LMI-Based Neural Observer for State and Nonlinear Function Estimation

Woongsun Jeon, Ankush Chakrabarty, Ali Zemouche, and Rajesh Rajamani\*

Woongsun Jeon

School of Electrical and Electronics Engineering, Chung-Ang University, Seoul 06974, South Korea

Ankush Chakrabarty

Mitsubishi Electric Research Laboratories, Cambridge, MA 02139, USA

Ali Zemouche

University of Lorraine, IUT Henri Poincare de Longwy, CRAN CNRS UMR 7039, 54400 Cosnes et Romain, France

Rajesh Rajamani

University of Minnesota, Twin Cities, Minneapolis, MN 55455 USA

E-mail: rajamani@umn.edu

Abstract

This paper develops a neuro-adaptive observer for state and nonlinear function estimation in systems with partially modeled process dynamics. The developed adaptive observer is shown to provide exponentially stable estimation errors in which both states and neural parameters converge to their true values. When the neural approximator has an approximation error with respect to the true nonlinear function, the observer can be used to provide an bound on the estimation error. The paper does not require assumptions on the process dynamics or output equation being linear functions of neural network weights and instead assumes a reasonable affine parameter dependence in the process dynamics. A convex problem is formulated and an equivalent polytopic observer design method is developed. Finally, a hybrid estimation system that switches between a neuro-adaptive observer for system identification and a regular nonlinear observer for state estimation is proposed. The switched operation enables parameter estimation updates whenever adequate measurements are available. The performance of the developed adaptive observer is shown through simulations for a Van der Pol oscillator and a single link robot. In the application, no manual tuning of adaptation gains is needed and estimates of both the states and the nonlinear functions converge successfully.

Keywords: Observers, learning for control, nonlinear systems, neural networks, linear matrix inequalities, function approximation

1. Introduction

The state estimation problem for nonlinear systems has attracted much attention during the last ten years. Major nonlinear observer design methods that have been developed include the high gain observer method for systems in triangular form (or transformable to triangular form),1-3 LMI-based methods for monotonic nonlinear systems,4 LMI/LPV based methods for Lipschitz and bounded Jacobian systems,5 LMI-based methods for unknown input nonlinear systems,6 LMI-based methods for nonlinearities satisfying an incremental quadratic constraint.7,8 All of the developed nonlinear observer design methods rely on having a known analytical model of the nonlinear system. However, many applications involve experimentally measured data with no analytical model available for accurately capturing some of the measured behavior. For example, see problems on magnetic position estimation9,10 and state-of-charge estimation in batteries.11 For many such practical systems, it would be beneficial to develop a method that can find a model of the nonlinear functions in the dynamics, while also estimating the states of the system.

Several authors have previously used neural network-based approximations of nonlinear systems for observer design. Some of these approximations utilized output-error backpropagation for updating the weights of the neural network, while also providing stability certificates for these algorithms.12-16 In this previous literature, even though estimation errors are uniformly bounded, manual tuning of multiple sets of adaptive gains is required in order to keep estimation errors small. For instance, the weight adaptation/learning rules require manual tuning of the learning rates and the damping factors parameters and the gains for system state estimation.14 A neural state estimator for discrete-time nonlinear systems with noisy measurement channels was proposed.17 The method ensured the convergence properties of the neural estimator using a minmax tuning technique, but the learning process had to be offline. An estimator based on neural networks and nonlinear programming was proposed.18 The method computed observer gains using a Lyapunov-based technique and its neural weights were searched for via nonlinear programming. A recent conference publication used a Lyapunov-based technique and an adaptive algorithm.19 Instead of adaptation gains obtained from backpropagation ideas, the authors proposed LMI conditions to compute both observer gains and adaptation gains. However, it required an assumption on the process dynamics being a linear function of parameters.

System identification using deep neural networks has been recently studied by a number of authors.20,21 In these studies, deep state space models as an extension to classic state space models are presented for system identification.21 The deep state space model can capture a wide range of dynamics and output uncertainty due to the flexibility of deep neural networks. However, the learning has to be conducted with a dense training data set and the approach does not estimate the system states and nonlinear function simultaneously.

Having many measurement channels can provide good performance on adaptation of the neural network-based approximations. In practice, however, some of the measurements may not always be available. For example, a GPS signal is not available when a car is driving in tunnels or in urban canyons. Also, a Lidar measurement to a target cannot be obtained if occlusion occurs or if there is a large incident angle between the surface of the target and the transmitted laser beam. Some measurements may only be available on an instrumented test track and only a subset of the measurements available during regular operation. Therefore, it is valuable to design a hybrid observer that switches between a neural network-based observer and a regular nonlinear observer to deal with changes of measurement channel availability.

In this paper, we propose a neuro-adaptive observer for nonlinear systems with partially modeled process dynamics. The linear portion of the dynamics is assumed to be known. The proposed observer estimates both states and the unknown nonlinear functions simultaneously using shallow neural network approximators. Uncertainty in the modeled linear portion is compensated through adaptation in the estimated nonlinear function. We formulate convex problems and observer-design inequalities involving LMIs to systematically compute observer gains for both state estimation and neural network weight adaptation. The state and model estimation error dynamics of the proposed observer are guaranteed to be exponentially stable. We also present a neuro-adaptive observer design that guarantees performance in the presence of approximation errors between the neural approximator and the true nonlinear functions. As a result, the proposed observer does not require manual tuning and works for any feasible solution of the observer design LMIs. Furthermore, we propose a hybrid observer design technique that switches between the neuro-adaptive observer and a regular nonlinear state observer. The hybrid observer utilizes the neuro-adaptive observer to estimate states and to conduct system identification. But, if only a subset of the measurement channels is available, it switches to the nonlinear observer to estimate states of the system using the limited measurement channels based on the previous system identification results from the neuro-adaptive observer.

The contributions of this paper are as follows:

1. The developed neuro-adaptive observer is shown to provide exponentially stable estimation errors in which both states and neural parameters converge to their true values.
2. When the neural approximator has an approximation error with respect to the true nonlinear function, the observer provides an bound on the estimation error as a function of the approximation error.
3. The paper does not require assumptions on the process dynamics or output equations being linear functions of parameters.
4. The paper assumes an affine parameter dependence in the estimation error dynamics and develops an equivalent polytopic model formulation for observer design.
5. The paper develops a hybrid observer design that switches between the neuro-adaptive observer and a nonlinear state observer based on the same neural approximation structure.

This paper is organized as follows. In Section 2, the system description and neural approximator are provided. In Section 3, the proposed observer structure and LMIs for the observer gains are presented. In Section 4, LMIs for the observer gains in the presence of approximation errors and measurement noise are developed. In Section 5, a hybrid observer design method with an additional nonlinear state observer based on the neural approximator structure is proposed. Then, in Section 6, the performance of the proposed observer is demonstrated though numerical examples. Conclusions are presented in Section 7.

2. Problem Statement

2.1. System Description

We consider a class of nonlinear systems in which the process dynamics have unknown nonlinear functions. The class of systems is described as

|  |  |
| --- | --- |
| , | (1) |

where is the state vector, is the input vector, and is the output vector. The nonlinearity is the unknown nonlinear function in the system dynamics. , , and are constant matrices assumed to be known. The following assumption is stated:

**Assumption 1.** The state is uniformly bounded, that is .

Since Assumption 1 implies that is restricted to a compact subset of , the neural network approximability results12,22 can be utilized to represent the unknown nonlinearity.

2.2. Neural Approximator

The nonlinearity in the system (1) can be represented as

|  |  |
| --- | --- |
|  | (2) |

for where is the weight in the output layer of a neural network which is unknown, is the number of neurons utilized, is an activation function, is a constant matrix, and is the approximation error, which is bounded.22 The activation functions are chosen by a designer, and thus are completely known. The following assumptions on the neural approximator are stated:

**Assumption 2.** The weights are bounded as

|  |  |
| --- | --- |
|  | (3) |

for all and .

**Assumption 3.** The activation functions are uniformly bounded as

|  |  |
| --- | --- |
|  | (4) |

and are differentiable Lipschitz continuous functions with bounded Jacobians:

|  |  |
| --- | --- |
|  | (5) |

for every where .

By using the neural approximator (2), the system (1) can be represented as

|  |  |
| --- | --- |
|  | (6) |

for .

If the nonlinear function can be found, it would also correct for errors in the assumed linear portion ( and matrices). Thus, errors in the assumed linear portion can be automatically corrected by the estimated nonlinear portion of the model.

3. Neuro-Adaptive Observer

A data-driven adaptive observer is proposed to estimate the states and unknown nonlinearity of the system (1) as follows

|  |  |
| --- | --- |
|  | (7) |

for and where and are observer gain matrices.

In the derivation of estimation error dynamics, we will use following notations for providing more compact writing: is a identity matrix, without subscript is an identity matrix of appropriate dimensions, is a column vector of elements all set to one, 0 is a zero matrix of appropriate dimensions, and is the Kronecker product.

Let the estimation error be . Based on (6) and (7), the state estimation error dynamics are given by

|  |  |
| --- | --- |
|  | (8) |

for . By subtracting and adding a term

|  |  |
| --- | --- |
|  | (9) |

from and to the nonlinear term, the nonlinear term can be written as

|  |  |
| --- | --- |
|  | (10) |

where and .

Therefore, the state estimation error dynamics (8) can be represented in the following compact form:

|  |  |
| --- | --- |
| , | (11) |

where is the parameter estimation error, with

|  |  |
| --- | --- |
| *,*  *,*  *,*  *,*  *,*  . | (12) |

Using (12), the parameter estimation error dynamics can be represented as

|  |  |
| --- | --- |
|  | (13) |

where .

We introduce an augmented state vector as follows

|  |  |
| --- | --- |
| . | (14) |

Then, the estimation error dynamics become

|  |  |
| --- | --- |
|  | (15) |

where

|  |  |
| --- | --- |
| . | (16) |

It is noted that and are observer gains to be computed, and is time-varying matrix since consists of the activation functions.

From Assumption 3, the value of is in an interval:

|  |  |
| --- | --- |
| . | (17) |

The set of vertices of the polytopic function can be defined as

|  |  |
| --- | --- |
| . | (18) |

Hence, the time-varying matrix can be considered to be an affine parameter dependent model23 of :

|  |  |
| --- | --- |
| for all , | (19) |

where is the matrix with the constant entries of while setting the time-varying entries to zero, and is the matrix that entries are all zero except the one corresponding to is equal to unity. Then, the affine model (19) with can equivalently be viewed as a polytopic model where . Using this, we will construct a finite number of LMIs to compute the observer gains and .

**Remark 1.** The activation function needs to be designed over an interval excluding zero, i.e., for the observer gain computation since having zero in the interval implies local unobservability of the pair at . Such a condition is necessary to ensure the feasibility of the LMI conditions for the proposed neuro-adaptive observer. Under this condition, nonlinear functions can still be represented by , where , by carefully selecting the activation functions. For example, even if the system has a known linear portion without unknwon nonlinearities, i.e., , we can still represent it as

|  |  |
| --- | --- |
|  | (20) |

Further, if the nonlinearity is linear in known functions which cross zero, the exact linear map cannot be represented by the neural approximator due to the condition . However, the nonlinearity can still be respresented by , where , .

In a compact space, if the true system is at least piecewise continuous, then ReLU and (continuous) sigmoid functions are activations that have universal approximation properties almost everywhere (except on a set of measure zero). The number of these activations is typically a hyperparameter that one has to tune by trial and error.20,21 In some applications in which the function estimation does not have to be done in real-time and prediction/estimation errors are accessible, model selection criteria such as Bayesian information criterion24 and Akaike information criterion25 can be utilized to find the optimal model structure. Since the identification is being done real-time, we need to start with an adequate number of activation functions, usually informed by prior knowledge of expected properties of the nonlinearity e.g., in vehicles and buildings we expect smooth functions; as long as these are scaled down so that no individual component (at least initially) overwhelms the other activation functions, our proposed method will work in practical settings. As a performance measure, the convergence of neural weights can be used. When different model structures need to be compared, convergence speed and variance of the neural weights can be used as an indicator for the model selection. If absolutely nothing is known about the nonlinear function, then we typically choose the number of activation functions based on practical considerations e.g., what the speed of feedforward prediction needs to be on-chip, what memory is available to us for storing weights, and how big of an SDP can be solved on our machine.

First, we aim to determine the observer gains such that the augmented estimation error converges exponentially toward zero under the assumption that the approximation error is negligible, . i.e., the unknown nonlinearity can be exactly modeled using a neural approximator.

**Theorem 1.** Consider the nonlinear system (6) and observer (7). If there exist matrices , and of appropriate dimensions, and fixed scalars and such that

|  |  |
| --- | --- |
| , | (21) |

where

|  |  |
| --- | --- |
|  | (22) |

then the estimation error dynamics of the observer (7) is globally exponentially stable with observer gain

|  |  |
| --- | --- |
| . | (23) |

**Proof.** Consider the Lyapunov function candidate . If , the derivative of along the trajectories of the estimation error dynamics (15) is given by

|  |  |
| --- | --- |
| . | (24) |

Global exponential stability of the estimation error dynamics can be ensured by having the Lyapunov function candidate to satisfy the condition .26

|  |  |
| --- | --- |
| . | (25) |

Then, (25) can be represented in matrix form as

|  |  |
| --- | --- |
|  | (26) |

where

|  |  |
| --- | --- |
| . | (27) |

Using the differential mean value theorem, the activation function difference can be represented as

|  |  |
| --- | --- |
|  | (28) |

where . The augmented activation functions can be written as

|  |  |
| --- | --- |
| . | (29) |

Based on Assumption 3, and satisfy the following sector condition:

|  |  |
| --- | --- |
|  | (30) |

where

|  |  |
| --- | --- |
|  | (31) |

and the lower and upper bounds of the Jacobian (as denoted in Assumption 3) are

|  |  |
| --- | --- |
| and  . | (32) |

The sector condition (30) can be rewritten in matrix form as

|  |  |
| --- | --- |
| . | (33) |

A symmetric form of (33) is

|  |  |
| --- | --- |
| . | (34) |

By considering the augmented estimation error vector, we represent (34) as

|  |  |
| --- | --- |
|  | (35) |

where

|  |  |
| --- | --- |
|  | (36) |

and

|  |  |
| --- | --- |
| . | (37) |

Using the S-procedure lemma,27 the inequality condition for all augmented state vector such that can be accomplished if there exists a positive scalar such that

|  |  |
| --- | --- |
| . | (38) |

Then, (38) is satisfied when the following condition is satisfied

|  |  |
| --- | --- |
|  | (39) |

where

|  |  |
| --- | --- |
| . | (40) |

By using the Schur complement, the inequality (39) is equivalent to

|  |  |
| --- | --- |
| , | (41) |

or

|  |  |
| --- | --- |
| . | (42) |

By using the Young’s relation,5 we have the following inequality

|  |  |
| --- | --- |
| . | (43) |

From Assumption 2, we know that

|  |  |
| --- | --- |
|  | (44) |

with

|  |  |
| --- | --- |
| . | (45) |

Using (43)-(45), (42) becomes

|  |  |
| --- | --- |
| . | (46) |

By taking the Schur complement and introducing a new variable , we rewrite (46) as

|  |  |
| --- | --- |
|  | (47) |

where

|  |  |
| --- | --- |
| . | (48) |

Instead of solving the LMI for all the set of , is considered to be the affine parameter dependent model as (19). Then, from Assumption 3, the affine model can be converted to a polytopic model with the set of vertices defined in (18). Therefore, solving the LMI in (47) is equivalent to finding a feasible solution of the LMI for the set of the vertices defined in (18). ■

**Remark 2.** The standard adaptive observer-based techniques need some additional constraints/conditions (a kind of passivity condition), in addition to the persistence excitation condition, to be satisfied in order to ensure the global convergence of the system containing the estimation parameter and the state observer.28 Such conditions are often expressed in terms of equality constraints which turn to be very conservative in general, except for some particular classes of systems, like the Brunovsky canonical form with additional Hurwitz assumption on . On the other hand, with the proposed LMI-based design method, these additional constraints are not necessary. Instead, only the detectability condition (Remark 1) which is less conservative is required for the proposed adaptive observer.

4. Neuro-Adaptive Observer

Next, we consider the situation that the function approximation error is not negligible. Therefore, we aim to determine the observer gains such that the estimation error dynamics with performance output is stable with gain :

|  |  |
| --- | --- |
|  | (49) |

where is the disturbance attenuation level. This can be done by the following theorem.

**Theorem 2.** Consider the nonlinear system (6) and observer (7). For a positive scalar , if there exist matrices , of appropriate dimensions, and fixed scalars , and such that

|  |  |
| --- | --- |
| , | (50) |

with

|  |  |
| --- | --- |
|  | (51) |

then the estimation error dynamics of the observer (7) with observer gain

|  |  |
| --- | --- |
|  | (52) |

demonstrates performance (49) with the attenuation level .

**Proof.** Consider the Lyapunov function candidate . The derivative of along the trajectories of the estimation error dynamics (15) is

|  |  |
| --- | --- |
| . | (53) |

The performance criterion (49) is satisfied if the following inequality holds.23

|  |  |
| --- | --- |
| . | (54) |

Using the derivative of Lyapunov function, the inequality (54) becomes

|  |  |
| --- | --- |
| . | (55) |

The condition ensures the exponential convergence of the estimation error.26 Then, the inequality can be represented in matrix form as

|  |  |
| --- | --- |
|  | (56) |

with

|  |  |
| --- | --- |
| . | (57) |

Using the differential mean value theorem, the activation function difference can be represented as (28). Like the proof of Theorem 1, the sector condition in matrix form can be rewritten as (35) with (36) and (37). By using S-procedure lemma and Schur complements, and invoking Young’s relation akin to arguments made in the proof of Theorem 1, the LMI (50) can be constructed. ■

The LMIs for the case where measurement noise is explicitly considered can also be constructed along the same lines. Let the output equations be

|  |  |
| --- | --- |
|  | (58) |

where is the measurement noise term. We introduce an augmented error state vector as . As a result, the estimation error dynamics become

|  |  |
| --- | --- |
| , | (59) |

where

|  |  |
| --- | --- |
| and . | (60) |

**Corollary 1.** Consider the case where the process dynamics are described as in (6) and the output equations are given by (58). For a positive scalar , suppose that there exist matrices , of appropriate dimensions, and fixed scalars , and such that the LMI

|  |  |
| --- | --- |
| , | (61) |

is satisfied. Then the observer (7) with observer gain

|  |  |
| --- | --- |
|  | (62) |

results in error dynamics which is stable with the attenuation level for the performance output with respect to the approximation error and sensor noise .

**Proof.** The proof is along the same lines as the proof of Theorem 2. By replacing and with and in the proof of Theorem 2, LMI (61) to compute observer gains that guarantee performance in the presence of measurement noise can be constructed. ■

5. Hybrid Observer Design

In this section, we propose a hybrid observer design that switches between the neuro-adaptive observer and a nonlinear observer based on the neural approximation structure. The hybrid observer utilizes the neuro-adaptive observer to estimate states and to conduct system identification. If only a subset of the measurement channels is available, it switches to the nonlinear observer to estimate states of the system using the limited measurement channels based on the system identification results from the neuro-adaptive observer.

A nonlinear observer is proposed for the system with the neural approximator as

|  |  |
| --- | --- |
| . | (63) |

It is noted that is a known constant since the previous estimated value from the neuro-adaptive observer is used for the observer (63).

Assume that the approximation error is negligible, . Then, the estimation error dynamics from the nonlinear system (6) and observer (63) is

|  |  |
| --- | --- |
| , | (64) |

where

|  |  |
| --- | --- |
| . | (65) |

It is noted that is bounded since is bounded (Assumption 3) and is the constant vector. It is obvious that is proportional to and becomes zero if is zero.

We aim to construct an LMI to compute a gain matrix for the observer (63) which guarantees stability such that

|  |  |
| --- | --- |
|  | (66) |

where is the disturbance attenuation level.

**Theorem 3.** Consider the nonlinear system (6) and observer (63). For a positive scalar , if there exist matrices , of appropriate dimensions, and fixed scalars , and such that

|  |  |
| --- | --- |
|  | (67) |

where

|  |  |
| --- | --- |
|  | (68) |

then the estimation error dynamics of the observer (63) is exponentially stable with observer gain

|  |  |
| --- | --- |
|  | (69) |

if and an performance gain with respect to is satisfied with the attenuation level .

**Proof.** Consider the Lyapunov function candidate . The derivative of along the trajectories of the estimation error dynamics (64) is

|  |  |
| --- | --- |
| . | (70) |

The performance criterion (66) is satisfied if the following inequality holds.23

|  |  |
| --- | --- |
| . | (71) |

Using the derivative of the Lyapunov function candidate and adding the term for the exponential convergence of the estimation error,26 the inequality (71) can be represented in matrix form as

|  |  |
| --- | --- |
|  | (72) |

with

|  |  |
| --- | --- |
| . | (73) |

Using the differential mean value theorem, the activation function difference can be represented as (28). Like the proof of Theorem 1, the sector condition in matrix form can be rewritten as (34):

|  |  |
| --- | --- |
|  | (74) |

where

|  |  |
| --- | --- |
| , and  . | (75) |

or

|  |  |
| --- | --- |
|  | (76) |

Using the S-procedure lemma,27 the inequality condition for all state vector such that can be accomplished if there exists a positive scalar such that

|  |  |
| --- | --- |
| . | (77) |

Absorbing into the matrix to define a new positive definite matrix and introducing a new variable . Then, (77) is satisfied when the following condition is satisfied

|  |  |
| --- | --- |
| , | (78) |

where

|  |  |
| --- | --- |
| . | (79) |

Thus concludes the proof. ■

**Theorem 4.** Let Observer 1 be the neuro-adaptive observer with the gain obtained from Theorem 1 and let Observer 2 be the nonlinear observer with the gain obtained from Theorem 3. Assume that the approximation error is negligible, i.e., . The estimation error in Observer 2 can be bounded to be smaller than any desired value by choosing the dwell time in Observer 1 to be sufficiently large.

**Proof.** Consider a switching from the neuro-adaptive observer (Observer 1) to the nonlinear observer (Observer 2). Let Observer 1 start at and let the switching occur at time . At the time of switching, the value of the Lyapunov function of Observer 1 is

|  |  |
| --- | --- |
| . | (80) |

Since the neuro-adaptive observer guarantees exponential stability,

|  |  |
| --- | --- |
| . | (81) |

Integrate both sides to obtain a relationship between and :

|  |  |
| --- | --- |
| . | (82) |

Then, (82) becomes

|  |  |
| --- | --- |
| , | (83) |

which implies

|  |  |
| --- | --- |
| , | (84) |

and hence,

|  |  |
| --- | --- |
| . | (85) |

Since , (85) becomes

|  |  |
| --- | --- |
| . | (86) |

Then,

|  |  |
| --- | --- |
| . | (87) |

Thus, will become sufficiently small or zero if the neuro-adaptive observer (Observer 1) spends sufficient dwell time . Since Observer 2 guarantees stability (66), is bounded after the switching as

|  |  |
| --- | --- |
| . | (88) |

From (65), (88) can be

|  |  |
| --- | --- |
| . | (89) |

where . Since the activation functions are bounded (Assumption 3) and is a diagonal matrix, is bounded as

|  |  |
| --- | --- |
| , | (90) |

where is the maximum singular value of and is the maximum bound of for all . Hence, if Observer 1 has sufficiently large dwell time, the switching to Observer 2 from Observer 1 enables any sufficiently small bound on . Furthermore, if is zero at the switching time, will converge to zero in Observer 2. ■

**Remark 3.** Although our contributions focus on the single-hidden-layer neural approximator, the method proposed herein could potentially be extended to multi-layer (deeper) neural networks. Some challenges associated with such a formulation include: conservatism induced in representing nested activation function with sector bounds, requirement of large amount of data to learn a (typically) large number of neural weights, and (to the best of our knowledge) the absence of a universal approximation theorem that directs activation function selection or ensures that the induced functions are dense in the space of a family of important functions e.g., continuous functions, enabling robustness analysis during learning.

6. Simulation Examples

Simulation studies are conducted for the Van der Pol oscillation system29 and a single-link robot system12 to evaluate the proposed neuro-adaptive observer. The observer gain is obtained from solving the convex programs in Theorem 1 and 3. We solve the convex programs using YALMIP with SDPT solver.30,31 Also, unmodeled dynamics of the system are simultaneously estimated by using estimated weights and states via the neural approximator:

|  |  |
| --- | --- |
| . | (91) |

6.1. Van der Pol Oscillation System

Consider a Van der Pol oscillation system29 given by

|  |  |
| --- | --- |
|  | (92) |

The nonlinearity is assumed to be unknown in this simulation study.

The neuro-adaptive observer is utilized to estimate the states and unknown nonlinearity. A hyperbolic tangent function and two soft clipping functions32 are used for the observer as activation functions:

|  |  |
| --- | --- |
|  | (93) |

We solve the conditions of Theorem 1 with , , and , for the observer gain:

|  |  |
| --- | --- |
|  | (94) |

The initial condition of the system is . Since is measurable, the initial condition of the observer is set as . The initial conditions of the weight   is randomly sampled from a normal distribution with zero mean and 1 variance, and and are randomly sampled from a normal distribution with zero mean and 402 variance.

**Figure 1** shows the simulation results of the state and nonlinear function estimation. The proposed neuro-adaptive observer estimates simultaneously both the state and unknown nonlinearity well. Also, the reconstructed nonlinear function based on the estimated weight in **Figure 2** is shown to be close to true nonlinear function, as shown in **Figure 1**.



**Figure 1.** Van der Pol simulation results of states and nonlinear function estimation using the proposed neuro-adaptive observer.



**Figure 2.** Van der Pol simulation results of weight estimation using the proposed neuro-adaptive observer.

6.2. Single Link Robot

Consider a single-link robot arm system12 given by

|  |  |
| --- | --- |
|  | (95) |

where is the angle, is the input torque, is the moment of inertia of the link, is the mass of the link, is the length of the link, and is the acceleration due to gravity. The parameters of the robot system are defined as , , , and . The state vector is defined as , then the state-space equation of the system is given by

|  |  |
| --- | --- |
|  | (96) |

The nonlinearity is assumed to be unknown in this simulation study.

First, simulations are conducted under the conditions that the initial condition of the system is , and the input is . The system is operated in the range and (although we do not know this).

A sine function is utilized as an activation function for the neuro-adaptive observer and is assumed to be . We solve the conditions of Theorem 1 with , , and for the observer gains:

|  |  |
| --- | --- |
|  | (97) |

The initial condition of the observer is set as . The initial condition of the weight is randomly sampled from a normal distribution with zero mean and 1 variance.

**Figure 3** shows the simulation results. It is seen that the proposed neuro-adaptive observer shows good performance in both state and unknown nonlinearity estimation. Also, the weight parameter converges to its true value as shown in **Figure 4**.

****

**Figure 3.** Single link robot simulation results of states and nonlinear function estimation using the proposed neuro-adaptive observer.



**Figure 4.** Single link robot simulation results of weight estimation using the proposed neuro-adaptive observer.

Next, we conduct simulation studies of the case that the neuro-adaptive observer is initially used for both system identification and state estimation, and then the nonlinear observer is used under a limited measurement condition after the system identification: all the states are initially measurable of the robot system (96), and after 25 seconds, only is measurable.

The neuro-adaptive observer is first utilized to estimates the states and unknown nonlinearity. A hyperbolic tangent function and soft clipping functions are considered to be three activation functions for the observer:

|  |  |
| --- | --- |
|  | (98) |

We solve the conditions of Theorem 1 with , and to obtain the observer gains:

|  |  |
| --- | --- |
|  | (99) |

After 25 seconds, the states of the system are estimated by using the nonlinear observer (63) while only measuring . in (63) is obtained by computing the mean value of (estimated from the neuro-adaptive observer) after transient period. We solve the conditions of Theorem 3 with , and to obtain the observer gains with :

|  |  |
| --- | --- |
| . | (100) |

The initial condition of the system is , and the input is . The initial condition of the observer is set as . The initial conditions of the weight are randomly sampled from a normal distribution with zero mean and 1 variance.

**Figure 5** shows the simulation results. It is seen that the proposed neuro-adaptive observer shows good performance in both state and unknown nonlinearity estimation. **Figure 6** shows weight estimation. It is seen that weight estimates have time-varying behavior to compensate for the approximation error if the selected activation functions are limited in being able to approximate the nonlinear function.

From 25 seconds, the nonlinear observer using only measurement is utilized to estimate the states. As shown in **Figure 5**, the observer estimates the states with very small error. Also, the learning performance can be seen from the reconstructed nonlinear function and is shown to be good as seen in **Figure 5**.



**Figure 5.** Single link robot simulation results of states and nonlinear function estimation using the proposed neuro-adaptive observer and nonlinear observer.



**Figure 6.** Single link robot simulation results of weight estimation using the proposed neuro-adaptive observer with 3 activation functions.

We consider the situation with noisy measurements and validate the performance of the neuro-adaptive observer for handling sensor noise. The output equation with noise is where the measurement noise is and . First, a simulation study is conducted without using the neuro-adaptive observer. **Figure 7(a)** shows simulation results using the observer with the gains from (99) and (100). Due to the significant measurement noise, the nonlinear function estimation performance is poor. Then, the neuro-adaptive observer is utilized. The gains of the neuro-adaptive observer are computed by solving the LMIs in Corollary 1 with , , , and :

|  |  |
| --- | --- |
|  | (101) |

The simulation results using neuro-adaptive observer are illustrated in **Figure 7(b)**. We see that the neuro-adaptive observer provides significantly better estimation performance in spite of the measurement noise.



(a)



(b)

**Figure 7.** Single link robot simulation results of nonlinear function estimation with noisy measurement. (a) Neuro-adaptive observer. (b) neuro-adaptive observer.

7. Conclusions

This paper developed a data-driven methodology for estimating the states of a nonlinear dynamical system with unmodeled dynamics. A neural approximator utilizes the operational data of the system and learns the unmodeled dynamics while a fixed-gain observer updates the state estimate based on measured outputs and approximator predictions. We provide rigorous theoretical guarantees on the estimation quality despite the presence of uncertain dynamics. This paper does not require assumptions on the process dynamics or output equations being linear functions of parameters and instead assumes a reasonable affine parameter dependence in the estimation error dynamics. Furthermore, we consider the practical situation when data relevant to model adaptation is not available and only a subset of the data is available. In this setting, we proposed a hybrid estimation system that switches between neuro-adaptive observer for system identification and a regular nonlinear observer for state estimation. The switched operation enables parameter estimation updates whenever adequate measurements are available. The performance of the developed adaptive observer is presented through simulations for a Van der Pol oscillator and a single link robot. In the application, no manual tuning of adaptation gains is needed and the developed adaptive observer estimates both the states and the nonlinear functions successfully. The proposed type of estimation algorithm has recently been applied for state and tire model estimation on autonomous vehicles.33 In the future, we plan to apply the fundamental estimation algorithm developed herein to intelligent transportation system applications and to a position estimation for magnetic sensors problem.

Acknowledgements

This research was funded in part by the National Science Foundation, under grant PFI 1631133. This research was also supported in part by by the National Research Foundation of Korea(NRF) funded by Ministry of Science and ICT(RS-2022-00166150). The authors would like to thank Mouhacine Benosman at MERL for useful discussions and helpful comments.

References

1. T. Ahmed-Ali & F. Lamnabhi-Lagarrigue (2012). High Gain Observer Design for Some Networked Control Systems, *IEEE Transactions on Automatic Control*, vol. 57, no. 4, pp. 995-1000.
2. H. Khalil & L. Praly (2014). High Gain Observers in Nonlinear Feedback Control. *International Journal of Robust and Nonlinear Control,* vol. 24, no. 6, pp. 993-1015.
3. A. Alessandri, & A. Rossi (2015). Increasing-gain observers for nonlinear systems: Stability and design, *Automatica*, vol. 57, pp. 180-188.
4. M. Arcak & P. Kokotovic (2001). Nonlinear Observers: A Circle Criterion Design and Robustness Analysis. *Automatica,* vol. 37, pp. 1923-1930.
5. A. Zemouche, R. Rajamani, G. Phanomchoeng, B. Boulkroune, H. Rafaralahy & M. Zasadzinski (2017). Circle Criterion based H-Infinity Observer design for Lipschitz and Monotonic Nonlinear Systems - Enhanced LMI Conditions and Constructive Discussions. *Automatica,* vol. 85, pp. 412-425.
6. Q. P. Ha & H. Trinh (2004). State and input simultaneous estimation for a class of nonlinear systems, *Automatica*, Volume 40, Issue 10, Pages 1779-1785.
7. A. Chakrabarty, M. J. Corless, G. T. Buzzard, S. H. Żak & A. E. Rundell (2017). "State and Unknown Input Observers for Nonlinear Systems With Bounded Exogenous Inputs," in *IEEE Transactions on Automatic Control*, vol. 62, no. 11, pp. 5497-5510.
8. A. Chakrabarty, & Martin Corless (2010). "Estimating unbounded unknown inputs in nonlinear systems," *Automatica*, vol. 104, pp. 57-66.
9. R. Madson & R. Rajamani (2018). Magnetic Position Estimation in Ferromagnetic Systems Involving Significant Hysteresis. *IEEE/ASME Transactions on Mechatronics*, Vol. 23, No. 4, pp. 1555-1563.
10. S. Taghvaeeyan & R. Rajamani (2014). Two-Dimensional Sensor System for Automotive Crash Prediction. *IEEE Transactions on Intelligent Transportation Systems*, Vol. 15, No. 1, pp. 178-190.
11. H. Fang, Y. Wang, Z. Sahinoglu, T. Wada, & S. Hara (2014). State of charge estimation for lithium-ion batteries: An adaptive approach, *Control Engineering Practice*, vol. 25, pp. 45-54.
12. Y. H. Kim, F. L. Lewis, & C. T. Abdallah (1997). A dynamic recurrent neural-network-based adaptive observer for a class of nonlinear systems. *Automatica*, vol. 33, no. 8, pp. 1539–1543.
13. A. S. Poznyak, & W. Yu (2000). Robust asymptotic neuro‐observer with time delay term. *International Journal of Robust and Nonlinear Control*, vol. 10, iss. 7, pp. 535-559.
14. F. Abdollahi, H. A. Talebi & R. V. Patel (2006). A stable neural network-based observer with application to flexible-joint manipulators. *IEEE Transactions on Neural Networks*, vol. 17, no. 1, pp. 118-129.
15. D. Liu, Y. Huang, D. Wang, & Q. Wei (2013). Neural-network-observer-based optimal control for unknown nonlinear systems using adaptive dynamic programming. *International Journal of Control,* vol. 86, no. 9, pp. 1554-1566.
16. R. Rahimilarki, Z. Gao, A. Zhang & R. Binns (2019). Robust Neural Network Fault Estimation Approach for Nonlinear Dynamic Systems With Applications to Wind Turbine Systems. *IEEE Transactions on Industrial Informatics*, vol. 15, no. 12, pp. 6302-6312.
17. A. Alessandri, M. Baglietto, T. Parisini & R. Zoppoli (1999). A neural state estimator with bounded errors for nonlinear systems. *IEEE Transactions on Automatic Control*, vol. 44, no. 11, pp. 2028-2042.
18. A. Alessandri, C. Cervellera & M. Sanguineti (2007). Design of Asymptotic Estimators: An Approach Based on Neural Networks and Nonlinear Programming. *IEEE Transactions on Neural Networks*, vol. 18, no. 1, pp. 86-96.
19. A. Chakrabarty, A. Zemouche, R. Rajamani & M. Benosman (2019). Robust Data-Driven Neuro-Adaptive Observers With Lipschitz Activation Functions. *2019 IEEE 58th Conference on Decision and Control (CDC)*, Nice, France, pp. 2862-2867.
20. L. Ljung, C. Andersson, K. Tiels, & T. B. Schön (2020). Deep learning and system identification. In *Proc IFAC Congress, Berlin*.
21. D. Gedon, N. Wahlström, T. B. Schön, & L. Ljung (2020). Deep State Space Models for Nonlinear System Identification. *arXiv preprint arXiv:2003.14162*.
22. K. Hornik (1991). Approximation capabilities of multilayer feedforward networks. *Neural Networks*, vol. 4, no. 2, pp. 251–257.
23. C. Scherer, & S. Weiland (2000). Linear matrix inequalities in control. *Lecture Notes, Dutch Institute for Systems and Control, Delft, The Netherlands* 3, no. 2.
24. G. Schwarz (1978). Estimating the dimension of a model. *The Annals of Statistics*, vol. 6, no. 2, pp. 461–464.
25. H. Akaike (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, vol. 19, pp. 716-723.
26. H. K. Khalil (2002). *Nonlinear Systems*, Prentice Hall, Upper Saddle River, NJ, USA, 3rd edition.
27. S. Boyd, L. Ghaoui, E. Feron & V. Balakrishnan (1994). Linear Matrix Inequalities in Systems and Control Theory. *Society for Industrial and Applied Mathematics, SIAM Studies in Applied Mathematics*.
28. G. Besancon (2000). Remarks on nonlinear adaptive observer design. Systems & control letters, vol. 41, iss. 4, pp. 271-280.
29. H. Ichihara (2007). Observer design for polynomial systems using convex optimization. 2007 46th IEEE Conference on Decision and Control, New Orleans, LA, pp. 5347-5352.
30. J. Lofberg (2004). YALMIP: a toolbox for modeling and optimization in MATLAB. *2004 IEEE International Conference on Robotics and Automation (IEEE Cat. No.04CH37508)*, New Orleans, LA, pp. 284-289.
31. K. C. Toh, M. J. Todd, & R. H. Tütüncü (1999). SDPT3-a MATLAB software package for semidefinite programming, version 1.3. *Optimization methods and software* 11, no. 1-4, pp. 545-581.
32. K. D. Matthew, & M. Perelstein (2018). Neural Network-Based Approach to Phase Space Integration. *arXiv preprint arXiv:1810.11509*.
33. W. Jeon, A. Chakrabarty, A. Zemouche & R. Rajamani (2021). Simultaneous State Estimation and Tire Model Learning for Autonomous Vehicle Applications, *IEEE/ASME Transactions on Mechatronics*, vol. 26, no. 4, pp. 1941-1950.