Алкезуини М.М., Горбаченко В.И., Жуков М.В. Решение уравнений в частных производных на сетях радиальных базисных функций. // Проблемы информатики в образовании, управлении, экономике и технике: Сб. статей XV Междунар. научно-техн. конф. – Пенза: ПДЗ, 2015. – С. 58-66.

УДК 004.032.26

РЕШЕНИЕ УРАВНЕНИЙ В ЧАСТНЫХ ПРОИЗВОДНЫХ НА СЕТЯХ РАДИАЛЬНЫХ БАЗИСНЫХ ФУНКЦИЙ[1]

М.М. Алкезуини, В.И. Горбаченко, М.В. Жуков

SOLUTION PARTIAL DIFFERENTIAL EQUATIONS ON RADIAL BASIS FUNCTIONS NETWORKS

M.M. Alqezweeni, V.I. Gorbachenko, M.V. Zhukov

Аннотация. В обзоре рассматриваются методы решения дифференциальных уравнений в частных производных на сетях радиальных базисных функций. Основное внимание уделено обучению сетей.

Ключевые слова: дифференциальные уравнения в частных производных, сети радиальных базисных функций.

Abstract. The review considers methods for solving differential equations in partial derivatives on the networks of radial basis functions. The focus is on learning networks.

Keywords: partial differential equations**,** radial basis functions networks.

Для численного решения дифференциальных уравнений в частных производных, описывающих системы с распределенными параметрами, наибольшее распространение получили методы конечных разностей и конечных элементов. Эти методы требуют построения расчетных сеток. Несмотря на большое число работ, посвященных построению сеток, генерация сеток для двух и трехмерных областей сложной конфигурации остается сложной и трудоемкой задачей. Трудоемкость формирования сетки для реальных задач зачастую превосходит трудоемкость решения системы разностных уравнений [1, 2]. Большие вычислительные затраты приводят к использованию аппроксимаций низкого порядка, которые обеспечивают непрерывную аппроксимацию решения на сети, но не его частных производных. Моделирование объектов с распределенными параметрами методами конечных разностей и конечных элементов сводится к решению разреженных систем алгебраических уравнений очень большой размерности. Эти системы характеризуются плохой обусловленностью, что требует больших затрат на их решение. Восстановление решения по его дискретной аппроксимации является отдельной достаточно трудоемкой задачей.

Подходом, альтернативным сеточным методам, является использование различных бессеточных методов [3–4]. Большинство бессеточных методов относится к классу проекционных методов, т. е. методов, в которых аппроксимация решения ищется в виде взвешенной суммы базисных функций,

[1]Работа поддержана грантами РФФИ 14-01-00660 «Методы построения нейросетевых и гибридных математических моделей процессов и явлений в сложных технических системах» и 14-01-00733 «Информационные модели на основе иерархических гетерогенных нейронных сетей в исследовании влияния объектов транспортной инфраструктуры на окружающую среду».

а веса функций выбираются таким образом, чтобы приближенное решение удовлетворяло краевой задаче в выбранных пробных точках.

Среди бессеточных методов широкое распространение получили методы, использующие в качестве базисных функций радиальные базисные функции (РБ-функции, Radial Basis Functions – RBF) [2, 6–8]. Методы, использующие РБ-функции, позволяют получить дифференцируемое решение в любой точке области решения в виде функции, удовлетворяющей требуемым условиям гладкости, являются универсальными, позволяют работать со сложной геометрией расчетных областей, применимы для решения задач любой размерности. Методы, основанные на РБ-функциях, требуют при выбранных параметрах радиальных базисных функций нахождения вектора весов, так, чтобы получаемое приближенное решение обеспечивало с допустимой погрешностью выполнение уравнения и граничных условий на некотором наборе пробных точек. Главный недостаток использования РБ-функций – необходимость неформализуемого подбора параметров базисных функций.

От этого недостатка свободен особый вид нейронных сетей – сети радиальных базисных функций (РБФ-сети, Radial Basis Network) [8–16]. Главное достоинство РБФ-сетей состоит в использовании принципов обучения для формирования оптимальных параметров радиальных базисных функций.

РБФ-сеть – сеть, состоящая из двух слоев. Первый слой осуществляет нелинейное преобразование входного вектора $\mathbf{x} = (x_1, x_2, ..., x_d)$. Второй – производит линейное суммирование. В качестве функции преобразования используются РБ-функции. Выход сети описывается выражением $u(\mathbf{x}) = \sum_{m=1}^{M} w_m \varphi_m(\mathbf{x};\mathbf{p}_m)$, где M – количество РБФ функций; w_m, $\mathbf{p}_m$ – вес и параметры РБ-функции φ_m. Структура такой сети представлена на рис. 1.

Процесс решения прямых задач математической физики с помощью РБФ-сетей рассмотрим на примере задачи, заданной в операторном виде

$$Lu(\mathbf{x}) = f(\mathbf{x}),\ \mathbf{x} \in \Omega, \tag{1}$$

$$Bu(\mathbf{x}) = p(\mathbf{x}),\ \mathbf{x} \in \partial\Omega, \tag{2}$$

где u – искомое решение; L – дифференциальный оператор; оператор B задает граничные условия; Ω – область решения; $\partial\Omega$ – граница области; f и p – известные функции.

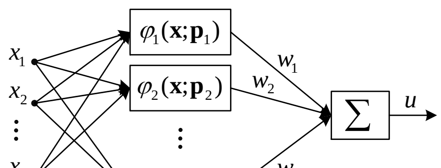

Рис. 1. Структура РБФ-сети

В области решения выберем множество узлов (пробных точек) $\Theta = \{(\mathbf{x}_i)|_{i=1,N-M} \subset \Omega, (\mathbf{x}_i)|_{i=N-M+1,N} \subset \partial\Omega\}$, где N – общее количество пробных точек; M – количество граничных пробных точек из $\partial\Omega$. Решение будем искать в виде

$$u_{RBF}(\mathbf{x}) = \sum_{j=1}^{N} w_j \varphi_j(\mathbf{x}), \qquad \mathbf{x} \in \bar{\Omega} = \Omega \cup \partial\Omega,$$

где φ_j – РБ-функция; w_j – неизвестные коэффициенты (w_j – вес функции φ_j), причём $\sum_{j=1}^{N} w_j = 0$.

Каждая РБ-функция определяется своими параметрами, например, Гауссиан определяется центром $\mathbf{c}$ и шириной a.

Решение формируется в процессе обучения сети, то есть необходимо так настроить w и p, чтобы функционал ошибки, представляющий собой сумму квадратов невязок в пробных точках для уравнения (1) и условия (2), принимал минимальное значение

$$J(\mathbf{w},\mathbf{p}) = \sum_{i=1}^{N} \left[Lu(\mathbf{x}_i) - f(\mathbf{x}_i) \right]^2 + \lambda \sum_{i=N+1}^{N+K} \left[Bu(\mathbf{x}_i) - p(\mathbf{x}_i) \right]^2 \to min, \tag{3}$$

где $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_N \in \Omega$, $\mathbf{x}_{N+1}, \mathbf{x}_{N+2}, ..., \mathbf{x}_{N+K} \in \partial\Omega$; λ – весовой множитель.

Известны два подхода к обучению РБФ-сети:

параметры РБ-функций выбираются, а в процессе обучения настраиваются только веса РБ-функций;

в процессе обучения сети настраиваются не только веса, но и центры и ширина РБ-функций.

При первом подходе [9] для линейных операторов в (1)–(2) задача сводится к решению линейной задачи наименьших квадратов, представляющей собой переопределенную систему линейных алгебраических уравнений (количество пробных точек больше количества РБ-функций). Подход позволяет эффективно, с точки зрения скорости и точности, решать линейные дифференциальные уравнения, однако, при решении нелинейных уравнений он требует применения квазилинеаризации, что резко ухудшает качественные характеристики подхода. Но главным недостатком подхода является необходимость трудно формализуемого подбора параметров РБ-функций. В результате теряются преимущества РБФ-сетей, как обучаемых нейронных сетей. По сути – это применение радиальных базисных функций.

Методы минимизации функционала ошибки, в которых настраиваются все параметры РБ-функций, являются итерационными. Можно выделить две группы методов. К первой группе относятся методы, в которых обучение сети проходит в два этапа: поочерёдно обучаются линейные (веса) и нелинейные параметры сети РБ-функций, причём для линейных и нелинейных параметров используются различные методы оптимизации [10, 16]. Ко второй группе относятся методы, в которых обучаются сразу все параметры функций [17].

Раздельное обучение линейных и нелинейных параметров в методах первой группы позволяет понизить размерность задачи, использовать для обучения весов высокоэффективные методы минимизации линейного квадратичного функционала. Методы хорошо подходят для решения линейных и линеаризованных задач. Каждая итерация этих методов состоит из двух шагов:

Шаг 1. Зафиксировать центры и ширину РБ-функций и выполнить несколько итераций метода градиентного спуска.

Шаг 2. Зафиксировать веса РБ-функций и выполнить несколько итераций метода градиентного спуска обучения ширины и центров РБ-функций. Причем параметр (скорость) обучения приходится подбирать.

В работе [16] предложен новый алгоритм метода сопряженных градиентов настройки весов РБФ-сети, отличающийся от известных учетом специфики решения краевых задач [16]. Алгоритм основан на матрично-векторном представлении функционала ошибки (3)

$$J = 0{,}5\left[(\mathbf{r}_1,\mathbf{r}_1) + \lambda(\mathbf{r}_2,\mathbf{r}_2)\right],$$

где $(\mathbf{r}_1,\mathbf{r}_1)$ и $(\mathbf{r}_2,\mathbf{r}_2)$ – скалярные произведения; $\mathbf{r}_1 = \mathbf{Mw} - \mathbf{f}$ – вектор невязки решения во внутренних пробных точках; $\mathbf{r}_2 = \mathbf{Nw} - \mathbf{p}$ – вектор невязки решения в граничных пробных точках; $\mathbf{M}$ – матрица $N \times m$, элементы которой являются значениями оператора $Lu(\mathbf{x}_i)$ во внутренних пробных точках; $\mathbf{N}$ – матрица $K \times m$, элементы которой являются значениями оператора $Bu(\mathbf{x}_i)$ в граничных пробных точках; N и K – количество внутренних и граничных пробных точек; $\mathbf{f}$ и $\mathbf{p}$ – векторы значений функций $f(\mathbf{x}_i)$ и $p(\mathbf{x}_i)$ в пробных точках.

Тогда функционал ошибки может быть представлен в виде

$$I = (\mathbf{Aw},\mathbf{w}) - 2(\mathbf{s},\mathbf{w}) + 0{,}5(\mathbf{f},\mathbf{f}) + 0{,}5\lambda(\mathbf{p},\mathbf{p}), \qquad (4)$$

где $\mathbf{A} = 0{,}5(\mathbf{M}^{\mathrm{T}}\mathbf{M} + \lambda\mathbf{N}^{\mathrm{T}}\mathbf{N})$; $\mathbf{s} = 0{,}5(\mathbf{M}^{\mathrm{T}}\mathbf{f} + \lambda\mathbf{N}^{\mathrm{T}}\mathbf{p})$; $\mathbf{A}$ – симметричная положительно определенная матрица.

Таким образом, задача настройки весов РБФ-сети сводится к задаче минимизации квадратичного функционала с симметричной положительно определенной матрицей. Эта задача легко решается классическим методом сопряженных градиентов, который значительно проще известных в обучении нейронных сетей методов Полака-Райбера и Флетчера-Ривса [18]. Предложенный алгоритм при решении стационарной задачи показал сокращение числа циклов обучения в 6 раз по сравнению с известным вариантом алгоритма сопряженных градиентов [16]. И это при том, что параметры РБ-функций (Гауссиана в данном случае) настраивались простейшим методом градиентного спуска.

Обучение РБФ-сети является плохо обусловленной задачей [18]. Наш подход заключается в исключении операций с плохо обусловленными матрицами $\mathbf{M}^{\mathrm{T}}\mathbf{M}$ и $\mathbf{N}^{\mathrm{T}}\mathbf{N}$. Например, в [19] предложена модификация метода сопряженных градиентов настройки весов РБФ-сети, основанная на идеях метода сопряженных градиентов для метода наименьших квадратов (CGLS – Conjugate Gradient Method for Least-Squares) [20]. Модификация исключает операции с плохо обусловленными матрицами $\mathbf{M}^{\mathrm{T}}\mathbf{M}$ и $\mathbf{N}^{\mathrm{T}}\mathbf{N}$ и дополнительно уменьшает время обучения сети.

Градиентные методы являются локально сходящимися, поэтому качество получаемого с их помощью решения во многом зависит от начальных параметров сети, при этом четких правил, которые бы указывали на "хорошие" начальные значения параметров, нет. Кроме того, градиентные методы не учитывают информацию о кривизне минимизируемой функции. Для преодоления указанных недостатков в [21] предложено использовать метод доверительных областей (МДО) для обучения РБФ-сетей с подвижным функциональным базисом, используемых для построения математических моделей систем с распределенными параметрами [17]. Ключевыми особенностями метода являются возможность одновременной оптимизации большого количества параметров; высокие показатели эффективности и сходимости даже для плохо обусловленных задач; за счет использования аппроксимационных моделей метод позволяет преодолевать локальные минимумы; позволяет минимизировать выпуклые функции, т.е. функции с отрицательно определенной матрицей Гессе; и, наконец, при использовании в роли аппроксимирующей функции разложения в ряд Тейлора второго порядка, задача минимизации целевой функции сводится к задаче минимизации квадратичного функционала. Все это делает МДО идеальным для решения задачи минимизации функционала (3), которая часто имеет и большое количество оптимизируемых параметров, и множество локальных минимумов, и плохо обусловленную матрицу Гессе.

Основная идея метода заключается в том, что на каждой итерации k минимизации функции $f(\mathbf{x})$, где $x \in \Omega \subseteq R^n$, Ω – область определения, функция f в некоторой доверительной области $B_k \subseteq \Omega$ заменяется аппроксимирующей её функцией m_k и ищется минимум m_k в B_k, который становится новым минимумом f. В зависимости от того, насколько уменьшение, предсказанное моделью, подтверждается целевой функцией, принимается решение о сужении либо расширении доверительной области.

МДО позволяет найти глобальный минимум лишь в некоторой локальной области. Чтобы найти минимум функционала (4), близкий к глобальному, необходимо либо разработать ряд рекомендаций по выбору "хорошей" начальной структуры РБФ-сети, либо использовать МДО совместно с глобальными методами оптимизации, такими как метод рестартов, генетический алгоритм, алгоритм имитации отжига и др. Альтернативный подход заключается в использовании иерархий РБФ-сетей для решения краевых задач. Его суть заключается в том, что сначала строится грубая аппроксимация решения, для этого используется РБФ-сеть с относительно небольшим количеством базисных функций. Затем полученное решение сглаживается с помощью какого-нибудь сглаживающего оператора, На каждой последующей итерации решение находится в виде суммы решения, полученного на предыдущей итерации, и выхода новой РБФ-сети с большим количество базисных функций.

Многие практически чрезвычайно важные классы реальных задач, например, задачи моделирования процессов разработки нефтяных месторождений, моделирования геофильтрации, теплофизики описываются как краевые задачи математической физики для неоднородных и нелинейных сред. Но известные методы, основанные на применении РБФ-сетей, позволяют моделировать только однородные среды. Проблема заключается в том, что вычисление невязки и градиента функционала требует вычисления производных от функции, описывающей среду. Причем эта функция известна экспериментально в некоторых точках. В [19] предложено использовать аппроксимацию функции среды с помощью РБФ-сети, что позволяет проводить аналитическое дифференцирование этой функции и использовать РБФ-сети для моделирования процессов в неоднородных средах.

Для моделирования процессов в нелинейных средах целесообразно использовать итерационный пересчет функции среды и решать задачу с зафиксированным значением функции на каждой итерации. Такой подход не требует линеаризации дифференциального оператора.

При решении нестационарных задач можно рассматривать время как одну из координат, и решать задачу рассмотренными ранее алгоритмами. Но увеличение размерности существенно усложняет задачу. Предпочтительнее использовать неявную конечно-разностную аппроксимацию по временной переменной и нейросетевую аппроксимацию по пространственным переменным. Таким образом, решение нестационарной задачи сводится к многократному решению стационарной на каждом временном слое.

При моделировании систем с распределенными параметрами часто приходится сталкиваться с ситуациями, когда необходимо определить не только распределение параметров, но и идентифицировать неизвестные свойства системы по дополнительным данным, заданным с погрешностью. Для построения устойчивых нейросетевых моделей таких систем необходимо использовать регуляризованные алгоритмы обучения РБФ-сетей. Один из таких алгоритмов предложен в работе [22], где для регуляризации решения используется условие Морозова.

В качестве перспективных направления исследований в области применения сетей радиальных базисных функций для решения уравнений в частных производных следует выделить:

- развитие методов обучения РБФ-сетей;
- разработка многоуровневых методов решения краевых задач с использованием иерархий РБФ-сетей;
- совершенствование нейросетевых методов решения задач для неоднородных и нелинейных сред, совершенствования нейросетевых методов решения нелинейных задач;
- совершенствование методов обучения РБФ-сетей при решении обратных задач, включая разработку способов регуляризации на основе регуляризации Тихонова, метода итераций Ландвебера, регуляризованного метода сопряженных направлений;
- совершенствование методов решения нестационарных задач, включая реакционно-диффузионные уравнения;
- разработка параллельных алгоритмов обучения РБФ-сетей, используемых для моделирования систем с распределенными параметрами.

Библиографический список

1. Толстых А.И., Широбоков Д.А. Бессеточный метод на основе радиальных базисных функций // Журнал вычислительной математики и математической физики. – 2005. – Т. 45. № 8. – С. 1498–1505.

2. Kansa E.J. Motivation for using Radial Basis Function to solve PDEs // [Электронный ресурс]. – Режим доступа: http://www.cityu.edu.hk/rbf-pde/files/overview-html.html

3. Meshless Methods: An Overview and Recent Developments / T. Belytschko, Y. Krongauz, D. Organ, M. Fleming, P. Krysl // Computers Methods in Applied Mechanics and Engineering. – 1996, Vol. 139. – Issue 1–4. – P. 3–47.

4. Liu G.R. Mesh free methods: moving beyond the finite element method. – CRC Press, 2003. – 712 p.

5. Meshfree Methods for Partial Differential Equations / Editors M. Griebel, Marc. A. Schweitzer. – Springer, 2008. – 412 p.

6. Buhmann M.D. Radial Basis Functions: Theory and Implementations. – Cambridge University Press, 2004. – 259 p.

7. Fasshauer G. E. Meshfree Approximation Methods with MATLAB. – World Scientific Publishing Company, 2007. – 520 p.

8. Chen W., Fu Z.-J. Recent Advances in Radial Basis Function Collocation Methods. – Springer, 2014. – 90 p.

9. Mai-Duy N., Tran-Cong T., Numerical solution of differential equations using multiquadric radial basis function networks // Neural Networks, 2001, vol. 14. – Issue 2. – P. 185–199.

10. Jianyu L., Siwei L., Yingjian Q., Yaping H. Numerical solution of elliptic partial differential equation by growing radial basis function neural networks // Neural Networks, 2003, vol. 16. – Issue 5–6. – P. 729–734.

11. Chen H., Kong L., Leng W., Numerical solution of PDEs via integrated radial basis function networks with adaptive training algorithm // Applied Soft Computing, 2011, vol. 11. – Issue 1. – P. 855–860.

12. Multilayer perceptrons and radial basis function neural network methods for the solution of differential equations: A survey / Manoj Kumar, NehaYadav // Computers and mathematics with applications 62 (2011). – November 2011. – P. 3796–3811.

13. Yadav N., Yadav A., Kumar M. An Introduction to Neural Network Methods for Differential Equations. – Springer, 2015. – 115 p.

14. Васильев А. Н., Тархов Д. А. Нейросетевое моделирование: Принципы. Алгоритмы. Приложения. – СПб.: Издательство Политехнического университета, 2009. – 527 с.

15. Тархов Д. А. Нейросетевые модели и алгоритмы : справочник. – М.: Радиотехника, 2014. – 352 с.

16. Горбаченко В. И., Артюхина Е. В. Бессеточные методы и их реализация на радиально-базисных нейронных сетях // Нейрокомпьютеры: разработка, применение. – 2010. – № 11. – С. 4–10.

17. Горбаченко В.И., Жуков М. В. Обучение сетей радиальных базисных функций методом доверительных областей для решения уравнения Пуассона // Информационные технологии. – 2013. – №9. – С. 65–70.

18. Хайкин С. Нейронные сети: полный курс. – М.: Вильямс, 2006. – 1104 с.

19. Горбаченко В. И., Артюхина Е. В. Метод сопряженных градиентов для задачи наименьших квадратов и его применение для обучения весов радиальных базисных нейронных сетей при решении дифференциальных уравнений в частных производных // Нейроинформатика-2011: сб. научных трудов XIII Всероссийской научно-технической конференции. –. Ч. 3. – М.: НИЯУМИФИ, 2010. – С. 219–228.

20. Vorst van der, H. Iterative Krylov Methods for Large Linear Systems. – Cambridge: Cambridge University Press, 2003. – 232 p.

21. Conn A. R., Gould N. M., Toint P. L. Trust regions methods. – MPS-SIAM series on optimization, 2000. – 959 p.

22. Жуков М. В. Решение коэффициентных обратных задач математической физики с помощью сетей радиальных базисных функций // Нейрокомпьютеры: разработка и применение. – 2014. – №2. – С.32–39.

Алкезуини Мухи Муртада
Пензенский государственный университет, г. Пенза, Россия
E-mail: mohieit@mail.ru

Alqezweeny M.M.
Penza State University,
Penza, Russia

Горбаченко Владимир Иванович
Пензенский государственный
университет, г. Пенза, Россия
E-mail: gorvi@mail.ru

Gorbachenko V.I.
Penza State University,
Penza, Russia

Жуков Максим Валерьевич
Пензенский государственный
университет, г. Пенза, Россия
E-mail: maxim.zh@gmail.com

Zhukov M.V.
Penza State University,
Penza, Russia