Approximation of Discrete-Time State-Space Trajectories Using Dynamic Recurrent Neural Networks

Liang Jin, Peter N. Nikiforuk, and Madan M. Gupta

Abstract—In this note, the approximation capability of a class of discrete-time dynamic recurrent neural networks (DRNN’s) is studied. Analytical results presented show that some of the states of such a DRNN described by a set of difference equations may be used to approximate uniformly a state-space trajectory produced by either a discrete-time nonlinear system or a continuous function on a closed discrete-time interval. This approximation process, however, has to be carried out by an adaptive learning process. This capability provides the potential for applications such as identification and adaptive control.

I. INTRODUCTION

Recently, much success has been achieved in the use of feedforward neural networks for identification and control. It has been shown that this type of neural networks, which describes the static input-output mapping, is capable of approximating not only a continuous function but also its derivatives to an arbitrary degree of accuracy. There is, however, increasing interest in studying the approximation capability of dynamic recurrent neural networks (DRNN’s) which are represented by nonlinear dynamic systems, so that they are applicable for control, robotics, and pattern recognition.

A dynamic recurrent neural network (DRNN) contains both a huge number of feedforward and feedback synaptic connections and has complex dynamics [4], [5], [17], [18]. The power of the DRNN’s for dealing with the problem of associative memories is well known. Also, the DRNN’s have been proved to be capable of processing time-varying spatio-temporal information. From the computational point of view, a dynamic neural structure which contains a state feedback may provide more computational advantages than that of a purely feedforward neural structure. For some problems, a small feedback system is equivalent to a large and possibly infinite feedforward system [9]. A well-known example is that an infinite number of feedforward logic gates are required to emulate an arbitrary finite-state machine. Also, an infinite-order finite impulse response (FIR) filter is required to emulate a single-pole infinite impulse response (IIR) filter [9].

Most theoretical studies on DRNN’s have been mainly concentrated on the stability and convergence of the network trajectory to the equilibria. It is worth mentioning, however, some reported studies on the approximation theory using DRNN’s. Funahashi and Nakamura [3] studied the approximation of continuous-time dynamic systems using a Hopfield-type DRNN. Li [13] studied the problem in the discrete-time domain and showed that a discrete-time trajectory on a closed finite interval may be represented exactly using a discrete-time DRNN. He also proved the approximation capability of a class of continuous-time DRNN’s. To implement such a neural approximation process, dynamic learning algorithms for updating the weights of the DRNN’s have been proposed by Williams and Zipser [19] for discrete-time case, and by Pearlmutter [16], Narendra and Parthasarthy [14] for continuous-time systems. Most recently, Olutorni [15] presented a weight learning algorithm for the DRNN’s.

 Manuscript received November 7, 1994; revised January 30, 1995.

The authors are with the Intelligent Systems Research Laboratory, College of Engineering, University of Saskatchewan, Saskatoon, Saskatchewan, Canada S7N 0W0.

IEEE Log Number 9411658.

which has a feedforward complexity. The successful applications of identification, control, and filtering using the DRNN’s are closely related to the approximation capability of the DRNN’s.

The main objective of this note is to exploit the approximation capability of a class of discrete-time DRNN’s. The proofs used in this note are based on the well-known results of the universal approximation of multilayered feedforward neural networks provided by Cybenko [1], Funahashi [2] and Hornik, Stinchcombe, and White [6], [7]. The model of a discrete-time DRNN is given in Section II. Section III presents some preliminaries and a method of embedding approximately a special class of discrete-time nonlinear systems into the higher dimensional DRNN’s. Section IV shows that such a discrete-time DRNN is capable of uniformly approximating a discrete-time state-space trajectory which is defined on a closed interval and produced by either a nonlinear dynamic system or a continuous function to an arbitrary precision. Section V contains some conclusions.

II. DYNAMIC RECURRENT NEURAL NETWORKS

A dynamic recurrent neural network (DRNN) is a complex nonlinear dynamic system described by a set of nonlinear differential or difference equations with extensive connection weights. In this note, only discrete-time version of analog DRNN’s with time-varying inputs is discussed. A general expression of this type of DRNN’s with N neural units is given by the following discrete-time nonlinear system:

\[ \begin{align*}
  \dot{x}(k+1) &= -ax(k) + f(Ax(k), B, u(k)) \\
  y(k) &= Cx(k)
\end{align*} \]

where \( x \in \mathbb{R}^N, y \in \mathbb{R}^m, \) and \( u \in \mathbb{R}^n \) are the neural state, output, and input vectors, respectively. \( A \in \mathbb{R}^{N \times N}, B \in \mathbb{R}^{N \times m}, \) and \( C \in \mathbb{R}^{m \times N} \) are the connection weight matrices associated with the neural state, input, and output vectors, respectively. \( a \) is a fixed constant for controlling state decay and is chosen as \( -1 < a < 1, \) and \( f: \mathbb{R}^N \times \mathbb{R}^m \rightarrow \mathbb{R}^m \) is appropriately chosen vector-valued nonlinear function.

Some commonly used choices of the nonlinear function \( f \) are given as follows:

i) modified Hopfield type: \( f(Ax, B, u) = Ax + Bu \)

ii) modified Pineda type I: \( f(Ax, B, u) = \sigma(Ax + Bu) \)

iii) modified Pineda type II: \( f(Ax, B, u) = \sigma(Ax) + Bu \)

where \( \sigma(z) \) may be chosen as a continuous and differentiable nonlinear sigmoidal function satisfying the following conditions: i) \( \sigma(z) \rightarrow \pm 1 \) as \( x \rightarrow \pm \infty, \) ii) \( \sigma(z) \) is bounded with the upper bound 1 and the lower bound -1, iii) \( \sigma(z) = 0 \) at a unique point \( z = 0, \) iv) \( \sigma'(z) > 0 \) and \( \sigma'(z) \rightarrow 0 \) as \( z \rightarrow \pm \infty, \) and v) \( \sigma'(z) \) has a global maximal value \( \mu > 0. \) Typical examples of such a function \( \sigma(.) \) are

\[ \sigma(\mu) = \tanh \left( \frac{1 - e^{-\mu x}}{1 + e^{-\mu x}} \right) + \frac{2}{\pi} \tan^{-1} \left( \frac{\pi \mu x}{2} \right) \]

where \( \mu > 0 \) is a constant which determines the slope or activation gain of \( \sigma(z) \). In this note, one assumes that the sigmoidal function is the hyperbolic tangent function; that is, \( \sigma(z) = \tanh(z) \).

These dynamic neural structures incorporated with some adaptive learning processes have been used effectively to solve the problems of identification and control [10], [12], [14]. A basic question often arising from these applications, however, is the approximation capability of such neural models. In fact, this capability not only

1995 IEEE
defines the suitability of the models but it is also essential to ensure the success of the applications. In this note, a discrete-time DRN is proposed as follows

\[
\begin{align*}
\dot{x}(k+1) &= -\alpha x(k) + A\sigma(x(k)) + Bu(k), \quad x \in \mathbb{R}^n \\
y(k) &= Cx(k), \\
\end{align*}
\]

which may be viewed as a dynamic neural network with single hidden layer which contains N dynamic neurons as basic information processing units. In the following discussions, the issue of approximation capability of such a dynamic neural system will be studied extensively.

III. EMBEDDING THE NONLINEAR SYSTEMS INTO THE DRN'S

To exploit the approximation problem of the DRN (2), some preliminaries will be provided in this section. Let both \( S \subset \mathbb{R}^n \) and \( U \subset \mathbb{R}^m \) be open subsets. A mapping \( f: S \times U \rightarrow \mathbb{R}^n \) is said to be Lipschitz in \( x \) on \( S \times U \) if there exists a constant \( L \) such that for all \( x_1, x_2 \in S \), and any \( u \in U \), and \( L \) is a Lipschitz constant of \( f(x, u) \) in \( x \). \( f \) is locally Lipschitz in \( x \) if each point of \( S \) has a neighborhood \( S_0 \subset S \) such that the restriction \( f \) to \( S_0 \times U \) is Lipschitz in \( x \).

**Lemma 1:** Let \( S \subset \mathbb{R}^n \) and \( U \subset \mathbb{R}^m \) be open sets and a mapping \( f: S \times U \rightarrow \mathbb{R}^n \) be Lipschitz continuous mappings, \( L \) be a Lipschitz constant of \( f(x, u) \) in \( x \) on \( S \times U \), and for all \( x \in S \) and \( u \in U \)

\[
||f(x, u) - f(x, u)|| \leq L||x_1 - x_2||
\]

for all \( x_1, x_2 \in S \), and any \( u \in U \), and \( L \) is a Lipschitz constant of \( f(x, u) \) in \( x \). \( f \) is locally Lipschitz in \( x \) if each point of \( S \) has a neighborhood \( S_0 \subset S \) such that the restriction \( f \) to \( S_0 \times U \) is Lipschitz in \( x \).

**Proof:** See Hirsch and Smale [8].

**Lemma 2:** Let \( S \subset \mathbb{R}^n \) and \( U \subset \mathbb{R}^m \) be open sets, \( f, f_1: S \times U \rightarrow \mathbb{R}^n \) be Lipschitz continuous mappings, \( L \) be a Lipschitz constant of \( f(x, u) \) in \( x \) on \( S \times U \), and for all \( x \in S \) and \( u \in U \)

\[
||f(x, u) - f(x, u)|| < \epsilon.
\]

If \( x(k) \) and \( z(k) \) are, respectively, solutions of the following difference equations

\[
x(k+1) = f(x(k), u(k))
\]

and

\[
z(k+1) = f(z(k), u(k))
\]

with an initial condition \( x(0) = z(0) \in S \), then

\[
||x(k) - z(k)|| < \epsilon a_{k-1}, \quad k \geq 0
\]

where \( a_k = 1 + L a_{k-1} \) with \( a_0 = 0 \).

**Proof:** First, one shows that (3) is true when \( k = 1 \) since

\[
||x(1) - z(1)|| = ||f(x(0), u(0)) - f(z(0), u(0))||
\]

\[
< ||f(x, u) - f(x, u)|| < \epsilon.
\]

Let (3) be true for \( (k - 1) \); that is

\[
||x(k) - z(k)|| < \epsilon a_{k-1}.
\]

Next, one needs to show that (3) is true for \( k \). In fact

\[
||x(k) - z(k)||
\]

\[
= ||f(x(k - 1), u(k - 1)) - f(z(k - 1), u(k - 1))||
\]

\[
= ||f(x(k - 1), u(k - 1)) - f(z(k - 1), u(k - 1)) + f(z(k - 1), u(k - 1)) - f(z(k - 1), u(k - 1))||
\]

\[
< \epsilon + ||f(x(k - 1), u(k - 1)) - f(z(k - 1), u(k - 1))||
\]

Note that for an arbitrary \( x(0) = z(0) \in S \), one has \( x(k) \in S \) and \( z(k) \in S \), and \( f \) is Lipschitz in \( x \) for all \( x \in S \) and \( u \in U \), thus

\[
||f(x(k - 1), u(k - 1)) - f(z(k - 1), u(k - 1))|| < L ||x(k - 1) - z(k - 1)||
\]

implies

\[
||x(k) - z(k)|| < \epsilon + L a_{k-1} = \epsilon a_k.
\]

Thus the lemma is proved.

Next, a method of embedding approximately a special class of discrete-time nonlinear systems into the higher dimensional systems with the form of the DRN's will be provided. First, the following lemma presents a basic method of approximating a constant vector using a DRN with an appropriate initial condition.

**Lemma 3:** Let \( \theta = [\theta_1, \theta_2, \ldots, \theta_n]^T \) be a constant vector. Then, there exists an \( n \)-th order discrete-time dynamic system of the form

\[
x(k + 1) = -\alpha x(k) + \theta = \theta
\]

with \(-1 \leq \alpha \leq 1\) and an appropriate initial condition \( x(0) \) such that for a given number \( \epsilon > 0 \)

\[
||x(k) - \theta|| < \epsilon, \quad \text{for all } k \geq 0
\]

**Proof:** When \( \alpha = -1 \), one may design a trivial \( n \)-th order linear system

\[
x(k + 1) = x(k), \quad x(0) = \theta
\]

and all the eigenvalues of the Jacobian matrix of system (4) at this point are located inside the unit circle on the complex plane. Obviously, there exist multiple solutions of such a \( W \). Next, one will find one of the solutions. Let \( W = \text{diag}[w_{i1}, w_{i2}, \ldots, w_{in}] \); that is, system (4) is fully decoupled. Without loss of generality, let all \( \theta_i \neq 0 \), \( 1 \leq i \leq n \). One may simply select

\[
w_{ii} = (1 + \alpha) - \theta_i \tanh (\theta_i) > 0.
\]

Then \( \theta \) is an equilibrium point of the system with such a weight matrix \( W \). Furthermore, the Jacobian of the \( i \)-th equation of the system (4) at this equilibrium point is

\[
\frac{\partial J_i}{\partial x_i} \bigg|_{x_i = \theta_i} = (-\alpha + w_{ii} \frac{\sigma'(x_i)}{\cosh^2(x_i)}) \bigg|_{x_i = \theta_i} = -\alpha + (1 + \alpha) \frac{2\theta_i}{\sinh(2\theta_i)}
\]

which implies

\[
\frac{\partial J_i}{\partial x_i} \bigg|_{x_i = \theta_i} < 1.
\]

Thus \( \theta \) is an asymptotically stable equilibrium point. Given a small number \( \epsilon > 0 \), one may find a \( \delta = \delta(\epsilon) > 0 \) and a neighborhood \( K = \{x \in \mathbb{R}^n : ||x - \theta|| < \delta\} \) of the point \( \theta \) in \( \mathbb{R}^n \) such that for any initial condition \( x(0) \in K \) the solution of the system (4) satisfies

\[
||x(k) - \theta|| < \epsilon \quad \text{for } k \geq 0.
\]
Lemma 4: Consider an nth-order discrete-time nonlinear system of the form
\[
\begin{align*}
\eta_{k+1} &= -a_0 \eta(k) + W_1 \sigma(W_2 \eta(k) + W_3 u(k) + \theta), \\
\eta(0) &= \eta_0,
\end{align*}
\] (6)
where \(-1 \leq a_0 \leq 1\), \(W_1 \in \mathbb{R}^{N_f \times n}, W_2 \in \mathbb{R}^{N_f \times n}, W_3 \in \mathbb{R}^{N_f \times n}\), and \(\theta \in \mathbb{R}^{N_f}\). For a given number \(\epsilon > 0\) and an integer \(0 < I < +\infty\), there exist an integer \(N\) and a DRNN of the form (2) with an appropriate initial state \(\eta(0) = \eta_0\) such that for any bounded input \(u(k)\)
\[
||\eta(k) - \eta_0|| < \epsilon, \quad 0 \leq k \leq I.
\]

Proof: For the case of \(k = -1\), using a linear coordinate transformation \(\eta_1(k) = \eta_1(k) + W_1 \sigma(\eta_2(k) + W_3 u(k)), \eta_1(k) \in \mathbb{R}^{N_f}\) with an initial condition
\[
\eta_1(0) = W_2 \eta_0 + \theta = W_2 \eta_0 + \theta.
\]

Introducing an \((n + N_f)\)-dimensional augmented coordinate vector \(x = [\eta_1^T, \eta_2^T] \in \mathbb{R}^{(n+N_f)}\), (6) and (7) may then be rewritten as follows
\[
x(k + 1) = x(k) + A x(k) + Bu(k), \quad x \in \mathbb{R}^{(n+N_f)}
\] (8)
where the weight matrices are given by
\[
A = \begin{bmatrix} 0 & W_1 \\ W_2 & W_3 \end{bmatrix}, \quad B = \begin{bmatrix} W_3 \\ W_3 \end{bmatrix}
\]
and an initial condition is given by
\[
x(0) = \begin{bmatrix} \eta_0 \\ W_2 \eta_0 + \theta \end{bmatrix}.
\]

If an output equation
\[
y(k) = C x(k), \quad C = [I, 0]
\]
is employed, one may see
\[
y(k) = \eta_1(k), \quad k \geq 0.
\]

In the case of \(-1 < a_0 \leq 1\), one denotes that \(a_0 = 0\) and \(a_k = 1 + ||W_1||\sqrt{n_{a_k-1}}\). Then \(a_k \leq a_{k+1}\) for \(k \geq 0\). Using Lemma 3, it can be seen that there exists a dynamic system of the form
\[
x_{k+1} = -a x_k + W_4 \sigma(x_k), \quad x_k \in \mathbb{R}^{N_f}
\] (9)
with \(W_4 \in \mathbb{R}^{N_f\times N_f}\) and an initial condition \(x_0(0) \in \mathbb{R}^{N_f}\) such that for a given number \(\epsilon > 0\)
\[
||x_0(0) - \theta|| < \epsilon, \quad \epsilon = \frac{\epsilon}{\sqrt{n_{a_{k-1}}}}, \quad \text{for } k \geq 0.
\]
Next, consider an nth-order dynamic system of the form
\[
\begin{align*}
\eta_{k+1} &= -a_0 \eta(k) + W_1 \sigma(W_2 \eta(k) + W_3 u(k) + \theta), \\
\eta(0) &= \eta_0,
\end{align*}
\] (10)

Denote the right-hand side function of the above system as \(\phi(\eta_1, W_1, W_2, W_3, x_1) = -a_0 \eta(k) + W_1 \sigma(W_2 \eta(k) + W_3 u(k) + x_1)\).

Then \(\phi(\eta_1, W_1, W_2, W_3, \theta) = \phi(\eta_1, W_1, W_2, W_3, \theta)\) is the right-hand side function of the system (6). It is easily verified that
\[
||\phi(\eta_1, W_1, W_2, W_3, x_1) - \phi(\eta_1, W_1, W_2, W_3, \theta)||
\]
\[
= ||W_1 \sigma(W_2 \eta + W_3 u + x_1) - W_1 \sigma(W_2 \eta + W_3 u + \theta)||
\]
\[
< ||W_1|| \sqrt{||x_1 - x_2||}
\]
\[
= \epsilon, \quad \epsilon = \frac{\epsilon}{\sqrt{n_{a_{k-1}}}}
\]

Thus it is clear by using Lemma 2 that
\[
||\eta(k) - \eta_0|| < \epsilon, \quad 0 \leq k \leq I.
\]

Next, introducing a linear transformation \(x_2 = W_1 \eta_1 + x_1\) yields
\[
x_{k+1} = -a_0 x_{k+1} + W_1 \sigma(x_{k+1} + W_3 u(k)) + W_4 \sigma(x_{k+1}) + x_2 \in \mathbb{R}^N
\] (11)
with an initial condition
\[
x_{2}(0) = W_2 \eta_1(0) + x_1(0) = W_2 \eta_0(0) + x_1(0).
\]

Furthermore, introducing an \((n + 2 N_f)\)-dimensional augmented coordinate vector \(x = [\eta_2^T, \eta_3^T] \in \mathbb{R}^{(n+2N_f)}\), then (9)-(11) can be rewritten as follows
\[
x_{k+1} = -a_0 x_{k+1} + A x_{k+1} + Bu(k), \quad x \in \mathbb{R}^{(n+2N_f)}
\]
where the weight matrices are given by
\[
A = \begin{bmatrix} 0 & 0 & W_1 \\ W_4 & W_3 & W_1 \end{bmatrix}, \quad B = \begin{bmatrix} W_3 \\ 0 \\ W_3 \end{bmatrix}
\]
and an initial condition for the above system is given by
\[
x(0) = \begin{bmatrix} \eta_2(0) \\ \eta_3(0) \\ x_1(0) \end{bmatrix}
\]
If an output \(y \in \mathbb{R}^n\) is produced at the instant \(k\) as follows
\[
y(k) = C x(k), \quad C = [I, 0, 0]
\]
one obtains
\[
||y(k) - \eta_0|| < \epsilon, \quad 0 \leq k \leq I.
\]

The above proof shows that if the self-feedback coefficient \(a = -1\), the embedding procedure discussed in Lemma 4 may be exactly realized.

IV. APPROXIMATION OF STATE-SPACE TRAJECTORIES USING DRNN’S

In this section, the approximation capability of the DRNN for the state trajectories of the discrete-time nonlinear systems will be studied. The main results will be presented in the following two theorems.

Theorem 1: Let \(S \subset \mathbb{R}^n\) and \(U \subset \mathbb{R}^m\) be open sets, \(D \subset S\) and \(D_u \subset U\) be compact sets, \(Z \subset D_u\) be an open set, and \(f: S \times U \rightarrow \mathbb{R}^n\) be a continuous vector-valued function. For a discrete-time nonlinear system of the form
\[
z(k + 1) = f(z(k), u(k)), \quad z \in \mathbb{R}^n, u \in \mathbb{R}^m
\] (12)
with an initial state \(z(0) \in Z\), whose solution \(z(k) \in D_u\), then, for an arbitrary \(\epsilon > 0\) and an integer \(0 < I < +\infty\), there exist an integer \(N\) and a DRNN of the form (2) with an appropriate initial state \(\eta(0)\) such that for any bounded input \(u: \mathbb{R}^+ = [0, +\infty) \rightarrow D_u\)
\[
\max_{0 \leq k \leq I} ||z(k) - y(k)|| < \epsilon.
\]

(13)
Proof: Using Lemma 1, one knows that \( f(z, u) \) is Lipschitz in \( z \) on \( D_x \times D_u \) with a constant \( L \). One defines

\[
 a_k = 1 + L a_k - 1, \quad \text{with} \ a_0 = 0.
\]

Obviously, \( a_k \neq 0 \) and \( a_k \leq a_{k+1} \) for \( k \geq 1 \). Given an arbitrary number \( \epsilon > 0 \), define

\[
 \epsilon_1 = \frac{\epsilon}{a_1}.
\]

Using the universal approximation theorem of multilayered feedforward neural networks, one knows that there exist an integer \( N \) and weight matrices \( W_1 \in \mathbb{R}^{n \times N}, W_2 \in \mathbb{R}^{2 \times N}, W_3 \in \mathbb{R}^{m \times N} \) and a threshold vector \( \theta \in \mathbb{R}^{N} \) such that

\[
 \| f(z, u) + \alpha z - W_1 \sigma(W_2 z + W_3 u + \theta) \| < \epsilon_1
\]

for all \( z \in D_x \) and \( u \in D_u \). Define a continuous vector-valued function \( g: \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^n \) as follows

\[
 g(z, u) = -\alpha z + W_1 \sigma(W_2 z + W_3 u + \theta)
\]

then (14) may be rewritten as

\[
 \max_{x \in D_x \cap \eta} \| f(z, u) - g(z, u) \| < \epsilon_1.
\]

Assume that \( z \in D_x \subset S \) and \( \eta \in \mathbb{R}^n \) are the solutions of the following difference equations

\[
 z(k+1) = f(z(k), u(k))
\]

\[
 \eta(k+1) = g(\eta(k), u(k))
\]

with an initial condition \( z(0) = \eta(0) = x_0 \in Z \). By the assumption, all solutions \( z(k) \in D_x \). Thus, by using Lemma 2

\[
 \| z(k) - \eta(k) \| < \epsilon a_k \leq \epsilon a_1 = \epsilon
\]

for any \( 0 \leq k \leq I \). Next, one considers the following dynamic system:

\[
 \eta(k+1) = -\alpha \eta(k) + W_1 \sigma(W_2 \eta(k) + W_3 u(k) + \theta)
\]

\[
 \eta(0) = x_0.
\]

The remaining part of the proof is easily obtained by using Lemma 4.

Proof: Since the vector-valued function \( f \) of the system (2) is continuous, for any initial condition \( z(0) \), the set of discrete-time trajectories of the nonlinear dynamic system (19) defined as

\[
 Z = \{ z(k) \in \mathbb{R}^n : z(0) \in D_x, 0 \leq k \leq I \}
\]

is a compact subset of \( S \). Thus from Theorem 1, one obtains the results.

The above proofs reflect a constructive way to form a discrete-time DRNN with the form (2) which has universal approximation capability for state-space trajectories of discrete-time nonlinear dynamic systems. There is no information, however, regarding the exact number of hidden neural units which is needed to achieve such a task of approximation. If the state equation of the DRNN (2) is decomposed as follows

\[
 \begin{aligned}
 z_1(k+1) = & -\alpha z_1(k) + A_1 \sigma(z_1(k) + B_1 u(k)) \\
 z_2(k+1) = & -\alpha z_2(k) + A_2 \sigma(z_2(k) + B_2 u(k)) \\
 z_3(k+1) = & -\alpha z_3(k) + A_3 \sigma(z_3(k) + B_3 u(k))
 \end{aligned}
\]

where \( z_1 \in \mathbb{R}^n, z_2 \in \mathbb{R}^{N-n}, \) and

\[
 A = \begin{bmatrix} A_1 & A_2 & A_3 \end{bmatrix}
\]

\[
 B = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}
\]

Then, the approximation (20) in Theorem 2 may be rewritten as

\[
 \max_{0 \leq k \leq I} \| z(k) - z_1(k) \| < \epsilon
\]

that is, the first \( n \) states of the DRNN (21) may be used to uniformly approximate the state-space trajectory \( z(k) \) of the dynamic system (19). On the other hand, the weight matrices \( A \) and \( B \) involved such a DRNN have to be determined through an adaptive learning process according to the target dynamic system. Next, some consequences of the above theorem will be discussed.

Corollary 1: Let \( S \subset \mathbb{R}^n \) be an open set, \( D_x \subset S \) be a compact set, and \( f: S \times \mathbb{R} \rightarrow \mathbb{R}^m \) be a continuous vector-valued function which defines the following nonautonomous nonlinear system

\[
 z(k+1) = f(x(k), k), \quad z \in \mathbb{R}^n
\]

with an initial state \( z(0) \). Then, for an arbitrary number \( \epsilon > 0 \) and an integer \( 0 < I < +\infty \), there exist an integer \( N \) and a DRNN of the form

\[
 \begin{aligned}
 x(k+1) = & -\alpha x(k) + A \sigma(x(k)), \quad x \in \mathbb{R}^N \\
 y(k) = & C x(k), \quad y \in \mathbb{R}^n
 \end{aligned}
\]

with an appropriate initial state \( x(0) \) such that

\[
 \max_{0 \leq k \leq I} \| x(k) - y(k) \| < \epsilon
\]

Proof: Let \( z_{n+1} = k \) and an \( (n+1) \)-dimensional vector \( \tilde{x} = [\tilde{x}, \tilde{z}_{n+1}] \). Then, the nonlinear system (22) may be represented as the following autonomous form

\[
 \tilde{z}(k+1) = \tilde{f}(\tilde{z}(k))
\]

where

\[
 \tilde{f}(\tilde{z}) = \begin{bmatrix} f(z, z_{n+1}) \\ 1 + z_{n+1} \end{bmatrix}
\]

with an initial condition \( \tilde{z}(0) = [x^T(0), 0] \). Also, it is obviously that for \( 0 < I < +\infty \), the set of trajectories defined as

\[
 \tilde{Z} = \{ \tilde{z} \in \mathbb{R}^{n+1} : \tilde{z}(0) = [x^T(0), 0], z(0) \in D_x, 0 \leq k \leq I \}
\]

is a compact set of \( \mathbb{R}^{n+1} \). Using Theorem 2, there exists an \( N \)-dimensional DRNN with \( (n+1) \)-dimensional output vector \( y \in \mathbb{R}^{n+1} \).
which can uniformly approximate the state $z$ of the augmented system (25) on the finite closed interval $0 \leq k \leq I < +\infty$; that is

$$\|z(k) - y(k)\| < \epsilon.$$ 

Thus the first $n$ elements of $y$, which are represented by the vector $y \in \mathbb{R}^n$, can uniformly approximate $z$ on the interval $0 \leq k \leq I < +\infty$; that is

$$\|z(k) - y(k)\| < \epsilon.$$ 

\[\square\]

**Corollary 2:** Let $f: \mathbb{R} \rightarrow \mathbb{R}^n$ be a continuous function, and $f(k)$ $(0 \leq k \leq I < +\infty)$ be discrete-time trajectories. Then, for an arbitrary number $\epsilon > 0$, there exist an integer $N$ and a DRNN of the form (23) with an appropriate initial state $z(0)$ such that

$$\max_{0 \leq k \leq I} \|f(k) - y(k)\| < \epsilon.$$ 

**Proof:** One may construct an $(n+1)$-dimensional discrete-time dynamic system with the form

$$\begin{align*}
    z(k+1) &= f(1 + z_{n+1}(k)) \\
    z_{n+1}(k+1) &= 1 + z_{n+1}(k)
\end{align*}$$

where $z \in \mathbb{R}^n$ and an initial condition is

$$\begin{align*}
    z(0) &= f(0) \\
    z_{n+1}(0) &= 0.
\end{align*}$$

It is easy to verify that

$$z(k) = f(k), \quad \text{for } 0 \leq k \leq I < +\infty.$$ 

Next, one may apply the same method used in the proof for Corollary 1 to obtain the corollary. \[\square\]

Corollaries 1 and 2 indicate that a DRNN with the form of (23) is capable of uniformly approximating a discrete-time trajectory which is produced from either a discrete-time system or a continuous function. This fact is useful for practical applications such as robot trajectory generation or data fit problem.

V. CONCLUSION

It has been proved in this note that a discrete-time DRNN may be used to uniformly approximate a discrete-time state-space trajectory which is produced by either a dynamic system or a continuous-time function to any degree of precision. The analytical results show that some of hidden units of such a DRNN may be selected as output units of the network and the states of these output units may be used to uniformly approximate a desired state-space trajectory. The proof used in this note is constructive and may be extended to study the approximation issue for other types of DRNN’s. Also, it has been indicated that this approximation process has to be carried out by an adaptive learning process. A successful approximation procedure contains the following three important steps: i) determining the appropriate numbers of the hidden units, ii) choosing the adequate weight learning algorithms, and iii) using the learning signals which contain sufficient information.

REFERENCES


On Infinity Norms as Lyapunov Functions for Linear Systems

Andrzej Polański

**Abstract**—The note presents a construction of the Lyapunov function defined by $\|z\|_\infty$, vector norm for both continuous- and discrete-time linear systems. In contrast to the previous results, no restrictions on the positions or multiplicity of system matrix eigenvalues are required.

**I. INTRODUCTION**

The most frequently used Lyapunov functions are quadratic ones. With the use of quadratic forms, one can state necessary and sufficient Manuscript received June 9, 1994; revised January 16, 1995. This work was supported in part by Polish KBN Grant 3 P030 010 06. The author is with the Department of Automatic Control, Silesian Technical University at Akademia Rzeczypospolitej Polskich, Poland.

IEEE Log Number 9411663.