Projection methods for the training of neural networks

Veit Elser
Cornell University, NY, USA

A weekly webinar on "Projection Methods in Feasibility, Superiorization and Optimization – Theory and Practice"organized by the division of Optimization of the Fraunhofer-Institut für Techno- und Wirtschaftsmathematik ITWM in Kaiserslautern, Germany

Abstract

Projection methods are an attractive alternative to gradient descent in the training of neural networks [1] because the computation distributes over both the network and data. They also have a unique advantage, in that activation functions may be replaced by *activation constraints*. An example of the latter is a constraint on the input-output pair (x, y) of a binary-valued neuron, with output $y \in \{0, 1\}$, but where the input is constrained to avoid a "gray area," say $|x| < 1$, where the conventional activation function switches between the two output values. By projecting to such activation constraints, during training, the network is forced to learn weight parameters such that all the neurons have an unambiguous output, promoting a semantic interpretation. As a simple demonstration of the method we train an autoencoder tasked with discovering the rule for expressing $0, \ldots, 2^n - 1$ as a string of n bits.

References

[1] Veit Elser, *Learning without loss*, arXiv:1911.00493.