Использование тропической математики в задачах аппроксимации функций

Айнабеков Захар, гр. 422
Выпускная квалификационная работа

Санкт-Петербургский государственный университет
Математико-механический факультет
Кафедра статистического моделирования

Научный руководитель: д. ф.-м. н., профессор Н. К. Кривулин
Рецензент: к. ф.-м. н., доцент А. Ю. Пономарева

2023 г.

Постановка задачи

Пусть имеется функция $f : [a, b] \subset \mathbb{R} \to \mathbb{R}$. Пусть имеется некоторый набор пар точек (x_k, y_k) таких, что $a \le x_1 < \dots < x_K \le b$ и $y_k = f(x_k)$, где $k = 1, \dots, K$.
Рассмотрим семейство функций

$$F_\Theta = \{f_\theta : [a, b] \subset \mathbb{R} \to \mathbb{R} | \theta \in \Theta\},$$

где Θ — множество параметров.
Требуется выбрать функцию f_{θ_*} так, чтобы выбранное расстояние d между вектором значений функций $\boldsymbol{y} = (y_k)_{k=1,\dots,K}$ и оценок $f_{\theta_*}(\boldsymbol{x}) = (f_{\theta_*}(x_k))_{k=1,\dots,K}$ было минимальным, то есть

$$\theta_* = \underset{\theta}{\operatorname{argmin}}\ d(\boldsymbol{y}, f_{\theta_*}(\boldsymbol{x})).$$

Постановка задачи

Рассмотрим различные критерии близости:

1. Расстояние Чебышева

$$\forall \boldsymbol{a} = (a_1, \dots, a_n), \boldsymbol{b} = (b_1, \dots, b_n) \in \mathbb{R}^n \quad d(\boldsymbol{a}, \boldsymbol{b}) = \max_{1 \le i \le n} |a_i - b_i|.$$

   Рассмотрим мотивирующие примеры:

   1. задача регрессии с равномерно распределенными шумами,
   2. задача наилучшего равномерного приближения функции.

   Алгоритмическое решение задачи аппроксимации функции в метрике Чебышева можно получить сведением задачи к задаче линейного программирования.

2. Лог-чебышевское расстояние с параметром $p > 1$

$$\forall \boldsymbol{a} = (a_1, \dots, a_n), \boldsymbol{b} = (b_1, \dots, b_n) \in \mathbb{R}_+{}^n \quad d(\boldsymbol{a}, \boldsymbol{b}) = \max_{1 \le i \le n} |\log_p a_i - \log_p b_i|.$$

Идемпотентное полуполе (Kolokoltsov и Maslov 1997)

Рассмотрим множество $\mathbb{X}$ с двумя операциями $\oplus$ и $\otimes$.

1. $(\mathbb{X}, \oplus)$ — идемпотентная коммутативная полугруппа с нейтральным элементом $\mathbb{0}$,
2. $(\mathbb{X}, \otimes)$ — абелева группа с нейтральным элементом $\mathbb{1}$,
3. операции $\oplus$ и $\otimes$ связаны дистрибутивностью и законом поглощения.

Структура $(\mathbb{X}, \mathbb{0}, \mathbb{1}, \oplus, \otimes)$ называется *идемпотентным полуполем*.
Примерами вещественных идемпотентных полуполей являются следующие структуры:

$$\mathbb{R}_{\max,+} = (\mathbb{R} \cup \{-\infty\}, -\infty, 0, \max, +), \quad \mathbb{R}_{\min,+} = (\mathbb{R} \cup \{+\infty\}, +\infty, 0, \min, +),$$
$$\mathbb{R}_{\max,\times} = (\mathbb{R}_{+} \cup \{0\}, 0, 1, \max, \times), \quad \mathbb{R}_{\min,\times} = (\mathbb{R}_{+} \cup \{+\infty\}, +\infty, 1, \min, \times).$$

Для указанных полуполей можно ввести операцию возведения в степень в смысле заданной в полуполе операции умножения.

Рассмотрим матрицы с элементами из $\mathbb{X}$.

$$\forall \boldsymbol{A} = (a_{ij}) \in \mathbb{X}^{m\times n},\ \boldsymbol{B} = (b_{ij}) \in \mathbb{X}^{m\times n},\ \boldsymbol{C} = (c_{ij}) \in \mathbb{X}^{n\times l}$$

Операции сложения и умножения матриц, а также операция умножения на скаляр $x \in \mathbb{X}$ определены обычным образом

$$\{\boldsymbol{A} \oplus \boldsymbol{B}\}_{ij} = a_{ij} \oplus b_{ij},\ \{\boldsymbol{BC}\}_{ij} = \bigoplus_{k=1}^{n} b_{ik} c_{kj},\ \{x\boldsymbol{A}\}_{ij} = x a_{ij}.$$

Множество вектор-столбцов размера n обозначается $\mathbb{X}^n$.
Для любой ненулевой матрицы $\boldsymbol{A} = (a_{ij}) \in \mathbb{X}^{m\times n}$ можно определить *мультипликативно сопряженную матрицу* $\boldsymbol{A}^{-} = (a_{ij}^{-}) \in \mathbb{X}^{n\times m}$ с элементами

$$a_{ij}^{-} = \begin{cases} a_{ji}^{-1}, & \text{если } a_{ji} \neq \mathbb{0}, \\ \mathbb{0}, & \text{если } a_{ji} = \mathbb{0}. \end{cases}$$

Расстояние от вектора до множества

Для любых векторов $\boldsymbol{a} = (a_i) \in \mathbb{X}^n$ и $\boldsymbol{b} = (b_i) \in \mathbb{X}^n$ без нулевых элементов можно определить расстояние между ними (метрику)

$$d(\boldsymbol{a}, \boldsymbol{b}) = \bigoplus_{1 \leq i \leq n} (b_i^{-1} a_i \oplus a_i^{-1} b_1) = \boldsymbol{b}^{-} \boldsymbol{a} \oplus \boldsymbol{a}^{-} \boldsymbol{b}.$$

Для полуполя $\mathbb{R}_{\max,+}$ расстояние, введенное таким образом, примет вид

$$d(\boldsymbol{a}, \boldsymbol{b}) = \max_{1 \leq i \leq n} |a_i - b_i|.$$

Для полуполя $\mathbb{R}_{\max,\times}$ расстояние примет вид

$$d(\boldsymbol{a}, \boldsymbol{b}) = \max_{1 \leq i \leq n} \max(a_i b_i^{-1}, a_i^{-1} b_i).$$

Рассмотрим $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_m \in \mathbb{X}^n$, пусть $\mathcal{A} = \operatorname{span}\{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_m\}$ и $\boldsymbol{A} = (\boldsymbol{a}_1, \ldots, \boldsymbol{a}_m)$. Определим расстояние от некоторого вектора $\boldsymbol{b} \in \mathbb{X}^n$ до $\mathcal{A}$.

$$d(\mathcal{A}, \boldsymbol{b}) = \min_{\boldsymbol{x} \in \mathbb{X}^m} d(\boldsymbol{A}\boldsymbol{x}, \boldsymbol{b}).$$

Расстояние от вектора до множества

Матрица называется *регулярной*, если она не имеет нулевых строк и столбцов. Вектор называется *регулярным*, если он не имеет нулевых координат.

Пусть матрица $\boldsymbol{A}$ и вектор $\boldsymbol{b}$ являются регулярными. Введем обозначение

$$\Delta = (\boldsymbol{A}(\boldsymbol{b}^{-}\boldsymbol{A})^{-})^{-}\boldsymbol{b}.$$

Рассмотрим следующий результат из работы (Кривулин 2009).

Лемма

Для любой регулярной матрицы $\boldsymbol{A}$ и регулярного вектора $\boldsymbol{b}$ выполняется равенство

$$d(\mathcal{A}, \boldsymbol{b}) = \sqrt{\Delta}.$$

При этом минимум величины $d(\boldsymbol{A}\boldsymbol{x}, \boldsymbol{b})$ достигается на векторе

$$\boldsymbol{x} = \sqrt{\Delta}(\boldsymbol{b}^{-}\boldsymbol{A})^{-}.$$

Аппроксимация тропическими полиномами

Пусть $d_1, \ldots, d_m$ – заданные рациональные числа, $\theta_1, \ldots, \theta_m$ – неизвестные параметры. Рассмотрим в качестве аппроксимирующей функции тропический полином (Krivulin 2020; Krivulin 2021)

$$P_\theta(x) = \bigoplus_{i=1}^{m} \theta_i x^{d_i}.$$

Предположим, что задан набор пар (x_k, y_k) для всех $k = 1, \ldots, K$. Будем подбирать значения параметров $\theta_1, \ldots, \theta_m$ так, чтобы обеспечить минимальное расхождение между обеими частями уравнений

$$y_k = P_\theta(x_k), \qquad k = 1, \ldots, K.$$

Аппроксимация тропическими полиномами

С использованием матрично-векторных обозначений

$$\boldsymbol{X} = \begin{pmatrix} x_1^{d_1} & x_1^{d_2} & \dots & x_1^{d_m} \\ \vdots & \vdots & & \vdots \\ x_K^{d_1} & x_K^{d_2} & \dots & x_K^{d_m} \end{pmatrix}, \quad \boldsymbol{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_K \end{pmatrix}, \quad \boldsymbol{\theta} = \begin{pmatrix} \theta_1 \\ \vdots \\ \theta_K \end{pmatrix}$$

скалярные уравнения можно записать в векторной форме

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

Из леммы следует, что минимум расстояния между $\boldsymbol{y}$ и $\boldsymbol{X}\boldsymbol{\theta}$ обеспечивается вектором параметров, который имеет вид:

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

где величина

$$\Delta_* = (\boldsymbol{X}(\boldsymbol{y}^-\boldsymbol{X})^-)^-\boldsymbol{y}$$

является (тропическим) квадратом ошибки (погрешности) аппроксимации.

Пример $y = x^x$

Для функции $y = x^x$ найдем аппроксимацию тропическим полиномом $P_\theta(x)$ в полуполе $\mathbb{R}_{\max,+}$ по $K = 21$ точкам $0, 0.1, \ldots, 1.9, 2$.

- Для $m = 3$ и степеней $-3, 2, 7$ значение квадрата ошибки аппроксимации $\Delta \approx 0.9054139$,
- Для $m = 5$ и степеней $-3, -0.5, 2, 4.5, 7$ значение квадрата ошибки аппроксимации $\Delta \approx 0.411732$,
- Для $m = 9$ и степеней $-3, -1.75, -0.5, 0.75, 2, 3.25, 4.5, 5.75, 7$ значение квадрата ошибки аппроксимации $\Delta \approx 0.09757421$.

Пример $y = x^x$

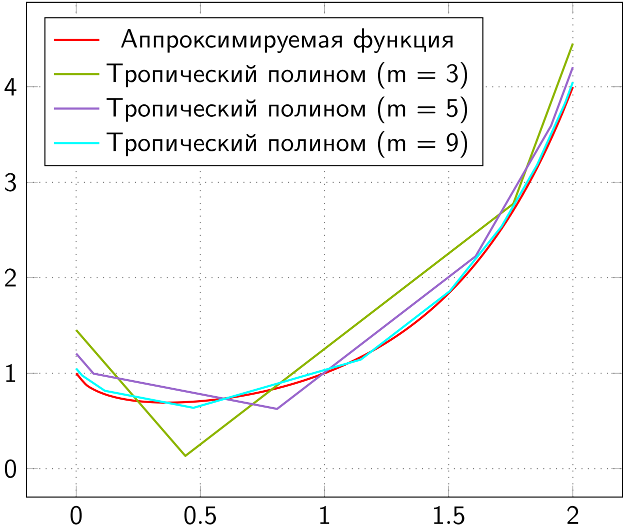

Рис. 1: Приближение функции $y = x^x$ тропическими полиномами.

Пример $y = \frac{\sin(x)}{x} + 1$

Для функции $y = \frac{\sin(x)}{x} + 1$ найдем аппроксимацию тропическим полиномом $P_\theta(x)$, заданным в терминах идемпотентного полуполя $\mathbb{R}_{\max,\times}$, по $K = 21$ точкам $2, 2.3, \ldots, 7.7, 8$.

- Для $m = 3$ и степеней $-1, 0, 1$ значение квадрата ошибки аппроксимации $\Delta \approx 1.132922$,
- Для $m = 5$ и степеней $-1, -0.5, 0, 0.5, 1$ значение квадрата ошибки аппроксимации $\Delta \approx 1.123068$,
- Для $m = 9$ и степеней $-1, -0.75, -0.5, -0.25, 0, 0.25, 0.5, 0.75, 1$ значение квадрата ошибки аппроксимации $\Delta \approx 1.078352$.

Пример $y = \frac{\sin(x)}{x} + 1$

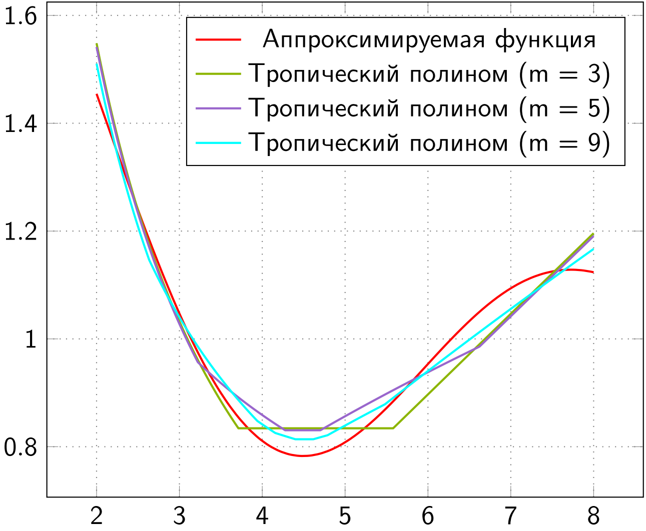

Рис. 2: Приближение функции $y = \frac{\sin(x)}{x} + 1$ тропическими полиномами.

Анализ матрицы задачи аппроксимации

Рассмотрим следующий результат из работы (Кривулин 2009).

Лемма

Вектор $\boldsymbol{b}$ принадлежит линейной оболочке столбцов матрицы $\boldsymbol{A}$, то есть является линейно зависимым от них тогда и только тогда, когда $\Delta(\boldsymbol{A}, \boldsymbol{b}) = \mathbb{1}$. При этом $\boldsymbol{b} = \boldsymbol{A}\boldsymbol{x}$, где $\boldsymbol{x} = (\boldsymbol{b}^{-}\boldsymbol{A})^{-}$.

Утверждение

Пусть $d_1 < d_2 < \cdots < d_{m-1} < d_m$ и $x_1 < x_2 < \cdots < x_{K-1} < x_K$, тогда в идемпотентном полуполе с зафиксированной операцией сложение $\oplus = \max$ столбцы матрицы

$$\boldsymbol{X} = \begin{pmatrix} x_1^{d_1} & x_1^{d_2} & \dots & x_1^{d_m} \\ \vdots & \vdots & & \vdots \\ x_K^{d_1} & x_K^{d_2} & \dots & x_K^{d_m} \end{pmatrix}.$$

являются линейно независимыми.

Алгоритм аппроксимации со случайным выбором степеней

При фиксированном наборе точек (x_k, y_k), где $k = 1, \ldots, K$, и фиксированном диапозоне набора степеней m рассмотрим величину Δ в качестве функции от вектора параметров $\boldsymbol{d} = \{d_1, \ldots, d_m\}$. Введем следующее обозначение

$$\boldsymbol{X}(\boldsymbol{d}) = \begin{pmatrix} x_1^{d_1} & x_1^{d_2} & \ldots & x_1^{d_m} \\ \vdots & \vdots & & \vdots \\ x_K^{d_1} & x_K^{d_2} & \ldots & x_K^{d_m} \end{pmatrix}.$$

Рассмотрим величину ошибки аппроксимации как функцию от вектора степеней

$$\Delta(\boldsymbol{d}) = (\boldsymbol{X}(\boldsymbol{d})(\boldsymbol{y}^{-}\boldsymbol{X}(\boldsymbol{d}))^{-})^{-}\boldsymbol{y}$$

и оптимизируем её случайным поиском.

Моделирование

Для оптимизируемой функции, вычисляемой в полуполе $\mathbb{R}_{\max,+}$, где значения y_k, вычисленных как $y = x^x$ по точкам $x_k = 0, 0.1, \ldots, 1.9, 2$, получены следующие результаты. Границами для параметров были выбраны -3 и 7. Минимальные значения получены

- для $m = 3$ на степенях $-0.4261115, 1.319608, 4.6146209$,
- для $m = 5$ на степенях $-1.4149564, 0.2025834, 1.307882, 2.777649, 5.4154603$,
- для $m = 9$ на степенях $-2.5115925, -0.5099779, 0.0452132, 0.6148764, 1.2978206, 1.7312447, 2.7698633, 4.3415426, 6.2598756$.

m	MinΔ	MeanΔ	MaxΔ	Значение Δ из примера
3	0.1987399	0.2463088	0.308695	0.9054139
5	0.0728836	0.1071197	0.1411797	0.411732
9	0.0224856	0.0431884	0.0602196	0.0975742

Таблица 1: Описательная статистика значений оптимизируемой функции.

Заключение

Основные результаты работы состоят в следующем:

1. рассмотрена задача аппроксимации функции,
2. изучены основные понятия и результаты тропической математики,
3. задача аппроксимации функции сформулирована в терминах тропической математики как задача минимизации расстояния от вектора до линейной оболочки столбцов матрицы решения,
4. предложено решение задачи с использованием в качестве аппроксимирующих функций тропических полиномов с фиксированным набором степеней,
5. для матрицы решения доказано утверждение о линейной независимости столбцов,
6. реализована модификация алгоритма аппроксимации со случайным выбором степеней.