

ESTIMATION OF SIGNAL PARAMETERS IN EXOGENOUS PROCESS MODELS

J. STÖCKLE

Institut für Prozeßmeßtechnik und Prozeßleittechnik,
Universität Karlsruhe,
German Federal Republic

Summary

This paper gives a survey of the methods which can be used to determine the parameters of certain exogenous process models. The least squares estimation method is described explicitly and properties of the estimator are mentioned from the view of digital filter theory. It is further shown how the dynamic properties of time invariant linear systems can be improved by the use of the described estimation procedure.

Introduction

In this paper we shall be concerned with sampled signals which can be represented by the discrete time function

$$y(k) = \sum_{i=1}^n b_i q_i^k + e(k). \quad (1)$$

The variables $b_i, q_i, i = 1, \dots, n$, are assumed to be non-random, where $q_i \neq q_j$ for $i \neq j$ and $e(k)$ will be a zero mean stochastic process. Eq. (1) represents a special case of an exogenous process model.

Let us consider for a moment the homogeneous n -th order difference equation

$$x(k) + a_1 x(k-1) + \dots + a_n x(k-n) = 0. \quad (2)$$

The z -transform of the coefficients a_i is

$$A(z) = \sum_{i=0}^n a_i z^{-i}$$

with $a_0 = 1$. Assuming that the n zeros z_1, \dots, z_n of $A(z)$ are different, we obtain for the general solution of the difference equation

$$x(k) = \sum_{i=1}^n \alpha_i z_i^k$$

with arbitrary coefficients α_i . Therefore the signal $y(k)$ defined by Eq. (1) satisfies the relation

$$y(k) = x(k) + e(k) \quad (3)$$

where $x(k)$ is the solution of an n -th order homogeneous difference equation with the property

$$A(q_i) = 0, \quad i = 1, \dots, n.$$

Combining Eq. (2) and Eq. (3) we get

$$y(k) = - \sum_{i=1}^n a_i x(k-i) + e(k).$$

Substituting $x(k-i) = y(k-i) - e(k-i)$ leads to

$$y(k) = - \sum_{i=1}^n a_i [y(k-i) - e(k-i)] + e(k)$$

and thus

$$\sum_{i=0}^n a_i y(k-i) = \sum_{i=0}^n a_i e(k-i) \quad (4)$$

with $a_0 = 1$.

Equation (4) is called a special case of an ARMA (autoregressive moving average) model, because the AR and MA parameters are the same. The determination of the parameters b_i and q_i in the exogenous process model will be performed in two steps. First the coefficients of the ARMA model are estimated. Then one determines the zeros of $A(z)$ and thus gets estimates for q_1, \dots, q_n . The second step is to estimate b_1, \dots, b_n with the knowledge of q_1, \dots, q_n . Sometimes the q_i 's are known and fixed so that their estimation can be omitted.

In the following we will give two examples for processes which can be described with an exogenous process model. Let us first consider a process which consists of p harmonic components and an additive white noise [1].

$$y(k) = \sum_{i=1}^p \hat{y}_i \sin(\omega_i k T + \Phi_i) + e(k),$$

$$E\{e(k) e(j)\} = \sigma_e^2 \delta_{kj}.$$

By use of Euler's formula we get

$$y(k) = \sum_{i=1}^p \frac{\hat{y}_i}{2j} [e^{j(\omega_i k T + \Phi_i)} - e^{-j(\omega_i k T + \Phi_i)}] + e(k)$$

and after some algebra

$$y(k) = \sum_{i=1}^n b_i q_i^k + e(k)$$

where

$$\begin{aligned} n &= 2p \\ q_i &= e^{j\omega_i T}, \quad q_{p+i} = q_i^* \\ b_i &= \frac{\hat{y}_i}{2j} e^{j\Phi_i}, \quad b_{p+i} = b_i^* \\ i &= 1, \dots, p \end{aligned}$$

Thus $y(k)$ and $e(k)$ satisfy Eq. (4) with $a_0 = 1$ and $A(q_i) = 0, i = 1, \dots, n$. It can be shown [1], that

$$\Phi_y \mathbf{a} = \sigma_e^2 \mathbf{a}, \quad (5)$$

where Φ_y is the Toeplitz autocorrelation matrix of the process $y(k)$ and

$$\mathbf{a}^T = [a_0, a_1, \dots, a_n].$$

Equation (5) is an eigenproblem in which the coefficient vector \mathbf{a} is the normalized ($a_0 = 1$) eigenvector of Φ_y belonging to the minimum eigenvalue σ_e^2 [1]. Therefore an estimate for \mathbf{a} can be obtained by estimating Φ_y and determining the eigenvector corresponding to the minimum eigenvalue. Computing the zeros of $A(z)$ gives estimates for q_1, \dots, q_n .



Fig. 1. Linear system

Another application of the exogenous process model is to improve the dynamic behavior of slow measurement systems. Consider the system shown in Fig. 1, where $G(s)$ is the transfer function of a linear, time invariant measurement system, $u(t)$ is the input, $y(t)$ the output of the system. If we apply the step signal $u(t) = u_0 \sigma(t)$ to the system we receive a characteristic transient behavior. Fig. 2 shows two typical step responses. For the sake of simplicity we have assumed that the system is settled, when the step occurs and

$$E = G(0) = 1,$$

where E is the sensitivity of the system.

Examples for measurement systems with relatively long rise times are certain types of balances. Here the step signal corresponds to the replacement of the mass. In order to improve the dynamical behavior of the system first we

must describe the output signal mathematically. The relation between input and output is given by the convolution integral

$$y(t) = \int_{-\infty}^{\infty} g(t-\tau) u(\tau) d\tau.$$

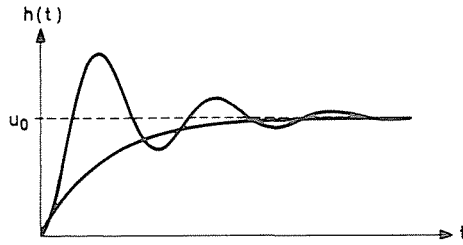


Fig. 2. Typical step responses

Applying $u(t) = u_0, t \geq 0$ leads to

$$y(t) = u_0 \int_0^{\infty} g(t-\tau) d\tau + \int_{-\infty}^0 g(t-\tau) u(\tau) d\tau. \quad (6)$$

The first term is the step response of the system, the second term depends on the behavior of $u(t)$ in the past and cannot be ignored in general. In the case of a rational transfer function we have

$$G(s) = \frac{\sum_{i=0}^m b_i s^i}{\sum_{i=0}^n a_i s^i} = \frac{b_m \prod_{i=1}^m (s - \beta_i)}{a_n \prod_{i=1}^n (s - \alpha_i)}$$

where $\alpha_1, \dots, \alpha_n$ are the poles and β_1, \dots, β_m the zeros of $G(s)$. Under the assumption that the poles are different and $m < n$ we can perform partial fraction expansion which leads to

$$G(s) = \sum_{i=1}^n \frac{K_i}{s - \alpha_i}. \quad (7)$$

Taking into account that $E = G(0) = 1$ we have

$$\sum_{i=1}^n \frac{K_i}{\alpha_i} = -1. \quad (8)$$

Inverse Laplace-transform of Eq. (7) yields

$$g(t) = \sum_{i=1}^n K_i e^{z_i t}.$$

Inserting this into Eq. (6) and integrating leads to

$$y(t) = \sum_{i=1}^n e^{z_i t} K_i \left[L_i + \frac{u_0}{\alpha_i} \right] - u_0 \sum_{i=1}^n \frac{K_i}{\alpha_i}$$

where

$$L_i = \int_{-\infty}^0 e^{-z_i \tau} u(\tau) d\tau.$$

Using Eq. (8) and the abbreviations

$$b_0 = u_0$$

$$b_i = K_i \left[L_i + \frac{u_0}{\alpha_i} \right]; \quad i = 1, \dots, n$$

we have

$$y(t) = b_0 + \sum_{i=1}^n b_i e^{z_i t}.$$

Equidistant sampling leads to

$$y(k) = b_0 + \sum_{i=1}^n b_i e^{z_i k T}$$

where T is the sample rate.

Denoting $q_i = e^{z_i T}$ we have

$$y(k) = b_0 + \sum_{i=1}^n b_i q_i^k$$

and with $q_0 = 1$

$$y(k) = \sum_{i=0}^n b_i q_i^k$$

for $k \geq 0$.

In practice we usually have to deal with measurement errors and we write

$$y(k) = \sum_{i=0}^n b_i q_i^k + e(k) \quad (9)$$

which of course is an exogenous process model.

In order to make the system faster we can try to get an estimate for b_0 as soon as possible. To do so we need to know q_1, \dots, q_n (see next section).

Because $y(k)$ is not stationary as in the case of harmonic components, the procedure described there cannot be used. Other methods have been established which can be used to estimate the coefficients of the corresponding difference equation [2].

Least Squares Estimation Procedure

Now we want to estimate the parameters b_1, \dots, b_n in the exogenous process model

$$y(k) = \sum_{i=1}^n b_i q_i^k + e(k)$$

where we assume that q_1, \dots, q_n are real and known exactly. First we make N measurements $y(0), \dots, y(N-1)$ which can be written in matrix form

$$\begin{bmatrix} y(0) \\ \vdots \\ y(N-1) \end{bmatrix} = \begin{bmatrix} 1 & \dots & 1 \\ q_1 & \dots & q_n \\ & \vdots & \\ q_1^{N-1} & \dots & q_n^{N-1} \end{bmatrix} \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} + \begin{bmatrix} e(0) \\ \vdots \\ e(N-1) \end{bmatrix}$$

or with the above abbreviations

$$\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{e}. \quad (10)$$

Often it is advantageous to write \mathbf{X} in the partitioned form

$$\mathbf{X} = [\mathbf{q}_1, \dots, \mathbf{q}_n]$$

where

$$\mathbf{q}_i^T = [1, q_i, \dots, q_i^{N-1}], \quad i = 1 \dots n.$$

The least squares estimate for \mathbf{b} is given by

$$\hat{\mathbf{b}} = [\mathbf{X}^T \mathbf{X}]^{-1} \mathbf{X}^T \mathbf{y}.$$

Because of $q_i \neq q_j, i \neq j$, the columns of \mathbf{X} are linear independent. Thus \mathbf{X} has maximum rank and therefore $[\mathbf{X}^T \mathbf{X}]^{-1}$ always exists. If we use the abbreviation

$$\mathbf{D} = [\mathbf{X}^T \mathbf{X}]^{-1} \mathbf{X}^T$$

we get for the estimator

$$\hat{\mathbf{b}} = \mathbf{D}\mathbf{y}. \quad (11)$$

From the definition of \mathbf{D} easily follows

$$\mathbf{DX} = \mathbf{I}. \quad (12)$$

The least squares estimator has some properties which we will summarize in the following (see for example [3]).

- The LS-estimator is a linear estimator. This follows at once from Eq. (11).
- Because of Eq. (12) the LS-estimator is unbiased under the condition $E\{\mathbf{e}\} = \mathbf{0}$, i.e.

$$E\{\hat{\mathbf{b}}\} = \mathbf{b}.$$

This follows easily from Eq. (11). Taking expectations we have

$$E\{\hat{\mathbf{b}}\} = \mathbf{D} E\{\mathbf{y}\}$$

Using Eq. (10), (12) and $E\{\mathbf{e}\} = \mathbf{0}$ we get the desired result.

- Under the condition, that $e(k)$ is discrete white noise, i.e.

$$E\{e(i)e(k)\} = \sigma_e^2 \delta_{ik},$$

the LS-estimator is the best linear estimator in the sense $E\{(\hat{b}_i - b_i)^2\} = \min.$, $i = 1, \dots, n$.

- The LS-estimator is the best estimator as such in the sense $E\{(\hat{b}_i - b_i)^2\} = \min.$, $i = 1, \dots, n$, if the errors $e(k)$ are jointly gaussian distributed.
- If the errors $e(k)$ form discrete white noise we get for the estimators' covariance matrix

$$\mathbf{V}_{\hat{\mathbf{b}}} = E\{(\hat{\mathbf{b}} - \mathbf{b})(\hat{\mathbf{b}} - \mathbf{b})^T\} = \sigma_e^2 [\mathbf{X}^T \mathbf{X}]^{-1}. \quad (13)$$

In order to get some further insight into the working of the LS-estimator in our special case we write Eq. (11) row by row

$$\hat{b}_i = \mathbf{d}_i^T \mathbf{y}, \quad i = 1, \dots, n$$

where

$$\mathbf{d}_i^T = [d_{i0}, \dots, d_{i,N-1}]$$

is the i -th row of \mathbf{D} .

Executing the matrix operation yields

$$\hat{b}_i = \sum_{k=0}^{N-1} d_{ik} y(k).$$

Defining

$$f_i(N-1-k) = d_{ik}$$

we have

$$\hat{b}_i = \sum_{k=0}^{N-1} f_i(N-1-k) y(k),$$

and thus

$$\hat{b}_i = \sum_{k=0}^{N-1} f_i(k) y(N-1-k).$$

Therefore \hat{b}_i can be considered as the output of a digital filter with impulse response $f_i(k)$ at time instant $N-1$. The z -transform of the impulse response $f_i(k)$ is given by

$$F_i(z) = \sum_{k=0}^{N-1} f_i(k) z^{-k}.$$

$F_i(z)$ is called a FIR (Finite Impulse Response) filter because its impulse response is of finite length. $F_i(z)$ shows the property

$$F_i(q_j) = \frac{1}{q_j^{N-1}} \delta_{ij} \quad i, j = 1, \dots, n. \quad (14)$$

Proof: By partitioning \mathbf{D} and \mathbf{X} in rows and columns respectively we get from Eq. (12):

$$\mathbf{d}_i^T \mathbf{q}_j = \delta_{ij} \quad i, j = 1, \dots, n$$

or in an explicit form

$$\sum_{k=0}^{N-1} d_{ik} q_j^{-k} = \delta_{ij}.$$

Using $f_i(N-1-k) = d_{ik}$ leads to

$$\sum_{k=0}^{N-1} f_i(N-1-k) q_j^{-k} = \delta_{ij}$$

and after some algebra

$$q_j^{N-1} \sum_{k=0}^{N-1} f_i(k) q_j^{-k} = \delta_{ij}.$$

With the definition of $F_i(z)$ Eq. (14) follows immediately.

$F_i(z)$ can be written in the form

$$F_i(z) = \sum_{k=0}^{N-1} f_i(k) z^{-k} = \frac{\sum_{k=0}^{N-1} f_i(k) z^{N-1-k}}{z^{N-1}} = \frac{f_i(0) \prod_{k=1}^{N-1} (z - z_k)}{z^{N-1}}$$

where the $N-1$ zeros of $F_i(z)$ are denoted with z_1, \dots, z_{N-1} .

From Eq. (14) we have

$$F_i(q_i) = \frac{1}{q_i^{N-1}}.$$

Therefore $f_i(0)$ is defined uniquely by the zeros of $F_i(z)$

$$f_i(0) = \frac{1}{\prod_{k=1}^{N-1} (q_i - z_k)}.$$

Another conclusion from Eq. (14) is that $q_j, j = 1, \dots, n, j \neq i$, are zeros of $F_i(z)$. This leads to

$$F_i(z) = \frac{f_i(0)}{z^{N-1}} \prod_{\substack{k=1 \\ k \neq i}}^n (z - q_k) \prod_{k=n}^{N-1} (z - z_k).$$

Summarizing the above-said up to this point we see that the filter $F_i(z)$ is determined, with exception of $N - n$ zeros, exclusively through the property of unbiasedness of the LS-estimator. Because of the uniqueness of the estimator the remaining zeros must result from the demand 'least squares'.

Numerical examples showed that these zeros are arranged in a characteristic manner in the z -plane. In the case $q_i = 1$ it was found that all remaining zeros are on the unit circle [4]. That this is actually true could be proved mathematically by the author. The proof is rather lengthy and therefore will be omitted here. Further it can be shown that all remaining zeros are outside the unit circle if $0 < q_i < 1$ or inside the unit circle if $q_i > 1$.

Another property of the estimator may be derived from the definition

$$\mathbf{D} = [\mathbf{X}^T \mathbf{X}]^{-1} \mathbf{X}^T.$$

If we denote $[\mathbf{X}^T \mathbf{X}]^{-1} = [\alpha_{ij}]$ and make use of the partitioned forms of \mathbf{D} and \mathbf{X} we get

$$\mathbf{d}_i = \alpha_{i1} \mathbf{q}_1 + \dots + \alpha_{in} \mathbf{q}_n.$$

Therefore

$$d_{ik} = \alpha_{i1} q_1^{-k} + \dots + \alpha_{in} q_n^{-k}$$

or with $f_i(N - 1 - k) = d_{ik}$

$$f_i(k) = \alpha_{i1} q_1^{N-1-k} + \dots + \alpha_{in} q_n^{N-1-k}.$$

Using the abbreviation $a_{ij} = \alpha_{ij} q_j^{N-1}$ we get the remarkable result

$$f_i(k) = a_{i1} q_1^{-k} + \dots + a_{in} q_n^{-k} = \sum_{j=1}^n a_{ij} q_j^{-k}, \quad i = 1, \dots, n. \quad (15)$$

The n coefficients $a_{ij}, j = 1, \dots, n$, can be determined in the above manner from $[\mathbf{X}^T \mathbf{X}]^{-1}$ or they follow from Eq. (14). Taking z -transform of Eq. (15) and

making use of the geometrical summation formula we obtain

$$F_i(z) = \sum_{j=1}^n a_{ij} \frac{1 - (q_j z)^{-N}}{1 - (q_j z)^{-1}}.$$

Thus we can decompose the nonrecursive filter $F_i(z)$ into n recursive filters. This is a very interesting fact especially under the aspect of an economical realization of the filter.

As shown above, the estimate \hat{b}_i can be considered as the output of a digital filter at time instant $N - 1$. If we observe the output of this filter all the time and denote this signal with $x_i(k)$ we have

$$x_i(k) = \sum_{j=0}^{N-1} f_i(j) y(k-j).$$

Inserting $y(k-j) = \sum_{l=1}^n b_l q_l^{k-j} + e(k-j)$ yields

$$x_i(k) = \sum_{l=1}^n b_l q_l^k \sum_{j=0}^{N-1} f_i(j) q_l^{-j} + \sum_{j=0}^{N-1} f_i(j) e(k-j), \quad k \geq N-1,$$

which can be written as

$$x_i(k) = \sum_{l=1}^n b_l q_l^k F_i(q_l) + \sum_{j=0}^{N-1} f_i(j) e(k-j)$$

using the definition of $F_i(z)$.

With the help of Eq. (14) we have the result

$$x_i(k) = b_i q_i^{k-N+1} + \sum_{j=0}^{N-1} f_i(j) e(k-j).$$

Taking expectations yields

$$E\{x_i(k)\} = b_i q_i^{k-N+1}. \quad (16)$$

From Eq. (16) we see that $x_i(k)$ gives an unbiased estimate for b_i at time instant $N - 1$. In the special case $q_i = 1$, b_i is estimated unbiased over all time.

Example

To illustrate the above-mentioned properties of the estimation procedure we will discuss a simple example. Let us consider a first order linear time invariant system with time constant T_s . If the input signal is $u(t) = u_0$, $t \geq 0$, we

can write for the sampled output signal (see Eq. (9))

$$y(k) = b_0 + b_1 q^k + e(k)$$

where

$$b_0 = u_0$$

$$b_1: \text{dependent on } u(t), t < 0$$

$$q = e^{-T/T_s}.$$

For the sake of simplicity we assume that $e(k)$ is discrete white noise, i.e.

$$E\{e(i)e(k)\} = \sigma_e^2 \delta_{ik}.$$

Now we want to estimate the coefficients b_0, b_1 of the exogenous process model. The impulse responses of the estimators for b_0 and b_1 are

$$f_0(k) = \frac{1}{N \frac{1+q^N}{1+q} - \frac{1-q^N}{1-q}} \left\{ \frac{1+q^N}{1+q} - q^{N-1-k} \right\}$$

$$f_1(k) = - \frac{1}{N \frac{1-q^{2N}}{1-q^2} - \frac{(1-q^N)^2}{(1-q)^2}} \left\{ \frac{1-q^N}{1-q} - N q^{N-1-k} \right\}$$

$$k = 0, \dots, N-1$$

which corresponds to Eq. (15). If we define the measuring time $T_M = (N-1)T$ we have

$$q = e^{-\frac{T_M}{T_s} \frac{1}{N-1}}.$$

Therefore the estimators depend but on the relation T_M/T_s and the filter order N . For filter order $N = 10$ and $T_M/T_s = 1$, i.e. $q = 0.895$, the impulse responses are plotted in Fig. 3 and Fig. 4.

The associate amplitude responses are shown in Fig. 5 and Fig. 6.

Because of the fact that the frequency response is the z -transform on the unit circle we deduce from Fig. 5 that $F_0(z)$ has exactly 8 zeros on the unit circle. This is actually true as we see from Fig. 7 where the exact positions of the zeros in the z -plane are plotted. We recognize that the 9-th zero of $F_0(z)$ is $q = 0.895$ which is also in accordance with the properties mentioned in the preceding section. Considering the zeros of $F_1(z)$ we recognize from Fig. 8 first that $F_1(z)$ has the real zero $z = 1$ as expected. The remaining 8 zeros are obviously slightly outside the unit circle which is due to the fact that $q < 1$.

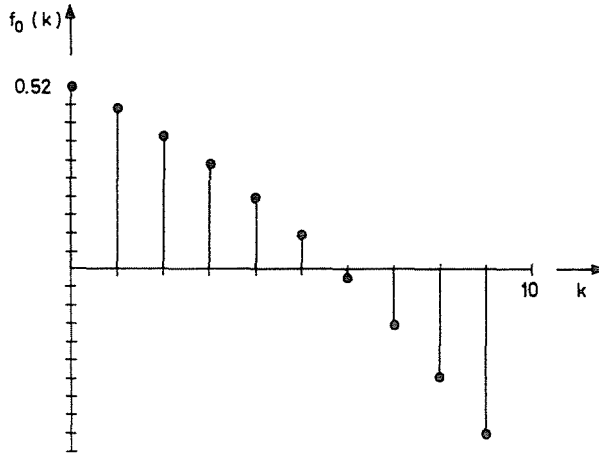


Fig. 3. Impulse response of the estimator for b_0

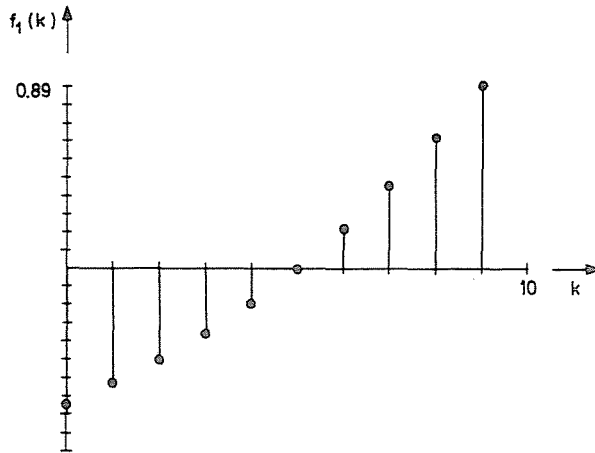


Fig. 4. Impulse response of the estimator for b_1

So far we didn't consider the quality of the estimation. If we evaluate the covariance matrix $V_{\hat{b}}$ (see Eq. (13)) in the special case of our example we get relations for the standard deviations of \hat{b}_0 and \hat{b}_1 in dependence of the relation T_M/T_s and filter order N . In Fig. 9 these functions are plotted for the cases $N = 10$ and $N = 100$.

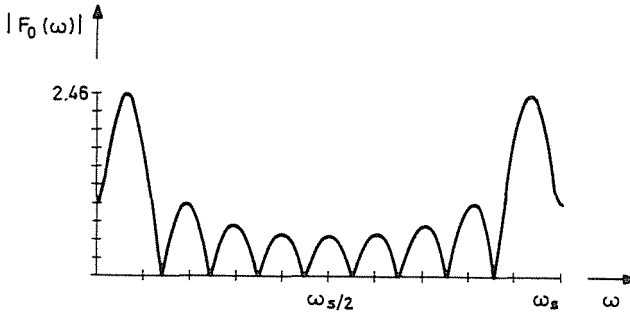


Fig. 5. Amplitude response $|F_0(\omega)|$. $\omega_s = \frac{2\pi}{T}$: sample frequency

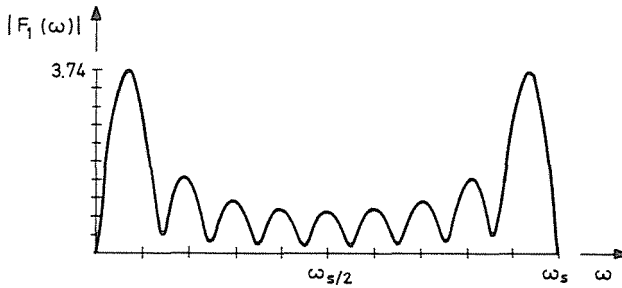


Fig. 6. Amplitude response $|F_1(\omega)|$

We see that the reduction of the error improves with increasing measuring time and filter order, respectively. For the above values $\left(\frac{T_M}{T_s} = 1, N = 10\right)$ we get

$$\frac{\sigma_{\hat{b}_0}}{\sigma_e} = 1.1 \quad \frac{\sigma_{\hat{b}_1}}{\sigma_e} = 1.6. \quad (17)$$

Often we are only interested in \hat{b}_0 . As we see from Eq. (17) the error is about the same if we wait until the system is settled or if we estimate with measuring time $T_M = T_s$ and filter order $N = 10$. Figure 10 shows the original step response of the system and the corresponding estimation signal where the estimator works all the time. For simplicity disturbances are neglected. Obviously the estimation procedure improves the dynamics of the system.

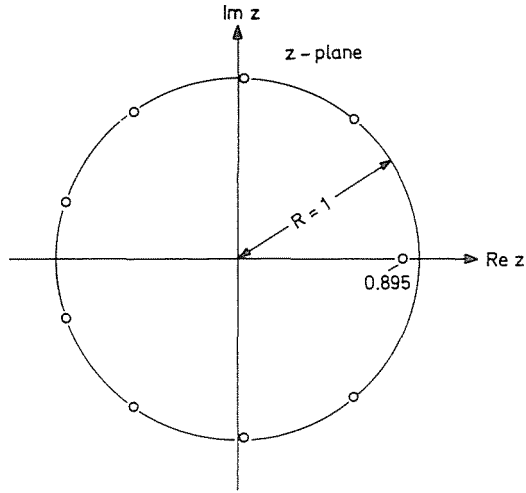


Fig. 7. Zero-plot of $F_0(z)$

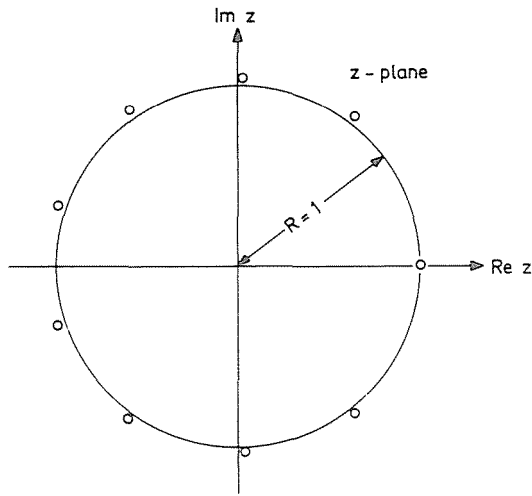


Fig. 8. Zero-plot of $F_1(z)$

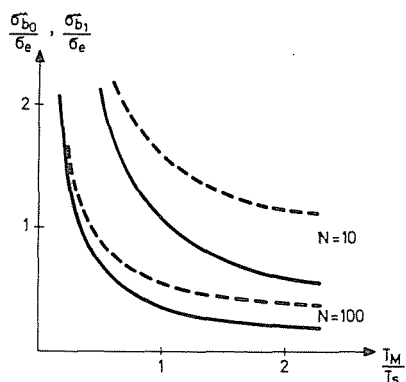


Fig. 9. Standard deviations of b_0 (—) and b_1 (---)

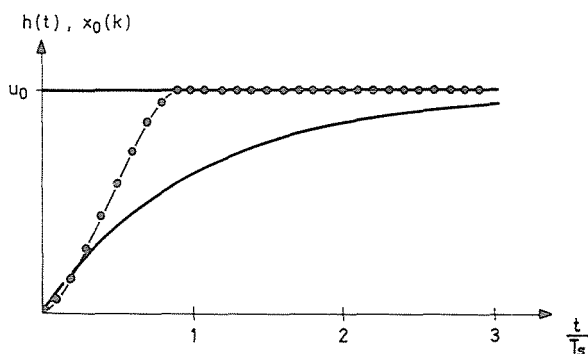


Fig. 10. Step response $h(t)$ (—) and estimation signal $x_0(k)$ (---○---)

References

1. HAYKIN, S.: Nonlinear Methods of Spectral Analysis. Springer Verlag, Berlin, 1979.
2. PROSS, D.: Approximations for the Bias and the Variance of a Class of Nonrecursive Autoregressive Estimators. Period. Polytechn. El. Eng. 28, 2—3 (1984)
3. DEUTSCH, R.: Estimation Theory. Prentice Hall, Englewood Cliffs, 1965.
4. PÜHLER, J.: Berechnung und Erprobung von FIR-Filtern zur näherungsweise Kompensation des Übertragungsverhaltens linearer zeitinvarianter dynamischer Systeme. Diplomarbeit am Institut für PMT/PLT, Universität Karlsruhe, 1981.
5. BAUMANN, U.: Schätzen der Parameter von Eingangssignalen in linearen Systemen. Dissertation am Institut für PMT/PLT, Universität Karlsruhe, 1975.

Joachim STÖCKLE Institut für Prozeßmeßtechnik und Prozeßleittechnik
 Universität Karlsruhe Hertzstr. 16. D-7500 Karlsruhe 21,
 German Federal Republic