Individual Sections of the Book

Inverse Problems: Exercices

With mathematica, matlab, and scilab solutions

Albert Tarantola¹
Université de Paris, Institut de Physique du Globe
4, place Jussieu; 75005 Paris; France
E-mail: albert.tarantola@ipgp.jussieu.fr

March 12, 2007

¹© A. Tarantola, 2006. Students and professors are invited to freely use this text.
5.4 Fitting Curves to Points with Two-Axis Uncertainty Bars

A car has constant acceleration. The initial position $x_0$, the initial velocity $v_0$, and the acceleration $a$ are unknown. We measure the positions $x(t)$ of the car at different times $t$, with measurement uncertainties in both, the times and the positions. Figure 5.7 displays the observations. Uncertainties can be modeled using the Gaussian distribution.

Because the car has uniform acceleration the theory models the positions as

$$x(t) = x_0 + v_0 t + \frac{1}{2} a t^2 . \quad (5.30)$$

Should the “horizontal” uncertainty bars not be present, we would take as data vector and as model parameters vector

$$d = \begin{pmatrix} x(t_1) \\ x(t_2) \\ x(t_3) \\ x(t_4) \\ x(t_5) \\ x(t_6) \end{pmatrix} ; \quad m = \begin{pmatrix} x_0 \\ v_0 \\ a \end{pmatrix} , \quad (5.31)$$

and the theoretical relation between the two,

$$x(t_i) = x_0 + v_0 t_i + \frac{1}{2} a t_i^2 , \quad (5.32)$$

could have been expressed by the linear relation

$$d = G m , \quad (5.33)$$

with the matrix

$$G = \begin{pmatrix} 1 & t_1 & t_1^2 / 2 \\ 1 & t_2 & t_2^2 / 2 \\ 1 & t_3 & t_3^2 / 2 \\ 1 & t_4 & t_4^2 / 2 \\ 1 & t_5 & t_5^2 / 2 \\ 1 & t_6 & t_6^2 / 2 \end{pmatrix} . \quad (5.34)$$
The standard theory of linear least-squares would apply, and we would obtain the solution on one step.

The simplest way to take into account the fact that the times are also uncertain is to consider that they are additional model parameters, using as "a priori values" the measured values, and using a "a priori uncertainties" the experimental uncertainties. In this case, we take as data vector and as model parameters vector

\[
\mathbf{d} = \begin{pmatrix} x(t_1) \\ x(t_2) \\ x(t_3) \\ x(t_4) \\ x(t_5) \\ x(t_6) \end{pmatrix} ; \quad \mathbf{m} = \begin{pmatrix} x_0 \\ v_0 \\ a \\ t_1 \\ t_2 \\ t_3 \\ t_4 \\ t_5 \\ t_6 \end{pmatrix}
\] (5.35)

The theoretical relation between data and model parameters is that in equation 5.32 but this relation is no more linear on our model parameters, as there are products of them (the parameters \(v_0\) and \(a\) appear multiplied by the parameters \(t_i\)). We then are in the general case where the relation can be written

\[
\mathbf{d} = \mathbf{g}(\mathbf{m})
\] (5.36)

and the theory of nonlinear least-squares applies. This, of course, necessitates the evaluation of the partial derivatives

\[
G^i_\alpha = \frac{\partial d^i}{\partial m^\alpha}
\] (5.37)

that are immediate to obtain.

Once the vectors \(\mathbf{d}_{\text{obs}}\) and \(\mathbf{m}_{\text{prior}}\) have been introduced, as well as the two covariance matrices \(\mathbf{C}_{\text{obs}}\) and \(\mathbf{C}_{\text{prior}}\), the a posteriori model can obtained —as explained in class— via the iterative Newton algorithm

\[
\mathbf{m}_{k+1} = \mathbf{m}_k - (G^T_k \mathbf{C}_{\text{obs}}^{-1} G_k + C_{\text{prior}}^{-1})^{-1} (G^T_k \mathbf{C}_{\text{obs}}^{-1} (\mathbf{d}_k - \mathbf{d}_{\text{obs}}) + C_{\text{prior}}^{-1} (\mathbf{m}_k - \mathbf{m}_{\text{prior}}))
\] (5.38)

where \(\mathbf{d}_k = \mathbf{g}(\mathbf{m}_k)\). As shown in the mathematica notebook below, the algorithm (initiated at point \(\mathbf{m}_0 = \mathbf{m}_{\text{prior}}\)) converges in five iterations. Using the posterior values obtained for \(x_0, v_0,\) and \(a\), produces the line displayed in figure 5.8.

Letting \(\mathbf{G}\) be the matrix with the final values of the partial derivatives, the posterior uncertainties are —as explained in class— those represented by the posterior covariance matrix

\[
\mathbf{C}_{\text{post}} = (G^T \mathbf{C}_{\text{obs}}^{-1} \mathbf{G} + C_{\text{prior}}^{-1})^{-1}
\] (5.39)

In the enclosed mathematica notebook, the posterior uncertainties are analyzed in the standard way. One says that the acceleration \(a\) is better resolved than the initial position \(x_0\) and the initial velocity \(v_0\).

As the correlations between the three parameters \(x_0, v_0,\) and \(a\), are strong, I choose to sample the (three-dimensional) posterior marginal for these parameters, obtaining six samples. The corresponding regression lines are displayed in figure 5.9.

Figures 5.10 and 5.11 respectively show the associated velocities and accelerations.
Figure 5.8: The optimal solution corresponds to $x_0 = (1.8 \pm 3.0) \text{ m}$, $v_0 = (0.0 \pm 1.4) \text{ m/s}$, and $a = (0.54 \pm 0.28) \text{ m/s}^2$.

Figure 5.9: Six samples of the posterior solution (i.e., samples of the marginal of the posterior Gaussian for the parameters $x_0$, $v_0$, and $a$).
Figure 5.10: The six velocity functions $v_0 + at$ associated to the six regression lines in figure 5.9.

Figure 5.11: The six accelerations associated to the six regression lines in figure 5.9. Let us focus on some extreme solution, the yellow one, for instance. The initial position (figure 5.9) and velocity (figure 5.10) are low. But, then, the acceleration (figure 5.11) is high. These correlations can be also identified with a look at the posterior covariance matrix (see Mathematica notebook).
5.4.1 Mathematica Notebook

Executable notebook at
http://www.ipgp.jussieu.fr/~tarantola/exercices/chapter_05/TwoAxes.nb

Note: I have not yet had time to clean the code and insert it in the text.