Науковий вісник НЛТУ України
Scientific Bulletin of UNFU
http://nv.nltu.edu.ua
https://doi.org/10.15421/40270435
Article received 12.05.2017 р.
Article accepted 24.05.2017 р.
УДК 004.02:519.63
ISSN 1994-7836 (print)
ISSN 2519-2477 (online)

@ ✉ Correspondence author
N. O. Semenyshyn
xa4abu@ukr.net

Н. О. Семенишин

Національний лісотехнічний університет України, м. Львів, Україна

РОЗВ'ЯЗУВАННЯ ДВОВИМІРНОЇ ЗАДАЧІ ТЕПЛОПРОВІДНОСТІ РЕКУРЕНТНОЮ НЕЙРОННОЮ МЕРЕЖЕЮ ДЖОРДАНА

Розглянуто метод для розв'язання двовимірного рівняння теплопровідності в ізотропних матеріалах із граничними умовами першого типу, використовуючи штучну нейронну мережу (ШНМ) Джордана. Побудовано функцію вартості ШНМ, в основі якої знаходиться метод Кранка-Ніколсона. При досягненні функцією вартості мінімуму, виходи мережі дають розв'язок рівняння теплопровідності. Ця функція вартості мінімізується методом Флетчера-Рівза за синаптичними вагами. Щоб знайти частинні похідні першого порядку від функції вартості за ваговими коефіцієнтами мережі, використано розширення стандартного алгоритму зворотного поширення, названого "Зворотним поширенням у часі за епохами".

Підібрано архітектуру мережі з урахуванням специфіки розв'язуваної задачі. Оптимальні можливості апроксимації отримано з використанням двох шарів мережі без використання функцій активації через лінійність рівняння. Наведено результати моделювання на двох тестових задачах і здійснено порівняння результату з іншими числовими методами. Показано, що результати розрахунків з використанням цього підходу дають добре наближення до точних рішень. Також отримано задовільний розв'язок за межами часового діапазону, для якого відбувалось навчання. Показано стійкість та збіжність цього підходу при значеннях кроку за часом, для якого явні різницеві методи є чисельно нестабільними.

Представлено методологію розв'язання крайових задач, а саме рівняння теплопровідності у двох вимірах з використанням штучної нейронної мережі. Гнучкість підходу полягає в можливості донавчання мережі для досягнення необхідної точності та здатності мережі "запам'ятовувати" рішення для різних початкових умов. Універсальність підходу полягає у використанні єдиної методології для різних фізичних параметрів задачі.

Ключові слова: рівняння теплопровідності; рекурентна нейронна мережа Джордана; метод Кранка-Ніколсона; метод Флетчера-Рівза; метод "Зворотного поширення в часі"; функція вартості (енергетична функція).

Вступ. Чисельним методам у багатьох дисциплінах, таких як фізика, прикладна математика, електротехніка, біохімія і т.д., останнім часом надають велику увагу як практичним методам для розуміння складних явищ, які майже неможливо розглядати в аналітичній формі. Багато методів розроблено до цього часу для вирішення диференціальних рівнянь у частинних похідних. Деякі з них дають розв'язок у вигляді масиву, який містить значення у деякій групі точок. Інші використовують базисні функції для представлення розв'язку в аналітичному вигляді і зведення початкової задачі до системи лінійних рівнянь (Zenkevich & Morgan, 1986; Farlou, 1985).

Існує також чимало методів з використанням ШНМ для цієї задачі. Тут можна виділити кілька груп методів, зокрема: методи побудовані на основі методу зважених нев'язок, без необхідності навчання з попереднім обчисленням коефіцієнтів B-сплайнів, якими апроксимується розв'язок (Meade & Fernandez, 1994), та ШНМ прямого поширення з навчанням і подальшою заміною ними базисних функцій для апроксимації розв'язку (Aarts & van der Veer, 2001; Lagaris, Likas &Papageorgiou, 2000; Lagaris, Likas & Fotiadis, 1998; Mall & Chakraverty, 2013). Великий клас РБФ мереж з можливістю підбору кількості, положення та розмірів радіально-базисних функцій (Vasilev & Tarhov, 2009). Методи, побудовані на основі теорії стійкості (метод Ляпунова), в яких зміна стану нейрона мінімізує енергетичну функцію, сформовану за скінченно-різницевим шаблоном (Lee & Kang, 1990). І подібні до останнього високо паралельні методи на клітинних нейронних мережах, в яких значення нейрона наближається до розв'язку рівняння за певною ітераційною схемою (Gorbachenko, 2003; Novotarskyi & Nesterenko, 2004).

Попри велику кількість літератури із ШНМ, все ж розв'язування рівнянь параболічного типу на нейронних мережах розкрите недостатньо. У цій роботі здійснено спробу створити гнучку та універсальну методологію для розв'язування двовимірних рівнянь теплопровідності. Гнучкість підходу полягає у можливості донавчати мережу для досягнення необхідної точності та здатності мережі "запам'ятовувати" розв'язки для різних початкових умов. Універсальність методу полягає у застосуванні єдиного підходу для розв'язання цієї задачі для різних фізичних параметрів задачі.

Опис двовимірної задачі теплопровідності. Розглянемо двовимірну задачу теплопровідності для ізотропних матеріалів із граничними умовами першого типу у тілі квадратної форми

Цитування за ДСТУ: Семенишин Н. О. Розв'язування двовимірної задачі теплопровідності рекурентною нейронною мережею Джордана. Науковий вісник НЛТУ України. 2017. Вип. 27(4). С. 166–169.
Citation APA: Semenyshyn, N. O. (2017). Solving Heat Equation in two Dimensions Using Jordan Recurrent Neural Network. Scientific Bulletin of UNFU, 27(4), 166–169. https://doi.org/10.15421/40270435

$$\frac{\partial u}{\partial t} = \alpha\left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}\right), 0 < x < L,\ 0 < y < H,\ 0 < t < T$$

$$\begin{cases} u(t,0,y) = C_1 \\ u(t,L,y) = C_2 \\ u(t,x,0) = C_3 \\ u(t,x,H) = C_4 \end{cases}, 0 < t < T \qquad (1)$$

$$u(0,x,y) = C_0(x,y),\ 0 < x < L,\ 0 < y < H,$$

де: u – температура тіла в точці x, y у певний момент часу t (градусів С); x і t – незалежні змінні; α – коефіцієнт температуропровідності дорівнює $\alpha = k/c\rho$ (м²/с); звідки c – теплоємність матеріалу стрижня (Дж/кг·°С); r – густина матеріалу тіла (кг/м³) і k – коефіцієнт теплопровідності (Вт/(м·К)); $C_0(x,y)$ – початковий розподіл температури у квадратному тілі; C_1, C_2, C_3, C_4 температура на лівій, правій, нижній та верхній границях тіла відповідно.

Рівномірно розбиваємо досліджувану прямокутну ділянку по t, x, y $\{(t_i, x_j, y_k)\}_{i=0,\,j=0,k=0}^{M_t M_x M_y}$, де $t_i = i\Delta t$; $x_j = j\Delta x$; $y_k = k\Delta y$; $\Delta t = T / M_t$; $\Delta x = L / M_x$; $\Delta y = H / M_y$, $\Delta t, \Delta x, \Delta y$ – кроки розбиття по відповідних осях, M_t, M_x, M_y – кількість кроків часу та кількість точок розбиття по x та y відповідно.

Побудова та навчання мережі. ШНМ для одновимірної задачі теплопровідності створено у попередній роботі (Zenkevich & Morgan, 1986; Semenyshyn, 2016), у якій роботі був істотний недолік – надто велика кількість синаптичних ваг (кожен нейрон наступного шару містив усі нейрони попереднього), що робило процес навчання дуже тривалим. При переході до двовимірної задачі за такого підходу обчислювальна складність дорівнювала б $O(2M_x^2 M_y^2)$, де M_x – кількість нейронів, що відповідають осі x, а M_y – кількість нейронів, що відповідають осі y, множник 2 означає два шари мережі. Тому у цій роботі вирішили взяти кількість ваг для кожного нейрона прихованого та вихідного шарів – 5, що відповідає кількості точок, які входять до скінченно-різницевого шаблону неявного методу Кранка-Ніколсона.

Як і в роботі (Semenyshyn, 2016), кожній дискретній точці досліджуваної області ставимо у відповідність нейрон вихідного шару, і так само кожному часовому шару (кроку) нашої рекурентної мережі ставимо у відповідність часовий крок задачі, що моделюється. Значення нейрона вихідного шару з індексом j, k для i-го кроку часу рекурентної мережі дорівнює $u_{ijk} = u(t_i, x_j, y_k)$, розраховуємо так:

$$u_{ijk} = \sum_{l=1}^{M_x}\sum_{r=1}^{M_y} a_{jklr} \sum_{c=1}^{M_x}\sum_{p=1}^{M_y} w_{lrcp} u_{i-1cp},\ 0 < i < M_t. \qquad (2)$$

Як бачимо, ваги є чотиривимірними масивами, що при обраному різницевому шаблоні зводиться до п'яти ваг для кожного нейрона індексом j, k у прихованому та вихідному шарах. За такого підходу потрібно використовувати розріджені масиви, щоб економити пам'ять. Для цієї мережі відсутня функція активації нейрона, оскільки ця крайова задача лінійна. Розмір прихованого шару дорівнює розміру видного шару.

Щоб почати навчання, потрібно мати енергетичну функцію, яку будемо оптимізувати. Для цього використаємо метод Кранка-Ніколсона для двовимірної задачі теплопровідності (Vasilev & Tarhov, 2009; Farlou, 1985)

$$J(u) = \sum_{i=1}^{M_t}\sum_{j=1}^{M_x}\sum_{k=1}^{M_y}\left(p_1(u_{ijk}) - \alpha \cdot p_2(u_{ijk}) - \alpha \cdot p_3(u_{ijk})\right)^2. \qquad (3)$$

де: $$p_1(u) = \frac{u_{i,j,k} - u_{i-1,j,k}}{\Delta t};$$

$$p_2(u) = \frac{u_{i,j-1,k} - 2u_{i,j,k} + u_{i,j+1,k} + u_{i-1,j-1,k} - 2u_{i-1,j,k} + u_{i-1,j+1,k}}{2\Delta x^2};$$

$$p_3(u) = \frac{u_{i,j,k-1} - 2u_{i,j,k} + u_{i,j,k+1} + u_{i-1,j,k-1} - 2u_{i-1,j,k} + u_{i-1,j,k+1}}{2\Delta y^2}.$$

Як видно із (3), чим ближчий $J(u)$ до нуля, тим більший вихід мережі задовольняє рівняння (1). Тому задача навчання нашої ШНМ зводиться до задачі оптимізації (3) за синаптичними вагами. Для цього використовуємо метод Флетчера-Рівза (Hameed, 2016)

$$w_{k+1}(l,c) = w_k(l,c) + \eta p_k(l,c),\ p_k(l,c) = g_k(l,c) + \beta p_{k-1}(l,c),$$
$$p_0(l,c) = g_0(l,c),\ \beta_k(l,c) = \frac{(g_k(l,c), g_k(l,c))}{(g_{k-1}(l,c), g_{k-1}(l,c))}, \qquad (4)$$

де $g_k = -\frac{\partial J}{\partial w}$, k – номер глобальної ітерації навчання нейронної мережі.

Оскільки g_k є чотиривимірною матрицею, тому знаходити β_k можемо тільки для матриці похідних від виразу (3) за синаптичними вагами, що входять до нейрону з індексом (l,c). А вираз $(g_k(l,c), g_k(l,c))$ – означає норму матриці $g_k(l,c)$, з фіксованими першими двома індексами l і c відповідно. Аналогічні обчислення виконують і для матриці a. Для пошуку оптимального параметра η можна використати метод Хорд. Умовою зупинки процесу оптимізації може бути: мале значення (3), мала зміна вихідних значень мережі, мале значення норми матриць (8) або комбінація з перелічених вище.

Для того, щоб обчислити частинні похідні функції вартості за синаптичними вагами мережі, використовуємо алгоритм Зворотного поширення похибки в часі за епохами (Epochwise Back-propagation through Time) (Khajkin, 2006; Werbos, 1990). Цього можна отримати шляхом розгортання часових операцій мережі в багатошарову мережу прямого поширення, топологія якої розширюється на один шар для кожного кроку часу, тобто для кожного часового шару визначено одні і ті ж вагові коефіцієнти.

У цьому алгоритмі спочатку виконується пряма передача даних мережею для M_t часових кроків, при цьому стан мережі запам'ятовується. Далі відбувається зворотне проходження даних, починаючи з кроку M_t за рекурентними формулами:

$$J\_\hat{u}_{ijk} = \frac{\partial J}{\partial u_{ijk}}; \qquad (5)$$

$$J\_u_{ijk} = J\_\hat{u}_{ijk} + \sum_{l=1}^{M_x}\sum_{p=1}^{M_y} w_{lpjk} J\_x_{i+1lp},\ M_x > j > 1,\ M_y > k > 1; \qquad (6)$$

$$J\_x_{ilp} = \sum_{j=1}^{M_x}\sum_{k=1}^{M_y} a_{jklp} J\_u_{ijk},\ M_x > l > 1,\ M_y > p > 1. \qquad (7)$$

Формули (6) і (7) обчислюють похідну згідно з ланцюговим правилом диференціювання від (3), щоб визначити вплив кожного нейрона на формування вихідного сигналу, префікс J_ у формулах означає похідну від функції (3) по змінній, яка йде після цього префікса. І, насамкінець, обчислюємо похідні від (3) по вагах

$$J\_a_{cjkl} = \sum_{i=1}^{M_t} J\_u_{icj} x_{ikl},\quad J\_w_{cjkl} = \sum_{i=1}^{M_t - 1} J\_x_{i+1cj} u_{ikl}, \qquad (8)$$

де x_{ikl} – це значення нейрона з індексом k, l вихідного шару на кроці часу i.

Результати моделювання. Розглянемо дві задачі теплопровідності (1) на проміжку $0<=x<=1$, $0<=y<=1$ та $0<=t<=0{,}05$ з коефіцієнтом температуропровідності (α) = 1/8 (м²/c) – для першої задачі та (α) = 1/2 (м²/с) – для другої. Кроки $\Delta x, \Delta y$ дорівнюють 0,05, а $\Delta t = 0{,}005$. Початкові умови першої задачі $u(0, x, y)=0{,}5$, $0{,}5<y<1$, $0<x<1$. Розв'язок мережі на 11-му кроці часу (t=0,05) зображено на рис. 1,б. Величину похибки між точним і нейромережевим розв'язками визначали за формулою

$$\sigma = \sum_{j=1}^{M_x} \sum_{k=1}^{M_y} (u(T, x_j, y_k) - \bar{u}(T, x_j, y_k))^2 \,, \qquad (9)$$

де $\bar{u}$ – точний розв'язок задачі (1). У першому прикладі σ=0,0957, а значення σ між точним розв'язком і розв'язком, отриманим неявною схемою (зразок робочої програми взято з ресурсу[1]) МСР дорівнює 0,1126. При спробі екстраполювати розв'язок для часу 2T значення σ збільшилась до 0,3418.

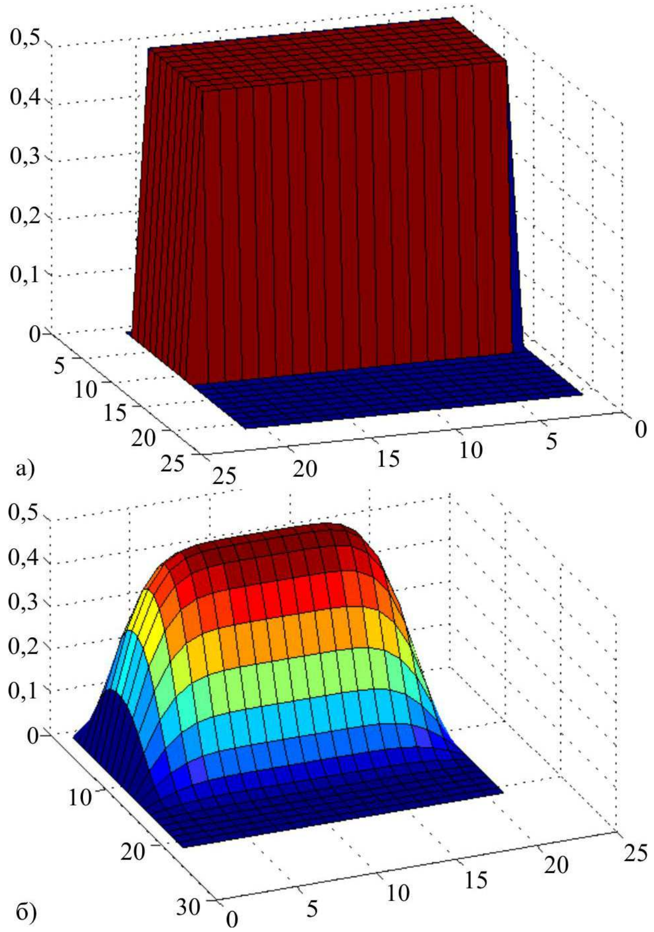

Рис. 1: а) початкові умови задачі 1; б) апроксимація задачі 1

Початкові умови другої задачі $u(0, x, y)=x(1-x)\, y(1-y)$, $0<y<1$, $0<x<1$. Результат мережі для цієї задачі за t=0,05 зображено на рис. 2,б. За явної схеми МСР розв'язок задачі нестійкий. Значення σ між точним і нейромережевим розв'язками дорівнює $6{,}2 \cdot 10^{-4}$ порівняно з неявною схемою $4{,}9 \cdot 10^{-5}$.

Навчена мережа одночасно для цих двох задач показала результат, який приблизно на два порядки гірший від точного розв'язку.

[1] https://www.mathworks.com/matlabcentral/fileexchange/38088-diffusion-in-1d-and2d? focused=5246996&tab=function

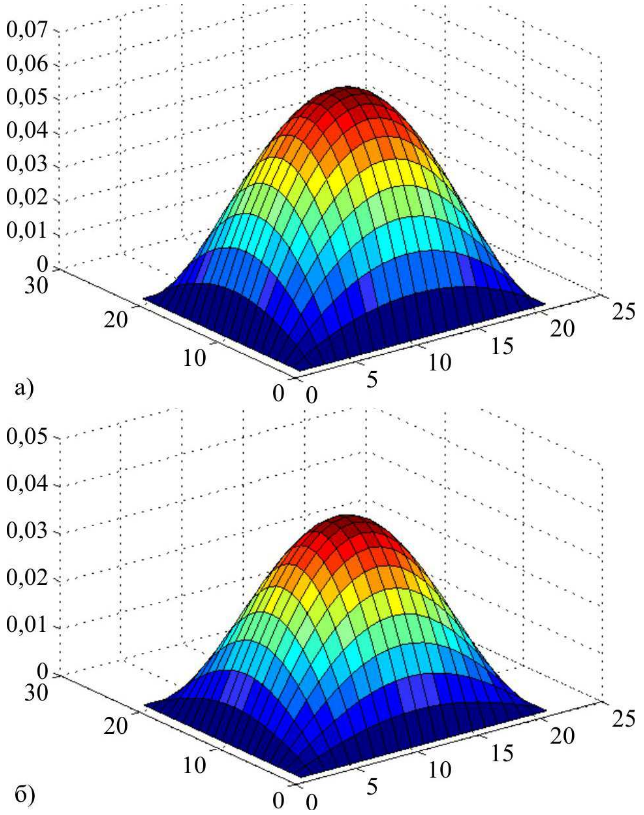

Рис. 2: а) початкові умови задачі 2; б) апроксимація задачі 2

Висновки. Розроблено методологію розв'язування двовимірного рівняння теплопровідності, використовуючи штучну нейронну мережу Джордана. Суть методології зводиться до корекції синаптичних ваг таким чином, щоб задовольнити енергетичну функцію, а також, щоб навчити мережу здійснювати відображення цього диференціального оператора, що дає змогу екстраполювати розв'язок мережею за часовою координатою та дає можливість навчання для різних наборів початкових і граничних умов. Перевагою підходу також є стійкість та менша чутливість до великих кроків за часом у сенсі критерію Куранта-Фрідріхса-Леві, на відміну від явних різницевих методів. Але недоліком можна вважати те, що процес навчання мережі для отримання точного розв'язку рівняння є дуже чутливим до початкових значень синаптичних ваг, про що свідчать різні результати моделювання. Також одночасно навчання мережі кількох розв'язків рівняння вимагає додаткового дослідження та розроблення точніших методів навчання.

Перелік використаних джерел

Aarts, L. P, & van der Veer, P. (2001). Solving nonlinear differential equations by a neural network method. *Proceedings computational science conference, 3*, 181–189. San Fransisco USA, May 2001. ed. / VN Alexandrov; JJ Dongarra.Berlin: Springer.

Farlou, S. (1985). Uravnenija s chastnymi proizvodnymi dlja nauchnyh rabotnikov i inzhenerov: per. s angl. Moscow: Mir, 384 p. [in Russian].

Gorbachenko, V. I. (2003). *Nejrokompjutery v reshenii kraevyh zadach teorii polja*. Vol. 10 Moscow: Izd-vo "Radiotehnika", 336 p. [in Russian].

Hameed, Dr. W. Abdul. (2016). *Fletcher-Reeves conjugate gradient neural network to solve systems of linear equations*. Department of Mathematics, School of Advanced Sciences, VIT University, Tamilnadu, India, 260 p.

Khajkin, S. (2006). *Nejronnye seti: polnyj kurs*: per. s angl. Izd. 2-oe, [pererab. i dop.]. Moscow: Izd. dom "Viljamsa", 1104 p. [in Russian].

Lagaris, I. E., Likas, A. C., & Fotiadis, D. I. (1998). Artificial neural networks for solving ordinary and partial differential equations. *IEEE Transactions on Neural Networks, 9*(5), 987–1000.

Lagaris, I. E., Likas, A. C., &Papageorgiou, D. G. (2000). Neural-Network Methods for Boundary Value Problems with Irregular Bounadaries. *IEEE Trans. Neural Networks, 11*(5), 1041–1049.

Lee, H., & Kang, I. S. (1990). Neural algorithm for solving differential equations. *Journal of Computational Physics, 91*(1), 110–131.

Mall, S., & Chakraverty, S. (2013). Comparison of Artificial Neural Network Architecture in Solving Ordinary Differential Equations. *Advances in Artificial Neural Systems, 3*, 1–12.

Meade, A. J. Jr., & Fernandez, A. A. (1994). Solution of nonlinear ordinary differential equations by feedforward neural networks. *Mathematical and Computer Modelling, 20*(9), 19–44.

Novotarskyi, M. A., & Nesterenko, B. B. (2004). Shtuchni neironni merezhi: obchyslennia. *Pratsi Instytutu matematyky HAH Ukrainy* (Vol. 50, pp. 137–142). Kyiv: In-t matematyky HAH Ukrainy, 408 p. [in Ukrainian].

Semenyshyn, N. O. (2016). Solving Heat Equation In One Dimension Using Jordan Recurrent Neural Network. *Scientific Bulletin of UNFU, 26*(7), 405–411. Retrieved from: http://nv.nltu.edu.ua/index.php/journal/article/view/575

Vasilev, A. N., & Tarhov, D. A. (2009). *Nejrosetevoe modelirovanie. Principy. Algoritmy. Prilozhenija*. Sankt_peterburg: Gosudarstvennyj Politehnicheskij Universitet, 360 p. [in Russian].

Werbos, P. J. (1990). Backpropagation through time: what it does and how to do it. *Proceedings of the IEEE, 78*(10), October, 1550–1560.

Zenkevich, O., & Morgan, K. (1986). *Konechnye jelementy i approksimacija*: per. s angl. Moscow: Mir, 318 p. [in Russian].

Н. О. Семенишин

Национальный лесотехнический университет Украины, г. Львов, Украина

РЕШЕНИЕ ДВУМЕРНОГО УРАВНЕНИЯ ТЕПЛОПРОВОДНОСТИ С РЕКУРРЕНТНОЙ НЕЙРОННОЙ СЕТЬЮ ДЖОРДАНА

Представлена методология решения краевых задач, а именно уравнений теплопроводности в двух измерениях с использованием искусственной нейронной сети. Сделана попытка создать гибкую и универсальную методологию для решения дифференциальных уравнений. Гибкость подхода заключается в возможности дообучения сети для достижения требуемой точности и способности сети "запоминать" решения для различных начальных условий. Универсальность подхода заключается в использовании единой методологии для различных физических параметров задачи.

Функция стоимости сети построена на 10-точечной схеме Кранка-Николсона, применяемой для всех точек в моделируемой пространственно-временной области. Как только функция стоимости достигает минимума, выходы нашей сети дают решение уравнения теплопроводности. Для оптимизации функции стоимости использован метод Флетчера-Ривза для поиска оптимальных синаптических весов сети. Чтобы найти частные производные первого порядка от функции стоимости по весовым коэффициентам сети, использовано расширение стандартного алгоритма обратного распространения, называемого "Обратным распространением во времени". Оптимальные возможности аппроксимации получены с использованием двух слоев сети без использования функции активации из-за линейности уравнения. Приведены результаты моделирования. Показано, что результаты расчетов с использованием этого подхода дают хорошее приближение к точным решениям. Результат сети является численно устойчивым для достаточно больших временных шагов (но удовлетворяющих условию Куранта-Фридрихса-Леви) по сравнению с явным методом конечных разностей. Экстраполировано решение уравнения теплопроводности для временного промежутка, вдвое большего, чем при обучении сети.

Ключевые слова: уравнение теплопроводности; рекуррентная нейронная сеть Джордана; метод Кранка-Николсона; метод Флетчера-Ривза; метод "обратного распространения-по-времени"; функция стоимости.

N. O. Semenyshyn

Ukrainian National Forestry University, Lviv, Ukraine

SOLVING HEAT EQUATION IN TWO DIMENSIONS USING JORDAN RECURRENT NEURAL NETWORK

The paper presents a methodology for solving boundary value problems, namely heat equation in two dimensions using Jordan artificial neural networks. This paper is also an attempt to create a flexible and universal methodology for solving differential equations. Flexibility of approach is the ability to retrain network to achieve the required accuracy and the ability to network "remember" solutions for different initial conditions. Universality of approach is to use a unified methodology for various physical parameters of the problem. The results of the study are as follows. Cost function of the network was based on the 10 points Crank-Nicolson numerical scheme applied for all points in spatio-temporal domain which are of interest. As soon as cost function approaches a minimum outputs of our network will give the solution of heat equation. To optimize cost function we use Fletcher-Reeves method with respect to synaptic weights of the network. To find first-order partial derivatives of a cost function with respect to networks weights we use an extension of the standart backpropagation algorithm referred to as "Back-propagation – through-time". Optimal approximation capabilities was obtained using two layers of the network without using activation function because of linear nature of equation. Initial weights of network were small enough and distributed uniformly. The results of simulation are presented. It is shown, that the results of calculation using this approach give good approximations to exact solutions. Result of the network is numerically stable for big enough time steps (but satisfying Courant – Friedrichs – Lewy condition) comparing with explicit finite difference method. We also were able to extrapolate solution of heat equation at time interval which was twice bigger than time interval used for training.

Keywords: heat equation; Jordan recurrent neural network; Crank-Nicolson method; Fletcher-Reeves method; "back-propagation – through-time" method; cost function.

Інформація про автора:

Семенишин Назар Олегович, аспірант, Національний лісотехнічний університет України, м. Львів, Україна.
Email: xa4abu@ukr.net