

令和 2 年 6 月 15 日現在

機関番号：32689

研究種目：基盤研究(C)（一般）

研究期間：2016～2019

課題番号：16K00060

研究課題名（和文）B-splineコピュラの相関構造とその推定

研究課題名（英文）Dependence properties and estimation of B-spline copulas

研究代表者

DOU XIAOLING (DOU, XIAOLING)

早稲田大学・データ科学センター・准教授（任期付）

研究者番号：10516868

交付決定額（研究期間全体）：（直接経費） 3,400,000円

研究成果の概要（和文）：B-spline基底関数を用いてB-splineコピュラを定義しその性質を調べ、推定を行う。B-spline基底関数をコピュラを構成できるようにnormalizationして、Bernsteinコピュラの一般化としてB-splineコピュラを定義した。B-splineコピュラはBernsteinコピュラより少ない基底関数で幅広い相関構造を表せることが分かった。また、相関を最大にするB-splineコピュラはFreshet-hoeffding boundに達することを証明した。EMアルゴリズムやグリッド法でB-splineコピュラのパラメータを推定しその有用性を実データで実証した。

研究成果の学術的意義や社会的意義

コピュラは金融工学やリスク、土木工学等多くの分野で活用されている。実際に様々なデータに対して、シミュレーションやモデリングするためには、柔軟性の高いコピュラが望ましい。B-splineコピュラを定義し、その相関構造の性質を調べることで、B-splineコピュラは従来のコピュラより高い柔軟性を持ち、多様なデータ構造を表現することができることが分かった。また、少ない基底関数で高い相関を表せることは計算コストも少なくなる。さらに、グリッド法やEMアルゴリズムの方法でB-splineコピュラを推定できるので、B-splineコピュラが幅広く利用できると思う。

研究成果の概要（英文）：Based on B-spline basis functions, we define a new class of copulas, called B-spline copulas. The B-spline copulas include the Bernstein copulas as a special case where the basis functions are generated without interior knots. By comparing the ranges of the maximum correlation of the B-spline copulas and the Bernstein copulas, we found that the B-spline copulas are more flexible. That is, with less basis functions, they can express higher correlations. We showed that the B-spline copulas of the maximum correlation case can attain the Frechet Hoeffding upper bound when the number of basis functions goes to infinity. We also proved the B-spline copulas of the maximum correlation case have total positivity of any order. We proposed to estimate the B-spline copulas by a grid method and some EM algorithms. These methods are examined with real data sets.

研究分野：統計科学

キーワード：コピュラ 多変量分布 分布の相関構造 分布の推定 EMアルゴリズム B-spline関数 B-splineコピュラ

科研費による研究は、研究者の自覚と責任において実施するものです。そのため、研究の実施や研究成果の公表等については、国の要請等に基づくものではなく、その研究成果に関する見解や責任は、研究者個人に帰属されます。

様式 C - 19, F - 19 - 1, Z - 19 (共通)

### 1. 研究開始当初の背景

変数間の相互依存関係を表す関数としてコピュラが金融工学やリスクなどの分野でよく用いられ、もっとも利用されているのが Gaussian コピュラである。しかし、Gaussian コピュラは相関構造が単純な時にはその同時分布を再現することができるが、実際には多様な相関構造を持つ実データをモデリングする時や、リスク分野のシミュレーション分析するためには、もっと柔軟性のあるコピュラが望ましい。近年、Bernstein 基底関数で構築された Bernstein コピュラの研究が進んでおり、基底関数が大きければ、Bernstein コピュラが任意の精度で同時分布を近似できることが知られている。しかし、多くの基底関数が必要である。Bernstein 関数は内部節点が存在しない時の B-spline 基底関数の特別な場合なので、一般的な B-spline 関数を用いれば、B-spline コピュラが定義でき、Bernstein コピュラよりも柔軟性が高くなる。

### 2. 研究の目的

Bernstein コピュラの一般化として B-spline コピュラを定義し、その相関構造の性質を調べる。また、周辺分布が与えられたときに、B-spline コピュラを推定する方法を提案し、実データに適用することを目的とした。

### 3. 研究の方法

Bernstein コピュラ密度は以下のように定義されている：

$$c^{\text{Bern}}(u, v; R) = mn \sum_{k=1}^m \sum_{l=1}^n r_{kl} b_{k-1, m-1}(u) b_{l-1, n-1}(v),$$

ここで、Bernstein 基底関数は

$$b_{k,n}(u) = \binom{n}{k} u^k (1-u)^{n-k}, \quad u \in [0, 1].$$

Bernstein 基底関数は非負で

$$b_{k-1, n-1}(u) \geq 0$$

与えられた次元数  $n$  に対して、生成された  $n$  個の Bernstein 関数の和が 1 になる性質を持つ：

$$1 = \sum_{k=1}^n b_{k-1, n-1}(u) \quad \text{for all } u \quad .$$

B-spline 基底関数もこれらの性質を持つことを確認し、それを用いて B-spline コピュラを定義する事にした。また、コピュラを構築するために B-spline 基底関数を基準化し、コピュラの周辺分布を一様分布になるようにする必要があった。このように B-spline コピュラは Bernstein コピュラの一般化として得られた。

同じ数の基底関数を持つ Bernstein コピュラと B-spline コピュラの最大相関係数を比較し、基底関数が増えるとともに、この二種類のコピュラの最大相関係数の変化を考察する。そのために、相関を最大にする B-spline コピュラの特長、証明、及び異なる次数について最大相関係数を計算する方法を研究する必要があった。B-spline 多項式の次数と等間隔に置かれた内部節点の数を決め、B-spline コピュラが表せる相関の範囲を計算した。また、Bernstein コピュラの場合と同じように、B-spline コピュラの相関構造の性質を調べた。

### 4. 研究成果

B-spline 基底関数は区分的多項式であり、一般的な B-spline 基底関数を用いるために、内部節点の数  $(p-1)$  と配置及び多項式の次数  $(d)$  を決める必要がある。B-spline 基底関数に対して検討し、Bernstein 基底関数と同じような性質を持つことが分かった：

$$N_{k-d-1}^d(x) \geq 0, \quad 1 = \sum_{k=1}^n N_{k-d-1}^d(x) \quad \text{for all } x, \quad \text{ここで } n := d+p.$$

B-spline 基底関数を基準化して、以下のように B-spline コピュラ  $c(u, v; R)$  を定義した：

$$c(u, v; R) = \sum_{k=1}^m \sum_{l=1}^n r_{kl} \phi_{k,m}(u) \phi_{l,n}(v)$$

$$\text{where } r_{kl} \geq 0, \quad \sum_{l=1}^n r_{kl} = q_{k,m}, \quad \sum_{k=1}^m r_{kl} = q_{l,n},$$
$$\phi_{k,m}(x) = \frac{1}{q_{k,m}} N_{k-d-1}^d(x), \quad q_{k,m} = \int_0^1 N_{k-d-1}^d(x) dx.$$

また、内部節点はない時に、B-spline コピュラは Bernstein コピュラになることも示した。

相関を最大にするコピュラが表せられる相関係数の範囲を調べるために、相関を最大にする B-spline コピュラを特定した。Bernstein コピュラの場合と同じように、相関を最大にする B-spline コピュラもパラメータ行列  $R$  が対角行列の場合であることも証明した。つまり、

$$R = (r_{kl}) = \text{diag}(q_{k,n})_{1 \leq k \leq n}.$$

同じ数の基底関数を持つ Bernstein コピュラと B-spline コピュラの最大相関係数を比較し、基底関数が多くするとともに、この二種類のコピュラの最大相関係数の変化を考察する。そのために、異なる次数について最大相関係数を計算する方法を研究する必要があった。私たちは、Stirling number of the second kind,  $S(n, k)$ , を用いて、B-spline 関数のモーメントの公式

$$\int t^h N_i^d(t) dt = \begin{cases} \sum_{\ell=0}^h i^\ell \binom{h}{\ell} \frac{S(h+d+1-\ell, d+1)}{\binom{h+d+1-\ell}{d+1}}, & i \geq 0, \\ \frac{i+d+1}{d+1} \frac{S(h+i+d+1, i+d+1)}{\binom{h+d+1}{d+1}}, & -d \leq i \leq 0, \\ 0, & i < -d, \end{cases}$$

で内部節点が等間隔に置かれた時の最大相関の公式を導出した:

$$1 - \frac{d+1}{(n-d)^2} + \frac{d(d+3)(2d+3)}{5(d+2)(n-d)^3}.$$

これで異なる次数の最大相関係数の以下のように求められる。

Table: Maximum correlations

	Bernstein	$d = 0$	$d = 1$	$d = 2$	$d = 3$
$n = 2$	0.333	0.75	0.333	NA	NA
$n = 3$	0.5	0.889	0.667	0.5	NA
$n = 4$	0.6	0.938	0.827	0.688	0.6
$n = 5$	0.667	0.96	0.896	0.796	0.72
$n = 6$	0.714	0.972	0.931	0.867	0.796
$n = 7$	0.75	0.980	0.951	0.908	0.851
$n = 8$	0.778	0.984	0.963	0.933	0.892
$n = 9$	0.8	0.988	0.971	0.949	0.919
$n = 10$	0.818	0.99	0.977	0.960	0.937
$n$	$1 - \frac{2}{n+1}$	$1 - \frac{1}{n^2}$	$1 - \frac{2(3n-5)}{3(n-1)^3}$	$1 - \frac{6n-19}{2(n-2)^3}$	$1 - \frac{2(50n-231)}{25(n-3)^3}$

この表で Bernstein コピュラと比べると、B-spline コピュラは比較的少ない基底関数で大きな相関を表すことができ、より高い柔軟性が持っていることが分かった。例えば、10 次元の Bernstein コピュラは  $[-0.818, 0.818]$  の相関を表せる事に対して、10 次元の B-spline コピュラは  $[-0.99, 0.99]$  の範囲の相関を表すことができる。

B-spline コピュラは Bernstein 基底関数の代わりに、もっと一般的な B-spline 基底関数を用いる形で、Bernstein コピュラの優良な性質も受け継ぐことが分かった。例えば、相関を最大にする B-spline コピュラは基底関数の数が無限大になる時に、Frechet-Hoeffding upper bound に達する。また、相関を最大にする B-spline コピュラは任意オーダの total positivity を持つ。これはもっとも強い正の相関を表す概念である。それは B-spline 基底関数は order-complete weak Tchebycheff system という性質を使えば、B-spline 関数行列の行列式は非負で B-spline コピュラのパラメータ行列の要素も非負なので、その積となる任意次元の B-spline コピュラも非負となることによって証明される。つまり：

$$\det(c^*(u_i, v_j)) = \sum_{n \geq k_1 \geq \dots \geq k_r \geq 1} \left( \prod_{i=1}^n q_{ki} \right) \underbrace{\det(\phi_{k_j, n}(u_i))}_{\geq 0} \underbrace{\det(\phi_{k_j, n}(v_i))}_{\geq 0} \quad \text{が非負である。}$$

また、B-spline コピュラの推定に関しては、Bernstein コピュラを推定するための EM アルゴリズムを B-spline コピュラの定義に求められる条件を満たせるように改めて作成し、新しいコピュラの EM アルゴリズム法を提案した。実データへ適用できることも確認した。

## 5. 主な発表論文等

〔雑誌論文〕 計4件（うち査読付論文 4件/うち国際共著 2件/うちオープンアクセス 3件）

1. 著者名 Xiaoling Dou, Shingo Shirahata, Hiroki Sugimoto	4. 巻 13(5)
2. 論文標題 Functional clustering of mouse ultrasonic vocalization data	5. 発行年 2018年
3. 雑誌名 PlosOne	6. 最初と最後の頁 1,26
掲載論文のDOI（デジタルオブジェクト識別子） <a href="https://doi.org/10.1371/journal.pone.0196834">https://doi.org/10.1371/journal.pone.0196834</a>	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -
1. 著者名 Xiaoling Dou, Satoshi Kuriki, Gwo Dong Lin, Donald Richards	4. 巻 -
2. 論文標題 Dependence Properties of B-Spline Copulas	5. 発行年 2019年
3. 雑誌名 Sankhya A	6. 最初と最後の頁 -
掲載論文のDOI（デジタルオブジェクト識別子） <a href="https://doi.org/10.1007/s13171-019-00179-y">https://doi.org/10.1007/s13171-019-00179-y</a>	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 該当する
1. 著者名 Gwo Dong Lin, Xiaoling Dou, Satoshi Kuriki	4. 巻 81
2. 論文標題 The bivariate lack-of-memory distributions	5. 発行年 2019年
3. 雑誌名 Sankhya A: The Indian Journal of Statistics	6. 最初と最後の頁 273,297
掲載論文のDOI（デジタルオブジェクト識別子） <a href="https://doi.org/10.1007/s13171-017-0119-1">https://doi.org/10.1007/s13171-017-0119-1</a>	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 該当する
1. 著者名 Xiaoling Dou	4. 巻 82
2. 論文標題 An investigation of a generalized least squares estimator for non-linear time series model	5. 発行年 2019年
3. 雑誌名 Scientiae Mathematicae Japonicae	6. 最初と最後の頁 29,44
掲載論文のDOI（デジタルオブジェクト識別子） <a href="http://www.jams.or.jp/scm/contents/e-2017-1/2017-8.pdf">http://www.jams.or.jp/scm/contents/e-2017-1/2017-8.pdf</a>	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

〔学会発表〕 計18件（うち招待講演 9件 / うち国際学会 14件）

1. 発表者名 Xiaoling Dou
2. 発表標題 EM algorithms for estimating B-spline copulas
3. 学会等名 Computational and Financial Econometrics (CFE 2019) & 12th International Conference of the ERCIM WG on Computational and Methodological Statistics (CMStatistics 2019) (国際学会)
4. 発表年 2019年

1. 発表者名 Xiaoling Dou
2. 発表標題 EM algorithms for estimating B-spline copula
3. 学会等名 New Developments in Statistics and its Applications (国際学会)
4. 発表年 2019年

1. 発表者名 Xiaoling Dou
2. 発表標題 Baker's distribution, Bernstein copula and B-spline copulas
3. 学会等名 日本数学会2019年度秋季総合分科会 (招待講演)
4. 発表年 2019年

1. 発表者名 Xiaoling Dou
2. 発表標題 Dependence Properties of B-Spline Copulas
3. 学会等名 28th South Taiwan Statistics Conference (招待講演) (国際学会)
4. 発表年 2019年

1. 発表者名 Xiaoling Dou
2. 発表標題 Testing for Granger Causality by Use of Box-Cox Transformations
3. 学会等名 Workshop on Causal Inference in Complex Marine Ecosystems (招待講演) (国際学会)
4. 発表年 2019年

1. 発表者名 Xiaoling Dou
2. 発表標題 Dependent structures of the B-spline copulas including the Bernstein ones
3. 学会等名 Statistical Methods and Models for Complex Data (招待講演) (国際学会)
4. 発表年 2019年

1. 発表者名 Xiaoling Dou
2. 発表標題 Dependent structures of the B-spline copulas
3. 学会等名 Mini Workshop on TDA, Time Series & Statistics (国際学会)
4. 発表年 2019年

1. 発表者名 Xiaoling Dou
2. 発表標題 A historical note on the Stirling and Eulerian numbers
3. 学会等名 Kinosaki Seminar on Data Science & Causality (国際学会)
4. 発表年 2019年

1. 発表者名 Xiaoling Dou, Hsien-Kuei Hwang
2. 発表標題 The Stirling and Eulerian numbers in the Edo Period
3. 学会等名 シンポジウム「融合する統計科学」(金沢大学 2018) (招待講演)
4. 発表年 2018年

1. 発表者名 Xiaoling Dou, Hsien-Kuei Hwang
2. 発表標題 The Stirling and Eulerian numbers in the Edo Period
3. 学会等名 ISI-ISM-ISSAS Joint Meeting (招待講演) (国際学会)
4. 発表年 2019年

1. 発表者名 Xiaoling Dou, Hsien-Kuei Hwang
2. 発表標題 A historical note on the Stirling and Eulerian numbers
3. 学会等名 Kinosaki International Semina 2019 (国際学会)
4. 発表年 2019年

1. 発表者名 Xiaoling Dou
2. 発表標題 A functional nonparametric unsupervised classification of mouse ultrasonic vocalization data
3. 学会等名 Ise-Shima International Sminar (国際学会)
4. 発表年 2017年

1. 発表者名 Xiaoling Dou
2. 発表標題 A generalized least square estimator for non-linear time series model
3. 学会等名 Keio International Symposium (国際学会)
4. 発表年 2017年

1. 発表者名 Xiaoling Dou
2. 発表標題 An investigation of Ochi's estimator
3. 学会等名 1st International Conference on Econometrics and Statistics(Ecosta 2017) (招待講演) (国際学会)
4. 発表年 2017年

1. 発表者名 Xiaoling Dou
2. 発表標題 A nonparametric functional clustering of mouse ultrasonic vocalization data
3. 学会等名 2017年度統計関連学会連合大会
4. 発表年 2017年

1. 発表者名 Xiaoling Dou
2. 発表標題 A nonparametric functional clustering of mouse ultrasonic vocalization data
3. 学会等名 ISI-ISM-ISSAS joint conference Tokyo 2017 (国際学会)
4. 発表年 2017年

1. 発表者名 Xiaoling Dou
2. 発表標題 B-spline copula and its estimation
3. 学会等名 The 10th ICOSA International Conference on Global Growth of Modern Statistics in the 21st Century (招待講演) (国際学会)
4. 発表年 2016年

1. 発表者名 Xiaoling Dou
2. 発表標題 Baker distribution, Bernstein copula and B-spline copula
3. 学会等名 9th Seminar of Data Science, Shiga University (招待講演)
4. 発表年 2016年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究協力者	栗木 哲  (Kuriki Satoshi)		
研究協力者	林 国棟  (Lin Gwo Dong)		
研究協力者	リチャーズ ドナルド  (Richards Donald)		