INT. J. CONTROL, 1986, VOL. 43, NO. 1, 1-12

A hamiltonian formulation of risk-sensitive linear/quadratic/gaussian control

P. WHITTLE† and J. KUHN†

We consider the optimal-control problem with time-homogeneous linear (and in general non-Markov) dynamics and risk-sensitive criterion. Appeal to the extremal principle prescribed by the risk-sensitive certainty equivalence principle (RSCEP) yields a symmetric equation system, indicating that the extended hamiltonian formulation generalizes naturally to the risk-sensitive case. The conjugate variable of a hamiltonian formulation now has an interpretation in terms of forecasted process noise, and the RSCEP in fact provides a stochastic maximum principle for which all variables have a clear interpretation and the desired measurability properties. In the infinite-horizon case (meaningful under generalized controllability conditions) optimal control is determined explicitly in terms of a canonical factorization. For the case of imperfect process observation, the RSCEP leads to coupled-equation systems that can again be solved in terms of canonical factorizations in the time-invariant (stationary) case.

1. Introduction

This paper considers the simultaneous generalization in two directions of the conventional linear/quadratic/gaussian (LQG) control problem. It could be regarded as a study of those aspects of the conventional problem that generalize naturally. The fact that these generalizations can be both natural and unevident illuminates even the conventional case.

It is usual these days to study the LQG problem in state-structured form, so that both process and observation relations have the linear Markov form given in (6) and (7) of this paper. However, there are many general features of the problem that do not depend upon state structure, and indeed economists in particular habitually find themselves forced to use models that are not in state form. There are some further points made in § 3 which collectively lead us to adopt general time-invariant linear dynamics of the form given in (8) and (9). With the introduction of appropriate auxiliary variables, we can then phrase the optimality conditions as the determination of the saddle-point of a quadratic form. These conditions yield two systems of coupled equations, each characterized by a symmetric operator matrix. Under stationary operation (and the useful possibility of this implies generalized controllability/observability hypotheses), optimal controls and process variable estimates are determined explicitly in terms of the canonical factors of these two operator matrices.

It is this characterization we refer to as a `hamiltonian formulation'. It is a formulation recurrently hinted at in the literature, in both control and estimation contexts. Attack on a state-structured deterministic control problem by the maximum principle yields a hamiltonian formulation immediately. Graves and Telser (1967) and Telser and Graves (1968, 1972) indicated how this formalism would extend to the

Received 30 April 1985.
† Statistical Laboratory, Department of Pure Mathematics and Mathematical Statistics, University of Cambridge, 16 Mill Lane, Cambridge

non-state case. In the estimation context, one sees manifestations of the approach in, for example, the papers by Ljung *et al.* (1976), Friedlander *et al.* (1976), Friedlander *et al.* (1977), van Dooren (1981) and Adams *et al.* (1984).

However, most of these treatments involve the hamiltonian formulation somewhat incidentally, and do not exhibit its full structure. For example, structure is generally lost by consideration of some reduced form (e.g. by elimination of the control variable) or, indeed, by an unnecessary restriction to state-structured models.

The full-blooded hamiltonian formulation, for both control and estimation, was stated in Whittle (1983) and the various reductions of it which yield the conventional Riccati equations, etc., indicated. Essentially, the formulation provides a stochastic maximum principle for the LQG case without the assumption of state structure and with imperfect observation. It should be noted that the canonical factorization spoken of is not, for estimation, the familiar factorization associated with optimal (Wiener) filtration. It is a factorization dual to the Wiener factorization in that it concerns information matrices rather than convariance matrices, and is associated with a characterization of estimates by maximum likelihood rather than by minimum variance properties.

The second generalization we consider is the replacement of the usual expected quadratic-cost criterion by a risk-sensitive criterion of exponential-quadratic form. This is a relaxation desirable for several reasons, for which results in special cases have been obtained, notably by Jacobson (1973, 1977) and also by Speyer *et al.* (1974), Speyer (1976) and Kumar and van Schuppen (1981). However, the general case of a state-structured model with imperfect state observation resisted solution until Whittle (1981) showed that the conventional risk-neutral treatment had a natural (although unevident) generalization. The conclusions of that paper have been transferred to the case of continuous time by Bensoussan and van Schuppen (1984), although in such a way as to imply that they are giving the first general solution.

What now emerges in the present case is that the hamiltonian treatment of the non-state case also extends naturally when the criterion is modified to the risk-sensitive form given in (2). This rather remarkable conclusion in fact illuminates the conventional (risk-neutral) case. A natural extremal principle is provided by a risk-sensitive form of the certainty equivalence principle (see §2). Furthermore, the conjugate or auxiliary variables formerly introduced as Lagrangian multipliers now turn out to be related to effective estimators at a given time of the course of process and observation noise (see the comments after the proof of Theorem 4.1).

One can say indeed that the risk-sensitive certainty equivalence principle (CEP) provides the natural stochastic maximum principle for this model, with no problems of interpretation or measurability.

The calculations in this paper determine optimal procedures over a finite time interval: i.e. assuming a finite past for observation and a finite future for action. The formal limit versions over infinite time are clear, and it is these which have a point, in view of their succinct characterization in terms of canonical factors. However, the validity of this limit will hold only if the minimal average cost (with the appropriate risk-sensitive interpretation of this term) has a limit, and optimal estimation and control procedures become time invariant. This will be true only if generalized versions of controllability/observability hypotheses are satisfied, ensuring that average cost can indeed be held to a bounded value. An analysis of these points would constitute a major investigation on its own, with a number of features that would emerge only when one allows risk-sensitivity. These we hope to address later; but for the moment we deduce only the main structural features of the analysis, which are substantial enough in themselves.

2. The risk-sensitive certainty equivalence principle (RSCEP)

The validity of an appropriate certainty equivalence principle turns out to be crucial. The classic risk-neutral principle (due originally to Theil 1957) has a direct but unobvious risk-sensitive version.

State structure is irrelevant; we give the principle in its most general finite-horizon form. We make the following assumptions.

(*a*) The actions $\mathbf{u}_t$ to be taken over a finite horizon $(0 \leqslant t \leqslant h)$ take values in finite-dimensional vector spaces.

(*b*) The action $\mathbf{u}_t$ can be a function only of observables W_t at t: previous actions $\mathbf{U}_{t-1} = \{\mathbf{u}_0, \mathbf{u}_1, \ldots, \mathbf{u}_{t-1}\}$ and observation history $\mathbf{Y}_t = \{\mathbf{y}_0, \mathbf{y}_1, \ldots, \mathbf{y}_t\}$, $0 \leqslant t \leqslant h$.

(*c*) The cost function $\mathscr{C}$ is a quadratic function $\mathbf{Q}(\mathbf{U}_{h-1}, \xi)$ of the control sequence $\mathbf{U}_{h-1} = \{\mathbf{u}_0, \mathbf{u}_1, \ldots, \mathbf{u}_{h-1}\}$ and an exogeneous noise vector ξ, positive definite in $\mathbf{U}_{h-1}$ for all ξ.

(*d*) ξ is normally distributed with known parameters, independent of policy.

(*e*) Each $\mathbf{y}_t$ is a policy-independent linear function of ξ.

The vector ξ is considered to embody all exogenous stochastics of the problem: process noise, observation noise and the stochastics of a reference path that one may wish the controlled system to follow. The actual observations $\mathbf{y}$ will in general depend upon control actions, but it is enough that these can be corrected to exhibit observations that are effectively linear functions of ξ.

The conventional (risk-neutral) criterion is that one chooses a policy π to minimize $\mathbf{E}_\pi(\mathscr{C})$, where $\mathbf{E}_\pi$ is the expectation operator induced by π. The classic certainty equivalence principle is then that the optimal value of $\mathbf{u}_t$ is determined by minimizing $\mathbf{Q}(\mathbf{U}_{h-1}, \xi^{(t)})$ with respect to $\mathbf{u}_t, \mathbf{u}_{t+1}, \ldots, \mathbf{u}_{h-1}$, where

$$\xi^{(t)} = \mathbf{E}(\xi | \mathbf{Y}_t) \tag{1}$$

is the optimal estimator of ξ based upon W_t ('optimal' in that it has minimum mean-square error; it is also the maximum likelihood estimate).

Note that the certainty equivalence principle has the following two features.

(i) *Conversion to free form*. A minimization with respect to functions $\mathbf{u}_\tau(W_\tau)$, $(\tau \geqslant t)$, constrained in that $\mathbf{u}_\tau$ may depend upon only W_τ, is replaced by a free minimization with respect to constants $\mathbf{u}_\tau$.

(ii) *Separation*. Optimization of estimation and control are separated in that the estimate (1) is derived without reference to the determination of $\mathbf{u}_t$, and $\mathbf{u}_t$ are determined as they would be in the full-information case, with the simple substitution of $\xi^{(t)}$ for ξ.

Suppose now that the criterion that $\mathbf{E}_\pi(\mathscr{C})$ be minimized is modified to the criterion that

$$\gamma_\pi(\theta) = -\frac{2}{\theta} \log\, [\mathbf{E}_\pi \exp - \theta\mathscr{C}/2] \tag{2}$$

be minimized with respect to π. Here θ is a scalar parameter, the *risk-sensitivity* parameter. The case $\theta = 0$ corresponds to the *risk-neutral* case

$$\gamma_\pi(0) = \mathbf{E}_\pi(\mathscr{C})$$

In the case $\theta > 0$ the optimizer is *risk-seeking*; he is more concerned to reduce the frequent occurrence of moderate values of $\mathscr{C}$ than the occasional occurrence of large values. In the case $\theta < 0$ the reverse is the case, and the optimizer is *risk-averse*.

We shall term a problem in which one minimizes criterion (2) under linear/gaussian dynamics an LEQG problem, the EQ standing for 'exponential–quadratic'.

It is remarkable that the features of the well-known risk-neutral (LQG) case have a version for the risk-sensitive (LEQG) case, although these are transformed sufficiently that they are not immediately evident. To take the analysis beyond the point reached by Jacobson (1973, 1977) required a certainty equivalence principle, which Whittle proved, first for the state-structured case (1981, 1982, 1983), and then for the general case (1985). However, the line of proof for the general case is quite clear from those proofs already published for the state-structured case.

Suppose that the exogenous noise vector $\boldsymbol{\xi}$ is normally distributed with zero mean and covariance matrix $\mathbf{V}$. Define

$$\mathscr{D} = \xi' \mathbf{V}^{-1} \xi \tag{3}$$

recognizable as occurring in the exponent of the $\boldsymbol{\xi}$-density, and define the *total stress*,

$$\mathscr{S} = \mathscr{C} + \frac{1}{\theta}\mathscr{D} \tag{4}$$

This is a spontaneously-occurring combination when one evaluates the expectation in (2). $\mathscr{C}$ is the component of stress owing to cost (e.g. to departures of $\mathbf{u}$ from zero) and $\mathscr{D}$ the component owing to implausibility, (i.e. to departures of $\boldsymbol{\xi}$ from zero).

We shall use the term *extremization* to denote an operation that is minimization when $\theta \geqslant 0$ and maximization when $\theta < 0$. The following theorem is quoted from Whittle (1985); see Whittle (1981) for a proof in the state-structured case.

Theorem 2.1. *The risk-sensitive certainty equivalence principle*

Suppose that one wishes to choose a policy π to minimize criterion (2). Then under assumptions (a)–(e) above, the optimal value of $\mathbf{u}_t$ is determined by simultaneously minimizing $\mathscr{S}$ with respect to $\mathbf{u}_t, \mathbf{u}_{t+1}, \ldots, \mathbf{u}_{h-1}$ and extremizing it with respect to $\mathbf{y}_{t+1}, \mathbf{y}_{t+2}, \ldots, \mathbf{y}_h$. In words: one minimizes stress with respect to all decisions currently unmade and extremizes it with respect to all quantities currently unobservable.

This principle really does reduce to the risk-neutral version as $\theta \downarrow 0$. The stress-extremizing value of unobservables tends to that minimizing $\mathscr{D}$ for a given value of $\mathbf{Y}_t$; this leads exactly to the estimate $\boldsymbol{\xi}^{(t)}$ of $\boldsymbol{\xi}$.

One can ask in what sense this is a certaintly equivalence principle. It certainly has the property (i) above, of conversion to free form. That is, minimization of γ_π with respect to policies π, i.e. with respect to functions $\mathbf{u}_\tau(W_\tau)$, has been replaced by a free minimization/extremization of stress with respect to relevant decisions/unobservables.

It does not have property (ii), of separation. Determination of optimal control $\mathbf{u}_t$ and of an effective current estimate of $\boldsymbol{\xi}$ are intertwined in the minimization/extremization of stress.

This fact is inevitable. If 'separation' is a meaningful concept at all, it must surface in another and less evident form. This it does, as demonstrated by Whittle (1981). In the state-structured case, one can evaluate extremal values of past stress and future stress at time t separately, conditional on the true (but in general unknown) value of current state $\mathbf{x}_t$. This achieves separation, in that one has two decoupled calculations of the familiar recursive form that separately yield an optimal condensation of data and an optimal determination of control, both parametrized by the 'pivot' $\mathbf{x}_t$. Let the optimal control thus determined be $\mathbf{u}_t(\mathbf{x}_t)$. One now recouples these two calculations by choosing $\mathbf{x}_t$ to extremize total stress, already determined parametrically in terms of $\mathbf{x}_t$. If the estimate thus yielded is $\mathbf{x}_t^{(t)}$, then the optimal control is $\mathbf{u}_t(\mathbf{x}_t^{(t)})$. This is very obviously a certainty-equivalence statement. Moreover, separation holds in the sense that optimization of estimation and control have been decoupled by parametrization of those sub-problems in terms of $\mathbf{x}_t$.

The effect of risk-sensitivity is that the optimizer behaves optimistically or pessimistically according as $\theta > 0$ or $\theta < 0$, in that he behaves as if unobservables would take values to his advantage or disadvantage, respectively. For values of θ less than a negative critical value, he is so pessimistic that $\gamma_\pi(\theta)$ is infinite for all π—a phenomenon that one might characterize as neurotic breakdown.

It scarcely needs saying that the definition (4) of stress is not arbitrary; the validity of the RSCEP is sufficient evidence of this.

3. Formulation

We suppose process variable $\mathbf{x}$, control variable $\mathbf{u}$ and observation $\mathbf{y}$ have respective values $\mathbf{x}_t$, $\mathbf{u}_t$ and $\mathbf{y}_t$ at time t. We shall also suppose time to be discrete, so that t takes values in Z, the set of signed integers. The quantities $\mathbf{x}$, $\mathbf{u}$ and $\mathbf{y}$ are taken as column vectors of appropriate fixed dimension. We shall assume the conventional quadratic cost function

$$\mathscr{C} = \sum_{t=h_1}^{h_2-1} (\mathbf{x}'\mathbf{R}\mathbf{x} + \mathbf{u}'\mathbf{S}\mathbf{x} + \mathbf{x}'\mathbf{S}'\mathbf{u} + \mathbf{u}'\mathbf{Q}\mathbf{u})_t + (\mathbf{x}'\mathbf{\Pi}\mathbf{x})^{h_2} \tag{5}$$

where the notation $(\ )_t$ implies that all quantities inside the bracket are evaluated at time t. The times h_1 and h_2 are then respectively the initial point and the horizon point, at which costing and decision respectively begin and cease.

In the conventional state-strucutred case, one assumes process and observation relations of the form

$$\mathbf{x}_t = \mathbf{A}\mathbf{x}_{t-1} + \mathbf{B}\mathbf{u}_{t-1} + \boldsymbol{\varepsilon}_t \tag{6}$$

$$\mathbf{y}_t = \mathbf{C}\mathbf{x}_{t-1} + \boldsymbol{\eta}_t \tag{7}$$

where the process/observation noise process $\{(\boldsymbol{\varepsilon}, \boldsymbol{\eta})_t\}$ is white. The aim of the present paper is to extend the treatment to the case of general linear time-invariant dynamics, when (6) and (7) are replaced by

$$\mathbf{A}(\mathbf{T})\mathbf{x}_t + \mathbf{B}(\mathbf{T})\mathbf{u}_t = \boldsymbol{\varepsilon}_t \tag{8}$$

$$\mathbf{y}_t + \mathbf{C}(\mathbf{T})\mathbf{x}_t = \boldsymbol{\eta}_t \tag{9}$$

Here $\mathbf{T}$ is the backwards translation operator (e.g. $\mathbf{T}^j\mathbf{x}_t = \mathbf{x}_{t-j}$) and the matrix operator coefficients have the form

$$\left.\begin{aligned} \mathbf{A}(\mathbf{T}) &= \sum_{j=1}^{\infty} \mathbf{A}_j\mathbf{T}^j \\ \mathbf{B}(\mathbf{T}) &= \sum_{j=1}^{\infty} \mathbf{B}_j\mathbf{T}^j \\ \mathbf{C}(\mathbf{T}) &= \sum_{j=0}^{\infty} \mathbf{C}_j\mathbf{T}^j \end{aligned}\right\} \tag{10}$$

We assume $\mathbf{A}_0 = \mathbf{I}$, so that (8) really does determine $\mathbf{x}$ recursively in terms of past $\mathbf{x}$, $\mathbf{u}$ and process noise. Note that we effectively assume that $\mathbf{B}_0 = 0$. We again assume the process $\{(\boldsymbol{\varepsilon}, \boldsymbol{\eta})_t\}$ to be white, with instantaneous covariance matrix

$$\operatorname{cov}(\varepsilon_t, \eta_t) = \begin{bmatrix} \mathbf{N} & \mathbf{L} \\ \mathbf{L}' & \mathbf{M} \end{bmatrix}$$

although both this assumption and the corresponding assumption (5) on the structure of the cost function will be relaxed in § 6.

We are essentially assuming operations over the finite time interval $h_1 \leqslant t \leqslant h_2$, in that the cost function (9) is specified over this period, and $\mathbf{u}_t$ must be chosen for $h_1 \leqslant t < h_2$. Initial conditions must be specified; these must take the form of the specification of the joint distribution of the process variable $\mathbf{x}_t$ for $t < h_1$. The natural formulation will turn out to be the specification of the actual gaussian density

$$f(\mathbf{x}_\tau;\ \tau < h_1) \propto \exp\left[-\frac{1}{2}\sum_{j,k<h_1}\sum (\mathbf{x}_j - \bar{\mathbf{x}}_j)'\mathbf{P}_{jk}(\mathbf{x}_k - \bar{\mathbf{x}}_k)\right] \tag{11}$$

That is, specification of information matrices rather then covariance matrices.

The conventional state-structured model is, of course, that for which the sums in (10) terminate at $j = 1$. If they terminate at $j = p$, one may say that one has p^{th} order dynamics.

The first justification for adopting the more general model is that this indeed may be the right model. Economists habitually work in terms of models with high-order dynamics. It is, of course, true that a model with dynamics of finite order p can be reduced to one with first-order dynamics. However, to do so not only multiplies the dimensionality of the 'description' space by p, but in fact loses some of the structure of the problem. The reduced problem still has special structure as a Markov problem, in that noise is injected in a special way and state-derivations costed in a special way, and one ignores this if one simply appeals to the well-established formalism of the general Markov case.

A second justification is a mathematical one, which should be taken as a serious general guide by even the most calloused of practitioners. Just as Markov models have their natural recursive formalism, so do time-invariant linear models have their own natural generating function formalism, especially effective in the infinite-horizon (stationary) limit. Moreover, this can be put into an especially symmetric and succinct form—what we term below the hamiltonian formulation. Yet more—and this perhaps surprises—the hamiltonian formulation accommodates the generalization to risk-sensitivity in the most natural way possible. This fact on its own justifies an account of the analysis.

Actually, vestiges of a Markov formulation remain, in that we suppose process and observation noise to be white and the cost function to have the simple form given in (5). It is in fact natural to relax these assumptions; see § 6. However, the argument is simplified if they are retained until then.

4. Optimization with perfect observation

We now specialize the formulation of § 3 by assuming perfect observation of the process variable, in that (9) reduces to

$$\mathbf{y}_t = \mathbf{x}_{t-1} \tag{12}$$

The assumption $\mathbf{y}_t = \mathbf{x}_t$ might seem to correspond more to 'perfect observation', but the formalism works out more naturally under assumption (12); observation of the process variable at unit lag. The same is of course true for the Kalman filter in the state-structured case; the filter has its familiar form only under the assumption of a unit lag in observation, in that the observation relation has the form (7). We return to this point at the end of the section.

Let us write the operator $\mathbf{A}(\mathbf{T}) = \Sigma\, \mathbf{A}_j \mathbf{T}^j$ simply as $\mathbf{A}$, and the associated operator $\mathbf{A}(\mathbf{T}^{-1})' = \Sigma\, \mathbf{A}_j' \mathbf{T}^{-j}$ as $\bar{\mathbf{A}}$. The process equation (8) can then be written

$$(\mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u} - \varepsilon)_t = 0 \tag{13}$$

We have

$$\mathscr{S} = \mathscr{C} + \frac{1}{\theta}\sum_{h_1}^{h_2} (\varepsilon' \mathbf{N}^{-1} \varepsilon)_t \tag{14}$$

where $\mathscr{C}$ is given by (5).

At time t one obtains the best provisional determination of undetermined quantities by minimizing stress with respect to decisions as yet unmade and extremizing it with respect to random variables as yet unobserved. These are indeed 'best' in virtue of the RSCEP. Let us for greater explicitness refer to them as 'minimal-stress' determinations, although they are in fact obtained by combined minimization/extremization. The minimal-stress determination of a quantity ξ at time t will be denoted $\xi^{(t)}$. If ξ is a known function of W_t then $\xi^{(t)}$ will be simply ξ. We shall extend this notation to vectors, so that, for example,

$$\begin{bmatrix} \mathbf{x} \\ \mathbf{u} \\ \lambda \end{bmatrix}_\tau^{(t)} = \begin{bmatrix} \mathbf{x}_\tau^{(t)} \\ \mathbf{u}_\tau^{(t)} \\ \lambda_\tau^{(t)} \end{bmatrix}$$

Theorem 4.1.

The optimal value $\mathbf{u}_t(W_t)$ of $\mathbf{u}_t$ is the quantity $\mathbf{u}_t^{(t)}$ determined by solution of the equations

$$\begin{bmatrix} \mathbf{R} & \mathbf{S}' & \bar{\mathbf{A}} \\ \mathbf{S} & \mathbf{Q} & \bar{\mathbf{B}} \\ \mathbf{A} & \mathbf{B} & -\theta\mathbf{N} \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \mathbf{u} \\ \lambda \end{bmatrix}_\tau^{(t)} = 0 \quad t \leqslant \tau < h_2 \tag{15}$$

with terminal conditions

$$\lambda_{h_2}^{(t)} + \mathbf{\Pi}\mathbf{x}_{h_2}^{(t)} = 0 \tag{16}$$

$$\lambda_\tau^{(t)} = 0 \quad \tau > h_2 \tag{17}$$

The quantity $\lambda_\tau^{(t)}$ has the interpretation

$$\varepsilon_\tau^{(t)} = \theta \mathbf{N} \lambda_\tau^{(t)} \tag{18}$$

The shift operator $\mathbf{T}$ occurring in the operators $\mathbf{A}$, $\bar{\mathbf{A}}$ etc. is considered to operate upon the subscript τ, not the superscript t, so that the equation system (15) written in full would be

$$\left.\begin{aligned} \mathbf{R}\mathbf{x}_\tau^{(t)} + \mathbf{S}'\mathbf{u}_\tau^{(t)} + \sum_j \mathbf{A}_j'\boldsymbol{\lambda}_{\tau+j}^{(t)} &= 0 \\ \mathbf{S}\mathbf{x}_\tau^{(t)} + \mathbf{Q}\mathbf{u}_\tau^{(t)} + \sum_j \mathbf{B}_j'\boldsymbol{\lambda}_{\tau+j}^{(t)} &= 0 \\ \sum_j \mathbf{A}_j\mathbf{x}_{\tau-j}^{(t)} + \sum_j \mathbf{B}_j\mathbf{u}_{\tau-j}^{(t)} - \theta\mathbf{N}\boldsymbol{\lambda}_\tau^{(t)} &= 0 \end{aligned}\right\} \quad (15)$$

Proof

Substitute for $\boldsymbol{\varepsilon}$ into expression (14) for the stress from the process equation (13). The stationarity conditions for $\mathscr{S}$ with respect to unobservables $\mathbf{x}$ and unformed decisions $\mathbf{u}_\tau$, $t \leqslant \tau < h_2$, then provide respectively the first two relations of (15), with $\boldsymbol{\varepsilon}_\tau$ expressed rather in terms of $\boldsymbol{\lambda}_\tau$ by (18). The final relation of (15) follows from the process equation (13). Relation (16) expresses a stationarity condition with respect to the final unobservable $\mathbf{x}_{h_2}$, and relations (16) are effective terminal conditions implied by the fact that $\boldsymbol{\varepsilon}_\tau$ does not occur in (14) for $\tau > h$. (Alternatively, the sum in (14) could have been continued, but the stationarity condition would have yielded the estimate $\boldsymbol{\varepsilon}_\tau^{(t)} = 0$ for $\tau > h_2$.) □

Relations (15) are the symmetric set of relations associated with the 'hamiltonian formulation' deduced in the risk-neutral case by Whittle (1983) and, in some special risk-neutral cases, by other authors (see the references to Telser and Graves). In this risk-neutral case $\theta = 0$, $\boldsymbol{\varepsilon}_\tau^{(t)} = \mathbf{E}(\boldsymbol{\varepsilon}_\tau | W_t) = 0$ $(\tau \geqslant t)$ and relation (18) yields no interpretation of λ. In fact, λ then has the interpretation of a lagrangian multiplier associated with a noise-free process equation (noise-free for $\tau \geqslant t$ by the CEP). It is interesting that this multiplier, introduced for formal reasons, has the interpretation (18) in the risk sensitive case. Indeed, the risk-neutral case now appears as degenerate in that this interpretation fails, although the theorem remains valid.

There are indeed treatments of risk-neutral models in the literature that express estimates of noise etc. in terms of dual (Lagrange multiplier) variables; see for example Adams *et al.* (1984). These are 'smoothing' models, in that observations have been taken later in time that give information on the noise variables in question. In the present case this is not so: the noise is future noise, and could be predicted only by its zero mean value in a risk-neutral treatment. It is only in a risk-sensitive treatment that cost pressures supplement statistical information to produce the less trivial prediction (18).

Relations (15)–(17) indeed constitute the optimality equations of a stochastic maximum principle. Being linear, their solutions are unique except in degenerate cases, and will yield the optimal solution for all non-negative θ and for all θ for which stress possesses a saddle-point (i.e. for all θ exceeding the breakdown value). There is no difficulty about measurability: all quantities $\boldsymbol{\xi}^{(t)}$ are determined as functions of W_t by these relations. Moreover, at least for $\theta \neq 0$ the conjugate variables $\boldsymbol{\lambda}_\tau$ (or rather their determinations $\boldsymbol{\lambda}_\tau^{(t)}$) have the clear interpretation (18). Further, rather than the principle's being derived by location of the saddle-point of a formal lagrangian form, it is derived from the minimization/extremization of stress. Indeed, the RSCEP *is* the stochastic maximum principle for this LEQG case.

Let us now write (15) as

$$\boldsymbol{\phi}(\mathbf{T})\boldsymbol{\zeta}^{(t)} = 0 \quad t \leqslant \tau < h_2 \qquad (19)$$

implicity thus defining the vector $\boldsymbol{\zeta}_\tau$ and the matrix of distributed lag operators $\boldsymbol{\phi}(\mathbf{T})$. One would hope that in the infinite horizon case $h_2 \to +\infty$, relation (19) would continue to hold as

$$\boldsymbol{\phi}(\mathbf{T})\zeta_\tau^{(t)} = 0 \quad t \geqslant t \tag{20}$$

with effective terminal condition

$$\zeta_\tau^{(t)} \to 0 \tag{21}$$

at some suitable rate as $\tau \to \infty$. That is $\{\zeta_\tau^{(t)}; \tau \geqslant t\}$ is the minimal-stress estimate of the future of the process as seen from time t, and relation (21) implies that this *can* be brought down to the ideal value, zero, at a suitable rate. This assumption can be regarded as a generalized controllability condition. One would of course like more explicit conditions on the model that would ensure this property. To find these constitutes a substantial study, beyond the scope of this paper. However, note that if this controllability assumption is justified, it implies as explicit a stationary solution of the optimization problem as one can expect in a general context.

Suppose that $\boldsymbol{\phi}$ has a canonical factorization

$$\boldsymbol{\phi}(z) = \boldsymbol{\phi}^-(z)\boldsymbol{\phi}^+(z) \tag{22}$$

where $\boldsymbol{\phi}^+(z)$ has a unilateral expansion in non-negative powers of z

$$\boldsymbol{\phi}^+(z) = \sum_0^\infty \boldsymbol{\phi}_j^+ z^j$$

valid on $|z| = 1$, as also does its inverse, and $\boldsymbol{\phi}^-$ and its inverse have corresponding expansions in non-positive powers. (In fact, the symmetric character of $\boldsymbol{\phi}$ implies that $\boldsymbol{\phi}^- = \bar{\boldsymbol{\phi}}^+$.) From (20), (21) one could then legitimately draw the conclusion

$$\boldsymbol{\phi}^+(\tau)\zeta_\tau = 0 \quad \tau \geqslant t \tag{23}$$

This relation at $\tau = t$ will give an explicit expression for $\mathbf{u}_t$ in terms of past $\mathbf{x}$ and $\mathbf{u}$: the closed-loop expression for the optimal control at time t. Past values of $\boldsymbol{\lambda}$ do not occur, because they do not occur in (15).

Whether $\boldsymbol{\phi}$ possesses a canonical factorization (22) is related to the question of the validity of (20) and (21).

If $\mathbf{x}_t$ is itself observable at time t, then the first relation of (15) is valid only for $\tau > t$, and the treatment requires minor modifications. That it is more natural to suppose a unit lag in observation is already familiar from the conventional treatment of the Kalman filter.

5. Optimization with imperfect observation

We now suppose the process and observation equations (8) and (9) with white-noise inputs and suppose the prior information on $\mathbf{x}_\tau, \tau < h_1$, expressed by specification of the gaussian prior density (11). In the total stress $\mathscr{C} + \frac{1}{\theta}\mathscr{D}$ the component $\mathscr{C}$ now has evaluation (5) and $\mathscr{D}$ has the evaluation

$$\mathscr{D} = \sum_{j,k<h_1}\sum \mathbf{x}_j'\mathbf{P}_{jk}\mathbf{x}_k + \sum_{\tau=h_1}^{h_2} \begin{bmatrix} \varepsilon \\ \boldsymbol{\eta} \end{bmatrix}_\tau' \begin{bmatrix} \mathbf{N} & \mathbf{L}' \\ \mathbf{L} & \mathbf{M} \end{bmatrix}^{-1} \begin{bmatrix} \varepsilon \\ \boldsymbol{\eta} \end{bmatrix}_\tau \tag{24}$$

Theorem 5.1.

Under the assumptions described above, the optimal value of control $\mathbf{u}_t$ is equal to $\mathbf{u}_t^{(t)}$, which quantity is determined by solution of two simultaneous equation systems. The first of these is (15) with terminal conditions (16) and (17), and the second of these is

$$\begin{bmatrix} -\theta\mathbf{N} & -\theta\mathbf{L} & \mathbf{A} \\ -\theta\mathbf{L}' & -\theta\mathbf{M} & \mathbf{C} \\ \bar{\mathbf{A}} & \bar{\mathbf{C}} & \mathbf{R} \end{bmatrix} \begin{bmatrix} \boldsymbol{\lambda} \\ \boldsymbol{\mu} \\ \mathbf{x} \end{bmatrix}_\tau^{(t)} + \begin{bmatrix} \mathbf{Bu} \\ \mathbf{y} \\ \mathbf{S}'\mathbf{u} \end{bmatrix}_\tau = 0 \quad h_1 \leqslant \tau < t \tag{25}$$

with initial and terminal conditions

$$\sum_{k<h_1} \mathbf{P}_{\tau k}\mathbf{x}_k^{(t)} + \bar{\mathbf{A}}\boldsymbol{\lambda}_\tau^{(t)} + \bar{\mathbf{C}}\boldsymbol{\lambda}_\tau^{(t)} = 0 \quad \tau < h_1 \tag{26}$$

$$\boldsymbol{\mu}_\tau^{(t)} = 0 \qquad \tau \geqslant t \tag{27}$$

The quantities $\boldsymbol{\lambda}$ and $\boldsymbol{\mu}$ have the interpretations implied by

$$\theta \begin{bmatrix} \mathbf{N} & \mathbf{L} \\ \mathbf{L}' & \mathbf{M} \end{bmatrix} \begin{bmatrix} \boldsymbol{\lambda} \\ \boldsymbol{\mu} \end{bmatrix}_\tau^{(t)} = \begin{bmatrix} \boldsymbol{\varepsilon} \\ \boldsymbol{\eta} \end{bmatrix}_\tau^{(t)} \tag{28}$$

Proof

Make the substitution

$$\begin{bmatrix} \boldsymbol{\varepsilon} \\ \boldsymbol{\eta} \end{bmatrix}_\tau' \begin{bmatrix} \mathbf{N} & \mathbf{L} \\ \mathbf{L}' & \mathbf{M} \end{bmatrix}^{-1} \begin{bmatrix} \boldsymbol{\varepsilon} \\ \boldsymbol{\eta} \end{bmatrix}_\tau = \max_{\lambda,\mu} \left[2\boldsymbol{\lambda}'\boldsymbol{\varepsilon} + 2\boldsymbol{\mu}'\boldsymbol{\eta} - \begin{bmatrix} \boldsymbol{\lambda} \\ \boldsymbol{\mu} \end{bmatrix}' \begin{bmatrix} \mathbf{N} & \mathbf{L} \\ \mathbf{L}' & \mathbf{M} \end{bmatrix} \begin{bmatrix} \boldsymbol{\lambda} \\ \boldsymbol{\mu} \end{bmatrix} \right]_\tau \tag{29}$$

in expression (24), and substitute for $\boldsymbol{\varepsilon}$ and $\boldsymbol{\eta}$ in terms of $\mathbf{x}$, $\mathbf{u}$ and $\mathbf{y}$ from (8) and (9). Writing down the stationarity conditions of the resulting expression for stress with respect to controls $\mathbf{u}_\tau$, $t \leqslant \tau < h_2$, and unobservables $\mathbf{x}_\tau$, $\tau \leqslant h_2$, and $\mathbf{y}_\tau$, $t \leqslant \tau \leqslant h_2$, we deduce the equation systems (15) and (25) and associated boundary conditions of the theorem. Relations (28) follow from the stationarity conditions with respect to $\boldsymbol{\lambda}$ and $\boldsymbol{\mu}$ implicit in (29). □

Note that the system (25) can be written in the alternative form

$$\begin{bmatrix} \mathbf{N} & \mathbf{L} & \mathbf{A} \\ \mathbf{L}' & \mathbf{M} & \mathbf{C} \\ \bar{\mathbf{A}} & \bar{\mathbf{C}} & -\theta\mathbf{R} \end{bmatrix} \begin{bmatrix} \theta\boldsymbol{\lambda} \\ \theta\boldsymbol{\eta} \\ \mathbf{x} \end{bmatrix}_\tau^{(t)} + \begin{bmatrix} \mathbf{Bu} \\ \mathbf{y} \\ \theta\mathbf{S}'\mathbf{u} \end{bmatrix}_\tau = 0 \quad h_1 \leqslant \tau < t \tag{30}$$

in which form its duality to (15) is more evident. Both systems have a 'hermitian' matrix. In the risk-neutral case $\theta \to 0$ (treated by Whittle, 1983), one can regard the two systems as decoupled, system (15) being concerned with control optimization and system (25) (or (30)) being concerned with optimization of the estimate of $\mathbf{x}_\tau$, $\tau < t$. In this case, the terms $\theta\mathbf{N}$ and $\theta\mathbf{R}$, $\theta\mathbf{S}'$ drop out in (15) and (25) respectively, and the variables $\theta\boldsymbol{\lambda}_\tau$ and $\theta\boldsymbol{\mu}_\tau$ have a zero limit for $\tau \geqslant t$ (but non-zero for $\tau < t$). Thus system (30) can be solved autonomously (i.e. estimation is independent of control) and system (15) can then be solved to give the certainty-equivalent form of control.

In the risk-sensitive case, a partial degree of decoupling can still be achieved, as demonstrated for the Markov form of the problem by Whittle (1981, 1982 and 1983).

Suppose that the system is of degree p, in that $\mathbf{A}_j = 0$ and $\mathbf{C}_j = 0$ for $j > p$. The equation systems (15)–(17) for $\tau \geqslant t$ and (25)–(27) for $\tau < \tau - p$ are then decoupled, and one can use these equations to solve for variables outside the range $t - p \leqslant \tau < t$ in terms of variables inside this range. Substituting these expressions into relations (25) for $t - p \leqslant \tau < t$, one has then a set of equations for the variables in this range of p time points, from which $\mathbf{u}_t^{(t)}$ can be determined. This procedure is equivalent to the separate determination of extremal 'past stress' and 'future stress' for given $\mathbf{x}_{t-1}, \mathbf{x}_{t-2}, \ldots, \mathbf{x}_{t-p}$, followed by determination of minimal stress estimation of these 'pivotal' quantities by extremization of the sum of partially-extremized past and future stresses (see Whittle 1981).

One hopes that under appropriate regularity conditions (to be interpreted as generalized controllability/observability conditions), one can properly make the passage $h_1 \to -\infty, h_2 \to +\infty$ and that systems (15) and (25) will then hold for $\tau \geqslant t$ and $\tau < t$, respectively, with appropriate behaviour of variables at $\pm\infty$. If we write these equations as

$$\phi\zeta_\tau^{(t)} = 0 \quad \tau \geqslant t \tag{31}$$

and

$$\psi\omega_\tau^{(t)} + \varrho_\tau = 0 \quad \tau < t \tag{32}$$

and assume canonical factorizations

$$\phi = \phi^- \phi^+$$

and

$$\psi = \psi^+ \psi^-$$

then the assumed behaviour at $\pm\infty$ will imply the validity of the reductions

$$\phi^+ \zeta_\tau^{(t)} = 0 \qquad \tau \geqslant t \tag{33}$$

and

$$\psi^- \omega_\tau^{(t)} = -(\psi^+)^{-1} \varrho_\tau \quad \tau < t) \tag{34}$$

In this form, solutions for variables outside the range $t - p \leqslant \tau < t$ are immediate, and the problem of determining $\mathbf{u}_t^{(t)}$ then reduces to solution of a system of equations over these p time points.

6. Further generalizations

If one incorporates deterministic disturbances in the process equation, or changes $\mathscr{C}$ so that $\mathbf{x}$ is required to follow a known reference signal rather than be stabilized to zero, then relation (31) becomes, like (32), inhomogeneous:

$$\phi\zeta_\tau^{(t)} + \chi_\tau = 0 \quad \tau \geqslant t$$

Here, χ_τ is calculable in terms of these known inputs. The reduction

$$\phi^+ \zeta_t^{(t)} + (\phi^-)^{-1} \chi_t = 0$$

then expresses optimal $\mathbf{u}_t$ in combined feed-back feed-forward form (i.e. as a function of past $\mathbf{x}$ and $\mathbf{u}$, and of future disturbances).

There is no necessity to assume

$$\begin{bmatrix} \varepsilon \\ \eta \end{bmatrix}_t$$

to be white; if one assumes the existence of auto-covariance generating functions

$$\mathbf{g}_{\varepsilon\eta}(\mathbf{z}) = \sum_s \mathbf{z}^s \operatorname{cov}(\varepsilon_t, \eta_{t-s})$$

etc. then relations (15) and (25) formally remain valid with the substitution of $\begin{bmatrix} \mathbf{g}_{\varepsilon\varepsilon}(\mathbf{T}) & \mathbf{g}_{\varepsilon\eta}(\mathbf{T}) \\ \mathbf{g}_{\eta\varepsilon}(\mathbf{T}) & \mathbf{g}_{\eta\eta}(\mathbf{T}) \end{bmatrix}$ for $\begin{bmatrix} \mathbf{N} & \mathbf{L} \\ \mathbf{L}' & \mathbf{M} \end{bmatrix}$. Correspondingly, if the cost component $(\mathbf{x}'\mathbf{R}\mathbf{x} + \mathbf{u}'\mathbf{S}\mathbf{x} + \mathbf{x}'\mathbf{S}'\mathbf{u} + \mathbf{u}'\mathbf{Q}\mathbf{u})_t$ is modified to $(\mathbf{x}_t\mathbf{R}(\mathbf{T})\mathbf{x}_t + \ldots)$, then relations (15) and (25) formally remain valid with the substitution of $\begin{bmatrix} \mathbf{R}(\mathbf{T}) & \mathbf{S}(\mathbf{T}^{-1})' \\ \mathbf{S}(\mathbf{T}) & \mathbf{Q}(\mathbf{T}) \end{bmatrix}$ for $\begin{bmatrix} \mathbf{R} & \mathbf{S}' \\ \mathbf{S} & \mathbf{Q} \end{bmatrix}$.

References

Adams, M. B., Willsky, A. G., and Levy, B. C., 1984, *I.E.E.E. Trans. autom. Control*, **29**, 803.

Bensoussan, A., and van Schuppen, J. H., 1984, *Proc. 23rd I.E.E.E. Conf. on Decision and Control*, Las Vegas, NV, 1473.

Friedlander, B., Kailath, T., and Ljung, L., 1976, *J. Franklin Inst.* **301**, 71.

Friedlander, B., Verghese, G., and Kailath, T., 1977, Scattering theory and linear least squares estimation. Part III: The scattering variables and the estimates. *Proc. 1977 I.E.E.E. Decision and Control Conf.*, New Orleans.

Graves, R. L., and Telser, L. G., 1967, *Econometrica*, **35**, 234.

Jacobson, D. H., 1973, *I.E.E.E. Trans. autom. Control*, **18**, 124; 1977, *Extensions of Linear-Quadratic Control, Optimization and Matrix Theory* (London: Academic Press).

Kumar, P. R., and van Schuppen, J. H., 1981, *J. math. Analysis Applic.*, **80**, 312.

Ljung, L., Kailath, T., and Friedlander, B., 1976, *Proc. Inst. elect. electron. Engrs*, **64**, 131.

Speyer, J. L., 1976, *I.E.E.E. Trans. autom. Control*, **21**, 371.

Speyer, J. L., Deyst, J., and Jacobson, D. H., 1974, *I.E.E.E. Trans. autom. Control*, **19**, 358.

Telser, L. G., and Graves, R. L., 1968, *Rev. econ. Stud.*, **35**, 307; 1972, *Functional Analysis in Mathematical Economics* (Chicago: University of Chicago Press).

Theil, H., 1957, *Econometrica*, **25**, 346.

van Dooren, P., 1981, *SIAM J. Sci. Stat. Computing*, **2**, 121.

Whittle, P., 1981, *Adv. appl. Probab.*, **13**, 764; 1982, *Optimization over Time*, Vol. 1 (Chichester: Wiley Interscience); 1983, *Prediction and Regulation by Linear Least Square Methods*, 2nd edn. (Minneapolis: University of Minnesota Press); 1985, The risk-sensitive certainty equivalence principle. *Essays in Time Series and Allied Processes* (London: Applied Probability Trust).