Arrangements of Lines and Hyperplanes

Subhash Suri

October 29, 2019

1 Arrangements and Geometric Duality

- Geometric Duality plays an important role in CG. The connection comes from the Cartesian idea that a point in $\mathbb{R}^2$ is specified by two coordinates $(x, y)$, as is a line $y = mx + c$, specified by its slope and intercept $(m, c)$.

- Similarly, points and hyperplanes in $d$-space require $d$ parameters, and therefore can be mapped into each other.

- Often, a problem about points becomes much easier to solve by viewing it as a problem about lines (or hyperplanes), and vice versa.

- Degeneracy Testing: A simple example of the power of duality is deciding whether an input set of points is non-degenerate, meaning no 3 are collinear. What is the most efficient algorithm for it?

- The trivial upper bound is $O(n^3)$, and it is not clear if a better algorithm is possible. We will show that an $O(n^2)$ time algorithm using duality and line arrangements.

- The complexity of this innocent looking problem is a major open problem in theoretical computer science.

- There are many different duality transforms, but a simple one is the following, called standard projective duality.

  - Each point $p = (a, b)$ maps to the dual line $y = ax - b$, and we denote the dual as $p^*$.

  - Each line $L$ with equation $y = mx + c$ maps to the dual point $(m, -c)$, which we denote as $L^*$.

  - That is, the $x$ coordinate maps to the slope, and the $y$ coordinate maps to the (negative) intercept.
• Geometric Properties of the Duality.

  – **Incidence Preserving:** Suppose \( p \) is a point, \( L \) a line, and \( p \) lies on \( L \). Then, the dual point \( L^* \) lies on the dual line \( p^* \).

  – **Order Preserving:** Suppose \( p \) is a point, \( L \) a line, and \( p \) lies above the line \( L \). Then, the dual point \( L^* \) lies above the dual line \( p^* \).

• These conditions are easy to check algebraically.

  – Suppose \( p = (a, b) \) and \( L : y = mx + c \). Then, incidence in primal space means that \( b = ma + c \).

  – The dual point for \( L \) is \((m, -c)\), and dual line for \( p \) is \( y = ax - b \). Their incidence in dual means \(-c = am - b\), which implies \( b = ma + c \). Done!

  – Similarly, the “aboveness” in primal space means \( b > ma + c \), while in dual space it means \(-c > am - b\), which are the same.

• Therefore, under the projective transform, if there are \( k \) collinear points in the input, then their “dual” lines must all intersect in a common point.

• We show that checking for multiple lines meeting in a common point can be done in \( O(n^2) \) time.

**Line Arrangement.**

• The subdivision of the plane induced by a finite set of lines \( \mathcal{L} \) is called the arrangement \( \mathcal{A}(\mathcal{L}) \).

• This subdivision consists of faces (convex polygons), edges (straight line segments, or half-rays), and vertices (points).

• The line arrangement is called **simple** if no two lines are parallel and no three meet in a common point.

• Although lines are unbounded, we can regard the arrangement bounded by conceptually placing them inside a sufficiently large rectangular box. (Such a box can be computed in \( O(n \log n) \) time. How?)

• The arrangement can also be viewed as a planar graph (by adding a vertex at infinity) and for algorithmic purposes we assume a DCEL representation.

• **Size Lemma:** A simple arrangement of \( n \) lines in \( \mathbb{R}^2 \) has \( \binom{n}{2} \) vertices, \( n^2 \) edges, and \( \binom{n}{2} + n + 1 \) faces.
• **Proof.** The number of vertices and edges is easy to count—each vertex involves a pair of lines, and each line is split into $n$ segments or half rays. For the number of faces, we use a sweepline argument. Imagine a vertical sweepline placed at $-\infty$. It intersects exactly $n + 1$ (unbounded) faces, meeting each line exactly once. As we sweep the line, we swap one old face for a new one, exactly when we sweep over a vertex, which happens \( \binom{n}{2} \) times, thus giving the bound for the faces.

• **Non-Simple Arrangements.** The complexity of an arrangement is maximum when it is a simple arrangement.

**Other Applications**

• The complexity of hyperplane arrangements is also relevant in many other settings, e.g. linear classifiers.

• In data classification and machine learning, we employ linear classifier rules, which geometrically means a hyperplane test: classify a data point $x$ as “+” if $hx > 0$, and “−” otherwise.

• Figure of classification.

• If we have a $n$ such classifiers (hyperplanes), each + on one side and − on the other, how many “dichotomies” do we get?

• In general, $n$ binary rules create $2^n$ size partition.

• But if the rules are hyperplanes, then the size (given by the arrangement) is only $O(n^d)$.

• **Linearization of non-linear functions.** Discrimination rules of higher order can be “lifted” linear rules in higher dimension.

• For instance, if you order $k$ polynomials in $d$ variables, then there are $\binom{k+1}{d}$ terms. So, we can linearize this as a $O(k^d)$-dimensional hyperplane.

**Horizon Theorem**

• An important step in finding an optimal algorithm for constructing the line arrangement is the following Horizon Theorem, which is of independent interest.

• Let $\mathcal{A}(\mathcal{L})$ be the arrangement of $n$ lines.

• Consider some line $l$ (not necessarily in $\mathcal{L}$), and let $h(l)$ be the total size of all the faces that intersect $l$.

• The size of a face is the number of edges bounding it (combinatorial complexity).
• The horizon meets $O(n)$ faces, and any single face can have size $\Theta(n)$, so the naive upper bound for $h(l)$ is $O(n^2)$.

• The Horizon Theorem proves that $h(l)$ is in fact $O(n)$, which is optimal.

• **Horizon Theorem.** The total number of edges in the horizon of $l$ (not necessarily from $\mathcal{L}$) is at most $6n$.

**Proof of the Horizon Theorem (ETH notes)**

• Assume that $l$ is horizontal, and none of the other lines are horizontal. (Otherwise, rotate the coordinate axes.)

• Split the boundary of each face in the horizon at its top and bottom vertices, and orient all edges from bottom to top.

• The edges that have a horizon cell to their right are called *left-bounding* for that cell. Similarly, define the right-bounding edges.

• We will show there are $\leq 3n$ left-bounding edges, by induction on $n$.

• The base case $n = 1$ is trivial: $1 \leq 3$. Assume it holds for $n - 1$.

• Consider the *rightmost line* $r \in \mathcal{L}$ that intersects $l$.

• By induction, in the reduced arrangement $\mathcal{A}(\mathcal{L} \setminus r)$, the horizon of $l$ has at most $3(n - 1)$ left-bounding edges.

• Adding back $r$ creates at most 3 new left-bounding edges for the horizon of $l$.

  1. Two of these edges belong to the rightmost cell in $l$’s horizon in $\mathcal{A}(\mathcal{L} \setminus r)$—at most two edges (call them $a$ and $b$) of the rightmost cell are intersected by $r$, and split into two, both of which may be left-bounding.

  2. The third edge is contributed by $r$ itself. The line $r$ cannot contribute a left-bounding edge to any cell other than the rightmost: to the left of $r$, the edges induced by $r$ form right-bounding edges only; and to the right of $r$, all other cells touched by $r$ are shielded away from $l$ by $a$ or $b$.

• Thus, the total number of left-bounding edges in the horizon is $3 + (3n - 3) = 3n$. QED.
Another Proof

- We count only the number of edges in the cells that lie above the line \( l \).
- For each such cell of the horizon, classify its edges as follows.
  - Floor: this is the edge defined by the horizon base line \( l \)
  - Roof: the two edges incident to the highest vertex of \( f \), where the height is measured from \( l \).
  - The remaining edges of \( f \) are divided in two “left walls” and “right walls.”
- The total contribution of \( f \) to \( h(l) \) is (1) floor, (2) roof, and (3) left and right walls.
- **Lemma:** Every line in \( L \) appears at most once as a left wall in \( h(l) \) and at most once as a right wall in \( h(l) \).
- **Proof.** By contradiction.
  - Suppose there exists a line \( g \) that appears at least twice as a left wall, appearing as edge \( e_1 \) in face \( f_1 \), and \( e_2 \) in face \( f_2 \).
  - Since \( e_1 \) is a left wall, consider the “left roof” edge \( f_1 \), namely, the edge that comes before the top vertex of \( f_1 \).
  - Let \( g \) be the line defining this left root edge.
  - It is easy to see \( g \) must “shield” \( e_2 \) away from \( l \), preventing it from appearing on the horizon.
- Thus, \( h(l) \leq 2n + (n + 1) + (2n - 2) \), where the first term accounts for left and right walls (each line appearing at most once), the second term accounts for the floor (at most \( n + 1 \) edges), and the third accounts for the roofs (each of the \( n - 1 \) faces have at most 2 roof lines).

Constructing the Line Arrangement

- The Horizon Theorem provides an easy “incremental” method for constructing the line arrangement.
- By induction, assume the arrangement for the first \( n - 1 \) lines has been built.
- We then add the \( n \)th line \( l \), modifying the arrangement in time proportional to the horizon of \( l \).
- The arrangement not touching the horizon is unaffected by the new addition.
• Since the horizon size is $O(n)$, we can afford to linearly scan it and modify the DCEL representation of the arrangement in $O(n)$ time.

• The total time to add $n$ lines is $O(n^2)$.

Applications of the Line Arrangement: Degeneracy Testing.

• Dualize the $n$ points, and construct the arrangement $A(\mathcal{L})$.

• There are $O(n^2)$ vertices and $O(n^2)$ edges.

• Using the DCEL representation, we can check how many edges incident to each vertex. If any vertex has more than 4 edges, we have a degeneracy.

Application: Smallest Area Triangle

• Given $n$ points, find the smallest-area triangle formed by a triple. (Degeneracy is the special case of zero area.)

• There are $\Theta(n^3)$ distinct triangles. Can we find the smallest of those without inspecting them all?

• We use the projective duality again, and use the fact that duality preserves “vertical distances” between points and lines.

• Suppose $p = (a, b)$ is a point, and $L : y = mx + c$ a line.

• The “vertical” distance from $p$ to $L$ is $b - (ma + c)$.

• In the dual space, the vertical distance from the dual point $(m, -c)$ to the line $y = ax - b$ is $-c - (am - b) = b - (ma + c)$.

• Instead of triangles formed by all “triples” of points independently, let’s consider all triangles formed by triples whose “base edge” is defined by the pair $p_1, p_2$.

• Suppose the apex of one such triangle is $z$.

• Then, the area of this triangle is $\frac{1}{2} \times \text{base} \times \text{height}$ .

• If $\triangle p_1p_2z$ is the smallest-area triangle with base $p_1p_2$, then the strip defined by the lines $p_1p_2$ and its parallel translate through $z$ does not contain any other point of the input.

• Therefore, among all the points of the input, the point $z$ must have the minimum vertical distance to the line defined by $p_1p_2$. 
Our $O(n^2)$ time algorithm, therefore, has the following form: (1) consider each of the $O(n^2)$ pairs of points. (2) for each such pair $p_1, p_2$, find the point $z$ (among the rest) whose vertical distance to the line $p_1p_2$ is smallest. (3) keep track of the smallest area triangles among these $O(n^2)$ candidates.

- We do this computation in the dual space. Form the arrangement defined by the dual lines of the $n$ input points.
- Each line defined by an input pair of points $(p_1, p_2)$ becomes a “vertex” in the arrangement of dual lines.
- The apex $z$ with minimum vertical distance to a line $(p_1, p_2)$ is the “line” in the dual space whose vertical distance to the “vertex” dual of $(p_1, p_2)$ is minimum.
- We can do this by simple finding the line-directly-above-each vertex and the line-directly-below-each vertex of the arrangement.
- The arrangement has size $O(n^2)$, and the entire computation takes $O(n^2)$ time.

Levels in Arrangements and Ham Sandwich Theorem

- Let $R$ and $B$ be two disjoint point sets in the plane, in general position.
- There exists a line $l$ that simultaneously bisects both $R$ and $B$, that is, each (open) halfplane defined by $l$ contains at most half the points of each color.
- We may assume that both $R$ and $B$ contain an odd number of points; otherwise, arbitrarily discard one point, and the bisector of the reduced set remains a valid bisector for the original.
- Assume that no two points in $R \cup B$ have the same $x$-coordinate; otherwise, perform an appropriate rotation of the plane.
- Let $R^*$ and $B^*$ denote the dual lines corresponding to our points.
- In the arrangement of the red lines $\mathcal{A}(R^*)$, consider the median level. this is the level containing an equal number of lines above and below it.
- By the order-preserving property of the duality, each point on this level is dual to a line that bisects the red point set $R$.
- The duality maps $x$-coordinates of points to the slopes of lines, and so the median level records the bisection lines as the slope of the line goes from 0 to $\pi$. 
Clearly, at the two extremes, $0 = \pi$, the bisection line is the same, and thus the median level belongs to the same line in $R$, namely, one whose $x$-coordinate is the median in the input set.

Similarly, we have the median level for the blue line arrangement $A(B^*)$, also with a same blue line defining the start and the end of the median level.

However, since no two points in $R \cup B$ have the same $x$-coordinate, these two lines $l_r, l_b$ are not parallel, and must intersect.

As a result, we can conclude that the median levels of $A(R^*)$ and $A(B^*)$, which are piecewise linear continuous functions, also intersect.

Any point in their common intersection is dual to a line that simultaneously bisects both $R$ and $B$.

2 Ham Sandwich Theorem $d$ Dimensions

Given $d$ measurable objects in $d$-dimensional Euclidean space, there always exists a $(d - 1)$-dim hyperplane that simultaneously bisects each of the objects. That is, each side of the halfplane contains exactly half the volume of each object.

In 3-dim, single knife cut split bread, ham, and cheese evenly.

Proof using Borsuk–Ulam theorem, which says that any continuous function $f : S^d \rightarrow R^{d-1}$ has two antipodal points $p, q \in S^d$ such that $f(p) = f(q)$.

Let $A_1, A_2, \ldots, A_d$ denote the $d$ objects that we wish to simultaneously bisect.

Let $S$ be the unit $(n-1)$-sphere embedded in $d$-dimensional Euclidean space $R^n$, centered at the origin.

For each point $p$ on the surface of the sphere $S$, we can define a continuum of oriented affine hyperplanes (not necessarily centred at $O$) perpendicular to the (normal) vector from the origin to $p$, with the positive side of each hyperplane defined as the side pointed to by that vector (i.e. it is a choice of orientation).

By the intermediate value theorem, every family of such hyperplanes contains at least one hyperplane that bisects the bounded object $A_d$: at one extreme translation, no volume of $A_d$ is on the positive side, and at the other extreme translation, all of $A_d$’s volume is on the positive side, so in between there must be a translation that has half of $A_d$’s volume on the positive side.
• If there is more than one such hyperplane in the family, we can pick one canonically by choosing the midpoint of the interval of translations for which \( A_d \) is bisected.

• Thus we obtain, for each point \( p \) on the sphere \( S \), a hyperplane \( \pi(p) \) that is perpendicular to the vector from the origin to \( p \) and that bisects \( A_d \).

• Now we define a function \( f \) from the \((n-1)\)-sphere \( S \) to \((n-1)\)-dimensional Euclidean space \( \mathbb{R}^{d-1} \) as follows:

\[
f(p) = (x_1, x_2, \ldots, x_{d-1}),
\]

where \( x_i = \text{vol of } A_i \) on the positive side of \( \pi(p) \).

• This function \( f \) is continuous. By the Borsuk–Ulam theorem, there are antipodal points \( p \) and \( q \) on the sphere \( S \) such that \( f(p) = f(q) \).

• Antipodal points \( p, q \) correspond to hyperplanes \( \pi(p), \pi(q) \) that are equal except that they have opposite positive sides.

• Thus, \( f(p) = f(q) \) means that the volume of \( A_i \) is the same on the positive and negative side of \( \pi(p) \) (or \( \pi(q) \)), for \( i = 1, 2, \ldots, d-1 \).

• Thus, \( \pi(p) \) (or \( \pi(q) \)) is the desired ham sandwich cut that simultaneously bisects the volumes of \( A_1, A_2, \ldots, A_d \).

2.1 Generalizations

• The original theorem works for at most \( d \) collections, where \( d \) is the number of dimensions. If we want to bisect a larger number of collections without going to higher dimensions, we can use, instead of a hyperplane, an algebraic surface of degree \( k \), i.e., an \((d-1)\)-dimensional surface defined by a polynomial function of degree \( k \). We then have the following generalization:

• Given \( \binom{k+n}{n} - 1 \) measures in an \( d \)-dim space, there exists an algebraic surface of degree \( k \) which bisects them all.

• The generalization is proved by mapping the \( d \)-dim plane into a \( \binom{k+n}{n} - 1 \) dimensional plane, and then applying the original theorem.

• For instance, if \( d = 2 \) and \( k = 2 \), then the 2-dim plane is mapped to a 5-dim plane via \((x, y) \rightarrow (x, y, x^2, y^2, xy)\).