KYBERNETIKA — VOLUME 28 (1992), NUMBER 6, PAGES 444–453

A CLASSIFICATION OF NONLINEAR REGRESSION MODELS AND PARAMETER CONFIDENCE REGIONS[1]

ANDREJ PÁZMAN

This is mainly a survey paper with a new look on comparison of confidence regions. It is divided into two parts. In the first part we compare classes of nonlinear regression models having some linear-like property (intrinsically linear models, models with constant information matrices, models with zero Riemannian curvature). In the second part we discuss four kinds of regions as candidate for being confidence regions for parameters: the elliptical region, the likelihood region, the linear inference region, and finally a confidence region proposed recently by the author.

1. INTRODUCTION

In this paper we consider a nonlinear regression model

$$\begin{aligned} y &= \eta(\vartheta) + \varepsilon; \qquad (\vartheta \in \Theta) \\ \varepsilon &\sim N(0, \sigma^2 W) \end{aligned} \tag{1.1}$$

with the observed vector $y \in \mathbb{R}^N$, the vector of unknown parameters $\vartheta \in \Theta \subseteq \mathbb{R}^m$, $m < N$. The matrix W is known, and will be supposed to be regular. It will be supposed that the boundary points of the (known) parameter space Θ are accessible as limit points of $\operatorname{int}\Theta$, in symbols $\Theta \subseteq \overline{\operatorname{int}\Theta}$. (For example, $\Theta = [a_1, b_1] \times \ldots \times [a_m, b_m]$). The (known) mapping $\vartheta \in \Theta \longrightarrow \eta(\vartheta) \in \mathbb{R}^N$ is supposed to be continuous, with continuous third order derivatives on $\operatorname{int}\Theta$ and with rank $[\partial\eta(\vartheta) / \partial\vartheta^{\mathrm{T}}] = m$ on $\operatorname{int}\Theta$.

A particular case of (1.1) is the linear model, with $\eta(\vartheta)$ being linear in ϑ. Statistical methods are much better elaborated in linear models than in nonlinear ones where such methods are only in development. One may say that the closer is the investigated model to the linear model, the better are the methods of statistical inference. Therefore our aim is to present different classes of models which have some properties of linear models (they are "flat" in some sense) and to show some consequences of this flatness on constructions of confidence regions for ϑ.

[1]Presented at PROBASTAT'91, Bratislava, August 26–30, 1991.

2. CLASSES OF NONLINEAR MODELS

We shall consider four classes of nonlinear regression models, which have some features in common with linear models. They are labelled by (A) - (D) in Figure 1 below, where also the relations of inclusions between the classes are indicated, i. e. class (A) $\subseteq$ class (B) $\subseteq$ class (D), class (A) $\subseteq$ class (C) $\subseteq$ class (D).

(A) - linear
$\partial^2\eta(\vartheta) \,/\, \partial\vartheta_i\partial\vartheta_j = 0$

$\cap$ $\qquad$ $\cap$

(B) - const. inform. m.
$P(\vartheta)\partial^2\eta(\vartheta) \,/\, \partial\vartheta_i\partial\vartheta_j = 0$

(C) - intrinsic. linear
$(I - P(\vartheta))\,\partial^2\eta(\vartheta) \,/\, \partial\vartheta_i\partial\vartheta_j = 0$

$\cap$ $\qquad$ $\cap$

(D) - zero Riemannian curv.

$\cap$

model (1.1)

Fig. 1.

The classes (A) - (D) can be characterized by different ways:
- by the geometry of the expectation surface

$$\mathcal{E} := \{\eta(\vartheta) : \vartheta \in \Theta\}$$

  of model (1.1)
- by some curvature measures of nonlinearity
- by some relations between the first and the second order derivatives of $\eta(\vartheta)$
- by the existence of some "linearizing reparameterizations".

We present these characterizations in details, starting by the class (C).

Class (C) - intrinsically linear models

This class can be defined geometrically as the one having a planar expectation surface $\mathcal{E}$. That is, $\mathcal{E}$ is a subset of an m-dimensional plane in $\mathbb{R}^N$. We have: *model (1.1) is intrinsically linear iff there exists a (regular) reparameterization $\beta = \beta(\vartheta)$ such that the new model is parameterized linearly by β.* Hence in many sense such model is close to a linear model with constraints on parameters. We have the following statements:

Proposition 2.1. Model (1.1) is intrinsically linear iff for every $\vartheta \in \operatorname{int}\Theta$ and for every i, j the equality

$$(I - P(\vartheta))\left[\partial^2\eta(\vartheta) / \partial\vartheta_i\partial\vartheta_j\right] = 0$$

holds. Here $P(\vartheta)$ is the matrix

$$P(\vartheta) := \frac{\partial\eta(\vartheta)}{\partial\vartheta^{\mathrm{T}}} M^{-1}(\vartheta)\frac{\partial\eta^{\mathrm{T}}(\vartheta)}{\partial\vartheta} W^{-1} \tag{2.1}$$

(the orthogonal projector onto the tangent plane to $\mathcal{E}$), and $M(\vartheta)$ is

$$M(\vartheta) := \frac{\partial\eta^{\mathrm{T}}(\vartheta)}{\partial\vartheta} W^{-1}\frac{\partial\eta(\vartheta)}{\partial\vartheta^{\mathrm{T}}}$$

(the Fisher information matrix for $\sigma = 1$).

The proposition can be found without proof and in a modified form in [5]. A proof of this proposition, as well as of the other propositions, should be presented in [14].

Proposition 2.1 gives the characterization of class (C) by a relation between first and second order derivatives of $\eta(\vartheta)$, since $P(\vartheta)$ is expressed in terms of the first order derivatives.

Proposition 2.2. Model (1.1) is intrinsically linear iff the intrinsic curvature of model (1.1)

$$K_{\mathrm{int}}(\vartheta) := \sup_{v\in\mathbb{R}^m\setminus\{0\}} \left\| v^{\mathrm{T}}(I - P(\vartheta))\frac{\partial^2\eta(\vartheta)}{\partial\vartheta\,\partial\vartheta^{\mathrm{T}}}v\right\|_W \Big/ \left(v^{\mathrm{T}}M(\vartheta)\,v\right)$$

is identically equal to zero on $\operatorname{int}\Theta$.

The curvature $K_{\mathrm{int}}(\vartheta)$ is a curvature in a usual geometric sense, and it has been proposed as a measure of nonlinearity by Bates and Watts [1]. The used expression for $K_{\mathrm{int}}(\vartheta)$ is derived in [9].

Class (B) – models with a constant information matrix

Such models can be defined as models having a constant (Fisher) information matrix

$$M(\vartheta) = M = \text{ const }; \qquad (\vartheta \in \Theta)$$

Geometrically it means that the parameter space Θ is a part of an m-dimensional Euclidean space having the inner product

$$a^{\mathrm{T}} M^{-1} b; \qquad (a, b \in \mathbb{R}^m)$$

defined by this information matrix. The expression $\sigma^{-1}\left[(\vartheta - \bar{\vartheta})^{\mathrm{T}}M^{-1}(\vartheta - \bar{\vartheta})\right]^{1/2}$ is well known in statistics as the Mahanolobis distance between ϑ and $\bar{\vartheta}$.

Since $M^{-1}(\vartheta)$ is the asymptotic variance matrix of the least squares estimator of ϑ, models of class (B) are studied also as "stabilized variance models" (cf. Kass [10] and Hougaard [8], where also Proposition 2.4 is presented; cf. also Pázman [11]).

Proposition 2.3. Model (1.1) is with a constant infomation matrix iff its parameter effect curvature

$$K_{\mathrm{par}}(\vartheta) := \sup_{v \in \mathbb{R}^m \setminus \{0\}} \left\| v^T P(\vartheta) \frac{\partial^2 \eta(\vartheta)}{\partial \vartheta \, \partial \vartheta^{\mathrm{T}}} v \right\|_W / \, (v^{\mathrm{T}} M(\vartheta) \, v)$$

is identically equal to zero.

This gives a characterization of class (B) in terms of a curvature measure of nonlinearity. The parameter effect curvature $K_{\mathrm{par}}(\vartheta)$ has been introduced by Bates and Watts [1] as a complementary measure to $K_{\mathrm{int}}(\vartheta)$.

Proposition 2.4. Model (1.1) is with a constant information matrix iff for every $\vartheta \in \operatorname{int} \Theta$, and for every i, j, k the equality

$$\frac{\partial \eta^{\mathrm{T}}(\vartheta)}{\partial \vartheta_k} W^{-1} \frac{\partial^2 \eta(\vartheta)}{\partial \vartheta_i \, \partial \vartheta_j} = 0$$

holds.

This is a characterization of class (B) by a relation between the first and the second order derivatives of $\eta(\vartheta)$.

The proofs of Propositions 2.3 and 2.4 are obtained essentially from $\partial M(\vartheta) \, / \, \partial \vartheta = 0$.

Class (A) – linear regression models

Here belong models having a linear $\eta(\vartheta)$

$$\eta(\vartheta) = F\vartheta + f$$

F and f being a known matrix and a known vector. Geometrically, the expectation surface $\mathcal{E}$ is an m-dimensional plane in $\mathbb{R}^N$, and at the same time, the parameter space Θ is an m-dimensional Euclidean space with the Mahanolobis distance as the norm in Θ. So we have: class (A) $\subseteq$ class (B) $\cap$ class (C). We have yet more, as given by the following proposition.

Proposition 2.5. The equality

$$\text{class (A)} = \text{class (B)} \cap \text{class (C)}$$

holds. Model (1.1) is linear iff

$$K_{\mathrm{int}}(\vartheta) = 0, \qquad K_{\mathrm{par}}(\vartheta) = 0$$

identically on int Θ.

Proof. From Propositions 2.1 and 2.4 we obtain that in a model from class (B) $\cap$ class (C) one has

$$\frac{\partial^2 \eta(\vartheta)}{\partial \vartheta_i \, \partial \vartheta_j} = 0$$

for every $\vartheta \in \operatorname{int} \Theta$ and every i, j, hence $\eta(\vartheta)$ is linear in ϑ. The last statement of Proposition 2.5 follows from Propositions 2.2 and 2.3. □

Class (D) – models with zero Riemannian curvature

The definition is given in terms of the Riemannian curvature tensor with components defined by the equality

$$R_{hijk}(\vartheta) := \frac{\partial^2 \eta^{\mathrm{T}}(\vartheta)}{\partial\vartheta_h\, \partial\vartheta_j} W^{-1}[I - P(\vartheta)] \frac{\partial^2 \eta(\vartheta)}{\partial\vartheta_i\, \partial\vartheta_k} - \frac{\partial^2 \eta^{\mathrm{T}}(\vartheta)}{\partial\vartheta_h\, \partial\vartheta_k} W^{-1}[I - P(\vartheta)] \frac{\partial^2 \eta(\vartheta)}{\partial\vartheta_i\, \partial\vartheta_j}$$

Model (1.1) is in the class (D) iff $R_{hijk}(\vartheta)$ is zero for every $\vartheta \in \Theta$, and for every h, i, j, k.

Proposition 2.6. There is a reparametrization of model (1.1) making the information matrix constant iff the Riemannian curvature tensor is identically equal to zero on int Θ.

The proof of this proposition can be taken from differential geometry by identifying the information matrix $M(\vartheta)$ with a Riemannian metric tensor in Θ (cf. Eisenhart [6]).

Corollary. We have class (B) $\subseteq$ class (D), and class (C) $\subseteq$ class (D).

The first inclusion follows from Proposition 2.6, the second inclusion follows from Proposition 2.1.

Proposition 2.7. Every model (1.1), such that $\dim(\vartheta) = 1$, is in class (D). A reparameterization making the information matrix constant is given by

$$\beta = \beta(\vartheta) = \int_{\vartheta_0}^{\vartheta} \|\mathrm{d}\eta(t) \,/\, \mathrm{d}t\|_W$$

where $\vartheta_0 \in \operatorname{int} \Theta$ is an arbitrary point from int Θ.

The proof is obtained by a direct verification that $M(\vartheta)$ is constant.

This proposition clearly shows that there are models in the class (D) which are not intrinsically linear. Hence class (C) $\neq$ class (D). Taking a model with $\dim(\vartheta) = 1$ and such that $\|\mathrm{d}\eta(\vartheta) \,/\, \mathrm{d}\vartheta\|_W \neq$ const, we see also that there are models in (D) which do not have a constant information matrix. In symbols, class (B) $\neq$ class (D). The importance of models with zero Riemannian curvature in small sample nonlinear regression has been demonstrated in Pázman [11, 12, 13].

An example of a two dimensional regression model with zero Riemannian curvature is the classical Michaelis–Menten enzyme kinetic function

$$\eta_i(\vartheta) = \vartheta_1 x_i \,/\, (\vartheta_2 + x_i)\,; \qquad (i = 1, \ldots, N)$$

where $x_1, \ldots, x_N \in \mathbb{R}$ are fixed design points.

3. CONFIDENCE REGIONS FOR ϑ

In this paper we shall consider four kind of regions, which may be considered as confidence regions for ϑ under some assumptions. We shall present them first under the assumption that σ is known.

Region (I) – the elliptical region

$$\left\{\vartheta \in \Theta : [\hat{\vartheta} - \vartheta]^T M(\hat{\vartheta}) [\hat{\vartheta} - \vartheta] \leq \sigma^2 \chi^2_m (1-\alpha)\right\}$$

where $\hat{\vartheta}$ is the maximum likelihood estimate

$$\hat{\vartheta} := \hat{\vartheta}(y) := \underset{\vartheta \in \Theta}{\arg\min} \; \| y - \eta(\vartheta) \|^2_W \tag{3.1}$$

and $\chi^2_m(\beta)$ is the β-quantile of the χ^2_m distribution.

Region (II) – the likelihood region

$$\left\{\vartheta \in \Theta : \| y - \eta(\vartheta) \|^2_W - \| y - \eta(\hat{\vartheta}) \|^2_W \leq \sigma^2 \chi^2_m (1-\alpha)\right\}$$

which has a contour of constant likelihood.

Region (III) – the tangent elliptical region

$$\left\{\vartheta \in \Theta : \| P(\hat{\vartheta})[\eta(\hat{\vartheta}) - \eta(\vartheta)] \|^2_W \leq \sigma^2 \chi^2_m (1-\alpha) \;\&\; \| \eta(\hat{\vartheta}) - \eta(\vartheta) \|_W < r\right\}$$

where $r > 0$ is chosen so that

$$\begin{aligned} r &< \inf_{\vartheta} [K_{\text{int}}(\vartheta)]^{-1} \\ r^2/2 &< \sigma^2 \chi^2_m (1-\beta) \end{aligned} \tag{3.2}$$

with β much smaller than α. If there is no $r \in \mathbb{R}$ satisfying both inequalities, then it means that σ is too large to give a meaningful estimate of ϑ.

Region (IV) – the linear inference region

$$\{\vartheta \in \Theta : \| P(\vartheta)[y - \eta(\vartheta)] \|^2_W \leq \sigma^2 \chi^2_m (1-\alpha)\}$$

Although this region has formal similarities with region (III), there is the fundamental difference in that the projector $P(\vartheta)$ is used instead of $P(\hat{\vartheta})$.

Some general properties of the regions

Proposition 3.1. The region (I) is parameterically dependent. On the other hand, the regions (II) - (IV) are equivariant with respect to any regular reparameterization.

In the proof of the second statement it is used that $\eta(\vartheta)$ an $P(\vartheta)$ are invariant. Hence by setting a reparameterization $\vartheta = \vartheta(\beta)$ in the regions (III) - (IV) we obtain the same expressions, as by making first the reparameterization in model (1.1) and computing the confidence regions for β.

Proposition 3.2. All four regions coincide in the case that model (1.1) is linear (class (A)). The regions (II), (III) and (IV) coincide if the model (1.1) is intrinsically linear (class (C)).

Proof. The first statement is verified by a direct calculation. The second follows from Proposition 3.1, and from the first part of Proposition 3.2. □

Proposition 3.3. If model (1.1) is in the class (B), then the region (I) has the form

$$\left\{\vartheta \in \Theta : [\hat{\vartheta} - \vartheta]^T M [\hat{\vartheta} - \vartheta] \le \sigma^2 \chi^2_m (1-\alpha)\right\}$$

and the regions (III) and (IV) are both of the form

$$\left\{\vartheta \in \Theta : [\hat{\vartheta} - \vartheta]^T M [\hat{\vartheta} - \vartheta] + O(\| \hat{\vartheta} - \vartheta \|^4) \le \sigma^2 \chi^2_m (1-\alpha)\right\}.$$

Essentially, the proof is obtained using the Taylor expansion for $\eta(\vartheta)$ at the point $\hat{\vartheta}$, and Proposition 2.4.

Statistical properties of the regions

Region (I) is exact and optimal in linear models, and theoretically it is justified also in general models by the asymptotic normality of $\hat{\vartheta}$

$$\hat{\vartheta} \sim N\left(\vartheta, \sigma^2 M^{-1}(\vartheta)\right)$$

and by a certain consistency of $M(\hat{\vartheta})$ when considered as an estimator of $M(\vartheta)$ (cf. Johansen [9]). However, in finite sample applications, region (I) gives sometimes misleading confidence regions (cf. Bates and Watts [2], p. 65), with an overestimated confidence. This is true also in the particular case of intrinsically linear models. On the other hand, Ratkowsky [15] has shown empirically that reducing of the parameter effect curvature may approach considerably the probability density of $\hat{\vartheta}$ to the normal density, hence may make the region (I) valuable as confidence regions. This follows also from our Proposition 3.3 which implies that when the information matrix is constant (i. e. the parameter effect curvature is zero, Proposition 2.3), then the region (I) may be close to region (IV) which is known to be an exact confidence interval in any model (1.1).

The region (II) is the most popular in the applications of nonlinear regression. It is exact, and it can be considered optimal in intrinsically linear models (class (C)). It is used as an approximate confidence region also in some other models, however, correcting terms must be used. As is well known, Beale [3] proposed to multiply the term $\sigma^2\chi^2_m(1-\alpha)$ in region (II) by a term depending on components of the intrinsic curvature of model (1.1). (Cf. Bates and Watts [1] for an explicit and numerically accessible formula for this term.)

If ϑ is the true value of the parameter, and P is any (fixed) orthogonal projector onto an m-dimensional subspace of $\mathbb{R}^N$,then the random variable

$$\| P[y-\eta(\vartheta)] \|^2_W$$

is distributed χ^2_m. This is a well known result from linear models, which has nothing to do with nonlinear regression. Now, if we take $P = P(\vartheta)$ (which is fixed when ϑ is a fixed hypothetical parameter value), we obtain that the variable

$$\| P(\vartheta)[y-\eta(\vartheta)] \|^2_W$$

is distributed χ^2_m as well. This implies that region (IV) is an exact confidence region in any model (1.1), however, it is not recommended (cf. [2]).

The region (III) has been proposed recently by the author for models of class (D) (cf. [13]). If the number β can be neglected when compared with α, then, as proved in Pázman [12], the random variable

$$\sigma^{-2} \| P(\hat{\vartheta})[\eta(\hat{\vartheta})-\eta(\vartheta)] \|^2_W$$

is distributed as χ^2_m. The assumptions on the number r given in (3.2) ensure that we can neglect the probability that

$$\| \eta(\hat{\vartheta})-\eta(\vartheta) \|_W \geq r.$$

The condition $\| \eta(\hat{\vartheta})-\eta(\vartheta) \|_W < r$ must enter into the definition of the confidence region (III) because $\| P(\hat{\vartheta})[\eta(\hat{\vartheta})-\eta(\vartheta)] \|_W$ can be small also for those $\vartheta \in \Theta$ giving a large distance $\| \eta(\hat{\vartheta})-\eta(\vartheta) \|_W$, which is false.

There can be computational difficulties with the establishment of the number r, namely with the computation or evaluation of $K_{\text{int}}(\vartheta)$ for different values of ϑ(cf. Bates and Watts [1] for an algorithm). However, any serious analysis of regression models requires the computation of some curvature measures. In simple cases it is just sufficient to find a bound for $K_{\text{int}}(\vartheta)$ to verify the inequalities (3.2).

We note that if the region (III) is composed from several disjoint sets (despite of the condition $\| \eta(\hat{\vartheta})-\eta(\vartheta) \|_W < r$) it means that model (1.1) is overlapping, i.e. there are several nonnegligible relative minima in (3.1).

Confidence regions in the case of unknown σ

If we have the possibility to estimate σ^2 from an independent experiment in a standard way, so to obtain an estimator s^2 such that ks^2/σ^2 is distributed χ^2_k, then in all regions (I) – (IV) we write the term

$$ms^2 F_{m,k}(1-\alpha)$$

instead of the term $\sigma^2\chi^2_m(1-\alpha)$. Here $F_{m,k}(1-\alpha)$ is the $1-\alpha$ quantile of the F distribution with m and k degrees of freedom. In the inequalities (3.2) we can use s^2 instead of σ^2.

If we have no such estimator of σ^2, we have to estimate σ^2 somehow from y. Different methods are possible for different classes of models and different regions.

The estimator

$$s_1^2 = \frac{\|y-\eta(\hat{\vartheta})\|^2_W}{N-m}$$

with $k = N-m$ is used in intrinsically linear models (class (C)). In these models $(N-m)\sigma^{-2}s_1^2$ is distributed independently from $\hat{\vartheta}$, and according to χ^2_{N-m}. This is no more true in class (B) or (D). In those classes we can use the term

$$s_2^2 = \frac{\|[I-P(\hat{\vartheta})]\,[y-\eta(\vartheta)\|^2_W}{N-m} \tag{3.3}$$

and $k = N-m$. This expression is not an estimator of σ^2 since it depends on ϑ. However, the use of random variables depending of ϑ is allowed in confidence regions. In models from class (D), satisfying the inequalities (3.2), s_2^2 is independent of $\hat{\vartheta}$, and the random variable $(N-m)\sigma^{-2}s_2^2$ is distributed χ^2_{N-m}(cf. Pázman [13] for the use in region (III)).

In the linear inference region (region (IV)) one uses that for the true ϑ the random variable $\|\, [I-P(\vartheta)][y-\eta(\vartheta)]\,\|^2_W$ is independent from $\|\, P(\vartheta)[y-\eta(\vartheta)]\,\|^2_W$ and that $\sigma^{-2}\,\|\, [I-P(\vartheta)][y-\eta(\vartheta)]\,\|^2_W$ is distributed χ^2_{N-m}. So in the region (IV) we set

$$s_3^2 = \frac{\|[I-P(\vartheta)]\,[y-\eta(\vartheta)]\|^2_W}{N-m}$$

and $k = N-m$. The so obtained confidence region, although exact in every model (1.1), is not recommended (cf. [2] p. 223 for an explicit opinion of this sort). One important reason is that s_3^2 so computed is large exactly when ϑ is on one of the very disjoint part of the confidence region. This makes these distant (and intuitively false) parts of region (IV) erroneously very important.

So making the recapitulation, regions (I), (II), (IV) are not adequate for models from class (D). The region

$$\left\{\vartheta : \frac{(N-m)\,\|P(\hat{\vartheta})\,[\eta(\hat{\vartheta})\eta(\vartheta)]\|^2_W}{m\|[I-P(\hat{\vartheta}]\,[y-\eta(\vartheta)]\|^2_W} < F_{m,N-m}(1-\alpha)\ \&\ \|\,\eta(\hat{\vartheta})-\eta(\vartheta)\,\|^2_W < r\right\}$$

is a confidence region for ϑ when model (1.1) is in the class (D).

Comparison with a modificatication of the likelihood region

Some authors approximate the likelihood region (II) by sets having countours which are numerically simpler to compute (cf. Hamilton et al. [7], Clarke [4], To-Van-Ban [16]). In our context it is necessary to stop at Hamilton [7] where a projection of the likelihood region onto the tangent space at $\hat{\vartheta}$ and a quadratic Taylor formula are used. The region (III) considered in this paper is related to the tangent space at $\hat{\vartheta}$ as well, however, its origin has nothing to do with the likelihood region, and the region (III) is not equal to the region of Hamilton et al. [7].

(Received September 18, 1991.)

REFERENCES

[1] D. M. Bates and D. G. Watts: Relative curvature measures of nonlinearity (with discussion). J. Roy. Statist. Soc. *B 42* (1980), 1–25.

[2] D. M. Bates and D. G. Watts: Nonlinear Regression Analysis and Its Applications. Wiley, New York – Chichester 1988.

[3] E. M. L. Beale: Confidence regions in nonlinear estimation (with discussion). J. Roy. Statist. Soc. *B 22* (1960), 41–88.

[4] G. P. Y. Clarke: Approximate confidence limits for a parameter function in nonlinear regression. J. Amer. Statist. Assoc. *82* (1987), 221–230.

[5] E. Z. Demidenko: Optimization and Regression (in Russian). Nauka, Moskva 1989.

[6] L. P. Eisenhart: Riemannian Geometry. Princeton, University Press, Princeton, N. J. 1960.

[7] D. C. Hamilton, D. G. Watts and D. M. Bates: Accounting for intrinsic nonlinearity in nonlinear regression parameter inference regions. Ann. Statist. *10* (1982), 386–393.

[8] P. Hougaard: Covariance stabilizing transformations in nonlinear regression. Scand. J. Statist. *13* (1986), 207–210.

[9] S. Johansen: Functional Relations, Random Coefficients and Nonlinear Regression with Application to Kinetic Data. Lectures Notes in Statistics, Springer-Verlag, Berlin – Heidelberg – New York 1984.

[10] R. Kass: Canonical parameterizations and zero parameter effect curvature. J. Roy. Statist. Soc. *B 46* (1984), 86–92.

[11] A. Pázman: Flat Gaussian nonlinear regression models. In: Model Oriented Data Analysis (H. Läuter and V. Fedorov, eds.). (Lecture Notes in Economics and Mathematical Systems 297). Springer-Verlag, Berlin – Heidelberg – New York 1987, pp. 120–124.

[12] A. Pázman: Almost exact distribution of estimators II – Flat nonlinear regression models. Statistics *21* (1990), 21–33.

[13] A. Pázman: Pivotal variables and confidence regions in flat nonlionear regression models with unknown σ. Statistics *22* (1991), 177–189.

[14] A. Pázman: Nonlinear Statistical Models of Experiments (1993 to be published).

[15] D. A. Ratkowsky: Nonlinear Regression Modeling. Marcel Dekker, New York 1983.

[16] To-Van-Ban: Confidence Regions in a Nonlinear Regression Model (in Slovak). Ph. D. Dissertation, Comenius University, Bratislava 1990.

Andrej Pázman, Mathematical Institute, Slovak Academy of Sciences, Štefánikova 49, CS-814 73 Bratislava. Present address: *Department of Probability and Statistics, Faculty of Mathematics and Physics, Comenius University, Mlynská dolina, CS-842 16 Bratislava. Czechoslovakia.*