Tạp chí Tin học và Điều khiển học, T. 12, S. 4 (1996) (45–57)

MỞ RỘNG CẤU TRÚC VÀ HÀM LIAPUNOV CHO MẠNG NƠ RON HOPFIELD

NGUYỄN QUANG HOAN

Abstract. The existence of the Hopfield neural network and the Liapunov function guarantees the convergence to some local minimum of these network with linear inputs [1, 6]. But standard Hopfield architecture [3, 14] is limited in its capacity of linear classifiers. In the present investigation we developed high-order Hopfield network architectures with their corresponding Liapunov functions $V(x)$ for more sophisticated classifiers with higher performances. We also regarded some assumptions on these architectures and proved the their stability.

1. GIỚI THIỆU MẠNG NƠ RON PHẢN HỒI

Mạng nơ ron có thể được xem như là mạng của các phần tử tính (norocomputing) phỏng theo nơ ron sinh học rất được chú trọng trong lĩnh vực tính cũng như trong điều khiển. Có hai khuynh hướng nghiên cứu đối với mạng nơ ron hiện nay. Thứ nhất, phát triển và hoàn thiện các cấu trúc cũng như thuật học cho các mạng nơ ron nhân tạo. Thứ hai, ứng dụng các kết quả vào các lĩnh vực khác nhau với các đặc tính trội của nó. Nghiên cứu này nhằm mở rộng cấu trúc Hopfield - một trong những mạng của nhóm phản hồi và xây dựng các hàm Liapunov tương ứng cho lĩnh vực thứ nhất. Mạng phản hồi gồm n phần tử với các đầu ra y_i được nối với tất cả các đầu vào của mạng. Đối với phần tử nơ ron i có thể xem như một hệ thống nhiều đầu vào và một đầu ra [6].

Cohen và Grossberg [11] đã nêu mô hình tổng quát về mạng phản hồi và chứng minh ổn định của mạng với định lý sau:

Định lý 1 (Cohen và Grossberg về hội tụ của các quỹ đạo). *Trong 1 hệ bất kỳ*

$$x_i = a_i(x_i)\left[b_i(x_i) - \sum_{j=1}^{n} c_{ij} d_j(x_j)\right] \tag{1}$$

sao cho

a) Ma trận $\|c_{ij}\|$ đối xứng và tất cả $c_{ij} \geq 0,\ i = 1, 2, ..., n;$

b) hàm $a_i(\xi),\ b_i(\xi)$ là liên tục cho $\xi > 0,\ i = 1, 2, ..., n;$

c) hàm $a_i(\xi) > 0$ *cho* $\xi > 0$*; hàm* $d_i(\xi) \geq 0$ *với tất cả* ξ, $i = 1, 2, \ldots, n$*;*

d) hàm $d_i(\xi)$ *khả vi và đơn điệu không giảm cho* $\xi \geq 0$, $i = 1, 2, \ldots, n$*;*

e) $\limsup\limits_{\xi \to \infty} [b_i(\xi) - c_{ii}d_i(\xi)] < 0$ *cho mọi* $i = 1, 2, \ldots, n$*;* (2)

f) $\lim\limits_{\xi \to 0^+} b_i(\xi) = \infty$, $i = 1, 2, \ldots, n$*;* (3)

hoặc $\lim\limits_{\xi \to 0^+} b_i(\xi) < \infty$*;* (4)

và $\int\limits_0^\varepsilon \frac{d\xi}{a_i(\xi)} = \infty$ *với* $\varepsilon > 0$. (5)

Tất cả các quỹ đạo chấp nhận đạt đến tập M bất biến lớn nhất chứa trong tập E (chứng minh ở phụ lục A)

$$E\{y \in R^n : \frac{d}{dt}V(y) = 0,\ y \geq 0\} \tag{6}$$

Ở đây,

$$V(x) = -\sum_{i=1}^{n} \int_0^{x_i} b_i(\xi)d_i(\xi_i)d\xi_i + \frac{1}{2}\sum_{i,j=1}^{n} C_{ij}d_i(x_i)d_j(x_j) \tag{7}$$

$$\frac{d}{dt}V(x) = -\sum_{i=1}^{n} a_i(x_i)d_i(x_i)\left[b_i(x_i) - \sum_{j=1}^{n} c_{ij}d_j(x_j)\right]^2 \tag{8}$$

Cohen và Grossberg chỉ ra rằng khi bổ đề 1 (phụ lục A) đảm bảo thì $V(x)$ là hàm Liapunov [11] tức $V(x)$ xác định dương và $\dot{V}(x) \leq 0$.

Hopfield (1984) cũng độc lập đưa ra mô hình mạng và mô tả bằng tập các phương trình vi phân (9) và hàm ra (10)

$$C_i x_i = -\frac{x_i}{R_i} + I_i + \sum_{j=1}^{n} T_{ij} y_j \tag{9}$$

$$y_j = g_j(x_j) \tag{10}$$

$$x_i = g_i^{-1}(y_i) \tag{11}$$

Phương trình (9) là phương trình mô tả luật tác động, trong đó C_i, R_i là các hằng số, I_i là ngưỡng, T_{ij} là trọng liên kết phần tử nơ ron thứ j tới nơ ron thứ i, x_i là trạng thái thứ i của mạng, $g_i(\cdot)$ là hàm sigmoid với hàm ngược của nó $x_i = g_i^{-1}(\cdot)$.

Hopfield cũng đưa ra hàm Liapunov cho mạng:

$$V(x) = \sum_{i=1}^{n}(1/R_i)\int_0^{y_i} g_i^{-1}(\varsigma)d\varsigma - \sum_{i=1}^{n} I_i y_i - \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} T_{ij} y_i y_j \tag{12}$$

Grossberg (1988) [11] cũng chỉ ra rằng mô hình (9) có thể xem là một trường hợp đặc biệt của (1) với các giả thiết:

$$a_i(x_i) = \frac{1}{C_i} = \text{const} \tag{13}$$

$$b_i(x_i) = -\frac{1}{R_i}x_i + I_i \tag{14}$$

$$C_{ij} = -T_{ij}\,,\quad d_j(x_j) = g_j(x_j) \tag{15}$$

trong đó $g_j(x_j)$ là hàm sigmoid, dương, khả vi, đơn điệu tăng, đồng thời đảm bảo hàm ngược của nó $x_i = g_i^{-1}(\cdot) = d'(x_i) \geq 0$ và tồn tại hàm Liapunov (12),

Từ những cơ sở và giả thiết trên, rõ ràng các hàm và hệ số của mô hình Hopfield hoàn toàn đảm bảo các giả thiết của Grossberg. Từ đó có thể phát biểu định lý sau cho mô hình Hopfield khi áp dụng các giả thiết (12) - (15).

Định lý 2. *Trong 1 hệ (Hopfield chuẩn)*

$$C_i x_i = -\frac{x_i}{R_i} + I_i + \sum_{j=1}^{n} T_{ij} g_j(x_j) \tag{16}$$

$$y_j = g_j(x_j) \tag{17}$$

sao cho

a) ma trận $\|T_{ij}\|$ *đối xứng và tất cả* $T_{ij} \leq 0,\ i,\ j = 1, 2, \dots, n;$ (18)

b) hàm $g_j(x)$ *khả vi, dương, đơn điệu không giảm, với mọi* $x_i > 0$ *và là hàm sigmoid, và hàm ngược của nó* $X_j = g_j^{-1}(y_j)$ *cho mọi* $i = 1, 2, \dots, n;$ (19)

c) $\limsup\limits_{\xi\to\infty}[-x_i/R_i + I_i + T_{ii}y_i(x_i)] < 0$ *cho mọi* $i = 1, 2, \dots, n$ (20)

d) $\lim\limits_{\xi\to 0^+}(-X_i(\xi)/R_i + I_i) = \infty;$ (21)

hoặc $\lim\limits_{\xi\to 0^+}(-X_i(\xi)/R_i + I_i) < \infty;$ (22)

và $\int_0^{\varphi} C_i dx = \infty$ *với* $\varphi > 0;$ (23)

thì tồn tại

$$V(x) = \sum_{i=1}^{n} \frac{1}{R_i} \int_0^{g_i} g_i^{-1}(\xi) d\xi - \sum_{i=1}^{n} I_i y_i - \frac{1}{2} \sum_{i,j=1}^{n} T_{ij} y_i y_j \tag{24}$$

xác định dương và

$$\dot{V}(x) = -\sum C_i g_i^{-1}(y_i)(y_i')^2 \leq 0 \tag{25}$$

Chứng minh. Với giả thiết (19) biểu thức (24) có thể viết

$$V(x) = \sum_{i=1}^{n} \frac{1}{R_i} \int_0^{g_i} (x_i - I_i) d\xi - \frac{1}{2} \sum_{i,j=1}^{n} T_{ij} y_i y_j \tag{24'}$$

nếu giả thiết (18), (21) đảm bảo, $T_{ij} \leq 0$ thì hệ (16) là trường hợp riêng của (1) và (24) là trường hợp riêng của (7) nên hoàn toàn áp dụng được bổ đề 1, tức là đảm bảo quỹ đạo dương và (24) xác định dương. Biểu thức (25) đã được Hopfield chứng minh $V(x) \leq 0$ [3], do đó $V(x)$ là hàm Liapunov [13].

Đối với hệ này tồn tại nhiều điểm cân bằng ứng với mức năng lượng cực tiểu (gọi là đáy năng lượng) trên một siêu phẳng năng lượng của siêu diện n chiều. Giả sử đưa vào hệ các mẫu vào (input paterns) tùy ý, hệ sẽ đạt đến điểm cân bằng gần nhất ứng với điểm ổn định [4].

Cấu trúc của mô hình Hopfield ấn định quan hệ giữa phần tử i và j là quan hệ tuyến tính. Số hạng cuối cùng của (9) cho thấy tổng một bậc phản hồi. Mạng như vậy gọi là mạng bậc 1. Nếu tăng số lớp tổng lên hai và các phần tử của nó là bậc hai thì quan hệ giữa các phần tử quan hệ bậc 2 (second-order). Mạng với các phần tử này (không chứa các thành phần bậc nhất) gọi theo tài liệu là mạng bậc hai. Tương tự, có thể tồn tại mạng bậc cao. Mạng gồm cả bậc một và hai chúng ta gọi là mạng tổng quát bậc hai. Tuy nhiên, mạng Hopfield tuyến tính có mặt hạn chế về khả năng phân lớp các mẫu tín hiệu vào và hạn chế đó sẽ được làm sáng tỏ như dưới đây.

2. KHẢ NĂNG PHÂN LỚP

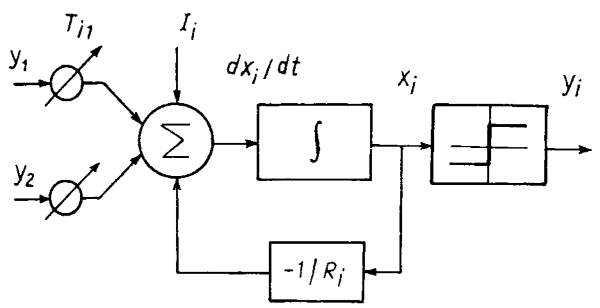
Khả năng phân lớp của mạng Hopfield tuyến tính

Sử dụng (9) với 2 đầu vào phản hồi y_1, y_2 là các trị nhị phân trong không gian véc tơ mẫu vào và $g(\cdot)$ là hàm signum (hình 1a). Các phần tử nơ ron cùng với các giá trị trọng tương ứng chia không gian này thành 2 nhóm giá trị, âm và dương. Đường phân cách giữa chúng được xác định khi trạng thái đạt giá trị ổn định ($dx/dt = 0$) theo điều kiện ngưỡng tới hạn:

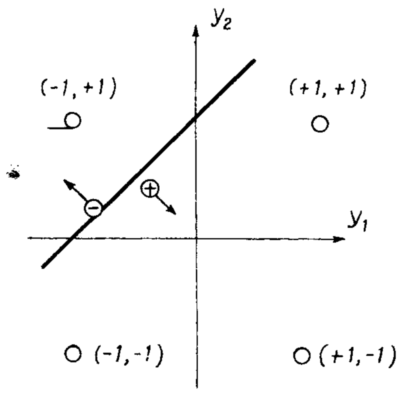
$$C_i X_i = I_i + T_{i1} Y_1 + T_{i2} Y_2 = 0 \tag{26}$$

$$y_2 = -(T_{i1}/T_{i2})y_1 - (I_i/T_{i2}) \tag{27}$$

Hình 1b cho thấy rằng: bằng ba trọng: T_{i1}, T_{i2}, I_i ta có thể tạo được đường thẳng với độ dốc $-(T_{i1}/T_{i2})$ và điểm cắt $-(I_i/T_{i2})$. Mặt trên đường thẳng ứng với tốc độ thay đổi trạng thái âm, mặt dưới dương. Nếu số trọng bằng 4, sự phân cách sẽ tạo được một mặt phẳng, và nếu số trọng lớn hơn 4, giới hạn phân cách là một siêu diện phân cách. (Ở ví dụ trên, để đơn giản cho R_i đủ lớn), $x_i/R_i \approx 0$ khi $dx_i/dt = 0$ nên có thể bỏ qua số hạng đầu của (24')).



Hình 1a



Hình 1b

Khả năng phân lớp của mạng Hopfield phi tuyến

Trong nhiều trường hợp, chúng ta không xác định được đường phân cách tuyến tính bằng một phần tử nơ ron, ví dụ hàm Exclusive NOR với hai đầu vào [5]

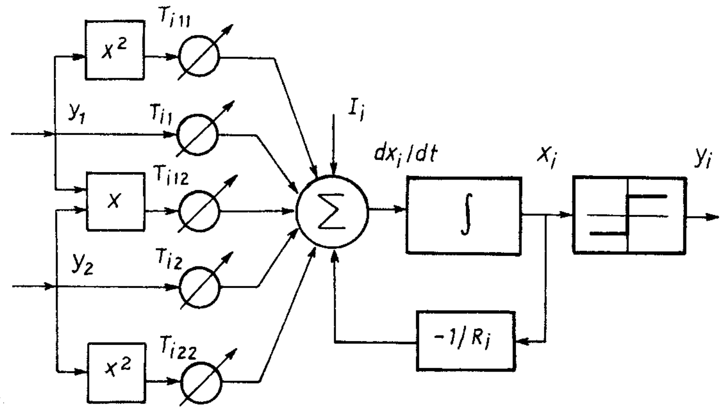
$$\begin{aligned}(+1, +1) &\to +1\\ (+1, -1) &\to +1\\ (-1, -1) &\to +1\\ (-1, +1) &\to -1\end{aligned}$$

Khi số đầu vào càng lớn, khả năng không nhận biết càng lớn hơn. Để khắc phục điều này có hai giải pháp [5] hoặc tăng số phần tử nơ ron để có thêm đường phân cách, tách thêm phần tử mà một nơ ron không đảm dương được. Giải pháp khác là xây dựng một đường cong (mô tả bằng đa thức) bao được các mẫu theo một hàm yêu cầu đặt ra. Đường cong bậc hai ở hình 2b trong ví dụ mô tả giải pháp thứ 2 thu được bằng cách cho thêm một lớp liên kết trọng bậc hai (hình 2a) vào phương trình (9). Cũng lý luận như trên ta có:

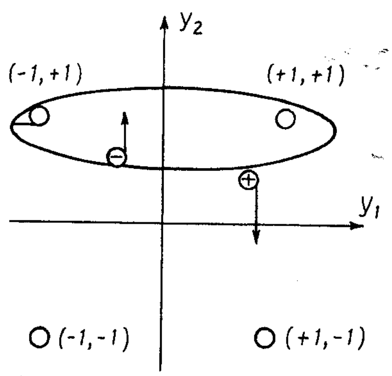
$$C_iX_i = I_i + T_{i1}Y_i + T_{i2}Y_2 + T_{i11}Y_1^2 + T_{i22}Y_2^2 + T_{i12}Y_1Y_2 = 0 \tag{28}$$

Nếu chọn các giá trị trọng thích hợp có thể tạo được một elip như hình 2b (để đơn giản trong sơ đồ không vẽ tham số C_i) đặng có thể bao được các điểm theo yêu cầu. Đây là lời giải cho hàm exclusive NOR.

Việc áp dụng các liên kết phi tuyến như vậy đã được Spetcht [7] áp dụng thành công cho bài toán phân tích và nhận mẫu cũng như đã được Barron [8] và Ivanknenko [9] tiến hành. Từ những nhận thức trên có thể phát biểu rằng việc nghiên cứu những mô hình bậc cao là điều có ý nghĩa thiết thực. Nội dung bài này nhằm mở rộng cấu trúc Hopfield chuẩn có thể suy rộng ra mạng bậc cao và mạng tổng quát bậc cao hơn, đồng thời sẽ chứng minh tồn tại các hàm Liapunov cho chúng. Đối với bài toán điều khiển và nhận dạng tham số, việc quan trọng là đảm bảo mạng ổn định. Do đó, việc nghiên cứu tính ổn định của mạng này là một khâu cần thiết cho một ứng dụng cụ thể và có định hướng.



Hình 2a



Hình 2b

3. MẠNG HOPFIELD TỔNG QUÁT BẬC CAO

Từ nhận định trên, chúng ta xây dựng mạng bậc cao dựa trên mô hình chuẩn Hopfield. Thêm các thành phần bậc cao cho (9) có thể nhận được mạng Hopfield tổng quát bậc cao

$$C_i x_i = -\frac{x_i}{R_i} + I_i + \sum_{j=1}^{n} T_{ij} y_j + \sum_{j=1}^{n}\sum_{k=1}^{n} T_{ijk} y_j y_k + \cdots$$

$$+ \sum_{j=1}^{n} \cdots \sum_{z=1}^{n} T_{ijk\ldots z} y_j \ldots y_z\,,\quad y_k = g_k(x_k),\ i,\ j,\ k, \ldots, z = 1, 2, \ldots, n \qquad (29)$$

Có thể cho hàm Liapunov đối với (29)

$$V(x) = \sum_{i=1}^{n}(1/R_i)\int_0^{y_i} g_i^{-1}(\varsigma)d\varsigma - \sum_{i=1}^{n} I_i y_i - \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} T_{ij}y_iy_j$$

$$-\frac{1}{3}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n} T_{ijk}y_iy_jy_k - \cdots - \frac{1}{m}\sum_{i=1}^{n}\sum_{j=1}^{n}\cdots\sum_{z=1}^{n} T_{ij\ldots z}y_iy_j\ldots y_z \qquad (30)$$

trong đó m là số biến tham gia trong tổng cuối của (30).

Giả sử các điều kiện của định lý 2 đảm bảo và kèm thêm điều kiện bổ sung cho các thành phần bậc cao, có thể tồn tại Liapunov đối với (29).

Mệnh đề 2. *Nếu định lý 2 đảm bảo tức*

a) ma trận $\|T_{ij}\|, \ldots, \|T_{ij\ldots z}\|$ *đối xứng và tất cả* $T_{ij}, \ldots, T_{ij\ldots z} \leq 0$; (31)

b) hàm $g_i(\xi)$ *sigmoid, khả vi, dương, đơn điệu không giảm với mọi* $\xi > 0$ *và hàm ngược của nó* $x_i = g_i^{-1}(y_i)$; (32)

c) $\limsup\limits_{\xi\to\infty}[-x_i/R_i + I_i + T_{ii}y_i(x_i) + T_{iii}y_i(x_i)y_i(x_i) + \cdots] < 0$ *cho mọi* $i = 1, 2, \ldots, n$; (33)

d) $\lim\limits_{\xi\to 0^+}(-X_i(\xi)/R_i + I_i) = \infty$, (34)

hoặc $\lim\limits_{\xi\to 0^+}(-X_i(\xi)/R_i + I_i) < \infty$, (35)

và $\int_0^{\varphi} C_i dx = \infty$ *với* $\varphi > 0$; (36)

thì tồn tại $V(x)$ *của (30) xác định dương.*

Chứng minh. Điều kiện (33) được dùng để ấn định giới hạn vì:

$$-\frac{x_i}{R_i} + I_i + \sum_{j=1}^{n} T_{ij}y_j + \sum_{j=1}^{n}\sum_{k=1}^{n} T_{ijk}y_jy_k + \cdots + \sum_{j=1}^{n}\cdots\sum_{z=1}^{n} T_{ij\ldots z}y_j\ldots y_z \geq$$

$$[-x_i/R_i + I_i + T_{ii}y_i(x_i) + T_{iii}y_i(x_i)y_i(x_i) + \cdots + T_{i\ldots i}y_i(x_i)\ldots y_i(x_i)] \qquad (33')$$

Khi các giả thiết (31) - (36) đảm bảo thì bổ đề 1 (phụ lục A) đảm bảo, ba số hạng đầu của (30) dương. Cần xác định để các thành phần bậc cao của (30) cũng dương. Thật vậy, vì các ma trận $T_{ij}, \ldots, T_{ij\ldots z} \leq 0$, các y_j, $y_k, \ldots, y_z$ không âm, do đó các số hạng sau của chuỗi (30) không âm, nên (30) xác định dương.

Mệnh đề 3. *Nếu hàm*

$$V(x) = \sum_{i=1}^{n}(1/R_i)\int_0^{y_i} g_i^{-1}(\varsigma)d\varsigma - \sum_{i=1}^{n} I_i y_i - \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} T_{ij}y_iy_j$$

$$-\frac{1}{3}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n} T_{ijk}y_iy_jy_k - \cdots - \frac{1}{m}\sum_{i=1}^{n}\sum_{j=1}^{n}\cdots\sum_{z=1}^{n} T_{ij\ldots z}y_iy_j\ldots y_z \qquad (30)$$

xác định dương với (29) thì thỏa mãn $\frac{d}{dt}V(x(t)) \leq 0$ *trên quỹ đạo chấp nhận.*

Chứng minh. Lấy đạo hàm V theo t

$$\frac{dV}{dt} = -\sum_{i=1}^{n}\Big[\frac{1}{2}\Big(\sum_{j=1}^{n} T_{ji}y_i\dot{y}_j\dot{x}_j + \sum_{j=1}^{n} T_{ij}y_j\dot{y}_i\dot{x}_i\Big) - \frac{1}{R_i}g_i^{-1}(y_i)\dot{y}_i\dot{x}_i + I_i\dot{y}_i\dot{x}_i$$
$$+ \frac{1}{3}\Big(\sum_{j,k=1}^{n} T_{ijk}y_iy_k\dot{y}_j\dot{x}_j + \sum_{j,k=1}^{n} T_{ijk}y_jy_k\dot{y}_i\dot{x}_i + \sum_{j,k=1}^{n} T_{ijk}y_iy_j\dot{y}_k\dot{x}_k\Big) + \cdots$$
$$+ \frac{1}{m}\Big(\sum_{j,k,\dots,z=1}^{n} T_{ijk..z}y_iy_k\dots y_z\dot{y}_j\dot{x}_j + \cdots + \sum_{j,k,\dots,z=1}^{n} T_{ijk\dots z}y_iy_j\dots\dot{y}_z\dot{x}_z\Big)\Big] \tag{37}$$

Trong công thức trên, chúng ta ký hiệu các chỉ số dưới của tổng ứng với các tổng cho ngắn gọn. Do tính đối xứng của các ma trận T, chúng ta hoàn toàn có thể đổi chỗ các chỉ số lẫn nhau do đó các số hạng trong dấu ngoặc đơn là bằng nhau. Từ nhận xét trên, sau khi rút gọn ta có:

$$\frac{dV}{dt} = -\sum_{i=1}^{n}\dot{y}_i\dot{x}_i\Big[\sum_{j=1}^{n} T_{ij}y_j - \frac{1}{R_i}g_i^{-1}(y_i) + I_i + \sum_{j,k=1}^{n} T_{ijk}y_jy_k + \cdots$$
$$+ \sum_{j,k,\dots,z=1}^{n} T_{ijk\dots z}y_jy_k\dots y_z\Big] \tag{38}$$

Thay $x_i = g_i^{-1}(y_i)$ vào (38)

$$\frac{dV}{dt} = -\sum_{i=1}^{n}\dot{y}_i\dot{x}_i\Big[\sum_{j=1}^{n} T_{ij}y_j - \frac{X_i}{R_i} + I_i + \sum_{j,k=1}^{n} T_{ijk}y_jy_k + \cdots + \sum_{j,k,\dots,z=1}^{n} T_{ijk\dots z}y_jy_k\dots y_z\Big] \tag{39}$$

Để ý rằng y' không âm do y dương và đơn điệu không giảm, $C_i = \text{const}$ và biểu thức trong dấu ngoặc vuông chính là x' (39) nên:

$$\frac{dV}{dt} = -\sum_{i=1}^{n} C_i\dot{y}_i\dot{x}_i^2 \leq 0 \tag{40}$$

đó là điều cần thiết đã được chứng minh.

Từ định lý 2 và mệnh đề 2, 3 có thể phát biểu định lý sau:

Định lý 3. *Trong hệ Hopfield mở rộng*

$$C_i x_i = -\frac{x_i}{R_i} + I_i + \sum_{j=1}^{n} T_{ij}y_j + \sum_{j=1}^{n}\sum_{k=1}^{n} T_{ijk}y_jy_k + \cdots + \sum_{j=1}^{n}\cdots\sum_{z=1}^{n} T_{ij\dots z}y_j\dots y_z$$

sao cho

a) ma trận $\|T_{ij}\|, \ldots, \|T_{ij\ldots z}\|$ *đối xứng và tất cả* $T_{ik}, \ldots, T_{ik\ldots z} \leq 0$ *cho mọi* $i = 1, 2, \ldots, n;$ (31)

b) hàm $g_i(\xi)$ *khả vi, đơn điệu không giảm, với mọi* $\xi > 0$ *và hàm ngược của nó* $X_i = g_i^{-1}(y_i)$ *cho mọi* $i = 1, 2, \ldots, n;$ (32)

c) $\limsup_{\xi\to\infty}[-x_i/R_i + I_i + T_{ii}y_i(x_i)] < 0$ *cho mọi* $i = 1, 2, \ldots, n;$ (33)

d) $\lim_{\xi\to 0^+}(-X_i(\xi)/R_i + I_i) = \infty,$ (34)

hoặc $\lim_{\xi\to 0^+}(-X_i(\xi)/R_i + I_i) < \infty,$ (35)

và $\int_0^{\varphi} C_i d\xi = \infty$ *với* $\varphi > 0;$ (36)

thì tồn tại hàm Liapunov

$$V(x) = \sum_{i=1}^{n}(1/R_i)\int_0^{y_i} g_i^{-1}(\varsigma)d\varsigma - \sum_{i=1}^{n} I_i y_i - \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} T_{ij}y_iy_j$$
$$- \frac{1}{3}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n} T_{ijk}y_iy_jy_k - \cdots - \frac{1}{m}\sum_{i=1}^{n}\sum_{j=1}^{n}\cdots\sum_{z=1}^{n} T_{ij\ldots z}y_iy_j\ldots y_z \qquad (30)$$

xác định dương và

$$\frac{dV(x)}{dt} = -\sum_{i=1}^{n} C_i\dot{y}_i\dot{x}_i^2 \leq 0 \qquad (40)$$

Việc chứng minh đã tiến hành từ mệnh đề 2 và 3.

Hoàn toàn có thể suy luận rằng tồn tại các cấu trúc khác dựa theo (29), trong đó có thể bỏ bớt một hoặc một số thành phần của bậc nào đó và cũng tồn tại các hàm Liapunov tương ứng. Như vậy, chúng ta có không chỉ một mà là một lớp các cấu trúc mạng theo kiểu Hopfield cũng như tồn tại các hàm Liapunov tương ứng đảm bảo điều kiện cho mạng ổn định. Đây là cơ sở cho khả năng áp dụng chúng rộng hơn vào các ứng dụng thực tế như các bài toán điều khiển chẳng hạn.

4. KẾT LUẬN VÀ CÁC ĐỊNH HƯỚNG NGHIÊN CỨU

- Ý nghĩa và mục đích lớn nhất của bài này là chứng minh ổn định của mạng Hopfield làm cơ sở để giải quyết các bài toán nhận dạng và điều khiển trong các hệ tuyến tính hoặc phi tuyến hoặc các bài toán khác [10].
- Mô hình Grossberg [11, 12] mang tính tổng quát với 1 lớp mạng rộng hơn, nhưng rất khó thực hiện trong thực tế. Trong khi đó mạng Hopfield đơn giản,

có thể thực hiện bằng mạch điện [2] và đã được dùng để giải nhiều bài toán điều khiển và tối ưu nên việc mở rộng này có ý nghĩa thiết thực.

- Bài này cũng góp phần lý luận trong việc mở rộng cấu trúc mạng Hopfield ở dạng tổng quát bậc cao và chứng minh tính ổn định của nó. Tùy những bài toán cụ thể có thể lựa chọn mức độ phức tạp khi bỏ bớt các hạng của (29). Chúng tôi không xét ổn định từng bậc vì chứng minh đã mang tính tổng quát.
- Về khả năng phân lớp tốt hơn như đã phân tích ở mục 2.
- Về khả năng lưu mẫu: mạng Hopfield bậc hai có khả năng lưu mẫu N^2 trong khi mạng Hopfield chuẩn chỉ đạt 0,046 N^2 [2]. Với mạng bậc cao khả năng này chắc sẽ còn lớn hơn và cũng là vấn đề mở.
- Tuy nhiên, mạng bậc cao sẽ có độ phức tạp lớn hơn, tốc độ xử lý cũng có thể chậm hơn. Đối với bài toán nhận mẫu cũng như bài toán điều khiển việc áp dụng mạng Hopfield nhiều lớp chắc sẽ tốt hơn theo nghĩa nào đó. Những nghiên cứu đó là các định hướng tiếp theo.

PHỤ LỤC A

Để chứng minh V(x) là hàm Liapunov, đầu tiên chứng minh quỹ đạo (1) dương bị chặn với dữ kiện ban đầu dương [11].

Bổ đề 1 (Cohen, Grossberg về tính chặn và xác định dương).

Tính chặn: Với $i = 1, 2, \ldots, n$, giả sử:

$$\limsup_{\xi \to \infty} [b_i(\xi) - c_{ii} d_i(\xi)] < 0 \qquad (A1)$$

Xác định cho quỹ đạo với $i = 1, 2, \ldots, n$, giả sử hoặc là

$$\lim_{\xi \to 0^+} b_i(\xi) = \infty \qquad (A2)$$

hoặc là

$$\lim_{\xi \to 0^+} b_i(\xi) < \infty \qquad (A3)$$

và

$$\int_0^{\varphi} \frac{d\xi}{a_i(\xi)} = \infty \quad \text{với } \varphi > 0 \qquad (A4)$$

Thì điều kiện ban đầu dương tạo thành các quỹ đạo dương bị chặn gọi là quỹ đạo chấp nhận.

Chứng minh. Tính chặn được chứng minh, sử dụng điều kiện (A1)

$$b_i(x_i) - \sum_{k=1}^{n} c_{ik}d_k(x_k) \le b_i(x_i) - c_{ii}d_i(x_i) \tag{A5}$$

là hoàn toàn đúng vì mọi c_{ik} và d_k là không âm. Vì $a_i(x_i)$ dương khi x_i lớn, từ (A5) cho thấy rằng $(d/dt)x_i < 0$ khi x_i lớn. Quả vậy, cho điều kiện đầu dương, tồn tại một $L_i < \infty$ sao cho $x_i(t) \le L_i$ khi t đủ lớn với $i = 1, 2, ..., n$.

Điều kiện (A2) được áp dụng cho xác định dương bởi: mỗi hạng trong tổng $\sum_{k=1}^{n} c_{ik}d_k(x_k)$ bị chặn nếu mọi $x_k \le L_k$ với $k = 1, 2, ..., n$, do vậy $b_i(x_i) - \sum_{k=1}^{n} c_{ik}d_k(x_k)$ bị chặn nếu mọi $x_k \le L_x$ với $k = 1, 2, ..., n$ khi $x_i \to 0^+$. Vì $a_i(x_i) > 0$ khi $x_i > 0$, $(d/dt)x_i > 0$ trước khi đạt đến 0, $(d/dt)x_i < 0$ trong khoảng $x \in (0^+ \to \infty)$ do đó x_i không khi nào đạt 0. Khi $(d/dt)x_i$ đổi dấu từ âm sang dương khi x_i thay đổi từ $+\infty$ đến 0^+; trong khoảng $+\infty$ đến 0^+ ắt hẳn phải có một điểm x_i nào đó để $(d/dt)x_i = 0$. Khi $(d/dt)x_i = 0$ thì x_i không còn chuyển động để đạt về 0 (giải thích của tác giả bài này).

Điều kiện (A3) và (A4) nếu đảm bảo để xác định quỹ đạo dương thì khi ấy (giả sử ở thời điểm $t = T$ mà $x_i(T) = 0$ sẽ xảy ra mâu thuẫn)

$$-\infty = \int_{X_{i(0)}}^{0} \frac{d\xi}{a_{i(\xi)}} = \int_{0}^{T} [b_i(x_i(t)) - \sum_{k=1}^{n} c_{ik}d_k(x_k(t))]dt > -\infty \tag{A6}$$

sẽ mâu thuẫn (tích phân trong ngoặc vuông là hữu hạn).

Như vậy $x_i(T)$ luôn đảm bảo dương cho mọi $t \ge 0$.

PHỤ LỤC B (Cohen, Grossberg)

Cũng cần chứng minh $V(x)$ giới hạn và liên tục để đảm bảo các nguyên tắc cho ổn định $V(x)$ [11] như sau:

Mệnh đề 1 (Cohen, Grossberg). *Nếu giả thiết của bổ đề 1 đảm bảo thì $V(x)$ được giới hạn và liên tục trên các quỹ đạo chấp nhận.*

Chứng minh. Nếu (A3) đảm bảo thì tích phân

$$\int_0^{x_i} b_i(\xi_i)d_i(\xi_i)d\xi_i \tag{B1}$$

là chặn bởi vì các quỹ đạo chấp nhận cũng bị chặn. Đồng thời các số hạng

$$\sum_{k=1}^{n} c_{jk} d_j(x_j) d_k(x_k) \tag{B2}$$

cũng bị chặn bởi hàm $d_j(x_j)$ là hàm liên tục của các biến bị chặn. Nếu (A2) đảm bảo, bởi vì

$$\lim_{\xi \to 0^+} |b_i(\xi) d_i(\xi)| \leq \infty \tag{B3}$$

thì lý luận trên là đúng. Nếu (A2) không đảm bảo thì tích phân $\int_0^{x_i}$ có thể được thay thế bằng tích phân $\int_{\lambda_i}^{\varphi_i}$, ở đó λ_i là 1 hằng số dương được chọn theo cách sẽ trình bày dưới đây. Việc chọn như vậy rất có thể xảy ra khi một số tương tác kết hợp với nhau. Mỗi một d_k là hàm không âm và đơn điệu không giảm của biến $0 \leq x_k \leq L_k$ với t đủ lớn, $k = 1, 2, \ldots, n$. Như vậy, tồn tại L hữu hạn dương sao cho

$$\sum_{k=1}^{n} c_{ik} d_k(x_k) \leq L \tag{B4}$$

trên các quỹ đạo dương chấp nhận với thời gian đủ lớn.

Do (A2) đảm bảo, một khoảng $[0, 2\lambda_i]$ nào đó tồn tại sao cho

$$b_i(x_i) - \sum_{k=1}^{n} c_{ik} d_k(x_k) \geq L \tag{B5}$$

như vậy

$$x_i \geq L a_i(x_i) \tag{B6}$$

khi $0 < x_i \leq 2\lambda_i$ trên quỹ đạo chấp nhận tùy ý với thời gian đủ lớn. Vì hàm a_i dương trong khoảng $[x_i(T), 2\lambda_i]$, ở đó $x_i(T) > 0$, do đó a_i có cận dưới dương trong khoảng này. Từ (B6) rút ra T được chọn đủ lớn để đảm bảo (B4) đối với $t \geq T$ thì $x_i(t)$ tăng ít nhất với tốc độ tuyến tính cho đến khi $x_i(t)$ vượt quá λ_i và vẫn lớn hơn λ_i sau. Vì điều này đảm bảo cho 1 quỹ đạo chấp nhận tùy ý, việc chọn λ_i trong tích phân $\int_{\lambda_i}^{x_i}$ được coi là thỏa mãn.

TÀI LIỆU THAM KHẢO

1. K. J. Hunt and Others, *Neural Networks for Control System - A Survey*, Automatica, **28**, No. 6 (1992), 1083 - 1112.

2. Mats Bengtsson, *Higher Order Artificial Neural Networks*, Automatica, **28**, No. 6, (1992), 169-174.

3. Hopfield J. J., *Neurons with graded response have collective computational proprerties like those of two-state neurons*, Proc. of the National Academy of Sciences, **81** (1984) 3088-3092.

4. Behnam Bavaran, *Introduction to Neural Networks for Intelligent Control*, IEEE Control System Magazine, April, 1988.

5. Widrow B., *30 years of Adaptive Neural Networks: Perceptron, Madaline and Backpropagation*, Proc. of the IEEE, Vol. 70, September, 1990.

6. Nicolaos B., Karayiannis Anastasios, N. Venetsanopoulos, *Artificial Neural Networks. Learning Algorithms, Performance Evaluation and Applications*, Kluwer Academic Publishers, Boston-Dodrecht-London, 1993.

7. D. F. Spetch, *Generation of Polynomial Discriminant Function for Partern Recognition*, IEEE Trans. Electron. Comput., **EC-16** (1967), 308-319.

8. A. R. Barron and R. L. Barron, *Statistical Learning Network: An Unifying view*, Sym. on the Interface: Statistics and Computing Science, pp. 192-203, Resion, Va. Apr. 21-23, 1998.

9. A. C. Ivakhnenko, *Polinomial Theory of Complex System*, IEEE Tran. Syst. Man, Cybernetics, SMC-1, p. 364-378, 1971.

10. Yun-Qui LEI, Zhang Zhi, and Liang-Qui Zhang, *A Neural Adaptive PID Controller*, Dep. of Automatic Control. Changaha Institute of Technology.

11. Michael A. Cohen and Stephen Grossberg, *Absolute Stability of Global Pattern Formation and Parallel Memory Storage by Competitive Neural Networks*, IEEE on System, Man and Cybernetics, Vol. 13, No. 5, 1983.

12. Grossberg S., *Neural Networks: Principles, Mechanism, and Architectures*, Neural Networks, **1** (1988), 17-61.

13. Katsuhiko Ogata, *Modern Control Engineering*, Prentic-Hall International Editions.

14. J. J. Hopfield, *Neural Networks and Physical System with Emergent Collective Computational Abilities*, Proc. of the National Academy of Sciences, **70** (1982), 2554-2558.

Phân viện Tự động,
Viện Công nghệ thông tin

Nhận bài ngày 6-12-1995