Modelado y Simulación

Julián Jiménez Cárdenas

Facultad de Matemáticas
FUKL

August 5, 2024

Introduction to Modeling I

Mathematical modeling is a powerful tool that:

- Translates real-world problems into mathematical language
- Provides insights into complex systems
- Enables predictions and decision-making
- Facilitates optimization and efficiency

Key applications include:

- Physics and engineering
- Economics and finance
- Biology and medicine
- Climate science and ecology

Elaborating in some applications, we have

Introduction to Modeling II

- **Celestial Mechanics**
  - Understanding motion of stars, planets, comets
  - Applications in agriculture, navigation, calendars
  - Modern astrophysics: universe formation and development
- **Energy Supply**
  - Modeling energy consumption trends
  - Planning for future energy needs
- **Everyday Scenarios**
  - E.g., safe driving distances at various speeds
- **Climate Change**
  - Modeling global warming impacts
  - Predicting changes in weather patterns
  - Assessing human impact on climate

Introduction to Modeling III

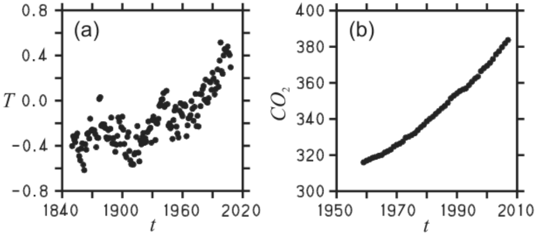

Figure: An illustration of global warming. (a) The HadCRUT3 global temperature anomaly data (consisting of annual differences from 1961-90 normals) are given in °C (Rayner et al, 2003, Brohan et al, 2006); (b) the Mauna Loa data (Tans 2008) of atmospheric CO_2 concentrations are given in ppmv.

Linear Models I

First we introduce the simplest modeling approach: **the linear functions**.

- Suppose we aim to set a relation between two variables, **x** and **y**, and we have paired data from these variables of the form

$$\mathcal{D} = \{(x_i, y_i) \ : \ i = 1, 2, \ldots, n\}.$$

- Thus, we can choose any pair of these points $(x_i, y_i), (x_j, y_j) \in \mathcal{D}$ with $i \neq j$ to find the line that passes through these, which algebraically can be written as

$$y(x) = y_i \frac{x - x_j}{x_i - x_j} + y_j \frac{x - x_i}{x_j - x_i}.$$

- Note that $y(x_i) = y_i$ and $y(x_j) = y_j$, as expected.

Linear Models II

Example (U.S. energy comsumption)

The following table relates the U.S. energy consumption C (in 10^15 Btu) in time t (U.S. Dept, of Energy 2008).

t	C	t	C
1950	34.616	1980	78.122
1955	40.208	1985	76.491
1960	45.087	1990	84.652
1965	54.017	1995	91.173
1970	67.844	2000	98.975
1975	71.999	2005	100.506

- Fit a line passing through the data.
- Graph the error $e = (y - y^{(mod)})/y^{(mod)}$ of the line with respect to the real data. $y^{(mod)}$ is the value predicted by the model.

Exponential Model

Example (Orbital period and distance)

Consider the following data.

Table: Orbital periods and mean distances of planets from the Sun (World Almanac 2010). Here r is the mean distance from the sun in 10^9 km. T_p is the period in earth years $a = 365.25$ days.

Planet	Mercury	Venus	Earth	Mars	Jupiter	Saturn	Uranus	Neptune
r	0.0579	0.1082	0.1496	0.2280	0.7785	1.4335	2.8718	4.4948
T_p	0.2408	0.6152	1.0000	1.8808	11.8618	29.4566	84.0107	164.7858

- Fit a straight line for this data. Is it adequate?
- Assume $T_p = Ar^B$ and use logarithms to transform this relation to a linear one. Fit a straight line. Is it adequate?

Polynomial methods I

The linear approach is useful because of its simplicity, but it has a rather limited range of applicability, as most models cannot be developed in terms of linear functions. Thus, we will consider polynomial models.

- **Linear Polynomials**. There is a unique line passing through a pair $(x_1, y_1), (x_2, y_2)$ of points. Algebraically, let $y = a_0 + a_1 x$ be this line. The requirement of this line passing throught these points imply the following relations

$$y_1 = a_0 + a_1 x_1, \quad y_2 = a_0 + a_1 x_2.$$

This system can be written as

$$\begin{pmatrix} 1 & x_1 \\ 1 & x_2 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix},$$

Polynomial methods II

and this can be inverted (almost always) in order to find a_0, a_1, *a.k.a.*, the line, which is

$$P_1(x) = y_1 \frac{x - x_2}{x_1 - x_2} + y_2 \frac{x - x_1}{x_2 - x_1}.$$

- **Non-Linear Polynomials**. Now, if we have a set $\mathcal{P} = \{(x_, y_0), (x_1, y_1), \ldots, (x_n, y_n)\}$ of $n+1$ points, there is (almost always) a polynomial of degree n passing through these points, and it has the form

$$P_n(x) = y_0 L_0(x) + y_1 L_1(x) + \cdots + y_n L_n(x),$$

with

$$L_k(x) = \prod_{i=0, i\neq k}^{n} \frac{x - x_i}{x_k - x_i}.$$

Polynomial methods III

One of the main advantages of exact polynomial models are that they are easily differentiated and integrated. Nevertheless, they are often not very useful, because slight deviations of the data tend to change the trend of the polynomial.

The solution to this problem usually implies working with reduced polynomial models, like **linear models**.

Polynomial methods IV

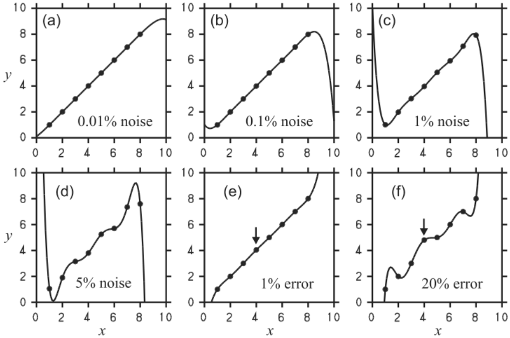

Figure: Exact polynomial models of 7th order for six cases of data points. The data points are given by dots. The polynomial models are given by lines. All the curves pass exactly through the data points. The arrows in (e) and (f) indicate the position of incorrect values.

Polynomial methods V

Consider the development of the world population in time from 1804–2050 according to the Decennial Censuses, U.S. Census Bureau, U.S. Dept. of Commerce (World Almanac 2010). The population P is measured in 10^9 and t refers to the year. The last two population values are projections.

t	1804	1927	1960	1974	1987	1999	2009	2025	2050
P	1.0	2.0	3.0	4.0	5.0	6.0	6.77	7.95	9.32

a) Use the data from 1804 to 2009 to define an exact polynomial of sixth order. Graph this polynomial and the data. Comment on the suitability of this model.

b) Use the data from 1960 to 2009 to define an exact polynomial of fourth order. Graph this polynomial and the data. Comment on the suitability of this model.

Polynomial methods VI

c) Use the data at 1960, 1987, and 2009 to define a polynomial of second order. Graph this polynomial and the data. Comment on the suitability of this model.

d) Use the 1960 and 2009 data to define a polynomial of first order. Graph this polynomial and the data. Comment on the suitability of this model.

The Best Line I

Now that we are again interested in lines, we look for the **best possible line** fitting a set of data $\mathcal{D} = \{(\boldsymbol{x}_i, y_i) \ : \ i = 1, 2, \ldots, n\}$, with $\boldsymbol{x}_i \in \mathbb{R}^m$. Such a line should minimize the error function

$$f = (\boldsymbol{y} - \boldsymbol{y}^{(mod)})^T(\boldsymbol{y} - \boldsymbol{y}^{(mod)}),$$

where $\boldsymbol{y} = (y_1, \ldots, y_n)$ is the vector of the dependent variable, and $\boldsymbol{y}^{(mod)}$ is the vector of predicted values. Now, we take a linear model, so that

$$\boldsymbol{y}^{(mod)} = X\boldsymbol{a},$$

where $X \in M_{n\times(m+1)}$ has in each row the $\boldsymbol{x}_i$'s with a one appended (accounting for the bias term), while $\boldsymbol{a} \in \mathbb{R}^{m+1}$ is the vector of parameters of the model.

The Best Line II

Then, we can show that the best selection of parameters (minimizing f) satisfies

$$a = \left(X^T X\right)^{-1} X^T \boldsymbol{y}.$$