The Rate of Convergence of AdaBoost

Indraneel Mukherjee, Cynthia Rudin, Robert E. Schapire

(Submitted on 29 Jun 2011)

The AdaBoost algorithm was designed to combine many "weak" hypotheses that perform slightly better than random guessing into a "strong" hypothesis that has very low error. We study the rate at which AdaBoost iteratively converges to the minimum of the "exponential loss." Unlike previous work, our proofs do not require a weak-learning assumption, nor do they require that minimizers of the exponential loss are finite. Our first result shows that at iteration t, the exponential loss of AdaBoost's computed parameter vector will be at most ϵ more than that of any parameter vector of ℓ_1-norm bounded by B in a number of rounds that is at most a polynomial in B and $1/\epsilon$. We also provide lower bounds showing that a polynomial dependence on these parameters is necessary. Our second result is that within C/ϵ iterations, AdaBoost achieves a value of the exponential loss that is at most ϵ more than the best possible value, where C depends on the dataset. We show that this dependence of the rate on ϵ is optimal up to constant factors, i.e., at least $\Omega(1/\epsilon)$ rounds are necessary to achieve within ϵ of the optimal exponential loss.

Comments:	A preliminary version will appear in COLT 2011
Subjects:	**Optimization and Control (math.OC)**; Artificial Intelligence (cs.AI); Machine Learning (stat.ML)
Cite as:	**arXiv:1106.6024 [math.OC]**
	(or **arXiv:1106.6024v1 [math.OC]** for this version)

Submission history

From: Indraneel Mukherjee [view email]
[v1] Wed, 29 Jun 2011 18:53:46 GMT (44kb)

Which authors of this paper are endorsers?