SIAM J. APPL. MATH.
Vol. 15, No. 2, March, 1967
Printed in U.S.A.

THE NUMBER OF LINEARLY INDUCIBLE ORDERINGS OF POINTS IN d-SPACE*

THOMAS M. COVER†

1. Introduction and summary. Consider a collection of n points $x_1, x_2, \cdots, x_n$ in Euclidean d-space E^d which are ordered according to orthogonal projection onto a reference vector $w \in E^d$. If π is a permutation of the set of integers $\{1, 2, \cdots, n\}$, we shall say that $w \in E^d$ induces the ordering π if

$$w \cdot x_{\pi(1)} > w \cdot x_{\pi(2)} > \cdots > w \cdot x_{\pi(n)}. \tag{1}$$

Conversely, the ordering π will be said to be *linearly inducible* if there exists such a w.

In this paper we demonstrate that there are precisely $Q(n, d)$ linearly inducible orderings of n points in general position in E^d, where $Q(n, d)$ satisfies the recurrence relation

$$Q(n + 1, d) = Q(n, d) + nQ(n, d - 1). \tag{2}$$

Since $n \geqq 2$ points can always be ordered in only two ways on a line, and since two points can be ordered in only two ways in $d \geqq 1$ dimensions, we see that

$$\begin{aligned} Q(n, 1) &= 2, \qquad && n \geqq 2, \\ Q(2, d) &= 2, && d \geqq 1, \end{aligned} \tag{3}$$

which, by iteration of (2), yields

$$Q(n, d) = 2 \sum_{k=0}^{d-1} {}_nS_k = 2\,[1 + \sum_{2 \leqq i \leqq n-1} i + \sum_{2 \leqq i < j \leqq n-1} ij + \cdots]\ (d \text{ terms}), \tag{4}$$

where ${}_nS_k$ is the sum of the ${}_{n-2}C_k = (n - 2)!/(n - 2 - k)!\,k!$ possible products of numbers taken k at a time without repetition from the set $\{2, 3, \cdots, n - 1\}$.

Thus we have found $Q(n, d)$, the number of ways that an art judge can rank n paintings, each having d numerical attributes, by forming weighted averages of the attributes. Our interest in this problem stems from work [1], [2], [3] on classification of vector-valued patterns by means of linear discriminants.

Notice that the number of linearly inducible orderings is independent of configuration (up to general position). Two examples, however, will show

* Received by the editors June 17, 1966.
† Stanford Electronics Laboratory, Stanford University, California.

that the "texture" of these orderings is not. In the first example, consider four points in the plane forming the vertices of a quadrangle as shown in Fig. 1A. (The w shown here induces the ordering (2, 3, 1, 4).) Any one of these points may be ranked first (or last) by an appropriate orientation of the weighting vector w. In the second example, let one point lie in the center of a triangle formed by three others, as shown in Fig. 1B. In this case, no linear weighting can rank the center point first or last. But in both cases, precisely $Q(4, 2) = 12$ of the 4! possible orderings of four points are linearly inducible. So the number of orderings is configuration-free, but the set of orderings is not, even under relabelling of the points.

We shall establish (2) and discuss some of the properties of $Q(n, d)$ in the next two sections. In the last section, general orderings induced by indexed families of nonlinear surfaces will be counted.

2. Theorem and proof.

DEFINITION.[1] A set of points is in *general position* in E^d if there exists no k-flat, $k < d$, containing $k + 2$ points, that is, there are no three points in a line, four points in a plane, etc.

THEOREM. *There are $Q(n, d)$ linearly inducible orderings of n points in general position in E^d.*

Proof. For a given set of n points $\{x_1, x_2, \cdots, x_n\}$ there is defined the open set $W(\pi)$ (a polyhedral convex cone) of all vectors w in E^d inducing the permutation π, where

$$W(\pi) = \{w: w\cdot x_{\pi(1)} > w\cdot x_{\pi(2)} > \cdots > w\cdot x_{\pi(n)}\}. \tag{5}$$

The theorem states that there are precisely $Q(n, d)$ nonempty sets of this form.

Equivalently, each difference vector $x_i - x_j$ defines a normal hyperplane

$$(x_i - x_j)^\perp = \{w: w\cdot(x_i - x_j) = 0\}, \tag{6}$$

and the collection of hyperplanes

$$\mathcal{H}_n = \{(x_i - x_j)^\perp : 1 \leqq i < j \leqq n\} \tag{7}$$

partitions E^d into $Q(n, d)$ nonempty cones, the nonempty $W(\pi)$'s. Each such cone is the equivalence class of vectors $w \in E^d$ inducing a given ordering. Thus the number of nonempty cones is the number of linearly inducible orderings.

Consider a new vector x_{n+1} such that $x_1, x_2, \cdots, x_n, x_{n+1}$ are in general position in E^d. Let $Q(n, d)$ denote the number of regions into which E^d is

[1] Compare this definition of general position with the definition which arises most frequently in geometrical considerations: A set of points is in general position in E^d if there exists no k-flat *through the origin*, $k < d$, containing $k + 1$ points.

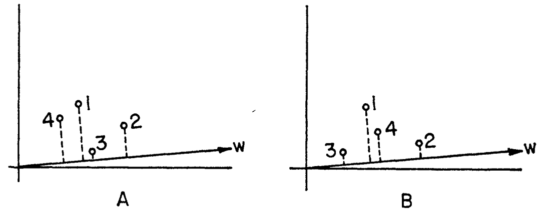

FIG. 1. *Two configurations, each having* 12 *linearly inducible orderings*

divided by $\mathcal{H}_n$. Let $n \geqq 2$. Assume that $Q(n', d')$ has been shown to be independent of configuration for $n' = 1, 2, \cdots, n$ and $d' = 1, 2, \cdots, d$. We shall find a relation for $Q(n + 1, d)$ and establish incidentally that this number is independent of configuration also.

The proof will follow when we have established the following three statements.

1. Each of the hyperplanes $(x_1 - x_{n+1})^\perp, (x_2 - x_{n+1})^\perp, \cdots, (x_n - x_{n+1})^\perp$ intersects precisely $Q(n, d - 1)$ regions created by $\mathcal{H}_n$.

2. No two of these hyperplanes intersect one another in the interior of a region created by $\mathcal{H}_n$. (The intersection of two such hyperplanes is contained in the set of boundaries of the regions into which E^d is partitioned by $\mathcal{H}_n$.)

3. Hence $nQ(n, d - 1)$ additional regions are formed, yielding $Q(n, d) + nQ(n, d - 1)$ in all.

Statement 1. Each plane $(x_i - x_j)^\perp$ in $\mathcal{H}_n$ intersects $(x_1 - x_{n+1})^\perp$ in a $(d - 2)$-space $(x_i - x_j)^\perp \cap (x_1 - x_{n+1})^\perp$. This $(d - 2)$-dimensional subspace of the $(d - 1)$-space $(x_1 - x_{n+1})^\perp$ has a normal in $(x_1 - x_{n+1})^\perp$ given by $(\hat{x}_i - \hat{x}_j)$, where $\hat{x}$ is defined to be the orthogonal projection of x into $(x_1 - x_{n+1})^\perp$. Thus $\mathcal{H}_n$ and $\hat{\mathcal{H}}_n$ induce the same partition of $(x_1 - x_{n+1})^\perp$.

Moreover, $\hat{x}_1, \hat{x}_2, \cdots, \hat{x}_n$ lie in general position in the $(d - 1)$-space $(x_1 - x_{n+1})^\perp$. Thus $\hat{\mathcal{H}}_n$ partitions $(x_1 - x_{n+1})^\perp$ into $Q(n, d - 1)$ cells. But we have shown that $\mathcal{H}_n$ and $\hat{\mathcal{H}}_n$ induce the same partition of $(x_1 - x_{n+1})^\perp$, and hence $\mathcal{H}_n$ partitions $(x_1 - x_{n+1})^\perp$ into $Q(n, d - 1)$ $(d - 1)$-dimensional cells. Since each cell into which $(x_1 - x_{n+1})^\perp$ has been partitioned serves as a boundary that divides into two cells one of the cells generated by $\mathcal{H}_n$ in d-space, we find that $Q(n, d - 1)$ new regions have been added to the $Q(n, d)$ old regions.

Statement 2. We shall now show that each new $(x_i - x_{n+1})^\perp$ creates precisely $Q(n, d - 1)$ new regions when added to $\mathcal{H}_n$ and the previously added $(x_j - x_{n+1})^\perp$'s. We are interested in the number of cells into which $(x_k - x_{n+1})^\perp$ is partitioned by the union of $\mathcal{H}_n$ and $(x_1 - x_{n+1})^\perp, (x_2 - x_{n+1})^\perp, \cdots, (x_{k-1} - x_{n+1})^\perp$. We note immediately from (6) that if $w \in (x_i - x_{n+1})^\perp$ and $w \in (x_k - x_{n+1})^\perp$, then

$$(8)\qquad \begin{aligned} w\cdot(x_i - x_{n+1}) &= 0, \\ w\cdot(x_k - x_{n+1}) &= 0, \end{aligned}$$

from which we see

$$(9)\qquad \begin{aligned} w\cdot x_i = w\cdot x_{n+1} &= w\cdot x_k\,, \\ w\cdot(x_i - x_k) &= 0. \end{aligned}$$

That is, $w \in (x_i - x_k)^\perp$. Thus $(x_i - x_{n+1})^\perp \cap (x_k - x_{n+1})^\perp$ is contained in $(x_i - x_k)^\perp$, which in turn is contained in the collection $\mathcal{H}_n$. Evidently, the new hyperplanes $(x_i - x_{n+1})^\perp$ and $(x_k - x_{n+1})^\perp$ intersect one another only in the boundaries of the regions previously formed by $\mathcal{H}_n$. Thus no regions can be formed by the intersection of $(x_k - x_{n+1})^\perp$ with $\mathcal{H}_n \cup \{x_1 - x_{n+1}\}^\perp \cup \cdots \cup \{x_{k-1} - x_{n+1}\}^\perp$ which could not already be formed by the intersection of $(x_k - x_{n+1})^\perp$ with $\mathcal{H}_n$ alone. And, as was argued in Statement 2 for $k = 1$, precisely $Q(n, d - 1)$ new regions are formed by intersecting $(x_k - x_{n+1})^\perp$ with $\mathcal{H}_n$ (and hence with $\mathcal{H}_n \bigcup_{i=1}^{k-1} (x_i - x_{n+1})^\perp$).

Statement 3. Each of the n additional planes $(x_1 - x_{n+1})^\perp$, $(x_2 - x_{n+1})^\perp$, $\cdots$, $(x_n - x_{n+1})^\perp$ creates $Q(n, d - 1)$ new regions. Hence,

$$(10)\qquad Q(n + 1, d) = Q(n, d) + nQ(n, d - 1).$$

Finally, (4) follows from the boundary conditions in (3).

The first few values of $Q(n, d)$ are summarized in Table 1.

3. Properties of $Q(n, d)$. We note that the terms ${}_nS_k$ of (4) are the coefficients in the generating function

$$(11)\qquad S_n(t) = \prod_{j=2}^{n-1} (1 + jt) = \sum_{k=0}^{\infty} {}_nS_k t^k.$$

Thus, for $d \geqq n - 1$,

$$(12)\qquad Q(n, d) = 2 \sum_{k=0}^{\infty} {}_nS_k = 2S_n(1) = n!.$$

TABLE 1

The number of linearly inducible orderings of n points in E^d

	n	d				
		1	2	3	4	5
$Q(n, d)$	2	2	2	2	2	2
	3	2	6	6	6	6
	4	2	12	24	24	24
	5	2	20	72	120	120
	6	2	30	172	480	720

Hence we see from (12) that, for $d \geqq n - 1$, all possible orderings of n points are linearly inducible. This, of course, is easily seen from simpler considerations.

The reader will observe that the terms ${}_nS_k$ are similar in definition to the Stirling numbers defined by

$$(13) \qquad {}_ns_k = \sum_{1 \leqq i_1 < i_2 < \cdots < i_k \leqq n} i_1 i_2 \cdots i_k .$$

However, we have not found a natural expression of $Q(n, d)$ in terms of the Stirling numbers and thus are denied an intriguing analogy between the number of linearly inducible orderings of n points in E^d with, for example, the number of permutations of n elements having fewer than k cycles.

$Q(n, d)$ may be given a probabilistic interpretation. Assume that each of the $n!$ permutations of $\{1, 2, \cdots, n\}$ is equiprobable and that a permutation π is drawn at random. Then the probability $P(n, d)$ that π is linearly inducible is given by

$$(14) \qquad P(n, d) = Q(n, d)/n! = \sum_{k=0}^{d-1} {}_nP_k ,$$

where the ${}_nP_k$ are coefficients of the generating function

$$(15) \quad P_n(t) = \left(\frac{1}{3} + \frac{2}{3}t\right)\left(\frac{1}{4} + \frac{3}{4}t\right) \cdots \left(\frac{1}{n} + \frac{(n-1)}{n}t\right) = \sum_{k=0}^{\infty} {}_nP_k t^k.$$

Now, $P_n(t)$ is the product of characteristic functions and hence is the characteristic function for the sum of $n - 2$ *independent* binary-valued random variables. Therefore, $P(n, d)$ may be interpreted as the probability that there are no more than $d - 1$ tails in $n - 2$ independent flips of coins having individual probabilities of heads $\frac{1}{3}, \frac{1}{4}, \cdots, 1/n$.

Finally, we remark that the number of linearly inducible orderings is related to the number of consistent solutions to a system of linear inequalities. Of the 2^n partitions of $x_1, x_2, \cdots, x_n$ (d-dimensional and in general position) into two subsets, exactly

$$(16) \quad C(n, d) = 2 \sum_{k=0}^{d-1} {}_{n-1}C_k = 2 \sum_{k=0}^{d-1} (n-1)!/(n-1-k)!\,k!$$

can be separated by a hyperplane through the origin [4], [5], [6].

4. General orderings. Suppose $x_1, x_2, \cdots, x_n$ are ranked, not according to their projections on a line, but according to their Euclidean distances from an arbitrary point $p \in E^d$. How many different orderings are induced as p ranges over E^d? Since

$$(17) \qquad \| x_{\pi(1)} - p \|^2 > \| x_{\pi(2)} - p \|^2 > \cdots > \| x_{\pi(n)} - p \|^2$$

is equivalent to

$$-p \cdot x_{\pi(1)} + \tfrac{1}{2}\| x_{\pi(1)} \|^2 > -p \cdot x_{\pi(2)} + \tfrac{1}{2}\| x_{\pi(2)} \|^2 > \cdots > -p \cdot x_{\pi(n)} + \tfrac{1}{2}\| x_{\pi(n)} \|^2,$$

we see that p induces the (distance) ordering π if and only if the augmented weighting vector $\tilde{w} = (-p, 1) \in E^{d+1}$ linearly induces the ordering π on the augmented vectors $\tilde{x}_i = (x_i, \tfrac{1}{2}\| x_i \|^2)$, $i = 1, 2, \cdots, n$; that is, if and only if

$$\tilde{w} \cdot \tilde{x}_{\pi(1)} > \tilde{w} \cdot \tilde{x}_{\pi(2)} > \cdots > \tilde{w} \cdot \tilde{x}_{\pi(n)} . \tag{18}$$

Thus n points in d-space having the property that no four points lie on a circle, no five points lie on a sphere, etc., may be ordered in $Q(n, d + 1)$ ways[2] with respect to their distances from an arbitrary point p.

This result is simply obtained as a special case of a more general point of view. If we define a general mapping ϕ from Euclidean d-space E^d into $E^{d'}$, we may define an ordering π to be *ϕ-linearly inducible* if there exists $w \in E^{d'}$ such that

$$w \cdot \phi(x_{\pi(1)}) > w \cdot \phi(x_{\pi(2)}) > \cdots > w \cdot \phi(x_{\pi(n)}). \tag{19}$$

Then, if $\phi(x_1), \phi(x_2), \cdots, \phi(x_n)$ are in general position in $E^{d'}$, there are $Q(n, d')$ such orderings. The distance (or spherical) ordering is obtained as a special case for $\phi(x) = (x, \tfrac{1}{2}\| x \|^2)$. Generalizations to such "nonlinear" orderings for a different class of problems—that of linearly separating two sets of vectors—are discussed in [6], but the same generalizations carry over to the linear-ordering problem.

REFERENCES

[1] H. D. Block, *The perceptron: a model for brain functioning, I*, Rev. Modern Phys., 34 (1962), pp. 123–135.

[2] N. J. Nilsson, *Learning Machines: Foundations of Trainable Pattern Classifying Systems*, McGraw-Hill, New York, 1965.

[3] B. Widrow, *Generalization and information storage in networks of Adaline 'neurons,'* Self-Organizing Systems, Spartan Books, Washington, D. C., 1962, pp. 442–459.

[4] L. Schläfli, *Gesammelte Mathematische Abhandlungen I*, Verlag Birkhäuser, Basel, Switzerland, 1950, pp. 209–212.

[5] J. Steiner, J. Reine Angew. Math. (Crelle), 1 (1826), pp. 349–364.

[6] T. M. Cover, *Geometrical and statistical properties of systems of linear inequalities with applications in pattern recognition*, IEEE Trans. Electronic Computers, EC-14 (1965), pp. 326–334.

[2] We count orderings according to increasing and decreasing distances from p as different orderings. Thus the $(d + 1)$th coordinate of w may be either positive or negative, and w may range over the entire space E^{d+1}.