Applied Stochastic Processes and Control for Jump-Diffusions: Modeling, Analysis and Computation

Floyd B. Hanson
University of Illinois
Chicago, Illinois, USA

Chapter 7 Stochastic Dynamic Programming

May 24, 2006

Chapter 7

Stochastic Optimal Control - Stochastic Dynamic Programming

It was the owl that shriek'd, the fatal bellman,
Which gives the stern'st good-night.
—William Shakespeare (1564-1616) in *Macbeth*.

But the principal failing occurred in the sailing,
And the Bellman, perplexed and distressed,
Said he had hoped, at least, when the wind blew due East,
That the ship would not travel due West!
—Lewis Carroll (1832-1898) in *The Bellman's Speech*.

7.1 Stochastic Optimal Control Problem

This main chapter introduces the optimal stochastic control problem. For many application systems, solving a SDE, or for that matter an ODE, to obtain its behavior is only part of the problem. The SDE is, in fact, a stochastic ordinary differential equation (SODE). Another, very significant part is finding out how to control the SDE or ODE as a model for controlling the application system.

Thus, the general jump-diffusion SDE (5.72) is reformulated with an additional process, the **vector control process** $\mathbf{U}(t) = [U_i(t)]_{n_u \times 1}$ on some n_u-dimensional control space $\mathcal{D}_u$,

$$d\mathbf{X}(t) \overset{\text{sym}}{=} \mathbf{f}(\mathbf{X}(t), \mathbf{U}(t), t)dt + g(\mathbf{X}(t), \mathbf{U}(t), t)d\mathbf{W}(t) \\ +h(\mathbf{X}(t), \mathbf{U}(t), t, \mathbf{Q})d\mathbf{P}(t; \mathbf{X}(t), \mathbf{U}(t), t) \,, \tag{7.1}$$

when $t_0 \le t \le t_f]$ subject to a given initial state $\mathbf{X}(t_0) = \mathbf{x}_0$, where again $\mathbf{X}(t) = [X_i(t)]_{n_x \times 1}$ is the **vector state process** on some n_x-dimensional state space $\mathcal{D}_x$. The stochastic processes are the n_w-dimensional vector Wiener process or diffusion process $\mathbf{W}(t) = [W_i(t)]_{n_w \times 1}$.

The symbolic notation for the n_p-dimensional vector state-dependent compound Poisson process or jump process can better be defined as in Chapt. 5 in two ways,

$$h(\mathbf{X}(t),\mathbf{U}(t),t,\mathbf{Q})d\mathbf{P}(t;\mathbf{X}(t),\mathbf{U}(t),t) \equiv \int_{\mathcal{Q}} h(\mathbf{X}(t),\mathbf{U}(t),t,\mathbf{q})\mathcal{P}(\mathbf{dt},\mathbf{dq})$$
$$= \left[\sum_{k=1}^{n_p}\sum_{k=1}^{dP_j(t)} h_{i,j}(\mathbf{X}(T_{j,k}^-),\mathbf{U}(T_{j,k}^-),T_{j,k}^-,Q_{j,k})\right]_{n_x\times 1},$$

with $d\mathbf{P}(t;\mathbf{X}(t),\mathbf{U}(t),t) = [dP_i(t;\mathbf{X}(t),\mathbf{U}(t),t)]_{n_p\times 1}$ and

$$\mathrm{E}[d\mathbf{P}(t;\mathbf{X}(t),\mathbf{U}(t),t)|\mathbf{X}(t)=\mathbf{x},\mathbf{U}(t)=\mathbf{u}] = \boldsymbol{\lambda}(t;\mathbf{x},\mathbf{u},t)dt,$$

where $\boldsymbol{\lambda}(t;\mathbf{x},\mathbf{u},t)$ is the jump rate vector and $T_{j,k}^-$ is the kth jump time of the jth differential Poisson process.

The coefficient functions are the $n_x \times 1$ plant function $\mathbf{f}(\mathbf{x},\mathbf{u},t)$, having the same dimension as the state $\mathbf{x}$, the $n_x \times n_w$ volatility function $g(\mathbf{x},\mathbf{u},t)$ or square root of the variance of the diffusion term, and the $n_x \times n_p$ jump amplitude of the jump term $h(\mathbf{x},\mathbf{u},t,\mathbf{Q})$, where $\mathbf{Q}$ is the underlying jump amplitude random mark process, the space part of the space-time Poisson process.

The optimization objective functional for a control formulation may be the combination of a final cost at time t_f and cumulative instantaneous costs, given the initial data $(\mathbf{x}_0,t_0)$. For instance,

$$V[\mathbf{X},\mathbf{U}](\mathbf{x}_0,t_0) = \int_{t_0}^{t_f} C(\mathbf{X}(s),\mathbf{U}(s),s)ds + S(\mathbf{X}(t_f),t_f) \tag{7.2}$$

is a functional of the processes $\mathbf{X}(t)$ and $\mathbf{U}(t)$, where $C(\mathbf{x},\mathbf{u},t)$ is the scalar instantaneous or **running cost function** on the **time horizon** $[t_0,t_f]$ and $S(\mathbf{x},t)$ is the **final cost function**; both are assumed continuous. This is the Bolza form of the objective. The objective $V[\mathbf{X},\mathbf{U}](\mathbf{x}_0,t_0)$ is a functional of the state $\mathbf{X}$ and control process $\mathbf{U}$, i.e., a function of functions, while also dependent on the values of the initial data $(\mathbf{x}_0,t_0)$. The optimal control objective, in this case, is to minimize the expected total costs with respect to the control process on $[t_0,t_f)$. The feedback control of the multibody stochastic dynamical system (7.1) is illustrated in the block diagram displayed in Figure 7.1.

Prior to the optimization step, an averaging step, taking the conditional expectation, conditioned on some initial state, is **essential** to avoid the ill-posed problem of trying to optimize an uncertain, fluctuating objective. It is further assumed here that the running and terminal cost functions permit a unique minimum, subject to stochastic differential dynamics in the multi-dimensional jump-diffusion case (7.1). Hence, the optimal, expected cost for (7.2) is

$$v^*(\mathbf{x}_0,t_0) \equiv \min_{\boldsymbol{U}[t_0,t_f)}\left[\mathop{\mathrm{E}}_{(\boldsymbol{dW},d\boldsymbol{P})[t_0,t_f)}\left[V[\mathbf{X},\mathbf{U}](\mathbf{x}_0,t_0)\right.\right.$$
$$\left.\left.|\mathbf{X}(t_0)=\mathbf{x}_0,\mathbf{U}(t_0)=\mathbf{u}_0\right]\right], \tag{7.3}$$

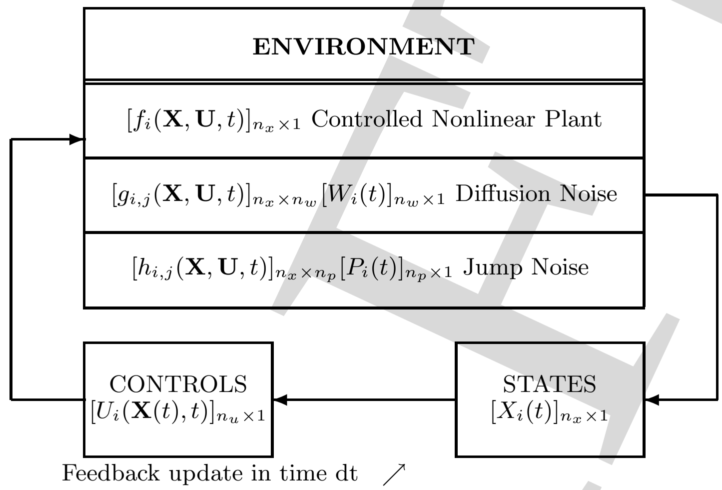

Figure 7.1. *Multibody Stochastic Dynamical System Under Feedback Control.*

with the expectation preceding the minimization so that the minimization problem is better-posed by smoothing random fluctuations through averaging. In the optimization in (7.5), it is implicit that the stochastic dynamical system (7.1) is a constraint. The minimization over $\mathbf{U}[t_0, t_f)$ denotes the minimization over the control path $\mathbf{U}(t)$ for $t \in [t_0, t_f)$ and similarly the expectation over $\{W, P\}[t_0, t_f)$ denotes expectation over the joint stochastic pair $\{W(t), P(t)\}$ for $t \in [t_0, t_f)$.

Recall that the maximum problem, as in the maximization of profits or portfolio returns, is an equivalent problem since

$$\max_{\mathbf{U}}[V[\mathbf{X},\mathbf{U}](x_0,t_0)] = -\min_{\mathbf{U}}[-V[\mathbf{X},\mathbf{U}](\mathbf{x}_0,t_0)]\,,$$

upon reversing the value.

In order to implement the dynamic part of dynamic programming, the fixed initial condition $\mathbf{X}(t_0) = \mathbf{x}_0$ needs to be replaced by a more arbitrary start, $\mathbf{X}(t) = \mathbf{x}$, so that the start can be analytically manipulated. This is a small but important step to produce a time-varying objective amenable to analysis. Hence, let the SDE dynamics be rewritten as

$$\begin{aligned} d\mathbf{X}(t) \overset{\text{sym}}{=} &\ \mathbf{f}(\mathbf{X}(t),\mathbf{U}(t),t)dt + g(\mathbf{X}(s),\mathbf{U}(t),t)d\mathbf{W}(t) \\ &+h(\mathbf{X}(t),\mathbf{U}(t),t,\mathbf{Q})d\mathbf{P}(t;\mathbf{X}(s),\mathbf{U}(t),t)\,, \end{aligned} \tag{7.4}$$

on $t_0 \le t \le t_f$ and the optimal expected value as

$$v^*(\mathbf{x},t) \equiv \min_{\boldsymbol{U}[t,t_f)}\left[\operatorname*{E}_{(\boldsymbol{dW},d\boldsymbol{P})[t,t_f)}[V[\mathbf{X},\mathbf{U}](\mathbf{x},t)|\,\mathbf{X}(t)=\mathbf{x}\,,\mathbf{U}(t)=\mathbf{u}]\right]. \tag{7.5}$$

Since the running cost integral vanishes when $t = t_f$, leaving only the terminal cost term conditioned on $\mathbf{X}(t_f) = \mathbf{x}$ and $\mathbf{U}(t_f) = \mathbf{u}$, a simple final condition for the optimal expected cost follows:

$$v^*(\mathbf{x}, t_f) = S(\mathbf{x}, t_f)\,, \tag{7.6}$$

for any $\mathbf{x}$ in the state domain $\mathcal{D}_x$, assuming that the terminal cost function $S(\mathbf{x}, t_f)$ is a deterministic function. This final condition is the first clue meaning that dynamic programming will use a backward program in time.

7.2 Bellman's Principle of Optimality

The basic assumption is that the optimization and expectation can be decomposed over increments in time. Bellman's Principle of Optimality can be systematically derived from optimization in time step proceeding backward from the final increment to the initial increment. Also, in the Markov processes case here, the independent increment properties of the Wiener and Poisson processes permit the decomposition of the expectation over time. This decomposition conveniently complements the decomposition of the optimization over time as in the deterministic case presented in Section 6.4.

The semi-close-opened time interval $[t, t_f)$ in the optimal expected cost formulation (7.5), given the state at time t, can be decomposed into disjoint increments $[t, t+\delta t)$ and $[t+\delta t, t_f)$ for fixed δt in $t < t+\delta t < t_f$. Symbolically, the **decomposition rules** are written:

Rules 7.1. ***Decomposition for Time, Integration, Expectation and Minimumization:***

- ***Time domain Decompostion into Subintervals:***

$$[t, t_f) = [t, t+\delta t) + [t+\delta t, t_f),$$

needs to be further decomposed for discrete approximations into sufficiently small increments Δt_i for $i = 0\!:\!n+1$, where $n \equiv n_1 + n_2 + 1$ such that

$$t_{i+1} = t_i + \sum_{j=1}^{i} \Delta t_j,$$

$t_0 = t$, $t_{n_1+1} = t + \delta t$ and $t_{n+1} = t_f$. For simplicity, Δt_i can be constant on the original pair of subintervals, i.e.,

$$\Delta t_i = \left\{ \begin{array}{ll} \overline{\Delta t}_1 = \delta t/(n_1+1), & i = 0 : n_1 \\ \overline{\Delta t}_{n+1} = (t_f - t - \delta t)/(n_2+1), & i = n_{1+1} : n \end{array} \right\}.$$

Recall that the approximation to the stochastic dynamics is (7.4)

$$\begin{aligned} \mathbf{X}_{i+1} &= \mathbf{X}_i + \int_{t_i}^{t_i+\Delta t_i} d\mathbf{X}(s) \\ &\simeq \mathbf{X}_i + \mathbf{f}_i \Delta t_i + g_i \Delta \mathbf{W}_i + h_i \Delta \mathbf{P}_i, \end{aligned}$$

where, for example, $\mathbf{f}_i \equiv \mathbf{f}(\mathbf{X}_i, \mathbf{U}_i, t_i)$, for sufficiently small Δt_i, so that the change from $\mathbf{X}_i$ to $\mathbf{X}_{i+1}$ is due to the control $\mathbf{U}_i$ and random fluctuations $(\Delta\mathbf{W}_i, \Delta\mathbf{P}_i)$ determined from a prior stage.

- ***Integration Additive Decomposition Rule:***

$$\int_t^{t_f} C(\mathbf{X}(s), \mathbf{U}(s), s)ds = \int_t^{t+\delta t} C(\mathbf{X}(s),\mathbf{U}(s), s)ds + \int_{t+\delta t}^{t_f} C(\mathbf{X}(s),\mathbf{U}(s), s)ds, \quad (7.7)$$

for the cumulative running costs by the regular additivity property of regular or Riemann-type integrals, or in terms of small increments in simplified notation. Let

$$V = \int_t^{t_f} C ds + S(\mathbf{X}(t_f), t_f) \simeq \sum_{i=0}^{n+1} \widehat{C}_i,$$

be the forward approximation, where $\widehat{C}_i = C_i \equiv C_i(\mathbf{X}_i, \mathbf{U}_i, t_i)$ for $i = 0 : n$ and $\widehat{C}_{n+1} \equiv S(\mathbf{X}(t_f), t_f)$.

- ***Expectation Operator Multiplication Decomposition Rule:***

$$\mathop{\mathrm{E}}_{(\boldsymbol{dW},\boldsymbol{dP})[t,t_f)}[V|\mathcal{C}(t)] = \mathop{\mathrm{E}}_{(\boldsymbol{dW},\boldsymbol{dP})[t,t+\delta t)}\left[\mathop{\mathrm{E}}_{(\boldsymbol{dW},\boldsymbol{dP})[t+\delta t,t_f)}[V|\mathcal{C}(t+\delta t)]\,|\mathcal{C}(t)\right],$$

where V is an objective function and $\mathcal{C}(t)$ is the conditioning at time t. This decomposition relies on the corresponding decomposition of the Markov processes $\mathbf{W}(t)$ and $\mathbf{P}(t)$ into independent increments, so that the expectation over $\{\mathbf{W}(s), \mathbf{P}(s)\}$ for $s \in [t, t_f)$ is the product of expectation over $\{\mathbf{W}(s), \mathbf{P}(s)\}$ for $s \in [t, t + \delta t)$ and expectation over $\{\mathbf{W}(r), \mathbf{P}(r)\}$ for $r \in [t + \delta t, t_f)$. In order to compute the expectation over the path of a Markov process, we need to approximate the process by a sum of n independent increments for sufficiently large n to obtain sufficiently small Δt_i and then take the product of the expectations with respect to each of these independent increments, and finally taking the limit as $n \to \infty$ relying on mean square convergence in the result as in the first two chapters. In simple notation,

$$\overline{V} = \mathrm{E}[V|(\mathbf{X}, \mathbf{U})(t) \simeq \mathrm{E}\left[\sum_{i=0}^{n+1} \widehat{C}_i \Delta t_i|(\mathbf{X}_0, \mathbf{U}_0)\right],$$

where $\mathrm{E}[\widehat{C}_0|(\mathbf{X}_0, \mathbf{U}_0)] = C_0$,

$$\mathrm{E}[\widehat{C}_1|(\mathbf{X}_0, \mathbf{U}_0)] = \mathop{\mathrm{E}}_{(\boldsymbol{\Delta W}_0,\boldsymbol{\Delta P}_0)}\left[\widehat{C}_1\middle|(\mathbf{X}_0, \mathbf{U}_0)\right] \equiv \mathrm{E}_0[\widehat{C}_1],$$

$$\begin{aligned}\mathrm{E}[\widehat{C}_2|(\mathbf{X}_0, \mathbf{U}_0)] &= \mathop{\mathrm{E}}_{(\boldsymbol{\Delta W}_0,\boldsymbol{\Delta P}_0)}\left[\mathop{\mathrm{E}}_{(\boldsymbol{\Delta W}_1,\boldsymbol{\Delta P}_1)}\left[\widehat{C}_2\middle|(\mathbf{X}_1, \mathbf{U}_1)\right]\middle|(\mathbf{X}_0, \mathbf{U}_0)\right] \\ &\equiv \Pi_{j=0}^{1}\mathrm{E}_j\left[\widehat{C}_2\right],\end{aligned}$$

so in general,

$$\mathrm{E}[\widehat{C}_i|(\mathbf{X}_0,\mathbf{U}_0)] = \Pi_{j=0}^{i-1}\mathrm{E}_j\left[\widehat{C}_i\right],$$

with

$$\mathrm{E}_j\left[\widehat{C}_i\right] \equiv \mathop{\mathrm{E}}_{(\boldsymbol{\Delta W}_j,\boldsymbol{\Delta P}_j)}\left[\widehat{C}_i\middle|(\mathbf{X}_j,\mathbf{U}_j)\right]$$

for $j = 0 : i-1$ *and finally,*

$$\begin{aligned}\overline{V} = \simeq \sum_{i=0}^{n+1}\Pi_{j=0}^{i-1}\mathrm{E}_j\left[\widehat{C}_i\Delta t_i\right] &= \sum_{i=0}^{n_1}\Pi_{j=0}^{i-1}\mathrm{E}_j\left[\widehat{C}_i\Delta t_i\right] + \sum_{i=n_1+1}^{n+1}\Pi_{j=0}^{i-1}\mathrm{E}_j\left[\widehat{C}_i\Delta t_i\right]\\ \to \mathop{\mathrm{E}}_{(\boldsymbol{dW},\boldsymbol{dP})[t,t+\delta t)}&\left[\int_t^{t+dt} C ds\right.\\ &+\left.\mathop{\mathrm{E}}_{(\boldsymbol{dW},\boldsymbol{dP})[t+\delta t,t_f)}\left[\int_{t+dt}^{t_f} C ds + S(\mathbf{X}(t_f),t_f)\middle|(\mathbf{X},\mathbf{U})(t+\delta t)\right]\middle|(\mathbf{X},\mathbf{U})(t)\right],\end{aligned}$$

as $n \to \infty$ $(n_1 \to \infty, n_2 \to \infty)$ *confirming the construction, assuming mean square convergence.*

- ***Minimization Operator Multiplication Decomposition Rule:***

$$\overline{V}^* = \min_{\boldsymbol{U}[t,t_f)}\left[\ \overline{V}\ \right] = \min_{\boldsymbol{U}[t,t+\delta t)}\left[\min_{\boldsymbol{U}[t+\delta t,t_f)}\left[\ \overline{V}\ \right]\right], \tag{7.8}$$

where $\overline{V}$ *is the expected value of an objective so that the decomposition rule is analogous to the use in deterministic dynamic programming. This decomposition depends on the reasonable heuristic idea that given a minimum on the later interval* $[t+\delta t, t_f)$*, taking the minimum of the given minimum over the small earlier interval* $[t, t+\delta t)$ *yields the minimum over the longer interval* $[t, t_f)$*. In terms of the small increments* (Δt_i) *construction,*

$$\begin{aligned}\overline{V}^* &\simeq \sum_{i=0}^{n+1}\min_{\boldsymbol{U}[t,t_f)}\Pi_{j=0}^{i-1}\mathrm{E}_j\left[\widehat{C}_i\Delta t_i\right] = \sum_{i=0}^{n+1}\Pi_{j=0}^{i-1}\min_{\boldsymbol{U}_j}\mathrm{E}_j\left[\widehat{C}_i\Delta t_i\right]\\ &= \sum_{i=0}^{n+1}\Pi_{j=0}^{i-1}\mathrm{ME}_j\left[\widehat{C}_i\Delta t_i\right]\end{aligned}$$

where

$$\mathrm{ME}_0 \equiv \min_{\boldsymbol{U}_0}\left[\mathop{\mathrm{E}}_{(\boldsymbol{\Delta W}_0,\boldsymbol{\Delta P}_0)}\left[\widehat{C}_0\Delta t_0\middle|(\mathbf{X}_0,\mathbf{U}_0)\right]\right]$$

and

$$\mathrm{ME}_j \equiv \min_{\boldsymbol{U}_j}\left[\mathop{\mathrm{E}}_{(\boldsymbol{\Delta W}_j,\boldsymbol{\Delta P}_j)}\left[\widehat{C}_i\Delta t_i\middle|(\mathbf{X}_j,\mathbf{U}_j)\right]\right]$$

for $j = 0 : i-1$. As $n \to \infty$ with $n = n_1 + n_2 + 1$ splitting,

$$\overline{V}^* \to \min_{\boldsymbol{U}[t,t+\delta t)} \Bigg[\mathop{\mathrm{E}}_{(\boldsymbol{dW},\boldsymbol{dP})[t,t+\delta t)} \Bigg[\int_t^{t+dt} C ds + \min_{\boldsymbol{U}[t+\delta t,t_f)} \Bigg[\mathop{\mathrm{E}}_{(\boldsymbol{dW},\boldsymbol{dP})[t+\delta t,t_f)} \Bigg[\int_{t+dt}^{t_f} C ds + S(\mathbf{X}(t_f),t_f) \Bigg| (\mathbf{X},\mathbf{U})(t+\delta t) \Bigg]\Bigg] \Bigg| (\mathbf{X},\mathbf{U})(t) \Bigg]\Bigg].$$

The optimal decomposition seems to work for many examples. However, for empirical counterexamples, see Rust [236].

Thus, **optimal expected cost** (7.5) can be decomposed as follows:

$$\begin{aligned} v^*(\mathbf{x},t) &= \min_{\boldsymbol{U}[t,t+\delta t)} \Bigg[\mathop{\mathrm{E}}_{(\boldsymbol{W},\boldsymbol{P})[t,t+\Delta t)} \Bigg[\int_t^{t+\delta t} C(\mathbf{X}(s),\mathbf{U}(s),s) ds \\ &\quad + \min_{\boldsymbol{U}(t+\delta t,t_f]} \Bigg[\mathop{\mathrm{E}}_{(\boldsymbol{W},\boldsymbol{P})[t+\Delta t,t_f)} \Bigg[\int_{t+\delta t}^{t_f} C(\mathbf{X}(s),\mathbf{U}(s),s) ds + S(\mathbf{X}(t_f),t_f) \\ &\quad \Bigg| \{\mathbf{X}(t+\delta t),\mathbf{U}(t+\delta t)\} \Bigg]\Bigg] \Bigg| \mathbf{X}(t)=\mathbf{x}\,,\mathbf{U}(t)=\mathbf{u} \Bigg]\Bigg] \\ &= \min_{\boldsymbol{U}[t,t+\delta t)} \Bigg[\mathop{\mathrm{E}}_{(\boldsymbol{W},\boldsymbol{P})[t,t+\Delta t)} \Bigg[\int_t^{t+\delta t} C(\mathbf{X}(s),\mathbf{U}(s),s) ds \\ &\quad + v^*(\mathbf{X}(t+\delta t),t+\delta t) \Bigg| \mathbf{X}(t)=\mathbf{x}\,,\mathbf{U}(t)=\mathbf{u} \Bigg]\Bigg], \end{aligned} \tag{7.9}$$

where the definition (7.5) for v^* has been reused with the arguments shifted by the time-step δt, since the inner part of the decomposition that is on $(t+\delta t, t_f]$ is precisely the definition of v^* in (7.5) but with arguments shifted from $(\mathbf{x},t)$ to $(\mathbf{X}(t+dt), t+dt)$. Thus, Eq. (7.9) is a backward recursion relation for v^*. The subscript notation $\mathbf{U}[t,t+\delta t)$ under the min operator means that the minimum is with respect to $\mathbf{U}$ in the range $[t,t+\delta t)$, with similar subscript notation $\{\mathbf{W},\mathbf{P}\}[t,t_f)$ for the expectation operator. Thus, we have formally shown:

Lemma 7.2. ***Bellman's Principle of Optimality:***
Under the assumptions of the decomposition rules (7.8, 7.8, 7.7) and the properties of jump-diffusions,

$$\begin{aligned} v^*(\mathbf{x},t) = \min_{\boldsymbol{U}[t,t+\delta t)} \Bigg[\mathop{\mathrm{E}}_{(\boldsymbol{W},\boldsymbol{P})[t,t+\Delta t)} \Bigg[\int_t^{t+\delta t} C(\mathbf{X}(s),\mathbf{U}(s),s) ds \\ + v^*(\mathbf{X}(t+\delta t),t+\delta t) \Bigg| \mathbf{X}(t)=\mathbf{x}\,,\mathbf{U}(t)=\mathbf{u} \Bigg]\Bigg]. \end{aligned} \tag{7.10}$$

The argument of the minimum when it exists, within the control domain $\mathcal{D}_u$, is the optimal control $\mathbf{u}^* = \mathbf{u}^*(\mathbf{x},t)$. Although the SDE is a forward differential equation integrated forward from the initial condition, the optimal control problem is a backward general or functional equation integrated backward from the final time. The backward equation is quite basic, when one has a final objective, here optimal costs. Then the primary question is where to start initially to get that optimum. People do backward calculations all the time, such as when going to a scheduled meeting or a class, the meeting time is fixed and the problem is to estimate what time one should leave to get to the meeting. However, when economic decisions are made, the decision makers may not behave according to Bellman's principle of optimality according to the studies of Rust [236].

In general, **capital letters** are used for stochastic processes and **lower case letters** for conditioned or realized variables.

7.3 Hamilton-Jacobi-Bellman (HJB) Equation of Stochastic Dynamic Programming

Using the **Principle of Optimality** (7.10) and by taking the limit of small δt, replacing δt by dt, we can systematically derive the partial differential equation of stochastic dynamic programming, also called the stochastic **Hamilton-Jacobi-Bellman** (HJB) equation, for the general, multi-dimensional Markov dynamics case. From the increment form of the state differential $d\mathbf{X}(t) = \mathbf{X}(t+dt) - \mathbf{X}(t)$, we consider the expansion of the state argument

$$\mathbf{X}(t+dt) = \mathbf{X}(t) + d\mathbf{X}(t)$$

about $\mathbf{X}(t)$ for small $d\mathbf{X}(t)$ and about the explicit time argument $t+dt$ about t in the limit of small time increments dt, using an extension of Taylor approximations extended to include discontinuous (i.e, Poisson) and non-smooth (i.e., Wiener) processes. Sufficient differentiability of the optimal value function $v^*(\mathbf{x}, t)$, at least to first order in time and second order in state, is assumed except when its state argument has Poisson jumps. The spirit of the derivation of the multi-dimensional chain rule (5.88) is applied to the Principle of Optimality (7.10), except that the mean square limit substitution for the bilinear Wiener $W_i(t)W_j(t)$ process is not needed here because of the pre-optimization expectation operation. Then neglecting $\text{o}(dt)$ terms as $dt \to 0^+$ (strictly, we are really working with finite increments δt) and substituting for the conditioning on $\mathbf{X}(t)$ and $\mathbf{U}(t)$, an intermediate reduction of the optimal expected value is

$$\begin{aligned} v^*(\mathbf{x},t) \overset{\text{dt}}{=} \min_{\mathbf{u}} \Bigg[\mathop{\mathrm{E}}_{(\boldsymbol{dW},\boldsymbol{dP})(t)} \Bigg[& C(\mathbf{x},\mathbf{u},t)dt + v^*(\mathbf{x},t) + v_t^*(\mathbf{x},t)dt \\ & + \nabla_{\mathbf{x}}^\top[v^*](\mathbf{x},t) \cdot (\mathbf{f}(\mathbf{x},\mathbf{u},t)dt + g(\mathbf{x},\mathbf{u},t)d\mathbf{W}(t)) \\ & + \frac{1}{2} d\mathbf{W}^\top(t) g^\top(\mathbf{x},\mathbf{u},t) \nabla_{\mathbf{x}}[\nabla_{\mathbf{x}}^\top[v^*]](\mathbf{x},t)(g(\mathbf{x},\mathbf{u},t)d\mathbf{W}(t)) \\ & + \sum_{j=1}^{n_p} \int_{\mathcal{Q}} \left(v^*(\mathbf{x} + \widehat{\boldsymbol{h}}_j(\mathbf{x},\mathbf{u},t,q_j),t) - v^*(\mathbf{x},t) \right) \mathcal{P}_j(dt, dq_j; \mathbf{x},\mathbf{u},t) \Bigg] \Bigg], \end{aligned} \tag{7.11}$$

where it has been assumed that the random mark variables $Q_j = q_j$ are pair-wise independently distributed and the jump amplitude is separable in the marks. So

$$h(\mathbf{x}, \mathbf{u}, t, \mathbf{q}) = [h_{i,j}(\mathbf{x}, \mathbf{u}, t, q_j)]_{n_x \times n_p} , \tag{7.12}$$

with a corresponding multiplicative factoring of the Poisson random measure. Recall from Chapter 5 (5.90) that the jth vector component of the jump amplitude is

$$\widehat{\boldsymbol{h}}_j(\mathbf{x}, \mathbf{u}, t, q_j) = [h_{i,j}(\mathbf{x}, \mathbf{u}, t, q_j)]_{n_x \times 1} , \tag{7.13}$$

for $j = 1 : n_p$, corresponding to the jth Poisson process

$$dP_j(t; \mathbf{x}, \mathbf{u}, t) = \int_{\mathcal{Q}} \mathcal{P}_j(dt, dq_j; \mathbf{x}, \mathbf{u}, t) ,$$

in terms of the jth Poisson mark-time random measure $\mathcal{P}_j$. Note that the first t argument of dP_j is the time implicit to the Poisson process, while the second t argument is an explicit time corresponding to the implicit state and control parametric dependence.

The next step is to take the conditional expectation over the now isolated differential Wiener and Poisson processes, but done by expanding them in components to facilitate understanding of the step and suppressing some arguments for simplicity,

$$\begin{aligned}
v^*(\mathbf{x}, t) \overset{\text{dt}}{=} {}& v^*(\mathbf{x}, t) + v^*_t(\mathbf{x}, t)dt + \min_{\boldsymbol{u}} \Big[C(\mathbf{x}, \mathbf{u}, t)dt \\
&+ \nabla^\top_{\mathbf{x}}[v^*](\mathbf{x}, t) \cdot \left(\mathbf{f}(\mathbf{x}, \mathbf{u}, t)dt + \sum_{i=1}^{n_w} g_i(\mathbf{x}, \mathbf{u}, t)\mathrm{E}_{dW_i}[dW_i(t)]\right) \\
&+ \frac{1}{2}\sum_{i=1}^{n_w}\sum_{j=1}^{n_w} \mathrm{E}_{dW_i, dW_j}\left[dW_i(t)dW_j(t)\right]\left[g^\top(\mathbf{x}, \mathbf{u}, t)\nabla_{\mathbf{x}}[\nabla^\top_{\mathbf{x}}[v^*]]g(\mathbf{x}, \mathbf{u}, t)\right]_{i,j} \\
&+ \sum_{j=1}^{n_p} \int_{\mathcal{Q}} \left(v^*(\mathbf{x} + \widehat{\boldsymbol{h}}_j(\mathbf{x}, \mathbf{u}, t, q_j), t) - v^*(\mathbf{x}, t)\right)\mathrm{E}_{\mathcal{P}_j}\left[\mathcal{P}_j(dt, dq_j; \mathbf{x}, \mathbf{u}, t)\right]\Big] \\
\overset{ind}{\underset{inc}{=}} {}& v^*(\mathbf{x}, t) + v^*_t(\mathbf{x}, t)dt + \min_{\boldsymbol{u}} \Bigg[C(\mathbf{x}, \mathbf{u}, t)dt + \nabla^\top_{\mathbf{x}}[v^*](\mathbf{x}, t)\left(\mathbf{f}(\mathbf{x}, \mathbf{u}, t)dt + 0\right) \\
&+ \frac{1}{2}\sum_{i=1}^{n_w}\left[1 + \sum_{j=1}^{n_w}\rho_{i,j}(1 - \delta_{i,j})\right]\left[g^\top(\mathbf{x}, \mathbf{u}, t)\nabla_{\mathbf{x}}\left[\nabla^\top_{\mathbf{x}}[v^*]\right](\mathbf{x}, t)g(\mathbf{x}, \mathbf{u}, t)\right]_{i,j} dt \\
&+ \sum_{j=1}^{n_p} \lambda_j(t; \mathbf{x}, \mathbf{u}, t) \int_{\mathcal{Q}} \left(v^*(\mathbf{x} + \widehat{\boldsymbol{h}}_j(\mathbf{x}, \mathbf{u}, t, q_j), t) - v^*(\mathbf{x}, t)\right) \\
&\qquad \cdot \Phi_{Q_j}(dq_j; \mathbf{x}, \mathbf{u}, t)dt\Bigg] ,
\end{aligned} \tag{7.14}$$

where we have used the expectations

$$\mathrm{E}[dW_i(t)] = 0, \quad \mathrm{E}[dW_i(t)dW_i(t)] = \delta_{i,j} + \rho_{i,j}(1 - \delta_{i,j})$$

with correlation coefficient $\rho_{i,j}$ and

$$\begin{aligned}\mathrm{E}[\mathcal{P}_j(dt, dq_j; \mathbf{x}, \mathbf{u}, t)] &= \lambda_j(t; \mathbf{x}, \mathbf{u}, t) dt \Phi_{Q_j}(dq_j; \mathbf{x}, \mathbf{u}, t) \\ &= \lambda_j(t; \mathbf{x}, \mathbf{u}, t) \phi_{Q_j}(q_j; \mathbf{x}, \mathbf{u}, t) dq_j dt\,.\end{aligned}$$

Also, with sufficiently small dt, $\mathbf{U}(t, t+dt]$ has been replaced by the conditioned control vector $\mathbf{u}$ at t.

Note that the $v^*(\mathbf{x}, t)$ value on both sides of the equation cancel and then the remaining common multiplicative factors of dt also cancel, so the **HJB equation** has been derived for this general case:

Theorem 7.3. ***Hamilton-Jacobi-Bellman Equation (HJBE) for Stochastic Dynamic Programming (SDP)***
If $v^*(\mathbf{x}, t)$ is twice differentiable in $\mathbf{x}$ and once differentiable in t, while the operator decomposition rules (7.8-7.7) are valid, then

$$0 = v_t^*(\mathbf{x}, t) + \min_{\mathbf{u}} \left[\widetilde{\mathcal{H}}(\mathbf{x}, \mathbf{u}, t) \right] \equiv v_t^*(\mathbf{x}, t) + \widetilde{\mathcal{H}}^*(\mathbf{x}, t) \tag{7.15}$$

where the ***Hamiltonian*** *(technically, a pseudo-Hamiltonian) functional is given by*

$$\begin{aligned}\widetilde{\mathcal{H}}(\mathbf{x}, \mathbf{u}, t) \equiv\ & C(\mathbf{x}, \mathbf{u}, t) + \nabla_{\mathbf{x}}^{\top}[v^*](\mathbf{x}, t) \cdot \mathbf{f}(\mathbf{x}, \mathbf{u}, t) \\ & + \frac{1}{2} \left(g R' g^{\top} \right)(\mathbf{x}, \mathbf{u}, t) \,:\, (\mathbf{x}, t) \\ & + \sum_{j=1}^{n_p} \lambda_j(t; \mathbf{x}, \mathbf{u}, t) \int_{\mathcal{Q}} \left[v^* \left(\mathbf{x} + \widehat{\boldsymbol{h}}_j(\mathbf{x}, \mathbf{u}, t, q_j), t \right) - v^*(\mathbf{x}, t) \right] \\ & \quad \cdot \phi_{Q_j}(q_j; \mathbf{x}, \mathbf{u}, t) dq_j\,,\end{aligned} \tag{7.16}$$

where the correlation modified indentity R' is defined in (5.85) as

$$R' \equiv \left[\delta_{i,j} + \rho_{i,j}(1 - \delta_{i,j}) \right]_{n_w \times n_w}\,, \tag{7.17}$$

and where the correlation coefficient between i and j components is

$$\rho_{i,j} dt = \mathrm{Cov}[dW_i(t), dW_j(t)]\,, \tag{7.18}$$

provided $j \neq i$ for $i, j = 1 : n_x$. The double-dot product $A : B$ is defined in (5.89).

The optimal control, if it exists, is given by

$$\mathbf{u}^* = \mathbf{u}^*(\mathbf{x}, t) = \operatorname*{argmin}_{\mathbf{u}} \left[\widetilde{\mathcal{H}}(\mathbf{x}, \mathbf{u}, t) \right]\,, \tag{7.19}$$

subject to any control constraints.

This **HJB equation** (7.15) is no ordinary PDE, but but has the following properties or attributes:

Properties 7.4.

- *The HJBE is a functional PDE or PIDE due to the presence of the minimum operator* min *and the Poisson integral term (the last term) with steps in the state argument of the optimal value function* v^* *due to the jump amplitude.*

- *The HJBE is a scalar valued equation, but has a* $(nu+1)$***-dimensional solution*** *consisting of the scalar optimal value function* $v^* = v^*(\mathbf{x}, t)$ *and the optimal control vector* $\mathbf{u}^* = \mathbf{u}^*(\mathbf{x}, t)$ *as well. These* ***dual solutions*** *are generally tightly coupled in functional dependence. In general, this tight coupling requires a number of iterations between* v^* *and* $\mathbf{u}^*$ *to obtain a reasonable approximation to the* $(nu+1)$*-dimensional solution. However, it should be noted that the optimal control* $\mathbf{u}(\mathbf{x}, t)$ *in (7.19) is deterministic and if the* $\mathbf{x}$ *dependence is genuine then it is also feedback optimal control. In fact, the HJB equation is a deterministic equation as well.*

- *A further complication in this functional PDE or PIDE is that the HJB equation (7.15) has* ***global state dependence*** *due to the Poisson jump functional integral term, whereas the HJB equation for purely Gaussian or Wiener processes is essentially a diffusion equation that has only* ***local state dependence*** *since it depends only on the values* $v^*(\mathbf{x}, t)$*,* $\mathbf{u}^*(\mathbf{x}, t)$*, the gradient vector* $\nabla_{\mathbf{x}}[v^*](\mathbf{x}, t)$*, and the Hessian matrix of 2nd order derivatives* $\nabla_{\mathbf{x}}[\nabla_{\mathbf{x}}^\top[v^*]](\mathbf{x}, t)$ *at* $(\mathbf{x}, t)$*. Contrast this with the random noise case including the Poisson random measure disturbance, with local dependence at* $\mathbf{x}$*, but global dependence on a range of points at* $\mathbf{x} + \widehat{\boldsymbol{h}}_j(\mathbf{x}, \mathbf{u}, t, q_j)$ *depending on the Poisson mark distribution.*

While letting $C^*(\mathbf{x}, t) \equiv C(\mathbf{x}, \mathbf{u}^*, t)$, $\mathbf{f}^*(\mathbf{x}, t) \equiv \mathbf{f}(\mathbf{x}, \mathbf{u}^*, t)$, $g^*(\mathbf{x}, t) \equiv g(\mathbf{x}, \mathbf{u}^*, t)$, $\widehat{\boldsymbol{h}}_j^*(\mathbf{x}, t, q_j) \equiv \widehat{\boldsymbol{h}}_j(\mathbf{x}, \mathbf{u}^*, t, q_j)$, and so forth for all control-dependent functions, then the HJB equation (HJBE) takes the form of a backward parabolic partial differential equation except that it has an additional integral term:

$$\begin{aligned} 0 &= v_t^*(\mathbf{x}, t) + \widetilde{\mathcal{H}}(\mathbf{x}, \mathbf{u}^*(\mathbf{x}, t), t) \\ &= v_t^*(\mathbf{x}, t) + C^*(\mathbf{x}, t) + \nabla_{\mathbf{x}}^\top[v^*](\mathbf{x}, t) \cdot \mathbf{f}^*(\mathbf{x}, t) \\ &\quad + \frac{1}{2}\left(g^* R' g^{*\top}\right)(\mathbf{x}, t) : \nabla_{\mathbf{x}}\left[\nabla_{\mathbf{x}}^\top[v^*]\right](\mathbf{x}, t) \\ &\quad + \sum_{j=1}^{n_p} \lambda_j^*(t; \mathbf{x}, t) \int_{\mathcal{Q}} \Delta_j[v^*](\mathbf{x}, t, q_j) \phi_{Q_j}^*(q_j; \mathbf{x}, t) dq_j \,, \end{aligned} \tag{7.20}$$

where the jth jump increment is defined as

$$\Delta_j[v^*](\mathbf{x}, t, q_j) \equiv v^*\left(\mathbf{x} + \widehat{\boldsymbol{h}}_j^*(\mathbf{x}, t, q_j), t\right) - v^*(\mathbf{x}, t) \tag{7.21}$$

and the double-dot product (:) is defined in (5.89). The final condition is given by $v^*(\mathbf{x}, t_f) = S(\mathbf{x}, t_f)$.

The Hamilton-Jacobi-Bellman name of the equation comes from the fact that Bellman [25, 26] was the founding developer of dynamic programming and the fact that the general evolution equation, $v_t^*(\mathbf{x},t) + \widetilde{\mathcal{H}}^*(\mathbf{x},t) = 0$, is called a Hamilton-Jacobi equation and where $\widetilde{\mathcal{H}}(\mathbf{x},\mathbf{u},t)$ is like a classical Hamiltonian. Sometimes, the HJB equation (7.15) is called simply the Bellman equation, or the stochastic dynamic programming equation or the PDE of stochastic dynamic programming, or in particular, the PIDE of stochastic dynamic programming where PIDE denotes a partial integral differential equation).

7.4 Linear Quadratic Jump-Diffusion (LQJD) Problem

The linear quadratic jump-diffusion (LQJD) problem is also called a linear quadratic Gaussian Poisson (LQGP) problem or jump linear quadratic Gaussian (JLQG) problem. The Markov property of the jump-diffusion processes described in this book leads to an analogous dynamic programming formulation to dynamic programming for deterministic processes as in the deterministic linear quadratic (LQ) problem of Subsection 6.4.4. In this chapter, the LQJD problem is presented in more generality than in Chapter 6.

The linear quadratic problem in both state and control leads to a quadratic decomposition of the optimal value function with respect to the state and a linear or feedback decomposition of the optimal control. However, first the LQJD problem is examined for a special case that is linear quadratic in control only to show how much an advantage is gained by the control dependence alone. For many applications it is not appropriate to have the problem linear quadratic in the state.

7.4.1 LQJD in Control Only (LQJD/U) Problem

A general variant of the LQJD problem is the LQJD/U problem that is LQJD in control only. Just having a control problem linear quadratic in control retains an important feature of the full linear quadratic control problem in that the optimal control can be solved for exactly in terms of the optimal value, even though the state decomposition property does not follow. The restricted linear quadratic problem in the control only will be treated first to examine how far the analysis can be taken before treating the full linear quadratic problem in the state and the control. In many control problems, the state dependence of the plant function $\mathbf{f}(\mathbf{x},\mathbf{u},t)$ is dictated by the application and may be significantly nonlinear, but the control dependence of the dynamics is up to the control designer who might chose to make the control simple, e.g., linear, so that the control process will be manageable for the control manager. Hence, the LQ problem in control only, may be more appropriate for some applications. In the past, linear systems were preferred since linear methods were well-known, but now nonlinear methods and problems have become more prevalent as we try to make more realistic models for applications.

Let the jump-diffusion linear quadratic model, in the control only, be given with the plant function for the deterministic or non-noise dynamics term,

$$\mathbf{f}(\mathbf{x},\mathbf{u},t) = \mathbf{f}_0(\mathbf{x},t) + f_1(\mathbf{x},t)\mathbf{u}\,, \tag{7.22}$$

with the diffusion term,

$$g(\mathbf{x}, \mathbf{u}, t) = g_0(\mathbf{x}, t) \,, \tag{7.23}$$

assumed control-independent for simplicity, with a jump term decomposition corresponding to independent sources of n_p-type jumps

$$h(\mathbf{x}, \mathbf{u}, t, \mathbf{q}) = h_0(\mathbf{x}, t, \mathbf{q}) = [h_{0,i,j}(\mathbf{x}, t, q_j)]_{n_x \times n_p} \,, \tag{7.24}$$

also assumed control-independent along with the very simplified Poisson noise

$$d\mathbf{P}(t; \mathbf{x}, \mathbf{u}, t) = d\mathbf{P}(t; \mathbf{x}, t) \,, \ \mathrm{E}[d\mathbf{P}(t; \mathbf{x}, t)] = \boldsymbol{\lambda}(t; \mathbf{x}, t)dt \,, \tag{7.25}$$

and finally with the quadratic running cost function,

$$C(\mathbf{x}, \mathbf{u}, t) = C_0(\mathbf{x}, t) + \mathbf{C}_1^\top(\mathbf{x}, t)\mathbf{u} + \frac{1}{2}\mathbf{u}^\top C_2(\mathbf{x}, t)\mathbf{u} \,. \tag{7.26}$$

It is assumed that all right hand side coefficients are commensurate in multiplication and that the product is the same type at that on the left hand side. A crucial assumption in case of a minimum objective is that the quadratic control $C_2(\mathbf{x}, t)$ is positive definite, but $C_2(\mathbf{x}, t)$ can be assumed to be symmetric without loss of generality by the symmetric property of quadratic forms (0.133).

Thus, the pseudo-Hamiltonian is quadratic in the control,

$$\widetilde{\mathcal{H}}(\mathbf{x}, \mathbf{u}, t) = \widetilde{\mathcal{H}}_0(\mathbf{x}, t) + \widetilde{\mathcal{H}}_1^\top(\mathbf{x}, t)\mathbf{u} + \frac{1}{2}\mathbf{u}^\top \widetilde{\mathcal{H}}_2(\mathbf{x}, t)\mathbf{u} \,, \tag{7.27}$$

where the scalar coefficient is

$$\begin{aligned} \widetilde{\mathcal{H}}_0(\mathbf{x}, t) = & \left[C_0 + \mathbf{f}_0^\top \nabla_{\mathbf{x}}[v^*] + \frac{1}{2} g_0 g_0^\top : \nabla_{\mathbf{x}}[\nabla_{\mathbf{x}}[v^*]] \right](\mathbf{x}, t) \\ & + \sum_{j=1}^{n_p} \lambda_j(t; \mathbf{x}, t) \int_{\mathcal{Q}_j} \Delta_j[v^*](\mathbf{x}, t, q_j)\phi_{\mathcal{Q}_j}(q_j)dq_j \,, \end{aligned} \tag{7.28}$$

where the double dot product is $GG^\top : A = \mathrm{Trace}[G^\top AG]$, while the jump increment is

$$\Delta_j[v^*](\mathbf{x}, t, q_j) \equiv v^* \left(\mathbf{x} + \widehat{\boldsymbol{h}}_j(\mathbf{x}, t, q_j), t \right) - v^*(\mathbf{x}, t) \,,$$

the linear control coefficient n_u-vector is

$$\widetilde{\mathcal{H}}_1(\mathbf{x}, t) = \left[\mathbf{C}_1 + f_1^\top \nabla_{\mathbf{x}}[v^*] \right] (\mathbf{x}, t) \,, \tag{7.29}$$

and the quadratic control coefficient $n_u \times n_u$-matrix is simply

$$\widetilde{\mathcal{H}}_2(\mathbf{x}, t) = C_2(\mathbf{x}, t) \,, \tag{7.30}$$

where $\widetilde{\mathcal{H}}_2(\mathbf{x}, t)$ is assumed to be symmetric along with $C_2(\mathbf{x}, t)$. If the minimum cost is the objective, then $\widetilde{\mathcal{H}}_2(\mathbf{x}, t)$ is positive definite since $C_2(\mathbf{x}, t)$ is assumed to be positive definite.

Thus, in search of a regular control minimum, the critical points of the pseudo-Hamiltonian $\widetilde{\mathcal{H}}(\mathbf{x},\mathbf{u},t)$ is considered by examining the zeros of its gradient,

$$\nabla_{\mathbf{u}}[\widetilde{\mathcal{H}}](\mathbf{x},\mathbf{u},t) = \widetilde{\mathcal{H}}_1(\mathbf{x},t) + \widetilde{\mathcal{H}}_2(\mathbf{x},t)\mathbf{u} = \mathbf{0}\,, \tag{7.31}$$

yielding the regular control,

$$\begin{aligned}\mathbf{u}^{(\mathrm{reg})}(\mathbf{x},t) &= -\widetilde{\mathcal{H}}_2^{-1}(\mathbf{x},t)\widetilde{\mathcal{H}}_1(\mathbf{x},t)\\ &= -C_2^{-1}(\mathbf{x},t)\left(\mathbf{C}_1 + f_1^\top\nabla_{\mathbf{x}}[v^*]\right)(\mathbf{x},t)\,,\end{aligned} \tag{7.32}$$

with the existence of the inverse being guaranteed by positive definiteness. The fact that the regular control can be solved for exactly in terms of the optimal value $v^*(\mathbf{x},t)$ is a major benefit of having an LQJD problem that is just quadratic in the control. If the usual LQ assumption it made that the control is unconstrained, then the regular control is also the optimal control:

$$\mathbf{u}^*(\mathbf{x},t) = \mathbf{u}^{(\mathrm{reg})}(\mathbf{x},t) \tag{7.33}$$

and the optimal Hamiltonian using (7.32) is

$$\begin{aligned}\widetilde{\mathcal{H}}^*(\mathbf{x},t) &\equiv \widetilde{\mathcal{H}}(\mathbf{x},\mathbf{u}^*,t)\\ &= \left[\widetilde{\mathcal{H}}_0 - \widetilde{\mathcal{H}}_1^\top\widetilde{\mathcal{H}}_2^{-1}\widetilde{\mathcal{H}}_1 + \frac{1}{2}\widetilde{\mathcal{H}}_1^\top\widetilde{\mathcal{H}}_2^{-\top}\widetilde{\mathcal{H}}_2\widetilde{\mathcal{H}}_2^{-1}\widetilde{\mathcal{H}}_1\right](\mathbf{x},t)\\ &= \left[\widetilde{\mathcal{H}}_0 - \frac{1}{2}\widetilde{\mathcal{H}}_1^\top\widetilde{\mathcal{H}}_2^{-1}\widetilde{\mathcal{H}}_1\right](\mathbf{x},t)\,,\end{aligned} \tag{7.34}$$

where by symmetry the inverse transpose $\widetilde{\mathcal{H}}_2^{-\top} = \widetilde{\mathcal{H}}_2^{-1}$. Since the difference of the quadratic $\widetilde{\mathcal{H}}$ in control from the designated minimum using the Taylor approximation form and the critical condition (7.31) is

$$\begin{aligned}\widetilde{\mathcal{H}}(\mathbf{x},\mathbf{u},t) - \widetilde{\mathcal{H}}^*(\mathbf{x},t) &= \widetilde{\mathcal{H}}_0 - \widetilde{\mathcal{H}}^*(\mathbf{x},t) + (\mathbf{u}-\mathbf{u}^*)^\top\nabla_{\mathbf{u}}[\widetilde{\mathcal{H}}](\mathbf{x},\mathbf{u}^*,t)\\ &\quad + \frac{1}{2}(\mathbf{u}-\mathbf{u}^*)^\top\nabla_{\mathbf{u}}[\nabla_{\mathbf{u}}^\top[\widetilde{\mathcal{H}}]](\mathbf{x},\mathbf{u}^*,t)(\mathbf{u}-\mathbf{u}^*)\\ &= \frac{1}{2}\widetilde{\mathcal{H}}_1^\top\widetilde{\mathcal{H}}_2^{-1}\widetilde{\mathcal{H}}_1 + \frac{1}{2}(\mathbf{u}-\mathbf{u}^*)^\top\widetilde{\mathcal{H}}_2(\mathbf{u}-\mathbf{u}^*)\\ &= \frac{1}{2}\left(\widetilde{\mathcal{H}}_1^\top\widetilde{\mathcal{H}}_2^{-1}\widetilde{\mathcal{H}}_1\right)(\mathbf{x},t) + \frac{1}{2}(\mathbf{u}-\mathbf{u}^*)^\top\widetilde{\mathcal{H}}_2(\mathbf{x},t)(\mathbf{u}-\mathbf{u}^*)\\ &\ge \frac{1}{2}\left(\widetilde{\mathcal{H}}_1^\top\widetilde{\mathcal{H}}_2^{-1}\widetilde{\mathcal{H}}_1\right)(\mathbf{x},t) \ge 0\,,\end{aligned} \tag{7.35}$$

it is always possible to solve the optimal control in the minimum problem if $C_2(\mathbf{x},t)$ and thus $\widetilde{\mathcal{H}}_2(\mathbf{x},t)$ are symmetric, positive definite. This corresponds to the minimum principle discussed for deterministic optimal control problems in Chapter 6.

Within the generality of this linear quadratic problem in control only, the optimal control will generally be nonlinear in the state, so the corresponding HJB equation,

$$v_t^*(\mathbf{x},t) + \widetilde{\mathcal{H}}^*(\mathbf{x},t) = 0\,, \tag{7.36}$$

will be highly nonlinear in the state, with $\widetilde{\mathcal{H}}^*(\mathbf{x},t)$ given by (7.34) and coefficients (7.28, 7.29, 7.30). This requires careful solution by numerical PDE or PIDE methods or the computational methods of Chapter 9.

These LQJD/U derived results are summarized in the following theorem:

Theorem 7.5. ***LQJD/U Equations:***
Let the problem be the LQJD in control only problem, so that the deterministic plant function $\mathbf{f}(\mathbf{x},\mathbf{u},t)$ *is linear in the control as given in (7.22), the coefficient* $g(\mathbf{x},\mathbf{u},t)$ *of the Wiener process* $d\mathbf{W}(t)$ *is given in (7.23), the jump amplitude* $h(\mathbf{x},\mathbf{u},t,\mathbf{q})$ *of the Poisson jump process* $d\mathbf{P}(t)$ *is given by (7.24), and the quadratic running cost* $C(\mathbf{x},\mathbf{u},t)$ *is given in (7.26).*

Then, the Hamiltonian $\widetilde{\mathcal{H}}(\mathbf{x},\mathbf{u},t)$ *is a quadratic in the control (7.27) with coefficients* $\{\widetilde{\mathcal{H}}_0(\mathbf{x},t),\widetilde{\mathcal{H}}_1(\mathbf{x},t),\widetilde{\mathcal{H}}_2(\mathbf{x},t)\}$ *given in (7.28, 7.29, 7.30), respectively. The optimal control vector, in absence of control constraints, has the linear feedback control form,*

$$\mathbf{u}^*(\mathbf{x},t) = \mathbf{u}^{(\mathrm{reg})}(\mathbf{x},t) = -C_2^{-1}(\mathbf{x},t)\left[\mathbf{C}_1 + f_1^\top \nabla_{\mathbf{x}}[v^*]\right](\mathbf{x},t)\,, \tag{7.37}$$

as long as the quadratic control coefficient $C_2(\mathbf{x},t)$ *is positive definite in case of a minimum expected objective and in absence of constraints on the control. Assuming that an optimal value* $v^*(\mathbf{x},t)$ *solution exists, then* $v^*(\mathbf{x},t)$ *satisfies the Hamilton Jacobi Bellman equation,*

$$v_t^*(\mathbf{x},t) + \left(\widetilde{\mathcal{H}}_0 - \frac{1}{2}\widetilde{\mathcal{H}}_1^\top \widetilde{\mathcal{H}}_2^{-1}\widetilde{\mathcal{H}}_1\right)(\mathbf{x},t) = 0\,. \tag{7.38}$$

The solution $v^*(\mathbf{x},t)$ *is subject to the final condition*

$$v^*(\mathbf{x},t_f) = S(\mathbf{x},t_f)\,, \tag{7.39}$$

and any necessary boundary conditions.

For solutions of LQJD/U problems, computational methods are quite essential; see Hanson's 1996 chapter [108] or Chapter 9.

7.4.2 LLJD/U or the Case $C_2 \equiv 0$:

If the quadratic cost coefficient $C_2(\mathbf{x},t) \equiv 0$, then

$$\widetilde{\mathcal{H}}(\mathbf{x},\mathbf{u},t) = \widetilde{\mathcal{H}}_0(\mathbf{x},t) + \widetilde{\mathcal{H}}_1^\top(\mathbf{x},t)\mathbf{u}\,, \tag{7.40}$$

the linear linear jump-diffusion (LLJD/U) problem in control only. The minimum with respect to the control depends on the linear cost coefficient

$$\widetilde{\mathcal{H}}^*(\mathbf{x},t) = \min_{\mathbf{u}}\left[\widetilde{\mathcal{H}}_0(\mathbf{x},t) + \widetilde{\mathcal{H}}_1^\top(\mathbf{x},t)\mathbf{u}\right] = \widetilde{\mathcal{H}}_0(\mathbf{x},t) + \min_{\mathbf{u}}\left[\widetilde{\mathcal{H}}_1^\top(\mathbf{x},t)\mathbf{u}\right]. \tag{7.41}$$

Since this is a problem of linear or singular control, it makes sense only if the control is constrained, e.g., component-wise constraints,

$$U_i^{(\min)} \le u_i \le U_i^{(\max)}\,. \tag{7.42}$$

For this type of constraint the minimum is separable by component and the optimal control ia a n_u-dimensional bang-bang control

$$\begin{aligned}\widetilde{\mathcal{H}}^*(\mathbf{x},t) &= \widetilde{\mathcal{H}}_0(\mathbf{x},t) + \sum_{i=1}^{n_u} \min\left[\widetilde{\mathcal{H}}_{1,i}(\mathbf{x},t)u_i\right] \\ &= \widetilde{\mathcal{H}}_0(\mathbf{x},t) + \sum_{i=1}^{n_u} \left\{ \begin{array}{ll} \widetilde{\mathcal{H}}_{1,i}(\mathbf{x},t)U_i^{(\max)}, & \widetilde{\mathcal{H}}_{1,i}(\mathbf{x},t) < 0 \\ 0, & \widetilde{\mathcal{H}}_{1,i}(\mathbf{x},t) = 0 \\ \widetilde{\mathcal{H}}_{1,i}(\mathbf{x},t)U_i^{(\min)}, & \widetilde{\mathcal{H}}_{1,i}(\mathbf{x},t) > 0 \end{array} \right\} \\ &= \widetilde{\mathcal{H}}_0(\mathbf{x},t) + \frac{1}{2}\widetilde{\mathcal{H}}_1(\mathbf{x},t). * \Big[\mathbf{U}^{(\min)}. *(\mathbf{1}+\mathbf{sgn}_1) \\ &\quad +\mathbf{U}^{(\max)}. *(\mathbf{1}-\mathbf{sgn}_1)\Big]\,, \end{aligned} \tag{7.43}$$

where $\mathbf{1} \equiv [1]_{n_u\times 1}$, $\mathbf{sgn}_1 \equiv [\mathrm{sgn}(\widetilde{\mathcal{H}}_{1,i}(\mathbf{x},t)]_{n_u\times 1}$,

$$\mathrm{sgn}(x) \equiv \left\{ \begin{array}{ll} -1, & x<0 \\ 0, & x=0 \\ +1, & x>0 \end{array} \right\} \tag{7.44}$$

is the sign or signum function, $\mathbf{U}^{(\min)} \equiv [U_i^{(\min)}]_{n_u\times 1}$, $\mathbf{U}^{(\max)} \equiv [U_i^{(\max)}]_{n_u\times 1}$, and $\mathbf{v}.*\mathbf{u} \equiv [v_i u_i]_{n_u\times 1}$ is the dot-star or element-by-element product. The optimal control is undefined for components for which $\widetilde{\mathcal{H}}_{1,i}(\mathbf{x},t) = 0$, but otherwise is given in composite form:

$$u_i^*(\mathbf{x},t) = \left\{ \begin{array}{ll} U_i^{(\max)}, & \widetilde{\mathcal{H}}_{1,i}(\mathbf{x},t) < 0 \\ U_i^{(\min)}, & \widetilde{\mathcal{H}}_{1,i}(\mathbf{x},t) > 0 \end{array} \right\}. \tag{7.45}$$

If the components of $\widetilde{\mathcal{H}}_1$ change sign often, then that can lead to **chattering control**.

7.4.3 Canonical LQJD Problem

The standard or canonical LQJD problem is linear in the dynamics and quadratic in the costs with respect to both state and control vectors. This LQJD problem is a special case of the LQJD problem in control only and results in substantial simplifications of the solution with a quadratic state decomposition of the optimal value function and the a linear or feedback decomposition of the optimal control vector. The decomposition of optimal value and control is similar to that of the deterministic LQ problem, but here the more general quadratic state and linear control decompositions is presented.

Let the more general jump-diffusion linear quadratic model be given with the plant function for the deterministic or non-noise dynamics term and be linear in both state $\mathbf{X}(t)$ and $\mathbf{U}(t)$,

$$\mathbf{f}(\mathbf{x},\mathbf{u},t) = \mathbf{f}_0(t) + f_{1,1}^\top(t)\mathbf{x} + f_{1,2}^\top(t)\mathbf{u}\,, \tag{7.46}$$

the first subscript indicating the degree and the subsequent subscripts, if present, indicating either state (1) or control (2), with the diffusion term,

$$g(\mathbf{x},\mathbf{u},t) = g_0(t)\,, \tag{7.47}$$

assumed state-independent and control-independent for simplicity, and with the jump term,

$$h(\mathbf{x},\mathbf{u},t,\mathbf{q}) = h_0(t,\mathbf{q})\,, \tag{7.48}$$

also assumed state-independent and control-independent for simplicity. The current form of the linear SDE (7.4) is written here as

$$d\mathbf{X}(s) \overset{\text{sym}}{=} \mathbf{f}(\mathbf{X}(s),\mathbf{U}(s),s)ds + g_0(s)d\mathbf{W}(s) + h_0(s,\mathbf{Q})d\mathbf{P}(s;s)\,, \tag{7.49}$$

on $t \le s \le t_f$, with $\mathrm{E}[d\mathbf{P}(t;t)] = [\lambda_{0,j}(t)dt]_{n_p\times 1}$.

The quadratic running cost function is

$$\begin{aligned} C(\mathbf{x},\mathbf{u},t) &= C_0(t) + \mathbf{C}_{1,1}^\top(t)\mathbf{x} + \mathbf{C}_{1,2}^\top(t)\mathbf{u} \\ &\quad + \frac{1}{2}\mathbf{x}^\top C_{2,1,1}(t)\mathbf{x} + \mathbf{x}^\top C_{2,1,2}(t)\mathbf{u} + \frac{1}{2}\mathbf{u}^\top C_{2,2,2}(t)\mathbf{u} \end{aligned} \tag{7.50}$$

and the terminal cost also has a general quadratic form

$$S(\mathbf{X}(t_f),t_f) = S_0(t_f) + \mathbf{S}_1^\top(t_f)\mathbf{X}(t_f) + \frac{1}{2}\mathbf{X}^\top(t_f)S_2(t_f)\mathbf{X}(t_f)\,, \tag{7.51}$$

in the state vector. It is assumed that all right hand side coefficients are commensurate in multiplication and the product is the same type as that on the left hand side. It is assumed that all coefficients are well-defined, but in particular that $C_{2,2,2}(t)$ is positive definite for the minimum problem, a crucial assumption, and symmetric due to the quadratic form, while $C_{2,1,1}(t)$ and $C_{2,1,2}(t)$ need to be positive semi-definite. Also, $S_2(t_f)$ is symmetric, positive semi-definite.

As in the deterministic LQ problem in Section 6.4.4, a quadratic function of the state vector is sought. However, due to the extra linear terms in the quadratic cost beyond the pure quadratic form in (6.125) a more general quadratic decomposition is heuristically assumed for the optimal value,

$$v^*(\mathbf{x},t) = V_0(t) + \mathbf{V}_1^\top(t)\mathbf{x} + \frac{1}{2}\mathbf{x}^\top V_2(t)\mathbf{x}\,, \tag{7.52}$$

where the optimal value coefficients $\{V_0(t),\mathbf{V}_1(t),V_2(t)\}$ are compatible in multiplication and any product is scalar valued. Without loss of generality, the quadratic coefficient $V_2(t)$ is taken to be symmetric. Consequently, the partial derivative with respect to time is

$$v_t^*(\mathbf{x},t) = \dot{V}_0(t) + \dot{\mathbf{V}}_1^\top(t)\mathbf{x} + \frac{1}{2}\mathbf{x}^\top \dot{V}_2(t)\mathbf{x}\,,$$

where $\dot{V}_j(t)$ denotes the time derivative of $V_j(t)$ for $j = 0:2$, the state gradient is

$$\nabla_{\mathbf{x}}[v^*](\mathbf{x},t) = \mathbf{V}_1(t) + V_2(t)\mathbf{x}\,,$$

the state Hessian is

$$\nabla_{\mathbf{x}}\left[\nabla_{\mathbf{x}}^{\top}[v^*]\right](\mathbf{x},t) = V_2(t)$$

and the jump increment is

$$\Delta_j[v^*](\mathbf{x},t,q_j) = \mathbf{V}_1^{\top}(t)\widehat{\boldsymbol{h}}_{0,j}(t,q_j) + \frac{1}{2}\widehat{\boldsymbol{h}}_{0,j}^{\top}(t,q_j)V_2(t)\widehat{\boldsymbol{h}}_{0,j}(t,q_j) \\ +\mathbf{x}^{\top}V_2(t)\widehat{\boldsymbol{h}}_{0,j}(t,q_j)\,,$$

where

$$\widehat{\boldsymbol{h}}_{0,j}(t,q_j) = [h_{0,i,j}(t,q_j)]_{n_x\times 1}$$

for $j = 1 : n_p$.

With the proposed general quadratic decomposition (7.52) of $v^*(\mathbf{x},t)$, the pseudo-Hamiltonian has a quadratic decomposition in both state and control vectors like the cost coefficient $C(\mathbf{x},\mathbf{u},t)$ decomposition (7.50),

$$\widetilde{\mathcal{H}}(\mathbf{x},\mathbf{u},t) = \widetilde{\mathcal{H}}_0(t) + \widetilde{\mathcal{H}}_{1,1}^{\top}(t)\mathbf{x} + \widetilde{\mathcal{H}}_{1,2}^{\top}(t)\mathbf{u} \\ +\frac{1}{2}\mathbf{x}^{\top}\widetilde{\mathcal{H}}_{2,1,1}(t)\mathbf{x} + \mathbf{x}^{\top}\widetilde{\mathcal{H}}_{2,1,2}(t)\mathbf{u} + \frac{1}{2}\mathbf{u}^{\top}\widetilde{\mathcal{H}}_{2,2,2}(t)\mathbf{u}\,, \tag{7.53}$$

where the scalar coefficient is

$$\widetilde{\mathcal{H}}_0(t) = C_0(t) + \mathbf{f}_0^{\top}(t)\mathbf{V}_1(t) + \frac{1}{2}\left(g_0 g_0^{\top}\right)(t) : V_2(t) \\ +\mathbf{V}_1^{\top}(t)\overline{h}_0(t).*\boldsymbol{\lambda}_0(t) + \frac{1}{2}V_2(t).*\overline{(h_0\Lambda h_0)}(t)\,, \tag{7.54}$$

where

$$\overline{h}_0(t) \equiv \left[\int_{\mathcal{Q}_j} h_{0,i,j}(t,q_j)\phi_{Q_j}(q_j;t)dq_j\right]_{n_x\times n_p}, \tag{7.55}$$

$$\boldsymbol{\lambda}_0(t) \equiv [\lambda_{0,i}(t)]_{n_p\times 1}\,, \tag{7.56}$$

$$\Lambda_0(t) \equiv [\lambda_{0,i}(t)\delta_{i,j}]_{n_p\times n_p}\,, \tag{7.57}$$

$$\overline{h_0\Lambda_0 h_0^{\top}}(t) \equiv \left[\sum_{k=1}^{n_p}\lambda_{0,k}\int_{\mathcal{Q}_j} h_{0,i,k}(t,q_k)h_{0,j,k}(t,q_k)\phi_{Q_k}(q_k;t)dq_k\right]_{n_x\times n_x}, \tag{7.58}$$

the linear state coefficients is

$$\widetilde{\mathcal{H}}_{1,1}(t) = C_{1,1}(t) + f_{1,1}(t)\mathbf{V}_1(t) + V_2(t)\mathbf{f}_0(t) + V_2(t)\overline{h}_0(t).*\boldsymbol{\lambda}_0(t)\,, \tag{7.59}$$

the linear control coefficient is

$$\widetilde{\mathcal{H}}_{1,2}(t) = C_{1,2}(t) + f_{1,2}(t)\mathbf{V}_1(t)\,, \tag{7.60}$$

and the quadratic coefficients are

$$\widetilde{\mathcal{H}}_{2,1,1}(t) = C_{2,1,1}(t) + 2f_{1,1}(t)\mathbf{V}_2(t)\,, \tag{7.61}$$

$$\widetilde{\mathcal{H}}_{2,1,2}(t) = C_{2,1,2}(t) + V_2^{\top}(t)f_{1,2}^{\top}(t)\,, \tag{7.62}$$

$$\widetilde{\mathcal{H}}_{2,2,2}(t) = C_{2,2,2}(t)\,. \tag{7.63}$$

Since quadratic forms only operate on the symmetric part of the quadratic coefficient (0.133), $\widetilde{\mathcal{H}}_{2,2,2}(t)$ will be symmetric, positive definite with $C_{2,2,2}(t)$.

The optimal control is the same as the regular control in the absence of control constraints, so the zero of

$$\nabla_{\mathbf{u}}[\widetilde{\mathcal{H}}](\mathbf{x},\mathbf{u},t) = \widetilde{\mathcal{H}}_{1,2}(t) + \widetilde{\mathcal{H}}_{2,1,2}^{\top}(t)\mathbf{x} + \widetilde{\mathcal{H}}_{2,2,2}(t)\mathbf{u}$$

results in

$$\begin{aligned}\mathbf{u}^*(\mathbf{x},t) &= -\widetilde{\mathcal{H}}_{2,2,2}^{-1}(t)\left(\widetilde{\mathcal{H}}_{1,2}(t) + \widetilde{\mathcal{H}}_{2,1,2}^{\top}(t)\mathbf{x}\right)\\ &= -C_{2,2,2}^{-1}(t)\left(C_{1,2}(t) + f_{1,2}(t)\mathbf{V}_1(t)\right.\\ &\quad \left. + \left(C_{2,1,2}^{\top}(t) + f_{1,2}(t)V_2(t)\right)\mathbf{x}\right) . \end{aligned} \tag{7.64}$$

Hence, the optimal control vector is a linear or affine function of the state vector, the general form of linear feedback control. This completes the preliminary work on the LQJD problem for the feedback control state dependence.

Upon substituting the preliminary reduction of the linear optimal control (7.64) into the Hamilton Jacobi Bellman equation (7.36), then the HJB equation becomes

$$\begin{aligned}0 = {}& \dot{V}_0(t) + \dot{\mathbf{V}}_1^{\top}(t)\mathbf{x} + \frac{1}{2}\mathbf{x}^{\top}\dot{V}_2(t)\mathbf{x} + \widetilde{\mathcal{H}}_0(t) + \widetilde{\mathcal{H}}_{1,1}^{\top}(t)\mathbf{x}\\ & -\widetilde{\mathcal{H}}_{1,2}^{\top}(t)\widetilde{\mathcal{H}}_{2,2,2}^{-1}(t)\left(\widetilde{\mathcal{H}}_{1,2}(t) + \widetilde{\mathcal{H}}_{2,1,2}^{\top}(t)\mathbf{x}\right)\\ & +\frac{1}{2}\mathbf{x}^{\top}\widetilde{\mathcal{H}}_{2,1,2}(t)\widetilde{\mathcal{H}}_{2,2,2}^{-1}(t)\mathbf{x} - \mathbf{x}^{\top}\widetilde{\mathcal{H}}_{2,1,2}(t)\widetilde{\mathcal{H}}_{2,2,2}^{-1}(t)\left(\widetilde{\mathcal{H}}_{1,2}(t) + \widetilde{\mathcal{H}}_{2,1,2}^{\top}(t)\mathbf{x}\right)\\ & +\frac{1}{2}\left(\widetilde{\mathcal{H}}_{1,2}^{\top}(t) + \mathbf{x}^{\top}\widetilde{\mathcal{H}}_{2,1,2}(t)\right)\widetilde{\mathcal{H}}_{2,2,2}^{-1}(t)\left(\widetilde{\mathcal{H}}_{1,2}(t) + \widetilde{\mathcal{H}}_{2,1,2}^{\top}(t)\mathbf{x}\right) . \end{aligned} \tag{7.65}$$

Next, separating this LQJD form of the HJBE (7.65) into purely quadratic terms, purely linear terms and state-independent terms leads to a set of three uni-directionally coupled ordinary matrix differential equations for the optimal control coefficients $V_2(t)$, $\mathbf{V}_1(t)$ and $V_0(t)$ which are summarized in the following theorem which we have just derived.

Theorem 7.6. ***LQJD Equations:***
Let the $n_x \times 1$ jump-diffusion state process $\mathbf{X}(t)$ satisfy dynamics linear in both the state and the $n_u \times 1$ control $\mathbf{U}(t)$ with $n_x \times 1$ linear deterministic plant term

$$\mathbf{f}(\mathbf{x},\mathbf{u},t) = \mathbf{f}_0(t) + f_{1,1}^{\top}(t)\mathbf{x} + f_{1,2}^{\top}(t)\mathbf{u}$$

from (7.46), with $n_x \times n_w$ state and control independent diffusion coefficient $g_0(t)$ of the $n_w \times 1$ Wiener process $d\mathbf{W}(t)$, and with $n_x \times n_p$ state and control independent jump amplitude $h_0(t,q)$ (7.47) of the $n_p \times 1$ Poisson process $d\mathbf{P}(t)$. Let the scalar quadratic running cost be

$$\begin{aligned}C(\mathbf{x},\mathbf{u},t) = {}& C_0(t) + \mathbf{C}_{1,1}^{\top}(t)\mathbf{x} + \mathbf{C}_{1,2}^{\top}(t)\mathbf{u}\\ & +\frac{1}{2}\mathbf{x}^{\top}C_{2,1,1}(t)\mathbf{x} + \mathbf{x}^{\top}C_{2,1,2}(t)\mathbf{u} + \frac{1}{2}\mathbf{u}^{\top}C_{2,2,2}(t)\mathbf{u}\end{aligned}$$

and terminal cost be

$$S(\mathbf{X}(t_f),t_f) = S_0(t_f) + \mathbf{S}_1^\top(t_f)\mathbf{X}(t_f) + \frac{1}{2}\mathbf{X}^\top(t_f)S_2(t_f)\mathbf{X}(t_f)\,.$$

Then the optimal stochastic control problem admits a solution quadratic in the state vector

$$v^*(\mathbf{x},t) = V_0(t) + \mathbf{V}_1^\top(t)\mathbf{x} + \frac{1}{2}\mathbf{x}^\top V_2(t)\mathbf{x}\,,$$

with optimal control vector that in linear in the state vector

$$\mathbf{u}^*(\mathbf{x},t) = -C_{2,2,2}^{-1}(t)\left(C_{1,2}(t) + f_{1,2}(t)\mathbf{V}_1(t) + \left(C_{2,1,2}^\top(t) + f_{1,2}(t)V_2(t)\right)\mathbf{x}\right)\,.$$

The optimal value $v^*(\mathbf{x},t)$ coefficients satisfy a uni-directionally coupled set of matrix ordinary differential equations, which are solved starting from the $n_x \times n_x$ quadratic coefficient equation

$$\begin{aligned} 0 = {} & \dot{V}_2(t) + C_{2,1,1}(t) + 2f_{1,1}(t)V_2(t) \\ & - \left(C_{2,1,2}(t) + V_2(t)f_{1,1}^\top(t)\right)C_{2,2,2}^{-1}(t)\left(C_{2,1,2}^\top(t) + f_{1,1}(t)V_2(t)\right) \end{aligned} \tag{7.66}$$

for $V_2(t)$, then the $n_x \times 1$ linear coefficient equation

$$\begin{aligned} 0 = {} & \dot{\mathbf{V}}_1(t) + \mathbf{C}_{1,1}(t) + f_{1,1}(t)\mathbf{V}_1(t) \\ & - \left(C_{2,1,2}(t) + V_2(t)f_{1,1}^\top(t)\right)C_{2,2,2}^{-1}(t)\left(\mathbf{C}_{1,2}(t) + f_{1,2}(t)\mathbf{V}_1(t)\right) \\ & + V_2(t)\overline{h}_0(t)\boldsymbol{\lambda}_0(t) \end{aligned} \tag{7.67}$$

for $\mathbf{V}_1(t)$ using the existing solution for $V_2(t)$, and finally the scalar state-independent coefficient equation

$$\begin{aligned} 0 = {} & \dot{V}_0(t) + C_0(t) + \mathbf{f}_0^\top(t)\mathbf{V}_1(t) + \frac{1}{2}g_0(t)g_0^\top(t) : V_2(t) \\ & - \frac{1}{2}\left(\mathbf{C}_{1,2}^\top(t) + \mathbf{V}_1^\top(t)f_{1,2}(t)\right)C_{2,2,2}^{-1}(t)\left(\mathbf{C}_{1,2}(t) + f_{1,2}(t)\mathbf{V}_1(t)\right) \\ & + \mathbf{V}_1^\top(t)\overline{h}_0(t). * \boldsymbol{\lambda}_0(t) + \frac{1}{2}\overline{\left(h_0\Lambda_0 h_0^\top\right)}(t) : V_2(t). \end{aligned} \tag{7.68}$$

Remarks 7.7.

- *The nonlinear differential equation (7.66) for the quadratic coefficient $V_2(t)$ is called a* ***matrix Riccati equation*** *due to the quadratic linearity in $V_2(t)$. Since $V_2(t)$ can be assumed to be symmetric without loss of generality since it is defined as the coefficient of a quadratic from, computational effort can be reduced to just finding the upper or lower triangular part, i.e., just $n_x(n_x+1)/2$ elements.*

- *Once $V2(t)$ is known or a reasonable approximation is found, the equation (7.67) for the linear coefficient $V_1(t)$ will be a linear matrix equation which is relatively simpler to solve than the matrix Riccati equation.*

- *Similarly, once both $V_2(t)$ and $\mathbf{V}_1(t)$ are found to reasonable approximations, then equation (7.68) for the state-independent coefficient $V_0(t)$ will be a linear scalar equation.*
- *Once the solutions to the time-dependent coefficients $V_2(t)$, $\mathbf{V}_1(t)$ and $V_0(t)$ are obtained, then the optimal value $v^*(\mathbf{x}, t)$ quadratic decomposition (7.52) is justified, at least heuristically.*

7.5 Exercises

1. For the linear jump-diffusion dynamics

$$dX(t) = (\mu_0 X(t) + \beta_0 U(t))dt + \sigma_0 dW(t) + \nu_0 X(t) dP(t),$$

for $0 \le t \le t_f$ and state $X(0) = x_0 > 0$ and the control process $-\infty < U(t) < +\infty$ is unconstrained. The coefficients $\mu_0 \neq 0$, $\beta_0 \neq 0$, $\sigma_0 > 0$, $\nu_0 \neq 0$ and $\lambda_0 > 0$ are constants, where $\mathrm{E}[dP(t)] = \lambda_0 dt$ (note that the jump process here is a discrete, Poisson process, since there is no mark process). The costs are quadratic, i.e.,

$$V[X, U](X(t), t) = \frac{1}{2} \int_t^{t_f} \left(q_0 X^2(s) + r_0 U^2(s) \right) ds + \frac{1}{2} S_f X^2(t_f)$$

for $q_0 > 0$, $r_0 > 0$, and $S_f > 0$. Let the optimal, expected value be

$$v^*(x, t) = \min \left[\mathrm{E} \left[V[X, U](X(t), t) \,|\, X(t) = x, U(t) = u \right] \right].$$

   (a) Write down the proper Hamilton-Jacobi-Bellman equation for the optimal, expected value and control;

   (b) Obtain an LQJD (LQGP) solution form for $v^*(x, t)$ and a linear feedback control law for $u^*(x, t)$;

   (c) Derive the corresponding Riccati equation for this special case.

2. Derive the modifications necessary in the set of Riccati-like equations for the Linear-Quadratic Jump-Diffusion (LQJD, LQGP or JLQG) problem when the dynamics are scalar and linear (affine), i.e.,

$$dX(t) = f(X(t), U(t), t)dt + g(X(t), U(t), t)dW(t) + h(X(t), U(t), t)dP(t)\,,$$

where

$$\mathrm{E}[dP(t)] = \lambda(t)dt\,,$$

$$f(x, u, t) = f_{0,0}(t) + f_{1,1}(t)x + f_{1,2}(t)u\,,$$

$$g(x, u, t) = g_{0,0}(t) + g_{1,1}(t)x\,,$$

$$h(x, u, t) = h_{0,0}(t) + h_{1,1}(t)x\,,$$

the jump amplitude being independent of any mark process. The running and terminal costs for a maximum objective are quadratic,

$$C(x,u,t) = C_{0,0}(t) + C_{1,1}(t)x + C_{1,2}(t)u + 0.5C_{2,1,1}(t)x^2 + C_{2,1,2}(t)xu \\ +0.5C_{2,2,2}(t)u^2\,,$$

where $C_{02}(t) < 0$, and

$$S(x,t) = S_0(t) + S_1(t)x + 0.5 * S_2(t)x^2\,,$$

where $S_2(t) < 0$.

If the objective is to maximize the expected total utility in the unconstrained control case, then find the coefficient functions $v_0(t)$, $v_1(t)$, $v_2(t)$, $u_0(t)$ and $u_1(t)$ in the solutions

$$v^*(x,t) = v_0(t) + v_1(t)x + 0.5v_2(t)x^2$$

and

$$u^*(x,t) = u_0(t) + u_1(t)x$$

explicitly in terms of the dynamical and cost coefficient functions. Do not try to solve the Riccati equation system for $\{v_0(t), v_1(t), v_2(t)\}$.

Suggested References for Further Reading

- Cont and Tankov, 2004 [59].
- Fleming and Rishel, 1975 [85].
- Gihman and Skorohod, 1979 [95].
- Hanson, 1996 [108].
- Jazwinski, 1970 [151].
- Kushner, 1967 [169].
- Lewis, 1986 [180]
- Øksendal and Sulem, 2005 [219].
- Runggaldier, 2003 [235].
- Stengel, 1994 [253].
- Yong and Zhou, 1999 [283].