Probability and Graphs

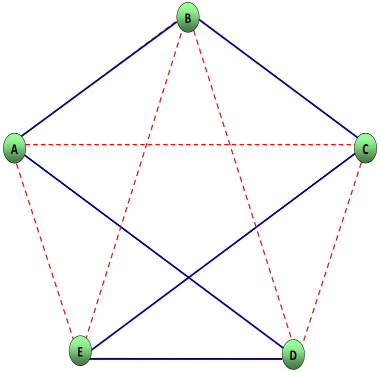
Arnab Basu

Quantitative Methods and Information Systems
Indian Institute of Management Bangalore
Bangalore 560076, India.

Ramsey Number

- *Ramsey Number* $R(k, l)$ is the smallest integer n such that in any two-colouring of the edges of a complete graph on n vertices K_n by red and blue, either there is a red K_k or there is a blue K_l.

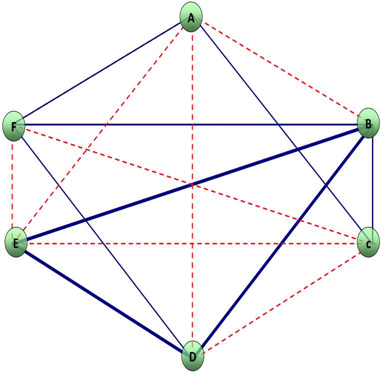
Example (K_5)



K_5 need not have a monochromatic triangle.

- Ramsey (1929) showed that $R(k, l)$ is finite for any two integers k and l.

Example ($R(3, 3) = 6$)



K_6 will have a monochromatic triangle.

- We propose to obtain a lower bound on the diagonal Ramsey Numbers $R(k, k)$.

- We now proceed to prove, step by step, that

$$R(k,k) > \lfloor 2^{\frac{k}{2}} \rfloor, \ \forall k \geq 3.$$

- Let S denote a fixed set of k vertices. Let A_S denote the event that the induced subgraph of K_n on S be monochromatic; then

$$P\left[A_S\right] = 2^{1-\binom{k}{2}}.$$

- Note that there are $\binom{n}{k}$ choices for such an S.

- So the total probability $q(n,k)$ of the event that at least one induced subgraph of k vertices on K_n is monochromatic is

$$q(n,k) \equiv \binom{n}{k} 2^{1-\binom{k}{2}}.$$

- Suppose, we indeed choose n and $k \geq 3$ such that $q(n, k) < 1$.
- Then, with positive probability, none of the A_S's occur i.e., there is a two-colouring of K_n without a monochromatic K_k i.e.,

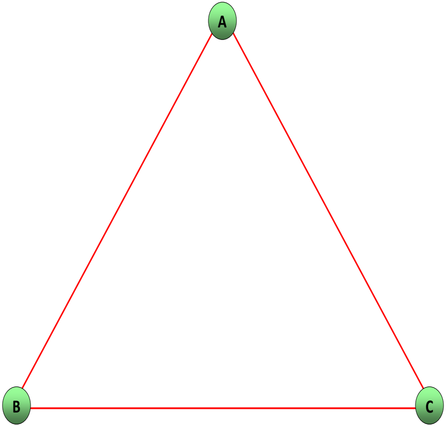
$$R(k, k) > n.$$

- Let the choice of n and $k \geq 3$ be $n = \lfloor 2^{\frac{k}{2}} \rfloor$.
- Then, $q(n, k) < \frac{2^{1+\frac{k}{2}}}{k!} \left(\frac{n}{2^{\frac{k}{2}}} \right)^k < 1$.
- So, $R(k, k) > \lfloor 2^{\frac{k}{2}} \rfloor, \ \forall k \geq 3$.

Crossing Number and Szemerédi-Trotter Theorem

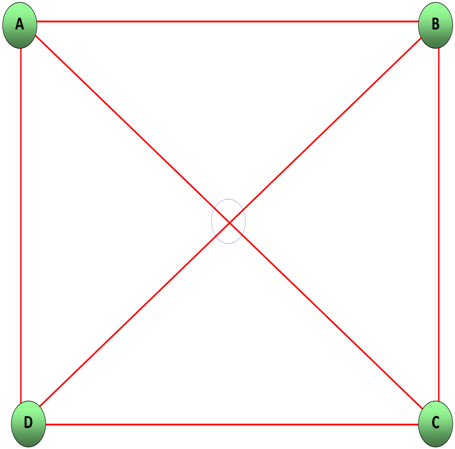
- An embedding of a graph $G = (V, E)$ in the plane is a planar representation of it, where each vertex is represented by a point in the plane, and each edge (u, v) is represented by a curve connecting the points corresponding to the vertices u and v.
- The crossing number of such an embedding is the number of pairs of intersecting curves that correspond to pairs to edges with no common endpoints.
- The crossing number $cr(G)$ of G is the minimum possible crossing number in an embedding of it in the plane.

Example (K_3)



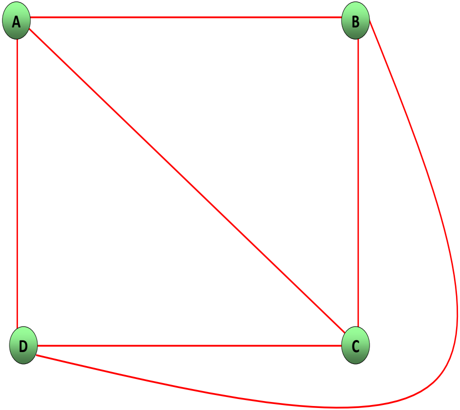
In every planar embedding the graph K_3 has crossing number 0. Hence it is a planar graph.

Example (K_4)



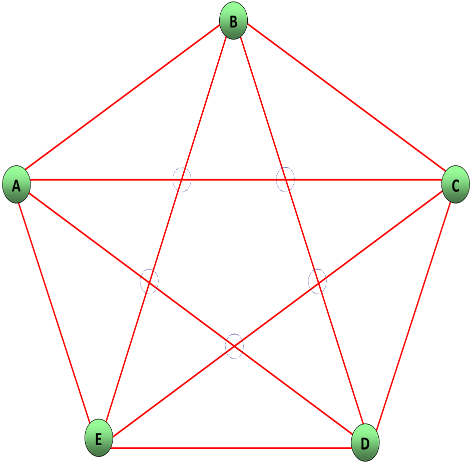
The graph K_4 has crossing number 1 !!!

Example (K_4)



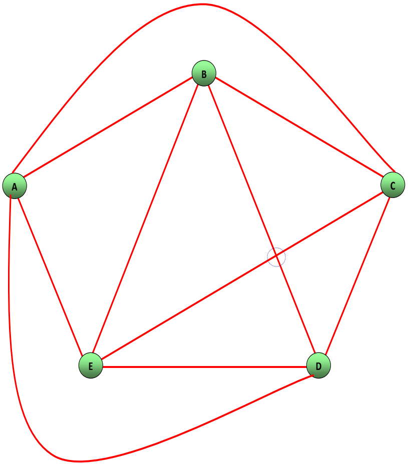
The graph K_4 actually has crossing number 0 !!! It is a planar graph.

Example (K_5)



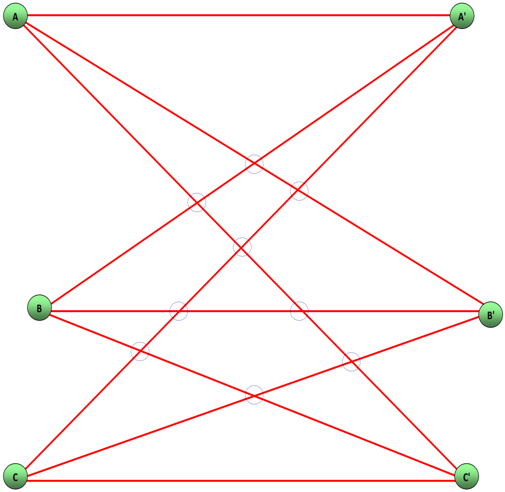
The graph K_5 has crossing number 5 !!!

Example (K_5 has crossing number 1 !!!)



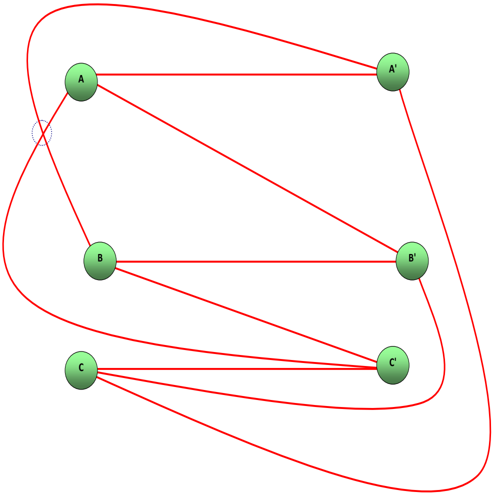
In every planar embedding the graph K_5 has at least a pair of edges crossing. Hence, it is a non-planar graph.

Example ($K_{3,3}$)



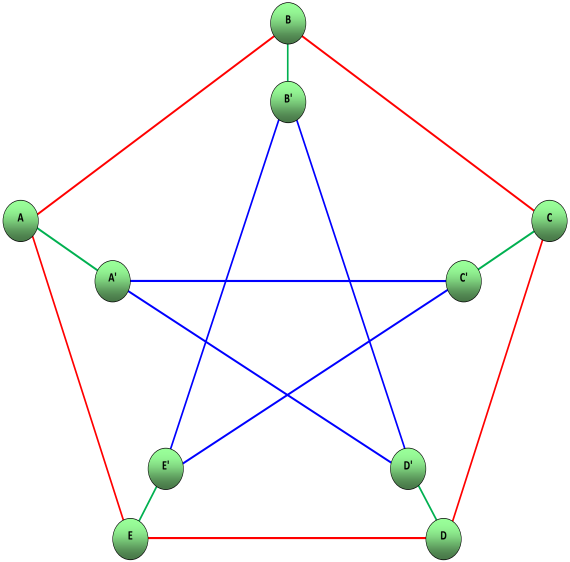
The crossing number of $K_{3,3}$ is 9 !!!

Example ($K_{3,3}$ has crossing number 1)



Hence, it is a non-planar graph.

Example (Petersen Graph)



Famous example of a non-planar graph

- Theorem (Kuratowski, 1930): A graph is planar iff it has no subgraph homeomorphic to K_5 or $K_{3,3}$.

- The following Crossing Number Theorem was proved by Ajtai, Chvátal, Newborn and Szemerédi (1982) and independently, by Leighton:

  The crossing number of any simple (i.e., with no multi-edges or no self-loops) graph $G = (V, E)$ with $|E| \geq 4|V|$ is at least $\frac{|E|^3}{64|V|^2}$.

- Let us describe a short probabilistic proof of this theorem.

- Euler's Formula: For any spherical polyhedron, with V vertices, E edges and F faces,
$$V - E + F = 2.$$

- Any maximal planar (i.e., one to which no edge can be added without losing planarity) graph will have triangular faces implying

$$3F = 2E.$$

- Hence for any simple planar graph with $V = n \geq 3$ vertices, we have

$$E = V + F - 2 \leq V + \frac{2}{3}E - 2 \Rightarrow E \leq 3n - 6,$$

  implying that it has at most $3n$ edges.

- Therefore, the crossing number of any simple graph with n vertices and m edges is at least $m - 3n$.

- Let $G = (V, E)$ be a graph with $|E| \geq 4|V|$ embedded in the plane with $t = cr(G)$ crossings.

- Let H be the random induced subgraph of G obtained by picking each vertex of G, randomly and independently, to be a vertex of H with probability p (to be chosen later).

- Then, the expected number of vertices in H is $p|V|$, the expected number of edges is $p^2|E|$, and the expected number of crossings (in its given embedding) is $p^4 t$.

- Therefore, we have

$$p^4 t \geq p^2|E| - 3p|V|,$$

  implying

$$t \geq \frac{|E|}{p^2} - 3\frac{|V|}{p^3}.$$

- Substituting $p = \frac{4|V|}{|E|}$ (≤ 1), we get the result.

- Now we state the famous Szemerédi-Trotter Theorem in Combinatorial Geometry:

  Let P be a set of n distinct points in the plane, and let L be a set of m distinct lines. Then the number of incidences between the members of P and those of L (i.e., the number of pairs (p, l) with $p \in P,\ l \in L,\ p \in l$) is at most $c(m^{\frac{2}{3}} n^{\frac{2}{3}} + m + n)$, for some absolute constant $c > 0$.

- We shall now give a step-by-step proof using probabilistic arguments. This proof is due to Székely (1997).

- We may and shall assume that every line in L is incident with one of the points of P.

- Denote the number of such incidences by I.

- Form a graph $G = (V, E)$ with $V = P$, where for $p, q \in P$, $(p, q) \in E$ iff they are consecutive points of P on some line in L.

- Clearly, $|V| = n$, and $|E| = \sum_{j=1}^{m}(k_j - 1) = \sum_{j=1}^{m} k_j - m = I - m$, where k_j is the number of points of P on line $j \in L$.

- Note that G is already embedded in the plane where the edges are represented by segments of the corresponding lines in L.

- In this embedding, every crossing is an intersection point of two members of L, implying

$$cr(G) \leq \binom{m}{2} \leq \frac{1}{2}m^2.$$

- By the Crossing Number Theorem, either $I - m = |E| < 4|V| = 4n$, that is,

$$I \leq m + 4n$$

or

$$\frac{m^2}{2} \geq cr(G) \geq \frac{(I-m)^3}{64n^2},$$

implying

$$I \leq (32)^{\frac{1}{3}} m^{\frac{2}{3}} n^{\frac{2}{3}} + m.$$

- In both cases,

$$I \leq 4\left(m^{\frac{2}{3}} n^{\frac{2}{3}} + m + n\right).$$

Discrepancy Methods in Graphs

- Consider a set system (hypergraph) $G(V, \mathcal{S})$, with n vertices ($|V| = n$) and a set $\mathcal{S}$ of m k-hyperedges (subsets of V of size k).

- For all $e \in S$, define

$$\chi(e) \stackrel{\text{def}}{=} \sum_{v \in V : v \in e} \chi(v),$$

where $\chi(v) \in \{1 = \text{blue}, -1 = \text{red}\}$ is the colour assigned to vertex v.

- The discrepancy $\mathcal{D}(\mathcal{S})$ of the system is defined as

$$\mathcal{D}(\mathcal{S}) \stackrel{\text{def}}{=} \min_{\chi : V \to \{1, -1\}} \max_{e \in \mathcal{S}} |\chi(e)|.$$

- Before we proceed further, let us state the following famous theorem due to Chernoff (1952):

  Let $X_i,\ i = 1, \dots, n$ be mutually independent random variables with

  $$P[X_i = +1] = P[X_i = -1] = \frac{1}{2},$$

  and let $S_n = \sum_{i=1}^{n} X_n$. Let $a > 0$. Then

  $$P[S_n > a] < e^{-\frac{a^2}{2n}}.$$

- Using symmetry arguments, we immediately get the following corollary:

  $P[|S_n| > a] < 2e^{-\frac{a^2}{2n}}.$

- The following theorem gives an upper bound on the discrepancy $\mathcal{D}(\mathcal{S})$ of such a set system S:

$\mathcal{D}(\mathcal{S}) \leq \sqrt{2n\ln(2m)}$.

- Let us prove this step by step.

- For $A \subset V$, and for random $\chi : V \to \{1, -1\}$, let X_A be the indicator of the event $\{|\chi(A)| > \alpha\}$, where $\alpha \stackrel{\text{def}}{=} \sqrt{2n\ln(2m)}$.

- If $|A| = k$, then by our choice of α, we have, by the above corollary of Chernoff's Theorem,

$$E[X_A] = P[|\chi(A)| > \alpha] < 2e^{-\frac{\alpha^2}{2k}} \leq 2e^{-\frac{\alpha^2}{2n}} = \frac{1}{m}.$$

- Let X be the number of A with $\{|\chi(A)| > \alpha\}$, so that $X = \sum_{A\in\mathcal{S}} X_A$.

- Hence, we have $E[X] = \sum_{A\in\mathcal{S}} E[X_A] < |\mathcal{S}|\left(\frac{1}{m}\right) = 1$.

- Thus, for some χ, we must have $X = 0$, implying

$$\chi(A) \leq \alpha, \ \forall A \in \mathcal{S},$$

  implying

$$\max_{e\in\mathcal{S}} |\chi(e)| \leq \alpha.$$

- Hence we have

$$\mathcal{D}(\mathcal{S}) = \min_{\chi: V\to\{1,-1\}} \max_{e\in\mathcal{S}} |\chi(e)| \leq \alpha = \sqrt{2n\ln(2m)}.$$

Bibliography

Ram - Ramsey, F. P. (1929); *On a Problem of Formal Logic*, Proc. Lond. Math. Soc., 30(2): 264-286.

Kur - Kuratowski, K. (1930); *Sur le problème des courbes gauches en topologie*, Fund. Math., 15: 271-183.

ST - Szemerédi, E. and Trotter (Jr.), W. (1983); *A Combinatorial Distinction between Euclidean and Projective Planes*, European J. Comb., 4: 385-394.

Sze - Székely, L. (1997); *Crossing Numbers and Hard Erdös Problems in Discrete Geometry*, Combin. Prob. Comp., 6: 353-358.

M - Matoušek, J. (1999); *Geometric Discrepancy: An Illustrated Guide*, Springer, Algorithms and Combinatorics, 18.

C - Chernoff, H. (1952); *A Measure of the Aymptotic Efficiency for Tests of a Hypothesis Based on the Sum of Observations*, Ann. Math. Stat., 23: 493-509.

THANK YOU