ONLINE CONVEX OPTIMIZATION ALGORITHMS AND SOME APPLICATIONS

D.B. Rokhlin

We describe general classes of gradient methods for online convex optimization problems. In fact, these problems are more gerenal than stochastic convex optimization problems due to the online-to-batch conversion scheme. Starting from the online gradient descent algorithm, we mention several incentives to study other online learning algorithms: using variable step sizes for unbounded domains; improvements of the estimates for narrower classes of convex functions; improvements of the constants in the estimates by using non-Euclidean norms; incorporating predictions of the feedbacks; finding scale-invariant and parameter-free algorithms. Following the recent paper [1], we consider Ada-FTRL (Adaptive Follow the Regularized Leared), Ada-MD (Adaptive Mirror Descent) classes of algorithms, and show, for instance, that the Exponentiated Gradient (EG), Online Newton Step (ONS) and optimistic learning algorithms are their particular cases.

Two applications, concerning the online portfolio selection and the choice of incentive transfer prices, are considered in more detail. One common goal that we have here is the developing of algorithms which do not require the parameter tuning. For online portfolio selection problem we consider the family of simple prediction algorithms [2] based on price averaging. After the discretization of the decay parameter we get the finite family of trivial investment strategies, based on these predictions. Regarding these strategies as experts, we apply the parameter-free coin-betting algorithm of [3], and, for several standard data sets, compare the results to the OLMAR algorithm proposed in [2].

In the second problem we consider a firm producing and selling d commodities, and consisting from n production and m sales divisions. The firm manager tries to stimulate the best division performance by sequentially selecting internal commodity prices (transfer prices). In the static problem under general strong convexity and compactness assumptions we show that the SOLO FTRL algorithm of [4] gives the estimates of order $T^{-1/4}$ in the number T of iterations for optimality and feasibility residuals. This algorithm uses only the information on division reactions to current prices. It does not depend on any parameters and requires no information on the production and cost functions. In the dynamic problem we assume that these functions depend on an i.i.d. sequence of random variables. It is shown that the same algorithm over a hypercube ensures no-regret learning with respect to the best possible plan sequence, and the average regret is stochastically bounded by a quantity of order $T^{-1/4}$. In this case the algorithm requires the knowledge of upper bounds for optimal transfer prices.

REFERENCES

1. *Joulani P., György, Szepesvári* A modular analysis of adaptive (non-)convex optimization: Optimism, composite objectives, variance reduction, and variational bounds, *Theoretical Computer Science*, **808**, 2020, 108-138.

2. *Li B., Hoi S.C.H., Sahoo D., Liu Z.-Y.* Moving average reversion strategy for on-line portfolio selection, *Artificial Intelligence*, **222**, 2015, 104-123.

3. *Orabona F., Pál D.,* Coin betting and parameter-free online learning. In D.D. Lee, M. Sugiyama, U. von Luxburg, I. Guyon, R. Garnett (Eds.), Advances in Neural Information Processing Systems 29, NIPS 2016, Curran Associates, Inc. (2016), pp. 577-585

The research is supported by the Russian Science Foundation, project 17-19-01038.

4. *Orabona F., Pál D.,* Scale-free online learning, *Theoretical Computer Science*, **716**, 2018, 50-69.