MATH 262A: DISCRETE GEOMETRY NOTES

RAY TSAI

Contents

1. Sums vs Product

Definition 1.1. The *crossing number* of a graph G, denoted $\mathrm{cr}(G)$, is the minimum number of crossing pair of edges over all possisble drawings of G in the plane.

Lemma 1.2 (Crossing Lemma)**.** *Let $G = (V, E)$ be a graph. If $|E| \geqslant 4|V|$, then*

$$\mathrm{cr}(G) \geqslant \frac{|E|^3}{64|V|^2}.$$

Theorem 1.3. *Let A be a set of n distinct real numbers. Then* $\max\{|A + A|, |A \cdot A|\} = \Omega(n^{5/4})$.

Proof. Denote $A + A = \{s_1, s_2, \ldots, s_x\}$ and $A \cdot A = \{p_1, p_2, \ldots, p_y\}$. Let L be the set of lines $v = a_i(u - a_j)$ for $a_i, a_j \in A$. Construct the graph $G = (V, E)$ with $V = (A + A) \times (A \cdot A)$ and $\{(s_i, p_i), (s_j, p_j)\} \in E$ if and only if there exists a line $l \in L$ such that (s_i, p_i) and (s_j, p_j) are consecutive points on l. Notice that each line passes through at least $n - 1$ points in V, so $|E| \geqslant (n - 1)|L| = \Omega(n^3)$. If $|E| < 4|V|$, then

$$4|A + A| \cdot |A \cdot A| = 4|V| > |E| = \Omega(n^3).$$

But then either $|A + A| = \Omega(n^{3/2})$ or $|A \cdot A| = \Omega(n^{3/2})$. Thus we may assume $|E| \geqslant 4|V|$. By the crossing lemma,

$$\frac{|E|^3}{64|V|^2} \leqslant \mathrm{cr}(G) \leqslant |L|^2 \leqslant n^4.$$

Rearranged, we have

$$|V|^2 \geqslant \frac{|E|^3}{64n^4} = \Omega(n^5).$$

The result now follows. ∎

2. Crossing Lemma

In this section we prove the Crossing lemma mentioned in the previous section.

Lemma 2.1. *Let $G = (V, E)$ be a graph. Then* $\operatorname{cr}(G) \geqslant |E| - 3|V|$.

Proof. Suppose not. We may assume $|E| \geqslant 3|V|$, otherwise we are done. Remove edges from each crossing until we have a planar graph. Since $\operatorname{cr}(G) < |E| - 3|V|$, we removed less than $|E| - 3|V|$ edges. But then the planar graph has more than $|E| - (|E| - 3|V|) = 3|V|$ edges, contradicting Euler's theorem. □

Lemma 2.2 (Crossing Lemma)**.** *Let $G = (V, E)$ be a graph. If $|E| \geqslant 4|V|$, then*

$$\operatorname{cr}(G) \geqslant \frac{|E|^3}{64|V|^2}.$$

Proof. For any graph H, define $X_H = \operatorname{cr}(H) - |E(H)| + 3|V(H)|$. By the crossing lemma we know $X_H \geqslant 0$. Consider the drawing of G in $\mathbb{R}^2$ with $\operatorname{cr}(G)$ crossings. Let $S \subseteq V$ be a set vertices where each vertex is chosen independently with probability $p \in [0, 1]$. Let $G' = G[S]$ be the induced subgraph on S. Then

$$\mathbb{E}[X_{G'}] = \mathbb{E}[\operatorname{cr}(G')] - \mathbb{E}[|E(G')|] + 3\mathbb{E}[|V(G')|] = \mathbb{E}[\operatorname{cr}(G')] - p^2|E| + 3p|V| \geqslant 0.$$

Let $C_{G'}$ be the number of crossings in the drawing of G' inherited from G. Obviously, $\mathbb{E}[\operatorname{cr}(G')] \leqslant \mathbb{E}[C_{G'}]$. Since each crossing pair has a probability of p^4 of being in G', we have $\mathbb{E}[C_{G'}] = p^4 \operatorname{cr}(G)$, and thus

$$p^4 \operatorname{cr}(G) \geqslant \mathbb{E}[\operatorname{cr}(G')] \geqslant p^2|E| - 3p|V|.$$

By setting $p = 4|V|/|E|$, we have

$$\operatorname{cr}(G) \geqslant \frac{|E|}{p^2} - \frac{3|V|}{p^3} \geqslant \frac{|E|^3}{64|V|^2}.$$

□

3. Szemerédi-Trotter Theorem

Definition 3.1. Let P be a set of n points and L be a set of m lines in the plane. We call a pair (p, l) *incidence* if $p \in P$, $l \in L$, and $p \in l$. Define $I(P, L)$ as the number of incidences between P and L, and define $I(m, n)$ as the maximum number of incidences between any m lines and n points.

Definition 3.2. Let P be a set of n points. A line is *generated by* P if it contains at least 2 points from P.

Definition 3.3. For $k \geqslant 2$ and a set of points P, a line l is k-rich if it contains at least k points from P.

Theorem 3.4 (Szemerédi-Trotter Theorem)**.** *For all $m, n \geqslant 1$, we have $I(m, n) = O(m^{2/3}n^{2/3} + m + n)$.*

Proof. We will adopt the same strategy as the proof of Theorem 1.3, which constructs a graph and double counts the number of crossings in it.

Let P be the set of n points in $\mathbb{R}^2$ and L be the set of m lines in $\mathbb{R}^2$. Define graph $G = (V, E)$ where $V = P$ and E is the set of consecutive pairs of vertices along some line in L. We may assume each line in L contains at least one point from P. For $l \in L$, let $|l|$ denote the number of points in P which lies in l. Observe that

$$|E| = \sum_{l \in L} |l| - 1 = |I(P, L)| - m.$$

Hence, it suffices to show that $|E| = O(m^{2/3}n^{2/3} + n)$. We may assume $|E| \geqslant 4|V|$, otherwise we are done. Note that the construction of G gives a natural drawing with points P and lines P in the plane, so we may define C as the number of crossings in this drawing. By the crossing lemma, we have

$$\frac{|E|^3}{64n^2} \leqslant \operatorname{cr}(G) \leqslant C \leqslant \binom{m}{2} = O(m^2).$$

It now follows that

$$|E| = O(n^{2/3}m^{2/3}).$$

This completes the proof. $\square$

Corollary 3.5. *Let P be a set of n points. Then P generates $O(n^2/k^3 + n/k)$ k-rich lines.*

Proof. Let L_k be the set of k-rich lines generated by P. By the Szemerédi-Trotter theorem,

$$k|L_k| \leqslant I(P, L_k) = c(|L_k|^{2/3}n^{2/3} + |L_k| + n),$$

for some constant c. We may assume $k \geqslant 4c$, otherwise we are done as $|L_k| = O(n^2)$. If $n + |L_k| \geqslant |L_k|^{2/3}n^{2/3}$. Then

$$k|L_k| \leqslant 2c(|L_k| + n) = 2cm + 2c|L_k|.$$

Rearranged,

$$|L_k| \leqslant \frac{2cm}{k - 2c} \leqslant O(m/k).$$

Now suppose $n + |L_k| < |L_k|^{2/3}n^{2/3}$. Then

$$k|L_k| \leqslant 2c|L_k|^{2/3}n^{2/3},$$

and so

$$|L_k| = O(n^2/k^3).$$

$\square$

4. The Cutting Lemma

Lemma 4.1 (Cutting Lemma)**.** *Let L be a set of m lines in $\mathbb{R}^2$ and let $r \in (1, m)$. Then the plane can be subdivied into $t = O(r^2)$ generalized triangles (intersections of three half planes) $\Delta_1, \Delta_2, \ldots, \Delta_t$ such that the interior of each Δ_i is intersected by at most m/r lines of L.*

Lemma 4.2. *Let L be a set of m lines in $\mathbb{R}^2$ and let $r \in (1, m)$. Then the plane can be subdivied into $t = O(r^2 \log^2 n)$ generalized triangles $\Delta_1, \Delta_2, \ldots, \Delta_t$ such that the interior of each Δ_i is intersected by at most m/r lines of L.*

Proof. Put $s = 6r \ln m$. Select a random set of lines $S \subset L$ by making s independent random draws with replacement. Consider the line arrangement of S. Partition any cell that is not a generalized triangle further by adding diagonals that connect vertices. To this end, $\mathbb{R}^2$ is partitioned into t generalized triangles. Consider a box B that contains all bounded triangles Δ_i. Since each line crosses through B two times and each two consecutive lines around B determine an unbounded triangle, the number of unbounded triangles is at most $2s$. Now consider the bounded triangles. View each intersecting point of two lines in S as a vertex of a graph, and each bounded triangle as a face. Let V denote the set of vertices and F the set of faces. We know that $|V| \leqslant \binom{s}{2} = O(s^2)$. By Euler's formula, we have

$$3|F| \leqslant \sum_{f \in F} \deg f = 2|E| = 2(|V| + |F| - 2),$$

and thus

$$|F| \leqslant 2|V| - 4 = O(s^2).$$

Hence, we have $t = O(s^2)$.

We call a (generalized) triangle *horny* if its interior intersects at least m/r lines of L. For any horny triangle T, the probability that no line in S intersects the interior of T is at most $(1 - 1/r)^s$. Using the inequality $1 - x \leqslant e^{-x}$, we have $(1 - 1/r)^s \leqslant e^{-6 \ln m} = m^{-6}$.

Now call a triangle *interesting* if it can appear in a triangulation for some sample $S \subset L$. Notice that each vertex of an interesting triangle is an intersecting point of two lines in the arrangement of L, and thus there are at most $\binom{m}{2}^3 < m^6$ such triangles.

But then the expected number of horny Δ_i's is less than $m^{-6} \cdot m^6 = 1$. It now follows that there exists a set of $S \subseteq L$ such that each Δ_i is intersected by at most m/r lines. $\square$

5. An Aliter for the Szemerédi-Trotter Theorem

Theorem 5.1 (Kővári-Sós-Turán Theorem)**.** *For $s, t \geqslant 2$, let G be an $m \times n$ bipartite graph that does not contain a complete bipartite graph $K_{s,t}$ where the s vertices are from the part of size m. Then,*

$$|E(G)| = O(nm^{1-1/t} + m) \quad \text{and} \quad |E(G)| = O(mn^{1-1/s} + n).$$

Proof. Let M, N be the two parts of the bipartite graph G, with $|M| = m$ and $|N| = n$. Notice that no set of s vertices in M has more than $t-1$ common neighbors in N, so

$$\sum_{v \in M} \binom{d(v)}{t} \leqslant \binom{n}{t}(s-1) \leqslant \frac{sn^t}{t!}.$$

By Jensen's inequality, we have

$$\sum_{v \in M} \binom{d(v)}{t} \geqslant m \binom{\frac{1}{m}\sum_{v \in M} d(v)}{t} \geqslant \frac{m(2|E(G)|/m - t)^t}{t!}.$$

The result now follows from the two inequalities. ☐

Corollary 5.2. *$|I(m,n)| \leqslant O(n\sqrt{m} + m)$ and $|I(m,n)| \leqslant O(m\sqrt{n} + n)$.*

Proof. Let P be the set of n points and L be the set of m lines in $\mathbb{R}^2$. Let $G = (P, L)$ be the bipartite graph with parts P and L and (p, l) is an edge if and only if $p \in l$. Since no two points lie on the same line, G is $K_{2,2}$-free. The resulting bounds now follows from the Kővári-Sós-Turán theorem. ☐

We give an alternative proof of a case of the Szemerédi-Trotter theorem with n points and n lines, using the Cutting lemma and the Kővári-Sós-Turán theorem.

Aliter for Theorem 3.4. Let P be the set of n points and L be the set of n lines in $\mathbb{R}^2$. We need to show that there are at most $O(n^{4/3})$ incidences between P and L. We apply the cutting lemma with $r = n^{1/3}$, which divides the plane into $t = O(n^{2/3})$ generalized triangles $\Delta_1, \Delta_2, \ldots, \Delta_t$.

Let V be the points that lie on the vertex of some Δ_i. Since $|V| \leqslant 3t = O(n^{2/3})$, Corollary 5.2 gives us $|I(V, L)| = O(n^{2/3}\sqrt{n} + n^{2/3}) = O(n^{4/3})$.

Let $|L'|$ be the set of lines that borders some triangle Δ_i. Then $|L'| \leqslant 3t = O(n^{2/3})$, and Corollary 5.2 again gives us $|I(P_0, L')| = O(n^{2/3}\sqrt{n} + n^{2/3}) = O(n^{4/3})$.

It remains to count the incidences that occur at the interior of some triangle. Let P_i be the set of points in P that lies in the interior of Δ_i. Let L_i be the set of lines intersecting the interior of Δ_i. By the cutting lemma, $|L_i| \leqslant n/r = O(n^{2/3})$. Hence,

$$\sum_{i=1}^{t} I(P_i, L_i) \leqslant \sum_{i=1}^{t} I(P_i, n^{2/3}) = \sum_{i=1}^{t} O(|P_i| n^{1/3} + n^{2/3}) = O(n^{4/3}).$$

$\square$

6. Beck's Theorem

Theorem 6.1 (Beck's Theorem). *Given a set of n points P, there exists $\epsilon \in (0,1)$ such that either P contains ϵn points on a line or P generates at least ϵn^2 distinct lines.*

Proof. We may assume n is large, otherwise we the problem is trivial. Let P be a set of n points in $\mathbb{R}^2$. For $b > a \geqslant 2$, let $L_{[a,b]}$ be the set of lines generated by P with least a but less than b points on it. By Corollary 3.5, $L_{[a,b]} = O(n^2/a^3)$. We first make the following two observations:

For $k \leqslant \sqrt{n}$,

$$\#\{\{p_1,p_2\} : p_1,p_2 \in l,\, l \in L_{[k,\sqrt{n}]}\} \leqslant \sum_{i=0}^{\log_2 \frac{\sqrt{n}}{k}} |L_{[2^i k, 2^{i+1}k]}| \binom{2^{i+1}k}{2} = \sum_{i=0}^{\log_2 \frac{\sqrt{n}}{k}} O(n^2/2^i k) = O(n^2/k).$$

Hence, for $k < \sqrt{n}$, there are $O(n^2/k)$ pair of points in P that lies on a line with at least k but at most $\sqrt{n}$ points.

For $K > \sqrt{n}$,

$$\#\{\{p_1,p_2\} : p_1,p_2 \in l,\, l \in L_{[\sqrt{n},K]}\} \leqslant \sum_{i=0}^{\log_2 \frac{K}{\sqrt{n}}} |L_{[2^i\sqrt{n}, 2^{i+1}\sqrt{n}]}| \binom{2^{i+1}\sqrt{n}}{2} = \sum_{i=0}^{\log_2 \frac{K}{\sqrt{n}}} O(2^i n^{3/2}) = O(Kn).$$

Hence, there are $O(Kn)$ pairs of points from P that lies on a line with at least $\sqrt{n}$ but at most K points.

We now prove the theorem. Let $\epsilon \in (0,1)$ and set $\epsilon' = 4\sqrt{\epsilon}$. Assume that no $\epsilon' n$ points in P are colinear. Let $K = \epsilon' n$ and note that $K > \sqrt{n}$. Then the number of pairs of points in P that lies on a line with at least $\sqrt{n}$ but at most K points is $O(Kn) \leqslant c\epsilon' n^2 \leqslant n^2/10$, for some constant c and suffciently small ϵ. Now let $k = 1/\epsilon'$ and note that $k \leqslant \sqrt{n}$. Then the number of pairs of points in P that lies on a line with at least k but at most $\sqrt{n}$ points is $O(n^2/k) \leqslant c'\epsilon' n^2 \leqslant n^2/10$, for some constant c' and ϵ suffciently small. But then the number of pairs of points in P that lies in a k-rich line is at most $n^2/10 + n^2/10 = n^2/5$. Thus there are at least $\binom{n}{2} - n^2/5 \geqslant n^2/4$ pairs in P that lies on a line with at most k points, and so there are at least $\frac{n^2/4}{\binom{k}{2}} \geqslant \epsilon m^2$ distinct lines generated by P. $\square$

7. Simplicial Partition

Theorem 7.1 (Simplicial Partition). *Let P be n points in $\mathbb{R}^2$. There exists partition $P = P_1 \sqcup P_2 \sqcup \cdots \sqcup P_{2r}$ and generalized triangles $\Delta_1, \Delta_2, \ldots, \Delta_{2r}$, with $P_i \subset \Delta_i$, $|P_i| = n/2r$ for $i < 2r$ and $|P_{2r}| \leqslant n/2r$, such that for any line l generated by P, l will cross the interior of $O(\sqrt{r})$ number of Δ_i's.*

Proof. Pick $r > (\log n)^2$. Let L be the set of lines generated by P. Let $\Delta'_1 \cup \Delta'_2 \cup \cdots \cup \Delta'_r$ be the generalized trianges yielded by the cutting lemma on L with parameter $t = r$. By the pigeonhole principle, there exists Δ_i that contains $\geqslant n/r$ points from P. Let P_1 be some $n/2r$ points selected from Δ_i excluding the corners, and let $\Delta_1 = \Delta'_i$. Set $P' = P \backslash P_1$. For each line that crosses the interior of Δ_1, we double it by creating a copy of the line close to it, and let L' be all the lines after this process. Note that by the cutting lemma, the number of lines that cross the interior of Δ_1 is $c|L|/\sqrt{r}$ for some $c > 0$, and so

$$|L'| \leqslant |L| + \frac{c|L|}{\sqrt{r}} = \left(1 + \frac{c}{\sqrt{r}}\right)|L|.$$

Now apply the cutting lemma again to L' with parameter $t = r(1 - 1/2r)$, and we get a generalized triangle Δ''_i wtih $\geqslant |P'|/t = \frac{|P'|}{r(1-1/2r)} = n/r$ points from P' that lies in Δ''_i. Set P_2 be some $n/2r$ points of P' in Δ''_i excluding the corners, and let $\Delta_2 = \Delta''_i$. Set $P'' = P' \backslash P_2$ and note taht $|P''| = (1 - 1/r)n$. For any line that crosses the interior of Δ_2, we double again it, and let L'' be all the lines after this process. By the same argument,

$$|L''| \leqslant |L'| + \frac{c|L'|}{\sqrt{r(1-1/2r)}} = \left(1 + \frac{c}{\sqrt{r(1-1/2r)}}\right)|L'| \leqslant \left(1 + \frac{c}{\sqrt{r}}\right)\left(1 + \frac{c}{\sqrt{r(1-1/2r)}}\right)|L|.$$

Repeat the above process, and after k iterations we get point sets $P_1, P_2, \ldots, P_k$ and generalized triangles $\Delta_1, \Delta_2, \ldots, \Delta_k$. Set $P^{(k)} = P \backslash (P_1 \cup P_2 \cup \cdots \cup P_k)$. Again, let $L^{(k)}$ be the set of lines after doubling the lines that cross the interior of some $\Delta_i^{(k)}$'s. Then

$$|P^{(k)}| = |P| - \frac{kn}{2r} = \left(1 - \frac{k}{2r}\right)n.$$

$$|L^{(k)}| \leqslant \left(1 + \frac{c}{\sqrt{r}}\right)\left(1 + \frac{c}{\sqrt{r - 1/2}}\right)\cdots\left(1 + \frac{c}{\sqrt{r - (k-1)/2}}\right)|L| \leqslant |L| \exp\left(c \sum_{j=0}^{2r-1} \frac{1}{\sqrt{r - j/2}}\right).$$

Iterate this process until there are $< n/2r$ points left, and let P_{2r} be the remaining points and Δ_{2r} be some generalized triangle that contains P_{2r}.

It remains to show that any line $l \in L$ crosses the interior of $O(\sqrt{r})$ Δ_i's. Let x be the number of Δ_i's that some line l crosses. Notice that by the end of the process above,

$$2^x \leqslant \#\text{copies of } l \leqslant |L^{(2r)}| \leqslant |L| \exp\left(c \sum_{j=0}^{2r-1} \frac{1}{\sqrt{r - j/2}}\right) \leqslant n^2 e^{O(\sqrt{r})} = 2^{O(\sqrt{r})}.$$

This proves the theorem. $\square$

8. Triangle Removal Lemma

Definition 8.1. The *density* of edges between two vertex sets A and B is

$$d(A,B) := \frac{|E(A,B)|}{|A||B|}.$$

Definition 8.2. Let $\epsilon \in (0,1)$. The pair of vertex sets (A,B) is *ϵ-regular* if for all $A' \leqslant A$ and $B' \leqslant B$ such that $|A'| \geqslant \epsilon|A|$ and $|B'| \geqslant \epsilon|B|$, we have

$$|d(A',B') - d(A,B)| \leqslant \epsilon.$$

Definition 8.3. Given a graph $G = (V,E)$, a partition $V = V_1 \sqcup V_2 \sqcup \cdots \sqcup V_k$ is a *ϵ-regular* if

$$\sum_{(i,j)\in[k]^2,\, (V_i,V_j) \text{ not } \epsilon\text{-regular}} |V_i||V_j| \leqslant \epsilon|V|^2.$$

Note that we are only interested in dense graphs. This is because if $|E(A,B)| = o(|A||B|)$, the density of 0 and so the pair (A,B) is trivially ϵ-regular.

Theorem 8.4 (Szemerédi's Regularity Lemma)**.** *For all $\epsilon > 0$, there exists $k = k(\epsilon)$ such that for any graph $G = (V,E)$, there exists an ϵ-regular partition $V = V_1 \sqcup \cdots \sqcup V_k$.*

Lemma 8.5 (Counting Lemma)**.** *Let $G = (V,E)$ be a graph, such that V is partitioned into $X \sqcup Y \sqcup Z$ where each pair of them are ϵ-regular, and $d(X,Y) = \alpha, d(X,Z) = \beta, d(Y,Z) = \gamma$, with $\alpha, \beta, \gamma > 2\epsilon$. Then*

$$\#\{K_3 \subseteq G\} \geqslant (1-2\epsilon)(\alpha-\epsilon)(\beta-\epsilon)(\gamma-\epsilon)|X||Y||Z|.$$

Proof. For $x \in X$, denote $d_Y(x) = d(x) \cap Y$ and $d_Z(x) = d(x) \cap Z$. We claim that $d_Y(x) < (\alpha-\epsilon)|Y|$ for at most $\epsilon|X|$ vertices in X. Suppose otherwise. Let $X' \subseteq X$ be the set of vertices with $d_Y(x) < (\alpha-\epsilon)|Y|$. Since (X,Y) is ϵ-regular, $|d(X',Y) - d(X,Y)| \leqslant \epsilon$, and so

$$\alpha - \epsilon < d(X',Y) = \frac{|E(X',Y)|}{|X'||Y|} \leqslant \frac{(\alpha-\epsilon)|X'||Y|}{|X'||Y|} = \alpha - \epsilon.$$

This contradiction proves the claim. By the same argument, we also know that $d_Z(x) < (\gamma-\epsilon)|Y|$ for at most $\epsilon|X|$ vertices in X.

Let $x \in X$ with $d_Y(x) \geqslant (\alpha-\epsilon)|Y|$ and $d_Z(x) \geqslant (\gamma-\epsilon)|Z|$. Let $|Y'| = N(x) \cap Y$ and $|Z'| = N(x) \cap Z$. Then

$$\#\{K_3 \subseteq G, x \in K_3\} = |E(Y',Z')|.$$

Since $|d(Y',Z') - d(Y,Z)| < \epsilon$, we have

$$\beta - \epsilon < d(Y',Z') = \frac{|E(Y',Z')|}{|Y'||Z'|}.$$

Rearranging gives us

$$\#\{K_3 \subseteq G, x \in K_3\} =\geqslant (\beta - \epsilon)|Y'||Z'| \geqslant (\beta - \epsilon)(\alpha - \epsilon)(\gamma - \epsilon)|Y||Z|.$$

Since there are at least $(1 - 2\epsilon)$ such x's in X,

$$\#\{K_3 \subseteq G\} \geqslant (1 - 2\epsilon)(\alpha - \epsilon)(\beta - \epsilon)(\gamma - \epsilon)|X||Y||Z|.$$

$\square$

9. Roth's Theorem

Theorem 9.1 (Roth's Theorem)**.** *For all $\delta \in (0,1)$, there exists n_0 such that for all $n > n_0$, any subset of $[n]$ of size $\geqslant \delta n$ contains a 3-term arithmetic progression.*

10. Solymosi's Theorem

Theorem 10.1. *Let P be a set of n points and L be a set of n lines in $\mathbb{R}^2$, and let r be a parameter. If the arrangement of P and L does not contain a triangle, then $|I(P,L)| = O(n^{4/3}/\log^* n) = o(n^{4/3})$, where $\log^*$ is the iterated logarithm.*