ADAPTIVE INPUT ESTIMATION

Anders Ahlén and Mikael Sternad

Automatic Control and Systems Analysis Group
Department of Technology, Uppsala University
P O Box 534, S-751 21 Uppsala, Sweden

Abstract: An adaptive algorithm for estimating the input to a linear system is presented. This explicite self-tuning filter is based on the identification of an innovations model. From that model, input and measurement noise descriptions are decomposed. Identifiability results guarantee a unique decomposition. Main tools in the algorithm are the solution of two linear systems of equations. The basic algorithm can be used for input signals described by ARMA-models and moving average measurement noise. An extension of the algorithm involves use of model reduction and spectral factorization. Simulation experiments illustrate the filtering performance.

Keywords: Input estimation; deconvolution; identification; adaptive filtering;

1 Introduction

The need to restore signals observed through linear systems and contaminated by noise arises frequently. In digital communication, the *equalization* problem is fundamental. Due to intersymbol interference caused by the communication channel, transmitted sequences have to be restored. Numerous papers have been written on this topic. See eg Qureshi (1985) and the references therein. In *seismology*, reflection coefficients representing hidden layers in the ground are sought. Deconvolution algorithms are important tools in this investigation, cf Mendel (1983). *Numerical differentiation* of noisy data involves a tradeoff in order to limit noise amplification. In such problems, a noise corrupted output of an integrator is measured, and the input is sought, see eg Ahlén (1986) and Carlsson et al (1987). In *control systems*, slow transducers often cause problems. If the transducer dynamics is eliminated, the total phase lag in the control system may be reduced. Estimates of transducer inputs could be used as artificial measurements.

In order to find the desired signal, some kind of inverse filtering is needed. Usually, such filtering is known as *deconvolution* or *input estimation*. The problem has been discussed by many authors, see eg Fitch and Kurz (1975), Demoment (1983), Commenges (1984), Deng (1985), Ahlén and Sternad (1985, 1989). A close correspondence between deconvolution and feedforward control has also been observed, see Sternad and Ahlén (1988).

In the areas described above, it is reasonable to assume some crucial part of the system to be known a priori or determined in advance by experiments. For example, the wavelet in seismology, the transducer in control problems or the integrator in numerical differentiation are known. The input and measurement noise properties may, however, vary with time. Ahlén (1986) and Moir et al (1987) have suggested an adaptive approach based on an innovations model in order to estimate the input. In this paper, we will develope this approach, and focus on the problem of estimating the input when the characteristics of the input and measurement noise vary with time.

2 Preliminaries

Consider the linear stochastic discrete-time system

$$y(t) = q^{-d}\frac{B(q^{-1})}{A(q^{-1})}u(t) + w(t) \tag{1}$$

where q^{-1} denotes the backward shift operator. The input $u(t)$ and measurement noise $w(t)$ are assumed to be accurately described by the ARMA-processes

$$u(t) = \frac{C(q^{-1})}{D(q^{-1})}e(t) \quad ; \quad w(t) = \frac{M(q^{-1})}{N(q^{-1})}v(t) \tag{2}$$

$$Ee(t)^2 = \lambda_e \qquad Ev(t)^2 = \lambda_v \qquad \rho = \lambda_v/\lambda_e$$

where $e(t)$ and $v(t)$ are mutually uncorrelated. They are stationary white and zero mean stochastic sequences.

Using measurements of the output $y(t)$, we seek the stable linear time-invariant estimator of the input

$$\hat{u}(t|t-m) = \frac{Q(q^{-1})}{R(q^{-1})}y(t-m) \tag{3}$$

which minimzes the mean square estimation error

$$Ez(t)^2 \triangleq E(u(t) - \hat{u}(t|t-m))^2 \tag{4}$$

See Figure 1. Depending on the sign of m, the estimator is a predictor, a filter or a fixed lag smoother.

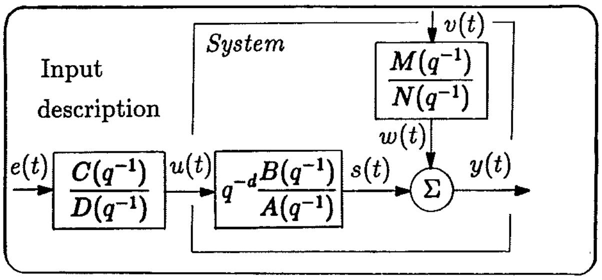

Figure 1: The input estimation problem. $u(t)$ is to be estimated from measurements of the output $y(t)$.

In Ahlén and Sternad (1985, 1989), a procedure for optimizing input estimators in the transfer function form (3) was derived. Compared to Kalman filtering, this approach is simple for scalar signals, and is well suited for self-tuning applications. It is based on the following assumptions:

- The signal $y(t)$ and input $u(t)$ can be described by linear models with structure (1), (2). All polynomial in (1) and (2), with degrees na, nb etc, are known.
- The denominators $D(q^{-1})$, $A(q^{-1})$ and $N(q^{-1})$ may have zeros inside or on, but not outside, the stability limit. The polynomials $C(q^{-1})B(q^{-1})N(q^{-1})$ and $M(q^{-1})A(q^{-1})D(q^{-1})$ have no common factors with zeros on the unit circle.

For any polynomial $P = P(q^{-1})$, let $P_* = P(q)$ and $\overline{P} = q^{-np}P_*$. The optimal input estimator is then given by

$$\hat{u}(t|t-m) = \frac{Q}{R}y(t-m) = \frac{Q_1NA}{\beta}y(t-m) \tag{5}$$

where β is the stable monic solution to a spectral factorization equation

$$r\beta\beta_* = CBNC_*B_*N_* + \rho MADM_*A_*D_* \tag{6}$$

with r being a scalar. The polynomial Q_1, together with a polynomial L_*, is the unique solution to the linear polynomial equation

$$q^{m+d}C_*B_*N_*C = r\beta_*Q_1 + qDL_* \tag{7}$$

with polynomial degrees

$$\begin{aligned} nQ_1 &= \max(nc-m-d, nd-1) \\ nL &= \max(nc+nb+nn+m+d, n\beta)-1 \end{aligned} \tag{8}$$

With z substituted for q, the minimal mean square error is

$$Ez(t)^2_{min} = \frac{\lambda_e}{2\pi j}\oint \frac{LL_* + \rho CMAC_*M_*A_*}{r\beta\beta_*}\frac{dz}{z} \tag{9}$$

From a practical point of view, we can hardly expect all polynomials to be known a priori. In many applications, as indicated in the previous section, it is however reasonable to assume $q^{-d}B(q^{-1})/A(q^{-1})$ to be known. In Ahlén (1986, 1988), the identifiability properties of the input estimation problem was investigated under the following asssumptions

- $A(q^{-1})$, $B(q^{-1})$, d and polynomial degrees nc, nd, nm, nn are known a priori.
- All polynomials in (1),(2) are asymptotically stable except the B-polynomial, which may be unstable.
- The polynomial pairs

$$\begin{matrix} (A,B),(C,D),(M,N),(B,D) \\ (\overline{B},A),(\overline{B},D),(A,N),(N,D) \end{matrix} \tag{10}$$

  are all coprime.
- The only measurable information is $\phi_y(e^{i\omega})$, the spectral density of the output.

Introducing the numbers

$$\Delta D \triangleq nd - nc \quad ; \quad \Delta N \triangleq nn - nm \tag{11}$$

$$k \triangleq \max(na + \Delta D, nb + \Delta N) \quad ; \quad \ell \triangleq 2\min(nn, nd)$$

it was found that

$$k \geq 1 \tag{12a}$$

is necessary, and

$$k > \ell \tag{12b}$$

is sufficient for parameter identifiability.

If the N-polynomial is known a priori, the number of unknown parameters decrease by nn. We thus set $N(q^{-1}) = 1$ and $nn = 0$ in (10)-(12), which implies that $k \geq 1$ is both necessary and sufficient for parameter identifiability.

3 A self-tuning input estimation algorithm

From (7), we find that C, B, N, D and β are required, but only B is known. We thus have to estimate (C, D, N, β) in some way using the output measurements only. Consider an innovations model of the output $y(t)$. Assuming A, D and N to be stable, the innovations model is given by

$$y(t) = \frac{\beta}{ADN}\tilde{y}(t)$$

where β is the stable spectral factor from (6) and $\tilde{y}(t)$ is the innovation sequence. The A-polynomial is known and can be filtered out. Thus, β/DN, may be estimated. Unfortunately, C and M are related to β through the nonlinear equation (6). Note, however that only the product CC_* is needed in (7). A key idea is to formulate (6) as a *linear* system of equations in the coefficients of CC_* and MM_*. Since we have an estimate of β, the system can be solved. Assume the N-polynomial to be known a priori or equal to unity. This leads to the algorithm below.

3.1 The basic algorithm

When the measurement noise is described by a moving average process or ARMA-process with known AR-part, use the following algorithm.

Algorithm 1.

> Assume $A(q^{-1})$, $B(q^{-1})$, $N(q^{-1})$, d, $n\beta$, nd, nc and nm to be known. For each data,
>
> 1. Generate the signal $x(t) = A(q^{-1})N(q^{-1})y(t)$.
> 2. Update recursive estimates of the β and D parameters from the data $x(t)$ using a prediction error method (PEM) or extended least squares (ELS). Call these estimates $\hat{\beta}$ and $\hat{D}$.
> 3. Solve the overdetemined linear system of equations in the coefficients of CC_* and MM_*, using $\hat{\beta}$ and $\hat{D}$,
>
> $$AA_*\hat{D}\hat{D}_*(MM_*) + BB_*NN_*(CC_*) = \hat{\beta}\hat{\beta}_* \qquad (13)$$
>
> with the least squares method, using e.g. singular value decomposition. Call these estimates $\hat{C}\hat{C}_*$ and $\hat{M}\hat{M}_*$. Before the solution is used in the next step, check the condition number of the Sylvester matrix $\mathbf{S}(AA_*\hat{D}\hat{D}_*, BB_*NN_*)$.
>
> 4. Solve the linear polynomial equation (7) using the estimates $\hat{\beta}$, $\hat{D}$ and $\hat{C}\hat{C}_*$
>
> $$q^{m+d}\hat{C}\hat{C}_*B_*N_* = \hat{\beta}_*Q_1 + q\hat{D}L_*$$
>
> with respect to Q_1 and L_*, using the degrees (8).
>
> 5. Perform the filtering (5), or, alternatively
>
> $$\hat{u}(t|t-m) = \frac{Q_1(q^{-1})}{\hat{\beta}(q^{-1})}x(t-m)$$

Some comments and interpretations are now in order.

1. We assume the system to be parameter identifiable, i.e. (10) and (12b) to hold. With a known $N(q^{-1})$, the chances to have parameter identifiability are good. The correct polynomial degrees are assumed known. This restrictive assumption can be avoided. Over-parametrized models is discussed in Section 3.3.

2. For time-invariant systems, the polynomials $\hat{\beta}$ and $\hat{D}$ in the innovations model can always be correctly estimated asymptotically, by using PEM-identification. See e.g. Ljung and Söderström (1983). Time-variable parameters may be tracked by means of a forgetting factor.

3. The estimate $\hat{\beta}$ is monic. If we let $\lambda_\varepsilon \triangleq E\tilde{y}(t)^2$, Step 3 gives (using a monic $\hat{\beta} = \beta$), $\hat{C}\hat{C}_* = (\lambda_\epsilon/\lambda_\varepsilon)CC_*$. From (6), it is seen that $\lambda_\epsilon/\lambda_\varepsilon = 1/r$. Step 4 will thus be the same as solving (7), with L scaled by the factor $1/r$.

4. In the algorithm, there is no need to perform any spectral factorization like (6), since the innovations model is estimated directly. Instead, the linear system of equations in Step 3 must be solved. In the transient phase, when $\hat{\beta}$ and $\hat{D}$ approaches β and D, Step 3 computes the least squares solution, the solution closest to CC_* and MM_*. Asymptotically, when $\hat{\beta} = \beta$ and $\hat{D} = D$, there is a unique and exact solution given by $\hat{C}\hat{C}_* = CC_*$ and $\hat{M}\hat{M}_* = MM_*$.

5. To reduce the computational requirements, the linear system defined by (13) can be transformed into a minimal order system with $n\beta + 1$ equations and $nc + nm + 2$ unknowns. This is achieved by eliminating rows and columns which are redundant due to symmetry. See Ahlén (1986).

6. Although the system is known to be identifiable when (10) and (12) holds, it may happen that the estimate $\hat{D}$ and BN sometimes have almost common factors as $\hat{D}$ converges towards D. This will cause a rank deficiency of $\mathbf{S}(AA_*\hat{D}\hat{D}_*, BB_*NN_*)$ in Step 3. The singular values of $\mathbf{S}$ must be checked. (When singular value decomposition is used for solving the LS problem, this requires no extra computations.) If the condition number is large, the Q_1-parameters are not updated.

The computational complexity of Algorithm 1 is presented in Table 1. The approximate number of mult-add operations required per sample, assuming all degrees $na, nb, \ldots$ etc $= n$, is estimated. A smoothing lag $\ell \triangleq -m \geq 0$ is assumed. [1] With $n = 1$ and $\ell = 2$, approximately 700 floating point mults+adds are required per sample.

Table 1. The computational complexity of Algorithm 1.

1:	Prefiltering	$2n$
2:	Identification	$30n^2 + 23n$
3:	LS solution	$56n^3 + 152n^2 + 136n$
4:	Linear system (7)	$\frac{1}{3}(4n+\ell+2)^3 + \frac{3}{2}(4n+\ell+2)^2$
5:	Filtering	$6n + \ell$

It should be noted that there is no need to recalculate the filter (steps 3 and 4) for each sample. Typically, it can be recalculated every 5'th to 10'th sample.

3.2 An extended algorithm

If $N(q^{-1})$ is unknown and the noise cannot be described with sufficient accuracy by a low order MA-model, adaptive input estimation becomes more complex. For general unknown ARMA-noise, we thus suggest the following algorithm.

Algorithm 2.

> Assume $A(q^{-1})$, $B(q^{-1})$, d, $n\beta$, nd, nc, nm and nn to be known. For each data,
>
> 1 Generate the signal $x(t) = A(q^{-1})y(t)$.
>
> 2 Identification as in Algorithm 1. Estimate $\hat{\beta}/\widehat{DN}$.
>
> 3 Solve the over-determined linear system of equations in the coefficients of CNC_*N_* and DMD_*M_*
>
> $$AA_*(DMD_*M_*) + BB_*(CNC_*N_*) = \hat{\beta}\hat{\beta}_*$$
>
> with the least squares method. Call these estimates $\widehat{DMD_*M_*}$ and $\widehat{CNC_*N_*}$.
>
> 4' This step is new compared to Algorithm 1. Perform spectral factorization on $\widehat{CNC_*N_*}$, obtained in step 3, in order to get $\widehat{CN}$. Use model reduction in in order to eliminate the common factor $\hat{N}$ of $\widehat{CN}$ and $\widehat{DN}$.
>
> 4-5 Identical to Algorithm 1. Use $\hat{N}$ instead of N in step 4, and $Q_1\hat{N}$ instead of Q_1 in the filter.

[1] In Step 3 of Table 1, the least squares solution is assumed to be computed by singular value decomposition, using the Golub-Reinsch algorithm (Golub and Van Loan 1983). Redundancies due to symmetry in the linear system are eliminated before computing the solution.

Compared to Algorithm 1, we have introduced a spectral factorization and model reduction. The price to be paid for less a priori information is a more complex algorithm.

3.3 Robustifying modifications

Two modifications of the Algorithms 1 and 2 are required in order to obtain a safe behaviour.

1. Common factors in the innovations model $\hat{\beta}/\widehat{ND}$ should be eliminated by *model reduction. Overparametrization* can thus be handeled, and the requirement of known polynomial degrees can be relaxed. Model reduction via balanced realizations, see eg Moore (1981) is one of many alternatives.

2. Rank deficiency in Step 3, caused by (nearly) common factors in $\hat{D}$ and B or $\hat{D}$ and N, must be detected.

With model reduction implemented, we only need to know the relative degrees ΔD and ΔN defined in (11). ~~For reduced models, it is possible to check the conditions (12) for parameter identifiability.~~

It should be mentioned that β, being an ARMA model numerator, is sometimes rather difficult to estimate. The estimates are often noisy when exponential forgetting is used. The variation of the estimator coefficients $(Q_1,\hat{\beta})$ can be reduced substantially by using low-pass filtered $\hat{\beta}$-polynomial coefficients in the design calculations.

If the input properties in certain frequency bands are of interest, a filtered input $\bar{u}(t) = (S(q^{-1})/T(q^{-1}))u(t)$ can be estimated, instead of $u(t)$ itself, with simple substitutions in the algorithm. See Ahlén and Sternad (1989). Low pass filtering will reduce the high-frequency gain of the estimation filter (5).

4 A numerical example

We illustrate Algorithm 1 by a simple example.

Example 1. Assume the true system to be given by

$$\begin{aligned} A(q^{-1}) &= 1-0.9q^{-1} & B(q^{-1}) &= 1+0.25q^{-1} \\ C(q^{-1}) &= 1+0.5q^{-1} & D(q^{-1}) &= 1-0.7q^{-1} \\ M(q^{-1}) &= 1.25 & d &= 1 \end{aligned}$$

An optimal one lag smoother ($m=-1$) is to be designed. The system is parameter identifiable according to (10)-(12). Assume correct model orders and the identification to give consistent estimates $\hat{\beta}=\beta$ and $\hat{D}=D$. Thus

$$\begin{aligned} \hat{\beta}(q^{-1}) &= 1-0.4314q^{-1}+0.1739q^{-2} \quad ; \quad r=6.38 \\ A\hat{D}A_*\hat{D}_* &= 0.63q^2-2.608q+3.9569-2.608q^{-1}+0.63q^{-2} \\ BB_* &= 0.25q+1.125+0.25q^{-1} \end{aligned}$$

According to Step 3, we have to solve

$$\begin{bmatrix} 0.63 & 0.25 & 0 & 0 \\ -2.608 & 1.125 & 0.25 & 0 \\ 3.9569 & 0.25 & 1.125 & 0.25 \\ -2.608 & 0 & 0.25 & 1.125 \\ 0.63 & 0 & 0 & 0.25 \end{bmatrix} \begin{bmatrix} \overline{m}_o \\ \overline{c}_1 \\ \overline{c}_o \\ \overline{c}_1 \end{bmatrix} = \begin{bmatrix} 0.1739 \\ -0.5064 \\ 1.2164 \\ -0.5064 \\ 0.1739 \end{bmatrix}$$

The least squares solution, in this case exact, is

$$[\overline{m}_o \;\; \overline{c}_1 \;\; \overline{c}_o \;\; \overline{c}_1]^T = [0.2449 \;\; 0.0784 \;\; 0.1959 \;\; 0.0784]^T$$

which correspond to $\hat{\lambda}_e = 0.1567$, $\hat{C} = 1+0.5q^{-1}$, $\hat{m}_o^2 = \hat{\lambda}_v = 0.25$. Note that $\hat{\lambda}_e = 1/r$ and $\lambda_v = m_o^2/r$.

As the left hand side of (7), we obtain

$$\hat{C}\hat{C}_*B_*N_* = 0.0196q^2+0.1273q+0.2155+0.0784q^{-1}$$

Since $nQ_1 = nC = 1$ and $nL = 1$ according to (8), the polynomial equation in Step 4 is found to be

$$0.0196q^2+0.1273q+0.2155+0.0784q^{-1} =$$

$$(0.1739q^2-0.4314q+1)(Q_o+Q_1q^{-1})+q(1-0.7q^{-1})(\ell_1q+\ell_o)$$

Multiplying both sides by q^{-2} and evaluating for equal powers of q^{-1} gives

$$\begin{bmatrix} 1 & 0 & 0.1739 & 0 \\ -0.7 & 1 & -0.4314 & 0.1739 \\ 0 & -0.7 & 1 & -0.4314 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \ell_1 \\ \ell_o \\ Q_o \\ Q_1 \end{bmatrix} = \begin{bmatrix} 0.0196 \\ 0.1273 \\ 0.2155 \\ 0.0784 \end{bmatrix}$$

The solution is

$$\begin{aligned} Q_1(q^{-1}) &= 0.4322+0.0784q^{-1} \\ L_1(q^{-1}) &= 0.2613-0.0556q^{-1} \end{aligned}$$

Thus, the optimal one lag smoothing input estimator (5) is

$$\hat{u}(t|t+1) = \frac{0.4322-0.3106q^{-1}-0.0705q^{-2}}{1-0.4314q^{-1}+0.1739q^{-2}}y(t+1)$$

The corresponding loss is found to be $Ez(t)^2 = 0.82$.

5 Simulations

Consider Algorithm 1 described in Section 3.1. Three examples will be presented in order to illustrate its behaviour.

Example 2: The system is described in Example 1. Correct model orders are assumed and SNR = 100. Identification with ELS is used from $t = 0-200$ and PEM from $t = 201-1000$. A forgetting factor $\lambda(t)$, starting with $\lambda(0) = 0.95$ and increasing to unity, is used.

Example 3: Same as in Example 2 up to $t = 500$. At $t = 500$, $D(q^{-1})$ changes abruptly to $D(q^{-1}) = 1-0.9q^{-1}$. (The SNR then changes from 100 to 590.) The forgetting factor is increasing towards 0.98.

Example 4: Same as in Example 2. The innovations model is overparametrized. It is assumed that $n\hat{\beta} = 3$ and $n\hat{d} = 2$. The forgetting factor is increasing towards unity.

In all examples described above, one future data was used ($m=-1$). The initial parameter estimates were $\hat{D} = 1$ and $\hat{\beta} = 1$. The covariance matrix was initialized as a unit matrix.

Figure 2a,b show the true and estimated input for Example 2, $t = 0-200$ and $800-1000$, respectively. In Figure 2c, the corresponding input estimation filter parameters are displayed.

Reasonable estimates of the input are obtained already after 100 data. Beyond $t = 800$, the estimator perform very well. This can be verified by comparing the estimated parameters $\hat{Q}_1(1000)$, $\hat{\beta}(1000)$ with the ones derived in Example 1. We have $(\beta_1 \;\; \beta_2 \;\; Q_o \;\; Q_1) = (-0.431 \;\; 0.174 \;\; 0.432 \;\; 0.078)$ and $(\hat{\beta}_1 \;\; \hat{\beta}_2 \;\; \hat{Q}_o \;\; \hat{Q}_1)_{1000} = (-0.345 \;\; 0.185 \;\; 0.416 \;\; 0.157)$. The loss (4) was estimated from $\widehat{Ez(t)}^2 = \frac{1}{900}\sum_{t=101}^{1000}(u(t) - \hat{u}(t|t+1))^2$ using both true and estimated parameters. This gave $\widehat{Ez(t)}^2_{est} = 0.96$ and $\widehat{Ez(t)}^2_{true} = 0.85$, which should be compared with the theoretical minimal variance $Ez(t)^2 = 0.82$ from (9). The close values verify the performance.

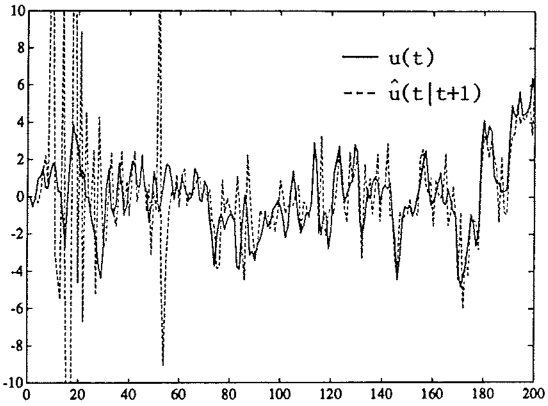

Figure 2a. Example 2. True and estimated input, $t = 0 - 200$, correct model order, SNR=100. The forgetting factor $\lambda(t) \to 1$.

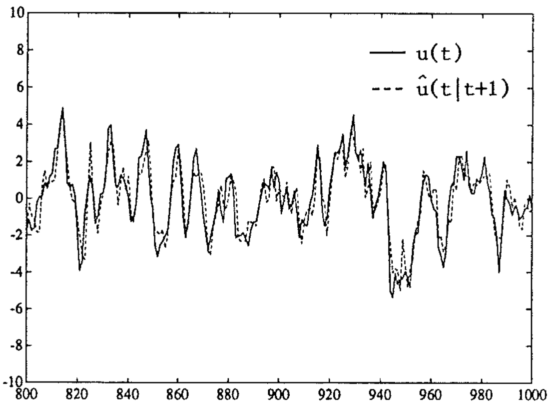

Figure 2b. Same as in Figure 2a, but for $t = 800 - 1000$.

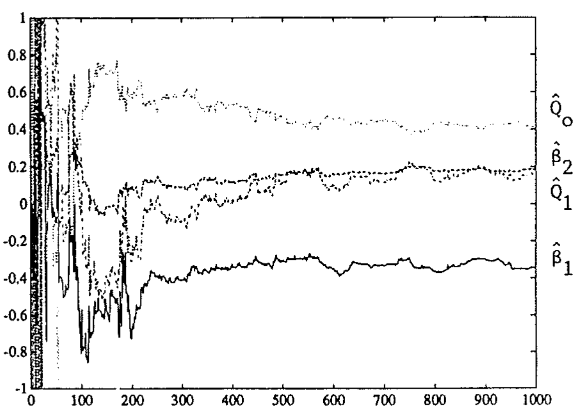

Figure 2c. Example 2. Input estimation filter parameters $\hat{\beta}(t)$, $\hat{Q}_1(t)$, $t = 0 - 1000$, with correct model order and SNR=100. The forgetting factor $\lambda(t) \to 1$.

In Figure 3a-3b, the true and estimated input for Example 3 is shown for $t = 450 - 650$ and $t = 800 - 1000$. Figure 3c displays the filter coefficiens $\hat{\beta}(t)$ and $\hat{Q}_1(t)$.

Regarding Figures 3, we conclude that the input estimator performs well. As can be seen from Figure 3c, the parameter estimates follows the underlying parameter change in D. The $\hat{\beta}$-parameters were slightly low pass filtered in order to obtain smooth variations in $\hat{Q}_1$. Although there is not a dramatically improvement in the obtained input estimate, see Figure 3a, it can be detected. From $t = 550$, there is an improved fit in the peaks.

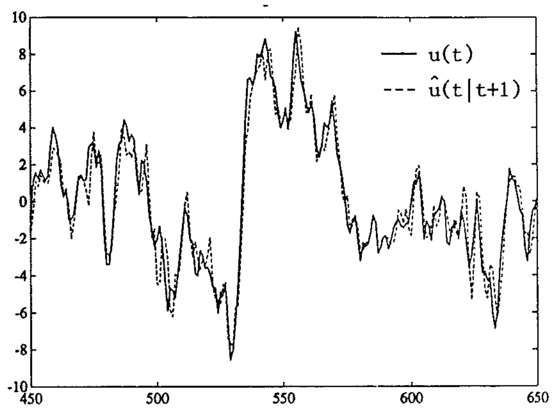

Figure 3a. Example 3. True and estimated input, $t = 450 - 650$. Correct model order assumed. At time $t = 500$, $D(q^{-1})$ changes abruptly from $1 - 0.7q^{-1}$ to $1 - 0.9q^{-1}$. The SNR thus increases from 100 to 590. $\lambda(t) \to 0.98$.

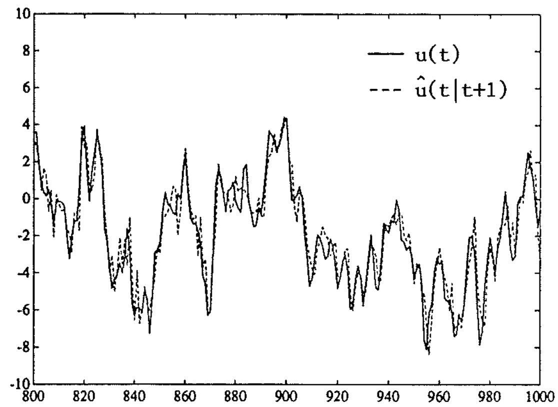

Figure 3b. Same as in Figure 3a, but for $t = 800 - 1000$.

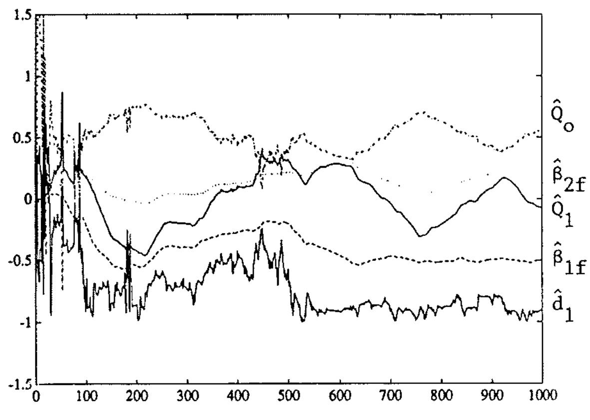

Figure 3c. Example 3. $\hat{D}(t)$ and input estimating filter parameters $\hat{\beta}(t)$, $\hat{Q}_1(t)$, for $t = 0 - 1000$, correct model order. The forgetting factor $\lambda(t) \to 0.98$.

Example 4 with Figures 4a and 4b illustrates the difficulties with overparametrization. In Figure 4a, a typical interval of the simulated data is shown. The result is not very nice and should deter every serious user. However, this may be avoided, as pointed out in Section 3.3. Figure 4b shows the trajectories of the superfluous zeros of $\hat{D}$ and $\hat{\beta}$. From $t = 500$, they follow each other closely. There is almost a common factor in the estimated transfer function $\hat{\beta}/\hat{D}$. It can be eliminated by model reduction. The model order is then reduced to that of Example 2 and we can expect a nice behaviour.

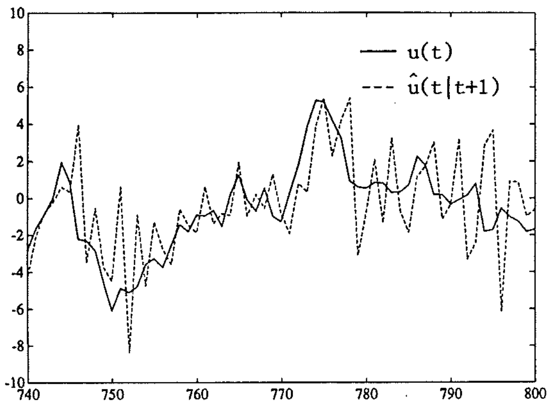

Figure 4a. Example 4. True and estimated input. Overparametrized innovations model, $n\hat{\beta} = 3$, $n\hat{d} = 2$. The a priori known relative degree of C/D leads to $n\hat{c} = 2$. SNR=100, $t = 740 - 800$. The forgetting factor $\lambda(t) \to 1$.

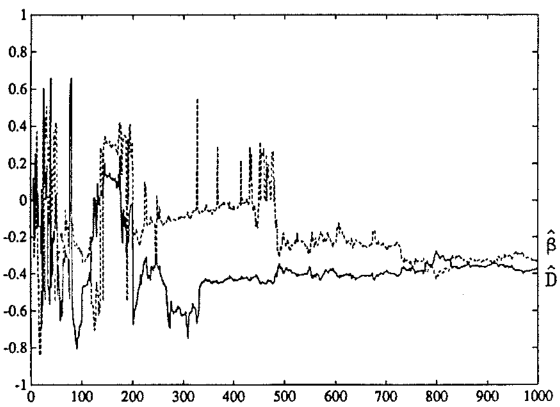

Figure 4b. Trajectories of superfluous zeros of $\hat{\beta}$ and $\hat{D}$. Overparametrized innovations model $n\hat{\beta} = 3$, $n\hat{d} = 2$.

6 Conclusion

The input estimation problem has been considered in an adaptive framework. We have presented an algorithm based on identification of an innovations model. Identifiability results guarantee a unique decomposition of input description and measurement noise. Two linear systems of equations are the main tools for obtaining the input estimate. The estimator may be a predictor, a filter or a smoother. The adaptive algorithm was illustrated by a numerical example and simulation experiments. Simulations with correct model order behaved well. An example illustrated difficulties with overparametrization of the innovations model. It was concluded that the difficulties may be avoided by model reduction. Since only second order statistics is used, non-minimum phase properties cannot be estimated from output data only. This is a limitation in some applications, such as digital channel equalization. Presently, work is carried out on the problem of eliminating slow transducer dynamics by means of input estimation.

7 References

Ahlén A and M Sternad (1985) Optimal input estimation: a polynomial approach. *Report UPTEC 85103R*, Department of Technology, Uppsala University, Sweden.

Ahlén A (1986) Input estimation with application to differentiation. *PhD dissertation*, Uppsala University, Sweden.

Ahlén A (1988) Identifiability of the deconvolution problem. *Proc of IFAC Symposium on Identification and System Parameter Estimation*, Beijing, China, 1387-1392.

Ahlén A and M Sternad (1989) Optimal deconvolution based on polynomial methods. *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol ASSP-37, February 1989.

Carlsson B, T Söderström and A Ahlén (1987) Digital differentiating filters. *Report UPTEC 8792R*, Department of Technology, Uppsala University, Sweden.

Commenges D (1984) The deconvolution problem: Fast algorithms including the preconditioned conjugate gradient to compute a MAP estimator. *IEEE Transactions on Automatic Control*, vol-29, 229-243.

Demoment G and R Reynaud (1985) Fast minimum variance deconvolution. *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol ASSP-33, 1324-1326.

Deng Z L (1985) White-noise filter and smoother with application to seismic data deconvolution. *IFAC Symposium on Identification and System Parameter Estimation*, York, UK, 621-624.

Fitch S M and L Kurz (1975) Recursive equalization in data transmission - A design procedure and performance evaluation. *IEEE Transactions on Communication*, vol COM-23, 546-550.

Golub G H and C F Van Loan (1983) *Matrix computations.* North Oxford Academic, Oxford.

Ljung L and T Söderström (1983). *Theory and Practice of Recursive Identification.* MIT Press, Cambridge, Mass.

Mendel J M (1983) *Optimal Seismic Deconvolution. An Estimation-based Approach.* Academic Press, New York.

Moir T J, T G Vishwanath and D R Campbell. Real-time self-tuning deconvolution filter and smoother. *International Journal of Control*, vol 45, 969-985, 1987.

Moore B C (1981) Principal component analysis in linear systems: controllability, observability and model reduction. *IEEE Trans. on Automatic Control*, vol AC-26, 17-27.

Qureshi S U H (1985) Adaptive equalization. *Proceeding of the IEEE*, vol 73, 1349-1387.

Sternad M and A Ahlén (1988) A correspondence between input estimation and feedforward control. *International Journal of Control*, to appear.