1

Introduction

The word system, (from the Greek $\sigma\upsilon\sigma\tau\eta\mu\alpha$) is used in a diversified range of contexts and has thus assumed different meanings. The most common definition of system concerns a group of parts linked by some form of interaction. In the context of System Theory a system can be defined as a slice of reality whose evolution in time can be described by a certain number of measurable attributes.

The measuring of system attributes introduces the problem of establishing quantitative relationships between them, i.e. to construct abstract (mathematical) models. The role played by mathematical models in science and technology is today of paramount importance and has attracted increased attention on model construction procedures. Identification is one of these procedures and consists in deducing mathematical models from the observations performed on a system. This approach involves a large variety of choices that lead to many possible schemes; some of these schemes are described in the following with particular reference of their application to multivariable processes.

1.1 SYSTEM MODELS

The etymological roots of model come from the Latin word *modus* and its diminutive *modulus* both meaning measure; its initial use in science and technology can be associated to scale representations used by architects to reproduce the shape of buildings before their actual construction. In this case the models are physical systems that reproduce approximately the aesthetic properties of other physical systems before their realization.

An important advancement was achieved with the introduction of models that were still small scale reproductions of physical systems but were used however to investigate their behavior either before construction or under operating conditions otherwise impossible or too expensive. Well known examples include hydrological models and models of boats that were used to evaluate their hydrodynamical properties. In such cases very complex phenomena are reproduced without any knowledge of the mathematical relations linking the relevant variables (water flows, velocity profiles, boat speed and hydrodynamical resistance). It is also necessary however, to scale properly the variables in order to assure an accurate reproduction of the case to be studied; laminar flows, for example, must not become turbulent flows and *vice–versa*. A further step was taken by reproducing the behavior of a system on another system that could be more easily studied, taking advantage of the different physical laws that can be described by formally equal relations. A typical example is given by analog computers, structured as flexible electrical networks that, properly interconnected (programmed), reproduce the behaviors of other systems (mechanical, hydraulic, economic etc.) less suitable for direct experiments. These models could be defined as analog or, to avoid any confusion with the common use of this term (analog to denote quantities whose measure can be performed with continuity, as opposite to digital used to denote quantizations), models based on analogy laws.

It can be observed that these models offer greater flexibility than previous ones where the only degree of freedom was the scale factor, since the physical nature of the model is also a matter of choice. However, the use of models based on analogy laws requires the knowledge of the laws describing the behavior of the system to be studied, since it is necessary to select, construct or configure another system governed by analog laws by studying its behavior from suitable initial conditions. What is, on the contrary, not required, is the capability of constructing a complete mathematical model and of using it to determine the system behavior.

The last step in the evolution of models is the development of abstract models i.e. mathematical models that describe the links established by the system between its measurable attributes. The important role played currently by mathematical models in science and technology is due to the availability of both the abstract tools offered by System Theory and numerical computers that allow their effective use.

The evolution in use of models fits well the criterion of growth for civilization given by Toynbee (1934–61) and (1962–64). According to this author, this criterion does not consist neither in increasing command in the human environment nor in increasing command over the physical environment, but in "etherialization". This concept, that Toynbee derives from Heard, is rather complex but oversimplifying, involves a transfer of emphasis from some lower sphere of action to a higher sphere; an example could be the media used to record information that from flint and sheepskin have evolved to paper, magnetic and optical media.

1.2 EVOLUTION OF MATHEMATICAL MODELS

Despite the comparatively recent growth in the role of mathematical models, their origin can be traced back (in western culture) to Aristotle (384–322 B.C.) who recognized the importance of numerical and geometrical relations within science indicating also mechanics, astronomy, optics and harmonics as fields where mathematical relationships linking physical quantities are particularly important.

Mathematical models were used to describe the motion of planets by Ptolemy (85–165), who also observed that different models can be constructed to describe the same astronomical observations. Galileo (1564–1642) formulated the law of falling bodies as a mathematical model obtained on the basis of experimental and theoretical work. Copernicus (1473–1543), Kepler (1571–1630), Newton (1642–1727), Halley (1656–1742) are among the best known scientists who used mathematical models as interpretation tools for the physical world.

Mathematical models have been defined as sets of relations among the measurable attributes of a system, describing the links established by the system among these quantities. This limits the descriptive capability of mathematical models to attributes that can be expressed by means of numbers and furthermore shows that these models constitute, in any case, only an approximate description of reality. The intrinsic approximation performed by introducing models can be better evaluated in the context of a classification based on the purposes of modeling.

1.2.1 Interpretative models

The rationale of these models lies in satisfying scientific curiosity and rationalizing the behavior of observed processes. They can also be seen as ways to extract essential information from complex experiments or to substitute (large amounts of) data with a data generating mechanism. Models of this kind have been developed by Ptolemy, Copernicus, Kepler, Galileo, Newton and Halley to describe the motion of physical objects. The purpose of interpretative models is to increase the understanding of a slice of reality existing behind the observed phenomena; they must thus "interpret" sets of collected data but they don't necessarily have any capability to generate other (future) sets of data (that will be) generated by the same system.

Interpretative models are used in a large number of disciplines like econometrics, ecology, life sciences, agriculture, physics. Most physical laws can be seen as models of this kind. Ptolemy's observation on the possibility of describing the same observations with different models highlights that the interpretation essentially concerns some measurable attributes of phenomena and not (necessarily) their actual nature. Another important observation concerns the approximations of interpretative models and/or their limited range of validity. So Newton's law of motion, giving a simple relation between the force acting on a mass and its acceleration, leads to large errors for speeds approaching the speed of light.

Example 1.2.1 – Sunspot cycle

The plot of Figure 1.2.1 shows the yearly mean sunspot count from 1749 to 1983, computed from daily relative sunspot numbers evaluated on the basis of more than fifty observing stations around the world.

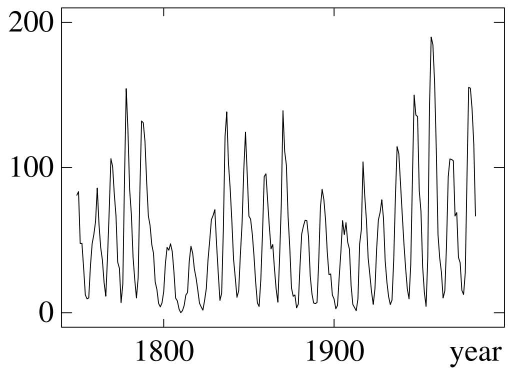


Figure 1.2.1 – Yearly mean sunspot count from 1749 to 1983

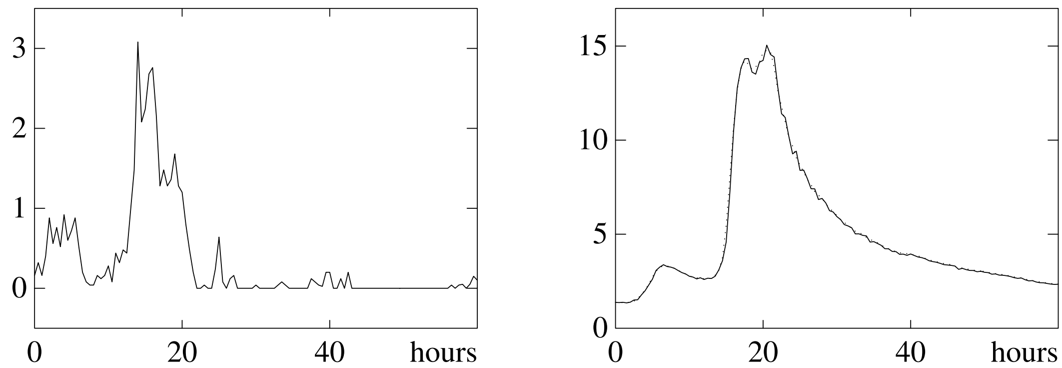
Estimating from this sequence (after subtracting its mean value) a second order autoregressive (AR) model with the least squares algorithm, we obtain the model $y(t+2) = 1.3873\, y(t+1) - 0.6937\, y(t)$ whose poles, $p_{1,2} = 0.6936 \pm i\, 0.4611$, indicate a periodicity of 10.71 years for the phenomenon. This "law", obtained by means of a mathematical elaboration of the observations, compares well with the commonly assumed period of 11 years and with the approximate evaluation that can be directly obtained from the plot.

1.2.2 Predictive models

The rationale of predictive models is forecasting the future behavior of a system i.e. interpolating available observations into the future. This is probably the most frequent use of mathematical models, with applications in many different fields (e.g. forecasting demands of specific products, weather conditions, population growth, the future state of an ecosystem or of a plant). The predictions obtained in this way are often used to manipulate the inputs of the considered system to achieve specific objectives like the desired attitude of an aircraft or of a missile, the position of a robot arm, the degree of purity of the output of a distillation column or the inflation rate. Other less obvious applications of predictive models concern speech and image processing to reduce bandwidth requirements in transmission and recording. A model can, of course, be at the same time an interpretative model and a predictive model. As observed by Norton (1986), when Halley in 1704 realized that the observations of 1531, 1607 and 1682 referred to the same comet and computed its orbit, he constructed an interpretative model that predicted accurately the subsequent return of 1758.

Example 1.2.2 – Forecast of a river flow

Since 1975 the Welsh Water Authority operates a real–time flow forecasting system on the River Dee as part of extensive water supply and flood control schemes for the catchment. The River Hirnant's subcatchment, with an area of 33.9 km^2, is situated west of Bala Lake, in North Wales. It is composed mainly of rocks, providing very little storage for rainfall. Furthermore, because of its steep slopes, it causes a fast streamflow response to rainfall. Figure 1.2.2 reports a rainfall recording over a period of 60 hours and Figure 1.2.3 the corresponding streamflow measured at Plas Rhiwaedog; the data are sampled at half–hourly intervals.



Figures 1.2.2 and 1.2.3 – Rainfall on the catchment and River Hirnant streamflow

Also the forecast of a model, obtained with identification techniques, whose input is given by rainfall, is reported (dotted line) in Figure 1.2.3. While short–term forecasts (few hours), useful for early flood warning, can rely on the available rainfall measures, long–term forecasts, useful for water resources management, must rely on weather forecasts and are, consequently, less accurate.

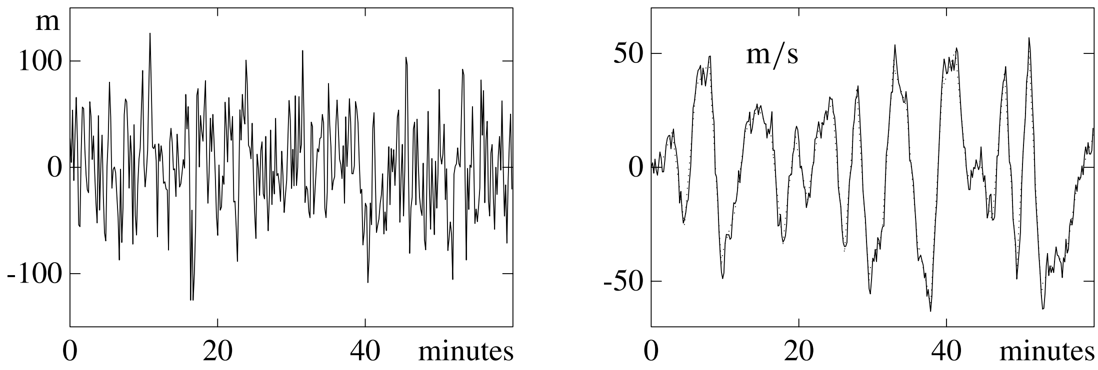
1.2.3 Models for filtering and state estimation

The rationale is here the extraction of some external variables (output) from noisy measurements performed on the system and/or the estimation of some internal variables (state) from external measures affected by errors.

Applications concern the reception and processing of radio signals (e.g. telemetry and pictures sent from a spacecraft), transmission of digital data over noisy channels (e.g. telephone lines), processing of radar signals, analysis of electrocardiographic and electroencephalographic signals, geophysical data processing, monitoring of industrial plants and of natural systems, demography. An application frequently cited is the Kalman filter used to estimate the state (position and velocity) of the spacecraft (Apollo 11) in the first manned lunar mission; all other space missions to Mercury, Venus, Mars and beyond relied, even more heavily, on these techniques.

Example 1.2.3 – Tracking of a maneuvering target

The altitude of a maneuvering target, given every 10 s by a radar system is affected by an error with a standard deviation $\sigma_a = 49$ m. The actual altitude is estimated by means of a Kalman filter that reduces the error standard deviation to $\hat{\sigma}_a = 43$ m (Figure 1.2.4). The state of the Kalman filter gives also an estimate of the vertical speed of the target, reported in Figure 1.2.5 against its actual value.



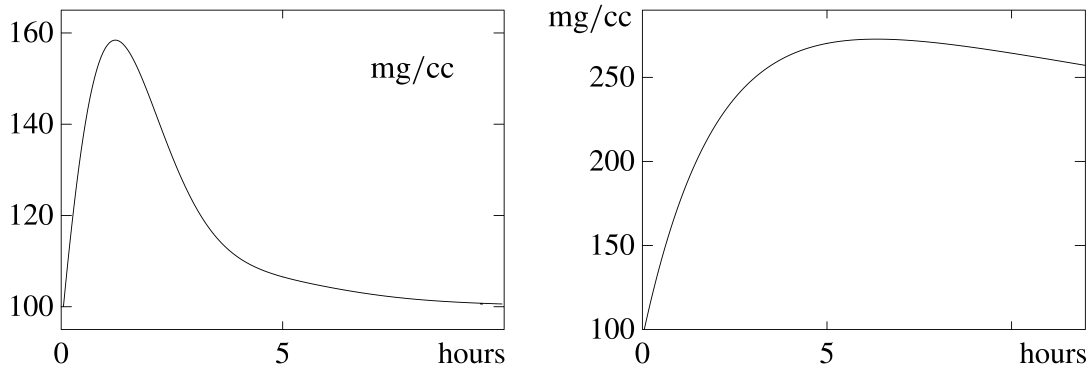
Figures 1.2.4 and 1.2.5 – Altitude error and estimated vertical speed of the target

1.2.4 Models for diagnosis

The computation of the specific model best fitting a set of data collected on a process allows its general behavior to be compared with a class of behaviors established as a reference, evaluating abnormal conditions like a sensor fault in an industrial process or a disease in a patient. The sulfobromophthalein (BSP) and the glucose tolerance (IVGT) tests are routinely used in the medical practice as aids in the assessment of hepatobiliary and pancreas diseases. In both cases the test starts with intravenous injections of these substances and is followed by measures of their plasmatic concentrations at specific intervals. Values beyond certain limits indicate a slow metabolism that could be associated to hepatitis or diabetic conditions.

Example 1.2.4 – Intravenous glucose tolerance test (IVGT)

The control of blood sugar levels in the human body is carried out by the insulin secreted by the pancreas when the sugar level exceeds the physiological equilibrium value.



Figures 1.2.6 and 1.2.7 – Response of normal and diabetic patients to a glucose loading

The rate of change of blood sugar levels after a glucose injection gives a reliable description of the efficiency of this regulation mechanism as follows from the comparison of Figure 1.2.6, reporting the response of a normal individual, with Figure 1.2.7 regarding the response of a diabetic. The measures performed on a patient, compared in Figure 1.2.8 with the standard response, allow to diagnose the presence of abnormal conditions.

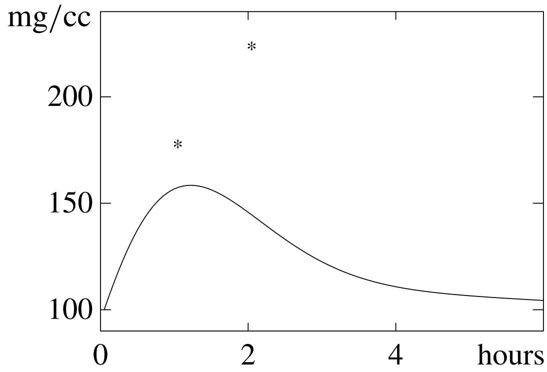


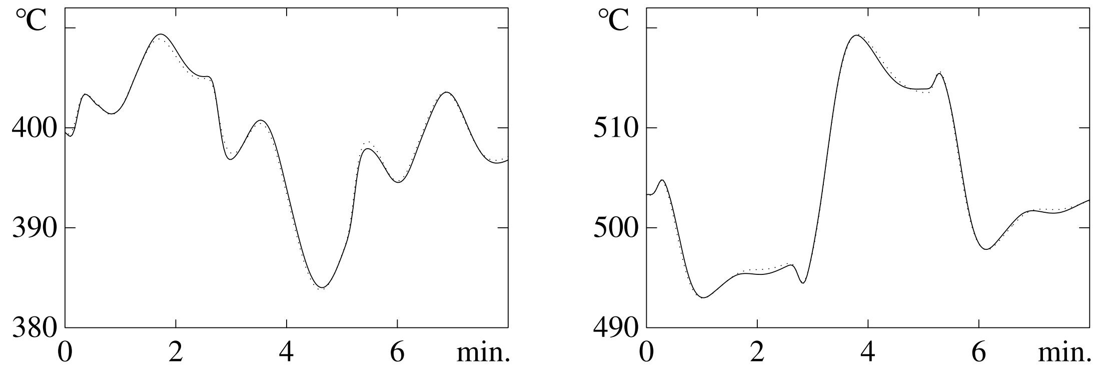
Figure 1.2.8 – Standard response and abnormal measures obtained on a patient

1.2.5 Models for simulation

The rationale is here the substitution of real systems with their models to evaluate their response to assumed control policies (inputs). A substitution of this kind can be very rewarding from an economic point of view and can also allow performing operations that would have been otherwise impossible or risky on real systems (e.g. demographic studies, the responses in a national economy to changes in interest rates, pilot training, major nuclear reactor incidents etc.). Of course the usefulness of simulations depends on the accuracy of the model in reproducing the behavior of the actual system; the etymology of simulation (the Latin *simulare* = to pretend) seems to suggest the possible ambiguity of this substitution.

Example 1.2.5 – Simulation of a sodium heat exchanger

PEC is a LMFBR (Liquid Metal Fast Breeder Reactor) with a thermal power of 120 MW, designed to test experimental fuel elements developing powers up to 3 MW in the thermal and neutron flux conditions that are met in large fast breeder nuclear reactors. The cooling of the core is performed by means of a double sodium primary loop and sodium–sodium heat exchanger, a secondary loop and sodium–air heat exchangers. The dynamical behavior of the reactor in emergency situations (e.g. failure of the pump in one of the primary loops) has been investigated by means of a large simulation package which includes the models of every part of the plant. This model is, however, unsuitable for real–time simulations because of its size. A reduced–order model obtained with identification techniques has been developed for real–time simulations regarding both operator training and process control.



Figures 1.2.9 and 1.2.10 – Output temperatures of the PEC sodium heat exchangers

Figures 1.2.9 and 1.2.10 show the output temperatures of the primary and secondary sodium heat exchangers given by the model (dotted line) against the true values for variations of the inputs (primary and secondary sodium flows and input temperatures) of approximately 20%. The limited error given by this model is fully compatible with its planned use.

1.3 MODELS AS APPROXIMATIONS OF REALITY

It has already been observed that mathematical models limit their description to the quantitative links established by real systems between their measurable attributes so that they constitute, in any case, only partial descriptions. The asymptotic evolution of science has however also canceled the illuministic illusions on the possibility of achieving exact descriptions of reality. Even the relations accepted as laws of nature can be considered, at most, as models *not yet* falsified. Newton's law of motion is a good example of the extended acceptance of a mathematical relation as an absolute description of a phenomenon before its falsification, but it's also a good example of the excellent accuracy of a simple model in describing a very large range of situations.

Many phenomena are simply too complex to be described in detail by manageable models and/or are not ruled by any definite law of nature (e.g. national economies). The construction of mathematical models should thus be ruled by *usefulness* criteria more than by (always relative) *truth* criteria. The inherent approximations associated to models outline that different models of the same system can be used for different purposes (interpretation, prediction, filtering, diagnosis, simulation) optimizing their performance for these tasks. The criteria to compare and select models have, consequently, both philosophical and practical importance. A well known criterion is the "razor of Occam", due to William of Occam (1290–1350) establishing that the simpler among the models accounting for the same phenomenon must be preferred. This principle certainly helped the acceptance of the model proposed, for the solar system, by Copernicus who, prudently, emphasized that his heliocentric model should have been considered only as an exercise to obtain in a simpler way the results of the officially accepted Ptolemaic model.

A different description of the parsimony principle can be found in the work of Popper (1934), (1963). According to this author, among the models that explain the available observations, the model explaining as little else as possible (the most powerful unfalsified model) is to be preferred. The parsimony principle is supported not only by philosophical arguments (and by common sense) but also by mathematical arguments that show how increasing model complexity leads, when the models are deduced from uncertain data, to corresponding increases in the uncertainty of their parameters.

1.4 MODEL CONSTRUCTION

The unavoidable association of the concepts of model and approximation puts modeling in a twilight zone that does not belong to pure science due to the lack of unicity postulates but must, in any case, rely on results and methodologies offered by the abstract science of mathematics.

Moreover the limited capability of models to solve any specific problem leads to the opportunity of constructing application–oriented models; the rationales behind the different uses of models that have been described previously become thus rationales also behind the construction of special–purpose models for the same system.

1.4.1 Deducing models from other models: physical modeling

The procedures to obtain mathematical models are usually classified into physical modeling and identification. Physical modeling is based on the partition of a system into subsystems and on their description by means of known laws. The model is then obtained joining such relations into a whole. This approach requires a general knowledge of the structure or design of the considered system and of the "laws" describing their behaviors. Since physical laws are, in turn, models obtained from observations or from unfalsified speculations, physical modeling consists of constructing a whole model joining together simple and already established models. The advantages of physical modeling consist in the possibility of using, in the model construction procedure, *a priori* information on the system and in the physical meaning of the model variables. This procedure cannot be applied, however, to systems whose internal structure is not known, whose behaviors are not described by established relations or whose complexity would lead to unmanageable models where most parameters would only marginally influence the aspects to be reproduced of the system behavior.

1.4.2 Deducing models from data: identification

Identification consists in the selection of a specific model in a specified class on the only basis of observations performed on the system to be described and of a selection criterion. The whole procedure makes no reference neither to the physical nature of the modeled system nor to the *a priori* knowledge of the modeller; only the data speak.

The internal variables of identified models may lack any physical meaning and the same can be true of the model parameters. Such models are, on the other side, simple, accurate and can extract from complex frameworks only some relevant aspects.

It is not difficult to recognize that often physical laws have been obtained as a result of identification procedures; the data collected by Galileo in his experiments on falling bodies, for instance, led him to recognize that a simple model could explain every experiment and could consequently be considered a law.

1.5 IDENTIFICATION STEPS

The identification of a dynamical system can be partitioned, from a logical standpoint, into the following steps:

Collecting the data (observing the system)
The purpose of identification exercises is to substitute collections of data with a data–generating mechanism (the model). Since this operation is exclusively based on the information contained in the data, it is useless to carry out the procedure when the data is garbage. The first step in identification consists thus, in collecting observations of the process variables (inputs and outputs if a partition of this kind has been introduced) and, when possible, in applying to the system inputs particularly suitable for identification purposes.

Selecting a model set
A specific model can be selected, on the basis of observations, only inside an assumed family of models. These *a priori* assumptions on the set of admissible models can be guided by the modeller's experience and result even improper.

Choosing a selection criterion
Any model only approximates the true system behavior; the planned use of the model will thus determine the selection criterion in order to optimize its desired features. Different models can thus be extracted from the same data.

Computing model parameters
The computation of the model parameters can be seen as an optimization problem (the selection of the "best" model in the considered class) or, equivalently, as a way of "tuning" a model on the data.

Validating the model
Validation is a trial to falsify the model, better by using collections of data different from those used for identification. If the model remains unfalsified, it is considered validated, i.e. acceptable for the planned purposes.

These choices are never univocal and can also involve some degree of *a priori* knowledge on the system, for instance in the choice of the class of models or in the design of the experiments required to collect the input–output sequences. When the model does not pass the final validation step, it is necessary to reconsider the choices that have been made and to restart the identification procedure from the beginning or from an intermediate step.

1.6 IDENTIFIABILITY

Once a class of models and a penalty function for the model behavior have been selected, identification reduces, for any set of available data, to an optimization problem whose solution consists in selecting the model associated to the minimal value of the penalty function. The problem is well defined if and only if the range of the penalty function contains a single absolute minimum; in this case the considered process is identifiable. Identifiability derives from the class of models that has been selected, from the penalty function and from the data, not from the system to be identified.

1.7 CLASSES OF MODELS FOR IDENTIFICATION

Many different classes of models can be considered. The most relevant, from an identification standpoint, are: oriented and non oriented, algebraic and dynamical, causal and non causal, lumped and distributed, constant and time–varying, linear and nonlinear, deterministic and stochastic, single–input single–output (SISO), multi–input single–output (MISO) and multi–input multi–output (MIMO), parametric and non parametric, continuous and discrete.

1.7.1 Oriented and non oriented models

Once the measurable attributes of a system have been defined, it is usual to partition them in two classes: inputs and outputs or also, in econometrics, exogenous and endogenous variables. The inputs can often be seen as the action of the surrounding environment on the system and the outputs as its reaction. It is also possible, following Willems (1986), to describe the outputs as the variables explained by the model and the inputs as those left unexplained by the model. In some cases it can be desirable to avoid any *a priori* orientation of the system and treating all variables in a symmetric way. It is important to note that the orientation can be imposed by the external environment instead than by the system itself.

Example 1.7.1

The measurable attributes of the electrical dipole in Figure 1.7.1 are the voltage and the current at his terminals.

If this dipole is connected to a voltage generator, it will be natural to consider the voltage as input and the current as output while a connection to a current source will reverse the situation. In any other case both orientations are equivalent.

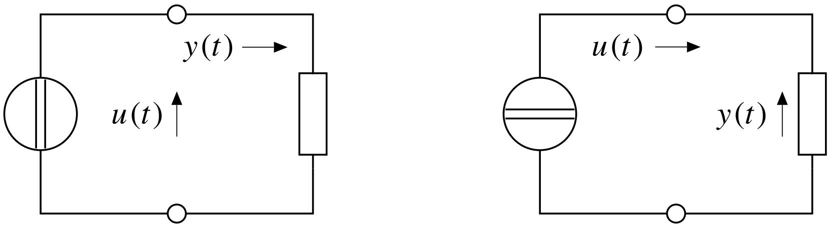


Figure 1.7.1 – Possible orientations of a dipole

1.7.2 Algebraic and dynamical models

Algebraic systems establish an instantaneous link between their variables and are described by sets of algebraic equations. Dynamical systems establish, on the contrary, a link between the values assumed by their attributes at different times and are described by sets of differential or difference equations. Algebraic systems can be treated in a comparatively simple way and describe a very limited slice of the real world; system identification is thus, almost implicitly, considered as dynamic system identification and makes reference to dynamical models. The identification of models for algebraic processes can, however, be more tricky than commonly assumed and present also conditions of non identifiability.

1.7.3 Causal and non causal models

An oriented model is defined as causal when its output at time t is not affected by future input values. While all real systems are causal and can be properly described by models of this kind, it is also possible, from a mathematical standpoint, to introduce non causal models.

1.7.4 Purely dynamic and non purely dynamic models

An oriented dynamic model is defined as purely dynamic when its input at time t does not affect its output at t, i.e. when the system does not establish any instantaneous (algebraic) link between its input and its output. If this condition is not satisfied the model is defined as non purely dynamic. This property interacts with other properties and with the planned use of the model. Thus a non oriented model is necessarily non purely dynamic because a purely dynamic model is intrinsically oriented and would even become non causal for other orientations. A predictive model must, on the contrary, be purely dynamic because (considering discrete systems) the output at time $t+1$ must be predicted at time t on the only basis of measures performed until t.

1.7.5 Lumped and distributed models

Most aspects of reality concern phenomena that don't take place in a single point of the space but affect areas or volumes (e.g. heat transmission, electromagnetic phenomena, energy exchanges, mechanical systems etc.). Such phenomena would require distributed models given, for dynamical systems, by sets of partial differential (or difference) equations, i.e. by distributed models. Lumped models, given by sets of ordinary differential (difference) equations refer to simplified schemes that assume constant values of the system attributes in some, properly defined, space regions.

1.7.6 Constant and time–varying models

Time–varying models, whose parameters are, in general, functions of time, can describe systems whose behavior changes with time. Time–invariant models are described by sets of constant parameters and can describe constant systems. Time–invariant systems, sometimes, look as time–varying because of lack of knowledge on some of their inputs.

1.7.7 Linear and nonlinear models

Linear models describe systems where the superposition principle is valid. Most real systems are nonlinear but can be described accurately by a linear model in the neighborhood of a working condition.

1.7.8 Deterministic and stochastic models

Real systems are always affected by disturbances (noise entering into the system and/or affecting the measures, unknown inputs, quantization errors etc.). These disturbances or their global effect can be described by means of noise acting on the input, state and output of the model which is called, in this case, stochastic. Often the global effect of disturbances is modeled as the output of a filter driven by white noise, and is added to the output of the deterministic part of the model which is thus decomposed into a deterministic and a stochastic part. Depending on the applications it can be sufficient to identify the deterministic part of the model (e.g. diagnosis) or it can be necessary to identify both parts (e.g. prediction).

1.7.9 SISO, MISO and MIMO models

Single–input single–output (SISO), multi–input single–output (MISO) and multi–input multi–output (MIMO) models have self–explanatory names. It can be observed that while SISO models have a limited usefulness (the world is multivariable), a multivariable model can be decomposed into a collection of MISO models. As discussed in the following, an approach of this kind, frequently followed to avoid the more complex tools required by the use of truly multivariable models, has many conceptual and practical limits.

1.7.10 Parametric and non parametric models

Some classes of models are given by sets of equations described by a certain number of parameters (parametric models) while other models are given without assigning any parameter, for instance in a graphical form (e.g. impulse or step responses and frequency response for linear systems).

1.7.11 Continuous and discrete models

Continuous models describe systems whose measurable attributes evolve with continuity in time while discrete models establish quantitative links only between the values assumed by the variables at discrete (sampling) times. While the intrinsic nature of all natural systems and of many technological systems is continuous, the widespread introduction of digital systems emphasizes the use of discrete models that can describe accurately, when the variables are properly sampled, also continuous systems.

1.7.12 Free and non free models

In some cases no action of the environment on the considered system exists or, more likely, can be observed. Systems and models without any input are defined as free; their outputs are called time series.

1.8 STRUCTURE OF THIS BOOK

Despite a wide range of possibilities and complexities, most identification procedures refer to a single class of models: lumped, linear, discrete, SISO or MISO constant models.

The use of lumped models is a common practice also outside the identification field because models of this kind combine, when their order is suitable, the possibility of giving accurate descriptions also for distributed processes and an easy use. The choice of linear models relies, on one side, on the necessity to describe most real systems, usually nonlinear, only in the neighborhood of specific working conditions where the behavior is essentially linear and, on the other side, on the powerful tools available for linear systems. When several working conditions are possible, a set of linear models can be preferred to a single nonlinear model. Discrete models are imposed by the general availability of sampled measures also for continuous systems and by the widespread use of digital controllers. With the exception of limited classes of time–dependent systems (e.g. periodic systems), constant models are often used also for time–dependent systems, adopting, in these cases, on–line identification schemes which tune the model on system variations.

This book also adopts previous choices and limits its content to equation error models i.e. to the sector that still constitutes the mainstream approach to system identification. Differently from most works on identification, the identification of minimally parametrised multivariable models is treated in detail. To properly evaluate the importance of identifying true multivariable models it is sufficient to recall that a collection of MISO models can be reduced to a MIMO model with the lowest possible complexity, i.e. without redundancies, only when the single MISO models are known *exactly*. In all other cases it is impossible to recognize the duplicated dynamics and insert them only once into the model; the reduction of a collection of uncertain MISO models to a single reduced–order MIMO model involves approximations and the introduction of specific criteria (many available) that can alter in a substantial way the optimality of the final model. The internal structure of MIMO systems is, however, substantially richer than that of MISO systems and, as a consequence, the use of MIMO models in identification requires a clear understanding of their properties.

The whole exposition follows a *model–oriented* instead than the more frequent *algorithm–oriented* approach; specific algorithms and their properties are introduced only when necessary to develop identification procedures for the models under study. This choice derives from many years of experience in teaching this subject.

The German physician and erudite Georg Bauer, best known as Georgius Agricola, stated, in the introduction of his monumental work *De Re Metallica* (1556), that remained an unsurpassed textbook and guide for miners and metallurgists for 180 years "Io non ho scritto cosa niuna la quale non abbia veduta, o letta o con accuratissima diligenza esaminata, quando che da altrui mi sia stata raccontata" (I have not written anything about what I have not seen personally or very carefully examined when reported by others). Any comparison between this book and the magnificent *De Re Metallica*, whose 290 carvings alone make it an exceptional artwork, would be out of place. I can however state that no parts of it concern schemes, algorithms and procedures not directly tested and applied to data obtained from simulated and real processes.