ID1

Introduction

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.

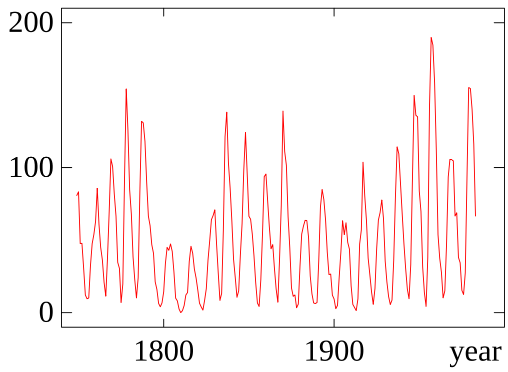


Figure 1.2.1 – Yearly mean sunspot count from 1749 to 1983

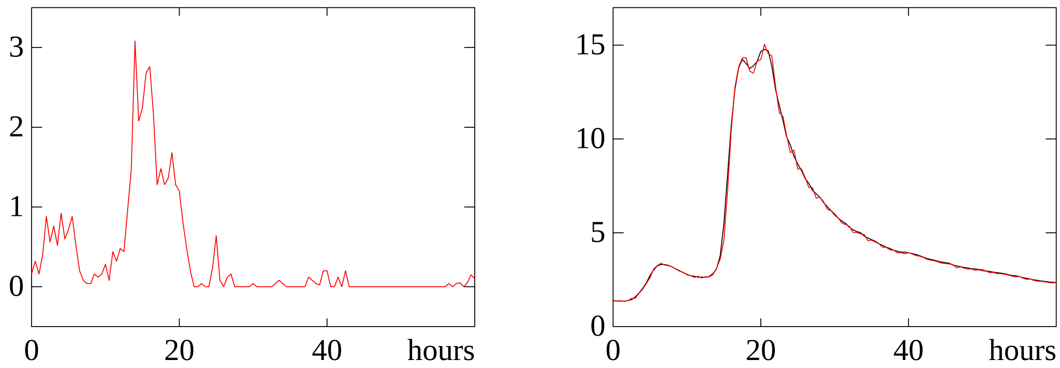
Estimating from this sequence (after subtracting its mean value) a second order autoregressive (AR) model with the least squares algorithm, we obtain the model $y(t+2) = 1.3873\, y(t+1) - 0.6937\, y(t)$ whose poles, $p_{1,2} = 0.6936 \pm i\, 0.4611$, indicate a periodicity of 10.71 years for the phenomenon. 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 (black 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.

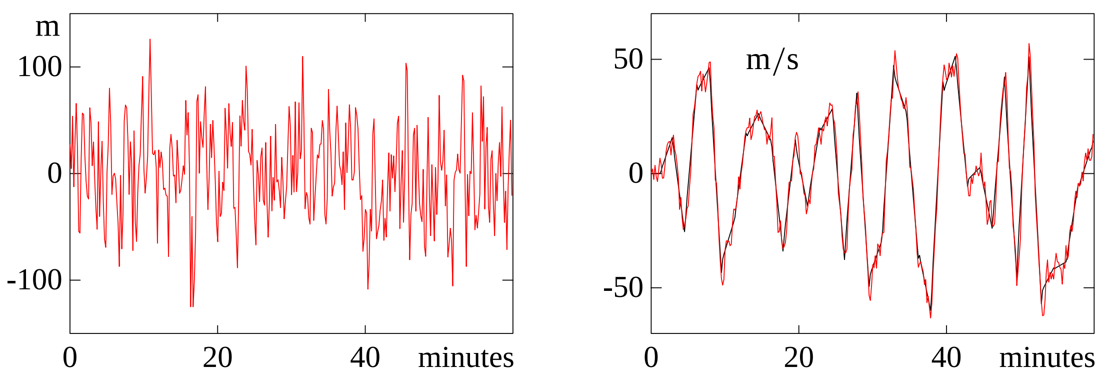
1.2.3 Models for filtering and state estimation

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

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.



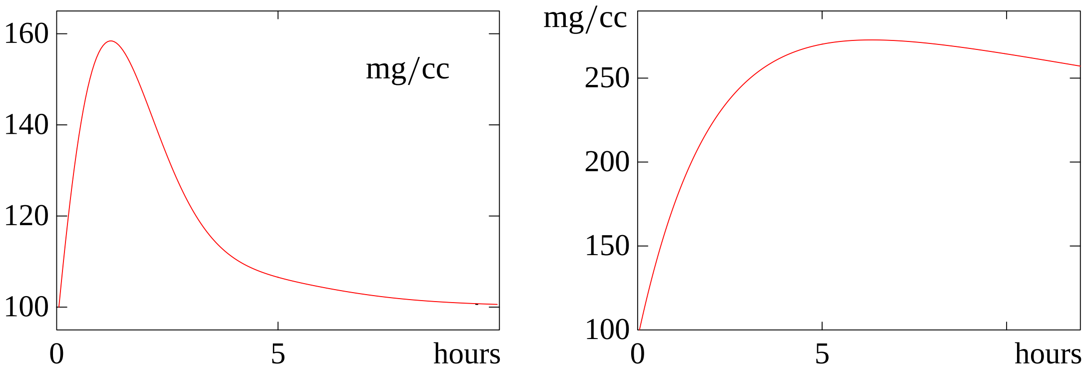
Figures 1.2.4 and 1.2.5 – Altitude error and estimated vertical speed of the target

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. 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.



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

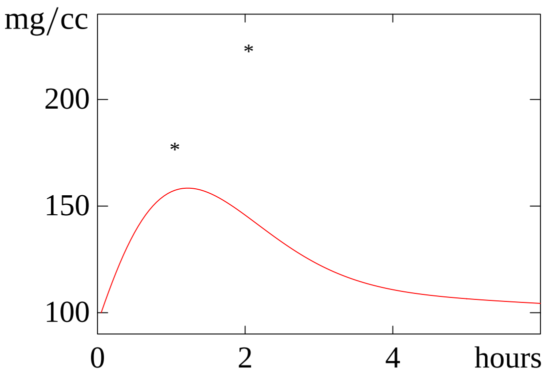


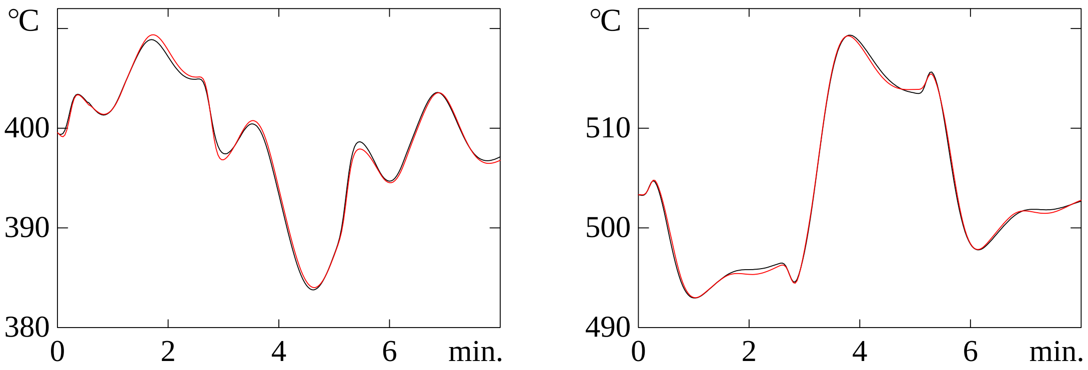
Figure 1.2.8 – Standard response and abnormal measures obtained on a patient

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 (black 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.