Lubomír Kubáček; Ludmila Kubáčková; Eva Tesaříková

Exponential regression

Acta Universitatis Palackianae Olomucensis. Facultas Rerum Naturalium. Mathematica, Vol. 40 (2001), No. 1, 125--149

Persistent URL: http://dml.cz/dmlcz/120426

Acta Univ. Palacki. Olomuc., Fac. rer. nat.,
Mathematica **40** (2001) 125–149

Exponential Regression *

Lubomír KUBÁČEK [1], Ludmila KUBÁČKOVÁ [1], Eva TESAŘÍKOVÁ [2]

[1] *Department of Mathematical Analysis and Applications of Mathematics*
Faculty of Science, Palacký University
Tomkova 40, 779 00 Olomouc, Czech Republic
e-mail: kubacekl@risc.upol.cz
[2] *Department of Algebra and Geometry, Faculty of Science, Palacký University*
Tomkova 40, 779 00 Olomouc, Czech Republic
e-mail: tesariko@risc.upol.cz

(Received January 3, 2001)

Abstract

In nonlinear regression models several procedures have been used in order to simplify an estimation of regression parameters. A comparison of different approach is demostrated in the case of exponential regression.

Key words: Nonlinear regression model, linearization, quadratization, nonlinear least squares estimator, design of experiment.

2000 Mathematics Subject Classification: 62J05, 62F10

1 Introduction

Solution of basic statistical problems are essentially more complicated in nonlinear regression models than in linear ones. Therefore several approaches were developed how to solve these problems by linear methods. It is to be mentioned two important approaches, i.e. a transformation of the measured data and a linearization of the model by the first term of Taylor series.

The aim of the paper is to compare different approaches in the case of an exponential regression $y = \beta_1 \exp(-\beta_2 x)$.

In the first step let us study this problem in an extremaly simple case in order to demonstrate a behaviour of a linear, quadratic and nonlinear least squares estimators.

*Supported by the grant No. 201/99/0327 of the Grant Agency of the Czech Republic and by the Council of the Czech Government J 14/98: 153 100011.

2 Motivation example

Let us consider the model $Y \sim N_1(\beta^2, \sigma^2)$, $\beta \in R^1$, where the dispersion σ^2 is known. The linearized version at the point β_0 of the model is $Y - \beta_0^2 \sim N_1(2\beta_0\delta\beta, \sigma^2)$, $\delta\beta \in R^1$ and the quadratized version $Y - \beta_0^2 \sim N_1(2\beta_0\delta\beta + \delta\beta^2, \sigma^2)$.

The $BLUE$ (best linear unbiased estimator) of β in the linearized model is

$$\hat{\beta} = \beta_0 + \delta\hat{\beta},$$
$$\delta\hat{\beta} = \frac{Y - \beta_0^2}{2\beta_0} \sim N_1\left(\delta\beta + \frac{\delta\beta^2}{2\beta_0}, \frac{\sigma^2}{4\beta_0^2}\right).$$

Thus MSE (mean square error) of this estimator is

$$MSE(\hat{\beta}) = \operatorname{var}(\hat{\beta}) + (\text{bias})^2(\hat{\beta}) = \frac{\sigma^2}{4\beta_0^2} + \frac{\delta\beta^4}{4\beta_0^2}.$$

The Bates and Watts [2] measures of nonlinearity in this case are

$$K^{(int)}(\beta_0) = 0, \qquad K^{(par)}(\beta_0) = \frac{\sigma}{2\beta_0^2}.$$

If $\delta\beta^2 C \leq \frac{2c}{K^{(par)}(\beta_0)}$, then $b(\delta\hat{\beta}) \leq c\sqrt{C^{-1}} = c\sqrt{\operatorname{var}(\delta\hat{\beta})}$, where $C = 4\beta_0^2/\sigma^2$ (the "matrix" of the normal equation). In detail cf. [17] and Statement 3.3.3. Let $c = \frac{1}{2}$; thus

$$|\delta\beta| \leq \frac{1}{\sqrt{CK^{(par)}(\beta_0)}} = \sqrt{\sigma/2} = h_{lin.} \Rightarrow |b(\delta\hat{\beta})| \leq \frac{1}{2}\sqrt{\operatorname{var}(\delta\hat{\beta})},$$

$$b(\delta\hat{\beta}) = E(\delta\hat{\beta}) - \delta\beta = \frac{\delta\beta^2}{2\beta_0}, \qquad \operatorname{var}(\delta\hat{\beta}) = \frac{\sigma^2}{4\beta_0^2}.$$

The 0.95-confidence interval is

$$\{\delta\beta : |\delta\hat{\beta} - \delta\beta| \leq 1.96\,\sigma/(2|\beta_0|) = h_{conf.}\}.$$

As a consequence of the mentioned relations we obtain

$$\sqrt{bCb} = \delta\beta^2/\sigma.$$

A corrected quadratic estimator is

$$\tilde{\beta} = \beta_0 + \delta\tilde{\beta}, \qquad \delta\tilde{\beta} = \delta\hat{\beta} - \frac{\delta\hat{\beta}^2}{2\beta_0} + \frac{\sigma^2}{8\beta_0^3}$$

and

$$E(\delta\tilde{\beta}) - \delta\beta = -\frac{\delta\beta^3}{2\beta_0^2} - \frac{\delta\beta^4}{8\beta_0^3},$$
$$\operatorname{var}(\delta\tilde{\beta}) = \frac{\sigma^2}{4\beta_0^2} + \frac{\sigma^4}{32\beta_0^6} - \frac{\sigma^2\delta\beta}{\beta_0^3} + \frac{\sigma^2\delta\beta^3}{4\beta_0^5} + \frac{\sigma^2\delta\beta^4}{16\beta_0^6}.$$

In the following tables a comparison of the variances, biases and MSEs of the estimators $\delta\hat{\beta}$ and $\delta\tilde{\beta}$ is given.

σ	$K^{(par)}$	$h_{lin.}$	$h_{conf.}$	$\delta\beta$	$\mathrm{var}(\delta\hat{\beta})$	$b(\delta\hat{\beta})$	$\sqrt{MSE(\delta\hat{\beta})}$
0.1	0.0005	0.224	0.00098	−0.01	0.000025	0.000005	0.005
				−0.30	0.000025	0.004500	0.0067
1.0	0.005	0.707	0.098	−0.1	0.00250	0.000500	0.050
				−0.7	0.00250	0.0245	0.0557
10.0	0.05	2.236	0.980	−1.0	0.250	0.050	0.5025
				−2.5	0.250	0.3125	0.58963
				−5.0	0.250	1.250	1.34629
100.0	0.5	7.071	9.8	−2.0	25	0.2	5.0040
				−7.0	25	2.45	5.56799
				−10.0	25	5	7.07106

Table 2.1
Variances and biases of linear estimators ($\beta_0 = 10$)

σ	$K^{(par)}$	$h_{lin.}$	$h_{conf.}$	$\delta\beta$	$\mathrm{var}(\delta\tilde{\beta})$	$b(\delta\tilde{\beta})$	$\sqrt{MSE(\delta\tilde{\beta})}$
0.1	0.0005	0.224	0.00098	−0.01	0.000025	0.00000	0.005
				−0.30	0.000026	0.000136	0.00515
1.0	0.005	0.707	0.098	−0.1	0.00255	0.000005	0.0505
				−0.7	0.00285	0.00174	0.05341
10.0	0.05	2.236	0.980	−1.0	0.30057	0.005125	0.54826
				−2.5	0.37952	0.08301	0.62162
				−5.0	0.47297	0.70312	0.98354
100.0	0.5	7.071	9.8	−2.0	37.935	0.042	6.15928
				−7.0	56.050	2.015	7.7531
				−10.0	59.375	6.25	9.9217

Table 2.2
Variances and biases of quadratic estimators ($\beta_0 = 10$)

It is to be referred to relations among $h_{lin.}$, $h_{conf.}$ and $\delta\beta$. As far as the parametric measure of nonlinearity $K^{(par)}$ is relatively small (except the last row of the tables) and $\delta\beta$ is inside the linearization region and at the same time inside the confidence interval, then the quadratic corrections is practically of no use and the linearized model gives practically the same results as the nonlinear least squares estimator as is shown in the following.

The nonlinear least squares estimator is

$$\hat{\hat{\beta}} = \sqrt{Y}.$$

If $\sqrt{\mathrm{var}(Y)}/E(Y) = \sigma/\beta^2 < 0.1$ we can use the approximate formula for the bias and the variance of $\hat{\hat{\beta}}$ (in detail [9])

$$E(\hat{\hat{\beta}}) = \beta - \frac{\sigma^2}{8\beta^3} - \frac{15\sigma^4}{128\beta^7} + \dots, \qquad \mathrm{var}(\hat{\hat{\beta}}) = \frac{\sigma^2}{4\beta^2} + \frac{7\sigma^4}{32\beta^6} + \dots$$

The following Table 2.3 can serve for a comparison with Tables 2.1. and 2.2 (here the value $\sigma = 100$ is omitted, since in this case $\sigma/\beta^2 = 1 \gg 0.1$).

σ	$\mathrm{var}(\hat{\hat{\beta}}$	$b(\hat{\hat{\beta}})$	$\sqrt{MSE(\hat{\hat{\beta}})}$
0.1	0.000025	0.000000	0.00500
1.0	0.002500	0.000125	0.05000
10.0	0.252187	0.012617	0.50234

Table 2.3

Variances and biases of the least squares estimators ($\beta_0 = 10$)

3 The function $y = \beta_1 \exp(-\beta_2 x)$; different approaches to an estimation of the parameters

Let values of the function $y = \beta_1 \exp(-\beta_2 x)$ be measured at the points $x = x_1, \dots x_n$. The results of measurements are a realization of the observation vector $\mathbf{Y} \sim N_n(\mathbf{f}(\beta), \mathbf{\Sigma})$, where $\{\mathbf{f}(\beta)\}_i = \beta_1 \exp(-\beta_2 x_i), i = 1, \dots, n$, and $\mathbf{\Sigma} = \sigma^2 \mathbf{I}$ is a known covariance matrix.

In general it can be written in the form

$$\mathbf{Y} \sim_n (\mathbf{f}(\beta), \mathbf{\Sigma}). \tag{1}$$

3.1 Reparametrization by a transformation of the data

Let $\eta_i = \ln Y_i,\ i = 1, \dots, n$[1]. Then

$$\begin{aligned}
\eta_i &= \ln \beta_1 - \beta_2 x_i + \delta_i, \quad i = 1, \dots, n, \\
\delta_i &= \ln\left(1 + \frac{\varepsilon_i}{\mu_i}\right) = \frac{\varepsilon_i}{\mu_i} - \frac{\varepsilon_i^2}{2\mu_i^2} + \frac{\varepsilon_i^3}{3\mu_i^3} - \dots, \quad i = 1, \dots, n, \\
\mu_i &= \beta_1 \exp(-\beta_2 x_i), \quad i = 1, \dots, n, \\
E(\delta_i) &= \frac{-\sigma^2}{2\mu_i^2} - \frac{3\sigma^4}{4\mu_i^4} + \dots, \quad i = 1, \dots, n, \\
\mathrm{var}(\delta_i) &= \frac{\sigma^2}{\mu_i^2} + \frac{5\sigma^4}{2\mu_i^4} + \dots, \quad i = 1, \dots, n.
\end{aligned}$$

OLSE (ordinary least squares estimator) of $\ln \beta_1$ and β_2 is

$$\begin{pmatrix} \widehat{\ln \beta_1} \\ \hat{\beta}_2 \end{pmatrix} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\eta = \begin{pmatrix} n, & -\sum_{i=1}^n x_i \\ -\sum_{i=1}^n x_i, & \sum_{i=1}^n x_i^2 \end{pmatrix}^{-1} \begin{pmatrix} \sum_{i=1}^n \eta_i \\ -\sum_{i=1}^n x_i \eta_i \end{pmatrix},$$

[1]To say it exact, the normal random variable $\{\mathbf{Y}\}_i = Y_i$ cannot be transformed in this way. However if $\sigma/E(Y_i)$ is sufficiently small, e.g. < 0.1, then the following results are realistic—in more detail cf. [9].

where

$$\mathbf{X} = \begin{pmatrix} 1, & -x_1 \\ \cdots & \cdots \\ 1, & -x_n \end{pmatrix}, \qquad \eta = \begin{pmatrix} \eta_1 \\ \vdots \\ \eta_n \end{pmatrix}$$

and

$$\begin{pmatrix} \widehat{\ln\beta_1} \\ \hat\beta_2 \end{pmatrix} \sim_2 \left[\begin{pmatrix} \ln\beta_1 \\ \beta_2 \end{pmatrix} + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'E(\boldsymbol{\delta}), (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\operatorname{var}(\boldsymbol{\delta})\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1} \right],$$
$$\boldsymbol{\delta} = (\delta_1, \ldots, \delta_n)'.$$

Lemma 3.1.1 *Let*

$$(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' = \begin{pmatrix} \mathbf{L}_1' \\ \mathbf{L}_2' \end{pmatrix}, \quad \mathbf{L}_1' = (L_{1,1}, \ldots, L_{1,n})', \quad \mathbf{L}_2' = (L_{2,1}, \ldots, L_{2,n})'.$$

Then

$$\begin{aligned}
E(\widehat{\ln\beta_1}) &= \ln\beta_1 + \sum_{i=1}^n L_{1,i}\left(-\frac{\sigma^2}{2\mu_i^2} - \frac{3\sigma^4}{4\mu_i^4} + \ldots\right), \\
E(\hat\beta_2) &= \beta_2 + \sum_{i=1}^n L_{2,i}\left(-\frac{\sigma^2}{2\mu_i^2} - \frac{3\sigma^4}{4\mu_i^4} + \ldots\right), \\
\operatorname{var}(\widehat{\ln\beta_1}) &= \sum_{i=1}^n L_{2,i}^2\left(\frac{\sigma^2}{\mu_i^2} + \frac{5\sigma^4}{2\mu_i^4} + \ldots\right), \\
\operatorname{var}(\hat\beta_2) &= \sum_{i=1}^n L_{2,i}^2\left(\frac{\sigma^2}{\mu_i^2} + \frac{5\sigma^4}{2\mu_i^4} + \ldots\right), \\
\operatorname{cov}(\widehat{\ln\beta_1}, \hat\beta_2) &= \sum_{i=1}^n L_{1,i}L_{2,i}\left(\frac{\sigma^2}{\mu_i^2} + \frac{5\sigma^4}{2\mu_i^4} + \ldots\right).
\end{aligned}$$

Proof Since it is obvious however tedious, it is omitted. □

Lemma 3.1.2 *If the estimators of β_1 and β_2 are considered in the form*

$$\tilde\beta_1 = \exp(\widehat{\ln\beta_1}), \qquad \tilde\beta_2 = \hat\beta_2,$$

then

$$\begin{aligned}
E(\tilde\beta_1) = \beta_1 + \beta_1 &\left[\sum_{i=1}^n \frac{\sigma^2}{\mu_i^2}\left(-\frac{1}{2}L_{1,i} + \frac{1}{2}L_{1,i}^2\right)\right. \\
&+ \sum_{i=1}^n \frac{\sigma^4}{\mu_i^4}\left(-\frac{3}{4}L_{1,i} + \frac{11}{8}L_{1,i}^2 - \frac{3}{4}L_{1,i}^3 + \frac{1}{8}L_{1,i}^4\right)
\end{aligned}$$

$$+\sum_{i\neq j}\frac{\sigma^4}{\mu_i^2\mu_j^2}\frac{1}{4}\left(L_{1,i}L_{1,j}-L_{1,i}^2L_{1,j}+L_{1,i}^2L_{1,j}^2\right)\Bigg]+\dots$$

$$E(\tilde{\beta}_2)=\beta_2+\sum_{i=1}^{n}L_{2,i}\left(-\frac{\sigma^2}{2\mu_i^2}-\frac{3\sigma^4}{4\mu_i^4}+\dots\right),$$

$$\operatorname{var}(\tilde{\beta}_1)=\beta_1^2\left[\sum_{i=1}^{n}\frac{\sigma^2}{\mu_i^2}L_{1,i}^2+\sum_{i=1}^{n}\frac{\sigma^4}{\mu_i^4}\left(\frac{5}{2}L_{1,i}^2-4L_{1,i}^3+\frac{3}{2}L_{1,i}^4\right)\right.$$

$$\left.+\sum_{i\neq j}\frac{\sigma^4}{\mu_i^2\mu_j^2}(3L_{1,i}^2L_{1,j}^2-L_{1,i}^2L_{1,j})\right]+\dots$$

$$\operatorname{cov}(\tilde{\beta}_1,\tilde{\beta}_2)=\beta_1\left\{\sum_{i=1}^{n}\frac{\sigma^2}{\mu_i^2}L_{1,i}L_{1,j}+\right.$$

$$+\sum_{i=1}^{n}\frac{\sigma^4}{\mu_i^4}\left(\frac{5}{2}L_{1,i}L_{2,i}-2L_{1,i}^2L_{2,i}+\frac{1}{2}L_{1,i}^3L_{2,i}\right)$$

$$+\sum_{i\neq j}\frac{\sigma^4}{\mu_i^2\mu_j^2}\left[-\frac{1}{2}L_{1,i}L_{1,j}(L_{2,i}+L_{2,j})\right.$$

$$\left.\left.+\frac{1}{2}L_{1,i}^2L_{1,j}L_{2,j}+\frac{1}{4}L_{1,i}^2L_{2,j}\right]\right\}+\dots,$$

$$\operatorname{var}(\tilde{\beta}_2)=\sum_{i=1}^{n}L_{2,i}^2\left(\frac{\sigma^2}{\mu_i^2}+\frac{5\sigma^4}{2\mu_i^4}+\dots\right).$$

Proof Since it is obvious however tedious, it is omitted. □

3.2 Reparametrization and weighting the transformed data

The approximate value of the $\operatorname{var}(\delta_i)$ is the first term σ^2/μ_i^2. With respect to the inequality

$$\operatorname{var}\begin{pmatrix}\widehat{\ln\beta_1}\\ \hat{\beta}_2\end{pmatrix}=(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\operatorname{var}(\boldsymbol{\delta})\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\geq(\mathbf{X}'\operatorname{var}^{-1}(\boldsymbol{\delta})\mathbf{X})^{-1},$$

(cf. [6, pp. 14–15]) it is reasonable to use *GLSE* (generalized least squares estimator), i.e.

$$\begin{pmatrix}\widehat{\widehat{\ln\beta_1}}\\ \hat{\hat{\beta}}_2\end{pmatrix}=(\mathbf{X}'\operatorname{var}^{-1}(\boldsymbol{\delta})\mathbf{X})^{-1}\mathbf{X}'\operatorname{var}^{-1}(\boldsymbol{\delta})\boldsymbol{\eta},$$

where $\operatorname{var}(\boldsymbol{\delta}) = \operatorname{Diag}(\sigma^2/\mu_1^2, \dots, \sigma^2/\mu_n^2)$ (the notation Diag() means the matrix with the entries on the diagonal equal to the values given in the brackets and with other entries equal to zero), instead of the OLSE (in more detail cf. [1]). Since in the variance

$$\operatorname{var}(\delta_i) = \sigma^2/\mu_i^2$$

the value μ_i is unknown, it is necessary to use approximate value $\mu_i^{(0)} = \beta_1^{(0)} \exp(-\beta_2^{(0)} x_i)$ of the quantity $\mu_i = \beta_1 \exp(-\beta_2 x_i)$, or to use some iterations in order to obtain the estimators $\widehat{\widehat{\ln \beta_1}}$ and $\hat{\hat{\beta}}_2$. Some investigation of an analogous situation cf. in [4].

Let

$$\mathbf{W} = \sigma^2 \begin{pmatrix} 1/(\mu^{(0)})_1^2, & 0, & \dots, 0, & 0 \\ 0, & 1/(\mu^{(0)})_2^2, & \dots, 0, & 0 \\ \dots & \dots & \dots & \dots \\ 0, & 0, & \dots, 0, & 1/(\mu^{(0)})_n^2 \end{pmatrix}$$

and

$$(\mathbf{X}'\mathbf{W}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{W}^{-1} = \begin{pmatrix} \mathbf{k}_1' \\ \mathbf{k}_2' \end{pmatrix}, \qquad \begin{pmatrix} \widehat{\widehat{\ln \beta_1}} \\ \hat{\hat{\beta}}_2 \end{pmatrix} = \begin{pmatrix} \mathbf{k}_1' \\ \mathbf{k}_2' \end{pmatrix} \boldsymbol{\eta}.$$

Then the bias and the variance of the estimator $\hat{\hat{\beta}}_1$ and $\hat{\hat{\beta}}_2$, respectively, can be expressed analogously as in Section 3.1. The vectors $\mathbf{L}_1, \mathbf{L}_2$ must be substituted by the vectors $\mathbf{k}_1, \mathbf{k}_2$. Thus we obtain

$$\begin{aligned} E(\widehat{\widehat{\ln \beta_1}}) &= \ln \beta_1 + \sum_{i=1}^n k_{1,i} \left(-\frac{\sigma^2}{2\mu_i^2} - \frac{3\sigma^4}{4\mu_i^4} + \dots \right), \\ E(\hat{\hat{\beta}}_2) &= \beta_2 + \sum_{i=1}^n k_{2,i} \left(-\frac{\sigma^2}{2\mu_i^2} - \frac{3\sigma^4}{4\mu_i^4} + \dots \right), \\ \operatorname{var}(\widehat{\widehat{\ln \beta_1}}) &= \sum_{i=1}^n k_{2,i}^2 \left(\frac{\sigma^2}{\mu_i^2} + \frac{5\sigma^4}{2\mu_i^4} + \dots \right), \\ \operatorname{var}(\hat{\hat{\beta}}_2) &= \sum_{i=1}^n k_{2,i}^2 \left(\frac{\sigma^2}{\mu_i^2} + \frac{5\sigma^4}{2\mu_i^4} + \dots \right), \\ \operatorname{cov}(\widehat{\widehat{\ln \beta_1}}, \hat{\hat{\beta}}_2) &= \sum_{i=1}^n k_{1,i} k_{2,i} \left(\frac{\sigma^2}{\mu_i^2} + \frac{5\sigma^4}{2\mu_i^4} + \dots \right). \end{aligned}$$

If the estimators of β_1 and β_2 are considered in the form

$$\tilde{\hat{\beta}}_1 = \exp(\widehat{\widehat{\ln \beta_1}}), \qquad \tilde{\hat{\beta}}_2 = \hat{\hat{\beta}}_2,$$

then

$$E(\tilde{\hat{\beta}}_1) = \beta_1 + \beta_1 \left[\sum_{i=1}^n \frac{\sigma^2}{\mu_i^2} \left(-\frac{1}{2} k_{1,i} + \frac{1}{2} k_{1,i}^2 \right) \right.$$

$$
\begin{aligned}
&+\sum_{i=1}^{n}\frac{\sigma^4}{\mu_i^4}\left(-\frac{3}{4}k_{1,i}+\frac{11}{8}k_{1,i}^2-\frac{3}{4}k_{1,i}^3+\frac{1}{8}k_{1,i}^4\right)\\
&\left.+\sum_{i\neq j}\frac{\sigma^4}{\mu_i^2\mu_j^2}\frac{1}{4}\left(k_{1,i}k_{1,j}-k_{1,i}^2k_{1,j}+k_{1,i}^2k_{1,j}^2\right)\right]+\dots
\end{aligned}
$$

$$
\begin{aligned}
E(\tilde{\tilde{\beta}}_2) &= \beta_2+\sum_{i=1}^{n}k_{2,i}\left(-\frac{\sigma^2}{2\mu_i^2}-\frac{3\sigma^4}{4\mu_i^4}+\dots\right),\\
\mathrm{var}(\tilde{\tilde{\beta}}_1) &= \beta_1^2\left[\sum_{i=1}^{n}\frac{\sigma^2}{\mu_i^2}k_{1,i}^2+\sum_{i=1}^{n}\frac{\sigma^4}{\mu_i^4}\left(\frac{5}{2}k_{1,i}^2-4k_{1,i}^3+\frac{3}{2}k_{1,i}^4\right)\right.\\
&\left.\quad+\sum_{i\neq j}\frac{\sigma^4}{\mu_i^2\mu_j^2}(3k_{1,i}^2k_{1,j}^2-k_{1,i}^2k_{1,j})\right]+\dots\\
\mathrm{cov}(\tilde{\tilde{\beta}}_1,\tilde{\tilde{\beta}}_2) &= \beta_1\left\{\sum_{i=1}^{n}\frac{\sigma^2}{\mu_i^2}k_{1,i}k_{1,j}\right.\\
&\quad+\sum_{i=1}^{n}\frac{\sigma^4}{\mu_i^4}\left(\frac{5}{2}k_{1,i}k_{2,i}-2k_{1,i}^2k_{2,i}+\frac{1}{2}k_{1,i}^3k_{2,i}\right)\\
&\quad+\sum_{i\neq j}\frac{\sigma^4}{\mu_i^2\mu_j^2}\left[-\frac{1}{2}k_{1,i}k_{1,j}(k_{2,i}+k_{2,j})\right.\\
&\quad\left.\left.+\frac{1}{2}k_{1,i}^2k_{1,j}k_{2,j}+\frac{1}{4}k_{1,i}^2k_{2,j}\right]\right\}+\dots,\\
\mathrm{var}(\tilde{\tilde{\beta}}_2) &= \sum_{i=1}^{n}k_{2,i}^2\left(\frac{\sigma^2}{\mu_i^2}+\frac{5\sigma^4}{2\mu_i^4}+\dots\right).
\end{aligned}
$$

3.3 Linearization of the model

If approximate values $\beta_1^{(0)}$ and $\beta_2^{(0)}$ of the parameters β_1 and β_2, respectively, are known, then in some cases the model

$$
\mathbf{Y}-\mathbf{f}_0\sim_n(\mathbf{F}\delta\boldsymbol{\beta},\boldsymbol{\Sigma}) \tag{2}
$$

can be used. Here

$$
\begin{aligned}
\mathbf{f}_0 &= \left(\beta_1^{(0)}\exp(-\beta_2^{(0)}x_1),\dots,\beta_1^{(0)}\exp(-\beta_2^{(0)}x_n)\right)',\\
\{\mathbf{F}\}_{j,1} &= \partial\left[\beta_1\exp(-\beta_2x_j)\right]/\partial\beta_1|_{\beta=\beta^{(0)}},\quad j=1,\dots,n,\\
\{\mathbf{F}\}_{j,2} &= \partial\left[\beta_1\exp(-\beta_2x_j)\right]/\partial\beta_2|_{\beta=\beta^{(0)}},\quad j=1,\dots,n.
\end{aligned}
$$

Some caution is necessary in using the linearization. It is useful to analyze a linearization regions for some statistical inference in the model. The theory is developed in [11] and [13].

The Bates and Watts measures of curvatures ([2]) of the model at the point β_0 are

$$K^{(int)}(\beta_0) = \sup\left\{ \frac{||\mathbf{M}_F^{\Sigma^{-1}}\boldsymbol{\kappa}_s||_{\Sigma^{-1}}}{||\mathbf{F}\mathbf{s}||^2_{\Sigma^{-1}}} : \mathbf{s} \in R^k \right\},$$

(the intrinsic curvature) and

$$K^{(par)}(\beta_0) = \sup\left\{ \frac{||\mathbf{P}_F^{\Sigma^{-1}}\boldsymbol{\kappa}_s||_{\Sigma^{-1}}}{||\mathbf{F}\mathbf{s}||^2_{\Sigma^{-1}}} : \mathbf{s} \in R^k \right\},$$

(the parametric curvature). Here $\mathbf{F} = \partial\mathbf{f}(\mathbf{u})/\partial\mathbf{u}'|_{u=\beta_0}$, $\boldsymbol{\kappa}_s = (\mathbf{s}\mathbf{H}_1'\mathbf{s},\ldots,\mathbf{s}\mathbf{H}_n'\mathbf{s})'$, $\mathbf{H}_i = \partial^2 f_i(\mathbf{u})/\partial\mathbf{u}\partial\mathbf{u}'|_{u=\beta_0}$, $i = 1,\ldots,n$, $\mathbf{M}_F^{\Sigma^{-1}} = \mathbf{I} - \mathbf{P}_F^{\Sigma^{-1}}$ and $\mathbf{P}_F^{\Sigma^{-1}} = \mathbf{F}(\mathbf{F}'\boldsymbol{\Sigma}^{-1}\mathbf{F})^{-1}\mathbf{F}'\boldsymbol{\Sigma}^{-1}$. The matrix $\mathbf{F}$ is assumed to be of the full rank in columns, i.e. $r(\mathbf{F}) = k < n$ and the matrix $\boldsymbol{\Sigma}$ is assumed to be positive definite.

The model (1) in the quadratized form is written as

$$\mathbf{Y} \sim N_n(\mathbf{f}_0 + \mathbf{F}\delta\beta + \frac{1}{2}\boldsymbol{\kappa}_{\delta\beta}, \boldsymbol{\Sigma}). \tag{3}$$

Here $\mathbf{f}_0 = \mathbf{f}(\beta_0)$.

The linear estimator in the model (2) is

$$\hat{\beta} = \beta_0 + \delta\hat{\beta}, \qquad \delta\hat{\beta} = (\mathbf{F}'\boldsymbol{\Sigma}^{-1}\mathbf{F})^{-1}\mathbf{F}'\boldsymbol{\Sigma}^{-1}(\mathbf{Y} - \mathbf{f}_0)$$

and with respect to (3)

$$\delta\hat{\beta} \sim N_k\left(\delta\beta + \frac{1}{2}(\mathbf{F}'\boldsymbol{\Sigma}^{-1}\mathbf{F})^{-1}\mathbf{F}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\kappa}_{\delta\beta}, (\mathbf{F}'\boldsymbol{\Sigma}^{-1}\mathbf{F})^{-1}\right).$$

In the following the influence of nonlinearity of the model is demonstrated in testing the consistency of measured data with the linearized model, in the bias of linear estimators of the parameter β, in the bias of the linear function of the parameter β and in a deterioration of the estimator of the variance of a linear function of the parameter β.

Let $\mathbf{h}$ be any k-dimensional vector and let

$$\begin{aligned}
b_h^*(\delta\beta) &= E[\mathbf{h}'\widehat{\delta\beta}(\mathbf{Y},\mathbf{0})|\delta\beta] - \mathbf{h}'\delta\beta = \mathbf{h}'\mathbf{b}(\delta\beta),\ \delta\beta \in \mathcal{O}(\beta_0),\\
d_h^*(\delta\beta) &= \mathrm{var}[\mathbf{h}'\widehat{\delta\beta}(\mathbf{Y},\delta\beta)|\boldsymbol{\Sigma}] - \mathrm{var}[\mathbf{h}'\widehat{\delta\beta}(\mathbf{Y},\mathbf{0})|\boldsymbol{\Sigma}],\ \delta\beta \in \mathcal{O}(\beta_0),\\
U_h^*(\delta\beta) &= \mathbf{h}'\widehat{\delta\beta}(\mathbf{Y},\delta\beta) - \mathbf{h}'\widehat{\delta\beta}(\mathbf{Y},\mathbf{0}),\ \delta\beta \in \mathcal{O}(\beta_0),\\
u_h^*(\delta\beta) &= \mathbf{h}'\widehat{\delta\beta}(\mathbf{y},\delta\beta) - \mathbf{h}'\widehat{\delta\beta}(\mathbf{y},\mathbf{0}),\ \delta\beta \in \mathcal{O}(\beta_0),
\end{aligned}$$

where $\mathbf{y}$ is a realization of the observation vector $\mathbf{Y}$,

$$\begin{aligned}
\mathbf{b}(\delta\beta) &= E[\widehat{\delta\beta}(\mathbf{Y},\mathbf{0})|\delta\beta] - \delta\beta,\\
\widehat{\delta\beta}(\mathbf{Y},\mathbf{0}) &= (\mathbf{F}'\boldsymbol{\Sigma}^{-1}\mathbf{F})^{-1}\mathbf{F}'\boldsymbol{\Sigma}^{-1}(\mathbf{Y} - \mathbf{f}_0),\\
E[\widehat{\delta\beta}(\mathbf{Y},\mathbf{0})|\delta\beta] &= (\mathbf{F}'\boldsymbol{\Sigma}^{-1}\mathbf{F})^{-1}\mathbf{F}'\boldsymbol{\Sigma}^{-1}(\mathbf{F}\delta\beta + \frac{1}{2}\boldsymbol{\kappa}_{\delta\beta}),
\end{aligned}$$

$$\begin{aligned}
\widehat{\delta\beta}(\mathbf{Y},\delta\beta) &= (\mathbf{F}_1'\mathbf{\Sigma}^{-1}\mathbf{F}_1)^{-1}\mathbf{F}_1'\mathbf{\Sigma}^{-1}(\mathbf{Y}-\mathbf{f}_1),\\
\mathbf{F}_1 &= \partial\mathbf{f}(\mathbf{u})/\partial\mathbf{u}'|_{u=\beta_0+\delta\beta},\\
\mathbf{f}_1 &= \mathbf{f}(\beta_0+\delta\beta),\\
\operatorname{var}[\mathbf{h}'\widehat{\delta\beta}(\mathbf{Y},\delta\beta)|\mathbf{\Sigma}] &= \mathbf{h}'(\mathbf{F}_1'\mathbf{\Sigma}^{-1}\mathbf{F}_1)^{-1}\mathbf{h}.
\end{aligned}$$

Let

$$\begin{aligned}
d_h(\delta\beta) &= \frac{\partial d_h^*(\delta\beta)}{\partial(\delta\beta')}\delta\beta, \delta\beta \in \mathcal{O}(\beta_0),\\
U_h(\delta\beta) &= \frac{\partial U_h^*(\delta\beta)}{\partial(\delta\beta')}\delta\beta, \delta\beta \in \mathcal{O}(\beta_0),\\
u_h(\delta\beta) &= \frac{\partial u_h^*(\delta\beta)}{\partial(\delta\beta')}\delta\beta, \delta\beta \in \mathcal{O}(\beta_0).
\end{aligned}$$

Let ε and c (> 0) be such constants and

$$\mathcal{O}_a(\beta_0),\ \mathcal{O}_b(\beta_0),\ \mathcal{O}_c(\beta_0),\ \mathcal{O}_d(\beta_0),\ \mathcal{O}_e(\beta_0) \text{ and } \mathcal{O}_f(\beta_0)$$

such neighbourhoods that

(a) $P\{\mathbf{v}'\mathbf{\Sigma}^{-1}\mathbf{v} \geq \chi^2(0;1-\alpha)\} \leq \alpha+\varepsilon, \quad \delta\beta \in \mathcal{O}_a(\beta_0),$

(b) $\mathbf{b}'(\delta\beta)\mathbf{C}\mathbf{b}(\delta\beta) \leq a^2\chi_k^2(0;1-\alpha), \quad \delta\beta \in \mathcal{O}_b(\beta_0),$

(c) $|b_h^*(\delta\beta)| \leq c\sqrt{\mathbf{h}'\mathbf{C}^{-1}\mathbf{h}}, \quad \delta\beta \in \mathcal{O}_c(\beta_0),$

(d) $|d_h(\delta\beta)| \leq c^2\mathbf{h}'\mathbf{C}^{-1}\mathbf{h}, \quad \delta\beta \in \mathcal{O}_d(\beta_0),$

(e) $\operatorname{var}[U_h(\delta\beta)|\mathbf{\Sigma}] \leq c^2\mathbf{h}'\mathbf{C}^{-1}\mathbf{h}, \quad \delta\beta \in \mathcal{O}_e(\beta_0),$

(f) $|u_h(\delta\beta)| \leq c\sqrt{\mathbf{h}'\mathbf{C}^{-1}\mathbf{h}}, \quad \delta\beta \in \mathcal{O}_f(\beta_0),$

respectively. Here

$$\begin{aligned}
\mathbf{C} &= \mathbf{F}'\mathbf{\Sigma}^{-1}\mathbf{F},\\
\mathbf{v} &= \mathbf{M}_F^{\Sigma^{-1}}(\mathbf{Y}-\mathbf{f}_0) \sim N_n\left(\frac{1}{2}\mathbf{M}_F^{\Sigma^{-1}}\boldsymbol{\kappa}_{\delta\beta}, \mathbf{M}_F^{\Sigma^{-1}}\mathbf{\Sigma}\left(\mathbf{M}_F^{\Sigma^{-1}}\right)'\right),
\end{aligned}$$

$a^2\chi_k^2(0;1-\alpha) = c^2$ and $\chi_k^2(0;1-\alpha)$ is the $(1-\alpha)$-quantile of the central chi-square distribution with k degrees of freedom.

Definition 3.3.1 The model (3) is c-linearizable with respect to a function $h(.)$ in the set $\mathcal{O}_c(\beta_0)$, $\mathcal{O}_d(\beta_0)$, $\mathcal{O}_e(\beta_0)$ and $\mathcal{O}_f(\beta_0)$, respectively, if the inequalities $(c), (d), (e)$ and (f), respectively, are satisfied.

The model (3) is ε-linearizable with respect to a compatibility of data with model, if the inequality (a) is satisfied.

The model (3) is c-linearizable with respect to the bias of the parameter estimator $\delta\hat{\beta}$, if the inequality (b) is satisfied.

The sets $\mathcal{O}_a(\beta_0),\dots,\mathcal{O}_f(\beta_0)$ are called the linearization regions.

The following statements are proved in [5] and [7],[11], respectively.

Statement 3.3.2 *If*

$$\delta\boldsymbol{\beta} \in \mathcal{O}_a(\boldsymbol{\beta}_0) = \left\{\delta\boldsymbol{\beta} : (\delta\boldsymbol{\beta})'\mathbf{C}\delta\boldsymbol{\beta} \le \frac{2\sqrt{\delta_{\max}}}{K^{(int)}(\boldsymbol{\beta}_0)}\right\},$$

where $\delta_{\max}$ *is given by the equation*

$$P\{\chi^2_{n-k}(\delta_{\max}) \ge \chi^2_{n-k}(0; 1-\alpha)\} = \alpha + \varepsilon,$$

the model is ε*-linearizable at the point* $\boldsymbol{\beta}_0$*. Here* $\chi^2_{n-k}(\delta)$ *is noncentral chi-square random variable with* $n-k$ *degrees of freedom and with the parameter of noncetrality equal to* δ*.*

Statement 3.3.3 *Let* $c = a\sqrt{\chi^2_k(0; 1-\alpha)}$*. If*

$$\delta\boldsymbol{\beta} \in \mathcal{O}_b(\boldsymbol{\beta}_0) = \left\{\delta\boldsymbol{\beta} : \delta\boldsymbol{\beta}'\mathbf{F}'\boldsymbol{\Sigma}^{-1}\mathbf{F}\delta\boldsymbol{\beta} \le \frac{2c}{K^{(par)}(\boldsymbol{\beta}_0)}\right\},$$

then

$$\forall\{\mathbf{h} \in R^k\}|b^*_h(\delta\boldsymbol{\beta})| \le c\sqrt{\mathbf{h}'\mathbf{C}^{-1}\mathbf{h}}.$$

Statement 3.3.4 *Let* $\mathbf{L}_h = \mathbf{h}'\mathbf{C}^{-1}\mathbf{F}'\boldsymbol{\Sigma}^{-1} = (\{\mathbf{L}_h\}_1, \ldots, \{\mathbf{L}_h\}_n)$ *and*

$$\mathbf{B}_0 = \sum_{i=1}^{n}\{\mathbf{L}_h\}_i\frac{1}{2}\mathbf{H}_i = \sum_{j=1}^{s}\lambda_j\mathbf{g}_j\mathbf{g}'_j$$

be the spectral decomposition and let

$$\mathbf{B} = \sum_{j=1}^{s}|\lambda_j|\mathbf{g}_j\mathbf{g}'_j.$$

Then

$$\delta\boldsymbol{\beta} \in \mathcal{O}_c(\boldsymbol{\beta}_0) = \left\{\delta\boldsymbol{\beta} : \delta\boldsymbol{\beta}'\mathbf{B}\delta\boldsymbol{\beta} \le c\sqrt{\mathbf{h}'\mathbf{C}^{-1}\mathbf{h}}\right\} \Rightarrow |b^*_h(\delta\boldsymbol{\beta})| \le c\sqrt{\mathbf{h}'\mathbf{C}^{-1}\mathbf{h}}.$$

Let

$$\mathbf{H}^*_i = \begin{pmatrix} \mathbf{e}'_i\mathbf{H}_1 \\ \vdots \\ \mathbf{e}'_i\mathbf{H}_n \end{pmatrix}, \quad i = 1, \ldots, k,$$

where $\mathbf{e}_i \in R^k$, $\mathbf{e}_i = (0, \ldots, 0, 1_i, 0, \ldots, 0)'$ and

$$\mathbf{K}^{(h)}_1 = \begin{pmatrix} \mathbf{h}'\mathbf{C}^{-1}(\mathbf{H}^*_1)'\boldsymbol{\Sigma}^{-1} \\ \vdots \\ \mathbf{h}'\mathbf{C}^{-1}(\mathbf{H}^*_k)'\boldsymbol{\Sigma}^{-1} \end{pmatrix} \quad \text{and} \quad \mathbf{K}^{(h)}_2 = \begin{pmatrix} \mathbf{L}'_h\mathbf{H}^*_1 \\ \vdots \\ \mathbf{L}'_h\mathbf{H}^*_k \end{pmatrix}.$$

Remark 3.3.5 Let

$$\mathbf{h}'\hat{\beta}(\mathbf{y},\mathbf{u}) = \mathbf{h}'\mathbf{u} + \left[(\partial\mathbf{f}(\mathbf{u})/\partial\mathbf{u}')\,\mathbf{\Sigma}^{-1}(\partial\mathbf{f}(\mathbf{u})/\partial\mathbf{u}')\right]^{-1}(\partial\mathbf{f}(\mathbf{u})/\partial\mathbf{u}')'\,\mathbf{\Sigma}^{-1}(\mathbf{y}-\mathbf{f}(\mathbf{u})).$$

Then

$$\partial\mathbf{h}'\hat{\beta}(\mathbf{y},\mathbf{u})/\partial\mathbf{u}|_{u=\beta_0} = \mathbf{h} + \mathbf{K}_1^{(h)}\mathbf{v} - \mathbf{K}_2^{(h)}\delta\hat{\beta}(\mathbf{y},\mathbf{0}),$$

where $\mathbf{v} = \mathbf{y} - \mathbf{f}(\beta_0) - \mathbf{F}\delta\hat{\beta}(\mathbf{y},\mathbf{0})$. (In Lemma 2.11 [5] a little different formula is given.) Thus

$$\mathbf{h}'(\beta_0+\delta\beta) + \mathbf{h}'\delta\hat{\beta}(\mathbf{y},\delta\beta) \approx \mathbf{h}'\beta_0 + \mathbf{h}'\delta\hat{\beta}(\mathbf{y},\mathbf{0}) + (\delta\beta)'\left[\mathbf{K}_1^{(h)}\mathbf{v} - \mathbf{K}_2^{(h)}\delta\hat{\beta}(\mathbf{y},\mathbf{0})\right]$$

and

$$u_h(\delta\beta) \approx (\delta\beta)'\left[\mathbf{K}_1^{(h)}\mathbf{v} - \mathbf{K}_2^{(h)}\delta\hat{\beta}(\mathbf{y},\mathbf{0})\right].$$

Statement 3.3.6 *If in the model (3) $\delta\beta \in \mathcal{O}_d(\beta_0)$, where*

$$\mathcal{O}_d(\beta_0) = \left\{\delta\beta : \delta\beta'\delta\beta \le c^2\frac{\mathbf{h}'\mathbf{C}^{-1}\mathbf{h}}{\sqrt{\mathbf{h}'\mathbf{C}^{-1}(\mathbf{K}_1^{(h)}\mathbf{F}+\mathbf{K}_2^{(h)})'(\mathbf{K}_1^{(h)}\mathbf{F}+\mathbf{K}_2^{(h)})\mathbf{C}^{-1}\mathbf{h}}}\right\},$$

then $|d_h(\delta\beta)| \le c^2\mathbf{h}'\mathbf{C}^{-1}\mathbf{h}$.

Statement 3.3.7 *If the power of components of the vector $\delta\beta$ greater than two is neglected, then in the model (3)*

$$U_h(\delta\beta) \sim N_1\left\{\mathbf{h}'\mathbf{b}(\delta\beta), \delta\beta'\mathbf{W}^{(h)}\delta\beta\right\},$$

where

$$\mathbf{W}^{(h)} = \left[\mathbf{K}_1^{(h)}(\mathbf{\Sigma}-\mathbf{F}\mathbf{C}^{-1}\mathbf{F}')(\mathbf{K}_1^{(h)})' + \mathbf{K}_2^{(h)}\mathbf{C}^{-1}(\mathbf{K}_2^{(h)})'\right].$$

Statement 3.3.8 *Let the notation $\mathbf{W}^{(h)}$ from Statement 3.3.7 be used.*
If

$$\delta\beta \in \mathcal{O}_e(\beta_0) = \{\delta\beta : \delta\beta'\mathbf{W}^{(h)}\delta\beta \le c^2\mathbf{h}'\mathbf{C}^{-1}\mathbf{h}\},$$

then $\sqrt{\mathrm{var}[U_h(\delta\beta)|\mathbf{\Sigma}]} \le c\sqrt{\mathbf{h}'\mathbf{C}^{-1}\mathbf{h}}$.

Remark 3.3.9 If the criterion from Statement 3.3.8 is too restrictive for some realization $\widehat{\delta\beta}(\mathbf{y},\mathbf{0})$ of the random variable $\widehat{\delta\beta}(\mathbf{Y},\mathbf{0})$, i.e., if a realization $\mathbf{v}_{real}$ of the rezidual $\mathbf{v}$ and the vector $\widehat{\delta\beta}(\mathbf{y},\mathbf{0})$ makes the value

$$u_h(\delta\beta) = \delta\beta'[\mathbf{K}_1^{(h)}\mathbf{v}_{real} - \mathbf{K}_2^{(h)}\widehat{\delta\beta}(\mathbf{y},\mathbf{0})]$$

small, then it is reasonable to calculate the value

$$\frac{c\sqrt{\mathbf{h}'\mathbf{C}^{-1}\mathbf{h}}}{\sqrt{[\mathbf{K}_1^{(h)}\mathbf{v}_{real} - \mathbf{K}_2^{(h)}\widehat{\delta\beta}(\mathbf{y},\mathbf{0})]'[\mathbf{K}_1^{(h)}\mathbf{v}_{real} - \mathbf{K}_2^{(h)}\widehat{\delta\beta}(\mathbf{y},\mathbf{0})]}} = T.$$

If $\sqrt{\delta\beta'\delta\beta} \leq T$, then $|u_h(\delta\beta)| \leq c\sqrt{\mathbf{h}'\mathbf{C}^{-1}\mathbf{h}}$.

If the region

$$\mathcal{O}_f(\beta_0) = \{\delta\beta : \delta\beta'\delta\beta \leq T^2\}$$

covers the region

$$\mathcal{O}_e(\beta_0) = \{\delta\beta : \delta\beta'\mathbf{W}^{(h)}\delta\beta \leq c^2\mathbf{h}'\mathbf{C}^{-1}\mathbf{h}\},$$

then in the actual case the value T is to be preferred. Also a comparison of the regions $\mathcal{O}_c(\beta_0)$ and $\mathcal{O}_f(\beta_0)$ can be interesting and important in practice.

From the above given relationships it is clear that the first important information on the nonlinearity are the measures $K^{(int)}(\beta_0)$ and $K^{(par)}(\beta_0)$. In the following table some values of those measures are given (in more deatil cf. [17]).

$\binom{\beta_1}{-\beta_2}$	$\binom{10.0}{-0.1}$	$\binom{9.0}{-0.2}$	$\binom{8.0}{-0.3}$	$\binom{7.0}{-0.4}$	$\binom{6.0}{-0.5}$	$\binom{5.0}{-0.6}$
$K^{(int)}(\beta)$	0.020582	0.055846	0.138862	0.250937	0.326234	0.359022
$K^{(par)}(\beta)$	0.057598	0.119664	0.207130	0.303481	0.474110	0.750977

Table 3.3.1

$x = 1, 2, 3, 7, 8, 9, \quad \sigma = 0.5$

3.4 Quadratic corrections of the linear estimator

The following notations will be used in this section.

$$\begin{aligned}
\mathbf{Y} &\sim N_n(\mathbf{f}(\beta), \sigma^2\mathbf{I}), \quad \sigma \in (0, \infty), \\
\mathbf{f}_0 &= \mathbf{f}(\beta_0), \qquad \mathbf{F} = \partial\mathbf{f}(\mathbf{u})/\partial\mathbf{u}'|_{u=\beta_0}, \\
\mathbf{H}_i &= \partial^2 f_i(\mathbf{u})/\partial\mathbf{u}\partial\mathbf{u}'|_{u=\beta_0}, \quad i = 1, \dots, n, \\
\mathbf{e}_i &\in R^n, \qquad \{\mathbf{e}_i\}_j = \delta_{i,j} \text{ (Kronecker delta)}, \\
\mathbf{B}_{0,i} &= \sum_{j=1}^{n} \{\mathbf{e}_i'\mathbf{C}^{-1}\mathbf{F}'\mathbf{\Sigma}^{-1}\}_j \frac{1}{2}\mathbf{H}_j, \quad i = 1, \dots, k, \\
\mathbf{C} &= \mathbf{F}'\mathbf{\Sigma}^{-1}\mathbf{F} (= \sigma^{-2}\mathbf{F}'\mathbf{F}), \qquad \mathbf{b} = \begin{pmatrix} \delta\beta'\mathbf{B}_{0,1}\delta\beta \\ \vdots \\ \delta\beta'\mathbf{B}_{0,k}\delta\beta \end{pmatrix}, \\
\delta\beta &= \sqrt{\frac{\chi_k^2(0; 1-\alpha)}{\mathbf{u}'\mathbf{C}\mathbf{u}}}\mathbf{u}, \quad \mathbf{u} \text{ an arbitrary non zero vector,}
\end{aligned}$$

($\delta\beta$ is located at the boundary of the $(1-\alpha)$-confidence ellipse).

The corrected estimator of β_i is

$$\begin{aligned}
\tilde{\tilde{\beta}}_i &= \beta_{0,i} + \delta\tilde{\tilde{\beta}}_i, \\
\delta\tilde{\tilde{\beta}}_i &= \delta\hat{\beta}_i - \delta\hat{\beta}'\mathbf{B}_{0,i}\delta\hat{\beta} + \mathrm{Tr}(\mathbf{B}_{0,i}\mathbf{C}^{-1}), \quad i = 1, \dots, k, \\
\delta\hat{\beta} &= \mathbf{C}^{-1}\mathbf{F}'\mathbf{\Sigma}^{-1}(\mathbf{Y} - \mathbf{f}_0).
\end{aligned}$$

The following statements given in this section are proved in [12]

Mean square error of the linear estimator

$$MSE(\hat{\beta}_i) = \{\mathbf{C}^{-1}\}_{i,i} + (\delta\boldsymbol{\beta}'\mathbf{B}_{0,i}\delta\boldsymbol{\beta})^2.$$

Mean square error of the quadratic estimator

$$\begin{aligned} MSE(\tilde{\tilde{\beta}}_i) &= \{\mathbf{C}^{-1}\}_{i,i} + 2\,\mathrm{Tr}(\mathbf{B}_{0,i}\mathbf{C}^{-1}\mathbf{B}_{0,i}\mathbf{C}^{-1}) - 4\mathbf{e}_i'\mathbf{C}^{-1}\mathbf{B}_{0,i}\delta\boldsymbol{\beta} \\ &+ 4\delta\boldsymbol{\beta}'\mathbf{B}_{0,i}\mathbf{C}^{-1}\mathbf{B}_{0,i}\delta\boldsymbol{\beta} - 4\mathbf{e}_i'\mathbf{C}^{-1}\mathbf{B}_{0,i}\mathbf{b} + 8\mathbf{b}'\mathbf{B}_{0,i}\mathbf{C}^{-1}\mathbf{B}_{0,i}\delta\boldsymbol{\beta} \\ &+ 4\mathbf{b}'\mathbf{B}_{0,i}\mathbf{C}^{-1}\mathbf{B}_{0,i}\mathbf{b} + 4(\mathbf{b}'\mathbf{B}_{0,i}\delta\boldsymbol{\beta})^2 + 4\mathbf{b}'\mathbf{B}_{0,i}\delta\boldsymbol{\beta}\mathbf{b}'\mathbf{B}_{0,i}\mathbf{b} + (\mathbf{b}'\mathbf{B}_{0,i}\mathbf{b})^2. \end{aligned}$$

In the following (cf. 5. Numerical example) the MSE of the both estimators will be calculated for

$$\delta\boldsymbol{\beta} = \sqrt{\frac{\chi_k^2(0;1-\alpha)}{\mathbf{u}'\mathbf{C}\mathbf{u}}}\,\mathbf{u}, \quad \mathbf{u} \text{ an arbitrary nonzero vector.}$$

Upper bound for MSE of the linear estimator at the boundary of the $(1-\alpha)$-confidence ellipsoid

$$MSE(\hat{\beta}_i) \le \{\mathbf{C}^{-1}\}_{i,i} + \mathrm{Tr}[(\mathbf{B}_{0,i}\mathbf{C}^{-1})^2][\chi_k^2(0;1-\alpha)]^2, \quad i = 1,\dots,k.$$

Upper bound for MSE of the quadratic estimator at the boundary of the $(1-\alpha)$-confidence ellipsoid

$$\begin{aligned} MSE(\tilde{\tilde{\beta}}_i) \le{}& \{\mathbf{C}^{-1}\}_{i,i} + 2\,\mathrm{Tr}(\mathbf{B}_{0,i}\mathbf{C}^{-1}\mathbf{B}_{0,i}\mathbf{C}^{-1}) \\ &+ 4\sqrt{\mathrm{Tr}[(\mathbf{B}_{0,i}\mathbf{C}^{-1})^2]}\sqrt{\{\mathbf{C}^{-1}\}_{i,i}}\sqrt{\chi_k^2(0,1-\alpha)} \\ &+ 4\sqrt{\mathrm{Tr}[(\mathbf{B}_{0,i}\mathbf{C}^{-1})^4]}\chi_k^2(0,1-\alpha) \\ &+ 2\sqrt{\mathrm{Tr}[(\mathbf{B}_{0,i}\mathbf{C}^{-1})^2]}\sqrt{\{\mathbf{C}^{-1}\}_{i,i}}K^{(par)}(\boldsymbol{\beta}_0)\chi_k^2(0;1-\alpha) \\ &+ 4\sqrt{\mathrm{Tr}[(\mathbf{B}_{0,i}\mathbf{C}^{-1})^4]}K^{(par)}(\boldsymbol{\beta}_0)[\chi_k^2(0;1-\alpha)]^{3/2} \\ &+ \sqrt{\mathrm{Tr}[(\mathbf{B}_{0,i}\mathbf{C}^{-1})^4]}[K^{(par)}(\boldsymbol{\beta}_0)]^2[\chi_k^2(0;1-\alpha)]^2 \\ &+ \mathrm{Tr}[(\mathbf{B}_{0,i}\mathbf{C}^{-1})^2][K^{(par)}(\boldsymbol{\beta}_0)]^2[\chi_k^2(0;1-\alpha)]^3 \\ &+ \frac{1}{2}\mathrm{Tr}[(\mathbf{B}_{0,i}\mathbf{C}^{-1})^2][K^{(par)}(\boldsymbol{\beta}_0)]^3[\chi_k^2(0;1-\alpha)]^{7/2} \\ &+ \frac{1}{16}\mathrm{Tr}[(\mathbf{B}_{0,i}\mathbf{C}^{-1})^2][K^{(par)}(\boldsymbol{\beta}_0)]^4[\chi_k^2(0;1-\alpha)]^4. \end{aligned}$$

Let

$$\delta\boldsymbol{\beta} = \sqrt{\frac{\chi_k^2(0;1-\alpha)}{\{\mathbf{C}^{-1}\mathbf{B}_{0,i}\mathbf{C}\mathbf{B}_{0,i}\mathbf{C}^{-1}\}_{i,i}}}\{\mathbf{B}_{0,i}\mathbf{C}^{-1}\}_{\cdot,i}$$

(this vector is directed as the gradient $\partial MSE(\tilde{\tilde{\beta}}_i)/\partial\delta\boldsymbol{\beta}$ of the function $MSE(\tilde{\tilde{\beta}}_i)$ and it is located at the boundary of the $(1-\alpha)$-confidence ellipsoid).

Then

$$(\delta\boldsymbol{\beta}'\mathbf{B}_{0,i}\delta\boldsymbol{\beta})^2 = \left(\frac{\chi_k^2(0;1-\alpha)}{\{\mathbf{C}^{-1}\mathbf{B}_{0,i}\mathbf{C}\mathbf{B}_{0,i}\mathbf{C}^{-1}\}_{i,i}}\right)^2 \left(\{\mathbf{C}^{-1}\mathbf{B}_{0,i}^3\mathbf{C}^{-1}\}_{i,i}\right)^2$$

(square of the bias of the linear estimator) and

$$4|-\mathbf{e}_i'\mathbf{C}^{-1}\mathbf{B}_{0,i}\delta\boldsymbol{\beta}| = 4\{\mathbf{C}^{-1}\mathbf{B}_{0,i}^2\mathbf{C}^{-1}\}_{i,i}\sqrt{\frac{\chi_k^2(0;1-\alpha)}{\{\mathbf{C}^{-1}\mathbf{B}_{0,i}\mathbf{C}\mathbf{B}_{0,i}\mathbf{C}^{-1}\}_{i,i}}}$$

(the value of the linear term in $MSE(\tilde{\tilde{\beta}}_i)$).

If

$$4\{\mathbf{C}^{-1}\mathbf{B}_{0,i}^2\mathbf{C}^{-1}\}_{i,i}\sqrt{\frac{\chi_k^2(0;1-\alpha)}{\{\mathbf{C}^{-1}\mathbf{B}_{0,i}\mathbf{C}\mathbf{B}_{0,i}\mathbf{C}^{-1}\}_{i,i}}} >$$
$$> \left(\frac{\chi_k^2(0;1-\alpha)}{\{\mathbf{C}^{-1}\mathbf{B}_{0,i}\mathbf{C}\mathbf{B}_{0,i}\mathbf{C}^{-1}\}_{i,i}}\right)^2 \left(\{\mathbf{C}^{-1}\mathbf{B}_{0,i}^3\mathbf{C}^{-1}\}_{i,i}\right)^2,$$

then the linear estimator is to be preferred.

3.5 Nonlinear least squares estimator

In this section the model

$$\mathbf{Y} \sim N_n(\mathbf{f}(\boldsymbol{\beta}), \boldsymbol{\Sigma}), \quad \boldsymbol{\beta} \in \underline{\beta},$$

where $\underline{\beta}$ means the parametric space (in our case of the exponential regression $\underline{\beta} = R^2$).

The following notations will be used in this section (in more detail cf. [15]).

$\boldsymbol{\beta}^*$ is the actual value of the parameter $\boldsymbol{\beta}$,

$\mathrm{int}(\underline{\beta})$ means the topological interior (in Euclidean sense) of the parametric space $\underline{\beta}$.

The nonlinear least squares estimators of the parameters β_1 and β_2 are given by the solution of the equation

$$[\partial\mathbf{f}'(\boldsymbol{\beta})/\partial\boldsymbol{\beta}]\,\boldsymbol{\Sigma}^{-1}[\mathbf{y}-\mathbf{f}(\boldsymbol{\beta})] = \mathbf{0},$$

i.e.

$$\begin{pmatrix} e^{-\beta_2 x_1}, & \dots, & e^{-\beta_2 x_n} \\ -\beta_1 x_1 e^{-\beta_2 x_1}, & \dots, & -\beta_1 x_n e^{-\beta_2 x_n} \end{pmatrix} \boldsymbol{\Sigma}^{-1} \begin{pmatrix} y_1 - \beta_1 e^{-\beta_2 x_1} \\ \vdots \\ y_1 - \beta_n e^{-\beta_2 x_n} \end{pmatrix} = \mathbf{0}.$$

In the following let $\boldsymbol{\Sigma} = \sigma^2\mathbf{W}$.

With respect to [15, Proposition 7.1.1], the probability density function of the nonlinear least squares estimators $\hat{\beta}_1$ and $\hat{\beta}_2$ of the parameters β_1 and β_2 is

$$p_{\beta^*}(\hat{\beta}) = C^* q(\hat{\beta}|\beta^*) E^*\{\det[\mathbf{I} + \mathbf{D}(\mathbf{b}, \hat{\beta})\mathbf{Q}^{-1}(\hat{\beta}, \beta^*)]\},$$

$$q(\hat{\beta}|\beta^*) = \frac{\det[\mathbf{Q}(\hat{\beta}, \beta^*)]}{(2\pi)^{\frac{k}{2}}\sigma^k\{\det[\mathbf{C}_0(\hat{\beta})]\}^{\frac{1}{2}}} \exp\Big\{-\frac{1}{2\sigma^2}\|\mathbf{P}_F^{W^{-1}}(\hat{\beta})[\mathbf{f}(\hat{\beta}) - \mathbf{f}(\beta^*)]\|^2_{W^{-1}}\Big\},$$

$$\{\mathbf{Q}(\hat{\beta}, \beta^*)\}_{i,j} = \{\mathbf{C}_0(\hat{\beta})\}_{i,j} + [\mathbf{f}(\hat{\beta}) - \mathbf{f}(\beta^*)]'\mathbf{W}^{-1}[\mathbf{I} - \mathbf{P}_F^{W^{-1}}(\hat{\beta})]\partial_i\partial_j \mathbf{f}(\hat{\beta}),$$

$$\{\mathbf{D}(\mathbf{b}, \hat{\beta})\}_{i,j} = -\sum_{s=1}^{n-k} b_s \boldsymbol{\omega}_s(\hat{\beta})\mathbf{W}^{-1}\partial_i\partial_j \mathbf{f}(\hat{\beta}), \quad i = 1, \dots, k.$$

Here $\mathbf{C}_0 = \mathbf{F}'\mathbf{W}^{-1}\mathbf{F}$, E^* means the mean value with respect to a truncated normal density with $(n-k)$ variables

$$f^*(\mathbf{b}) = \begin{cases} (2\pi\sigma^2)^{(n-k)/2}\exp\big\{-\frac{1}{2\sigma^2}\mathbf{b}'\mathbf{b}\big\}\,, & \text{if } \mathbf{b}'\mathbf{b} \le (r_0/2)^2, \\ 0, & \text{if } \mathbf{b}'\mathbf{b} \ge (r_0/2)^2, \end{cases}$$

the symbol C^* means the norming constant and it is either equal to 1 (if the density $p_{\beta^*}(\cdot)$ should be normed to the probability of the restricted parameter space $\underline{\beta}(r_0)$), i.e.

$$\int_{\underline{\beta}(r_0)} p_{\beta^*}(\beta)\mathrm{d}\beta = \Pr\{\mathcal{T}^*(r_0)\},$$

where

$$\begin{aligned} \underline{\beta}(r_0) = \{\beta \in \operatorname{int}(\underline{\beta}) : \|\mathbf{P}_F^{W^{-1}}(\beta)[\mathbf{f}(\beta) - \mathbf{f}(\beta^*)]\|_{W^{-1}} \le r_0/2 \\ \text{and } \|\mathbf{f}(\beta) - \mathbf{f}(\beta^*)\|_{W^{-1}} \le r_0\}. \end{aligned}$$

or it is equal to

$$C^* = [\Pr\{\mathcal{T}^*(r_0)\}]^{-1},$$

if the density $p_{\beta^*}(\cdot)$ should be normed to one, i.e.

$$\int_{\underline{\beta}(r_0)} p_{\beta^*}(\beta)\mathrm{d}\beta = 1.$$

Further n-dimensional vectors

$$\boldsymbol{\omega}_s(\hat{\beta}), \quad s = 1, \dots, n-k$$

are $\mathbf{W}^{-1}$-orthonormal and $\mathbf{W}^{-1}$-orthogonal to the columns of the matrix $\mathbf{F}(\hat{\beta})$ and the number r_0 defines so called restricted sample space $\mathcal{T}^*(r_0)$

$$\begin{aligned} \mathcal{T}^*(r_0) &= \bigcup_{\beta \in \Theta(r_0)} \mathcal{Z}(\beta, r_0/2), \\ \mathcal{Z}(\beta, r_0/2) &= \{\mathbf{y} : \langle \mathbf{y} - \mathbf{f}(\beta), \mathbf{F}(\beta)\rangle_{W^{-1}} = 0 \\ &\qquad \text{and } \|\mathbf{y} - \boldsymbol{\psi}_{\beta^*}(\beta)\|_{W^{-1}} \le r_0/2\}, \\ \boldsymbol{\psi}_{\beta^*}(\beta) &= \mathbf{P}_F^{W^{-1}}(\beta)[\mathbf{f}(\beta) - \mathbf{f}(\beta^*)] + \mathbf{f}(\beta^*). \end{aligned}$$

If $K_0^{(int)}(\beta_0)$ is defined as follows

$$K_0^{(int)}(\beta_0) = \sup\left\{\frac{||\mathbf{M}_F^{W^{-1}}\boldsymbol{\kappa}_s||_{W^{-1}}}{||\mathbf{F}\mathbf{s}||^2_{W^{-1}}} : \mathbf{s} \in R^k\right\},$$

then the so called assumption of bounded curvature must be satisfied:

$$\forall\{\beta \in \underline{\beta}(r_0)\}\ r_0 < \frac{1}{K_0^{(int)}(\beta)}.$$

Further the so called assumption of non-overlapping must be also satisfied: There is no $\mathbf{y} \in \mathcal{T}(r_0)$, such that the normal equation

$$[\partial\mathbf{f}'(\beta)/\partial\beta]W^{-1}[\mathbf{y} - \mathbf{f}(\beta)] = \mathbf{0}$$

has two solutions $\beta^{(1)}, \beta^{(2)} \in \mathrm{int}(\underline{\beta})$, and that

$$||\mathbf{y} - \mathbf{f}(\beta^{(i)})||_{W^{-1}} < r_0, i = 1, 2.$$

Under these assumptions the following inclusions hold

$$\mathcal{G}(r_0/2) \subseteq \mathcal{T}^*(r_0) \subseteq \mathcal{G}(r_0),$$

where $\mathcal{G}(r) = \{\mathbf{y} \in R^n : ||\mathbf{y} - \mathbf{f}(\beta^*)||_{W^{-1}} < r\}$.

The number r_0 must be chosen so that the conditions are satisfied. If r_0 is too large, then the assumptions are violated; if it is too small, then the accuracy of the approximate density is unsatisfactory.

If both assumptions hold, then for every $\mathbf{y} \in \mathcal{T}^*(r_0)$ the estimate $\hat{\beta}(\mathbf{y})$ is the unique solution of the normal equation which belongs to the set $\underline{\beta}(r_0)$. (Further detail cf. [15].)

4 Design of experiment

In the following two versions of the optimum design of experiment (in more detail cf. [14]) will be considered, i.e. the D-optimum design for the model after logarithmic transformation and the D-optimum design for the model after linearization.

Let the set of the experimental points be $\mathcal{E} = \{x_1, , \dots, x_r\}$. In the case of the model after the logarithmic transformation

$$\boldsymbol{\eta} \sim \left(\begin{pmatrix} 1, & -x_1 \\ \multicolumn{2}{c}{\dots\dots} \\ 1, & -x_r \end{pmatrix}\begin{pmatrix} \ln\beta_1 \\ \beta_2 \end{pmatrix}, \begin{pmatrix} \sigma^2/\mu_1^2, & \dots, & 0 \\ \multicolumn{3}{c}{\dots\dots\dots\dots} \\ 0, & \dots, & \sigma^2/\mu_r^2 \end{pmatrix}\right),$$

the information matrix for the design $\boldsymbol{\zeta}$ is

$$\mathbf{M}(\zeta) = \sum_{i=1}^{r} \mu_i^2 \begin{pmatrix} 1 \\ -x_i \end{pmatrix} (1, -x_i)\zeta_i.$$

Here $\zeta_i \geq 0, i = 1, \ldots, r$, and $\sum_{i=1}^{r} \zeta_i = 1$.

Let

$$\begin{aligned} d_i &= \mu_i^2(1, -x_i)\mathbf{M}^{-1}(\zeta)\begin{pmatrix} 1 \\ -x_i \end{pmatrix} \\ &= \mu_i^2(1, -x_i)\left(\sum_{j=1}^{r} \mu_j^2 \zeta_j \begin{pmatrix} 1, & -x_j \\ -x_j, & x_j^2 \end{pmatrix}\right)^{-1}\begin{pmatrix} 1 \\ -x_i \end{pmatrix}. \end{aligned}$$

The D-optimum design $\boldsymbol{\zeta}_D^*$ is defined as follows

$$\det[\mathbf{M}(\boldsymbol{\zeta}^*)] = \max\{\det[\mathbf{M}(\boldsymbol{\zeta})] : \boldsymbol{\zeta} \in \Delta\},$$

where Δ is the set of all designs. In [14] the following statement is proved.

Statement 4.1 Let $\zeta^{(0)} = 1/l, e_i \in \mathrm{Sp}(\zeta^{(0)}) = \{e_i : \zeta_i^{(0)} > 0\}$. If

$$d_{i_{s+1}^*} = \max\left\{\mu_i^2(1, -x_i)\mathbf{M}^{-1}(\zeta^{(s)})\begin{pmatrix} 1 \\ -x_i \end{pmatrix} : i = 1, \ldots, r\right\}$$

and

$$\zeta_i^{(s+1)} = \begin{cases} \frac{l+s}{l+s+1}\zeta_i^{(s)}, & i \neq i_{s+1}^*, \\ \frac{l+s}{l+s+1}\zeta_i^{(s)} + \frac{1}{l+s+1}, & i = i_{s+1}^*. \end{cases}$$

for $s = 0, 1, \ldots$, then $\lim_{s\to\infty} \mathbf{M}(\boldsymbol{\zeta}^{(s)}) \to \mathbf{M}(\boldsymbol{\zeta}_D^*)$.

In this way we obtain the D-optimum design for the model after logarithmization when simultaneously the weighting is taken into account.

The following formula is useful in the process of iteration.

$$\begin{aligned} &\mathbf{M}^{-1}(\boldsymbol{\zeta}^{(s+1)}) = \\ &= \frac{l+s+1}{l+s}\left[\mathbf{M}^{-1}(\boldsymbol{\zeta}^{(s)}) - \frac{\mu_{i_{s+1}^*}^2 \mathbf{M}^{-1}(\boldsymbol{\zeta}^{(s)}\begin{pmatrix} 1, & -x_{i_{s+1}^*} \\ -x_{i_{s+1}^*}, & x_{i_{s+1}^*}^2 \end{pmatrix}\mathbf{M}^{-1}(\boldsymbol{\zeta}^{(s)})}{l + s + \mu_{i_{s+1}^*}^2(1, -x_{i_{s+1}^*})\mathbf{M}^{-1}(\boldsymbol{\zeta}^{(s)})\begin{pmatrix} 1 \\ -x_{i_{s+1}^*} \end{pmatrix}}\right]. \end{aligned}$$

The iteration procedure can be stopped when

$$d_{i_t^*} - k \leq \varepsilon \quad \text{(chosen by a user)}.$$

In our case of the exponential regression $k = 2$.

Quite analogously the D-optimum design for the linearized model can be obtained.

Let

$$\mathbf{f}_i' = (\partial f_i(\boldsymbol{\beta})/\partial\beta_1, \partial f_i(\boldsymbol{\beta})/\partial\beta_2)\,|_{\boldsymbol{\beta}=\boldsymbol{\beta}_0} = (\mathrm{e}^{-\beta_2^{(0)}x_i}, -\beta_1^{(0)}x_i\mathrm{e}^{-\beta_2^{(0)}x_i}),\ i = 1, \ldots, r,$$

where $\boldsymbol{\beta}_0$ is an approximate value of the vector $\boldsymbol{\beta}^*$. If the vector $(1, -x_i)$ in the given procedure is substituted by the vector $\mathbf{f}_i'$ and instead of μ_i the value 1 is used, we obtain the D-optimum design for the linearized model at the point $\boldsymbol{\beta}_0$.

Remark 4.2 It is to be said that in the model

$$\eta \sim \left(\begin{pmatrix} 1, & -x_1 \\ \cdots & \cdots \\ 1, & -x_r \end{pmatrix} \begin{pmatrix} \ln \beta_1 \\ \beta_2 \end{pmatrix}, \sigma^2 \mathbf{I} \right),$$

the D-optimum design is

$$\zeta_i = \begin{cases} \frac{1}{2}, & x_{\min} = \min\{x_j : j = 1, \ldots, r\} \text{ or } x_{\max} = \max\{x_j : j = 1, \ldots, r\}, \\ 0, & \text{otherwise.} \end{cases}$$

5 Numerical example

Let $\mathcal{E} = \{1, 2, 3, 4, 5, 6\}$, $\boldsymbol{\beta}_0 = \binom{8}{-0.3}$ and $\sigma = 1$. Then the D-optimum design for the model

$$\ln y_i = \ln \beta_1 - \beta_2 x_i + \delta_i, \quad \delta_i \sim_1 (0, \sigma^2/\mu_i^2),$$

is

x	1	2	3	4	5	6
ζ_{100}	0.009434	0.009434	0.471698	0.009434	0.009434	0.490566

Table 5.1

In this case $d_{100,\max} = 2.02644$ and

$$\mathbf{M}^{-1}(\zeta_{100}) = \begin{pmatrix} 0.021568, & 0.003742 \\ 0.003742, & 0.000672 \end{pmatrix},$$

$\det[\mathbf{M}^{-1}(\zeta_{100})] = 491 \times 10^{-9}$ in a comparison with the starting design (Table 5.2)

x	1	2	3	4	5	6
ζ_0	1/6	1/6	1/6	1/6	1/6	1/6

Table 5.2

where

$$\mathbf{M}^{-1}(\zeta_0) = \begin{pmatrix} 0.018783, & 0.003553 \\ 0.003553, & 0.000717 \end{pmatrix},$$

$\det[\mathbf{M}^{-1}(\zeta_0)] = 843.5 \times 10^{-9}$.

In the case of the linearized model

$$y_i = \beta_1^{(0)} \exp(-\beta_2^{(0)} x_i) + (\exp(\beta_2^{(0)} x_i, -\beta_1^{(0)} x_i \exp(-\beta_2^{(0)} x_i))\delta\boldsymbol{\beta} + \varepsilon_i,$$

for the same set $\mathcal{E}$, the D-optimum design is identical with Table 5.1. The information matrices for the starting design (Table 5.2) and the D-optimum one (after 100 iterations) are

$$\mathbf{M}^{-1}(\zeta_0) = \begin{pmatrix} 1.202100, & 0.028423 \\ 0.028423, & 0.000717 \end{pmatrix}, \quad \det[\mathbf{M}^{-1}(\zeta_0)] = 54.039 \times 10^{-6},$$

$$\mathbf{M}^{-1}(\zeta_{100}) = \begin{pmatrix} 1.380400, & 0.029936 \\ 0.029936, & 0.000672 \end{pmatrix}, \quad \det[\mathbf{M}^{-1}(\zeta_{100})] = 31.465 \times 10^{-6}.$$

The D-optimum design cannot be used in practice for a small number of measurements. Thus let us try some approximations of the optimum design for $n = 6$.

If

x	3	6
ζ	1/2	1/2

Table 5.3

then for $\sigma = 1$

$$\mathbf{C}^{-1} = \begin{pmatrix} 20.266000, & -0.676350 \\ -0.676350, & 0.024680 \end{pmatrix}, \quad \det(\mathbf{C}^{-1}) = 0.042716$$

and $K^{(int)}(\beta_0) = 0$, $K^{(par)}(\beta_0) = 1.77497$. (The value of $K^{(par)}(\beta_0)$ is extremaly large.)

If

x	1	3	4	6
ζ	1/6	1/3	1/6	1/3

Table 5.4

then for $\sigma = 1$

$$\mathbf{C}^{-1} = \begin{pmatrix} 3.250400, & -0.130580 \\ -0.130580, & 0.007496 \end{pmatrix}, \quad \det(\mathbf{C}^{-1}) = 0.007314$$

and $K^{(int)}(\beta_0) = 0.142282$, $K^{(par)}(\beta_0) = 0.212473$.

In the further the last design is used.

Since

x	1	3	4	6
$\beta_1^{(0)} \exp(-\beta_2^{(0)} x)$	5.926	3.253	2.410	1.322

Table 5.5

the value of σ is chosen as 0.3. Even in this case

$$\sigma / \left[\beta_1^{(0)} \exp(-\beta_2^{(0)} 6)\right] = 0.227,$$

what does not mean a precise measurement usual in metrology. (It is to be said that the smaller value of σ is chosen the linearazation is better.) Then

$$\mathbf{C}^{-1} = \begin{pmatrix} 0.292540, & -0.011752 \\ -0.011752, & 0.000675 \end{pmatrix}, \tag{4}$$

$K^{(int)}(\beta_0) = 0.042684$, $K^{(par)}(\beta_0) = 0.063742$.

In Figs. 5.1–5.8 the linearization regions are given for $c = 0.5$, $\alpha = 0.05$ and $\varepsilon = 0.04$ (cf. Definition 3.3.1).

The 0.95 confidence ellipse is the same in all figures however its graphical demonstration is different. The aim is to show boundaries of linearization regions and that is why scales on the axis $db1$ and $db2$, respectively, are different in different figures.

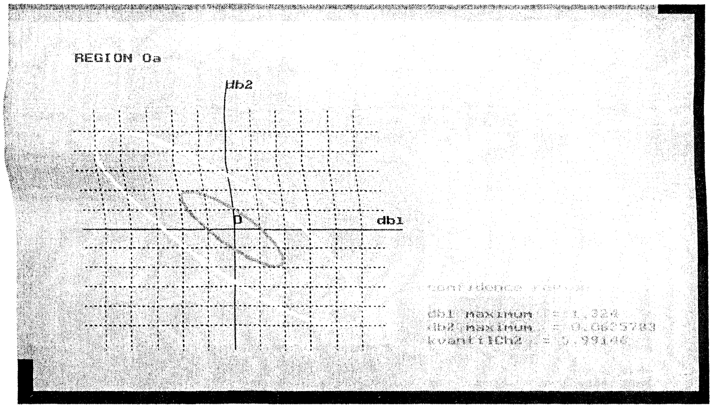

Fig. 5.1: Region $\mathcal{O}_a$

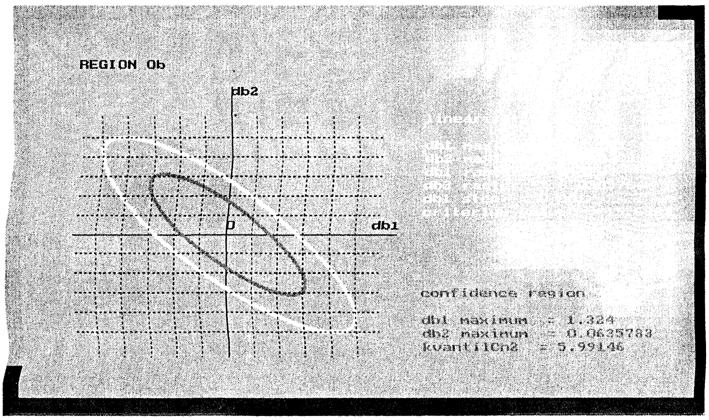

Fig. 5.2: Region $\mathcal{O}_b$

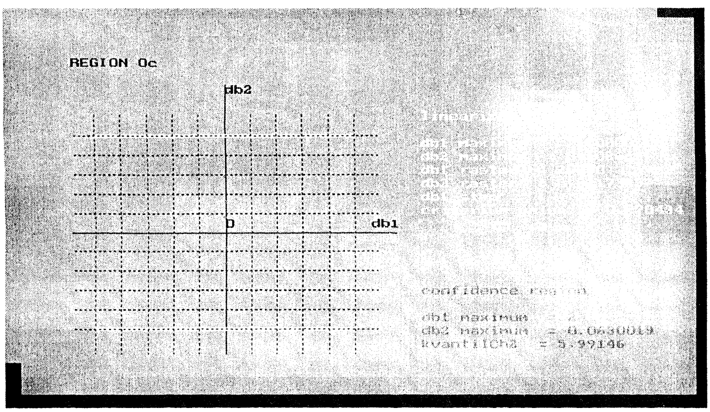

Fig. 5.3: Region $\mathcal{O}_c$ for $\mathbf{h} = (1, 0)'$

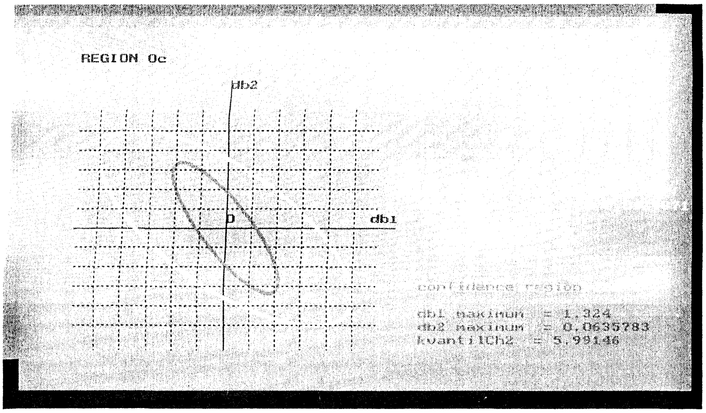

Fig. 5.4: Region $\mathcal{O}_c$ for $\mathbf{h} = (0, 1)'$

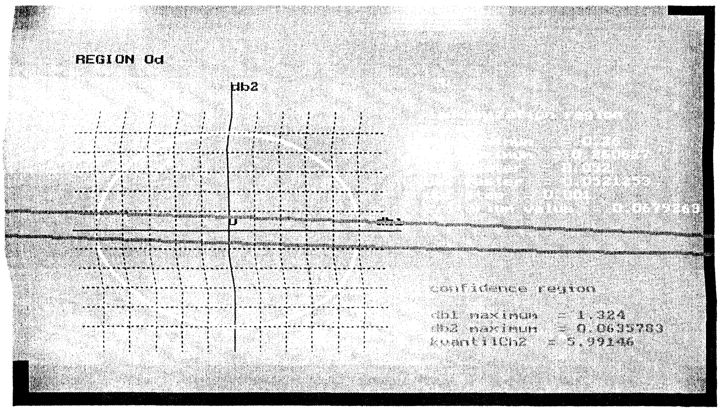

Fig. 5.5: Region $\mathcal{O}_d$ for $\mathbf{h} = (1, 0)'$

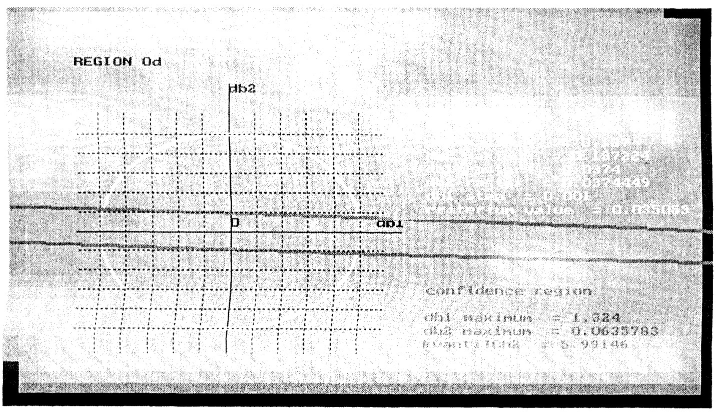

Fig. 5.6: Region $\mathcal{O}_d$ for $\mathbf{h} = (0, 1)'$

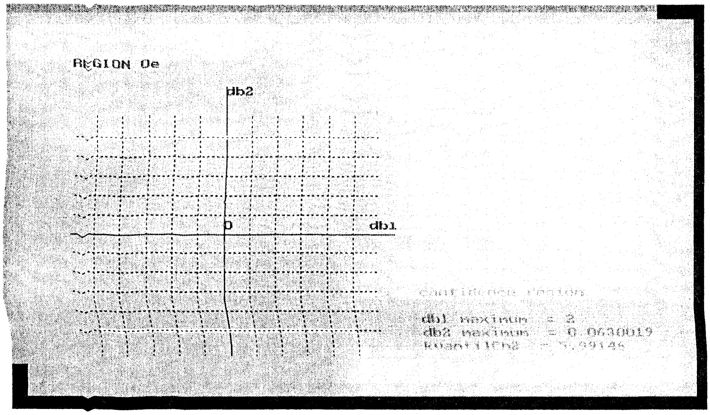

Fig. 5.7: Region $\mathcal{O}_e$ for $\mathbf{h} = (1,0)'$

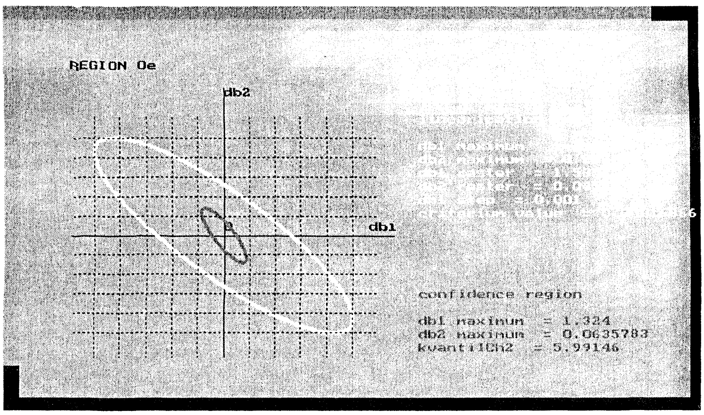

Fig. 5.8: Region $\mathcal{O}_e$ for $\mathbf{h} = (0,1)'$

A comparison of the behaviour of the linear and quadratic estimators is given at the boundary of the 0.95-confidence ellipse in the following Tables 5.6 and 5.7.

	$\delta\boldsymbol{\beta} = \binom{0.725}{0}$	$\delta\boldsymbol{\beta} = \binom{0}{0.035}$	$\delta\boldsymbol{\beta} = \binom{0.144}{0.029}$
$MSE(\hat{\beta}_1)$	0.292538	0.293311	0.292903
$MSE(\tilde{\tilde{\beta}}_1)$	0.293016	0.253871	0.260243
$MSE(\hat{\beta}_2)$	0.000675	0.000687	0.000683
$MSE(\tilde{\tilde{\beta}}_2)$	0.000559	0.000511	0.000515

Table 5.6: $\sigma = 0.3$

The upper bounds of the *MSE* (*UBMSE*) at the boundary of the 0.95--confidence ellipse of the estimators are in this case

$$UBMSE(\hat{\beta}_1) = 0.301110,$$
$$UBMSE(\tilde{\tilde{\beta}}_1) = 0.388105,$$
$$UBMSE(\hat{\beta}_2) = 0.000699,$$
$$UBMSE(\tilde{\tilde{\beta}}_2) = 0.000920.$$

	$\delta\beta = \binom{2.417}{0}$	$\delta\beta = \binom{0}{0.116}$	$\delta\beta = \binom{-0.814}{-0.081}$
$MSE(\hat{\beta}_1)$	3.250430	3.345800	3.777260
$MSE(\tilde{\tilde{\beta}}_1)$	3.309380	1.891950	1.621550
$MSE(\hat{\beta}_2)$	0.007496	0.008994	0.010144
$MSE(\tilde{\tilde{\beta}}_2)$	0.003818	0.002893	0.003585

Table 5.7: $\sigma = 1$

The upper bounds of the MSE at the boundary of the 0.95-confidence ellipse of the estimators are in this case

$$UBMSE(\hat{\beta}_1) = 4.308630,$$
$$UBMSE(\tilde{\tilde{\beta}}_1) = 8.615380,$$
$$UBMSE(\hat{\beta}_2) = 0.010517,$$
$$UBMSE(\tilde{\tilde{\beta}}_2) = 0.021315.$$

If the ln-transformation of data (i.e. Lemmas 3.1.1 and 3.1.2) is used, then we obtain

OLSE			GLSE		
$E(\ln(\hat{\beta}_1)$	=	2.08974	$E(\ln(\hat{\beta}_1)$	=	2.08296
$\mathrm{var}(\ln(\hat{\beta}_1)$	=	0.001631	$\mathrm{var}(\ln(\hat{\beta}_1)$	=	0.000715
$E(\hat{\beta}_1)$	=	8.13214	$E(\hat{\beta}_1)$	=	8.04683
$\mathrm{var}(\hat{\beta}_1)$	=	0.849031	$\mathrm{var}(\hat{\beta}_1)$	=	0.302976
$E(\hat{\beta}_2)$	=	−0.294127	$E(\hat{\beta}_2)$	=	−0.29643
$\mathrm{var}(\hat{\beta}_2)$	=	0.001631	$\mathrm{var}(\hat{\beta}_2)$	=	0.000715

If these values are compared with (4), we can see that the best results can be obtained in the case of linearization (if it is possible) or by quadratization, then by GLSE and the worst results are in the case of OLSE, which is the most frequently used procedure in practice.

References

[1] Anděl, J., Zvára, K.: *On a nonlinear regression.* Ekon. mat. obzor **21** (1985), 444–453 (in Czech).

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

[3] Jenčová, A.: *A comparison of linearization and quadratization domains.* Applications of Mathematics **42** (1997), 279–291.

[4] Klička, R., Kubáček, L.: *Statistical properties of linearization of the Arrhenius equation via the logarithmic transformation.* Chemometrics and intelligent laboratory systems **39** (1997), 69–75.

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

[6] Kubáček, L., Kubáčková, L., Volaufová, J.: Statistical Models with Linear Structures. *Veda, Bratislava,* 1995.

[7] Kubáček, L.: *Models wit a low nonlinearity.* Tatra Mountains Math. Publ. **7** (1996), 149–155.

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

[9] Kubáček, L.: *Nonlinear error propagation law.* Applications of Mathematics **41** (1996), 329–345.

[10] Kubáček, L.: *Corrections of estimators in linearized models.* Acta Univ. Palacki. Olomuc., Fac. rer. nat., Math. **37** (1998), 69–80.

[11] Kubáček, L., Kubáčková, L.: *Regression Models with a weak Nonlinearity.* Technical Reports, University of Stuttgart, 1998, 1–64.

[12] Kubáček, L.: *Linear versus quadratic estimators in linearized models.* Submitted to Applications of Mathematics.

[13] Kubáček, L., Kubáčková, L.: Statistika a metrologie. *Vyd. Univ. Palackého, Olomouc,* 2000.

[14] Pázman, A.: Foundations of Optimum Experimental Design. *Reidel, Dordrecht — Veda, Bratislava,* 1985.

[15] Pázman, A.: Nonlinear Statistical Models. *Kluwer Academic Publishers, Dordrecht–Boston–London and Ister Science Press, Bratislava,* 1993.

[16] Rao, C. R., Mitra, S. K.: Generalized Inverse of the Matrix and Its Applications. *J. Wiley, New York,* 1971.

[17] Tesaříková, E., Kubáček, L.: *How to deal with regression models with a weak nonlinearity.* To appear in Discussiones Mathematicae: Algebra and Stochastic Methods