Kendall の順位相関係数を固定した下での最小情報コピュラ

助田 一晟 *　　清智也 †

2022 年 11 月 18 日

概要

コピュラは複数の確率変数の従属性を記述するモデルであり, 応用上も盛んに利用されている. 最小情報コピュラ（最大エントロピーコピュラ）は所与の制約条件下で最も独立に近いコピュラである. 制約条件としては Spearman の順位相関係数をはじめとするはじめとして 1 次のものが主に考察されてきた. 一方で, 2 次の制約式を与えた状況下での同様のコピュラについてはその性質がほとんど知られていない. 本研究では, 従来の最小情報コピュラの亜種として, 最も代表的な 2 次の制約条件である Kendall の順位相関係数を固定した下での情報量最小化問題を離散近似手法を通して考察し, その最適解の性質について複数の特徴付けを行う.

1　はじめに

2 つの確率変数の従属性を測る代表的な指標として Spearman の順位相関係数と Kendall の順位相関係数がある. 一般に周辺分布が区間 $[0,1]$ 上の一様分布である多次元分布をコピュラと呼ぶ (詳細は Nelsen (2006) [6] が詳しい). 本研究では簡単のため 2 次元のコピュラを考えることとする. *1 Spearman や Kendall の順位相関係数はコピュラから一意的に定まる量である.

一方で, 同じ値の順位相関係数を持つコピュラは無数にあり, 順位相関係数を固定した状況下でどのようなコピュラ族が現れるかという方向を考えることもできる. 中でも, Kullback-Leibler 情報量の意味で最も一様コピュラ（独立コピュラ）に近いコピュラについての研究がなされており, 最初に Meeuwissen and Bedford (1997) [5] で最適化問題として取り上げられ，解の一意存在性が示された. そして, 一般的な制約条件に対し同様の問題が Bedford and Wilson (2014) [1] により整理され, 最小情報コピュラと呼ばれるようになった. また連続版の最小情報コピュラを扱いやすくするため離散化して考えることも多く（たとえば Piantadosi et al. (2012) [9]）, グリッドで離散化した各領域において区分的に一様な密度を持つコピュラは最小情報チェス盤コピュラと呼ばれる. これにより本質的には分割表の問題となり, 有限次元の最適化問題に帰着する.

情報量は分布のエントロピーの -1 倍であり, 情報量を最小化することはエントロピーを最大化することと等価である. このことから最小情報コピュラは, 最大エントロピーコピュラと呼ばれることもある. Spearman の順位相関係数を制約条件として課した下でのエントロピー最大化問題は, Piantadosi et al. (2007) [8] や Perrone et al. (2019) [7] で離散化をした上で扱われており, チェス盤コピュラ全体が成す Birkhoff 多面体に対する考察をはじめとした幾何学的なアプローチが取られ, 数値解も与えられている. Perrone et al. (2019) [7] ではウルトラモジュラという関数クラスと, それに対応する多面体へ拡張が行われている.

本研究では, 2 次の制約条件の一例として Kendall の順位相関係数を固定した下での最小情報コピュラの性質について考察する. 従来の最小情報コピュラに関する研究では 1 次の制約条件, すなわちある関数の期待値として表されるものについてのみ対象とされてきた. Meeuwissen and Bedford (1997) [5] や Piantadosi et al. (2007) [8] で扱われている Spearman の順位相関係数も 1 次の母数である. 一方で対比されることの多い Kendall の順位相関係数は 2 次の母数であり, これまでの最小情報コピュラの枠組み [1] では扱うことができない. このような 2 次の制約条件を課した設定下では従来の最小情報コピュラと異なり最適解の形が不明であるだけではなく, 最適化問題も非凸となり扱

* 東京大学大学院情報理工学系研究科数理情報学専攻
† 東京大学大学院情報理工学系研究科数理情報学専攻
*1 先行研究においても 2 次元で議論されていることが多く, そのまま多次元分布に拡張される議論も少なくない.

いが難しくなる．そこでアプローチを変更し, チェス盤コピュラ（分割表）上の成分移動操作を導入することにより停留条件を導出した．この操作はチェス盤コピュラ全体から成る空間の斜交基底を考えることに相当する．また, 従属性に関するその他の特徴付けも紹介する．

本論文の構成は以下の通りである．第 2 節では最小情報コピュラについて, 第 3 節ではその離散化版である最小チェス盤コピュラについて紹介する．第 4 節では本題である Kendall の順位相関係数を固定した条件下での最小情報コピュラを定式化し, 離散近似によって得られる分割表とその性質を詳しく議論する．このコピュラは密度関数が陽に表されていないが, 第 5 節では数値計算により得られた最適解の形状を紹介する．第 6 節で今後の展望を述べる．

2　最小情報コピュラ

一般に d 次元コピュラとは全ての周辺分布が $[0,1]$ 上の一様分布である d 次元の分布関数を指し, 任意の d 次元分布は Sklar の定理によりコピュラと周辺分布を用いて表現することができることから, コピュラは 2 変数間の従属性を完全に記述しているとされる．たとえばある 2 次元分布に対し確率密度 $p(x,y)$ が存在する場合, 確率密度 $p(x,y)$ がコピュラであることの条件は以下のように表される:

$$\int_0^1 p(x,y)dy = 1 \ (x \in [0,1]) \tag{2.1}$$

$$\int_0^1 p(x,y)dx = 1 \ (y \in [0,1]) \tag{2.2}$$

所与の制約条件下で, 独立コピュラ ($[0,1]^2$ 上の一様密度を持つコピュラ) に Kullback-Leibler 情報量の意味で最も近いコピュラを**最小情報コピュラ**と呼ぶ．このコピュラは以下の最適化問題の解として与えられるものである．

$$\text{minimize} \quad \int_0^1 \int_0^1 p(x,y) \log p(x,y) dxdy \tag{2.3}$$

$$\text{s.t.} \quad \int_0^1 p(x,y)dy = 1, \ \int_0^1 p(x,y)dx = 1 \tag{2.4}$$

$$\int_0^1 \int_0^1 h_k(x,y) p(x,y) dxdy = \mu_k \tag{2.5}$$

目的関数 (2.3) が独立コピュラとの Kullback-Leibler 情報量で測った時の近さ（分布の情報量）となっており, 分布の微分エントロピー（の -1 倍）とも言い換えられる．制約式 (2.4) は前述のコピュラであるための条件である．制約式 (2.5) は $h_k(x,y)$ という k 個の関数に対し, その期待値 μ_k を所与とする制約である．情報量最小化はエントロピー最大化と等価であることから, エントロピー最大化原理に従うという意味で最も自然な問題設定と言える．

この問題に対しては, 解の存在性と一意性, 解の形が知られている [5] [1].

Theorem 1 (Bedford and Wilson (2014), Theorem 2 および Theorem 3)**.** 最適化問題 (2.3) – (2.5) が内点許容解を持つならば，最小情報コピュラは存在し，一意的である．また解は次の形で表される:

$$p(x,y) = A(x)B(y) \exp\left(\sum_{k=1}^{K} \theta_k h_k(x,y) \right) \tag{2.6}$$

$A(x), B(y) (> 0)$ は可測関数.
$h_1, \ldots, h_K$ は所与の関数.
$\theta = (\theta_1, \ldots, \theta_K)$ は未知パラメータ.

なお, 式 (2.5) の $A(x)$ と $B(y)$ は, $p(x,y)$ がコピュラの条件を満たすための正規化関数で, その選び方には定数倍の任意性がある．

最小情報コピュラは指数分布族に似た分布族であり数学的に見通しが良い．また, 統計モデルとしては制約式に登場する関数 $h_k(x,y)$ の与え方によって, 柔軟なモデリングが可能であるといった長所がある．一方で, 一般には正規化関数 $A(x)$ や $B(y)$ は通常解析的に求まらないという扱いの難しさもある．この問題に対し, 後述する離散近似やスコアマッチング [2] 等のアプローチが研究されている．

3 最小情報チェス盤コピュラ

次に最小情報コピュラの離散近似を考える．これにより問題が扱いやすくなるだけではなく, 情報幾何学や行列スケーリングとの関係も示唆されている [10].

3.1 チェス盤コピュラ

定義域がグリッドにより離散化され, 各領域において密度関数が一様であるようなコピュラは**チェス盤コピュラ**と呼ばれる． チェス盤コピュラは一般のグリッドサイズで定義されるが, 簡単のため本発表では 2 次元で, グリッドのサイズが $n \times n$ の時のみを扱うこととする． すなわち各領域は $D_{ij} = ((i-1)/n, i/n) \times ((j-1)/n, j/n)$ $(1 \leq i \leq n, 1 \leq j \leq n)$ と表される.

チェス盤コピュラの密度関数を $p_\pi(x,y) = \frac{\pi_{ij}}{1/n^2}$ とおくと, コピュラであることの条件は

$$\sum_{j=1}^{n} \pi_{ij} = \frac{1}{n},\ \sum_{i=1}^{n} \pi_{ij} = \frac{1}{n}$$

となり, 周辺分布が離散一様な 2 元分割表 π_{ij} と同一視することができる．以下本発表では, チェス盤コピュラと 2 元分割表 π_{ij}, それに対応するサイズ $n \times n$ の行列を同一視する.

3.2 最小情報チェス盤コピュラ

最小情報コピュラの離散近似版として**最小情報チェス盤コピュラ**が考えられている．最小情報コピュラと同様に, エントロピー最大化原理に則った枠組みで定式化を行うことで, 最小情報チェス盤コピュラは以下の最適化問題の解と捉えることができる.

$$\text{minimize} \quad \sum_{i=1}^{n} \sum_{j=1}^{n} \pi_{ij} \log \pi_{ij} \tag{3.1}$$

$$\text{s.t.} \quad \sum_{j=1}^{n} \pi_{ij} = \frac{1}{n}, \sum_{i=1}^{n} \pi_{ij} = \frac{1}{n} \tag{3.2}$$

$$\sum_{i=1}^{n} \sum_{j=1}^{n} h_{k,ij} \pi_{ij} = \mu_k \tag{3.3}$$

ただし, $h_1, \ldots, h_K$ は所与の関数, $h_{k,ij}$ は領域 D_{ij} の中心での $h_k(x,y)$ の値, $\theta = (\theta_1, \ldots, \theta_K)$ は未知パラメータである．この最適化問題は有限次元凸最適化問題となっており, 扱いが比較的容易である．詳細は割愛するが, 特に Theorem 1. に対応して次の定理が成り立つ.

Theorem 2 (Piantadosi et al. (2012) [9], Proposition 2)**.** 最適化問題 (3.1) – (3.3) が 内点許容解（制約式を満たし全ての i, j に対して $\pi_{ij} > 0$ となるもの）を持つならば，最適解が存在し，一意的である. またそれは次の形で表される:

$$\pi_{ij} = A_i B_j \exp\left(\sum_{k=1}^{K} \theta_k h_{k,ij}\right) \tag{3.4}$$

ただし, $A_i > 0, B_j > 0, \theta_k \in \mathbb{R}$ である.

ただし, A_i, B_j は正規化定数であり, θ の関数として（定数倍を除いて）一意的に定まる.

4 2次の最小情報コピュラへの拡張

4.1 問題設定

従来の最小情報コピュラは所与の制約として 1 次の母数（ある関数の期待値の形で表されるもの）のみを扱っていた．たとえば, Spearman の順位相関係数

$$12\left(\int_0^1\int_0^1 xyp(x,y)dxdy\right)-\frac{1}{4}=\int_0^1\int_0^1 12(x-\frac{1}{2})(y-\frac{1}{2})p(x,y)dxdy \tag{4.1}$$

は 1 次の母数であり, Spearman の順位相関係数を固定した状況は従来の最小情報コピュラの枠組みで扱うことができる [8] [7].

一方, Spearman の順位相関係数と比較されることが多い Kendall の順位相関係数

$$\int_0^1\int_0^1\int_0^1\int_0^1 \mathrm{sgn}(x-\tilde{x})\mathrm{sgn}(y-\tilde{y})p(x,y)p(\tilde{x},\tilde{y})dxdyd\tilde{x}d\tilde{y} \tag{4.2}$$

は 2 次の母数であるため, Kendall の順位相関係数を固定した状況は従来の最小情報コピュラの枠組みで扱うことができない．そこで本研究では所与の制約式を 2 次の母数である Kendall の順位相関係数としたときの最適解（最小情報チェス盤コピュラ）を考察する．

従来の最小情報チェス盤コピュラと同様に, Kendall の順位相関係数を $\tau(\in[-1,1])$ に固定した下での最小情報チェス盤コピュラは以下のように最適化問題の解として定式化される．

$$(MP)\ \text{minimize}\quad \sum_{i=1}^{n}\sum_{j=1}^{n}\pi_{ij}\log\pi_{ij} \tag{4.3}$$

$$\text{s.t.}\quad \sum_{j=1}^{n}\pi_{ij}=\frac{1}{n},\ \sum_{i=1}^{n}\pi_{ij}=\frac{1}{n} \tag{4.4}$$

$$1-\mathrm{Tr}(\Xi\Pi\Xi\Pi^{\mathrm{T}})=\tau,\ \Pi=(\pi_{ij}) \tag{4.5}$$

ただし, $\Xi=\begin{pmatrix}1 & 0 & \dots & 0\\ 2 & 1 & \dots & 0\\ \dots & & & \\ 2 & & 2 & 1\end{pmatrix}$ であり, 制約式 (4.5) はチェス盤コピュラに対する Kendall の順位相関係数を表す（Durrlemann et al.(2000) [3] を参照）．Tr は行列のトレースを表す．

この問題 (MP) の最適解を本発表では, 誤解を恐れず **2 次の最小情報チェス盤コピュラ**と呼ぶこととする．最適解の存在性はコンパクト性から保証される．制約式が 1 次の母数であった場合と異なり, この問題は制約式が非凸となっていることに注意する．

4.2 成分移動操作の導入

以下, 2 次の最小情報チェス盤コピュラの特徴付けを行う．次節では, 最適化問題 (MP) の停留条件を考える．そのために本節ではチェス盤コピュラ上の各領域の確率質量（分割表モデルにおける各マス目の値）を移動させる操作を以下のように導入する．

まず, チェス盤コピュラ全体の空間を一様コピュラ $U=\frac{1}{n^2}I$ を原点とするベクトル空間 (の部分空間) とみなし, 以下のような $(n-1)^2$ 本から成る斜交基底 $(T_{ij})_{i=1,\dots,n-1,\ j=1,\dots,n-1}$ を考える:

$$T_{ij}:=\begin{pmatrix}O & O & O & O\\ O & +1 & -1 & O\\ O & -1 & +1 & O\\ O & O & O & O\end{pmatrix}. \tag{4.6}$$

すなわち, T_{ij} は i 行 j 列と $i+1$ 行 $j+1$ 列が $+1$, i 行 $j+1$ 列と $i+1$ 行 j 列が -1, それ以外が 0 となっている行列である.

任意のチェス盤コピュラ Π は以下のように一意な係数 a_{ij} を用いて表記できる.

$$\Pi = U + \sum_{i,j} a_{ij} T_{ij}, (i = 1, \dots, n-1, j = 1, \dots, n-1) \tag{4.7}$$

定義から, T_{ij} を加算するという操作は, チェス盤コピュラに対応する分割表中の 4 成分の間で成分を移動させることに相当している.

4.3 不変量の存在

次に, 本節では具体的に最適化問題 (MP) の停留条件を導出する. そのために, 前述の成分移動操作を行なった際の (1) 制約条件 (4.5) の左辺である Kendall の順位相関係数の変分と (2) 目的関数 (4.3) の変分を導出する. なお, コピュラの制約条件は任意の成分移動操作に対し常に満たされることが明らかなので考えない.

まず, 成分移動操作と Kendall の順位相関係数との間に以下の関係式が成立する.

Lemma 1. 任意のチェス盤コピュラ Π と任意の成分移動操作 T_{ij} と $T_{i'j'}$ に対し,

$$\Pi' := \Pi + \epsilon T_{ij} - r\epsilon T_{i'j'}, r := \frac{\pi_{i,j} + \pi_{i+1,j} + \pi_{i+1,j} + \pi_{i+1,j+1}}{\pi_{i',j'} + \pi_{i'+1,j'} + \pi_{i'+1,j'} + \pi_{i'+1,j'+1}} \tag{4.8}$$

と定める. このとき, Π と Π' の Kendall の順位相関係数は等しい:

$$1 - \mathrm{Tr}(\Xi\Pi\Xi\Pi^{\mathrm{T}}) = 1 - \mathrm{Tr}(\Xi\Pi'\Xi\Pi'^{\mathrm{T}}) \tag{4.9}$$

また逆も成立する.

Proof. $\Pi = (\pi_{i,j})$ はサイズ $n \times n$ の行列である. p を Π のすべての成分を添字の辞書付順序で並べた長さ n^2 のベクトルとする. Π' に対してベクトル p', T_{ij} に対し t_{ij} も同様に定義する.

$$\begin{aligned}
1 - \mathrm{Tr}(\Xi\Pi\Xi\Pi^{\mathrm{T}}) &= 1 - \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n}\sum_{l=1}^{n} \pi_{i,j}\pi_{k,l}\xi_{i,k}\xi_{l,j} \\
&= 1 - \sum_{i=k,j=l} \pi_{i,j}\pi_{k,l} - \sum_{i=k,j<l} 2\pi_{i,j}\pi_{k,l} - \sum_{i>k,j=l} 2\pi_{i,j}\pi_{k,l} - \sum_{i>k,j<l} 4\pi_{i,j}\pi_{k,l} \\
&= 1 - p^{\mathrm{T}} W p
\end{aligned}$$

ただし, W は以下のサイズ $n^2 \times n^2$ の行列し, I は成分が全て 1 の行列とする:

$$W = \begin{pmatrix} I & \Xi & \dots & \Xi \\ \Xi^{\mathrm{T}} & I & \dots & \Xi \\ \dots & & & \\ & & & I \end{pmatrix}.$$

従って, 同様の計算を用いると

$$\begin{aligned}
1 - \mathrm{Tr}(\Xi\Pi'\Xi\Pi'^{\mathrm{T}}) &= 1 - p'^{\mathrm{T}} W p' \\
&= 1 - (p + \epsilon t_{ij} - r\epsilon t_{i'j'})^{\mathrm{T}} W (p + \epsilon t_{ij} - r\epsilon t_{i'j'}) \\
&= 1 - p^{\mathrm{T}} W p - 2p^{\mathrm{T}} W (\epsilon t_{ij} - r\epsilon t_{i'j'}) - (\epsilon t_{ij} - r\epsilon t_{i'j'})^{\mathrm{T}} W (\epsilon t_{ij} - r\epsilon t_{i'j'}) \\
&= 1 - p^{\mathrm{T}} W p \\
&= 1 - \mathrm{Tr}(\Xi\Pi\Xi\Pi^{\mathrm{T}})
\end{aligned}$$

□

一方, 目的関数 $\sum\sum \pi_{ij} \log \pi_{ij}$ の変分を計算すると以下を得る.

Lemma 2. チェス盤コピュラ $\Pi = (\pi_{ij})$ に対し, 任意の成分移動操作 T_{ij} を微小に行なった時の情報量の変分は

$$\log \frac{\pi_{i,j}\pi_{i+1,j+1}}{\pi_{i+1,j}\pi_{i,j+1}}$$

Proof. $\frac{\partial}{\partial \pi_{ij}}(\sum_i \sum_j \pi_{ij}\log \pi_{ij}) = \log \pi_{ij} + 1$ なので, 変分は

$$(\log \pi_{ij} + 1) + (\log \pi_{i+1,j+1} + 1) - (\log \pi_{i+1,j} + 1) - (\log \pi_{i,j+1} + 1) = \log \frac{\pi_{i,j}\pi_{i+1,j+1}}{\pi_{i+1,j}\pi_{i,j+1}}.$$

□

以上の 2 つの補題を組み合わせることにより, 以下の主張を得る.

Theorem 3 (2 次の最小情報チェス盤コピュラの不変量)**.** 2 次の最小情報チェス盤コピュラでは任意の (i,j) の組み $(i,j = 1,\ldots,n-1)$ に対し, 以下の値が一定.

$$\frac{1}{\pi_{i,j} + \pi_{i+1,j} + \pi_{i+1,j} + \pi_{i+1,j+1}} \log \frac{\pi_{i,j}\pi_{i+1,j+1}}{\pi_{i+1,j}\pi_{i,j+1}}$$

Proof. Lemma 1. で考察した Kendall の順位相関係数を一定に保つ任意の成分移動操作に対して目的関数の変分が 0 であるという停留条件を考えると, 任意の $\epsilon, (i,j), (i',j')$ に対し,

$$\epsilon \log \frac{\pi_{i,j}\pi_{i+1,j+1}}{\pi_{i+1,j}\pi_{i,j+1}} - \frac{\pi_{i,j} + \pi_{i+1,j} + \pi_{i+1,j} + \pi_{i+1,j+1}}{\pi_{i',j'} + \pi_{i'+1,j'} + \pi_{i'+1,j'} + \pi_{i'+1,j'+1}} \epsilon \log \frac{\pi_{i',j'}\pi_{i'+1,j'+1}}{\pi_{i'+1,j'}\pi_{i',j'+1}} = 0.$$

移項により,

$$\frac{1}{\pi_{i,j} + \pi_{i+1,j} + \pi_{i+1,j} + \pi_{i+1,j+1}} \log \frac{\pi_{i,j}\pi_{i+1,j+1}}{\pi_{i+1,j}\pi_{i,j+1}} = \frac{1}{\pi_{i',j'} + \pi_{i'+1,j'} + \pi_{i'+1,j'} + \pi_{i'+1,j'+1}} \log \frac{\pi_{i',j'}\pi_{i'+1,j'+1}}{\pi_{i'+1,j'}\pi_{i',j'+1}}$$

を得るが, 左辺が (i,j) のみに, 右辺が (i',j') のみによる同じ形となっていることに注意する. 任意の $(i,j), (i',j')$ の組に対し上式が成立することから主張を得る. □

4.4 緩和問題の最適解との一致

次に (MP) の緩和問題を検討する.

$$\begin{aligned}
(RP)\ \text{minimize} \quad & \sum_{i=1}^{n}\sum_{j=1}^{n} \pi_{ij}\log \pi_{ij} \\
\text{s.t.} \quad & \sum_{j=1}^{n} \pi_{ij} = \frac{1}{n},\ \sum_{i=1}^{n} \pi_{ij} = \frac{1}{n} \\
& 1 - \mathrm{Tr}(\Xi\Pi\Xi\Pi^{\mathrm{T}}) \geq \tau
\end{aligned}$$

なお, 逆方向への緩和は一様コピュラが実行可能解に含まれるようになり自明なので考慮しない.

Definition 4 (相対的内点)**.** $\mathrm{aff}(S) \cap B(x,\epsilon) \subseteq S$ を満たす開球 $B(x,\epsilon)$ が存在する x 全体を相対的内部という. 相対的内部に含まれる点を相対的内点という.

Theorem 5. (RP) の最適解は相対的内点ではない.

Proof. $S = \{P \mid p_{ij} \geq 0, \sum_{j=1}^{n} p_{ij} = \frac{1}{n}, \sum_{i=1}^{n} p_{ij} = \frac{1}{n}\}$ とする. $f(P) = \sum_{i=1}^{n}\sum_{j=1}^{n} p_{ij}\log p_{ij}$ は S 上で凸関数であり, 停留点は $U = \frac{1}{n^2}I$ のみ.
大域最適解 p^* が相対的内点であるとする. p^* を中心とする開球（と実行可能領域の共通部分集合）を $B(p^*,\epsilon)$ とする. p^* は明らかに停留点ではないことから, $\exists p' \in B(p^*,\epsilon), s.t. f(p^*) > f(p')$. p' は実行可能領域に含まれるので p^* の最適性に矛盾する. よって p^* は相対的内点ではない. □

(MP) と (RP) の最適解は Theorem 5. より一致することから, (MP) の最適解の代わりに緩和問題 (RP) を解けば十分であると言える.

4.5　従属性

最後に最小情報チェス盤コピュラが有する従属性に関して, d-TP_2（density-TP_2）という正の従属性を有する関数クラスに属することを以下説明する.

Definition 6 (Total Positivity of Order Two, TP_2 [4])**.** $f : \mathbb{R}^2 \to \mathbb{R}$ が TP_2 に属する.
$:\Leftrightarrow f(x,y) \geq 0$ かつ $\begin{vmatrix} f(x,y) & f(x,y') \\ f(x',y) & f(x',y') \end{vmatrix} \geq 0 \ (x < x', y < y')$

特にコピュラに対しては, $C(x,y)C(x',y') - C(x,y')C(x',y) \geq 0$ を満たす時に TP_2, $c(x,y)c(x',y') - c(x,y')c(x',y) \geq 0$ を満たす時に d-TP_2 と言う. また, 不等号が strict に成立する場合, Strict Total Positivity (STP) という.

Theorem 7. 2 次の最小情報チェス盤コピュラは d-TP_2 に属する. また, d-STP_2 には属さない.

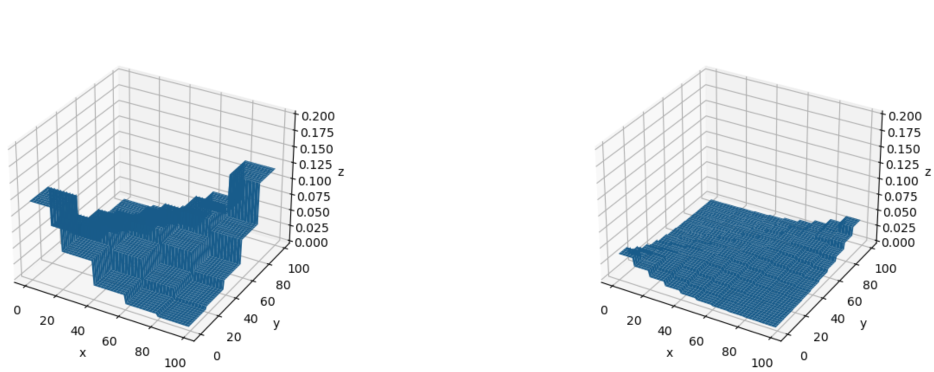
Proof. (RP) の最適解が TP_2 でないとする. このとき, Π のある 2×2 部分行列 $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ が存在して, $ad - bc < 0$. この 2×2 部分行列に対し 成分移動操作を行うことで, Kendall の順位相関係数は増大し, 目的関数である情報量は

$$ad - bc < 0 \Leftrightarrow \log \frac{ad}{bc} < 0$$

のため減少する. これは最適性に反する.
一方, 同じ領域 D_{ij} 内に含まれる 2 点 $(x,y), (x',y')$ に対しては明らかに $\begin{vmatrix} f(x,y) & f(x,y') \\ f(x'y) & f(x',y') \end{vmatrix} = 0$ であるので, d-STP_2 には属さない. □

5　数値実験結果

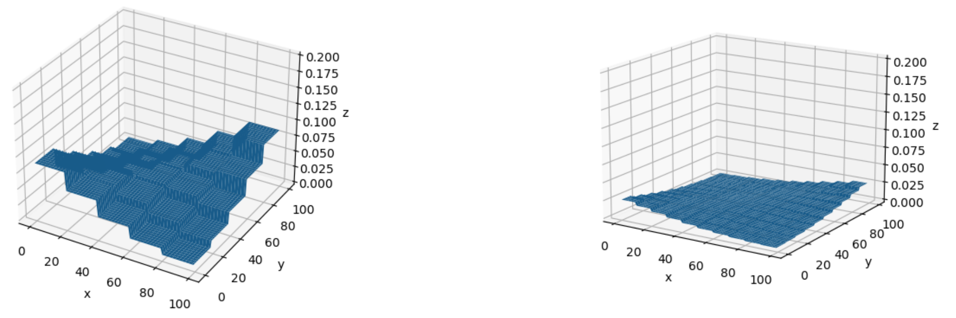
4.4 節で 2 次の最小情報チェス盤コピュラを求めるには, 緩和問題 (RP) を解けば十分であることを確認した. 以下は Python の科学計算ライブラリ scipy を用いて緩和問題 (RP) を解いた出力結果である. $\tau = 0.5$ とし, グリッドサイズは 5×5（図 1(a)）と 10×10（図 1(b)）とした. また比較対象として, Spearman の順位相関係数を $\rho = 0.5$ に固定した下での 1 次の最小情報チェス盤コピュラについても同様の結果をグリッドサイズ 5×5（図 2(a)）と 10×10（図 2(b)）に対し図示した.



(a) $n = 5, \tau = 0.5$　　**(b)** $n = 10, \tau = 0.5$

図 1: Kendall の τ を与えた時の最小情報チェス盤コピュラ

2 次の最小情報チェス盤コピュラには特徴的な階段状の密度関数を持つことや, 1 次の場合に比べて $(0,0)$ や $(1,1)$ 付近の裾の密度が重くなっている様子が観察される. また, 4 章で述べた主要な特徴付けに関する主張に対し, これら



(a) $n = 5, \rho = 0.5$　　**(b)** $n = 10, \rho = 0.5$

図 2: Spearman の ρ を与えた時の最小情報チェス盤コピュラ

の数値解が確かに満たしていることも確認される.

6　今後の方針

今後の展望としては, 幾何学的特徴の考察や, それに基づく一意性についての考察, 到達可能性（後述）に関する議論等が考えられる. チェス盤コピュラ全体は Birkhoff 多面体を成すことが知られている. すなわち, (MP) の実行可能領域は, Kendall の順位相関係数が一定という曲面と Birkhoff 多面体の共通部分である. そしてその面上で最も一様コピュラに Kullback-Leibler ダイバージェンスの意味で近いのが最適解（2 次の最小情報チェス盤コピュラ）である. しかしながらこれらの幾何学的な形状の特徴付けはまだできていない. これらの考察を通して最適解の一意性を保証できれば, τ をパラメタに持つ新たな分布族を提案でき, 従属性を捉えるための統計モデリングに活用できる可能性がある.

また. チェス盤コピュラにおける成分移動操作は分割表上のマルコフ基底と類似した概念であり, (4.7) における任意の係数 a_{ij} が非負であるという条件を到達可能性として定義することができる. 最適解は一様コピュラからこの意味で到達可能だと予想しており, 引き続き検証が必要である.

2 次の最小情報コピュラとして最も基本的な Kendall の順位相関係数を固定した条件下の解に対しこのような性質が明らかになれば, より一般の制約式への拡張が今後の課題となる.

参考文献

[1] Bedford, T. and Wilson, K. J., On the construction of minimum information bivariate copula families, Ann. Inst. Stat. Math., 66, 703–723, 2014.

[2] Chen, Y. and Sei, T., A proper scoring rule for minimum information copulas, arXiv, 2022.

[3] Durrleman, V., Nikeghbali, A. and Roncalli, T., Copulas Approximation and New Families, Available at SSRN, 2000.

[4] Karlin, S., Total Positivity, Vol. I, Stanford U.P., Calif., 1968.

[5] Meeuwissen, A. M. H. and Bedford, T., Minimally informative distributions with given rank correlation for use in uncertainty analysis, J. Statist. Comput. Simul., 57, 143–174, 1997.

[6] Nelsen, R. B., An Introduction to Copulas, 2nd ed., Springer, 2006.

[7] Perrone, E., Solus, L., and Uhler, C., Geometry of discrete copulas. J. Multivar. Anal. 172, 162–179, 2019.

[8] Piantadosi, J., Howlett, P., and Boland, J., Matching the grade correlation coefficient using a copula with maximum disorder. Journal of Industrial and Management Optimization, 3 (2), 305–312, 2007.

[9] Piantadosi, J., Howlett, P., and Boland, J., Copulas with maximum entropy, Optim. Lett., 6, 99–125, 2012.

[10] 清 智也, 最小情報コピュラとその周辺, 日本統計学会誌, 2021-2022, 51 巻, 1 号, p.75–99, 2021.