

Computationally feasible greedy algorithms for neural nets

Andrew Barron
Yale University

Abstract. Previously, greedy algorithms have been shown to have good approximation and estimation properties for superpositions of a sigmoidal function or other ridge activation functions. In each step the parameters of a new sigmoid are fit to the residuals of the previous sigmoids. However, it remains a challenging task to produce guarantees for the parameter search for each sigmoid. Here we discuss developments of two algorithms for this task. One is an implementation of adaptive annealing, in which internal modifications of the sigmoid is made, allowing stability of the adaptive annealing solution. The other algorithm is a convergent nonlinear modification of the tensor methods of Anandkumar and her colleagues which allows optimization of the inner product of the residuals with certain sigmoids and related activation functions. This work is joint with Jason Klusowski.