Extremal Graph Theory and the Shifting Method

Ivy Yao, Julia Zhou (mentor Tomasz Slusarczyk)

April 2024

1 Introduction

In graph theory, a *graph* is a pair of sets $G = (V, E)$ with the vertex set V and edge set E connecting the vertices. Two vertices x, y of G are *neighbors*, if xy is an edge of G. Collectively, all neighbors of a vertex v is called the *neighborhood* of v, $N(v)$. The *degree* of a vertex v, which is the number of edges in E that is incident at v, is denoted by $\deg(v)$.

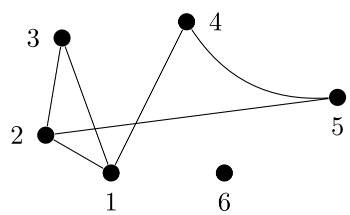

Figure 1. The graph on $V = 1, \ldots, 7$ with edge set $E =$ [1,2][1,3][1,4][2,3][2,5][4,5]

A specific study under Graph theory is *extremal graph theory.* In extremal graph theory, one focuses on finding the maximum and minimum conditions that either ensures or inhibits a pattern. The starting point of extremal graph theory was this question: what is the maximum number of edges in an n-vertex triangle-free graph? This question was answered by Willem Mantel in the early 1900's via Mantel's Theorem.

In this paper, we will introduce two proofs of Mantel's theorem using inequality of means. We will also present proofs of the inequality of means with shifting. Lastly, we will use the shifting method to prove some example problems.

2 Mantel's Theorem

Theorem 1 *Mantel's Theorem states that every n-vertex, triangle free graph contains at most* $\lfloor \frac{n^2}{4} \rfloor$ *edges.*

Proof 1. Let G be a triangle-free graph with n vertices and m edges. Since the graph is triangle-free, the endpoint vertices of any edge $xy \in E(G)$ have no shared neighbors. This means that $\deg(x) + \deg(y) \leq n$. Taking the sum of all edges in G gives the following inequality:

$$\sum_{xy \in E(G)} \deg(x) + \deg(y) \leq mn \tag{1}$$

Observe that the term $\deg(x)$ will appear in every edge incident to x, so $\deg(x)$ appears $\deg(x)$ times. Thus, the previous inequality can be simplified by considering each vertex instead of each edge:

$$\sum_{x \in V(G)} \deg(x)^2 \leq mn \tag{2}$$

The inequality between quadratic and arithmetic means is as follows:

Lemma 1 *For any reals $x_1, \dots, x_n$ the following inequality holds*

$$\sqrt{\frac{x_1^2 + x_2^2 + \ldots + x_n^2}{n}} \geq \frac{x_1 + x_2 + \ldots + x_n}{n}.$$

A proof of this inequality is covered in section 3 using shifting. Algebraic manipulation on this inequality converts the quadratic (left hand) side to match the form of eq. (2):

$$x_1^2 + x_2^2 + \ldots + x_n^2 \geq (\frac{x_1 + x_2 + \ldots + x_n}{n})^2 \tag{3}$$

Lemma 2 (The Handshaking Lemma) *The sum of all degrees in a graph is equal to twice the number of edges.*

Combining the above 2 lemmas, eq. (2) generates the following inequality:

$$\sum_{x \in V(G)} \deg(x)^2 \geq \frac{1}{n}(\sum_{x \in V(G)} \deg(x))^2 = \frac{(2m)^2}{n} \tag{4}$$

Recall that the right-hand side of this inequality originated from inequality 1, so $\frac{(2m)^2}{n} \leq mn$ and $m \leq \frac{n^2}{4}$. Lastly, since m, the number of edges in a graph, must be an integer, the floor function can be added to become $m \leq \lfloor \frac{n^2}{4} \rfloor$, proving Mantel's theorem.

Proof 2. Let G be a triangle-free graph. Let v be a vertex with maximum degree. Because G is a triangle-free graph, the neighborhood of v, $N(v)$, is a set of vertices without edges between. Let $V = A \cup B$ where $N(v) = A, B = V \setminus A$. Because there are no edges in A, every edge must have at least one endpoint in B, thus:

$$|E| \leq \sum_{x \in B} \deg(x) \tag{5}$$

Because v is the vertex in G with maximum degree, among all $x \in V$, $\deg(x) \leq \deg(v) = |A|$. Thus:

$$|E| \leq \sum_{x \in B} \deg(x) \leq |B| \max_{x \in B} \deg(x) \leq |B| \deg(v) = |B||A| \tag{6}$$

The AM-GM inequality which states the arithmetic mean is always larger than the geometric mean is as follows:

Lemma 3 *For any reals $x_1, \ldots, x_n$ the following inequality holds*

$$(x_1 x_2 x_3 \cdots x_n)^{\frac{1}{n}} \leq \frac{x_1 + x_2 + x_3 \cdots + x_n}{n}$$

A proof of AM-GM inequality using shifting is covered in Section 3. Algebraic manipulation of this inequality gives us the below:

$$x_1 x_2 x_3 \cdots x_n \leq (\frac{x_1 + x_2 + x_3 \cdots + x_n}{n})^n \tag{7}$$

eq. (6) and eq. (7) then gives the following:

$$|E| \leq |A| * |B| \leq (\frac{|A| + |B|}{2})^2 = \frac{n^2}{4} \tag{8}$$

Since the number of edges in a graph must be an integer, the floor function can be added to become $|E| \leq \lfloor \frac{n^2}{4} \rfloor$, proving Mantel's theorem.

3 Proving inequalities with shifting

A technique called shifting can be used to prove all parts of the Cauchy-Schwarz inequality.

The quadratic mean and arithmetic mean (QM-AM) inequality is as follows:

$$\sqrt{\frac{x_1^2 + x_2^2 + \cdots + x_n^2}{n}} \geq \frac{x_1 + x_2 + \cdots + x_n}{n} \tag{9}$$

For this proof, we will shift values of x to decrease the left-hand side while keeping the right-hand side constant. Let A equal the current arithmetic mean of the numbers. Without loss of generality, we will choose two terms, x_1 and x_2 such that $x_1 < A < x_2$. Next, let s be the minimum of $A - x_1$ and $x_2 - A$ and shift the value of x_1 to $x_1 + s$ and x_2 to $x_2 - s$. This moves x_1 and x_2 closer to A (and makes one value equal to A).

We can show that this shift will decrease the QM of the set by substituting the new values of x_1 and x_2 into the original formula: ‘

$$QM' = \sqrt{\frac{x_1^2 + x_2^2 + \cdots + x_n^2}{n}} \tag{10}$$

We want to prove that the QM' is indeed $\leq QM$. The inequality $QM' \leq QM$ simplifies to become $0 \leq (x_1)s - (x_2)s + s^2$ and $x_2 - x_1 \leq s$, which must be true because of the definition we used for s. Thus, shifting values does decrease the QM without changing the AM. We continue shifting until $x_1 = x_2 = \ldots = x_n = A$. For this equality case, $QM' = \sqrt{\frac{n(A^2)}{n}} = A$ and $AM = \frac{n(A^2)}{n} = A$, so $QM' = AM$ after shifting. However, QM' is smaller than the original QM, proving that $QM \leq AM$.

Similarly, shifting can also be used to prove the AM-GM inequality, which is as follows:

$$\frac{x_1 + x_2 + x_3 \cdots + x_n}{n} \geq (x_1 x_2 x_3 \cdots x_n)^{\frac{1}{n}} \tag{11}$$

For this proof, we will shift values of x to increase the right-hand side (RHS) while keeping the left-hand side (LHS) constant. Let A equal the current arithmetic mean of the numbers. We will choose two terms, x_1 and x_2 such that $x_1 < A < x_2$. Next, let s be the minimum of $A - x_1$ and $x_2 - A$ and shift the value of x_1 to $x_1' = x_1 + s$ and x_2 to $x_2' = x_2 - s$. This moves x_1 and x_2 closer. After shifting, LHS and RHS are as follows:

$$LHS = \frac{(x_1 + s) + (x_2 - s) + \cdots + x_n}{n} = \frac{x_1 + x_2 + \cdots + x_n}{n} \tag{12}$$

$$RHS = \sqrt[n]{(x_1 + s)(x_2 - s) \cdots x_n} \geq \sqrt[n]{x_1 x_2 \cdots x_n} \tag{13}$$

The LHS remains constant after shifting while the RHS, which is the geometric mean, increased after shifting. We can continue shifting until equality case is reached, where all $x = A$. In this case:

$$LHS = \frac{nA}{n} = A \tag{14}$$

$$RHS = \sqrt[n]{A^n} = A \tag{15}$$

$$LHS = RHS \tag{16}$$

The geometric mean increased after shifting while arithmetic did not, thus arithmetic mean is greater than or equal to geometric mean.

4 More questions using shifting

1. Prove that a K_5 free graph on n vertices has $\leq \frac{n^3}{16}$ triangles. This is equivalent to proving Zykov's theorem, as shown below, for $k = 3$, $l = 5$.

Theorem 2 (Zykov 1949) *Let $l > k \geq 2$ be integers. Any n-vertex graph without any K_l has at most*

$$\frac{n^2}{(l-1)^2} \binom{l-1}{k}$$

copies of K_k.

Let $t(u)$ be the number of triangles containing vertex u, and $t(uv)$ be the number of triangles containing edge uv. Let also $t(G)$ be the number of triangles in a graph G. Let $G = (V, E)$ be such that G has no K_5's and the most triangles. Any edge not contained in any triangles can be removed without changing the number of triangles, so $t(uv) \geq 1$ for all $uv \in E$.

Lemma 4 *For any $uv \notin E$, $t(u) = t(v)$.*

We will prove this lemma by contradiction. Consider vertices u and v such that $uv \notin E$ and $t(u) < t(v)$. Then remove u and replace it with a new vertex v', such that $N(v') = N(v)$, to create G'.

$$t(G') = t(G) - t(u) + t(v) > t(G) \tag{17}$$

If G' contains a K_5, then $N(v)$ contains a K_4, so G contained a K_5. Thus G' is K_5-free and has more triangles than G, which is a contradiction.

Lemma 5 *If $uv, vw \notin E$ then $uw \notin E$.*

We will also prove this lemma by contradiction. Consider vertices u, v, w such that $uv, vw \notin E$. Replace u with v' and w with v'', such that $N(v') = N(v'') = N(v)$, to create G':

$$\begin{aligned} t(G') &= t(G) - t(u) - t(w) + t(uw) + 2t(v) && (18)\\ &= t(G) + t(uw) && (19)\\ &> t(G). && (20) \end{aligned}$$

Similarly as before, G' must be K_5-free, and it has more triangles than G, which is a contradiction.

Fact 1 *If an graph $G = (V, E)$ is such that for any $uv, vw \notin E$ we have $uw \notin E$, then G must be a complete multipartite graph.*

By the previous lemma, G is a complete multipartite graph. G is K_5 free, so it is a 4-partite graph. Let a, b, c, d be the numbers of vertices in 4 parts. Now, $t(G) = abc + abd + acd + bcd$. Since we are proving that any K_5-free graph on n vertices has $\leq \frac{n^3}{16}$ triangles, we can write the following inequality for any $a, b, c, d \geq 0$:

$$abc + abd + acd + bcd \leq \frac{(a + b + c + d)^3}{16}. \tag{21}$$

If $a = b = c = d$, the equality holds. We can use the shifting method to achieve this condition. Let $M = \frac{a+b+c+d}{4}$ be the mean of a, b, c, d. First, we pick a and b such that $a < M < b$. Next, we shift 2. (a, b) so that $(a' = a + x,\ b' = b - x)$ where $x = \min((M - a), (b - M))$.

After shifting, the left-hand side of eq. (21) increases while the right-hand side stayed the same:

$$
\begin{align}
LHS &= (a+x)(b-x)c + (a+x)(b-x)d + (a+x)cd + (b-x)cd \tag{22}\\
&= abc + abd + acd + bcd + (b-a-x)(c+d)x \tag{23}\\
&> abc + abd + acd + bcd \tag{24}
\end{align}
$$

$$
RHS = \frac{((a+x)+(b-x)+c+d)^3}{16} = \frac{(a+b+c+d)^3}{16}. \tag{25}
$$

We keep shifting until we reach the equality case $a = b = c = d$ when left-hand side = right-hand side. Because the left-hand side increased and the right-hand side did not, so, at the beginning, left-hand side $\leq$ right-hand side. Thus, it is proved that a K_5 free graph on n vertices has $\leq \frac{n^3}{16}$ triangles.

2. Assume G is an oriented graph such that every edge has a direction and edges xy and yx cannot simultaneously exist. Every edge has weight 1 or 2, and for any edges xy and yz, $w(xy) + w(yz) \leq 2$ where w is the weight of the edge. Show that the total weight of all edges is at most $\frac{n^2}{2}$.

We start by discarding all isolated vertices. We will then partition the vertices into three sets. Set X will contain all vertices with an outgoing edge of weight 2, set Z will contain all vertices with an incoming edge of weight 2, and set Y will contain all other vertices. Y is disjoint from X and Z by definition and X and Z are disjoint because there a vertex with an incoming edge of weight 2 cannot have any outgoing edge, and vice versa. All edges are of the form set X to set Z with weight 2, X to Y with weight 1, Y to Z with weight 1, or Y to a different vertex in Y with weight 1.

For any two sets A and B, we will notate the number of edges from A to B as ab. We will also notate the number of vertices in any set A as a. We can find the total weight of all edges and set up an inequality:

$$
2xz + xy + yz + \frac{y(y-1)}{2} \leq \frac{(x+y+z)^2}{2} \tag{26}
$$

When $x = y = z$, the equality case holds, in which both the left-hand side (LHS) and the right-hand side (RHS) is equal to $\frac{9x^2}{2}$. To achieve that equality case, we can use the shifting method. Let $M = \frac{x+y+z}{3}$ be the mean of x, y, z. First, we pick x and y such that $x < M < y$. Next, we shift x and y such that $x' = x + a$ and $y' = y - a$ where $a = \min((M - a), (b - M))$. After shifting, the LHS increased while the RHS remained constant:

$$LHS = 2(x+a)z + (x+a)(y-a) + (y-a)z + \frac{(y-a)(y-a-1)}{2} \tag{27}$$

$$= 2xz + xy + yz + a(y-x+z) - a^2 + \frac{(y-a)^2-(y-a)}{2} \tag{28}$$

$$\geq 2xz + xy + yz + \frac{y(y-1)}{2} \tag{29}$$

$$RHS = \frac{((x+a)+(y-a)+z)^2}{2} = \frac{(x+y+z)^2}{2} \tag{30}$$

We keep shifting until we reach the equality case $x = y = z$ when LHS=RHS. The LHS side increased and the RHS did not, so, at the beginning, LHS $\leq$ RHS. Thus, it is prove the total weight of all edges is at most $\frac{n^2}{2}$.