HKUST Theoretical Computer Science Center Research Report HKUST-TCSC-2003-07

The Convex Hull for Random Lines in the Plane.

Mordecai Golin[1], Stefan Langerman[2⋆], and William Steiger[3]

[1] Computer Science, University of Science and Technology, Hong Kong
[2] Département d'Informatique, Université Libre de Bruxelles
[3] Computer Science, Rutgers University

Abstract. An arrangement of n lines chosen at random from R^2 has a vertex set whose convex hull has constant (expected) size.

1 Introduction and Summary.

Let $L = \{\ell_1, \ldots, \ell_n\}$ be a set of lines in general position in R^2. The vertex set $V = \{\ell_i \cap \ell_j, i < j\}$ of this arrangement has size $O(n^2)$ and we are interested in |Conv(V)|, the number of extreme points of its convex hull. As observed by Atallah [1],

$$|\mathrm{conv(V)}| \leq 2\mathrm{n},$$

a fact that sparked algorithmic interest in the hull of line arrangements [2], [3], [4].

Suppose the lines are *chosen uniformly at random*. The specific model we use is that the lines in L are the duals of n points chosen uniformly and independently from $[0,1]^2$, under the familiar duality that maps a point $P = (x, y)$ to the line $TP = \{(u, v) : v = xu + y\}$ and maps the non-vertical line $\ell = \{(x, y) : y = mx+b\}$ to the point $T\ell = (-m, b)$. To get n randomly chosen lines $\ell_1, \ldots, \ell_n$, we start with points $P_i = (x_i, y_i)$, $i = 1, \ldots, n$ chosen uniformly and independently from $[0,1]^2$ and then take

$$\ell_i = \{(u, v) : v = x_i u + y_i\}, i = 1, \ldots, n.$$

We give a simple proof of the following statement.

Theorem 1 *Let L be a set of n lines chosen uniformly at random. There is a constant $c > 0$ so that*

$$E(|\mathrm{Conv(V)}|) < \mathrm{c}; \tag{1}$$

A similar statement holds when the lines are dual to points chosen uniformly from other convex polygons. We have not tried to estimate c carefully, but we believe it is smaller than 10.

Devroye and Toussaint [4] proved the same result when the lines are polar duals to points chosen at random from a wide range of radially symmetric distributions. The two models for random lines are quite different, and both are natural. Our proof is simple and elementary. Much more is needed to establish the statement in [4].

⋆ Chargé de recherches du FNRS

2 The Proof

Choose $P_1 = (x_1, y_1), \ldots, P_n = (x_n, y_n)$ uniformly and independently from the unit square and numbered so $x_i < x_{i+1}$. We may assume that the points chosen are in general position in the sense that no three points lie on a common line and no two points have the same x coordinate, because these degeneracies occur with zero probability. The random lines are $\ell_i = \{(u, v) : v = x_i u + y_i\}$, $i = 1, \ldots, n$, and the vertex set is $V = \{\ell_i \cap \ell_j, i < j\}$. It is better to consider Conv(V) in the primal. A vertex $\ell_i \cap \ell_j \in V$ is an extreme point of Conv(V) only if $j = (i \bmod n) + 1$, so we seek the convex hull of the n vertices formed by the lines in L with successive slopes (in the radial ordering of the lines by slope). In the primal we seek lines through successive points P_i, P_{i+1} which are part of the upper or lower envelope of these lines. Specifically let r_i be the line joining P_i and P_{i+1}, $i = 1, \ldots, n-1$, and $r_i(t)$ the y-coordinate of the point on r_i with x-coordinate t. Write $U(t) = \max_i r_i(t)$ and $L(t) = \min_i r_i(t)$ for the upper and lower envelopes of the r_i. Then

$$|\mathrm{Conv(V)}| = |\mathrm{UP}| + |\mathrm{DN}|$$

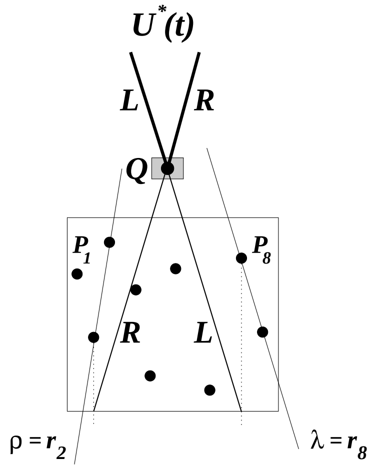
where we write **UP** for the set $\{i : r_i \text{ has a segment in } U(t)\}$ and **DN** for the set $\{i : r_i \text{ has a segment in } L(t) \text{ but not in } U(t)\}$. We only show how to bound the expected size of **UP**, the argument for **DN** being similar.

Let A_i be the event that r_i meets $U(t)$ and Z_i, its indicator. We will show that $E(Z_i) \le c/n$ for an appropriate constant $c > 0$.

We cover A_i by simpler events whose probability is easier to estimate. Let λ be the $r_i, i > 3n/4$, of minimum slope, let ρ be the $r_i, i \le n/4$, of max slope, and

$$\sigma = \min\left[|slope(\lambda)|, |slope(\rho)|\right]$$

Let L be the line through $(x_{3n/4}, 0)$ of slope $-\sigma$, R the line through $(x_{n/4}, 0)$ of slope σ, and $U^*(t) = \max_t [L(t), R(t)]$. Clearly $U^*(t) \le U(t)$. Write $Q = (x, y) = L \cap R$ for the intersection of L and R (see figure).



Let $E_1, \dots, E_{n+1}$ be i.i.d. standard exponential random variables with partial sums

$$S_i = E_1 + \cdots + E_i,$$

and let $y_1, \dots, y_n$ be i.i.d. uniforms on $[0,1]$. It is familiar that the joint distribution of $S_1/S_{n+1}, \dots, S_n/S_{n+1}$ is the same as that of $x_{(1)}, \dots, x_{(n)}$, the order statistics of a sample $x_1, \dots, x_n$ of i.i.d. uniforms. We think of the points $P_1, \dots, P_n$ ordered by x-coordinate, with $P_i = (S_i, y_i)$, $i = 1, \dots, n$. In this way the unit square is replaced by the random rectangle with corners at $(0,0)$ and $(S_{n+1}, 1)$. The law of large numbers implies that as $j \to \infty$

$$\mathrm{Prob}[|S_j - j| < \varepsilon j] \geq 1 - 1/j$$

for any $\varepsilon > 0$, and that the point $Q = (x, y)$ where L and R meet satisfies $|x - n/2| < \varepsilon n$ and $|y - \sigma n/4| < \varepsilon n$ with probability at least $1 - 1/n$. We therefore assume that x and y, the coordinates of Q, satisfy those inequalities.

For $i \leq n/4$, $A_i \subseteq B_i \cup C_i$, where B_i is the event that r_i has slope less than $-\sigma$ and C_i the event that r_i is above Q; if neither B_i nor C_i occur, r_i does not meet $U^*(t)$, so it can't meet $U(t)$. To estimate the probabilities of B_i and C_i, note that the line r_i joining P_i and P_{i+1} has slope $s_i = (y_{i+1} - y_i)/E_{i+1}$. The numerator has density $f(t) = 1 - |t|$, $t \in [-1, 1]$. Therefore for $t > 0$

$$\begin{aligned} P[s_i \leq -t] = P[s_i \geq t] &= \int_0^{1/t} P[y_{i+1} - y_i \geq ts] e^{-s} ds \\ &= \int_0^{1/t} \frac{(1-ts)^2}{2} e^{-s} ds = t^2(1 - e^{-1/t}) + \frac{1}{2} - t \qquad (2) \\ &= \frac{1}{6t} - \frac{1}{24t^2} + \frac{1}{120t^3} + \cdots \end{aligned}$$

Denote this function by $g(t)$.

Write $M = \max(s_2, s_4, \dots, s_{n/4})$ and note that $\rho = \max_{i \leq n/4} s_i \geq M$. Also, because the random variables $s_{2j}, j = 1, \dots, n/8$ are independent, we have

$$P[M \leq t] = P[s_{2i} \leq t]^{n/8} = (1 - P[s_{2i} \geq t])^{n/8} = [1 - g(t)]^{n/8}.$$

Similarly, writing $m = \min(s_{3n/4+2}, s_{3n/4+4}, \dots, s_n)$, $\lambda = \min_{i > 3n/4} s_i \leq m$, and because the even slopes are independent,

$$P[m \geq -t] = P[s_{2i} \geq -t]^{n/8} = (1 - P[s_{2i} \leq -t])^{n/8} = [1 - g(t)]^{n/8}.$$

These combine to show

$$P[\sigma \leq t] \leq P[min(M, |m|) \leq t] = 2[1 - g(t)]^{n/8} - [1 - g(t)]^{n/4}. \qquad (3)$$

The intuition from (2) and (3) is that M has median $\Theta(n)$ and s_i exceeds this with probability $O(1/n)$.

More formally, $C_i = \{r_i \text{ above } Q\} \subset \{s_i \geq \sigma/2 - \epsilon\}$ for some small $\epsilon > 0$, an event with probability at most

$$P[\sigma \leq K] + \int_K^\infty P[t/2 - \epsilon \leq s_i \leq t] h(t) dt,$$

for any positive K, where we write $h(t)$ for the density of $\min(M, |m|)$ obtained by differentiating the right hand side of (3). Therefore

$$P[C_i] \leq P[\sigma \leq K] + \int_{K-2\epsilon}^\infty P[s_i \geq t/2] h(t) dt.$$

Note that for large t, $1/(7t) < g(t) < 1/(5t)$ and for $K = an/\log n$, $P[\sigma \leq K] \leq 2e^{-ng(K)/8} \leq 1/n$. Applying these estimates,

$$\begin{aligned} P[C_i] &\leq \frac{1}{n} + \int_K^\infty g(t/2) n e^{-ng(t)/8}/(6t^2) dt \\ &\leq \frac{1}{n} + \int_K^\infty \frac{n}{120 t^3 e^{n/(40t)}} dt, \end{aligned}$$

an expression bounded by c/n.

We also have $P[B_i] = P[s_i < -\sigma] = P[s_i > \sigma] < P[C_i]$, and so $P[A_i] < 2P[C_i] < 2c/n$, for $i = 1, \ldots, n/4$. The same is true for $i = 3n/4 + 1, \ldots, n$ by symmetry. Finally, when $|i - n/2| \leq n/4$, r_i meets $U(t)$ only if $s_i > \sigma - \epsilon$ or $s_i < -\sigma + \epsilon$, and both these events have probability less than $P[C_i] < c/n$. $\square$

Acknowledgements

We thank an anonymous referee for valuable suggestions.

References

1. M. Atallah. Computing the Convex Hull of Line Intersections. *J. Algorithms 7*, 285-288 (1986).
2. B. Battacharya, H. Everett, and G. Toussaint. A Counter-Example to a Dynamic Algorithm for Convex Hulls of Line Arrangements. *Pattern Rec. Letters 12*, 145-147 (1991).
3. J. Boreddy. An Incremental Computation of Convex Hull of Planar Line Intersections. *Pattern Rec. Letters 11*, 541-543 (1990).
4. L. Devroye and G. Toussaint. Convex Hulls for Random Lines. *J. Algorithms 14*, 381-394 (1993).