Panagiotis Mertikopoulos, Department of Mathematics, NKUA

TITLE:
From Robbins-Monro to artificial intelligence: 70 years of stochastic approximation and the road ahead

ABSTRACT:
In this talk, I will present a (biased) sample of results from the theory of stochastic approximation, from the seminal ideas of Robbins-Monro and Kiefer-Wolfowitz in the 1950's, to the general theory of asymptotic pseudotrajectories by Benaïm & Hirsch and its modern applications to machine learning and artificial intelligence. We will begin by discussing how the asymptotic behavior of a wide range of stochastic iterative algorithms for solving non-convex minimization and min-max problems can be understood by means of a dynamical system in continuous time. In the case of gradient-like, minimization problems, this methodology allows us to derive a range of conditions guaranteeing convergence to sets of critical points while avoiding unstable saddle points and other undesirable solutions. However, in the min-max case, the overall situation is considerably more involved, and exhibits convergence to lower-dimensional manifolds that are in no way critical (or otherwise min-max optimal). We show that "spurious attractors" of this type may arise even in simple two-dimensional problems with polynomial losses of degree 4, a fact which highlights the fundamental gap between minimization and min-max problems - and thus, in the context of machine learning, the reason that generative models are so much harder to train than discriminative ones.