

Multivariable adaptive control†

H. T. TOIVONEN‡

Keywords: *multivariable adaptive control.*

In recent years there has been an extensive interest in adaptive and self-tuning controllers, and there is a vast literature on various adaptive algorithms. The purpose of the present paper is to review some common approaches for multivariable adaptive control. The presentation concentrates on procedures which are based on stochastic controller design methods, but some close connections with other design techniques are also indicated.

1. Introduction

Adaptive and self-tuning controllers have been proposed as a means for tuning digital controllers for industrial processes. The idea is to use on-line identification in combination with a controller design method. The procedure is applied on-line to recompute the controller parameters at each sampling time using the identified model. In practice such a procedure is useful for systems with slowly time-varying parameters for keeping the controller properly tuned when the process dynamics change. The method can also be applied to time-invariant systems, for which manual tuning may be difficult. In this case an adaptive controller can first be used for tuning the controller parameters, and then be removed after the parameters have converged to proper values.

The present interest in self-tuning controllers was initiated by the work of Åström and Wittenmark (1973), who applied recursive least squares estimation and a minimum variance strategy to obtain a self-tuning minimum variance regulator. Similar approaches had previously been proposed by Kalman (1958) and Peterka (1970). There is presently a vast literature on adaptive and self-tuning controllers obtained by combining various on-line identification methods and different controller design procedures. Good surveys of various areas in the field are available, see for example Wittenmark (1975), Åström *et al.* (1977), Åström (1983), Landau (1979), and the books edited by Narendra and Monopoli (1980) and Harris and Billings (1981).

Stochastic control problems form one area of process control problems for which the adaptive control methods have proved useful. The original self-tuning regulator of Åström and Wittenmark (1973), which was designed to achieve minimum output variance, has been successfully applied to a number of industrial processes (Åström *et al.* 1977). Adaptive controllers have also been designed for more general stochastic control problems (Åström *et al.* 1977).

It would seem that in practice adaptive controllers could be particularly useful for multivariable plants with many interacting loops, as it is difficult to use manual tuning in such cases. It also requires sophisticated and time-consuming off-line

Received 27 August 1983.

† Part of this work was done during the author's stay at the Norwegian Institute of Technology, Division of Engineering Cybernetics, Trondheim, Norway.

‡ Department of Chemical Engineering, Åbo Akademi, SF-20500 Turku (Åbo), Finland.

experiments to determine a process model, on which the controller design could be based. A number of multivariable adaptive controllers have been considered in the literature (for example Peterka and Åström 1973, Borisson 1975, 1979, Koivo 1980, Keviczky and Kumar 1981). Very few applications have been reported, however. The purpose of the present paper is to review procedures for multivariable adaptive control, which seem promising for practical applications. The discussion is restricted to methods based on stochastic controller design. The primary control criterion is here taken to control the system in such a way that the input and output variances are as small as possible when the disturbances which affect the system are described as stochastic processes. An important class of industrial quality control problems can be formulated in this way (Åström 1970, 1978, Mäkilä, Westerlund and Toivonen 1984). The generality of the stochastic approach considered here is further increased by the fact that many common design problems, such as reference signal tracking and pole-placement design, are conveniently handled in a stochastic framework.

The purpose here is not to attempt to give a complete survey of various techniques for multivariable stochastic adaptive control. Instead the approach taken is to consider the basic procedures and to indicate their applicability. Useful modifications of the basic algorithms are also considered. The structure of the paper is as follows. The basic principles of the stochastic adaptive control problem are considered in § 2. In § 3 adaptive controllers based on linear quadratic gaussian design are considered. In these algorithms a Riccati equation is solved on-line. Some possibilities to reduce the on-line computing requirements are discussed. The problem of selecting the design weights in the quadratic loss function properly to correspond to a well-defined design criterion is also considered, and an algorithm for adapting the weights on-line is described. This is a particularly important problem in the multivariable case as it may be difficult to tune the weights manually in such a way that the variances of all variables are jointly acceptable.

In linear quadratic gaussian design it is straightforward to take the uncertainty of the parameter estimates into account by solving an optimal control problem for a system with uncertain parameters. This results in a cautious adaptive controller (Wittenmark 1975). Peterka and Åström (1973) have given an algorithm of this type using a square-root implementation of the least squares method for parameter estimation. In § 3 the equations are given which are obtained when the commonly used recursive least squares equations with exponential forgetting are applied.

In § 4 adaptive minimum variance control of stably invertible systems is considered. The basic multivariable algorithm was given by Borisson (1975, 1979) for systems with an equal number of inputs and outputs, and with a restrictive condition on the system time delays. For systems where the number of inputs exceeds the number of outputs the minimum variance strategy is in general not uniquely defined. This case can, for example, be handled by determining the strategy which gives minimum output variance with the least input energy (Goodwin, McInnis and Wang 1982). The case when the system has arbitrary but known time delays has been discussed in Goodwin and Dugard (1983). A modification of the approach is considered in § 4.

In § 5 a modification of the adaptive minimum variance controller is considered, which is based on a generalized single-step loss function. The approach was originally proposed in Clarke and Gawthrop (1975, 1979) for single-input single-output systems, and has been generalized to the multivariable case in Koivo (1980) and Keviczky and Kumar (1981). These approaches are restricted to systems with an equal number of inputs and outputs. A convenient generalization of the method to systems having an

arbitrary number of inputs and outputs was described and analysed in Toivonen (1983 c), and this algorithm is considered in § 5.

There are many useful modifications of the methods based on minimum variance control and generalized single-step optimal control. For example, it is straightforward to include reference signal tracking. In this way the methods can be related to model reference adaptive control. It is also possible to select the single-step loss function in such a way that the dynamic behaviour of the closed-loop system is specified. These issues are discussed in §§ 4 and 5.

Simulated examples which illustrate some of the adaptive algorithms are given in § 6.

2. Stochastic adaptive control

2.1. System description

Consider a linear sampled stochastic system described by

$$\left. \begin{aligned} x(t+1) &= Ax(t) + Bu(t) + w(t) \\ y(t) &= Cx(t) + v(t) \end{aligned} \right\} \quad (2.1)$$

where x is the n -dimensional state vector, u is the p -dimensional input, y is the r -dimensional output, and $\{w(t)\}$ and $\{v(t)\}$ are gaussian zero mean white noise sequences. The matrices A , B and C and the noise covariances are assumed constant or slowly time-varying.

From optimal filtering theory it follows that the stochastic process $\{y(t)\}$ has the alternative representation (Åström 1970)

$$\left. \begin{aligned} \hat{x}(t+1) &= A\hat{x}(t) + Bu(t) + Ke(t) \\ y(t) &= C\hat{x}(t) + e(t) \end{aligned} \right\} \quad (2.2)$$

where $\hat{x}(t) = E[x(t) | y(t-1), y(t-2), \dots, u(t-1), u(t-2), \dots]$ is the state estimate, K is the steady-state Kalman filter gain, and $\{e(t)\}$ is a gaussian white noise sequence of prediction errors with zero mean value and covariance R_e .

An input-output description of eqn. (2.2) is given by the vector difference equation

$$A(q^{-1})y(t) = B(q^{-1})u(t-L-1) + C(q^{-1})e(t) \quad (2.3)$$

where q^{-1} is the backwards shift operator ($q^{-1}y(t) = y(t-1)$ etc.), L represents a time delay, and $A(\cdot)$ ($r \times r$), $B(\cdot)$ ($r \times p$) and $C(\cdot)$ ($r \times r$) are matrix polynomials given by

$$\left. \begin{aligned} A(z) &= I + A_1z + \dots + A_mz^m \\ B(z) &= B_0 + B_1z + \dots + B_{m-1}z^{m-1} \\ C(z) &= I + C_1z + \dots + C_mz^m \end{aligned} \right\} \quad (2.4)$$

where m is the observability index of the pair (C, A) . The polynomials $A(\cdot)$, $B(\cdot)$ and $C(\cdot)$ are in general not unique. For obtaining the input-output representation (2.3) from the representation (2.2), see Furuta and Paquet (1975).

2.2 The stochastic controller design problem

The basic control problem considered here is to control the system described by any of the alternative representations (2.1) to (2.3) against the stochastic disturbances in such a way that the variances of the outputs y_i , $i=1, \dots, r$, and the inputs u_j , $j=1, \dots, p$ are as small as possible. This control problem has been studied extensively in the literature (Athans 1971, MacGregor 1973), and it has been shown that it is a useful design method in many industrial control problems (Åström 1970, 1978, Mäkilä, Westerlund and Toivonen 1984).

The optimal control laws are obtained by minimizing quadratic loss functions of the form

$$V = \lim_{N \rightarrow \infty} E \frac{1}{N} \sum_{t=1}^N y(t)^T Q_y y(t) + u(t)^T Q_u u(t) \quad (2.5)$$

where Q_y and Q_u are positive (semi)definite weight matrices to be chosen. This is the well known linear quadratic gaussian control problem. The controllers obtained by minimizing eqn. (2.5) have the property that it is not possible to reduce any of the closed-loop variances of the outputs y_i or the inputs u_j by changing the control law, without simultaneously increasing the variance of at least one other output or input. The loss function (2.5) thus gives a convenient parametrization of the optimal control strategies in terms of the weight matrices Q_y and Q_u . The selection of Q_y and Q_u is made in accordance with the designer's preference ordering to give a satisfactory combination of closed-loop variances.

For a multivariable system it may not be easy to find proper values of the design weights by a simple trial and error procedure, and systematic methods for selecting the weights have therefore been discussed (Athans 1971, Mäkilä, Westerlund and Toivonen 1984).

2.3. Adaptive controllers

An adaptive controller for the control problem described in § 2.2. can be designed by combining an on-line parameter estimator for estimating the parameters of eqn. (2.3) and a part for computing the optimal control law for the identified model (Åström *et al.* 1977). The adaptive controllers considered here can be described by the following general algorithm.

Algorithm 2.1. Adaptive controller

Step 1. Parameter estimation. At time instant t , estimate the parameters of the system (2.3) by an on-line identification method based on the measured outputs and the inputs up to time t .

Step 2. Control law computation. Compute the optimal control strategy for the model obtained in step 1, and determine the corresponding optimal control signal $u^0(t)$.

Step 3. Compute the new input

$$u(t) = u^0(t) + \eta(t) \quad (2.6)$$

where $\eta(t)$ is an input excitation signal, e.g. a PRBS-signal or white noise.

Step 4. Determine the input applied to the process at time t as

$$u(t) = \text{sat}(u(t); \beta, \alpha) \quad (2.7)$$

where $\beta_i \leq \alpha_i$ and

$$\text{sat}_i(z; \beta, \alpha) = \begin{cases} \beta_i, & \text{if } z_i < \beta_i \\ z_i, & \text{if } z_i \in [\beta_i, \alpha_i] \\ \alpha_i, & \text{if } z_i > \alpha_i \end{cases} \quad (2.8)$$

Repeat from step 1 at each sampling instant.

The adaptive algorithms have been classified into *explicit* and *implicit* algorithms depending on how the identification in step 1 is organized. In the explicit methods the parameters of the system (2.3) are estimated explicitly. In the implicit procedures the system is parametrized in such a way that the controller parameters are obtained from the estimated parameters in a trivial way. There are many on-line identification procedures that can be used in the algorithm, and these will be discussed briefly below.

Step 2 of Algorithm 2.1 consists of a computation of the optimal control strategy which minimizes a quadratic loss function, in accordance with the stochastic controller design problem considered in § 2.2. In order to reduce the computational effort a finite-time loss function is often used instead of (2.5). In particular, if $N=L+1$, where L is the system time delay, cf. eqn. (2.3), it is possible to parameterize the model in such a way that the design calculations are trivial. It is also possible to use the parameter uncertainties in the control signal computation in various ways (Wittenmark 1975). Various algorithms which are obtained are discussed in § 3 to 5.

The excitation signal $\eta(t)$ in step 3 is added to the input in order to preserve parameter identifiability in closed-loop operation (Gustavsson, Ljung and Söderström 1977). It can also be used when the parameter estimates are poor in order to obtain better estimates.

Step 4 is introduced in order to avoid unacceptably large inputs, for example reflecting the saturation effects which are always present in practice. It has also been proposed that linear input constraints can be used as design variables offering a method to reduce control signal variations to an appropriate level (Mäkilä 1982).

Parameter estimation

There are many on-line identification methods which can be used in the adaptive algorithms. Extensive surveys of the various techniques can be found in Söderström, Ljung and Gustavsson (1978) and Goodwin and Payne (1977). Only a brief summary is given here.

A large class of on-line parameter estimators can be described by the recursive equations

$$\Theta^T(t+1) = \Theta^T(t) + K(t+1)\epsilon(t+1; \Theta(t))^T \quad (2.9 a)$$

$$K(t+1) = P(t)\psi(t)/(\lambda(t+1) + \psi(t)^T P(t)\psi(t)) \quad (2.9 b)$$

$$P(t+1) = \frac{1}{\lambda(t+1)} \left[P(t) - \frac{P(t)\psi(t)\psi(t)^T P(t)}{\lambda(t+1) + \psi(t)^T P(t)\psi(t)} \right] \quad (2.9 c)$$

$$\epsilon(t+1; \Theta(t)) = y(t+1) - \Theta(t)\varphi(t) \quad (2.9 d)$$

where

$$\Theta(t) = [\hat{A}_1(t), \dots, \hat{B}_0(t), \dots, \hat{C}_1(t), \dots] \quad (2.10)$$

is the estimate of the parameters in (2.4) and

$$\varphi(t) = [-y(t)^T, \dots, u(t-L)^T, \dots, \epsilon(t; \Theta(t-1))^T, \dots]^T \quad (2.11)$$

An alternative to the recursive equation (2.9) is to use numerically stable square root methods (Bierman 1977, Peterka 1975), in which the square root of the matrix $P(t)$ is propagated rather than $P(t)$.

In the estimation procedure, estimates of the prediction errors $e(t)$ are needed for estimating the C -parameters of (2.3). In the recursive algorithm (2.9), the prediction errors are approximated by recursive evaluation of (2.9 *d*), using the current parameter estimate at each step.

Various estimation procedures correspond to different choices of the vector $\psi(t)$. In the recursive extended least square method $\psi(t) = \varphi(t)$ is used. In the recursive maximum likelihood method (Söderström, Ljung and Gustavsson 1978), $\psi(t)$ is obtained by filtering $\varphi(t)$ according to

$$\psi(t) = \hat{C}_t^{-1}(q^{-1})\varphi(t) \quad (2.12)$$

where $\hat{C}_t(q^{-1})$ is formed from the estimated C -parameters at time t .

In the case when the matrix polynomial $C(q^{-1}) = I$, (2.9) reduces to the recursive least squares method. In this case

$$\Theta(t) = [\hat{A}_1(t), \dots, \hat{B}_0(t), \dots] \quad (2.13 a)$$

$$\psi(t) = \varphi(t) = [-y(t)^T, \dots, u(t-L)^T, \dots]^T \quad (2.13 b)$$

This method is particularly convenient since the recursive algorithm generates the optimal estimates which could be obtained by the corresponding off-line method.

In (2.9) $\lambda(\cdot)$ is an exponential weighting factor which makes it possible to track slowly time-varying parameters. When $\lambda = 1$, all observations are weighted equally, while $\lambda < 1$ corresponds to exponential forgetting of past data. In practice it may be useful to apply variable weighting factors, which adapt to changing disturbance levels by retaining a constant amount of information in the estimator (Fortescue, Kershenbaum and Ydstie 1981, Hägglund 1983 a, b).

In some applications it is of interest to consider methods based on stochastic approximation. In these methods the estimates are obtained recursively from

$$\Theta^T(t+1) = \Theta^T(t) + p(t+1)\psi(t)\epsilon(t+1; \Theta(t))^T \quad (2.14)$$

where $p(t)$ is a scalar, e.g. c/t or $1/\text{tr } P(t)^{-1}$. The modification reduces the computational effort considerably, but in return the convergence rate of the estimates is in general slower than with the algorithm (2.9).

Convergence analysis

Simulations can give valuable insight into the behaviour of the adaptive algorithms. It may, however, be difficult to make decisive conclusions about convergence properties from simulations only. For example, it is difficult to determine the exact convergence point of an algorithm, or to decide whether the algorithm actually converges or oscillates around a point (Ljung and Wittenmark 1974). Theoretical analysis can then give valuable insight in the study of adaptive algorithms.

The closed-loop system obtained when controlling (2.3) by an adaptive algorithm as Algorithm 2.1 is non-linear, time-varying and stochastic, which makes it quite difficult to analyse. Some results are available, however, for the case when the system (2.3) is time-invariant. When the recursive parameter estimator gives convergence to the true parameters (which is, e.g. the case for the least squares method), then it follows that the possible convergence points of the algorithm give the required control law, if the system order is not underestimated (Åström *et al.* 1977). This does not imply that the closed-loop under adaptive control is stable, however. There are also commonly used identification methods, such as the extended least squares method, which are known not to converge to the true parameters for all systems (Ljung 1977 a, Söderström, Ljung and Gustavsson 1978). A further complication arises as many adaptive schemes use a reparametrization of (2.3) in order to make the design calculations simple.

Goodwin, Ramadge and Caines (1981) apply martingale theory to study an algorithm based on a modified stochastic approximation identification procedure and a minimum variance design. It was shown that subject to a positive realness condition the inputs and outputs are mean square bounded, and that the algorithm gives convergence to the optimal minimum variance controller.

Ljung (1977 b) has developed a general procedure for analysing recursive stochastic algorithms, which is useful in the study of the adaptive controllers considered here (Åström *et al.* 1977, Ljung 1980). Assume that $\lambda(t) \rightarrow 1$ in the recursive scheme (2.9). Then Ljung (1977 b) has shown that the ordinary differential equation

$$\left. \begin{aligned} \frac{\partial}{\partial \tau} \Theta^T(\tau) &= R(\tau)^{-1} h(\Theta) \\ \frac{\partial}{\partial \tau} R(\tau) &= G(\Theta) - R(\tau) \end{aligned} \right\} \quad (2.15)$$

is associated with the asymptotic behaviour of (2.9) as $t \rightarrow \infty$. Here

$$\left. \begin{aligned} h(\Theta) &= \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{t=1}^N E \Psi(t; \Theta) \epsilon(t; \Theta)^T \\ G(\Theta) &= \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{t=1}^N E \Psi(t; \Theta) \Psi(t; \Theta)^T \end{aligned} \right\} \quad (2.16)$$

where $\Psi(t; \Theta)$ and $\epsilon(t; \Theta)$ are the stochastic processes obtained for the closed-loop system when the parameter matrix Θ , cf. eqns. (2.10) and (2.13 a), which is used in Algorithm 2.1, is constant.

Equation (2.15) can be used to study the asymptotic behaviour of the adaptive algorithm. A restriction of this approach is that stability of the closed-loop system must be assumed for the limits (2.16) to exist. Equation (2.15) can be analysed only in simple cases (Ljung 1977 a, 1980). A multivariable adaptive algorithm is analysed using (2.15) in Toivonen (1983 c). When the analysis of (2.15) is infeasible, the results can be used heuristically by solving the ordinary differential equation (2.15) numerically (Åström *et al.* 1977, Ljung 1980). In this way valuable information of the asymptotic behaviour of the adaptive controller can be gained (Åström and Wittenmark 1974, Ljung and Wittenmark 1974, Toivonen 1983 a).

3. Adaptive controllers based on linear quadratic gaussian design

In this section some specific adaptive algorithms based on linear quadratic gaussian design are considered. In this approach the parameters of the system (2.3) are estimated recursively, and the control law is then determined by solving a Riccati equation. Algorithms of this form have been considered for example by Peterka and Åström (1973), Lam (1980), Åström and Zhao-Ying (1982).

Consider the model

$$\hat{A}(q^{-1})y(t) = \hat{B}(q^{-1})u(t-1) + \hat{C}(q^{-1})\epsilon(t) \quad (3.1)$$

which corresponds to (2.3) with $L=0$. This implies no restriction since the leading matrix coefficients of the B -polynomial can always be set equal to zero. A possible, non-minimal state-space representation of (3.1) is

$$\hat{z}(t+1) = \hat{A}\hat{z}(t) + \hat{B}u(t) + K\epsilon(t+1) \quad (3.2 a)$$

$$y(t) = [-I \ 0 \ \dots]z(t) \quad (3.2 b)$$

with

$$\hat{z}(t) = [-y(t)^T, -y(t-1)^T, \dots, u(t-1)^T, \dots, \epsilon(t)^T, \dots]^T \quad (3.3)$$

and

$$\hat{A} = \begin{bmatrix} F \\ M \end{bmatrix}, \quad \hat{B} = \begin{bmatrix} \hat{B}_0 \\ N \end{bmatrix} \quad (3.4)$$

where

$$F = [\hat{A}_1, \hat{A}_2, \dots, \hat{B}_1, \dots, \hat{C}_1, \dots] \quad (3.5)$$

and the matrices M , N and K consist of zeros and ones in positions which follow from the structure (3.2), (3.3).

The basic adaptive algorithm based on linear quadratic gaussian control takes the following form.

Algorithm 3.1. Linear quadratic gaussian control

Step 1. Estimate the parameters of (3.1) by an on-line parameter estimation method to obtain the estimates \hat{A}_i , \hat{B}_i , \hat{C}_i .

Step 2. Determine the control law which minimizes the loss function (2.5) for the model (3.1), i.e.

$$u^0(t) = -L_t \hat{z}(t) \quad (3.6)$$

where

$$L_t = (\hat{B}^T S_t \hat{B} + Q_2)^{-1} \hat{B}^T S_t \hat{A} \quad (3.7 a)$$

$$S_t = \hat{A}^T S_t \hat{A} - \hat{A}^T S_t \hat{B} (\hat{B}^T S_t \hat{B} + Q_2)^{-1} \hat{B}^T S_t \hat{A} + Q_x \quad (3.7 b)$$

$$Q_x = \begin{bmatrix} Q_y & 0 \\ 0 & 0 \end{bmatrix} \quad (3.7 c)$$

Steps 3 and 4 as in Algorithm 2.1.

Repeat from step 1 at each sampling instant.

There are many possible modifications of Algorithm 3.1. One possibility is to use a state space representation which is of lower order than (3.2), (3.3). In that case the

state vector will depend explicitly on the estimated parameters \hat{A}_i , \hat{B}_i , \hat{C}_i . Some other modification will be discussed below.

The Riccati equation (3.7) is conveniently solved iteratively, starting the iteration using the Riccati matrix S_{t-1} from the previous sampling period. In order to reduce the computational effort required, it has been proposed that only one iteration of (3.7) is performed at each sampling instant (Bartolini *et al.* 1982), i.e.

$$S_t = \hat{A}^T S_{t-1} \hat{A} - \hat{A}^T S_{t-1} \hat{B} (\hat{B}^T S_{t-1} \hat{B} + Q_2)^{-1} \hat{B}^T S_{t-1} \hat{A} + Q_1 \quad (3.8)$$

This procedure can be expected to work well only if the estimates do not vary too much. When the estimates change from step to step, (3.8) may not be able to keep track of these changes, and after a while the control law (3.6) may give poor control performance, even in cases when the estimated parameters are quite acceptable all the time. This phenomenon has been observed in simulations. In these cases the number of iterations can be increased, and it is straightforward to include a test which controls the number of iterations, for example by requiring that $\|S^{(k+1)} - S^{(k)}\| / \|S^{(k)}\| \leq \delta$, where $S^{(k)}$ denotes the Riccati matrix at the k th iteration and δ is a small positive scalar.

Use of least squares models

The least squares method is a convenient estimation procedure, and it is therefore often used in practice. It is worth studying to what extent it is possible to base the adaptive algorithm on a least squares model,

$$\hat{A}(q^{-1})y(t) = \hat{B}(q^{-1})u(t-1) + \epsilon(t) \quad (3.9)$$

even in cases, when the system is appropriately described by (2.3) with $C(q^{-1}) \neq I$. Such an approach can be expected to work well, since (3.9) can be considered as an approximation to (2.3) obtained by approximating the inverse $C(q^{-1})^{-1}$ by a matrix polynomial of finite order. In general, high orders for $\hat{A}(\cdot)$ and $\hat{B}(\cdot)$ in (3.9) may have to be used in order to obtain a good approximation, however.

For the adaptive algorithms operating in closed-loop there are, however, some experimental results showing that use of the least squares model (3.9) may give quite good performance even with low order models (Åström and Wittenmark 1974, Toivonen 1981). Some insight can be obtained by the following heuristic reasoning. When the system (3.1) with the state space representation (3.2) is controlled by a time-invariant control law of the form (3.6), the state vector $\hat{z}(t)$ can be constructed in terms of $u:s$ and $y:s$ only using (3.2) and (3.6), eliminating the residuals in (3.2) by the relation

$$\epsilon(t+1) = y(t+1) - F\hat{z}(t) - \hat{B}_0 u(t)$$

Thus, when using a constant feedback law, as eqn. (3.6), the output of the closed-loop system is correctly predicted by a least squares model of the form (3.9), which holds only for the given closed-loop conditions. The order of the least squares model is determined by the observability index of the pair $(L, \hat{A} - KF)$. For these reasons an adaptive controller based on a least squares model may converge to the required strategy when the design is based on a single step loss function, see § 5. Numerical examples have shown that good results can also be obtained in cases when the design is based on a multistep loss function. Several numerical examples are presented in Åström and Wittenmark (1974) for minimum variance control of non-minimum phase plants and in Toivonen (1981) for the variance constrained optimal control

problem (3.10) (see below). It was found that an adaptive algorithm based on (3.9), when applied to a system described by (2.3), in many cases converged to a control law giving only a few percent larger values for the loss function than the optimal strategy. In many cases the non-optimality of the algorithm could only be discovered by studying the associated differential equation (2.15). An example is shown in § 6. In view of these results the use of the least squares model (3.9) is clearly an interesting possibility.

Selection of the weight matrices

In the loss function (2.5) the weight matrices Q_y and Q_u should be properly chosen to achieve the desired closed-loop performance. The proper values for the weights depend on the system dynamics, and in a truly adaptive procedure the weights should therefore also be adjusted on-line. For single-input single-output systems the adjustment can be done manually (Clarke and Gawthrop 1981, Åström 1983), but in the multivariable case with several inputs and outputs this may not always be easy. It is therefore also of interest to study methods for tuning the design weights adaptively. In Toivonen (1983 a, b) a procedure is considered where the control problem is formulated as an optimal stochastic control problem with explicit variance restrictions. In this formulation, the solution is found to the constrained problem

$$\left. \begin{aligned} \text{Minimize } V &= \lim_{N \rightarrow \infty} E \frac{1}{N} \sum_{t=1}^N y(t)^T Q_y y(t) \\ \text{subject to } \lim_{N \rightarrow \infty} E \frac{1}{N} \sum_{t=1}^N u(t)^T R_i u(t) &\leq c_i^2, \quad i=1, \dots, q \end{aligned} \right\} \quad (3.10)$$

For unstable plants the constraints c_i^2 should be chosen so that the problem has a solution. The formulation (3.10) in stochastic controller design has been applied successfully to industrial process control (Westerlund 1980). The off-line design procedure has been discussed extensively in Mäkilä, Westerlund and Toivonen (1984). The design of an adaptive controller for the variance constrained optimal control problem (3.10) results in a procedure for on-line adaptation of the weight matrices in the quadratic loss function. This is seen by considering the Lagrangian function of the constrained minimization problem (3.10),

$$V_L = \lim_{N \rightarrow \infty} E \frac{1}{N} \sum_{t=1}^N y(t)^T Q_y y(t) + u(t)^T Q_u^* u(t) \quad (3.11)$$

where

$$Q_u^* = \sum_{i=1}^q \lambda_i R_i \quad (3.12)$$

and $\lambda_i (\geq 0)$ are Lagrange multipliers. The necessary optimality conditions are

$$\lambda_i \left(\lim_{N \rightarrow \infty} E \frac{1}{N} \sum_{t=1}^N u(t)^T R_i u(t) - c_i^2 \right) = 0, \quad i=1, \dots, q \quad (3.13)$$

A two-phase adaptive algorithm for the variance constrained optimal control problem (3.10) can now be obtained as follows (Toivonen 1983 a, b)

1. Apply an adaptive controller based on linear quadratic gaussian design for minimizing the loss function V_L , eqn. (3.11), where the weight matrix Q_u^* is given by (3.12) (Algorithm 3.1).

2. At each sampling instant, adjust the Lagrange multipliers λ_i by applying the Robbins–Monro stochastic approximation scheme to the eqns. (3.13):

$$\lambda_i(t+1) = \lambda_i(t) + \mu_t \lambda_i(t) (u(t)^T R_i u(t) - c_i^2) / c_i^2, \quad i=1, \dots, q \quad (3.14)$$

where μ_t is a small positive scalar.

Cautious control

Algorithm 3.1 is a certainty equivalence adaptive controller (Wittenmark 1975), i.e. the control signal is determined in step 2 as if the estimated parameters were the correct ones. When the parameter estimates are poor the approach may not work well because too much confidence is put on the current estimates. It could then be useful to take the uncertainty of the parameter estimates into account in some way. One way to achieve this is by cautious control (Wittenmark 1975). In this approach the loss function

$$V_N = E \left[\sum_{i=t}^{t+N} y(i)^T Q_y y(i) + u(i)^T Q_u u(i) \mid t \right] \quad (3.15)$$

is minimized at each step under the assumption that future measurements are not used for improving the current parameter estimates. The first control $u^0(t)$ in the optimal control sequence is then applied to the system. This procedure is repeated at each sampling time. In the minimization of (3.15) the uncertainty of the current estimate $\Theta(t)$ is taken into account. An optimal control problem for a system with uncertain parameters (Åström 1977, Panossian and Leondes 1983) is thus solved at each sampling time. The variances of the estimates $\Theta(t)$ are obtained approximately from the recursive estimator (2.9). This approach is not optimal, since the fact that the control will influence future estimates and their accuracy is not taken into account. In this approach the effect of parameter uncertainties is simply to make the controller more cautious, since less confidence is put on the parameter estimates which are available at each sampling instant.

Peterka and Åström (1973) give a cautious controller, in which the least squares procedure is used for parameter estimation. A square-root implementation of the least square method was used. The corresponding algorithm obtained when using the recursive least squares method of eqns. (2.9), (2.13) has been considered in Toivonen (1981). The procedure is a straightforward generalization of Algorithm 3.1.

In the method of least squares, the matrix $P(t)$ of eqn. (2.9) is proportional to the covariance of the estimates, and it can be shown (Peterka 1975, Toivonen 1981) that for any known matrix S

$$E[(\Theta(t) - \Theta_0)^T S (\Theta(t) - \Theta_0)] = P(t) \operatorname{tr} S R_e \quad (3.16)$$

where Θ_0 is the true parameter matrix, and R_e is the covariance of the white noise e , eqn. (2.3):

$$R_e = E e(t) e(t)^T \quad (3.17)$$

In practice R_e is not known, and it can be estimated. The maximum likelihood estimate of R_e is obtained recursively from (Peterka 1975)

$$\gamma(t+1) \hat{R}_e(t+1) = \lambda(t+1) [\gamma(t) \hat{R}_e(t) + e(t+1; \Theta(t)) e(t+1; \Theta(t))^T / (\lambda(t+1) + \varphi(t)^T P(t) \varphi(t))] \quad (3.18 a)$$

$$\gamma(t+1) = \lambda(t)\gamma(t) + 1 \quad (3.18 b)$$

and $\varphi(t)$ is given by (2.13 b). These relations can now be used to obtain the following algorithm for cautious control.

Algorithm 3.2. Cautious controller

Step 1. Estimate the parameters of (3.9) by the recursive least squares method using (2.9), (3.18) to obtain $\Theta(t)$, $P(t)$, and $\hat{R}_e(t)$. It is convenient to reorder the φ -vector and to use

$$\left. \begin{aligned} \psi(t) = \varphi(t) &= [u(t)^T, -y(t)^T, -y(t-1)^T, \dots, u(t-1)^T, \dots]^T \\ \Theta(t) &= [\hat{B}_0, \hat{A}_1, \hat{A}_2, \dots, \hat{B}_1, \dots]^T \end{aligned} \right\} \quad (3.19)$$

instead of (2.13).

Step 2. Computation of cautious control law. Use the estimates and their uncertainties for determining the control sequence which minimizes (3.15) under the assumption that future measurements are not used for improving the current estimates. This is an optimal control problem for a system with uncertain parameters, which can be solved by dynamic programming (Åström 1977). Introducing the result (3.16) and the state space representation (3.2) gives the solution (Toivonen 1981)

$$u^0(t) = -L(t)z(t) \quad (3.20)$$

where

$$\left. \begin{aligned} L(k) &= [\hat{B}^T S(k+1)\hat{B} + P_{BB}(t) \operatorname{tr}(S_{11}(k+1)\hat{R}_e(t)) + Q_2]^{-1} [\hat{B}^T S(k+1)\hat{A} \\ &\quad + P_{BF}(t) \operatorname{tr}(S_{11}(k+1)\hat{R}_e(t))] \\ S(k) &= \hat{A}^T S(k+1)\hat{A} + P_{FF}(t) \operatorname{tr}(S_{11}(k+1)\hat{R}_e(t)) \\ &\quad - [\hat{A}^T S(k+1)\hat{B} + P_{BF}^T(t) \operatorname{tr}(S_{11}(k+1)\hat{R}_e(t))] L(k) + Q_x \\ S(t+N) &= Q_x, \quad k = t+N-1, \dots, t \end{aligned} \right\} \quad (3.21)$$

where \hat{A} and \hat{B} are given by (3.4) and Q_x by (3.7 c). The Riccati matrix S is partitioned as

$$S = \begin{bmatrix} S_{11} & S_{12} \\ \underbrace{S_{12}^T}_r & S_{22} \end{bmatrix}$$

and the matrix $P(t)$ of eqn. (2.9) as (cf. (3.19))

$$P(t) = \begin{bmatrix} P_{BB}(t) & P_{BF}(t) \\ \underbrace{P_{BF}^T(t)}_p & P_{FF}(t) \end{bmatrix}$$

Steps 3 and 4 as in Algorithm 2.1.

Repeat from step 1 at each sampling instant.

A well known phenomenon in cautious control is the so-called turn-off (Wittenmark 1975). This means that when the estimates are poor, the covariances will cause

the control action to be cautious, and the control signal may become very small. As the control signal fails to excite the system the estimates can become still worse in the future. In this way the control may be turned off for some period of time, until the noise excites the system in such a way that better estimates are obtained. This phenomenon can be avoided in Algorithm 3.2 by applying the input excitation in step 3 when the estimates are poor.

4. Adaptive minimum variance control

A particularly simple adaptive controller can be designed for the case when the criterion is to achieve minimum variance of the outputs, i.e. when $Q_u = 0$ in the loss function (2.5). In this case there is a class of systems for which it is possible to parametrize the model in such a way that the design calculations are trivial. It is also straightforward to include reference signal tracking and feedforward from measured signals in this approach. The basic self-tuning minimum variance regulator for single-input single-output systems was given in Åström and Wittenmark (1973) and Wittenmark (1973). A multivariable generalization has been given by Borisson (1975, 1979). Here the method is presented with some generalizations.

Consider a stochastic system described by

$$A(q^{-1})y(t) = B(q^{-1})u(t-L-1) + D(q^{-1})n(t-L-1) + C(q^{-1})e(t) \quad (4.1)$$

where $n(t)$ is a known signal. It is assumed that the number r of outputs does not exceed the number p of inputs, that $\text{rank } B(0) = r$, and that the system is stably invertible. In the case when $r = p$ the last condition implies that $\det B(z)$ should have all zeros outside the unit disc (minimum phase assumption).

Now consider the loss function

$$V_{MV} = \lim_{N \rightarrow \infty} E \frac{1}{N} \sum_{t=1}^N [y(t+L+1) - y_r(t+L+1)]^T Q_y [y(t+L+1) - y_r(t+L+1)] \quad (4.2)$$

where $y_r(\cdot)$ is a reference signal to be tracked, and assumed known at least $L+1$ steps ahead. The control strategy which minimizes (4.2) for the system (4.1) is obtained from the $(L+1)$ -step predictive model (Borisson 1975, 1979)

$$y(t+L+1) = \tilde{C}(q^{-1})^{-1} [\tilde{G}(q^{-1})y(t) + \tilde{F}(q^{-1})B(q^{-1})u(t) + \tilde{F}(q^{-1})D(q^{-1})n(t) + F(q^{-1})e(t+L+1)] \quad (4.3)$$

where the matrix polynomials $\tilde{C}(\cdot)$, $\tilde{G}(\cdot)$, $\tilde{F}(\cdot)$ and $F(\cdot)$ are defined by the identities

$$\left. \begin{aligned} C(z) &= A(z)F(z) + z^{L+1}G(z), \quad \deg F(z) = L \\ \tilde{F}(z)G(z) &= \tilde{G}(z)F(z), \quad \det \tilde{F}(z) = \det F(z), \quad \tilde{F}(0) = F(0) = I \\ \tilde{C}(z) &= \tilde{F}(z)A(z) + z^{L+1}\tilde{G}(z), \quad \det \tilde{C}(z) = \det C(z) \end{aligned} \right\} \quad (4.5)$$

Since the last term in (4.3) is uncorrelated with $y(t)$, $y(t-1)$, ..., $u(t-1)$, ..., $n(t)$, $n(t-1)$, ..., the strategy which minimizes (4.2) is obtained by setting the first expression in (4.3) equal to $y_r(t+L+1)$, i.e.

$$\tilde{F}(q^{-1})B(q^{-1})u(t) + \tilde{G}(q^{-1})y(t) + \tilde{F}(q^{-1})D(q^{-1})n(t) - \tilde{C}(q^{-1})y_r(t+L+1) = 0 \quad (4.6)$$

The output of the closed-loop system is then

$$y(t+L+1) = y_r(t+L+1) + F(q^{-1})e(t+L+1) \quad (4.7)$$

The minimum variance strategy is seen to be independent of the weight Q_y . The strategy (4.6) can be written compactly as

$$B_0 u(t) + \Theta \varphi(t) = y_r(t+L+1) \quad (4.8)$$

where $B_0 = B(0)$,

$$\varphi(t) = [y(t)^T, \dots, u(t-1)^T, \dots, n(t)^T, \dots, y_r(t+L)^T, \dots]^T$$

and Θ consists of the matrix coefficients of the polynomials $\tilde{G}(z)$, $\tilde{F}(z)B(z) - B(0)$, $\tilde{F}(z)D(z)$ and $I - \tilde{C}(z)$.

In the case when the number of inputs and outputs are equal, $r=p$, (4.6), or (4.8), defines the minimum variance strategy uniquely:

$$u_{MV}(t) = -B_0^{-1}[\Theta \varphi(t) - y_r(t+L+1)] \quad (4.9)$$

In the case when the number of inputs exceeds the number of outputs, $p > r$, the minimum variance strategy is not unique. The non-uniqueness can be exploited for example by selecting the control signal which gives minimum output variance with least control energy at each step (Goodwin, McInnis and Wang 1982), i.e. by solving

$$\min u(t)^T P u(t) \quad (P > 0) \quad \text{subject to (4.8)} \quad (4.10)$$

which gives the control law

$$u_{MV}(t) = -P^{-1} B_0^T (B_0 P^{-1} B_0^T)^{-1} [\Theta \varphi(t) - y_r(t+L+1)] \quad (4.11)$$

Another possibility is to determine the input which gives minimum output variance with the smallest changes of the control signal, by solving

$$\min [u(t) - u(t-1)]^T P [u(t) - u(t-1)] \quad (P > 0) \quad \text{subject to (4.8)} \quad (4.12)$$

This gives the control law

$$u_{MV}(t) = u(t-1) - P^{-1} B_0^T (B_0 P^{-1} B_0^T)^{-1} [\Theta \varphi(t) - y_r(t+L+1) + B_0 u(t-1)] \quad (4.13)$$

The structure of the minimum variance strategy (4.6) suggests that an adaptive minimum variance controller can be designed by estimating the parameters of a prediction model corresponding to (4.3), and setting the predicted output equal to the reference value $y_r(t+L+1)$ (Åström and Wittenmark 1973, Borisson 1975, 1979). The following algorithm is obtained.

Algorithm 4.1. Minimum variance control ($\dim u \geq \dim y$)

Step 1. Estimate the parameters of the predictive model

$$y(t+L+1) = \mathcal{A}(q^{-1})y(t) + \mathcal{B}(q^{-1})u(t) + \mathcal{C}(q^{-1})y_r(t+L) + \mathcal{D}(q^{-1})n(t) + \epsilon(t+L+1) \quad (4.14)$$

by the method of least squares.

Step 2. Determine the control signal $u_{MV}(t)$ from the minimum variance strategy

$$\mathcal{A}(q^{-1})y(t) + \mathcal{B}(q^{-1})u(t) + \mathcal{C}(q^{-1})y_r(t+L) + \mathcal{D}(q^{-1})n(t) = y_r(t+L+1) \quad (4.15)$$

This leads to (4.9), (4.11) or (4.13) depending on the situation.

Steps 3 and 4 as in Algorithm 2.1.

Repeat from step 1 at each sampling instant.

Note that the reference signal $y_r(\cdot)$ can be determined as the output from a reference model

$$A_r(q^{-1})y_r(t+L+1) = B_r(q^{-1})u_r(t)$$

This relates Algorithm 4.1 with model reference adaptive control (Landau 1979). The procedure is still limited to stably invertible systems.

Goodwin, Ramadge and Caines (1981) have applied martingale theory to establish global convergence of a modified version of Algorithm 4.1, in which a modified stochastic approximation procedure is used for parameter estimation. It is shown that, provided a certain transfer function associated with the system (4.1) is strictly positive real, the inputs and the outputs are mean square bounded, and the algorithm converges to the minimum variance controller. A global convergence proof of Algorithm 4.1 with least squares estimation is not available. However, if it is assumed that the signals remain bounded, the associated ordinary differential equation (2.15) can be used to study the asymptotic behaviour of the algorithm. The results of Ljung (1978) applied to Algorithm 4.1 show that convergence to the minimum variance strategy (4.6) is obtained if the transfer function

$$\tilde{C}(q^{-1})^{-1} - \frac{1}{2}I$$

is strictly positive real, i.e. if the matrix

$$\text{Re} [\tilde{C}(\exp(i\omega))^{-1} + \tilde{C}(\exp(i\omega))^{-T} - I]$$

is strictly positive definite for all ω .

Systems with arbitrary time delays

Algorithm 4.1 is restricted to the case when the matrix $B_0 = B(0)$ has full rank. This implies a restriction on the time delays of the system. It is, however, straightforward to handle systems with arbitrary but known time delays. Goodwin and Dugard (1983) give a procedure which involves the system interactor matrix. Here a modification of the method is given, which is more straightforward in the case when there are different time delays at the inputs.

Let the system be given by the transfer function representation

$$y(t) = q^{-L}H(q^{-1})u(t) + N(q^{-1})e(t) \quad (4.16)$$

where

$$H(q^{-1}) = \begin{bmatrix} q^{-L_{1j}-1} & \frac{b_{1j}(q^{-1})}{a_{1j}(q^{-1})} \end{bmatrix}, \quad a_{ij}(0) = 1, \quad b_{ij}(0) \neq 0, \quad \min L_{ij} = 0 \quad (4.17)$$

Then there are two diagonal matrices

$$D_1(q) = \text{diag}(q^{k_1}, \dots, q^{k_r}), \quad \min k_i = 0 \quad (4.18 a)$$

$$D_2(q) = \text{diag}(q^{l_1}, \dots, q^{l_p}), \quad \min l_i = 0 \quad (4.18 b)$$

such that

$$\lim_{q^{-1} \rightarrow 0} qD_1(q)H(q^{-1})D_2(q) = K \quad (4.19)$$

is finite and has no zero rows or columns. It is assumed that the matrix K formed in this way has full rank. Introduce

$$\tilde{H}(q^{-1}) = D_1(q)H(q^{-1})D_2(q) \quad (4.20)$$

and the signals

$$\tilde{y}(t) = D_1(q)y(t) \quad (4.21 a)$$

$$\tilde{u}(t) = q^{l_{\max}} D_2(q^{-1})u(t), \quad l_{\max} = \max l_i \quad (4.21 b)$$

From (4.16) is now obtained

$$\tilde{y}(t) = q^{-\tilde{L}} \tilde{H}(q^{-1})\tilde{u}(t) + D_1(q^{-1})N(q^{-1})e(t) \quad (4.22)$$

where $\tilde{L} = L + l_{\max}$. Goodwin and Dugard (1983) show that using (4.22) a predictive model analogous to (4.3) can be derived for the signal $\tilde{y}(t)$, giving

$$\tilde{y}(t + \tilde{L} + 1) = \tilde{C}(q^{-1})^{-1} [\mathcal{A}^0(q^{-1})y(t) + \mathcal{B}^0(q^{-1})\tilde{u}(t)] + F(q^{-1})e(t + \tilde{L} + 1) \quad (4.23)$$

where $\mathcal{B}^0(0) = K$. A minimum variance strategy corresponding to (4.6) is now obtained by setting the predicted $\tilde{y}(t + \tilde{L} + 1)$ equal to $D_1(q)y_r(t + \tilde{L} + 1)$:

$$\mathcal{B}^0(q^{-1})\tilde{u}(t) + \mathcal{A}^0(q^{-1})y(t) - \tilde{C}(q^{-1})D_1(q)y_r(t + \tilde{L} + 1) \quad (4.24)$$

It is clear from (4.23) and (4.24) that Algorithm 4.1 can be generalized to systems with arbitrary but known time delays by replacing the predictive model (4.14) of the algorithm by

$$\begin{aligned} \tilde{y}(t + \tilde{L} + 1) = & \mathcal{A}(q^{-1})y(t) + \mathcal{B}(q^{-1})\tilde{u}(t) + \mathcal{C}(q^{-1})(D_1(q)y_r(t + \tilde{L})) + \mathcal{D}(q^{-1})n(t) \\ & + \epsilon(t + \tilde{L} + 1) \end{aligned} \quad (4.25)$$

It is observed that when there are different delays in the outputs only, i.e. when $D_2(q) = I$, (4.23) corresponds to using different prediction times for the various outputs, and (4.24) gives the optimal minimum variance strategy which minimizes (4.2). Mäkilä (1982) illustrates the situation in this case by a numerical example in which a self-tuning algorithm based on (4.25) is used. When there are different delays at the inputs, (4.21 b) corresponds to introducing additional time delays, since for the i th input $\tilde{u}_i(t) = u_i(t + l_{\max} - l_i)$, i.e., the signal $\tilde{u}_i(t)$ is determined at time t and applied to the process at the later time $t + l_{\max} - l_i$. This case has been considered in Tantt and Koivo (1983). In this case the approach is not optimal, since the fact that some of the inputs affect the system outputs with a delay less than \tilde{L} is not exploited. The optimal strategy is, however, much more complex, and requires the solution of a steady-state Riccati equation. An alternative procedure for designing adaptive controllers for systems with arbitrary time delays is the MUSMAR-algorithm (Menga and Mosca 1980), which is described in § 5. In that approach it is not required that the time delays are known.

5. Adaptive algorithms based on single-step optimal control

Minimum variance control has the drawback that the strategy may generate extensively large control signals (MacGregor 1973). Linear quadratic gaussian design can then be used, but this results in more complex algorithms. A useful modification, which preserves the computational simplicity of the approach based on minimum variance control, is then obtained by basing the design on the generalized single-step loss function (Clarke and Gawthrop 1975, 1979, Koivo 1980, Keviczky and Kumar 1981)

$$\begin{aligned} V_1 = E[(y(t + L + 1) - y_r(t + L + 1))^T Q_y (y(t + L + 1) - y_r(t + L + 1)) \\ + (P(q^{-1})u(t))^T Q_u (P(q^{-1})u(t)) | y(t), \dots, u(t), \dots, n(t), \dots] \end{aligned} \quad (5.1)$$

where $P(q^{-1})$ is a (matrix) polynomial, which can be chosen to affect the closed-loop response.

Now consider a system described by (4.1). The predictive equation (4.3) gives the control law which minimizes the single-step loss function (5.1) at each step as (cf. Koivo 1980, Keviczky and Kumar 1981, Toivonen 1983 c)

$$B_0^T Q_y [\mathcal{A}^0(q^{-1})y(t) + \mathcal{B}^0(q^{-1})u(t) + \mathcal{D}^0(q^{-1})n(t)] \\ + c(q^{-1})P(0)^T Q_u P(q^{-1})u(t) - B_0^T Q_y c(q^{-1})y_r(t+L+1) = 0 \quad (5.2)$$

where

$$\left. \begin{aligned} c(q^{-1}) &= \det \tilde{C}(q^{-1}) = \det C(q^{-1}) \\ \mathcal{A}^0(q^{-1}) &= \text{adj} [\tilde{C}(q^{-1})] \tilde{G}(q^{-1}) \\ \mathcal{B}^0(q^{-1}) &= \text{adj} [\tilde{C}(q^{-1})] \tilde{F}(q^{-1}) B(q^{-1}), \quad \mathcal{B}^0(0) = B(0) = B_0 \\ \mathcal{D}^0(q^{-1}) &= \text{adj} [\tilde{C}(q^{-1})] D(q^{-1}) \end{aligned} \right\} \quad (5.3)$$

In (5.2) it is meaningful to assume that B_0 has full rank, whereas the number of inputs and outputs are arbitrary. Systems with arbitrary time delays can be treated as was described in § 4, cf. eqn. (4.21).

The strategy (5.2) is not optimal for the problem considered in § 2.2, which involves the steady state input and output variances. For single-input single-output systems it has been shown, however, that in many cases good steady-state performance can be obtained by basing the design on a single-step loss function (Modén and Söderström 1982). The strategy (5.2) then gives a simple suboptimal method for reducing the input variances.

It is straightforward to construct adaptive controllers based on single-step optimal control. For the single-input single-output case an adaptive algorithm has been given by Clarke and Gawthrop (1975, 1979), and generalizations to multivariable systems have been considered by Koivo (1980) and Keviczky and Kumar (1981). In these papers the observation is used that when $r=p$ the strategy (5.2) can be considered as a minimum variance strategy for the signal

$$\phi(t+L+1) = y(t+L+1) + B_0(B_0^T Q_y B_0)^{-1} P(0)^T Q_u P(q^{-1})u(t) \quad (5.3)$$

This is readily seen from eqns. (4.3), (4.6) and (5.2). Algorithm 4.1 can then be applied to achieve minimum variance control of the auxiliary signal ϕ , using in step 1 the predictive model.

$$\phi(t+L+1) = \mathcal{A}(q^{-1})y(t) + \mathcal{B}(q^{-1})u(t) + \mathcal{C}(q^{-1})y_r(t+L) \\ + \mathcal{D}(q^{-1})n(t) + \epsilon(t+L+1) \quad (5.4)$$

and in step 2 the control law (4.15). Apart from being restricted to systems with an equal number of inputs and output, this approach has the drawback that there is no convenient way of estimating the matrix B_0 in the algorithm, as this parameter appears non-linearly in the equations, cf. eqn. (5.3). Therefore B_0 is usually assumed known (Koivo 1980, Keviczky and Kumar 1981).

A more direct approach, which does not have the above restrictions, is discussed in Toivonen (1983 c). The following algorithm is obtained.

Algorithm 5.1. Single-step optimal control ($\dim y$, $\dim u$ arbitrary)

Step 1. Estimate the parameters of the predictive model (4.14) by the recursive least squares method.

Step 2. Determine the control signal to minimize the loss function (5.1) for the model obtained in step 1, i.e. determine $u(t)$ from

$$\mathcal{B}_0^T Q_y [\mathcal{A}(q^{-1})y(t) + \mathcal{B}(q^{-1})u(t) + \mathcal{C}(q^{-1})y_r(t+L) + \mathcal{D}(q^{-1})n(t)] + P(0)^T Q_u P(q^{-1})u(t) = \mathcal{B}_0^T Q_y y_r(t+L+1) \quad (5.5)$$

where $\mathcal{B}_0 = \mathcal{B}(0)$.

Steps 3 and 4 as in Algorithm 2.1.

Repeat from step 1 at each sampling instant.

In Toivonen (1983 c) the convergence of Algorithm 5.1 was studied using the approach based on the associated differential equation (2.15). It was shown for the case when B_0 is known that convergence to the required strategy (5.2) is obtained if the transfer function

$$\frac{1}{c(q^{-1})} - \frac{1}{2}$$

is strictly positive real.

A more general single-step cost function is given by (Keviczky and Kumar 1981)

$$V_1 = E[(W_y(q^{-1})y(t+L+1) - W_r(q^{-1})y_r(t+L+1))^T \times Q_y(W_y(q^{-1})y(t+L+1) - W_r(q^{-1})y_r(t+L+1)) + (W_u(q^{-1})u(t))^T W_u(q^{-1})u(t) | y(t), \dots, u(t), \dots, n(t), \dots] \quad (5.6)$$

where $W_y(\cdot)$, $W_r(\cdot)$ and $W_u(\cdot)$ are matrix fraction descriptions. This case can be handled as follows. Let $W_u(\cdot)$ be expressed by the right prime matrix description

$$W_u(q^{-1}) = P(q^{-1})W_{ud}^{-1}(q^{-1})$$

and introduce the filtered signals

$$\left. \begin{aligned} y_F(t) &= W_y(q^{-1})y(t) \\ y_{r,F}(t) &= W_r(q^{-1})y_r(t) \\ u_F(t) &= W_{ud}^{-1}(q^{-1})u(t) \end{aligned} \right\} \quad (5.7)$$

The loss function (5.6) then takes the form (5.1) in terms of the filtered signals $y_F(t)$ and $u_F(t)$. From (4.1) a predictive model of the form (4.3) can be derived for the filtered signals. See Keviczky and Kumar (1981) for details. Algorithm 5.1 can then be applied to the filtered signals (5.7). In (5.6) the design weights can be used to affect the closed-loop response. For example, for stably invertible systems with $\dim y \leq \dim u$, the choice $W_u(q^{-1}) = 0$ gives the closed-loop response (cf. eqns. (4.7) and (5.7))

$$W_y(q^{-1})y(t) = W_r(q^{-1})y_r(t) + \epsilon(t) \quad (5.8)$$

The approach can thus be used as an adaptive method to obtain a desired closed-loop behaviour.

The MUSMAR-algorithm

The design based on the single-step loss function (5.1) has the advantage that it is possible to obtain a simple implicit adaptive algorithm, as Algorithm 5.1. The procedure is, however, not optimal for the stochastic control problem described in § 2.2, and it does not work well for systems with unknown time delays or other non-minimum phase properties. In general, better control performance can be obtained by minimizing a multi-step loss function. The procedures described in § 3 can be used in this case, but the computational advantages of Algorithm 5.1 are then lost. A method to overcome the restrictions of Algorithm 5.1 is proposed by Menga and Mosca (1980). In their approach (MUSMAR, *Multistep Multivariable Adaptive Regulator*), an implicit adaptive algorithm of the form of Algorithm 5.1 is designed for minimizing a multi-step loss function of the form (3.15) by using several predictive models simultaneously.

If a constant feedback law is applied to the system (2.3) at $t, t+1, \dots$, the future outputs $y(t+1), \dots, y(t+N)$ and the inputs $u(t+1), \dots, u(t+N)$ can be predicted at the time instant t . Menga and Mosca (1980) introduce the predictive models

$$\left. \begin{aligned} y(t+i) &= \mu_i u(t) + \Theta_i \varphi(t) + w_i(t+i) \\ u(t+i) &= \phi_i u(t) + \psi_i \varphi(t) + v_i(t+i), \quad i=1, \dots, N \end{aligned} \right\} \quad (5.9)$$

where $\varphi(t)$ is given by (2.13 *b*) and $w_i(t+i), v_i(t+i), i=1, \dots, N$ are uncorrelated with $u(t)$ and $\varphi(t)$. Minimization of an N -step loss function of the form (3.15) using (5.9) gives the control law

$$u(t) = - \left[Q_u + \sum_{i=1}^N (\mu_i^T Q_y \mu_i + \phi_i^T Q_u \phi_i) \right]^{-1} \times \sum_{i=1}^N (\mu_i^T Q_y \Theta_i + \phi_i^T Q_u \psi_i) \varphi(t) \quad (5.10)$$

An adaptive controller can be designed in analogy with Algorithm 5.1, by estimating the parameters of the predictive models (5.9) using the least squares method, and by determining the control signal according to (5.10). Simulation studies (Menga and Mosca 1980, Mosca, Zappa and Manfredi 1982) have shown that the algorithm performs favourably. Some theoretical results for the method have also been reported (Mosca and Zappa 1980).

6. Simulated examples

Some simulated examples are now given in order to illustrate the use of the adaptive controllers. More examples of the various algorithms can be found in the references.

Example 1

Consider the system

$$y(t) + A_1 y(t-1) = B_0 u(t-1) + B_1 u(t-2) + e(t) \quad (6.1)$$

where

$$A_1 = \begin{bmatrix} -0.8 & 0.5 \\ 0.1 & -0.4 \end{bmatrix}, \quad B_0 = \begin{bmatrix} 2 & 0 \\ 1 & 0 \end{bmatrix}, \quad B_1 = \begin{bmatrix} 1.2 & -1 \\ 0.5 & -2 \end{bmatrix}, \quad R_e = E e(t) e(t)^T = 0.25I$$

The system is difficult to control due to the structure of the B_0 -matrix. In minimum variance control the procedure described in § 4 can be used to handle the time delays. The approach is suboptimal in this case, as an additional time delay is introduced in the first input, cf. eqn. (4.21 b). The approach also requires that the time delays are known.

For these reasons it is appropriate to use a method based on a multi-step loss function. Here Algorithm 3.1 is used with the weights $Q_y = 100I$, $Q_u = I$. The Riccati equation was solved on-line by performing one iteration at each step, cf. eqn. (3.8). This method of solving the Riccati equation required that the estimates do not vary too much. Therefore, in order to obtain reasonably good initial estimates, the inputs applied during the first 20 steps were PRBS-sequences of unit amplitude.

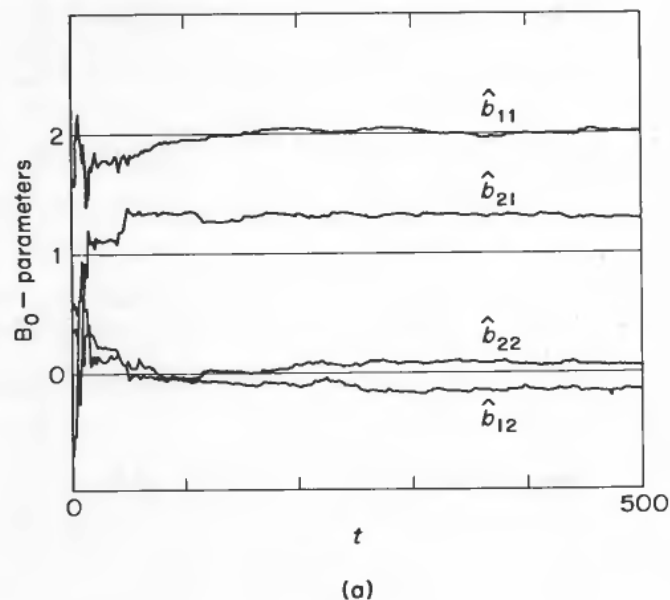
The estimated parameters obtained when the algorithm was simulated for the system (6.1) are shown in Fig. 1. Figure 2 shows the resulting feedback coefficients in the control law

$$u(t) = [L_1(t) \ L_2(t)] \begin{bmatrix} y(t) \\ u(t-1) \end{bmatrix} \quad (6.2)$$

It is seen that the use of (3.8) to update the feedback law gives good performance when the parameter estimates do not change too much from step to step.

Example 2

In this example a variance-constrained optimal control problem of the form (3.10) is considered. The example also illustrates what happens when Algorithm 3.1 is based on the least squares model (3.9) in a case when the model is not compatible with the system.



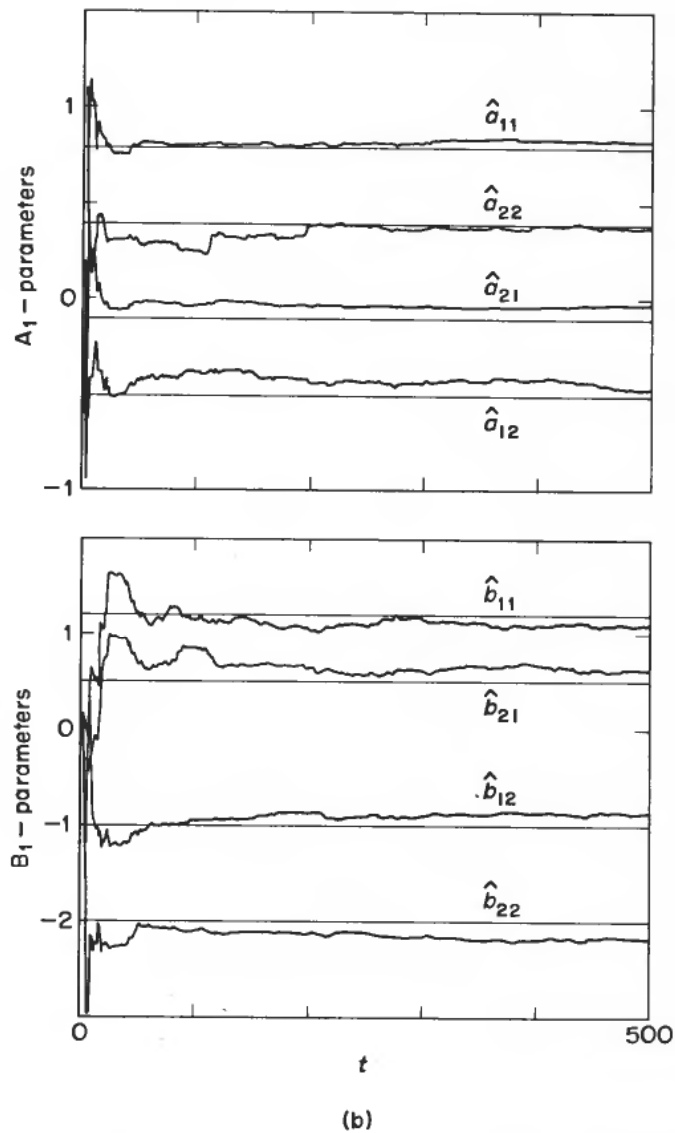


Figure 1. Estimated parameters in Example 1. The true system parameters are indicated by straight lines.

The following model has been applied to describe the dynamic behaviour of an industrial cement kiln (Westerlund 1980)

$$y(t) + A_1 y(t-1) = B_0 u(t-1) + e(t) + C_1 e(t-1) \quad (6.3)$$

where

$$A_1 = \begin{bmatrix} -0.917 & -0.0846 \\ 0.132 & -0.915 \end{bmatrix}, \quad B_0 = \begin{bmatrix} 2.06 & -0.0746 \\ -0.108 & -0.0192 \end{bmatrix}$$

$$C_1 = \begin{bmatrix} -0.0449 & -0.216 \\ 0.0256 & 0.841 \end{bmatrix}, \quad R_e = \begin{bmatrix} 0.0639 & 0.00188 \\ 0.00188 & 0.0233 \end{bmatrix}$$

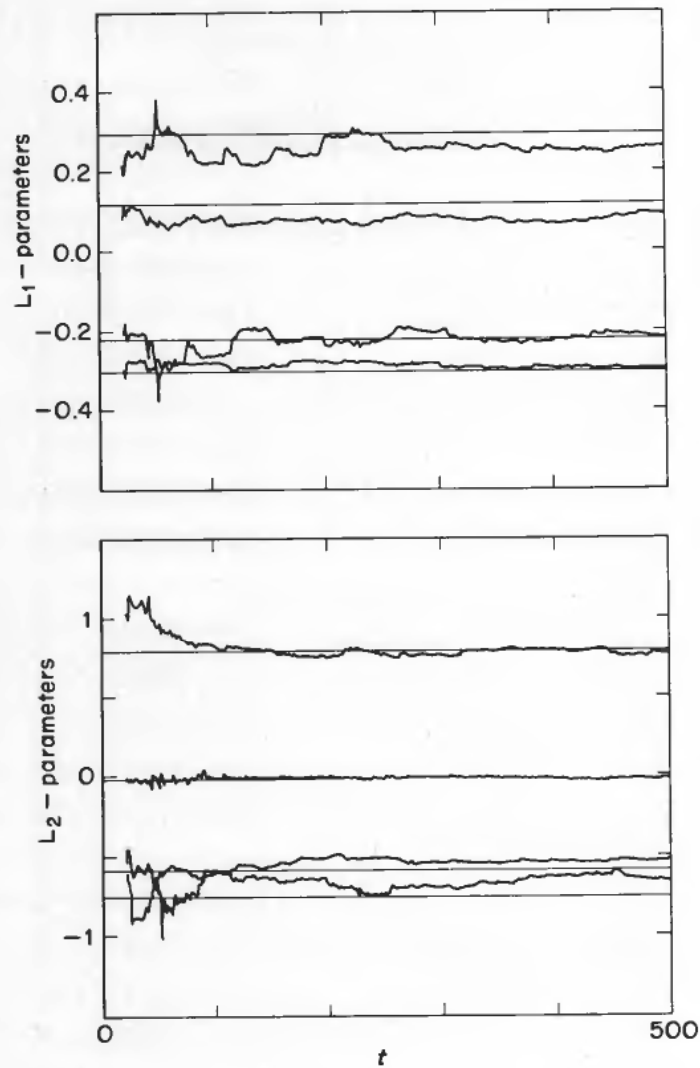


Figure 2. Control law parameters in Example 1.

The criterion for control is to minimize the loss function

$$V = \lim_{N \rightarrow \infty} E \frac{1}{N} \sum_{t=1}^N y_1(t)^2 + y_2(t)^2 = r_{y_1} + r_{y_2} \quad (6.4 a)$$

where the r_y denote the steady state variances of the outputs, subject to the variance constraints

$$\left. \begin{aligned} r_{u_1} &= \lim_{N \rightarrow \infty} E \frac{1}{N} \sum_{t=1}^N u_1(t)^2 \leq 0.004 \\ r_{u_2} &= \lim_{N \rightarrow \infty} E \frac{1}{N} \sum_{t=1}^N u_2(t)^2 \leq 1.5 \end{aligned} \right\} \quad (6.4 b)$$

The optimal variance constrained strategy for (6.3) gives the loss and the variances

$$V_{\min}=0.338, \quad r_{y_1}=0.093, \quad r_{y_2}=0.245, \quad r_{u_1}=0.004, \quad r_{u_2}=1.5 \quad (6.5)$$

Algorithm 3.1 was applied using the least squares model

$$y(t) + \mathcal{A}_1 y(t-1) = \mathcal{B}_0 u(t-1) + \mathcal{B}_1 u(t-2) + \epsilon(t) \quad (6.6)$$

The input weights were adapted according to eqns. (3.12) and (3.14). The value $\mu=0.02$ was used in (3.14). Figure 3 shows the result of a simulation of the adaptive algorithm for the system described by (6.3). The initial parameter values were

$$\mathcal{A}_1(0)=0, \quad \mathcal{B}_1(0)=0, \quad \mathcal{B}_0(0)=\begin{bmatrix} 2 & 0 \\ 0 & -0.02 \end{bmatrix}, \quad Q_u=\begin{bmatrix} 10 & 0 \\ 0 & 0.01 \end{bmatrix}$$

The sample averages obtained when using the adaptive controller are given in the Table. For comparison, the Table also shows the sample averages obtained when using the optimal variance constrained strategy which can be determined for (6.3).

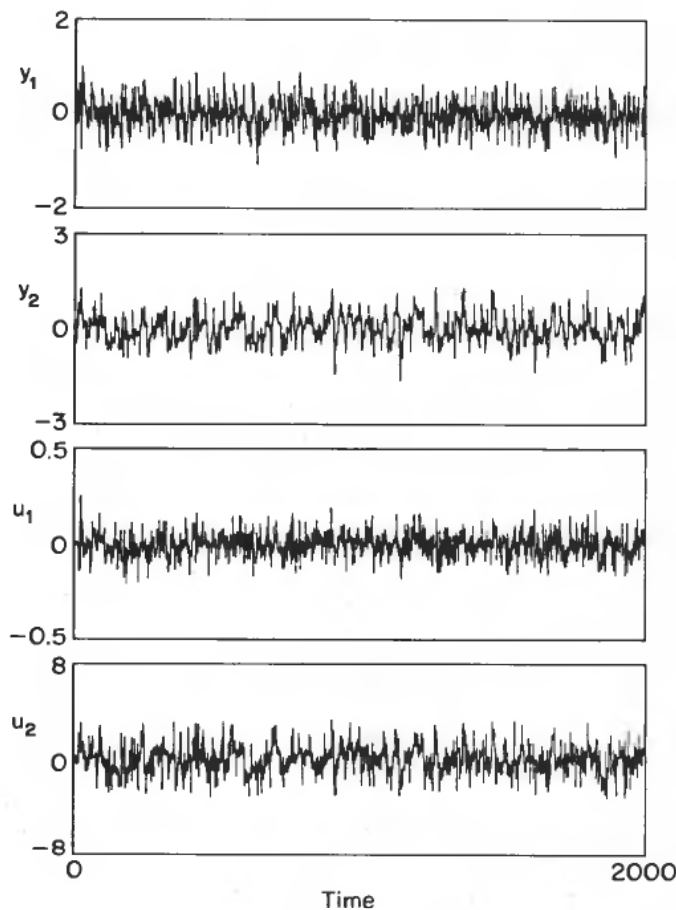


Figure 3. Inputs and outputs of the system (6.3) in Example 2 when using the adaptive controller based on (6.6).

| Sample averages | Adaptive controller | Optimal strategy |
|-----------------|---------------------|------------------|
| V | 0.284 | 0.281 |
| r_{y_1} | 0.089 | 0.085 |
| r_{y_2} | 0.195 | 0.197 |
| r_{u_1} | 0.00401 | 0.00367 |
| r_{u_2} | 1.44 | 1.32 |

Sample averages obtained for the system (6.3) in a 2000 step simulation when using the adaptive controller and the optimal variance constrained strategy. The same noise realization was used in the simulations.

As the model (6.6) is incompatible with (6.3) the adaptive controller, when based on (6.6), does not converge to the optimal variance constrained strategy. The convergence of the algorithm was studied by solving the associated differential equation (cf. eqn. (2.15)) numerically. When the initial parameters were taken from the above simulation at $t=2000$, the differential equation converged to a feedback law giving the variances

$$V=0.350, \quad r_{y_1}=0.119, \quad r_{y_2}=0.231, \quad r_{u_1}=0.004, \quad r_{u_2}=1.5, \quad (6.7)$$

which can be compared with (6.5). The performance obtained in simulations, and the near-optimality of the convergence point show that good results are obtained with the least squares model (6.6). This is encouraging from a practical point of view, since least squares models are often used in practice.

7. Conclusion

Techniques for multivariable stochastic adaptive control have been described. The basic control problem considered is to control the system in such a way that the steady state input and output variances are as small as possible. The adaptive algorithms have been classified according to the underlying design method into algorithms based on linear quadratic gaussian design, adaptive minimum variance controllers, and algorithms based on single-step optimal control.

The optimal off-line design method for the stochastic control problems leads to linear quadratic gaussian design. The adaptive controllers based on linear quadratic gaussian design involve the on-line solution of a steady state Riccati equation, which increases the computational requirements of the algorithm. In return, the approach has the advantage that it can be applied to systems with unknown time delays and other non-minimum phase properties.

The adaptive minimum variance controller is designed for the case when the criterion for control is to achieve minimum output variance. The algorithm is appealing due to its simplicity. A major restriction of the procedure is the condition that the system should be stably invertible. It has been shown how system time delays can be handled in this approach if they are known.

The adaptive algorithms based on single-step optimal control are generalizations of the adaptive minimum variance controller, preserving its computational simplicity, but having design parameters by which the closed loop behaviour can be affected. The approach can for example be applied to systems which are not stably invertible, by

selecting the design weights of the generalized single-step loss function properly. The problem is, however, that it is not always straightforward to find proper values for the design parameters in such cases. The design method is also suboptimal since a single-step loss function is used.

A generalization of the procedure based on single-step optimal control is the so-called MUSMAR-algorithm, in which several predictive models with different prediction times are used. The control signal is then determined to minimize a multi-step loss function. As a multistep loss function is used, the procedure, which is moderately simple, can be applied to systems with unknown time delays and other non-minimum phase properties.

Many of the adaptive control techniques which have been considered are relatively new, and they are still being developed. Therefore, very few applications of the methods have been reported. However, in recent years there has been an increasing interest in industrial applications of adaptive controllers designed for single-input single-output systems (Åström *et al.* 1977, Narendra and Monopoli 1980, Fjeld and Willems 1981, Åström 1983). The multivariable adaptive controllers seem promising for industrial applications as well.

REFERENCES

- ÅSTRÖM, K. J. (1970). *Introduction to Stochastic Control Theory* (New York: Academic Press).
- ÅSTRÖM, K. J. (1977). Control of systems with uncertain parameters, Report TFRT 7115, Dep. of Automatic Control, Lund Institute of Technology, Lund, Sweden.
- ÅSTRÖM, K. J. (1978). Stochastic control problems. In *Mathematical Control Theory*, Lecture Notes in Mathematics, edited by W. A. Coppel (Berlin: Springer).
- ÅSTRÖM, K. J. (1983). Theory and applications of adaptive control—a survey. *Automatica*, **19**, 471–486.
- ÅSTRÖM, K. J., BORISSON, U., LJUNG, L., and WITTENMARK, B. (1977). Theory and applications of self-tuning regulators. *Automatica*, **13**, 457–476.
- ÅSTRÖM, K. J., and WITTENMARK, B. (1973). On self-tuning regulators. *Automatica*, **9**, 185–199.
- ÅSTRÖM, K. J., and WITTENMARK, B. (1974). Analysis of a self-tuning regulator for non-minimum phase systems. *Proceedings, IFAC Symposium on Stochastic Control*, Budapest.
- ÅSTRÖM, K. J., and ZHAO-YING, Z. (1982). A linear quadratic Gaussian self-tuner. *Proceedings, Workshop on Adaptive Control*, Florence, Italy.
- ATHANS, M. (1971). The role and use of the stochastic linear-quadratic-gaussian problem in control system design. *I.E.E. Trans. Autom. Control*, **AC-16**, 529–552.
- BARTOLINI, G., CASALINO, G., DAVOLI, F., MINCIARDI, R., and ZOPPOLI, R. (1982). Efficient algorithms for multivariable adaptive control: the ICOF scheme. *Control and Computers*, **10**, 58–62.
- BIERMAN, G. (1977). *Factorization Methods for Discrete Estimation* (New York: Academic Press).
- BORISSON, U. (1975). Self-tuning regulators—industrial applications and multivariable theory, Report 7513, Dep. of Automatic Control, Lund Institute of Technology, Lund, Sweden.
- BORISSON, U. (1979). Self-tuning regulators for a class of multivariable systems. *Automatica*, **15**, 209–215.
- CLARKE, D. W., and GAWTHROP, P. J. (1975). Self-tuning controller. *Proc. I.E.E.*, **122**, 929–934.
- CLARKE, D. W., and GAWTHROP, P. J. (1979). Self-tuning control. *Proc. I.E.E.*, **126-D**, 633–640.
- CLARKE, D. W., and GAWTHROP, P. J. (1981). Implementation and application of micro-processor-based self-tuners. *Automatica*, **17**, 233–244.

- FJELD, M., and WILLEMS, R. G. (1981). Self-tuning regulators—the software way. *Control Engineering*, 99–102.
- FORTESCUE, T. R., KERSHENBAUM, L. S., and YDSTIE, B. E. (1981). Implementation of self-tuning regulators with variable forgetting factors. *Automatica*, 17, 831–835.
- FURUTA, K., and PAQUET, J. G. (1975). Determination of matrix transfer functions in the form of matrix fractions from input-output observations, *I.E.E.E. Trans. Autom. Control*, AC-20, 392–396.
- GOODWIN, G. C., and DUGARD, L. (1983). Stochastic adaptive control with known and unknown interactor matrices. *Preprints of the IFAC Workshop on Adaptive Systems in Control and Signal Processing*, San Francisco.
- GOODWIN, G. C., McINNIS, B. C., and WANG, J. C. (1982). Model reference adaptive control for systems having non-square transfer functions. *Proceedings, 21st I.E.E.E. Conference on Decision and Control*, Orlando.
- GOODWIN, G. C., and PAYNE, R. L. (1977). *Dynamic System Identification: Experiment Design and Data Analysis* (New York: Academic Press).
- GOODWIN, G. C., RAMADGE, P. J., and CAINES, P. E. (1981). Discrete time stochastic adaptive control. *SIAM J. Control Optim.*, 19, 829–853.
- GUSTAVSSON, I., LJUNG, L., and SÖDERSTRÖM, T. (1977). Identification of processes in closed-loop—identifiability and accuracy aspects. *Automatica*, 13, 59–75.
- HÄGGLUND, T. (1983 a). Recursive least squares identification with forgetting of old data, Report TFRT-7254, Dep. of Automatic Control, Lund Institute of Technology, Lund, Sweden.
- HÄGGLUND, T. (1983 b). The problem of forgetting old data in recursive estimation. *Preprints of the IFAC Workshop on Adaptive Systems in Control and Signal Processing*, San Francisco.
- HARRIS, C. J., and BILLINGS, S. A. (Editors) (1981). *Self-Tuning and Adaptive Control: Theory and Application* (London: Peter Peregrinus Ltd.).
- KALMAN, R. E. (1958). Design of a self-optimizing control system. *Am. Soc. Mech. Engr. Trans.*, 80, 468–478.
- KEVICZKY, L., and KUMAR, K. S. P. (1981). Multivariable self-tuning regulator with generalized cost function. *Int. J. Control*, 33, 913–921.
- KOIVO, H. N. (1980). A multivariable self-tuning controller. *Automatica*, 16, 351–366.
- LAM, K. P. (1980). D.Phil. thesis, Oxford University.
- LANDAU, I. D. (1979). *Adaptive Control—The Model Reference Approach* (New York: Marcel Dekker).
- LJUNG, L. (1977 a). On positive real transfer functions and the convergence of some recursions, *I.E.E.E. Trans. Autom. Control*, AC-22, 539–551.
- LJUNG, L. (1977 b). Analysis of recursive stochastic algorithms. *I.E.E.E. Trans. Autom. Control*, AC-22, 551–575.
- LJUNG, L. (1978). Convergence of an adaptive filter algorithm. *Int. J. Control*, 27, 673–693.
- LJUNG, L. (1980). The ODE approach to the analysis of adaptive control systems—possibilities and limitations. *Proceedings, 1980 Joint Automatic Control Conference*, San Francisco.
- LJUNG, L., and WITTENMARK, B. (1974). Asymptotic properties of self-tuning regulators, Report 7404, Dep. of Automatic Control, Lund Institute of Technology, Lund, Sweden.
- MACGREGOR, J. F. (1973). Optimal discrete stochastic control theory for process application. *Can. J. Chem. Engng*, 51, 468–478.
- MÄKILÄ, P. M. (1982). Self-tuning control with linear input constraints. *Optimal Control Appl. & Methods*, 3, 337–353.
- MÄKILÄ, P. M., WESTERLUND, T., and TOIVONEN, H. T. (1984). Constrained linear quadratic gaussian control with process applications. *Automatica*, 20, 15–29.
- MENGA, G., and MOSCA, E. (1980). MUSMAR: multivariable adaptive regulators based on multistep cost functionals. In *Advances in Control*, edited by D. G. Lainiotis and N. S. Tzannes (D. Reidel Publishing Company).
- MODÉN, P. E., and SÖDERSTRÖM, T. (1982). Stationary performance of linear stochastic systems under single-step optimal control. *I.E.E.E. Trans. Autom. Control*, AC-23, 214–216.
- MOSCA, E., and ZAPPA, G. (1980). MUSMAR: basic convergence and consistency properties. In *Lecture Notes in Control*, 28, (Springer).

- MOSCA, E., ZAPPA, G., and MANFREDI, C. (1982). Robust adaptive prediction and control by MUSMAR techniques. *Proceedings, 6th IFAC Symposium on Identification and System Parameter Estimation*, Arlington.
- NARENDARA, K. S., and MONOPOLI, R. V. (editors) (1980). *Applications of Adaptive Control* (New York: Academic Press).
- PANOSSIAN, H., and LEONIDES, C. T. (1983). On discrete-time Riccati-like matrix difference equations with random coefficients. *Int. J. Systems Sci.*, **14**, 385-407.
- PETERKA, V. (1970). Adaptive digital regulation of noisy systems. *IFAC Symposium on Identification and Process Parameter Estimation*, Prague.
- PETERKA, V. (1975). A square root filter for real time multivariate regression. *Kybernetika*, **11**, 53-67.
- PETERKA, V., and ÅSTRÖM, K. J. (1973). Control of multivariable systems with unknown but constant parameters. *Proceedings, 3rd IFAC Symposium on Identification and System Parameter Estimation*, The Hague.
- SÖDERSTRÖM, T., LJUNG, L., and GUSTAVSSON, I. (1978). A theoretical analysis of recursive identification methods. *Automatica*, **14**, 231-244.
- TANTTU, J. T., and KOIVO, H. (1983). A multivariable self-tuning controller: different delays in the control signal. *Proceedings of the IASTED Symposium*, Copenhagen.
- TOIVONEN, H. T. (1981). Ph.D. thesis, Dep. Chem. Eng., Åbo Akademi, Turku, Finland.
- TOIVONEN, H. T. (1983 a). Variance constrained self-tuning control. *Automatica*, **19**, 415-418.
- TOIVONEN, H. T. (1983 b). Self-tuning regulators with explicit variance restrictions. *Int. J. Control*, **38**, 229-235.
- TOIVONEN, H. T. (1983 c). A self-tuning controller for systems with non-square transfer functions. Report 83-3, Process Control Lab., Åbo Akademi, Turku, Finland.
- WESTERLUND, T. (1980). Identification and control of an industrial dry process cement kiln. *Proceedings, 1980 Joint Automatic Control Conference*, San Francisco.
- WITTENMARK, B. (1973). A self-tuning regulator. Report 7311, Dep. of Automatic Control, Lund Institute of Technology, Lund, Sweden.
- WITTENMARK, B. (1975). Stochastic adaptive control methods—a survey. *Int. J. Control*, **21**, 705-730.