九州大学学術情報リポジトリ
Kyushu University Institutional Repository

ON BIAS REDUCTION IN ESTIMATION OF THE MAGNITUDE-SQUARED COHERENCE FUNCTION

Lee, Jae C.
Department of Statistics, Korea University

https://doi.org/10.5109/13350

出版情報：Bulletin of informatics and cybernetics. 20 (3/4), pp.107-114, 1983-03. Research Association of Statistical Sciences
バージョン：
権利関係：

Bulletin of Informatics and Cybernetics Vol. 20, No. 3~4, 1983

ON BIAS REDUCTION IN ESTIMATION OF THE MAGNITUDE-SQUARED COHERENCE FUNCTION*

By

Jae C. LEE**

Abstract

Magnitude-squared coherence (MSC) function between two stochastic processes measures the linear relationship between the processes. The usual estimators of MSC are biased. A Jackknife estimator of MSC is introduced to reduce the bias in estimation.

1. Introduction

A measure of linear relationship between two stochastic processes is given by the magnitude-squared coherence function (MSC) defined by

$$\rho=\frac{|S_{12}(f)|^2}{S_{11}(f)S_{22}(f)}$$

where $S_{12}(f)$ is the cross spectral density at frequency f between two zero-mean stationary random processes $X(t)$ and $Y(t)$ with auto spectra $S_{11}(f)$ and $S_{22}(f)$.

The usual estimator of ρ given by

$$\frac{|\hat{S}_{12}|^2}{\hat{S}_{11}\cdot\hat{S}_{22}},$$

where $\hat{S}_{11}$, $\hat{S}_{12}$ and $\hat{S}_{22}$ are estimates of S_{11}, S_{12} and S_{22} based on the whole sample, has the variance which does not decrease as the sample size increases. Hence this does not possess the usually desired property of sample statistic that the variance of estimate is inversely proportional to the sample size.

Current techniques in MSC estimation rely on partitioning the whole sample into n equal segments and averaging the segment-wise estimates of spectral and cross spectral densities to obtain a reduction in variance of the estimate. When the partitioning is disjoint, properties and approximate distribution function of the MSC estimator have been studied by many authors. Goodman (1957) derived the probability density function of the statistic under the Gaussian assumption, Carter and Nuttall (1972) showed the bias and variance of the estimate as functions of true MSC and the number of segments.

---

* This work was done during the author's visit to US Naval Coastal Systems Center, Panama City, FL, sponsored by US-ONR and ASEE, in the summer of 1981.

** Department of Statistics, Korea University, Seoul, Korea.

In many practical applications it has been suggested that overlapped segments in place of disjoint ones reduce both bias and variance of the estimate. Carter, Knapp and Nuttall (1973) investigated the propertes of overlapped estimator by simulation and showed an advantage of it over disjoint segments in terms of reduced bias and variance. The reduction in bias and variance actually depends on the true MSC and proportion of overlapping as well as the number of segments. The overlapping methods increase the computational cost as the percentage of overlap gets highter. Also the additional reduction in bias and variance seems very small beyond 65 percent of overlapping. For example, when the true MSC is 0.3 and the overlap is 50 percent, the variance is 55 percent of the non-overlapped estimator, and the bias is 50 percent as large. Their study included the cases of $\rho=0$ and $\rho=0.3$, where ρ is the true MSC. In a more recent study by Lugannani (1981) he claimed that, under some restrictions, approximate probability density function (asymptotic) depends on the effective number of degrees of freedom (EDF) and true MSC.

2. Lugannani's Results

Lugannani considered the sample MSC

$$\hat{\rho}=\frac{|\hat{S}_{12}(f)|^2}{\hat{S}_{11}(f)\hat{S}_{12}(f)}=\frac{\xi_3^2+\xi_4^2}{\xi_1\xi_2}$$

where $\xi_1=\hat{S}_{11}(f)$, $\xi_2=\hat{S}_{22}(f)$, $\xi_3=Re\{\hat{S}_{12}(f)\}$ and $\xi_4=Im\{\hat{S}_{12}(f)\}$. He also assumed that the overlapping of the type defined by

$$K=MQ$$

where K is the segment legnth, M is the shift and Q is an integer factor of overlapping. For example, $Q=2$ gives the 50 percent overlapping.

Unlike other previous studies where the Guassian assumption was given to the original processes he assumed that $\{\xi_1, \xi_2, \xi_3, \xi_4\}$ has a joint Gaussian distribution. The asymptotic distribution of $\hat{\rho}$ under these assumptions was given without details by

$$P_0(\hat{\rho})=\frac{1}{2\pi i}\int_{-i\infty}^{i\infty} D^{-1/2}(v)\{\lambda P^2(v)+R(v)\}\exp(-\lambda E(v))dv$$

where $P(v)$, $R(v)$, $D(v)$ and $E(v)$ are claimed to be rational functions of v that depend upon $\hat{\rho}$ and true MSC ρ but not explicitly upon M and Q. He claims that $P(\hat{\rho})$ depends on N and Q only through λ which is defined by

$$\lambda(N, Q)=\frac{6N^2Q}{2N(2Q^2+1)-Q(Q^2-1)}.$$

Hence the overlapped estimator with factor Q has $((N-1)Q+1)$ overlapped segments and EDF

$$\lambda'=\lambda\frac{6Q(NQ-Q+1)^2}{2N(NQ-Q+1)(2Q^2+1)-NQ(Q^2-1)}.$$

Since $\lambda'\leqq(3/2)\lambda$ with equality attained only in the limit as $N\to\infty$ and $Q\to\infty$, he claims

that bias and variance cannot be reduced to less than 2/3 of their non-overlap values (except the case of $\rho=0$ where it is 4/9).

For specially chosen values of N, M and Q the approximated bias and variance of overlap MSC estimator were compared to the simulated results of Carter, Knapp and Nuttall (1973) to see the closeness of approximation.

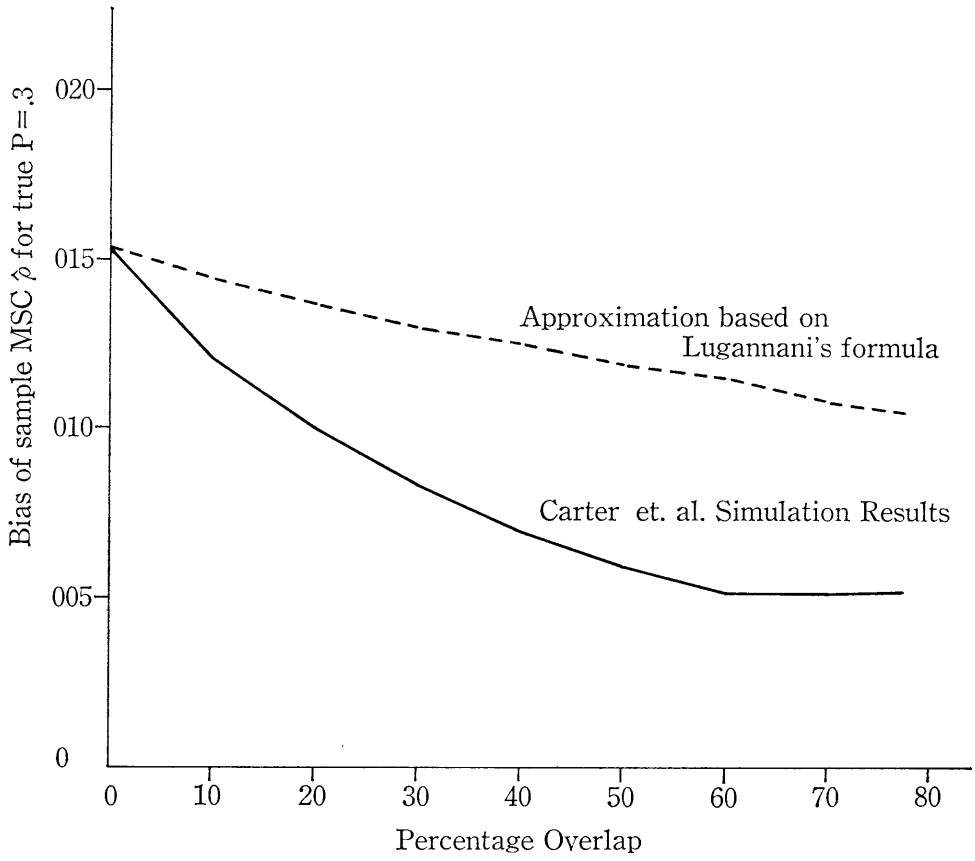

Fig. 1

It seems that the rate of convergence may be slower perhaps than what is desired to be practical for ordinary sample sizes. However, it is of great intrest since his result is the first of its kind under overlapping. If the rate of convergence is reasonable for practical cases, Lugannani's result will be very useful in approximating the bias and variance of sample MSC for construction of its desired confidence interval. Also the use of table prepared by Amos and Koopmans (1963) can be extended to overlap cases with proper equivalent degrees of freedom.

3. Jackknife Estimator

In most of the previous studies the Gaussian assumption was made in order to derive an approximate distribution of the sample MSC. There has not been a study on the robustness of estimate, i. e. sensitivity to departure of the given assumption.

It is our desire to find a practical method to establish a confidence interval of a sample MSC. A method which has been successful in reducing bias in ordinary sample statistics and asympotically robust is Jackknife. (or generalized Jackknife estimator.)

In this section we will briefly review the properties of these estimators and their asymptotic distributions. In the following section Jackknife MSC estimator will be formulated.

Let $\hat{\theta}$ be an estimator defined on the random sample $X_1, X_2, \cdots, X_n$. Partition this sample into N subsets of size M so that $NM=n$, and form a new random sample by deleting a subset of size M from the original sample. Then we define the estimator $\hat{\theta}^i$ to be the estimator $\hat{\theta}$ defined on the sub-sample which arises when the ith subset of size M has been deleted.

Now define pseudovalues of the Jackknife by

$$J_i(\hat{\theta})=N\hat{\theta}-(N-1)\hat{\theta}^i, \qquad i=1, \cdots, N.$$

Then the Jackknife estimator is defined to be

$$\begin{aligned} J(\hat{\theta}) &= \frac{1}{N}\sum_{i=1}^{N} J_i(\hat{\theta}) \\ &= N\hat{\theta}-(N-1)\bar{\hat{\theta}}^i \end{aligned}$$

where $\bar{\hat{\theta}}=\frac{1}{N}\sum_{i=1}^{N}\hat{\theta}^i$, i. e. the average of segment-wise estimators.

It is easier to understtand the Jackknife as a special case of the Generalized Jackknife which is defined as follows. (See Gray and Schucany (1972).) Let $\hat{\theta}_1$ and $\hat{\theta}_2$ be estimators for parameter θ. Then for any $R\neq 1$ we define the generalized Jackknife $G(\hat{\theta}_1, \hat{\theta}_2)$ by

$$G(\hat{\theta}_1, \hat{\theta}_2)=\frac{\hat{\theta}_1-R\hat{\theta}_2}{1-R}.$$

Clearly, if $\hat{\theta}_1$, and $\hat{\theta}_2$ are consistent, then $G(\hat{\theta}_1, \hat{\theta}_2)$ is also a consistent estimator of θ. To see the bias elimination property of G, suppose that we have $\hat{\theta}_1$, and $\hat{\theta}_2$ satisfying the following.

$$E(\hat{\theta}_k)=\theta+b_k(n, \theta), \qquad k=1, 2$$

where $b_k(n, \theta)\neq 0$. Now take

$$R=\frac{b_1(n, \theta)}{b_2(n, \theta)}\neq 1.$$

Then,

$$E(G(\hat{\theta}_1, \hat{\theta}_2))=\frac{E(\hat{\theta}_1)-R\cdot E(\hat{\theta}_2)}{1-R}$$

$$=\frac{\theta(1-R)+b_1(N,\ \theta)-R\cdot b_2(n,\ \theta)}{1-R}$$

$$=\theta\,.$$

This shows that the generalized Jackknife estimator can remove bias when we know the biases of both $\hat{\theta}_1$ and $\hat{\theta}_2$.

Further in a more general assumption of biases $G(\hat{\theta}_1,\ \hat{\theta}_2)$ removes, even though not all, some part of bias. If

$$E(G(\hat{\theta}_k)=\theta+\sum_{i=1}^{\infty} b_{ki}(n,\ \theta)\,,\quad k=1,\ 2\,,\qquad \text{and}\ \ R=\frac{b_{11}(n,\ \theta)}{b_{21}(n,\ \theta)}\,,$$

then

$$E(G(\hat{\theta}_1,\ \hat{\theta}_2))=\theta+\frac{\sum_{i=2}^{\infty} b_{1i}(n,\ \theta)-R\sum_{i=2}^{\infty} b_{2i}(n,\ \theta)}{1-R}\,.$$

It shows that in many cases, although $G(\hat{\theta}_1,\ \hat{\theta}_2)$ is not unbiased, it does contain less bias than either $\hat{\theta}_1$ or $\hat{\theta}_2$.

Now returning to ordinary Jackknife $J(\hat{\theta})$, we consider $\hat{\theta}_1=\hat{\theta}$ and

$$\hat{\theta}_2=\frac{1}{N}\sum_{i=1}^{N}\hat{\theta}^i=\bar{\hat{\theta}}^i\,.$$

Pick $R(N)=(N-1)/N$ and suppose

$$E(\hat{\theta})=\theta+\sum_{i=1}^{\infty}\frac{a_i}{n^i}=\theta+\sum_{i=1}^{\infty}\frac{a_i}{N^iM^i}\,,$$

where a_i may be function of θ but not of n. Then

$$E(\bar{\hat{\theta}}^i)=\theta+\sum_{i=1}^{\infty}\frac{a_i}{(n-m)^i}=\theta+\sum_{i=1}^{\infty}\frac{a_i}{(N-1)^iM^i}$$

and hence

$$E(J(\hat{\theta}))=\theta-\frac{a_2}{M^2N(N-1)}-\frac{a_3(2N-1)}{M^3N^2(N-1)^2}-\cdots$$

$$=\theta-\frac{a_2}{n(n-M)}-\frac{a_3(2n-M)}{n^2(n-M)^2}-\cdots$$

Thus the bias in $J(\hat{\theta})$ is of order n^{-2} while that of $\hat{\theta}$ is of order n^{-1} when $a_1\neq 0$. Moreover, if $a_2=a_3=\cdots=0$, as in the previous example it is unbiased.

Under the assumption that the pseudo-values are independent, the statistic

$$\frac{(J(\hat{\theta})-\theta)\sqrt{N(N-1)}}{\sqrt{\sum_{i=1}^{N}[J_i(\hat{\theta})-J(\hat{\theta})]^2}}\xrightarrow{\mathcal{L}}N(0,\ 1)\qquad \text{as}\ (N\to\infty)$$

where the convergences in law.

For practical use in construction of confidence intervals or hypothesis testing, we may work with

$$\frac{(J(\hat{\theta})-\theta)\sqrt{N(N-1)}}{\sqrt{\sum_{i=1}^{N}[J_i(\hat{\theta})-J(\theta)]^2}}\cdot(C)$$

as if it were distributed as student t with $(N-1)$ degrees of freedom. The factor C will be discussed later with a specific estimator. Hence the Jackknife estimator can be used to construct its approximate confidence interval with much less assumption on the distribution of sample.

4. Jackknife Estimator of MSC

Let $X_1(t)$ and $X_2(t)$ be real processes which are jointly wide sense stationary with power spectral densities $S_{jl}(f)$, $j, l=1, 2$. Let $\hat{S}_{jl}(f)$ be the estimates of S_{jl} based on discrete fourier transforms. The discrete transforms are

$$X_j(f, n)=\sum_{k=0}^{K-1} x_j(k+(m-1)K)\cdot\exp(-2\pi ifk)$$

$j=1, 2$ and $n=1, 2, \cdots, n_d$, where K is the length of a segment and n_d is the number of segments.

Then

$$\hat{S}_{jl}(f)=\frac{1}{n_d k}\sum_{n=1}^{n_d} X_j(f:n)\cdot X_l^*(f:m);\ j, l=1, 2,$$

and

$$\hat{\rho}(f)=\frac{|\hat{S}_{12}(f)|^2}{\hat{S}_{11}(f)\hat{S}_{22}(f)}.$$

Hence, $\hat{\rho}(f)$ is an estimator of $\rho(f)$ based on all n_d segments of the sample. We define $\hat{\rho}^n(f)$ to be the estimator $\hat{\rho}(f)$ defined on the subsample which arises when the nth segment of length K has been deleted. That is,

$$\hat{S}_{jl}^n(f)=\frac{1}{(n_d-1)K}\sum_{h\neq n} X_j(f:h)\cdot X_l^*(f:h);\qquad j, l=1, 2$$

where the sum is over $h=1$ to $h=n_d$ except $h=n$, and

$$\hat{\rho}^n(f)=\frac{|\hat{S}_{12}^n(f)|^2}{\hat{S}_{11}^n(f)\hat{S}_{22}^n(f)}.$$

Let the pseudo values of the Jackknife be

$$J_n(\hat{\rho})=n_d\hat{\rho}-(n_d-1)\hat{\rho}^n,\qquad n=1, 2, \cdots, n_d,$$

and the Jackknife estimator be

$$\begin{aligned}J(\hat{\rho})&=\frac{1}{n_d}\sum_{n=1}^{n_d} J_n(\hat{\rho})\\&=n_d\hat{\rho}-(n_d-1)\bar{\hat{\rho}}^n.\end{aligned}$$

If we assume an approximate bias of the estimator $\hat{\rho}$ to be

$$b(\hat{\rho} : n_d) = \frac{1}{n_d}(1-\rho)^2$$

and hence

$$b(\hat{\rho}^n : n_d - 1) = \frac{1}{n_d - 1}(1-\rho)^2 .$$

Then we have

$$\begin{aligned} E(J(\hat{\rho})) &= n_d E(\hat{\rho}) - (n_d - 1) E(\bar{\hat{\rho}}^n) \\ &= \rho . \end{aligned}$$

Hence the Jackknife estimator is unbiased in this case. However, since the better approximation of bias of $\hat{\rho}$ has more terms beyond what is given here, the Jackknife here reduces the bias of term $(1/n_d)$.

We can also find an approximate distribution of $J(\hat{\rho})$. That is

$$\frac{(J(\hat{\rho})-\rho)\sqrt{n_d(n_d-1)}}{\sqrt{\sum_{n=1}^{n_d}(J_n(\hat{\rho})-J(\hat{\rho}))^2}} \longrightarrow t_{(n_d-1)\,df}$$

as $n_d \to \infty$.

As in other approximations of the distribution it is assumed here that intraclass correlation is zero. It actually assumes that each pseudo value is independent of others. Since this assumption is not reasonable in our case, we can make some adjustment following Walsh (1947).

To correct the t statistic when each pair of variables in the sample have correlation, γ, the multiplying factor is

$$C = \sqrt{\frac{1-\gamma}{1+(n-1)\gamma}}$$

i.e. when the n values $X_1, X_2, \cdots, X_n$ represent a single observation of a multivariate Gaussian population for which each of n variables has mean μ, variance σ^2, and the correlation between two variables γ, then

$$\frac{(\bar{X}-\mu)\sqrt{n(n-1)}}{\sqrt{\sum_{i=1}^{n}(X_i-\bar{X})^2}} \cdot \sqrt{\frac{1-\gamma}{1+(n-1)\gamma}}$$

has the student t distribution with $(n-1)$ degrees of freedom.

Suppose that $X_1, X_2, \cdots$ are independently, identically distributed as $N(0, 1)$. Let

$$\hat{\theta}_1(X_1, \cdots, X_n) = \frac{\sum_{i=1}^{n} X_i}{n} \quad \text{and} \quad \hat{\theta}_2(X_{m+1}, \cdots, X_{n+m}) = \frac{\sum_{i=m+1}^{n+m} X_i}{n}, \quad m < n .$$

$\hat{\theta}_1$ and $\hat{\theta}_2$ are sample means based on $(n-m)$-many overlapped sample points. Then $\mathrm{Corr}(\hat{\theta}_1, \hat{\theta}_2) = (1/n) - (m/n^2)$. Since $\hat{\rho}^{n_1}$ and $\hat{\rho}^{n_2}$ share (n_d-2)-many $\hat{S}_{jl}$'s, it can be considerd as an overlapping with $m=2$. Hence the approximate correlation between $\hat{\rho}^{n_1}$ and $\hat{\rho}^{n_2}$ $(n_1 \neq n_2)$ is

$$\mathrm{Corr}\,(\hat{\rho}^{n_1},\ \hat{\rho}^{n_2})=\frac{n_d-3}{(n_d-1)^2}\ .$$

This gives the factor in Walsh

$$C=\sqrt{\frac{1-\gamma}{1+(n-1)\gamma}}=\sqrt{\frac{n_d^2-3n_d+4}{2n_d^2-6n_d+4}}\simeq\frac{1}{\sqrt{2}}$$

for large n_d.

References

[1] AMOS, D. E. and KOOPMANS, L. H.: *Tables of the Distribution of the Coefficient of Coherence for Stationary Bivariate Gaussian Processes*, Sandia Corp., Mono. SCR-483 (1963).

[2] CARTER, G. C., KNAPP, C. H. and NUTTALL, A. H.: *Estimation of the Magnitude-Squared Coherence Function Via Overlapped Fast Fourier Transform Processing*, IEEE Trans. Audio Electroacoust. **AU-21**, (1973), 337-344.

[3] CARTER, G. C. and NUTTALL, A. H.: *Statistics of the Estimate of Coherence*, Proc. IEEE (Lett.), **60**, (1972), 465-466.

[4] GOODMAN, N. R.: *On the Joint Estimation of the Spectra, Cospectrum, and Quadrature Spectrum of a Two-Dimensional Stationary Gaussian Process*, New York Univ., NY, Sci. Paper 10 (AD 134919), (1957).

[5] GRAY, H. L. and SCHUCANY, W. R.: *The Generalized Jacknife Statistics*, Marcel Dekker, Inc., NY. (1972).

[6] LUGANNANI, R.: *Distribntion of the Sample Magnitude-Squared Coherence Obtained Using Overlapped Fourier Transforms*, Proc. 1981 IEEE, (1981), 1243-1246.

[7] WALSH, J. E.: *Concerning the Effect of Intraclass Correlation on Certain Significance Tests*, Ann. Math. Stat. **18**, (1947), 88-96.

Communicated by Ch. Asano
Received November 22, 1982