ON THE MINIMAL VARIANCE CONTROL STRATEGIES

By

L. KEVICZKY and J. HETTHÉSSY*

Department of Automation, Technical University, Budapest

(Received September 13, 1973)

Presented by Prof. Dr. F. CSÁKI

Recently, as the problems of the computer aided process control have come to the foreground, the identification of the discrete time (sampled) systems and algorithms of the optimal control strategies based on the identification have great importance in the control engineering.

At present several identification methods are known that can be successfully applied to the solution of the parameter estimation of the linear discrete time systems [1]. These methods have proved their applicability and excellent filtering properties in practical use, and the identification of those sort of models can be considered to be solved. The models are suitable for designing optimal digital controller to a process (controlled plant) by some criteria. The function of the digital controller is performed by the process control computer.

One of the most important functions of the direct digital control (DDC) is to substitute the classical analogue (PI, PID etc.) constant value (and other) controllers by digital computer. The latest results clearly show that the discrete versions of the classical controllers are not sufficient to reach the theoretically best control for constant value, if the system is influenced not only by deterministic disturbances. If the purpose of the control is to minimize the oscillation of the controlled signal around the desired value, then a more complicated signal formation is needed than by the classical controllers, but this complicated signal formation can simply be performed by a digital computer. The oscillation of the controlled signal due to the various disturbances can be reduced by the minimal variance control. By this control strategy the expectable value of the square of the control error will be minimal in a stationary case.

First of all, in this paper the algorithm of the minimal variance control is discussed on the basis of K. J. ASTRÖM's works [1], when the identification of the controlled plant is previously made and so the system parameters can be assumed to be known for the design of the controller.

In the next section the adaptive control algorithm of V. PETERKA is considered [7], namely, a "self-tuning" control strategy can be established by

* Institute for Electrical Power Research.

the least squares method, in which the apriori knowledge of the process parameters is not necessary.

Then it is shown how the self-tuning (adaptive) control can be realized by the generalized least squares method.

Finally expressive simulation examples are presented in order to compare the various control strategies.

1. Minimal variance control of a constant and known system

Let the system, restricted to single-input single-output discrete time systems, be described by a linear time-invariant difference equation. It is assumed that the additive disturbances reduced to the output can be characterized by stochastic process with a rational spectral density [1]:

$$y(t) = \frac{B(z^{-1})}{A(z^{-1})}\, u(t-d) + \lambda \frac{C(z^{-1})}{A(z^{-1})}\, e(t);\ (t = 0, 1, 2, \ldots) \tag{1}$$

where $B(z^{-1})/A(z^{-1})$ is the pulse transfer function of the process, $\lambda C(z^{-1})/$ $/A(z^{-1})$ is the pulse transfer function of the disturbance referred to the process output, $u(t)$ is the control signal, $y(t)$ is the output of the noisy system, $e(t)$ is a sequence of independent normal variables with zero mean value and variance 1 (white noise). Further

$$\left.\begin{aligned} A(z^{-1}) &= 1 + a_1 z^{-1} + \ldots + a_n z^{-n}, \\ B(z^{-1}) &= b_0 + b_1 z^{-1} + \ldots + b_m z^{-m}, \\ C(z^{-1}) &= 1 + c_1 z^{-1} + \ldots + c_k z^{-k}, \end{aligned}\right\} \tag{2}$$

where $n \geq m$ and $n \geq k$ for physically realizable systems. It is assumed that the polynomials $A(z^{-1})$ and $C(z^{-1})$ have all their zeros inside the unit circle (in the z plain). According to the interpretation of the backward shift operator z^{-1} the difference equation of the system is:

$$\begin{aligned} y(t) = b_0 u(t-d) + b_1 u(t-d-1) + \ldots + b_m u(t-d-m) - a_1 y(t-1) - \\ - a_2 y(t-2) - \ldots - a_n y(t-n) + e(t) + c_1 e(t-1) + \ldots + c_k e(t-k). \end{aligned} \tag{3}$$

It can be seen that the process has a time delay d. The problem is to determine a sequence of the control signal $u(t)$ on the basis of the knowledge of the system parameters and the observations $y(t)$, in such a way that

$$E\left\{y^2(t)\right\} \tag{4}$$

would be as small as possible. $E\{\ldots\}$ denotes mathematical expectation. Let us write the earliest value of y influenced by $u(t)$ on the basis of Eq. (1):

$$y(t+d) = \frac{B(z^{-1})}{A(z^{-1})} u(t) + \lambda \frac{C(z^{-1})}{A(z^{-1})} e(t+d). \tag{5}$$

The second term of the right side is a linear function of $e(t+d), \ldots, e(t+1), e(t), e(t-1), \ldots$. Since $e(t), e(t-1), \ldots$ can be computed exactly from the observations by (1), but $e(t+d), \ldots, e(t+1)$ are independent of the observations, it is expedient to make this separation as follows:

$$y(t+d) = \frac{B(z^{-1})}{A(z^{-1})} u(t) + \lambda \left[F(z^{-1}) + z^{-d} \frac{G(z^{-1})}{A(z^{-1})} \right] e(t+d), \tag{6}$$

where

$$\left. \begin{aligned} F(z^{-1}) &= 1 + f_1 z^{-1} + \ldots + f_{d-1} z^{-(d-1)}, \\ G(z^{-1}) &= g_0 + g_1 z^{-1} + \ldots + g_{n-1} z^{-(n-1)} \end{aligned} \right\} . \tag{7}$$

The coefficients of the polynomials $F(z^{-1})$ and $G(z^{-1})$ can be computed from the following equality by comparison of the coefficients:

$$C(z^{-1}) = A(z^{-1}) F(z^{-1}) + z^{-d} G(z^{-1}) \cdot \tag{8}$$

Using the identity $z^{-d} e(t+d) = e(t)$ the equation (6) can be written in the following form:

$$y(t+d) = \frac{B(z^{-1})}{A(z^{-1})} u(t) = \frac{G(z^{-1})}{A(z^{-1})} e(t) + \lambda F(z^{-1}) e(t+d). \tag{9}$$

Eliminating $e(t)$ from (1) and replacing it into (9), and taking into account (8), we obtain:

$$y(t+d) = \frac{F(z^{-1}) B(z^{-1})}{C(z^{-1})} u(t) + \frac{G(z^{-1})}{C(z^{-1})} y(t) + \lambda F(z^{-1}) e(t+d). \tag{10}$$

In consequence of the separation discussed above the third term of the right side is independent of the first and second term, hence

$$E\{y^2(t+d)\} = E\left\{ \left[\frac{F(z^{-1}) B(z^{-1})}{C(z^{-1})} u(t) + \frac{G(z^{-1})}{C(z^{-1})} y(t) \right]^2 \right\} + \\ + E\{[\lambda F(z^{-1}) e(t+d)]^2\}. \tag{11}$$

It can be seen that only the first term of the right side, which is a non-negative expression, depends on $u(t)$, thus the whole right side is minimal by $u(t)$ only then, if the first term vanishes, from which the optimal control strategy is [1, 2]:

$$u(t) = -\frac{G(z^{-1})}{B(z^{-1})F(z^{-1})}\,y(t). \tag{12}$$

By such a control strategy the output is the following on the basis of (10):

$$y(t+d) = F(z^{-1})e(t+d) = \lambda[e(t+d) + f_1 e(t+d-1) + \ldots + f_{d-1}e(t+1)]. \tag{13}$$

On the basis of (13) we can say that the expectable value of the output is zero, thus the expectable value of the square of the output, that is the minimal value of the loss function, is equal to the variance of the output:

$$\begin{aligned}\operatorname{var}\{y(t+d)\} = E\{y^2(t+d)\} = E\{\lambda^2[e(t+d) + f_1 e(t+d-1) + \ldots \\ \ldots + f_{d-1}e(t+1)]^2\} = \lambda^2(1 + f_1^2 + \ldots + f_{d-1}^2),\end{aligned} \tag{14}$$

because the expectable value of the mixed terms vanishes, as $e(t+d), \ldots, e(t+1)$ are independent of $y(t), y(t-1), \ldots$ and $u(t-1), u(t-2), \ldots$.

By taking into account the previous considerations it can be seen that the parameters of the controller in (12) can be derived unanimously in the knowledge of the system parameters. The output, which can be regarded to be an error signal, is described by the moving average stochastic process by Eq. (13).

1.1. *The minimal variance control strategy as a prediction problem*

Make a prediction for $y(t+d)$ based on the available observations at the time t in such a way that

$$E\{[y(t+d) - \hat{y}(t+d \mid t)]^2\} \tag{15}$$

would be as small as possible, where $\hat{y}(\cdot \mid t)$ denotes the predicted value based on the available observations at the time t. Let us use the expression of $y(t+d)$ in Eq. (10) to Eq. (15), where the values before and after time t are well separated:

$$\begin{aligned}E\{[y(t+d) - \hat{y}(t+d \mid t)]^2\} = E\left\{\left[\frac{F(z^{-1})B(z^{-1})}{C(z^{-1})}u(t) + \frac{G(z^{-1})}{C(z^{-1})}y(t) + \right.\right. \\ \left.\left. + \lambda F(z^{-1})e(t+d) - \hat{y}(t+d \mid t)\right]^2\right\} = E\left\{\left[\frac{F(z^{-1})B(z^{-1})}{C(z^{-1})}u(t) + \right.\right. \\ \left.\left. + \frac{G(z^{-1})}{C(z^{-1})}y(t) - \hat{y}(t+d \mid t)\right]^2\right\} + E\left\{\lambda^2\left[F(z^{-1})e(t+d)\right]^2\right\}.\end{aligned} \tag{16}$$

Here it is used again as $e(t+d), \ldots, e(t+1)$ are independent. The observations available at the time t appear only in the first term, which is a non-negative expression, thus it is minimal if it vanishes, from which follows that

$$\hat{y}(t+d \mid t) = \frac{F(z^{-1})B(z^{-1})}{C(z^{-1})} u(t) + \frac{G(z^{-1})}{C(z^{-1})} y(t). \tag{17}$$

The prediction error is given by the following moving average stochastic process:

$$h(t+d) = y(t+d) - \hat{y}(t+d \mid t) = \lambda F(z^{-1}) e(t+d) =$$
$$= \lambda [e(t+d) + f_1 e(t+d-1) + \ldots + f_{d-1} e(t+1)] . \tag{18}$$

So the prediction error will be a white noise, if and only if, $d = 1$.
It follows from Eq. (17) that the predicted value of the output is a function of $u(t)$, thus with a proper choice of $u(t)$ it can always be achieved that the predicted value of the output should be a value given in advance. To satisfy the criterion

$$E\{y^2(t+d)\} = \min.$$

the desired value of the output must obviously be

$$\hat{y}(t+d \mid t) = 0 . \tag{19}$$

Substituting (19) into Eq. (17) we get again Eq. (12), and the control error $y(t+d)$ will be equal to the prediction error, as it is shown in Eq. (18).

1.2. *Control to a constant value*

Find the control strategy, which minimizes the following loss function:

$$E\{[y(t) - R]^2\} = E\{r^2(t)\}. \tag{20}$$

In Eq. (20) $R =$ const. denotes the desired value of the output and $r(t)$ is the error signal. By Eq. (17) we can write that

$$\hat{y}(t+d \mid t) = R = \frac{F(z^{-1})B(z^{-1})}{C(z^{-1})} u(t) + \frac{G(z^{-1})}{C(z^{-1})} y(t), \tag{21}$$

whereof the optimal control strategy is:

$$u(t) = \frac{C(z^{-1})R - G(z^{-1})y(t)}{B(z^{-1})F(z^{-1})} . \tag{22}$$

2. Minimal variance control of constant but unknown system using the least squares method

Previously we have seen that the parameters of the optimal controller can be determined unanimously in the knowledge of the system parameters. Since the system parameters are unknown the control step should be preceeded by an identification step. In the following section a method will be discussed, which determines the parameters of the controller directly and avoids the complicated dual control.

Consider the following system equation:

$$A(z^{-1})y(t) = B(z^{-1})u(t-d) + \lambda C(z^{-1})e(t)\,. \tag{23}$$

Since $y(t-1), \ldots, y(t-d+1)$ can be expressed by means of $y(t-d), \ldots, y(t-d-n+1)$; $u(t-d-1), \ldots, u(t-2d-m+1)$; $e(t-1), \ldots, e(t-d-k+1)$, therefore, a linear transformation always exists, and (23) can be written in the following form:

$$y(t) = T(z^{-1})y(t-d) + S(z^{-1})u(t-d) + \lambda R(z^{-1})e(t)\,, \tag{24}$$

where

$$\left.\begin{aligned} T(z^{-1}) &= t_0 + t_1 z^{-1} + \ldots + t_{n-1} z^{-(n-1)} \\ S(z^{-1}) &= s_0 + s_1 z^{-1} + \ldots + s_{m+d-1} z^{-(m+d-1)} \\ R(z^{-1}) &= 1 + r_1 z^{-1} + \ldots + r_{k+d-1} z^{-(k+d-1)}\,. \end{aligned}\right\} \tag{25}$$

According to Eq. (24) the value of $y(t+d)$ is given by

$$y(t+d) = T(z^{-1})y(t) + S(z^{-1})u(t) + \lambda R(z^{-1})e(t+d)\,. \tag{26}$$

To separate the values before and after time t, let us decompose the polynomial $R(z^{-1})$ as follows:

$$R(z^{-1}) = R_1(z^{-1}) + z^{-d}R_2(z^{-1})\,, \tag{27}$$

where

$$\begin{aligned} R_1(z^{-1}) &= 1 + r_1 z^{-1} + \ldots + r_{d-1} z^{-(d-1)}. \\ R_2(z^{-1}) &= r_d + r_{d+1} z^{-1} + \ldots + r_{k+d-1} z^{-(k-1)} \end{aligned} \tag{28}$$

Then

$$y(t+d) = T(z^{-1})y(t) + S(z^{-1})u(t) + \lambda R_2(z^{-1})e(t) + \lambda R_1(z^{-1})e(t+d)\,. \tag{29}$$

Expressing $e(t)$ from Eq. (24) and substituting it into Eq. (29), by means of (27) we get:

$$y(t+d) = \frac{T(z^{-1})R_1(z^{-1}) + R_2(z^{-1})}{R(z^{-1})}\,y(t) + \frac{S(z^{-1})R_1(z^{-1})}{R(z^{-1})}\,u(t) + \lambda R_1(z^{-1})e(t+d). \tag{30}$$

Compare this with Eq. (10) and we obtain:

$$F(z^{-1}) = R_1(z^{-1}) \,, \tag{31}$$

and the optimal control strategy is given by

$$u(t) = - \frac{T(z^{-1})R_1(z^{-1}) + R_2(z^{-1})}{S(z^{-1})R_1(z^{-1})} y(t) = - \frac{G(z^{-1})}{B(z^{-1})F(z^{-1})} y(t), \tag{32}$$

that is the optimal control strategy is the quotient of the "coefficients" of $y(t)$ and $u(t)$ in both cases. The other way around, by the feedback

$$u(t) = - \frac{P(z^{-1})}{Q(z^{-1})} y(t), \tag{33}$$

taking again the prediction error $\lambda R_1(z^{-1})e(t+d)$ in the prediction equation (30), choosing $P(z^{-1})$ as the coefficient of $y(t)$ and $Q(z^{-1})$ as the coefficient of $u(t)$, we will show that Eq. (33) gives the minimal variance control strategy. According to this:

$$y(t+d) = P(z^{-1})y(t) + Q(z^{-1})u(t) + \lambda R_1(z^{-1})e(t+d) \,. \tag{34}$$

Comparing this with Eq. (29) the following equation must hold:

$$T(z^{-1})y(t) + S(z^{-1})u(t) + \lambda R_2(z^{-1})e(t) = P(z^{-1})y(t) + Q(z^{-1})u(t) \,. \tag{35}$$

Taking into the consideration the feedback (33), as a limiting condition, Eq. (35) has the following form:

$$T(z^{-1})y(t) + S(z^{-1}) \left[- \frac{P(z^{-1})}{Q(z^{-1})} y(t) \right] + \lambda R_2(z^{-1})e(t) = 0. \tag{36}$$

Eliminating $e(t)$ from (24) and writing it to (36), using (27) we obtain:

$$[(R_1(z^{-1})T(z^{-1}) + R_2(z^{-1}))Q(z^{-1}) - R_1(z^{-1})S(z^{-1})P(z^{-1})]y(t) = 0. \tag{37}$$

(37) holds only then, if

$$\frac{P(z^{-1})}{Q(z^{-1})} = \frac{R_1(z^{-1})T(z^{-1}) + R_2(z^{-1})}{R_1(z^{-1})S(z^{-1})} \,. \tag{38}$$

According to Eq. (32) the right side is really an optimal controller. So we have shown that the minimal variance controller can be searched by means of such a prediction equation as (34). Moreover, it can be seen from Eq. (34) that the least squares estimation for the coefficients of the polynomials $P(z^{-1})$ and $Q(z^{-1})$ will be unbiased, since the residual is independent of the first two terms of the right side [3]. Comparing (38) with (32) it can be seen that the polynomials $P(z^{-1})$ and $Q(z^{-1})$ must be of order $(n-1)$ and $(m-1+d)$, respectively. Normalizing the controller based on Eq. (33):

$$\begin{aligned} P(z^{-1}) &= p_0 + p_1 z^{-1} + \dots + p_{n-1} z^{-(n-1)} \\ Q(z^{-1}) &= 1 + q_1 z^{-1} + \dots + q_{m+d-1} z^{-(m+d-1)} . \end{aligned} \tag{39}$$

Introduce the parameter vector

$$\mathbf{p}^T = [p_0, p_1, \dots, p_{n-1}, q_1, q_2, \dots, q_{m+d-1}] \tag{40}$$

and the observation vector

$$\begin{aligned} \mathbf{x}^T(t) = [&y(t), y(t-1), \dots, y(t-n+1), u(t-1), u(t-2), \dots, \\ &u(t-d-m+1)] . \end{aligned} \tag{41}$$

By Eq. (34) we can write that

$$y(t+d) = \mathbf{p}^T \mathbf{x}(t) + u(t) + \varepsilon(t+d) , \tag{42}$$

where

$$\varepsilon(t+d) = \lambda R_1(z^{-1}) e(t+d) \tag{43}$$

is the independent residual. The loss function to be minimized

$$\begin{aligned} E\{y^2(t+d)\} &= E\{[\mathbf{p}^T \mathbf{x}(t) + u(t) + \varepsilon(t+d)]^2\} = \\ &= E\{[\mathbf{p}^T \mathbf{x}(t) + u(t)]^2\} + E\{\varepsilon^2(t+d)\} \end{aligned} \tag{44}$$

is minimal if

$$u(t) = -\mathbf{p}^T \mathbf{x}(t) , \tag{45}$$

which is equivalent to Eq. (33).

Introduce the following notations for the least squares estimation of the parameter vector $\mathbf{p}$:

$$\left. \begin{aligned} \tau &= t + d \\ s(\tau) &= y(\tau) - u(\tau - d) . \end{aligned} \right\} \tag{46}$$

By these notations Eq. (42) takes the following form:

$$s(\tau) = \mathbf{p}^T \mathbf{x}(\tau - d) + \varepsilon(\tau) . \tag{47}$$

Determine the parameter vector $\mathbf{p}$ in such a way that

$$\sum_{\tau=0}^{t} [w^{t-\tau} \varepsilon(\tau)]^2 = \sum_{\tau=0}^{t} [w^{t-\tau}]^2 [s(\tau) - \mathbf{p}^T \mathbf{x}(\tau - d)]^2 \tag{48}$$

will be minimal. By means of value $|w| =$ const. < 1 appeared in the weighted least squares solution above, a so-called exponential forgetting strategy can be realized. It means that the observations referring to the old adjustment of the controller play a smaller role in the sum, than the observations referring to the new adjustment of the controller. Consider the following $(t + 1)$ equations:

$$s(\tau) = \mathbf{p}^T \mathbf{x}(\tau - d) + \varepsilon(\tau); \ \tau = 0, 1, 2, \ldots, t\,. \tag{49}$$

Writing (49) in a vector form we get

$$\mathbf{s}_t = \mathbf{X}_{t-d}\mathbf{p} = \varepsilon_t\,, \tag{50}$$

where

$$\left.\begin{aligned} \mathbf{s}_t^T &= [s(0),\ s(1),\ \ldots,\ s(t)]\,; \\ \varepsilon_t^T &= [\varepsilon(0),\ \varepsilon(1),\ \ldots,\ \varepsilon(t)]\,; \end{aligned}\right\} \tag{51}$$

$$\mathbf{X}_{t-d} = \begin{bmatrix} \mathbf{x}^T(-d) \\ \mathbf{x}^T(1-d) \\ \vdots \\ \mathbf{x}^T(t-d) \end{bmatrix}. \tag{52}$$

Introducing the weighting matrix

$$\mathbf{W}_t = \mathbf{diag}[w^t,\ w^{t-1},\ \ldots,\ w,\ 1] \tag{53}$$

the loss function (48) can be written as

$$[\mathbf{W}_t \varepsilon_t]^T [\mathbf{W}_t \varepsilon_t] \to \min. \tag{54}$$

Substituting (50) into (54) we obtain

$$[\mathbf{W}_t(\mathbf{s}_t - \mathbf{X}_{t-d}\mathbf{p})]^T [\mathbf{W}_t(\mathbf{s}_t - \mathbf{X}_{t-d}\mathbf{p})] \to \min_{\mathbf{p}}. \tag{55}$$

whereof the place of the minimum is by differentiation:

$$\hat{\mathbf{p}}_t = [(\mathbf{W}_t\mathbf{X}_{t-d})^T (\mathbf{W}_t\mathbf{X}_{t-d})]^{-1} \mathbf{W}_t\mathbf{X}_{t-d}\mathbf{s}_t\,. \tag{56}$$

Using the relations of the recursive least squares estimation [3, 5, 7] $\hat{\mathbf{p}}_t$ can be derived from $\hat{\mathbf{p}}_{t-1}$ by a recursive relation:

$$\hat{\mathbf{p}}_t = \hat{\mathbf{p}}_{t-1} + \mathbf{H}_t \mathbf{x}(t-d)[s(t) - \mathbf{x}^T(t-d)\hat{\mathbf{p}}_{t-1}], \tag{57}$$

where

$$\mathbf{H}_t = \frac{1}{w^2}\left\{ \mathbf{H}_{t-1} - \frac{[\mathbf{H}_{t-1}\mathbf{x}(t-d)][\mathbf{H}_{t-1}\mathbf{x}(t-d)]^T}{w^2 + \mathbf{x}^T(t-d)\mathbf{H}_{t-1}\mathbf{x}(t-d)} \right\}. \tag{58}$$

If $d = 1$ then the expression of $\hat{\mathbf{p}}_t$ is more simple. By (45) and (57) we can write that

$$\hat{\mathbf{p}}_t = \hat{\mathbf{p}}_{t-1} + \mathbf{H}_t\mathbf{x}(t-1)[y(t) - u(t-1) - \mathbf{x}^T(t-1)\hat{\mathbf{p}}_{t-1}] = \hat{\mathbf{p}}_{t-1} +$$
$$+ \mathbf{H}_t\mathbf{x}(t-1)[y(t) - u(t-1) + u(t-1)] = \hat{\mathbf{p}}_{t-1} + \mathbf{H}_t\mathbf{x}(t-1)y(t)\,. \tag{57'}$$

Thus in the knowledge of $\hat{\mathbf{p}}_t$ the equation of the controller is [7]:

$$u(t) = -\hat{\mathbf{p}}^T\mathbf{x}(t)\,. \tag{59}$$

3. Minimal variance control of constant but unknown system using the generalized least squares method

Studying the literature of system identification we can ever more often meet such kinds of identification methods, which estimate the disturbance model besides the process model. Such a method is the CLARKE's generalized least squares method. Now we will investigate how to control a system described by equation

$$A(z^{-1})y(t) = B(z^{-1})u(t-d) + \lambda C(z^{-1})e(t) \tag{60}$$

in order to minimize the variance of output, if the identification is made by the generalized least squares method.

In the previous section it could be seen that the system described by Eq. (60) could always be transformed into the following form:

$$y(t+d) = T(z^{-1})y(t) + S(z^{-1})u(t) + \lambda R(z^{-1})e(t+d)\,. \tag{61}$$

Introduce the parameter vector $\mathbf{p}$ and observation vector $\mathbf{x}(t)$ as follows:

$$\left.\begin{aligned} \mathbf{p}^T &= [t_0, t_1, \ldots, t_{n-1}, s_0, s_1, \ldots, s_{m+d-1}]\,, \\ \mathbf{x}^T(t) &= [y(t), y(t-1), \ldots, y(t-n+1), u(t), u(t-1), \ldots, u(t-m-d+1)]. \end{aligned}\right\} \tag{62}$$

On the basis of Eq. (61) we can write:

$$y(t) = \mathbf{x}^T(t-d)\mathbf{p} + v(t)\,, \tag{63}$$

where the residual

$$v(t) = \lambda R(z^{-1})e(t) \tag{64}$$

contains components before time t which are in correlation with observation. For N measurements Eq. (63) has the following vector form:

$$\mathbf{y}_t = \mathbf{X}_{t-d}\mathbf{p} + \mathbf{v}_t\,, \tag{65}$$

where

$$\begin{aligned} \mathbf{y}_t^T &= [y(t), y(t-1), \ldots, y(t-N)]\,, \\ \mathbf{X}_{t-d}^T &= [\mathbf{x}(t-d), \mathbf{x}(t-d-1), \ldots, \mathbf{x}(t-d-N)]\,, \\ \mathbf{v}_t^T &= [v(t), v(t-1), \ldots, v(t-N)]\,. \end{aligned} \tag{66}$$

The least squares estimation of $\mathbf{p}$

$$\hat{\mathbf{p}}_t = [\mathbf{X}_{t-d}^T\mathbf{X}_{t-d}]^{-1}\mathbf{X}_{t-d}\mathbf{y}_t \tag{67}$$

will be biased, because $v(t)$ is not independent of the observations. Approximate $\lambda/R(z^{-1})$ with a polynomial $H(z^{-1})$ of finite order:

$$\frac{\lambda}{R(z^{-1})} \cong H(z^{-1}) = 1 + h_1 z^{-1} + \ldots + h_r z^{-r}\,. \tag{68}$$

By Eq. (64) we can write that

$$H(z^{-1})v(t) = v(t) + \sum_{i=1}^{N} v(t-i)h_i = e_1(t) \cong e(t)\,, \tag{69}$$

where $e_1(t)$ is not exactly a white noise, because of the approximation of finite order, but the least squares estimation, which minimizes the loss function

$$\sum_{t=1}^{N} e_1^2(t) \tag{70}$$

will have a small bias. Substitute (69) into (61):

$$y(t) = T(z^{-1})y(t-d) + S(z^{-1})u(t-d) + \frac{1}{H(z^{-1})}\,e_1(t)\,. \tag{71}$$

Multiplying by $H(z^{-1})$ and using the commutativity of the shift operation we obtain:

$$[H(z^{-1})y(t)] = T(z^{-1})[H(z^{-1})y(t-d)] + S(z^{-1})[H(z^{-1})u(t-d)] + e_1(t). \quad (72)$$

Thus, the terms in angular brackets can be generated by an autoregressive smoothing. Denote the filtered values with superscript F:

$$\left.\begin{aligned} y^F(t) = H(z^{-1})y(t); \qquad y^F(t-d) &= H(z^{-1})y(t-d); \\ u^F(t-d) &= H(z^{-1})u(t-d)\,. \end{aligned}\right\} \quad (73)$$

The residual vector $\mathbf{v}_t$ can be computed by means of $\hat{\mathbf{p}}_t$ deriving from the biased estimation according to Eq. (67):

$$\mathbf{v}_t = \mathbf{y}_t - \mathbf{X}_{t-d}\hat{\mathbf{p}}_t\,. \quad (74)$$

Let

$$\mathbf{f}^T(t) = [-v(t-1), -v(t-2), \ldots, -v(t-r)]$$

and

$$\mathbf{h}^T = [h_1, h_2, \ldots, h]\,, \quad (75)$$

then

$$v(t) = \mathbf{f}^T(t)\mathbf{h} + e_1(t)\,. \quad (76)$$

Thus, the least squares estimation for h based on N measurements is:

$$\mathbf{h}_t = [\mathbf{F}_t^T\mathbf{F}_t]^{-1}\mathbf{F}_t^T\mathbf{v}_t\,, \quad (77)$$

where

$$\mathbf{F}_t^T = [\mathbf{f}(t), \mathbf{f}(t-1), \ldots, \mathbf{f}(t-N)]\,. \quad (78)$$

According to the CLARKE's generalized least squares iterative method, which gradually decreases the estimation error, the next step starts with the smoothing by (73) and the estimation of the system parameters is repeated on the basis of the filtered observation vectors:

$$\mathbf{x}^F(t) = H(z^{-1})\mathbf{x}(t); \qquad t = 1, 2, \ldots, N\,. \quad (79)$$

The estimation of the system parameters and the filter $H(z^{-1})$ can be written in a recursive form as are Eq. (57) and Eq. (58):

$$\hat{\mathbf{p}}_t = \hat{\mathbf{p}}_{t-1} + \mathbf{G}_t\mathbf{x}^F(t-d)[y^F(t) - \mathbf{x}^{TF}(t-d)\hat{\mathbf{p}}_{t-1}]\,, \quad (80)$$

where

$$\mathbf{G}_t = \frac{1}{w_1^2}\left\{\mathbf{G}_{t-1} - \frac{[\mathbf{G}_{t-1}\mathbf{x}^F(t-d)][\mathbf{G}_{t-1}\mathbf{x}^F(t-d)]^T}{w_1^2 + \mathbf{x}^{FT}(t-d)\mathbf{G}_{t-1}\mathbf{x}^F(t-d)}\right\} \tag{81}$$

and

$$\hat{\mathbf{h}}_t = \hat{\mathbf{h}}_{t-1} + \mathbf{D}_t\mathbf{f}(t)[v(t) - \mathbf{f}^T(t)\hat{\mathbf{h}}_{t-1}], \tag{82}$$

where

$$\mathbf{D}_t = \frac{1}{w_2^2}\left\{\mathbf{D}_{t-1} - \frac{[\mathbf{D}_{t-1}\mathbf{f}(t)][\mathbf{D}_{t-1}\mathbf{f}(t)]^T}{w_2^2 + \mathbf{f}^T(t)\mathbf{D}_{t-1}\mathbf{f}(t)}\right\}. \tag{83}$$

Here w_1 and w_2 are forgetting factors. Determine $u(t)$ based on $\hat{\mathbf{p}}_t$ and $\hat{\mathbf{h}}_t$ in such a way that

$$E\{y^2(t+d)\} \tag{84}$$

should be minimal. Introduce the notation

$$\widetilde{H}(z^{-1}) = h_1 z^{-1} + h_2 z^{-2} + \ldots + h_r z^{-r}. \tag{85}$$

According to Eq. (72) and (73) $y(t+d)$ is given by

$$y(t+d) = T(z^{-1})y^F(t) + S(z^{-1})u^F(t) - \widetilde{H}(z^{-1})y(t+d) + e_1(t+d). \tag{86}$$

Investigate the third term more precisely:

$$\widetilde{H}(z^{-1})y(t+d) = h_1 y(t+d-1) + \ldots + h_{d-1}y(t+1) + h_d y(t) + \ldots$$
$$\ldots + h_r y(t+d-r). \tag{87}$$

Let us express $y(t+1), \ldots, y(t+d-1)$ by (86):

$$y(t+1) = T(z^{-1})y^F(t-d+1) + S(z^{-1})u^F(t-d+1) - h_1 y(t) - \ldots$$
$$\ldots - h_r y(t-r+1) + e_1(t+1) = \hat{y}(t+1 \mid t) + e_1(t+1), \tag{88}$$

where $\hat{y}(t+1 \mid t)$ is the least squares estimation for $y(t+1)$ based on the observations available at the time t.

On the basis of Eq. (86) and (88):

$$y(t+2) = T(z^{-1})y^F(t-d+2) + S(z^{-1})u^F(t-d+2) - h_1 y(t+1) -$$
$$- h_2 y(t) - \ldots - h_r y(t-r+2) + e_1(t+2) = T(z^{-1})y^F(t-d+2) +$$
$$+ S(z^{-1})u^F(t-d+2) - h_1\hat{y}(t+1 \mid t) - h_2 y(t) - \ldots - h_r y(t-r+2) -$$
$$- h_1 e_1(t+1) + e_1(t+2) = \hat{y}(t+2 \mid t) - h_1 e_1(t+1) + e_1(t+2). \tag{89}$$

Going on from this procedure

$$y(t+d) = T(z^{-1})y^F(t) + S(z^{-1})u^F(t) - H_2(z^{-1})y(t) - \\ - H_1(z^{-1})y(t+d-1 \mid t) + L(z^{-1})e_1(t+d)\,, \tag{90}$$

where

$$\begin{aligned} H_1(z^{-1}) &= h_1 + h_2 z^{-1} + \ldots + h_{d-1} z^{-(d-2)}, \\ H_2(z^{-1}) &= h_d + h_{d+1} z^{-1} + \ldots + h_r z^{-(r-d)}, \\ L(z^{-1}) &= 1 + l_1 z^{-1} + \ldots + l_{d-1} z^{-(d-1)}. \end{aligned} \tag{91}$$

The coefficients of the polynomial $L(z^{-1})$ can be derived from the coefficients of the polynomial $H(z^{-1})$ by the following relation:

$$\mathbf{g}_{i+1} = \begin{bmatrix} \mathbf{0}^T \\ \cdots \\ \mathbf{I}_{d-2} \end{bmatrix} \mathbf{g}_i + \begin{bmatrix} \mathbf{g}_i^T \\ \cdots \\ \mathbf{0} \end{bmatrix} \mathbf{h}_1\,, \tag{92}$$

where

$$\begin{aligned} \mathbf{g}_1^T &= [1, 0, 0, \ldots, 0]\,, \\ \mathbf{g}_{d-1}^T &= [l_1, l_2, \ldots, l_{d-1}]\,, \\ \mathbf{h}_1 &= [-h_1, -h_2, \ldots, -h_{d-1}] \end{aligned} \tag{93}$$

and $\mathbf{I}_{d-2}$ is a unit matrix of order $(d-2)$.

To determine the control law, let us introduce the following notations:

$$u^F(t) = u(t) + \sum_{i=1}^{r} u(t-i)h_i = u(t) + \tilde{u}^F(t) \tag{94}$$

$$\tilde{S}(z^{-1}) = s_1 + s_2 z^{-1} + \ldots + s_{m+d-1} z^{-(m+d-2)}\,. \tag{95}$$

Using the notations above in Eq. (90) we obtain:

$$y(t+d) = [T(z^{-1})y^F(t) + \tilde{S}(z^{-1})u^F(t-1) + s_0\tilde{u}^F(t) + s_0 u(t) - \\ - H_2(z^{-1})y(t) - H_1(z^{-1})\hat{y}(t+d-1 \mid t)] + [L(z^{-1})e_1(t+d)]\,. \tag{96}$$

After raising to the second power, taking mathematical expectation in Eq. (96) and using the fact that $e_1(t)$ is asymptotically independent, we get:

$$E\{y^2(t+d)\} = E\{[T(z^{-1})y^F(t) + \tilde{S}(z^{-1})u^F(t-1) + s_0\tilde{u}^F(t) + s_0 u(t) - \\ - H_2(z^{-1})y(t) - H_1(z^{-1})\hat{y}(t+d-1 \mid t)]^2\} + \{E[L(z^{-1})e_1(t+d)]^2\}\,. \tag{97}$$

By means of $u(t)$ the first, non-negative term can be minimized, thus

$$u(t) = \frac{-1}{s_0}\left[T(z^{-1})y^F(t) + \tilde{S}(z^{-1})u^F(t-1) + s_0\tilde{u}^F(t) - H_2(z^{-1})y(t) - \right.$$
$$\left. - H_1(z^{-1})\hat{y}(t+d-1 \mid t)\right] \quad (98)$$

is an optimal strategy, by which

$$E\{y^2(t+d)\} = E\{[L(z^{-1})e_1(t+d)]^2\} = E\{e_1^2(t+d)\}(1 + h_1^2 + \ldots + h_{d-1}^2)\,. \quad (99)$$

4. Simulation results

In order to investigate the properties of these mentioned control algorithms we have simulated the situation in Fig. 1 by digital computer "Odra 1204". The difference equation of the system to be controlled was the following:

$$y(t) = 1.5y(t-1) - 0.54\,y(t-2) + 2\,u(t-1) - 1.8\,u(t-2) +$$
$$+ 3[e(t) + 0.2\,e(t-1) - 0.48\,e(t-2)] \quad (100)$$

that is a second order system was simulated with $d = 1$ and $\lambda = 3$.

We have investigated the following cases:

— uncontrolled case (UC);

— minimal variance control of constant and known system (KPC)

— minimal variance control of constant but unknown system using the least squares method (LSC);

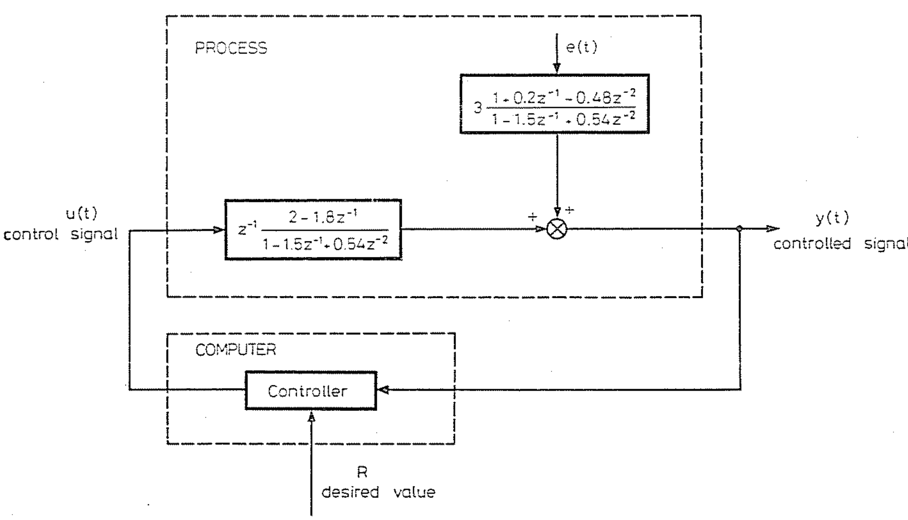


Fig. 1

— minimal variance control of constant but unknown system using the generalized least squares method (GLSC).

The desired value of the output (R) was counted as zero. The time curves for every case can be seen in Fig. 2, and the variance of the output (σ_y^2) is presented in Fig. 3.

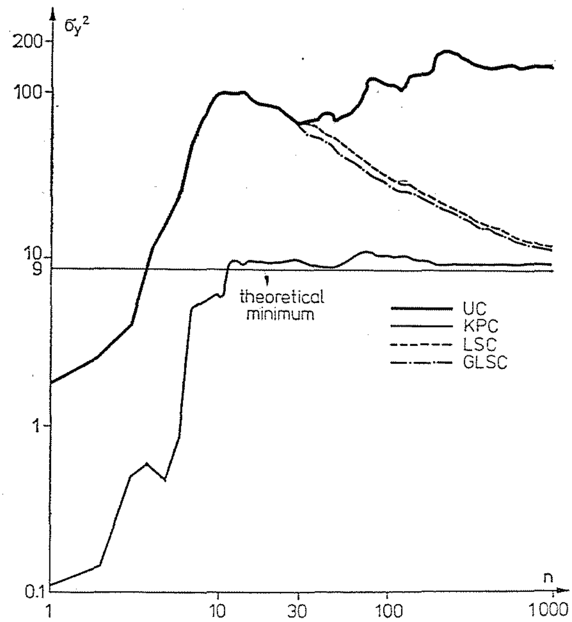


Fig. 3

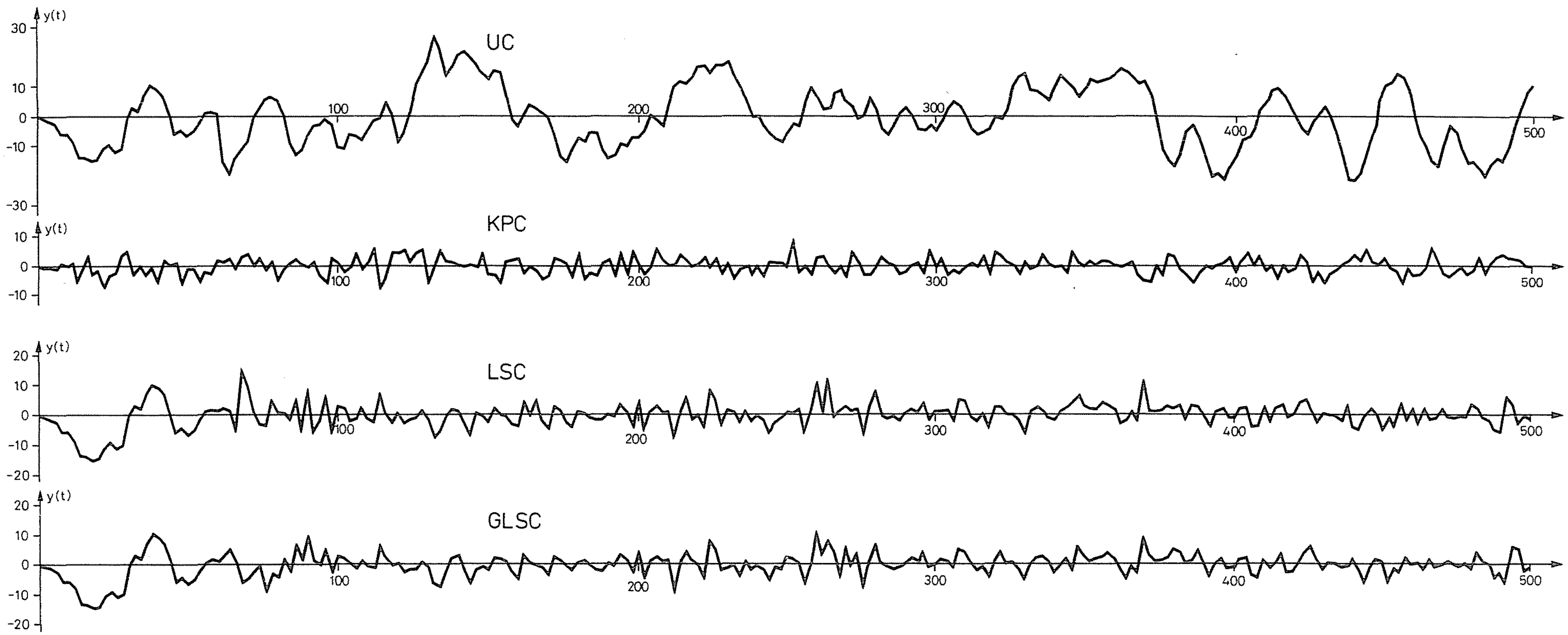
In the uncontrolled case $u(t)$ should be chosen in such a way that the mean value of the process output should be equal to the desired value of the controlled signal (R). Since the transfer factor of the process referring to the mean value is $\sum_{i=0}^{m} b_i \Big/ \left(1 + \sum_{i=1}^{n} a_i\right)$, therefore, the control signal should be equal to

$$u(t) = \frac{1 + \sum_{i=1}^{n} a_i}{\sum_{i=0}^{m} b_i} R = \text{constant.}$$

Choosing $R = 0$ the control signal $u(t)$ will be zero in the uncontrolled case.

The equation of the optimal controller can be computed on the basis of Eq. (8) and (12) as follows:

$$u(t) = -\frac{0.85 - 0.51z^{-1}}{1 - 0.9z^{-1}} y(t) \tag{102}$$



y(t)
30
UC
10
0
-10
-30
100
200
300
400
500
t
y(t)
KPC
10
0
-10
y(t)
20
LSC
-20
y(t)
GLSC

Fig. 2

and the control error by (18) is

$$h(t + d) = h(t + 1) = \lambda e(t + d) = 3e(t + 1) . \tag{103}$$

The variance of the controlled signal by Eq. (14) is $\sigma_y^2 = \lambda^2 = 9$.

At the control based on the least squares method (LSC) the estimation of $\mathbf{p}$ was started with a zero initial value and carried out by recursive relation (57). It means that we did not make an off-line estimation in order to get an apriori initial value. This is why the estimated value of $\mathbf{p}$ differs from its exact value during the first steps, so the adjustment of the controller is not optimal at all. Thus, it is worthwhile starting the control only after a few estimation steps. Up to that time the control signal corresponding to the uncontrolled case should be given for the system. When the simulation program was used, the control was started from the 30th step. It was also made at the generalized least squares method. We have seen that by this method the estimated system parameters are also needed for the estimation of the parameters of the filter. Therefore, in the case of effective generalized least squares method the estimation of the filter parameters should be started only if the estimation of the system parameters gives a relatively stationary value. At the simulation the estimation of the fourth order filter $H(z^{-1})$ was started at the 80th step.

5. Conclusions

It can be seen from the simulation results that the minimal variance control strategies are very effective, the variance of the control error can be considerably decreased by them (in our example the output variance decreased with one order). It is noteworthy that after a few steps the control of an unknown system has similarly nice qualitative properties (mean value, variance), as the control of a known system. It seems that the generalized least squares method (GLSC) suggested by us has the most complicated structure, but according to our experiences its asymptotical properties are the most advantageous.

On the basis of this paper we can establish that in the knowledge of system parameters the memory and computation time needed for the realization of optimal control algorithm by computer are very small and the program can easily be made. The adaptive control of the unknown systems needs far more store capacity and computation time, whereas the optimal control is always able to adopt to the process.

Studying the literature of the minimal variance control we can say that it is a subject worth dealing with. Several papers have already given account of practical applications and the general opinion is that these algorithms will constitute one of the most important field of direct digital control. It is so much the more expectable because there is no particular difficulty to extend these methods for multivariable systems.

Summary

In this paper the minimal variance control of discrete time, linear, time-invariant systems is discussed. The disturbances of the system are considered to be a stochastic process with rational spectral density referred to the output. In the different sections algorithms are shown for known and unknown systems. Besides the known methods a new method is suggested for the adaptive solution of the minimal variance control, where the generalized least squares method is used for the adaptive estimation of the parameters of the optimal controller.

References

1. ASTRÖM, K. J.: Introduction to Stochastic Control Theory. Academic Press, New York, 1970.
2. ASTRÖM, K. J.: On Self-Tuning Regulators. Automatica **9**, pp. 185, (1973).
3. ASTRÖM, K. J.—EYKHOFF, P.: System Identification — A Survey. Automatica **7**, pp. 123 (1971).
4. ASTRÖM, K. J.—WITTENMARK, B.: Problems of Identification and Control. Journal of Mathematical Analysis and Applications **34**, 1 (1971).
5. CLARKE, D. W.: Generalized-least-squares Estimation of the Process Parameters of a Dynamic Model. IFAC Symp. Prague (1967).
6. PETERKA, V.: On Steady State Minimum Variance Control Strategy. Kybernetika **8**, 3 (1972).
7. PETERKA, V.: Adaptive Digital Regulation of Ncisy Systems. IFAC Symp. Prague (1970).
8. TURTLE, D. P.—PHILLIPSON, P. H.: Simultaneous Identification and Control. Automatica **7**, pp. 445 (1971).
9. WIESLANDER, J.—WITTENMARK, B.: An Approach to Adaptive Control Using Real Time Identification. IFAC Symp. Prague (1970).

Dr. László KEVICZKY
Jenő HETTHÉSSY
H-1521 Budapest