Cornell University Library

We gratefully acknowledge support from the Simons Foundation and member institutions

arXiv.org > math > arXiv:1105.5041

Search or Article-id

(Help | Advanced search)

All papers   Go!

Mathematics > Statistics Theory

Lipschitz Bandits without the Lipschitz Constant

Sébastien Bubeck (CRM), Gilles Stoltz (DMA, GREGH, INRIA Paris - Rocquencourt), Jia Yuan Yu (DMA, GREGH, INRIA Paris - Rocquencourt)

(Submitted on 25 May 2011 (v1), last revised 14 Jul 2011 (this version, v2))

We consider the setting of stochastic bandit problems with a continuum of arms. We first point out that the strategies considered so far in the literature only provided theoretical guarantees of the form: given some tuning parameters, the regret is small with respect to a class of environments that depends on these parameters. This is however not the right perspective, as it is the strategy that should adapt to the specific bandit environment at hand, and not the other way round. Put differently, an adaptation issue is raised. We solve it for the special case of environments whose mean-payoff functions are globally Lipschitz. More precisely, we show that the minimax optimal orders of magnitude $L^{d/(d+2)} \, T^{(d+1)/(d+2)}$ of the regret bound against an environment f with Lipschitz constant L over T time instances can be achieved without knowing L or T in advance. This is in contrast to all previously known strategies, which require to some extent the knowledge of L to achieve this performance guarantee.

Subjects: **Statistics Theory (math.ST)**

Journal reference: 22nd International Conference on Algorithmic Learning Theory --- (ALT 2011), Espoo : Finland (2011)

Cite as: **arXiv:1105.5041 [math.ST]**

(or **arXiv:1105.5041v2 [math.ST]** for this version)

Submission history

Which authors of this paper are endorsers?