Tropical Geometry:

Theory, Applications, and Open Problems

(notes in progress)

George D. Torres

M394C - Spring 2019 - Ngoc Tran

These are lecture notes following M394C, Stochastic Tropical Geometry, taught by Ngoc Tran at UT Austin. The prerequisite for following these notes is nothing more than a comfort with linear algebra and graph theory. However, familiarity with linear programming, algebraic geometry, economics, and related topics will make the material more rewarding. Included are exercises and sources for further reading as well as solutions to selected exercises in the Appendix. Exercises that are more extensive or challenging will be marked with a "$*$". Please email `gdavtor@math.utexas.edu` with any corrections.

Last updated October 29, 2019

1. Tropical Origins

❖

This is a course on the basic theory of tropical geometry and its applications to other fields of mathematics. From the course description:

> "Tropical geometry is the interface of matroid theory, combinatorial optimization and algebraic geometry. It has found numerous applications in auction theory, mechanism design, game theory, complexity theory, discrete convex analysis. Tropical geometry is a variational version of combinatorial optimization, where it examines the combinatorics and geometry of the entire space of input parameters, as opposed to analyzing the solution of a particular instance. It is an excellent tool for constructing examples and counter-examples in economics, game theory, network optimization, matroid theory and algebraic geometry."

This section is meant to introduce us to the basics of tropical theory, with a view towards how it is applied to other fields.

1.1 The tropical semirings ❖

Lecture 01/22

Let $\mathbb{R}_\infty = \mathbb{R} \cup \{\infty\}$ and define the operations $a \underline{\oplus} b := \min(a,b)$ and $a \odot b = a + b$ for $a, b \in \mathbb{R}_\infty$. Then $(\mathbb{R}_\infty, \underline{\oplus}, \odot)$ is a semiring and is called the *min-plus* tropical semiring. It is an idempotent semiring (because of the idempotent operation $\underline{\oplus}$). There is also the *max-plus* tropical semiring, whose operations are $a \overline{\oplus} b := \max(a,b)$ and $a \odot b = a + b$ and underlying set is $\mathbb{R}_\infty = \mathbb{R} \cup \{-\infty\}$, and the *max-times* tropical semiring, whose operations are $a \overline{\oplus} b = \max(a,b)$ and $a \odot b = ab$. The essential idea behind tropical geometry is to replace our usual arithmetic with tropical arithmetic. We will be vocal about which of these three rings we are using (when not stated, assume it is max-plus).

Examples:

- Vector addition: we can define vectors in $\mathbb{R}^3_\infty$ in the usual way and add them component-wise:

$$(10, -1, -2) \overline{\oplus} (-4, \pi, -1) = (10, \pi, -1)$$

- Matrix multiplication: suppose that A, B are matrices of appropriate dimensions whose entries are in $\mathbb{R}_\infty$. Then their product in tropical and standard notation is:

$$(A \odot B)_{ij} = \bigoplus_{k=1}^{n} A_{jk} \odot B_{kj} = \max_{k \in [1,\dots,n]} A_{jk} + B_{kj}$$

- Eigenvectors: given $\lambda \in \mathbb{R}$, an eigenvector for $x \in \mathbb{R}^n_\infty$ of a matrix A is:

$$A \odot x = \lambda \odot x \iff \max_{j \in [1,\dots,n]} A_{ij} + x_j = \lambda + x_i \ \ (\forall i \in [1, \dots, n])$$

Do the following exercises to get a better sense of tropical arithmetic. Solutions to selected exercises will be provided in the Appendix.

Exercise 1.1.

1. Prove the distributive law: $(x \oplus y) \odot z = (x \odot z) \oplus (y \odot z)$.
2. Expand $(x \oplus 3)^{\odot 2}$.
3. Prove the Freshman's dream: $(x \oplus y)^{\odot n} = x^{\odot n} \oplus y^{\odot n}$.

Exercise 1.2. Show that the polynomial $x^2 + 15x + 2 \equiv x^{\odot 2} \oplus 15 \odot x \oplus 2$ is reducible and find a factorization.

Exercise 1.3. Write the classical matrix-vector equation:

$$\begin{pmatrix} 10 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 10x_1 + 2x_2 \\ 3x_1 + 4x_2 \end{pmatrix}$$

as a tropical equation.

Some more difficult exercises are:

Exercise* 1.4. Prove the fundamental theorem of tropical algebra: if f is a tropical polynomial in one variable, then there exists another tropical polynomial g such that $g(x) = f(x)$ for all $x \in \mathbb{R}_\infty$ and g is a product of tropical linear terms.

Exercise* 1.5. Is the function $f(x, y) = x^2 \odot 3xy \odot y^2 \oplus 4$ linearly factorizable? In other words, does it factorize into polynomials of the form:

$$L(x, y) = c_0 \oplus c_1 \odot x \oplus c_2 \odot y$$

Hint: Use Theorem 1.47. Can it be written as $f = g - h$, where g and h are linearly factorizable tropical polynomials? (Such functions are called rationally factorizable). *Hint:* see [LT17].

Given a set of m-multi-indices A, a tropical polynomial in m variables is anything of the form $f(x_1, ..., x_m) = \bigoplus_{a \in A} c_a \odot x^{\odot a} = \max_{a \in A}(c_a + \langle a, x \rangle)$. The brackets $\langle a, x \rangle$ denotes the inner product $a_1 x_1 + ... + a_m x_m$.

Definition 1.6. The *Newton polytope* associated to a polynomial is the convex hull $\mathrm{Conv}(A)$.

Definition 1.7. A function $f : \mathbb{R}^n \to \mathbb{R}$ is *convex* if for all $x, y \in \mathbb{R}^n$, we have:

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda) f(y) \quad \forall \lambda \in [0, 1]$$

It is called *concave* if $-f(x)$ is convex.

Exercise 1.8. Show that tropical polynomials in one variable (in max-plus) are convex.

1.2 First applications ❖

To get a sense for why these operations are useful, we present here some toy problems that show how large-scale operational problems can be reformulated in tropical arithmetic.

Example 1.9. Consider the problem of airport transfers. Suppose that there are three airports, A, B and C and there are two flights $\alpha : A \to C, \beta : B \to C$. These flights depart at times x_α, x_β, respectively and each have durations d_α and d_β. From airport C, there are two possible transfer flights, γ_α and γ_β. Let A_{ij} be the transfer time between arriving from flight j and going to flight γ_i (see Figure 1.1). Then we would like to know the earliest possible departure times for the flights $\gamma_\alpha, \gamma_\beta$. Let these times be b_1 and b_2, respectively. Then in classical notation, the solution is:

$$b_1 = \max(x_\alpha + d_\alpha + A_{\alpha\alpha}, x_\beta + d_\beta + A_{\alpha\beta})$$
$$b_2 = \max(x_\alpha + d_\alpha + A_{\beta\alpha}, x_\beta + d_\beta + A_{\beta\beta})$$

In max-plus notation, we can define the static matrix $C_{ij} = A_{ij} + d_j$ so that the above equations become $b = C \odot x$, where $b = (b_1, b_2)$ and $x = (x_1, x_2)$. Importantly, this is a linear equation.

Example 1.10. Let G be weighted flow network (a weighted directed graph with a distinguished source and sink). This can represent the workflow of a project, where the nodes are tasks and the weights represent the amount of time it takes a one task to feed into another. Each task must have all feeding nodes completed before it can begin. Let A_{ij} be the adjacency matrix of G, with $A_{ij} = -\infty$ when there is no edge between i and j. Let x_i denote the earliest time that node i can start its activity, with the source node x_1 starting at time u. Then the equation that describes this is:

$$x_i = \max_j A_{ij} + x_j \quad (x_1 = u)$$

As a tropical linear problem, we can write this as $x = A \odot x$. Therefore we have reduced this to an the tropical eigenvector problem.

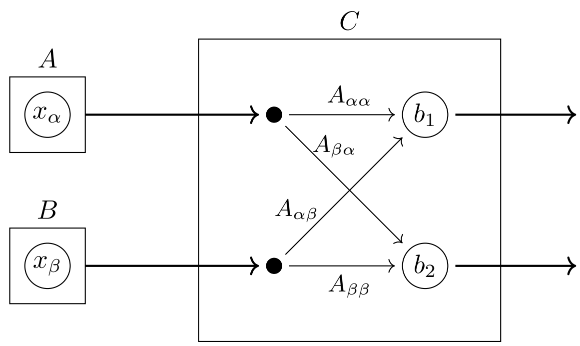


Figure 1.1: Scheduling flights with transfer times.

1.3 Tropical Linear Algebra and Network Flows ❖

Lecture 01/24

Tropical linear algebra is fundamentally tied to network flow problems. To start, we start with a square matrix C with entries in $\mathbb{R}_\infty$. This can be seen as the adjacency matrix for a directed graph G. Recall if two nodes are not connected, the value of C_{ij} is ∞ (in min-plus).

Example 1.11. We can find network (directed graph) interpretations of the following:

i) $(C \odot C)_{ij}$ is the shortest path of length 2 from i to j.

ii) $(C^{\odot n})_{ij} = (C \odot C \odot \dots \odot C)_{ij}$ is the shortest path of length n from i to j.

iii) $(C \oplus C^2 \oplus C^2 + \dots \oplus C^m)_{ij}$ is the shortest path of length at most m from i to j.

Definition 1.12. Given $C \in \mathbb{R}_\infty^{n \times n}$, the *shortest path matrix* of C is $C^+ := C \oplus C^2 \oplus C^3 \oplus \cdots$.

Lemma 1.13. *Some basic facts about the shortest path matrix are:*

1. $C^+_{ij} > -\infty$ *for all i, j if and only if there is no directed cycle in the graph defined by C with negative length. (The length of a path is the sum of the edge weights in the path).*
2. $C^+_{ij} = \infty$ *if and only if there is no directed path from i to j in the graph defined by C.*

Proof:

Exercise.

□

Let $C \in \mathbb{R}_\infty^{m \times n}$ (no longer square) and consider the equation $y = C \underline{\odot} x$. The matrix can be seen as the adjacency matrix of a weighted bipartite graph and the equation can be seen as the solution to a discrete optimal transport problem on this graph, which is a network flow problem. Generally, this falls into the class of a linear programming problem.

Definition 1.14. Given a matrix A and vectors c, b, a *linear program* is the problem of finding a minimum of $c^T x$ subject to the constraint $Ax \leq b$ (non-tropically). The *dual linear program* is to maximize $b^T y$ subject to $A^T y \leq c$. These are equivalent problems, which is known as LP duality.

What LP duality means conceptually is that minimizing cost is the same as maximizing profit. We can see this in tropical algebra by rewriting the equation $y = C \odot x$ classically:

$$
\begin{aligned}
y_i = \min_k (C_{ij} + x_k) &\iff y_i \leq C_{ik} + x_k \qquad \text{and equality must be achieved for some } k \\
&\iff x_k \geq y_i - C_{ik} \qquad \text{``} \qquad\qquad\qquad\qquad\qquad\qquad \text{''} \\
&\implies x_k = \max_i ((-C_{ik}) + y_i) \\
&\iff x = (-C)^T \overline{\odot} y
\end{aligned}
$$

The y variable is seen as cost and the x variable is seen as profit; we have just shown that minimizing cost is (almost) equivalent to maximizing profit. It is a homework exercise to understand why the third line is not quite an if and only if.

1.3.1 The Eigenvector Problem

We return to the eigenvalue problem. Let C be a square matrix with entries in $\mathbb{R}_\infty$ and consider $C \odot x = \lambda \odot x$. The meaning of λ as a transport problem is the following theorem:

Theorem 1.15 ([Cg62])**.** *In max-plus, the eigenvalue λ is the maximum mean weighted cycle in the weighted graph determined by C, i.e. $\lambda = \max_{\text{cycles } \sigma} C(\sigma)/|\sigma|$, where $C(\sigma)$ is the sum of all edges in σ and $|\sigma|$ is the cycle length.*

Proof:

In max-plus, the eigenvector equation means $\max_j (C_{ij} + x_j) = \lambda + x_i$ for all i. Therefore:

$$C_{ij} + x_j - x_i \le \lambda \qquad \forall i, j$$

If we sum this quantity over all edges $i \to j$ in σ, we get:

$$\sum_{(i \to j) \in \sigma} C_{ij} + x_j - x_i \le |\sigma| \lambda$$

$$\Rightarrow C(\sigma) \le |\sigma| \lambda$$

Now we wish to show the other direction of the inequality. We only need to exhibit one cycle for which equality holds. For each i, there must exist some least index $\phi(i)$ where $C_{i\phi(i)} + x_{\phi(i)} = \lambda + x_i$. Therefore $\phi : [n] \to [n]$ is a well-defined function. Since $[n]$ is a finite set, the graph defined by ϕ must have at least once cycle. Let this cycle be σ^*, so that $C(\sigma^*) = |\sigma^*| \lambda$.

□

Remark 1.16. If $C \in \mathbb{R}_\infty^{n \times n}$ can have at most one eigenvalue.

Exercise 1.17. What is an efficient way to compute λ? We will discuss this in a future lecture.

An amusing application of the eigenvalue problem comes from currency trading. In this example we work in max-times. Currency exchange rates can be represented by a square matrix whose i, j entry is the cost of the currency in i expressed in j dollars per unit. Are there ways to gain in the currency trading system by exchanging in cycles? If $\lambda(C) > 1$, then yes by the above theorem.

Now we return to min-plus and define the matrix $\overline{C}_{ij} = C_{ij} - \lambda$. Then $\min_k (\overline{C}_{ik} + x_k) = x_i$ means that $x_i - x_k \le \overline{C}_{ik}$. This means that the flow from i to k cannot exceed the "capacity" value $\overline{C}_{ik}$.

Exercise 1.18. Let $C = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$. Find $\lambda(C)$ (in max-plus) and find an eigenvector. Try to find all of them! Try the same with:

$$C = \begin{pmatrix} 0 & 0 & 0 \\ -2 & -2 & 1 \\ 4 & 1 & -1 \end{pmatrix}$$

Lecture 01/29

In the previous exercise, it should have been clear that trying to bash out the inequalities directly (algebraically) to find eigenvectors and eigenvalues isn't very efficient. We present two possible ways to improve this:

1. (Symbolic approach) Consider the diagonal entries of the matrix $D = C \oplus C^2 \oplus ... \oplus C^n$. This represents the shortest weighted cycle of length at most n from a node to itself.

2. (Numeric approach) Another way is to reduce this to a linear program:

$$\min \lambda \text{ subject to } C_{ij} + x_j - x_i \geq \lambda$$

where the variables are $\tilde{x} = (\lambda, x_1, ..., x_n)$. Our cost vector is $c = (1, 0, 0, ..., 0)$ and our constraint matrix is the $n^2 \times n^2$ matrix that corresponds to the n^2 inequalities above. The solution to this linear program is an eigenvalue λ and an eigenvector $(x_1, ..., x_n)$.

1.4 Tropical Polytopes and Eigenspaces ❖

The geometry of eigenspaces will lead us to the theory of tropical polytopes. To start, we assume that we have solved for λ in the eigenvector problem. Our goal is to characterize all of its eigenvectors. Defining $\bar{C}_{ij} = C_{ij} - \lambda$, we can reduce to the case of $\lambda(C) = 0$. The inequalities are:

$$C_{ij} + x_j \geq x_i \text{ and each } i \text{ has to have some } j \text{ achieving equality}$$

The inequalities define a polyhedron and the equality condition restricts feasible region to a subset of the boundary. To approach this type of problem, we introduce some terms:

Definition 1.19. For an $n \times n$ matrix, define $\text{Eig}(C)$ be the tropical eigenspace of C.

Definitions 1.20. A set S is *tropically convex* if, for all $p, q \in S$ and $a, b \in \mathbb{R}$ we have $a \odot p \oplus b \odot q \in S$. A *tropical line segment* between $p, q \in \mathbb{R}^n$ is $[p, q] := \{a \odot p \oplus b \odot q \mid a, b \in \mathbb{R}\}$. Therefore a set is tropically convex if it contains the tropical line segment between any two points in that set.

Definition 1.21. A set S is a *tropical polytope* if there exists finitely many points $v_1, ..., v_n \in S$ such that S is the tropical convex hull of $v_1, ..., v_n$.

Remark 1.22. The tropical line segment $[p, q]$ doesn't include the usual (tropicalized) condition $a \oplus b = 0, a, b \geq -\infty$ for a somewhat subtle reason. The latter condition is always true here and the former condition implies that at least one of a, b is zero (which comes from the ideompotency property). The geometry of idempotency is weird and not natural for us.

Exercise 1.23. Draw a tropical line segment $[p, q]$ for $p, q \in \mathbb{R}^2$. You should find that it is a compact polytope plus the span of $(1, 1)$.

Exercise 1.24. Show that the tropical eigenspace $\text{Eig}(C)$ is tropically convex.

Theorem 1.25. *Suppose that $C \in \mathbb{R}_\infty^{n\times n}$ and $\lambda(C) > -\infty$ in min-plus. Let C^+ be the shortest path matrix of $\overline{C} = C - \lambda$. Then:*

1. $C^+ = \overline{C} \oplus \overline{C}^2 \oplus \cdots \oplus \overline{C}^n$.
2. $C^+_{ii} = 0$ *if and only if i belongs to a critical cycle.*
3. *If i is in a critical cycle, then the ith column of C^+ is a tropical eigenvector of C.*
4. $\text{Eig}(C) = \text{Eig}(\overline{C}) =$ *tropical convex hull of* $\{$*column i of C^+, where i is a critical cycle*$\}$.

A dual statement holds for max-plus as well.

Proof:

The fact that $\lambda(C) > -\infty$ shows that all minimal paths must be simple paths, and hence we only need to consider paths of length up to n. This shows 1). We leave 2) as a homework exercise. For 3), we can check this directly by multiplying. Let $C^* = I \oplus C^+$, where I is the tropical identity matrix (zeros on diagonal and $-\infty$ off the diagonals). Then $\overline{C} \odot C^* = C^+$, and so for all i where $C^+_{ii} = 0$, we have:

$$\overline{C} \odot C^+_i = C^+_i$$

Therefore this implies that C^+_i is a tropical eigenvector of $\overline{C}$. Finally, we will prove that $\text{Eig}(C) = \text{Eig}(\overline{C}) = \text{Eig}(C^+)$. Then we will show $\text{Eig}(C^+)$ is the tropical convex hull of the critical columns (call this set S), hence 4). Given $x \in \mathrm{Eig}(C)$, we have $Cx = \lambda x \iff (C-\lambda)x = x \iff x \in \mathrm{Eig}(\overline{C})$; so $\mathrm{Eig}(C) = \mathrm{Eig}(\overline{C})$. Now let $x \in \mathrm{Eig}(\overline{C})$. Then $\overline{C}x = x$, and hence $(\overline{C})^m x = x \; \forall m > 0$. Therefore:

$$C^+x = (\overline{C} \oplus \overline{C}^2 \oplus ... \oplus \overline{C}^n)x = x \oplus x \oplus ... \oplus x = x$$

So $x \in \mathrm{Eig}(C^+)$. Given $x \in \mathrm{Eig}(C^+)$, then $C^+x = x$ and so $(C^+ \oplus I)x = x$. Therefore $\overline{C}x = \overline{C}(C^+ \oplus I)x = C^+x = x$. Therefore $x \in \mathrm{Eig}(\overline{C})$ and hence $\mathrm{Eig}(\overline{C}) = \mathrm{Eig}(C^+)$.

It follows from 3) that $S \subset \mathrm{Eig}(C^+)$, since each critical column is an eigenvector. Suppose then that $x \in \mathrm{Eig}(C^+)$, so that $\min_j C^+_{ij} + x_j = x_i$ for all i. Define:

$$\Lambda = \{i \in \{1, ..., n\} \mid i \in \text{ a critical cycle}\}$$

For each $i \in \{1, ..., n\}$, let $\Phi(i)$ be the set of j where $C^+_{ij} + x_j = x_i$. Our goal is to show that $\Phi(i) \cap \Lambda \neq \emptyset$ for all i. This ensures that we can get the last equality below (where $x \in \mathrm{Eig}(C^+)$):

$$x = C^+x = \bigoplus_{k=1}^{n} x_k \odot C^i_k = \left(\bigoplus_{k \in \Lambda} x_k \odot C^+_k\right) \oplus \left(\bigoplus_{\ell \notin \Lambda} x_\ell \odot C^+_\ell\right) = \bigoplus_{k \in \Lambda} x_k \odot C^+_k$$

This equality would mean that the eigenspace of C^+ is the convex hull of columns in Λ. We proceed case-by-case analysis. Fix $i \in \{1, ..., n\}$. If $i \in \Lambda$, then $C_i i^+ = 0$ and so $x_i = C^+_{ii} + x_i \oplus \min_{j \neq i} C^+ij + x_j = x_i$. Therefore i achieves the minimum, hence $i \in \Phi(i)$. If $i \notin \Lambda$, then $C^+_{ii} > 0$. Then $x_i = (C^+_{ii} + x_i) \oplus \min_{j \neq i}(C^+_{ij} + x_j)$ and since $C_{ii} > 0$, the minimum must be achieved for some $j \neq i$. In other words, $x_i = C^+_{ij} + x_j$ for some $i \neq j$ and hence $j \in \Phi(i)$. If $j \in \Lambda$, we're done. So assume $j \notin \Lambda$. Then C^+_{ij} is the value of the shortest path from i to j. Either this path is direct or it has intermediate nodes. We claim that we can always assume it is direct. If not, suppose k an intermediate node. Then $C^+_{ij} + x_j = (C^+_{ik} + C^+_{kj}) + x_j = (C^+_{ik} + x_k) + (C^+_{kj} + x_j - x_k)$. The first term has to be $\geq x_i$ and the second must be positive. Therefore $C^+_{ij} x_j \geq x_i$. But since j achieves the minimum, this must be a strict equality and hence $C^+_{ik} + x_k = x_i$. Then $k \in \Phi(i)$. Therefore we can replace j with k. Using the same trick, repeatedly, we can show that there must exist some node j' in Λ lying in the critical path. Thus we have shown $\Phi(i) \cap \Lambda \neq \emptyset$ for all i.

□

Corollary 1.26. $\mathrm{Eig}(C)$ *is a tropical polytope.*

Lecture 01/31

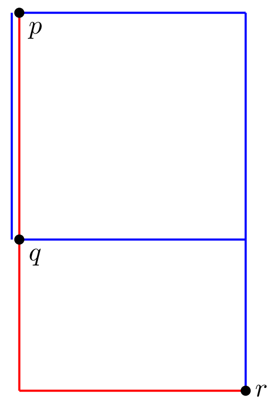
Exercise 1.27. Solve Exercise 1.18 by computing $\lambda(C), \Lambda$ first, and then computing C^+ to find all of the eigenvectors.

How do we visualize $\mathrm{Eig}(C)$? The key is the previous theorem; that it is the tropical convex hull of some collection of vectors. So how do we visualize the tropical convex hull of vectors? The trick is to think of $p \in \mathbb{R}^n$ and all of its tropical multiples as a single point. Our convention will be to normalize the first coordinate to 0 and drop it (we can think of this as an analogue of projectivization):

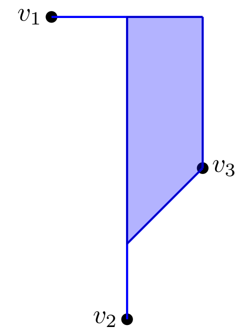
$$(x_0, ..., x_n) \in \mathbb{R}^{n+1} \mapsto (x_1 - x_0, x_2 - x_0, ..., x_n - x_0) \in \mathbb{R}^n$$

Example 1.28. Consider the column span of $C = \begin{pmatrix} 0 & 0 & 0 \\ -2 & -2 & -2 \\ 4 & 1 & -1 \end{pmatrix}$. This is the tropical convex hull of the columns. We can visualize this space in $\mathbb{R}^2$ by drawing all line segments between the columns and projectivizing down to $\mathbb{R}^2$. For example, let p and q be the first two columns. Then $a \odot p \oplus b \odot q$ for $a, b \in \mathbb{R}$ can be one of three things:

$$a \odot p \oplus b \odot q = \begin{pmatrix} a \\ a-2 \\ 4+a \end{pmatrix} \text{ or } \begin{pmatrix} a \\ a-2 \\ b+1 \end{pmatrix} \text{ or } \begin{pmatrix} b \\ b-2 \\ b+1 \end{pmatrix}$$



(a) Tropical convex hulls in min-plus (red) and max-plus (blue) of p,q, and r.



(b) A more general convex hull in max-plus.

Figure 1.2: Tropical convex hulls in $\mathbb{R}^2$

Subtracting out the first entry of each and plotting these in $\mathbb{R}^2$ gives the red line between p and q shown in Figure 1.2. The remaining lines can be filled in and the same can be done in max-plus to get Figure 1.2.

1.4.1 H and V representations of polytopes

Lecture 02/05

Given a tropical polytope $\mathcal{V}$, we have shown that it is of the form $\mathcal{V} = \text{Conv}\{v_1, ..., v_m \mid v_i \in \mathbb{R}^n\}$ (by definition). This is the so-called "V representation" of $\mathcal{V}$. An alternative representation is the "H representation" (the hyperplane representation).

Definition 1.29. A *min-plus hyperplane* $\underline{\mathcal{H}}_a$ in $\mathbb{R}^n$ *supported* by $a \in \mathbb{R}^n$ is:

$$\underline{\mathcal{H}}_a = \{x \mid \langle x, a \rangle = 0\} = \{x \mid \min_i (a_i + x_i) \text{ is achieved at least twice}\}$$

A *max-plus hyperplane* $\overline{\mathcal{H}}_a$ in $\mathbb{R}^n$ supported by $a \in \mathbb{R}^n$ is:

$$\overline{\mathcal{H}}_a = \{x \mid \langle x, a \rangle = 0\} = \{x \mid \max_i (a_i + x_i) \text{ is achieved at least twice}\}$$

These definitions are the tropical versions of the hyperplanes, which are classically defined as the zero set of a linear form. Examples of tropical hyperplanes are shown in Figure 1.3. Generally, the point $x = -a$ in $\overline{\mathcal{H}}_a$ or $\underline{\mathcal{H}}_a$ is the "apex" of the hyperplane.

Definition 1.30. Given a subset $I \subset \{1, ..., n\}$, the *half-space* associated to a hyperplane $\mathcal{H}_a$ is (in max-plus and min-plus):

$$\overline{\mathcal{H}}_{a,I} = \{x \mid \langle a, x \rangle \leq \langle a_I, x \rangle\}$$
$$\underline{\mathcal{H}}_{a,I} = \{x \mid \langle a, x \rangle \geq \langle a_I, x \rangle\}$$

Intuitively, the halfspace associated to a subset I is the set of points for which the min/max is achieved at indices in I. In $\mathbb{R}^2$, the hyperplane divides the plane into three regions, each individual region corresponding to $I = \{0\}, \{1\}$ and $\{2\}$. The intersection of a collection $(\mathbb{A}, \mathcal{I})$ of max-plus half-spaces is:

$$\bigcap_{a \in \mathbb{A}, I \in \mathcal{I}} \overline{\mathcal{H}}_{a,I} = \{Ax \leq A_I x\}$$

where A is the matrix whose rows are the supporting vectors $a \in \mathbb{A}$ and $A_{\mathcal{I}}$ is the matrix whose rows are just the entries of a_I indicated by $I \in \mathcal{I}$ (with $-\infty$ in other places). An example of an intersection of three half-spaces is shown in Figure 1.4. Note that the region of intersection can equally be described as the convex hull of the labeled points p_1, p_2, and p_3.

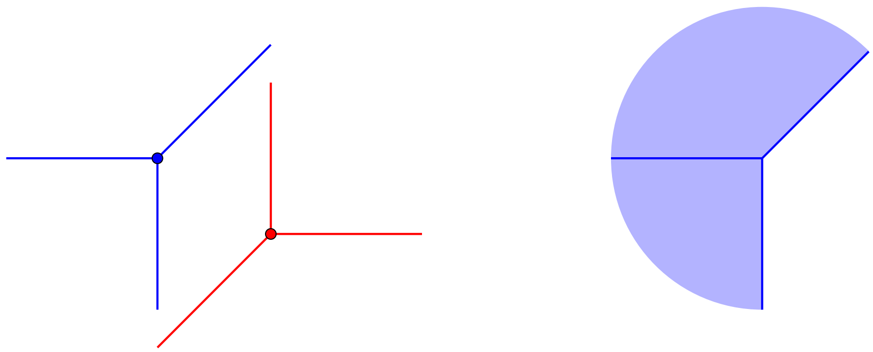


Figure 1.3: (Left) Tropical hyperplanes and half spaces in min-plus (red) and max-plus (blue) projected down to $\mathbb{R}^2$. The apices (or supports) are at the confluence of the three rays. (Right) The half space $\overline{\mathcal{H}}_{a,\{0,2\}}$.

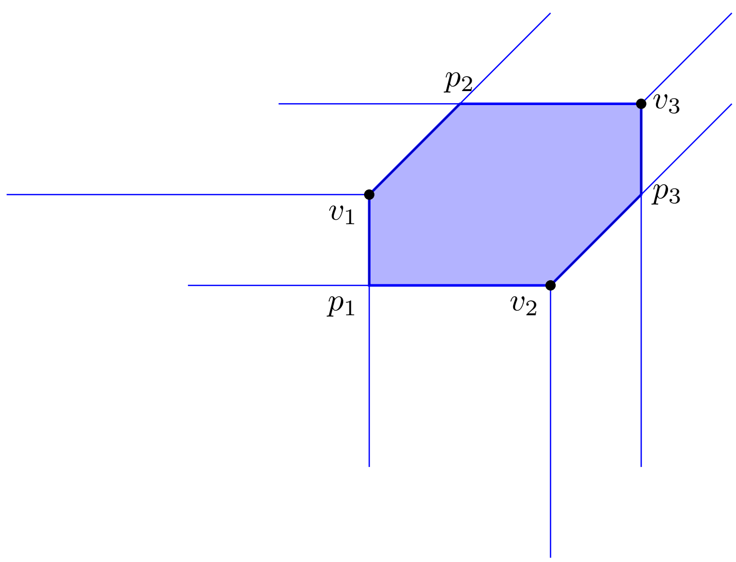


Figure 1.4: An intersection of three tropical hyperplanes supported at $\{v_1, v_2, v_3\}$ max-plus (blue).

Theorem 1.31 ([GA92]). *A set $\mathcal{V}$ is a max-plus tropical polytope if and only if it is the intersection of finitely many max-plus halfspaces.*

This theorem says that sets like $\{x \mid Ax \leq \tilde{A}x\}$, where $A, \tilde{A}$ are matrices can be written as $\{Cw\}$ for some matrix C. The former is the H representation of a polytope and the latter is the V representation of the same polytope. An example of this is when $\tilde{A} = \lambda$, an eigenvalue. The matrix C turned out to be the submatrix of $\overline{A}^{+}$ with zeros on the diagonal.

Theorem 1.32 ([DS04]). *Given a tropical matrix C, there is a row/column and max/min duality: there exists $\mathcal{I}$ such that we have an isomorphism:*

$$\{C \underline{\odot} x\} \cong \{(-C)^T \overline{\odot} y \leq (-C)^T_{\mathcal{I}} \overline{\odot} y\}$$

Theorem 1.33 ([DS04]). *The combinatorial types of a tropical polytope in $\mathbb{R}^n$ generated by m points are in 1-1 bijection with regular subdivisions of $\triangle_{m-1} \times \triangle_{n-1}$ which are in 1-1 bijection with combinatorial types of tropical hyperplane arrangements of m planes in $\mathbb{R}^n$.*

1.4.2 Tropical Hyperplane Arrangements

Lecture 02/12

To understand and prove Theorem 1.33, we need to know what tropical hyperplane arrangements are. A classical hyperplane arrangement in $\mathbb{R}^n$ is a collection of hyperplanes in $\mathbb{R}^n$. One example of such an arrangement is a Braid arrangement B_n given by:

$$H_{ij} = \{x \in \mathbb{R}^n \mid x_i - x_j = 0\}$$

for $i, j \in \{1, ..., n\}$. When $n = 2$, there is only one such hyperplane; when $n = 3$, there are three such hyperplanes and they intersect at the line spanned by $(1, 1, 1)$. In general, hyperplane arrangements give us families of polytopes by intersecting their half spaces. Each of these polyhedra and has an associated binary covector. The jth entry of this vector says if the region is above or below the jth hyperplane.

What is the set of all possible covectors of a given hyperplane arrangement? Suppose there are m hyperplanes in $\mathbb{R}^n$; then certainly there are at most 2^m such covectors, since the covectors are binary of length m. However, for most hyperplane arrangements, the number is much smaller because some covectors give an empty region.

Lemma 1.34. *The number of covectors of the Braid arrangement B_n is $n!$.*

Proof:

> Each entry of a covector is a pairwise comparison between pairs (i, j) such that this comparison is consistent across all entries. Therefore the covectors are in bijection with total orders on n elements. This is known to have cardinality $n!$.

□

We can do the same for tropical hyperplane arrangements. In this case, there is more than 2 chambers that a single hyperplane determines; there are n such chambers (where n is the ambient dimension). Therefore the covectors are subsets of $\{1, ..., n\}^m$. Sometimes we represent the covectors as n by m binary matrices.

Example 1.35. The covector of the blue polytope in Figure 1.4 is $(2, 3, 1)$. As a matrix, this is:

$$\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$$

The ith row is the covector of v_i.

Definition 1.36. The *combinatorial type* of a tropical arrangement $\mathcal{H} = \{\mathcal{H}_1, ..., \mathcal{H}_m\}$ is the set of all covectors of all of its cells.

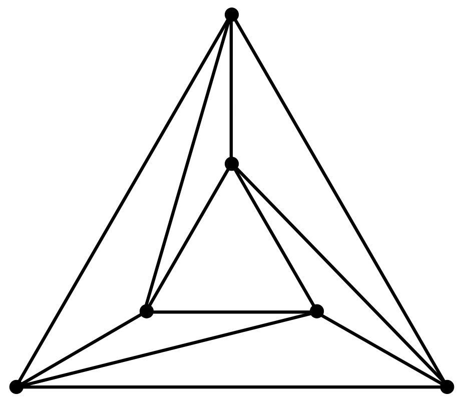


Figure 1.5: A subdivision of 6 points that is not regular.

Remark 1.37. We often denote a hyperplane arrangement $\{\mathcal{H}_1, ..., \mathcal{H}_m\}$ by their supporting vectors $\{t^1, ..., t^m\}$.

Lemma 1.38. *The combinatorial type of a tropical hyperplane arrangement $\mathcal{H} = \{t^1, ..., t^m\}$ is the set of all possible optimal transport plans on a bipartite graph on (m, n) nodes with edge $i \to j$ weighted to t^i_j.*

Remark 1.39. Recall from the beginning of §1.3 that a pair (x, y) satisfying $y = C \odot x$ is an optimal transport pair (a solution to the optimal transport problem determined by C). An *optimal transport plan* is a matching on the bipartite graph corresponding to the solution (x, y).

1.4.3 Regular Subdivisions

Now we turn our attention to regular subdivisions of simplices, which is the second part of Theorem 1.33. Loosely, a subdivision of a set of points $P \subset \mathbb{R}^n$ is a subdivision of the convex hull of P into polytopes with vertices at points in P. A subdivision Δ is called *regular* if there exist $c_p \in \mathbb{R}$ for each $p \in P$ such that Δ is the projection of the lower convex hull of the set $\{(p, c_p)\} \subset \mathbb{R}^{n+1}$.

Exercise 1.40. Show that all subdivisions of a finite set of points $P \subset \mathbb{R}$ are regular.

Exercise* 1.41. Show that the subdivision shown in Figure 1.5 is not regular.

When we write $\triangle_{m-1} \times \triangle_{n-1}$, we think of it as a set of points:

$$\{e_i \in \mathbb{R}^m \mid i \in \{1, ..., m\}\} \times \{e_j \in \mathbb{R}^n \mid j \in \{1, ..., n\}\}$$

The vectors e_i are standard unit vectors in $\mathbb{R}^n$ or $\mathbb{R}^m$. These are the vertices of $\triangle_{m-1} \times \triangle_{n-1}$. When we say a regular subdivision of $\triangle_{m-1} \times \triangle_{n-1}$, we mean a regular subdivision of this set in $\mathbb{R}^{n+m}$. The bijection in Theorem 1.33 is as follows. Given $C \in \mathbb{R}^{m \times n}$, define Δ_C to be the regular subdivision of $\triangle_{m-1} \times \triangle_{n-1}$ defined by the lifts of (e_i, e_j) to height C_{ij}. We can also consider the tropical hyperplane arrangement $\mathcal{H}_C$ in $\mathbb{R}^n$ whose supports are the rows of C. The Theorem says that:

$$\Delta_C = \Delta_{C'} \iff \text{combinatorial type of } \mathcal{H}_C = \text{combinatorial type of } \mathcal{H}_{C'}$$

Definition 1.42. Given a matrices C, C' as above, we say that $C \sim_\Delta C'$ if $\Delta_C = \Delta_{C'}$. We say that $C \sim_{\mathcal{H}} C'$ if the hyperplane arrangements induced by C and C' have the same combinatorial type.

With the above definition, yet another restatement of Theorem 1.33 is that $\sim_\Delta$ and $\sim_{\mathcal{H}}$ are the same equivalence relation.

Lecture 02/14 ♡

To proceed with the proof, we will use the Caley trick. This will say that regular subdivisions of $\triangle_{m-1} \times \triangle_{n-1}$ are in bijection with regular mixed subdivisions of $m \cdot \triangle_{n-1}$. A precise definition and discussion of mixed subdivision can be found in Chapter 4 of [Jos]. Roughly, it has to do with subdivisions of the Minkowski sum of finite sets $A+B$. If point $p = a + b \in A + B$ can be expressed as $p = a' + b'$ for $a \neq a'$ and $b \neq b'$, we give p two distinct labels. Given $A' \subset A$ and $B' \subset B$, we define the *mixed cell* with label set $A' \times B'$ to be:

$$M(A', B') := \mathrm{Conv}\{a + b \mid a \in A', b \in B'\}$$

A polyhedral subdivision of $A+B$ is *mixed* if it is formed from mixed cells. The *Cayley embedding* of the two point configurations A, B in $\mathbb{R}^d$ is the point configuration:

$$\mathcal{C}(A, B) = \{(a, e_1) \mid a \in A\} \cup \{(b, e_2) \mid b \in B\} \subset \mathbb{R}^{d+2}$$

Any polytope of the form $\mathrm{Conv}(\mathcal{C}(A', B'))$ for subsets A', B' is called a *Caley cell*. Intersecting $\mathcal{C}(A', B')$ with the hyperplane $\{x_{d+2} = x_{d+1}\}$ yields the Minkowski cell $M(A', B')$ (up to scaling by a factor of 2). Let Σ be a polyhedral subdivision of $\mathcal{C}(A, B)$ and let:

$$M(\Sigma) = \{M(A', B') \mid \mathrm{Conv}(\mathcal{C}(A', B')) \in \Sigma\}$$

Note that $M(\Sigma)$ is a subdivision of $\frac{1}{2}(A+B)$ and we call it the *mixed subdivision* associated to Σ. More explicitly, M takes a subdivision of $\mathcal{C}(A, B)$ and returns a (mixed) subdivision of $\frac{1}{2}(A+B)$ given by intersecting with the hyperplane $\{x_{d+2} = x_{d+1}\}$. For some small examples of these constructions, see Chapter 4 of [Jos].

Theorem 1.43 (Cayley's Trick)**.** *The map M from the set of subdivisions of $\mathcal{C}(A, B)$ to the set of mixed subdivisions of $A+B$ is a bijection that preserves refinement.*

A version of this theorem was first proved by Sturmfels in 1992. We will not prove it in this class.

A quick application of mixed subdivisions is to tropical polynomial factorization. Recall that a multivariate tropical polynomial associated to a finite set $A \subset \mathbb{N}^n$ is:

$$f = \bigoplus_{a \in A} c_a \odot x^{\odot a} = \bigoplus_{a \in A} c_a \odot \langle x, a \rangle$$

Definition 1.44. The *tropical hypersurface* $V(f) \subset \mathbb{R}^n$ associated to f is the set of x such that the minimum is achieved at least twice.

Exercise 1.45. Suppose $f = f_1 \odot f_2$. Show that $V(f) = V(f_1) \cup V(f_2)$. Show also that Δ_f is a mixed subdivision of N_f with respect to N_{f_1}, N_{f_2}.

Observe that f induces a regular subdivision of its newton polytope $N_f = \mathrm{Conv}(A)$ defined by lifting the point $a \in A$ to height c_a. We denote this by Δ_f.

Lemma 1.46. *$V(f)$ is dual to Δ_f as a polytope.*

Theorem 1.47. *Δ_f is a regular mixed subdivision of N_f with respect to $A = N_{f_1}, B = N_{f_2}$ if and only if $V(f) = V(f_1) \cup V(f_2)$ (i.e. $f = f_1 \odot f_2$).*

This shows that tropical polynomial factorization is achieved by finding a regular mixed subdivisions of a polytope.

Exercise 1.48. Show that the polynomial $3 \odot x^3 \odot y^3 \oplus 4$ is irreducible.

Lecture 02/19

A more general version of the Caley trick is that regular mixed subdivisions of P with respect to $(P_1, ..., P_m)$ is in 1-1 correspondence with regular subdivisions of $(P_1, e_1) \times ... \times (P_m, e_m)$.

Proof of Theorem 1.33:

> The combinatorial type of a hyperplane arrangement $\mathcal{H}(c_1, ..., c_m)$ is exactly the set of possible labels of vertices of $\triangle_{n-1} + ... + \triangle_{n-1}$ as mixed cells (because the dual of a hyperplane is the simplex $\triangle_{n-1}$). But the labels of vertices determines the labels of all cells in Δ_f, where f is the product of all the linear polynomials in the arrangement. Therefore C and C' have the same combinatorial type if and only if the labels of all the vertices of Δ_C and $\Delta_{C'}$ are the same. But this is true if and only if all the cells of Δ_C and $\Delta_{C'}$ are the same, which is true if and only if $\Delta_C = \Delta_{C'}$.

□

1.5 Gröbner Bases and Integer Programming ❖

Recall a standard linear program has primal and dual forms:

$$\min c^T x, Ax = b, x \geq 0 \tag{P}$$

$$\max y^T b, yA \leq c \tag{D}$$

We denote these by $LP_{A,c}(b)$ and $LP^*_{A,b}(c)$.

Definition 1.49. We say $c \sim_{LP} c'$ if and only if $LP_{A,c}(b)$ and $LP_{A,c'}(b)$ have the same set of optimal solutions for all b.

One can check that this is a proper equivalence relation.

Theorem 1.50. *$\mathbb{R}^N / \sim_{LP}$ is the secondary fan of $A \in \mathbb{R}^{d \times N}$. In other words:*

1. *$\mathbb{R}^N / \sim_{LP}$ is a fan.*
2. *This fan is the formal fan of a polytope, and*
3. *this polytope can be read off with a formula from the columns of A.*

We can view c_i as the lift of the ith column of A as a vector in $\mathbb{R}^d$, and hence c can be associated to a regular subdivision Δ_c of the set of columns of A. Then:

Theorem 1.51. *$c \sim_{LP} c'$ if and only if $\Delta_c = \Delta_{c'}$.*

1.5.1 The Transport Program

Lecture 02/21

Our goal today is to give a linear-programmatic interpretation of tropical hyperplane arrangements. The following lemma links the primal and dual linear programs:

Lemma 1.52 (Complementary slack-ness condition)**.** *Assume that (P) and (D) are finite, let x_0 be a feasible solution to (P), and let y_0 be a feasible solution to (D). Then the following are equivalent:*

1. *x_0 and y_0 are optimal solutions to their respective programs.*
2. *$c^T x_0 = y_0^T b$.*
3. *If $(x_0)_i > 0$, then $(y_0 A)_i = c_i$. (in other words $x_0^T(c - y_0 A) = 0$).*

The program we care about is the *transport program*. Let $c \in \mathbb{R}^{m \times n}$

$$\min \left(\sum_{i \in [m], j \in [n]} x_{ij} c_{ij} \right) \text{ s.t. } \sum_{j \in [n]} x_{ij} = b_i \text{ and } \sum_{i \in [m]} x_{ij} = \tilde{b}_j,\ x_{ij} \geq 0 \tag{P}$$

$$\max\left(\sum_i \tilde{y}_i b_i - \sum_j \tilde{z}_j b_j\right) \text{ s.t. } \tilde{y}_i - \tilde{z}_j \le c_{ij} \tag{D}$$

One can check that this this is actually a linear program by writing it in standard form. To see why it is called the transport program, consider a (m,n) (left-to-right directed) bipartite graph whose $(i \to j)$ edge is weighted by c_{ij}. We can interpret b_i as the amount of stuff that outgoing node i wants to sell and $\tilde{b}_j$ as the amount of stuff that ingoing node j wants to buy. The values x_{ij} represent the amount of materials sent from i to j (the transport plan). This is the transport solution that must meet the supply and demand criteria. The (P) linear program expresses that we want to choose a transport plan that minimizes transport costs coming from c_{ij} and meets supply and demand. The dual program can be seen as maximizing profit subject to the constraint that profits must be at most the cost of business (i.e. no net profit, since we assume we are in market equilibrium). What does complementary slack-ness mean in the transport program? It says that, given optimal transport solutions x^*, y^*, maximizing profit is equivalent to minimizing cost. It also says that if you are transporting a nonzero amount on $(i \to j)$ (i.e. $x^*_{ij} > 0$), then $y^*_i - z_j = c_{ij}$, which says that there is no profit happening.

The tropical connection is as follows. The dual constraint $y_i \le c_{ij} + z_j$ implies that $y \le C \odot z$, where C is the matrix whose entries are c_{ij}. Let $Q_C = \{(y,z) \mid y_i - z_j \le c_{ij}\}$. A key observation is that each cell of the tropical hyperplane arrangement whose supports are the rows of C is the projection onto the z coordinate of the face of Q_C supported by some $(b, \tilde{b})$. From this observation, we can see how it might be plausible that $c \sim_{LP} c'$ if and only if $C \sim_\triangle C'$.

Lecture 02/26

Given $c, c' \in \mathbb{R}^{m\times n}$, we define the an equivalence $\sim_{IP}$ similarly to $\sim_{LP}$. We say $c \sim_{IP} c'$ if and only if the transport program has the same set of optimal integer solutions for all integer vectors b. We will show today $c \sim_{IP} c'$ if and only if $c \sim_{\mathcal{H}} c'$ (where $\sim_{\mathcal{H}}$ is the equivalence of hyperplane arrangements). The reference we used for this lecture is [Th98].

Consider the ring $\mathbb{R}[x] \equiv \mathbb{R}[x_1, ..., x_M]$. Every $u \in \mathbb{N}^M$ determines a monomial $x^u = \prod_{i=1}^M x_i^{u_i}$. Given a matrix A, we define the *toric ideal:*

$$\begin{aligned} I_A &= \langle x^{u_+} - x^{u_-} \mid Au_+ = Au_-, \text{ for } u_+, u_- \in \mathbb{N}^M \rangle \\ &= \langle x^{u_+} - x^{u_-} \mid u = u_+ - u_- \in \ker_{\mathbb{Z}} A \rangle \end{aligned}$$

where in the second presentation we are decomposing u into a difference of nonnegative vectors $u_\pm$. The toric ideal of the transport program is the toric ideal associated to the matrix of the primal transport program:

$$I_T = \langle x^{u_+} - x^{u_-} \mid (u_+)_{i\cdot} = (u_-)_{i\cdot}\ \forall i \in [m] \text{ and } (u_+)_{\cdot j} = (u_-)_{\cdot j}\ \forall j \in [n] \rangle$$

where $(u)_{i\cdot}$ denotes summing over the second index. Here $(u_+)_{ij}$ is the number of times the edge $(i \to j)$ is used in this transport problem defined by u_+. So, a pair (u_+, u_-) is a pair of plans with the same in/out flow at each node.

Definition 1.53. Let $c \in \mathbb{R}^M$. The term ordering $\succ_c$:

$$x^u \succ_c x^v \iff c \cdot u > c \cdot v$$

is a partial order on the monomials in $\mathbb{R}[x]$.

Definition 1.54. Given $c \in \mathbb{R}^M$, for a polynomial $f = \sum_u a_u x^u \in \mathbb{R}[x]$, its *initial form* $\mathrm{in}_c(f)$ is the sum of all terms $a_u x^u$ with maximal order under $\succ_c$. For the toric ideal I_T, the *initial ideal* is $\mathrm{in}_c(I_T) = \langle \mathrm{in}_c(f) \mid f \in I \rangle$.

Lemma 1.55. *$c \sim_{IP} c'$ if and only if $\mathrm{in}_c(I_T) = \mathrm{in}_{c'}(I_T)$.*

Lemma 1.56. *For the transport problem, $\mathrm{in}_c(I_T)$ is an ideal of "bad plans," i.e. $x^u \in \mathrm{in}_c(I_T) \iff u$ is not an optimal transport plan in the program with cost c.*

Lemma 1.55 follows from Lemma 1.56. From the argument in the proof of Theorem 2.3 [Th98], we can claim that $c \sim_{\mathcal{H}} c'$ if and only if for all b the support of solutions to $LP_{A,c}(b)$ is the same as the support of solutions to $LP_{A,c'}(b)$. What we need is that the solution sets are the same, which is stronger. The key is to go to integer programs and use the toric ideal. There are some missing parts here that I didn't catch. Apologies.

Definition 1.57. A *circuit* of A is a primitive, non-zero vector $u \in \ker_{\mathbb{Z}}(A)$ such that its support (index set of nonzero terms) is minimal with respect to inclusion.

Lemma 1.58. *For the transport program, the circuits are all cyclic shifts.*

Theorem 1.59. *Circuits generate I_T.*

Lecture 02/07

1.6 Open questions in Tropical Linear Algebra ❖

Areas of current research and sources of open problems in tropical linear algebra abound. Some such problems are below. For more details and sources on these, see the suggested projects handout on the course website.

1. *Commuting matrices.* Suppose $A, B \in \mathbb{R}_{\infty}^{n\times n}$. When does $A \odot B = B \odot A$? We denote $C(A) = \{B \in \mathbb{R}_{\infty}^{n\times n} \mid A \odot B = B \odot A\}$. It can be shown that $C(A)$ is a tropical polyotpe. How do we characterize the V and H representations of $C(A)$? What does the geometry of $C(A) \cap C(B)$ look like? Some primers on this material can be found on the first homework.

2. *H and V Representations.* We saw that the H and V representations are two equivalent ways of describing a tropical polytope. What are efficient algorithms that can transfer between these two representations? There is literature and research in going from H to V by Gaubert, Allamigeon, and Goubalt. A related problem is finding a tropical analogue of Gaussian elimination.

3. *Max-linear Bayesian networks.* A Gaussian graphical model is a graph representing conditional independence on a set of Gaussian random variables $X_1, ..., X_n$. These models represent complicated interactions of systems, but often fail in extreme events. The solution is a Max-linear Bayesian model. These boil down to observing variables X, Z and finding C such that $X = C \odot X \oplus Z$.

4. *Tropical Tensors.* We have a notion of a tropical matrix, which is a tensor. Is there a way to define a more general tropical tensor? If so, what is a generalization of the eigenvector problem? A reason we might care about this is that there is a geometric way to interpolate between tropical polytopes and hyperplane arrangements. In turn, hyperplane arrangements have a combinatorial connection to regular subdivisions of $\triangle_{m-1} \times \triangle_{n-1}$.

5. *Perron-Frobenius and tropical eigenvectors.* Perron Frobenius theory deals with matrices having positive entries. They have a unique positive eigenvalue λ and an associated eigenvector generating the eigenspace. In the max-times algebra, a matrix with positive entries has a special eigenvalue λ_{max} such that, in an appropriate limit, $\lambda \to \lambda_{max}$. If the max-times eigenspace is 1-dimensional, then in the same limit the eigenspaces coincide as well. However, if the max-times eigenspace is higher dimensional, it is not known what the behavior is. There are some nice numerical approaches to this.

6. *Pairwise ranking.* This is the problem of finding an ordering on n items by only asking for pairwise comparisons. Such comparisons are encoded in a matrix A, where $A_{ij} > 0$ is how more preferable i is than j. It must satisfy $A_{ij} = 1/A_{ji}$. A ranking algorithm takes A as input and returns a ranking vector $x \in \mathbb{R}^n$. Some popular methods for doing this use Perron-Frobenius eigenvectors, max-times P-F eigenvectors, and geometric row means. A lot of what is not understood yet stems from our lack of understanding of the theory from5.

7. *Spectral theory of random matrices.* Given a matrix $A \in \mathbb{R}^{n\times n}$ with entries iid to some distribution F. What can we say about the distribution of $\lambda_+(A)$, the min-plus eigenvalue, and $v_+(A)$, the min-plus eigenvector? Some known cases are with $F = \exp(1), \mathrm{Ber}(1/2)$, and the uniform distribution on $\{1, ..., m\}$. However, this question hasn't been fully answered even in these cases.

8. *Interior point methods in classical LP.* In 2018, Allamigeon, Benchimol, Gaubert, and Joswig disproved the continuous Hirsch conjecture using tropical methods. This was widely influential in showing that tropical methods are promising. Some good work to be done here is to make their methods more accessible.

9. *Identities in the semigroup of tropical matrices.* Suppose a, b are 2 by 2 upper triangular tropical matrices with indeterminate entries. An identity is something of the form like $abbaba = aabbaab$ (and the equality must hold for all matrices a, b). When we have an identity, we get three polynomial equalities in the indeterminates. It has been shown that the shortest nontrivial identity has length 10 and is:

   $$abbaaababba = abbabaabba$$

   This was found by constructing polytopes in $\mathbb{R}^2$ out of the two words. Then two words are equal if and only if their corresponding polytopes are equal. A good source of open questions in this topic are listed in [JT18].

10. *Tropical geometry of deep neural networks.* The authors in [Zh18] showed that tropical rational functions are exactly neural networks with ReLU activation functions. In this paper, the authors raise some questions that have yet to be answered. An example of such a question is: when are two tropical rational functions that have different algebraic expressions equal to each other? Equivalently, when are two neural networks with different architectures and weights actually the same function? What about approximately the same?

2. Tropical Linear Spaces and Matroid Theory

❖

Lecture 02/28

A matroid is an abstraction of linear algebra. In linear algebra, we often encode a matrix E as a list of columns $\{e_i\} \subset \mathbb{R}^m$. A basis for E can be thought of as a maximal independent subset of $\{e_i\}$.

Definition 2.1. A *matroid* is a pair $(E, \mathcal{B})$, where E is a finite set and $\mathcal{B}$ is a collection of subset of E called *bases*. It must satisfy the basis exchange axiom: given $B, B' \in \mathcal{B}$ distinct, for each $i \in B \setminus B'$, there exists $j \in B' \setminus B$ such that $(B \setminus \{i\}) \cup j \in \mathcal{B}$.

If $E = \{e_1, ..., e_n\}$ is a collection of vectors in $\mathbb{R}^n$, then taking $\mathcal{B}$ to be the set of linearly independent subsets of E gives us a matroid.

Definition 2.2. A *rank* on a set E is a function $\rho : 2^E \to \mathbb{N}$ such that:

1. $\rho(\emptyset) = 0$.
2. $\rho(A) \leq \rho(A + i)$ for all $A \subset E, i \in E$.
3. $\rho(A) \leq |A|$.
4. $\rho(A) + \rho(B) \geq \rho(A \cap B) + \rho(A \cup B)$.

An equivalent definition of a matroid is:

Definition 2.3. A matroid (E, ρ) is a set E and a rank ρ on E.

This is equivalent to our first definition as follows. Given a matroid (E, ρ), the induced basis $\mathcal{B}$ is the set of subsets $b \subset E$ such that $\rho(b)$ is maximal. Conversely, given a matroid $(E, \mathcal{B})$, we can define a rank ρ as $\rho(A) = \max_{B \in \mathcal{B}} |B \cap A|$. One can verify that this rank satisfies the required 4 axioms.

In linear algebra, a circuit is a collection of minimally dependent sets of vectors. Equivalently, a circuit C is a minimal collection $\{e_i\} \subset E$ such that $\sum_i a_i e_i = 0$ has a nontrivial solution for a_i. To generalize this, we can present a third definition of a matroid via independence:

Definition 2.4. A matroid $(E, \mathcal{I})$ is a set E with a collection $\mathcal{I}$ of subsets satisfying:

1. $\emptyset \in \mathcal{I}$.
2. If $I' \subseteq I$ and $I \in \mathcal{I}$, then $I' \in \mathcal{I}$.
3. If $I_1, I_2 \in \mathcal{I}$ and $|I_1| < |I_2|$, then there exists $e \in I_2 - I_1$ such that $I_1 \cup e \in \mathcal{I}$.

Given a matroid as above, we can construct a set of bases $\mathcal{B}$ given by the set of maximal independent sets. In other words, $\mathcal{B} = \{I \in \mathcal{I} \mid I \not\subset I' \; \forall \, I' \in \mathcal{I}\}$. Equivalently, $\mathcal{B}$ is the set of independent sets with $\rho(I) = |I|$, where ρ is a rank on E.

Definition 2.5. The *closure* on a matroid (E, ρ) is the function $\mathrm{cl} : 2^E \to 2^E$ given by:

$$\mathrm{cl}(A) = \{e \in E \mid \rho(A + e) = \rho(A)\}$$

Definition 2.6. The *circuits* $\mathcal{C}$ of E are the collection of subsets of E satisfying $X \in \mathcal{C} \Rightarrow X$ is the smallest nonempty set such that $\forall x \in X, x \in \mathrm{cl}(X \setminus x)$. A *circuit* is an element of $\mathcal{C}$.

A fourth equivalent way to define a matroid is to specify E and $\mathcal{C}$. The subsets of the set $E_n = [n]$ can be identified with the cube $\{0,1\}^n$. Given a matroid $([n], \mathcal{B})$, the *matroid base polytope* $P_{\mathcal{B}}$ is the convex hull of $\{b \in \{0,1\}^n \mid b \in \mathcal{B}\}$. An example of a matroid on E_3 is shown in Figure 2.1.

Exercise 2.7. Find all circuits of the matroid in Figure 2.1.

Lecture 03/07

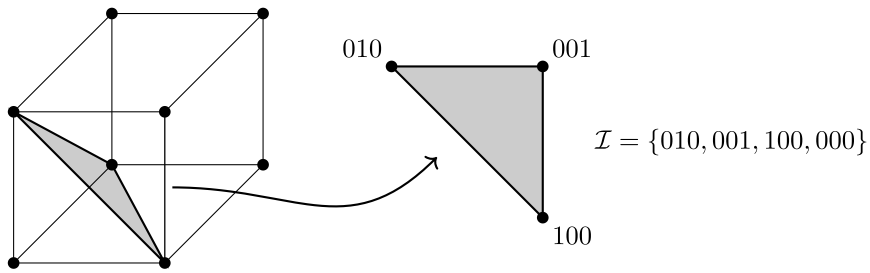


Figure 2.1: The base polytope of a matroid on E_3 inside of the cube. The corresponding independent set is on the right.

2.1 Geometry of Matroids

Given a graph $G = (V, E)$, we can define a matroid by taking the elements to be the edges E and the circuits to be minimal cycles of G. A minimal cycle is a cycle that can't be written as the union of two cycles.

Exercise 2.8. Given a graphical matroid M as above, characterize the bases, independent sets, and rank function of M.

Recall that, for matroids on $([n], \mathcal{B})$, we defined the matroid base polytope $P_{\mathcal{B}}$. Our goal is to read off the characterizing rank function ρ, circuits $\mathcal{C}$, and independent sets $\mathcal{I}$ of the matroid from the geometry of $P_{\mathcal{B}}$.

Lemma 2.9. *Given a matroid $([n], \mathcal{B})$, the matroid base polytope*

$$P_{\mathcal{B}} = \{x \in \mathbb{R}^n \mid 0 \leq x_i \leq 1, \langle x, S\rangle \leq \rho(S)\ \forall S \subseteq [n] \textit{ with equality when } S = [n]\}$$

where $\langle x, S\rangle := \sum_{i \in S} x_i$ for a subset $S \subset [n]$.

This gives the H-representation of the matroid base polytope in terms of the rank function ρ. In the example of Figure 2.1, the stated inequalities are:

$$\begin{aligned} x_1 &\leq \rho(100) = 1 \\ x_2 &\leq \rho(010) = 1 \\ x_3 &\leq \rho(001) = 1 \\ x_1 + x_2 &\leq \rho(110) = 2 \\ x_2 + x_3 &\leq \rho(011) = 2 \\ x_1 + x_3 &\leq \rho(101) = 2 \\ x_1 + x_2 + x_3 &\leq 2 \end{aligned}$$

We see that this indeed cuts out the pictured triangle (all but the last inequality is superfluous).

Proof of Lemma 2.9:

If $x \in P_{\mathcal{B}}$, we want to show that it satisfies the stated inequalities. It suffices to check these inequalities on the vertices of $P_{\mathcal{B}}$ (i.e. the bases). Let $x \in \mathcal{B}$. Then this is independent, and so any $T \subseteq x$ is independent. Thus for any $S \subseteq [n]$:

$$\langle x_i, S\rangle = \sum_{i \in S} x_i \leq \max_{T \subseteq S, T \in \mathcal{I}} |T| = \rho(S)$$

This follows because $x \cap S$ is an independent set. Then $|S \cap x| = \langle x_i, S\rangle$. The RHS follows from the definition of rank. In particular, if $S = [n]$, then $\langle x_i, S\rangle = |x| = \rho([n])$. This proves one direction.

Now we show the other direction. If x is a vertex of the RHS of Lemma 2.9 (call this set Q), then we claim $x_i \in \{0, 1\}$ for all $i \in [n]$. To show this, observe that if x is a vertex of Q, then it must satisfy $n-1$ linearly independent equalities and that these are linearly independent from the equality $\langle x, [x]\rangle =$

$\rho([n])$. Suppose these equalities are satisfied at $S_1, ..., S_{n-1}$ for $S_i \subseteq [n]$:

$$\langle x, S_i \rangle = \rho(S_i)$$

This is a matrix equation. We can produce binary rows $T_1, ..., T_{n-1}$ that are mutually disjoint such that $\langle x, T_i \rangle = \rho(T_i)$ (we are thinking of T_i also as subsets of $[n]$).[a] Since the T_i cover $[n]$ and are disjoint, by the Pigeon Hole principle, then T_i are either all singletons or all singletons and a doubleton. However, the second case cannot happen because $\langle x, T_i \rangle = \rho(T_i)$ are independent relations. Therefore $|T_i| = 1$ for all $i < n$, and hence $x_i = \rho(T_i)$ for $i < n$. This shows $x_1, ..., x_{n-1}$ are integers. Moreover, x_n is an integer by the relation:

$$x_n = \rho([n]) - \sum_{i=1}^{n-1} \rho(\{i\})$$

Therefore $x_i \in \mathbb{N}$ for all i and therefore $x_i \in \{0, 1\}$. Thus $P_{\mathcal{B}}$ is the convex hull of basis points in E.

[a] We will justify this claim later.

□

Theorem 2.10. *Let $P \subseteq [0,1]^n$ be a polytope. Then $P = P_{\mathcal{B}}$ for some matroid $([n], \mathcal{B})$ if and only if the edges of P are parallel to $e_i - e_j$, where e_i is the ith unit vector in $\mathbb{R}^n$.*

This is due to Murota circa 2003 and later by Gelfand et. al in 2008. See also [Mc15].

Definition 2.11. For $w \in \mathbb{R}^n$, let $P_{\mathcal{B},w} = \mathrm{Conv}(\{b \in P_{\mathcal{B}} \mid \langle w, b \rangle = \max_{a \in P_{\mathcal{B}}} \langle w, a \rangle\}$.

Corollary 2.12. *For $w \in \mathbb{R}^n$, then $P_{B,w}$ is the base polyotpe of some matroid M_w.*

The easiest way to prove this is to use Theorem 2.10. The matroid M_w is called the initial matroid of w for M.

Returning to Exercise 2.8, the independent sets are the trees of G (subgraphs with no cycles), the bases are spanning trees, and the rank of the matroid is the number of edges minus 1. If we choose a bijection between the edges E and $[n]$, we can consider the matroid base polytope $P_{\mathcal{B}}$. This is the convex hull of indicator vectors of spanning trees. Given $w \in \mathbb{R}^n$, representing the weights of each edge, the initial matroid $P_{\mathcal{B},m}$ is the convex hull of max-weight spanning trees.

Proposition 2.13. *For each circuit C of M, let $C_w := \{j \in C \mid w_j = \min_{i \in C} w_i\}$. Then the circuits of M_w are all minimal such sets (with respect to inclusion) over all circuits of M.*

Proof (done on 3/12):

We denote $M^{\mathcal{C}}$ to be the matroid whose circuits are those defined in the Proposition. We wish to show that this matroid is the same as the basis-defined initial matroid M_w.

Claim 1: If $B \in \mathcal{B}(M)$ but $B \notin \mathcal{B}(M_w)$, then $B \notin \mathcal{B}(M^{\mathcal{C}})$. Let $W = \max_{b \in B(M)} \langle w, b \rangle$. Suppose B_1 does not achieve this maximum weight (i.e. $B_1 \notin \mathcal{B}(M_w)$. Our goal will be to show that B_1 is not an independent set, with respect to the matroid $M^{\mathcal{C}}$. Pick $B_2 \in \mathcal{B}(M_w)$ such that $B_1 \cap B_2$ has the smallest possible weight. Let $i \in B_1$ a maximal weight element not contained in B_2. Now we can perform a basis exchange with some $j \in B_2$ to get $B_3 = B_2 - j + i$. Then:

$$w(B_3 - B_1) < w(B_2 - B_1)$$

$$\begin{aligned}\Rightarrow w(B_3) &= w(B_1) + w(B_3 - B_1) \\ &< w(B_1) + w(B_2 - B_1) = w(B_2)\end{aligned}$$

Therefore $w_i < w_j$. Now consider $B_1 + j$. This is not a basis of M, since $j \notin B_1$. Therefore there exists a circuit C of M inside of $B_1 + j$, and hence C must contain j. This circuit induces $C_w = \{j \in C \mid w_j =$

$\min_{i \in C} w_i\}$. Since $w_j > w_i$, so $j \notin C_w$. Therefore $C_w \subset B_1$ and hence B_1 is not a basis in $\mathcal{M}^{\mathcal{C}}$.

Claim 2: If $B \in \mathcal{B}(M_w)$, then $B \in \mathcal{B}(M^{\mathcal{C}})$. It is sufficient to prove that if $B \in \mathcal{B}(M_w)$, then B has no circuits in $M^{\mathcal{C}}$ (i.e. $B \in \mathcal{I}(M^{\mathcal{C}})$. This follows because:

$$\text{circuits of } M^{\mathcal{C}} \subseteq \text{ circuits of } M$$

$$\Rightarrow \text{ bases of } M^{\mathcal{C}} \subseteq \text{ independent sets of } M$$

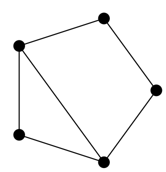
We proceed by contradiction. Take $B \in \mathcal{B}(M_w)$ and suppose that B is not independent; i.e. that it contains a circuit C_w of $M^{\mathcal{C}}$. This circuit comes from some $C \in \mathcal{C}(M)$ with $C \not\subset B$. Pick C_w such that $C - B$ contains only one element, say $r \in C$. Let $i \in C_w, i \notin C - B$. Then $B' = B - i + r$. Then B' is a basis of M because the circuit C_w is minimal. But $w_i = \min_{j \in C} w_j < w_r$. Therefore $w(B') = w(B) - w_i + w_r > w(B)$. This contradicts the fact that B is of maximal weight.

□

2.1.1 Some exercises

To get a good sense of how to work with matroids, try the following exercises:

Exercise 2.14. Let G be the graph drawn below. Find the $\mathcal{C}, \mathcal{I}, \mathcal{B}$, and ρ representations of the induced matroid.



Exercise 2.15. Compute the rank function ρ, the independent sets $\mathcal{I}$, and the set of bases $\mathcal{B}$ for the graphical matroid $M(K_4)$ (where K_4 is the complete graph on 4 vertices).

Exercise 2.16. List all possible matroid base polytopes on the set $[3] = \{1, 2, 3\}$. For each polytope, write down the rank function of the corresponding matroid.

2.2 Tropical Linear Spaces and Dressians ❖

Definition 2.17. Let $M = (\{0, ..., n\}, \mathcal{C})$ be a matroid defined by circuits. Then we define:

$$\operatorname{trop}_0(M) := \left\{ w \in \mathbb{R}^{n+1} \;\middle|\; \min_{i \in C} w_i \text{ is achieved at least twice } \forall C \in \mathcal{C} \right\}$$

Exercise 2.18. If M is the matroid associated to a graph G, show that that $\operatorname{trop}_0(M)$ is the set of w such that M_w has no loops (circuits of length 1). This gives a geometric description of the tropicalized variety $\operatorname{trop}_0(M)$.

Lecture 03/12

To motivate tropical linear spaces, suppose we have a graphical matroid $M = (E, B)$ and a function $\omega : B \to \mathbb{R}$. A special case of this is when ω is linear, so that the weight of a tree is the sum of the weights of its edges. This was the case we were considering above. It is not obvious that the set of max-weight spanning trees under ω is a matroid base polytope (i.e. it must satisfy the basis exchange axioms). Given two distinct max-weight trees T, T', for each $(i \to j) \in T$, there must exist $(k \to \ell) \in T'$ such that swapping $(i \to j)$ and $(k \to \ell)$ also gives max-weight trees.

Remark 2.19. Note that any function $\omega : \mathcal{B} \to \mathbb{R}$ determines a regular subdivision Δ_ω on $P_{\mathcal{B}}$ by lifting each vertex $b \in \mathcal{B}$ to $\omega(b)$.

Definition 2.20. Given a matroid $M = ([n], \mathcal{B})$, then the *Dressian* $\mathrm{Dress}(M)$ of M is the set of $\omega : \mathcal{B} \to \mathbb{R}$ such that Δ_ω subdivides $P_\mathcal{B}$ into smaller matroid polytopes (i.e. sub-polytopes of the form $P_{\mathcal{B},w}$). We say that a cell of Δ_ω is *supported* by $p \in \mathbb{R}^n$ if this cell is of the form $\{b \in \mathcal{B} \mid \omega(b) - \langle p, b\rangle = \max_{b' \in \mathcal{B}} \omega(b') - \langle p, b'\rangle\}$.

We can think of a cell of Δ_ω supported by $p \in \mathbb{R}^n$ is the set of spanning trees with maximal "profit" at "edge price" p. Note that when ω is linear, the subdivision Δ_ω is trivial.

Example 2.21. The *uniform matroids* $U_{d+1,n+1}$ are matroids on $n+1$ elements $[n+1] = \{0, ..., n\}$ whose bases are all $d+1$ element subsets of $[n+1]$. For $n = 2$, the matroid base polytopes are given by intersecting the cube with hyperplanes $x_0 + x_1 + x_2 = d+1$. There are three valid such hyperplanes. The circuits are the set of bases strictly above these hyperplanes (i.e. all size $d+2$ or more subsets). One can show that $\mathrm{trop}_0(U_{1,3})$ is a point.

Exercise 2.22. Show that $\mathrm{trop}_0(U_{d+1,n+1})$ is the union of positive orthants spanned by d-subsets of $(e_0, ..., e_n)$.

Lecture 03/14 π

Just as with $\mathrm{trop}_0(M)$, there is an algebraic description of the Dressian:

Theorem 2.23. *For the uniform matroid $U_{2,n}$, the Dressian is:*

$$\mathrm{Dress}(U_{2,n}) = \left\{ W \in \mathbb{R}^{n\times n} \;\middle|\; \begin{array}{c} W_{ij} = W_{ji}, W_{ii} = 0 \\ \forall i,j,k,\ell \;\; \min(W_{ij} + W_{k\ell}, W_{il} + W_{kj}, W_{ik} + W_{jl}) \text{ is achieved at least twice} \end{array} \right\}$$

Remark 2.24. The algebraic geometer will notice that the second condition is a tropical equivalent of Plücker relations. Phylogeneticists call it the four point condition.

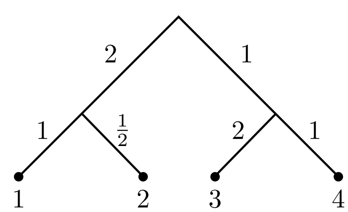
These are related to phylogenetics and tree metrics.

Definition 2.25. We say that $d \in \mathbb{R}^{n\times n}$ is a *tree metric* if there exists an edge-weighted tree T on n leaves with internal edge lengths such that d_{ij} is the sum of all edges in the unique tree-geodesic from i to j.

Example 2.26. Let d be the matrix:

$$d = \begin{pmatrix} 0 & 1.5 & 6 & 5 \\ 1.5 & 0 & 5.5 & 4.5 \\ 6 & 5.5 & 0 & 3 \\ 5 & 4.5 & 3 & 0 \end{pmatrix}$$

A tree that realizes this as a tree metric is:



Theorem 2.27. *$d \in \mathbb{R}^{n\times n}$ is a tree metric on $[n]$ if and only if $-d \in \mathrm{Dress}(U_{2,n})$.*

Definition 2.28. A weighted tree is called *equidistant* if the distance from a leaf to the root is the same for all leaves.

Theorem 2.29. *$d \in \mathbb{R}^{n\times n}$ is an equidistant tree on $[n]$ if and only if $-d \in \mathrm{trop}_0(M_{K_n})$, where K_n is the complete graph on n vertices and M_{K_n} is the graphical matroid associated to this graph. Equivalently d is an equidistant tree on $[n]$ if and only if:*

$$d \in \{w \in \mathbb{R}^{n\times n} \mid w_{ij} = w_{ji} = 0, \text{ and } \forall i,j,k, \; \max d_{ij}, d_{ik}, d_{jk} \text{ is achieved at least twice}\}$$

Lemma 2.30. *If $d \in \mathbb{R}^{n\times n}$ is a tree metric, then there exists $d^* \in \mathbb{R}^{n\times n}$ and $v \in \mathrm{span}\left(\left\{\sum_{j=1}^n e_{ij}\right\}_{i\in[n]}\right)$ such that $d = d^* + v$ and d^* is a balanced tree metric (i.e. its induced tree is equidistant).*

Lecture 03/26

2.3 Dressians and Greedy Algorithms ❖

Recall in our proof of Lemma 2.9, we wanted to show that $Q \subset P_{\mathcal{B}}$. The proof we left as incomplete was actually incorrect. To supplement it, consider the linear program:

$$\max(c \cdot x) \text{ such that } x \in Q$$

for some $c \in \mathbb{R}^n$. We claim that we can solve this program greedily. The greedy approach (shown below) is to think of c_i as the weight of item $i \in [n]$. Then we recursively add max-value items to our basket until we cannot do it anymore.

- **Algorithm 1:**

  1. Start with $S = \emptyset$.
  2. Look at all indices $i \notin S$ and pick i^* with maximal c-value that does not exceed the rank constraint $\langle x, S\rangle \leq \rho(S)$. If no such index exists, return x^* where $x_i^* = 1$ if and only if $i \in S$.
  3. Then update: $S \leftarrow S \cup \{i^*\}$.
  4. Return to 1).

We note that the vector obtained by solving this linear program is contained in $P_{\mathcal{B}}$ and is a vertex of Q. Therefore the vertices of Q are contained in $P_{\mathcal{B}}$, which implies $Q \subset P_{\mathcal{B}}$. This completes the proof of Lemma 2.9.

Remark 2.31. We can reorder the items in $[n]$ with decreasing c values. Let π be this sorting permutation, so that $c_{\pi(i)} \geq c_{\pi(j)}$ when $i < j$. Let $S_i = \{\pi(1)..., \pi(i)\}$. Then the above greedy algorithm results in:

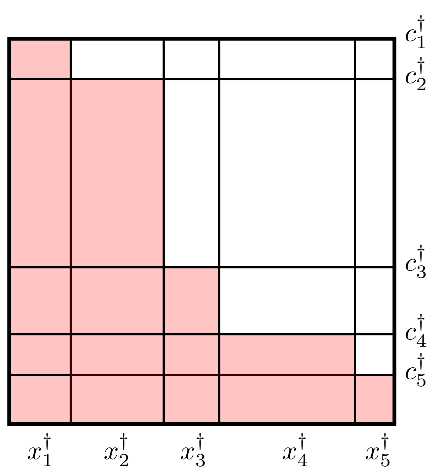
$$x_i^* = \rho(S_i) - \rho(S_{i-1}) = \begin{cases} 1 & \text{adding } \pi(i) \text{ into } S_{i-1} \text{ will increase rank.} \\ 0 & \text{otherwise} \end{cases}$$

Proof of Algorithm 1:

For any $x \in [0,1]^n$, we have:

$$\sum_{i=1}^{n} c_i x_i = \left(\sum_{i=1}^{n-1} (c_i^\dagger - c_{i+1}^\dagger)\langle x, S_i\rangle\right) + c_n^\dagger \langle x, S_n\rangle$$

where $S_i = \{\pi(1), ..., \pi(i)\}$ as in the remark and $c_i^\dagger = c_{\pi(i)}$. This equality can be seen by rearranging the Riemann sum below into a Lebesgue sum:



For $x \in Q$, then we have $\langle x, S_i \rangle \leq \rho(S_i)$ and hence:

$$\begin{aligned}\sum_i c_i x_i &= \left(\sum_i (c_i^\dagger - c_{i+1}^\dagger)\langle x, S_i\rangle\right) + c_n^\dagger \langle x, S_n\rangle \\ &\leq \left(\sum_i (c_i^\dagger - c_{i+1}^\dagger)\rho(S_i)\right) + c_n^\dagger \rho([n]) \\ &= \left(\sum_i c_i^\dagger(\rho(S_i) - \rho(S_{i+1}))\right) + c_n^\dagger \rho([n]) = c \cdot x^*\end{aligned}$$

□

Therefore, we have shown that the linear program:

$$\begin{gathered}\max(c^T x)\\ x \in P_{\mathcal{B}}\end{gathered}$$

can be solved by a greedy algorithm. A similar proof shows that the program:

$$\begin{gathered}\max(c^T x)\\ x \in P_{\mathcal{I}} := \mathrm{Conv}\{S : S \in \mathcal{I}\} = \{x \in [0,1]^n : \langle x, S\rangle \leq \rho(S)\ \forall S\}\end{gathered}$$

can also be solved with a greedy algorithm.

Theorem 2.32. *Consider the linear program:*

$$\begin{gathered}\max(c^T x)\\ x \in P_{\mathcal{S}} = \text{ convex hull of a subset } \mathcal{S} \text{ of vertices of } [0,1]^n.\end{gathered} \tag{P_0}$$

Then the greedy algorithm solves (P_0) *for all* $c \in \mathbb{R}^n$ *if and only if* $P_{\mathcal{S}}$ *is an independent set polytope for some matroid.*

Corollary 2.33. *Fix* $\eta : \{0,1\}^n \to \mathbb{R}$. *Consider the optimization problem:*

$$\begin{gathered}\max(\eta(x) - c^T x)\\ x \in P_{\mathcal{S}}\end{gathered}$$

Then the greedy algorithm solves this problem for all $c \in \mathbb{R}^n$ *if and only if* $P_{\mathcal{S}} = P_{\mathcal{I}}$, *where* $\mathcal{I}$ *is the independent set of some matroid and* η *induces a regular subdivision of* $P_{\mathcal{I}}$ *into more matroid independence polytopes.*

A special case of the above corollary is with the program:

$$\begin{gathered}\max(\eta(x) - c^T x)\\ x \in P_{\mathcal{B}}\end{gathered}$$

The greedy algorithm will solve this for all $c \in \mathbb{R}^n$ if and only if $\eta \in \mathrm{Dress}([n], \mathcal{B})$.

Example 2.34. Given a function $v : \{0,1\}^n \to \mathbb{R}$ and a cost vector $c \in \mathbb{R}^n$, we can think of this as describing a team of n workers, where c_i is the salary of worker i an $v(S)$ is the productivity value of a team of workers $S \subset [n]$. Suppose the goal is to maximize $v(S) - \sum_{i \in S} c_i$. We would like to use a greedy approach.

1. A *unit demand function* is a function of the form $v(S) = \max_{i \in S} v(\{i\})$.
2. An *additive demand function* is one of the form $v(S) = \sum_{i \in S} v(\{i\})$.

Do either of these types of functions induce a matroidal subdivision of $[0,1]^n$? It is not too hard to show that additive demand functions induce a trivial subdivision Δ_v. Therefore the greedy algorithm will work to solve our maximization problem. For unit demand functions, it is not quite as easy to show. Consider a $(n,1)$ bipartite graph with edge weights $w_1, ..., w_n$. Then v can be realized as matching function for this bipartite graph (defined below). This is in fact sufficient for Δ_v to be matroidal.

Definition 2.35. Let $A \in \mathbb{R}^{m\times n}$ be a matrix (representing the adjacency of a bipartite graph). Define a function $v : \{0,1\}^n \to \mathbb{R}$ by identifying $\{0,1\}^n$ with the power set of columns of A and setting $v(J) = |A_J|_{\min}$, where $|\cdot|_{\min}$ is the tropical determinant of a square matrix:

$$|A|_{\min} := \min_{\sigma \in S_n} \left(\sum_{i=1}^{n} A_{i\sigma(i)} \right)$$

This is called a *matching function* for the associated bipartite graph. It gives a minimal matching of the J left hand nodes to the right hand nodes.

Lecture 04/02

Today, we will describe an alternative characterization to the optimality of the greedy algorithm involving discrete derivatives.

Definition 2.36. Given $v : \{0,1\}^n \to \mathbb{R}$ and $i \in [n]$, we define the derivative function $\partial_i v : \{0,1\}^{n-\{i\}} \to \mathbb{R}$ by:

$$\partial_i v(S) := v(S+i) - v(S)$$

Here, $\{0,1\}^{n-\{i\}}$ denotes the vertices of the cube $\{1,0\}^n$ such that their ith coordinate is zero. The secoind derivative $\partial_{ij} : \{1,0\}^{n-\{i,j\}} \to \mathbb{R}$ is defined by:

$$\partial_{ij} v(S) := v(S+i+j) + v(S) - (v(S+j) + v(S+i))$$

Remark 2.37. A function $v : \{0,1\}^n \to \mathbb{R}$ is submodular if and only if $\partial_{ij} v(S) \le 0$ for all $S \subset [n]$ and $i, j \notin S$.

Theorem 2.38. *Let $v : \{0,1\}^n \to \mathbb{R}$. Then:*

$$\Delta_v \text{ is matroidal} \iff \begin{array}{c} \text{For each given } S \subset [n] \text{ and } i,j,k \notin S, \\ \max\left(\partial_{ij}v(S), \partial_{ik}v(S), \partial_{jk}v(S)\right) \text{ is achieved as least twice and is} \le 0 \end{array}$$

This theorem implies that for fixed S, the matrix d^S defined by $d^S_{ij} = \partial_{ij} v(S)$ is an ultrametric with nonpositive weights. An open problem is: can we easily determine the subdivision Δ_v from the associated ultrametric d^S for some S?

Remark 2.39. In 2005, Hatfield and Milgrom conjectured that Δ_v is matroidal if and only if v came from a bipartite matching (the tropical determinant of some bibpartite matching c.f. Definition 2.35). We saw that the unit demand function satisfied this property in Example 2.34.

Exercise 2.40. Disprove this conjecture (this will require no more than two lines using the above ideas).

Lecture 04/04

2.3.1 Economics in Tropical language

It is not clear why the Hatfield-Milgrom conjecture should be false. To get an intuition for why it could be true, we take a short detour into economics. An *economy* consists of: m agents $j \in [m]$and n types of goods and one good of each type. Each agent has a valuation $v^j : \{0,1\}^n \to \mathbb{R}_+ \cup \{-\infty\}$ which represents the value of a set of goods to agent $j \in [m]$. The *utility agent* of j is the polynomial:

$$f_{v^j}(p) := \bigoplus_{a \in \{0,1\}^n} v^j(a) \odot (-p)^{\odot a} = \max_{a \in \{0,1\}^n} (v^j(a) - \langle p, a \rangle)$$

Given p, the *demand set* $D_{v^j}(p)$ of agent j is the set of $a^* \in \{0,1\}^n$ such that the maximum is achieved. The *aggregated utility* is the polynomial:

$$f_v(p) := f_{v^1}(p) \odot \ldots \odot f_{v^m}(p) = \bigoplus_{a \in \{0,1,\ldots,n\}^n} V(a) \odot (-p)^{\odot a}$$

The coefficient $V(a)$ can be worked out to be:

$$V(a) = \bigoplus_{\substack{a^1,\ldots,a^m \\ a^1 \odot \ldots \odot a^n = n}} \left[v^1(a^1) \odot \ldots \odot v^m(a^m)\right] = \max\left(v^1(a^1) + \ldots + v^m(a^m)\right)$$

where the max is taken over the same set as the $\bigoplus$. The aggregated demand set is the demand set of the aggregated utility:

$$D_v(p) = \{a^* \in \{0,1,\ldots,n\}^n \mid V(a^*) - \langle p, a^* \rangle = f_v(p)\}$$

In economics terms, $V(a)$ is the max possible valuation to partition a into $a^1, .., a^n$ and assign to agent j. An issue is that there is only one copy of each good. Let $\mathbf{1} = (1, \ldots, 1) \in \mathbb{R}^n$. An economy is said to have a *Walrasian equilibrium* if there exists $p \in \mathbb{R}^n$ such that $\mathbf{1} \in D_v(p)$.

Example 2.41. Let $n = 2$. Agent 1 has valuation $v^1(10) = v^1(11) = -\infty, v^1(00) = 0$ and $v^1(01) = q$. The subdivision Δ_{v^1} of the square $\{0,1\}^2$ is just a vertical line. Let agent 2 have valuation v^2 that is the same as v^1 except $v^2(10) = q'$ and $v^2(01) = -\infty$. The subdivision Δ_{v^2} is the horizontal line in the cube. The tropical variety $T_{f_{v^1}}$ is a horizontal line and $T_{f_{v^2}}$ is vertical. Therefore $T_{f_{v^1} \odot f_{v^2}}$ is the union of these lines, and hence $\Delta_{v^1} + \Delta_{v^2}$ is the trivial subdivision of the square. In particular, 11 is a mixed vertex of this subdivision and hence a Walrasian equilibrium exists.

Lecture 04/09

A question we might ask is for which sets of $\{v^i\}$ does a Walrasian equilibrium exist? There isn't a complete answer to this, but some special cases are known. In 1970, Kelso and Crawford asked this question in the case where the allocation $(a_1, \ldots, a_m)$ can be computed using a greedy procedure.

Definition 2.42. The *greedy allocation* starts with a price $p \in \mathbb{R}^n$ and computing $D_v(p)$ (the aggregated demand set at price p). We check if $\mathbf{1} \in D_v(p)$, and if not, look at what each agent demands and draw it as a bipartite graph between $[m]$ and $[n]$. An edge between agent j and a number i exists if and only if the agent demands item i at current price p. If an item has more than one incident edge, then increase the price component p_i by ϵ; if it has no incident edges, then decrease the price component p_i by ϵ. Then continue from the beginning.

Definition 2.43. A function $v : \{0,1\}^n \to \mathbb{R}$ is a *gross substitute* if for all $p \in \mathbb{R}^n$, if $p' \leq p$ (coordinate wise), then there exists $S' \in D_v(p')$ such that:

$$S \cap \{j : p_j = p'_j\} \subseteq S'$$

where $S = D_v(p)$.

The intersection $S \cap \{j : p_j = p'_j\}$ above is the set of items whose price do not go up, and this being a subset of S' says that these items are still demanded after a price increase.

Theorem 2.44. *If $v^1, \ldots, v^n$ are gross substitute functions, then the output of the greedy allocation c.f. Definition 2.42 is an optimal allocation (i.e. satisfies Walrasian equilibrium).*

This theorem becomes useful for us with the following result:

Theorem 2.45. *A function $v : \{0,1\}^n \to \mathbb{R}$ is a gross substitute if and only if Δ_v is matroidal.*

2.3.2 Valuated Matroids

Consider the matroid $U_{n,d}$, the uniform matroid on $[n]$ of rank d. The bases $\mathcal{B}$ are denoted $\binom{[n]}{d}$. A function $\omega : \mathcal{B} \to \mathbb{R}$ is a valuated matroid if $\omega \in \mathrm{Dress}(U_{n,d})$.

Theorem 2.46. *A function $v : \{0,1\}^n \to \mathbb{R}$ is a gross substitute if and only if $\omega : \binom{[2n]}{n} \to \mathbb{R}$ defined by $\omega(S) = v(S \cap \{1, \ldots, n\})$ is in the dressian $\mathrm{Dress}(U_{2n,n})$.*

The following theorem shows that the Hatfield-Milgrom conjecture is false.

Theorem 2.47. *For $n \geq 3$, there exists a gross substitute $v : \{0,1\}^n \to \mathbb{R}$ but v is not a bipartite matching function.*

Proof:

Let v be a bipartite matching function; then there exists a matrix $A \in \mathbb{R}^{n\times d}$ such that $v(J) = |A_J|$. Consider ω defined from v via Theorem 2.46. We claim that ω lies in the Stiefel linear space of $U_{2n,n}$. To see this, introduce n new dummy edges on the left side of the bipartite graph associated to A and $n-d$ on the right side. The new nodes are fully connected with edge weights 0. This gives us a $(2n, n)$ bipartite graph G. Now, a max perfect matching of G consists of assigning each $S \in \binom{[2n]}{n}$ to $|G_S|$, the max matching of "real nodes" in S (i.e. $S \cap \{1, ..., n\}$) and "real nodes" on the right hand side (the first d notes). This is exactly $v(S \cap \{1, ..., n\})$. Therefore $\omega(S) = |G_S|$, so by definition ω lies in the Stiefel tropical linear space.

Now all we need to show is that the Stiefel tropical linear space of $U_{2n,n}$ is a strict subset of $\mathrm{Dress}(U_{2n,n})$. Since the tropical grassmannian $\mathrm{Gr}(U_{2n,n})$ is known to be a strict subset of $\mathrm{Dress}(U_{2n,n})$ and contains the Stiefel tropical linear space in question, we are done.

□

Exercise 2.48. For $n = 3$, explicitly exhibit an element of the tropical grassmannian of $U_{6,3}$ which is not in $\mathrm{Dress}(U_{6,3})$.

2.3.3 How to compute the Walrasian price

Returning to Walrasian equilibrium, finding the associated price p will require solving an integer program:

$$\max\left(\sum_{i=1}^{m}\sum_{S\subseteq[n]} x_{iS}\cdot v^i(S)\right) \tag{W_{IP}}$$

such that
- $\displaystyle\sum_{\substack{S \subseteq [n] \\ j \in S}}\sum_{i=1}^{m} x_{iS} = 1 \quad \forall j \in [n]$
- $\displaystyle\sum_{S\subseteq[n]} x_{iS} = 1 \quad \forall i \in [m]$
- $x_{iS} \in \{0,1\} \quad \forall i \in [m], S \subseteq [n]$

The first constraint is that each item is only assigned to one agent and the second is that each agent gets something. The third is the integer constraint that items cannot be split. This makes it an integer linear program. A solution $(x_{iS})_{i\in[n]}$ is a binary vector, where $x_{iS} = 1$ implies that agent i recieves subset S of items. The relaxed program (W_{LP}) exchanges the third constraint with $x_{iS} \in [0, 1]$.

Theorem 2.49. *The solution to (W_{IP}) is the solution to (W_{LP}) if and only if a Walrasian equilibrium exists.*

Lecture 04/11

To actually find the Walrasian price, we consider the dual program:

$$\min\left(\sum_{i=1}^{m} u_i + \sum_{j=1}^{n} p_j\right) \tag{DW_{LP}}$$

$$\text{such that } u_i \geq v^i(S) - \sum_{j\in S} p_j \quad \forall i = 1, ..., m \text{ and } S \subseteq [n]$$
$$p_j \geq 0 \quad \forall j = 1, ..., n$$
$$u_i \geq 0 \quad \forall i = 1, ..., m$$

The variables are the $n + m$ quantities u_i, p_j. The following lemma follows from duality of linear programs:

Lemma 2.50. *If (p, u) is optimial in (DW_{LP}), then p is a Walrasian price.*

Proposition 2.51. *The set of Walrasian prices is a tropical polytope.*

Proof:

We will show that checking if an assignment (x_{iS}) (also written $i \to S_i$) is optimal is equivalent to computing the min-plus tropical eigenvalue of a certain graph. For $v^1, ..., v^m : \{0,1\}^n \to \mathbb{R}$, define $w^1, ..., w^m : \binom{[2n]}{n} \to \mathbb{R}$ by:

$$w^i(S) = v^i(S \cap [n])$$

Each assignment $i \to S_i$ is trivially extended to $S'_i \subset \binom{[2n]}{n}$ via $S'_i \cap [n] = S_i$. By abuse of notation we write S'_i as S_i. Now we define a weighted, directed, graph G on $2n$ nodes. The weight of the edge $(j \to k)$ is $w^i(S_i) - w^i(S_i - j + k)$ where S_i is the unique set containing j. We think of G as the adjacency matrix.

Claim: If $q \in \mathbb{R}^{2n}$ satisfies $G \underline{\odot} q = q$, then $(q_1, ..., q_n)$ is a Walrasian price and $(q_{n+1}, ..., q_{2n}) = (0, ..., 0)$.

Proof: Let q be in the tropical eigenspace $\mathrm{Eig}(G)$. Then for each $j \in [n]$ we have $\min_k G_{jk} + q_k = q_j$. In particular, for all $k = 1, ..., 2n$ we have:

$$G_{jk} + q_k \geq q_j \implies w^i(S_i) - q_j \geq w^i(S_i - j + k) - q_k \tag{2.3.1}$$

For $S' \neq S_i$ and $S' \in \binom{[2n]}{n}$ then by the basis exchange algorithm, there exists a sequence of swaps $j_1 \to k_1, ..., j_r \to k_r$ with $j_1, ..., j_r \in S_i$ and $k_1, ..., k_r \in S' \setminus S_i$ taking us from S to S'. We can choose this sequence such that:

$$v(S) - v(S') = (v(S) - v(S - j_1 + k_1)) + ... + (v(S) - v(S - j_r + k_r)) \tag{2.3.2}$$

$$= rv(S) - \sum_{\ell=1}^{r} v(S - j_\ell + k_\ell) \tag{2.3.3}$$

This is an "independence property" of these swaps. Such a sequence of swaps can be found by associating it to a minimal weight bipartite matching of a graph. Now we sum equation (2.3.1) over all such swaps and use (2.3.2):

$$\sum_{\ell=1}^{r} w^i(S_i) - q_{j_\ell} \geq \sum_{\ell=1}^{r} w^i(S_i - j_\ell + k_\ell)$$

$$\implies rw^i(S_i) - \sum_{\ell=1}^{r} q_{j_\ell} \geq (r-1)w(S) + w(S') - \sum_{\ell=1}^{r} q_{k_\ell}$$

$$\implies w^i(S_i) - \sum_{\ell=1}^{r} q_{j_\ell} \geq w(S') - \sum_{\ell=1}^{r} q_{k_\ell}$$

This shows that S_i is an optimal assignment for agent i under price q. The only thing left to check is that the dummy items have price zero. This can be checked by using the equation $G \underline{\odot} q = q$. □

Conversely, we also claim:

1. If $\lambda(G) \neq 0$, then $S_1, ..., S_m$ is not optimal.
2. If $\lambda(G) = 0$, then any $q \in \mathrm{Eig}(G)$ satisfies, for $k = 1, ..., n$, $p_k := q_k - \min_{j \in [n]} q_j$ is a Walrasian price then $S_1, ..., S_m$ is optimal.

These claims follow easily by reversing the above argument. Therefore $(S_1, ..., S_m)$ is optimal if and only if $\lambda(G) = 0$ and the set of Walrasian prices is a polytope (in fact, $\mathrm{Eig}(G)$).

□

3. Tropical Algebraic Geometry

There is a way of associating to any affine variety $X \subset \mathbb{A}^n$ a *tropical variety* $\mathrm{trop}(X)$ through the process of tropicalizing its defining polynomials. Namely, if $f \in k[x_1, ..., x_n]$ is a multivariate polynomial, then let $\hat{f}$ denote its tropicalization (replacing all sums and products by tropical sums and products). Then the hypersurface $V(\hat{f})$ c.f Definition 1.44 is a subset of tropical n-space. Thus we define:

$$\mathrm{trop}(X) := \bigcap_{f \in I(X)} V(\hat{f})$$

Here $I(X)$ denotes the vanishing ideal of X. Strictly speaking, there isn't a scheme-theoretic interpretation of $\mathrm{trop}(X)$; it is simply a subset of $\mathbb{R}^n_\infty$. Thus, even though we call it a tropical algebraic *variety*, it isn't a variety in a rigorous sense of the word.

3.1 The Tropical Grassmannian

3.2 Stiefel Tropical Linear Spaces

This subsection is based on a lecture given by Austin Alderete on 3/28/2019.[1] Above we defined the tropical grassmannian $\mathcal{G}\mathrm{r}_d(n)$ as the tropicalization of the classical grassmannian $\mathrm{Gr}_d(n)$. It is a well-known result that the classical Stiefel map:

$$\pi : \mathbb{R}_0^{n \times d} \to \mathrm{Gr}_d(n)$$

sending a full-rank matrix $M \in \mathbb{R}_0^{n\times d}$ to its column space is a surjection; in other words, every subspace has a basis. While this statement might be trivial for the classical Stiefel map, it is non-trivial for the tropicalized version and, in fact, it is false.

Definition 3.1. Let G be a bipartite graph on $[n] \sqcup \mathcal{A}$ for some finite set $\mathcal{A}$. Recall that a *matching* is a subset of edges of G such that each edge is adjacent to at most one element of $\mathcal{A}$. It is called *maximal* if the number of edges in the matching is $\min(n, |\mathcal{A}|)$. See Figure 3.1.

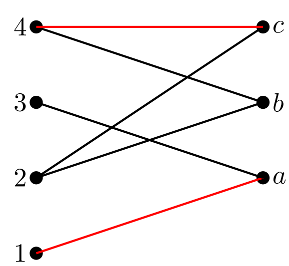


Figure 3.1: A non-maximal matching (red) on the a bipartite graph on $[4] \sqcup \{a, b, c\}$.

Given a graph as above, we can define a matroid $\mathcal{M}[\mathcal{A}]$ as follows. The base set E is $[n]$, the left hand side vertices, and the independent sets $\mathcal{I}(\mathcal{M}[\mathcal{A}])$ are the sets of left-hand nodes in matchings on G. Consequently, the bases are the maximal matchings on the graph. We call any matroid *transversal* if it arises as the matroid of some bipartite graph in this fashion.

Exercise 3.2. Show that two non-isomorphic bipartite graphs can produce the same matroid in the above construction.

[1]A copy of his lecture slides can be found at `web.ma.utexas.edu/users/aaldere`

An equivalent way to encode a bipartite graph is to give an $n \times m$ adjacency matrix; therefore this matrix also represents a transversal matroid. The matrix of the transversal matroid from Figure 3.1 is $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \end{pmatrix}$.

Appendix

A.1 Solutions to Selected Exercises ❖

Exercise 1.2: The polynomial $x^2 + 15x + 2$ in classical notation is $\max(2x, 15 + x, 2)$. There are two "roots" of this polynomial, which are the discontinuity points of $f'(x)$. These happen at $x = -13$ and $x = 15$. Therefore we claim that $(-13 \oplus x) \odot (15 \oplus x)$ is a factorization of $x^2 + 15x + 2$. We can verify this by expanding:

$$\begin{aligned}(-13 \oplus x) \odot (15 \oplus x) &= -13 \odot 15 \oplus -13 \odot x \oplus 15 \odot x \oplus x^{\odot 2} \\ &= 2 \oplus (-13 \oplus 15) \odot x \oplus x^{\odot 2} \\ &= 2 \oplus 15 \odot x \oplus x^{\odot 2}\end{aligned}$$

Therefore this is a factorization. Question: is factorization unique? This will be answered in the next exercise.

Exercise 1.4: We define an equivalence relation on the set of all tropical polynomials by $f \sim g$ if and only if $f(x) = g(x)$ for all $x \in \mathbb{R}$. In other words, we consider two polynomials equivalent if their graphs coincide. We will show that for any polynomial f, there is a polynomial in its equivalence class which is expressable as a tropical product of linear terms. Given $f(x) = a_0 \oplus a_1 x \oplus a_2 x^2 \oplus ... \oplus a_n x^n$, we define the ith *resolvent* to be $r_i = a_{i-1} - a_i$ for $1 \leq i \leq n$. We say that f is *resolved* if $r_1 \leq r_2 \leq ... \leq r_n$. The fundamental theorem of tropical algebra will follow from a few lemmata.

Lemma A.1. *Any tropical polynomial f is equivalent to a resolved tropical polynomial.*

Proof:

Since the graph of f is piecewise linear, there are a finite number of points where $f'(x)$ is not continuous. If d is one such point, then we let the *multiplicity* of this point be:

$$\lim_{x \to d^+} f'(x) - \lim_{x \to d^-} f'(x)$$

This is a positive integer because $f'(x)$ is nondecreasing and integer valued. The number of discontinuities counting multiplicity is $n = \deg(f)$. We denote these points in nondecreasing order and with multiplicity by $\{d_i\}$ for $1 \leq i \leq n$. Let $c_0 = \lim_{x \to -\infty} f(x)$ and define $c_i = c_0 - d_1 - d_2 - ... - d_i$ for $i > 0$. Then by construction $c_{i-1} - c_i = d_i$ and hence the polynomial $g(x) = \bigoplus_{i=0}^n c_i \odot x^i$ is resolved. We claim that $f \sim g$. The discontinuities of $g'(x)$ happen when the lines $y = c_i + ix$ and $y = c_{i+1} + (i+1)x$ intersect for every $0 \leq i < n$:

$$c_i + ix = c_{i+1} + (i+1)x \iff x = c_i - c_{i+1} = d_{i+1}$$

Thereore the graphs of $f'(x)$ and $g'(x)$ have the same points of discontinuity. They also have the same multiplicities. Since the graphs of g and f have the same constant term (i.e. c_0) and the same jumps with multiplicity, they must be the same graph.

□

Lemma A.2. *Any resolved tropical polynomial f can be factored into linear terms.*

Proof:

Let $f(x) = \bigoplus_{i=0}^n a_i \odot x^i$ be resolved. As we saw in the above proof, the points of discontinuity of $f'(x)$ were the resolvents r_i (roots of f in this case). Now we claim that the following is a linear factorization of f:

$$a_n(x \oplus r_1)(x \oplus r_2)(x \oplus r_3)...(x \oplus r_n)$$

> To see this, we expand it out:
>
> $$a_n(x \oplus r_1)(x \oplus r_2)...(x \oplus r_n) = \bigoplus_{i=0}^{n} a_n \odot e_i(r_1, ..., r_n) \odot x^i$$
>
> where $e_i(r_1, ..., r_n)$ is the ith elementary symmetric polynomial in the r_j. Since f is resolved, these polynomials are easily computed:
>
> $$\begin{aligned} e_n &= 0 \\ e_{n-1} &= r_1 \oplus r_2 \oplus ... \oplus r_n = r_n = a_{n-1} - a_n \\ e_{n-2} &= (r_1 \odot r_2) \oplus (r_2 \odot r_3) \oplus ... \oplus (r_{n-1} \odot r_n) = r_{n-1} \odot r_n = a_{n-2} - a_n \\ &\vdots \\ e_0 &= r_1 \odot r_2 \odot ... \odot r_n = a_0 - a_n \end{aligned}$$
>
> In other words, $e_i = a_i - a_n$. Therefore:
>
> $$a_n(x \oplus r_1)(x \oplus r_2)...(x \oplus r_n) = \bigoplus_{i=0}^{n} a_n \odot (a_i - a_n) \odot x^i = \bigoplus_{i=0}^{n} a_i \odot x^i = f(x)$$

□

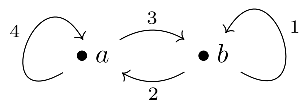
Putting these together, every tropical polynomial is equivalent to a polynomial that factors into linear parts.

As a supplement to this solution, below is an $\mathcal{O}(n^2)$ algorithm that computes the roots of an arbitrary tropical polynomial. The language is Python and the input is the list of coefficents $\{a_0, ..., a_n\}$ of f as floats. It returns a list of roots and their respective multiplicities.

```python
def rootfind(coeffs):
    roots=[]
    n=len(coeffs)
    i=0

    while i<(n-1):
        #list of possible roots at i
        xj=[(coeffs[i]-coeffs[j])/(j-i) for j in range(i+1,n)]

        #root at i
        d=min(xj)

        #highest index achieving d in xj
        xj.reverse()
        k=len(xj)-xj.index(d)-1

        #multiplicity of d
        k+=1
        i+=k

        roots.append([d,k])

    return roots
```

Listing 1: Tropical root finding in python.

Exercise 1.18: We will find λ both algebraically and using mean weighted cycles. The latter is easier to start with. The graph of C is:



If σ_a is the cycle $a \to a$ and σ_b is the cycle $b \to b$, then all cycles starting at a will be of the form $\sigma_a^{n_1} \circ (a \to b) \circ \sigma_b^{n_2} \circ (b \to a) \circ ... \circ \sigma_a^{n_e}$ (and similarly with all cycles starting from b). This has weight $4n_1 + 3 + n_2 + 2 + 4n_3 + ... + 4n_e$ and length $2(e-1) + \sum_{i=1}^{e} n_i$. By inspection, the mean weight of this cycle is always less than 4. Therefore the highest possible mean weight is exactly 4, which is achieved by the cycle σ_a. A similar analysis will show that all cycles starting from b must have mean weight strictly less than 4. Therefore $\lambda(C) = 4$.

Algebraically, the eigenvalue equation is:

$$\max(x_1 + 1, x_2 + 2) = \lambda + x_1$$

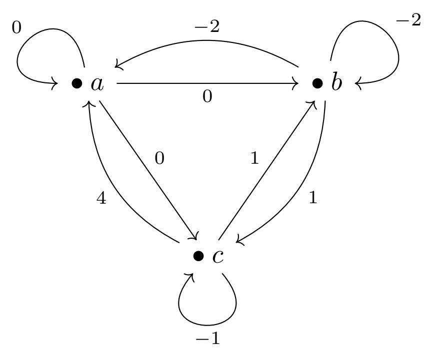
$$\max(x_1 + 3, x_2 + 4) = \lambda + x_2$$

We assume that (x_1, x_2) is a solution with $x_1 > x_2 + 1$. Then the first equation dictates that $x_1 + 1 = \lambda + x_1 \Rightarrow \lambda = 1$. Since $x_1 > x_2 + 1$, we have $x_1 + 3 < x_2 + 4$ and so the second equation tells us $x_1 + 3 = 1 + x_2 \Rightarrow x_2 = x_1 + 2$. This contradicts our assumption that $x_1 > x_2 + 1$, therefore we must have $x_1 \leq x_2 + 1$. Then the second equation tells us $x_2 + 4 = \lambda + x_2 \Rightarrow \lambda = 4$. Then the first implies that $x_2 = 2 + x_1$. Therefore $\lambda = 4$ and the eigenvectors are $(x_1, 2 + x_1)$ for any $x_1 \in \mathbb{R}_\infty$.

Remark A.3. Doing the same exercise in min-plus, we are looking for the smallest mean weight cycle. The solution is $\lambda = 1$, using the same graph analsyis. Now define $\bar{C}_{ij} = C_{ij} - \lambda$. We can get two inequalities $x_1 - x_2 \leq 1$ and $x_2 - x_1 \leq 2$ (with equality achieved in one of them). One can plot the two graphs in $\mathbb{R}^2$ and see that the solutions are $(0, 2) + c(1, 1)$ and $(-1, 0) + c(1, 1)$. This is a specific case of the process given to us by Theorem 1.25.

Now we do the same for the 3×3 matrix. The matrix and graph of C are:

$$C = \begin{pmatrix} 0 & 0 & 0 \\ -2 & -2 & 1 \\ 4 & 1 & -1 \end{pmatrix}$$



We begin by considering the cycle $(a \to c \to a)$. This has weight 4 and mean weight 2. We claim that this is the cycle with the maximal mean weight, since all other edges have weight less than or equal to 1. Therefore $\lambda(C) = 2$. To find the eigenvectors, we write out the eigenvalue equation:

$$\max(x_1, x_2, x_3) = 2 + x_1$$

$$\max(x_1 - 2, x_2 - 2, x_3 + 1) = 2 + x_2$$

$$\max(x_1 + 4, x_2 + 1, x_3 - 1) = 2 + x_3$$

The first equation tells us that $x_1 < x_2, x_3$, which simplifies the second equation to $\max(x_2 - 2, x_3 + 1) = 2 + x_2$. We first assume that $x_2 > x_3$, which implies by the first equation that $x_2 = x_1 + 2$. Then the second equation is $\max(x_1, x_3 + 1) = x_1 + 4$. If $x_1 < x_3 + 1$, then this implies that $x_3 + 1 = x_2 + 2 = x_1 + 3$. This contradicts our first assumption that $x_2 > x_3$. Therefore $x_1 \geq x_3 + 1$. The second equation then implies $x_1 = x_1 + 4$, a contradiction. Therefore our initial assumption that $x_2 > x_3$ was false.

We have now determined that $x_3 \geq x_2$, from which the first equation says $x_3 = x_1 + 2$. The second equation is then $\max(x_2 - 2, x_1 + 3) = x_2 + 2$. If we assume that $x_2 - 2 > x_1 + 3$, then we have $x_2 - 2 = x_2 + 2$, which is a contradiction. Therefore $x_2 - 2 \leq x_1 + 3$ and hence $x_1 + 3 = 2 + x_2$. To check that this satisfies the third equation, we substitute $x_3 = x_1 + 2$ and $x_2 = x_1 + 1$:

$$\max(x_1 + 4, x_1 + 2, x_1 + 1) = x_1 + 4$$

which is indeed true. Therefore the eigenvectors are $(x_1, x_1 + 1, x_1 + 2)$.

Exercise 1.27: For the 3 by 3 matrix C in Exercise 1.18, the eigenvalue was $\lambda = 2$, where the critical cycle was $(a \to c \to a)$. Therefore $\Lambda = \{a, c\}$. We can compute C^+ either by hand or by using the graph of $\overline{C} = C - \lambda$. It is:

$$C^+ = \begin{pmatrix} 0 & -2 & -2 \\ 1 & -1 & -1 \\ 2 & 0 & 0 \end{pmatrix}$$

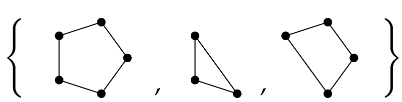
Therefore columns 1 and 3 are eigenvectors (call them v_1, v_2) and the eigenspace is $\mathrm{Eig}(C) = \mathrm{Eig}(C^+) = \mathrm{Conv}(v_1, v_2)$. Since v_1 is a tropical multiple of v_2, this is the space spanned by just v_1. Indeed, this is the same set of eigenvectors that we found by bashing inequalities.

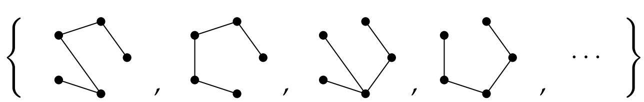
If we do the same for the 2 by 2 matrix, we find $\Lambda = \{a\}$ and:

$$C^+ = \begin{pmatrix} 0 & -1 \\ -2 & -3 \end{pmatrix}$$

and indeed the first column is the eigenvector we found before.

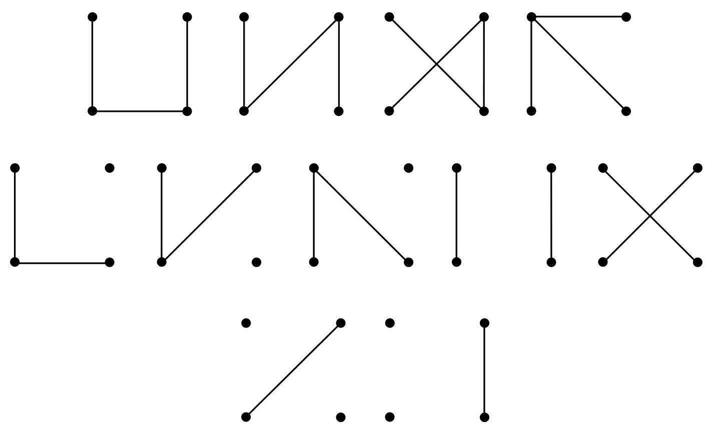
Exercise 2.14: The circuits of G, which are the minimal cycles, and the bases, which are the spanning trees, are easily seen to be:

$\mathcal{C} =$ 

$\mathcal{B} =$ 

The rank function is defined as usual: $\rho(A) = \max_{B \in \mathcal{B}} |A \cap B|$.

Exercise 2.15: We represent K_4 using four vertices arranged in a square. We can rotate any subtree of K_4 by 90 degrees to get a new subtree. This gives us a $\mathbb{Z}/4\mathbb{Z}$ action on $\mathcal{B}$ and $\mathcal{I}$. Below are orbit representatives of $\mathcal{B}$ and $\mathcal{I}$ (the top row being the representatives of $\mathcal{B}$).



In total, there are 16 elements of $\mathcal{B}$ and 37 elements of $\mathcal{I}$. The rank function is then defined to be $\rho(A) = \max_{B\in\mathcal{B}} |B \cap A|$, where A is a subset of edges of K_4.

Exercise 2.16: We identify subsets of [3] with binary strings of length 3 (which lie on the corners of the unit cube in $\mathbb{R}^3$). The matroid base polytopes of the uniform matroids $U_{i,3}$ for $0 \le i \le 3$ are all intersections of the cube with the hyperplanes $x_1 + x_2 + x_3 = i$. Geometrically, these are two triangles and two points. The remaining matroid base polytopes come from taking subsets of the corners of the uniform base polytopes. That is, if $P_{U_{i,3}}$ is a uniform matroid base polytope with corner set C, then we get another family of matroid base polytopes $\{\mathrm{Conv}(S)\}_{S\subset C}$. All matroid base polytopes on [3] are of this form.

We will write the rank functions on each matroid $([3], \mathcal{B})$ using binary AND/OR notation: if $a, b \in \{0, 1\}$, then we define $a \wedge b$ to be the OR of the bits and $a \vee b$ to be the AND of the bits. A rank function is a function on the cube $\{0, 1\}^3$ (i.e. a function of three variables x,y, and z) and is always of the form $\rho(x, y, z) = \max_{b\in\mathcal{B}} |b \cap (xyz)|$.

Basis of Matroid	Rank function $\rho(x, y, z)$
$\mathcal{B} = \{010, 001\}$	$y \wedge z$
$\mathcal{B} = \{100, 001\}$	$x \wedge z$
$\mathcal{B} = \{100, 010\}$	$x \wedge y$
$\mathcal{B} = \{100, 010, 001\}$	$x \wedge y \wedge z$
$\mathcal{B} = \{110, 011\}$	$y + x \wedge z$
$\mathcal{B} = \{110, 101\}$	$x + y \wedge z$
$\mathcal{B} = \{011, 101\}$	$z + x \wedge y$
$\mathcal{B} = \{110, 101, 011\}$	$x + y + z - (x \vee y \vee z)$
$\mathcal{B} = \{abc\}$	$ax + by + cz$

The last row is the family of bases that are comprised of one subset abc.

References

[Cg62] Cuninghame-Green, R. A. *Describing industrial processes with interference and approximating their steady-state behaviour*. Operations Research Quarterly, 13, 95-100. 1962.

[DS04] Develin, M., Sturmfels, B. *Tropical convexity*. Doc. Math., 9:1-27, 2004. (Erratum pp. 205-206). `arXiv:math.MG/0308254`.

[GA92] Gaubert, S. *Théorie des systémes linéaires dans les diöldes. Thése* , Ecole des Mines de Paris, July 1992.

[JT18] Johnson, M., Tran, N. *Geometry and algorithms for upper triangular tropical matrix identities.* `arXiv:1806.01835`, 2018.

[Jos] Joswig, M. Chapter 4 printout

[Jos16] Joswig, M. *The Cayley Trick for Tropical Hypersurfaces with a view toward Ricardian Economics.* `arXiv:1606.0916v2`. 28 Oct. 2016.

[Mc15] Maclagan, D., Sturmfels, B. "Introduction to Tropical Geometry". Graduate Studies in Mathematics vol. 161. 2015.

[LT17] Lin, B., Tran, N. M. *Linear and Rational Factorization of Tropical Polynomials*. `arXiv:1707.03332v2`. 2017.

[Th98] Thomas, R.R. "Gröbner Bases in Integer Programming." Handbook of Combinatorial Optimization (Vol. 1). 1998 Kluwer Academic Publishers.

[Zh18] Zhang, L., Naitzat, G., Lim, L-H. *Tropical Geometry of Deep Neural Networks*. arXiv:1805.07091 submitted 18 May, 2018.