Successive Approximation Approach to the Hamilton-Jacobi-Bellman Equation via Quasilinearization for Nonlinear Control Systems

Akira Ohsumi (Prof. Emeritus of Kyoto Institute of Technology*)

Abstract—For the optimal control of nonlinear systems, an approach to solve the Hamilton-Jacobi-Bellman equation which is given by a nonlinear partial differential equation of second- or first-order is proposed based on the quasilinearization, and a successive approximation algorithm is obtained.

Key Words: Hamilton-Jacobi-Bellman equation, nonlinear control system, quasilinearization, successive approximation

1 Introduction

It is prerequisite to solve the Hamilton-Jacobi-Bellman (HJB) equation in order to execute the optimal control for both linear and nonlinear systems1). In the context of LQG (linear-quadratic-Gaussian) or LQ framework, its solution is given in a quadratic form accompanied by the matrix Riccati differential equation2). However, for nonlinear systems, we have to face up to solve numerically the HJB equation which is given as a nonlinear partial differential equation of first-order (in the deterministic case) or of second-order (in the stochastic case). As for the nonlinear control systems, several investigations have been done up to the present time, mainly, to obtain the stationary control for deterministic systems (e.g., see Ref. 3)-5) and references cited therein).

In this paper, a feasible method is proposed to realize the optimal control for both stochastic and deterministic systems by introducing the quasilinearization to the nonlinear HJB equations.

2 Hamilton-Jacobi-Bellman Equation

Let $x(t)$ be an $n$-vector process of a nonlinear stochastic system connected affinely with a control term. The system is described by the Itô stochastic differential equation2),

$$dx(t) = f[t, x(t)] dt + C[t, x(t)]u(t) dt$$
$$+ G(t, x(t))dw(t), \quad x(0) = x_0 \quad (0 \leq t \leq T),$$

(1)

where $u(t) \in R^l$ ($l \leq n$) is a control to be determined; $w(t) \in R^d$ is a Wiener process with zero-mean and covariance matrix $Q(t)$, $E\{dw(t)[dw(t)]^T\} = Q(t)dt$; $f(t, \cdot)$ is an $n$-vector-valued nonlinear function in $x$ and $C(t, \cdot)$, $G(t, \cdot)$ are matrix-valued functions satisfying the Itô conditions, viz., Lipschitz and growth conditions2).

The process $\{x(t)\}$ should be controlled by selecting $\{u(t), 0 \leq t \leq T\}$ in order to minimize the cost functional,

$$J(u) = \mathcal{E}\left\{ \phi[x(T)] + \int_0^T \{L[t,x(t)] + u^T(t)N(t)u(t)\} dt \right\}$$

(2)

in which $\phi(\cdot)$ and $L(t, \cdot)$ are positive-definite functions, and $N(t)$ is an $\ell \times \ell$-symmetric and positive-definite matrix. For this cost functional, let the minimal cost functional (value function) be

$$S(t,x) = \min_{u(\tau)} \mathcal{E}\left\{ \phi[x(T)] + \int_t^T \{L[\tau,x(\tau)] + u^T(\tau)N(\tau)u(\tau)\} d\tau \mid x(t) = x \right\},$$

(3)

where $\mathcal{E}\{* \mid x(t) = x\}$ denotes the conditional expectation of $*$ conditioned that $x(t)$ has taken its realization $x$ at $t$. Then, via the Bellman’s principle of optimality, the control problem leads us to solve the dynamic programming equation2),

$$- \frac{\partial S(t,x)}{\partial t} = \min_{u(t)} \left[ L(t,x) + u^T(t)N(t)u(t) \right.$$
$$+ \left( \frac{\partial S(t,x)}{\partial x} \right)^T \left[ f(t,x) + C(t,x)u(t) \right]$$
$$\left. + \frac{1}{2} \text{tr}\left[ G(t,x)Q(t)G^T(t,x) \frac{\partial}{\partial x}\left( \frac{\partial S(t,x)}{\partial x} \right)^T \right] \right].$$

(4)

The optimal control $u^o(t)$ is given by performing the minimization of the R.H.S. of (4) as follows:

$$u^o(t) = - \frac{1}{2} N^{-1}(t)C^T(t,x) \frac{\partial S(t,x)}{\partial x},$$

(5)

where the gradient is defined here as $\partial S/\partial x = [\partial S/\partial x_1, \cdots, \partial S/\partial x_n]^T$. Substituting (5) into (4) leads us to solve the following nonlinear partial

---

* 56 Nakonodacho, Ichijoji, Sakyo, Kyoto 606-8163, Japan (e-mail: akiraspika@nifty.com)
differential equation of second-order,
\[
\begin{aligned}
\frac{\partial S(t,x)}{\partial t} &= L(t,x) + f^T(t,x) \frac{\partial S(t,x)}{\partial x} \\
+ \frac{1}{2} &\text{tr}\left\{G(t,x)Q(t)G^T(t,x) \frac{\partial}{\partial x} \left(\frac{\partial S(t,x)}{\partial x}\right)^T\right\} \\
- \frac{1}{4} &\left(\frac{\partial S(t,x)}{\partial x}\right)^T C(t,x) N^{-1}(t) C^T(t,x) \frac{\partial S(t,x)}{\partial x} \\
S(T,x) &= \phi(x), \quad x \in \mathbb{R}^n.
\end{aligned}
\] (6)

In the sequel, we assume that \(G(t,x)Q(t)G^T(t,x) > 0\) and \(C(t,x)N^{-1}(t)C^T(t,x) > 0\) for all \((t,x) \in [0,T] \times \mathbb{R}^n\). The case where \(Q(t) \equiv 0\) (i.e., the deterministic case) is investigated in Section 6.

At this stage, there seem mainly two approaches to solve the HJB equation (6): (i) One is to solve it by some iteration method; and (ii) the other is to convert it to a linear one via the logarithmic Cole-Hopf transformation and to represent its solution as path integrals via the Feynman-Kac formula. In this paper, the first approach is investigated, while the second one is presented in a companion paper \(^9\).

Remark 1: In (2), suppose that \(f(t,x) \equiv f(x), C(t,x)N^{-1}(t)C^T(t,x) = R(x), L(t,x) \equiv L(x), \phi(x) \equiv 0\) and \(T \to \infty\). Then, for the deterministic case \((Q(t) \equiv 0)\), the value function \(S(t,x)\) becomes stationary, viz. \(\partial S(t,x)/\partial t = 0\), so (6) reduces to
\[
L(x) + f^T(x)p - \frac{1}{4} p^T R(x)p = 0
\]
in which \(p = \partial S(x)/\partial x\). Such a stationary HJB equation is treated in Ref. 4) under the assumptions \(f(x) = Ax + O(\|x\|^2)\) and \(L(x) = x^T M x + O(\|x\|^3)\).

3 Successive Approximation to the Solution of HJB Equation

First, recall that the fourth term in the R.H.S. of (6) is quadratic in \(\partial S(t,x)/\partial x\), so that there seems an analogy between the second-order nonlinear partial differential equation (6) and the matrix Riccati differential equation appearing in the LQG/LQ control problem \(^1\) or the Kalman filter. \(^1\) The solution of the Riccati equation can be obtained using the Bellman’s quasilinearization method \(^1\). In order to solve

\(^1\) The problem formulation of the Kalman filter to obtain the optimal estimation of the state \(x(t)\) of the linear system is set as \(^2\)

(System) \(dx(t) = Ax(t) dt + G dw(t)\)

(Observation) \(dy(t) = Hx(t) dt + dw(t)\)

(w(t) \~ N[0, Q dt], v(t) \~ N[0, R dt]). The estimation error covariance satisfies the Riccati differential equation,

\[
P(t) = AP(t) + P(t)A^T + QG^T - P(t)H^T R^{-1} H P(t), \quad P(0) = P_0.
\]

In case of \(H \equiv 0\) (no observation data) this reduces to

\[
P(t) = AP(t) + P(t)A^T + QG^T, \quad P_L(0) = P_{Lo}
\]

which is the Lyapunov differential equation (whose stationary algebraic form is the familiar Lyapunov equation), and it holds that \(P(t) \leq P_L(t)\) for all \(t\) if \(P_0 = P_{Lo}(\geq 0)\).

(6), let us extend the idea of this quasilinearization to the second-order nonlinear partial differential equation. To do this, consider
\[
\begin{aligned}
- \frac{\partial S_{n+1}(t,x)}{\partial t} &= L(t,x) + f^T(t,x) \frac{\partial S_{n+1}(t,x)}{\partial x} \\
+ \frac{1}{2} &\text{tr}\left\{Q_0(t,x) \frac{\partial}{\partial x} \left(\frac{\partial S_{n+1}(t,x)}{\partial x}\right)^T\right\} \\
- \frac{1}{4} &\left(\frac{\partial S_{n+1}(t,x)}{\partial x}\right)^T N_0(t,x) \frac{\partial S_{n+1}(t,x)}{\partial x} \\
- \frac{1}{4} &\left(\frac{\partial S_0(t,x)}{\partial x}\right)^T N_0(t,x) \frac{\partial S_0(t,x)}{\partial x} \\
S_{n+1}(T,x) &= \phi(x), \quad n = 0, 1, 2, \ldots
\end{aligned}
\] (7)

where \(Q_0(t,x) = G(t,x)Q(t)G^T(t,x)\) and \(N_0(t,x) = C(t,x)N^{-1}(t)C^T(t,x)\). Note that (7) is a linear equation for \(S_{n+1}(t,x)\). The initial value \(S_0(t,x)\) is set as \(S_0(t,x) = 0\) for all \((t,x) \in [0,T] \times \mathbb{R}^n\).

**Theorem:** Assume that \(Q_0(t,x)\) and \(N_0(t,x)\) are bounded. Then, the quasilinearized partial differential equation (7) for \(S_{n+1}(t,x)\) constitutes a convergent sequence \(\{S_n(t,x)\}_{n=0,1,2,\ldots}\) such that
\[
\lim_{n \to \infty} S_n(t,x) = S(t,x).
\] (8)

As an initial value \(S_0(t,x)\), we employ the solution of the linear partial differential equation,
\[
\begin{aligned}
- \frac{\partial S_0(t,x)}{\partial t} &= L(t,x) + f^T(t,x) \frac{\partial S_0(t,x)}{\partial x} \\
+ \frac{1}{2} &\text{tr}\left\{Q_0(t,x) \frac{\partial}{\partial x} \left(\frac{\partial S_0(t,x)}{\partial x}\right)^T\right\} \\
S_0(T,x) &= \phi(x).
\end{aligned}
\] (9)

The proof of Theorem is lengthy, so it is given in the next section.

4 Proof of Theorem

A. Proof of Theorem In this section the theorem stated in the previous section is proved according to the following procedure showing (i) first that \(S_n(t,x) \geq S(t,x)\), (ii) second, \(S_{n+1}(t,x) \leq S_n(t,x)\), and (iii) \(\lim_{n \to \infty} S_n(t,x) = S(t,x)\).

(i) Proof of \(S_n(t,x) \geq S(t,x)\): Recall that
\[
\left(\frac{\partial S}{\partial x} - \frac{\partial S_n}{\partial x}\right)^T N_0 \left(\frac{\partial S}{\partial x} - \frac{\partial S_n}{\partial x}\right) \geq 0
\]
since \(N_0(t,x) > 0\). So, we have
\[
\begin{aligned}
\left(\frac{\partial S}{\partial x}\right)^T N_0 \frac{\partial S}{\partial x} &\geq \left(\frac{\partial S_n}{\partial x}\right)^T N_0 \frac{\partial S_n}{\partial x} \\
+ \left(\frac{\partial S}{\partial x}\right)^T N_0 \frac{\partial S_n}{\partial x} &\geq \left(\frac{\partial S_n}{\partial x}\right)^T N_0 \frac{\partial S_n}{\partial x}.
\end{aligned}
\] (10)
Using this inequality in (6), we get
\[-\frac{\partial S}{\partial t} \leq L(t,x) + f^T(t,x) \frac{\partial S}{\partial x} + \frac{1}{2} \text{tr} \left\{ Q_0(t,x) \left( \frac{\partial S}{\partial x} \right)^T \right\} \]
\[-\frac{1}{4} \left( \frac{\partial S_T}{\partial x} \right)^T N_0(t,x) \frac{\partial S}{\partial x} + \left( \frac{\partial S}{\partial x} \right)^T N_0(t,x) \frac{\partial S_T}{\partial x} \]
\[-\left( \frac{\partial S_T}{\partial x} \right)^T N_0(t,x) \frac{\partial S}{\partial x} \right], \quad S(T,x) = \varphi(x). \quad (11)\]

Subtracting (7) from (11) yields
\[-\frac{\partial (S - S_{\nu+1})}{\partial t} \leq \frac{1}{2} \text{tr} \left\{ Q_0(t,x) \left( \frac{\partial (S - S_{\nu+1})}{\partial x} \right)^T \right\} \]
\[-\frac{1}{4} \left( N_0(t,x) \frac{\partial S_T}{\partial x} - 2 f(t,x) \right)^T \frac{\partial (S - S_{\nu+1})}{\partial x} \]
\[+ \left( \frac{\partial (S - S_{\nu+1})}{\partial x} \right)^T \left( N_0(t,x) \frac{\partial S_T}{\partial x} - 2 f(t,x) \right)]. \]

Here, define \( \Sigma(t,x) := S(t,x) - S_{\nu+1}(t,x) \). Then, we have
\[-\frac{\partial \Sigma(t,x)}{\partial t} \leq \frac{1}{2} \text{tr} \left\{ Q_0(t,x) \left( \frac{\partial \Sigma(t,x)}{\partial x} \right)^T \right\} \]
\[-\frac{1}{4} \left( h^T(t,x) \frac{\partial \Sigma(t,x)}{\partial x} \right) + \left( \frac{\partial \Sigma(t,x)}{\partial x} \right)^T h(t,x) \]
\[\Sigma(T,x) = 0 \quad (x \in \mathbb{R}^n), \quad (12)\]

where \( h(t,x) = N_0(t,x) (\partial S_T(x)/\partial x) - 2 f(t,x) \).

Pick a nonnegative-definite function \( E(t,x) \geq 0 \) such that
\[-\frac{\partial \Sigma(t,x)}{\partial t} = \frac{1}{2} \text{tr} \left\{ Q_0(t,x) \left( \frac{\partial \Sigma(t,x)}{\partial x} \right)^T \right\} \]
\[-\frac{1}{4} \left( h^T(t,x) \frac{\partial \Sigma(t,x)}{\partial x} \right) + \left( \frac{\partial \Sigma(t,x)}{\partial x} \right)^T h(t,x) \]
\[-E(t,x), \quad \Sigma(T,x) = 0. \quad (13)\]

For this, let \( \tau = T - t \). Then,
\[-\frac{\partial \Sigma(t,x)}{\partial t} = -\frac{\partial \Sigma(T-t,x)}{\partial \tau} \quad (d\tau > 0) \]
and (13) is converted to its time-reversal form,
\[-\frac{\partial \bar{\Sigma}(\tau,x)}{\partial \tau} = \frac{1}{2} \text{tr} \left\{ Q_0(T-\tau,x) \left( \frac{\partial \bar{\Sigma}(\tau,x)}{\partial x} \right)^T \right\} \]
\[-\frac{1}{4} \left( h^T(T-\tau,x) \frac{\partial \bar{\Sigma}(\tau,x)}{\partial x} \right) + \left( \frac{\partial \bar{\Sigma}(\tau,x)}{\partial x} \right)^T h(T-\tau,x) \]
\[-E(T-\tau,x), \quad \bar{\Sigma}(0,x) = 0 \quad (0 \leq \tau \leq T), \quad (14)\]
where \( \bar{\Sigma}(\tau,x) := \Sigma(T-\tau,x) \). The solution of (14) is given as \( \bar{\Sigma}(\tau,x) = \int_{0}^{\infty} \Gamma(\tau,x;0,\xi) \bar{\Sigma}(0,\xi) \, d\xi \)
\[-\int_{0}^{T} \int_{\mathbb{R}^n} \Gamma(\tau,x;\sigma,\xi) E(\sigma,\xi) \, d\xi \, d\sigma \]
\[-\int_{0}^{T} \int_{\mathbb{R}^n} \Gamma(\tau,x;\sigma,\xi) E(\sigma,\xi) \, d\xi \, d\sigma \quad (15)\]
in which \( \bar{\Sigma}(0,\xi) \equiv 0 \) by its initial condition, and \( \Gamma \) is a fundamental solution to the homogeneous equation (14) with \( E(T-\tau,x) \equiv 0 \) for all \( (\tau,x) \in [0,T] \times \mathbb{R}^n \). Since \( E(\sigma,\xi) \geq 0 \) for all \( (\sigma,\xi) \) and the fundamental solution \( \Gamma \) is positive and bounded from below,\(^7\) Chap. 2-54, it follows that \( \bar{\Sigma}(\tau,x) \leq 0 \), or \( \bar{\Sigma}(\tau,x) \leq 0 \) for all \( (\tau,x) \in [0,T] \times \mathbb{R}^n \). This fact implies that \( \Sigma(t,x) = S(t,x) - S_{\nu+1}(t,x) \leq 0 \), or
\[S(t,x) \leq S_{\nu+1}(t,x) \quad \text{for } \nu = 1, 2, \ldots. \quad (16)\]

(ii) Proof of \( S_{\nu+1}(t,x) \leq S_{\nu}(t,x) \): Noting that
\[-\frac{\partial S_{\nu-1}}{\partial x} + \frac{\partial S_{\nu}}{\partial x} \right) N_0 \left( \frac{\partial S_{\nu-1}}{\partial x} - \frac{\partial S_{\nu}}{\partial x} \right) \geq 0, \]
we get the inequality,
\[-\frac{\partial S_{\nu}}{\partial x} \right) N_0 \left( \frac{\partial S_{\nu-1}}{\partial x} \right) \frac{\partial S_{\nu}}{\partial x} \quad -\frac{1}{4} \left( \frac{\partial S_{\nu}}{\partial x} \right) N_0(t,x) \frac{\partial S_{\nu}}{\partial x} \quad (17)\]

Replacing the subscript \( \nu + 1 \) by \( \nu \) in (7) and applying the inequality (17) to the fourth term in the R.H.S. of the result (7) for \( S_{\nu}(t,x) \), we have
\[-\frac{\partial S_{\nu}}{\partial t} \geq L(t,x) + f^T(t,x) \frac{\partial S_{\nu}}{\partial x} \]
\[+ \frac{1}{2} \text{tr} \left\{ Q_0(t,x) \left( \frac{\partial S_{\nu}}{\partial x} \right)^T \right\} \]
\[-\frac{1}{4} \left( \frac{\partial S_{\nu}}{\partial x} \right) N_0(t,x) \frac{\partial S_{\nu}}{\partial x} \quad S_{\nu}(T,x) = \varphi(x). \quad (18)\]

Subtracting (18) from (7), we get
\[-\frac{\partial (S_{\nu+1} - S_{\nu})}{\partial t} \leq \frac{1}{2} \text{tr} \left\{ Q_0(t,x) \left( \frac{\partial (S_{\nu+1} - S_{\nu})}{\partial x} \right)^T \right\} \]
\[-\frac{1}{4} \left( \frac{\partial S_{\nu+1}}{\partial x} \right) N_0(t,x) \frac{\partial S_{\nu}}{\partial x} \right) + \left( \frac{\partial S_{\nu}}{\partial x} \right) N_0(t,x) \frac{\partial S_{\nu+1}}{\partial x} \]
\[-2 \left( \frac{\partial S_{\nu}}{\partial x} \right) N_0(t,x) \frac{\partial S_{\nu}}{\partial x} \quad (19)\]

Let \( \Sigma_0(t,x) := S_{\nu+1}(t,x) - S_{\nu}(t,x) \). Then, the above inequality is expressed as
\[-\frac{\partial \Sigma_0(t,x)}{\partial t} \leq \frac{1}{2} \text{tr} \left\{ Q_0(t,x) \left( \frac{\partial \Sigma_0(t,x)}{\partial x} \right)^T \right\} \]
\[-\frac{1}{4} \left( h^T(t,x) \frac{\partial \Sigma_0(t,x)}{\partial x} \right) + \left( \frac{\partial \Sigma_0(t,x)}{\partial x} \right)^T h(t,x) \]
\[\Sigma_0(T,x) = 0. \quad (19)\]
This is the same type inequality as (12). Hence, the same argument will be applicable to show that \( S_0(t, x) \leq 0 \) for all \((t, x)\). Therefore, we have
\[
S_{\nu+1}(t, x) \leq S_0(t, x)
\]  
which shows that the sequence \(\{S_\nu(t, x)\}_{\nu=1,2,\ldots} \) is monotonically decreasing and the existence of its limit \( S_\infty(t, x) \).

(iii) Proof of \( \lim_{\nu \to \infty} S_\nu(t, x) = S(t, x) \): Whenever \( \partial S_\nu/\partial t, \partial S_\nu/\partial x \) and \((\partial/\partial x)(\partial S_\nu/\partial x)^T \) are equicontinuous and equibounded, then \( \partial S_\nu/\partial t \to \partial S_\infty/\partial t, \partial S_\nu/\partial x \to \partial S_\infty/\partial x, (\partial/\partial x)(\partial S_\nu/\partial x)^T \to (\partial/\partial x)(\partial S_\infty/\partial x)^T \), respectively, via the Ascoli-Arzelà theorem; and hence, by letting \( \nu \to \infty \) in (7), we get
\[
-\frac{\partial S_\infty(t, x)}{\partial t} = L(t, x) + f^T(t, x) \frac{\partial S_\infty(t, x)}{\partial x}
+ \frac{1}{2} \text{tr}\left\{ Q_0(t, x) \frac{\partial}{\partial x} \left( \frac{\partial S_\infty(t, x)}{\partial x} \right)^T \right\}
- \frac{1}{4} \left( \frac{\partial S_\infty(t, x)}{\partial x} \right)^T N(t, x) \frac{\partial S_\infty(t, x)}{\partial x},
\]  
whose solution is expressed as
\[
\Sigma(\tau, x) = \frac{1}{4} \int_0^\tau \int_{\mathbb{R}^n} \Gamma_0(\tau, x; \sigma, \xi)
\left( \frac{\partial S(\sigma, \xi)}{\partial \xi} \right)^T N_0(\sigma, x) \frac{\partial S(\sigma, \xi)}{\partial \xi} d\xi d\sigma
\]  
in which \( \Gamma_0 \) is the fundamental solution to
\[
\frac{\partial \Sigma(\tau, x)}{\partial \tau} = \frac{1}{2} \text{tr}\left\{ Q_0(T-\tau, x) \frac{\partial}{\partial x} \left( \frac{\partial \Sigma(\tau, x)}{\partial x} \right)^T \right\}
+ f^T(T-\tau, x) \frac{\partial \Sigma(\tau, x)}{\partial x}.
\]
Noting in (23) that \( \Gamma_0 \geq 0 \) and the quadratic form \((\partial S/\partial x)^T N(t, x) (\partial S/\partial x) \geq 0 \) for all \((\tau, x)\). Therefore, we have \( \Sigma(t, x) \geq 0 \), that is \( S_0(t, x) \geq S(t, x) \).

Remark 2: As for the relation between the initial value \( S_0(t, x) \) and the first \( S_1(t, x) \), no one can say which one is smaller or larger.

5 Computation of Optimal Control

A. Computation of the Gradient \( \partial S(t, x)/\partial x \) To compute the optimal control given by (5), the gradient \( \partial S(t, x)/\partial x \) is required. Here, let us investigate how it can be computed. To do this, first let us consider the scalar case (i.e., \( n = \ell = d = 1 \)) for simplicity, and let both \( Q_0 \) and \( N_0 \) be constant.

Differentiate the HJB equation (6) with respect to \( x \), we have
\[
-\frac{\partial}{\partial x} \left( \frac{\partial S(t, x)}{\partial t} \right) = \frac{\partial L(t, x)}{\partial x} + \frac{\partial}{\partial x} \left( f(t, x) \frac{\partial S(t, x)}{\partial x} \right)
+ \frac{1}{2} Q_0 \frac{\partial}{\partial x} \left( \frac{\partial^2 S(t, x)}{\partial x^2} \right) - \frac{1}{4} N_0 \frac{\partial}{\partial x} \left( \frac{\partial S(t, x)}{\partial x} \right)^2.
\]
Noting that
\[
\frac{\partial}{\partial x} \left( \frac{\partial S(t, x)}{\partial x} \right) = \frac{\partial S(t, x)}{\partial x}, \quad \frac{\partial}{\partial x} \left( \frac{\partial^2 S(t, x)}{\partial x^2} \right) = \frac{\partial^2 S(t, x)}{\partial x^2},
\]  
and denoting \( \partial S(t, x)/\partial x = s(t, x) \), the above equation is rewritten as follows:
\[
-\frac{\partial s(t, x)}{\partial t} = \frac{1}{2} Q_0 \frac{\partial^2 s(t, x)}{\partial x^2} - \frac{1}{2} N_0 s(t, x) \frac{\partial s(t, x)}{\partial x}
+ f(t, x) s(t, x) + \frac{\partial L(t, x)}{\partial x},
\]
which is a Burgers-like or a Navier-Stokes-like equation. Here, consider the Cole-Hopf transformation
\[
s(t, x) = -Q_0 \frac{\partial}{\partial x} \ln \mu(t, x).
\]
Then, the L.H.S. of (25) is expressed as
\[
- \frac{\partial s(t, x)}{\partial t} = Q_0 \frac{\partial^2 s(t, x)}{\partial x^2} \ln \mu(t, x)
= Q_0 \frac{\partial}{\partial x} \ln \frac{\mu(t, x)}{\mu(t, x)}.
\]
where $\mu_t = \partial \mu / \partial t$. As for the R.H.S. of (25), note that it is expressed as

$$\frac{\partial}{\partial x} \left( \frac{1}{2} Q_0 \frac{\partial s}{\partial x} - \frac{1}{4} N_0 s^2 + f(t,x)s + L(t,x) \right)$$

and

$$s(t,x) = -Q_0 \frac{\mu_x(t,x)}{\mu(t,x)}$$

$$\frac{\partial s(t,x)}{\partial x} = -Q_0 \frac{\partial}{\partial x} \left( \frac{\mu_x}{\mu} \right)$$

$$= -Q_0 \frac{1}{\mu^2} (\mu_{xx} - \mu_x^2),$$

where $\mu_x$ and $\mu_{xx}$ are the first and second derivatives (gradient and Hessian) of $\mu$, respectively; so that

R.H.S. of (25) = $Q_0 \frac{\partial}{\partial x} \left[ \frac{1}{\mu} \left( \frac{1}{2} Q_0 \mu_{xx} + \frac{1}{2} Q_0 (1 - \frac{1}{2} N_0) \mu_x^2 - f(t,x) \mu_x + Q_0^{-1} L(t,x) \mu \right) \right].$

Assume that $\frac{1}{2} N_0 = 1$ under which the term of $\mu_x^2$ disappears. Then,

R.H.S. of (25) = $Q_0 \frac{\partial}{\partial x} \left[ \frac{1}{\mu} \left( \frac{1}{2} Q_0 \mu_{xx} - f(t,x) \mu_x + Q_0^{-1} L(t,x) \mu \right) \right].$

Consequently, by equating both sides, we get

$$\mu_t = -Q_0 \mu_{xx} - f(t,x) \mu_x + Q_0^{-1} L(t,x) \mu$$

or

$$\frac{\partial \mu(t,x)}{\partial t} = \frac{1}{2} Q_0 \frac{\partial^2 \mu(t,x)}{\partial x^2} - f(t,x) \frac{\partial \mu(t,x)}{\partial x}$$

$$+ Q_0^{-1} L(t,x) \mu(t,x).$$

(27)

The terminal condition is given from the relations

$$\frac{\partial S(T,x)}{\partial x} = s(T,x) = -Q_0 \frac{\partial}{\partial x} \ln \mu(t,x) \Big|_{t=T}$$

and $S(T,x) = \phi(x)$ as follows:

$$\mu(T,x) = \exp \{-Q_0^{-1} \phi(x)\}.$$  \hspace{1cm}  (28)

For the vector case, (27) is written under the assumptions $Q_0 = I_n$ and $\frac{1}{2} N_0 = I_n$ as follows:

$$\left\{ \begin{array}{l}
\frac{\partial \mu(t,x)}{\partial t} = -\frac{1}{2} \mu^T \left( \frac{\partial \mu(t,x)}{\partial x} \right) \\
- f^T(t,x) \frac{\partial \mu(t,x)}{\partial x} + L(t,x) \mu(t,x)
\end{array} \right.$$  \hspace{1cm}  (29)

$$\mu(T,x) = \exp \{-\phi(x)\}.$$

Thus the nonlinear equation (25) for $s(t,x)$ has been transformed to the linear equation (27) or (29) under proper conditions.

B. LQG Case

Here, it will be worth observing that the Riccati equation is derived from (29) in the LQG framework. To see this, let $f(t,x) = Ax, L(t,x) = x^T M x, \phi(x) = x^T F x$, and let $C, G$ be constant. In this case, it is well-known that the solution of the HJB equation (6) is given as $S(t,x) = x^T \Pi(t) x + \beta(t)$, where $\Pi(t)$ and $\beta(t)$ are the solutions to $Q_0^2[\delta^2]_2$.

$$\begin{align*}
\dot{\Pi}(t) + \Pi(t) A + A^T \Pi(t) + M \\
- \Pi(t) CN^{-1}C^T \Pi(t) = 0, \quad \Pi(T) = F
\end{align*}$$

(30)

$$\dot{\beta}(t) + tr \{GQG^T \Pi(t)\} = 0, \quad \beta(T) = 0.$$

From the definition of $s(t,x) = \partial S(t,x)/\partial x$ and (26) with $Q_0 = I_n$, we have $\partial S(t,x)/\partial x = -\partial \ln \mu(t,x)/\partial x$, or $\mu(t,x) = e^{-s(t,x)}$ which is just the (inverse) logarithmic Cole-Hopf transformation. Assume that

$$\mu(t,x) = e^{-[x^T \Pi(t) x + \beta(t)]}$$

(31)

Then, substituting this into (29), we see that $\dot{\Pi}(t)$, $\dot{\alpha}(t)$ and $\dot{\beta}(t)$ satisfy

$$\begin{align*}
\dot{\Pi}(t) + \Pi(t) A + A^T \Pi(t) + M \\
- 2\dot{\Pi}^2(t) = 0, \quad \dot{\Pi}(T) = F
\end{align*}$$

(32)

$$\dot{\alpha}(t) + [A^T - 2\Pi(t)] \dot{\alpha}(t) = 0, \quad \alpha(T) = 0$$

$$\dot{\beta}(t) - \frac{1}{2} \alpha^T(t) \dot{\alpha}(t) + tr \dot{\Pi}(t) = 0, \quad \beta(T) = 0.$$

Noting that the homogeneous linear equation for $\dot{\alpha}$ is subject to $\dot{\alpha}(0) = 0$ and that its solution is $\dot{\alpha}(t) = 0$ for all $t$. Therefore, (32) is just (30) under the assumptions $GQG^T = Q_0 = I_n$ and $\frac{1}{2} CN^{-1}C^T = \frac{1}{2} N_0 = I_n$.

C. Numerical Computation of Optimal Control

Thus, the nonlinear equation (25) has been converted to the linear partial differential equation (27) or (29) with (28). Its solution can be obtained à la stochastical representation using the Feynman-Kac formula $^6$. However, after having solved (26) numerically, the transformation (26) seems rather cumbersome for the numerical computations. So, instead, the convergent solution $S_{\nu}(t,x)$ obtained in Section 3 will be, rather, commendable to compute the gradient $\partial S(t,x)/\partial x$ to get the optimal control. Once we have obtained a sufficiently convergent $S_{\nu}(t,x)$ which can be evaluated for a given $\delta > 0$ by

$$\sup_{(t,x) \in [0,T] \times R^n} \left| \frac{S_{\nu}(t,x) - S_{\nu-1}(t,x)}{S_{\nu-1}(t,x)} \right| \leq \delta,$$

the gradient $\partial S(t,x)/\partial x$ can be approximated by using $S_{\nu}(t,x + \Delta x)$ and $S_{\nu}(t,x - \Delta x)$ computed at the spatial point $x = x(t)$ with a sufficiently small $\Delta x (> 0)$ as

$$\frac{\partial S(t,x)}{\partial x} \approx \frac{S_{\nu}(t,x + \Delta x) - S_{\nu}(t,x - \Delta x)}{2 \Delta x}.$$  \hspace{1cm}  (33)
if we employ the familiar central difference scheme. Hence, the optimal control \( u^*(t) \) is computed by

\[
u^*(t) = -\frac{1}{2}N^{-1}(t)C^T(t, x) \frac{\partial S(t, x)}{\partial x} \bigg|_{x=x(t)}.
\]

(34)

6 Deterministic Case

For the deterministic nonlinear control system, the solution to the HJB equation can be obtained mutatis mutandis by the similar way. Note that the HJB equation for the deterministic system follows by letting formally \( Q(t) \equiv 0 \) (or \( Q_0(t, x) \equiv 0 \)). In the deterministic case, the minimal cost functional is given (with no expectation operation) by

\[
S(t, x) = \min_{u(\tau)} \left[ \phi(x(T)) + \int_t^T \{ L[\tau, x(\tau)] + u(\tau)\} d\tau \right]
\]

under the process

\[
\dot{x}(t) = f[t, x(t)] + C[t, x(t)]u(t), \quad x(0) = x_0.
\]

(35)

The quasilinearization equation is given by

\[
\begin{cases}
\frac{\partial S_{\nu+1}(t, x)}{\partial t} = L(t, x) + f^T(t, x) \frac{\partial S_{\nu+1}(t, x)}{\partial x} \\
- \frac{1}{4} \left( \frac{\partial S_{\nu+1}(t, x)}{\partial x} \right)^T N_0(t, x) \frac{\partial S_{\nu+1}(t, x)}{\partial x} \\
+ \left( \frac{\partial S_{\nu}(t, x)}{\partial x} \right)^T N_0(t, x) \frac{\partial S_{\nu+1}(t, x)}{\partial x} \\
- \left( \frac{\partial S_{\nu+1}(t, x)}{\partial x} \right)^T N_0(t, x) \frac{\partial S_{\nu}(t, x)}{\partial x}
\end{cases}
\]

(36)

and Theorem still holds.

However, its proof should be carefully done. The equation (13) reduces in this case to the first-order partial differential equation for \( \Sigma(t, x) = S(t, x) - S_{\nu+1}(t, x) \),

\[
\frac{\partial \Sigma(t, x)}{\partial t} = \frac{1}{4} \left[ h^T(t, x) \frac{\partial \Sigma(t, x)}{\partial x} \right] + E(t, x), \quad \Sigma(T, x) = 0
\]

(37)

or in componentwise expression,

\[
\frac{\partial \Sigma(t, x)}{\partial t} - \sum_{i=1}^n \frac{1}{2} h_i(t, x) \frac{\partial \Sigma(t, x)}{\partial x_i} = E(t, x).
\]

(38)

For the first-order partial differential equation (38), the associated characteristic differential equations are given by

\[
\frac{dt}{1} = -\frac{1}{2} h_1(t, x), \cdots, \frac{dx_n}{1} = -\frac{1}{2} h_n(t, x) = \frac{d\Sigma}{E(t, x)}.
\]

(39)

The integrals of the equations \( dt = -dx_i/\frac{1}{2} h_i(t, x) \) are \( x_i = -\frac{1}{2} \int h_i(t, x) dt + \text{const} \). Substituting these into the last expression of (39), we have

\[
d\Sigma/E(t, x_1, \cdots, x_n) \text{ or } d\Sigma = E(t, x_1, \cdots, x_n) dt.
\]

Integrating this from \( t \) to \( T \) and noting \( \Sigma(T, x) \equiv 0 \) and \( E(\sigma, x) \geq 0 \) for all \( (\sigma, x) \in [0, T] \times \mathbb{R}^n \), we get

\[
\Sigma(t, x) = \int_t^T E(\sigma, x_1, \cdots, x_n) d\sigma \leq 0.
\]

(40)

Hence, \( S(t, x) \leq S_{\nu+1}(t, x) \) follows. Also, \( S_{\nu+1}(t, x) \leq S_{\nu}(t, x) \) and \( \lim_{\nu \to \infty} S_{\nu}(t, x) = S(t, x) \) are also shown.

As for the initial value \( S_0(t, x) \), we can employ the solution of (9) with \( Q_0(t, x) \equiv 0 \). Similarly, \( S(t, x) \leq S_0(t, x) \) holds.

7 Conclusion

A practicable algorithm for solving the HJB equation described by a nonlinear partial differential equation has been derived. The key to success is to have noticed the analogy between the nonlinear HJB equation and the Riccati differential equation which can be solved using the quasilinearization as shown in Ref. 1, §8.6).

References