

Decomposition Techniques for Multi Layer Perceptrons Training

Luigi Grippo, Andrea Manno and Marco Sciandrone

March 27, 2015

Abstract

In this work we consider the learning problem of Multi Layer Perceptrons (MLPs) formulated as the problem of minimizing a smooth error function. As well-known the learning problem of MLPs can be a difficult nonlinear nonconvex optimization problem. Typical difficulties can be the presence of extensive flat regions and steep sided valleys in the error surface, and the possible large number of training data and of free network parameters. We define a wide class of batch learning algorithms for MLP, based on the use of block decomposition techniques in the minimization of the error function. The learning problem is decomposed into a sequence of smaller and structured minimization problems in order to advantageously exploit the structure of the objective function. Theoretical convergence results are established and a specific algorithm is constructed and evaluated through an extensive numerical experimentation. The comparisons with state-of-the-art learning algorithms show the effectiveness of the proposed techniques.