

Литература: 1. *Campbell S.L.* Singular Systems of Differential Equations. Pitman, 1980. 176 p. 2. *Marz R.* On linear differential-algebraic equations and linearizations // Applied Numerical Mathematics, 1994. P.279–292. 3. *Самойленко А.М., Шкіль М.І., Яковець В.П.* Лінійні системи диференціальних рівнянь з виродженнями: Навч. посіб., К.: Вища шк., 2000. 294 с. 4. *Руткас А.Г.* Задача Коши для уравнения

$A \frac{dx}{dt} + Bx(t) = f(t)$ // Дифференциальные уравнения. 1975. №11. С.1996–2010. 5. *Favini A., Plazzi P.* Some results concerning the abstract nonlinear equation $D_t \mu(t) + Lu(t) = f(t, Ku(t))$ // Circuits, Systems, Signal Processing, 1986. P. 261–274. 6. *Rutkas A.* The solvability of a nonlinear differential equation in a Banach space // Spectral and evolutionary problems. Proceedings of Sixth Crimean Fall Mathematical School-Symposium (Simferopol), 1996. P.317–320. 7. *Favini A., Rutkas A.* Existence and uniqueness of solution of some abstract degenerate nonlinear equation // Differential Integral Equations. 1999. No.12. P.373–394. 8. *Rutkas A.G., Vlasenko L.A.* Existence of solutions of degenerate nonlinear differential operator equations //

Nonlinear Oscillations. 2001. No12. P.252–263. 9. *Власенко Л.А.* Эволюционные модели с неявными и вырожденными дифференциальными уравнениями // Системные технологии. 2006. С.273. 10. *Руткас А.Г., Худошин И.Г.* Глобальная разрешимость одного вырожденного полулинейного дифференциально-операторного уравнения // Нелінійні коливання. 2004. Т.7, №3. С.414–429. 11. *Радбель Н.И.* О начальном многообразии и диссипативности задачи Коши для уравнения $Ax'(t) + Bx(t) = 0$ // Дифференциальные уравнения. 1979. №6. 12. *Худошин И.Г.* Начальная задача для некоторых квазилинейных дифференциально-алгебраических уравнений // Вісник Харківського університету. Математика, прикладна математика і механіка. 1999. №458. С.159–164. 13. *Шварц Л.* Анализ. Т.1. М.: Мир, 1972. С.824.

Поступила в редколлегию 16.09.2008

Рецензент: д-р физ.-мат. наук, проф. Руткас А.Г.

Худошин Илья Григорьевич, ст. преподаватель каф. ММ и ПО ХНУ им. В.Н.Каразина. Адрес: Украина, 61045, г. Харьков, пер. Шекспира, 7, кв. 69, тел. (050)4043498.

УДК 517.922+517.958

О СВОЙСТВАХ ДЕСКРИПТОРНЫХ НЕЙРОННЫХ СЕТЕЙ

РУТКАС А.А.

Конструируется дескрипторная нейронная сеть как определенная комбинация динамических и статических нейронов. Ее функционирование описывается системой разностных и чисто алгебраических уравнений, как в случае дескрипторных систем управления. Исследуются условия разрешимости системы, в частности, находятся необходимые ограничения на вектор входных сигналов сети. Подробно рассматривается дескрипторная сеть с линейными активационными функциями нейронов.

1. Введение

Динамические системы с дискретным временем как математические модели инженерных объектов и естественных процессов могут содержать одновременно с разностными уравнениями еще и чисто алгебраические, статические уравнения связей между динамическими переменными в один и тот же момент времени. В качестве лишь нескольких примеров можно указать макроэкономическую динамическую модель Леонтьева в дискретной форме [1], математические модели биологии и физики [2], вычислительные схемы решения дифференциально-алгебраических уравнений (ДАЕ_S) [3, 4], неявные и вырожденные дискретные динамические системы [5] и др.

Здесь рассматривается простейший класс разностно-алгебраических систем, допускающих преобразование к следующему виду:

$$\left. \begin{aligned} x_1(k+1) &= \psi_1[w_{11}x_1(k) + \dots + w_{1n}x_n(k) + \theta_1(k)] \\ &\dots \\ x_m(k+1) &= \psi_m[w_{m1}x_1(k) + \dots + w_{mn}x_n(k) + \theta_m(k)] \end{aligned} \right\} (1)$$

$$\left. \begin{aligned} x_{m+1}(k) &= \psi_{m+1}[w_{m+1,1}x_1(k) + \dots + w_{m+1,n}x_n(k) + \theta_{m+1}(k)] \\ &\dots \\ x_n(k) &= \psi_n[w_{n1}x_1(k) + \dots + w_{nn}x_n(k) + \theta_n(k)], \\ k &= 0, 1, 2, \dots \end{aligned} \right\} (2)$$

Уравнения подсистемы (1) являются разностными уравнениями первого порядка, остальные $(n - m)$ уравнений в подсистеме (2) – чисто алгебраическими, связывающими переменные $x_1(k), \dots, x_n(k)$ в один и тот же момент времени k . Числовые значения w_{ji} , $\theta_j(k)$ и функции одной переменной $\psi_j(u)$ считаются заданными ($j, i = 1, \dots, n; k = 0, 1, 2, \dots$).

Мы покажем, что уравнения (1), (2) описывают динамику искусственной дискретной нейронной сети с m динамическими нейронами типа Хопфилда [6, 7] и $(n - m)$ статическими нейронами Маккалоха-Питтса [8, 9]. Будем называть такую комбинированную ИНС *дескрипторной*, по аналогии с терминологией теории управления, где дескрипторными называются системы управления с дифференциально-алгебраическими уравнениями состояний [10, 11]. Дескрипторная ИНС имеет ряд свойств, которые не реализуются в чисто динамических или в чисто статических сетях. Изучение этих «дескрипторных» свойств представляет интерес для теории ИНС и для приложений этой теории к анализу объектов и процессов, в математических моделях которых комбинируются разностные (динамические) и алгебраические (статические) уравнения.

2. Модельная дескрипторная ИНС с двумя нейронами

Сеть, изображенная на рис. 1, содержит один динамический нейрон с активационной функцией ψ_1 и один статический нейрон с активационной функцией ψ_2 .

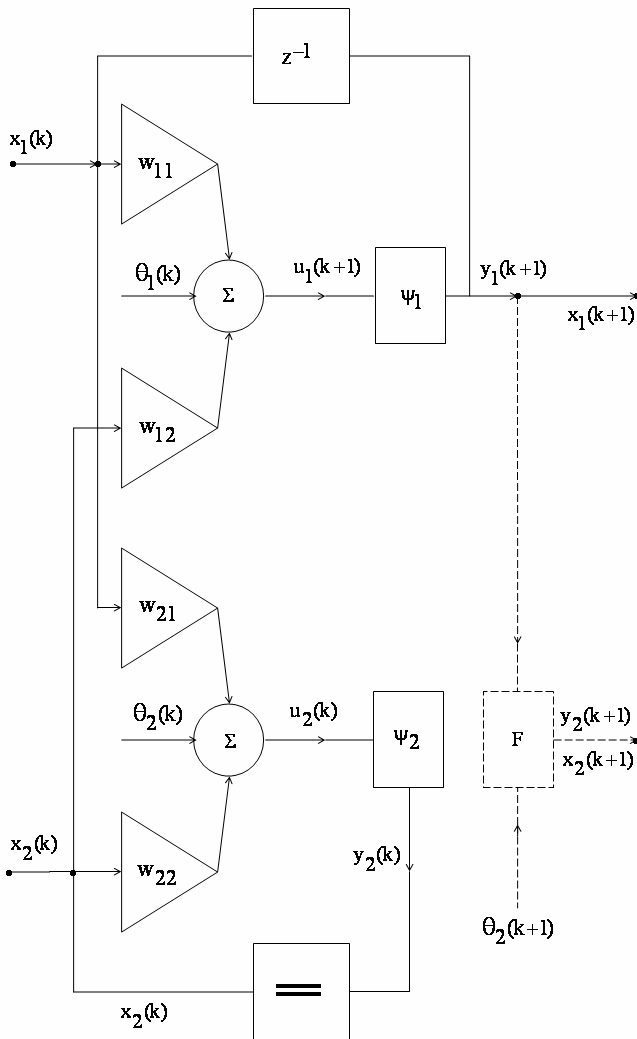


Рис. 1. Дескрипторная сеть с динамическим и статическим нейронами

В момент времени k на вход сети подается два сигнала $x_1(k), x_2(k)$, выходы $y_1(k+1), y_2(k)$ отвечают разным моментам времени $(k+1), k$ соответственно и вычисляются по формулам

$$y_1(k+1) = \psi_1(w_{11}x_1(k) + w_{12}x_2(k) + \theta_1(k)), \quad (3)$$

$$y_2(k) = \psi_2(w_{21}x_1(k) + w_{22}x_2(k) + \theta_2(k)). \quad (4)$$

Для получения входных сигналов в момент $(k+1)$

$$x_1(k+1) = y_1(k+1), \quad x_2(k+1) = y_2(k+1)$$

следует воспользоваться динамическим уравнением (3) и статическим уравнением (4) в следующий момент $(k+1)$:

$$x_1(k+1) = \psi_1(w_{11}x_1(k) + w_{12}x_2(k) + \theta_1(k)), \quad (5)$$

$$x_2(k+1) = \psi_2(w_{21}x_1(k+1) + w_{22}x_2(k+1) + \theta_2(k+1)) \quad (6)$$

Предположим, соотношение (6) можно разрешить явно относительно переменной $x_2(k+1)$:

$$x_2(k+1) = F[x_1(k+1), \theta_2(k+1)]. \quad (7)$$

Тогда путем подстановки правой части (5) вместо $x_1(k+1)$ в (7) можно получить выход $y_2(k+1) = x_2(k+1)$ как функцию от входов $x_1(k), x_2(k)$, смещения $\theta_1(k)$ в момент k и смещения $\theta_2(k+1)$ в следующий момент $(k+1)$:

$$x_2(k+1) = F\{\psi_1[w_{11}x_1(k) + w_{12}x_2(k) + \theta_1(k)], \theta_2(k+1)\} \equiv \Pi[x_1(k), x_2(k), \theta_1(k), \theta_2(k+1)]. \quad (8)$$

Естественно, что наличие динамического и статического нейронов в сети предопределяет использование сигналов смещения $\theta_1(k)$ и $\theta_2(k+1)$ в разные моменты времени для получения полного выходного состояния $y_1(k+1), y_2(k+1)$ в момент $k+1$.

Подставляя правую часть (5) в (6) вместо $x_1(k+1)$, получаем функциональное уравнение относительно $x_2(k+1)$:

$$x_2(k+1) = \Phi(x_2(k+1)), \quad \Phi \equiv \psi_2\{w_{21} \cdot \psi_1[w_{11}x_1(k) + w_{12}x_2(k) + \theta_1(k)] + w_{22} \cdot x_2(k+1) + \theta_2(k+1)\} \quad (9)$$

Функция Π (8) может быть найдена непосредственно (без использования функции F (7)) из соотношения (9). Возможность получить единственное явное решение $x_2(k+1)$ уравнения (9) предполагает определенные согласования между параметрами сети – синаптическими весами w_{ik} , активационными функциями ψ_k и сигналами смещения $\theta_1(k), \theta_2(k+1)$. В частности, подходящими являются такие условия согласования, которые обеспечивают сжимаемость отображения Φ (9) как функции от переменной $x_2(k+1)$ в нетривиальной замкнутой ограниченной области значений этой переменной.

3. Ограничения на входные векторы, распознаваемые дескрипторной ИНС

Равенство входа $x_2(k)$ и выхода $y_2(k)$ (4) в статическом нейроне порождает на функциональной схеме рис.1 цикл, содержащий блок отождествления $\square =$. Это означает, что входные сигналы $x_1(k), x_2(k)$ не могут выбираться произвольно на любом k -м шаге (итерации), а должны удовлетворять соотношению, вытекающему из уравнения (4):

$$x_2(k) = \psi_2[w_{21}x_1(k) + w_{22}x_2(k) + \theta_2(k)], \quad k = 0, 1, 2, \dots \quad (10)$$

Следовательно, вектор $x(k) = (x_1(k), x_2(k))^T$ воспринимается или распознается дескрипторной нейронной сетью (см. рис.1) в том случае, когда он принадлежит определенному одномерному многообразию Λ_k в двухмерном пространстве входов $X = \mathbb{R}^2, \dim \Lambda_k = 1$. В частности, допустимые вход-

ные векторы $x(0)$ начального шага $k = 0$ должны принадлежать одномерному *начальному многообразию*

$$\Lambda_0 = \{(x_1; x_2)^T : \psi_2[w_{21} \cdot x_1 + w_{22} \cdot x_2 + \theta_2(0)] - x_2 = 0\} \quad (11)$$

Приведем достаточные условия разрешимости функционального уравнения (10) относительно $x_2(k)$, которые позволяют указать конструктивное правило построения отображений F (7), Π (8).

Теорема. Пусть для каждого $k = 0, 1, 2, \dots$, каждого x из открытого множества $\Omega \subset \mathbb{R}$ и каждого y из сегмента $y_1 \leq y \leq y_2$ активационная функция $\psi_2(u)$ непрерывно дифференцируема при всех соответствующих значениях своего аргумента $u = u(x, y, k) = w_{21} \cdot x + w_{22} \cdot y + \theta_2(k)$. Предположим, для каждого фиксированного k функция $g(u) = \frac{d\psi_2(u)}{du} \cdot w_{22} - 1$ отлична от нуля и не меняет свой знак в области $\{(x, y) = \Omega \times [y_1, y_2]\}$, а на концах сегмента $[y_1, y_2]$ одновременно для всех $x \in \Omega$ выполнено знаковое условие

$$\text{Sgn}[\psi_2(u(x, y_1, k)) - y_1] + \text{Sgn}[\psi_2(u(x, y_2, k)) - y_2] = 0 \quad (12)$$

Тогда при фиксированном k и заданном смещении $\theta_2(k)$ для каждого значения $x_1 \in \Omega$ существует единственное $x_2(k) \in [y_1, y_2]$, такое что равенство (10) выполняется, а соответствующая явная функция $x_2(k) = f_k(x_1)$ непрерывна и дифференцируема по x_1 . Размерность многообразия Λ_k соответствующих векторов $(x_1, x_2(k))^T$ равна $\dim \Omega = 1$.

Доказательство. Функция

$$v_k(x, y) = \psi_2[w_{21} \cdot x + w_{22} \cdot y + \theta_2(k)] - y$$

имеет непрерывные частные производные по x и y .

Производная $\frac{\partial v_k(x, y)}{\partial y} = \frac{d\psi_2}{du} \cdot w_{22} - 1$ всюду в области $\Omega \times [y_1, y_2]$ отлична от нуля и сохраняет знак, так что по переменной $y \in [y_1, y_2]$ функция $v_k(x, y)$ является строго монотонной. Вследствие (12) $\text{Sgn} v_k(x, y_1) = -\text{Sgn} v_k(x, y_2)$, так что по теореме о промежуточных значениях непрерывной функции существует нуль y функции $v_k(x, y)$ и при каждом фиксированном x он является единственным нулем на интервале (y_1, y_2) : $v_k(x, y) = 0, y = y(x) \in (y_1, y_2)$. Таким образом, глобально на области Ω определена функция нулей $y = f_k(x) : \Omega \rightarrow [y_1, y_2]$ такая, что $v_k(x, f_k(x)) = 0, \forall x \in \Omega$. Из локальной теоремы о не-

явной функции следует, что явная функция $y = f_k(x)$ непрерывна и дифференцируема. Соотношение (10) удовлетворяется для каждой пары точек $x_1(k) = x_1; x_2(k) = f_k(x_1), x_1 \in \Omega$. Теорема доказана.

4. Предельные режимы функционирования дескрипторной нейронной сети

Для каждого входного сигнала x_i существует такой интервал его «малых» значений, что после синаптического усиления $w_{ji}x_i$ он все же не преодолевает порога чувствительности j -го нейрона и потому не воспринимается j -м нейроном. При большом разбросе значений синаптических весов w_{ji} входной сигнал x_i может не восприниматься j -м нейроном, но восприниматься другим, k -м нейроном. Отметим три таких предельных режима в дескрипторной сети (см. рис. 1).

I. *Режим автономного функционирования динамической и статической составляющих.*

Если в уравнениях (1), (2) величины $w_{21}x_1(k), w_{12}x_2(k)$ малы, то заменяя их нулями, получаем несвязанные уравнения

$$\begin{aligned} y_1(k+1) &= \psi_1(w_{11}x_1(k) + \theta_1(k)); \\ y_2(k) &= \psi_2(w_{22}x_2(k) + \theta_2(k)) \end{aligned}$$

динамического и статического нейронов.

II. *Режим относительно слабой собственной синаптической связи в статическом нейроне.*

Пусть влияние собственного сигнала $w_{22}x_2(k)$ в статическом нейроне не ощущается по сравнению с влиянием $w_{21}x_1(k)$ со стороны динамического нейрона. В частности, это верно при любых входах, если $w_{22} = 0$. Тогда необходимость решения неявного функционального уравнения (6) относительно $x_2(k+1)$ отпадает, использование достаточных условий типа теоремы также излишне. Из (5), (6) с учетом $w_{22}x_2(k+1) = 0$ получаем явное выражение для отображения Π (8):

$$x_2(k+1) = \psi_2\{w_{21} \cdot \psi_1[w_{11}x_1(k) + w_{12} \cdot x_2(k) + \theta_1(k)] + \theta_2(k+1)\} \quad (13)$$

III. *Режим активационной компенсации собственного синаптического усиления в статическом нейроне.*

Рассматривается дескрипторная сеть (см. рис. 1), в которой активационная функция статического нейрона имеет вид

$$\psi_2(u_2) = \frac{1}{w_{22}}u_2 + b_2 + o(u_2), \quad (14)$$

где нелинейная часть $o(u_2)$ мала в рабочем диапазоне изменения внутреннего сигнала u_2 , а скорость изменения линейной части является постоянной, обратной к синаптическому весу w_{22} . Подстановка в (6) активационной функции $\psi_2(u_2)$ вида (14) с внутренним сигналом

$$u_2 = u_2(k+1) = w_{21} \cdot x_1(k+1) + w_{22} \cdot x_2(k+1) + \theta_2(k+1) \quad (15)$$

приводит к соотношению

$$w_{21} \cdot x_1(k+1) = -\theta_2(k+1) - w_{22} \cdot b_2 - w_{22} \cdot o(u_2). \quad (16)$$

После исключения $x_1(k+1)$ в (16) и (5) мы получаем

$$\theta_2(k+1) + w_{21} \cdot \psi_1[w_{11} \cdot x_1(k) + w_{12} \cdot x_2(k) + \theta_1(k)] + w_{22} \cdot b_2 + w_{22} \cdot o(u_2(k+1)) = 0. \quad (17)$$

Отбрасывая малую нелинейность $o(u_2(k+1))$, в качестве *главного линейного приближения* получаем *правило упреждающего выбора* смещения $\theta_2(k+1)$ на статическом нейроне:

$$-\theta_2(k+1) = w_{21} \cdot \psi_1[w_{11} \cdot x_1(k) + w_{12} \cdot x_2(k) + \theta_1(k)] + w_{22} \cdot b_2 \quad (18)$$

Правило (18) есть условие корректного функционирования дескрипторной сети (см. рис.1) в режиме активационной компенсации (14) собственного синаптического усиления w_{22} в статическом нейроне.

Замечание. Рассмотренный режим «компенсации» является исключительным в том смысле, что он параметрически неустойчив. Действительно малое возмущение $\frac{1}{w_{22}} \rightarrow \left(\frac{1}{w_{22}} + \varepsilon\right)$ линейной части в (14) меняет соотношение (16) на равенство

$$\varepsilon \cdot w_{22} \cdot x_2(k+1) + \left(\frac{1}{w_{22}} + \varepsilon\right) \theta_2(k+1) + \left(\frac{1}{w_{22}} + \varepsilon\right) \cdot w_{21} \cdot \psi_1[w_{11} \cdot x_1(k) + w_{12} \cdot x_2(k) + \theta_1(k)] + b_2 + o(u_2(k+1)) = 0.$$

Пренебрегая величиной $o(u_2)$, мы вместо упреждающего значения $\theta_2(k+1)$ получаем формулу для нахождения $x_2(k+1)$ при произвольном задании смещения $\theta_2(k+1)$.

5. Линейная дескрипторная ИНС

В случае линейных активационных функций нейронов $\psi_k(u) = a_k u + b_k$ уравнения (5), (6) имеют вид

$$x_1(k+1) = a_1[w_{11} \cdot x_1(k) + w_{12} \cdot x_2(k) + \theta_1(k)] + b_1, \quad (19)$$

$$(1 - a_2 w_{22}) \cdot x_2(k+1) = a_2 w_{21} \cdot x_1(k+1) + a_2 \theta_2(k+1) + b_2 \quad (20)$$

Кроме самостоятельного значения, исследование линейных дескрипторных сетей представляет интерес для качественной оценки эволюции нелинейных дескрипторных сетей, у которых активационные функции нейронов имеют участки линейности или участки удовлетворительной линейной аппроксимации.

Исключительный режим компенсации синаптического усиления w_{22} и активации ψ_2 реализуется, если $a_2 w_{22} = 1$. В этом случае упреждающее значение смещения $\theta_2(k+1)$ единственным образом вычисляется по формуле

$$-\theta_2(k+1) = a_1 w_{21} [w_{11} \cdot x_1(k) + w_{12} \cdot x_2(k) + \theta_1(k)] + w_{21} b_1 + w_{22} \cdot b_2. \quad (21)$$

В нормальном, параметрически устойчивом режиме функционирования дескрипторной нейронной сети, когда $a_2 \cdot w_{22} \neq 1$, из (19), (20) находим

$$x_2(k+1) = (1 - a_2 w_{22})^{-1} \{a_2 w_{21} a_1 [w_{11} \cdot x_1(k) + w_{12} \cdot x_2(k) + \theta_1(k)] + a_2 w_{21} b_1 + b_2 + a_2 \theta_2(k+1)\} \quad (22)$$

Равенства (19), (22) определяют отображение рекурсии

$$S_k : x(k) \rightarrow x(k+1), \quad x(k) = (x_1(k), x_2(k))^T. \quad (23)$$

Если смещения θ_1, θ_2 не зависят от k , то отображение $S_k = S$ также не зависит от k , а в однородном случае нулевых смещений $\theta_j(k) = 0$ и $b_j = 0$ отображение S линейно и имеет матричную форму

$$\hat{S} = a_1 \begin{bmatrix} w_{11} & w_{12} \\ r w_{11} & r w_{12} \end{bmatrix}, \quad r = \frac{a_2 w_{21}}{1 - a_2 w_{22}}. \quad (24)$$

Многообразие $\Lambda_k = \{x(k)\}$ допустимых входных векторов, распознаваемых дескрипторной ИНС и удовлетворяющих соотношению (10), в данном случае определяется как

$$\Lambda_k = \{x(k) : (1 - a_2 w_{22}) x_2(k) = a_2 w_{21} \cdot x_1(k) + a_2 \cdot \theta_2(k) + b_2\} \quad (25)$$

Формально отображение рекурсии S_k (23) можно применять к любым векторам $x \in \mathbb{R}^2$, однако только векторы $x(k) \in \Lambda_k (\subset \mathbb{R}^2)$ образуют распознаваемые входы k -й итерации, а отображение S_k переводит их в допустимые входы $(k+1)$ -й итерации:

$$S_k(\Lambda_k) = \Lambda_{k+1}, \quad k = 0, 1, 2, \dots \quad (26)$$

Этот факт следует из равенства (20). Учитывая (26), с помощью суперпозиции отображений S_j можно по-

лучать распознаваемые векторы входов $x(k)$ и строить многообразие Λ_k , исходя из допустимых начальных векторов $x(0)$ и начального многообразия Λ_0 на нулевом шаге $k = 0$:

$$\begin{aligned} x(k) &= S_{k-1} \circ S_{k-2} \circ \dots \circ S_0(x(0)); \\ \Lambda_k &= S_{k-1} \circ S_{k-2} \circ \dots \circ S_0(\Lambda_0) \end{aligned} \quad (27)$$

При этом начальное многообразие Λ_0 описывается в виде

$$\Lambda_0 = \{x(0) = (x_1, x_2)^{tr} : x_2 = (1 - a_2 w_{22})^{-1} \cdot [a_2 w_{21} \cdot x_1 + a_2 \theta_2(0) + b_2], \forall x_1 \in \mathbb{R}\} \quad (28)$$

В частности, если $\theta_2(0) = b_2 = 0$, то начальное многообразие $\hat{\Lambda}_0 = \text{span}\{(1, r)^{tr}\}$ есть одномерное линейное подпространство в \mathbb{R}^2 , а вследствие структуры рекурсивного отображения \hat{S} (24) оно отображает $\hat{\Lambda}_0$ на $\hat{\Lambda}_0$. В этом случае все реализуемые векторы входов-выходов $x(k)$ лежат в начальном многообразии $\hat{\Lambda}_0$ и $\hat{\Lambda}_k = \hat{\Lambda}_0$ для всех k .

6. Дескрипторная ИНС с произвольным числом нейронов

В общем случае разностно-алгебраическая система (1), (2) описывает дескрипторную нейронную сеть с m динамическими и $(n - m)$ статическими нейронами. Принципиальная логическая схема этой сети представлена на рис.2, где входные состояния $x_j(k)$ в k -й момент времени сгруппированы в векторы

$$x_D(k) = \begin{bmatrix} x_1(k) \\ \vdots \\ x_m(k) \end{bmatrix}; x_C(k) = \begin{bmatrix} x_{m+1}(k) \\ \vdots \\ x_n(k) \end{bmatrix}; x(k) = \begin{bmatrix} x_D(k) \\ x_C(k) \end{bmatrix}.$$

Здесь и далее индексом D помечаются векторы переменных или параметров, отвечающих динамическим нейронам, индексом C – статическим нейронам. В частности, внутренние состояния нейронов u_j на k -м шаге объединяются в два вектора:

$$\begin{aligned} u_D(k+1) &= (u_1(k+1), \dots, u_m(k+1))^{tr}, \\ u_j(k+1) &= \sum_{i=1}^n w_{ji} \cdot x_i(k) + \theta_j(k), \quad j = 1, \dots, m. \\ u_C(k) &= (u_{m+1}(k), \dots, u_n(k))^{tr}, \\ u_j(k) &= \sum_{i=1}^n w_{ji} \cdot x_i(k) + \theta_j(k), \quad j = m+1, \dots, n. \end{aligned}$$

С помощью $m \times n$ -матрицы D и $(n - m) \times n$ -матрицы C : $D = [E_m; 0]$, $C = [0; E_{n-m}]$ матрица синаптических весов $W(n \times n)$ разбивается на четыре блока соответствующих размерностей, присутствующих на блок-схеме рис.2:

$$W = \begin{bmatrix} DWD^* & | & DWC^* \\ - & - & - \\ CWD^* & | & CWC^* \end{bmatrix} : \mathbb{R}^m \oplus \mathbb{R}^{n-m} \rightarrow \mathbb{R}^m \oplus \mathbb{R}^{n-m}.$$

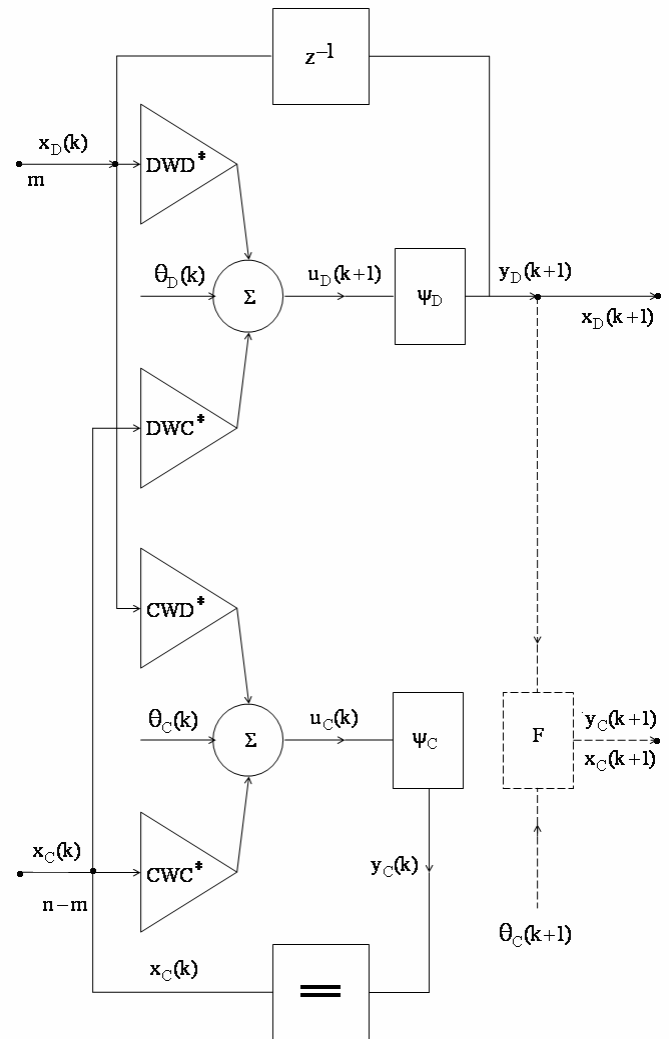


Рис. 2. Блочная схема дескрипторной сети с m динамическими и $(n - m)$ статическими нейронами

Обозначим

$$\begin{aligned} \Psi_D(u_D) &= \begin{bmatrix} \psi_1(u_1) \\ \vdots \\ \psi_m(u_m) \end{bmatrix}; \Psi_C(u_C) = \begin{bmatrix} \psi_{m+1}(u_{m+1}) \\ \vdots \\ \psi_n(u_n) \end{bmatrix}; \\ \theta_D(k) &= \begin{bmatrix} \theta_1(k) \\ \vdots \\ \theta_m(k) \end{bmatrix}; \theta_C(k) = \begin{bmatrix} \theta_{m+1}(k) \\ \vdots \\ \theta_n(k) \end{bmatrix}. \end{aligned}$$

Тогда уравнения (1), (2) записываются в векторной форме:

$$x_D(k+1) = \Psi_D[DWx(k) + \theta_D(k)], \quad (29)$$

$$x_C(k) = \Psi_C[CWx(k) + \theta_C(k)]. \quad (30)$$

Статическое уравнение (30) связи между входными сигналами $x_j(k)$ перепишем детальнее для $(k + 1)$ -го шага:

$$\begin{aligned} x_C(k+1) &= \Psi_C[CWD^* \cdot x_D(k+1) + \\ &+ CWC^* \cdot x_C(k+1) + \theta_C(k+1)]. \end{aligned} \quad (31)$$

Пусть уравнение (31) разрешимо относительно вектора $x_C(k+1)$:

$$x_c(k+1) = F\{x_D(k+1), \theta_c(k+1)\} = F\{\psi_D[DW \cdot x(k) + \theta_D(k)], \theta_c(k+1)\}. \quad (32)$$

Функция F явно выражает статическую векторную компоненту $x_c(k+1)$ через динамическую $x_D(k+1)$ и смещение $\theta_c(k+1)$, что отражено на блок-схеме дескрипторной сети рис.2 пунктирным блоком F .

Соотношения (29), (32) определяют отображение рекурсии (ср.(23))

$$S_k : x(k) \rightarrow \begin{bmatrix} x_D(k+1) \\ x_c(k+1) \end{bmatrix}; S_k = S_k(\theta_D(k), \theta_c(k+1)). \quad (33)$$

Распознаваемые сетью на k -м шаге входные векторы $x(k) = x_D(k) \oplus x_c(k)$ образуют в \mathbb{R}^n многообразие

$$\Lambda_k = \{x(k) : x_c(k) = \psi_c[CWD^* \cdot x_D(k) + CWC^* \cdot x_c(k) + \theta_c(k)]\}.$$

В случае *линейных* активационных функций нейронов $\psi_k(u_k) = a_k \cdot u_k + b_k$ примем обозначения векторов и матриц:

$$\psi(u) = a \cdot u + b; u = \begin{bmatrix} u_D \\ u_c \end{bmatrix}, b = \begin{bmatrix} b_1 \\ \dots \\ b_n \end{bmatrix}, a = \text{diag}\{a_j\}_{j=1}^n.$$

Уравнения (1), (2) имеют векторную форму:

$$\begin{aligned} x_D(k+1) &= DAD^*DW \cdot x(k) + DAD^*\theta_D(k) + Db, \\ x_c(k) &= CAC^*[CWD^* \cdot x_D(k) + CWC^* \cdot x_c(k) + \theta_c(k)] + C \cdot b. \end{aligned} \quad (34)$$

Из последнего уравнения можно выразить $x_c(k)$ через $x_D(k)$:

$$\begin{aligned} x_c(k) &= V^{-1}\{CAC^*[CWD^* \cdot x_D(k) + \theta_c(k)] + C \cdot b\}, \\ V &= E_c - CAC^*CWC^*, \end{aligned}$$

если $\det V \neq 0$. Отсюда входы статических нейронов $x_c(k+1)$ на $(k+1)$ -м шаге образуют вектор

$$x_c(k+1) = V^{-1}\{CAC^* \cdot CWD^*[DAD^*DW \cdot x(k) + DAD^*\theta_D(k) + Db] + CAC^* \cdot \theta_c(k+1) + Cb\}. \quad (35)$$

Отображение рекурсии (33) задается равенствами (34), (35), а в однородном случае ($b_j = 0, \theta_j(k) = 0, \forall j, k$) оно не зависит от k и принимает форму блочной матрицы:

$$\hat{S} = \begin{bmatrix} DAD^*DW & \\ V^{-1}CAC^*CWD^*DAD^*DW & \end{bmatrix} : x(k) \rightarrow \begin{pmatrix} x_D(k+1) \\ x_c(k+1) \end{pmatrix}.$$

В последней ситуации многообразии распознаваемых сетью входных векторов одинаково для всех шагов ($k = 0, 1, 2, \dots$), является линейным подпространством $\hat{\Lambda}$ в \mathbb{R}^n :

$$\begin{aligned} \hat{\Lambda} &= \{x_D \oplus x_c : x_c = V^{-1}CAC^*CWD^* \cdot x_D, \forall x_D \in \mathbb{R}^m\}, \\ \dim \hat{\Lambda} &= m. \end{aligned}$$

7. Выводы

Рассмотрена искусственная нейронная сеть, сконструированная из динамических нейронов типа Хопфилда и статических нейронов типа Маккалоха-Питтса. Функционирование сети описано разностными и алгебраическими уравнениями, векторная форма которых содержит вырожденную матрицу при полном векторе разностей. Подобные уравнения встречаются в дескрипторных системах управления, в связи с чем указанная нейронная сеть названа дескрипторной. Изучены особенности нейронной сети, обусловленные ее дескрипторностью – комбинированием динамических и статических нейронов. В частности, не всякий набор входных сигналов (входной вектор) распознается сетью. Исследованы многообразия допустимых входных векторов, их зависимость от смещений и активационных функций нейронов, от конструкции сети. Описано три предельных режима функционирования дескрипторной нейронной сети, определяемых специальными выборами синаптических весов либо активационных функций нейронов. Предложенные нейросетевые модели могут применяться для анализа и синтеза физических и вычислительных систем, описываемых неявными разностными и алгебраическими уравнениями.

Литература: 1. *Leontief W.W.* Input-output Economics, London-New York, 1966. 2. *Campbell S.L.* Singular Systems of Differential Equations – San Francisco, London, Melbourne: Pitman Publishing, Research Notes in Mathematics; I – Vol.40, 1980, 176 p.; II – Vol.61, 1982. 234 p. 3. *Бояринцев Ю.Е., Данилов В.А., Логинов А.А.* Численные методы решения сингулярных систем. Новосибирск: Наука, 1989. 223 с. 4. *Kunkel P., Mehrmann V.* Differential-Algebraic Equations: Analysis and Numerical Solutions, Zurich, European Mathematical Society Publishing House. 5. *Бондаренко М.Ф., Руткас А.Г.* Признаки детерминированности неявных дискретных неавтономных систем // Доп.НАН України. 2001. № 2. С.7-11. 6. *Hopfield J.I.* Neural Networks and Physical Systems with Emergent Collective Computational Abilities, Proc. of the National Academy of Science. 1982. 79. P.2554-2558. 7. *Бодянский Е.В., Руденко О.Г.* Искусственные нейронные сети: архитектуры, обучение, применения. Харьков, ТЕЛЕТЕХ, 2004. 372 с. 8. *Mc Culloch W.S., Pitts W.* A Logical Calculus of the Ideals Immanent in Nervous Activity, Bulletin of Mathematical Biophysics. 1943. № 5. P.115-133. 9. *Руденко О.Г., Бодянский Е.В.* Искусственные нейронные сети. Харьков: «СМИТ», 2005. 408 с. 10. *Bender D.I., Laub A.* The linear-quadratic optimal regulator for descriptor systems, IEE Transactions on Automatic Control. 1987. Vol.AC-32, № 6. P.2062-2077. 11. *Campbell S.L.* Nonregular descriptor systems with delays, IMA I.Math.Control and Information. 1995. V.12.P.57-67.

Поступила в редколлегию 11.08.2008

Рецензент: д-р техн. наук, проф. Кривуля Г.Ф.

Руткас Андрей Анатольевич, аспирант ХНУРЭ. Научные интересы: машинный перевод, искусственные нейронные сети, динамические системы. Увлечение и хобби: электронное и математическое обеспечение систем GPS, системы безопасности и слежения. Адрес: Украина, 61001, Харьков, ул. Плехановская 2/5, кв. 29, тел.: (057) 732 28 35.