行政院國家科學委員會專題研究計畫　期中進度報告

循環相關係數的分配(1/2)

計畫類別： 個別型計畫
計畫編號： NSC94-2118-M-032-003-
執行期間： 94 年 08 月 01 日至 95 年 10 月 31 日
執行單位： 淡江大學數學系

計畫主持人： 林千代

計畫參與人員： 吳正新,黃彥龍

報告類型： 精簡報告
報告附件： 出席國際會議研究心得報告及發表論文
處理方式： 本計畫可公開查詢

中　華　民　國 95 年 5 月 10 日

行政院國家科學委員會專題研究計畫成果報告
The Distribution of the Circular Correlation Coefficient
循環相關係數的分配 (1/2)

計畫編號: NSC 94 - 2118- M- 032- 003
執行期限: 94年8月1日至95年7月31日

主持人: 林千代
執行機構及單位名稱: 淡江大學數學系
計畫參與人員: 吳正新及黃彥龍
執行機構及單位名稱: 淡江大學數學系

中文摘要

我們利用 Huffer and Lin (2001) 的演算法來計算虛無假設下循環相關係數檢定統計量的分配.

關鍵詞: 循環相關係數, 留間隔.

Abstract

A computing algorithm developed by Huffer and Lin (2001) is employed to evaluate the null distribution of the circular correlation coefficient detection statistic discussed in Pakula and Kay (1986).

Keywords: Circular correlation coefficient; Spacings.

1 Introduction

The serial correlation coefficient and circular correlation coefficient are statistics for detecting the presence of a signal in noise. The goal of this project is to compute the null distribution of the circular correlation coefficient detection statistic (see Pakula and Kay, 1986). Let $X_1, X_2, \ldots, X_n, Y_1, Y_2, \ldots, Y_n$ be the i.i.d. $N(0,1)$ random variables. Define $W_t = X_t + jY_t$ for $t = 1, \ldots, n$ where $j = \sqrt{-1}$. Then $W_1, W_2, \ldots, W_n$ is the complex white Gaussian noise. (This represents "noise" at time $t = 1, 2, \ldots, n$.) Moreover, the circular correlation coefficient is defined as

$$R_c = \frac{\sum_{t=1}^{n} W_t^* W_{t+1}}{\sum_{t=1}^{n} |W_t|^2}$$

where $W_{n+1} \equiv W_1$ (this makes it "circular"), W_t^* is the complex conjugate, and $|W_t|^2$ is the squared length. Therefore, the detection problem can be addressed: A signal is detected if $|R_c|^2 > z$, where z is chosen so that $P(|R_c|^2 > z) = .05$ under the null hypothesis of "white noise". Hence, our computational goal is to find $P(|R_c|^2 > z) = .05$ for arbitrary z under the null hypothesis.

2 Connection with Spacings

Our approach to this problem is mainly based on the connection with spacings via a hermitian form in complex normal random variables. Let $\boldsymbol{W} = (W_1, W_2, \ldots, W_n)'$. Then

$$\begin{aligned} R_c &= \frac{\sum_{t=1}^n W_t^* W_{t+1}}{\sum_{t=1}^n |W_t|^2} \\ &= \frac{(\boldsymbol{W}^*)'\boldsymbol{H}\boldsymbol{W}}{\sum_{t=1}^n |W_t|^2} \\ &\stackrel{d}{=} \frac{\sum_{t=1}^n e^{j2\pi t/n}|W_t|^2}{\sum_{t=1}^n |W_t|^2} \\ &\stackrel{d}{=} \frac{\sum_{t=1}^n e^{j2\pi t/n} Z_t}{\sum_{t=1}^n Z_t} \\ &\stackrel{d}{=} \sum_{t=1}^n e^{j2\pi t/n} S_t^{(n-1)}, \qquad (1) \end{aligned}$$

where $j = \sqrt{-1}$, $\boldsymbol{S}^{(n)} = (S_1, S_2, \ldots, S_{n+1})'$, and $Z_1, \ldots, Z_n$ are i.i.d. exponential random variables. Define $T = \sum_{i=1}^n Z_i$ and $\boldsymbol{Z} = (Z_1, Z_2, \ldots, Z_n$. Thus, (1) can be obtained by applying the fact that

$$\boldsymbol{Z}/T \stackrel{d}{=} \boldsymbol{S}^{(n)}.$$

One can separate R_c into its real and imaginary parts to have $R_c = (U, V)'$. The result

$$R_c \stackrel{d}{=} \sum_{t=1}^n e^{j2\pi t/n} S_t^{(n-1)}$$

then becomes

$$\begin{pmatrix} U \\ V \end{pmatrix} \stackrel{d}{=} \boldsymbol{A}\boldsymbol{S}^{(n-1)},$$

where $c_k = \cos(k \cdot 2\pi/n)$, $s_k = \sin(k \cdot 2\pi/n)$ for $k = 1, \ldots, n$, and $\boldsymbol{A}$ is the matrix

$$\begin{pmatrix} c_1 & c_2 & \ldots & c_n \\ s_1 & s_2 & \ldots & s_n \end{pmatrix}.$$

Hence, the problem in which we are interested can be further rephrased as to compute the probability $P(|\boldsymbol{A}\boldsymbol{S}^{(n-1)}|^2 > z)$.

The current problem differs in several ways from the ones tackled by the earlier other algorithms proposed by Huffer and Lin (1997b, 1999, 2001):

- The entries of $\boldsymbol{A}$ are irrational (except in some special cases).
- The region of interest is circular instead of being rectangular.
- The problem possesses rotational symmetry because

  $$|\boldsymbol{A}\boldsymbol{S}^{(n-1)}|^2 = |\boldsymbol{\Gamma}\boldsymbol{A}\boldsymbol{S}^{(n-1)}|^2$$

  for any 2×2 orthogonal matrix $\boldsymbol{\Gamma}$.

3 The Basic Recursion

For $r \geq 1$, let $\boldsymbol{A}$ be an $r \times (n+1)$ real matrix. Suppose $\boldsymbol{c} = (c_1, c_2, \ldots, c_{n+1})'$ satisfies $\sum_{i=1}^{n+1} c_i = 1$. Let $\boldsymbol{A}_i$ be the $r \times (n+1)$ matrix obtained by replacing the i-th column of $\boldsymbol{A}$ by $\boldsymbol{A}\boldsymbol{c}$. Then

$$P(\boldsymbol{A}\boldsymbol{S}^{(n)} \in B) = \sum_{i=1}^{n+1} c_i \, P(\boldsymbol{A}_i \boldsymbol{S}^{(n)} \in B) \qquad (2)$$

for any measurable set $B \subset \mathcal{R}^r$.

Based on the basic recursion, our algorithm reduces the distribution of $|R_c|^2$ down to simpler distributions, each involving only two spacings. Define (with $j = \sqrt{-1}$)

$$T_{k,n}(z) = P(|\boldsymbol{S}_1^{(n-1)} + e^j 2\pi k/n \boldsymbol{S}_2^{(n-1)}|^2 \leq z).$$

This can be expressed as a one-dimensional integral and computed numerically.

For brevity and write $T_{k,n}(z) \equiv T(k)$, suppressing the dependence on n and z.

The last page gives expressions for the cumulative distribution function of $|R_c|^2$ for signals of lengths $n = 9$ and $n = 15$. These are MAPLE expressions which can be read directly into MAPLE.

References

Choudhuri, N., Bayesian bootstrap credible sets for multidimensional mean functional. *Ann. Statist.* **26** (1998) 2104–2127.

Gasparini, M., Exact multivariate Bayesian bootstrap distributions of moments. *Ann. Statist.* **23** (1995) 762–768.

Huffer, F., Divided differences and the joint distribution of linear combinations of spacings, *J. Appl. Prob.* **25** (1988) 346–354.

Huffer, F., and C.T. Lin, Approximating the distribution of the scan statistic using moments of the number of clumps, *J. Amer. Stat. Assoc.* **92** (1997a) 1466–1475.

Huffer, F., and C.T. Lin, Computing the exact distribution of the extremes of sums of consecutive spacings. *Comput. Stat. Data Analysis* **26** (1997b) 117–132.

Huffer, F., and C.T. Lin, An approach to computations involving spacings with applications to the scan statistic. In *Scan Statistics and Applications* (edited by J. Glaz and N. Balakrishnan) (1999) 141–163. Birkhäuser, Boston.

Huffer, F. W. and Lin, C. T. (2001). Computing the joint distribution of general linear combinations of spacings or exponential variates, *Statistica Sinica* **11**, 1141–1157.

Lin, C.T., The computation of probabilities which involve spacings, with applications to the scan statistic, Ph.D. Dissertation (Dept. of Statistics, Florida State University, Tallahassee, FL, 1993).

Pakula, L., and Kay, S. M., Detection performance of the circular correlation coefficient receiver, *IEEE Trans. Acoustics, Speech, and Signal Processing* **34** (1986) 399–404.

Watson, G. S., On the joint distribution of the circular serial correlation coefficients, *Biometrika* **43** (1956) 161–168.

The cdf of $|R_c|^2$

For $n = 9$:

```
(2/9+44/9*cos(1/9*Pi)-2/3*cos(2/9*Pi))*T(1)
+(44/9*cos(2/9*Pi)+2/9-38/9*cos(1/9*Pi))*T(4)
+1/3*T(3)
+(-2/3*cos(1/9*Pi)+2/9-38/9*cos(2/9*Pi))*T(2);
```

For $n = 15$:

```
 (644/15*cos(1/5*Pi)+199/15)*T(3)
+(-644/15*cos(1/5*Pi)+521/15)*T(6)
+(-28*cos(1/15*Pi)-964/15*cos(2/15*Pi)
        -488/15*cos(1/5*Pi)+13/15)*T(2)
+(28*cos(2/15*Pi)+484/5*cos(1/5*Pi)
       -713/15-964/15*cos(1/15*Pi))*T(4)
+(964/15*cos(2/15*Pi)+28*cos(1/15*Pi)
         -624/5*cos(1/5*Pi)+223/15)*T(7)
+(-77/5+908/15*cos(1/5*Pi)
+964/15*cos(1/15*Pi)-28*cos(2/15*Pi))*T(1)
+1/5*T(5);
```

A typical term in conventional notation:

$$\left(\frac{2}{9} + \frac{44}{9}\cos\left(\frac{\pi}{9}\right) - \frac{2}{3}\cos\left(\frac{2\pi}{9}\right)\right) T(1)$$