**SOME BASIC PROPERTIES OF CROSS-CORRELATION FUNCTIONS OF N-DIMENSIONAL VECTOR TIME SERIES**

**BY**

**USORO, ANTHONY E.**

**Department of Mathematics and Statistics, Akwa Ibom State University, Mkpat Enin, Akwa Ibom State.**

**Email: toskila2@yahoo.com**

**ABSTRACT**: In this work, cross-correlation function of multivariate time series was the interest. The design of cross-correlation function at different lags was presented. is the matrix of the cross-covariance functions, and are the variances of and vectors respectively. Vector cross-correlation function was derived as . A statistical package was used to verify the vector cross correlation functions, with trivariate analysis as a special case. From the results, some properties of vector cross-correlation functions were established.

**KEY WORDS**: Vector time series, cross-covariance function and cross-correlation function.

1.INTRODUCTION

In statistics, the term cross-covariance is sometimes used to refer to the covariance corr(X,Y) between two random vectors and , (where In signal processing, the cross-covariance is often called cross-correlation and is a measure of similarity of two signals, commonly used to find features in an unknown signal by comparing it to a known one. It is a function of the relative time between the signals, and it is sometimes called the sliding dot product. In univariate time series, the autocorrelation of a random process describes the correlation between values of the process at different points in time, as a function of the two times or of the time difference. Let be some repeatable process, and be some point in time after the start of that process. ( may be an integer for a discrete-time process or a real number for a continuous-time process.) Then is the value (or realization) produced by a given run of the process at time . Suppose that the process is further known to have defined values for mean and variance for all times . Then the definition of the autocorrelation between times and is

,

where “E” is the expected value operator. It is required to note that the above expression is not well-defined for all time series or processes, because the variance may be zero. If the function is well-defined, its value must lie in the range [-1,1], with 1 indicating perfect correlation and -1 indicating perfect anti-correlation. If is a second-order stationary process then the mean and the variance are time-independent, and further the autocorrelation depends only on the difference between and : the correlation depends only on the time-distance between the pair of values but not on their position in time. This further implies that the autocorrelation can be expressed as a function of the time-lag, and that this would be an even function of the lag , which implies . This gives the more familiar form,

=

where and are time series process at lag k time difference. Hence, autocovariance coefficient at lag k, measures the covariance between two values and , a distance k apart. The autocorrelation coefficient is defined as the autocovariance at lag k divided by variance . The plot of against lag k is called the autocovariance function (, while the plot of against lag k is called the autocorrelation function (Box and Jenkins 1976).

In multivariate time series, cross-correlation or covariance involves more than one process. For instance, and are two processes of which could be cross-correlated with at lag k. The lag k value return by estimates the correlation between , Venables and Ripley (2002). Storch and Zwiers (2001) described cross-correlation in signal processing and time series. In signal processing, cross-correlation is a measure of similarity of two waveforms as a function of a time lag applied to one of them. This is also known as a sliding dot product or sliding inner-product. It is commonly used for searching a -duration signal for a shorter known feature. It also has application in pattern recognition, signal particle analysis, electron tomographic averaging, cryptanalysis and neurophysiology. In autocorrelation, which is the cross-correlation of a signal with itself, there is always a peak at a lag of zero unless the signal is a trivial zero signal. In probability theory and Statistics, correlation is always used to include a standardising factor in such a way that correlations have values between -1 and 1. Let represent a pair of stochastic process that are jointly wide sense stationary. Then the cross covariance given by Box et al (1984) is , where and are the means of and respectively. The cross-correlation function is the normalized cross-covariance function. Therefore,

Where and are the standard deviation of processes and respectively. If = for all t, then the cross-correlation function is simply the autocorrelation function for a discrete process of length n defined as { which known mean and variance, an estimate of the autocorrelation may be obtained as

for any positive integer k<n, Patrick (2005). When the true mean and variance are known, the estimate is unbiased. If the true mean, this estimate is unbiased. If the true mean and variance of the process are not known, there are several probabilities:

i. if and are replaced by the standard formulas for sample mean and sample variance, then this is a biased estimate.

ii. if n-k in the above formula is replaced with n, the estimate is biased. However, it usually has a smaller mean square error, Priestly (1982) and Donald and Walden (1993).

iii. if is stationary process, then the following are true

for all t,s and , where T=s-t, is the lag time or the moment of time by which the signal has been shifted. As a result, the autocovariance becomes

=, where represents the autocorrelation in the signal processing sense.

,Hoel (1984).

For and , the following properties hold:

1.

2.

3.

4.

Mardia and Goodall (1993) defined separable cross-correlation function as where is a positive definite matrix and is a valid correlation function. Goulard & Voltz (1992); Wackernage (2003); Ver Hoef and Barry (1998) implied that the cross- covariance function is

, for an integer 1 where is a full rank matrix and are valid stationary correlation functions. Apanasovich and Genton (2010) constructed valid parametric cross-covariance functions. Apanasovich and Genton proposed a simple methodology based on latent dimensions and existing covariance models for univariate covariance, to develop flexible, interpretable and computationally feasible classes of cross-covariance functions in closed forms. They discussed estimation of the models and performed a small simulation study to demonstrate the models. The interest in this work is to extend cross-correlation functions beyond a-two variable case, present the multivariate design of vector cross-covariance and correlation functions and therefore establish some basic properties of vector cross-correlation functions from the analysis of vector cross-correlation functions.

2. THE DESIGN OF CROSS-COVARIANCE CROSS-CORRELATION FUNCTIONS

The matrix of cross-covariance functions is as shown below:

. .

. .

. .

The above matrix is a square matrix, and could be reduced to the form,

From the above cross-covariance matrix, ,

Where, is the matrix of the cross-covariance functions, and

are the variances of and vectors respectively. Given the above matrix, it is required to note that two vector processes and can only be cross-correlated at different lags, if either of or of has a fixed value zero. That is can be cross-correlated with , or can be cross-correlated with

3. ANALYSIS OF THE CROSS-CORRELATION FUNCTIONS

Given two processes , is the cross-correlation between , while, is the cross-correlation between , Box et al (1984). In this work, three vector processes are used to carry out the cross-correlation analysis. For the following results were obtained with a software:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Lag k |  |  |  |  |  |  |
| -4 | -0.172 | 0.572 | -0.102 | 0.643 | -0.427 | -0.350 |
| -3 | -0.517 | 0.405 | -0.501 | 0.410 | -0.076 | 0.042 |
| -2 | -0.611 | 0.098 | -0.662 | 0.067 | 0.327 | 0.399 |
| -1 | -0.605 | -0.290 | -0.674 | -0.303 | 0.659 | 0.697 |
| 0 | -0.506 | -0.506 | -0.578 | -0.578 | 0.900 | 0.900 |
| 1 | -0.290 | -0.605 | -0.303 | -0.674 | 0.697 | 0.659 |
| 2 | 0.098 | -0.611 | 0.067 | -0.662 | 0.399 | 0.327 |
| 3 | 0.405 | -0.517 | 0.410 | -0.501 | 0.042 | -0.076 |
| 4 | 0.572 | -0.172 | 0.643 | -0.102 | -0.350 | -0.427 |

From the above analysis, the following properties were established:

1. a.

b.

1. a.

b.

1. a.

b.

CONCLUSION:

The motivation behind this research work was to carry out cross-correlation functions of multivariate time series. Ordinarily, cross-correlation compares two series by shifting one of them relative to the other. In the case of and variables, the variable may be cross-correlated at different lags of , and vice versa. In this work, and were used as vector time series, using trivariate as a special case of multivariate cross-correlation functions. The design of the cross-covariance functions has been displayed in a matrix form. Estimates obtained revealed some basic properties of vector cross-correlation functions.

REFERENCES:

1. Apanasovich, Tatiyana V. And Genton, Marc G. (2010): Cross-Covariance functions for multivariate random fields based on latent dimensions. *Biometrika*, 97,1,pp.15-30.

2. Box, G. E. P. And Jenkins, G. M. (1976): Time Series Analysis; Forecasting and Control 1st Edition, Holden-day, san Francisco.

3. Box, G. E. P., G. M. Jenkins, and G. C. Reinsel (1994): Time Series Analysis: Forecasting and Control. 3rd ed. Upper Saddle River, NJ: Prentice-Hall

4. Goulard, M. & VolTz, M. (1992). Linear Coregionalization Model: tools for estimation and choice of croo-variogram matrix. Math. Geol. 24, 269-82.

5. Hoel, P. G(1984): Mathematical Statistics, Wiley, New York.

6. Mardia, K. V. And Goodall, C. R. (1993): Spartial-temporal analysis of multivariate environmental monitoring data. Multivariate environmental Statistics, Northytolland Series in Statistics & Probability 6. Amsterdam, North-Holland.

7. Pattrick F. Dunn(2005): Measurement and Data Analysis for Engineering and Science, New York: McGraw-Hill. ISBN 0-07-282538-3

8. Percival, Donald B.; Andrew T. Walden (1993). Spectral Analysis for Physical Applications: Multitaper and Conventional Univariate Techniques. Cambridge University Press. PP. 190-195. ISBN 0-521-43541-2

9. Priestly, M. B (1982): Spectral Analysis and Time Series. Academic press, London.

10. Storch, H. F. and W Zwiers (2001): Statistical Analysis in climate research. Cambridge University Press. ISBN 0-521-01230-9.

11. Venables, W. N. And Ripley, B. D. (2002): Auto and Cross-Covariance and Correlation Function Estimatiom. Modern Applied Statistics, Fourth Edition, Springer-Verlag.

12. Ver Hoef, J. M. & Barry, R. P. (1998). Constructing and Fitting models for Cokriging and multivariate spatial prediction. *Journal Statistics Plan Inference*; 69.275-94.

13. Wackernagel, H. (2003): Multivariate Geostatistics: An introduction with Application. 2nd ed Berlin, Springer.