Stochastic Control Problems

Åström, Karl Johan

1977

Document Version:
Publisher's PDF, also known as Version of record

CODEN : LUTFD2/(TFRT-3147)/1-068/(1977)

STOCHASTIC CONTROL PROBLEMS

K.J ÅSTRÖM

Department of Automatic Control
Lund Institute of Technology
November 1977

Dokumentutgivare
Lund Institute of Technology
Handläggare Dept of Automatic Control
K J Aström
Författare
K J Aström

Dokumentnamn
REPORT
Utgivningsdatum
Nov. 1977

Dokumentbeteckning
LUTFD2/(TFRT-3147)/1-68/(1977)
Ärendebeteckning

Dokumenttitel och undertitel

Stochastic Control Problems

Referat (sammandrag)

This report is the write-up of six two-hour-lectures of stochastic control theory. The treatment is limited to discrete time systems. Both linear and nonlinear systems are covered. The material can be used as an introduction at a graduate level. It is also suitable material for an introduction to stochastic adaptive control.

Referat skrivet av
Author
Förslag till ytterligare nyckelord
stochastic control, prediction, filtering, minimum variance control, linear quadratic gaussian control, Kalman filtering, separation theorem, duality, certainty equivalence, caution, dual control, self-tuning regulators, Bellman equation.
Klassifikationssystem och -klass(er)

Indextermer (ange källa)

Omfång
68 pages
Övriga bibliografiska uppgifter
Språk
English
Sekretessuppgifter
ISSN
ISBN
Dokumentet kan erhållas från
Department of Automatic Control
Lund Institute of Technology
Box 725, S-220 07 LUND 7, Sweden
Mottagarens uppgifter
Pris

STOCHASTIC CONTROL PROBLEMS

K.J. Åström

This work was partly supported by the Swedish Board of Technical Development under Contract No. 76-3804.

Table of Contents

Chapter 1 - Introduction 1

Chapter 2 - Minimum Variance Control 5
1. Introduction 5
2. Mathematical Models 5
3. Optimal Prediction 8
4. Minimum Variance Control 11
5. Applications 21
6. References 22

Chapter 3 - Linear Quadratic Gaussian Control 23
1. Introduction 23
2. Mathematical Models 23
3. Kalman Filtering and Prediction 25
4. Optimal Control 29
5. Comparison with Minimum Variance Control 35
6. Applications 36
7. References 36

Chapter 4 - Control of Markov Chains 38
1. Introduction 38
2. Mathematical Models 38
3. Optimal Filtering 39
4. Optimal Control 40
5. An Example 43
6. References 45

Chapter 5 - Nonlinear Stochastic Control 46
1. Introduction 46
2. Mathematical Models 46
3. Optimal Filtering 48
4. Optimal Control 50
5. Linear Systems with Random Parameters 51
6. References 58

Chapter 6 - Self-Tuning Regulators 59
1. Introduction 59
2. Mathematical Model 59
3. A Simple Self-Tuning Regulator 60
4. Analysis 63
5. Conclusions 65
6. References 65

CHAPTER 1 - INTRODUCTION

The purpose of these lectures is to present some basic stochastic control problems and to present mathematical theory that is useful in solving the problems. To provide a red line in the lectures they are focused on a specific problem, namely to understand feedback mechanisms which is a fundamental problem of control engineering.

A schematic picture of a process with feedback is shown in Fig. 1. The process is characterized by *inputs*, i.e. variables which can be manipulated, *outputs*, i.e. variables that can be measured, and *disturbances*. The disturbances describe the interaction between the environment and the process. It is assumed that this interaction is such that the environment influences the process but that the process does not influence the environment. The feedback mechanism receives information about the process and the environment through the measurements and it generates appropriate control actions so that the closed loop system behaves appropriately in spite of the disturbances from the environment. A common example of a feedback law is the PI regulator which is described by

$$\begin{cases} u(t) = u_{ref}(t) + K\left[e(t) + \frac{1}{T}\int^{t} e(s)\, ds\right] \\ e(t) = y_{ref}(t) - y(t), \end{cases} \tag{1}$$

where t is time, u is the input signal, y the output signal, u_{ref} and y_{ref} are reference values for the input and the output. It is very fortunate for the control engineer that many processes can be controlled very successfully using a PI-regulator provided that the parameters K and T are chosen appropriately. This fact is of course less fortunate for the control theoretician.

Feedback processes were first explored purely empirically in connection with technical systems like centrifugal governors and electronic amplifiers. It has later been found that feedback processes also play an important role in economical, biological, environmental, and social systems.

Many attempts have been made to develop mathematical theory which will help to understand and to design feedback systems. Classical control theory was largely analytical in its nature. It gave tools for analysing a given feedback system. There was a great emphasis on stability theory. Synthesis and design problems were dealt with by repeated analysis. Over the past 30 years theory which aims directly

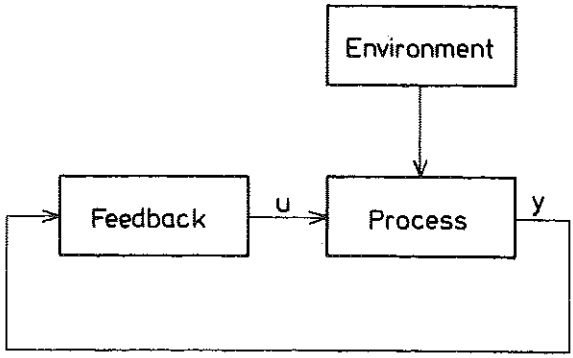

Fig. 1 - Schematic diagram of a feedback system.

at design and synthesis has been formulated. Optimal control theory is one idea. It answers the following problem. Given a description of the process to be controlled and a criterion which characterizes the desired behaviour of the closed loop system. Find the best feedback. One drawback with optimal control theory is that it does not necessarily give a feedback solution. This is shown by the following example.

EXAMPLE 1.1
Consider a process described by

$$\frac{dy}{dt} = u$$

with initial conditions

$$y(0) = a .$$

Assume that it is desirable to control the process in such a way that the performance of the system evaluated by the criterion

$$J = \int_0^\infty [y^2(t) + u^2(t)]\, dt$$

is as small as possible. It is easy to show that $J \geq a^2$ and that the minimum is achieved for the control signal

$$u(t) = - y(0)\, e^{-t}, \tag{1.1}$$

or for the feedback law

$$u(t) = - y(t), \tag{1.2}$$

or for any combination like

$$u(t) = -\alpha\, y(0)\, e^{-t} - (1-\alpha)\, y(t), \qquad \alpha \leq 1.$$

The control signal given by (1.1) is called a control program or an open loop solution because it requires only the knowledge at the measured output at time $t = 0$. Equation (1.2) gives a proper feedback law because the value of the control signal at time t is a function of the measured output at time t. It is clear that the solution (1.2) is more robust than the solution (1.1) because it will give a smaller value of the loss function if there are perturbations in the description of the model.

□

The example shows clearly that in order to get a feedback solution to an optimal control problem it is necessary to introduce disturbances and uncertainties in process descriptions. In stochastic control theory the disturbances are described as stochastic processes. Stochastic control theory will give valuable insights into the properties of feedback systems. It will give the structure of optimal feedback laws and it will e.g. tell when it is motivated to use a feedback law that is more complicated than a PI regulator. It will also in some cases give practical design tools.

From a mathematical point of view stochastic control theory is a combination of the theory of stochastic processes with the theory of differential and difference equations, calculus of variations, and optimal control theory.

The purpose of these lectures is to give an exposé of some ideas in stochastic control theory. The material for the lectures has been chosen in order to give some feel for the variety of the theory. Insight into the nature of feedback processes has been chosen as a unifying theme. For simplicity discrete time systems are treated throughout. Section 2 deals with a linear problem minimum variance control. The main virtue of this problem is that the theory is very simple and the ideas transparent. The feedback laws obtained are linear. It is thus a good starting point. The models used to describe the process and its environment are controlled ARMA (autoregressive moving average) processes or CARMA processes for short. The theory thus has strong ties to parametric time series analysis. The criterion is to minimize the variance of process outputs. Minimum variance control is of interest for control of some industrial processes where the purpose is to keep certain important quality variables as close as possible to prescribed limits. The theory will tell when and why it is useful to use a feedback law that is more complicated than the PI-

-regulator. The theory will also show the close relationships between minimum variance control and optimal predictor theory. The models used in the minimum variance control theory also occur in macroeconomics. There the models are referred to as "the reduced form" of the equations describing a macroeconomy.

The process models used in Chapter 2 are pure external descriptions. In Chapter 3 the linear stochastic control problem is approached from a different point of view. The main difference is that the model of process and its environment are now characterized by internal descriptions or state models. The criterion is again to minimize the expected value of a quadratic form. The problem statement is somewhat more general than the minimum variance problem. The major results of this theory are the Kalman filtering problem and the so called separation theorem or certainty equivalence problem which again will give important insights into the nature of the feedback control problem.

The feedback laws obtained in Chapters 2 and 3 are all linear feedbacks. In Chapter 4 we turn to models that will give nonlinear feedback laws. To keep the mathematics simple a problem with finite states is discussed. The nonlinear problem can then be solved and the solution will provide valuable insight into the properties of feedback control.

In Chapter 5 the results of Chapter 4 are generalized. The problem formulation will e.g. include the problems discussed in Chapter 2 with the additional complication of the process models now being unknown. The analysis will lead to discussion of notions of dual control, certainty equivalence, caution, and probing.

The control laws obtained in Chapter 5 are extremely complex. They can not be implemented with computing power available today. In Chapter 6 we therefore discuss simplifications that will have nice asymptotic properties. This leads to the notion of self-tuning regulators.

The books [1] and [2] listed below are useful supplementary reading. A reader interested in the continuous time problems can consult [3].

References

[1] H Kushner: Introduction to Stochastic Control. Holt, Rinehart and Winston, New York 1971.

[2] K J Åström: Introduction to Stochastic Control Theory. Academic Press, New York 1970.

[3] W H Fleming and R W Rishel: Deterministic and Stochastic Optimal Control. Springer-Verlag, New York 1975.

CHAPTER 2 - MINIMUM VARIANCE CONTROL

1. INTRODUCTION

A very simple stochastic control problem is discussed in this section. It is assumed that the process dynamics can be described by an external description in the form of a difference equation with constant coefficients and that the disturbances can be characterized as ARMA processes. It is assumed that the purpose of the control is to find a feedback law such that the fluctuations in the process output are as small as possible as measured by the output variance. The mathematical models for the process, its environment, and the criterion are discussed in Section 2. It turns out that there is a close relationship between minimum variance control and optimal prediction. The prediction problem being somewhat simpler is therefore first discussed in Section 3. The minimum variance problem is then formulated and solved in Section 4.

2. MATHEMATICAL MODELS

The mathematical models used to describe the process dynamics and its environment will now be discussed. Single-input single-output systems are first treated. It is found that a generic model called a CARMA process (Controlled ARMA process) can be obtained. The multivariable version of this process is then given.

Process Dynamics

Consider a system described by Fig. 1. Assume that there is one input and one output only. It is assumed that the relation between the measured output y and the control variable u can be described by the difference equation

$$y(t) + a_1' y(t-1) + \ldots + a_m' y(t-m) = b_0' u(t-k) + \ldots + b_m' u(t-k-m).$$

This is the case for example if the process can be described by an ordinary linear differential equation with constant coefficients and a time delay and if the input signal is assumed constant over sampling intervals of unit length. Introduce the backward shift-operator q^{-1} and the polynomials

$$A_1(q^{-1}) = 1 + a_1' q^{-1} + \ldots + a_m' q^{-m}$$

$$B_1(q^{-1}) = b_0' + b_1' q^{-1} + \ldots + b_m' q^{-m}.$$

The model can then be written as

$$y(t) = \frac{B_1(q^{-1})}{A_1(q^{-1})} u(t-k). \tag{2.1}$$

This model is often a reasonable approximation of many engineering processes that are being operated close to equilibrium conditions.

The Environment

It is assumed that the action of the environment on the process can be described by a disturbance n acting on the output. Adding a disturbance n to the output y of (2.1) gives

$$y(t) = \frac{B_1(q^{-1})}{A_1(q^{-1})} u(t-k) + n(t). \tag{2.2}$$

There may in fact be many different disturbances acting on the process. Under the linearity assumption it is possible to use the superposition principle to reduce all disturbances to an equivalent disturbance n on the output. The disturbance n thus has physical interpretation as the output that would be observed if there is no control i.e. $u = 0$. Moreover it is assumed that the disturbance n can be represented by

$$n(t) = \frac{C_1(q^{-1})}{A_2(q^{-1})} \varepsilon(t), \tag{2.3}$$

where $\{\varepsilon(t), t = 0, \pm 1, \pm 2, \ldots\}$ is a sequence of independent normal random variables and $C_1(q^{-1})$ and $A_2(q^{-1})$ are polynomials in the backward shift operator. Such a representation is certainly possible if n is a stationary stochastic process with a rational spectral density. The representation (2.3) will however not necessarily require a stationarity assumption. Non-stationary processes can be handled by letting the polynomial $A_2(\xi)$ be unstable i.e. have zeros inside the unit disc.

The CARMA Model

A combination of the equations (2.2) and (2.3) gives the following description of the process

$$y(t) = \frac{B_1(q^{-1})}{A_1(q^{-1})} u(t-k) + \frac{C_1(q^{-1})}{A_2(q^{-1})} \varepsilon(t).$$

By introducing the polynomials $A = A_1A_2$, $B = B_1A_2$, and $C = C_1A_1$,

this description can be simplified to

$$A(q^{-1})\ y(t) = B(q^{-1})\ u(t-k) + C(q^{-1})\ \varepsilon(t), \tag{2.4}$$

where

$$A(q^{-1}) = 1 + a_1 q^{-1} + \ldots + a_n q^{-n}$$

$$B(q^{-1}) = b_0 + b_1 q^{-1} + \ldots + b_n q^{-n}, \qquad b_0 \neq 0$$

$$C(q^{-1}) = 1 + c_1 q^{-1} + \ldots + c_n q^{-n}.$$

There is no loss in generality in assuming that all polynomials are of degree n because we can always put trailing coefficients equal to zero.

The mathematical model (2.4) will be called a CARMA (controlled ARMA) process, because without the control i.e. $u = 0$ the model is identical to the ARMA process which is commonly used in time series analysis. Notice also that without the disturbance the model is a simple rational transfer function model which is commonly used in engineering.

Notice that it is always possible to assume that the polynomial $C(\xi)$ has all its zeros outside the unit disc or on the unit circle. This is seen as follows. The polynomial $C(q^{-1})$ only enters the system description in the description of a disturbance

$$v(t) = C(q^{-1})\ \varepsilon(t)$$

The signal v is completely characterized by its covariances.

$$r_v(k) = E\ v(t)\ v(t+k) = \sum_{i=0}^{n-k} C_i\ C_{i+k}.$$

The covariance $r_v(k)$ is also given as the coefficient of the term q^k or q^{-k} in the Laurent series of the function $C(\xi)\ C(\xi^{-1})$. But by factoring the polynomial and sorting the factors differently it is always possible to find a polynomial $\tilde{C}(\xi)$ with all zeros outside the unit disc or on the unit circle such that

$$C(\xi)\ C(\xi^{-1}) = \tilde{C}(\xi)\ \tilde{C}(\xi^{-1}).$$

A simple example will serve as an illustration.

EXAMPLE 2.1
Consider the random process $\{v(t)\}$

$$v(t) = \varepsilon(t) + c\varepsilon(t-1) = (1 + cq^{-1})\ \varepsilon(t),$$

where $c > 1$ and $\text{var}\ \varepsilon(t) = 1$. Hence

$$C(q^{-1})C(q) = (1+cq^{-1})(1+cq) = (c+q^{-1})(c+q) = c^2(1+c^{-1}q^{-1})(1+c^{-1}q).$$

The stochastic process $\{v(t)\}$ can thus also be represented as

$$v(t) = (1+\frac{1}{c}q^{-1})\,c\varepsilon(t) = \varepsilon'(t) + \frac{1}{c}\,\varepsilon'(t-1),$$

where $\text{var}\ \varepsilon'(t) = c^2$.

□

Multivariable Generalizations

The CARMA model can easily be generalized to the multivariable case. The description (2.4) still holds provided that $y(t)$, $u(t)$, and $\varepsilon(t)$ are interpreted as vectors and that $A(q^{-1})$, $B(q^{-1})$, and $C(q^{-1})$ are interpreted as matrix polynomials. The vectors y and ε can be chosen to be of the same dimension. The matrix polynomial $C(q^{-1})$ can always be chosen in such a way that $\det C(\xi)$ will always have its zeros outside the unit disc or on the unit circle.

The multivariable CARMA model can be used to represent input output relations for multivariable industrial regulation problems. This model is also used in economics to represent the so called reduced form of a macro economic model.

3. OPTIMAL PREDICTION

The optimal prediction problem will now be discussed as a preliminary to solve the minimum variance control problem. The main result is given by

THEOREM 3.1
Let $\{y(t), t = 0, \pm1, \pm2, \ldots\}$ be a normal stochastic process with the representation

$$A(q^{-1})\ y(t) = C(q^{-1})\ \varepsilon(t), \tag{3.1}$$

where $\{\varepsilon(t), t = 0, \pm1, \pm2, \ldots\}$ is a sequence of independent normal $(0, R)$ random variables. Assume that the polynomial $\det C(\xi)$ has all its zeros outside the unit disc. Then the k-step predictor which minimizes the variance of the prediction error in steady state is given by

$$\hat{y}(t+k|t) = G(q^{-1})\ C^{-1}(q^{-1})\ y(t), \tag{3.2}$$

where

$$A^{-1}(q^{-1})C(q^{-1}) = C(q^{-1})A^{-1}(q^{-1}) = F(q^{-1}) + q^{-k}\, G(q^{-1})A^{-1}(q^{-1}) \tag{3.3}$$

and the polynomial $F(q^{-1})$ is of degree k-1:

$$F(q^{-1}) = I + F_1 q^{-1} + \ldots + F_{k-1}\, q^{-k+1}. \tag{3.4}$$

The error of the optimal predictor is a moving average of order k

$$\tilde{y}(t+k|t) = \varepsilon(t+k) + F_1\, \varepsilon(t+k-1) + \ldots + F_{k-1}\, \varepsilon(t+1) \tag{3.5}$$

and the covariance of the prediction error is

$$\operatorname{cov}[\tilde{y},\tilde{y}] = R + F_1\, R\, F_1^T + \ldots + F_{k-1}\, R\, F_{k-1}^T. \tag{3.6}$$

Proof:
The proof is straightforward and constructive. Equations (3.1) and (3.3) give

$$y(t+k) = CA^{-1}\varepsilon(t+k) = F\varepsilon(t+k) + GA^{-1}\varepsilon(t).$$

Substitution of ε by y in the last term using (3.1) gives

$$y(t+k) = F\varepsilon(t+k) + GC^{-1}y(t).$$

The expression $GC^{-1}y(t)$ exist because $\det \mathcal{C} = \det C$ and it was assumed that all zeros of $\det C(\xi)$ were outside the unit disc. Now let $\hat{y}$ be an arbitrary function of $y(t), y(t-1), \ldots$ Consider the prediction error

$$\tilde{y}(t+k|t) = y(t+k) - \hat{y} = F\varepsilon(t+k) + [GC^{-1}y(t) - \hat{y}]. \tag{3.7}$$

Let a be an arbitrary vector. Then

$$E[a^T\tilde{y}(t+k|t)]^2 = E[a^TF\varepsilon(t+k)]^2 + E\{a^T[GC^{-1}y(t) - \hat{y}]\}^2 + \\ + 2E\{a^TF\varepsilon(t+k)\, a^T[GC^{-1}y(t) - \hat{y}]\}. \tag{3.8}$$

The last term vanishes because $\varepsilon(t+k), \varepsilon(t+k-1), \ldots \varepsilon(t+1)$ are all independent of $y(t), y(t-1), \ldots$ and then also independent of $\hat{y}$. The predictor (3.2) thus gives the minimum value of the prediction error for all a. It then follows from (3.7) that the prediction error is given by (3.5). A simple calculation based on $\{\varepsilon(t)\}$ being independent then gives (3.6).

□

Remark 1
Notice that the best predictor is linear. The linearity does not

depend critically on the minimum variance criterium. Since $\{y(t)\}$ is normal, the result would be the same for all criteria of the form $E\, h\{a^T[y(t+k) - \hat{y}]\}$ provided that h is symmetrical.

Remark 2

The assumption that $\varepsilon(t)$ and $\varepsilon(s)$ are independent for $t \neq s$ is crucial for the argument that the last term in (3.8) will vanish. If the stochastic variables $\varepsilon(t)$ and $\varepsilon(s)$ are not independent it is in general not true that the product of $\varepsilon(t+\tau)$ and an arbitrary function of $y(t), y(t-1), \ldots$ will vanish. However if the predictor is restricted to be linear functions of $y(t), y(t-1), \ldots$ then it is sufficient to assume $\varepsilon(t)$ and $\varepsilon(s)$ uncorrelated for the proof to hold. This situation is typical for linear problems with quadratic criteria.

Remark 3

Notice that it follows from (3.5) that

$$\tilde{y}(t+1|t) = y(t+1) - \hat{y}(t+1|t) = \varepsilon(t+1).$$

The stochastic variables $\{\varepsilon(t)\}$ can thus be interpreted as the innovations of the stochastic process $\{y(t)\}$. It is straightforward to calculate the predictor. The polynomials A and C such that $CA^{-1} = A^{-1}C$ are first determined. The polynomials F and G are then obtained as the quotient of degree $k-1$ and the remainder obtained when dividing C by A.

Notice that the predictor is a dynamical system (3.2) whose dynamics is governed by the matrix polynomial $C(\xi)$. The assumption in the theorem that $\det C(\xi)$ has all its zeros outside the unit disc guarantees that the predictor is stable. The initial conditions chosen for the predictor are thus immaterial. It was shown in Section 2 that the model could always be chosen in such a way that $\det C(\xi)$ has all its zeros outside the unit disc or on the unit circle. The Theorem 3.1 thus assumes away the case when $\det C(\xi)$ has zeros on the unit circle. This case requires special treatment because the optimal predictor is timevarying. A simple example illustrates what happens.

EXAMPLE 3.1

Consider the following scalar process

$$y(t) = \varepsilon(t) - \varepsilon(t-1).$$

In this case the polynomial $C(\xi) = 1 - \xi$ has apparently a zero on the unit circle. The one-step predictor is given by

$$\hat{y}(t+1|t) = -\varepsilon(t).$$

Attempting to calculate $\varepsilon(t)$ from $y(t), y(t-1), \ldots$ as was done before we get

$$\varepsilon(t) = \sum_{k=t_0+1}^{t} y(k) + \varepsilon(t_0) = z(t) + \varepsilon(t_0).$$

The presence of the term $\varepsilon(t_0)$ whose influence does not vanish as $t_0 \to -\infty$ clearly shows the consequences of the equation

$$C(q^{-1})\ \varepsilon(t) = y(t)$$

being unstable. The initial condition $\varepsilon(t_0)$ can be estimated by

$$\hat{\varepsilon}(t_0) = -\frac{1}{t-t_0} \sum_{k=t_0+1}^{t} z(k) = -\frac{1}{t-t_0} \sum_{k=t_0+1}^{t} \sum_{i=t_0+1}^{k} y(i) =$$

$$= -\frac{1}{t-t_0} \sum_{k=t_0+1}^{t} (t+1-k)\ y(k).$$

This estimate will converge to $\varepsilon(t_0)$ as the number of terms in the series increases towards infinity. The predictor for y then becomes

$$\hat{y}(t+1|t) = -\sum_{k=t_0+1}^{t} \frac{k-1-t_0}{t-t_0}\, y(k) = -\sum_{i=1}^{t-t_0} \frac{t-t_0-i}{t-t_0}\, y(t+1-i). \qquad (3.9)$$

This predictor is clearly not a linear time-invariant system. Notice that the predictor (3.9) has a variance that approaches $E\varepsilon^2$ as $t-t_0 \to \infty$. A formal application of Theorem 3.1 gives the predictor

$$\hat{\hat{y}}(t+1|t) = -\sum_{-\infty}^{t} y(k),$$

which gives a prediction error with variance $2E\varepsilon^2$. □

The result of this example can be extended to the general case. See Hannan [5].

4. MINIMUM VARIANCE CONTROL

Having solved the prediction problem for the ARMA process we will now return to the CARMA process defined by equation (2.4) where A, B, and C are now regarded as matrix polynomials. To formulate the control problem it is necessary to define a criterion and the admissible controls.

The Criterion

It is assumed that the criterion for the control problem is to control the system in such a way that the steady state variance of the output is as small as possible. This criterion is a fairly good model for steady state control of important quality variables in industrial processes. The situation is illustrated in Fig. 2. Because of the fluctuations in the process output it is necessary to choose the reference value for the regulator above the test limit to make sure that a given percentage of the production is acceptable. By reducing the variance in the output it is then possible to operate closer to the test limit. This gives a gain which can be capitalized as increased production or reduction of raw materials used. For processes with a large volume of production even very moderate reductions in variance can give very substantial benefits.

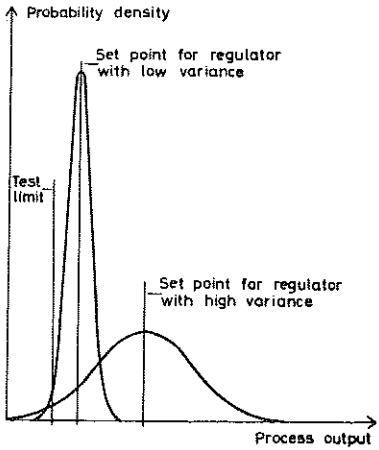

Fig. 2 - Illustrates that a decrease of the variance of the output signal makes it possible to move the set point closer to the test limit.

For single output systems the criterion will thus be taken as to minimize the following loss function

$$V_1 = E\, y^2(t). \tag{4.1}$$

It will be shown that the same results will be obtained for the loss-function

$$V_\infty = \lim_{N\to\infty} E \frac{1}{N} \sum_{t=1}^{N} y^2(t). \tag{4.2}$$

In the vector case the corresponding lossfunctions are

$$V_1 = E\, y^T(t)\, Q\, y(t) \tag{4.3}$$

and

$$V_\infty = \lim_{N\to\infty} \frac{1}{N} E \sum_{t=1}^{N} y^T(t)\, Q\, y(t) \tag{4.4}$$

respectively.

Admissible Controls

It is assumed that the admissible control laws are such that $u(t)$, i.e. the value of the control signal at time t, is a function of $y(t), y(t-1), \ldots$ and $u(t-1), u(t-2), \ldots$. By restricting the function to be linear the assumption on $\{\varepsilon(t)\}$ in the CARMA model can be relaxed from $\varepsilon(t)$ and $\varepsilon(s)$ being independent for $t \neq s$ to being uncorrelated.

The Minimum Variance Control Problem

The problem of controlling a CARMA process in such a way that the minimum variance criterion (4.1) or (4.3) is minimized will now be discussed. The solution is given by

THEOREM 4.1

Consider a CARMA process given by (2.4) where $\{\varepsilon(t)\}$ is a sequence of independent stochastic vectors with zero means and covariances R. Assume that the number of inputs and outputs are the same and that the polynomials $\det C(\xi)$ and $\det B(\xi)$ have all their zeros outside the unit disc. Let the matrix polynomials $F(q^{-1})$ of degree $k-1$ and $G(q^{-1})$ of degree $n-1$ be defined by

$$A^{-1}(q^{-1})\, C(q^{-1}) = F(q^{-1}) + q^{-k}\, A^{-1}(q^{-1})\, G(q^{-1}). \tag{4.5}$$

Then the control law

$$u(t) = -B^{-1}(q^{-1})\, G(q^{-1})\, F^{-1}(q^{-1})\, y(t) = -B^{-1}(q^{-1})\, G(q^{-1})\, \varepsilon(t) \tag{4.6}$$

minimizes the criterion (4.3) in the steady state and the steady output of the controlled system becomes

$$y(t) = F(q^{-1})\, \varepsilon(t) = \varepsilon(t) + F_1\varepsilon(t-1) + \ldots + F_{k-1}\varepsilon(t-k+1). \tag{4.7}$$

Proof:

A change of the control signal at time t will be noticeable in the output at first at time $t+k$. Because the matrix B_0 was assumed regular it is also possible to change all components of the output at time $t+k$ arbitrarily. It follows from (2.4) and (4.5) that

$$y(t+k) = F(q^{-1})\ \varepsilon(t+k) + A^{-1}(q^{-1})\ [B(q^{-1})\ u(t) + G(q^{-1})\ \varepsilon(t)].$$

For simplicity the polynomial $A(q^{-1})$ will now simply be written as A. Using (2.4) to eliminate ε in the last term we get

$$y(t+k) = F\ \varepsilon(t+k) + A^{-1}Bu(t) + A^{-1}GC^{-1}Ay(t) - A^{-1}GC^{-1}Bu(t-k) =$$

$$= F\ \varepsilon(t+k) + FC^{-1}Bu(t) + A^{-1}GC^{-1}Ay(t), \qquad (4.8)$$

where the equality is obtained by applying (4.5) to the terms containing $u(t)$. To proceed notice that

$$G(AF)^{-1}\ C = C(AF)^{-1}\ G, \qquad (4.9)$$

because it follows from (4.5) that

$$G(AF)^{-1}\ C = q^k(C-AF)\,(AF)^{-1}C = q^k[C(AF)^{-1}C - C]$$

$$C(AF)^{-1}\ G = q^kC(AF)^{-1}(C-AF) = q^k[C(AF)^{-1}C - C].$$

Equations (4.8) and (4.9) give

$$y(t+k) = F(q^{-1})\ \varepsilon(t+k) + F(q^{-1})C^{-1}(q^{-1})\,[B(q^{-1})u(t) + G(q^{-1})F^{-1}(q^{-1})y(t)].$$

The two terms of the right member are independent because of the definition of admissible strategies, because the polynomial $\det C(q^{-1})$ is stable too, and because of $\varepsilon(t+k)$ being independent of $y(t), y(t-1), \ldots$ for $k > 0$. It thus follows that

$$E\ y^T(t+k)Q\ y(t+k) \geq E\,[F(q^{-1})\varepsilon(t+k)]^TQ[F(q^{-1})\varepsilon(t+k)] =$$

$$= \operatorname{tr}\,[Q + F_1^T Q F_1 + \ldots + F_{k-1}^T Q F_{k-1}]\ R,$$

where equality is obtained for

$$B(q^{-1})\ u(t) + G(q^{-1})\ \varepsilon(t) = 0.$$

Then also

$$y(t) = F(q^{-1})\ \varepsilon(t).$$

A combination of these equations gives the control law (4.6). To see

the transient behaviour of the system introduce the control law (4.6) into the system description (2.4). Hence

$$[A(q^{-1}) + q^{-k}G(q^{-1})F^{-1}(q^{-1})]\ y(t) = C(q^{-1})\ \varepsilon(t).$$

Equation (4.5) gives

$$C(q^{-1})\ [F^{-1}(q^{-1})y(t) - \varepsilon(t)] = 0. \qquad (4.10)$$

Since the polynomial $C(\xi)$ was assumed to have all its zeros outside the unit disc, this implies that the expression in brackets will converge to zero exponentially at a rate governed by the zeros of $\det C(\xi)$.

Remark 1
The theorem still holds if $\varepsilon(t)$ and $\varepsilon(s)$ are only assumed uncorrelated for $t \neq s$ if a linear control law is postulated.

Remark 2
A comparison with the solution of the prediction problem shows that the control error under minimum variance control equals the error in predicting the process k steps. The minimum variance control law can thus be interpreted as doing the following. Predict output k steps ahead where k is the time it takes before a control action is noticeable in the output. Choose a control signal which makes the predicted value equal to the desired output.

Remark 3
The control error is a moving average of order k. This is easy to test and useful for diagnosis.

Remark 4
It follows from (4.10) that the poles of the closed loop system are given by

$$\det C(z^{-1}) = 0.$$

Remark 5
Notice that the control law (4.6) does not depend on the matrix Q. The control law will thus simultaneously minimize the variances in all components of the output. This motivates the name minimum variance strategy.

Control Effectiveness

Without control the output becomes

$$y_0(t) = A^{-1}(q^{-1})\ C(q^{-1})\ \varepsilon(t)$$

and under minimum variance control the output becomes

$$y_{mv}(t) = F(q^{-1})\ \varepsilon(t).$$

Hence

$$y_{mv}(t) = F(q^{-1})\ C^{-1}(q^{-1})\ A(q^{-1})\ y_0(t).$$

The reduction of the fluctuations in the output can thus be characterized by the transfer function

$$H(q^{-1}) = F(q^{-1})\ C^{-1}(q^{-1})\ A(q^{-1}).$$

A simple example illustrates what can happen.

EXAMPLE

Consider a first order scalar system with $k = 1$ and

$$A(q^{-1}) = 1 + a\,q^{-1}$$

$$C(q^{-1}) = 1 + c\,q^{-1}.$$

Hence

$$H(q^{-1}) = \frac{1 + aq^{-1}}{1 + cq^{-1}}$$

$$|H(e^{-i\omega})| = \frac{1 + a^2 + 2a\cos\omega}{1 + c^2 + 2c\cos\omega}.$$

A graph of the function $|H|$ is shown in Fig. 3. The graph shows that in the particular case the action of the minimum variance is to reduce the low-frequency components and to increase the high frequency components in the output.

The Minimum Phase Assumption

In Theorem 4.2 it was assumed that the polynomial $\det B(\xi)$ has all its zeros outside the unit disc. This assumption is called the minimum phase condition because it implies that the input-output relation given by (2.4) for $\varepsilon = 0$ is a nonminimum phase system. If this condition is violated the control law given by (4.6) still gives the

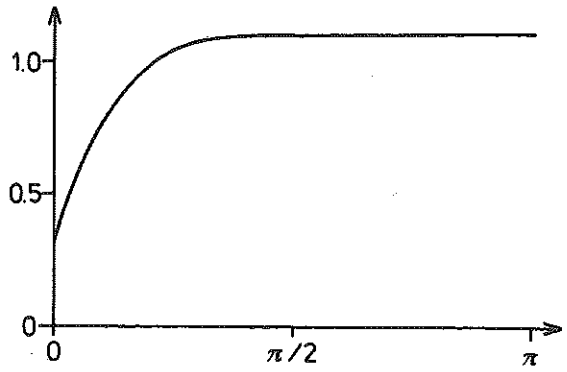

Fig. 3 - Amplitude curve of the transfer function H which shows how the minimum variance control law attenuates disturbances of different frequencies.

smallest variance of the outputs. It follows, however, from (4.6) that the control signal u is given by

$$u(t) = -B^{-1}(q^{-1})\ G(q^{-1})\ \varepsilon(t).$$

If the polynomial $\det B(\xi)$ has zeros inside the unit disc this difference equation will be unstable and the control signal u will grow exponentially. This will not have any influence on the output y because the exponential components of u will be cancelled by the operator $B(q^{-1})$ which operates on y in the system model. The cancellation will of course only be possible if the control law was calculated from a precise model of the system. Small perturbations in the model implies that the exponentially growing components will be transmitted to the output. This is illustrated in Fig. 4 which shows the results of a simulation.

From a practical point of view it is thus clear that the control law (4.6) is useless if the polynomial $\det B(\xi)$ has zeros inside the unit disc. There are several different possibilities to circumvent this problem. One possibility is to include a penalty on the control actions, i.e. to change the criterion to

$$E[y^2 + \rho\ u^2].$$

Control laws with the property that $u(t)$ will be very large will then be excluded. In several cases it may, however, be unrealistic to assign a proper value of ρ and we will therefore investigate the problem with $\rho = 0$. It turns out in fact that the problem of minimizing (4.1) has several local minima if the polynomial $\det B(\xi)$ has

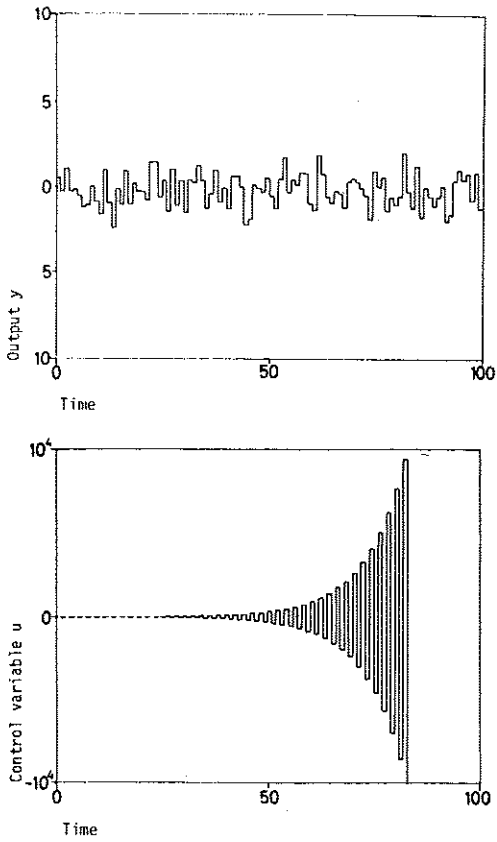

Fig. 4 - Simulation of a nonminimum phase system with minimum variance control. The system is described by $y(t) - 1.7\,y(t-1) + 0.7\,y(t-2) = 0.9\,u(t-1) + u(t-2) + e(t) - 0.7\,e(t-1)$.

zeros inside the unit disc. Before stating the main result the notion of reciprocal polynomial is introduced. Let $B(\xi)$ be a polynomial

$$B(\xi) = b_0 + b_1\xi + \ldots + b_n\xi^n \ ;$$

then the reciprocal polynomial is defined by

$$\tilde{B}(\xi) = b_0\xi^n + b_1\xi^{n-1} + \ldots + b_n.$$

In the single-input single-output case we have

THEOREM 4.2

Consider a single-input single-output system described by (2.4). Assume that the polynomial $C(\xi)$ has all its zeros outside the unit disc and that the polynomial $B(\xi)$ has zeros both inside and outside the unit disc. Let B be factored as

$$B(q^{-1}) = B_1(q^{-1})\, B_2(q^{-1}), \tag{4.11}$$

where $B_2(\xi)$ has all its zeros inside the unit disc and $B_2(0) = 1$. Let $H(q^{-1})$ and $K(q^{-1})$ be defined by the partial fraction expansion

$$\frac{C(q^{-1})\, \tilde{B}_2(q^{-1})}{A(q^{-1})\, q^{-k} B_2(q^{-1})} = \frac{H(q^{-1})}{q^{-k} B_2(q^{-1})} + \frac{K(q^{-1})}{A(q^{-1})}, \tag{4.12}$$

where

$$\deg H(q^{-1}) = k - 1 + \deg B_2(q^{-1}). \tag{4.13}$$

Then the variance of the output has a local minimum for the control law

$$u(t) = -\frac{K(q^{-1})}{H(q^{-1}) B_1(q^{-1})}\, y(t) = -\frac{K(q^{-1})}{B_1(q^{-1}) \tilde{B}_2(q^{-1})}\, \varepsilon(t) \tag{4.14}$$

and the corresponding output is given by

$$y(t) = \frac{H(q^{-1})}{\tilde{B}_2(q^{-1})}\, \varepsilon(t). \tag{4.15}$$

Proof:

Equation (2.4) gives

$$y(t+k) = \frac{B}{A}\, u(t) + \frac{C}{A}\, \varepsilon(t+k) = \frac{B_2}{\tilde{B}_2}\, w(t+k),$$

where

$$w(t+k) = \frac{B_1 \tilde{B}_2}{A}\, u(t) + \frac{C\, \tilde{B}_2}{q^{-k} A B_2}\, \varepsilon(t).$$

The signals y and w have the same variances because B_2 and $\tilde{B}_2$ are reciprocal polynomials. Equation (4.12) then gives

$$w(t+k) = \frac{B_1 \tilde{B}_2}{A}\, u(t) + \frac{H}{q^{-k} B_2}\, \varepsilon(t) + \frac{K}{A}\, \varepsilon(t).$$

Now use (2.4) to eliminate $\varepsilon(t)$ in the last term, then

$$w(t+k) = \frac{H}{q^{-k}B_2}\,\varepsilon(t) + \frac{B_1\tilde{B}_2}{A}\,u(t) + \frac{K}{A}\left[\frac{A}{C}\,y(t) - \frac{B}{C}\,u(t-k)\right] =$$

$$= \frac{H}{q^{-k}B_2}\,\varepsilon(t) + \left[\frac{B_1\tilde{B}_2}{A} - \frac{q^{-k}KB}{AC}\right] u(t) + \frac{K}{C}\,y(t) =$$

$$= \frac{H}{q^{-k}B_2}\,\varepsilon(t) + \frac{q^{-k}B_1B_2}{C}\left[\frac{C\tilde{B}_2}{Aq^{-k}B_2} - \frac{K}{A}\right] u(t) + \frac{K}{C}\,y(t) =$$

$$= \frac{H}{q^{-k}B_2}\,\varepsilon(t) + \frac{B_1H}{C}\,u(t) + \frac{K}{C}\,y(t),$$

where the last equality follows from (4.12). Hence

$$w(t+k) = \frac{H(q^{-1})}{B_2(q^{-1})}\,\varepsilon(t+k) + \left[\frac{B_1(q^{-1})H(q^{-1})}{C(q^{-1})}\,u(t) + \frac{K(q^{-1})}{C(q^{-1})}\,y(t)\right]. \qquad (4.16)$$

Because the polynomial $B_2(\xi)$ has all its zeros inside the unit disc and because of (4.13) the first term of the right member of (4.16) can be written as the converging series

$$\frac{H(q^{-1})}{q^{-k}B_2(q^{-1})}\,\varepsilon(t) = \varepsilon(t+1) + \alpha_1\varepsilon(t+2) + \ldots \quad .$$

Since the polynomial $C(\xi)$ has all its zeros outside the unit disc, the second term can be expanded as a converging series in $y(t), y(t-1),\ldots\ u(t), u(t-1),\ldots$ The two terms of (4.16) are thus independent and the smallest variance is obtained for the control law (4.14). The output of the controlled system is then given by (4.15).

Remark

The control signal u defined by (4.14) is bounded if $B_2(\xi)$ is chosen as the factor which contains all zeros of B inside the unit disc.

Consider a system described by (2.4) where the polynomial $B(\xi)$ has zeros inside the unit disc. According to Theorem 4.1 there is an absolute minimum to the variance of the output given by

$$E[F(q^{-1})\varepsilon(t)]^2. \qquad (4.16)$$

The control law which realizes this minimum is given by

$$u(t) = -\frac{G(q^{-1})}{B(q^{-1})F(q^{-1})}\,y(t) = -\frac{G(q^{-1})}{B(q^{-1})}\,\varepsilon(t).$$

The variance of the control signal will clearly be infinite because $B(q^{-1})$ is unstable. Let B_2 be an unstable factor of B. According to Theorem 4.2 there is then another local minimum of the loss function given by

$$E\left[\frac{H(q^{-1})}{q^{-k}\,\tilde{B}_2(q^{-1})}\,\varepsilon(t)\right]^2. \tag{4.17}$$

It is easy to show that

$$\frac{H(q^{-1})}{q^{-k}\,\tilde{B}_2(q^{-1})}\,\varepsilon(t) = F(q^{-1})\,\varepsilon(t+k) + \frac{L_1(q^{-1})}{\tilde{B}_2(q^{-1})}\,\varepsilon(t),$$

where L_1 is defined by the partial fraction expansion

$$\frac{G(q^{-1})B_2(q^{-1})}{A(q^{-1})B_2(q^{-1})} = \frac{L_1(q^{-1})}{B_2(q^{-1})} + \frac{L_2(q^{-1})}{A(q^{-1})} \qquad \deg L_1 < \deg B_2.$$

The loss function (4.17) is thus always larger than (4.16) and the term

$$E\left[\frac{L_1(q^{-1})}{\tilde{B}_2(q^{-1})}\,\varepsilon(t)\right]^2$$

represents the increase in the loss function required in order to avoid having the factor $B_2(q^{-1})$ in the equation (4.14).

When solving the minimum variance problem B_2 should be chosen to contain all factors of B which have zeros inside the unit disc except those factors which are also factors of A.

5. APPLICATIONS

The minimum variance control theory tells that the feedback law of Fig. 1 is a time invariant dynamical system. The theory also gives a possibility to interpret the action of the feedback law. In the simplest case action of the feedback can be described as follows: Predict the output k steps ahead. Choose a control signal such that the predicted value is equal to the desired value. The complexity of the feedback law is uniquely determined by the mathematical model of the process. For simple models the regulator may be equivalent to the common PI regulator (1) but the regulator may also be much more complicated.

The minimum variance regulator has been applied to a number of different industrial control problems. The problems have typically been steady state regulation of industrial processes. The benefits obtained have been determined from the arguments illustrated in Fig. 2. The successful applications have the property that moderate reductions of the variances will give rise to substantial economic gains. It is thus motivated to spend the extra effort required to develop the models and to obtain the control laws. The determination of the mathematical model (2.4) is the major difficulty when trying to apply minimum variance control. The model (2.4) can rarely be obtained from apriori physical data. Instead the model has to be estimated from data obtained from an experiment on the process. In a typical experiment the input signal is perturbed and the resulting variations in the output are recorded. When the model is obtained it is possible to tell the results that can be expected from minimum variance control. This is a substantial advantage because it can then be decided if the effort is worthwhile and if it is justified to use a control law which is more complicated than the simple PID regulator.

6. REFERENCES

The scalar version of the minimum variance control law is discussed in [1]. This reference also contains many references. The solution to the nonminimum phase problem was given in [2]. The spectral factorization idea is given in [3]. An application of the minimum variance strategy to paper machine control is described in [2]. Applications to metal rolling are described in [4].

[1] K J Åström: Introduction to Stochastic Control Theory. Academic Press, New York 1970.

[2] V Peterka: On Steady State Minimum Variance Control Strategy. Kybernetika 8 (1972) 219-232.

[3] N Wiener: Extrapolation, Interpolation and Smoothing of Stationary Time Series with Engineering Applications. Wiley, New York 1949.

[4] D Watanapongse and N A Robbins: Application of Modern Control Techniques to Computerized Setup for Effective Operation of Inland Strip Mills. Report, Inland Steel Company, East Chicago, Indiana, USA, 1976.

[5] E J Hannan: Multiple Time Series. Wiley, New York 1970.

CHAPTER 3 - LINEAR QUADRATIC GAUSSIAN CONTROL

1. INTRODUCTION

The theory described in this chapter is called linear quadratic gaussian (LQG) control theory because the process dynamics is characterized by linear equations, the criterion is a quadratic function, and the disturbances are gaussian. In the previous chapter the process dynamics was also described by an external model. This model was a difference equation which related the process output to its input and the disturbances. In this chapter the process model will instead be described by an internal model. A set of variables which completely specifies the past development of the system called *state variables* are thus introduced. The mathematical model is then a difference equation which describes the future development of the state variables. The model is still assumed to be linear. The criterion is the expected value of a general quadratic form. The problem is thus slightly more general than the problem discussed in the previous chapter. The development of the theory is also analogous to that of the previous chapter. The mathematical models used are discussed in Section 2. The prediction problem is then solved in Section 3 and the control problem is solved in Section 4.

When external models are used it was natural to use the theory of polynomials and rational functions. For the internal models it is instead natural to use matrix theory. The relations beween the two approaches to the problem are discussed in Section 5 and Section 6 deals briefly with applications. Since the LQG theory is covered in detail in text books, the treatment is here kept fairly brief.

2. MATHEMATICAL MODELS

Process Dynamics and Disturbances

For simplicity only discrete time systems will be considered. It is assumed that time T is the set of integers $\{\ldots -1, 0, 1, \ldots\}$. Let the input u, the state x, and the output y be vector valued time functions of dimensions p, n, and r. It is assumed that the system and its environment can be described by the linear difference equations

$$\begin{aligned} x(t+1) &= Ax(t) + Bu(t) + v(t) \\ y(t) &= Cx(t) + e(t), \qquad t \in T, \end{aligned} \tag{2.1}$$

where the "process noise" $\{v(t), t \in T\}$ and the "measurement noise" $\{e(t), t \in T\}$ are sequences of independent gaussian random vectors. It is assumed that $\{v(t)\}$ and $\{e(t)\}$ have zero mean values and that their covariances are given by

$$\operatorname{cov}\,[v(t), v(s)] = \begin{cases} R_1 & t = s \\ 0 & t \neq s \end{cases}$$

$$\operatorname{cov}\; v(t), e(s) = 0$$

$$\operatorname{cov}\; e(t), e(s) = \begin{cases} R_2 & t = s \\ 0 & t \neq s\,. \end{cases} \qquad (2.2)$$

The initial condition $x(t_0)$ of (2.1) is assumed to be gaussian with mean value m and covariance R_0. It is also assumed that the initial condition is independent of $\{v(t)\}$ and $\{e(t)\}$ or equivalently that $x(t)$ is independent of $\{v(t)\}$ and $\{e(t)\}$.

The model (2.1) tells that given the values of the state x and the control u at time t then the conditional distribution of the state at time $t+1$ is gaussian with mean value $Ax(t) + Bu(t)$ and covariance R_1. The equation (2.1) also tells that the conditional distribution of the measurement $y(t)$ given $x(t)$ is gaussian with mean $Cx(t)$ and covariance R_2.

Notice that it is frequently necessary to introduce extra variables in order to arrive at a model having the form (2.1). For example if the environment is characterized by a disturbance having the spectral density

$$\phi(\omega) = \frac{1}{1 + a^2 - 2a \cos \omega}$$

it can be characterized by the difference equation

$$\xi(t+1) = a\xi(t) + n(t),$$

where $\{n(t)\}$ is white noise. It is then necessary to include ξ as a component of the state vector. Similarly a constant disturbance acting on the system can be described by the difference equation

$$d(t+1) = d(t).$$

Such a disturbance can be included by augmenting the state vector.

The Criterion

In the linear quadratic gaussian problem it is assumed that the purpose of the control can be expressed as to minimize the loss function

$$V_1 = \min E \left\{ x^T(N)Q_0x(N) + \sum_{t=t_0}^{N-1} x^T(t)Q_1x(t) + u^T(t)Q_2u(t) \right\}. \qquad (2.3)$$

Time Varying Models

The matrices A, B, C, R_1, R_2, Q_1, and Q_2 may vary with time t.

3. KALMAN FILTERING AND PREDICTION

The filtering problem will be solved before the optimal control problem is discussed. It is assumed that the outputs $y(t_0),\dots,y(t)$ have been observed and the problem is to predict $x(t+1)$ as well as possible. Let $\mathcal{Y}_t$ denote the σ-algebra generated by $y(t),\dots,y(t_0)$. The prediction problem is clearly solved if the conditional distribution of $x(t+1)$ given $\mathcal{Y}_t$ can be determined. The solution is given by the following theorem.

THEOREM 3.1
Let the gaussian process $\{x(t)\}$ be generated by (2.1) with $u = 0$. The conditional distribution of $\{x(t+1)\}$ given $\mathcal{Y}_t$ is gaussian $(\hat{x}(t+1), P(t+1))$ where

$$\begin{cases} \hat{x}(k+1) = A\hat{x}(k) + K(k)[y(k) - C\hat{x}(k)], & k = t_0,\dots,t \\ \hat{x}(t_0) = m \end{cases} \qquad (3.1)$$

$$K(k) = AP(k)C^T[CP(k)C^T + R_2]^{-1} \qquad (3.2)$$

$$\begin{cases} P(k+1) = AP(k)A^T + R_1 - AP(k)C^T[CP(k)C^T + R_2]^{-1} CP(k)A^T = \\ \qquad = [A - K(k)C]\, P(k)A^T + R_1, \qquad k = t_0,\dots,t \\ P(t_0) = R_0. \end{cases} \qquad (3.3)$$

Proof:
The proof consists of a repeated use of the following well known property of gaussian random variables. If the vector

$$\begin{pmatrix} x \\ y \end{pmatrix}$$

is gaussian with mean value

$$\begin{pmatrix} m_x \\ m_y \end{pmatrix}$$

and covariance

$$\begin{pmatrix} R_x & R_{xy} \\ R_{yx} & R_y \end{pmatrix},$$

then the conditional mean of x given y is

$$E[x|y] = m_x + R_{xy} R_y^{-1}(y-m_y).$$

Full details are given in the references.

□

Remark 1

The theorem has a strong intuitive appeal. The term $A\hat{x}(k)$ is the apriori estimate of $x(k+1)$ and the correction to the prior $K(k)[y(k) - C\hat{x}(k)]$ is proportional to the deviation of the measurement $y(k)$ from its prior $C\hat{x}(k)$.

Remark 2

The covariance $P(k)$ does not depend on the measurements.

Remark 3

The result of the theorem can easily be extended to include a control signal different from zero in (2.1). If $u(t)$ is measurable with respect to y_t for each t then it is easily shown that the conditional distribution of $x(t+1)$ given y_t is gaussian $(\hat{x}(t+1), P(t))$ where

$$\hat{x}(t+1) = A\hat{x}(t) + Bu(t) + K(t)[y(t) - C\hat{x}(t)] \qquad (3.4)$$

and $K(t)$ and $P(t)$ are given by (3.2) and (3.3).

Remark 4

The theorem can be extended to the case when the random processes $\{v(t)\}$ and $\{e(t)\}$ are assumed to be second order processes only. The best *linear* prediction is then given by $\hat{x}(t+1)$.

Innovations Representations

Theorem 3.1 allows for an alternative representation of the stochastic process $\{y(t)\}$. It follows from the proof of Theorem 3.1 that the variables

$$\tilde{y}(t) = y(t) - C\hat{x}(t) \qquad (3.4)$$

are gaussian random variables with zero mean values and the covariances

$$E\,\tilde{y}(t)\,\tilde{y}^T(s) = \begin{cases} R = [CP(t)C^T + R_2] & t = s \\ 0 & t \neq s\,. \end{cases} \tag{3.5}$$

Since $\tilde{y}(t)$ is gaussian it then follows that $\{\tilde{y}(t), t \in T\}$ is a sequence of independent gaussian random variables. The following theorem is then obtained.

THEOREM 3.2
Consider the stochastic process $\{y(t)\}$ defined by (2.1) where $u(t)$ is measurable with respect to y_t. The process $\{y(t)\}$ then has the representation

$$\begin{cases} \hat{x}(t+1) = A\hat{x}(t) + Bu(t) + K(t)\,\tilde{y}(t) \\ y(t) = C\hat{x}(t) + \tilde{y}(t), \end{cases} \tag{3.6}$$

where $\{\tilde{y}(t)\}$ is a sequence of independent gaussian $(0, R)$ random variables where $K(t)$ is given by (3.2) and R by (3.5).

Duality

Let x and y be gaussian random vectors. The space obtained by introducing the scalar product

$$\langle x, y \rangle = E\,x^T y$$

can be shown to be the dual of a Euclidean space. By using this concept of duality it can be shown that the Kalman filtering problem is the dual of a deterministic control problem.

To see this consider the problem of estimating $a^T x(t_1)$ linearly in $y(t_1-1), \ldots, y(t_0)$ and m in such a way that the criterion

$$E[a^T x(t_1) - a^T\hat{x}(t_1)]^2 \tag{3.7}$$

is minimal.

As the estimate is linear we have

$$a^T\hat{x}(t_1) = -\sum_{t=t_0}^{t_1-1} u^T(t)\,y(t) + b^T m\,. \tag{3.8}$$

The minus sign is introduced in order to obtain the final result in a nice form. The estimation problem is thus a problem of determining the vectors $b, u(t_1-1), u(t_1-2), \ldots, u(t_0)$. Now determine the u:s in such a way that the criterion (3.7) is minimal. To do so, introduce the vectors $z(t)$ defined recursively from

$$z(t) = A^T z(t+1) + C^T u(t+1) \tag{3.9}$$

with the initial condition

$$z(t_1-1) = a.$$

Hence

$$a^T x(t_1) = z^T(t_1-1)x(t_1) = z^T(t_0-1)x(t_0) + \sum_{t=t_0}^{t_1-1} [z^T(t)x(t+1) - z^T(t-1)x(t)]. \tag{3.10}$$

It follows from (2.1) and (3.9) that

$$z^T(t)x(t+1) = z^T(t)Ax(t) + z^T(t)v(t)$$

$$z^T(t-1)x(t) = z^T(t)Ax(t) + u^T(t)Cx(t).$$

Introducing this in (3.10), we find

$$a^T x(t_1) = z^T(t_0-1)x(t_0) + \sum_{t=t_0}^{t_1-1} [z^T(t)v(t) - u^T(t)Cx(t)]. \tag{3.11}$$

Equations (2.1) and (3.8) give

$$a^T \hat{x}(t_1) = -\sum_{t=t_0}^{t_1-1} u^T(t)y(t) + b^T m = -\sum_{t=t_0}^{t_1-1} [u^T(t)Cx(t) + u^T(t)e(t)] + b^T m. \tag{3.12}$$

Hence

$$a^T x(t_1) - a^T \hat{x}(t_1) = z^T(t_0-1)x(t_0) - b^T m + \sum_{t=t_0}^{t_1-1} [z^T(t)v(t) - u^T(t)e(t)].$$

Squaring and taking mathematical expectations, the criterion (3.7) can be expressed as follows:

$$E[a^T x(t_1) - a^T \hat{x}(t_1)]^2 = [(z(t_0-1) - b)^T m]^2 + z^T(t_0-1)R_0 z(t_0-1) + \sum_{t=t_0}^{t_1-1} [z^T(t)R_1 z(t) + u^T(t)R_2 u(t)]. \tag{3.13}$$

To minimize the criterion, the parameter b must be chosen equal to $z(t_0-1)$ and the u:s should be determined in such a way that the function

$$z^T(t_0-1)R_0 z(t_0-1) + \sum_{t=t_0}^{t_1-1} [z^T(t)R_1 z(t) + u^T(t)R_2 u(t)] \tag{3.14}$$

is as small as possible.

It has now been shown that the problem of finding a linear predictor which minimizes (3.7) is equivalent to finding a control signal u for the system (3.9) such that the criterion (3.14) is minimal.

4. OPTIMAL CONTROL

Having solved the prediction problem we will now return to the optimal control problem. A system described by (2.1) is considered. The problem is to find an admissible control such that the criterion (2.3) is minimal. The following result is useful in the solution of the problem.

LEMMA 4.1
Consider a system described by the difference equation

$$x(t+1) = Ax(t) + Bu(t) + v(t). \tag{4.1}$$

Assume that the difference equation

$$S(t) = A^T S(t+1)A + Q_1 - A^T S(t+1)B[Q_2 + B^T S(t+1)B]^{-1} B^T S(t+1)A \tag{4.2}$$

with the initial condition

$$S(N) = Q_0 \tag{4.3}$$

has a solution $S(t)$ which is non-negative definite for $t_0 \le t \le N$ and such that

$$Q(t) = Q_2 + B^T S(t+1)B \tag{4.4}$$

is non-singular for all t. Let

$$L(t) = [Q_2 + B^T S(t+1)B]^{-1} B^T S(t+1)A. \tag{4.5}$$

Then

$$\begin{aligned} x^T(N)Q_0x(N) + \sum_{t=t_0}^{N-1} x^T(t)Q_1x(t) + u^T(t)Q_2u(t) &= x^T(t_0)S(t_0)x(t_0) + \\ &+ \sum_{t=t_0}^{N-1} [u(t) + L(t)x(t)]^T[B^TS(t+1)B + Q_2][u(t) + L(t)x(t)] + \\ &+ \sum_{t=t_0}^{N-1} \Big\{ v^T(t)S(t+1)[Ax(t) + Bu(t)] + \\ &+ [Ax(t) + Bu(t)]^T S(t+1)v(t) + v^T(t)S(t+1)v(t) \Big\}. \end{aligned} \tag{4.6}$$

Proof:
The proof is straightforward. We have the following

$$x^T(N)Q_0x(N) = x^T(N)S(N)x(N) = x^T(t_0)S(t_0)x(t_0) +$$
$$+ \sum_{t=t_0}^{N-1} [x^T(t+1)S(t+1)x(t+1) - x^T(t)S(t)x(t)].$$

Consider the different terms of the sum. We have

$$x^T(t+1)S(t+1)x(t+1) = [Ax(t) + Bu(t) + v(t)]^T S(t+1)[Ax(t) + Bu(t) + v(t)]$$

and

$$x^T(t)S(t)x(t) = x^T(t)\{A^TS(t+1)A + Q_1 - L^T(t)[B^TS(t+1)B + Q_2]L(t)\}x(t).$$

Hence

$$x^T(N)Q_0x(N) = x^T(t_0)S(t_0)x(t_0) + \sum_{t=t_0}^{N-1} \{[Ax(t) + Bu(t)]^T S(t+1)v(t) +$$
$$+ v^T(t)S(t+1)[Ax(t) + Bu(t)] + v^T(t)S(t+1)v(t)\} +$$
$$+ \sum_{t=t_0}^{N-1} \{u^T(t)[B^TS(t+1)B + Q_2]u(t) +$$
$$+ u^T(t)B^TS(t+1)Ax(t) + x^T(t)A^TS(t+1)Bu(t) +$$
$$+ x^T(t)L^T(t)[B^TS(t+1)B + Q_2]L(t)x(t) - x^T(t)Q_1x(t) -$$
$$- u^T(t)Q_2u(t)\},$$

where the term u^TQ_2u has been added and subtracted in the last sum. Rearrangement of the terms now completes the proof of the lemma. □

The Lemma 4.1 is a useful tool for solving the optimal control problem because it shows directly how the loss function is influenced by the value of the control signal at time t. The optimal control problem will now be solved for some different choices of the admissible controls.

Complete State Information

It is first assumed that the admissible controls are such that $u(t)$ is a function of $x(t)$. The solution to the optimal control problem is then given by

THEOREM 4.1

Consider a system described by (2.1). Let the admissible controls be such that $u(t)$ is a function of $x(t)$. Assume that the equation (4.2) with initial conditions (4.3) has a non-negative solution such that $Q(t)$ defined by (4.4) is positive definite for $t_0 \leq t \leq N$. Then the criterion (2.3) is minimal for the control law

$$u(t) = - L(t)x(t), \tag{4.7}$$

where L is given by (4.5). The minimal loss is

$$\min V = m^T S(t_0) m + \operatorname{tr} S(t_0) R_0 + \sum_{t=t_0}^{N-1} \operatorname{tr} S(t+1) R_1(t). \tag{4.8}$$

Proof:

Let x be gaussian (m,R). Then

$$Ex^T Q x = m^T Q m + E(x-m)^T Q (x-m) = m^T Q m + E \operatorname{tr}(x-m)^T Q (x-m) =$$

$$= m^T Q m + E \operatorname{tr} Q(x-m)(x-m)^T = m^T Q m + \operatorname{tr} QR .$$

It follows from Lemma 4.1 that

$$E\left[x^T(N) Q_0 x(N) + \sum_{t=t_0}^{N-1} x^T(t) Q_1 x(t) + u^T(t) Q_2 u(t) \right] =$$

$$= m^T S(t_0) m + \operatorname{tr} S(t_0) R_0 + \sum_{t=t_0}^{N-1} \operatorname{tr} S(t+1) R_1(t) +$$

$$+ \Sigma \, [u(t) + L(t)x(t)]^T Q(t) [u(t) + L(t)x(t)], \tag{4.9}$$

because $v(t)$ is independent of $x(t)$ and $u(t)$.

Since $Q(t)$ was assumed to be positive definite the right hand side is minimal for the control law (4.7) and the proof is completed. □

Incomplete State Information

The admissible controls are now assumed to be such that $u(t+1)$ is a function of $y(t),\ldots,y(t_0)$ or more precisely for each t $u(t+1)$ is assumed to be measureable with respect to the σ-algebra Y_t generated by $y(t),\ldots,y(t_0)$. To obtain the result in this case a measure selection theorem will be used.

Let x and y be random variables which take values in R^n and R^p. Let $\ell(x,y,u)$ be a loss function ℓ: $R^{n+p+r} \to R$. We have

LEMMA 4.2
Let $E[\cdot|y]$ denote the conditional mean given y. Assume that

$$f(y,u) = E[\ell(x,y,u)|y]$$

has a unique minimum with respect to u, attained for $u = u^0(y)$. Then the minimum of $E\ell(x,y,u(y))$ with respect to all $u(y)$ which are measurable with respect to y is given by

$$\min E\ell(x,y,u(y)) = E_y f(y,u^0(y)).$$

Proof:
The proof is given in Åström [2], p. 261. □

The solution to the control problem with incomplete state information is now given by

THEOREM 4.2 (Separation Theorem)
Consider a system described by (2.1). Let the admissible controls be such that for each t $u(t+1)$ is measurable with respect to Y_t. Assume that the equation (4.2) with initial conditions (4.3) has a non-negative solution such that $Q(t)$ given by (4.4) is positive definite for $t_0 \le t \le N$. Then the criterion (2.3) is minimized for the control law

$$u(t) = -L(t)\hat{x}(t) \tag{4.10}$$

where L is given by (4.5) and $\hat{x}(t)$ is the conditional mean of $x(t)$ given Y_{t-1}, given by the Kalman filter (3.4). The minimal loss is given by

$$\min E\, V = m^T S(t_0) m + \operatorname{tr} S(t_0) R_0 + \sum_{t=t_0}^{N-1} \operatorname{tr} S(t+1) R_1(t) + \sum_{t=t_0}^{N-1} \operatorname{tr} P(t) L^T(t) Q(t) L(t). \tag{4.11}$$

Proof:
It follows from Remark 3 of Theorem 3.1 that the conditional distribution of $x(t+1)$ given Y_t is given by (3.4) where the conditional covariance does not depend on the control signal.

Proceed in the same way as for the proof of Theorem 4.1 to obtain equation (4.9). Use Lemma 4.2 to minimize the right hand side of (4.9). The minimum (4.11) is then obtained for the control law (4.10).

□

Remark 1

Notice that the different terms in the minimal loss function all have a nice physical interpretation. The term $m^T S(t_0) m$ is the contribution due to the off set of the initial state. The term $\operatorname{tr} S(t_0) R_0$ is the contribution due to the initial uncertainty of the initial state. The term $\Sigma \operatorname{tr} S(t+1) R_1(t)$ depends on the process noise that is acting on the system and the term $\Sigma \operatorname{tr} P(t) L^T(t) Q(t) L(t)$ depends on the uncertainty in the state estimation. A calculation of the relative magnitudes of the different terms will give good information about the nature of the difficulties in solving the control problem.

Remark 2

Notice that the optimal control law is a linear feedback from the conditional mean. The linear feedback gain L is the same as for the problem of Theorem 4.1 with complete state information. This motivates the name *certainty equivalence theorem* which is sometimes given to Theorem 4.2.

Theorem 4.2 gives valuable insight into the nature of the optimal feedback. The feedback can be thought of as being composed of two parts. See Fig. 5. One part is a dynamical system (a Kalman filter) which generates the conditional mean of the state vector from the measured process outputs. The other part is a static linear system which simply generates the control as a linear function (4.10) of the estimated state variables. See Fig. 5. Notice that the matrix $L(t)$ in (4.10) only depends on A, B, Q_1, and Q_0 and that it is independent of the stochastic elements of the model. The gain K of the Kalman filter depends on A, C, R_0, R_1, and R_2 but it is independent of the loss function. This motivates the name *separation theorem*, which expresses the fact that the control problem can be split up into two parts: a deterministic control problem to obtain L and a Kalman filtering problem to obtain K. Also notice that the conditional covariance does not depend on the measured data.

To use Theorem 4.1 or Theorem 4.2 it must be asserted that the equations (3.3) and (4.2) have solutions. The conditions Q_2 and R_2 being positive definite and the system (2.1) being completely reachable and completely observable are sufficient to ensure this.

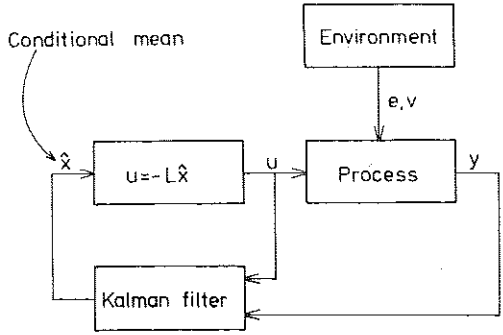

Fig. 5 - Block diagram which illustrates the feedback given by Theorem 4.2.

Stationary Solutions

In many cases the matrices A, B, C, Q_1, Q_2, R_1, and R_2 which appear in the problem formulation are constant. Under weak additional assumptions it can then be shown that if $N \to \infty$ then the Kalman filter gain K and the feedback gain L will converge to unique constant solutions. A sufficient condition is that Q_2 and R_2 are positive and that the system (2.1) is completely reachable and completely observable. In such a case the optimal feedback shown in Fig. 5 is simply a linear time invariant dynamical system. There are, however, cases where the stationary solutions are not unique. An example is given below.

EXAMPLE 4.1

Consider the system

$$x(t+1) = \begin{pmatrix} -a & 1 \\ 0 & 1 \end{pmatrix} x(t) + \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} u(t)$$

with the loss function

$$V = \sum_{1}^{N} x_1^2(t).$$

It is easy to show that if $|b_2| > |b_1|$ then the equation (4.2) has the following two positive solutions as $N \to \infty$.

$$S_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \qquad S_2 = \begin{pmatrix} s_1 & s_2 \\ s_2 & s_3 \end{pmatrix}$$

$$s_1 = 1 + \frac{a^2(b_2^2 - b_1^2)}{(ab_1 - b_2)^2}$$

$$s_2 = -\frac{a(b_2^2 - b_1^2)}{(ab_1 - b_2)^2}$$

$$s_3 = \frac{b_2^2 - b_1^2}{(ab_1 - b_2)^2}.$$

The corresponding feedback gains are

$$L_1 = \frac{1}{b_1}\begin{pmatrix} -a & 1 \end{pmatrix}$$

and

$$L_2 = \frac{ab_2 - b_1}{b_2(ab_1 - b_2)}\begin{pmatrix} -a & 1 \end{pmatrix}.$$

□

5. COMPARISON WITH MINIMUM VARIANCE CONTROL

The problem discussed in Chapter 2 can be regarded as a special case of the linear quadratic control problem. To see this consider a system with one input and one output described by (2.1). Let the criterion be

$$V = \min E \frac{1}{N} \sum_{t=1}^{N} y^2(t).$$

Change the coordinate system in such a way that the matrix A is in companion form. Applying the Theorem 3.1 it is then found that the equation (2.1) can be written as

$$\hat{x}(t+1) = \begin{pmatrix} -a_1 & 1 & \dots & 0 \\ -a_2 & 0 & \dots & 0 \\ \vdots & & & \\ -a_{n-1} & 0 & \dots & 1 \\ -a_n & 0 & \dots & 0 \end{pmatrix} \hat{x}(t) + \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_{n-1} \\ b_n \end{pmatrix} u(t) + \begin{pmatrix} k_1 \\ k_2 \\ \vdots \\ k_{n-1} \\ k_n \end{pmatrix} \varepsilon(t)$$

$$y(t) = \hat{x}_1(t) + \varepsilon(t). \qquad (5.1)$$

It is easy to show by direct comparison that the relation between the input u and the output y can be written as

$$y(t) + a_1 y(t-1) + \ldots + a_n y(t-n) = b_1 u(t-1) + \ldots + b_n u(t-n) + \varepsilon(t) +$$

$$+ c_1 \varepsilon(t-1) + \ldots + c_n \varepsilon(t-n), \qquad (5.2)$$

where

$$c_i = a_i + k_i, \qquad i = 1, 2, \ldots, n. \qquad (5.3)$$

Equation (5.2) is, however, a CARMA model and the equivalence is thus obvious.

6. APPLICATIONS

The linear quadratic gaussian theory is frequently referenced in engineering literature and sometimes also in economics. It is, however, difficult to find good straightforward applications of the theory. Apart from the cases where minimum variance criteria apply it is not easy to find examples where the quadratic criterion (2.3) is well motivated physically. One rare case is the steering of ships where the average increase in drag due to steering can be expressed as

$$\frac{\Delta R}{R} = \frac{k}{T} \int_0^T [\psi^2(t) + \lambda \delta^2(t)] \, dt,$$

where ψ is the heading deviation and δ the rudder angle.

Another difficulty is to obtain appropriate models for the process dynamics and the environment. In spite of this it is frequently attempted to use the LQG theory to solve control problems because the structure of the solution is very appealing intuitively.

7. REFERENCES

The linear quadratic gaussian theory is well covered in textbooks [1], [2], [3], and [4]. A special issue [5] of the IEEE Transactions is entirely devoted to the theory and its applications. Application of the theory to steering of ships is discussed in [6]. Examples of the application of the theory to economic problems are found in [3].

[1] B D O Anderson and J B Moore: Linear Optimal Control. Prentice Hall, Englewood Cliffs, N J, 1971.

[2] K J Åström: Introduction to Stochastic Control Theory. Academic Press, New York, 1970.

[3] G C Chow: Analysis and Control of Dynamic Economic Systems. John Wiley & Sons, New York, 1975.

[4] H Kwakernaak and R Sivan: Linear Optimal Control Systems. Wiley--Interscience, New York, 1972.

[5] M Athans (editor): Special Issue on the Linear-Quadratic-Gaussian Estimation and Control Problem. IEEE Trans AC-16 (1971), no. 6.

[6] K J Åström: Some Aspects on the Control of Large Tankers. Proc Colloques IRIA Analyse de Systèmes et ses Orientations Nouvelles. Versailles-Rocquencourt, December 1976, to appear in Springer Lecture Notes.

CHAPTER 4 - CONTROL OF MARKOV CHAINS

1. INTRODUCTION

The previous chapters have dealt with linear systems only. In this chapter a nonlinear problem will be discussed. To make the analysis simple a model where the process dynamics is approximated by a markov chain will be investigated. The analysis follows the pattern of the previous chapters. The mathematical models of the process and its environment are discussed in Section 2. The solution to the filtering problem is given in Section 3 and the optimal control problem is solved in Section 4.

2. MATHEMATICAL MODELS

Internal descriptions on state models will be investigated. If it is attempted to generalize the model given by equation (2.1) in Chapter 3 to the nonlinear case it will be necessary to describe the conditional probability distribution of $x(t+1)$ given $x(t)$ and of the measurement $y(t)$ given $x(t)$. Such a description will be simplified if the state space is simple. It will therefore be assumed that the state vector x and the measurements y can assume finitely many values only. The stochastic process $\{x(t), t = 0, 1, \dots\}$ thus becomes a markov chain.

It is assumed that the initial probability distribution of the states is given by

$$p_i^0 = \Pr\{x(0) = i\}, \quad i = 1,2,\dots,n. \tag{2.1}$$

The dynamic development of the state is described by the transition probability

$$p_{ij}(u,t) = \Pr\{x(t+1) = j \mid x(t) = i\}, \quad i,j = 1,2,\dots,n. \tag{2.2}$$

The transition probabilities may depend on time t and the control u. The transition probabilities have the properties

$$p_{ij}(u,t) \geq 0, \qquad \sum_{j=1}^{n} p_{ij}(u,t) = 1. \tag{2.3}$$

The measurement process $\{y(t), t = 0, 1, \dots\}$ is similarly characterized by the probabilities

$$q_{ij}(u,t) = \Pr\{y(t) = j \mid x(t) = i\}, \tag{2.4}$$

where

$$q_{ij}(u,t) \geq 0, \quad \sum_{j=1}^{m} q_{ij}(u,t) = 1 . \tag{2.5}$$

The dynamics of the process and its environment are thus characterized by the matrices $P = \{p_{ij},\ i,j = 1,\ldots,n\}$ and $Q = \{q_{ij},\ i = 1,\ldots,n,\ j = 1,\ldots,m\}$ and by the initial distribution of the states.

It is assumed that the purpose is to control the system in such a way that the following loss function is as small as possible.

$$J = E \sum_{t=0}^{N} g[x(t),u(t),t],$$

where g is a function which assumes real values.

The admissible controls are assumed to be such that $u(t)$ is a function of $y_t = [y(t),y(t-1),\ldots,y(0)]$ for each t.

3. OPTIMAL FILTERING

To solve the prediction problem we require the conditional probability distribution

$$w_i(t) = \Pr\{x(t) = i \mid y_t\}. \tag{3.1}$$

If this probability distribution is known then many different predictors like the conditional mean, the value with highest probability etc. can easily be determined. In analogy with the linear case a recursive equation will be given for the predictor. This recursion is given by the following result.

THEOREM 3.1
Introduce the linear maps A_j defined by

$$(A_j w)_i = \sum_{k=1}^{n} q_{ij}\, p_{ki}\, w_k , \qquad j = 1,\ldots,m \tag{3.2}$$

and introduce the norm

$$\| A_j w \| = \sum_{i=1}^{n} (A_j w)_i . \tag{3.3}$$

Then the conditional distribution $w(t)$ defined by (3.1) satisfies

$$w(t+1) = \frac{A_{y(t+1)}\, w(t)}{\| A_{y(t+1)}\, w(t) \|} \tag{3.4}$$

and

$$\| A_j w(t) \| = \Pr \{y(t+1) = j \mid Y_t\}. \tag{3.5}$$

Proof:
It follows from the multiplication rule for probabilities

$$\Pr \{x(t+1) = i \mid Y_{t+1}\} = \Pr \{x(t+1) = i \mid Y_t,\ y(t+1)\} =$$

$$= \frac{\Pr \{x(t+1) = i, y(t+1) = j \mid Y_t\}}{\Pr \{y(t+1) = j \mid Y_t\}}.$$

Furthermore the equations (2.2) and (2.4) give

$$\Pr \{x(t+1) = i, y(t+1) = j \mid Y_t\} = \sum_{k=1}^{n} q_{ij}\, p_{ki}\, w_k(t) = (A_j w)_i$$

and

$$\Pr \{y(t+1) = j \mid Y_t\} = \sum_{i=1}^{n} \sum_{k=1}^{n} q_{ij}\, p_{ki}\, w_k(t) = \| A_j w(t) \|$$

and the proof is complete. □

4. OPTIMAL CONTROL

Having solved the filtering problem the optimal control problem will now be discussed. A functional equation which characterizes the optimal solution will first be derived. The properties of the functional equation will then be discussed.

The Bellman Equation

Assume that the control u can take values in a finite set U only. The minimum of the loss function will then always exist. Introduce the function $V: R^n \to R$ defined by

$$V_t(w(t)) = \min_{u(t), u(t+1), \ldots} E \left\{ \sum_{k=t}^{N} g(x(k), u(k), k) \mid Y_t \right\}. \tag{4.1}$$

Then

$$V_t(w(t)) = \min_u \left\{ \sum_{i=1}^{n} g(i,u,t) w_i(t) + \min_{u(t+1),\ldots} E\left[\sum_{t+1}^{N} q(x(k),u(k),k) \,|\, y_t \right] \right\} =$$

$$= \min_u \left\{ < g(u,t), w(t) > + E\left[V_{t+1}(w(t+1)) \,|\, y_t \right] \right\} .$$

where $<g,w>$ denotes a scalar product and $g(u,t)$ is a vector with components $g(i,u,t)$. It follows from Theorem 3.1 that conditioned on y_t, $w(t+1)$ can assume m different values

$$w(t+1) = \frac{A_j w(t)}{\| A_j w(t) \|} \qquad j = 1,\ldots,m$$

with probabilities $\| A_j w(t) \|$. Hence

$$V_t(w(t)) = \min_u \left[< g(u,t), w(t) > + \sum_{j=1}^{m} \| A_j w(t) \| \, V_{t+1}(A_j w(t) / \| A_j w(t) \|) \right]. \tag{4.2}$$

We now have

THEOREM 4.1
A necessary condition for the minimum is that the function $V_t(w(t))$ satisfies the Bellman equation (4.2).

Proof:
It has been shown that if u is a minimizing feedback then there is a function V which satisfies (4.2). Conversely let v be an admissible control $v: R^n \to U$. Introduce the function W defined by

$$W_t(w(t)) = E\left\{ \sum_{k=t}^{N} g(x(k), v(w(k)), k) \,|\, y_t \right\}.$$

Then W satisfies the recursion

$$W_t(w(t)) = < g(v,t), w(t) > + \sum_{j=1}^{m} \| A_j w(t) \| \, W_{t+1}(A_j w(t) / \| A_j w(t) \|).$$

It is now straightforward to show that

$$W_t(w) \geq V_t(w).$$

This is obviously true for $t = N$ and it follows for $t < N$ by induction. □

The equation (2.4), which is called the *Bellman equation*, plays the role of the Hamilton-Jacobi equation in stochastic control theory. When solving the Bellman equation two functions are obtained:

$V: R^n \to R$

and

$u: R^n \to U$ = {the set of possible controls}.

Structure of Optimal Feedback

The function u is a map from the conditional probability distributions to the controls. The structure of the optimal feedback is thus as shown in Fig. 6. It is composed of a filter which computes the conditional probability density of the states given the past controls and the past measurements. The filter is described by the equation (3.4). The other part of the feedback is the function u which is obtained from the solution of the Bellman equation.

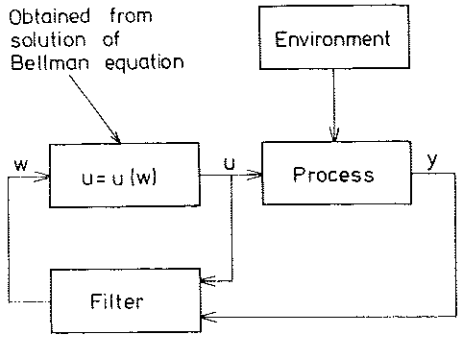

Fig. 6 - Block diagram of the optimal regulator. The filter computes the conditional distribution of the states given at past data Y_t.

Properties of the Bellman Equation

We have

THEOREM 4.2
The solution V of (4.2) is concave.

Proof:
An outline is given below. The full details are given in Åström [1].
The function

$$V_N = \min_u \langle g(u,t), w(t) \rangle$$

is concave because it is the minimum of linear functions. Now use induction. Assume V_{t+1} concave. Then V_t is also concave because it is obtained by adding concave functions with positive weights and minimizing.

□

Computational Aspects

The Bellman equation can rarely be solved analytically. It is thus necessary to resort to numerical solutions. This is not trivial. To solve the equation (4.2) numerically it is necessary to store the functions V_t. This is a substantial burden if the number of states is large. Assume that there are n states. Because the components of w are probabilities the argument of the function can be characterized by $n-1$ variables in the range $0 \leq w_i \leq 1$. Assuming that the components of w are quantized in 10 levels each. It is then necessary to use 10^{n-1} cells to store the function V_t. For $n = 11$ the number is prohibitive even for the largest computers available today. It is then necessary to find good approximations of the function which are more economical storage wise.

The solution to the stochastic control problem is thus useful in the sense that it gives valuable insight into the structure of the optimal feedback. The solution is, however, not very practical in the sense that the computational effort to obtain the optimal feedback is prohibitive if the number of states is large.

5. AN EXAMPLE

An example is used to illustrate the properties of the solution. Consider a case where the transition and observation matrices are defined by

$$P = \begin{pmatrix} 0.9 & 0.1 \\ 0.1 & 0.9 \end{pmatrix}, \quad Q = \begin{pmatrix} 0.8 & 0.2 \\ 0.2 & 0.8 \end{pmatrix}, \quad u = 1$$

$$P = \begin{pmatrix} 0.8 & 0.2 \\ 0.2 & 0.8 \end{pmatrix}, \quad Q = \begin{pmatrix} 0.85 & 0.15 \\ 0.15 & 0.85 \end{pmatrix}, \quad u = 2$$

$$P = \begin{pmatrix} 0.6 & 0.4 \\ 0.4 & 0.6 \end{pmatrix}, \quad Q = \begin{pmatrix} 0.9 & 0.1 \\ 0.1 & 0.9 \end{pmatrix}, \quad u = 3$$

$$P = \begin{pmatrix} 0.4 & 0.6 \\ 0.6 & 0.4 \end{pmatrix}, \quad Q = \begin{pmatrix} 0.9 & 0.1 \\ 0.1 & 0.9 \end{pmatrix}, \quad u = 4$$

$$P = \begin{pmatrix} 0.2 & 0.8 \\ 0.8 & 0.2 \end{pmatrix}, \quad Q = \begin{pmatrix} 0.85 & 0.15 \\ 0.15 & 0.85 \end{pmatrix}, \quad u = 5$$

$$P = \begin{pmatrix} 0.1 & 0.9 \\ 0.9 & 0.1 \end{pmatrix}, \quad Q = \begin{pmatrix} 0.8 & 0.2 \\ 0.2 & 0.8 \end{pmatrix}, \quad u = 6.$$

The loss function is assumed to be given by

$g = (1 \quad 0)$.

It is thus desired to keep the process in the second state. The conditional distribution can be chosen as the conditional probability for the process to be in state $\neq 1$. The Bellman equation was in this case solved numerically by quantizing this probability into 10 steps. The control law $u(w)$ obtained is given below.

step \ w	0.05	0.15	0.25	0.35	0.45	0.55	0.65	0.75	0.85	0.95
10	1	1	1	1	1	6	6	6	6	6
9	1	1	2	2	3	4	5	5	6	6
8	1	1	2	2	3	4	5	5	6	6
7	1	1	2	2	3	4	5	5	6	6
⋮										
1	1	1	2	2	3	4	5	5	6	6

The solution obtained for a finer quantization of 20 steps is given by

t \ w	0.025	0.075	0.125	0.175	0.225	0.275	0.325	0.375	0.425	0.475	0.525	0.575	0.625	0.675	0.725	0.775	0.825	0.875	0.925	0.975
10	1	1	1	1	1	1	1	1	1	1	6	6	6	6	6	6	6	6	6	6
9	1	1	1	1	2	2	2	2	3	3	4	4	5	5	5	5	6	6	6	6
8	1	1	1	2	2	2	2	3	3	3	4	4	4	5	5	5	5	6	6	6
7	1	1	1	2	2	2	2	2	3	3	4	4	5	5	5	5	5	6	6	6
6	1	1	1	2	2	2	2	3	3	3	4	4	4	5	5	5	5	6	6	6
5	1	1	1	2	2	2	2	2	3	3	4	4	5	5	5	5	5	6	6	6
4	1	1	1	2	2	2	2	2	3	3	4	4	5	5	5	5	5	6	6	6
3	1	1	1	2	2	2	2	2	3	3	4	4	5	5	5	5	5	6	6	6
2	1	1	1	2	2	2	2	2	3	3	4	4	5	5	5	5	5	6	6	6
1	1	1	1	2	2	2	2	2	3	3	4	4	5	5	5	5	5	6	6	6

In the last step $t = 10$ the solution is obvious if the probability of being in state # is less than 0.5, then the control # 1 is chosen. This means that the probability of the state being unchanged is as large as possible. Otherwise control # 6 is chosen which means that the probability for a switch is maximized.

For step 9 the same policy is used provided that the probability of being in state # 1 is very small or very large. When the probability of being in state # is between 0.2 and 0.4 the optimal control is however $u = 2$. This means that the probability for a switch is higher than for $u = 1$. The measurements will, however, be more accurate which will benefit the conditional probabilities in the next step. If the probability of being in the state # 1 is between 0.4 and 0.5 it is beneficial to choose $u = 3$ which gives an even better measurement accuracy.

This example clearly illustrates some interesting properties of the solution to the nonlinear stochastic control problem. The control law may generate control actions that will drive the process away from its target provided that this will result in a more accurate estimate of the state. This property is called *dual control*.

Also notice in the tables above that the control law converges after a few steps only.

6. REFERENCES

The probabilistic model used in this section was introduced and analysed in [1]. The model is discussed extensively in [2]. The possibilities of approximating a system with a continuous state space by a markov chain is discussed in [3].

[1] K J Åström: Optimal Control of Markov Processes with Incomplete State Information, I and II. J Math Anal Appl 10 (1965) 174-205 and 26 (1969) 403-406.

[2] H J Kushner: Introduction to Stochastic Control. Holt, Rinehart and Winston, New York, 1971.

[3] R W Brockett: Stationary Covariance Generation with Finite State Markov Processes. Paper TA26-12:30, Joint Automatic Control Conference 1977, pp 1057-1060.

CHAPTER 5 - NONLINEAR STOCHASTIC CONTROL

1. INTRODUCTION

In this chapter the results of Chapter 4 will be extended to systems where the statespace is continuous. It will be shown that the problem formulation includes many interesting control problems. For example it is possible to treat adaptive control systems using these models. The theoretical results that are available are unfortunately fairly weak. Very little is known about existence. The order of presentation is the same as has been used in the previous chapters. The mathematical models used are discussed in Section 2. It is shown that the interesting cases of linear systems with drifting parameters and linear systems with constant but unknown parameters are included as special cases. The filtering problem is analysed in Section 3. A recursive equation is derived for the conditional probability density of the state variables. It is shown that the conditional densities are gaussian in particular cases. In Section 4 the control problem is investigated. The Bellman equation is derived formally. Unfortunately neither the recursive equation for the conditional density nor the Bellman equation are suitable to solve practical problems because of the excessive computational requirements. The analysis gives, however, interesting insight into the nature of the optimal solution. This insight can then be exploited to obtain different useful approximations. Some approximations are discussed in Section 5.

2. MATHEMATICAL MODELS

When analysing nonlinear problems it is frequently easier to work with internal descriptions. It is assumed as in Chapter 3 that the state x, the input u, and the output y take values in R^n, R^p, and R^r respectively. A general nonlinear generalization of the linear model discussed in Section 2 of Chapter 3 is then given by

$$\begin{cases} x(t+1) = f(x(t), u(t), v(t)) \\ y(t) = g(x(t), u(t), e(t)), \end{cases} \tag{2.1}$$

where $\{v(t)\}$ and $\{e(t)\}$ are sequences of random variables. The probability distribution of the initial state is characterized by

$$p^0(x)\, dx = \Pr\{x(t_0) \in x + dx\}, \tag{2.2}$$

where dx is an infinitesimal neighborhood of x.

To complete the characterization of the model it is necessary to specify the probability distributions of the disturbances v and e. If the model (2.1) should represent a state model in the sense that the conditional distributions of $x(t+1)$ and $y(t)$ given $x(t)$ are the same as the conditional distributions of $x(t+1)$ and $y(t)$ given $x(t), x(t-1), \ldots$ then $\{v(t)\}$ and $\{e(t)\}$ must be sequences of independent random variables. The stochastic process $\{x(t)\}$ is then a Markov process. Instead of using the description (2.1) it is then natural to work directly with the probabilities

$$\begin{cases} p(\xi,x)dx = \Pr\{x(t+1) \in x+dx \mid x(t) = \xi\} \\ q(\xi,y)dy = \Pr\{y(t) \in y+dy \mid x(t) = \xi\}. \end{cases} \tag{2.3}$$

It is assumed that these densities exist. The densities p and q will also depend on t and $u(t)$. This dependence is suppressed to simplify the notations. If the probability distributions for $v(t)$ and $e(t)$ are known the densities p and q can be determined. In the sequel it is therefore assumed that p and q are known.

The model (2.1) or (2.3) includes several special cases that are of great interest. Some will be discussed below.

EXAMPLE 2.1 (Linear Systems with Stochastic Parameters)
Consider a linear system characterized by the input-output relation

$$y(t+1) + a_1(t)y(t) + \ldots + a_n(t)y(t-n+1) = b_1(t)u(t) + \ldots + b_n(t)u(t-n+1) + e(t). \tag{2.4}$$

Introduce

$$\theta_1(t) = a_1(t), \ldots, \theta_n(t) = a_n(t), \theta_{n+1}(t) = b_1(t), \ldots, \theta_{2n}(t) = b_n(t)$$

and assume that the parameters are governed by

$$\theta(t+1) = \Phi\theta(t) + v(t), \tag{2.5}$$

where $\{v(t)\}$ is a sequence of uncorrelated gaussian random variables and the initial state $\theta(t_0)$ is gaussian (m, R_0). Introducing the vector

$$\varphi(t) = [-y(t) \ldots -y(t-n+1) \; u(t) \; u(t-1) \ldots u(t-n+1)]$$

the model (2.4) can then be written as

$$y(t+1) = \varphi(t)\theta(t) + e(t). \tag{2.4'}$$

The system described by (2.4) and (2.5) is clearly of the form (2.1). A special case is when $\Phi = I$ which means that the parameters $a_i(t)$ and $b_i(t)$ are discrete Wiener processes. In specific

applications it may not be realistic to assume that a model like (2.4) can be given for the parameter fluctuations. The case $\Phi = I$ can however serve as a generic case. □

EXAMPLE 2.2 (Systems with Constant but Unknown Parameters)
It is frequently possible to assume that the equation which describes the process dynamics can be described by models having parameters which are constant but unknown. Such systems can be included into the form (2.1) by introducing the unknown parameters θ as extra state variables governed by the state equation

$$\theta(t+1) = \theta(t).$$

Consider for example a linear system governed by (2.4) where the parameters are constant. Such a system can be described by equations (2.4) and (2.5) with $\Phi = I$. The initial distribution of the state of (2.5) reflects the prior knowledge of the parameters. □

3. OPTIMAL FILTERING

The optimal filtering problem will now be solved for the model (2.3). Since it is not possible to find a universally acceptable criterion the full conditional probability distribution of the state will be determined. It is thus assumed that the outputs $y(t), y(t-1), \ldots, y(t_0)$ have been observed. The problem is to determine the conditional probability distribution

$$\Pr\{x(t) \in A | Y_t\}, \tag{3.1}$$

where Y_t is the values of all observed outputs or more precisely the σ-algebra generated by $y(t), y(t-1), \ldots, y(t_0)$ and $x(t_0)$. Assuming that the distribution (3.1) has a density denoted as

$$w(x,t)dx = \Pr\{x(t) \in x + dx | Y_t\},$$

we find from the properties of conditional densities that

$$w(x,t+1) = \frac{q(x,y(t+1)) \int p(\xi,x) w(\xi,t) d\xi}{\iint q(x,y(t+1)) p(\xi,x) w(\xi,t) dx d\xi} \tag{3.2}$$

$$w(x,t_0) = p^0(x).$$

The expression can be simplified somewhat if the linear positive operator A is defined by

$$(A_\eta\, w)(x) = q(x,\eta) \int p(\xi,x) w(\xi,t)\, d\xi \,. \tag{3.3}$$

Notice that A depends on u and t. Define the norm of a positive function w as

$$\| w(t) \| = \int w(x,t)\, dx \,. \tag{3.4}$$

Then the formula for updating the conditional density can be written as

$$w(\cdot,t+1) = A_{y(t+1)} w(\cdot,t) / \| A_{y(t+1)} w(\cdot,t) \| \,. \tag{3.5}$$

This formula is clearly a generalization of equation (3.4) in Chapter 4 for the markov chain case and we get

THEOREM 3.1
Consider a stochastic process defined by (2.2) and (2.3). The conditional density of the state $x(t)$ given past data y_t is given by the recursive equation (3.5) with initial condition

$$w(t_0) = p^0 \,.$$

Furthermore

$$\Pr \{ y(t+1) \in \eta + d\eta \mid y_t \} = \| A_\eta w(\cdot,t) \| d\eta \,. \tag{3.6}$$

Even if the equation (3.5) looks innocent it requires extensive computations. Since w is a probability density over R^n the problem of storing the function is substantial. It is therefore interesting to consider special cases where the conditional density is simple. Such cases are discussed below.

THEOREM 3.2
Consider a linear system with random parameters described by (2.4) and (2.5) where $\{v(t)\}$ and $\{e(t)\}$ are sequences of independent gaussian $(0,R_1)$ and $(0,R_2)$ random variables. Let the initial parameter distribution be gaussian (m,R_0). Then the conditional distribution of the parameters $\theta(t)$ given y_t is gaussian $(\hat{\theta}(t), P(t))$ where

$$\begin{aligned}
\hat{\theta}(t+1) &= \Phi\hat{\theta}(t) + K(t)[y(t+1) - \varphi(t)\hat{\theta}(t)] \\
K(t) &= \Phi P(t)\varphi^T(t)[\varphi(t)P(t)\varphi^T(t) + R_2]^{-1} \\
P(t+1) &= [\Phi - K(t)\varphi(t)]P(t)\Phi^T + R_1 \\
\hat{\theta}(t_0) &= m \\
P(t_0) &= R_0 \,.
\end{aligned} \tag{3.7}$$

Furthermore the conditional density of $y(t+1)$ given $\mathcal{Y}_t$ is gaussian $(\varphi(t)\hat{\theta}(t), \sigma^2(t))$ where

$$\sigma^2(t) = R_2 + \varphi(t)P(t)\varphi^T(t). \tag{3.8}$$

Proof:
The system given by (2.4) and (2.5) is clearly a special case of (2.3). It follows from (2.4) that the density $p(\xi,x)$ is gaussian $(\varphi\xi, R_1)$ and that the density $q(\xi,y)$ is gaussian $(\varphi\xi, R_2)$. It was furthermore assumed that the prior density was gaussian (m, R_0). Repeated use of (3.2) then shows that the conditional density is also gaussian. The formulas (3.7) and (3.8) are verified simply by computing the density in the same way as was done when deriving the Kalman filter theorem in Section 3 of Chapter 3. □

Remark 1
Notice that the distribution of $y(t)$ is not gaussian.

Remark 2
A similar result can be obtained for vector difference equations (2.3). In particular this means that if the parameters appear as elements of the matrices A and B in the model

$$x(t+1) = Ax(t) + Bu(t) + v(t),$$

then the conditional density of the parameters given $x(t), x(t-1), \ldots$ is also gaussian.

4. OPTIMAL CONTROL

The optimal control of processes described by equations (2.2) and (2.3) will now be investigated. It is assumed that the purpose of the control can be expressed as to minimize the loss function

$$J = E \sum_{t=t_0}^{N} h(x(t), u(t), t), \tag{4.1}$$

where $h: R^{n+p+1} \to R$. The admissible controls are assumed to be such that $u(t)$ is a function of $y(t), y(t-1), \ldots$ or more precisely that $u(t)$ is measurable with respect to the σ-algebra $\mathcal{Y}_t$ generated by $y(t), y(t-1), \ldots, y(t_0)$ and $x(t_0)$.

Proceeding in the same way as in the previous chapter it is first found that it is difficult to show existence of the minimum. It is therefore assumed that the minimum exists and we will proceed formally.

Introduce

$$V_t(w(\cdot,t)) = \min_u E\left\{\sum_{k=t}^{n} h(x(k), u(k), k) \mid y_t\right\}. \tag{4.2}$$

We find

$$V_t(w(\cdot,t)) = \min_u \left\{\int h(\xi, u, t) w(\xi, t) d\xi + E\left[V_{t+1}(w(\cdot,t+1)) \mid y_t\right]\right\}.$$

It follows from Theorem 3.1 that

$$V_t(w) = \min_u \left[\langle h, w\rangle + \int \| A_\eta w \| \; V_{t+1}(A_\eta w / \| A_\eta w \|) d\eta\right], \tag{4.3}$$

where

$$\langle h, w\rangle = \int h(\xi, u, t) w(\xi, t) d\xi.$$

The minimum is thus characterized by the Bellman equation (4.3). Notice that the argument of the Bellman equation is a density of a probability distribution over R^n. Even a numerical solution is thus not possible in the general case.

The convexity of V can be established in the same way as was done for systems with finite states in Chapter 4.

Even if the Bellman equation can not be solved the analysis shows that even in the general nonlinear case the structure of the feedback is that shown in Fig. 6 of Chapter 4. The optimal feedback can thus be thought of as being composed of two parts. One part is a nonlinear filter which computes the conditional density of the state vector given all observations. The other part is a nonlinear function which maps the conditional density on the control variables. This function can be precomputed from the Bellman equation.

5. LINEAR SYSTEMS WITH RANDOM PARAMETERS

In the general case the Bellman equation can neither be solved analytically nor computationally. Some special cases will therefore be investigated. The linear system with stochastic parameters given in Example 2.1 will be investigated.

It is assumed that the criterion is to minimize the loss function

$$J_N = E\left\{\sum_{t=t_0}^{N} [y(t) - y_r(t)]^2\right\}, \tag{5.1}$$

where $y_r(t)$ is a given reference value. The admissible controls are

assumed to be such that $u(t)$ is a function of $Y_t = y(t), \ldots, y(t_0)$. The reason for choosing the criterion (5.1) is that the solution to the control problem is known if the parameters θ of the model are known. In the case of constant parameter systems the solution was e.g. given in Chapter 2. The analysis will thus illustrate the effects of fluctuations in the model parameters.

Even in this simple case it is not easy to establish existence of the minimum.

The Filtering Problem

The filtering problem will first be solved. The system is described by equations (2.4) and (2.5). It can be brought to standard form (2.3) by introducing a state composed of the vectors $\theta(t)$ and the vector $\tilde{\varphi}$ which is defined by

$$\tilde{\varphi}(t) = [-y(t) \ \ldots \ -y(t-n+1) \ \ 0 \ \ u(t-1) \ \ldots \ u(t-n+1)]. \tag{5.2}$$

The conditional distribution of $\tilde{\varphi}(t)$ given Y_t is a point distribution. It was shown in Theorem 3.2 that the conditional distribution of $\theta(t)$ given Y_t is gaussian $(\hat{\theta}(t), P(t))$ for linear systems with random parameters. The conditional distribution of the state of the system is called the *hyperstate*. It can be characterized by the triple

$$\xi(t) = [\tilde{\varphi}(t), \hat{\theta}(t), P(t)]. \tag{5.3}$$

The equation for updating $\tilde{\varphi}$ follows directly from (5.2). The equations for updating $\hat{\theta}$ and P are given by Theorem 3.2.

The Control Problem

The filtering problem is thus easily solved for the particular model structure chosen. To discuss the control problem we will first solve the problem in the case the parameters are known. The problem will then be solved for the special case when the criterion (5.1) only contains one term. The solution of the complete problem is finally discussed.

Systems with Known Parameters

If the parameters of the system (2.4) are known it is easy to obtain the optimal feedback. Introducing the vector $\tilde{\varphi}$ defined by (5.2) the equation (2.4) can be written as

$$y(t+1) = \varphi(t)\theta(t) + e(t+1) = b(t)u(t) + \tilde{\varphi}(t)\theta(t) + e(t+1), \tag{5.4}$$

where we have introduced $b(t) = b_1(t)$. The optimal feedback is then

given by

$$u(t) = \frac{y_r(t+1) - \tilde{\varphi}(t)\theta(t)}{b(t)} . \tag{5.5}$$

The minimal loss is given by

$$\min J_N = (N+1-t_0)R_2.$$

Notice that it is necessary to impose the condition $b(t) \neq 0$ otherwise the control law (5.5) does not make sense. It is also necessary to require that the parameters $b_i(t)$ are such that the difference equation

$$b_1(t)u(t-1) + b_2(t)u(t-2) + \ldots + b_n(t)u(t-n) = 0$$

are asymptotically stable because otherwise the control signal will become infinitely large. Compare with the discussion in Section 4 of Chapter 2. A simple example is used as an illustration.

EXAMPLE 5.1
Consider a process described by

$$y(t+1) = y(t) + b(t)u(t) + e(t+1). \tag{5.6}$$

This is a discrete time version of a continuous time system whose output is the time integral of its input. Let the reference value y_r be zero. If the parameter b is known then the control which minimizes J_1 or J_N is given by

$$u(t) = -\frac{y(t)}{b(t)} . \tag{5.7}$$

The optimal feedback is thus a proportional regulator with gain $1/b(t)$. □

Certainty Equivalence Control

When the parameters θ of the system (2.4) are not known it is tempting to replace the control law (5.5) by the control law

$$u(t) = \frac{y_r(t+1) - \tilde{\varphi}(t)\hat{\theta}(t)}{\hat{b}(t)} , \tag{5.8}$$

where $\hat{\theta}(t)$ is the conditional mean of the values $\theta(t)$ of the unknown parameters given Y_t.

EXAMPLE 5.2
The certainty equivalence control for the process (5.6) is given by

$$u(t) = -\frac{1}{\hat{b}(t)} y(t). \tag{5.9}$$

□

Cautious Control or One-Step Control

The special case when the criterion (5.1) has one term only is first discussed. According to Theorem 3.2 the conditional distribution of $y(t+1)$ given y_t is gaussian with mean $\varphi(t)\hat{\theta}(t)$ and covariance $R_2 + \varphi(t)P(t)\varphi^T(t)$. Then

$$E\{[y_r(t+1) - y(t+1)]^2 | y_t\} = [y_r(t+1) - \varphi(t)\hat{\theta}(t)]^2 + R_2 + \varphi(t)P(t)\varphi^T(t). \tag{5.10}$$

To see how the right hand side depends on the control $u(t)$, introduce equation (5.4). Then

$$E\{[y_r(t+1) - y(t+1)]^2 | y_t\} = R_2 + [y_r(t+1) - \hat{b}(t)u(t) - \tilde{\varphi}(t)\hat{\theta}(t)]^2 +$$
$$+ u^2(t)p_b(t) + 2u(t)\tilde{\varphi}(t)P(t)\ell + \tilde{\varphi}(t)P(t)\tilde{\varphi}^T(t), \tag{5.11}$$

where ℓ is a column-vector which selects the $(n+1)$:th row of the matrix P i.e.

$$\ell = \text{col}\ (\underbrace{0 \ldots 0}_{n}\ 1\ \underbrace{0 \ldots 0}_{n-1}).$$

Minimization of (5.11) with respect to $u(t)$ gives

$$\min_{u(t)} E\{[y_r(t+1) - y(t+1)]^2 | y_t\} = R_2 + [y_r(t+1) - \tilde{\varphi}(t)\hat{\theta}(t)]^2 + \tilde{\varphi}(t)P(t)\tilde{\varphi}^T(t)$$
$$+ \frac{[y_r(t+1) - \hat{b}(t)\tilde{\varphi}(t)\hat{\theta}(t) - \tilde{\varphi}(t)P(t)\ell]^2}{\hat{b}^2(t) + p_b(t)}, \tag{5.12}$$

where the minimum is attained for the control law

$$u(t) = \frac{y_r(t+1) - \hat{b}(t)\tilde{\varphi}(t)\hat{\theta}(t) - \tilde{\varphi}(t)P(t)\ell}{\hat{b}^2(t) + p_b(t)}. \tag{5.13}$$

This control law is called *one-step control*, because it minimizes the expected loss over one step only. Notice that in the case of known parameters the minimal loss is a constant. This means that the one-step control is optimal also for the N-step criterion. This is, however, not the case when the parameters are unknown.

EXAMPLE 5.3

The one-step control for the process (5.6) is given by

$$u(t) = -\frac{\hat{b}(t)}{\hat{b}^2(t) + p_b(t)}\, y(t) = -\frac{\hat{b}^2(t)}{\hat{b}^2(t) + p_b(t)} \cdot \frac{1}{\hat{b}(t)}\, y(t). \tag{5.14}$$

□

For the control laws, ((5.5) known parameters), ((5.8) certainty equivalence), and ((5.13) one-step), the input-output relation can be expressed as

$$u(t) = \alpha_1(t)y(t) + \ldots + \alpha_n y(t-n+1) + \beta_1(t)u(t-1) + \ldots + \beta_n(t)u(t-n+1).$$

When the parameters are known (5.5) the coefficients $\alpha_i(t)$ and $\beta_i(t)$ are simply functions of time. But when the parameters are stochastic the parameters $\alpha_i(t)$ and $\beta_i(t)$ will depend also on past observations y_t.

Notice that the one-step control law (5.13) reduces to the certainty equivalence control if the conditional variance $p_b(t)$ of the estimate $\hat{b}(t)$ is zero.

The examples 5.2 and 5.3 clearly illustrate the differences between the one-step and the certainty equivalence controls. In these examples both regulators are simply proportional controllers. The one-step control (5.13) has a gain which is a factor

$$\hat{b}^2(t) / [\hat{b}^2(t) + p_b(t)]$$

lower than the gain of the certainty equivalence control. The effect of the parameter uncertainty is thus to reduce the gain. For this reason the one-step control is also called the *cautious* regulator.

The cautious regulator does not have the dual property discussed in Section 5 of Chapter 4. To obtain such a regulator it is necessary to solve the multistep optimization problem.

Multistep Optimization

To solve the multistep optimization problem it is necessary to solve the Bellman equation for the stochastic control problem. The Bellman equation can be simplified because the conditional distributions are gaussian.

Assume that the minimum exists and recall that the conditional distribution of the state given the measurements can be characterized by the triplet (5.3). Introduce

$$V(\tilde{\varphi}(t), \hat{\theta}(t), P(t), t) = \min_u E\left\{ \sum_{k=t+1}^{N} [y(k) - y_r(k)]^2 | y_t \right\}.$$

The following recursive equation is then obtained.

$$V(\tilde{\varphi}(t), \hat{\theta}(t), P(t), t) = \min_{u(t)} E\left\{ [y(t+1) - y_r(t+1)]^2 + V(\tilde{\varphi}(t+1), \hat{\theta}(t+1), P(t+1), t+1) | y_t \right\}. \qquad (5.15)$$

To proceed it is necessary to have the equations for the updating of the hyperstate (5.3). It follows from (2.4') and (5.2) that

$$\tilde{\varphi}(t+1) = \tilde{\varphi}^T(t)C + f\,\varphi(t)\hat{\theta}(t) + f\,\sigma(t)\varepsilon(t+1),$$

where ε is the normalized innovation

$$\varepsilon(t+1) = [y(t+1) - \varphi(t)\hat{\theta}(t)] \,/\, \sigma(t),$$

$$\sigma^2(t) = \varphi(t)P(t)\varphi^T(t) + R_2 ,$$

and

$$C = \begin{pmatrix} S_1 & 0 \\ 0 & S_1 \end{pmatrix}, \quad S_1 = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & & & & \\ 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & 0 & \dots & 0 \end{pmatrix}, \quad f = \begin{pmatrix} 1 & 0 \dots 0 \end{pmatrix}.$$

It also follows from (3.7) that

$$\hat{\theta}(t+1) = \Phi\hat{\theta}(t) + K(t)\sigma(t)\varepsilon(t+1).$$

The conditional densities of $\tilde{\varphi}(t+1)$ and $\hat{\theta}(t+1)$ given y_t are thus gaussian. Furthermore the conditional distribution of $P(t+1)$ given y_t is a point distribution with all mass in

$$P(t+1) = [\Phi - K(t)\varphi^T(t)]P(t)\Phi^T + R_1 ,$$

where

$$K(t) = \Phi P(t)\varphi^T(t)\sigma^{-2}(t).$$

The functional equation (5.15) can then be written as

$$V(\varphi, \hat{\theta}, P, t) = \min_u \left\{ [y_r(t+1) - \varphi\hat{\theta}]^2 + \sigma^2 + \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\varepsilon^2/2} \cdot \right. \tag{5.16}$$

$$\left. \cdot\, V[\tilde{\varphi}C + f\varphi\hat{\theta} + f\sigma\varepsilon,\ \Phi\hat{\theta} + K\sigma\varepsilon,\ (\Phi - K\varphi^T)P\Phi^T + R_1,\ t+1]d\varepsilon \right\},$$

where

$$\varphi\hat{\theta} = \tilde{\varphi}\hat{\theta} + \hat{b}u$$

$$K = \Phi P\varphi^T\sigma^{-2}$$

$$\sigma^2 = \varphi P\varphi^T + R_2 .$$

Notice that φ, K, P, and σ depend on the control u.

The equation (5.16) is the Bellman equation for the problem. The equation can not be solved analytically. The equation can be solved numerically if the order of the system is sufficiently small.

EXAMPLE 5.4

For the system described by (5.6) the conditional distribution can be characterized by three variables $y(t)$, $\hat{b}(t)$, and $P(t)$. Assuming that $\Phi = 1$ and $R_2 = r_2$ the Bellman equation (5.16) then becomes

$$V(y,\hat{b},P,t) = \min_u \left\{ [y_r(t+1) - y - \hat{b}u]^2 + r_2 + u^2P + \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\varepsilon^2/2} \cdot \right.$$

$$\left. \cdot\, V\left(y + \hat{b}u + \sqrt{r_2+u^2P}\,\varepsilon,\ \hat{b} + \frac{uP}{\sqrt{r_2+u^2P}}\,\varepsilon,\ \frac{r_2P}{r_2+u^2P} + r_1,\ t+1\right)d\varepsilon \right\}. \tag{5.17}$$

□

The Bellman equation has been solved numerically for special low order examples. The solutions have given insight into the nature of the optimal strategies. In particular it has been found that the character of the multistep optimization is different from the cautious control in the sense that the optimal feedback is *dual*. The optimal feedback will thus generate control actions which will improve the accuracy of future estimates at the expense of increased short term loss. The properties of the optimal feedback are thus similar to those found for the markov chain example in Section 5 of Chapter 4. Another interesting property is that the optimal feedback may be discontinuous in $P(t)$. From a theoretical point of view there are, however, many important problems that still remain unsolved. The existence of stationary policies as $N \to \infty$, the existence of optimal solutions are typical examples.

Because of the difficulty of solving the Bellman equation several suboptimal control strategies have been proposed. They have mainly been investigated by simulations.

6. REFERENCES

The concept of dual control was introduced by Feldbaum [1]. The Bellman equation can be regarded as a natural extension of classical Hamilton-Jacobi theory to stochastic control problems. See [2]. This chapter is largely based on the paper [3] which also contains an example of a dual control law obtained by solving the Bellman equation. Further examples of this type are given in [4] and [5]. Different suboptimal controls are discussed in [6] which contains many additional references and several simulation examples.

[1] A A Feldbaum: Dual Control Theory I-IV. Automation and Remote Control 21 (1961) 874-880, 1033-1039, and 22 (1961) 1-12, 109-121.

[2] R Bellman: Adaptive Control Processes - A Guided Tour. Princeton Univ Press, Princeton, 1961.

[3] K J Åström and B Wittenmark: Problems of Identification and Control. J Math Anal Appl 34 (1971) 90-113.

[4] T Bohlin: Optimal Dual Control of a Simple Process with Unknown Gain. Technical Paper PT 18.196, IBM Nordic Laboratory, Lidingö, Sweden, 1969.

[5] O L R Jacobs and J W Patchell: Caution and Probing in Stochastic Control. Int J on Control 16 (1972) 189-199.

[6] J Sternby: Topics in Dual Control. PhD dissertation, Dept of Automatic Control, Lund Institute of Technology, CODEN: LUTFD2/(TFRT-1012)/1-135/(1977).

CHAPTER 6 - SELF-TUNING REGULATORS

1. INTRODUCTION

It was shown in Chapters 2 and 3 that optimal stochastic control problems could be formulated and solved at least for linear systems with quadratic criteria. It is, however, a difficulty from the point of view of applications that the models describing the process and its environment are rarely known apriori. Nonlinear stochastic control problems were formulated in Chapters 4 and 5. It was shown that linear systems with unknown parameters could fit into the problem formulation provided that the state was measured exactly or that the dynamics could be described by a controlled autoregression. The theory developed in Chapter 5 can thus be applied to generalize some linear problems to the case of unknown parameters. The results of Chapter 5 give interesting insight into the structure of the optimal feedback. The results are, however, discouraging from a practical point of view because of the computations required to obtain the optimal feedback. It is thus meaningful to attempt a reformulation of the control problem which will lead to practical solutions. The following is one possibility. Consider a system with constant but unknown parameters and a criterion. Find a control law, which only operates on past input-output data, which does not require knowledge of the system parameters, and which converges to the optimal regulator that could be designed if the parameters of the process were known. Such a regulator is called a *self-tuning* regulator. There are many self-tuning regulators. The optimal dual controller is clearly self-tuning. It will be shown in this chapter that there are indeed self-tuning regulators at least for simple problems. The mathematical model of the process and its environment is discussed in Section 2. To keep the analysis simple only a simple first order system is treated. The problem formulation is also given in Section 2. In Section 3 it is shown that a solution is given by a comparatively simple feedback law. The properties of this feedback are analysed in Section 4.

2. MATHEMATICAL MODEL

It is assumed that the dynamics of the process and its environment can be described by the simple first order system

$$y(t) + ay(t-1) = bu(t-1) + e(t) + ce(t-1), \tag{2.1}$$

where u is the control variable, y the output and $\{e(t)\}$ a sequence of independent gaussian random variables. It is furthermore assumed that the criterion is to minimize the quadratic loss function

$$J = \lim_{N\to\infty} E \frac{1}{N} \sum_{t=1}^{N} y^2(t). \tag{2.2}$$

The admissible controls are assumed such that $u(t)$ is a function of all past outputs $y(t), y(t-1), \ldots$. If the parameters are known it follows from Theorem 4.1 of Chapter 2 that the optimal control is the proportional feedback

$$u(t) = \frac{a-c}{b} y(t). \tag{2.3}$$

The problem of finding a self-tuning regulator can be stated as follows. Find a feedback law which does not depend on knowledge of the parameters a, b, and c which converges to the control law (2.3) as time increases.

3. A SIMPLE SELF-TUNING REGULATOR

If it is attempted to solve the problem formulated in Section 2 using the methods discussed in Chapter 5 the filtering problem must first be solved. The state of the system is $y(t)$ and the parameters a, b, and c. For the filtering problem a prior distribution for a, b, and c must be assumed. Recursive equations for the conditional distribution of a, b, and c given Y_t can then be obtained. This distribution can be simplified a little by observing that the conditional distribution of a and b given c and Y_t is gaussian. It can be characterized by two mean values $(\hat{a}, \hat{b})$ and three covariances (p_a, p_{ab}, and p_b). The conditional distribution of a, b, and c given Y_t can thus be characterized by the conditional distribution of c and 5 real variables ($\hat{a}$, $\hat{b}$, p_a, p_{ab}, and p_b). The problem is simplified considerably if the parameter c is known because the conditional density of a and b given Y_t is then gaussian as was shown in Theorem 3.2 of Chapter 5.

Assuming that the filtering problem is solved it can then be attempted to solve the Bellman equation. This may perhaps be done computationally in the case c is known because the state can then be characterized by 5 real variables. It is thus clear that even in a simple case like this it is not possible to compute the optimal dual control law if c is unknown.

Since the parameters of the process are constant it can be expected that the conditional distributions of the parameters a, b, and c given Y_t will converge to point distributions. For large t it can then be expected that reasonably good self-tuning strategies can be obtained from control laws that are computed from parameter estimates only.

It will be shown that there are indeed simple self-tuning control laws. One possibility is given by the following control law

$$u(t) = \hat{\theta}(t)y(t), \tag{3.1}$$

where $\hat{\theta}(t)$ is the least squares estimate of the parameter θ in the model

$$y(t) + \theta y(t-1) = u(t-1) + e(t) \tag{3.2}$$

based on data available up to time t i.e. $y(t), y(t-1), \ldots, y(1)$, $u(t-1), u(t-2), \ldots, u(1)$. The least squares estimate $\hat{\theta}$ is given by

$$\hat{\theta}(t) = -\Big[\sum_{k=1}^{t-1} [y(k+1) - u(k)]\, y(k)\Big] \Big/ \sum_{k=1}^{t-1} y^2(k). \tag{3.3}$$

The control algorithm given by (3.1) and (3.3) can be expected to work nicely for the system (2.1) if $c = 0$ and $b = 1$. In this case the least squares estimate $\hat{\theta}$ will converge to a as $t \to \infty$ and the control law (3.1) will converge to

$$u(t) = ay(t),$$

which is the desired control law.

It is a remarkable property of the feedback law described by (3.1) and (3.3) that it will converge to the optimal law (2.3) also when $c \neq 0$. This is illustrated in the following example.

EXAMPLE 3.1

Consider a system (2.1) with $a = -1$, $c = -0.7$, and $b = 1$. With these numerical values the optimal feedback is

$$u(t) = -0.3\, y(t).$$

Fig. 7 shows a simulation of the regulator (3.1), (3.3) applied to the system. From the simulation it appears that the parameter estimate converges to the value $\hat{\theta} = -0.3$ (which is the gain of the optimal feedback).

To compare the self-tuning regulator (3.1) (3.3) with the optimal regulator in the case of known parameters the accumulated loss defined by

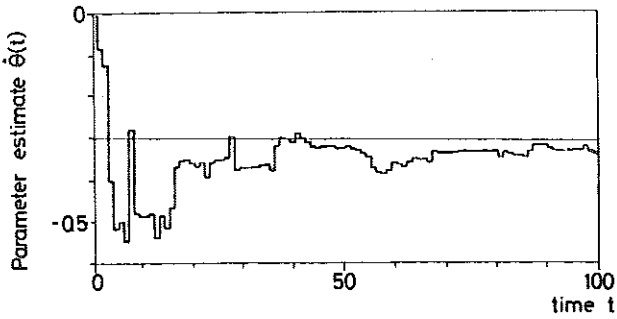

Fig. 7 - Parameter estimate $\hat{\theta}$ obtained in a simulation of the self-tuning regulator (3.1), (3.3) applied to the system (2.1) with $a = -1$, $c = -0.7$, and $b = 1$.

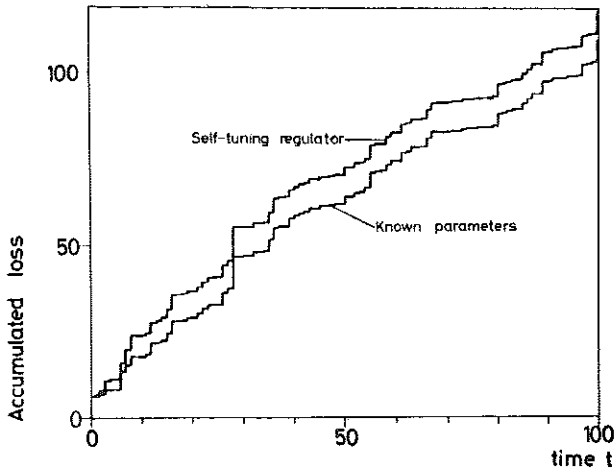

Fig. 8 - Accumulated loss functions for the self-tuning regulator (3.1), (3.3) and the optimal regulator based on known parameters.

$$V(t) = \sum_{k=1}^{t} y^2(k)$$

has been calculated for the self-tuning regulator and the optimal regulator

$$u(t) = -0.3\, y(t).$$

The results are shown in Fig. 8. It is clear from this figure that there is not a large difference between the performance of the two

regulators. In particular it is seen in Fig. 8 that the loss functions will be virtually the same if the first 20 steps are neglected. □

The example shows that the simple self-tuning feedback (3.1), (3.3) will perform very well. After a short transient it will give almost the same performance as a regulator based on exact knowledge of the system parameters. It is therefore of interest to analyse the performance of the simple regulator described by the equations (3.1) and (3.3).

4. ANALYSIS

The control law (3.1), (3.3) is a nonlinear feedback. It will now be analysed what happens when the regulator is applied to a process described by

$$y(t+1) + ay(t) = u(t) + n(t), \tag{4.1}$$

where n is a disturbance.

If the disturbance is bounded in the sense

$$\frac{1}{t}\sum_{k=1}^{t} n^2(k), \tag{4.2}$$

then the mean square output of the closed loop system

$$\frac{1}{t}\sum_{k=1}^{t} y^2(k) \tag{4.3}$$

is also bounded.

This statement is shown by contradiction. Assume therefore that (4.3) is not bounded. Equations (3.3) and (4.1) give

$$\hat{\theta}(t+1) = -\Big[\sum_{k=1}^{t}[y(k+1) - u(k)]\,y(k)\Big] \Big/ \sum_{k=1}^{t} y^2(k) = a - \sum_{k=1}^{t} n(k)y(k) \Big/ \sum_{k=1}^{t} y^2(k).$$

Schwartz' inequality gives

$$|\hat{\theta}(t+1) - a| \le \sqrt{\Big[\frac{1}{t}\Sigma n^2(k)\Big]\Big[\frac{1}{t}\Sigma y^2(k)\Big]} \Big/ \frac{1}{t}\Sigma y^2(k) = \sqrt{\frac{1}{t}\Sigma n^2(k)} \Big/ \sqrt{\frac{1}{t}\Sigma y^2(k)}.$$

Since (4.2) is bounded and since it was assumed that (4.3) was unbounded then $\hat{\theta}$ is arbitrarily close to a for large t. The closed loop system is then given by

$$y(t+1) + [a - \hat{\theta}(t)]\,y(t) = n(t).$$

The solution to this linear equation can be written as

$$y(t) = \psi(t,t_0)\, y(t_0) + \sum_{k=t_0}^{t-1} \psi(t,k+1)\, n(k),$$

where the fundamental solution ψ is given by

$$\psi(t,k) = \begin{cases} 1 & k = t \\ \prod_{i=k}^{t-1} [a - \hat{\theta}(i)] & k < t . \end{cases}$$

Since $\hat{\theta}(t)$ is arbitrarily close to a for large t, we have

$$|\psi(t,k)| < \varepsilon^{t-k}, \qquad t > k .$$

If n is bounded in the mean square sense then y is also bounded in the mean square and we have a contradiction.

The self-tuning regulator (3.1), (3.3) will thus always stabilize the system (4.1) in the mean square sense. It can be shown in this case that the parameter estimate $\hat{\theta}(t)$ will also converge. If $\hat{\theta}(t)$ converges as $t \to \infty$ it is easy to find the convergence point. The normal equations can be written

$$\sum_{k=1}^{t} y(k+1)y(k) = \hat{\theta}(t+1) \sum_{k=1}^{t} y^2(k) + \sum_{k=1}^{t} y(k)u(k) .$$

Equation (3.1) now gives

$$\frac{1}{t} \sum_{k=1}^{t} y(k+1)y(k) = \frac{1}{t} \sum_{k=1}^{t} [\hat{\theta}(t+1) - \hat{\theta}(k)]\, y^2(k) .$$

The right member converges to zero as $t \to \infty$ because $\hat{\theta}(t)$ converges and (4.3) is bounded. It is thus shown that if the parameter estimates converge then

$$\lim_{t\to\infty} \frac{1}{t} \sum_{k=1}^{t} y(k+1)y(k) = 0. \tag{4.4}$$

The self-tuning regulator (3.1), (3.3) thus attempts to make the correlation of the closed system output zero at lag 1. Assuming that the process to be controlled is given by (2.1) it now follows that it is only one value of $\hat{\theta}$ for which (4.3) is bounded and (4.4) holds, namely

$$\theta = a - c.$$

It has thus been established that the regulator (3.1), (3.3) is self-tuning for the system (2.1) and the minimum variance criterion. The analysis can be extended to the case when $b \neq 1$. The condition required is that $0 < b < 2$. The results can be extended to control

of an n:th order CARMA process. Additional conditions are then required both for stability and convergence. There are also cases where the parameter estimates do not converge.

5. CONCLUSIONS

The self-tuning regulator given by (3.1) and (3.3) is much simpler than the optimal dual regulator. The performance of the dual regulator will be better than the self-tuning regulator. It will, however, be worse than the performance of the regulator based on exact knowledge of the system parameters. In the simple example there is a difference in the transient performances of the different regulators say in the first 20 steps in Fig. 8. After about 20 steps there is, however, little difference between the accumulated loss of the self-tuning regulator and the regulator based on exact knowledge of the parameters and consequently little room for improvement.

There are many different possibilities to design self-tuning regulators. Other recursive parameter estimation schemes than least squares can be used. They can be combined with many different procedures for control design. It is also possible to take uncertainties of the estimates into account and also to incorporate some approximative dual control features. Some progress has been made towards the analysis and understanding of such control laws. There are, however, still many open problems.

6. REFERENCES

Self-tuning regulators have been used to control industrial processes. Several applications are given in the paper [1] which also contains a review of the theory and many references. The convergence analysis is given in [2].

[1] K J Åström, U Borisson, L Ljung, and B Wittenmark: Theory and Applications of Self-Tuning Regulators. Automatica 13 (1977) 457-476.

[2] L Ljung: Analysis of Recursive Stochastic Algorithms. IEEE Trans AC-22 (1977), 551-575.

Tryckbaren
Lund 1978