Yugoslav Journal of Operations Research
3 (1993), Number 1, 21-31

A NEW APPROACH TO MODELING OF STOCHASTIC PROCESSES BASED ON PRINCIPAL COMPONENTS OF HANKEL MATRIX

Dragan RADOJEVIĆ

Mihajlo Pupin Institute
P.O.Box 15, 11000 Belgrade, Yugoslavia

Abstract. This paper describes a new method of modeling of stochastic processes by simultaneously determining the order and parameters of model. Process information is contained in observed time series – the system output or covariation of the output. The new method is based on singular value decomposition (SVD) of Hankel matrix of covariation of process output. It differs from a previous similar approach in a more adequate definition of optimization problem. Noniterative algorithm of time series modeling is realized. The capabilities of the algorithm are illustrated through an example and are compared to the known approach.

Key words: Stochastic processes, time series, modeling, Hankel matrix of covariances, singular value decomposition (SVD), principal components of Hankel matrix.

1. INTRODUCTION

This paper presents a new approach to solving the stochastic realization problem – the modeling of a time-stationary stochastic process described in terms of observed time series or, indirectly, in terms of system output covariance.

The problem reduces to summarizing adequately the information about the past of the process, represented theoretically by an infinite length vector, into a state vector whose dimension corresponds to the process order. The adequacy is achieved indirectly through the chosen criterion function and the problem reduces formally to an optimization problem. In the known approach, [7], the measure of correlation between the random vectors of the past and future was proposed to be the criterion function. The solution of the problem stated in this way are the principal components of process output covariance. It was shown that this approach has its shortcomings resulting form inadequate norming.

In this paper we propose a new statement of the optimization problem and give a solution that is operationalized in a noniterative algorithm for simultaneously determining the model order and model parameter estimates. The capabilities of the algorithm are illustrated by an example and compared to the results obtained by using the known procedure of the principal components of Hankel matrix.

Some notions required for the statement and solving of the modeling problem are given in Section 2 of this paper.

Section 3 describes briefly the existing approach to modeling by using the principal components of Hankel matrix and the main shortcomings of this approach.

A new statement of the modeling problem, the solution of the stated problem and the realized noniterative algorithm are presented in Section 4.

2. BASIC NOTIONS

We give here some notions needed for presenting the results obtained by modeling based on the analysis of the influence of the past on the future by using Hankel matrix of covariance.

The past of a stochastic process at a considered instant t is defined as the vector:

$$Y^{-}(t) = [y(t), y(t-1), \dots]^{T} \tag{1}$$

The future of a stochastic process is defined by the following vector:

$$Y^{+}(t) = [y(t+1), y(t+2), \dots]^{T} \tag{2}$$

The innovations model in the state space:

$$x(t+1) = Fx(t) + Tv(t) \tag{3a}$$

$$y(t) = hx(t) + v(t) \tag{3b}$$

where:

$x(t)$ – an $n \times 1$ state vector of stochastic process at instant t,
$y(t)$ – the scalar value of observation at instant t,
$v(t)$ – the scalar value of innovation at instant t,
F – an $n \times n$ transfer matrix,
T – an $n \times 1$ vector of innovation gain,
h – an $1 \times n$ vector of state gain.

By analogy between the definitions of the past and future, on the basis of stochastic process output it is possible to define the past and the future of a linear stochastic process in terms of appropriate innovation as well.

The past defined in terms of innovation at t:

$$V^{-}(t) = [\nu(t), \nu(t-1), \ldots]^{\mathrm{T}}. \tag{4}$$

The future defined in terms of innovation at t:

$$V^{+}(t) = [\nu(t+1), \nu(t+2), \ldots]^{\mathrm{T}}. \tag{5}$$

The observability matrix – Θ, of model (3):

$$\Theta := \begin{bmatrix} h \\ hF \\ hF^2 \\ hF^3 \\ \vdots \end{bmatrix}, \tag{6}$$

The controllability matrix – C, of model (3):

$$C := [T \;\; FT \;\; F^2T \;\; F^3T \;\; \ldots] \tag{7}$$

Hankel matrix of impulse response coefficients:

$$\mathcal{H} = \Theta\, C \tag{8}$$

The future vector estimate, based on the past vector expressed in terms of innovation, having the property of the minimum error variance, reads:

$$Y^{+} \,|\, V^{-} = \mathcal{H}\, V^{-}. \tag{9}$$

The state of stochastic process at instant t represents a vector of minimum dimensions that summarizes the relevant information about the entire past for prediction purposes.

$$\begin{aligned} x(t) &= [\, T \;\; FT \;\; F^2T \;\; F^3T \ldots] \, V^{-}(t) \\ &= C\, V^{-}(t). \end{aligned} \tag{10}$$

The state of minimum-phase model may also be interpreted as the summarizing of the past expressed in terms of system output, $Y^{-}(t)$, in a way similar to that in the original system, (3):

$$x(t) = [\, T \ (F - Th) T \ \ (F - Th)^2 T \ . \ . \,] \, Y^-(t)$$

$$= \Psi \, Y^-(t). \tag{11}$$

The minimum–phase model has all zeros in the unit circle.

The normal projection of vector $Y^+(t)$ onto $Y^-(t)$, of the future onto the past, respectively, of stochastic process on the basis of the results of multivariate statistics, [3], is:

$$Y^+ \,|\, Y^- = E(\, Y^+ \, Y^{-\mathrm{T}}\,) \, [\, E(\, Y^- \, Y^{-\mathrm{T}}\,)\,]^{-1} \, Y^- \tag{12a}$$

or

$$Y^+(t) \,|\, Y^-(t) = H R^{-1} \, Y^-(t). \tag{12b}$$

where:

R – Toeplitz matrix of output covariances:

$$R := E(\, Y^- \, Y^{-\mathrm{T}}\,)$$

H – Hankel matrix of output covariances:

$$H := E(\, Y^+ \, Y^{-\mathrm{T}}\,)$$

Note: The analyzed stochastic process is weak stationary and the following relation applied to it:

$$E(\, Y^- \, Y^{-\mathrm{T}}\,) = E(Y^+ \, Y^{+\mathrm{T}}\,)$$

Model order – state vector order

After the factorization of expression $H R^{-1}$ in Θ and Ψ, the following relation is obtained:

$$H R^{-1} \ Y^-(t) = \Theta \, \Psi \, Y^-(t)$$

$$= \Theta \, x(t), \tag{13}$$

therefore, the rank of $H R^{-1}$ is equal to the order of state vector x, i.e. the space defined by $\mathrm{Span}(H R^{-1} \, Y^-)$ has dimensions equal to the order of the state vector.

3. MODELING PROCEDURE

Since any base of space $\mathrm{Span}(H R^{-1} Y^-)$ can be the state vector x, the stochastic realization problem reduces to the choice of the most suitable basis for space $\mathrm{Span}(H R^{-1} Y^-)$, [1].

Summarizing of the past expressed in terms of system state vector is not unique. In the approach based on the principal components of Hankel matrix, the measure of correlation between the past and the future is maximized, without normalizing either the vector of the past or the vector of the future.

Formal statement of the optimization problem is:

Determine a matrix Ψ, whose rank is p (it has p columns) such that the following is achieved:

$$\underset{\Psi\Psi^{\mathrm{T}}=I_p}{\text{maximize}}\left\|E\left(Y^{+}\left(Y^{-}\right)^{\mathrm{T}}\Psi^{\mathrm{T}}\right)\right\|_F = \underset{\Psi\Psi^{\mathrm{T}}=I_p}{\text{maximize}}\left\|\left(H\,\Psi^{\mathrm{T}}\right)\right\|_F$$

The solution of the stated optimization problem is:

$$\Psi = U_1^{\mathrm{T}} \tag{14}$$

where: U_1 are the principal components of Hankel matrix:

$$H = V\Sigma U = V_1\Sigma_1 U_1^{\mathrm{T}} + V_2\Sigma_2 U_2^{\mathrm{T}}$$

The observability matrix, for the previously determined Ψ, is defined, on the basis of (13) and (14), by the following matrix equation:

$$\Theta = H\Psi^{\mathrm{T}}(\Psi R\Psi^{\mathrm{T}})^{-1} = H U_1 (U_1^{\mathrm{T}} R U_1)^{-1}\ . \tag{15}$$

The innovation model in the state space, on the basis of matrices Ψ and Θ, is determined by applying the modified Ho–Kalman algorithm, [4],:

$h = 1$. the 1st row of matrix Θ,

$T = 1$. the 1st column of matrix Ψ and

$F = \Theta_1^{\dagger}\,\Theta_2$,

where: $\Theta_1^{\dagger} := (\Theta_1^{\mathrm{T}}\Theta_1)^{-1}$; Θ_1 – a matrix obtained by omitting the last row from Θ and Θ_2 – a matrix obtained by omitting the first row from Θ.

Shortcomings of the existing approach

Although the procedure described was widely used in the mid–eighties [2] – [6], the shortcomings of approximation to the principal components of Hankel matrix were identified in paper [5]. The property of maximum correlation with the future measured in the proposed way is not adequate, because this measure is affected by a quantity such as the variance of state vector x. This becomes particularly prominent when this procedure is applied to the vector case where the variance of different series differ greatly. The inadequacy of scalar case is illustrated by the example given in this paper.

4. NEW MODELING PROCEDURE

Contrary to the previous, the approach we propose consists of searching for the most suitable orthonormed directions in the space of the future that determine the subspace - the future which contains maximum information about the past. The matrix of such p principal components is the observability matrix. The matrix of summarizing the past into a state is obtained simply, on the basis of the observability matrix. Formal statement of optimization problem is:

Determine the observability matrix, Θ, having the property that:

$$\underset{\Theta^{\mathrm{T}}\Theta=I_p}{\operatorname{maximize}} \left\| E\left(\Theta Y^{+} \left(Y^{-}\right)^{\mathrm{T}} \right) \right\|_F = \underset{\Theta^{\mathrm{T}}\Theta=I_p}{\operatorname{maximize}} \left\| (\Theta H) \right\|_F$$

The solution of the stated optimization problem is

$$\Theta = V_1 \tag{16}$$

where: V_1 are the first p principal components of Hankel matrix:

$$H = V \Sigma U = V_1 \Sigma_1 U_1^{\mathrm{T}} + V_2 \Sigma_2 U_2^{\mathrm{T}}.$$

The matrix of summarizing the past is:

$$\Psi = V_1^{\mathrm{T}} H R^{-1} \tag{17}$$

PROOF:

$$\left\| E\left(\Theta Y^{+} \left(Y^{-}\right)^{\mathrm{T}} \right) \right\|_F = \left\| (\Theta H) \right\|_F = tr\, \Theta H H^{\mathrm{T}} \Theta^{\mathrm{T}}$$

Incorporating a matrix U, such that $U^{\mathrm{T}} U = U U^{\mathrm{T}} = I$, expression (18) can be written as:

$$tr\, \Theta H H^{\mathrm{T}} \Theta^{\mathrm{T}} = tr\, \Theta H U^{\mathrm{T}} U H^{\mathrm{T}} \Theta^{\mathrm{T}} \tag{19}$$

The maximum value of expression (19) under the conditions of the imposed constraint is achieved with that Θ which maximizes the following expression:

$$\underset{\substack{\Theta^{\mathrm{T}}\Theta=I_p \\ U^{\mathrm{T}}U=I}}{\operatorname{maximize}}\ \Theta H U^{\mathrm{T}}$$

The solution of this problem is achieved by the principal components of the eigenvalues of matrix H:

$$H = V \Sigma U^{T} = V_1 \Sigma_1^{T} U_1 + V_2 \Sigma_2 U_2^{T},$$

wherefrom:

$$\Theta = V_1 .$$

Since, on the basis of (13), $\Theta \Psi = H R^{-1}$, it follows that:

$$\Psi = V_1^{T} H R^{-1}.$$

End of proof. ■

Taking into account all that has been stated, the algorithm for time series modeling based on the system output covariance consists of the following three steps:

ALGORITHM

Step 1: Perform SVD of matrix $H = V \Sigma U^{T}$ and determine the number of singular values p which is statistically significantly different from zero (p is the model order).

Step 2: Determine: the observability matrix $\Psi = V_1$ and the matrix of summarizing $\Psi = V_1 H R^{-1}$.

Step 3: Determine the parameters of innovation model in the state space

h - 1st row of matrix Θ,

T - 1st column of matrix Ψ,

$F = \Theta_1^{\dagger} \Theta_2$.

Where: $\Theta_1^{\dagger} := (\Theta_1^{T} \Theta_1)^{-1}$; Θ_1 - a matrix obtained by omitting the last row from Θ and Θ_2 - a matrix obtained by omitting the first row from Θ.

EXAMPLE 1: Realize a model in the state space in the innovation form for a stochastic process whose output (time series) covariance is given, Figure 1, by applying the old and the new method. Compare the functions of output covariance that correspond to the former and the later model.

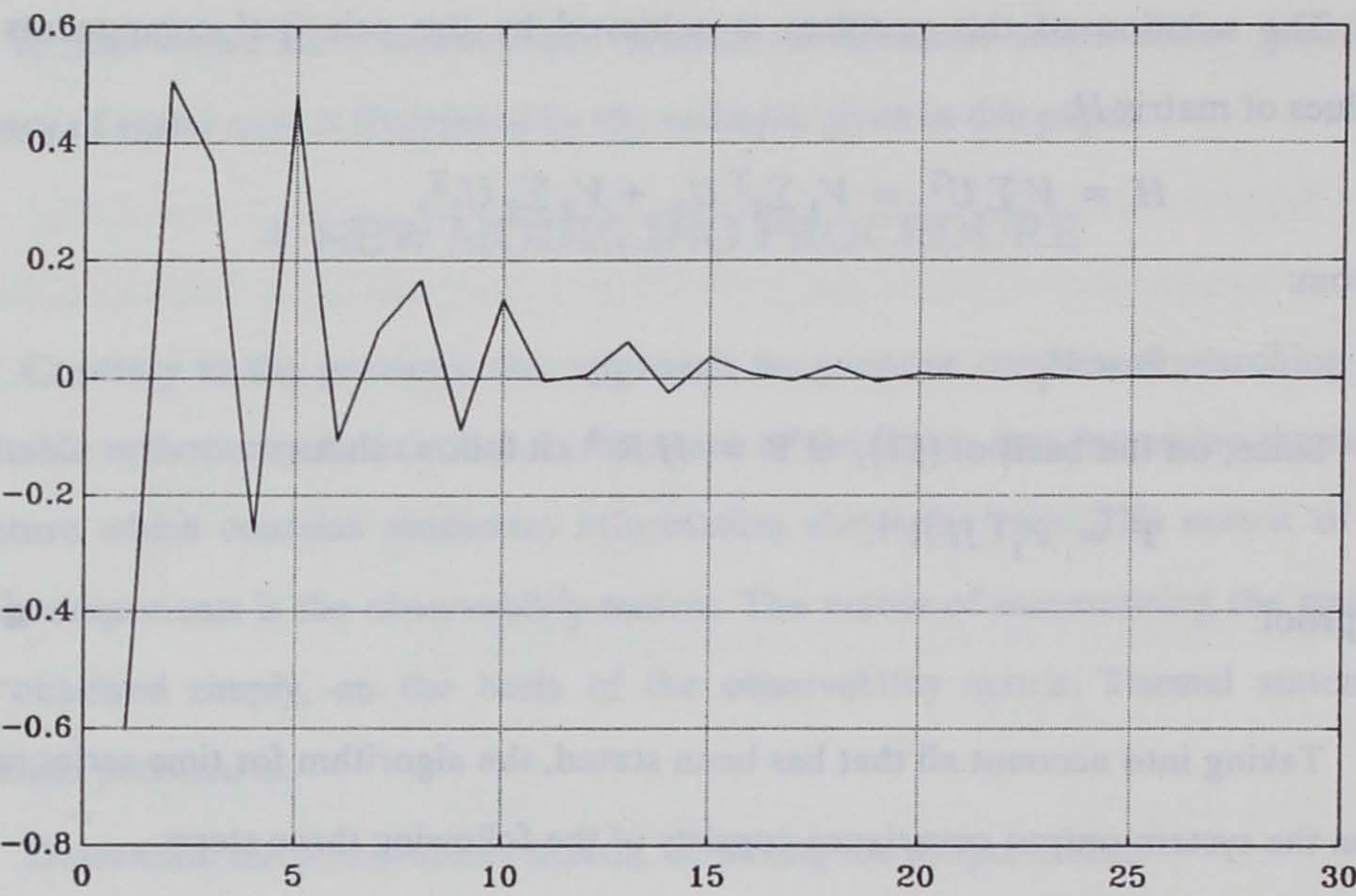

Figure 1. Time series covariance

The results of modeling are:

<u>The singular values of Hankel matrix</u> covariances are presented in the Figure 2:

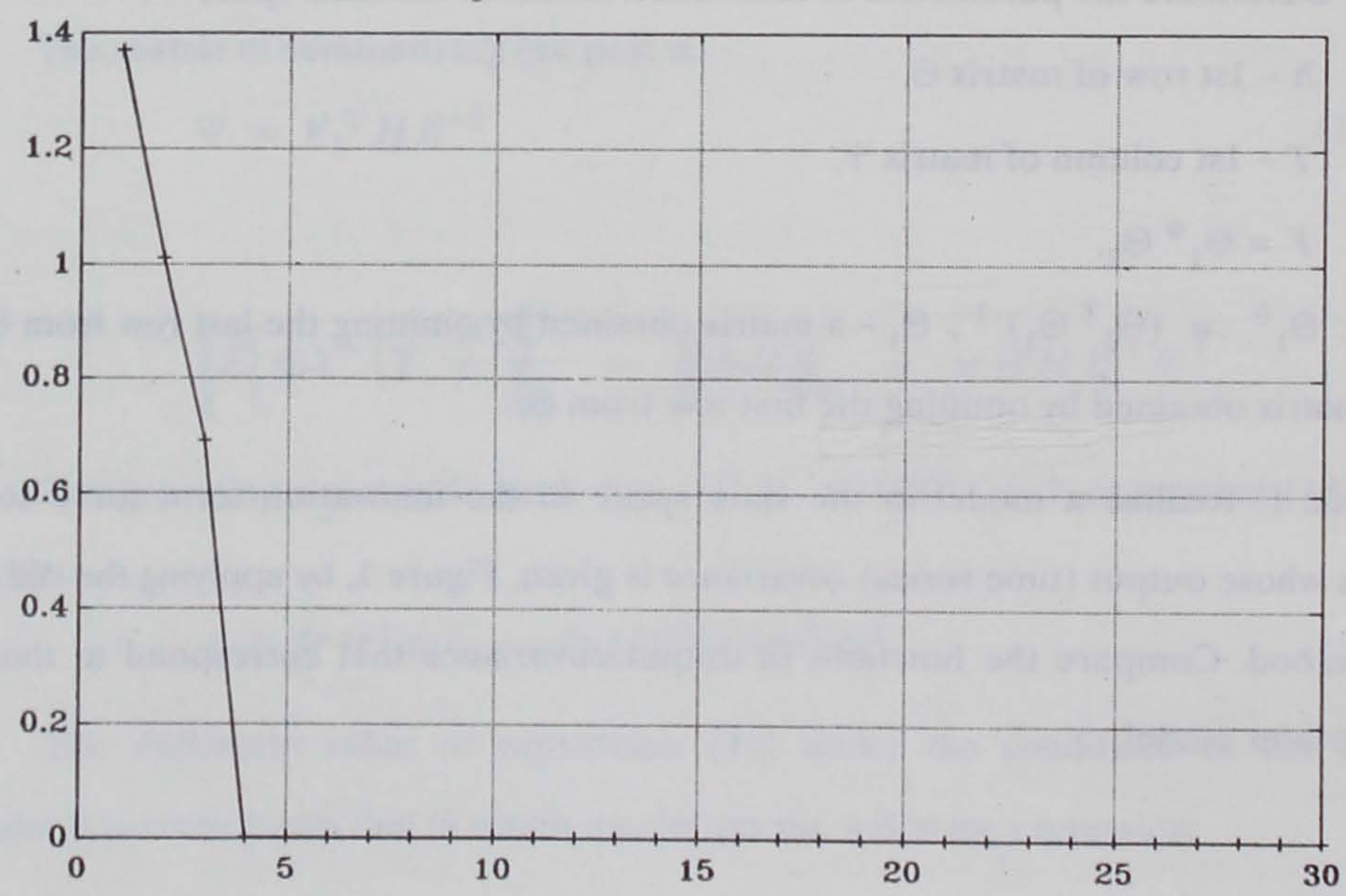

Figure 2. SV of Hankel matrix

It is concluded from the Figure 2 that the model order is $p = 3$.

System matrices are:

(a) for the old approach:

$$F = \begin{bmatrix} -0.519 & 0.436 & -0.231 \\ -0.746 & -0.657 & -0.208 \\ -0.008 & -0.489 & 0.776 \end{bmatrix}; \qquad T = \begin{bmatrix} -0.740 \\ -0.007 \\ -0.462 \end{bmatrix}$$

$$h = [-0.358 \quad 0.006 \quad -0.1721]$$

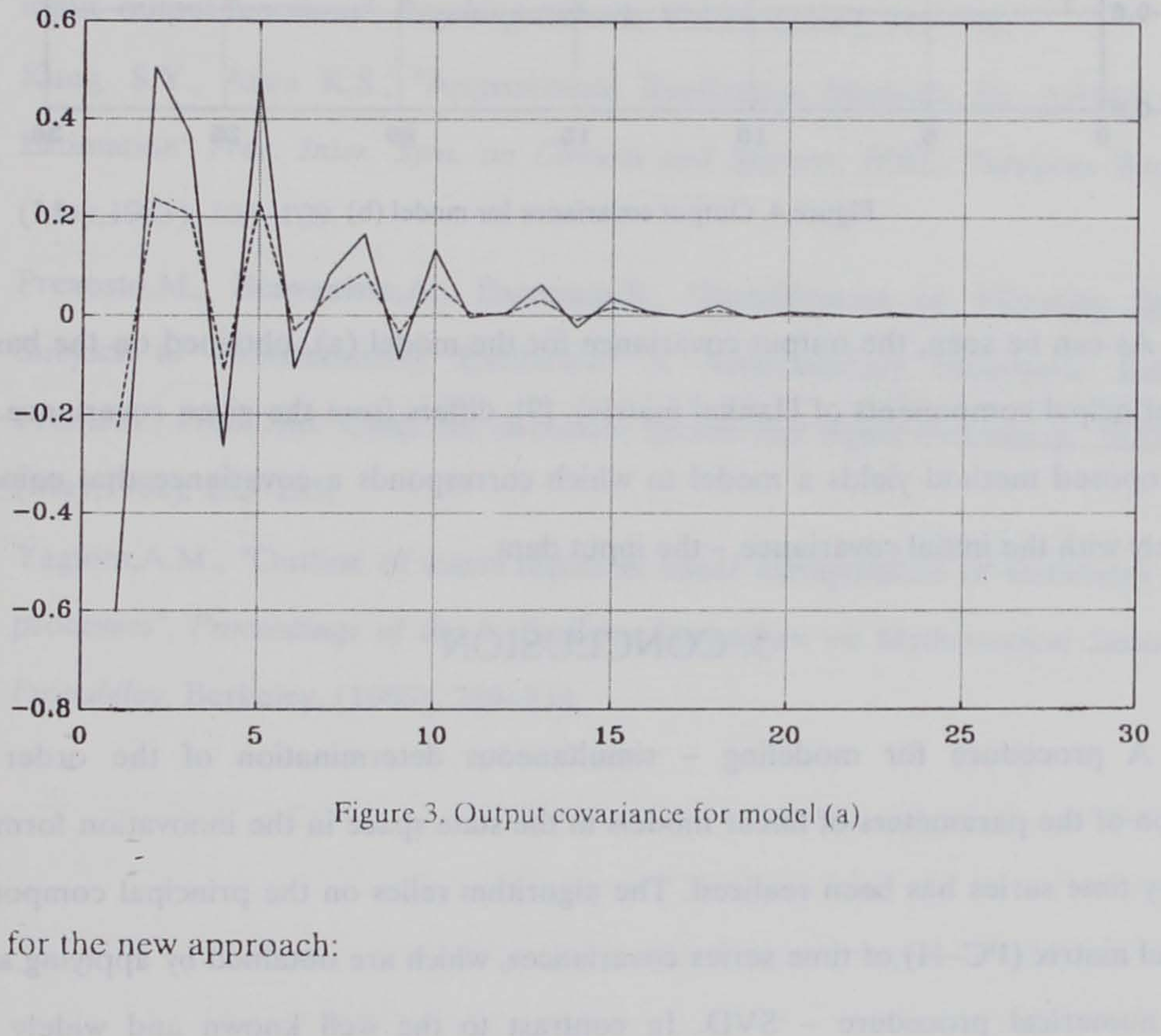

Figure 3. Output covariance for model (a)

(b) for the new approach:

$$F = \begin{bmatrix} -0.458 & -0.660 & 0.131 \\ 0.487 & -0.676 & -0.480 \\ -0.066 & -0.328 & 0.735 \end{bmatrix}; \qquad T = \begin{bmatrix} 0.694 \\ 0.228 \\ -0.250 \end{bmatrix}$$

$$h = [-0.740 \quad 0.007 \quad -0.462]$$

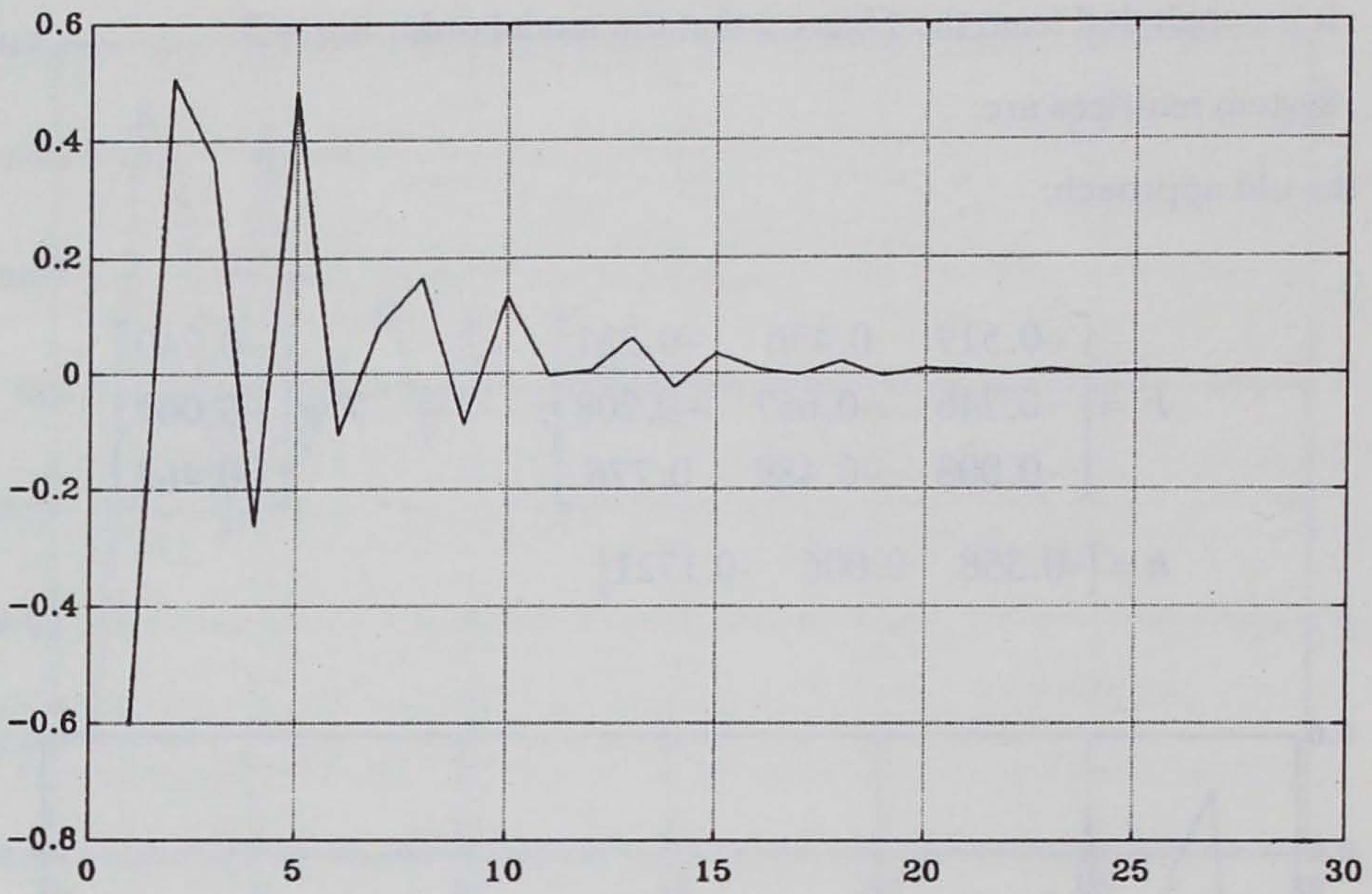

Figure 4. Output covariance for model (b)

As can be seen, the output covariance for the model (a), obtained on the basis of PC–H (principal components of Hankel matrix), [9], differs from the given covariance. The newly proposed method yields a model to which corresponds a covariance that coincides completely with the initial covariance – the input data.

5. CONCLUSION

A procedure for modeling – simultaneous determination of the order and estimation of the parameters of linear models in the state space in the innovation form – of stationary time series has been realized. The algorithm relies on the principal components of Hankel matrix (PC–H) of time series covariances, which are obtained by applying a very reliable numerical procedure – SVD. In contrast to the well known and widely used approach, which is also based on the PC–H, in the newly proposed approach the problem of norming is solved adequately by defining a new optimization problem. A scalar case is considered in this paper, and the expansion to a vector case is direct. The vector case and a balanced realization relying on the use of this procedure will be treated in detail in a forthcoming paper.

REFERENCES

[1] Akaike,H., "A New Look at the Statistical Model Identification", *IEEE Tr. on Aut. Cont.*, Vol. AC–19, No. 6, (Dec.,1974), 716–723.

[2] Cadzow,J.A., "Spectral Estimation: An Overdetermined Rational Model Equation Approach", *Proc. IEEE*, Vol.70, No.9, (Septem.,1982), 907–939.

[3] Dillon,W.R., Goldstein,M., *Multivariate Analysis – Methods and Applications*, John Wiley & Sons, 1984.

[4] Ho,B.L., Kalman,R.E., "Effective construction of linear, state–variable models from input–output functions", *Regelungstechnik*, Vol.14, (1966), 545–548.

[5] Kung, S.Y., Arun K.S., "Approximate Realization Methods for ARMA Spectral Estimation" *Proc. Inter. Sym. on Circuits and Systems, IEEE*, Newport Beach, CA, (May,1983), 105–109.

[6] Prevosto,M., Benveniste,A., Barnouin,B., "Identification of Vibrating Structures Subject to Nonstationary Excitation: A Nonstationary Stochastic Realization Problem", *Proc. Intl. Conf. on Acoustics Speech and Signal Processing, IEEE*, Paris, (May,1982), 252–255.

[7] Yaglom,A.M., "Outline of some topics in linear extrapolation of stationary random processes", *Proceedings of the 5. Berkeley Symposium on Mathematical Statistics and Probability*, Berkeley, (1965), 259–278.