Stochastic Optimal Control

Bangyan Liao
liaobangyan@westlake.edu.cn

Nov 2024

WESTLAKE UNIVERSITY

Outline

Stochastic Optimal Control

Quadratic-regularized State Cost SOC

Outline

Stochastic Optimal Control

Definition (Stochastic Optimal Control)

Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathcal{P})$ be a fixed filtered probability space on which is defined a Brownian motion $W = (W_t)_{t\geq 0}$. We consider the control-affine problem

$$\min_{u\in\mathcal{U}} \mathbb{E}\Big[\int_0^T w(X_t^u, u_t, t)\,\mathrm{d}t + g(X_T^u)\Big],$$
$$\text{where } \mathrm{d}X_t^u = (b(X_t^u, t) + \sigma(t)u(X_t^u, t))\,\mathrm{d}t + \sqrt{\lambda}\sigma(t)\mathrm{d}W_t, \qquad X_0^u \sim p_0. \tag{1}$$

and where $X_t^u \in \mathbb{R}^d$ is the state, $u : \mathbb{R}^d \times [0, T]$ is the feedback control and belongs to the set of admissible controls $\mathcal{U}$, w is the state cost, $g : \mathbb{R}^d \to \mathbb{R}$ is the terminal cost, $b : \mathbb{R}^d \times [0, T] \to \mathbb{R}^d$ is the base drift, and $\sigma : [0, T] \to \mathbb{R}^{d\times d}$ is the invertible diffusion coefficient and $\lambda \in (0, +\infty)$ is the noise level.

Definition (Cost Functional and Value Function)

The *cost functional* for the control u, point x and time t is defined as $J(u; x, t) := \mathbb{E}\big[\int_t^T w(X_t^u, u_t, t)\,\mathrm{d}t + g(X_T^u)\big| X_t^u = x\big]$. That is, the cost functional is the expected value of the control objective restricted to the times $[t, T]$ with the initial value x at time t. The *value function* or *optimal cost-to-go* at a point x and time t is defined as the minimum value of the cost functional across all possible controls:

$$V(x, t) := \inf_{u \in \mathcal{U}} J(u; x, t). \tag{2}$$

Definition (HJB Optimality Condition for SOC)

If we define the infinitesimal generator $\mathcal{L} := \frac{\lambda}{2}\sum_{i,j=1}^{d}(\sigma\sigma^\top)_{ij}(t)\partial_{x_i}\partial_{x_j} + \sum_{i=1}^{d} b_i(x,t)\partial_{x_i} + \sum_{i=1}^{d}\sigma_i(t)u_i(x,t)\partial_{x_i}$, the value function solves the following Hamilton-Jacobi-Bellman (HJB) partial differential equation:

$$\frac{\partial V(x,t)}{\partial t} + \min_{u\in\mathcal{U}}\{\mathcal{L}V(x,t) + w(x,u,t)\} = 0,\ V(x,T) = g(x). \tag{3}$$

Proof.

Recall the Itô Lemma for SDE $\mathrm{d}X_t^u = (b(X_t^u, t) + \sigma(t)u(X_t^u, t))\,\mathrm{d}t + \sqrt{\lambda}\sigma(t)\mathrm{d}W_t$:

$$
\begin{aligned}
\mathrm{d}V(X_t^u, t) &= \frac{\partial V(X_t^u, t)}{\partial t}\mathrm{d}t + \frac{\partial V(X_t^u, t)}{\partial x}\mathrm{d}X_t^u + \frac{1}{2}\frac{\partial^2 V(X_t^u, t)}{\partial x^2}(\mathrm{d}X_t^u)^2 \\
&= \frac{\partial V_t}{\partial t}\mathrm{d}t + \mathcal{L}V(x, t)\mathrm{d}t + \nabla V_t(x) \cdot \sqrt{\lambda}\sigma(t)\mathrm{d}W_t
\end{aligned} \tag{4}
$$

where the $\mathcal{L}V(x, t)$ is the generator which defines as

$$
\mathcal{L}V(x, t) = \nabla V_t(x) \cdot (b(X_t^u, t) + \sigma(t)u(X_t^u, t)) + \frac{\lambda}{2} Trace\left[\sigma(t)\sigma(t)^\top \nabla^2 V_t(x)\right] \tag{5}
$$

Proof.

Recall the proof of HJB equation in the optimal control section. The key step is

$$
\begin{aligned}
V(s,\mathbf{z}) &= \inf_{\boldsymbol{\theta}} \left\{ \int_s^{s+\Delta s} L(t,\mathbf{x}(t),\boldsymbol{\theta}(t))\mathrm{d}t + V(s+\Delta s,\mathbf{x}(s+\Delta s)) \right\} \\
&\approx \inf_{\boldsymbol{\theta}} \left\{ L(s,\mathbf{z},\boldsymbol{\theta}(s))\Delta s + V(s+\Delta s,\mathbf{x}(s+\Delta s)) \right\} \\
&\approx \inf_{\boldsymbol{\theta}} \{ L(s,\mathbf{z},\boldsymbol{\theta}(s))\Delta s + V(s,\mathbf{x}(s)) \\
&\qquad + \partial_s V(s,\mathbf{z})\Delta s + [\nabla_{\mathbf{z}} V(s,\mathbf{z})]^\top f(s,\mathbf{z},\boldsymbol{\theta}(s))\Delta s \} \\
\dot{\mathbf{x}}(t) &= f(t,\mathbf{x}(t),\boldsymbol{\theta}(t)), \quad t\in[s,\tau], \quad \mathbf{x}(s)=\mathbf{z}
\end{aligned}
\tag{4}
$$

Proof.

Similarly, we can derive the HJB equation for SOC as:

$$
\begin{aligned}
V(X_s^u, s) &= \inf_u \mathbb{E}\left\{\int_s^{s+\Delta s} w(X_t^u, u, t)\mathrm{d}t + V(X_{s+\Delta s}^u, s+\Delta s)\right\} \\
&\approx \inf_u \mathbb{E}\left\{w(X_s^u, u, s)\Delta s + V(X_{s+\Delta s}^u, s+\Delta s)\right\} \\
&\approx \inf_u \mathbb{E}\{w(X_t^u, u, t)\Delta s + V(X_s^u, s) \\
&\quad + \partial_s V(\mathbf{z}, s)\Delta s + \mathcal{L}V(\mathbf{z}, s)\Delta s + \nabla V_s(\mathbf{z}) \cdot \sqrt{\lambda}\sigma(s)\Delta dW_s\}
\end{aligned}
\tag{4}
$$

$$
\mathrm{d}X_t^u = \left(b(X_t^u, t) + \sigma(t)u(X_t^u, t)\right)\mathrm{d}t + \sqrt{\lambda}\sigma(t)\mathrm{d}W_t, \quad t \in [s, \tau], \quad X_s^u = \mathbf{z}
$$

□

Outline

Stochastic Optimal Control

Quadratic-regularized State Cost SOC

- HJB Optimality Condition
- Path Integral Control
- Forward and Backward SDEs
- Verification Theorem

Quadratic-regularized State Cost SOC

Definition (Quadratic-regularized State Cost)

Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathcal{P})$ be a fixed filtered probability space on which is defined a Brownian motion $W = (W_t)_{t\geq 0}$. We consider the control-affine problem

$$\min_{u\in\mathcal{U}} \mathbb{E}\Big[\int_0^T \big(\frac{1}{2}\|u(X_t^u, t)\|^2 + f(X_t^u, t)\big)\,\mathrm{d}t + g(X_T^u)\Big],$$
$$\text{where } \mathrm{d}X_t^u = \big(b(X_t^u, t) + \sigma(t)u(X_t^u, t)\big)\,\mathrm{d}t + \sqrt{\lambda}\sigma(t)\mathrm{d}W_t, \qquad X_0^u \sim p_0. \tag{5}$$

and where $X_t^u \in \mathbb{R}^d$ is the state, $u : \mathbb{R}^d \times [0, T]$ is the feedback control and belongs to the set of admissible controls $\mathcal{U}$, $f : \mathbb{R}^d \times [0, T] \to \mathbb{R}$ is the state cost, $g : \mathbb{R}^d \to \mathbb{R}$ is the terminal cost, $b : \mathbb{R}^d \times [0, T] \to \mathbb{R}^d$ is the base drift, and $\sigma : [0, T] \to \mathbb{R}^{d\times d}$ is the invertible diffusion coefficient and $\lambda \in (0, +\infty)$ is the noise level.

Definition (HJB equation fot Quadratic-regularized State Cost)

Since the unique optimal control is given in terms of the value function as $u^*(x,t) = -\sigma(t)^\top \nabla V(x,t)$. If we define the infinitesimal generator $L := \frac{\lambda}{2}\sum_{i,j=1}^{d}(\sigma\sigma^\top)_{ij}(t)\partial_{x_i}\partial_{x_j} + \sum_{i=1}^{d} b_i(x,t)\partial_{x_i}$, the value function solves the following Hamilton-Jacobi-Bellman (HJB) partial differential equation:

$$
\begin{aligned}
&(\partial_t + L)V(x,t) - \frac{\lambda}{2}\|(\sigma^\top \nabla V)(x,t)\|^2 + f(x,t) = 0, \\
&V(x,T) = g(x).
\end{aligned}
\tag{6}
$$

Lemma (Path-integral representation of the optimal control)

$$
\begin{aligned}
u^*(x,t) &= \lambda\sigma(t)^\top \nabla_x \log \mathbb{E}\Big[\exp\Big(-\lambda^{-1}\int_t^T f(X_s,s)\,\mathrm{d}s - \lambda^{-1} g(X_T)\Big)\Big| X_t = x\Big] \\
V(x,t) &= -\lambda \log \mathbb{E}\Big[\exp\Big(-\lambda^{-1}\int_t^T f(X_s,s)\,\mathrm{d}s - \lambda^{-1} g(X_T)\Big)\Big| X_t = x\Big],
\end{aligned}
\tag{7}
$$

where X_t is generated by the uncontrolled process. The optimal control and the value function are related to each other by $u^*(x,t) = -\sigma(t)^\top \nabla V(x,t)$.

Proof. (Path-integral Control).

Let us recall the HJB optimality condition

$$\begin{aligned}
&(\partial_t + L)V(x,t) - \frac{\lambda}{2}\|(\sigma^\top \nabla V)(x,t)\|^2 + f(x,t) = 0,\\
&L = \frac{\lambda}{2}\sum_{i,j=1}^{d}(\sigma\sigma^\top)_{ij}(t)\partial_{x_i}\partial_{x_j} + \sum_{i=1}^{d} b_i(x,t)\partial_{x_i} \qquad (8)\\
&V(x,T) = g(x).
\end{aligned}$$

and perform the Cole-Hopf transform $V(x,t) = -\lambda \ln \Psi(x,t)$.

Proof. (Path-integral Control).

and perform the Cole-Hopf transform $V(x,t) = -\lambda \ln \Psi(x,t)$.

$$
\begin{aligned}
&-\lambda \frac{\partial_t \Psi + L\Psi}{\Psi}(x,t) + \frac{\lambda^3}{2}\left\|\frac{\sigma^\top \nabla \Psi}{\Psi}(x,t)\right\|^2 - \frac{\lambda^3}{2}\left\|\frac{\sigma^\top \nabla \Psi}{\Psi}(x,t)\right\|^2 + f(x,t) = 0 \\
&L = \frac{\lambda}{2}\sum_{i,j=1}^{d} (\sigma\sigma^\top)_{ij}(t)\partial_{x_i}\partial_{x_j} + \sum_{i=1}^{d} b_i(x,t)\partial_{x_i} \\
&\Psi(x,T) = \exp(-\lambda^{-1} g(x)).
\end{aligned}
\tag{8}
$$

Proof. (Path-integral Control).

After some canceling processes, we have

$$
\begin{aligned}
&\partial_t \Psi(x,t) + L\Psi(x,t) - \lambda^{-1}\Psi(x,t) f(x,t) = 0 \\
&L = \frac{1}{2}\sum_{i,j=1}^{d} (\sigma\sigma^\top)_{ij}(t)\partial_{x_i}\partial_{x_j} + \sum_{i=1}^{d} b_i(x,t)\partial_{x_i} \\
&\Psi(x,T) = \exp(-\lambda^{-1} g(x)).
\end{aligned}
\tag{8}
$$

Then, let us recall the Feynman-Kac formulation:

$$
\begin{cases}
\frac{\partial u(x,t)}{\partial t} + \mu(x,t)\frac{\partial u(x,t)}{\partial x} + \frac{1}{2}\sigma^2(x,t)\frac{\partial^2 u(x,t)}{\partial x^2} - q(x,t)u(x,t) = -g(x,t) \\
u(x,T) = f(x)
\end{cases}
\tag{9}
$$

with its conclusion

Proof of Path-integral Control

Quadratic-regularized State Cost SOC

Proof. (Path-integral Control).

$$u(x,t) = \mathbb{E}\left[f(\xi_T) e^{-\int_t^T q(\theta,\xi_\theta) d\theta} + \int_t^T g(s,\xi_s) e^{-\int_t^T q(\theta,\xi_\theta) d\theta} ds \,|\, \xi_t = x \right] \tag{8}$$

Then, substitute it into the original formula,

$$\Psi(x,t) = \mathbb{E}\left[\exp(-\lambda^{-1} g(x)) \exp(-\lambda^{-1} \int_t^T f(s,X_s) ds) | X_t = x \right] \tag{9}$$

□

Consider the pair of SDEs

$$\begin{aligned} \mathrm{d}X_t &= b(X_t, t)\,\mathrm{d}t + \sqrt{\lambda}\sigma(t)\mathrm{d}B_t, \qquad X_0 \sim p_0, \\ \mathrm{d}Y_t &= (-f(X_t, t) + \frac{1}{2}\|Z_t\|^2)\,\mathrm{d}t + \sqrt{\lambda}\langle Z_t, \mathrm{d}B_t\rangle, \qquad Y_T = g(X_T). \end{aligned} \tag{10}$$

where $Y : \Omega \times [0, T] \to \mathbb{R}$ and $Z : \Omega \times [0, T] \to \mathbb{R}^d$ are progressively measurable random processes. It turns out that Y_t and Z_t defined as $Y_t := V(X_t, t)$ and $Z_t := \sigma(t)^\top \nabla V(X_t, t) = -u^*(X_t, t)$ satisfy the HJB optimality condition.

Definition (Verification Theorem fot Quadratic-regularized State Cost SOC)

The *verification theorem* states that if a function V solves the HJB equation above and has certain regularity conditions, then V is the value function (2) of the problem (5). An implication of the verification theorem is that for every $u \in \mathcal{U}$,

$$V(x,t) + \mathbb{E}\big[\frac{1}{2}\int_t^T \|\sigma^\top \nabla V + u\|^2 (X_s^u, s)\,\mathrm{d}s \,\big|\, X_t^u = x\big] = J(u,x,t). \tag{11}$$

Equation (11) can be deduced by integrating the HJB equation (6) over $[t, T]$, and taking the conditional expectation with respect to $X_t^u = x$.

Proof of Verification Theorem

Quadratic-regularized State Cost SOC

Proof. (Verification Theorem).

By Itô Lemma, we have that

$$V(X_T^u, T) - V(X_t^u, t) = \int_t^T \Big(\partial_s V(X_s^u, s) + \langle b(X_s^u, s) + \sigma(X_s^u, s)u(X_s^u, s), \nabla V(X_s^u, s)\rangle \\ + \frac{\lambda}{2}\sum_{i,j=1}^{d}(\sigma\sigma^\top)_{ij}(X_s^u, s)\partial_{x_i}\partial_{x_j}V(X_s^u, s)\Big)\,\mathrm{d}s + S_t^u, \tag{12}$$

where $S_t^u = \sqrt{\lambda}\int_t^T \nabla V(X_s^u, s)^\top \sigma(X_s^u, s)\,\mathrm{d}B_s$. Note that by (6),

$$\partial_s V(X_s^u, s) + \langle b(X_s^u, s) + \sigma(X_s^u, s)u(X_s^u, s), \nabla V(X_s^u, s)\rangle \\ + \frac{\lambda}{2}\sum_{i,j=1}^{d}(\sigma\sigma^\top)_{ij}(X_s^u, s)\partial_{x_i}\partial_{x_j}V(X_s^u, s) \tag{13}$$

Proof. (Verification Theorem).

$$
\begin{aligned}
&= \frac{1}{2}\|(\sigma^\top \nabla V)(X_s^u, s)\|^2 - f(X_s^u, s) + \langle \sigma(X_s^u, s)u(X_s^u, s), \nabla V(X_s^u, s)\rangle \\
&= \frac{1}{2}\|(\sigma^\top \nabla V)(X_s^u, s) + u(X_s^u, s)\|^2 - \frac{1}{2}\|u(X_s^u, s)\|^2 - f(X_s^u, s),
\end{aligned}
\tag{12}
$$

and this implies that

$$
\begin{aligned}
g(X_T^u) - V(X_t^u, t) = \int_t^T \big(&\frac{1}{2}\|(\sigma^\top \nabla V)(X_s^u, s) + u(X_s^u, s)\|^2 - \\
&\frac{1}{2}\|u(X_s^u, s)\|^2 - f(X_s^u, s)\big)\,\mathrm{d}s + S_t^u
\end{aligned}
\tag{13}
$$

Since $\mathbb{E}[S_t^u \mid X_t^u = x] = 0$, rearranging and taking the conditional expectation with respect to X_t^u yields the final result. □

Reference

- Stochastic Optimal Control Matching
- An optimal control approach to particle filtering
- Stochastic Optimal Control