Harvard University
School of Engineering
and Applied Sciences

Tropical mathematics

(and why you should care)

Manos Theodosis

Outline

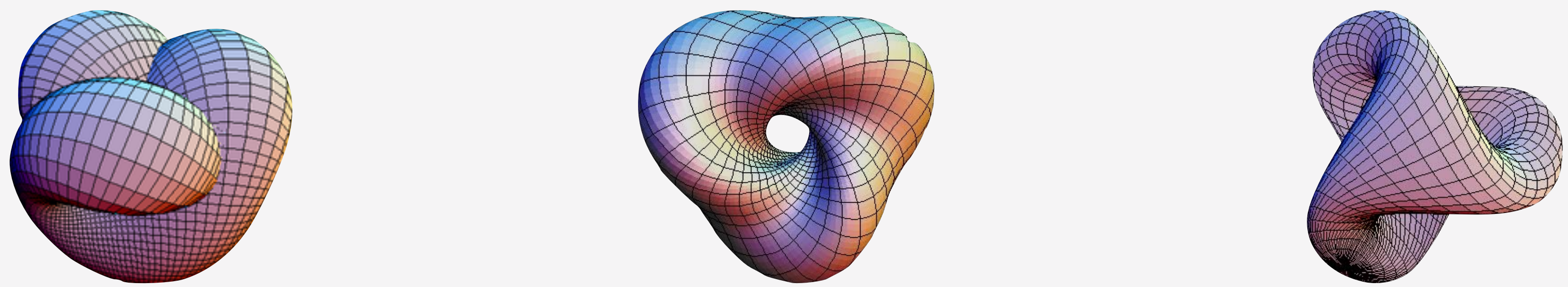
Introduction
Motivation (algebra and geometry)
General definitions
Eigenvalues and polynomials
Dilations and erosions
Curves and surfaces

Applications
Neural networks
Curve fitting
Spectral analysis
Parameter reduction

Introduction

Motivation (geometry)

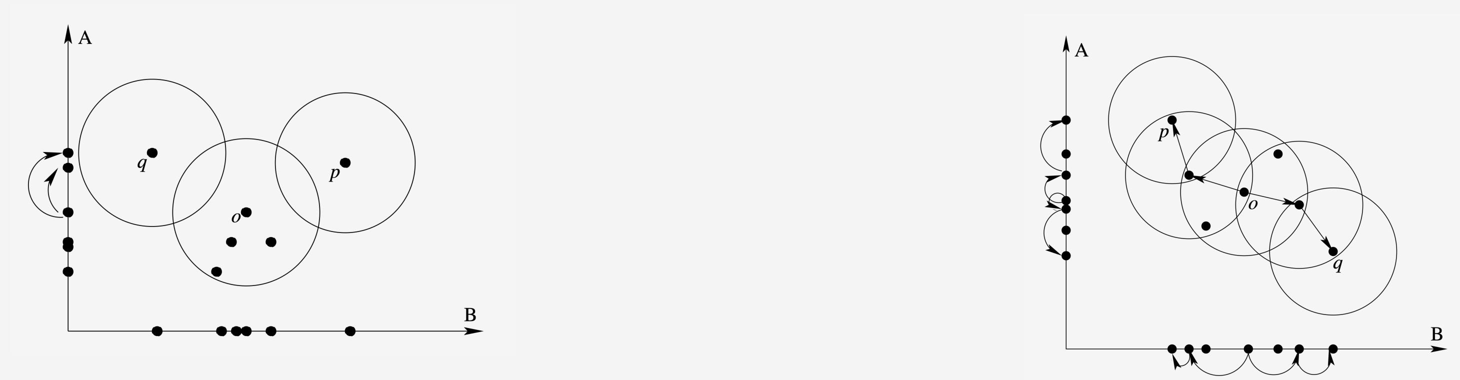
Multidimensional manifolds



(a) (b) (c)

Figure 1
Complex manifolds.

Geometric algorithms

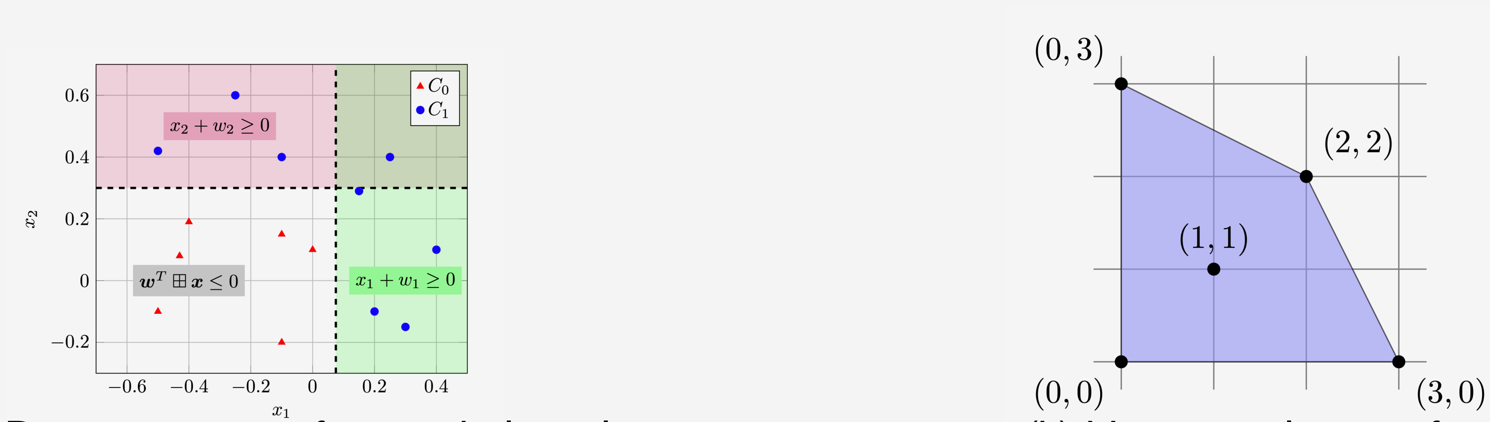


(a) Unconnected components.

(b) Connected components.

Figure 2
DBSCAN algorithm.

Neural networks



(a) Decision region of a morphological neural network.

(b) Newton polytope of a tropical polynomial.

Figure 3
Tropical neural networks and their regions.

Harvard University
School of Engineering
and Applied Sciences

Motivation (algebra)

Shortest path problems

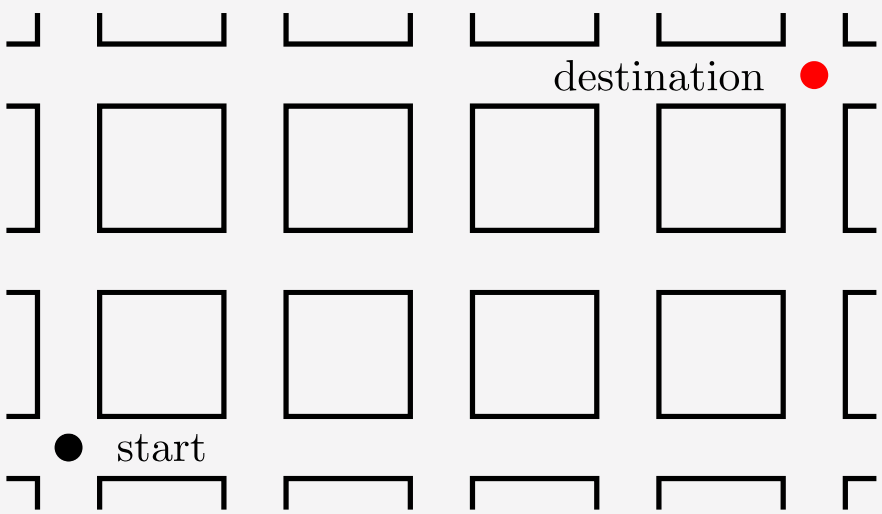


Figure 4
General formulation of a shortest path problem.

Scheduling problems

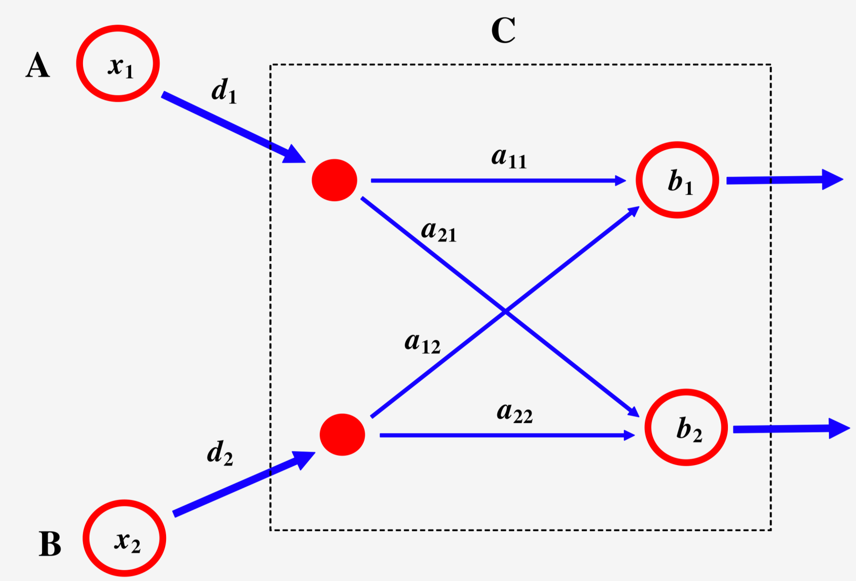


Figure 5
Scheduling problem of connecting flights.

Bonus:
Matrix representation (vector spaces)
Piecewise linear solution spaces

General definitions

Similar to linear algebra, but replace the pair $(+\,,\times)$ with $(\min\,,+)$.

Matrix-vector multiplication becomes

$$\boxed{\left(\mathbf{A}\mathbf{x}\right)_j = \sum_i (A_{ji}\cdot x_i) \quad \Leftrightarrow \quad \left(\mathbf{A} \boxplus' \mathbf{x}\right)_j = \min_i (A_{ji} + x_i)}$$

In general two variants; min-plus algebra and max-plus algebra.

Example

$$\mathbf{A} \boxplus' \mathbf{b} = \begin{bmatrix} 3 & 4 \\ 1 & -2 \\ 6 & 2 \end{bmatrix} \boxplus' \begin{bmatrix} 1 \\ 3 \end{bmatrix} = \begin{bmatrix} \min(3+1,\ 4+3) \\ \min(1+1,\ -2+3) \\ \min(6+1,\ 2+3) \end{bmatrix} = \begin{bmatrix} 4 \\ 1 \\ 5 \end{bmatrix}$$

vs

$$\mathbf{A}\mathbf{b} = \begin{bmatrix} 3 & 4 \\ 1 & -2 \\ 6 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \end{bmatrix} = \begin{bmatrix} (3\cdot 1) + (4\cdot 3) \\ (1\cdot 1) + (-2\cdot 3) \\ (6\cdot 1) + (2\cdot 3) \end{bmatrix} = \begin{bmatrix} 15 \\ -5 \\ 12 \end{bmatrix}$$

Eigenvalues and polynomials

Eigenvalues

Two very important matrices

$$\Gamma(\mathbf{A}) = \min(\mathbf{A}, \mathbf{A}^2, \ldots) = \min_{n=1}^{\infty} \mathbf{A}^n, \quad \Delta(\mathbf{A}) = \min(\mathbf{I}, \mathbf{A}, \mathbf{A}^2, \ldots) = \min_{n=0}^{\infty} \mathbf{A}^n$$

Solutions to the generalized (sub)eigenvalue problems

$$\mathbf{A} \boxplus' \mathbf{x} = \lambda + \mathbf{x}, \quad \mathbf{A} \boxplus' \mathbf{x} \geq \lambda + \mathbf{x}$$

Polynomials

Exponents change meaning

$$\boxed{x^a = x \cdot x \cdot \ldots \cdot x \quad \Leftrightarrow \quad x^a = x + x + \ldots + x = a \cdot x}$$

linear **tropical**

General form of a tropical polynomial

$$y = \min_i a_i x + c_i \Rightarrow y = \min_i \mathbf{a}_i^T \mathbf{x}$$

We can allow for non integer (or negative) exponents

$$y = \min(0.3x + 4, -5x - 5, \frac{1}{5}x)$$

Dilations and erosions

Distributivity

Erosions distribute over infima, dilations over suprema

$$\bigwedge_i \boldsymbol{\varepsilon}(X_i) = \boldsymbol{\varepsilon}\left(\bigwedge_i X_i\right), \quad \bigvee_i \boldsymbol{\delta}(X_i) = \boldsymbol{\delta}\left(\bigvee_i X_i\right)$$

Formulation

A unique adjunction pair

$$\boldsymbol{\delta}(X) \leq Y \Leftrightarrow X \leq \boldsymbol{\varepsilon}(Y)$$

Fundamental property

For any dilation, erosion, lattice element

$$\boldsymbol{\delta}(\boldsymbol{\varepsilon}(X)) \leq X \leq \boldsymbol{\varepsilon}(\boldsymbol{\delta}(X))$$

Tropical erosions and dilations

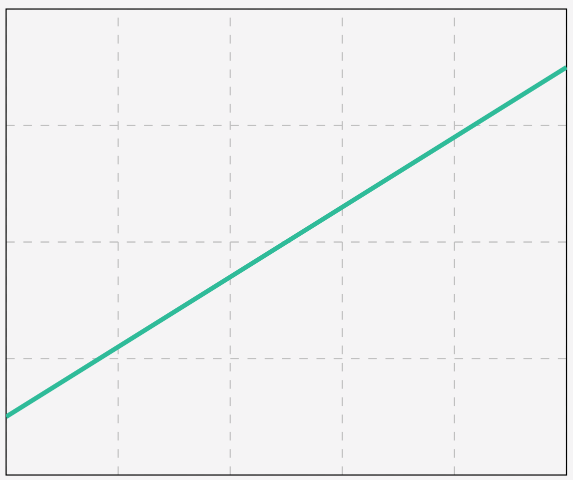
In our context

$$\boxed{\boldsymbol{\delta}_{\mathbf{A}}(\mathbf{x}) = \mathbf{A} \boxplus \mathbf{x}, \quad \boldsymbol{\varepsilon}_{\mathbf{A}}(\mathbf{x}) = \mathbf{A}^* \boxplus' \mathbf{x}}$$

where $\mathbf{A}^* = -\mathbf{A}^T$ and $\boxplus$ is the max-plus multiplication.

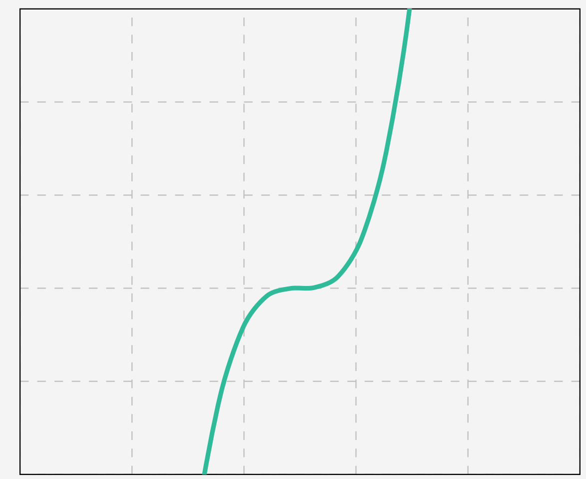
Curves: Polynomials

Euclidean curves



(a) $y = ax + b$

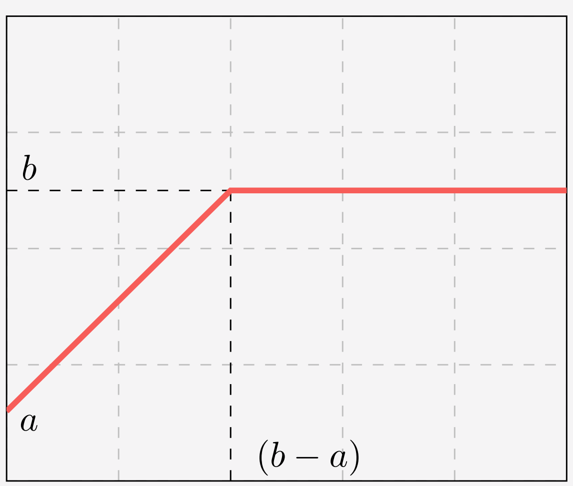
(b) $y = ax^2 + bx + c$



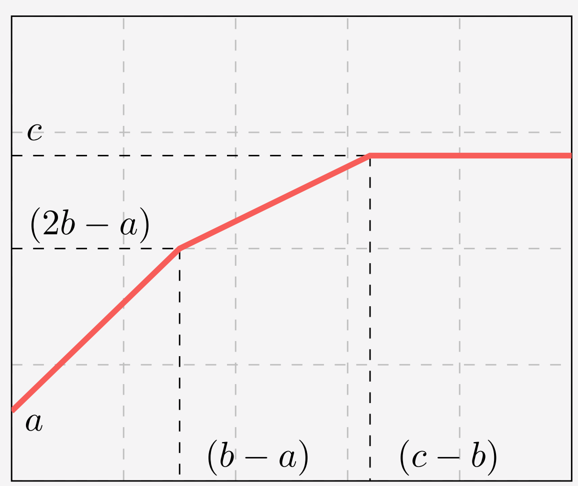
(c) $y = ax^3 + bx^2 + cx + d$

Figure 6
Euclidean polynomials.

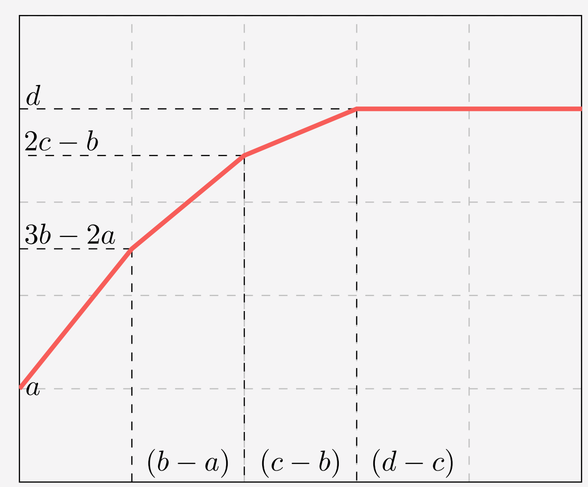
Tropical curves



(a) $y = \min(a + x, b)$



(b) $y = \min(a + 2x, b + x, c)$

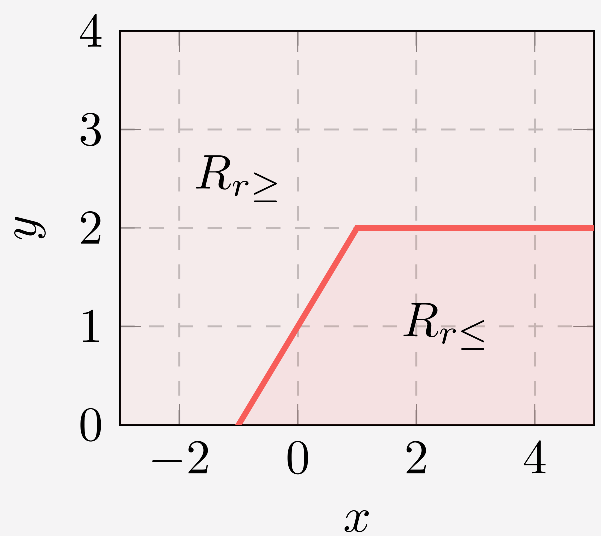


(c) $y = \min(a + 3x, b + 2x, c + x, d)$

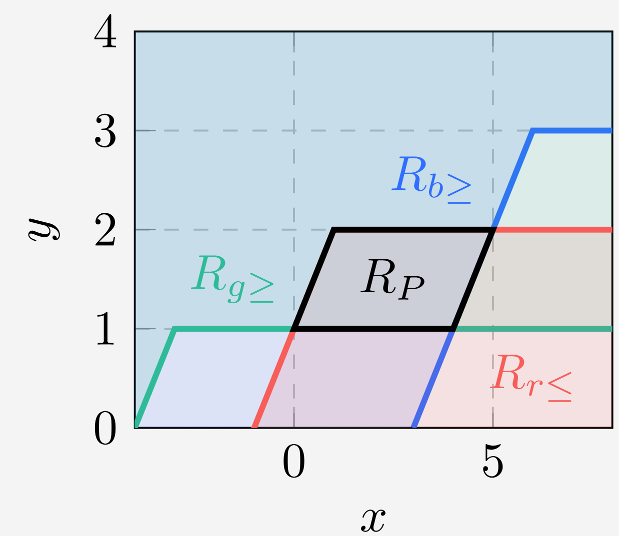
Figure 7
Tropical polynomials.

Curves: Halfspaces and polytopes

1D Halfspaces



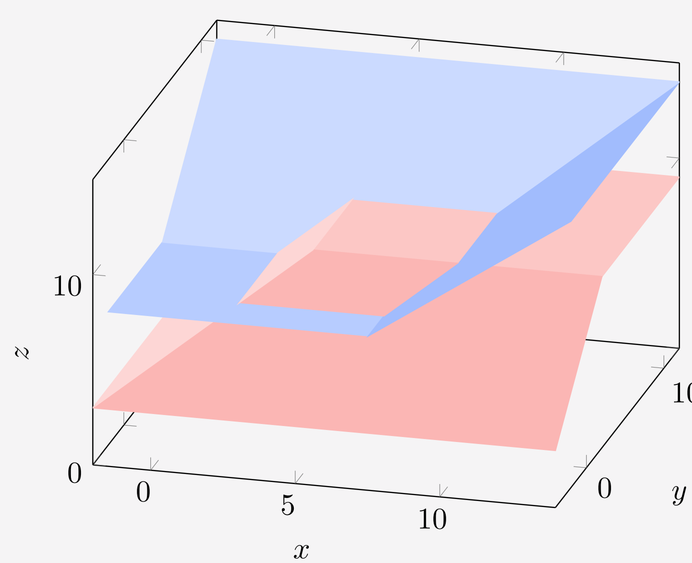
(a) Tropical line and its halfspaces.



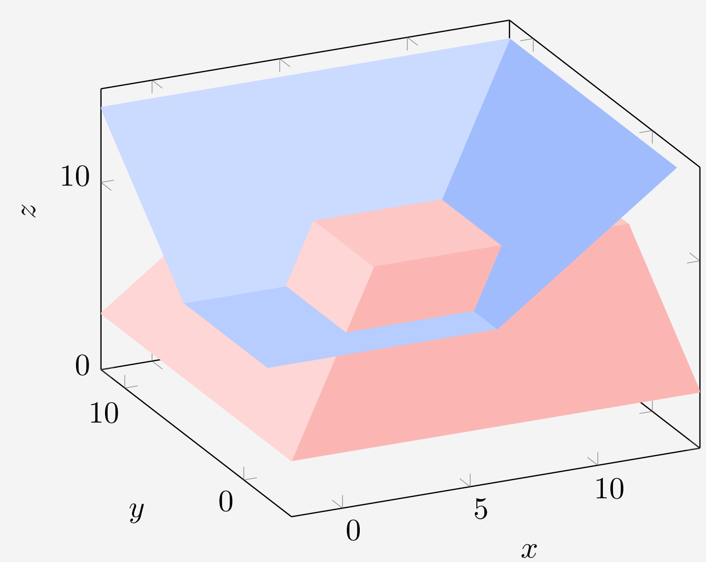
(b) Tropical polytope; intersection of tropical halfspaces.

Figure 8
Halfspaces and polytopes in 1D. For example, the red line is $y = \min(1 + x, 2)$.

2D Halfspaces



(a) First view.



(b) Second view.

Figure 9
Tropical polytopes in 2D. Blue is $z_1 = \max(0 + x, 2 + y, 7)$ and red is $z_2 = \min(5 + x, 7 + y, 9)$.

Curves: Tropical Loci

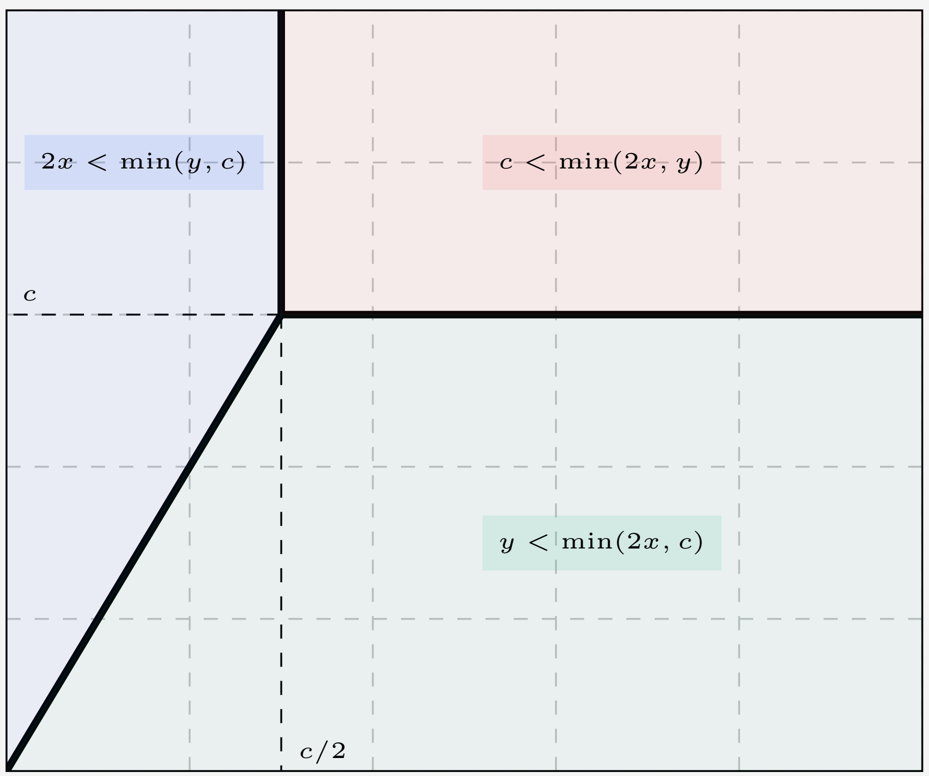


Figure 10
Tropical locus of the equation $y = \min(2x, c)$ and the resulting space clustering.

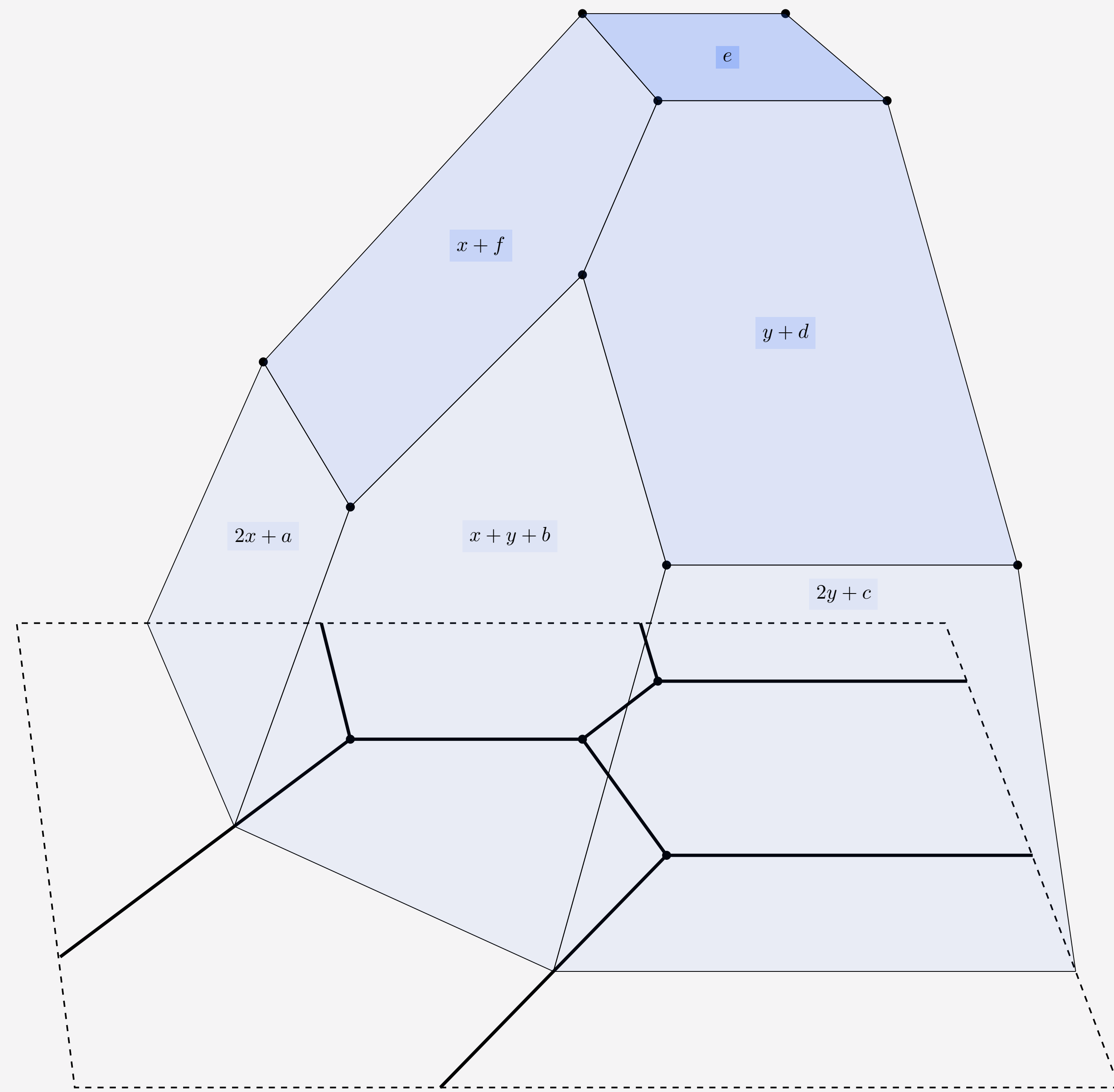


Figure 11
Tropical locus of the general quadratic tropical 2-D polynomial $z = \min(2x + a, x + y + b, 2y + c, y + d, x + f, e)$.

Applications

Neural networks

Neural networks

Traditional morphological perceptron

$$\tau(\mathbf{x}) = \max_i w_i + x_i = \mathbf{w}^T \boxplus \mathbf{x}$$

ReLUs and maxout units are tropical polynomials

$$\boxed{\mathrm{ReLU}(\mathbf{x}) = \max(0, \mathbf{w}^T\mathbf{x} + b), \quad \mathrm{maxout}(\mathbf{x}) = \max_j(\mathbf{W}_j^T\mathbf{x} + b_j)}$$

Idea: find a bound the number of linear regions/vertices of the solution space.

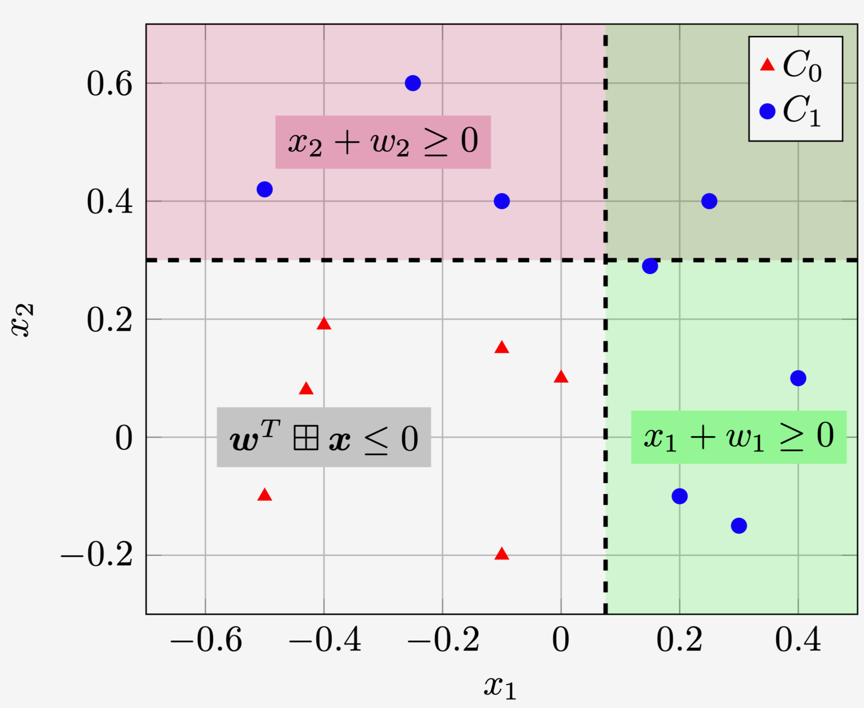


Figure 12
Regions of a morphological perceptron for binary classification.

Neural networks: Linear regions

Tropical polynomials

Remember the definition

$$\boxed{y = \min_i \mathbf{a}_i^T \mathbf{x}}$$

Newton polytope

Defined as the convex hull of the coefficient vectors $\mathbf{a}_i$. How do operations on polynomials affect the Newton polytope?

- $\mathrm{Newt}(y_1 + y_2) = P_1 \oplus P_2$ **Minkwoski sum**
- $\mathrm{Newt}(\max(y_1, y_2)) = \mathrm{conv}(P_1, P_2)$

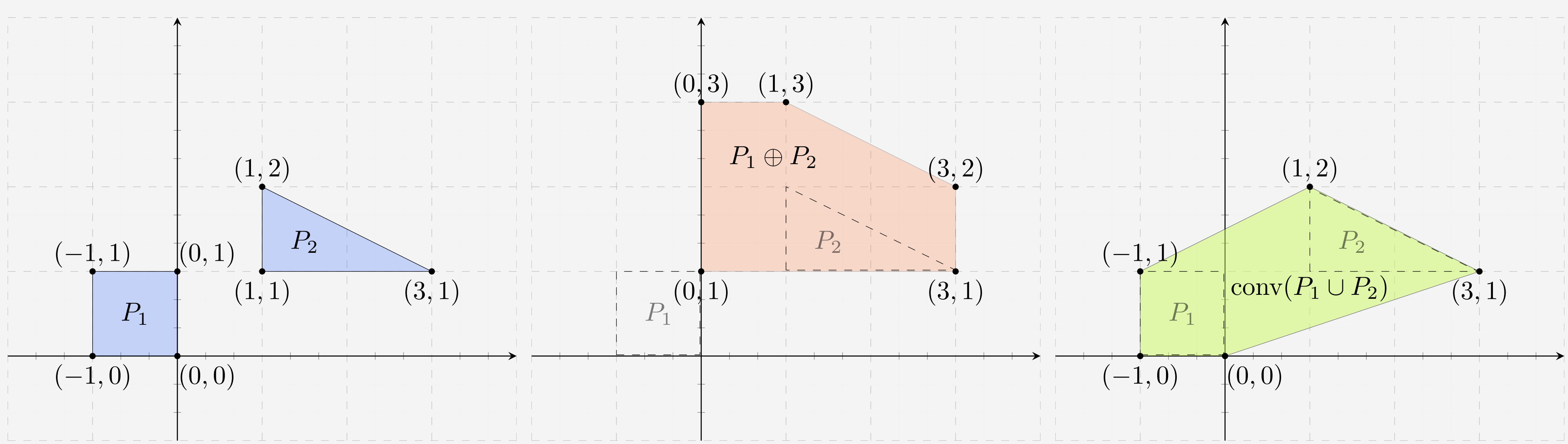


Figure 13
Newton polytopes of the tropical polynomials $z_1 = \max(-x, -x + y, y, 0)$ and $z_2 = \max(x + y, 3x + y, x + 2y)$, their Minkwoski sum, and their convex hull.

Harvard University
School of Engineering
and Applied Sciences

Curve fitting

Modeling
Suppose we fit a model

$$\boxed{y = \max(\mathbf{a}_1^T\mathbf{x} + w_1, \ldots, \mathbf{a}_k^T\mathbf{x} + w_k)}$$ **hyperplanes**

For more data

$$\mathbf{X} \boxplus \mathbf{w} \leq \mathbf{b}$$

Fundamental property
Remember that

$$\boldsymbol{\delta}(\boldsymbol{\varepsilon}(X)) \leq X \leq \boldsymbol{\varepsilon}(\boldsymbol{\delta}(X))$$

In the tropical case

$$\boldsymbol{\delta}_{\mathbf{X}}(\boldsymbol{\varepsilon}_{\mathbf{X}}(\mathbf{b})) \leq \mathbf{b} \Leftrightarrow \mathbf{X} \boxplus (-\mathbf{X}^T \boxplus' \mathbf{b}) \leq \mathbf{b}$$

Optimal solutions
Optimal solution for $\mathbf{w}$

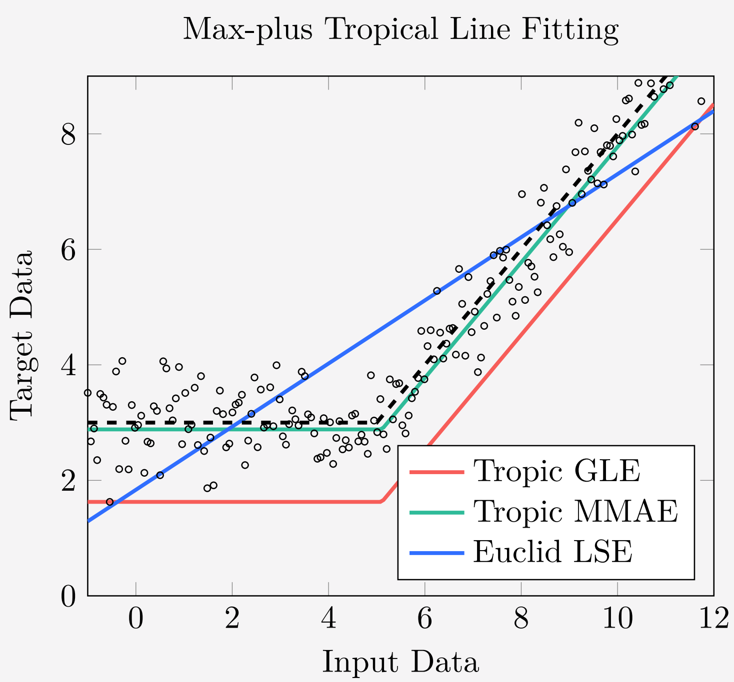
$$\boxed{\hat{\mathbf{w}} = -\mathbf{X}^T \boxplus' \mathbf{b}}$$

Optimal unconstrained solution

$$\boxed{\hat{\mathbf{w}}_\infty = -\mathbf{X}^T \boxplus' \mathbf{b} + \mu} \qquad \mu = \frac{1}{2}\|\mathbf{X} \boxplus \hat{\mathbf{w}} - \mathbf{b}\|_\infty$$

Curve fitting: 1D examples

Tropical models



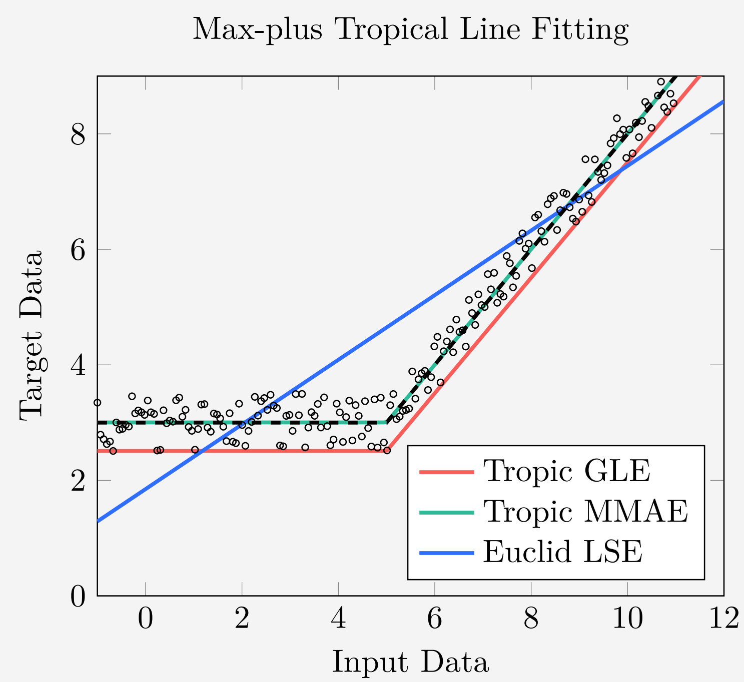


Figure 14
Euclidean vs tropical fitting for the max-plus line $y = \max(x - 2, 3)$ under different noises.

Arbitrary models

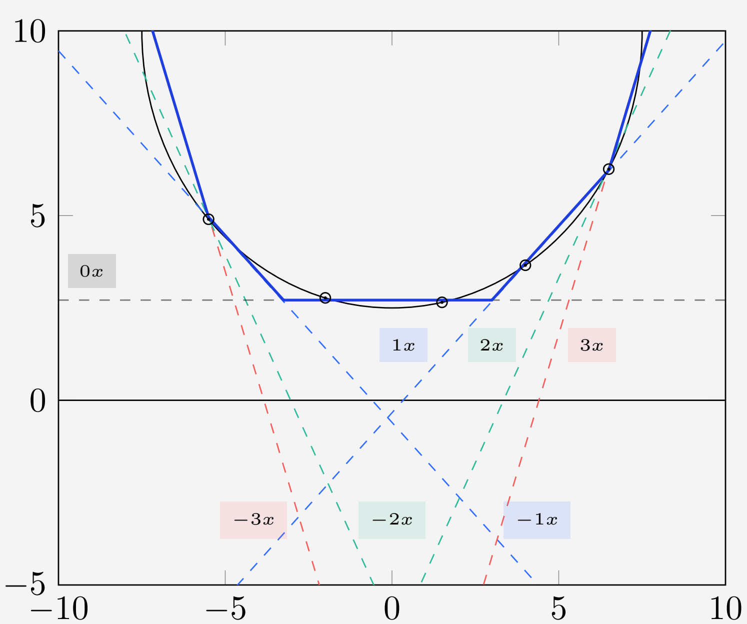
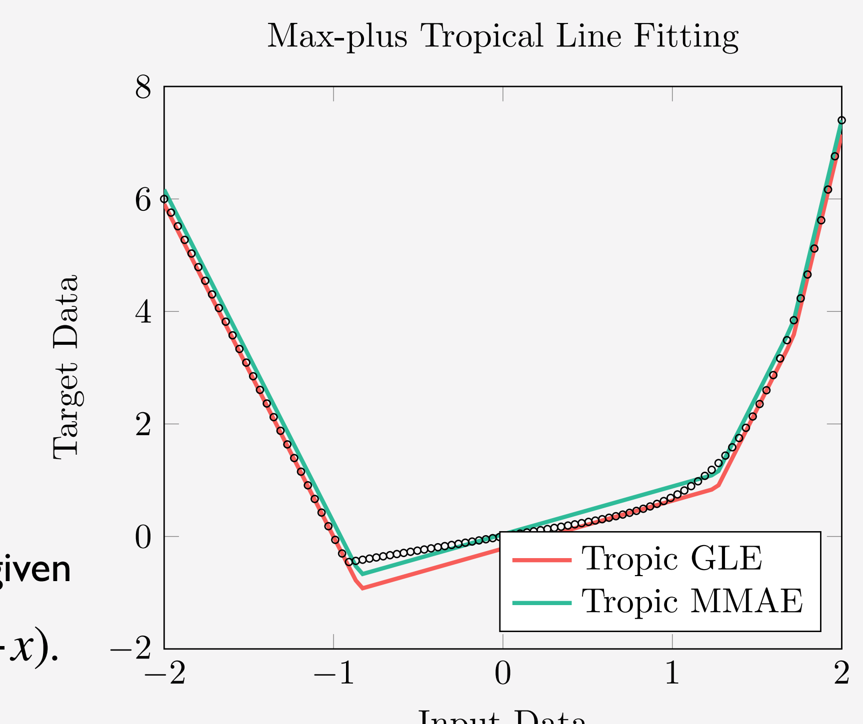


Figure 15
Toy circle fitting; highlights the various tropical lines.

Figure 16
Optimal fitting of the max-afine curve given by $y = \max(-6x - 6, \frac{1}{2}x, \frac{1}{5}x^5 + \frac{1}{2}x)$.



Harvard University
School of Engineering
and Applied Sciences

Curve fitting: 2D examples

2D surfaces

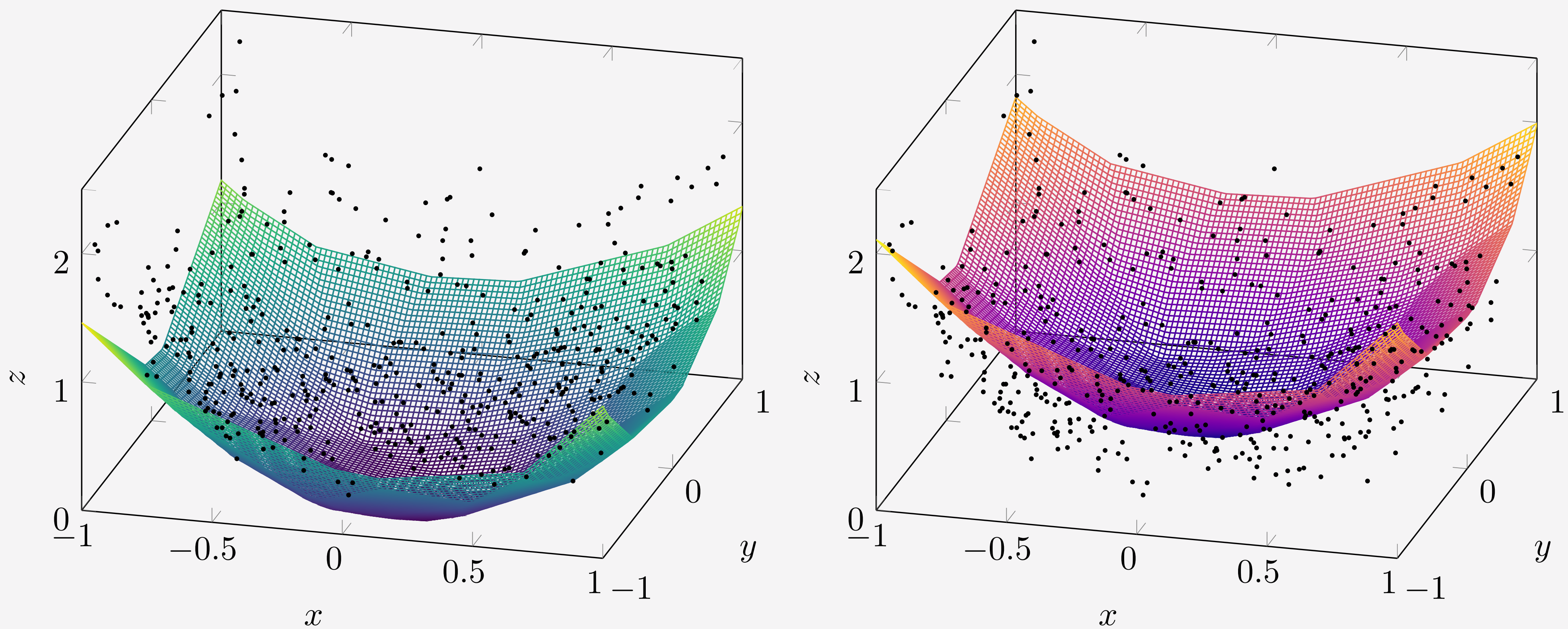


Figure 17
Optimal fitting (left) of parabolic data and the corresponding unconstrained solution (right).

Questions

Number of terms?

Complexity?

How to compute the slopes?

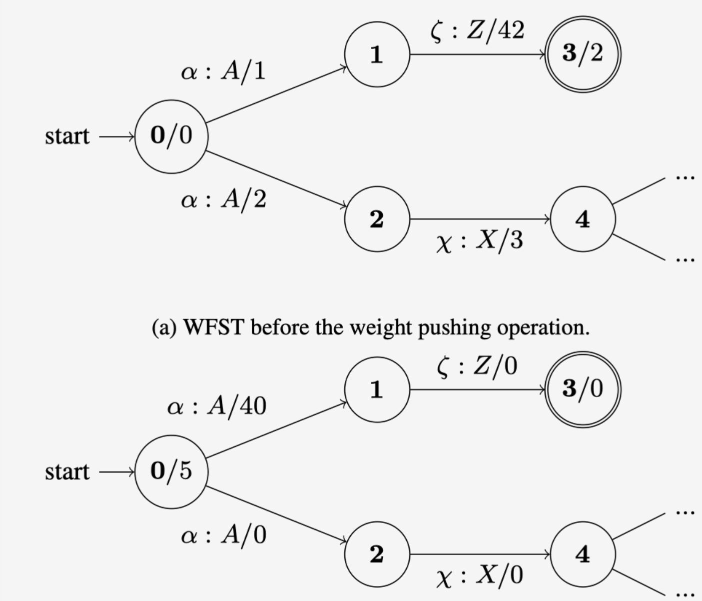
Spectral Analysis of WFSTs

Weight pushing

Can be written in tropical algebra

$$\boxed{\boldsymbol{\lambda}' = \boldsymbol{\Lambda} \boxplus' \mathbf{v}_\infty, \quad \boldsymbol{\rho}' = \mathbf{P} \boxplus' (-\mathbf{v}_\infty), \quad \mathbf{A}' = \mathbf{V}^- \boxplus' \mathbf{A} \boxplus' \mathbf{V}^+}$$

with $\mathbf{v}_\infty = \Delta(\mathbf{A}) \boxplus' \boldsymbol{\rho}$.



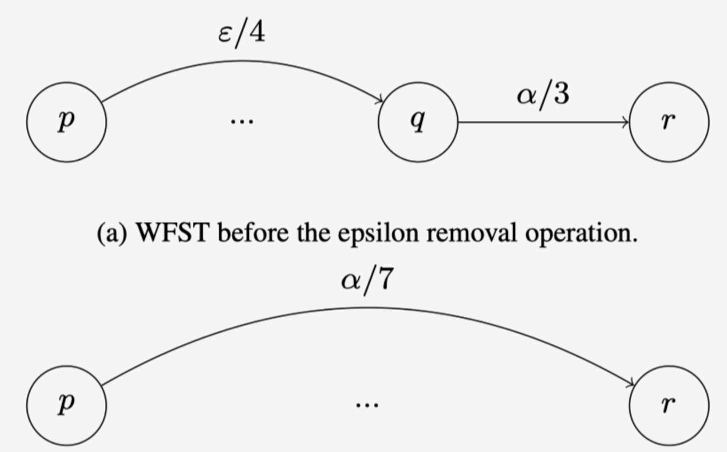
(a) WFST before the weight pushing operation.

Figure 18
Example showcasing the weight pushing algorithm for WFSTs.

Epsilon removal

Can be written in tropical algebra

$$\boxed{\mathbf{A}' = \Delta(\mathbf{E}) \boxplus' \mathbf{A}_\varepsilon, \quad \boldsymbol{\rho}' = \Delta(\mathbf{E}) \boxplus' \boldsymbol{\rho}}$$



(a) WFST before the epsilon removal operation.

(b) WFST after the epsilon removal operation.

Figure 19
Example showcasing the epsilon removal algorithm for WFSTs.

Do they remind you something?

Spectral Analysis of WFSTs

Eigenproblems

$\Delta(\mathbf{A})$ provides the solution to

$$\boxed{\mathbf{A} \boxplus' \mathbf{x} \geq \lambda + \mathbf{x}}$$

Solution set is

$$\boxed{V^*(\mathbf{A}, \lambda) = \{\Delta(-\lambda + \mathbf{A}) \boxplus' \mathbf{u}, \quad \mathbf{u} \in \mathbf{R}^n\}}$$

Interpretation

For weight pushing $\mathbf{v}_\infty = \Delta(\mathbf{A}) \boxplus' \rho$; eigenvalue of zero.

For epsilon removal $\mathbf{A}' = \Delta(\mathbf{E}) \boxplus' \mathbf{A}_\varepsilon, \quad \rho' = \Delta(\mathbf{E}) \boxplus' \rho$; eigenvalues of zero.

Idea: what about other eigenvalues?

Parameter Reduction

Euclidean

What's the minimum number of hyperplanes required for a closed region?

$d = 2$: three lines

$d = 3$: four planes

$d = n$: (?) $n + 1$ hyperplanes

Proof: 1. $d + 1$ hyperplanes can bound d-dimensional space (axes + a d-dimensional hyperplane passing through d points)

2. d hyperplanes can't bound the d-dimensional space (induction)

How many parameters? d parameters per hyperplane; $d(d + 1)$ total parameters.

Tropical

2 hyperplanes of d parameters each; $2d$ total parameters.

Common ground

For a polytope of 2^d vertices

· Euclidean needs $2d^2$ parameters.

· Tropical needs $2d$ parameters.

Thanks!