Three notions of tropical rank for symmetric matrices

Dustin Cartwright, Melody Chan

To cite this version:

HAL Id: hal-01186294
https://inria.hal.science/hal-01186294

Submitted on 24 Aug 2015

Three notions of tropical rank for symmetric matrices

Dustin Cartwright [1] and Melody Chan[1†]

[1]*Dept. of Mathematics, University of California, Berkeley, CA 94720, USA*

Abstract. We introduce and study three different notions of tropical rank for symmetric matrices and dissimilarity matrices in terms of minimal decompositions into rank 1 symmetric matrices, star tree matrices, and tree matrices. Our results provide a close study of the tropical secant sets of certain nice tropical varieties, including the tropical Grassmannian. In particular, we determine the dimension of each secant set, the convex hull of the variety, and in most cases, the smallest secant set which is equal to the convex hull.

Résumé. Nous introduisons et étudions trois notions différentes de rang tropical pour des matrices symétriques et des matrices de dissimilarité, en utilisant des décompositions minimales en matrices symétriques de rang 1, en matrices d'arbres étoiles, et en matrices d'arbres. Nos résultats donnent lieu à une étude détaillée des ensembles des sécantes tropicales de certaines jolies variétés tropicales, y compris la grassmannienne tropicale. En particulier, nous déterminons la dimension de chaque ensemble des sécantes, l'enveloppe convexe de la variété, ainsi que, dans la plupart des cas, le plus petit ensemble des sécantes qui est égal à l'enveloppe convexe.

Resumen. Introducimos y estudiamos tres nociones diferentes de rango tropical para matrices simétricas y matrices de disimilaridad, utilizando las decomposiciones minimales en matrices simétricas de rango 1 en matrices de árboles estrella y en matrices de árboles. Nuestros resultados brindan un estudio detallado de conjuntos de secantes de ciertas variedades tropicales clásicas, incluyendo la grassmanniana tropical. En particular, determinamos la dimensión de cada conjunto de dichas secantes, la cápsula convexa de la variedad, y también, en la mayoridad de los casos, el conjunto más pequeño de secantes que coincide con la cápsula convexa.

Keywords: tropical geometry, tropical convexity, secant varieties, rank, symmetric matrices, hypergraph coloring

1   Introduction

In this paper, we study tropical secant sets and rank for symmetric matrices. Our setting is the *tropical semiring* $(\mathbb{R} \cup \{\infty\}, \oplus, \odot)$, where tropical addition is given by $x \oplus y = \min(x, y)$ and tropical multiplication is given by $x \odot y = x + y$. The kth *tropical secant set* of a subset V of $\mathbb{R}^N$ is defined to be the set of points

$$\{x \in \mathbb{R}^N : x = v_1 \oplus \cdots \oplus v_k, \ v_i \in V\},$$

where $\oplus$ denotes coordinate-wise minimum. This set is typically not a tropical variety and thus we prefer the term "secant set" to "secant variety," which has appeared previously in the literature. The *rank* of a

point $x \in \mathbb{R}^N$ with respect to V is the smallest integer k such that x lies in the kth tropical secant set of V, or ∞ if there is no such k.

In [3], Develin, Santos, and Sturmfels define the Barvinok rank of a matrix, not necessarily symmetric, to be the rank with respect to the subset of $n \times n$ rank 1 matrices, and their definition serves as a model for ours. In addition, they define two other notions of rank, Kapranov rank and tropical rank, for which there are no analogues in this paper. Further examination of min-plus ranks of matrices can be found in the review article [1].

We give a careful examination of secant sets and rank with respect to three families of tropical varieties in the space of symmetric matrices and the space of dissimilarity matrices. By a $n \times n$ *dissimilarity matrix* we simply mean a function from $\binom{[n]}{2}$ to $\mathbb{R}$, which we will write as a symmetric matrix with the placeholder symbol $*$ along the diagonal. There is a natural projection from $n \times n$ symmetric matrices to $n \times n$ dissimilarity matrices which we denote by π. For example,

$$M = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad \text{and} \quad \pi(M) = \begin{bmatrix} * & 1 & 0 & 0 \\ 1 & * & 0 & 0 \\ 0 & 0 & * & 1 \\ 0 & 0 & 1 & * \end{bmatrix} \tag{1}$$

are a symmetric matrix and dissimilarity matrix respectively.

Our first family is the *tropical Veronese* of degree 2, which is the tropicalization of the classical space of symmetric matrices of rank 1. It is a classical linear subspace of the space of symmetric matrices consisting of those matrices which can be written as $v^T \odot v$ for some row vector v. The rank of a matrix with respect to the tropical Veronese is called *symmetric Barvinok rank*, because it is the symmetric analogue of Barvinok rank.

Second, we consider the *space of star trees*, which is the image of the tropical Veronese under the projection π. Equivalently, it can be obtained by first projecting the classical Veronese onto its off-diagonal entries and then tropicalizing. The classical variety and its secant varieties were studied in a statistical context in [5]. The tropical variety is a classical linear subspace of the space of dissimilarity matrices, and we call the corresponding notion of rank *star tree rank*. The name reflects the fact that the matrices with star tree rank 1 are precisely those points of the tropical Grassmannian which correspond to trees with no internal edges, i.e. star trees, in in the identification below.

Third, we consider the *tropical Grassmannian* $G_{2,n}$, which is the tropicalization of the Grassmannian of 2-dimensional subspaces in an n-dimensional vector space, and was first studied in [9]. It consists of exactly those dissimilarity matrices arising as the distance matrix of a weighted tree with n leaves in which internal edges have negative weights. Therefore, we call the points in the tropical Grassmannian *tree matrices*, and call rank with respect to the tropical Grassmannian the *tree rank*. Note that our definition of tree rank differs from that in [7, Ch. 3], which uses a different notion of mixtures.

We use our examples of M and $\pi(M)$ from (1) to illustrate our three notions of rank. Proposition 4 tells us that the symmetric Barvinok rank of M is 4. Theorem 8 tells us that the star tree rank of $\pi(M)$ is 2. Explicitly, we have

$$\pi(M) = \begin{bmatrix} * & 1 & 0 & 0 \\ 1 & * & 2 & 2 \\ 0 & 2 & * & 1 \\ 0 & 2 & 1 & * \end{bmatrix} \oplus \begin{bmatrix} * & 1 & 2 & 2 \\ 1 & * & 0 & 0 \\ 2 & 0 & * & 1 \\ 2 & 0 & 1 & * \end{bmatrix}.$$

Finally, the tree rank of $\pi(M)$ is 1 by Proposition 13. This example shows that all three of our notions of rank can be different.

However, for any $n \times n$ symmetric matrix M, we have

$$\text{symmetric Barvinok rank}(M) \geq \text{star tree rank}(\pi(M)) \geq \text{tree rank}(\pi(M)). \tag{2}$$

The first inequality follows from the fact that the set of dissimilarity matrices of star tree rank 1 is the projection of the set of matrices of symmetric Barvinok rank 1. The second inequality follows from the fact that the space of star trees is contained in the tropical Grassmannian.

The rest of the paper is organized as follows. In Section 2, we present a technique for proving lower bounds on rank. We introduce a hypergraph whose chromatic number is a lower bound on rank. We examine symmetric Barvinok rank, star tree rank, and tree rank in Sections 3, 4, and 5 respectively. We prove upper bounds on the rank in each case, and with the exception of tree rank, our upper bounds are sharp. We show that the symmetric Barvinok rank of an $n \times n$ symmetric matrix can be infinite, but even when the rank is finite it can exceed n, and in fact can grow quadratically in n (Theorem 5). For each notion of rank, the set of matrices with rank at most k is a union of polyhedral cones, and we compute the dimension of these sets, defined as the dimension of the largest cone. In each case, the dimension of the tropical secant set equals the dimension of the clasical secant variety, confirming Draisma's observation that tropical geometry provides useful lower bounds for the dimensions of classical secant varieties [4]. We also give a combinatorial characterization of each notion of rank for a $0/1$ matrix in terms of graph covers. Finally, in Section 6, we explicitly characterize the stratification of the 5×5 dissimilarity matrices by star tree rank and tree rank respectively, and show that the lower bounds from the chromatic number in Section 2 are exact in these cases.

2 Lower bounds on rank via hypergraph coloring

We begin by giving a general combinatorial construction: a hypergraph whose chromatic number yields a lower bound on rank.

Recall that a *hypergraph* consists of a ground set, called vertices, and a set of subsets of the ground set, called hyperedges. The *chromatic number* of a hypergraph H, denoted $\chi(H)$, is the smallest number r such that the vertices of H can be partitioned into r color classes with no hyperedge of H monochromatic. In particular, if H contains a hyperedge of size 1, then $\chi(H)$ is ∞.

Now, suppose we have a tropical prevariety $V \subseteq \mathbb{R}^N$. Recall that a tropical polynomial

$$p(x_1, \ldots, x_N) = \bigoplus_{i=1}^{t} a_i \odot x_1^{c_{i1}} \odot \cdots \odot x_N^{c_{iN}} \tag{3}$$

defines a tropical hypersurface consisting of those vectors $x \in \mathbb{R}^N$ such that the minimum in evaluating $p(x)$ is achieved at least twice. A *tropical prevariety* is the intersection of finitely many tropical hypersurfaces, and we call the set of tropical polynomials defining the prevariety V a *defining set*.

Now, given a point $w \in \mathbb{R}^N$ and a defining set S for V, we construct a hypergraph on ground set $[N]$ as follows. Let p from (3) be a tropical polynomial in S, with all exponents $c_{ij} \geq 0$. If the minimum is achieved uniquely when p is evaluated at w, then we add a hyperedge E whose elements correspond to the coordinates that appear with non-zero exponent in the unique minimal term. The *deficiency hypergraph*

of w with respect to V and S consists of hyperedges coming from all polynomials in S with a unique minimum at w. In particular, the deficiency hypergraph has no hyperedges (and thus has chromatic number 1) if and only if w is in V.

Proposition 1 *If H is the deficiency hypergraph constructed above, then the rank of $w \in \mathbb{R}^N$ with respect to $V \subseteq \mathbb{R}^N$ is at least $\chi(H)$.*

Corollary 2 *If the deficiency hypergraph H has a hyperedge of size 1, then the rank of w with respect to V is infinite.*

The lower bound in Proposition 1 may be strict, such as with $S = \{xz \oplus y^2, xw \oplus yz\}$ and $w = (0, 0, 0, 1)$. The rank of w with respect to the variety defined by S is infinite, but the deficiency hypergraph is 2-colorable.

For the varieties considered in this paper, we will take quadratic tropical bases as our defining tropical polynomials, and thus the deficiency hypergraph will always be a graph (possibly with loops). Accordingly, we will call it the *deficiency graph*.

3   Symmetric Barvinok rank

The symmetric Barvinok rank of a symmetric matrix M is the smallest number r such that M can be written as the sum of r rank 1 symmetric matrices. The 2×2 minors $x_{ij}x_{kl} \oplus x_{il}x_{kj}$ of M for $i \neq k$ and $l \neq j$ form a tropical basis for the variety of rank 1 symmetric matrices. We will always construct our deficiency graph with respect to this tropical basis.

Our first observation is that the symmetric Barvinok rank of a matrix can be infinite. More precisely,

Proposition 3 *If M is a symmetric matrix and $2M_{ij} < M_{ii} + M_{jj}$ for some i and j, then the symmetric Barvinok rank of M is infinite.*

Proof: The tropical polynomial $x_{ij}^2 \oplus x_{ii}x_{jj}$ is in the tropical basis, so if $2M_{ij} < M_{ii} + M_{jj}$ for some i and j, then the deficiency graph for M has a loop at the node ij. Therefore, M has infinite rank by Corollary 2. $\square$

In fact, the converse to Proposition 3 is also true; see Theorem 5.

Next, we give a graph-theoretic characterization of the symmetric Barvinok rank of $0/1$-matrices. We define a *clique cover* of a simple graph G to be a collection of r complete subgraphs such that every edge and every vertex of G is in some element of the collection. Given an $n \times n$ symmetric $0/1$ matrix M with zeroes on the diagonal, define G_M to be the graph whose vertices are the integers $[n]$, and which has an edge between i and j if and only if $M_{ij} = 0$.

Proposition 4 *Suppose M is a symmetric $0/1$ matrix with zeroes on the diagonal. Then the symmetric Barvinok rank of M is the size of a smallest clique cover of G_M.*

On the other hand, suppose that M is a symmetric $0/1$ matrix with at least one entry of 1 on the diagonal. If there exist i and j such that $M_{ii} = 1$ and $M_{ij} = 0$, then the symmetric Barvinok rank of M is infinite. Otherwise, let M' be the maximal principal submatrix with zeroes on the diagonal. The symmetric Barvinok rank of M is one greater than the symmetric Barvionk rank of M'.

This characterization gives us two families of matrices which have rank n and $\lfloor n^2/4 \rfloor$ respectively, namely those corresponding to the trivial graph with n isolated vertices and the complete bipartite graph

$K_{\lfloor n/2 \rfloor, \lceil n/2 \rceil}$. In the latter case, $K_{\lfloor n/2 \rfloor, \lceil n/2 \rceil}$ is triangle-free, so no clique can consist of more than one edge. On the other hand, there are $\lfloor n^2/4 \rfloor$ edges in the graph, so $\lfloor n^2/4 \rfloor$ cliques are needed in a cover. In fact, these two examples have the maximum possible rank for $n \times n$ matrices.

Theorem 5 *Suppose that M is a symmetric $n \times n$ matrix with $M_{ii} + M_{jj} \leq 2M_{ij}$ for all i and j. Then the symmetric Barvinok rank of M is at most $\max\{n, \lfloor n^2/4 \rfloor\}$, and this bound is tight. Thus, every matrix with finite rank has rank at most $\max\{n, \lfloor n^2/4 \rfloor\}$.*

The next theorem shows that the dimensions of the tropical secant sets and their classical secant varieties agree.

Theorem 6 *The dimension of the space of symmetric $n \times n$ matrices of symmetric Barvinok rank at most r is $\binom{n+1}{2} - \binom{n-r+1}{2}$, which is the dimension of the classical secant variety, i.e. the space of classical symmetric matrices of classical rank at most r.*

Proof: Let $D = \binom{n+1}{2} - \binom{n-r+1}{2}$. The tropical secant set is contained in the tropicalization of the classical secant variety, so the dimension is at most D, by the Bieri-Groves Theorem [2, Thm. A]. Thus, it is sufficient to find an open neighborhood in which the tropical secant set has dimension D. For i from 1 to r, let $v_i = (C, \ldots, C, v_{i,i}, \ldots, v_{i,n})$ be a vector with C for the first $i-1$ entries. Choose the coordinates $v_{i+1,j}$ to be smaller than all the $v_{i,j}$ and let C be very large. Then,

$$v_1^T \odot v_1 \oplus \cdots \oplus v_r^T \odot v_r = \begin{bmatrix} 2v_{11} & v_{11} + v_{12} & \cdots & v_{11} + v_{1n} \\ v_{11} + v_{12} & 2v_{22} & \cdots & v_{22} + v_{2n} \\ \vdots & \vdots & & \vdots \\ v_{11} + v_{1n} & v_{22} + v_{2n} & \cdots & 2v_{rn} \end{bmatrix}$$

This matrix is an injective function of the vector entries v_{ij} for $i \leq r$ and $j \geq i$. Thus, it defines a neighborhood of the rth secant set of the desired dimension

$$n + (n-1) + \cdots + (n-r+1) = \binom{n+1}{2} - \binom{n-r+1}{2} = D.$$

$\square$

In the case $n = 3$, we can explicitly describe the stratification of symmetric matrices by symmetric Barvinok rank. By Theorem 5, the symmetric Barvinok rank of a 3×3 matrix is at most 3, and the locus of rank 1 matrices is the tropical variety defined by the 2×2 minors, so it suffices to characterize the matrices of rank at most 2. Following [3], we call a square matrix *tropically singular* if it lies in the tropical variety of the determinant.

Proposition 7 *Let M be a symmetric 3×3 matrix. Then the following are equivalent:*

1. *M has symmetric Barvinok rank at most 2;*

2. *The deficiency graph of M is 2-colorable;*

3. *M is tropically singular and $M_{ii} + M_{jj} \leq 2M_{ij}$ for all $1 \leq i, j \leq 3$.*

We remark that for larger matrices, the symmetric Barvinok rank does not have as simple a characterization as the third condition in Proposition 7. A necessary condition for a symmetric $n \times n$ matrix to have rank at most r is that $M_{ii} + M_{jj} \leq 2M_{ij}$ and all the $(r+1) \times (r+1)$ submatrices are tropically singular, but one can show that this condition is not sufficient for $n \geq 4$.

4 Star tree rank

A star tree matrix is a dissimilarity matrix which can be written as $\pi(v^T \odot v)$ for $v \in \mathbb{R}^n$ a row vector. The star tree matrices form a classical linear space in the space of $n \times n$ dissimilarity matrices defined by the tropical polynomials

$$x_{ij}x_{kl} \oplus x_{ik}x_{jl} \qquad \text{for } i, j, k, \text{ and } l \text{ distinct integers.} \tag{4}$$

In this section, the deficiency graph will always be taken with respect to this tropical basis.

Unlike the case of symmetric Barvinok rank, the star tree rank is always finite.

Theorem 8 *For n at least 3, the star tree rank of a $n \times n$ dissimilarity matrix M is at most $n - 2$, and this bound is sharp. In particular, the dissimilarity matrix defined by $M_{ij} = \min\{i, j\}$ has star tree rank $n - 2$.*

Note that the matrix with maximal star tree rank in Theorem 8 is in fact in the Grassmannian, i.e. it has tree rank 1. Indeed, one may check that the four-point condition (5) holds.

We can also give a graph-theoretic characterization of the star tree rank of $0/1$-matrices. For M a $0/1$ dissimilarity matrix, we define G_M to be the graph whose edges correspond to the zeroes of M. As in the case of symmetric Barvinok rank, we can characterize the star tree rank of M in terms of covers of G_M, this time by both cliques and star trees. We will also say that a cover of G_M by cliques and star trees is a *solid cover* if for every pair of distinct vertices i and j either:

1. there is an edge between i and j,

2. either i or j belongs to a clique in the cover,

3. either i or j is the center of a star tree in the cover, or

4. both i and j are leaves of the same star tree.

Proposition 9 *Let M be a $0/1$ dissimilarity matrix. Let r be the minimal number of graphs in a cover of G_M by cliques and star trees, such that every edge (but not necessarily every vertex) is in some element of the cover. Then M has star tree rank either r or $r + 1$.*

Moreover, if G_M has a solid cover by r graphs, then M has star tree rank r.

In contrast to symmetric Barvinok rank, the upper bound of $n - 2$ on the star tree rank of an $n \times n$ dissimilarity matrix cannot be achieved by a $0/1$ matrix for large n. Recall that the *Ramsey number* $R(k, k)$ is the smallest integer such that any graph on at least $R(k, k)$ vertices has either a clique or a independent set of size k. Then we have the following stronger bound on the star tree rank of a $0/1$ matrix.

Proposition 10 *For $n \geq R(k, k)$, any $n \times n$ $0/1$ dissimilarity matrix has star tree rank at most $n - k + 1$.*

Proof: By the assumption on n, the graph G_M has either a clique of size k or an independent set of size k. In the former case, we can cover G_M by a star tree centered at each vertex not part of the clique, together with the clique itself. This gives a solid cover by $n - k + 1$ subgraphs. In the latter case, we can just take the star trees centered at the vertices not in the independent set, giving a cover of G_M by $n - k$ subgraphs. In either case, Proposition 9 shows that M has rank at most $n - k + 1$. □

Corollary 11 *For $n \geq 18$, every $n \times n$ 0/1 dissimilarity matrix has star tree rank at most $n - 3$.*

Proof: The Ramsey number $R(4, 4)$ is 18 [8]. □

In [5, Theorem 2], Drton, Sturmfels, and Sullivant prove a dimension theorem for the secant varieties of the classical Veronese projected to off-diagonal entries. Here, we prove the tropical analogue of their result.

Theorem 12 *Let r and n be positive integers. Then the dimension of the space of dissimilarity $n \times n$ matrices of star tree rank at most r is*

$$\min \left\{ \binom{n+1}{2} - \binom{n-r+1}{2}, \binom{n}{2} \right\}.$$

In fact, the difficult part of Theorem 2 in [5] is proving the lower bound on the dimension of the classical secant variety. Our computation of the dimension of the tropical secant set, combined with the Bieri-Groves Theorem [2, Theorem A], provides an alternative proof of this lower bound.

5 Tree rank

The tropical Grassmannian $G_{2,n}$ is the tropical variety defined by the 3-term Plücker relations:

$$x_{ij}x_{k\ell} \oplus x_{ik}x_{j\ell} \oplus x_{i\ell}x_{jk} \qquad \text{for all } i < j < k < \ell. \tag{5}$$

This condition is equivalent to coming from the distances along a weighted tree which has negative weights along the internal edges [9, Sec. 4]. In this section, we will always take the deficiency graph to be with respect to the Plücker relations in (5).

As with the previous notions of rank, the tree rank of a 0/1 matrix can be characterized in terms of covers of graphs. For any disjoint subsets $I_1, \ldots, I_k \subset [n]$ (not necessarily a partition), the *complete k-partite graph* is the graph which has an edge between the elements of I_i and I_j for all $i \neq j$. Complete k-partite graphs are characterized by the property that among vertices which are incident to some edge, the relation of having a non-edge is a transitive relation.

Note that the complete k-partite graphs defined above are exactly those graphs whose edge set forms the set of bases of a rank 2 matroid on n elements. The transitivity of being a non-edge is equivalent to the basis exchange axiom. Alternatively, each of the sets $I_1, \ldots, I_k$ partition the set of non-loops in the matroid into parallel classes. See [6] for definitions of these terms. In the following proposition, we will see that the Plücker relations imply the basis exchange axiom for the 0 entries of a non-negative tree matrix.

Proposition 13 *Let M be an $n \times n$ 0/1 dissimilarity matrix and let r be smallest size of a cover of G_M by complete k-partite subgraphs. As in Proposition 9, we only require every edge to be in the cover, not necessarily every vertex. If G_M has at most one isolated vertex then M has tree rank r. Otherwise, M has tree rank $r + 1$.*

Note that by taking the I_i in the definition of k-partite graph to be singletons, we get complete graphs, and by taking $k = 2$ with I_1 a singleton and I_2 any set disjoint from I_1, we get star trees. Together with Propositions 9 and 13, this confirms, for 0/1-matrices, the second inequality in (2).

Again, we can show that the tropical secant sets and classical secant varieties agree in dimension:

n	maximum tree rank	example
3	1	
4	2	
5	3	0/1 matrix corresponding to 5-cycle
6	3	
7	4	
8	5	
9	6	M in (6)
10	6 or 7	Any extension of M in (6)
$9k$	between $6k$ and $9k - 3$	M_k from discussion following (6)

Tab. 1: Maximum possible tree rank of an $n \times n$ dissimilarity matrix, to the best of our knowledge. The upper bounds come from Theorems 8 and 15. The examples have the largest tree ranks that are known to us. The omitted examples can be provided by taking a principal submatrix of a larger example, by Lemma 16.

Proposition 14 *The dimension of the set of dissimilarity $n \times n$ matrices of tree rank at most r is the dimension of the classical secant variety,*

$$\binom{n}{2} - \binom{n - 2r}{2} \quad \textit{if } r \leq \frac{n}{2},$$
$$\binom{n}{2} \quad \textit{if } r \geq \frac{n-1}{2}.$$

Unlike the cases of symmetric Barvinok rank and star tree rank, we do not know the maximum tree rank of a $n \times n$ dissimilarity matrix for large n. We have an upper bound of $n - 2$ by Theorem 8, and we can improve on this slightly:

Theorem 15 *For $n \geq 6$, an $n \times n$ dissimilarity matrix M has tree rank at most $n - 3$.*

Beginning with $n = 10$, we don't know whether or not the bound in Theorem 15 is sharp. For the following 9×9 matrix, found by random search, the deficiency graph was computed to have chromatic number 6:

$$M = \begin{bmatrix} * & 1 & 6 & 7 & 2 & 3 & 8 & 9 & 6 \\ 1 & * & 2 & 7 & 9 & 7 & 5 & 7 & 1 \\ 6 & 2 & * & 6 & 0 & 6 & 1 & 7 & 1 \\ 7 & 7 & 6 & * & 3 & 3 & 8 & 5 & 3 \\ 2 & 9 & 0 & 3 & * & 5 & 7 & 5 & 7 \\ 3 & 7 & 6 & 3 & 5 & * & 9 & 3 & 9 \\ 8 & 5 & 1 & 8 & 7 & 9 & * & 2 & 3 \\ 9 & 7 & 7 & 5 & 5 & 3 & 2 & * & 8 \\ 6 & 1 & 1 & 3 & 7 & 9 & 3 & 8 & * \end{bmatrix} \quad (6)$$

Together with Theorem 15, this computation shows that M has tree rank 6. For any $k \geq 1$, we can form an $9k \times 9k$ matrix M_k by putting M in blocks along the diagonal and setting all other entries to 10. The deficiency graph of M_k includes k copies of the deficiency graph of M, and all edges between distinct copies. Therefore, the chromatic number, and thus the tree rank, are at least $6k$.

On the other hand, in order to provide examples of an $n \times n$ matrix with tree rank $n - 3$ for all $n \leq 9$, we have the following lemma.

Lemma 16 *Let M be an $n \times n$ matrix. If any $(n - m) \times (n - m)$ principal submatrix has tree rank r, then M has tree rank at most $r + m$.*

Proof: Fix a decomposition of the $(n - m) \times (n - m)$ principal submatrix into r tree matrices. We can extend each tree matrix to an $n \times n$ tree matrix by adding leaf edges with large positive weights. For each index i not in the principal submatrix, define v_i to be the vector which is $C + M_{ij}$ in the jth entry and $-C$ in the ith entry, where C is a large real number. Then, the extended tree matrices, together with $\pi(v_i^T \odot v_i)$ for all i not in the principal submatrix, give a decomposition of M into $r + m$ tree matrices, as desired. □

These results on the maximum tree rank are summarized in Table 1.

6 Star tree rank and tree rank for $n = 5$

In this section, we characterize the secant sets of the space of star trees and of the tropical Grassmannian in the case $n = 5$. Both give examples where the lower bound of Proposition 1 is actually an equality.

6.1 Star tree rank for $n = 5$

From Theorem 8, we know that the maximum star tree rank of a 5×5 matrix is 3. On the other hand, the set of dissimilarity matrices of star tree rank 1 is defined by the 2×2 minors. Thus, our task is to describe the second secant set of the space of star trees, i.e. the set of dissimilarity matrices of star tree rank 2.

First, we recall the defining ideal of the classical secant variety. The space of star trees is the tropicalization of the projection of the rank 1 symmetric matrices onto their off-diagonal entries. Its second secant variety is a hypersurface in $\mathbb{C}^{10}$ defined by the following 12-term quintic, known as the pentad [5]:

$$x_{12}x_{13}x_{24}x_{35}x_{45} - x_{12}x_{13}x_{25}x_{34}x_{45} - x_{12}x_{14}x_{23}x_{35}x_{45} + x_{12}x_{14}x_{25}x_{34}x_{35}$$
$$+ x_{12}x_{15}x_{23}x_{34}x_{45} - x_{12}x_{15}x_{24}x_{34}x_{35} + x_{13}x_{14}x_{23}x_{25}x_{45} - x_{13}x_{14}x_{24}x_{25}x_{35}$$
$$- x_{13}x_{15}x_{23}x_{24}x_{45} + x_{13}x_{15}x_{24}x_{25}x_{34} - x_{14}x_{15}x_{23}x_{25}x_{34} + x_{14}x_{15}x_{23}x_{24}x_{35}$$

Note that the 12 terms of the pentad correspond to the 12 different cycles on 5 vertices. The second secant set of the space of star trees is contained in the tropicalization of the pentad, but the containment is proper. Nonetheless, the terms of the pentad play a fundamental role in characterizing matrices of rank at most 2.

Theorem 17 *Let M be a 5×5 dissimilarity matrix. The following are equivalent:*

1. *M has star tree rank at most 2;*

2. *The deficiency graph of M is 2-colorable;*

3. *The minimum of the terms of the pentad is achieved at two terms which satisfy the following:*

   (a) *The terms differ by a transposition;*

   (b) *Assuming, without loss of generality, that the minimized terms are $x_{12}x_{23}x_{34}x_{45}x_{15}$ and $x_{13}x_{23}x_{24}x_{45}x_{15}$, then we have that $M_{14} + M_{23} \leq M_{12} + M_{34} = M_{13} + M_{24}$.*

The proof is similar in spirit to the proof of Theorem 18 below, and we omit it.

6.2 Tree rank for $n = 5$

We now turn our attention to tree rank of 5×5 dissimilarity matrices. As in the case of star tree rank, the maximum tree rank is 3, and so it suffices to characterize 5×5 dissimilarity matrices of tree rank at most 2. Unlike the previous case, the second classical secant variety is already all of $\mathbb{C}^{10}$, so there is no classical polynomial whose tropicalization gives us a clue to the tropical secant set. However, the tropical pentad again shows up in our characterization.

First, here is a simple example of a 5×5 0/1 dissimilarity matrix with tree rank 3. Consider the 0/1 matrix corresponding to the 5-cycle C_5. Now, C_5 cannot be covered by fewer than 3 k-partite graphs, and so the matrix has tree rank at least 3 by Proposition 13. On the other hand, it has tree rank at most 3 by Theorem 8 and the inequality in (2). We will see in Remark 1 that this matrix is, in a certain sense, the only such example.

Let P be the tropical polynomial in variables $\{x_{ij} : 1 \leq i < j \leq 5\}$ which is the tropical sum of the 22 tropical monomials of degree 5 in which each $i \in \{1, \ldots, 5\}$ appears in a subscript exactly twice. Thus P has 12 monomials of the form $x_{12}x_{23}x_{34}x_{45}x_{15}$, forming the terms of the pentad, and 10 new monomials of the form $x_{12}x_{23}x_{31}x_{45}^2$. Let us call terms of the former kind *pentagons*, and terms of the latter kind *triangles*.

Theorem 18 *Let M be a 5×5 dissimilarity matrix. Then the following are equivalent:*

1. *M has tree rank at most 2;*

2. *The deficiency graph is 2-colorable;*

3. *The tropical polynomial P achieves its minimum at a triangle.*

Proof: First, (1) implies (2) by Proposition 1.

For (2) implies (3), we prove the contrapositive. Suppose the minimal terms of P are all pentagons; without loss of generality, we assume that $x_{12}x_{23}x_{34}x_{45}x_{15}$ is a minimal term. Since $x_{14}x_{45}x_{15}x_{23}^2$ is not minimal, we have $M_{12} + M_{34} < M_{14} + M_{23}$. Similarly, we have,

$$M_{12} + M_{23} + M_{34} + M_{45} + M_{15} < 2M_{15} + M_{23} + M_{34} + M_{24}, \text{ and}$$
$$M_{12} + M_{23} + M_{34} + M_{45} + M_{15} < 2M_{45} + M_{12} + M_{23} + M_{13}.$$

Adding these together and cancelling, we get $M_{12} + M_{34} < M_{13} + M_{24}$. Thus, 12 and 34 are adjacent in the deficiency graph. By similar reasoning, we have adjacencies $12 - 34 - 15 - 23 - 45 - 12$ in the deficiency graph, so it has a five cycle and is not 2-colorable.

Finally, we prove that (3) implies (1). Assume without loss of generality that $x_{34}x_{35}x_{45}x_{12}^2$ is among the terms minimizing P. This implies that $x_{12}x_{34}$, $x_{12}x_{35}$, and $x_{12}x_{45}$ are each minimal terms in their respective Plücker equations. Then we can use Lemmas 19 and 20, whose proofs we omit, to obtain a decomposition of M into two tree matrices.

Lemma 19 *For any 5×5 dissimilarity matrix M such that $x_{12}x_{34}$, $x_{12}x_{35}$, and $x_{12}x_{45}$ are each minimal terms in their respective Plücker equations, there exists some 5×5 tree matrix T, such that for every $ij \in \binom{[5]}{2}$, we have $T_{ij} \geq M_{ij}$, with equality if $ij \in \{12, 13, 14, 15, 23, 24, 25\}$.*

Lemma 20 *For any 5×5 dissimilarity matrix M, there exists some 5×5 tree matrix T' such that for every pair of indices i and j, we have $T'_{ij} \geq M_{ij}$, with equality if $ij \in \{34, 35, 45\}$.*

We can now finish the proof of Theorem 18. Let T be as given in Lemma 19 and let T' be as given in Lemma 20. Then $M = T \oplus T'$ and so M has tree rank at most 2. □

In fact, we can describe precisely which subgraphs of the Petersen graph arise as deficiency graphs Δ_M for $n = 5$. There are 5 tropical Plücker relations on a 5×5 matrix, each containing 3 terms. Each term is the tropical product of terms with disjoint entries. Thus, Δ_M is a subgraph with at most 5 edges of the Petersen graph.
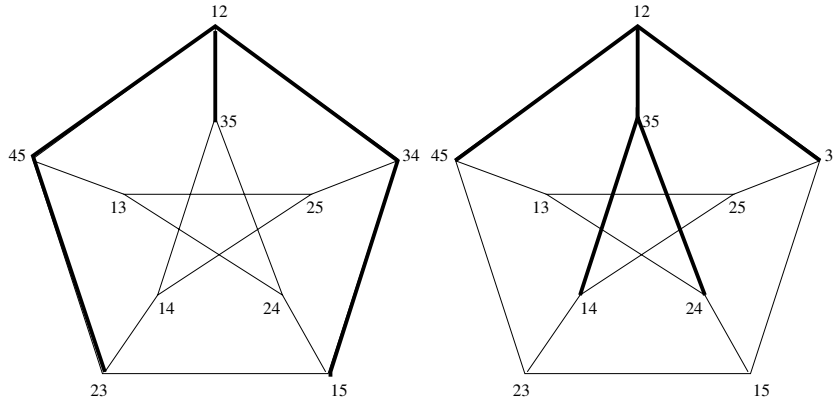

Fig. 1: The two 2-colorable possibilities for Δ_M.

Theorem 21 *Let M be a 5×5 dissimilarity matrix. Then the deficiency graph Δ_M is precisely one of the following:*

1. *The trivial graph, in which case M has tree rank* 1.

2. *A non-trivial graph with fewer than* 5 *edges, in which case M has tree rank* 2.

3. *Up to relabeling, either of the two graphs in Figure 1, in which case M has tree rank* 2.

4. *A 5-cycle, in which case M has tree rank* 3.

Proof: The matrix M is a tree matrix if and only if the four-point condition holds for all 4-tuples, i.e. if and only if Δ_M is trivial. This is the first case.

Now suppose that Δ_M is a non-trivial graph with at most 4 edges. Then, at least one four-point condition holds, so Lemma 16 implies that M has tree rank at most 2. However, at least one four-point condition is violated, so M must have tree rank exactly 2. We omit the case analysis that shows that, up to relabeling, the only two possibilities for Δ_M, assuming that it is 2-colorable, are those in Figure 1.

Finally, if Δ_M is not 2-colorable, then it must have an odd cycle. The Petersen graph has no 3-cycles, so Δ_M must be a 5-cycle, since it has at most 5 edges. □

Remark 1 *If M is the $0/1$ matrix corresponding to the 5-cycle C_5, then Δ_M is also a 5-cycle by Theorem 21. Explicitly, Δ_M has an edge for each non-adjacent pair of edges in the graph C_5. Moreover, Theorem 21 tells us that any other matrix N with tree rank 3 must have the same deficiency graph (up to relabeling). In this sense, M is the only example of a 5×5 dissimilarity matrix with tree rank 3.*

Acknowledgements

References

[1] M. Akian, S. Gaubert, A. Guterman, Linear independence over tropical semirings and beyond, preprint, arXiv:0812.3496, 2008.

[2] R. Bieri, J. Groves, The geometry of the set of characters induced by valuations, J. Reine. Angew. Math. 347:168–195, 1984.

[3] M. Develin, F. Santos, B. Sturmfels, On the rank of a tropical matrix, in "Discrete and Computational Geometry" (E. Goodman, J. Pach and E. Welzl, eds), MSRI Publications, Cambridge University Press, 2005.

[4] J. Draisma, A tropical approach to secant dimensions, Journal of Pure and Applied Algebra, 212(2):349–363, 2008.

[5] M. Drton, B. Sturmfels, S. Sullivant, Algebraic factor analysis: tetrads, pentads and beyond, Probability Theory and Related Fields 138(3/4): 463–493, 2007.

[6] J. Oxley, Matroid Theory, Oxford Univ. Press, New York, 1992.

[7] L. Pachter, B. Sturmfels, Algebraic Statistics for Computational Biology, Cambridge University Press, Cambridge, 2005.

[8] S. Radziszowski, Small Ramsey Numbers, Electronic Journal of Combinatorics, Dynamic survey DS1. updated 2009.

[9] D. Speyer, B. Sturmfels, The tropical Grassmannian, Adv. Geom., 4(3):389–411, 2004.