COLLOQUIUM MATHEMATICUM

LI DÉDIÉ À M. STANISŁAW HARTMAN 1987

# ON RANDOM CONVEX HULLS

BY

JAN MYCIELSKI (BOULDER, COLORADO)

Let $S^{d-1}$ be the sphere $x_0^2 + \ldots + x_d^2 = 1$ in $R^d$. Let $A, B \subseteq S^{d-1}$ be two finite sets of "data points" given by some experiment. The perceptron learning procedure (see [2]–[5]) is an algorithm for finding a linear subspace $L$ of $R$, of dimension $d-1$, which separates $A$ from $B$, whenever such a subspace exists. The following problem arises. Suppose that the procedure succeeded and found a separating $L$; is this a significant fact or merely a chance phenomenon? We will show that, under appropriate conditions, the existence of $L$ is very improbable, whence the result can be quite significant.

In general, the answer depends on the probability distribution (in $S^{d-1}$) which rules the choice of the points. If, e.g., there was a constraint $A \cup B \subseteq \{(x_1, \ldots, x_d): x_1 \geq \frac{1}{2}\}$, then the cardinalities of $A$ and $B$ and their relative sizes are essential and little can be said except to suggest that the probability in question could perhaps be estimated by a Monte Carlo test. However, we will settle the problem for the case when the probability distribution of the points of $A$ and $B$ is uniform over $S^{d-1}$, or, more generally, when

(*) every $(d-2)$-dimensional great circle in $S^{d-1}$ has probability 0 and the probability is invariant under reflection of $S^{d-1}$ through its center.

A natural case of the problem, namely when $B$ is empty, was solved under those conditions by J. G. Wendel [6]. Our theorem is an easy corollary of his, but, for convenience of the reader, we produce a complete proof (which is somewhat different from his).

THEOREM. *If a probability measure $\mu$ in $S^{d-1}$ satisfies (*) and $x_1, \ldots, x_n$ are chosen in $S^{d-1}$ independently according to $\mu$, then the probability $p$ that $\{x_1, \ldots, x_i\}$ is linearly separable from $\{x_{i+1}, \ldots, x_n\}$ does not depend on $i$ and is given by*

$$p = \frac{1}{2^{n-1}} \sum_{k=0}^{d-1} \binom{n-1}{k}. \tag{1}$$

*Moreover the probability that the center* $0$ *of* $S^{d-1}$ *is not in the convex hull of* $\{x_1, \ldots, x_n\}$ *also equals* $p$.

Remark. Of course, by (1) and the theorem of de Moivre–Laplace,

$$p \approx \Phi\left(\frac{d-n/2}{\sqrt{n}/2}\right),$$

where $\Phi$ is the standard normal distribution. Hence $p$ tends to 0 iff $\frac{\sqrt{n}}{2} - \frac{d}{\sqrt{n}}$ tends to $\infty$.

Proof of the Theorem. The fact that the probability does not depend on $i$ follows immediately from three obvious facts:

(a) $\{x_1, \ldots, x_i\}$ is linearly separable from $\{x_{i+1}, \ldots, x_n\}$ iff the set $\{-x_1, \ldots, -x_i, x_{i+1}, \ldots, x_n\}$ is linearly separable from the empty set $\emptyset$.

(b) The distribution of $\{-x_1, \ldots, -x_i, x_{i+1}, \ldots, x_n\}$ is the same as the distribution of $\{x_1, \ldots, x_n\}$. (This follows from $(*)$.)

(c) 0 is not in the convex hull of $\{x_1, \ldots, x_n\}$ iff $\{x_1, \ldots, x_n\}$ is linearly separable from $\emptyset$.

At this point the theorem of Wendel [6] yields (1), but, for convenience of the reader, we continue the proof (in a slightly different way).

Let $S \subseteq \{1, \ldots, n\}$ and $p$ be the probability that $\{x_i\colon i \in S\}$ is linearly separable from $\{x_i\colon i \notin S\}$. By the above remarks $p$ does not depend on $S$. Hence

$$p = \Pr[(x_1, \ldots, x_n, S) \text{ is such that } \{x_i\colon i \in S\} \text{ is linearly separable from } \{x_i\colon i \notin S\}],$$

where $x_i$ are chosen as in the Theorem and $S$ is chosen uniformly among the $2^n$ subsets of $\{1, \ldots, n\}$. Hence, by $(*)$, the equation (1) of the Theorem follows immediately from the theorem of Fubini and the following Lemma.

Lemma. *Let* $x_1, \ldots, x_n \in S^{d-1}$ *be in general position (i.e., no* $k$ *of them are in the same* $(k-2)$*-dimensional great circle) and let* $f_d(n)$ *be the number of partitions of* $\{x_1, \ldots, x_n\}$ *into two sets linearly separable from each other. Then*

$$f_d(n) = \sum_{k=0}^{d-1} \binom{n-1}{k}. \tag{2}$$

(This Lemma is an analog of Proposition 2.7 of T. H. Brylawski [1].)

Proof. If $n \leqslant d$, then it is easy to check that

$$f_d(n) = 2^{n-1}. \tag{3}$$

Let $S$ be the $(d-2)$-dimensional great circle in a hyperplane perpendicular to the vector $x_1$. Let $X$ be the projection of $\{x_2, \ldots, x_n\}$ into $S$ (along meridians with pole $x_1$). Now we will prove that

$$f_d(n) = f_d(n-1) + f_{d-1}(n-1). \tag{4}$$

In fact $f_d(n)$ equals the number of all the partitions of $\{x_2, \ldots, x_n\}$ which are linearly separable plus the number of those partitions of $\{x_2, \ldots, x_n\}$ which are separable by a hyperplane containing 0 and $x_1$. Clearly the first number is $f_d(n-1)$ and the second equals the number of partitions of $X$ which are linearly separable, which is $f_{d-1}(n-1)$. This proves (4).

It is routine to derive (2) from (3) and (4).

PROBLEM (**P 1298**). Suppose that $\mu$ is uniform over $S^{d-1}$. Can one estimate the expected value of the volume of the convex hull of $\{x_1, \ldots, x_n\}$?

**Acknowledgement.** I am indebted to T. H. Brylawski, A. Ehrenfeucht and R. B. Schnabel for information and discussions on this subject.

*REFERENCES*

[1] T. H. Brylawski, *A combinatorial perspective on the Radon convexity theorem*, Geometriae Dedicata 5 (1976), p. 459–466.

[2] R. D. Duda and P. E. Hardt, *Pattern Classifications and Scene Analysis*, Wiley, New York 1973.

[3] M. Minsky and S. Papert, *Perceptrons*, Massachussets Institute of Technology Press, Cambridge, 1972.

[4] J. Mycielski, *Linear dynamic approximation theory*, Journal of Approximation Theory 25 (1979), p. 369–383.

[5] –, *A learning theorem*, Abstracts of Papers Presented to the American Mathematical Society 1 (1980), p. 340.

[6] J. G. Wendel, *A problem in geometric probability*, Mathematica Scandinavica 11 (1962), p. 109–111.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF COLORADO
BOULDER, COLORADO, U.S.A.

*Reçu par la Rédaction le 06. 04. 1984*