Robust and Adaptive Methods for Sequential Decision Making

by Wouter M. Koolen (CWI)

Machine learning systems for sequential decision making continuously improve the quality of their actions. Our new adaptive methods learn to improve as fast as possible in a wide range of applications.

Many practical problems can be cast as sequential decision making tasks, ranging from time series prediction, data compression and portfolio management to online versions of routing and ranking. Yet other problems may be reduced to sequential decision making, in particular batch learning from large data sets. The main challenge is the design of generic, efficient, robust and adaptive learning methods.

Following [1], we model the learning process as a series of interactions between the learner and its environment. Each round the learner picks an action. Then the environment reveals the quality of all available actions by means of a "loss function". Based on this feedback the learner updates its internal state, with the goal of picking better future actions.

Robustness is traditionally achieved by taking a game-theoretic perspective. We interpret the learning problem as a sequential game, and regard the environment as an adversary that chooses the loss functions to maximally frustrate the learner. One of the accomplishments of the online learning community is the design of learners that are provably successful in such games. It is amazing that generic learning methods can be designed for such a rich class of problems.

Unfortunately, the robust game-theoretic methods do not always fare well in practice, in the sense that for some problems they are outperformed significantly by special-purpose methods. The reason is that, in practice, environments are not maximally evil, and may admit more aggressive learning techniques which are unsafe in general. This realisation has spurred the search for reasonable assumptions that describe the additional useful structure present in practical problems.

We studied two popular types of such assumptions. First, one may require that the loss functions exhibit curvature, either in every direction (strong convexity) or along the gradient (exp-concavity). For either scenario with known curvature, methods have been developed that increase performance dramatically. Yet in practice the curvature often is a feature of the data, and hence unknowable a-priori. We aim to design methods that adapt to any exploitable curvature presented.

Second, one may consider the relation between loss functions issued by the environment over the course of multiple rounds. A simple (but rich and often practically reasonable) class of assumptions is obtained by modelling the sequence of loss functions as an independent and identically distributed stochastic process. We may then characterise the simplicity of the distribution by its "Bernstein exponent" (generalisation of the Tsybakov margin condition used in classification). For distributions with a known Bernstein exponent it is possible to design learning algorithms that perform optimally. But how does one deal with and adapt to an unknown Bernstein exponent?

Online learning methods maintain internal parameters that are updated every round based on the newly observed loss function. To exploit the aforementioned structural properties of the environment (curvature and/or Bernstein exponent) it turns out that the magnitude of this update needs to be tuned carefully. The friendlier the problem, the larger the step size required (whence the title of the NIPS workshop series "Learning Faster from Easy Data").

Tuning the step size, a single scalar, appropriately for some unknown parameters might sound like a simple learning problem in itself. Yet no
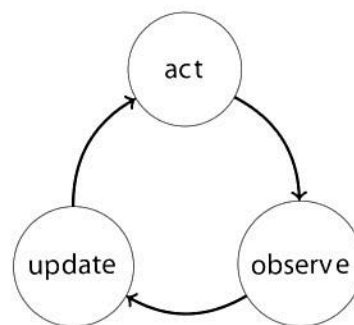

Figure 1: Sequential decision making.

existing learning method would apply: the overhead for learning the step size would dwarf the benefit of setting it right. The key contribution of our new method, called MetaGrad, is a novel approach for learning the step size from the data.

We proved that MetaGrad has a certain performance guarantee of second-order form [2]. This bound implies in particular worst-case safety (robustness), adaptivity to curvature and adaptivity to the Bernstein exponent [3]. This establishes MetaGrad as the new state-of-the art adaptive method for online convex optimisation. Since MetaGrad only has

modest computational overhead over earlier methods, we also expect it to be highly relevant in practical applications.

The research team consisted of Wouter M. Koolen (CWI), Tim van Erven (Leiden University) and Peter Grünwald (CWI and Leiden University). The author is funded by an NWO Veni grant. The results will be presented at the 30th NIPS conference in December 2016. Our reference implementation of MetaGrad is publicly available [L1].

Link:
[L1] https://bitbucket.org/wmkoolen/metagrad

References:
[1] S. Shalev-Shwartz: "Online learning and online convex optimization", Foundations and Trends in Machine Learning, 4(2):107–194, 2012.
[2] T. van Erven, W. M. Koolen: "MetaGrad: Multiple learning rates in online learning", accepted to Advances in Neural Information Processing Systems (NIPS) 29, 2016, Pre-print. https://arxiv.org/abs/1604.08740.
[3] W. M. Koolen, P. D. Grünwald, T. van Erven: "Combining adversarial guarantees and stochastic fast rates in online learning", accepted to Advances in Neural Information Processing Systems (NIPS) 29, 2016. Pre-print https://arxiv.org/abs/1605.06439.

Please contact:
Wouter M. Koolen, CWI, The Netherlands.
+31 20 592 4244
W.M.Koolen-Wijkstra@cwi.nl