ИНФОРМАТИКА

2006 январь-март № 1

МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ

УДК 519.5

А.В. Картынник, А.Д. Линкевич

ГЛОБАЛЬНЫЙ АТТРАКТОР И ДОСТАТОЧНЫЕ УСЛОВИЯ ДЛЯ ВОССТАНОВЛЕНИЯ ОБРАЗОВ НЕЙРОСЕТЯМИ КОЭНА – ГРОССБЕРГА

Аналоговые нейронные сети Коэна – Гроссберга изучаются в общем случае, когда межнейронные (синаптические) связи могут быть несимметричными и, следовательно, отсутствует аналог гамильтониана для системы в полном фазовом пространстве (глобальная функция Ляпунова). Доказывается, что при определенных условиях существует глобальная притягивающая область Q, такая, что все аттракторы лежат внутри Q и, кроме того, область Q является аттрактором системы. Находится верхняя оценка времени, необходимого для достижения предписанной окрестности области Q. Получаются достаточные условия, при выполнении которых нейросеть асимптотически сходится к ближайшему стационарному состоянию (запомненному образу) и достигает заданной окрестности аттрактора системы за определенное (конечное) время. Находится верхняя граница времени, необходимого для восстановления запомненного образа с заданной точностью.

Введение

Достигнутый в последние годы прогресс в изучении функционирования нейронных сетей как устройств автоассоциативной памяти в значительной степени обязан предположению, что межнейронные (синаптические) связи симметричны. Это свойство позволяет ввести аналог гамильтониана системы (глобальную функцию Ляпунова) и трактовать ее поведение как релаксацию в «энергетическом ландшафте» в направлении к ближайшему локальному минимуму, представляющему запомненный образ [1–3]. Тогда такая система может функционировать как автоассоциативная память, если матрица синаптических связей подобрана подходящим образом в течение процесса обучения. Однако правила обучения для аналоговых нейросетей дают, вообще говоря, несимметричные синаптические связи [4–8]. Кроме того, симметрия синаптических связей не является достаточно правдоподобной с точки зрения биологии: связи между реальными нервными клетками несимметричны (см., например, [9]). Для того чтобы расширить возможности аппаратурной реализации нейронных сетей, желательно также устранить условие симметричности.

Проблема, однако, состоит в том, что нейросети с несимметричными синаптическими связями могут демонстрировать сложные типы поведения, включая хаотические, даже если число степеней свободы равно 3 или 4 (см., например, [10, 11]). Тем не менее, для случая аналоговых нейросетей Хопфилда [3] два метода (метод локальных функций Ляпунова и метод локальных сжимающих отображений) были предложены для исследования асимптотического поведения систем и для получения условий, которые гарантируют сходимость нейросети к ближайшему стационарному состоянию [8]. В настоящей работе развиваются предыдущие результаты авторов, а именно модель Коэна – Гроссберга [2], которая переходит в модель Хопфилда [3] при специальном выборе функций $a_i(\cdot)$ и $b_i(\cdot)$ в уравнении (1). В разд. 1 показано, что все аттракторы нейросети Коэна – Гроссберга расположены внутри притягивающей области Q. Разд. 2 посвящен изучению запомненных образов, т. е. сходимости состояния нейросети к стационарному состоянию динамики.

1. Глобальная структура фазового пространства нейросети

Рассмотрим систему, состоящую из N нейроноподобных элементов, описываемых переменными состояния $u_1(t),...,u_N(t) \in \boldsymbol{R}$, которые эволюционируют во времени в соответствии со следующей системой дифференциальных уравнений [2]:

$$\dot{u}_i = -a_i\left(u_i\right)\left[b_i\left(u_i\right) + \sum_{j=1}^{N} T_{ij} f_j\left(u_j\right)\right], \quad i = \overline{1,N}\,. \tag{1}$$

Здесь $a_i\left(u_i\right)$ есть произвольные неотрицательные вещественные функции и функции $b_i\left(u_i\right)$ обычно выбираются так, чтобы существовала глобальная функция Ляпунова

$$L = -\frac{1}{2}\sum_{i,j=1}^{N} T_{ij} f_i\left(u_i\right) f_j\left(u_j\right) + \sum_{i=1}^{N}\int_0^{u_i} dz f_i^{/}\left(z\right) b_i\left(z\right),$$

если синаптическая матрица T_{ij} симметрична. Передаточные функции $f_i\left(u_i\right)$ имеют строго монотонную сигмоидную форму.

В отличие от работы [2] в настоящей статье рассматривается случай, когда синаптическая матрица T_{ij} может быть несимметричной. Удобно ввести обозначения

$$S_j = \sup_{u_j \in R}\left|f_j\left(u_j\right)\right| < \infty;$$

$$M_i = \sup_{\boldsymbol{u} \in R^N}\left(\sum_{j=1}^{N} T_{ij} f_j\left(u_j\right)\right) < \infty;$$

$$m_i = \inf_{\boldsymbol{u} \in R^N}\left(\sum_{j=1}^{N} T_{ij} f_j\left(u_j\right)\right) < -\infty,$$

где $\boldsymbol{u} = \left(u_1, \ldots, u_N\right)$.

Основные результаты видны из следующих трех теорем.

Теорема 1. *Предположим, что для каждой функции $b_i\left(u_i\right)$ существуют конечные величины α_i и β_i, такие, что $b_i\left(u_i\right)$ ограничена в некоторых малых окрестностях этих точек и выполняется одно из следующих двух условий:*

$$\begin{cases} b_i\left(u_i\right) < \min\left\{0; m_i\right\}, & \forall u_i \le \alpha_i, \\ b_i\left(u_i\right) > \max\left\{0; M_i\right\}, & \forall u_i \ge \beta_i; \end{cases} \tag{2а}$$

$$\begin{cases} b_i\left(u_i\right) > \max\left\{0; M_i\right\}, & \forall u_i \le \alpha_i, \\ b_i\left(u_i\right) < \min\left\{0; m_i\right\}, & \forall u_i \ge \beta_i. \end{cases} \tag{2б}$$

Тогда справедливы следующие утверждения:

1. Все аттракторы системы принадлежат N-мерному гиперпараллелепипеду $Q = \prod_{i=1}^{N}\left[\alpha_i; \beta_i\right]$.

2. Если для каждого $i = \overline{1,N}$ справедливы только условия (2а), то гиперпараллелепипед Q является аттрактором системы, т. е. любая фазовая траектория, начинающаяся где-либо вне области Q, сходится асимптотически со временем к границе области Q и все траектории, начинающиеся где-либо внутри гиперпараллелепипеда Q, никогда не покидают эту область.

3. Если для каждого $i = \overline{1,N}$ выполняются только условия (2б), то гиперпараллелепипед Q является репеллером системы, т. е. любая фазовая траектория, начинающаяся вне облас-

ти Q, движется прочь от области Q и любая траектория, начинающаяся внутри гиперпараллелепипеда Q, навсегда покидает эту область, если она достигает границы области Q.

Доказательство:

1. Пусть $\boldsymbol{\xi} \in R^N$ есть стационарное состояние системы, т. е.

$$a_i\left(\xi_i\right)\left(b_i\left(\xi_i\right)-\sum_j T_{ij} f_j\left(\xi_j\right)\right)=0, \quad i=\overline{1, N},$$

или

$$b_i\left(\xi_i\right)-\sum_j T_{ij} f_j\left(\xi_j\right)=0, \quad i=\overline{1, N}, \tag{3}$$

где учтено, что $a_i\left(\xi_i\right)>0$. Следовательно,

$$m_i=\inf\left(\sum_j T_{ij} f_j\left(u_j\right)\right) \leq b_i\left(\xi_r\right) \leq \sup\left(\sum_j T_{ij} f_j\left(u_j\right)\right)=M_i, \quad i=\overline{1, N}, \tag{4}$$

где инфимум и супреум определяются для всех $u_j \in R$. Если предположить (от противного), что $\boldsymbol{\xi} \notin Q$, т. е. $\xi_i<\alpha_i$ или $\xi_i<\beta_i$ для по крайней мере одного значения $i \in\{1, \ldots, N\}$, то в силу условия (2а) или (2б) одно из неравенств (4) не справедливо. Следовательно, $\boldsymbol{\xi} \in Q$.

2. Из уравнений (1) и (3) вытекает оценка

$$\dot{u}_i \geq-a_i\left(u_i\right)\left(b_i-\left(u_i-m_i\right)\right);$$

$$\dot{u}_i \geq-a_i\left(u_i\right)\left(b_i-\left(u_i-M_i\right)\right), \quad i=\overline{1, N}. \tag{5}$$

Предположим, что условие (2а) удовлетворяется и $u_i(t)<\alpha_i$ для некоторого значения i в некоторый момент времени $t = t_1$. Тогда $u_i(t)<0$ и, следовательно, $u_i(t)$ убывает и траектория сходится к границе $u_i=\beta_i$ области Q. Наоборот, если условие (2б) справедливо (вместо (2а)), то в случае, когда $u_i(t)<\alpha_i$, имеем $u_i(t)<0$ и фазовая траектория движется прочь от ребра $u_i=\alpha_i$. Если неравенство $u_i(t)>\beta_i$ удовлетворяется в некоторый момент $t = t_1$, то имеем $u_i(t)>0$ для всех $t \geq t_1$ и траектория уходит прочь от гиперпараллелепипеда Q.

Следовательно, если условие (2а) удовлетворяется для всех $i=\overline{1, N}$, то расстояние $d(\boldsymbol{u}(t), Q)=\inf\limits_{\boldsymbol{q} \in Q}\|\boldsymbol{u}(t)-\boldsymbol{q}\|$ от текущего положения $\boldsymbol{u}(t)$ изображающей точки в фазовом пространстве системы до гиперпараллелепипеда Q стремится к нулю при $t \rightarrow \infty$. Это означает, что область Q является аттрактором нейросети и ее область притяжения совпадает с полным пространством вокруг Q. Наоборот, если условие (2б) выполняется вместо (2а) для всех $i=\overline{1, N}$, то расстояние $d(\boldsymbol{u}(t), Q) \rightarrow \infty$ при $t \rightarrow \infty$, т. е. гиперпараллелепипед Q представляет собой репеллер системы.

2. Восстановление запомненных образов нейросетью

Перейдем к изучению поведения системы внутри гиперпараллелепипеда Q, т. е. предположим, что начальное состояние системы $\boldsymbol{u}\,(t_0)$ принадлежит области Q. Для того чтобы нейросеть могла трактоваться как автоассоциативная память, запоминаемые образы должны быть стационарными состояниями динамики нейросети. Для этого синаптическая матрица T_{ij} должна быть подобрана так, чтобы выполнялись условия

$$b_i\left(\xi_i^\mu\right)-\sum_i T_{ij} f_j\left(\xi_j^\mu\right)=0, \quad i=\overline{1,N}\,, \quad \mu=\overline{1,p}, \tag{6а}$$

или

$$\sum_j T_{ij} y_j^\mu = z_i^\mu, \quad i=\overline{1,N} \quad \mu=\overline{1,p}. \tag{6б}$$

Здесь $\boldsymbol{\xi}^\mu=\left(\xi_1^\mu,\ldots,\xi_N^\mu\right)$ есть состояние нейросети, кодирующее μ-й запоминаемый образ, и $y_i^\mu=f_i\left(\xi_i^\mu\right), \quad z_i^\mu=b_i\left(\xi_i^\mu\right)$. Указанная цель достигается посредством какого-либо алгоритма обучения.

Предположим, что нейросеть (1) обучена так, что она обладает множеством заданных стационарных состояний $\xi^\mu, \quad \mu=\overline{1,p}$. Будем рассматривать динамику системы в области вокруг одного стационарного состояния, скажем $\boldsymbol{\xi}^\mu$, и обозначим через $\boldsymbol{x}(t)=\boldsymbol{u}(t)-\boldsymbol{\xi}^\mu$ отклонение состояния нейросети $\boldsymbol{\xi}^\mu$ и через $S\left(\boldsymbol{\xi}^\mu,r\right)=\left\{\boldsymbol{u}\in R^N:\left\|\boldsymbol{u}-\boldsymbol{\xi}^\mu\right\|<r\right\}$ открытую сферическую область с центром в точке $\boldsymbol{\xi}^\mu$ и радиусом *r*. Символ $\|\cdot\|$ означает обычную евклидову норму вектора, т. е. $\|\boldsymbol{x}\|=\left(\sum_{i=1}^N x_i^2\right)^{1/2}$ для произвольного *N*-мерного вектора $\boldsymbol{x}$.

Теорема 2. *1. Пусть существует открытая сферическая область* $S\left(\boldsymbol{\xi}^\mu,r\right)$ *с Q вокруг стационарного состояния* $\boldsymbol{\xi}^\mu$*, такая, что неравенство*

$$\sum_{i=1}^N x_i a_i\left(u_i\right)\left[b_i\left(u_i\right)-b_i\left(\xi_i^\mu\right)-\sum_j T_{ij}\left(f_j\left(u_j\right)-f_j\left(\xi_i^\mu\right)\right)\right]>0 \tag{7}$$

справедливо для всех $\boldsymbol{u}\in S\left(\boldsymbol{\xi}^\mu,r\right)$. *Тогда эта точка* $\boldsymbol{\xi}^\mu$ *асимптотически устойчива и* $\boldsymbol{u}\to\boldsymbol{\xi}^\mu$ *при* $t\to\infty$, *если начальное состояние* $\boldsymbol{u}(0)\in S\left(\boldsymbol{\xi}^\mu,r\right)$.

2. Предположим, что вместо неравенства (7) более сильное условие

$$\sum_{i=1}^N x_i a_i\left(u_i\right)\left[\delta\left(b_i\left(u_i\right)-b_i\left(\xi_i^\mu\right)\right)-\sum_j T_{ij}\left(f_j\left(u_j\right)-f_j\left(\xi_j^\mu\right)\right)\right]>0 \tag{8}$$

выполняется с некоторой положительной константой $\delta<1$ *для всех* $\boldsymbol{u}\in S\left(\boldsymbol{\xi}^\mu,r\right)$ *и каждая функция* $b_i\left(u_i\right)$, $i=\overline{1,N}$, *является дифференцируемой на интервале* $R_i^\mu=\left[\xi_i^\mu-r,\xi_i^\mu+r\right]$, *а ее производная положительно определена, т. е.*

$$B_i=\inf_{u_i\in R_i^\mu} b_i^{/}\left(u_i\right)>0, \quad i=\overline{1,N}\,. \tag{9}$$

Тогда если начальное состояние нейросети $\boldsymbol{u}(0)\in S\left(\boldsymbol{\xi}^\mu,r\right)$, *то состояние* $\boldsymbol{u}(t)$ *достигает произвольно малую* ε*-окрестность* $S\left(\boldsymbol{\xi}^\mu,\varepsilon\right)$ *стационарной точки* $\boldsymbol{\xi}^\mu$ *после конечного времени* $t\le t^*$, *где*

$$t^* = \left[(1-\delta) \min_i B_i \min_{i, u_i \in R_i^{\mu}} a_i(u_i) \right]^{-1} \ln(r/\varepsilon). \tag{10}$$

Доказательство. Для каждого $i = \overline{1, N}$ умножим уравнение (6) на $a_i(u_i)$ и вычтем результат из уравнения (1). В результате получим систему дифференциальных уравнений

$$\dot{x}_i = -a_i(u_i)\left[b_i(u_i) - b_i(\xi_i^{\mu}) - \sum_j T_{ij}\left(f_j(u_j) - f_j(\xi_j^{\mu}) \right) \right], \quad i = \overline{1, N}, \tag{11}$$

где $u_i = x_i + \xi_i^{\mu}$.

1. Рассмотрим функцию

$$L(t) = \left\| \boldsymbol{x}(t) \right\|^2 = \sum_{i=1}^{N} x_i^2(t), \tag{12}$$

где $\boldsymbol{x} \in S(\boldsymbol{0}, r)$ и компоненты $x_1, \ldots, x_N$, эволюционируют во времени в соответствии с уравнениями (11). Очевидно, что $L(t) \geq 0$ тогда и только тогда, когда $\boldsymbol{x} = \boldsymbol{0}$. Кроме того, при выполнении условия (7) временная производная

$$\frac{dL}{dt} = 2\sum_{i=1}^{N} x_i \dot{x}_i = -2\sum_{i=1}^{N}\left[b_i(u_i) - b_i(\xi_i^{\mu}) - \sum_j T_{ij}\left(f_j(\xi_i^{\mu}) \right) \right] x_i \tag{13}$$

является отрицательной для $\forall \boldsymbol{x} \in S(\boldsymbol{0}, r)$. Таким образом, $L(t)$ есть функция Ляпунова системы (11) в сферической области $S(\boldsymbol{0}, r)$. Следовательно, нулевое решение $\boldsymbol{x} \equiv \boldsymbol{u} - \boldsymbol{\xi}^{\mu} = 0$ асимптотически устойчиво, т. е. $\boldsymbol{x}(t) \to 0$ при $t \to \infty$ для любого $\boldsymbol{x}(0)$ из шара $S(\boldsymbol{0}, r)$ или, иначе говоря, $\boldsymbol{u}(t) \to \boldsymbol{\xi}^{\mu}$ при $t \to \infty$ для любого начального состояния $\boldsymbol{u}(0)$ из сферической окрестности $S(\boldsymbol{\xi}^{\mu}, r)$ стационарного состояния $\boldsymbol{\xi}^{\mu}$.

2. Для каждой функции $b_i(u_i)$, рассматриваемой на интервале, целиком расположенном внутри области $R_i^{\mu} = \left[\xi_i^{\mu} - r, \xi_i^{\mu} + r \right]$, справедливы условия теоремы Лагранжа о конечных приращениях (см., например, работу [12]). В соответствии с этим утверждением существует точка η_i, такая, что

$$b_i(u_i) - b_i(\xi_i^{\mu}) = b_i^{/}(\eta_i)(u_i - \xi_i^{\mu}) = b_i^{/}(\eta_i) x_i, \tag{14}$$

где $\eta_i \in \left] \xi_i^{\mu}, u_i \right[$, если $u_i > \xi_i^{\mu}$, и $\eta_i \in \left] u_i, \xi_i^{\mu} \right[$, если $u_i < \xi_i^{\mu}$. (Заметим, что в обоих случаях $\eta_i \in R_i^{\mu}$, если $\boldsymbol{u} \in S(\boldsymbol{\xi}^{\mu}, r)$.) Из уравнений (13), (14) при выполнении условий (8), (9) получаем

$$\begin{gathered} \frac{dL}{dt} = -2\sum_{i=1}^{N} a_i(u_i)(1-\delta)\left(b_i(u_i) - b_i(\xi^{\mu}) \right) x_i = \\ = -2(1-\delta)\sum_{i=1}^{N} a_i(u_i) b_i^{/}(\eta_i) x_i^2 \leq -2k\sum_{i=1}^{N} x_i^2 = -2kL, \end{gathered} \tag{15}$$

где $$k = (1-\delta)\min_i \left[B_i \min_{i, u_i \in R_i^{\mu}} a_i(u_i) \right].$$

Поскольку $L(t) \neq 0$ для любого ненулевого $\boldsymbol{x}$, то из (15) следует, что

$$dL/L \leq -2kdt, \quad \forall \boldsymbol{x} \in S(\boldsymbol{0}, r), \quad \boldsymbol{x} \neq 0. \tag{16}$$

Взяв начальное состояние $\boldsymbol{x}(0)$ системы (11) внутри шара $S(\boldsymbol{0}, r)$, проинтегрируем последнее неравенство по переменной t на произвольном конечном интервале $[0, t]$. Получаем

$$L(t) \leq L(0)\exp(-2kt), \quad \forall t \in [0, \infty[. \tag{17}$$

В силу определения функции $L(t)$

$$\|\boldsymbol{x}(t)\| \leq \|\boldsymbol{x}(0)\| \exp(-kt) \leq r \exp(-kt), \quad \forall t \in [0, \infty[. \tag{18}$$

(Заметим, что эта оценка может быть получена путем использования теоремы сравнения Чаплыгина (см., например, работы [13, 14]).) Легко видеть, что состояние нейросети $\boldsymbol{u}(t)$ достигает произвольную заданную ε-окрестность стационарного состояния $\boldsymbol{u}^*$ (т. е. $\|\boldsymbol{x}(t)\| = \|\boldsymbol{u}(t) - \boldsymbol{u}^*\| \leq \varepsilon$) после времени $t \leq t^*$, где

$$t^* = \frac{1}{k}\ln\left(\frac{r}{\varepsilon}\right) = \left[(1-\delta)\min_i B_i \min_{i, u_i \in R_i^\mu} a_i(u_i)\right]^{-1} \ln\left(\frac{r}{\varepsilon}\right).$$

В первом утверждении теоремы говорится, что состояние нейросети $\boldsymbol{u}(t)$ приближается к стационарному состоянию $\boldsymbol{u}^*$, если начальное состояние $\boldsymbol{u}(0)$ достаточно близко к $\boldsymbol{u}^*$ (более точно, если оно лежит в шаре $S(\boldsymbol{u}^*, r)$ с некоторым радиусом r, где неравенство (7) выполняется). Однако функционирование нейросети как автоассоциативной памяти не сводится к такому свойству асимптотической сходимости состояния нейросети к стационарному состоянию, кодирующему запомненный образ. Действительно, первое утверждение теоремы гарантирует только совпадение состояния нейросети $\boldsymbol{u}(t)$ со стационарным состоянием $\boldsymbol{u}^*$ после неограниченно большого интервала времени. Очевидно, что это неприемлемо с практической точки зрения. С другой стороны, любая обработка информации является приближенной. Следовательно, точность восстановления хранимой информации задается заранее и нейросеть должна функционировать так, чтобы достичь предписанную точность в ограниченное время (чем меньше это время, тем лучше). Под точностью восстановления ε понимается максимальное расстояние от состояния нейросети $\boldsymbol{u}(t)$ (для любого момента времени t после некоторой величины t_r, трактуемой как время восстановления) до стационарного состояния $\boldsymbol{u}^*$, т. е.

$$\varepsilon = \max_{t \geq t_r} \|\boldsymbol{u}(t) - \boldsymbol{u}^*\|.$$

Условие первой части рассматриваемой теоремы не запрещает ситуации, когда время t_r, необходимое для восстановления образа $\boldsymbol{u}^*$ с предписанной точностью ε, является очень большим. Это нежелательное свойство исключается наложением более сильного условия (8) вместо (7) и дополнительного условия на функции $b_i(u_i)$ (см. (9)). При соблюдении указанных условий восстановление хранимой информации с произвольной заданной (конечной) точностью ε всегда выполняется после конечного времени и из уравнения (10) находится верхняя граница этого времени, т. е. $t_r \geq t^*$.

Теорема 2 дает достаточные условия, которые гарантируют, что области некоторых конечных размеров существуют вокруг стационарных состояний, где имеет место сходимость состояния нейросети к стационарным состояниям. Это означает способность нейросети восстанавливать хранимую информацию. Следующая задача состоит в том, что необходимо проанализировать условия теоремы. Требования к функции $b_i\left(u_i\right)$ достаточно ясные и не очень ограничительные. Так, могут быть использованы любые возрастающие дифференцируемые функции. С другой стороны, трудно проверить аналитически, удовлетворяются ли условия (7) или (8) для всех $\boldsymbol{u} \in S\left(\boldsymbol{\xi},r\right)$. По этой причине имеет смысл нахождение более сильных и более легко проверяемых условий, что достигается утверждением следующей теоремы.

Теорема 3. *Условия второй части теоремы 2 удовлетворяются в некоторой открытой сферической области* $S\left(\boldsymbol{\xi},r\right)$, *если каждая функция* $f_i\left(u_i\right)$, $i=\overline{1,N}$, *рассматриваемая на интервале* $R_i^{\mu}=\left[\xi_i^{\mu}-r,\xi_i^{\mu}+r\right]$, *имеет ограниченную производную* $f_i^{/}\left(u_i\right)$, *а коэффициенты* $T_{ij}, i,j=\overline{1,N}$, *подчиняются неравенству*

$$\max_{i,j}\left|T_{ij}\right| \le \delta \min_i\left[B_i \cdot \min_{u_i\in R_i^{\mu}} a_i\left(u_i\right)\right] \cdot \left\{N\left[\max_{u_i\in R_i^{\mu}} a_i\left(u_i\right)\cdot \max_{z_i\in R_i^{\mu}}\left|f_i^{/}\left(z_i\right)\right|\right]\right\}^{-1}. \tag{19}$$

Доказательство. 1. Для начала докажем следующее вспомогательное утверждение: если в некоторый момент времени t состояние системы $\boldsymbol{x}\left(t\right)\in S\left(\boldsymbol{0},r\right)$, то существует величина $\Delta t>0$, такая, что решение $\boldsymbol{x}\left(t\right)$ уравнения (11) лежит внутри шара $S\left(\boldsymbol{0},r\right)$ для всех моментов из интервала $\left[t,t+\Delta t\right]$.

Действительно, если $\boldsymbol{x}\left(t\right)$ принадлежит открытому шару $S\left(\boldsymbol{0},r\right)$, то существует такой малый шар $S\left(\boldsymbol{x}\left(t\right),\Delta r\right)$ с центром в точке $\boldsymbol{x}\left(t\right)$ и радиусом Δr, что $S\left(\boldsymbol{x}\left(t\right),\Delta r\right)\subset S\left(\boldsymbol{0},r\right)$. Кроме того, производные $\dot{x}_i\left(t\right)$ ограничены в силу условий, наложенных на функции $a_i\left(u_i\right)$, $b_i\left(u_i\right)$, $f_i\left(u_i\right)$:

$$\left|\dot{x}_i\left(t\right)\right| \le a_i\left(u_i\right)\left|b_i\left(u_i\right)-b_i\left(\xi_j^{\mu}\right)-\sum_j T_{ij}\left(f_j\left(u_j\right)-f_j\left(\xi_j^{\mu}\right)\right)\right| \le$$
$$\le 2\max_{u_i\in R_i^{\mu}} a_i\left(u_i\right)\left(\max_{u_i\in R_i^{\mu}} a_i\left(u_i\right)+\max\left\{\left|i\right|,\left|m_i\right|\right\}\right)<\infty .$$

В силу определения $\varepsilon-\delta$ производной функции для данного Δr заключаем, что существует величина $\Delta t>0$, такая, что $\left\|\boldsymbol{x}\left(t^{/}\right)-\boldsymbol{x}\left(t\right)\right\|\le\Delta r$ для всех $t^{/}\in\left[t,t+\Delta t\right]$. Это означает, что решение $\boldsymbol{x}\left(t\right)$, рассмотренное на временном интервале $\left[t+\Delta t\right]$, принадлежит шару $S\left(\boldsymbol{x}\left(t\right),\Delta r\right)$, который содержится в шаре $S\left(\boldsymbol{0},r\right)$.

2. Следующее промежуточное утверждение состоит в том, что если в некоторый момент времени t_0 (трактуемый как начальный момент) состояние нейросети $\boldsymbol{u}\left(t_0\right)\in S\left(\boldsymbol{\xi},r\right)$ (или $\boldsymbol{x}\left(t_0\right)\in S\left(\boldsymbol{0},r\right)$), то при выполнении условия (19) неравенство (8) справедливо для всех моментов времени из интервала $\left[t_0,t_0+\Delta t\right]$, где величина Δt определяется так, как описано выше в первой части доказательства. Действительно, для каждой функции $f_j\left(u_j\right)$, рассмотренной на любом интервале, который расположен целиком внутри области $R_i^{\mu}=\left[\xi_i^{\mu}-r,\xi_i^{\mu}+r\right]$, выпол-

няются условия теоремы Лагранжа о конечных приращениях (см., например, [12]). Следовательно, существует точка z_j, такая, что

$$f_j(u_j) - f_j(\xi_j^\mu) = f_j^/(z_j)(u_j - \xi_j^\mu) = f_j^/(z_j)x_j, \tag{20}$$

где $z_i \in]\xi_j^\mu, u_j[$, если $u_j > \xi_j^\mu$, и $z_j \in]u_j, \xi_j^\mu[$, если $u_j < \xi_j^\mu$. В обоих случаях $z_j \in R_j^\mu$ на протяжении интервала времени $[t_0, t_0 + \Delta t]$, поскольку $\boldsymbol{u}(t_0) \in S(\boldsymbol{0}, r)$ для всех этих величин в силу первой части настоящего доказательства.

Следовательно, неравенство (8) может быть переписано в силу (9) и (20) в виде

$$\sum_{i=1}^N a_i(u_i)\left(\delta b_i^/(\eta_i)x_i^2 - \sum_j T_{ij} f_j^/(z_j)x_i x_j\right) \geq 0\,. \tag{21}$$

Далее для всех $\boldsymbol{x} \in S(\boldsymbol{0}, r)$ из условия (9) следует, что

$$\sum_{i=1}^N a_i(u_i)\delta b_i^/(\eta_i)x_i^2 \geq \delta \min_i B_i \min_{i, u_i \in R_i^\mu} a_i(u_i)\|\boldsymbol{x}\|^2. \tag{22}$$

С другой стороны, для всех $\boldsymbol{x} \in S(\boldsymbol{0}, r)$ имеем

$$\begin{aligned}\left|\sum_{i=1}^N a_i(u_i)\sum_j T_{ij} f_j^/(z_j)x_i x_j\right| &\leq \max_{i, u_i \in R_i^\mu} a_i(u_i)\max_{i,j}|T_{ij}|\max_{j, z_j \in R_j^\mu}\left|f_j^/(z_j)\right|\sum_{i=1}^N \sum_j |x_i||x_j| \leq \\ &\leq N\max_{i, u_i \in R_i^\mu} a_i(u_i)\max_{i,j}|T_{ij}|\max_{i, z_j \in R_j^\mu}\left|f_i^/(z_i)\right|\|x\|^2,\end{aligned} \tag{23}$$

поскольку

$$\sum_{i=1}^N \sum_j |x_i||x_j| \leq \left(\sum_{i=1}^N |x_i|\right)\cdot\left(\sum_j |x_j|\right) \leq \left(\sum_{i=1}^N |x_i|\right)^2 \leq N\sum_{i=1}^N x_i^2 = N\|x\|^2 \tag{24}$$

в силу известного неравенства Коши – Буняковского

$$\left(\sum_{i=1}^N a_i b_i\right)^2 \leq \left(\sum_{i=1}^N a_i^2\right)\left(\sum_{i=1}^N b_i^2\right)$$

с $a_i = |x_i|$, $b_i = 1$ для всех $i = \overline{1, N}$. (Заметим, что в первой строке выражения (24) получается равенство, если суммирование по *j* осуществляется от 1 до *N*, включая все значения, в то время как неравенство имеет место в случае, когда условие $j \neq i$ накладывается на сумму по *j*.)

Из выражений (22), (23) следует, что неравенства (21) и, следовательно, (8) выполняются для всех $t \in [t_0, t_0 + \Delta t]$, если условие

$$\delta \min_i B_i \min_{i, u_i} a_i(u_i) \geq N\max_{i,j} a_i(u_i)\max_{i,j}|T_{ij}|\max_{i, z_i}\left|f_i^/(z_i)\right|$$

справедливо (здесь $u_i, z_i \in R_i^\mu$). Последнее неравенство эквивалентно неравенству (19).

3. Очевидно, что момент времени $(t + \Delta t)$ может быть взят как новый начальный момент времени (вместо t_0) в приведенных выкладках и т. д. Таким образом, заключаем по индукции,

что неравенство (8) удовлетворяется для всех $t > 0$, если начальное состояние нейросети $\boldsymbol{u}(0)$ принадлежит шару $S\left(\mathbf{u}^{\mu}, r\right)$ и условие (19) справедливо.

Для того чтобы проиллюстрировать справедливость данного теоретического исследования, было проведено компьютерное моделирование поведения нейронных сетей, описываемых уравнениями (1) в частном случае, когда все $a_i(u_i) = 1,\ b_i(u_i) = \gamma_i u_i$, где $\gamma_i = const$ (в этом случае модель Коэна – Гроссберга совпадает с моделью Хопфилда). Передаточная функция нейрона $f_i(u_i)$ бралась, как обычно, в виде $th(\lambda_i u_i)$. Параметр λ_i варьировался в интервале $1 \le \lambda_i \le 4$, в то время как коэффициенты $\gamma_i = 1$ для всех i. Число нейронов N выбиралось в интервале от 8 до 128, а число запомненных образов было равно 2, 3, …, вплоть до 0,2 N.

Запомненные образы ξ^{μ} генерировались случайным образом так, что их компоненты ограничивались условиями $\ell_{\min} \le \left|\xi_i^{\mu}\right| \le \ell_{\max}$ для всех μ и всех i с параметрами $\ell_{\min}$ и $\ell_{\max}$, варьируемыми в процессе моделирования. Величина $\ell_{\min}$ бралась равной 2., 3., 5. или 10., а величина $\ell_{\max}$ бралась равной 6., 8., 15. или 20. Процесс обучения осуществлялся с использованием алгоритма [8].

После обучения начальное состояние нейросети $\mathbf{u}(t_0)$ выбиралось случайным образом. Можно сразу же отметить, что не было обнаружено случаев, когда нейросеть не сходилась к запомненному образу, но условие (19) выполнялось. Таким образом, неравенство (19) действительно является достаточным условием для восстановления запомненных образов для любого начального состояния $\mathbf{u}(t_0) \in S\left(\mathbf{u}^{\mu}, r\right)$.

Полученные условия являются достаточными, но, вообще говоря, необязательными. Действительно, такие случаи были обнаружены, когда синаптическая матрица не подчиняется условию (19), но состояние нейросети сходится к запомненному образу.

Заключение

В настоящей работе рассматриваются аналоговые нейросети Коэна – Гроссберга [2] в случае, когда матрица синаптических связей несимметрична. Доказано, что при выполнении определенных условий существует глобальный аттрактор, который содержит все аттракторы системы. Найдена верхняя оценка времени, необходимого для достижения заданной окрестности глобального аттрактора. Получены достаточные условия, при выполнении которых нейросеть эволюционирует по направлению к ближайшему стационарному состоянию и достигает заданной окрестности стационарного состояния за конечное время. Если нейросеть обучена подходящим образом, то указанное выше стационарное состояние кодирует один из запомненных образов. Получена верхняя оценка времени, необходимого для восстановления запомненного образа с заданной точностью.

Авторы выражают благодарность А.М. Кроту и В.И. Кувшинову за полезные замечания к работе.

Список литературы

1. Hopfield. J.J. Neural Networks and Physical Systems with Emergent Collective Computational Abilities // Proc. Natl. Acad. Sci. USA. – 1984. – Vol. 79. – P. 2554–2558.

2. Cohen M.A., Grossberg S. Absolute Stability of Global Pattern Formation and Parallel Memory Storage by Competitive Neural Networks // IEEE Trans. Syst., Man, Cybern. – 1983. – Vol. SMS-13. – P. 815–826.

3. Hopfield J.J. Neurons with Graded Response Have Collective Computational Properties Like those of Two. – State Neurons // Proc. Natl. Acad. Sci. USA. – 1984. – Vol. 81. – P. 3088–3092.

4. Pineda F.J. Generalization of Back – Propagation to Recurrent Neural Networks // Phys. Rev. Lett. – 1987. – Vol. 59. – P. 2229–2232.

5. Linkevich A.D. Learning neural networks with the aid of projection techniques // Proc. Second Seminar «Nonlinear Phenomena in Complex Systems», Polatsk, February 15–17, 1993. – Sant-Peterburg, 1993. – P. 373–382.

6. Linkevich A.D. A sequential learning rule for analogous neural networks // Optical Memory and Neural Networks. – 1993. – Vol. 2. – № 2. – P. 111–116.

7. Linkevich A.D. Projection learning algorithms for analog neural networks without self-action // Proc. Workshop «Mind, Brain and Neurocomputers», Minsk, November 16–18, 1993 // Advances in Synergetics. Vol. 1. – Minsk: Belarusian State Univ. Press, 1994. – P. 147–153.

8. Kartynnick A.V., Linkevich A.D. Retrieval of memorized patterns by non-symmetric neural networks: Local Lyapunov functions and local contracting maps methods // Optical Memory and Neural Networksю – 1994. – Vol. 3. – № 3. – P. 329–342.

9. Eccles J.C. Physiology of Synapses. – Berlin: Springer, 1964.

10. Das II P.K., Schieve W.C., Zeng Z.J. Chaos in an Effective Four-Neuron Neural Networks // Phys. Lett A. – 1991. – Vol. 161. – P. 60–66.

11. Linkevich A.D. Mathematical Methods and Models for Investigation of Neurodynamical Mechanisms of Cognitive Processes. – Minsk: IEC NAS RB, 2001.

12. Шилов Г.Е. Математический анализ (функции одного переменного). Ч. 1. – М.: Наука, 1969.

13. Чаплыгин С.А. Собрание работ. Т. 1. – М.–Л.: Гостехиздат, 1948.

14. Арнольд В.И. Обыкновенные дифференциальные уравнения. – М.: Наука, 1971.

Поступила 09.08.05

Полоцкий государственный университет,
Новополоцк, Блохина, 29
e-mail: adlinkevich@tut.by

A.V. Kartynnick, A.D. Linkevich

A GLOBAL ATTRACTOR AND SUFFICIENT CONDITIONS FOR RETRIEVAL OF PATTERNS BY NEURAL NETWORKS OF COHEN – GROSSBERG

The Cohen – Grossberg model of analogous neural networks is studied in the general case when interneuronal (synaptic) connections can be non-symmetric and consequently there is no usual analogue of the Hamiltonians for the systems in complete phase space. It is proved that under certain conditions there exists a global attracting region Q so that all attractors of the system lie inside the Q and, besides, the domain Q is an attractor of the system as well. The upper bound has been estimated for time to reach a prescribed neighbourhood of the region Q. Sufficient conditions have been found under which the network converges asymptotically to a nearest fixed point (memorized pattern) and reaches a given neighbourhood of this attractor in a certain time. An upper bound is given for time needed to retrieve a memorized pattern with a prescribed accuracy.