

Distance matrices and determinants

Distance = number of edges in **shortest path** between vertices

▪ **Examples** of distance matrices:

$$D = \begin{pmatrix} 0 & 1 & 2 & 3 & 4 \\ 1 & 0 & 1 & 2 & 3 \\ 2 & 1 & 0 & 1 & 2 \\ 3 & 2 & 1 & 0 & 1 \\ 4 & 3 & 2 & 1 & 0 \end{pmatrix} \quad D = \begin{pmatrix} 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 2 & 2 & 2 \\ 1 & 2 & 0 & 2 & 2 \\ 1 & 2 & 2 & 0 & 2 \\ 1 & 2 & 2 & 2 & 0 \end{pmatrix}$$

Submatrices

▪ **Example:** Tree shown in Figure 1, S = subset of leaf vertices.

$$D = \begin{pmatrix} 0 & 1 & 2 & 1 & 1 & 2 & 3 \\ 1 & 0 & 1 & 2 & 2 & 1 & 2 \\ 2 & 1 & 0 & 3 & 3 & 2 & 1 \\ 1 & 2 & 3 & 0 & 2 & 3 & 4 \\ 1 & 2 & 3 & 2 & 0 & 3 & 4 \\ 2 & 1 & 2 & 3 & 3 & 0 & 3 \\ 3 & 2 & 1 & 4 & 4 & 3 & 0 \end{pmatrix}, \quad D[S] = \begin{pmatrix} 0 & 2 & 3 & 4 \\ 2 & 0 & 3 & 4 \\ 3 & 3 & 0 & 3 \\ 4 & 4 & 3 & 0 \end{pmatrix},$$

$\det D = 192$ $\det D[S] = -252$

Application: Phylogenetics

A **phylogenetic tree** describes the evolutionary history connecting living organisms

▪ **Problem:** What is the most likely phylogenetic tree, given observed biological data?

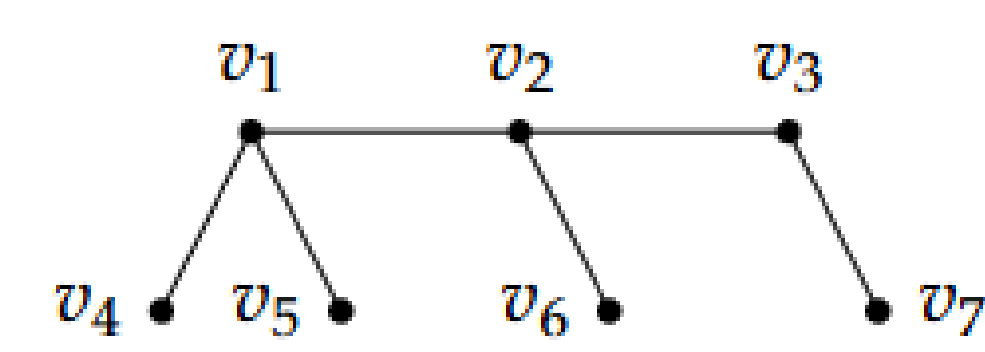


Figure 1. Example tree with 4 leaves.

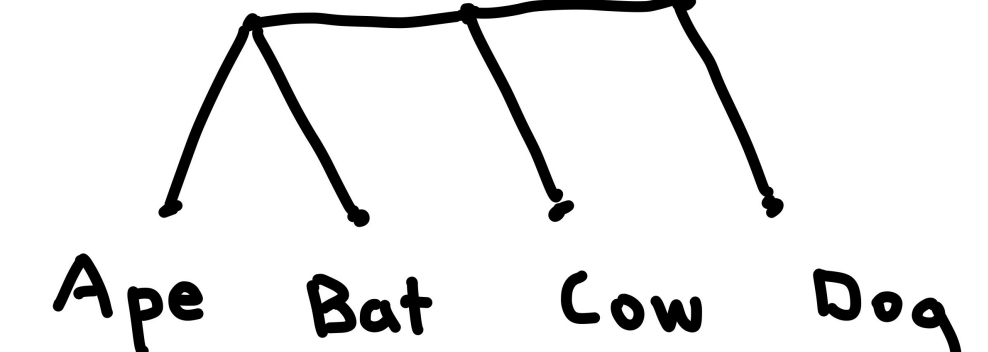


Figure 2. Possible phylogenetic tree.

Potential theory

Problem: How do particles “distribute” within a region, given repulsive potential?

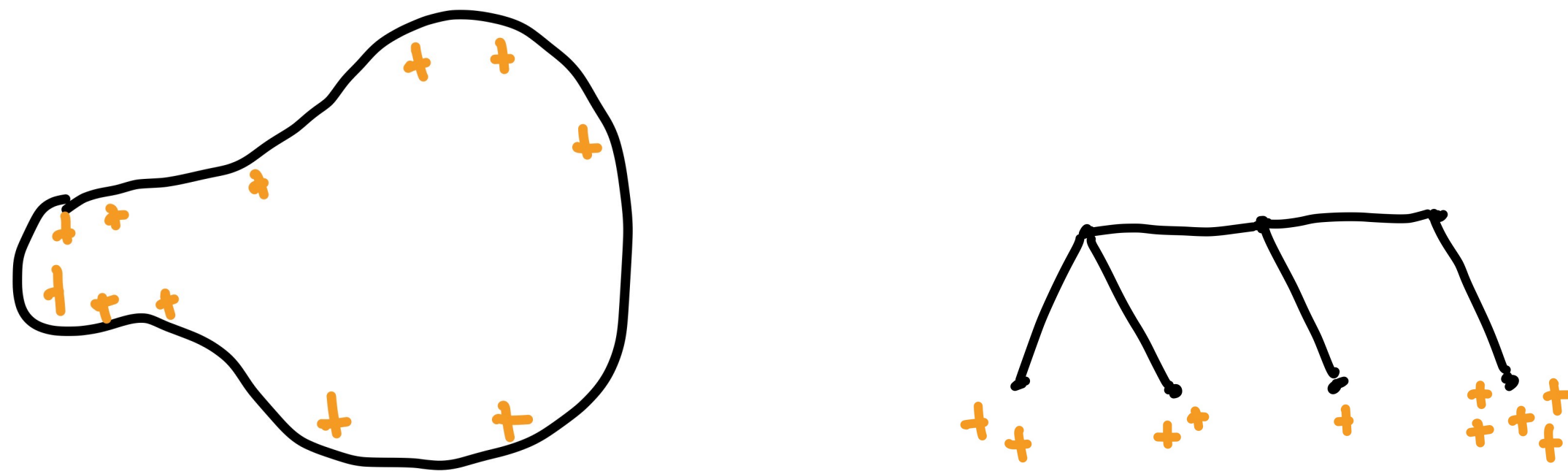


Figure 3. Charged particles on a 2D surface, and on a 1D space.

- Consider “energy” $\mathcal{E}(\mathbf{u}) = -\frac{1}{2}\mathbf{u}^\top D[S]\mathbf{u}$
- Equilibrium reached when energy minimized
- Conservation of mass: constrained to $\mathbf{u}$ with $\mathbf{1}^\top \mathbf{u} = 1$

Rooted spanning forests

Equilibrium vector $\mathbf{u}^*$ can be expressed in terms of **combinatorial** quantities

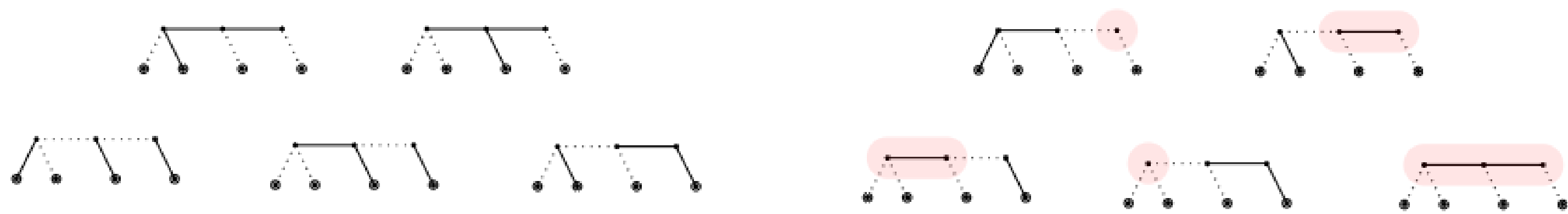


Figure 4. Some S -rooted and $(S, *)$ -rooted spanning forests

- $\mathcal{F}_1(G; S) = S$ -rooted spanning forests of G
- $\mathcal{F}_2(G; S) = (S, *)$ -rooted spanning forests
- $\kappa_1(G; S)$ and $\kappa_2(G; S)$ = number of respective spanning forests

Theorem (Bapat–Sivasubramanian, 2011)

“Equilibrium” vector $\mathbf{u}^*$ satisfying $D[S]\mathbf{u}^* = \lambda \mathbf{1}$ and $\mathbf{1}^\top \mathbf{u}^* = 1$ is

$$\mathbf{u}_i^* = \frac{1}{2\kappa_1(G; S)} \sum_{F \in \mathcal{F}_1(G; S)} (2 - \deg^o(F, v_i)).$$

Bapat–Sivasubramanian used above to prove a combinatorial identity for $\text{cof } D[S]$

- Why does $2 - \deg$ appear? Identity with Laplacian matrix L :

$$(LD)_{i,j} = \begin{cases} 2 - \deg(v_i) & \text{if } i \neq j \\ -\deg(v_i) & \text{if } i = j \end{cases}$$

Problem: What is the **determinant** of the distance matrix of a tree?

- For the two examples, $\det D = 32$.

The fact that both matrices have the same determinant is not a coincidence!

Theorem (Graham–Pollack, 1971)

For any tree on n vertices, $\det D = (-1)^{n-1} 2^{n-2} (n-1)$.

Key Observations: (cf. facts on signature of $D[S]$, due to Bapat)

- $\min \{\mathcal{E}(\mathbf{u}) : \mathbf{1}^\top \mathbf{u} = 1\} = -\frac{1 \det D[S]}{2 \text{cof } D[S]}$
- Minimum $\mathcal{E}(\mathbf{u})$ occurs at $D[S]\mathbf{u}^* = \lambda \mathbf{1}$

Here $\text{cof } A$ is the **sum of cofactors**, i.e. $\mathbf{1}^\top A^{-1} \mathbf{1} = \text{cof } A / \det A$

Consequence of potential theory perspective: **monotonicity** property

Theorem (R–Shokrieh–Wu, 2025+)

For any tree G , if $A \subset B \subset V(G)$, then

$$\frac{\det D[A]}{\text{cof } D[A]} \leq \frac{\det D[B]}{\text{cof } D[B]}.$$

Theorem (R–Shokrieh–Wu, 2025+)

For any tree G and any vertex subset $S \subset V(G)$, we have

$$\det D[S] = (-1)^{|S|-1} 2^{|S|-2} \left((n-1) \kappa_1(G; S) - \sum_{F \in \mathcal{F}_2(G; S)} (\deg^o(F, *) - 2)^2 \right).$$

By above and an identity for $\text{cof } D[S]$ by Bapat–Sivasubramanian (2011),

$$\frac{\det D[S]}{\text{cof } D[S]} = \frac{1}{2} \left((n-1) - \frac{\sum_{F \in \mathcal{F}_2(G; S)} (\deg^o(F, *) - 2)^2}{\kappa_1(G; S)} \right).$$

Symanzik polynomials

Symanzik polynomials of the first and second kind, ψ_G and ϕ_G , appear in quantum field theory. For a tree with **edge lengths** $\{\alpha_e : e \in E(G)\}$,

$$\frac{\det D[S]}{\text{cof } D[S]} = \frac{1}{2} \left(\sum_{E(G)} \alpha_e - \frac{\phi_{(G/S)}(p_{can}; \underline{\alpha})}{\psi_{(G/S)}(\underline{\alpha})} \right), \quad \text{where } p_{can}(v) = \deg(v) - 2.$$

Further questions

- q -distance analogues,

$$[n]_q = 1 + q + q^2 + \cdots + q^{n-1}$$

q -analogue of Graham–Pollack found by Bapat–Lal–Pati and Yan–Yeh

- k -Steiner distances: for a choice of k vertices, count number of edges in subtree spanned by chosen vertices. Explored by Cooper–Du and Azimi–Sivasubramanian
- “Combinatorial” proof of main identity? Gutierrez–Lillo