Алкезуини М.М., Горбаченко В.И., Жуков М.В. Обучение сетей радиальных базисных функций для решения краевых задач математической физики. // Проблемы информатики в образовании, управлении, экономике и технике: Сб. статей XVI Междунар. научно-техн. конф. – Пенза: ПДЗ, 2016. – С. 103-109.

УДК 004.032.26

ОБУЧЕНИЕ СЕТЕЙ РАДИАЛЬНЫХ БАЗИСНЫХ ФУНКЦИЙ ДЛЯ РЕШЕНИЯ КРАЕВЫХ ЗАДАЧ МАТЕМАТИЧЕСКОЙ ФИЗИКИ

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

SOLUTION OF BOUNDARY VALUE PROBLEMS OF MATHEMATICAL PHYSICS ON THE NETWORKS OF RADIAL BASIS FUNCTIONS

M.M. Alqezweeny, V.I. Gorbachenko, M.V. Zhukov

Аннотация. Рассмотрено решение краевых задач математической физики на сетях радиальных базисных функций. Предложен метод обучения сетей на основе метода доверительных областей, позволяющий упростить процесс выбора структуры сети и существенно сократить временные затраты на настройку её параметров. Разработан алгоритм обучения на основе метода доверительных областей.

Ключевые слова: краевые задачи математической физики, сети радиальных базисных функций, обучение нейронных сетей, метод доверительных областей.

Abstract. A network of training method based on the method of trust region, to simplify the process of selecting the network structure and significantly reduce the time required for setting its parameters. A learning algorithm based on trust region method.

Keywords: boundary value problems of mathematical physics, the network of radial basis functions neural network training method trust region.

Сеть радиальных базисных функций (РБФ-сеть) – сеть, состоящая из двух слоев [1]. Первый слой осуществляет нелинейное преобразование входного вектора $\mathbf{x} = (x_1, x_2, \ldots, x_d) \in \mathbb{R}^d$. Второй слой производит линейное суммирование. В качестве функции преобразования используются радиальные бпзисные функции (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 РБ-функций, w_m – вес РБ-функции φ_m, $\mathbf{p}_m$ – вектор параметров РБ-функции. В качестве РБ-функций используют функции Гаусса, мультиквадрики, обратные мультиквадрики и др. [1].

Процесс решения краевых задач с помощью РБФ-сетей рассмотрим на примере краевой задачи, заданной в операторной форме,

$$Lu(\mathbf{x}) = f(\mathbf{x}),\ \mathbf{x} \in \Omega, \qquad Bu(\mathbf{x}) = p(\mathbf{x}),\ \mathbf{x} \in \partial\Omega,$$

где u – искомое решение; L – дифференциальный оператор; B – оператор граничных условий; Ω – область решения; $\partial\Omega$ – граница области; f и p – известные функции. Решение u аппроксимируется с помощью РБФ-сети, для этого:

1. Из множеств Ω и $\partial\Omega$ выбираются N внутренних и K граничных пробных точек.

2. Определяется структура сети: тип сети, количество РБ-функций M, тип РБ-функций, задаются начальные значения: $\mathbf{w} = (w_1, w_2, \ldots, w_M)$ – веса РБ-функций,

$\mathbf{p}=(\mathbf{p}_1,\mathbf{p}_2,...,\mathbf{p}_M)$ – параметры РБ-функций (структура элементов вектора **p** зависит от вида функции, например для двухмерной функции Гаусса $\mathbf{p}_i=(a_i,c_i^1,c_i^2)$, где $i=\overline{1...M}$, a_i – параметр формы (ширина) *i*-ой РБ-функции, (c_i^1,c_i^2) – положение её центра).

3. Выполняется обучение сети, т.е. подбираются такие значения **w** и **p**, чтобы функционал ошибки, представляющий собой сумму квадратов невязок в пробных точках, принимал минимальное значение

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

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 – подбираемый штрафной множитель, учитывающий вклад невязок в граничных пробных точках.

Обученная сеть при подаче на вход координат произвольной точки из области решения или границы формирует решение в этой точке.

Эффективность нейросетевого метода решения краевых задачзависит от эффективности метода решения задачи минимизации функционала ошибки (1). Для решения задачи (1) перспективен метод доверительных областей (МДО) [2], модифицированный для обучения РБФ-сетей [3–4]. Особенностями метода являются возможность одновременной оптимизации большого количества параметров и высокие показатели эффективности и сходимости даже для плохо обусловленных задач.

Основная идея МДО заключается в том, что на каждой итерации *k* минимизации функции $f(\mathbf{x})$, где $\mathbf{x}\in\Omega\subseteq\mathbb{R}^n$ (Ω – область определения, $\mathbb{R}$ – множество действительных чисел), функция f в некоторой доверительной области $B_k\subseteq\Omega$ заменяется аппроксимирующей ее функцией m_k и вычисляется 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. Формальное описание алгоритма, реализующего метод доверительных областей, имеет вид:

Шаг 1. Инициализация. Задаются предварительное положение $\mathbf{x}_0$; радиус доверительной области Δ_0; пороговые значения оценок точности модели μ_1 и μ_2 такие, что $0<\mu_1\le\mu_2<1$; коэффициенты преобразования доверительной области γ_1 и γ_2 ($0<\gamma_1\le\gamma_2<1$); порядковый номер итерации $k=0$.

*Шаг 2***.** Аппроксимация f. Выбирается норма $\|\cdot\|$, строится функция m_k, аппроксимирующая функцию f. в области B_k.

Шаг 3. Минимизация m_k. Выбирается метод условной минимизации m_k. С его помощью находится такой шаг $\mathbf{s}_k$, что точка $\mathbf{x}_k+\mathbf{s}_k$ является глобальным минимумом m_k в области B_k.

Шаг 4. Оценка точности модели p_k. Вычисляется

$$p_k=\frac{f(\mathbf{x}_k)-f(\mathbf{x}_k+\mathbf{s}_k)}{m_k(\mathbf{x}_k)-m_k(\mathbf{x}_k+\mathbf{s}_k)}.$$

Если $p_k\ge\mu_1$, то $\mathbf{x}_{k+1}=\mathbf{x}_k+\mathbf{s}_k$, иначе $\mathbf{x}_{k+1}=\mathbf{x}_k$.

Шаг 5. Изменение радиуса доверительной области.

$$\Delta_{k+1} \in \begin{cases} [\Delta_k; \infty), \text{если } p_k \geq \mu_2, \\ [\gamma_2 \Delta_k; \Delta_k), \text{если } p_k \in [\mu_1, \mu_2), \\ [\gamma_1 \Delta_k; \gamma_2 \Delta_k), \text{если } p_k < \mu_1. \end{cases}$$

Шаг 6. Увеличить порядковый номер итерации $k = k+1$. Если достигнута требуемая точность решения или k равен максимальному количеству итераций, или радиус доверительной области слишком мал, то завершить обучение, иначе перейти к шагу 2.

Поскольку функционал ошибки сети – дважды дифференцируемая функция, то для построения m_k применим разложение J по формуле Тейлора второго порядка, а в качестве нормы – евклидову норму. Это приводит к необходимости вычисления матрицы Гессе, что требует больших вычислительных затрат. Вместо точного значения матрицы Гессе будем использовать ее приближенное значение, представляющее собой произведение матриц Якоби $\mathbf{H}(\mathbf{x}) \approx \mathbf{G}(\mathbf{x}) = [\mathbf{J}(\mathbf{x})]^T \mathbf{J}(\mathbf{x})$, где Якобиан векторфункции ошибки имеет вид

$$\mathbf{J}(\mathbf{q}) = \begin{bmatrix} \dfrac{\partial r_1}{\partial q_1} & \dfrac{\partial r_1}{\partial q_2} & \cdots & \dfrac{\partial r_1}{\partial q_{M\cdot(1+d+l)}} \\ \dfrac{\partial r_2}{\partial q_1} & \dfrac{\partial r_2}{\partial q_2} & \cdots & \dfrac{\partial r_2}{\partial q_{M\cdot(1+d+l)}} \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial r_{N+K}}{\partial q_1} & \dfrac{\partial r_{N+K}}{\partial q_2} & \cdots & \dfrac{\partial r_{N+K}}{\partial q_{M\cdot(1+d+l)}} \end{bmatrix},$$

где $r_i(\mathbf{q}) = \begin{cases} Lu(\mathbf{x}_i; \mathbf{q}) - f(\mathbf{x}_i), \ 1 \leq i \leq N, \\ \sqrt{\lambda}\left[Bu(\mathbf{x}_i; \mathbf{q}) - p(\mathbf{x}_i)\right], \ N < i \leq N+K \end{cases}$, $\mathbf{q} = (w^1, p_1^1, ..., p_l^1, ..., w^M, p_1^M, ..., p_l^M)$.

Использование разложения по формуле Тейлора второго порядка приводит к необходимости решения условной задачи минимизации квадратичного функционала. Для этого предлагается использовать метод Стайхауга [5]. Метод представляет собой модификацию предобусловленного метода сопряженных направлений, учитывающую при минимизации функционала m_k ограничения на решение (решение должно лежать в B_k) и позволяющую работать с отрицательно определенной матрицей Гессе.

В качестве рассмотрим решение краевой задачи для двухмерного уравнения Пуассона

$$\frac{\partial^2 u(x,y)}{\partial x^2} + \frac{\partial^2 u(x,y)}{\partial y^2} = f(x,y), \ (x,y) \in \Omega, \quad u(x,y) = p(x,y), \qquad (x,y) \in \partial\Omega.$$

Областью решения является квадрат, ограниченный точками (0,0) и (1,1), а $f(x,y) = \sin(\pi x)\sin(\pi y)$, $p(x,y) = 0$. Данная задача имеет аналитическое решение $u_a(x,y) = -\frac{1}{2\pi^2}\sin(\pi x)\sin(\pi y)$, что позволит оценить точность полученного численного решения. Функционал ошибки этой задачи имеет вид:

$$J(\mathbf{w}, \mathbf{p}) = \sum_{i=1}^{N} \left[\frac{\partial^2 u_{RBF}(x_i, y_i; \mathbf{w}, \mathbf{p})}{\partial x^2} + \frac{\partial^2 u_{RBF}(x_i, y_i; \mathbf{w}, \mathbf{p})}{\partial y^2} - \sin(\pi x_i)\sin(\pi y_j) \right]^2 +$$
$$+ \lambda \sum_{i=N+1}^{N+K} \left[u_{RBF}(x_i, y_i; \mathbf{w}, \mathbf{p}) \right]^2.$$

Проведено два эксперимента с различными значениями параметров сети. В обоих экспериментах для обучения сети использовались 144 случайно расположенные контрольные точки, 100 из которых располагались в области решения Ω, 44 –

на границе $\partial\Omega$. В качестве РБ-функций использовалась функция Гаусса. В первом эксперименте РБФ-сеть состояла из 16 нейронов, центры которых предварительно были случайным образом расположены в квадратной области, ограниченной точками $(-0,2;-0,2)$ и $(1,2;1,2)$. Ширина РБ-функций выбиралась случайным образом из интервала [0,3; 0,6], предварительные значения весов РБ-функций также выбирались случайным образом из интервала [−0,05; 0,05]. Штрафной параметр $\lambda = 1000$ Параметры МДО $\Delta_0 = 2$; $\mu_1 = 0,2$; $\mu_2 = 0,7$; $\gamma_1 = 0,5$; $\gamma_2 = 0,7$.

Обучение сети завершилось на 8 итерации со значением погрешности решения по сравнению с аналитическим решением, равной $1,7 \cdot 10^{-4}$. Погрешность решения рассчитывалась по формуле относительной среднеквадратической погрешности

$$\varepsilon = \sqrt{\sum_{i=1}^{N+K} \left(u_{RBF}(x_i, y_i; \mathbf{w}, \mathbf{p}) - u_a(x_i, y_i)\right)^2} \Bigg/ \sqrt{\sum_{i=1}^{N+K} u_a(x_i, y_i)^2}.$$

График погрешности представлен на рис. 1.

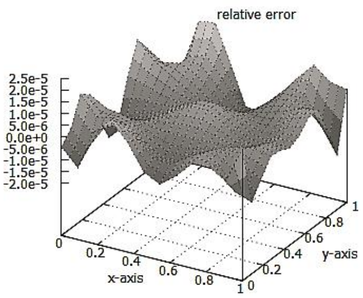

Рис. 1. Эксперимент №1. Погрешность решения

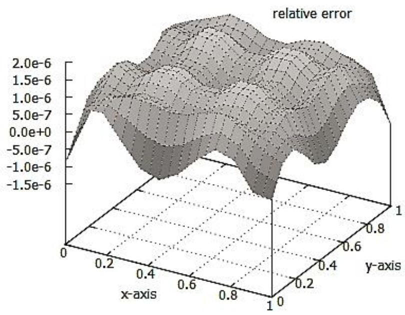

Рис. 2. Эксперимент №2. Погрешность решения

Во втором эксперименте РБФ-сеть состояла из 64 нейронов, центры которых располагались в узлах равномерной квадратной сетки, ограниченной точками (−0,2; −0,2) и (1,2; 1,2), ширина нейронов была постоянной и равнялась 0,5, предварительные значения весов РБ-функций равнялись 0. За 15 итерации была достигнута погрешность решения равная $5,3 \cdot 10^{-5}$. График погрешности представлен на рис. 2.

Как и следовало ожидать, чем больше нейронов в сети, тем точнее она аппроксимирует искомое решение. Однако процесс обучения такой сети требует значительно больше времени (для обучения сети из первого эксперимента потребовалась почти в два раза меньше итераций, чем для обучения сети из второго эксперимента).

Таким образом, в работе на основе метода доверительных областей разработан метод обучения РБФ-сетей с настраиваемым функциональным базисом. Для повышения его быстродействия используются приближенные значения матрицы Гессе. Метод позволяет сократить время обучения сетей и уменьшить погрешность нейросетевых моделей.

Библиографическийсписок

1. Haykin S.O. Neural Networks and Learning Machines. Pearson, 2008. 936 p.

2. Conn A.R., Gould N.I.M., Toint P.L. Trust-Region Methods. MPS-SIAM, 1987. 979 p.

3. Горбаченко В. И., Жуков М. В. Подходы и методы обучения сетей радиальных базисных функций для решения задач математической физики // Нейрокомпьютеры: разработка, применение. 2013. № 9. С.12–18.

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

5. Staihaug T. The conjugate gradient method and trust region in large scale optimization // SIAM Journal on Numerical Analysis. 1983, Vol 20. No 3. P. 626–637.

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