Pergamon

PII: S0021-8928(96)00071-8

J. Appl. Maths Mechs, Vol. 60, No. 4, pp. 561–566, 1996

0021-8928/96 $24.00+0.00

SOLUTION OF THE PROBLEM OF SYNTHESIZING A STOCHASTIC OPTIMAL CONTROL USING NON-LINEAR PROBABILISTIC CRITERIA†

S. V. SOKOLOV

Rostov-on-Don

(*Received* 24 *February* 1995)

A method that enables control laws for non-linear stochastic objects to be synthesized exactly is considered. The control is optimal in the sense of probabilistic criteria of a general form. The advantages of the method over the traditional methods are demonstrated and an example of its practical application is considered. Copyright © 1996 Elsevier Science Ltd.

Available methods for synthesizing stochastic controls that optimize the mathematical expectation of some given functional [1, 2] preclude the use of more general functionals as optimizing probabilistic criteria, namely, functionals that depend non-linearly on the probability density of the parameters of the state of the object, such as the minimum of the entropy of the state vector, Kullback's criterion, and so on.

In this connection the problem arises of developing an approach to the synthesis of optimal controls for stochastic objects, which would enable an exact control law to be constructed for the most general optimization criterion, depending non-linearly on the probability density of the state vector.

1. FORMULATION OF THE PROBLEM

Suppose that a stochastic object is described by a non-linear vector differential equation of dimension N in symmetrized form

$$\mathbf{X}^{\cdot} = \mathbf{f}(\mathbf{X},t) + f_0(\mathbf{X},t)\mathbf{V}_t + \mathbf{U}(\mathbf{X},t) \tag{1.1}$$

where $\mathbf{f} = (f_1, f_2, \ldots, f_N)$, $f_0 = \| f_{0ik} \|$ are known non-linear vector and matrix functions, $\mathbf{V}_t$ is white Gaussian normalized vector-noise of dimension N_1 and $\mathbf{U} = (U_1, \ldots, U_N)$ is the required control vector.

It is required to determine the control vector $\mathbf{U}(X, t)$, defined in the time interval $T = [t_0, t_k]$, that will minimize a probabilistic functional J defined over a bounded set $\mathbf{X}_* \in [\mathbf{X}_{\min}, \mathbf{X}_{\max}]$ and which depends non-linearly both on the control $\mathbf{U}(\mathbf{X}, t)$ and on the probability density $\rho(\mathbf{X}, t)$ of the process $\mathbf{X}_t$

$$J = \int_T \int_{\mathbf{X}_*} \Phi[\rho(\mathbf{X},t), \mathbf{U}(\mathbf{X},t)]\, d\mathbf{X} dt = \int_T W dt \tag{1.2}$$

where Φ is a known non-linear function, representing, in the general case, possible analytic constraints on the control vector.

Since the density $\rho(\mathbf{X}, t)$ of the process $\mathbf{X}_t$, which occurs in the functional (1.2) and is subject to control, is described by the Fokker–Planck–Kolmogorov equation

$$\begin{gathered}\frac{\partial \rho(\mathbf{X},t)}{\partial t} = L\{\mathbf{a}, b, \rho(\mathbf{X},t)\} \\ \mathbf{a} = (a_1, a_2, \ldots, a_N), \quad b = \| b_{ij} \|, \quad i,j = 1,2,\ldots,N\end{gathered} \tag{1.3}$$

$$L\{\mathbf{a}, b, \rho(\mathbf{X},t)\} = -\sum_i \frac{\partial}{\partial X_i}\{a_i(\mathbf{X},t)\rho(\mathbf{X},t)\} + \frac{1}{2}\sum_i \sum_j \frac{\partial^2}{\partial x_i \partial x_j}\{b_{ij}(\mathbf{X},t)\rho(\mathbf{X},t)\}$$

†*Prikl. Mat. Mekh.* Vol. 60, No. 4, pp. 564–569, 1996.

$$a_i(\mathbf{X},t) = f_i(\mathbf{X},t) + U_i(\mathbf{X},t) + \frac{1}{4}\sum_k\sum_j f_{0_{jk}}(\mathbf{X},t)\frac{\partial}{\partial x_j} f_{0_{ik}}(\mathbf{X},t) =$$

$$= u_i(\mathbf{X},t) + a_{1i}(\mathbf{X},t), \quad b_{ij}(\mathbf{X},t) = \frac{1}{2}\sum_k f_{0_{ik}}(\mathbf{X},t)\, f_{0_{jk}}(\mathbf{X},t)$$

which is a partial differential equation, it follows that the solution of our problem will necessitate using methods of optimal control theory for systems with distributed parameters. Throughout this paper summation over i, j goes from 1 to N, and over k, from 1 to N_1.

2. SYNTHESIS OF THE OPTIMAL CONTROL

We shall assume that the optimal control is a member of the class of bounded continuous functions with values in an open domain $\mathbf{U}_*$. To construct the control we will use dynamic programming, according to which the problem reduces to solving the functional equation [3]

$$\min_{\mathbf{U}\in\mathbf{U}_*}\left\{\frac{dV}{dt} + W\right\} = 0 \tag{2.1}$$

subject to the terminal condition $V(t_k) = 0$, where V is an optimal functional that depends parametrically on the time $t \in T$ and is defined on the set of functions ρ that satisfy Eq. (1.3).

For linear systems, the functional is sought as an integral quadratic form [3]

$$V = \int_{\mathbf{X}} v(\mathbf{X},t)\rho^2(\mathbf{X},t)d\mathbf{X}$$

whence we obtain

$$\frac{dV}{dt} + W = \int_{\mathbf{X}}\left\{\frac{dv}{dt}\rho^2 + 2v\rho L\{\mathbf{a}_1, b, \rho\} + \Phi[\rho,\mathbf{U}] - 2v\rho\sum_i\left(\frac{\partial u_i}{\partial x_i}\rho + u_i\frac{\partial \rho}{\partial x_i}\right)\right\}d\mathbf{X} \tag{2.2}$$

$$\mathbf{a}_1 = (a_{11}, a_{12}, \ldots, a_{1N})$$

Analysis of this expression shows that the determination of the vector $\mathbf{U}(\mathbf{X}, t)$ in a solution of the functional equation (2.1) reduces to a classical problem: to find a vector function that minimizes the definite integral (2.2). Moreover, the vector function $\mathbf{U}(\mathbf{X}, t)$ solving this problem must satisfy the system of Euler equations

$$-\frac{\partial}{\partial x_i}[2v\rho^2] - \frac{\partial\Phi}{\partial u_i} + 2v\rho\frac{\partial\rho}{\partial x_i} = 0, \quad i = 1, 2, \ldots, N$$

or

$$\left(\frac{\partial\Phi}{\partial\mathbf{U}}\right)^T = -\rho\frac{\partial}{\partial\mathbf{X}}[2v\rho]^T \tag{2.3}$$

In the general case, Eq. (2.3) is a system of non-linear equations in the components of the control vector, and it can be solved only in certain special cases, e.g. when the function $\Phi[\rho, \mathbf{U}]$ has the form

$$\Phi = \Phi_0[\rho] + \sum_i \psi_i(u_i)$$

where ψ_y are analytic functions with invertible first derivatives, or

$$\Phi = \Phi_0[\rho] + [\mathbf{U}(\mathbf{X},t) - \mathbf{U}_0(\mathbf{X},t)]^T D_U(\mathbf{X},t)[\mathbf{U}(\mathbf{X},t) - \mathbf{U}_0(\mathbf{X},t)]$$

where $D_U(\mathbf{X}, t)$ is a known symmetric square matrix and $\mathbf{U}_0(\mathbf{X}, t)$ is a known vector.

In the latter case the assumption that the function is a quadratic form in the vector $\mathbf{U}$ is dictated, as a rule, by the need to minimize the deviation of a form involving the unknown control from the given vector $\mathbf{U}_0$, and this in turn depends on the possibility of the technical implementation of the control. Equation (2.3) can be solved for $\mathbf{U}$ as follows:

$$\mathbf{U}_{\text{opt}} = \mathbf{U}_0 - \rho D_U^{-1}\left[\frac{\partial}{\partial \mathbf{X}}(2\upsilon\rho)\right]^T \tag{2.4}$$

Substitution of this optimal control law into (1.3) and (2.2) (assuming that the condition $dV/dt + W = 0$ holds for the optimal control $\mathbf{U}_{\text{opt}}$ [3]) yields a system of partial differential equations

$$\begin{aligned}
&\frac{\partial \rho}{\partial t} = L\{\mathbf{a}_1, b, \rho\} - A(\rho, \upsilon), \quad \rho(\mathbf{X}, t_0) = \rho_0 \\
&A(\rho, \upsilon) = \sum_i \frac{\partial}{\partial x_i}\left\{\left(\mathbf{U}_0 - \rho D_U^{-1}\left[\frac{\partial}{\partial \mathbf{X}}(2\upsilon\rho)\right]^T\right)_{(i)} \rho\right\} \\
&\frac{\partial \upsilon}{\partial t} = -2\upsilon\rho^{-1}L\{\mathbf{a}_1, b, \rho\} - \Phi_0[\rho]\rho^{-2} - \left[\frac{\partial}{\partial \mathbf{X}}(2\upsilon\rho)\right] D_U^{-1}\left[\frac{\partial}{\partial \mathbf{X}}(2\upsilon\rho)\right]^T + 2\upsilon\rho^{-1}A(\rho, \upsilon) \\
&\upsilon(\mathbf{X}, t_k) = 0
\end{aligned} \tag{2.5}$$

The subscript i is introduced to denote the ith vector component.

In principle, the solution of this system exhausts the solution of our problem.

3. ANALYSIS OF METHODS FOR THE COMPUTATIONAL IMPLEMENTATION OF THE OPTIMAL CONTROL

From the standpoint of the accuracy and complexity with which the optimal control is constructed, a comparative analysis of the solution of system (2.5) and the system of adjoint equations obtained using the traditional approach [1, 2] may be of interest. While the optimal control in the case of (2.5) depends on ρ through a proportional-differential relationship, the control produced by the traditional method depends on ρ through integration, which implies, given an N-dimensional state vector, the need for simultaneous integration of the $(2N + 1)$-dimensional system of adjoint ordinary integro-differential equations and a partial integro-differential equation for the N-dimensional function ρ. This is far more complicated than solving a system of partial differential equations like (2.5). Nevertheless, the difficulties involved in solving system (2.5), for which there are at present no general methods for exact analytic solution, are not lessened. There are a great many approximate methods for solving the problem, all guided by some compromise between the necessary accuracy and the available computational resources; we shall not dwell on them here. Rather, we shall focus on one possible solution method, based on expanding the functions υ and ρ in series in some system of orthonormal functions of a vector argument

$$\upsilon(\mathbf{X}, t) = \sum_\mu \alpha_\mu(t)\varphi_\mu(\mathbf{X}) = \boldsymbol{\varphi}^T\boldsymbol{\alpha}$$

$$\rho(\mathbf{X}, t) = \sum_\mu \beta_\mu(t)\varphi_\mu(\mathbf{X}) = \boldsymbol{\varphi}^T\boldsymbol{\beta}$$

where μ takes a set of values from $(0, \ldots, 0)$ to $(N_2, \ldots, N_2)$ [3], φ is a vector of orthonormal functions of $\mathbf{X}$, and $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are the coefficient vectors of the expansions.

In that case, the solution reduces to solving the two-point boundary-value problems of integrating the following system of ordinary differential equations

$$\boldsymbol{\beta}^{\cdot} = \int_{\mathbf{X}} \boldsymbol{\varphi} L[\mathbf{a}_1, b, \boldsymbol{\varphi}^T\boldsymbol{\beta}]d\mathbf{X} - \int_{\mathbf{X}} \boldsymbol{\varphi} A(\boldsymbol{\varphi}^T\boldsymbol{\beta}, \boldsymbol{\varphi}^T\boldsymbol{\alpha})d\mathbf{X}$$

$$\alpha^{\cdot} = \int_X \varphi\Big\{-2\varphi^T\alpha(\varphi^T\beta)^{-1}L[\mathbf{a}_1,b,\varphi^T\beta]-\Phi_0[\varphi^T\beta](\varphi^T\beta)^{-2}-$$

$$-\left[\frac{\partial}{\partial \mathbf{X}}(2\varphi^T\alpha\varphi^T\beta)\right]D_U^{-1}\left[\frac{\partial}{\partial \mathbf{X}}(2\varphi^T\alpha\varphi^T\beta)\right]^T+2\varphi^T\alpha(\varphi^T\beta)^{-1}A(\varphi^T\beta,\varphi^T\alpha)\Big\}d\mathbf{X} \tag{3.1}$$

subject to the boundary conditions $\alpha(t_k) = 0$, $\beta(t_0) = \beta_0$, where the values of the components β_0 are determined by expanding the function $\rho(\mathbf{X}, t_0) = \rho_0$.

From the standpoint of practical implementation, it turns out to be easier to integrate system (3.1) with its boundary conditions than to integrate (2.5), but as far as real-time organization of the estimation process is concerned, the problem remains no less difficult. Moreover, the desirability of such a direct approach is questionable, for the following reasons. First, the necessary time and computational resources turn out to be large; second, there is no possibility of adjusting the vector $\mathbf{U}$ in real time, and, third, in the process of instrumental realization it is generally not possible to store the given values of $\mathbf{U}$ with sufficient accuracy.

Thus, in this case we are quite justified in solving problem (3.1) by means of approximate methods, and one of these is the invariant embedding method [4], which yields the desired approximate solution in real time.

The use of this method presupposes that all the components of the unknown vector are given in differential form. Therefore, to permit real-time synthesis of the vector $\mathbf{U}$ by this method, we introduce a fictitious variable ϑ, which will enable us in what follows to incorporate (2.4) as a differential equation

$$\vartheta^{\cdot} = \mathbf{U}_{opt}(\varphi^T\alpha,\ \varphi^T\beta)$$

considered together with Eqs (3.1) as a single system. The use of the invariant embedding method in this case yields the following system of equations

$$\begin{vmatrix}\hat{\vartheta}^{\cdot}\\ \hat{\beta}^{\cdot}\end{vmatrix} = \begin{vmatrix}\mathbf{U}_0\\ \int_X \varphi B(\mathbf{a}_1,b,\varphi^T\hat{\beta},\mathbf{U}_0)d\mathbf{X}\end{vmatrix} - D\int_X \varphi\Phi_0[\varphi^T\hat{\beta}](\varphi^T\hat{\beta})^{-2}d\mathbf{X}$$

$$D^{\cdot} = 2\int_X \varphi\left\{\frac{\partial}{\partial\hat{\beta}}B(\mathbf{a}_1,b,\varphi^T\hat{\beta},\mathbf{U}_0)\right\}d\mathbf{X}\cdot D + D\int_X \varphi\{2\varphi^T(\varphi^T\hat{\beta})^{-1}B(\mathbf{a}_1,b,\varphi^T\hat{\beta},\mathbf{U}_0)\}d\mathbf{X}+$$

$$+\int_X \varphi\left\{\sum_i \frac{\partial}{\partial x_i}\left(\left[(\varphi^T\hat{\beta})D_U^{-1}\frac{\partial}{\partial \mathbf{X}}(2\varphi^T[\varphi^T\hat{\beta}])\right]_{(i)}\varphi^T\hat{\beta}\right)\right\}d\mathbf{X}-$$

$$-2D\int_X \varphi\left\{\frac{\partial}{\partial\hat{\beta}}[\Phi_0[\varphi^T\hat{\beta}](\varphi^T\hat{\beta})^{-2}]\right\}d\mathbf{X}\cdot D$$

$$B(\mathbf{a}_1,b,\varphi^T\hat{\beta},\mathbf{U}_0) = L[\mathbf{a}_1,b,\varphi^T\hat{\beta}] - \sum_i \frac{\partial}{\partial x_i}[u_{0(i)}\varphi^T\hat{\beta}]$$

Since the matrix D in the invariant embedding method plays the role of a weight matrix relative to the deviation from optimum of the approximate solution vector, the components corresponding to the variables β_i in D characterize the degree of deviation from the expansion coefficients of the true density (the components of D, accordingly, are the deviations of the parameters at the starting time). A considerable merit of this approach, irrespective of the formation of the approximate solution, is the possibility of a real-time synthesis of an optimal control $\mathbf{U}_{opt}$.

4. EXAMPLE

The effective use of the proposed approach may be illustrated by the following example.

A controlled system is described by an equation

$$x^{\cdot} = -ax^3 + u + V,\quad x(t_0) = 0$$

where $V(t)$ is centred white Gaussian noise of intensity D_V.

The control u will be synthesized subject to the condition that the deviation of the coordinate x from its initial state $x(t_0)$ over an interval $[t_0, t_k]$ must be a minimum and that the control must be produced at minimum cost. The traditional approach enables us to solve the problem on the basis of the root-mean-square criterion [2]

$$J_1 = M\left\{\int_{t_0}^{t_k} [x^2(\tau) + K^2u^2(\tau)]d\tau\right\}, \quad K = \text{const}$$

while the criterion underlying the approach proposed here is the maximum probability of the existence of x in a given neighbourhood $x_* = [X_{\min}, X_{\max}]$ of x_0 (by Chebyshev's inequality, this potentially yields high accuracy in the control of x), i.e. the minimization of the criterion

$$J_2 = \int_{x_*}\int_{t_0}^{t_k} [-\rho(x,\tau) + K^2u^2(x,\tau)]\,d\tau dx$$

In the first case, the optimal control is defined as [2]

$$u_{\text{opt}} = \frac{1}{2K^2}M[\lambda] = \frac{1}{2K^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\lambda\rho(x,\lambda,\tau)d\lambda dx$$

where λ is the adjoint variable and the system of canonical adjoint equations, derived from the stochastic Hamiltonian, is

$$\dot{x} = -ax^3 + u_{\text{opt}} + V_t, \quad x(t_0) = x_0$$

$$\dot{\lambda} = 3a\lambda x^2 + 2x, \quad \lambda(t_k) = 0$$

so that one can write down the following equation for the density $\rho = \rho(x, \lambda, t)$

$$\frac{\partial\rho}{\partial t} = -\left\{\frac{\partial}{\partial x}\left[\left(\frac{1}{2K^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\lambda\rho dx d\lambda - ax^3\right)\rho\right] + \frac{\partial}{\partial\lambda}[(3a\lambda x^2 + 2x)\rho]\right\} + \frac{D_V}{2}\frac{\partial^2\rho}{\partial x^2} \tag{4.1}$$

Consequently, exact solutions of the problem reduces in this case to solving a two-dimensional integro-differential equation (4.1), but with indeterminate boundary conditions (as the variables x and λ are undetermined at times t_0 and t_k). Because of this last detail it is in principle impossible, using the traditional approach, to synthesize an exact statistically optimal control (unlike the approach in this paper). The classical way to eliminate this contradiction is Gaussian approximation of the density $\rho(x, \lambda, t)$, which in this case leads to the adjoint system for estimates of the means

$$\dot{\hat{x}} = -a\hat{x}^3 + \frac{1}{2K^2}\hat{\lambda}, \quad \hat{x}(t_0) = M(x_0) = 0 \tag{4.2}$$

$$\dot{\hat{\lambda}} = 3a\hat{\lambda}\hat{x}^2 + 2\hat{x}, \quad \hat{\lambda}(t_k) = 0$$

Integration of system (4.2) is equivalent to solving a two-point boundary-value problem, so that in order to construct the control in real time one must use invariant embedding, which yields equations for the approximate estimate $\hat{x}_*$

$$\dot{\hat{x}}_* = -a\hat{x}_*^3 + 2D\hat{x}_*, \quad \dot{D} = 4D^2 - 9aD\hat{x}_*^2 - \frac{1}{2K^2}$$

from which one can determine the approximate control law

$$\hat{u}_{\text{opt}} = 2\hat{x}_*D$$

The alternative approach proposed here, in turn, enables one to construct an approximate control law as the right-hand side of the equation in the variable ϑ, which makes up a single system together with the equations in the approximate coefficients $\hat{\beta}$ of the expansion of the density $\rho(x, t)$ in the orthonormal system of functions φ

$$\begin{vmatrix} \dot{\hat{\vartheta}} \\ \dot{\hat{\beta}} \end{vmatrix} = \begin{vmatrix} 0 \\ \int_{x_*}\varphi\{B_1(\varphi,x)\hat{\beta} + 3x^2a\varphi^T\hat{\beta}\}dx \end{vmatrix} + D\int_{x_*}\varphi(\varphi^T\hat{\beta})^{-1}dx$$

$$D = 2\int_{x_*}\boldsymbol{\varphi}\{B_1(\boldsymbol{\varphi},x) + 3x^2 a\boldsymbol{\varphi}^T\}dxD + 2D\int_{x_*}\boldsymbol{\varphi}\{\boldsymbol{\varphi}^T(B_1(\boldsymbol{\varphi},x)\hat{\boldsymbol{\beta}}(\boldsymbol{\varphi}^T\hat{\boldsymbol{\beta}})^{-1} + 3ax^2)\}dx +$$

$$+\frac{2}{K^2}\int_{x_*}\boldsymbol{\varphi}\left[\frac{\partial}{\partial x}\left\{(\boldsymbol{\varphi}^T\hat{\boldsymbol{\beta}})^2\frac{\partial}{\partial x}(\boldsymbol{\varphi}^T[\boldsymbol{\varphi}^T\hat{\boldsymbol{\beta}}])\right\}\right]dx - 2D\int_{x_*}\boldsymbol{\varphi}\boldsymbol{\varphi}^T(\boldsymbol{\varphi}^T\hat{\boldsymbol{\beta}})^{-2}dxD$$

$$B_1(\boldsymbol{\varphi},x) = \frac{D_V}{2}\frac{\partial^2\boldsymbol{\varphi}^T}{\partial x^2} + ax^3\frac{\partial\boldsymbol{\varphi}^T}{\partial x}$$

For a comparative evaluation of the accuracy of both approaches, a numerical simulation was carried out of a suboptimal control (obtained by the invariant embedding method) for the above object with

$$a = K = 1;\ D_V = 1.7;\ x_0 = 0,\ x_* = [-2.3, 2.3];\ T = [0, 100]\ \text{s}$$

$$\boldsymbol{\varphi} = (\cos(\omega_0 x), \sin(\omega_0 x), \cos(2\omega_0 x), \sin(2\omega_0 x)),\ \omega_0 = \pi/2.3$$

the control being implemented for 30 stochastic trajectories of the object.

The equations were integrated by a third-order Runge–Kutta procedure with step size 0.03 s. The accuracy of the control was estimated by averaging over the ensemble of realizations of the mean-modulus deviations of the trajectories, taken separately, from the boundaries of the interval x_* in time T. By the end of the simulation it was established that the accuracy of the control based on the approach proposed here exceeds that of the traditional approach by a factor of more than two, which in turn implies that the method can be used to good effect to synthesize controls for practical stochastic systems.

REFERENCES

1. AFANAS'YEV V. N., KOLMANOVSKII V. B. and NOSOV V. R., *Mathematical Theory of Control System Design*. Vysshaya Shkola, Moscow, 1989.
2. KAZAKOV I. Ye., *Statistical Theory of Control Systems in State Space*. Nauka, Moscow, 1975.
3. SIRAZETDINOV T. K., *Optimization of Systems with Distributed Parameters*. Nauka, Moscow, 1977.
4. PERVACHEV S. V. and PEROV A. I., *Adaptive Filtering of Messages*. Radio i Svyaz', Moscow, 1991.

Translated by D.L.