Proceedings of the 31st Conference
on Decision and Control
Tucson, Arizona • December 1992

Suboptimal Control for Nonlinear Stochastic Systems

Fei-Yue Wang† and George N. Saridis‡
†Department of Systems and Industrial Engineering
University of Arizona, Tucson, Arizona 85721, USA
‡Department of Electrical, Computer, and Systems Engineering
Rensselaer Polytechnic Institute, Troy, New York 12180, USA

Abstract

Theoretical procedures are developed for comparing the performance of arbitrarily selected admissible feedback controls among themselves with that of the optimal solution of a nonlinear optimal stochastic control problem. Iterative design schemes are proposed for successively improving the performance of a controller until a satisfactory design is achieved. Specifically, the exact design procedure is based on the generalized Hamilton-Jacobi-Bellman equation for the value function of nonlinear stochastic systems, and the approximate design procedure for nonlinear stochastic regulator problem with infinite horizon is developed by using the upper and lower bounds to the value functions. For a given controller, both the upper and lower bounds to its value function can be obtained by solving a partial differential inequality. In particular, the upper and lower bounds to the optimal value function, which may be used as measure to evaluate the acceptability of suboptimal controllers, can be constructed without actually knowing the optimal controller.

1 Introduction

The problem of controlling a stochastic, dynamic system so that its behavior is optimal with respect to a performance index has received considerable attention over the past two decades. From a practical point of view, it is often desirable to obtain a feedback solution to the optimal control problem. In situations of linear stochastic systems with additive white Gaussian noise and quadratic performance indices (LQG problems), the separation principle is directly applicable, and the optimal control theory is well established with a high level of maturity [6]. However, due to the mathematical difficulties involved with stochastic processes, only fragmentary results are available for the optimal control of general especially, nonlinear, stochastic systems. While the optimal control theory for deterministic systems is at a respectable level of maturity, the corresponding theory for stochastic systems needs further developments for practical implementation and meaningful applications.

The objective of this paper is to develop an approximation theory that may be used to find some feasible, practical solutions to the optimal control of nonlinear stochastic systems. To this end, the problem of stochastic control is addressed from an inverse point of view: given an arbitrary selected admissible feedback control, how does it measure with respect to a given performance index, to other feedback controls, and how can it be successively improved to converge to the optimal ? This approach toward optimal control has been widely studied for nonlinear deterministic systems [1, 4, 5], and appeared more promising than the linearization type approximation methods which have met with limited success for highly nonlinear systems [2, 3].

This paper presents theoretical procedures for developing suboptimal feedback controllers for stochastic nonlinear systems as an extension of the Approximation Theory of Optimal Control developed by Saridis and Lee [4] for deterministic nonlinear systems.

2 Problem Formulation

For the purpose of obtaining explicit expressions, but without loss of generality since the results are immediately generalizable, consider a nonlinear stochastic control system described by the following stochastic differential equation,

$$dx = f(t,x)dt + b(t,x)udt + g(t,x)dw, \qquad t \in I \equiv [t_0, T] \quad (1)$$

where $x \in R^n$ is a vector of state of the stochastic system, $u \in \Omega_u \subset R^m$ is a control vector, Ω_u is a specified compact set of admissible controls, and $w \in R^k$ is a separable Wiener process. $f : I \times R^n \to R^n$, $b : I \times R^n \to R^{n\times m}$, and $g : I \times R^n \to R^{n\times k}$ are measurable system functions. It is assumed that feedback control $u(t,x)$ of Ω_u satisfies the following conditions, i) Linear Growth Condition:

$$\|f(t,x) + b(t,x)u(t,x)\| + \|g(t,x)\| \le a(1 + \|x\|)$$

ii) Uniform Lipschitz Condition:

$$\|(f - bu)(t,x) - (f - bu)(t,y)\| + \|g(t,x) - g(t,y)\| \le a\|x - y\|$$

where $(t,x),(t,y) \in I \times R^n$, $\|\cdot\|$ is Euclid normal operator, and a is some constant.

For a given *initial* state $x(t_0)=x_0$ (deterministic) and feedback control $u(t,x)$, the performance index of the system (1) is defined as,

$$J(u;t_0,x_0) = E\{\int_{t_0}^{T}[G + \|u\|^2]dt + h[T,x(T)]/x(t_0) = x_0\} \quad (2)$$

with $G : I \times R^n \to R^1$ and $h : I \times R^n \to R^1$ as non-negative functions. J is also called the *value function* of the system (1).

The infinitesimal generator of the stochastic process specified by (1) is defined to be,

$$\mathcal{L}_u\phi \equiv \frac{1}{2}tr[g^T(t,x)\phi_{xx}g(t,x)] + \phi_x^T[f(t,x) + b(t,x)u] \quad (3)$$

where $\phi : I \times R^n \to R^1$ has compact support and is continuous upon to all its second order derivatives, and $(\cdot)^T$ and $tr(\cdot)$ are transpose and trace operators, respectively. The differential operators are defined as,

$$()_t = \frac{\partial()}{\partial t}, \qquad ()_x = \frac{\partial()}{\partial x}, \qquad ()_{xx} = \frac{\partial^2()}{\partial x^2}$$

The *pre-Hamiltonian* function of the system with respect to the given performance index (4) and a control law $u(t,x)$ is defined as,

$$\mathcal{H}(x,\phi_x,\phi_{xx},u,t) = G(t,x) + \|u\|^2 + \mathcal{L}_u\phi \quad (4)$$

The optimal control of stochastic systems now can be stated as the following:

Optimal Stochastic Control Problem: For a given initial condition $(t_0,x_0) \in I \times R^n$, and the performance index (4), find $u^* \in \Omega_u$ such that

$$V^*(t_0,x_0) \equiv J(u^*;t_0,x_0) = \inf_{u\in\Omega_u} J(u;t_0,x_0) \quad (5)$$

If it is assumed that the optimal control law, $u^*(x,t)$, exists and if the corresponding value function, $V^*(x,t)$, is sufficiently smooth, then u^* and V^* may be found by solving the well-known Hamilton-Jacobi-Bellman equation,

$$V_t^* + \min\{\mathcal{L}_{u^*}V^* + G(t,x) + \|u^*\|^2\} = 0,$$
$$V[T,x(T)] = h[T,x(T)] \quad (6)$$

Unfortunately, except in the case of linear quadratic Gaussian controls, where the problem has been well solved [6], a closed-loop form solution of the Hamilton-Jacobi-Bellman for solving the optimal stochastic control problem cannot be obtained in general when the system of (1) is nonlinear.

Therefore, one may instead consider the optimal control problem relaxed to that of finding an admissible feedback control law, $u(x,t)$, that has an *acceptable (not necessarily optimal) value function*. This gives rise to a *suboptimal stochastic control problem* that could conceivably be solved with less difficulty than the original optimal stochastic control problem. The exact conditions for acceptability of a given value function are, of course, to be determined from practical considerations for the specific problem.

3 An Approximation Theory of Optimal Stochastic Control

This section contains the main results of the approximation theory for the solution of nonlinear stochastic control problems. Two theorems, one for the evaluation of performance of control laws and the other for the construction of lower and upper bounds of value functions, are established first. Then theoretical procedures which can lead to the iterative design of suboptimal controls are developed based on those two theorems.

Theorem 1 *Assume $V : I \times R^n \to R^1$ be an arbitrary function with continuous V, V_t, V_x, and V_{xx} and satisfy the condition:*

$$\|V\| + \|V_t\| + \|x\|\|V_x\| + \|x\|^2\|V_{xx}\| < b(1 + \|x\|^2) \tag{7}$$

where b is a suitable constant. Then the necessary and sufficient conditions for $V(t, x)$ to be the value function of an admissible fixed feedback control law $u(t, x) \in \Omega_u$, i.e.,

$$V(t,x) = E\{\int_t^T [G(t,x) + \|u\|^2]dt + h[T, x(T)]/x(t) = x\}, \tag{8}$$

are:

$$V_t + \mathcal{L}_u V + G(t,x) + \|u\|^2 = 0 \tag{9}$$

$$V[T, x(T)] = h[T, x(T)] \tag{10}$$

Proof: From (7), it follows from *Itô's integration formula* that,

$$V(t,x) = E\{V[T, x(T)] - \int_t^T [\mathcal{L}_u V(\tau, x(\tau)) + V_\tau(\tau, x(\tau))]d\tau/x(t) = x\}, t \in I$$

Therefore,

$$J(u;t,x) - V(t,x) = E\{h[T,x(T)] - V[T,x(T)] + \int_t^T [\mathcal{L}_u V(\tau,x(\tau)) + V_\tau(\tau,x(\tau)) + G(\tau,x) + \|u(\tau,x)\|^2]d\tau/x(t) = x\}, t \in I$$

the sufficient condition can be seen from the above equation immediately.

For necessary condition, assume $V(t,x) = J(u;t,x)$. Then from above equation, for $t = T$,

$$V[T, x(T)] = h[T, x(T)]$$

Therefore,

$$E\{\int_t^T [\mathcal{L}_u V(\tau,x(\tau)) + V_\tau(\tau,x(\tau)) + G(\tau,x) + \|u(\tau,x)\|^2]d\tau/x(t) = x\} = 0$$

has to be true for all $(t,x) \in I \times R^n$, hence,

$$V_t + \mathcal{L}_u V + G(t,x) + \|u\|^2 = 0$$

which proves the necessary condition.

Q.E.D.

Since it is generally difficult to find the exact value functions satisfying (9) and (10) of Theorem 1, the following theorem introduces a method of constructing the lower and upper bounds of value functions. This method can be used for the design of simpler suboptimal controllers based only on the upper bounds to value functions.

Theorem 2 (Lower and upper bounds of value function) *For an admissible fixed feedback control law $u(t,x) \in \Omega_u$ and a continuous function $s(t,x)$ with $|s(t,x)| < \infty$ for all $(t,x) \in I \times R^n$. If function $V(t,x)$ satisfies (7) with continuous V, V_t, V_x, and V_{xx}, and*

$$V_t + \mathcal{L}_u V + G(t,x) + \|u\|^2 \equiv \nabla V \le s(t,x) \le 0 (\ge s(t,x) \ge 0) \tag{11}$$

$$V[T,x(T)] \ge h[T,x(T)] \quad (\le h[T,x(T)]) \tag{12}$$

then $V(t,x)$ is an upper (or a lower) bound to the value function of system (1). That is,

$$V(t,x) \ge J(u;t,x) \quad (\le J(u;t,x)), \qquad \forall (t,x) \in I \times R^n \tag{13}$$

Proof: By the similar procedure used in the proof of Theorem 1, it can be shown that

$$J(u;t,x) - V(t,x) = E\{h[T,x(T)] - V[T,x(T)] + \int_t^T [\mathcal{L}_u V(\tau,x(\tau)) + V_\tau(\tau,x(\tau)) + G(\tau,x) + \|u(\tau,x)\|^2]d\tau/x(t) = x\}$$
$$E\{h[T,x(T)] - V[T,x(T)] + \int_t^T \nabla V(\tau,x(\tau))d\tau/x(t) = x\}$$

Therefore, from equations (11)-(12), it follows that,

$$J(u;t,x) - V(t,x) \le (\ge) E\{\int_t^T \nabla V(\tau,x(\tau))d\tau/x(t) = x\}$$
$$\le (\ge) E\{\int_t^T s(\tau,x(\tau))d\tau/x(t) = x\} \le 0 \quad (\ge 0)$$

for all $(t,x) \in I \times R^n$, which completes the proof.

Q.E.D.

Having established the two theorems for the evaluation of performance of a given feedback control law, now it is necessary to develop algorithms to improve the control law. In the followings Theorems 3-5 provide a theoretical procedure for designing the suboptimal feedback controllers based on the Theorem 1, while Theorem 6 presents a result for constructing upper and lower bounds to the optimal value function, which can be used to evaluate the acceptability of suboptimal controllers.

Theorem 3 *Given admissible controls $u_1 \in \Omega_u$ and $u_2 \in \Omega_u$, with $V_1(t,x)$ and $V_2(t,x)$ be the corresponding functions satisfying (7) and (8) for u_1 and u_2, respectively, define the Hamiltonian functions for $i = 1$ and 2:*

$$\mathcal{H}_{imin} = \mathcal{H}(x, V_{ix}, V_{ixx}, u_i^*, t) = G(t,x) + \|u_i^*\|^2 + \mathcal{L}_{u_i^*} V_i \tag{14}$$

where

$$u_i^*(t,x) = -\frac{1}{2} b^T(t,x) V_{ix}(t,x) \tag{15}$$

It can be shown that

$$V_1 \ge V_2 \tag{16}$$

when

$$V_{1t} + \mathcal{H}_{1min} \le V_{2t} + \mathcal{H}_{2min} \tag{17}$$

Proof: Let

$\Delta V = V_2 - V_1, \Delta V_t = V_{2t} - V_{1t}, \Delta V_x = V_{2x} - V_{1x}, \Delta V_{xx} = V_{2xx} - V_{1xx}$

then,

$$V_{2t} + \mathcal{H}_{2min} = V_{1t} + \Delta V_t + G(t,x) + \frac{1}{2}tr[g^T(t,x)V_{1xx}g(t,x)] + \frac{1}{2}tr[g^T(t,x)\Delta V_{xx}g(t,x)] +$$

$$V_{1x}^T f(t,x) + \Delta V_x^T f(t,x) - \frac{1}{4}\|b^T V_{1x}\|^2 - \frac{1}{4}\|b^T \Delta V_x\|^2 - \frac{1}{2}V_{1x}^T b b^T \Delta V_x$$
$$= V_{1t} + \mathcal{H}_{1min} - \frac{1}{4}\|b^T \Delta V_x\|^2 + \Delta V_t + \frac{1}{2}tr[g^T(t,x)\Delta V_{xx}g(t,x)] + \Delta V_x^T f(t,x) - \Delta V_x^T b(\frac{1}{2}b^T V_{1x})$$
$$= V_{1t} + \mathcal{H}_{1min} - \frac{1}{4}\|b^T \Delta V_x\|^2 + \Delta V_t + \mathcal{L}_{u_1^*}\Delta V$$

Therefore,

$$\Delta V_t + \mathcal{L}_{u_1^*}\Delta V = (V_{2t} + \mathcal{H}_{2min}) - (V_{1t} + \mathcal{H}_{1min}) + \frac{1}{4}\|b^T \Delta V_x\|^2 \tag{18}$$

which implies, from the assumption (18), that,

$$\Delta V_t + \mathcal{L}_{u_1^*}\Delta V \ge 0$$

In addition, from (10) of Theorem 1,

$$\Delta V(T,x) = V_2(T,x) - V_1(T,x) = 0$$

However, applying *Itô's integration formula* to $\Delta V(t,x)$ along the trajectory generated by control u_1^*, it follows that,

$$\Delta V = -E\{\int_t^T [\Delta V_\tau(\tau, x(\tau)) + \mathcal{L}_{u_1^*}\Delta V(\tau, x(\tau))]d\tau / x(t) = x\} \leq 0$$

Hence,

$$V_2(t,x) \leq V_1(t,x), \qquad \forall(t,x) \in I \times R^n$$

Q.E.D.

A combination of Theorems 1 and 3, in which case a $V(t,x)$ represents the value function of the system (1) when driven by control $u(t,x)$, yields an inequality as a basis of suboptimal control algorithms, to iteratively reduce the value of the performance of the system. This is summarized in the following theorem.

Theorem 4 *Assume that there exist a control $u_1 \in \Omega_u$ and a corresponding function $V_1(t,x)$ satisfying (7) and (8) of Theorem 1. If there exists a function $V_2(t,x)$ satisfying the same conditions of Theorem 1, of which the associated control $u_2 \in \Omega_u$ has been selected to satisfy*

$$\|u_2 + \frac{1}{2}b^T V_{2x}\| \leq \|u_1 + \frac{1}{2}b^T V_{1x}\| \tag{19}$$

then,

$$V_1 \geq V_2 \tag{20}$$

Proof: Since control u_1 and the corresponding value function V_1 must satisfy (9) and (10), according to Theorem 1, it follows that for every $(t,x) \in I \times R^n$,

$$V_{1t} + G(t,x) + V_{1x}^T(f + bu_1) + \frac{1}{2}tr(g^T V_{1xx} g) + \|u_1\|^2 = 0$$

This can be rewritten as,

$$V_{1t} + \mathcal{H}_{1min} + \|u_1 + \frac{1}{2}b^T V_{1x}\|^2 = 0,$$

that is,

$$V_{1t} + \mathcal{H}_{1min} = -\|u_1 + \frac{1}{2}b^T V_{1x}\|^2 \tag{21}$$

Similarly, one can find,

$$V_{2t} + \mathcal{H}_{2min} = -\|u_2 + \frac{1}{2}b^T V_{2x}\|^2 \tag{22}$$

Since

$$\|u_2 + \frac{1}{2}b^T V_{2x}\| \leq \|u_1 + \frac{1}{2}b^T V_{1x}\|,$$

it follows from equations (22) and (23) that,

$$V_{2t} + \mathcal{H}_{2min} \geq V_{1t} + \mathcal{H}_{1min}$$

Hence, according to Theorem 2,

$$V_2(t,x) \leq V_1(t,x), \qquad \forall(t,x) \in I \times R^n$$

which proves the theorem.

Q.E.D.

Based on Theorems 3 and 4, the following theorem establishes a sequence of feedback controls which are successively improved and converge to the optimal feedback control.

Theorem 5 *Let a sequence of pairs $\{u_i, V_i\}$ satisfy (7)-(8) of Theorem 1, and u_i be obtained by minimizing the pre-Hamiltonian function corresponding to the previous value function V_{i-1}, that is,*

$$u_i = -\frac{1}{2}b^T V_{i-1x}, \quad i = 1, 2, \ldots \tag{23}$$

then the corresponding value functions V_i satisfy the inequality,

$$V_{i-1} \geq V_i, \qquad i = 1, 2, \ldots \tag{24}$$

Thus by selecting the pairs $\{u_i, V_i\}$ in the above manner sequentially, the resulting sequence $\{u_i\}$ converges to the optimal control u^*, and the corresponding sequence $\{V_i\}$ converges monotonically to the optimal value function V^* associated with u^*.

Proof: Since control u_i of (24) and the corresponding value function V_i satisfy (9) and (10) of Theorem 1, it follows from (22) of Theorem 4 that,

$$\begin{aligned} V_{it} + \mathcal{H}_{imin} &= -\|u_i + \frac{1}{2}b^T V_{ix}\|^2 \\ &= -\|\frac{1}{2}b^T \Delta V_{ix}\|^2 \end{aligned}$$

where $\Delta V_i = V_i - V_{i-1}$.

Therefore, application of (19) of Theorem 2 leads to,

$$\begin{aligned} \Delta V_{it} + \mathcal{L}_{u_{i-1}^*}\Delta V_i = (V_{it} + \mathcal{H}_{imin}) - (V_{i-1t} + \mathcal{H}_{i-1min}) + \\ \frac{1}{4}\|\Delta V_{ix}^T b\|^2 = -\frac{1}{4}\|\Delta V_{ix}^T b\|^2 + \frac{1}{4}\|\Delta V_{i-1x}^T b\|^2 + \\ \frac{1}{4}\|\Delta V_{ix}^T b\|^2 = \frac{1}{4}\|\Delta V_{i-1x}^T b\|^2 \geq 0 \end{aligned}$$

From (10),

$$\Delta V_i(T,x) = V_i(T,x) - V_{i-1}(T,x) = 0,$$

hence, *Itô's integration formula* applied to ΔV_i along the trajectory generated by u_{i-1}^* leads to the inequality,

$$\Delta V_i = -E\{\int_t^T [\Delta V_{i\tau} + \mathcal{L}_{u_{i-1}^*}\Delta V_i]d\tau / x(t) = x\} \leq 0$$

that is,

$$V_{i-1} \geq V_i$$

which proves (25).

To show the convergence of the sequence, note that $\{V_i\}$ is a non-negative and monotonically decreasing sequence and satisfies (7), therefore, the following limits exist,

$$\lim_{i\to\infty} V_i(t,x) = V^o(t,x), \qquad \text{and} \tag{25}$$

$$\lim_{i\to\infty} V_{ix}(t,x) = V_x^o(t,x) \tag{26}$$

for all t and x, where V^o is the limit of value functions.

The corresponding limit of control sequence $\{u_i\}$ can be identified from (24) as,

$$u^o(t,x) = \lim_{i\to\infty} u_i = \lim_{i\to\infty}(-\frac{1}{2}b^T V_{i-1x}) = (-\frac{1}{2}b^T V_x^o)(t,x) \tag{27}$$

Clearly, u^o and V^o thus obtained still satisfy (9) and (10) of Theorem 1. However, from the construction of control sequence $\{u_i\}$, u^o minimizes the pre-Hamiltonian function associated with the value function V^o. In other words, u^o and V^o satisfy the Hamilton-Jacobi-Bellman equation for the optimal control of stochastic system (1),

$$V_t + \min_{u\in\Omega_u}\{\mathcal{L}_u v + G(t,x) + \|u\|^2\} = 0 \tag{28}$$

Hence,

$$u^o(t,x) = u^*(t,x), \quad V^o(t,x) = V^*(t,x), \forall(t,x) \in I \times R^n \tag{29}$$

are the optimal control and the optimal value function of the stochastic control problem (5).

Q.E.D.

Finally, the following theorem presents a method for the construction of an upper (or a lower) bound of the optimal value function $V^*(t,x)$. Since the optimal value function itself is extremely difficult to be found, its upper (or lower) bounds therefore can provide a practical measure to evaluate the effectiveness of suboptimal controllers.

Theorem 6 *Assume that there exists a function $V^s(t,x)$ satisfying the condition (7) of Theorem 1, for which the associated control*

$$u^s = -\frac{1}{2}b^T V_x^s(t,x) \tag{30}$$

is an admissible one. Then $V^s(t,x)$ is an upper (or a lower) bound to the optimal value function $V^*(t,x)$ of system (1) if it satisfies the following conditions,

$$V_t^s + \mathcal{L}_{u^s}V^s + G(t,x) + \|u^s\|^2 = s(t,x) \le 0 \quad (\ge 0) \tag{31}$$
$$V^s[T,x(T)] \ge h[T,x(T)] \quad (\le h[T,x(T)]) \tag{32}$$

where $s(t,x)$ is continuous and $|s| < \infty$, for all (t,x).

Proof: From Theorem 5 and the Hamilton-Jacobi-Bellman equation, it is obviously that for the optimal control and the optimal value function,

$$V_t^* + \mathcal{H}_{min}^* = V_t^* + G(t,x) + \|u^*\|^2 + \mathcal{L}_{u^*}V^* = 0, \tag{33}$$

and similarly, for u^s and V^s,

$$V_t^s + \mathcal{H}_{min}^s = V_t^s + G(t,x) + \|u^s\|^2 + \mathcal{L}_{u^s}V^s = s(t,x) \tag{34}$$

For $s(t,x) \le 0$, subtraction of (35) from (36) leads to,

$$\Delta V_t + \mathcal{L}_{u^s}\Delta V = -s(t,x) + \frac{1}{4}\|b^T\Delta V_x\|^2 \ge 0$$

where $\Delta V = V^*(t,x) - V^s(t,x)$. From assumption (34),

$$\Delta V[T,x(T)] = V^* - V^s = h[T,x(T)] - V^s[T,x(T)] \le 0$$

Therefore, application of *Itô's integration formula* to $\Delta V(t,x)$ along the trajectory generated by control u^s obtains,

$$\Delta V(t,x) \le -E\{\int_t^T [\Delta V_\tau + \mathcal{L}_{u^s}\Delta V]d\tau / x(t) = x\} \le 0$$

So $V^s(t,x)$ is an upper bound to the optimal value function $V^*(t,x)$

For $s(t,x) \ge 0$, subtraction of (36) from (35) leads to,

$$\Delta V_t + \mathcal{L}_{u^*}\Delta V = s(t,x) + \frac{1}{4}\|b^T\Delta V_x\|^2 \ge 0$$

where $\Delta V = V^s(t,x) - V^*(t,x)$. In the same way, one can show by using condition (34) and *Itô's integration formula* that,

$$\Delta V(t,x) \le -E\{\int_t^T [\Delta V_\tau + \mathcal{L}_{u^*}\Delta V]d\tau / x(t) = x\} \le 0$$

So in this case $V^s(t,x)$ is a lower bound to the optimal value function $V^*(t,x)$.

Q.E.D.

Theorems which can lead to the design of simpler suboptimal controllers based on the upper and lower bounds to value functions can also be constructed.

4 Stochastic Regulator Problem with Infinite Horizon

The stochastic regulator problem with infinite horizon is defined as a control problem for nonlinear stochastic system (1) with the (infinite-time) terminal state manifold taken as zero state, i.e.,

$$x(t) \to 0, \qquad \text{as } t \to \infty$$

and all state trajectories generated by admissible controls in Ω_u must be bounded uniformly in $I \times R^n$.

For the stochastic regulator problem with infinite horizon, the performance index of the system is defined to be,

$$J(u;t_0,x_0) = E\{\int_{t_0}^\infty [G(t,x) + \|u\|^2]dt / x(t_0) = x_0\}, \tag{35}$$

and *Itô's integration formula* applied to this case becomes,

$$V(t,x) = -E\{\int_t^\infty [\mathcal{L}_u V + V_\tau]d\tau / x(t) = x\}, \quad t \in I \tag{36}$$

where $V(t,x)$ satisfies $V(t,0) = 0$ and (8) of Theorem 1 for all the possible state trajectories (which is true for all $u \in \Omega_u$).

All the theorems developed in the previous section are still valid for the stochastic regulator problem with infinite horizon, except that all the terminal conditions at $t = T$ in those theorems are no longer required. However, in this case, theorems can be constructed which can lead to the iterative design of simpler suboptimal controls based only on the upper and lower bounds to value functions. Since in general the upper and lower bounds can be obtained without solving the partial differential equation (9) of Theorem 1, those theorems have a great potential for application. Corresponding to Theorems 3 and 4, two of such theorems are given in the sequel.

Theorem 7 *Given admissible controls u_1 and $u_2 \in \Omega_u$ with $J_1(t,x)$ and $J_2(t,x)$ be their corresponding value functions defined by (37), if there exist function pairs $\{V_1(t,x), s_1(t,x) \ge 0\}$ and $\{V_2, s_2 \le 0\}$ satisfying (11) of Theorem 2 of u_1 and u_2, respectively, then*

$$J_1 \ge J_2 \tag{37}$$

when

$$V_{1t} + \mathcal{H}_{1min} \le V_{2t} + \mathcal{H}_{2min} \tag{38}$$

Proof: Following the same procedure used in the proof for Theorem 3, one can show that

$$\Delta V_t + \mathcal{L}_{u_1^*}\Delta V = (V_{2t} + \mathcal{H}_{2min}) - (V_{1t} + \mathcal{H}_{1min}) + \frac{1}{4}\|b^T\Delta V_x\|^2 \ge 0$$

where $\Delta V = V_2 - V_1$. Thus *Itô's integration formula* (38) yields,

$$\Delta V(t,x) = -E\{\int_t^\infty [\Delta V_\tau + \mathcal{L}_{u_1^*}\Delta V]d\tau / x(t) = x\} \le 0,$$

hence,

$$V_2(t,x) \le V_1(t,x), \qquad \forall (t,x) \in I \times R^n$$

which implies that

$$J_2(t,x) \le V_2(t,x) \le V_1(t,x) \le J_1(t,x).$$

Q.E.D.

The next theorem is a counterpart of Theorem 4 and its proof can be carried out by the same procedure used in Theorem 4.

Theorem 8 *Assume that there exist a control $u_1 \in \Omega_u$ and a function pair $\{V_1(t,x), s_1(t,x) \ge 0\}$ satisfying (11) of Theorem 2. If there exists a function pair $\{V_2(t,x), s_2(t,x) \le 0\}$ satisfying the same condition of Theorem 2, of which the associated control $u_2 \in \Omega_u$ has been selected to satisfy*

$$\|u_2 + \frac{1}{2}b^T V_{2x}\| \le \|u_1 + \frac{1}{2}b^T V_{1x}\| \tag{39}$$

then,

$$J_1 \ge J_2 \tag{40}$$

J_1 and J_2 are the value functions of u_1 and u_2, respectively.

5 Design of Suboptimal Controllers

The optimal feedback control $u^*(t,x)$ and its associated $V^*(t,x)$ satisfying the Hamilton-Jacobi-Bellman equation equation (6), obviously satisfy all the theorems developed in Section 3. However, in most of cases of nonlinear stochastic control systems, the optimal solution is very difficult, if not impossible, to implement either because the solution is unavailable or because some of the states are not available for measure. In both cases the theory developed in Section 3 may serve to obtain controllers which can make the system stable, and then be successively modified to approximate the optimal solution. Upper and lower bounds of the value function of the nonlinear stochastic system may be used to evaluate the effectiveness of the approximation.

5.1 Exact Design Procedure

This approach, based on the assumption that the value function $V(t,x)$ for a control $u(t,x)$ can be found to satisfy (9) and (10) of Theorem 1, may be implemented according to Theorems 3-4 by the following procedure to control system (1).

1. Select a feedback control law $u_0(t,x)$ for system (1), set $i=0$.
2. Find a $V_i(t,x)$ to satisfy Theorem 1 for u_i.
3. Obtain a $u_{i+1}(t,x)$ and a $V_{i+1}(t,x)$ to satisfy Theorem 1, and Theorems 3 or 4 for u_i and V_i. u_{i+1} is an improved controller.
4. From Theorem 6, find a lower bound $V_L(t,x)$ to the optimal value function $V^*(t,x)$, and then use $V_{i+1}-V_L$ as a measure to evaluate $u_{i+1}(t,x)$ as an approximation to the optimal control $u^*(t,x)$. If acceptable, stop.
5. If the approximation is not acceptable, repeat Step 2 by increasing index i by one and continue

Example 1 (Linear stochastic systems): As the first example, the design procedure is applied to a linear stochastic system described by,

$$dx = A(t)xdt + B(t)udt + E(t)dw.$$

The cost function of the system has the quadratic form,

$$J = E\{\int_{t_0}^{T}[x^T M(t)x + \|u\|^2]dt + x(T)^T Qx(T)/x(t_0) = x_0\}$$

The infinitesimal generator of the linear stochastic process is,

$$\mathcal{L}_u V(t,x) = \frac{1}{2}tr[E(t)^T V_{xx}E(t)] + V_x^T[A(t) + B(t)u]$$

Assume first a linear control,

$$u_1(t,x) = -K_1(t)x$$

where $K_1(t)$ is a feedback matrix. The corresponding value function is assumed to be

$$V_1(t,x) = p_1(t) + x^T P_1(t)x,$$

where p_1 and P_1 can be found by solving (9) and (10) of Theorem 1, i.e.,

$$\frac{dP_1}{dt} + (A - BK_1)^T P_1 + P_1^T(A - BK_1) + M + K_1^T K_1 = 0, P_1(T) = Q,$$

and,

$$p_1(t) = \int_t^T tr[E(s)^T P_1(s)E(s)]ds$$

The feedback law can be improved by using Theorem 5. From equation (24),

$$u_i(t,x) = -\frac{1}{2}B^T(t)V_{i-1x} = -B^T P_{i-1}(t)x = -K_i(t)x, \qquad i \geq 2$$

and the corresponding value function is still assumed to be

$$V_i(t,x) = p_i(t) + x^T P_i(t)x, \qquad i \geq 2$$

where p_i and P_i are determined by solving

$$\frac{dP_i}{dt} + (A - BK_i)^T P_i + P_i^T(A - BK_i) + M + K_i^T K_i = 0, P_i(T) = Q,$$

and,

$$p_i(t) = \int_t^T tr[E(s)^T P_i(s)E(s)]ds$$

As $i \to \infty$, $P_i(t)$ approaches P, the solution of the matrix Riccati equation, i.e.,

$$\frac{dP}{dt} + (A - BK)^T P + P^T(A - BK) + M + K^T K = 0, K = B^T P$$

and correspondingly, the control approaches to,

$$u(t,x) = -B^T(t)P(t)x = -K(t)x$$

which is the optimal control for the linear stochastic systems with quadratic performance criterion [6].

Example 2: The second example illustrates the design method by the following nonlinear first-order stochastic system,

$$dx = -xdt + udt + \frac{x}{2}dw,$$

with a cost function,

$$J(u;t_0,x_0) = E\{\int_{t_0}^{\infty}[10x^2 + x^4 + u^2]dt/x(t_0) = x_0\}.$$

The infinitesimal generator of the stochastic process becomes,

$$\mathcal{L}_u V(t,x) = \frac{x^2}{4}V_{xx} + (-x+u)V_x$$

First assume a linear control law,

$$u_1(x) = -ax, \qquad a > 0$$

the corresponding value function is assumed to be,

$$V_1(x) = p_1x^2 + p_2x^4.$$

Equation (9) of Theorem 1 leads to,

$$[10 + a^2 - p_1(2a + \frac{7}{4})]x^2 + [1 - p_2(4a + \frac{5}{2})]x^4 = 0$$

which is true for

$$p_1 = \frac{40 + 4a^2}{8a + 7}, \qquad p_2 = \frac{2}{8a + 5}, \qquad a > 0$$

Next, select a higher order control law,

$$u_2(x) = -ax - bx^3, b > 0$$

the corresponding value function is assumed to be,

$$V_2(x) = q_1x^2 + q_2x^4.$$

To satisfy (9) in this case, it must have:

$$\begin{aligned}[10 + a^2 - q_1(2a + \frac{7}{4})]x^2 + [1 + 2ab - 2bq_1 \\ -q_2(4a + \frac{5}{2})]x^4 + (b^2 - 4bq_2)x^6 = 0\end{aligned} \qquad (41)$$

which is true for

$$q_1 = \frac{40 + 4a^2}{8a + 7}, q_2 = \frac{16a + 14}{675 - 16a}, b = 4q_2, 0 < a < \frac{675}{16} = 42.1875$$

To satisfy Theorem 4, controllers u_1 and u_2 must satisfy

$$\|u_2 + q_1x + 2q_2x^3\| \leq \|u_1 + p_1x + 2p_2x^3\|$$

which yields

$$0 < a \leq \frac{\sqrt{689} - 7}{8} = 2.4061$$

Their corresponding value functions can be compared as,

$$\begin{aligned}
V_1(x) &= \frac{40 + 4a^2}{8a + 7}x^2 + \frac{2}{8a + 5}x^4, \\
V_2(x) &= \frac{40 + 4a^2}{8a + 7}x^2 + \frac{16a + 14}{675 - 16a}x^4, \\
\Delta V(x) &= V_2(x) - V_1(x) = \frac{32(4a^2 + 7a - 40)}{(8a + 5)(675 - 16a)}x^4 \geq 0.
\end{aligned}$$

If Theorem 5 is to be used for the above u_1 and V_1, u_2 must be selected according to,

$$u_2(x) = -\frac{V_{1x}}{2} = -p_1x - 2p_2x^3,$$

and a V_2 satisfying (9) of Theorem 1 exists if

$$V_2(x) = q_1x^2 \quad + \quad q_2x^4,$$

$$q_1 = \frac{40 + 4p_1^2}{8p_1 + 7} = 2.7800, \qquad q_2 = \frac{16p_1 + 14}{675 - 16p_1} = 0.1393,$$
$$a = 0.2723, \qquad p_1 = 4.3904.$$

Comparing the values functions, one finds that

$$\begin{aligned} V_1(x) &= 4.3904x^2 + 0.2786x^4, \\ V_2(x) &= 2.7800x^2 + 0.1393x^4, \\ \Delta V(x) &= V_2(x) - V_1(x) = -1.6104x^2 - 0.1393x^4. \end{aligned}$$

5.2 Approximate Design Procedure for Regulator Problems

In many cases the selection of a $V(t, x)$ to satisfy (9) and (10) of Theorem 1 is a very difficult task. In such a case approximate design procedures, which use the upper and lower bounds of the value function obtained through Theorem 2 can be constructed. For the infinite time stochastic regulator problem, the following design procedure is proposed based on Theorems 7 and 8 in Section 4.

1. Select a feedback control law $u_0(t, x)$ for system (1), set $i = 0$.
2. For a $s_i(t, x) \geq 0$, find a $V_i(t, x)$ for u_i to satisfy Theorem 2 for a lower bound.
3. Obtain a $u_{i+1}(t, x)$, and for a $s_{i+1} \leq 0$, find a $V_{i+1}(t, x)$ for u_{i+1} to satisfy Theorem 2 for a upper bound. $u_{i+1}(t, x)$, s_{i+1}, and $V_{i+1}(t, x)$ found should also satisfy conditions (40) of Theorem 7 or (41) of Theorem 8 for the improvement of performance.
4. Using a lower bound to the optimal value function, which is determined according to Theorem 6, the approximation of the optimal control can be measured. If acceptable, stop.
5. If the approximation is not acceptable, repeat Step 2 by increasing index i by one and continue.

Example 3: The design method is illustrated with the following nonlinear first-order stochastic regulator problem,

$$dx = x^3 dt + u dt + \frac{x}{2} dw,$$

with a value function,

$$J(u; t_0, x_0) = E\{\int_{t_0}^{\infty} [10x^2 + x^4 + u^2] dt / x(t_0) = x_0\}.$$

The infinitesimal generator of the stochastic process is,

$$\mathcal{L}_u V(t, x) = \frac{x^2}{4} V_{xx} + (x^3 + u) V_x$$

For a linear control law,

$$u_1(x) = -a_1 x, \qquad a_1 > 0$$

the lower bound to its value function is assumed to be,

$$V_1(x) = p_1 x^2 + p_2 x^4.$$

And application of (11) of Theorem 2 leads to,

$$[10 + a_1^2 - p_1(2a_1 - \frac{1}{4})]x^2 + [1 +$$
$$2p_1 - p_2(4a_1 - \frac{3}{2})]x^4 + 4p_2 x^6 \geq 0,$$

which is satisfied by

$$p_1 = \frac{40(1 - c_1) + 4a_1^2}{8a_1 - 1}, p_2 = \frac{2(1 - c_2) + 4p_1}{8a_1 - 3}, a_1 > \frac{3}{8}$$

for any $1 \geq c_1, c_2 > 0$.

For a higher order control law,

$$u_2(x) = -a_2 x - b_2 x^3, \qquad a_2 > 0, \quad b_2 > 0,$$

the upper bound to its value function is assumed to be,

$$V_2(x) = q_1 x^2 + q_2 x^4.$$

And application of (11) yields in this case,

$$[10 + a_2^2 - q_1(2a_2 - \frac{1}{4})]x^2 + [1 + 2q_1 - 2(q_1 -$$
$$a_2)b_2 - q_2(4a_2 - \frac{3}{2})]x^4 + [b_2^2 - 4q_2(b_2 - 1)]x^6 \leq 0,$$

which is true for

$$q_1 = \frac{40(1 + d_1) + 4a_2^2}{8a_2 - 1}, q_2 = \frac{(1 + d_2)b_2^2}{4b_2 - 4},$$
$$1 + 2q_1 - 2(q_1 - a_2)b_2 - q_2(4a_2 - \frac{3}{2}) \leq 0, a_2 > \frac{3}{8}, b_2 > 1$$

where $d_1, d_2 > 0$ are arbitrary.

Improvement of performance $\Delta V \leq 0$ occurs if (40) or (41) is satisfied, which leads to

$$\begin{aligned} p_1 = \frac{40(1 - c_1) + 4a_1^2}{8a_1 - 1} &\geq q_1 = \frac{40(1 + d_1) + 4a_2^2}{8a_2 - 1} \\ p_2 = \frac{2(1 - c_2) + 4p_1}{8a_1 - 3} &\geq q_2 = \frac{(1 + d_2)b_2^2}{4b_2 - 4} \end{aligned}$$

which, with the rest of the inequalities, produce acceptable values for a_1, a_2 and b_2. For example, one can show that

$$a_1 = 10, \qquad a_2 = \frac{3}{2}, \qquad b_2 = 2, \qquad c_1 = c_2 = d_1 = d_2 = 0.1$$

is a set of the acceptable values. The lower and upper bounds of the value functions in this case are found as,

$$\begin{aligned} V_1(x) &= 5.5190x^2 + 0.3101x^4, \\ V_2(x) &= 4.8182x^2 + 0.0688x^4, \\ \Delta V(x) &= V_2(x) - V_1(x) = -0.7008x^2 - 0.2413x^4. \end{aligned}$$

Note that in this case the actual value function of control u_2 cannot be found by using the method applied in Example 2.

References

[1] R. J. Leake and R.-W. Liu, "Construction of Suboptimal Control Sequences," *SIAM J. Control*, Vol.5, No.1, pp54-63, 1967.

[2] D. L. Lukes, "Optimal Regulation of Nonlinear Dynamical Systems," *SIAM J. Control*, Vol.7, No.1, pp75-100, 1969.

[3] Y. Nishikawa, N. Sannomiya, and H. Itakura, "A Method for Suboptimal Design of Nonlinear Feedback Systems," *Automatica*, Vol.7, No.6, pp703-712.

[4] G. N. Saridis and C. S. G. Lee, "An Approximation Theory of Optimal Control for Trainable Manipulators," *IEEE Trans. on Systems, Man, and Cybernetics*, Vol.SMC-9, No.3, pp152-159, March 1979.

[5] G. N. Saridis and J. Balaram, "Suboptimal Control for Nonlinear Systems," *Control-Theory and Advanced Technology*, Vol.2, No.3, pp547-562, 1986.

[6] W. M. Wonham, "Random Differential Equations in Control Theory," In *Probabilistic Methods in Applied Mathematics*, edited by A. T. Bharucha-Reid, Academic Press, New York, 1970.