Rigorous Bounds to Retarded Learning

Using an elegant approach, Herschkowitz and Opper [1] established rigorous bounds on the information inferable (learnable) from a set of data (m points $\in \Re^N$) when the latter are drawn from a distribution $P(\mathbf{x}) = P_0(\mathbf{x}) \times \exp[-V(\lambda)]$, where $P_0(\mathbf{x})$ is a spherical normal law and $\exp[-V(\lambda)]$ is a modulation along an unknown anisotropy axis $\lambda = \mathbf{w} \cdot \mathbf{x}$, for some direction $\mathbf{w}$. They show in particular that if $P(\mathbf{x})$ has zero mean, it is impossible to learn the direction of anisotropy below a critical fraction of data α^*, and claim that $\alpha^* = \alpha_{lb} \equiv (1 - \overline{\lambda^2})^{-1}$ *only* depends on $\overline{\lambda^2}$, the second moment of the distribution along λ, $P(\lambda) \equiv e^{-\lambda^2/2} \exp[-V(\lambda)]/\sqrt{2\pi}$.

The authors reach this conclusion by an expansion at small q of the upper bound to ΔR, the difference between the trivial risk and the cumulative Bayes risk. In the thermodynamic limit ($m \to \infty, N \to \infty$, with $\alpha = m/N$ finite) the upper bound ($\min_q$ erroneously stands for $\max_q$ in Eq. (7) of [1]) is given by $\max_q G_\alpha(q) = G_\alpha(q^*)$, with $G_\alpha(q) = \ln(1 - q^2)^{1/2} + \alpha \ln F(q)$ where

$$F(q) = \iint Dx\, Dy\, e^{[-V(x) - V(xq + y\sqrt{1-q^2})]}, \qquad (1)$$

and $Dx = e^{-x^2/2} dx/\sqrt{2\pi}$. The authors show that $q = 0$ is always a maximum for $\alpha \leq \alpha_{lb}$. In the case of highly anisotropic data distributions, when the learning task is simple enough that only the variance matters, this leads to $\alpha^* = \alpha_{lb}$. However, they disregarded the possibility of having other extrema, which we expect to exist [2,3] if there is some structure in the data along λ. We show here that the global maximum may jump from $q^* = 0$ to a *finite* value $q^* \equiv q_1$, at $\alpha^* = \alpha_1 < \alpha_{lb}$. Consider data with components along λ drawn according to

$$P(\lambda) = \frac{1}{2\sigma\sqrt{2\pi}} [e^{-(\lambda-\rho)^2/2\sigma^2} + e^{-(\lambda+\rho)^2/2\sigma^2}]. \qquad (2)$$

A straightforward calculation shows that $G_\alpha(q)$ has indeed a maximum at $q_1 > 0$, which may overcome the one at $q = 0$ for some values of ρ and σ. This *first order* phase transition of the upper bound signals the onset of a phase where learning is possible at $\alpha_1 < \alpha_{lb}$. In Fig. 1 we represent α_1 and α_{lb} as a function of ρ, for $\sigma = 0.5$. It may be seen that $\alpha_1(\rho)$ leaves $\alpha_{lb}(\rho)$ with a discontinuous slope at $\rho = 0.7023(7)$ [$\alpha_1 = \alpha_{lb} = 15.177(9)$, $q_1 = 0.900(5)$] [the inset shows the two maxima of $G(q)$], but smoothly at $\rho = 1.338(1)$ ($q_1 = 0, \alpha_1 = \alpha_{lb} = 0.96$). In the latter case, both the second and fourth order coefficients of the q expansion of $G_\alpha(q)$ vanish at the transition. The other inset represents $G_\alpha(q)$ for $\rho = 1$; the transition occurs at $\alpha_1 = 4.477(5) < \alpha_{lb} = 16$, at which q^* jumps from 0 to $q_1 \sim 0.876(4)$.

In some limiting cases, the first order transition may occur with a jump of q^* from 0 directly to $q^* = 1$ at

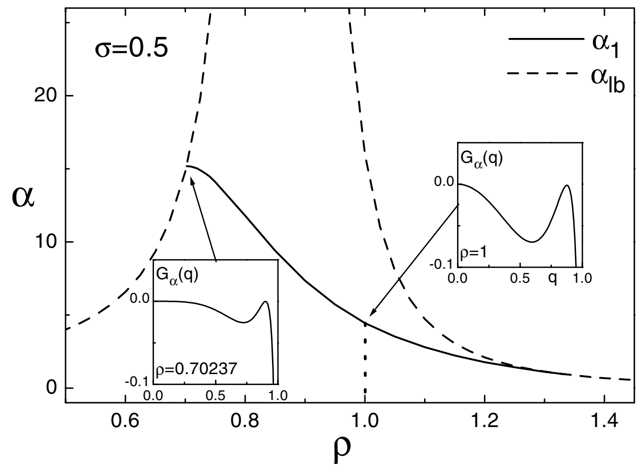

FIG. 1. Lower bound to the fraction of examples α below which learning is impossible, as a function of ρ for $\sigma = 0.5$. Insets: $G(q)$ showing the maxima at $q = 0$ and at $q_1 > 0$, for the two particular values of (α_1, ρ) indicated by the arrows.

$\alpha^* = \alpha_1 = 1$. This arises, for example, for $P(\lambda) = \frac{a}{2} \sum_{\tau=\pm 1} \delta(\lambda - \tau\rho) + \frac{1-a}{2} \sum_{\tau=\pm 1} \delta(\lambda - \tau\rho')$, when $a = 0.2$, $\rho = 0.5$, $\rho' = 1.4$.

One of the main conclusions of Ref. [1], based on this upper bound, is that retarded learning exists whenever $\overline{\lambda} = 0$. This conclusion is not invalidated by the present analysis: although there is no simple and general expression for α^*, it can be shown that $0 < \alpha^* \leq \alpha_{lb}$.

This work was done at the ZiF, Bielefeld.

Arnaud Buhot
Theoretical Physics
University of Oxford
1 Keble Road
Oxford, OX1 3NP, United Kingdom

Mirta B. Gordon
DRFMC CEA Grenoble
17 rue des Martyrs
38054 Grenoble Cedex 9, France

Jean-Pierre Nadal
LPS ENS*
24 rue Lhomond
75231 Paris Cedex 05, France

Received 6 July 2001; published 12 February 2002
DOI: 10.1103/PhysRevLett.88.099801
PACS numbers: 87.10.+e

*Laboratory associated with CNRS (UMR 8550) and the Universities Paris VI and Paris VII.

[1] D. Herschkowitz and M. Opper, Phys. Rev. Lett. **86**, 2174–2177 (2001).
[2] M. B. Gordon and A. Buhot, Physica (Amsterdam) **257A**, 85–98 (1998).
[3] A. Buhot and M. B. Gordon, Phys. Rev. E **57**, 3326–3333 (1998).

