

DIMENSION REDUCTION FOR STATISTICAL INFERENCE VIA (DIMENSIONAL)

MATTHEW LI

In this work, we propose and analyse a continuous-time model for stochastic gradient descent with momentum. This model is a piecewise-deterministic Markov process that represents the particle movement by an underdamped dynamical system and the data subsampling through a stochastic switching of the dynamical system. In our analysis, we investigate longtime limits, the subsampling-to-no-subsampling limit, and the momentum-to-no-momentum limit. We are particularly interested in the case of reducing the momentum over time: intuitively, the momentum helps to overcome local minimisers in the initial phase of the algorithm, but prohibits fast convergence to a global minimiser later. Under convexity assumptions, we show convergence of our dynamical system to the global minimiser when reducing momentum over time and let the subsampling rate go to infinity.