УДК 517.96

УСТОЙЧИВОСТЬ ДВУХСЛОЙНЫХ РЕКУРСИВНЫХ НЕЙРОННЫХ СЕТЕЙ[1]

С.А. Иванов[2]

Получены численные критерии устойчивости двухслойных дискретных нейронных сетей. Построены области устойчивости в пространстве параметров для таких сетей. Задача сводится к проблеме устойчивости матричных разностных уравнений высоких порядков с запаздыванием. Основным средством решения проблемы являются конусы устойчивости.

Ключевые слова: нейронные сети, разностные матричные уравнения, устойчивость разностных уравнений, двухслойные сети.

Введение

В статье рассмотрены двухслойные нейронные сети с одинаковыми запаздываниями во взаимодействии между нейронами в сети. Такие модели имеют широкое применение в различных областях знаний.

Связи двухслойной сети с тремя нейронами в каждом слое изображены на рис. 1.

В результате линеаризации вокруг стационарного решения уравнений двухслойной нейронной сети получается линейное матричное разностное уравнение

$$x_s = \gamma I x_{s-1} + B x_{s-k},\ s = 1,2\ldots, \qquad (1)$$

где x_s – вектор сигналов нейронов в момент s. Вектор x_s размерности $2n$ характеризует отклонения сигналов нейронов от стационарных, I единичная $2n\times 2n$ матрица, $\gamma (0\le\gamma<1)$ коэффициент затухания колебаний нейронов, B – матрица размера $2n\times 2n$, характеризующая взаимодействия между нейронами в сети, k – запаздывание во взаимодействии между нейронами, n – число нейронов в каждом слое.

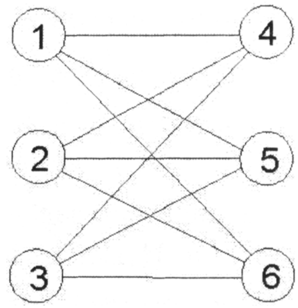

Рис. 1: Двухслойная нейронная сеть

Уравнение (1) принадлежит классу матричных разностных уравнений вида:

$$x_s = A x_{s-1} + B x_{s-k},\ s = 1,2\ldots \qquad (2)$$

которые обладают важным для нас свойством: матрицы A, B могут быть приведены к треугольному виду одним преобразованием. Поэтому мы имеем возможность применить метод конуса устойчивости [4] для устойчивости этих уравнений. На основе этого метода были изучены другие нейронные сети стандартных конфигураций [1, 2, 5]. Непрерывные модели исследованы в [3].

Матрица B, например, двухслойной сети, состоящей из шести нейронов, имеет следующий вид:

$$B = \begin{pmatrix} 0 & 0 & 0 & a & a & a \\ 0 & 0 & 0 & a & a & a \\ 0 & 0 & 0 & a & a & a \\ b & b & b & 0 & 0 & 0 \\ b & b & b & 0 & 0 & 0 \\ b & b & b & 0 & 0 & 0 \end{pmatrix}, \qquad (3)$$

где a – сила воздействия нейронов первого слоя на второй, b - сила обратного воздействия.

Мы ставим задачу изучить область устойчивости системы (1) в пространстве параметров a,b при разных значениях γ, n и k.

---

[1] Работа поддержана грантом Министерства образования и науки 1.1711.2011 и грантом для аспирантов Челябинского государственного педагогического университета.

[2] Иванов Сергей Александрович – аспирант, кафедра математического анализа, Челябинский государственный педагогический университет.

E-mail: ivanovlord@yandex.ru

Конус устойчивости для диагностирования устойчивости нейронных сетей

В работах [4, 6] введены конусы устойчивости для диагностирования устойчивости систем вида (2) с матрицами A, B, одновременно приводимыми к треугольному виду. Аналогичные конусы устойчивости для дифференциальных уравнений введены в [7]. Для решения задачи устойчивости двухслойных нейронных сетей нам понадобится техника конусов устойчивости, которую мы здесь изложим.

Определение 1. Конусом устойчивости для уравнения вида (2) для данного k мы называем множество точек $M=(u_1,u_2,u_3)\in R^3$, такое, что

$$u_1+iu_2=\exp(ik\omega)-h\exp(i(k-1)\omega), u_3=h, \tag{4}$$

где параметры h,ω связаны соотношениями

$$0\le h\le \frac{\sin k\omega}{\sin(k-1)\omega}, -\frac{\pi}{k}\le\omega\le\frac{\pi}{k}. \tag{5}$$

Теорема 1 [4]. Пусть $A,B,S\in R^{2n\times 2n}$ и $S^{-1}AS=A_l, S^{-1}BS=B_T$, где A_l, B_T – треугольные матрицы с диагональными элементами λ_j, μ_j соответственно $(1\le j\le 2n)$. Построим точки $M_j=(u_{1j},u_{2j},u_{3j})\in R^3$ $(1\le j\le 2n)$, так что

$$u_{1j}+iu_{2j}=\mu_j\exp(-ik\arg\lambda_j),\ u_{3j}=\left|\lambda_j\right|. \tag{6}$$

Тогда уравнение (2) асимптотически устойчиво, если и только если все точки M_j лежат внутри конуса устойчивости (4), (6) для данного k. Если некоторая точка M_j лежит вне конуса устойчивости, то уравнение (2) неустойчиво.

Теорема 1 сводит задачу диагностирования устойчивости системы (2) порядка $(2n\times 2n)$ к геометрической задаче в R^3: асимптотическая устойчивость системы равносильна условию, что все точки $M_j(1\le j\le 2n)$ лежат внутри конуса устойчивости (4), (6) для данного k.

Для применения теории конусов устойчивости необходимо знать собственные числа матрицы B. Для матрицы B порядка $2n$ собственные числа равны $\mu_1=\mu_2=\ldots=0$; $\mu_{2n-1}=-n\sqrt{ab}$, $\mu_{2n}=n\sqrt{ab}$.

Диагностирование устойчивости двухслойной сети

Определение 2. Овалом устойчивости для уравнений вида (2) для запаздывания $k>1$ и параметра γ мы называем кривую $M(\omega)=(u_1(\omega),u_2(\omega))$, такую, что

$$u_1(\omega)+iu_2(\omega)=\exp(ik\omega)-|\gamma|\exp(i(k-1)\omega),$$

где $\omega\in(-\omega_1,\omega_1)$, где ω_1 есть наименьший положительный корень уравнения

$$|\gamma|=\frac{\sin k\omega}{\sin(k-1)\omega}.$$

Овал устойчивости для данного запаздывания k и данного γ – это сечение конуса устойчивости (см. Определение 1) плоскостью $u_3=|\gamma|$. На основании Теоремы 1 и свойств матрицы B для диагностирования устойчивости уравнения (1) достаточно проверить две точки $M(u_{1j},u_{2j})=u_{1j}+iu_{2j}=\pm n\sqrt{ab}$ $(1\le j\le 2)$. Поэтому имеют место следующие теоремы.

Теорема 2. Пусть даны произвольные $n,k\in Z_+, k>1$. Пусть $0\le\gamma\le 1$. Построим в R^2 овал устойчивости (см. Определение 2) для данных k,γ. Построим точки $M_j=(u_{1j},u_{2j})\in R^2$ $(1\le j\le 2)$ так, что

$$u_{1j}+iu_{2j}=\pm n\sqrt{ab}.$$

Если обе точки $M_j(1\le j\le 2)$ лежат внутри овала устойчивости, то система (1) асимптотически устойчива. В противном случае система (1) неустойчива.

Теорема 3.

1. Если $\gamma > 1$, то система (1) неустойчива.

2. Если $\gamma \le 1$ и $0 \le ab < \left(\frac{1-\gamma}{n}\right)^2$, то система (1) асимптотически устойчива при любом запаздывании k. Если $\gamma \le 1$ и $ab > \left(\frac{1-\gamma}{n}\right)^2$, то система (1) неустойчива при любом запаздывании k.

3. Если $\gamma \le 1$ и $ab < 0$ и $|ab| < \left(\frac{F(\gamma,k)}{n}\right)^2$, то система (1) асимптотически устойчива при данном значении k. Если $\gamma \le 1$ и $ab < 0$ и $|ab| > \left(\frac{F(\gamma,k)}{n}\right)^2$, то система неустойчива при данном запаздывании k. Здесь $F(\gamma,k) = \frac{\sin\omega(\gamma)}{\cos(k-1)\omega(\gamma)}$, где $\omega(\gamma)$ есть наименьший неотрицательный корень уравнения $\gamma = \frac{\cos k\omega}{\cos(k-1)\omega}$.

Области устойчивости системы (1) отражены на рис. 2, 3.

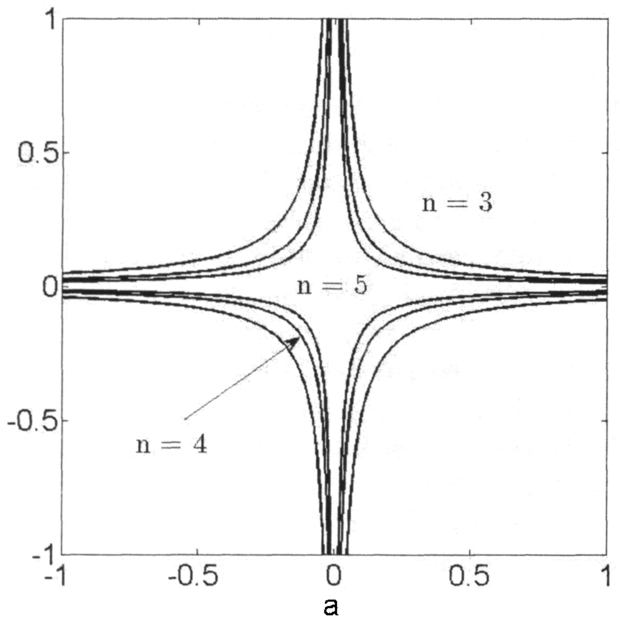

Рис. 2. Область устойчивости системы (1) в плоскости (a,b) при фиксированных $\gamma = 0{,}4, k = 3$ и переменном числе нейронов n

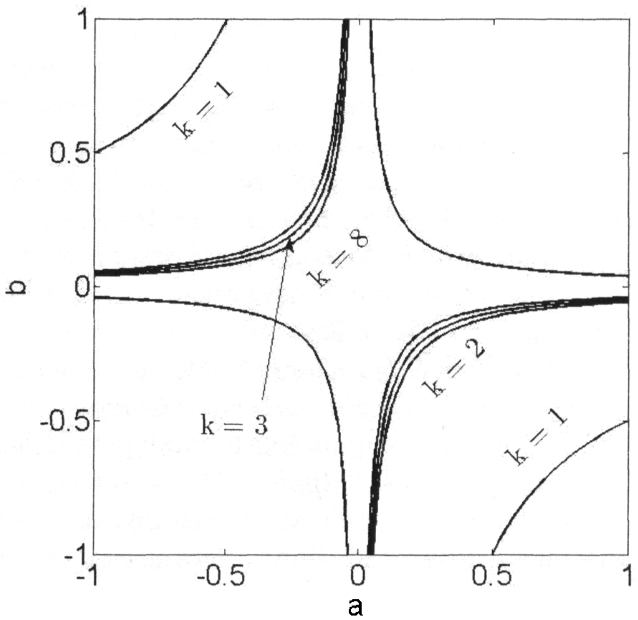

Рис. 3. Область устойчивости системы (1) в плоскости (a,b) при фиксированных $\gamma = 0{,}4, n = 3$ и переменном запаздывании k

Вывод о динамике областей устойчивости в пространстве параметров таков. С ростом числа нейронов в сети область устойчивости стягивается в крест. Но при фиксированном количестве нейронов $2n$ имеется область в пространстве параметров, в которой гарантируется устойчивость независимо от запаздывания (delay-independent stability).

Литература

1. Иванов, С.А. Область устойчивости в пространстве параметров рекурсивных нейронных сетей с топологией многомерного куба / С.А. Иванов // Вестник ЮУрГУ. Серия «Математика. Механика. Физика». – 2012. – Вып. 7. – №34(293). – С. 157–160.

2. Иванов, С.А. Устойчивость рекурсивных нейронных сетей со звездной топологией связей / С.А. Иванов // Естественные и технические науки. – 2012. – №6(62). – С. 21–25.

3. Khokhlova, T.N. Stability of a ring and linear neural networks with a large number of neurons / T.N. Khokhlova, M.M. Kipnis // Applied Mathematics and Computation. – 2012. – P. 1–14.

4. Ivanov, S.A. The stability cone for a difference matrix equation with two delays / S.A. Ivanov, M.M. Kipnis, V.V. Malygina // ISRN J. Applied Mathematics. – 2011. – P. 1–19. ID 910936.

5. Ivanov, S.A. Stability analysis of discrete-time neural networks with delayed interactions: torus, ring, grid, line / S.A. Ivanov, M.M. Kipnis // International Journal of Pure and Applied Math. – 2012. – Vol. 78, № 5. – P. 691–709.

6. Kipnis, M.M. The stability cone for a matrix delay difference equation / M.M. Kipnis, V.V. Malygina // International Journal of Mathematics and Mathematical Sciences. – 2011. – P. 1–15. ID 860326.

7. Khokhlova, T.N The stability cone for a delay differential matrix equation / T.N. Khokhlova, M.M. Kipnis, V.V. Malygina // Applied Math. Lett. – 2011. – Vol. 24. – P. 742–745.

STABILITY OF TWO-LAYER RECURSIVE NEURAL NETWORKS

S.A. Ivanov[1]

The stability conditions are described for the discrete neural networks. Stability regions are constructed in the parameter space for these networks. The problem reduces to the stability problem for the matrix difference equations of higher order with delay. The main tool is the stability cone.

Keywords: neural networks, difference matrix equations, stability of difference equations, two-layer network.

References

1. Ivanov S.A. Oblast' ustoychivosti v prostranstve parametrov rekursivnykh neyronnykh setey s topologiey mnogomernogo kuba (The stability domain in the parameters space of recursive neural networks with hypercube topology). *Vestnik YuUrGU. Seriya "Matematika. Mekhanika. Fizika".* 2012. Issue 7. no. 34(293). pp. 157–160. (in Russ.).

2. Ivanov S.A. Ustoychivost' rekursivnykh neyronnykh setey so zvezdnoy topologiey svyazey (Stability of recursive neural networks with star topology). *Estestvennye i tekhnicheskie nauki.* 2012. no. 6(62). pp. 21–25. (in Russ.).

3. Khokhlova T.N., Kipnis M.M. Stability of a ring and linear neural networks with a large number of neurons. *Applied Mathematics and Computation.* 2012. pp. 1–14.

4. Ivanov S.A., Kipnis M.M., Malygina V.V. The stability cone for a difference matrix equation with two delays. *ISRN J. Applied Mathematics.* 2011. pp. 1–19. ID 910936.

5. Ivanov S.A., Kipnis M.M. Stability analysis of discrete-time neural networks with delayed interactions: torus, ring, grid, line. *International Journal of Pure and Applied Math.* 2012. Vol. 78, no. 5. pp. 691–709.

6. Kipnis M.M., Malygina V.V. The stability cone for a matrix delay difference equation. *International Journal of Mathematics and Mathematical Sciences.* 2011. pp. 1–15. ID 860326

7. Khokhlova T.N., Kipnis M.M., Malygina V.V. The stability cone for a delay differential matrix equation. *Applied Math. Lett.* 2011. Vol. 24. pp. 742–745.

Поступила в редакцию 4 июня 2013 г.

[1] Ivanov Sergey Alexandrovich is Post-graduate Student, Mathematical Analysis Department, Chelyabinsk State Pedagogical University
E-mail ivanovlord@yandex.ru