Hypersimplices, tropical geometry and finite metric spaces

Michael Joswig

TU Berlin & MPI MiS, Leipzig

FPSAC, Bochum, 26 July 2024

joint w/ Laura Casabella
Lars Kastner
and others

1. Hypersimplices and Their Subdivisions
   - Tropicalized linear spaces
   - Hypersimplices
   - Computations

2. Finite Metric Spaces
   - Phylogenetic trees
   - Metric cones and metric fans
   - Beyond split decomposition, by example

Tropicalizing the row space of a matrix

- consider matrix with coefficients in $\mathbb{C}\{\{t\}\}$

$$A = \begin{pmatrix} t^2 & -2t & t^6 & 3 & -2t^7 \\ t^4 & 5t^6 & -t^3 & -4t^5 & t^3 \end{pmatrix}$$

- form *(ordinary) Plücker vector* of maximal minors

$$p(A) = (\, 2t^5 + 5t^8,\ -t^5 - t^{10},\ -3t^4 - 4t^7,\ t^5 + 2t^{11},\ \dots)$$

- take the *lower degree* coefficientwise to get *tropicalized Plücker vector*

$$\pi(A) = (\, 5, 5, 4, 5, 4, 6, 4, 3, 9, 3\,)$$

Tropical Grassmannians

Fix (k, n) with $2 \leq k \leq \lfloor n/2 \rfloor$.

Definition (Tropical Grassmannian)

$\mathrm{TGr}(k, n)$ = set of tropicalized Plücker vectors of $k \times n$-matrices (over $\mathbb{C}\{\{t\}\}$)

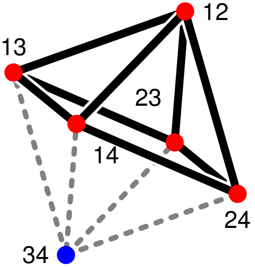
Theorem (Speyer & Sturmfels, 2004)

$\mathrm{TGr}(k, n)$ *is the tropicalization of the ordinary Grassmannian* $\mathrm{Gr}(k, n)$

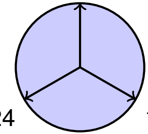
- Speyer & Sturmfels (2004): $(3, 6)$ and $(2, n)$ for arbitrary $n \geq 4$
- Herrmann, Jensen, J. & Sturmfels (2009): $(3, 7)$
- Bendle, Böhm, Ren & Schröter (2024): $(3, 8)$

Subdividing the hypersimplex $\Delta(k,n) = \{x \in [0,1]^n \mid \sum x_i = k\}$

- hypersimplex $\Delta(k,n)$
  - conv(0/1-vectors of length n with k ones)
  - vertices = bases of uniform matroid $U_{k,n}$
- (k,n)-tropicalized Plücker vector
  $\rightsquigarrow$ height function on vertices of $\Delta(k,n)$
  - induces *regular subdivision*
  - in fact, cells are matroid base polytopes
- *secondary fan* = polyhedral fan, where each relatively open cone $\mathrm{sec}(S)$ collects height functions inducing same subdivision S



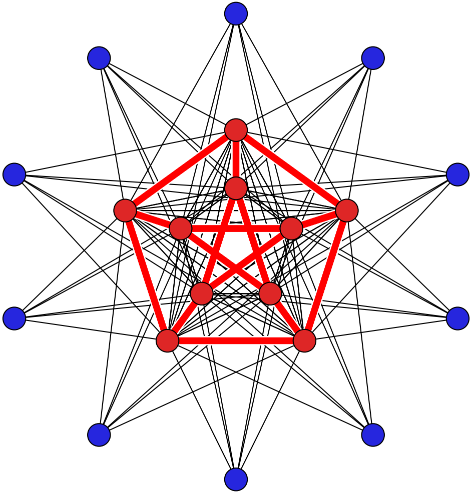
$$\mathrm{TGr}(k,n) \subseteq \underbrace{\mathrm{Dr}(k,n)}_{\text{matroidal}} \leq \underbrace{\mathrm{SecFan}(\Delta(k,n))}_{=:\Sigma(k,n)}$$



$\Sigma(2,4) \subseteq \mathbb{R}^6/\mathbb{R}^4$

Tropical Grassmannian $\mathrm{TGr}(2,5)$ and secondary fan $\Sigma(2,5)$

$\mathrm{TGr}(2,n) = \mathrm{Dr}(2,n)$



$\dim \mathrm{TGr}(2,5) = 2 \qquad \dim \Sigma(2,5) = 10 - 5 = 5$

The secondary fan $\Sigma(2, 7)$

Theorem (Casabella, J. & Kastner 2024+)

1GB of data @ `doi:10.5281/zenodo.12685857`

Theorem (Casabella, J. & Kastner 2024+)

The hypersimplex $\Delta(2, 7)$ has exactly 153,209,697,210 regular triangulations in 30,485,496 orbits, with respect to the natural $\mathrm{Sym}(7)$ action.

Moreover, it has exactly 13,147 orbits of regular coarsest subdivisions.

- Sturmfels & Yu, 2004: $\Sigma(2, 6)$
- Casabella, J. & Kastner, 2024+: $\Sigma(3, 6)$

Software experiments

computing secondary fans

- De Loera: `PUNTOS` (1995)
- Rambau: `TOPCOM` v1.1.2 (2002–2023)
- Jordan, J. & Kastner: `mptopcom` v1.4 (2018–2024)
    - Gawrilow, J. & others: `polymake` v4.12 (1997–2024)

confirmable workflows in computer algebra

- The `MaRDI` Consortium, *Research data management planning in mathematics*, `doi:10.5281/zenodo.10018246`
- J., Kastner & Lorenz, in: Decker et al., The Computer Algebra System `OSCAR`: Algorithms and Examples, Springer 2024.

No computer required, save LaTeX

Sort vertices of $\Delta(k,n)$ in descending lexicographic order. For $1 \leq i \leq \binom{n}{k}$ let

$$\lambda(i) \;=\; \lambda_{k,n}(i) \;=\; \begin{cases} 1 & 1 \leq i \leq n-k; \\ 0 & \text{otherwise,} \end{cases}$$

which induces a (lower) regular subdivision $\Delta(k,n)^{\lambda}$.

Example ($\Delta(2,4)$)

$\lambda_{2,4}(1100) = \lambda_{2,4}(1010) = 1$, and $\lambda_{2,4}(1001) = \cdots = \lambda_{2,4}(0011) = 0$

Proposition (Casabella, J. & Kastner, 2024+)

The regular subdivision $\Delta(2,n)^{\lambda}$ is a non-matroidal coarsest subdivision of spread n.

- Speyer, 2008: for matroidal subdivision of $\Delta(2,n)$ spread $\leq n-2$

Dissimilarity map on n points = height function on $\Delta(2, n)$

- *dissimilarity map* on n points is a map $D : \binom{[n]}{2} \to \mathbb{R}_{\geq 0}$
- for $i \neq j$, identify $\{i, j\} \subset [n]$ with vertex $e_i + e_j$ of $\Delta(2, n)$
- lower envelope of $-D$ becomes

$$\mathcal{E}_{-D}(\Delta(2,n)) \;=\; \left\{ x \in \mathbb{R}^n_{\geq 0} \,\middle|\, x_i + x_j \geq D(i,j) \text{ for all } i \neq j \right\} \;.$$

- *tight span* of D = complex of bounded cells of $\mathcal{E}_{-D}$, dual to $\Delta(2, n)^{-D}$

Theorem (Isbell 1964; Dress 1984; Speyer & Sturmfels, 2004)

The tight span of D is a tree on n labeled leaves if and only if $-D \in \mathrm{TGr}(2, n)$.

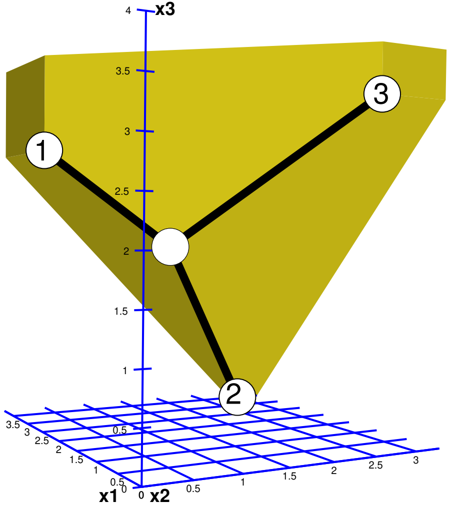
Example

$n = 3$

$D(1,2) = 2.5$
$D(1,3) = 3$
$D(2,3) = 3.5$

$$\mathcal{E}_{-D} = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbb{R}^3 \;\middle|\; \begin{array}{l} x_1, x_2, x_3 \geq 0 \\ x_1 + x_2 \geq 2.5 \\ x_1 + x_3 \geq 3 \\ x_2 + x_3 \geq 3.5 \end{array} \right\}$$

- bounded subcomplex is a tree



Splits

Let $\emptyset \subsetneq A \subsetneq [n]$.

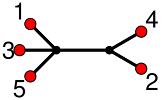
Definition (split pseudometric)

$$D_{A,[n]\setminus A}(i,j) = \begin{cases} 1 & \text{if } \#(\{i,j\} \cap A) = 1 \\ 0 & \text{otherwise.} \end{cases}$$

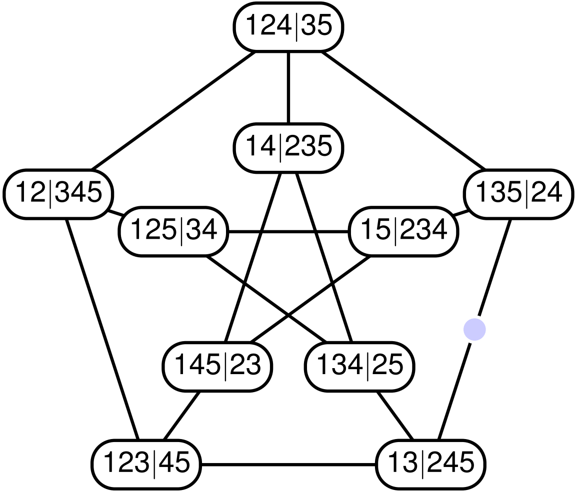
- denote orbit type as $D_{\ell,n-\ell}$, where $\ell = \#A$

Example ($n = 5$)

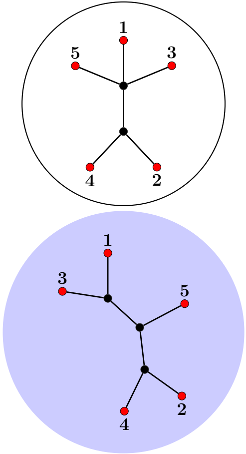
tight span of $D_{135,24}$ looks like



$\mathrm{TGr}(2,5) = \mathrm{Dr}(2,5)$



splits correspond to rays



The metric cone

- $D : \binom{[n]}{2} \to \mathbb{R}_{\geq 0}$ *pseudo-metric* if it satisfies the triangle inequality

$$D(i,k) \leq D(i,j) + D(j,k) \quad \text{for all } i,j,k \in [n]$$

Definition (metric cone)

$$\mathrm{MC}(n) = \left\{ D \in \mathbb{R}_{\geq 0}^{\binom{n}{2}} \,\middle|\, \text{D pseudometric} \right\}$$

Deza and Dutour-Sikirić, 2018: computing $\mathrm{MC}(n)$

n	3	4	5	6	7	8
rays	3	7	25	296	55,226	119,269,588
orbits	1	2	3	8	46	3,918

The metric fan

- *secondary metric cone* of height function $\delta : \binom{[n]}{2} \to \mathbb{R}$:

$$\mathrm{MC}(\delta) \ := \ \mathrm{sec}(\Delta(2,n)^{-\delta}) \cap \mathrm{MC}(n)$$

Definition (metric fan)

$\mathrm{MF}(n)$ = polyhedral fan partitioned by secondary metric cones, supported on $\mathrm{MC}(n)$

Proposition

The metric fan $\mathrm{MF}(n)$ *has the following types of rays:*

1. *negatives of rays of the secondary fan* $\Sigma(2,n)$*;*
2. *one additional orbit, corresponding to the n split pseudo-metrics of type* $D_{1,n-1}$*.*

“prime metrics” of Koolen, Moulton & Tönges, 2000

Classification of finite metric spaces

Theorem (Koolen, Moulton & Tönges, 2000; Sturmfels & Yu, 2004)

The metric fan $\mathsf{MF}(6)$ *has exactly* 14 *orbits of rays, with respect to the natural* $\mathsf{Sym}(6)$ *action.*

Theorem (Casabella, J. & Kastner, 2024+)

The metric fan $\mathsf{MF}(7)$ *has exactly* $13{,}147 + 1$ *orbits of rays, with respect to* $\mathsf{Sym}(7)$.

Multiple alignment of DNA sequences

A.andrenof	...atttctacatgaataatatttatatttcaagagtcaaattca...
A.mellifer	...atttccacatgatttatatttatatttcaagaatcaaattca...
A.dorsata	...atttcaacatgaataatattaatatttcaagaatcaaattca...
A.cerana	...atttctacatgattcatatttatgtttcaagaatcaaattca...
A.florea	...atttctacatgaataatatttatatttcaagagtcaaattca...
A.koschev	...atttctacatgaataatatttatatttcaagaatcaaactca...

$\rightsquigarrow$ *editing distance* = genetic distance

Huson & Bryant 2006:
SplitsTree, example file bees.nex,
dissimilarity β from DNA sequences of length 677
(out of $\approx 250 \cdot 10^6$)

Split decomposition and beyond

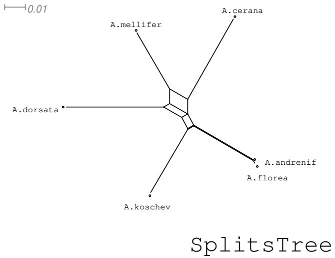
Let D be a dissimilarity on n points.

Theorem (Bandelt & Dress, 1992)

There is a coherent decomposition

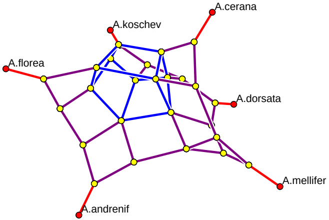
$$D = D_0 + \sum_{S \text{ split of } [n]} \lambda_S D_S ,$$

where D_0 is split prime, and decomposition is unique.



Hirai, 2006; Herrmann & Joswig, 2008:

- replace splits by an arbitrary set R of rays of $\mathrm{MF}(n)$
- D_0 not in a secondary cone spanned by rays in R
- coefficients depend on ordering of R



Negatives of the rays of the secondary cone of $\Delta(2,6)^{-\beta}$

ray	coordinates	spread	coherency index
s_1	$(1,1,1,0,1,0,0,1,0,0,1,0,1,0,1)$	2	0.03175776
s_2	$(1,1,0,0,0,0,1,1,1,1,1,1,0,0,0)$	2	0.00886262
s_3	$(1,0,1,0,0,1,0,1,1,1,0,0,1,1,0)$	2	0.00664697
s_4	$(0,1,0,0,1,1,0,0,1,1,1,0,0,1,1)$	2	0.00147710
s_5	$(1,1,1,0,0,0,0,1,1,0,1,1,1,1,0)$	2	0.00516987
r_1	$(1,2,2,0,1,1,1,1,2,2,2,1,2,1,1)$	5	0.00369276
r_2	$(1,2,2,1,2,1,1,2,3,2,3,2,1,2,1)$	7	$2 \cdot 10^{-9}$
r_3	$(1,2,2,1,1,1,1,2,2,2,3,1,1,1,2)$	6	$2.5 \cdot 10^{-9}$
r_4	$(1,2,2,1,1,1,1,2,2,2,1,1,1,1,0)$	5	$5 \cdot 10^{-10}$

Six orbits, separated by extra line space; sorted by descending coherency index.

Conclusion

Computing $\Sigma(2,7)$ yields

- new families of coarsest subdivisions of $\Delta(k,n)$;
- classification of finite metric spaces on 7 points;
- better resolution for analyzing data.

Laura Casabella, Michael Joswig, and Lars Kastner, *Subdivisions of hypersimplices: with a view toward finite metric spaces*, 2024, Preprint `arXiv:2402.17665`.

Sven Herrmann and Michael Joswig, *Splitting polytopes*, Münster J. Math. **1** (2008), 109–141. MR 2502496

Charles Jordan, Michael Joswig, and Lars Kastner, *Parallel enumeration of triangulations*, Electron. J. Combin. **25** (2018), no. 3, Paper 3.6, 27. MR 3829292