Systems & Control Letters 72 (2014) 36–43

Contents lists available at ScienceDirect

Systems & Control Letters

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

Explicit solution of relative entropy weighted control

Joris Bierkens *, Hilbert J. Kappen

Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen, The Netherlands

ARTICLE INFO

Article history:
Received 22 November 2012
Received in revised form
30 July 2014
Accepted 7 August 2014
Available online 8 September 2014

Keywords:
Stochastic optimal control
Itô calculus
Brownian motion
Malliavin calculus
Relative entropy
Path integral control
Monte-Carlo sampling

ABSTRACT

We consider the minimization over probability measures of the expected value of a random variable, regularized by relative entropy with respect to a given probability distribution. In the general setting we provide a complete characterization of the situations in which a finite optimal value exists and the situations in which a minimizing probability distribution exists. Specializing to the case where the underlying probability distribution is Wiener measure, we characterize finite relative entropy changes of measure in terms of square integrability of the corresponding change of drift. For the optimal change of measure for the relative entropy weighted optimization, an expression involving the Malliavin derivative of the cost random variable is derived. The theory is illustrated by its application to several examples, including the case where the cost variable is the maximum of a standard Brownian motion over a finite time horizon. For this example we obtain an exact optimal drift, as well as an approximation of the optimal drift through a Monte-Carlo algorithm.

1. Introduction

In certain situations in stochastic optimal control theory, the dynamic programming or Hamilton–Jacobi–Bellman equations may be transformed, through the Hopf-transform, into linear equations [1], [2, Chapter VI]. In the past years, within the applied control and machine learning community, there has been a significant amount of interest in this class of 'path integral control problems' (see e.g. [3–5]). This class of problems also occurs in risk sensitive control theory (see [2]) and the theory of large deviations (see [6]), and it occurs in modified form (constrained to equivalent martingale measures) in mathematical finance, in particular as the dual problem for a portfolio optimization problem [7]. It is the goal of this paper to review and extend the mathematical underpinning of this optimization problem, as well as showcase some new results within this context.

The problem we consider is a minimization problem over probability measures that are absolutely continuous with respect to a given probability measure (referred to as the 'uncontrolled measure'). The functional we wish to minimize is the sum of (i) the expectation of a given random variable with respect to any probability measure, and (ii) the relative entropy of that probability measure with respect to the uncontrolled measure. The density of the optimal probability measure with respect to the uncontrolled distribution is readily available through an explicit expression in terms of the cost random variable. The challenge is then to understand this probability measure within the context of the underlying problem.

In particular, in the special case in which we are interested in this paper, the uncontrolled distribution is Wiener measure on the space of continuous sample paths. Absolutely continuous change of distribution then corresponds, by the Girsanov theorem, to a change of drift, which we will interpret as control process. The regularizing relative entropy corresponds to squared control cost. The questions we wish to answer in this paper are: (i) under what conditions does there exist an optimal change of drift corresponding to a given cost functional and probability measure, and (ii) how can it be computed?

There is a close relation to existing theory within the field of large deviations theory and stochastic optimal control [6,8]. We should also mention the work of Föllmer [9,10]. For an excellent self-contained review of these results, see [11]. The main aim of this paper is to review the application of the mentioned results in a mathematical control context. We also illustrate the use of Monte-Carlo methods; for recent work on the use of Monte-Carlo methods in relative entropy weighted control, see [12].

Furthermore, for the reader who is familiar with the available literature, theoretical contributions of our paper include:

(i) Relaxing the usual boundedness assumptions to a condition that guarantees finite relative entropy of the optimal change of measure (condition (FE) of Section 2);

* Corresponding author. Tel.: +31 24 36 14 230.
E-mail addresses: j.bierkens.04@cantab.net, j.bierkens@science.ru.nl (J. Bierkens), b.kappen@science.ru.nl (H.J. Kappen).

http://dx.doi.org/10.1016/j.sysconle.2014.08.001
0167-6911/

(ii) Solution of the problem where the cost random variable is the maximum of a standard Brownian motion with controlled drift over a finite time horizon.

1.1. Outline

In Section 2, we consider the general relative entropy weighted optimization problem, and completely characterize the different situations that may arise. The situation with finite relative entropy is most useful, and problems for which the optimal change of measure has finite relative entropy are easily characterized in terms of conditions on the cost functional and the probability measure. Then in Section 3, the finite relative entropy case is further investigated within the context of a Wiener process. It is shown that a change of measure with finite relative entropy corresponds to a square integrable drift, which in particular is the case for the optimal density. In Section 3.2 we show how the optimal drift may be computed through the Clark–Ocone formula. To illustrate the use of this approach, and as an interesting result in its own right, we compute the optimal drift for the case where the cost functional is the maximum of a one dimensional Wiener process with controlled drift on a finite time horizon (Section 3.3). We also provide a Monte-Carlo algorithm for the approximation of such a solution, which is easily extended to other problems.

1.2. Notation

As is common in probability theory, we will allow random variables to assume their values within the extended reals $[-\infty, \infty]$. Resulting formal expressions may be interpreted as follows: $\log 0 = -\infty$, $\log\infty = \infty$, $\exp(-\infty) = 0$, $\exp(\infty) = \infty$ and $\infty \exp(-\infty) = 0$. For any $a \in \mathbb{R}$, we write $(a)^+ := a \vee 0$ and $(a)^- := -a \vee 0$ for the positive and negative parts of a. The euclidean norm of $x \in \mathbb{R}^d$ is denoted by $|x|$. For an adapted process θ and a continuous local martingale M, both with values in $\mathbb{R}^d$, we write $\int_0^t \langle\theta_s, dM_s\rangle$ to indicate $\sum_{i=1}^d \int_0^t \theta_s^i \, dM_s^i$. If $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space we write $\mathbb{E}^{\mathbb{P}}$ for expectation with respect to the probability measure $\mathbb{P}$. Lebesgue measure will be denoted by Leb.

2. Relative entropy weighted optimization

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. The probability measure $\mathbb{P}$ will be referred to as the *uncontrolled (probability) measure*. Let C be a random variable assuming values in $[-\infty, \infty]$. The random variable C indicates a cost we wish to minimize, as explained below.

Let $\mathcal{P}$ denote the set of probability measures on $(\Omega, \mathcal{F})$. We wish to find a probability measure $\mathbb{Q} \in \mathcal{P}$ that

(i) is absolutely continuous with respect to $\mathbb{P}$ (denoted by $\mathbb{Q} \ll \mathbb{P}$),
(ii) reduces the expected cost $\mathbb{E}^{\mathbb{Q}}C$, but
(iii) has small deviation from $\mathbb{P}$. We take the relative entropy

$$\mathcal{H}(\mathbb{Q};\mathbb{P}) = \int_\Omega \log\left(\frac{d\mathbb{Q}}{d\mathbb{P}}\right) d\mathbb{Q} = \mathbb{E}^{\mathbb{Q}}\left[\log\left(\frac{d\mathbb{Q}}{d\mathbb{P}}\right)\right]$$

as a measure of this deviation (see e.g. [8, Section 1.4] for preliminary results on relative entropy). Recall that $\mathcal{H}(\mathbb{Q};\mathbb{P}) \geq 0$ for any $\mathbb{Q}, \mathbb{P} \in \mathcal{P}$, and $\mathcal{H}(\mathbb{Q};\mathbb{P}) = 0$ if and only if $\mathbb{Q} = \mathbb{P}$.

Note that (i) is a constraint and (ii) and (iii) are conflicting optimization targets.

Let $\mathcal{P}_0 := \left\{\mathbb{Q} \in \mathcal{P} : \mathbb{E}^{\mathbb{Q}}[(C)^+] < \infty \text{ and } \mathcal{H}(\mathbb{Q};\mathbb{P}) < \infty\right\}$ denote the set of admissible probability measures, and note that $\mathcal{P}_0$ is convex. Define the cost functional $J : \mathcal{P} \to \mathbb{R}$ by

$$J(\mathbb{Q}) := \begin{cases} \mathbb{E}^{\mathbb{Q}}C + \mathcal{H}(\mathbb{Q};\mathbb{P}) \\ = \mathbb{E}^{\mathbb{Q}}\left[C + \log\left(\frac{d\mathbb{Q}}{d\mathbb{P}}\right)\right] & \text{if } \mathbb{Q} \in \mathcal{P}_0, \\ \infty & \text{otherwise.} \end{cases} \tag{1}$$

The definition of $\mathcal{P}_0$ is as non-restrictive as possible, so that for $\mathbb{Q} \in \mathcal{P}_0$ the value $J(\mathbb{Q})$ is well defined within the interval $[-\infty, \infty)$.

We arrive at the following problem:

Problem 2.1 (*Relative Entropy Weighted Optimization*)**.** Compute $J^\star = \inf_{\mathbb{Q}\in\mathcal{P}_0} J(\mathbb{Q})$, and if it exists, a minimizer $\mathbb{Q}^\star \in \mathcal{P}_0$ such that $J(\mathbb{Q}^\star) = J^\star$.

The solution of this problem is well known for the case in which $\mathbb{P}(|C| < K) = 1$ for some $K > 0$, see e.g. [8, Proposition 1.4.2] or [6, Proposition 2.5]. The purpose of this section is to provide a complete characterization of the existence of solutions of Problem 2.1 in terms of conditions on $\mathbb{P}$ and C. To achieve this goal, we will consider the following further conditions on C and $\mathbb{P}$.

$$\begin{aligned} &\textit{finite (relative) entropy}: \ \mathbb{P}(C < +\infty) > 0 \quad \text{and} \\ &\quad \mathbb{E}^{\mathbb{P}}[\exp(-C)|C|] < \infty. \end{aligned} \tag{FE}$$

$$\textit{integrability}: \ 0 < \mathbb{E}^{\mathbb{P}}[\exp(-C)] < \infty. \tag{I}$$

Condition (FE) will earn its name ('finite relative entropy') below: as we will see, it is a necessary and sufficient condition for the 'optimal' probability distribution to have finite relative entropy with respect to $\mathbb{P}$. The implication

(FE) $\implies$ (I)

holds as a result of the estimates $\exp(-x)\mathbb{1}_{\{x\leq-1\}} \leq \exp(-x)|x|$ and $\exp(-x)\mathbb{1}_{\{x>-1\}} \leq \frac{1}{e}$.

Example 2.2. The following examples may serve to illustrate the conditions (FE) and (I).

(i) $\Omega = [0, \infty)$, with $\mathbb{P}$ having density $f(\omega) = \exp(-\omega)$ with respect to Lebesgue measure; $C(\omega) = -\omega : \mathbb{P}(C < +\infty) > 0$ but (I) does not hold.
(ii) $\Omega = (-\infty, \infty)$, $\frac{d\mathbb{P}}{d\mathrm{Leb}}(\omega) = \frac{\exp(-|\omega|)}{k(1+\omega^2)}$, with k a normalization constant, and $C(\omega) = -|\omega|$. Then $\mathbb{P}$ is a probability distribution,

$$\mathbb{E}^{\mathbb{P}}[\exp(-C)] = \frac{1}{k}\int_{-\infty}^{\infty} \frac{1}{1+x^2}\,dx = \frac{\pi}{k},$$

and $\mathbb{E}^{\mathbb{P}}[|C|\exp(-C)] = \infty$. So (I) holds but (FE) does not.

We have the following observation.

Lemma 2.3. *If $\mathbb{P}(C < +\infty) > 0$ then $\mathcal{P}_0$ is non-empty.*

Proof. Under the assumption, there exists an $M > 0$ such that $\mathbb{P}(C < M) > 0$. Let $Z = \mathbb{1}_{\{C\leq M\}}/\mathbb{P}(C \leq M)$. Then $\mathbb{E}^{\mathbb{P}}Z = 1$, so that $d\mathbb{Q}/d\mathbb{P} = Z$ defines a valid probability measure. Furthermore $\mathbb{E}^{\mathbb{Q}}(C)^+ \leq M < \infty$ and $\mathbb{E}^{\mathbb{Q}}[|\log(Z)|] < \infty$. We may conclude that $\mathbb{Q} \in \mathcal{P}_0$. □

If (I) holds, then

$$\frac{d\mathbb{Q}^\star}{d\mathbb{P}} = Z^\star := \frac{\exp(-C)}{\mathbb{E}^{\mathbb{P}}[\exp(-C)]} \tag{2}$$

defines a probability measure $\mathbb{Q}^\star$ that is absolutely continuous with respect to $\mathbb{P}$.