

Military Technical College
Kobry El-Kobba
Cairo, Egypt



12-th International Conference
on
Aerospace Sciences &
Aviation Technology

THE USE OF DISCRETE KARHUNEN-LOEVE TRANSFORM TO CHARACTERIZE THE SIGNAL AND NOISE PORTIONS IN MODERN STATISTICAL SIGNAL PROCESSING

Ashraf Mamdouh A. Aziz *

ABSTRACT

Frequently in communication and radar signal processing problems it is useful to expand a function as a linear combination of orthogonal basis functions. The orthogonality of the basis functions makes the representation efficient and mathematically convenient. In signal processing applications it also guarantees that the components of the signals with respect to the basis functions do not interfere with each other. For example, in communication problems the use of orthogonal functions permits several different signals to share the same transmission medium, such as a cable or fiber optic link. This paper shows that the use of orthogonal basis functions, in reconstructing an original signal from a received noisy observation, increases the signal-to-noise ratio and is counterpart for the nonstationary noise case to eliminating the out-of-band noise. The proposed signal representation is based on the discrete Karhunen-Loeve transform. The proposed filtering method reconstructs the original signal from a received noisy observation based on real data and does not depend on any prior statistical knowledge. Numerical results show that the proposed filtering method is very efficient.

KEYWORDS

Signal processing; Maximization of SNR; Optimum filtering; Modern statistical signal processing

* Associate Professor, Elec. Eng. Dept., Military Technical College, Cairo, Egypt

I. INTRODUCTION

A basic issue in the design of receivers is that of detecting a signal embedded in a background of additive noise. The purpose of such receivers is to establish the presence or absence of signal in noise. In order to enhance the strength of the signal relative to that of the noise, and thereby facilitate the detection process, a detection system usually consists of a predetection filter followed by a decision device. When the additive noise is white, that is, the power spectral density of the noise is constant, it turns out that the optimum solution to the prediction filter is a matched filter [1]. The matched filter is so-called because its characterization is matched to that of the signal component in the received signal. A matched filter is optimum in the sense that it maximizes the output signal-to-noise ratio (defined in a special way). The matched filter is useful in the design of digital communication systems where the concern is to enhance the received pulses. This can be done by maximizing the signal to noise ratio of the receiver output. Another approach of this basic optimization problem is based on probabilistic criterion directly related to performance ratings of most digital communication systems in which we are interested. In this case the filter optimization criterion is based on minimization of the average probability of error (involving the two types of errors: error when deciding a symbol 0 when symbol 1 is true and error when deciding a symbol 1 when symbol 0 is true). The probabilistic approach yields the so-called correlation receiver, which involves a correlation of the received signal with a stored replica of the transmitted signal. It is well known that the two receiver structures (matched filter and correlation receiver) are equivalent for the case of additive white Gaussian noise [2].

In these applications we are primarily interested in improving the ability to recognize a pulse signal in the presence of additive noise and not in preserving the fidelity of the pulse shape. It is also assumed that the signal and noise are both stationary with known power spectral densities. If the random signals are not stationary, then the power spectral densities are not defined and the solution is not so clear [3-5].

In this paper, we propose a method to reconstruct an original signal from a received noisy observation in case of non-stationary random signals, i.e. if the power spectral densities are not known. The proposed method is depending on the discrete Karhunen-Loeve transform [6-8]. The discrete Karhunen-Loeve transform expresses a random signal as a set of orthogonal basis functions. The paper is organized as follows. Diagonalization of the correlation matrix of a random signal by unitary transformation is addressed in section II. The discrete Karhunen-Loeve transform is presented in section III. The proposed filtering method and the most efficient representation of the random process, based on the discrete Karhunen-Loeve transform, is also presented in section III. Simulation results are considered in section IV. Section V contains conclusions.

II. DIAGONALIZATION OF THE CORRELATION MATRIX BY UNITARY TRANSFORMATION

A sequence $x(n)$ defined on the interval $0 \leq n \leq N-1$ can be represented as a vector \mathbf{X} with components $x(n)$. A real random vector \mathbf{X} will be denoted by the column vector

$$\mathbf{x} = \begin{bmatrix} x(0) \\ x(1) \\ \dots \\ \dots \\ x(N-1) \end{bmatrix}, \dots\dots\dots(1)$$

where the components of \mathbf{x} are real-valued random variables. Let $\psi(\mathbf{x})$ represent any quantity derived from the random vector \mathbf{x} . This quantity may be a scalar, vector, or matrix. The expectation of ψ is denoted by $E\{\psi(\mathbf{x})\}$ and defined by the operation

$$E\{\psi(\mathbf{x})\} = \int_{-\infty}^{\infty} \psi(\mathbf{x}) f_x(\mathbf{x}) d\mathbf{x}, \dots\dots\dots(2)$$

where $f_x(\mathbf{x})$ is the probability density function of the random vector \mathbf{x} . The correlation matrix of \mathbf{x} represents the complete set of second moments for the random vector and is defined by [9-11]

$$\mathbf{R}_x = E\{\mathbf{x} \mathbf{x}^T\}. \dots\dots\dots(3)$$

For a real random vector this matrix has the form

$$\mathbf{R}_x = \begin{bmatrix} E\{(x_0)^2\} & E\{x_0 x_1\} & \dots\dots\dots & E\{x_0 x_{N-1}\} \\ E\{x_1 x_0\} & E\{(x_1)^2\} & \dots\dots\dots & E\{x_1 x_{N-1}\} \\ \dots\dots\dots & \dots\dots\dots & \dots\dots\dots & \dots\dots\dots \\ \dots\dots\dots & \dots\dots\dots & \dots\dots\dots & \dots\dots\dots \\ E\{x_{N-1} x_0\} & E\{x_{N-1} x_1\} & \dots\dots\dots & E\{(x_{N-1})^2\} \end{bmatrix} \dots\dots\dots(4)$$

Note that the correlation matrix is a symmetric, that is

$$\mathbf{R}_x = \mathbf{R}_x^T. \dots\dots\dots(5)$$

It is useful to estimate the expected values when the density of the random vector is not known but instead there are a number of samples $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \dots\dots\dots, \mathbf{x}^{(K)}$ of the random vector \mathbf{x} . The estimate for the correlation matrix takes the form [7, 12-13]

$$\hat{\mathbf{R}}_x = \frac{1}{K} \sum_{k=1}^K \mathbf{x}^{(k)} (\mathbf{x}^{(k)})^T = \frac{1}{K} \mathbf{X}^T \mathbf{X}, \dots\dots\dots(6)$$

where

$$\mathbf{X} = \begin{bmatrix} - \mathbf{x}^{(1)T} - \\ - \mathbf{x}^{(2)T} - \\ \dots \\ \dots \\ - \mathbf{x}^{(K)T} - \end{bmatrix} \dots\dots\dots(7)$$

Sometimes it is desirable to transform a random vector x to another random vector x' whose components have the orthogonality property

$$E\{x'_k x'_l\} = 0; \quad k \neq l \dots\dots\dots(8)$$

This provides the advantages and simplicity of working with orthogonal random variables. Since the components are orthogonal the correlation matrix for the new random vector is a diagonal matrix. In this case the components of the random vector are uncorrelated.

If \mathbf{R}_x is the correlation matrix, an eigenvector \mathbf{e} and an eigenvalue λ satisfy the relation

$$\mathbf{R}_x \mathbf{e} = \lambda \mathbf{e} \dots\dots\dots(9)$$

That is, \mathbf{R}_x when regarded as a linear transformation, maps the eigenvector \mathbf{e} into a scaled version of itself. Because the matrix \mathbf{R}_x is symmetric it is possible to find N orthogonal vectors $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_N$ and a corresponding set of eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_N$ which are all real (some of them may have the same value). Let \mathbf{e}_l and \mathbf{e}_k be any two such eigenvectors from this orthonormal set. Then it follows that

$$\mathbf{e}_l^T \mathbf{R}_x \mathbf{e}_k = \lambda_k \mathbf{e}_l^T \mathbf{e}_k = \begin{cases} \lambda_k & \text{if } l = k \\ 0 & \text{if } l \neq k \end{cases} \dots\dots\dots(10)$$

Now define a matrix \mathbf{E} whose columns are the eigenvectors

$$\mathbf{E} = \begin{bmatrix} \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \mathbf{e}_1 & \mathbf{e}_2 & & \mathbf{e}_N \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \end{bmatrix} \dots\dots\dots(11)$$

From (10) it follows that

$$\mathbf{E}^T \mathbf{R}_x \mathbf{E} = \begin{bmatrix} \cdot \mathbf{e}_1^T \cdot \\ \cdot \mathbf{e}_2^T \cdot \\ \dots\dots\dots \\ \cdot \mathbf{e}_N^T \cdot \end{bmatrix} \mathbf{R}_x \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_N \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} = \begin{bmatrix} \lambda_1 & & 0 \\ & \lambda_2 & \\ & & \dots \\ 0 & & & \lambda_N \end{bmatrix} = \mathbf{\Lambda} \dots\dots\dots(12)$$

It is clear that $\mathbf{\Lambda}$ is the correlation matrix for a random vector \mathbf{x}' defined by

$$\mathbf{x}' = \mathbf{E}^T \mathbf{x} = \begin{bmatrix} -\mathbf{e}_1^T & - \\ -\mathbf{e}_2^T & - \\ \dots & \\ \dots & \\ -\mathbf{e}_N^T & - \end{bmatrix} \mathbf{x} \dots\dots\dots(13)$$

and since Λ is diagonal, the components of \mathbf{x}' are orthogonal. Further, the eigenvalues, which are the second moments of the components x'_k , are always positive (or possibly zero). Since the columns of \mathbf{E} are orthogonal, the matrix \mathbf{E}^T represents a unitary transformation, i.e.,

$$\mathbf{E}^T \mathbf{E} = \mathbf{E} \mathbf{E}^T = \mathbf{I} \dots\dots\dots(14)$$

The correlation matrix is

$$\mathbf{R}_{x'} = \mathbf{E}^T \mathbf{R}_x \mathbf{E} = \Lambda \dots\dots\dots(15)$$

Note that the eigenvalue problem (9) can be written in the alternative matrix form

$$\mathbf{R}_x \mathbf{E} = \Lambda \mathbf{E} \dots\dots\dots(16)$$

Equation (15) can be pre- and post-multiplied by \mathbf{E} and \mathbf{E}^T ; since \mathbf{E} is a unitary matrix the result is

$$\mathbf{R}_x = \mathbf{E} \Lambda \mathbf{E}^T \dots\dots\dots(17)$$

Further, taking the inverse of this equation and noting that

$$\mathbf{R}_x^{-1} = \mathbf{E} \Lambda^{-1} \mathbf{E}^T \dots\dots\dots(18)$$

This equation is useful for inverting the correlation matrix since Λ^{-1} is simply a real diagonal matrix with diagonal element $1/\lambda_j$. In addition since \mathbf{E} is orthonormal, the determinant and trace of \mathbf{R}_x are the same as the determinant and trace of Λ . Thus it follows that [14-18]

$$|\mathbf{R}_x| = |\Lambda| = \prod_{j=1}^N \lambda_j \dots\dots\dots(19)$$

$$\text{tr} \mathbf{R}_x = \text{tr} \Lambda \dots\dots\dots(20)$$

$$\text{tr} \Lambda = \sum_{j=1}^N \lambda_j \dots\dots\dots(21)$$

In most cases a computer program will be used to find the eigenvectors and eigenvalues. However it is worthwhile to review a procedure for doing it manually. To find the eigenvectors of a matrix, we can write (9) in the equivalent form

$$(\mathbf{R}_x - \lambda \mathbf{I}) \mathbf{e} = \mathbf{0} \dots\dots\dots(22)$$

In order for nontrivial solutions of this equation to exist, it is required that

$$|\mathbf{R}_x - \lambda \mathbf{I}| = 0 \dots\dots\dots(23)$$

Equation (22) is the characteristic equation corresponding to (9); it is a polynomial in λ whose roots are the eigenvalues of \mathbf{R}_x . Once these eigenvalues are found, (22) can be used to determine the corresponding eigenvectors.

III. DEVELOPMENT OF THE DKLT AND ITS OPTIMAL REPRESENTATION

Consider a segment of a random sequence $\{x(n); n = 0, 1, \dots, N - 1\}$. This segment can be expanded in any set of orthonormal basis functions $\phi_i(n)$ as

$$x(n) = k_1 \phi_1(n) + k_2 \phi_2(n) \dots + k_N \phi_N(n), \dots \dots \dots (24)$$

where the k_i are coefficients in the expansion. By orthonormal it is meant that

$$\sum_{n=0}^{N-1} \phi_i(n) \phi_j(n) = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases} \dots \dots \dots (25)$$

It follows from (24) and (25) that the coefficients are given by

$$k_i = \sum_{n=0}^{N-1} \phi_i(n) x(n) \dots \dots \dots (26)$$

It is desired to find a particular orthonormal set of functions such that

$$E\{k_i k_j\} = \begin{cases} \delta_j^2 & i = j \\ 0 & i \neq j \end{cases} \dots \dots \dots (27)$$

Therefore, if the random process has zero mean, which is the usual assumption, the coefficients are uncorrelated. Define the vector of coefficients

$$\mathbf{K} = \begin{bmatrix} k_1 \\ k_2 \\ \dots \\ \dots \\ k_N \end{bmatrix}, \dots \dots \dots (28)$$

and the matrix

$$\mathbf{\Phi} = \begin{bmatrix} | & & & | \\ | & & & | \\ \mathbf{\Phi}_1 & \mathbf{\Phi}_2 & & \mathbf{\Phi}_N \\ | & & & | \\ | & & & | \end{bmatrix}; \quad \mathbf{\Phi}_i = \begin{bmatrix} \phi_i(0) \\ \phi_i(1) \\ \dots \\ \dots \\ \phi_i(N-1) \end{bmatrix}, \quad i = 1, 2, \dots, N \dots \dots (29)$$

From (29) and (25), it is clear that the columns of $\mathbf{\Phi}$ are a set of orthonormal vector satisfying

$$\mathbf{\Phi}_i^T \mathbf{\Phi}_j = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}, \dots \dots \dots (30)$$

therefore, Φ is a unitary matrix. Equations (24) and (26) can be written using the matrix formulation

$$\mathbf{x} = \begin{bmatrix} | & & | \\ | & & | \\ \Phi_1 & \Phi_2 & \Phi_N \\ | & & | \\ | & & | \end{bmatrix} \begin{bmatrix} k_1 \\ k_2 \\ \dots \\ \dots \\ k_N \end{bmatrix} = \Phi \mathbf{K}, \dots\dots\dots(31)$$

and

$$\mathbf{K} = \begin{bmatrix} - - \Phi_1^T - - \\ - - \Phi_2^T - - \\ \dots \\ \dots \\ - - \Phi_N^T - - \end{bmatrix} \mathbf{x} = \Phi^T \mathbf{x}, \dots\dots\dots(32)$$

where since Φ is unitary, Φ^T is the inverse of Φ .

Recall from section II that if the φ_i are chosen as the eigenvectors of the correlation matrix, then the resulting k_i satisfy (27). The desired set of basis functions are thus determined by the eigenvectors of the correlation matrix. These functions are called eigenfunctions of the random process and satisfy this equation

$$\sum_{k=0}^{N-1} R_x(l-k) \varphi_i(k) = \lambda_i \varphi_i(l), \quad i = 1, 2, \dots, N. \dots\dots\dots(33)$$

The last equation has the matrix form

$$\mathbf{R}_x \varphi_i = \lambda_i \varphi_i. \dots\dots\dots(34)$$

With this choice of basis functions, the transformation (26) is the DKLT, and the representation (24) is called Karhunen-Loeve expansion for the random process [7]. The DKLT is also unique in that it is the only transformation that results in (27). To show this, consider any set of orthonormal basis functions (not necessarily those corresponding to the DKLT). Since the functions and the corresponding vectors are orthonormal, the requirement (27) can be written as

$$\begin{aligned} E\{k_i k_j\} &= E\{\varphi_i^T \mathbf{x} \mathbf{x}^T \varphi_j\} = \varphi_i^T E\{\mathbf{x} \mathbf{x}^T\} \varphi_j \\ &= \varphi_i^T \mathbf{R}_x \varphi_j = \begin{cases} \delta_j^2 & i = j \\ 0 & i \neq j \end{cases}. \dots\dots\dots(35) \end{aligned}$$

Consider

$$\varphi_i^T \mathbf{u}_j = \begin{cases} \delta_j^2 & i = j \\ 0 & i \neq j \end{cases}, \dots\dots\dots(36)$$

where

$$\mathbf{u}_j = \mathbf{R}_x \boldsymbol{\varphi}_j = \lambda_j \boldsymbol{\varphi}_j, \dots\dots\dots(37)$$

is whatever vector results from multiplying $\boldsymbol{\varphi}_j$ by the correlation matrix. Condition (37) states that \mathbf{u}_j is orthogonal to all of the $\boldsymbol{\varphi}_i$ for $i \neq j$. But since the $\boldsymbol{\varphi}_i$ form orthonormal set, \mathbf{u}_j must be equal to a constant times $\boldsymbol{\varphi}_j$, that is,

$$\mathbf{u}_j = \mathbf{R}_x \boldsymbol{\varphi}_j = \lambda_j \boldsymbol{\varphi}_j. \dots\dots\dots(38)$$

This states that $\boldsymbol{\varphi}_j$ must be an eigenvector and λ_j must be the corresponding eigenvalue. Since this is true for any choice of j , it follows that the unique set of $\{\boldsymbol{\varphi}_i\}$ that result in (27) are the eigenvectors of the correlation matrix. Then from equations (35) and (38) it also follows that

$$\delta_j^2 = \lambda_j. \dots\dots\dots(39)$$

An important property of the DKLT is that it is the most efficient representation of the random process if the expansion is truncated to use fewer than N orthonormal basis functions. If the random process consists of a signal in additive noise, it can turn out that by using a truncated expansion, a significant part of the noise is eliminated while most of the signal is kept intact.

Consider approximating the random sequence in terms of some number $M < N$ of basis functions as

$$\hat{x}(n) = \sum_{i=1}^M k_i \varphi_i(n); \quad M < N, \dots\dots\dots(40)$$

and define the error sequence

$$e(n) = x(n) - \hat{x}(n). \dots\dots\dots(41)$$

The problem is to find the appropriate set of basis functions to minimize the average energy in the error process

$$\varepsilon = E \left\{ \sum_{n=0}^{N-1} e^2(n) \right\}. \dots\dots\dots(42)$$

This quantity is referred to as the mean-square error. We can write

$$\mathbf{x} = \sum_{i=1}^N k_i \boldsymbol{\varphi}_i(n) = \sum_{i=1}^M k_i \boldsymbol{\varphi}_i(n) + \sum_{i=M+1}^N k_i \boldsymbol{\varphi}_i(n) = \hat{\mathbf{x}} + \boldsymbol{\varepsilon}, \dots\dots\dots(43)$$

where the vector quantities $\hat{\mathbf{x}}$ and $\boldsymbol{\varepsilon}$ defined here correspond to the sequence $\hat{x}(n)$ and $e(n)$. Equation (42) can then be expressed as

$$\varepsilon = E\{\boldsymbol{\varepsilon}^T \boldsymbol{\varepsilon}\} = E \left\{ \left(\sum_{i=M+1}^N k_i \boldsymbol{\varphi}_i^T \right) \left(\sum_{j=M+1}^N k_j \boldsymbol{\varphi}_j \right) \right\} = \sum_{i=M+1}^N E\{|k_i|^2\}, \dots\dots\dots(44)$$

where the last step follows because the $\boldsymbol{\varphi}_i$ satisfy the orthonormal condition of equation (30). Equation (35) can be used to write the mean-square error in compact form as

$$\varepsilon = \sum_{i=M+1}^N \boldsymbol{\varphi}_i^T \mathbf{R}_x \boldsymbol{\varphi}_i \quad \dots\dots\dots(45)$$

The problem is now to minimize (45) subject to the constraints

$$\boldsymbol{\varphi}_i^T \boldsymbol{\varphi}_i = 1; \quad i = M + 1, M + 2, \dots, N. \quad \dots\dots\dots(46)$$

This problem can be easily solved by using Lagrange multipliers. The gradient can be expressed as

$$\mathbf{R}_x \boldsymbol{\varphi}_i = \lambda_i \boldsymbol{\varphi}_i, \quad i = M + 1, \dots, N. \quad \dots\dots\dots(47)$$

Equation (45) can then be reexpressed using (47) as

$$\varepsilon = \sum_{i=M+1}^N \boldsymbol{\varphi}_i^T \mathbf{R}_x \boldsymbol{\varphi}_i = \sum_{i=M+1}^N \boldsymbol{\varphi}_i^T (\lambda_i \boldsymbol{\varphi}_i) = \sum_{i=M+1}^N \lambda_i. \quad \dots\dots\dots(48)$$

Since the eigenvalues of a correlation matrix are never negative, we should use the $N - M$ smallest eigenvalues of the correlation matrix in order to minimize the mean-square error. Since these smallest eigenvalues correspond to the terms left out of the expansion, the optimal basis functions for the truncated expansion of equation (40) correspond to the eigenvectors of \mathbf{R}_x with the M largest eigenvalues.

IV. SIMULATION RESULTS

In simulation, we consider the problem of DKLT expansion a signal, and its use in recovering data corrupted by noise. Let us consider a discrete finite-length deterministic signal which is observed in additive noise [19-20]. The observed sequence is

$$x(n) = s(n) + w(n), \quad \dots\dots\dots(49)$$

where $w(n)$ is the noise sequence. We assume the noise to be zero mean, white, independent on $s(n)$. The goal is to see how we can use the DKLT effectively to best reconstruct the original signal $s(n)$ from the observations $x(n)$. We assume that the signal sequence is

$$s(n) = \cos\left(2\pi \frac{n}{40}\right) - 0.5 \sin\left(2\pi \frac{n}{60}\right) + 0.3 \cos\left(2\pi \frac{n}{55} + \frac{\pi}{3}\right). \quad \dots\dots\dots(50)$$

Samples of 1024 data points (N=1024) of the noise sequence $w(n)$ and the signal sequence $s(n)$ are generated. They are added together to obtain the observed sequence $x(n)$. The signal to noise ratio is assumed to be -3 dB. The original sequence $s(n)$ (without adding noise) is shown in Fig.1. The noise sequence $w(n)$ is shown in Fig.2. The observed sequence $x(n)$ (signal plus noise) is shown in Fig.3.

The correlation between any two samples of a random process $x(n_1)$ and $x(n_2)$ is expressed by the correlation function (or the autocorrelation function)

$$R_x(n_1, n_2) = E\{x(n_1)x(n_2)\}. \quad \dots\dots\dots(51)$$

For a stationary random process, the correlation function can be defined as

$$R_x(\ell) = E\{x(n)x(n - \ell)\}. \quad \dots\dots\dots(52)$$

It is clear that the correlation matrix is completely specified by the correlation function for the random process.

The autocorrelation sequence $R_x(\ell)$ for the whole sequence is calculated and plotted in Fig.4. It is clear that it is an even function. Now we want to reconstruct the first 128 points of the original sequence $s(n)$ based on $x(n)$, $n = 1, 2, \dots, 128$ and the corresponding autocorrelation function $R_x(\ell)$, thus we obtain the correlation matrix \mathbf{R}_x . Also we calculate the eigenvectors and the eigenvalues of the correlation matrix. The first 128 points of the observed sequence $x(n)$ and its correlation function are shown in Fig.5 and Fig.6 respectively. The eigenvalues of \mathbf{R}_x are plotted in Fig.7. Note that most of them are very small. We reconstruct the sequence using the eigenvectors corresponding to the largest eigenvalues (we call them the largest eigenvectors). Fig.8 compares the original sequence $s(n)$ and the noisy estimated sequence. Fig.9 compares the original sequence $s(n)$ and the estimated sequence $\hat{s}(n)$ based on the largest 4-eigenvectors, i.e.,

$$\hat{x}(n) = \sum_{i=1}^4 k_i \varphi_i(n); \quad \dots\dots\dots(53)$$

where the $\varphi_i(n)$ correspond to the largest 4-eigenvectors .

It is clear that by using only 4-eigenvectors we can estimate the sequence with high accuracy despite of the noise sequence. Figures 10-14 compare the original sequence $s(n)$ and the estimated sequence $\hat{s}(n)$ based on the largest 5, 6, 9, 20, and 40 eigenvectors, respectively. It is clear that when we use more eigenvalues including some with small eigenvalues, the estimated sequence $\hat{s}(n)$ is not tracking the original sequence $s(n)$, i.e., the reconstructed sequence differs from the original sequence, especially, when we use more than 6 eigenvectors. Thus to reduce the mean square error in the approximation, we must use the smallest number of the dominant eigenvectors corresponding to the largest eigenvalues.

V. CONCLUSIONS

The problem of reconstructing an original signal from a received noisy observation in case of non-stationary random signals, i.e. if the power spectral densities of the random signals are not known, has been considered. The proposed filtering method is based on the discrete Karhunen-Loeve transform, which expresses a random signal as a set of orthogonal basis functions. It has been shown that a random signal can be represented in terms of the eigenvectors and the eigenvalues of the correlation matrix of the random signal. The most efficient representation which maximizes the signal-to-noise ratio is the representation that uses the dominant eigenvectors corresponding to the largest eigenvalues. Simulation results show that when we use only four or five eigenvectors we can estimate the original signal with high accuracy despite of the noise sequence. When we use more eigenvalues (more than five) including some with small eigenvalues, the reconstructed sequence differs from the original sequence. The proposed filtering method reconstructs the original

signal from a received noisy observation based on real data and does not depend on any prior statistical knowledge. Simulation results show that the proposed filtering method is very efficient.

REFERECNCES

- [1] Simon Haykin. An Introduction to Analog and Digital Communications. John Wiley & Sons. New York, (1989).
- [2] Simon Haykin. Adaptive Filter Theory. Prentice-Hall Inc., Englewood Cliffs, New Jersey, (1991).
- [3] William A. Grander. Introduction to Random Processes with Applications to Signals and systems. MacMillan, New York, (1986).
- [4] Harry L. Van Trees. Detection, Estimation, and Modulation Theory, Part I. John Wiley & Sons, New York, (1971).
- [5] Harry L. Van Trees. Detection, Estimation, and Modulation Theory, Part III. John Wiley & Sons, New York, (1971).
- [6] Harry L. Van Trees. Detection, Estimation, and Modulation Theory, Part II. John Wiley & Sons, New York, (1968).
- [7] Charles W. Therrien. Discrete Random Signals and Statistical Signal Processing. Prentice-Hall Inc., Englewood Cliffs, New Jersey, (1992).
- [8] Louis L. Scharf. Statistical Signal processing: Detection, Estimation, and Time Series Analysis. Addison-Wesley, Reading, Massachusetts, (1991).
- [9] Athanasios Papoulis. The Fourier Integral and Its Applications. McGraw-Hill, New York, (1962).
- [10] Alan V. Oppenheim and Ronald W. Schafer. Discrete-Time Signal processing. Prentice Hall, Inc., Englewood Cliffs, New Jersey, (1989).
- [11] Alan V. Oppenheim and Ronald W. Schafer. Digital Signal processing. Prentice-Hall, Inc., Englewood Cliffs, New Jersey, (1975).
- [12] Athanasios Papoulis. Probability, Random Variables, and Stochastic Processes, 3rd ed. McGraw-Hill, New York, (1991).
- [13] Car W. Helstrom. Probability and Stochastic Process for Engineers. MacMillan, New York, 2nd ed., (1991).
- [14] Leo Breiman. Probability and Stochastic Processes, With a View Toward Applications. The Scientific Press, Palo Alto, California, 2nd ed., (1986).
- [15] Alvin W. Drake. Fundamentals of Applied Probability Theory. McGraw-Hill, New York, (1991).
- [16] Howard Anton and Chris Rorres. Elementary Linear Algebra, 6th ed., John Wiley & Sons, New York, (1991).
- [17] Charles L. Lawson and Richard J. Hanson. Solving Least Squares Problems. Prentice-Hall Inc., Englewood Cliffs, New Jersey, (1974).
- [18] Athanasios Papoulis. Probability and Statistics. Prentice-Hall Inc., Englewood Cliffs, New Jersey, (1990).
- [19] Leon W. Couch. Digital and Analog Communication Systems, 4th ed. Prentice-Hall Inc., Upper Saddle River, New Jersey, (1993).
- [20] B. P. Lathi. Modern Digital and Analog Communication Systems. 2nd ed. Rinehart and Winston Inc., Orlando, Florida, (1989).

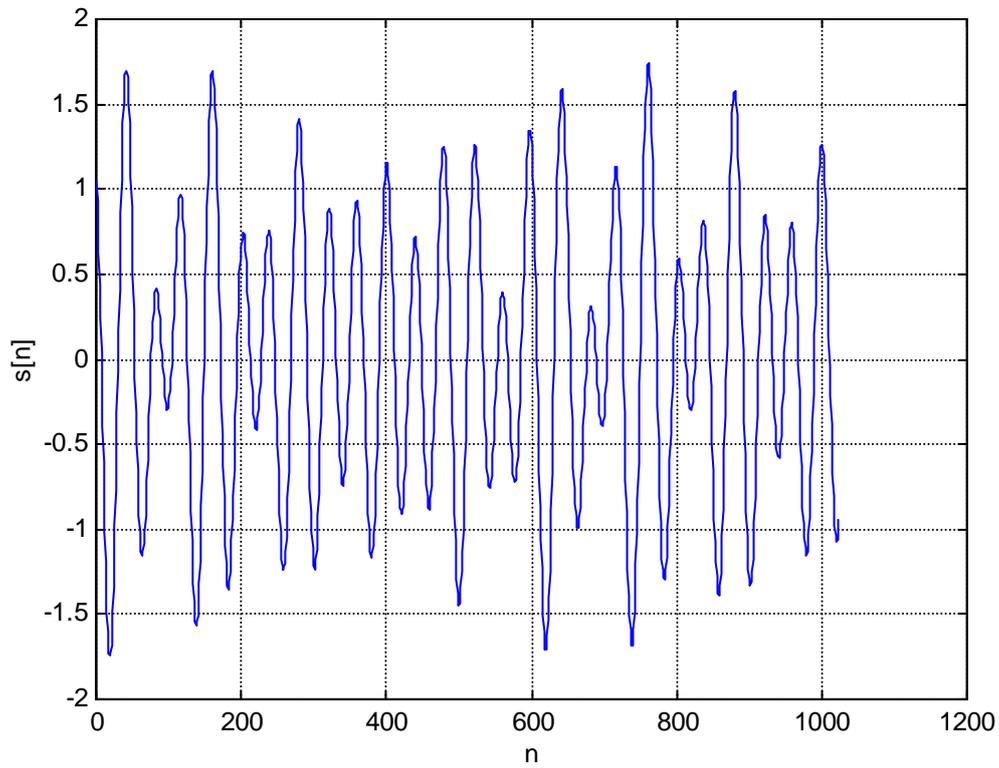


Fig.1 The Original Sequence $s(n)$

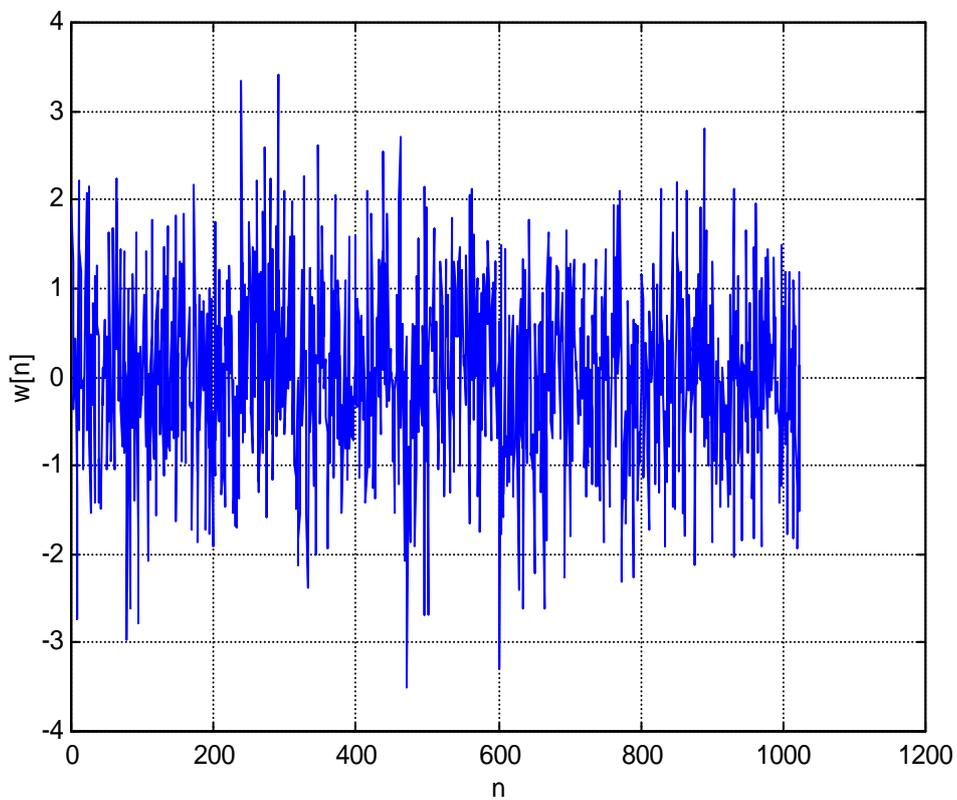


Fig.2 The Noise Sequence $w(n)$

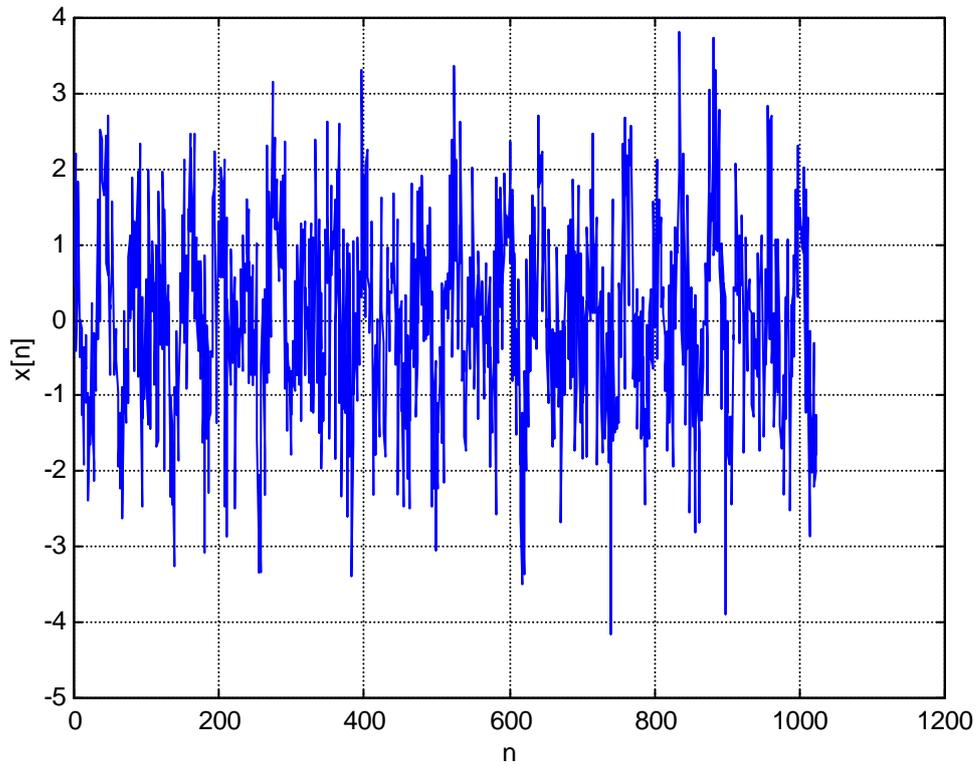


Fig.3 The Observed Sequence $x(n)=s(n)+w(n)$

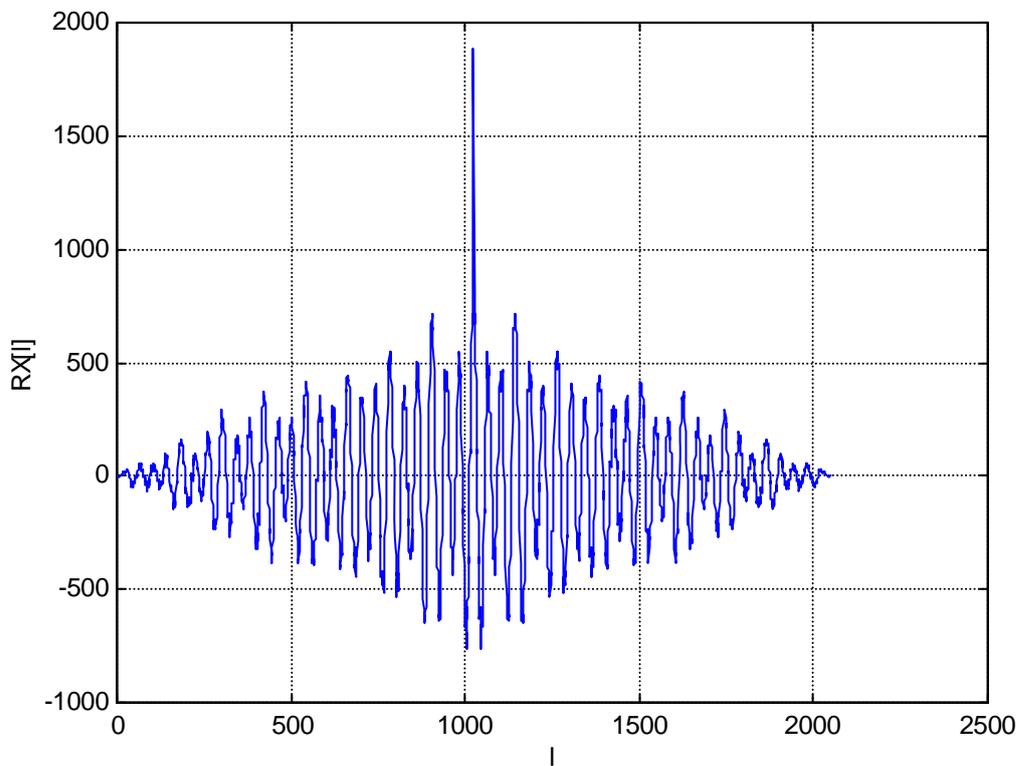


Fig.4 The Autocorrelation Function of the Observed Process

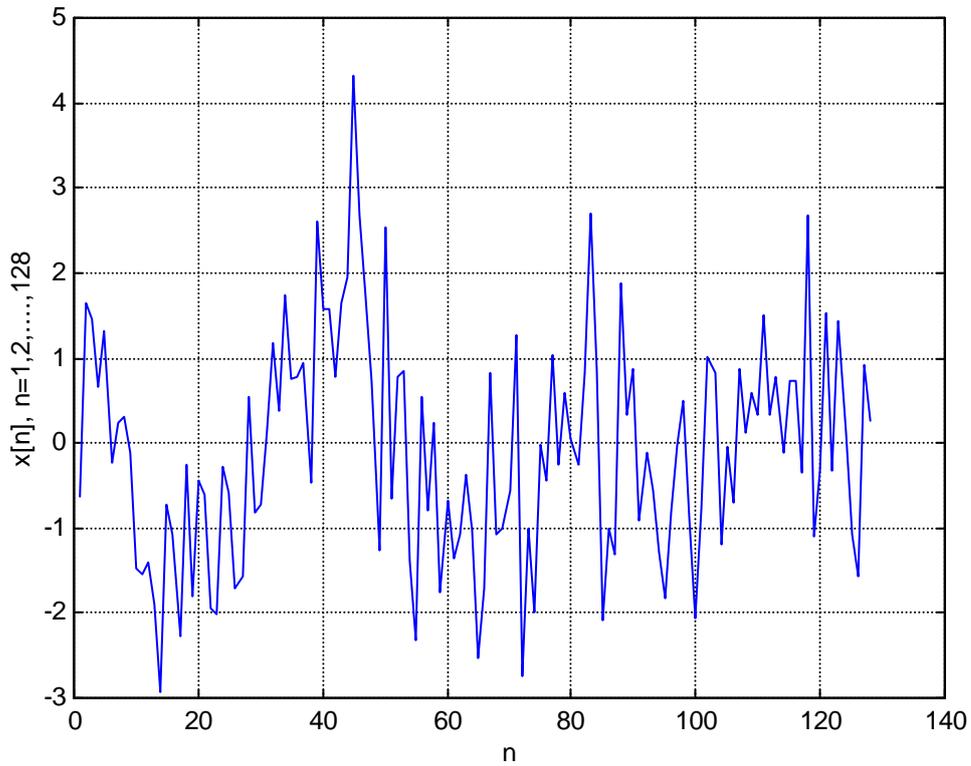


Fig.5 Segment of 128 Points of the Observed Sequence

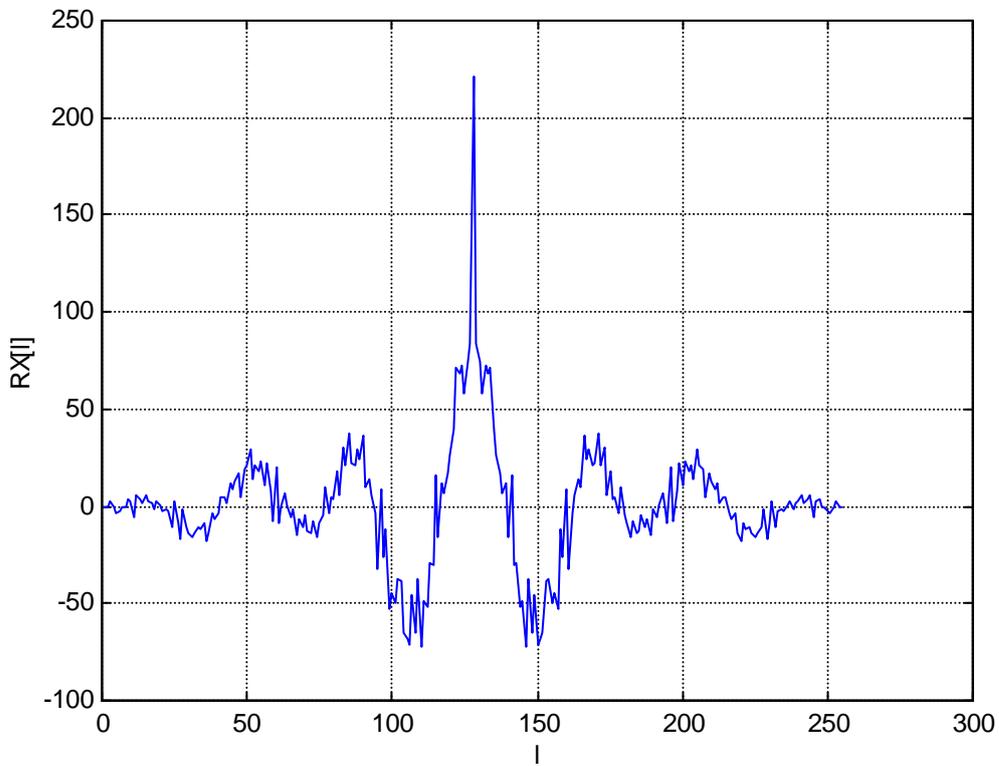


Fig.6 The Autocorrelation Function of the 128 Points of $x(n)$

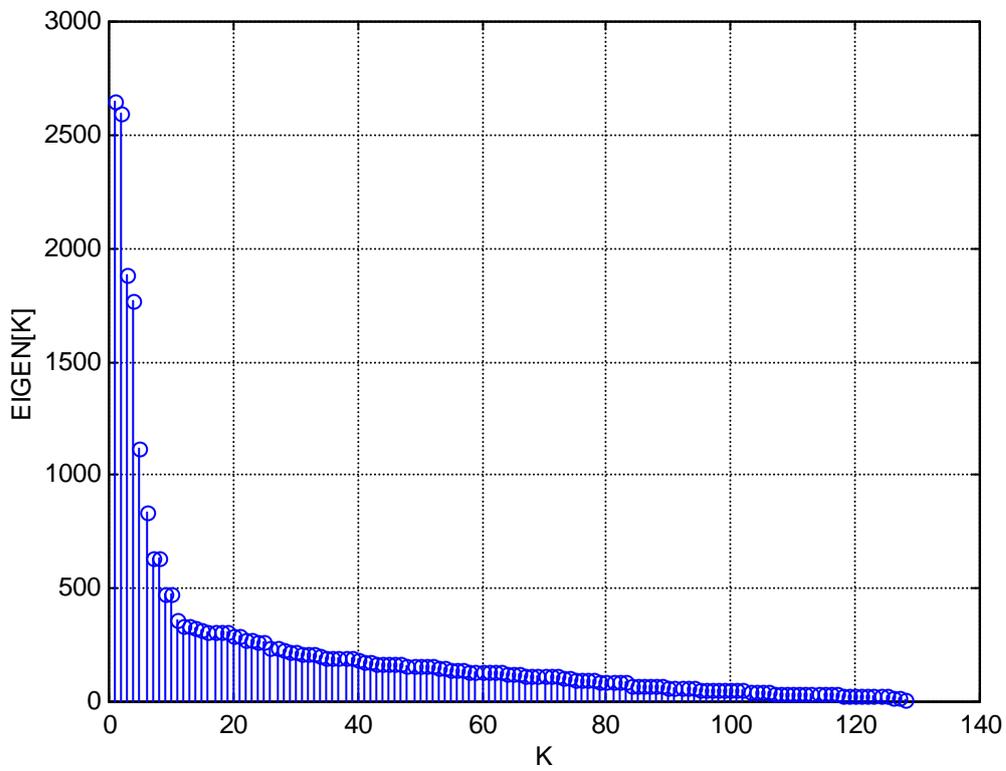


Fig.7 The Eigen Values in Descending Order

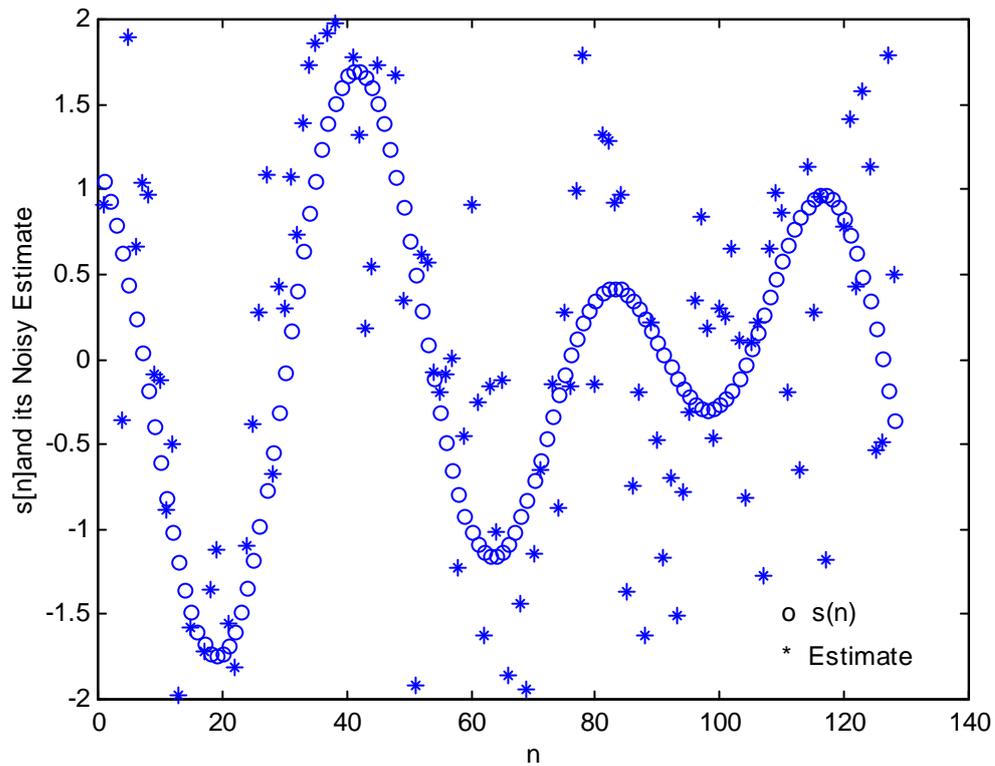


Fig.8 $s(n)$ (o) and its Noisy Estimate (*)

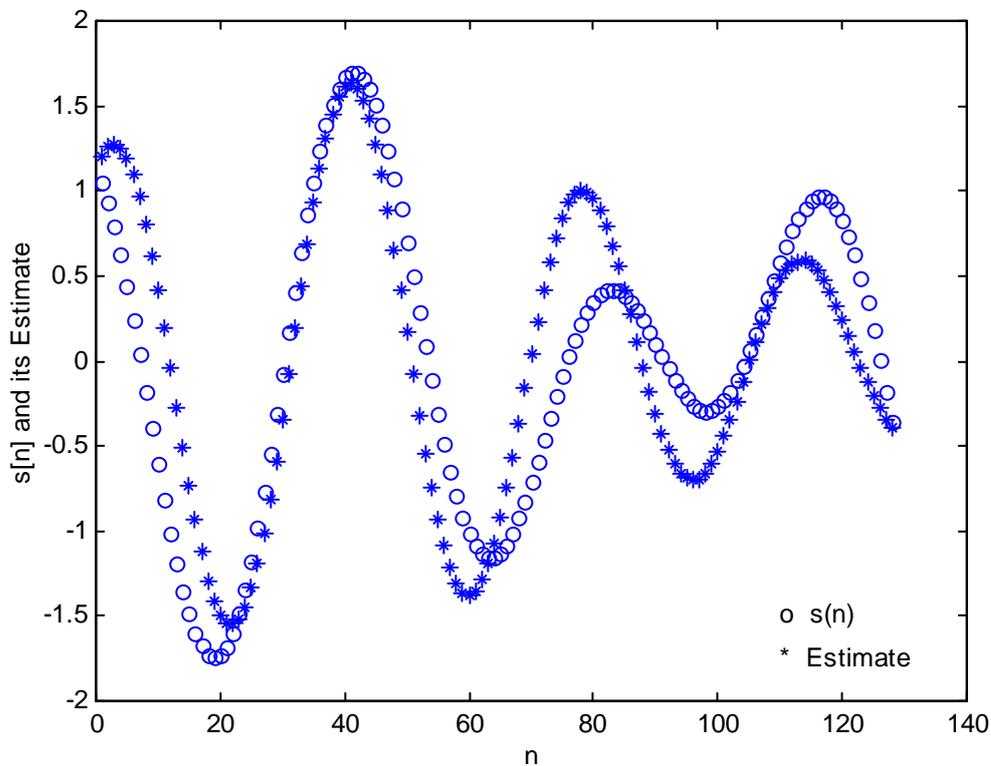


Fig.9 $s(n)$ (o) and its Estimate (*) Using 4-Dominant Eigen Vectors

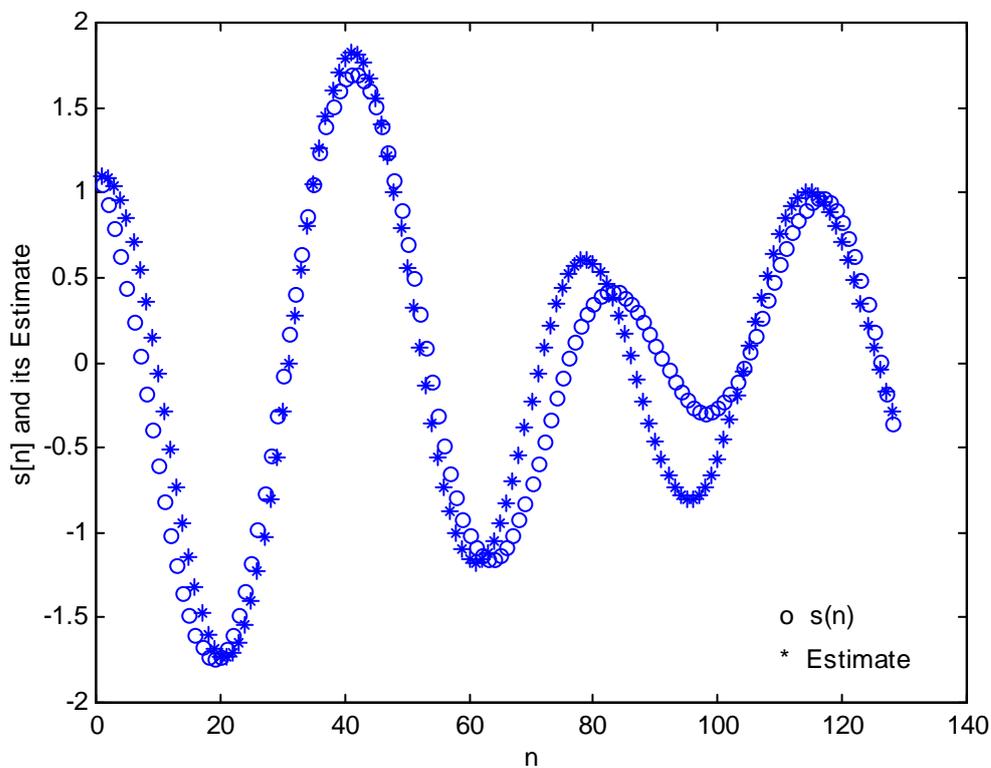


Fig.10 $s(n)$ (o) and its Estimate (*) Using 5-Dominant Eigen Vectors

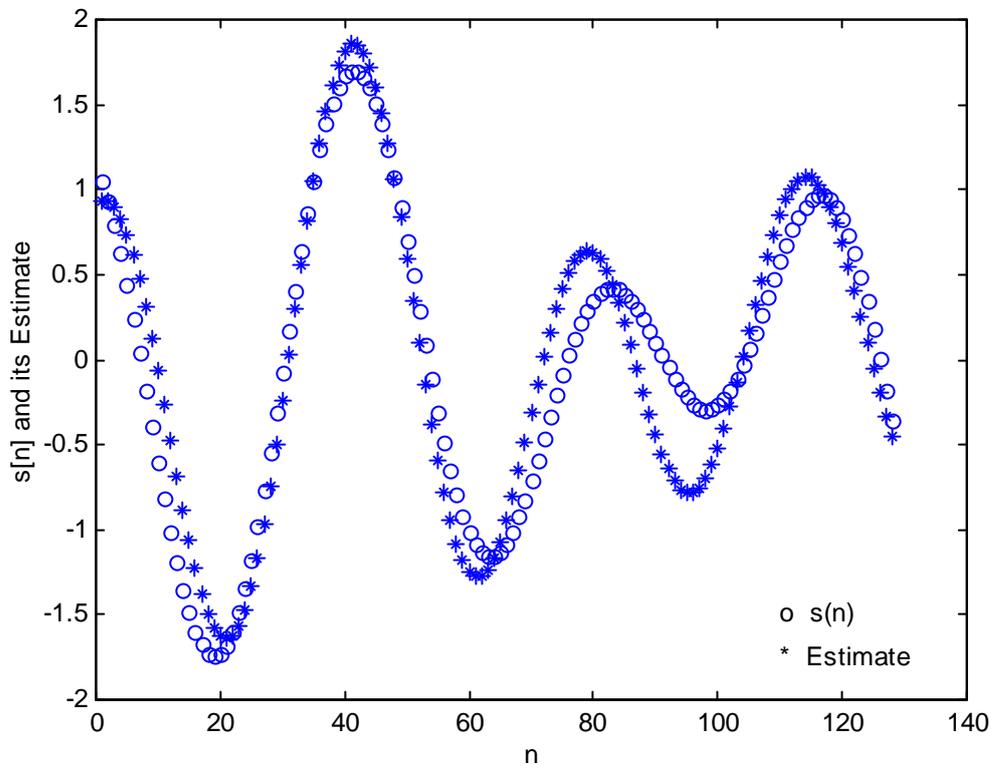


Fig.11 $s(n)$ (o) and its Estimate (*) Using 6-Dominant Eigen Vectors

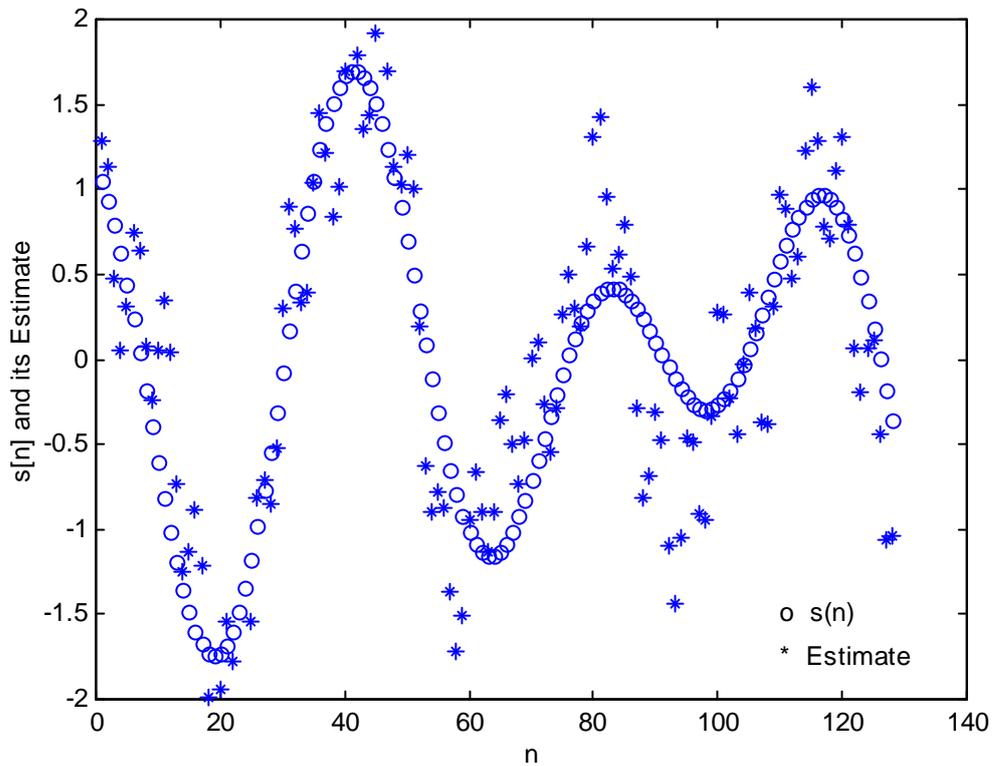


Fig.12 $s(n)$ (o) and its Estimate (*) Using 9-Dominant Eigen Vectors

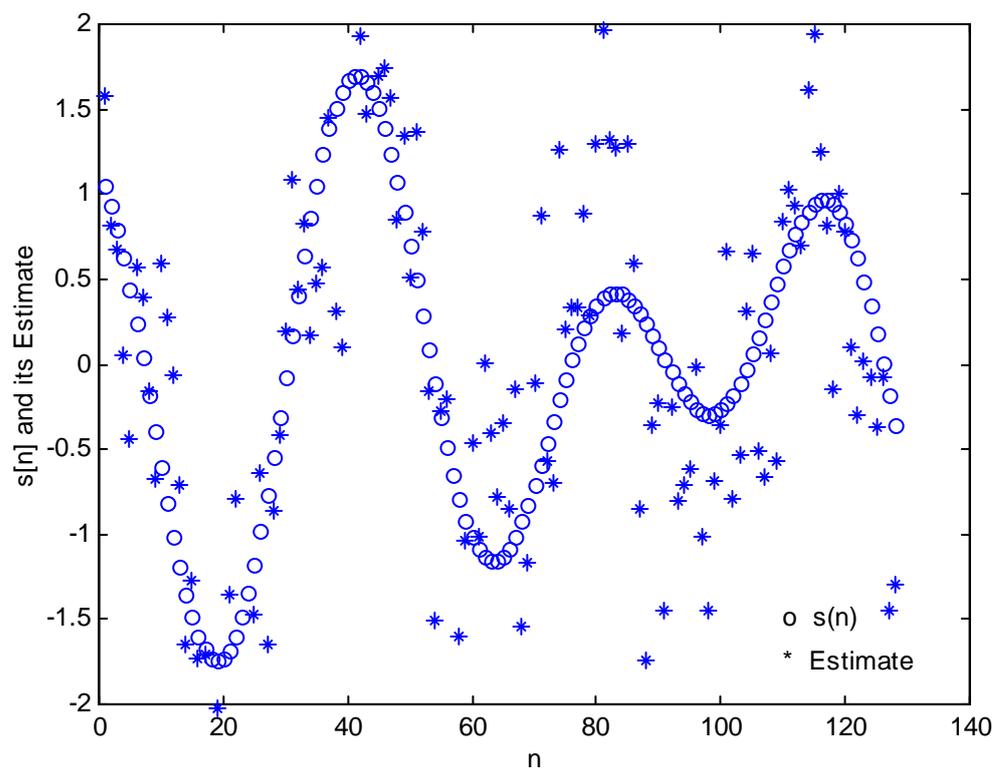


Fig.13 $s(n)$ (o) and its Estimate (*) Using 20-Dominant Eigen Vectors

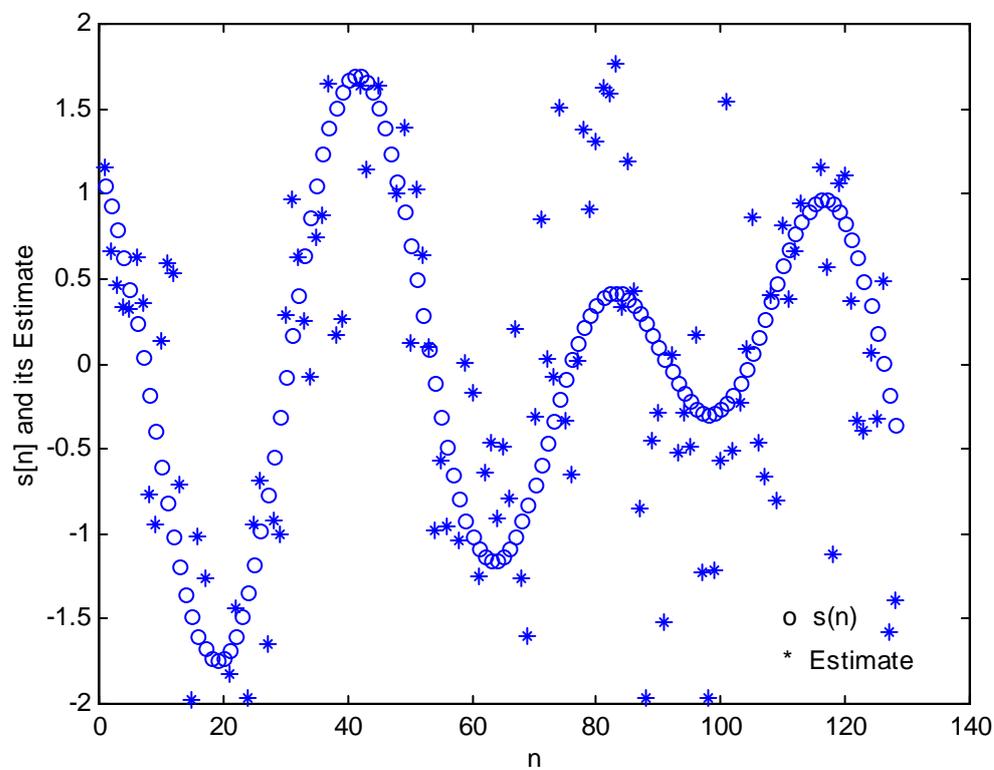


Fig.14 $s(n)$ (o) and its Estimate (*) Using 40-Dominant Eigen Vectors