A parameter estimation method for a special class of systems of ordinary differential equations

Carla Seatzu*
Department of Electrical and Electronic Engineering
Universiy of Cagliari
Piazza d'Armi — 09123 Cagliari, Italy

Abstract

In this paper a special class of systems of ordinary differential equations is considered. This class is particularly common both in biological and medical field and is denoted as *S-Systems*.

The problem we deal with is the estimation of a set of unknown parameters given a set of observational data. The procedure we now propose arises from the requirement to overcome the main difficulties typical of the iterative gradient based methods. The main idea of the method is that of approximating each state variable by a fitting process and then splitting the overall estimation problem into a set of simpler independent problems, thus lessening the difficulty concerning great parameter vector dimensions. Each subproblem reduces to the minimization of a differential residual and in particular cases it simply requires the solution of an overdetermined algebraic linear system.

1 Introduction

Various differential equation models have been proposed in the literature to represent population growth in biology, with different degrees of complexity and sophistication, but a canonical approach to complex system modeling, namely the S-System approach, has recently gained great interest and looks particularly attractive (Voit, 1991). The approach is based on a formalism for the analysis of complex biological systems called the *Power Law Formalism*. Such a formalism, resulted from a combination of ideas stemming from biochemistry and network theory, proved to be particularly suitable for expressing synergetic and saturation phenomena in biological systems, from which the term S-System derives. It assumes the following regularly structured system of nonlinear ordinary differential equations as modeled in the *S-System canonical form*:

$$\dot{x}_i(t) = \alpha_i \prod_{k=1}^{n} x_k(t)^{g_{ik}} - \beta_i \prod_{k=1}^{n} x_k(t)^{h_{ik}}, \quad i = 1, \cdots, n, \quad t \in [t_0, t_f]. \tag{1}$$

It represents the dynamics of a system in terms of the time derivatives of the components $x_i > 0$ relevant to the system behaviour, which are expressed as the difference between all the influences that increase x_i and all the influences that decrease x_i. More specifically, referring to dynamic growth models, the two right-hand side terms of (1) represent the metabolic and catabolic terms respectively.

*Email: seatzu@diee.unica.it.

The S-System canonical form (1) is only one of the many possible realization within the power-law formalism. A particularly suited canonical form for dynamic growth models, which presents monotonicity in rates (Torsella *et al.*, 1991), is the *Half-System canonical form*:

$$\dot{x}_i(t) = \gamma_i \prod_{k=1}^{n} x_k(t)^{l_{ik}}, \quad i = 1, \cdots, n, \quad t \in [t_0, t_f] \tag{2}$$

where $x_i > 0$ and γ_i and l_{ik} are real constants $\forall i, k$.

All different types of canonical forms can be reduced to each other by means of proper recasting techniques (Voit, 1991).

The constant parameters α_i, β_i, γ_i, g_{ik}, h_{ik}, l_{ik}, concisely introduced above, are called *state parameters* and their estimation is crucial for guaranteeing reliable quantitative representation of the dynamic behaviour of the real systems modeled, when resorting to computer simulation.

In this paper we deal with the problem of estimating these parameters when a set of observational data on state variables is available.

The most popular estimation methods consist in the minimization of an integral residual with respect to the unknown parameters. The integral residual is defined as the sum of the square differences between the solution of the system of ordinary differential equations (ODE's) at a discrete set of points and the sampling data. Such a minimization is generally performed recurring to iterative gradient based algorithms (Bard, 1974; Corriga *et al.*, 1997; Varah, 1982; Voit, 1996). Even if in some cases they produce satisfactory results, all these methods present the following drawbacks: the convergence is local and a criterion of effective choice of the initial parameter vector does not exist; when the parameters are constrained, the computational complexity of the iterative methods excessively increases; they do not allow to simplify the problem in case of linearity with respect to parameters; since the computational complexity of such methods excessively increases with the number of unknown parameters, they can only be useful when solving problems with small dimensions.

The method we now propose aims to overcome the difficulties mentioned above. It is inspired by the identification procedure proposed by Varah in (Varah, 1982). It mainly consists of two phases. In the first one observational data are fitted in a finite-dimensional space which provides both an evaluation of the initial state and a data noise filtering. Since fitting functions are known analytically, they can be differentiated and evaluated, together with their derivatives, in an arbitrary set of time instants in $[t_0, t_f]$. At this point, the whole estimation problem can be decoupled into n independent estimation problems: one for each state equation. The i-th unknown vector of parameters can then be evaluated by minimizing a differential residual defined as the sum of the square differences between the two terms of the i-th state equation when state variables are replaced by fitting curves evaluated in an arbitrarily chosen set of time instants.

The main advantages of the proposed estimation method can be summarized as follows: it enables us to split the overall estimation problem into a set of simpler independent problems with reduced parameter vector dimensions, thus lessening the difficulty of large scale problems; it is also possible to take into account the parameter constraints with a modest computational effort; whenever the model can be linearized with respect to the unknown parameters, as always happens in the Half-System case, the estimation problem reduces to the solution of an overdetermined algebraic linear system.

In order to test the performance of the proposed method with respect to the prediction reliability, several simulations have been carried out both on Half-Systems and S-Systems.

B-splines have been considered for fitting.

The robustness of the numerical procedure has also been tested affecting data with different noise levels.

The paper is structured as follows. In Section 2 the fitting based method is described in detail. The results of numerical simulations are presented in Section 3. In the Appendix a brief review of B-splines, used as basis functions for data fitting, is firstly given. Finally, the Nelder-Mead simplex algorithm, used for multivariate minimization, is summarized.

2 The fitting based method

Let $\boldsymbol{p}_i = (\alpha_i, g_{i1}, \cdots, g_{in}, \beta_i, h_{i1}, \cdots, h_{in})^T$, $i = 1, \cdots, n$, be the vector of unknown parameters relative to the i-th ordinary differential equation of system (1). Furthermore, let $\boldsymbol{y}_i \in I\!R^N$, $i = 1, \cdots, n$, be a data vector corresponding to the i-th state variable x_i. The problem is to determine the parameter matrix

$$\boldsymbol{P} = [\boldsymbol{p}_1, \boldsymbol{p}_2, \cdots, \boldsymbol{p}_n] = \begin{bmatrix} \alpha_1 & \cdots & \alpha_i & \cdots & \alpha_n \\ g_{11} & \cdots & g_{i1} & \cdots & g_{n1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ g_{1n} & \cdots & g_{in} & \cdots & g_{nn} \\ \beta_1 & \cdots & \beta_i & \cdots & \beta_n \\ h_{11} & \cdots & h_{i1} & \cdots & h_{n1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ h_{1n} & \cdots & h_{in} & \cdots & h_{nn} \end{bmatrix},$$

such that the associated solution $\{x_1(t, \boldsymbol{P}), \cdots, x_n(t, \boldsymbol{P})\}$ of system (1) is in good agreement with data matrix $\boldsymbol{Y} = (\boldsymbol{y}_1, \cdots, \boldsymbol{y}_n)$.

The estimation parameter method we propose consists in the following two phases:

1. Choose a finite-dimensional space $\Phi = span\{\varphi_0, \varphi_1, \cdots, \varphi_m\}$, which approximates each state variable. Then, using data matrix $\boldsymbol{Y}$, a fitting for each state function is performed in Φ providing both an evaluation of the initial state and a data noise filtering. It associates an analytical function ψ_i to each data vector $\boldsymbol{y}_i$, $i = 1, \cdots, n$.

2. Take a set of $M > 2(n+1)$ values $t_0 \leq \bar{t}_1 < \bar{t}_2 < \cdots \bar{t}_M \leq t_f$ and minimize in $\mathbb{R}^{2(n+1)}$ the n following differential residuals:

$$R_D(\boldsymbol{p}_i) = \sum_{r=1}^{M} [\dot{\psi}_i(\bar{t}_r) - f_i(\bar{t}_r, \psi_i(\bar{t}_r, \boldsymbol{p}_i))]^2, \quad i = 1, \cdots, n, \tag{3}$$

   where $f_i(\cdot)$ is the right hand side of the i-th equation of system (1) in which $x_i(t)$ and $\dot{x}_i(t)$ have been replaced by $\psi_i(t)$ and $\dot{\psi}_i(t)$ respectively, both evaluated at $\bar{t}_r$.

As will be shown in detail when dealing with the numerical simulations, B-splines (de Boor, 1978) provide a valid tool for state variable fitting. Note that other basis functions, in particular Legendre polynomials (Isaacson, 1966) and Wavelets (Chui, 1992), can be effective as well.

We remark that state variables approximation enables us to realize a decoupling of the whole system into a set of n uncorrelated ODE's. As a result, instead of estimating a matrix in $\mathbb{R}^{2(n+1)\times n}$, we have to estimate n vectors in $\mathbb{R}^{2(n+1)}$ thus lessening the difficulty of large scale problems.

Furthermore, this method is particularly effective whenever the model, even if nonlinear with respect to the state variables, can be linearized with respect to the unknown parameters, as quite always happens in the Half-Systems. More specifically, let us suppose that state variables in system (2) are monotone (a common property of many biological state functions) and so, without loosing generality, we can assume that all γ_i are strictly positive, which implies that linearization with respect to the unknown parameters can be simply obtained by taking the logarithms of both the left and right end sides of the system equations. Substituting in system (2) the approximating functions $\psi_i(t)$ obtained via fitting, and resorting to logarithms, we obtain the following algebraic linear system:

$$\Delta_i(t) = \Gamma_i + \sum_{k=1}^{n} X_k(t) l_{ik} \qquad i = 1, \cdots, n$$

where

$$\Delta_i(t) = ln(\dot{\psi}_i(t)), \qquad \Gamma_i = ln(\gamma_i), \qquad X_k(t) = ln(\psi_k(t)).$$

The unknowns are $\{\Gamma_i\}$ and $\{l_{ik}\}$. In such a case, minimizing $R_D(\boldsymbol{p}_i)$, for each $i = 1, \cdots, n$, is equivalent to solve, in the least squares sense, the n overdetermined systems

$$\Delta_i(\bar{t}_r) = \Gamma_i + \sum_{k=1}^{n} X_k(\bar{t}_r) l_{ik}, \qquad r = 1, \cdots, M > n+1, \tag{4}$$

for $i = 1, \cdots, n$.

Our numerical experiments suggest that a good choice for M is $M = 2(n+1)$. The solution of system (4) by a QR technique enables us to compute Γ_i (therefore $\gamma_i = e^{\Gamma_i}$) and l_{ik} for $k, i = 1, \cdots, n$.

Some a priori information on parameters, as the non-negativity or the variation between their lower and upper bounds, can also be easily taken into account. The minimization of each one of the n differential residuals $R_D(\boldsymbol{p}_i)$ can be performed by a standard method of quadratic programming. In our numerical simulations, the routine `qp` of the *Minimization Toolbox* of MATLAB (Matlab, 1997) has been used.

3 Numerical results

The effectiveness of the described procedure has been widely checked by means of a lot of simulations concerning several S-Systems. In this paper we specifically refer to two different test cases: an Half-System and an S-System.

To assess the numerical accuracy of the results, we chose a parameter matrix and an initial state vector $\boldsymbol{x}(t_0)$ for each model. Then state variables evolution is obtained by solving the ODE's system under consideration by means of the routine `ode45` of MATLAB (Matlab, 1997). In order to estimate the robustness of the method we considered both noiseless and noisy state data. In all the cases examined, we only considered white noise, so that for each state variable, the errors $\{\varepsilon_i\}$ on data are assumed to be uncorrelated random values, each one with zero mean value and a common variance σ^2.

The numerical reliability of the method, with respect to state variables $\bar{\boldsymbol{x}}_i, \;\; i = 1, \cdots, n$, has been estimated in terms of the parameter

$$\varrho = \frac{\sigma \bar{N}^{1/2}}{\sum_{i=1}^{n} \| \bar{\boldsymbol{x}}_i \|_2}$$

where $\bar{\boldsymbol{x}}_i$ is a vector which gives the values of the state variable $x_i(t)$ at $\bar{N}$ distinct time instants and σ is the associated standard deviation. In our experiments, chosen $\bar{N}$ and σ, the errors $\{\varepsilon_i\}$ have been generated by means of the *Simulink Toolbox* of MATLAB (Matlab, 1997).

Now, let us denote as $\boldsymbol{P}$ the parameter matrix to be identified and $\hat{\boldsymbol{P}}$ the recovered matrix. Furthermore, let $\boldsymbol{x}(t, \boldsymbol{P})$ and $\hat{\boldsymbol{x}}(t, \hat{\boldsymbol{P}})$ be the corresponding state variable vectors. Note that $\hat{\boldsymbol{x}}(t, \hat{\boldsymbol{P}})$ has been generated assuming as initial state the value of the fitting curves at $t = t_0$, i.e., $\hat{x}_i(t_0) = \psi_i(t_0)$, $i = 1, \cdots, n$. Even if both $\boldsymbol{x}(t, \boldsymbol{P})$ and $\hat{\boldsymbol{x}}(t, \hat{\boldsymbol{P}})$ depend from $\boldsymbol{x}(t_0)$ and $\hat{\boldsymbol{x}}(t_0)$ respectively, for sake of simplicity, we do not make explicit this fact.

A significant measure of the identification error, i.e., of the distance between the true matrix $\boldsymbol{P}$ and the recovered one $\hat{\boldsymbol{P}}$, is given by the function

$$E(\boldsymbol{P}) = \frac{\| \boldsymbol{P} - \hat{\boldsymbol{P}}||_F}{||\boldsymbol{P} \|_F}$$

where the subscript F denotes the Frobenius norm [1].

A special point is to be stressed: in some models, even if the recovered matrix $\hat{\boldsymbol{P}}$ is far from the true matrix $\boldsymbol{P}$, state variables $x_i(t, \boldsymbol{P})$ and $\hat{x}_i(t, \hat{\boldsymbol{P}})$, for $i = 1, \cdots, n$, are approximately the same in the time interval $[t_0, t_f]$. This is not a drawback of the proposed estimation procedure. It simply arises from the fact that the solution is not unique in general, but specifically depends on the optimality criteria adopted. A unique solution can only be obtained by means of specific analysis on physical constraints and/or the influence of each parameter on the system behaviour (Corriga *et al.*, 1997; Torsella *et al.*, 1991; Voit, 1988). As a result, we cannot limit ourselves to compare $\boldsymbol{P}$ and $\hat{\boldsymbol{P}}$. We note that in real applications we are interested in obtaining a parameter matrix $\hat{\boldsymbol{P}}$ such that the corresponding state variables are in good agreement with the real state variables in a wider time interval than $[t_0, t_f]$. For this reason, given a model and a data time interval, we assume

$$E_P(\boldsymbol{x}) = \sum_{i=1}^{n} \frac{||x_i(\boldsymbol{t}, \boldsymbol{P}) - \hat{x}_i(\boldsymbol{t}, \hat{\boldsymbol{P}})||_2}{||x_i(\boldsymbol{t}, \boldsymbol{P})||_2}$$

as a measure of the effectiveness of the identification process, where $\boldsymbol{t} = (t_1, t_2, \cdots, t_{100})$ is a vector of 100 equally spaced points in the extended time interval $[t_0, \frac{3}{2} t_f]$.

3.1 First model

In this subsection, we consider the following Half-System already treated in (Sanna *et al.*, 1998, Seatzu, to appear, Torsella *et al.*, 1991):

$$\begin{cases} \dot{x}_1(t) = \gamma_1 x_1(t)^{l_{11}} x_2(t)^{l_{12}} x_3(t)^{l_{13}} \\ \dot{x}_2(t) = \gamma_2 x_1(t)^{l_{21}} x_2(t)^{l_{22}} x_3(t)^{l_{23}} \qquad\qquad t \in [t_0, t_f]. \\ \dot{x}_3(t) = \gamma_3 x_1(t)^{l_{31}} x_2(t)^{l_{32}} x_3(t)^{l_{33}} \end{cases}$$

Such a system is particularly suited to describe dynamic growth models which present monotonicity in rates. As a consequence, we assume that parameters γ_1, γ_2 and γ_3 are positive.

We further assume that observational data are generated by matrix

$$\boldsymbol{P} = (\boldsymbol{p}_1, \boldsymbol{p}_2, \boldsymbol{p}_3) = \begin{bmatrix} 10 & 10 & 10 \\ 1 & -1 & -0.75 \\ -1 & 1 & -0.5 \\ -1 & -1 & 1 \end{bmatrix},$$

[1] Given a matrix $\boldsymbol{A} = A_{ij}$, $i = 1, \cdots, m$, $j = 1, \cdots, n$, $||\boldsymbol{A}||_F = (\sum_{i=1}^{n} \sum_{j=1}^{n} |a_{ij}|^2)^{1/2}$.

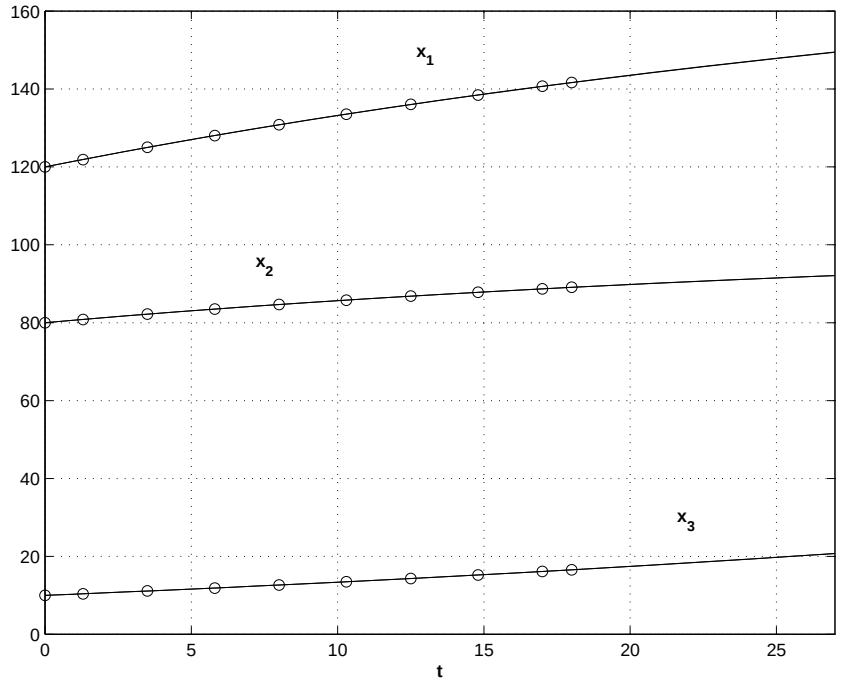

Figure 1: *Half-System: comparison between* $x_i(t, \boldsymbol{P})$ *and* $\hat{x}_i(t, \hat{\boldsymbol{P}})$, $i = 1, 2, 3$, *corresponding to the* 1*-st row data of Table 1.*

which is therefore the parameter matrix we want to recover.

We consider both noiseless and noisy data. In particular we introduce two different noise levels.

Furthermore, we assume that state variables are available at 10 equally spaced points in the time interval $[0, 18]$. In fact, by several numerical simulations it has been observed that the accuracy of the results does not increase significantly when the observational data exceed in number such a value.

Thanks to the regularity of the actual growth curves, the choice of cubic B-splines of degree $m = 3$ resulted to be satisfactory.

The main numerical results are summarized in Table 1 where the a priori information on parameters, the noise levels and the values of the quality indexes $E_P(\boldsymbol{x})$ and $E(\boldsymbol{P})$ are reported.

Furthermore, as in (Sanna *et al.*, 1998; Seatzu, to appear) we assume known the sign of $\{l_{ij}\}$. Numerical simulations have been performed even in absence of such an hypothesis. Results show that even if a perfect state reconstruction is obtained in a time interval wider than that of observational data, the estimated parameters are quite different from the exact ones.

To illustrate the results of the estimation procedure we compare $\boldsymbol{x}(t, \boldsymbol{P})$ with $\hat{\boldsymbol{x}}(t, \hat{\boldsymbol{P}})$. Figure 1 shows the results corresponding to the first row in Table 1, i.e. when the only a priori information on parameters concern the sign of l_{ij}, $i, j = 1, 2, 3$, and sampled data are noiseless. Circles are representative of the available data. As it can be seen, the two set of curves are practically coincident in the whole time interval $[t_0, \frac{3}{2}t_f]$.

Finally, the accuracy of the results relative to the case reported in the last row of Table 1 can be evaluated by looking at Figure 2. In such a case, accordingly with the value of $E_P(\boldsymbol{x})$ in Table 1, the two sets of curves are not perfectly coincident outside the time interval of observational data. However, differences are quite negligible.

A priori information	ρ	$E_P(\boldsymbol{x})$	$E(\boldsymbol{P})$
$sign(l_{ij}) \quad i,j=1,2,3$	0	$0.21 \cdot 10^{-5}$	1.12
$sign(l_{ij}) \quad i,j=1,2,3$	10^{-2}	$0.36 \cdot 10^{-2}$	0.89
$sign(l_{ij}) \quad i,j=1,2,3$	$5 \cdot 10^{-2}$	$0.71 \cdot 10^{-2}$	0.89
$sign(l_{ij}) \quad i,j=1,2,3$ $8 \leq \gamma_i \leq 12, \quad j=1,2,3$	0	$8.67 \cdot 10^{-5}$	0.20
$sign(l_{ij}) \quad i,j=1,2,3$ $8 \leq \gamma_i \leq 12, \quad j=1,2,3$	10^{-2}	$0.34 \cdot 10^{-2}$	0.23
$sign(l_{ij}) \quad i,j=1,2,3$ $8 \leq \gamma_i \leq 12, \quad j=1,2,3$	$5 \cdot 10^{-2}$	$0.61 \cdot 10^{-2}$	0.24
$sign(l_{ij}) \quad i,j=1,2,3$ $\gamma_i = 10, \quad j=1,2,3$	0	0	0
$sign(l_{ij}) \quad i,j=1,2,3$ $\gamma_i = 10, \quad j=1,2,3$	10^{-2}	$0.30 \cdot 10^{-2}$	0.10
$sign(l_{ij}) \quad i,j=1,2,3$ $\gamma_i = 10, \quad j=1,2,3$	$5 \cdot 10^{-2}$	$0.61 \cdot 10^{-2}$	0.12

Table 1: *Half-System: fitting based method with B-Splines of degree* $m = 3$.

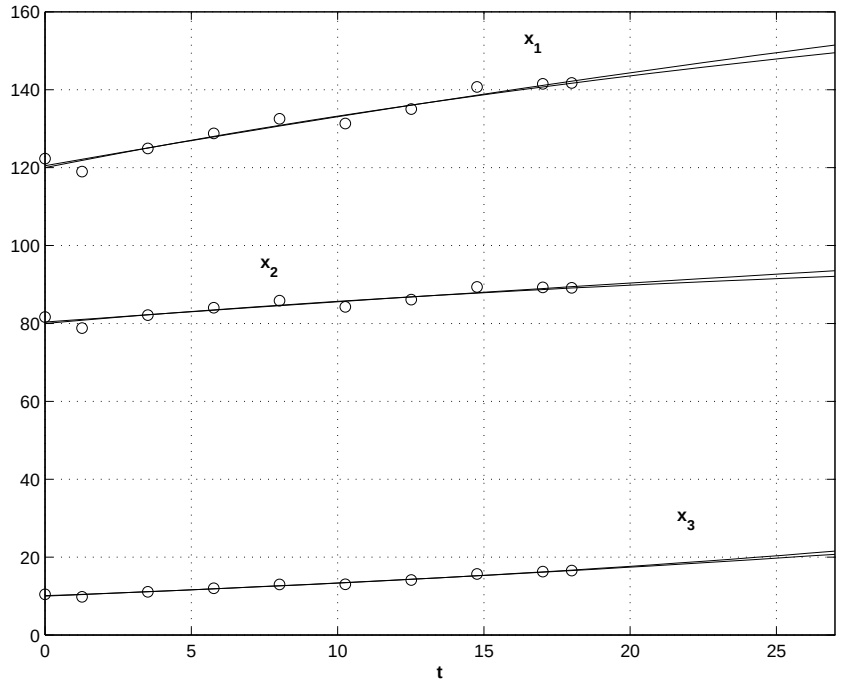

Figure 2: *Half-System: comparison between* $x_i(t, \boldsymbol{P})$ *and* $\hat{x}_i(t, \hat{\boldsymbol{P}})$, $i = 1,2,3$, *corresponding to the last row data of Table 1.*

3.2 Second model

Now, let us consider the following S-System (Seatzu, to appear):

$$\begin{cases} \dot{x}_1(t) = \alpha_1 x_1(t)^{g_{11}} x_2(t)^{g_{12}} x_3(t)^{g_{13}} x_4(t)^{g_{14}} x_5^{g_{15}} - \beta_1 x_1(t)^{h_{11}} x_2(t)^{h_{12}} x_3(t)^{h_{13}} x_4(t)^{h_{14}} x_5^{h_{15}} \\ \dot{x}_2(t) = \alpha_2 x_1(t)^{g_{21}} x_2(t)^{g_{22}} x_3(t)^{g_{23}} x_4(t)^{g_{24}} x_5^{g_{25}} - \beta_2 x_1(t)^{h_{21}} x_2(t)^{h_{22}} x_3(t)^{h_{23}} x_4(t)^{h_{24}} x_5^{h_{25}} \\ \dot{x}_3(t) = \alpha_3 x_1(t)^{g_{31}} x_2(t)^{g_{32}} x_3(t)^{g_{33}} x_4(t)^{g_{34}} x_5^{g_{35}} - \beta_3 x_1(t)^{h_{31}} x_2(t)^{h_{32}} x_3(t)^{h_{33}} x_4(t)^{h_{34}} x_5^{h_{35}} \\ \dot{x}_4(t) = \alpha_4 x_1(t)^{g_{41}} x_2(t)^{g_{42}} x_3(t)^{g_{43}} x_4(t)^{g_{44}} x_5^{g_{45}} - \beta_4 x_1(t)^{h_{41}} x_2(t)^{h_{42}} x_3(t)^{h_{43}} x_4(t)^{h_{44}} x_5^{h_{45}} \\ \dot{x}_5(t) = 0 \end{cases}$$

$$t \in [0, 14].$$

The parameter values to be recovered have been taken from (Voit, 1996) and the matrix to be identified is the following:

$$\boldsymbol{P}' = \begin{bmatrix} 9.5 \cdot 10^{-5} & -1 & 0.72 & 0 & 0.72 & 1 & 6 \cdot 10^{-5} & 0 & 1 & 1 & 1 & 0 \\ 3.2 & 0.5 & 0.57 & 0 & 0 & 0 & 0.2 & 0 & 1 & 0 & 0 & 0 \\ 5 & 0.5 & 0.7 & 0 & 0 & 0 & 0.2 & 0 & 0 & 0.7 & 0 & 0 \\ 0.36 & -0.4 & 1 & 0 & 0 & 0 & 0.75 & -0.13 & 0 & 0 & 0.8 & 0 \\ 0 & ? & ? & ? & ? & ? & 0 & ? & ? & ? & ? & ? \end{bmatrix}$$

where the numbers at the question marks are irrelevant.

In this subsection we only present the main results of a wider series of experiments carried out at different noise levels, different number of observational data and different a priori information on parameters. In particular Table 2 shows the results of the estimation when data are both noiseless and noisy. Both situations have been examined when considering $N = 10$ and $N = 25$.

In the case at hand $[t_0, t_f] = [0, 14]$ and data are equally spaced within it.

The regularity of the state variable curves justifies once again the choice of $m = 3$.

As it can be deduced from Table 2, the accuracy of the estimation does not increase significantly with the number of available data. It is worthwhile to note that in this case the method reveals less robustness with respect to data noise than in the Half-System case. This fact can be justified by considering that now linearization with respect to unknown parameters has not been possible. In addition, a greater number of unknown parameters has to be identified.

As in the previous case, to illustrate the results of the estimation procedure we compare the state variable behavior corresponding to the true matrix $\boldsymbol{P}$ and the initial state vector $\boldsymbol{x}(t_0)$ with that obtained by means of the estimated matrix $\hat{\boldsymbol{P}}$ and the initial state vector $\hat{\boldsymbol{x}}(t_0) = \boldsymbol{\psi}(t_0)$. Figure 3 shows the results relative to the second row of Table 2, i.e., when no a priori information on parameters is available and data are affected by noise characterized by $\rho = 10^{-3}$ and $N = 10$. In both cases, the two set of curves are not coincident neither in the time interval $[t_0, t_f]$. This is in accordance with the value of $E(\boldsymbol{x})$ reported in the last column of Table 2. However, differences are not significant when evaluated in percentage.

It is important to remark that in such a case more classical procedures, which do not enable us to divide the whole problem into n independent sub-problems, cannot be usefulness since the number of unknown parameters is excessive.

4 Conclusions

In this paper we dealt with a special class of ODE's systems particularly common in biological and medical field and denoted as *S-Systems*.

N	A priori information	ρ	$E_P(\boldsymbol{x})$	$E(\boldsymbol{P})$
10	–	0	$2.54 \cdot 10^{-2}$	0.89
10	–	10^{-3}	$8.04 \cdot 10^{-2}$	1.01
25	–	0	$2.79 \cdot 10^{-2}$	0.83
25	–	10^{-3}	$5.41 \cdot 10^{-2}$	0.95
10	$\alpha_i, \beta_i, \ \ i = 1, \cdots, 5$	0	$4.86 \cdot 10^{-2}$	0.19
10	$\alpha_i, \beta_i, \ \ i = 1, \cdots, 5$	10^{-3}	$5.84 \cdot 10^{-2}$	1.05
25	$\alpha_i, \beta_i, \ \ i = 1, \cdots, 5$	0	$3.94 \cdot 10^{-2}$	0.12
25	$\alpha_i, \beta_i, \ \ i = 1, \cdots, 5$	10^{-3}	$4.92 \cdot 10^{-2}$	0.97

Table 2: *S-system: fitting based method with B-Splines of degree* $m = 3$.

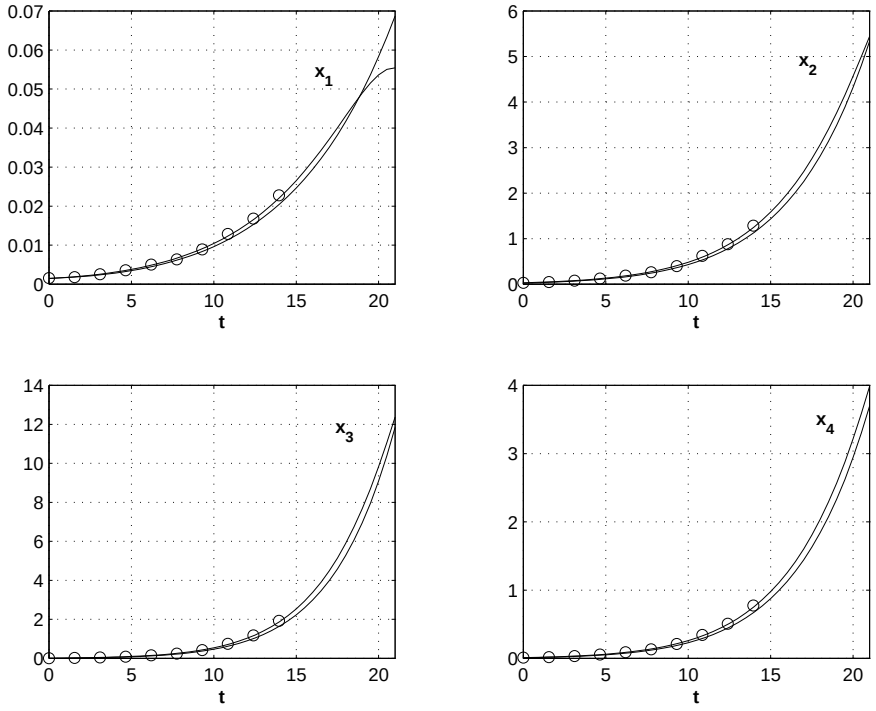

Figure 3: *S-System: comparison between* $x_i(t, \boldsymbol{P})$ *and* $\hat{x}_i(t, \hat{\boldsymbol{P}})$, $i = 1, 2, 3$, *corresponding to the* 2*-nd row data of Table 2.*

A method for the estimation of the unknown parameters in such a class of dynamical systems, given a set of observational data, has been proposed. Its effectiveness has been widely checked. The main idea of the method is that of fitting data in a finite dimensional space so that data knowledge is transformed into the knowledge of analytical functions. In such a way, the whole estimation problem can be decoupled into a set of simpler subproblems.

Each subproblem is independent from the others and can be solved by minimizing an appropriate differential residual. Many advantages, a part from the minor computational complexity, arises when dealing with large scale problems and when dealing with systems that can be linearized with respect to the unknown parameters.

The proposed method revealed to be effective in all the numerical examples considered.

Appendix

B-splines properties

Let Π_m be the space of polynomials of degree at most m and $\boldsymbol{\tau} \equiv \{\tau_i\}$ a nondecreasing sequence of real numbers. The i-th forward B-spline of degree m for the knot sequence $\boldsymbol{\tau}$, $B_{i,m}(t)$, is characterized by the following properties (de Boor, 1978):

- $B_{i,m}(t) = 0$ for $t \notin (\tau_i, \tau_{i+m+1})$ and $B_{i,m}(t) > 0$ for $t \in (\tau_i, \tau_{i+m+1})$;
- $B_{i,m}(t) \in \Pi_m$ for $t \in (\tau_j, \tau_{j+1})$, $\forall j$;
- $B_{i,m}(t) \in C^{m-1}(\mathbb{R})$.

Hence $B_{i,m}(t)$ is a piecewise polynomial of degree m whose support is (τ_i, τ_{i+m+1}) and whose first $m+1$ derivatives are continuous everywhere. As a result, each interval (τ_i, τ_{i+1}) belongs to the support of the B-splines $B_{i-m,m}(t), \cdots, B_{i,m}(t)$. Furthermore the B-splines, which are strictly positive on their support, can be computed by the following three-term recurrence relationship (de Boor, 1978):

$$\begin{aligned} B_{i,0}(t) &= \begin{cases} 1, & \tau_i \le t \le \tau_{i+1} \\ 0, & \text{elsewhere} \end{cases} \\ B_{i,k}(t) &= \frac{t-\tau_i}{\tau_{i+k+1}-\tau_i} B_{i,k-1}(t) + \frac{\tau_{i+k}-t}{\tau_{i+k}-\tau_{i+1}} B_{i+1,k-1}(t) \qquad k = 1, \cdots, m. \end{aligned} \tag{5}$$

This relationship is numerically stable, as it enables us to compute $B_{i,k}(t)$ by repeatedly forming linear combinations of positive quantities, starting with the definition of $B_{i,0}(t)$.

As usual we denote by $S_{\tau,m}$ the space of B-splines of degree m relative to the knot sequence $\boldsymbol{\tau} = \{\tau_i\}$. Note that if m is the degree of B-splines and $M \ge m+1$ is the number of knots of $\boldsymbol{\tau}$, the dimension of the linear space $S_{\tau,m}$ is $M^* = M - m - 1$. Hence each function $s \in S_{\tau,m}$ can be expressed as follows

$$s(t) = \sum_{j=1}^{M^*} d_j B_{j,m}(t).$$

Conversely, each function spanned as before belongs to $S_{\tau,m}$.

In our numerical simulations, assuming that $[t_0, t_f]$ is the common domain of the state variables and denoting by m the degree of B-splines, we choose knots as follows

$$\tau_i = t_0 + [i - (m+1)](t_f - t_0) \qquad i = 1, 2, \cdots, 2(m+1).$$

Furthermore, let $||\cdot||_2$ be the Euclidean norm in $\mathbb{R}^N$. The best approximation, in the least squares sense, of $\hat{\boldsymbol{f}} = (f_1, \cdots, f_N)^T$ in $S_{\tau,m}$ is

$$s^*(t) = \sum_{j=1}^{m+1} \alpha_j^* B_{j,m}(t)$$

where $\boldsymbol{\alpha}^* = (\alpha_1^*, \cdots, \alpha_{m+1}^*)^T$ is the minimizer in $\mathbb{R}^{m+1}$ of the quadratic function

$$Q(\boldsymbol{\alpha}) = ||\hat{\boldsymbol{f}} - \hat{\boldsymbol{s}}||_2^2 = \sum_{i=1}^{N} (f_i - s(t_i))^2 = ||\boldsymbol{B}\boldsymbol{\alpha} - \hat{\boldsymbol{f}}||_2^2$$

where $(\boldsymbol{B})_{ij} = B_{j,m}(t_i)$, $i = 1, \cdots, N$, $j = 1, 2, \cdots, m+1$ and $\hat{\boldsymbol{s}} = (s(t_1), \cdots, s(t_N))^T$.

Nelder-Mead algorithm

In numerical simulations we used the routine `fmins` of MATLAB (Matlab, 1997) to minimize the differential residual. It is an implementation of the Nelder-Meade simplex algorithm, described in detail in (Dennis *et al.*, 1987). It is a direct search method for finding a multivariate function's minimum. Let $f(x_1, x_2, \cdots, x_n)$ be the function to be minimized. Then, at each iteration, the new research direction is determined by evaluating f at $n+1$ appropriately selected points $\hat{\boldsymbol{x}}_1, \hat{\boldsymbol{x}}_2, \cdots, \hat{\boldsymbol{x}}_{n+1} \in \mathbb{R}^n$. These points are thought as vertices of an n-dimensional simplex. A trial point is accepted or rejected depending on the value of f at that point compared with the values of f at the vertices of the simplex. Fixed a positive value ε, iteration continues until the diameter of the simplex is $\leq \varepsilon$.

References

de Boor, C. (1978). "A Practical Guide to Spines," *Applied Mathematical Sciences*, **27**, Springer Verlag, New York.

Bard, Y. (1974). *Non Linear Parameter Estimation,* Academic Press, New York.

Chui, C. (1992). *An Introduction to Wavelets,* Academic Press, San Diego.

Corriga, G., S. Sanna, E. Usai and G. Usai (1997). "Some problems in the identification of a class of dynamical models for forest simulation," *Simulation Practice and Theory*, **5**, pp. 719–731.

Dennis, J.E. Jr. and D.J. Woods (1987). "New Computing environments: Microcomputers in Large-Scale computing," *SIAM*, Philadelphia, pp. 116–122.

Isaacson E. and H. Keller (1966). "Analysis of Numerical Methods," Wiley, New York.

Matlab Reference Guide, (1997). Version 5.0, The Math Works Inc..

Sanna, S., C. Seatzu and E. Usai (1998). "An effective procedure for growth models parameter estimation," *Proc. Comput. Eng. in Systems Applications '98*, Hammamet.

Seatzu C. (to appear). "A fitting based method for parameter estimation in S-Systems," *Dynamic Systems and Applications.*

Torsella, J.A. and A.B. Razali (1991). "An Analysis of Foresty Date," in *Canonical Nonlinear Modelling. S-System Approach to Understanding Complexity*, E.O. Voit, ed., Van Nostrand Reinhold, cap. 10, pp. 181–199.

Varah, J.M. (1982). "A spline least square squares method for numerical parameter estimation in differential equations," *SIAM J. Sci. Stat. Comput.*, **3**, pp. 28–46.

Voit, E.O. (1988). "Dynamics of Self-Thinning Plant Stands," *Annuals of Botany*, **62**, pp. 67–78.

Voit, E.O. (1988). *Canonical Nonlinear Modelling. S-System Approach to Understanding Complexity,* Van Nostrand Reinhold.

Voit, E.O. (1996). "Modelling forest growth. II. Biomass partitioning in scots pine," *Ecological Modelling*, **86**, pp. 73–89.