Adaptive Bayesian Method and its Application to Stochastic Control

Tao Chen* Jiyoun Myung†

Abstract

In this paper... **Key words:**

1 Introduction

In this paper...

2 Bayesian Control Problem

Maybe briefly discuss existing results of parametric Bayesian control in the section.

As mentioned above,

3 Nonparametric Bayesian Control

We extend the work in [1] to the nonparametric setup. Consider a market model consists of a risk-free asset with a constant interest rate r, and a risky asset $\{S_t\}$ with the corresponding return from time t to $t+1$ denoted by $Y_{t+1} = S_{t+1}/S_t$. We assume that $Y_t,\ t = 1, \ldots, T$, is an i.i.d. sequence of random variables. The process $\{Y_t\}$ is observed but its true distribution $\mathbb{P}^*$ is unknown.

The dynamics of the wealth process in the market produced by a self-financing trading strategy is given by

$$V_{t+1} = V_t(1 + r + \varphi_t(Y_{t+1} - r - 1)), \quad t = 0, \ldots, T-1, \tag{3.1}$$

with initial wealth $V_0 = v_0$, and where $\{\varphi_t\} \in \mathcal{U}$ is an adapted process such that φ_t is the proportion of the portfolio wealth invested in the risky asset from time t to $t+1$. That means φ_t, $t = 0, \ldots, T-1$, takes values in the interval $[0, 1]$.

*Department of Statistics and Applied Probability, University of California Santa Barbara, 5519 South Hall Santa Barbara, CA 93106-3110

†Department of Statistics and Applied Probability, University of California Santa Barbara, 5513 South Hall Santa Barbara, CA 93106-3110

Denote by $\mathbb{P}(\mathbb{R})$ and $\mathbb{M}(\mathbb{R})$ the set of probability and finite measures on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, respectively, where $\mathcal{B}(\mathbb{R})$ is the Borel σ-algebra. We consider a process D such that D_t, $t \in \mathcal{T}$, is a random probability measure on the space of probability measures $\mathbb{P}(\mathbb{R})$ which is equipped with the Borel σ-algebra generated by weak convergence. We take D_0 as the Dirichlet process with parameter α and adopt the notations in [4] to write that $D_0 = \mathscr{D}(\alpha)$ where $\alpha \in \mathbb{M}(\mathbb{R})$ with full support. Correspondingly, $D_t = \mathscr{D}(\alpha + \sum_{s=1}^{t} \delta_{Y_s})$, $t = 1, \ldots, T-1$, where δ is the Dirac measure. It is well-known that the support of D_t with respect to weak convergence is the set of all distributions on $\mathbb{R}$ (since α has full support).

In order to proceed, we denote by $C^\alpha = \alpha(\mathbb{R})$ and introduce the following process of probability measures

$$A_t := \frac{\alpha + \sum_{s=1}^{t} \delta_{Y_s}}{C^\alpha + t} = \frac{\alpha + \sum_{s=1}^{t} \delta_{Y_s}}{\left\{\alpha + \sum_{s=1}^{t} \delta_{Y_s}\right\}(\mathbb{R})}, \quad A_0 = \frac{\alpha}{C^\alpha}. \tag{3.2}$$

It is obvious that process $\{A_t\}$ has the dynamics

$$A_t = \frac{(C^\alpha + t - 1)A_{t-1} + \delta_{Y_t}}{C^\alpha + t}, \quad t = 1, \ldots, T,$$

and the process $\{D_t\}$ can be written as

$$D_t = \mathscr{D}((C^\alpha + t)A_t), \quad t = 0, \ldots, T.$$

By adopting a similar idea presented in [3] (need to add more citations here), we consider the augmented state process $X_t = (V_t, A_t)$, $t = 0, \ldots, T$, and the augmented state space

$$E_X = \mathbb{R} \times \mathbb{P}(\mathbb{R}).$$

The process X has the following dynamics,

$$X_{t+1} = \boldsymbol{G}(t+1, X_t, \varphi_t, Y_{t+1}), \quad t = 0, \ldots, T-1,$$

where $\boldsymbol{G}$ is defined as

$$\boldsymbol{G}(t, x, u, y) = \left(v(1 + r + u(y - r - 1)), \frac{(C^\alpha + t - 1)P + \delta_y}{C^\alpha + t} \right), \tag{3.3}$$

where $x = (v, P) \in E_X$.

Lemma 3.1. *The mapping $\boldsymbol{G}(t, \cdot, \cdot, \cdot)$ is continuous.*

Proof. It is enough to show that $R(t, P, y) := \frac{(C^\alpha + t - 1)P + \delta_y}{C^\alpha + t}$, $P \in \mathbb{P}(\mathbb{R})$, $y \in \mathbb{R}$, is continuous with respect to A and y for fixed t.

Assume that $P_n \to P$ weakly and $y_n \to y$ where $P, P_n \in \mathbb{P}(\mathbb{R})$, $y, y_n \in \mathbb{R}$, $n = 1, 2, \ldots$. Take $B \subset \mathbb{R}$ such that

$$\left\{ \frac{(C^\alpha + t - 1)P + \delta_y}{C^\alpha + t} \right\} (\partial B) = 0.$$

Then, B satisfies that $P(\partial B) = 0$ and $y \notin \partial B$. According to Portmanteau theorem, we have $P_n(B) \to P(B)$ and $\delta_{y_n}(B) \to \delta_y(B)$. It is implied that

$$\lim_{n\to\infty} \left\{ \frac{(C^\alpha + t - 1)P_n + \delta_{y_n}}{C^\alpha + t} \right\} (B) = \left\{ \frac{(C^\alpha + t - 1)P + \delta_y}{C^\alpha + t} \right\} (B).$$

Continuity of $R(t, \cdot, \cdot)$ follows immediately. □

Lemma 3.2. *Fix $t = 0, \ldots, T$. Let P and P_n, $n = 1, \ldots,$ be probability measures on $\mathbb{R}$ such that $P_n \to P$ weakly. Then, $\mathscr{D}((C^\alpha + t)P_n) \to \mathscr{D}((C^\alpha + t)P)$ weakly.*

The proof of this lemma follows directly from [5, Theorem 3.2.6] so that we omit it here.

We proceed to define the transition probability for the process $\{X_t\}$, that is, for fixed t, and each $(x, u) \in E_X \times [0, 1]$, we define a probability measure on (the Borel σ-algebra) $\mathcal{E}_X$:

$$Q(B|t, x, u) = \int \mathbb{P}\left(Y_{t+1} \in \{y : \boldsymbol{G}(t + 1, x, u, y) \in B\}\right) \pi(d\mathbb{P}), \quad \pi \in \mathscr{D}((C^\alpha + t)P),$$

for any $B \in \mathcal{E}_X$. It is clear that the above definition and be rewritten as

$$Q(B|t, x, u) = \widetilde{\mathbb{P}}^\pi \left(Y_{t+1} \in \{y : \boldsymbol{G}(t + 1, x, u, y) \in B\}\right), \quad \pi \in \mathscr{D}((C^\alpha + t)P), \tag{3.4}$$

with $\widetilde{\mathbb{P}}^\pi$ being the Bayesian estimator of $\mathbb{P}^*$. Then, for fixed $\left(v_0, \frac{\alpha}{C^\alpha}\right) \in E_X$ and every control process $\{\varphi_t\} \in \mathcal{U}$, we define the probability measure $\mathbb{Q}^\varphi$ on the canonical space E_X^{T+1}:

$$\mathbb{Q}^\varphi(B_0, B_1, \ldots, B_T) = \int_{B_0} \int_{B_1} \cdots \int_{B_T} \prod_{t=1}^{T} Q(dx_t|t - 1, x_{t-1}, \varphi_{t-1}(h_{t-1}))\delta_{\left(v_0, \frac{\alpha}{C^\alpha}\right)}(dx_0),$$

where $\{h_t\}$ is a realization of the history process $H_t := (X_0, X_1, \ldots, X_t)$:

$$h_t = (x_0, x_1, \ldots, x_t) = (v_0, P_0, v_1, P_1, \ldots, v_t, P_t).$$

Note that $h_0 = x_0 = \left(v_0, \frac{\alpha}{C^\alpha}\right)$.

Remark 3.3. The measure $\mathbb{Q}^\varphi$ in fact depends on (v_0, α). However, we fix such pair throughout and for simplicity of notations we write $\mathbb{Q}^\varphi$ instead of $\mathbb{Q}^\varphi_{v_0,\alpha}$.

We are now ready to define the *nonparametric Bayesian control* problem:

$$\sup_{\{\varphi_t\}\in\mathcal{U}} \mathbb{E}_{\mathbb{Q}^\varphi}[\ell(V_T)]. \tag{3.5}$$

4 Solution to Bayesian Control Problem

The main result in this section is to show that problem (3.5) is solved by the following Bellman equations:

$$\begin{aligned} W_T(x) &= \ell(v), \\ W_t(x) &= \inf_{u\in[0,1]} \mathbb{E}_{\widetilde{\mathbb{P}}^\pi}\left[W_{t+1}(\boldsymbol{G}(t + 1, x, u, Y_{t+1}))\right], \quad \pi \in \mathscr{D}((\alpha(\mathbb{R}) + t)P), \end{aligned} \tag{4.1}$$

for $t = T-1, \ldots, 1, 0$. In addition, we prove that optimal selectors exist for equations (4.1). In other words, the optimal control φ_t^* at time t, $t = 0, 1, \ldots, T-1$, is a measurable function of $x = (v, P) \in E_X$:

$$W_t(x) = \mathbb{E}_{\widetilde{\mathbb{P}}^\pi}\left[W_{t+1}(\boldsymbol{G}(t+1, x, \varphi_t^*(x), Y_{t+1}))\right].$$

To proceed, we postulate that the loss function ℓ is continuous and bounded. Then, we have the following result.

Proposition 4.1. *The functions W_t, $t = T, T-1, \ldots, 0$, are lower semi-continuous (l.s.c.), and the optimal selectors φ_t^*, $t = T, T-1, \ldots, 0$, in* (4.1) *exist.*

Proof. By assumption and Lemma 3.1, $W_T(\boldsymbol{G}(T, \cdot, \cdot, y))$ is continuous. Then, the function

$$w_{T-1}(x, u) = \mathbb{E}_{\widetilde{\mathbb{P}}^\pi}\left[W_T((T, x, u, Y_T))\right]$$

is continuous and hence l.s.c..

By adopting the notations of [2, Proposition 7.33], we let

$$\begin{aligned} \mathrm{X} &= E_X = \mathbb{R} \times \mathbb{P}(\mathbb{R}), \quad \mathrm{x} = (v, P), \\ \mathrm{Y} &= [0,1], \quad \mathrm{y} = u, \\ \mathrm{D} &= E_X \times [0,1], \\ f(\mathrm{x}, \mathrm{y}) &= -w_{T-1}(v, P, u). \end{aligned}$$

It is clear that X is metrizable, Y is compact, and D is closed. It is also trivial to verify that $\mathrm{proj}_{\mathrm{X}}(D) = E_X$ and $\mathrm{D}_x = [0,1]$ for any $x \in E_X$. Then, according to [2, Proposition 7.33],

$$W_{T-1}(x) = \inf_{u \in [0,1]} w_{T-1}(x, u)$$

is l.s.c., and there exists a Borel-measurable function $\varphi_{T-1}^* : E_X \to [0,1]$ such that

$$W_{T-1}(x) = w_{T-1}(x, \varphi_{T-1}^*(x)).$$

The rest of the proof follows analogously. □

In order to present the main result of the section, we first introduce some useful notations. Let

$$\mathbb{H}_t = \underbrace{E_X \times E_X \times \cdots \times E_X}_{\text{t+1 times}}$$

be the space from which H_t takes its values. We define the Borel probability measure $\mathbb{Q}_{h_t}^{\varphi^t}$ on $\mathbb{H}_t$:

$$\mathbb{Q}_{h_t}^{\varphi^t}(B_{t+1}, \ldots, B_T) = \int_{B_{t+1}} \int_{B_{t+2}} \cdots \int_{B_T} \prod_{s=t+1}^{T} Q(dx_s | s-1, x_{s-1}, \varphi_{s-1}(h_{s-1})),$$

where $\varphi^t := (\varphi_t, \varphi_{t+1}, \ldots, \varphi_{T-1})$.

Proposition 4.2. *The process $\{\varphi_t^*\}$ constructed from the selectors in Proposition 4.1 is the solution of the nonparametric Bayesian control problem* (3.5)*:*

$$\inf_{\{\varphi_t\}\in\mathcal{U}} \mathbb{E}_{\mathbb{Q}^\varphi}[\ell(V_T)] = \mathbb{E}_{\mathbb{Q}^\varphi}[\ell(V_T(\varphi^*))]. \tag{4.2}$$

Proof. We prove the result via backward induction in $t = T, T-1, \ldots, 1, 0$:

$$W_t(x_t) = \inf_{\varphi^t\in[0,1]^{T-t}} \mathbb{E}_{\mathbb{Q}_{h_t}^{\varphi^t}}[\ell(V_T)].$$

Take $t = T$, and it is clear that $U_T(h_T) = W_T(x_T)$. For $t = T-1$ we have

$$\begin{aligned}
\inf_{\varphi^{T-1}\in[0,1]} \mathbb{E}_{\mathbb{Q}_{h_{T-1}}^{\varphi^{T-1}}}[\ell(V_T)] &= \inf_{\varphi_{T-1}\in[0,1]} \mathbb{E}_{\mathbb{Q}_{h_{T-1}}^{\varphi^{T-1}}}[\ell(V_T)] \\
&= \inf_{u\in[0,1]} \mathbb{E}_{\widetilde{\mathbb{P}}^\pi}[\ell(\boldsymbol{G}(T, x_{T-1}, u, Y_T))] \\
&= \inf_{u\in[0,1]} \mathbb{E}_{\widetilde{\mathbb{P}}^\pi}[W_T(\boldsymbol{G}(T, x_{T-1}, u, Y_T))] = W_{T-1}(x_{T-1}).
\end{aligned}$$

For $t = T-1, \ldots, 1, 0$, we have by induction

$$\begin{aligned}
\inf_{\varphi^t\in[0,1]^{T-t}} \mathbb{E}_{\mathbb{Q}_{h_t}^{\varphi^t}}[\ell(V_T)] &= \inf_{\varphi^t\in[0,1]^{T-t}} \mathbb{E}_{\mathbb{Q}_{h_t}^{\varphi^t}}[\ell(V_T)] \\
&= \inf_{\varphi^t=(\varphi_t,\varphi^{t+1})\in[0,1]^{T-t}} \int_{E_X} \mathbb{E}_{\mathbb{Q}_{h_{t+1}}^{\varphi^{t+1}}}[\ell(V_T)] Q(dx_{t+1}|t, x_t, \varphi_t(h_t)) \\
&\geq \inf_{\varphi^t=(\varphi_t,\varphi^{t+1})\in[0,1]^{T-t}} \int_{E_X} U_{t+1}(h_{t+1}) Q(dx_{t+1}|t, x_t, \varphi_t(h_t)) \\
&= \inf_{\varphi_t\in[0,1]} \int_{E_X} W_{t+1}(x_{t+1}) Q(dx_{t+1}|t, x_t, \varphi_t(h_t)) = W_t(x_t).
\end{aligned}$$

Next, fix $\varepsilon > 0$, and let $\varphi^{t+1,\varepsilon}$ be the ε-optimal control process at time $t+1$, namely,

$$\mathbb{E}_{\mathbb{Q}_{h_{t+1}}^{\varphi^{t+1,\varepsilon}}}[\ell(V_T)] \leq \inf_{\varphi^{t+1}\in[0,1]^{T-t-1}} \mathbb{E}_{\mathbb{Q}_{h_{t+1}}^{\varphi^{t+1}}}[\ell(V_T)] + \varepsilon.$$

Then, we have that

$$\begin{aligned}
\inf_{\varphi^t\in[0,1]^{T-t}} \mathbb{E}_{\mathbb{Q}_{h_t}^{\varphi^t}}[\ell(V_T)] &= \inf_{\varphi^t=(\varphi_t,\varphi^{t+1})\in[0,1]^{T-t}} \int_{E_X} \mathbb{E}_{\mathbb{Q}_{h_{t+1}}^{\varphi^{t+1}}}[\ell(V_T)] Q(dx_{t+1}|t, x_t, \varphi_t(h_t)) \\
&\leq \inf_{\varphi_t\in[0,1]} \int_{E_X} \mathbb{E}_{\mathbb{Q}_{h_{t+1}}^{\varphi^{t+1,\varepsilon}}}[\ell(V_T)] Q(dx_{t+1}|t, x_t, \varphi_t(h_t)) \\
&\leq \inf_{\varphi_t\in[0,1]} \int_{E_X} U_{t+1}(h_{t+1}) Q(dx_{t+1}|t, x_t, \varphi_t(h_t)) + \varepsilon \\
&= \inf_{\varphi_t\in[0,1]} \int_{E_X} W_{t+1}(x_{t+1}) Q(dx_{t+1}|t, x_t, \varphi_t(h_t)) + \varepsilon = W_t(x_t) + \varepsilon.
\end{aligned}$$

Because ε is arbitrary, we conclude that $W_t(x_t) = \inf_{\varphi^t\in[0,1]^{T-t}} \mathbb{E}_{\mathbb{Q}_{h_t}^{\varphi^t}}[\ell(V_T)]$ for all $t = T, T-1, \ldots, 0$. Let $t = 0$, we therefore get (4.2). $\square$

5 Numerical Example

In this section we will use an example of dynamic optimal portfolio selection to illustrate our method. To this end, we take the CRRA utility function $U(x) = \frac{x^{1-\gamma}}{1-\gamma}$, for $x > 0$, and some $\gamma \neq 1$. Then, we attempt to solve the following

References

[1] N. Bäuerle and U. Rieder, *Markov Decision Processes with Applications to Finance*, Universitext, Springer-Verlag Berlin Heidelberg, 2011.

[2] D. P. Bertsekas and S. Shreve, *Stochastic Optimal Control: The Discrete-Time Case*, Academic Press, 1978.

[3] T. Bielecki, T. Chen, I. Cialenco, A. Cousin, and J. M., *Adaptive robust control under model uncertainty*, SIAM J. Control Optim., 57 (2019).

[4] T. Ferguson, *Prior distributions on spaces of probability measures*, The Annals of Statistics, 2 (1974), pp. 615–629.

[5] J. Ghosh and R. Rmamoorthi, *Bayesian Nonparametrics*, Springer Series in Statistics, Springer-Verlag, 2003.