C. R. Acad. Sci. Paris, Ser. I 352 (2014) 707–713

Contents lists available at ScienceDirect

C. R. Acad. Sci. Paris, Ser. I

www.sciencedirect.com

Optimal control/Calculus of variations

Dynamic programming for mean-field type control

Programmation dynamique pour les problèmes de contrôle à champs moyen

Mathieu Laurière, Olivier Pironneau

LJLL, Université Pierre-et-Marie-Curie (Paris-6), 4, place Jussieu, 75005 Paris, France

ARTICLE INFO

Article history:
Received 13 April 2014
Accepted 17 July 2014
Available online 6 September 2014

Presented by Alain Bensoussan

ABSTRACT

For mean-field type control problems, stochastic dynamic programming requires adaptation. We propose to reformulate the problem as a distributed control problem by assuming that the PDF ρ of the stochastic process exists. Then we show that Bellman's principle applies to the dynamic programming value function $V(\tau,\rho_\tau)$, where the dependency on ρ_τ is functional as in P.-L. Lions' analysis of mean-field games (2007) [10]. We derive HJB equations and apply them to two examples, a portfolio optimization and a systemic risk model.

RÉSUMÉ

Pour les problèmes de contrôle stochastique à champs moyen, la programmation dynamique ne s'applique pas sans adaptation ; mais si l'on reformule le problème avec l'équation de Fokker–Planck, on peut le faire en utilisant une fonctionnelle valeur $\{\tau,\rho_\tau(\cdot)\}\to V(\tau,\rho_\tau)$ comme dans l'analyse des problèmes de jeux à champs moyen par P.-L. Lions (2007) [10]. Les résultats sont appliqués à un problème d'optimisation de portefeuille et à un problème de risque systémique.

1. Introduction

Stochastic control is an old topic [5,12,14,15], which benefits from a renewed interest in economy and finance due to mean-field games [6,7,9,13]. They lead, among other things, to stochastic control problems which involve statistics of the Markov process like means and variance. Optimality conditions for these are derived either by stochastic calculus of variation [1] or by stochastic dynamic programming in the quadratic case [2,3], but not in the general case for the fundamental reason that Bellman's principle does not apply in its original form on the stochastic trajectories of say X_t if those depend upon statistics of X_t like its mean value. As noticed earlier in [10] and in [4],[1] there seems to be no such restriction if one works with the probability measure of X_t and uses the Fokker–Planck equation.

E-mail addresses: Mathieu.Lauriere@UPMC.fr (M. Laurière), Olivier.Pironneau@UPMC.fr (O. Pironneau).

[1] This preprint came to our knowledge after the submission of this note.

http://dx.doi.org/10.1016/j.crma.2014.07.008
1631-073X/

In this note we apply the dynamic programming argument to the value functional $V(\tau, \rho_\tau(\cdot))$, where ρ_τ is the PDF of X_τ. Of course, this is at the cost of several regularity assumptions; in particular, it requires the existence of PDF at all times.

Once the problem is reformulated with the Fokker–Planck equation, it becomes a somewhat standard exercise to find the optimality necessary conditions by a calculus of variations. So the note begins likewise. Then a similar result is obtained by using dynamic programming, and the connection with the previous approach and with stochastic dynamic programming is established, with the advantage that sufficient conditions for optimality are obtained. Finally, we apply the method to two mean-field type control problems stated in [1] and [6].

2. The problem

Let $d, s, r \in \mathbf{N}^+$. Consider a stochastic differential equation

$$dX_t = u(X_t, t)dt + \sigma\big(X_t, t, u(X_t, t)\big)dW_t \tag{2.1}$$

where $T > 0$, $u : \mathbf{R}^d \times (0, T) \to \mathbf{R}^d$, $\sigma : \mathbf{R}^d \times (0, T) \times \mathbf{R}^d \to \mathbf{R}^{d\times d}$ and W_t is a d-vector of independent Brownian motions. We make the usual assumptions for X_t to exist once X_0 is known [14].

Let $\tilde{H} : \mathbf{R}^d \times (0, T) \times \mathbf{R}^d \times \mathbf{R}^r \to \mathbf{R}$, $\tilde{h} : \mathbf{R}^d \times (0, T) \times \mathbf{R}^d \to \mathbf{R}^r$, $G : \mathbf{R}^d \times \mathbf{R}^s \to \mathbf{R}$, $g : \mathbf{R}^d \to \mathbf{R}^s$. Assume also that ρ_0 is positive with unit measure on $\mathbf{R}^d$.

Let $\mathcal{V}_d \subset \mathbf{R}^d$, $\mathcal{U}_d = \{u \in (L^\infty(\mathbf{R}^d \times \mathbf{R}))^d \ : \ u(x,t) \in \mathcal{V}_d \ \forall x, t\}$ and consider the problem

$$\min_{u\in\mathcal{U}_d} J := \int_0^T \mathbf{E}\big[\tilde{H}\big(X_t, t, u(X_t, t), \mathbf{E}\big[\tilde{h}\big(X_t, t, u(X_t, t)\big)\big]\big)\big]dt + \mathbf{E}\big[G\big(X_T, \mathbf{E}\big[g(X_T)\big]\big)\big]$$

$$\text{subject to (2.1) and such that } \rho_0 \text{ is the PDF of } X_0 \tag{2.2}$$

Andersson et al. [1] analyzed this problem using stochastic calculus of variations, claiming rightly that dynamic programming is not possible unless $\tilde{h} = 0$, $g = 0$. Yet denoting $Q = \mathbf{R}^d \times (0, T)$ and $\mu_{ij} = \frac{1}{2}\sum_k \sigma_{ik}\sigma_{jk}$, with sufficient regularity, namely if X_t has a PDF ρ_t (for weaker hypotheses see [11]), the problem is equivalent to

$$\min_{u\in\mathcal{U}_d} J = \int_Q H\big(x, t, u(x,t), \rho_t(x), \chi(t)\big)\rho_t(x)dxdt + \int_{\mathbf{R}^d} G(x, \xi)\rho_{|T}dx$$

$$\text{where } \chi(t) = \int_{\mathbf{R}^d} h\big(x, t, u(x,t), \rho_t(x)\big)\rho_t(x)dx, \ \xi = \int_{\mathbf{R}^d} g(x)\rho_T(x)dx \text{ and } \rho_t$$

$$\text{s.t.} \quad \partial_t\rho + \nabla\cdot(u\rho) - \nabla\cdot\nabla\cdot(\mu\rho) = 0, \quad \rho|_0 = \rho_0(x), \quad x \in \mathbf{R}^d \tag{2.3}$$

where $\tilde{H} = H$, $\tilde{h} = h$ if these are not functions of $\rho_t(x)$.

Hypothesis 1. Assume that all data are continuously differentiable with respect to u and ρ and have additional regularity so that the solution to the Fokker–Planck equation is unique and uniformly continuously differentiable with respect to u and μ.

3. Calculus of variations

Proposition 1. *Let $A : B = trace(A^{\mathrm{T}}B)$. A control u is optimal for (2.3) only if*

$$\int_{\mathbf{R}^d}\left(H'_u + h'_u\int_{\mathbf{R}^d} H'_\chi\rho dx + \nabla\rho^* - \mu'_u : \big(\nabla\nabla\rho^*\big)\right)(v-u)\rho dx \geqslant 0 \quad \forall t, \forall v \in \mathcal{U}_d \tag{3.4}$$

$$\text{with } \partial_t\rho^* + u\nabla\rho^* + \mu : \nabla\nabla\rho^* = -\left[H'_\rho\rho + H + \big(h'_\rho\rho + h\big)\int_{\mathbf{R}^d} H'_\chi\rho dx\right], \ \rho^*_{|T} = g\int_{\mathbf{R}^d} G'_\xi\rho_{|T}dx + G \tag{3.5}$$

Proof. See http://hal.archives-ouvertes.fr:hal-01018361. □

4. Dynamic programming

For notational clarity, let us consider the more general case where H, G are functionals of $\rho_t(\cdot)$. For any $\tau \in [0, T]$ and any $\rho_\tau \geqslant 0$ with unit measure on $\mathbf{R}^d$, let: