УДК 602.1:519.85:53.082.9:616-07
DOI: 10.20535/SRIT.2308-8893.2019.2.07

ЕКСПОНЕНЦІАЛЬНА ОЦІНКА ДЛЯ РЕКУРЕНТНОЇ НЕЙРОННОЇ МЕРЕЖІ З ДИСКРЕТНИМ ЗАПІЗНЕННЯМ

В.П. МАРЦЕНЮК, А.С. СВЕРСТЮК

Анотація. Розроблено та застосовано метод розрахунку швидкості експоненціального згасання для моделі рекурентної нейронної мережі на основі диференціальних рівнянь із дискретним запізненням. Експоненціальну оцінку отримано на основі різницевої нерівності для функціонала Ляпунова. Розглянуто приклад експоненціального оцінювання для моделі рекурентної нейронної мережі з трьома нейронами.

Ключові слова: рекурентна нейронна мережа, диференціальні рівняння із запізненням, експоненціальна стійкість, функціонал Ляпунова.

ВСТУП

Протягом останнього часу дослідження в галузі штучних нейронних мереж є одними з пріоритетних у науці та техніці. Застосування нейронних мереж, пов'язане з діагностикою [1–3], розпізнаванням образів [4, 5], обробленням сигналів [6], має важливе значення для вирішення багатьох прикладних завдань. Для вирішення найскладніших завдань (наприклад, моделювання та оброблення неперервного рукописного тексту, мовлення і мови) пропонують використовувати рекурентні нейронні мережі [7, 8].

З метою дослідження широкого кола проблем аналізу стійкості, синхронізації та оцінювання збіжності застосовують моделі рекурентних нейронних мереж на основі систем диференціальних рівнянь із часовими запізненнями. Неперервно або дискретно розподілені запізнення, спричинені часом, що затрачається на обчислення на попередніх вузлах, спостерігаються в різних нейронних мережах. Як приклади можна навести нейронні мережі Хопфілда, клітинні нейронні мережі та двонапрямлені нейронні мережі з асоціативною пам'яттю. При цьому запізнення можуть впливати на хід коливань, нестійкість та продуктивність мережі [9–11].

За допомогою моделей на основі диференціальних рівнянь із запізненням можуть бути досліджені динамічні характеристики рекурентних нейронних мереж із запізненнями часу [12, 13].

Ураховуючи той факт, що глобальна стійкість є однією з найважливіших динамічних властивостей нейронних мереж, протягом останнього часу дослідження стійкості для нейронних мереж із запізненням стосувалися глобальної асимптотичної стійкості [9, 10, 14, 15], глобальної експоненціальної стійкості [11, 16] і робастної стійкості [17].

Експоненціальна стійкість є часто використовуваною для дослідження багатьох систем, оскільки вона описується чіткими показниками швидкості згасання експоненціальної оцінки. Для отримання таких оцінок застосовується ряд методів, серед яких доцільно виокремити непрямі [18, 19] та прямі методи. Прямі методи грунтуються на побудові функціонала Ляпунова та оцінюванні його загальної похідної. Експоненціальну оцінку можна знайти як розв'язок різницево-диференціальної нерівності щодо функціонала Ляпунова [20]. Утім, у праці [21] для лінійної нестаціонарної системи розроблено метод, який грунтується на отриманні різницевої нерівності для функціонала Ляпунова. Цей метод складніший для побудови нерівності для функціонала Ляпунова, але пропонує більш простий розв'язок цієї нерівності. У свою чергу, швидкість згасання може бути розрахована як результат числового розв'язання нелінійного алгебричного рівняння.

Мета роботи — розроблення методу побудови оцінки експоненціального згасання для розв'язання систем диференціальних рівнянь, що використовуються як моделі рекурентних нейронних мереж. Таку оцінку можна застосовувати для вибору параметрів нейронної мережі, що забезпечували б її оптимальну збіжність до стійкого стану.

У роботі використано такі позначення:

$\left|\varphi(\bullet)\right|_{\tau} = \sup\limits_{\theta\in[-\tau,0],\, i=\overline{1,n}} \left|\varphi_i(\theta)\right|$ — норма вектора-функції, де функції $\varphi \in C^1[-\tau,0]$ є неперервно диференційованими на $[-\tau,0]$;

$\|M\|$ — довільна матрична норма для матриці $M \in R^{n\times n}$;

$\|x\|$ — евклідова норма для вектора $x \in R^n$;

$\lambda_{\max}(\bullet)$ — максимальне власне значення матриці;

$\tau_{\max} = \max\{\tau_1,...,\tau_r\}$.

МАТЕМАТИЧНА МОДЕЛЬ РЕКУРЕНТНОЇ НЕЙРОМЕРЕЖІ

Розглядається рекурентна нейронна мережа, описана системою диференціальних рівнянь з дискретними запізненнями:

$$\dot{x}(t) = -Ax(t) + \sum_{m=1}^{r} W_m g(x(t-\tau_m(t))), \tag{1}$$

де $x(t)\in R^n$ — вектор стану; $A = \operatorname{diag}(a_1, a_2, ..., a_n)$ — діагональна матриця з додатними елементами; $a_i > 0$, $W_m = (w_{ij}^m)_{n\times n}$, $m=\overline{1,r}$ — матриці ваг; $g(x(t)) = [g_1(x(t)), g_2(x(t)),..., g_n(x(t))]^{\mathrm{T}} \in R^n$, $x(t)$ позначає функції активації нейрона, які є монотонно обмеженими, неспадними, $g_j(0)=0$ і виконуються умови

$$0 \le \frac{g_j(\xi_1) - g_j(\xi_2)}{\xi_1 - \xi_2} \le l_j; \tag{2}$$

$$\xi_1, \xi_2 \in R, \quad \xi_1 \ne \xi_2, \quad j = 1, 2, ..., n.$$

Обмежені функції $\tau_m(t)$ є дискретними запізненнями системи, причому $0 \le \tau_m(t) \le \tau_M$, $\tau_m(t) \le \tau_D(t) < 1$, $m=\overline{1,r}$.

Початкові умови системи (1) мають вигляд

$$x_i(s) = \varphi_i(s), \quad s \in [-\tau_M, 0]. \tag{3}$$

У формулі (3) $\varphi_i(s)$ є неперервними дійсно значущими функціями, якщо $s \in [-\tau_M, 0]$.

Як показано у праці [22], під час виконання припущення (2) розв'язок системи (1) існує для всіх $t \ge 0$ і є єдиним.

РОЗРОБЛЕННЯ МЕТОДУ ЕКСПОНЕНЦІАЛЬНОГО ОЦІНЮВАННЯ

Основна ідея методу експоненціального оцінювання розв'язків (1) полягає в побудові різницевої нерівності для квадратичної функції Ляпунова $\upsilon(\varphi) \in C^1[-\tau, 0]$ у вигляді

$$\upsilon(t) \le d\upsilon(t - 2\tau), \quad t \ge 2\tau,$$

для деякої сталої $d \in (0,1)$.

Вихідною є нерівність для похідної функціонала Ляпунова у вигляді $\dot{\upsilon}(t) \le -\mu \left\| x(t)^2 \right\|$ для деякого $\mu > 0$.

Для отримання експоненціальної оцінки необхідно виконати чотири основні кроки, зокрема побудувати:

1) нижню оцінку для розв'язку $x(t)$, а саме функцію $M(t)$, що залежить від значення функціонала Ляпунова $\upsilon(t)$ таку, що для довільного достатньо великого t існує $s \in [t - \tau, t]$ таке, що $\|x(s)\| \ge M(t)$;

2) верхню оцінку для $\dot{x}(t)$, а саме функцію $L(t)$ (залежить від значення $\upsilon(t - \tau)$), що за довільного достатньо великого t отримуємо $\|\dot{x}(t)\| \le L(t)$;

3) нижню оцінку для розв'язку $x(t)$, а саме функцію $M(t)$ (залежить від значення функціонала Ляпунова $\upsilon(t)$), що для довільного достатньо великого t існує такий інтервал $I \subset [t - \tau, t]$, що $\|x(s)\| \ge M(t)/2$ для будь-якого $s \in I$;

4) різницеву нерівність для квадратичного функціонала Ляпунова $\upsilon(\varphi)$ за допомогою інтегрування на інтервалі $s \in [t - \tau, t]$ оцінки $\dot{\upsilon}(t) \le -\mu \|x(t)\|^2$ для деякого $\mu > 0$.

ОЦІНКА ДЛЯ ПОХІДНОЇ ФУНКЦІОНАЛА ЛЯПУНОВА

Розглянемо функціонал Ляпунова у вигляді

$$\upsilon(\varphi) = \varphi^{\mathrm{T}}(0)\varphi(0) + \sum_{m=1}^{r} \int_{-\tau_m}^{0} g^{\mathrm{T}}(\varphi(s)) V_m(s) g(\varphi(s)) ds, \tag{4}$$

де $V_m(s) \in \mathrm{R}^{n \times n}$, $s \in [-\tau_m, 0]$ — деякі додатно визначені матричні функції.

Обчислимо похідну функціонала вздовж розв’язків системи (1):

$$\dot{\upsilon}(\varphi)=\left(-x^{\mathrm{T}}(t)A^{\mathrm{T}}+\sum_{m=1}^{r}g^{\mathrm{T}}(x(t-\tau_m))W_m^{\mathrm{T}}\right)x(t)+$$

$$+x^{\mathrm{T}}(t)\left(-Ax(t)+\sum_{m=1}^{r}W_m g(x(t-\tau_m))\right)+$$

$$+\sum_{m=1}^{r}(g^{\mathrm{T}}(x(t))V_m(t)g(x(t))-g^{\mathrm{T}}(x(t-\tau_m))V_m(t-\tau_m)g(x(t-\tau_m)))=\xi^{\mathrm{T}}\Xi\xi, \quad (5)$$

де $\xi=(x(t),g(x(t)),g(x(t-\tau_1)),...,g(x(t-\tau_r)))^{\mathrm{T}}$;

$$\Xi=\begin{bmatrix}-2A & \theta & W_1 & W_2 & \cdots & W_r\\ \theta & \sum_{m=1}^{r}V_m(t) & \theta & \theta & \cdots & \theta\\ W_1^{\mathrm{T}} & \theta & -V_1(t-\tau_1) & \theta & \cdots & \theta\\ W_2 & \theta & \theta & -V_2(t-\tau_2) & \cdots & \theta\\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots\\ W_r^{\mathrm{T}} & \theta & \theta & \theta & \cdots & -V_r(t-\tau_r)\end{bmatrix},$$

де $\theta\in R^{n\times n}$ — матриця з нульовими елементами.

Лема 1. Згідно з працею [21] для довільних векторів $y,z\in R^n$ та матриці $C\in R^{n\times n}$ маємо

$$2y^{\mathrm{T}}C^{\mathrm{T}}z\le z^{\mathrm{T}}P^{-1}z+y^{\mathrm{T}}C^{\mathrm{T}}PCy,$$

де P — довільна додатно визначена матриця.

Розглянемо складову $2g^{\mathrm{T}}(x(t-\tau_m))W_m^T x(t)$ у формулі (5). Для будь-якого фіксованого розв’язку x і фіксованого t можна застосувати лему 1, де замість C, y та z покладемо W_m, $g(x(t-\tau_m))$ і $x(t)$ відповідно. Отримуємо

$$2g^{\mathrm{T}}(x(t-\tau_m))W_m{}^{\mathrm{T}}x(t)\le x^{\mathrm{T}}(t)P_m{}^{-1}x(t)+$$

$$+g^{\mathrm{T}}(x(t-\tau_m))W_m{}^{T}P_mW_m\, g(x(t-\tau_m))$$

для довільної додатно визначеної матриці P_m.

Застосувавши верхню оцінку у формулі (5), дістанемо

$$\dot{\upsilon}\le x^{\mathrm{T}}(t)\left(-2A+\sum_{m=1}^{r}P_m^{-1}\right)x(t)+\sum_{m=1}^{r}g^{\mathrm{T}}(x(t))V_m(t)g(x(t))+$$

$$+\sum_{m=1}^{r}g^{\mathrm{T}}(x(t-\tau_m))(W_m^{\mathrm{T}}P_mW_m-V_m(t-\tau_m))g(x(t-\tau_m))=\xi^{\mathrm{T}}\Xi_1\xi, \quad (6)$$

де

$$\Xi_1 = \begin{bmatrix} -2A+\sum_{m=1}^{r} P_m^{-1} & \theta & \theta & \dots & \theta \\ \theta & V_m(t) & \theta & \cdots & \theta \\ \theta & \theta & W_1^{\mathrm{T}} P_1 W_1 - V_1(t-\tau_1) & \cdots & \theta \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \theta & \theta & \theta & \cdots & W_r^{\mathrm{T}} P_r W_r - V_r(t-\tau_r) \end{bmatrix}.$$

Теорема 1. Нехай система (1) є такою, що існують додатно визначені матриці $P_1,...,P_r$, для яких матриця

$$\Xi_2 = -2A + \sum_{m=1}^{r} (P_m^{-1} + \mathrm{diag}\,(l^2) W_m^{\mathrm{T}} P_m W_m)$$

є від’ємно визначеною; тут $\mathrm{diag}\,(l^2) = \mathrm{diag}(l_1^2,...,l_r^2)$.

Тоді існує $\mu_1 < 0$ таке, що $\dot{\upsilon}_{(1)} \le \mu_1 \| x(t) \|^2$.

Доведення. Продовжуючи формулу (6) та використовуючи припущення (2) для $g(\bullet)$, можна записати:

$$\dot{\upsilon}_{(1)} \le x^{\mathrm{T}}(t)\left(-2A + \sum_{m=1}^{r} P_m^{-1}\right) x(t) + \sum_{m=1}^{r} x^{\mathrm{T}}(t)\,\mathrm{diag}\,(l) V_m(t) \mathrm{diag}(l) x(t) +$$

$$+ \sum_{m=1}^{r} g^{\mathrm{T}}(x(t-\tau_m))(W_m^{\mathrm{T}} P_m W_m - V_m(t-\tau_m)) g(x(t-\tau_m)).$$

Замінивши $W_m^{\mathrm{T}} P_m W_m$ на V_m, отримаємо

$$\dot{\upsilon}_{(1)} \le x^{\mathrm{T}}(t)\left(-2A + \sum_{m=1}^{r} P_m^{-1} + \sum_{m=1}^{r} \mathrm{diag}(l) W_m^{\mathrm{T}} P_m W_m \mathrm{diag}\,(l)\right) x(t) \le$$

$$\le \lambda_{\max}(\Xi_2) \| x(t) \|^2 = \mu_1 \| x(t) \|^2 .$$

Теорему доведено.

Скалярний випадок. Розглянемо нерівність

$$\dot{\upsilon}_1 \le \left(-2A + \sum_{m=1}^{r} P_m^{-1} + W_m^2 P_m l^2\right) x^2(t).$$

Права частина нерівності має мінімальне значення, якщо $W_m \ne 0$ і $P_m = 1/|W_m|$, $m = \overline{1, r}$. У цьому випадку

$$\dot{\upsilon}_1 \le \left(-2A + (1 + l^{\mathrm{T}} l) \sum_{m=1}^{r} | W_m |\right) x^2(t). \qquad (7)$$

РІЗНИЦЕВА НЕРІВНІСТЬ ДЛЯ ФУНКЦІОНАЛА ЛЯПУНОВА

Використаємо допоміжні результати [21], отримані для лінійної нестаціонарної системи із запізненням, виходячи з нерівності для похідної функціонала Ляпунова (4).

Лема 2. Для довільного $t \geq \tau_{\max}$ існує $s \in [t-\tau_{\max}, t]$ таке, що $\| x(s) \| \geq M(t)$, де

$$M(t) = \left(\frac{\upsilon(t)}{1 + l^{\mathrm{T}} l \sum_{m=1}^{r} | \lambda_{\max}(V_m(\bullet)) |_{\tau_m} \tau_m} \right)^{1/2}.$$

Лема 3. Якщо $\dot{\upsilon}_{(1)} \leq 0$, тоді $\| x(t) \| \leq L(t)$, $t \geq \tau_{\max}$, де $L(t) = \left(\| A \| + \| l \| \sum_{m=1}^{r} \| W_m \| \right) (\upsilon(t - 2\tau_{\max}))^{1/2}$.

Лема 4. Для довільного $t \geq 2\tau_{\max}$ існує $s \in [t-\tau_{\max}, t]$ таке, що для будь-якого $\theta \in I$ співвідношення $\| x(\theta) \| \geq \dfrac{M}{2}$ виконується, якщо $I = [s - M/2L_1; s + M/2L_1] \cap [t - \tau_{\max}, t]$, де M використано з леми 2 та $L_1 = \left(\| A \| + \| l \| \sum_{m=1}^{r} \| W_m \| \right) (\upsilon(t - 2\tau_{\max}))^{1/2}$.

Теорема 2. Вважатимемо, що припущення теореми 1 виконується. Тоді існує додатна стала $d < 1$ така, що

$$\upsilon(t) \leq d\upsilon(t - \tau_{\max}), \; t \geq 2\tau_{\max}.$$

Доведення. Із теореми 1 випливає, що

$$\dot{\upsilon}_{(1)} \leq \lambda_{\max}(\Xi_2) \| x(t) \|^2 < 0. \tag{8}$$

Якщо $\upsilon(t^* - 2_{T\max}) = 0$ для деякого t^*, то $\upsilon(t^*) = 0$ і твердження є справедливими для t^*.

Припустімо, що $\upsilon(t^* - 2\tau_{\max}) > 0$ для деякого t^*. Від супротивного матимемо

$$\upsilon(t^*) > d\upsilon(t^* - \tau_{\max}). \tag{9}$$

Із виразу (8) випливає, що

$$0 < \upsilon(t^*) < \upsilon(t^* - \tau_{\max}) + \lambda_{\max}(\Xi_2) \int_{t^* - \tau_{\max}}^{t^*} \| x(\theta) \|^2 d\theta \leq \upsilon(t^* - \tau_{\max}) +$$

$$+ \lambda_{\max}(\Xi_2) \int_I \left(\frac{M(t^*)}{2} \right)^2 d\theta \leq$$

$$\leq \upsilon(t^* - \tau_{\max}) + \lambda_{\max}(\Xi_2) \left(\frac{M(t^*)}{2} \right)^2 \min\left(\frac{\tau_{\max}}{2}; \frac{M(t^*)}{2L_1(t^*)} \right).$$

Ураховуючи (9) для $M(t^*)/L_1(t^*)$, маємо таку оцінку:

$$\frac{M(t^*)}{L_1(t^*)}=\left(\frac{\upsilon(t^*)}{\upsilon(t^*-2\tau_{\max})}\right)^{1/2}\frac{1}{\left(1+l^{\mathrm{T}}l\sum_{m=1}^{r}\left|\lambda_{\max}(V_m(\bullet))\right|_{\tau m}\tau_m\right)^{1/2}}\frac{1}{\|A\|+\|l\|\sum_{m=1}^{r}\|W_m\|}>$$

$$>\frac{1}{\left(1+l^{\mathrm{T}}l\sum_{m=1}^{r}\|W_m\|\lambda_{\max}(V_m(\bullet))\big|_{\tau_m}\tau_m\right)^{1/2}}\frac{1}{\|A\|+\|l\|\sum_{m=1}^{r}\|W_m\|}d^{1/2}.$$

Визначаємо константу ρ таким чином:

$$0<\rho:=\frac{1}{(1+l^{T}l\sum_{m=1}^{r}\left|\lambda_{\max}(V_m(\bullet))\right|_{\tau m}\tau_m)^{1/2}+\|A\|+\|l\|\sum_{m=1}^{r}\|W_m\|}\;\cdot$$

Тоді $\dfrac{M(t^*)}{L_1(t^*)}>\rho d^{1/2}$.

Отже, $\upsilon(t^*)\le\upsilon(t^*-\tau_{\max})+\lambda_{\max}(\Xi_2)\left(\dfrac{M(t^*)}{2}\right)^2\min\left(\dfrac{\tau_{\max}}{2};\dfrac{\rho d^{1/2}}{2}\right)$.

Спочатку припустімо, що $\tau_{\max}>\rho d^{1/2}$. Звідси

$$\upsilon(t^*)\le\upsilon(t^*-\tau_{\max})+\lambda_{\max}(\Xi_2)\frac{\upsilon(t^*)}{1+l^{\mathrm{T}}l\sum_{m=1}^{r}(\lambda_{\max}(V_m(\bullet)))_{\tau_m}\tau_m}\frac{\rho d^{1/2}}{2}.$$

Тобто $\upsilon(t^*)\le\upsilon(t^*-\tau_{\max})\dfrac{1}{1-\dfrac{\lambda_{\max}(\Xi_2)}{8}+\dfrac{\rho d^{1/2}}{1+l^{\mathrm{T}}l\sum_{m=1}^{r}\left|\lambda_{\max}(V_m(\bullet))\right|_{\tau_m}\tau_m}}$.

У цьому випадку нехай d є єдиним (додатним) розв’язком рівняння

$$\frac{1}{1-\dfrac{\lambda_{\max}(\Xi_2)}{8}+\dfrac{\rho d^{1/2}}{1+l^{\mathrm{T}}l\sum_{m=1}^{r}\left|\lambda_{\max}(V_m(\bullet))\right|_{\tau_m}\tau_m}}=d\ ,$$

або

$$d-\frac{\lambda_{\max}(\Xi_2)}{8}\frac{\rho d^{1/2}}{1+l^{\mathrm{T}}l\sum_{m=1}^{r}\left|\lambda_{\max}(V_m(\bullet))\right|_{\tau_m}\tau_m}d^{3/2}-1=0. \qquad (10)$$

Використовуючи (10), можна записати $\upsilon(t^*)\le d\upsilon(t^*-\tau_{\max})$, що є суперечністю.

З іншого боку, припустімо, що $\tau_{\max}\le\rho d^{1/2}$.

Тоді $\upsilon(t^*) \le \upsilon(t^* - \tau_{\max}) + \lambda_{\max}(\Xi_2)\dfrac{1}{4}\dfrac{\upsilon(t^*)}{1 + l^{\mathrm{T}} l \sum\limits_{m=1}^{r} \left|\lambda_{\max}\left(V_m(\bullet)\right)\right|_{\tau_m} \tau_m} \dfrac{\tau_{\max}}{2}$.

Тобто $\upsilon(t^*) \le \upsilon(t^* - \tau_{\max}) \dfrac{1}{1 - \dfrac{\lambda_{\max}(\Xi_2)}{8} \dfrac{\tau_{\max}}{1 + l^{\mathrm{T}} l \sum\limits_{m=1}^{r} \left|\lambda_{\max}(V_m(\bullet))\right|_{\tau_m} \tau_m}} + \dfrac{\tau_{\max}}{2}$.

У цьому випадку нехай d матиме такий вигляд:

$$d = \frac{1}{1 - \dfrac{\lambda_{\max}(\Xi_2)}{8} \dfrac{\tau_{\max}}{1 + l^{\mathrm{T}} l \sum\limits_{m=1}^{r} \left|\lambda_{\max}(V_m(\bullet))\right|_{\tau_m} \tau_m}}.$$

Тоді $\upsilon(t^*) \le d\upsilon(t^* - \tau_{\max})$, що також є суперечністю.

Наслідок 1. Припускаючи, що умови теореми 2 виконуються, константу $d < 1$ можна подати у вигляді:

$$d = \begin{cases} \dfrac{1}{1 - \dfrac{\lambda_{\max}(\Xi_2)}{8} \dfrac{\tau_{\max}}{1 + l^{\mathrm{T}} l \sum_{m=1}^{r} \left|W_m^{\mathrm{T}} P_m W \lambda_{\max}(V_m(\bullet))\right|_{\tau_m} \tau_m}}, & \text{якщо } \tau_{\max} \le \rho d^{1/2}, \\ d_0, & \text{якщо } \tau_{\max} > \rho d^{1/2}, \end{cases}$$

де d_0 — єдиний додатний розв’язок рівняння

$$\frac{\lambda_{\max}(\Xi_2)}{8} \frac{\rho}{1 + l^{\mathrm{T}} l \sum\limits_{m=1}^{r} \left|\lambda_{\max}\left(W_m^{\mathrm{T}} P_m W_m\right)\right| \tau_m} d^{3/2} - d + 1 = 0. \qquad (11)$$

У формулі (11)

$$\rho := \frac{1}{\left(1 + l^{\mathrm{T}} l \sum\limits_{m=1}^{r} \lambda_{\max}\left(W_m^{\mathrm{T}} P_m W_m\right) \tau_m\right)^{1/2}} \frac{1}{\|A\| + \|l\| \sum_{m=1}^{r} \|W_m\|}.$$

Доведення випливає з підставляння $W_m^{\mathrm{T}} P_m W_m$ замість V_m.

Теорема 3. Припустімо, що виконуються припущення теореми 2.

Тоді існують константи $\lambda > 0$ і $k > 1$ такі, що $\upsilon(t) \le \upsilon(0) k e^{-\lambda t}$, $t \ge 2\tau_{\max}$.

Доведення. Нехай λ і k мають вигляд

$$0 < \lambda = -\frac{\log d}{2\tau_{\max}}, \; 1 < k = \frac{1}{d}.$$

Припустімо, що $t \ge 2\tau_{\max}$ — довільний момент часу з інтервалу $t \in [2i\tau_{\max}, 2(i+1)\tau_{\max})$. Тоді

$$\upsilon(t) \le \upsilon(t - 2\tau_{\max})d \le \upsilon(t - 4\tau_{\max})d^2 \le \ldots \le \upsilon(t - 2i\tau_{\max})d^i \le$$

$$\le \upsilon(0)\frac{1}{d}d^{i+1} \le \upsilon(0)ke^{-\lambda 2\tau_{\max}(i+1)} \le \upsilon(0)ke^{-\lambda t}.$$

Теорема 4. Нехай виконуються припущення теореми 2.

Тоді існують константи $k > 1$ і $\lambda > 0$ такі, що

$$|x_t|_{\tau_{\max}} \le (\upsilon(0)k)^{1/2} e^{(\lambda/2)\tau_{\max}} e^{-(\lambda/2)t}, \quad t \ge 3\tau_{\max}.$$

Доведення випливає безпосередньо з нерівностей $\|x(t)\|^2 \le \upsilon(t)$, $(|x_t|_{\tau_{\max}})^2 \le \upsilon(t - \tau_{\max})$ і теореми 3.

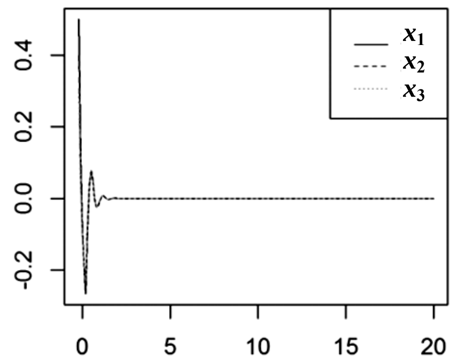
Приклад. Розглядається рекурентна нейронна мережа з трьома нейронами, що використовувалася як приклад у праці [23]. При цьому

$$A = \begin{pmatrix} 6 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 6 \end{pmatrix}; \; W_1 = \begin{pmatrix} -3 & -1 & -1 \\ -1 & -3 & -1 \\ -1 & -1 & -3 \end{pmatrix}; \qquad (12)$$

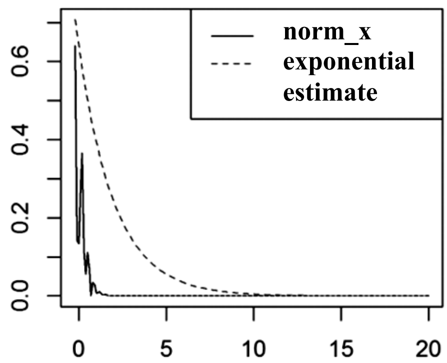
$$g_1(x) = g_2(x) = g_3(x) = \tanh(x), \quad x \in R^3, \; \tau_1 = 0{,}2.$$

Застосуванням результату теореми 4 встановлено значення експоненційного згасання $\lambda = 0{,}12$.

Траєкторії розв'язків показано на рисунку, *а*, а експоненціальна оцінка розв'язків — на рисунку, *б*. Із рисунка, *а* бачимо, що тривіальний розв'язок системи є стійким фокусом.



а



б

Траєкторії розв'язків (*а*) та експоненціальна оцінка розв'язків (*б*) системи диференціальних рівнянь із запізненням, заданої параметрами (12), що використано як модель рекурентної нейронної мережі з трьома нейронами [23]

ВИСНОВКИ

У роботі досліджено модель рекурентної нейронної мережі у вигляді системи диференціальних рівнянь із запізненням. Побудовано експоненціальну оцінку розв'язку диференціальних рівнянь зі швидкістю згасання, що залежить від часу запізнення.

З'ясовано, що традиційні методи побудови експоненціальних оцінок розв'язків систем диференціальних рівнянь із запізненням грунтуються на диференціальних або диференціально-різницевих нерівностях, що включають функції (функціонали) Ляпунова і полягають в аналізі відповідних лінійних матричних нерівностей. Водночас метод на основі різницевої нерівності для функціоналів Ляпунова дає змогу отримати оцінки експоненціального згасання як розв'язок нелінійного алгебричного рівняння.

У подальших дослідженнях отримані результати можна використати для вирішення проблем оптимізації рекурентних нейронних мереж у практичних застосуваннях.

ЛІТЕРАТУРА

1. *Amato Filippo.* Artificial neural networks in medical diagnosis / Filippo Amato, Alberto López, Eladia María Peña-Méndez et al. // Journal of Applied Biomedicine. — 2013. — Vol. 11, Issue 2. — P. 47–58. — https://doi.org/10.2478/v10136-012-0031-x.
2. *Jančíková Z.K.* Review on Artificial Intelligence Applications in Material Diagnostics and Technology / Z.K. Jančíková, P. Koštial, M. Heger et al. — (2018) MATEC Web of Conferences, 210, art. no. 04030. DOI: 10.1051/matecconf/201821004030.
3. *Pliego MarugĂˇn Alberto.* A survey of artificial neural network in wind energy systems / Alberto Pliego MarugĂˇn, Fausto Pedro GarcĂ-a MĂˇrquez, Jesus MarĂ-a Pinar Perez, Diego Ruiz-HernĂˇndez // Applied Energy. — 2018. — Vol. 228. — P. 1822–1836. — https://doi.org/10.1016/j.apenergy.2018.07.084.
4. *Sieniutycz Stanislaw.* A Review of Applications Optimizing Thermal, Chemical, and Environmental Systems / Stanislaw Sieniutycz, Zbigniew Szwast // Elsevier. — 2018. — P. 109–120. — https://doi.org/10.1016/B978-0-12-813582-2.00004-5.
5. *Qin C.* Computer-aided detection in chest radiography based on artificial intelligence: A survey / C. Qin, D. Yao, Y. Shi, Z. Song // BioMedical Engineering Online. — 2018. — **17** (1), art. no. 113. — DOI: 10.1186/s12938-018-0544-y
6. *Rajendra Acharya U.* Deep convolutional neural network for the automated detection and diagnosis of seizure using EEG signals / U. Rajendra Acharya, Shu Lih Oh, Yuki Hagiwara et al. // Computers in Biology and Medicine. — 2018. — Vol. 100. — P. 270–278. — https://doi.org/10.1016/j.compbiomed.2017.09.017.
7. *Deng Hongli.* Feature memory-based deep recurrent neural network for language modeling / Hongli Deng, Lei Zhang, Xin Shu // Applied Soft Computing. — 2018. — Vol. 68. — P. 432–446. — https://doi.org/10.1016/j.asoc.2018.03.040.
8. *Plappert Matthias.* Learning a bidirectional mapping between human whole-body motion and natural language using deep recurrent neural networks / Matthias Plappert, Christian Mandery, Tamim Asfour // Robotics and Autonomous Systems. — 2018. — Vol. 109. — P. 13–26. — https://doi.org/10.1016/j. robot.2018.07.006.
9. *Park J.H.* On global stability criterion for neural networks with discrete and distributed delays / J.H. Park // Chaos, Solitons & Fractals. — 2006. — Vol. 30, N 4. — P. 897–902. — Available at: http://dx.doi.org/10.1016/j.chaos.2005.08.147.
10. *Park J.H.* A delay-dependent asymptotic stability criterion of cellular neural networks with time-varying discrete and distributed delays / J.H. Park, H.J. Cho // Chaos, Solitons & Fractals. — 2007. — Vol. 33, N 2. — P. 436–442. — Available at: http://dx. doi.org/10.1016/j.chaos.2006.01.015.

11. *Liao X.* Delay-dependent exponential stability analysis of delayed neural networks: An LMI approach / X. Liao, G. Chen, E.N. Sanchez // Neural Networks. — 2002. — Vol. 15, N 7. — P. 855–866. — Available at: http://dx.doi.org/10.1016/S0893-6080(02)00041-2.
12. *Haykin Simon.* Neural Networks and Learning Machines: A Comprehensive Foundation (3rd Edition) / Simon Haykin. — 2011. — 936 p.
13. *Marcus C.M.* Stability of analog neural networks with delay / C.M. Marcus, R.M. Westervelt // Physical Review A. — 1989, 39.1: 347.
14. *He Y.* Delay-range-dependent stability for systems with time-varying delay / Y. He, Q.G. Wang, C. Lin, M. Wu // Automatica. — 2007. — Vol. 43, N 2. — P. 371–376. — Available at: http://dx.doi.org/10.1016/j.automatica.2006.08.015.
15. *Lien C.-H.* Global asymptotic stability for cellular neural networks with discrete and distributed time-varying delays / C.-H. Lien, L.-Y. Chung // Chaos, Solitons & Fractals. — 2007. — Vol. 34, N 4. — P. 1213–1219. — Available at: http://dx.doi.org/ 10.1016/j.chaos.2006.03.121.
16. *Zhang Q.* Stability of delayed cellular neural networks / Q. Zhang, X. Wei, J. Xu // Chaos, Solitons & Fractals. — 2007. — Vol. 31, N 2. — P. 514–520. — Available at: http://dx. doi.org/10.1016/j.chaos.2005.10.003.
17. *Singh V.* New global robust stability results for delayed cellular neural networks based on norm-bounded uncertainties / V. Singh // Chaos, Solitons & Fractals. — 2006. — Vol. 30, N 5. — P. 1165–1171. — Available at: http://dx.doi.org/10.1016/j.chaos.2005.08.183.
18. *Martsenyuk V.* On an indirect method of exponential estimation for a neural network model with discretely distributed delays / V. Martsenyuk // Electronic Journal of Qualitative Theory of Differential Equations. — 2017. — N 23. — P. 1–16.
19. *Indirect* method of exponential convergence estimation for neural network with discrete and distributed delays // Electronic Journal of Differential Equations. — 2017.
20. *Khusainov D.* Two-side estimates of solutions of linear systems with delay / D. Khusainov, V. Marzeniuk // Reports of Ukr. Nat. Acad. Sciences. — 1996. — P. 8–13.
21. *Kertesz V.* Stability investigations and exponential estimations for functional differential equations of retarded type / V. Kertesz // Acta Mathematica Hungarica. — 1990. — Vol. 55, N 3–4. — P. 365–378.
22. *Hale J.K.* Introduction to functional differential equations / J.K. Hale, S.M.V. Lunel // Springer Science & Business Media. — 2013. — Vol. 99.
23. *Cao J.* Absolute exponential stability of recurrent neural networks with Lipschitz-continuous activation functions and time delays / J. Cao, J. Wang // Neural Netw. — 2004. — **17**, N 3. — P. 379–390.

Надійшла 07.02.2019