Numerical Methods for Optimal Controls for Nonlinear Stochastic Systems With Delays, and Applications to Internet Regulation

Harold J. Kushner *

* *Applied Math, Brown University, Providence RI 02912 USA, hjk@dam.brown.edu*

Abstract: The Markov chain approximation method, a primary approach for computing optimal values and controls for stochastic systems, was extended to nonlinear diffusions with delays in a recent book. The convergence of many forms of algorithms was proved. The path, control and/or reflection terms can all be delayed. Reflection terms occur in communications models, where they correspond to buffer overflows. If the control and/or reflection terms are delayed, the memory requirements can make the problem intractable. Recasting the problem in terms of a "wave equation" yields practical algorithms with much reduced computational needs. We outline the approach, concentrating on forms motivated by applications to communications, and give data illustrating the potential.

Keywords: Nonlinear stochastic delay systems, optimal control, numerical methods, communications applications, Markov chain approximations.

1. INTRODUCTION

The Markov chain approximation methods Kushner and Dupuis (2001) are commonly used to compute optimal values and controls for stochastic systems. The approach is simple: Approximate the process by a controlled finite-state Markov chain that satisfies certain minimal properties, then solve the Bellman equation. The proofs of convergence are probabilistic and do not depend on analytical properties of the Bellman equation. This makes the methods robust and converge under weak conditions, and is essential when there are delays, where little is known about the properties of the Bellman equation. In the proof, one interpolates the optimally controlled chain to a continuous-time process and then shows that this converges to an optimal diffusion. Many ways of getting the approximations are in Kushner and Dupuis (2001).

The methods were extended to controlled general nonlinear delayed diffusion models in Kushner (2008). If the control and/or reflection terms are delayed, then the memory requirements can make even simple problems intractable. Then we took a promising approach where the delay equation was represented in terms of a stochastic "wave equation" whose numerical solution yields the optimal costs and controls. Algorithms were developed with much reduced memory needs. Convergence theorems were proved, but little attention was given to applications. This was partially remedied in Kushner (2009) which adapted them to concrete models, and gave data illustrating the potential power. The problems did not have delayed reflection terms, which occur in communications models, where they correspond to buffer overflows, and where this data is sent to the admissions controller via a transportation delay. We continue the development, emphasizing models arising in communications, involving both delayed controls and reflection terms. For clarity, some details overlap those of Kushner (2009), which did not cover delayed reflection terms, a main concern here. Numerical data illustrates the type of information that could be obtained. Such problems would be intractable by other current method: e.g., using time discretizations over the delay interval. The model and assumptions are in Section 2. Section 4 represents the solution in terms of a stochastic wave equation. The Markov chain approximation method is reviewed for the no-delay case in Section 5 and extended to the delay case in Section 6. Numerical data is in Section 7.

2. THE MODEL, NOTATION, AND ASSUMPTIONS

Let $I\!R^r$ = Euclidean r-space. The $I\!R^r$-valued path process $x(\cdot) = \{x_i(\cdot), i \leq r\}$ is confined to a convex polyhedron $G \in I\!R^r$. by the boundary reflection process $z(\cdot)$. Let $y_i(\cdot)$ =component of $z(\cdot)$ due to reflection from the ith face of G, with reflection direction d_i. Then $z(t) = \sum_i d_i y_i(t)$. If a component y_i models a buffer over/underflow, then d_i is an inward normal. See the figure in Section 3. The model is the diffusion equation, where $\bar\theta$ =maximum delay and $w(\cdot)$ is a Wiener process:

$$dx(t) = dt \int_{-\bar\theta}^{0} b(x(t+\theta), u(t+\theta), \theta) d\mu_a(\theta) + \sigma(x(t))dw(t) + c(x(t), u(t))dt + dz(t) + dt \int_{\theta=-\bar\theta}^{0} p(\theta) d_\theta y(t+\theta). \tag{1}$$

* Supported by NSF DMS-0804822, ARO contract FA9550-09-1-0378, AFOSR contract W911NF-09-1-0155

The last term models the delayed reflection or buffer under/overflow. We use $\bar{\theta}$ for all components. If the max delay depends on component, there is only a notational change. The initial conditions are the delayed components over the delay interval: e.g., $\hat{x} = \{x(s), -\bar{\theta} \leq s \leq 0\}$, etc. If there is no reflection term, but the process is stopped on reaching a target boundary, add a termination cost and drop all $z(\cdot), y(\cdot)$ terms.

Assumptions. The non-restrictive standard conditions on $\{d_i\}$ are those in (Kushner and Dupuis, 2001, Section 5.7), to which the reader is referred.[1] The controls take values in a compact set U. The functions $b(\cdot), c(\cdot), p(\cdot), \sigma(\cdot), k(\cdot)$ are bounded and continuous, and there is a unique weak-sense solution to (1) for each initial condition and control. All functions of θ are zero for $\theta < -\bar{\theta}$ and $\theta > 0$. $\mu_a(\cdot)$ is a finite measure on $[-\bar{\theta}, 0]$ with $\mu_a([-t, 0]) \to 0$ as $t \to 0$. To assure that $x(\cdot)$ is well defined, there is $\bar{\theta}_0 \in (0, \bar{\theta}]$ such that $p(\theta) = 0$ for $\theta \geq -\bar{\theta}_0$.

The cost function. With $\hat{x}, \hat{u}, \hat{y}$ denoting the canonical values of the (path, control, reflection) segments on $[-\bar{\theta}, 0]$, $\beta > 0$, vector q, and control $u(\cdot)$, we concentrate on the cost function $W(\hat{x}, \hat{u}, \hat{y}, u)$ defined by

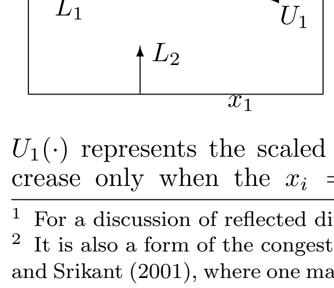
$$E^u_{\hat{x},\hat{u},\hat{y}} \int_0^\infty e^{-\beta t} \left[k(x(t), u(t))dt + q' dy(t) \right]. \tag{2}$$

$E^u_{\hat{x},\hat{u},\hat{y}}$ =expectation. Set $V(\hat{x}, \hat{u}, \hat{y}) = \inf_u W(\hat{x}, \hat{u}, \hat{y}, u)$.

3. AN EXAMPLE FROM COMMUNICATIONS

The example is an AIMD-type (additive increase, multiplicative decrease) internet rate control model from Altman and Kushner (2005). It is illustrative of problems where the memory requirements could be intractable. Although derived as a heavy traffic limit, its features are typical of other models. There is a set of sources whose rates are controlled, and many uncontrolled sources with short transmissions. After a delay, the packets are received by a "bottleneck router." Buffer overflow packets from the controlled sources are not acknowledged, causing a decrease in their transmission rate. Let $x_2(\cdot)$ =scaled controlled source rate, $x_1(\cdot)$ =scaled content of the bounded buffer, $\bar{\theta}$ the round-trip transportation delay, and q_0 the scaled router processing rate. The model is[2]

$$\begin{aligned} dx_1(t) &= [x_2(t) - q_0]\, dt + \sigma dw(t) + dL_1(t) - dU_1(t), \\ dx_2(t) &= q_1 dt + q_2 u(t - \bar{\theta}) dt \\ &\quad - q_3 dU_1(t - \bar{\theta}) x_2(t - \bar{\theta}) + dL_2(t). \end{aligned} \tag{3}$$



The set G and reflection directions.

$U_1(\cdot)$ represents the scaled buffer overflow. $L_i(\cdot)$ can increase only when the $x_i = 0$, assuring non-negativity. $w(\cdot)$ is a Wiener process representing the scaled centered uncontrolled disturbances.[3] The control $u(t)$ is used to signal the source to reduce or increase the rate. The source rate is controlled by the constant increase rate $q_1 > 0$, and the delayed overflow $dU_1(t - \bar{\theta})$ and control $u(t - \bar{\theta})$. The variables are measured and control determined at the router. Its values are sent to the sources and the result is seen at the router after the roundtrip delay $\bar{\theta}$. The centering that yields $w(\cdot)$ is accounted for by the choice of the net processor rate q_0. Cost functions would penalize buffer overflow and queue length, and reward rate.

Systems like (3), if uncontrolled, have large buffer overflows. Present controls depend only on the current state at the router. Instability arises since, at each t, they do not account for the signals sent on $[t - \bar{\theta}, t)$, whose effect has not yet been seen at the router, leading to overcontrol. The data in Section 7 shows that the use of controls that take the past actions into account can improve the performance considerably.

4. REPRESENTING $X(\cdot)$ BY A WAVE EQUATION

The use of a wave eqn. to model delay systems is not new but little use has been made of it for nonlinear systems. The algorithm is based on the following equation, justified by Theorem 4.1. For $-\bar{\theta} < \theta \leq 0$, define $\chi^0(\cdot)$ and $\chi^1(\cdot)$ by (d_t, d_θ are differentials with respect to t, θ)

$$\begin{aligned} d\chi^0(t) &= \chi^1(t, 0)dt + c(\chi^0(t), u(t))dt + \sigma(\chi^0(t))dw(t) + dz^0(t), \\ d_t \chi^1(t, \theta) &= -d_\theta \chi^1(t, \theta) + p(\theta) dy^0(t) \\ &\quad + b(\chi^0(t), u(t), \theta) \left[\mu_a(\theta + dt) - \mu_a(\theta) \right]. \end{aligned} \tag{4}$$

The boundary condition is $\chi^1(t, -\bar{\theta}) = 0$. The reflection term $z^0(\cdot)$ is for $\chi^0(\cdot)$, taking values in G. (4) is formal, but will be the basis of a convergent algorithm. The solution will be well-defined by (7). Initial conditions are $\chi^0(0) = x(0)$ and $\chi^1(0, \theta)$ equal to

$$\int_{-\bar{\theta}}^{\theta} b(x(\gamma - \theta), u(\gamma - \theta), \gamma) d\mu_a(\gamma) + \int_{-\bar{\theta}}^{\theta} p(\gamma) d_\gamma y^0(\gamma - \theta). \tag{5}$$

The cost function is (2) with $\chi^0(t)$ replacing $x(\cdot)$.

A semigroup representation. A key role is played by the semigroup of the wave equation $d_t \chi^1(t, \theta) = -d_\theta \chi^1(t, \theta)$. Define $\Phi(\cdot)$, acting on functions of θ:

$$(\Phi(t) f(\cdot))(\theta) = \begin{cases} f(\theta - t), & -\bar{\theta} \leq \theta - t \leq 0, \\ 0, & \text{otherwise.} \end{cases} \tag{6}$$

Then the solution to (4) is well-defined by

$$\begin{aligned} \chi^1(t, \theta) &= \Phi(t)\chi^1(0, \theta) + \int_0^t \Phi(t - s) p(\theta) dy^0(s) + \\ &\int_0^t \Phi(t - s) b(\chi^0(s), u(s), \theta) \left[\mu_a(\theta + ds) - \mu_a(\theta) \right]. \end{aligned} \tag{7}$$

The next theorem justifies the use of (4).

Theorem 4.1. $\chi^0(\cdot) = x(\cdot)$. (The proof is in Kushner (2008).)

[1] For a discussion of reflected diffusions see Kushner (2001).

[2] It is also a form of the congestion control method RED Kunniyur and Srikant (2001), where one marks packets for notifications to their sources with a probability depending on the route state.

[3] The Wiener process in Altman and Kushner (2005) is due to the central limit theorems used to get the approximation. It also represents the unpredictability of the system.