On a corrected contrast estimator in the implicit nonlinear functional relation model

Alexander Kukush
University of Kiev

Silvelyn Zwanzig
University of Hamburg

October 1997
Preprint 97-11
Institute for Mathematical Stochastics
University of Hamburg
Bundesstr.55 D-20146 Hamburg, Germany

Abstract

For the implicit nonlinear functional relation model a new corrected contrast estimation procedure is proposed, which is based on the deconvolution idea and eliminates the nuisance parameters in the usual minimum contrast function. The sufficient conditions for consistency are given. Several examples are considered including L_1 and L_2 methods.

Contents

1 Introduction

The aim of this paper is to present a new corrected minimum contrast function in the implicit functional relation model.

The functional relation model is characterized by an increasing number of nuisance parameters and a fixed number of parameters of interest. It is an old fact that in models with an increasing number of nuisance parameters usual estimation procedures may fail, compare for instant the classical paper of Neyman, Scott (1948), [?]. The main problem is the elimination of the nuisance parameters in the estimating function, in such a way that the estimating procedure delivers consistent estimators also in cases when the nuisance parameters are not consistently estimable. In linear models the least squares approach fulfills that and gives consistent and efficient estimators for the parameter of interest. The situation is different in nonlinear models. When the conditions are such that the nuisance parameters are consistently estimable, then the least squares estimator is consistent, compare Zwanzig (1990), [?], otherwise under no restriction on the nuisance parameters the least squares approach delivers inconsistent estimates, this is shown for arbitrary explicit nonlinear regression models in Kukush, Zwanzig (1996),[?] and for more complicated error distributions in Fazekas, Kukush and Zwanzig (1997), [?].

Stefanski (1985), [?], discussed the effect of measurement errors on M-estimation and proposed estimators with a smaller bias. In Stefanski (1989), [?], and Stefanski (1989), [?] he continued this approach of finding unbiased score functions for estimation in the nonlinear measurement error model. In [?] a general theorem basing on Fourier transforms is given for the construction of unbiased score functions independent on the nuisance parameters for normally distributed errors. Nakamura (1990), [?], independently developed the same approach to measurement error problems and applied it to generalized nonlinear measurement error models. In [?] Buzas and Stefanski (1996) extended the corrected-score method studied by Nakamura (1990), [?], and Stefanski (1989), [?], to a large class of generalized linear measurement error models. They assumed normal errors and an expansion of the known link functions together with a lemma of Stein (1981), [?]. Hanfelt and Kuing-Yee Liang (1997), [?], proposed a conditional quasi-likelihood function for the generalized linear model.

In Kukush, Zwanzig (1996) [?] the analogous idea was used to correct the naive least squares estimation criterion and an alternative estimator was proposed. Fazekas and Kukush (1996), [?] showed for that alternative estimator a large deviation result. Unfortunately this estimator has the disadvantage to correct the least squares idea asymmetrically and delivers an asymptotically inefficient estimator, but which is also consistent under conditions, where the least squares estimator is not.

Here we consider the implicit model to emphasize the symmetrical character of the observations, which makes no difference between independent and dependent variables. We start with an usual contrast estimator depending on the unknown nuisance parameter. We propose an elimination procedure for the nuisance parameters in the minimum contrast criterion such that we obtain asymptotically the same estimation procedure as the starting one. When we start from the naive estimator we obtain the alternative estimators as above in Kukush, Zwanzig (1996), [?], otherwise when we begin with the least squares approach respecting to both types of parameters we get a symmetric correcting procedure, which in the linear case coincides with the orthogonal regression. That fact gives us the hope to obtain also asymptotic efficient estimation procedures under normally distributed errors.

Like in the papers of Stefanski , [?], [?], [?] the correcting approach is based on the deconvolution problem. The key point as discussed by Stefanski is the exact solution of the deconvolution integral equations. He also proposed approximate solutions by linearization of the regression functions. Note, that all authors in [?], [?], [?], [?], [?] considered the asymmetric correcting procedure only.

In this paper we consider a general approximate solutions of the deconvolution equations and introduce approximate corrected minimum contrast estimators (ac-MCE). Under conditions where the approximation error in the deconvolution equations is decreasing with increasing sample size we show the consistency of the approximate corrected minimum contrast estimator (ac-MCE).

2 The implicit model and the main idea of correcting

2.1 The model assumptions

Consider the implicit model

$$G(\eta_i, \beta) = 0, \quad i = 1, .., n \tag{1}$$

where $G : \mathrm{R}^q \times \mathrm{R}^p \longrightarrow \mathrm{R}$ is a known smooth function and

$$\eta_i \in \mathcal{D} \subset \mathrm{R}^q, \eta = (\eta_1, ..., \eta_n) \in \mathcal{D}^n \quad \text{nuisance parameter} \tag{2}$$

and

$$\beta \in \mathcal{B} \subset \mathrm{R}^p \text{ parameter of interest.} \tag{3}$$

Sometimes we shall use also the denotation

$$\theta = \begin{pmatrix} \eta \\ \beta \end{pmatrix} \in \Theta \subset \mathcal{D}^n \times \mathcal{B}, \quad \Theta \text{ compact,} \tag{4}$$

for the common parameter consisting of the nuisance parameter and the parameter of interest. Then the parameter set is given by

$$\Theta = \left\{ \theta = \begin{pmatrix} \eta \\ \beta \end{pmatrix} \in \mathcal{D}^n \times \mathcal{B} : G(\eta_i, \beta) = 0, i = 1, .., n \right\}. \tag{5}$$

Note that the number of parameters included in θ is $\dim \theta = nq + p$ and increases with the sample size. We observe the $q-$dimensional vector Z_i

$$Z_i = \eta_i + \varepsilon_i, \tag{6}$$

with ε_i are *i.i.d.* with expected value zero and covariance matrix Γ, denoted by

$$\varepsilon_i \sim (0, \Gamma) \quad i.i.d.. \tag{7}$$

The explicit functional relation model is included for $q \geq 2$,

$$\eta_i = \begin{pmatrix} \eta_{1i} \\ \eta_{2i} \end{pmatrix}, \ \eta_{1i} \in \mathrm{R}^1, \ \eta_{2i} \in \mathrm{R}^{q-1} \tag{8}$$

and

$$G(\eta_i, \beta) = \eta_{1i} - g(\eta_{2i}, \beta). \tag{9}$$

Then we have for

$$Z_i = \begin{pmatrix} y_i \\ x_i \end{pmatrix}$$

the usual explicit model

$$\begin{pmatrix} y_i \\ x_i \end{pmatrix} = \begin{pmatrix} g(\xi_i, \beta) \\ \xi_i \end{pmatrix} + \begin{pmatrix} \varepsilon_{1i} \\ \varepsilon_{2i} \end{pmatrix}, \tag{10}$$

with the unknown design points

$$\xi = (\xi_1, ..., \xi_n) \in \mathcal{F}^{(n)} \subset \left(\mathrm{R}^{q-1}\right)^n.$$

Note, in (**??**) we may have the impression that the y_i act like dependent variables and the x_i play the role of independent variables of a regression model. But that is not, the explicit model as a particular case of the implicit model, where both variables have the symmetric influence.

2.2 The minimum contrast approach

We call a nonrandom positive real function $C_n : \theta \in \Theta \to \mathrm{R}_+$ a *contrast for θ at*

$$\theta^0 = \begin{pmatrix} \eta^0 \\ \beta^0 \end{pmatrix} \tag{11}$$

iff it is lower semicontinuous and for some semimetric $d(.,.)$ on Θ and for strictly increasing $\rho : \mathrm{R}_+ \to \mathrm{R}_+$, with $\rho(0) = 0$,

$$\Delta C_n(\theta) = C_n(\theta) - C_n\left(\theta^0\right)$$

fulfills the following separation condition:

Con For all $n \geq n_0$ for all $\theta \in \Theta$

$$\Delta C_n(\theta) \geq \rho\left(d\left(\theta, \theta^0\right)\right). \tag{12}$$

Note that under (**??**)

$$\theta^0 = \arg\min_{\theta \in \Theta} C_n(\theta). \tag{13}$$

The contrast may depend on the sample size n. For instance we can choose for some weights $w_i = w_i(n) \geq 0$ and some $r \geq 1$

$$C_n(\theta) = \sum_{i=1}^{n} w_i \left|G(\eta_i, \beta)\right|^r. \tag{14}$$

For the explicit model (**??**) examples are the empirical L_r-contrast

$$C_n(\theta) = \sum_{i=1}^{n} w_i \left|g\left(\xi_i^0, \beta\right) - g\left(\xi_i^0, \beta^0\right)\right|^r, \tag{15}$$

or the asymptotic L_r-contrast

$$C_n(\theta) = C(\beta) = \int \left|g(x, \beta) - g\left(x, \beta^0\right)\right|^r dG(x). \tag{16}$$

The contrast in (**??**) depends on the unknown design points $\xi \in \mathcal{F}^{(n)}$, the other one in (**??**) depends on an asymptotic design G. Under an unique parameterization of the regression function each distance measure d for functions $g(.,\beta)$ seems to be a useful contrast at β^0, namely $C_n(\beta) = d\left(g(.,\beta), g(.,\beta^0)\right)$. For the explicit model (**??**) let us introduce also

$$C_n(\theta) = \frac{1}{n}\sum_{i=1}^{n}\left|g(\xi_i,\beta) - g\left(\xi_i^0,\beta^0\right)\right|^2 + \left\|\xi_i^0 - \xi_i\right\|^2. \tag{17}$$

We call a measurable function

$$\widetilde{C}_n(.,.) : \mathrm{R}^{qn} \times \Theta :\rightarrow \mathrm{R}_+ \tag{18}$$

a *contrast function* and require that it is lower semicontinuous with respect to θ. In the general statistical experiment, which includes random processes and random fields as well, Liese and Vajda (1995), [**?**], introduced a more general concept and called the function corresponding to (**??**) a contrast principle. Note, in order to simplify the denotation we will suppress the dependence on the sample and let the tilde hints to this: $\widetilde{C}_n(Z,\theta) =: \widetilde{C}_n(\theta)$. Then we define the corresponding estimator as follows.

A measurable solution $\widetilde{\theta} : \mathrm{R}^n \times \mathrm{R}^n \rightarrow \Theta$ is called *a minimum contrast estimator* (MCE) iff

$$\widetilde{\theta} \in \arg\min_{\theta\in\Theta} \widetilde{C}_n(\theta). \tag{19}$$

Under the above introduced model assumptions the existence of minimum contrast estimators are given by Lemma 2 of Liese and Vajda (1995), [**?**]. Consider the differences of the contrast and of the contrast function,

$$\Delta C_n(\theta) = C_n(\theta) - C_n\left(\theta^0\right) \quad \text{and} \quad \Delta\widetilde{C}_n(\theta) = \widetilde{C}_n(\theta) - \widetilde{C}_n\left(\theta^0\right). \tag{20}$$

Then the connection between the consistency of the minimum contrast estimator and the uniform consistent approximation of the contrast $C_n(\theta)$ by the contrast function $\widetilde{C}_n(\theta)$ is given by the following inequality, here quoted from Zwanzig (1997), [**?**]:

$$\forall\epsilon > 0 \quad P\left(d\left(\widetilde{\theta},\theta^0\right) > \epsilon\right) \leq P\left(\sup_{\theta\in\Theta(\epsilon)} \frac{\left|\Delta C_n(\theta) - \Delta\widetilde{C}_n(\theta)\right|}{\rho\left(d\left(\theta,\theta^0\right)\right)} \geq 1\right), \tag{21}$$

where $\Theta(\epsilon) = \Theta \cap \{\theta : d(\theta,\theta^0) > \epsilon\}$. It is a version of an "argmin" result, like the argmax theorem for i.i.d. experiments in van der Vaart and Wellner (1996), [**?**].

Note, the expression $\rho\left(d\left(\theta,\theta^0\right)\right)$ comes from the separation condition (**??**) of the contrast. The rate of convergence of the minimum contrast estimator in (**??**) depends mainly on the separation property of the contrast (**??**) and the semimetric $d\left(.,.\right)$ chosen in (**??**).

The main problem of that approach for the functional relation model is that the dimension of the common parameter $\dim\theta = nq + p$ grows in the same order as the sample size. To obtain a nice consistency result from (**??**) requires an uniform approximation of the contrast by the contrast function on such semiparametric set. This approach is hopeless under no additional assumptions on the nuisance parameter set. Otherwise the estimating procedure in (**??**) includes also the consistent estimation of the nuisance parameters.

The key point now is to find a corrected contrast function to the same contrast, which does not depend on the unknown nuisance parameters.

2.3 The correction of the estimating criterion

Usually like in the examples (**??**), (**??**), (**??**) the contrast has an average structure:

$$C_n\left(\theta\right) = \frac{1}{n}\sum_{i=1}^{n} c_i\left(\theta\right). \tag{22}$$

Here we have much more because each term $c_i\left(\theta\right)$ depends only on the $i'th$ component η_i of the nuisance parameter η. That means

$$c_i\left(\theta\right) =: c\left(\eta_i,\beta\right)$$

and that's why

$$C_n\left(\theta\right) = \frac{1}{n}\sum_{i=1}^{n} c\left(\eta_i,\beta\right). \tag{23}$$

The average structure of (**??**) and the independence of $\left(Z_i\right)_{i=1,..,n}$, introduced in (**??**), with common distribution

$$P_{\eta^o} = \prod_{i=1}^{n} P_{\eta_i^o}, \tag{24}$$

imply an average structure of the contrast function. We set

$$\widetilde{C}_n\left(Z,\theta\right) = \frac{1}{n}\sum_{i=1}^{n}\widetilde{c}\left(Z_i,\theta\right), \tag{25}$$

with

$$E_{\eta_i^0}\widetilde{c}\left(Z_i,\theta\right)=c_i\left(\theta\right). \tag{26}$$

By suppressing the dependence on the sample we write

$$\widetilde{C}_n\left(\theta\right)=\frac{1}{n}\sum_{i=1}^{n}\widetilde{c}_i\left(\theta\right). \tag{27}$$

The specific property $c_i\left(\theta\right)=:c\left(\eta_i,\beta\right)$ leads to

$$\widetilde{C}_n\left(\theta\right)=\frac{1}{n}\sum_{i=1}^{n}\widetilde{c}\left(\eta_i,\beta\right), \tag{28}$$

where $\widetilde{c}\left(\eta_i,\beta\right)$ is given by

$$E_{\eta_i^0}\widetilde{c}\left(Z_i,\eta_i,\beta\right)=c\left(\eta_i,\beta\right). \tag{29}$$

Then the approximation problem in (**??**) relies on the uniform convergence results of a sum of independent r.v. with expected value zero

$$\widetilde{C}_n\left(\theta\right)-C_n\left(\theta\right)=\frac{1}{n}\sum_{i=1}^{n}\widetilde{c}\left(\eta_i,\beta\right)-E_{\eta_i^0}\widetilde{c}\left(\eta_i,\beta\right). \tag{30}$$

The correction proposal is now to take instead of $\widetilde{c}\left(Z_i,\eta_i,\beta\right)$ another function $q\left(Z_i,\beta\right)$, which is independent of the nuisance parameter η_i and has the same expected value, such that

$$\forall i=1,..,n\ \forall\eta_i\in\mathcal{D}\ \forall\beta\in\mathcal{B}\quad E_{\eta_i}q\left(Z_i,\beta\right)=c\left(\eta_i,\beta\right). \tag{31}$$

It is sufficient to solve (**??**) for $i=1$ only. Rewrite it as integral deconvolution equation

$$\forall\beta\in\mathcal{B}\quad\forall\eta_1\in\mathcal{D}\qquad\int q\left(\eta_1+\varepsilon,\beta\right)p\left(\varepsilon\right)d\varepsilon=c\left(\eta_1,\beta\right), \tag{32}$$

where $p\left(\varepsilon\right)$ is the density of the i.i.d. errors in (**??**). The correction by deconvolution requires the complete knowledge of the density $p\left(\varepsilon\right)$ and depends mainly on $p\left(\varepsilon\right)$. Only for polynomial models the knowledge of the higher moments will be sufficient. We introduce the corrected contrast function by

$$\widetilde{C}_{cor,n}\left(\beta\right)=\frac{1}{n}\sum_{i=1}^{n}q\left(Z_i,\beta\right). \tag{33}$$

The new *corrected minimum contrast estimator (c- MCE)* for the parameter of interest $\widetilde{\beta}$ is then defined as a measurable solution of

$$\widetilde{\beta} \in \arg\min_{\beta\in\mathcal{B}} \frac{1}{n}\sum_{i=1}^{n} q\left(Z_i, \beta\right). \tag{34}$$

In Section 5 and Section 7 we present examples, where (**??**) is explicitly solvable. That are for instance the polynomial model, the exponential explicit model and arbitrary nonlinear smooth implicit models with Laplace error distribution.

The explicit solution of the deconvolution integral equation (**??**) is not given in every case. That's why we require the existence of an approximate solution. We call $q_\mu\left(., \beta\right), \mu > 0$ an approximate solution of (**??**) iff

$$\exists c_0 > 0 \;\forall \mu > 0 \quad \sup_{\beta\in\mathcal{B}}\sup_{\eta_1\in\mathcal{D}}\left|\int q_\mu\left(\eta_1+\varepsilon,\beta\right)p\left(\varepsilon\right)d\varepsilon - c\left(\eta_1,\beta\right)\right| \leq c_0\mu. \tag{35}$$

We introduce the approximate corrected estimation criterion

$$\widetilde{C}_\mu\left(\beta\right) = \frac{1}{n}\sum_{i=1}^{n} q_\mu\left(Z_i, \beta\right). \tag{36}$$

The related *approximate corrected minimum contrast estimator (ac-MCE)* for the parameter of interest $\widetilde{\beta}_\mu$ is then defined as the measurable solution of

$$\widetilde{\beta}_\mu \in \arg\min_{\beta\in\mathcal{B}} \frac{1}{n}\sum_{i=1}^{n} q_\mu\left(Z_i, \beta\right). \tag{37}$$

3 The corrected naive estimator in the explicit model

Consider the explicit model in (**??**) with regression function g. The interpretation of naive estimating is that the error-in-variables effect is neglected and the researcher naively tries to use the regression ideas.

The corrected naive estimator is based on the naive estimation of the unknown design points ξ_i^0 by the observations x_i. In the estimation criterion in (**??**)

$$\widetilde{C}_n\left(\theta\right) = \frac{1}{n}\sum_{i=1}^{n} \widetilde{c}\left(\eta_i, \beta\right),$$

the ξ_i are replaced by x_i. Then we have for the starting estimating procedure:

$$\widetilde{C}_{naive}(\theta) = \frac{1}{n}\sum_{i=1}^{n} \widetilde{c}_{naive}(x_i, \beta),$$

The correction equations get now an asymmetrical form. We adjust only with respect to the distribution of the ε_{2i} and that's why we need only the complete knowledge of the distribution of ε_{2i}.

Remember (**??**) and introduce for the explicit model $\widetilde{c}(y_i, \xi_i^0, \beta)$by

$$\int \widetilde{c}_{naive}\left(\xi_i^0, \beta\right) p(x_i)\, dx_i = \widetilde{c}\left(y_i, \xi_i^0, \beta\right).$$

Then correction proposal is now to take instead of $c(Z_i, \eta_i^0, \beta)$ another function $q(Z_i, \beta)$, which is independent of the nuisance parameters and has the same expected value, such that

$$\forall \beta \in \mathcal{B} \quad E_{\theta^0} \int q(Z_i, \beta)\, p(x_i)\, dx_i = E_{\theta^0} \widetilde{c}\left(y_i, \xi_i^0, \beta\right). \tag{38}$$

For mutually independent x_i, y_i rewrite it as integral deconvolution equation

$$\forall \beta \in \mathcal{B}\ \forall \xi_1\ \forall y \ \int q(y, \xi_1 + \varepsilon, \beta)\, p(\varepsilon)\, d\varepsilon = \widetilde{c}(y, \xi_1, \beta). \tag{39}$$

where $p(\varepsilon)$ is the density of ε_{21} in (**??**). Respectively to above we define

$$\widetilde{C}_{cor,n}(\beta) = \frac{1}{n}\sum_{i=1}^{n} q(Z_i, \beta).$$

Note that is exactly the proposal of Stefanski (1989) in [**?**].

Here we consider also an approximate solution $q_\mu(., \beta), \mu > 0$ of (**??**):

$$\sup_{y \in \mathrm{R}} \sup_{\beta \in \mathcal{B}} \sup_{\eta_1 \in \mathcal{D}} \left| \int q_\mu(y, \xi_1 + \varepsilon, \beta)\, p(\varepsilon)\, d\varepsilon - \widetilde{c}(y, \xi_1, \beta) \right| \le c_0 \mu. \tag{40}$$

Then we introduce the approximate corrected estimation criterion

$$\widetilde{C}_{ac,\mu}(\beta) = \frac{1}{n}\sum_{i=1}^{n} q_\mu(Z_i, \beta).$$

The related approximate naive corrected minimum contrast estimator for the parameter of interest $\widetilde{\beta}_{naive,\mu}$ is then defined as the measurable solution of

$$\widetilde{\beta}_{naive,\mu} \in \arg\min_{\beta \in \mathcal{B}} \frac{1}{n}\sum_{i=1}^{n} q_\mu(Z_i, \beta). \tag{41}$$

3.1 The corrected least squares estimator

The corrected least squares estimator is the name for the estimator in (**??**) with

$$\widetilde{c}_{naive}\left(x_i,\beta\right)=\widetilde{c}\left(Z_i,x_i,\beta\right)=\left(y_i-g\left(x_i,\beta\right)\right)^2$$

and

$$\widetilde{c}\left(y_i,\xi_i^0,\beta\right)=\left(y_i-g\left(\xi_i^0,\beta\right)\right)^2. \tag{42}$$

The contrast with respect to the least squares estimator in the model (**??**),

$$c\left(\xi_1,\beta\right)=w_{11}\left(g\left(\xi_1^0,\beta^0\right)-g\left(\xi_1^0,\beta\right)\right)^2+w_{12}\left(\xi_1^0-\xi_1\right)^2, \tag{43}$$

or respectively

$$c\left(\xi_1^0,\beta\right)=w_{11}\left(g\left(\xi_1^0,\beta^0\right)-g\left(\xi_1^0,\beta\right)\right)^2$$

corresponds to the naive estimator, because in (**??**) we have

$$E_{\theta^0}\widetilde{c}\left(y_i,\xi_i^0,\beta\right)=\left(g\left(\xi_1^0,\beta^0\right)-g\left(\xi_1^0,\beta\right)\right)^2+Var\left(\varepsilon_{1i}\right).$$

That means the naive correcting of the least squares methods leads to the same estimator as the symmetrical correcting in (**??**) with respect to (**??**).

For the naive correcting with (**??**) the deconvolution equation (**??**) becomes much more convenient. The key assumption for the new estimator is the requirement of the existence of two continuous functions f and h , Borel measurable with respect to the first argument:

Ex

$$\forall\xi_1\exists f\left(\xi_1,.\right)\in C\left(U\right)\ \forall\beta\in\mathcal{B}:E\,f\left(\xi_1+\varepsilon_1,\beta\right)=g\left(\xi_1,\beta\right) \tag{44}$$

$$\forall\xi_1\,\exists h\left(\xi_1,.\right)\in C\left(U\right)\ \forall\beta\in\mathcal{B}:E\,h\left(\xi_1+\varepsilon_1,\beta\right)=\left(g\left(\xi_1,\beta\right)\right)^2\ . \tag{45}$$

The *corrected least squares estimator* $\widetilde{\beta}_{lse}$ (c-l.s.e.) for β_0 in (**??**) is defined as a measurable solution of the optimization problem:

$$\widetilde{\beta}_{lse}\in\arg\min_{\beta\in\mathcal{B}}\frac{1}{n}\sum_{i=1}^{n}\left[\left(y_i-f\left(x_i,\beta\right)\right)^2+h\left(x_i,\beta\right)-\left(f\left(x_i,\beta\right)\right)^2\right]. \tag{46}$$

Let us denote the new optimization criterion by

$$\widehat{C}_{lse}\left(\beta\right)=\frac{1}{n}\sum_{i=1}^{n}\left[\left(y_i-f\left(x_i,\beta\right)\right)^2+h\left(x_i,\beta\right)-\left(f\left(x_i,\beta\right)\right)^2\right]. \tag{47}$$

The nuisance parameters $\xi_1, .., \xi_n$ are eliminated, the criterion function depends on the parameter β of interest only. The first term of (**??**) can be interpreted as the least squares criterion in the transformed (projected) regression model

$$y_i = f(x_i, \beta_0) + u_i, \text{ with } u_i = f(x_i, \beta_0) - g\left(\xi_i^0, \beta_0\right) + \varepsilon_{1i} \tag{48}$$

and the other terms of (**??**) are a correction for the covariance between the $f(x_i, \beta_0)$ and the transformed error term u_i. We have

$$Cov_\theta\left(u_i,\ f(x_i, \beta)\right) = E_\theta\left(f(x_i, \beta) - g(\xi_i, \beta)\right)^2 = E_\theta(f(x_i, \beta)^2) - g(\xi_i, \beta)^2 . \tag{49}$$

3.2 The Fourier transform method

The main point is the construction of the auxiliary functions f and h in **Ex.** The following lemma gives some hint, how to derive them. We will denote the Fourier transform of an arbitrary function $f \in L^2(R)$ by $\widehat{f}$, such that $\widehat{p}$ denotes the characteristic function of ε_{21} :

$$\widehat{f}(t) := \int \exp(itx) f(x)\, dx \quad \text{and} \quad f(x) = \frac{1}{2\pi} \int \exp(-itx) \widehat{f}(t)\, dt. \tag{50}$$

Let us suppose further for the Fourier transform of the regression functions $g(x, \beta)$ and of the squared regression functions $(g(x, \beta))^2 = G(x, \beta)$:

ExFour

$$\forall\, \beta \in \Theta : \int \left(\frac{\widehat{g}(t, \beta)}{\widehat{p}(t)}\right)^2 dt < \infty \quad \text{and} \quad \int \left(\frac{\widehat{G}(t, \beta)}{\widehat{p}(t)}\right)^2 dt < \infty . \tag{51}$$

This assumption **ExFour** is really strong. We will give an example later.

Lemma 3.1 *Suppose for all* $\beta \in \Theta : g(., \beta) \in L^4(\mathrm{R}^q)$ *and* ε_{21} *has a known density* p, *with* $p(\varepsilon_{21}) = p(-\varepsilon_{21})$.

Then under **ExFour** *the transformations* $f(x_1, \beta)$ *and* $h(x_1, \beta)$ *in* **Ex** *are given by*

$$f(x_1, \beta) = \frac{1}{2\pi} \int \exp(-itx_1) \frac{\widehat{g}(t, \beta)}{\widehat{p}(t)} dt. \tag{52}$$

and

$$h(x_1,\beta) = \frac{1}{2\pi}\int \exp(-itx_1)\,\frac{\widehat{G}(t,\beta)}{\widehat{p}(t)}dt. \tag{53}$$

Proof.

Inside of that proof we will suppress the indices 1 and 21. The formula (**??**) describes the convolution

$$f * p = g \tag{54}$$

or otherwise

$$\forall\,\beta\in\Theta : \int f(\xi-\varepsilon,\beta)\,p(-\varepsilon)\,d\varepsilon = \int f(\xi-\varepsilon,\beta)\,p(\varepsilon)\,d\varepsilon = g(\xi,\beta)\ .$$

From (**??**) it follows by the theorem on the Fourier transform of convolution that

$$\forall\,\beta\in\Theta : \widehat{f}(t,\beta)\ \widehat{p}(t) = \widehat{g}(t,\beta)\ . \tag{55}$$

Now we calculate the transformation $f(x,\beta)$ by the inversion of its Fourier transform $\widehat{f}(t,\beta)$ in (**??**)

$$f(x,\beta) = \frac{1}{2\pi}\int \exp(-itx)\,\frac{\widehat{g}(t,\beta)}{\widehat{p}(t)}dt.$$

The same is done for $h(x,\beta)$. ■

Note, we use here the same approach as in the deconvolution problem, compare for instance, Stefanski, Carroll (1990), [**?**].

4 The consistency

In this section we give a formal consistency proof for the approximate corrected minimum contrast estimator (ac-MCE).

The correction procedure leads to a usual parametric estimation problem with a finite number of parameters. The difficulties with the nuisance parameters are shifted to the derivation of the new approximate corrected contrast functions $\widetilde{C}_\mu(\beta)$. That will be done for special cases in the following sections.

We require the separation condition on the contrast only for fixed η^0, given (**??**), and for the Euclidean distance in R^p :

Sep For $\rho : \mathrm{R}_+ \to \mathrm{R}_+$ be strictly increasing with $\rho(0) = 0$, and for all $\beta \in \mathcal{B}$

$$C_n\left(\eta^0, \beta\right) - C_n\left(\eta^0, \beta^0\right) \geq \rho\left(\left\|\beta - \beta^0\right\|\right).$$

It is possible to derive an analogous result to (**??**).

Lemma 4.1 *Let C_n be a contrast with* **Sep.** *Then for the ac-MCE $\widetilde{\beta}_\mu$ defined in (**??**) with (**??**) it holds for*

$$\mathcal{B}(\tau) = \left\{\beta \in \mathcal{B} : \left\|\beta - \beta^0\right\| > \tau\right\}$$

that

$$\forall \tau > 0 \quad P_{\theta^0}\left(\left\|\widetilde{\beta}_\mu - \beta^0\right\| > \tau\right) \leq \tag{56}$$

$$\leq P_{\theta^0}\left(\sup_{\beta \in \mathcal{B}(\tau)} \left|\Delta \widetilde{C}_{ac,\mu}(\beta) - E_{\theta^0} \Delta \widetilde{C}_{ac,\mu}(\beta)\right| \geq \rho(\tau) - c_0 \mu\right).$$

□

Proof. From **Sep** it follows for $\Delta C_n(\eta^0, \beta) = C_n(\eta^0, \beta) - C_n(\eta^0, \beta^0)$ that for all $\beta \in \mathcal{B}(\tau)$ holds

$$\Delta C_n\left(\eta^0, \beta\right) \geq \rho\left(\left\|\beta - \beta^0\right\|\right) \geq \rho(\tau) > 0.$$

Thus

$$\inf_{\beta \in \mathcal{B}(\tau)} \Delta C_n\left(\eta^0, \beta\right) \geq \rho(\tau). \tag{57}$$

From (**??**) we have

$$\left|\Delta C_n\left(\eta^0, \beta\right) - E_{\theta^0} \Delta \widetilde{C}_{ac,\mu}(\beta)\right| \leq c_0 \mu.$$

Thus (**??**) implies

$$\inf_{\beta \in \mathcal{B}(\tau)} E_{\theta^0} \Delta \widetilde{C}_{ac,\mu}(\beta) \geq \rho(\tau) - c_0 \mu. \tag{58}$$

Under $\widetilde{\beta}_\mu \in \mathcal{B}(\tau)$ we have $\widetilde{C}_{ac,\mu}\left(\widetilde{\beta}_\mu\right) \leq \widetilde{C}_{ac,\mu}(\beta^0)$, thus

$$\Delta \widetilde{C}_{ac,\mu}\left(\widetilde{\beta}_\mu\right) \leq 0. \tag{59}$$

From (**??**) and (**??**) the following chain of inequalities is valid

$$\rho(\tau) - c_0\mu \leq \inf_{\beta\in\mathcal{B}(\tau)} E_{\theta^0}\Delta\widetilde{C}_{ac,\mu}(\beta) - \Delta\widetilde{C}_{ac,\mu}\left(\widetilde{\beta}_\mu\right) \tag{60}$$

$$\leq \sup_{\beta\in\mathcal{B}(\tau)} \left|E_{\theta^0}\Delta\widetilde{C}_{ac,\mu}(\beta) - \Delta\widetilde{C}_{ac,\mu}(\beta)\right|,$$

which yields the statement of the lemma. ■

Note, the structure of $\Delta\widetilde{C}_{ac,\mu}(\beta) - E_{\theta^0}\Delta\widetilde{C}_{ac,\mu}(\beta)$ is a sum of independent not identically distributed random values with expected value zero:

$$\Delta\widetilde{C}_{ac,\mu}(\beta) - E_{\theta^0}\Delta\widetilde{C}_{ac,\mu}(\beta) = \frac{1}{n}\sum_{i=1}^{n}\left[\Delta q_\mu(Z_i,\beta) - E_{\theta^0}\Delta q_\mu(Z_i,\beta)\right]. \tag{61}$$

Denote

$$S_n(\beta) = \frac{1}{n}\sum_{i=1}^{n}\left(q_\mu(Z_i,\beta) - E_{\theta^0}q_\mu(Z_i,\beta)\right).$$

We know from the law of large numbers that $S_n(\beta) \to 0$ in probability. The uniform convergence in (**??**) is required with respect to a set of fixed dimension p.

Under additional assumptions on the moments of $q_\mu(Z_i,\beta)$ we obtain the convergence of (**??**) and thus the consistence of the new estimator.

The first assumption is a moment condition on the error distribution with respect to $q_\mu(Z_i,\beta)$.

QM There exist a constant c and a real number k, $k \geq 2$ and a function $\gamma_1(k) > 0$, such that for all $\mu \leq \mu_0$

$$\sup_{\beta\in\mathcal{B}} \frac{1}{n}\sum_{i=1}^{n}\left(E_{\theta^0}\left|q_\mu(Z_i,\beta) - E_{\theta^0}q_\mu(Z_i,\beta)\right|^{2k}\right)^{\frac{1}{k}} \leq \frac{c}{\mu^{\gamma_1(k)}}.$$

The next assumption is a Lipschitz type condition on $q_\mu(Z_i,\beta)$.

QL There exist a random variable $M_{(n)}$, a constant c and a real number $k \geq 2$ and a function $\gamma_2(k) > 0$, such that for all n and for all $\beta, \beta' \in \mathcal{B}$

$$\frac{1}{n}\sum_{i=1}^{n}\left|q_\mu(Z_i,\beta) - E_{\theta^0}q_\mu(Z_i,\beta) - \left(q_\mu(Z_i,\beta') - E_{\theta^0}q_\mu(Z_i,\beta')\right)\right|^2$$

$$\leq M_{(n)} \left\| \beta - \beta' \right\|^2,$$

with

$$E_{\theta^0} {M_{(n)}}^k \leq \frac{c}{\mu^{\gamma_2(k)}}.$$

First let us remind the following result:

Lemma 4.2 *Suppose* **QM , QL.**

Then there exists a constant const, such that for all ϵ

$$P_{\theta^0} \left(\sup_{\beta \in \mathcal{B}} \left| \frac{1}{n} \sum_{i=1}^{n} \left(q_\mu \left(Z_i, \beta \right) - E_{\theta^0} q_\mu \left(Z_i, \beta \right) \right) \right| \geq \epsilon \right) \tag{62}$$

$$\leq const\, \epsilon^{-2k} \mu^{-\frac{\gamma_1(k)(2k-p)+\gamma_2(k)p}{2k}} n^{-k+\frac{p}{4}}.$$

□

Proof.

Because $\mathcal{B}$ is a compact subset of R^p, $\mathcal{B}$ is bounded. There exists a constant R for which $\mathcal{B} \subset K\left(\beta^0, R\right) = \left\{\beta : \left\|\beta - \beta^0\right\| \leq R\right\}$. For each d there exist balls with radius d and center β^l

$$K\left(\beta^l, d\right) = \left\{\beta : \left\|\beta - \beta^l\right\| \leq d\right\},$$

$$l = 1, ..., N\left(d\right), \quad N\left(d\right) \leq \left(\left[\frac{2R}{d} \right] + 1 \right)^p,$$

where [.] denotes the Gaussian brackets, which cover $\mathcal{B}$:

$$\mathcal{B} \subset \bigcup_{l=1}^{N(d)} K\left(\beta^l, d\right).$$

We have

$$P_{\theta^0} \left(\sup_{\beta \in \mathcal{B}} \left| S_n\left(\beta\right) \right| \geq \epsilon \right) \leq$$

$$\leq \sum_{l=1}^{N(d)} P_{\theta^0} \left(\sup_{\beta \in K\left(\beta^l, d\right)} \left| S_n\left(\beta\right) \right| \geq \epsilon \right)$$

$$\leq \sum_{l=1}^{N(d)} P_{\theta^0}\left(\sup_{\beta\in K\left(\beta^l,d\right)}\left|S_n(\beta)-S_n\left(\beta^l\right)\right|\geq\frac{\epsilon}{2}\right)+P_{\theta^0}\left(\left|S_n\left(\beta^l\right)\right|\geq\frac{\epsilon}{2}\right). \quad (63)$$

From the Hölderian inequality and the Lipschitz condition **QL** we get

$$\left|S_n(\beta)-S_n\left(\beta^l\right)\right|^2$$

$$=\left|\frac{1}{n}\sum_{i=1}^{n}\left(q_\mu(Z_i,\beta)-q_\mu\left(Z_i,\beta^l\right)\right)+\left(E_{\theta^0}\left(q_\mu(Z_i,\beta)-q_\mu\left(Z_i,\beta^l\right)\right)\right)\right|^2$$

$$\leq M_{(n)}\left\|\beta-\beta^l\right\|^2\leq M_{(n)}d^2.$$

Now we apply the Chebychev inequality and **QL** on the first terms in (**??**)

$$\sum_{l=1}^{N(d)} P_{\theta^0}\left(\sup_{\beta\in K\left(\beta^l,d\right)}\left|S_n(\beta)-S_n\left(\beta^l\right)\right|\geq\frac{\epsilon}{2}\right)$$

$$\leq \sum_{l=1}^{N(d)} P_{\theta^0}\left(\sup_{\beta\in K\left(\beta^l,d\right)}\left|S_n(\beta)-S_n\left(\beta^l\right)\right|^2\geq\frac{\epsilon^2}{4}\right).$$

$$\leq N(d)\,P_{\theta^0}\left(M_{(n)}\geq\frac{\epsilon^2}{d^2 4}\right)\leq N(d)\frac{4^k d^{2k} E M_{(n)}^k}{\epsilon^{2k}}\leq const\frac{d^{2k-p}}{\epsilon^{2k}\mu^{\gamma_2(k)}}.$$

For the second term in (**??**) we apply the Whittle inequality, compare Whittle (1960), [**?**], for a real number $s, s\geq 2$, that is

$$E\left|\sum_{j=1}^{n} b_j X_j\right|^s\leq\frac{2^{\frac{s}{2}+s}}{\sqrt{\pi}}\Gamma\left(\frac{s+1}{2}\right)\left(\sum_{j=1}^{n} b_j^2\left(E\left|X_j\right|^s\right)^{\frac{2}{s}}\right)^{\frac{s}{2}},$$

and **QM**

$$\sum_{l=1}^{N(d)} P_{\theta^0}\left(\left|S_n\left(\beta^l\right)\right|\geq\frac{\epsilon}{2}\right)\leq const\, d^{-p}n^{-k}\mu^{-\gamma_1(k)}\epsilon^{-2k}.$$

Summarizing we obtain

$$P_{\theta^0}\left(\sup_{\beta\in\mathcal{B}_0}\left|\frac{1}{n}\sum_{i=1}^{n}\left(q_\mu(Z_i,\beta)-E_{\theta^0}q_\mu(Z_i,\beta)\right)\right|\geq\epsilon\right)\leq const\,\left[\frac{d^{2k-p}}{\epsilon^{2k}\mu^{\gamma_2(k)}}+\frac{d^{-p}}{\epsilon^{2k}\mu^{\gamma_1(k)}}n^{-k}\right]. \quad (64)$$

Now we choose the d in an optimal way, that it minimizes the right hand side of (**??**). Consider the function

$$F(d) = Ad^{2k-p} + Bd^{-p},\ d > 0.$$

For

$$d = const\left(\frac{B}{A}\right)^{\frac{1}{2k}}$$

the minimum of $F(d)$ is attained and that is

$$\min_{d} F(d) = const A^{\frac{p}{2k}} B^{\frac{2k-p}{2k}}.$$

Applying this to (**??**) for

$$A = \frac{1}{\epsilon^{2k}\mu^{\gamma_2(k)}}, \quad B = \frac{1}{\epsilon^{2k}\mu^{\gamma_1(k)}}n^{-k}$$

we obtain (**??**).

■

Choosing now the constants in a convenient form we get the following auxiliary result:

Lemma 4.3 *Suppose* **QM** *with* $k \geq 2$ **,** **QL** *with* $k \geq 2$**.** *Then for*

$$k = \frac{p}{2} + 2\delta_1, \quad \delta_1 > 0 \tag{65}$$

and δ_1 *such that*

$$\mu = const\, n^{-r}, \quad r > 0 \tag{66}$$

with

$$r = \frac{2k(\delta_1 - \kappa)}{\gamma_1(k)(2k-p) + \gamma_2(k)p} \tag{67}$$

there exists a constant const, such that for all ϵ

$$P_{\theta^0}\left(\sup_{\beta\in\mathcal{B}_0}\left|\frac{1}{n}\sum_{i=1}^{n}\left(q_\mu(Z_i,\beta) - E_{\theta^0}q_\mu(Z_i,\beta)\right)\right| \geq \epsilon\right) \leq \epsilon^{-2k} const\, n^{-\kappa}. \tag{68}$$

□

Proof. The constants μ, k are chosen such that

$$\mu^{-\frac{\gamma_1(k)(2k-p)+\gamma_2(k)p}{2k}} n^{-k+\frac{p}{4}} \leq n^{-\kappa}.$$

■

Combining the results of both lemmata and using the lemma of Borel Cantelli for $\kappa = 1 + \delta$ we obtain the strong consistency of the new estimator ac-MCE, defined in (**??**).

Theorem 4.4 *1. Suppose in (**??**) that*

$$\mu = \mu(n) \to 0 \tag{69}$$

and **Sep**, *with* $\rho(\varepsilon) \geq c\varepsilon$, **QM, QL** *with* k *given in (**??**), (**??**), (**??**) with*

$$\kappa > 0.$$

Then

$$\forall \tau > 0 \quad \lim_{n\to\infty} \sup_{\beta^0 \in \mathcal{B}} P_{\theta^0}\left(\left\|\widetilde{\beta}_{\mu(n)} - \beta^0\right\| > \tau\right) = 0. \tag{70}$$

2. Suppose **Sep**, *with* $\rho(\varepsilon) \geq c\varepsilon$,**QM, QL** *with* k *given in (**??**), (**??**), (**??**) with*

$$\kappa > 1$$

Then

$$\widetilde{\beta}_{\mu(n)} \to \beta^0 \quad a.s.. \tag{71}$$

□

Proof. From Lemma **??** it follows

$$\forall \tau > 0 \quad P_{\theta^0}\left(\left\|\widetilde{\beta}_{\mu} - \beta^0\right\| > \tau\right) \leq P_{\theta^0}\left(\sup_{\beta \in \mathcal{B}(\tau)} |S_n(\beta)| \geq \rho(\tau) - \mu\right)$$

$$\leq P_{\theta^0}\left(\sup_{\beta \in \mathcal{B}} |S_n(\beta)| \geq \rho(\tau) - \mu\right).$$

For $\epsilon = \rho(\tau) - \mu$ we apply Lemma **??** and Lemma **??**

$$P_{\theta^0}\left(\sup_{\beta \in \mathcal{B}} |S_n(\beta)| \geq \rho(\tau) - \mu\right) \leq const\, n^{-\kappa}.$$

If we choose $\kappa = 1 + \delta$ then we obtain by Borel Cantelli the strong consistency result (**??**). ■

We close this section by formulating the consistency result for the exact corrected MCE. Introduce the related versions of **QM** and **QL** for the exact corrected estimation function q, satisfying (**??**).

QM0 There exist a constant c and a real number k, $k \geq 2$

$$\sup_{\beta\in\mathcal{B}} \frac{1}{n}\sum_{i=1}^{n}\left(E_{\theta^0}\left|q\left(Z_i,\beta\right) - E_{\theta^0}q\left(Z_i,\beta\right)\right|^{2k}\right)^{\frac{1}{k}} \leq c.$$

QL0 There exist a random variable $M_{(n)}$, a constant c and a real number $k \geq 2$, such that for all n and for all $\beta, \beta' \in \mathcal{B}$

$$\frac{1}{n}\sum_{i=1}^{n}\left|q\left(Z_i,\beta\right) - E_{\theta^0}q\left(Z_i,\beta\right) - \left(q\left(Z_i,\beta'\right) - E_{\theta^0}q\left(Z_i,\beta'\right)\right)\right|^2$$

$$\leq M_{(n)}\left\|\beta - \beta'\right\|^2,$$

with

$$E_{\theta^0}{M_{(n)}}^k \leq c.$$

Theorem 4.5

1. *Suppose* **Sep**, *with* $\rho\left(\varepsilon\right) \geq c\varepsilon$, **QM0** , **QL0** *with* $k \geq 2$ *and*

$$k = \frac{p}{2} + \delta_{1,} \quad \delta_1 > 0.$$

Then

$$\forall\tau > 0 \quad \lim_{n\to\infty}\sup_{\beta^0\in\mathcal{B}} P_{\theta^0}\left(\left\|\widetilde{\beta} - \beta^0\right\| > \tau\right) \leq const\,\tau^{-2k}n^{-k+\frac{p}{4}}. \tag{72}$$

2. *Suppose* **Sep**, *with* $\rho\left(\varepsilon\right) \geq c\varepsilon$,**QM0** , **QL0** *with* $k \geq 2$ *and*

$$k > \frac{p}{2} + 1$$

Then

$$\widetilde{\beta} \to \beta^0 \quad a.s.. \tag{73}$$

□

Proof.

Proof goes along the line above. We can interpret the assumptions **QM0 , QL0** as **QM** with $\gamma_1(k) = 0$, **QL** with $\gamma_2(k) = 0$ Because of $\gamma_1(k) = 0$ and $\gamma_2(k) = 0$ we have in (**??**)

$$P_{\theta^0}\left(\sup_{\beta\in\Theta_0}\left|\frac{1}{n}\sum_{i=1}^{n}\left(q\left(Z_i,\beta\right) - E_{\theta^0}q\left(Z_i,\beta\right)\right)\right| \geq \epsilon\right)$$
$$\leq const\,\epsilon^{-2k}n^{-k+\frac{p}{4}},$$

which delivers together with Lemma **??** the statement (**??**). For $k > \frac{p}{4} + 1$ the rate in (**??**) is more than n^{-1}. Then (**??**) follows from the Lemma of Borel Cantelli. ■

5 Special cases for the corrected least squares estimator

In this section we summarize special models, where we can derive explicitly the functions f and h in the estimation criterion of $\widetilde{\beta}_{lse}$

$$\widetilde{\beta}_{lse} \in \arg\min_{\beta\in\mathcal{B}^c}\frac{1}{n}\sum_{i=1}^{n}\left[\left(y_i - f\left(x_i,\beta\right)\right)^2 + h\left(x_i,\beta\right) - \left(f\left(x_i,\beta\right)\right)^2\right].$$

There is also one example in Subsubsection 5.4.1, where we derive for f an approximate solution f_μ.

5.1 The polynomial functional relation model

Consider $q = 2$ and for the regression function a polynomial of k'th order:

$$g\left(\xi_1,\beta\right) = \sum_{r=0}^{k}\beta_l\left(\xi_1\right)^r. \tag{74}$$

Note, because it is nonlinear in the nuisance parameter ξ_1, it is nonlinear model in this context. Further we need the complete knowledge of all moments up to the $2k'$th order of the error-in-variables distribution:

Mom For some k the moments

$$\mu_r = E\left(\varepsilon_{21}\right)^r, r = 1, ..., 2k$$

are known.

The correction for a real valued monomial $z^r$of order r

$$Ef_r\left(z+\varepsilon\right) = z^r, \tag{75}$$

is given by

$$f_0\left(z\right) = 1, \quad f_1\left(z\right) = z \tag{76}$$

and for all $r \geq 2$

$$f_r\left(z\right) = z^r - \sum_{l=0}^{r-2} \binom{r}{l} f_l\left(z\right) \mu_{r-l}. \tag{77}$$

Compare for instance Zwanzig (1996) [**?**] for more properties of f_r. The statement (**??**) is easily checked by the Binomial formula

$$\left(z+\varepsilon\right)^r = \sum_{l=0}^{r} \binom{r}{l} z^l \varepsilon^{r-l},$$

because:

$$Ef_r\left(z+\varepsilon\right) = E\left(z+\varepsilon\right)^r - \sum_{l=0}^{r-2} \binom{r}{l} Ef_l\left(z+\varepsilon\right) \mu_{r-l}$$

$$= E\left(\sum_{l=0}^{r} \binom{r}{l} z^l \varepsilon^{r-l}\right) - \sum_{l=0}^{r-2} \binom{r}{l} Ez^l E\varepsilon^{r-l} = z^r.$$

Please do not mix the moments μ_r and the parameter for the approximation in (**??**);also the here introduced f_r are not the function introduced in **Ex.**

For the model (**??**) the functions introduced in **Ex** are

$$f\left(\xi_1, \beta\right) = \sum_{r=0}^{k} \beta_r f_r\left(\xi_1\right) \tag{78}$$

and

$$h\left(\xi_1, \beta\right) = \sum_{r_1=0}^{k} \sum_{r_2=0}^{k} \beta_{r_1} \beta_{r_2} f_{r_1+r_2}\left(\xi_1\right). \tag{79}$$

In the application of polynomials models to astrometric plate reduction we have two dimensional polynomials in $\xi_1 = (\xi_{11}, \xi_{12})$ of third order. The correction formulas corresponding to (**??**) and (**??**) for the special astrometric setting up are given in Zwanzig (1997), [**?**].

5.2 The exponential model

Consider the model

$$g(\xi_1, \beta) = \beta_0 + \beta_1 \exp(\beta_2 \xi_1). \tag{80}$$

Assume the knowledge of the exponential moments of the error-in-variables distribution.

ExpMom

$$\varepsilon_{2i} \quad i.i.d. \text{ and } E \exp \alpha \varepsilon_{21} = m(\alpha).$$

We have

$$\int \exp(\beta_2 (\xi_1 + \varepsilon_{21})) p(\varepsilon_{21}) d\varepsilon_{21} = \exp(\beta_2 \xi_1) m(\beta_2).$$

For the model (**??**) the functions introduced in **Ex** are

$$f(x_1, \beta) = \beta_0 + \beta_1 m(\beta_2)^{-1} \exp(\beta_2 x_1) \tag{81}$$

and

$$h(x_1, \beta) = \beta_0^2 + 2\beta_1 \beta_0 m(\beta_2)^{-1} \exp(\beta_2 x_1) + \beta_1^2 m(2\beta_2)^{-1} \exp(2\beta_2 x_1). \tag{82}$$

5.3 Laplace distribution

Assume the error-in-variables distribution is the Laplace one. Using the results given in Appendix we can derive the corrected least squares criterion for a great class of smooth regression functions.

La1

$$\varepsilon_{2i} \quad i.i. \text{ Laplace distributed with density } p,$$

$$p(u) = \prod_{i=1}^{q-1} \frac{1}{2} \exp(-|u_i|).$$

For $q = 2$ we get for regressions functions $g(., \beta)$ satisfying **G1** and **G2** that under Laplace distribution **La1** the functions introduced in **Ex** from Lemma **??**:

$$f(x_1, \beta) = g(x_1, \beta) - g^{\xi\xi}(x_1, \beta) \tag{83}$$

and

$$h(x_1, \beta) = g(x_1, \beta)^2 - 2\left(g^{\xi}(x_1, \beta)\right)^2 - 2g(x_1, \beta)\, g^{\xi\xi}(x_1, \beta)\,. \tag{84}$$

For $q > 2$ we obtain the correspondent functions from Lemma **??**.

5.3.1 An approximate corrected least squares estimator

Let us give at least one example for an corrected least squares estimator. We assume **La1** with $q = 2$ and for the regression function in the explicit model (**??**),

$$g(\xi_1, \beta) = \beta\,|\xi_1|\,, \quad \beta \in [-d, d]\,.$$

We get the formal solution of the first deconvolution equation (**??**) in **Ex** from Lemma **??** (We set in Lemma **??** $y = 0$ and $g(\xi) = \xi$). That means

$$f(x_1, \beta) = \beta\,(|x_1| - 2\delta(x_1))\,.$$

We introduce a nonnegative, normalized, smooth kernel function $\varpi : \mathrm{R} \to \mathrm{R}$ with bounded support,

$$\sup p(\varpi) \subseteq [-1, 1] \quad \varpi \in C^3(\mathrm{R})\,, \quad \int \varpi(t)\, dt = 1. \tag{85}$$

Then we approximate the $\delta-$function by

$$\delta_\mu(t) := \frac{1}{\mu}\varpi\left(\frac{t}{\mu}\right). \tag{86}$$

Hence we get for the approximate criterion

$$f_\mu(x_1, \beta) = \beta\,(|x_1| - 2\delta_\mu(x_1))\,.$$

The second function h in (**??**) in **Ex** is related to $g(\xi)^2 = \beta^2\xi^2$, which satisfies the conditions **G1,G2** of Lemma **??**. We get

$$h(x_1, \beta) = \beta\left(x^2 - 2\right).$$

Summarizing we have the approximate corrected estimation criterion (**??**) concerning to (**??**):

$$C_\mu(\beta) = \sum_{i=1}^{n} (y_i - \beta(|x_i| - 2\delta_\mu(x_i)))^2 - 2\beta^2\left(-2|x_i|\delta_\mu(x_i) + 2\delta_\mu^2(x_i) + 1\right)$$

$$= \sum_{i=1}^{n}\left(y_i^2 - 2y_i f_\mu(x_i,\beta) + \beta\left(x_i^2 - 2\right)\right).$$

It remains to show (**??**). We have

$$|Ef_\mu(x_1,\beta) - Ef(x_1,\beta)| = |Ef_\mu(\xi_1 + \varepsilon_{11},\beta) - \beta|\xi_1||$$

$$\leq 2|\beta|\frac{\mu}{2} \leq const\,\mu.$$

5.4 Application of the Fourier transform method

Because of the strong condition **ExFour** the application of the Fourier transform method is useful for heavy tails error distributions. Let us consider the Gamma distribution.

Gam Let ε_{2i} be $i.i.$ centered Gamma distributed with parameters $\vartheta > 0, r$ natural number and density p,

$$p(u) = \begin{cases} \frac{\vartheta^r}{(r-1)!}\left(u - \frac{r}{\vartheta}\right)^{r-1} \exp\left(-\vartheta\left(u - \frac{r}{\vartheta}\right)\right) & \text{for } u \geq \frac{r}{\vartheta} \\ 0 & \text{else} \end{cases}.$$

Further we need that the regression function is smooth enough.

Diff

$$g(.,\beta) \in S(\mathrm{R}), \text{ set of primary function for tempered distribution.}$$

It follows from Lemma **??**, by using the formulas for the characteristic function

$$\widehat{p}(\lambda) = \frac{1}{\sqrt{2\pi}} e^{i\lambda m} \frac{\vartheta^r}{(\vartheta - i\lambda)^r},$$

that the functions in **Ex** are

$$f(x,\beta) = \frac{1}{\vartheta^r}\sum_{k=0}^{r}\binom{k}{r}C_r^k\vartheta^{-k}g^{(k)}\left(x-\frac{r}{\vartheta},\beta\right)$$

and

$$h(x,\beta) = \frac{1}{\vartheta^r}\sum_{k=0}^{r}\binom{k}{r}\vartheta^{-k}\left(g^2\left(x-\frac{r}{\vartheta},\beta\right)\right)^{(k)},$$

where $g^{(k)}(x,\beta)$ denotes the k'th derivatives with respect to the first argument.

6 On L1-type approximate corrected minimum contrast estimator under Laplace error distribution

Assume model (**??**) with $q = 2$. In this section we consider two special cases for approximate correcting, both related to the L_1-norm, but in a different way.

The first estimator is based on the naive correcting (**??**) of

$$\widetilde{c}(y_1,\xi_1,\beta) = |y_1 - g(\xi_1,\beta)|.$$

The respected contrast

$$E_{\theta^0}|y_1 - g(\xi_1,\beta)|$$

has a lower bound in terms of the L_2- norm. Compare for instance Zwanzig (1997) [**?**].

The second estimator is given through the approximate corrected estimating criterion for the contrast function in (**??**) with $r = 1$

$$c(\xi_1,\beta) = \left|g(\xi_1,\beta) - g\left(\xi_1,\beta^0\right)\right|. \tag{87}$$

The deconvolution equations of both estimators are only approximately solvable in the sense of (**??**).

6.1 The approximate corrected naive L_1-estimator

Consider the explicit model (**??**) and assume **La1** with $q = 2$**.**

We define the *approximate corrected naive L_1-estimator* as a measurable solution of

$$\widetilde{\beta}_\mu \in \arg\min_{\beta\in\mathcal{B}} \frac{1}{n}\sum_{i=1}^{n} q_\mu\left(y_i, x_i, \beta\right)$$

with $q_\mu\left(.,.,\beta\right)$ such that:

$$\sup_{y_1\in\mathrm{R}}\sup_{\beta\in\mathcal{B}}\sup_{\xi_1\in[-a,a]}\left|\int q_\mu\left(y_1, \xi_1+\varepsilon, \beta\right)p\left(\varepsilon\right)d\varepsilon - \left|y_1 - g\left(\xi_1,\beta\right)\right|\right| \le c_0\mu. \quad (88)$$

The idea is to approximate $\left|y_1 - g\left(\xi_1,\beta\right)\right|$ by a smooth function and to apply Lemma **??**. The first step is formulated in the following Lemma.

Lemma 6.1 *Define*

$$A_\mu(t) = \int_0^t \Delta_\mu\left(s\right)ds, \qquad t\in\mathrm{R}, \quad (89)$$

with

$$\Delta_\mu\left(s\right) = 2\int_0^s \delta_\mu\left(v\right)dv, \quad (90)$$

where

$$\delta_\mu\left(v\right) = \frac{1}{\mu}\varpi\left(\frac{v}{\mu}\right), \quad (91)$$

and ϖ is a nonnegative,even, normalized smooth kernel function with support $[-1,1]$ and $\int\varpi = 1$.

Then

$$\forall\, t\in\mathrm{R} \quad \left|A_\mu(t) - |t|\right| \le \mu. \quad (92)$$

Proof. Note, both functions $A_\mu(t)$ and $|t|$ are even, thus it is enough to consider $t\ge 0$. For $\mu < t$ it holds

$$\left|A_\mu(t) - |t|\right| = \left|\int_0^t \Delta_\mu\left(s\right)ds - t\right| = \left|\int_0^\mu \Delta_\mu\left(s\right)ds + \int_\mu^t \Delta_\mu\left(s\right)ds - t\right|.$$

Further for $s > \mu$

$$\Delta_\mu\left(s\right) = \int_0^s 2\frac{1}{\mu}\varpi\left(\frac{v}{\mu}\right)dv = 2\int_0^\mu \frac{1}{\mu}\varpi\left(\frac{v}{\mu}\right)dv = 2\int_0^1 \varpi\left(v\right)dv = \int_{-1}^1 \varpi\left(v\right)dv = 1.$$

Thus

$$\left|\int_0^\mu \Delta_\mu(s)\,ds + \int_\mu^t ds - t\right| = \left|\int_0^\mu \Delta_\mu(s)\,ds - \mu\right|.$$

From $\Delta_\mu(s) \le 1$ follows that

$$|A_\mu(t) - |t|| = \left|\int_0^\mu \Delta_\mu(s)\,ds - \mu\right| \le \max\left(\int_0^\mu \Delta_\mu(s)\,ds, \mu\right) = \mu.$$

For $0 < t \le \mu$ we have

$$|A_\mu(t) - |t|| = \left|\int_0^t \Delta_\mu(s)\,ds - t\right| \le \max\left(\int_0^t \Delta_\mu(s)\,ds, t\right) = t \le \mu.$$

■

In order to show (**??**) it remains to show that

$$Eq_\mu(y_1, x_1, \beta) = A_\mu(y_1 - g(x_1, \beta)).$$

That will be done in the proof of the following theorem with help of Lemma **??**.

We use a further smoothness conditions on the regression function.

G1

$$\int |g(t)| \exp(-|t|)\,dt < \infty,$$

$$\int |g''(t)| \exp(-|t|)\,dt < \infty$$

G2

$$\lim_{|t|\to\infty} g(t) \exp(-|t|) = 0$$

$$\lim_{|t|\to\infty} g'(t) \exp(-|t|) = 0$$

Theorem 6.2 *Assume for all $\beta \in \Theta$ the regression function $g(.,\beta) : \mathrm{R} \longrightarrow \mathrm{R}$, $g(.,\beta) \in C^2(\mathrm{R})$, $g(.,\beta)$ satisfies* **G1,G2** *uniformly with respect to $\beta \in \Theta$.*

Then the function

$$q_\mu(y_1, x_1, \beta)$$

$$= A_\mu(y_1 - g(x_1, \beta)) + g^{\xi\xi}(x_1, \beta)\,\Delta_\mu(y_1 - g(x_1, \beta)) - 2\left(g^\xi(x_1, \beta)\right)^2 \delta_\mu(y_1 - g(x_1, \beta))$$

*satisfies (**??**), where A_μ is given in (**??**), Δ_μ in (**??**) and $\delta\mu$ in (**??**).*□

Proof. Under **La1** we rewrite

$$Eq_\mu\left(y_1, \xi_1+\varepsilon_{21}, \beta\right)=\int \frac{1}{2} q_\mu\left(y_1, \xi_1+\varepsilon_{21}, \beta\right) \exp\left(-\left|\varepsilon_{21}\right|\right) d\varepsilon_{21.}$$

The approximation $A_\mu(y_1-g\left(\xi_1, \beta\right))$ we consider as function of ξ_1 :

$$A_\mu(y_1-g\left(\xi_1, \beta\right))=b_\mu\left(\xi_1\right).$$

We have to solve

$$\int \frac{1}{2} q_\mu\left(y_1, \xi_1+\varepsilon_{21}, \beta\right) \exp\left(-\left|\varepsilon_{21}\right|\right) d\varepsilon_{21}=b_\mu\left(\xi_1\right).$$

Under **G1, G2** the solution is given in Lemma **??** by

$$q_\mu\left(y_1, x_1, \beta\right)=b_\mu\left(x_1\right)-b_\mu^{\prime\prime}\left(x_1\right).$$

where

$$b_\mu^{\prime\prime}\left(x_1\right)=-A_\mu^{\prime}(y_1-g\left(x_1, \beta\right)) g^{\xi\xi}\left(x_1, \beta\right)+A_\mu^{\prime\prime}(y_1-g\left(x_1, \beta\right))\left(g^{\xi}\left(x_1, \beta\right)\right)^2,$$

with

$$A_\mu^{\prime}=\Delta_\mu,\ A_\mu^{\prime\prime}=2\delta_\mu.$$

■

6.2 The approximate symmetrical corrected L_1-norm estimator

We consider here a symmetrical L_1-norm approach with respect to (**??**) under Laplace distributed errors . In difference to the above section we require Laplace distribution for both errors $\varepsilon_{1i}, \varepsilon_{2i}$.

La2

$$\varepsilon_{ji}, j=1,2 \quad i.i. \text{ Laplace distributed with density } p,$$

$$p\left(u\right)=\frac{1}{2}\exp\left(-\left|u\right|\right).$$

We are now looking for an estimation function $q_\mu(x_i, y_i, \beta)$fulfilling

$$\sup_{\beta\in\Theta} \sup_{\xi_i^0\in[-a,a]} \left| E_{\theta^0} q_\mu(x_i, y_i, \beta) - \left| g\left(\xi_i^0, \beta^0\right) - g\left(\xi_i^0, \beta\right)\right|\right| \leq const\, \mu. \tag{93}$$

The way is similar to above. We start with a smooth approximation of $|g(\xi_i^0, \beta^0) - g(\xi_i^0, \beta)|$ by Lemma **??** and but then we apply the Lemma **??**, instead of Lemma **??**, because we required the deconvolution with respect to both error terms $\varepsilon_{1i}, \varepsilon_{2i}$.

Theorem 6.3 *Assume for all* $\beta \in \Theta$ *the regression function* $g(.,\beta) : \mathrm{R} \longrightarrow \mathrm{R}$, $g(.,\beta) \in C^2(\mathrm{R})$, $g(.,\beta)$ *satisfies*

$$\int \left| g^\xi(\xi,\beta)\right| \exp(-|\xi|)\, d\xi < \infty, \quad \int \left| g^{\xi\xi}(\xi,\beta)\right| \exp(-|\xi|)\, d\xi < \infty \tag{94}$$

and

$$\lim_{|\xi|\to\infty} g^{\xi\xi}(\xi,\beta) \exp(-|\xi|) = 0, \quad \lim_{|\xi|\to\infty} g^{\xi\xi}(\xi,\beta) \exp(-|\xi|) = 0 \tag{95}$$

uniformly with respect to $\beta \in \Theta$ *. Then under* **La2** *the approximate corrected criterion for a* L_1*-norm type estimator is*

$$\widetilde{C}_\mu(\beta) = \frac{1}{n} \sum_{i=1}^n q_\mu(x_i, y_i, \beta),$$

with

$$\begin{aligned} & q_\mu(x_i, y_i, \beta) \qquad (96) \\ &= A_\mu(y_i - g) + g^{\xi\xi} \Delta_\mu(y_i - g) - 2\left(\left(g^\xi\right)^2 + 1\right) \delta_\mu(y_i - g) \\ & -2g^{\xi\xi} \delta'_\mu(y_i - g) + 2\left(g^\xi\right)^2 \delta''_\mu(y_i - g), \end{aligned}$$

where g *and the derivatives* $g^\xi, g^{\xi\xi}$ *are taken at* (x_i, β) *and* A_μ *is given in (**??**),* Δ_μ *in (**??**) and* $\delta\mu$ *in (**??**).*□

Proof. We have to show (**??**). Because of Lemma **??**, it remains to search the solution of

$$E_{\theta^0} q_\mu(x_i, y_i, \beta) = A_\mu\left(g\left(\xi_i^0, \beta^0\right) - g\left(\xi_i^0, \beta\right)\right).$$

Under **La2** we rewrite

$$E_{\theta^0} q_\mu (x_i, y_i, \beta) = \int\int q_\mu \left(\xi_i^0 + \varepsilon_{2i}, \eta^0 + \varepsilon_{1i}, \beta\right) \frac{1}{4} \exp\left(-|\varepsilon_{1i}| - |\varepsilon_{2i}|\right) d\varepsilon_{1i} d\varepsilon_{2i},$$

with

$$\eta^0 = g\left(\xi_i^0, \beta^0\right).$$

Further we consider the left hand side as a function of expected values of x_i and y_i,such that

$$A_\mu \left(\eta^0 - g\left(\xi_i^0, \beta\right)\right) = B_\mu \left(\eta^0, \xi^0\right).$$

Under (**??**) and (**??**) the conditions **F1** and **F2** of Lemma **??** are satisfied. We apply this lemma to

$$\int\int q_\mu (\xi + \varepsilon_2, \eta + \varepsilon_1, \beta) \frac{1}{4} \exp\left(-|\varepsilon_1| - |\varepsilon_2|\right) d\varepsilon_1 d\varepsilon_2 = B_\mu (\eta, \xi)$$

and obtain

$$q_\mu (\xi, \eta, \beta) = B_\mu (\eta, \xi) - \frac{\partial^2}{\partial \eta^2} B_\mu (\eta, \xi) - \frac{\partial^2}{\partial \xi^2} B_\mu (\eta, \xi) + \frac{\partial^4}{\partial \eta^2 \xi^2} B_\mu (\eta, \xi).$$

Especially we have

$$\frac{\partial}{\partial \eta} B_\mu (\eta, \xi) = A_\mu^{'} (\eta - g(\xi_i, \beta))$$

$$\frac{\partial}{\partial \xi} B_\mu (\eta, \xi) = -A_\mu^{'} (\eta - g(\xi_i, \beta)) g^\xi (\xi_i, \beta)$$

and

$$\frac{\partial^2}{\partial \eta^2} B_\mu (\eta, \xi) = A_\mu^{''} (\eta - g(\xi_i, \beta))$$

$$\frac{\partial^2}{\partial \xi^2} B_\mu (\eta, \xi) = A_\mu^{'} (\eta - g(\xi_i, \beta)) g^{\xi\xi} (\xi_i, \beta) - A_\mu^{''} (\eta - g(\xi_i, \beta)) \left(g^\xi (\xi_i, \beta)\right)^2$$

and

$$\frac{\partial^4}{\partial \xi^2 \eta^2} B_\mu (\eta, \xi) = A_\mu^{(3)} (\eta - g(\xi_i, \beta)) g^{\xi\xi} (\xi_i, \beta) - A_\mu^{(4)} (\eta - g(\xi_i, \beta)) \left(g^\xi (\xi_i, \beta)\right)^2.$$

with

$$A_\mu' = \Delta_\mu, \quad A_\mu'' (t) = 2\delta_\mu, \quad A_\mu''' (t) = 2\delta_\mu^{'}, \quad A_\mu^{(4)} (t) = 2\delta_\mu''.$$

Summarizing we obtain the result of the Theorem. ■

7 The corrected L_2-estimator for the implicit model

7.1 On a corrected-minimum contrast estimator in explicit models with regression function of quotient form

The special case of an explicit model (**??**) with $q = 2$ and

$$y_i = \frac{g_1(\xi_i, \beta)}{g_2(\xi_i, \beta)} + \varepsilon_{1i} \tag{97}$$

$$x_i = \xi_i + \varepsilon_{2i},$$

with

$$\inf_{\xi_1} \inf_{\beta} g_2(\xi_i, \beta) \geq d > 0, \tag{98}$$

we can reformulate to an implicit model where

$$\eta_{1i} = \frac{g_1(\xi_i, \beta)}{g_2(\xi_i, \beta)} \quad \text{and} \quad \eta_{2i} = \xi_i.$$

Then under (**??**)

$$G(\eta_i, \beta) = \eta_{1i} - \frac{g_1(\eta_{2i}, \beta)}{g_2(\eta_{2i}, \beta)} = 0$$

is equivalent to

$$G_0(\eta_i, \beta) = (g_2(\eta_{2i}, \beta)\, \eta_{1i} - g_1(\eta_{2i}, \beta))^2 = 0.$$

A contrast function as proposed in (**??**) may be

$$C_n(\theta) = \frac{1}{n} \sum_{i=1}^{n} (g_2(\eta_{2i}, \beta)\, \eta_{1i} - g_1(\eta_{2i}, \beta))^2 . \tag{99}$$

The needed separation condition **Sep** for $C_n(\theta)$ is fulfilled under a Jennrich type contrast condition on the regression function g

$$\frac{1}{n} \sum_{i=1}^{n} \left(g\left(\xi_i^0, \beta^0\right) - g\left(\xi_i^0, \beta\right)\right)^2 \geq const \left\|\beta^0 - \beta\right\|^2,$$

because under (**??**) we have

$$C_n\left(\eta^0,\beta\right)-C_n\left(\eta^0,\beta^0\right)=\frac{1}{n}\sum_{i=1}^{n}\left(g_2\left(\eta_{2i}^0,\beta\right)\eta_{1i}^0-g_1\left(\eta_{2i}^0,\beta\right)\right)^2$$

$$\geq d\frac{1}{n}\sum_{i=1}^{n}\left(g\left(\xi_i^0,\beta^0\right)-g\left(\xi_i^0,\beta\right)\right)^2.$$

Models of the special form (**??**) occur in the biometry, for instance in the context of growth curves. We will give for one special case the corrected MCE related to the contrast (**??**).

7.1.1 A c-MCE for a growth curve

Consider the regression function

$$g\left(\xi_1,\beta\right)=\frac{\exp\left(\beta_{1,1}\xi_1\right)}{g_2\left(\xi_1,\beta\right)},$$

where g_2 is a polynomial of K'th order,

$$g_2\left(\xi_1,\beta\right)=\sum_{k=0}^{K}\beta_k\xi_1^k,$$

satisfying (**??**). Assume the existence of exponential moments $m\left(\beta\right)$ with respect to the error distribution **ExpMom.** Then the estimation criterion for the c-MCE related to the contrast in (**??**) is given by

$$q\left(y_1,x_1,\beta\right)=m\left(2\beta_{1,1}\right)^{-1}\exp\left(2\beta_{1,1}x_1\right)+\left(y_1^2-\sigma_1^2\right)h\left(x_1,\beta\right)-2y\exp\left(\beta_{1,1}x_1\right)f\left(x_1,\beta\right)$$

where the function f and h

$$f\left(x_1,\beta\right)=\sum_{k=0}^{K}\beta_k f_k^*\left(x_1\right)$$

and

$$h\left(x_1,\beta\right)=\sum_{k_1=0}^{K}\sum_{k_2=0}^{K}\beta_{k_1}\beta_{k_2}f_{k_1+k_1}^*\left(x_1\right)$$

relate to the polynomial $g_2(\xi_1, \beta)$. The auxiliary function $f_k^*(x_1)$ fulfill

$$Ef_k^*(x_1)\exp(\beta_{11}x_1) = \exp(\beta_{11}\xi_1)\,\xi_1^k$$

It corresponds to the auxiliary function introduced in (??), (??) and (??) with $\mu_k^* = E\left(\exp(\beta_{1,1}\varepsilon_{21})\,\varepsilon_{21}^k\right)$, that means

$$m(\beta_{11})\,f_0^*(z) = 1$$

and for all $r \geq 2$

$$m(\beta_{11})\,f_r^*(z) = z^r - \sum_{l=0}^{r-2} \begin{pmatrix} r \\ l \end{pmatrix} f_l^*(z)\,\mu_{r-l}^*. \tag{100}$$

7.2 The corrected L$_2-$estimator for the implicit polynomial model

Consider the implicit polynomial model of order k

$$G(\eta_1, \beta) = \sum_{j_1, j_2, \ldots j_q} a_{j_1 j_2 \ldots j_q} \eta_{11}^{j_1} \ldots \eta_{1q}^{j_q} \tag{101}$$

with

$$\beta \in \left\{ \left(a_{j_1 j_2 \ldots j_m}\right)_{j_1, j_2, \ldots j_m} : \sum_{l=1}^{m} j_l \leq k \right\} \in \mathrm{R}^p,$$

and

$$Z_i = \eta_i^0 + \varepsilon_i,$$

ε_i *i.i.d.* and the components of ε_1 are mutually independent.

The moments of the components of $\varepsilon_1 = (\varepsilon_{11}, \ldots, \varepsilon_{1q})$are known up to the order $2k$. Let us denote them by

$$m_{jr} = E(\varepsilon_{1j})^r .$$

We have to solve the L_2-deconvolution equation

$$E_{\eta_1, \beta} Q(\eta_1 + \varepsilon, \beta) = G(\eta_1, \beta)^2 . \tag{102}$$

That is possible by the same approach as for the explicit polynomial model. Remind the correction (**??**) for a real valued monomial $z^r$of order r given in (**??**), (**??**). Then we obtain for the solution of (**??**)

$$Q\left(Z_1,\beta\right)=\sum_{j_1,j_2,\ldots j_q} b_{j_1j_2..j_m} f_{1j_1}\left(Z_{11}\right)\ldots f_{qj_q}\left(Z_{1q}\right)$$

where

$$G^2\left(\eta_1,\beta\right)=\sum_{j_1,j_2,\ldots j_q} b_{j_1j_2..j_q}\eta_{11}^{j_1}\ldots\eta_{1q}^{j_q}$$

and

$$Ef_{jr}\left(z_j+\varepsilon_{ij}\right)=\left(z_j\right)^r, \quad j=1,\ldots,q, \quad r=0,\ldots,2k.$$

7.2.1 Application to chemistry

The implicit polynomial models play an important role in chemistry especially in the theory of polymerization. We shall argue it for the example of the copolymerization, compare for instance Keeler and Reilly (1992), [**?**]. Copolymerization is called the reaction of two monomers M_1, M_2 to a polymer consisting of both. The concentration of the monomer j at time point t is denoted by $M_j\left(t\right)$. The experiments are carried out for the determination of the copolymerization parameter,

$$r_1=\frac{k_{11}}{k_{12}}, \quad r_2=\frac{k_{22}}{k_{21}},$$

where k_{jj} is the reaction rate of the polymerization of the monomer M_j and k_{jk} is the reaction rate when the polymerization changes from the monomer M_j to M_k. The copolymerization equation is given by

$$\frac{dM_1\left(t\right)}{dM_2\left(t\right)}=\frac{M_1\left(t\right)}{M_2\left(t\right)}\frac{r_2\left(\frac{M_1(t)}{M_2(t)}\right)+1}{\left(\frac{M_1(t)}{M_2(t)}\right)+r_1}. \tag{103}$$

The experiment is repeated for n different mixtures i of the concentration of the monomers in the solvent solution and constant reaction time, where

$$\eta_{1i}=\frac{dM_1\left(t\right)}{dM_2\left(t\right)} \quad \text{and} \quad \eta_{2i}=\frac{M_1\left(t\right)}{M_2\left(t\right)}$$

are measured with errors. Therefore the copolymerization can also be described by a nonlinear functional relation model,

$$\eta_{1i}\left(\eta_{2i}+r_2\right)=\eta_{2i}\left(r_1\eta_{2i}+1\right)$$

which is of kind (**??**). We apply the solution (**??**) to that model and propose the corrected L_2−estimator $\widetilde{r}_1, \widetilde{r}_2$ for r_1, r_2, defined by

$$\left(\widetilde{r}_1,\widetilde{r}_2\right)\in\arg\min\frac{1}{n}\sum_{i=1}^{n}q\left(Z_i,r_1,r_2\right)$$

with

$$\begin{gathered}q\left(Z_i,r_1,r_2\right)=\\ \left(r_1\right)^2 f_{24}\left(Z_{2i}\right)-2r_1r_2\left(Z_{2i}^2-\sigma_2^2\right)Z_{1i}+\left(r_2\right)^2\left(Z_{1i}^2-\sigma_1^2\right)\\ +2r_1f_{23}\left(Z_{2i}\right)\left(1-Z_{1i}\right)+2r_2\left(Z_{1i}^2-Z_{1i}-\sigma_1^2\right)Z_{2i}.\end{gathered}$$

8 Appendix

Lemma 8.1 *Consider a function $g:\mathrm{R}\longrightarrow\mathrm{R}$, $g\in C^2\left(\mathrm{R}\right)$ and g is strictly increasing. The inverse function of g is denoted by g^{-1}. Assume* **G1, G2.**

Then for all $x,y\in\mathrm{R}$

$$\frac{1}{2}\int q\left(x+t\right)\exp\left(-\left|t\right|\right)dt=\left|y-g\left(x\right)\right|,$$

with

$$q\left(x\right)=\left|y-g\left(x\right)\right|-g''\left(x\right)sign\left(x-g^{-1}\left(y\right)\right)-2g'\left(x\right)\delta\left(x-g^{-1}\left(y\right)\right),$$

where δ denotes the δ-function, formally here defined by

$$\int g'\left(x+t\right)\delta\left(x+t-g^{-1}\left(y\right)\right)\exp\left(-\left|t\right|\right)dt=g'\left(g^{-1}\left(y\right)\right)\exp\left(-\left|x-g^{-1}\left(y\right)\right|\right).\tag{104}$$

□

Proof. The proof is done directly by computing the respected integral with partial integration. Denote

$$p(t) = \frac{1}{2} \exp(-|t|)$$

then for $t \neq 0$

$$p'(t) = -p(t)\, sgn(t) \tag{105}$$

and

$$\lim_{t \to -0} p'(t) = p(0) = -\lim_{t \to +0} p'(t) \tag{106}$$

We have

$$-\int g''(x+t)\, sign\left(x+t-g^{-1}(y)\right) \exp(-|t|)\, dt$$

$$= \int_{-\infty}^{g^{-1}(y)-x} \exp(-|t|)\, dg'(x+t) - \int_{g^{-1}(y)-x}^{\infty} \exp(-|t|)\, dg'(x+t)$$

$$= \exp(-|t|)\, g'(x+t)|_{-\infty}^{g^{-1}(y)-x} - \exp(-|t|)\, g'(x+t)|_{g^{-1}(y)-x}^{\infty} + J$$

$$= 2g'\left(g^{-1}(y)\right) \exp\left(-\left|x-g^{-1}(y)\right|\right) + J$$

with

$$J = \int_{-\infty}^{g^{-1}(y)-x} sign(t) \exp(-|t|)\, g'(x+t)\, d(t) - \int_{g^{-1}(y)-x}^{\infty} sign(t) \exp(-|t|)\, g'(x+t)\, dt$$

$$= -\int_{-\infty}^{0} \exp(-|t|)\, g'(x+t)\, d(t) - \int_{0}^{g^{-1}(y)-x} \exp(-|t|)\, g'(x+t)\, dt - \int_{g^{-1}(y)-x}^{\infty} \exp(-|t|)\, g'(x+t)\, dt$$

$$= -2g(x) + 2g\left(g^{-1}(y)\right) \exp\left(-\left|g^{-1}(y)-x\right|\right)$$

$$+ \int_{-\infty}^{g^{-1}(y)-x} \exp(-|t|)\, g(x+t)\, d(t) - \int_{g^{-1}(y)-x}^{\infty} \exp(-|t|)\, g(x+t)\, d(t).$$

Summarizing all we get the result. ■

Lemma 8.2 *Consider a function* $g : \mathrm{R} \longrightarrow \mathrm{R}$, $g \in C^2(\mathrm{R})$. *Assume* **G1,G2.**

Then for all $x \in \mathrm{R}$

$$\frac{1}{2} \int \left[g(x+t) - g''(x+t)\right] \exp(-|t|)\, dt = g(x).$$

□

Proof. The proof is done directly by computing the respected integrals with partial integration and using (**??**) and (**??**) and the assumptions. ■

We also need a multivariate version of Lemma **??**. Suppose

$$G : \mathrm{R}^q \to \mathrm{R}, \quad G \in C^{2q}\left(\mathrm{R}^q\right)$$

and

F1 For all partial derivatives $D^\alpha, \alpha \in \mathrm{Z}^q$,$\alpha_i \geq 0$,up to the order $\sum \alpha_i \leq 2q$

$$\int ... \int \left|D^\alpha G\left(t\right)\right| \exp\left(-\sum_{i=1}^{q} \left|t_i\right|\right) dt_1...dt_q < \infty,$$

and

F2 for all $i = 1, ..., q$

$$\lim_{\left|t_i\right| \to \infty} D^\alpha G\left(t\right) \exp\left(-\left|t_i\right|\right) = 0.$$

Define the operators

$$L_i = \frac{\partial^2}{\partial t_i^2}$$

and

$$L = L_1 L_2 ... L_q.$$

That is

$$L = I - \sum_{i=1}^{q} \frac{\partial^2}{\partial t_i^2} + \sum_{1 \leq i < j \leq q} \frac{\partial^2}{\partial t_i^2} \frac{\partial^2}{\partial t_j^2} - \sum_{1 \leq i < j < k \leq q} \frac{\partial^2}{\partial t_i^2} \frac{\partial^2}{\partial t_j^2} \frac{\partial^2}{\partial t_k^2} + ... + (-1)^q \frac{\partial^2}{\partial t_1^2} \frac{\partial^2}{\partial t_2^2} ... \frac{\partial^2}{\partial t_q^2},$$

where I denotes the identity operator.

Lemma 8.3 *Consider a function* $G : \mathrm{R}^q \to \mathrm{R}, \quad G \in C^{2q}\left(\mathrm{R}^q\right)$. *Assume* **F1,F2.**

Then for all $x \in \mathrm{R}^q$

$$\frac{1}{2^q} \int .. \int \left(LG\right)\left(x + t\right) \exp\left(-\sum_{i=1}^{q} \left|t_i\right|\right) dt_1...dt_q = G\left(x\right).$$

□

Proof. We apply Fubini and use Lemma **??** for each component t_i of t . ■