Proceedings of the
44th IEEE Conference on Decision and Control, and
the European Control Conference 2005
Seville, Spain, December 12-15, 2005

A New Characterization on the Approximation of Nonlinear Functions via Neural Networks: An Adaptive Control Perspective

Tomohisa Hayakawa
CREST, Japan Science and Technology Agency, Saitama, 332-0012, JAPAN
tomohisa_hayakawa@ipc.i.u-tokyo.ac.jp

Abstract

A novel neural network approximation scheme that is especially appropriate for adaptive control of nonlinear dynamical systems is proposed. In light of the new function approximation characterization, a neuro adaptive control framework for continuous- and discrete-time nonlinear uncertain dynamical systems is also presented. Specifically, the proposed neural network control framework is Lyapunov-based and, unlike standard neural network controllers guaranteeing ultimate boundedness, the framework guarantees partial asymptotic stability of the closed-loop system, that is, asymptotic stability with respect to part of the closed-loop system states associated with the system plant dynamics. The neuro adaptive controllers are constructed without requiring explicit knowledge of the system dynamics other than the assumption that the plant dynamics are continuous and piecewise continuously differentiable, and that the approximation error of uncertain system nonlinearities lie in a small gain-type norm bounded conic sector. This allows us to show that the standard neural network controllers are in fact capable of achieving partial asymptotic stability around the system equilibrium point for continuous- and discrete-time uncertain systems.

1. Introduction

There has been a tremendous amount of effort to develop adaptive control framework using neural networks (see, for example, [1–3] for continuous-time cases and [4–6] for discrete-time cases, to cite but a few) ever since it was shown that the neural network is a universal approximator [7, 8]. One of the key ideas behind the neuro adaptive control is to cancel uncertain nonlinear elements in the plant dynamics via neural networks online and generate an additional control signal based on the known part of the system dynamics.

In approximating general nonlinear functions via *finite* linear combination of activation functions, it is inevitable that there remains residual approximation error between the actual and the approximated nonlinear functions. The approximation error is conventionally evaluated in the sense of infinity norm, that is, the worst approximation error over the domain of interest. In the field of neural network control, this fact naturally forces us to construct a positive-definite (Lyapunov-like) function for a given dynamical system and to show that the Lyapunov derivative along the closed-loop system trajectories is negative *outside* an open set that contains the targeted system equilibrium point. As a consequence, this uncertainty characterization yields *ultimate boundedness* [9, 10] of the closed-loop system, instead of *stability*.

To claim ultimate boundedness (even for noiseless systems), it is necessary to add to weight update laws 'damping terms' such as σ- or e-modification terms which prevent the neural network weighting gains from blowing up. One can easily surmise that these additional terms make the computation of Lyapunov derivatives complicated (especially for discrete-time systems) and induce conservativeness. On top of that, that is not to say that the Lyapunov methods provide only sufficient conditions for judging stability of nonlinear systems.

In recent papers [11, 12], a novel characterization for approximating nonlinear functions was proposed. The approach was distinct from the ones in the conventional neural network literature in that the nonlinear function was transformed into the form of a linear function with a state-dependent coefficient matrix, and the coefficient matrix was then approximated via neural networks. Consequently, the approximation error was assessed in the sense of *Lipschitz norm*, that is, the size of the small gain-type conic sector. This uncertainty characterization certainly allows us to achieve asymptotic stability with respect to the plant states without involving the additional damping terms in the weight update laws. The control laws and update laws of [11, 12], however, involve Kronecker products, which make the dimension of adaptive weighting gains substantially large.

In this paper we take a similar approach to the one given in [11, 12] in characterizing approximation errors. In particular, our formulation builds on the results given in [13] which shows that the multilayer neural network is capable of approximating nonlinear functions as well as their derivatives. By approximating the *derivative* of the nonlinear function rather than the function itself, we are now allowed to show that the approximation error can be made artitrarily small in the sense of Lipschitz norm and is thus contained in a small gain-type *norm bounded* conic sector with arbitrarily small sector bound. Accordingly, we can characterize neuro adaptive control laws that guarantee partial asymptotic stability of the closed-loop system, that is, Lyapunov stability of the overall closed-loop states and convergence with respect to the plant state. The neuro adaptive controllers are constructed *without* requiring explicit knowledge of the system dynamics other than the fact that the plant dynamics are continuous and piecewise continuously differentiable.

We emphasize in this paper that we take the *conventional* form of the control laws and update laws without damping (modification) terms in the update laws. The difference is that the stability property is drawn by viewing the neural network as an approximator of nonlinear functions in a certain way that has not been considered in the literature. To focus our attention to this point, we consider a simple problem in this paper, that is, neural network control with the full-state feedback and matched uncertainties. It is expected that our approach can be easily extended to deal with a variety of more general cases since the control architecture is the same as in the literature.

The notation used in this paper is fairly standard. Specifically, $\mathbb{R}$ denotes the set of real numbers, $\mathbb{N}_0$ denotes the set of nonnegative integers, $(\cdot)^{\mathrm{T}}$ denotes transpose, $\mathrm{tr}(\cdot)$ denotes the trace operator, $\|\cdot\|$ denotes the Euclidean vector norm, and $\lambda_{\max}(\cdot)$ (resp., $\lambda_{\min}(\cdot)$) denotes the maximum (resp., minimum) eigenvalue of a Hermitian matrix. Furthermore, we write $\varphi'(x)$ for the Fréchet derivative of φ at x and $\overset{\circ}{\mathcal{D}}$ for the interior of the set $\mathcal{D}$.

2. Nonlinear Function Approximation

In this section we propose a new treatment of approximating nonlinear functions suitable for developing neuro adaptive control laws that can guarantee *asymptotic stability* of nonlinear uncertain dynamical systems. To this end, we first review some of the important results given in [13] which is necessary for our main theorems of this paper. For the statements of this section let $\mathrm{C}(\mathbb{R}^n)$ be the set of all functions $f:\mathbb{R}^n\to\mathbb{R}$ such that $f(\cdot)$ is continuous on $\mathbb{R}^n$. Furthermore, for $l\in\mathbb{N}_0$, we define $\mathrm{C}^l(\mathbb{R}^n)\triangleq\{f:\mathbb{R}^n\to\mathbb{R}:\mathrm{D}^\alpha f\in\mathrm{C}(\mathbb{R}^n),\ |\alpha|\le l\}$, where $\alpha\triangleq(\alpha_1,\ldots,\alpha_n)$ denotes a multi-index with respect to the differential operator D^α defined by

$$\mathrm{D}^\alpha\triangleq\frac{\partial^{|\alpha|}}{\partial^{\alpha_1}x_1\cdots\partial^{\alpha_n}x_n},\qquad(1)$$

with order $|\alpha|\triangleq\alpha_1+\cdots+\alpha_n$.

Definition 2.1 [13]. Let $l\in\mathbb{N}_0$ be given. Then a function $\sigma:\mathbb{R}^n\to\mathbb{R}$ is called *l-finite* if $\sigma\in\mathrm{C}^l(\mathbb{R})$ and $0<\int_{\mathbb{R}^n}\|\mathrm{D}^l\sigma(x)\|\lambda(\mathrm{d}x)<\infty$, where λ is the Lebesgue measure on the measure space $(\mathbb{R}^n,\mathcal{B}(\mathbb{R}^n))$ and $\mathcal{B}(\mathbb{R}^n)$ is the Borel σ-algebra generated by the open subsets of $\mathbb{R}^n$.

The following theorem states that the single-hidden-layer feedforward neural network (see Figure 2.1) is capable of approximating nonlinear functions as well as their derivatives.

Theorem 2.1 [13]. Let $\delta:\mathbb{R}^n\to\mathbb{R}$ be a measurable function and let $\gamma\in\mathbb{R}$ be a given positive constant. Then for any compact set $\mathcal{D}_{\mathrm{c}}\subset\mathbb{R}^n$, there exist an l-finite activation function $\sigma:\mathbb{R}^n\to\mathbb{R}^s$ and a weighting matrix $W\in\mathbb{R}^s$ such that

$$\|\mathrm{D}^\alpha\delta(x)-\mathrm{D}^\alpha(W^{\mathrm{T}}\sigma(x))\|<n^{-\frac{1}{2}}\gamma^{-1},\quad\text{a.e. }x\in\mathcal{D}_{\mathrm{c}},\quad|\alpha|\le l,\qquad(2)$$

provided that the dimension of activation functions, s, is sufficiently large.

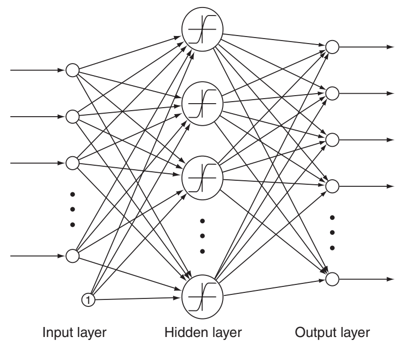

Figure 2.1: Single-hidden-layer feedforward neural network

In Theorem 2.1, $\mathrm{D}^\alpha\delta(x)$ is defined as a distributional (generalized) derivative [14]. The class of functions that are l-finite includes the usual squashing functions such as logistic, hyperbolic tangent, radial basis functions, splines, etc., which are commonly used in the neural network control literature.

For our purposes, we are particularly interested in Theorem 2.1 with the special case of $l=1$. In this case, the restatement of Theorem 2.1 is as follows: for any compact set $\mathcal{D}_{\mathrm{c}}\subset\mathbb{R}^n$, there exist an 1-finite activation function $\sigma:\mathbb{R}^n\to\mathbb{R}^s$ and a weighting matrix $W\in\mathbb{R}^s$ such that $\varphi(x)\triangleq\delta(x)-W^{\mathrm{T}}\sigma(x)$ satisfies

$$\|\varphi'(x)\|<\gamma^{-1},\quad\text{a.e. }x\in\mathcal{D}_{\mathrm{c}},\qquad(3)$$

provided that the dimension of activation functions, s, is sufficiently large.

Using the special case of Theorem 2.1 described by (3), we show that given a measurable function, there always exist a weighting matrix and an activation function such that approximation error is contained in a small gain-type norm bounded conic sector with arbitrarily small sector bound. The precise statement is given in the following theorem.

Theorem 2.2. Let $\delta:\mathbb{R}^n\to\mathbb{R}$ be a function such that $\delta(\cdot)$ is continuous and piecewise continuously differentiable, and let $\gamma\in\mathbb{R}$ be a given positive constant. Then for any convex compact set $\mathcal{D}_{\mathrm{c}}\subset\mathbb{R}^n$ such that $0\in\overset{\circ}{\mathcal{D}}_{\mathrm{c}}$, there exist a smooth activation function $\sigma:\mathbb{R}^n\to\mathbb{R}^s$ and a weighting matrix $W\in\mathbb{R}^s$ such that $\varphi(x)\triangleq\delta(x)-W^{\mathrm{T}}\sigma(x)$ satisfies

$$\varphi^{\mathrm{T}}(x)\varphi(x)\le\gamma^{-2}x^{\mathrm{T}}x,\quad x\in\mathcal{D}_{\mathrm{c}},\qquad(4)$$

provided that the number of activation functions, s, is sufficiently large.

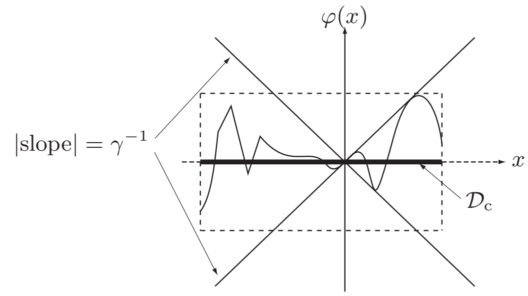

Figure 2.2: Visualization of function $\varphi(\cdot)$ for $n = 1$

Proof: First, note that there exist $W_1^{\mathrm{T}} \in \mathbb{R}^{s_1}$ and a 1-finite function $\sigma_1 : \mathbb{R}^n \to \mathbb{R}^{s_1}$ such that $\delta(0) = W_1^{\mathrm{T}}\sigma_1(0)$. Furthermore, it follows from Theorem 2.1 that there exist $W_2 \in \mathbb{R}^{s_2}$ and a 1-finite function $\sigma_2 : \mathbb{R}^n \to \mathbb{R}^{s_2}$ such that $\sigma_2(0) = 0$ and (3) holds with $W = [W_1^{\mathrm{T}}, W_2^{\mathrm{T}}]^{\mathrm{T}}$ and $\sigma(x) = [\sigma_1^{\mathrm{T}}(x), \sigma_2^{\mathrm{T}}(x)]^{\mathrm{T}}$. Next, consider the function $\varphi(\nu x)$, $\nu \in [0, 1]$, $x \in \mathcal{D}_c$. Since $\varphi(\cdot)$ is piecewise continuously differentiable, it follows that

$$\frac{\mathrm{d}}{\mathrm{d}\nu}\varphi(\nu x) = \varphi'(\nu x)x, \quad \text{a.e. } x \in \mathcal{D}_c. \tag{5}$$

Note that since $\mathcal{D}_c$ is convex, $\nu x \in \mathcal{D}_c$ for all $x \in \mathcal{D}_c$ and $\nu \in [0, 1]$. Now, integrating both sides of (5) in ν from 0 to 1 yields

$$\varphi(x) - \varphi(0) = \int_0^1 \varphi'(\nu x)\mathrm{d}\nu \cdot x, \quad x \in \mathcal{D}_c, \tag{6}$$

where the integral in (6) is well defined. Hence, since $\varphi(0) = 0$, it follows from (3) and the Cauchy-Schwartz inequality that

$$\begin{aligned} \varphi^{\mathrm{T}}(x)\varphi(x) &= x^{\mathrm{T}}\left[\int_0^1 \varphi'(\nu x)\mathrm{d}\nu\right]^{\mathrm{T}}\left[\int_0^1 \varphi'(\nu x)\mathrm{d}\nu\right]x \\ &\le \left[\int_0^1 \|\varphi'(\nu x)\|\mathrm{d}\nu\right]^2 x^{\mathrm{T}}x \\ &\le \left[\int_0^1 \gamma^{-1}\mathrm{d}\nu\right]^2 x^{\mathrm{T}}x \\ &\le \gamma^{-2}x^{\mathrm{T}}x, \quad x \in \mathcal{D}_c, \end{aligned} \tag{7}$$

which completes the proof. □

Remark 2.1. The approximation error induced in Theorem 2.2 corresponds to a nonlinear small gain-type norm bounded uncertainty characterization (see Figure 2.2). Note that in the proof of Theorem 2.2 we used Theorem 2.1 for the case where the first derivative (gradient) of $\delta(\cdot)$ can be arbitrarily well approximated; in the conventional neural network control it is assumed that the unknown function itself (zeroth derivative of the function) is approximated arbitrarily well. This conceptual difference leads to the substantial difference between the results of this paper (asymptotic stability) and the conventional results (ultimate boundedness) in the context of neural network adaptive control.

Using Theorem 2.2, in the following sections we construct neuro adaptive control laws for continuous- and discrete-time systems, respectively, that guarantee asymptotic stability with respect to the plant states.

3. Stable Neuro Adaptive Control for Nonlinear Uncertain Systems

Based on the discussion in Section 2 on the nonlinear function approximation, in this section we characterize neural adaptive feedback control laws for nonlinear uncertain dynamical systems that guarantee asymptotic stability, rather than ultimate boundedness, with respect to the plant states. Specifically, consider the controlled nonlinear uncertain dynamical system $\mathcal{G}$ given by

$$\dot{x}(t) = f(x(t)) + G(x(t))u(t), \quad x(0) = x_0, \quad t \ge 0, \tag{8}$$

where $x(t) \in \mathbb{R}^n$ is the state vector, $u(t) \in \mathbb{R}^m$ is the control input, $f : \mathbb{R}^n \to \mathbb{R}^n$ is continuous and piecewise continuously differentiable on $\mathbb{R}^n$ with $f(0) = 0$, and $G : \mathbb{R}^n \to \mathbb{R}^{n\times m}$. For the nonlinear uncertain system $\mathcal{G}$ we assume that the required properties for the existence and uniqueness of solutions are satisfied, that is, $f(\cdot)$, $G(\cdot)$, and $u(\cdot)$ satisfy sufficient regularity conditions such that (8) has a unique solution forward in time.

In this paper, we assume that $f(\cdot)$ is an unknown function, and $f(\cdot)$ and $G(\cdot)$ are given by

$$f(x) = Ax + \Delta f(x), \tag{9}$$
$$G(x) = BG_n(x), \tag{10}$$

where $A \in \mathbb{R}^{n\times n}$ and $B \in \mathbb{R}^{n\times m}$ are known matrices, $G_n : \mathbb{R}^n \to \mathbb{R}^{m\times m}$ is a known matrix function such that $\det G_n(x) \ne 0$, $x \in \mathbb{R}^n$, and $\Delta f : \mathbb{R}^n \to \mathbb{R}^n$ is an uncertain function belonging to the uncertainty set $\mathcal{F}$ given by

$$\begin{aligned} \mathcal{F} = \{&\Delta f : \mathbb{R}^n \to \mathbb{R}^n : \Delta f(0) = 0, \\ &\Delta f(x) = B\delta(x),\, x \in \mathbb{R}^n\}, \end{aligned} \tag{11}$$

where $\delta : \mathbb{R}^n \to \mathbb{R}^m$ is an uncertain, continuous, piecewise continuously differentiable function on $\mathbb{R}^n$. It is important to note that since $\delta(\cdot)$ is continuously differentiable and $\delta(0) = 0$, it follows from Theorem 2.2 that there exist an activation function $\sigma : \mathbb{R}^n \to \mathbb{R}^s$ and the (optimal) weighting matrix $W \in \mathbb{R}^{s\times m}$ such that

$$\varphi(x) \triangleq \delta(x) - W^{\mathrm{T}}\sigma(x), \tag{12}$$

satisfies (4).

Theorem 3.1. Consider the nonlinear uncertain dynamical system $\mathcal{G}$ given by (8) where $f(\cdot)$ and $G(\cdot)$ are given by (9) and (10), respectively, and $\Delta f(\cdot)$ belongs to $\mathcal{F}$. Assume that (A, B) is stabilizable and let $K \in \mathbb{R}^{m\times n}$ be such that $A_s \triangleq A + BK$ is Hurwitz. Furthermore, for a given $\gamma > 0$, assume there exist positive-definite matrices $P \in \mathbb{R}^{n\times n}$ and $R \in \mathbb{R}^{n\times n}$ such that

$$0 = A_s^{\mathrm{T}}P + PA_s + \gamma^{-2}PBB^{\mathrm{T}}P + I_n + R. \tag{13}$$

Finally, let $Q \in \mathbb{R}^{m\times m}$ and $Y \in \mathbb{R}^{s\times s}$ be positive definite. Then the neural adaptive feedback control law

$$u(t) = G_{\mathrm{n}}^{-1}(x(t))\Big[Kx(t) - \hat{W}^{\mathrm{T}}(t)\sigma(x(t))\Big], \quad (14)$$

where $\hat{W}^{\mathrm{T}}(t) \in \mathbb{R}^{m\times s}$ and $\sigma : \mathbb{R}^n \to \mathbb{R}^s$ is a given basis function such that $\sigma(0) = 0$, with update law

$$\dot{\hat{W}}^{\mathrm{T}}(t) = QB^{\mathrm{T}}Px(t)\sigma(x(t))^{\mathrm{T}}Y, \quad \hat{W}^{\mathrm{T}}(0) = \hat{W}_0^{\mathrm{T}}, \quad (15)$$

guarantees that there exists a compact, positively invariant set $\mathcal{D}_\alpha \subset \mathbb{R}^n \times \mathbb{R}^{m\times s}$ such that $(0, W^{\mathrm{T}}) \in \mathcal{D}_\alpha$, where $W^{\mathrm{T}} \in \mathbb{R}^{m\times s}$, and the solution $(x(t), \hat{W}^{\mathrm{T}}(t))$ of the closed-loop system given by (8), (14), (15) is Lyapunov stable and $x(t) \to 0$ as $t \to \infty$ for all $\Delta f(\cdot) \in \mathcal{F}$ and $(x_0, \hat{W}_0^{\mathrm{T}}) \in \mathcal{D}_\alpha$.

Proof. The proof is omitted due to space limitations. □

Remark 3.1. Note that the conditions in Theorem 3.1 imply partial asymptotic stability of the closed-loop system, that is, the solution $(x(t), \hat{W}^{\mathrm{T}}(t)) \equiv (0, W^{\mathrm{T}})$ of the overall closed-loop system is Lyapunov stable and $x(t) \to 0$ as $t \to \infty$. Hence, it follows from (15) that $\dot{\hat{W}}^{\mathrm{T}}(t) \to 0$ as $t \to \infty$.

Remark 3.2. In the case where the neural network approximation in the sense of (3) holds on $\mathbb{R}^n$, since the Lyapunov function used in the proof of Theorem 3.1 is a class $\mathcal{K}_\infty$ function, it follows that the control law (14) ensures global asymptotic stability with respect to x. If the uncertain function $\delta(\cdot)$ possesses a linear growth rate, then it is likely to satisfy (3) on $\mathbb{R}^n$ since in this case $\delta'(\cdot)$ is uniformly bounded over the whole space $\mathbb{R}^n$. To guarantee global asymptotic stability, this assumption is much weaker than the assumption that the uncertain function is perfectly approximated over the whole space $\mathbb{R}^n$, which has to be made in the conventional uncertainty error characterization. However, the existence of a global neural network approximator for an uncertain nonlinear map cannot in general be established. Yet, for a given arbitrarily large compact set $\mathcal{D}_{\mathrm{c}} \subset \mathbb{R}^n$, there exists an approximator of the derivative of the unknown nonlinear map up to a desired accuracy, and hence our neuro adaptive controller guarantees *semiglobal* partial asymptotic stability.

Remark 3.3. Note that the neuro adaptive controller (14) and (15) can be constructed to guarantee partial asymptotic stability using standard *linear* H_∞ theory. Specifically, it follows from standard H_∞ theory [15] that $\|G(s)\|_\infty < \gamma$, where $G(s) = E(sI_n - A_{\mathrm{s}})^{-1}B$ and E is such that $E^{\mathrm{T}}E = I_n + R$, if and only if there exists a positive-definite matrix P satisfying the bounded real Riccati equation (13). It is well known that (13) has a positive-definite solution if and only if the Hamiltonian matrix

$$\mathcal{H} = \begin{bmatrix} A_{\mathrm{s}} & \gamma^{-2}BB^{\mathrm{T}} \\ -E^{\mathrm{T}}E & -A_{\mathrm{s}}^{\mathrm{T}} \end{bmatrix}, \quad (16)$$

has no purely imaginary eigenvalues.

It is important to note that the adaptive control law (14) and (15) does not require the explicit knowledge of the optimal weighting matrix W. Furthermore, no specific structure on the nonlinear dynamics $f(x)$ is required to apply Theorem 3.1 other than the fact that $f(x)$ is continuous and piecewise continuously differentiable. However, if (8) is in normal form [16], then we can always construct a neuro adaptive control law *without* requiring knowledge of the system dynamics $f(x)$. To see this, assume that the nonlinear uncertain system $\mathcal{G}$ is generated by

$$q_i^{(r_i)}(t) = f_{\mathrm{u}i}(q(t)) + \sum_{j=1}^{m} G_{\mathrm{s}(i,j)}(q(t))u_j(t), \quad t \geq 0, \quad i = 1, \cdots, m, \quad (17)$$

where $q = [q_1, \cdots, q_1^{(r_1-1)}, \cdots, q_m, \cdots, q_m^{(r_m-1)}]^{\mathrm{T}}$, $q(0) = q_0$, $q_i^{(r_i)}$ denotes the r_ith derivative of q_i, and r_i denotes the relative degree with respect to the output q_i. Here we assume that the square matrix function $G_{\mathrm{s}}(q)$ composed of the entries $G_{\mathrm{s}(i,j)}(q)$, $i, j = 1, \cdots, m$, is such that $\det G_{\mathrm{s}}(q) \neq 0$, $q \in \mathbb{R}^{\hat{r}}$, where $\hat{r} = r_1 + \cdots + r_m$ is the (vector) relative degree of (17) and $\hat{r} = n$. Furthermore, we assume that $f_{\mathrm{u}i}(\cdot)$ is continuous and piecewise continuously differentiable on $\mathbb{R}^n$, and $f_{\mathrm{u}i}(0) = 0$.

Next, define $x_i \triangleq \Big[q_i, \cdots, q_i^{(r_i-2)}\Big]^{\mathrm{T}}$, $i = 1, \cdots, m$, $x_{m+1} \triangleq \Big[q_1^{(r_1-1)}, \cdots, q_m^{(r_m-1)}\Big]^{\mathrm{T}}$, and $x \triangleq \big[x_1^{\mathrm{T}}, \cdots, x_{m+1}^{\mathrm{T}}\big]^{\mathrm{T}}$, so that (17) can be described by (8) with

$$A = \begin{bmatrix} A_0 \\ 0_{m\times n} \end{bmatrix}, \quad \Delta f(x) = \begin{bmatrix} 0_{(n-m)\times 1} \\ f_{\mathrm{u}}(x) \end{bmatrix}, \quad G(x) = \begin{bmatrix} 0_{(n-m)\times m} \\ G_{\mathrm{s}}(x) \end{bmatrix}, \quad (18)$$

where $A_0 \in \mathbb{R}^{(n-m)\times n}$ is a known matrix of zeros and ones capturing the multivariable controllable canonical form representation [17], $f_{\mathrm{u}} : \mathbb{R}^n \to \mathbb{R}^n$ is an unknown function and satisfies $f_{\mathrm{u}}(0) = 0$, and $G_{\mathrm{s}} : \mathbb{R}^n \to \mathbb{R}^{m\times m}$. Note that $\Delta f(\cdot) \in \mathcal{F}$ with $B = [0_{m\times(n-m)}, I_m]^{\mathrm{T}}$ and $\delta(x) = f_{\mathrm{u}}(x)$. In this case, $G_{\mathrm{n}}(x) \equiv G_{\mathrm{s}}(x)$. Furthermore, since A is in multivariable controllable canonical form, we can always construct K such that $A + BK$ is Hurwitz.

4. Stable Neuro Adaptive Control for Discrete-Time Nonlinear Uncertain Systems

In this section, we develop a similar framework to the framework presented in Section 3 for *discrete-time* nonlinear uncertain systems. Specifically, consider the controlled nonlinear uncertain dynamical system $\mathcal{G}$ given by

$$x(k+1) = f(x(k)) + G(x(k))u(k), \quad x(0) = x_0, \quad k \in \mathbb{N}_0, \quad (19)$$

where $x(k) \in \mathbb{R}^n$ is the state vector, $u(k) \in \mathbb{R}^m$ is the control input, $f : \mathbb{R}^n \to \mathbb{R}^n$ is continuous, piecewise continuously differentiable, and satisfies $f(0) = 0$, and $G : \mathbb{R}^n \to \mathbb{R}^{n\times m}$.

As in Section 3, we assume that $f(\cdot)$ is an unknown function and $f(\cdot)$ and $G(\cdot)$ are given by (9) and (10) where $A \in \mathbb{R}^{n\times n}$ and $B \in \mathbb{R}^{n\times m}$ are known matrices, $G_{\rm n} : \mathbb{R}^n \to \mathbb{R}^{m\times m}$ is a known matrix function such that $\det G_{\rm n}(x) \neq 0$, $x \in \mathbb{R}^n$, and $\Delta f : \mathbb{R}^n \to \mathbb{R}^n$ is an uncertain function belonging to the uncertainty set $\mathcal{F}$ given by (11) where $\delta : \mathbb{R}^n \to \mathbb{R}^m$ is an uncertain, continuous, piecewise continuously differentiable function such that $\delta(0) = 0$. As discussed in Section 3, since $\delta(x)$ is piecewise continuously differentiable and $\delta(0) = 0$, it follows that there exist an activation function $\sigma : \mathbb{R}^n \to \mathbb{R}^s$ and the (optimal) weighting matrix $W \in \mathbb{R}^{s\times m}$ such that the approximation error $\varphi(x)$ defined by (12) satisfies (4).

Theorem 4.1. Consider the nonlinear uncertain dynamical system $\mathcal{G}$ given by (19) where $f(\cdot)$ and $G(\cdot)$ are given by (9) and (10), respectively, and $\Delta f(\cdot)$ belongs to $\mathcal{F}$ given by (11). Assume that (A, B) is stabilizable and, for a given $\gamma > 0$, let $P \in \mathbb{R}^{n\times n}$ be the positive-definite solution to the Riccati equation

$$P = A^{\rm T}PA - A^{\rm T}PB(B^{\rm T}PB)^{-1}B^{\rm T}PA + (\alpha+\beta)I_n + R, \quad (20)$$

where $R \in \mathbb{R}^{n\times n}$ is positive definite, $\alpha > 0$, and β satisfies

$$\beta \geq \gamma^{-2}\Big(\lambda_{\max}(B^{\rm T}PB) + a\frac{1 + x^{\rm T}Px}{c + \sigma^{\rm T}(x)\sigma(x)}\Big), \quad x \in \mathcal{D}_{\rm c}, \quad (21)$$

$$a = \max\{c, n/\lambda_{\min}(P)\}B^{\rm T}PB\Big(I_m + \tfrac{1}{\alpha\gamma^2}B^{\rm T}PB\Big), \quad (22)$$

and $c > 0$. Then the neural adaptive feedback control law

$$u(k) = G_{\rm n}^{-1}(x(k))\Big[Kx(k) - \hat{W}^{\rm T}(k)\sigma(x(k))\Big], \quad (23)$$

where $\hat{W}^{\rm T}(k) \in \mathbb{R}^{m\times s}$, $\sigma : \mathbb{R}^n \to \mathbb{R}^s$ is a given basis function, and

$$K = -(B^{\rm T}PB)^{-1}B^{\rm T}P \quad (24)$$

with update law

$$\hat{W}^{\rm T}(k+1) = \hat{W}^{\rm T}(k) + \tfrac{1}{c+\sigma^{\rm T}(x(k))\sigma(x(k))}B^{\dagger}[x(k+1) - A_{\rm s}x(k)]\sigma^{\rm T}(x(k)), \quad \hat{W}^{\rm T}(0) = \hat{W}_0^{\rm T}, \quad (25)$$

guarantees that there exists a compact, positively invariant set $\mathcal{D}_\alpha \subset \mathbb{R}^n \times \mathbb{R}^{m\times s}$ such that $(0, W^{\rm T}) \in \mathcal{D}_\alpha$, where $W^{\rm T} \in \mathbb{R}^{m\times s}$, and the solution $(x(k), \hat{W}^{\rm T}(k))$ of the closed-loop system given by (19), (23), (25) is Lyapunov stable and $x(k) \to 0$ as $k \to \infty$ for all $\Delta f(\cdot) \in \mathcal{F}$ and $(x_0, \hat{W}_0^{\rm T}) \in \mathcal{D}_\alpha$.

Proof. The proof is omitted due to space limitations. □

Remark 4.1. The conditions in Theorem 4.1 imply partial asymptotic stability the closed-loop system, that is, the solution $(x(k), \hat{W}^{\rm T}(k)) \equiv (0, W^{\rm T})$ of the overall closed-loop system is Lyapunov stable and $x(k) \to 0$ as $k \to \infty$. Hence, it follows from (25) that $\hat{W}^{\rm T}(k+1) - \hat{W}^{\rm T}(k) \to 0$ as $k \to \infty$.

5. Illustrative Numerical Example

In this section we present a numerical example to demonstrate the utility of the proposed neuro adaptive control framework for adaptive stabilization. Specifically, consider the uncertain controlled Liénard system given by

$$\ddot{q}(t) + c(q(t))\dot{q}(t) + k(q(t)) = bu(t), \quad q(0) = q_0, \quad \dot{q}(0) = \dot{q}_0, \quad t \geq 0, \quad (26)$$

where $c : \mathbb{R} \to \mathbb{R}$ and $k : \mathbb{R} \to \mathbb{R}$ are unknown, continuously differential functions. Note that with $x_1 = q$ and $x_2 = \dot{q}$, (26) can be written in state space form (8) and (4) with $x = [x_1, x_2]^{\rm T}$, $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$, $\Delta f(x) = [0, -c(x_1)x_2 - k(x_1)]^{\rm T}$, $B = [0, b]^{\rm T}$, and $G_{\rm n}(x) = 1$. Here, we assume that the unknown function $\Delta f(x)$ can be written as $\Delta f(x) = B\delta(x)$, where $\delta(x) = \frac{1}{b}[-c(x_1)x_2 - k(x_1)]$ is an unknown, continuously differentiable function. Next, let $K = \frac{1}{b}[k_1, k_2]$, where k_1, k_2 are arbitrary scalars, so that $A_{\rm s} = A + BK = \begin{bmatrix} 0 & 1 \\ k_1 & k_2 \end{bmatrix}$. Now, with the proper choice of k_1 and k_2, it follows from Theorem 3.1 that if there exists $P > 0$ satisfying (13), then the neuro adaptive feedback controller (14) guarantees that $x(t) \to 0$ as $t \to \infty$. Specifically, here we choose $k_1 = -1$, $k_2 = -1$, $\gamma = 3$, and $R = I_2$, so that P satisfying (13) is given by

$$P = \begin{bmatrix} 3.1586 & 1.0627 \\ 1.0627 & 2.3765 \end{bmatrix}. \quad (27)$$

With $c(x_1) = 2(x_1^4 - 1)$, $k(x_1) = x_1 + \tanh(x_1)$, $b = 3$, $Q = 1$, $Y = 0.1I_{12}$, $\sigma(x) = \Big[\frac{1}{1+e^{-a_1x_1}}, \cdots, \frac{1}{1+e^{-3a_1x_1}}, \frac{1}{1+e^{-a_2x_2}}, \cdots, \frac{1}{1+e^{-3a_2x_2}}\Big]$, where $a_1 = a_2 = 0.5$, and initial conditions $x(0) = [1, 1]^{\rm T}$ and $\hat{W}(0) = 0_{12\times1}$, Figure 5.3 shows the phase portrait of the controlled and uncontrolled system. Note that the neuro adaptive controller is switched on at $t = 10$ sec. Figure 5.4 shows the state trajectories versus time and the control signal versus time. Finally, Figure 5.5 shows the neural network weighting functions versus time.

6. Conclusion

A novel approach of characterizing nonlinear function approximation was presented. As a result, the standard neuro adaptive control framework for continuous- and discrete-time nonlinear uncertain dynamical systems was shown to guarantee asymptotic stability, rather than ultimate boundedness, with respect to the plant

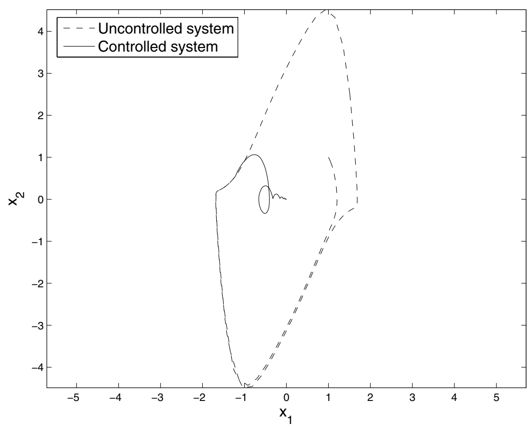

Figure 5.3: Phase portrait of controlled and uncontrolled Liénard system

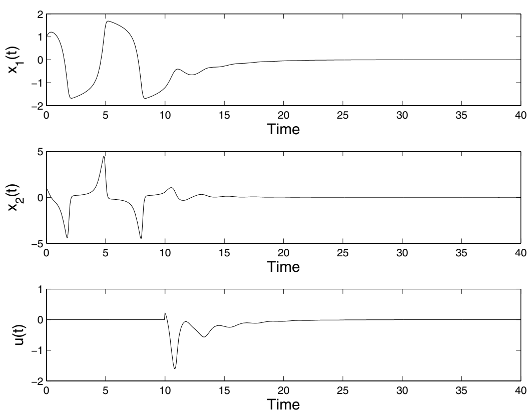

Figure 5.4: State trajectories and control signal versus time

states. Furthermore, in the case where the nonlinear system is represented in normal form, the neuro adaptive controllers were constructed without requiring knowledge of the system dynamics other than the fact that the plant dynamics are continuous and piecewise continuously differentiable. The proposed uncertainty characterization considerably reduces the conservativeness of the Lyapunov approach.

Acknowledgment. The author would like to thank Dr. E. Lavretsky for several stimulating discussions on approximation theory.

References

[1] F. L. Lewis, S. Jagannathan, and A. Yesildirak, *Neural Network Control of Robot Manipulators and Nonlinear Systems.* London, UK: Taylor & Francis, 1999.

[2] S. S. Ge, C. C. Hang, T. H. Lee, and T. Zhang, *Stable adaptive neural network control.* Boston, MA: Kluwer Academic Publishers, 2001.

[3] J. Spooner, M. Maggiore, R. Ordonez, and K. Passino, *Stable Adaptive Control and Estimation for Nonlinear Systems: Neural and Fuzzy Approximator Techniques.* New York, NY: John Wiley & Sons, 2002.

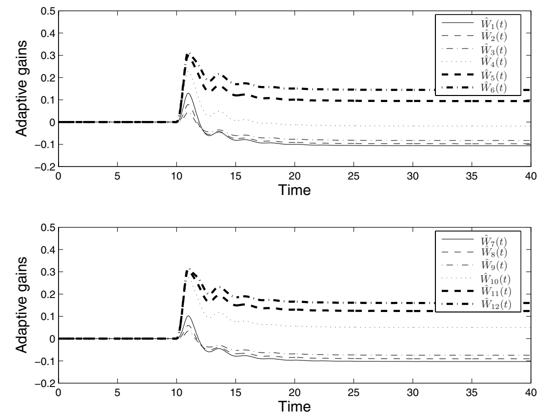

Figure 5.5: Neural network weighting functions versus time

[4] F. C. Chen and H. K. Khalil, "Adaptive control of a class of nonlinear discrete-time systems using neural networks," *IEEE Trans. Autom. Contr.*, vol. 40, no. 5, pp. 791–801, 1995.

[5] S. Jagannathan and F. L. Lewis, "Discrete-time neural net controller for a class of nonlinear dynamical systems," *IEEE Trans. Autom. Contr.*, vol. 41, no. 11, pp. 1693–1699, 1996.

[6] S. S. Ge, T. H. Lee, G. Y. Li, and J. Zhang, "Adaptive NN control for a class of discrete-time non-linear systems," *Int. J. Contr.*, vol. 76, no. 4, pp. 334–354, 2003.

[7] K. Funahashi, "On the approximate realization of continuous mappings by neural networks," *Neural Networks*, vol. 2, no. 3, pp. 183–192, 1989.

[8] K. Hornik, M. Stinchcombe, and H. White, "Multilayer feedforward networks are universal approximators," *Neural Networks*, vol. 2, no. 5, pp. 359–366, 1989.

[9] H. K. Khalil, *Nonlinear Systems.* Upper Saddle River, NJ: Prentice-Hall, 3 ed., 2002.

[10] W. M. Haddad, V. Chellaboina, Q. Hui, and T. Hayakawa, "Neural network adaptive control for discrete-time nonlinear nonnegative dynamical systems," in *Proc. IEEE Conf. Dec. Contr.*, (Maui, HI), December 2003.

[11] T. Hayakawa, W. M. Haddad, and N. Hovakimyan, "Neural network adaptive control for nonlinear uncertain dynamical systems with asymptotic stability guarantees," in *Proc. Amer. Contr. Conf.*, (Portland, OR), pp. 1301–1306, June 2005.

[12] T. Hayakawa, W. M. Haddad, and N. Hovakimyan, "A new characterization of stable neural network control for discrete-time uncertain systems," in *Proc. IFAC World Cong.*, (Prague, Czech Republic), July 2005.

[13] K. Hornik, M. Stinchcombe, and H. White, "Universal approximation of an unknown mapping and its derivatives using multilayer feedforward networks," *Neural Networks*, vol. 3, pp. 551–560, 1990.

[14] E. H. Lieb and M. Loss, *Analysis.* Providence, RI: American Mathematical Society, 2001.

[15] J. C. Willems, "Least squares stationary optimal control and the algebraic Riccati equation," *IEEE Trans. Autom. Contr.*, vol. 16, no. 6, pp. 621–634, 1971.

[16] A. Isidori, *Nonlinear Control Systems.* New York, NY: Springer, 1995.

[17] C.-T. Chen, *Linear System Theory and Design.* New York: Holt, Rinehart, and Winston, 1984.