Separating Bi-Chromatic Points by Parallel Lines

Tetsuo Asano John Hershberger János Pach Eduardo Sontag
Diane Souvaine Subhash Suri

March 24, 2001

Abstract

Given a 2-coloring of the vertices of a regular n-gon P, how many parallel lines are needed to separate the vertices into monochromatic subsets? We prove that $\lfloor n/2 \rfloor$ is a tight upper bound, and also provide an $O(n \log n)$ time algorithm to determine the direction that gives the minimum number of lines. If the polygon is a non-regular convex polygon, then $n-3$ lines may be necessary, while $n-2$ lines always suffice. This problem arises in machine learning and has implications about the representational capabilities of some neural networks.

1 Introduction

Let P be a regular polygon on n vertices. Given a 2-coloring of the vertices of P, how many parallel lines do we need to separate the vertices of P into monochromatic subsets? In other words, what is the smallest number $m(n)$ of parallel lines that always suffices to partition the vertices of P into $m(n)+1$ monochromatic subsets? Observe that an adversary picks the coloring of the vertices of P and we (the algorithm) choose the direction of lines. We show in this paper that $m(n) = \lfloor \frac{n}{2} \rfloor$. This bound is clearly tight since if the vertices of P are colored alternately *black* and *white*, then $n/2$ parallel lines are necessary. We present two different proofs of this fact. The first proof is somewhat shorter but it is non-constructive. The second proof also leads to an $O(n \log n)$ time algorithm.

This problem arises in machine learning. Briefly, for the class of neural nets having a single hidden layer consisting of threshold units and two inputs, but allowing direct connections from inputs to outputs, we show that the Vapnik-Chervonenkis (VC) dimension of the class is at least $4k+3$, where k is the number of hidden units. This provides an improvement by a factor of two over the corresponding result for nets that allow no such direct connections (Baum 88). In the context of learning theory, this has implications about the representational capabilities of the two types of nets. A different, worse-case, comparison, is given in (Sontag 89), which also shows a factor of two, and relates all this to sigmoidal responses.

A natural extension of the geometric problem is to consider the 2-coloring of arbitrary points in the plane. We show that even if the points form the vertices of a convex (but non-regular) polygon, $n-3$ parallel lines may be sometimes necessary. The upper bound is $n-2$ unless all the points are collinear, in which case $n-1$ lines are necessary and sufficient.

This abstract is organized as follows. In Section 2, we give the main result of this paper: a proof of the $\lfloor n/2 \rfloor$ bound for a regular n-gon along with an algorithm to find the direction of these lines. Section 3 presents the application of the geometric problem to neural-net theory. Section 4 outlines some extensions and generalizations of the geometric problem and makes some concluding remarks.

2 The Main Section

Let P be a convex regular n-gon in the plane with vertices $p_1, p_2, \ldots, p_n$. Let $\vec{l}$ be a directed line through the center of P, with orientation in the range from 0 to π such that the orthogonal projections of the p_i's onto $\vec{l}$ are distinct. Let $\pi(\vec{l})$ denote the permutation of the p_i's determined by the order of their projections along $\vec{l}$. Denoting the set of all distinct permutations $\pi(\vec{l})$ by Π, we clearly have $|\Pi| = n$.

Let $c : \{p_1, p_2, \ldots, p_n\} \to \{B, W\}$ be a 2-coloring of the vertices with colors black and white and let b represent the number of black vertices, and w the number of white.

2.1 The First Proof

For a given ordered pair of points p_i, p_j, there is exactly one permutation in which p_j immediately succeeds p_i. The total number of monochromatic ordered pairs over all the permutations is $b(b-1) + w(w-1)$, or $b^2 + w^2 - n$. The expected number of pairs of adjacent elements with the same color in an arbitrary permutation is $\frac{b^2+w^2-n}{n}$. Thus, at least one line $\vec{l}$ defines a permutation $\pi(\vec{l})$ that has at least $\left\lceil \frac{b^2+w^2}{n} \right\rceil - 1$ pairs of adjacent elements with the same color. But then, taking any collection of $n-1$ lines perpendicular to $\vec{l}$ and separating the vertices of the polygon from each other, we can remove at least $\left\lceil \frac{b^2+w^2}{n} \right\rceil - 1$ of them without creating a strip that contains vertices of both colors. Thus at most

$$n - 1 - (\left\lceil \frac{b^2 + w^2}{n} \right\rceil - 1) \;\le\; \left\lfloor \frac{n}{2} \right\rfloor$$

separating lines are needed.

2.2 The Second Proof

Given a line $\vec{l}$, the subscripts of the points in the permutation $\pi(\vec{l})$ are given by two interleaved sequences, one ascending and one descending. Away from the ends of the subscript permutation, the subscripts follow the pattern $\ldots, i, k-i, i+1, k-i-1, \ldots$, where k is a parameter that completely characterizes $\pi(\vec{l})$. (All subscript arithmetic is modulo n.) The ends of the subscript permutation are given by the equation $2i \equiv \{k \text{ or } k+1\} \pmod{n}$, depending on whether k is odd or even. The number of color differences between neighbors in $\pi(\vec{l})$ is the same as the number of lines perpendicular to $\vec{l}$ needed to partition the polygon vertices into monochromatic slabs.

To simplify our counting arguments, we consider the palindromic permutation $\pi(\vec{l})\pi(-\vec{l})$ given by concatenating $\pi(\vec{l})$ with its reversal. This sequence is composed of two full-length (length n) ascending and descending subscript sequences, interleaved as specified by the parameter k. In particular, the junction between $\pi(\vec{l})$ and its reversal follows the pattern of the interior of $\pi(\vec{l})$. The palindromic sequence has exactly twice as many color differences between neighbors as $\pi(\vec{l})$, even when considered as a cyclic sequence.

Let $D(i,j)$ be a function that is 1 if $c(p_i) \neq c(p_j)$ and 0 otherwise. Then the number of neighbor differences in $\pi(\vec{l})\pi(-\vec{l})$ is

$$\sum_{i=1}^{n} D(i,k-i) + \sum_{i=1}^{n} D(i,k+1-i),$$

that is, the sum of two convolutions. Note that the average value of a convolution is $2bw/n$:

$$\sum_{k=1}^{n}\sum_{i=1}^{n} D(i,k-i) = \sum_{i=1}^{n}\sum_{k=1}^{n} D(i,k-i) = 2bw.$$

Hence there is some k such that the number of neighbor differences in $\pi(\vec{l})\pi(-\vec{l})$ is at most $4bw/n$. It follows that the number of parallel lines needed to partition the polygon vertices is at most $\lfloor 2bw/n \rfloor \leq \lfloor n/2 \rfloor$.

2.3 The Algorithm

We can use the Fast Fourier Transform algorithm to compute the sum $\sum_{i=1}^{n} D(i,k-i)$ for every value of k in $O(n \log n)$ time, then find the value of k that gives the minimum number of slabs in additional $O(n)$ time.

In particular, let

$$a_i = \begin{cases} +1, & \text{if } c(p_i) = B, \\ -1, & \text{otherwise.} \end{cases}$$

Then $D(i,j) = \frac{1}{2}(1 - a_i a_j)$, and

$$\sum_{i=1}^{n} D(i,k-i) = \frac{n}{2} - \frac{1}{2}\sum_{i=1}^{n} a_i a_{k-i}.$$

We compute the convolution $\sum a_i a_{k-i}$ for every value of k using the FFT (see (AHU 74)), where the underlying field is the integers $\bmod p$, for some prime $p > 2n$ of the form $p = rn+1$. Linnik's "large sieve" proves that such a prime exists and has only $O(\log n)$ bits (and hence is suitable for computation); see (Linnik 41) or (Bombieri 74).

3 Application to Neural Networks

Let N be a positive integer (soon to be restricted to $N = 2$), and let $\mathcal{H} : \mathbf{R} \to \mathbf{R}$ denote the "hardlimiter" or "Heaviside" function: $\mathcal{H}(x) = 0$ if $x \leq 0$ and $\mathcal{H}(x) = 1$ if $x > 0$. We say that $f : \mathbf{R}^n \to \mathbf{R}$ is a (single hidden layer) *k-neuron net* (with possible direct connections from inputs to outputs, and N input neurons) if there are real numbers $w_0, w_1, \ldots, w_k, \tau_1, \ldots, \tau_k$ and vectors $v_0, v_1, \ldots, v_k \in \mathbf{R}^N$ such that, for all $x \in \mathbf{R}^N$,

$$f(x) \;=\; w_0 \;+\; v_0 \cdot x \;+\; \sum_{i=1}^{k} w_i\, \mathcal{H}(v_i \cdot x - \tau_i)$$

where the dot indicates inner product.

Theorem 3.1 *Fix $N = 2$. Then there is some set S of $4k + 3$ points in the plane with the following property: for each possible 2-coloring of S, there exists some k-neuron net f with the property that $f(x) > 0$ on the elements of one color and $f(x) < 0$ on the elements of the other color.*

Theorem 3.1 says that the VC (Vapnik-Chervonenkis) dimension of the set of all k-neuron nets with possible direct connections is at least $4k+3$; see for instance (Baum and Haussler 89) for the basic terminology on VC dimension. We conjecture that for general N a lower bound of about $2kN$ is achievable.

The proof of Theorem 3.1 is quite simple at this stage. Pick any k, and consider the particular set S consisting of the vertices of the convex regular $(4k + 3)$-gon in the plane. Assume that a 2-coloring of S is given. ¿From the main result we can conclude that there is some unit vector ν (for instance, the one that gives the orientation of $\vec{l}$) so that the real numbers $\nu \cdot x$ are arranged into at most $2k + 2$ monochromatic intervals (the "color" $\nu \cdot x$ is that of x). This reduces our problem to the case of finding a k-net with one input (that is, $N = 1$,) which separates $2k + 2$ bicolored disjoint intervals, and this is done in turn in (Sontag 89).

4 Extensions and Concluding Remarks

A natural extension of our geometric problem is to consider arbitrary point sets. We can construct a non-regular convex n-gon for which $m(n) \geq n-3$. Our construction uses exponentially spaced points on a parabola. This bound is close to the worst possible, since if n points are in general position, we always have $m(n) \leq n - 2$. (If all points are collinear, $m(n) = n - 1$.) We believe that $n - 2$ is the true upper bound for points in general position.

An interesting open question is to extend the regular n-gon construction to three or more dimensions. We now want to find a set for which any 2-coloring can be separated into monochromatic subsets by a small number of parallel hyperplanes. In three dimensions, we would like to know if there is a configuration S of n points such that for any 2-coloring of S, $\lfloor n/3 \rfloor$ parallel planes suffice to separate the points into monochromatic slabs. (This would give as a corollary a VC dimension of at least $6k + 5$ when $N = 3$.)

References

(AHU 74) Aho, A. V., Hopcroft, J. E., and Ullman, J. D., *The Design and Analysis of Computer Algorithms*, Addison-Wesley, 1974.

(Baum 88) Baum, E. B., "On the capabilities of multilayer perceptrons," *J.Complexity* **4** (1988): 193–215.

(Baum and Haussler 89) Baum, E. B., and D. Haussler, "What size net gives valid generalization?," *Neural Computation* **1** (1989): 151–160.

(Bombieri 74) E. Bombieri, "le Grand Crible," *Asterisque* **18**, 1974.

(Linnik 41), Y. V. Linnik, *Dokl. Akad. Nauk SSSR* **30** (1941): 292–294.

(Sontag 89) "Sigmoids distinguish better than Heavisides," *Neural Computation* **1** (1989): 470–472.