89-115 77

INT. J. CONTROL, 1990, VOL. 51, NO. 5, 1133–1146

# QSVD approach to on- and off-line state-space identification

MARC MOONEN† and JOOS VANDEWALLE†

Moonen *et al.* (1989 a), presented an SVD-based identification scheme for computing state-space models for multivariable linear time-invariant systems. In the present paper, this identification procedure is reformulated making use of the quotient singular value decomposition (QSVD). Here the input–output error covariance matrix can be taken into account explicitly, thus extending the applicability of the identification scheme to the case were the input and output data are corrupted by coloured noise. It turns out that in practice, due to the use of various pre-filtering techniques (anti-aliasing, etc.), this latter case is most often encountered. The extended identification scheme explicitly compensates for the filter characteristics and the consistency of the identification results follows from the consistency results for the QSVD. The usefulness of this generalization is demonstrated. The development is largely inspired by recent progress in total least-squares solution techniques (Van Huffel 1989) for the identification of static linear relations. The present identification scheme can therefore be viewed as the analogous counterpart for identifying dynamic linear relations.

## 1. Introduction

Identification aims at finding a mathematical model from the measurement record of inputs and outputs of a system. 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.

An elegant identification scheme was presented in an earlier paper (Moonen *et al.* 1989 a). The main step consists in computing the singular value decomposition (SVD) of a block Hankel matrix, constructed with I/O data. This procedure clearly resembles well known realization algorithms that compute a state-space model from the SVD of a block Hankel matrix, this time constructed with Markov parameters (Kung 1978, Zeiger and McEwen 1974). These realization algorithms suffer however from severe model inconsistency when there is noise on the data, due to loss of the Hankel structure when the 'noise singular values' are implicitly set to zero. Moreover, the sequence of Markov parameters might be hard to obtain in some applications. Instead, the above-mentioned identification scheme computes a state-space model immediately from the I/O data, which in practice turns out to be its main advantage. Even using a Hankel matrix, as well, the identification procedure can be shown to provide consistent results if both the input and output data are corrupted by additive white (measurement) noise (Fig. 1).

In practice, however, the I/O data are mostly corrupted by coloured noise, due

---

Received 18 May 1989.

† ESAT Katholieke Universiteit Leuven, K. Mercierlaan 94, 3030 Heverlee, Belgium.

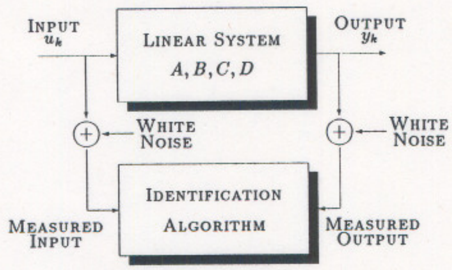


Figure 1.

to the use of various pre-filtering techniques (e.g. anti-aliasing, or bandpass filtering if a model for a limited frequency range is sought) (Fig. 2). Under these conditions it appears to be reasonable to assume that the noise colouring is completely defined through the filter characteristics, so that an error covariance matrix (up to a factor of proportionality) can be computed from the filter impulse response (see below for an example). Following recent progress in total least-squares solution techniques (Van Huffel 1987, 1989) for the identification of static linear relations, the above identification scheme can be reformulated in a QSVD framework (De Moor and Golub 1989). Then any input–output error covariance matrix (possibly even rank-deficient) can be taken into account explicitly. In this way consistent model estimates can be computed, even for the coloured noise case.

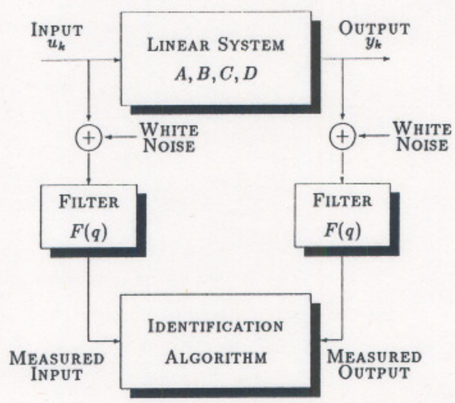


Figure 2.

In § 2 the original identification scheme is briefly reviewed and shown to provide consistent estimates in the white noise case. In § 3 the generalized procedure for the coloured noise case is presented. It is illustrated by practical examples in § 4. Finally, § 5 gives an outline of an adaptive version of the QSVD-based identification scheme.

## 2. SVD-based system identification: white noise case

For the time being we consider time-invariant linear, discrete-time, multivariable systems with the state-space representation

$$x_{k+1} = A \cdot x_k + B \cdot u_k$$
$$y_k = C \cdot x_k + D \cdot u_k$$

where $u_k$, $y_k$ and $x_k$ denote the input ($m$-vector), output ($l$-vector) and state vector 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 recorded I/O sequences $u_k, u_{k+1}, \ldots$ and $y_k, y_{k+1}, \ldots$.

### 2.1. *Identification scheme*

Moonen *et al.* (1989) showed how a state vector sequence can be computed from I/O measurements only, as follows. Let $H_1$ and $H_2$ be defined as

$$H_1 = \begin{bmatrix} u_k & u_{k+1} & \cdots & u_{k+j-1} \\ y_k & y_{k+1} & \cdots & y_{k+j-1} \\ u_{k+1} & u_{k+2} & \cdots & u_{k+j} \\ y_{k+1} & y_{k+2} & \cdots & y_{k+j} \\ \vdots & \vdots & & \vdots \\ u_{k+i-1} & u_{k+i} & \cdots & u_{k+j+i-2} \\ y_{k+i-1} & y_{k+1} & \cdots & y_{k+j+i-2} \end{bmatrix}$$

$$H_2 = \begin{bmatrix} u_{k+i} & u_{k+i+1} & \cdots & u_{k+i+j-1} \\ y_{k+i} & y_{k+i+1} & \cdots & y_{k+i+j-1} \\ u_{k+i+1} & u_{k+i+2} & \cdots & u_{k+i+j} \\ y_{k+i+1} & y_{k+i+2} & \cdots & y_{k+i+j} \\ \vdots & \vdots & \cdots & \vdots \\ u_{k+2i-1} & u_{k+2i} & \cdots & u_{k+2i+j-2} \\ y_{k+2i-1} & y_{k+2i} & \cdots & y_{k+2i+j-2} \end{bmatrix}$$

$$j \gg i$$

and let the state vector sequence $\chi$ be defined as

$$\chi = [x_{k+i} \quad x_{k+i+1} \quad \cdots \quad x_{k+i+j-1}]$$

Then, under certain conditions (Moonen *et al.* 1989)

$$\mathrm{span}_{\mathrm{row}}(\chi) = \mathrm{span}_{\mathrm{row}}(H_1) \cap \mathrm{span}_{\mathrm{row}}(H_2)$$

so that any basis for this intersection constitutes a valid state vector sequence $\chi$ with the basis vectors as the consecutive row vectors.

Once $\chi$ is known, the system matrices can be identified by solving an (overdetermined) set of linear equations:

$$\begin{bmatrix} x_{k+i+1} & \cdots & x_{k+i+j-1} \\ y_{k+i} & \cdots & y_{k+i+j-2} \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \cdot \begin{bmatrix} x_{k+i} & \cdots & x_{k+i+j-2} \\ u_{k+i} & \cdots & u_{k+i+j-2} \end{bmatrix}$$

The above results constitute the heart of a two-step identification scheme. First a state vector sequence is realized as the intersection of the row spaces of two block Hankel matrices, constructed with I/O data. Then the system matrices are obtained at once from the least-squares solution of a set of linear equations.

2.2. *Computational details*

The following derivation (which is slightly different from the one in Moonen *et al.* (1989)) shows how these computations can be carried out quite easily, resulting in a consistent double SVD identification algorithm.

As a first step the intersection of the row spaces spanned by $H_1$ and $H_2$, can be recovered from the SVD of the concatenation

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

$$H = \begin{bmatrix} H_1 \\ H_2 \end{bmatrix} = U_H \cdot S_H \cdot V_H^t = \begin{bmatrix} U_{11} & U_{12} \\ U_{21} & U_{22} \end{bmatrix} \cdot \begin{bmatrix} S_{11} & 0 \\ 0 & 0 \end{bmatrix} \cdot V_H^t$$

$$\dim(U_{11}) = (mi + li) \times (2mi + n)$$

$$\dim(U_{12}) = (mi + li) \times (2li - n)$$

$$\dim(U_{21}) = (mi + li) \times (2mi + n)$$

$$\dim(U_{22}) = (mi + li) \times (2li - n)$$

$$\dim(S_{11}) = (2mi + n) \times (2mi + n)$$

(see Moonen *et al.* 1989 for details). From

$$U_{12}^t \cdot H_1 = -U_{22}^t \cdot H_2$$

it follows that the row space of $U_{12}^t \cdot H_1$ is equal to the required intersection. However, $U_{12}^t \cdot H_1$ contains $2li - n$ row vectors, only $n$ of which are linearly independent (i.e. $n$ is the dimension of the intersection). Thus it remains to select $n$ suitable combinations of these row vectors.

Making use of a *CS* decomposition (Golub *et al.* 1983), one can easily show that

$$U_{12} = [U_{12}^{(1)} \quad U_{12}^{(2)} \quad U_{12}^{(3)}] \cdot \begin{bmatrix} I_{(li-n)\times(li-n)} & & \\ & C_{n\times n} & \\ & & 0_{(li-n)\times(li-n)} \end{bmatrix} \cdot V_*^t$$

$$U_{22} = [U_{22}^{(1)} \quad U_{22}^{(2)} \quad U_{22}^{(3)}] \cdot \begin{bmatrix} O_{(li-n)\times(li-n)} & & \\ & S_{n\times n} & \\ & & I_{(li-n)\times(li-n)} \end{bmatrix} \cdot V_*^t$$

$$C = \operatorname{diag}(c_1, \ldots, c_n)$$

$$S = \operatorname{diag}(s_1, \ldots, s_n)$$

$$I_{n\times n} = C^2 + S^2$$

where $U_{12}^{(1)}$ then constitutes the $(li - n)$-dimensional orthogonal complement of $H_1$. Clearly, only $U_{12}^{(2)}$ delivers useful combinations for the computation of the intersection, and we can take

$$\chi = U_{12}^{(2)t} \cdot H_1$$

The above expressions for $U_{12}$ and $U_{22}$ are in themselves SVDs of these matrices, and can be computed as such. It thus suffices to compute, for example, the SVD of $U_{12}$. Computation of the required intersection then reduces to computation of two successive SVDs for $H$ and $U_{12}$.

Note that the above derivation is nothing more than a double SVD approach to computing the QSVD of the matrix pair $(H_1, H_2)$, following from the constructive QSVD proof of Paige and Saunders (1981). From this last remark, one might be tempted to apply immediately a one-stage QSVD procedure to the matrix pair, as developed by Paige (1986). This latter method would, however, compute the exact intersection of the row spaces, which in the presence of noise turns out to be completely absent (generically). The outcome of applying Paige's algorithm would then be a zero-dimensional intersection, as could have been guessed beforehand. The difference between these methods turns out to be the intermediate rank decision after the first SVD in the first approach (double SVD) that fixes the dimension of the approximate intersection to be computed next. Although this (possibly difficult) intermediate rank decision has been a main motive for developing a one-stage QSVD algorithm, for our purpose it is somehow inevitable.

In the second step, the system matrices are to be identified from a set of linear equations. Much as was done by Moonen *et al.* (1989), it can be shown straightforwardly that the system matrices can be computed from the following reduced set as well (obtained after discarding the common orthogonal factor $V_H$).

$$\begin{bmatrix} U_{12}^{(2)t} \cdot U_H(m+l+1:(i+1)(m+l), 1:2mi+n) \cdot S_{11} \\ U_H(mi+li+m+1:(m+l)(i+1), 1:2mi+n) \cdot S_{11} \end{bmatrix}$$

$$= \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} U_{12}^{(2)t} \cdot U_H(1:mi+li, 1:2mi+n) \cdot S_{11} \\ U_H(mi+li+1:mi+li+m, 1:2mi+n) \cdot S_{11} \end{bmatrix}$$

where $U_H(r:s, v:w)$ is a submatrix of $U_H$ at the intersection of rows $r, r+1, \ldots, s$ and columns $v, v+1, \ldots, w$.

### 2.3. *Consistency*

The identification procedure is proven to be consistent if the number of columns in $H$ tends to infinity and if the input–output measurements are corrupted with additive white measurement noise, or in other words, if the columns in $H$ are subject to independently and identically distributed errors with zero-mean and common error covariance matrix equal to the identity matrix, up to a factor of proportionality. For that case it can indeed be shown (De Moor 1988 a) that the left singular basis $U_H$ can be computed consistently (as opposed to the singular values $S_H$ and the right singular basis $V_H$). As the system matrices are next computed essentially from $U_H$ only (see the above set of equations), the model estimate is clearly consistent. (The matrix $S_H$ in this set imposes weights on the different equations. This does not influence the outcome if the set of equations can be solved exactly, which is the case under the assumed conditions.) The corresponding noise model is depicted in Fig. 1.

### 2.4. *Adaptive identification*

The above algorithm is easily converted into an adaptive one, where model updating should account for time variance. Every time step, a new input–output measurement becomes available, defining a new column to be appended to the matrix $H$. On the other hand, older measurements should be discarded, e.g. by exponential weighting. The off-line algorithm of the previous section is then applied to the updated $H$-matrix.

Since only $U_H$ and $S_H$ in the SVD of $H$ are needed for further computations, $H$ need not be constructed explicitly since the weighting can be applied to $S_H$ as well. It then suffices to update only $U_H$ and $S_H$ in every time step, as outlined in the following algorithm.

*Algorithm* 1

Initialize

$$U_{H^{(0)}} = I_{(2mi+2li)\times(2mi+2li)}$$

$$S_{H^{(0)}} = O_{(2mi+2li)\times(2mi+2li)}$$

($m$ and $l$ being the number of inputs and outputs, respectively, $2i$ being the number of block rows in the fictitious matrix $H$). For $k = 1, \dots$ do the following.

*Step* 1

Construct new column *column* to be added to $H^{(k-1)}$, using the $2i$ latest I/O measurements.

*Step* 2

Calculate SVD

$$U_{H^{(k)}} \cdot S_{H^{(k)}} \cdot V^t_{H^{(k)}} = [\alpha \cdot U_{H^{(k-1)}} \cdot S_{H^{(k-1)}}\ column]$$

and partition

$$U_{H^{(k)}} = \begin{bmatrix} U_{11} & U_{12} \\ U_{21} & U_{22} \end{bmatrix} \quad \dim(U_{11}) = (mi+li)\times(2mi+n)$$

$$S_{H^{(k)}} = \begin{bmatrix} S_{11} & 0 \\ 0 & S_{22} \end{bmatrix} \quad \dim(S_{11}) = (2mi+n)\times(2mi+n)$$

*Step* 3

Calculate the SVD:

$$U_{12} = [U^{(1)}_{12}\quad U^{(2)}_{12}\quad U^{(3)}_{12}] \cdot \begin{bmatrix} I_{(li-n)\times(li-n)} & & \\ & C_{n\times n} & \\ & & 0_{(li-n)\times(li-n)} \end{bmatrix} \cdot V^t_*$$

*Step* 4

Solve the set of linear equations

$$\begin{bmatrix} U^{(2)t}_{12} \cdot H_{H^{(k)}}(m+l+1:(i+1)(m+l), 1:2mi+n) \cdot S_{11} \\ U_{H^{(k)}}(mi+li+m+1:(m+l)(i+1), 1:2mi+n) \cdot S_{11} \end{bmatrix}$$

$$= \begin{bmatrix} A^{(k)} & B^{(k)} \\ C^{(k)} & D^{(k)} \end{bmatrix} \begin{bmatrix} U^{(2)t}_{12} \cdot U_{H^{(k)}}(1:mi+li, 1:2mi+n), \cdot S_{11} \\ U_{H^{(k)}}(mi+li+1:mi+li+m, 1:2mi+n) \cdot S_{11} \end{bmatrix}$$

Clearly, the model-updating boils down to SVD-updating (Step 2), followed by a limited number of additional operations. The overal efficiency therefore closely depends on the efficiency with which the SVD-updating can be carried out. Efficient parallel algorithms for updating the singular value decomposition are dealt with by Moonen *et al.* (1989 b).

## 3. QSVD-based system identification: coloured noise case

Let us proceed to the case where the I/O data are corrupted by coloured noise. For instance, in Fig. 2, the use of pre-filtering ($F(q)$) does not change the dynamic relation between input and output, but it introduces a colouring of the additive (measurement) noise.

Assume that the columns in the concatenated matrix

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

are subject to identically distributed errors with zero mean and common error covariance matrix $\Delta$ up to a factor of proportionality, where

$$\Delta = R_\Delta \cdot R_\Delta^t$$

is the Cholesky factorization of $\Delta$ ($R_\Delta$ lower triangular).

One can easily verify that the columns in the transformed matrix $R_\Delta^{-1} \cdot H$ have an error covariance matrix equal to the identity matrix up to a factor of proportionality. One way of carrying out the identification would then consist of having the identification based on the SVD of $R_\Delta^{-1} \cdot H$ (with a consistent computation of the left singular basis, see § 2) instead of $H$, and including a kind of retransformation with $R_\Delta^{-1}$. The overall identification scheme would then deliver a consistent estimate. However, if $R_\Delta$ is singular or ill-conditioned, the matrix inverse $R_\Delta^{-1}$ should not be computed explicitly. Instead, one should make use of the quotient singular value decomposition (QSVD) of the matrix pair $(H, R_\Delta)$ (which in the non-singular case indeed reduces to the SVD of $R_\Delta^{-1} \cdot H$). We can now show how a double QSVD identification scheme can be designed, analogous to the double SVD scheme for the white noise case (the latter being a special case of the former where every single QSVD reduces to an SVD, as can easily be verified).

The QSVD of $(H, R_\Delta)$ is a simultaneous reduction of the two matrices (having the same number of rows) by two orthogonal matrices $Q_H$ and $Q_{R_\Delta}$ and a non-singular matrix $X$ (Van Loan 1976, Paige and Saunders 1981).

$$X^t \cdot H \cdot Q_H = \Sigma_H$$
$$X^t \cdot R_\Delta \cdot Q_{R_\Delta} = \Sigma_{R_\Delta}$$

where

$$\Sigma_H = \operatorname{diag}(\alpha_1, \ldots, \alpha_{2li+2mi})$$
$$\Sigma_{R_\Delta} = \operatorname{diag}(\beta_1, \ldots, \beta_{2li+2mi})$$

$$\frac{\alpha_1}{\beta_1} > \frac{\alpha_2}{\beta_2} > \ldots > \frac{\alpha_{2li+2mi}}{\beta_{2li+2mi}}$$

$(\alpha_i, \beta_i)$ is called a quotient singular value pair, whereas $\alpha_i/\beta_i$ is a quotient singular value.

Alternatively, one can write

$$H = X^{-t} \cdot \Sigma_H \cdot Q_H^t$$
$$R_\Delta = X^{-t} \cdot \Sigma_{R_\Delta} \cdot Q_{R_\Delta}^t$$

Although the above QSVD form is not the most desirable one in terms of either numerical stability or efficient parallel implementation, it is quite expository and instructive for our purpose. We therefore stick to this representation, keeping in mind that in actual practice one should preferably make use of an alternative QSVD computation scheme (see Paige 1986).

A geometrical interpretation for the matrix $X$ is as follows. If the error covariance matrix $\Delta$ is equal to the identity matrix (up to a factor), then so does $R_\Delta$, and one can easily verify that $X$ corresponds to $U_H$ in the SVD of $H$ ($H = U_H \cdot S_H \cdot V_H^t$), up to a possible scaling of its column vectors. The column vectors of $X$ then define directions in which the *oriented signal energy* in the column space of $H$ is extremal (De Moor *et al.* 1988). Similarly, in the general case where $\Delta \neq I$, the column vectors of $X$ define directions in which the *oriented signal-to-noise ratio* is extremal (De Moor *et al.* 1988). Much as was done for the white noise case, where the intersection of the row spaces of $H_1$ and $H_2$ was computed using the directions of minimal oriented signal energy for

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

that is

$$\begin{bmatrix} U_{12} \\ U_{22} \end{bmatrix}$$

(remember $U_{12}^t \cdot H_1 = -U_{22}^t \cdot H_2$), we can now compute the intersection making use of the directions of minimal oriented signal-to-noise ratio

$$\begin{bmatrix} X_{12} \\ X_{22} \end{bmatrix}$$

to be defined next.

It is instructive to consider first the noise-free case (error covariance proportional to $\Delta$, but with a zero factor of proportionality), and then demonstrate that the derivations still hold if there is a non-zero error contribution. If the data are noise-free, then from the above QSVD definition it follows that

$$H = \begin{bmatrix} H_1 \\ H_2 \end{bmatrix} = X^{-t} \cdot \Sigma_H \cdot Q_H^t = \begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix}^{-t} \cdot \begin{bmatrix} \Sigma_{11} & 0 \\ 0 & 0 \end{bmatrix} \cdot Q_H^t$$

$$\dim(X_{11}) = (mi + li) \times (2mi + n)$$

$$\dim(X_{12}) = (mi+li) \times (2li-n)$$
$$\dim(X_{21}) = (mi+li) \times (2mi+n)$$
$$\dim(X_{22}) = (mi+li) \times (2li-n)$$
$$\dim(\Sigma_{11}) = (2mi+n) \times (2mi+n)$$

Again, from

$$X_{12}^{t} \cdot H_1 = -X_{22}^{t} \cdot H_2$$

it follows that the row space of $X_{12}^{t} \cdot H_1$ is equal to the required intersection. As $X_{12}^{t} \cdot H_1$ contains $2li-n$ row vectors, only $n$ of which are linear independent (due to the dimension of the intersection), it remains to select $n$ suitable combinations of these row vectors.

Making use of a QSVD, one can easily show that

$$X_{12} = [X_{12}^{(1)} \quad X_{12}^{(2)} \quad X_{12}^{(3)}] \cdot \begin{bmatrix} I_{(li-n)\times(li-n)} & & \\ & C_{n\times n} & \\ & & 0_{(li-n)\times(li-n)} \end{bmatrix} \cdot T_*^{t}$$

$$X_{22} = [X_{22}^{(1)} \quad X_{22}^{(2)} \quad X_{22}^{(3)}] \cdot \begin{bmatrix} O_{(li-n)\times(li-n)} & & \\ & S_{n\times n} & \\ & & I_{(li-n)\times(li-n)} \end{bmatrix} \cdot T_*^{t}$$

$$C = \operatorname{diag}(c_1, \ldots, c_n)$$
$$S = \operatorname{diag}(s_1, \ldots, s_n)$$
$$I_{n\times n} = C^2 + S^2$$

Clearly, only $X_{12}^{(2)}$ delivers useful combinations for the computation of the intersection, and we can take

$$\chi = X_{12}^{(2)t} \cdot H_1$$

Note that in the white noise case, this last QSVD reduced to a *CS* decomposition and could then be computed from a single SVD, resulting in an overall *double SVD scheme* for the computation of the intersection. In the general case, the computation of this intersection is carried out in a *double QSVD scheme*.

In the second step, the system matrices can be computed from the following reduced set of equations (obtained after discarding the common orthogonal factor $Q_H$)

$$\begin{bmatrix} X_{12}^{(2)t} \cdot X^{-t}(m+l+1:(i+1)(m+l),\ 1:2mi+n) \cdot \Sigma_{11} \\ X^{-t}(mi+li+m+1:(m+l)(i+1),\ 1:2mi+n) \cdot \Sigma_{11} \end{bmatrix}$$
$$= \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} X_{12}^{(2)t} \cdot X^{-t}(1:mi+li,\ 1:2mi+n) \cdot \Sigma_{11} \\ X^{-t}(mi+li+1:mi+li+m,\ 1:2mi+n) \cdot \Sigma_{11} \end{bmatrix}$$

where $X^{-t}(r:s, v:w)$ is a submatrix of $X^{-t}$ at the intersection of rows $r, r+1, \ldots, s$ and columns $v, v+1, \ldots, w$.

It remains to show that the above identification scheme delivers consistent results if the number of columns in $H$ tends to infinity, and if the columns in $H$ are subject to identically distributed errors with zero mean and a common error covariance matrix equal to $\Delta$, up to a factor of proportionality. For that case, it can again be shown (De Moor 1988) that the matrix $X$ in the QSVD can be computed consistently. As the system matrices are next computed essentially from $X$ only (the matrix $\Sigma_H$ in the above set of equations again imposes weights that do not influence the solution in the case considered, see § 2), the model estimate is clearly consistent.

A final remark concerns trivial quotient singular values $\alpha_i/\beta_i = 0/0$. As these correspond to vectors in the orthogonal complement of $H$ ($\alpha_i = 0$)—and additionally correspond to noise-free directions as $\beta_i = 0$—they should be treated as such. In other words, the column vectors in $X$ corresponding to trivial quotient singular values should be reckoned in

$$\begin{bmatrix} X_{12} \\ X_{22} \end{bmatrix}$$

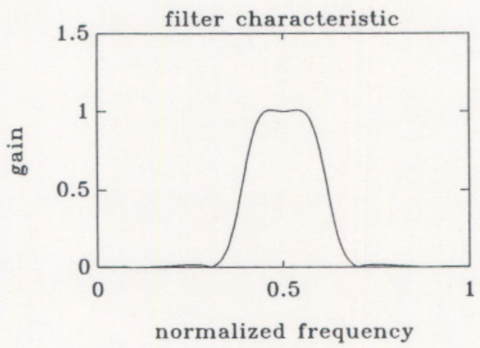


Figure 3.

## 4. Examples

The following small examples illustrate the usefulness of the generalized identification procedure.

*Example* 1

First let us consider a first-order single-input single-output system, with state-space equations

$$x_{k+1} = 0{\cdot}5x_k + u_k$$

$$y_k = x_k$$

The general set-up is depicted in Fig. 2. An I/O sequence was generated using a random input sequence, with additive white noise superimposed on both the input and the output sequence next. Finally, the noise-corrupted I/O sequence was passed through a linear filter $F(q)$, with filter characteristics as shown in Fig. 3. The noise colouring, or equivalently the noise error covariance matrix for the columns in $H$ (to be introduced next), can be determined from the filter impulse response $\delta_0, \delta_1, \delta_2, \ldots,$

as follows (symmetric Toeplitz matrix):

$$\Delta_{(2li+2mi)\times(2li+2mi)} = \begin{bmatrix} E\{f_k^{\text{out}} f_k^{\text{out}}\} & 0 & E\{f_k^{\text{out}} f_{k+1}^{\text{out}}\} & 0 & \cdots \\ 0 & E\{f_k^{\text{out}} f_k^{\text{out}}\} & 0 & E\{f_k^{\text{out}} f_{k+1}^{\text{out}}\} & 0 \\ E\{f_k^{\text{out}} f_{k+1}^{\text{out}}\} & 0 & E\{f_k^{\text{out}} f_k^{\text{out}}\} & 0 & \cdots \\ 0 & E\{f_k^{\text{out}} f_{k+1}^{\text{out}}\} & 0 & E\{f_k^{\text{out}} f_k^{\text{out}}\} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

where $E\{\cdot\}$ is the expectation operator, and $f_k^{\text{out}}$ is the output of the filter for a white noise input sequence $F_k^{\text{in}}$. The zero entries are due to the fact that the additive noise sequences on the input and output ($u_k$ and $y_k$) are assumed to be uncorrelated.

$E\{f_k^{\text{out}} f_{k+l}^{\text{out}}\}$ can be computed as follows (Papoulis 1980):

$$E\{f_k^{\text{out}} f_{k+l}^{\text{out}}\} = \sigma_{\text{noise}}^2 \sum_{r=0}^{\infty} \delta_r \delta_{r+1}$$

($\sigma_{\text{noise}}^2$ is a factor of proportionality that need not be known.) Making use of the Cholesky factor of $\Delta$, one can then apply the QSVD-based identification procedure. Figure 4 shows the identified pole as a function of the number of columns in $H$, both for the QSVD-based identification scheme and for the SVD-based scheme. The latter cannot compensate for the noise colouring. For small noise contributions (1%) the difference between the two results turns out to be relatively small (Fig. 4 (a)). For larger values of the signal-to-noise ration (10% in Fig. 4 (b)), only the QSVD scheme delivers useful results. Furthermore, the estimate clearly improves as the number of columns in $H$ increases (consistency).

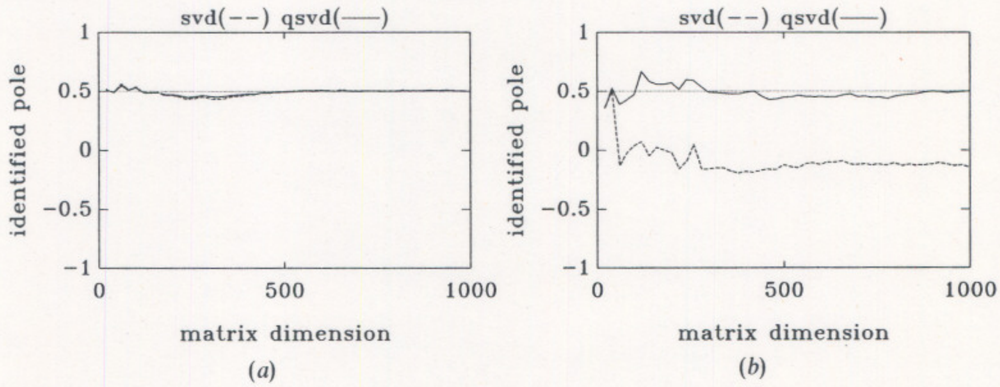


Figure 4.

*Example* 2

As a second example, we again consider the same first-order SISO system. If only the output is corrupted by additive white noise (Fig. 5) the error covariance matrix clearly equals

$$\Delta_{(2li+2mi)\times(2li+2mi)} = \begin{bmatrix} 0 & 0 & 0 & 0 & \dots \\ 0 & 1 & 0 & 0 & \dots \\ 0 & 0 & 0 & 0 & \dots \\ 0 & 0 & 0 & 1 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

(which is a rank deficient matrix). Figure 6 shows the identified pole as a function of the number of columns in $H$, both for the QSVD-based identification scheme and for the SVD-based scheme. Again, only the QSVD scheme delivers consistent results.

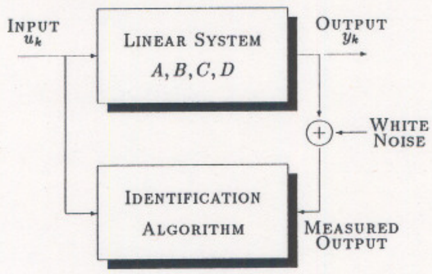


Figure 5.

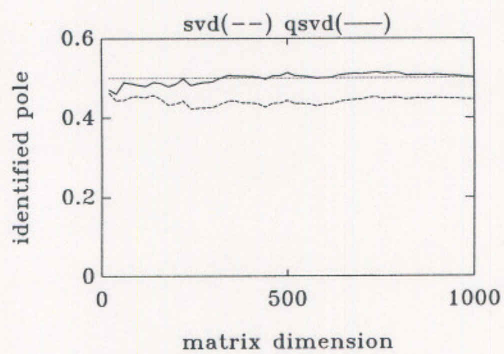


Figure 6.

## 5. Adaptive QSVD-based identification

Much as for the SVD-based scheme, the above QSVD-based algorithm is easily converted into an adaptive one for time-varying systems, making use of QSVD-updating and exponential weighting. If the error covariance matrix is time-invariant, the following algorithm straightforwardly applies. Again, $H$ is never constructed explicitly but its factors $\Sigma_H$ and $X$ are stored and updated.

*Algorithm* 2

Initialize

$$X_{(0)} = I_{(2mi+2li)\times(2mi+2li)}$$

$$\Sigma_{H^{(0)}} = 0_{(2mi+2li)\times(2mi+2li)}$$

($m$ and $l$ being the number of inputs and outputs respectively, $2i$ the number of block rows in the fictitious matrix $H$).

For $k = 1, \ldots$

*Step* 1

Construct new column *column* to be added to $H^{(k-1)}$, using the $2i$ latest I/O measurements.

*Step* 2

Calculate QSVD

$$X_{(k)}^{-t} \cdot \Sigma_{H^{(k)}} \cdot Q_{H^{(k)}}^{t} = [\alpha \cdot X_{(k-1)}^{-t} \Sigma_{H^{(k-1)}}\ column]$$

$$X_{(k)}^{-t} \cdot \Sigma_{R_\Delta^{(k)}} \cdot Q_{R_\Delta^{(k)}}^{t} = X_{(k-1)}^{-t} \Sigma_{R_\Delta^{(k-1)}}$$

and partition

$$X_{(k)} = \begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix} \quad \dim(X_{11}) = (mi + li) \times (2mi + n)$$

$$\Sigma_{H^{(k)}} = \begin{bmatrix} \Sigma_{11} & 0 \\ 0 & \Sigma_{22} \end{bmatrix} \quad \dim(\Sigma_{11}) = (2mi + n) \times (2mi + n)$$

*Step* 3

Calculate QSVD

$$X_{12} = [X_{12}^{(1)} \quad X_{12}^{(2)} \quad X_{12}^{(3)}] \cdot \begin{bmatrix} I_{(li-n)\times(li-n)} & & \\ & C_{n\times n} & \\ & & 0_{(li-n)\times(li-n)} \end{bmatrix} \cdot T_*^{t}$$

$$X_{22} = [X_{22}^{(1)} \quad X_{22}^{(2)} \quad X_{22}^{(3)}] \cdot \begin{bmatrix} O_{(lo\|n)\times(li-n)} & & \\ & S_{n\times n} & \\ & & I_{(li-n)\times(li-n)} \end{bmatrix} \cdot T_*^{t}$$

*Step* 4

Solve the set of linear equations

$$\begin{bmatrix} X_{12}^{(2)t} \cdot X_{(k)}^{-t}(m + l + 1 : (i+1)(m+l), 1 : 2mi + n) \cdot \Sigma_{11} \\ X_{(k)}^{-t}(mi + li + m + 1 : (m+l)(i+1), 1 : 2mi + n) \cdot \Sigma_{11} \end{bmatrix}$$

$$= \begin{bmatrix} A^{(k)} & B^{(k)} \\ C^{(k)} & D^{(k)} \end{bmatrix} \begin{bmatrix} X_{12}^{(2)t} \cdot X_{(k)}^{-t}(1 : mi + li, 1 : 2mi + n) \cdot \Sigma_{11} \\ X_{(k)}^{-t}(mi + li + 1 : mi + li + m, 1 : 2mi + n) \cdot \Sigma_{11} \end{bmatrix}$$

Again, the model-updating boils down to QSVD-updating (Step 2), followed by a limited number of additional operations. The overall efficiency therefore closely depends on the efficiency with which the QSVD-updating is carried out. Efficient parallel algorithms for updating the quotient singular value decomposition are dealt with in Moonen *et al.* (1989 b).

## 6. Conclusions

A double SVD scheme for state-space identification that applies to the case where the available data are subject to additive white (measurement) noise has been generalized to a double QSVD scheme for the coloured noise case. By use of examples the practical relevance of this identification scheme has been demonstrated. Finally, much like the SVD scheme, the QSVD scheme can easily be converted into an adaptive algorithm for on-line model updating.

ACKNOWLEDGMENT

This text presents research results of the Belgian programme on interuniversity attraction poles initiated by the Belgian State Prime Minister's Office—Science Policy Programming. Scientific responsibility is assumed by its authors.

Marc Moonen is a research assistant with the N.F.W.O. (Belgian National Fund for Scientific Research).

REFERENCES

DE MOOR, B., 1988, Mathematical concepts and techniques for modelling static and dynamic systems. Doctoral dissertation, Katholieke Universiteit, Leuven.

DE MOOR, B., VANDEWALLE, J., and STAAR, J., 1988, Oriented energy and oriented signal-to-signal ratio concepts in the analysis of vector sequences and time series. *Singular Value Decomposition and Signal Processing*, edited by E. Deprettere (Amsterdam: North-Holland).

DE MOOR, B., and GOLUB, G. H., 1989, Generalized singular value decompositions: a proposal for a standardized nomenclature. Internal Report, Department of Computer Science, Stanford University, U.S.A.

GOLUB, G. H., and VAN LOAN, C. F., 1983, *Matrix Computations* (Baltimore: Johns Hopkins University Press).

KUNG, S. Y., 1978, A new identification and model reduction algorithm via singular value decomposition. *Proceedings of the* 12*th Asilomar Conference on Circuits, Systems and Computers* Pacific Grove, pp. 705–714.

MOONEN, M., DE MOOR, B., VANDENBERGHE, L., and VANDEWALLE, J., 1989 a, On- and off-line identification of linear state space models. *International Journal of Control*, **49,** 219–232.

MOONEN, M., VAN DOOREN, P., and VANDEWALLE, J., 1989 b, Parallel algorithms for updating the (generalized) singular value decomposition. Internal Report, ESAT Katholieke Universiteit, Leuven.

PAIGE, C. C., 1986, Computing the generalized singular value decomposition. *SIAM Journal on Scientific and Statistical Computing*, **7,** 1126–1146.

PAIGE, C. C., and SAUNDERS, M. A., 1981, Towards a generalized singular value decomposition. SIAM Journal on Numerical Analysis, **18,** 398–405.

PAPOULIS, A., 1980, *Circuits and Systems—a Modern Approach* (Holt, Rinehart and Winston).

VAN HUFFEL, S., 1987, Analysis of the total least squares problem and its use in parameter estimation. Doctoral dissertation, Katholieke Universiteit, Leuven; 1989, Analysis and properties of the generalized total least squares problem $AX \approx B$ when some or all columns in $A$ are subject to error. *SIAM Journal on Matrix Analysis and Applications*, to be published.

VAN LOAN, C. F., 1976, Generalizing the singular value decomposition. *SIAM Journal on Numerical Analysis*, **13,** 76–83.

ZEIGER, H. P., and MCEWEN, A. J., 1974, Approximate linear realizations of given dimensions via Ho's algorithm. *I.E.E.E. Transactions on Automatic Control*, **19,** 153.