UNBIASED MINIMUM VARIANCE ESTIMATION OF CORRELATION FUNCTIONS OF RANDOM SIGNALS

Chong–Yung Chi and Wu–Ton Chen

Department of Electrical Engineering
National Tsing Hua University
Hsinchu, Taiwan, Republic of China

ABSTRACT

In this paper, we present the unbiased minimum variance estimation of the correlation function, $r_{xx}(k)$, of a wide–sense stationary random signal x(k). However, the obtained theoretical minimum variance estimator, $\tilde{r}_{xx}(k)$, for $r_{xx}(k)$ is a function of not only x(k) but also the unknown $r_{xx}(k)$ and thus is not computable. Additionally, $\tilde{r}_{xx}(k)$ is not computationally efficient. We, therefore, propose a modified estimator, denoted $\hat{r}_{xx}(k)$, implemented by a 3–step computationally efficient algorithm without need of $r_{xx}(k)$. Finally, we show some simulation examples using Capon's minimum variance spectral estimator in which $\hat{r}_{xx}(k)$ is used for the correlation function estimate to indicate that $\hat{r}_{xx}(k)$ leads to very good results on spectral estimation.

I. INTRODUCTION

Various signal processing algorithms are designed based on the correlation function $r_{xx}(k)$ of a given set of random data x(0), x(1), ..., x(N–1), in order to extract the information of interest from data. For instance, $r_{xx}(k)$ is needed in linear prediction coding in speech processing [1], beamforming in sonar array processing [2], direction–finding and low–angle tracking in radar array processing [2], linear predictive deconvolution [3] in seismology and spectral estimation [4]. Similar cases can also be found in the fields of imaging processing, adaptive signal processing, radio astronomy and oceanography. However, $r_{xx}(k)$ can only be estimated from data. The estimation accuracy of the used estimate of $r_{xx}(k)$ is surely very important to the performance of the associated signal processing algorithms. A well–known biased estimate of $r_{xx}(k)$ of a stationary random signal x(n) is given by

$$\tilde{r}^1_{xx}(k) = \begin{cases} \frac{1}{N} \sum_{n=0}^{N-1-k} x(n)\, x(n+k), & 0 \le k \le N-1 \\ \tilde{r}^1_{xx}(-k), & -N<k<0 \end{cases} \quad (1)$$

which was popularly used in various areas mentioned previously. The reason for this may be that $E[x(n)x(n+k)]$ is intuitively replaced by the time average of $x(n)\cdot x(n+k)$. Although $\tilde{r}^1_{xx}(k)(N/N-|k|)$ is an unbiased estimate of $r_{xx}(k)$, it is not a minimum variance estimate of $r_{xx}(k)$. There are also cases that an estimate of $r_{xx}(k)$, said $\tilde{r}^1_{xx}(k)$, is prewindowed and then used in the following signal processing such as Blackman–Tukey spectral estimation [4].

In this paper, we propose an unbiased minimum variance estimator $\tilde{r}_{xx}(k) = \tilde{r}_{xx}(-k)$ (see (2) below) for $r_{xx}(k)$ which linearly processes $x(n)\cdot x(n+k)$ for n=0, 1, ...,N–1–k, and k≥0. However, the obtained theoretical minimum variance estimator $\tilde{r}_{xx}(k)$ is a function of not only x(k) but also the unknown $r_{xx}(k)$, and thus is not computable. Additionally, it is not computationally efficient. We, therefore, propose a modified estimator, denoted $\hat{r}_{xx}(k)$, implemented by a 3–step computationally efficient algorithm without need of $r_{xx}(k)$. Then we show some simulation examples using Capon's minimum variance spectral estimator in which $\hat{r}_{xx}(k)$ is used for the correlation function estimate to indicate that $\hat{r}_{xx}(k)$ leads to very good results on spectral estimation. Finally, we briefly discuss the positive definiteness of $\hat{r}_{xx}(k)$ and draw some conclusions.

II. UNBIASED MINIMUM VARIANCE CORRELATION FUNCTION ESTIMATION

Assume that x(n) is a wide–sense stationary real Gaussian random process with zero mean. The desired unbiased minimum variance estimate, $\tilde{r}_{xx}(k)$, of the correlation function $r_{xx}(k)$ of x(n) is expressed as:

$$\tilde{r}_{xx}(k) = \begin{cases} \sum_{n=0}^{N-1-k} h_k(n)\, y_k(n), & 0 \le k \le N-1 \\ \tilde{r}_{xx}(-k) & -N<k<0 \end{cases} \quad (2)$$

where

$$y_k(n) = x(n)\cdot x(n+k). \quad (3)$$

Note that $\tilde{r}^1_{xx}(k)$ is a special case of (2) with $h_k(n) = 1/N$ for all k and n. $h_k(n)$ for $0 \le n \le N-1-k$ is to be determined such that $\tilde{r}_{xx}(k)$ is unbiased with minimum variance. $\tilde{r}_{xx}(k)$ can also be expressed as the following vector form:

$$\tilde{r}_{xx}(k) = \underline{h}_k'\, \underline{y}_k \quad (4)$$

where

$$\underline{y}_k = (y(0), y(1), \cdots, y(N-k-1))' \quad (5)$$

and

$$\underline{h}_k = (h_k(0), h_k(1), \cdots, h_k(N-k-1))'. \quad (6)$$

CH2825–8/90/0000–0281/$01.00

The unbiasedness requires

$$E[\tilde{r}_{xx}(k)] = \underline{h}_k' E[\underline{y}_k] = r_{xx}(k) (\underline{h}_k' \underline{1}) = r_{xx}(k) \quad (7)$$

where $\underline{1}=(1,1, \cdots,1)'$, which leads to the following constraint

$$\underline{h}_k' \underline{1} = 1. \quad (8)$$

From (2) and the assumption that x(k) is zero–mean Gaussian, one can show that the variance of $\tilde{r}_{xx}(k)$ is given by

$$\mathrm{var}[\tilde{r}_{xx}(k)] = E[(\tilde{r}_{xx}(k))^2] - (r_{xx}(k))^2 = \underline{h}_k' R_k \underline{h}_k \quad (9)$$

where R_k is an $(N-k)\times(N-k)$ symmetric Toeplitz matrix whose (i,j) component is given by

$$[R_k]_{ij} = (r_{xx}(i-j))^2 + r_{xx}(i-j-k)\, r_{xx}(i-j+k),\ 1\leq i,j\leq N-k. \quad (10)$$

Minimizing $\mathrm{var}[\tilde{r}_{xx}(k)]$ under the constraint given by (8) can be done by minimizing

$$J = \underline{h}_k' R_k \underline{h}_k + \lambda (\underline{h}_k' \underline{1} - 1) \quad (11)$$

where λ is the Lagrange multiplier. The optimal $\underline{h}_k$ can be easily shown to be

$$\underline{\tilde{h}}_k = \frac{1}{\underline{1}'\ R_k^{-1}\ \underline{1}} R_k^{-1} \underline{1}. \quad (12)$$

Substituting (12) back into (9) yields

$$\mathrm{Var}[\tilde{r}_{xx}(k)] = \frac{1}{\underline{1}'\ R_k^{-1}\ \underline{1}}. \quad (13)$$

Some observations regarding (12) are given in the following. First of all, when x(n) is white, R_k reduces to $2[r_{xx}(0)]^2 I$ for k=0 and $[r_{xx}(0)]^2 I$ for $k\geq 1$ where I is an identity matrix with a proper dimension. Thus, $\tilde{h}_k(n)=1/(N-k)$, or $\tilde{r}_{xx}(k)= [N/(N-k)]\ \tilde{r}^1_{xx}(k)$ for $k\geq 0$. Secondly, R_k^{-1} is both symmetric and persymmetric since R_k is symmtric Toeplitz. Thus, $\tilde{h}_k(n) = \tilde{h}_k(N-k-1-n)$. Thirdly, the dimension, $(N-k)\times(N-k)$ of R_k is large when N is large and k is small. Therefore, computing $\underline{\tilde{h}}_k$ using (12) is not practical. Finally, $\underline{\tilde{h}}_k$ is not computable because R_k is unknown.

Based on the previous observations associated with $\tilde{r}_{xx}(k)$, we propose a modified estimator, denoted $\hat{r}_{xx}(k)$ (see (18) below), implemented by a 3–step computationally efficient algorithm without need of $r_{xx}(k)$. From (10) we see that R_k can be expressed as

$$R_k = a_0 I + A \quad (14)$$

where

$$A = \begin{bmatrix} 0 & a_1 & a_2 & \dots & \\ a_1 & 0 & a_1 & a_2 & \\ a_2 & a_1 & & & a_2 \\ \vdots & & & & a_1 \\ & \dots & & a_2 & a_1 & 0 \end{bmatrix} \quad (15)$$

and

$$a_i = (r_{xx}(i))^2 + r_{xx}(k-i)\cdot r_{xx}(k+i). \quad (16)$$

When $|\lambda_i| \ll a_0$ for all the eigenvalues λ_i 's of the matrix A, the following approximation to R_k^{-1} is reasonable:

$$R_k^{-1} = \frac{1}{a_0} [I + \frac{A}{a_0}]^{-1} \simeq \frac{1}{a_0} [I - \frac{A}{a_0}]. \quad (17)$$

The proposed algorithm for obtaining $\hat{r}_{xx}(k)$ is given as follows:

S1: Prewhiten x(n) with an M–th order prediction error filter H(z). Let e(n) denote the prewhitened data.

S2: Estimate $\tilde{r}_{ee}(k)$ for k=0, 1, ..., N_e-1, using (4), (12) and (17) where $r_{xx}(i)$ is replaced by $\tilde{r}^1_{ee}(i)$ (see (1)) and $a_i=0$ for $i>N_a$.

S3: Compute

$$\hat{r}_{xx}(k) = [w(k)\cdot\tilde{r}_{ee}(k)] * r_{vv}(k) \quad (18)$$

where w(k) is a window function, $r_{vv}(k) = v(k) * v(-k)$ and v(k) is the impulse response of the inverse filter 1/H(z).

Notice that, in **S1**, we prewhiten the data so that the problem of estimating $r_{xx}(k)$ is converted into the problem of estimating $r_{ee}(k)$ which is much narrower than $r_{xx}(k)$. The parameters M, N_e and N_a must be chosen large enough for the reasonable approximation to R_k^{-1} given by (17). We then estimate $\tilde{r}_{ee}(k)$ in **S2** with $r_{ee}(k)$ replaced by $\tilde{r}^1_{ee}(k)$. Finally, we compute $\hat{r}_{xx}(k)$ from the windowed $\tilde{r}_{ee}(k)$. The reason for this is that the variance of $\tilde{r}_{ee}(k)$ is larger for a larger k because the total number of $e(n)\cdot e(n+k)$ used, N–k, is smaller for a larger k. The window w(k) such as $w(k) = (N-|k|)/N$ (a triangular window) is used to compensate for this effect.

III. COMPUTER SIMULATIONS

The simulations were performed as follows. we generated a pseudo white Gaussian sequence u(k) with unity variance and let it pass through an autoregressive moving average (ARMA) filter to provide data x(k). The data length was N=256. We then estimated $\hat{r}_{xx}(k)$ using the 3–step algoirthm presented in Section II where Burg's prediction error filter [4] with order M=10, $w(k)=(N-|k|)/N$, $N_a=70$ and $N_e = 50$ were used. Capon's minimum variance spectral estimator is known to compute the estimate, $\hat{P}_{mv}(f)$, of the power spectral density (PSD) of x(k) as follows:

$$\hat{P}_{mv}(f) = \frac{p}{\underline{v}'\ \hat{R}_{xx}^{-1}\ \underline{v}} \quad (19)$$

where $\underline{v}=(1\ \exp(j2\pi f)\ \cdots\ \exp(j2\pi f(p-1)))'$, and $\hat{R}_{xx}$ is an estimate of the p×p autocorrelation matrix

$$R_{xx}=\begin{bmatrix} r_{xx}(0) & r_{xx}(-1) & \dots\ r_{xx}[-(p-1)] \\ r_{xx}(1) & r_{xx}(0) & \dots\ r_{xx}[-(p-2)] \\ \vdots & \vdots & \vdots \\ r_{xx}(p-1) & r_{xx}(p-2) & \dots\ r_{xx}(0) \end{bmatrix}. \quad (20)$$

We then computed $\hat{P}_{mv}(f)$ with $r_{xx}(k)$ in R_{xx} replaced by $\hat{r}_{xx}(k)$ and p=50. The average of 20 estimates is shown together with the true PSD. Also, the 20 estimates are plotted in an overlaid fashion to indicate the variability of $\hat{P}_{mv}(f)$. Finally, we also compared our simulation results with the corresponding results when $\hat{R}_{xx}$ is given by

$$[\hat{R}_{xx}]_{ij} = \frac{1}{2(N-p)} \left[\sum_{n=p}^{N-1} x(n-i)\, x(n-j) + \sum_{n=0}^{N-1-p} x(n+i)\, x(n+j) \right] \quad (21)$$

which is used by the modified covariance method [4] for autoregressive (AR) spectral estimation and is, to the authors' knowledge, currently best for minimum variance spectral estmation.

A case of broadband spectrum and a case of narrowband spectrum taken from [4] were simulated and the simulation results are shown in Figures 1 and 2, respectively. The results shown in Figures 1a and 2a were obtained using $\tilde{r}^1_{xx}(k)$. The results shown in Figures 1b and 2b were obtained using $\hat{r}_{xx}(k)$. One can see, from Figures 1a and 1b, that the results are very comparable because they are unbiased with similar variances. However, from Figure 2, one can see that the results shown in Figure 2a are biased with a much larger variance than those associated with Figures 2b, whereas the results shown in Figure 2b are very good. We also performed the same simulations using $\hat{R}_{xx}$ given by (21). However, the results are quite similar to those shown in Figures 1b and 2b, and thus are not shown here. These simulation results also indicate that our $\hat{r}_{xx}(k)$ leads to very good results on spectral estimation. We also performed other simualtions for different parameters, $M>10$, $30<N_a<70$ and $30<N_e<70$. The results are very similar to those shown in Figures 1b and 2b because $M=10$ and $N_a=70$ and $N_e=50$ are large enough for a good approximation to R_k^{-1} (see (17)).

IV. DISCUSSION AND CONCLUSIONS

In this paper, we presented the unbiased minimum variance estimation of the autocorrelation function of a wide–sense stationary randon signal $x(k)$. The obtained estimate $\tilde{r}_{xx}(k)$ of the autocorrelation function $r_{xx}(k)$ is given by (4) where $\underline{h}_k=\underline{\tilde{h}}_k$ is given by (12). Unfortunately, $\tilde{r}_{xx}(k)$ is not computable since the unknown $r_{xx}(k)$ is needed in computing $\underline{\tilde{h}}_k$ and computing $\underline{\tilde{h}}_k$ using (12) is not practical from the computational point of view. We also proposed a modified estimator $\hat{r}_{xx}(k)$ given by (18) which is implemented by a 3–step computatioanlly efficient algorithm without need of $r_{xx}(k)$ described in Section II. One can show, from (18), that when $\tilde{r}_{ee}(k)$ as well as $w(k)$ are positive definite $\hat{r}_{xx}(k)$ is positive definite. With our

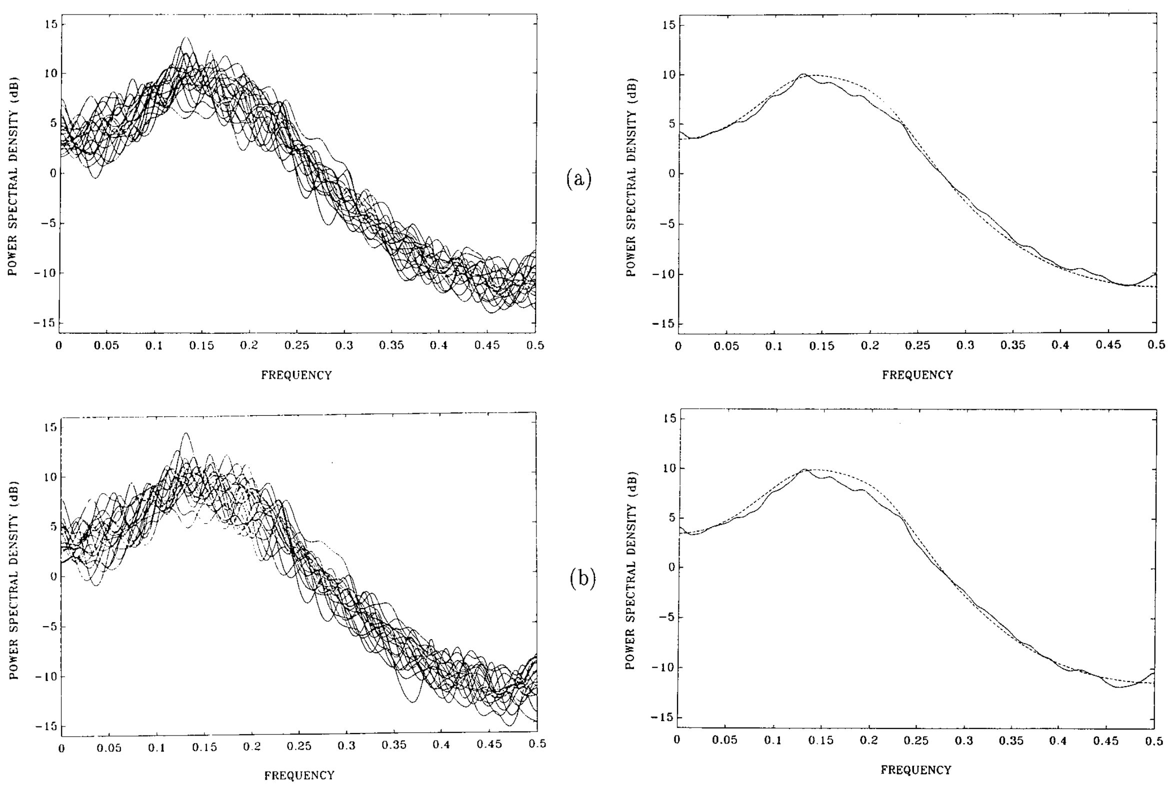

Figure 1. Overlaid realizations (left part), average of realizations (solid line) and the true PSD (dashed line) (right part) associated with (a) $\tilde{r}^1_{xx}(k)$ and (b) $\hat{r}_{xx}(k)$, respectively, for the broadband case.

experience, $\tilde{r}_{ee}(k)$ is almost surely positive definite when the order M of the prediction error filter, N_a and N_e (see S2) are large enough for a good approximation to the associated R_k^{-1} (see (17)). We then showed some simulation results using Capon's minimum variance power spectrum estimator $\hat{P}_{mv}(f)$ (see (19)) in which $\tilde{r}^1_{xx}(k)$ (see (1)) and $\hat{r}_{xx}(k)$ were used, respectively. The corresponding results associated with $\hat{R}_{xx}$ given by (21) which, we believe, are currently best results, are quite similar to those associated with $\hat{r}_{xx}(k)$ and thus were not provided. These simulation results indicate that $\hat{r}_{xx}(k)$ is a very good estimator for $r_{xx}(k)$. Potentially, with an appropriate selection of the window w(k), M, N_a, and N_e, our 3–step algorithm could yield a $\hat{r}_{xx}(k)$ which leads to better results than other estimates of $r_{xx}(k)$ in the following associated signal processing such as minimum variance spectral estimation. The performance of applying $\hat{r}_{xx}(k)$ to other signal processing problems is under study. As a final remark, the corresponding results for the case of complex random signals can be similarly obtained.

ACKNOWLEDGEMENT

The research described in this paper was performaned at the Department of Electrical Engineering, National Tsing Hua University, Hsinchu, Taiwan, and was supported by the National Science Council under Grant NSC79–0404–E–007–18.

REFERENCES

[1] L. R. Rabiner and R. W. Schafer, Digital Processing of Speech Signals, Prentice–Hall, Englewood Cliffs, New Jersey, 1978.

[2] S. Haykin, Array Signal Processing, Prentice–Hall, Englewood Cliffs, New Jersey, 1985.

[3] M. Silvia and E. Robinson, Deconvolution of Geophysical Time Series in the Exploration for Oil and Natural Gas, Elsevier Scientfic Publishing Company, 1979.

[4] Steven M. Kay, Modern Spectral Estimation, Prentice–Hall, Englewood Cliffs, New Jersey, 1988.

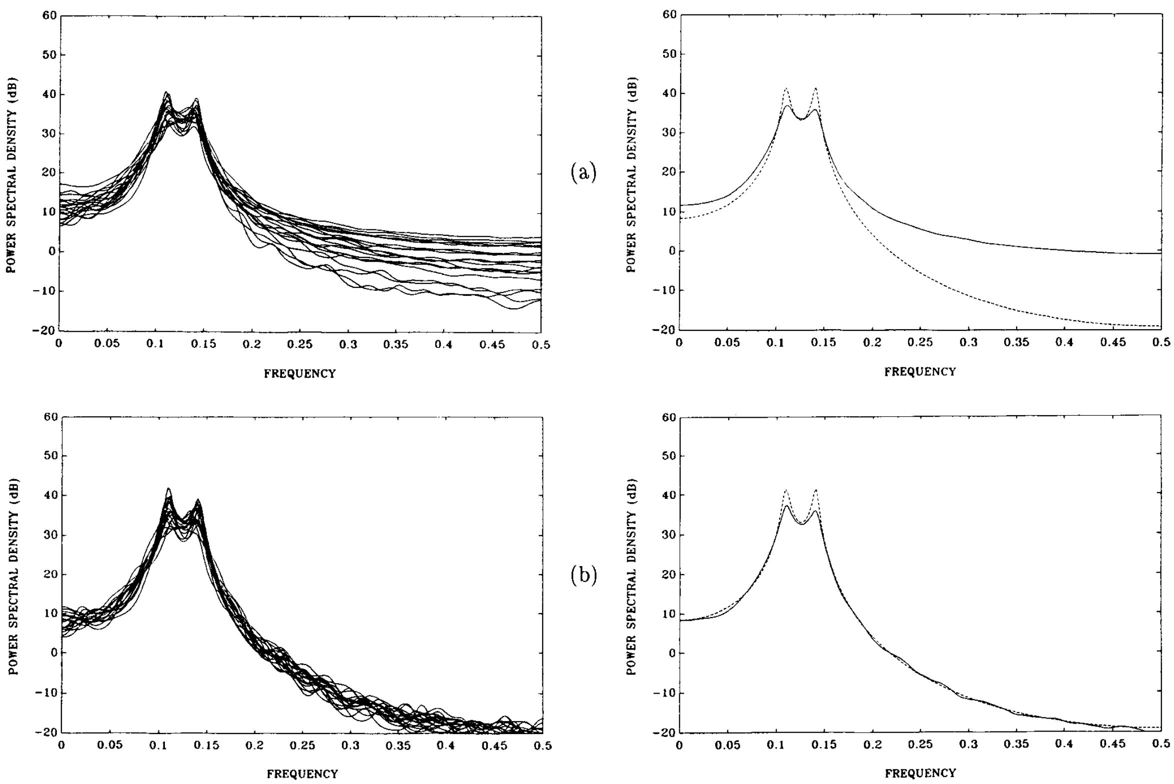

Figure 2. Overlaid realizations (left part), average of realizations (solid line) and the true PSD (dashed line) (right part) associated with (a) $\tilde{r}^1_{xx}(k)$ and (b) $\hat{r}_{xx}(k)$, respectively, for the narrowband case.