Spectral Collocation-based Optimization in Parameter Estimation for Nonlinear Time-Varying Dynamical Systems

Venkatesh Deshmukh

Center for Nonlinear Dynamics and Control
Department of Mechanical Engineering
Villanova University

ABSTRACT

A constructive optimization algorithm using Chebyshev spectral collocation and quadratic programming is proposed for unknown parameter estimation in nonlinear time-varying dynamic system models to be constructed from available data. The parameters to be estimated are assumed to be identifiable from the data which also implies that the assumed system models with known parameter values have a unique solution corresponding to every initial condition and parameter set. The nonlinear terms in the dynamic system models are assumed to have a known form, and the models are assumed to be parameter affine. Using an equivalent algebraic description of dynamical systems by Chebyshev spectral collocation and data, a residual quadratic cost is set up which is a function of unknown parameters only. The minimization of this cost yields the unique solution for the unknown parameters since the models are assumed to have a unique solution for a particular parameter set. An efficient algorithm is presented step-wise and is illustrated using suitable examples. The case of parameter estimation with incomplete or partial data availability is also illustrated with an example.

Keywords: NLTV systems, parameter estimation, Chebyshev spectral collocation

Department of Mechanical Engineering
800 Lancaster Avenue, Villanova, PA 19085.
E-mail: venkatesh.deshmukh@villanova.edu
1. Introduction

Parameter estimation in nonlinear ordinary differential equations with general time-varying coefficients is the central theme that is encountered in biological and engineering systems modeling. The experimental or clinical data is usually available and, the linear system matrix and the forms of nonlinearities appearing in the models are known from mechanistic or physical viewpoint. It is required to calculate or determine the coefficients that accompany the linear or nonlinear terms of variables (states) in the model. These coefficients have to be determined so that the available data is expressed as the solution of the assumed models, with parameter values in place. An ideal situation is considered here: the models are parameter affine or are linear in terms of the parameters and there is no disturbance acting on the system. To the author’s knowledge, there has not been an attempt as yet reported for parameter estimation in nonlinear models with time-varying coefficients in the literature.

Parameter estimation in nonlinear time-invariant nonlinear differential equations from experimental or clinical data that represent the mechanistic behavior of a biological system or a process is an extensively researched area in bioinformatics [Voit]. Models of spread of diseases (epidemiology) such as West Nile virus [Pacheko, Bowman] and AIDS [Capasso], models of disease progression in a single human for HIV AIDS [Bajaria], diabetes [Kansal, Cobelli, Topp], models of circadian rhythms [Goldbeter], gene regulatory networks [Chen, Smolen], metabolism [Covert, Bradshaw] and biochemical networks [Mendes1] and competition systems [de Mottoni] are all nonlinear ordinary differential equations with unknown parameters multiplying linear and nonlinear terms of the model states. Every major pharmaceutical company has a modeling group that is involved in drug discovery [Butcher] and computer simulation of drug absorption [Grass] in human intestine based on clinical data. The parameter estimation methods in biological systems include Bayesian estimation [Tan], \textit{brute-force} simulation by numerical integration with Latin Hypercube Sampling [Bajaria] and nonlinear optimization programs using global optimization methods [Mendes2]. There are private companies dedicated only to the process modeling (and parameter estimation) of diseases and have their own platforms (application software) to facilitate modeling. Academia also has contributed by creating software programs such as COPASI and GEPASI [Mendes3], MetMAP [Vera].

Health monitoring and damage detection for civil, mechanical and electro-mechanical engineering systems involves tracking the variations of certain parameters. Parameter variation estimation is performed by the available methods such as scalar tracking metric [Chalidze] and monitoring bifurcation morphing [Yin], and the research work reported therein. An exhaustive survey is avoided here since the method presented in the present manuscript is not an attractor-based or sensitivity-based method.

The abovementioned approaches have been demonstrated for nonlinear differential equation models with constant coefficients. These approaches do have the potential to be extended to time-varying nonlinear models but it has not been demonstrated yet. There are instances in biological and engineering systems modeling, when the model contains the nonlinear terms that have time periodic or general varying coefficients. The method proposed in the present manuscript for parameter estimation is applicable to generic nonlinear time-varying systems and uses Chebyshev spectral
collocation followed by quadratic programming. There are certain assumptions made to
guarantee the uniqueness of parameter determination, without the loss of generality. The
proposed method does not use any feedback to change the fixed point of the nonlinear
system [Epureanu] or iterative parameter variations, instead is a “one-shot” method for
estimation. Also, spectral collocation is used for the first time in an estimation setting.
Chebyshev collocation is used to convert a dynamic system (model) into an algebraic
system with unknown parameters to yield the residual vector which is a function of the
unknown parameters. A cost function is formulated using the square of the $L_2$-norm of
the residual vector, which is quadratic in unknown parameters. Solving a quadratic
program yields the solution for unknown parameters. The method is illustrated using
three examples of benchmark nonlinear systems with full data (entire state vector)
availability and one example of partial data availability for parameter identification.

2. Algebraic Description of Dynamical Systems

2.1 Chebyshev Spectral Collocation

Consider a local description of a nonlinear system given by

$$[h(x,t) + \sum_{i=1}^{q} \eta_i \, k_i(x,t)] \ddot{x} = A(t) \dot{x} + f(x,t) + \sum_{i=1}^{p} \lambda_i \, g_i(x,t)$$  \hspace{1cm} (1)

where $x(t)$ is an $n$-dimensional state vector, $A(t)$ is an $n \times n$ bounded, time-varying
system matrix, $h(x,t)$ is an $n \times n$ nonlinear inertia matrix of bounded time variation,
$k_i(x,t), \ i=1,2,...,q$ are $n \times n$ nonlinear matrices of bounded time variation, $f(x,t)$ and
g_i(x,t), \ i=1,2,...,p$ are $n$-dimensional vector nonlinearities of the state vector having
known forms and of bounded variation in time, $\eta_i, \ i=1,2,...,q; \ \lambda_i, \ i=1,2,...,p$ are the
unknown constant parameters. The nonlinear matrix $h(x,t)$ and $k_i(x,t)$ include the
purely time-varying bounded (and constant) matrices which are independent of the state
in addition to linear and nonlinear functions of the state. The vector functions $g_i(x,t)$
include the terms that are linear functions of the state and functions that are independent
of the state and of bounded time variation, in addition to the nonlinear terms. For brevity,
a particular parameter set $\{\eta_1, \eta_2,...,\eta_q; \lambda_1, \lambda_2,...,\lambda_p\}$ is represented by a point $(\bar{\eta}, \bar{\lambda})$ in
an $(p+q)$-dimensional space of parameters.

The following assumptions are made.

A. The state vector is available in the form of experimental or clinical data at sampling
points of time. The data is represented by $x(t_0,t_f)$ which also specifies the initial time $t_0$
and the terminal time $t_f$. There are $N_d$ unequally spaced sampling points in $[t_0,t_f]$. 

B. The form of the nonlinear vector functions \( f(x,t), k_i(x,t), i=1,2,...,q \) and \( g_i(x,t), i=1,2,...,p \) is known which could be polynomial, trigonometric or any Lipschitz continuous functions of the state.

C. The unknown parameters are constant.

D. There exists a unique solution of system (1) with known parameters corresponding to every initial condition \( x(0) \) and parameter set \( \lambda \). This assumption can also be stated in terms of the identifiability [Tunali] of system (1) with the system state as the system output. For the sake of completeness, this property is stated as follows.

System (1) is \( x(0) \)-identifiable at \((\tilde{\eta}, \tilde{\lambda})\) if there exists an open set \( P^0 \subset P \) containing \((\tilde{\eta}, \tilde{\lambda})\), such that for any two \((\tilde{\eta}_1, \tilde{\lambda}_1)\), \((\tilde{\eta}_2, \tilde{\lambda}_2)\) \( \in P^0 \), \((\tilde{\eta}_1, \tilde{\lambda}_1) \neq (\tilde{\eta}_2, \tilde{\lambda}_2)\), the solutions \( x(t,x(0),\tilde{\eta}_1, \tilde{\lambda}_1) \) and \( x(t,x(0),\tilde{\eta}_2, \tilde{\lambda}_2) \) exist on \([0,\delta]\), \( 0 < \delta \leq T \), and their corresponding outputs (states) satisfy on \( t \in [0,\delta] \),

\[
y(t,x(0),\tilde{\eta}_1, \tilde{\lambda}_1) = x(t,x(0),\tilde{\eta}_1, \tilde{\lambda}_1) \neq x(t,x(0),\tilde{\eta}_2, \tilde{\lambda}_2) = y(t,x(0),\tilde{\eta}_2, \tilde{\lambda}_2).
\]

The purpose of this paper is to present a simple computational algorithm to compute \((\tilde{\eta}, \tilde{\lambda})\) so that \( \tilde{x}(t_0,t_f) \) is a solution of system (1), with the above-mentioned assumptions. The task is accomplished by using Chebyshev spectral collocation to get the algebraic description of the dynamical system model (1) in terms of the unknown \( \lambda \) in order to set up a residual. This residual forms an integral part of the cost function to be minimized later on. Therefore, in this section, the process of arriving at the residual is explained in detail.

Gauss-Chebyshev-Lobatto points, or Chebyshev extreme points [Trefethen], or merely Chebyshev points, for brevity, are the points in the interval \([t_*, t_* + \tau]\) defined by

\[
t^j = t_* + \left(\frac{\tau}{2}\right)[\cos(j\pi/(N-1)) + 1] \quad j = 0,1,2,...,N-1
\]

Note that \( t_* + \tau = t^0 > t^1 > ... > t^{N-1} = t_* \), a standard ordering for this method [Trefethen]. For this set of \( N \) Chebyshev points we also have an \( N \times N \) Chebyshev spectral differential matrix [Trefethen]. For each \( N \geq 1 \), let the rows and columns of the \( N \times N \) Chebyshev spectral differentiation matrix \( D_N \) be indexed from 1 to \( N \). The entries of this matrix are

\[
(D_N)_{00} = \frac{2(N-1)^2 + 1}{6}, \quad (D_N)_{NN} = -\frac{2(N-1)^2 + 1}{6}
\]  

(3-a)

\[
(D_N)_{jj} = \frac{-t_j}{2(1-t_j)^2}, \quad j = 1,...,N-2
\]  

(3-b)
\begin{align*}
(D_N)_{ij} &= \frac{c_i(-1)^{i+j}}{c_j(t_i - t_j)}, \quad i \neq j, \quad j = 0, \ldots, N - 1 \\
c_i &= \begin{cases} 
2 & i = 0 \text{ or } N - 1 \\
1 & \text{otherwise}
\end{cases}
\end{align*}

The dimension of $D_N$ is $N \times N$, where $N$ is the number of Chebyshev points. If $n$ is the dimension and $N$ is the number of points considered in a given interval, and $I_n$ is the $n \times n$ identity matrix, then the differential operator $d/dt$ is defined using the Kronecker product as

$$
\overline{D} = D_N \otimes I_n
$$

The algebraic description is obtained by writing system (1) at the Chebyshev points (2) as follows.

$$
\begin{bmatrix}
h(x(t^0),t^0) \\
\vdots \\
h(x(t^{N-2}),t^{N-2}) \\
\end{bmatrix} + \sum_{i=1}^{n} \eta_i 
\begin{bmatrix}
k_i(x(t^0),t^0) \\
\vdots \\
k_i(x(t^{N-2}),t^{N-2}) \\
\end{bmatrix}
\begin{bmatrix}
x(t^0) \\
\vdots \\
x(t^{N-2}) \\
\end{bmatrix} =
\begin{bmatrix}
0 \\
\vdots \\
0
\end{bmatrix}
$$

System of equations (5) is an $Nn$ dimensional system and is called the algebraic description of system (1) over an arbitrary interval $[t, t + \tau]$ in the real time. The last vector appearing in representation (5) has the provision for the initial point data to be included in the description which guarantees compatibility between the two adjacent intervals.

2.2 The Residual and Residual Cost

Equation (5) provides an extremely useful tool for the construction of the quadratic cost in terms of the parameters. By assumption, the entire state vector is
available in the form of data over an arbitrary interval \([t_*, t_* + \tau]\). The only point to be noted at this stage is that the data is not available at the Chebyshev points given by equation (2). Therefore, the data \(x(t_*, t_* + \tau)\) is interpolated at Chebyshev points by using an appropriate interpolation algorithm. Let \(z(t_*, t_* + \tau)\) be the interpolated data at Chebyshev points, then substituting it in equation (5) should satisfy equation (5) exactly, if the parameter values were known. Substituting \(z(t_*, t_* + \tau)\) in (5) yields

\[
\begin{bmatrix}
  h(z(t^0), t^0) \\
  \vdots \\
  h(z(t^{N-2}), t^{N-2}) \\
  I_n
\end{bmatrix}
+ \sum_{i=1}^{q} \eta_i 
\begin{bmatrix}
  k_i(z(t^0), t^0) \\
  \vdots \\
  k_i(z(t^{N-2}), t^{N-2}) \\
  0
\end{bmatrix} 
= \begin{bmatrix}
  \otimes \\
  \otimes \\
  \otimes \\
  \otimes
\end{bmatrix}
\begin{bmatrix}
  z(t^0) \\
  z(t^{N-2}) \\
  z(t^{N-1}) \\
  IC
\end{bmatrix} 
= \begin{bmatrix}
  A(t^0) \\
  \vdots \\
  A(t^{N-2}) \\
  0
\end{bmatrix}
\begin{bmatrix}
  z(t^0) \\
  z(t^{N-2}) \\
  z(t^{N-1}) \\
  0
\end{bmatrix} 
+ \sum_{i=1}^{p} \lambda_i 
\begin{bmatrix}
  f_i(z(t^0), t^0) \\
  f_i(z(t^{N-2}), t^{N-2}) \\
  0 \\
  0
\end{bmatrix} 
+ \begin{bmatrix}
  g_i(z(t^0), t^0) \\
  g_i(z(t^{N-2}), t^{N-2}) \\
  0 \\
  IC
\end{bmatrix} \Rightarrow (6)
\]

with \(t_* + \tau = t^0 > t^1 > ... > t^{N-1} = t_*\) and \(IC = z(t_*)\).

Equation (6) is exactly satisfied for only one set of values of \(\lambda = [\lambda_1, \lambda_2, \ldots, \lambda_p]^T\) because of the identifiability of parameters assumption (D). Thus equation (6) written in the form
\[
\begin{bmatrix}
    h(z(t^0), t^0) \\
    \vdots \\
    h(z(t^{N-2}), t^{N-2}) \\
    I_n
\end{bmatrix}
+ \sum_{i=1}^{p} \eta_i
\begin{bmatrix}
    k_i(z(t^0), t^0) \\
    \vdots \\
    k_i(z(t^{N-2}), t^{N-2}) \\
    0
\end{bmatrix}
\begin{bmatrix}
    z(t^0) \\
    \vdots \\
    z(t^{N-2}) \\
    z(t^{N-1})
\end{bmatrix}
\begin{bmatrix}
    \overline{D}_n \otimes I_{non} \\
    \vdots \\
    \overline{D}_n \otimes I_{non}
\end{bmatrix}
\]

\[
\begin{bmatrix}
    A(t^0) \\
    \vdots \\
    A(t^{N-2}) \\
    0
\end{bmatrix}
\begin{bmatrix}
    z(t^0) \\
    \vdots \\
    z(t^{N-2}) \\
    z(t^{N-1})
\end{bmatrix}
- \begin{bmatrix}
    f(z(t^0), t^0) \\
    \vdots \\
    f(z(t^{N-2}), t^{N-2}) \\
    0
\end{bmatrix}
- \sum_{i=1}^{p} \lambda_i
\begin{bmatrix}
    g_i(z(t^0), t^0) \\
    \vdots \\
    g_i(z(t^{N-2}), t^{N-2}) \\
    0
\end{bmatrix}
- \begin{bmatrix}
    0 \\
    \vdots \\
    0
\end{bmatrix}
\begin{bmatrix}
    IC
\end{bmatrix}
= \varepsilon(\overline{\eta}, \overline{\lambda})
\]

is a residual vector that assumes zero norm when the actual parameter values are substituted. It should also be noted that the residual is a linear function of the parameters since the system is parameter affine. The residual cost for the interval \([t_0, t_0 + \tau]\) is written as

\[
E(\overline{\eta}, \overline{\lambda}) = (\varepsilon(\overline{\eta}, \overline{\lambda}))^T \varepsilon(\overline{\eta}, \overline{\lambda})
\]  

(8)

The residual cost is a quadratic function of the parameters.

3. The Estimation Algorithm

Now, the estimation algorithm is formulated. The data \(\overline{x}(t_0, t_f)\) is available for \([t_0, t_f]\).

1. The interval \([t_0, t_f]\) is partitioned into \(n_e\) intervals of time \([t_k, t_k + \tau]\), \(k = 0, 1, ..., n_e - 1\) with \(t_k + \tau = t_{k+1}\) and \(t_f = t_{n_e} + \tau\).
2. For each interval, the data \(\overline{x}(t_k, t_k + \tau)\) at \(N_d\) equally spaced points is interpolated to get \(z(t_k, t_k + \tau)\) at \(N\) Chebyshev points using a suitable efficient interpolation algorithm.
3. For each interval, the residual \(\varepsilon_k(\overline{\eta}, \overline{\lambda})\) is constructed using equation (7).
4. The master cost for the entire interval \([t_0, t_f]\) is constructed as
The master cost is quadratic and positive definite function of \((\overline{\eta}, \overline{\lambda})\).

5. A quadratic program is solved to compute the values of the unknown parameter set \((\overline{\eta}, \overline{\lambda})\).

The estimation algorithm can be implemented efficiently in any of the programming languages available. It should also be noted at this point that if parameter values were known and if the solution of equations (1) were desired, the solution can be computed interval-wise by setting up a residual for equation (5) and using it to set up the cost (8) and then minimizing with respect to \(x\). The minimization is a nonlinear program and can be solved on a PC for smaller benchmark nonlinear systems. In fact, in the following section, even before the estimation algorithm is put to the test, the solution of the nonlinear problem with known parameters is computed by collocation-based optimization and compared to the numerically integrated solution. There is an excellent agreement between the two for the system models under consideration.

The algorithm is also applicable to second order mechanical or structural systems of the form

\[
E_m = \sum_{k=1}^{n_e} e_k^T e_k = \sum_{k=1}^{n_e} E_k(\overline{\eta}, \overline{\lambda})
\]

\[E_m = \sum_{k=1}^{n_e} e_k^T e_k = \sum_{k=1}^{n_e} E_k(\overline{\eta}, \overline{\lambda}) \quad (9)\]

with the just displacement \(x(t)\) available as data and not the velocity. The velocity however, can be estimated by using a spectral differentiation matrix (4) on the displacement data and then, system (10) written in the state-space form exactly resembles system (1). Therefore, the same algorithm with some modifications can be utilized for a wide spectrum of system models. Now, by first order error analysis, the following proposition is stated and proven.

**Proposition 1**: With the assumptions stated in section 2, the estimation algorithm uniquely determines the unknown parameters.

Proof: There are two distinct errors possible in the algorithm.
1. Error in interpolation of data from equally spaced points to Chebyshev points
2. Error in Chebyshev collocation-based cost function and its optimization.
It is shown that these errors are controlled.
1. The error in data interpolation from \( N_d \) equally spaced points to \( N \) Chebyshev points is theoretically zero. The reason is that there exists a \( N_d - 1 \) degree polynomial that exactly fits the equally spaced data. Therefore, evaluating the polynomials that exactly fit the data \( \bar{x}(t_k, t_k + \tau) \) to produce \( z(t_k, t_k + \tau) \) at Chebyshev points is in fact, exact.

2. Now, the products in equation (1) are considered term by term for the data at Chebyshev points given by \( z(t_k, t_k + \tau) \). Let \( er(y) \) denote the absolute error in the variable \( y \). Note that if \( y \) is a vector, \( er(y) \) is also a vector.

The first term is \( h(z,t)\dot{z} \) and

\[
er(h(z,t)\dot{z}) = \sum_{j=1}^{n} er(h_{j}(z,t))\dot{z}_j + \sum_{j=1}^{n} h_{j}(z,t) er(\dot{z}_j) \quad (A)
\]

Now, the error \( er(h_{j}(z,t)) \) is the error in the functions \( h_{j}(z,t) \) due to the error in the data \( z \) at Chebyshev points. In turn, \( er(h_{j}(z,t)) \) can be expressed in the first order approximation as

\[
er(h_{j}(z,t)) = \frac{\partial h_{j}}{\partial z} er(z) \quad (B)
\]

Because \( er(t) \) is zero. Also, error estimates for the other terms are

\[
er(k_{j}(z,t)\dot{z}) = \sum_{j=1}^{n} er(k_{j}(z,t))\dot{z}_j + \sum_{j=1}^{n} k_{j}(z,t) er(\dot{z}_j) \quad (D)
\]

\[
er(A(t)\dot{z}) = \sum_{j=1}^{n} er(A_{j}(t))\dot{z}_j + \sum_{j=1}^{n} A_{j}(t) er(\dot{z}_j) \quad (F)
\]

\[
er(f(z,t)) = \frac{\partial f}{\partial z} er(z) \quad \text{and} \quad er(g_{i}(z,t)) = \frac{\partial g_{i}}{\partial z} er(z) \quad (G)
\]

Equations (A) – (G) show that, the error estimates for the terms in equation (1) and therefore the residual vector is a linear function of the error estimates \( er(z(t)) \) and \( er(\dot{z}(t)) \) for a given data set \( z \). It has been proved that the choice of Chebyshev points is the best choice from the point of view of \( er(z(t)) \) and \( er(\dot{z}(t)) \) being equal to \( C O(-N) \) [Trefethen, Chapter 5, Theorem 5 and 6, respectively], where \( O \) stands for the exponential order of magnitude and \( C \) is a constant independent of \( z \) and \( N \). As far as the choice of collocation is concerned, it has been shown that the implicit Runge-Kutta and Galerkin methods are equivalent to collocation methods, for initial [Hulme, de Boor]
and boundary value problems [Weiss]. Therefore, accuracy and convenience are the two compelling factors to use Chebyshev spectral collocation. The error in the residual is given by

\[ er(\varepsilon) = er(h \dot{z}) + \sum_{i=1}^{q} \eta_i \, er(k_i \dot{z}) - er(A(t) \dot{z}) - er(f) - \sum_{i=1}^{p} \lambda_i \, er(g_i) \]  

which is minimum at Chebyshev points than any other distribution of points [Trefethen] for a fixed \( N \). The residual equation (J) when written at Chebyshev points in a particular interval yields equation (8) and subsequently, equation (9). The residual is a linear function of the unknown parameters and by the “identifiability of the parameters” assumption (assumption D), there exists a unique parameter set to every data set that is a solution of the model (of the form (1)) which is then uniquely computed by a QP.

4. Examples

In this section, the proposed estimation algorithm is put to the test in benchmark nonlinear systems of engineering importance. Since the experimental data is not available here, the general route taken in illustrating the proposed method is as follows.

A. Simulation module: Firstly, the nonlinear system is simulated over a certain interval of time using known parameter values to generate the required data. It is in this module that the effectiveness of collocation-based optimization in computing solutions to the general nonlinear time-varying systems is also illustrated.

B. Estimation module: Secondly, the estimation module which is independent of the simulation module is used to estimate the parameters which are designated as unknowns from the available data. The data generated by the numerical integrator is interpolated at Chebyshev points to be used by the quadratic program.

Example 1: Double inverted pendulum – parameter estimation

A double inverted pendulum with a follower force has equation of motion given by

\[
\begin{bmatrix}
3 & \cos(x_2 - x_1) \\
\cos(x_2 - x_1) & 1
\end{bmatrix}
\begin{bmatrix}
\ddot{x}_1 \\
\ddot{x}_2
\end{bmatrix}
+ \begin{bmatrix}
2k & -k \\
-k & k
\end{bmatrix}
\begin{bmatrix}
x_1 \\
x_2
\end{bmatrix}
+ \begin{bmatrix}
-\sin(x_2 - x_1) \dot{x}_2^2 - \bar{Q} \sin(x_1 - \gamma x_2) \\
\sin(x_2 - x_1) \dot{x}_1^2 - \bar{Q} \sin((1 - \gamma)x_2)
\end{bmatrix} = 0
\]

(11)
where $m$ is an inertial parameter, $\bar{k} = k/l^2$ is the spring stiffness, $\bar{Q} = [Q_1 + Q_2 \cos(\sqrt{l}/l)]/l$ is a follower force with irrational frequency, $x_1$ and $x_2$ are the absolute angular joint displacement variables. The inertial parameter $m$ is an unknown that is to be estimated by the estimation algorithm. The system is written in the state-space form as

\[
\begin{bmatrix}
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 3m & m \cos(x_2 - x_1) \\
0 & 0 & m \cos(x_2 - x_1) & m
\end{bmatrix}
\begin{bmatrix}
x_1 \\
x_2 \\
x_3 \\
x_4
\end{bmatrix}
= \begin{bmatrix}
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1 \\
-2\bar{k} & \bar{k} & 0 & 0 \\
\bar{k} & -\bar{k} & 0 & 0
\end{bmatrix}
\begin{bmatrix}
x_1 \\
x_2 \\
x_3 \\
x_4
\end{bmatrix}
+ \begin{bmatrix}
f_1 \\
f_2
\end{bmatrix}
\] (12)

\[
\begin{bmatrix}
f_1 \\
f_2
\end{bmatrix}
= \begin{bmatrix}
-\sin(x_2 - x_1)x_3^2 - \bar{Q}\sin(x_1 - \gamma x_2) \\
\sin(x_2 - x_1)x_3^2 - \bar{Q}\sin((1 - \gamma)x_2)
\end{bmatrix}
\] (13)

It should be noted that the form of equation (12) is slightly different from the original system considered in equation (1).

\[
[H(x,t) + \sum_{i=1}^{q} \eta_i k_i(x,t)] \dot{x} = A(t) x + f(x,t)
\] (14)

A. Simulation: Figure 1 shows the time trajectories for the state of equation (12) by MATLAB routine ode45 and collocation-based optimization. The parameter set is chosen as $\bar{Q}_1 = 0.1, \bar{Q}_2 = 0.6, \bar{k} = 1.0, \gamma = 1.0$. The value of the inertial parameter is $m = 1$. The initial condition is chosen as $x(0) = [0.22, 0.22, 0, 0]^T$, the number of Chebyshev points $N = 30$.

B. Estimation: The number of estimation intervals $n_e = 5$ and the estimation horizon $\tau = 3.0982$. There was no bias while choosing the parameter values. From the data generated in figure (1), the residual cost (9) is set up for system (12) with $m$ as an unknown parameter. The estimation algorithm computes $m = 1.000833883$ independent of the simulation module.

Example 2: Double inverted pendulum – multi-parameter estimation

This example illustrates the multi-parameter estimation for a local model of the double inverted pendulum. The state-space equation of motion for the local dynamics is given by
\[
\begin{bmatrix}
\dot{x}_1 \\
\dot{x}_2 \\
\dot{x}_3 \\
\dot{x}_4 
\end{bmatrix}
= \begin{bmatrix}
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1 \\
(-3\overline{k} + \overline{Q})/2 & \overline{k} - \overline{Q}/2 & 0 & 0 \\
(-5\overline{k} - \overline{Q})/2 & -2\overline{k} + (1.5 - \gamma)\overline{Q} & 0 & 0
\end{bmatrix}
\begin{bmatrix}
x_1 \\
x_2 \\
x_3 \\
x_4 
\end{bmatrix}
+ \begin{bmatrix}
f_1 \\
f_2 
\end{bmatrix}
\]

with

\[
f_1 = 0.5(x_1 - x_2)(x_3^2 + x_4^2) - (1/12)\overline{Q}[(x_1 - \gamma x_2)^3 - (1 - \gamma)^3 x_2^3]
+ 0.25(x_1 - x_2)^2[(4\overline{k} - \overline{Q})x_1 + (-3\overline{k} + (2 - \gamma)\overline{Q})x_2] \\
f_2 = -0.5(x_1 - x_2)(3x_3^2 + x_4^2) - (1/12)\overline{Q}[(x_1 - \gamma x_2)^3 - 3(1 - \gamma)^3 x_2^3]
+ 0.25(x_1 - x_2)^2[(7\overline{k} - 2\overline{Q})x_1 + (5\overline{k} + (\gamma - 3)\overline{Q})x_2]
\]

where \( m \) is an inertial parameter, \( \overline{k} = k / ml^2 \) is the spring stiffness, \( \overline{Q} = [Q_1 + Q_2 \cos(\sqrt{11}t)] / ml \) is a follower force with irrational frequency, \( x_1 \) and \( x_2 \) are the absolute angular joint displacements, \( x_3 \) and \( x_4 \) are absolute joint velocities. The stiffness parameter \( \overline{k} = k / ml^2 \) and constant follower force \( Q_2 / ml \) are the unknown parameters that are to be estimated by the estimation algorithm.

A. Simulation: Figure 2 shows the time trajectories for the state of equation (15) by MATLAB routine ode45 and collocation-based optimization. The parameter set is chosen as \( m = 1.0, Q_2 = 0.6, \gamma = 1 \). The value of the stiffness parameter is chosen as \( \overline{k} = 1 \) and the constant follower force is chosen as \( Q_2 / ml = 0.1 \). The initial condition is chosen as \( x(0) = [1.02 1.02 0 0]^T \), the number of Chebyshev points \( N = 30 \).

B. Estimation: The number of estimation intervals \( n_e = 5 \) and the estimation horizon \( \tau = 3.0982 \). There was no bias while choosing the parameter values. From the data generated in figure 2, the residual cost as in (9) is set up for system (15) with \( \overline{k} \) and \( \overline{Q}_2 = Q_2 / ml \) as unknown parameters. The estimation algorithm computes \( \overline{k} = 1.000040518, \overline{Q}_2 = 0.600225099 \).

Example 3: Coupled pendulum – parameter estimation and sensitivity

In this section, the effect of parametric sensitivity on parameter estimation is examined. A coupled pendulum example is considered here which points the pitfall in the estimation algorithm due to numerical reasons. The state-space equation of motion of two coupled pendulums with follower forces on both is given by
\[
\begin{bmatrix}
\dot{x}_1 \\
\dot{x}_2 \\
\dot{x}_3 \\
\dot{x}_4 \\
\end{bmatrix} =
\begin{bmatrix}
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1 \\
-(\omega_1^2 - q_1 \cos(\sqrt{2}t)) & b & 0 & 0 \\
b & -(\omega_2^2 - q_2 \cos(\sqrt{2}t)) & 0 & 0 \\
\end{bmatrix}
\begin{bmatrix}
x_1 \\
x_2 \\
x_3 \\
x_4 \\
\end{bmatrix} +
\begin{bmatrix}
0 \\
f_1 \\
f_2 \\
\end{bmatrix}
\]
\[(17)\]

with

\[
\begin{align*}
f_1 &= -q_1 \cos(\sqrt{2}t)x_1^3 / 6 + c(x_1 - x_2)^2 + d(x_1 - x_2)^3 \\
f_2 &= -q_2 \cos(\sqrt{2}t)x_2^3 / 6 + c(x_1 - x_2)^2 + d(x_1 - x_2)^3 \\
\end{align*}
\]
\[(18)\]

Case I: \(c\) and \(b\) as unknown parameters

A. Simulation: System model (14) is simulated with the parameter set \(\omega_1^2 = 2.1, \omega_2^2 = 2.5, c = 1.2, d = 1.03, q_1 = 1.3, q_2 = 1.2, b = 1.0\) and initial condition \(x(0) = [0.31 \ 0.31 \ 0 \ 0]^T\) to generate the data. The number of Chebyshev points \(N = 24\).

B. Estimation: The estimation algorithm is applied to the data generated by the simulation module with \(c\) and \(b\) as unknown parameters. The cost minimization yields the solution \(c = 1.24788488, b = 1.00388293\).

Case II: \(c\) and \(d\) as unknown parameters

A. Simulation: System model (15) is simulated with the parameter set \(\omega_1^2 = 2.1, \omega_2^2 = 2.5, c = 1.2, d = 1.03, q_1 = 1.3, q_2 = 1.2, b = 1.0\) and initial condition \(x(0) = [0.31 \ 0.31 \ 0 \ 0]^T\) to generate the data. The number of Chebyshev points \(N = 24\).

B. Estimation: The estimation algorithm is applied to the data generated by the simulation module with \(c\) and \(b\) as unknown parameters. The cost minimization yields the solution \(c = 1.24488754, d = 1.00386268\), thus diverging slightly from the actual value of \(d\).

It is evident from the results that parameter sensitivity, i.e., the dependence of the system solution on the parameter has a profound effect on the success of the estimation algorithm. The reason is that the master cost is minimized only up to a point where it is smaller than a predetermined tolerance which is a small number and not exactly equal to zero. This is because of the numerical nature of the process of quadratic optimization. The unknown parameters in case II are \(c\) and \(d\), out of which \(d\) is a coefficient of the non-dominant cubic nonlinearity. The cubic nonlinearity associated with \(d\) is having a non-dominant effect on the solution of the system (17). The dominant candidate here is quadratic nonlinearity associated with \(c\) and hence, that is the parameter which is determined accurately. In case I, both \(c\) and \(b\) are dominant parameters and hence, both are computed accurately. Also, the smallness of parameter has an adverse effect on the
system identification unless the system is at the critical bifurcation point. These effects are not included in the present study.

Example 4: Double inverted pendulum with partial data

Consider a double inverted pendulum equations (15) and (16) as in example 2 with only \([x_1, x_2]^T\) available as data. The stiffness parameter \(k = k/\text{ml}^2\) and the states \([x_3, x_4]^T\) are the unknowns that are to be estimated by the estimation algorithm. The estimation algorithm is run for one interval at a time.

A. Simulation: Figure 3 shows the time trajectories for the state of equation (15) by MATLAB routine ode45 and collocation-based optimization. The parameter set is chosen as \(m = 1.0, Q_2 = 0.6, \gamma = 1\). The value of the stiffness parameter is chosen as \(k = 1\) and the constant follower force is chosen as \(Q_1/\text{ml} = 0.1\). The initial condition is chosen as \(x(0) = [1.02, 1.02, 0, 0]^T\), the number of Chebyshev points \(N = 30\).

B. Estimation: The number of estimation intervals \(n_e = 5\) and the estimation horizon \(\tau = 3.0982\). There was no bias while choosing the parameter values. From the data generated in figure 2, the residual cost as in (9) is set up for system (15) with \(\bar{k}\) and \([x_3, x_4]^T\) as unknown parameters. The estimation algorithm is not a QP anymore, it is in fact an NLP that has to be run as many times as the number of intervals, in this case, 5. The algorithm computes

\[
\bar{k} = 1.00015758, 0.99993321, 1.00058832, 0.99987161, 1.00140496
\]  

over these 5 intervals. Figure 3 which is identical to figure 2 is also included because it shows the agreement of estimated states \([x_3, x_4]^T\) with the actual ones obtained by numerical simulation in A.

5. Conclusions

A methodology for estimating unknown parameters in nominal nonlinear time-varying system models from the available data which can be complete or incomplete in terms of the state vector is proposed and illustrated. The assumptions made are not stringent and the estimation algorithm suggested is efficiently implemented. The algorithm is based on Chebyshev spectral collocation and quadratic optimization. The pitfalls of the algorithm are also pointed out via a suitable example. Such an algorithm paves the way for a unified approach to system identification and control. If the problem at hand is control of linear or nonlinear time-varying systems, the estimation algorithm is used on the closed loop system with the controller gains as unknown parameters. The other possible extensions are in the estimation and control of delay differential equation models that are nonlinear and time-varying and partial differential equation models for data fields. Differential algebraic constraints, robust estimation and control, structural
systems with non-smooth nonlinearities are also important from the completeness perspective. In general, it is envisaged to use the algorithm for engineering, biological, epidemiological and financial systems for the purposes of parameter estimation.

6. Acknowledgments

The financial support provided by NSWCCD contract no. N65540-05-C-0028 through CENDAC is gratefully acknowledged

7. Reference


[Mendes3] www.vbi.vt.edu/mendes
Figure 1. The comparison of state trajectories for system (12), ___ ode45, ..... collocation-based optimization
Figure 2. The comparison of state trajectories for system (15), ___ ode45, ..... collocation-based optimization
Figure 3. The comparison of state trajectories for system (15), ode45, collocation-based optimization for incomplete data (state vector) availability