Individual Sections of the Book

Inverse Problems: Exercices

With *mathematica*, *matlab*, and *scilab* solutions

Albert Tarantola
Université de Paris, Institut de Physique du Globe
4, place Jussieu; 75005 Paris; France
E-mail: albert.tarantola@ipgp.jussieu.fr

March 12, 2007

© A. Tarantola, 2006. Students and professors are invited to freely use this text.
2.7 Robust Fitting (Number of Humans and the Logistic Model)

Disclaimer: This is just a methodological exercise, and nothing serious should be inferred from the results shown here.

2.7.1 Data

Using data from the U.S. Census Bureau and from the United Nations, I have found the following estimates of the Human World population, including estimates for the years 2020 and 2050 (values represented in figure 5.7):

<table>
<thead>
<tr>
<th>year</th>
<th>population</th>
</tr>
</thead>
<tbody>
<tr>
<td>1800</td>
<td>$1.0 \times 10^9$</td>
</tr>
<tr>
<td>1850</td>
<td>$1.3 \times 10^9$</td>
</tr>
<tr>
<td>1900</td>
<td>$1.7 \times 10^9$</td>
</tr>
<tr>
<td>1950</td>
<td>$2.5 \times 10^9$</td>
</tr>
<tr>
<td>1970</td>
<td>$3.7 \times 10^9$</td>
</tr>
<tr>
<td>1990</td>
<td>$5.3 \times 10^9$</td>
</tr>
<tr>
<td>2005</td>
<td>$6.5 \times 10^9$</td>
</tr>
<tr>
<td>2020</td>
<td>$7.6 \times 10^9$</td>
</tr>
<tr>
<td>2050</td>
<td>$9.1 \times 10^9$</td>
</tr>
</tbody>
</table>

Figure 2.29: Estimates of human population.

2.7.2 The Verhulst model

In the Logistic function article of Wikipedia I have found the following text.

A typical application of the logistic equation is a common model of population growth, which states that:

- The rate of reproduction is proportional to the existing population, all else being equal.
- The rate of reproduction is proportional to the amount of available resources, all else being equal. Thus the second term models the competition for available resources, which tends to limit the population growth.

Letting $P$ represent population size and $t$ represent time, this model is formalized by the differential equation

$$\frac{dP}{dt} = \tau P(t) \left(1 - \frac{P(t)}{P_1}\right),$$

(2.68)
where the constant $\tau$ defines the growth rate and $P_1$ is the carrying capacity. The general solution to this equation is a logistic function. The solution can be written ($P_0$ being the population at time $t_0$) as

$$P(t) = \frac{P_0 P_1 e^{(t-t_0)/\tau}}{P_0 (e^{(t-t_0)/\tau} - 1) + P_1}, \quad (2.69)$$

and one has

$$\lim_{t \to -\infty} P(t) = 0 \quad ; \quad \lim_{t \to +\infty} P(t) = P_1. \quad (2.70)$$

The Verhulst function is represented in figure 2.30.

Figure 2.30: The Verhulst function.

2.7.3 Fitting the data using the Laplacian model of uncertainties

Bla, bla, bla, and one arrives at the probability density represented in figure 2.31. The two marginal probability densities are represented in figure 2.32.

Figure 2.31: The probability density for the parameters. Note the multimodality.

Figure 2.32: The two marginal probability densities of the joint probability density in figure 2.31.

See results in figure 2.33.
Figure 2.33: Sample of the posterior distribution. Old data don’t adjust to the Verhust model, so, if we wish to believe in that model, we should forget about the older data. In this case, the Verhust model predicts that in year 2200 the Human population will have a value between $10^{10}$ and $1.75 \times 10^{10}$, with a large probability that it already will have attained its maximum.
Although this is not very useful, we can have a look at the two maxima (the global one and the secondary one). Figure 2.34 shows the associated Verhust curves. The one in blue fits well recent data, and discards old data. This suggests that older data don’t adjust to the Verhust model, so, if we wish to believe in that model, we should forget about the older data (the Laplacian model of uncertainties is making that automatically for us). In this case, the Verhust model predicts that the Human population will peak at the value $10.6 \cdot 10^9$, that would be essentially attained around the year 2100.

![Figure 2.34: The blue curve is the best in the least-absolute values sense. It fits well recent data, and discards old data. This suggests that older data don’t adjust to the Verhust model, so, if we wish to believe in that model, we should forget about the older data. In this case, the Verhust model predicts that the Human population will peak at the value $10.6 \cdot 10^9$, that would be essentially attained around the year 2100. The Laplacian model of uncertainties used here makes that the optimum solution fits two of the data points, and leaves as much positive residuals as negative residuals.](image)

2.7.4 Mathematica Code

Executable notebook at http://www.ipgp.jussieu.fr/~tarantola/exercices/chapter_02/PopulationExercice.nb

Note: I have not yet had time to insert these lines of code into the text, and to comment them. For the time being, here is a raw copy of the lines.

```mathematica
(* Defining the Verhult function *)
Verhulst[t_, t0_, tau_, P0_, P1_] := (P0 P1 Exp[(t - t0)/tau])/(P1 + P0(Exp[(t - t0)/tau] - 1))

(* Our special case of Verhulst function *)
V[t_, tau_, P1_] := Verhulst[t, 2005, tau, 6.5 \cdot 10^{-9}, P1]

(* Data values *)
data = {{1800, 1 \cdot 10^{-9}}, {1850, 1.3 \cdot 10^{-9}}, {1900, 1.7 \cdot 10^{-9}}, {1950, 2.5 \cdot 10^{-9}}, {1970, 3.7 \cdot 10^{-9}}, {1990,}}```


5.3 \times 10^{-9}, \{2005, 6.5 \times 10^{-9}\}, \{2020, 7.6 \times 10^{-9}\}, \{2050, 9.1 \times 10^{-9}\} );
\sigma = \{10^{-8}, 10^{-8}, 10^{-8}, 10^{-8}, 10^{-8}, 10^{-8}, 10^{-8}, 10^{-8},
10^{-8}\};
NumberOfData = 9;

(* Prior information *)
fprior[\tau, P1_] := 1

(* Expressing the observations *)
goobs[p1_, p2_, p3_, p4_, p5_, p6_,
p7_, p8_, p9_] := \text{Exp}[-\text{Abs}[p1 - \text{data}[1, 2]]/\text{sigma}[1]] \cdot \text{Exp}[-\text{Abs}[p2 - \text{data}[2, 2]]/\text{sigma}[2]] \cdot \text{Exp}[-\text{Abs}[p3 - \text{data}[3, 2]]/\text{sigma}[3]] \cdot \text{Exp}[-\text{Abs}[p4 - \text{data}[4, 2]]/\text{sigma}[4]] \cdot \text{Exp}[-\text{Abs}[p5 - \text{data}[5, 2]]/\text{sigma}[5]] \cdot \text{Exp}[-\text{Abs}[p6 - \text{data}[6, 2]]/\text{sigma}[6]] \cdot \text{Exp}[-\text{Abs}[p7 - \text{data}[7, 2]]/\text{sigma}[7]] \cdot \text{Exp}[-\text{Abs}[p8 - \text{data}[8, 2]]/\text{sigma}[8]] \cdot \text{Exp}[-\text{Abs}[p9 - \text{data}[9, 2]]/\text{sigma}[9]]

(* Likelihood function *)
L[\tau, P1_] := goobs[V[data[[1, 1]], \tau, P1], V[data[[2, 1]], \tau, P1], V[data[[3, 1]], \tau, P1], V[data[[4, 1]], \tau, P1], V[data[[5, 1]], \tau, P1], V[data[[6, 1]], \tau, P1], V[data[[7, 1]], \tau, P1], V[data[[8, 1]], \tau, P1], V[data[[9, 1]], \tau, P1]]

(* Posterior volumetric probability *)
fpost[\tau, P1_] := fprior[\tau, P1] \cdot L[\tau, P1]

(* Marginal for the parameter P1 *)
margP1[P1_] := NIntegrate[fpost[\tau, P1], {\tau, 28, 90}]

(* Marginal for the parameter \tau (note the factor \text{1}/P1) *)
margtau[\tau_] := NIntegrate[fpost[\tau, P1]/P1, {P1, 9 \times 10^{-9}, 40 \times 10^{-9}]}]

(* Estimating the maximum value of the likelihood function *)
max = L[33.64, 1.061 \times 10^{-10}]

(* Sampling the posterior distribution in the model parameter space *)
SeedRandom[123]
Do[
\{\tau = \text{Random}[\text{Real}, \{28, 70\}],
\text{LogP1} = \text{Random}[\text{Real}, \{\text{Log}[9 \times 10^{-9}], \text{Log}[20 \times 10^{-9}]\}],
P1 = \text{Exp}[\text{LogP1}],
\text{chi} = L[\tau, P1]/\text{max},
\text{phi} = \text{Random}[\text{Real}, \{0, 1\}],
\text{If}[\text{phi} < = \text{chi}, \text{Print}["\tau_0 = \tau", \tau, "\text{, P1}_0 = P1", P1]]
\}, \{3000\}]