The space of phylogenetic trees
and the tropical geometry of flag varieties

FPSAC, Nagoya, August 3, 2012
Christopher Manon
www.math.berkeley.edu/~manonc
supported by NSF fellowship DMS-0902710
$G$ a reductive group over $\mathbb{C}$ (e.g. $SL_n(\mathbb{C})$, $SP_{2n}(\mathbb{C})$)

Finite dimensional irreducible representations are indexed by the lattice points $\lambda$ in a cone $\Delta_G$.

Example: $\Delta_{SL_n(\mathbb{C})} = \{(a_1, \ldots, a_{n-1}) | a_i \geq a_j \geq 0, \ i < j\}$

$\omega_m = (1, \ldots, 1, 0, \ldots, 0)$

$V(\omega_m) = \Lambda^m(\mathbb{C}^n)$
A flag variety $G/P$ is a complete quotient of $G$ by a parabolic subgroup $P \subset G$.

Any flag variety is the orbit through the highest weight vector in $\mathbb{P}(V(\lambda))$ for some $\lambda \in \Delta$.

$$G/P \cong G \circ [v_\lambda] \in \mathbb{P}(V(\lambda))$$

In particular $G/P$ is projective, cut out by the homogeneous ideal $I_\lambda$, with projective coordinate ring

$$R_\lambda = \bigoplus_{N \geq 0} H^0(G/P, L_\lambda^\otimes N) = \bigoplus_{N \geq 0} V(N\lambda^*)$$

-Borel-Bott-Weil Theorem.
Let $P_{m,n}$ be the parabolic subgroup of $SL_n(\mathbb{C})$ of the form

$$\begin{bmatrix} A & B \\ 0 & C \end{bmatrix}$$

where $A$ is $m \times m$, and $C$ is $n - m \times n - m$.

$SL_n(\mathbb{C})/P_{m,n} \cong Gr_m(\mathbb{C}^n) = SL_n(\mathbb{C}) \circ [z_1 \wedge \ldots \wedge z_m] \subset \mathbb{P}(\bigwedge^m(\mathbb{C}^n))$

$$R_{\omega_m} = \bigoplus_{N \geq 0} V(N\omega^*_m) = \mathbb{C}[\ldots z_{i_1 \ldots i_m} \ldots]/I_{m,n}$$

This algebra is known as the Plücker algebra, and $I_{m,n}$ is the Plücker ideal.
Reminder: tropical varieties

\[ T = \mathbb{R} \cup \{-\infty\} \]

\[ a \oplus b = \max\{a, b\} \]
\[ a \otimes b = a + b. \]

\[ X = \{x_1, \ldots, x_n\}, \ f(X) = \sum C_\vec{m} \vec{x}^\vec{m} \in \mathbb{C}[X] \]

\[ T(f) = \bigoplus_{\vec{m} \neq 0} (\otimes_{i=1}^{n} m_i x_i) = \max\{\ldots, \sum_{i=1}^{n} m_i x_i, \ldots\} \]

\[ tr(f) = \{\vec{p} \in \mathbb{T}^n | T(f) \text{ has two maxima at } \vec{p}\} \]

\[ I \subset \mathbb{C}[X], \ tr(I) = \bigcap_{f \in I} tr(f) \]
There is always a finite set $M \subset I$ such that $\text{tr}(I) = \bigcap_{f \in M} \text{tr}(f)$.

-Bogart, Jensen, Speyer, Sturmfels, Thomas
weighted (phylogenetic) trees

Tree $T$ with $n$ leaves labeled $\{1, \ldots, n\}$

$P_T = \{w : Edge(T) \to \mathbb{R} \mid \geq 0 \text{ on internal edges}\}$

![Figure 1:](image)
The space of trees $\mathcal{T}^n$

For $\psi : T' \to T$ a map of $n$-trees there is a map $\psi^* : P_T \to P_{T'}$.

![Figure 2:](image)

We define $\mathcal{T}^n = \bigsqcup_{|\text{Leaf}(T)|=n} P_T / \sim$

This space was studied by Billera, Holmes and Vogtman.
Dissimilarity vectors

For $1 < m < n$, and $\{i_1, \ldots, i_m\} \subset \{1, \ldots, n\}$ define

$$d_{i_1, \ldots, i_m}(T, w) = \sum_{e \in C(i_1, \ldots, i_m)} w(e).$$

Figure 3: $d_{134}(T, w) = 14.26$
Dissimilarity vectors

Definition:
- We call \( d^m(T, w) = (\ldots d_{i_1,\ldots,i_m}(T, w)\ldots) \) the \( m \)-dissimilarity vector of \((T, w)\).
- We call \( d^m : \mathcal{T}^n \rightarrow \mathbb{R}^{\binom{n}{m}} \) the \( m \)-dissimilarity map.

\[
d^3(2.5 \rightarrow 1, 1.5 \rightarrow 3.5, 2 \rightarrow 2) = (8.5, 7, 9.5, 8)
\]
2-dissimilarity vectors

The map $d^2 : \mathcal{T}^n \rightarrow \mathbb{R}^{\binom{n}{2}}$ is 1 – 1.

Figure 4: $d_{34} = 13.35$
$d^2$ and the tropical Grassmannian

\[ I_{2,n} = \langle \{ z_{ij}z_{kl} - z_{ik}z_{jl} + z_{il}z_{jk} | 1 \leq i < j < k < l \leq n \} \rangle \subset \mathbb{C}[\ldots z_{ij} \ldots], \]

\[ \mathbb{C}[\ldots z_{ij} \ldots]/I_{2,n} = R_{\omega_2} \]

**Theorem [Speyer, Sturmfels]:**
The image of $d^2$ coincides with $tr(I_{2,n})$.

Plücker relations form a tropical basis of $I_{2,n}$, so the $d^2(T) \in \mathbb{R}^{n\choose 2}$ satisfy

\[ \max\{d_{ij} + d_{kl}, d_{ik} + d_{jl}, d_{il} + d_{jk}\}. \]
Conjecture [Cools, Pachter, Speyer]:

**Theorem [Iriarte-Giraldo, M]:**

\[ d_m(T^n) \subset tr(I_{m,n}) \subset \mathbb{R}^{n \choose m} \]

This implies \( d_m(T, w) \) satisfies \( T(f) \) for all \( f \in I_{m,n} \).

*Plücker relations no longer make a tropical basis when \( m > 2 \).*
An irreducible representation of $GL_n(\mathbb{C})$ is given by a Young diagram.

The representation $V(\lambda)$ has a basis given by semi-standard fillings $\tau$ of $\lambda$ by the indices $\{1, \ldots, n\}$. 

\begin{tabular}{|c|c|c|c|c|c|}
\hline
1 & 1 & 2 & 4 & 5 \\
\hline
3 & 3 & 3 & 6 \\
\hline
4 & 5 \\
\hline
5 \\
\hline
\end{tabular}
Example: $V(\omega_m) = \bigwedge^m(\mathbb{C}^n)$

$\bigwedge^m(\mathbb{C}^n)$ has a basis of elements $z_{i_1} \wedge \ldots \wedge z_{i_m}$, with $i_1 < \ldots < i_m$. This determines a semi-standard filling of a column of $m$ boxes.
We have seen that $z_{i_1} \wedge \ldots \wedge z_{i_m}$ tropicalizes to the dissimilarity function $d_{i_1,\ldots,i_m} : \mathcal{T}^n \rightarrow \mathbb{R}$. What about other semi-standard tableaux?

Definition:
We define $d_\tau : \mathcal{T}^n \rightarrow \mathbb{R}$ to be $\sum d_{I_k}$, where $I_k$ are the columns of $\tau$.

This construction gives a function on $\mathcal{T}^n$ for any basis member of a representation of $GL_n(\mathbb{C})$. 
Yes.

Theorem [M]:
Let $I_{\lambda}$ be the ideal which cuts the flag variety $GL_n(\mathbb{C})/P$ out of $\mathbb{P}(V(\lambda))$. There is a map of complexes $d^\lambda : T^n \to tr(I_{\lambda})$ where $d^\lambda = (d_{\tau_1}, \ldots, d_{\tau_t})$.
Let $A$ be an algebra over $\mathbb{C}$. A valuation $\nu : A \rightarrow \mathbb{T}$ is a function which satisfies the following.

\[ \nu(ab) = \nu(a) + \nu(b) = \nu(a) \otimes \nu(b) \]

\[ \nu(a + b) \leq \max\{\nu(a), \nu(b)\} = \nu(a) \oplus \nu(b) \]

\[ \nu(C) = 0 \text{ for } 0 \neq C \in \mathbb{C} \]

\[ \nu(0) = -\infty. \]
Tropical theory: lifting

For $A$ an algebra over $\mathbb{C}$, let $\mathcal{V}_T(A)$ be the set of valuations of $A$ into $\mathbb{T}$ over $\mathbb{C}$.

**Theorem [Payne]:**
For any presentation

\[
0 \longrightarrow I \longrightarrow \mathbb{C}[X] \longrightarrow A \longrightarrow 0
\]

there is a surjective map $\pi_X : \mathcal{V}_T(A) \rightarrow tr(I)$ given by $\pi_X(v) = (\ldots v(x_i) \ldots)$.

Instead of the ideal $I_\lambda$, we consider the algebra $R_\lambda$
Valuations from chains of groups

**Theorem [M]:**
For any commutative $G$–algebra $A$, and every chain of subgroups

$$G_1 \xrightarrow{\phi_1} \cdots \xrightarrow{\phi_{k-1}} G_{k-1} \xrightarrow{\phi_k} G,$$

there is a cone of valuations $D_{\vec{\phi}}$ in $\mathbb{V}_T(A)$.

These cones glue together into a polyhedral complex $\mathcal{D}(G)$ of valuations on $A$.

**Strategy:** Find a way to turn trees into chains of subgroups of $GL_n(\mathbb{C})$. 
Trees and subgroups of $GL_n(\mathbb{C})$

Rule: to an edge $e \in E(T)$ assign the group $GL_k(\mathbb{C})$ where $k$ is the number of leaves "above" $e$. 

$1 \rightarrow GL_1(\mathbb{C}) \times GL_1(\mathbb{C}) \times GL_1(\mathbb{C}) \rightarrow GL_1(\mathbb{C}) \times GL_2(\mathbb{C}) \rightarrow GL_3(\mathbb{C})$
Trees and subgroups of $GL_n(\mathbb{C})$
More general groups

For any Dynkin Diagram $\Gamma$, let $G(\Gamma)$ be the associated simply connected, semi-simple group over $\mathbb{C}$.

The diagram $\Gamma$ also gives a hyperplane arrangement, we let $B_\Gamma$ denote the Bergman fan of the associated matroid. This space was studied by Ardilla, Klivans and Williams.
Dynkin Diagrams

- $A_n$
- $B_n$
- $C_n$
- $D_n$
- $E_6$
- $F_4$
- $G_2$
Faces of $B_{\Gamma}$ are indexed by "tubings" of $\Gamma$. This is related to the graph associahedron of the Dynkin diagram.

Tubings also correspond to chains of Levi subgroups in $G(\Gamma)$. 
For $\Gamma$ a Dynkin diagram, $B_\Gamma$ the Bergman Fan of its associated hyperplane arrangement, and $tr(I_\lambda)$ the tropical variety associated to an ideal $I_\lambda$ which cuts out a flag variety $G(\Gamma)/P$, we have the following:

**Theorem [M]:**
There is a map of complexes

$$\pi_\lambda : B_\Gamma \to tr(I_\lambda)$$
Other groups

Just as semi-standard tableaux give functions on $\mathcal{T}^n$, basis members (ie canonical basis, standard monomials, etc) of $V(\lambda)$ give functions on $B_\Gamma$.

<table>
<thead>
<tr>
<th>$GL_n(\mathbb{C})$</th>
<th>$G(\Gamma)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mathcal{T}^n$</td>
<td>$B_\Gamma$</td>
</tr>
<tr>
<td>$GL_{k_1}(\mathbb{C}) \times \ldots \times GL_{k_m}(\mathbb{C}) \subset GL_n(\mathbb{C})$</td>
<td>$L \subset G(\Gamma)$</td>
</tr>
<tr>
<td>semi-standard tableaux</td>
<td>standard monomials</td>
</tr>
</tbody>
</table>
Other Generalizations: Buildings

The complex $\mathcal{D}(G)$ has a $G$–action. An element $g \in G$ takes a cone $D_\phi \subset \mathcal{D}(G)$ to $D_{g \circ [\phi]}$

$$1 \rightarrow H \xrightarrow{\phi} G \xrightarrow{\text{Ad}_g} G$$

Applying this to the cone defined by $1 \subset T \subset G$ for a maximal torus $T$ yields a copy of (a cone over) the spherical building of $G$ inside $\mathcal{D}(G)$. For $A$ a $G$–algebra we have,

$$\mathcal{B}(\mathbb{C}, G) \subset \mathcal{D}(G) \rightarrow \mathbb{V}_\mathbb{C}(A)$$

Results like this have been obtained by Berkovich, and Remy, Thuillier, Werner in the context of flag varieties.
In fact, the product complex $\mathcal{B}(\mathcal{C}, G(\Gamma)) \times \mathcal{B}(\Gamma)$ is a subcomplex of $\mathcal{D}(G(\Gamma))$. 


B. Giraldo: Dissimilarity vectors of trees are contained in the tropical Grassmannian, The Electronic Journal of Combinatorics 17, no. 1, 2010


Thankyou!