MASSACHUSETTS INSTITUTE OF TECHNOLOGY

RE-49

A NEW ALGORITHM FOR SUBOPTIMAL STOCHASTIC CONTROL

by

Renwick W. Curry

September 1968

EXPERIMENTAL ASTRONOMY LABORATORY
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
CAMBRIDGE 39, MASSACHUSETTS

RE-49

A NEW ALGORITHM FOR SUBOPTIMAL STOCHASTIC CONTROL

by

Renwick W. Curry

September 1968

EXPERIMENTAL ASTRONOMY LABORATORY
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
CAMBRIDGE, MASSACHUSETTS, 02139

Approved: W. Markey
Director
Experimental Astronomy Laboratory

This report was presented at the Sixth Allerton Conference on Circuit and System Theory, University of Illinois, Urbana, Ill., Oct. 2-4, 1968.

A NEW ALGORITHM FOR SUBOPTIMAL STOCHASTIC CONTROL*

RENWICK E. CURRY**
Experimental Astronomy Laboratory
Massachusetts Institute of Technology
Cambridge, Massachusetts

ABSTRACT

An apparently new stochastic control algorithm, herein called M-measurement-optimal feedback control, is described for discrete time systems. This scheme incorporates M future measurements into the control computations: when M is zero, it reduces to the well-known open-loop-optimal feedback control; when M is the actual number of measurements remaining in the problem, it becomes the truly optimal stochastic control. This new algorithm may also be used to simplify computations when the plant is nonlinear, when the controls are constrained, or when the cost is nonquadratic. Simulation results are presented which show the superiority of the new algorithm over the open-loop-optimal feedback control.

I. INTRODUCTION

In most cases of practical interest it is exceedingly difficult to obtain the solution to the optimal stochastic control problem (or combined estimation and control problem) if the plant and/or measurement equations are nonlinear, if the control is constrained, or if the cost is non-quadratic. Different suboptimal control algorithms have evolved as a compromise between computational effort and a desire to incorporate realistic modeling of the uncertainties involved. One widely used algorithm is the so-called open-loop-optimal feedback control [1], in which the control action is computed under the assumption that no measurements are to be taken in the future.

The algorithm presented here, called M-measurement-optimal feedback control, assumes that M measurements are to be taken in the future. This control is more akin to Fel'dbaum's "dual control" theory [2] than the open-loop-optimal feedback control, since the control not only "directs" the state, but can also "probe" or learn about the state. The same concept can be used to treat other difficult-to-handle constraints such as nonlinear state/measurement equations, constrained controls, and nonquadratic cost criteria.

First the M-measurement-optimal feedback control is described and discussed, and the special case of one future measurement is examined in more detail. The paper concludes with an example and Monte Carlo simulation results of the stochastic control of a linear plant with coarsely quantized (nonlinear) measurements.

II. M-MEASUREMENT-OPTIMAL FEEDBACK CONTROL

Description of the Control Algorithm

The algorithm proceeds as follows: measurements and control actions are recorded up to (say) time t_k, and an estimate of the state vector (and,

* Supported by NASA Grant NsG 254-62. This work is based on a portion of the author's Ph.D. Dissertation submitted to the Department of Aeronautics and Astronautics, Massachusetts Institute of Technology, Cambridge, Mass.

** Now with the School of Electrical Engineering, Cornell University, Ithaca, New York.

perhaps, its conditional probability density function) is made. It is temporarily assumed that M measurements are to be taken in the future (at preassigned times), and the control sequence $\{u_k, u_{k+1}, \dots\}$ is computed to minimize the cost to complete the process taking M measurements along the way. This control sequence is the M-measurement-optimal control sequence. Only u_k, the first member of the sequence, is actually implemented; this advances the time index, a new measurement is taken, a new estimate is computed, and the process of computing a new M-measurement-optimal control sequence is repeated.

Remarks

Limiting Forms of the Control The M-measurement-optimal feedback control has two limiting forms: when M is zero, no future measurements are included in the control computations and it becomes the well-known open-loop-optimal feedback control [1] ; when M is the actual number of measurements remaining in the problem, it becomes the truly optimal stochastic control [2] .

Qualitative Description All available past and present measurements are used in the control computations, hence the name feedback control.

The primary advantage of the M-measurement-optimal feedback control over other suboptimal control algorithms is that one or more future observations are incorporated into the control computations. This allows for the "probing" or "test signal" type of control action which is characteristic of the truly optimal stochastic control [2] . The open-loop-optimal feedback control cannot take this form since future measurements are ignored.

Extensions The concept of incorporating M future measurements rather than the actual number of measurements may be used to simplify computations involving constrained controls, nonlinear plants, and nonquadratic cost criteria as well. For example, if the actual controls are constrained at all future times, it can be assumed for the purposes of control computation that they are constrained only at M future times and unconstrained at all other times. Similarly, it can be assumed that the plant is nonlinear only at M future times but linear at all other times, or that the cost is nonquadratic at M future times, but quadratic at all other times.

In all of these situations the suboptimal control approaches the optimal control as M approaches the number of stages remaining in the problem. Generally speaking, the computational effort will increase as M increases, but the author knows of no analysis to show that the performance will improve as M increases.

III. ONE-MEASUREMENT-OPTIMAL FEEDBACK CONTROL

Problem Statement

Given:

a) The state equation

$$x_{i+1} = f_i(x_i, u_i, w_i) \quad i = 0,1,\dots,N \tag{1}$$

where x_i is the state vector at time t_i, u_i is the control vector at time t_i, and w_i is the process noise at time t_i.

b) The measurement equation

$$z_i = h_i(x_i, v_i) \qquad i = 1, 2, \ldots, N \tag{2}$$

where v_i is the observation noise at time t_i.

c) The cost criterion

$$J = \mathcal{E}\left\{\sum_{i=0}^{N} L_i(x_i, u_i) + \phi(x_{N+1})\right\} \tag{3}$$

d) The probability density functions $p(x_0)$, $p(w_i)$, $p(v_i)$. It is assumed that the vectors x_0, $\{w_i\}$, and $\{v_i\}$ are independent.

e) For notational convenience let

$$Z_k = \{z_1, z_2, \ldots z_k\}$$

$$U_k = \{u_0, u_1, \ldots, u_k\}$$

Open-Loop-Optimal Feedback Control

The cost function to be minimized by the open-loop control at time t_k is given by

$$J_{OL,k} = \mathcal{E}\left[\sum_{j=k}^{N} L_j(x_j, u_j) + \phi(x_{N+1}) \,\Big|\, Z_k, U_{k-1}\right] \tag{4}$$

This is minimized by the control sequence $\{u_k, \ldots, u_N\}$ subject to the state constraint (1). The first member of this sequence, u_k, is the open-loop-optimal feedback control at time t_k.

One-Measurement-Optimal Feedback Control

This portion treats a special case of M-measurement-optimal feedback control. If only one future measurement is incorporated in the control computation at time t_k, and if this measurement occurs at time t_{k+n}, then the one-measurement cost function becomes

$$J_{OM,k} = \mathcal{E}\left\{\sum_{j=k}^{k+n-1} L_j(x_j, u_j) + \mathcal{E}\left[J^o_{OL,k+n} \big| Z_k, U_{k+n-1}\right] \,\Big|\, Z_k, U_{k-1}\right\} \tag{5}$$

In this equation, the optimum open-loop cost function $J^o_{OL,k+n}$ is similar to (4) after minimization, but it depends on the measurements $\{Z_k, z_{k+n}\}$ and the controls U_{k+n-1}. The first member of the control sequence $\{u_k, \ldots, u_N\}$ that minimizes (5) is the one-measurement-optimal feedback control at time t_k.

The solution to (5) may still be difficult to find in practice. Even if the explicit dependence of the open-loop cost function on the open-loop-optimal controls $\{u_{k+n}, \ldots u_N\}$ can be removed, the solution of (5) still involves an n-stage two point boundary problem. If, however, the future measurement occurs at the "next" time, then (5) reduces to a parameter optimization over u_k, i.e., the one-measurement cost function becomes

$$J_{OM,k} = \mathcal{E}\left\{ L_k(x_k,u_k) + \mathcal{E}\left[J^o_{OL,k+1} \mid Z_k, U_k \right] \Big| Z_k, U_{k-1} \right\} \qquad (6)$$

In this equation $J^o_{OL,k+1}$ depends on the measurements Z_{k+1} and the controls U_k.

Linear Systems, Quadratic Cost, and Nonlinear Measurements

This section contains some specialized results when the state equation is linear

$$x_{i+1} = \Phi_i x_i + G_i u_i + w_i \qquad i = 0,\ldots,N \qquad (7)$$

and when the cost is quadratic

$$J = \mathcal{E}\left\{ \sum_{i=0}^{N} x_i^T A_i x_i + u_i^T B_i u_i + x_{N+1}^T A_{N+1} x_{N+1} \right\} \qquad (8)$$

Here the superscript T denotes the transpose, and the weighting matrices $\{A_i\}$, $\{B_i\}$ are positive definite. The measurements may still be nonlinear.

It has been shown [3, Appendix D] that the optimum open-loop cost function in this case is given by

$$J^o_{OL,k+1} = \mathcal{E}(x_{k+1}^T S_{k+1} x_{k+1} \mid Z_{k+1}, U_k)$$

$$+ \operatorname{tr}\left[E_{k+1}(F_{k+1} - S_{k+1}) \right] + \text{const.} \qquad (9)$$

where tr is the trace of a matrix, and

$$S_i = A_i + \Phi_i^T \left[S_{i+1} - S_{i+1} G_i (B_i + G_i^T S_{i+1} G_i)^{-1} G_i^T S_{i+1} \right] \Phi_i \qquad (10)$$

$$S_{N+1} = A_{N+1}$$

$$F_i = \Phi_i^T F_{i+1} \Phi_i + A_i$$

$$F_{N+1} = A_{N+1} \qquad (11)$$

E_{k+1} = covariance of x_{k+1} conditioned on Z_{k+1} and U_k.

A by-product in the derivation of (9) is the open-loop-optimal feedback control. At time t_k it is given by [3, p.80]

$$u^o_{OL,k} = -(B_k + G_k^T S_{k+1} G_k)^{-1} G_k^T S_{k+1} \Phi_k \mathcal{E}(x_k \mid Z_k, U_{k-1}) \qquad (12)$$

This is the precise form of the truly optimal control when the measurements are linear, the plant is linear, and the cost is quadratic.

The equation for the one-measurement cost function with the future measurement occuring at the next time, Eq.(6), becomes

$$J_{OM,k} = \mathcal{E}\Big[x_k^T A_k x_k + u_k^T B_k u_k + x_{k+1}^T S_{k+1} x_{k+1} + \mathcal{E}\Big\{ tr\big[E_{k+1}(F_{k+1} - S_{k+1})\big] \Big| Z_k, U_k \Big\} \Big| Z_k, U_{k-1} \Big] + \text{const.} \quad (13)$$

The value of u_k that minimizes (13) is the one-measurement-optimal feedback control at time t_k for linear systems, quadratic cost, and nonlinear measurements. This is computationally equivalent to solving for the first control action in a two stage process [3, p.67]. Nonlinear control action results from the fact that the conditional covariance E_{k+1} in (13) is a function of u_k. (See [3] for some numerical results of the two stage control process.) It has also been shown in [3, Appendix D] that the weighting matrix $F_{k+1} - S_{k+1}$ is at least positive semidefinite. If it were not, the control would minimize the cost by degrading, not improving, the knowledge of the state.

IV. SIMULATIONS

System Configuration

For definiteness, let the scalar state satisfy the equation

$$x_{i+1} = x_i + u_i + w_i \qquad i = 0,\dots,N \quad (14)$$

where x is the state, u is the scalar control, and w is Gaussian process noise. The initial conditions also have a Gaussian distribution. The cost function is quadratic in nature.

$$J = \mathcal{E}\Big\{ \sum_{i=0}^{N} A_i x_i^2 + B_i u_i^2 + A_{N+1} x_{N+1}^2 \Big\} \quad (15)$$

Observations are taken through a three level quantizer shown in Fig. 1. Observation noise is not included in this example.

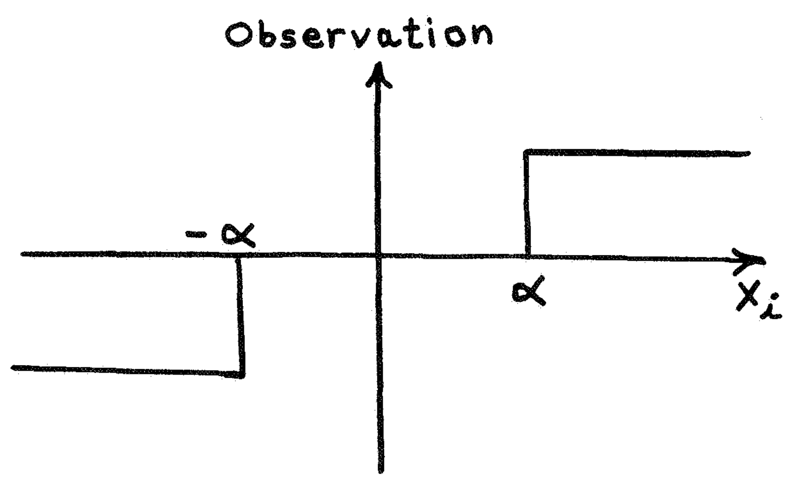

Fig. 1. Three Level Quantizer for Observations

Estimation Procedure

Obtaining numerical solutions to the stochastic control problem requires extensive use of the probability density function of the state conditioned on past measurements and control actions. Here it is assumed that the distribution just prior to a quantized measurement is Gaussian. This algorithm has been extensively treated in [3], and appears to be a good approximation for linear systems and quantized measurements. The details of calculating moments conditioned on quantized measurements can also be found in [3].

Simulation Description

Both the open-loop-optimal feedback control and the one-measurement-optimal feedback control have been simulated on the digital computer. The latter algorithm assumes that the one future measurement occurs at the next time instant. The control algorithms used the same realizations of the initial conditions and process noise, although the state trajectories, in general, were different. Each case was run 50 times and the ensemble averages were approximated by the numerical average of these 50 trials. In all cases the weighting coefficients for the control effort $\{B_i\}$ are unity, and the process noise variance is constant, but varies from case to case.

Numerical Results

The results of a seven stage stochastic process are described. Quantitative results are displayed for a terminal control problem, and the results of other simulations are discussed.

Terminal Control In the terminal control problem, all state weighting coefficients are unity except the last one, A_7, which is 100. This means, roughly, that the RMS terminal error is ten times "more important" than the other quantities in the cost function.

Fig. 2 shows the results when the measurements are taken through a three level quantizer: the ensemble mean square state and ensemble average of the (approximate) conditional covariance of the state are plotted as a function of time. For this case, the variance of the process noise is 0.2, and the quantizer switch points are at $\pm$ 1.

The most noticeable difference between the two control laws is that the one-measurement control acts to reduce the conditional covariance of the state estimate. Note that the ensemble average of the conditional covariance is about half the average of the conditional covariance for the open-loop control. The one-measurement control is able to do this by centering the conditional distribution of the measurement near the quantizer switch point. This is reflected in the curves for the mean square value of the state, which stays in the neighborhood of 1.0 (the switch point) for the one-measurement control, but gradually goes to zero for the open-loop control. The control effort (not shown) for the one-measurement control is higher, and it requires a large control action at the last application to bring the state from the vicinity of the quantizer switch point to the origin.

The performance penalty of the open-loop-optimal feedback control over the one-measurement-optimal feedback control is 17.1 percent for this case. Other simulations revealed that the performance penalty ranged as high as 44 percent when observations were taken through a two level quantizer.

Other Simulations Cost functions other than the terminal control type were simulated: the state deviations were weighted more heavily as time progressed, or else the weightings were constant. The performance advantage of the one-measurement control was always less than 10 percent in these cases. This arises from the fact that the one-measurement control tries to move the state around to gain information, but these movements are restricted by the heavy weighting on the state deviations.

Thus a qualitative assessment, at least for linear systems and nonlinear measurements, is that incorporating future measurements in the control computations will yield the greatest return when the cost function is such that the state and/or control is free to reduce uncertainty in the estimate. In other situations, the open-loop control is quite attractive, especially because of its computational simplicity.

V. CONCLUSIONS

A new suboptimal stochastic control algorithm is presented which incorporates one or more future measurements into the control computations. The two limiting forms of the algorithm are the well-known open-loop-optimal feedback control and the truly optimal stochastic control. This concept can be extended to simplify the computations for constrained controls, nonlinear plants, and nonquadratic cost criteria. A linear system with quantized measurements was simulated to compare the one-measurement-optimal feedback control to the open-loop-optimal feedback control. The greatest improvement over the open-loop algorithm occurs in those situations where the cost function gives the control and state some freedom to reduce the uncertainty in the state estimate.

ACKNOWLEDGEMENT

The author wishes to acknowledge the helpful discussions of C. F. Price of the Experimental Astronomy Laboratory, MIT.

REFERENCES

[1] Dreyfus, S., "Some Types of Optimal Control of Stochastic Systems", J. SIAM Control, Ser. A, Vol.2, pp.120-134.

[2] Fel'dbaum, A., "Dual Control Theory, I-IV", Automation and Remote Control, 21, 1960, p.874, p.1033; 22, 1961, p.1, p.109.

[3] Curry, R. E., "Estimation and Control with Quantized Measurements", Ph.D. Dissertation, Dept. Aeronautics and Astronautics, MIT, Cambridge, Mass., May, 1968. Also Report TE-23, Experimental Astronomy Laboratory, MIT, Cambridge, Mass.

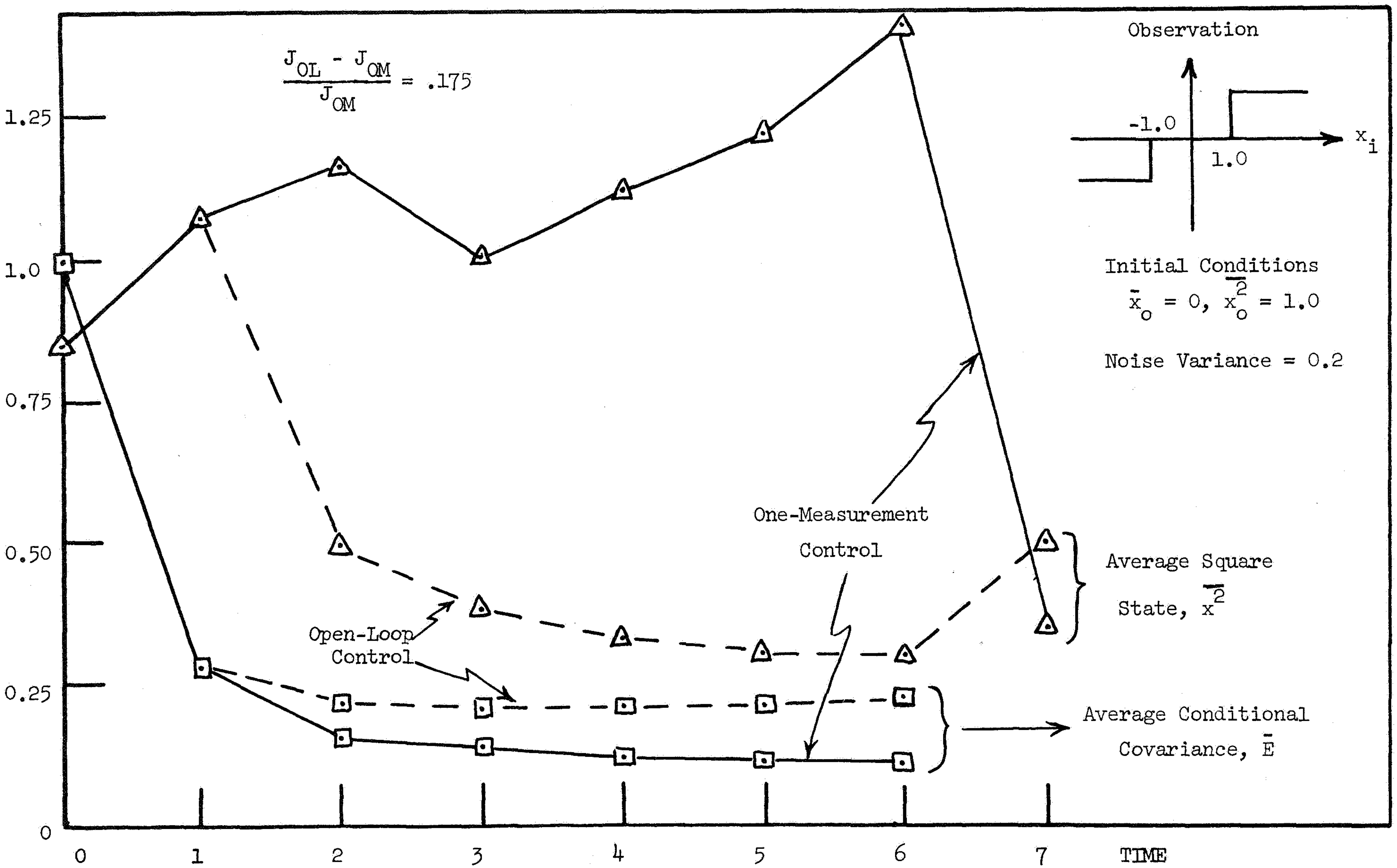

Fig. 2. Seven Stage Terminal Stochastic Control with Observations Quantized to Three Levels