ЗАСТОСУВАННЯ МЕТОДУ РЕКУРЕНТНОЇ АПРОКСИМАЦІЇ ДО СИНТЕЗУ НЕЙРОННОЇ МЕРЕЖІ ВИЗНАЧЕННЯ ГІДРОДИНАМІЧНИХ ХАРАКТЕРИСТИК ПІДВОДНИХ АПАРАТІВ

УДК 517.962.27:517.55:004.032.26:629.584

ТРУНОВ Олександр Миколайович
к.т.н., доцент кафедри медичних приладів та систем, перший проректор Чорноморського державного університету ім. Петра Могили.
Наукові інтереси: робототехніка, сенсорна техніка, автоматизовані системи керування, математичне моделювання нелінійних об'єктів та систем.
e-mail: ant@kma.mk.ua, trunovalexandr@gmail.com.

ВСТУП

Розвиток програмного забезпечення моделювання динаміки підводних апаратів передбачає розширення можливостей його застосування для різних видів робіт, технологічного та допоміжного обладнання, типів корпусів, умов експлуатації. Інтеграція додаткових модулів визначення гідродинамічних характеристик, до складу ПЗ моделювання динаміки [1-5], дозволяє гнучко переходити за заданими критеріями від режимів наближеного аналізу до режимів, що враховують особливості геометричних характеристик корпусів, наявності отворів та порожнин, а також впливу збурення поверхневих хвиль на коефіцієнти приєднаних мас та демпфірування підводного апарата у цілому. Перехід від баз даних гідродинамічних характеристик, побудованих для окремих типів корпусів та умов, до баз знань є однією з багатообіцяючих перспектив. Успіхи застосування нейронних мереж (НМ) до створення баз знань в загалі є особливо приваблюючим підходом [8-13]. Однак, існуючи методи навчання таких мереж не є такими, що дозволяють швидко визначати синаптичні коефіцієнти ваги, тому потребують свого удосконалення. Більшість існуючих методів навчання [8-13] стримується математичними проблемами пошуку точного розв'язку системи нелінійних алгебраїчних рівнянь. Однак, успіхи спроб, що здійснювались застосуванням квазілінеаризації та методу рекурентної апроксимації (МРА) [3-5] дозволяють звезти її розв'язок до послідовності, яка швидко збігається [4, 5, 7]. Такі результати стимулюють подальше застосування МРА до постановки та розв'язку нових типів задач [1, 2]. Особливо це становиться привабливим з введенням ідеї про використання вектора індикатора та інтелектуалізації процесу, що відкриває нові можливості МРА навіть для негладких осцилюючих функцій [14]. Дослідження такого комбінованого підходу на сьогоднішній день обмежується тільки скалярною функцією одного аргументу. Застосовність ідеї вектора індикатора до розкладу скалярної або вектор-функції векторного аргументу та подальша інтелектуалізація процесу пошуку аналітичного розв'язку системи нелінійних рівнянь, які утворено негладкими функціями є *головною з невивчених та нерозв'язаних задач*, що буде обмежувати створення НМ.

Мета статті – дослідити можливості аналітичного навчання НМ для функцій активації типу функцій розподілу Фермі, отримати аналітичні вирази для синаптичних коефіцієнтів ваги, продемонструвати швидкість збіжності та точність розв'язків.

ПОСТАНОВКА ЗАДАЧІ АНАЛІТИЧНОГО НАВЧАННЯ НЕЙРОННОЇ МЕРЕЖІ

Для прозорості викладу загальну задачу проектування структури та навчання НМ визначення гідродинамічних характеристик розділимо на два випадки: одношаровий штучний нейрон та багато шарова нейрона мережа.

Одношаровий штучний нейрон. Розглянемо штучний нейрон, який містить чотири входи та один вихід.

Останній є стандартним елементом для конструювання багатошарової НМ. Припустимо, що на всі чотири входи нейрону подано від задатчиків особливостей корпусу та умов обтікання вхідний чотирьох компонентній вектор $\overline{X}$. Після підсилення кожної з компонент, тобто множення на відповідні коефіцієнти синаптичних ваг ω_i та зсуву на задану величину ω_0 формується вхід

$$S = \omega_1 x + \omega_2 y + \omega_3 z + \omega_4 t + \omega_0,$$

а за допомогою наперед заданої активаційної функції $H(S)$ вихід $Y = H(S)$.

Покладемо, що вихідний сигнал є скалярною величиною. Припустимо, що в результаті вивчення бази даних, що зібрано а наслідок експериментальної обробки та узагальнення [3,5] або побудованої за алгоритмами [2, 3] обрано п'ять достовірних еталонних прикладів значень гідродинамічних характеристик приєднаних мас або демпфірування. Для цих еталонів зібрано інформацію про значення всіх компонент вектору входів та виходу. Поставимо задачу знайти величини синаптичних коефіцієнтів ваги, які забезпечують повну відповідність зразковим еталонам або задовольняють вимозі мінімуму суми квадратів відхилень від них.

Введемо шости вимірний простір у який відображається вхідний вектор $\overline{X}$ та вихід Y, тоді відповідно до структури нейрону аргументом функції активації буде S - скалярна, лінійна комбінація входів зсунутих відносно початку відліку і яка теж відображається у цей простір:

$$S_m = \omega_1 x_m + \omega_2 y_m + \omega_3 z_m + \omega_4 t_m + \omega_0 \qquad (1)$$

Припустимо, що активаційну функцію, яка за вимогою [15] є неперервною разом із її похідними та диференційованою, яку загально позначимо $h(S)$ обрано, тоді задача аналітичного навчання [15-16] зведеться до задачі знаходження коренів системи п'яти скалярних рівнянь виду:

$$\delta_m = 0; \qquad m = \overline{1,5}, \qquad (2)$$

де позначено δ_m – похибка виходу для еталонного значення Y_m, яка дорівнює

$$\delta_m = Y_m - h(S_m), \qquad m = \overline{1,5}, \qquad (3)$$

а $h(S_m)$ – значення активаційної функції вектора входу для еталону, належність до якого позначено нижнім індексом m. Розкладемо рівняння системи (3), що складено за умовами (2) у ряди [4,15-16] :

– за лінійною схемою наближення

$$\delta_{mn} = Y_m - h(S_{mn}) - \nabla_\omega^T \left[h(S_{mn}) \right] \Delta\overline{\omega}_n - \\ -\frac{1}{2}\Delta\overline{\omega}_n^T \nabla_\omega \left\{ \nabla_\omega \left[h(S_{mn}) \right] \right\} \Delta\overline{\omega}_{n-1} = 0, \quad ; (4) \\ m = \overline{1,5}$$

– за квадратичною схемою наближення

$$\delta_{mn} = Y_m - h(S_{mn}) - \nabla_\omega^T \left[h(S_{mn}) \right] \Delta\overline{\omega}_n - \\ -\frac{1}{2}\Delta\overline{\omega}_n^T \nabla_\omega \left\{ \nabla_\omega \left[h(S_{mn}) \right] \right\} \Delta\overline{\omega}_n = 0, \quad , (5) \\ m = \overline{1,5}$$

де додатковими нижніми індексами n та $n-1$ позначено номери наближення.

Тепер розглянемо випадок, коли кількість еталонів *M*, яким необхідно обов'язково задовольнити, більше за кількість коефіцієнтів ваги. Тоді задача про мінімізацію суми квадратів похибок сформулюється наступним чином

$$\min_{\overline{\omega}} \sum_{m=1}^{M} \left[Y_m - h(S_m) \right]^2, m = \overline{1,M}, \qquad (6)$$

яка є еквівалентною до задачі про знаходження коренів системи:

$$L_i(\overline{\omega}) = \sum_{m=1}^{M} \delta_m \frac{\partial(S_m)}{\partial \omega_i} \frac{\partial h(S_m)}{\partial S} = 0, \qquad (7) \\ i = \overline{0,4}, \qquad m = \overline{1,M}.$$

Розкладемо рівняння системи (7), що складено за умовами (6) у ряди [4,15-16]:

– за лінійною схемою наближення

$$L_i(\overline{\omega}) = L_i(\overline{\omega}_n) + \nabla_\omega^T \left[L_i(\overline{\omega}_n) \right] \Delta\overline{\omega}_n + \\ +\frac{1}{2}\Delta\overline{\omega}_n^T \nabla_\omega \left\{ \nabla_\omega \left[L_i(\overline{\omega}_n) \right] \right\} \Delta\overline{\omega}_{n-1} = 0, \quad (8) \\ m = \overline{1,5};$$

– за квадратичною схемою наближення

$$L_i(\overline{\omega}) = L_i(\overline{\omega}_n) + \nabla^T_{\omega}\left[L_i(\overline{\omega}_n)\right]\Delta\overline{\omega}_n +$$
$$+\frac{1}{2}\Delta\overline{\omega}_n^T\nabla_{\omega}\left\{\nabla_{\omega}\left[L_i(\overline{\omega}_n)\right]\right\}\Delta\overline{\omega}_n = 0, \quad . \quad (9)$$
$$m = \overline{1,5}$$

ЗАСТОСУВАННЯ ІНТЕЛЕКТУАЛЬНОЇ ОБРОБКИ ВЕЛИЧИНИ ДО РОЗКЛАДУ У РЯДИ СКАЛЯРНОЇ ФУНКЦІЇ АБО ВЕКТОР-ФУНКЦІЇ ВІД ВЕКТОРА

Введемо правило спрацьовування три рівневого компаратору. Покладемо, що на необмеженій множині дійсних чисел $\forall Y \in (-\infty, \infty)$ існує набір трьох еталонів величин Y_1, Y_2, Y_3, тоді введемо компаратор, який реалізуватиме предикат виду

$$D_1(Y_1, Y_2, Y_3) = \begin{cases} -1, & если \quad Y < Y_1, \\ 0, & если \quad Y = Y_2, \\ 1, & если \quad Y > Y_3, \end{cases} \quad (10)$$

де Y_1, Y_2, Y_3 величини приймаються як такі, що допускають вимірювання і які покладені за еталони. Їх значення можуть варіюватися, а у окремих випадках навіть приймаються рівними. Покладемо також, що це правило справджується не тільки для фізичної величини, а й у тому числі для її похідних. Введемо три типи величин індикаторів, які утворено в наслідок обробки за допомогою компаратора типу (10), але у якому усі три значення величин еталонів покладено рівними та які також дорівнюють нулеві:

$$V1 = D\left[L(\overline{x}_n)\right]; \qquad V2_i = D\left[\frac{\partial L(\overline{x}_n)}{\partial x_i}\right];$$
$$V3_{ji} = D\left[\frac{\partial^2 L(\overline{x}_n)}{\partial x_j \partial x_i}\right] \quad . \quad (11)$$

Розклад $L(\overline{x} + \Delta\overline{x})$ – скалярної функції векторного аргументу у ряд Тейлора подамо тепер у вигляді, добутку величини індикаторів стану у тому числі і частинних похідних

$$L(\overline{x}_p + \Delta\overline{x}_p) =$$
$$= L(\overline{x}_p)V1 + \Delta\overline{x}_p^T \begin{bmatrix} b_1 \\ \dots \\ b_i \\ \dots \\ b_n \end{bmatrix} + \Delta\overline{x}_p^T \left\|c_{ij}\right\| \frac{\Delta\overline{x}_p}{2}, \quad (12)$$

де елементи відповідних матриць визначено:

$$b_i = \left|\frac{\partial L(\overline{x}_n)}{\partial x_i}\right| V2_i; \qquad c_{ji} = \left|\frac{\partial^2 L(\overline{x}_n)}{\partial x_j \partial x_i}\right| V3_{ji} \quad . \quad (13)$$

Додаткове введення величин

$$V1_i = D\left[L_i(\overline{x}_n)\right]; \qquad a_i = \left|L_i(\overline{x}_n)\right| V1_i$$
$$b_{ij} = \left|\frac{\partial L_i(\overline{x}_n)}{\partial x_j}\right| V2_{ij}; \qquad c_{ij} = \left|\frac{\partial^2 L_i(\overline{x}_n)}{\partial x_j \partial x_i}\right| V3_{ij} \quad (14)$$

дозволяє узагальнити та подати розклад $L(\overline{x} + \Delta\overline{x})$ - вектор функції векторного аргументу у ряд Тейлора у вигляді, добутку величин індикаторів стану у тому числі і частинних похідних

$$L(\overline{x}_p + \Delta\overline{x}_p) = \begin{bmatrix} a_1 \\ \dots \\ a_i \\ \dots \\ a_n \end{bmatrix} + \left\|b_{ij}\right\| \Delta\overline{x}_p + \Delta\overline{x}_p^T \left\|c_{ij}\right\| \frac{\Delta\overline{x}_p}{2} \quad . \quad (15)$$

Таким чином, за рахунок введення додаткових величин за виразами (11) та (14) і попередньої обробки компаратором (10), з урахуванням виразів для компонент вектор функції приведемо задачу з негладкими функціями до розв'язку у вигляді рекурентної послідовності як це було запропоновано у роботі для скалярної функції одного скалярного аргументу при наявності локальних максимумів та мінімумів, тобто осциляцій [14]. Остання, після підстановок, трансформується до системи п'яти рівнянь.

РОЗВ'ЯЗОК ЗАДАЧІ АНАЛІТИЧНОГО НАВЧАННЯ НЕЙРОНУ

Задачі (2) або (4) або (5) зводяться до задачі знаходження кореня системи нелінійних рівнянь, оскільки вимога неперервності активаційної функції реалізується тільки як нелінійна неперервна функція. До цієї ж задачі зводиться ж задача пошуку синаптичних ваг, що поставлена як задача мінімізації суми квадратів похибок (7) або (8) або (9).

В роботі [1,2,3] продемонстровано, що за умов диференційованості вектор-функції, тобто правих частин системи (2) або (4) або (5) ефективним є метод рекурентної апроксимації. Припустимо, що усі функції активації є неперервними та $N+1$ раз диференційованими, запишемо розклад у ряд умови (4) відповідно до методу рекурентної апроксимації скориставшись кубічною апроксимацією:

– лінійна схема наближення

$$\left[D_{kjm}-h_{kj}^{(q)}\left(S_{kj}^{q}\right)\right]\Big|_{\omega_{kj}^{(q)}=a}-$$

$$-\sum_{j=1}^{N^q}\frac{\partial h_{kj}^{(q)}\left(S_{kj}^{q}\right)}{\partial\omega_{kj}^{(q)}}\Bigg|_{\omega_{kj}^{(q)}=a}\Delta\omega_{kjp+1}^{(q)}-$$

$$-\frac{1}{2}\sum_{j=1}^{N^q}\Delta\omega_{kjp+1}^{(q)}\frac{\partial}{\partial\omega_{kj}^{(q)}}\sum_{j=1}^{N^q}\Delta\omega_{kjp}^{(q)}\frac{\partial h_{kj}^{(q)}\left(S_{kj}^{q}\right)}{\partial\omega_{kj}^{(q)}}\Bigg|_{\omega_{kj}^{(q)}=a}-$$

$$-\frac{1}{6}\sum_{j=1}^{N^q}\Delta\omega_{kjp+1}^{(q)}\frac{\partial}{\partial\omega_{kj}^{(q)}}\sum_{j=1}^{N^q}\Delta\omega_{kjp}^{(q)}\times$$

$$\times\frac{\partial}{\partial\omega_{kj}^{(q)}}\sum_{j=1}^{N^q}\Delta\omega_{kjp}^{(q)}\frac{\partial h_{kj}^{(q)}\left(S_{kj}^{q}\right)}{\partial\omega_{kj}^{(q)}}\Bigg|_{\omega_{kj}^{(q)}=a}\ldots-\ldots=0;$$

$$a=\omega_{kjp}^{(q)};k=\overline{1,N_j^{(Q)}};\ j=\overline{1,N^{(Q)}};m=\overline{1,M};\tag{16}$$

квадратична схема наближення

$$\left[D_{kjm}-h_{kj}^{(q)}\left(S_{kj}^{q}\right)\right]\Big|_{\omega_{kj}^{(q)}=\omega_{kjp}^{(q)}}-$$

$$-\sum_{j=1}^{N^q}\frac{\partial h_{kj}^{(q)}\left(S_{kj}^{q}\right)}{\partial\omega_{kj}^{(q)}}\Bigg|_{\omega_{kj}^{(q)}=\omega_{kjp}^{(q)}}\Delta\omega_{kjp+1}^{(q)}-$$

$$-\frac{1}{2}\sum_{j=1}^{N^q}\Delta\omega_{kjp+1}^{(q)}\frac{\partial}{\partial\omega_{kj}^{(q)}}\sum_{j=1}^{N^q}\Delta\omega_{kjp+1}^{(q)}\times$$

$$\times\left[\begin{array}{l}\dfrac{\partial h_{kj}^{(q)}\left(S_{kj}^{q}\right)}{\partial\omega_{kj}^{(q)}}\Bigg|_{\omega_{kj}^{(q)}=\omega_{kjp}^{(q)}}+\\[2ex]+\dfrac{\partial}{3\partial\omega_{kj}^{(q)}}\displaystyle\sum_{j=1}^{N^q}\Delta\omega_{kjp}^{(q)}\frac{\partial h_{kj}^{(q)}\left(S_{kj}^{q}\right)}{\partial\omega_{kj}^{(q)}}\Bigg|_{\omega_{kj}^{(q)}=\omega_{kjp}^{(q)}}\end{array}\right]=0;\tag{17}$$

$$k=\overline{1,N_j^{(Q)}};\qquad j=\overline{1,N^{(Q)}};\qquad m=\overline{1,M}.$$

Розв'язок (16) подамо у вигляді рекурентної послідовності, визначаючи для $p+1$- шого значення наближення синаптичного коефіцієнту ваги з шару q - $\omega_{ksp+1}^{(q)}$ із кубічної апроксимації за лінійною схемою наближення:

$$\omega_{ksp+1}^{(q)}=\omega_{ksp}^{(q)}-$$

$$-\left\{\begin{array}{l}\left[h_{kj}^{(q)}\left(S_{kj}^{q}\right)-D_{kjm}\right]\Big|_{\omega_{kj}^{(q)}=a}+\displaystyle\sum_{j=1,j\neq s}^{N^q}\frac{\partial h_{kj}^{(q)}\left(S_{kj}^{q}\right)}{\partial\omega_{kj}^{(q)}}\Bigg|_{\omega_{kj}^{(q)}=a}\Delta\omega_{kjp+1}^{(q)}+\\[2ex]+\dfrac{1}{2}\displaystyle\sum_{j=1,j\neq s}^{N^q}\Delta\omega_{kjp+1}^{(q)}\frac{\partial}{\partial\omega_{kj}^{(q)}}\sum_{j=1}^{N^q}P_{kjp}^{(q)}+\\[2ex]+\dfrac{1}{6}\displaystyle\sum_{j=1,j\neq s}^{N^q}\Delta\omega_{kjp+1}^{(q)}\frac{\partial}{\partial\omega_{kj}^{(q)}}\sum_{j=1}^{N^q}\Delta\omega_{kjp}^{(q)}\frac{\partial}{\partial\omega_{kj}^{(q)}}\sum_{j=1}^{N^q}P_{kjp}^{(q)}\end{array}\right\}/$$

$$/\left[\begin{array}{l}\dfrac{\partial h_{kj}^{(q)}\left(S_{kj}^{q}\right)}{\partial\omega_{kj}^{(q)}}\Bigg|_{\omega_{kj}^{(q)}=a}+\dfrac{\partial}{2\partial\omega_{kj}^{(q)}}\displaystyle\sum_{j=1}^{N^q}P_{kjp}^{(q)}+\\[2ex]+\dfrac{\partial}{6\partial\omega_{kj}^{(q)}}\displaystyle\sum_{j=1}^{N^q}\Delta\omega_{kjp}^{(q)}\frac{\partial}{\partial\omega_{kj}^{(q)}}\sum_{j=1}^{N^q}P_{kjp}^{(q)}\end{array}\right]..;.$$

$$P_{kjp}^{(q)}=\Delta\omega_{kjp}^{(q)}\frac{\partial h_{kj}^{(q)}\left(S_{kj}^{q}\right)}{\partial\omega_{kj}^{(q)}}\Bigg|_{\omega_{kj}^{(q)}=a};$$

$$a=\omega_{kjp}^{(q)};k=\overline{1,N_j^{(Q)}};$$

$$j=\overline{1,N^{(Q)}};s=1,N^{(Q)};j\neq s;m=\overline{1,M}.\tag{18}$$

Слід зазначити, що при сумуванні необхідно враховувати умову $j \neq s$, яка виникає як наслідок переносу доданків з правої частини рівняння до лівої.

Для розв'язку задачі з використанням кубічної апроксимації за квадратичною схемою наближення утворимо рекурентну послідовність для визначення $p+1$ - шого значення наближення синаптичного коефіцієнту ваги $\omega_{ksp+1}^{(q)}$:

$$\omega_{kjp+1}^{(q)} = \omega_{kjp}^{(q)} - \frac{B_{kjp}^{(q)}}{2A_{kjp}^{(q)}} \pm \sqrt{\left(\frac{B_{kjp}^{(q)}}{2A_{kjp}^{(q)}}\right)^2 - \frac{F_{kjp}^{(q)}}{A_{kjp}^{(q)}}}, \quad (19)$$

де позначено

$$A_{kjp}^{(q)} = \frac{1}{2}\frac{\partial}{\partial \omega_{ks}^{(q)}}\left[\begin{array}{l} \left.\dfrac{\partial h_{kj}^{(q)}\left(S_{kj}^{q}\right)}{\partial \omega_{ks}^{(q)}}\right|_{\omega_{kj}^{(q)}=a} + \\ + \dfrac{\partial}{3\partial \omega_{ks}^{(q)}}\sum_{j=1}^{N^q} P_{kjp}^{(q)} \end{array}\right];$$

$$B_{kjp}^{(q)} = \left.\frac{\partial h_{kj}^{(q)}\left(S_{kj}^{q}\right)}{\partial \omega_{kj}^{(q)}}\right|_{\omega_{kj}^{(q)}=a};$$

$$F_{kjp}^{(q)} = \left[D_{kjm} - h_{kj}^{(q)}\left(S_{kj}^{q}\right)\right]\Big|_{\omega_{kj}^{(q)}=a} +$$

$$+\sum_{j=1, j\neq s}^{N^q} \left.\frac{\partial h_{kj}^{(q)}\left(S_{kj}^{q}\right)}{\partial \omega_{kj}^{(q)}}\right|_{\omega_{kj}^{(q)}=a} \Delta\omega_{kjp+1}^{(q)} +$$

$$+\frac{1}{2}\sum_{j=1, j\neq s}^{N^q} \Delta\omega_{kjp+1}^{(q)} \frac{\partial}{\partial \omega_{kj}^{(q)}} \sum_{j=1, j\neq s}^{N^q} \Delta\omega_{kjp+1}^{(q)} \times$$

$$\times\left[\begin{array}{l} \left.\dfrac{\partial h_{kj}^{(q)}\left(S_{kj}^{q}\right)}{\partial \omega_{kj}^{(q)}}\right|_{\omega_{kj}^{(q)}=a} + \\ + \dfrac{\partial}{3\partial \omega_{kj}^{(q)}}\sum_{j=1}^{N^q} P_{kjp}^{(q)} \end{array}\right];$$

$$k = \overline{1, N_j^{(Q)}}; \quad ; \quad j = \overline{1, N^{(Q)}}; \quad j \neq s; \quad m = \overline{1, M}.$$

Слід аналогічно зазначити, що при сумуванні необхідно враховувати умову $j \neq s$, яка виникає як наслідок переносу доданків з правої частини рівняння до лівої.

Розв'язок задачі про мінімізацію суми квадратів похибок здійснимо теж за допомогою методу рекурентної апроксимації [4]. Для його реалізації попередньо диференціюємо умову (6) за кожним з коефіцієнтів синаптичних ваг, а далі розкладемо ліві частини у ряди за методом рекурентної апроксимації:

- лінійна схема наближення

$$\sum_{j=1}^{N^q} T_{kjp}^{q} + \sum_{j=1}^{N^q} \Delta\omega_{kjp+1}^{(q)} \frac{\partial}{\partial \omega_{kj}^{(q)}} \sum_{j=1}^{N^q} T_{kjp}^{q} +$$

$$+\frac{1}{2}\sum_{j=1}^{N^q} \Delta\omega_{kjp+1}^{(q)} \frac{\partial}{\partial \omega_{kj}^{(q)}} \sum_{j=1}^{N^q} \Delta\omega_{kjp}^{(q)} \frac{\partial}{\partial \omega_{kj}^{(q)}} \sum_{j=1}^{N^q} T_{kjp}^{q} +$$

$$+\frac{1}{6}\sum_{j=1}^{N^q} \Delta\omega_{kjp+1}^{(q)} \frac{\partial}{\partial \omega_{kj}^{(q)}} \sum_{j=1}^{N^q} \Delta\omega_{kjp}^{(q)} \frac{\partial}{\partial \omega_{kj}^{(q)}} \sum_{j=1}^{N^q} \Delta\omega_{kjp}^{(q)} \times$$

$$\times\sum_{j=1}^{N^q} \frac{\partial}{\partial \omega_{kj}^{(q)}} T_{kjp}^{q} = 0;$$

$$T_{kjp}^{q} = \left.\left\{\frac{\partial h_{kj}^{(q)}\left(S_{kj}^{q}\right)}{\partial \omega_{kj}^{(q)}}\left[D_{kjm} - h_{kj}^{(q)}\left(S_{kj}^{q}\right)\right]\right\}\right|_{\omega_{kj}^{(q)}=a}$$

$$k = \overline{1, N_j^{(Q)}}; \qquad j = \overline{1, N^{(Q)}}; \qquad m = \overline{1, M};$$

- квадратична схема наближення

$$\sum_{j=1}^{N^q} T_{kjp}^{q} + \sum_{j=1}^{N^q} \Delta\omega_{kjp+1}^{(q)} \frac{\partial}{\partial \omega_{kj}^{(q)}} \sum_{j=1}^{N^q} T_{kjp}^{q} +$$

$$+\frac{1}{2}\sum_{j=1}^{N^q} \Delta\omega_{kjp+1}^{(q)} \frac{\partial}{\partial \omega_{kj}^{(q)}} \sum_{j=1}^{N^q} \Delta\omega_{kjp+1}^{(q)} \frac{\partial}{\partial \omega_{kj}^{(q)}} \sum_{j=1}^{N^q} T_{kjp}^{q} +$$

$$+\frac{1}{6}\sum_{j=1}^{N^q}\Delta\omega_{kjp+1}^{(q)}\frac{\partial}{\partial\omega_{kj}^{(q)}}\sum_{j=1}^{N^q}\Delta\omega_{kjp+1}^{(q)}\times$$

$$\times\sum_{j=1}^{N^q}\frac{\partial}{\partial\omega_{kj}^{(q)}}\Delta\omega_{kjp}^{(q)}\frac{\partial}{\partial\omega_{kj}^{(q)}}\sum_{j=1}^{N^q}T_{kjp}^{q}=0;$$

$$k=\overline{1,N_j^{(Q)}};\ j=\overline{1,N^{(Q)}};\ m=\overline{1,M}.$$

ОЦІНКА ЗБІЖНОСТІ

Позначимо для спрощення і компактності викладу та запису ліві частини умов (4) або (5) або (7) загальним виразом $L(x)$. Побудовані алгоритми розв'язку цих рівнянь є послідовністю, яка за умов існування її збіжності взагалі має властивості двохсторонньої та квадратичної [4,6,7]:

для лінійної схеми наближення

$$x_{n+1}=x_n-\frac{L(x_n)}{\sum_{k=1}^{N}\frac{1}{(k)!}\left.\frac{d^kL(x)}{d\Delta^k}\right|_{\Delta_n=0,x=x_n}(\Delta_n)^{k-1}}.$$

Позначимо корінь до якого прямує розв'язок x^*, тоді похибка на $n+1$ - шому наближенні визначиться

$$L(x_n)+\Delta_{n+1}\sum_{k=1}^{N}\frac{1}{(k)!}\left.\frac{d^kL(x)}{d\Delta^k}\right|_{\Delta_n=0,x=x_n}(\Delta_n)^{k-1}=0;$$

$$x_{n+1}-x^*=x_n-\frac{L(x_n)}{\sum_{k=1}^{N}\frac{1}{(k)!}\left.\frac{d^kL(x)}{d\Delta^k}\right|_{\Delta_n=0}(\Delta_n)^{k-1}}-x^*=$$

$$=x_n-\frac{L(x_n)}{\sum_{k=1}^{N}\frac{1}{(k)!}\left.\frac{d^kL(x)}{d\Delta^k}\right|_{\Delta_n=0,x=x_n}(\Delta_n)^{k-1}}- \quad (20)$$

$$-\left\{x^*-\frac{L(x^*)}{\sum_{k=1}^{N}\frac{1}{(k)!}\left.\frac{d^kL(x^*)}{d\Delta^k}\right|_{\Delta_n=0,x=x_n}(\Delta_n)^{k-1}}\right\}=$$

$$=\psi(x_n)-\psi(x^*),$$

де позначено

$$\psi(x_n)=x_n-\frac{L(x_n)}{\sum_{k=1}^{N}\frac{1}{(k)!}\left.\frac{d^kL(x)}{d\Delta^k}\right|_{\Delta_n=0,x=x_n}(\Delta_n)^{k-1}},$$

.

$$\psi(x^*)=\left\{x^*-\frac{L(x^*)}{\sum_{k=1}^{N}\frac{1}{(k)!}\left.\frac{d^kL(x^*)}{d\Delta^k}\right|_{\Delta_n=0,x=x_n}(\Delta_n)^{k-1}}\right\}$$

Розклавши праву частину (20) у ряд Тейлора, обмежившись трьома першими доданками запишемо:

$$x_{n+1}-x^*=\psi(x_n)-\psi(x^*)=$$

$$=\left(x_{n+1}-x^*\right)\psi'(x^*)+\frac{\left(x_{n+1}-x^*\right)^2}{2}\psi''(\theta);$$

$$x_0\le x_n\le\theta\le x^*.$$

Похідна від $\psi(x_n)$ по x_n у точці кореня дорівнює нулеві, оскільки прямим диференціюванням маємо

$$\psi'(x_n)=1-$$

$$-\frac{\left\{\begin{aligned}&L'(x_n)\sum_{k=1}^{N}\frac{1}{(k)!}\left.\frac{d^kL(x)}{d\Delta^k}\right|_{\Delta_n=0,x=x_n}(\Delta_n)^{k-1}-\\&-L(x_n)\left[\sum_{k=1}^{N}\frac{(k-1)}{(k)!}\left.\frac{d^kL(x)}{d\Delta^k}\right|_{\Delta_n=0,x=x_n}(\Delta_n)^{k-2}\right]\end{aligned}\right\}}{\left\{\sum_{k=1}^{N}\frac{1}{(k)!}\left.\frac{d^kL(x)}{d\Delta^k}\right|_{\Delta_n=0,x=x_n}(\Delta_n)^{k-1}\right\}^2},$$

а $L(x^*)=0,\qquad \Delta_n\Big|_{x=x^*}=0$.

Таким чином, друга похідна від $\psi(x_n)$ по x_n у точці кореня визначає швидкість збіжності, а значить третя від $L(x^*)$ по x_n у точці кореня визначає швидкість збіжності, а сама збіжність є двохсторонньою та квадратичною.

Тепер розглянемо квадратичну схему наближення [4,6]. Доведення квадратичної збіжності здійснюється двома шляхами. Перший з них, полягає у знаходженні границі, а потім дослідженні її поведінки. Другий здійснюється обчисленням значень функції рекурентного розкладу

$$L(x_n)+\Delta_{n+1}\frac{d\ L\ (x)}{d\Delta}\bigg|_{\Delta_n=0,x=x_n}+$$

$$+\Delta_{n+1}^2\sum_{k=2}^{N}\frac{1}{(k)!}\frac{d^kL\ (x)}{d\Delta^k}\bigg|_{\Delta_n=0,x=x_n}(\Delta_n)^{k-2}=0;$$

Так позначимо

$$B(x_n)=2\sum_{k=2}^{m}\frac{\partial^k L(x)}{\partial\Delta^k}\bigg|_{\Delta_n=0,x=x_n}\frac{\Delta_{n-1}^{k-2}}{k!}$$

та припустимо, що при $B(x)\to 0$ існують границі

$$\lim_{B(x)\to 0}\frac{L'(x)}{B(x)},$$

разом з

$$\lim_{B(x)\to 0}\frac{L(x)}{B(x)},$$

причому $\lim\limits_{B(x)\to 0}\dfrac{2L(x)B(x)}{\left[L'(x)\right]^2}\ll 1.$

За цих умов для квадратичної схеми наближення маємо

$$x_{n+1}=x_n-\frac{L'(x_n)}{B(x_n)}\left[1\mp\sqrt{1-\frac{2L(x_n)B(x_n)}{\left[L'(x_n)\right]^2}}\right],$$

а за додаткових умов припущення про опуклість функції, останнє спрощується і має границю:

$$\lim_{B(x)\to 0}x_{n+1}=x_n-\frac{L(x_n)}{L'(x_n)}.$$

Таким чином, для границі маємо вигляд лінійної схеми наближення лінійної апроксимації швидкість збіжності якої як показано в [4,6] є двохсторонньої та квадратичної.

Доведемо за іншим підходом, що для квадратичної схеми наближення у загальному випадку має місце двохстороння та квадратична збіжність.

Введемо позначення

$$x_{n+1}-x^*=x_n-$$

$$-\frac{L'(x_n)}{B(x_n)}\pm\sqrt{\left[\frac{L'(x_n)}{B(x_n)}\right]^2-\frac{2L(x_n)}{B(x_n)}}-x^*=$$

$$=x_n-\frac{L'(x_n)}{B(x_n)}\pm\sqrt{\left[\frac{L'(x_n)}{B(x_n)}\right]^2-\frac{2L(x_n)}{B(x_n)}}-\qquad(21)$$

$$-\left\{x^*-\frac{L'(x^*)}{B(x^*)}\pm\sqrt{\left[\frac{L'(x^*)}{B(x^*)}\right]^2-\frac{2L(x^*)}{B(x^*)}}\right\}=$$

$$=\psi(x_n)-\psi(x^*)$$

де позначено

$$\psi(x_n)=x_n-\frac{L'(x_n)}{B(x_n)}\pm\sqrt{\left[\frac{L'(x_n)}{B(x_n)}\right]^2-\frac{2L(x_n)}{B(x_n)}},$$

.

$$\psi(x^*)=x^*-\frac{L'(x^*)}{B(x^*)}\pm\sqrt{\left[\frac{L'(x^*)}{B(x^*)}\right]^2-\frac{2L(x^*)}{B(x^*)}}$$

Розклавши праву частину (21) у ряд Тейлора, обмежившись трьома першими доданками запишемо:

$$x_{n+1} - x^* = \psi(x_n) - \psi(x^*) =$$

$$= \left(x_{n+1} - x^*\right)\psi'(x^*) + \frac{\left(x_{n+1} - x^*\right)^2}{2}\psi''(\theta);$$

$$x_0 \le x_n \le \theta \le x^*.$$

Застосуємо безпосереднє диференціювання функції $\psi\left(x_n\right)$ по x_n та обчислимо її значення у точці кореня x^*

$$\psi'\left(x_n\right) = 1 - \left(\frac{L'(x)}{B(x)}\right)' \pm$$

$$\pm \frac{\left(\frac{L'(x)}{B(x)}\right)\left(\frac{L'(x)}{B(x)}\right)' - \frac{L'(x)B(x) - L(x)B'(x)}{B(x)^2}}{\sqrt{\left[\frac{L'(x)}{B(x)}\right]^2 - \frac{2L(x)}{B(x)}}}.$$

Останнє з урахуванням, що $L(x^*) = 0$ дозволяє обчислити границю похідної у точці кореня x^*

$$\lim_{x \to x^*} \psi'\left(x_n\right) =$$

$$= 1 - \left(\frac{L'(x)}{B(x)}\right)' + \left(\frac{L'(x)}{B(x)}\right)' - 1 = 0'$$

що дозволяє стверджувати; - при квадратичній схемі наближення має місце двохстороння та квадратична збіжність, а її швидкість залежить від властивостей функції активації та виду поставленої задачі та меншою мірою від схеми наближення. Так для задачі навчання у якій забезпечується задане значення похибки вона визначається максимальною величиною третьої та мінімальною четвертої похідної від функції активації, а задачі мінімізації суми квадратів похибки максимумом четвертої та мінімумом п'ятої її похідної.

ВИСНОВКИ

1. За допомогою методу рекурентної апроксимації розв'язано задачу про аналітичне налаштування, яку зведено до системи нелінійних алгебраїчних рівнянь розв'язок якої подано як рекурентну послідовність.
2. Швидкість збіжності утвореної послідовності розв'язків визначається величиною оцінки похідних функції активації.

ЛІТЕРАТУРА:

1. Trunov O.M. Rozvitok metodiv ocinki efektivnosti sistem upravlinnja robotizovanimi kompleksami u glibokovodnih tehnologijah //Vestnik Hersonskogo nacional'nogo tehnicheskogo universiteta. – 2013. – №1 (46). – S.328-337.
2. Trunov O. M., Udoskonalennja PZ modeljuvannja dinamiki pidvodnih aparativ za umov reguljarnih poverhnevih hvil' /Trunov O.M., Novosadovs'kij O.O. //Materiali IV Mizhnarodnoї naukovo-tehnichnoї konferencії «Innovacії v sudnobuduvanni ta okeanotehnici»: Materiali mizhnarodnoї naukovo-tehnichnoї konferencії. – Mikolaїv: NUK im. S.O. Makarova, 2013. – S.412-413.
3. Trunov A.N. Matematicheskaja model' podvodnogo apparata s izmenjajushhejsja geometriej korpusa, //Naukovi praci Mikolaїvs'kogo Derzhavnogo gumanitarnogo universitetu im. Petra Mogili, Naukovo-metodichnij zhurnal. – 2005. – t.41, vip.28. – S.22-31.
4. Trunov A.N. Rekurentna aproksimacija u zadachah modeljuvannja ta proektuvannja: Monografija. Mikolaїv, 2012. – 270 s.
5. Trunov A.N. Opredelenie parametrov dvizhitelej i bortovyh priborov izmeritel'nogo kompleksa podvodnogo aparata monitoringa vodnoj akvatorii na stadijah predjeskiznogo proektirovanija //Naukovi praci Mikolaїvs'kogo Derzhavnogo gumanitarnogo universitetu im. Petra Mogili, Naukovo- metodichnij zhurnal. – 2004. – t.31, vip.18. – S.163-176.
6. Trunov O.M. Doslidzhennja zbizhnosti rozv'jazkiv zadach nelinijnogo programuvannja ta optimal'nogo proektuvannja /O.M. Trunov //Naukovi praci Mikolaїvs'kogo Derzhavnogo gumanitarnogo universitetu im. Petra Mogili, Naukovo- metodichnij zhurnal. – 2009. – T.117. Vip.104. «Komp'juterni tehnologii». – S.68-79.
7. Bellman R.E. Quasilinearization and nonlinear boundary-value problems /Bellman R.E, Kalaba R.E: New York, 1965. – 183 p.
8. Croall I.F., Mason J.P., Industrial Applications of Neural Networks (research reports Esprit, I.F. Croall, J.P. Mason).
9. Learning internal representations by error propagation by Rumelhart, Hinton and Williams (1986).
10. Alkon, D.L, Memory Storage and Neural Systems, Scientific American, July, 1989, pp.42-50.
11. Minsky and Papert, Perceptrons, An introduction to computational geometry, MIT press, expanded edition, 1969.

12. Trunov A. Robust Fault Diagnosis of State and Sensor Faults in Nonlinear Multivariable Systems /Trunov A., Polycarpov M. //Proceedings of the 1999 American Control Conference, pp.608-612, June 1999.
13. Trunov A., Robust Nonlinear Fault Diagnosis: Application to Robotic Systems /Trunov A., Polycarpov M. //Proceedings of the 1999 IEEE Conference on Control Applications, pp.1424-1429, August 1999.
14. Trunov A.N. Intellectualization of the models' transformation process to the recurrent sequence, European Applied Sciences, Ort Publishing, # 9 – 2013, vol. 1, p.123-130.
15. Trunov O.M. Zastosuvannja metodu rekurentnoї aproksimacії do analitichnogo navchannja nejronnoї merezhi viznachennja gidrodinamichnih harakteristik pidvodnih aparativ /Trunov O.M., Novosadovs'kij O.O., Kihtenko D.P. //Materiali IV Mizhnarodnoї naukovo-tehnichnoї konferencії «Innovacії v sudnobuduvanni ta okeanotehnici»: Materiali mizhnarodnoї naukovo-tehnichnoї konferencії. – Mikolaïv: NUK im. S.O. Makarova, 2014. – S.389-392.
16. Trunov A. Application of recurrent approximation to the synthesis of a eural network for process control physical rehabilitation /Trunov A. //Australian science review, Melbourne, 2014, No.1. – P.101-116.

Рецензент: *д.т.н., проф. Фисун Н.Т.,*
Черноморский государственный университет им. П. Могилы.