СИСТЕМНЫЙ АНАЛИЗ, УПРАВЛЕНИЕ И ОБРАБОТКА ИНФОРМАЦИИ

УДК 519.677

РЕШЕНИЕ ОБРАТНЫХ ОПТИМИЗАЦИОННЫХ ЗАДАЧ ДЛЯ НЕЙРОСЕТЕВЫХ ИНТЕЛЛЕКТУАЛЬНЫХ МОДЕЛЕЙ НА ОСНОВЕ ЭПСИЛОН-ЛИПШИЦЕВЫХ МЕТОДОВ

Новикова С.В., Чернышевский П.А.
КНИТУ КАИ имени А.Н. Туполева, г. Казань

Поступила в редакцию 30.03.2022, после переработки 05.05.2022.

Использование методов интеллектуального анализа не всегда позволяет ответить на все вопросы, которые могут быть сформулированы в рамках рассматриваемой математической модели. В данной работе показано, как некоторые из таких запросов могут быть представлены в виде задачи глобальной оптимизации непрерывной нейросетевой функции. Нахождение глобального минимума функции, заданной нейросетевой моделью, в некоторых случаях затрудняется сложностью доказательства ее липшицевости и вычисления константы Липшица, поскольку наличие непрерывности не гарантирует в общем случае выполнение неравенства Липшица. В свою очередь, это затрудняет применение классических подходов. В данной работе предложено использовать для приближенного нахождения минимума модифицированные методы на основе использования понятия ε-липшицевости, так как для их работы требуется лишь свойство непрерывности. В качестве примера рассмотрена нейросетевая модель расчета концентрации металлов в биосредах населения в зависимости от их содержания в питьевой воде, составлена соответствующая оптимизационная задача и приведены результаты её численного решения с помощью обобщенного метода Стронгина.

Ключевые слова: нейросетевое моделирование, интеллектуальный анализ, непрерывная функция, глобальная оптимизация, обобщенный метод Стронгина.

Вестник ТвГУ. Серия: Прикладная математика. 2022. № 2. С. 74–83.
https://doi.org/10.26456/vtpmk639

Введение

Развитие различных методов интеллектуального анализа данных способствует увеличению эффективности экологического мониторинга и позволяет строить достаточно точные модели прогноза с целью отслеживания всевозможных рисков [1].

В частности, существует ряд задач, где рассматривается зависимость количества вредных веществ в организме от содержания этих веществ в различных средах, например, в питьевой воде, и некоторых физиологических характеристиках самого человека [2, 3].

С формальной точки зрения задача сводится к построению модели, описывающей прямую зависимость выходного вектора $Y = (y_1, y_2, ..., y_m)$ от входного вектора $X = (x_1, x_2, ..., x_n)$. Имеется множество различных и хорошо зарекомендовавших себя способов реализовать такую зависимость на практике, например, построить нейронную сеть, применить методы кластеризации, получить регрессионную зависимость и т.д. В результате удается в том или ином виде получить функцию $Y = f(X)$, которая и понимается как искомая прямая зависимость.

Однако, наличие такой функции не всегда позволяет решить все вопросы, которые можно задать в рамках рассматриваемой модели. Как будет показано далее на примере, постановка вопроса может включать поиск такого значения входа X, при котором достигается некоторое фиксированное выходное значение Y^*. В такой постановке задачу можно отнести к классу так называемых «обратных задач», решение которых в общем случае нетривиально [4].

В данной работе предлагается метод решения обратной задачи путем преобразования ее к задаче глобальной оптимизации. Подход, сводящий обратную задачу к оптимизационной, применяется на практике давно. Однако в данном случае ставится задача поиска именно глобального минимума, что в условиях слабой устойчивости прямой модели является весьма сложной задачей.

1. Сведение исходной задачи к задаче глобальной оптимизации

Введем следующие обозначения:

$X = (x_1, x_2, ..., x_n) \in D_X \subseteq \mathrm{R}^n$ – входной вектор;

$Y = (y_1, y_2, ..., y_m) \in D_Y \subseteq \mathrm{R}^m$– выходной вектор, причем считаем, что в R^m задана некоторая метрика ρ;

$f : D_X \to D_Y$ – известная функция, являющаяся результатом моделирования и представляющая собой прямую зависимость выхода Y от входа X.

$Y^* \in D_Y$ – некоторый фиксированный вектор.

Построим функцию $F : \mathrm{R}^n \to \mathrm{R}$:

$$F(X) = \rho(f(X), Y^*).$$

Тогда задача может быть сформулирована следующим образом:

$$\arg\min F(X)\, s.t.\, X \in D_X. \tag{1}$$

Задача (1) представляет собой задачу глобальной оптимизации целевой функции F, которая в зависимости от вида этой функции и вида ограничений может быть решена соответствующими методами.

Способов обеспечения поиска именно глобального минимума разработано немного. Так, наиболее применимыми являются методы, опирающиеся на свойство липшицевости целевой функции и априорном знании константы Липшица, используемой во всех таких алгоритмах в качестве параметра. В частности, если функция F липшицева на множестве D_X, а само множество представимо в виде гиперинтервала $D_X = \{X : a_i \le x_i \le b_i\}$, то для решения (1) можно применить диагональный подход из [5], а если область поиска представляет собой отрезок, то применимы алгоритмы Пиявского, Евтушенко и Стронгина [6-8]. В случае, если функция непрерывна, можно использовать обобщение некоторых «липшицевых» алгоритмов, базирующееся на введенном в [9] свойстве ε-липшицевости. В той же работе доказана теорема о том, что всякая непрерывная на выпуклом компактном множестве функция обладает этим свойством и наоборот. Для одномерного случая в [9-12] предложены обобщения всех вышеупомянутых алгоритмов для липшицевых функций, а для случая функции многих переменных можно реализовать алгоритм из [13]. Необходимо отметить, что все перечисленные алгоритмы позволяют гарантированно решить задачу (1).

Далее мы опишем математическую модель из [3] и сформулируем для неё вид задачи (1).

2. Модель прогнозирования уровня металлов в биосредах человека

В [3] предложена система, позволяющей прогнозировать уровень металлов в биосредах организма человека по их содержанию в питьевой воде. Исходный набор данных представляет собой 404 кортежа обезличенных данных, включающих такие параметры как вес, рост, возраст, пол, площадь поверхности тела пациента, концентрация вредных металлов (цинк, хром, железо, стронций, медь, свинец) в крови, моче, волосах и в питьевой воде.

В качестве параметров входного вектора X рассматривается уровень заданного металла в питьевой воде (мкг/мл, x_1) и площадь поверхности тела, далее ППТ (м^2,x_2), вычисленная по формуле Мостеллера на основе значений веса и роста. Параметрами выходного вектора Y являются уровень металла в крови, в волосах и в моче. Однако в силу неполноты информации в кортежах данных и особенностей модели далее будет использован только один параметр, а именно, уровень металла в крови (мкг/мл, y) в предположении, что металл в организме распространяется током крови.

Целью исследования ставится определение допустимый уровень x_1^* заданного металла в питьевой воде, чтобы он не превышал предельно допустимый y^* и был безопасен для детей в возрасте от 1 года до 12 лет. Для решения этой задачи в данной работе предлагаются два этапа:

1. Построение и обучение на основе исходных данных нейронной сети (персептрона) для установления зависимости f_{ANN} вектора Y от X [14]. Архитектура нейронной сети представляет собой персептрон с двумя входными нейронами, скрытым слоем из трех нейронов и одним выходным слоем. В качестве функции активации взята сигмоида.

2. Сведение исходной задачи к задаче глобальной оптимизации вида (1). Для этого сделаны следующие дополнительные предположения. Во-первых, для уровня металла в питьевой воде установлены максимальные b и минимальные a значения для каждого из металлов исходя из исходных данных, т.е. $a \le x_1 \le b$. Во-вторых, определены 4 возрастные группы, дети 2, 9, 10 и 12-13 лет относительно которых взято известное значение c индекса ППТ.

Следовательно, значение x_2 далее будет фиксировано для каждой группы. Таким образом, задача (1) имеет вид:

$$\arg\min_{X \in D} F(X), \tag{2}$$

$$F(X) = \rho(f_{ANN}(X), y^*),$$

$$D = \left\{X \in \mathrm{R}^2 : a \le x_1 \le b, x_2 = c\right\}, \ \ \rho(x, y) = \|x - y\|_2 .$$

Применение классических методов глобальной оптимизации упирается в необходимость доказательства липшицевости функции $F(X)$, а значит и функции f_{ANN}, и необходимости определения значения оценки константы Липшица для нее. Так как функция задана в виде нейронной сети, оба условия выполнить довольно затруднительно.

Однако можно показать, что функция f_{ANN} непрерывна. Действительно, так как функция активации непрерывна, а сама функция f_{ANN} представляет собой линейную комбинацию активационной функции и фиксированных значений весов, то целевая функция из (2) также непрерывна на области D. Кроме того, изменяется только переменная x_1, следовательно, задачу (2) можно рассматривать как задачу одномерной глобальной оптимизации непрерывной функции на отрезке $[a; b]$.

Для подобных задач, как это отмечено в пункте 1, разработаны специальные модифицированные алгоритмы, позволяющие гарантировать нахождение глобального минимума одномерной функции, опираясь на понятие ε-липшицевости. Для работы этих метод так же, как и в случае липшицевых функций, необходимо знание оценки $L(\varepsilon)$ минимальной ε-постоянной Липшица при фиксированном $\varepsilon > 0$, которая существует в силу теоремы из [9]. В работе [12] предложено обобщение алгоритма Стронгина, который строит эту оценку в ходе работы метода, что позволяет требовать от целевой функции только её непрерывности на заданном отрезке. Шаги алгоритма и доказательство нахождения таким способом приближенного решения именно глобального минимума оптимизируемой функции подробно описаны в той же работе.

В следующем разделе приведены результаты численного решения (2) упомянутым методом для двух металлов – стронция и хрома.

3. Результаты вычислений

Задача А. Определение допустимого уровня стронция в питьевой воде для фиксированных значений физиологических параметров. Известно, что предельное значение по нормативу составляет 0.12 мкг/мл стронция в крови. Минимальные и максимальные значения стронция в питьевой воде по результатам измерений $a = 0.297$ и $b = 0.693$ мкг/мл. Результаты вычислений для различных возрастных групп и соответствующих каждой группе значений ППТ представлены в Таблице 1.

Для взрослых максимальное допустимое содержания стронция согласно СанПиН 2.1.4.559-96 составляет 7 мкг/мл, что заметно больше полученных результатов.

Таблица 1: Результаты вычислений для Задачи А (стронций)

Возрастная группа	Значение ППТ (м^2)	Значение стронция в воде, соответствующее предельному (мкг/мл)
2 года	0.5	0.3920318260091045
9 лет	1.07	0.4377102576832432
10 лет	1.14	0.43679181833828207
12-13 лет	1.33	0.43383055799245107

Задача Б. Определение допустимого уровня хрома в питьевой воде для фиксированных значений физиологических параметров. Предельное значение по нормативу составляет 0.075 мкг/мл хрома в крови. Минимальные и максимальные значения хрома в питьевой воде по результатам измерений $a = 0.0005$ и $b = 0.022$ мкг/мл. Результаты вычислений представлены в Таблице 2.

Таблица 2: Результаты вычислений для Задачи Б (хром)

Возрастная группа	Значение ППТ (м^2)	Значение стронция в воде, соответствующее предельному (мкг/мл)
2 года	0.5	0.0050507274076464015
9 лет	1.07	0.008302587283067214 (будет в крови 0.059)
10 лет	1.14	0.010263598770334932 (будет в крови 0.070)
12-13 лет	1.33	0.018721770865570222 (будет в крови 0.073)

Для трех возрастных групп получено значении концентрации хрома в воде, при котором достигается наибольшее значение хрома в крови, не превышающие ПДК. Предельное же содержание хрома в воде по СанПиН 2.1.4.559-96 составляет от 0.05 до 0.5 мкг/мл, что достаточно хорошо согласуется с полученными результатами.

Заключение

В данной работе показана применимость алгоритмов глобальной оптимизации на основе ε-липшицевости, в частности, обобщенного алгоритма Стронгина, для функций, описываемых интеллектуальными нейросетевыми структурами. Приведен пример конкретной математической модели на основе многослойного нейросетевого персептрона с непрерывной активационной функцией, для которой составлена и приближенно решена такая задача обобщенным методом Стронгина. Результаты вычислений согласуются с экспериментальными данными.

Список литературы

[1] Дмитриев В.Г. Оценка экологического риска. Аналитический обзор публикаций // Арктика и Север. 2014. № 14. С. 126–147.

[2] Тунакова Ю.А., Новикова С.В., Шагидуллин А.Р., Валиев В.С. Подходы для обеспечения техносферной безопасности городской среды с помощью методов нейросетевого моделирования // XXI век. Техносферная безопасность. 2020. Т. 5, № 1(17). С. 21–28.

[3] Кремлева Э.Ш., Новикова С.В., Шагидуллин А.Р. Интегральная оценка состояния окружающей среды на основе автоматического нейросетевого распознавателя // Сборник трудов международной научной конференции (школа молодых ученых) "Химия и инженерная экология - XIX". 2019. С. 224–227.

[4] Tarantola A. Inverse Problem Theory Methods for Data Fitting and Model Parameter Estimation. Amsterdam: Elseveir, 1987. 644 p.

[5] Сергеев Я.Д., Квасов Д.Е. Диагональные методы глобальной оптимизации. М.: Физматлит, 2008. 352 с.

[6] Пиявский С.А. Один алгоритм отыскания абсолютного экстремума функций // Журнал вычислительной математики и математической физики. 1972. Т. 12, № 4. С. 885–896.

[7] Евтушенко Ю.Г. Численный метод поиска глобального экстремума функции (перебор на неравномерной сетке) // Журнал вычислительной математики и математической физики. 1971. Т. 11, № 6. С. 1390–1403.

[8] Стронгин Р.Г. Численные методы в многоэкстремальных задачах. М.: Наука, 1978. 240 с.

[9] Vanderbei R.J. Extension of Piyavskii's Algorithm to Continuous Global Optimization // Journal of Global Optimization. 1999. Vol. 14. Pp. 205–216.

[10] Заботин В.И., Чернышевский П.А. Две модификации обобщенного метода Пиявского поиска глобального минимума непрерывной на отрезке функции и их сходимость // Вестник ТвГУ. Серия: Прикладная математика. 2021. № 3. С. 70–85. https://doi.org/10.26456/vtpmk624

[11] Арутюнова Н.К. Метод Евтушенко поиска глобального минимума ε-липшицевой функции и его приложения // Вестник КГТУ им. А.Н. Туполева. 2013. № 2. С. 154–157.

[12] Zabotin V.I., Chernyshevskij P.A. Extension of Strongin's Global Optimization Algorithm to a Function Continuous on a Compact Interval // Computer Research and Modeling. 2019. Vol. 11, № 6. Pp. 1111–1119.

[13] Arutyunova N.K., Dulliev A.M., Zabotin V.I. Global optimization of multivariable functions satisfying the Vanderbei condition // Journal of Applied Mathematics and Computing. 2022. Vol. 68, № 2. Pp. 1135–1161. https://doi.org/10.1007/s12190-021-01563-4

[14] Ростовцев В.С. Искусственные нейронные сети. Лань, 2019. 216 с.

Образец цитирования

Новикова С.В., Чернышевский П.А. Решение обратных оптимизационных задач для нейросетевых интеллектуальных моделей на основе эпсилон-липшицевых методов // Вестник ТвГУ. Серия: Прикладная математика. 2022. № 2. С. 74–83. https://doi.org/10.26456/vtpmk639

Сведения об авторах

1. **Новикова Светлана Владимировна**
профессор кафедры Прикладной математики и информатики Казанского национального исследовательского технического университета им А.Н. Туполева – КАИ.

Россия, 420015, г. Казань, ул. Большая Красная, д. 55, КНИТУ-КАИ.
E-mail: sweta72@bk.ru

2. **Чернышевский Павел Андреевич**
ассистент кафедры прикладной математики и информатики им. Ю.В. Кожевникова Казанского национального исследовательского технического университета им. А.Н. Туполева КАИ.

Россия, 420015, Республика Татарстан, г. Казань, улица Большая Красная, д. 55, 7-е учебное здание КНИТУ-КАИ. E-mail: pavelcomm@mail.ru

INVERSE OPTIMIZATION PROBLEM SOLVING FOR ANN DATA MINING MODELS BASED ON THE EPSILON-LIPSCHITZ APPROACH

Novikova Svetlana Vladimirovna
Professor at the Department of Applied Mathematics and Computer Science, Kazan National Research Technical University named after A.N. Tupolev – KAI
Russia, 420015, Kazan, 55 Bolshaya Krasnaya str., KNRTU-KAI.
E-mail: sweta72@bk.ru

Chernyshevskij Pavel Andreevich
Teaching assistant at Applied Mathematics and Informatics department, Kazan National Research Technical University named after A. N. Tupolev - KAI
Russia, 420015, Tatarstan, Kazan, 55 Bolshaya Krasnaya Str., Academic Building No. 7.
E-mail: pavelcomm@mail.ru

Received 30.03.2022, revised 05.05.2022.

Data mining techniques in particular cases cannot give us answers to all questions appeared in terms of the concerned simulation model. In this paper we show how some of such questions can be formulated as global optimization problem with continuous ANN function. Difficulties with proving an ANN based function Lipschitz continuity and Lipschitz constant estimating in some cases makes searching for the global minimum problematic since continuity does not guarantee us Lipschitz inequality holding. As a result, we are not able to apply conventional techniques. In this paper we propose the use of modified methods based on the ε- Lipschitz property for finding the global minimum because it requires only objective function continuity. As the example we analyze an ANN based prediction model for calculating metal level in human depending on metal level in drinking water, obtain associated optimization problem and show numerical results based on extended Strongin algorithm.

Keywords: ANN modeling, data mining, continuous function, global optimization, extended Strongin algorithm.

Citation

Novikova S.V., Chernyshevskij P.A., “Inverse optimization problem solving for ANN data mining models based on the epsilon-Lipschitz approach”, *Vestnik TvGU. Seriya: Prikladnaya Matematika [Herald of Tver State University. Series: Applied Mathematics]*, 2022, № 2, 74–83 (in Russian). https://doi.org/10.26456/vtpmk639

References

[1] Dmitriev V.G., "Environmental risk assessment. Analytical review of publications", *Arktika i Sever [The Arctic and the North]*, 2014, № 14, 126–147 (in Russian).

[2] Tunakova Yu.A., Novikova S.V., Shagidullin A.R., Valiev V.S., "Approaches for ensuring the technosphere safety of the urban environment using neural network modeling methods", *KhKhI vek. Tekhnosfernaya bezopasnost [XXI century. Technosphere safety]*, **5**:1(17) (2020), 21–28 (in Russian).

[3] Kremleva E.Sh., Novikova S.V., Shagidullin A.R., "Integrated assessment of the state of the environment based on an automatic neural network recognizer", *Sbornik trudov mezhdunarodnoj nauchnoj konferentsii (shkola molodykh uchenykh) Khimiya i inzhenernaya ekologiya - XIKh [Proceedings of the International Scientific Conference (School of Young Scientists) Chemistry and Engineering Ecology - XIX]*, 2019, 224–227 (in Russian).

[4] Tarantola A., *Inverse Problem Theory Methods for Data Fitting and Model Parameter Estimation*, Elseveir, Amsterdam, 1987, 644 pp.

[5] Sergeev Ya.D., Kvasov D.E., *Diagonal methods of global optimization*, Fizmatlit Publ., Moscow, 2008 (in Russian), 352 pp.

[6] Piyavskij S.A., "An algorithm for finding the absolute extremum of a function", *USSR Computational Mathematics and Mathematical Physics*, **12**:4 (1972), 57–67.

[7] Yevtushenko Yu.G., "Numerical methods for finding global extrema (Case of a non-uniform mesh)", *USSR Computational Mathematics and Mathematical Physics*, **11**:6 (1971), 38–54.

[8] Strongin R.G., *Numerical methods in multiextremal problems*, Nauka Publ., Moscow, 1978 (in Russian), 240 pp.

[9] Vanderbei R.J., "Extension of Piyavskii's Algorithm to Continuous Global Optimization", *Journal of Global Optimization*, **14** (1999), 205–216.

[10] Zabotin V.I., Chernyshevskij P.A., "Two modifications of extension of piyavskii's global optimization algorithm to a function continuous on a compact interval and its convergence", *Vestnik TvGU. Seriya: Prikladnaya Matematika [Herald of Tver State University. Series: Applied Mathematics]*, 2021, № 3, 70–85 (in Russian), https://doi.org/10.26456/vtpmk624.

[11] Arutyunova N.K., "Yevtushenko's method for finding ε-Lipschitzian function global minimum and its application", *Vestnik KGTU im. A.N. Tupoleva [Herald of the KSTU-KAI Named After A.N. Tupolev]*, 2013, № 2, 154–157 (in Russian).

[12] Zabotin V.I., Chernyshevskij P.A., "Extension of Strongin's Global Optimization Algorithm to a Function Continuous on a Compact Interval", *Computer Research and Modeling*, **11**:6 (2019), 1111–1119.

[13] Arutyunova N.K., Dulliev A.M., Zabotin V.I., "Global optimization of multivariable functions satisfying the Vanderbei condition", *Journal of Applied Mathematics and Computing*, **68**:2 (2022), 1135–1161, https://doi.org/10.1007/s12190-021-01563-4.

[14] Rostovtsev V.S., *Iskusstvennye nejronnye seti [Artificial neural networks]*, Lan, 2019 (in Russian), 216 pp.