Solution of Tropical Best Approximation Problems

Nikolai Krivulin

Faculty of Mathematics and Mechanics
St. Petersburg State University
St. Petersburg, Russia

E-mail: nkk<at>math.spbu.ru
URL: http://www.math.spbu.ru/user/krivulin/

Polynomial Computer Algebra 2024 (PCA '2024)
Euler International Mathematical Institute, St. Petersburg, Russia
April 15-20, 2024

Outline

Introduction

- We consider an approximation problem of an unknown function $y = f(x)$ given a set of samples (x_i, y_i) of function input/output
- Let $F(x; \boldsymbol{\theta})$ be an approximating function that depends on the vector $\boldsymbol{\theta}$ of unknown parameters that are to be determined
- The purpose is to find a best minimax approximate solution, which is defined in the sense of a distance function d as follows:

$$\boldsymbol{\theta}_* = \arg\min_{\boldsymbol{\theta}} \max_i d(F(x_i; \boldsymbol{\theta}), y_i)$$

- We formulate the problem in the framework of a tropical semifield (a semiring with idempotent addition and invertible multiplication)
- As approximating functions, we use tropical Puiseux polynomials (which allow rational exponents) and tropical rational functions
- We apply the results to the best Chebyshev approximation of real functions with piecewise linear functions taken as approximants

Idempotent Algebra: Idempotent Semifield

Idempotent Semifield

- Idempotent semifield: the algebraic system $\langle \mathbb{X}, \mathbb{0}, \mathbb{1}, \oplus, \otimes \rangle$
- The carrier set $\mathbb{X}$ has neutral elements, zero $\mathbb{0}$ and identity $\mathbb{1}$
- The binary operations $\oplus$ and $\otimes$ are associative and commutative
- Multiplication $\otimes$ distributes over addition
- Addition $\oplus$ is idempotent: $x \oplus x = x$ for all $x \in \mathbb{X}$
- Multiplication $\otimes$ is invertible: for each nonzero $x \in \mathbb{X}$, there exists an inverse $x^{-1} \in \mathbb{X}$ such that $x \otimes x^{-1} = \mathbb{1}$
- Linear order: the order $x \leq y \Longleftrightarrow x \oplus y = y$ is a total order
- Algebraic completeness: the equation $x^p = a$ is solvable for any $a \in \mathbb{X}$ and integer p (there exist powers with rational exponents)
- Notational convention: the multiplication sign $\otimes$ is omitted

Semifield $\mathbb{R}_{\max,+}$ (Max-Plus Algebra)

- Max-Plus algebra: $\mathbb{R}_{\max,+} = \langle \mathbb{R} \cup \{-\infty\}, -\infty, 0, \max, + \rangle$
- Carrier set: $\mathbb{X} = \mathbb{R} \cup \{-\infty\}$; *zero and one:* $\mathbb{0} = -\infty$, $\mathbb{1} = 0$
- Binary operations: $\oplus = \max$ and $\otimes = +$
- Idempotent addition: $x \oplus x = x$ for each $x \in \mathbb{X}$ ($= \max(x, x)$)
- Multiplicative inverse: there exists x^{-1} for each $x \in \mathbb{R}$ ($= -x$)
- Power notation: x^y is well-defined for any $x, y \in \mathbb{R}$ $(= x \times y)$
- More real idempotent semifields:

$$\mathbb{R}_{\min,+} = \langle \mathbb{R} \cup \{+\infty\}, +\infty, 0, \min, + \rangle,$$
$$\mathbb{R}_{\max} = \langle \mathbb{R}_{+} \cup \{0\}, 0, 1, \max, \times \rangle,$$
$$\mathbb{R}_{\min} = \langle \mathbb{R}_{+} \cup \{+\infty\}, +\infty, 1, \min, \times \rangle,$$

where $\mathbb{R}_{+} = \{x \in \mathbb{R} | x > 0\}$

Algebra of Matrices and Vectors

- Matrix operations follow the standard entrywise rules, where addition and multiplication are replaced by $\oplus$ and $\otimes$
- A matrix without zero rows and columns is called regular
- A matrix of a single column (row) is a column (row) vector
- If a vector has no zero elements, it is called regular
- For a column vector $\boldsymbol{x} = (x_j)$, its conjugate is a row vector $\boldsymbol{x}^- = (x_j^-)$ with $x_j^- = x_j^{-1}$ if $x_j \neq \mathbb{0}$, and $x_j^- = \mathbb{0}$ otherwise
- For regular vectors $\boldsymbol{x} = (x_j)$ and $\boldsymbol{y} = (y_j)$, a metric is given by

$$d(\boldsymbol{x}, \boldsymbol{y}) = \bigoplus_j \left(x_j y_j^{-1} \oplus x_j^{-1} y_j \right) = \boldsymbol{y}^- \boldsymbol{x} \oplus \boldsymbol{x}^- \boldsymbol{y}.$$

- In the context of $\mathbb{R}_{\max,+}$, this metric is the Chebyshev metric

$$d_\infty(\boldsymbol{x}, \boldsymbol{y}) = \max_j |x_j - y_j| = \max_j \max(x_j - y_j, y_j - x_j)$$

Polynomials and Rational Functions

Tropical Puiseux Polynomials

- A tropical Puiseux polynomial of n monomials is given by

$$P(x) = \bigoplus_{j=1}^{n} \theta_j x^{p_j}, \qquad x \neq \mathbb{0},$$

  where $p_j \in \mathbb{Q}$ are exponents and $\theta_j \in \mathbb{X}$, $\theta_j \neq \mathbb{0}$, are coefficients

- When defined in the context of $\mathbb{R}_{\max,+}$ (max-plus algebra), a polynomial is represented in terms of the usual operations as

$$P(x) = \max_{1 \leq j \leq n} (p_j x + \theta_j),$$

  and therefore specifies a piecewise-linear convex function on $\mathbb{R}$

Tropical Rational Functions

- A tropical rational function is defined by two polynomials as

$$R(x) = \frac{P(x)}{Q(x)}, \quad P(x) = \bigoplus_{j=1}^{n} \theta_j x^{p_j}, \quad Q(x) = \bigoplus_{k=1}^{l} \sigma_k x^{q_k}, \quad x \neq \mathbb{0}$$

- In the framework of $\mathbb{R}_{\max,+}$, the rational function can be written as

$$R(x) = P(x) - Q(x) = \max_{1 \leq j \leq n} (p_j x + \theta_j) - \max_{1 \leq k \leq l} (q_k x + \sigma_k),$$

  which presents a difference of piecewise linear convex functions

- We observe that any arbitrary continuous function can be represented as the difference of two convex functions

Best Approximate Solutions: One-Sided Equation

- Given an $(m \times n)$-matrix $\boldsymbol{A}$ and m-vector $\boldsymbol{b}$, the problem is to find regular n-vectors $\boldsymbol{x}$ that solve the one-sided equation

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$$

- The next result offers a best approximate solution to the equation

Theorem (K. 2004, 2009, 2012)

Let $\boldsymbol{A}$ be a regular matrix, $\boldsymbol{b}$ regular vector and $\Delta = (\boldsymbol{A}(\boldsymbol{b}^{-}\boldsymbol{A})^{-})^{-}\boldsymbol{b}$. Then, the following statements hold:

1. *The best approximate error for the equation is equal to $\sqrt{\Delta}$;*
2. *The best approximate solution of the equation is given by*

$$\boldsymbol{x}_{*} = \sqrt{\Delta}(\boldsymbol{b}^{-}\boldsymbol{A})^{-};$$

3. *If $\Delta = \mathbb{1}$, then $\boldsymbol{x}_{*} = (\boldsymbol{b}^{-}\boldsymbol{A})^{-}$ is an exact (the maximum) solution*

Two-Sided Equation

- Let $\boldsymbol{A}$ and $\boldsymbol{B}$ be given matrices of order $(m \times n)$ and $(m \times l)$
- Consider the problem to find regular vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ of order n and l, which are solutions of the two-sided equation

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{B}\boldsymbol{y}$$

- To obtain a best approximate solution of the two-sided equation, we implement an alternating algorithm (K. 2023)
- The algorithm applies the result of Theorem to solve a series of one-sided equations derived from the two-sided equation
- The equations are obtained from the two-sided equation where the left and right sides are alternately replaced by constant vectors

Alternating Algorithm

- Given a vector x_0, the algorithm examines the equations

$$Ax_0 = By_1, \quad Ax_2 = By_1, \quad Ax_2 = By_3, \quad Ax_4 = By_3, \quad \ldots$$

as one-sided equations in the unknowns y_1, x_2, y_3, x_4, $\ldots$

- The alternating algorithm successively calculate

$$\begin{aligned} y_1 &= \sqrt{\Delta_0}((Ax_0)^- B)^-, & \Delta_0 &= (B((Ax_0)^- B)^-)^- Ax_0, \\ x_2 &= \sqrt{\Delta_1}((By_1)^- A)^-, & \Delta_1 &= (A((By_1)^- A)^-)^- By_1, \\ y_3 &= \sqrt{\Delta_2}((Ax_2)^- B)^-, & \Delta_2 &= (B((Ax_2)^- B)^-)^- Ax_2, \quad \ldots \end{aligned}$$

- One can verify that the sequence $\Delta_0, \Delta_1, \ldots$ has a limit $\Delta_* \geq \mathbb{1}$
- If $\Delta_* = \mathbb{1}$, then the two-sided equation has a solution (x_*, y_*)
- Otherwise the pair (x_*, y_*) specifies a best approximate solution

Tropical Approximation: Polynomial Approximation

- Suppose there are m samples (x_i, y_i) where x_i and $y_i = f(x_i)$ are input and output of an unknown function $f : \mathbb{X} \to \mathbb{X}$
- Consider the problem of best approximation of these sample data by polynomials that have n monomials and are given by

$$P(x) = \bigoplus_{j=1}^{n} \theta_j x^{p_j}$$

  with known exponents $p_j \in \mathbb{Q}$ and unknown coefficients $\theta_j \in \mathbb{X}$
- The problem consists in finding the unknown coefficients that make the following equations hold exactly or approximately:

$$P(x_i) = \theta_1 x_i^{p_1} \oplus \cdots \oplus \theta_n x_i^{p_n} = y_i \qquad i = 1, \dots, m$$

Vector Representation

- Consider the system of scalar equations

$$\theta_1 x_i^{p_1} \oplus \cdots \oplus \theta_n x_i^{p_n} = y_i \qquad i = 1, \ldots, m$$

- With the matrix-vector notation

$$\boldsymbol{X} = \begin{pmatrix} x_1^{p_1} & \ldots & x_1^{p_n} \\ \vdots & & \vdots \\ x_m^{p_1} & \ldots & x_m^{p_n} \end{pmatrix}, \qquad \boldsymbol{y} = \begin{pmatrix} y_1, \\ \vdots \\ y_m \end{pmatrix}, \qquad \boldsymbol{\theta} = \begin{pmatrix} \theta_1 \\ \vdots \\ \theta_n \end{pmatrix},$$

we transform the system into the one-sided vector equation

$$\boldsymbol{X}\boldsymbol{\theta} = \boldsymbol{y}$$

- Application of Theorem yields the squared approximation error Δ^* and corresponding vector $\boldsymbol{\theta}_*$ of coefficients, which are given by

$$\Delta_* = (\boldsymbol{X}(\boldsymbol{y}^- \boldsymbol{X})^-)^- \boldsymbol{y}, \qquad \boldsymbol{\theta}_* = \sqrt{\Delta_*}(\boldsymbol{y}^- \boldsymbol{X})^-$$

Rational Approximation

- Consider a rational function as an approximant, which is given by

$$R(x) = \frac{P(x)}{Q(x)}, \qquad P(x) = \bigoplus_{j=1}^{n} \theta_j x^{p_j}, \qquad Q(x) = \bigoplus_{k=1}^{l} \sigma_k x^{q_k}.$$

- Given samples $x_i, y_i \in \mathbb{X}$ for $i = 1, \ldots, m$, we find the coefficients in polynomials $P(x)$ and $Q(x)$ to best approximate the equations

$$R(x_i) = y_i, \qquad i = 1, \ldots, m.$$

- We rewrite these equations as

$$\theta_1 x_i^{p_1} \oplus \cdots \oplus \theta_n x_i^{p_n} = y_i(\sigma_1 x_i^{q_1} \oplus \cdots \oplus \sigma_l x_i^{q_l}), \qquad i = 1, \ldots, m$$

Vector Representation

- To represent the problem in vector form, we introduce the notation

$$\boldsymbol{X} = \begin{pmatrix} x_1^{p_1} & \dots & x_1^{p_n} \\ \vdots & & \vdots \\ x_m^{p_1} & \dots & x_m^{p_n} \end{pmatrix}, \qquad \boldsymbol{Y} = \begin{pmatrix} y_1 & & \mathbb{0} \\ & \ddots & \\ \mathbb{0} & & y_m \end{pmatrix},$$

$$\boldsymbol{Z} = \begin{pmatrix} x_1^{q_1} & \dots & x_1^{q_l} \\ \vdots & & \vdots \\ x_m^{q_1} & \dots & x_m^{q_l} \end{pmatrix}, \qquad \boldsymbol{\theta} = \begin{pmatrix} \theta_1 \\ \vdots \\ \theta_n \end{pmatrix}, \qquad \boldsymbol{\sigma} = \begin{pmatrix} \sigma_1 \\ \vdots \\ \sigma_l \end{pmatrix}$$

- Then, the system takes the form of the two-sided vector equation

$$\boldsymbol{X}\boldsymbol{\theta} = \boldsymbol{Y}\boldsymbol{Z}\boldsymbol{\sigma}$$

- We use Alternating Algorithm to find the squared approximation error Δ_* and obtain the vectors of coefficients $\boldsymbol{\theta}_*$ and $\boldsymbol{\sigma}_*$

Examples: Approximation in Max-Plus Algebra

- We consider examples in terms of $\mathbb{R}_{\max,+}$ (max-plus algebra)
- In this setting, both tropical polynomials and rational functions can be represented as conventional piecewise linear functions
- We assume the polynomials have a fixed number of monomials, while the exponents of these monomials are not given in advance
- We combine random search to fix exponents, with the best approximation to find coefficients of monomials in the polynomials
- To reduce the feasible set of exponents in random search, we consider only polynomials with integer exponents

Approximation of Convex Function

- We consider $m = 21$ input/output points (x_i, y_i) from the function
$$f(x) = x^2 - 3x^{1/3} + 5/2, \qquad 0 \le x \le 2,$$
where $x_i = (i-1)/10$ and $y_i = f(x_i)$ for $i = 1, \dots, 21$
- We approximate the data by polynomials with $n = 7$ monomials
- The exponents for the monomials are produced by random sampling from the discrete uniform distribution over $[-15, 5]$
- For each sample set of exponents, coefficients of monomials are found using Theorem to minimize the approximation error
- After examining 10,000 sample sets, the minimum squared error is $\Delta_* = 0.0481$ and the polynomial (in the conventional form) is

$$P_*(x) = \max(2.5240 - 15x, 1.4096 - 3x, 0.8736 - x, 0.3503, \\ -0.4760 + x, -1.6720 + 2x, -3.2853 + 3x)$$

Approximation of Convex Function

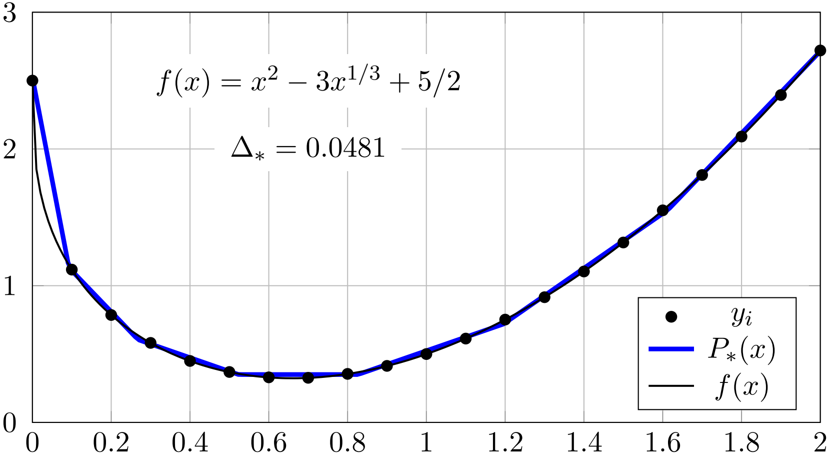

Figure: Approximation of $f(x)$ by a polynomial $P_*(x)$ with $n = 7$ terms

Approximation of Non-Convex Function

- Suppose that data points (x_i, y_i) are obtained from the function

$$g(x) = 3(x-1)^2 \sin(x) + 1/4, \qquad 0 \leq x \leq 2$$

- We approximate $g(x)$ by rational functions given by

$$R(x) = P(x)/Q(x),$$

  where $P(x)$ and $Q(x)$ are polynomials of $n = 6$ and $l = 4$ terms
- Random sampling of 10,000 pairs of sets of exponent in $[-10, 10]$ and using Alternating Algorithm yield $\Delta_* = 0.0701$
- The approximating function is $R_*(x) = P_*(x) - Q_*(x)$, where

$$P_*(x) = \max(6.9455 - 3x, 6.0860 - 2x, 4.9978 - x, 3.7461, \\ 0.7639 + 2x, -2.6361 + 4x),$$

$$Q_*(x) = \max(6.6880 - 5x, 6.2962 - 3x, 5.8009 - 2x, 2.4211)$$

Approximation of Non-Convex Function

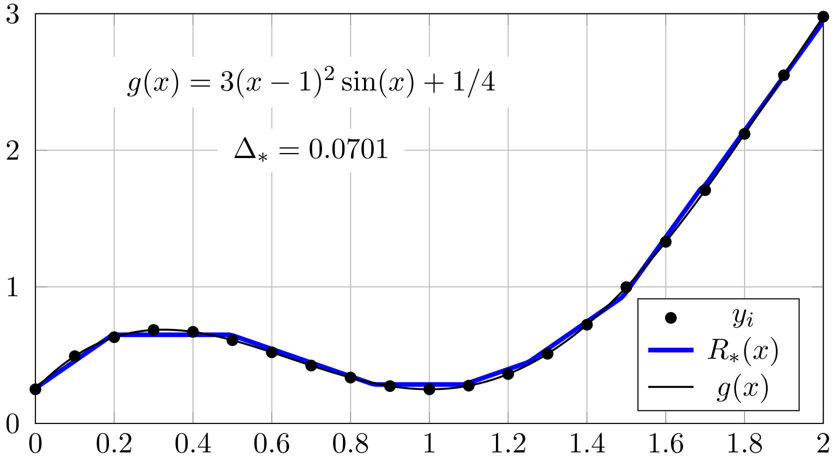

Figure: Approximation of $g(x)$ by a rational function $R_*(x)$ with $n = 6$, $l = 4$

Conclusion

- We considered a best approximation problem in which a function is approximated from a set of input/output samples
- We formulated the problem in terms of a tropical semifield where addition is idempotent and multiplication is invertible
- We transformed the problem into solving tropical linear vector equations with an unknown vector on one side or on both sides
- We have derived an exact best approximate solution of the one-sided equation, obtained in direct analytical form
- To obtain a best approximate solution of the two-sided equation, we have used an iterative alternating algorithm
- As illustration, we presented results of best discrete Chebyshev approximation of real functions by piecewise linear functions