IDENTIFICATION, ESTIMATION, AND RESOLUTION OF MATHEMATICAL MODELS

J.M. van den Hof
Researcher at Statistics Netherlands, Voorburg, The Netherlands

J.H. van Schuppen
Senior researcher, Centre for Mathematics and Computer Science, Amsterdam, and the Department of Mathematics, Vrije Universiteit, Amsterdam, The Netherlands

Keywords: Dynamicsystem,identification,realization,stochasticrealization, entifiability, approximation, least-squares, maximum likelihood method, subspace identification algorithm.

Contents

Summary

Identification is the procedure to determine for an observed phenomenon a mathematical model in the form of a dynamic system with numerical values of its parameters. Examples of phenomena for which identification has been applied include: modeling of nitrate concentration in parts of the human body, modeling of glycolys in the parasite T. brucei, and modeling of freeway traffic flow. In this paper a five step procedure is formulated for identification consisting of: (1) Selection of a subclass of dynamic systems. (2) Data collection. (3) Identifiability. (4) Approximation. (5) Evaluation. The procedure is illustrated for the classes of dynamic systems: finite-dimensional linear systems and finite-dimensional Gaussian systems. Identification for several other classes of dynamic systems is discussed. Finally research problems of the area of identification are mentioned.

1. Introduction

Identification is the research topic of the construction of a dynamic system from an observed time series. The system is expected to represent the observations as well as possible according to an approximation criterion. In this paper a five step procedure for the identification problem is presented and illustrated. The character of this article is tutorial; it is an introduction for readers not familiar with the subject. It is not an exhaustive survey.

Examples of identification problems are the modeling of nitrate concentrations in parts of the human body, of traffic flow on a motorway, of a robot arm in space, etc. Mathematical models in the form of dynamic systems are needed for many purposes. Examples of such purposes are detection of events (faults in long gas lines), prediction (of electric power demand), and control, directly (control of an aircraft) and indirectly (setting standards for the uptake of potentially dangerous chemical substances).

Identification as a research topic has been developed in engineering (electrical, mechanical, civil), in statistics (where it is known as time series analysis), and in econometrics. Each branch of these sciences or technologies has its own models of dynamic systems, its analysis tools, and its books and conference series. Future problems of identification are likely to come from biology and medicine, from economics and management science, and from new areas of engineering.

Below a five step procedure for identification is formulated and illustrated for two classes of dynamic systems. The steps of the procedure are: (1) Selection of a subclass of dynamic systems. (2) Data collection. (3) Identifiability. (4) Approximation. (5) Evaluation. This way the reader can follow the steps to determine a dynamic system from a time series. Subsequently the identification problem for other classes of dynamic systems is briefly discussed. Not covered is the topic of the combination of identification and control.

Research problems which require attention are briefly mentioned.

2. Problem of Identification

The purpose of this section is to formulate and to discuss the problem of system identification.

Problems of identification arise in many areas of life support systems. Examples of such problems and areas with which the authors are partly familiar include: models for freeway traffic flow; concentrations of toxic chemical substances in physiological entities of the human body; models for the distribution of pollutants in the atmosphere; models for underground water flow; models for chemical substances in rivers and estuaries; and models for biochemical reaction networks in cells of humans, animals, or plants.

The purpose of a mathematical model is usually a combination of the following: understanding of the inherent dynamics of the phenomenon; estimation of the full state of the dynamic systems which is often not measured directly; prediction of the state or of observed variables; control of the state or of related variables, either directly or indirectly.

In the next section an approach to the identification problem is sketched which consists of the following five steps: (1) Selection of a subclass of dynamic systems. (2) Data collection. (3) Parameterization and check for identifiability. (4) Approximation. (5) Evaluation.

The basic problem of identification is to make a trade-off between the accuracy of the approximation versus the complexity of the class of dynamic systems. By enlarging the class of systems, say with a higher state space dimension and with more parameters, one can obtain a system which has a smaller distance to the time series according to the approximation criterion. But this goes at the cost of a higher complexity. Conversely, reducing the class of dynamic systems hence reducing complexity, one obtains a higher value of the approximation criterion. Making the trade-off is an iterative process.

Researchers interested in applying identification to a phenomenon may want to concentrate on the following aspects. The value of the criterion is minimized best by selection of a subclass of dynamic systems which adequately models the phenomenon underlying the data. This aspect is best dealt with by modeling using the laws of the domain of the phenomenon, for example, based on laws from physics, chemistry, biology, engineering, or economics. Such researchers can then request support from engineers and/or mathematicians for the Steps 3 and 4 of the above defined identification procedure.

3. Identification Procedure

The purpose of this section is to present a general procedure for identification. As example the phenomenon of modeling the nitrate flow in the human body will be discussed. As special cases two classes of dynamics systems will be used for illustration, that of finite-dimensional linear systems and that of finite-dimensional Gaussian systems. These two special cases occur most frequently in engineering. In Section 4 identification problems are discussed for which other classes of dynamic systems are appropriate models.

In several subsubsections that follow the exposition is necessarily mathematically technical. In each such subsubsection the reader interested in only the main development will be advised to read only the main definitions and results. This can be done with only a minor loss of understanding. The reader interested in concepts and theorems may read the full text of the paper, or only those parts of interest, or return at a later time to the technical parts.

The starting point of the identification procedure is the phenomenon to be modeled. Examples of phenomena are the behavior of an inverted pendulum in a control laboratory, a robot arm, the flow of benzo-a-pyrene in the body of an animal, or the gross domestic product of a nation. The model is expected to describe the inherent dynamics of the phenomenon. The model is to be based on time series of observations taken of the phenomenon.

3.1. Selection of Subclass of Dynamic Systems

The first step of the identification procedure is to select a subclass of dynamic systems with which to model the phenomenon. This is the most creative but also the most difficult part of the procedure. The dynamic system has to be based on laws of the domain of the phenomenon. Thus a combination of laws of physics, chemistry, engineering, biology, or economics has to be used.

As an example, a model for freeway traffic at a macroscopic level is formulated in terms of the state variables of density (in vehicles per km per lane) and of speed of the flow (in km per hour). The dynamics is based on a law of conservation of mass and on an impulse-like law formulated in terms of the behavior of drivers.

The modeler thus has to select: (1) the scope of the model; (2) the state variables of the model, which may be in a finite-dimensional space or an infinite-dimensional one; (3) the dynamics in the form of an ordinary differential equation, a difference equation, or a partial differential equation. The differential equation may be linear or nonlinear.

The concept of a dynamic system has been defined by R.E. Kalman in system theory, see the bibliography for references. In a dynamic system one distinguishes the state, the input signal, and the output signal. There are then axioms on how to go from an initial state with an input on a finite time interval to a final state and an output on the same time interval. In a stochastic dynamic system the definition differs from that of a deterministic dynamic system in that the map to the final state and the output maps to the distributions of these variables rather than the variables themselves.

It is important to note that dynamic systems are best represented in state space form. In many text books and papers dynamic systems are presented in forms that differ from state space form and this makes identification much more difficult, if not practically impossible.

Below two often used classes of dynamic systems are described.

Definition 3.1 *A time-invariant continuous-time* finite-dimensional linear system *is described by the linear ordinary differential equation,*

$$\dot{x}(t) = Ax(t) + Bu(t), \;\; x(t_0) = x_0, \tag{1}$$

$$y(t) = Cx(t) + Du(t), \tag{2}$$

$$\mathbf{U} \subseteq \{u : T \to U \mid a \text{ function}\},$$

where $n, m, k \in \mathbb{Z}_+$, $X = \mathbb{R}^n$ *is called the* state space, $U = \mathbb{R}^m$ *is called the* input space, $Y = \mathbb{R}^k$ *is called the* output space, $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times m}$, $C \in \mathbb{R}^{k \times n}$, $D \in \mathbb{R}^{k \times m}$, $x_0 \in X$ *is called the* initial state, $T \subseteq \mathbb{R}_+$ *is called the* time-index set, $u \in \mathbf{U}$ *is called the* input signal, $x : T \to X$ *is called the* state trajectory, *it is defined as the unique solution of the above differential equation, and* $y : T \to Y$ *is called the* output trajectory, *and it is defined by the above relation.*

The corresponding discrete-time system is defined similarly but with representation,

$$x(t+1) = Ax(t) + Bu(t), \;\; x(t_0) = x_0, \tag{3}$$

$$y(t) = Cx(t) + Du(t). \tag{4}$$

Denote by $LS(n, m, k)$ *the class of finite-dimensional linear systems as specified above and the parameters of any system in this class by* $(A, B, C, D, x_0) \in LSP(n, m, k)$.

Often extra modeling is required to formulate a model of the phenomenon in the form of a linear system.

The second class of systems is that of stochastic systems. Consider a probability space (Ω, F, P) consisting of a set Ω, a σ-algebra *F*, and a probability measure *P*. A random variable $x_0 : \Omega \to \mathbb{R}^n$ is said to be a *Gaussian random variable* if it has a Gaussian probability distribution. Denote by $x_0 \in G(m_0, Q_0)$ that x_0 has a Gaussian probability distribution function with mean value $m_0 \in \mathbb{R}^n$ and variance $Q_0 \in \mathbb{R}^{n \times n}$.

A *stochastic process* $x : \Omega \times T \to \mathbb{R}^n$ is a function such that for all $t \in T$, $x(.,t) : \Omega \to \mathbb{R}^n$ is a random variable. A *Gaussian stochastic process* or a *Gaussian process* is a stochastic process $x : \Omega \times T \to \mathbb{R}^n$ such that for all $k \in Z_+$ and $t_1, \ldots, t_k \in T$ the random variables $x_{t_1}, \ldots, x_{t_k}$ have a joint probability distribution function which is Gaussian. A *stationary stochastic process* is a stochastic process on either $T = \mathbb{R}$ or $T = \mathbb{Z}$ such that for any $k \in Z_+$, $t_1, \ldots, t_k \in T$, and $s \in T$, the probability distribution functions of the tuples $(x_{t_1}, \ldots, x_{t_k})$ and $(x_{t_1+s}, \ldots, x_{t_k+s})$ are equal. A *discrete-time Gaussian white noise process* is a stationary Gaussian process on $T = \mathbb{Z}$ such that for all $k \in Z_+$ and $t_1, \ldots, t_k \in T$ the random variables $(x_{t_1}, \ldots, x_{t_k})$ are independent, and for all $t \in T$, x_t has a Gaussian probability distribution function, denoted by $x_t \in G(0,V)$, with $V \in \mathbb{R}^{n \times n}$, $V = V^T \geq 0$.

Definition 3.2 *A discrete-time time-invariant* finite-dimensional Gaussian system (*for short, a Gaussian system*) *is a stochastic system described by the recursion and the equation,*

$$x(t+1) = Ax(t) + Bu(t) + M\upsilon(t), \;\; x(t_0) = x_0, \tag{5}$$

$$y(t) = Cx(t) + Du(t) + N\upsilon(t), \tag{6}$$

where, in addition to the previous definition, the initial state $x_0 : \Omega \to X$ *is a random variable with a Gaussian probability distribution function and the stochastic process* $\upsilon : \Omega \times T \to \mathbb{R}^r$ *is Gaussian white noise process. The initial state and the Gaussian white noise process are assumed to be independent. Moreover, the state process* $x : \Omega \times T \to X$ *is defined by the above recursion and the output process* $y : \Omega \times T \to Y$ *is defined by the above relation. Notation:* $(A, B, C, D, M, N, V) \in FDGSP(n, m, r, k)$.

Continuous-time Gaussian stochastic systems will not be described in this paper because they require the introduction of stochastic integrals and stochastic differential equations and thus more advanced mathematics than used in this article.

Example 3.3 *Nitrate model*. The model concerns the uptake and dispersion of nitrate in the human body. It has been developed at the Rijksinstituut voor Volksgezondheid en Milieuhygiëne (National Institute for Public Health and Environmental Hygiene) in Bilthoven, The Netherlands. The purpose of the model is to help setting standards for a new toxidity standard for the exposure to nitrate.

The model is a compartmental system. A compartmental system is a biological model consisting of a finite number of subsystems, which are called compartments. Each compartment is assumed to be kinetically homogeneous, any material entering the compartment is assumed to be instantly mixed with the material already in the compartment. The compartments interact via transportation and diffusion processes. The dynamics is based on the mass conservation law and, up to first order, the dynamics is described by a linear differential equation.

Four compartments of the body have been selected for inclusion in the model: nitrate (NO_3^-) in the stomach, in the body pool, and in the saliva, and nitrite (NO_2^-) in the saliva. The dynamics can be described by the following differential equations,

$$\dot{x}_1(t) = -K_a x_1(t) + \frac{b}{V_s} x_3(t) + u_1(t), \tag{7}$$

$$\dot{x}_2(t) = K_a x_1(t) - (K_2 + K_T) x_2(t) + K, \tag{8}$$

$$\dot{x}_3(t) = K_2 x_2(t) - (K_1 + \frac{b}{V_s}) x_3(t), \quad (9)$$

$$\dot{x}_4(t) = K_1 x_3(t) - \frac{b}{V_s} x_4(t). \quad (10)$$

in which x_1, x_2, x_3, x_4 denote the amounts of nitrate or nitrite in the compartments listed above, respectively. The input u_1 denotes the uptake of nitrate via food. The nitrate is converted to nitrite in the saliva. The concentrations of NO_3^- in the body pool and in the saliva, and the concentrations of NO_2^- in the saliva can all be observed.

The dynamic system is then transformed into a structured linear system of the form,

$$\dot{x}(t) = A(p)x(t) + B(p)u(t), \; x(t_0) = x_0(p), \quad (11)$$

$$y(t) = C(p)x(t) + D(p)u(t), \quad (12)$$

$$x = (x_1 \; x_2 \; x_3 \; x_4)^T, \; p = (K_a, K_2, K_1, b/V_s, 1/V_s, x_{0,1}, x_{0,2}, x_{0,3}, x_{0,4})^T, \quad (13)$$

the values of the constants K_T, V_d, are assumed known,

$$A(p) = \begin{pmatrix} -p_1 & 0 & p_4 & 0 \\ p_1 & -p_2 - K_T & 0 & 0 \\ 0 & p_2 & -p_3 - p_4 & 0 \\ 0 & 0 & p_3 & -p_4 \end{pmatrix}, \; B(p) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}, \quad (14)$$

$$C(p) = \begin{pmatrix} 0 & 1/V_d & 0 & 0 \\ 0 & 0 & p_5 & 0 \\ 0 & 0 & 0 & p_5 \end{pmatrix}, \; x_0(p) = \begin{pmatrix} p_6 \\ p_7 \\ p_8 \\ p_9 \end{pmatrix}, \quad (15)$$

$$P = \{p \in \mathbb{R}_+^9\}, \quad (16)$$

$$f(p) = (A(p), B(p), C(p), 0, x_0(p)) \in LSP(4,2,3), \text{ is a polynomial map.} \quad (17)$$

The parameters in the differential equations are constants and their names are omitted. The model is a positive linear system since the state variables take values in the positive real numbers, $\mathbb{R}_+$, and the input signal is positive also.

3.2. Data Collection

The second step of the identification procedure is to collect data in the form of time series.

One distinguishes the case in which the phenomenon can be excited by an input signal and that in which no input signal is possible or allowed. A robot arm with an input voltage to a motor is an example of a phenomenon with input signal. The rainfall in a region and solar activity are examples of phenomena which are not subject to input signals which can be influenced by humans. A national economy with as input variables the taxation rate and central bank rate is a phenomenon where input signals are possible yet almost never consciously used for obtaining observations for mathematical models. For biological and economic phenomena the time series are often short and the costs to collect the data are high. Experiments with human beings or animals are other examples. For many engineering models the time series can be as long as one can handle.

In case the modeler can supply an input signal to the dynamic system for data generation there is a problem: What input signal to choose? The simplest approach is to provide a small nonzero input signal at a particular time and to record the response of the output signal of the system. An input sequence generated by a pseudo-random number generator is often used as input signal. But the modeler often wants more; the input signal should be the best possible for the task. There is a circularity at this point. To determine an input signal which excites the phenomenon such that the dynamics can be well identified, one needs to use a realistic model of the phenomenon. But to get a realistic dynamic system one has to choose an input signal. In practice one applies a procedure with iterations. A simple random input signal is applied to the system. A first estimate of the dynamic system is made. Based on the first estimate of the dynamic system one then selects a more appropriate input signal, which excites the system in the frequency range or in the modes of interest. Then a second and more refined model is computed. Experiences with this are quite positive.

The observed time series has to be preprocessed. Thus computation of averages and variances is useful. Detection of malfunctioning of measurement devices is necessary. Time series with outliers need treatment. An outlier is a data point far outside the expected range of observed data. The cause of this outlier has to be determined and if it is due to a malfunctioning device then the time series is discarded or the data point replaced by a smoothed data point.

-
-
-

Bibliography

[1]T. Bohlin. (1976) Four cases of identification of changing systems. In R.K. Mehra and D.G. Lainiotis,

editors, System Identification – Advances and case studies, pages 441–518. Academic Press, New York.

[2]T. Bohlin. (1991) Interactive system identification: Prospects and pitfalls. Springer Verlag., Berlin.

[3]F.M. Callier and C.A. Desoer. (1991) Linear system theory. Springer-Verlag, New York. [A basic reference on linear systems.]

[4]Chi-Tsong Chen. (1999) Linear system theory and design (3rd Ed.). Oxford University Press, Oxford. [Another reference on linear systems.]

[5]P.Faurre. Stochastic realization algorithms. In R.K. Mehra and D.G. Lainiotis, editors, (1976) System Identification – Advances and Case Studies, pages 1–25. Academic Press, New York. [An article on the weak stochastic realization problem.]

[6]P. Faurre, M. Clerget, and F. Germain. (1979) Opérateurs rationnels positifs (in French). Dunod, Paris. [A book on the weak stochastic realization problem for stationary Gaussian processes and on the dissipativity of linear systems.]

[7]G.C. Goodwin and R.L. Payne. (1977) Dynamic system identification – Experiment design data analysis. Academic Press, New York.

[8]E.J. Hannan and M. Deistler. (1988) The statistical theory of linear systems. John Wiley & Sons, New York, 1988. [This book presents concepts and theorems on realization and on parameterization for finite-dimensional Gaussian systems.]

[9]R.E. Kalman. (1963) Mathematical description of linear dynamical systems. SIAM J. Control, 1:152–192. [The main article on realization of finite-dimensional linear systems.]

[10]R.E. Kalman, P.L. Falb, and M.A. Arbib, (1969) Topics in mathematical systems theory. McGraw-Hill Book Co., New York. [This book presents the main concepts of dynamic systems and theorems on realization of linear systems.]

[11]T. Katayama and G. Picci. (1999) Realization of stochastic systems with exogenous inputs and subspace identification methods. Automatica, 35:1635–1652. [Subspace identification algorithms with exogeneous inputs.]

[12]W.E. Larimore. (1983) System identification, reduced-order filtering and modeling via canonical variate analysis. In Proceedings 1983 American Control Conference, pages 445–451. [An early reference on the subspace identification algorithm.]

[13]A. Lindquist and G. Picci. (1996) Canonical correlation analysis, approximate covariance extension, and identification of stationary time series. Automatica, 32:709–733. [Subspace identification algorithms seen from stochastic realization theory.]

[14]A. Lindquist and G. Picci. (1996) Geometric methods for state space identification. In S. Bittanti and G. Picci, editors, Identification, adaption, learning, pages 1–69. Springer, London. [The theory of strong stochastic realization of stationary Gaussian processes.]

[15]L. Ljung. (1987) System identification: Theory for the user. Prentice-Hall, Inc., Englewood Cliffs, NJ. [A main reference on identification as considered in electrical engineering.]

[16]J.W. Polderman and J.C. Willems. (1997) Introduction to mathematical system theory – A behavioral approach. Number 26 in Texts in Applied Mathematics. Springer, New York. [The behavioral approach to system theory in which observations are distinguished into inputs and outputs.]

[17]G. Ruckebusch. (1980) Théorie géométrique de la représentation markovienne. Ann. Inst. HenriPoincaré, 16:225–297. [Major contribution to the theory of strong stochastic realization of stationary Gaussian processes in a Hilbert space framework.]

[18]T. Söderström and P. Stoica. (1989) System identification. Prentice Hall, New York. [A major text book addressed to electrical engineers.]

[19]E.D. Sontag. (1998) Mathematical control theory: Deterministic finite dimensional systems (2nd. Ed.). Number 6 in Graduate Text in Applied Mathematics. Springer, New York, 1998. [The main reference on control and system theory, addressed to mathematicians.]

[20]A.A. Stoorvogel and J.H. van Schuppen. (1996) System identification with information theoretic

criteria. In S. Bittanti and G. Picci, editors, Identification, adaptation, learning, pages 289–338. Springer, Berlin. [Discussion of information theoretic criteria for identification.]

[21]J.M. van den Hof. (1998) Structural identifiability of linear compartmental systems. IEEE Trans. Automatic Control, 43:800–818. [Major source for structural identifiability of linear and of positive linear systems.]

[22]P. van Overschee and B.L.R. De Moor. (1996) Subspace identification for linear systems. Kluwer Academic Publishers, Dordrecht. [Book on the subspace identification algorithm.]

Biographical Sketches

Jacqueline van den Hof, is affiliated with Statistics Netherlands in Voorburg, The Netherlands, working on the National Accounts, in particular the Social Acconting Matrix and the sector households. She studied at the department of Mathematics of the University of Groningen, The Netherlands and was awarded a Ph.D diploma by the department of Mathematics and Natural Sciences of the University of Groningen in 1996. In the period 1992-1996 she was affiliated with the research institute Centrum voor Wiskunde en Informatica (CWI) in Amsterdam, The Netherlands, and worked at system identification of compartmental systems with Jan H. van Schuppen as research advisor.

Jan H. van Schuppen, is affiliated with the Stichting Centrum voor Wiskunde en Informatica and its research institute Centrum voor Wiskunde en Informatica (CWI) in Amsterdam, The Netherlands. In the period January - June 2002 he is on sabbatical leave at the Department of Electrical and Computer Enginering of the University of Illinois in Urbana-Champaign, IL, U.S.A. He has a part-time affiliation, for one day a week, with the Department of Mathematics of the Vrije Universiteit (Free University) in Amsterdam. He has studied at the Department of Applied Physics of Delft University of Technology and was awarded a Ph.D. diploma by the Department of Electrical Engineering and Computer Science of the University of California at Berkeley, CA, U.S.A. in 1973 where Pravin Varaiya was his research advisor.Van Schuppen's research interests are control and system theory, in particular, they include realization theory, identificationn stochastic control, control of discrete-event systems, and control of hybrid systems. In applied research his interests include engineering problems of control of motorway traffic, of communication networks, and modeling, identification, and control in the life sciences and biotechnology. He was the research advisor of 11 Ph.D. students and several post-doctoral researchers. He regularly teaches graduate courses in control and system theory.He is Editor-in-Chief of the journal 'Mathematics of Control, Signals, and Systems'. He has been Associate Editor-at-Large of the journal 'IEEE Transactions Automatic Control' (1998-2001) and Department Editor of the Journal 'Discrete Event Dynamic Systems' (1990-2000). He was Co-Chairman of the International Symposium on the Mathematical Theory of Networks and Systems 1989 (MTNS '89), of the Workshop Hybrid Systems - Computation Control 1998 (HSCC1998), and organizer of several other workshops and summer schools.He is the coordinator of the Project System Identification (EU.TMR ERBFMRXCT 980206) that is financially supported by the European Commission and has been the coordinator of two earlier projects in the same topic. He has participated in seven projects by the European Commission.