

К.Е. БУРНАЕВ, БелГТУ им. Шухова, Белгород, Россия

ПРИБЛИЖЕННОЕ РЕШЕНИЕ УРАВНЕНИЙ ТЕПЛОПРОВОДНОСТИ С ИСПОЛЬЗОВАНИЕМ НЕЙРОННЫХ СЕТЕЙ

У статті розглядається можливість отримання наближеного рішення диференціальних рівнянь у частинних похідних за допомогою одношарових нейронних мереж із зворотними зв'язками. Досліджуються питання стабільності нейронної мережі. Представлений простий приклад застосування мереж вказаної структури для вирішення диференціального рівняння. Проведено порівняння результатів рішення з результатами, одержаними методом кінцевих елементів.

The possibility of solving partial derivative differential equations by using neural networks with specific structure is investigated in the article. The simple single-layer neural network ability to approximate the solution of PDE is proved. The issue of network stability is examined. A simple example of described method application is given and the simple result comparison with the finite element method is made.

Постановка задачи. Общий вид уравнения теплопроводности для стационарных тепловых процессов в n -мерном пространстве определяется уравнением Пуассона (или его частным случаем при $n = 0$ – уравнением Лапласа):

$$\Delta u = -g, \quad (1)$$

где функции $u(\vec{x})$, $g(\vec{x})$ определены на области $\Omega \subset E^n$. Задача формулируется следующим образом: найти такую функцию u , которая бы удовлетворяла уравнению Пуассона и некоторым условиям, заданным на границе Γ области Ω : $u(x) = v_1(x)$, $\left. \frac{\partial u}{\partial \vec{n}} \right|_{\Gamma} = v_2(x)$, $\left. \frac{\partial u}{\partial \vec{n}} \right|_{\Gamma} = \alpha(v_3(x) - u(x))$.

Заметим, что даже относительно простые задачи, возникающие на практике, могут содержать краевые условия всех трех типов и, как правило, имеют Ω «неудобной» – в частности, не прямоугольной – формы [1 – 5].

Задачи такого вида возникают при решении многих технических задач – расчете охлаждающих систем, пресс-форм и литьевых форм, печатных плат и тому подобного. Как правило, при решении таких задач пользователя главным образом интересует получение численного решения.

Обзор существующих методов решения задачи. Методы решения дифференциальных уравнений в частных производных можно разбить на две большие группы: аналитические и численные. Аналитические методы, такие как метод разделения переменных, метод функции Грина или методы, использующие интегральные преобразования, дают возможность получить

аналитическое выражение решения. Однако все эти методы, к сожалению, не могут быть применены для решения произвольной задачи, так как они накладывают жесткие ограничения на вид граничных условий задачи и (или) форму области, в которой идет поиск решения. В некоторых случаях с помощью искусственных методик возможно привести некоторые задачи к требуемому виду, однако, в общем случае это невозможно [3 – 5]. Отсутствие формальных правил решения произвольной задачи мешает применению аналитических методов в системах автоматизированного проектирования.

Численные методы, в частности, методы конечных разностей и конечных элементов, являются в достаточной степени универсальными методами. При их применении процесс сводится к дискретизации задачи и поиску решения системы линейных алгебраических уравнений. Как правило, системы автоматизированного проектирования используют именно численные методы. Однако проблема такого подхода состоит в том, что такие алгоритмы в силу своей последовательной природы плохо распараллеливаются между несколькими вычислительными устройствами. Учитывая все большее распространение распределенных вычислительных систем, возникает потребность в алгоритме, имеющем достоинства упомянутых метода конечных разностей, метода конечных элементов и позволяющем использовать преимущества параллельных вычислений [2].

Применение нейронных сетей. Создадим нейронную сеть, нейроны которой имеют линейные активационные функции $f(x) = x$ и каждый нейрон имеет $2^n + 1$ связей – с непосредственными соседями и самим собой. Обозначим множество номеров нейронов, являющихся непосредственными соседями i -го нейрона как N_i , номера нейронов, предшествующих i -му нейрону по j -й координате как P_{ji} , следующих по j -й координате как S_{ji} . Веса связей выберем, основываясь на следующих правилах:

$$\begin{aligned} w_{ii} &= b, \forall i, b \in (0,1); \\ w_{ij} &= \frac{1-b}{2^n}, i \neq j; \\ w_{ij} &= w_{ji}, \end{aligned} \quad (2)$$

где w_{ij} – вес связи, идущей от i -го нейрона к j -му. Проведем дискретизацию области Ω , покрыв ее прямоугольной сеткой прямых $x_{ji} = P_{ji}h$, где P_{ji} принимает все целые значения из интервала $\left[0, \frac{D_i}{h}\right]$.

Обозначим множество координат узлов сетки дискретизации как X . Поставим в соответствие значению искомой функции в узле сетки выходное значение нейрона $f_i = u(x_i)$, $x_i \in X$.

Заметим, что в стабильном состоянии нейронной сети будет выполняться неравенство $|f_i^{(t+1)} - f_i^{(t)}| < \varepsilon$, где ε – сколь угодно малая величина. Преобразуя неравенство очевидным образом, получаем:

$$\left| f_i^{(t+1)} - \left(b f_i^{(t)} + \sum_{j \in N_i} \frac{1-b}{2^n} f_j^{(t)} \right) \right| < \varepsilon, \quad (3)$$

или, перегруппировав слагаемые с учетом того, что $|N_i| = 2^n$ получим

$$\left| \frac{1-b}{2^n} \sum_{j \in N_i} (f_i^{(t)} - f_j^{(t)}) \right| < \varepsilon, \quad (4)$$

и далее

$$\left| \sum_{k=1}^n (f_{P_{ik}}^{(t)} - 2f_i^{(t)} + f_{S_{ik}}^{(t)}) \right| < \varepsilon. \quad (5)$$

Как легко видеть, под знаком суммы стоит приближенное выражение для второй производной функции по k -й переменной в узле i . Произведя тривиальные преобразования, получаем соотношение

$$\sum_i \left(\sum_{k=1}^n (f_{P_{ik}}^{(t)} - 2f_i^{(t)} + f_{S_{ik}}^{(t)}) \right)^2 \approx 0, \quad (6)$$

являющееся дискретной формой выражения

$$\int_{\Omega} \Delta u(x)^2 dx \approx 0. \quad (7)$$

Таким образом, стабильное состояние нейронной сети будет удовлетворять уравнению Лапласа. Рассмотрим, как можно перейти от уравнения Лапласа, не включающего в себя информацию о внутренних источниках тепла, к общему уравнению Пуассона. Используем активационную функцию нейронов общего вида

$$f_i(x) = x + g_i, \quad (8)$$

где $g_i = g(\vec{x}_i)$ – значение функции, стоящей в правой части уравнения Пуассона, в i -м узле сетки. Применяя аналогичные (4) – (6) преобразования, приходим к выражению

$$\left| \sum_{k=1}^n (f_{P_{ik}}^{(t)} - 2f_i^{(t)} + f_{S_{ik}}^{(t)}) + g_i \right| < \varepsilon, \quad (9)$$

или

$$\sum_i \left(\sum_{k=1}^n (f_{P_{ik}}^{(t)} - 2f_i^{(t)} + f_{S_{ik}}^{(t)}) + g_i \right)^2 \approx 0 \quad (10)$$

и

$$\int_{\Omega} (\Delta u(x) + g(x))^2 dx \approx 0. \quad (11)$$

Таким образом, определяя величину единичного смещения, подаваемого на каждый из нейронов сетки, получаем решения произвольного уравнения Пуассона. Рассмотрим, как в выбранной модели могут задаваться краевые условия задачи. Выделим три основных вида краевых условий:

$$u(x) = v(x), x \in \Gamma; \quad (12)$$

$$\left. \frac{\partial u}{\partial \vec{n}} \right|_x = v(x), x \in \Gamma; \quad (13)$$

$$\left. \frac{\partial u}{\partial \vec{n}} \right|_x = \alpha(u(x) - v(x)), x \in \Gamma. \quad (14)$$

Очевидно, что часть нейронов будет соответствовать узлам сетки дискретизации, находящимся на границе Γ области Ω . Для задания краевых условий используем нейроны с измененными активационными функциями

$$f_i(x) = v_i = const, \quad (15)$$

$$f_i(x) = k(x + v_i); \quad k = \frac{2^n}{|N_i|} \text{ – коэффициент, зависящий от формы границы в узле}$$

в узле

$$f_i(x) = x + \alpha(x - v_i). \quad (16)$$

Покажем, что граничные условия в стабильном состоянии нейронной сети будут выполняться для каждого нейрона, соответствующего граничному узлу сетки дискретизации.

Для первого случая доказательство не требуется в силу очевидности.

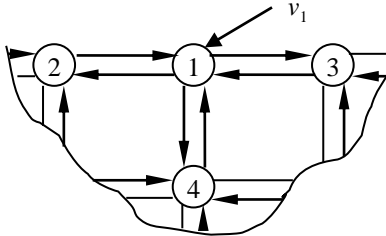


Рис. 1.

Для второго случая сначала рассмотрим двумерный случай,

приведенный на рис. 1. Здесь

$$f_1 = \frac{f_2 + f_3 + f_4 + v}{3}. \quad (17)$$

Тогда, в силу (5)

$$\left. \frac{\partial u}{\partial \bar{n}} \right|_{x_i} \approx (f_1 - f_2) + (f_1 - f_3) + (f_1 - f_4). \quad (18)$$

После тривиальной подстановки (17) получаем

$$\left. \frac{\partial u}{\partial \bar{n}} \right|_{x_i} \approx v_i. \quad (19)$$

После обобщения выражений (17)–(18) на произвольное число измерений n получаем следующее выражение для k -го узла (считаем, что границу в данном узле можно определить $(n-1)$ -мерной гиперплоскостью, заданной уравнением $x_n = c$ и $S_{kn} = \emptyset$):

$$f_k = \frac{1}{2^n - 1} \left(\sum_{i=1}^{n-1} (f_{P_{ki}} + f_{S_{ki}}) + f_{P_{kn}} + v_k \right); \quad (20)$$

$$\left. \frac{\partial u}{\partial \bar{n}} \right|_{x_k} \approx \sum_{i=1}^{n-1} ((f_k - f_{P_{ki}}) + (f_k - f_{S_{ki}})) + (f_k - f_{P_{kn}}) \approx v_k. \quad (21)$$

Покажем, что нейронная сеть выбранной структуры стабильна и имеет единственное стабильное состояние. Нестабильность нейронной сети может быть вызвана одной из двух причин [6, 7]:

1. Существование в сети циклов с положительными обратными связями. Эта проблема характерна для нейронных сетей, использующих неограниченные активационные функции нейронов.

2. «Осцилляция» сети, то есть ее колебание между несколькими состояниями. Эта проблема, как правило, характерна для синхронных нейронных сетей, в которых все нейроны изменяют свое выходное значение одновременно.

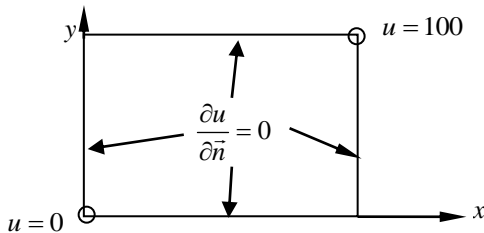


Рис. 2.

В силу условия (2), наложенного на веса связей сети, в сети не может существовать связей с весом, большим единицы. Учитывая кроме этого то, что активационные функции нейронов являются линейными с коэффициентами усиления меньше или равными единице, получаем, что суммарный выходной сигнал нейронов не будет превышать суммарный входной сигнал. Таким образом, возможно гарантировать отсутствие положительных обратных связей в сети.

Пример решения задачи. В качестве иллюстрации выбрана достаточно простая задача, условие которой приведено на рис. 2. Для решения задачи была использована нейронная сеть размером 10×10 указанной структуры. Коэффициент b , задающий теплопроводность материала, был выбран равным 0.1. Стабильным считалось состояние нейронной сети, в котором максимальное по абсолютному значению изменение выходного сигнала нейрона было меньше 10^{-2} . Было произведено 100 запусков со случайными начальными состояниями нейронной сети. Результат сравнивался с решением аналогичной задачи методом конечных элементов с треугольными элементами и сеткой дискретизации 10×10 и 100×100 .

При сравнении с МКЭ получены следующие результаты: для сетки МКЭ 10×10 порядок максимального отклонения – 10^{-2} (0.01 – 0.1% в отдельно взятом узле); для сетки МКЭ 100×100 порядок максимального отклонения – 10^{-1} (0.1 – 1% в отдельно взятом узле).

Выводы. Таким образом, можно получить приближенное решение дифференциального уравнения с помощью нейронной сети указанной структуры. Заметим, что хотя мы говорили о решении уравнения теплопроводности, описанная модель применима для решения произвольных эллиптических дифференциальных уравнений в частных производных.

Список литературы: 1. Адамар Ж. Задача Коши для линейных уравнений с частными производными гиперболического вида. – М.: Наука, 1978. – 351 с. 2. Арфкен Г. Математические методы в физике. – М.: Атомиздат, 1969. – 712 с. 3. Беллман Р., Энджел Э. Динамическое программирование и уравнения в частных производных / Под ред. А.М. Летова. – М.: Мир, 1974. – 207 с. 4. Зоммерфельд А. Дифференциальные уравнения в частных производных физики. – М.: Издательство иностранной литературы, 1950. – 456 с. 5. Несис Е. И. Методы математической физики. – М.: Просвещение, 1977. – 99 с. 6. Горбань А. Н. Обучение нейронных сетей. – М.: СП ПараГраф, 1990. 7. Круглов В., Борисов В. Искусственные нейронные сети. Теория и практика. – М.: Горячая Линия – Телеком, 2002. – 377 с.

Поступила в редакцию 29.09.2004