Systems & Control Letters 61 (2012) 203–211

Contents lists available at SciVerse ScienceDirect

Systems & Control Letters

journal homepage: www.elsevier.com/locate/sysconle

# Efficient a-posteriori error estimation for nonlinear kernel-based reduced systems

D. Wirtz *, B. Haasdonk

*Institute of Applied Analysis and Numerical Simulation, University of Stuttgart, Pfaffenwaldring 57, D-70569 Stuttgart, Germany*

ARTICLE INFO

*Article history:*
Received 18 February 2011
Received in revised form
14 October 2011
Accepted 25 October 2011
Available online 3 December 2011

*Keywords:*
Nonlinear dynamical systems
Model reduction
Kernel methods
a-posteriori error estimates
Offline/online decomposition
Subspace projection

ABSTRACT

In this paper, we consider the topic of model reduction for nonlinear dynamical systems based on kernel expansions. Our approach allows for a full offline/online decomposition and efficient online computation of the reduced model. In particular, we derive an a-posteriori state-space error estimator for the reduction error. A key ingredient is a local Lipschitz constant estimation that enables rigorous a-posteriori error estimation. The computation of the error estimator is realized by solving an auxiliary differential equation during online simulations. Estimation iterations can be performed that allow a balancing between estimation sharpness and computation time. Numerical experiments demonstrate the estimation improvement over different estimator versions and the rigor and effectiveness of the error bounds.

© 2011 Elsevier B.V. All rights reserved.

## 1. Introduction

Nowadays, modeling of real world processes like biochemical reactions or electric circuits naturally leads to a formulation as dynamical system with inputs and outputs. Mathematically, they are described by systems of ordinary differential equations (ODEs) and can be roughly categorized into linear and nonlinear types. Even though computational power has significantly increased over the past years, high-resolution models often result in large-scale dynamical systems that are expensive to simulate. In this context, the need for fast simulation is particularly evident in many-query scenarios. They comprise parameter studies or inverse problems where multiple simulations have to be performed for different inputs or initial state configurations. Additionally, some dynamical systems model processes in a real-time setting like control components and thus also require fast computation without strong hardware. Within all those settings fast and rigorous error estimation procedures are important in order to quantify the model errors introduced by the reduction process.

Consequently, model reduction techniques for dynamical systems are nowadays subject to intensive research. Work has been done for various types of linear systems, e.g. time invariant [1], time-variant [2] or parameterized systems [3]. For an overview, we refer the reader to [4–6].

Reduction techniques for nonlinear dynamical systems have been less investigated, not at least because of the arising difficulties. As with linear systems, most of them involve projection of the system into a lower dimensional subspace. One well known method for nonlinear model reduction is the trajectory piecewise linear (TPWL) approach [7], which is extended by moment matching techniques in [8] or to a piecewise-polynomial scheme in [9]. Various extensions of the balanced truncation procedure [10] to nonlinear systems are investigated in [11–13]. An "approximate reduction" method is introduced in [14] and model reduction for weakly nonlinear systems by bilinearization has been discussed in [15], for example. In addition, the special class of bilinear quadratic nonlinearities has recently been investigated in [16]. A further promising technique for nonlinear model reduction of dynamical systems is by means of discrete empirical interpolation [17]. Finally, computing subspaces by proper orthogonal decomposition (POD) of presumably statistically representative trajectories [18] is an expensive but well-established method. This method is also known as principal component analysis (PCA); see [19] and references therein for an overview.

In this work, we adopt principles of the reduction technique proposed in [20], extend the class of usable kernels and provide novel, efficient a-posteriori error estimators for the resulting reduced systems. The derived error estimators are tested using a

* Corresponding author. Tel.: +49 711 685 65549.
*E-mail addresses:* daniel.wirtz@mathematik.uni-stuttgart.de (D. Wirtz), bernard.haasdonk@mathematik.uni-stuttgart.de (B. Haasdonk).
*URLs:* http://www.agh.ians.uni-stuttgart.de/orga/people/wirtz (D. Wirtz), http://www.agh.ians.uni-stuttgart.de/orga/people/haasdonk (B. Haasdonk).

0167-6911/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.sysconle.2011.10.012

synthetic kernel-based example system which resembles typical real system's properties like varying inputs, and the benefits of different estimator versions are shown. Possible applications of the reduction technique involve electric circuits [20] or biochemical systems [21], or more generally, any systems where a good approximation of the nonlinearity with few kernel components can be found. Of course any system given by a kernel expansion can be straightforwardly reduced with rigorous error bounds.

In Section 2 we introduce our base dynamical system and discuss aspects of the reduction process. Section 3 is concerned with a-posteriori error estimates and introduces our new local Lipschitz constant estimation method. Next, Section 4 presents numerical experiments for synthetic dynamical systems and we conclude with Section 5. Some auxiliary mathematical details are presented in Appendix.

## 2. Reduction of kernel based systems

The starting point of our investigations is the reduction method introduced in [20], which combines subspace projection with kernel methods. In model reduction, the latter are used to approximate the nonlinearities of dynamical systems in kernel spaces and promise a large potential in the field. Kernel methods comprise applications from machine learning like support vector regression [22,23] as well as kernel interpolation with corresponding theoretical foundation [24,25]. However, those approximation techniques are out of the scope of this article. Instead we focus on nonlinear systems whose inner dynamics are already given by a kernel expansion and investigate the projection behavior and the resulting a-posteriori error estimators.

### 2.1. The base dynamical system

Our central assumption is to have a nonlinear kernel expansion

$$f(x) = \sum_{i=1}^{N} c_i \Phi(x, x_i), \tag{1}$$

where $X := \{x_1, \ldots, x_n\} \subset \mathbb{R}^d$ are the centers or *support vectors* of the expansion and $c_i \in \mathbb{R}^d$ the coefficient vectors. The function $\Phi$ is a *kernel*, which is basically a symmetric function $\Phi: \mathbb{R}^d \times \mathbb{R}^d \longrightarrow \mathbb{R}$. We omit further characterizations of kernels and mention here that the key analytical property is that for some open $\Omega \subset \mathbb{R}^d$ each symmetric positive definite kernel spans a unique Hilbert space $\mathcal{N}_\Phi(\Omega)$ of functions. Those Hilbert spaces have a special reproducing property and are commonly referred to as reproducing kernel Hilbert spaces (RKHS). In our context those RKHS serve as a base class for the dynamical systems functions $f$, so (1) implies $f \in (\mathcal{N}_\Phi(\Omega))^d$. For more details on kernels, RKHS and their applications we refer to [26,24,27], for example.

Now, the base system considered throughout this paper is of the form

$$x'(t) = f(x) + Bu(t), \tag{2}$$

$$x_0 = x_0, \qquad y(t) = Cx(t), \tag{3}$$

where $x(t) \in \mathbb{R}^d$ denotes the state of the system at times $t \in [0, T]$, $0 \le T < \infty$, $x_0 \in \mathbb{R}^d$ the initial condition, $u: [0, T] \to \mathbb{R}^m$ an input function with $B \in \mathbb{R}^{d\times m}$ and output $y(t) \in \mathbb{R}^k$ with $C \in \mathbb{R}^{k\times d}$.

Additionally, in some applications different norms than the standard Euclidean norm are taken for the state space. Thus, let $G \in \mathbb{R}^{d\times d}$ be a symmetric positive definite matrix. Then $G$ defines a scalar product $\langle x, y\rangle_G := x^t G y$ on $\mathbb{R}^d$ with induced norm $||x||_G := \sqrt{\langle x, x\rangle_G}$. Choosing $G := I_d$ will yield the standard Euclidean inner product $\langle \cdot, \cdot \rangle$ and 2-norm $||\cdot||$, where $I_d$ denotes the $d$-dimensional identity matrix. For example, $G$ is usually chosen as the Gram-matrix of the finite element/finite volume basis for dynamical systems obtained from partial differential equation discretizations.

Having introduced the underlying dynamical system (2)–(3) we now detail the reduction approach using subspace projection. We assume to have two biorthogonal projection matrices $V, W \in \mathbb{R}^{d\times r}$, $W^t V = I_r$ with $r \ll d$ denoting the reduced system's dimension. The reduced system is then obtained by applying a Petrov–Galerkin projection to the full system:

$$z'(t) = W^t f(Vz(t)) + W^t Bu(t), \tag{4}$$

$$z(0) = W^t x_0, \qquad y^r(t) = CVz(t), \tag{5}$$

where $z(t) \in \mathbb{R}^r$ now denotes the reduced system's state and $y^r(t)$ the approximate output. Let $U := \langle v_1, \ldots, v_r\rangle \subset \mathbb{R}^d$ be the space spanned by the columns $v_i \in \mathbb{R}^d$ of $V$. Consequently, the reconstructed approximate solution and output are given by $x^r(t) := Vz(t) \in U$ and $y^r(t) = Cx^r(t)$. The requirements on $V, W$ are now to yield a good approximation $x^r(t) \approx x(t)$ and thus $y^r(t) \approx y(t)$ for different initial values and inputs. For the remainder of this work we will assume the matrix computation method for $V, W$ as a black-box since our method is applicable using any basis/projection matrix generation method.

The rest of this section is concerned with the different projection aspects of the reduction technique for kernel-based dynamical systems. We will see that the structure of the kernel expansion allows for dramatic reduction of the computational costs using certain kernels.

### 2.2. System projection

At first, the projection into $\mathbb{R}^r$ by $W^t$ can be applied directly to the coefficient vectors $c_i$ via $c_i^r := W^t c_i$, $i = 1 \cdots N$. This results in a reduced function $f^r(z) := \sum_{i=1}^N c_i^r \Phi(Vz, x_i)$, with new coefficient vectors $c_i^r \in \mathbb{R}^r$.

Unfortunately, the reduced system (4) is still expensive to simulate given that $f$ is evaluated at $x^r(t) \in \mathbb{R}^d$. In order to avoid input arguments of high dimension $d$, we consider two special classes of kernels.

#### 2.2.1. Inner product kernels

The first type of kernels that allow efficient argument evaluations are the inner product kernels also mentioned in [20]. It is assumed that $\Phi(x, y) = \phi(\langle x, y\rangle_G)$ for some scalar function $\phi: \mathbb{R} \to \mathbb{R}$. Then $\Phi(Vz, x_i) = \phi(\langle Vz, x_i\rangle_G) = \phi(\langle z, z_i\rangle) =: \Phi^r(z, z_i)$ for $z_i := V^t G x_i \in \mathbb{R}^r$, $i = 1 \cdots N$. This way, it is sufficient to project the centers $x_i$ into the low-dimensional subspace and the evaluations of $\Phi$ can be computed efficiently and *loss-less* during the reduced simulation via $\Phi^r$. Some examples for those kernels are the linear kernel $\Phi(x, y) = \langle x, y\rangle_G$ or polynomial kernels $\Phi(x, y) = (1 + \langle x, y\rangle_G)^p$ of degree $p \in \mathbb{N}$.

#### 2.2.2. Translation- and rotation invariant kernels

In extension to [20] we present here a further class of kernels that allow for efficient *loss-less* argument evaluations: The translation and rotation-invariant kernels $\Phi(x, y) := \phi(||x - y||_G)$ for some scalar function $\phi: \mathbb{R}_0^+ \to \mathbb{R}$. We impose the additional requirement $x_i \in U$, i.e. $x_i = Vz_i$ for some $z_i \in \mathbb{R}^r$, $i = 1 \cdots N$. Then we obtain $\Phi(Vz, x_i) = \phi(||Vz - Vz_i||_G) = \phi(||z - z_i||_{V^t GV}) =: \Phi^r(z, z_i)$ with $V^t GV$ being a small $\mathbb{R}^{r\times r}$ matrix inducing a new norm on $\mathbb{R}^r$. Note that the assumption $x_i \in U$ is of a technical nature. We either extend $U$ by the span of the $x_i$, or, if the kernel expansion is created with knowledge of $U$, one can choose $x_i \in U$ in the first place. Those kernels are also commonly referred to as *radial basis functions*, of which the Gaussian kernel $\Phi(x, y) = \exp(-||x - y||^2/\beta^2)$ is probably the most popular example. For further examples and characterizations of the above mentioned kernels we refer to [22,27].

Download English Version:

https://daneshyari.com/en/article/751976

Download Persian Version:

https://daneshyari.com/article/751976

Daneshyari.com