Analysis of subspace methods for System Identification

Seyed Iman Shekarforoush s151362

Danish Technical University, DTU - Electrical Department

Arturo Arranz Mateo s160412

Danish Technical University, DTU - Electrical Department

Abstract

In this report the performance and limitations of subspace methods for system identification under noisy environments are studied.

Keywords: Subspace methods, N4SID, realization theory, State Space models,

1. Introduction

Traditional input-output methods for system identification such as *prediction error methods*(PEMs) or *Intrumental variable methods*(IV) do not deal satisfactorily with MIMO systems (Katayama (2006)).Input-output methods are based on optimization techniques, which can get stuck in local minima, not finding global solutions.
Alternatively there are the *subspace identification methods* (SIMs), which are based on algebra properties and matrices decomposition such as SVD. Most SIMs fall into the unifying theorem (Van Overschee and De Moor (1995)) among which are canoncial variate analysis(CVA), N4SID Van Overschee and De Moor (1994) and MOESP Verhaegen and Dewilde (1992).
Another attractive advantage of SIMs are the state space form which is very convenient for estimation, filtering, prediction and control. However SIMs methods count with several drawbacks. In general the estimates of SIMs are not as accurate as input-output methods. Moreover, it is not until recently which SIMs are suitable for closed-loop identification, necessary or many applications (Qin (2006)).
The structure of the report is as follow. In the section 2, the notation and assumptions are presented. In the section 3 an overview of the subspace algorithms is explained. In sections 4,5 and 6 the N4SID algorithm is explained for a deterministic, stochastic and deterministic-stochastic systems, respectively. Finally in section 7, experimental results are discussed for each of the aforementioned cases.

2. Notation and Assumptions

In this section we introduce the notation of the Hankel matrices, extended observability matrix, reversed controlability matrix et.al. Hankel matrices play and important role in the subspace identification methods. These matrices can be build using the input-output data. Input Hankel matrices are defined as :

$$U_{0|2i-1} := \left(\begin{array}{ccccc} u_0 & u_1 & u_0 & \cdots & u_0 \\ u_1 & u_2 & u_0 & \cdots & u_0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ u_{i-1} & u_i & u_{i+1} & \cdots & u_{i+j-2} \\ \hline u_i & u_{i+1} & u_{i+2} & \cdots & u_{i+j-1} \\ u_{i+1} & u_{i+2} & u_{i+3} & \cdots & u_{i+j} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ u_{2i-1} & u_{2i} & u_{2i+1} & \cdots & u_{2i+j-2} \end{array}\right)$$
$$:= \left(\frac{U_{0|i-1}}{U_{i|2i-1}}\right) := \left(\frac{U_p}{U_f}\right)$$

$$U_{0|2i-1} := \left(\begin{array}{ccccc} u_0 & u_1 & u_0 & \cdots & u_0 \\ u_1 & u_2 & u_0 & \cdots & u_0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ u_{i-1} & u_i & u_{i+1} & \cdots & u_{i+j-2} \\ u_i & u_{i+1} & u_{i+2} & \cdots & u_{i+j-1} \\ \hline u_{i+1} & u_{i+2} & u_{i+3} & \cdots & u_{i+j} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ u_{2i-1} & u_{2i} & u_{2i+1} & \cdots & u_{2i+j-2} \end{array}\right)$$
$$:= \left(\frac{U_{0|i}}{U_{i+1|2i-1}}\right) := \left(\frac{U_p^+}{U_f^-}\right)$$

where :

- The number of block rows, i, is defined by the user. It should be defined larger than the maximum order of the system one want to identify.
- The number of columns, j is defined such as all the sampled data is used. This implies that $j = s - 2i + 1$, where s is the number of samples.
- The subscripts in $U_{0|i}, U_{i|2i-1}, U_{i|2i-1}$ indicate the first and last element in the first column of the Hankel matrix. The subscript "p" and "f" indicates

"past" and "future". This definition of past and future data is somehow loose, since in both sides we find shared data. However, it is useful for explaining the algorithm.

The output block Hankel matrices, $Y_{i|2i-1}, Y_p, Y_f$ are defined in the same way that the input. The block Hankel Matrix, W_p, consist on the inputs and outputs :

$$W_p = \begin{pmatrix} U_p \\ Y_p \end{pmatrix}$$

The state space sequence which the system have follow during the sequence of input-output is defined as :

$$X_i := \begin{pmatrix} x_i & x_{i+1} & \dots & x_{i+j-2} & x_{i+j-1} \end{pmatrix} \in \mathbb{R}^{nxj}$$

the subscript "i" indicates the first element in the state space sequence.
The extended (i>n) observabilty matrix is defined as :

$$\Gamma_i := \begin{pmatrix} C \\ CA \\ CA^2 \\ \dots \\ CA^{i-1} \end{pmatrix}$$

We assume that A,C to be observable, which implies that the rank of Γ_i is n. The reversed extended controllability matrix, Δ_i, is defined as :

$$\Delta_i := \begin{pmatrix} A^{i-1}B & A^{i-2} & \dots & AB & B \end{pmatrix} \in \mathbb{R}^{nxmi}$$

We assume that the pair A,B, are controllable. The lower triangular Toepliz matrix, H_i is defined as :

$$H_i := \begin{pmatrix} D & u_1 & u_0 & \dots & 0 \\ CB & D & u_0 & \dots & 0 \\ CAB & CB & D & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ CA^{i-2}B & CA^{i-3}B & CA^{i-4}B & \dots & D \end{pmatrix}$$

3. Algorithm overview and Unifying Theorem

The unifying theorem Van Overschee and De Moor (1995) established a framework for most of the subspace methods. Most of the subspace methods such as CVA, MOESP or N4SID, consist in two steps. In the first step a certain characteristic subspace is calculated directly from the input-output data. This subspace finds to be the extended observability matrix, Γ_i. We can also infer the order of the system, n, which is equal to the dimension of Γ_i. The aforementioned subspace algorithms coincide on this first step.
However, differences are found in the second step which is getting the system matrices : A,B,C and D. N4SID, the preferred algorithm during this report, makes a reconstruction of state space sequence, X_i, which the system follow during the input-output sequence. Then the system matrices follow directly. Given the following system, the algorithms flow is presented in Figure 1.

$$\begin{aligned} x_{k+1} &= Ax_k + Bu_k + w_k \\ y_k &= Cx_k + Du_k + v_k \end{aligned} \tag{1}$$

with

$$E\Big[\begin{pmatrix} w_k \\ v_k \end{pmatrix} \begin{pmatrix} w_k & v_k \end{pmatrix} \Big] = \begin{pmatrix} Q & S \\ S^T & R \end{pmatrix} \delta``0 \tag{2}$$

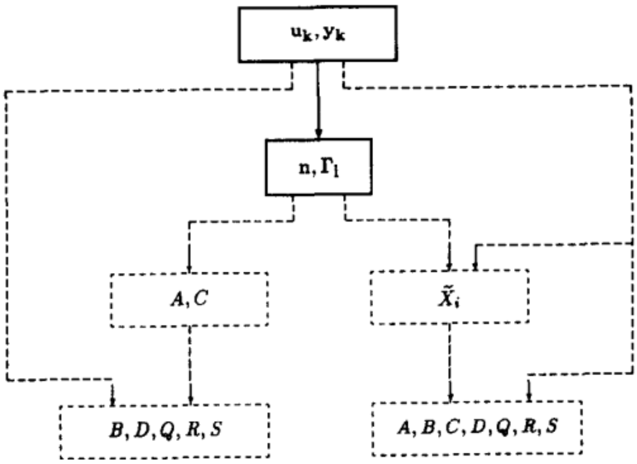

FIGURE 1: SIDs algorithms flow

4. Deterministic Systems

4.1. Problem description

Given : s measurements of input $u_k \in \mathbb{R}^m$ and output $y_k \in \mathbb{R}^l$ generated by the pure deterministic system :

$$\begin{aligned} x_{k+1} &= Ax_k + Bu_k \\ y_k &= Cx_k + Du_k \end{aligned} \tag{3}$$

we need to **determine** the order of the system, n, and the system matrices $A \in \mathbb{R}^{nxn}, B \in \mathbb{R}^{nxm}, C \in \mathbb{R}^{lxn}, D \in \mathbb{R}^{lxm}$.

4.2. Geometric properties

Thanks to the Hankel block matrices defined in the notation section we can define the following system of equations :

$$Y_p = \Gamma_i X_p + H_i U_p \tag{4a}$$
$$Y_f = \Gamma_i X_f + H_i U_f \tag{4b}$$
$$X_f = A^i X_p + \Delta_i U_p \tag{4c}$$

The vectors in the row space of Y_f, which is known data, are obtained as a sum of linear combinations of vectors in the row space of U_f, also known, plus a linear combinations of the row space of the state space sequence X_f. By means of oblique projections of this Hankel matrices, the state space sequence and the extended observaiblity matrix can be found, as explained the the next section. These geometrical properties are shown in the Figure 8

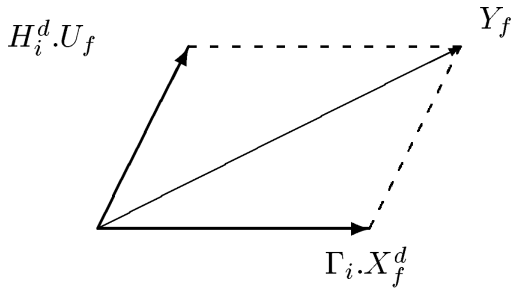

FIGURE 2: Oblique projection of the output

4.3. Main theorem

Under the assumptions that :

- *The input u_k is persistently exiting of order 2i (Lung (1987))*
- *The intersection of the row space of U_f (future inputs) and the row space of X_p (the past states) is empty.*

And with $\mathcal{O}_i$ defined as the oblique projection :

$$\mathcal{O}_i := Y_f/_{U_f}W_p \tag{5}$$

which means the oblique projection of the row space of Y_f over the row space of W_p along the direction U_f. We define the SVD decomposition of $\mathcal{O}_i$ as :

$$\mathcal{O} = USV^T \tag{6}$$

then we have :

1. The matrix $\mathcal{O}_i$ is equal to the product of the extended observability matrix and the states :

$$\mathbb{O}_i = \Gamma_i X_f \tag{7}$$

2. The order of the system is equal to the number of singular values in the SVD decomposition different from zero.
3. The extended observability matrix Γ_i is equal to :

$$\Gamma_i = US^{1/2} \tag{8}$$

4. The part of the state sequence X_f can be recovered from :

$$X_f = S^{1/2}V^T \tag{9}$$

The main theorem explained in this section is the common denominator for most of the subspace methods which allow to complete the first step : calculating n ,Γ_i and X_f. Even though that the theorem suffer from some modifications in the stochastic case, the underlying idea of projections over Hankel matrices is maintained. The proofs can be found at (Van Overschee and De Moor (2012))

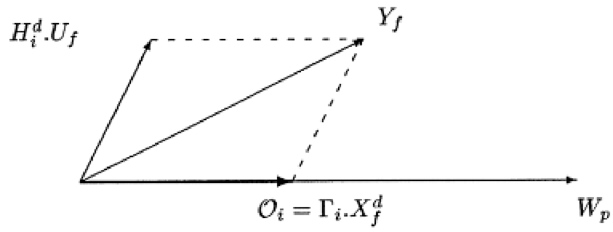

FIGURE 3: Oblique projection of Y_f over W_p

4.4. Computing the System Matrices

With same proof as in the main theorem the following holds :

$$\mathcal{O}_{i-1} := Y_f^- /_{U_f^-} W_p^+ \quad \Gamma_{i-1}X_{i+1} \tag{10}$$

and X_{i+1} is calculated as :

$$X_{i+1} = \Gamma_{i+1}^{\dagger}\mathcal{O}_{i-1} \tag{11}$$

With just the input-output data we have been able to calculate X_{i+1} and X_i. Now it is possible the build the augmented system :

$$\underbrace{\begin{pmatrix} X_{i+1} \\ Y_{i|i} \end{pmatrix}}_{known} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \underbrace{\begin{pmatrix} X_i \\ U_{i|i} \end{pmatrix}}_{known} \tag{12}$$

Note how the system matrices can be compute as a linear set of equations in one step. Here conclude the algorithm explanation for a deterministic system.

5. Stochastic Systems

5.1. Problem description

In this section,the state space representation of a stochastic system will be driven based on the available output data.The stochastic subspace problem can be illustrated by Figure 4 .In this problem,only the output data is measured and the states of system are unknown,while the states will be calculated as a result of subspace system identification algorithm.

Given : N sample of output $y_k \in \Re^l$ generated by the unknown stochastic system with unknown order of n :

$$x_{k+1} = Ax_k + w_k$$

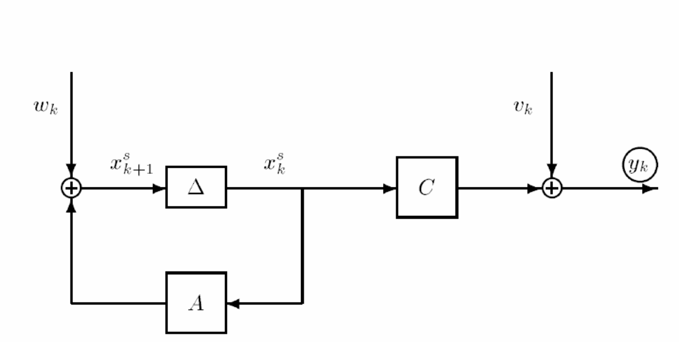

FIGURE 4: A LTI stochastic system.The symbol Δ represents a delay.

$$y_k = Cx_k + v_k$$

where the mean of stochastic vector w_k and v_k are zero,the covariance matrix :

$$E[\begin{pmatrix} w_p \\ v_p \end{pmatrix} \begin{pmatrix} w_p^T & v_p^T \end{pmatrix}] = \begin{pmatrix} Q & S \\ S^T & R \end{pmatrix}$$

Determine :

- The stochastic system order
- The system matrices A,C,Q,S,R such that the second moment of the model output and given data are equal.

5.2. *Properties of the stochastic systems*

In order to use the subspace Identification,some assumption need to be assumed,firstly the state process should be stationary which implies that the mean value and covariance matrix of states must converge to the finite value.For example,

$$E[x_k] = 0,$$

$$E[x_k(x_k)^T] = \Sigma,$$

Where the state covariance matrix Σ is not function of time instance k.Here another observation is that the system matrix A need to be stable which means the eigenvalues of the matrix are strictly inside the unit circle in Z plane.

Moreover,the state sequence is independent(uncorrelated)fr both white noises w_k and v_k.The mathematical description coul be written as below :

$$E[x_k(v_k)^T] = 0,$$

$$E[x_k(w_k)^T] = 0,$$

Now,the Lyapunov equation for state update is :

$$E[x_{k+1}(x_{k+1})^T] = \Sigma,$$

$$= E[(Ax_k + w_k)(Ax_k + w_k)^T]$$

$$= AE[x_k(x_k)^T]A^T + E[w_k(w_k)^T]$$

$$= A\Sigma A^T + Q$$

and now the output covariance(correlation) matrix is :

$$\Lambda_i = E[y_{k+i}(y_k)^T]$$

and the Λ_0 :

$$\Lambda_0 = E[y_k(y_k)^T]$$

$$= E[(Cx_k + v_k)(Cx_k + v_k)^T]$$

$$= CE[x_k(x_k)^T]C^T + E[v_k(v_k)^T]$$

$$= C\Sigma C^T + R$$

and the correlation between state update and current measurement :

$$G = E[x_{k+1}(y_k)^T],$$

$$= E[(Ax_k + w_k).(Cx_k + v_k)^T]$$

$$= AE[x_k(x_k)^T]C^T + E[w_k(v_k)^T]$$

$$= A\Sigma C^T + S$$

5.3. *GEOMETRIC PROPERTIIES OF STOCHASTIC SYSTEMS*

In this section the main theorem of the stochastic subspace identification problem will be illustrated.

5.4. *Main Theorem*

Base on this theorem,row space of state sequence $\hat{X}_i$ and the extended observability matrix Γ_i directly can be computed from the given data,without having any information about system matrices.The system matrices can be calculated from $\hat{X}_i$ or Γ_i.The procedure is depicted in figure 5.

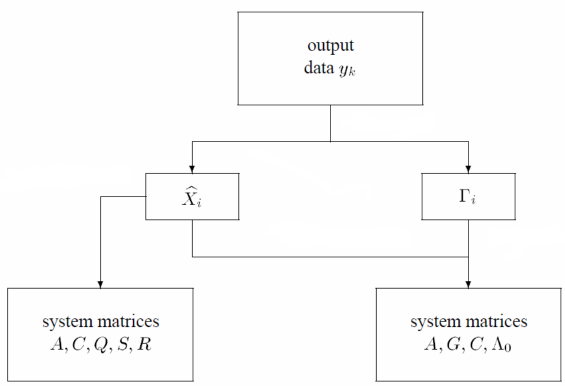

FIGURE 5: An overview of the stochastic subspace method..

Theorem of stochastic identification :

assumption :

- The process noise w_k and measurement noise v_k are not zero at the same time.
- The number of samples tends to infinity $j \to \infty$

Now the future data (Y_f) is projected to the past data (Y_p) and the result is :

$$O_i = Y_f/Y_p,$$

and once this matrix available the singular value decomposition gives :

$$O_i = \begin{pmatrix} U_1 & U_2 \end{pmatrix} \begin{pmatrix} S_1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} V_1^T \\ V_2^T \end{pmatrix} = U_1 S_1 V_1^T,$$

1- and we can prove that the matrix O_i is equal to the multiplication of the extended observability matrix and kalman filter states :

$$O_i = \Gamma_i.\hat{X}_i,$$

2- and by utilizing the SVD the order of the system in nothing but the rank of the singular value matrix and simply the number of singulars values which are not zero,and it can be proven based on the matrices dimension match.

3- The extended observability matrix and an extended controllability matrix Δ_i are :

$$\Gamma_i = U_1 S_1^{1/2}.T,$$

$$\Delta_i = \Gamma_i^{+}.\Phi_{[Y_f,Y_p]}.$$

4- The estimation of the state sequence can be found from the SVD :

$$\hat{X}_i = T^{-1}.S_1^{1/2} V_1^T.$$

Hence by far the states and the extended observability matrix calculated, The next is to calculate the system matrices.

5.5. COMPUTING THE SYSTEM MATRICES

In this section the system matrices A,C and Q,S,R can be calculated based on the states estimation sequence and extended observability matrix.

There are some algorithms that systematically calculate the system matrix according to the system states and the extended observability matrix,but we pick up only one of them and the explanation of the algorithm goes here :

stochastic algorithm :

- Solve the set of linear equation for A and C :

$$\begin{pmatrix} \hat{X_{i+1}} \\ Y_{i|i} \end{pmatrix} = \begin{pmatrix} A \\ C \end{pmatrix} \hat{X}_i + \begin{pmatrix} \varphi_w \\ \varphi_v \end{pmatrix}$$

  where $Y_{i|i}$ is a block Hankel matrix with only one row of the sampled data.The left side of this equation is completely known,because the state sequence already calculated and the residual vectors $\begin{pmatrix} \varphi_w \\ \varphi_v \end{pmatrix}$ are known,because they can be computed from kalman estimator equation.Thus,the only unknowns in this equation are the system matrices A and C,and since the set of residuals are uncorrelated from state sequence the natural way of solving this optimization problem is least square method.The solution of the cost function which tends to find the estimation of the parameters that minimize the variance of the prediction error.The solution is :

$$\begin{pmatrix} A \\ C \end{pmatrix} = \begin{pmatrix} \hat{X_{i+1}} \\ Y_{i|i} \end{pmatrix} \hat{X_i}^{+}$$

- Determine Q,S and R from :

$$\begin{pmatrix} Q & S \\ S^T & R \end{pmatrix} = E[\begin{pmatrix} \varphi_w \\ \varphi_v \end{pmatrix} . \begin{pmatrix} \varphi_w^T & \varphi_v^T \end{pmatrix}]$$

- The following step need to be done in order to find Σ,G and Λ_0 :

$$\Sigma = A\Sigma A^T + Q$$

$$\Lambda_0 = C\Sigma C^T + R$$

$$G = A\Sigma C^T + S$$

- Now we are able to calculate covariance of the state estimate and stationary Kalman gain :

$$P = APA^T + (G - APC^T)(\Lambda_0 - CPC^T)^{-1}(G - APC^T)^T$$

$$K = (G - APC^T)(\Lambda_0 - CPC^T)^{-1}$$

- And the forward innovation model is :

$$x_{k+1} = Ax_{k+1} + Ke_k$$

$$y_k = Cx_k + e_k$$

$$E[e_k(e_k)^T] = R$$

5.6. Conclusion

In this chapter we identified the space state representation by using the subspace method of the stochastic system.The states of the system and the extended observability matrix calculated by the projection theorem by means of the system output data and followed by calculation of system matrices.The least square method utilized in order to find the decision variables of cost function which are the system parameters.The success of subspace method will be examined at the end of the document.

5.7. COMBINED DETERMINISTIC-STOCHASTIC IDENTIFICATION

In this section we will describe the subspace identification for a system which the state space representation of system has deterministic control input u_k and both process noise w_k and measurement noise v_k.We utilize the results of the previous sections to derive the main theorem.,which illustrates that how the Kalman states can be calculated from the given input-output data.

The problem statement goes here :

Combined identification problem :

Given : j measurements of the control input u_k and output y_k generated by the unknown system of order n :

$$x_{k+1} = Ax_{k+1} + Bu_k + w_k$$

$$y_k = Cx_k + Du_k + v_k$$

where the process noise and measurement noise assumed to be white the covariance matrix :

$$E[\begin{pmatrix} w_p \\ v_p \end{pmatrix} \begin{pmatrix} w_p^T & v_p^T \end{pmatrix}] = \begin{pmatrix} Q & S \\ S^T & R \end{pmatrix}$$

Determine :

- The stochastic system order
- The system matrices A,B,C,D,Q,S,R such that the second moment of the model output and given data are equal.

5.8. Problem Description

combined subspace identification algorithm calculates state space models from given input-output data.Figure 6 graphically shows the combined subspace problem with unknown matrices.

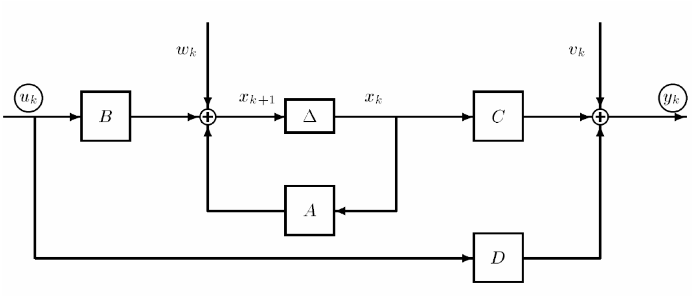

Figure 6: A LTI combined deterministic-stochastic system with input u_k,output y_k and statesx_k,described by the matrices A,B,C,D and covariance matrices Q,S,R.The symbol Σ illustrates a unit delay.

5.9. Notation

The system is divided in a deterministic and a stochastic subsystem,and due to a linear dynamic of the system, we can split the state(x_k) and outputy_k in a deterministic and stochastic component :
$x_k = x_k^d + x_k^s$,
$y_k = y_k^d + y_k^s$.
The deterministic state(x_k^d) and output(y_k^d) change by means of the deterministic inputu_k on the deterministic output :

$$x_{k+1}^d = Ax_k^d + Bu_k$$

$$y_k^d = Cx_k^d + Du_k$$

The stochastic state (x_k^s) and output(y_k)driven from stochastic system,which describes the effect of noise on the system output :

$$x_{k+1}^s = Ax_k^s + w_k$$
$$y_k^s = Cx_k^s + v_k$$

The state sequence is defined as :

$$X_i = \begin{pmatrix} x_i & x_{i+1} & \cdots & x_{i+j-2} & x_{i+j-1} \end{pmatrix}$$

The deterministic state sequence X_i^d and stochastic state sequence X_i^s are defined as :

$$X_i^d = \begin{pmatrix} x_i^d & x_{i+1}^d & \cdots & x_{i+j-2}^d & x_{i+j-1}^d \end{pmatrix}$$

$$X_i^s = \begin{pmatrix} x_i^s & x_{i+1}^s & \cdots & x_{i+j-2}^s & x_{i+j-1}^s \end{pmatrix}$$

in a similar way the pas and future deterministic and stochastic state sequences are defined as :

$$X_p^d = X_0^d \ , X_f^d = X_i^d,$$

$$X_p^s = X_0^s \ , X_f^s = X_i^s,$$

6. GEOMETRIC PROPERTIES OF COMBINED SYSTEMS

6.1. Matrix input-output equation

The matrix input-output equations for the combined system can be defined as following :

Theorem Combined matrix input-output equations

$$Y_p = \Gamma_i X_p^d + H_i^d U_p + Y_p^s,$$
$$Y_f = \Gamma_i X_f^d + H_i^d U_f + Y_f^s,$$
$$X_f^d = A^i X_p^d + \Delta_i^d U_p,$$

6.2. Main Theorem

This theorem calculates the sequence of state estimation and the extended observability matrix directly from the input-output data without having information of system matrices.The system matrices can be computed from $\hat{X}_i$ and Γ_i.The procedure of the theorem is illustrated in the figure 7.

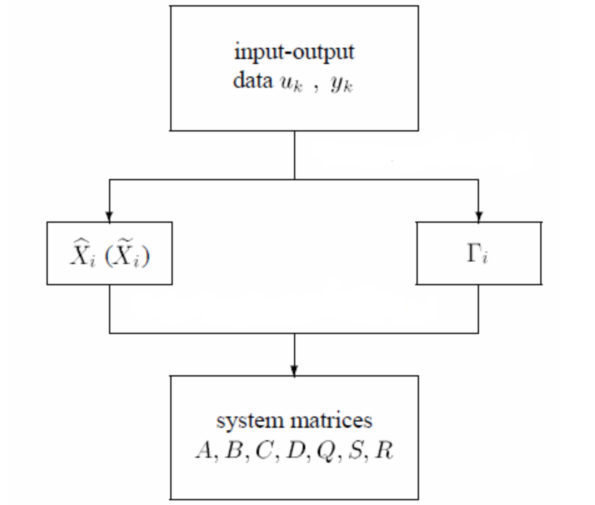

FIGURE 7: An overview of the combined deterministic-stochastic subspace identification procedure.

In order to explain the theorem there are some assumptions :

Assumptions :

- The deterministic control input u_k is uncorrelated with stochastic process noise w_k and measurement noise v_k .
- The number of measurement tends to infinity $j \rightarrow \infty$
- The process noise w_k and v_k are not zero identically.

and with O_i defined as the projection :

$$O_i = Y_f / U_f^{\perp}$$

and the singular value decomposition :

$$O_i = \begin{pmatrix} U_1 & U_2 \end{pmatrix} \begin{pmatrix} S_1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} V_1^T \\ V_2^T \end{pmatrix} = U_1 S_1 V_1^T,$$

Then the sequence of the calculating the states and extended observability matrix is :

1- The matrix O_i is equal to multiplication of the kalman filter states and extended observability matrix :

$$O_i = \Gamma_i . \hat{X}_i$$

2- The order of the system is equal of the most significant singular values,To put it simply,the minimum realization of the system is the most significant values of the singular values,due to noise some singular values are noisy and need to canceled out from the system because they are not the part of the system.

3- The extended observability matrix Γ_i can be calculated as follow :

$$\Gamma_i = U_1 S_1^{1/2} . T,$$

4- The estimation of the state sequence can be found from the SVD :

$$\hat{X}_i = T^{-1} . S_1^{1/2} V_1^T .$$

By far the states and extended observability matrix calculated and the next step is to fined the system matrices.

6.3. COMPUTING THE SYSTEM MATRICES

The algorithm of finding the system matrices explained in a brief here :

1- Calculate the projection of given data :

$$O_i = Y_f / U_f^{\perp}$$
$$O_{i+1} = Y_f^- / U_f^{\perp}$$

2- Calculate the **SVD** of O_i :

$$O_i = U_1 S_1 V_1^T,$$

3- Determine the minimum system realization by considering the most significant singular values in S matrix to obtain U_1 and S_1.

4- Determine Γ_i and Γ_{i-1} as :

$\Gamma_i = U_1 S_1^{1/2}$
$\Gamma_{i-1} = \Gamma_i$

5- Determine the estimation of the states :

$$\hat{X}_i = \Gamma_i^{+} . O_i$$
$$\hat{X_{i+1}} = \Gamma_{i-1}^{+} . O_{i+1}$$

6- Solve the set of linear equations for A,B,C and D :

$$\begin{pmatrix} \hat{X_{i+1}} \\ Y_{i|i} \end{pmatrix} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} \hat{X}_i \\ U_{i|i} \end{pmatrix} + \begin{pmatrix} \varphi_w \\ \varphi_v \end{pmatrix}$$

The only thing here is that the parameter of the linear regression could be calculated from the least square sense.

7- And the last step is to find Q,S and R from the kalman filter residuals :

$$E[\begin{pmatrix} w_p \\ v_p \end{pmatrix} \begin{pmatrix} w_p^T & v_p^T \end{pmatrix}] = \begin{pmatrix} Q & S \\ S^T & R \end{pmatrix}$$

6.4. CONCLUSIONS

In this chapter we considered the subspace identifications of a combined system with both stochastic and deterministic inputs.The system matrices are estimated based on the least square sense and having the knowledge of Kalman filter states and extended observability matrix.The algorithm merely utilizes the input-output data to estimate the states and extended observability matrix and follows the instructions to reach the system matrices.The simulation will be done in the experimental chapter.

7. Simulation

In this section we will perform some simulation for different type of the systems, as we described in previous chapters and discuss the results and compare the results with ARX estimation.The order of simulations are as same as the materials presented in this report.

7.1. Deterministic simulation

Firstly We performed the deterministic simulation for a system with ARMAX representation and the noise set to the zero.The system description is :

$$A(q^{-1})y_k = B(q^{-1})u_k + C(q^{-1})e_k$$

The Set of system parameters defined as :

$$A = [1, -0.98, 0.4, -0.4],$$

$$B = [0, 2],$$

$$C = [1, 0.3],$$

and the input stochastic assumed to be white with zero average and variance of σ^2.This is the system we will use for driven the data and the data will use in order to estimate the system parameters in input-output description with ARMAX command and also the state space description will be estimated base on **N4SID** algorithm.

In the deterministic simulation the noise variance set to the zero and the system output is only driven based on the deterministic PRBS control input which takes values of -1 and 1.By applying this input to the system and measuring the output we can build a data for this system and given data will be stored in a vector.

The subspace algorithm only gives us one description of state space representation,because the system can be written in an infinity state space form.By using the **N4SID** command we get the system matrices as following :

$$A = \begin{pmatrix} 0,4496 & 0,0436 & -0,47 \\ 0,8657 & 0,1241 & 0,0640 \\ 0,19764 & -0,8732 & 0,4062 \end{pmatrix}$$

$$B = \begin{pmatrix} -0,0277 \\ 0,0172 \\ -0,0057 \end{pmatrix}$$

$$C = \begin{pmatrix} -158,6776 & -110,7672 & 84,2580 \end{pmatrix}$$

$$D = 0$$

$$K = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

The system representation in state space form based on the **N4SID** is as below in general :

$$x_{k+1} = Ax_k + Bu_k + Ke_k$$
$$y_k = Cx_k + Du_k + e_k$$

where **A** is the system matrix,**B** is the input matrix , **K** is the Kalman gain,**C** is the output matrix,**D** is direct input matrix and e_k is the measurement noise which assumed to be white and uncorrelated from the states.

As it was expected the kalman gain estimated as a zero vector because we have a deterministic system. The comparison between the model which is given from subspace method and system model with defined parameter is performed in order to show the accuracy of the estimation and validation of the system matrices which estimated by the subspace algorithm.The result of comparison is here :

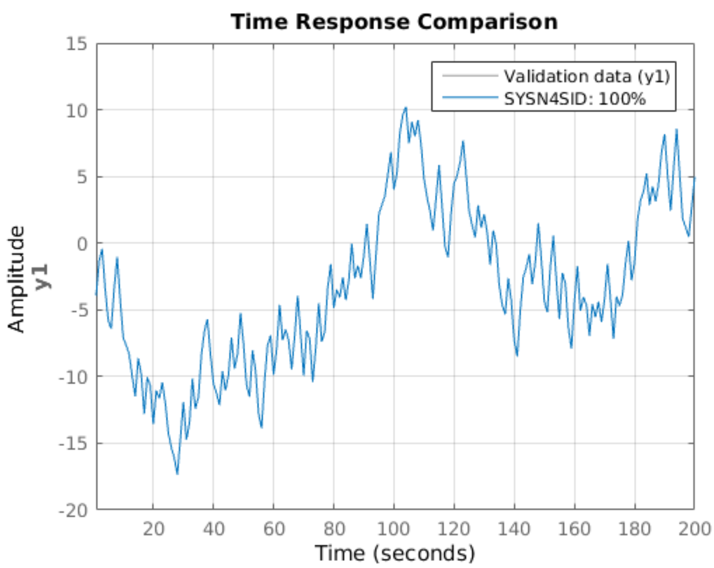

FIGURE 8: Validation of the estimated system matrices with test data in order to compare the real system output and estimated model output.

It should be noted that there are 2 set of data.One set used so as to find the system matrices namely train or estimation set and the second one used in order to check the consistency and validity of the model namely test data.As it could be expected the PRBS is a random input that uncorrelated with states and subsequently uncorrelated from the output,as far as we perform another experiment we will get another state space representation which could be transformed to the similar transfer function.

In summary,although the estimated space state of the real system would vary in every realization,the model transfer function remains same and as it could be seen the output of the system and model are match in all of the samples with probability of 100 percent.

7.2. Stochastic simulation

In this section a pure stochastic systems is examined. The input, u, is vanished, resulting in the following ARMA model :

$$A(q^{-1})y_k = C(q^{-1})e_k$$

$$A = [1, -0.98, 0.4, -0.4],$$

$$C = [1, 0.3],$$

In this case the only input to the system is the noise which is normal distributed with a variance of 0.1. Again, the N4SID algorithm is used to extract a state space model. The result is the following system matrices :

$$A = \begin{pmatrix} 0,4604 & 0,2353 & 0.4227 \\ -0.7648 & -0.1035 & 0.618 \\ 0.4513 & -0.382 & 0.623 \end{pmatrix}$$
$$B = \begin{pmatrix} 0 \\ -7.16^{-16} \\ 2.584e^{-16} \end{pmatrix}$$
$$C = \begin{pmatrix} -9.55 & 0.7507 & 7.606 \end{pmatrix}$$

$$D = \begin{pmatrix} 0 \end{pmatrix}$$

$$K = \begin{pmatrix} -2.957 \\ 1.22 \\ -0.644 \end{pmatrix}$$

The results obtained in the B matrices can be considered as zero since the order of magnitude is e^{-16}. This result is expected as the simulated system had no input. On the other hand, as opposed to the deterministic case, the matrix K is non-zero since the processed noise is modeled.

The performance of the estimation, 81% ,is reflected in the figure 9. For assessing this performance, the estimated system is fed with exactly the same simulated noise and then compared with the simulated output. One can expect to have the same performance as in the deterministic case, since we are using a noise signal that we already know. However, N4SID works differently for each case. Here an optimal fitting is needed, but not in the deterministic system.

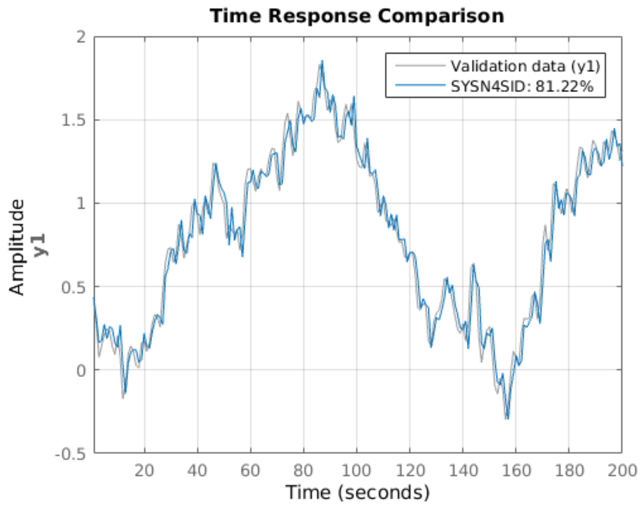

FIGURE 9: Stochastic with s2 = 0.01

7.3. Combined Stochastic-Deterministic Simulation

In this section the system has two inputs,one system is a deterministic control input and the other is the stochastic input.It should be noted that the current system identification is performed in open loop and thus there is no correlation between control input and states and system output.Knowing this fact,the system realization carried out through a given description as :

$$A(q^{-1})y_k = B(q^{-1})u_k + C(q^{-1})e_k$$

The Set of system parameters defined as :

$$A = [1, -0.98, 0.4, -0.4],$$

$$B = [0, 2],$$

$$C = [1, 0.3],$$

and the input stochastic assumed to be white with zero average and variance of σ^2.This is the system we will use for driven the data and the data will use in order to estimate the system parameters in input-output description with ARMAX command and also the state space description will be estimated base on **N4SID** algorithm.The σ^2 is set to different values and the effect of noise on the estimation will examine in detail.

In combined system it is expected the **B** and **K** vectors are not going to be zero,since the system output driven by both deterministic and stochastic input with a variance different from zero.

We performed a simulation for a given system with PRBS deterministic input and stochastic input with $\sigma^2 = 0.01$ and repeat the experiment for different value of the σ^2

to examine the effect of noise on the system estimation.It is also assumed the level of control input is considerably high from level of noise,and the intuition for this is that we prefer to excite the system in direction of control input to understand the effect of the control input on the system output,while if the level of noise is much than control than the existence of the control is meaningless du to the fact the noise is able to bring the system to any position and the level of control is not capable to change the dynamic of the noise in direction of our objection.In other words,if the noise variance is high,we do not have any control so to speak and the output behave randomly.Hence,by using this fact,we limit the noise variance up to 0.1 and perform the subspace identification.The results are here :
$\sigma^2 = 0.01$

$$A = \begin{pmatrix} 0,9639 & 0,1688 & -0,0163 \\ -0,0877 & -0,2919 & -0,763 \\ -0,2514 & 0,6055 & 0,30941 \end{pmatrix}$$

$$B = \begin{pmatrix} 0,0088 \\ 0,0198 \\ 0,0245 \end{pmatrix}$$

$$C = \begin{pmatrix} 190,2519 & 11,3890 & 3,4551 \end{pmatrix}$$

$$D = 0$$

$$K = \begin{pmatrix} 0,0057 \\ -0,0093 \\ 0,0098 \end{pmatrix}$$

As it can be seen there are three poles in the system description and the order of the system in state space representation is also 3.The Kalman gain is not zero due to noise.Another observation done to check the consistency of estimation .The result are in figure 10.

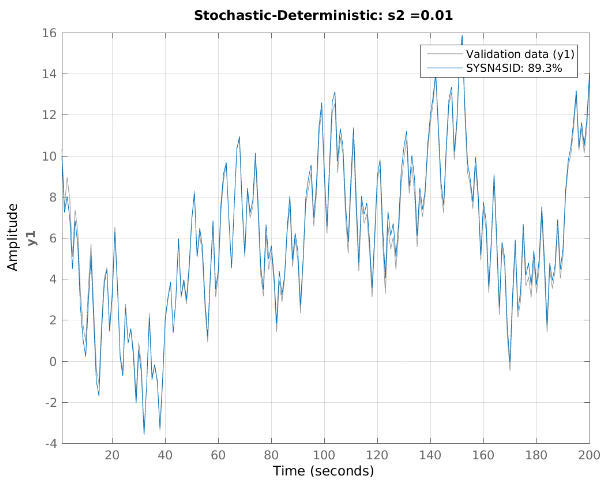

FIGURE 10: compare test data and model output for s2 = 0.01

It should be noted that due to the low level of the noise the match between model output and test data is around 90 percent of the total samples.

Another experiment performed for $\sigma^2 = 0.05$ and the results are :

$\sigma^2 = 0.05$

$$A = \begin{pmatrix} 0,9591 & 0,1944 & 0,0394 \\ -0,1091 & -0,2815 & 0,72899 \\ 0,2458 & -0,6641 & 0,3139 \end{pmatrix}$$

$$B = \begin{pmatrix} 0,0096 \\ 0,0221 \\ -0,0209 \end{pmatrix}$$

$$C = \begin{pmatrix} 175,5991 & 13,0678 & -0,6968 \end{pmatrix}$$

$$D = 0$$

$$K = \begin{pmatrix} 0,006 \\ -0,0028 \\ -0,0115 \end{pmatrix}$$

Hence,again the system estimated with 3 states and due to the noise and deterministic input both **B** and **K** matrix are not identically zero.

The comparison performed so as to examine the estimation.The result in figure 11.

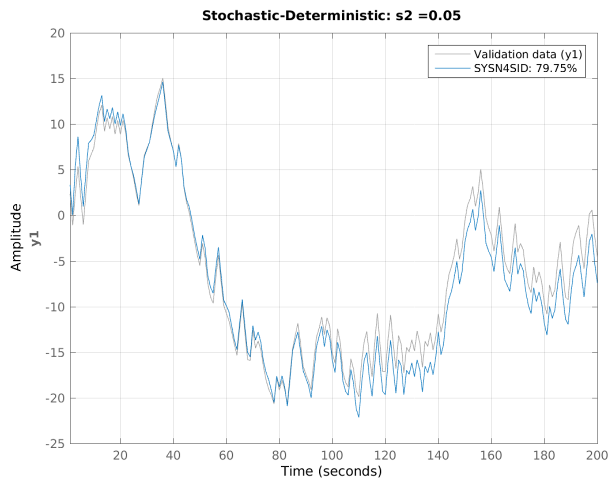

FIGURE 11: compare test data and model output for s2 = 0.05

It is expected that the level of noise affect the estimation of the parameter. Because from least square sense the optimal estimation can not in charge of the noise and thus the number of samples which are match together is going to be decrease which could be seen from figure refs2 = 0.05.

The last experiment have been done for variance of the noise equal to 0.1.In this case we expect that the estimation of the system matrices affects highly with noise.Because the level of deterministic input and noise is in some point equal.The estimation goes here :

$$\sigma^2 = 0.1$$

$$A = \begin{pmatrix} 0,9609 & -0,1631 & -0,0038 \\ 0,0624 & -0,2528 & 0,8537 \\ -0,2479 & -0,53590,2564 \end{pmatrix}$$

$$B = \begin{pmatrix} -0,0085 \\ 0,0166 \\ -0,0245 \end{pmatrix}$$

$$C = \begin{pmatrix} -196,0541 & 10,4169 & -4,91893825957348 \end{pmatrix}$$

$$D = 0$$

$$K = \begin{pmatrix} -0,0056 \\ -0,0095 \\ -0,010 \end{pmatrix}$$

And the comparison between test data and model out put shown in figure 12

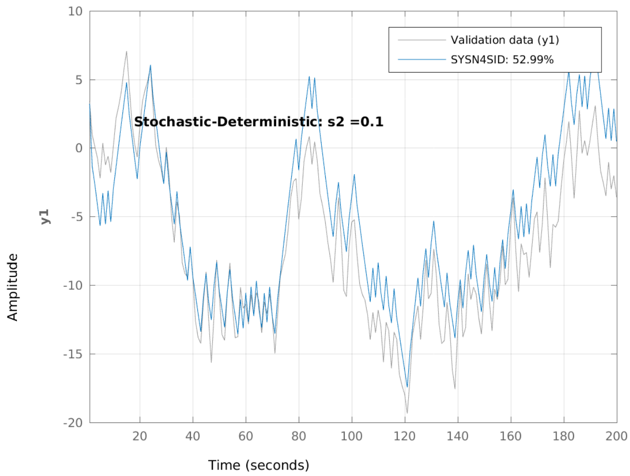

Figure 12: Combined simulation with s2 = 0.1

As it Expected the fifty percent of sampled between model and test data are coincide to each other.The reason is that the effect of noise on the data is huge and we would like to extract the data without noise and this task not easy because the deterministic input and stochastic one excite the system with a same level and the information will spread out over the noise and deterministic direction and extracting the true matrices is affected with noise.

8. Conclusion

In this report the subspace identification algorithm examined for different systems including deterministic,stochastic and combined system. In all case the subspace has solution for system matrices, and the system matrices estimated only base on the given data. The level of the noise will affect the estimation, because the system matrices are the solution of the set of liner equation which are affected by kalman filter prediction error, and once the noise variance is high it implies that the prediction error will increase and that followed by the fact that the measurements are noisy and we should rely on the model rather than the output itself, and that is why the output from model and measurements has big difference.

Références

Katayama, T., 2006. Subspace methods for system identification. Springer Science & Business Media.

Lung, L., 1987. System identification : Theory for the user.

Qin, S. J., 2006. An overview of subspace identification. Computers & chemical engineering 30 (10), 1502–1513.

Van Overschee, P., De Moor, B., 1994. N4sid : Subspace algorithms for the identification of combined deterministic-stochastic systems. Automatica 30 (1), 75–93.

Van Overschee, P., De Moor, B., 1995. A unifying theorem for three subspace system identification algorithms. Automatica 31 (12), 1853–1864.

Van Overschee, P., De Moor, B., 2012. Subspace identification for linear systems : Theory—Implementation—Applications. Springer Science & Business Media.

Verhaegen, M., Dewilde, P., 1992. Subspace model identification part 1. the output-error state-space model identification class of algorithms. International journal of control 56 (5), 1187–1210.