづらい「近づく」ということを表現している。

定理 B.16 $(X, d_X), (Y, d_Y)$ を距離空間、f を $f : X \to Y$ なる写像、a を X の点とする。このとき

$$f \text{ が } a \text{ で連続} \quad \leftrightarrow \quad \forall\epsilon, \exists\delta \qquad f(V_\delta(a)) \subset V_\epsilon(f(a))$$

である。

(proof)

fがaで連続であるとき。$f(a)$ の任意の近傍 N_Y に対して $f(N_X) \subset N_Y$ となる a の近傍 N_X が存在する。よって、補題 B.15 より、任意の $\forall\epsilon > 0$ について

$$f(N_X) \subset V_\epsilon(f(a))$$

となる a の近傍 N_X が存在する。このとき、補題 B.15 よりある δ が存在して

$$V_\delta(a) \subset N_X$$

である。よって、任意の ϵ について

$$f(V_\delta(a)) \subset f(N_X) \subset V_\epsilon(f(a))$$

である。

逆に　$\forall\epsilon, \exists\delta, f(V_\delta(a)) \subset V_\epsilon(f(a))$ であるとき。任意の $f(a)$ の近傍 N_Y について、補題 B.15 より、ある $\phi > 0$ が存在して $V_\phi(f(a)) \subset N_Y$ となる。よって、条件よりある δ が存在して

$$f(V_\delta(a)) \subset V_\phi(f(a)) \subset N_Y$$

であり、補題 B.15 から $V_\delta(a)$ は a の近傍である。よって示された。 証明終

定理 B.17 $(X, d_X), (Y, d_Y)$ を距離空間、f を $f : X \to Y$ なる写像、a を X の点とする。このとき

$$f \text{ が } a \text{ で連続} \quad \leftrightarrow \quad \lim_{n\to\infty} x_n = a \text{ なる任意の点列 } \{x_n\} \text{ について } \lim_{x\to\infty} f(x_n) = f(a)$$

である。

(proof)

f が a で連続であるとき。定理 B.16 より

$$\forall\epsilon, \exists\delta \qquad f(V_\delta(a)) \subset V_\epsilon(f(a))$$

である。これを言い換えると

$$\forall d(a,x) < \delta ならば\, d(f(a), f(x)) < \epsilon$$

である。$\lim_{n\to\infty} x_n = a$ なので、n を十分大きくとれば $d(a,x) < \delta$ とできる。よってこのとき任意の ϵ について $d(f(a), f(x)) < \epsilon$ となるのだから

$$\lim_{x\to\infty} f(x_n) = f(a)$$

である。

逆に、$\lim_{n\to\infty} x_n = a$ なる任意の点列 $\{x_n\}$ について $\lim_{n\to\infty} f(x_n) = f(a)$ であるとき。f が a で連続でないとすると、定理 B.16 より

$$\forall\delta, \exists\epsilon \quad f(V_\delta(a)) \not\subset V_\epsilon(f(a))$$

である。つまり、任意の δ に対して、ある ϵ が存在して $b \in V_\delta(a)$ かつ $f(b) \notin V_\epsilon(f(a))$ となる $b \in X$ が存在する。よって、$b_n \in V_{\frac{1}{n}}(a)$ かつ $f(b_n) \notin V_\epsilon(f(a))$ であるような点列 $\{b_n\}$ が存在する。これについては $d(b_n, a) < \frac{1}{n}$ であるから、

$$\lim_{n\to\infty} b_n = a$$

よって、条件より

$$\lim_{n\to\infty} f(b_n) = f(a)$$

である。つまり、n を十分大きくとれば

$$d(f(b_n), f(a)) < \epsilon$$

となるということだが、これは $f(b_n) \notin V_\epsilon(f(a))$ に矛盾する。よって f が a で連続である。 証明終

定理 B.18 距離関数は連続である。

(proof)

距離空間を (X,d) とする。$\forall a,b \in X$ について、任意の $x_n \to a, y_n \to b$ なる点列 $\{x_n\}\{y_n\}$ を考えたとき、三角不等式より

$$\begin{aligned}
d(x_n, y_n) &\leq d(x_n, a) + d(a, y_n) \\
&\leq d(x_n, a) + d(y_n, b) + d(a, b) \\
d(x_n, y_n) - d(a,b) &\leq d(x_n, a) + d(y_n, b)
\end{aligned}$$

であり、また

$$\begin{aligned}
d(a,b) &\leq d(x_n, a) + d(b, x_n) \\
&\leq d(x_n, a) + d(y_n, b) + d(x_n, y_n) \\
-(d(x_n, y_n) - d(a,b)) &\leq d(x_n, a) + d(y_n, b)
\end{aligned}$$

であるから

$$|d(x_n, y_n) - d(a,b)| \leq d(x_n, a) + d(y_n, b)$$

である。したがって、条件より $n \to \infty$ とすると $d(x_n, a) \to 0, d(y_n, b) \to 0$ であるから

$$\lim_{n\to\infty} d(x_n, y_n) = d(a,b)$$

となる。$a, b \in X$ は任意だったので、定理 B.17 より距離関数 d は連続である。 証明終

定理 B.19 写像 $f: X \to Y$ が連続であることと、任意の Y の開集合 O について $f^{-1}(O)$ が X の開集合であることは、同値である。

(proof)

写像 $f: X \to Y$ が連続であるとき。O を Y の任意の開集合とする。$f^{-1}(O)$ が空集合ならば、これは確かに開集合である。$f^{-1}(O) \neq \phi$ であるとき。$\forall x \in f^{-1}(O)$ について、$f(x) \in O$ であり、O は $f(x)$ の近傍であるから、f の連続性より

$$f(N_x) \subset O \to N_x \subset f^{-1}(O)$$

となる x の近傍 N_x が存在する。よって、定理 B.14 より $f^{-1}(O)$ は開集合である。

逆に、任意の Y の開集合 O について $f^{-1}(O)$ が X の開集合であるとき。$\forall x \in X$ について、$f(x)$ の任意の近傍 N_Y を考えると、近傍の定義より $f(x) \in O \subset N_Y$ なる開集合 O が存在する。このとき、$x \in f^{-1}(O)$ であり、条件から $f^{-1}(O)$ は開集合なので $f^{-1}(O)$ は x の近傍である。また、$O \subset N_Y$ より

$$f^{-1}(O) \subset f^{-1}(N_Y)$$

であるから、$f(x)$ の任意の近傍 N_Y に対して、上記の関係を満たす x の近傍 $f^{-1}(O)$ が存在するということであり、任意の $x \in X$ について f は x で連続である。つまり、f は連続である。 証明終

B.5.1 連続写像で保たれる性質

定理 B.20 連続写像 $f: X \to Y$ について、X がコンパクトならば $f(X)$ もコンパクトである。

(proof)

$\{O_\gamma\}_{\gamma\in\Gamma}$ を $f(X)$ の開被覆とする。定理 B.19 より、$f^{-1}(O_\gamma)$ は開集合である。また、$\forall x \in X$ について

$$x \in f(X) \subset \bigcup_{\gamma\in\Gamma} O_\gamma$$

であるから

$$\exists\gamma' \in \Gamma \quad f(x) \subset O'_\gamma \leftrightarrow x \in f^{-1}(O'_\gamma)$$

である。$\therefore x \in f^{-1}(O'_\gamma) \subset \bigcup_{\gamma\in\Gamma} f^{-1}(O_\gamma)$ つまり

$$X \subset \bigcup_{\gamma\in\Gamma} f^{-1}(O_\gamma)$$

であり、$\{f^{-1}(O_\gamma)\}_{\gamma\in\Gamma}$ は X の開被覆である。X がコンパクトなので、このうちの有限個の $\{f^{-1}(O_1), \cdots, f^{-1}(O_m)\}$ によって

$$\begin{aligned} X &\subset \bigcup_{n=1}^{m} f^{-1}(O_n) \\ &= f^{-1}\left(\bigcup_{n=1}^{m} O_n\right) \quad \because m < \infty \end{aligned}$$

と覆われる。このとき

$$f(X) \subset f\left(f^{-1}\left(\bigcup_{n=1}^{m} O_n\right)\right) \subset \bigcup_{n=1}^{m} O_n \quad \because f(f^{-1}(A)) \subset A$$

であるから、もとの開被覆から選んだ $\{O_n\}$ は $f(X)$ の有限被覆である。よって、$f(X)$ はコンパクトである。 証明終

C　ユークリッド空間

定義 C.1 $\boldsymbol{R}$ を実数体とする。$\boldsymbol{R^n}$ に

$$\boldsymbol{x} = (x_1, \cdots, x_n), \boldsymbol{y} = (y_1, \cdots, y_n) \in \boldsymbol{R^n} \quad ; \quad d(\boldsymbol{x}, \boldsymbol{y}) = \sqrt{\sum_{k=1^n} (x_k - y_k)^2}$$

という距離を導入すると、$(\boldsymbol{R^n}, d)$ は距離空間[22]となる。これを**ユークリッド空間**と言い、d を**標準距離・ユークリッド距離**と称す。

距離空間であるので、定義 B.11 及び B.12 により**開集合・閉集合**を定義できる。他の距離空間・位相空間に関する事項も同様である。

C.1　積位相空間による開集合の定義

$\boldsymbol{R}$ 自身も距離空間なので、$\boldsymbol{R^n}$ の開集合族を $\boldsymbol{R}$ の積位相空間 (p50) から定義することもできる、すなわち

$$O = O_1 \times \cdots \times O_n \subset \boldsymbol{R^n}\text{が開集合} \Leftrightarrow O_1, \cdots, O_n\text{がすべて } \boldsymbol{R} \text{ の開集合}$$

と定義すれば、これは定理 B.2 より位相空間の開集合族の条件を満たす。こうして得られた、積位相空間としての開集合は、ユークリッド距離による開集合と一見異なるが、実のところ、この二つは同値である。

定理 C.1 $O = O_1 \times \cdots \times O_n \subset \boldsymbol{R^n}$ が、積位相空間としての $\boldsymbol{R^n}$ の開集合であることと、ユークリッド距離空間としての開集合であることは、同値である。

(proof)
$O = O_1 \times \cdots \times O_n \subset \boldsymbol{R^n}$ が、積位相空間としての $\boldsymbol{R^n}$ の開集合であるとき。まず

$$O_i\text{が } \boldsymbol{R} \text{ の開集合である} \quad \Leftrightarrow \quad \forall a \in O_i\text{に対して適当な正数}\epsilon\text{をとると } V_\epsilon(a) \subset O_i\text{となる}$$

である。ここでは $V_\epsilon(a) = \{x; |x - a| < \epsilon\}$ であるから、O が積位相空間としての $\boldsymbol{R^n}$ の開集合であるとは、

$\forall \boldsymbol{a} = (a_1, \cdots, a_n) \in O$ に対して適当な正数 $\epsilon_1, \cdots, \epsilon_n$ をとると
$\{(x_1, \cdots, x_n); |x_1 - a_1| < \epsilon_1, \cdots, |x_n - a_n| < \epsilon_n\} \subset O$ となる

ことである。

このとき、$\epsilon \equiv min(\epsilon_1, \cdots, \epsilon_n)$ とすると

$$\{(x_1, \cdots, x_n); |x_1 - a_1| < \epsilon, \cdots, |x_n - a_n| < \epsilon\} \subset O \tag{48}$$

である。$\forall \boldsymbol{y} \in V_\epsilon(\boldsymbol{a})$ について

$$|y_i - a_i|^2 \leq \sum_{k=1}^{n} |y_k - a_k|^2 = d(\boldsymbol{y}, \boldsymbol{a})^2 < \epsilon^2$$

よって $\boldsymbol{y} \in \{(x_1, \cdots, x_n); |x_1 - a_1| < \epsilon, \cdots, |x_n - a_n| < \epsilon\}$ であるからつまり

$$\begin{aligned} V_\epsilon(\boldsymbol{a}) &\subset \{(x_1, \cdots, x_n); |x_1 - a_1| < \epsilon, \cdots, |x_n - a_n| < \epsilon\} \\ &\subset O \quad \because (48) \end{aligned}$$

であり、任意の $\boldsymbol{a} \in O$ について適当な $\epsilon > 0$ をとると $V_\epsilon(\boldsymbol{a}) \subset O$ が成り立つので、O はユークリッド距離空間の意味でも開集合である。

[22] 付録 53p 参照。証明略。

逆に、O がユークリッド距離空間の意味で開集合であるとき。任意の $\boldsymbol{a} \in O$ について適当な $\epsilon > 0$ をとると $V_{\epsilon\sqrt{n}}(\boldsymbol{a}) \subset O$ が成り立つ。$\forall \boldsymbol{x} \in \{(x_1, \cdots, x_n); |x_1 - a_1| < \epsilon, \cdots, |x_n - a_n| < \epsilon\}$ について

$$\begin{aligned} d(\boldsymbol{x}, \boldsymbol{a})^2 &= \sum_{i=1}^{n} |x_i - a_i|^2 \\ &< \sum_{i=1}^{n} \epsilon^2 = n\epsilon^2 \\ d(\boldsymbol{x}, \boldsymbol{a}) &< \epsilon\sqrt{n} \end{aligned}$$

よって、$\boldsymbol{x} \in V_{\epsilon\sqrt{n}}$ である。つまり

$$\{(y_1, \cdots, y_n); |y_1 - a_1| < \epsilon, \cdots, |y_n - a_n| < \epsilon\} \subset V_{\epsilon\sqrt{n}} \subset O$$

であり、任意の $\boldsymbol{a} \in O$ について適当な $\epsilon > 0$ をとると $\{(y_1, \cdots, y_n); |y_1 - a_1| < \epsilon, \cdots, |y_n - a_n| < \epsilon\} \subset O$ が成り立つので、O は $\boldsymbol{R}$ の積位相空間の意味でも開集合である。 証明終

C.2 半開区間・開区間・閉区間

定義 C.2 $\boldsymbol{a}, \boldsymbol{b} \in \boldsymbol{R^n}$ について

$$[\boldsymbol{a}, \boldsymbol{b}) \equiv [a_1, b_1) \times \cdots \times [a_n, b_n) = \{(x_1, \cdots, x_n); a_1 \leq x_1 < b_1, \cdots, a_n \leq x_n < b_n\}$$

と定義し、これを**半開区間**という。同様に**閉区間**もしくは $\boldsymbol{n}$ **次元長方形**を

$$[\boldsymbol{a}, \boldsymbol{b}] \equiv [a_1, b_1] \times \cdots \times [a_n, b_n] = \{(x_1, \cdots, x_n); a_1 \leq x_1 \leq b_1, \cdots, a_n \leq x_n \leq b_n\}$$

と定義し、また**開区間**を

$$(\boldsymbol{a}, \boldsymbol{b}) \equiv (a_1, b_1) \times \cdots \times (a_n, b_n) = \{(x_1, \cdots, x_n); a_1 < x_1 < b_1, \cdots, a_n < x_n < b_n\}$$

と定義する。また、これらについて、その大きさを

$$|[\boldsymbol{a}, \boldsymbol{b})| \equiv |[\boldsymbol{a}, \boldsymbol{b}]| \equiv |(\boldsymbol{a}, \boldsymbol{b})| \equiv |a_1 - b_1| \cdots |a_n - b_n|$$

と定義する。

定義 C.3 $\boldsymbol{R^n}$ の部分集合 A が**有界**であるとは、適切な閉区間 I をとったとき $A \subset I$ となることをいう。

定理 C.2 半開区間の共通部分は半開区間である。

(proof)
定義を考えると

$$\begin{aligned} &[b_1, a_1) \times \cdots \times [b_n, a_n) \cap [b'_1, a'_1) \times \cdots \times [b'_n, a'_n) \\ =&[min(b_1, b'_1) - max(a_1, a'_1)) \times \cdots \times [min(b_n, b'_n) - max(a_n, a'_n)) \end{aligned} \tag{49}$$

である。よって示された。 証明終

定理 C.3 半開区間 $I_1 \subset I_2$ について $|I_1| \leq |I_2|$

(proof)
各次元について $|b_i - a_i|$ を考えればよい。容易。詳細略。 証明終

補題 C.4 開区間は開集合である。

(proof)
$\boldsymbol{R}$ では、$(a,b) = V_{\frac{|a-b|}{2}}(\frac{a+b}{2})$ なので、補題 B.12 より、開集合である。

$\boldsymbol{R^n}$ の開区間は、$\boldsymbol{R}$ の開区間の直積なので、積位相空間の意味では開集合である。よって、定理 C.1 より、$\boldsymbol{R^n}$ の開区間は開集合である。 $\boxed{\text{証明終}}$

C.3 コンパクト

コンパクトの定義等については、p51 を参照。

定理 C.5 $\boldsymbol{R}$ の有界閉区間はコンパクトである。

(proof)
有界閉区間 $[a,b]$ と、その開被覆 $\{O_\gamma\}_{\gamma\in\Gamma}$ を考える。

$$A = \{x | [a,x] \text{ が } \{O_\gamma\}_{\gamma\in\Gamma} \text{の有限部分被覆で覆われる }\}$$

とおく。開被覆 $\{O_\gamma\}_{\gamma\in\Gamma}$ の中には、必ず a を含むものがあるので、それを持ってくることによって $[a,a] = \{a\}$ は有限部分被覆で覆われる。すなわち $a \in A$ であり、A は空集合ではない。また、x が十分大きい時、$[a,x]$ は $[a,b]$ の開被覆では覆えない。よって、A は上に有界である。したがって、$\boldsymbol{R}$ は順序完備 (付録 p44) なので、上限 $\sup A$ が存在する。

ここで、$\sup A < b$ と仮定する。$a \le \sup A < b$ なので、$\sup A$ は、開被覆 $\{O_\gamma\}_{\gamma\in\Gamma}$ のある開集合 O' に含まれる。O' が開集合なので、適当な正数 ϵ をとると、$V_\epsilon(\sup A) \subset O'$ となる。よって、$\sup A < d < \sup A + \epsilon$ となる d を選んでおくと

$$(\sup A, d] \subset (\sup A - \epsilon, \sup A + \epsilon) = V_\epsilon(\sup A) \subset O'$$

である。A の定義より $[a, \sup A]$ は開被覆 $\{O_\gamma\}_{\gamma\in\Gamma}$ の有限部分被覆 $\{O_1, \cdots, O_s\}$ によって覆われる。このとき、

$$[a,d] = [a, \sup A] + (\sup A, d] \subset O_1 \cup \cdots \cup O_s \cup O'$$

となる。すなわち $[a,d]$ が開被覆 $\{O_\gamma\}_{\gamma\in\Gamma}$ の有限部分被覆で覆われるということであり、これは $\sup A$ が A の上限であることに矛盾する。したがって、$b \le \sup A$ である。このとき $[a,b] \subset [a, \sup A]$ であり、$[a, \sup A]$ は $[a,b]$ の開被覆 $\{O_\gamma\}_{\gamma\in\Gamma}$ の有限部分被覆で覆われるので、$[a,b]$ もその有限部分被覆によって覆われる。よって、$[a,b]$ はコンパクトである。 $\boxed{\text{証明終}}$

定理 C.6 $\boldsymbol{R^n}$ の有界閉集合はコンパクトである。

(proof)
$\boldsymbol{R^n}$ の有界閉集合を F とおく。有界なので十分大きな閉区間 $[-\boldsymbol{a}, \boldsymbol{a}]$ をとれば $F \subset [-\boldsymbol{a}, \boldsymbol{a}] = [-a,a]^n$ となる。定理 C.5 より $[-a,a]$ はコンパクトであり、したがって、定理 B.4 より $[-\boldsymbol{a}, \boldsymbol{a}] = [-a,a]^n$ もコンパクトである。F はその部分閉集合なので、定理 B.3 より F もコンパクトである。 $\boxed{\text{証明終}}$

逆も成り立つ。

定理 C.7 $\boldsymbol{R^n}$ のコンパクトな集合は、有界閉集合である。

(proof)
 C を $\boldsymbol{R^n}$ のコンパクト集合とする。補題 C.4 より、開区間は開集合なので、$(-a,a)^n \quad (a=1,2,\cdots\to\infty)$ の全体は A の開被覆であるが、C はコンパクトなのでそのうちの有限個によって覆われる。この有限部分被覆のうち最大のものを $(-N,N)^n$ とすると、$C\subset(-N,N)^n\subset[-N-1,N+1]^n$ であるので、C は有界である。

 また、任意の $x\in C^c$ をとり、x を中心とする閉球の補集合による集合族 $\{\overline{V}_{\frac{1}{m}}(x)^c\}_{m\in\boldsymbol{Z}}$ を考える。$\forall y\in C$ について、補題 B.5 より $d(x,y)>0$ である。よって $d(x,y)>\frac{1}{M}>0$ なる $M\in\boldsymbol{Z}$ をとれば

$$y\notin\overline{V}_{\frac{1}{M}}(x)\quad\Leftrightarrow\quad y\in\overline{V}_{\frac{1}{M}}(x)^c\quad\therefore C^c\subset\bigcup_{m\in\boldsymbol{Z}}\overline{V}_{\frac{1}{m}}(x)^c$$

となる。さらに、補題 B.13 より $\overline{V}_{\frac{1}{m}}(x)$ は閉集合なので、$\overline{V}_{\frac{1}{m}}(x)^c$ は開集合である。よって、$\{\overline{V}_{\frac{1}{m}}(x)^c\}_{m\in\boldsymbol{Z}}$ は C の開被覆である。C はコンパクトなので、この中から C の有限部分被覆をとることができる。そのうち最大のもの、すなわち m が最小のものを $\overline{V}_{\frac{1}{m'}}(x)^c$ とすれば、これは有限部分被覆の他のどの開集合も含むので

$$C\subset\overline{V}_{\frac{1}{m'}}(x)^c\quad\Leftrightarrow\quad\overline{V}_{\frac{1}{m'}}(x)\subset C^c$$

となる。すなわち、任意の $x\in C^c$ に関して、正数 $\frac{1}{m'}$ をとれば $\overline{V}_{\frac{1}{m'}}(x)\subset C^c$ となるので、C^c は開集合である。よって C は閉集合である。 証明終

補題 C.8 閉区間はコンパクトであり、有界閉集合である。

(proof)
 閉区間を $[a_1,b_1]\times\cdots\times[a_n,b_n]$ とする。定理 C.5 より、各 $[a_1,b_1],\cdots,[a_n,b_n]$ はコンパクトである。よって、定理 B.4 より閉区間 $[a_1,b_1]\times\cdots\times[a_n,b_n]$ はコンパクトである。したがって、定理 C.7 よりこれは有界閉集合である。 証明終

C.4 最大値・最小値

定理 C.9 $\boldsymbol{R}$ の有界閉集合は、最大値・最小値をもつ。

(proof)
 任意の $\boldsymbol{R}$ の有界閉集合Fについて、有界であることよりF $\subset[-a.a]$ が成立する。$[-a.a]$ の n 等分点

$$P_n(k)\equiv\frac{2a}{n}\left(k-\frac{n}{2}\right)\quad(k=0,\cdots,n)$$

をとる。このとき

$$\sum_{k=0}^{n}\overline{V}_{\frac{a}{n}}(P_n(k))=[-a-\frac{a}{n},a+\frac{a}{n}]\supset F$$

である。よって $\overline{V}_{\frac{a}{n}}(P_n(k))\cap F\neq\phi$ となる k が存在する。その k のなかで最大のものを $K(n)$ とし、$x_n\equiv P_n(K(n))$ とおく。$x_n\in F$ であるから、無限実数列 $\{x_n\}$ は有界である。よって定理 A.9 より、収束部分列 $\{x_{\varphi(n)}\}$ を持つ。この極限を c と置くと、$x_n\in F$ で F が閉集合なので、定理 B.11 より $c\in F$ となる。

 $\forall f\in F\subset\sum_{k=0}^{n}\overline{V}_{\frac{a}{n}}(P_n(k))$ について $\exists k'_n,f\in\overline{V}_{\frac{a}{n}}(P_n(k'_n))$ である。このとき $\overline{V}_{\frac{a}{n}}(P_n(k'))\cap F\neq\phi$ であるから

$$k'_n\leq K(n)\quad\therefore P_n(k'_n)\leq P_n(K(n))=x_n$$

である。また $f\in\overline{V}_{\frac{a}{n}}(P_n(k'_n))$ より

$$|f-P_n(k'_n)|\leq\frac{a}{n}$$

なので $\lim_{n\to\infty} P_n(k'_n) = f$ である。ところで $P_n(k'_n) \leq x_n$ であるから、$n \to \infty$ とすると

$$f \leq c$$

である。$f \in F$ は任意だったので、c は F の最大値である。よって示された。最小値も同様。 証明終

定理 C.10 $\boldsymbol{R}$ の上に有界な閉集合は、最大値を持ち、下に有界な閉集合は、最小値を持つ。

(proof)
$\boldsymbol{R}$ の上に有界な閉集合を F とする。上に有界なので $\forall f \in F, f \leq a$ である。ここで適当な $f' \in F$ をとり

$$F' \equiv F \cap [f', a]$$

とすると、定理 C.8 より $[f', a]$ は閉集合なので、定理 B.1 より F' は閉集合である。また、

$$F' \subset [f', a] \subset [-max(|f'|, |a|), max(|f'|, |a|)]$$

なので、F' は有界である。よって定理 C.9 より F' は最大値 M をもつ。

$\forall f \in F$ について、$f \in F'$ ならば M は F' の最大値より $f \leq M$ である。$f \notin F'$ ならば $f \notin [f', a]$ であるが、$f \leq a$ なので $f < f' \leq M$ である。よって $M \in F' \subset F$ であるから M は F の最大値である。よって示された。最小値も同様。 証明終

定理 C.11 コンパクトな位相空間 X と連続関数 $f : X \to \boldsymbol{R}$ については、$f(X) \subset \boldsymbol{R}$ に最大値・最小値が存在する。

(proof)
定理 B.20 より、$f(X)$ はコンパクトである。よって、定理 C.7 より $f(X)$ は有界閉集合である。ゆえに、定理 C.9 より、$f(X)$ に最大値・最小値が存在する。 証明終

索 引

参考文献

[1] 最適化法 田村明久・村松正和著 工系数学講座 17 共立出版

[2] 最適化法 藤田宏・今野浩・田邉國士著 岩波講座応用数学 [方法 7] 岩波書店

[3] 数理計画法 東京大学工学部計数工学科授業 室田一雄教授 土谷隆非常勤講師

[4] 離散凸解析 室田一雄著 現代数学の潮流 共立出版

[5] 数学の基礎 集合・数・位相 齋藤正彦著 基礎数学 14 東京大学出版会

[6] 解析入門 I 小平邦彦著 岩波書店

[7] 位相への 30 講 志賀浩二 朝倉書店

[8] OR の基礎 AHP から最適化まで 加藤豊・小沢正典著 実教出版

[9] COOLEE のホームページ http://coolee.at.infoseek.co.jp/convex.html

[10] 多変数の初等解析入門 落合卓四郎・高橋勝雄著 東京大学出版会