Approximation of Intrinsic Curvature in One Dimensional Nonlinear Regression Model by Moments of Prior Distribution of Parameter

K. Hornišová

Institute of Measurement Science of Slovak Academy of Sciences,
Dúbravská 9, 841 04 Bratislava, Slovakia
umerhorn@savba.sk

Abstract. *Taylor linearization methods are often used in nonlinear regression models to simplify statistical inferences. Criteria for correct use of such methods are often build on intrinsic curvature of the original model. When the prior distribution of parameter of expectation is known, the linearization by smoothing can be used instead of Taylor linearization. Criteria for its use should be based on some alternative measure of intrinsic curvature. We propose such a measure in the case of one dimensional expectation parameter.*

Keywords: *nonlinear regression model, intrinsic curvature, prior, eigenvalues*

1. Introduction

We consider a nonlinear regression model

$$y = \eta(\theta) + \varepsilon;\ \theta \in \Theta \subseteq \mathbb{R}^p, \qquad (1)$$
$$\varepsilon \sim N(0, \sigma^2 I_{N\times N}),$$

where $\eta(.) : \Theta \to \mathbb{R}^N$ is measurable mapping, $y \in \mathbb{R}^N$ are measurements with errors ε, θ are unknown parameters, I is identity matrix, and $\sigma > 0$ is unknown number.

Some of the properties of the model (1) can be studied by means of differential geometric properties of its expectation surface, i.e. of the set

$$\mathcal{E}_\eta := \{\eta(\theta); \theta \in \Theta\} \qquad (2)$$

(see e.g. [7]). One of the most important characteristics is the intrinsic curvature $K_{\text{int}}(\theta)$ of the model (1) in the point θ:

$$K_{\text{int}}(\theta) := \sup_{u \in \mathbb{R}^p \setminus \{\bar{0}\}} \frac{\|[I - P(\theta)]u^\top H(\theta) u\|}{u^\top M(\theta) u}, \theta \in \text{int}\,\Theta, \qquad (3)$$

where

$$P(\theta) := \frac{\partial \eta(\theta)}{\partial \theta^\top} \left(\frac{\partial \eta^\top(\theta)}{\partial \theta} \frac{\partial \eta(\theta)}{\partial \theta^\top}\right)^{-1} \frac{\partial \eta^\top(\theta)}{\partial \theta}$$

is the projector onto the tangent space of manifold $\mathcal{E}_\eta$ in the point θ, $M(\theta) := \frac{\partial \eta^\top(\theta)}{\partial \theta} \frac{\partial \eta(\theta)}{\partial \theta^\top}$ and $H(\theta) := \frac{\partial^2 \eta(\theta)}{\partial \theta \partial \theta^\top}$ is $3-$dimensional array of the second derivatives. For example, some functions of this quantity can serve as criteria of admissibility of Taylor linearization of model (1) in some a priori given point θ^0 with regard to various kinds of statistical inferences, see [5] or [4].

If the prior information on parameter θ is in the form of prior distribution π on Θ, linearization by smoothing of model (1) (see [8]), i.e. the model

$$y = A\theta + a + \varepsilon;\ \theta \in \Theta \subseteq \mathbb{R}^p, \qquad (4)$$
$$\varepsilon \sim N(0, \sigma^2 I_{N\times N}),$$

where $A := Cov_\pi(\eta,\theta)Var_\pi^{-1}\theta$ and $a = E_\pi\eta - AE_\pi\theta$, can be more appropriate than Taylor linearization. However, it is obvious that criteria of admissibility of (4) should have the form different from those in [4]. In particular, it is necessary to find some suitable analogue of (3). One such analogue in the case of $1-$dimensional parameter θ will be given in the next part. Now we review some known facts about the so-called intrinsically linear models, i.e. models (1) with the property $K_{\text{int}}(\theta) = 0$ for all $\theta \in int\Theta$ (see [7]).

LEMMA 1. *([7]) Model* (1) *je intrinsically linear iff* $\mathcal{E}_\eta$ *is relatively open set in some* $k-$*dimensional affine subspace in* $\mathbb{R}^N$*, where* $k \leq p$.

Since in intrinsically linear models some statistical procedures have better properties than in general model (1), it is a natural question whether the given model (1) can be sufficiently precisely approximated by an intrinsically linear one.

If a prior π is given on the parameter space Θ, the intrinsically linear approximation of model (1) with the minimal prior mean squared error is described in the following proposition:

PROPOSITION 1. *([2]) Let* $\mathcal{E}_\eta$ *is a* $p-$*dimensional manifold, let* π *be regular probability distribution on* Θ*. Then the optimal intrinsic linearization of the model* (1)*, i.e. the solution of the problem*

$$C_1 := \min_{\substack{A\in\mathbb{R}^{N\times p}\\ h(A)=p\\ a\in\mathbb{R}^N\\ \beta(.):\Theta\to\mathbb{R}^p}} E_\pi\{[\eta(\theta) - (A\beta(\theta)+a)][\eta(\theta) - (A\beta(\theta)+a)]^\top\}, \qquad (4a)$$

or equivalently

$$c_1 := \min_{\substack{A\in\mathbb{R}^{N\times p}\\ h(A)=p\\ a\in\mathbb{R}^N\\ \beta(.):\Theta\to\mathbb{R}^p}} E_\pi[\|\eta(\theta) - (A\beta(\theta)+a)\|^2] \qquad (4b)$$

is

$$\begin{aligned} A &= (u_1,\dots,u_p), \qquad (5)\\ a &= E_\pi\eta\ ,\\ \beta(\theta) &= (A^\top A)^{-1}A^\top(\eta(\theta)-a),\\ C_1 &= (I-P)Var_\pi\eta(I-P),\\ c_1 &= trC_1 = \sum_{i=p+1}^N \lambda_i, \end{aligned}$$

where $u_1,\dots,u_N$ *are orthonormal eigenvectors, corresponding respectively to eigenvalues* $\lambda_1 \geq \dots \geq \lambda_N \geq 0$ *of the matrix* $Var_\pi(\eta)$, $P := P_U := U(U^\top U)^{-1}U^\top$, $U := \begin{pmatrix} u_1 & \dots & u_p\end{pmatrix}$*. If the model* (1) *is regular then* $k = p$.

Optimal intrinsically linear approximation of the model (1) from the proposition 1 is

$$y = A\beta(\theta) + a + \varepsilon = A(A^\top A)^{-1}A^\top(\eta(\theta) - E_\pi\eta) + E_\pi\eta + \varepsilon,$$
$$\varepsilon \sim N(0, \sigma^2 I), \tag{6}$$

where A is the same as in proposition 1.

The following lemma gives another characterization of intrinsically linear models:

LEMMA 2. *([3]) Let $\varepsilon\eta$ is a $p-$dimensional manifold. Let π be regular probability distribution on Θ. The model* (1) *is intrinsically linear with $\pi-$ probability* 1 *iff* $\mathrm{h}(Var_\pi\eta) \leq p$.

2. 2. Approximation of intrinsic curvature by prior moments

Proposition 1 suggests that minimal square "distance"

$$c_1 := \sum_{i=k+1}^{N} \lambda_i \tag{7}$$

from linearized model (6) from the original model (1) can be considered as a measure of intrinsic nonlinearity of the model (1). However, for sequence of priors π contracting to singular distribution concentrated in θ^0, this expression converges to zero, so it cannot be used directly as approximation of intrinsic curvature $K_{\text{int}}(\theta^0)$ but only after some suitable standardization. In what follows we derive such standardization for case $p = 1$. We shall utilize some known propositions.

In the first one, the approximation of matrix of first derivatives of the regression function $\eta(.)$, based on moments of prior π and utilizable for π concentrated in a small neighbourhood of θ^0, is given:

PROPOSITION 2. *([8]) Let $\eta(.)$ be $2-$times continuously differentiable. Let $\{\pi_n; n \in \mathbb{N}\}$ is a sequence of regular prior distributions with supports that are subsets of some common compact subset of Θ. Let for $\forall n$ there exist finite moments $E_{\pi_n}(\theta)$, $E_{\pi_n}(\eta)$, $Var_{\pi_n}(\theta)$, $Cov_{\pi_n}(\eta,\theta)$, $Var_{\pi_n}(\eta)$, let $Var_{\pi_n}(\theta)$ are regular, let $Cov_{\pi_n}(\eta,\theta)$ have the full rank. Let*

$$\pi_n \overset{weakly}{\longrightarrow} \pi_0,$$

where π_0 is distribution concentrated in the point θ^0. Then

$$\lim_{n\to\infty} E_n\theta = \theta^0, \qquad \lim_{n\to\infty} E_n\eta = \eta(\theta^0),$$
$$\lim_{n\to\infty} A_{\pi_n} := \lim_{n\to\infty} Cov_n(\eta,\theta)Var_n^+\theta = \frac{\partial\eta(\theta^0)}{\partial\theta^\top}. \tag{8}$$

Similarly, the approximation of the second derivatives of $\eta(.)$ can be derived:

PROPOSITION 3. *([3]) Let the assumptions of proposition 2 hold. Let $\eta(.)$ is three times continuously differentiable on Θ. Let*

$$"\partial^2\eta(\theta^0)" = \left(\partial^2\eta_{1,1}(\theta^0), \partial^2\eta_{1,2}(\theta^0), \ldots, \partial^2\eta_{1,p}(\theta^0), \partial^2\eta_{2,2}(\theta^0), \partial^2\eta_{2,3}(\theta^0), \ldots, \partial^2\eta_{2,p}(\theta^0), \ldots, \partial^2\eta_{p,p}(\theta^0)\right)$$

is $N\times(p+\binom{p}{2})$ matrix. Let $\theta^{\odot 2} = \left(\theta_1^2, \theta_1\theta_2, \ldots, \theta_1\theta_p, \theta_2^2, \theta_2\theta_3, \ldots, \theta_2\theta_p, \ldots, \theta_p^2\right)^\top$ is $(p+\binom{p}{2})-$dimensional vector, let F_2 is a $\binom{p+1}{2} \times \binom{p+1}{2}$ diagonal matrix with $i_1! \ldots i_p!$ as its $(i_1, \ldots, i_p)-$th (in lexicographical ordering) diagonal element, where $i_u \geq 0$; $u = 1, \ldots, p$, are integers such that $i_1 + \cdots + i_p = 2$. Then

$$\lim_{n\to\infty}[Cov_n(\eta,\theta^{\odot 2}) - Cov_n(\eta,\theta)Var_n^+\theta Cov_n(\theta,\theta^{\odot 2})][Var_n\theta^{\odot 2} - Cov_n(\theta^{\odot 2},\theta)Var_n^+\theta Cov_n(\theta,\theta^{\odot 2})]^+F_2 =$$
$$="\partial^2\eta(\theta^0)".$$

If prior π is concentrated in a small neighbourhood of the point θ^0, then the expectation surface of optimal intrinsic linearization is close to the affine tangent space of expectation surface of the original model in point θ^0. It means that the space generated by eigenvectors $u_1, \ldots, u_k$ of the matrix $Var_\pi \eta$ is near to the tangent space of expectation surface $\mathcal{E}_\eta$ of the model (1) in the point $\eta(\theta^0)$, i.e. the projectors onto these spaces are close to each other:

THEOREM 1. *([3]) Let us consider the function $\eta(.)$ and a sequence of priors π_n with the same properties as in proposition 2. Let $\lambda_1^n \geq \cdots \geq \lambda_N^n$ be eigenvalues and $u_1^n, \ldots, u_N^n$ are corresponding orthonormal eigenvectors of the matrix $Var_{\pi_n}\eta$. Let $\eta(.)$ be two times continuously differentiable on Θ_0. Let $\forall$ limit point of the sequence $\{\frac{Var_{\pi_n}\theta}{\|Var_{\pi_n}\theta\|}\}_{n=1}^\infty$ be a regular matrix. Then*

$$\lim_{n\to\infty} P_{(u_1^{(n)},\ldots,u_k^{(n)})} = P := P_{\frac{\partial\eta(\theta^0)}{\partial\theta^\top}}, \qquad (9)$$
$$\lim_{n\to\infty} P_{(u_{k+1}^{(n)},\ldots,u_N^{(n)})} = I - P.$$

COROLLARY 1. *Let the assumptions of theorem 1 hold. Then the sets of limits of all convergent susequences of sequences*

$$\frac{Var_{\pi_n}\eta}{\|Var_{\pi_n}\theta\|} \text{ and } \partial\eta\frac{Var_{\pi_n}\theta}{\|Var_{\pi_n}\theta\|}\partial\eta^\top \qquad (14)$$

are identical.

Let $q := \dim([\{(I-P)\frac{\partial^2\eta(\theta^0)}{\partial\theta_j\partial\theta_i}; i, j = 1, \ldots, p\}])$.

THEOREM 2. *Let the assumptions of theorem 1 hold, let $p = 1$. Moreover, let the following assertions about limits hold:*

$$\limsup_{n\to\infty} \frac{Var_{\pi_n}[(\theta-\theta^0)^2]}{\|Var_{\pi_n}\theta\|^2}$$

is finite and positive,

$$\limsup_{n\to\infty} \frac{E_{\pi_n}[(\theta-\theta^0)^2]}{\|Var_{\pi_n}\theta\|} < \infty,$$

$$\limsup_{n\to\infty} \frac{Var_{\pi_n}\theta^2 - Cov_{\pi_n}(\theta^2,\theta)Var_{\pi_n}^{-1}\theta Cov_{\pi_n}(\theta,\theta^2)}{\|Var_{\pi_n}\theta\|^2}$$

is finite
and

$$\limsup_{n\to\infty} \frac{E_{\pi_n}[\partial^3\eta(\theta^*)(\theta-\theta^0)^3(\theta-E_{\pi_n}\theta)]}{\|Var_{\pi_n}\theta\|^2}$$

is finite.

Then

$$\lim_{n\to\infty} P_{(u_{k+1}^{(n)},\ldots,u_{k+q}^{(n)})} = P_{(I-P)\partial^2\eta(\theta^0)}, \qquad (15)$$

where

$$(I-P)\partial^2\eta(\theta^0) := (I-P)\begin{pmatrix} \frac{\partial^2\eta(\theta^0)}{\partial\theta_1\partial\theta_1} & \frac{\partial^2\eta(\theta^0)}{\partial\theta_2\partial\theta_1} & \ldots & \frac{\partial^2\eta(\theta^0)}{\partial\theta_p\partial\theta_p}\end{pmatrix}.$$

PROOF.

$$
\begin{aligned}
&(I-P)Var_\pi\eta(I-P)=(I-P)E_\pi[(\eta(\theta)-E_\pi\eta)(\eta(\theta)-E_\pi\eta)^\top](I-P)=\\
=&(I-P)E_\pi[(\eta(\theta^0)-E_\pi\eta+\partial\eta(\theta^0)(\theta-\theta^0)+\frac{1}{2}\partial^2\eta(\theta^0)(\theta-\theta^0)^2+\\
&\qquad+\frac{1}{6}\partial^3\eta(\theta^*)(\theta-\theta^0)^3)(\eta(\theta)-E_\pi\eta)^\top](I-P)=\\
=&(I-P)\{\partial\eta(\theta^0)Cov_\pi(\theta,\eta)+\frac{1}{2}\partial^2\eta(\theta^0)E_\pi[(\theta-\theta^0)^2(\eta(\theta)-E_\pi\eta)^\top]+\\
&\qquad+\frac{1}{6}E_\pi[\partial^3\eta(\theta^*)(\theta-\theta^0)^3)(\eta(\theta)-E_\pi\eta)^\top]\}(I-P)=\\
=&(I-P)\{\frac{1}{2}\partial^2\eta(\theta^0)E_\pi[(\theta-\theta^0)^2(\eta(\theta)-E_\pi\eta)^\top]+\frac{1}{6}E_\pi[\partial^3\eta(\theta^*)(\theta-\theta^0)^3)(\eta(\theta)-E_\pi\eta)^\top]\}(I-P).
\end{aligned}
\tag{16.0}
$$

Since it holds that

$$
\begin{aligned}
E_\pi[(\theta-\theta^0)^2(\eta(\theta)-E_\pi\eta)^\top]&=Cov_\pi(\theta^2,\eta)-2\theta^0Cov_\pi(\theta,\eta)=\\
&=Cov_\pi(\theta^2,\eta)-Cov_\pi(\theta^2,\theta)Var_\pi^{-1}\theta Cov_\pi(\theta,\eta)+\\
&+[Cov_\pi(\theta^2,\theta)Var_\pi^{-1}\theta-2\theta^0]Cov_\pi(\theta,\eta),
\end{aligned}
\tag{16}
$$

$$
\begin{aligned}
Cov_\pi(\eta,\theta)=\partial\eta(\theta^0)Var_\pi\theta+\frac{1}{2}\partial^2\eta(\theta^0)Cov_\pi(\theta^2,\theta)-\frac{1}{2}\partial^2\eta(\theta^0)2\theta^0Var_\pi\theta+\\
+\frac{1}{6}E_\pi[\partial^3\eta(\theta^*)(\theta-\theta^0)^3(\theta-E_\pi\theta)],
\end{aligned}
\tag{17}
$$

and

$$
\begin{aligned}
Var_\pi[(\theta-\theta^0)^2]=&Var_\pi(\theta^2)-Cov_\pi(\theta^2,\theta)Var_\pi^{-1}\theta Cov_\pi(\theta,\theta^2)+\\
&+[Cov_\pi(\theta^2,\theta)Var_\pi^{-1}\theta-2\theta^0]Var_\pi\theta[Var_\pi^{-1}\theta Cov_\pi(\theta,\theta^2)-2\theta^0],
\end{aligned}
\tag{18}
$$

it also holds that

$$
\begin{aligned}
&E_\pi[(\theta-\theta^0)^2(\eta(\theta)-E_\pi\eta)^\top]=\\
=&\{[Var_\pi(\theta^2)-Cov_\pi(\theta^2,\theta)Var_\pi^{-1}\theta Cov_\pi(\theta,\theta^2)]\frac{1}{2}\partial^2\eta^\top(\theta^0)+\\
+&\frac{1}{2}[Var_\pi(\theta^2)-Cov_\pi(\theta^2,\theta)Var_\pi^{-1}\theta Cov_\pi(\theta,\theta^2)][2[Var_\pi(\theta^2)-Cov_\pi(\theta^2,\theta)Var_\pi^{-1}\theta Cov_\pi(\theta,\theta^2)]^{-1}.\\
&.[Cov_\pi(\theta^2,\eta)-Cov_\pi(\theta^2,\theta)Var_\pi^{-1}\theta Cov_\pi(\theta,\eta)]-\partial^2\eta^\top(\theta^0)]+\\
+&[Cov_\pi(\theta^2,\theta)Var_\pi^{-1}\theta-2\theta^0]\{[Cov_\pi(\theta,\theta^2)-Var_\pi\theta2\theta^0]\frac{1}{2}\partial^2\eta^\top(\theta^0)+Var_\pi\theta\partial\eta^\top(\theta^0)+\\
+&\frac{1}{6}E_\pi[\partial^3\eta(\theta^*)(\theta-\theta^0)^3(\theta-E_\pi\theta)]\}=\\
=&[Cov_\pi(\theta^2,\theta)Var_\pi^{-1}\theta-2\theta^0]Var_\pi\theta\partial\eta^\top(\theta^0)+Var_\pi[(\theta-\theta^0)^2]\frac{1}{2}\partial^2\eta^\top(\theta^0)+\\
+&\frac{1}{2}[Var_\pi(\theta^2)-Cov_\pi(\theta^2,\theta)Var_\pi^{-1}\theta Cov_\pi(\theta,\theta^2)][2[Var_\pi(\theta^2)-Cov_\pi(\theta^2,\theta)Var_\pi^{-1}\theta Cov_\pi(\theta,\theta^2)]^{-1}.\\
&.[Cov_\pi(\theta^2,\eta)-Cov_\pi(\theta^2,\theta)Var_\pi^{-1}\theta Cov_\pi(\theta,\eta)]-\partial^2\eta^\top(\theta^0)]+\\
+&[Cov_\pi(\theta^2,\theta)Var_\pi^{-1}\theta-2\theta^0]\frac{1}{6}E_\pi[\partial^3\eta(\theta^*)(\theta-\theta^0)^3(\theta-E_\pi\theta)].
\end{aligned}
$$

Further, proposition 3 implies that

$$\lim_{n\to\infty} Cov_{\pi_n}(\theta^2,\theta)Var_{\pi_n}^{-1}\theta = 2\theta^0. \tag{19}$$

Then we get

$$\begin{aligned}
&(I-P)\frac{1}{2}\partial^2\eta(\theta^0)E_\pi[(\theta-\theta^0)^2(\eta(\theta)-E_\pi\eta)^\top](I-P) = \\
&=(I-P)\frac{1}{2}\partial^2\eta(\theta^0)\{[Cov_\pi(\theta^2,\eta)-Cov_\pi(\theta^2,\theta)Var_\pi^{-1}Cov_\pi(\theta,\eta)]+ \\
&+[Cov_\pi(\theta^2,\theta)Var_\pi^{-1}\theta-2\theta^0]Cov_\pi(\theta,\eta)\}(I-P) = \\
&=(I-P)\{\frac{1}{4}\partial^2\eta(\theta^0)[Var_\pi(\theta^2)-Cov_\pi(\theta^2,\theta)Var_\pi^{-1}Cov_\pi(\theta,\theta^2)]\partial^2\eta^\top(\theta^0)+ \\
&+\frac{1}{2}\frac{1}{2}\partial^2\eta(\theta^0)[Var_\pi(\theta^2)-Cov_\pi(\theta^2,\theta)Var_\pi^{-1}\theta Cov_\pi(\theta,\theta^2)]. \\
&\quad .[2[Var_\pi(\theta^2)-Cov_\pi(\theta^2,\theta)Var_\pi^{-1}\theta Cov_\pi(\theta,\theta^2)]^{-1}. \\
&\quad .[Cov_\pi(\theta^2,\eta)-Cov_\pi(\theta^2,\theta)Var_\pi^{-1}\theta Cov_\pi(\theta,\eta)]-\partial^2\eta^\top(\theta^0)]+ \\
&+\frac{1}{2}\partial^2\eta(\theta^0)[Cov_\pi(\theta^2,\theta)Var_\pi^{-1}\theta-2\theta^0]Cov_\pi(\theta,\eta)\}(I-P) =
\end{aligned}$$

$$\begin{aligned}
= (I-P)\{&\frac{1}{4}\partial^2\eta(\theta^0)[Var_\pi(\theta^2)-Cov_\pi(\theta^2,\theta)Var_\pi^{-1}Cov_\pi(\theta,\theta^2)]\partial^2\eta^\top(\theta^0)+ \\
&+\frac{1}{2}\frac{1}{2}\partial^2\eta(\theta^0)[Var_\pi(\theta^2)-Cov_\pi(\theta^2,\theta)Var_\pi^{-1}\theta Cov_\pi(\theta,\theta^2)]. \\
&\quad .[2[Var_\pi(\theta^2)-Cov_\pi(\theta^2,\theta)Var_\pi^{-1}\theta Cov_\pi(\theta,\theta^2)]^{-1}. \\
&\quad .[Cov_\pi(\theta^2,\eta)-Cov_\pi(\theta^2,\theta)Var_\pi^{-1}\theta Cov_\pi(\theta,\eta)]-\partial^2\eta^\top(\theta^0)]+ \\
&+\frac{1}{2}\partial^2\eta(\theta^0)[Cov_\pi(\theta^2,\theta)Var_\pi^{-1}\theta-2\theta^0]\{[Cov_\pi(\theta,\theta^2)-Var_\pi\theta 2.\theta^0]\frac{1}{2}\partial^2\eta^\top(\theta^0)+ \\
&+Var_\pi\theta.\partial\eta^\top(\theta^0)+\frac{1}{6}E_\pi[\partial^3\eta(\theta^*)(\theta-\theta^0)^3(\theta-E_\pi\theta)]\}\}(I-P) = \\
&=(I-P)\frac{1}{4}\partial^2\eta(\theta^0)Var_\pi[(\theta-\theta^0)^2]\partial^2\eta(\theta^0)(I-P)+ \\
&+(I-P)\frac{1}{2}\frac{1}{2}\partial^2\eta(\theta^0)[Var_\pi(\theta^2)-Cov_\pi(\theta^2,\theta)Var_\pi^{-1}\theta Cov_\pi(\theta,\theta^2)]. \\
&\quad .[2[Var_\pi(\theta^2)-Cov_\pi(\theta^2,\theta)Var_\pi^{-1}\theta Cov_\pi(\theta,\theta^2)]^{-1}. \\
&\quad .[Cov_\pi(\theta^2,\eta)-Cov_\pi(\theta^2,\theta)Var_\pi^{-1}\theta Cov_\pi(\theta,\eta)]-\partial^2\eta^\top(\theta^0)](I-P)+ \\
&+\frac{1}{2}\partial^2\eta(\theta^0)[Cov_\pi(\theta^2,\theta)Var_\pi^{-1}\theta-2\theta^0]\frac{1}{6}E_\pi[\partial^3\eta(\theta^*)(\theta-\theta^0)^3(\theta-E_\pi\theta)](I-P).
\end{aligned}$$

Then

$$\begin{aligned}
&\lim_{n\to\infty}(I-P)\frac{Var_{\pi_n}\eta}{\|Var_{\pi_n}\theta\|^2}(I-P) = \\
&=\frac{1}{4}(I-P)\partial^2\eta(\theta^0)\lim_{n\to\infty}\frac{Var_{\pi_n}[(\theta-\theta^0)^2]}{\|Var_{\pi_n}\theta\|^2}\partial^2\eta^\top(\theta^0)(I-P)+ \\
&+(I-P)\frac{1}{2}\frac{1}{2}\partial^2\eta(\theta^0)\lim_{n\to\infty}\frac{Var_{\pi_n}(\theta^2)-Cov_{\pi_n}(\theta^2,\theta)Var_{\pi_n}^{-1}\theta Cov_{\pi_n}(\theta,\theta^2)]}{\|Var_{\pi_n}\theta\|^2}. \\
&.\{2[Var_{\pi_n}(\theta^2)-Cov_{\pi_n}(\theta^2,\theta)Var_{\pi_n}^{-1}\theta Cov_{\pi_n}(\theta,\theta^2)]^{-1}. \\
&.[Cov_{\pi_n}(\theta^2,\eta)-Cov_{\pi_n}(\theta^2,\theta)Var_{\pi_n}^{-1}\theta Cov_{\pi_n}(\theta,\eta)]- \\
&-\partial^2\eta^\top(\theta^0)\}(I-P)+
\end{aligned}$$

$$+ (I-P)\frac{1}{2}\partial^2\eta(\theta^0)\lim_{n\to\infty}[Cov_{\pi_n}(\theta^2,\theta)Var^{-1}_{\pi_n}\theta - 2\theta^0]\frac{1}{6}\frac{E_{\pi_n}[\partial^3\eta(\theta^*)(\theta-\theta^0)^3(\theta-E_{\pi_n}\theta)]}{\|Var_{\pi_n}\theta\|^2}(I-P)+$$
$$+ (I-P)\frac{1}{6}\lim_{n\to\infty}\frac{E_{\pi_n}[\partial^3\eta(\theta^*)(\theta-\theta^0)(\theta-\theta^0)^2(\eta(\theta)-E_{\pi_n}\eta)^\top](I-P)}{\|Var_{\pi_n}\theta\|^2} =$$
$$= \frac{1}{4}(I-P)\partial^2\eta(\theta^0)\lim_{n\to\infty}\frac{Var_{\pi_n}[(\theta-\theta^0)^2]}{\|Var_{\pi_n}\theta\|^2}\partial^2\eta^\top(\theta^0)(I-P).$$

The rest of the proof is then the same as the proof for the previous theorem. □ □

COROLLARY 2.

$$\lim_{n\to\infty}(I-P_n)\frac{Var_{\pi_n}\eta}{\|Var_{\pi_n}\theta\|^2}(I-P_n) = \lim_{n\to\infty}\frac{1}{4}(I-P_n)\partial^2\eta\frac{Var_{\pi_n}((\theta-\theta^0)^2)}{\|Var_{\pi_n}\theta\|^2}\partial^2\eta^\top(I-P_n). \quad (20)$$

By means of $Var_\pi\eta$ we can approximate also the curvature $K_{\text{int}}(\theta^0)$ of one dimensional models: Let us denote

$$\kappa_n = \frac{Var_{\pi_n}((\theta-\theta^0)^2)}{\|Var_n\theta\|^2}.$$

For example, if π_α is uniform distribution on interval $(\theta^0-\alpha,\theta^0+\alpha)$, then

$$\lim_{\alpha\to 0^+}\kappa_\alpha = \lim_{\alpha\to 0^+}\frac{\frac{4\alpha^4}{45}}{\frac{\alpha^4}{9}} = \frac{4}{5}.$$

THEOREM 3. *Let the assumptions of theorem 2 hold. If $p=1$ then*

$$\lim_{n\to\infty}\sqrt{\frac{4}{\kappa_n}}\frac{\sqrt{\lambda_2^{(n)}}}{\lambda_1^{(n)}} = K_{\text{int}}(\theta^0). \quad (21)$$

PROOF. Theorem 1 implies that in case $k=1$ it holds that

$$\lim_{n\to\infty}u_1^n = \pm\frac{\partial\eta(\theta^0)}{\|\partial\eta(\theta^0)\|}.$$

This condition also implies that $q=0$ or $q=1$. If $q=0$, then the model is intrinsically linear and then according to proposition 1 and lemma 2 $\lambda_2^n=0$, so

$$\lim_{n\to\infty}\frac{\sqrt{\lambda_2^n}}{\lambda_1^n} = 0 = K_{\text{int}}(\theta^0).$$

If $k=1$ and $q=1$ then according to theorem 2

$$\frac{\nu(\lim_{n\to\infty}u_1^n)}{\|\nu(\lim_{n\to\infty}u_1^n)\|} = \lim_{n\to\infty}\pm u_2^n,$$

where $\nu(u)$ is a vector of curvature of $\mathcal{E}_\eta$ in the direction of tangential vector u. Then utilizing corollaries

1 and 2

$$\lim_{n\to\infty}\sqrt{\frac{4}{\kappa_n}}\frac{\sqrt{\lambda_2^n}}{\lambda_1^n} = \lim_{n\to\infty}\sqrt{\frac{4}{\kappa_n}}\frac{\sqrt{u_2^{n\top}Var_{\pi_n}\eta u_2^n}}{u_1^{n\top}Var_{\pi_n}\eta u_1^n} = \lim_{n\to\infty}\sqrt{\frac{4}{\kappa_n}}\frac{\sqrt{u_2^{n\top}\frac{Var_{\pi_n}\eta}{\|Var_{\pi_n}\theta\|^2}u_2^n}}{u_1^{n\top}\frac{Var_{\pi_n}\eta}{\|Var_{\pi_n}\theta\|}u_1^n} =$$

$$= \lim_{n\to\infty}\sqrt{\frac{4}{\kappa_n}}\frac{\sqrt{\frac{v(u_1^{n\top})^\top}{\|v(u_1^{n\top})\|}\frac{Var_{\pi_n}\eta}{\|Var_{\pi_n}\theta\|^2}\frac{v(u_1^n)}{\|v(u_1^n)\|}}}{u_1^{n\top}\frac{Var_{\pi_n}\eta}{\|Var_{\pi_n}\theta\|}u_1^n} = \lim_{n\to\infty}\sqrt{\frac{4}{\kappa_n}}\frac{\sqrt{\frac{\partial^2\eta^\top(\theta^0)(I-P)\frac{Var_{\pi_n}\eta}{\|Var_{\pi_n}\theta\|^2}(I-P)\partial^2\eta(\theta^0)}{\|(I-P)\partial^2\eta(\theta^0)\|^2}}}{u_1^{n\top}\partial\eta(\theta^0)\frac{Var_{\pi_n}\theta}{\|Var_{\pi_n}\theta\|}\partial\eta^\top(\theta^0)u_1^n} =$$

$$= \lim_{n\to\infty}\sqrt{\frac{4}{\kappa_n}}\frac{\sqrt{\frac{\frac{1}{4}\partial^2\eta^\top(\theta^0)(I-P)\partial^2\eta(\theta^0)\frac{Var_{\pi_n}((\theta-\theta^0)^2)}{\|Var_{\pi_n}\theta\|^2}\partial^2\eta^\top(\theta^0)(I-P)\partial^2\eta(\theta^0)}{\|(I-P)\partial^2\eta(\theta^0)\|^2}}}{\frac{\partial\eta^\top(\theta^0)}{\|\partial\eta(\theta^0)\|}\partial\eta(\theta^0).1.\partial\eta^\top(\theta^0)\frac{\partial\eta(\theta^0)}{\|\partial\eta(\theta^0)\|}} =$$

$$= \frac{\|(I-P)\partial^2\eta(\theta^0)\|}{\|\partial\eta(\theta^0)\|^2} = K_{\text{int}}(\theta^0).$$

□ □

Moreover, from the proof it can be seen that for $p = 1$

$$\lim_{n\to\infty}\frac{\lambda_1^n}{\|Var_{\pi_n}\theta\|} = \|\partial\eta(\theta^0)\|^2$$

and

$$\lim_{n\to\infty}\sqrt{\frac{4\lambda_2^n}{Var_{\pi_n}[(\theta-\theta^0)^2]}} = \|(I-P)\partial^2\eta(\theta^0)\|.$$

Approximation (21) equals to the intrinsic curvature $K_{\text{int}}(\theta^0)$ in the limit for prior distribution π concentrated in θ^0. A differently defined intrinsic curvature $\tilde{K}_{\text{int}}(\theta^0)$ of the model (1) with prior distribution π is in [P, 92] and [P, 93]. $\tilde{K}_{\text{int}}(\theta^0)$ involves derivatives and equals to $K_{\text{int}}(\theta^0)$ for prior π uniform on Θ.

3. Examples

Theorem 3 will be further illustrated on examples with different dimensions p, q and N:

EXAMPLE 1 ellipse ($N = 2, p = 1, q = 1$)

$$\eta(\theta) = \begin{pmatrix} a\cos\theta \\ b\sin\theta \end{pmatrix},$$

where $a \neq 0 \neq b$ are known, π is uniform on $\Theta_\alpha := <\theta^0 - \alpha, \theta^0 + \alpha>$, $\theta^0 = \frac{\pi}{2}$, $\alpha \in (0, \frac{\pi}{2})$. Then eigenvalues of the matrix $Var_\pi\eta$ are

$$\lambda_1 = \lambda_1(\alpha) = a^2(\frac{1}{2} - \frac{1}{2}\frac{\sin\alpha}{\alpha}\cos\alpha)$$
$$\lambda_2 = \lambda_2(\alpha) = b^2(\frac{1}{2} + \frac{1}{2}\frac{\sin\alpha}{\alpha}\cos\alpha - \frac{\sin^2\alpha}{\alpha^2}).$$

It holds:

$$\lim_{\alpha\to 0}\frac{\lambda_1(\alpha)}{Var_{\pi_\alpha}\theta} = \lim_{\alpha\to 0}\frac{\lambda_1(\alpha)}{\frac{\alpha^2}{3}} = a^2.1,$$
$$\lim_{\alpha\to 0}\frac{\lambda_2(\alpha)}{(Var_{\pi_\alpha}\theta)^2} = \lim_{\alpha\to 0}\frac{\lambda_2(\alpha)}{\frac{\alpha^4}{9}} = b^2.\frac{1}{5}.$$

λ_1 is „greater" than λ_2 in the sense that for $\forall a, b > 0; \exists \epsilon > 0; \forall \alpha \in (0, \epsilon); \lambda_1(\alpha) > \lambda_2(\alpha)$. Then

$$\lim_{\alpha\to 0} \frac{\sqrt{5}\sqrt{\lambda_{k+1}(\alpha)}}{\lambda_k(\alpha)} = \lim_{\alpha\to 0} \frac{\sqrt{5}\sqrt{\lambda_2(\alpha)}}{\lambda_1(\alpha)} = \frac{\sqrt{5}b}{a^2} \lim_{\alpha\to 0} \frac{\sqrt{\frac{1}{2}+\frac{1}{2}\frac{\sin\alpha}{\alpha}\cos\alpha - \frac{\sin^2\alpha}{\alpha^2}}}{\frac{1}{2}-\frac{1}{2}\frac{\sin\alpha}{\alpha}\cos\alpha} =$$

$$= \frac{\sqrt{5}b}{a^2} \lim_{\alpha\to 0} \frac{\sqrt{\frac{\frac{1}{2}+\frac{1}{2}\frac{\sin\alpha}{\alpha}\cos\alpha - \frac{\sin^2\alpha}{\alpha^2}}{\frac{\alpha^4}{9}}}}{\frac{\frac{1}{2}-\frac{1}{2}\frac{\sin\alpha}{\alpha}\cos\alpha}{\frac{\alpha^2}{3}}} = \frac{b}{a^2} = K_{\text{int}}(\theta^0).$$

EXAMPLE 2 helix wound around the cylinder, with ellipse from example 1 as the base; parameter c determines the density of windings (the greater c, the smaller density) $(N = 3, p = 1, q = 1)$

$$\eta(\theta) = \begin{pmatrix} a\cos\theta \\ b\sin\theta \\ c\theta \end{pmatrix},$$

where $c \neq 0$ is known and other assumptions are as in example 1. Then

$$\lambda_1 = \frac{a^2(\frac{1}{2}-\frac{1}{2}\frac{\sin\alpha}{\alpha}\cos\alpha) + c^2\frac{\alpha^2}{3} + \sqrt{(a^2(\frac{1}{2}-\frac{1}{2}\frac{\sin\alpha}{\alpha}\cos\alpha) - c^2\frac{\alpha^2}{3})^2 + 4a^2c^2(\cos\alpha - \frac{\sin\alpha}{\alpha})^2}}{2}$$

$$\lambda_2 = b^2(\frac{1}{2}+\frac{1}{2}\frac{\sin\alpha}{\alpha}\cos\alpha - \frac{\sin^2\alpha}{\alpha^2})$$

$$\lambda_3 = \frac{a^2(\frac{1}{2}-\frac{1}{2}\frac{\sin\alpha}{\alpha}\cos\alpha) + c^2\frac{\alpha^2}{3} - \sqrt{(a^2(\frac{1}{2}-\frac{1}{2}\frac{\sin\alpha}{\alpha}\cos\alpha) - c^2\frac{\alpha^2}{3})^2 + 4a^2c^2(\cos\alpha - \frac{\sin\alpha}{\alpha})^2}}{2}.$$

It holds:

$$\lim_{\alpha\to 0} \frac{\lambda_1(\alpha)}{Var_{\pi_\alpha}\theta} = \lim_{\alpha\to 0} \frac{\lambda_1(\alpha)}{\frac{\alpha^2}{3}} = a^2 + c^2,$$

$$\lim_{\alpha\to 0} \frac{\lambda_2(\alpha)}{(Var_{\pi_\alpha}\theta)^2} = \lim_{\alpha\to 0} \frac{\lambda_2(\alpha)}{\frac{\alpha^4}{9}} = b^2.\frac{1}{5}$$

$$\lim_{\alpha\to 0} \frac{\lambda_3(\alpha)}{(Var_{\pi_\alpha}\theta)^3} = \lim_{\alpha\to 0} \frac{\lambda_3(\alpha)}{\frac{\alpha^6}{27}} = \frac{3}{25.7}\frac{a^2c^2}{a^2+c^2}.$$

Again

$$,,\lambda_1 > \lambda_2 > \lambda_3"$$

in the sense defined in example 1.

$$\lim_{\alpha\to 0} \frac{\sqrt{5}\sqrt{\lambda_{k+1}(\alpha)}}{\lambda_k} = \lim_{\alpha\to 0} \frac{\sqrt{5}\sqrt{\lambda_2(\alpha)}}{\lambda_1} = \frac{b}{a^2+c^2} = K_{\text{int}}(\theta^0).$$

Acknowledgement
The work was supported by Vega grants no. 2/4026/04 and 1/3016/06.

References

[1] Harville, D. A. (2000) *Matrix algebra from a statistician's perspective*. Springer, New York. Corrected third printing.

[2] Hornišová, K. (2004). *Intrinsic linearization of nonlinear regression by principal components method*. AMUC **LXXIII**, 207 – 216.

[3] Hornišová, K. (2004). *Linear and polynomial approximations of nonlinear regression models utilizing the prior distribution of the parameter*. Thesis, Faculty of mathematics, physics and informatics, Comenius University, Bratislava.

[4] Jenčová, A. (2000) *Linearization conditions for regression models with unknown variance parameter*. Appl. Math. **45**, 1, 61 – 78.

[5] Kubáček, L. (1995) *On a linearization of regression models*. Appl. Math. **40**, 1, 61 – 78.

[6] Pázman, A. (1992) *Geometry of the nonlinear regression with prior*. AMUC **LXI** 263 – 276.

[7] Pázman, A. (1993). *Nonlinear statistical models*. Kluwer, Dordrecht.

[8] Pázman, A. (2001). *Linearization of nonlinear regression models by smoothing*. Tatra Mt. Math. Publ. **22**, 13 – 25.