2003 American Control Conference
http://www.cds.caltech.edu/~murray/papers/2002m_dm03-acc.html

Observability and Local Observer Construction for Unknown Parameters in Linearly and Nonlinearly Parameterized Systems

D. Del Vecchio and R. M. Murray
Control and Dynamical Systems
California Institute of Technology
Pasadena, CA 91125

Abstract

Using geometric concepts from observability theory for nonlinear systems, we propose an approach for parameter estimation for linearly and nonlinearly parameterized systems. The proposed approach relies on extending a parameter estimation problem to a state estimation problem by introducing the parameters as auxiliary state variables. Applying tools from geometric nonlinear control theory we establish an observability check for parameters, and we construct a local observer with established speed of convergence in the observable sets of the extended system.

1 Introduction

The problem of parameter estimation for linearly parameterized dynamical systems is well established (see [4] for example). Most of the update laws designed for parameter estimation guarantee parameter convergence if the persistence of excitation condition (PE) is verified [3, 9, 8]. The check of the persistence of excitation condition involves the computation of an integral on a given interval of time for all times. Such an integral must be positive definite to allow parameter convergence of a gradient-like adaptation law [15, 16]. In such a case the parameters converge to their true values and in [13] a Lyapunov function for the parameter's error dynamics is found. Unfortunately there is no constructive procedure for determining which inputs are persistently exciting and it is well known [4] that, for estimating the parameters in linear systems (DARMA model), the PE degree of the input has to be at least equal to the number of the parameters to be estimated.

As described in [2], for example, there are in principle no differences between parameter estimation and state estimation. Also [1] relates the stability properties of differential equations arising in adaptive identification with observability properties of a related dynamical system (see also [16].) A parameter estimation problem can be extended to a state estimation problem by introducing the parameters as auxiliary state variables. This idea is not new to the literature on estimation of parameters in linear stochastic systems where the extended Kalman filter has been used for joint parameter and state estimation problems [10].

Here we adopt a nonlinear observability point of view, and construct a joint state-parameter observer with local Lyapunov function-based convergence proof. The trade-off is that the analysis is local and holds only in the observable sets of the extended system. In the following sections we develop the theoretical framework, relying on well known results in nonlinear observers [6, 17, 5, 12, 18]. We first give an observability check for parameters: the extended system can be observable, locally observable or unobservable. In the latter case there does not exist any parameter observer since in any portion of the state space the complete evolution of the system can be generated by different dynamical systems with different parameters as well. In the first and second case there exists a parameter observer and we construct one with convergence proof. In particular, in the case of locally observable systems, the convergence proof relies on the assumption that the system evolves in an observable set (see for example [12].) Finally, to show the applicability of the proposed approach, we provide some simulation examples.

2 Problem Statement

Consider the system

$$\begin{cases} \dot{\overline{x}} = \overline{f}(\overline{x}, \theta, u(t)) \\ y = \overline{h}(\overline{x}, \theta, u(t)) \end{cases} \tag{1}$$

with $\overline{x} \in \mathbb{R}^{\overline{n}}$, $\theta \in \mathbb{R}^p$ unknown constant parameter vector, $u = (u_1, ..., u_l) \in \mathbb{R}^l$ exogenous input, $y \in \mathbb{R}^m$ measured output, $\overline{f}$ and $\overline{h}$ smooth functions. We assume that θ is lying in a known compact set Ω_θ. The problem is the joint estimation of the state $\overline{x}$ and the unknown constant parameters θ, which may occur linearly or nonlinearly in the function $\overline{f}$. Considering the extended state $x := (\overline{x}, \theta) \in \mathbb{R}^n$ with $n = \overline{n} + p$, system (1) becomes

$$\begin{cases} \dot{x} = f(x, u(t)) \\ y = h(x, u(t)) \end{cases} \quad (2)$$

where $f(x, u(t)) := (\overline{f}(\overline{x}, \theta, u(t)), 0_p)$. We pose the question of what conditions on system (2) allow to reconstruct x and therefore θ from the observation of y. We then recall some results on nonlinear observability (see [5, 17, 6, 12] for example). Let $h(x, u) = (h_1(x, u), ..., h_m(x, u))$, $\mathbf{u} = (u_1, ..., u_1^{(n_1-1)}, ..., u_l, ..., u_l^{(n_l-1)})$ with $\sum_{i=1}^{l} n_i = n_u$, and

$$\varphi_i^0 = h_i,$$

$$\varphi_i^j = L_f \varphi_i^{j-1} = \frac{\partial \varphi_i^{j-1}}{\partial x} f + \sum_{k=0}^{j-1} \frac{\partial \varphi_i^{j-1}}{\partial u^{(k)}} u^{(k+1)} = y_i^{(j)}.$$

Define the map $\Phi(x, \mathbf{u}) : \mathbb{R}^n \times \mathbb{R}^{n_u} \to \mathbb{R}^n$ to be

$$\Phi(x, \mathbf{u}) = (h_1, \varphi_1^1, ..., \varphi_1^{k_1-1}, ..., h_m, \varphi_m^1, ..., \varphi_m^{k_m-1})$$

for some k_is such that $\sum_{i=1}^{m} k_i = n$.

Definition 1 System (2) is said to be *locally observable* if there exist a non-empty set $\chi \times \mathcal{U} \subset \mathbb{R}^n \times \mathbb{R}^{n_u}$, such that the map $\Phi(x, \mathbf{u})$, for some k_is, is invertible with respect to x and its inverse is smooth for all $(x, \mathbf{u}) \in \chi \times \mathcal{U}$, that is

$$\text{rank}(\frac{\partial \Phi(x, \mathbf{u})}{\partial x}) = n \quad (3)$$

for all $(x, \mathbf{u}) \in \chi \times \mathcal{U}$. Then we say that $\chi \times \mathcal{U}$ is an observable set. If $\chi \times \mathcal{U} = \mathbb{R}^n \times \mathbb{R}^{n_u}$, the system is *observable*, and if $\chi \times \mathcal{U} = \emptyset$, the system is not observable.

Notice that in practice the observable sets can be computed by computing the set where the determinant of $(\partial \Phi(x, \mathbf{u})/\partial x)$ is zero, i.e. $\{(x, \mathbf{u}) | det(\partial \Phi(x, \mathbf{u})/\partial x) = 0\}$, and then taking the complement to $\mathbb{R}^n \times \mathbb{R}^{n_u}$.

Let $z = (y_1, y_1^{(1)}, ..., y_1^{(k_1-1)}, ..., y_m, ..., y_m^{(k_m-1)})^T$, so that $z = \Phi(x, \mathbf{u})$. Assuming that system (2) is locally observable and that $(x, \mathbf{u}) \in \chi \times \mathcal{U}$ we can estimate z and then invert Φ with respect to x so to have an estimate for x. The dynamics for z can be written as

$$\begin{aligned} \dot{z} &= Az + \rho(z, \mathbf{u}) \\ y &= Cz \end{aligned} \quad (4)$$

where $A = \text{block-diag}(A_1, ..., A_m)$ with $A_i \in \mathbb{R}^{k_i \times k_i}$ in the form

$$A_i = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ \vdots & \cdots & \cdots & \cdots & 0 \\ 0 & \cdots & \cdots & \cdots & 1 \\ 0 & 0 & \cdots & \cdots & 0 \end{pmatrix},$$

$C = (c_1, ..., c_m)^T$, with $c_i = (0, ..., 0, 1, 0, ..., 0)^T \in \mathbb{R}^n$ with the 1 in position $k_1 + ... + k_{i-1} + 1$ for $i > 1$ and $c_1 = (1, 0, ..., 0)^T$, $\rho(z, \mathbf{u}) = (\rho_1(z, \mathbf{u})^T, ..., \rho_m(z, \mathbf{u})^T)^T$, with $\rho_i(z, \mathbf{u}) = (0, ..., 0, L_f \varphi_i^{(k_i-1)}|_{x=\Phi^{-1}(z,\mathbf{u})})^T \in \mathbb{R}^{k_i}$. Then we focus on the estimation problem for z, and show with the following theorem how to construct an observer. The basic idea of such a theorem can be found in many references on nonlinear observers, here we follow the framework found in [12] with some modifications.

Theorem 2 *Given system (2), assume that* $(x, \mathbf{u}) \in \Omega_x \times \Omega_u \subset \chi \times \mathcal{U}$*, where* $\chi \times \mathcal{U}$ *is an observable set. Perform the nonlinear change of coordinates* $z = \Phi(x, \mathbf{u})$*, and construct the observer*

$$\dot{\hat{z}} = \text{Proj}(A\hat{z} + \rho(\hat{z}, \mathbf{u}) + S^{-1} K_o (y - C\hat{z}), \hat{z}), \quad (5)$$

where $\text{Proj}(y, \hat{z})$ *is the Lipschitz continuous function defined as*

$$\text{Proj}(y, \hat{z}) = \begin{cases} y & \text{if } p(\hat{z}) \leq 0 \\ y & \text{if } p(\hat{z}) \geq 0 \text{ and } (\nabla p(\hat{z}), y) \leq 0 \\ y_p & \text{if } p(\hat{z}) > 0 \text{ and } (\nabla p(\hat{z}), y) > 0 \end{cases} \quad (6)$$

where

$$p(\hat{z}) = \frac{\|\hat{z} - z_0\|^2 - r_\Omega^2}{\alpha^2 + 2\alpha r_\Omega},$$

and

$$y_p = \left[I - \overline{S}^{-1} P^{-1} \overline{S}^{-1} \frac{p(\hat{z}) \nabla p(\hat{z}) \nabla p(\hat{z})^T}{\nabla p(\hat{z})^T \overline{S}^{-1} P^{-1} \overline{S}^{-1} \nabla p(\hat{z})} \right] y$$

with α *an arbitrarily small positive constant,* r_Ω *and* z_0 *the radius and the center of the region* $\Omega_z \in \mathbb{R}^n$ *respectively, with* $\Phi(\Omega_x \times \Omega_u) \subset \Omega_z \subset \Phi(\chi \times \mathcal{U})$*,* K_o *designed such that* $(A - K_o C)$ *is Hurwitz,* P *solution of the Lyapunov equation* $P(A - K_o C) + (A - K_o C)^T P = -Q$*,* $Q > 0$*,* $S = \textit{block-diag}(S_1, ..., S_m)$*,* $S_i = \textit{diag}(\nu, \nu^2, ..., \nu^{k_i})$*, with* $\nu \in \mathbb{R}$*, and* $\overline{S} = \textit{block-diag}(\overline{S}_1, ..., \overline{S}_m)$ *with* $\overline{S}_i = \textit{diag}(1/\nu^{k_i-1}, ..., 1)$ *for all* i*. Then observer (5) guarantees for* $\hat{z}(0) \in \Omega_z$ *that*

(i) $\|\hat{z}(t) - z_0\| \leq r_\Omega + \alpha$ *for all times;*

(ii) $\lim_{t \to \infty} \|\hat{z}(t) - z(t)\| = 0.$

Proof: The introduced projection operator is a variant of the standard projection algorithm presented in [14], and it has the same properties, namely

1) $\|\hat{z}(t) - z_0\| \leq r_\Omega + \alpha$ for all times;

2) $\text{Proj}(y, \hat{z})$ is Lipschitz continuous;

3) $\|\text{Proj}(y, \hat{z})\| \leq \|y\|$;

4) $\tilde{z}^T\overline{S}P\overline{S}\text{Proj}(y,\hat{z}) \geq \tilde{z}^T\overline{S}P\overline{S}y$;

where $\tilde{z} = z - \hat{z}$. The proof of (i) descends directly from property 1) of the projection operator. To prove (ii) consider the dynamics of the error $\tilde{z}$, which can be derived by equations (4) and (5):

$$\dot{\tilde{z}} = Az+\rho(z,\mathbf{u})-\text{Proj}(A\hat{z}+\rho(\hat{z},\mathbf{u})+S^{-1}K_o(y-C\hat{z}),\hat{z}) \quad (7)$$

We consider the Lyapunov function for system (7):

$$V = \tilde{z}^T\overline{S}P\overline{S}\tilde{z},$$

where $\overline{S} = \text{block-diag}(\overline{S}_1,...,\overline{S}_m)$ with $\overline{S}_i = \text{diag}(1/\nu^{k_i-1},...,1)$ for all i, and with K_o such that

$$P(A-K_oC)+(A-K_oC)^TP = -Q, \quad (8)$$

for some arbitrary $Q > 0$. Taking the derivative of V with respect to time we have

$$\begin{aligned}\dot{V} &= \tilde{z}\overline{S}P\overline{S}\big(Az+\rho(z,\mathbf{u}) \\ &\quad - \text{Proj}(A\hat{z}+\rho(\hat{z},\mathbf{u})+S^{-1}K_o(y-C\hat{z}),\hat{z})\big) \\ &\quad + \big(Az+\rho(z,\mathbf{u}) \\ &\quad - \text{Proj}(A\hat{z}+\rho(z,\hat{\mathbf{u}})+S^{-1}K_o(y-C\hat{z}),\hat{z})\big)^T\overline{S}P\overline{S}\tilde{z}\end{aligned}$$

and recalling property 4) we obtain

$$\begin{aligned}\dot{V} &\leq \tilde{z}^T\overline{S}P\overline{S}\big(Az+\rho(z,\mathbf{u})\big)- \\ &\quad \tilde{z}\overline{S}P\overline{S}\big(A\hat{z}+\rho(\hat{z},\mathbf{u})+S^{-1}K_o(y-C\hat{z})\big) \\ &\quad + \big(Az+\rho(z,\mathbf{u})\big)^T\overline{S}P\overline{S}\tilde{z} \\ &\quad - \big(A\hat{z}+\rho(\hat{z},\mathbf{u})+S^{-1}K_o(y-C\hat{z})\big)^T\overline{S}P\overline{S}\tilde{z},\end{aligned}$$

which leads to

$$\begin{aligned}\dot{V} &\leq \tilde{z}^T(\overline{S}P\overline{S}(A-S^{-1}K_oC)+(A-S^{-1}K_oC)^T\overline{S}P\overline{S})\tilde{z}+ \\ &\quad \tilde{z}^T\overline{S}P\overline{S}(\rho(z,\mathbf{u})-\rho(\hat{z},\mathbf{u}))+(\rho(z,\mathbf{u})-\rho(\hat{z},\mathbf{u}))^T\overline{S}P\overline{S}\tilde{z}.\end{aligned}$$

Let $\tilde{w} = \overline{S}\tilde{z}$, then given the structure of $\overline{S}$ and S, the latter expression becomes

$$\begin{aligned}\dot{V} &\leq \tilde{w}^T\frac{1}{\nu}[P(A-K_oC)+(A-K_oC)^TP]\tilde{w} \\ &\quad - 2\tilde{w}P(\rho(z,\mathbf{u})-\rho(\hat{z},\mathbf{u})).\end{aligned}$$

Since $\hat{z}$, z and $\mathbf{u}$ are bounded in compact sets Ω_z, $\Phi(\Omega_x \times \Omega_u)$, and Ω_u, and $\rho(\cdot,\cdot)$ is continuous, it has a Lipschitz constant uniformly in $\mathbf{u}$ that we call k_L. Then by using (8) also, we get for $\nu < 1$

$$\dot{V} \leq -\frac{1}{\nu}\tilde{w}^TQ\tilde{w}+2\|P\|k_L\|\tilde{w}\|^2,$$

which finally leads to

$$\dot{V} \leq -\lambda_0\|\tilde{w}\|^2, \quad (9)$$

with $\lambda_0 = \lambda_{min}(Q)\frac{1}{\nu} - 2\|P\|k_L > 0$, which depends on the choice of K_o and $\nu < 1$, and leads to the proof of (ii). By integrating equation (9) we can obtain

$$\|\tilde{z}(t)\| \leq a\|\tilde{z}(0)\|e^{-bt}$$

where $a = 1/\nu^k$, with $k = \max_i\{k_i - 1\}$, and $b = \lambda_0/2\lambda_M(P)$, with $\lambda_M(P)$ the maximum eigenvalue of P. Thus since $\lambda_0 = \lambda_{min}(Q)\frac{1}{\nu} - 2\|P\|k_L$, we can decrease ν so to have the desired speed of convergence. By increasing the speed of convergence also a increases so to cause a larger transient, but the projection algorithm will assure that $\hat{z}$ remains bounded in the desired set. ■

Since α is arbitrarily small, and $\Omega_z \subset \Phi(\chi \times \mathcal{U})$, then by virtue of (i) of Theorem 2 $\hat{z} \in \Phi(\chi \times \mathcal{U})$ for all time. Therefore $\hat{x} = \Phi^{-1}(\hat{z},\mathbf{u})$.

Remark 3 Theorem 2 relies on the assumption that $(x,\mathbf{u})$ is in the observable set $\chi \times \mathcal{U}$ during the entire evolution of the system. Since x is not available for measurement, the condition $(x,\mathbf{u}) \in \{(v,w) \mid \det(\partial\Phi(v,w)/\partial v) = 0\}$ in general cannot be checked directly or established *a priori*. However we can check whether $(\dot{x},\mathbf{u}) \in \{(v,w) \mid \det(\partial\Phi(v,w)/\partial v) = 0\}$. Therefore in practice we can relax the projection to project $\hat{z}$ into a bigger set Ω_z not necessarily contained in $\Phi(\chi \times \mathcal{U})$, and choose $\hat{z}(0) \in \Phi(\chi \times \mathcal{U})$. Then we can let $\hat{x}(0) = \Phi^{-1}(\hat{z}(0),\mathbf{u}(0))$ and monitor whether $(\hat{x},\mathbf{u})$ is approaching the set $\{(v,w) \mid \det(\partial\Phi(v,w)/\partial v) = 0\}$. When $(\hat{x},\mathbf{u})$ becomes too close to this set, then we turn the observer off, so that $\|\tilde{z}\| \leq a\|\tilde{z}(0)\|e^{-bT}$, where T is the time at which the observer was turned off. This way we loose asymptotic convergence, but we can still have a small enough estimation error if T is sufficiently big (Example 1 will show this case.)

Remark 4 In the case when the input u of system (1) is not exogenous and contains feedback terms, as occurs in a feedback loop, it is not possible to differentiate it directly to obtain the vector $\mathbf{u}$. We need to provide an estimate for all the needed input derivatives. It can be shown that it is possible to construct an observer for the input derivatives with linear dynamics [12], so that the closed loop system is still stable. However in this case it is possible to show that the estimates error will converge into an arbitrarily small neighborhood of the origin, but not to the origin exactly.

An alternative solution that still guarantees asymptotic convergence to zero of the observer error $\tilde{z}$ can be proposed as well. In fact it is possible to augment system (2) with n_u integrators on the input side, and then solve the control problem for the augmented system. For example in the case that system (2) is single input, single

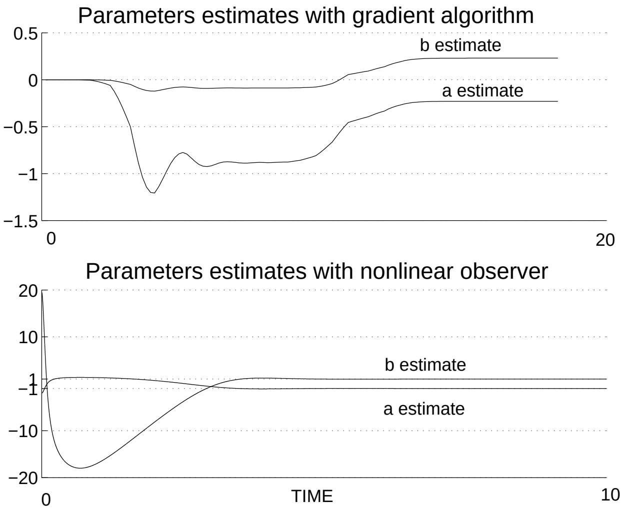

Figure 1: Parameters estimates with a gradient algorithm (upper plot) and with the nonlinear observer approach (lower plot).

output, with $f(0,0) = 0$, and that it is stabilizable by a static function of x, we can consider the augmented system

$$\begin{aligned} \dot{x} &= f(x, \omega_1) \\ \dot{\omega}_1 &= \omega_2 \\ &\vdots \\ \dot{\omega}_{n_u} &= v, \end{aligned}$$

where v is the new control. This system is still stabilizable by a control $v = v(x, \omega)$, and the vector ω is known since it is the state of the controller. Thus we do not need to estimate $\mathbf{u} = \omega$ for constructing the observer for z, and Theorem 2 applies unchanged. Moreover it can be shown that the closed loop system with $v = v(\hat{x}, \omega)$ is stable (see [11]).

Remark 5 In the case that the system (1) is affected by noise, as long as it is bounded, the results of Theorem 2 still hold with the modification that the error $\tilde{z}$ will not be asymptotically zero, but it will be trapped into a ball around zero whose amplitude depends on noise amplitude and on the values of ν and K_o.

3 Examples and Simulations

In this section we provide some examples to illustrate the applicability of the proposed procedure for the estimation of parameters, and show the performance in the case of lack of PE condition (persistence of excitation). We report also one example when adding white noise to the input channel.

Example 1: We want to estimate parameters a and b of the following system

$$\begin{aligned} \dot{x} &= ax + b \\ y &= x \end{aligned}$$

by measuring $y(t)$. Assume $a = -1$ and $b = 1$. We first consider the extended nonlinear system by letting $x_1 = x$, $x_2 = a$, $x_3 = b$ be states:

$$\begin{aligned} \dot{x} &= f(x) \\ y &= Cx \end{aligned}$$

with $f(x) = (x_1 x_2 + x_3, 0, 0)^T$, and $C = (1, 0, 0)$, so that $h(x) = x_1$. The problem is then to estimate the state $x = (x_1, x_2, x_3)^T$ by the output measurements y. We perform first the observability check

$$\text{rank}(dh, dL_f h, dL_f^2 h) = 3,$$

which is verified if $ax + b \neq 0$. Consider the change of coordinates $z_1 = x_1$, $z_2 = \dot{x}_1$, $z_3 = \ddot{x}_1$, and let the system evolve from initial condition $x_1(0) = z_1(0) = 10$. We then construct the observer in the new coordinates z by choosing $K_0 = (15, 75, 125)^T$ so to have eigenvalues $\lambda_i = -5$, with initial conditions $\hat{z}(0) = (10, 1, 1)^T$ that lie in the observable set, and obtain the estimates behavior reported in Figure 1 (bottom). In this example we did not implement the projection algorithm and we monitored the behavior of $\hat{a}(t)\hat{x}(t) + \hat{b}(t)$. Since $ax + b = 0$ is an attracting set for the system, the observer also approaches asymptotically the set $\hat{a}(t)\hat{x}(t) + \hat{b}(t) = 0$. According to Remark 3 the z observer was turned off when $\hat{a}(t)\hat{x}(t) + \hat{b}(t)$ was too close to zero, i.e. at $t = 10$, so that after that time the a and b estimates remain constant. To have a term of comparison, we report also in Figure 1 (top) the performance of a standard gradient-like adaptive law, i.e. $\dot{\hat{a}} = \gamma_1 \tilde{x} \hat{x}$ and $\dot{\hat{b}} = \gamma_2 \tilde{x}$, $\dot{\hat{x}} = \hat{a}\hat{x} + \hat{b} + k\tilde{x}$, where γ_1, γ_2 are the adaptation gains chosen to be 500 and 400 respectively, $k = 100$, and $\tilde{x} = x - \hat{x}$.

Example 2: We want to find the values of a_{11}, a_{12}, a_{21}, a_{22} for the system

$$\begin{aligned} \dot{x}_1 &= a_{11} x_1 + a_{12} x_2 + u_1 \\ \dot{x}_2 &= a_{21} x_1 + a_{22} x_2 + u_2 \\ y &= (x_1, x_2)^T \end{aligned}$$

by observing the output $y(t)$, with $u_1 = 2$, $u_2 = 0$, $a_{11} = -1$, $a_{12} = -2$, $a_{21} = 1$, $a_{22} = 1$. We then consider the related nonlinear system with $x_1 = x_1$, $x_2 = x_2$, $x_3 = a_{11}$, $x_4 = a_{12}$, $x_5 = a_{21}$, $x_6 = a_{22}$, so that the problem of estimating the parameters of the above system is analogous to estimating the state $x \in \mathbb{R}^6$ of the system

$$\begin{aligned} \dot{x} &= f(x) \\ y &= h(x) \end{aligned} \tag{10}$$

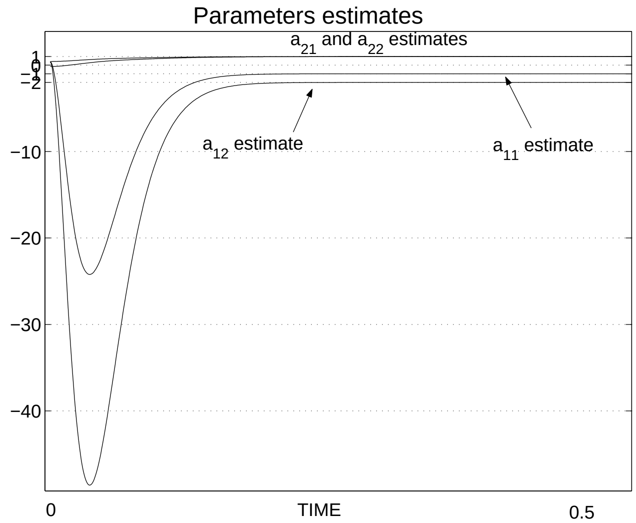

Figure 2: Parameters estimates for the nonlinear observer approach: estimation of four parameters.

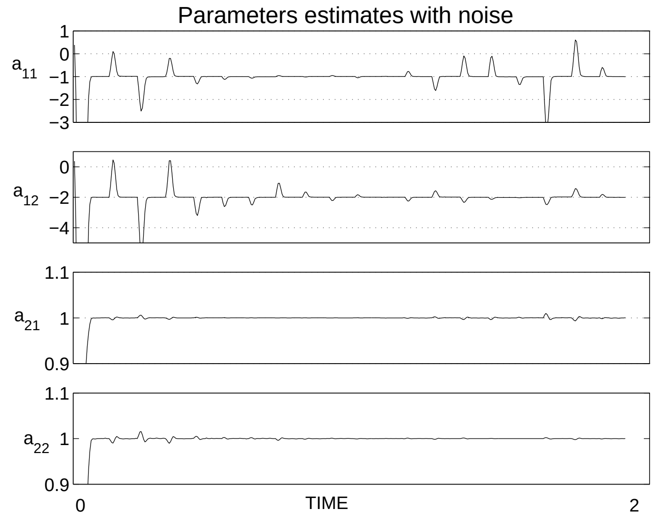

Figure 3: Parameters estimates for the nonlinear observer approach: estimation of four parameters in presence of noise.

with $f(x) = (x_3x_1 + x_4x_2 + 2, x_5x_1 + x_6x_2, 0, 0, 0, 0)^T$ and $h(x) = (x_1, x_2)^T$. We proceed by doing the observability check according to equation (3), which is verified in those regions of the state space where

$$\dot{x}_1x_2 - x_1\dot{x}_2 \neq 0,$$

so that the system is locally observable. Perform the change of coordinates $z_1 = x_1$, $z_2 = \dot{x}_1$, $z_3 = \ddot{x}_1$, $z_4 = x_2$, $z_5 = \dot{x}_2$, $z_6 = \ddot{x}_2$ so to apply results of Theorem 2. We let the system evolve from the initial condition $x_1(0) = 20$ and $x_2(0) = -10$. For the observer we choose initial conditions $(20, 6, 6, -10, 10, 1)^T$ that lie in the observable set, and set the eigenvalues of the observer to -50 and $\nu = 0.1$. The set $\dot{\hat{x}}_1\hat{x}_2 - \hat{x}_1\dot{\hat{x}}_2 = 0$ was never approached, so that the observer was never turned off. In fact system (10) always evolves in the observable set $\dot{x}_1x_2 - x_1\dot{x}_2 < 0$. We obtain good estimates of the a_{ij}'s within 0.5 seconds from the beginning of the simulation as shown in Figure 2. Moreover we report in Figure 3 the result of parameter estimation for the same system in the case in which we add noise on the input channel of the x_1 dynamics. The noise is band limited white noise of amplitude 0.3. As we can see in the figure the parameters estimates are still stable and converge to a neighborhood of the true values

4 Conclusion

We have proposed the use of nonlinear geometric tools for the estimation of parameters in linearly and nonlinearly parameterized systems. We set the problem as an observability problem for an enlarged dynamical system containing the parameters as states. A joint parameter-state observer which is exponentially stable with arbitrary convergence speed was constructed. This approach allows to obtain good parameter estimates also in cases that lack persistence of excitation (see Example1.) Unfortunately Theorem 2 holds only if the state of the enlarged system is guaranteed to evolve in an observable set for all time. If the state is not assured to stay in such regions for all time, we proposed, in Remark 3, a possible way in which one could implement the observer in practice. This is shown in Example 1. The examples proposed suggest that the transient behavior of a system may contain enough information to estimate with high accuracy the values of the parameters from output measurements.

The possibility of designing a control law that guarantees that the system evolves in an observable set for all time needs to be investigated. The observer proposed can be shown to be robust with respect to bounded disturbances, but we did not address the issue of robustness with respect to model error (as it is done in [7] for example), which may be investigated in future work. One interesting application of the proposed approach is the estimation of parameters in piecewise linear systems, where each dynamical system, possibly driven from a step input, never reaches a steady state because a switch to another dynamical model occurs. In this case we cannot rely on asymptotic results, but we need tools for parameter estimation that can extract information from the transient of the system to find an accurate estimate of its parameters. The applicability of the nonlinear geometric approach to this problem needs to be explored in further work.

5 Acknowledgments

This project has been funded in part by the NSF Engineering Research Center for Neuromorphic Systems Engineering (CNSE) at Caltech (NSF9402726). The authors would like to acknowledge the reviewers for their comments.

References

[1] Brian. D. O. Anderson. Exponential stability of linear equations arising in adaptive identification. *IEEE Trans. on Automatic Control*, AC-22(1):83–88, February 1977.

[2] K. J. Astrom and P. Eykhoff. System identification-a survey. *Automatica*, 7:123–162, 1971.

[3] K. J. Astrom and B. Wittenmark. *Adaptive Control*. Addison-Wesley, 1995.

[4] G.C. Goodwin and K S Sin. *Adaptive Filtering Prediction and Control*. Prentice-Hall,NJ, 1984.

[5] R. Hermann and A. J. Krener. Nonlinear controllability and observability. *IEEE Trans. on Automatic Control*, 22:728–740, October 1977.

[6] A. Isidori. *Nonlinear Control Systems*. Springer, London, 1995.

[7] Z. P. Jiang and L. Praly. Design of robust adaptive controllers for nonlinear systems with dynamic uncertainties. *Automatica*, 34:825–840, 1998.

[8] H.K. Khalil. *Nonlinear Systems*. Prentice Hall, New Jersey, 1996.

[9] M. Kristic, I. Kanellakopoulos, and P. Kokotovic. *Nonlinear and Adaptive Control Design*. John Wiley and Sons, 1995.

[10] L. Ljung. Asymptotic behavior of the extended kalman filter as parameter estimator for linear systems. *IEEE Trans. on Automatic Control*, 24(1):36–50, February 1979.

[11] M. Maggiore and K. Passino. Output feedback control of stabilizable and incompletely observable systems: Theory. In *Proc. of the 2000 American Control Conference*, pages 3641–3645, Chicago, IL, 2000.

[12] M. Maggiore and K. Passino. Robust output feedback control of incompletely observable nonlinear systems without input dynamic extension. In *Proc. of 39th CDC*, pages 2902–2908, Sidney, Australia, 2000.

[13] R. Marino and G. Santosuosso. Robust adaptive observers for nonlinear systems with bounded disturbances. In *Proc. of 38th CDC*, pages 5200–5205, Phoenix, Arizona, 1999.

[14] J. Pomet and L. Praly. Adaptive nonlinear regulation: Estimation from the Lyapunov equation. *IEEE Trans. on Automatic Control*, 37:729–740, October 1992.

[15] S. Sastry. *Nonlinear Systems*. Springer, New York, 1999.

[16] S. Sastry and M. Bodson. *Adaptive Control: Stability, Convergence, Robustness*. Prentice-Hall Advanced Reference Series, 1989.

[17] E. D. Sontag. *Mathematical Control Theory*. Springer, New York, 1998.

[18] A. Tornambé. High-gain observers for non-linear systems. *International Journal of Systems Science*, 23:1475–1489, 1992.