On the Wasserstein Distance Between k-Step Probability Measures on Finite Graphs

Sophia Benjamin[a] Arushi Mantri[b] Quinn Perian[c]
Mentor: Pakawut Jiradilok

[a]North Carolina School of Science and Mathematics, Durham, NC
[b]Jesuit High School, Portland, OR
[c]Stanford Online High School, Palo Alto, CA

PRIMES Conference

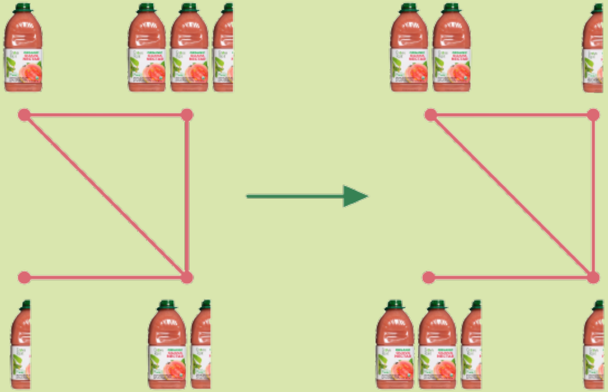
October 16 2021

Guavas?

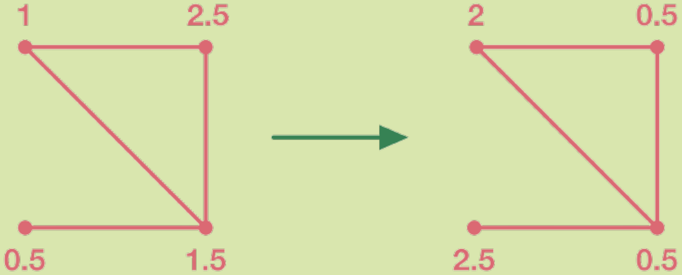
Guava Juice!


Grown Right
ORGANIC
GUAVA
NECTAR
ORGANIC GUAVA NECTAR
MADE WITH FRUIT PUREE
NOT FROM CONCENTRATE

Transporting Guava Juice

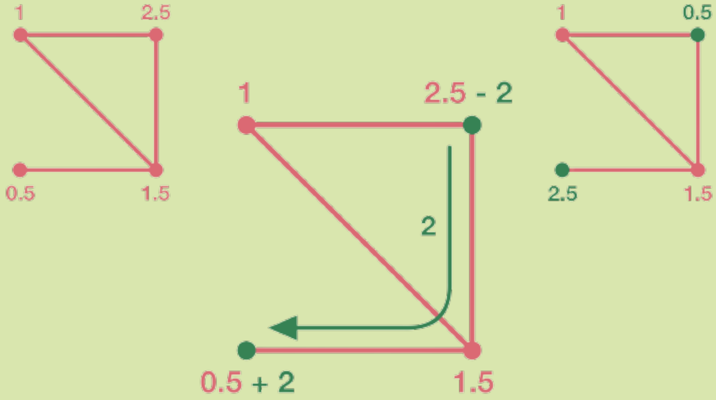


We must transport guava juice stored in warehouses from the first distribution to the second distribution via roads.



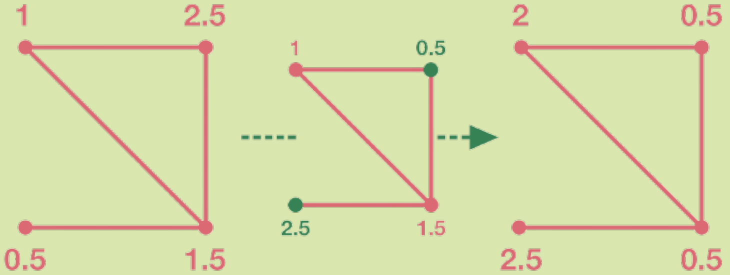
Transporting 1 gallon of guava juice along 1 road costs $1.
Let's try transporting the juice and see how much it costs!

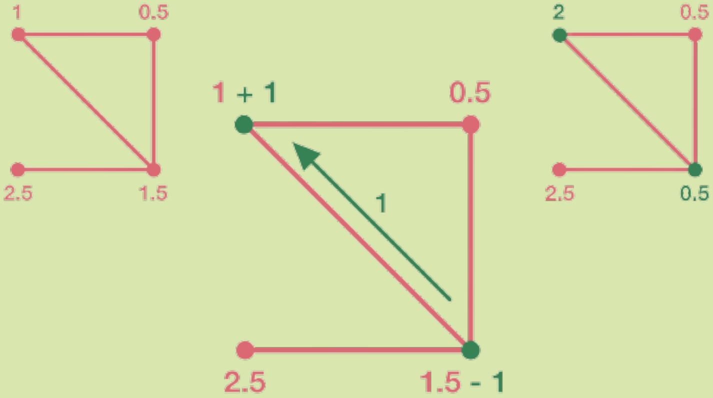
Transporting Guava Juice



cost = $2 \cdot 2$

Transporting Guava Juice

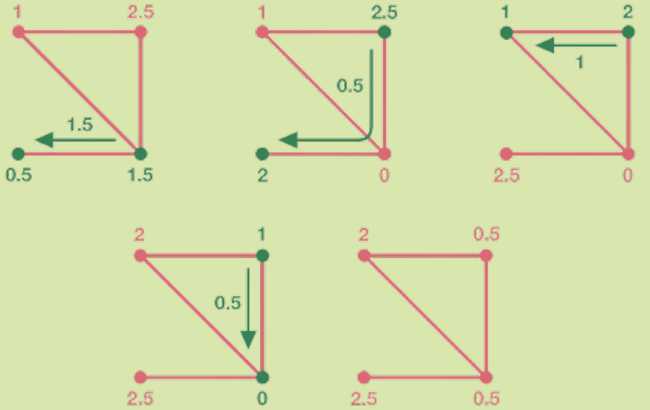




cost = 2 · 2 + 1 · 1 = 5

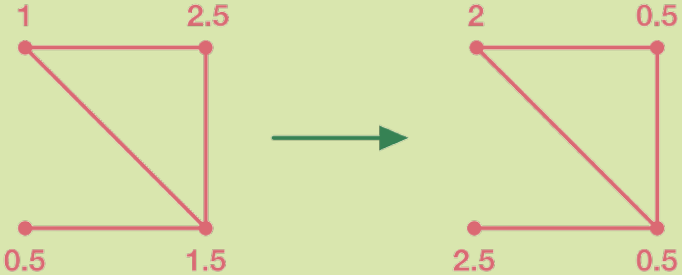
Transporting Guava Juice

Here's a more cost-effective way of transporting the guava juice:



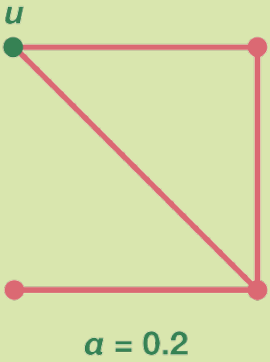
$$\text{cost} = 1.5 \cdot 1 + 0.5 \cdot 2 + 1 \cdot 1 + 0.5 \cdot 1 = 4$$

A natural question:
What is the most cost-effective way of transporting the juice?



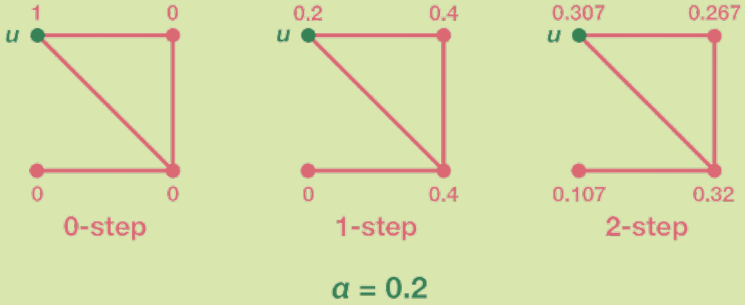
Wasserstein distance = minimum cost of transportation.

We study the Wasserstein distance between the k-step probability distributions of **random walks with laziness** on a finite graph.



- starting vertex u
- laziness $\alpha \in [0, 1]$: probability of staying put

We study the Wasserstein distance between the ***k*-step probability distributions** of random walks with laziness on a finite graph.



At each vertex, the proportion α of the mass stays while the rest of the mass splits evenly among its neighbors.

The following definition captures our object of study.

Definition

We define a **Guvab** to be a tuple (G, u, v, α, β) where G is a finite simple connected graph, $u, v \in V(G)$, and $\alpha, \beta \in [0, 1]$ with $\alpha \leq \beta$.

Given a Guvab and a nonnegative integer k, consider the k-step probability distributions of the two random walks with starting vertices u,v and lazinesses α, β, respectively. We denote by W_k the Wasserstein distance between these two k-step probability distributions.

Our Project

Motivation:

- W_1 is used to determine Lin-Lu-Yau-Ollivier-Ricci curvature ([LLY11])
- Applications in drug design, cancer networks, and economic risk ([SGR$^+$15], [SGT16], [WX21])

Our Question:

- What about W_k as k gets larger and larger?
- Does it converge? When? To what? How fast?

Main Result #1: Classifying End Behavior

When $\lim_{k\to\infty} W_k$ is well-defined, call it W.

Theorem (Classifying End Behavior)

All Guvabs fit into one of four categories, and we know when they fit into each category:

1. $W = 1$ *and* $\alpha, \beta < 1$
   - *G bipartite,* $\alpha = \beta = 0$*,* $d(u, v)$ *is odd*
2. $W = \frac{1}{2}$ *and* $\alpha, \beta < 1$
   - *G bipartite,* $\alpha = 0 < \beta < 1$
3. $W = 0$ *and* $\alpha, \beta < 1$
   - *all other Guvabs with* $\alpha, \beta < 1$
4. $\beta = 1$

Main Result #2: Exponential Convergence

For any Guvab, $\lim_{k\to\infty} W_{2k}$ and $\lim_{k\to\infty} W_{2k+1}$ are well-defined (due to Main Result 1).

Theorem (Exponential Convergence of W-Dist)

For any Guvab, we have that:

- *either* $\{W_{2k}\}$ *is eventually constant, or there exists a constant* $\lambda_{even} \in (-1, 1)$ *and a positive constant* $c_{even} > 0$ *such that* $|W_{2k} - \lim_{k\to\infty} W_{2k}| \sim c_{even} \cdot |\lambda_{even}|^{2k}$
- *either* $\{W_{2k+1}\}$ *is eventually constant, or there exists a constant* $\lambda_{odd} \in (-1, 1)$ *and a positive constant* $c_{odd} > 0$ *such that* $|W_{2k+1} - \lim_{k\to\infty} W_{2k+1}| \sim c_{odd} \cdot |\lambda_{odd}|^{2k+1}$

Main Result #3: Characterization of Constancy

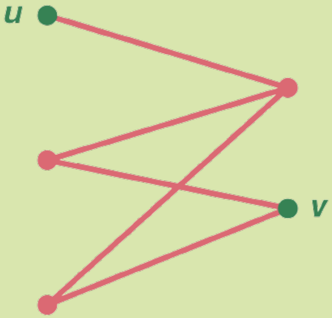
Theorem (Characterization of Constancy)

When $\alpha, \beta < 1$, we have that $\{W_k\}$ is eventually constant if and only if one of the following holds:

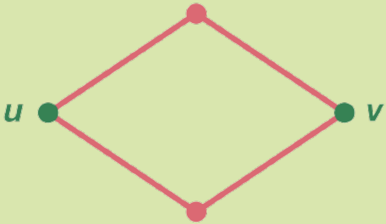
1. $\alpha = \beta = 0$, *G is bipartite, and $d(u, v)$ is odd (here $W = 1$),*
2. $\alpha = 0$, $\beta = \frac{1}{2}$, *and G is bipartite (here $W = \frac{1}{2}$),*
3. $\alpha = \beta = 0$ *and $N(u) = N(v)$ (here $W = 0$),*
4. $\alpha = \beta = \dfrac{1}{\deg u + 1}$, *the edge $(u, v) \in E(G)$, and if the edge (u, v) were removed from $E(G)$ then u, v would have $N(u) = N(v)$ (here $W = 0$),*
5. $\alpha = \beta$ *and $u = v$ (here $W = 0$).*

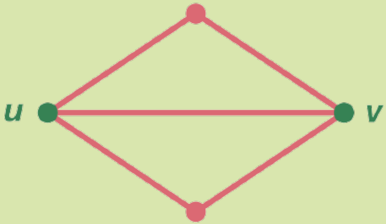
Main Result #3: Characterization of Constancy

1. $\alpha = \beta = 0$, G is bipartite, and $d(u, v)$ is odd (here $W = 1$),
2. $\alpha = 0$, $\beta = \frac{1}{2}$, and G is bipartite (here $W = \frac{1}{2}$),



3. $\alpha = \beta = 0$ and $N(u) = N(v)$ (here $W = 0$),
4. $\alpha = \beta = \dfrac{1}{\deg u + 1}$, the edge $(u, v) \in E(G)$, and if the edge (u, v) were removed from $E(G)$ then u, v would have $N(u) = N(v)$ (here $W = 0$),





5. $\alpha = \beta$ and $u = v$ (here $W = 0$).

Acknowledgements

Big thank you to:

- Our mentor, Pakawut Jiradilok,
- Dr. Supanat Kamtue,
- The PRIMES-USA program,
- Our families and friends,
- And the Guvabs we found along the way!

Thanks for listening! Any questions?

References

Yong Lin, Linyuan Lu, and Shing-Tung Yau.
Ricci curvature of graphs.
Tohoku Math. J., 63(4):605–627, 2011.

Romeil Sandhu, Tryphon Georgiou, Ed Reznik, Liangjia Zhu, Ivan Kolesov, Yasin Senbabaoglu, and Allen Tannenbaum.
Graph curvature for differentiating cancer networks.
Sci. Rep., 5(1):1–13, 2015.

Romeil S Sandhu, Tryphon T Georgiou, and Allen R Tannenbaum.
Ricci curvature: An economic indicator for market fragility and systemic risk.
Sci. Adv., 2(5):e1501495, 2016.

JunJie Wee and Kelin Xia.
Ollivier persistent ricci curvature-based machine learning for the protein–ligand binding affinity prediction.
J Chem Inf Model, 61(4):1617–1626, 2021.