

# 通信路容量を達成する入力シンボルについての一考察

## On the Input Symbols that Achieve the Channel Capacity

中川 健治\*  
Kenji Nakagawa

渡部 康平\*  
Kohei Watabe

**Abstract**— We first consider the minimum enclosing circle in the Euclidean space, and then, based on the investigation, we make an algorithm to calculate the channel capacity of a discrete memoryless channel.

**Keywords**— channel capacity, minimum enclosing circle, Kullback-Leibler divergence, information geometry

### 1 はじめに

離散的無記憶通信路の通信路容量  $C$  は相互情報量の最大値として定義される。 $C$  はまた、Kullback-Leibler (KL) 情報量に関する min max 問題としても定式化できる [3],[7]。この min max 問題を KL 情報量による幾何学に基づいて考察すると  $C$  を計算するアルゴリズムを見いだすことができる。

通信路容量の問題の類似として、ユークリッド空間の有限点集合に対する最小包含円の問題を考えることができる [2],[8]。最小包含円の問題はユークリッド距離に関する min max 問題として定式化でき、通信路容量の問題と類似した幾何構造を持つ。

本論文では、まず、ユークリッド空間における最小包含円について考察し、その考察をもとに通信路容量を計算する組み合わせ的でないアルゴリズムを作ることを目的とする。

### 2 ユークリッド空間における最小包含円

$n$  次元ユークリッド空間  $\mathbb{R}^n$  内の有限個の点  $P_1, \dots, P_m$  を考える。これらの点を内部または境界上を含む球のうち半径が最小のものを**最小包含円**といい、 $\Gamma(P_1, \dots, P_m)$  と表す。与えられた点  $P_1, \dots, P_m$  に対して、最小包含円は唯一つ存在する。最小包含円  $\Gamma(P_1, \dots, P_m)$  は次のように定式化される。すなわち、点  $P_i$  と  $Q \in \mathbb{R}^n$  とのユークリッド距離を  $d(P_i, Q)$  とすると、

$$\min_{Q \in \mathbb{R}^n} \max_{1 \leq i \leq m} d(P_i, Q) \quad (1)$$

を達成する  $Q = Q^*$  が  $\Gamma(P_1, \dots, P_m)$  の中心であり、 $\max_{1 \leq i \leq m} d(P_i, Q^*)$  がその半径である。

最小包含円に関して多くの従来研究がある。本論文では、特に理論的な観点からの研究として [2],[8] を参照した。

我々は、与えられた点  $P_1, \dots, P_m$  に対して  $\Gamma(P_1, \dots, P_m)$  を決定するアルゴリズムを作りたい。まず、 $n = 2$ ,  $m = 3$  の場合を考える。この場合、以下の3つの場合

に分かれる。 $P_1, P_2, P_3$  を頂点とする三角形を  $\triangle P_1 P_2 P_3$  とする。

- (I)  $\triangle P_1 P_2 P_3$  が鋭角三角形
- (II)  $\triangle P_1 P_2 P_3$  が鈍角三角形
- (III) 3点  $P_1, P_2, P_3$  が一直線上にある

(I) の場合、 $\triangle P_1 P_2 P_3$  の外心、すなわち  $P_1, P_2, P_3$  からの等距離点を  $Q^*$  とすると、 $Q^*$  は  $\triangle P_1 P_2 P_3$  の内部にある。 $Q^*$  が最小包含円  $\Gamma(P_1, P_2, P_3)$  の中心であり、 $d(P_1, Q^*) = d(P_2, Q^*) = d(P_3, Q^*)$  がその半径である。

次に、(II) の場合、例えば、頂点  $P_2$  の角が鈍角とすると、 $\Gamma(P_1, P_2, P_3)$  の中心は辺  $P_1 P_3$  の中点である。 $P_1, P_2, P_3$  からの等距離点（すなわち  $\triangle P_1 P_2 P_3$  の外心）を  $Q_0$  とすると、 $Q_0$  は  $\triangle P_1 P_2 P_3$  の内部または周上ではなく、辺  $P_1 P_3$  に関して  $P_2$  と反対側にある。そこで、最小包含円  $\Gamma(P_1, P_2, P_3)$  の中心  $Q^*$  を次の計算アルゴリズムで得ることができる。

#### Algorithm-1

- (i) 方程式  $d(P_1, Q) = d(P_2, Q) = d(P_3, Q)$  を解いて等距離点  $Q = Q_0$  を得る。
- (ii)  $Q_0$  が（例えば）辺  $P_1 P_3$  に関して  $P_2$  と反対側にあることを確認する。
- (iii) 点  $P_2$  を除外して、 $P_1, P_3$  に対する最小包含円  $\Gamma(P_1, P_3)$  の中心  $Q^*$  を得る。
- (iv)  $Q^*$  が最小包含円  $\Gamma(P_1, P_2, P_3)$  の中心である。

(III) の場合は、もし  $P_2$  が線分  $P_1 P_3$  上にあるとすると、線分  $P_1 P_3$  の中点  $Q^*$  が  $\Gamma(P_1, P_2, P_3)$  の中心である。この場合、 $P_1, P_2, P_3$  からの等距離点は存在しないのだが、 $P_1, P_2, P_3$  の位置関係を知らずに  $Q^*$  を得るためのアルゴリズムはどのようにしたらよいだろうか？

すぐに考えられる方法は、3点のうちの2点ずつ調べて、 $\Gamma(P_1, P_2), \Gamma(P_1, P_3), \Gamma(P_2, P_3)$  のうち半径が最大のものが求める最小包含円  $\Gamma(P_1, P_2, P_3)$  である。しかし、この方法では、点の個数が多くなり、空間の次元が大きくなると計算量が大きく、一般化は困難である。

#### 2.1 点の次元を上げる方法

そこで、上記 (III) の場合を再考する。図 1 のように  $P_1, P_2, P_3$  が一直線上にあるときでも、 $P_2$  を少しだけずらして  $\tilde{P}_2$  とすると  $\triangle P_1 \tilde{P}_2 P_3$  は鈍角三角形になり、 $P_1, \tilde{P}_2, P_3$  からの等距離点  $Q_0$  が存在する。そこで、Algorithm-1 が適用できる。 $Q_0$  が辺  $P_1 P_3$  に関して  $\tilde{P}_2$  と反対側にあることを確認して  $P_2$  を除外して最小包含円  $\Gamma(P_1, P_3)$  を得る。ずらす量が十分小さければ  $\Gamma(P_1, P_2, P_3) = \Gamma(P_1, \tilde{P}_2, P_3) = \Gamma(P_1, P_3)$  となる。

\* 長岡技術科学大学 〒940-2188 新潟県長岡市上富岡町 1603-1 Nagaoka University of Technology, Nagaoka, Niigata, 940-2188, Japan, E-mail: nakagawa@, k.watabe@vos.nagaokaut.ac.jp

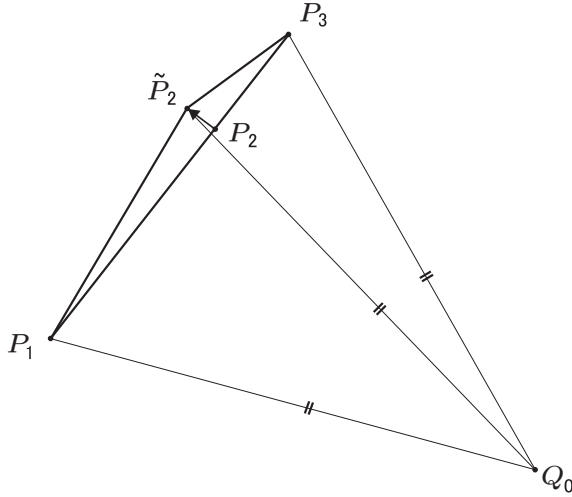


図 1: 点をずらして等距離点を得る

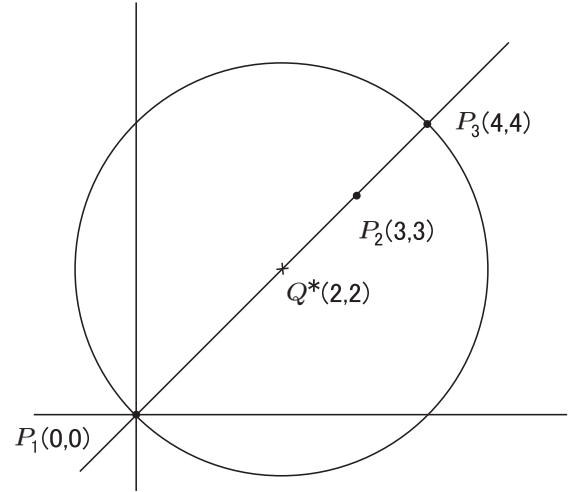


図 2: 与えられた点が直線上にある場合

## 2.2 重心座標

等距離点の位置を表すのに重心座標を利用する。 $\mathbb{R}^n$ において一般の位置にある  $m$  個の点  $P_1, \dots, P_m$  に対して、任意の点  $Q \in \mathbb{R}^n$  の  $P_1, \dots, P_m$  に関する重心座標とは

$$\vec{OQ} = \lambda_1 \vec{OP_1} + \dots + \lambda_m \vec{OP_m} \quad (2)$$

$$\sum_{i=1}^m \lambda_i = 1 \quad (3)$$

を満たす実数の組  $\lambda = (\lambda_1, \dots, \lambda_m)$  のことである。これ以降、簡単のため (2) を

$$Q = \lambda_1 P_1 + \dots + \lambda_m P_m \quad (4)$$

と表す。

$\mathbb{R}^2$  における 3 点  $P_1, P_2, P_3$  の場合、点  $Q$  が  $\triangle P_1 P_2 P_3$  の内部にあることと  $\lambda_1 > 0, \lambda_2 > 0, \lambda_3 > 0$  は同値である。また、直線  $P_1 P_3$  上で  $\lambda_2 = 0$  であり、直線  $P_1 P_3$  に関して  $P_2$  と同じ側の領域で  $\lambda_2 > 0$ 、反対側の領域で  $\lambda_2 < 0$  である。図 1 では  $Q_0$  が直線  $P_1 P_3$  に関して  $\tilde{P}_2$  と反対側にあり  $\lambda_2 < 0$  となっている。そこで、 $\lambda_2 < 0$  ならば  $P_2$  は最小包含円の内部にあるといえる。

## 2.3 例題 1

ここで、例題として、図 2 のように  $\mathbb{R}^2$  内で一直線上にある 3 点  $P_1(0,0), P_2(3,3), P_3(4,4)$  を考える。最小包含円  $\Gamma(P_1, P_2, P_3)$  は点  $Q^*(2,2)$  を中心とする半径 2 の円である。 $P_1$  と  $P_3$  が円周上にあり  $P_2$  が円の内部にあるので、 $P_2$  は最小包含円の決定に寄与しない。この寄与しない点  $P_2$  を組み合わせ的でない方法で除外するアルゴリズムを作りたい。

提案するアルゴリズムでは、各点を同等に扱って、各点をそれぞれ異なる次元の方向に少しずつずらす。例題 1 では、点の個数が 3 個なので、問題の次元を 3 上げて 5 次元空間  $\mathbb{R}^5$  内で考える。 $\mathbb{R}^5$  内で各点を少しずつ、例

えば、0.1 ずつずらして、次のように  $\tilde{P}_1, \tilde{P}_2, \tilde{P}_3$  を得る。

$$\tilde{P}_1(0, 0, 0.1, 0, 0) \quad (5)$$

$$\tilde{P}_2(3, 3, 0, 0.1, 0) \quad (6)$$

$$\tilde{P}_3(4, 4, 0, 0, 0.1) \quad (7)$$

もとの  $\mathbb{R}^2$  では  $P_1, P_2, P_3$  は一直線上にあるので、 $P_1, P_2, P_3$  からの等距離点は存在しないが、次元を上げて点を少しずつずらしたことによって、 $\mathbb{R}^5$  において  $\tilde{P}_1, \tilde{P}_2, \tilde{P}_3$  からの等距離点が存在する。 $\tilde{Q}_0 \in \mathbb{R}^5$  を等距離点とし、その  $\tilde{P}_1, \tilde{P}_2, \tilde{P}_3$  に関する重心座標を  $\lambda = (\lambda_1, \lambda_2, \lambda_3)$  とすると、

$$d(\tilde{P}_1, \tilde{Q}) = d(\tilde{P}_2, \tilde{Q}) = d(\tilde{P}_3, \tilde{Q}) \quad (8)$$

$$\tilde{Q} = \lambda_1 \tilde{P}_1 + \lambda_2 \tilde{P}_2 + \lambda_3 \tilde{P}_3 \quad (9)$$

$$\lambda_1 + \lambda_2 + \lambda_3 = 1 \quad (10)$$

が成り立つ。実際、(8),(9),(10) の唯一の解が存在して、

$$\lambda = (\lambda_1, \lambda_2, \lambda_3) = (46.6, -184.3, 138.7) \quad (11)$$

となる。 $P_2$  が除外すべき点であることが  $\lambda_2 < 0$  によって判別できる。つまり、与えられた点の位置関係についての知識に依らずに除外すべき点  $P_2$  がわかる。そして、 $P_2$  を除外して、 $P_1, P_3$  に対して

$$\tilde{P}_1(0, 0, 0.1, 0) \quad (12)$$

$$\tilde{P}_3(4, 4, 0, 0.1) \quad (13)$$

からの等距離点の重心座標  $(\lambda_1, \lambda_3) = (1/2, 1/2)$  を得る。その結果、求める最小包含円  $\Gamma(P_1, P_2, P_3)$  の中心が  $(1/2)P_1 + (1/2)P_3 = (2, 2)$  とわかる。

## 3 高次元化による等距離点の重心座標

一般に、 $\mathbb{R}^n$  内の  $m$  個の点  $P_1, \dots, P_m$  に対して、以下のように次元を上げて等距離点が存在するようにする。

各点  $P_i \in \mathbb{R}^n$  ( $i = 1, \dots, m$ ) の座標ベクトルを  $\mathbf{x}_i = (x_{i1}, \dots, x_{in})$  とし,  $\mathbb{R}^m$  の第  $i$  基本ベクトル  $\mathbf{e}_i \in \mathbb{R}^m$  を

$$\mathbf{e}_i = (0, \dots, 0, \overset{i \text{ 番目}}{1}, 0, \dots, 0), \quad i = 1, \dots, m \quad (14)$$

とする。  $P_i \in \mathbb{R}^n$  に対して  $\tilde{P}_i \in \mathbb{R}^{n+m}$  を次の座標ベクトル  $\tilde{\mathbf{x}}_i$  を持つ点として定義する。

$$\tilde{\mathbf{x}}_i = (\mathbf{x}_i, \varepsilon \mathbf{e}_i), \quad \varepsilon \in \mathbb{R}, \quad i = 1, \dots, m \quad (15)$$

(15) は成分で  $\tilde{\mathbf{x}}_i = (x_{i1}, \dots, x_{in}, 0, \dots, 0, \varepsilon, 0, \dots, 0)$  と表せる。  $\mathbb{R}^n$  内の点  $Q$  に対して,  $\mathbf{y} = (y_1, \dots, y_n)$  をその座標ベクトルとする。また,  $\mathbb{R}^m$  のベクトル  $\boldsymbol{\delta} = (\delta_1, \dots, \delta_m)$  を考える。点  $Q \in \mathbb{R}^n$  に対して  $\tilde{Q} \in \mathbb{R}^{n+m}$  を座標ベクトル  $\tilde{\mathbf{y}} = (\mathbf{y}, \boldsymbol{\delta})$  を持つ点として定義する。

$P_1, \dots, P_m$  からの等距離点が存在しない場合でも, 上記のように次元を上げることによって,  $\varepsilon \neq 0$  ならば必ず  $\tilde{P}_1, \dots, \tilde{P}_m$  からの等距離点が存在する。それを以下に示す。

等距離点  $\tilde{Q}$  は

$$d(\tilde{P}_i, \tilde{Q}) = d(\tilde{P}_1, \tilde{Q}), \quad i = 2, \dots, m \quad (16)$$

を満たす。(16) の両辺を 2 乗してベクトルで表すと,

$$2(\mathbf{x}_i - \mathbf{x}_1)^t \mathbf{y} + 2\varepsilon(\delta_i - \delta_1) = \|\mathbf{x}_i\|^2 - \|\mathbf{x}_1\|^2 \quad (17)$$

を得る。ここで,  ${}^t \mathbf{y}$  はベクトル  $\mathbf{y}$  の転置を表し,  $\|\mathbf{x}_i\|^2$  はベクトル  $\mathbf{x}_i$  のノルムの 2 乗, すなわち  $\|\mathbf{x}_i\|^2 = \mathbf{x}_i^t \mathbf{x}_i = \sum_{j=1}^n x_{ij}^2$  である。

さて,  $m \times n$  行列  $A$ ,  $(m-1) \times m$  行列  $J$ , および  $m$  次元行ベクトル  $\mathbf{c}$  を次のように定義する。

$$A = \begin{pmatrix} \mathbf{x}_1 \\ \vdots \\ \mathbf{x}_m \end{pmatrix} = \begin{pmatrix} x_{11} & \cdots & x_{1n} \\ \vdots & & \vdots \\ x_{m1} & \cdots & x_{mn} \end{pmatrix} \quad (18)$$

$$J = \begin{pmatrix} -1 & 1 & 0 & \cdots & 0 \\ -1 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -1 & 0 & 0 & \cdots & 1 \end{pmatrix} \quad (19)$$

$$\mathbf{c} = (\|\mathbf{x}_1\|^2, \dots, \|\mathbf{x}_m\|^2) \quad (20)$$

このとき (17) は次のように書ける。

$$2JA^t \mathbf{y} + 2\varepsilon J^t \boldsymbol{\delta} = J^t \mathbf{c} \quad (21)$$

次に,  $m \times (n+m)$  行列  $\tilde{A}$  を  $\tilde{A} = (A, \varepsilon I_m)$  と定義する。ここで,  $I_m$  は  $m \times m$  単位行列である。等距離点  $\tilde{Q}$  の  $\tilde{P}_1, \dots, \tilde{P}_m$  に関する重心座標を  $\boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_m) \in \mathbb{R}^m$  とする。 $\tilde{Q}$  の座標ベクトルは  $\tilde{\mathbf{y}} = (\mathbf{y}, \boldsymbol{\delta})$  であり,  $\tilde{P}_i$  の座標ベクトルは (15) より  $\tilde{\mathbf{x}}_i = (\mathbf{x}_i, \varepsilon \mathbf{e}_i)$  なので,

$$\tilde{\mathbf{y}} = \boldsymbol{\lambda} \tilde{A} \quad (22)$$

$$\boldsymbol{\lambda}^t \mathbf{1}_m = 1 \quad (23)$$

を得る。ここで,  $\mathbf{1}_m = (1, \dots, 1) \in \mathbb{R}^m$  とおいた。(22) より

$$\mathbf{y} = \boldsymbol{\lambda} A, \quad \boldsymbol{\delta} = \varepsilon \boldsymbol{\lambda} \quad (24)$$

となる。(24) を (21) に代入して

$$\boldsymbol{\lambda} (2A^t A + 2\varepsilon^2 I_m)^t J = \mathbf{c}^t J \quad (25)$$

となる。ここで,  $M = 2A^t A + 2\varepsilon^2 I_m$  とおくと, (25) は

$$(\boldsymbol{\lambda} M - \mathbf{c})^t J = \mathbf{0} \quad (26)$$

と書ける。任意の  $\varepsilon \neq 0$  に対して  $M$  は対称な正定値行列であり, 特に正則である。ベクトル  $\mathbf{x}_i$  の次元を上げて  $\tilde{\mathbf{x}}_i = (\mathbf{x}_i, \varepsilon \mathbf{e}_i)$  とした効果で  $M$  が正則になっている。

そこで,  $\hat{\boldsymbol{\lambda}} = \mathbf{c} M^{-1}$  とおくと, 連立方程式 (23), (26) の解  $\boldsymbol{\lambda}$  が計算によって次のように得られる。

$$\boldsymbol{\lambda} = \hat{\boldsymbol{\lambda}} + \frac{1 - \hat{\boldsymbol{\lambda}}^t \mathbf{1}_m}{\mathbf{1}_m^t M^{-1} \mathbf{1}_m} \mathbf{1}_m M^{-1} \quad (27)$$

この (27) が等距離点  $\tilde{Q}$  の重心座標である。

以上のようにして  $\tilde{P}_1, \dots, \tilde{P}_m$  に対して唯一つ存在する等距離点  $\tilde{Q}$  を

$$\tilde{Q} = \mathcal{F}(\tilde{P}_1, \dots, \tilde{P}_m) \quad (28)$$

と表し,  $\tilde{Q}$  の  $\tilde{P}_1, \dots, \tilde{P}_m$  に関する重心座標 (27) を

$$\boldsymbol{\lambda} = \boldsymbol{\lambda}(\tilde{Q}; \tilde{P}_1, \dots, \tilde{P}_m) \quad (29)$$

と表す。

#### 4 最小包含円についての解析

$\mathbb{R}^n$  内の  $m$  個の点  $P_1, \dots, P_m$  に対する最小包含円  $\Gamma(P_1, \dots, P_m)$  は (1) の min max 問題を解くことによって得られる。(1) は距離  $d(P_i, Q)$  によって問題を表現していて図形的に理解しやすく, したがって, 幾何的な考察によって解法アルゴリズムを開発することができる。実際に本論文では (1) に基づいて新しいアルゴリズムを考えている。

しかしながら, アルゴリズムが正しいことを証明するのに (1) のような二重最適化問題を利用することは困難である。そこで, (1) と等価な単純最適化問題としての凸最適化問題 (convex programming) が以下のように得られる。

$m$  次元シンプレックス  $\bar{\Delta}^m$  を

$$\bar{\Delta}^m = \left\{ \boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_m) \mid \lambda_i \geq 0, \quad i = 1, \dots, m, \quad \sum_{i=1}^m \lambda_i = 1 \right\}$$

と定義する。 $\bar{\Delta}^m$  は  $m$  次元確率ベクトルの全体といってもよい。そこで,  $\boldsymbol{\lambda} \in \mathbb{R}^m$  に対して

$$f(\boldsymbol{\lambda}) = \boldsymbol{\lambda}^t \mathbf{c} - \boldsymbol{\lambda} A^t A \boldsymbol{\lambda} \quad (30)$$

とおく。 $f: \mathbb{R}^m \rightarrow \mathbb{R}$  は微分可能で上に凸な関数である。下記 (31) の凸最適化問題は最小包含円問題 (1) と等価である [8]。

$$\max_{\lambda \in \Delta^m} f(\lambda) \quad (31)$$

(31) を達成する  $\lambda = \lambda^*$  が最小包含円  $\Gamma(P_1, \dots, P_m)$  の中心  $Q^*$  の  $P_1, \dots, P_m$  に関する重心座標であり、 $\sqrt{f(\lambda^*)}$  が  $\Gamma(P_1, \dots, P_m)$  の半径である。

(31) の形の凸最適化問題に関して次の Kuhn-Tucker 条件が成り立つ [4]。

Kuhn-Tucker 条件

$\mathbb{R}^n$  の点  $P_1, \dots, P_m$  に関する  $Q^* \in \mathbb{R}^n$  の重心座標を  $\lambda^* = (\lambda_1^*, \dots, \lambda_m^*) \in \Delta^m$  とする。ある  $d_0$  が存在して、

$$d(P_i, Q^*) \begin{cases} = d_0, & \text{for } i \text{ with } \lambda_i^* > 0 \\ \leq d_0, & \text{for } i \text{ with } \lambda_i^* = 0 \end{cases} \quad (32)$$

を満たすことが、 $Q^*$  が最小包含円  $\Gamma(P_1, \dots, P_m)$  の中心であるための必要十分条件であり、 $d_0$  がその半径である。

逆に、最小包含円問題の解は (32) を満たす。すなわち、(1), (31), (32) は同値である。最小包含円の証明には (32) が使い易い。

## 5 最小包含円を計算するアルゴリズム [予想]

次の例題を考える。

### 5.1 例題 2

$\mathbb{R}^2$  の 4 点  $P_1(0, 0), P_2(10, 10), P_3(11, 0), P_4(0, 8)$  を考える (図 3)。 $\varepsilon = 0.1$  として  $\tilde{P}_1, \dots, \tilde{P}_4$  からの等距離点  $\tilde{Q}^0$  の重心座標  $\lambda^0$  を (29) によって計算すると

$$\lambda^0 = (122.7, 105.9, -95.8, -131.8) \quad (33)$$

を得る。 $\lambda_3^0 = -95.8 < 0, \lambda_4^0 = -131.8 < 0$  であるが、これらに対応する点  $P_3, P_4$  を除外してしまうと正しい結果が得られない。そこで、最小 (すなわち、負で絶対値が最大) の  $\lambda_4^0$  に対応する  $P_4$  を除外して、 $P_1, P_2, P_3$  に対して、 $\varepsilon = 0.1$  として (29) によって等距離点  $\tilde{Q}^1$  の重心座標  $\lambda^1$  を計算すると  $\lambda^1 = (0.46, 0.45, 0.09)$  となり、成分はすべて正である。したがって、Kuhn-Tucker 条件より  $\tilde{Q}^1 = \lambda^1 \tilde{A}$  が最小包含円  $\Gamma(\tilde{P}_1, \tilde{P}_2, \tilde{P}_3)$  の中心であることがわかる。

そこで、元の問題において初めから  $P_4$  を除外して、 $\Gamma(P_1, P_2, P_3)$  の中心  $Q^*$  を (28) によって ( $\varepsilon = 0$  として) 計算すると  $Q^* = \lambda^1 A = (5.5, 4.5)$  を得る (図 3)。そして、最後に

$$d(P_4, Q^*) = \sqrt{85/2} < \sqrt{101/2} = d(P_1, Q^*) \quad (34)$$

が確認できるので、Kuhn-Tucker 条件によって  $Q^* = \lambda^1 A = (5.5, 4.5)$  が求める最小包含円  $\Gamma(P_1, P_2, P_3, P_4)$  の中心であることがわかる。

この例題に基づいて下記の計算アルゴリズムを提案する。しかし、提案するアルゴリズムによって必ず最小包含円が得られることは証明できていない。

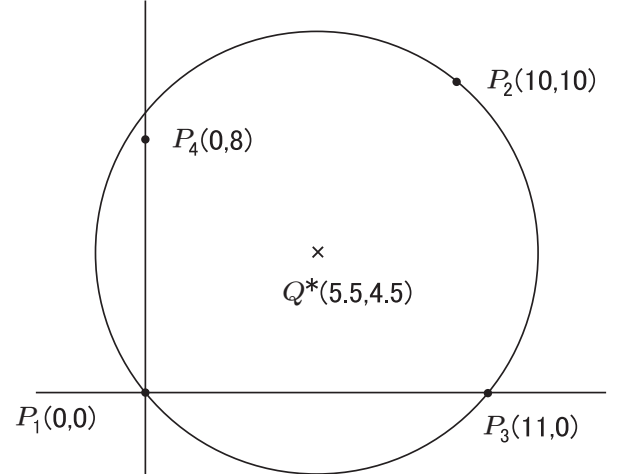


図 3: 負の重心座標をすべて除外できない例

### 5.2 最小包含円を計算するアルゴリズム [予想]

与えられた点  $P_1, \dots, P_m$  に対して、それらの次元を上げた点  $\tilde{P}_1, \dots, \tilde{P}_m$  からの等距離点を  $\tilde{Q}^0 = \mathcal{F}(\tilde{P}_1, \dots, \tilde{P}_m)$  とし、 $\tilde{Q}^0$  の重心座標を  $\lambda^0 = \lambda(\tilde{Q}^0; \tilde{P}_1, \dots, \tilde{P}_m)$  とする。 $\lambda^0$  を成分で表して  $\lambda^0 = (\lambda_1^0, \dots, \lambda_m^0)$  とする。

[Case 1]  $\lambda_i^0 \geq 0, i = 1, \dots, m$  の場合

この場合、Kuhn-Tucker 条件によって  $\tilde{Q}^0$  が最小包含円  $\Gamma(\tilde{P}_1, \dots, \tilde{P}_m)$  の中心であり、全ての点  $\tilde{P}_1, \dots, \tilde{P}_m$  が  $\Gamma(\tilde{P}_1, \dots, \tilde{P}_m)$  の周上にある。 $\varepsilon \neq 0$  で  $|\varepsilon|$  が十分小さければ、 $\Gamma(\tilde{P}_1, \dots, \tilde{P}_m)$  は  $\Gamma(P_1, \dots, P_m)$  に近く、 $P_1, \dots, P_m$  がすべて最小包含円に寄与する。

[Case 2]  $\lambda_i^0, i = 1, \dots, m$  の中に負のものがある場合  
この場合、(必要なら順番を並べ替えて)  $\lambda_1^0 \geq \dots \geq \lambda_m^0$  とする。このとき、 $\lambda_m^0$  は最小 (すなわち負で絶対値が最大) である。番号  $m$  を除外して  $\tilde{P}_1, \dots, \tilde{P}_{m-1}$  を考え、

$$\tilde{Q}^1 = \mathcal{F}(\tilde{P}_1, \dots, \tilde{P}_{m-1}), \lambda^1 = \lambda(\tilde{Q}^1; \tilde{P}_1, \dots, \tilde{P}_{m-1})$$

とする。 $\lambda^1$  のすべての成分が 0 以上であれば、上記の Case 1 となり、最小包含円が決まり、アルゴリズムが終了する。そうでない場合は Case 2 となり、最小の成分の番号を除外して (必要なら順番を並べ替えて)  $\tilde{P}_1, \dots, \tilde{P}_{m-2}$  を考える。以下同様にして、Case 2 が続くとして、 $k = 0, 1, \dots$  として

$$\tilde{Q}^k = \mathcal{F}(\tilde{P}_1, \dots, \tilde{P}_{m-k}), \lambda^k = \lambda(\tilde{Q}^k; \tilde{P}_1, \dots, \tilde{P}_{m-k})$$

を得る。 $k = m-2$  のとき 2 点となり  $\lambda^{m-2} = (1/2, 1/2)$  となるのでアルゴリズムは高々  $m-2$  ステップで終了する。

以上のアルゴリズムによって必ず正しく最小包含円が得られると予想する。

[予想] ある  $\varepsilon$  に対して、上記のアルゴリズムが終了したときの  $\tilde{Q}^k$  が最小包含円  $\Gamma(\tilde{P}_1, \dots, \tilde{P}_m)$  の中心であり、 $\lambda^k$  がその重心座標である。 $\tilde{P}_1, \dots, \tilde{P}_{m-k}$  が最小包含円の決定に寄与する点である。

この予想が正しいことを証明するためには、下記のことを証明すればよい。

$k$  ステップでアルゴリズムが終了しているの、その段階で点  $\tilde{P}_1, \dots, \tilde{P}_{m-k}$  が残っていて、それらからの等距離点  $\tilde{Q}^k$  が得られている。 $\tilde{Q}^k$  の  $\tilde{P}_1, \dots, \tilde{P}_{m-k}$  に関する重心座標  $\lambda^k = (\lambda_1^k, \dots, \lambda_{m-k}^k)$  は  $\lambda_i^k \geq 0, i = 1, \dots, m-k$  を満たしている。そして、 $d(\tilde{P}_1, \tilde{Q}^k) = \dots = d(\tilde{P}_{m-k}, \tilde{Q}^k)$  が成り立っている。 $\tilde{P}_1, \dots, \tilde{P}_{m-k}$  以外の点、つまり、 $0 \sim k-1$  ステップで除外された  $k$  個の点に改めて番号を付けて  $\tilde{P}_{m-k+1}, \dots, \tilde{P}_m$  とする。このとき、

$$d(\tilde{P}_i, \tilde{Q}^k) \leq d(\tilde{P}_1, \tilde{Q}^k), i = m-k+1, \dots, m \quad (35)$$

であることが証明できればよい。

### 5.3 例題 3 ( $\varepsilon$ の値によってはうまくいかない例)

4 点  $P_1(-10, -9), P_2(6, 5), P_3(-9, -10), P_4(6, -7)$  に対する最小包含円  $\Gamma(P_1, P_2, P_3, P_4)$  の中心は  $Q^*(-2, -2)$  であり、 $P_1, P_2, P_3$  が周上にあり、 $P_4$  が内部にある。ここで、 $\varepsilon = 0.1$  とすると  $\lambda = (-91.6, 9.8, 98.3, -15.5)$  となり、周上にある  $P_1$  を除外してしまうことになる。しかし、この例題でも、 $\varepsilon = 1.8$  とすると、 $\lambda = (0.02, 0.62, 0.57, -0.21)$  となり、 $P_4$  が除外されてうまく行く。

$n = 2, m = 4$  の場合に、 $-10$  以上  $10$  以下の整数を座標成分とする 4 点をランダムに生成して試したところ、 $\varepsilon = 0.1$  のとき 1000 パターン中 18 パターンで失敗した。また、 $\varepsilon = 1$  のとき、1000 パターン中 7 パターンで失敗した。

## 6 通信路容量を達成する入力シンボルについて

$X$  を入力情報源とし、 $Y$  を出力情報源とする離散的無記憶通信路を考える。入力記号を  $\{a_1, \dots, a_m\}$  とし、出力記号を  $\{b_1, \dots, b_n\}$  とする。 $a_i$  を送信したとき  $b_j$  が受信される条件付き確率を

$$P_{ij} = P(Y = b_j | X = a_i), i = 1, \dots, m, j = 1, \dots, n$$

として、通信路行列を

$$P = \begin{pmatrix} P_1 \\ \vdots \\ P_m \end{pmatrix} = \begin{pmatrix} P_{11} & \dots & P_{1n} \\ \vdots & & \vdots \\ P_{m1} & \dots & P_{mn} \end{pmatrix} \quad (36)$$

$$P_i = (P_{i1}, \dots, P_{in}), i = 1, \dots, m \quad (37)$$

とする。

$\mathbf{p} = (p_1, \dots, p_m) \in \bar{\Delta}^m$  を入力分布とし、 $\mathbf{q} = \mathbf{p}P, \mathbf{q} = (q_1, \dots, q_n) \in \bar{\Delta}^n$  を出力分布とする。このとき、相互情報量  $I(\mathbf{p}, P) = \sum_i p_i \sum_j P_{ij} \log P_{ij}/p_i q_j$  に対して通信路容量  $C$  が

$$C = \max_{\mathbf{p} \in \bar{\Delta}^m} I(\mathbf{p}, P) \quad (38)$$

と定義される [4]。また、 $C$  は次のように KL 情報量によって定式化される [3]。

$$C = \min_{\mathbf{q} \in \bar{\Delta}^n} \max_{1 \leq i \leq m} D(P_i \| \mathbf{q}) \quad (39)$$

### Kuhn-Tucker 条件

入力確率分布  $\mathbf{p}$  が通信路容量  $C$  を達成するための必要十分条件は、ある値  $C_0$  が存在して、

$$D(P_i \| \mathbf{p}P) \begin{cases} = C_0, & \text{for } i \text{ with } p_i > 0 \\ \leq C_0, & \text{for } i \text{ with } p_i = 0 \end{cases} \quad (40)$$

を満たすことである。このとき、 $C_0$  は  $C$  に等しい。

上記の (38), (39), (40) は、最小包含円の問題における (31), (1), (32) にそれぞれ対応している。

### 6.1 通信路行列の次元を上げる方法

通信路行列  $P$  に対して、その次元を上げた通信路行列

$$\tilde{P} = \begin{pmatrix} \tilde{P}_1 \\ \vdots \\ \tilde{P}_m \end{pmatrix}, \tilde{P}_i \in \bar{\Delta}^{n+2m}, i = 1, \dots, m \quad (41)$$

を次のように定義する。

$$\tilde{P}_i = \left( \frac{P_{i1}}{2m+1}, \dots, \frac{P_{in}}{2m+1}, \overbrace{\frac{1}{2m+1}, \dots, \frac{1}{2m+1}}^{2i-2 \text{ 個}}, \right. \\ \left. \frac{1+\varepsilon}{2m+1}, \frac{1-\varepsilon}{2m+1}, \overbrace{\frac{1}{2m+1}, \dots, \frac{1}{2m+1}}^{2m-2i \text{ 個}} \right) \quad (42)$$

ここで、 $\varepsilon$  は  $|\varepsilon|$  が十分小さい実数で  $\varepsilon \neq 0$  とする。 $m \times 2m$  行列  $R$  を

$$R = \frac{1}{2m} \begin{pmatrix} 1+\varepsilon & 1-\varepsilon & 1 & 1 & \dots & 1 & 1 \\ 1 & 1 & 1+\varepsilon & 1-\varepsilon & \dots & 1 & 1 \\ \vdots & \vdots & & \ddots & \ddots & \vdots & \vdots \\ 1 & 1 & 1 & 1 & \dots & 1+\varepsilon & 1-\varepsilon \end{pmatrix} \quad (43)$$

と定義し  $R$  の第  $i$  行を  $R_i, i = 1, \dots, m$  とおくと (42) は

$$\tilde{P}_i = \left( \frac{1}{2m+1} P_i, \frac{2m}{2m+1} R_i \right), i = 1, \dots, m \quad (44)$$

と書ける。

$\mathbf{q} = (q_1, \dots, q_n) \in \bar{\Delta}^n, \boldsymbol{\delta} = (\delta_1, \dots, \delta_{2m}) \in \bar{\Delta}^{2m}$  に対して、 $\tilde{\mathbf{q}} \in \bar{\Delta}^{n+2m}$  を

$$\tilde{\mathbf{q}} = \left( \frac{1}{2m+1} \mathbf{q}, \frac{2m}{2m+1} \boldsymbol{\delta} \right) \quad (45) \\ = \left( \frac{q_1}{2m+1}, \dots, \frac{q_n}{2m+1}, \frac{2m\delta_1}{2m+1}, \dots, \frac{2m\delta_{2m}}{2m+1} \right)$$

と定義する。 $\tilde{\mathbf{q}}$  を  $\tilde{P}_1, \dots, \tilde{P}_m$  からの KL 情報量が等しい分布とする。すなわち、

$$D(\tilde{P}_i \| \tilde{\mathbf{q}}) = D(\tilde{P}_1 \| \tilde{\mathbf{q}}), i = 2, \dots, m \quad (46)$$

である。計算により

$$D(\tilde{P}_i \| \tilde{q}) = \frac{1}{2m+1} \left\{ -H(P_i) - \sum_{j=1}^n P_{ij} \log q_j - \sum_{k=1}^{2m} \log(2m\delta_k) - \hat{H}(\epsilon) - \epsilon(\log \delta_{2i-1} - \log \delta_{2i}) \right\} \quad (47)$$

となる。ここで、 $H(P_i) = -\sum_{j=1}^n P_{ij} \log P_{ij}$ ,  $\hat{H}(\epsilon) = -(1+\epsilon) \log(1+\epsilon) - (1-\epsilon) \log(1-\epsilon)$  とおいた。

(46),(47) より

$$\sum_{j=1}^n (P_{ij} - P_{1j}) \log q_j + \epsilon(\log \delta_{2i-1} - \log \delta_{2i}) - \epsilon(\log \delta_1 - \log \delta_2) = -H(P_i) + H(P_1) \quad (48)$$

$$i = 2, \dots, m$$

を得る。

次に、 $\tilde{q}$  の  $\tilde{P}_1, \dots, \tilde{P}_m$  に関する重心座標を  $\mathbf{p} = (p_1, \dots, p_m)$  とする。すなわち、

$$\tilde{q} = \sum_{i=1}^m p_i \tilde{P}_i = \mathbf{p} \tilde{P}, \quad (49)$$

$$\sum_{i=1}^m p_i = 1 \quad (50)$$

である。 $p_i \geq 0, i = 1, \dots, m$  ならば  $\mathbf{p} \in \bar{\Delta}^n$  となり、 $\mathbf{p}$  は入力確率分布になる。

(44),(45),(49) より

$$\mathbf{q} = \mathbf{p} \mathbf{P}, \quad \delta = \mathbf{p} \mathbf{R} \quad (51)$$

となる。(48),(51) より

$$\sum_{j=1}^n (P_{ij} - P_{1j}) \log \sum_{i'=1}^m p_{i'} P_{i'j} + \epsilon \frac{1 + \epsilon p_i}{1 - \epsilon p_i} - \epsilon \frac{1 + \epsilon p_1}{1 - \epsilon p_1} = -H(P_i) + H(P_1), \quad i = 2, \dots, m \quad (52)$$

を得る。

連立方程式 (50),(52) は未知数が  $p_1, \dots, p_m$  の  $m$  個であり、方程式が  $m$  個なので解くことができる。

## 6.2 通信路容量を達成する入力シンボルを計算するアルゴリズム [予想]

通信路容量を求める問題においても第5節で提案した計算アルゴリズムが適用できると予想する。

## 6.3 例題 4

通信路行列

$$\mathbf{P} = \begin{pmatrix} P_1 \\ P_2 \\ P_3 \end{pmatrix} = \begin{pmatrix} 0.1 & 0.9 \\ 0.7 & 0.3 \\ 0.8 & 0.2 \end{pmatrix} \quad (53)$$

に対して  $\epsilon = 0.1$  として方程式 (50),(52) を解いて

$$\mathbf{p} = (1.03, -3.44, 3.41) \quad (54)$$

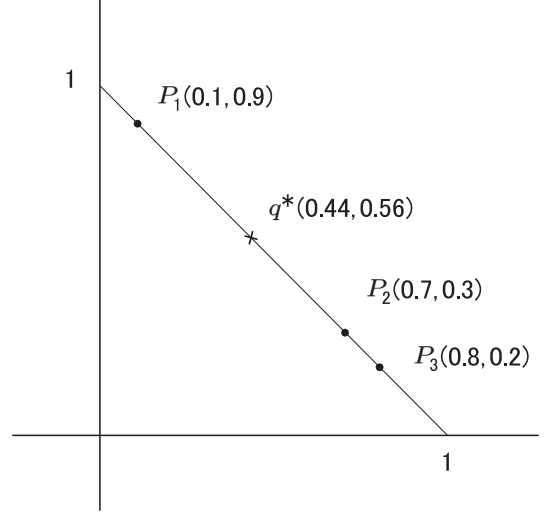


図 4:  $n = 2, m = 3$  の通信路の例

を得る。 $p_2 = -3.44 < 0$  なので、 $P_2$  が通信路容量の決定に寄与しないと判断できる。実際、図 4 より、通信路容量を達成する  $\mathbf{q}^*$  は  $D(P_1 \| \mathbf{q}^*) = D(P_3 \| \mathbf{q}^*)$  を満たす分布で、計算によって  $\mathbf{q}^*(0.44, 0.56)$  となる。 $P_2$  は  $C$  の達成に寄与せず、 $P_1$  と  $P_3$  が通信路容量を決定する。

## 7 まとめ

ユークリッド空間における最小包含円  $\Gamma$  および離散的無記憶通信路の通信路容量  $C$  を求める問題を幾何的に考察して両者に類似の構造を示した。その幾何的な考察に基づいて  $\Gamma$  および  $C$  の決定に寄与する点あるいは入力記号を得るためのアルゴリズムを提案したが、その正しさが証明できていない。今後の研究課題としたい。

## 参考文献

- [1] Arimoto, S., "An algorithm for computing the capacity of arbitrary discrete memoryless channels," IEEE Trans. Inform. Theory, vol.IT-18, pp.14-20, 1972.
- [2] Cheng, D., Hu, X and Martin, C., "On the Smallest Enclosing Balls," Communications in Information and Systems, vol.6, no.2, pp.137-160, 2006.
- [3] Csizár, I. and Körner, J., *Information Theory: Coding Theorems for Discrete Memoryless Systems*, Academic Press, Orlando, 1982.
- [4] Gallager, G., *Information Theory and Reliable Communication*, Wiley, New York, 1968.
- [5] 北山義明, 川端勉, "縦長通信路の通信路容量計算法について," SITA2013 予稿集, pp.87-91, 2013.
- [6] Muroga, S., "On the capacity of a discrete channel. I," J. Phys. Soc. Japan, vol.8, pp.484-494, 1953.
- [7] Nakagawa, K. and Kanaya, F., "A New Geometric Capacity Characterization of a Discrete Memoryless Channel," IEEE Trans. Inform. Theory, vol.34, no.2, pp.318-321, 1988.
- [8] Schönherr, S., "Quadratic Programming in Geometric Optimization: Theory, Implementation, and Applications," Dissertation of Swiss Federal Institute of Technology, <http://www.inf.ethz.ch/personal/emo/DoctThesisFiles/schoenherr02.pdf>