Горбаченко В.И., Артюхина Е.В. Обучение весов радиально – базисных нейронных сетей для решения краевых задач математической физики. // Проблемы информатики в образовании, управлении, экономике и технике: Сб. статей IX Междунар. научно-техн. конф. – Пенза: ПДЗ, 2009. – С. 71-74.

ОБУЧЕНИЕ ВЕСОВ РАДИАЛЬНО-БАЗИСНЫХ НЕЙРОННЫХ СЕТЕЙ ДЛЯ РЕШЕНИЯ КРАЕВЫХ ЗАДАЧ МАТЕМАТИЧЕСКОЙ ФИЗИКИ[1]

В.И. Горбаченко, Е.В. Артюхина

Пензенский государственный педагогический университет
им. В.Г. Белинского,
г. Пенза, Россия, gorvi@mail.ru

Разработаны градиентные алгоритмы обучения весов радиально-базисной нейронной сети для решения краевых задач математической физики. Экспериментально показано, что алгоритм сопряженных градиентов для минимизации квадратичного функционала позволяет достичь большей точности за меньшее время.

Gorbachenko V.I., Artyuhina E.V. Instruction the scales in radial-basis neural networks for solving boundary problems of mathematical physics.

The gradient algorithms of scales' teaching the radial-basis neural networks for solving boundary problems of mathematical physics are worked out. It is experimentally shown that the algorithm of conjugate gradients for minimization of quadratic functional permits the achievement of a greater precision for less time.

В настоящее время интенсивно развиваются методы решения ДУЧП с применением радиально-базисных нейронных сетей (RBFNN) [1]. Целью данной работы является исследование различных алгоритмов обучения весов RBFNN. В качестве модельной задачи для сравнительного анализа с [1] взято уравнение Пуассона

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = f(x,y), (x,y) \in \Omega, \; u = p(x,y), (x,y) \in \partial\Omega, \qquad (1)$$

где $\partial\Omega$ – граница области; f и p – известные функции (x,y).

Выбирая в качестве радиально-базисной функции Гауссиан, определяемый как $\varphi_k(x,y) = \exp\left(-r_k^2 / a_k^2\right)$, рассмотрим RBFNN как аппроксиматор функции $u(x) = \sum_{k=1}^{m} w_k \varphi_k(r_k)$, где m – число радиально-базисных функций (скрытых нейронов); w_k – веса сети; $r_k = \sqrt{(x - c_{xk})^2 + (y - c_{yk})^2}$; (c_{xk}, c_{yk})– координаты центра нейрона k; a_k – ширина нейрона k.

Обучение сети сводится к настройке весов, расположения центров и ширины нейронов, минимизирующих функционал качества (функционал ошибки), представляющий собой сумму квадратов невязок в контрольных точках

[1] Работа выполнена по тематическому плану научно-исследовательских работ Пензенского государственного педагогического университета, проводимых по заданию Федерального агентства по образованию.

$$I = 0,5(\mathbf{r}_1, \mathbf{r}_1) + 0,5\lambda(\mathbf{r}_2, \mathbf{r}_2), \qquad (2)$$

где λ – штрафной множитель; $\mathbf{r}_1$, $\mathbf{r}_2$ – векторы невязки решения во внутренних и граничных контрольных точках соответственно. С учетом вида радиально-базисных функций компоненты векторов $\mathbf{r}_1$ и $\mathbf{r}_2$ рассчитываются по формулам

$$\mathbf{r}_1 = \mathbf{Mw} - \mathbf{f}\ ,\ \mathbf{r}_2 = \mathbf{Nw} - \mathbf{p}, \qquad (3)$$

где $\mathbf{M}$ – матрица $N \times m$ с элементами $m_{ik} = 4\exp\left(-r_{ik}^2 / a_k^2\right)\left(r_{ik}^2 - a_k^2\right) / a_k^4$; $\mathbf{N}$ – матрица $K \times m$ с элементами $n_{ik} = \exp\left(-r_{ik}^2 / a_k^2\right)$; $\mathbf{w}$ – вектор весов; N и K – количество внутренних и граничных контрольных точек; $\mathbf{f}$ и $\mathbf{p}$ – векторы, компоненты которых равны значениям функций $f(x, y)$ и $p(x, y)$ в контрольных точках; i и k – номера контрольных точек и радиально-базисных функций соответственно.

Для обучения выходного слоя сети существенно большую скорость сходимости, чем метод скорейшего спуска, обеспечивает метод сопряженных градиентов, причем он требует всего порядка $O(m)$ арифметических операций. Классический метод сопряженных градиентов обучения нейронных сетей [2] может быть упрощен, если учесть линейный характер выходного слоя RBFNN. Для этого запишем функционал (2) в матричной форме. С учетом (3) имеем

$$\begin{aligned} I = {} & 0,5\left(\left(\mathbf{M}^{\mathrm{T}}\mathbf{M} + \lambda\mathbf{N}^{\mathrm{T}}\mathbf{N}\right)\mathbf{w}, \mathbf{w}\right) - \left(\left(\mathbf{M}^{\mathrm{T}}\mathbf{f} + \lambda\mathbf{N}^{\mathrm{T}}\mathbf{p}\right), \mathbf{w}\right) + \\ & + 0,5(\mathbf{f}, \mathbf{f}) + 0,5\lambda(\mathbf{p}, \mathbf{p}) \end{aligned}. \qquad (4)$$

Введем обозначения $\mathbf{A} = 0,5\left(\mathbf{M}^{\mathrm{T}}\mathbf{M} + \lambda\mathbf{N}^{\mathrm{T}}\mathbf{N}\right)$, $\mathbf{s} = 0,5\left(\mathbf{M}^{\mathrm{T}}\mathbf{f} + \lambda\mathbf{N}^{\mathrm{T}}\mathbf{p}\right)$, где $\mathbf{A}$ – симметричная положительно определенная матрица. Тогда (4) примет вид

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

Из (5) видно, что наша задача является задачей минимизации квадратического функционала с симметричной положительно определенной матрицей. Применяя метод сопряженных градиентов [3] для минимизации (5), получим алгоритм:

1. По заданному начальному приближению весов $\mathbf{w}^{(0)}$ вычисляется невязка $\mathbf{r}^{(0)} = \mathbf{s} - \mathbf{Aw}^{(0)}$. В качестве направления движения выбирается $\mathbf{p}^{(0)} = \mathbf{r}^{(0)}$.

2. Вычисляется номер текущей итерации $k = k + 1$.

3. Находится новое приближение решения $\mathbf{w}^{(k+1)} = \mathbf{w}^{(k)} + \alpha_k\mathbf{p}^{(k)}$, где $\alpha_k = \left(\mathbf{r}^{(k)}, \mathbf{r}^{(k)}\right) \big/ \left(\mathbf{Ap}^{(k)}, \mathbf{p}^{(k)}\right)$.

4. Вычисляется новая невязка $\mathbf{r}^{(k+1)} = \mathbf{r}^{(k)} - \alpha_k\mathbf{Ap}^{(k)}$.

5. Проверяется условие окончания итерационного процесса $\left\|\mathbf{r}^{(k+1)}\right\| \big/ \left\|\mathbf{s}\right\| \le \varepsilon$.

6. Определяется новое направление движения $\mathbf{p}^{(k+1)} = \mathbf{r}^{(k+1)} + \beta_k\mathbf{p}^{(k)}$, где $\beta_k = \left(\mathbf{r}^{(k+1)}, \mathbf{r}^{(k+1)}\right) \big/ \left(\mathbf{r}^{(k)}, \mathbf{r}^{(k)}\right)$.

7. Переход на шаг 2.

Из (5) видно, что рассмотренный алгоритм эквивалентен решению системы линейных алгебраических уравнений $\mathbf{Aw} = \mathbf{s}$. Метод сопряженных градиентов в случае положительно определенной матрицы $\mathbf{A}$ размером $n \times n$ при отсутствии ошибок округления позволяет получить решение не более чем за n итераций.

Экспериментальное исследование проводилось на примере модельной задачи (1) для $f(x,y) = \sin(\pi x) \cdot \sin(\pi y)$, $p(x,y) = 0$. Данная задача имеет аналитическое решение, для оценки погрешности решения рассчитывалась относительная среднеквадратическая погрешность решения.

В процессе исследования приведенных алгоритмов получены следующие результаты: достигнуто значение относительной среднеквадратической погрешности решения 0,0005, что лучше, чем 0,005 в [1], абсолютная погрешность по сравнению с аналитическим решением не превышает 0,00003. Алгоритм сопряженных градиентов для минимизации квадратичного функционала позволяет сократить время решения задачи в 6 раз по сравнению с классическим методом сопряженных градиентов, сокращение времени более значимое по сравнению с методом скорейшего спуска и градиентным методом с подбираемым коэффициентом обучения для весов.

Библиографический список

1. Numerical solution of elliptic partial differential equation using radial basis function neural networks / L. Jianyu, L. Siwei, Q. Yingjiana, H. Yapinga // Neural Networks. – 2003. – 16(5/6). – P. 729 – 734.

2. Оссовский С. Нейронные сети для обработки информации. – М. : Финансы и статистика, 2002. – 344 с.

3. Амосов А.А., Дубинский Ю.А., Копченова Н.В. Вычислительные методы. – М. : Изд. дом МЭИ, 2008. – 672 с.