ON FRACTIONAL REALIZATIONS OF TOURNAMENT SCORE SEQUENCES

by

Kaitlin S. Murphy

A thesis submitted in partial fulfillment
of the requirements for the degree

of

MASTER OF SCIENCE

in

Mathematics

Approved:

David E. Brown, Ph.D. Major Professor	Andreas Malmendier, Ph.D. Committee Member
Brynja Kohler, Ph.D. Committee Member	Richard S. Inouye, Ph.D. Vice Provost for Graduate Studies

UTAH STATE UNIVERSITY
Logan, Utah

2019

ABSTRACT

On Fractional Realizations of Tournament Score Sequences

by

Kaitlin S. Murphy, Master of Science

Utah State University, 2019

Major Professor: David E. Brown, Ph.D.
Department: Mathematics and Statistics

The problem of determining whether a list of nonnegative integers is the score sequence of some round robin tournament, sometimes referred to as the Tournament Score Sequence Problem (or TSSP), can be proposed in the form of an integer program and was determined fully by mathematician H.G. Landau in the 1950's. In this thesis, we examine a generalization of tournaments which allow for fractional arc-weightings; we introduce several related polytopes as well as the new notion of probabilization and prove several results about them.

Fractional scores of a tournament are discussed in the context of relaxing the constraints on the aforementioned integer program to obtain a linear program. The feasible solution space of this linear program forms an n-dimensional polytope. We will prove that the vertices of this polytope are those that correspond to tournaments with integral scores. These results complement the work of M. Barrus in "On fractional realizations of graph degree sequences", *Electronic Journal of Combinatorics* **21** (2014), no. 2, Paper #P2.18.

The intersection of digraph theory, polyhedral combinatorics, and linear programming is a relatively new branch of graph theory. These results pioneer research in this field.

(53 pages)

PUBLIC ABSTRACT

On Fractional Realizations of Tournament Score Sequences

Kaitlin S. Murphy

Contrary to popular belief, we can't all be winners.

Suppose 6 people compete in a chess tournament in which all pairs of players compete directly and no ties are allowed; *i.e.*, 6 people compete in a 'round robin tournament'. Each player is assigned a 'score', namely the number of games they won, and the 'score sequence' of the tournament is a list of the players' scores. Determining whether a given potential score sequence actually is a score sequence proves to be difficult. For instance, $(0, 0, 3, 3, 3, 6)$ is not feasible because two players cannot both have score 0. Neither is the sequence $(1, 1, 1, 4, 4, 4)$ because the sum of the scores is 16, but only 15 games are played among 6 players. This so called 'tournament score sequence problem' (TSSP) was solved in 1953 by the mathematical sociologist H. G. Landau. His work inspired the investigation of round robin tournaments as directed graphs.

We study a modification in which the TSSP is cast as a system of inequalities whose solutions form a polytope in n-dimensional space. This relaxation allows us to investigate the possibility of fractional scores. If, in a 'round-robin'-ish tournament, Players A and B play each other 3 times, and Player A wins 2 of the 3 games, we can record this interaction as a 2/3 score for Player A and a $1/3$ score for Player B. This generalization greatly impacts the nature of possible score sequences. We will also entertain an interpretation of these fractional scores as probabilities predicting the outcome of a true round robin tournament.

The intersection of digraph theory, polyhedral combinatorics, and linear programming is a relatively new branch of graph theory. These results pioneer research in this field.

ACKNOWLEDGMENTS

"If I have seen further than others, it is by standing on the shoulders of giants."
- Sir Isaac Newton

I am extremely grateful to all of the individuals who have supported me in any manner during my graduate studies, but I would like to mention a few of the giants that have enabled me to succeed.

To my father, Greg, my mother, Becky, and my brothers, Christopher, Matthew, and Andrew, who have never let me get away with mediocrity. My brothers form a superhero team of compassion: Chris with his wit and wisdom, Matt with his empathy and emotion, and Andrew with his courage and confidence. Thank you for being there for me, even during your hard times.

For their continued guidance and support I would like to thank my advisor Dr. Dave Brown, and my committee members Dr. Brynja Kohler and Dr. Andreas Malmendier. They have introduced me to opportunities that have changed my education in the form of research, conferences, and life advice. I would also like to recognize Dr. Michael Barrus of the University of Rhode Island for laying the groundwork for my findings.

A great deal of credit goes to Dr. Brent Thomas, who has very selflessly advised and revised my work throughout my master's program. Not only did he teach me almost everything I know about graph theory, he taught me most everything I know about researching.

In addition, I am grateful for the wonderful faculty and staff of the USU Math and Stat department. In particular, Linda Skabelund and Gary Tanner, for making sure I graduated and had a job, but also for the fun and games.

Lastly, to my many friends for eating with me, crying with me, and laughing with me. In particular, I relied heavily on the shoulders of Camille Wardle, Tyler Bowles, Brandon Ashley, Sam Schwartz, and Jessie Whittaker. May all the greatest things come to you.

Kaitlin S. Murphy

CONTENTS

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

The tournament score sequence problem (TSSP) of determining which lists of integers coincide with tournaments was completely determined by a mathematician by the name of H. G. Landau when he provided necessary and sufficient conditions characterizing such lists. The work presented in this thesis will use techniques from linear programming and fractional graph theory to investigate the feasible region obtained when viewing the TSSP as a system of linear inequalities, essentially allowing fractional scores in a tournament.

The motivation for this research came mostly from the recent work of Dr. Michael Barrus in a paper published in 2013 [2]. In this paper, Barrus approaches realizations of graphic degree sequences from a degree-based perspective while allowing fractional weightings on edges. This is achieved by relaxing the conditions on an integer programming interpretation of a realization of a degree sequence to a linear program. The feasible region of the associated linear program is the intersection of a finite number of halfspaces, hence a convex polytope. The findings presented in this thesis are complementary to the work of Barrus, but lie instead in the realm of directed graphs.

The concept of fractional tournaments has been studied in the past from a matrix perspective as opposed to a degree perspective (like the one taken in this paper). In [9], a generalized tournament matrix is defined as an $n \times n$ matrix P with nonnegative entries for which the property $P + P^{tr} = J - I$ holds where J denotes the matrix of 1's and I the identity matrix. This paper proposes several methods for ranking players in a tournament and possible handicapping measures that could be taken. A similar matrix theory approach is discussed briefly in [12] by Bryan Shader.

In Chapter 2, foundational material is presented on the basics of graph theory, optimization, and fractional graph theory. Two fundamental optimization problems are investigated from both the integer programming and linear programming perspective to demonstrate the usefulness of relaxing integer constraints. The motivating work of Dr. Michael Barrus in "On Fractional Realizations of Tournament Score Sequences" (2013) is introduced.

Chapter 3 consists of the novel fractional analogues of directed graphs, tournaments, and score sequences which will serve as the basis for the main results of this thesis presented in the following chapters.

Two of the polytopes in question are defined and studied in Chapter 4. We show that if a score sequence is of the form $(0, 1, 2, ..., n-1)$, there is a unique fractional realization of the sequence. It is also shown that a point of the polytope of possible arc weightings for a given sequence is a vertex of the polytope if and only if all weightings are integral.

In Chapter 5 the arc weightings of fractional complete directed graphs are interpreted as probabilities that may, in a sense, "predict" the outcome of a round robin tournament between the vertices. This concept of an expected outcome tournament and an associated effective score sequence is developed and an associated polytope is studied.

Chapter 6 concludes this work with a brief foray into possible future research directions.

CHAPTER 2

PRELIMINARIES

2.1 Graphs, Digraphs, and Tournaments

2.1.1 Graphs

A *graph* G is an ordered pair (V, E) in which $V = V(G)$ is a set of vertices and $E = E(G)$ is a set of edges disjoint from V together with an incidence function $\psi_G : E \to \binom{V}{2}$ that associates each edge with an unordered pair of vertices. Note that some authors may allow $\psi_G : E \to \binom{V}{2} \cup V$. Such graphs may contain *loops*, i.e. edges joining a vertex to itself. We assume *loopless* graphs, so each edge is assigned to an unordered pair of distinct vertices. For ease of notation, we use uv (equivalently vu) to represent the unordered pair $\{u, v\}$. Graphs are commonly visualized as vertices and edges, such as in Example 1.

Example 1. Let $G = (V, E)$ with $V = \{v_1, v_2, v_3, v_4, v_5\}$ and $E = \{e_1, e_2, e_3, e_4, e_5\}$ where ψ_G is given by

$$\psi_G(e_1) = v_1v_2, \quad \psi_G(e_2) = v_1v_3, \quad \psi_G(e_3) = v_1v_5, \quad \psi_G(e_4) = v_2v_4, \quad \psi_G(e_5) = v_3v_5$$

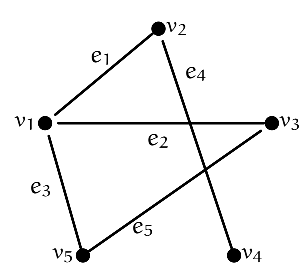


Fig. 2.1: A visual representation of G.

We will often identify an edge with its image under the incidence function. In Example 2.1, we may refer to the edge e_1 as the edge v_1v_2 since $\psi_G(e_1) = v_1v_2$.

A vertex v_i is said to be *adjacent* to a vertex v_j in a graph G if v_iv_j is in the image of ψ_G. A vertex v_i is said to be *incident* to an edge if there exists an $e \in E(G)$ such that $\psi_G(e) = v_iv_k$ for some $v_k \in V$. For a vertex $v_i \in V(G)$ we may refer to the set of all vertices adjacent to v_i in G as the *neighborhood* of v_i, denoted $N_G(v_i)$. Note that in a loopless graph $v_i \notin N_G(v_i)$. The *degree of a vertex* v_i, denoted $d_G(v_i)$, is the number of vertices adjacent to v_i, so $d_G(v_i) = |N_G(v_i)|$. For example, in Figure 2.1 the degree of v_1 in G is $d_G(v_1) = 3$. The subscript serves to clarify the graph in which we are determining the degree of the vertex and may be omitted if the context clearly determines the graph in question.

A *simple graph* is a loopless graph in which the incidence function is injective (one to one). Note that the graph in Example 1 is simple. A simple graph on n vertices is *complete* if the associated incidence function is a bijection. The complete graph on n vertices is unique up to isomorphism and is commonly notated as K_n.

If two edges in the edge set have the same image under the incidence function, the resulting graph is called a *multigraph*.

A *degree sequence* is a nondecreasing sequence of nonnegative numbers representing the degrees of the vertices in a graph G. For example, the degree sequence of G in Figure 2.1 is given by $d = (1, 2, 2, 2, 3)$.

2.1.2 Digraphs

Let $V = \{v_i\}_{i \in I}$ be a set. Define the set $V \bowtie V = \{(v_i, v_j) \mid i, j \in I, i \neq j\}$.
Note that $V \bowtie V$ is a subset of $V \times V$.

Example 2. Let $V = \{v_1, v_2, v_3\}$. Then we have the three following sets:

$$\binom{V}{2} = \{\{v_1, v_2\}, \{v_1, v_3\}, \{v_2, v_3\}\}$$

$$V \times V = \{(v_1, v_1), (v_1, v_2), (v_1, v_3), (v_2, v_1), (v_2, v_2), (v_2, v_3), (v_3, v_1), (v_3, v_2), (v_3, v_3)\}$$

and

$$V \bowtie V = \{(v_1, v_2), (v_1, v_3), (v_2, v_1), (v_2, v_3), (v_3, v_1), (v_3, v_3)\}.$$

A *directed graph* or *digraph* is an ordered pair $D = (V, A)$ where $V = V(D)$ is a set of vertices and $A = A(D)$ is a set of arcs together with an incidence function $\psi_D : A \to V \times V$. If ψ maps instead into the restricted codomain $V \bowtie V$, the digraph is called *loopless*. For our purposes, all digraphs being considered will be loopless. Once again, we omit parentheses, so uv is understood to represent the ordered pair (u, v). Digraphs are commonly represented as vertices and arrows as in Example 3.

Example 3. Let $D = (V, A)$ with $V = \{v_1, v_2, v_3, v_4, v_5\}$ and $A = \{a_1, a_2, a_3, a_4, a_5\}$ where ψ_D is given by

$$\psi_D(a_1) = v_1v_2, \quad \psi_D(a_2) = v_1v_3, \quad \psi_D(a_3) = v_1v_5, \quad \psi_D(a_4) = v_2v_4, \quad \psi_D(a_5) = v_5v_3$$

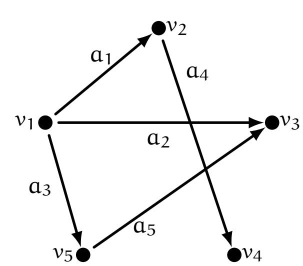


Fig. 2.2: A visual representation of D.

Again we may identify an arc with its image under the incidence function. In Figure 3, arc a_1 may be identified as v_1v_2. The *outdegree* of a vertex v in a digraph D, denoted $d_D(v)$, is the number of outgoing arcs from vertex v. In Figure 3, $d_D(v_5) = 1$ and $d_D(v_3) = 0$.

2.1.3 Tournaments

Given an undirected graph, an *orientation* of the graph is an assignment of exactly one direction to each of the edges of the graph. An orientation may be thought of as a map from $\binom{V}{2}$ to $V \bowtie V$. A *tournament* $T = (V, A)$ is an orientation of a complete loopless undirected graph. One can think of a tournament as a directed graph whose incidence function ψ_T satisfies the following properties: for any two vertices $v_i, v_j \in V(T)$, either $v_i v_j \in \operatorname{Im} \psi_T$ or $v_j v_i \in \operatorname{Im} \psi_T$ as in Example 4.

Example 4. Let $T = (V, A)$ with $V = \{v_1, v_2, v_3, v_4, v_5\}$ and
$A = \{a_1, a_2, a_3, a_4, a_5, a_6, a_7, a_8, a_9, a_{10}\}$ with ψ_T given by

$$\psi_T(a_1) = v_1v_2, \quad \psi_T(a_2) = v_1v_3, \quad \psi_T(a_3) = v_1v_4, \quad \psi_T(a_4) = v_1v_5, \quad \psi_T(a_5) = v_2v_4$$

$$\psi_T(a_6) = v_2v_5, \quad \psi_T(a_7) = v_3v_2, \quad \psi_T(a_8) = v_3v_4, \quad \psi_T(a_9) = v_4v_5, \quad \psi_T(a_{10}) = v_5v_3.$$

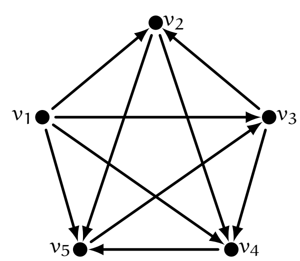


Fig. 2.3: A visual representation of T.

If $v_i v_j \in \operatorname{Im} \psi_T$, we say that vertex v_i *beats* vertex v_j or v_j is beaten by vertex v_i. In a tournament T, we refer to the outdegree of a vertex v as the *score* of vertex v, denoted $s_T(v)$. The score of vertex v_3 in T in Figure 2.3 is $s_T(v_3) = 2$ since v_3 beats v_2 and v_3 beats v_4. A *score sequence* is a nondecreasing sequence of nonnegative integers representing the scores of vertices in a tournament. For example, the score sequence of T in Figure 2.3 is given by $\overrightarrow{s} = (1, 1, 2, 2, 4)$. In the case of labeled digraphs, we may refer to a *score vector* in which the i^{th} entry of the vector is the score of vector v_i. Note that the score vector of T is identical to the score sequence, but that is not always the case.

The incidence function of a *transitive* tournament has the additional constraint that if $v_iv_j \in \operatorname{Im}\psi$ and $v_jv_k \in \operatorname{Im}\psi$ then $v_iv_k \in \operatorname{Im}\psi$. A transitive tournament on n vertices has score sequence $(0, 1, 2, ..., n-1)$.

2.2 Optimization Motivation

Given a condition or parameters, it is natural to seek for a best or "optimal" outcome or solution. Presented below are two well-known optimization problems from the field of graph theory. The first problem is known as "graph coloring" and is a quintessential example of optimization, often introduced in entry level graph theory and combinatorics courses. Many believe that the coloring problem kick-started the entire field of graph theory. It is included here to demonstrate the usefulness of integer programming and linear programming in furthering the understanding and study of even the most cherished problems.

The second example provided is that of biclique covering numbers. This example highlights a deficiency in integer programming and combinatorial methods that can be overcome by the use of linear programming and fractional techniques.

2.2.1 Graph Coloring

A *coloring* of a graph G is an assignment of labels, referred to as *colors*, to the vertices of G. If the colors are assigned so that adjacent vertices get different colors, the coloring is a *proper coloring.*

Example 5. Consider the following proper colorings of planar representations of the cube. Note that on the left, three colors are used, while on the right only two colors are used.

The optimization question then becomes, what is the least number of colors needed to properly color a graph G? This least number is referred to as the *chromatic number* of G, notated as $\chi(G)$.

Obviously, a graph $G = (V, E)$ may be properly colored by assigning every vertex a different color, so $1 \leq \chi(G) \leq |V|$. In general, it is very difficult to determine the chromatic

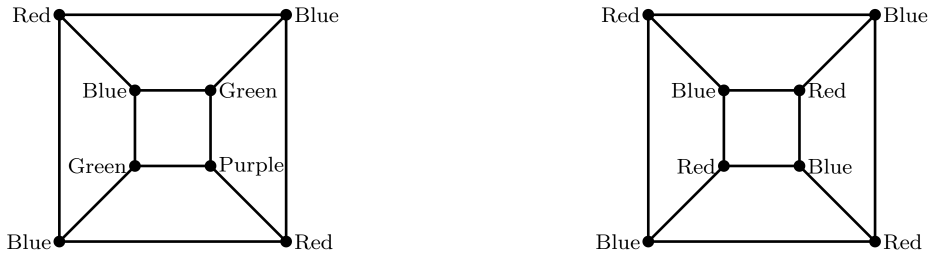


Fig. 2.4: Two colorings of the cube.

number of a graph. If $\chi(G)$ is determined, the typical proof demonstrates a coloring with $\chi(G)$ colors and provides a proof as to why $\chi(G) - 1$ colors is insufficient.

Proposition 1. The cube H in Example 5 has $\chi(H) = 2$.

Proof. Since the graph is connected, $\chi(H) > 1$, and the second coloring from Figure 2.4 demonstrates a coloring using two colors; therefore $\chi(H) = 2$. □

Proposition 2. The complete graph on n vertices, K_n, is the only graph on n vertices with $\chi = n$.

Proof. Note that the degree of any vertex in K_n is $n - 1$, thus n colors are necessary and sufficient for coloring the graph. So $\chi(K_n) = n$. For any graph H on n vertices that is not complete, there are at least two vertices that are not adjacent that can be assigned the same color. Thus, $\chi(H) \leq n - 1$. □

A *bipartite graph* is a nonempty graph (i.e., a graph with edges) in which its vertices can be partitioned into two nonempty sets so that any two vertices in the same partite set are not adjacent.

Proposition 3. A graph H is a bipartite graph if and only if $\chi(H) = 2$.

Proof. Let H be a bipartite graph with parts P_1 and P_2. Assign all vertices in P_1 color 1 and all vertices in part P_2 color 2. This coloring is clearly proper since all vertices with color 1 are nonadjacent as are those vertices with color 2.

Conversely, let H satisfy $\chi(\mathsf{H}) = 2$. Then H has edges and is hence nonempty. All vertices with color 1 are nonadjacent and may be regarded as comprising a partite set; similarly all vertices with color 2 may comprise a partite set. □

The study of graph coloring originated in the 1800s while cartographers attempted to color maps. It was conjectured that four colors was sufficient to color a map so that bordering regions were assigned different colors. A South African mathematician, Francis Guthrie, is credited with postulating this problem, eventually referred to as "The Four Color Problem."

A map of this type is equivalent to a *planar graph*, a graph that has an embedding in the plane with no edges crossing.

Conjecture 1 (The Four Color Problem)**.** A planar graph has $\chi \leq 4$.

For the next 150 years the question remained unsolved, except for a brief stint in the late 1800s in which Alfred Kempe published a proof, only to have it discredited a decade later. The conjecture was eventually proved to be true in the 1970s by mathematicians Kenneth Appel and Wolfgang Haken [1]. To prove the conjecture, Appel and Haken supposed that a planar graph exists with $\chi = 5$ with the minimum number of vertices that such graphs ever have. From this assumption they deduce a set of 1,482 unavoidable forbidden subgraphs for the hypothetical graph, proving that no minimal planar graph with $\chi = 5$ exists, so no planar graph has $\chi = 5$.

The proof is extraordinary in more ways than one. Aside from solving a famous unsolved problem, the proof was one of the first to take advantage of computational power in a major way. In fact, the authors even remarked on the reception of the proof in their paper.

> ...mathematicians...who were not aware of the developments leading to the proof are rather dismayed by the result because the proof made unprecedented use of computers; the computations of the proof make it longer than has traditionally been considered acceptable.

Indeed, the proof of the Four Color Problem was a great achievement, but one would be shortsighted to overlook ingenuity and techniques applied in attempts to solve the Four Color Problem. Although most were unsuccessful, they still had a large influence on the study of graph theory, particularly in computer-aided techniques and algorithms.

In 1889, when Kempe's proof was discredited, Heawood took the opportunity to use the technique to prove what is now called the Five Color Theorem [7].

Theorem 1 (Five Color Theorem, Heawood 1890). *Every planar graph has* $\chi \leq 5$.

It would take over 100 years for this result to be improved. [13]

2.2.2 Biclique Covering

A *biclique* is a graph whose vertices can be partitioned into two bipartite sets, P_1 and P_2, such that no vertices in the same bipartite set are adjacent, but every pair of vertices from different bipartite sets are adjacent. A *biclique cover* of a graph G is a collection of bicliques such that every edge of G is contained in atleast one biclique.

The *biclique cover number* of a graph G, notated $\mathrm{bc}(G)$ is the smallest integer k such that there is a biclique cover of G with k bicliques.

2.3 Fractional Graph Theory

A *linear program* (LP) is an optimization problem expressed in the form:

$$\min \vec{c}^T\vec{x}, \quad \text{subject to} \quad A\vec{x} \geq \vec{b}$$

or

$$\max \vec{c}^T\vec{x}, \quad \text{subject to} \quad A\vec{x} \leq \vec{b}$$

where $A\vec{x} \geq \vec{b}$ is used to mean that each component of $A\vec{x}$ is greater than or equal to its corresponding component in $\vec{b}$ and $A\vec{x} \leq \vec{b}$ is used to mean that each component of $A\vec{x}$ is less than or equal to its corresponding component in $\vec{b}$. An *integer program* (IP) is a linear program with the additional restraint that the components $\vec{x}$ are integral.

It is common to notate linear programs using summations instead of inner products of vectors and systems of inequalities rather than arrays. The indexing set of these summations will be elements of sets which satisfy certain properties. Consider the following integer programming interpretations of the graph coloring and biclique covering problems presented previously.

2.3.1 Coloring Graphs via Integer Programming

We now show that the computation of the chromatic number of a graph can be viewed as computing the optimal solution to a certain integer program (IP).

A graph G is said to have a proper k-coloring if $\chi(G) \geq k$. Let $G = (V, E)$ be a graph with $\chi(G) \leq k$ and $f : V \to \{1, 2, ..., k\}$ a proper k-coloring; define sets $C_1, ..., C_k$ where $C_i = \{v \in V \mid f(v) = i\}$ for each $1 \leq i \leq k$. We refer to each C_i as a *color class*. Since f corresponds to a proper coloring, the proposed color classes partition V. Furthermore, each C_i represents an independent set of vertices in G, since no two adjacent vertices receive the same color assignment under the proper coloring f.

Let $\mathcal{A}$ be the set of all independent sets in G and let $w : A \to \{0, 1\}$. Consider the following IP:

$$\min \sum_{X \in \mathcal{A}} w(X)$$

$$\text{s.t.} \sum_{X : v \in X} w(X) \geq 1 \text{ for all } v \in V$$

We argue, via Propositions 4 and 5 below, that the optimal solution to the IP is the chromatic number of G. Essentially, think of w as a selection function with $w(x) = 1$ if independent set x is to be selected. The constraint ensures that each vertex v in the graph is contained in at least one of the independent sets selected by w. Therefore, the solution to the IP identifies a minimally weighted cover of the vertices of G by independent sets. The propositions below argue that this cover corresponds to a partition into independent sets that is minimal.

Proposition 4. If w is a feasible solution to the IP that does not correspond to a proper coloring, then either w is equivalent to a solution w' that corresponds to a proper coloring or w is not optimal.

Proof. Let w be a feasible solution to the IP that does not correspond to a proper coloring.
Case 1: For some set $X \in \mathcal{A}$ with $w(X) = 1$, there exists a proper subset $X_0 \subset X$ such that $w(X_0) = 1$.
Define an assignment function $w' : \mathcal{A} \to \{0, 1\}$ such that if Y is a proper subset of some $X \in \mathcal{A}$ such that $w(X) = 1$ then $w'(Y) = 0$, else $w'(Y) = w(Y)$. Then w' is a feasible solution such that $\sum_{X\in\mathcal{A}} w'(X) < \sum_{X\in\mathcal{A}} w(X)$. Thus, w is not optimal.
Case 2: For all $X, Y \in \mathcal{A}$ such that $w(X) = w(Y) = 1$ neither X nor Y are proper subsets of the other, but for some $X_1, Y_1 \in \mathcal{A}$ with $w(X_1) = w(Y_1) = 1$, $X_1 \cap Y_1 \neq \emptyset$.
Let $v \in X_1 \cap Y_1$ where $w(X_1) = w(Y_1) = 1$. Since $Y_1 \backslash \{v\}$ is a proper subset of Y_1, $w(Y_1 \backslash \{v\}) = 0$. Define a new assignment function $w' : \mathcal{A} \to \{0, 1\}$ such that $w'(Y_1) = 0$, $w'(Y_1 \backslash \{v\}) = 1$, and $w'(X) = w(X)$ for all $X \neq Y_1$. Note that $\sum_{X\in\mathcal{A}} w'(X) = \sum_{X\in\mathcal{A}} w(X)$. This process can be iterated for any vertex in the intersection of any two independent sets chosen until all pairwise intersections are empty. The resultant function $w^{(n)}$ is a feasible solution such that $\sum_{X\in\mathcal{A}} w^{(n)}(X) = \sum_{X\in\mathcal{A}} w(X)$ and corresponds to a proper coloring of the graph G.

□

Proposition 5. If z is the minimal value of the IP, then $\chi(G) = z$.

Proof. Suppose z is the minimal value of the IP and let $\chi(G) = y$. Clearly $y \leq z$ since z corresponds to a proper coloring of the graph G. If $y < z$, then y corresponds to a solution of the IP that assigns exactly y independent sets a value of 1, contradicting that z is the minimal value. Thus, $z = y = \chi(G)$ as proposed. □

2.3.2 Biclique Coverings via Integer Programming

Let $\mathcal{B}$ be the set of all bicliques in a graph G. We can associate a biclique cover with an assignment function $w : \mathcal{B} \to \{0, 1\}$ where the output signifies if the biclique is included in the cover (value 1) or the biclique is excluded from the cover (value 0). We can formulate an IP as follows:

$$\min \sum_{B \in \mathcal{B}} w(B)$$

$$\text{s.t.} \sum_{B : e \in B} w(B) \geq 1 \text{ for all } e \in E.$$

We minimize the number of bicliques included in the cover with $|E|$ constraints ensuring that each edge $e \in E$ is included in at least one chosen biclique.

Proposition 6. A function w is a feasible solution to the IP if and only if its corresponding set of bicliques covers G.

Proof. ($\Rightarrow$) Suppose $w : \mathcal{B} \to \{0, 1\}$ is a feasible solution to the IP and define the set $A = \{B \in \mathcal{B} \mid w(B) = 1\}$. For each $e \in E$, the constraints require that there exists some biclique $B_e \in A$ such that $e \in B_e$. Thus, every edge of G is contained in a biclique, making A a biclique cover of G.

($\Leftarrow$) Let $\mathcal{B}$ be the set of all bicliques of a graph G and $\mathcal{A}$ be a biclique cover of G. Consider the weighting function $w : \mathcal{B} \to \{0, 1\}$ given by

$$w(B) = \begin{cases} 1 & B \in \mathcal{A} \\ 0 & B \notin \mathcal{A} \end{cases}.$$

Note that for every $e \in E$, there exists an $A_e \in \mathcal{A}$ such that $e \in A_e$ and $w(A_e) = 1$. It follows that all $|E|$ constraints are satisfied and w is a feasible solution to the IP as proposed. $\square$

2.3.3 Optimality in Integer and Linear Programming

An optimization problem has the form

$$\min f_0(\vec{x})$$

$$\text{s.t.} \quad f_i(\vec{x}) \leq b_i, \qquad i = 1, ..., m$$

where the vector $\vec{x} = (x_1, ..., x_n)$ is the optimization variable of the problem, the function $f_0 : \mathbb{R}^n \to \mathbb{R}$ is the objective function, the functions $f_i : \mathbb{R}^n \to \mathbb{R}, i = 1, ..., m$ are the inequality constraint functions, and the constants $b_1, ..., b_m$ are the limits for the constraints. A vector $\vec{x^*}$ is called optimal if it has the smallest objective value among all vectors that satisfy the constraints. The optimization problem is an abstraction of the problem of making the best possible choice of a vector in $\mathbb{R}^n$ from a set of candidates and has applications in many different areas including engineering, electronic design automation, automatic control systems, and optimal design problems arising in various fields of engineering. Optimization is also widely used in the areas of finance, network design and operation, and scheduling.

In general, solving many kinds of optimization problems is still a very daunting task, however there exist some promising approaches for linear programs in particular. These approaches as well as further theory in the area of linear programming can be found in *Convex Optimization* and *Combinatorial Optimization: Algorithms and Complexity* [[3], [11]]. One popular method relies on the concept of duality.

Given a linear program of the form

$$\max \vec{c}^T \vec{x}, \text{ subject to } A\vec{x} \leq b$$

we refer to this LP as the *primal LP* and define its dual as the LP given by

$$\min \vec{b}^T \vec{y}, \text{ subject to } A^T \vec{y} \geq c.$$

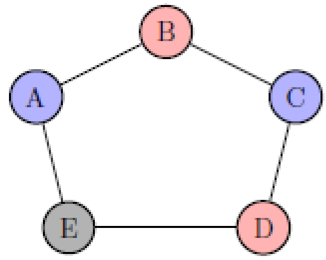
The properties of a primal LP and its dual LP have been studied extensively. For a more in-depth look at their behavior and methods for finding solutions I recommend [11]. The principles of weak vs. strong duality are covered in the recommended literature and will be useful in our study of the fractional chromatic number and fractional biclique cover number below. Linear programs satisfy the principle of strong duality; that is, if the primal LP is bounded from above, then the dual LP is bounded from below and the optimal solution to the dual LP will be equal to the optimal value of the primal LP and vice versa. However, in the case of an integer program, strong duality may not hold, in which case the dual IP does not always have an equal optimal value to the primal LP. This gap, referred to as the duality gap, is difficult to classify or calculate, and thus the dual of a graph property cannot necessarily be used to determine the optimal primal quantity. These calculations are not completely useless though, the principle of weak duality dictates that dual problems place a bound on their primal counterparts. Thus, relaxing the constraints of our initial integer programs to consider linear programs can allow for a more in depth study of their optimal values by allowing more flexibility in the study of their duals and other properties and methods.

2.3.4 Fractional Graph Coloring

The study of fractional graph theory considers the effects of relaxing constraints, such as those proposed in the two previous integer programs, to allow for real-valued values. The fractional chromatic number is commonly used to demonstrate the usefulness of such a relaxation.

Consider a situation in which n committees regularly meet for one hour on the first of each month, but some individuals are members of multiple committees. A schedule must be created where each committee meets for one hour with all of their members. This problem can be visualized as a conflict graph where a vertex represents a committee and vertices are adjacent if they have a common member. If we let our colors represent time slots, a proper coloring of this graph G corresponds to an acceptable schedule. Therefore, $\chi(G)$ is the fewest number of one hour time-slots needed to accommodate all of the committees.

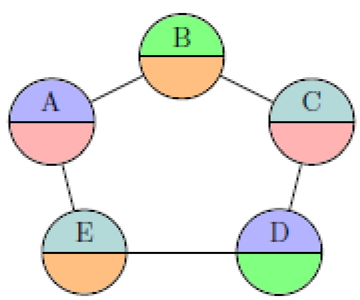
Consider a set of five committees $\{A, B, C, D, E\}$ with a cyclical conflict graph shown below.



This conflict graph may coincide with a schedule such as the following.

Committee	Meeting Time
A, C (Blue)	9:00 - 10:00
B, D (Red)	10:00 - 11:00
E (Gray)	11:00 - 12:00

Note that the conference rooms are being utilized for three hours, but one room sits empty while committee E is meeting. If we allow the meetings to be broken up into two halves, we can use this space more efficiently. This can be represented as a fractional coloring of the conflict graph where each vertex is assigned two colors and no adjacent vertices have colors in common. An example of such a coloring is given here.



This graph may correspond to a schedule like so:

Committee	Meeting Time
A, D (Blue)	9:00 - 9:30
A, C (Red)	9:30 - 10:00
E, C (Teal)	10:00 - 10:30
E, B (Orange)	10:30 - 11:00
D, B (Green)	11:00 - 11:30

Here, committees A, B, C, and E are given hour long time slots with no interruption and committee D has an hour time slot with a break from 9:30 to 11:00. In this scenario the conference rooms are both being used at all times for two and a half hours, the optimal solution.

This fractional chromatic number example provides a very accessible example of the usefulness of studying fractional properties of graphs. Now that we have a general idea of why this study might be useful to us, let's take a look under the hood at the mechanics of relaxing these integer constraints.

Allowing the committees meetings to split into two halves is analogous to relaxing the integer constraint on w in the integer program presented in Section 2.3.1 to obtain a fractional weighting function $w_f : \mathcal{A} \to [0, 1]$ and the following linear program:

$$\min \sum_{X \in \mathcal{A}} w_f(X)$$

$$\text{s.t.} \sum_{X : v \in X} w_f(X) \geq 1 \text{ for all } v \in V.$$

This linear program certainly has some inconvenient aspects. For one, it assigns values to independent sets of vertices, which may be hard to enumerate in some larger or irregular graphs. To overcome these hardships, we will investigate the dual linear program. Let $d_f : V \to [0, 1]$ be a weighting function and consider the dual LP given by

$$\max \sum_{v \in V} d_f(v)$$

$$\text{s.t.} \sum_{v:v\in X} d_f(v) \leq 1 \text{ for all } X \in \mathcal{A}$$

where $v : v \in X$ indicates that one should sum over all vertices $v \in V$ that are in some independent set X. Here each independent set contributes a constraint, as opposed to each vertex contributing a constraint in the primal LP. Also, the dual LP assigns values to vertices instead of independent sets, which alleviates much of the hardship found in the primal LP. As an example, we study the fractional chromatic number of C_n, the cycle on n vertices.

Lemma 1. *The largest independent set in a C_n has $\frac{n}{2}$ vertices if n is even and $\frac{n-1}{2}$ if n is odd.*

Proof. Consider the case where n is even. Take every other vertex along the path to be part of the independent set. This is done without loss of generality since the cycle is vertex transitive. Note that the size of this chosen set is $\frac{n}{2}$ and the addition of any other vertex would yield the set not independent. Thus, the set is maximal. When n is odd, consider C_{n+1} (where $n+1$ is even by assumption of n odd) and find its maximal independent set. This set will have size $\frac{n+1}{2}$. Delete one vertex in the independent set and make their neighbors adjacent. These neighbors were not in the independent set, so the remaining $\frac{n+1}{2} - 1 = \frac{n-1}{2}$ vertices form a maximal independent set. □

Proposition 7. The fractional chromatic number for C_n is given by

$$\chi_f(C_n) = \begin{cases} 2 & n \text{ is even} \\ \frac{2n}{n-1} & n \text{ is odd} \end{cases}.$$

Proof. Let z be the optimal solution to the fractional dual LP. By the principle of strong duality for linear programs, $\chi_f(C_n) = z$. Since C_n is vertex transitive, the weighting function $d_f : V \to [0,1]$ is a constant function, say $d_f(v) = \alpha$ for every $v \in V$. Now each independent set imposes a constraint on the system and we have for each independent set X a constraint of the form

$$\sum_{v:v\in X} d_f(v) = \sum_{v:v\in X} \alpha = |X|\alpha \geq 1.$$

Since the dual LP is a maximization of the sums of d_f values, the optimal solution will be the independent set with the most vertices, hence the above lemma. Thus, let

$$\alpha = \begin{cases} 2 & n \text{ is even} \\ \frac{2n}{n-1} & n \text{ is odd} \end{cases}.$$

These α values are maximal and satisfy all constraints imposed by independent sets and it follows that

$$\chi_f(C_n) = z = \sum_{v\in V} d_f(v) = \sum_{v\in V} \alpha = n\alpha = \begin{cases} 2 & n \text{ is even} \\ \frac{2n}{n-1} & n \text{ is odd} \end{cases}$$

as proposed. □

Note that a cycle on an even number of vertices is a bipartite graph, and this result is consistent with the previous finding on bipartite graphs. Also, this result is consistent with our committee meeting example when $n = 5$.

2.3.5 Fractional Biclique Coverings

Similarly, we can investigate the fractional relaxation of the biclique cover number introduced earlier. The natural fractional analog can be obtained by relaxing the constraint that $w : \mathcal{B} \rightarrow \{0, 1\}$ to $w_f : \mathcal{B} \rightarrow [0, 1]$. Thus we allow the fractional inclusion of bicliques in the cover with the restraint that the total weight of the bicliques covering any edge in the graph must be at least one. This generates an LP

$$\min \sum_{B\in\mathcal{B}} w_f(B)$$

$$\text{s.t.} \sum_{B: e \in B} w_f(B) \geq 1 \text{ for all } e \in E.$$

Unfortunately, the set $\mathcal{B}$ of all bicliques of a graph has proven to be a very difficult set to determine given a graph, yielding the above LP rather impractical. To sidestep this issue, we will investigate the dual LP. To prepare, it is helpful to reformulate the original or primal LP in vector form. Since G is finite, the set $\mathcal{B}$ of all biclique subgraphs of G is finite. Let $|\mathcal{B}| = n$. Without loss of generality, index the bicliques in $\mathcal{B}$ with the values $1, 2, ..., n$. Also, let $|E(G)| = m$ and index the edges of G with the values $1, 2, ..., m$. Define $\vec{w_f}$ as a vector with n components where the ith component is the value of the ith biclique under the assignment function w_f.

Construct the matrix $A_{m \times n}$ where entry $a_{ij} = \begin{cases} 1 & e_i \in B_j \\ 0 & e_i \notin B_j \end{cases}$. Note that each row of A corresponds to an edge of G, so we can reformulate our primal LP as

$$\min \vec{1}^T \vec{w_f}$$

$$\text{s.t.} \quad A\vec{w_f} \geq 1.$$

The dual LP is then given by

$$\max \vec{1}^T \vec{d_f}$$

$$\text{s.t.} \quad A\vec{d_f} \leq 1.$$

Note that A^T is an $n \times m$ matrix, thus $\vec{d_f}$ is an $m \times 1$ column vector. Consider the ith entry of $\vec{d_f}$ to be the value assigned to the ith edge via the function d_f, $d_f(e_i) = \vec{d_f}_i$. With this interpretation, each row of A corresponds to a biclique, so each biclique contributes a constraint. So, we can write $d_f : E \to [0, 1]$ and the following reformulated dual LP

$$\max \sum_{e \in E} d_f(e)$$

$$\text{s.t.} \sum_{e: e \in B} d_f(e) \leq 1 \quad \text{for all } B \in \mathcal{B}.$$

This dual LP is more practical since it allows us to weight edges instead of bicliques and in general this is an easier task as there are usually fewer edges than bicliques to consider. To demonstrate the usefulness of this approach, we will calculate the fractional biclique cover number of the complete graph on n vertices.

Observation. A complete bipartite graph $K_{a,b}$ has ab edges.

Lemma 2. K_n *contains all possible biclique subgraphs* $K_{a,b}$ *where* $a + b \leq n$.

Proof. Let $a + b \leq n$. Choose a vertices of the K_n and b of the remaining vertices to be the vertex set of a subgraph K_n'. Let the edge set of K_n' be all edges connecting one of the a vertices to one of the b vertices. Note that since K_n is complete, each of the a vertices are connected to every one of the b vertices and vice versa. Also, no edges between two of the a vertices or between two of the b vertices was included, making K_n' a complete bipartite graph, $K_{a,b}$. □

Fact. The number of edges in a K_n is $\frac{n(n-1)}{2}$.

Proposition 8. The fractional biclique cover number of the complete graph on n vertices is given by

$$bc_f(K_n) = \begin{cases} \frac{2(n-1)}{n} & n \text{ is even} \\ \frac{2n}{n+1} & n \text{ is odd} \end{cases}.$$

Proof. By the strong duality principle of linear programs, if z is the optimal solution to the fractional dual we developed, then $bc_f(K_n) = z$. It suffices to find the optimal solution to the dual fractional linear program proposed above. Since K_n is edge transitive, the function $d_f : A \to [0, 1]$ must assign the same weight to each edge, say $d_f(e) = \alpha$ for all $e \in E(K_n)$.

Note that in the dual LP there is a restriction for each biclique of K_n. By the lemma above, we know that for any a and b such that $a + b \leq n$, $K_{a,b}$ is a biclique subgraph of K_n with ab edges. Therefore we have the constraint,

$$\sum_{e:e\in K_{a,b}} d_f(e) = \sum_{e:e\in K_{a,b}} \alpha = (ab)\alpha \leq 1. \tag{2.1}$$

By the lemma, if n is even, then $K_{\frac{n}{2},\frac{n}{2}}$ is a biclique of K_n and if n is odd then $K_{\frac{n+1}{2},\frac{n-1}{2}}$ is a biclique and these are both maximal since they both contain all n vertices. Furthermore, they maximize ab such that $a+b=n$. By Equation 2.1, if n is even we have the constraint $\left(\frac{n}{2}\right)\left(\frac{n}{2}\right)\alpha = \frac{n^2}{4}\alpha \leq 1$. Similarly, if n is odd we have the constraint $\left(\frac{n+1}{2}\right)\left(\frac{n-1}{2}\right)\alpha = \frac{n^2-1}{4}\alpha \leq 1$. The optimal solution maximizes α, hence the optimal solution is

$$\alpha = \begin{cases} \frac{4}{n^2} & n \text{ is even} \\ \frac{4}{n^2-1} & n \text{ is odd} \end{cases}. \tag{2.2}$$

It follows then from the above fact that,

$$\mathrm{bc}_f(K_n) = z = \sum_{e\in E} d_f(e) = \sum_{e\in E} \alpha = \frac{n(n-1)}{2}\alpha = \begin{cases} \frac{2(n-1)}{n} & n \text{ is even} \\ \frac{2n}{n+1} & n \text{ is odd} \end{cases}.$$

□

2.4 Degree and Score Sequences

2.4.1 Relevant Theorems and Results

A nonincreasing sequence of nonnegative integers is called *graphic* if it is the degree sequence of at least one graph. The following two well-known results provide necessary and sufficient conditions for a nonincreasing sequence of nonnegative integers to be graphic.

Theorem 2 (Erdős - Gallai 60 [4]). *A nonincreasing sequence of nonnegative integers* $S : d_1, d_2, ..., d_p$ *with* $p \geq 2$ *is graphic if and only if for every* k *in* $1 \leq k \leq n$

$$\sum_{i=1}^{k} d_i \leq k(k-1) + \sum_{i=k+1}^{p} \min(d_i, k)$$

and $\sum_{i=1}^{p} d_i$ *is even.*

Theorem 3 (Havel 55 [6] and Hakimi 62 [5]). *A nonincreasing sequence of nonnegative integers* $S : d_1, ..., d_p$ *with* $p \geq 2$ *is graphical if and only if the sequence* $S' : d_2 - 1, ..., d_{d_1+1} - 1, d_{d_1+2}, ..., d_p$ *is graphical.*

Proof. ($\Leftarrow$) If a sequence $S' : d_2 - 1, ..., d_{d_1+1} - 1, d_{d_1+2}, ..., d_p$ is graphic and G' is a graph that realizes S', label the vertices of G' so that $d(v_2) = d_2 - 1, d(v_3) = d_3 - 1, ..., d(v_{d_1+1}) = d_{d_1+1} - 1, d(v_{d_1+2}) = d_{d_1+2}, ...,$ and $d(v_p) = d_p$. Then the graph G obtained by taking the graph G' and appending a vertex v_1 adjacent to the first d_1 vertices in order of the labeling is a graph with degree sequence $S = d_1, ..., d_p$, making S graphic.

($\Rightarrow$) Suppose $S = d_1, ..., d_p$ is a graphic sequence. Among all graph with degree sequence S choose the graph G such that

1. $V(G) = \{v_1, ..., v_p\}$ and $d(v_i) = d_i$ for $1 \leq i \leq p$ and

2. the sum of the vertex degree of vertices adjacent to v_1 is maximum.

If v_1 is adjacent to the vertices of degree $d_2, ..., d_{d_1+1}$ then the induced subgraph G' obtained by removing vertex v_1 is a graph with degree sequence S' and the statement holds. Otherwise, suppose there exist two vertices v_i and v_j such that $d_j > d_i$ where $v_1 v_i \in E(G)$ but $v_1 v_j \notin E(G)$. Since $d_j > d_i$, there must exist a vertex v_k such that $v_j v_k \in E(G)$ but $v_i v_k \notin E(G)$. Performing a two switch (deleting the edges $v_1 v_i$ and $v_j v_k$ and adding the edges $v_1 v_j$ and $v_i v_k$ creates a new graph G' with degree sequence S. However, in G' the sum of the vertex degrees of the vertices adjacent to v_1 is greater than in G, contradicting the maximum condition imposed in the choice. Thus, v_1 must be adjacent to the vertices of degree $d_2, ..., d_{d_1+1}$ and the statement holds. □

If G is a graph that has degree sequence d, then G is called a *realization* of the graphic degree sequence d.

Landau's Theorem provides necessary and sufficient conditions for a sequence of nondecreasing integers to be the score sequence of some tournament. The following proof can be found in [10]

Theorem 4 (Landau 53 [8]). *A score sequence* $\overrightarrow{s} = (s_1, s_2, ..., s_n)$ *with* $s_1 \leq s_2 \leq \cdots \leq s_n$ *is the score vector of some tournament* T_n *if and only if*

$$\sum_{i=1}^{k} s_i \geq \binom{k}{2}, \tag{2.3}$$

for $k = 1, 2, ..., n$ *with equality holding when* $k = n$.

Proof. Any k nodes of a tournament are joined by $\binom{k}{2}$ arcs, by definition. Consequently, the sum of the scores of any k nodes of a tournament must be at least $\binom{k}{2}$. This shows the necessity of (2.3).

The sufficiency of (2.3) when $n = 1$ is obvious. The proof for the general case will be by induction. Let j and k be the smallest and largest indices less than n such that $s_j = s_{s_n} = s_k$. Consider the set of integers $(s_1', s_2', ..., s_{n-1}')$ defined as follows

$$\begin{aligned} s_i' = s_i \quad & \text{if} \quad i = 1, 2, ..., j-1 \quad \text{or} \\ & i = k - (s_n - j), ..., k-1, k; \\ s_i' = s_i - 1 \quad & \text{if } i = j, j+1, ..., k - (s_n - j) - 1 \quad \text{or} \\ & i = k+1, k+2, ..., n-1. \end{aligned}$$

From this definition, it follows that

$$s_1' \leq s_2' \leq \cdots \leq s_{n-1}',$$

that $s_1' = s_i$ for s_n values of i, and that $s_i' = s_i - 1$ for $(n-1) - s_n$ values of i. Consequently,

$$\sum_{i=1}^{n-1} s_i' = \sum_{i=1}^{n} s_i - (n-1) = \binom{n-1}{2}.$$

If there exists a tournament T_{n-1} with score vector $(s_1', s_2', ..., s_{n-1}')$, then there certainly exists a tournament T_n with score vector $(s_1, s_2, ..., s_n)$; namely, the tournament consisting of T_{n-1} plus the node p_n, where p_n dominates the s_n nodes p_i such that $s_i' = s_i$ and is dominated by the remaining nodes. Therefore, we need only show that the inequality

$$\sum_{i=1}^{h} s_i' < \binom{h}{2} \tag{2.4}$$

is impossible for every integer h such that $1 < h < n-1$ in order to complete the proof by induction.

Consider the smallest value of h for which inequality (2.4) holds, if it ever holds. Since

$$\sum_{i=1}^{h-1} s_i' \geq \binom{h-1}{2},$$

it follows that $s_h \leq h$. Furthermore, $j \leq h$, since the first $j-1$ scores were unchanged. Hence,

$$s_h = s_{h+1} = \cdots = s_f$$

if we let $f = \max(h, k)$.

Let t denote the number of values of i not exceeding h such that $s_i' = s_i - 1$. Then it must be that

$$s_n \leq f - t. \tag{2.5}$$

Therefore,

$$\begin{aligned}
\binom{n}{2} &= \sum_{i=1}^{h} s_i' + \sum_{i=h+1}^{f} s_i + \sum_{i=f+1}^{n-1} s_i + s_n + t \\
&< \binom{h}{2} + (f-h)s_h + \sum_{i=f+1}^{n-1} s_i + f \\
&\leq \binom{h}{2} + h(f-h) + \sum_{i=f+1}^{n-1} s_i + f \\
&\leq \binom{f}{2} + f(n-f) \\
&\leq \binom{n}{2}.
\end{aligned}$$

Consequently, inequality (2.4) cannot hold and the theorem is proved.

□

2.4.2 Polytopes

A polytope is a generalization of a polygon. The following introduction to polytopes is extracted from [11]. An affine subspace of $\mathbb{R}^d$ of dimension $d-1$ is called a hyperplane. Alternatively, a hyperplane is a set of points x satisfying

$$a_1x_1 + a_2x_2 + \cdots + a_dx_d = b$$

where not all a_i are zero.

A hyperplane determines two halfspaces, namely the sets of points satisfying, respectively,

$$a_1x_1 + a_2x_2 + \cdots + a_dx_d \leq b$$

$$a_1x_1 + a_2x_2 + \cdots + a_dx_d \geq b.$$

The bounded and nonempty intersection of a finite number of halfspaces is called a polytope. Note that halfspaces are convex, thus a polytope is convex as the intersection of finitely many convex sets.

The extreme points of a polytope are known as vertices. A vertex is the unique point satisfying at least $\binom{n}{2}$ of the constraints (hyperplane equations).

For the linear programs of interest in this paper the inequality constraints define a *feasible region* of solutions. This feasible region is the intersection of halfspaces determined by these inequality constraints, thus forming a polytope.

Typically in the study of linear programming, one would attempt to optimize an objective function over the feasible region defined by the constraints of a linear program. Optimization techniques, such as the simplex method, have been a rich field of study, especially with the help of technological advances. An overview of various methods for optimization over convex polytopes can be found in [3].

For our purposes, we will simply be investigating the structure and properties of the feasible region of the polytopes constructed, foregoing the optimization of some objective function.

2.4.3 Fractional Graph Degree Sequences

In [2] a realization of a degree sequence, d, is associated with a solution to the linear program

$$\sum_i x_{ij} = d_j, \qquad 1 \leq j \leq n$$

$$x_{ij} \in \{0, 1\}, \qquad 1 \leq i \leq j \leq n,$$

where $x_{ij} = 1$ is interpreted to mean the edge ij is present in the graph and $x_{ij} = 0$ otherwise. This integer program is then relaxed to a linear program by allowing $x_{ij} \in [0, 1]$ seen below.

$$\sum_i x_{ij} = d_j, \qquad 1 \leq j \leq n$$

$$0 \leq x_{ij} \leq 1, \qquad 1 \leq i < j \leq n.$$

We say that these conditions describe "fractional" realizations of degree sequences.

A polytope, $\mathcal{P}(d)$ consisting of all vectors (x_{ij}) whose coordinates are lexicographically indexed that satisfy the above conditions for a given degree sequence is defined. Given a point x in the polytope $\mathcal{P}(D)$, the *fractional realization of d corresponding to x* is defined to be the labeling of the edges of the complete graph on n vertices such that the edge ij receives the label x_{ij} for all pairs i, j of distinct elements in $\{1, ..., n\}$. This point x is sometimes referred to as the *characteristic vector* of the fractional realization. Consider the following fractional realizations of the degree sequence $d = (1, 1, 1, 1, 1, 1)$.

Barrus notes that the extreme points of this polytope are, in some sense, generalizations of the realizations of d. Note that each integral point in $\mathcal{P}(D)$ (in other words a 0/1-point) is a vertex of the polytope since it satisfies $\binom{n}{2}$ of the conditions of the linear program with equality. For some degree sequences however, there may exist non-integral vertices in the

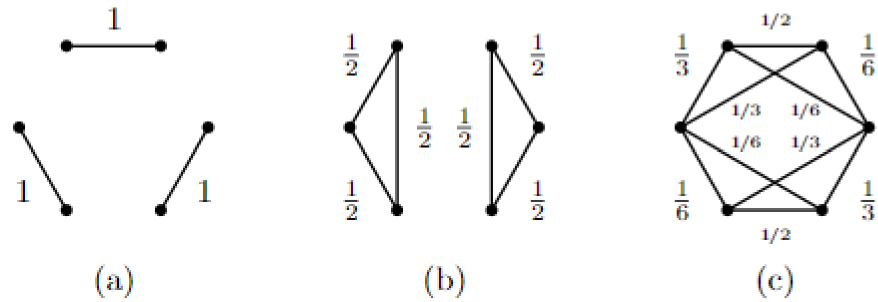


Fig. 2.5: Fractional realizations of (1,1,1,1,1,1).

polytope $\mathcal{P}(\mathsf{D})$. For example, for the degree sequence $\mathsf{d} = (1, 1, 1, 1, 1, 1)$, the characteristic vector for the realization in Figure 2.5(b) is a non-integral vertex of the polytope $\mathcal{P}(\mathsf{D})$. Thus, the vertices of $\mathcal{P}(\mathsf{D})$ may or may not correspond to simple graph realizations of d.

The following result characterizing the vertices of the polytope $\mathcal{P}(\mathsf{D})$ is proven.

Theorem 5 (Barrus 13 [2]). *Given a graphic list d, let h be a point of $\mathcal{P}(\mathsf{d})$,and let H be the fractional realization of d corresponding to h. The point h is a vertex of $\mathcal{P}(\mathsf{D})$ if and only if the edges of H labeled with nonintegral coordinates of h form vertex disjoint odd cycles. Furthermore, there are an even number of these cycles. and the nonintegral coordinates of h all equal 1/2.*

It is noted that fractional realizations of a degree sequence form a special case in the study of b-matchings of a graph. The theorem above is essentially a reformation of some previous known results from the study of fractional perfect b-matchings into the language of degree sequences.

A sequence is said to be *decisive* if $\mathcal{P}(\mathsf{D})$ is a 0/1-polytope. The remaining results of [2] are dedicated to characterizing the decisive sequences and graphs via a number of different techniques including the development of a set of 70 minimal forbidden induced subgraphs for decisive graphs.

The first characterization presented focuses on a particular pattern of adjacencies and non-adjacencies referred to as a (3,3)-blossom.

Theorem 6 (Barrus 13 [2]). *For a graphic sequence d, the following are equivalent:*

(1) d *is a decisive sequence;*

(2) No integral realization of d *contains an integral* (k, ℓ)*-blossom for any odd* $k, \ell \geq 3$*;*

(3) No integral realization of d *contains an integral (3,3)-blossom.*

The second characterization focuses on 70 potential induced subgraphs and being able to partition the vertex set of a graph into three sets that satisfy certain adjacency properties.

The third characterization does so in terms of the numerical values of the terms in the degree sequence and draws inspiration from the Erdős-Gallai conditions.

Theorem 7 (Barrus 13 [2])**.** *Let* $d = (d_1, ..., d_n)$ *be a graphic list in weakly decreasing order. Let* k *be the largest integer such that* d *satisfies the* k*th Erdős-Gallai inequality with equality. The list* d *is a decisive sequence if and only if one of the following is true:*

(1) $k = \max\{i : d_i \geq i - 1\}$*;*

(2) the number $\ell = \max\{i : d_i \geq k$ *and* $i > k\}$ *exists and satisfies one of*

  (i) $\ell - k \leq 5$*;*

  (ii) $(d_{k+1} - k, ..., d_\ell - k)$ *is one of*

$$(4, 2, 2, 2, 2, 2), (3, 3, 3, 3, 3, 1), (m, 1^{(m+2)}), ((m + 1)^{m+2}, 2)$$

  where $m \geq 3$*.*

It is also observed that if d is a threshold sequence, then d has a unique fractional realization.

CHAPTER 3

DIRECTED ANALOGUES TO FRACTIONAL GRAPH THEORY

3.1 Fractional Directed Graphs

Define a *fractional directed graph* as an ordered quadruple $(V, A, \psi, \psi_\alpha)$ consisting of a vertex set V with $|V| = n$, arc set A with $|A| = 2\binom{n}{2}$, an incidence bijection $\psi : A \to V \bowtie V$, and a weighting function $\psi_\alpha : A \to [0, 1]$ with the restriction that if $\psi(a_i) = v_k v_l$ and $\psi(a_j) = v_l v_k$ then $\psi_\alpha(a_i) + \psi_\alpha(a_j) = 1$. For clarity of notation, if $\psi(a_i) = v_k v_l$, we refer to $\psi_\alpha(a_i)$ as α_{kl}. Thus, $\alpha_{kl} + \alpha_{lk} = 1$ for all $k \neq l \in [n]$.

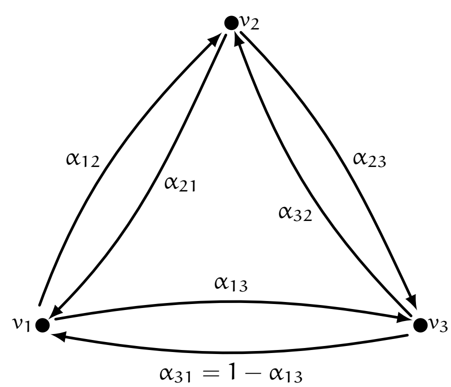


Fig. 3.1: A fractional directed graph.

Let $\overrightarrow{\alpha}' \in [0, 1]^{2\binom{n}{2}}$ be a list of all α_{ij} associated with a fractional directed graph. Since $\alpha_{ij} + \alpha_{ji} = 1$ for all $i \neq j \in [n]$, all of the information in $\overrightarrow{\alpha}'$ can be coded into a vector $\overrightarrow{\alpha} \in [0, 1]^{\binom{n}{2}}$ where the weightings are indexed lexicographically. This can significantly simplify the representative figures of the graphs as in Example 6.

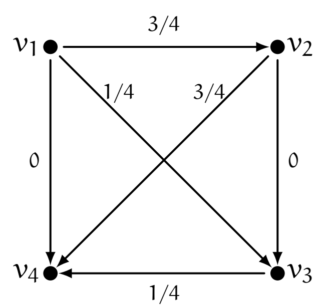
Example 6. Let $V = \{v_1, v_2, v_3, v_4\}$, $A = \{a_1, ..., a_{12}\}$,

$$\psi(a_1) = v_1v_2, \quad \psi(a_2) = v_2v_1, \quad \psi(a_3) = v_1v_3, \quad \psi(a_4) = v_3v_1,$$
$$\psi(a_5) = v_1v_4, \quad \psi(a_6) = v_4v_1, \quad \psi(a_7) = v_2v_3, \quad \psi(a_8) = v_3v_2,$$
$$\psi(a_9) = v_2v_4, \quad \psi(a_{10}) = v_4v_2, \quad \psi(a_{11}) = v_3v_4, \quad \psi(a_{12}) = v_4v_3$$

and

$$\psi_\alpha(a_1) = \alpha_{12} = \frac{3}{4}, \quad \psi_\alpha(a_2) = \alpha_{21} = \frac{1}{4}, \quad \psi_\alpha(a_3) = \alpha_{13} = \frac{1}{4}, \quad \psi_\alpha(a_4) = \alpha_{31} = \frac{3}{4},$$
$$\psi_\alpha(a_5) = \alpha_{14} = 0, \quad \psi_\alpha(a_6) = \alpha_{41} = 1, \quad \psi_\alpha(a_7) = \alpha_{23} = 0, \quad \psi_\alpha(a_8) = \alpha_{32} = 1,$$
$$\psi_\alpha(a_9) = \alpha_{24} = \frac{3}{4}, \quad \psi_\alpha(a_{10}) = \alpha_{42} = \frac{1}{4}, \quad \psi_\alpha(a_{11}) = \alpha_{34} = \frac{1}{4}, \quad \psi_\alpha(a_{12}) = \alpha_{43} = \frac{3}{4}.$$

The $\overrightarrow{\alpha} = \left(\frac{3}{4}, \frac{1}{4}, 0, 0, \frac{3}{4}, \frac{1}{4}\right)$ and we can represent this fractional digraph as follows.



Define the *fractional score* of a vertex, v_i, of a fractional directed graph to be the sum of the arc weights of outgoing arcs, $s_i^f = \sum_{j:j\neq i} \alpha_{ij}$. We analogously define a *fractional score sequence* to be a sequence $\overrightarrow{s}^f = (s_1^f, s_2^f, ..., s_n^f)$ of n nonnegative real values in nondecreasing order for which there exists at least one fractional digraph on n vertices such that the score of vertex $v_i = s_i^f$ for all $i \in [n]$.

3.2 Fractional Tournaments and Fractional Score Sequences

Associate an integer realization of a tournament score sequence with a solution to an integer programming problem as follows: Let $\overrightarrow{s}$ be a score sequence of length n and consider a vertex set of size n. We associate a value $\alpha_{ij} \in \{0, 1\}$ to each unordered pair (v_i, v_j) of distinct vertices. We interpret $\alpha_{ij} = 1$ to mean that the arc originating at vertex v_i ending at vertex v_j is present in the tournament. If the arc from vertex v_i to vertex v_j is not present in the tournament, then $\alpha_{ij} = 0$ and it must be that $\alpha_{ji} = 1$. We can associate realizations of $\overrightarrow{s}$ with vector solutions to the integer program:

$$\sum_{j:i \neq j} \alpha_{ij} = s_i, \qquad 1 \leq i \leq n$$

$$\alpha_{ij} + \alpha_{ji} = 1 \qquad 1 \leq i, j \leq n, \qquad i \neq j$$

$$\alpha_{ij} \in \{0, 1\} \qquad 1 \leq i, j \leq n, \qquad i \neq j.$$

CHAPTER 4

FRACTIONAL REALIZATIONS OF SCORE SEQUENCES

A fractional analogue is constructed by relaxing the constraints to allow α_{ij} to be a real number between 0 and 1. Consider a complete loopless digraph on n vertices. We associate a nonnegative weighting α_{ij} to the arc originating at v_i terminating at v_j as presented previously. Due to the close relationship of α_{ij} and α_{ji}, it suffices to consider (and show) only the arcs associated with the α_{ij} whose coordinates are lexicographically indexed with $i < j$.

The points $\overrightarrow{\alpha} = (\alpha_{ij}) \in \mathbb{R}^{\binom{n}{2}}$ of arc weightings whose coordinates are lexicographically indexed and their related digraphs that satisfy the following equations for a specific score sequence $\overrightarrow{s}$ are *fractional realizations of the score sequence* $\overrightarrow{s}$.

$$\begin{aligned} \sum_{j:i<j} \alpha_{ij} + \sum_{j:i>j} (1 - \alpha_{ji}) = s_i \qquad 1 \le i \le n \\ 0 \le \alpha_{ij} \le 1 \qquad 1 \le i < j \le n \end{aligned} \tag{4.1}$$

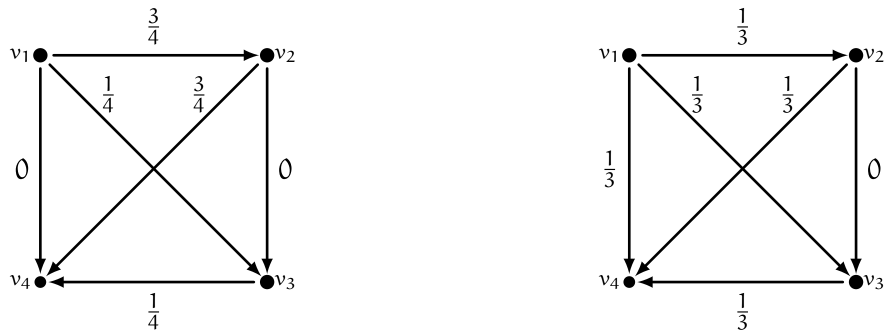
4.1 The Polytope $\mathrm{Frac}_{\vec{\alpha}}(\vec{s})$

Define the polytope $\mathrm{Frac}_{\overrightarrow{\alpha}}(\overrightarrow{s})$ to be the points $\overrightarrow{\alpha}$ that satisfy Equation 4.1 for a given sequence $\overrightarrow{s}$.

All $\overrightarrow{\alpha} \in \mathrm{Frac}_{\overrightarrow{\alpha}}(\overrightarrow{s})$ are fractional realizations of the score sequence $\overrightarrow{s}$. We refer to the complete loopless digraph with arc weightings defined by $\overrightarrow{\alpha}$ as the *fractional realization of* $\overrightarrow{s}$ *corresponding to* $\overrightarrow{\alpha}$.

For clarity, we will refer to a realization in the traditional sense, where $\overrightarrow{\alpha} \in \{0, 1\}^{\binom{n}{2}}$ as an *integer realization*. Notice that integer realizations of the score sequence $\overrightarrow{s}$ correspond to fractional realizations of $\overrightarrow{s}$ where either α_{ij} or α_{ji} is 1.

We will also use the following notion throughout the paper.



(a) D_1: A fractional realization of $\overrightarrow{s}$ corresponding to $\overrightarrow{\alpha} = (\frac{3}{4}, \frac{1}{4}, 0, 0, \frac{3}{4}, \frac{1}{4})$

(b) D_2: A fractional realization of $\overrightarrow{s}$ corresponding to $\overrightarrow{\alpha} = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3}, 0, \frac{1}{3}, \frac{1}{3})$

Fig. 4.1: Two distinct fractional realizations of $\vec{s} = (1, 1, 2, 2)$

Definition. A *fractional tournament matrix* A is a nonnegative matrix that satisfies $A + A^T = J - I$ where J is the all ones matrix and I is the identity matrix. In the case that all entries are either 0 or 1, we refer to the matrix A as simply a *tournament matrix.*

Given a score sequence $\overrightarrow{s}$, we consider vectors $\overrightarrow{\alpha}$ which satisfy the linear program (1) to be points of the polytope $\mathrm{Frac}_\alpha(\overrightarrow{s})$, thus $\mathrm{Frac}_\alpha(\overrightarrow{s}) \subset \mathbb{R}^{\binom{n}{2}}$. We can view solutions of this linear program as a matrix problem. A possible vector $\overrightarrow{\alpha}$ would have to satisfy the following matrix equation

$$\left[\begin{array}{c|c|c|c}
1 \;\; \ldots \;\; 1 & 0 \;\; \ldots & & 0 \\ \cline{1-1}
 & 1 \;\; \ldots \;\; 1 & 0 \;\; \ldots & 0 \\ \cline{2-2}
 & & 1 \;\; \ldots & 0 \\ \cline{3-3}
-I_{n-1} & & & \vdots \\
 & -I_{n-2} & \ldots & 1 \\
 & & & -1
\end{array}\right]
\begin{bmatrix} \alpha_{12} \\ \vdots \\ \alpha_{1,n} \\ \alpha_{2,3} \\ \vdots \\ \alpha_{n-1,n} \end{bmatrix}
+
\begin{bmatrix} 0 \\ 1 \\ 2 \\ \\ \vdots \\ n-2 \\ n-1 \end{bmatrix}
=
\begin{bmatrix} s_1 \\ s_2 \\ \vdots \\ s_n \end{bmatrix}$$

where

$$D_1 \overrightarrow{\alpha} + b_1 = \sum_{j>1} \alpha_{1j} = s_1$$

$$D_2\vec{\alpha} + b_2 = \sum_{j>2} \alpha_{2j} + (-\alpha_{12}) + 1 = s_2$$

$$\vdots$$

$$D_k\vec{\alpha} + b_k = \sum_{j>k} \alpha_{kj} + \left(-\sum_{j<k} \alpha_{jk}\right) + (k-1) = s_k$$

$$\vdots$$

$$D_n\vec{\alpha} + b_n = \left(-\sum_{j<n} \alpha_{jn}\right) + n - 1 = s_n.$$

Example 7. Consider the first fractional realization of the score sequence $\vec{s} = (1, 1, 2, 2)$ in Figure 4.1. Note that the alpha vector corresponding to this fractional realization is $\vec{\alpha} = (\alpha_{12}, \alpha_{13}, \alpha_{14}, \alpha_{23}, \alpha_{24}, \alpha_{34}) = (\frac{3}{4}, \frac{1}{4}, 0, 0, \frac{3}{4}, \frac{1}{4})$. Then we have,

$$\begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ -1 & 0 & 0 & 1 & 1 & 0 \\ 0 & -1 & 0 & -1 & 0 & 1 \\ 0 & 0 & -1 & 0 & -1 & -1 \end{bmatrix} \begin{bmatrix} \frac{3}{4} \\ \frac{1}{4} \\ 0 \\ 0 \\ \frac{3}{4} \\ \frac{1}{4} \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 2 \\ 2 \end{bmatrix}.$$

In a first attempt to understand the polytopes, observations were made about rather basic digraphs. The following are some of the findings.

Theorem 8. *If a score sequence $\vec{s}$ is of the form $(0, 1, 2, ..., n-1)$, there is only one fractional realization of the score vector which is an integer realization.*

Proof. The theorem will be proved via the principle of mathematical induction. Consider the case where $\vec{s} = (0, 1, 2)$. Without loss of generality say $s(v_1) = 0, s(v_2) = 1$, and $s(v_3) = 2$. Since v_1 has score 0, $\alpha_{12} = 0$ and $\alpha_{13} = 0$ and it follows that $\alpha_{21} = 1$ and $\alpha_{31} = 1$. Then, since v_2 has score 1 and $\alpha_{21} = 1$, the remaining arc α_{23} must be zero and it follows that $\alpha_{32} = 1$. Thus, for $\vec{s} = (0, 1, 2)$ the realization $\vec{\alpha} = (0, 0, 0)$ is forced and the statement holds when $n = 3$.

We proceed by induction on the number of vertices with the induction hypothesis that for any $m < n$, the score sequence $\overrightarrow{s} = (0, 1, 2, ..., m-1)$ has a unique realization.

Let $\overrightarrow{s} = (0, 1, 2, ..., n-1)$ where $s(v_1) = 0, s(v_2) = 1, ..., s(v_n) = n-1$ and consider the sequence of the induced subgraph gained by removing vertex v_n in a realization of $\overrightarrow{s}$. The sequence for the remaining sequence is $\overrightarrow{s}' = (0, 1, 2, ..., n-2)$ since $\alpha_{in} = 0$ for all $1 \leq i \leq n-1$ which is a sequence of the desired form for $n-1$ vertices. Thus, by the induction hypothesis, there exists a unique realization of $\overrightarrow{s}'$, call it S'. Now consider the fractional graph $S = S' \cup \{v_n\}$. Since $\alpha_{in} = 0$ for all $1 \leq i \leq n-1$, $\alpha_{ni} = 1$ for all $1 \leq i \leq n-1$. So, the score sequence for S is $\overrightarrow{s} = (0, 1, 2, ..., n-1)$ and is determined uniquely and the theorem is proved by the principle of mathematical induction. $\square$

Theorem 9. *If $\overrightarrow{\alpha}$ is a point in the polytope $Frac_{\overrightarrow{\alpha}}(\vec{s})$ for some integral score sequence $\overrightarrow{s}$, then $\overrightarrow{\alpha}$ is a vertex of the polytope if and only if $\overrightarrow{\alpha} \in \{0, 1\}^{\binom{n}{2}}$.*

Proof. Let $\overrightarrow{\alpha}$ be a vertex of $\text{Frac}_{\overrightarrow{\alpha}}(\overrightarrow{s})$. Since $\text{Frac}_{\overrightarrow{\alpha}}(\overrightarrow{s})$ is in $\binom{n}{2}$ space, $\overrightarrow{\alpha}$ must be the unique solution to $\binom{n}{2}$ hyperplane conditions. Thus, $\overrightarrow{\alpha}$ could be written as the unique solution to a matrix equation $A\alpha = b$.

Define a simple undirected multigraph $H_{\overrightarrow{\alpha}}$ by $V(H_{\overrightarrow{\alpha}}) = V(D_{\overrightarrow{\alpha}})$ and $E(H_{\overrightarrow{\alpha}}) = \{e_{ij} \mid \alpha_{ij} \in (0, 1)\}$. Let $G_{\overrightarrow{\alpha}}$ be the subgraph induced by the edges of $H_{\overrightarrow{\alpha}}$.

If $E(G_{\overrightarrow{\alpha}})$ is empty, then all $\alpha_{ij} \in \overrightarrow{\alpha}$ are integers and the statement holds. Suppose $E(G_{\overrightarrow{\alpha}})$ is nonempty. Then there exist at least two vertices $v_i, v_j \in V(G_{\overrightarrow{\alpha}})$ such that $v_i v_j \in E(G_{\overrightarrow{\alpha}})$, implying that the arc $v_i v_j \in A(D_{\overrightarrow{\alpha}})$ has nonintegral weighting $\alpha_{ij} \in (0, 1)$. Since $\alpha_{ij} + \alpha_{ji} = 1$, α_{ji} is also nonintegral, so $v_j v_i \in E(G_{\overrightarrow{\alpha}})$. Thus, there exist two edges between v_i and v_j in $G_{\overrightarrow{\alpha}}$. Also, since $\overrightarrow{\alpha} \in \text{Frac}_{\overrightarrow{\alpha}}(\overrightarrow{s})$ and the entries of $\overrightarrow{s}$ are integers, the score of any vertex in $D_{\overrightarrow{\alpha}}$ must be integral. So if $v_i v_j \in E(G_{\overrightarrow{\alpha}})$, implying v_i is the initial point of an arc with a nonintegral weighting, there must exist another vertex v_k such that $v_i v_k$ is an arc with a nonintegral weighting, α_{ik}, implying $v_i v_k \in E(G_{\overrightarrow{\alpha}})$ so v_i is adjacent to v_k in $G_{\overrightarrow{\alpha}}$. So, any vertex in $V(G_{\overrightarrow{\alpha}})$ is incident to at least two distinct vertices in $V(G_{\overrightarrow{\alpha}})$. Thus, $G_{\overrightarrow{\alpha}}$ consists of two-connected components.

Consider one of these two-connected components. This component necessarily contains some cycle $\mathcal{C}$. Without loss of generality, relabel the vertices so that $\mathcal{C}$ consists of vertices $v_1, v_2, ..., v_t$. Note that by the above argument, consecutive vertices are connected by two edges and each edge v_iv_j in $G_{\overrightarrow{\alpha}}$ corresponds to the arc v_iv_j with weighting α_{ij} in $D_{\overrightarrow{\alpha}}$.

Let γ be the minimum weighting α_{ij} such that v_iv_j is an edge in $\mathcal{C}$. Define a vector $\overrightarrow{\alpha}' \in \mathbb{R}^{\binom{n}{2}}$ componentwise by

$$\alpha'_{ij} = \begin{cases} \alpha_{ij} + \gamma & v_iv_j \text{ is an edge in } \mathcal{C}, \quad i \pmod t < j \pmod t \\ \alpha_{ij} - \gamma & v_iv_j \text{ is an edge in } \mathcal{C}, \quad i \pmod t > j \pmod t \\ \alpha_{ij} & v_iv_j \notin \mathcal{C}. \end{cases}$$

Let $D_{\overrightarrow{\alpha}'}$ be a digraph corresponding to $\overrightarrow{\alpha}'$. Note that by construction, the score of vertex $v_i \in D_{\overrightarrow{\alpha}'}$ is the same as the score of vertex v_i in $D_{\overrightarrow{\alpha}}$ for all $1 \leq i \leq n$, thus $\overrightarrow{\alpha}' \in \text{Frac}_{\overrightarrow{\alpha}}(\overrightarrow{s})$. Also, $\overrightarrow{\alpha}'$ satisfies the same hyperplane conditions as $\overrightarrow{\alpha}$. Thus, $\overrightarrow{\alpha}'$ is also a solution to the matrix equation presented above, $A\overrightarrow{\alpha} = A\overrightarrow{\alpha}' = b$, contradicting our assumption that $\overrightarrow{\alpha}$ is the unique solution to this matrix equation. Thus, $E(G_{\overrightarrow{\alpha}})$ must be empty, implying that all arc weightings in D_α are integral.

Assume that $\overrightarrow{\alpha} \in \{0,1\}^{\binom{n}{2}}$. Each $\alpha_{ij} \in \overrightarrow{\alpha}$ satisfies a constraint either of the form $\alpha_{ij} \leq 1$ or $\alpha_{ij} \geq 0$ with equality. So, $\overrightarrow{\alpha}$ satisfies $\binom{n}{2}$ boundary conditions. These $\binom{n}{2}$ boundary conditions are equations specifying each component of a vector, thus any vector satisfying all $\binom{n}{2}$ boundary conditions has the same components as $\overrightarrow{\alpha}$. Therefore, $\overrightarrow{\alpha}$ is the unique solution to these $\binom{n}{2}$ equations, making $\overrightarrow{\alpha}$ a vertex of $\text{Frac}_{\overrightarrow{\alpha}}(\overrightarrow{s})$ as proposed.

$\square$

4.2 The Polytope $\text{Frac}_{\vec{x}}(n)$

Theorem 10. *Let $\text{Frac}_{\vec{x}}(n)$ be the polytope of vectors $\vec{x}$ such that some fractional tournament realizes $\vec{x}$. Then the vector $\overrightarrow{s}^f$ is a vertex of $\text{Frac}_{\vec{x}}(n)$ if and only if $\overrightarrow{s}^f$ is a permutation of* $(0, 1, 2, ..., n-1)$.

Proof. The vectors $\vec{x}$ in question are those that satisfy Landau's conditions

$$\sum_{i \in S_k} x_i \geq \binom{k}{2}, \qquad 1 \leq k \leq n$$

for every subset S_k of size k of $[n]$, and with equality for $k = n$.

Suppose that $\overrightarrow{s}^f$ is a vertex of $\mathrm{Frac}_{\vec{x}}(n)$. Then $\overrightarrow{s}^f$ is the intersection of n hyperplanes of the form

$$\sum_{i \in S_k} s_i^f = \binom{k}{2}.$$

Let $R_1, R_2, ..., R_n$ be the defining sets of each of these hyperplanes, where the subscript doesn't necessarily correspond to the size of the set. Now suppose that two of the sets are the same size, say $r = |R_j| = |R_l|$ for some $j, l \in [n]$, and let $t = |R_j \cup R_l|$.

Then

$$\sum_{i \in R_j} s_i^f = \sum_{i \in R_l} s_i^f = \binom{r}{2},$$

which implies that

$$\binom{t}{2} \leq \sum_{i \in R_j \cup R_l} s_i^f$$

which can be expanded via the Principle of Inclusion Exclusion to

$$\begin{aligned} &= \sum_{i \in R_j} s_i^f + \sum_{i \in R_l} s_i^f - \sum_{i \in R_j \cap R_l} s_i^f \\ &\leq 2\binom{r}{2} - \binom{|R_j \cap R_l|}{2}. \end{aligned}$$

Note that by the principle of inclusion exclusion, we have $t = |R_j \cup R_l| = |R_j| + |R_l| - |R_j \cap R_l| = r + r - |R_j \cap R_l|$. Thus, $|R_j \cap R_l| = 2r - t$ and we have

$$\begin{aligned} &= 2\binom{r}{2} - \binom{2r-t}{2} \\ &= r(r-1) - \frac{(2r-t)(2r-t-1)}{2}. \end{aligned}$$

$$= -r^2 + 2rt - \frac{t^2 - t}{2}.$$

Completing the square yields,

$$\begin{aligned} &= -(t-r)^2 + \frac{t^2 - t}{2} \\ &= \frac{t(t-1)}{2} - (t-r)^2 \\ &= \binom{t}{2} - (t-r)^2. \end{aligned}$$

We have $\binom{t}{2} \leq \binom{t}{2} - (t-r)^2$ which implies that $(t-r)^2 = 0$. Therefore, $t = |R_j \cup R_l| = |R_j| = |R_l| = r$, so $R_j = R_l$ and $j = l$.

In other words, $R_1, R_2, ..., R_n$ are all distinct sizes, namely $1, ..., n$. Therefore, there exists exactly one i_1 such that $s^f_{i_1} = 0$, and consequently, exactly one i_2 such that $s^f_{i_2} = 1$, and so on inductively. That is, $\vec{s}^f$ is a permutation of $(0, 1, 2, ..., n-1)$ as claimed.

Conversely, any permutation of $(0, 1, 2, ..., n-1)$ will be in one hyperplane of the form

$$\sum_{i \in S_k} s^f_i = \binom{k}{2}$$

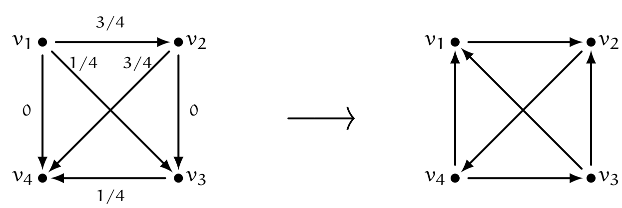
for each $k \in [n]$. Also, as discussed above, this is the unique vector in the intersection of those n hyperplanes, and is thus a vertex of $\mathrm{Frac}_{\vec{x}}(n)$. $\square$

CHAPTER 5

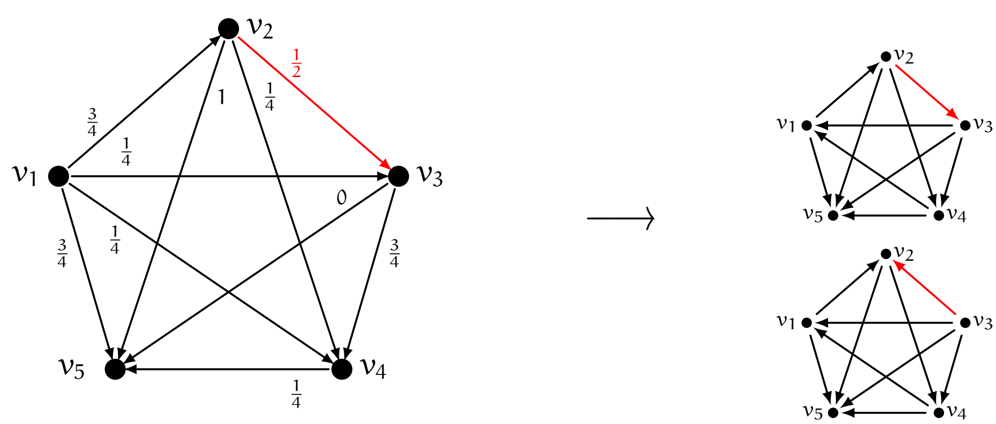
EXPECTED OUTCOME TOURNAMENTS

Let $D = (V, A)$ be a fractional realization of a score sequence $\vec{s}$. Associate D with a tournament $T = (V, A)$ where $V(T) = V(D)$ and an arc a_{ij} is included in $A(T)$ if $\alpha_{ij} = 1 - \alpha_{ji} < \frac{1}{2}$ in D. In the case that some $\alpha_{ij} = \frac{1}{2}$, we associate D with two tournaments, one in which arc a_{ij} is included, and one where arc a_{ji} is included. We refer to T as an *expected outcome tournament* of D.

Example 8. A fractional realization of the score sequence $\vec{s} = (1, 1, 2, 2)$ (left) and its associated expected outcome tournament (right).

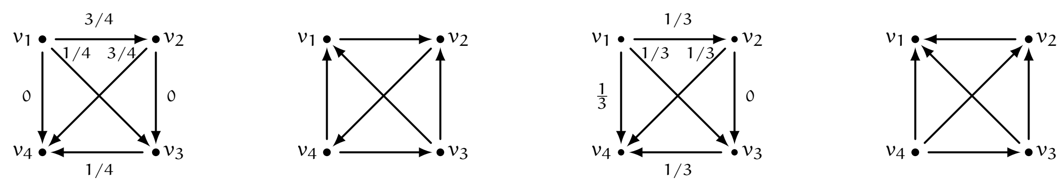


Example 9. A fractional realization of the score vector $\vec{s} = (2, 2, 2, 2, 2)$ and two expected outcome tournaments.



Let D be a fractional realization of a score sequence $\overrightarrow{s}$ and T be an expected outcome tournament of D. The score sequence $\vec{s'}$ of T is called an *effective score sequence* of $\overrightarrow{s}$. We say that $\overrightarrow{s}$ *probabilizes* $\vec{s'}$. A given score sequence $\overrightarrow{s}$ may probabilize many distinct sequences, as in the following example.

Example 10. Two distinct fractional realizations of the score sequence $\overrightarrow{s} = (1, 1, 2, 2)$ and their associated effective outcome tournaments.



We would then say that $\overrightarrow{s} = (1, 1, 2, 2)$ probabilizes the score sequences $\overrightarrow{s}_1 = (1, 1, 2, 2)$ and $\overrightarrow{s}_2 = (0, 1, 2, 3)$. Equivalently, it could be said that $\overrightarrow{s}_1 = (1, 1, 2, 2)$ and $\overrightarrow{s}_2 = (0, 1, 2, 3)$ are effective score sequences of $\overrightarrow{s} = (1, 1, 2, 2)$.

Alternatively, we could say that a score vector $\vec{x}$ probabilizes a score sequence $\overrightarrow{s}$ if there exists some tournament matrix A and some fractional tournament matrix A^f such that

$$A\vec{1} = \overrightarrow{s} \qquad A^f\vec{1} = \vec{x} \qquad A_{ij} = \begin{cases} 1 & A^f_{ij} > \frac{1}{2} \\ 0 & A^f_{ij} < \frac{1}{2} \end{cases}$$

where $\vec{1}$ is the all ones vector.

5.1 The Polytope $\mathrm{Prob}_{\vec{x}}(\vec{s})$

Define the polytope $\mathrm{Prob}_{\vec{x}}(\overrightarrow{s})$ to be the set of all vectors $\vec{x}$ that probabilize a score sequence $\overrightarrow{s}$.

Theorem 11. *For any score sequence $\overrightarrow{s}$, the vertices of $Prob_{\vec{x}}(\overrightarrow{s})$ are located at $\frac{1}{2}(\overrightarrow{s} + \vec{t})$ for each permutation $\vec{t}$ of $(0, 1, 2, ..., n-1)$. Furthermore, $\vec{x} \in Prob_{\vec{x}}(\overrightarrow{s})$ if and only if $2\vec{x} - \overrightarrow{s}$ satisfies Landau's conditions.*

Proof. For a given $n \in \mathbb{N}$, the set of vectors in $\mathrm{Frac}_{\vec{x}}(n)$ can be thought of as the set

$$\mathrm{Frac}_{\vec{x}}(n) = \left\{ D\overrightarrow{\alpha} + \vec{b} \mid \overrightarrow{\alpha} \in [0,1]^{\binom{n}{2}} \right\}$$

where

$$D = \left[\begin{array}{ccc|ccc|cc|c}
1 & \dots & 1 & 0 & \dots & & & & 0 \\
\cline{1-3}
 & & & 1 & \dots & 1 & 0 & \dots & 0 \\
\cline{4-6}
 & & & & & & 1 & \dots & 0 \\
\cline{7-8}
 & -I_{n-1} & & & & & & & \vdots \\
 & & & & -I_{n-2} & & & \dots & 1 \\
 & & & & & & & & -1
\end{array}\right] \quad \text{and} \quad b = \begin{bmatrix} 0 \\ 1 \\ 2 \\ \\ \vdots \\ n-2 \\ n-1 \end{bmatrix}.$$

Similarly, for a given sequence $\overrightarrow{s}$, we claim that

$$\mathrm{Prob}_{\vec{x}}(\overrightarrow{s}) = \left\{ \frac{1}{2} D(\overrightarrow{\alpha} + \vec{\beta}) + \vec{b} \;\middle|\; \overrightarrow{\alpha} \in [0,1]^{\binom{n}{2}}, \quad \beta \in \mathrm{Frac}_{\overrightarrow{\alpha}}(\overrightarrow{s}) \cap \{0,1\}^{\binom{n}{2}} \right\}.$$

It follows that

$$\begin{aligned}
\mathrm{Prob}_{\vec{x}}(\overrightarrow{s}) &= \left\{ \frac{1}{2} D(\overrightarrow{\alpha} + \vec{\beta}) + \vec{b} \mid \overrightarrow{\alpha} \in [0,1]^{\binom{n}{2}}, \quad \beta \in \mathrm{Frac}_{\overrightarrow{\alpha}}(\overrightarrow{s}) \cap \{0,1\}^{\binom{n}{2}} \right\} \\
&= \left\{ \frac{1}{2} \left(D\overrightarrow{\alpha} + D\vec{\beta} + 2\vec{b} \right) \right\}.
\end{aligned}$$

Since $\vec{\beta} \in \mathrm{Frac}_{\overrightarrow{\alpha}}(\overrightarrow{s})$, $D\overrightarrow{\alpha} + \vec{b} = \overrightarrow{s}$. So we have,

$$\begin{aligned}
&= \left\{ \frac{1}{2} \left(\overrightarrow{s} + D\overrightarrow{\alpha} + \vec{b} \right) \right\} \\
&= \left\{ \frac{1}{2} \left(\overrightarrow{s} + \vec{r} \right) \mid \vec{r} \in \mathrm{Frac}_{\vec{x}}(n) \right\}.
\end{aligned}$$

By Theorem 10, the vertices of $\mathrm{Frac}_{\vec{x}}(n)$ are permutations of the transitive sequence $(0, 1, 2, .., n-1)$. Thus, the vertex set of $\mathrm{Prob}_{\vec{x}}(\overrightarrow{s})$ is given by

$$\left\{ \frac{1}{2}\left(\overrightarrow{s} + \vec{t}\right) \;\middle|\; \vec{t} \text{ is a permutation of } (0, 1, 2, ..., n-1) \right\}.$$

Suppose $\vec{x} \in \mathrm{Prob}_{\vec{x}}(\overrightarrow{s})$ for some $\overrightarrow{s}$ with length n. Then $\vec{x} = \frac{1}{2}(\overrightarrow{s} + \vec{r})$ for some $r \in \mathrm{Frac}_{\vec{x}}(n)$ and it follows that $\vec{r} = 2\vec{x} - \overrightarrow{s} \in \mathrm{Frac}_{\vec{x}}(n)$. Thus, $2\vec{x} - \overrightarrow{s}$ satisfies Landau's conditions as claimed.

Conversely, if $\vec{r} = 2\vec{x} - \overrightarrow{s} \in \mathrm{Frac}_{\vec{x}}(n)$. Then $\vec{x} = \frac{1}{2}(\vec{r} + \overrightarrow{s}) \in \mathrm{Prob}_{\vec{x}}(\overrightarrow{s})$. $\square$

CHAPTER 6

FUTURE DIRECTIONS

The work presented here lends itself to a myriad of research directions. We have generalized the work of Barrus in the context of complete directed graphs (tournaments); we believe that the work may be generalized further to all directed graphs. In the study of fractional realizations of tournament score sequences, our work focused on investigating the feasible region of the linear program presented. While a characterization of the vertices of such a polytope is included here, there are other properties of the feasible region that may be of interest. For instance, given a vertex of a polytope $\text{Frac}_{\overrightarrow{\alpha}}(\overrightarrow{s})$ for some score sequence $\overrightarrow{s}$, is there a way to measure which other vertices are 'closest' to the given vertex and is there a systematic way to traverse the edges of the polytope to reach another vertex? Along the same lines, is there a meaningful way to partition this polytope to identify vertices with certain graph structures? This idea came about after pondering about the $1/2$ cases in our arc weightings, which seemed to represent some sort of tipping or critical point.

Theorem 8 proves that for a sequence $\vec{s}$ of the form $(0, 1, 2, ..., n-1)$, the polytope $\text{Frac}_{\overrightarrow{\alpha}}(\overrightarrow{s})$ is a single point, meaning there is a unique fractional realization of the score sequence. Given a score sequence $\vec{s}$, can one determine the dimension of the polytope $\text{Frac}_{\overrightarrow{\alpha}}(\overrightarrow{s})$ and interpret this in a meaningful way?

Theorem 9 proves that if an objective function attains an optimal value over the polytope $\text{Frac}_{\overrightarrow{\alpha}}(\overrightarrow{s})$, it will be attained at one of the vertices which correspond to tournaments with all integral weightings. It may be the case that the optimization of certain objective functions over this polytope may provide an interesting way to rank players in a tournament or gain information about tournament structures.

The notation and vocabulary of expected outcome tournaments, effective score sequence, and probabilizations serves to facilitate much further research. In particular, in the study of expected outcome tournaments, among multiple fractional tournaments with the same expected outcome tournament is one of the fractional tournaments 'more likely' to yield the expected outcome tournament? Is there a way to associate some sort of 'confidence score' with expected outcome tournaments based on the arc weightings of the fractional tournament? During our research, we noted that a score sequence of the form $(0, 1, 2, ..., n-1)$ has the property that its only effective score sequence is itself, which begs the question: Are there other sequences $\overrightarrow{s}$ such that the set $\{\vec{x} \mid \vec{x} \text{ is an effective score sequence of } \overrightarrow{s}\}$ only contains the vector $\vec{s}$?

REFERENCES

[1] K. Appel and W. Haken. The Solution of the Four-Color-Map Problem. *Scientific American*, 237:108–121, October 1977.

[2] Michael D. Barrus. On fractional realizations of graph degree sequences. *Electron. J. Combin.*, 21(2):Paper 2.18, 20, 2014.

[3] Stephen Boyd and Lieven Vandenberghe. *Convex Optimization.* Cambridge University Press, 2004.

[4] P. Erdős. Graphs with prescribed degree of vertices. *Mat. Lapok.*, 11:264–274, 1960.

[5] S Louis Hakimi. On realizability of a set of integers as degrees of the vertices of a linear graph. i. *Journal of the Society for Industrial and Applied Mathematics*, 10(3):496–506, 1962.

[6] Václav Havel. A remark on the existence of finite graphs. *Casopis Pest. Mat.*, 80:477–480, 1955.

[7] P.J. Heawood. Map-Colour Theorems. *Quaterly Journal of Mathematics*, 24:332–338, 1890.

[8] H. G. Landau. On dominance relations and the structure of animal societies: Iii the condition for a score structure. *The bulletin of mathematical biophysics*, 15(2):143–148, Jun 1953.

[9] J. W. Moon and N. J. Pullman. On generalized tournament matrices. *SIAM Rev.*, 12:384–399, 1970.

[10] John W. Moon. *Topics on Tournaments in Graph Theory.* Dover Publications, Inc, 2015.

[11] Christos H. Papadimitriou and Kenneth Steiglitz. *Combinatorial Optimization: Algorithms and Complexity.* Dover Publications, Inc, 1998.

[12] Bryan L Shader. On tournament matrices. *Linear algebra and its applications*, 162:335–368, 1992.

[13] Richard J. Trudeau. *Introduction to Graph Theory.* Dover Publications, Inc, 1993.