2005 年情報論的学習理論ワークショップ
2005 Workshop on Information-Based Induction Sciences (IBIS2005)
Tokyo, Japan, November 9 - 11, 2005.

相関行列の群論的一般化について

On Group-Theoretical Generalization of Correlation Matrices

井手剛*
Tsuyoshi Idé

Abstract: We outline a new mathematical framework for generalizing the traditional covariance to include nonlinear correlations. Our key idea is to represent pairwise correlation patterns as the superposition of the irreducible representations of a point group. We explicitly give the definition of generalized correlation matrices and demonstrate their utility using an analytically solvable model.

Keywords: nonlinear correlations, group theory, irreducible representations

1 はじめに

相関解析は多変量解析における中心的な題材のひとつである。主成分分析や正準相関解析、比較的最近ではカーネル主成分分析など、自由度同士の絡み合いを記述する手法は非常に活発に研究されている。

因果関係の分析などの文脈で、主にカテゴリカルデータを対象にして、グラフィカルモデルと呼ばれる一連のモデリング手法がほぼ確立されている [7]。しかし、実数値データに関しては実用的な相関解析の手法は、事実上多変量正規分布を陰に陽にモデルとして仮定したものがほとんどである。

正規分布は、分散共分散行列、あるいはそれを規格化して（偏）相関係数行列を変数同士の絡み合いの尺度として採用する。しかし、よく知られているように、共分散が変数の相関をうまく記述できるのは、線形相関に近い場合である。逆に言えば、変数 x と y の間で共分散がゼロであっても、一般にはそれは相関がないことを意味しない。教科書的な例としては、点 (x, y) が、xy 平面において円形に分布する場合がある。この場合、明らかに二つの変数は強く相関しているが、相関係数はゼロとなってしまう。

非線形相関を扱えないという相関係数の弱点を補うために、既存の相関解析の手法を「カーネル化」する試みが数多くなされている。しかし残念なことに、実用的な場面において、それらの試みを因果分析や障害検知に適用するのは簡単ではない。最大の理由は、「どのような非線形相関が取り込まれたのか」が、カーネルトリックの陰に隠れてよくわからないという点にある。

まとめると、本論文の目的は次の通りである。

- 非線形相関を取り込めるように、また、実世界のデータのノイズに頑強であるように、伝統的な相関係数概念を拡張すること。
- ブラックボックス的ではないやり方で非線形パターンを取り込む理論的枠組みを与えること。

本論文では、まず第 2 節において、多変量系の変数同士の絡み合いを表す自然な量として 2 体交差キュムラントを用いることを提案する。次に、第 3 節において、最低次の 2 体交差キュムラントとしての伝統的な共分散が、C_{4v} 群の B_2 既約表現として理解できることを示す。この事実を手がかりに、第 4 節において、一般化相関係数を、C_{4v} の既約表現となるような、高次の 2 体交差キュムラントの 1 次結合として定義する。最後に、第 5 節において、解析的に解ける模型を使って、新たに定義された一般化相関係数の計算例を示す。

われわれの知る限り本研究は、群の既約表現を相関パターンと結びつけた初めての試みである[1]。画像認識の分野では、リー群に基づいて不変特徴抽出論が展開される [9]。そのカーネル拡張も最近議論されている [4]。しかし、点群の既約表現それ自体を明示的にパターンとして把握した研究は殆どないように思われる。機械学習の分野でも群論は強力な道具になりえると思われるが、明示的に群の性質を利用した研究は、Smola-Kondor [5] などの他はまだ少ないようである。

*IBM 東京基礎研究所, 242-8502 神奈川県大和市下鶴間 1623-14, e-mail goodidea@jp.ibm.com,
IBM Research, Tokyo Research Lab., 1623-14 Shimo-Tsuruma, Yamato, Kanagawa 242-8502, Japan

[1]本論文と密接に関連した研究を [1] においても発表予定である。

2　確率密度関数の近似

2.1　キュムラント母関数

系が n 次元の確率ベクトル $\boldsymbol{x} = (x_1, ..., x_n)^T \in \mathcal{D}_0$ で記述されるとし、その確率密度関数を $p(\boldsymbol{x})$ と表す。$\mathcal{D}_0$ は $p(\boldsymbol{x})$ の定義域である。系の内部構造に関する情報はすべてこの p の中に含まれていると考えられる。p の近似的表式を考える出発点として、キュムラント母関数 $\Psi(\boldsymbol{s})$ を考える:

$$\Psi(\boldsymbol{s}) \equiv \ln \int_{\mathcal{D}_0} d\boldsymbol{x}\; p(\boldsymbol{x}) \exp(\boldsymbol{s}^{\mathrm{T}}\boldsymbol{x}) = \ln \left\langle \exp(\boldsymbol{s}^{\mathrm{T}}\boldsymbol{x}) \right\rangle \quad (1)$$

ただし、$\langle\cdot\rangle$ は $p(\boldsymbol{x})$ に関する期待値 $\langle\cdot\rangle = \int_{\mathcal{D}_0} d\boldsymbol{x} p(\boldsymbol{x})\cdot$ を表す。多変数キュムラントは、$\boldsymbol{s}$ についてのテーラー展開の展開係数として定義される:

$$\begin{aligned} \Psi(\boldsymbol{s}) &= \sum_j s_j \langle x_j \rangle_{\mathrm{c}} + \ldots \qquad (2) \\ &+ \frac{1}{k!} \sum_{j_1,\cdots,j_k} s_{j_1} \cdots s_{j_k} \langle x_{j_1} \cdots x_{j_k} \rangle_{\mathrm{c}} + \ldots \end{aligned}$$

ここで、多変数キュムラントを表すために、**キュムラント平均** $\langle\cdot\rangle_{\mathrm{c}}$ の表記を用いた [3]。例えば、s_i および $s_i s_j$ の係数をそれぞれ $\langle x_i \rangle_{\mathrm{c}}$ および $\langle x_i x_j \rangle_{\mathrm{c}}$ とおくと、$\langle x_i \rangle_{\mathrm{c}} = \langle x_i \rangle$ および $\langle x_i x_j \rangle_{\mathrm{c}} = \langle x_i x_j \rangle - \langle x_i \rangle \langle x_j \rangle$ を得る。前者は平均、後者は共分散に他ならない[2]。

多変数キュムラントに関しては次の二つの性質がよく知られている [10]。

定理 1 正規分布について $k \geq 3$ のすべてのキュムラントはゼロである。

定理 2 $\langle\ \rangle_{\mathrm{c}}$ の中に統計的に独立な変数が入っていればキュムラントはゼロとなる。

正規分布は平均と共分散行列を与えれば一意に定まるから、共分散もしくは相関係数で変数同士の相関を記述するということは、暗に確率分布として正規分布を想定していることを意味する。定理 1 は、非線形相関を取り込むためには、必然的に高次のキュムラントを取り込まねばならないことを示している。逆に言えば、長い間正規分布が相関解析の主役であったことを思えば、$\Psi(\boldsymbol{s})$ を有限項で近似しておけば、実用的にはさほど悪くない結果を与えるだろうと予想できる。

[2]統計学の教科書ではたとえば $\langle x_i x_j \rangle_{\mathrm{c}}$ の代わりに、$\mathrm{cum}(x_i x_j)$ や $\kappa_{i,j}$ と記すことがある。また、$\langle\cdot\rangle$ の代わりに $\mathrm{E}(\cdot)$ 等の記号を使うことがある。しかし、多変数を扱う場合は本論文の記号の方が簡潔であろう。なお、定義から当然であるが、一般には $\langle\cdot\rangle_{\mathrm{c}} \neq \langle\cdot\rangle$ であることに注意されたい。

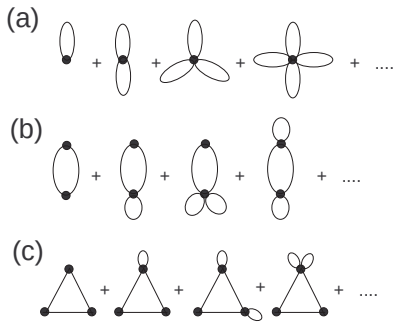

図 1: 図式的に表したクラスター展開。 (a) K_1, (b) K_2, (c) K_3.

2.2　SCA: 疎相関の近似

上記の定理をもとに、$\Psi(\boldsymbol{s})$ 自然な近似的表式について考えよう。まず式 (2) を次のように書き換えよう。

$$\Psi(\boldsymbol{s}) = K_1 + K_2 + ... + K_i + ... \qquad (3)$$

ここで K_i は i 個の異なった変数を持つ項の和を表す。たとえば K_2 は

$$K_2 = \sum_{i \neq j} \left[\frac{1}{2!} s_i s_j \langle x_i x_j \rangle_{\mathrm{c}} + \frac{1}{3!} {s_i}^2 s_j \langle {x_i}^2 x_j \rangle_{\mathrm{c}} + \ldots \right]$$

のようになる。各項は一般に複雑になりえるが、これをイメージするための便利な方法は、図式的に表示してみることである。図 1 に、K_1, K_2, K_3 を表した。頂点 $\bullet$ は x_1 のような変数、辺は $\langle\ \rangle_{\mathrm{c}}$ の中での積を表している。辺の数が k と等しくなるように、たとえば $\langle {x_1}^2 x_2 \rangle_{\mathrm{c}}$ は $\langle x_1 \cdot x_1 \cdot x_2 \cdot \rangle_{\mathrm{c}}$ のようにみなして、3 つの辺を与える。統計力学の用語を流用して、K_i を i 体のクラスターと呼ぶ。また、$i \geq 2$ に対して K_i の中の各項を i 体の交差キュムラントと呼ぶ。

今、多変数同士の相関が比較的疎である状況を考えよう。このようなときには、$\langle\ \rangle_{\mathrm{c}}$ の中に入る変数の種類が多ければ多いほど、定理 2 により、それがゼロになる確率も高くなる。極端な話、$\langle x_1 x_2 \cdots x_n \rangle_{\mathrm{c}}$ はほぼ確実にゼロであろう。そこで、非自明な最低次の近似として、

$$\Psi(\boldsymbol{s}) \simeq K_1 + K_2 \qquad (4)$$

とおき、これを、疎相関の近似（sparse correlation approximation; SCA) と呼ぶことにする。

数学的にこれが妥当であるためには、ひとつの変数が強く相関する変数の個数の期待値が 1 程度でなければならない。一般にはこれは成り立っているとは限らないが、仮に式 (4) が数学的に $\Psi(\boldsymbol{s})$ のよい近似となっていなくても、相関を表す基本量という意味で K_2 の地位は

order	cumulant	moment
2	$\langle xy\rangle_{\rm c}$	$\langle xy\rangle$
3	$\langle xy^2\rangle_{\rm c}$	$\langle xy^2\rangle$
3	$\langle x^2y\rangle_{\rm c}$	$\langle x^2y\rangle$
4	$\langle x^2y^2\rangle_{\rm c}$	$\langle x^2y^2\rangle - \langle x^2\rangle\langle y^2\rangle - 2\langle xy\rangle^2$
4	$\langle xy^3\rangle_{\rm c}$	$\langle xy^3\rangle - 3\langle xy\rangle\langle y^2\rangle$
4	$\langle x^3y\rangle_{\rm c}$	$\langle x^3y\rangle - 3\langle x^2\rangle\langle xy\rangle$

表 1: $\langle x\rangle = \langle y\rangle = 0$ におけるモーメントと 2 体交差キュムラントの関係 [10]。

変わらない。1 体のクラスター K_1 は個々の変数のみに関係するから、相関解析の上では K_2 が常に主要項となるからである。ゆえ、本論文の後半では、2 体の交差キュムラント $\langle {x_i}^\mu {x_j}^\nu\rangle_{\rm c}$ をどのように相関係数と関連付けるかを議論する（μ と ν は正整数)。参考のため表 1 に、2 体交差キュムラントとモーメントの関係を掲げた。

なお、定理 2 は、複数の変数がどのように独立であっても成り立つ。たとえば、二つの変数が定数分布として独立でも、非相関のガウス分布として独立でも、同じくゼロである。実データにはノイズがつき物であるが、この性質によりノイズへの頑強性がある程度保証される。

3 確率分布の対称性分解

3.1 状態ベクトル表現

前節では p を、$\boldsymbol{x} \in \mathcal{D}_0$ を引数とする実関数と考えた。後の議論の便宜上、本節では p をヒルベルト空間 $\mathcal{H}_0$ の元と考え、Dirac のブラ-ケット記法を用いて $|p\rangle$ と表す [2]。われわれは $\mathcal{H}_0$ を、位置演算子と呼ばれるエルミート演算子 $\hat{\boldsymbol{x}}$ の固有ベクトル $\{|\boldsymbol{x}\rangle | \boldsymbol{x} \in \mathcal{D}_0\}$ により張られる空間として定義する [6]。ここで位置固有ケット $|\boldsymbol{x}\rangle$ は、その固有値によりパラメトライズされているとする。すなわち、$\hat{\boldsymbol{x}}$ の第 i 成分 $\hat{x}_i$ は、$\boldsymbol{y} \in \mathcal{D}_0$ なる $|\boldsymbol{y}\rangle$ に対して、方程式 $\hat{x}_i|\boldsymbol{y}\rangle = y_i|\boldsymbol{y}\rangle$ を満たす。量子力学の言葉を用いて、$\mathcal{H}_0$ を状態空間、その元を状態ベクトルと呼ぶことにする。

仮定より、位置固有ケットは $\mathcal{H}_0$ において完全系をなし、$\forall |f\rangle \in \mathcal{H}_0$ は位置固有ケットの 1 次結合として形式的に

$$|f\rangle = \int_{\mathcal{D}_0} d\boldsymbol{x}\ |\boldsymbol{x}\rangle f(\boldsymbol{x}) \tag{5}$$

と表せる。ここで、$f(\boldsymbol{x})$ は状態ベクトル $|f\rangle$ の $|\boldsymbol{x}\rangle$ 成分である。この式から逆に、$\forall |f\rangle \in \mathcal{H}_0$ は、すべての $\boldsymbol{x} \in \mathcal{D}_0$ に対して $|\boldsymbol{x}\rangle$ 成分を与えることにより定義されることがわかる。この意味で、関数 $f(\boldsymbol{x})$ は、状態ベクトル $|f\rangle$ と 1 対 1 に対応している。

次に、$\mathcal{H}_0$ の双対空間として $\bar{\mathcal{H}_0}$ を考え、$|\boldsymbol{x}\rangle$ の対応物を $\langle\boldsymbol{x}|$ と表す。これを位置固有ブラと呼ぶ。$\mathcal{H}_0$ の元 $|f\rangle, |h\rangle$ の内積を、ブラとケットの積 $\langle f|h\rangle$ として定義する。* で複素共役を表すことにし、$\langle f|h\rangle = \langle h|f\rangle^*$ と約束する。これより、式 (5) に対応して、$\langle f| = \int_{\mathcal{D}_0} d\boldsymbol{x} f(\boldsymbol{x})^* \langle\boldsymbol{x}|$ が成り立つ。また、位置固有状態の完全性より、$\langle\boldsymbol{x}|\boldsymbol{x}'\rangle = \delta(\boldsymbol{x} - \boldsymbol{x}')$ が成り立たねばならないことがわかる。ここで $\delta(\cdot)$ は Dirac のデルタ関数である。

以上の議論から、$|f\rangle, |h\rangle \in \mathcal{H}$ の内積を、

$$\langle f|h\rangle = \int_{\mathcal{D}_0} d\boldsymbol{x}\ f(\boldsymbol{x})^* h(\boldsymbol{x}) \tag{6}$$

のように具体的に計算できることがわかる。また、確率密度関数 $p(\boldsymbol{x})$ には、$|\boldsymbol{x}\rangle$ と $|p\rangle$ の内積 $\langle\boldsymbol{x}|p\rangle$ という意味を与えることができる。今、$\boldsymbol{x} = (x_1, ..., x_n)^T$ から二つの変数を取り出し、簡単のため $\boldsymbol{r} = (x, y)^T$ とおく。これに対応して前と同様に位置固有ケット $\{|\boldsymbol{r}\rangle | \boldsymbol{r} \in \mathcal{D}\}$ を考え、それの張る空間を $\mathcal{H}$ と表す。$\mathcal{D} \subset \mathcal{D}_0$ は、$p(\boldsymbol{x})$ の周辺分布 $p(\boldsymbol{r})$ の定義域である。状態ベクトル表現では $p(\boldsymbol{r})$ は $\langle\boldsymbol{r}|p\rangle$ と書ける。次に、$\mathcal{H}$ の状態ベクトル $|x^\mu y^\nu\rangle$ を、その $\boldsymbol{r}$ 成分により、$\langle\boldsymbol{r}|x^\mu y^\nu\rangle = x^\mu y^\nu$ と定義する。これを用いると 2 体の交差モーメントは、$\langle x^\mu y^\nu\rangle = \langle p|x^\mu y^\nu\rangle$ と簡潔に書ける。なぜなら $\mathcal{H}$ の双対空間 $\bar{\mathcal{H}}$ における展開式 $\langle p| = \int d\boldsymbol{r}\ p(\boldsymbol{r})^* \langle\boldsymbol{r}| = \int d\boldsymbol{r}\ p(\boldsymbol{r})\langle\boldsymbol{r}|$ を用いると、

$$\langle p|x^\mu y^\nu\rangle = \int_{\mathcal{D}} d\boldsymbol{r} \langle p|\boldsymbol{r}\rangle\langle\boldsymbol{r}|x^\mu y^\nu\rangle = \int_{\mathcal{D}} d\boldsymbol{r} p(\boldsymbol{r}) x^\mu y^\nu$$

となるためである。特に、通常の共分散は $\langle p|xy\rangle$ となる。

3.2 $\mathrm{C}_{4\mathrm{v}}$ 群の導入

通常の共分散で現れる $|xy\rangle$ は、たとえば、x と y の値を入れ替えるような変換に対して不変である。このような対称性には何か深い意味があるように思われるので、その性質をより正確に記述するために、x, y 空間内で定義される対称操作の集合 G を考えよう。集合といってもさしあたり任意であるが、対称操作としては回転と鏡映を考えておけば十分一般的であろう。そして、対称操作の首尾一貫性を保証するために、G が群の公理を満たすことを要請する。すなわち、G において定義された積演算 $\circ$ に対し、次の性質を要請する：(1) **結合律**：任意の a, b, $c \in G$ に対し $(a \circ b) \circ c = a \circ (b \circ c)$。(2) **単位元**：任意の $g \in G$ に対し $e \circ g = g = g \circ e$ となる $e \in G$ が存在する。(3) **逆元**：任意の $g \in G$ に対して、$g' \circ g = g \circ g' = e$ となるような元 $g' \in G$ が存在する。

意外なことに、これらの公理を満たす集合は非常に限られ、実際、回転と鏡映を組みあわせてできる群（これを点群と称する）のうち、群の公理を満たすものは 32

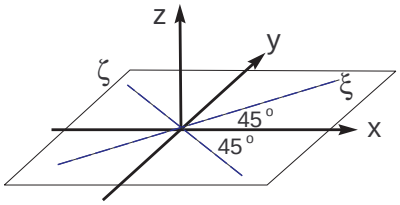

図 2: C_{4v} 群における対称軸。

個しかないことが知られている [8]。2 次元平面で定義される点群のうちでもっとも一般的なものは C_{4v} と呼ばれ、以下の 8 つの対称操作から成る:

$$C_{4v} = \{e, C_4, C_2, {C_4}^3, \sigma_x, \sigma_y, \sigma_\xi, \sigma_\eta\}$$

ここで、e は恒等演算（何もしない）を表し、C_4, C_2, および ${C_4}^3$ はそれぞれ z 軸周りの $\pi/2$-, π-, および $3\pi/2$-回転を表す。図 2 はこの群に現れる対称軸を図示している。鏡映操作は xz-, yz-, ξz- および ηz-面に関して定義され、それぞれ $\sigma_x, \sigma_y, \sigma_\xi, \sigma_\eta$ と表される。

これらの対称操作は、$\mathcal{D}$ 内の 1 点 $\boldsymbol{r}$ を別の 1 点 $\boldsymbol{r}'$ に移すから、$\mathcal{H}$ における $g \in C_{4v}$ の作用を、位置固有状態 $|\boldsymbol{r}\rangle$ を用いて、$g|\boldsymbol{r}\rangle = |\boldsymbol{r}'\rangle$ などと自然に定義できる。これを簡単に $g|\boldsymbol{r}\rangle = |g\boldsymbol{r}\rangle$ と表すことにすると、状態ベクトル $\forall |f\rangle \in \mathcal{H}$ に対する $g \in C_{4v}$ の作用は、その $\boldsymbol{r}$ 成分を用いて、$\langle \boldsymbol{r}|g|f\rangle = \langle g^{-1}\boldsymbol{r}|f\rangle = f(g^{-1}\boldsymbol{r})$ のように定義される[3]。

3.3 対称性分解定理

一般に、l 個の 1 次独立な $|\phi_1\rangle, .., |\phi_l\rangle \in \mathcal{H}$ が、$\forall g \in G$ に対し、

$$g|\phi_j\rangle = \sum_{i=1}^{l} |\phi_i\rangle D_{ij}(g) \tag{7}$$

を満たす時、$\{|\phi_1\rangle, .., |\phi_l\rangle\}$ で張られる l 次元の空間を、G の不変部分空間と呼ぶ。対称操作の後もこの空間からはみ出さない、という意味である。行列 $D_{ij}(g)$ は g の表現行列と呼ばれる。

$|\phi_i\rangle \otimes |\phi_j\rangle$ のような直積で張られる l^2 次元の空間も不変部分空間となることからも分かる通り、不変部分空間の次元に上限はない。では次元に下限はあるのだろうか。有限群論の教えるところでは [8]、不変部分空間には、既約不変部分空間と呼ばれるある最小単位が、所定の有限個存在する。より具体的に言えば、$\{|\phi_1\rangle, .., |\phi_l\rangle\}$ の 1 次結合をつくることより、表現行列をブロック対角化できる。そしてその時、異なるブロック γ, γ' に対応する $|\varphi_\cdot^{(\gamma)}\rangle, |\varphi_\cdot^{(\gamma')}\rangle \in \mathcal{H}$ について、

$$\langle \varphi_\cdot^{(\gamma)} | \varphi_\cdot^{(\gamma')} \rangle \propto \delta_{\gamma,\gamma'} \tag{8}$$

なる直交関係が成り立つ。$\delta_{\gamma,\gamma'}$ はクロネッカーのデルタである。したがって、$|p\rangle$ は γ により分解することができる[4]。$\boldsymbol{r}$-表示で表せば、

$$p(\boldsymbol{r}) = \sum_\gamma \pi^{(\gamma)}(\boldsymbol{r}). \tag{9}$$

今考えている C_{4v} 群に関しては、既約表現の個数は A_1, A_2, B_1, B_2, および E と呼ばれる 5 つしかないことが知られている。E 表現は 2 次元表現で、他は 1 次元表現である。γ =E の時は、上式の $\pi^{(E)}$ は $\pi_1^{(E)}$ と $\pi_2^{(E)}$ の 1 次結合と考える。

ここで、次の非常に興味深い定理を与える。

定理 3 $\{|xy\rangle\}$ は C_{4v} の B_2 既約不変部分空間を張る。

一般に、群の既約表現は表現行列の対角和（これを指標と呼ぶ）により分類される。C_{4v} の指標を表 2 にまとめた。容易に確かめられるように、$\langle \boldsymbol{r}|xy\rangle$ は、$\{e, C_2, \sigma_\xi, \sigma_\eta\}$ については不変だが、$\{C_4, {C_4}^3, \sigma_x, \sigma_y\}$ については符号を変える。したがって、表 2 と見比べることにより、$\mathcal{H}$ の部分空間 $\{|xy\rangle\}$ は、群 C_{4v} の B_2 既約不変部分空間を張っていることがわかる。これで定理が証明された。

この定理と式 (8), (9) によれば、$\langle xy|p\rangle = \langle xy|\pi^{(B2)}\rangle$ が成り立つ。すなわち、通常の共分散は $|p\rangle$ の中にある B_2 という対称性の成分を抽出したものと言える。これを一般化したものが次の定理である。

定理 4 既約表現 γ に属する $|\varphi^{(\gamma)}\rangle \in \mathcal{H}$ に対し $\langle \varphi^{(\gamma)}|p\rangle = \langle \varphi^{(\gamma)}|\pi^{(\gamma)}\rangle$ が成り立つ。すなわち、状態ベクトル $|\varphi^{(\gamma)}\rangle$ は γ 成分を抽出するフィルタとして作用する。

証明は既約表現の定義から明らかである。これを**対称性分解定理**と呼ぶことにする。定理 3 および 4 は、相関係数の一般化のための重要な指導原理となりえる。すなわち、$\mathcal{H}$ には 5 つの自然な対称性（既約表現）が存在するにもかかわらず、従来はそのうちのひとつ、B_2 表現に属する対称性フィルタしか使われてこなかった。だとすれば残りの 4 つに対応するものを定義すれば、それは対称性の観点から見て自然な一般化になるわけである。

[3] これは、$\langle \boldsymbol{r}|g|f\rangle$ を $\langle \boldsymbol{r}|g\cdot|f\rangle$ とみなして得られる

$$\langle \boldsymbol{r}|g|f\rangle = \int d\boldsymbol{r}' \langle \boldsymbol{r}|g|\boldsymbol{r}'\rangle\langle \boldsymbol{r}'|f\rangle = \int d\boldsymbol{r}' \delta(\boldsymbol{r} - g\boldsymbol{r}') f(\boldsymbol{r}')$$

に、変数変換 $\boldsymbol{r}'' = g\boldsymbol{r}'$ を施すと、$f(g^{-1}\boldsymbol{r})$ を得ることから了解されよう。C_{4v} の場合はこの変数変換のヤコビアンが常に 1 であることに注意。このことはまた、$g^{-1}|\boldsymbol{r}\rangle$ と $\langle \boldsymbol{r}|g$ が双対対応の関係にあることを示している。

[4] 別の群の例。単位元 e と空間反転 I からなる群（C_i と呼ばれる）は、偶と奇という 2 つの既約表現をもち、それぞれの基底は偶関数と奇関数である。偶関数と奇関数の内積が 0 というよく知られた事実は、群論的には本文で述べた直交性の表れである。

C_{4v}	e	$C_4, {C_4}^3$	C_2	σ_x, σ_y	σ_ξ, σ_η
A_1	1	1	1	1	1
A_2	1	1	1	−1	−1
B_1	1	−1	1	1	−1
B_2	1	−1	1	−1	1
E	2	0	−2	0	0

表 2: C_{4v} 群の指標 [8]。

4 相関係数の一般化

他の既約表現に属する $|x^\mu y^\nu\rangle$ の形の状態ベクトルを探そう。特定の既約表現を張る基底を得る手続きは、一般には射影演算子の方法 [8] を使ってなされる。しかしここではノイズへの頑強性の観点から、$0 < \mu, \nu \leq 3$ および $\mu+\nu \leq 4$ の範囲に限るものとし、低次のものから総当り的に対称性を調べてもよい。たとえば、$\mu+\nu = 3$ 次の状態ベクトル $|x^2y\rangle$ から考えよう。$\langle \boldsymbol{r}|C_4|x^2y\rangle = -y^2x$ であることから、$C_4|x^2y\rangle = -|xy^2\rangle$ であることがわかる。そこで、空間 $\{|x^2y\rangle, |xy^2\rangle\}$ を考えると、これは C_{4v} の不変部分空間を張っていることが確かめられる。たとえば、$C_4|xy^2\rangle = |x^2y\rangle$ であるから、確かにこの空間は C_4 に対して不変である。C_4 の指標が 0 であることもわかる。同様にして、他の演算子に対してもあらわに指標を求めることができて、指標の表と比較することにより、この空間が E 表現を張っていることがわかる。

$|xy^3\rangle, |x^2y^2\rangle$ に対して以上の操作を繰り返すことで、$|x^2y^2\rangle$ が A_1 表現、$|xy^3\rangle - |x^3y\rangle$ が A_2 表現であることを確認できる[5]。これらの結果を 2 体交差キュムラントと結びつけるために、下記の定理が成り立つことに注意する（証明略)。

定理 5 $\forall g \in C_{4v}$ に対して、$\langle x^\mu y^\nu\rangle_c$ は $\langle p|x^\mu y^\nu\rangle$ と同様に変換される。

以上より、一般化共分散として、われわれは次のものを定義する。

$$
\begin{aligned}
\mathsf{C}(B_2) &= \langle xy\rangle_c && (10)\\
\mathsf{C}(E_1) &= \left[\langle xy^2\rangle_c + \langle x^2y\rangle_c\right]/2 && (11)\\
\mathsf{C}(E_2) &= \left[\langle xy^2\rangle_c - \langle x^2y\rangle_c\right]/2 && (12)\\
\mathsf{C}(A_1) &= \langle x^2y^2\rangle_c && (13)\\
\mathsf{C}(A_2) &= \left[\langle xy^3\rangle_c - \langle x^3y\rangle_c\right]/2 && (14)
\end{aligned}
$$

これらは 4 次までの 2 体交差キュムラントの 1 次結合になっている[6]。ただし、簡単のため E 表現の二つの基底を、E の添え字で区別している。E は 2 次元表現であるから、上記の $\langle xy^2\rangle_c$ と $\langle x^2y\rangle_c$ の 1 次結合の仕方には、特に必然性はない。また、1 変数の場合に分散の値を歪度や尖度と直接比較するのが適切ではないのと同じく、異なる対称性を持つ一般化共分散の値同士を直接比べるのは適当ではない。実用上は、一般化共分散を $\left[\langle {x_i}^2\rangle\langle {x_j}^2\rangle\right]^{k/4}$ で割ることで無次元化するのが便利である。明らかにこの因子は A_1 にしたがって変換するため、式 (10)-(14) で指定した対称性を乱すことはない。規格化された一般化共分散を**一般化相関係数**と呼んでおく。

5 実験

ここでは、解析的に一般化相関係数が求まるモデルを使って、非線形相関を定量的に捉える様子を示す。強い相関を持つ時系列のモデルとして

$$
\begin{aligned}
x(t) &= \sqrt{2}\cos(\omega_1 t + \alpha)\\
y(t) &= \sqrt{2}\sin(\omega_2 t + \beta)
\end{aligned}
$$

を考える。時間領域では確率分布は定数と考える。明らかに、平均 $\langle x\rangle$ および $\langle y\rangle$ はゼロで、分散 $\langle x^2\rangle_c$ および $\langle y^2\rangle_c$ は 1 である。したがって、一般化共分散は一般化相関係数と一致する。a がゼロでない限り $\langle \sin(at+b)\rangle = 0$ となるという事実を用いると（b は任意の実数)、初等的な三角関数の変形を繰り返すことにより、次の結果を得る:

$$
\begin{aligned}
\mathsf{C}(B_2) &= \delta_{\omega_1,\omega_2}\sin\Omega_1^{\beta,\alpha}\\
\mathsf{C}(E_1) &= -\frac{\delta_{\omega_1,2\omega_2}}{\sqrt{2}}\cos\Omega_2^{\alpha,\beta} + \frac{\delta_{2\omega_1,\omega_2}}{\sqrt{2}}\sin\Omega_2^{\beta,\alpha}\\
\mathsf{C}(E_2) &= -\frac{\delta_{\omega_1,2\omega_2}}{\sqrt{2}}\cos\Omega_2^{\alpha,\beta} - \frac{\delta_{2\omega_1,\omega_2}}{\sqrt{2}}\sin\Omega_2^{\beta,\alpha}\\
\mathsf{C}(A_1) &= -\frac{\delta_{\omega_1,\omega_2}}{2}\left[1 + 2\sin^2(\Omega_1^{\alpha,\beta})\right]\\
\mathsf{C}(A_2) &= \frac{\delta_{\omega_1,3\omega_2}}{4}\sin\Omega_3^{\alpha,\beta} - \frac{\delta_{3\omega_1,\omega_2}}{4}\sin\Omega_3^{\beta,\alpha}
\end{aligned}
$$

ただし $\Omega_c^{a,b} = a - bc$ という記号を用いた。上式から一般化相関係数は振動数比が簡単な有理数でないとゼロになってしまうことが分かる。これは、もし振動数比が半端な値であれば $t \to \infty$ の極限で xy 平面を軌道が埋め尽くすことになり、結果として x と y が統計的に独立とみなせるためと解釈できる。

図 3 は、時系列の軌道と対応する一般化相関係数を、いくつかのパラメターの組について示したものである。この軌道は Lissajous 図形としてよく知られている。(a) は線形相関であるから、従来の相関係数注目 $\mathsf{C}(B_2)$ によ

[5] $|x^\mu y^\nu\rangle$ において μ と ν を非ゼロに限れば、4 次までの交差キュムラントで B_1 表現は現れない。

[6] 更に高次のキュムラントを用いれば、別の形の一般化共分散を定義できるが、ノイズへの頑強性という意味で、実用上は低次のキュムラントを用いるのが合理的である。

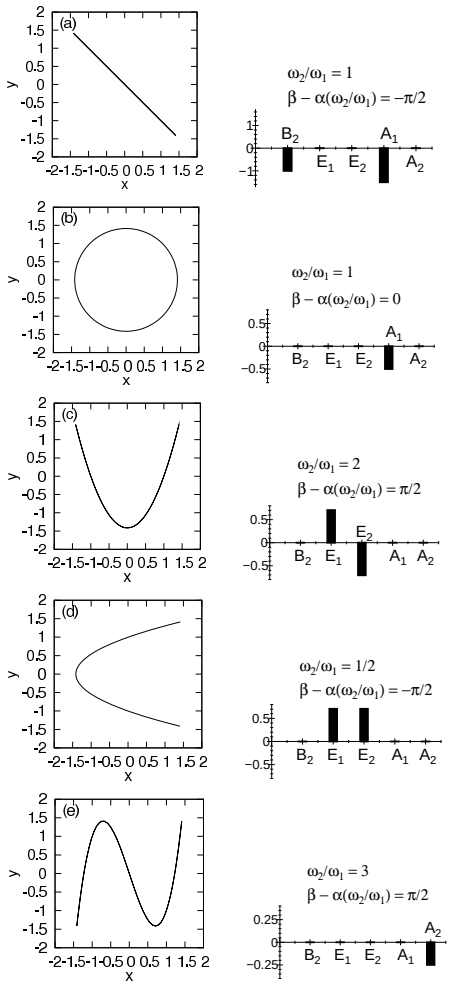

図 3: Lissajous 図形と、対応する一般化相関係数。

りその特徴を評価することができる。しかし (b)-(e) のような軌道には何の情報も与えない。一方、今回新たに定義した一般化共分散は、非線形相関をよく捉えていることわかる。たとえば、冒頭に挙げた円形の分布 (b) の特徴は、$\mathsf{C}(\mathrm{A}_1)$ により定量的に捉えられている。また、(c)-(e) のような非線形相関の場合も、 $\mathsf{C}(\mathrm{E}_1)$ や $\mathsf{C}(\mathrm{A}_2)$ が、定量的なよい評価尺度になっていることが分かる。

6 まとめ

本論文では非線形相関を題材に、パターン認識における群論的方法を新たに提案した。まず、多次元正規分布を拡張する出発点として、キュムラント母関数のクラスター展開を考えた。それに基づいて、変数間の非線形相関を記述する自然な量として、2 体の交差キュムラントを取るべきことを提案した。

次に、変数対の相関を 2 次元空間の幾何学的パターンとして捉え、それを群の既約表現を使って特徴付けることを提案した。また、伝統的な共分散の群論的性質を調べ、それが $\mathrm{C}_{4\mathrm{v}}$ 群の B_2 表現の基底として把握できることを示した。そしてこの事実から、一般化共分散を自然に定義できることを示した。すなわち、低次の 2 体交差キュムラントから、$\mathrm{C}_{4\mathrm{v}}$ 群の既約表現となるようなものを作ればよい。われわれの知る限り本研究は、相関解析、ひいてはパターン認識に、対称性の概念を持ち込んだ先駆的な試みのひとつであると思われる。

最後に、相関した時系列の可解モデルを使って、一般化相関係数の特徴を実験的に明らかにした。ここでは素朴なモデルについての結果を示したのみであるが、実際の時系列データに関する評価結果などは、別途発表の予定である。

謝辞

IBM 東京基礎研究所の吉田一星氏、宅間大介氏との議論は有益であった。謝意を表する。

参考文献

[1] T. Idé. Pairwise symmetry decomposition method for generalized covariance analysis. In *Proceedings of the fifth IEEE International Conference in Data Mining (ICDM 05)*, 2005. (to appear).

[2] R. Kondor and T. Jebara. A kernel between sets of vectors. In *Proc. the 20th International Conference on Machine Learning*, 2003.

[3] R. Kubo. Generalized cumulant expansion method. *Journal of the Physical Society of Japan*, 17(7):1100–1120, 1962.

[4] B. Schölkopf and A. J. Smola. *Learning with Kernels*. The MIT Press, 2002.

[5] A. Smola and R. Kondor. Kernels and regularization on graphs. In *Proc. of 16th Annual Conference on Computational Learning Theory and 7th Kernel Workshop*, pp. 144–158, 2003.

[6] サクライ. 現代の量子力学, 上. 吉岡書店, 1989.

[7] 宮川雅巳. グラフィカルモデリング. 朝倉書店, 1997.

[8] 犬井鉄郎, 田辺行人, 小野寺嘉孝. 応用群論. 裳華房, 1976.

[9] 大津展之, 栗田多喜夫, 関田巌. パターン認識. 朝倉書店, 1996.

[10] 竹村彰通, 谷口正信. 統計学の基礎 I. 岩波書店, 2003.