Published in:
Proc. 24th ACM Symp. on Theory of Computing (1992), 170–177.

LINEAR DECISION TREES: VOLUME ESTIMATES AND TOPOLOGICAL BOUNDS

Anders Björner
Royal Institute of Technology
Stockholm, Sweden S-100 44

László Lovász
Eötvös Loránd University, Budapest, Hungary H-1088;
Princeton University, Princeton, NJ 08544

Andrew C.C. Yao,
Princeton University, Princeton, NJ 08544

Abstract. We describe two methods for estimating the size and depth of decision trees where a linear test is performed at each node. Both methods are applied to the question of deciding, by a linear decision tree, whether given n real numbers, some k of them are equal. We show that the minimum depth of a linear decision tree for this problem is $\Theta(n\log(n/k))$. The upper bound is easy; the lower bound can be established for $k = O(n^{1/4-\varepsilon})$ by a volume argument; for the whole range, however, our proof is more complicated and it involves the use of some topology as well as the theory of Möbius functions.

1. Introduction. Let P be a set in $\mathbb{R}^n$. Given a point x, we want to test if $x \in P$. Our model of computation is a *linear decision tree*, a rooted ternary tree T where each node v is associated with a linear function $l_v(x) = \sum_i a_i x_i + b$, and the three edges connecting an interior node to its descendants are labelled "$<$", "$=$" and "$>$". Starting from the root, we move down the tree; at each internal node v, we check whether $l_v(x) >=< 0$ and follow the appropriately labelled edge. Leaves are labelled YES and NO, and arriving at a leaf we read off the answer to the question "is $x \in P$?".

We denote the number of YES leaves and NO leaves by $\ell^+(T)$ and $\ell^-(T)$, respectively, and we denote by $\ell^+(P)$ and $\ell^-(P)$ the minimum of $\ell^+(T)$ and $\ell^-(T)$ over all linear decision trees for P.

Linear decision trees are sometimes surprisingly powerful devices; we mention here the result that some NP-complete problems like knapsack have polynomial size linear decision trees [MH]. Lower bound results on the size or depth of linear decision trees usually depend on counting the number of connected components of P or $\mathbb{R}^n \setminus P$ (see [DL]). In this paper we develop two methods that can be applied to obtain lower bounds if these sets are connected.

Our "test problem" is the k-equal-problem: given n real numbers $x_1, \dots, x_n$, decide if some k of them are equal. This problem has a trivial $O(n\log n)$ linear decision tree: just sort the elements, and compare the ith element with the $(i+k-1)$-st for every i (note that a comparison "$x < y$?" is a special case of a linear test). For $k = 2$, the proof that at least $\Omega(n\log n)$ linear tests (in particular, at least $\Omega(n\log n)$ comparisons) are needed was one of the first applications of the lower bound argument mentioned above [DL].

For $k > 2$, this lower bound technique fails since both P (the union of $\binom{n}{k}$ linear subspaces of the form $x_{i_1} = \dots = x_{i_k}$) and its complement are connected. Note that for large values of k, the linear decision tree complexity of this problem does decrease. To show this, assume (for simplicity) that $n = 2^m k$. We start with determining the $(2^{m-1}k)$-th largest element; this takes $O(n)$ comparisons. Then we go on with finding $(2^{m-2}k)$-th largest elements among those smaller and also among those larger than this element (ties are broken arbitrarily). In the j-th phase, those elements found so far split all elements into blocks of size $2^{m-j}k$, and we find the element of each block which splits it into two equal parts (where each element is counted in the block immediately before it).

After m phases, we have found the k-th, $(2k)$-th, ..., $2^m k$-th largest elements. Now if there are k equal elements, then one of these special elements must occur among them; therefore it is enough to compare each of them with $2k$ other elements (in the blocks immediately before and after them) to see if indeed this is the case.

Each phase takes $O(n)$ comparisons, so the total number of comparisons needed is $O(nm) = O(n\log(n/k))$. This determines a linear decision tree with depth $O(n\log(n/k))$ and (consequently) with size $(n/k)^{O(n)}$. Our main result shows that this is essentially best possible:

Main Theorem. *Every linear decision tree for the k-equal-problem has size at least $(n/k)^{\Omega(n)}$ and (consequently) depth at least $\Omega(n\log(n/k))$.*

Our first method is based on volume estimates, and it proves this result in the range $k < n^{1/4-\varepsilon}$ ($\varepsilon > 0$). Our second method, based on computing certain topological invariants (and substantially more complicated) will establish the result for any n and k. For this part, we will assume some familiarity with basic algebraic topology and its connections with combinatorics. For the former see e.g. [Mu] and for the latter [Bj].

2. Volume arguments. In this section we assume that P has finite volume. Let T be a linear decision tree for P, and let U and W denote the set of internal nodes and leaves of T, respectively. Let W^- and W^+ be the sets of NO-leaves and YES-leaves, respectively. Let, for each $w \in W$, P_w denote the set of inputs leading to leaf w. Each set P_w is a convex subset of $\mathbb{R}^n$. Since P is the union of all cells P_w ($w \in W^+$), this implies the following:

Proposition 2.1 *Let V be the maximum volume of a convex subset of P. Then*

$$\ell^+ \geq \frac{\mathrm{vol}(P)}{V}.$$

The volume of an arbitrary convex subset of P may be quite difficult to estimate. Using the recent result [Ba] that every convex body with volume V contains an ellipsoid with volume $n!n^{-n/2}(n+1)^{-(n+1)/2}\pi^{n/2}\Gamma(1+n/2)^{-1}V > n^{-n/2}V$ (the bound is tight for the simplex), we obtain:

Proposition 2.2 *Let $V_0(P)$ denote the maximum volume of an ellipsoid contained in P. Then*

$$\ell^+ \geq n^{-n/2}\frac{\mathrm{vol}(P)}{V_0(P)}.$$

To apply this bound to the k-equal-problem, we choose P to be $M_{n,k} = [0,1]^n \setminus V_{n,k}$, where $V_{n,k}$ is the set of all points in $[0,1]^n$ with at least k equal coordinates (so $\mathrm{vol}(P) = 1$). We will show that

Theorem 2.3 *For some constant $c > 0$,*

$$V_0(M_{n,k}) \leq \left(\frac{ck(\log n)}{\sqrt{n}}\right)^{2n}.$$

Let $\varepsilon > 0$ be any fixed constant and $2 \leq k \leq n^{(1/4)-\varepsilon}$. It follows from Theorem 2.3 and Proposition 2.2 that

$$\ell^+ \geq n^{-n/2}\left(\frac{n}{c^2k^2(\log n)^2}\right)^n \geq \left(\frac{n^{2\varepsilon}}{c^2(\log n)^2}\right)^n.$$

Hence the depth of any linear decision tree for $M_{n,k}$ is at least

$$\begin{aligned}\log_3 \ell^+ &\geq n\bigl(2\varepsilon\log_3 n - \log_3\log n - c'\bigr)\\ &> \varepsilon n\log_3 n > \varepsilon n\log_3(n/k).\end{aligned}$$

Proof of Theorem 2.3. Without loss of generality, we assume that $n \geq 10^8$ and $2 \leq k \leq \sqrt{n}$. Let $\mathcal{E}$ be any ellipsoid contained in $M_{n,k}$. Let $z = (z_1, \ldots, z_n)$ be the center of $\mathcal{E}$ (where we may assume that $z_1 \leq z_2 \leq \ldots \leq z_n$), and $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$, the lengths of its semi-axes.

Let $m = n - \lceil \frac{n}{\log n} \rceil$, and consider the $(n-m)$-dimensional affine subspace L passing through z and containing the longest $n-m$ semi-axes. We will prove that there exists a point $x \in L \cap V_{n,k}$ such that

$$\|x - z\| \leq \frac{8k^2(\log n)^2}{\sqrt{n}}. \tag{2.1}$$

This will imply $\lambda_{n-m} \leq 8k^2(\log n)^2/\sqrt{n}$, since otherwise $\mathcal{E}$ would contain x, contradicting the assumption $\mathcal{E} \subseteq M_{n,k}$. It will then follow that

$$\begin{aligned}\mathrm{vol}(\mathcal{E}) &= \left(\prod_{i=1}^n \lambda_i\right)\frac{\pi^{n/2}}{\Gamma(1+n/2)}\\ &\leq (\sqrt{n})^{n-m-1}\left(\frac{8k^2(\log n)^2}{\sqrt{n}}\right)^{m+1}\frac{\pi^{n/2}}{\Gamma(1+n/2)}\\ &\leq \left(\frac{ck(\log n)}{\sqrt{n}}\right)^{2n}.\end{aligned}$$

To prove (2.1), write L as the solution set in variables $(x_1,\dots,x_n)$ of the equations

$$\begin{aligned} a_{11}(x_1-z_1)+\dots+a_{1n}(x_n-z_n)&=0\\ &\vdots\\ a_{m1}(x_1-z_1)+\dots+a_{mn}(x_n-z_n)&=0 \end{aligned} \tag{2.2}$$

Let $D=\{i_1,i_2,\dots,i_m\}\subseteq\{1,2,\dots,n\}$ be a set of m columns such that $|\det(a_{ij}|_D)|$ is maximum. Clearly $|\det(a_{ij}|_D)|\neq 0$. Let $\overline{D}=\{1,\dots,n\}\setminus D$. Set $a=\lceil 2k/\log n\rceil$, and divide the columns into $\lceil n/a\rceil$ consecutive blocks, all of size a except possibly the last one. Observe that there are at least

$$\left\lceil\frac{|D|-k\lfloor n/a\rfloor-a}{a-k}\right\rceil\geq\frac{n}{8k(\log n)^2}$$

blocks among these that contain more than k columns of $\overline{D}$. Among these blocks B, at least one must have

$$|\max_{i\in B}z_i-\min_{i\in B}z_i|\leq\frac{8k(\log n)^2}{n}.$$

Let us fix such a block B, and write $\overline{D}\cap B=\{j_1,\dots,j_k,\dots\}$. Define $y=(\max_{i\in B}z_i+\min_{i\in B}z_i)/2$. Then, for all $i\in B$,

$$|z_i-y|\leq\frac{4k(\log n)^2}{n}. \tag{2.3}$$

Set $b_r=-\sum_{p=1}^{k}a_{rj_p}(y-z_{j_p})$ $(r=1,\dots,m)$. Consider the point $x=(x_1,\dots,x_n)$ where $x_i=y$ for $i\in\{j_1,\dots,j_n\}$, x_i $(i\in D)$ are defined by the following equations:

$$\begin{aligned} a_{1i_1}(x_{i_1}-z_{i_1})+\dots+a_{1i_m}(x_{i_m}-z_{i_m})&=b_1,\\ &\vdots\\ a_{mi_1}(x_{i_1}-z_{i_1})+\dots+a_{mi_m}(x_{i_m}-z_{i_m})&=b_p, \end{aligned}$$

and all other $x_i=z_i$.

As $\det(a_{ij}|_D)$ is maximum, we have for $i_s\in D$ by (2.3),

$$\begin{aligned} |x_{i_s}-z_{i_s}|&=\left|\sum_{p=1}^{k}\frac{\det(i_1,\dots,j_p,\dots,i_m)}{\det(i_1,\dots,i_m)}(y-z_{j_p})\right|\\ &\leq\sum_{p=1}^{k}|y-z_{j_p}|\leq\frac{4k^2(\log n)^2}{n}. \end{aligned}$$

Hence

$$\begin{aligned} \|x-z\|^2&=\sum_{p=1}^{k}|y-z_{j_p}|^2+\sum_{i\in D}|x_i-z_i|^2\\ &\leq k\left(\frac{4k(\log n)^2}{n}\right)^2+m\left(\frac{4k^2(\log n)^2}{n}\right)^2\\ &\leq\frac{32k^4(\log n)^4}{n}. \end{aligned}$$

∎

3. Topological invariants. In the sequel, we restrict our considerations to closed sets P. It is clear that among these, only polyhedra (sets arising as unions of finitely many convex polyhedra) have linear decision trees, so we shall assume that P is a polyhedron. The following fairly simple inequality gives a lower bound on the number of leaves of a linear decision tree for membership in a polyhedron P in terms of its Euler characteristic. We recall that this may be defined as follows. Consider a finite family $\mathcal{F}$ of (non-empty, not necessarily bounded) convex polyhedra such that

(1) if $Q\in\mathcal{F}$ then every face of Q is in $\mathcal{F}$;

(2) the intersection of any two members of $\mathcal{F}$ is a face of both.

Such a family is called a *convex cell complex.* If $\cup\mathcal{F}=P$, we call $\mathcal{F}$ a *convex cell decomposition* of P. If P is bounded, then $\mathcal{F}$ is a regular CW-complex (see [Mu, Bj]) and we define its Euler characteristic $\chi(P)$ by

$$\chi(P)=\sum_{Q\in\mathcal{F}}(-1)^{\dim(Q)}.$$

It is a basic fact of algebraic topology that the Euler characteristic is independent of the choice of the subdivision, and invariant under a number of morphisms (e.g. homotopy equivalence). In particular, any contractible polyhedron has Euler characteristic 1.

For unbounded polyhedra, we compactify the space by adding a single point ω "at infinity", and let $\hat{P}=P\cup\{\omega\}$. Consider a convex cell decomposition $\mathcal{F}$ of P; then $\mathcal{F}\cup\{\omega\}$ is a CW decomposition of $\hat{P}$, and

$$\chi(\hat{P})=1+\sum_{Q\in\mathcal{F}}(-1)^{\dim(Q)}$$

(the first term comes from the point ω as a cell). In the following we set $\hat{P} = P$ for bounded polyhedra (for convenience), and define $\varepsilon_P = 1$ os P is unbounded and $= 0$ otherwise.

Theorem 3.1 *Let P be a polyhedron in $\mathbb{R}^n$. Then*

$$\ell^+(P) \geq |\chi(\hat{P}) - \varepsilon_P|,$$
$$\ell^-(P) \geq |\chi(\hat{P}) - \varepsilon_P + (-1)^{n-1}|.$$

Proof. Let T be a linear decision tree for P. Using notation from section 2, we note that each set P_w ($w \in W$) is convex and polyhedral, but not necessarily closed; in fact, P_w is open in its affine hull: the affine hull A_w of P_w is obtained as the intersection of those hyperplanes $l_u(x) = 0$ which tested with equality along the path from the root to w, and the remaining strict inequalities along this path define P_w. We denote by $\bar{P}_w$ the closure of P_w and by ∂P_w, the boundary of P_w in A_w. If P_w is bounded, then $\bar{P}_w$ is a ball (in fact a convex cell) and ∂P_w is its bounding sphere. If P_w is unbounded then there are two cases:

(i) $\hat{\bar{P}}_w$ is a ball and $\hat{\partial P_w}$ is its bounding sphere, if $P_w \neq A_w$,

(ii) $\hat{\bar{P}}_w$ is a sphere and $\partial \hat{P}_w = \omega$, if $P_w = A_w$.

We have $P = \cup\{P_w : w \in W^+\}$; but the polyhedra $\bar{P}_w$ (even together with their faces) do not form a convex cell decomposition in general. So to relate to the Euler characteristic, we consider the following finer decomposition. Our linear decision tree T determines an arrangement of affine hyperplanes $\mathcal{A}_T = \{H_u\}$, where $H_u = \{x \in \mathbb{R}^n : l_u(x) = 0\}$ for each inner node $u \in U$. These hyperplanes subdivide $\mathbb{R}^n$ to a number of relatively open convex (polyhedral) regions. These regions, together with their faces, partition $\mathbb{R}^n$ (points in the same class behave the same way in all tests on the tree). Let Δ denote the set of these classes. Their closures give a convex cell decomposition of $\mathbb{R}^n$.

For each leaf w, let Δ_w, $\bar{\Delta}_w$ and $\partial\Delta_w$ be the collections of cells in Δ contained in P_w, $\bar{P}_w$ and ∂P_w, respectively. Let Δ' denote the set of cells in Δ contained in P. Then the closures of cells in Δ' form a convex cell decomposition of P, and hence

$$\sum_{C \in \Delta'} (-1)^{\dim(C)} = \chi(\hat{P}) - \varepsilon_P.$$

Now we can partition this sum according to the YES-leaves:

$$\begin{aligned}
\chi(\hat{P}) - \varepsilon_P &= \sum_{w \in W^+} \sum_{C \in \Delta_w} (-1)^{\dim(C)} \\
&= \sum_{w \in W^+} \left(\sum_{C \in \bar{\Delta}_w} (-1)^{\dim(C)} - \sum_{C \in \partial\Delta_w} (-1)^{\dim(C)} \right) \\
&= \sum_{w \in W^+} \left((\chi(\hat{\bar{P}}_w) - \varepsilon_{P_w}) - (\chi(\hat{\partial P}_w) - \varepsilon_{P_w}) \right) \\
&= \sum_{w \in W^+} (-1)^{\dim(P_w)}.
\end{aligned}$$

In fact, $\bar{\Delta}_w$ forms a convex cell decomposition of $\bar{P}_w$, $\partial\Delta_w$ forms a convex cell decomposition of ∂P_w, and $\hat{\bar{P}}_w$ and $\hat{\partial P}_w$ are a ball and a sphere (or conversely) as stated in (i) or (ii) above. Therefore $\chi(\hat{\bar{P}}_w) - \chi(\hat{\partial P}_w) = (-1)^{\dim(P_w)}$ follows from $\chi(\text{ball}) = 1$ and $\chi(n\text{-sphere}) = 1 + (-1)^n$.

The bound on ℓ^+ follows immediately from this equation, as the right hand side is at most $W^+ = \ell^+(T)$. For ℓ^-, observe that the cells in Δ, together with ω, form a CW decomposition of the n-sphere, and so $\sum_{C \in \Delta} (-1)^{\dim(C)} = \chi(S^n) - 1$. Hence

$$\begin{aligned}
\sum_{C \in \Delta \setminus \Delta'} (-1)^{\dim(C)} &= \chi(S^n) - 1 - (\chi(\hat{P}) - \varepsilon_P) \\
&= -\chi(\hat{P}) + \varepsilon_P + (-1)^n.
\end{aligned}$$

From here the argument is just like that for ℓ^+, by partitioning the left-hand sum according to the NO-leaves. ∎

Remarks. 1. It is in general not easy to compute the Euler characteristic of a combinatorially presented polyhedron. In a special class of polyhedra representable as the union of affine subspaces, however, combinatorial tools are available to determine this Euler characteristic, as we shall show below.

2. The Euler characteristic of a polyhedron may be small even if its structure is very complicated. For example, if the polyhedron is star-shaped (i.e. it has a point v such that the segment connecting v to any other point is contained in the polyhedron), then its Euler characteristic is 1. We may get better bounds by intersecting P with a hyperplane H; it is easy to see that every linear

decision tree for P yields a linear decision tree for $P \cap H$ (as a polyhedron in H) of the same size.

3. It is a very natural problem to extend these results to algebraic decision trees (trees where linear tests are replaced by polynomial tests with bounded degree). The "component count" version of this method was initiated, for algebraic decision trees, by Steele and Yao [SY] and improved by Ben-Or [BO]. One may hope that the results of Milnor and Thom [Mi,Th] can be used in combination with the ideas of this paper.

4. Subspace arrangements. By an *affine subspace arrangement* we mean a finite family $\mathcal{A} = \{K_1, \ldots, K_m\}$ of affine subspaces in $\mathbb{R}^n$. Such an arrangement is called *central*, if each K_i is a linear subspace, i.e., it contains 0. We may assume that no K_i contains another K_j. In this section we derive combinatorial bounds on the size of a linear decision tree testing membership in the union of subspaces in an arbitrary arrangement.

For an arrangement $\mathcal{A} = \{K_1, \ldots, K_m\}$, central or affine, let $V_\mathcal{A} = \cup_{i=1}^m K_i$ and $M_\mathcal{A} = \mathbb{R}^n \setminus V_\mathcal{A}$. Then $V_\mathcal{A}$ is an unbounded polyhedron (if $\dim V_\mathcal{A} > 0$); let $\hat{V}_\mathcal{A}$ denote its compactification with an element ω. The set $M_\mathcal{A}$ is a d-dimensional manifold, which is connected in case $\dim K_i \leq n-2$ for all $1 \leq i \leq m$. We know from the previous section that $\ell^+(V_\mathcal{A}) \geq |\chi(\hat{V}_\mathcal{A}) - 1|$. Unfortunately, $\chi(\hat{V}_\mathcal{A})$ is still quite difficult to determine or estimate, and it will be useful to consider $V'_\mathcal{A}$, the intersection of $V_\mathcal{A}$ with a very large cube Q ("very large" means that it intersects every non-empty intersection of subspaces in $\mathcal{A}$). Note that

$$\ell^+(V_\mathcal{A}) \geq 3^{-n} \ell^+(V'_\mathcal{A})$$

(first test if $x \in Q$, and if so, test if $x \in V_\mathcal{A}$), and hence it suffices to find lower bounds on $\chi(V'_\mathcal{A})$. Here we use that an optimal tree for the closed n-cube has 3^n YES-leaves, corresponding to its non-empty faces.

The partially ordered set $L_\mathcal{A}$ of all intersections $K_{i_1} \cap \ldots \cap K_{i_j}, 1 \leq i_1 < \ldots < i_j \leq m$, ordered by reverse inclusion, is called the *intersection lattice* of $\mathcal{A} = \{K_1, \ldots, K_m\}$. This lattice has least element $\hat{0} = \mathbb{R}^n$, and greatest element $\hat{1} = \cap\mathcal{A} = K_1 \cap \ldots \cap K_m$. Its atoms (elements covering $\hat{0}$) are the subspaces K_i. Some very useful aspects of the topology of the spaces $V_\mathcal{A}$ and $M_\mathcal{A}$ are encoded into the lattice $L_\mathcal{A}$, as the following results show.

The *order complex* $\Delta(P)$ of a poset P is the simplicial complex whose vertex set is P and whose simplices are the chains $x_1 < \ldots < x_k$ in P. Note that if the elements of P are vertices of a $(|P|-1)$-dimensional simplex, and each chain is represented by the convex hull of its elements, then we get a convex cell complex, and the union of these simplices is a polyhedron, which we shall also denote by $\Delta(P)$.

Proposition 4.1 *Let $\mathcal{A}$ be an affine arrangement with $\cap\mathcal{A} = \emptyset$. Then $V_\mathcal{A}$, $V'_\mathcal{A}$ and $\Delta(L_\mathcal{A} - \{\hat{0}, \hat{1}\})$ are of the same homotopy type.*

Proof. The homotopy equivalence of $V_\mathcal{A}$ and $\Delta(L_\mathcal{A} - \{\hat{0}, \hat{1}\})$ follows from the nerve theorem [Bj, (10.7)] applied to the covering of $V_\mathcal{A}$ by the subspaces K_i, together with the cross-cut theorem [Bj, (10.8)] applied to the cross-cut of atoms K_i in $L_\mathcal{A}$. For $V'_\mathcal{A}$ instead of $V_\mathcal{A}$ this follows similarly. ∎

As a consequence, we obtain that for any affine arrangement $\mathcal{A}$ with $\cap\mathcal{A} = \emptyset$, we have

$$\begin{aligned} \ell^+(V_\mathcal{A}) &\geq 3^{-n} \ell^+(V'_\mathcal{A}) \geq 3^{-n} |\chi(V'_\mathcal{A})| \\ &= 3^{-n} |\chi(\Delta(L_\mathcal{A} - \{\hat{0}, \hat{1}\}))|. \end{aligned} \tag{4.1}$$

For arrangements with $\cap\mathcal{A} \neq \emptyset$ (equivalently, for central arrangements) this method does not give any information since in this case $V'_\mathcal{A}$ is star-shaped and its Euler characteristic is 1. We can apply, however, our lower bound to the intersection of $V'_\mathcal{A}$ with an appropriate hyperplane H. In this case, every linear decision tree for $V'_\mathcal{A}$ would yield a linear decision tree for $V'_\mathcal{A} \cap H$ (as a set in H) with the same number of YES-leaves and NO-leaves. Applying Proposition 4.1 to the intersections, we get

Corollary 4.2 *Let $\mathcal{A} = \{K_1, \ldots, K_m\}$ be a central arrangement, and H an affine hyperplane avoiding 0 whose translate through the origin contains $K_1 \cap \ldots \cap K_m$ but no other intersection in $L_\mathcal{A}$. Then $V_\mathcal{A} \cap H$, $V'_\mathcal{A} \cap H$ and $\Delta(L_\mathcal{A} - \{\hat{0}, \hat{1}\})$ are of the same homotopy type.* ∎

From here we obtain in the central case:

$$\begin{aligned} \ell^+(V_\mathcal{A}) &\geq 3^{-n} \ell^+(V'_\mathcal{A} \cap H) \geq 3^{-n} |\chi(V'_\mathcal{A} \cap H)| \\ &= 3^{-n} |\chi(\Delta(L_\mathcal{A} - \{\hat{0}, \hat{1}\}))|. \end{aligned} \tag{4.2}$$

The adventage of expressing the Euler characteristic of $V'_{\mathcal{A}}$ or $V'_{\mathcal{A}} \cap H$ in terms of the Euler characteristic of posets is that we can now invoke the powerful theory of Möbius functions (see [S], Chapter 3, for an introduction to this theory). The key is the following theorem of Ph. Hall: *the value of the Möbius function $\mu(x,y)$ for a pair $x < y$ in a poset P is one less than the Euler characteristic χ of the order complex of the open interval $(x,y) = \{z \in P : x < z < y\}$*; see [S, p. 120] or [Bj, (9.13)]. Therefore our results (4.1) and (4.2) can be summed up as follows:

Theorem 4.3 *Let $\mathcal{A}$ be an affine subspace arrangement in $\mathbb{R}^n$. Then*

$$\ell^+(V_{\mathcal{A}}) \geq 3^{-n} \left|\mu_{L_{\mathcal{A}}}(\hat{0},\hat{1})+1\right|. \qquad (4.3)$$

For central arrangements the Euler characteristic of $\hat{V}_{\mathcal{A}}$ can be computed in terms of the intersection lattice $L_{\mathcal{A}}$, using a formula of Goresky and MacPherson [GM, p.238] for the Betti numbers of $M_{\mathcal{A}}$ and Alexander duality:

$$\chi(\hat{V}_{\mathcal{A}}) = 1 + \sum_{x \in L_{\mathcal{A}},\ x > \hat{0}} (-1)^{\dim(x)-1} \mu_{L_{\mathcal{A}}}(\hat{0},x).$$

This leads, using Theorem 3.1, to the following bound:

Theorem 4.4 *Let $\mathcal{A}$ be a central subspace arrangement in $\mathbb{R}^n$. Then*

$$\ell^+(V_{\mathcal{A}}) \geq \left| \sum_{x \in L_{\mathcal{A}},\ x > \hat{0}} (-1)^{\dim(x)} \mu_{L_{\mathcal{A}}}(\hat{0},x) \right|. \quad (4.4)$$

Lower bounds for $\ell^-(V_{\mathcal{A}})$ similar to those in Theorems 4.3 and 4.4 can be derived from Theorem 3.1 by parallel reasoning.

We do not know if, for a central arrangement, either one of the lower bounds in Theorems 4.3 and 4.4 for ℓ^+ is always better than the other, although in all examples we have checked (4.4) is far better than (4.3). On the other hand, (4.3) is simpler to evaluate. Although the evaluation of the Möbius function is a difficult problem in general, its theory is fairly well developed and we shall illustrate the use of this theory in the analysis of the k-equal problem in the next section.

5. Application to the k-equal-problem. The motivation for this paper is to give lower bounds for the size of linear decision trees for the problem: "Given real numbers $x_1, x_2, \ldots, x_n$ decide whether k of them are equal". This is a special case of the problem of testing membership in the union of an arrangement: Let $\mathcal{A}_{n,k}$ denote the arrangement in $\mathbb{R}^n$ of the $\binom{n}{k}$ subspaces of dimension $n-k+1$ given by the equations $x_{i_1} = x_{i_2} = \ldots = x_{i_k}$, for $1 \leq i_1 < \ldots < i_k \leq n$. It is clear that the k-equal problem is to decide whether $x \in V_{\mathcal{A}_{n,k}}$ for points $x \in \mathbb{R}^n$. Therefore by Theorem 4.3 the Möbius function of the corresponding intersection lattice will provide lower bounds for the size of decision trees. Such intersection lattices have a very concrete combinatorial description.

Let Π_n denote the lattice of partitions of the set $\{1,2,\ldots,n\}$ ordered by refinement, and for $1 \leq k \leq n-1$ let $\Pi_{n,k}$ be the subposet of partitions with no block sizes in $\{2,3,\ldots,k\}$. In particular, $\Pi_{n,1} = \Pi_n$. We observe that $\Pi_{n,k}$ is itself a lattice, whose join-operation is the same as that of Π_n and whose meet-operation is that of Π_n (coarsest common refinement) followed by breaking all blocks of size $\leq k$ into singletons.

Proposition 5.1 *The intersection lattice of $\mathcal{A}_{n,k}$, $2 \leq k \leq n$, is isomorphic to the lattice $\Pi_{n,k-1}$. If under this isomorphism a subspace $x \in \mathcal{A}_{n,k}$ corresponds to a partition in $\Pi_{n,k-1}$ with j blocks, then* $\dim(x) = j$. ∎

We define $\mu_{n,k}$ for all $n,k \geq 1$ as follows. If $n > k$, let $\mu_{n,k} = \mu_k(\hat{0},\hat{1})$, i.e. the Möbius function μ_k of $\Pi_{n,k}$ computed over all of $\Pi_{n,k}$. Let also $\mu_{1,k} = 1$ and $\mu_{2,k} = \ldots = \mu_{k,k} = 0$. We denote by μ the Möbius function of Π_n.

We derive various formulas for the Möbius function μ_k; the last one of these will give a good enough estimate to settle the linear decision tree problem for all values of k up to a constant.

The following formula is a special case of a result of Crapo (see [S, p. 159]):

Lemma 5.2 *Let $n \geq k$, and let Y denote the set of partitions in Π_n with all classes of size k or less. Then $\mu_{n,k} = \sum_{y \in Y} \mu(y,1)$.* ∎

Let $S_k(n,j)$ denote the number of partitions of an n-set into j parts of size at most k. Then we have by the well-known formula for the Möbius function of the partition lattice,

Corollary 5.3 $\mu_{n,k} = \sum_{j=1}^{n} (-1)^{j-1}(j-1)! S_k(n,j)$.

(For $n \le k$, this follows from $S_k(n,j) = S(n,j)$ and the well-known identity

$$\sum_{j=1}^{n}(-1)^{j-1}(j-1)!S(n,j) = 0.)$$

Unfortunately, this alternating sum does not provide any easy way to see the order of magnitude of $\mu_{n,k}$, so we turn to generating functions. Let

$$F_k(x) = \sum_{n=1}^{\infty}\mu_{n,k}\frac{x^n}{n!} \quad \text{and} \quad p_k(x) = \sum_{i=0}^{k}\frac{x^i}{i!}.$$

By a well-known formula for the exponential generating function of partitions into parts with specified sizes, we have

Lemma 5.4 $\displaystyle\sum_{n=1}^{\infty}S_k(n,j)\frac{x^n}{n!} = \frac{1}{j!}(p_k(x)-1)^j.$ ∎

This leads to the following nice formula:

Theorem 5.5 $F_k(x) = \ln p_k(x)$.

Proof. By Corollary 5.3 and Lemma 5.4,

$$\begin{aligned}
F_k(x) &= \sum_{n=1}^{\infty}\mu_{n,k}\frac{x^n}{n!} \\
&= \sum_{n=1}^{\infty}\sum_{j=1}^{n}(-1)^{j-1}(j-1)!S_k(n,j)\frac{x^n}{n!} \\
&= \sum_{j=1}^{\infty}\sum_{n=j}^{\infty}(-1)^{j-1}(j-1)!S_k(n,j)\frac{x^n}{n!} \\
&= \sum_{j=1}^{\infty}(-1)^{j-1}(j-1)!\frac{1}{j!}(p_k(x)-1)^j = \ln p_k(x).
\end{aligned}$$

∎

For the derivative we get

$$F_k'(x) = \frac{p_k'(x)}{p_k(x)} = \frac{p_k(x)-(x^k/k!)}{p_k(x)} = 1-\frac{x^k}{k!p_k(x)}.$$

Let $\alpha_1,\ldots,\alpha_k$ be the roots of p_k. These are distinct, since

$$p_k'(\alpha_i) = p_k(\alpha_i) - \frac{\alpha_i^k}{k!} = -\frac{\alpha_i^k}{k!} \neq 0.$$

In terms of these numbers, we now have a formula for $\mu_{n,k}$ which will be good enough for our purposes.

Theorem 5.6 $\displaystyle\mu_{n,k} = -(n-1)!\sum_{i=1}^{k}\alpha_i^{-n}.$

Proof. We can write

$$\frac{1}{p_k(x)} = \sum_{i=1}^{k}\frac{A_i}{x-\alpha_i}, \qquad A_i = \frac{1}{p_k'(\alpha_i)} = \frac{-k!}{\alpha_i^k}.$$

Thus

$$\begin{aligned}
F_k'(x) &= 1 - x^k\sum_{i=1}^{k}\frac{-1}{\alpha_i^k(x-\alpha_i)} \\
&= 1 - x^k\sum_{i=1}^{k}\frac{1}{\alpha_i^{k+1}}\frac{1}{1-(x/\alpha_i)} \\
&= 1 - x^k\sum_{i=1}^{k}\frac{1}{\alpha_i^{k+1}}\sum_{h=0}^{\infty}\left(\frac{x}{\alpha_i}\right)^h \\
&= 1 - \sum_{n=k+1}^{\infty}x^{n-1}\sum_{i=1}^{k}\alpha_i^{-n}.
\end{aligned}$$

∎

This formula reveals that $\mu_{n,k}$ is not always large, e.g., $\mu_{n,2} = 0$ whenever $n \equiv 2 \pmod 4$. However, we now show that $\mu_{n,k}$ is large often enough.

Theorem 5.7 *For all n,k with $1 \le k \le n/2$ there exists an m such that $n-k+1 \le m \le n$ and $|\mu_{m,k}| > (m-1)!k^{-m-1}$.*

Proof. (This is actually a special case of Turán's principle [Tu].) We have $\min_i|\alpha_i| < k$, since $\prod_i \alpha_i = \pm k!$. Assume that $|\alpha_1| < k$. Write

$$q(x) = \prod_{i=2}^{k}(x-\alpha_i) = \sum_{j=0}^{k-1}b_jx^j.$$

Then

$$\begin{aligned}
\sum_{j=0}^{k-1}b_j\frac{\mu_{n-j,k}}{(n-1-j)!} &= -\sum_{j=0}^{k-1}\sum_{i=1}^{k}b_j\alpha_i^{-n+j} \\
&= -\sum_{i=1}^{k}q(\alpha_i)\alpha_i^{-n} = -q(\alpha_1)\alpha_1^{-n} \\
&= -k!p_k'(\alpha_1)\alpha_1^{-n} = \alpha_1^{-n+k}.
\end{aligned}$$

Hence there is a j, $0 \le j \le k-1$, such that

$$\left|b_j\frac{\mu_{n-j,k}}{(n-j-1)!}\right| > \frac{1}{k}|\alpha_1|^{-n+k}.$$

One can show by induction that

$$b_j = -k!\alpha_1^{-j-1}p_j(\alpha_1),$$

and hence it is not difficult to see that

$$|b_j| \le k^{k+1}|\alpha_1|^{-j-1}.$$

Thus

$$\begin{aligned}|\mu_{n-j,k}| &> \frac{1}{k}(n-1-j)!|\alpha_1|^{-n+k}|b_j|^{-1}\\ &\ge (n-1-j)!|\alpha_1|^{-n+j+1+k}k^{-k-2}\\ &> (n-1-j)!k^{-n+j-1}.\end{aligned}$$

∎

Let $\ell^+_{n,k} = \ell^+(V_{\mathcal{A}_{n,k}})$. Then we get:

Proposition 5.8 *For all n,k with $1 \le k \le n/4$, we have $\ell^+_{n,k+1} \ge (n-k)!(3k)^{k-n-2}$.*

Proof. We use the monotonicity property that $\ell_{n,k+1} \ge \ell_{m,k+1}$ if $n \ge m$. Choosing $n \ge m \ge n-k+1$ as in Theorem 5.7, we get

$$\begin{aligned}&\ell^+_{n,k+1} \ge \ell^+_{m,k+1} \ge 3^{-m}|\mu_{m,k}+1|\\ &\ge (m-1)!(3k)^{-m-1} \ge (n-k)!(3k)^{k-n-2}.\end{aligned}$$

(The last inequality uses that $v!(3k)^{-v}$ is increasing as a function of v for $v \ge 3k$). ∎

Proof of the Main Theorem: For $k \ge n/100$ we need to prove a bound of $\Omega(n)$ for the depth of the tree, which follows by elementary reasoning (or we could argue that for fixed k, the minimum depth increases in n). So we may assume that $k < n/100$. Then the linear decision tree complexity of the k-equal problem for n numbers is at least:

$$\begin{aligned}\log_3 \ell^+_{n,k} &\ge (n-k)(\log_3(n-k)-1)\\ &\quad -(n-k+2)\log_3 k-(n-k+2)\\ &\ge (n-k)\log_3\frac{n-k}{k}-2n \ge \frac{1}{10}n\log\frac{n}{k}\end{aligned}.$$

∎

References:

[Ba] K. Ball: Volume ratios and a reverse isoperimetric inequality (preprint 1991).

[Bj] A. Björner: Topological Methods, in: *Handbook of Combinatorics* (ed. R. Graham, M. Grötschel, L. Lovász), North-Holland, to appear (preprint 1989).

[BO] M. Ben-Or: Lower bounds for algebraic computation trees, *15th ACM STOC* (1983) 80–86.

[DL] D. Dobkin and R. Lipton, On the complexity of computations under varying sets of primitives, in: *Automata Theory and Formal Languages*, (ed. H. Bradhage), Lecture Notes in Computer Science **33**, Springer-Verlag 1975, 110–117.

[GM] M. Goresky and R. MacPherson: *Stratified Morse Theory*, Ergebnisse, Band 14, Springer-Verlag, Berlin (1988)

[Mi] J. Milnor: On the Betti numbers of real algebraic varieties, *Proc. Amer. Math. Soc.* **15** (1964) 275–280.

[Mu] J.R. Munkres: *Elements of Algebraic Topology*, Addison-Wesley, Menlo Park (1984).

[MH] F. Meyer auf der Heide: A polynomial linear search algorithm for the n-dimensional knapsack problem, *J. ACM* **31** (1984), 668–676.

[S] R.P. Stanley: *Enumerative Combinatorics*, Vol. 1, Wadsworth & Brooks/Cole, Monterey (1986).

[SY] M. Steele and A. Yao, Lower bounds for algebraic decision trees, *J. Algorithms* **3** (1982), 1–8.

[Th] R. Thom: Sur l'homologie des variétés algébriques réelles, in: *Differential and Algebraic Topology* (ed. S. S. Cairns), Princeton Univ. Press, Princeton (1965).

[Tu] P. Turán: *Eine neue Methode in der Analysis und deren Anwendungen*, Akad. Kiadó, Budapest (1953).