89-02C

89-02a

De Moor B., Moonen M., Vandenberghe L., Vandewalle J., ``Identification of linear state space models with singular value decomposition using canonical correlation concepts", in *'SVD and Signal Processing: Algorithms, Applications and Architectures'*, E. Deprettere (ed.), Elsevier Science Publishers B.V. (North-Holland), 1988, pp. 161-169., Lirias number: 181512.

IDENTIFICATION OF LINEAR STATE SPACE MODELS WITH SINGULAR VALUE DECOMPOSITION USING CANONICAL CORRELATION CONCEPTS

Bart De Moor, Marc Moonen*, Lieven Vandenberghe*, Joos Vandewalle
ESAT Katholieke Universiteit Leuven
K.Mercierlaan 94, 3030 Heverlee, Belgium
tel 016/22 09 31 telex 25941 elekul fax 32/16/221855

Abstract

In this paper, a geometrically inspired algorithm is derived for identification of state space models for multivariable linear time-invariant systems using noisy input-output measurements. The algorithm contains two conceptual steps which allow a robust implementation using SVD techniques :

1. Using canonical correlation, first a state vector sequence is realized.
2. The system matrices are then obtained at once from the total linear least squares solution of a set of linear equations.

These results are applied to the identification of an industrial plant.

1 Introduction

Selection and identification from input-output measurements of appropriate mathematical representations are of central importance in system analysis, design and control. A state space model is a most obvious choice for a mathematical representation because of its widespread use in system theory and control. Still, reliable general purpose state space identification schemes have not become standard tools so far, mostly due to the computational complexity involved.

The concept of canonical correlation however, together with the total linear least squares approach, allows us to set up a new most promising identification scheme, attractive due to its utmost simplicity. Since singular value decomposition is the main tool throughout all calculations, this identification scheme is expected to outperform most present-day used schemes, as far as robustness is concerned. Moreover, recent developments in fast algorithms for structured matrices, give promising perspectives to obtain a fast and yet numerically stable identification scheme.

In section 2, the main principles of canonical correlation are outlined. In section 3, useful properties of dynamic systems are summarized, which -together with the canonical correlation

*Supported by the N.F.W.O. (Belgian National Fund for Scientific Research)

concept- are used in section 4 to show how a sequence of state vectors can be calculated. The system matrices are then identified by solving a set of linear equations using TLLS (Section 5). Finally, section 6 gives an example of an industrial plant identification.

2 Canonical correlation

Canonical correlation analysis, introduced for the first time by H.Hotelling, essentially consists in computing the principal (canonical) angles between subspaces, a straightforward generalisation of angles between vectors [1] :

Let F and G be subspaces in $\mathcal{R}^m$

$$\begin{array}{l} p = \dim(F) \\ q = \dim(G) \end{array} \qquad p \geq q \geq 1$$

The principal angles $\theta_1, \ldots, \theta_q$ between F and G are defined recursively by

$$\cos(\theta_k) = \max_{u \in F} \max_{v \in G} u^t.v = u_k^t.v_k$$

subject to

$$\|u\|_2 = \|v\|_2 = 1$$

$$\begin{array}{rcll} u^t.u_i & = & 0 & i = 1, \ldots, k-1 \\ v^t.v_i & = & 0 & i = 1, \ldots, k-1 \end{array}$$

First, two vectors u and v are searched for, between which the angle is minimal (maximal cosine), resulting in u_1, v_1 and θ_1. $\mathrm{Cos}(\theta_1)$ equals 1 if the subspaces intersect. Next, the same is done for the deflated subspaces, etc. The vectors $\{u_1, \ldots, u_q\}$ and $\{v_1, \ldots, v_q\}$ are called principal vectors of the subspace pair (F, G).

The idea of using SVD to compute the principal angles and vectors is due to Bjorck and Golub (1973) and is summarized in the following theorem :
If the columns of Q_f and Q_g define orthonormal bases for F and G respectively, the principal angles and vectors follow from the SVD of $(Q_f)^t.Q_g$

$$Q_f^t.Q_g = Y.\mathrm{diag}(s_1, \ldots, s_q).Z^t$$

$$\begin{array}{rcll} [u_1, \ldots, u_q] & = & Q_f.Y & \\ [v_1, \ldots, v_q] & = & Q_g.Z & \\ \cos(\theta_k) & = & s_k & k = 1, \ldots, q \end{array}$$

Several alternative methods to obtain the required orthonormal bases are conceivable (QR-decomposition, SVD, generalized SVD, etc.), each of them having its own advantages and disadvantages. If only knowledge of the intersection is needed, as will be the case in the identification algorithm, an SVD of the concatenation of two matrices defining the subspaces can suffice : The intersection of the row spaces of two matrices F and G can be calculated from $P_1.F$, where

$$[P_1 \; P_2].\begin{bmatrix} F \\ G \end{bmatrix} = 0$$

matrices P_1 and P_2 being determined through the SVD. Results of an in progress comparative study will be reported in future publications.

3 Dynamic systems

We now state some general properties that will be used throughout the sequel.

3.1 State space model

We consider linear, discrete time, time-invariant systems with m inputs and l outputs, with state space representation :

$$x[k+1] = A.x[k] + B.u[k] \quad (1)$$
$$y[k] = C.x[k] + D.u[k] \quad (2)$$

Vectors $u[k], y[k]$ and $x[k]$ denote the input,output and state at time k, the dimension of $x[k]$ being the minimal system order n. A, B, C and D are the unknown system matrices to be identified, making use only of measured I/O-sequences $u[k], u[k+1], \ldots$ and $y[k], y[k+1], \ldots$

3.2 Input-output equation

Equation (2) can easily be extended to a general structured I/O-equation :

$$Y_h = \Gamma_i.X + H_t.U_h \quad (3)$$

Y_h is a block Hankel matrix (i block rows, j columns) containing the consecutive outputs : ($y[k]$ is a $l \times 1$ vector, where l is the number of outputs)

$$Y_h = \begin{bmatrix} y[k] & y[k+1] & \ldots & y[k+j-1] \\ y[k+1] & y[k+2] & \ldots & y[k+j] \\ y[k+2] & y[k+3] & \ldots & y[k+j+1] \\ \ldots & \ldots & \ldots & \ldots \\ y[k+i-1] & y[k+i] & \ldots & y[k+j+i-2] \end{bmatrix}$$

Γ_i is an extended observability matrix :

$$\Gamma_i = \begin{bmatrix} C \\ CA \\ CA^2 \\ \ldots \\ CA^{i-1} \end{bmatrix}$$

X contains consecutive state vectors :

$$X = [x[k]\ x[k+1]\ x[k+2]\ \ldots\ x[k+j-1]]$$

H_t is a lower triangular Toeplitz matrix containing the Markov parameters :

$$H_t = \begin{bmatrix} D & 0 & 0 & \ldots & 0 \\ CB & D & 0 & \ldots & 0 \\ CAB & CB & D & \ldots & 0 \\ CA^2B & CAB & CB & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ CA^{i-2}B & CA^{i-3}B & CA^{i-4}B & \ldots & D \end{bmatrix}$$

Finally U_h is a block Hankel matrix with the same block dimensions as Y_h, but now containing the consecutive inputs. ($u[k]$ is a $m \times 1$ vector, where m is the number of inputs)

$$U_h = \begin{bmatrix} u[k] & u[k+1] & \dots & u[k+j-1] \\ u[k+1] & u[k+2] & \dots & u[k+j] \\ u[k+2] & u[k+3] & \dots & u[k+j+1] \\ \dots & \dots & \dots & \dots \\ u[k+i-1] & u[k+i] & \dots & u[k+j+i-2] \end{bmatrix}$$

Proof : straightforward by repeated substitution.

□

Instead of going into details, we loosely state that i and j should be chosen "sufficiently large" so as to satisfy certain conditions, and in particular $j \gg \max(mi, li)$ ("very rectangular" block Hankel Matrices).

3.3 Rank property

Let H denote the concatenation of Y_h and U_h :

$$H = \begin{bmatrix} Y_h \\ U_h \end{bmatrix}$$

then, generically (see below), the following theorem holds.

Theorem :

$$\text{rank}(H) = \text{rank}(U_h) + n \tag{4}$$

where n is the system order (observable part). Or, when U_h has full row rank :

$$\text{rank}(H) = mi + n \tag{5}$$

Proof :
From equation (3) we derive :

$$Y_h.U_h^{\perp} = \Gamma_i.X.U_h^{\perp}$$

and then of course :

$$\text{rank}(Y_h.U_h^{\perp}) = \text{rank}(\Gamma_i.X.U_h^{\perp})$$

where the columns of $U_h^{\perp}$ span the kernel of U_h (not trivial since $j \gg mi$).
Since Γ_i has full column rank :

$$\text{rank}(\Gamma_i.X.U_h^{\perp}) = \text{rank}(X.U_h^{\perp})$$

The generic condition to be satisfied is :

$$\text{rank}(X.U_h^{\perp}) = \text{rank}(X)$$

indicating that the row spaces of X and U_h should not intersect. Finally, under the condition that all modes are sufficiently excited (persistant excitation), the columns of X span the minimal part of the state space :

$$\text{rank}(X) = n$$

Putting all above equations in a row, we obtain

$$\text{rank}(Y_h.U_h^{\perp}) = n$$

and this, in fact, means that the row space of Y_h adds n dimensions to the row space of U_h, which proves equation (4).

□

This theorem allows us to estimate the system order, prior to further identification of the system matrices.

4 Determination of a state vector sequence

We now demonstrate how a sequence of state vectors can be calculated as the intersection of the row spaces of two block Hankel matrices, constructed from input-output vectors.

Let H_1 and H_2 be the concatenation of Y_{h1}, U_{h1} and Y_{h2}, U_{h2} respectively

$$H_1 = \begin{bmatrix} Y_{h1} \\ U_{h1} \end{bmatrix}, \qquad H_2 = \begin{bmatrix} Y_{h2} \\ U_{h2} \end{bmatrix}$$

where

$$Y_{h1} = \begin{bmatrix} y[k] & y[k+1] & \dots & y[k+j-1] \\ y[k+1] & y[k+2] & \dots & y[k+j] \\ y[k+2] & y[k+3] & \dots & y[k+j+1] \\ \dots & \dots & \dots & \dots \\ y[k+i-1] & y[k+i] & \dots & y[k+j+i-2] \end{bmatrix}$$

$$Y_{h2} = \begin{bmatrix} y[k+i] & y[k+i+1] & \dots & y[k+i+j-1] \\ y[k+i+1] & y[k+i+2] & \dots & y[k+i+j] \\ y[k+i+2] & y[k+i+3] & \dots & y[k+i+j+1] \\ \dots & \dots & \dots & \dots \\ y[k+2i-1] & y[k+2i] & \dots & y[k+2i+j-2] \end{bmatrix}$$

and U_{h1}, U_{h2} similarly constructed.
Both matrix pairs satisfy the I/O-equation :

$$Y_{h1} = \Gamma_i.X_1 + H_t.U_{h1} \tag{6}$$

$$Y_{h2} = \Gamma_i.X_2 + H_t.U_{h2} \tag{7}$$

Theorem :
A realization of X_2 can be calculated as the intersection of the row spaces of H_1 and H_2 :

$$\begin{aligned} X_2 &= [x[k+i]\ x[k+i+1]\ \dots\ x[k+i+j-1]] \\ &= \text{span}_{\text{row}}(H_1) \cap \text{span}_{\text{row}}(H_2) \end{aligned}$$

Proof :
We first prove that the dimension of the intersection equals n. Then, the (n-dimensional) row space of X_2 is shown to lie within both row spaces, which completes the proof.

a. Making use of the rank property (5), we derive

$$\dim(H_1) = \dim(H_2) = mi + n$$

where $\dim(M)$ is a shorthand notation for the dimension of M's row space. This rank property holds equally well for the concatenation of H_1 and H_2 :

$$H = \begin{bmatrix} H_1 \\ H_2 \end{bmatrix}$$

$$\dim(H_1 + H_2) = \dim(H) = 2mi + n$$

Applying Grassmann's dimension theorem :

$$\begin{aligned} \dim(H_1 \cap H_2) &= \dim(H_1) + \dim(H_2) - \dim(H_1 + H_2) \\ &= mi + n + mi + n - 2mi - n \\ &= n \end{aligned}$$

b. From equation (7), we derive

$$X_2 = \Gamma_i^+ . Y_{h2} - \Gamma_i^+ . H_t . U_{h2} = [\Gamma_i^+ \quad -\Gamma_i^+ . Ht] . \begin{bmatrix} Y_{h2} \\ U_{h2} \end{bmatrix}$$

where Γ_i^+ is Γ_i's pseudo-inverse, which shows that X_2's row space lies within H_2's row space. Equally well, X_1's row space lies within H_1's row space.

On the other hand, X_1 and U_{h1} completely determine X_2 through

$$X_2 = A^i . X_1 + [A^{i-1} . B \quad \dots \quad A . B \quad B] . U_{h1}$$

and since X_1's row space lies within H_1's row space, the same holds true for X_2's row space.

□

In practice, it occurs that both row spaces do not intersect, due to noise, non-linearity, etc. An approximate intersection can be calculated though, using the n first principal vectors, n being determined from equation (5).

5 Identification of the system matrices

Once X_2 is known, the system matrices can be identified by solving a set of linear equations :

$$\begin{bmatrix} x[k+i+1] & x[k+i+2] & \dots \\ y[k+i] & y[k+i+1] & \dots \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \cdot \begin{bmatrix} x[k+i] & x[k+i+1] & \dots \\ u[k+i] & u[k+i+1] & \dots \end{bmatrix}$$

Solving the (overdetermined) set of linear equations, one is faced with the problem that the state vector sequence appears twice in the equations, the noise on all matrix elements thus not being independent. Solving the set of equations in a total least squares sense will not take this in account [2]. Still, the TLS solution gives rise to nice interpretations. For instance, it implicitly creates corrected I/O-sequences u_c and y_c, and two differently corrected sequences of state vectors x_c and x_{cc}, that satisfy

$$\begin{bmatrix} x_{cc}[h+1] \\ y_c[h] \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \cdot \begin{bmatrix} x_c[h] \\ u_c[h] \end{bmatrix}$$

where A, B, C and D are the identified system matrices. This can be rewritten as :

$$\begin{aligned} x_c[h+1] &= A.x_c[h] + B.u_c[h] + (x_c[h+1] - x_{cc}[h+1]) \\ y_c[h] &= C.x_c[h] + D.u_c[h] \end{aligned}$$

The corrected sequences u_c and y_c thus form a valid I/O-sequence for the identified model, when process noise equal to $(x_c - x_{cc})$ is taken into account.

6 Real life example

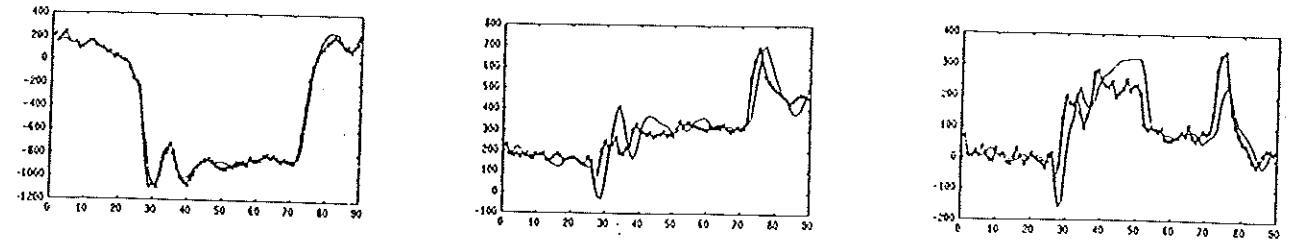
The performance of the algorithm has been evaluated on both simulated and industrial data sets. The following example is due to Prof. R.Guidorzi (University of Bologna). The I/O-sequence was obtained under normal operating conditions of a 120 MW power plant (Pont sur Sambre - France), a system with 5 inputs and 3 outputs. The identified models (for different system order estimates) were evaluated by comparing original and simulated outputs, using the original input signals and the identified model (see Figures). These simulations demonstrate the remarkable robustness of the identification scheme with respect to over- and underestimation of the system order.

7 Current research activities

As already mentioned earlier, several alternatives for the computation of the canonical correlation are possible and deserve further study. Also, since the identification scheme requires orthogonal decomposition of highly structured matrices with low displacement rank, implementation via so-called fast algorithms could be worth considering [3]. Exploiting the structure is also the actual issue in the updating and downdating mechanisms (adding and deleting measurements) in adaptive versions. Finally, an extension that deals with generalized state space models

$$\begin{aligned} E.x[k+1] &= A.x[k] + B.u[k] \\ F.y[k] &= C.x[k] + D.u[k] \end{aligned}$$

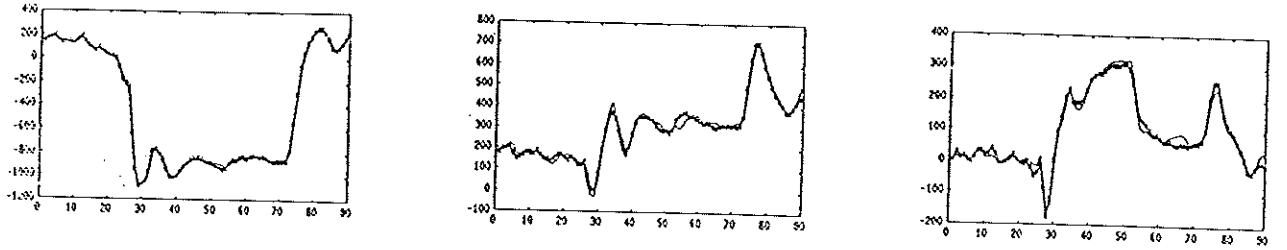
(where for instance E could be singular) is currently under investigation.



First order model



Third order model



Sixth order model

Figure 1 : Identification of a power plant : original and reconstructed outputs for different system order estimates.

8 Conclusion

A novel, strikingly simple strategy for state space identification from noisy I/O-measurements was presented. Only the main features were outlined, in expectation of wider study results. Future work will concentrate on several alternative computation schemes and efficient numerical implementation of adaptive singular value decomposition of structured matrices.

References

[1] G.H.Golub and C.F.Van Loan. Matrix computations. North Oxford Academic Publishing Co., Johns Hopkins University Press, 1983.

[2] S.Van Huffel. Analysis of the Total Least Squares problem and its use in parameter estimation. Doctoral Thesis, ESAT Dept. of Electrical Engineering, Katholieke Universiteit Leuven, 1987.

[3] B. Friedlander, M. Morf, T.Kailath, L. Ljung. New inversion formulas for matrices classified in terms of their distance from Toeplitz matrices. Linear Algebra and Its Applications, Vol. 27, pp. 31-60, 1979.