Corrigendum: A Random Sampling Algorithm for Learning an Intersection of Halfspaces

Santosh S. Vempala
School of Computer Science, Georgia Tech
Atlanta, USA
vempala@gatech.edu

***Abstract*—We correct a claim from [Vem97] and provide a status update.**

Learning convex sets is a fundamental topic in algorithms. The simplest convex concept, a single halfspace, can be PAC-learned in polynomial time via efficient linear programming algorithms. The complexity of PAC-learning an intersection of two halfspaces is open. Progress has been made on learning an intersection of halfspaces under restricted distributions and on learning using restricted hypothesis classes [She10].

In 1990, Baum [Bau90] gave an algorithm for learning an intersection of two homogeneous halfspaces over any origin-symmetric distribution. Baum's algorithm was recently shown to work for logconcave distributions [KLT09]. A few years after Baum's work, Blum and Kannan [BK93], [BK97] found a polynomial-time algorithm to learn the intersectio of a constant number of halfspaces under the uniform distribution in the unit ball. The running time, the number of examples required and the size of the hypothesis reported by their algorithm are all doubly exponential in k, namely $n^{2^{O(k)}}$.

In 1997, we presented an algorithm [Vem97], [Vem04] with running time and sample complexity $n^k(k/\epsilon)^{O(k)}$, i.e., singly exponential in k. The conference version [Vem97] claimed a fixed polynomial dependence on n and this was corrected in [Vem04]. The algorithm is based on (a) approximating the positive region using a large sample of examples (b) estimating the normal subspace of the positive region using one-dimensional random projections and (c) greedily choosing a subset of normal vectors from the normal subspace. The algorithm was shown to work for near-uniform distributions on the unit ball with the property that the density does not vary by more than a polynomial factor. The hypothesis learned is itself an intersection of $O(k\log(1/\epsilon))$ halfspaces. Recently, in the full journal version [Vem10b], this algorithm and its analysis were extended to any logconcave distribution in $\mathbb{R}^n$ with the same complexity bounds.

Klivans, O'Donnell and Servedio [KOS08] gave an algorithm based on approximating an intersection of k halfspaces with a low-degree polynomial threshold function. Their approach has time and sample complexity $n^{O(\log k/\epsilon^4)}$, works for Gaussian input distributions, and outputs a hypothesis that is a polynomial threshold function of degree $O(\log k/\epsilon^4)$. Thus, their complexity was a substantial improvement as a function of k, although worse as a function of the error parameter ϵ.

Our paper in this proceedings [Vem10a], based on PCA, achieves a complexity of

$$\mathrm{poly}(n,k,1/\epsilon) + n \cdot \min\ k^{O(\log k/\epsilon^4)}, (k/\epsilon)^{O(k)},$$

improving on both [Vem04], [KOS08] and achieving a fixed polynomial dependence on n for Gaussian input distributions. The algorithm from [Vem97] remains the state of the art for learning an intersection of halfspaces from input distributions other than Gaussians. All these results are summarized in the following table.

Reference	complexity	distribution
[Bau90] ($k = 2$)	$\mathrm{poly}(n, 1/\epsilon)$	origin-symmetric
[BK93], [BK97]	$(n/\epsilon)^{2^{O(k)}}$	uniform in ball
[Vem97], [Vem04]	$n^k(k/\epsilon)^{O(k)}$	nonconcentrated
[KLT09] ($k = 2$)	$\mathrm{poly}(n, 1/\epsilon)$	logconcave
[Vem10b]	$n^k(k/\epsilon)^{O(k)}$	logconcave
[KOS08]	$n^{O(\log k/\epsilon^4)}$	Gaussian
[Vem10a]	$\mathrm{poly}(n, k, 1/\epsilon)+$ $n \min\ k^{O(\log k/\epsilon^4)}, (k/\epsilon)^{O(k)}$	Gaussian

I. Acknowledgment

We thank the FOCS 2010 committee for the suggestion and opportunity to publish this note.

References

[Bau90] Eric B. Baum, *Polynomial time algorithms for learning neural nets*, COLT, 1990, pp. 258–272.
[BK93] Avrim Blum and Ravi Kannan, *Learning an intersection of k halfspaces over a uniform distribution*, FOCS, 1993, pp. 312–320.
[BK97] Avrim Blum and Ravindran Kannan, *Learning an intersection of a constant number of halfspaces over a uniform distribution*, J. Comput. Syst. Sci. **54** (1997), no. 2, 371–380.
[KLT09] Adam R. Klivans, Philip M. Long, and Alex K. Tang, *Baum's algorithm learns intersections of halfspaces with respect to logconcave distributions*, APPROX-RANDOM, 2009, pp. 588–600.
[KOS08] Adam R. Klivans, Ryan O'Donnell, and Rocco A. Servedio, *Learning geometric concepts via gaussian surface area*, FOCS, 2008, pp. 541–550.
[She10] Alexander A. Sherstov, *Optimal bounds for sign-representing the intersection of two halfspaces by polynomials*, STOC, 2010.
[Vem97] Santosh Vempala, *A random sampling based algorithm for learning the intersection of half-spaces*, FOCS, 1997, pp. 508–513.
[Vem04] Santosh S. Vempala, *The random projection method*, AMS, 2004.
[Vem10a] Santosh Vempala, *Learning convex concepts from gaussian distributions with pca*, FOCS, 2010.
[Vem10b] ______, *A random sampling based algorithm for learning the intersection of half-spaces*, JACM, to appear (2010).