数理解析研究所講究録
第1894巻 2014年 1-9

フーリエ変換による置換族の独立性解析

長岡技術科学大学　鈴木　孝*　NGUYEN THAI PHAT†　武井　由智‡

Takashi Suzuki, NGUYEN THAI PHAT, Yoshinori Takei
Nagaoka University of Technology

1 はじめに

最小値独立性は文書類似性判定アルゴリズムなどに応用をもつ置換族の確率的対称性の一種である. また, 置換の確率分布の対称性度合いを表す尺度としてラベル依存度がフーリエ解析の応用により提案されている.

本論文は, 最小値独立置換族, その強化であるRobust置換族, あるいは k 対独立性置換族のもつ確率的対称性をフーリエ解析の観点より特徴づけることを試みる. まずラベル依存度の基本性質である逆元操作, 共役操作に対する不変性を示す. 次に4次の最小値独立置換族でサイズが12であるものを全数調査し, そのラベル依存度の分布を特定する. 特にラベル依存度が0となる置換族が4次交代群およびその補集合として得られることを示す. また, 同一サイズの最小値独立性を有さない置換族と比較して, 最小値独立置換族は平均的に小さいラベル依存度を持つことを示す. その他, k 対独立性 (k-wise independence) およびRobust性がそれぞれYoung直交表現に基づくフーリエ係数の特定 "周波数"における値が0になることとして特徴づけられることを示す (定理5, 6).

2 準備

ここでは, 対称群の表現と, 対称群上の関数のフーリエ変換についての既知の事柄を述べる [2, 3, 7, 8, 9].

2.1 置換と対称群

n 個の対象の集合を, $[n] := \{1, 2, \ldots, n-1, n\}$ とすると, n 個の対象の置換の集合は,

$$S_n := \{\sigma : [n] \to [n] \mid \text{全単射}\}$$

で定義される. S_n は, 写像の結合を演算とした群であり, n 次対称群と呼ばれる.

S_n のサイズは, $|S_n| = n!$ である. $\sigma(j) = i_j$ を満たすような S_n の要素 σ は, two line notation により

$$\sigma = \begin{pmatrix} 1 & 2 & \ldots & n \\ i_1 & i_2 & \ldots & i_n \end{pmatrix} \tag{1}$$

のように表記される. 置換を構成する巡回置換の長さの組を巡回置換型 (cycle-type), あるいは単に型 (type) と呼ぶ.

後のため, $\sigma \in S_n$ に対する別の表記法を導入する.

$$\sigma^{-1} = \begin{pmatrix} 1 & 2 & \ldots & n \\ a_1 & a_2 & \ldots & a_n \end{pmatrix}$$

であるとする. すなわち, $\sigma(a_1) = 1, \sigma(a_2) = 2, \ldots, \sigma(a_n) = n$ が満たされるとする. このような σ を

$$\sigma = \langle a_1, a_2, \ldots, a_n \rangle$$

と書く. この記法を σ の逆像文字列表記と呼ぶ.

2.2 置換族, 対称群上の関数

対称群 S_n の部分集合を置換族と呼ぶ. 置換族 $F \subseteq S_n$ が与えられると, その特性関数

$$F(\pi) = \begin{cases} 1 & (\pi \in F) \\ 0 & (\pi \notin F) \end{cases} \tag{2}$$

*長岡技術科学大学大学院工学研究科電気電子情報工学専攻, tsuzuki@act-w.nagaokaut.ac.jp
†第1著者に同じ, dhphat@act-w.nagaokaut.ac.jp
‡長岡技術科学大学電気系, takei@act-w.nagaokaut.ac.jp

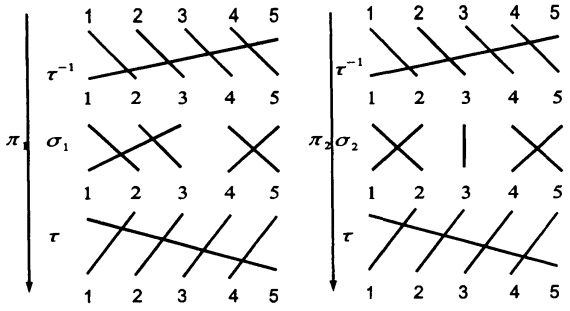


図 1: 1,2,3,4,5 を 5,1,2,3,4 とラベルを貼り替えた時の例

あるいは F からの一様ランダムな抽出に対する分布関数

$$f(\pi) = \begin{cases} \frac{1}{|F|} & (\pi \in F) \\ 0 & (\pi \notin F) \end{cases} \tag{3}$$

と同一視することが出来る. このことから, 置換族は後述のフーリエ解析の対象になり得る.

2.3 共役とラベルの貼り替え

対象 1,2,3,4,5 にそれぞれ 5,1,2,3,4 というラベルを貼る. すると置換の集合

$$\{\sigma_1, \sigma_2\} = \left\{ \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 3 & 1 & 5 & 4 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 1 & 3 & 5 & 4 \end{pmatrix} \right\}$$

は

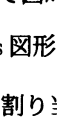
$$\{\pi_1, \pi_2\} = \left\{ \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 5 & 4 & 3 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 2 & 4 & 3 & 1 \end{pmatrix} \right\}$$

のように見える. その様子を図 1 に示す. ここで, $i = 1, 2$ に対し,

$$\pi_i = \tau\sigma_i\tau^{-1}, \tau = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 1 & 2 & 3 & 4 \end{pmatrix}.$$

$\pi, \sigma \in S_n$ に対し, $^\exists\tau, \pi = \tau\sigma\tau^{-1}$ のとき π と σ は共役と呼び, $^\forall\tau, \sigma \in S_n, f(\sigma) = f(\tau\sigma\tau^{-1})$ が成り立つ関数を類関数と呼ぶ. また, ある元 σ と互いに共役な元を全て集めた集合を, σ の共役類と呼ぶ.

2.4 tablaux と tabloid, standard tablaux

$\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_k)$, $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_k$, $\sum_{i=1}^{k} \lambda_i = n$ を n の分割と呼び $\lambda \vdash n$ と書く. たとえば $n = 4$ のとき $(2,1,1) =$  , $(2,2) =$  , $\cdots$ となる. このように, 整数の分割を $\lambda_1, \lambda_2, \ldots, \lambda_k$ 個のタイルの並びで図的に表現したものを Ferrers 図形と呼ぶ. Ferrers 図形に対して $[n]$ の元を  のように過不足なく割り当てたものを tablaux t と呼ぶ.

$\lambda \vdash n$ に対する 2 つの tablaux t, t' に対し, t と t' の対応する各行は, 集合としては $[n]$ の同一の部分集合であるとする. このことを $t \sim t'$ と書いたとすれば, これは $\lambda \vdash n$ に対する tablaux 間の同値関係となる. その同値類を $\{t\} = \{t'\}$ と書いて tabloid であるという. ある tabloid を代表する tablaux として, 各行が昇順にソートされた (右に行くにつれて要素が増加する) tablaux を取ることができる. また, ある tabloid を図示するとき, 列方向の並び順は重要でないことを示すために列の要素間の区切りを省略する. 例えば, $t =$  に対して

$$\{t\} = \left\{ \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 \\ \cline{1-1} \end{array}, \begin{array}{|c|c|}\hline 2 & 1 \\ \hline 3 \\ \cline{1-1} \end{array} \right\} = \begin{array}{|c|c|}\hline 1 & 2 \\ \hline \underline{3} \\ \cline{1-1} \end{array}$$

のようになる. また,

$$\begin{array}{|c|c|c|}\hline 1 & 2 & 3 \\ \hline \underline{4} \\ \cline{1-1} \end{array} = \begin{array}{|c|c|c|}\hline 2 & 3 & 1 \\ \hline \underline{4} \\ \cline{1-1} \end{array}$$

は $\lambda = (3,1) \vdash 4$ に対する同一の tabloid の 2 通りの表示である. さらに, 各数が一つの箱に必ず一回きり現れ, 各行と列について数が昇順に並んでいるものを標準ヤング盤又は standard tablaux と呼ぶ. これは, tablaux t と $\sigma \in S_n$ に対して tablaux $\sigma(t)$ が σ を t 内の各数に作用させることで自然に得られる. tabloids $\{t\}$ に対する $\sigma(\{t\})$ についても同様である.

2.5 群の表現

一般に群 G の表現は G から d_ρ 次複素正則行列への準同型写像すなわち, $\rho : G \to \mathbb{C}^{d_\rho \times d_\rho}$ であって,

$$\rho(xy) = \rho(x)\rho(y), \quad ^\forall x, y \in G \tag{4}$$

$$\rho(x^{-1}) = (\rho(x))^{-1}, \quad ^\forall x \in G \tag{5}$$

を満たすものを言う. とくに $\rho(e) = I$ が成立する. d_ρ は表現の次数と呼ばれる. 以下では $\rho(x)$ がユニタリ行列であるような表現のみを考える. 表現 ρ に対して, $d_\rho \times d_\rho$ 行列 Q と, 表現 $\rho_i = G \to \mathbb{C}^{d_{\rho_i} \times d_{\rho_i}} (i = 1, 2)$

が存在して $^\forall g \in G, \rho(g) = Q \left[\begin{array}{c|c} \rho_1(g) & 0 \\ \hline 0 & \rho_2(g) \end{array}\right] Q^{-1}$ となるとき，$\rho \cong \rho_1 \oplus \rho_2$ と書き，ρ は ρ_1 と ρ_2 の直和に分解されると言う．ρ がより次数の低い表現の直和に分解出来る時 ρ は可約表現と呼ばれ，逆に分解できない時 ρ は既約表現と呼ばれる．一般に，表現は一つ以上の既約表現の直和に分解する．

2.6 Young 直交表現 [2]

対称群の場合，群の既約表現として Young 直交表現 (YOR) が用いられる．YOR は整数の分割 $\lambda \vdash n$ によりラベル付けされる．λ に対する標準ヤング盤の数を d_λ とするとき YOR は既約表現 $\rho_\lambda : S_n \to \mathbb{R}^{d_\lambda \times d_\lambda}$ と書かれる．行列 $[\rho_\lambda(k,k+1)]^{d_\lambda \times d_\lambda}$ の要素は標準ヤング盤 t でインデックスされ，隣接互換 $(k,k+1)$ に対し，

$$[\rho_\lambda((k,k+1))]_{t,t} = 1/d_t(k,k+1). \tag{6}$$

また，$(k,k+1)(t)$ も標準ヤング盤の場合，(6) に加えて非対角要素

$$[\rho_\lambda((k,k+1))]_{t,(k,k+1)(t)} = \sqrt{1 - 1/d_t(k,k+1)^2} \tag{7}$$

があらわれる．他のすべての行列要素はゼロである．ここで $d_t(i,j) = c(j) - c(i)$，$c(x)$ は x の content と呼ばれる量であり，ヤング盤の中に置かれた x の列インデックスと行インデックスの差で定義される．

すべての置換は隣接互換の積で表されることから，式 (6),(7) によって $^\forall \sigma \in S_n$ の表現が定義できる．これにより定義された $\rho_\lambda(\sigma)$ は，

$$\rho_\lambda(\sigma)^{-1} = \rho_\lambda(\sigma)^\top \tag{8}$$

を満たす．すなわち，$\rho_\lambda(\sigma)$ は直交行列である．

また，ここでこの ρ_λ とは別の自然な表現を考える．各 $\lambda \vdash n$ に対して，λ 上の tabloids を基底とする線形空間上の表現 τ_λ が

$$[\tau_\lambda(\sigma)]_{ij} = \begin{cases} 1 & \sigma(\{t_j\}) = \{t_i\} \\ 0 & \text{otherwise} \end{cases}. \tag{9}$$

で定義される．この行列は，λ に対する各 tabloids $\{t_i\}$ で添字付けられており，$\sigma(\{t_j\}) = \{t_i\}$ とは $\{t_j\}$ の各行を構成する集合を σ で写した像が $\{t_i\}$ の対応行を構成する集合に一致することを表わす．

2 種類の表現 ρ_λ と τ_λ は次のように結びつけられる．

定義 1. $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_l)$，$\mu = (\mu_1, \mu_2, \ldots, \mu_m), \lambda, \mu \vdash n$ に対し，

$$\mu_1 + \mu_2 + \ldots + \mu_i \geq \lambda_1 + \lambda_2 + \ldots + \lambda_i, \ ^\forall i \geq 1.$$

ただし，$i > m$ に対する，μ_i は 0 と考える．$i > l$ に対する λ_i も同様である．が成り立つ時，$\mu \trianglerighteq \lambda$ と書き，μ は λ を支配すると言う． □

定理 1. *[9]* 各 $\lambda \vdash n$ に対する Young 直交表現 ρ_λ と *tabloids* 上の表現 τ_λ に対し，正則行列 C_λ が存在して，

$$C_\lambda^{-1} \tau_\lambda(\sigma) C_\lambda = \bigoplus_{\mu \trianglerighteq \lambda} \bigoplus_{l=1}^{K_{\lambda,\mu}} \rho_\mu(\sigma) \quad ^\forall \sigma \in S_n) \tag{10}$$

となる．$K_{\lambda,\mu}$ は *Kostka number* と呼ばれる． □

すなわち，τ_λ は $\mu \trianglerighteq \lambda$ なる μ に関する Young 直交表現の (重複を含む) 直和に分解する．

2.7 対称群上のフーリエ解析

$f, g : S_n \to \mathbb{R}$ に対し，これらの内積を

$$\langle f | g \rangle = (n!)^{-1} \sum_{\sigma \in S_n} f(\sigma) g(\sigma) \tag{11}$$

で定義する．

ρ_λ の (i,j) 成分と $\rho_{\lambda'}$ の (k,l) 成分は $\lambda = \lambda'$ かつ $(i,j) = (k,l)$ の時を除き直交する．また，$\lambda = \lambda'$ かつ $(i,j) = (k,l)$ の時は $1/d_\lambda$ となる．すなわち，

$$\langle \rho_{\lambda,(i,j)} | \rho_{\lambda',(k,l)} \rangle = \delta_{\lambda\lambda'} \delta_{ik} \delta_{jl} \frac{1}{d_\lambda} \tag{12}$$

が成立する．S_n 上の関数 $f : S_n \to \mathbb{R}$ の Young 直交表現に基づくフーリエ変換は

$$\hat{f}_{\rho_\lambda} = \sum_{\sigma \in S_n} f(\sigma) \rho_\lambda(\sigma), \quad \lambda \vdash n \tag{13}$$

で定義される．さらに，

$$\chi_\lambda(\sigma) = \mathrm{trace}(\rho_\lambda(\sigma)) \tag{14}$$

を λ における指標と定義する．以下が成立する:

$$(n!)^{-1} \mathrm{trace}(\hat{f}(\lambda)) = \langle f | \mathrm{trace}(\rho_\lambda) \rangle, \tag{15}$$

$$\forall \tau, \sigma \ \chi_\lambda(\tau\sigma\tau^{-1}) = \chi_\lambda(\sigma). \tag{16}$$

つまり指標は，共役 = ラベルの貼り替え で不変である．さらに，$\{\chi_\lambda\}_{\lambda \vdash n}$ は $f : S_n \to \mathbb{R}$ で共役不変な関数 (類関数) の空間の正規直交基底になっている．

また, 式 (10) より, τ_λ を Young 直交表現 ρ_λ の立場で用いて, 式 (13) と同様の Fourier 変換を

$$\hat{f}_{\tau_\lambda} = \sum_{\sigma \in S_n} f(\sigma)\tau_\lambda(\sigma) \tag{17}$$

で定義する. これは, $\tau_\lambda(\sigma)$ と同様に tabloids で添字付けられた行列であり, $\{t\}, \{t'\}$ 成分は 式 (9) から

$$[\hat{f}_{\tau_\lambda}]_{\{t\},\{t'\}} = \Pr_{\sigma \in_f S_n}[\sigma(\{t'\}) = \{t\}] \tag{18}$$

($\sigma \in_f S_n$ は分布 $\Pr[\sigma] = f(\sigma)$ で σ をランダムに抽出することを表す) であることに注意する.

例 1. (Diaconis [2, Chap. 3D], Huang et al. [9].) 2-wise(Pair-wise) の場合, 分割 $\lambda = (n-2,1,1)$ に対する tabloid は互いに異なる $x_1, x_2 \in [n]$ に対する $\{t'\} = \begin{array}{|c|}\hline [n]\setminus\{x_1,x_2\} \\ \hline \underline{x_1} \\ \underline{x_2} \\ \hline \end{array}$ の形となる. 同様に, $\{t\} = \begin{array}{|c|}\hline [n]\setminus\{y_1,y_2\} \\ \hline \underline{y_1} \\ \underline{y_2} \\ \hline \end{array}$ とおくと, 式 (18) から,

$$\begin{aligned}[\hat{f}_{\tau_\lambda}]_{\{t\},\{t'\}} &= \Pr_{\sigma}[\sigma(\{t'\}) = \{t\}] \\ &= \Pr_{\sigma \in_f S_n}[\bigwedge_{i=1}^{2} \sigma(x_i) = y_i]\end{aligned}$$

となり, 順序対 $(x_1,\ x_2)$ が $(y_1,\ y_2)$ に写像される確率を表す.

2.8 畳み込み

関数 $f, g : \mathbb{S}_n \to \mathbb{R}$ はそれぞれランダム置換 σ, τ の分布とする. つまり $f(\sigma) = \Pr[\sigma], g(\tau) = \Pr[\tau]$ とする. 2 つのランダム置換 σ, τ が独立であるとき, 2 つの合成 $\pi = \sigma\tau$(τ が先, σ が後に作用する) がいかなる確率分布に従うかを考える. $h(\pi) = \Pr[\pi]$ とおくとき, $[\![A]\!]$ で事象 A のインジケータ変数 (A が真のとき 1, 偽のとき 0) で表せば

$$\begin{aligned}h(\pi) &= \sum_{\sigma,\tau \in \mathbb{S}_n} f(\sigma)g(\tau)[\![\pi = \sigma\tau]\!] \\ &= \sum_{\tau \in \mathbb{S}_n} f(\pi\tau^{-1})g(\tau) \\ &= f * g(\pi). \end{aligned} \tag{19}$$

すなわち, h は畳み込み $f * g$ となる. 一般に $f * g$ について, そのフーリエ変換は簡明な性質をもつ, つまり, 畳み込み定理

$$\widehat{f * g} = \hat{f} \cdot \hat{g} \tag{20}$$

が成立する.

2.9 最小値独立置換族 (MWIPF)[1]

集合間の類似度を復元できるハッシュ値の構成法に関連して Broder ら [1] により $F \subseteq S_n$ が最小値独立置換族 (Min-Wise Independent Parmutation Family, MWIPF) であることが, 次の通り定義されている.

定義 2. $F \subseteq S_n$ が最小値独立であることは次が成立することをいう.

$$\forall X \neq \emptyset \subseteq [n], x \in X,\ \Pr_{\pi \in F}[\min \pi(X) = \pi(x)] = |X|^{-1}. \tag{21}$$

□

これについては次の特徴付けが与えられている [1, 5, 6].

定理 2. $F \subseteq S_n$ が最小値独立であることと, 次は同値である;

$$\begin{aligned}&{}^\forall k\ (0 \le k < n), X \in \binom{[n]}{k}, x \in [n]\setminus X \\ &\left[\Pr_{\pi \in F}[\pi(X) = \{1,2,\ldots,k\} \wedge \pi(x) = k+1]\right. \\ &\qquad \left. = \frac{1}{(n-k)\binom{n}{k}}\right]. \end{aligned} \tag{22}$$

□

この特徴づけを利用した最小値独立置換族の一般的構成法が [5] で与えられ, それにより特に $n = 4$ のときサイズが $\mathrm{lcm}(4,3,2,1) = 12$ を達成する最小値独立置換族は全部で 64 個存在することがわかる.

2.10 Robust Permutation Family

最小値独立性より更に強い置換族の確率的対称性として, 次の Robustness が定義されている [6].

定義 3. $F \subseteq S_n$ が *Robust* であるとは, 次が成立することを言う;

$$\forall \sigma \in S_n, X \subseteq [n]\ (X \neq \emptyset),\ x \in X$$
$$\Pr_{\pi \in F}[\min \sigma \circ \pi(X) = \sigma \circ \pi(x)] = \frac{1}{|X|}. \quad \square \quad (23)$$

すなわち, 任意の $[n]$ の全順序に関して F が最小値独立であるとき, F は Robust であると呼ばれる. この Robustness についても, 定理 2 と類似の特徴付けが与えられている;

定理 3. [6] $F \subseteq S_n$ が Robust であることと, 次は同値である;

$$\forall k\ (0 \leq k < n), X, Y \in \binom{[n]}{k},$$
$$x \in [n] \setminus X, y \in [n] \setminus Y$$
$$\left[\Pr_{\pi \in F}[\pi(X) = Y \wedge \pi(x) = y] = \frac{1}{(n-k)\binom{n}{k}}\right]. \quad (24)$$

□

2.11 k-wise Independent Permutation Family

最小値独立性や Robustness の他, よく知られた置換族の確率的対称性として k 対独立性がある.

定義 4. S_n 上の確率分布 $f : S_n \to \mathbb{R}$ が *k-wise independent* であるとは $\forall (x_1, x_2, ..., x_k)\ (x_i \in [n], x_i \neq x_j\ for\ i \neq j) \forall (y_1, y_2, ..., y_k)\ (y_i \in [n], y_i \neq y_j\ for\ i \neq j)$

$$\Pr_{\sigma \in_f S_n}[\bigwedge_{i=1}^{k} \sigma(x_i) = y_i] = \frac{(n-k)!}{n!} \quad (25)$$

を満すことを言う. ただし, $\sigma \in_f$ は, 確率分布 f に従って $\sigma \in S_n$ がランダムに抽出されることを表す. □

2.12 ラベル依存度 [4]

文献 [4] で, $f : S_n \to \mathbb{R}$ に対し, f の共役不変成分

$$(P(f))(\sigma) = \sum_{\lambda \vdash n} \langle f | \chi_\lambda \rangle \chi_\lambda(\sigma) \quad (26)$$

および, 第 2.3 節のラベルの貼り替えに対する関数 f の安定性の尺度としてラベル依存度 LD が,

$$\mathsf{LD}(f) := \frac{\langle f - P(f) | f - P(f) \rangle}{\langle f | f \rangle} = 1 - \frac{\langle P(f) | P(f) \rangle}{\langle f | f \rangle} \quad (27)$$

として定義されている. ラベル依存度は $0 \leq \mathsf{LD}(f) \leq 1$ の範囲で値をとる. 特に f がラベルの貼り替えで値が不変であることと, $\mathsf{LD}(f) = 0$ は同値である.

3 ラベル依存度を変化させない関数操作

関数 $f : S_n \to \mathbb{R}$ に対し, $\overline{f}$, f^τ を

$$\overline{f}(\sigma) = f(\sigma^{-1}), \quad (28)$$
$$f^\tau(\sigma) = f(\tau^{-1}\sigma\tau)\ (\tau \in S_n) \quad (29)$$

として定義する. すると, これらの操作 $(f \mapsto \overline{f}, f \mapsto f^\tau)$ はラベル依存度を変えないことが示せる.

命題 4.

$$\mathsf{LD}(f) = \mathsf{LD}(\overline{f}), \quad (30)$$
$$\mathsf{LD}(f^\tau) = \mathsf{LD}(f). \quad (31)$$

□

4 小さい $n(n = 4, 5)$ における最小サイズの最小値独立置換族のラベル依存度調査

4.1 4 次の最小サイズの最小値独立置換族のラベル依存度分布

S_4(4 次対称群) の部分集合で, 最小サイズ $|F| = \mathrm{lcm}(4, 3, 2, 1) = 12$ の最小値独立置換族は全 $2^6 = 64$ 通り存在する. この全 64 通りのラベル依存度抽出を行う. また, その結果から, 最小値独立置換族が持つ性質について考察する.

表 1: $n = 4$ の最小値独立置換族のラベル依存度分布

ラベル依存度	0.444	0.389	0.375	0.264
置換族の個数	12	12	24	8
ラベル依存度	0.222	0.167	0	
置換族の個数	2	4	2	

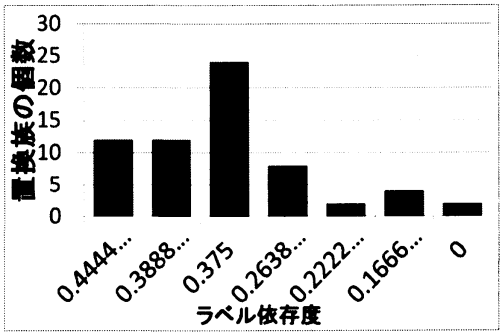


図 2: $n = 4$ の最小値独立置換族のラベル依存度分布

4.1.1 結果

$n = 4$ のとき, 最もサイズの小さい $\mathrm{lcm}(4, 3, 2, 1) = 12$ の最小値独立置換族は文献 [5] により全 64 通り存在することが知られている. その 64 通りそれぞれに対してラベル依存度を計算した. その結果を表 1 に示す.

特に $\mathrm{LD}(F) = 0$ である F が 2 個存在するが, その 1 つは偶置換の集合 A_4, 他方は $S_4 \backslash A_4$ であった. A_4 は S_4 の正規部分群で, $\forall \tau \in S_4, \tau A_4 \tau^{-1} = A_4$ が成立するため, $\mathrm{LD}(A_4) = 0$ であり, このことにより $\mathrm{LD}(S_4 \backslash A_4) = 0$ も出る.

また [6] により, $n \geq 4$ に対する交代群 A_n は最小値独立置換族であることが証明されている.

4.2 4 次の重複元の無いサイズ 12 の全置換族とのラベル依存度の比較

S_4 において, 重複元のないサイズ 12 の置換族全 ${}_{24}\mathrm{C}_{12} = 2704156$ 個の平均と分散, 最小値独立置換族全 64 個の平均と分散を表 (3) に示す. この結果より, ラベル依存度の観点からは, 最小値独立置換族は同サイズの重複元を持たない置換族よりも高い対称性を持っていることがわかる.

また, 重複元のないサイズ 12 の置換族全 ${}_{24}\mathrm{C}_{12} = 2704156$ 個の分布を図 3 と図 3 の $\mathrm{LD} = 0.28125$ 以下

表 2: $n = 4$ の最小値独立置換族と重複元のないサイズ 12 の置換族のラベル依存度平均と分散比較

	ラベル依存度の平均値	ラベル依存度の分散
S_4 のサイズ 12 の最小値独立置換族	0.347	0.00955
重複元のないサイズ 12 の置換族	0.413	0.00236

の分布を拡大した図を右上に示す. この図を見て分かる通り, ラベル依存度が小さい置換族は, 非常に少ない特別な存在であることが分かる.

サイズ 12 の最小値独立置換族の LD の中で最頻である $\mathrm{LD} = 0.375$ 以下の LD を持つ置換族の比率は, サイズ 12 の最小値独立置換族の場合, $40/64 = 0.625$, で, サイズ 12 の一般の置換族の場合は, $546616/{}_{24}\mathrm{C}_{12} = 0.202$ となった. この事実も, 最小値独立置換族のラベル依存度は同じサイズの置換族の中で小さい側に偏っているといえる.

4.3 5 次の最小値独立置換族のラベル依存度調査

$n = 5$ に対するサイズが最小値 ($|F| = \mathrm{lcm}(5, 4, 3, 2, 1) = 60$) の最小値独立置換族について調査した. このような F はおよそ 7^{20} 個存在すると見積もられ, 全数探索は難しいため, $6 \cdot 10^4$ 個の F をランダムに抽出して $\mathrm{LD}(F)$ を計算した.

その結果の平均は, 0.50, 最大は 0.62, 分散は $6.5 \cdot 10^{-4}$ であった. なお, 標本にはかからなかったものの, 5 次の最小サイズの最小値独立置換族の $\mathrm{LD}(F)$ の最小値は, $F = A_5$ とその補集合が $\mathrm{LD}(F) = 0$ を達成する. 同じサイズの必ずしも最小値独立でない置換族の 10^3 個のランダム標本 LD の平均は 0.64, 分散は $1.2 \cdot 10^{-3}$ であるから, 最小値独立置換族はその対称性により, 小さい LD を持つと考えられる.

なお, $n = 5$ に対して, サイズが $\mathrm{lcm}(5, 4, 3, 2, 1) = 60$ であり, $\mathrm{LD}(F) = 0$ を満たす最小値独立置換族は, A_5, A_5^c だけであることがわかる. また, $n = 6$ に対する最小値独立置換族の最も小さいサイズは, $\mathrm{lcm}(6, 5, 4, \ldots, 1) = 60$ だが, このサイズの置換族 $F \subseteq S_6$ で重複元を含まないものは, $\mathrm{LD}(F) = 0$ を達成できないことが, S_6 の共役類のサイズの考察からわかる.

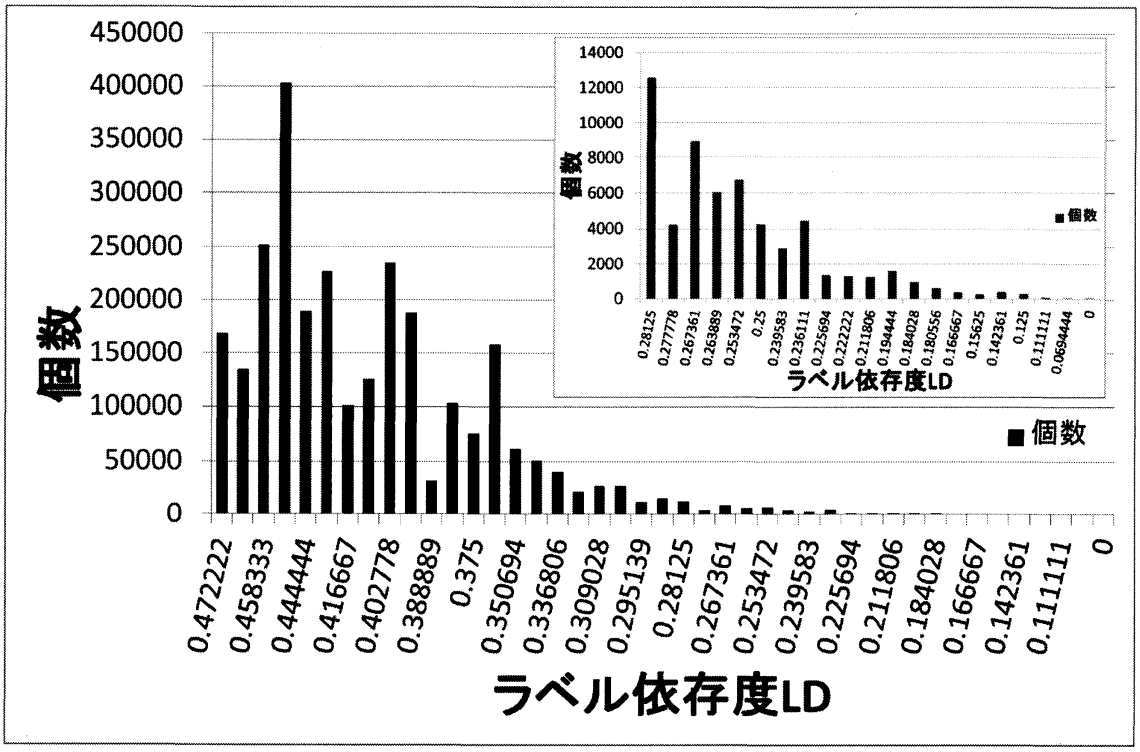


図 3: $n = 4$ の最小値独立置換族と重複元のないサイズ 12 の置換族のラベル依存度分布 (右上の図は LD = 0.28125 以下の拡大)

表 3: $n = 5$ の最小値独立置換族とランダムにサンプリングした重複元のないサイズ 60 の置換族のラベル依存度平均と分散比較

	ラベル依存度の平均値	ラベル依存度の分散
S_5 のサイズ 60 の最小値独立置換族	0.50	$6.5 \cdot 10^{-4}$
重複元のないサイズ 60 の置換族	0.64	$1.2 \cdot 10^{-3}$

5 確率的独立性のフーリエ係数による特徴付け

k-wise 独立性および Robust 独立性の定義 3, 4 を例 1 にあるような tabloid 相互の写り合い確率でとらえることで, Young 直交表現に基づくフーリエ解析の言葉で特徴づけられることを示す.

定理 5. $f : S_n \to \mathbb{R}$ を分布関数とするとき, f が *k-wise independent* であることと, 各 $\mu \trianglerighteq (n-k, 1, \ldots, 1)$ (1 は k 個), $\mu \neq (n)$ に対する *Young* 直交表現に基づくフーリエ係数行列 $\hat{f}_{\rho_\mu}$ が 0 であることは同値である.

定理 6. $F \subseteq S_n$, に対し, $f : S_n \to \mathbb{R}$ を F からの一様ランダム抽出を表す分布関数とする. F が *Robust Permutation Family* であることと, 各 $\mu \trianglerighteq \left(\lceil \frac{n-1}{2} \rceil, \lfloor \frac{n-1}{2} \rfloor, 1\right)$, $\mu \neq (n)$ に対する *Young* 直交表現に基づくフーリエ係数行列 $\hat{f}_{\rho_\mu}$ が 0 であることは同値である.

これらの定理の証明はほぼ同じようにでき, ここでは定理 6 に対する証明を示す.

5.1 Robust Family と Young 直交表現

以下, Robustness が $\lambda = (n-k-1, k, 1)$ に対する $\hat{f}_{\tau_\lambda}$ の言葉で自然に特徴付けられることを示す. さらに, Young 直交表現 ρ_λ $(\mu \trianglerighteq \lambda)$ に基づくフーリエ係数 $\hat{f}_{\rho_\mu}$ は (自明な表現 $\mu = (n)$ に対するものを除き) 消えることを示す.

証明. 定理 2 より, $F \subseteq S_n$ が Robust になることは,

$1 \le k < n$ に対して,

$$\forall X,\ Y \subseteq \binom{[n]}{k}, \quad x \in [n] \setminus X, \quad y \in [n] \setminus Y$$

$$\Pr_{\pi \in F}[\pi(X) = Y \wedge \pi(x) = y] = \frac{1}{(n-k)\binom{n}{k}} \quad (32)$$

となることであるが, X と $[n]\setminus X \setminus \{x\}$, Y と $[n]\setminus Y \setminus \{y\}$ の立場を交換することにより, $1 \le k < \lfloor \frac{n-1}{2} \rfloor$ に対して式 (32) が成立することが必要十分であることに注意する. このとき, $\lambda = (n-k-1, k, 1)$ に対する 2 つの tabloids $\{t_{X,x}\}, \{t_{Y,y}\}$ を

$$\{t_{X,x}\} = \begin{array}{|c|}\hline [n]\setminus X \setminus \{x\} \\ \hline X \\ \hline x \\ \hline \end{array}, \quad \{t_{Y,y}\} = \begin{array}{|c|}\hline [n]\setminus Y \setminus \{y\} \\ \hline Y \\ \hline y \\ \hline \end{array}$$

とおけば, 各 $\sigma \in S_n$ に対して

$$[\tau_\lambda(\sigma)]_{\{t_{Y,y}\},\{t_{X,x}\}} = \begin{cases} 1 & (\sigma(X) = Y \wedge \sigma(x) = y) \\ 0 & (\text{otherwise}) \end{cases} \quad (33)$$

が成立する. 置換 π を $F \subseteq S_n$ から一様分布に従って抽出する確率分布を

$$f(\pi) = \begin{cases} \frac{1}{|F|} & (\pi \in F) \\ 0 & (\pi \notin F) \end{cases}$$

とすれば,

$$\left[\hat{f}_{\tau_\lambda}\right]_{\{t_{Y,y}\},\{t_{X,x}\}} = \sum_{\sigma \in S_n} f(\sigma)\,[\tau_\lambda(\sigma)]_{\{t_{Y,y}\},\{t_{X,x}\}} = \Pr_{\pi \in F}[\pi(X) = Y \wedge \pi(x) = y] \quad (34)$$

が成立する. 従って, $F \subseteq S_n$ が Robust であれば, 各 $\lambda = (n-k-1, k, 1)$ に対して $\hat{f}_{\tau_\lambda}$ の全成分は一つの値 $\frac{1}{(n-k)\binom{n}{k}}$ となる.

次に $\hat{f}_{\rho_\lambda}$ を計算する. 定理 1 の

$$\tau_\lambda = C^{-1}\left[\bigoplus_{\mu \trianglerighteq \lambda} \bigoplus_{l=1}^{K_{\lambda\mu}} \rho_\mu\right] C$$

より,

$$\hat{f}_{\tau_\lambda} = C^{-1}\left[\bigoplus_{\mu \trianglerighteq \lambda} \bigoplus_{l=1}^{K_{\lambda\mu}} \hat{f}_{\rho_\mu}\right] C$$

が成立する. ここで, $\hat{f}' = \left[\bigoplus_{\substack{\mu \trianglerighteq \lambda \\ \mu \neq (n)}} \bigoplus_{l=1}^{K_{\lambda\mu}} \hat{f}_{\rho_\mu}\right]$ と置くと,

$$\begin{aligned} \hat{f}_{\tau_\lambda} &= \sum_{\sigma \in S_n} f(\sigma)\tau_\lambda(\sigma) \\ &= C^{-1}\left(\sum_{\sigma \in S_n} f(\sigma) \left[\begin{array}{c|c} \rho_{(n)}(\sigma) & 0 \\ \hline 0 & \bigoplus_{\substack{\mu \trianglerighteq \lambda \\ \mu \neq (n)}} \bigoplus_{l=1}^{K_{\lambda\mu}} \rho_\mu(\sigma) \end{array}\right]\right) C \\ &= C^{-1}\left(\left[\begin{array}{c|c} \hat{f}_{\rho_{(n)}} & 0 \\ \hline 0 & \hat{f}' \end{array}\right]\right) C. \end{aligned} \quad (35)$$

ここで f は確率分布であるから

$$\hat{f}_{\rho_{(n)}} = \sum_{\sigma \in S_n} f(\sigma) \cdot 1 = 1 \quad (36)$$

である. $\hat{f}_{\tau_\lambda}$ の全成分が $\frac{1}{(n-k)\binom{n}{k}}$ であるとき, その Rank は 1 となるため,

$$C\hat{f}_{\tau_\lambda}C^{-1} = \left(\left[\begin{array}{c|c} 1 & 0 \\ \hline 0 & \hat{f}' \end{array}\right]\right)$$

の Rank も 1 である. そのためには, $\hat{f}' = 0$ となる他はない. よって, 各 $k\left(1 \le k \le \lfloor \frac{n-1}{2} \rfloor\right)$ と $\mu \trianglerighteq (n-k-1, k, 1)$, $\mu \neq (n)$ に対して, $\hat{f}_{\rho_\mu}$ はゼロ行列である. $1 \le k' \le k \le \lfloor \frac{n-1}{2} \rfloor$ に対して $(n-k'-1, k', 1) \trianglerighteq (n-k-1, k, 1)$ であるから,

$$\begin{aligned} &\left(\exists k \left(1 \le k \le \lfloor \frac{n-1}{2} \rfloor\right) \mu \trianglerighteq (n-k-1, k, 1)\right) \\ &\quad \Leftrightarrow \mu \trianglerighteq \left(n - \lfloor \frac{n-1}{2} \rfloor - 1, \lfloor \frac{n-1}{2} \rfloor, 1\right) \\ &\quad = \left(\lceil \frac{n-1}{2} \rceil, \lfloor \frac{n-1}{2} \rfloor, 1\right) \end{aligned} \quad (37)$$

であるので,

$$F \subseteq S_n \ :\ \text{Robust} \Rightarrow \hat{f}_{\rho_\mu} = 0$$
$$\text{for} \quad \mu \trianglerighteq \left(n - \lfloor \frac{n-1}{2} \rfloor - 1, \lfloor \frac{n-1}{2} \rfloor, 1\right),\ \mu \neq (n) \quad (38)$$

と簡単に表される. そして逆に $\Rightarrow$ の右辺が成立すれば, 各 $\lambda = (n-k-1, k, 1)$, $\left(1 \le k \le \lfloor \frac{n-1}{2} \rfloor\right)$ に対して $\hat{f}_{\tau_\lambda}$ の全成分が $\frac{1}{(n-k)\binom{n}{k}}$ となり, f は Robust Family からの一様ランダム抽出を表す. □

5.2 最小値独立置換族と Young 直交表現

$F \subseteq S_n$ が最小値独立であることを Young 直交表現で特徴付けることは, 前節ほど単純ではない. ただ

し, F が最小値独立であるとき, 定理 2 の式 (22) を各 $X \in \binom{[n]\backslash\{x\}}{k}$ について足し合わせることで, F からの一様ランダム抽出は 1-wise independent であることが従うため, F が最小値独立であれば, 定理 5 によって, 対応する分布関数 f は

$$\hat{f}_{\rho_\mu} = 0 \quad \text{for} \quad \mu = (n-1, 1) \tag{39}$$

を満たすことが分かる.

6 まとめ

本稿では, 最小値独立性, Robust 性, k 対独立性といった置換族の確率的独立性を Young 直交表現に基づくフーリエ変換の言葉で解析した.

まず, ラベル依存度の基本性質 (逆元操作不変性, ラベル貼り替え不変性) を示した後, 最小サイズ $\mathrm{lcm}(n, n-1, \ldots, 1)$ を持つ最小値独立置換族を $n = 4, 5$ の場合について調査した. 4 次の最小値独立置換族でサイズが $\mathrm{lcm}(4, 3, 2, 1) = 12$ であるものを全数調査し, そのラベル依存度の分布を特定し, 特にラベル依存度が 0 となる置換族が 4 次交代群およびその補集合として得られることを示した. また, 同一サイズの最小値独立性を有さない置換族と比較して, 最小値独立置換族は平均的に小さいラベル依存度を持つことを示した. さらに, 4 次の最小値独立置換族に対し, ある種の重ね合わせ操作を行った場合ラベル依存度が減少することを全数調査により示した. 5 次の最小サイズ ($\mathrm{lcm}(5, 4, 3, 2, 1) = 60$) の $\mathsf{LD} = 0$ となる最小値独立置換族は A_5 とその補集合 A_5^c が満たすことを示した.

次に, Robust 性, k 対独立性といった置換族の確率的独立性を Young 直交表現に基づくフーリエ係数の一部がゼロ行列となること (即ち, "帯域制限されていること") として特徴づけた.

7 謝辞

本稿について議論頂いた 和田州平氏に感謝いたします.

参考文献

[1] A. Z. Broder., M. Charikar., A. M. Frieze., and M. Mitzenmacher, "Min-Wise Independent Permutations, " *STOC 1998*, pp. 327–336

[2] P. Diaconis, *Group Representations in Probability and Statistics,* Inst. Math. Stat. (1988)

[3] Kondor, R., Shervashidze, N. and Borgwardt, K. M., "The graphlet spectrum, " *Proc. 26th ICML*, pp. 529–536 (2009)

[4] 小川 浩明, 武井 由智, "対称群上の未知関数のラベル依存度の抽出, " 平成 24 年度 電子情報通信学会信越支部大会 講演論文集, 2C-2, Oct. 2012.

[5] Y. Takei and T. Itoh, "A General Construction of Min-Wise Independent Permutations, " *Trans. Fund. IEICE,* **E83-A(4)**, pp. 646–655 (2000)

[6] Andrei Z. Broder and Michael Mitzenmacher, "Completeness and robustness properties of min-wise independent permutations, " *Random Structures and Algorithms,* **Volume 18 Issue 1**, pp.18–30 (January 2001)

[7] Gordon James and Adalbert Kerber, *The Representation Theory of the Symmetric Group, ENCYCLOPEDIA OF MATHEMATICS AND ITS APPLICATIONS,* **Volume 16** , CAMBRIDGE UNIVERSITY PRESS (1985)

[8] Gordon James *Representations of General Linear Groups,*, CAMBRIDGE UNIVERSITY PRESS (1984)

[9] Jonathan Huang, Carlos Guestrin and Leonidas Guibas "Fourier Theoretic Probabilistic Inference over Permutations, " *Journal of Machine Learning Research*, **10**, pp. 997–1070 (2009)