LA-UR-21-27108

Approved for public release; distribution is unlimited.

Title: Tropical Matrix Factorization

Author(s): Ortiz, Jose Angel

Intended for: HPC Showcase, 2021-07-22 (Online, New Mexico, United States)

Issued: 2021-07-21

Tropical Matrix Factorization

Jose Ortiz

Kansas State University

Los Alamos NATIONAL LABORATORY

Introduction

- Replacing addition with maximum (or minimum) and swapping multiplication with addition gives the **tropical numbers** $\mathbb{T} = (\overline{\mathbb{R}}, \oplus, \odot)$.

Classical Setting ⟷ ***Tropical Setting***

Classical Setting	Tropical Setting
$1 + 0 = 1$	$1 \oplus 0 := \max(1, 0) = 1$
$1 + 1 = 2$	$1 \oplus 1 := \max(1, 1) = 1$
$1 + 2 = 3$	$1 \oplus 2 := \max(1, 2) = 2$
$1 \cdot 0 = 0$	$1 \odot 0 := 1 + 0 = 1$
$1 \cdot 1 = 1$	$1 \odot 1 := 1 + 1 = 2$
$3 \cdot 3 = 9$	$3 \odot 3 := 3 + 3 = 6$

Introduction

- Compared to addition and multiplication, comparison and addition are **faster**.

Introduction

- Compared to addition and multiplication, comparison and addition are **faster**.
- Comparison and addition are **robust** operations.

Introduction

- Compared to addition and multiplication, comparison and addition are **faster**.
- Comparison and addition are **robust** operations.
- Moving between classical and tropical settings is often **reversible**.
  - Classical $\rightarrow$ Tropical: *Easy*
  - Classical $\leftarrow$ Tropical: *Harder*

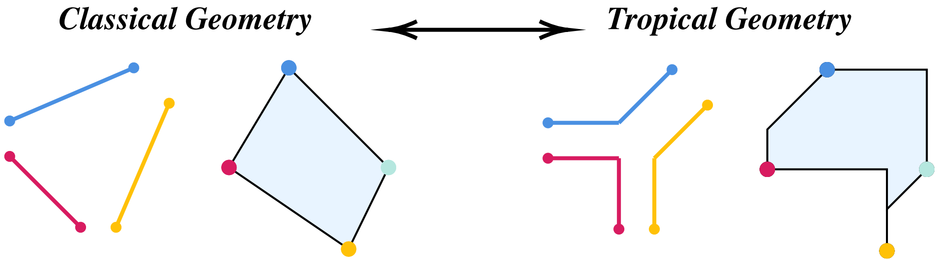
Los Alamos NATIONAL LABORATORY

Introduction

- **Tropical geometry** deals with how this change in arithmetic affects the underlying geometry and related structures.

Introduction

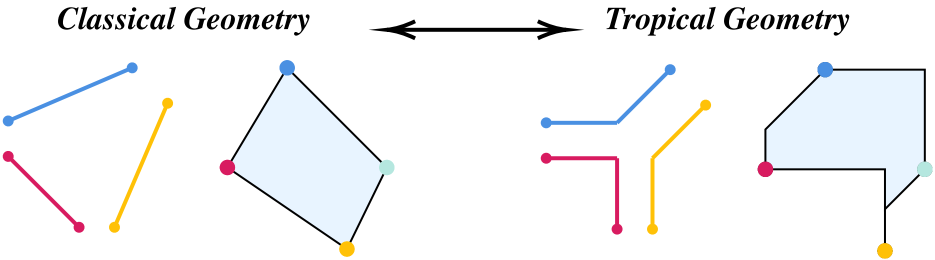
- **Tropical geometry** deals with how this change in arithmetic affects the underlying geometry and related structures.



$ax + by + c$ $\max(a + x,\ b + y,\ c)$

Introduction

- **Tropical geometry** deals with how this change in arithmetic affects the underlying geometry and related structures.



$ax + by + c$ $\max(a + x,\ b + y,\ c)$

- This project was an exploration of how tropical methods could be applied to inexact computing.

Los Alamos NATIONAL LABORATORY

Tropical Matrices

- We can extend tropical arithmetic to matrices

Tropical Matrices

- We can extend tropical arithmetic to matrices

$$\begin{bmatrix} \mathbf{1} & \mathbf{2} \\ 3 & 4 \end{bmatrix} \odot \begin{bmatrix} \mathbf{5} & 3 \\ \mathbf{2} & 0 \end{bmatrix} = \begin{bmatrix} \mathbf{max}(\mathbf{1}+\mathbf{5}, \mathbf{2}+\mathbf{2}) & \max(1+3, 2+0) \\ \max(3+5, 4+2) & \max(3+3, 4+0) \end{bmatrix} = \begin{bmatrix} \mathbf{6} & 4 \\ 8 & 6 \end{bmatrix}$$

Tropical Matrices

- We can extend tropical arithmetic to matrices

$$\begin{bmatrix} \mathbf{1} & \mathbf{2} \\ 3 & 4 \end{bmatrix} \odot \begin{bmatrix} \mathbf{5} & 3 \\ \mathbf{2} & 0 \end{bmatrix} = \begin{bmatrix} \mathbf{max}(\mathbf{1}+\mathbf{5}, \mathbf{2}+\mathbf{2}) & \max(1+3, 2+0) \\ \max(3+5, 4+2) & \max(3+3, 4+0) \end{bmatrix} = \begin{bmatrix} \mathbf{6} & 4 \\ 8 & 6 \end{bmatrix}$$

- Perturbing some of the entries may not affect the product:

Tropical Matrices

- We can extend tropical arithmetic to matrices

$$\begin{bmatrix} \mathbf{1} & \mathbf{2} \\ 3 & 4 \end{bmatrix} \odot \begin{bmatrix} \mathbf{5} & 3 \\ \mathbf{2} & 0 \end{bmatrix} = \begin{bmatrix} \mathbf{max(1+5,2+2)} & \max(1+3,2+0) \\ \max(3+5,4+2) & \max(3+3,4+0) \end{bmatrix} = \begin{bmatrix} \mathbf{6} & 4 \\ 8 & 6 \end{bmatrix}$$

- Perturbing some of the entries may not affect the product:

$$\begin{bmatrix} \mathbf{1} & \mathbf{3} \\ 3 & 4 \end{bmatrix} \odot \begin{bmatrix} \mathbf{5} & 3 \\ \mathbf{2} & 0 \end{bmatrix} = \begin{bmatrix} \mathbf{max(1+5,3+2)} & \max(1+3,3+0) \\ \max(3+5,4+2) & \max(3+3,4+0) \end{bmatrix} = \begin{bmatrix} \mathbf{6} & 4 \\ 8 & 6 \end{bmatrix}$$

Tropical Matrices

- We can extend tropical arithmetic to matrices

$$\begin{bmatrix} \mathbf{1} & \mathbf{2} \\ 3 & 4 \end{bmatrix} \odot \begin{bmatrix} \mathbf{5} & 3 \\ \mathbf{2} & 0 \end{bmatrix} = \begin{bmatrix} \mathbf{max}(\mathbf{1} + \mathbf{5}, \mathbf{2} + \mathbf{2}) & \max(1 + 3, 2 + 0) \\ \max(3 + 5, 4 + 2) & \max(3 + 3, 4 + 0) \end{bmatrix} = \begin{bmatrix} \mathbf{6} & 4 \\ 8 & 6 \end{bmatrix}$$

- Perturbing some of the entries may not affect the product:

$$\begin{bmatrix} \mathbf{1} & \mathbf{3} \\ 3 & 4 \end{bmatrix} \odot \begin{bmatrix} \mathbf{5} & 3 \\ \mathbf{2} & 0 \end{bmatrix} = \begin{bmatrix} \mathbf{max}(\mathbf{1} + \mathbf{5}, \mathbf{3} + \mathbf{2}) & \max(1 + 3, 3 + 0) \\ \max(3 + 5, 4 + 2) & \max(3 + 3, 4 + 0) \end{bmatrix} = \begin{bmatrix} \mathbf{6} & 4 \\ 8 & 6 \end{bmatrix}$$

- *Many* entries could be changed without affecting the product.

Los Alamos NATIONAL LABORATORY

Tropical Matrices

- We can extend tropical arithmetic to matrices

$$\begin{bmatrix} \mathbf{1} & \mathbf{2} \\ 3 & 4 \end{bmatrix} \odot \begin{bmatrix} \mathbf{5} & 3 \\ \mathbf{2} & 0 \end{bmatrix} = \begin{bmatrix} \mathbf{max(1+5, 2+2)} & \max(1+3, 2+0) \\ \max(3+5, 4+2) & \max(3+3, 4+0) \end{bmatrix} = \begin{bmatrix} \mathbf{6} & 4 \\ 8 & 6 \end{bmatrix}$$

- Perturbing some of the entries may not affect the product:

$$\begin{bmatrix} \mathbf{1} & \mathbf{3} \\ 3 & 4 \end{bmatrix} \odot \begin{bmatrix} \mathbf{5} & 3 \\ \mathbf{2} & 0 \end{bmatrix} = \begin{bmatrix} \mathbf{max(1+5, 3+2)} & \max(1+3, 3+0) \\ \max(3+5, 4+2) & \max(3+3, 4+0) \end{bmatrix} = \begin{bmatrix} \mathbf{6} & 4 \\ 8 & 6 \end{bmatrix}$$

- *Many* entries could be changed without affecting the product.
- The usual factorization techniques **don't work here**.

Los Alamos NATIONAL LABORATORY

Matrix Factorization

- Tropical matrix factorization problem:

  Given a matrix M, find matrices A, B so that $M = A \odot B$.

Matrix Factorization

- Tropical matrix factorization problem:

  Given a matrix M, find matrices A, B so that $M = A \odot B$.

- **Only two known methods** can tackle this problem:
  - `Tropical Matrix Factorization (1997)` [1]:
    - First known approach
    - Applied to discrete event systems
    - Convergence is guaranteed and is a polynomial-time method

Matrix Factorization

- Tropical matrix factorization problem:

$$\textit{Given a matrix } M, \textit{ find matrices } A, B \textit{ so that } M = A \odot B.$$

- **Only two known methods** can tackle this problem:
  - `Tropical Matrix Factorization (1997)` [1]:
    - First known approach
    - Applied to discrete event systems
    - Convergence is guaranteed and is a polynomial-time method
  - `Sparse Tropical Matrix Factorization (2021)` [3]:
    - Modification of `TMF` that introduces sparsity
    - Applied to matrix completion and prediction tasks
    - Convergence and complexity are inherited from `TMF`

Matrix Factorization

- Tropical matrix factorization problem:

$$\textit{Given a matrix } M\textit{, find matrices } A, B \textit{ so that } M = A \odot B.$$

- **Only two known methods** can tackle this problem:
  - `Tropical Matrix Factorization (1997)` [1]:
    - First known approach
    - Applied to discrete event systems
    - Convergence is guaranteed and is a polynomial-time method
  - `Sparse Tropical Matrix Factorization (2021)` [3]:
    - Modification of `TMF` that introduces sparsity
    - Applied to matrix completion and prediction tasks
    - Convergence and complexity are inherited from `TMF`
- Both are *very* slow, but great at capturing non-linear structures.

Los Alamos NATIONAL LABORATORY

Benefits of Working in This Setting

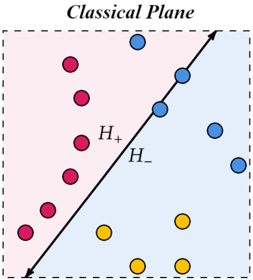
- Tropical arithmetic is **fast** and **robust**.

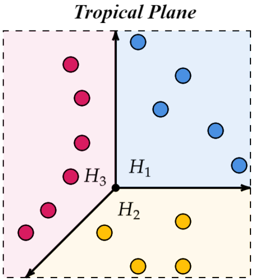
Benefits of Working in This Setting

- Tropical arithmetic is **fast** and **robust**.
- Tropical methods can **encode data with 'extreme' values**.

Benefits of Working in This Setting

- Tropical arithmetic is **fast** and **robust**.
- Tropical methods can **encode data with 'extreme' values**.
- Tropical hyperplanes separate the ambient space into $n+1$ regions–**encoding more information with fewer parameters**.

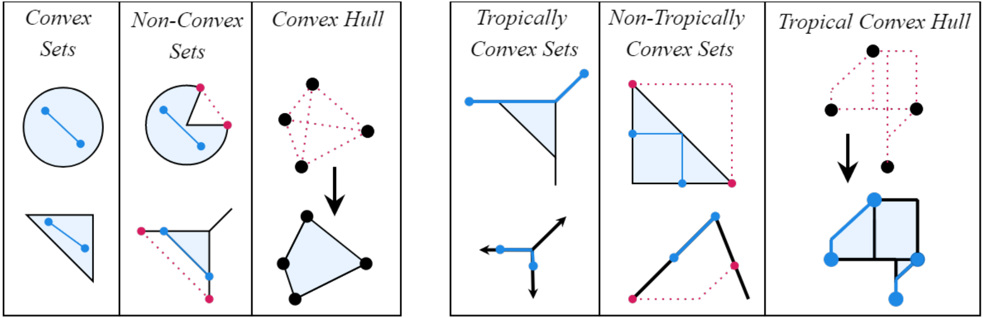




Los Alamos NATIONAL LABORATORY

A Geometric Approach to Matrix Factorization

- Key observation:
  - *The columns (or rows) of M can be identified with points that generate a* **tropical convex hull** *tconv*(M).



A Geometric Approach to Matrix Factorization

- This approach **leans on known methods** for computing convex hulls in the classical setting, which are very fast.

A Geometric Approach to Matrix Factorization

- This approach **leans on known methods** for computing convex hulls in the classical setting, which are very fast.
- The special case described in this talk allows for significantly **faster factorization**.

A Geometric Approach to Matrix Factorization

- This approach **leans on known methods** for computing convex hulls in the classical setting, which are very fast.
- The special case described in this talk allows for significantly **faster factorization**.
- The left factor A that results from this method is **stable under translations** of M; only B varies as M is perturbed.

A Geometric Approach to Matrix Factorization

- This approach **leans on known methods** for computing convex hulls in the classical setting, which are very fast.
- The special case described in this talk allows for significantly **faster factorization**.
- The left factor A that results from this method is **stable under translations** of M; only B varies as M is perturbed.
- This approach **can be modified** such as with `STMF` in order to be applied to prediction tasks.

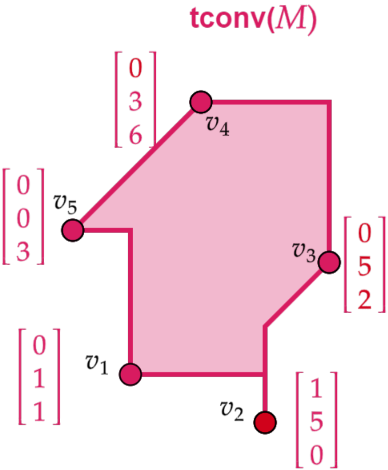
A Geometric Approach to Matrix Factorization

- This approach **leans on known methods** for computing convex hulls in the classical setting, which are very fast.
- The special case described in this talk allows for significantly **faster factorization**.
- The left factor A that results from this method is **stable under translations** of M; only B varies as M is perturbed.
- This approach **can be modified** such as with `STMF` in order to be applied to prediction tasks.
- This approach has a clear **geometric interpretation**.

Overview of the Geometric Method

- The columns of M can be identified with the tropical convex hull tconv(M).

$$M = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 5 & 5 & 3 & 0 \\ 1 & 0 & 2 & 6 & 3 \end{bmatrix} \quad \leftrightsquigarrow$$

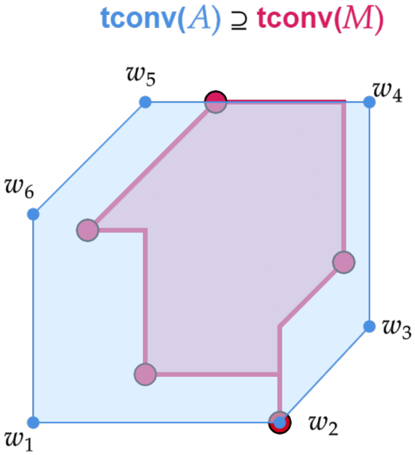


Overview of the Geometric Method

- The columns of M can be identified with the tropical convex hull tconv(M).
- Which is contained in some tconv(A).

$$M = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 5 & 5 & 3 & 0 \\ 1 & 0 & 2 & 6 & 3 \end{bmatrix} \leftrightsquigarrow$$

$$A = \begin{bmatrix} w_1 & w_2 & \cdots & w_6 \end{bmatrix} \leftrightsquigarrow$$



Los Alamos NATIONAL LABORATORY

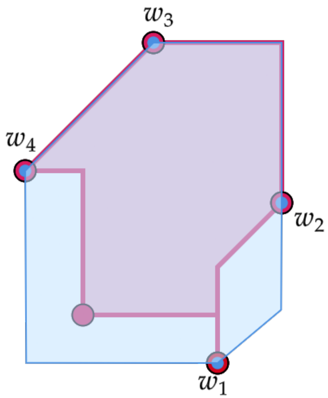
Overview of the Geometric Method

- The columns of M can be identified with the tropical convex hull tconv(M).
- Which is contained in some tconv(A)...

$$M = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 5 & 5 & 3 & 0 \\ 1 & 0 & 2 & 6 & 3 \end{bmatrix} \quad \leftrightsquigarrow$$

$$A = \begin{bmatrix} w_1 & w_2 & w_3 & w_4 \end{bmatrix} \quad \leftrightsquigarrow$$

tconv(A) ⊇ tconv(M)



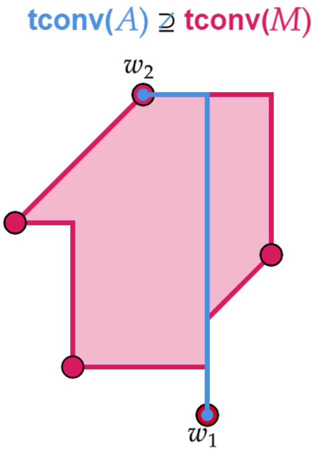
Los Alamos NATIONAL LABORATORY

Overview of the Geometric Method

- The columns of M can be identified with the tropical convex hull tconv(M).
- Which is contained in some tconv(A)......

$$M = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 5 & 5 & 3 & 0 \\ 1 & 0 & 2 & 6 & 3 \end{bmatrix} \longleftrightarrow$$

$$A = \begin{bmatrix} w_1 & w_2 \end{bmatrix} \longleftrightarrow$$



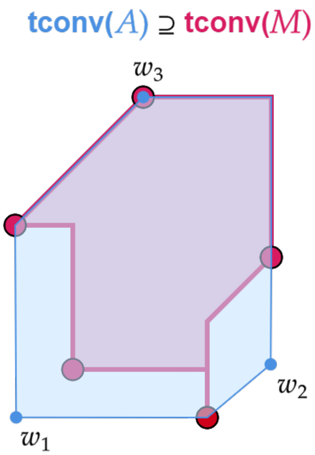
Los Alamos NATIONAL LABORATORY

Overview of the Geometric Method

- The columns of M can be identified with the tropical convex hull tconv(M).
- Which is contained in some tconv(A).
- There is some optimal choice for A depending on the inner dimension r.

$$M = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 5 & 5 & 3 & 0 \\ 1 & 0 & 2 & 6 & 3 \end{bmatrix} \quad \leftrightsquigarrow$$

$$A = [\, w_1 \;\; w_2 \;\; w_3 \,] = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 5 & 3 \\ 0 & 0 & 6 \end{bmatrix} \quad \leftrightsquigarrow$$



Los Alamos NATIONAL LABORATORY

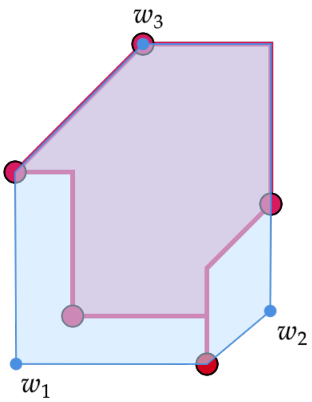
Overview of the Geometric Method

- The columns of M can be identified with the tropical convex hull tconv(M).
- Which is contained in some tconv(A).
- There is some optimal choice for A depending on the inner dimension r.
- For each v_i, compute the solutions to $A \odot b_i = v_i$ to form B.

$$A \odot B = M$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 1 & 5 & 3 \\ 0 & 0 & 6 \end{bmatrix} \odot \begin{bmatrix} b_1 & \cdots & b_5 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 5 & 5 & 3 & 0 \\ 1 & 0 & 2 & 6 & 3 \end{bmatrix}$$

tconv(A) $\supseteq$ tconv(M)



[2]

Los Alamos NATIONAL LABORATORY

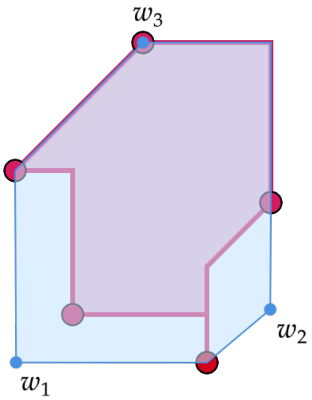
Overview of the Geometric Method

- The columns of M can be identified with the tropical convex hull tconv(M).
- Which is contained in some tconv(A).
- There is some optimal choice for A depending on the inner dimension r.
- For each v_i, compute the solutions to $A \odot b_i = v_i$ to form B.

$$A \odot B = M$$

$$\underset{w_1\, w_2\, w_3}{\begin{bmatrix} 1 & 0 & 0 \\ 1 & 5 & 3 \\ 0 & 0 & 6 \end{bmatrix}} \odot \begin{bmatrix} -1 & 0 & -1 & -1 & -1 \\ 4 & 0 & 0 & -2 & -5 \\ -5 & -6 & -4 & 0 & -3 \end{bmatrix} = \underset{v_1\, v_2\, v_3\, v_4\, v_5}{\begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 5 & 5 & 3 & 0 \\ 1 & 0 & 2 & 6 & 3 \end{bmatrix}}$$

tconv(A) ⊇ tconv(M)



Los Alamos NATIONAL LABORATORY

Performance Results

- Running TMF and GTMF 100 times on random 6×6 matrices:

	TMF	GTMF
Time	3.68884	0.00168
Error	5.18909	$1.21192 \cdot 10^{-14}$
Added Error	0.69285	0.00039
Added Time	$+1.21192 \cdot 10^{-14}$	$-5.41559 \cdot 10^{-5}$

Los Alamos NATIONAL LABORATORY

Performance Results

- Running TMF and GTMF 100 times on random 6×6 matrices:

	TMF	GTMF
Time	3.68884	0.00168
Error	5.18909	$1.21192 \cdot 10^{-14}$
Added Error	0.69285	0.00039
Added Time	$+1.21192 \cdot 10^{-14}$	$-5.41559 \cdot 10^{-5}$

- Perturbing entries of M improved speed in GTMF at minimal cost to accuracy.

Performance Results

- Running TMF and GTMF 100 times on random 6×6 matrices:

	TMF	GTMF
Time	3.68884	0.00168
Error	5.18909	$1.21192 \cdot 10^{-14}$
Added Error	0.69285	0.00039
Added Time	$+1.21192 \cdot 10^{-14}$	$-5.41559 \cdot 10^{-5}$

- Perturbing entries of M improved speed in GTMF at minimal cost to accuracy.
- Most of the error was contained to B in GTMF. A & B were affected in TMF.

Performance Results

- Running TMF and GTMF 100 times on random 6×6 matrices:

	TMF	GTMF
Time	3.68884	0.00168
Error	5.18909	$1.21192 \cdot 10^{-14}$
Added Error	0.69285	0.00039
Added Time	$+1.21192 \cdot 10^{-14}$	$-5.41559 \cdot 10^{-5}$

- Perturbing entries of M improved speed in GTMF at minimal cost to accuracy.
- Most of the error was contained to B in GTMF. A & B were affected in TMF.
- Questions?

Thank You!

Thanks to my mentor Laura Monroe for her guidance and to team members Vanessa Job, Nathan Kodama, Andrew Alexander, and Matthew Broussard for their helpful suggestions and questions that led me to this problem.

Email: joseaortiz@ksu.edu

- B. De Schutter, B. De Moor, *Matrix Factorization and Minimal State Space Realization in the max-plus algebra*, July 1997.
- P. Maragos, E. Theodosis, *Tropical Geometry and Piecewise-Linear Approximation of Curves and Surfaces on Weighted Lattices*, 2019, arXiv: 1912.03891 [cs.LG].
- A. Omanović, H. Kazan, P. Oblak, T. Curk, *Sparse Data Embedding and Prediction by Tropical Matrix Factorization*, Feb. 2021, (http://dx.doi.org/10.1186/s12859-021-04023-9).

Los Alamos NATIONAL LABORATORY

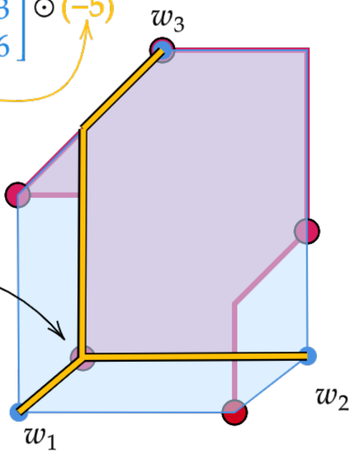
$$\underset{w_1\,w_2\,w_3}{\begin{bmatrix}1&0&0\\1&5&3\\0&0&6\end{bmatrix}} \odot \begin{bmatrix}-1\\-4\\-5\end{bmatrix} = \begin{bmatrix}1\\1\\0\end{bmatrix} \odot (-1) \oplus \begin{bmatrix}0\\5\\0\end{bmatrix} \odot (-4) \oplus \begin{bmatrix}0\\3\\6\end{bmatrix} \odot (-5)$$

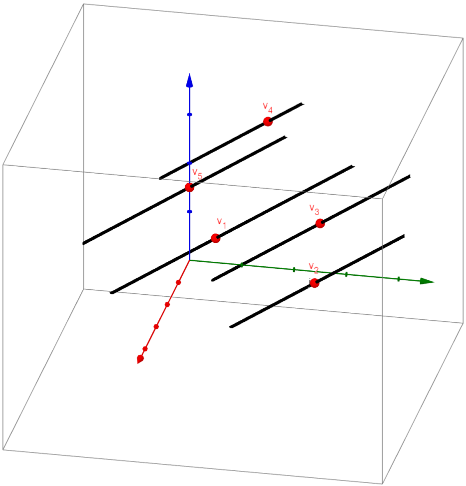
$$= \underset{v_1}{\begin{bmatrix}0\\1\\1\end{bmatrix}}$$

'Lattice distance'

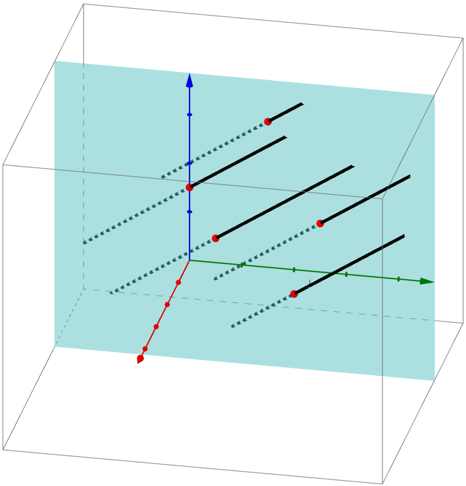
$$A \odot B = M$$

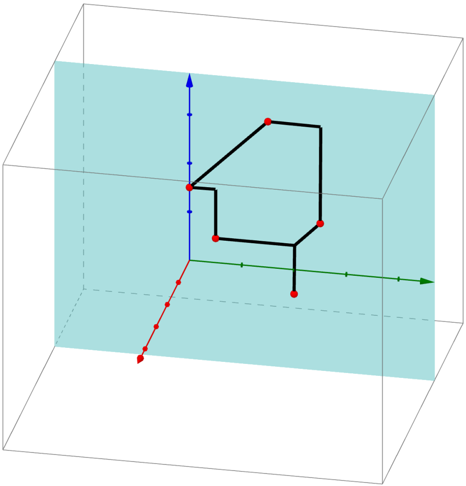
$$\underset{w_1\,w_2\,w_3}{\begin{bmatrix}1&0&0\\1&5&3\\0&0&6\end{bmatrix}} \odot \begin{bmatrix}-1&0&-1&-1&-1\\-4&0&0&-2&-5\\-5&-6&-4&0&-3\end{bmatrix} = \underset{v_1\;v_2\;v_3\;v_4\;v_5}{\begin{bmatrix}0&1&0&0&0\\1&5&5&3&0\\1&0&2&6&3\end{bmatrix}}$$

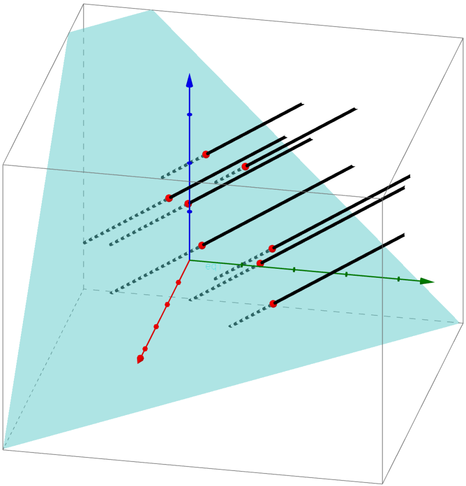


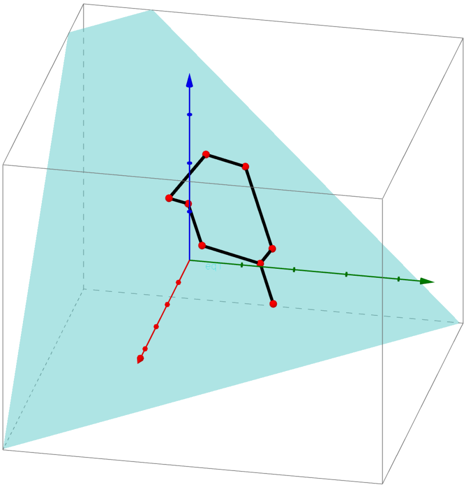


Los Alamos NATIONAL LABORATORY

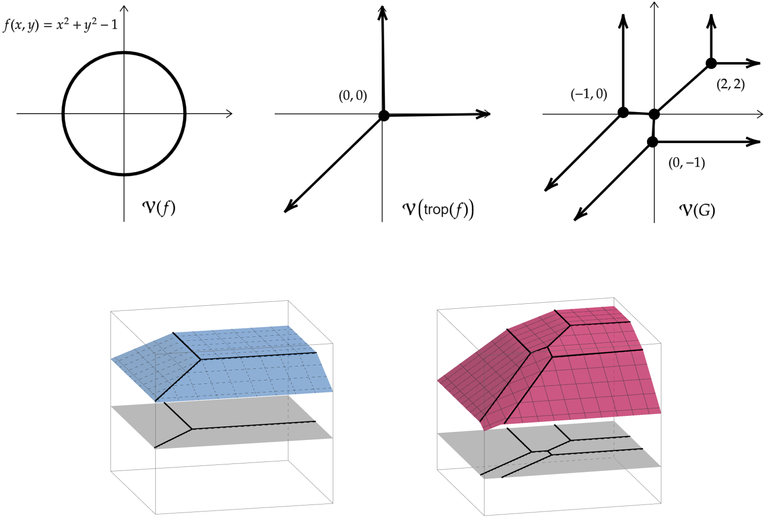








Los Alamos
NATIONAL LABORATORY


$f(x, y) = x^2 + y^2 - 1$
$\mathcal{V}(f)$
(0, 0)
$\mathcal{V}(\mathrm{trop}(f))$
(−1, 0)
(2, 2)
(0, −1)
$\mathcal{V}(G)$