Systems & Control Letters 12 (1989) 351–356
North-Holland

Reverse-time modeling, optimal control and large deviations *

Michael R. FRATER,
Rodney A. KENNEDY,
and Brian D.O. ANDERSON

Department of Systems Engineering, Australian National University, G.P.O. Box 4, Canberra, A.C.T. 2601, Australia

Received 6 December 1988
Revised 18 February 1989

Abstract: A connection between deterministic optimal control and large deviations theory has been known for a number of years, whereby the optimal performance index provides information about exit times of stable systems excited by noise, and the optimal trajectory provides information about exit trajectories asymptotically as the scalar multiplying the noise tends to zero. In this paper, a further connection is made to reverse-time modeling of stationary diffusions and linear, stationary, Gauss–Markov discrete-time systems, in which the drift part of the reverse-time model has the same trajectories as the closed loop system resulting from the solution of the same optimal control problem as used for the previous connection.

Keywords: Diffusion equations; Gauss–Markov processes; reverse-time processes; stochastic differential equations; optimal control; large deviations; exit problem.

1. Introduction

The use of optimal control techniques for obtaining large deviations results on the statistics of rare events in diffusion processes has been under development for a number of years (e.g. [1]). Consider a stationary diffusion process with the property that when the noise excitation is removed, all trajectories decay to the origin. Consider also a region surrounding the origin, and an initial condition for the diffusion process lying in this region. Large deviations results are concerned with identifying the expected time to exit from the region, and the most likely exit trajectory when the noise is small. It turns out that the solution of a deterministic optimal control problem for a system related to the diffusion process, where the task is to minimize control energy, provides information on the most likely path to be taken by a trajectory leaving some region containing the origin, and the expected time for this exit to occur. In fact, the trajectory of the closed loop system defines the most likely path for the diffusion, in the limit as noise tends to zero and the expected time to exit can be found from the optimal value of the performance index.

In this paper, we show that for linear stationary diffusions and linear stationary discrete-time Gauss–Markov processes the trajectory defining the solution of the optimal control problem associated with large deviations defines the mean path of the reverse-time model of the process [2]. We also show that for general stationary diffusions, if the input noise is small, then the solution of the optimal control control problem defines approximately the mean path of the reverse-time model.

Section 2 contains the derivation of the results for linear stationary diffusions, Section 3 contains the results for linear stationary Gauss–Markov processes. The general result for stationary diffusions is derived in Section 4. Section 5 presents a brief summary of the large deviations results relating the optimal cost of the optimal control problem to exit times.

2. Linear diffusion equations

In this section, we demonstrate that the construction of the reverse-time model of a linear stationary diffusion process corresponds exactly to solving a linear quadratic optimal control problem with the input energy as the performance index to be minimized.

Consider a system defined by the linear time-invariant stochastic differential equation

$$\mathrm{d}x = Ax\,\mathrm{d}t + B\,\mathrm{d}w, \tag{1}$$

* Work supported by Australian Telecommunications and Electronics Research Board (ATERB).

where $\mathrm{Re}\ \lambda_i(A) < 0$ for all i (i.e. all eigenvalues of the matrix A have negative real part, or the system $\dot{z} = Az$ is asymptotically stable), $[A, B]$ is controllable and w is a vector Wiener process such that $x(t)$ is independent of future increments $w(s) - w(t)$ for $s > t$ of w. The associated reverse-time model is defined by [2]

$$\begin{aligned} \mathrm{d}x &= \left(A + BB'P_R^{-1}\right)x\ \mathrm{d}t + B\ \mathrm{d}\overline{w} \\ &= -P_R A' P_R^{-1} x\ \mathrm{d}t + B\ \mathrm{d}\overline{w}, \end{aligned} \tag{2}$$

where $\overline{w}$ is a vector Wiener process such that $x(t)$ is independent of past increments $\overline{w}(s) - \overline{w}(t)$ for $s < t$ of $\overline{w}$, and P_R and $\mathrm{d}\overline{w}$ are defined by

$$P_R A' + A P_R = -BB', \tag{3}$$

$$\mathrm{d}\overline{w} = \mathrm{d}w - B'P_R^{-1}x\ \mathrm{d}t. \tag{4}$$

We will show that the mean trajectories of this reverse-time system are defined by the solution of the linear quadratic optimal control problem corresponding to the minimization of the control energy in moving a deterministic system from 0 to an arbitrary x_0. We consider the deterministic linear time-invariant system

$$\dot{x} = Ax + Bu \tag{5}$$

(which is obviously closely tied to (1)) and the performance index

$$V(x_0, u(\cdot)) = \frac{1}{2}\int_{-\infty}^{0} u(t)'u(t)\ \mathrm{d}t, \tag{6}$$

subject to the constraints $x(0) = x_0$ and $x(-\infty) = 0$. An open-loop solution to this problem of minimizing V is well known, see e.g. [3]. A form of closed-loop solution can be obtained a follows. Following [4], an associated optimal control problem is defined with the reverse time flow direction. The solution of this problem provides the optimal control as a constant linear feedback law that is stabilizing, and the law is expressible in terms of the solution of a linear matrix equation that is actually equivalent to a degenerate Riccati equation. When the time flow is reversed to recover a solution of the original problem, we find that at any intermediate point $x(t)$ on an optimal trajectory encountered at some time $t < 0$, the optimal control can be expressed in feedback form as [4]

$$u^*(t) = -B'Px(t) \tag{7}$$

where P is a positive definite matrix defined in terms of the optimal value of the performance index ($V^*(x_0)$):

$$-\tfrac{1}{2}x_0'Px_0 = V^*(x_0) = \min_{\substack{x(-\infty)=0,\\ x(0)=x_0}} V(x_0, u(\cdot)), \tag{8}$$

or in terms of the parameters of the original system (5), by the matrix Lyapunov equation

$$P^{-1}A' + AP^{-1} = BB', \tag{9}$$

and from (7), the closed loop system can be written as:

$$\begin{aligned} \dot{x} &= (A - BB'P)x \\ &= -P^{-1}A'Px. \end{aligned} \tag{10}$$

(Obviously, implementation of the optimal control in feedback form commencing infinitely far back in the past is not possible. However, if it is known that a certain state $x(t)$ lies on the optimal trajectory, from time t onwards, the optimal control can be implemented in feedback form, although in practice the instability of the closed loop system would likely preclude this.)

Now observe that (10) is precisely the mean trajectory of the reverse-time system above (2), because, by (3) and (9) there holds $P_R^{-1} = -P$. Hence the optimal control specified by requiring that the input energy to the system (5) be minimized gives a closed loop system corresponding to the reverse-time system associated with (1).

3. Linear discrete-time systems

In the previous section, it was shown that the two problems of finding the reverse-time model of a linear, stationary diffusion, and of finding a minimum-energy control are equivalent in the sense that the closed loop system associated with the minimum-energy control has the dynamics of the drift part of the reverse-time model. In this section, for completeness, a corresponding result is derived for linear, stationary, Gauss–Markov discrete-time systems.

We consider a system defined by the linear time-invariant stochastic difference equation

$$x(t+1) = Ax(t) + Bw(t), \tag{11}$$

where $|\lambda_i(A)| < 1$, $[A, B]$ is controllable, and for convenience, A is non-singular. The process $w(\cdot)$ is discrete, Gaussian, unit variance white noise, and such that $x(t)$ is independent of present and future values of w. The reverse-time model of this system is defined by [5]

$$\begin{aligned} x(t) &= A^{-1}\left(I - BB'P_R^{-1}\right)x(t+1) \\ &\quad + A^{-1}B\bar{w}(t) \\ &= P_R A' P_R^{-1} x(t+1) + A^{-1}B\bar{w}(t), \end{aligned} \tag{12}$$

where $\bar{w}$ is a Gaussian vector white-noise process such that $x(t)$ is independent of past values of $\bar{w}$ and, P_R and $\bar{w}$ are defined by:

$$P_R - AP_RA' = BB', \tag{13}$$

$$\bar{w}(t) = w(t) - B'P_R^{-1}x(t+1). \tag{14}$$

Consider also the linear time-invariant system

$$x(t+1) = Ax(t) + Bu(t) \tag{15}$$

and the performance index

$$V(x(0), u(\cdot)) = \frac{1}{2} \sum_{t=1-T}^{0} u'(t-1)u(t-1) \tag{16}$$

with the constraints $x(-T) = 0$ and $x(0) = x_0$. The aim is to minimize with respect to $u(\cdot)$ and T the cost of moving from $x(-T) = 0$ to $x(0) = x_0$. It is well known that this cost for fixed T, but minimized over $u(\cdot)$, is given by the inverse of a controllability Gramian (see e.g. [6] for the continuous problem), and, as in continuous time, the optimal cost when $T \to \infty$ and $u(\cdot)$) is obtained by using the infinite time Gramian. More precisely,

$$V^*(x_0) = -\tfrac{1}{2}x'Px \tag{17}$$

where

$$P^{-1} - AP^{-1}A' = -BB'. \tag{18}$$

The optimal control for fixed finite T is expressible in feedback form or open loop form. When $T \to \infty$, a formal expression for the control in feedback form is

$$u^*(t) = -(B'PB + I)^{-1}B'PAx(t), \tag{19}$$

but note that the same remarks regarding implementation apply here as in the last section.

Combining (15) and (19), the nominal closed loop system is

$$\begin{aligned} x(t+1) &= Ax(t) - B(B'PB+I)^{-1}B'PAx(t) \\ &= \left(I - B[B'PB+I]^{-1}B'P\right)Ax(t) \\ &= (I + BB'P)^{-1}Ax(t) \\ &= P^{-1}(AP^{-1}A')^{-1}Ax(t) \\ &= P^{-1}(A')^{-1}Px(t) \end{aligned} \tag{20}$$

and therefore,

$$x(t) = P^{-1}A'Px(t+1). \tag{21}$$

From (13) and (18), it can be seen that $P_R = -P^{-1}$, and hence from (21) that the solution of the optimal control problem defines the mean trajectory of the reverse time system given in (12).

4. Nonlinear diffusion equations

In this section, it is shown that the construction of the reverse-time model of a stationary nonlinear diffusion equation corresponds to the solution of the optimal control problem of minimizing the control energy when the noise in the diffusion is small. Consider a system and performance index to be minimized defined as follows:

$$\dot{x} = f(x) + g(x)u, \tag{22}$$

$$V(x(0), u(\cdot)) = \frac{1}{2}\int_{-\infty}^{0} u(t)'u(t)\,\mathrm{d}t, \tag{23}$$

and subject to the boundary conditions $x(0) = x_0$ and $x(-\infty) = 0$. We assume that $f(\cdot)$ and $g(\cdot)$ are continuously differentiable, $\dot{z} = f(z)$ is asymptotically stable to the origin, and (22) is controllable from any state to the origin and from the origin to any state. The optimal value of the cost (23) is $V^*(\cdot)$ parameterized by the terminal state, and satisfies the steady-state Hamilton–Jacobi equation

$$\left(\frac{\partial V^*}{\partial x}\right)' f(x) + \frac{1}{2}\left[\left(\frac{\partial V^*}{\partial x}\right)' g(x)\right]^2 = 0. \tag{24}$$

At any intermediate point $x(t)$ on an optimal trajectory, the optimal control $u^*(t)$ takes the value [4]

$$u^*(t) = g'(x(t)) \frac{\partial V^*(x(t))}{\partial x}. \tag{25}$$

Again, the remarks on implementation apply. In light of (25), however, we can define a nominal closed loop system by

$$\dot{x} = f(x) + g(x) g(x)' \frac{\partial V^*}{\partial x}. \tag{26}$$

Turning now to the problem of constructing reverse-time models, consider the stationary diffusion equation

$$\mathrm{d}x = f(x)\,\mathrm{d}t + \varepsilon g(x)\,\mathrm{d}w, \tag{27}$$

where w is a Wiener process such that $x(t)$ is independent of future increments $w(s) - w(t)$ for $s > t$ of w, and ε is a small scalar parameter; we wish to construct a reverse-time model

$$\mathrm{d}x = \bar{f}(x, t)\,\mathrm{d}t + \varepsilon \bar{g}(x)\,\mathrm{d}\bar{w}, \tag{28}$$

where $\bar{w}$ is a Wiener process such that $x(t)$ is independent of past increments $\bar{w}(s) - \bar{w}(t)$ for $s < t$ of $\bar{w}$. It is understood that the integral form of (28) involves a backward Ito integral. From [2], we know that $\bar{g}(\cdot) = g(\cdot)$ and that $\bar{f}$ is given by

$$\bar{f}^i = f^i - \frac{\varepsilon^2}{p} \sum_{j,k} \frac{\partial}{\partial x^j}\left[p g^{ik} g^{jk} \right] \tag{29}$$

where $p(x)$ is the stationary probability density of state x, which is guaranteed to exist whenever a strong solution exists to (27), $f(\cdot)$ and $g(\cdot)$ are twice continuously differentiable, their first order derivatives are bounded and the second order derivatives grow no faster than $\| x \|^m$ as $x \to \infty$ for some $m > 0$ [2]. (The state dependence of $f(\cdot)$, $g(\cdot)$ and $p(\cdot)$ is not explicitly stated in this and following equations in order to increase their clarity.) The density $p(x)$ obeys the Fokker–Planck equation (which is closely related to the forward Kolmogorov equation):

$$\begin{aligned} 0 = &-\sum_{i=1}^{n} \frac{\partial}{\partial x^i}\left[p f^i \right] \\ &+ \frac{1}{2} \sum_{i,j} \frac{\partial^2}{\partial x^i\, \partial x^j}\left[p \varepsilon^2 (g g')^{ij} \right]. \end{aligned} \tag{30}$$

Let $W_\varepsilon(x) = -\varepsilon^2 \log p(x)$. Note that

$$\lim_{\varepsilon \to 0} W_\varepsilon(x) = W_0(x), \quad \text{say,}$$

is an action function [7], and that W_0 is therefore positive, and does not tend to zero as $\varepsilon \to 0$. Substituting for $p(x)$ in (30),

$$\begin{aligned} 0 = &\sum_i \frac{1}{\varepsilon^2} \frac{\partial W_\varepsilon}{\partial x^i} f^i - \sum_i \frac{\partial f_i}{\partial x^i} - \sum_{i,j} \frac{1}{2} \frac{\partial^2 W_\varepsilon}{\partial x^i\, \partial x^j} (g g')^{ij} \\ &+ \sum_{i,j} \frac{1}{\varepsilon^2} \frac{1}{2} \frac{\partial W_\varepsilon}{\partial x^i} \frac{\partial W_\varepsilon}{\partial x^j} (g g')^{ij} \\ &- \sum_{i,j} \frac{1}{2} \frac{\partial W_\varepsilon}{\partial x^i} \frac{\partial}{\partial x^i} (g g')^{ij} \\ &- \sum_{i,j} \frac{1}{2} \frac{\partial W_\varepsilon}{\partial x^i} \frac{\partial}{\partial x^j} (g g')^{ij} \\ &+ \sum_{i,j} \frac{1}{2} \frac{\partial^2}{\partial x^i\, \partial x^j} (g g')^{ij}. \end{aligned} \tag{31}$$

Then if $\varepsilon \to 0$, formally we obtain (see [7] for a rigorous justification)

$$0 = \sum_i \frac{\partial W_0}{\partial x^i} f^i + \frac{1}{2} \sum_{i,j} \frac{\partial W_0}{\partial x^i} \frac{\partial W_0}{\partial x^j} (g g')^{ij}, \tag{32}$$

which is nothing other than the Hamilton–Jacobi equation (24) with $W_0(x)$ substituted for $V^*(x)$. This is guaranteed to have a unique positive definite solution under the conditions on the diffusion equation set out above [4]. Substituting for $p(x)$ in (29), we have

$$\begin{aligned} \bar{f}^i(x) = &f^i(x) + \left(g(x) g(x)' \frac{\partial W_\varepsilon}{\partial x} \right)^i \\ &- \varepsilon^2 \left(\frac{\partial}{\partial x} (g g') \right)^i. \end{aligned} \tag{33}$$

If ε is small, then $\bar{f}$ can be approximated by

$$\bar{f}(x) = f(x) + g(x) g(x)' \frac{\partial W_0(x)}{\partial x}, \tag{34}$$

and hence the reverse-time model of the system defined by (27) is

$$\begin{aligned} \mathrm{d}x = &\left(f(x) + g(x) g(x)' \frac{\partial W_0(x)}{\partial x} \right) \mathrm{d}t \\ &+ \varepsilon g(x)\,\mathrm{d}\bar{w}. \end{aligned} \tag{35}$$

It can be seen from this equation that the mean path of the reverse-time model is given by the solution of the optimal control problem described above with V^* replaced by W_0, which are one and the same. These calculations all assume ε small, and therefore are not as aesthetically appealing as those of the previous sections. Nevertheless, this captures an important class of problems, and we develop this theme briefly in the next section.

5. Large deviations and optimal control

Let D be a bounded region containing the origin with boundary ∂D. Define a stochastic process by

$$\mathrm{d}x = f(x)\,\mathrm{d}t + \varepsilon g(x)\,\mathrm{d}w, \tag{36}$$

where D is invariant for (36), $\dot{z} = f(z)$ is asymptotically stable and attracted to the origin from any point in D, w is a vector Wiener process, and $x(0) = x_0 \in D$. Denote the solution of (36) at any time t by $x^{x_0,\varepsilon}(t)$, and the first time for $x(\cdot)$ to reach the boundary ∂D by $\tau^{x_0,\varepsilon}$. Then we have

$$\tau^{x_0,\varepsilon} = \inf_{x^{x_0,\varepsilon}(t) \in \partial D} \{t \geq 0\}. \tag{37}$$

We now introduce a deterministic system:

$$\dot{z} = f(z) + g(z)u, \tag{38}$$

for which properties, amounting to nonlinear generalizations of controllability, are required [1]. For trajectories starting at x and ending at y, we define a performance index $V(x, y, u(\cdot), T)$, where x is the initial point and y the final point of a trajectory:

$$V(x, y, u(\cdot), T) = \frac{1}{2}\int_0^T |u(s)|^2\,\mathrm{d}s. \tag{39}$$

Define

$$V^* = \inf_{\substack{y \in \partial D,\\ u(\cdot), T}} V(0, y, u(\cdot), T) \tag{40}$$

and also

$$Y = \Big\{ y \in \partial D: \inf_{u(\cdot), T} V(0, y, u(\cdot), T) = V^* \Big\}. \tag{41}$$

Then for all $\eta > 0$,

$$\lim_{\varepsilon \downarrow 0} P\Big(\inf_{y \in Y} |x^{x_0,\varepsilon}(\tau^{x_0,\varepsilon}) - y| < \eta \Big) = 1, \tag{42}$$

$$\lim_{\varepsilon \downarrow 0} \varepsilon^2 \log E(\tau^{x_0,\varepsilon}) = V^*. \tag{43}$$

(42) says that trajectories of the diffusion process beginning at arbitrary $x_0 \in D$ (not just the origin) tend to exit at the point of least cost as $\varepsilon \to 0$. It is also true that exiting trajectories of the diffusion process tend to follow trajectories of least cost of the deterministic system [8]. Thus, if V has a unique minimum with respect to y, there is a unique trajectory around which all exit trajectories will cluster given small input noise. (43) relates the optimal value of the performance index to the expected time for an exit to occur.

Note that the deterministic system here and its performance index are the same as those used in the previous section. We have therefore shown, among other things, that given small input noise, the trajectories of a diffusion process that exit a region will cluster around a trajectory of the associated reverse-time model when the noise in that model is set to zero.

6. Conclusion

This paper has demonstrated a connection between large deviations and reverse-time modeling of random processes, and it has been shown that the mean trajectory of a reverse-time model is given by the solution of the same optimal control problem as provides information on exit times and exit paths for both diffusions and linear discrete-time systems. Further work needs to be done in order to establish whether or not these results can be extended to include more general classes of discrete-time systems. It would also be interesting to explore the application of these ideas to Large Deviations problems for other types of systems, e.g. finite-state Markov processes.

References

[1] J. Zabczyk, Exit problem and control theory, *Systems Control Lett.* **6** (1985) 165–172.

[2] B.D.O. Anderson, Reverse-time diffusion equation models, *Stochastic Process. Appl.* **12** (1982) 313–326.

[3] B.D.O. Anderson and J.B. Moore, *Linear Optimal Control* (Prentice-Hall, Englewood Cliffs, NJ, 1972).

[4] B.D.O. Anderson and P.V. Kokotovic, Optimal control problems over large time intervals, *Automatica* **23** (1987) 355–363.

[5] G. Verghese and T. Kailath, A further note on backwards markovian models, *IEEE Trans. Inform. Theory* **25** (1979) 121–124; Correction, p. 501.

[6] R.W. Brockett, *Finite Dimensional Linear Systems* (John Wiley and Sons, New York, 1970).

[7] M.R. James, Asymptotic Nonlinear Filtering and Large Deviations with Application to Observer Design, Ph.D. Thesis, University of Maryland (1988).

[8] M. Cottrell, J.C. Fort, and G. Malgouyres, Large deviations and rare events in the study of stochastic algorithms, *IEEE Trans. Automat. Control* **28** (1983) 907–918.

[9] S.M. Meerkov and T. Runolfsson, Residence time control, *IEEE Trans. Automat. Control* **33** (1988) 323–332.