

New error estimates for the Adam optimization algorithm

Steffen Dereich

University of Münster

`steffen.dereich@uni-muenster.de`

Coauthor(s): Arnulf Jentzen

Special session: Stochastic Computation and Complexity

The Adam optimization algorithm is a very popular training algorithm for artificial neural networks. It is a stochastic gradient descent algorithm that combines a momentum approach with a damping factor. Despite its popularity its theoretical analysis is still in its infancy. In this talk, we provide new error estimates for the Adam algorithm and we illustrate our findings in a simple example, where we estimate expectations of random variables.