Inverse Optimal Control for Discrete-Time Stochastic Nonlinear Systems Stabilization

Santiago Elvira-Ceja and Edgar N. Sanchez

Abstract—This paper presents an inverse optimal control approach for stabilization in probability of discrete-time stochastic nonlinear systems. With the proposed scheme, we avoid to solve the associated stochastic Hamilton-Jacobi-Bellman equation; additionally a cost functional is minimized. This stabilizing optimal controller is based on a discrete-time stochastic control Lyapunov function. The applicability of the proposed approach is illustrated via numerical simulations.

I. INTRODUCTION

In optimal stochastic nonlinear control, we deal with the problem of determining a stabilizing control law for a given system such that a criterion is minimized; in order to do so, it is required to solve the associated stochastic Hamilton-Jacobi-Bellman (SHJB), which is not an easy task [1]. To overcome this difficulty, Crandall and Lions [2] introduced the so-called viscosity solution, while Deng and Krstic [3] used the inverse optimal control approach (both for the continuous-time case).

To avoid the discrete-time SHJB equation solution, in this paper, we use the inverse optimal control approach for stabilization in probability of discrete-time stochastic nonlinear systems. In this approach, a stabilizing feedback control law, based on a priori knowledge of a discrete-time stochastic control Lyapunov function (DSCLF), is synthesized first, and then it is ensured that this law optimizes a cost functional.

For inverse optimal control of deterministic systems, we refer the reader to the results presented in [4], [5], [6], [7], [8] (the continuous-time setting) and [9], [10], [11], [12] (for the discrete-time one). Although, there already exists an important result on inverse optimal control for continuous-time stochastic nonlinear systems [3], to the best of the authors knowledge, the discrete-time case is not yet analyzed, in spite of its advantages for real-time implementation.

The paper is organized as follows. Section II summarizes results on the stability of stochastic systems. In Section III, we introduce the class of stochastic systems we are dealing with and present the control problem to be solved. In section IV, we discuss the main result of this paper, which consists in synthesizing an stabilizing inverse optimal control law. Section V illustrates the applicability of the proposed controller by means of a numerical example. Finally, conclusions are stated.

This work is supported by CONACYT Mexico, under project 131678.

Both authors are with Automatic Control Laboratory, CINVESTAV IPN Unidad Guadalajara, Av. del Bosque 1145, Col. El Bajío 45019, Zapopan, Jalisco, Mexico. [jelvira, sanchez]@gdl.cinvestav.mx

II. STOCHASTIC STABILITY

Let consider a stochastic difference equation described by

\[ x_{k+1} = F(x_k, w_k), \quad x_0 = \hat{x}_0 \]  

where \( F(0, w_k) = 0 \) and \( \{w_k\}_{k \geq 0} \) is a sequence of random variables, \( x_0 \) is the initial condition which is also a random variable, and \( k \in \mathbb{Z}_+ = \{0, 1, 2, \ldots\} \).

Definition 1 [13]: System (1) is stable in probability at the origin if and only if, for every \( 0 < \rho_1 < 1 \) and \( \epsilon > 0 \) there exist \( 0 < \rho_2 < 1 \) and \( \delta = \delta(\rho_1, \epsilon) > 0 \) so that

\[ \Pr \left[ \sup_{k} \|x_k\| \geq \epsilon \right] < \rho_1 \]  

for all \( \hat{x}_0 \) such that

\[ \Pr \left[ \|\hat{x}_0\| \geq \delta \right] < \rho_2 \]

Definition 2 [13]: Let be a sequence of random variables \( \{X_k\}_{k \geq 0} \) such that the conditional expectation fulfills:

\[ E[X_k | X_{k-1}, \ldots, X_0] \leq X_{k-1} \quad \text{and} \quad E[X_0] < \infty \]

then this sequence is called a supermartingale one. The notations \( E[\cdot] \) and \( E[\cdot|\cdot] \) stand for mathematical expectation and conditional expectation, respectively.

Theorem 1 [13]: Suppose there exists a continuous positive definite function \( V(x) \) satisfying

(i) \( V(0) = 0 \)

(ii) \( V(x) \to \infty \) as \( \|x\| \to \infty \)

(iii) \( V(x_k) \) is a supermartingale along the motion of (1), then (1) is stable in probability.

Thus, we consider as a discrete-time stochastic control Lyapunov function (DSCLF) \( V(x_k, u_k) \), if this function satisfies Theorem 1.

III. PROBLEM STATEMENT

The stochastic sequences \( \{x_k\}_{k \geq 0}, \{w_k\}_{k \geq 0} \) and \( \{u_k\}_{k \geq 0} \) are defined on a probability space \((\Omega, \mathcal{F}, \mathbb{P})\) and connected by the stochastic difference equation

\[ x_{k+1} = f(x_k) + g(x_k)u_k + h(x_k)w_k, \quad x_0 = \hat{x}_0 \]

where \( x_k \in \mathbb{R}^n \) is the system state, \( u_k \in \mathbb{R}^{m_1} \) is the control input and \( w_k \in \mathbb{R}^{m_2} \) is a sequence of independent normal random variables with mean value equal to zero and covariance \( \Phi \); the smooth mappings \( f(x_k) : \mathbb{R}^n \to \mathbb{R}^n \), \( g(x_k) : \mathbb{R}^n \to \mathbb{R}^{n \times m_1} \), and \( h(x_k) : \mathbb{R}^n \to \mathbb{R}^{n \times m_2} \) are assumed known, \( f(0) = h(0) = 0 \) and \( g(x_k) \neq 0 \) for all \( x_k \neq 0 \). The initial condition \( x_0 \) is a normal random variable with mean value equals to \( m \) and covariance \( \Psi \),
such that \( x_k \) and \( w_k \) are independent. It is also assumed that real symmetric matrices \( \Psi \) and \( \Phi \) are positive semidefinite\(^1\) of appropriate dimension.

The problem to be addressed is as follows: to synthesize a feedback control law for system (5), such that the following three requirements are satisfied.

1. The control structure is given by
   \[
   u_k^* = -\frac{1}{2} R^{-1} g^T(x_k) E \left[ \frac{\partial V(x_{k+1})}{\partial x_{k+1}} \right] x_k \tag{6}
   \]
   where \( V(\cdot) \) is a twice differentiable positive definite function.

2. The controlled system achieves stability in probability at the origin.

3. The resulting control law minimizes an expectation cost functional
   \[
   J = E \left[ \sum_{k=0}^{\infty} (\ell(x_k) + u_k^T R u_k) \right] \tag{7}
   \]
   where \( \ell \) is a positive semidefinite\(^2\) function and \( R \) is a real symmetric positive definite\(^3\) weighting matrix.

IV. INVERSE OPTIMAL CONTROL SYNTHESIS

In order to solve the stated problem, we use the inverse optimal control approach. In this approach, a stabilizing feedback control law is first developed, and then it is demonstrated that this control law optimizes a cost functional. We establish the following definition which allows the inverse optimal control solution.

**Definition 3:** The control law (6) is inverse optimal stabilizing in probability if:

1. It achieves stability in probability of \( x = 0 \) for system (5);
2. \( V(x_k) \) is a positive definite function such that
   \[
   \dot{V} := E[V(x_{k+1}) | x_k] - V(x_k) + u_k^T R u_k^* \leq 0 \quad \text{and} \quad E[V_0] < \infty \tag{8}
   \]
   are fulfilled.

To obtain the stabilizing control law, based on Definition 3, we propose the function

\[
V(x_k) = \frac{1}{2} x_k^T P x_k, \tag{9}
\]

where \( P = P^T \in \mathbb{R}^{n \times n} \) is a real positive definite matrix.

Considering one step ahead for (9) and evaluating (6), then
\[
\begin{align*}
\dot{u}_k &= -\frac{1}{2} R^{-1} g^T(x_k) E \left[ \frac{\partial V(x_{k+1})}{\partial x_{k+1}} \right] x_k \\
&= -\frac{1}{2} R^{-1} g^T(x_k) E \left[ x_{k+1} \right] x_k \\
&= -\frac{1}{2} R^{-1} g^T(x_k) P (f(x_k) + g(x_k) u_k^*),
\end{align*}
\]

Solving (10) for \( u_k^* \) results in
\[
\begin{align*}
u_k^* &= -\frac{1}{2} \left( R + \frac{1}{2} g^T(x_k) P g(x_k) \right)^{-1} g^T(x_k) P f(x_k) \tag{11}
\end{align*}
\]

For notation simplification, (11) is rewritten as follows
\[
\begin{align*}
u_k^* &= \alpha(x_k) \\
&= -\frac{1}{2} (R + P_2(x_k))^{-1} P_1(x_k) \tag{12}
\end{align*}
\]

where \( P_1(x_k) = g^T(x_k) P f(x_k) \) and \( P_2(x_k) = \frac{1}{2} g^T(x_k) P g(x_k) \).

Note that \( P_2(x_k) \) is a positive definite and symmetric matrix, which ensures that the inverse matrix in (12) exists.

Once we have proposed a quadratic positive definite function \( V(x_k) \), according to Definition 3, the main contribution of this paper is presented as the following theorem.

**Theorem 2:** Consider the stochastic system (5). If there exists a real symmetric positive definite matrix \( P \) such that the following inequality holds:
\[
\begin{align*}
&\frac{1}{2} \left( f^T(x_k) P f(x_k) - x_k^T P x_k + t r[\Phi h^T(x_k) \Phi h(x_k)] \right) \\
&\quad -\frac{1}{4} P_1^T(x_k) (R + P_2(x_k))^{-1} P_1(x_k) \leq 0,
\end{align*}
\]

then, the equilibrium point \( x = 0 \) is stabilized in probability by the control law (12).

Moreover, this control law is inverse optimal in the sense that it minimizes the cost functional given by
\[
J(x_k) = E \left[ \sum_{k=0}^{\infty} (\ell(x_k) + u_k^T R u_k) \right],
\]

where \( \ell(x_k) = -\dot{V} \mid u_k = \alpha(x_k) \).

**Proof:** First, we analyze stability. Stability in probability for the equilibrium point \( x = 0 \) of system (5) with (12) as input, is guaranteed if (8) is satisfied; then
\[
\begin{align*}
\dot{V} &= E[V(x_{k+1}) | x_k] - V(x_k) + \alpha(x_k)^T R \alpha(x_k) \\
&= \frac{1}{2} E[x_{k+1}^T P x_{k+1} | x_k] - \frac{1}{2} x_k^T P x_k + \alpha(x_k)^T R \alpha(x_k) \tag{15}
\end{align*}
\]

The conditional expectation of \( x_{k+1}^T P x_{k+1} \), given the state \( x_k \), is formulated as
\[
\begin{align*}
E[x_{k+1}^T P x_{k+1} | x_k] &= E\left[ (f(x_k) + g(x_k) \alpha(x_k) + h(x_k) w_k)^T P(f(x_k) + g(x_k) \alpha(x_k) + h(x_k) w_k) \right] \\
&= (f(x_k) + g(x_k) \alpha(x_k))^T P(f(x_k) + g(x_k) \alpha(x_k)) \\
&\quad + 2E\left[ (f(x_k) + g(x_k) \alpha(x_k))^T P h(x_k) w_k \right] \\
&\quad + E[w_k^T h(x_k)^T P h(x_k) w_k] \tag{16}
\end{align*}
\]
Taking into account that 
\[ E[w_k^T \zeta w_k] = E[w_k^T \zeta E[w_k] + \text{tr}[\Phi \zeta], \] 
where \( \zeta = h(x_k)^T P h(x_k) \), and that \( E[w_k] = 0 \), then
\[
E[x_{k+1}^T P x_{k+1} | x_k] = (f(x_k) + g(x_k)\alpha(x_k))^T P \times (f(x_k) + g(x_k)\alpha(x_k)) + \text{tr}(\Phi h(x_k)^T P h(x_k))
\]
\[
= f_k^T P f_k + 2f_k^T P g_k\alpha(x_k) + \alpha^T(x_k)g_k^T P \times g(x_k)\alpha(x_k) + \text{tr}(\Phi h(x_k)^T P h(x_k)) \tag{17}
\]

Since \( g^T(x_k)P f(x_k) \) and \( \frac{1}{2}g^T(x_k)P g(x_k) \) are represented by \( P_1(x_k) \) and \( P_2(x_k) \), respectively, then
\[
E[x_{k+1}^T P x_{k+1} | x_k] = tr[\Phi h^T(x_k)P h(x_k)] + f^T(x_k)P \times f(x_k) - 2\alpha^T(x_k)P_2(x_k)\alpha(x_k) - 4\alpha^T(x_k)R\alpha(x_k)
\]
\[ \tag{18} \]

Substituting (18) into \( \tilde{V} \), we have
\[
\tilde{V} = \frac{1}{2} (f^T(x_k)P f(x_k) - x_k^T P x_k + \text{tr}(\Phi h^T(x_k)P) \times h(x_k)) - \alpha^T(x_k)P_2(x_k)\alpha(x_k) - \alpha(x_k)^T R\alpha(x_k)
\]
\[
= \frac{1}{2} (f(x_k) + g(x_k)\alpha(x_k))^T P(f(x_k) + g(x_k)\alpha(x_k)) + \text{tr}(\Phi h(x_k)^T P h(x_k))
\]
\[ = \frac{1}{2} (f(x_k) + g(x_k)\alpha(x_k))^T P(f(x_k) + g(x_k)\alpha(x_k)) - \frac{1}{4} P_1^T(x_k)(R + P_2(x_k))^{-1} P_1(x_k) \tag{19} \]

Replacing (12) in (19), then
\[
\tilde{V} = \frac{1}{2} (f^T(x_k)P f(x_k) - x_k^T P x_k + \text{tr}(\Phi h^T(x_k)P) \times h(x_k)) - \frac{1}{4} P_1^T(x_k)(R + P_2(x_k))^{-1} P_1(x_k)
\]
\[ \tag{20} \]

Selecting \( P \) such that \( \tilde{V} \leq 0 \), we obtain
\[
E[V(x_{k+1}) | x_k] - V(x_k) \leq \tilde{V}, \tag{21} \]
which means that \( V(x_k) \) is a DSCLF.

According to Theorem 1, stability in probability of \( x = 0 \) is guaranteed. Thus, from (20) follows condition (13).

Now we establish optimality. To do so, the term \( \ell(x_k) \) of (14) is replaced by \( -\tilde{V} \)
\[
J(x_k) = E \left[ \sum_{k=0}^{\infty} (-\tilde{V} + u_k^T R u_k) \right] \tag{22} \]

Hence:
\[
J(x_k) = E \left[ \sum_{k=0}^{\infty} (V_k - \frac{1}{2} f^T(x_k)P f(x_k)) \right. 
- \frac{1}{2} \text{tr}(\Phi h^T(x_k)P h(x_k)) 
+ \frac{1}{4} P_1^T(x_k)(R + P_2(x_k))^{-1} P_1(x_k) + u_k^T R u_k \right] \tag{23} \]

Factorizing (23) and adding the identity matrix \( I = (R + P_2(x_k))(R + P_2(x_k))^{-1} \), we obtain
\[
J(x_k) = E \left[ \sum_{k=0}^{\infty} (V(x_k) - \frac{1}{2} f^T(x_k)P f(x_k)) \right. 
- \frac{1}{2} \text{tr}(\Phi h^T(x_k)P h(x_k)) - \frac{1}{4} P_1^T(x_k)(R + P_2(x_k))^{-1} P_1(x_k) + \frac{1}{4} P_1^T(x_k)(R + P_2(x_k))^{-1} P_1(x_k) + u_k^T R u_k \right] \tag{24} \]

Being \( \alpha(x_k) = -\frac{1}{2} (R + P_2(x_k))^{-1} P_1(x_k) \), then (24) becomes
\[
J(x_k) = E \left[ \sum_{k=0}^{\infty} (V(x_k) - \frac{1}{2} f^T(x_k)P f(x_k)) \right. 
- \frac{1}{2} \text{tr}(\Phi h^T(x_k)P h(x_k)) - \frac{1}{4} P_1^T(x_k)(R + P_2(x_k))^{-1} P_1(x_k) + \frac{1}{4} P_1^T(x_k)(R + P_2(x_k))^{-1} P_1(x_k) + u_k^T R u_k \right] \tag{25} \]

Let \( \bar{x}_{k+1} \) be the mean value of (5) with \( u_k = \alpha(x_k) \); hence
\[
E[V(x_{k+1})] = \frac{1}{2} \bar{x}_{k+1}^T P \bar{x}_{k+1}
\]
\[
= \frac{1}{2} (f(x_k) + g(x_k)\alpha(x_k))^T P(f(x_k) + g(x_k)\alpha(x_k))
\]
\[
= \frac{1}{2} (f^T(x_k)P f(x_k) + 2 f^T(x_k)P g(x_k)\alpha(x_k) + \alpha^T(x_k)g^T(x_k)P g(x_k)\alpha(x_k))
\]
\[
= \frac{1}{2} (f(x_k) + g(x_k)\alpha(x_k))^T P f(x_k) + \alpha^T(x_k)P_2(x_k)\alpha(x_k)
\]
\[
+ P_1^T(x_k)\alpha(x_k)
\]
\[ \tag{26} \]

Using (26), cost functional (25) is simplified as
\[
J(x_k) = E \left[ \sum_{k=0}^{\infty} (V(x_k) - \bar{V}(x_{k+1})) \right. 
- \frac{1}{2} \sum_{k=0}^{\infty} \text{tr}(\Phi h^T(x_k)P h(x_k)) \right. 
+ E \left[ \sum_{k=0}^{\infty} (u_k^T R u_k - \alpha^T(x_k)R\alpha(x_k)) \right] \tag{27} \]

4691
which can be written as
\[
J(x_k) = \sum_{k=0}^{\infty} (\bar{V}(x_k) - \bar{V}(x_{k+1})) - \frac{1}{2} \sum_{k=0}^{\infty} \text{tr} [\Phi h^T(x_k)P h(x_k)] \\
+ E \left[ \sum_{k=0}^{\infty} (u_k^T R u_k - \alpha^T(x_k) R \alpha(x_k)) \right] \\
= \sum_{k=2}^{\infty} (\bar{V}(x_k) - \bar{V}(x_{k+1})) - \frac{1}{2} \sum_{k=0}^{\infty} \text{tr} [\Phi h^T(x_k)P h(x_k)] \\
+ E \left[ \sum_{k=0}^{\infty} (u_k^T R u_k - \alpha^T(x_k) R \alpha(x_k)) \right] - \bar{V}_2 + \bar{V}_1 - \bar{V}_1 + \bar{V}_0 \\
(28)
\]

For notation convenience in (28), the upper limit \(\infty\) is expressed as \(N \to \infty\), and thus
\[
J(x_k) = -\bar{V}_N - \lim_{N \to \infty} \frac{1}{2} \sum_{k=0}^{N} \text{tr} [\Phi h^T(x_k)P h(x_k)] \\
+ \lim_{N \to \infty} E \left[ \sum_{k=0}^{N} (u_k^T R u_k - \alpha^T(x_k) R \alpha(x_k)) \right] + \bar{V}_0 \\
(29)
\]

Since \(u_k = \alpha(x_k)\) is stabilizing in probability, \(\lim_{N \to \infty} \bar{V}_N = 0\), and thus
\[
\min_u J(x) = \bar{V}_0 + \lim_{N \to \infty} \frac{1}{2} \sum_{k=0}^{N} \text{tr} [\Phi h^T(x_k)P h(x_k)] \\
(30)
\]
where \(\bar{V}_0 = \frac{1}{2} m P m + \frac{1}{2} \text{tr} [\Psi P]\). Since the control law (12) minimizes the cost functional (14), it is optimal.

Remark 1: Additionally, \(V(x_k)\) solves the following discrete-time stochastic Hamilton-Jacobi-Bellman equation:
\[
\ell(x_k) + E[V(x_{k+1})|x_k)] - V(x_k) \\
+ \frac{1}{4} \frac{\partial V^T(x_{k+1})}{\partial x_{k+1}} g(x_k) R^{-1} g^T(x_k) \frac{\partial V(x_{k+1})}{\partial x_{k+1}} = 0 \\
(31)
\]

V. SIMULATIONS

In this section, we present an example to illustrate the applicability of the proposed approach. Simulations are done using MATLAB, a trade mark of The MathWorks, Inc.

Let the discrete-time stochastic nonlinear system (5) be defined as
\[
f(x_k) = \begin{bmatrix} x_{1,k} x_{2,k} \\ x_{1,k} \end{bmatrix}, \quad g(x_k) = \begin{bmatrix} x_{1,k}^2 \\ x_{2,k} \end{bmatrix} \\
(32)
\]
and
\[
h(x_k) = \begin{bmatrix} \sin(x_{2,k}) \\ 0 \end{bmatrix}
\]

According to Theorem 2, the control law is formulated as
\[
\alpha(x_k) = -5 x_{1}(k) x_{2}(k) (x_{1}^2(k) + 1) \\
(1 + 5 x_{1}^2(k) + 5 x_{2}^2(k))^{-1} \\
(33)
\]

where the positive definite matrix \(P\) is selected as
\[
P = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}
\]
and \(R\) is a constant term, \(R = 2\).

Simulation results are presented as follows: Fig. 1 and Fig. 2 show the sample realizations for \(x_{1,k}\) and \(x_{2,k}\), respectively. The initial condition is selected as a random variable with normal distribution, which has mean value \(m = [3, -1]^T\) and corvariance \(\Psi = [1, 0; 0, 1]^T\).

Fig. 1. Stabilization of state \(x_1\)

Fig. 2. Stabilization of state \(x_2\)

Fig. 3 presents the control action, which achieves stability in probability. Fig. 4 displays the evaluation of the cost
functional for which the control law is optimal. Additionally, Fig. 1, Fig. 2 and Fig. 3 include an enlargement in order to better appreciate randomness.

The stochastic sequence $w_k$ is simulated based on a standard normal distribution.

VI. CONCLUSION

This paper has developed an inverse optimal control scheme for discrete-time stochastic nonlinear systems. This scheme avoids the solution of the associated Hamilton-Jacobi-Bellman equation, achieves stabilization in probability, and minimizes a cost function. Simulation results illustrate the applicability of the proposed scheme.

REFERENCES