A Recursive Multilevel Algorithm for Deep Learning

Tuesday, August 13, 2024 9:30 AM (30 minutes)

As the use cases for neural networks become increasingly complex, modern neural networks must also become deeper and more intricate to keep up, indicating the need for more efficient learning algorithms. Multilevel methods, traditionally used to solve differential equations using a hierarchy of discretizations, offer the potential to reduce computational effort.

In this talk, we combine both concepts and introduce a multilevel stochastic gradient descent algorithm that accelerates learning through a multilevel strategy. A gradient correction term is needed to establish first-order consistency.

We discuss convergence of the method in the case of a deterministic gradient correction as well as a stochastic gradient correction under additional conditions including step size regularization and an angle condition.

To demonstrate the usefulness of our approach, we apply it to residual neural networks in image classification. The resolution of the images is utilized to generate data sets of varying complexity, which are then used to build a hierarchy of neural networks with a decreasing number of variables. Additionally, we construct corresponding prolongation and restriction operators. Numerical results are presented.

Authors: JACOB, Isabel (TU Darmstadt); Prof. ULBRICH, Stefan (TU Darmstadt)

Presenter: JACOB, Isabel (TU Darmstadt)

Session Classification: MS 01: Optimal Control and Machine Learning

Track Classification: Minisymposia: MS 01: Optimal Control and Machine Learning