NONLINEAR-MODEL BASED SIGNAL PROCESSING - JUMPS AND COINCIDENCES

Adriaan van den Bos

Department of Applied Physics,
Delft University of Technology,
P.O. Box 5046, 2600 GA Delft,
The Netherlands,
email vdbos@tn.tudelft.nl

ABSTRACT

Different from the least squares criterion for linear models, that for nonlinear models may have relative minima. If, under the influence of the observations, a relative minimum becomes absolute, the solution becomes discontinuous in the observations: it jumps. Also, if under the influence of the observations the criterion becomes singular, its structure may change. Then, for example, solutions for different parameters may coincide exactly. Both phenomena, which have a substantial influence on the solutions, are discussed and explained in a number of numerical examples.

1. INTRODUCTION

Nonlinear model fitting criteria may have a structure essentially different from the well-known single minimum criterion associated with linear least squares. They may have a variety of stationary points: extrema, relative extrema and saddle-points. Among these, the absolute minimum is most important since it represents the solution for the model parameters. With nonlinear models, the absolute minimum may be discontinuous in the observations or the criterion may become singular. The purpose of this paper is to show the practical consequences of discontinuity and singularity for the model fitting solution.

Discontinuity occurs when a relative minimum of the criterion becomes absolute if the observations slightly change. Then the solutions for the parameters jump to new values. This implies that in the Euclidean space of the observations closely located points may correspond to essentially different solutions for the parameters. For example, a minor change in the noise disturbed observations of overlapping pulses of different amplitude, may abruptly change the relative position of the larger and the smaller pulse in the solution. Therefore, jumps may cause large systematic errors in the solutions for the parameters. Jumps correspond to a quantitative change of the criterion only. The structure of the criterion in the sense of its pattern of stationary points does not change.

Singularity causes structural change of the criterion under the influence of the observations. A minimum and a saddle point may merge to form a singular stationary point and subsequently vanish as the observations are changing. Then the structure of the criterion depends on the particular realization of the observations. Singularity may, for instance, cause the solutions for two parameters to coincide exactly. This may occur, for example, if the observations are overlapping time domain pulses, overlapping frequency domain spectral peaks, or multiexponentials with components with little differing decays [1]. Coincidence of location parameters or of decays means that only one pulse, spectral peak or exponential is found instead of two. This means that coincidences like these set a limit to the resolvability of the separate components [2].

The sets of points at which coincidences and jumps occur, divide the space of observations into two distinct parts, each corresponding to a particular type of solution. Jumping may be explained using standard calculus. For coincidence, use is made of singularity theory [3]. Then the set of observations for which the coincidence occurs is a bifurcation set.

In the next section, relevant aspects of the structure of the least squares criterion are discussed in a numerical example. Jumping and coinciding solutions are the subject of the remaining sections.

2. CRITERION STRUCTURE

Suppose that observations $w_1, \ldots, w_N$ made at the points $x_1, \ldots, x_N$ are available and that a model

$$a_1\, g\,(x_n; b_1) + a_2\, g\,(x_n; b_2) \qquad (1)$$

is fitted to these observations with respect to the parameters a_1, a_2, b_1 and b_2. This simple model will be used in what follows. However, the results presented usually also apply if a further function $f(x_n; c)$ is added to it and the resulting model is fitted with respect to the vector of parameters c and a_1, a_2, b_1 and b_2 . For example, $g\,(x; b_k)$ in Eq.(1) may be an exponential $e^{-b_k x}$. The function $f(x; c)$ may be a trend $c_1\, x + c_2$ and possibly contain further exponential terms. Alternatively, the functions $g\,(x; b_k)$ may be Gaussian peaks $e^{-\frac{1}{2}(x-b_k)^2}$ with location parameters b_k. The results presented usually also apply to models that are functions of the function described by Eq.(1).

Next suppose that the model defined by Eq.(1) is fitted in the least squares sense. Then, the criterion is described by

$$\sum_n \{w_n - a_1\, g\,(x_n; b_1) - a_2\, g\,(x_n; b_2)\}^2 \qquad (2)$$

This is equivalent to fitting $l\, a\, e^{-b_1\, x_n} + (1-l)\; a\, e^{-b_2\, x_n}$ with respect to a, b_1 and b_2 for all relevant values of l. This parameterization will be used since it is more suitable for the purposes of this paper. Notice that l fixes the ratio of a_1 to a_2. The criterion is quadratic in the linear parameter a. Its stationary points are all found in the parameter subspace where the derivative of the criterion with respect to a is equal to zero. If this linear equation is solved for a in terms of b_1 and b_2 and the solution is substituted for a, the criterion becomes a function of b_1 and b_2 only, suitable for plotting. This plot is parametric in l. To understand the structure of this criterion first suppose that the observations w_n are described exactly by $\lambda\, \alpha\, g\,(x_n; \beta_1) + (1-\lambda)\, \alpha\, g\,(x_n; \beta_2)$ and that the model fitted to it is of the same parametric family. For example, both the observations and the model fitted are biexponential.

Numerical example 1 Let the observations be described by $0.7\, e^{-x_n} + 0.3\, e^{-0.8 x_n}$, that is, $\alpha = 1, \lambda = 0.7$, $\beta_1 = 1$ and $\beta_2 = 0.8$ and let $x_n = 0.4 \times n$ with $n = 1, \ldots, 10$. Suppose that a model $l\, a\, e^{-b_1\, x_n} + (1-l)\; a\, e^{-b_2\, x_n}$ is fitted with respect to a, b_1 and b_2 to these exact observations for $l = 0.6$. Notice that this l is different from λ. Also suppose that the linear parameter a has been eliminated as described above. Then contours of the criterion are shown in Fig. 1. In this figure, the vertical and horizontal coordinate are $\lambda\, b_1 + (1-\lambda)\, b_2$ and $b_1 - b_2$, respectively. The reason why these coordinates have been chosen is that they expose the structure clearer. The plot, which is slightly asymmetrical, has two minima and a saddle point in between. This structure is characteristic of this problem for a broad class of component functions, criteria of goodness of fit and for all l on $(0, 1)$. The criterion value at the absolute minimum $(b_1, b_2) = (1.0178, 0.8242)$ is 7.96×10^{-10}; that at the relative minimum $(b_1, b_2) = (0.8586, 1.0651)$ is 6.67×10^{-9}. The criterion value at the saddle point $(b_1, b_2) = (0.9287, 0.9287)$ is 4.22×10^{-6}. This point is always located on the line $b_1 = b_2$ and can be shown to represent the least squares solution if a one component model, in this example a monoexponential, is fitted to the same observations.

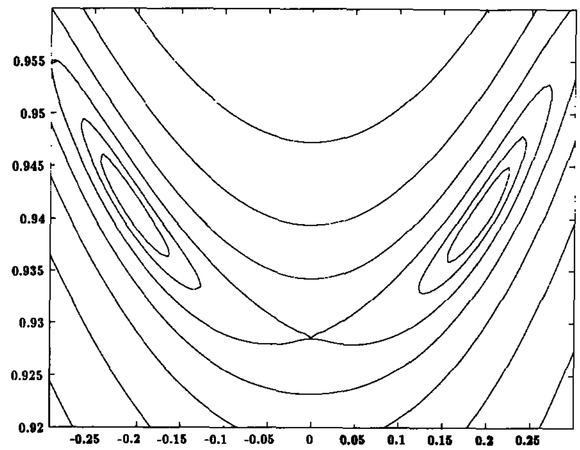

Figure 1

3. COINCIDENCES

The two minima one saddle structure found in Numerical example 1 occurred for observations exactly describable by the model fitted. It has been found that this structure may change if the observations are no longer exact as a result of additive errors or model errors [1].

Numerical example 2 The exact observations of Numerical example 1 are transformed into those of this example by adding the errors $v_2 = 0.01$ and $v_4 = 0.01$ to the second and fourth observation, respectively. The quantity l is kept equal to 0.6. Figure 2 shows the contours of the criterion for these new observations. The two minima one saddle structure has been replaced by a single minimum structure. This single minimum is located on $b_1 = b_2$ and has replaced the saddle point. Since at the minimum b_1 and b_2 coincide, a monoexponential solution is obtained from error corrupted biexponential observations.

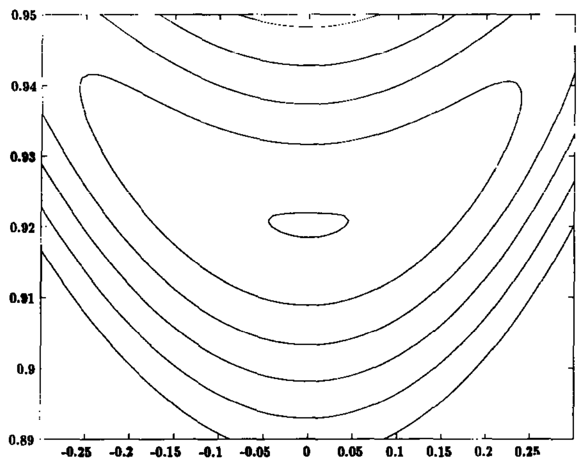

Figure 2

This numerical example shows that the structure of the criterion may be changed by the observations and that this causes the least squares solutions for b_1 and b_2 to coincide. This phenomenon has been discussed extensively in [1] where a singularity theoretic explanation is given. In this reference, it is also shown that in the Euclidean space of the observations $(w_1, ..., w_N)$, observations in the region corresponding to the one structure are separated from observations in the region corresponding to the other by a so-called hypersurface which is a subset of codimension one. Hypersurfaces have the property that they divide the space concerned into two distinct parts such as a line divides a plane. In singularity theory, this hypersurface is called bifurcation set. If the observations of Numerical example 2 are gradually changed and cross the bifurcation set, the corresponding solution for the parameters bifurcates. In [1], a discriminant for on which side of the bifurcation set a particular set of observations is located for a chosen model is also described. Below, the intersection of the bifurcation set and the (w_2, w_4) coordinate plane will be shown in Numerical example 4.

4. JUMPS

Different from coinciding solutions, jump discontinuities in the solutions for the parameters as a function of the observations do not require the structure of the criterion to change. Using Numerical example 1, the following heuristic explanation for their occurrence may be given. As mentioned in Section 2, the criterion value at the absolute minimum $(b_1, b_2) = (1.0178, 0.8242)$ is 7.96×10^{-10} and that at the relative minimum $(b_1, b_2) = (0.8586, 1.0651)$ is 6.67×10^{-9}. Intuitively, one may conclude that errors in the order of magnitude 10^{-4} might be sufficient to make the relative minimum absolute. It will be shown in the following numerical example that this is true.

Numerical example 3 The exact observations of Numerical Example 1 are modified by adding an error v_2 to the second observation. For a number of values of v_2 the relative and absolute minimum of the criterion are computed until the point on the v_2 axis is found where the relative minimum becomes absolute and, as a result, the solutions for a, b_1 and b_2 make a jump. Next the same is done with respect to v_4. The straight line connecting both points is subsequently found to accurately predict the linear combinations of v_2 and v_4 for which the jump occurs. This line is shown in Fig. 3. The figure shows that the order of magnitude of the errors causing the solution to jump from a solution with $b_1 > b_2$ to one with $b_1 < b_2$ is 10^{-4}.

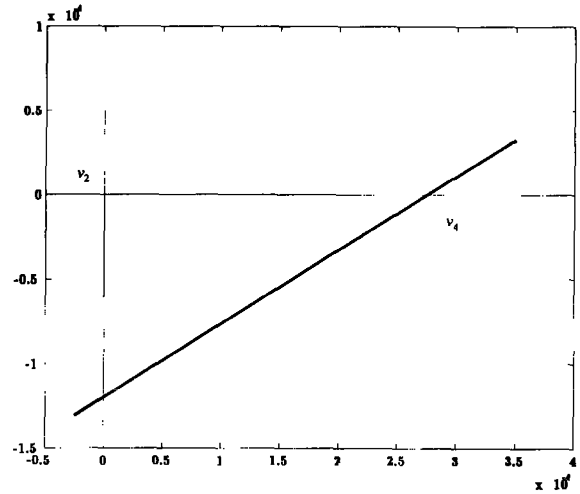

Figure 3

The coordinates a_1, b_{11} and b_{21} of the absolute minimum have to satisfy three normal equations. These are the nonlinear equations obtained by equating the first order partial derivatives of the criterion with respect to a, b_1 and b_2 to zero. The same applies to the coordinates a_2, b_{12} and b_{22} of the relative minimum. At the jump set, the criterion values at the absolute and the relative minimum must be equal which produces one further equation. Thus the $N + 6$ variables $a_1, b_{11}, b_{21}, a_2, b_{12}, b_{22}, w_1, ..., w_N$ have to satisfy 7 equations. From these 7 equations the first six variables could, hypothetically, be eliminated. This produces one equation in the N observations $w_1, ..., w_N$. This equation defines the jump set, the set of all sets of observations for which both minima are equivalent. Having codimension one it divides the Euclidean space of the observations in two complementary parts. The observations in the one part correspond to solutions with

$b_1 > b_2$ and those in the other to solutions with $b_1 < b_2$. Figure 3 shows the intersection of the jump set with the coordinate plane (v_2, v_4) which is equivalent to the coordinate plane (w_2, w_4) after translation.

5. JUMPS AND COINCIDENCES COMBINED

In this section, a numerical example is discussed showing that the jump set and the bifurcation set divide the Euclidean space of the observations into three distinct parts. Only in one of these parts the solutions for the parameters qualitatively agree with those of the model underlying the observations.

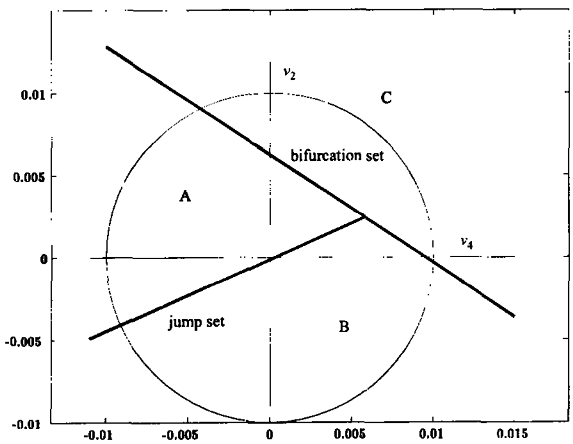

Figure 4

Numerical example 4 As in Numerical example 3, errors are added to the second and the fourth observation of the set of exact observations used in Numerical example 1. These errors are v_2 and v_4, respectively. Next the bifurcation set is computed expressed in these errors. For this computation used is made of a method described in [1]. Figure 4 shows this bifurcation set along with the jump set described in the previous section and shown around the origin in Fig. 3. The jump set stops at the bifurcation set since beyond the bifurcation set there is only one minimum and jumps cannot occur. In Fig. 4 observations disturbed by errors in the region A correspond qualitatively to exact observations since the solution b_1 for β_1 is larger than the solution b_2 for β_2. In the regions B and C the solutions satisfy $b_1 < b_2$ and $b_1 = b_2$, respectively. Next consider errors v_2 and v_4 on the circle $v_2^2 + v_4^2 = (0.01)^2$ drawn in Fig.4. Then starting at the topmost point of intersection with the bifurcation set, travelling anticlockwise along the circle and computing the solutions b_1 and b_2, first produces a bifurcation, then, upon crossing the jump set, a jump and then again a bifurcation. Figure 5 shows the corresponding solutions as a function of the rotation angle ψ. Since $\beta_1 = 1$ and $\beta_2 = 0.8$, the main conclusion to be drawn from the solutions shown is that seemingly small errors may cause relatively large errors of a perhaps unexpected nature.

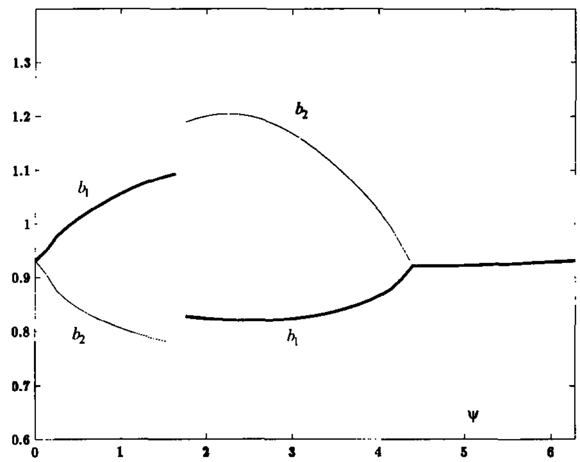

Figure 5

6. CONCLUSIONS

Jumps and coincidences in nonlinear model fittiing solutions have been discussed. They are caused by the nonlinearity of the model and are not revealed if the models are linearized as in asymptotical statistical model fitting theory [4]. It has also been shown that jumps and coincidences may have a considerable influence on the solutions and may be caused by seemingly small errors.

7. REFERENCES

[1] A. van den Bos and J.H. Swarte, "Resolvability of the parameters of multiexponentials and other sum models," *IEEE Trans. Signal Processing*, vol. 41, no. 1, pp. 313-322, 1993.

[2] A.J. den Dekker and A. van den Bos, "Resolution - A survey," *J. Opt. Soc. Am.*, vol. A 14, no. 3, pp. 547-557, 1997.

[3] V.I. Arnol'd, *Catastrophe Theory.* Berlin: Springer, 1992.

[4] R.I. Jennrich, *An Introduction to Computational Statistics.* Englewood Cliffs: Prentice-Hall, 1995.