

FROM THE THEORY OF (STOCHASTIC) CONTROL TO DEEP LEARNING AND BACK.

LUKASZ SZPRUCH

ABSTRACT. There is overwhelming empirical evidence that deep neural networks trained with stochastic gradient descent perform (extremely) well in the high dimensional settings. Nonetheless, a complete mathematical theory that would provide theoretical guarantees why and when these methods work so well has been elusive. In this mini-course, I will demonstrate how one may leverage control theory and the theory of statistical sampling to study the convergence of stochastic gradient algorithms used in deep learning. Conversely, I will show that machine learning perspective leads to new algorithms for (stochastic) control problems and offers a fresh perspective on classical quantitative finance problems.

The course will consist of 4 parts.

- *Mean-Field Langevin Dynamics.* In this part of the course, I will show that training of one hidden layer neural network can be framed as an optimisation problem over Wasserstein space. I will show that under appropriate conditions, the flow of marginal laws induced by the corresponding gradient flow converges to a stationary distribution, which is exactly the minimiser of the functional one aims to optimise. See [?] for more details.
- *Neural ODEs via Relaxed Optimal Control.* In this part, of course, I will demonstrate that by using the theory of relaxed optimal controls one can extend the analysis of one-hidden layer to neural ODEs that are a model for deep recurrent neural networks. Using calculus on Wasserstein space, it is possible to offer a perspective on the generalisation error. See [2] for more details.
- *Gradient flows for stochastic control problem.* While control theory provides a useful perspective when analysing deep learning, machine learning point of view leads to new algorithms for solving stochastic control problems. In this part of the course, I will extend the analysis developed for Neural ODEs to solve the regularised stochastic control problem. See [3] for more details.
- *Robust pricing and hedging with neural SDEs* In the final part of the mini-course, I will show that neural SDEs provide an attractive class of models that seamlessly integrate deep neural networks with classical quantitative finance models. In particular by using neural SDE one can: a) consistently calibrate these under the risk-neutral measure as well as the real-world measure; b) easily integrate additional market information; c) verify martingale property; d) build market generators. See [1] for more details.

REFERENCES

- [1] P. Gierjatowicz, M. Sabate-Vidales, D. Siska, L. Szpruch, and Z. Zuric. Robust pricing and hedging via neural sdes. Available at SSRN 3646241, 2020.
- [2] J.-F. Jabir, D. Šiška, and L. Szpruch. Mean-field neural odes via relaxed optimal control. arXiv:1912.05475, 2019.
- [3] D. Šiška and L. Szpruch. Gradient flows for regularized stochastic control problems. [arXiv preprint arXiv:2006.05956](#), 2020.