Algorithmes probabilistes pour les équations de Hamilton-Jacobi-Bellman en dimension élevée

Nicolas LANGRENÉ, CSIRO Data61 Melbourne

Côme HURÉ, Université Paris Diderot

Huyên PHAM, Université Paris Diderot

Achref Bachouch, Université d'Oslo

Nous considérons l'équation de Hamilton-Jacobi-Bellman générale de type

$$\frac{\partial v}{\partial t}(t,x)+\sup_{a\in A}\left\{b\left(x,a\right).D_x v(t,x)+\frac{1}{2}\mathrm{tr}\left(\sigma\sigma^\top\left(x,a\right)\right)D_x^2 v(t,x)\right)+f\left(x,a\right)\right\}=0,(t,x)\in[0,T)\times\mathbb{R}^d \quad (1)$$
$$v\left(T,x\right)=g\left(x\right)\ ,\ x\in\mathbb{R}^d\ .$$

Cette équation est directement associée au problème de contrôle stochastique suivant ([1])

$$v\left(t,x\right)=\sup_{\alpha\in\mathcal{A}}\mathbb{E}\left[\int_t^T f\left(X_s^\alpha,\alpha_s\right)ds+g\left(X_T^\alpha\right)|X_t^\alpha=x\right] \quad (2)$$
$$dX_s^\alpha=b\left(X_s^\alpha,\alpha_s\right)ds+\sigma\left(X_s^\alpha,\alpha_s\right)dW_s$$

dont une approximation en temps discret ($0=t_0<t_1<t_2<\ldots<t_N=T$) peut s'écrire

$$\hat{v}(t_N,x)=g(t_N,x)$$
$$\hat{v}(t_i,x)=\sup_{a\in A}\mathbb{E}\left[f\left(x,a\right)\left(t_{i+1}-t_i\right)+\hat{v}\left(t_{i+1},x+b\left(x,a\right)\left(t_{i+1}-t_i\right)+\sigma\left(x,a\right)\left(W_{t_{i+1}}-W_{t_i}\right)\right)\right]\ . \quad (3)$$

L'intérêt de l'approche probabiliste (2)-(3) pour résoudre numériquement l'équation de HJB (1) est qu'elle ouvre la voie à des algorithmes à base de simulations, lesquelles restent applicables lorsque la dimension d du problème est élevée.
Nombreux sont les domaines de recherche s'attaquant au problème en temps discret (3). [2] et [9] passent en revue et classifient les solutions proposées en recherche opérationnelle, en apprentissage par renforcement, en programmation stochastique et en programmation dynamique approchée. Pour l'essentiel, il s'agit de méthodes par force brute itérant de façon progressive l'équation de programmation dynamique (3) et approximant d'une façon ou d'une autre ce qui ne peut être calculé analytiquement.
Récemment, des améliorations algorithmiques venant de l'apprentissage automatique et de l'apprentissage profond ont permis de pousser de telles approches jusqu'à résoudre aisément des problèmes en dimension $d=100$, cf. [3], [10], [11].
Cette présentation se concentrera sur les récents travaux [12], [13], qui tirent profit des aptitudes pratiques de l'apprentissage profond au sein d'algorithmes plus classiques de Regression Monte Carlo ([1], [5], [8]) qui approximent les espérances conditionnelles et/ou les règles de décisions dans la boucle de programmation dynamique rétrograde (3) par des rseaux neuronaux. Les algorithmes proposés sont analysés théoriquement ([12]) et illustrés numériquement sur plusieurs exemples en dimension élevée ([13]).
Le fait que la résolution numérique des équations de HJB de type (1) soit désormais possible en dimension élevée a plusieurs conséquences pratiques, en premier lieu sur la modélisation des problèmes à résoudre. Avec ces nouveaux outils, le modélisateur n'a désormais plus à compromettre le réalisme d'un modèle pour le rendre numériquement soluble, les problèmes pratiques faisant effectivement intervenir des centaines de processus stochastiques différents tant rares. Désormais, le front de la recherche se situe sur l'extension de tels outils à des problèmes plus généraux que l'équation de HJB (1), tels que les problèmes de contrôle stochastique avec retard ([7]), les jeux stochastiques ([4]) ou les jeux à champ moyen ([6]).

Références

[1] Idris Kharroubi, Nicolas Langrené, Huyên Pham, *A numerical algorithm for fully nonlinear HJB equations: an approach by control randomization*, Monte Carlo Methods and Applications, 2014.

[2] Warren Powell, *Perspectives of approximate dynamic programming*, Annals of Operations Research, 2016.

[3] Weinan E, Jiequn Han, Arnulf Jentzen, *Deep learning-based numerical methods for high-dimensional parabolic partial differential equations and backward stochastic differential equations*, Communications in Mathematics and Statistics, 2017.

[4] René Aïd, Liangchen Li, Michael Ludkovski, *Capacity expansion games with application to competition in power generation investments*, Journal of Economic Dynamics and Control, 2017.

[5] Alessandro Balata, Jan Palczewski, *Regress-Later Monte Carlo for optimal control of Markov processes*, 2017.

[6] Clémence Alasseur, Imen Ben Tahar, Anis Matoussi, *An extended mean field game for storage in smart grids*, 2018.

[7] Thomas Lim, Idris Kharroubi, Vathana Ly-Vath, *Optimal exploitation of a resource with stochastic population dynamics and delayed renewal*, 2018.

[8] Michael Ludkovski, Aditya Maheshwari, *Simulation methods for stochastic storage problems: a statistical learning perspective*, Energy Systems, 2019.

[9] Warren Powell, *A unified framework for stochastic optimization*, European Journal of Operational Research, 2019.

[10] Christian Beck, Weinan E, Arnulf Jentzen, *Machine learning approximation algorithms for high-dimensional fully nonlinear partial differential equations and second-order backward stochastic differential equations*, 2019.

[11] Sebastian Becker, Patrick Cheridito, Arnulf Jentzen, *Deep optimal stopping*, 2019.

[12] Côme Huré, Huyên Pham, Achref Bachouch, Nicolas Langrené, *Deep neural networks algorithms for stochastic control problems on finite horizon: convergence analysis*, 2019.

[13] Achref Bachouch, Côme Huré, Nicolas Langrené, Huyên Pham, *Deep neural networks algorithms for stochastic control problems on finite horizon: numerical applications*, 2019.

Nicolas LANGRENÉ, Commonwealth Scientific and Industrial Research Organisation, Data61, RiskLab Australia, Door 34 Goods Shed North, Village Street, Docklands VIC 3008, Australia
nicolas.langrene@csiro.au
Côme HURÉ, LPSM, Université Paris Diderot, Bâtiment Sophie Germain, Avenue de France, 75013 Paris
hure@lpsm.paris
Huyên PHAM, LPSM, Université Paris Diderot, Bâtiment Sophie Germain, Avenue de France, 75013 Paris
pham@lpsm.paris
Achref Bachouch, Department of Mathematics, University of Oslo, Moltke Moes vei 35 Niels Henrik Abels hus, 0851 Oslo
achrefb@math.uio.no