Applications of Mathematics

Anna Jenčová

A comparison of linearization and quadratization domains

Applications of Mathematics, Vol. 42 (1997), No. 4, 279--291

Persistent URL: http://dml.cz/dmlcz/134359

42 (1997) APPLICATIONS OF MATHEMATICS No. 4, 279–291

A COMPARISON OF LINEARIZATION AND QUADRATIZATION DOMAINS

Anna Jenčová, Bratislava

(Received June 10, 1996)

Abstract. In a nonlinear model, the linearization and quadratization domains are considered. In the case of a locally quadratic model, explicit expressions for these domains are given and the domains are compared.

Keywords: nonlinear regression models, linear and quadratic estimators, linearization and quadratization domains

MSC 2000: 62F10

Introduction

In [K1], the problem of linearization of nonlinear regression models is solved and the linearization domains are defined. Quadratization domains are indicated in [K2].

The domains mentioned are of several kinds (with respect to bias, dispersion, etc.). It is not quite clear which of these domains is the most important. Regarding some experience in the analysis of several simple nonlinear regression models (see [Pu]), the linearization domain for the bias seems to be important.

If the linearization domain for the bias is not sufficiently large, then two possibilities occur. Either a quadratic estimator, or a procedure given in [Pa] must be used. Since the quadratic estimator defined in [K2] is simple, it is of some importance to recognize whether the quadratization domain (for the bias) is sufficiently large, or at least, whether it contains the linearization domain.

The aim of the paper is to find conditions under which the above mentioned inclusion occurs. To do this, it is necessary to find a suitable expression for the characterization of the domains. This characterization is given in Definition 1.1, the conditions can be found in Proposition 2.1

1. Definition of the domains

Let Y be an n-dimensional normally distributed random vector with $E[Y] = f(\beta)$, $\operatorname{Var}[Y] = \Sigma$, where $f\colon \mathbb{R}^k \to \mathbb{R}^n$ is a known function with continuous second derivatives, $\beta \in \mathbb{R}^k$ is an unknown parameter and Σ is a known positively definite (p.d.) matrix. We assume that the true value of the parameter lies in a neighbourhood $\mathcal{O}$ of a chosen point $\beta_0 \in \mathbb{R}^k$ (we suppose below that $\beta_0 = 0$) and that within $\mathcal{O}$ the function f has the form

$$f(\beta) = f_0 + F\beta + \frac{1}{2}\kappa_\beta$$

where

$$\kappa_\beta = \begin{pmatrix} \beta' H_1 \beta \\ \vdots \\ \beta' H_n \beta \end{pmatrix},$$

$f_0 = f(\beta_0)$, $F = \frac{\partial f(\beta)}{\partial \beta'}|_{\beta=\beta_0}$ is a full rank matrix in columns, $H_i = \frac{\partial^2 f_i(\beta)}{\partial\beta\partial\beta'}$, $i = 1,\dots,n$.

Let us consider the linear estimator of the parameter $\hat\beta = C^{-1}F'\Sigma^{-1}(Y - f_0)$, where $C = F'\Sigma^{-1}F$, and the quadratic estimators

$$\hat{\hat\beta} = C_S^{-1}F'(\Sigma+S)^{-1}\left\{Y - f_0 + \frac{1}{2}\begin{pmatrix} \operatorname{Tr}(H_1C^{-1}) - \hat\beta' H_1\hat\beta \\ \vdots \\ \operatorname{Tr}(H_nC^{-1}) - \hat\beta' H_n\hat\beta \end{pmatrix}\right\},$$

$$\tilde\beta = \hat\beta + \frac{1}{2}C^{-1}F'\Sigma^{-1}\begin{pmatrix} \operatorname{Tr}(H_1C^{-1}) - \hat\beta' H_1\hat\beta \\ \vdots \\ \operatorname{Tr}(H_nC^{-1}) - \hat\beta' H_n\hat\beta \end{pmatrix}$$

defined in [K2]; here $C_S = F'(\Sigma+S)^{-1}F$ and $\{S\}_{i,j} = \frac{1}{2}\operatorname{Tr}(H_iC^{-1}H_jC^{-1})$. Let $\hat b(\beta)$, $\hat{\hat b}(\beta)$ and $\tilde b(\beta)$ be the bias of the estimators $\hat\beta$, $\hat{\hat\beta}$ and $\tilde\beta$, respectively.

In [K1] and [K2], the linearization domains (with respect to the bias) and, respectively, the quadratization domains are defined as sets $\mathcal{O} \subset \mathbb{R}^k$ such that

(i)

$$\forall h \in \mathbb{R}^k \qquad |h'\hat b(\beta)| \leqslant c\sqrt{h'C^{-1}h}$$

(ii)

$$\forall h \in \mathbb{R}^k \qquad |h'\hat{\hat b}(\beta)| \leqslant c\sqrt{h'C_S^{-1}h}$$

(iii)

$$\forall h \in \mathbb{R}^k \qquad |h'\tilde b(\beta)| \leqslant c\sqrt{h'C^{-1}h}$$

hold for $\beta \in \mathcal{O}$. Here c is a criterion parameter, chosen by a statistician. The aim is to find these domains in the form of an ellipsoid, i.e. in the form $\{\beta\colon \|\beta\|_C \leqslant M(c)\}$, where $\|x\|_C = \sqrt{x'Cx}$. We will use the Bates-Watts parameter effect and intrinsic curvatures (see [Pa] for more detail), defined as

$$K^{\rm par} = \sup_h \frac{\|P_F \kappa_h\|_{\Sigma^{-1}}}{\|Fh\|^2_{\Sigma^{-1}}}$$

and

$$K^{\rm int} = \sup_h \frac{\|M_F \kappa_h\|_{\Sigma^{-1}}}{\|Fh\|^2_{\Sigma^{-1}}},$$

where $\|x\|^2_{\Sigma^{-1}} = x'\Sigma^{-1}x$, $P_F = FC^{-1}F'\Sigma^{-1}$ is the orthogonal projector onto the subspace $\mathcal{M}(F) \subset (\mathbb{R}^n, \|.\|_{\Sigma^{-1}})$, spanned by the columns of the matrix F, and $M_F = I - P_F$, I is the $n \times n$ identical matrix. From [K1] it follows that $M(c) = \sqrt{\frac{2c}{K^{\rm par}}}$ for the linearization domain.

First, we prove some auxiliary statements.

Lemma 1.1.
(a) $\hat{b}(\beta) = -\frac{1}{2}C^{-1}F'\Sigma^{-1}\kappa_\beta$ *and* $\mathrm{Var}[\hat{\beta}] = C^{-1}$.
(b)

$$\hat{\hat{b}}(\beta) = -C_S^{-1}F'(\Sigma + S)^{-1} \begin{pmatrix} \beta' H_1 \hat{b}(\beta) + \frac{1}{2}\hat{b}(\beta)' H_1 \hat{b}(\beta) \\ \vdots \\ \beta' H_n \hat{b}(\beta) + \frac{1}{2}\hat{b}(\beta)' H_n \hat{b}(\beta) \end{pmatrix}$$

and $\mathrm{Var}[\hat{\hat{\beta}}|\beta = 0] = C_S^{-1}$.
(c)

$$\tilde{b}(\beta) = -C^{-1}F'\Sigma^{-1} \begin{pmatrix} \beta' H_1 \hat{b}(\beta) + \frac{1}{2}\hat{b}(\beta)' H_1 \hat{b}(\beta) \\ \vdots \\ \beta' H_n \hat{b}(\beta) + \frac{1}{2}\hat{b}(\beta)' H_n \hat{b}(\beta) \end{pmatrix}$$

and $\mathrm{Var}[\tilde{\beta}|\beta = 0] = C^{-1} + C^{-1}F'\Sigma^{-1}S\Sigma^{-1}FC^{-1}$.

The proof of (b) and (c) can be found in [2], the proof of (a) is easy.

Lemma 1.2. *If* $\|\beta\|_C \leqslant M$, *then* $\|\hat{b}(\beta)\|_C \leqslant \frac{1}{2}K^{\rm par}M^2$.

P r o o f.

$$\sup_{\|\beta\|_C \leqslant M} \|\hat{b}(\beta)\|_C = \frac{1}{2} \sup_{\|\beta\|_C \leqslant M} \frac{\sqrt{\kappa'_\beta \Sigma^{-1}F'C^{-1}F\Sigma^{-1}\kappa_\beta}}{\beta' C \beta} \beta' C \beta$$
$$\leqslant \frac{1}{2} \sup_\beta \frac{\sqrt{\kappa'_\beta \Sigma^{-1}F'C^{-1}F\Sigma^{-1}\kappa_\beta}}{\beta' C \beta} M^2 = \frac{1}{2}K^{\rm par}M^2.$$

□

Lemma 1.3. *Let A be a symmetric and W a p.d. $k \times k$ matrix. Then*

$$\sup_{\|x\|_W=1, \|y\|_W=1} |x'Ay| \leqslant 2 \sup_{\|x\|_W=1} |x'Ax|.$$

P r o o f. First, let A be p.s.d. Then $A = BB'$ for some B and therefore

$$\begin{aligned}
\sup_{\|x\|_W=1, \|y\|_W=1} |x'Ay| &= \sup_{\|x\|_W=1, \|y\|_W=1} |x'BB'y| \\
&\leqslant \sup_{\|x\|_W=1, \|y\|_W=1} \sqrt{x'BB'xy'BB'y} \\
&= \sup_{\|x\|_W=1} x'Ax \\
&\leqslant 2 \sup_{\|x\|_W=1} x'Ax.
\end{aligned}$$

Now let A be an arbitrary symmetric matrix. Then there exists a nonsingular matrix R such that $R'AR = \Lambda = \operatorname{diag}(\lambda_1, \ldots, \lambda_k)$ and $W = (R^{-1})'R^{-1}$. Let $\lambda_1 \geqslant \ldots \geqslant \lambda_s \geqslant 0$ and $0 \geqslant \lambda_{s+1} \geqslant \ldots \geqslant \lambda_k$ and let us denote $z = R^{-1}x$, $u = R^{-1}y$, $\Lambda_1 = \operatorname{diag}(\lambda_1, \ldots, \lambda_s)$, $\Lambda_2 = \operatorname{diag}(\lambda_{s+1}, \ldots, \lambda_k)$. Then Λ_1 and $(-\Lambda_2)$ are obviously p.s.d. and

$$\begin{aligned}
\sup_{\|x\|_W=1, \|y\|_W=1} |x'Ay| &= \sup_{\|z\|_I=1, \|u\|_I=1} |z'\Lambda u| \\
&\leqslant \sup_{\|z\|_I=1, \|u\|_I=1} |z'\Lambda_1 u| + \sup_{\|z\|_I=1, \|u\|_I=1} |z'(-\Lambda_2)u| \\
&\leqslant \sup_{\|z\|_I=1} z'\Lambda_1 z + \sup_{\|z\|_I=1} z'(-\Lambda_2)z \\
&= \sup_{\|z\|_I=1, z_{s+1}=\ldots=z_k=0} |z'\Lambda z| + \sup_{\|z\|_I=1, z_1=\ldots=z_s=0} |z'\Lambda z| \\
&\leqslant \sup_{\|x\|_W=1} |x'Ax| + \sup_{\|x\|_W=1} |x'Ax| \\
&= 2 \sup_{\|x\|_W=1} |x'Ax|.
\end{aligned}$$

□

Lemma 1.4. *If $\|\beta\|_C \leqslant M$ then $\|\tilde{b}(\beta)\|_C \leqslant (K^{\mathrm{par}})^2 M^3 + \frac{1}{8}(K^{\mathrm{par}})^3 M^4$.*

P r o o f. According to Lemma 1.1 (c),

$$\begin{aligned}
&\sup_{\|\beta\|_C \leqslant M} \|\tilde{b}(\beta)\|_C \\
&\quad = \sup_{\|\beta\|_C \leqslant M} \left\| -C^{-1}F'\Sigma^{-1} \begin{pmatrix} \beta' H_1 \hat{b}(\beta) + \frac{1}{2}\hat{b}(\beta)' H_1 \hat{b}(\beta) \\ \vdots \\ \beta' H_n \hat{b}(\beta) + \frac{1}{2}\hat{b}(\beta)' H_n \hat{b}(\beta) \end{pmatrix} \right\|_C \\
&\quad \leqslant \sup_{\|\beta\|_C \leqslant M} \left\| -C^{-1}F'\Sigma^{-1} \begin{pmatrix} \beta' H_1 \hat{b}(\beta) \\ \vdots \\ \beta' H_n \hat{b}(\beta) \end{pmatrix} \right\|_C \\
&\qquad + \sup_{\|\beta\|_C \leqslant M} \frac{1}{2} \left\| -C^{-1}F'\Sigma^{-1} \begin{pmatrix} \hat{b}(\beta)' H_1 \hat{b}(\beta) \\ \vdots \\ \hat{b}(\beta)' H_n \hat{b}(\beta) \end{pmatrix} \right\|_C .
\end{aligned}$$

The expression $\hat{b}(\beta) = -\frac{1}{2}C^{-1}F'\Sigma^{-1}\kappa_\beta$ can be viewed as a function of β, hence

$$-\frac{1}{2}C^{-1}F'\Sigma^{-1} \begin{pmatrix} \hat{b}(\beta)' H_1 \hat{b}(\beta) \\ \vdots \\ \hat{b}(\beta)' H_n \hat{b}(\beta) \end{pmatrix} = \hat{b}(\hat{b}(\beta)).$$

According to Lemma 1.2,

$$\sup_{\|\beta\|_C \leqslant M} \|\hat{b}(\hat{b}(\beta))\|_C \leqslant \frac{1}{2}K^{\text{par}}\Big(\frac{1}{2}K^{\text{par}}M^2\Big)^2 = \frac{1}{8}(K^{\text{par}})^3 M^4.$$

Further, let $K^{(l)} = \sum_{i=1}^{n} \{l'C^{-1}F'\Sigma^{-1}\}_i H_i$, $l \in \mathbb{R}^k$. By the Schwarz inequality we have

$$\begin{aligned}
\left\| -C^{-1}F'\Sigma^{-1} \begin{pmatrix} \beta' H_1 b \\ \vdots \\ \beta' H_n b \end{pmatrix} \right\|_C &= \sup_{\|C^{-1}l\|_C = 1} \left| l'C^{-1}F'\Sigma^{-1} \begin{pmatrix} \beta' H_1 b \\ \vdots \\ \beta' H_n b \end{pmatrix} \right| \\
&= \sup_{\|C^{-1}l\|_C = 1} |\beta' K^{(l)} b|.
\end{aligned}$$

Hence, using Lemma 1.3, we obtain

$$
\begin{aligned}
&\sup_{\|\beta\|_C \leqslant M} \left\| -C^{-1}F'\Sigma^{-1} \begin{pmatrix} \beta' H_1 \hat{b}(\beta) \\ \vdots \\ \beta' H_n \hat{b}(\beta) \end{pmatrix} \right\|_C \\
&\quad = \sup_{\|\beta\|_C \leqslant M} \left\| -C^{-1}F'\Sigma^{-1} \begin{pmatrix} \beta' H_1 \hat{b}(\beta) \\ \vdots \\ \beta' H_n \hat{b}(\beta) \end{pmatrix} \right\|_C \frac{\|\beta\|_C \|\hat{b}(\beta)\|_C}{\|\beta\|_C \|\hat{b}(\beta)\|_C} \\
&\quad \leqslant \sup_{\|\beta\|_C \leqslant M} \frac{\left\| -C^{-1}F'\Sigma^{-1} \begin{pmatrix} \beta' H_1 \hat{b}(\beta) \\ \vdots \\ \beta' H_n \hat{b}(\beta) \end{pmatrix} \right\|_C}{\|\beta\|_C \|\hat{b}(\beta)\|_C} \frac{1}{2} K^{\mathrm{par}} M^3 \\
&\quad \leqslant \sup_{\|\beta\|_C = 1, \|b\|_C = 1} \left\| -C^{-1}F'\Sigma^{-1} \begin{pmatrix} \beta' H_1 b \\ \vdots \\ \beta' H_n b \end{pmatrix} \right\|_C \frac{1}{2} K^{\mathrm{par}} M^3 \\
&\quad = \sup_{\|\beta\|_C = 1, \|b\|_C = 1} \sup_{\|C^{-1}l\|_C = 1} |\beta' K^{(l)} b| \frac{1}{2} K^{\mathrm{par}} M^3 \\
&\quad \leqslant 2 \sup_{\|\beta\|_C = 1} \sup_{\|C^{-1}l\|_C = 1} |\beta' K^{(l)} \beta| \frac{1}{2} K^{\mathrm{par}} M^3 \\
&\quad = \sup_{\|\beta\|_C = 1} \frac{\left\| -C^{-1}F'\Sigma^{-1} \begin{pmatrix} \beta' H_1 \beta \\ \vdots \\ \beta' H_n \beta \end{pmatrix} \right\|_C}{\beta' C \beta} M^3 K^{\mathrm{par}} = M^3 (K^{\mathrm{par}})^2
\end{aligned}
$$

□

Lemma 1.5. *Let* $R^{(x)} = \sum_i H_i x_i$, $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$. *Then*

$$\frac{1}{2} \operatorname{Tr}(R^{(x)} C^{-1} R^{(x)} C^{-1}) = x' S x.$$

P r o o f.

$$
\begin{aligned}
\frac{1}{2} \operatorname{Tr}(R^{(x)} C^{-1} R^{(x)} C^{-1}) &= \frac{1}{2} \operatorname{Tr} \left(\sum_i x_i H_i C^{-1} \sum_j x_j H_j C^{-1} \right) \\
&= \frac{1}{2} \operatorname{Tr} \left(\sum_i \sum_j x_i x_j H_i C^{-1} H_j C^{-1} \right) \\
&= \sum_i \sum_j x_i x_j \frac{1}{2} \operatorname{Tr}(H_i C^{-1} H_j C^{-1}) = x' S x.
\end{aligned}
$$

□

Lemma 1.6. *Let A be a p.s.d. and B a p.d. $k \times k$ matrix. Then*

$$\sup_x \frac{x'Ax}{x'B^{-1}x} \leqslant \operatorname{Tr}(AB) \leqslant k \sup_x \frac{x'Ax}{x'B^{-1}x}.$$

P r o o f. Let L be a nonsingular matrix such that $B = LL'$. Then

$$\operatorname{Tr}(AB) = \operatorname{Tr}(ALL') = \operatorname{Tr}(L'AL) = \sum \eta_i,$$

where $\eta_1 \geqslant \ldots \geqslant \eta_k$ are the eigenvalues of the matrix $L'AL$. Further,

$$\sup_x \frac{x'Ax}{x'B^{-1}x} = \sup_x \frac{x'Ax}{x'(L')^{-1}L^{-1}x} = \sup_y \frac{y'L'ALy}{y'y} = \eta_1,$$

where $y = L^{-1}x$. Obviously $\eta_i \geqslant 0$, $i = 1, \ldots, k$, thus

$$\eta_1 \leqslant \sum \eta_i \leqslant k\eta_1.$$

□

Lemma 1.7.

$$(\Sigma + S)^{-1} = \sum_{i=1}^{n} \frac{1}{1+\lambda_i} p_i p_i',$$

$\sum p_i p_i' = \Sigma^{-1}$, and $\forall i \lambda_i \leqslant d_1$, where

$$d_1 = k^2((K^{\text{int}})^2 + (K^{\text{par}})^2).$$

P r o o f. Let $\Sigma = UU'$. Then

$$U^{-1}(\Sigma + S)(U')^{-1} = I + U^{-1}S(U')^{-1}.$$

Let $\lambda_1 \geqslant \ldots \geqslant \lambda_n \geqslant 0$ be the eigenvalues of the matrix $U^{-1}S(U')^{-1}$, and q_i, $i = 1, \ldots, n$ the corresponding eigenvectors. Then

$$I + U^{-1}S(U')^{-1} = \sum_i (1+\lambda_i) q_i q_i'$$

and

$$\begin{aligned} U'(\Sigma + S)^{-1}U &= (U^{-1}(\Sigma + S)(U')^{-1})^{-1} \\ &= (I + U^{-1}S(U')^{-1})^{-1} = \sum_i (1+\lambda_i)^{-1} q_i q_i'. \end{aligned}$$

Hence

$$(\Sigma + S)^{-1} = \sum_i (1+\lambda_i)^{-1}(U')^{-1} q_i q_i' U^{-1}.$$

Now, put $p_i = (U')^{-1} q_i$ and

$$\sum_i p_i p_i' = (U')^{-1}\Big[\sum_i q_i q_i'\Big]U^{-1} = (U')^{-1}U^{-1} = \Sigma^{-1}.$$

Moreover, $\lambda_i = q_i' U^{-1} S (U')^{-1} q_i$, $q_i' q_i = 1$. If z_i is a vector such that $q_i = U^{-1} z_i$, then

$$\lambda_i = z_i'(U')^{-1}U^{-1}S(U')^{-1}U^{-1}z_i = z_i'\Sigma^{-1}S\Sigma^{-1}z_i$$

and $z_i'\Sigma^{-1}z_i = q_i'q_i = 1$. Let $x \in \mathbb{R}^n$ be an arbitrary vector such that $\|x\|_{\Sigma^{-1}} = 1$. Using Lemmas 1.3, 1.5 and 1.6 we get

$$\begin{aligned}
x'\Sigma^{-1}S\Sigma^{-1}x &= \frac{1}{2}\operatorname{Tr}(R^{(\Sigma^{-1}x)}C^{-1}R^{(\Sigma^{-1}x)}C^{-1})\\
&\leqslant \frac{1}{2}k\sup_y \frac{y'R^{(\Sigma^{-1}x)}C^{-1}R^{(\Sigma^{-1}x)}y}{y'Cy} = \frac{1}{2}k\sup_y \frac{\operatorname{Tr}(R^{(\Sigma^{-1}x)}yy'R^{(\Sigma^{-1}x)}C^{-1})}{y'Cy}\\
&\leqslant k^2 \sup_y\sup_z \frac{1}{2}\frac{(z'R^{(\Sigma^{-1}x)}y)^2}{y'Cyz'Cz} = \sup_{\|y\|_C=1}\ \sup_{\|z\|_C=1} \frac{k^2}{2}(z'R^{(\Sigma^{-1}x)}y)^2\\
&\leqslant k^2 \sup_{\|y\|_C=1}(y'R^{(\Sigma^{-1}x)}y)^2 = k^2\sup_y \frac{(y'R^{(\Sigma^{-1}x)}y)^2}{(y'Cy)^2}\\
&= k^2\sup_y\frac{(x'\Sigma^{-1}\kappa_y)^2}{(y'Cy)^2} \leqslant k^2\sup_y \frac{\kappa_y'\Sigma^{-1}\kappa_y}{(y'Cy)^2}\\
&\leqslant k^2\Big\{\sup_y \frac{\kappa_y'\Sigma^{-1}P_F\kappa_y}{(y'Cy)^2} + \sup_y\frac{\kappa_y'\Sigma^{-1}M_F\kappa_y}{(y'Cy)^2}\Big\}\\
&= k^2((K^{\text{int}})^2 + (K^{\text{par}})^2).
\end{aligned}$$

□

Lemma 1.8. *Let $\|x\|_1 = x'\Sigma^{-1}x$ and $\|x\|_2 = x'(\Sigma+S)^{-1}x$. Let $d = \frac{1}{1+d_1}$. Then*

$$d\|x\|_1^2 \leqslant \|x\|_2^2 \leqslant \|x\|_1^2$$

Proof. According to Lemma 1.7,

$$\|x\|_1^2 = \sum_i x'p_ip_i'x,$$

$$\|x\|_2^2 = \sum_i \frac{1}{1+\lambda_i}x'p_ip_i'x$$

and $d \leqslant \frac{1}{1+\lambda_i} \leqslant 1$. The rest of the proof is obvious. □

Lemma 1.9. *If* $\|\beta\|_C \leqslant M$, *then* $\|\hat{\hat{b}}(\beta)\|_{C_S} \leqslant K_S^{\rm par}(K^{\rm par}M^3 + \frac{1}{8}(K^{\rm par})^2 M^4)$, *where* $K_S^{\rm par}$ *is the parameter effect curvature for the inner product* $\langle x, x\rangle = x'(\Sigma + S)^{-1}x$.

P r o o f. The statement of the lemma can be proved similarly as Lemma 1.4, using the fact that according to Lemma 1.8

$$\begin{aligned}\beta' C_S \beta = \beta' F'(\Sigma + S)^{-1} F\beta = \|F\beta\|^2_{(\Sigma+S)^{-1}} \\ \leqslant \|F\beta\|^2_{\Sigma^{-1}} = \beta' C\beta.\end{aligned}$$

□

Proposition 1.1.

(a) *If* $\|\beta\|_C \leqslant \tilde{M}(c)$ *where* $\tilde{M}(c) > 0$ *is the solution of*

$$(K^{\rm par})^2 M^3 + \frac{1}{8}(K^{\rm par})^3 M^4 = c,$$

then $|h'\tilde{b}(\beta)| \leqslant c\sqrt{h'C^{-1}h}\ \forall h \in \mathbb{R}^k$.

(b) *If* $\|\beta\|_C \leqslant \hat{\hat{M}}(c)$ *where* $\hat{\hat{M}}(c) > 0$ *is the solution of*

$$K_S^{\rm par}\Big(K^{\rm par}M^3 + \frac{1}{8}(K^{\rm par})^2 M^4\Big) = c,$$

then

$$|h'\hat{\hat{b}}(\beta)| \leqslant c\sqrt{h'C_S^{-1}h} \quad \forall h \in \mathbb{R}^k.$$

P r o o f. (a) The equivalence

$$|h'b| \leqslant c\sqrt{h'C^{-1}h}\ \forall h \quad \Leftrightarrow \quad \|b\|_C \leqslant c$$

following from the Schwarz inequality, will be used. Now, by Lemma 1.4, for $\|\beta\|_C \leqslant \tilde{M}(c)$ we have

$$\|\tilde{b}(\beta)\|_C \leqslant (K^{\rm par})^2 \tilde{M}(c)^3 + \frac{1}{8}(K^{\rm par})^3 \tilde{M}(c)^4 = c.$$

(b) The same as (a). □

Now, we can define the linearization and quadratization domains as follows:

Definition 1.1. The linearization domain is

$$\hat{\mathcal{O}}(c) = \Big\{\beta\colon \ \|\beta\|_C \leqslant \hat{M}(c) = \sqrt{\frac{2c}{K^{\rm par}}}\Big\}.$$

The quadratization domains are

$$\tilde{\mathcal{O}}(c) = \{\beta\colon \ \|\beta\|_C \leqslant \tilde{M}(c)\}$$

and

$$\hat{\hat{\mathcal{O}}}(c) = \{\beta\colon \ \|\beta\|_C \leqslant \hat{\hat{M}}(c)\}.$$

2. Linearization vs. quadratization

Now we have an expression for the domains, which allows us to compare easily the linearization and quadratization domains.

Proposition 2.1. *If* $0 < c < \frac{2(9-4\sqrt{5})}{K^{\rm par}}$, *then* $\hat{\mathcal{O}}(c) \subset \tilde{\mathcal{O}}(c)$.

P r o o f. It is obvious that $\hat{\mathcal{O}}(c) \subset \tilde{\mathcal{O}}(c)$ iff $\hat{M}(c) \leqslant \tilde{M}(c)$. The function $p(M) = (K^{\rm par})^2M^3 + \frac{1}{8}(K^{\rm par})^3M^4$ is increasing for $M > 0$, i.e. $\forall M$, $M \leqslant \tilde{M}(c) \ \Leftrightarrow \ p(M) \leqslant p(\tilde{M}(c)) = c$. Therefore it suffices to prove that $p(\hat{M}(c)) \leqslant c$ for $0 < c < \frac{2(9-4\sqrt{5})}{K^{\rm par}}$. Moreover,

$$p(\sqrt{2c/K^{\rm par}}) \leqslant c \ \Leftrightarrow \ 2\sqrt{2K^{\rm par}c} + \frac{1}{2}K^{\rm par}c \leqslant 1.$$

The last inequality implies that $1 - \frac{1}{2}K^{\rm par}c \geqslant 0$, i.e. $c \leqslant \frac{2}{K^{\rm par}}$. Further,

$$2\sqrt{2K^{\rm par}c} \leqslant 1 - \frac{1}{2}K^{\rm par}c \Leftrightarrow \frac{1}{4}(K^{\rm par})^2c^2 - 9K^{\rm par}c + 1 \geqslant 0.$$

The last inequality holds for $c \leqslant \frac{2(9-4\sqrt{5})}{K^{\rm par}}$. □

R e m a r k 2.1. It follows from the last proposition that if $K^{\rm par} \leqslant 2(9-4\sqrt{5}) \doteq 0.11$, the quadratization domain $\tilde{\mathcal{O}}$ is larger than the linearization domain for all $0 < c < 1$.

It is clear that the estimator $\tilde{\beta}$ is more suitable for practical purposes than $\hat{\hat{\beta}}$, because it is simpler and has a more convenient form $\hat{\beta}+$ correction terms. Moreover, we will see that if the nonlinearity of the model is low, the domains $\tilde{\mathcal{O}}$ and $\hat{\hat{\mathcal{O}}}$ seem to be nearly the same.

Lemma 2.1. *Let* $d = \frac{1}{1+d_1}$, *where* $d_1 = k^2((K^{\mathrm{int}})^2 + (K^{\mathrm{par}})^2)$. *Then*

$$d(K^{\mathrm{par}})^2 - (1-d)(K^{\mathrm{int}})^2 \leqslant (K_S^{\mathrm{par}})^2 \leqslant \frac{1}{d^2}((K^{\mathrm{par}})^2 + (1-d)(K^{\mathrm{int}})^2).$$

P r o o f. Let P_1 and P_2 be the orthogonal projectors onto the subspace $\mathcal{F}$ spanned by the columns of the matrix F in the space $(\mathbb{R}^n, \|\cdot\|_1)$ and $(\mathbb{R}^n, \|\cdot\|_2)$, respectively; let $M_1 = I - P_1$, $M_2 = I - P_2$. Then

$$K^{\mathrm{par}} = \sup_h \frac{\|P_1\kappa_h\|_1}{\|Fh\|_1^2}$$

and

$$K_S^{\mathrm{par}} = \sup_h \frac{\|P_2\kappa_h\|_2}{\|Fh\|_2^2}.$$

Let P be an orthogonal projector matrix in the space $(\mathbb{R}^n, \|\cdot\|)$ onto a subspace $\mathcal{L} \subset \mathbb{R}^n$. Then for each $x \in \mathbb{R}^n$, $\|x\|^2 = \|Px\|^2 + \|Mx\|^2$ and $\|Mx\|^2 = \|x - Px\|^2 = \inf\limits_{y\in\mathcal{L}} \|x-y\|^2$. Thus, by Lemma 1.8,

$$\|M_2x\|_2^2 = \inf_{y\in\mathcal{F}} \|x-y\|_2^2 \leqslant \inf_{y\in\mathcal{F}} \|x-y\|_1^2 = \|M_1x\|_1^2$$

and

$$\|M_2x\|_2^2 \geqslant d \inf_{y\in\mathcal{F}} \|x-y\|_1^2 = d\|M_1x\|_1^2.$$

Moreover,

$$\begin{aligned}\|P_2x\|_2^2 = \|x\|_2^2 - \|M_2x\|_2^2 \leqslant \|x\|_1^2 - d\|M_1x\|_1^2 \\ = \|P_1x\|_1^2 + (1-d)\|M_1x\|_1^2\end{aligned}$$

and

$$\|P_2x\|_2^2 \geqslant d\|x\|_1^2 - \|M_1x\|_1^2 = d\|P_1x\|_1^2 - (1-d)\|M_1x\|_1^2.$$

Hence

$$\frac{\|P_2\kappa_h\|_2^2}{\|Fh\|_2^4} \leqslant \frac{1}{d^2\|Fh\|_1^4}(\|P_1\kappa_h\|_1^2 + (1-d)\|M_1\kappa_h\|_1^2)$$

and thus

$$(K_S^{\mathrm{par}})^2 \leqslant \frac{1}{d^2}((K^{\mathrm{par}})^2 + (1-d)(K^{\mathrm{int}})^2).$$

On the other hand,

$$\frac{\|P_2\kappa_h\|_2^2}{\|Fh\|_2^4} \geqslant d\frac{\|P_1\kappa_h\|_1^2}{\|Fh\|_1^4} - (1-d)\frac{\|M_1\kappa_h\|_1^2}{\|Fh\|_1^4}$$

and

$$\sup \frac{\|P_2\kappa_h\|_2^2}{\|Fh\|_2^4} \geqslant d\frac{\|P_1\kappa_h\|_1^2}{\|Fh\|_1^4} - (1-d)\sup \frac{\|M_1\kappa_h\|_1^2}{\|Fh\|_1^4},$$

hence

$$(K_S^{\mathrm{par}})^2 \geqslant d(K^{\mathrm{par}})^2 - (1-d)(K^{\mathrm{int}})^2.$$

□

R e m a r k 2.2. We see that the first inequality in Lemma 2.1 makes sense only if

$$d(K^{\mathrm{par}})^2 - (1-d)(K^{\mathrm{int}})^2 \geqslant 0,$$

i.e.

$$(K^{\mathrm{par}})^2 \geqslant d_1(K^{\mathrm{int}})^2 = k^2((K^{\mathrm{int}})^2 + (K^{\mathrm{par}})^2)(K^{\mathrm{int}})^2.$$

It follows that $K^{\mathrm{int}} < \frac{1}{k}$ and $(K^{\mathrm{par}})^2 \geqslant \frac{k^2(K^{\mathrm{int}})^4}{1-k^2(K^{\mathrm{int}})^2}$. We suppose below that both these conditions are fulfilled.

Proposition 2.2. *Let*

$$r^2 = c^2d^2\frac{1}{1+(1-d)(\frac{K^{\mathrm{int}}}{K^{\mathrm{par}}})^2}$$

and

$$q^2 = c^2\frac{1}{d-(1-d)(\frac{K^{\mathrm{int}}}{K^{\mathrm{par}}})^2}.$$

Then $\tilde{\mathcal{O}}(r) \subseteq \hat{\tilde{\mathcal{O}}}(c) \subseteq \tilde{\mathcal{O}}(q)$.

P r o o f. Let $p(M) = K^{\mathrm{par}}M^3 + \frac{1}{8}(K^{\mathrm{par}})^2M^4$. It is sufficient to prove that $K_S^{\mathrm{par}}p(\tilde{M}(r)) \leqslant c$ and $K_S^{\mathrm{par}}p(\tilde{M}(q)) \geqslant c$. But according to Lemma 2.1,

$$(K_S^{\mathrm{par}}p(\tilde{M}(r)))^2 = r^2(\frac{K_S^{\mathrm{par}}}{K^{\mathrm{par}}})^2 \leqslant \frac{r^2}{d^2}\left(1+(1-d)\Big(\frac{K^{\mathrm{int}}}{K^{\mathrm{par}}}\Big)^2\right) = c^2$$

and

$$K_S^{\mathrm{par}}p(\tilde{M}(q)) \geqslant q^2\left(d-(1-d)\Big(\frac{K^{\mathrm{int}}}{K^{\mathrm{par}}}\Big)^2\right) = c^2.$$

It is obvious how to complete the proof. □

References

[K1] *L. Kubáček*: On a linearization of regression models. Applications of Mathematics *40* (1995), 61–78.

[K2] *L. Kubáček*: Quadratic regression models. Math. Slovaca *46* (1996), 111–126.

[Pa] *A. Pázman*: Nonlinear Statistical Models. Kluwer Acad. Publishers, Dordrecht-Boston-London and Ister Science Press, Bratislava, 1993.

[Pu] *A. Pulmannová*: Linear estimates in regression models. Diploma thesis, Faculty of Mathematics and Physics, Commenius University, Bratislava, 1994.

Author's address: *Anna Jenčová*, Mathematical Institute of the Slovac Academy of Sciences, Štefánikova 49, 814 73 Bratislava, Slovakia.