Optimisation stochastique pour de grandes masses de données

- Kevin Scaman (INRIA)

  Title: Introduction à l'optimisation stochastique pour des grandes masses de données.

- Raphaël Berthier (INRIA)

  Title: A Continuized View on Nesterov Acceleration for Stochastic Gradient Descent and Randomized Gossip

  Abstract: We introduce the "continuized" Nesterov acceleration, a close variant of Nesterov acceleration whose variables are indexed by a continuous time parameter. The two variables continuously mix following a linear ordinary differential equation and take gradient steps at random times. This continuized variant benefits from the best of the continuous and the discrete frameworks: as a continuous process, one can use differential calculus to analyze convergence and obtain analytical expressions for the parameters; and a discretization of the continuized process can be computed exactly with convergence rates similar to those of Nesterov original acceleration. We show that the discretization has the same structure as Nesterov acceleration, but with random parameters. We provide continuized Nesterov acceleration under deterministic as well as stochastic gradients, with either additive or multiplicative noise. Finally, using our continuized framework and expressing the gossip averaging problem as the stochastic minimization of a certain energy function, we provide the first rigorous acceleration of asynchronous gossip algorithms.

  Joint with Mathieu Even, Raphaël Berthier Francis Bach, Nicolas Flammarion, Pierre Gaillard, Hadrien Hendrikx, Laurent Massoulié, and Adrien Taylor

- Edwige Cyffers (INRIA)

  Title: Differentially Private Decentralized Learning with Random Walks

  Abstract: The popularity of federated learning comes from the possibility of better scalability and the ability for participants to keep control of their data, improving data security and sovereignty. Unfortunately, sharing model updates also creates a new privacy attack surface. In this work, we characterize the privacy guarantees of decentralized learning with random walk algorithms, where a model is updated by traveling from one node to another along the edges of a communication graph. Using a recent variant of differential privacy tailored to the study of decentralized algorithms, namely Pairwise Network Differential Privacy, we derive closed-form expressions for the privacy loss between each pair of nodes where the impact of the communication topology is captured by graph theoretic quantities. Our results further reveal that random walk algorithms tends to yield better privacy guarantees than gossip algorithms for nodes close from each other. We supplement our theoretical results with empirical

evaluation on synthetic and real-world graphs and datasets. Joint work with : Edwige Cyffers, Aurélien Bellet, Jalaj Upadhyay

- Aymeric Dieuleveut (École Polytechnique)