

通信路容量を達成する出力分布の射影アルゴリズムによる探索について

On the Search by the Projection Algorithm for the Output Distribution that Achieves the Channel Capacity

中川 健治*
Kenji Nakagawa

渡部 康平*
Kohei Watabe

佐武 拓斗*
Takuto Sabu

Abstract— We consider the search algorithm for the output distribution that achieves the channel capacity of a discrete memoryless channel. This algorithm is the iterated projections of an output distribution onto affine subspaces in the set of the output distributions. The problem of the channel capacity has a similar geometric structure as that of the smallest enclosing circle for finite number of points in the Euclidean space. The metric in the Euclidean space is the Euclidean distance and the metric in the space of the output distributions is the Kullback-Leibler divergence. Then, we first consider the smallest enclosing circle in the Euclidean space and develop an algorithm to find the center of the smallest enclosing circle. Based on the investigation, we will apply the obtained algorithm to the problem of the channel capacity.

Keywords— channel capacity, discrete memoryless channel, smallest enclosing circle, information geometry, projection algorithm

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 問題の解としても定式化される [5],[8]。通信路容量に関する min max 問題における KL 情報量をユークリッド距離に置き換えると、ユークリッド空間における類似の問題が得られる。それは、有限点集合に対する最小包含円の問題である。本論文では、ユークリッド空間における最小包含円問題を幾何的に考察し、ユークリッド距離に関する min max 問題の解を計算するアルゴリズムを開発する。そして得られたアルゴリズムを通信路容量に関する min max 問題に適用して、通信路容量を達成する出力分布を計算するアルゴリズムを作る。このようなアプローチをとる理由は、ユークリッド幾何は我々にとってなじみがある幾何なので、新しい幾何的なアルゴリズムを作りやすいからである。

通信路容量の問題は出力分布に関する最適化問題であり、考える空間はユークリッド空間ではなく確率分布全体がなす集合である。その集合上の幾何学は情報幾何である。ユークリッド幾何と情報幾何は甘利 [1] の α 幾何の立場から統一的に考えることができる。したがって、最小包含円問題の解法アルゴリズムを、幾何的な類似性を通して、通信路容量問題に適用できる。その際に重要なことは、最小包含円問題を考察するのにユークリッド

幾何のどんな性質を使ってもよいということではなく、通信路容量問題に適用可能な性質だけを使う必要がある、ということである。それが実現できれば、最小包含円問題で得られたアルゴリズムをほぼ自動的に通信路容量問題に移植することができる。本論文では、実際に両方の幾何に共通の性質として、距離、重心座標、内積、ピタゴラスの定理、アファイン部分空間への射影のみを使ってアルゴリズムを開発する。このアルゴリズムを本論文では「射影アルゴリズム」という。さらに、通信路行列の rank が m 未満のときには通信路容量の計算に組み合わせ的な困難さが生じるが、本論文ではその困難さを緩和する新たなアルゴリズムを提案し、数値的な結果を示す。

通信路容量の計算法は大きく分けて、Muroga [7] による方程式による解法と Arimoto [2] による逐次計算法がある。Muroga [7] では、Lagrange 未定乗数法から得られる方程式を直接解くことによって相互情報量を最大化する入力分布を計算する。その際に重要なことは、確率は 0 以上 1 以下であるという不等式条件である。単に方程式を解くと、確率が負となる解が現れてしまう。不等式条件が問題を難しいものになっているが、それ故に問題が魅力的なものになっている。Muroga [7] では、ある確率が負となる解が得られたときには、入力シンボルを 1 個ずつ除いて方程式を解いてすべての確率が非負になるまで繰り返し計算を続ける。したがって、すべての場合を尽くすとその総数は組み合わせ的なものになってしまう。Arimoto [2] による逐次近似法では任意の通信路行列に対して通信路容量を漸化式によって数値的に計算することができる。その方法では確率が負になることはないが、それに対応することとして、ある入力確率が 0 になると、逆向き通信路の計算において例外的な対応が必要になる。Muroga [7], Arimoto [2] のどちらも、どの入力シンボルの確率が正になるかを決定することが重要である。本論文ではそのことを主題としている。

1.1 通信路行列と通信路容量の定式化

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 $X \rightarrow Y$ を考える。入力アルファベットを $\{x_1, \dots, x_m\}$ とし、出力アルファベットを $\{y_1, \dots, y_n\}$ とする。 x_i を送信したときに y_j が受信される条件付き確率 P_{ij} を

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

* 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@nagaokaut.ac.jp

とし、行ベクトル P_i を

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

とする。通信路行列 Φ が

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

と定義される。

入力アルファベット $\{x_1, \dots, x_m\}$ 上の入力分布全体 $\bar{\Delta}^m$ を

$$\bar{\Delta}^m = \{\lambda = (\lambda_1, \dots, \lambda_m) | \lambda_i \geq 0, i = 1, \dots, m, \sum_{i=1}^m \lambda_i = 1\} \quad (2)$$

と定義する。同様に、出力アルファベット $\{y_1, \dots, y_n\}$ 上の出力分布全体 $\bar{\Delta}^n$ を

$$\bar{\Delta}^n = \{Q = (Q_1, \dots, Q_n) | Q_j \geq 0, j = 1, \dots, n, \sum_{j=1}^n Q_j = 1\} \quad (3)$$

と定義する。 $\lambda \in \bar{\Delta}^m$ を入力分布とするときの出力分布 $Q \in \bar{\Delta}^n$ が $Q = \lambda\Phi$, すなわち, $Q_j = \sum_{i=1}^m \lambda_i P_{ij}$, $j = 1, \dots, n$ と定義され, 相互情報量 $I(\lambda, \Phi)$ が

$$I(\lambda, \Phi) = \sum_{i=1}^m \sum_{j=1}^n \lambda_i P_{ij} \log \frac{P_{ij}}{Q_j} \quad (4)$$

と定義される。このとき、通信路容量 C が

$$C = \max_{\lambda \in \bar{\Delta}^m} I(\lambda, \Phi) \quad (5)$$

と定義される。

$Q = (Q_j)_{j=1, \dots, n}$, $Q' = (Q'_j)_{j=1, \dots, n} \in \bar{\Delta}^n$ に対して KL 情報量 $D(Q \| Q')$ が

$$D(Q \| Q') \equiv \sum_{j=1}^n Q_j \log \frac{Q_j}{Q'_j} \quad (6)$$

と定義される [5]。 C は KL 情報量によって次のように表すこともできる [5]。

$$C = \min_{Q \in \bar{\Delta}^n} \max_{1 \leq i \leq m} D(P_i \| Q) \quad (7)$$

通信路行列によっては (5) を達成する入力分布 λ は必ずしも一意的不是であるが, (7) を達成する出力分布 Q は一意的不である [5]。その一意性のために, (7) に基づいて, KL 情報量 $D(P \| Q)$ の幾何的性質を利用して通信路容量 C を達成する出力分布 Q を求める方法は考え易い。一方, 得られた出力分布が実際に C を達成することを証明するには (7) による幾何的考察よりも (5) の凸

最適化問題による方が考え易い。実際, (5) に関して次の Kuhn-Tucker 条件が利用できる [6]。

定理 (通信路容量問題に関する Kuhn-Tucker 条件) 入力分布 $\lambda^* = (\lambda_1^*, \dots, \lambda_m^*) \in \bar{\Delta}^m$ 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,

$$D(P_i \| \lambda^* \Phi) \begin{cases} = C_0, & \text{for } i \text{ with } \lambda_i^* > 0 \\ \leq C_0, & \text{for } i \text{ with } \lambda_i^* = 0 \end{cases} \quad (8)$$

を満たすことである。このとき, C_0 は C に等しい。

さて, (7) における $D(P_i \| Q)$ をユークリッド距離に置き換えると, \mathbb{R}^n における類似の問題を考えることができる。すなわち, $P_1, \dots, P_m \in \mathbb{R}^n$ に対して,

$$\min_{Q \in \mathbb{R}^n} \max_{1 \leq i \leq m} d(P_i, Q) \quad (9)$$

を考える。ここで, $d(P_i, Q)$ は点 P_i と Q とのユークリッド距離を表す。これは点集合 $\{P_1, \dots, P_m\}$ に対する最小包含円の問題である。本研究の目的は, 最小包含円問題を幾何的に考察して, そこから得られる最適解の探索アルゴリズムを, (7) と (9) の類似性を通して, 通信路容量問題に応用して通信路容量を達成する分布の探索アルゴリズムを得ることである。

(7) と (9) は形式的に類似しているだけでなく, $\bar{\Delta}^n$ と \mathbb{R}^n は甘利の情報幾何 [1] の観点から統一的に議論できる幾何構造を持っている。そこで, 両者に共通の性質だけを利用してアルゴリズムを開発し, それによって得られた結果を証明できれば, 一方の結果から他方が自動的に得られる。実際, そのようになっていることが本論文で明らかになる。

そこで, まず, ユークリッド空間における最小包含円問題について考えよう。

2 ユークリッド空間における最小包含円問題

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に含む球のうち半径が最小のものを**最小包含円**といい, これを $\Gamma(P_1, \dots, P_m)$ と表す。最小包含円 $\Gamma(P_1, \dots, P_m)$ は (9) によって定式化される。(9) を達成する $Q = Q^*$ が $\Gamma(P_1, \dots, P_m)$ の中心であり, $d^* = \max_{1 \leq i \leq m} d(P_i, Q^*)$ がその半径である。

2.1 等距離点と射影アルゴリズム

最も簡単な例として, \mathbb{R}^2 の 3 点 P_1, P_2, P_3 に対する最小包含円 $\Gamma = \Gamma(P_1, P_2, P_3)$ の中心 Q^* と半径 d^* について考える。 $\triangle P_1 P_2 P_3$ が鋭角三角形の場合と鈍角三角形の場合に分けて考える。

(I) $\triangle P_1 P_2 P_3$ が鋭角三角形の場合 (図 1 参照)

この場合, $\triangle P_1 P_2 P_3$ の外心, すなわち P_1, P_2, P_3 から等距離にある点 Q^0 が最小包含円 Γ の中心 Q^* であり, $d(P_1, Q^0) (= d(P_2, Q^0) = d(P_3, Q^0))$ がその半径 d^* である。

この例から, 与えられた点 P_1, \dots, P_m からの等距離点 Q^0 を考えることが有効であるといえる。

(II) $\triangle P_1 P_2 P_3$ が鈍角三角形の場合 (図 2 参照)

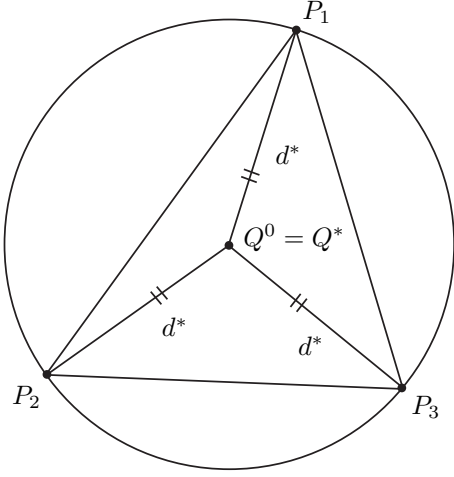


図 1: $\triangle P_1P_2P_3$ が鋭角三角形のときの最小包含円

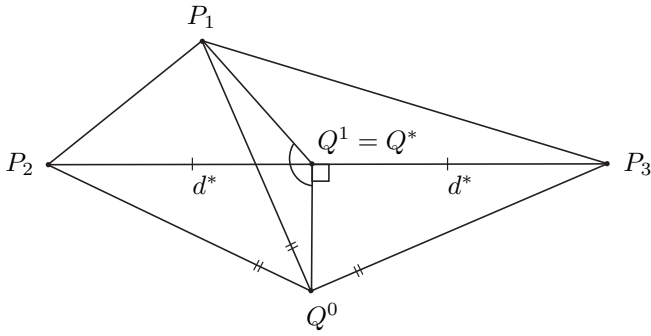


図 2: $\triangle P_1P_2P_3$ が鈍角三角形のときの最小包含円

この場合、例えば $\angle P_1$ が鈍角とすると、最小包含円 Γ の中心 Q^* は辺 P_2P_3 の中点で、半径 d^* は $d(P_2, Q^*) (= d(P_3, Q^*))$ に等しい。 P_1, P_2, P_3 からの等距離点 Q^0 は存在するが、それは $\triangle P_1P_2P_3$ の外部にあり、 Q^0 は最小包含円の中心とはならない。そこで、 P_2, P_3 を通る直線 $L(P_2, P_3)$ 上で Q^0 から最も近い点を Q^1 とし、それを $Q^1 = \pi(Q^0 | L(P_2, P_3))$ と表す。 Q^1 を Q^0 の $L(P_2, P_3)$ への射影という。 Q^1 は辺 P_2P_3 の中点であり、したがって、 $Q^* = Q^1$ となる。

この例から、 $\triangle P_1P_2P_3$ の外部に等距離点 Q^0 がある場合、 Q^0 からある集合への射影を考えることが有効であることがわかる。

このように、2 点を通る直線、または複数個の点で張られるアファイン部分空間への射影によって最小包含円の中心 Q^* あるいは通信路容量を達成する出力分布 Q^* を求める方法を本論文では射影アルゴリズムという。

2.2 重心座標

\mathbb{R}^n の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 (2.4 節参照) m 個の点 P_1, \dots, P_m が与えられているとする。 P_1, \dots, P_m で張られるアファイン部分空間、すなわち、 P_1, \dots, P_m を含む最小のアファイン部分空間を $L_0 = L(P_1, \dots, P_m)$ とする。 \mathbb{R}^n の原点を O とする。点 $Q \in L_0$ の P_1, \dots, P_m に関

する重心座標とは

$$\begin{cases} \overrightarrow{OQ} = \lambda_1 \overrightarrow{OP_1} + \dots + \lambda_m \overrightarrow{OP_m} \\ \sum_{i=1}^m \lambda_i = 1 \end{cases} \quad (10)$$

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 (10) を $Q = \lambda_1 P_1 + \dots + \lambda_m P_m$ と表す。 P_1, \dots, P_m からの等距離点 Q^0 の重心座標を 2.4 節で計算する。

最小包含円問題における重心座標 λ は通信路容量問題における入力分布 λ に相当する。

2.3 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 (9) の min max 問題を解くことによって得られる。(9) 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 (9) 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 (9) のような二重最適化問題は使いにくい。そこで、(9) と等価な単純最適化問題としての凸最適化問題 (convex programming) が以下のように与えられる。

P_1, \dots, P_m の座標を

$$P_i = (P_{i1}, \dots, P_{in}) \in \mathbb{R}^n, i = 1, \dots, m \quad (12)$$

とし、 P_i を並べてできる行列 Φ を

$$\Phi = \begin{pmatrix} P_1 \\ \vdots \\ P_m \end{pmatrix} = \begin{pmatrix} P_{11} & \dots & P_{1n} \\ \vdots & & \vdots \\ P_{m1} & \dots & P_{mn} \end{pmatrix} \in \mathbb{R}^{m \times n} \quad (13)$$

とする。 P_i のノルムを $\|P_i\| = (\sum_{j=1}^n P_{ij}^2)^{1/2}$ として、ベクトル \mathbf{a} を $\mathbf{a} = (\|P_1\|^2, \dots, \|P_m\|^2)$ と定義する。ここで、 $\lambda \in \mathbb{R}^m$ に対して関数 $f: \mathbb{R}^m \rightarrow \mathbb{R}$ を

$$f(\lambda) = \lambda^t \mathbf{a} - \lambda \Phi^t \Phi^t \lambda \quad (14)$$

と定義する。ここで、 t はベクトルや行列の転置を表す。 f は微分可能で上に凸な関数である。このとき、凸最適化問題

$$\max_{\lambda \in \Delta^m} f(\lambda) \quad (15)$$

は最小包含円問題 (9) と等価である [9]。(15) を達成する $\lambda = \lambda^*$ に対して $Q^* \equiv \lambda^* \Phi$ が最小包含円 $\Gamma(P_1, \dots, P_m)$ の中心であり、 $d^* \equiv \sqrt{f(\lambda^*)}$ がその半径である [9]。

(15) に対して次の Kuhn-Tucker 条件が成り立つ [6], [9]。
定理 (最小包含円問題に関する Kuhn-Tucker 条件) $\lambda^* = (\lambda_1^*, \dots, \lambda_m^*) \in \Delta^m$ が (15) を達成するための必要十分条件は、ある d_0 が存在して、

$$d(P_i, \lambda^* \Phi) \begin{cases} = d_0, & \text{for } i \text{ with } \lambda_i^* > 0 \\ \leq d_0, & \text{for } i \text{ with } \lambda_i^* = 0 \end{cases} \quad (16)$$

を満たすことである。このとき、 $Q^* = \lambda^* \Phi$ が最小包含円 $\Gamma(P_1, \dots, P_m)$ の中心であり、 $d^* = d_0$ がその半径に等しい。

2.4 P_1, \dots, P_m からの等距離点とその重心座標

P_1, \dots, P_m に対して、 Q^0 が P_1, \dots, P_m からの等距離点とは、 $d(P_1, Q^0) = \dots = d(P_m, Q^0)$ であり、かつ Q^0 が P_1, \dots, P_m で張られるアファイン部分空間に含まれることである。

\mathbb{R}^n の原点を O とする。以下、記号の簡単化のために、 $\overrightarrow{OP_i}$ を単に P_i と書く。そして、実数 λ_1, λ_2 に対して、 $\lambda_1 \overrightarrow{OP_1} + \lambda_2 \overrightarrow{OP_2}$ を $\lambda_1 P_1 + \lambda_2 P_2$ と書く。したがって、例えば、 $P_2 - P_1$ は $\overrightarrow{P_1 P_2}$ を意味する。場合に依りて、 P_i を \mathbb{R}^n の点と考えたり、座標 (P_{i1}, \dots, P_{in}) と考えたり、ベクトル $\overrightarrow{OP_i}$ またはその成分と考えたりする。

m 個の点 $P_1, \dots, P_m \in \mathbb{R}^n$ が一般の位置にある、とは、ベクトルの集合

$$\{P_2 - P_1, \dots, P_m - P_1\} \text{ が一次独立} \quad (17)$$

となることである。点 P_1, \dots, P_m で張られるアファイン部分空間を $L_0 = L(P_1, \dots, P_m)$ とおくと、(17) の条件のもとで $\dim L_0 = m - 1$ である。最小包含円問題は \mathbb{R}^n の直交変換によって不変であるから、適当な直交変換を用いて、 L_0 は初めから \mathbb{R}^{m-1} に等しいとして一般性を失わない。そこで、 $P_i = (P_{i1}, \dots, P_{i,m-1}) \in \mathbb{R}^{m-1}$, $i = 1, \dots, m$ とする。

$(m-1) \times (m-1)$ 行列 Ψ を

$$\Psi = \begin{pmatrix} P_2 - P_1 \\ \vdots \\ P_m - P_1 \end{pmatrix} \quad (18)$$

と定義すると、 P_1, \dots, P_m が一般の位置にあることは、(17) より

$$\det \Psi \neq 0 \quad (19)$$

と同値である。

$P_1, \dots, P_m \in \mathbb{R}^{m-1}$ が一般の位置にあるとき、 P_1, \dots, P_m からの等距離点 $Q^0 \in \mathbb{R}^{m-1}$ が存在して一意であることを示し、さらに Q^0 の P_1, \dots, P_m に関する重心座標 λ^0 を計算する。

2.4.1 等距離点の存在と一意性

等距離点 Q^0 は

$$d(P_i, Q^0) = d(P_1, Q^0), i = 2, \dots, m \quad (20)$$

を満たす点である。(20) の両辺を 2 乗してベクトルで表すと、

$$2(P_i - P_1)^t Q^0 = \|P_i\|^2 - \|P_1\|^2, i = 2, \dots, m \quad (21)$$

となる。ここで、 ${}^t Q^0$ はベクトル $Q^0 (= \overrightarrow{OQ^0})$ の転置を表す。 $\mathbf{b} = (\|P_2\|^2 - \|P_1\|^2, \dots, \|P_m\|^2 - \|P_1\|^2) \in \mathbb{R}^{m-1}$ と定義すると、(21) は

$$2\Psi {}^t Q^0 = {}^t \mathbf{b} \quad (22)$$

と書ける。(19) より

$$Q^0 = \frac{1}{2} \mathbf{b} {}^t \Psi^{-1} \quad (23)$$

となり、等距離点 $Q^0 \in \mathbb{R}^{m-1}$ が唯一つ存在することがわかる。

2.4.2 等距離点の重心座標

次に、等距離点 $Q^0 \in \mathbb{R}^{m-1}$ の $P_1, \dots, P_m \in \mathbb{R}^{m-1}$ に関する重心座標を $\lambda^0 = (\lambda_i^0)_{i=1, \dots, m} \in \mathbb{R}^m$ とおき、これを計算する。

重心座標は \mathbb{R}^{m-1} の平行移動によって不変なので、 λ^0 を計算するために $Q^0 = O$ と仮定してよい。このとき、 λ^0 は

$$\begin{cases} \sum_{i=1}^m \lambda_i^0 P_i = O \\ \sum_{i=1}^m \lambda_i^0 = 1 \end{cases} \quad (24)$$

$$\sum_{i=1}^m \lambda_i^0 = 1 \quad (25)$$

を満たす。ここで、各 $P_i \in \mathbb{R}^{m-1}$ に対して $\hat{P}_i \in \mathbb{R}^m$ を $\hat{P}_i = (1, P_i)$ と定義する。そして、 $m \times m$ 行列 $\hat{\Phi}$ を

$$\hat{\Phi} = \begin{pmatrix} \hat{P}_1 \\ \vdots \\ \hat{P}_m \end{pmatrix} \quad (26)$$

と定義する。 $\hat{\Phi}$ は行ベクトルが一次独立なので正則行列である。実際、 $\sum_{i=1}^m c_i \hat{P}_i = \mathbf{0}$ とすると、 $\sum_{i=2}^m c_i (P_i - P_1) = \mathbf{0}$ となり、(17) より $c_i = 0, i = 1, \dots, m$ となるからである。

また、 m 次元行ベクトル \mathbf{e}_1 を $\mathbf{e}_1 = (1, 0, \dots, 0) \in \mathbb{R}^m$ と定義する。(24), (25) をまとめて

$$\lambda^0 \hat{\Phi} = \mathbf{e}_1 \quad (27)$$

と表すことができるので、 $\hat{\Phi}$ が正則であることから

$$\lambda^0 = \mathbf{e}_1 \hat{\Phi}^{-1} \quad (28)$$

を得る。

さらに、 λ^0 の各成分 $\lambda_i^0, i = 1, \dots, m$ を具体的に計算する。 $(m-1) \times (m-1)$ 行列 $\hat{\Phi}^{(i)}, i = 1, \dots, m$ を

$$\hat{\Phi}^{(i)} = \begin{pmatrix} P_1 \\ \vdots \\ (P_i) \\ \vdots \\ P_m \end{pmatrix} \quad (29)$$

と定義する。ここで、 (P_i) は P_i を除外することを意味する。このとき、Laplace 展開によって $\hat{\Phi}^{-1}$ の第 1 行は

$$\frac{1}{\det \hat{\Phi}} \left(\det \hat{\Phi}^{(1)}, \dots, (-1)^{i-1} \det \hat{\Phi}^{(i)}, \dots, (-1)^{m-1} \det \hat{\Phi}^{(m)} \right) \quad (30)$$

となる。 $\det \hat{\Phi} = \sum_{i=1}^m (-1)^{i-1} \det \hat{\Phi}^{(i)}$ なので, (28) より

$$\lambda_i^0 = \frac{(-1)^{i-1} \det \hat{\Phi}^{(i)}}{\sum_{i'=1}^m (-1)^{i'-1} \det \hat{\Phi}^{(i')}} , i = 1, \dots, m \quad (31)$$

となる。これが等距離点 Q^0 の重心座標である。

2.5 \mathbb{R}^n における P_1, \dots, P_m からの等距離点 Q^0

前節では, P_1, \dots, P_m からの等距離点の存在を示すこととその重心座標を計算するために, \mathbb{R}^n の直交変換によって $L(P_1, \dots, P_m)$ が \mathbb{R}^{m-1} に等しいと仮定した。しかし, 与えられた P_1, \dots, P_m に対してそのような直交変換を求めることは通常困難なので, 本節では \mathbb{R}^n において P_1, \dots, P_m が一般の位置にあるとして等距離点 $Q^0 \in L(P_1, \dots, P_m)$ とその重心座標 λ^0 を求める。 m 個の点が \mathbb{R}^n において一般の位置にあるので $m \leq n+1$ である。

P_i の座標が (12) で定義され, $m \times n$ 行列 Φ が (13) で定義される。 Q^0 の座標を

$$Q^0 = (Q_1^0, \dots, Q_n^0) \in \mathbb{R}^n \quad (32)$$

とおき, さらに以下の定義を行う。

$$\begin{aligned} \hat{P}_i &= (1, P_i) \\ &= (1, P_{i1}, \dots, P_{in}) \in \mathbb{R}^{n+1}, i = 1, \dots, m \end{aligned} \quad (33)$$

$$\hat{Q}^0 = (1, Q^0) = (1, Q_1^0, \dots, Q_n^0) \in \mathbb{R}^{n+1} \quad (34)$$

$$\begin{aligned} \hat{\Phi} &= \begin{pmatrix} \hat{P}_1 \\ \vdots \\ \hat{P}_m \end{pmatrix} \\ &= \begin{pmatrix} 1 & P_{11} & \cdots & P_{1n} \\ \vdots & \vdots & & \vdots \\ 1 & P_{m1} & \cdots & P_{mn} \end{pmatrix} \in \mathbb{R}^{m \times (n+1)} \end{aligned} \quad (35)$$

$$J = \begin{pmatrix} -1 & 1 & 0 & \cdots & 0 \\ -1 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -1 & 0 & 0 & \cdots & 1 \end{pmatrix} \in \mathbb{R}^{(m-1) \times m} \quad (36)$$

$$\hat{\mathbf{a}} = (\|\hat{P}_1\|^2, \dots, \|\hat{P}_m\|^2) \in \mathbb{R}^m \quad (37)$$

$$\mathbf{1} = (1, \dots, 1) \in \mathbb{R}^m \quad (38)$$

λ^0 は Q^0 の P_1, \dots, P_m に関する重心座標なので,

$$\begin{cases} Q^0 = \lambda^0 \Phi \\ \lambda^0 \mathbf{1} = 1 \end{cases} \quad (39)$$

$$\lambda^0 \mathbf{1} = 1 \quad (40)$$

が成り立つ。(39), (40) をまとめて,

$$\hat{Q}^0 = \lambda^0 \hat{\Phi} \quad (41)$$

と書ける。

m 個の点 P_1, \dots, P_m からの等距離点 Q^0 は

$$d(P_i, Q^0) = d(P_1, Q^0), i = 2, \dots, m \quad (42)$$

を満たす。(42) の両辺を 2 乗してベクトルで表すと,

$$2(P_i - P_1)^t Q^0 = \|P_i\|^2 - \|P_1\|^2, i = 2, \dots, m \quad (43)$$

となり (33), (34), (43) より

$$2(\hat{P}_i - \hat{P}_1)^t \hat{Q}^0 = \|\hat{P}_i\|^2 - \|\hat{P}_1\|^2, i = 2, \dots, m \quad (44)$$

が成り立つ。(44) は

$$2J\hat{\Phi}^t \hat{Q}^0 = J^t \hat{\mathbf{a}} \quad (45)$$

と書けるので, (41) より

$$2J\hat{\Phi}^t \hat{\Phi}^t \lambda^0 = J^t \hat{\mathbf{a}} \quad (46)$$

となる。ここで, $M \equiv \hat{\Phi}^t \hat{\Phi} \in \mathbb{R}^{m \times m}$ とおくと, (46) は

$$J(2M^t \lambda^0 - \hat{\mathbf{a}}) = \mathbf{0} \quad (47)$$

と書ける。前節と同様に, $\hat{\Phi}$ の行ベクトルが一次独立であり, したがって $\text{rank } \hat{\Phi} = m$ となることがわかるので, $\text{rank } M = \text{rank } \hat{\Phi} = m$ となる。すなわち, M は正則行列である。そして, $M = \hat{\Phi}^t \hat{\Phi}$ より M は正定値対称行列である。また, $\text{Ker } J = \{\tau \mathbf{1} | \tau \in \mathbb{R}\}$ なので, (47) より, ある $\tau \in \mathbb{R}$ が存在して,

$$\lambda^0 = \frac{1}{2}(\hat{\mathbf{a}} + \tau \mathbf{1})M^{-1} \quad (48)$$

となる。(40) より

$$\mathbf{1} = \lambda^0 \mathbf{1} = \frac{1}{2}(\hat{\mathbf{a}} + \tau \mathbf{1})M^{-1} \mathbf{1} \quad (49)$$

となり, (49) より

$$\tau = \frac{2 - \hat{\mathbf{a}} M^{-1} \mathbf{1}}{\mathbf{1} M^{-1} \mathbf{1}} \quad (50)$$

となる。(50) を (48) に代入して, 結局, 等距離点 Q^0 の重心座標 λ^0 として

$$\lambda^0 = \frac{1}{2} \left(\hat{\mathbf{a}} + \frac{2 - \hat{\mathbf{a}} M^{-1} \mathbf{1}}{\mathbf{1} M^{-1} \mathbf{1}} \mathbf{1} \right) M^{-1} \quad (51)$$

を得る。

2.6 \mathbb{R}^n における内積, Pythagoras の定理, 射影

最小包含円の中心を射影アルゴリズムで求めるために重要な, \mathbb{R}^n における内積, Pythagoras の定理, 射影について述べる。

2.6.1 内積

\mathbb{R}^n の3点 $Q^1 = (Q_j^1)_{j=1,\dots,n}$, $Q^2 = (Q_j^2)_{j=1,\dots,n}$, $Q^3 = (Q_j^3)_{j=1,\dots,n}$ を考える。 Q^1, Q^2, Q^3 の Q^2 における内積 $(Q^1 - Q^2, Q^3 - Q^2)$ を

$$(Q^1 - Q^2, Q^3 - Q^2) = \sum_{j=1}^n (Q_j^1 - Q_j^2)(Q_j^3 - Q_j^2) \quad (52)$$

と定義する。これは、ベクトル $\overrightarrow{Q^2Q^1}$ と $\overrightarrow{Q^2Q^3}$ の内積である。

次の補題が成り立つ。

補題 1 \mathbb{R}^n の点 P_i ($i = 1, \dots, m$), Q, R に対して、内積 $\sigma_i = (P_i - Q, R - Q)$, $i = 1, \dots, m$ を考える。 $\sum_{i=1}^m \lambda_i = 1$ のとき

$$\sum_{i=1}^m \lambda_i \sigma_i = \left(\sum_{i=1}^m \lambda_i P_i - Q, R - Q \right)$$

証明 簡単な計算による。 \square

補題 2 任意の $P, Q \in \mathbb{R}^n$, $P \neq Q$ に対して

$$(P - Q, P - Q) > 0$$

証明 簡単な計算による。 \square

補題 3 任意の $P, Q, R \in \mathbb{R}^n$ に対して

$$2(P - Q, R - Q) = d^2(P, Q) + d^2(Q, R) - d^2(P, R)$$

証明 簡単な計算による。 \square

2.6.2 Pythagoras の定理

\mathbb{R}^n の3点 P, Q, R に対して、次のように Pythagoras の定理とその不等式版が成り立つ。

定理 1 (Pythagoras) $P, Q, R \in \mathbb{R}^n$ に対して

$$\begin{aligned} (P - Q, R - Q) &\geq 0 \\ \iff d^2(P, Q) + d^2(Q, R) &\geq d^2(P, R) \end{aligned} \quad (53)$$

証明 補題 3 より \square

2.6.3 射影

\mathbb{R}^n の点 Q^0 および \mathbb{R}^n の部分集合 L を考える。最小値 $\min_{Q \in L} d(Q, Q^0)$ を達成する点 $Q = Q^1$ を Q^0 の L への射影といい、 $Q^1 = \pi(Q^0|L)$ と表す。以下では、特に、部分集合 L として \mathbb{R}^n のアフィン部分空間だけを考える。

補題 4 L を \mathbb{R}^n のアフィン部分空間とする。任意の $Q^0 \in \mathbb{R}^n$ に対して $Q^1 = \pi(Q^0|L)$ は一意に存在する。 $Q^1 = \pi(Q^0|L)$ であることと、任意の $P \in L$ に対して $(P - Q^1, Q^0 - Q^1) = 0$ であることは同値である。

証明 [10] 参照。 \square

2.6.4 等距離点からの射影

\mathbb{R}^n において P_1, \dots, P_m が一般の位置にあるとする。 $L_0 = L(P_1, \dots, P_m)$ を P_1, \dots, P_m で張られるアフィン部分空間とする。 $Q^0 \in L_0$ を P_1, \dots, P_m からの等距離点とする。(23) からそのような Q^0 は唯一つ存在する。さらに $L_k = L(P_{k+1}, \dots, P_m)$, $k = 0, \dots, m-2$ を P_{k+1}, \dots, P_m で張られるアフィン部分空間とする。 $L_0 \supset L_1 \supset \dots \supset L_k \supset \dots$ は次元が1ずつ小さくなるアフィン部分空間の減少列である。

Q^0 の L_1 への射影を $Q^1 = \pi(Q^0|L_1)$ とする。以下、同様に、 $Q^k = \pi(Q^{k-1}|L_k)$, $k = 1, \dots, m-2$ とする。次の補題が成り立つ。

補題 5 $Q^k = \pi(Q^0|L_k)$, $k = 0, 1, \dots, m-2$ 。

証明 $k = 0, 1$ のときは定義より明らか。 $k = 2$ のときは証明する。補題 4 および定理 1 より、任意の $Q \in L_2 \subset L_1$ に対して $d^2(Q, Q^0) = d^2(Q, Q^1) + d^2(Q^1, Q^0)$ が成り立つ。したがって、 $Q \in L_2$ に関して $d^2(Q, Q^0)$ を最小化することと $d^2(Q, Q^1)$ を最小化することは同値である。補題 4 より射影 $\pi(Q^0|L_2)$ および $\pi(Q^1|L_2)$ は一意なので、 $\pi(Q^0|L_2) = \pi(Q^1|L_2)$ を得る。 $k \geq 3$ に対しては数学的帰納法によって証明できる。 \square

補題 6 $(Q^i - Q^k, Q^0 - Q^k) = (Q^i - Q^k, Q^i - Q^k)$, $i = 0, 1, \dots, k, k = 0, 1, \dots, m-2$ 。

証明 計算よって $(Q^i - Q^k, Q^0 - Q^k) = -(Q^k - Q^i, Q^0 - Q^i) + (Q^i - Q^k, Q^i - Q^k)$ となる。 $Q^k \in L_k \subset L_i$ なので、補題 4 より $(Q^k - Q^i, Q^0 - Q^i) = 0$ を得る。 \square

2.7 射影アルゴリズムによる Q^* の探索

2.1 節で考察したように、等距離点 Q^0 を考え、 Q^0 であるアフィン部分空間への射影する方法が有効である。この方法を射影アルゴリズムという。

与えられた点 P_1, \dots, P_m に対する最小包含円 $\Gamma = \Gamma(P_1, \dots, P_m)$ の中心を Q^* とし、半径を d^* とする。本節では、 P_1, \dots, P_m が一般の位置にあると仮定して、あるいくつかの状況で Q^* , d^* が射影アルゴリズムによって得られることを示す。

2.8 状況 1 [毎回の射影で負の重心座標成分がちょうど1個ある場合]

まず、次の例題を考える。

例題 1 \mathbb{R}^3 の4点 $P_1 = (-10, 1, -3)$, $P_2 = (-9, -2, 8)$, $P_3 = (-8, 10, -5)$, $P_4 = (4, -8, 8)$ を考える。 P_1, \dots, P_4 は一般の位置にある (2.4 節参照)。 P_1, \dots, P_4 からの等距離点は $Q^0 = (1.93, 4.59, 2.85)$ となり、その重心座標 λ^0 は (51) より

$$\lambda^0 = (-0.84, 0.04, 1.11, 0.69) \quad (54)$$

となる。 $\lambda_1^0 = -0.84 < 0$ に対する P_1 を除外して $L_1 = L(P_2, P_3, P_4)$ への Q^0 の射影 $Q^1 = \pi(Q^0|L_1)$ を計算する。 $Q^1 = (0.48, 1.45, -0.16)$ となりその重心座標 λ^1 は

$$\lambda^1 = (0, -0.31, 0.63, 0.68) \quad (55)$$

となる。 $\lambda_2^1 = -0.31 < 0$ に対する P_2 を除外して $L_2 = L(P_3, P_4)$ への Q^1 の射影 $Q^2 = \pi(Q^1|L_2)$ を計算する。

$Q^2 = (-2, 1, 1.5)$ となり, その重心座標 λ^2 は

$$\lambda^2 = (0, 0, 0.5, 0.5) \quad (56)$$

となる。このとき, $d(P_1, Q^2) = 9.18, d(P_2, Q^2) = 10.01, d(P_3, Q^2) = d(P_4, Q^2) = 12.62$ となり, Kuhn-Tucker 条件 (16) によって Q^2 が最小包含円の中心であること, すなわち $Q^* = Q^2$ がわかる。そして, $d^* = d(P_3, Q^2)$ である。

本節では, (54), (55) のように毎回の射影で負の重心座標成分がちょうど 1 個あり, (56) のように最後にすべての重心座標の成分が非負になる場合を「状況 1」として, 最小包含円を求める射影アルゴリズムを考える。

状況 1 の仮定 $P_1, \dots, P_m \in \mathbb{R}^n$ は一般の位置にあるとし, $L_k = L(P_{k+1}, \dots, P_m), k = 0, 1, \dots, m-2$ を P_{k+1}, \dots, P_m で張られるアファイン部分空間とする。 $Q^0 \in L_0$ を P_1, \dots, P_m からの等距離点として, $Q^k = \pi(Q^{k-1} | L_k), k = 1, \dots, m-2$ と定義する。 Q^k の P_1, \dots, P_m に関する重心座標を $\lambda^k = (\lambda_1^k, \dots, \lambda_m^k)$ とする。そこで, $K = 0, 1, \dots, m-2$ とし, $k = 0, 1, \dots, K-1$ のとき λ^k の成分のうち負のものがちょうど 1 個あり, $k = K$ のとき λ^K の成分はすべて非負であるとする。すなわち,

$$\begin{aligned} \lambda^0 &= (\lambda_1^0, \lambda_2^0, \dots, \lambda_m^0) = (-, +, \dots, +) \\ \lambda^1 &= (\lambda_1^1, \lambda_2^1, \lambda_3^1, \dots, \lambda_m^1) = (0, -, +, \dots, +) \\ \lambda^2 &= (\lambda_1^2, \lambda_2^2, \lambda_3^2, \lambda_4^2, \dots, \lambda_m^2) = (0, 0, -, +, \dots, +) \\ &\vdots \\ \lambda^{K-1} &= (\lambda_1^{K-1}, \dots, \lambda_{K-1}^{K-1}, \lambda_K^{K-1}, \dots, \lambda_m^{K-1}) \\ &= (0, \dots, 0, -, +, \dots, +) \\ \lambda^K &= (\lambda_1^K, \dots, \lambda_K^K, \lambda_{K+1}^K, \dots, \lambda_m^K) \\ &= (0, \dots, 0, +, \dots, +) \end{aligned}$$

と仮定するのである。ただし, $K = 0$ のときは $\lambda^0 = (\lambda_1^0, \lambda_2^0, \dots, \lambda_m^0) = (+, +, \dots, +)$ とみなす。これらをまとめて書くと,

$k = 0, 1, \dots, K-1$ のとき,

$$\lambda_i^k \begin{cases} = 0, & i = 1, \dots, k, \\ < 0, & i = k+1, \\ > 0, & i = k+2, \dots, m \end{cases} \quad (57a)$$

$$< 0, \quad i = k+1, \quad (57b)$$

$$> 0, \quad i = k+2, \dots, m \quad (57c)$$

となり, また $k = K$ のとき,

$$\lambda_i^K \begin{cases} = 0, & i = 1, \dots, K, \\ > 0, & i = K+1, \dots, m \end{cases} \quad (58a)$$

$$> 0, \quad i = K+1, \dots, m \quad (58b)$$

となる。

(状況 1 の仮定終わり)

さて, ここで, 内積 $\sigma_i \equiv (P_i - Q^K, Q^0 - Q^K), i = 1, \dots, m$ を考える。次の補題が成り立つ。

$$\text{補題 7 } \sigma_i \begin{cases} < 0, & i = 1, \dots, K, \\ = 0, & i = K+1, \dots, m. \end{cases}$$

証明 補題 5 より $Q^K = \pi(Q^0 | L_K), L_K = L(P_{K+1}, \dots, P_m)$ なので, $i = K+1, \dots, m$ に対して $P_i \in L_K$ だから補題 4 より $P_i - Q^K \perp Q^0 - Q^K$ となり, したがって,

$$\sigma_i = 0, \quad i = K+1, \dots, m \quad (59)$$

となる。次に, $\sigma_i < 0, i = 1, \dots, K$ を数学的帰納法によって, $i = K, K-1, \dots, 1$ の順に証明する。

(I) $\sigma_K < 0$ を証明する:

(57a), (59) より

$$\begin{aligned} \lambda_K^{K-1} \sigma_K &= \sum_{i=1}^m \lambda_i^{K-1} \sigma_i \\ &= \left(\sum_{i=1}^m \lambda_i^{K-1} P_i - Q^K, Q^0 - Q^K \right) \quad (\text{補題 1 より}) \\ &= (Q^{K-1} - Q^K, Q^0 - Q^K) \\ &= (Q^{K-1} - Q^K, Q^{K-1} - Q^K) \quad (\text{補題 6 より}) \\ &> 0 \quad (\text{補題 2 より}) \end{aligned}$$

となる。(57b) より $\lambda_K^{K-1} < 0$ なので $\sigma_K < 0$ を得る。
(II) $\sigma_K < 0, \sigma_{K-1} < 0, \dots, \sigma_{k+1} < 0$ を仮定して $\sigma_k < 0$ を証明する:

(57a), (59) より

$$\begin{aligned} \sum_{i=k}^K \lambda_i^{k-1} \sigma_i &= \sum_{i=1}^m \lambda_i^{k-1} \sigma_i \\ &= \left(\sum_{i=1}^m \lambda_i^{k-1} P_i - Q^K, Q^0 - Q^K \right) \\ &= (Q^{k-1} - Q^K, Q^0 - Q^K) \\ &= (Q^{k-1} - Q^K, Q^{k-1} - Q^K) \\ &> 0 \end{aligned}$$

となる。すなわち, $\lambda_k^{k-1} \sigma_k + \sum_{i=k+1}^K \lambda_i^{k-1} \sigma_i > 0$ である。帰納法の仮定より $\sigma_i < 0, i = k+1, \dots, K$ であり, また, (57b) より $\lambda_k^{k-1} < 0$ であり, (57c) より $\lambda_i^{k-1} > 0, i = k+1, \dots, K$ であるから $\sigma_k < 0$ を得る。
(III) 上記 (I), (II) より, $\sigma_i < 0, i = 1, \dots, K$ を得る。□
補題 8 $d(P_i, Q^K) < d(P_{K+1}, Q^K) = \dots = d(P_m, Q^K), i = 1, \dots, K$.

証明 補題 7 より $\sigma_i = 0, i = K+1, \dots, m$ なので, 定理 1 より $d(P_{K+1}, Q^K) = \dots = d(P_m, Q^K)$ である。また, $\sigma_i < 0, i = 1, \dots, K$ なので, 定理 1 より $i = 1, \dots, K$ に対して

$$\begin{aligned} d^2(P_i, Q^K) &< d^2(P_i, Q^0) - d^2(Q^K, Q^0) \\ &= d^2(P_{K+1}, Q^0) - d^2(Q^K, Q^0) \\ &= d^2(P_{K+1}, Q^K) \end{aligned}$$

を得る。□

定理 2 最小包含円の中心は $Q^* = Q^K$ であり半径は $d^* = d(P_{K+1}, Q^K)$ である。

証明 (58a), (58b) と補題 8 および Kuhn-Tucker 条件 (16) より。□

2.9 状況 2 [$m = 2, 3, 4$ の場合]

さて、次に \mathbb{R}^n において一般の位置にある m 個の点 P_1, \dots, P_m に対する最小包含円 $\Gamma = \Gamma(P_1, \dots, P_m)$ の中心 Q^* と半径 d^* を射影アルゴリズムによって計算する。目標はすべての m に対して計算するアルゴリズムを見出すことであるが、現時点では以下のように $m = 2, 3, 4$ の場合だけ解決できている。

2.9.1 $m = 2$ の場合

\mathbb{R}^n における異なる 2 点 P_1, P_2 に対する最小包含円 $\Gamma = \Gamma(P_1, P_2)$ について考える。線分 P_1P_2 の中点を Q^0 とすると、容易に次の定理が得られる。

定理 3 最小包含円 Γ の中心は $Q^* = Q^0$ であり半径は $d^* = d(P_1, Q^0)$ である。

2.9.2 $m = 3$ の場合

\mathbb{R}^n において一般の位置にある 3 点 P_1, P_2, P_3 に対する最小包含円 $\Gamma = \Gamma(P_1, P_2, P_3)$ について考える。 $L_0 = L(P_1, P_2, P_3)$ を P_1, P_2, P_3 で張られるアファイン部分空間とする。 $Q^0 \in L_0$ を P_1, P_2, P_3 からの等距離点とする。 Q^0 の P_1, P_2, P_3 に関する重心座標を $\lambda^0 = (\lambda_1^0, \lambda_2^0, \lambda_3^0)$ とする。

λ^0 の各成分の符号に基づいて、一般性を失うことなく、以下の場合分けをすることができる。

Case 1: $\lambda_1^0 \geq 0, \lambda_2^0 \geq 0, \lambda_3^0 \geq 0$

Case 2: $\lambda_1^0 < 0, \lambda_2^0 \geq 0, \lambda_3^0 \geq 0$

Q^0 は P_1, P_2, P_3 からの等距離点なので正の重心座標成分は必ず 2 個以上ある。したがって、Case 1 と Case 2 ですべての場合が尽くされている。それぞれの Case について、最小包含円 $\Gamma = \Gamma(P_1, P_2, P_3)$ の中心 Q^* と半径 d^* を決定する。

Case 1

この場合は容易である。Kuhn-Tucker 条件 (16) によって次の定理が得られる。

定理 4 最小包含円 Γ の中心は $Q^* = Q^0$ であり半径は $d^* = d(P_1, Q^0)$ である。

Case 2

2 点 P_2, P_3 を通る直線を $L_1 = L(P_2, P_3)$ とする。仮定 $\lambda_1^0 < 0$ より、等距離点 Q^0 は直線 L_1 に関して P_1 と反対側にある。 Q^0 の直線 L_1 への射影を $Q^1 = \pi(Q^0|L_1)$ とし、内積 $\sigma_i = (P_i - Q^1, Q^0 - Q^1)$, $i = 1, 2, 3$ を考える。次の補題が成り立つ。

補題 9 $\sigma_i \begin{cases} < 0, & i = 1, \\ = 0, & i = 2, 3. \end{cases}$

証明 $Q^1 = \pi(Q^0|L(P_2, P_3))$ より $i = 2, 3$ に対して $P_i - Q^1 \perp Q^0 - Q^1$ である。したがって、

$$\sigma_i = 0, \quad i = 2, 3 \quad (60)$$

となり、

$$\begin{aligned} \lambda_1^0 \sigma_1 &= \sum_{i=1}^3 \lambda_i^0 \sigma_i \\ &= \left(\sum_{i=1}^3 \lambda_i^0 P_i - Q^1, Q^0 - Q^1 \right) \\ &= (Q^0 - Q^1, Q^0 - Q^1) \\ &> 0 \end{aligned} \quad (61)$$

となる。(61) および Case 2 の仮定 $\lambda_1^0 < 0$ より $\sigma_1 < 0$ を得る。□

補題 10 $d(P_1, Q^1) < d(P_2, Q^1) = d(P_3, Q^1)$.

証明 (60) および定理 1 より $d(P_2, Q^1) = d(P_3, Q^1)$ である。また、 $\sigma_1 < 0$ なので、定理 1 より

$$\begin{aligned} d^2(P_1, Q^1) &< d^2(P_1, Q^0) - d^2(Q^1, Q^0) \\ &= d^2(P_2, Q^0) - d^2(Q^1, Q^0) \\ &= d^2(P_2, Q^1) \end{aligned}$$

を得る。□

定理 5 最小包含円 Γ の中心は $Q^* = Q^1$ であり半径は $d^* = d(P_2, Q^1)$ である。

証明 補題 10 および Kuhn-Tucker 条件 (16) より。□

2.9.3 $m = 4$ の場合

\mathbb{R}^n において一般の位置にある 4 点 P_1, P_2, P_3, P_4 に対する最小包含円 $\Gamma = \Gamma(P_1, P_2, P_3, P_4)$ の中心 Q^* と半径 d^* を求める。 $L_0 = L(P_1, \dots, P_4)$ を P_1, \dots, P_4 で張られるアファイン部分空間とし、 $Q^0 \in L_0$ を P_1, \dots, P_4 からの等距離点とする。 Q^0 の P_1, \dots, P_4 に関する重心座標を $\lambda^0 = (\lambda_i^0)_{i=1, \dots, 4}$ とする。一般性を失うことなく、以下の場合分けをすることができる。

Case 1: $\lambda_1^0 \geq 0, \lambda_2^0 \geq 0, \lambda_3^0 \geq 0, \lambda_4^0 \geq 0$

Case 2: $\lambda_1^0 < 0, \lambda_2^0 \geq 0, \lambda_3^0 \geq 0, \lambda_4^0 \geq 0$

Case 3: $\lambda_1^0 < 0, \lambda_2^0 < 0, \lambda_3^0 \geq 0, \lambda_4^0 \geq 0$

Case 1

Kuhn-Tucker 条件 (16) によって、次の定理を得る。

定理 6 最小包含円 Γ の中心および半径は $Q^* = Q^0$, $d^* = d(P_1, Q^0)$ である。

Case 2

$\lambda_1^0 < 0$ なので、Kuhn-Tucker 条件 (16) により、 Q^0 は最小包含円の中心ではない。そこで、 Q^0 のアファイン部分空間 $L_1 = L(P_2, P_3, P_4)$ への射影を $Q^1 = \pi(Q^0|L_1)$ とする。 Q^1 の P_1, \dots, P_4 に関する重心座標を $\lambda^1 = (\lambda_i^1)_{i=1, \dots, 4}$ とおく。 $P_1 \notin L_1$ より $\lambda_1^1 = 0$ である。 λ^1 の各成分の符号に基づいて、一般性を失うことなく、さらに以下の場合分けができる。

Case 2-1: $\lambda_1^1 = 0, \lambda_2^1 \geq 0, \lambda_3^1 \geq 0, \lambda_4^1 \geq 0$

Case 2-2: $\lambda_1^1 = 0, \lambda_2^1 < 0, \lambda_3^1 \geq 0, \lambda_4^1 \geq 0$

Case 2-1

内積 $\sigma_i = (P_i - Q^1, Q^0 - Q^1)$, $i = 1, \dots, 4$ を考える。
次の補題を得る。

$$\text{補題 11 } \sigma_i \begin{cases} < 0, & i = 1, \\ = 0, & i = 2, 3, 4. \end{cases}$$

証明 補題 9 の証明とほぼ同様なので、簡潔に記述する。
 $P_i - Q^1 \perp Q^0 - Q^1$, $i = 2, 3, 4$ なので、 $\sigma_i = 0$, $i = 2, 3, 4$ である。したがって、 $\lambda_1^0 \sigma_1 = \sum_{i=1}^4 \lambda_i^0 \sigma_i = (Q^0 - Q^1, Q^0 - Q^1) > 0$ となり、仮定 $\lambda_1^0 < 0$ より $\sigma_1 < 0$ を得る。□

補題 12 $d(P_1, Q^1) < d(P_2, Q^1) = d(P_3, Q^1) = d(P_4, Q^1)$.

証明 補題 10 の証明と同様。□

定理 7 最小包含円 Γ の中心および半径は $Q^* = Q^1$, $d^* = d(P_2, Q^1)$ である。

証明 補題 12 および Kuhn-Tucker 条件 (16) より。□

Case 2-2

この場合は、2.8 節「状況 1」と同じで、毎回の射影で負の成分がちょうど 1 個ある場合だから、すでに証明されている。すなわち、次の定理が成り立つ。

定理 8 $Q^2 = \pi(Q^0 | L(P_3, P_4))$ とおく。最小包含円 Γ の中心は $Q^* = Q^2$ であり半径は $d^* = d(P_3, Q^2)$ である。

これで Case 2 は完了した。次に Case 3 について考える。

Case 3

等距離点 $Q^0 \in L_0$ に対して、 $Q^{1(1)} = \pi(Q^0 | L(P_2, P_3, P_4))$, $Q^{1(2)} = \pi(Q^0 | L(P_1, P_3, P_4))$ として、 $Q^{1(1)}$, $Q^{1(2)}$ の P_1, \dots, P_4 に関する重心座標を、それぞれ、 $\lambda^{1(1)} = (\lambda_i^{1(1)})_{i=1, \dots, 4}$, $\lambda^{1(2)} = (\lambda_i^{1(2)})_{i=1, \dots, 4}$ とおく。 $P_1 \notin L(P_2, P_3, P_4)$, $P_2 \notin L(P_1, P_3, P_4)$ より

$$\lambda_1^{1(1)} = 0, \lambda_2^{1(2)} = 0 \quad (62)$$

である。そして、内積

$$\sigma_i^{1(1)} = (P_i - Q^{1(1)}, Q^0 - Q^{1(1)}), \quad i = 1, \dots, 4, \quad (63)$$

$$\sigma_i^{1(2)} = (P_i - Q^{1(2)}, Q^0 - Q^{1(2)}), \quad i = 1, \dots, 4 \quad (64)$$

を考える。次の補題が成り立つ。

補題 13

$$\sigma_i^{1(1)} \begin{cases} < 0, & i = 1, \\ = 0, & i = 2, 3, 4, \end{cases} \quad \sigma_i^{1(2)} \begin{cases} < 0, & i = 2, \\ = 0, & i = 1, 3, 4. \end{cases}$$

証明 $P_i - Q^{1(1)} \perp Q^0 - Q^{1(1)}$, $i = 2, 3, 4$ より $\sigma_i^{1(1)} = 0$, $i = 2, 3, 4$ である。したがって、

$$\begin{aligned} \lambda_1^0 \sigma_1^{1(1)} &= \sum_{i=1}^4 \lambda_i^0 \sigma_i^{1(1)} \\ &= (Q^0 - Q^{1(1)}, Q^0 - Q^{1(1)}) > 0 \end{aligned} \quad (65)$$

であり、仮定 $\lambda_1^0 < 0$ より $\sigma_1^{1(1)} < 0$ を得る。 $\sigma_i^{1(2)}$ についても同様に証明できる。□

補題 14 $\lambda_1^{1(2)} \sigma_1^{1(1)} + \lambda_2^{1(1)} \sigma_2^{1(2)} > 0$.

証明 補題 13 より

$$\begin{aligned} \lambda_1^{1(2)} \sigma_1^{1(1)} &= \sum_{i=1}^4 \lambda_i^{1(2)} \sigma_i^{1(1)} \\ &= (Q^{1(2)} - Q^{1(1)}, Q^0 - Q^{1(1)}) \end{aligned} \quad (66)$$

$$\begin{aligned} \lambda_2^{1(1)} \sigma_2^{1(2)} &= \sum_{i=1}^4 \lambda_i^{1(1)} \sigma_i^{1(2)} \\ &= (Q^{1(1)} - Q^{1(2)}, Q^0 - Q^{1(2)}) \end{aligned} \quad (67)$$

(66)+(67) より $\lambda_1^{1(2)} \sigma_1^{1(1)} + \lambda_2^{1(1)} \sigma_2^{1(2)} = (Q^{1(1)} - Q^{1(2)}, Q^{1(1)} - Q^{1(2)}) > 0$ を得る。□

さて、補題 13, 14 より、 $\lambda_1^{1(2)} < 0$ または $\lambda_2^{1(1)} < 0$ である。ここまで、点 P_1 と P_2 は全く同等に扱ってきたので、一般性を失うことなく、

$$\lambda_1^{1(2)} < 0 \quad (68)$$

と仮定して次のステップに進む。なお、ここで、 $\lambda^{1(2)}$ の各成分の符号を再確認すると

$$\lambda_1^{1(2)} < 0, \lambda_2^{1(2)} = 0, \lambda_3^{1(2)} \geq 0, \lambda_4^{1(2)} \geq 0 \quad (69)$$

である。

$\lambda^{1(1)}$ の各成分の符号に基づいて場合分けを行う。ここまで、点 P_3 と P_4 は全く同等に扱ってきたので、一般性を失うことなく、以下のように場合分けすることができる。

$$\text{Case 3-1: } \lambda_1^{1(1)} = 0, \lambda_2^{1(1)} \geq 0, \lambda_3^{1(1)} \geq 0, \lambda_4^{1(1)} \geq 0$$

$$\text{Case 3-2: } \lambda_1^{1(1)} = 0, \lambda_2^{1(1)} < 0, \lambda_3^{1(1)} \geq 0, \lambda_4^{1(1)} \geq 0$$

$$\text{Case 3-3: } \lambda_1^{1(1)} = 0, \lambda_2^{1(1)} \geq 0, \lambda_3^{1(1)} < 0, \lambda_4^{1(1)} \geq 0$$

Case 3-1

補題 13 より $\sigma_1^{1(1)} < 0$ なので、補題 10 の証明と同様に、 $d(P_1, Q^{1(1)}) < d(P_2, Q^{1(1)}) = \dots = d(P_4, Q^{1(1)})$ を得る。したがって、Kuhn-Tucker 条件 (16) によって、次の定理を得る。

定理 9 最小包含円 Γ の中心は $Q^* = Q^{1(1)}$ であり半径は $d^* = d(P_2, Q^{1(1)})$ である。

Case 3-2

$Q^2 \equiv \pi(Q^{1(1)} | L(P_3, P_4)) = \pi(Q^0 | L(P_3, P_4))$ として、内積 $\sigma_i^2 = (P_i - Q^2, Q^0 - Q^2)$, $i = 1, \dots, 4$ を考える。次の補題が成り立つ。

$$\text{補題 15 } \sigma_i^2 \begin{cases} < 0, & i = 1, 2, \\ = 0, & i = 3, 4. \end{cases}$$

証明 $P_i - Q^2 \perp Q^0 - Q^2$, $i = 3, 4$ より $\sigma_i^2 = 0$, $i = 3, 4$ である。次に $\sigma_1^2 < 0$, $\sigma_2^2 < 0$ を示す。 $Q^2 = \pi(Q^{1(2)} | L(P_3, P_4)) = \pi(Q^0 | L(P_3, P_4))$ に注意する。(62) より

$$\begin{aligned} \lambda_1^{1(2)} \sigma_1^2 &= \sum_{i=1}^4 \lambda_i^{1(2)} \sigma_i^2 \\ &= (Q^{1(2)} - Q^2, Q^0 - Q^2) \\ &= (Q^{1(2)} - Q^2, Q^{1(2)} - Q^2) \\ &> 0 \end{aligned}$$

となり、仮定 $\lambda_1^{(2)} < 0$ より $\sigma_1^2 < 0$ を得る。同様にし、 $\sum_{i=1}^4 \lambda_i^{(1)} \sigma_i^2$ を考えて、 $\sigma_2^2 < 0$ を得る。□

補題 16 $d(P_i, Q^2) < d(P_3, Q^2) = d(P_4, Q^2)$, $i = 1, 2$.

証明 補題 10 の証明と同様。□

定理 10 最小包含円 Γ の中心は $Q^* = Q^2$ であり半径は $d^* = d(P_3, Q^2)$ である。

証明 補題 16 および Kuhn-Tucker 条件 (16) より。□

Case 3-3

$Q^{1(3)} = \pi(Q^0 | L(P_1, P_2, P_4))$ として、内積 $\sigma_i^{1(3)} = (P_i - Q^{1(3)}, Q^0 - Q^{1(3)})$, $i = 1, \dots, 4$ を考える。次の補題が成り立つ。

補題 17 $\sigma_i^{1(3)} \begin{cases} > 0, & i = 3, \\ = 0, & i = 1, 2, 4. \end{cases}$

証明 $P_i - Q^{1(3)} \perp Q^0 - Q^{1(3)}$, $i = 1, 2, 4$ より $\sigma_i^{1(3)} = 0$, $i = 1, 2, 4$ となる。したがって、 $\lambda_3^0 \sigma_3^{1(3)} = \sum_{i=1}^4 \lambda_i^0 \sigma_i^{1(3)} = (Q^0 - Q^{1(3)}, Q^0 - Q^{1(3)}) > 0$ となる。仮定 $\lambda_3^0 > 0$ より $\sigma_3^{1(3)} > 0$ を得る。□

補題 18 $\lambda_1^{1(3)} < 0$.

証明 補題 17 より

$$\begin{aligned} \lambda_3^{1(1)} \sigma_3^{1(3)} &= \sum_{i=1}^4 \lambda_i^{1(1)} \sigma_i^{1(3)} \\ &= (Q^{1(1)} - Q^{1(3)}, Q^0 - Q^{1(3)}) \end{aligned} \quad (70)$$

となる。また、補題 13 より

$$\begin{aligned} \lambda_1^{1(3)} \sigma_1^{1(1)} &= \sum_{i=1}^4 \lambda_i^{1(3)} \sigma_i^{1(1)} \\ &= (Q^{1(3)} - Q^{1(1)}, Q^0 - Q^{1(1)}) \end{aligned} \quad (71)$$

となる。(70)+(71) より

$$\begin{aligned} \lambda_3^{1(1)} \sigma_3^{1(3)} + \lambda_1^{1(3)} \sigma_1^{1(1)} &= (Q^{1(1)} - Q^{1(3)}, Q^{1(1)} - Q^{1(3)}) \\ &> 0 \end{aligned} \quad (72)$$

となる。Case 3-3 の仮定より $\lambda_3^{1(1)} < 0$ 、補題 17 より $\sigma_3^{1(3)} > 0$ 、補題 13 より $\sigma_1^{1(1)} < 0$ だから、(72) より $\lambda_1^{1(3)} < 0$ を得る。□

ここで、 $\lambda^{1(1)}$ と $\lambda^{1(3)}$ の各成分の符号は次のようになっている。

$$\lambda_1^{1(1)} = 0, \lambda_2^{1(1)} \geq 0, \lambda_3^{1(1)} < 0, \lambda_4^{1(1)} \geq 0 \quad (73)$$

$$\lambda_1^{1(3)} < 0, \lambda_2^{1(3)} \geq 0, \lambda_3^{1(3)} = 0, \lambda_4^{1(3)} \geq 0 \quad (74)$$

点 P_2 と P_3 を入れ替えると (73) と (74) で第 2 座標と第 3 座標が入れ替わり、その結果、成分の符号が Case 3-2 と (69) の組み合わせと同じになる。Case 3-2 の証明では Q^0 の重心座標 λ^0 は使われず、 $\lambda^{1(1)}$ と $\lambda^{1(3)}$ のみで証明されているので、Case 3-2 の証明で P_2 と P_3 を入れ替えれば、それが Case 3-3 の証明になる。したがって、この Case 3-3 において、次の定理を得る。

定理 11 $Q^{2\dagger} = \pi(Q^0 | L(P_2, P_4))$ とおく。最小包含円 Γ の中心は $Q^* = Q^{2\dagger}$ であり半径は $d^* = d(P_2, Q^{2\dagger})$ である。

3 通信路容量の問題

我々は、第 2 章において最小包含円問題に関するいくつかの定理をユークリッド幾何を利用して導いた。特に、ユークリッド空間における距離、内積、Pythagoras の定理、射影を利用して、射影アルゴリズムによって最小包含円の中心 Q^* を探索する方法を得た。通信路容量の問題は (7) と (9) との類似性に基づいて、ユークリッド空間における最小包含円の問題と類似の幾何構造を持つ。本章では、第 2 章の結果に基づいて、通信路容量の問題を幾何的に考察して通信路容量を達成する出力分布 Q^* を探索する射影アルゴリズムを開発する。

3.1 情報幾何

通信路容量の問題の基盤となる幾何は、ユークリッド幾何ではなく「情報幾何」である [1]。ユークリッド幾何と情報幾何の異なる点は、ユークリッド幾何では 1 つの座標系を用いるのに対して、情報幾何では互いに双対な 2 つの座標系を用いることである。情報幾何は、実数 α によって α 幾何という統一的な幾何構造の族として考えられる。ユークリッド幾何は $\alpha = 0$ 、 Δ^n の幾何は $\alpha = \pm 1$ という特別な値に対応する幾何として捉えることができる。したがって、ユークリッド幾何も情報幾何の特別な場合と考えることもできる。 α 幾何では、 α divergence、内積、Pythagoras の定理、 α 射影が使える。最小包含円問題に関する定理の証明において利用したユークリッド幾何の性質が距離、内積、Pythagoras の定理、射影だけなので、そこで得られた定理やアルゴリズムは容易に通信路容量の問題に適用できると期待される。実際に、以下においてそれを示す。

3.2 Δ^n 上の幾何構造

Δ^n を出力アルファベット $\{y_1, \dots, y_n\}$ 上で全て正の確率からなる確率分布全体、すなわち、

$$\begin{aligned} \Delta^n &= \{Q = (Q_1, \dots, Q_n) | Q_j > 0, j = 1, \dots, n, \\ &\quad \sum_{j=1}^n Q_j = 1\} \end{aligned} \quad (75)$$

と定義する。 Δ^n 上に統計的な幾何構造が以下のように導入される [1]。

3.2.1 双対座標系

Δ^n 上に 2 つの座標系、言い換えると、 $Q = (Q_1, \dots, Q_n) \in \Delta^n$ を指定する 2 つの方法が与えられる。

η 座標: $\eta = (\eta_2, \dots, \eta_n)$, $\eta_j = Q_j$, $j = 2, \dots, n$,

θ 座標: $\theta = (\theta_2, \dots, \theta_n)$, $\theta_j = \log \frac{Q_j}{Q_1}$, $j = 2, \dots, n$

η 座標系と θ 座標系は互いに双対な座標系である [1]。

3.2.2 測地線

η 座標に関する直線を η 測地線という。すなわち、 $Q^1, Q^2 \in \Delta^n$ の η 座標を、それぞれ、 η^1, η^2 とするとき、 Q^1, Q^2 を通る η 測地線は

$$\eta(t) = (1-t)\eta^1 + t\eta^2 \in \Delta^n, t \in \mathbb{R} \quad (76)$$

と定義される。

また、 θ 座標に関する直線を θ 測地線という。すなわち、 $Q^1, Q^2 \in \Delta^n$ の θ 座標を、それぞれ、 θ^1, θ^2 とするとき、 Q^1, Q^2 を通る θ 測地線は

$$\theta(t) = (1-t)\theta^1 + t\theta^2 \in \Delta^n, t \in \mathbb{R} \quad (77)$$

と定義される。

3.2.3 内積

3点 $Q^1, Q^2, Q^3 \in \Delta^n$ を考える。 Q^1, Q^2, Q^3 の Q^2 における内積 $(Q^1 - Q^2, Q^3 \ominus Q^2)$ が次のように定義される。 Q^1, Q^2 の η 座標を、それぞれ、 $\eta^1 = (\eta_j^1)_{j=2, \dots, n}, \eta^2 = (\eta_j^2)_{j=2, \dots, n}$ とし、 Q^2, Q^3 の θ 座標を、それぞれ、 $\theta^2 = (\theta_j^2)_{j=2, \dots, n}, \theta^3 = (\theta_j^3)_{j=2, \dots, n}$ とする。このとき、内積が

$$(Q^1 - Q^2, Q^3 \ominus Q^2) = \sum_{j=2}^n (\eta_j^1 - \eta_j^2)(\theta_j^3 - \theta_j^2) \quad (78)$$

と定義される。これは、 Q^2 を通る2つの測地線 $\eta(t) = (1-t)\eta^2 + t\eta^1$ と $\theta(t) = (1-t)\theta^2 + t\theta^3$ の Q^2 における接ベクトル $d\eta(t)/dt|_{t=0} = \eta^1 - \eta^2$ と $d\theta(t)/dt|_{t=0} = \theta^3 - \theta^2$ の（通常の意味での）内積になっている。

3.2.4 直交性

$Q^1, Q^2, Q^3 \in \Delta^n$ に対して、 Q^1, Q^2 を通る η 測地線 $\eta(t) = (1-t)\eta^2 + t\eta^1$ と Q^2, Q^3 を通る θ 測地線 $\theta(t) = (1-t)\theta^2 + t\theta^3$ が、 Q^2 において直交する、とは

$$(Q^1 - Q^2, Q^3 \ominus Q^2) = 0 \quad (79)$$

となることである [1]。

3.2.5 アファイン部分空間

$P_1, \dots, P_m \in \Delta^n$ に対して、 P_1, \dots, P_m で張られるアファイン部分空間 $L(P_1, \dots, P_m) \subset \Delta^n$ が

$$L(P_1, \dots, P_m) = \left\{ \sum_{i=1}^m \lambda_i P_i \mid \sum_{i=1}^m \lambda_i = 1 \right\} \cap \Delta^n \quad (80)$$

と定義される。 $L(P_1, \dots, P_m)$ は η 座標に関して平坦な Δ^n の部分多様体である [1]。

3.2.6 KL 情報量による射影

$L \subset \Delta^n$ を任意の部分集合とし、 $Q^0 \in \Delta^n$ とする。 $\min_{Q \in L} D(Q \| Q^0)$ を達成する $Q = Q^1$ を Q^0 の L への射影といい、 $Q^1 = \pi(Q^0 | L)$ と表す。以下では、 L としてアファイン部分空間だけを考える。

以上の準備のもとで以下の補題が成り立つ。

補題 19 $P_i (i = 1, \dots, m), Q, R \in \Delta^n$ に対して、 $\sigma_i = (P_i - Q, R \ominus Q)$, $i = 1, \dots, m$ とする。 $\sum_{i=1}^m \lambda_i = 1$ のとき、

$$\sum_{i=1}^m \lambda_i \sigma_i = \left(\sum_{i=1}^m \lambda_i P_i - Q, R \ominus Q \right)$$

証明 内積の定義と簡単な計算による。 \square

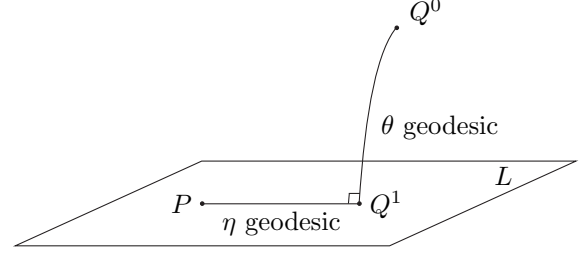


図 3: アファイン部分空間への射影

補題 20 任意の $P, Q, R \in \Delta^n$ に対して

$$(P - Q, R \ominus Q) = -(Q - P, R \ominus P) + (P - Q, P \ominus Q)$$

証明 簡単な計算による。 \square

補題 21 任意の $P, Q, R \in \Delta^n$ に対して

$$(P - Q, R \ominus Q) = D(P \| Q) + D(Q \| R) - D(P \| R)$$

証明 簡単な計算による。([1] 参照) \square

3.2.7 Pythagoras の定理

Δ^n の3点 P, Q, R に対して、次のように Pythagoras の定理とその不等式版が成り立つ。

定理 12 (Pythagoras) $P, Q, R \in \Delta^n$ に対して

$$\begin{aligned} (P - Q, R \ominus Q) &\geq 0 \\ \iff D(P \| Q) + D(Q \| R) &\geq D(P \| R) \end{aligned} \quad (81)$$

証明 補題 21 による。([1] 参照) \square

補題 22 $P_1, \dots, P_m \in \Delta^n$ に対して $L_0 = L(P_1, \dots, P_m)$ を P_1, \dots, P_m で張られるアファイン部分空間とする。このとき、任意の $Q^0 \in \Delta^n$ に対して射影 $Q^1 = \pi(Q^0 | L_0)$ が唯一存在する。

証明 [1] 参照。 \square

補題 23 $L \subset \Delta^n$ をアファイン部分空間とする。 $Q^0 \in \Delta^n$ に対して $Q^1 = \pi(Q^0 | L)$ であることと、任意の $P \in L$ に対して $(P - Q^1, Q^0 \ominus Q^1) = 0$ が成り立つことは同値である（下図参照）。

証明 [1] 参照。 \square

3.3 通信路容量を達成する出力分布について

式 (1) の通信路行列 Φ の行ベクトル $P_i, i = 1, \dots, m$ が Δ^n において一般の位置にある場合に通信路容量を達成する出力分布 Q^* および通信路容量 C を計算する。第2章では、最小包含円の問題を、点 P_1, \dots, P_m が一般の位置にあるとき、ユークリッド幾何に基づいて考察した。本章では、通信路行列の行ベクトル $P_i, i = 1, \dots, m$ が一般の位置にあるとき、通信路容量の問題を情報幾何に基づいて考察する。

3.4 P_1, \dots, P_m からの等距離点について

通信路行列 Φ に対して、行列 $\Psi \in \mathbb{R}^{(m-1) \times n}$ を

$$\Psi = \begin{pmatrix} P_2 - P_1 \\ \vdots \\ P_m - P_1 \end{pmatrix} = \begin{pmatrix} P_{21} - P_{11} & \cdots & P_{2n} - P_{1n} \\ \vdots & & \vdots \\ P_{m1} - P_{11} & \cdots & P_{mn} - P_{1n} \end{pmatrix} \quad (82)$$

と定義する。このとき、点 P_1, \dots, P_m が一般の位置にある、とは、

$$\text{rank } \Psi = m - 1 \quad (83)$$

となることである。 Ψ の第 1 列を削除した行列を $\Psi' \in \mathbb{R}^{(m-1) \times (n-1)}$ とする。すなわち、

$$\Psi' = \begin{pmatrix} P_{22} - P_{12} & \cdots & P_{2n} - P_{1n} \\ \vdots & & \vdots \\ P_{m2} - P_{12} & \cdots & P_{mn} - P_{1n} \end{pmatrix} \quad (84)$$

である。このとき、(83) より $\text{rank } \Psi' = m - 1$ である。なぜならば、もし Ψ' の第 1 行ベクトルが、その第 2 行から第 $m - 1$ 行ベクトルの線形和で表すことができたすると、 $\sum_{j=1}^n P_{ij} = 1$ に注意して、 Ψ の第 1 行ベクトルがその第 2 行から第 $m - 1$ 行ベクトルの線形和で表してしまうからである。

さて、最小包含円問題のときの類似として、 P_1, \dots, P_m との KL 情報量が等しい分布について考える。すなわち、

$$D(P_i \| \hat{Q}) = D(P_1 \| \hat{Q}), i = 2, \dots, m \quad (85)$$

を満たす $\hat{Q} \in \Delta^n$ を求める。 \hat{Q} の θ 座標を $\hat{\theta} = (\hat{\theta}_2, \dots, \hat{\theta}_n)$ として、 $H(P_i) = -\sum_{j=1}^n P_{ij} \log P_{ij}$ (P_i のエントロピー), $i = 1, \dots, m$ とおくと、(85) より簡単な計算によって、

$$\begin{aligned} D(P_i \| \hat{Q}) - D(P_1 \| \hat{Q}) &= -\sum_{j=2}^n (P_{ij} - P_{1j}) \hat{\theta}_j - H(P_i) + H(P_1) \\ &= 0, i = 2, \dots, m \end{aligned} \quad (86)$$

となる。ここで、 $\mathbf{b} = (-H(P_i) + H(P_1))_{i=2, \dots, m}$ とおくと、(86) は

$$\Psi' {}^t \hat{\theta} = {}^t \mathbf{b} \quad (87)$$

と書ける。 $\text{rank } \Psi' = m - 1$ なので、方程式 (87) は (必ずしも一意とは限らないが) 解 $\hat{\theta}$ を持つ。その解 $\hat{\theta} = (\hat{\theta}_2, \dots, \hat{\theta}_n)$ によって

$$\begin{cases} \hat{Q}_1 = \left(1 + \sum_{j=2}^n \exp \hat{\theta}_j \right)^{-1} \\ \hat{Q}_j = \hat{Q}_1 \exp \hat{\theta}_j, j = 2, \dots, n \end{cases} \quad (88)$$

とすると、 $\hat{Q} = (\hat{Q}_1, \dots, \hat{Q}_n) \in \Delta^n$ となり、 \hat{Q} の θ 座標は $\hat{\theta}$ となる。次の補題が成り立つ。

補題 24 $L_0 = L(P_1, \dots, P_m)$ とおく。方程式 (87) の任意の解 $\hat{\theta}$ によって (88) のように $\hat{Q} \in \Delta^n$ を作り $Q^0 = \pi(\hat{Q}|L_0)$ とおく。 Q^0 は一意である。

証明 (87), (88) を満たす $\hat{Q}, \hat{Q}' \in \Delta^n$ に対して $Q^0 = \pi(\hat{Q}|L_0)$, $Q^{0'} = \pi(\hat{Q}'|L_0)$ とおく。 $Q^0 = Q^{0'}$ を示せばよい。 $Q^0 = (Q_j^0)_{j=1, \dots, n}$, $Q^{0'} = (Q_j^{0'})_{j=1, \dots, n}$ とおくと、 \hat{Q}, \hat{Q}' はともに (85) を満たすので、定理 12 および補題 23 より

$$\sum_{j=1}^n P_{ij} \log \frac{P_{ij}}{Q_j^0} = \sum_{j=1}^n P_{1j} \log \frac{P_{1j}}{Q_j^0}, i = 2, \dots, m \quad (89)$$

$$\sum_{j=1}^n P_{ij} \log \frac{P_{ij}}{Q_j^{0'}} = \sum_{j=1}^n P_{1j} \log \frac{P_{1j}}{Q_j^{0'}}, i = 2, \dots, m \quad (90)$$

が成り立つ。(90) から (89) を引いて

$$\sum_{j=1}^n P_{ij} \log \frac{Q_j^0}{Q_j^{0'}} = \sum_{j=1}^n P_{1j} \log \frac{Q_j^0}{Q_j^{0'}}, i = 2, \dots, m \quad (91)$$

となる。(91) は $i = 1$ のときも自明に成り立つので、

$$\sum_{j=1}^n P_{ij} \log \frac{Q_j^0}{Q_j^{0'}} = \kappa (\text{定数}), i = 1, \dots, m \quad (92)$$

とおくことができる。 $Q^0, Q^{0'} \in L_0$ の P_1, \dots, P_m に関する重心座標を、それぞれ、 $\lambda^0 = (\lambda_i^0)_{i=1, \dots, m}$, $\lambda^{0'} = (\lambda_i^{0'})_{i=1, \dots, m}$ とおくと、(92) より

$$\kappa = \sum_{i=1}^m \lambda_i^0 \sum_{j=1}^n P_{ij} \log \frac{Q_j^0}{Q_j^{0'}} = D(Q^0 \| Q^{0'}) \quad (93)$$

$$\kappa = \sum_{i=1}^m \lambda_i^{0'} \sum_{j=1}^n P_{ij} \log \frac{Q_j^0}{Q_j^{0'}} = -D(Q^{0'} \| Q^0) \quad (94)$$

となる。(93) から (94) を引いて

$$0 = D(Q^0 \| Q^{0'}) + D(Q^{0'} \| Q^0) \quad (95)$$

となり、 $Q^0 = Q^{0'}$ を得る。□

補題 24 によって存在と一意性が保証されている $Q^0 \in L_0$ を P_1, \dots, P_m からの等距離点という。正しくは等 KL 情報量分布と呼ぶべきだろうが、最小包含円問題との類似で等距離点と呼ぶことにする。

さて、さらに $L_k = L(P_{k+1}, \dots, P_m)$, $k = 0, 1, \dots, m - 2$ とおく。 $L_0 \supset L_1 \supset \dots \supset L_k \supset \dots$ は次元が 1 ずつ小さくなる Δ^n のアファイン部分空間の減少列である。

Q^0 の L_1 への射影、すなわち、 $\min_{Q \in L_1} D(Q \| Q^0)$ を達成する点を Q^1 として、 $Q^1 = \pi(Q^0 | L_1)$ と表す。同様に、 $Q^k = \pi(Q^{k-1} | L_k)$, $k = 1, \dots, m - 2$ とする。以下の補題が成り立つ。

補題 25 $Q^k = \pi(Q^0 | L_k)$, $k = 0, 1, \dots, m - 2$.

証明 $k = 0, 1$ のときは定義から明らか。 $k = 2$ の場合を証明する。すなわち、 $Q^{2'} = \pi(Q^0|L_2)$ において $Q^{2'} = Q^2 \equiv \pi(Q^1|L_2)$ を示す。 $Q^1 = \pi(Q^0|L_1)$ なので、定理 12 および補題 23 より、任意の $Q \in L_1$ に対して Pythagoras の定理

$$D(Q\|Q^0) = D(Q\|Q^1) + D(Q^1\|Q^0) \quad (96)$$

が成り立つ。(96) より $Q \in L_2 (\subset L_1)$ に関して $D(Q\|Q^0)$ を最小化することと $D(Q\|Q^1)$ を最小化することは同値である。補題 22 より射影 $\pi(Q^1|L_2)$ および $\pi(Q^0|L_2)$ は一意的なので $Q^{2'} = Q^2$ を得る。 $k \geq 3$ に対しては数学的帰納法によって証明される。 \square

補題 26 $(Q^i - Q^k, Q^0 \ominus Q^k) = (Q^i - Q^k, Q^i \ominus Q^k)$, $i = 0, 1, \dots, k, k = 0, 1, \dots, m-2$.

証明 補題 20 より

$$\begin{aligned} & (Q^i - Q^k, Q^0 \ominus Q^k) \\ &= -(Q^k - Q^i, Q^0 \ominus Q^i) + (Q^i - Q^k, Q^i \ominus Q^k) \quad (97) \end{aligned}$$

となる。補題 25 より $Q^i = \pi(Q^0|L_i)$ であり、 $k \geq i$ より $Q^k \in L_i$ なので、補題 23 より $(Q^k - Q^i, Q^0 \ominus Q^i) = 0$ を得る。 \square

補題 27 任意の $P, Q \in \Delta^n$, $P \neq Q$ に対して $(P - Q, P \ominus Q) > 0$ である。

証明 補題 21 より $(P - Q, P \ominus Q) = D(P\|Q) + D(Q\|P) > 0$ を得る。 \square

3.5 射影アルゴリズムによる Q^* の探索

与えられた通信路行列 Φ に対して、通信路容量を達成する出力分布を Q^* とし通信路容量を C とする。本節では、通信路行列の行ベクトル P_1, \dots, P_m が一般の位置にあるとして、2.7 節と同様の場合に Q^* , C が射影アルゴリズムによって得られることを示す。

3.6 状況 1 [毎回の射影で負の座標成分がちょうど 1 個ある場合]

2.8 節と同様に下記を仮定する。

状況 1 の仮定 $P_1, \dots, P_m \in \Delta^n$ は一般の位置にあるとし、 $L_k = L(P_{k+1}, \dots, P_m)$, $k = 0, 1, \dots, m-2$ を P_{k+1}, \dots, P_m で張られるアファイン部分空間とする。 $Q^0 \in L_0$ を P_1, \dots, P_m からの等距離点として、 $Q^k = \pi(Q^{k-1}|L_k)$, $k = 1, \dots, m-2$ と定義する。 Q^k の P_1, \dots, P_m に関する重心座標を $\lambda^k = (\lambda_1^k, \dots, \lambda_m^k)$ とする。 $K = 0, 1, \dots, m-2$ とする。 $k = 0, 1, \dots, K-1$ のとき λ^k の成分のうち負のものがちょうど 1 個あり、 $k = K$ のとき λ^K の成分はすべて非負であるとする。すなわち、 $k = 0, 1, \dots, K-1$ のとき、

$$\lambda_i^k \begin{cases} = 0, & i = 1, \dots, k, \\ < 0, & i = k+1, \\ > 0, & i = k+2, \dots, m \end{cases} \quad (98a)$$

$$\lambda_i^k \begin{cases} < 0, & i = k+1, \\ > 0, & i = k+2, \dots, m \end{cases} \quad (98b)$$

$$\lambda_i^k \begin{cases} > 0, & i = k+2, \dots, m \end{cases} \quad (98c)$$

となり、また $k = K$ のとき、

$$\lambda_i^K \begin{cases} = 0, & i = 1, \dots, K, \\ > 0, & i = K+1, \dots, m \end{cases} \quad (99a)$$

$$\lambda_i^K \begin{cases} > 0, & i = K+1, \dots, m \end{cases} \quad (99b)$$

となる。ただし、 $K = 0$ のときは $\lambda_i^0 > 0$, $i = 1, \dots, m$ とみなす。

(状況 1 の仮定終わり)

ここで、内積 $\sigma_i \equiv (P_i - Q^K, Q^0 \ominus Q^K)$, $i = 1, \dots, m$ を考える。次の補題が成り立つ。

補題 28 $\sigma_i \begin{cases} < 0, & i = 1, \dots, K, \\ = 0, & i = K+1, \dots, m. \end{cases}$

証明 補題 25 より $Q^K = \pi(Q^0|L_K)$ であり、 $P_{K+1}, \dots, P_m \in L_K$ なので、補題 23 より

$$\sigma_i = 0, \quad i = K+1, \dots, m \quad (100)$$

を得る。次に、 $\sigma_i < 0$, $i = 1, \dots, K$ を数学的帰納法によって、 $i = K, K-1, \dots, 1$ の順に証明する。

(I) $\sigma_K < 0$ を証明する：

(98a), (100) より

$$\begin{aligned} \lambda_K^{K-1} \sigma_K &= \sum_{i=1}^m \lambda_i^{K-1} \sigma_i \\ &= (Q^{K-1} - Q^K, Q^0 \ominus Q^K) \\ &= (Q^{K-1} - Q^K, Q^{K-1} \ominus Q^K) \quad (\text{補題 26 より}) \\ &> 0 \quad (\text{補題 27 より}) \end{aligned} \quad (101)$$

となる。(98b) より $\lambda_K^{K-1} < 0$ なので、(101) より $\sigma_K < 0$ を得る。

(II) $\sigma_K < 0$, $\sigma_{K-1} < 0, \dots, \sigma_{k+1} < 0$ を仮定して $\sigma_k < 0$ を証明する：

(98a), (100) より

$$\begin{aligned} \lambda_k^{k-1} \sigma_k + \sum_{i=k+1}^K \lambda_i^{k-1} \sigma_i &= \sum_{i=1}^m \lambda_i^{k-1} \sigma_i \\ &= (Q^{k-1} - Q^K, Q^0 \ominus Q^K) \\ &= (Q^{k-1} - Q^K, Q^{k-1} \ominus Q^K) \\ &> 0 \end{aligned} \quad (102)$$

となる。帰納法の仮定より $\sigma_i < 0$, $i = k+1, \dots, K$, (98b) より $\lambda_k^{k-1} < 0$, (98c) より $\lambda_i^{k-1} > 0$, $i = k+1, \dots, K$ なので、(102) より $\sigma_k < 0$ を得る。

(III) 上記 (I), (II) より、 $\sigma_i < 0$, $i = 1, \dots, K$ を得る。 \square

補題 29 $D(P_i\|Q^K) < D(P_{K+1}\|Q^K) = \dots = D(P_m\|Q^K)$, $i = 1, \dots, K$.

証明 補題 28 および定理 12 より。 \square

定理 13 通信路容量を達成する出力分布は $Q^* = Q^K$ であり、通信路容量は $C = D(P_{K+1}\|Q^K)$ である。

証明 (99a), (99b) と補題 29 および Kuhn-Tucker 条件 (8) より。 \square

上記の補題 28, 29, 定理 13 とその証明は第 2 章の最小包含円問題における補題 7, 8, 定理 2 とその証明に類似していることがわかる。最小包含円問題と通信路容量問題が α 幾何に共通な性質だけを使って証明できる問題だからである。

3.7 状況 2 [$m = 2, 3, 4$ の場合]

一般の位置にある m 個の行ベクトル P_1, \dots, P_m を持つ通信路行列 Φ に対して、その通信路容量を達成する出力分布 Q^* と通信路容量 C を射影アルゴリズムによって求める。第 2 章と同様に、目標は任意の m に対して計算することであるが、現時点では以下に示すように $m = 2, 3, 4$ の場合だけ解決できている。

前節の結果からわかるように、補題や定理の証明は最小包含円問題と通信路容量問題の場合でほぼ同様である。本節においてもそれは同様なので、以下の補題と定理の証明は簡単に述べるかあるいは省略する。

3.7.1 $m = 2$ の場合

2 つの異なる行ベクトル $P_1, P_2 \in \Delta^n$ からなる通信路行列 Φ の通信路容量について考える。 $L_0 = L(P_1, P_2)$ とし、 $Q^0 \in L_0$ を P_1, P_2 からの等距離点とする。すなわち、 Q^0 は

$$D(P_1 \| Q^0) = D(P_2 \| Q^0), Q^0 \in L_0 \quad (103)$$

を満たす点である。 Q^0 の P_1, P_2 に関する重心座標を $\lambda^0 = (\lambda_1^0, \lambda_2^0)$ とおくと、 Q^0 が L_0 において P_1 と P_2 の間にあるので、 $\lambda_1^0 > 0, \lambda_2^0 > 0$ である。したがって、次の定理を得る。

定理 14 通信路容量を達成する出力分布は $Q^* = Q^0$ であり通信路容量は $C = D(P_1 \| Q^0)$ である。

証明 (103) と Kuhn-Tucker 条件 (8) による。 \square

3.7.2 $m = 3$ の場合

通信路行列 Φ が一般の位置にある 3 つの行ベクトル $P_1, P_2, P_3 \in \Delta^n$ からなるとする。 $L_0 = L(P_1, P_2, P_3)$ として、 $Q^0 \in L_0$ を P_1, P_2, P_3 からの等距離点とする。 Q^0 の P_1, P_2, P_3 に関する重心座標を $\lambda^0 = (\lambda_1^0, \lambda_2^0, \lambda_3^0)$ とする。ここで、一般性を失うことなく、以下の場合分けをすることができる。

$$\text{Case 1: } \lambda_1^0 \geq 0, \lambda_2^0 \geq 0, \lambda_3^0 \geq 0$$

$$\text{Case 2: } \lambda_1^0 < 0, \lambda_2^0 \geq 0, \lambda_3^0 \geq 0$$

最小包含円問題のときと同様に、Case 1, 2 ですべての場合が尽くされている。

Case 1

次の定理を得る。

定理 15 通信路容量を達成する出力分布は $Q^* = Q^0$ であり通信路容量は $C = D(P_1 \| Q^0)$ である。

証明 Kuhn-Tucker 条件 (8) による。 \square

Case 2

$L_1 = L(P_2, P_3)$ として $Q^1 = \pi(Q^0 | L_1)$ とする。内積 $\sigma_i = (P_i - Q^1, Q^0 \ominus Q^1)$, $i = 1, 2, 3$ を考える。次の補題が成り立つ。

$$\text{補題 30 } \sigma_i \begin{cases} < 0, & i = 1, \\ = 0, & i = 2, 3. \end{cases}$$

補題 31 $D(P_1 \| Q^1) < D(P_2 \| Q^1) = D(P_3 \| Q^1)$.

定理 16 通信路容量を達成する出力分布は $Q^* = Q^1$ であり、通信路容量は $C = D(P_2 \| Q^1)$ である。

上記の補題 30, 31, 定理 16 は補題 9, 10, 定理 5 と同様に証明できる。

3.7.3 $m = 4$ の場合

通信路行列 Φ が一般の位置にある 4 つの行ベクトル $P_1, P_2, P_3, P_4 \in \Delta^n$ からなるとする。 $L_0 = L(P_1, P_2, P_3, P_4)$ とする。 P_1, \dots, P_4 からの等距離点 $Q^0 \in L_0$ の P_1, \dots, P_4 に関する重心座標を $\lambda^0 = (\lambda_i^0)_{i=1, \dots, 4}$ とする。一般性を失うことなく、以下のように場合分けすることができる。

$$\text{Case 1: } \lambda_1^0 \geq 0, \lambda_2^0 \geq 0, \lambda_3^0 \geq 0, \lambda_4^0 \geq 0$$

$$\text{Case 2: } \lambda_1^0 < 0, \lambda_2^0 \geq 0, \lambda_3^0 \geq 0, \lambda_4^0 \geq 0$$

$$\text{Case 3: } \lambda_1^0 < 0, \lambda_2^0 < 0, \lambda_3^0 \geq 0, \lambda_4^0 \geq 0$$

Case 1

Kuhn-Tucker 条件 (8) によって次の定理を得る。

定理 17 通信路容量を達成する出力分布は $Q^* = Q^0$ であり、通信路容量は $C = D(P_1 \| Q^0)$ である。

Case 2

$L_1 = L(P_2, P_3, P_4)$ として、 $Q^1 = \pi(Q^0 | L_1)$ とし、 Q^1 の P_1, \dots, P_4 に関する重心座標を $\lambda^1 = (\lambda_i^1)_{i=1, \dots, 4}$ とおく。 $P_1 \notin L_1$ より $\lambda_1^1 = 0$ である。 $\lambda_i^1, i = 2, 3, 4$ の符号に基づいて、一般性を失うことなく、さらに以下の場合分けができる。

$$\text{Case 2-1: } \lambda_1^1 = 0, \lambda_2^1 \geq 0, \lambda_3^1 \geq 0, \lambda_4^1 \geq 0$$

$$\text{Case 2-2: } \lambda_1^1 = 0, \lambda_2^1 < 0, \lambda_3^1 \geq 0, \lambda_4^1 \geq 0$$

内積 $\sigma_i = (P_i - Q^1, Q^0 \ominus Q^1)$, $i = 1, \dots, 4$ を考える。

Case 2-1

次の補題と定理が成り立つ。

$$\text{補題 32 } \sigma_i \begin{cases} < 0, & i = 1, \\ = 0, & i = 2, 3, 4. \end{cases}$$

補題 33 $D(P_1 \| Q^1) < D(P_2 \| Q^1) = \dots = D(P_4 \| Q^1)$.

定理 18 通信路容量を達成する出力分布は $Q^* = Q^1$ であり、通信路容量は $C = D(P_2 \| Q^1)$ である。

上記の補題 32, 33, 定理 18 は補題 11, 12, 定理 7 と同様に証明できる。

Case 2-2

この場合は 3.6 節と同じ場合であり、すでに証明済みである。すなわち、次の定理が成り立つ。

定理 19 $Q^2 = \pi(Q^1 | L(P_3, P_4))$ とおく。通信路容量を達成する出力分布は $Q^* = Q^2$ であり、通信路容量は $C = D(P_3 \| Q^2)$ である。

以上で Case 2 は完了した。次に、Case 3 を考える。

Case 3

等距離点 $Q^0 \in L_0$ に対して、 $Q^{1(1)} = \pi(Q^0 | L(P_2, P_3, P_4))$, $Q^{1(2)} = \pi(Q^0 | L(P_1, P_3, P_4))$ とおく。 $Q^{1(1)}, Q^{1(2)}$ の P_1, \dots, P_4 に関する重心座標を、それぞれ、 $\lambda^{1(1)} =$

$(\lambda_i^{1(1)})_{i=1,\dots,4}, \lambda^{1(2)} = (\lambda_i^{1(2)})_{i=1,\dots,4}$ とおく。そして、内積

$$\sigma_i^{1(1)} = (P_i - Q^{1(1)}, Q^0 \ominus Q^{1(1)}), i = 1, \dots, 4, \quad (104)$$

$$\sigma_i^{1(2)} = (P_i - Q^{1(2)}, Q^0 \ominus Q^{1(2)}), i = 1, \dots, 4 \quad (105)$$

を考える。次の補題が成り立つ。

補題 34

$$\sigma_i^{1(1)} \begin{cases} < 0, & i = 1, \\ = 0, & i = 2, 3, 4, \end{cases} \quad \sigma_i^{1(2)} \begin{cases} < 0, & i = 2, \\ = 0, & i = 1, 3, 4. \end{cases}$$

補題 35 $\lambda_1^{1(2)} \sigma_1^{1(1)} + \lambda_2^{1(1)} \sigma_2^{1(2)} > 0$.

上記の補題 34, 35 は補題 13, 14 と同様に証明できる。

補題 34, 35 より $\lambda_1^{1(2)} < 0$ または $\lambda_2^{1(1)} < 0$ が成り立つので、一般性を失うことなく、 $\lambda_1^{1(2)} < 0$ として、 $\lambda^{1(1)}$ の各成分の符号によって Case 3 をさらに次のように場合分けする。

$$\text{Case 3-1: } \lambda_1^{1(1)} = 0, \lambda_2^{1(1)} \geq 0, \lambda_3^{1(1)} \geq 0, \lambda_4^{1(1)} \geq 0$$

$$\text{Case 3-2: } \lambda_1^{1(1)} = 0, \lambda_2^{1(1)} < 0, \lambda_3^{1(1)} \geq 0, \lambda_4^{1(1)} \geq 0$$

$$\text{Case 3-3: } \lambda_1^{1(1)} = 0, \lambda_2^{1(1)} \geq 0, \lambda_3^{1(1)} < 0, \lambda_4^{1(1)} \geq 0$$

Case 3-1

次の定理が成り立つ。

定理 20 通信路容量を達成する出力分布は $Q^* = Q^{1(1)}$ であり、通信路容量は $C = D(P_2 \| Q^{1(1)})$ である。

証明 定理 9 の証明と同様。 \square

Case 3-2

$Q^2 \equiv \pi(Q^{1(1)} | L(P_3, P_4)) (= \pi(Q^0 | L(P_3, P_4)))$ として、内積 $\sigma_i^2 = (P_i - Q^2, Q^0 \ominus Q^2), i = 1, \dots, 4$ を考える。次の補題と定理が成り立つ。

$$\text{補題 36 } \sigma_i^2 \begin{cases} < 0, & i = 1, 2, \\ = 0, & i = 3, 4. \end{cases}$$

補題 37 $D(P_i \| Q^2) < D(P_3 \| Q^2) = D(P_4 \| Q^2), i = 1, 2$.

定理 21 通信路容量を達成する出力分布は $Q^* = Q^2$ であり、通信路容量は $C = D(P_3 \| Q^2)$ である。

上記の補題 36, 37, 定理 21 は補題 15, 16, 定理 10 と同様に証明できる。

Case 3-3

$Q^{1(3)} \equiv \pi(Q^0 | L(P_1, P_2, P_4))$ として $Q^{1(3)}$ の P_1, \dots, P_4 に関する重心座標を $\lambda^{1(3)} = (\lambda_i^{1(3)})_{i=1,\dots,4}$ とする。内積 $\sigma_i^{1(3)} = (P_i - Q^{1(3)}, Q^0 \ominus Q^{1(3)}), i = 1, \dots, 4$ を考える。次の補題が成り立つ。

$$\text{補題 38 } \sigma_i^{1(3)} \begin{cases} > 0, & i = 3, \\ = 0, & i = 1, 2, 4. \end{cases}$$

補題 39 $\lambda_1^{1(3)} < 0$.

定理 22 $Q^{2\dagger} \equiv \pi(Q^0 | L(P_2, P_4))$ とおく。通信路容量を達成する出力分布は $Q^* = Q^{2\dagger}$ であり、通信路容量は $C = D(P_2 \| Q^{2\dagger})$ である。

上記の補題 38, 39, 定理 22 は補題 17, 18, 定理 11 と同様に証明できる。

4 任意の点配置に対する最適点の探索

これまでの章において、与えられた点 P_1, \dots, P_m が \mathbb{R}^n または Δ^n において一般の位置にあること、すなわち、

$$\text{rank} \begin{pmatrix} P_2 - P_1 \\ \vdots \\ P_m - P_1 \end{pmatrix} = m - 1 \quad (106)$$

であることを仮定していた。この仮定によって P_1, \dots, P_m からの等距離点 Q^0 が存在するので、 Q^0 をアファイン部分空間に射影する射影アルゴリズムによって最小包含円の中心 Q^* 、あるいは通信路容量を達成する点 Q^* を計算することができる。

しかし、任意に与えられた m 個の点 P_1, \dots, P_m が (106) を満たすことはほとんどないだろう。もし (106) が成り立つとすれば $m - 1 \leq n$ でなければならないから、例えば、 \mathbb{R}^2 において与えられた 4 点は (106) を満たさない。(106) が成り立たなければこれまでの章で考えてきた射影アルゴリズムは全く使うことができない。

そこで、本章では、与えられた点 P_1, \dots, P_m の配置を少し変形して (106) が成り立つようにして、これまでの射影アルゴリズムが適用できる方法を考える。本章で提案する方法によると、与えられた点の配置が (106) を満足するかどうかなどの点の位置関係について事前にチェックする必要がない。例えば、ある点が他のいくつかの点の凸包に含まれるとか、点全体が張る部分空間が何次元であるか、などを知る必要がない。ただし、与えられた元の問題を少し変形するので、得られた結果が元の問題に対する解と異なる心配があるが、それについては第 5 章で議論する。

図 4 の例を考える。 \mathbb{R}^2 において 3 点 P_1, P_2, P_3 が一直線上にある。この場合、 P_1 が P_2 と P_3 の間にある、などの位置関係を知らずに最小包含円 $\Gamma(P_1, P_2, P_3)$ を計算するにはどうしたらいいだろうか？すぐに考えられる方法は、3 点から 2 点ずつを選んで $\Gamma(P_1, P_2), \Gamma(P_1, P_3), \Gamma(P_2, P_3)$ の中ですべての点を包含するもののうち半径が最小のものが求める最小包含円である。しかし、このような組み合わせ的な方法では、点の個数が多くなると計算量が大きくなり、一般化は困難である。

そこで、 P_1 を少しずらして \tilde{P}_1 とすると、 \tilde{P}_1, P_2, P_3 からの等距離点 Q^0 が存在するので、前章までの射影アルゴリズムによって最小包含円 $\Gamma(\tilde{P}_1, P_2, P_3)$ が得られる。ずらす量が小さければ $\Gamma(\tilde{P}_1, P_2, P_3) = \Gamma(P_1, P_2, P_3)$ となるであろう。

以上のことを基本的なアイデアとして以下のアルゴリズムを与える。

4.1 \mathbb{R}^n の点の次元を上げる方法

\mathbb{R}^n において与えられた（必ずしも一般の位置にあるとは限らない） m 個の点 P_1, \dots, P_m に対して、それら

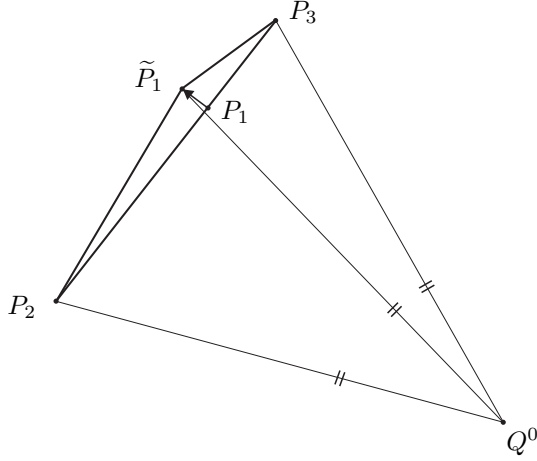


図 4: 点をずらして等距離点を得る

の次元を上げた点 $\tilde{P}_1, \dots, \tilde{P}_m$ を次のように定義する。

$$\begin{aligned} P_1(P_{11}, \dots, P_{1n}) &\rightarrow \tilde{P}_1(P_{11}, \dots, P_{1n}, \varepsilon, 0, \dots, 0) \\ P_2(P_{21}, \dots, P_{2n}) &\rightarrow \tilde{P}_2(P_{21}, \dots, P_{2n}, 0, \varepsilon, \dots, 0) \\ &\vdots \\ P_m(P_{m1}, \dots, P_{mn}) &\rightarrow \tilde{P}_m(P_{m1}, \dots, P_{mn}, 0, \dots, 0, \varepsilon) \end{aligned} \quad (107)$$

ただし, $\varepsilon \in \mathbb{R}, \varepsilon \neq 0$ で $|\varepsilon|$ は十分小さいとする。上記の対応 $P_i \rightarrow \tilde{P}_i$ は次のように書ける。すなわち, \mathbb{R}^m の第 i 基本ベクトル $e_i \in \mathbb{R}^m$ を

$$e_i = (0, \dots, 0, \overset{i \text{ 番目}}{\underset{\uparrow}{1}}, 0, \dots, 0), \quad i = 1, \dots, m \quad (108)$$

として,

$$\mathbb{R}^n \ni P_i \rightarrow \tilde{P}_i = (P_i, \varepsilon e_i) \in \mathbb{R}^{n+m}, \quad i = 1, \dots, m \quad (109)$$

とする。そこで,

$$\tilde{\Phi} = \begin{pmatrix} \tilde{P}_1 \\ \vdots \\ \tilde{P}_m \end{pmatrix} \in \mathbb{R}^{m \times (n+m)} \quad (110)$$

とおくと, $\varepsilon \neq 0$ より $\text{rank } \tilde{\Phi} = m$ であるから $\tilde{P}_1, \dots, \tilde{P}_m$ からの等距離点 $\tilde{Q}^0 \in L(\tilde{P}_1, \dots, \tilde{P}_m) \subset \mathbb{R}^{n+m}$ が (23) により存在する。 \tilde{Q}^0 の $\tilde{P}_1, \dots, \tilde{P}_m$ に関する重心座標を $\tilde{\lambda}^0 = (\tilde{\lambda}_i^0)_{i=1, \dots, m}$ とおく。 $\tilde{M} \equiv \tilde{\Phi}^t \tilde{\Phi} \in \mathbb{R}^{m \times m}$ とおくと, $\text{rank } \tilde{M} = \text{rank } \tilde{\Phi} = m$ より \tilde{M} は正則行列である。 $\tilde{\alpha} = (\|\tilde{P}_1\|^2, \dots, \|\tilde{P}_m\|^2) \in \mathbb{R}^m$ とおくと, (51) を導いたのと同様にして,

$$\tilde{\lambda}^0 = \frac{1}{2} \left(\tilde{\alpha} + \frac{2 - \tilde{\alpha} \tilde{M}^{-1} \mathbf{1}}{\mathbf{1}^t \tilde{M}^{-1} \mathbf{1}} \mathbf{1} \right) \tilde{M}^{-1} \quad (111)$$

を得る。そして, 第 2 章の射影アルゴリズムによって $\tilde{P}_1, \dots, \tilde{P}_m$ に対する最小包含円の中心 \tilde{Q}^* が得られたとすると, ε が十分小さい場合, \tilde{Q}^* は真に求めたい P_1, \dots, P_m に対する最小包含円の中心 Q^* に近い。そ

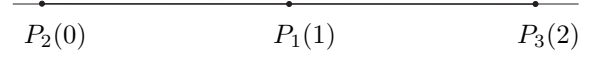


図 5: \mathbb{R} 上に与えられた 3 点の例

こで, \tilde{Q}^* の $\tilde{P}_1, \dots, \tilde{P}_m$ に関する重心座標を $\tilde{\lambda}^*$ として, $\tilde{\lambda}^*$ に含まれる ε を 0 としたものを $\tilde{\lambda}^*|_{\varepsilon=0}$ とすると, $Q^* = \tilde{\lambda}^*|_{\varepsilon=0} \Phi$ が元の最小包含円の中心となる可能性がある。以下に例を与える。

例題 2 \mathbb{R} の 3 点 $P_1(1), P_2(0), P_3(2)$ を考える (図 5 参照)。これらに対して, その次元を上げた点 $\tilde{P}_1, \tilde{P}_2, \tilde{P}_3 \in \mathbb{R}^4$ を以下のように定義する。

$$\begin{aligned} P_1(1) &\rightarrow \tilde{P}_1(1, \varepsilon, 0, 0) \\ P_2(0) &\rightarrow \tilde{P}_2(0, 0, \varepsilon, 0) \\ P_3(2) &\rightarrow \tilde{P}_3(2, 0, 0, \varepsilon) \end{aligned}$$

$\tilde{P}_1, \tilde{P}_2, \tilde{P}_3$ からの等距離点 \tilde{Q}^0 の $\tilde{P}_1, \tilde{P}_2, \tilde{P}_3$ に関する重心座標 $\tilde{\lambda}^0 = (\tilde{\lambda}_i^0)_{i=1, \dots, 3}$ は (111) より

$$\tilde{\lambda}^0 = \left(\frac{-2 + 2\varepsilon^2}{6\varepsilon^2}, \frac{1 + 2\varepsilon^2}{6\varepsilon^2}, \frac{1 + 2\varepsilon^2}{6\varepsilon^2} \right) \quad (112)$$

となる。 $|\varepsilon|$ が十分小さいとき $\tilde{\lambda}_1^0 = (-2 + 2\varepsilon^2)/(6\varepsilon^2) < 0$ となるので, \tilde{P}_1 を除外する。そして, 射影した点 $\tilde{Q}^1 = \pi(\tilde{Q}^0 | L(\tilde{P}_2, \tilde{P}_3))$ の $\tilde{P}_1, \tilde{P}_2, \tilde{P}_3$ に関する重心座標を $\tilde{\lambda}^1 = (\tilde{\lambda}_i^1)_{i=1, \dots, 3}$ として $\tilde{\lambda}^1 = (0, 1/2, 1/2)$ を得る。 $\tilde{\lambda}_i^1 \geq 0, i = 1, \dots, 3$ なので, Kuhn-Tucker 条件 (16) より $\tilde{\lambda}^1 = \tilde{\lambda}^*$ である。そこで, $\varepsilon = 0$ として, $\tilde{\lambda}^*|_{\varepsilon=0} = (0, 1/2, 1/2)$ となり,

$$\begin{aligned} Q^* &= \tilde{\lambda}^*|_{\varepsilon=0} \Phi \\ &= \left(0, \frac{1}{2}, \frac{1}{2} \right) \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} \\ &= 1 \end{aligned}$$

であるから正しく最小包含円 $\Gamma(P_1, P_2, P_3)$ の中心 Q^* が得られる。

例題 3 \mathbb{R}^2 の 4 点 $P_1(1, 2), P_2(0, 0), P_3(2, 0), P_4(1, 3)$ を考える (図 6 参照)。これら 4 点は一般の位置にはない。これらに対して, その次元を上げた点 $\tilde{P}_1, \tilde{P}_2, \tilde{P}_3, \tilde{P}_4 \in \mathbb{R}^6$ を以下のように定義する。

$$\begin{aligned} P_1(1, 2) &\rightarrow \tilde{P}_1(1, 2, \varepsilon, 0, 0, 0) \\ P_2(0, 0) &\rightarrow \tilde{P}_2(0, 0, 0, \varepsilon, 0, 0) \\ P_3(2, 0) &\rightarrow \tilde{P}_3(2, 0, 0, 0, \varepsilon, 0) \\ P_4(1, 3) &\rightarrow \tilde{P}_4(1, 3, 0, 0, 0, \varepsilon) \end{aligned}$$

$\tilde{P}_1, \tilde{P}_2, \tilde{P}_3, \tilde{P}_4$ からの等距離点 \tilde{Q}^0 の $\tilde{P}_1, \tilde{P}_2, \tilde{P}_3, \tilde{P}_4$ に関する重心座標 $\tilde{\lambda}^0 = (\tilde{\lambda}_i^0)_{i=1, \dots, 4}$ は (111) より

$$\tilde{\lambda}^0 = \left(\frac{-42 + 7\varepsilon^2 + 2\varepsilon^4}{54\varepsilon^2 + 8\varepsilon^4}, \frac{7 + 15\varepsilon^2 + 2\varepsilon^4}{54\varepsilon^2 + 8\varepsilon^4}, \frac{7 + 15\varepsilon^2 + 2\varepsilon^4}{54\varepsilon^2 + 8\varepsilon^4}, \frac{28 + 17\varepsilon^2 + 2\varepsilon^4}{54\varepsilon^2 + 8\varepsilon^4} \right)$$

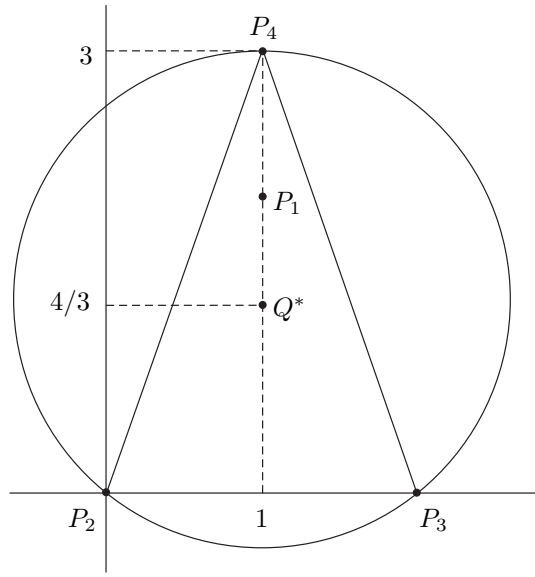


図 6: \mathbb{R}^2 上に与えられた 4 点の例

となる。 $|\varepsilon|$ が十分小さいとき $\tilde{\lambda}_1^0 < 0$ となるので、 \tilde{P}_1 を除外する。そして、射影した点 $\tilde{Q}^1 = \pi(\tilde{Q}^0 | L(\tilde{P}_2, \tilde{P}_3, \tilde{P}_4))$ の $\tilde{P}_1, \tilde{P}_2, \tilde{P}_3, \tilde{P}_4$ に関する重心座標を $\tilde{\lambda}^1 = (\tilde{\lambda}_i^1)_{i=1, \dots, 4}$ として (111) より

$$\tilde{\lambda}^1 = \left(0, \frac{5+\varepsilon^2}{18+3\varepsilon^2}, \frac{5+\varepsilon^2}{18+3\varepsilon^2}, \frac{8+\varepsilon^2}{18+3\varepsilon^2}\right) \quad (113)$$

を得る。 $\tilde{\lambda}_i^1 \geq 0, i = 1, \dots, 4$ なので、Kuhn-Tucker 条件 (16) より $\tilde{\lambda}^1 = \tilde{\lambda}^*$ である。ここで $\varepsilon = 0$ として、 $\tilde{\lambda}^*|_{\varepsilon=0} = (0, 5/18, 5/18, 8/18)$ となり、

$$\begin{aligned} Q^* &= \tilde{\lambda}^*|_{\varepsilon=0} \Phi \\ &= \left(0, \frac{5}{18}, \frac{5}{18}, \frac{8}{18}\right) \begin{pmatrix} 1 & 2 \\ 0 & 0 \\ 2 & 0 \\ 1 & 3 \end{pmatrix} \\ &= \left(1, \frac{4}{3}\right) \end{aligned}$$

であるから正しく最小包含円 $\Gamma(P_1, \dots, P_4)$ の中心 Q^* が得られる (図 6 参照)。

4.2 通信路行列の次元を上げる方法

4.1 節で \mathbb{R}^n の点 P_1, \dots, P_m に対して、それらの次元を上げた点 $\tilde{P}_1, \dots, \tilde{P}_m$ を定義して $\tilde{P}_1, \dots, \tilde{P}_m$ が一般の位置にあるようにした。本節では、通信路行列の行ベクトル P_1, \dots, P_m に対してそれらの次元を上げた行ベクトル $\tilde{P}_1, \dots, \tilde{P}_m$ を以下のように定義する。

\mathbb{R}^n の中心は原点 O だと考えられるから、原点 O の近傍の点 $\varepsilon \mathbf{e}_i$ を P_i に付加した。その類似で考えて、確率分布全体 Δ^n の中心は一様分布と考えると、一様分布の近傍をもとの P_i に付加する。確率分布は総和が 1 であるから、 ε を加算したら、どこかで ε を減算しなければ

ならない。そのように考えて、次のように $\tilde{P}_i \in \Delta^{n+2m}$ を定義する。

$$\tilde{P}_i = \left(\frac{P_{i1}}{2m+1}, \dots, \frac{P_{in}}{2m+1}, \overbrace{\frac{1}{2m+1}, \dots, \frac{1}{2m+1}}^{2i-2 \text{ 個}}, \frac{1+\varepsilon}{2m+1}, \frac{1-\varepsilon}{2m+1}, \overbrace{\frac{1}{2m+1}, \dots, \frac{1}{2m+1}}^{2m-2i \text{ 個}} \right), \quad (114)$$

ここで、 ε は $|\varepsilon|$ が十分小さい実数で $\varepsilon \neq 0$ とする。そして、

$$\tilde{\Phi} = \begin{pmatrix} \tilde{P}_1 \\ \vdots \\ \tilde{P}_m \end{pmatrix} \quad (115)$$

とおく。 $\varepsilon \neq 0$ なので、 $\text{rank } \tilde{\Phi} = m$ である。 $m \times 2m$ 行列 Ξ を

$$\Xi = \frac{1}{2m} \times \begin{pmatrix} 1+\varepsilon & 1-\varepsilon & 1 & 1 & \cdots & 1 & 1 \\ 1 & 1 & 1+\varepsilon & 1-\varepsilon & \cdots & 1 & 1 \\ \vdots & \vdots & & \ddots & \ddots & \vdots & \vdots \\ 1 & 1 & 1 & 1 & \cdots & 1+\varepsilon & 1-\varepsilon \end{pmatrix}$$

と定義し、 Ξ の第 i 行を $\Xi_i, i = 1, \dots, m$ とおくと $\Xi_i \in \Delta^{2m}$ であり、(114) は

$$\tilde{P}_i = \left(\frac{1}{2m+1} P_i, \frac{2m}{2m+1} \Xi_i \right), \quad i = 1, \dots, m \quad (116)$$

と書ける。

(114), (115), (116) で定義された通信路は次のようにモデル化される。すなわち、入力アルファベット $\{x_1, \dots, x_m\}$ は元のままとし、出力アルファベット $\{y_1, \dots, y_n\}$ に新たにダミーのアルファベット $\{y'_1, \dots, y'_{2m}\}$ を追加し、 $\{y_1, \dots, y_n\}, \{y'_1, \dots, y'_{2m}\}$ に値をとる確率変数を、それぞれ、 Y, Y' とする。そして、

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

$$P(Y' = y'_{j'} | X = x_i) = \Xi_{ij'}, \quad i = 1, \dots, m, j' = 1, \dots, 2m$$

と定義する。ここで、 $\Xi_{ij'}$ は行ベクトル Ξ_i の j' 番目の要素を表す。そこで、 Y と Y' を重み $(1/(2m+1), 2m/(2m+1))$ で混合した変数を \tilde{Y} とすると、 $X \rightarrow \tilde{Y}$ が (114), (115), (116) で定義された通信路のモデルである。 X と Y' との相互情報量 $I(\lambda, \Xi)$ はほぼ 0 であり、 X と \tilde{Y} との相互情報量 $I(\lambda, \Phi)$ は元の通信路容量 C に近い。ダミーのアルファベットは、通信路行列の行ベクトルの次元を上げてそれらが一般の位置にあるようにするために使われる。

さて, $Q = (Q_1, \dots, Q_n) \in \Delta^n$, $T = (T_1, \dots, T_{2m}) \in \Delta^{2m}$ に対して, $\tilde{Q} \in \Delta^{n+2m}$ を

$$\tilde{Q} = \left(\frac{1}{2m+1}Q, \frac{2m}{2m+1}T \right) \quad (117)$$

と定義する。 $\tilde{P}_1, \dots, \tilde{P}_m$ からの等距離点, すなわち, KL 情報量が等しい分布 \tilde{Q} を計算する。 \tilde{Q} は

$$D(\tilde{P}_i \| \tilde{Q}) = D(\tilde{P}_1 \| \tilde{Q}), \quad i = 2, \dots, m \quad (118)$$

を満たす。計算により

$$D(\tilde{P}_i \| \tilde{Q}) = \frac{-1}{2m+1} \left\{ \sum_{j=1}^n P_{ij} \log Q_j + \epsilon \log \frac{T_{2i-1}}{T_{2i}} + H(P_i) + \sum_{j'=1}^{2m} \log(2mT_{j'}) + \bar{H}(\epsilon) \right\} \quad (119)$$

となる。ここで, $H(P_i) = -\sum_{j=1}^n P_{ij} \log P_{ij}$, $\bar{H}(\epsilon) = -(1+\epsilon) \log(1+\epsilon) - (1-\epsilon) \log(1-\epsilon)$ とおいた。(118), (119) より

$$\begin{aligned} \sum_{j=1}^n (P_{ij} - P_{1j}) \log Q_j + \epsilon \log \frac{T_{2i-1}}{T_{2i}} - \epsilon \log \frac{T_1}{T_2} \\ = -H(P_i) + H(P_1), \quad i = 2, \dots, m \end{aligned} \quad (120)$$

となる。次に, \tilde{Q} の $\tilde{P}_1, \dots, \tilde{P}_m$ に関する重心座標を $\tilde{\lambda} = (\tilde{\lambda}_1, \dots, \tilde{\lambda}_m)$ とする。すなわち, (115) より $\tilde{Q} = \tilde{\lambda} \tilde{\Phi}$ である。(116), (117) より $Q = \tilde{\lambda} \Phi$, $T = \tilde{\lambda} \Xi$ となるので, これを (120) に代入して,

$$\begin{aligned} \sum_{j=1}^n (P_{ij} - P_{1j}) \log \sum_{i'=1}^m \tilde{\lambda}_{i'} P_{i'j} + \epsilon \log \frac{1 + \epsilon \tilde{\lambda}_i}{1 - \epsilon \tilde{\lambda}_i} \\ - \epsilon \log \frac{1 + \epsilon \tilde{\lambda}_1}{1 - \epsilon \tilde{\lambda}_1} = -H(P_i) + H(P_1), \quad i = 2, \dots, m \end{aligned} \quad (121)$$

を得る。 m 個の未知数 $\tilde{\lambda}_1, \dots, \tilde{\lambda}_m$ に対して $m-1$ 個の方程式 (121) と $\sum_{i=1}^m \tilde{\lambda}_i = 1$ によって解 $\tilde{\lambda}$ が得られる。解 $\tilde{\lambda}$ の存在と一意性は補題 24 で証明されている。

次の例題を考える。

例題 4 通信路行列

$$\Phi = \begin{pmatrix} P_1 \\ P_2 \\ P_3 \end{pmatrix} = \begin{pmatrix} 0.1 & 0.9 \\ 0.7 & 0.3 \\ 0.8 & 0.2 \end{pmatrix} \quad (122)$$

を考える (図 7 参照)。 Φ の次元を上げた通信路行列

$$\tilde{\Phi} = \begin{pmatrix} \tilde{P}_1 \\ \tilde{P}_2 \\ \tilde{P}_3 \end{pmatrix} \text{ が次のように与えられる。}$$

$$P_1(0.1, 0.9) \rightarrow \tilde{P}_1(0.1, 0.9, 1 + \epsilon, 1 - \epsilon, 1, 1, 1, 1)/7$$

$$P_2(0.7, 0.3) \rightarrow \tilde{P}_2(0.7, 0.3, 1, 1, 1 + \epsilon, 1 - \epsilon, 1, 1)/7$$

$$P_3(0.8, 0.2) \rightarrow \tilde{P}_3(0.8, 0.2, 1, 1, 1, 1, 1 + \epsilon, 1 - \epsilon)/7$$

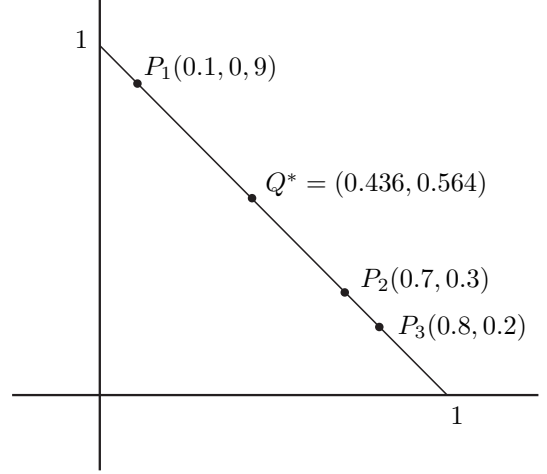


図 7: 入力シンボル数 $m = 3$, 出力シンボル数 $n = 2$ の例

$\epsilon = 0.05$ として方程式 (121) を解いて $\tilde{P}_1, \tilde{P}_2, \tilde{P}_3$ からの等距離点 \tilde{Q}^0 の $\tilde{P}_1, \tilde{P}_2, \tilde{P}_3$ に関する重心座標 $\tilde{\lambda}^0$ として

$$\tilde{\lambda}^0 = (2.39, -12.96, 11.57) \quad (123)$$

を得る。 $\tilde{\lambda}_2^0 = -12.96 < 0$ なので, \tilde{P}_2 を除外して, $\tilde{Q}^1 = \pi(\tilde{Q}^0 | L(\tilde{P}_1, \tilde{P}_3))$ として \tilde{Q}^1 の $\tilde{P}_1, \tilde{P}_2, \tilde{P}_3$ に関する重心座標を $\tilde{\lambda}^1$ とすると, $\tilde{\lambda}^* = \tilde{\lambda}^1 = (0.52, 0, 0.48)$ を得る。そして, $\tilde{\lambda}^*|_{\epsilon=0} = (0.52, 0, 0.48)$ となり,

$$\begin{aligned} Q^* &= \tilde{\lambda}^*|_{\epsilon=0} \Phi \\ &= (0.52, 0, 0.48) \begin{pmatrix} 0.1 & 0.9 \\ 0.7 & 0.3 \\ 0.8 & 0.2 \end{pmatrix} \\ &= (0.436, 0.564) \end{aligned}$$

を得る (図 7 参照)。通信路容量は $C = D(P_1 \| Q^*) = 0.398[\text{bit/symbol}]$ となる。

例題 5 通信路行列

$$\Phi = \begin{pmatrix} P_1 \\ P_2 \\ P_3 \\ P_4 \end{pmatrix} = \begin{pmatrix} 2/5 & 2/5 & 1/5 \\ 1/3 & 1/3 & 1/3 \\ 4/5 & 1/10 & 1/10 \\ 1/10 & 4/5 & 1/10 \end{pmatrix} \quad (124)$$

を考える (図 8 参照)。 Φ の次元を上げた通信路行列

$$\tilde{\Phi} = \begin{pmatrix} \tilde{P}_1 \\ \tilde{P}_2 \\ \tilde{P}_3 \\ \tilde{P}_4 \end{pmatrix} \text{が次のように与えられる。}$$

$$P_1(2/5, 2/5, 1/5)$$

$$\rightarrow \tilde{P}_1(2/5, 2/5, 1/5, 1+\varepsilon, 1-\varepsilon, 1, 1, 1, 1, 1)/9$$

$$P_2(1/3, 1/3, 1/3)$$

$$\rightarrow \tilde{P}_2(1/3, 1/3, 1/3, 1, 1, 1+\varepsilon, 1-\varepsilon, 1, 1, 1)/9$$

$$P_3(4/5, 1/10, 1/10)$$

$$\rightarrow \tilde{P}_3(4/5, 1/10, 1/10, 1, 1, 1, 1+\varepsilon, 1-\varepsilon, 1, 1)/9$$

$$P_4(1/10, 4/5, 1/10)$$

$$\rightarrow \tilde{P}_4(1/10, 4/5, 1/10, 1, 1, 1, 1, 1+\varepsilon, 1-\varepsilon)/9$$

$\varepsilon = 0.05$ として方程式 (121) を解いて $\tilde{P}_1, \tilde{P}_2, \tilde{P}_3, \tilde{P}_4$ からの等距離点 \tilde{Q}^0 の重心座標 $\tilde{\lambda}^0$ として

$$\tilde{\lambda}^0 = (-19.00, 7.98, 6.01, 6.01) \quad (125)$$

を得る。 $\tilde{\lambda}_1^0 = -19.00 < 0$ なので \tilde{P}_1 を除外して、射影 $\tilde{Q}^1 = \pi(\tilde{Q}^0 | L(\tilde{P}_2, \tilde{P}_3, \tilde{P}_4))$ を考える。 \tilde{Q}^1 の重心座標 $\tilde{\lambda}^1$ として

$$\tilde{\lambda}^1 = (0, -0.14, 0.57, 0.57) \quad (126)$$

を得る。 $\tilde{\lambda}_2^1 = -0.14 < 0$ なので、さらに \tilde{P}_2 を除外して $\tilde{Q}^2 = \pi(\tilde{Q}^1 | L(\tilde{P}_3, \tilde{P}_4))$ を考える。 \tilde{Q}^2 の重心座標 $\tilde{\lambda}^2$ として

$$\tilde{\lambda}^2 = (0, 0, 1/2, 1/2) \quad (127)$$

となり、 $\tilde{\lambda}^2 = \tilde{\lambda}^*$ を得る。したがって、

$$\begin{aligned} Q^* &= \tilde{\lambda}^*|_{\varepsilon=0} \Phi \\ &= (0, 0, 1/2, 1/2) \begin{pmatrix} 2/5 & 2/5 & 1/5 \\ 1/3 & 1/3 & 1/3 \\ 4/5 & 1/10 & 1/10 \\ 1/10 & 4/5 & 1/10 \end{pmatrix} \\ &= (9/20, 9/20, 1/19) \end{aligned}$$

を得る (図 8 参照)。通信路容量 C は $C = D(P_3 \| Q^*) = 0.540[\text{bit/symbol}]$ となる。

5 最小包含円を計算するヒューリスティックなアルゴリズム

第 2 章で与えた最小包含円の中心と半径を計算する射影アルゴリズムおよび第 4 章の次元を上げる方法を組み合わせ、任意の m, n および任意の点配置に対して適用できるヒューリスティックな射影アルゴリズムを提案する。第 2 章では、状況 1 と状況 2 という限定された

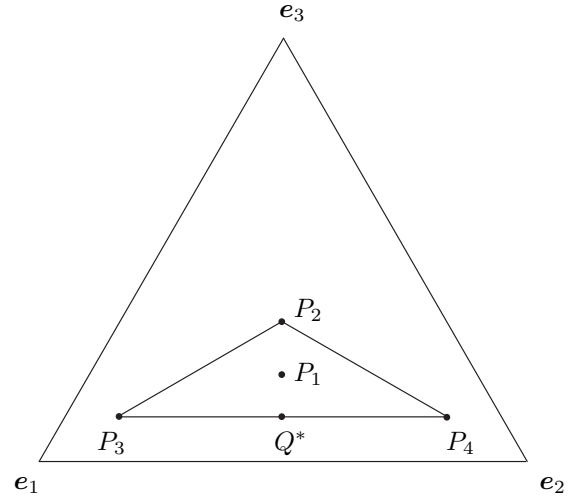


図 8: 入力シンボル数 $m = 4$, 出力シンボル数 $n = 3$ の例

場合に正しく最小包含円を計算するアルゴリズムを与えた。本章で提案する方法は、どのような点の個数や配置に対しても使えるが、得られた結果が必ずしも正しいとは限らない。以下において、ヒューリスティックアルゴリズムを説明し、具体的に乱数を用いて発生させた多数の最小包含円問題に対して得られた解が正しい解である割合を示す。

5.1 ヒューリスティックアルゴリズム HA

- (HA1) \mathbb{R}^n における m 個の点 P_1, \dots, P_m に対して、(109) によってそれらの次元を上げた点を $\tilde{P}_1, \dots, \tilde{P}_m$ とする。
- (HA2) $\tilde{P}_1, \dots, \tilde{P}_m$ からの等距離点 $\tilde{Q}^0 \in L(\tilde{P}_1, \dots, \tilde{P}_m)$ の重心座標 $\tilde{\lambda}^0 = (\tilde{\lambda}_i^0)_{i=1, \dots, m}$ を (111) によって計算する。
- (HA3) $\tilde{\lambda}_i^0 \geq 0, i = 1, \dots, m$ ならばアルゴリズムは終了し \tilde{Q}^0 を出力する。
- (HA4) $\tilde{\lambda}_i^0, i = 1, \dots, m$ に中に負のものがあれば、最小のもの、すなわち負で絶対値が最大のものに対応する点、例えば \tilde{P}_1 とする、を除外して $\tilde{P}_2, \dots, \tilde{P}_m$ を残す。
- (HA5) $m-1$ 個の点 $\tilde{P}_2, \dots, \tilde{P}_m$ に対して (HA1) から繰り返す。

上記のアルゴリズムに基づいてシミュレーションを実行した結果を次の表 1 に示す。シミュレーション方法は、 -1000 以上 1000 以下の整数を一樣ランダムに n 個発生させて \mathbb{R}^n の点の座標とする。それを m 回繰り返して、 \mathbb{R}^n の m 個の点を得る。得られた m 個の点を P_1, \dots, P_m として、ヒューリスティックアルゴリズム HA で最小包含円の中心 Q^* を計算する。以上の計算を 1 セットとして、表 1 に示す m, n のそれぞれの組に対して 10000 セット実行して、正しく Q^* が得られた割合を示す。表 1 全体では 20 万セットが実行されている。

5.2 計算量の比較

与えられた m 個の点 $P_1, \dots, P_m \in \mathbb{R}^n$ に対する最小包含円を計算する最も単純な方法は、すべての可能な点の組み合わせを調べる方法である [7]。この方法を総当たり法と呼ぶ。例えば、 \mathbb{R}^2 における 4 点の場合を考え

表 1: ヒューリスティックアルゴリズムの成功率

$m \backslash n$	2	3	10	20
3	100%	100%	100%	100%
4	98%	99%	100%	100%
5	96%	95%	100%	100%
8	94%	85%	99%	100%
10	93%	83%	98%	100%

る。 \mathbb{R}^2 において円は 2 点または 3 点で決定されるので、総当たり法におけるすべての可能な点の組み合わせの総数は ${}_4C_2 + {}_4C_3 = 10$ 通りである。同様に考えて、 \mathbb{R}^n における m 個の点の可能な組み合わせの総数 N_1 は、

$$N_1 = \sum_{\ell=0}^{\min(m,n+1)} m C_{\ell} \quad (128)$$

となる。1 つの組み合わせにおける計算量を 1 とすると、総当たり法の計算量は N_1 である。一方、提案のヒューリスティックアルゴリズムでは毎回必ず点が 1 つずつ減っていくので、計算量 N_2 は

$$N_2 = m - 2 \quad (129)$$

である。表 1 の場合の計算量の比 N_2/N_1 を表 2 に示す。

表 2: 総当たり法に対する提案法の計算量の比

$m \backslash n$	2	3	10	20
3	0.250	0.250	0.250	0.250
4	0.200	0.182	0.182	0.182
5	0.150	0.120	0.115	0.115
8	0.071	0.039	0.024	0.024
10	0.048	0.021	0.008	0.008

6 まとめ

ユークリッド幾何は我々になじみの幾何なので、アルゴリズムの開発とその証明がしやすい。それらは少々強引な方法でもできてしまうが、それはある意味でいいことである。そこで、それと同じ方法を情報幾何に移植しようとする、うまくいかない場合があった。そこで、それは強引で自然な方法ではなかったと気がつく。情報幾何に適用できるように考え直すと、逆にユークリッド幾何に還元できて、ユークリッド幾何で別の方法を発見することができた。つまり、ユークリッド幾何と情報幾何に共通の性質だけを使ったアルゴリズムが最も自然であり、実は最も簡単な方法である。これは非常に興味深い結果である。たぶん、初めから通信路容量の問題を情報幾何だけで考えていたのでは本論文の射影アルゴリズムは得られなかったと思う。ユークリッド幾何と情報幾何が有機的に結びついた結果だと思ふ。

7 今後の課題

1. 本論文では点の個数が $m = 2, 3, 4$ の場合にのみ最適点 Q^* を計算するアルゴリズムを得たが、これを任意の m に対して拡張する。
2. 本論文で利用した最小包含円問題と通信路容量問題の幾何学的な類似性を、レートひずみ関数の最適点を計算するアルゴリズムを開発する。
3. 通信路容量に関する有本アルゴリズムを、逆に最小包含円問題に移植して、最適点を探索するアルゴリズムを得る。

参考文献

- [1] S. Amari, *Differential-Geometrical Methods in Statistics*, Lecture Notes in Statistics, Springer-Verlag, 1985.
- [2] Arimoto, S., "An algorithm for computing the capacity of arbitrary discrete memoryless channels," IEEE Trans. Inform. Theory, vol.IT-18, pp.14-20, 1972.
- [3] Blahut, R.E., "Computation of channel capacity and rate-distortion functions," IEEE Trans. Inform. Theory, vol.IT-18, pp.460-473, 1972.
- [4] 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.
- [5] Csiszár, I. and Körner, J., *Information Theory: Coding Theorems for Discrete Memoryless Systems*, Academic Press, Orlando, 1982.
- [6] Gallager, G., *Information Theory and Reliable Communication*, Wiley, New York, 1968.
- [7] Muroga, S., "On the capacity of a discrete channel. I," J. Phys. Soc. Japan, vol.8, pp.484-494, 1953.
- [8] 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.
- [9] 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>
- [10] Strang, G., *Linear algebra and its applications*, Academic Press, New York, 1976.