Subspace identification with multiple data sets

Laurent Duchesne,* Eric Feron,† James D. Paduano‡ and Marty Brenner§

December 12, 1995

Abstract

Most existing subspace identification algorithms assume that a single input to output data set is available. Motivated by a real life problem on the F18-SRA experimental aircraft, we show how these algorithms are readily adapted to handle multiple data sets. We show by means of an example the relevance of such an improvement.

1   Introduction

Identification problems occur as soon as some practical engineering is done. For example, in control design, it is necessary to have a reliable model in order to design an efficient control law. Very often, time domain input to output data are available and a state space model can then be estimated by an identification algorithm.

Subspace identification methods have been initiated by the works of Kung [1] and Juang and Pappa [13]. A variety of new methods has then emerged [5], [6], [10] and [11] identifying the system in the time domain, and also [8] in the frequency domain. Currently available subspace identification algorithms assume that plant identification is based on a single experiment, where only a single input to output data set is available. There are, however, many cases for which data collection cannot be done all at once, and experiments must be segmented possibly over a period of several days, leading to the collection of many data sets all related to the same dynamical system, but with possibly different initial conditions. This is typically the case, for example, when attempting to identify the flexible dynamics of the F18 Systems Research Aircraft (SRA) at NASA Dryden Flight Research Center, where several data sets generated through many flights at the flight conditions (altitude, Mach number and dynamic pressure) are available. These data sets are extremely noisy, such that it is highly desirable to use all data sequences at once to obtain the best possible identified model.

In this paper, we describe how existing subspace identification algorithms may be readily adapted to handle multiple data sets. We then show by means of an example the efficiency of the proposed scheme, as compared to more ad hoc solutions, such as simply concatenating the data.

---

*Graduate student, Massachusetts Institute of Technology

†Assistant professor, Massachusetts Institute of Technology

‡Assistant professor, Massachusetts Institute of Technology

§Research engineer, NASA Dryden Flight Research Center

2 Subspace algorithm

2.1 Notations

The goal of subspace identification is to find a linear, time invariant, finite dimensional state space realization

$$x_{k+1} = Ax_k + Bu_k \tag{1}$$

$$y_k = Cx_k + Du_k,$$

where $A \in \Re^{n \times n}$, $B \in \Re^{n \times m}$, $C \in \Re^{l \times n}$, $D \in \Re^{l \times m}$, based on the knowledge of specific sequences $u = [u_1, ..., u_p]$, $y = [y_1, ..., y_p]$.

The following notation is used:

The block Hankel input and output matrices

$$Y_h(k,i,j) = \begin{bmatrix} y_k & y_{k+1} & \cdots & y_{k+j-1} \\ y_{k+1} & y_{k+2} & \cdots & y_{k+j} \\ \cdots & \cdots & \cdots & \cdots \\ y_{k+i-1} & y_{k+i} & \cdots & y_{k+j+i-2} \end{bmatrix}$$

and

$$U_h(k,i,j) = \begin{bmatrix} u_k & u_{k+1} & \cdots & u_{k+j-1} \\ u_{k+1} & u_{k+2} & \cdots & u_{k+j} \\ \cdots & \cdots & \cdots & \cdots \\ u_{k+i-1} & u_{k+i} & \cdots & u_{k+j+i-2} \end{bmatrix}.$$

We also introduce the extended observablity matrix

$$\Gamma = \begin{bmatrix} C \\ CA \\ \cdots \\ CA^{i-1} \end{bmatrix}.$$

j

We define the lower block triangular Toeplitz matrix

$$H_{tl} = \begin{bmatrix} D & 0 & 0 & \cdots & 0 \\ CB & D & 0 & \cdots & 0 \\ CAB & CB & D & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ CA^{i-2}B & CA^{i-3}B & CA^{i-4}B & \cdots & D \end{bmatrix}.$$

Finally, the state matrix is defined as

$$X = \begin{bmatrix} x_k & x_{k+1} & \cdots & x_{k+j-1} \end{bmatrix}.$$

It is then easy to see that

$$Y_h(k,i,j) = \Gamma X + H_{tl}U_h. \tag{2}$$

2.2 Step by step procedure

The algorithm to perform the identification with multiple data sets has similarities with the classical, single data set algorithm. Therefore, the step by step procedure of a subspace identification algorithm with one data set is now explained. The example of the deterministic identification (i.e. no noise is corrupting the data) is specified in more detail.

Step 1: find a matrix P that satisfies an equation of the form

$$P = \Gamma Q, \tag{3}$$

where Γ is the extended observability matrix and such that rank(P)=rank(Γ)=n.

In practice, the existence of noise makes it impossible to obtain equation (3) exactly. Any subspace method extracts a matrix P from the input to output data which is optimal in the sense defined by the method which depends mainly on the assumption made on the noise. Depending on the subspace method that is chosen, different computations of this matrix P are possible, all leading to different results.

In the case of a deterministic system, this can be done by post multiplying equation (2) by a matrix $U_h^\perp$ that satisfies $U_h U_h^\perp = 0$. We then obtain $P = Y_h U_h^\perp$. However, the rank of the matrix P may not be equal to the order of the system. This phenomenon is known as rank cancellation and its probability of occuring decreases when the number of rows in Y_h increases.

Step2: perform a singular value decomposition of P

$$P = USV,$$

where $S = \begin{pmatrix} S_1 & 0 \\ 0 & 0 \end{pmatrix}$ and $U = (U_1\ U_2)$ such that U_1 is the first n columns of U.

Note that S_1 is an $n \times n$ matrix. With equation (3), we can see that there must exist a full rank $n \times n$ matrix T such that

$$U_1 = \Gamma T.$$

Let us now use the following notation: if M is an $m \times n$ matrix, $\overline{M}$ (resp. $\underline{M}$) will be the matrix with a reduced number of rows, obtained from M by omitting the first (resp. last) l rows, where l is the number of output of the system.

Step 3: $A = \overline{U_1}^\dagger \underline{U_1}$ and C is equal to the first block of U_1, where $\overline{U_1}^\dagger$ denote the pseudo-inverse of $\overline{U_1}$.

Using the structure of the extended observability matrix, it is clear that

$$\overline{\Gamma} = \underline{\Gamma} A$$

$$\underline{U_1} = \underline{\Gamma} T \ , \ \overline{U_1} = \overline{\Gamma} T$$

$$\underline{U_1} T^{-1} = \overline{U_1} T^{-1} A.$$

This can also be written as

$$\underline{U_1} = \overline{U_1} \Psi \ , \ \Psi = T^{-1} A T.$$

We have proven that Ψ is a matrix similar to A which is what we wanted originally.

Step 4: Use a least square method to compute B and D.

We can pre multiply equation (2) by $\Gamma^\perp$ such that $\Gamma^\perp \Gamma = 0$, and post multiply it by the pseudo-inverse of U_h. By using the stucture of the matrix H_{tl}, we get

$$\Gamma^\perp Y_h U_h^\dagger = \Gamma^\perp \begin{bmatrix} I & 0 \\ 0 & \underline{\Gamma} \end{bmatrix} \begin{bmatrix} D \\ B \end{bmatrix},$$

leading to

$$\begin{bmatrix} D \\ B \end{bmatrix} = (\Gamma^{\perp} \begin{bmatrix} I & 0 \\ 0 & \underline{\Gamma} \end{bmatrix})^{\dagger} \Gamma^{\perp} Y_h U_h^{\dagger}.$$

3 Subspace algorithm with multiple data sets

We will now assume that we have collected two data sets (the generalization to n sets of data is very simple and is omitted for notation purposes), $u_1(k)$, $y_1(k)$ and $u_2(k)$, $y_2(k)$ and the following equations are satisfied

$$Y_1 = \Gamma X_1 + H U_1 \tag{4}$$

$$Y_2 = \Gamma X_2 + H U_2.$$

Let us now explain how does the original algorithm has to be modified in order to handle multiple data sets.

Step 1: Find two matrices P_1 and P_2 that satisfy $P_i = \Gamma Q_i$, for $i = 1, 2$, where Γ is the extended observability matrix.
Actually, this step is similar to the first step of initial algorithm, but we need to realize it for each data sets. For example, if we want to use the noise free method, we should proceed as follow

$$P_1 = Y_1 U_1^{\perp} = \Gamma(X_1 U_1^{\perp})$$

$$P_2 = Y_2 U_2^{\perp} = \Gamma(X_2 U_2^{\perp}).$$

The main modification of the algorithm is to compute an additional step at this point.
Step 1bis: Compute the matrix $\Phi = [P_1 \; P_2]$.
This matrix Φ satisfies

$$\Phi = \Gamma[Q_1 \; Q_2],$$

which is exactly the same property as the matrix P of the first step of the original algorithm.

The step 2 to 4 are exactly the same as in the original algorithm, where the matrix Φ replaces the matrix P.

3.1 Remarks

If we append the two data sets at the the beginning of the experiment and use the single data set algorithm, the Hankel matrix Y_h will have some rows that have no physical meanings. At the junction of the two data sets, it appears some rows that contain some data from the first experiment and some from the second one. Equation (2) would then not be satified anymore. If the classical algorithm were used, those rows would be considered as part of the dynamic of the system. On the other hand, the proposed method avoid this problem by removing those undesirable rows. The algorithm treats those data sets in parallel, and concatenate them only when performing a least square fit.
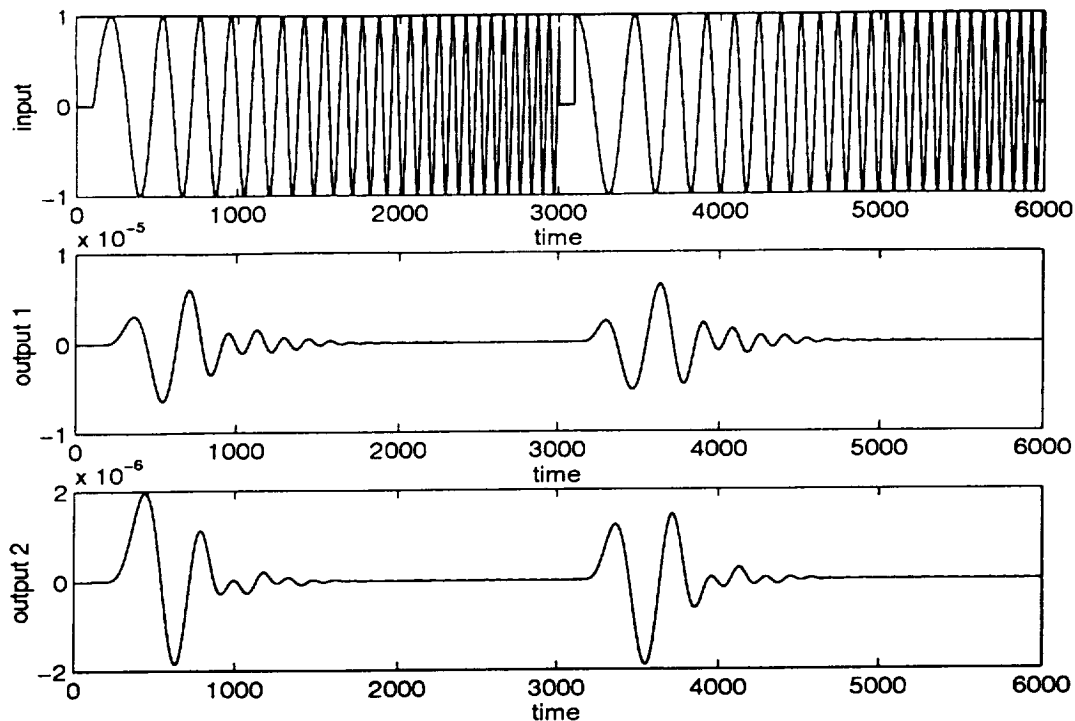
4

Figure 1: Concatenation of two simulations made on a 8 order system with two different inputs and no noise.

4 Application

An example has been computed to show the relevance of such an improvement. It is an order 8 system with one input and two outputs and whose state space representation can be found in the Appendix. The system has been excited separately by two sets of linear frequency sweeps. Here again, the choice of such inputs has been motivated by some practical concerns since linear sweeps were the only available excitations at our disposal to identify the structural dynamics of the F18-SRA. The following formula for the inputs has been used from $k = 100$ to 3000, the first 100 points were set to 0

$$e1(k) = \sin(2\pi(5 + 20k/3000)(k - 100)/3000)$$

$$e2(k) = \cos(2\pi(5 + 20k/3000)(k - 100)/3000).$$

The simulation of this system has been realized for each input and the two data sets were appended together. The plot of the input and outputs can be seen in figure (1) and we can notice that the discontinuity at the junction of the two data sets is very small. We then tried to identify the system with a subspace identification algorithm (we used N4SID) with a number of blocks i in the Hankel matrix equal to 14, 15 and 16. For $i = 15$, the original system was perfectly recovered. The problem came when we tried to use an $i = 14$ or 16 where some of the eigenvalues have become unstable as seen on table 1. Other i have been tested from 10 to 30 and the algorithm failed in about 70 % of the cases. Eventhough the identification was accurate for a certain value of i, this · remains a problem because this number does not have a real physical meaning since it is just an over estimation of the order of the system in order to obtain a sufficiant rank in the Hankel matrix. On figure 2, we have simulated the system with the concatenated input and plotted the outputs of this experiment. If we compare to the outputs shown on figure 1, we note that the difference between the two tests is very small. However, the identification with those data recovered the right
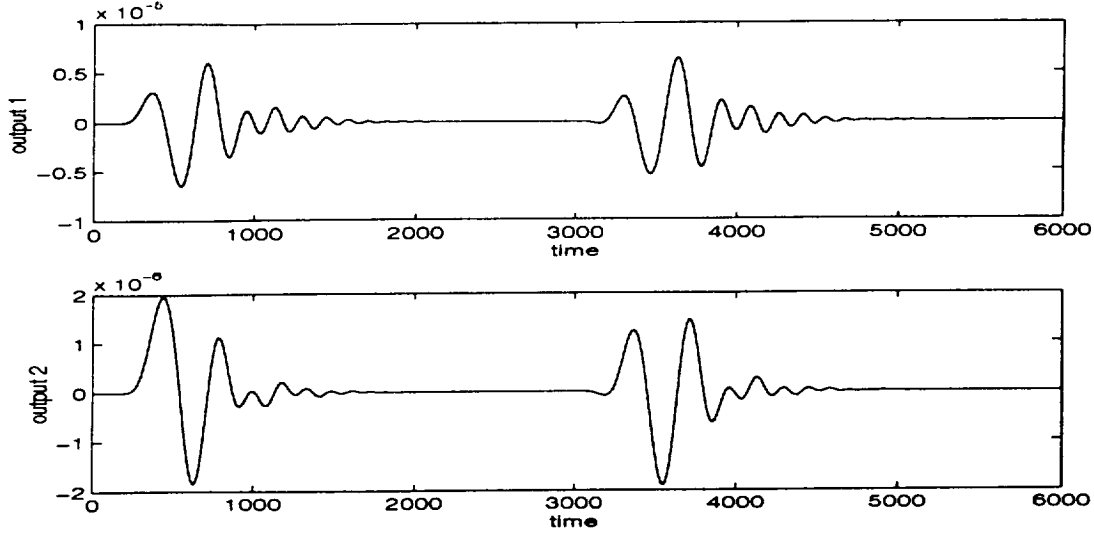
5

Figure 2: Simulation made with the same system as in figure 1 but with the concatenated input.

original eigenvalues	eigenvalues calculated by concatenating the two sets		eigenvalues calculated with this new algorithm
	$i = 14$	$i = 16$	
0.9893 + 0.0396i	.9977+.0100i	1.0133 + 0.0614i	0.9893 + 0.0396i
0.9893 - 0.0396i	.9977-.0100i	1.0133 - 0.0614i	0.9893 - 0.0396i
0.9799 + 0.0245i	.9960+.0200i	0.9969 + 0.0377i	0.9799 + 0.0245i
0.9799 - 0.0245i	.9960-.0200i	0.9969 - 0.0377i	0.9799 - 0.0245i
0.9949 + 0.0149i	.9944+.0386i	0.9985 + 0.0098i	0.9949 + 0.0149i
0.9949 - 0.0149i	.9944-.0386i	0.9985 - 0.0098i	0.9949 - 0.0149i
0.9753	.9454+.1431i	0.9976 + 0.0195i	0.9754
0.9851	.9454-.1431i	0.9976 - 0.0195i	0.9850

Table 1: Eigenvalues of the identified model.

eigenvalues. This shows that the identification procedure is very sensitive to data corruption due to concatenating the two data sets.

To show that this problem does not come from the kind of input that we have chosen, we tried to identify the system with each data sets separately. The original system was recovered with any i that we picked for both data sets.

Let us now apply the identification method explained in this paper to identify the exact same data. The modification of this algorithm has also been made on N4SID in order to show that the improvement of the results is only due to this modification. As shown in table 1, the result of this identification was very accurate. The eigenvalues has been fitted with an error lower than 0.1%.

The question of determining the order of the system is also a major issue in identification methods. In practice, the order is also an unknown that need to be calculated. In many subspace identification, the singular values of the matrix P (step 1) are plotted and the user has to decide the order of the system. If there is a jump in the singular values, the order is determined by the number of singular values to the left of this jump. If there is no detectable jump, then the user just has to guess, by his knowledge of the system, what the order is. Figure 2 shows the plots that are obtain
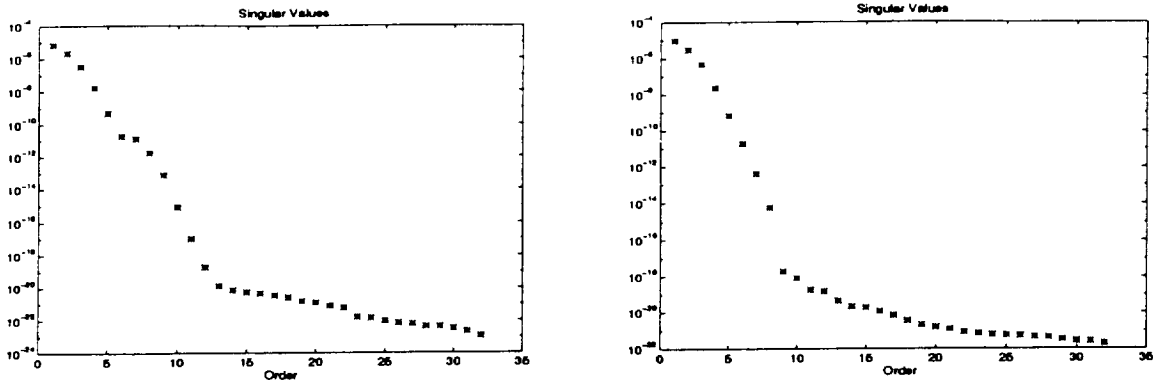
Figure 3: Singular values to estimate the order of the system. The left picture happens when concatenating the data, the right one is with the new scheme.

using both procedures with an number i of blocks in the Hankel matrix of 16. We can notice that it is impossible to determine the order of the system when the data has been concatenated. On the other hand, there is a gap of 3 orders of magnitude for the other procedure.

5 Conclusion

In this paper, motivated by a problem with simple concatenation of data sets using subspace identification algorithm, we described a way to handle multiple data sets when using subspace identification. The step by step procedure details more specifically a deterministic identification problem by the same idea that can be used for any subspace identification technique since the only assumption that is needed remains in the structure of the extended observability matrix.

6 Appendix

State space representation of the example chosen to show the relevance of the scheme described in this paper

$$
A = \begin{bmatrix}
0.89 & -1.5 & -13.1 & -81.9 & -353.5 & -1013.8 & -1957.5 & -1977.6 \\
0.005 & 1 & 0 & -.2 & -.9 & -2.6 & -5 & -5 \\
0 & 0.005 & 1 & 0 & 0 & 0 & -0.0084 & -0.0085 \\
0 & 0 & 0.005 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0.005 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0.005 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0.005 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0.005 & 1
\end{bmatrix}
$$

7

$$B = \begin{bmatrix} .0047 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$$C = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

$$D = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

References

[1] S. Kung. A new identification and model reduction algorithm via singular value decompositions. *12 th Asilomar conference on circuits, systems and computers*, pages p. 705–714, 1978.

[2] W. Larimore. Canonical variate analysis in identification filtering, and adaptative control. *Proceeding 29 th Conference on Decision and Control*, pages p. 596–604, 1990.

[3] B. De Moor. Numerical algorithms for state space subspace system identification.

[4] P. Van Overschee and B. De Moor . N4SID: Subspace algorithm for identification of combined deterministic stochastic systems. *Automatica*, 30(1):p. 75–93, 1994.

[5] Van Overschee P. Subspace Identification: Theory - Implementation - Applications. *PhD thesis, Department of Electrical Engineering, Katholieke Universiteit Leuven, Belgium*, February 1995.

[6] B. De Moor, J. Vandewalle, M. Moonen, L. Vandenberghe, and P. Van Mieghem. A geometrical strategy for the identification of state space models of linear multivariable systems with singular value decomposition. *Symposium on Identification and System Parameter Estimation*, pages p. 700–704, August 1988.

[7] P. Van Overschee and B. De Moor. Subspace algorithms for the stochastic identification problem. *Automatica*, 29(3):pp. 649–660, 1993.

[8] T. McKelvey and H. Akçay. An efficient frequency domain state-space identification algorithm. *Proceeding 33rd Conference on Decision and Control*, p3359-3364, Dec 1994.

[9] Y. M. Cho. Fast subspace based system identification: theory and practice. *PhD thesis Stanford University*, August 1993

[10] Y. M. Cho, G. Xu and T. Kailath. Fast identification of state space models via exploitation of displacement structures. *IEEE Transactions on Automatic Control*, p. 2094-2017, 1994

[11] K. Liu and R. E. Skelton. Q-markov covariance equivalent realization and its application to flexible structure identification. *AIAA Journal of Guidance, Control and Dynamics*, 16(2):308-319, 1993.

[12] L. Ljung *System Identification: Theory for the User*. Prentice-Hall, Englewood Cliffs, New Jersey, 1987.

[13] J. N. Juang and R. S. Pappa. An eigensystem realization algorithm for modal parameter identification and model reduction. *Journal of Guidance, Control and Dynamics*, 8(5):620-627, 1985.

[14] P. Faurre. Stochastic realization algorithm. In R. Mehra and D. Lainiotis System identification: advances and case studies Academic Press N.Y. (1976)

[15] H. Akaike. Markovian representation of stochastic process by canonical variables. *SIAM J. of Control* (13) p166-173, 1975.