УДК 681.142.37

# НЕЙРОМЕРЕЖІ ПРЯМОГО ПОШИРЕННЯ. ПРОБЛЕМИ СИНТЕЗУ ТА ВИКОРИСТАННЯ

**© Р. Ткаченко, Р. Когут**
Національний університет "Львівська політехніка"

***Описано проблеми, що виникають при синтезі та застосуванні нейронних мереж прямого поширення. Наведено результати порівняльного аналізу нейронних мереж з ітераційним алгоритмом навчання зворотного поширення похибки та нейронної мережі, побудованої на основі парадигми "Функціонал на множині табличних функцій".***

***In this paper the problems of synthesis and using feed forward neural networks are described. We present the results of analysis of using neural network with iterational learning (back propagation) and neural network in base of paradigm "functional on the tabular functions set".***

Останнім часом із нейромережами пов'язують реальні можливості успішного розв'язання складних завдань нової галузі інформатики, відомої під назвою "Видобуток даних - *Data mining*". Зазначимо, що вже у 1943 році Уорреном Маккалоком (*Warren McCullock*) та Уолтером Пітсом (*Walter Pitts*) здійснені перші спроби промоделювати функціонування біологічних нейронів та застосувати технічні моделі для задач опрацювання інформації [1]. За тривалий період інтенсивних досліджень прихильниками зазначеного напрямку розроблено вражаючу кількість нейроархітектур і нейропарадигм різноманітного призначення. Задля справедливості зазначимо, що серед цих розробок лише деякі отримали застосування на практиці. До базових в першу чергу належать штучні нейронні мережі зі зворотними зв'язками (*Back Forward*) та прямого поширення (*Feed Forward*). Останні, як більш універсальні та більш точні, переважно використовуються для розв'язку задач відображення вхідних наборів даних у вихідні.

Як відомо, першим з варіантів нейромереж прямого поширення став перцептрон Ф. Розенблата, який започаткував практичний напрямок у побудові систем штучного інтелекту. Історія розвитку перцептрона – від надмірного захоплення і перебільшення його можливостей і аж до повного заперечення – стала типовою для нейрокомп'ютерної техніки загалом. Спробуємо показати, що очевидні спади в розвитку теорії НМ безпосередньо пов'язані з абсолютизацією окремих моментів моделей НМ, запропонованих їх розробниками.

Проблема універсальності застосування НМ докладно описана у працях Мінскі, Пейперта [2], що довели неможливість відтворення за допомогою перцептрона деяких залежностей, зокрема логічної функції "сума за модулем два". Беззаперечні висновки авторитетних вчених-математиків надовго відвернули від проблеми увагу більшості дослідників. Однак наведені твердження справедливі за умови наявності в НМ лише одного прихованого шару синаптичних ваг, що налагоджуються в процесі навчання. На вказане обмеження перцептрона, що зовсім не є обов'язковим, звернули увагу лише через довгі роки.

Проблема навчання. В основі теорії навчання НМ лежить так зване правило Д. Гебба, запропоноване ним ще в 1949р.

Правило Гебба. При збудженні одночасно двох нейронів з виходами $(x_j, y_i)$ на $k$-тому кроці навчання вага синаптичного з'єднання між ними зростає, в інакшому випадку – зменшується, тобто

$$\Delta W_{ij}(k) = \eta\ x_j(k)\ y_i(k),$$

де $\eta$ – коефіцієнт швидкості навчання. Приймається також до уваги, що НМ будується на бінарних нейронах, що мають два стани – збудження та гальмування. Особливість передатних функцій подібних нейронів – відсутність похідних в точках переходів. Це стало причиною невдалих спроб створити надійні алгоритми навчання подібних НМ, що тривали протягом десятків років. І лише Вербосом, а пізніше Дж. Хінтоном запропонований варіант НМ, де використовуються нейрони з неперервними передатними функціями (наприклад, функція Фермі). Це, в свою чергу, дозволяє під час навчання керуватися поширенням сигналів помилки від виходів НМ до її входів у напрямку, зворотному прямому поширенню сигналів у звичайному режимі роботи. Даний алгоритм навчання ШНМ одержав назву методу зворотного поширення (*Back Propagation*), численні модифікації якого застосовуються до сьогодні.

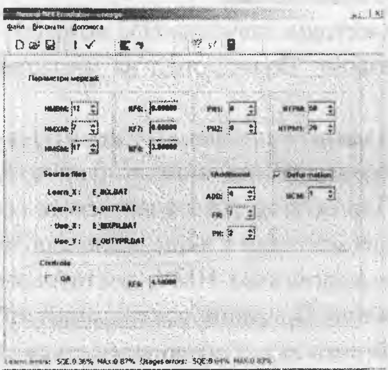


Рис 1. Діалогове вікно програмної НМ ФМТФ

По суті, описаний підхід зводить навчання НМ до вирішення задачі багатокритеріальної оптимізації. Отже, недоліки сучасних методів навчання НМ прямого поширення безпосередньо випливають з реальних проблем реалізації оптимізаційних методів. Наведемо лише деякі з них [3]:

1. Так званий *параліч мережі*. У процесі навчання мережі значення ваг можуть стати дуже великими, що може привести до того, що всі чи більшість нейронів будуть функціонувати при дуже великих значеннях виходів в області, де похідна передатної функції дуже мала. Оскільки помилка, що посилається назад у процесі навчання, пропорційна до цієї похідної, то процес навчання може практично зупинитись. Звичайно цього уникають зменшенням коефіцієнта швидкості навчання $\eta$, але це збільшує його тривалість.

2. *Локальні мінімуми*. Мережа може потрапити в локальний мінімум, коли поруч є набагато більший мінімум. У точці локального мінімуму всі напрямки ведуть нагору, тому мережа не здатна з нього вибратися.

3. *Часова нестійкість*. Процес навчання має бути таким, щоби мережа навчалася на всій навчальній вибірці. Необхідні зміни ваг повинні обчислюватися на всій множині реалізацій, а це вимагає додаткової пам'яті; після ряду таких навчальних циклів ваги зійдуться до мінімальної помилки. Цей метод може виявитися марним, якщо мережа знаходиться в постійно змінному зовнішньому середовищі. У цьому випадку процес навчання може ніколи не зійтися.

У результаті виявляється, що основна проблема НМ прямого поширення – тривалий час навчання, що нелінійно залежить від розміру мережі та об'єму навчальної вибірки. Як наслідок цього, реально вдається розв'язати завдання невеликої або середньої вимірності, що лише наближено можна віднести до проблем видобутку даних.

Задачу зменшення часу навчання НМ прямого поширення намагаються розв'язати шляхом використання неоднорідних НМ типу зустрічного поширення (*Counter propagation*), або НМ повнозв'язного типу [3]. В обох випадках прискорення навчання досягається за рахунок падіння точності функціонування, властивого названим НМ. Отже, по суті, сформувалася ще одна думка, що набула ваги аксіоми, згідно з якою скорочення часу на навчання НМ можливе лише за рахунок збільшення похибок функціонування. Очевидно, дана теза справедлива лише за умови використання традиційних підходів до навчання НМ, що, на наш погляд, є далеко не єдино можливими.

Як альтернатива до оптимізаційних, ітераційних методів навчання НМ може розглядатися просторово-подібна модель, відома під назвою "Функціонал на множині табличних функцій" (ФМТФ) [4].

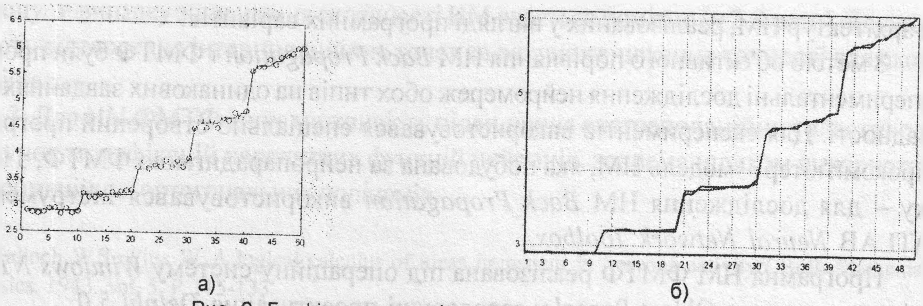
В основу базової моделі покладено концепцію, що множина точок-реалізацій об'єкта формує в n-вимірному евклідовому просторі над полем дійсних чисел тіло об'єкта, що може бути промодельоване за допомогою НМ. Тіло об'єкта розглядається як геометричне місце всіх його точок-реалізацій, де навчальна множина є лише відповідною підмножиною. Тіло об'єкта може бути відтворене за допомогою НМ з достатньою точністю після її навчання на основі навчальної підмножини. Принциповою відмінністю для НМ ФМТФ є алгоритми навчання, що є неітераційними за суттю (виконуються за визначене число кроків), незалежними від початково заданих випадкових наближень, як це існує для відомих методів. Останнє означає, що принципові обмеження на вимірність вирішуваних завдань практично знімаються і можуть залежати лише від наявного об'єму оперативної пам'яті комп'ютера. На базі моделі ФМТФ розроблено

*Таблиця 1*

| | № виходу | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|---|
| Back propagation | Середньо-квадратична зведена похибка,% | 0.0608 | 0.0809 | 0.0411 | 0.0039 | 0.3955 | 0.0127 | 0.0287 |
| | Максима-льна зведена похибка,% | 0.1340 | 0.1995 | 0.0848 | 0.0010 | 0.9181 | 0.0302 | 0.0596 |
| ФМТФ | Середньо-квадратична зведена похибка,% | 0.0574 | 0.0729 | 0.0367 | 0.0005 | 0.3662 | 0.0120 | 0.0232 |
| | Максимальна зведена похибка,% | 0.1242 | 0.1675 | 0.0693 | 0.0014 | 0.8780 | 0.0338 | 0.0538 |

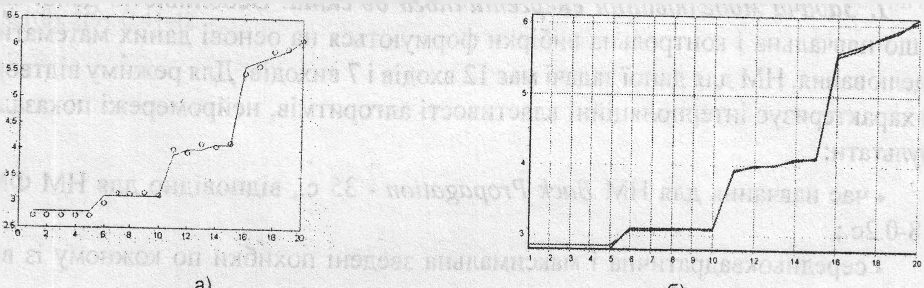
*Таблиця 2*

| | № виходу | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|---|
| Back propagation | Середньо-квадратична приведена похибка,% | 0.0813 | 0.1209 | 0.1360 | 0.0279 | 1.1465 | 0.0287 | 0.2724 |
| | Максима-льна приведена похибка,% | 0.1655 | 0.1807 | 0.2125 | 0.0474 | 2.001 | 0.0498 | 0.5342 |
| ФМТФ | Середньо-квадратична приведена похибка,% | 0.0914 | 0.1314 | 0.1545 | 0.0272 | 0.6422 | 0.0227 | 0.0318 |
| | Максимальна приведена похибка,% | 0.0813 | 0.1209 | 0.1360 | 0.0279 | 1.1465 | 0.0287 | 0.2724 |



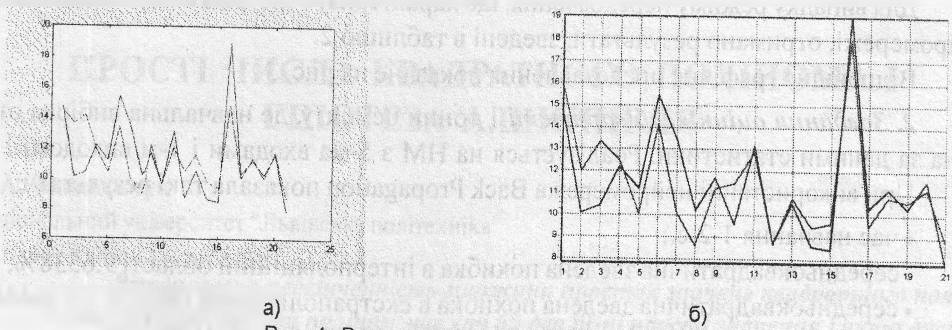
а) б)

Рис. 2. Графіки режиму відтворення, отримані для:
а) MATLAB Neural Network Toolbox; б) програмної НМ ФМТФ



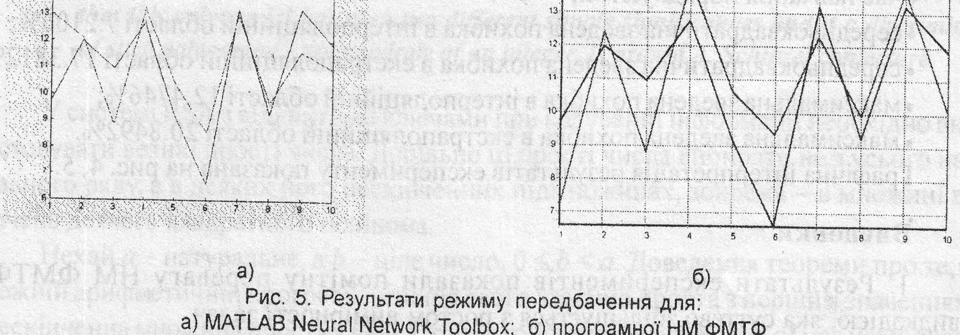
а) б)

Рис. 3. Графіки режиму передбачення, отримані для:
а) MATLAB Neural Network Toolbox; б) програмної НМ ФМТФ



а) б)

Рис. 4. Результати режиму відтворення для:
а) MATLAB Neural Network Toolbox; б) програмної НМ ФМТФ



а) б)

Рис. 5. Результати режиму передбачення для:
а) MATLAB Neural Network Toolbox; б) програмної НМ ФМТФ

ряд архітектур НМ, реалізованих у вигляді програмних варіантів.

З метою об'єктивного порівняння НМ *Back Propagation* і ФМТФ були проведені експериментальні дослідження нейромереж обох типів на однакових завданнях різної складності. Для експериментів використовувався спеціально створений програмний нейрокомп'ютер – модель НМ, яка побудована за нейропарадигмою ФМТФ, з іншого боку – для дослідження НМ *Back Propagation* використовувався інструментарій MATLAB *Neural Network Toolbox*.

Програмна НМ ФМТФ реалізована під операційну систему *Windows NT5.0* на мові пограмування *Object Pascal* у середовищі проектування *Delphi 5.0*.

Проаналізуємо деякі результати експериментальних досліджень.

***1. Задача моделювання енергетичного об'єкта.*** Особливістю даної задачі є те, що навчальна і контрольна вибірки формуються на основі даних математичного моделювання. НМ для даної задачі має 12 входів і 7 виходів. Для режиму відтворення, що характеризує інтерполяційні властивості алгоритмів, нейромережі показали такі результати:

• час навчання для НМ *Back Propagation* - 35 с., відповідно для НМ ФМТФ - 0.18-0.2с.;

• середньоквадратична і максимальна зведені похибки по кожному із виходів представлені в таблиці 1.

На рис. 2 наведено графіки порівняння величин, отриманих для 5-го виходу нейромереж з контрольними даними.

Для випадку режиму передбачення, що характеризує екстраполяційні можливості нейромережі, отримано результати, зведені в таблицю 2.

Відповідне графічне представлення показане на рис. 3.

***2. Завдання оцінки собівартості*** 1 тонни цементу, де навчальна вибірка отримана за даними статистики. Реалізується на НМ з 5-ма входами і 1-м виходом.

При використанні нейромережа Back Propagation показала такі результати:

• час навчання 1-2 с.;
• середньоквадратична зведена похибка в інтерполяційній області 9.6538%;
• середньоквадратична зведена похибка в екстраполяційній області 12.1064%;
• максимальна зведена похибка в інтерполяційній області 21.0526%;
• максимальна зведена похибка в екстраполяційній області 23.0769%.

Відповідно для НМ ФМТФ:

• час навчання порядку 0.1с.;
• середньоквадратична зведена похибка в інтерполяційній області 7.2106%;
• середньоквадратична зведена похибка в екстраполяційній області 17.3812%;
• максимальна зведена похибка в інтерполяційній області 12.4746%;
• максимальна зведена похибка в екстраполяційній області 20.8492%.

Графічна інтерпретація результатів експерименту показана на рис. 4, 5.

## Висновки

1. Результати експериментів показали помітну перевагу НМ ФМТФ за швидкодією, яка суттєво збільшується з ростом вимірності задач.

2. НМ ФМТФ забезпечують точність розв'язання задач співмірну з методом BP

або вищу. У випадку збільшення складності НМ вдається помітно (у 2-3 рази) підвищити точність відтворення інтерполяційних точок за рахунок зниження точності для екстраполяційних.

3. Для НМ ФМТФ існує можливість підвищення екстраполяційних властивостей за рахунок модифікацій передатних функцій нейронів, зокрема шляхом використання апроксимаційних ортогональних поліномів.

1. McCulloch W.S., Pitts W. A logical calculus of ideas immanent in nervous activity// Bull. Mathematical Biophysics. 1943. Vol. 5. P. 115-133.

2. Минский М., Пейперт С. Перцептроны. М., 1971.

3. Уоссермен Ф. Нейрокомпьютерная техника: Теория и практика. М., 1992.

4. Ткаченко Р.О., Неітераційне навчання штучних нейронних мереж прямого поширення // Технічні вісті. 1999. №1(8), 2(9). С.41-42

**УДК 511.217**

# ПРОСТІ ЧИСЛА КВАДРАТНОГО ПОЛІНОМА І ГІПОТЕЗА А.ШИНЦЕЛЯ

**© А. Ковальчук**
Національний університет "Львівська політехніка"

***Встановлено нескінченність множини простих значень квадратного полінома за умови, що цей поліном має хоч би два різні прості значення і якщо дискримінант такого квадратного полінома – не квадрат цілого числа. Доведено гіпотезу А.Шинцеля.***

***The infinity of a set of prime numbers of a square polynomial is established, provided that this polynomial has even two different simple significances and if a discriminant of such polynomial - not guadrate of an integer. Is proved A. Schinzel conjecture.***

У системі RSA з відкритими ключами при кодуванні інформації необхідно використовувати великі прості числа. Доцільно ці прості числа вибирати не з усього натурального ряду, а в деяких його нескінченних підмножинах, зокрема – в множині всіх значень деякого квадратного полінома.

Нехай $a$ – натуральне, а $b$ – ціле число, $0 \le b < a$. Доведення теореми про те, що в кожній арифметичній прогресії, різниця якої а взаємно проста з першим значенням $b$, є нескінченна множина простих чисел, тобто конгруенція $p \equiv b(\mathrm{mod}\ a)$ має нескінченну