AN APPROACH TO FUZZY NONLINEAR REGRESSION ANALYSIS

George SOTIROV

Systems Optimization Laboratory

Faculty of Automatics

Technical University, Sofia 1156, Bulgaria

ABSTRACT. In this preliminary study the fuzzy set theory is used as a framework for representation and manipulation of uncertain information in regression analysis. An approach for obtaining a nonlinear interval α-model which fits to given fuzzy output data is proposed.

1. INTRODUCTION

Regression analysis has been widely recognized in the last years as one of the most relevant area in system identification and its applications. Techniques for identification of linear and nonlinear models in the presence of noise are well established. The conventional approach [3] has some disadvantages despite its popularity. Statistical properties of the noise are commonly assumed to be known in advance or determinable from observations. A classical technique is maximum-likelihood estimation, where the parameter estimates are computed by maximizing a criterion deduced from the probability density function of the noise corrupting the data, assumed to be known or parametrized. Such statistical assumptions are questionable when information is limited or model structure and hence the modelling error are in doubt.

This paper addresses such a situation, where the data y are supposed to be described by a deterministic parametric model F corrupted by an additive noise e, so that the measured vector y $= (y_1, \ldots, y_n)^T$ satisfies

$$y = F(\theta) + e \qquad (1)$$

where $\theta = (\theta_1, \ldots, \theta_m)^T$ is the true value of the parameter vector to be estimated and $e = (e_1, \ldots, e_n)^T$ represents the unknown n-dimensional observation error. To perform the parameter estimation it is necessary that the structure (1) is identifiable and that an error model is available.

In recent years different error models have been proposed [1], [3], [5], [11], referring to stochastic, interval and fuzzy models. Tanaka, Uejima and Asai [6] have developed fuzzy regression analysis based on fuzzy set theory. Regarding the deviation between the data as a reflection of fuzziness on system parameters, the fuzzy structure is presented as a linear function with fuzzy parameters. In [7] a possibilistic interpretation of a fuzzy regression model is given and discussed in detail in references [8], [9]. The possibilistic linear regression analysis is formulated by fuzzy linear function as a model of possibility structure of the system we wish to model. The fuzzy linear function is defined from possibility distributions of the parameters based on extension principle [1] which can also be explained in terms of possibility measure.

In this contribution we assume that deviations between the actual values and the computed values of the dependent variable in nonlinear fuzzy regression (NLFR) are related to the fuzzy observation errors as opposed to being related to the soft structure of the system under consideration. Different algorithms for obtaining an interval α-model wich fits to given fuzzy output data are developed.

2. PROBLEM STATEMENT

The basic model considered here is the general multiple nonlinear regression model (1). Our basic assumption is that the residual or deviation of the estimated from (1) value $\hat{y}_i$ and observed value y_i, i.e. $e_i = \hat{y}_i - y_i$, $i = 1, \ldots, n$, where n is the number of samples, is related to observation error which is fuzzy variable. Therefore, the problem is to estimate the parameter vector θ given the input-output data (x_i, y_i) disturbed by fuzzy errors e_i, where x_i is nonfuzzy input vector and y_i is a fuzzy output. A fuzzy number e_i is represented by a fuzzy set defined on real field $R = [-\infty, +\infty]$ with membership function $\mu_{ei}(x)$ as the degree of belongingness of the element x to this set. At each sample the observation error is considered to be a fuzzy number having concave membership function with compact support. Given a particular α-level of all $\mu_{ei}(x)$, $0 \le \alpha \le 1$, all the deviations e_i become closed and bounded intervals

$e_i(\alpha) = [e_i^-(\alpha), e_i^+(\alpha)]$ for each α-level. Thus, we suggest to discretize their membership values by different $\alpha \epsilon [0,1]$ values, i.e. we use α-cut to get a nonfuzzy and nonstatistical description of the deviation in the form of bounds on its instantaneous values.

The generic NLFR model used herein consists of following bounded-error model, one for each α-level of corresponding membership function μ_e:

$$y_\alpha(i) = F(x_i, \theta) + e_i(\alpha), \; i = 1, \ldots, n \qquad (2)$$

For notational simplisity $y_\alpha(i)$ is assumed to be a scalar, but the results to be presented can be extended without eny difficulty to the multivariable case. Our study of NLFR is based on the following hypotheses:

1. The system is of the form (1).
2. The empirical input data are error free and that of the output are bounded at α-level by intervals, i.e. the set of observations

   $x_i = (x_{1i}, \ldots, x_{pi})^T$, $[y_\alpha^-(i), y_\alpha^+(i)]$ $\quad i = 1, \ldots, n$

   on the input and output variables are available. Note that $y_\alpha^-(i) = y_\alpha(i) - e_i^+(\alpha)$ and $y_\alpha^-(i) = y_\alpha(i) - e_i^-(\alpha)$.
3. A model $y = F(\theta)$ is acceptable at α-level if and only if the noise free output is consistent with the measurement $y(i)$ and bound over the observation error $[e_i^-(\alpha), e_i^+(\alpha)]$.

Estimating θ with hypotheses 1-3 then consists of the determination of a feasible parameter set at α-level.

3. PARAMETER ESTIMATION

The approach proposed here considers that the parameter vector belongs to a set $D_\alpha \epsilon R^m$ (where m is the number of parameters) which has to be characterized from the available information on the data and noise. Such estimators thus do not yield directly a single estimate in terms of a priori specified membership function of the parameter vector. We are looking for the set D_α of all admissible values of θ that are consistent with (1) and assumptions on the fuzzy error. D_α is therefore the set of solutions for θ of the 2n inequalities

$$y_\alpha^-(i) \le F(x_i,\theta) \le y_\alpha^+(i),\ i = 1, \ldots, n \qquad (3)$$

and also can be defined as the intersection of n hypersurfaces. Note that in general D_α is not necessarily convex and may not even be connected. This may result from the fact that the model is not uniquely identifiable. D_α may be empty (if the model and/or hypotheses on the error bounds are erroneous), but is usually not a singleton.

Obviously, D_α may assume a very complicated shape. For this reason here we will consider only a simplified approximation for D_α in form of a minimal outer box I_α bounding the set (3). To this extent the parameter uncertainty interval $I_{\alpha j}$ is determined as:

$$I_{\alpha j} = [\theta_{\alpha j}^-,\ \theta_{\alpha j}^+],\ j = 1, \ldots, m \qquad (4)$$

where

$$\theta_{\alpha j}^- = \min_{\theta_j \in D_\alpha} \theta_j,\ \theta_{\alpha j}^+ = \max_{\theta_j \in D_\alpha} \theta_j \qquad (5)$$

Finally, the det I_α is defined as the cartesian product of uncertainty intervals for each component of parameter vector:

$$I_\alpha = I_{\alpha 1} \times I_{\alpha 2} \times \ldots \times I_{\alpha m} \qquad (6)$$

Some basic properties of D_α, $I_{\alpha j}$ and I_α can be summarized as follows:

1. D_α is a set that in general may be non-convex and non-connected. It can also consist of infinite non-connected subsets.
2. $I_{\alpha j}$ are intervals that can be bounded or unbounded, not all the values within the interval can always be taken by the parameter. If so, sub-interval can be defined, any value of wich can be taken by the parameter.
3. $\theta_{\alpha j}^-$ and $\theta_{\alpha j}^+$ may not be unique, moreover local max and min values in the parameter space are possible.
4. Outer box I_α may not exist when the set D_α is unbounded.

It should be noted that $I_{\alpha j}$, $j = 1, \ldots, m$ may be considered as an α-cut of the fuzzy parameter θ_j for different $\alpha \in [0, 1]$. Therefore, some available methods can be used to identify the membership functions $\mu_{\theta j}$ from such data.

In the following section we consider briefly some algorithms for the computation of parameter uncertainty intervals for certain classes of nonlinear problems.

4. ALGORITHMS

Four different classes of algorithms will be considered to perform the computation of parameter uncertainty intervals.

The first one mainly refer to linearization procedures that can be applied whenever relation (1) is continuous and differentiable in the parameters θ. One may thing of linearizing the model around some value of the parameters estimated beforehand and then using any method for linear models [4]. Unfortunately, it seems impossible to evaluate the degree of approximation incurred and to know whether the bounds obtained contain D_α or not.

A second classe of algorithms consists in searching the boundary points of D_α in randomly selected directions, starting from a point within D_α [10]. In such a way it is possible to obtain a cloud of points belonging to the boundary of D_α. One may, however, argue that these boundary points are not easy to use for further investigations.

The third classe includes techniques for scanning the parametric space. The easiest one consists in seting the minimum amd maximum values of each parameter θ_j over D_α (resp. θ_{jmin}, θ_{jmax}). Note, that $\theta° = [\theta_1] \times \ldots \times [\theta_m]$ inclu-des D_α. N points are picked at random within $\theta°$ according to a uniform distribution. Then from N_D points belonging to D_α parameter uncertainty intervals can easily be computed.

Finally the problem can be tackled using constrained nonlinear programming algorithms. Obviously these algorithms cannot quarantee in general the convergence to the global minimum (maximum).

Thus, a general solution to identification of nonlinear fuzzy regression model using α-cuts is probably not possible since nonlinear problems can present so many different and not homogenous characteristic. It should be noted, that for some restricted but important classes of nonlinear systems some stronger results and algorithms can be obtained. For instance, if $F(\theta)$ is polinomial function in θ, optimization problems (5) are signomial ones. Though signomial problem is in general not convex, an iterative algorithm can be designed, which is able to evaluate at Ñach iteration lower and upper bounds of the global extremum. Moreover, the sequences of lower and upper bounds are guaranteed to converge monotonically to the global solution [2]. Note, that the polynomial hypothesis covers large class of problems of practical interest such as, for example, the identification of fuzzy multiexponential, ARMA and state spase discrete time models.

5. CONCLUSIONS

In this study attention has been paid to the problem of estimating the parameters of nonlinear regression model together with their uncertainties in the presence of fuzzy error. A nonfuzzy description of the error in the form of bounds on its instantaneous values for each α-level has been used in order to characterize the feasible parameter set. Several classes of algorithms for the computation of parameter uncertainty intervals have been discussed.

REFERENCES

1. D.Dubois and H.Prade, Fuzzy Sets and Systems: Theory and Applications (Academic Press, N.Y., 1980).
2. J.G.Ecker, Geometric programming: Methods, computations and applications, SIAM Rev. 1 (1980) 339-362.
3. J.P.Norton, An Introduction to Identification (Academic Press, London, 1986).
4. G.Sotirov, Model building in decision support systems for optimization problems under uncertainty, Proc. NAFIPS'90, Toronto, II (1990) 398-401.
5. G.Sotirov, Control of fuzzy systems based on α-ARMAX models, Proc. Intrn. Conf. Fuzzy Logic&Neural Networks, Iizuka, Japan (1990) 883-886.

6. H.Tanaka, S.Uejima and K.Asai, Linear regression analysis with fuzzy model, IEEE Trans. SMC-12 (1982), 903-907.

7. H.Tanaka, Fuzzy data analysis by possibilistic linear models, FSS 24(1987) 363-375.

8. H.Tanaka and J.Watada, Possibilistic linear systems and their application to the linear regression model, FSS 27(1988) 275-289.

9. H.Tanaka, I.Hayashi and J.Watada, Possibilistic linear regression analysis for fuzzy data, EJOR 40 (1989) 389-396.

10. E.Walter, H.Lahanier and J.Happel, Estimation of non uniquely identifiable parameters via exaustive modelling and membership set theory, Proc.1st IMACS Cong.Modelling and Sim. for Control of Lumped and Distributed Parameter Systems, Lille (1986) 313-317.

11. E.Walter and H.Piet-Lahanier, Estimation of parameter bounds from bounded-error data: a survey, Math. Comp. in Simulation 32 (1990) 449-468.