

Quentin Berthet:
Learning with Differentiable Perturbed Optimizers

Machine learning pipelines often rely on optimization procedures to make discrete decisions (e.g., sorting, picking closest neighbors, or shortest paths). Although these discrete decisions are easily computed, they break the back-propagation of computational graphs. In order to expand the scope of learning problems that can be solved in an end-to-end fashion, we propose a systematic method to transform optimizers into operations that are differentiable and never locally constant. Our approach relies on stochastically perturbed optimizers, and can be used readily together with existing solvers. Their derivatives can be evaluated efficiently, and smoothness tuned via the chosen noise amplitude. We also show how this framework can be connected to a family of losses developed in structured prediction, and give theoretical guarantees for their use in learning tasks. We demonstrate experimentally the performance of our approach on various tasks. (joint work with M. Blondel, O. Teboul, M. Cuturi, J-P. Vert, F. Bach)