Automatica 94 (2018) 249–257

Contents lists available at ScienceDirect

Automatica

journal homepage: www.elsevier.com/locate/automatica

Brief paper

Robust time-inconsistent stochastic control problems[☆]

Chi Seng Pun

School of Physical and Mathematical Sciences, Nanyang Technological University, Singapore

ARTICLE INFO

Article history:
Received 11 September 2017
Received in revised form 15 January 2018
Accepted 20 March 2018
Available online 5 May 2018

Keywords:
Stochastic control
Robustness
Time-inconsistent preference
State-dependence
Hamilton–Jacobi–Bellman–Isaacs equations
Robust dynamic mean–variance portfolio

ABSTRACT

This paper establishes a general analytical framework for continuous-time stochastic control problems for an ambiguity-averse agent (AAA) with time-inconsistent preference, where the control problems do not satisfy Bellman's principle of optimality. The AAA is concerned about model uncertainty in the sense that she is not completely confident in the reference model of the controlled Markov state process and rather considers some similar alternative models. The problems of interest are studied within a set of dominated models and the AAA seeks for an optimal decision that is robust with respect to model risks. We adopt a game-theoretic framework and the concept of subgame perfect Nash equilibrium to derive an extended dynamic programming equation and extended Hamilton–Jacobi–Bellman–Isaacs equations for characterizing the robust dynamically optimal control of the problem. We also prove a verification theorem to theoretically support our construction of robust control. To illustrate the tractability of the proposed framework, we study an example of robust dynamic mean–variance portfolio selection under two cases: 1. constant risk aversion; and 2. state-dependent risk aversion.

1. Introduction

Stochastic control theory has achieved great success in modeling and providing solutions to lots of physical, biological, economical, and financial problems, to name a few. Stochastic optimal control is the serial control variables that accomplish a desired goal for the controlled state process with minimum cost or with maximum reward in the presence of noises (risks). The most commonly used approaches in solving stochastic optimal control problems are Pontryagin's maximum principle and Bellman's dynamic programming. These two principal approaches and their relationship are documented in many classic reference books such as Yong and Zhou (1999).

While stochastic control deals with the existence of risk, robust stochastic control deals with the existence of ambiguity as well. Here, the ambiguity refers to the Knightian (model) uncertainty originating with Knight (1921), who clarified the subtle difference between risk and uncertainty. Ellsberg (1961) reveals the inadequacy of utility theory and argues that human beings are ambiguity-averse; thus our rational decisions should be made under conditions of ambiguity. Our lack of knowledge about the actual state process or estimation error unavoidably introduces ambiguity into the control problem and it has important implications for many critical aspects such as risk quantification. We call the agent fearing ambiguity as *ambiguity-averse agent (AAA)*. The AAA has certain confidence in a reference model but rather considers some alternative models. A key to dealing with ambiguity is to quantify the model misspecification given the AAA's historical data record; see Anderson, Hansen, and Sargent (2003) for a statistical method. Following the robust decision rule in Anderson et al. (2003), Maenhout (2004) and Wald (1945) derives robust portfolio rules in the context of Merton (1971)'s portfolios. Recently, Pun and Wong (2015) extend the analysis to a general time-consistent objective functional.

In recent years, there is a growing literature investigating time-inconsistent stochastic control problems, where the objective functional contains time-inconsistent terms such that Pontryagin's and Bellman's optimality principles are not applicable. A famous example is the financial mean–variance (MV) portfolio selection, pioneered by Markowitz (1952) and further studied in Li and Ng (2000) and Zhou and Li (2000) for dynamic settings. Some other examples include endogenous habit formulation in economics and equilibrium production economy; see Björk, Khapko, and Murgoci (2017) and Björk and Murgoci (2014). For a general time-inconsistent objective functional (see Eq. (4)), Lemma 2 below reveals the sources of time-inconsistency that violate Bellman's principle of optimality. In this paper, we contribute to incorporate the model uncertainty with time-inconsistency to study a general

☆ Chi Seng Pun gratefully acknowledges Nanyang Technological University, Start-Up Grant (Reference No: M4081840.110) and Ministry of Education (MOE), AcRF Tier 2 grant (Reference No: MOE2017-T2-1-044) for the funding of this research. The material in this paper was not presented at any conference. This paper was recommended for publication in revised form by Associate Editor Michael V. Basin under the direction of Editor Ian R. Petersen.

E-mail address: cspun@ntu.edu.sg.

https://doi.org/10.1016/j.automatica.2018.04.038
0005-1098/

class of robust time-inconsistent stochastic control problems. To the best of our knowledge, we are the first to establish a general, analytical, and tractable framework for such problems. Our model and objective settings nest many classes of well-known continuous-time problems, such as Heston (1993), Merton (1971) and Zhou and Li (2000), as special cases.

In the existing literature, several approaches to handling the time-inconsistency derive reasonable policies with different features, which include but are not limited to the followings:

- *Precommitment* policy: that optimizes the objective functional anticipated at the very beginning time point and the controller sticks with this policy over the whole control period. For example, Li and Ng (2000) and Zhou and Li (2000) introduced an embedding technique to solve for a precommitment MV portfolio.
- *Equilibrium (time-consistent)* policy: that consistently optimizes the objective functional anticipated at every time point in the similar manner of dynamic programming but using the concept of subgame perfect equilibrium. This idea is initiated in Goldman (1980) and Strotz (1955). Related papers include Basak and Chabakauri (2010) for time-consistent MV portfolio and Björk et al. (2017), Björk and Murgoci (2014) and Björk, Murgoci, and Zhou (2014) for more concrete examples.

Some recent alternatives are proposed in Cui, Li, and Shi (2017), Karnam, Ma, and Zhang (2017) and Pedersen and Peskir (2017) with different views on dynamic objectives. Loosely speaking, from the perspective of decision making, precommitment policy emphasizes on global optimality while time-consistent policy emphasizes on local optimality. However, precommitment policy is sometimes inferior because its strong commitment leads to time-inconsistency in efficiency (see Cui, Li, Wang, & Zhu, 2010 for details and a remedy) and its error-accumulation property brings huge estimation error (see Chiu, Pun, & Wong, 2017). Moreover, finding precommitment policy poses analytical challenges for general time-inconsistent stochastic control problems. In this paper, we attack the time-inconsistency with the second approach, together with robustness, resulting in robust time-consistent policy. A related work is Zeng, Li, and Gu (2016), which considered similar mean–variance optimization and it will be compared with this paper in Section 2.

In this paper, we assume a general continuous-time Markov stochastic process for the state under the reference model. Moreover, we define alternative models, which are equivalent to the reference model in terms of probability measure. By the Girsanov theorem, the link between the reference and alternative models is characterized by a stochastic process that acts as an adverse control variate. The agent has a time-inconsistent preference of a general form and aims to find a time-consistent policy that is robust with respect to the choice of the alternative model. We use a maximin formulation as in Wald (1945) to construct robust decision rule and use the concept of subgame perfect equilibrium to characterize the time-consistent policy. With an extended dynamic programming approach derived in this paper, we characterize the robust time-consistent policy using an extended Hamilton–Jacobi–Bellman–Isaacs (HJBI) system. Moreover, we introduce and discuss the choice of ambiguity preference function, which completes a general analytical framework for a large class of time-inconsistent stochastic control problems with model uncertainty. To illustrate the tractability of the proposed framework, we apply it to solve for robust dynamic mean–variance (MV) portfolio selection under two cases: 1. constant risk aversion; and 2. state-dependent risk aversion. For the latter case, we introduce nonhomogeneous Abel's differential equations to characterize the robust MV portfolio.

The contribution of this paper is threefold: first, we provide a rigorous mathematical definition of robust time-consistent policy and reveal its nature as the perfect equilibrium of subgames of maximin control problems, i.e. "games in subgames". Second, we prove a verification theorem that solving the proposed extended HJBI system is a sufficient condition for robust optimality. The extended dynamic programming approach and the verification theorem, derived in this paper, are innovative to the literature on robust control. Third, we apply the proposed framework to solve an open problem of robust dynamic mean–variance portfolio selection under the robustness rule of Anderson et al. (2003). The analyses cover two economically meaningful cases, which extend the results in Björk et al. (2014) to robust counterparts. Through this study, our discussion about the ambiguity preference function for the general case extends the results in Maenhout (2004) and Pun and Wong (2015).

The remainder of this paper is organized as follows. Section 2 introduces the reference and alternative models and the robust time-inconsistent stochastic control problems of our interest. The main results are in Section 3, where we present an extended dynamic programming equation for the robust value function and the extended HJBI system with the verification theorem while a generalized result and the proofs are documented in the extended version of this paper (Pun, 2018) due to the page limit. To facilitate the analysis, Section 3.3 provides a simplification of the HJBI system to an HJB system. Section 4 is devoted to robust MV portfolio analysis and the discussion of the choice of ambiguity preference function for general problems. Finally, Section 5 concludes and discusses the future research directions.

2. Problem formulation

We study the problem of our interest with a set of candidate models, which stem from a reference model with agent's preliminary knowledge. Specifically, we suppose the agent does not have complete confidence in the reference model, and she prefers to consider alternative models perturbed around the reference model and to make a robust decision with respect to the model risk.

2.1. The reference and alternative models

The *reference model* is defined over a filtered physical probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t^{\mathbb{P}}\}_{t\geq 0}, \mathbb{P})$, where the filtration $\{\mathcal{F}_t^{\mathbb{P}}\}_{t\geq 0}$ is generated by an m-dimensional standard $\mathbb{P}$-Brownian motion $W_t^{\mathbb{P}} = (W_{1t}^{\mathbb{P}}, \ldots, W_{mt}^{\mathbb{P}})'$. Hereafter, the transpose of a vector or matrix a is denoted by a'. We consider the p-dimensional controlled state process within a time horizon T, $\{X_t\}_{t\in[0,T]}$, driven by a stochastic differential equation (SDE)

$$dX_t = \mu(t, X_t, u_t)dt + \sigma(t, X_t, u_t)dW_t^{\mathbb{P}}, \tag{1}$$

where $\{u_t\}_{t\in[0,T]}$ is a k-dimensional control process with the constraint $u_t \in U(t, X_t)$ and $\mu \in \mathbb{R}^p$ and $\sigma \in \mathbb{R}^{p\times m}$ are the drift and diffusion coefficient functions in $(t, X_t, u_t) \in [0, T] \times \mathbb{R}^p \times U(t, X_t)$. The model (1) is general as m, p, k can be arbitrary natural numbers and it nests the stochastic volatility models (Heston, 1993) in finance and stochastic multi-factor models (Pun, Chung, & Wong, 2015) as special cases.

The *alternative models* are conceptually defined as the models that are "similar" to the reference model. In this paper, we employ the mathematical concept of measure equivalence to characterize the "similarity" between models; see Anderson et al. (2003). Specifically, the alternative measures (models) are induced by a class of probability measures equivalent to $\mathbb{P}$: $\mathcal{Q} := \{\mathbb{Q}|\mathbb{Q} \sim \mathbb{P}\}$. By the Girsanov theorem, for each $\mathbb{Q} \in \mathcal{Q}$, there is a $\mathbb{R}^m$-valued