Ukrainian Mathematical Journal, Vol. 53, No. 9, 2001

ON AN ADAPTIVE ESTIMATOR OF THE LEAST CONTRAST IN A MODEL WITH NONLINEAR FUNCTIONAL RELATIONS

A. G. Kukush[1] **and S. Zwanzig**[2]

UDC 519.21

We consider an implicit nonlinear functional model with errors in variables. On the basis of the concept of deconvolution, we propose a new adaptive estimator of the least contrast of the regression parameter. We formulate sufficient conditions for the consistency of this estimator. We consider several examples within the framework of the L_1- and L_2-approaches.

The aim of the present paper is to construct consistent estimators for the regression parameters in an implicit functional error-in-variables model. In this model, an increase in the number of observations leads to an increase in the number of nuisance parameters for a fixed number of estimated parameters. It was established long ago that standard procedures of estimation are inapplicable to such models [1]. In linear models, the least-squares method gives consistent and efficient estimates for regression parameters. In the nonlinear case, such estimates are inconsistent.

1. Adaptive Estimators of the Least Contrast

Consider a *model with implicit functional relations*

$$G(\zeta_i, \beta_0) = 0, \qquad i = 1, \ldots, n, \tag{1}$$

where $\zeta_i \in D \subset \boldsymbol{R}^s$, $\zeta = \zeta(n) = (\zeta_1, \ldots, \zeta_n) \in D^n$ is a *deterministic* nuisance parameter, β_0 is the regression parameter belonging to a compact set $B \subset \boldsymbol{R}^p$, and $G\colon D \times B \to \boldsymbol{R}$ is a given function. One observes s-dimensional vectors

$$Z_i = \zeta_i + \varepsilon_i, \qquad i = 1, \ldots, n, \tag{2}$$

where ε_i are independent and identically distributed and, furthermore,

$$\mathrm{E}\varepsilon_1 = 0 \quad \text{and} \quad \mathrm{E}\|\varepsilon_1\|^2 < \infty.$$

A *model with explicit functional relations* is involved in (2) for $s \ge 2$, $\zeta_i = (\zeta_{1i}; \zeta_{2i})$, $\zeta_{1i} \in \boldsymbol{R}$, $\zeta_{2i} \in \boldsymbol{R}^{s-1}$, and $G(\zeta_i, \beta_0) = \zeta_{1i} - g(\zeta_{2i}, \beta_0)$. Then the observations $Z_i = (y_i; x_i)$ obey the classical explicit model

$$(y_i; x_i) = (g(\xi_i, \beta_0); \xi_i) + (\varepsilon_{1i}; \varepsilon_{2i}) \tag{3}$$

[1] Shevchenko Kiev University, Kiev.

[2] Hamburg University, Germany.

Translated from Ukrains'kyi Matematychnyi Zhurnal, Vol. 53, No. 9, pp. 1204–1209, September, 2001. Original article submitted May 25, 2000.

0041–5995/01/5309–1445 $25.00

with unknown deterministic points of a plan

$$\xi_i \in F \subset \boldsymbol{R}^{s-1}, \quad i = 1, \dots, n.$$

For the implicit model (1), we introduce the function of contrast dependent on the independent nuisance parameters.

The function $C_n\colon D^n \times B \to \boldsymbol{R}_+$, $n = 1, 2, \dots$, is called a *contrast* for β at a point β_0 if, for a certain strictly increasing function $\rho\colon \boldsymbol{R}_+ \to \boldsymbol{R}_+$, $\rho(0) = 0$, the following separability condition is satisfied:

(i) $\exists n_0 \geq 1 \quad \forall n \geq n_0 \quad \forall \beta \in B\colon C_n(\zeta, \beta) - C_n(\zeta_0, \beta_0) \;\geq\; \rho(\|\beta - \beta_0\|)$.

In what follows, C_n is determined by averaging, namely,

$$C_n(\zeta, \beta) = \frac{1}{n}\sum_{i=1}^{n} c(\zeta_i, \beta), \tag{4}$$

where $c\colon D \times B \to \boldsymbol{R}_+$. The main problem is to find an estimator for $C_n(\zeta, \beta)$ independent of ζ and based on observations (2). The first idea is to substitute the observations $Z_1, \dots, Z_n$ for the nuisance parameters $\zeta_1, \dots, \zeta_n$ in (4). As a result, one obtains a so-called *naive estimator of the least contrast* β_{naive}, and

$$\beta_{\text{naive}} \in \arg\min_{\beta \in B} C_n(Z_1, \dots, Z_n; \beta).$$

However, this estimate is inconsistent. The main idea of the present paper is to propose an adaptive procedure. We seek a Borel function $q\colon \boldsymbol{R}^s \times B \to \boldsymbol{R}$ lower-semicontinuous in $\beta \in B$ and such that

$$\forall \zeta_1 \in D \;\; \forall \beta \in B\colon \; \mathrm{E}_{\zeta_1} q(Z_1, \beta) = c(\zeta_1, \beta). \tag{5}$$

A random vector

$$\beta_{\text{ad}} \in \arg\min_{\beta \in B} \frac{1}{n}\sum_{i=1}^{n} q(Z_i, \beta) \tag{6}$$

is called an *adaptive estimator of the least contrast* for the parameter β_0.

A measurable solution of this optimization problem always exists [2].

Estimate (6) resembles the Stefanski estimate [3] for the explicit model (3). Stefanski also used convolution equations of the type (5) for the construction of a test of estimation. However, this test is based on a consistent procedure of estimation in the classical scheme of nonlinear regression, i.e., in the case where the points of a plan ξ_i are known. In [4], Nakamura applied a similar approach to a generalized nonlinear error-in-variables model. Buzas and Stefanski [5] extended the method of an adaptive counting function to a broad class of generalized linear models. In [6], Hanfelt and Kuung-Yee Liang proposed to use the function of conditional quasimaximum likelihood in a generalized linear model. In [7], Fazekas and Kukush considered an adaptive test connected with the naive least-squares estimator in model (3). Approach (5), (6) proposed in the present paper is more general and is based on the contrast $C_n(\zeta, \beta)$.

To determine q from Eq. (5), it is necessary to know the distribution law of the errors ε_i. Only for polynomial models it suffices to know several moments of distribution. Equation (5) can be explicitly solved only in certain special cases, namely, for an implicit polynomial model, explicit exponential model, and nonlinear smooth implicit model with errors distributed according to the Laplace law. In [3], a solution in the form of a series is presented for the case of normally distributed errors. One can also use the method of Fourier transformation [8].

Since Eq. (5) is not always solvable in an explicit form, we need an approximate solution. A family of functions $q_\mu : \boldsymbol{R}^s \times B \to \boldsymbol{R}$, $\mu > 0$, is called an *approximate solution* of Eq. (5) if

$$\forall \mu > 0 \colon \sup_{\beta \in B} \sup_{\zeta_1 \in D} \left| \mathrm{E}_{\zeta_1} q_\mu(Z_1, \beta) - c(\zeta_1, \beta) \right| \le \mu . \tag{7}$$

Let q_μ be a Borel function lower-semicontinuous in β. Then an *approximate adaptive estimator of the least contrast* β_μ is defined as a random vector that satisfies the relation

$$\beta_\mu \in \arg \min_{\beta \in B} \frac{1}{n} \sum_{i=1}^{n} q_\mu(Z_i, \beta), \qquad \mu > 0. \tag{8}$$

2. Consistency

We introduce the following moment conditions for $q(Z_i, \beta)$:

(ii) $\exists C > 0 \ \exists k \in [1, +\infty) \ \forall \beta \in B \colon \ \frac{1}{n} \sum_{i=1}^{n} \left(\mathrm{E} \left| q(Z_i, \beta) - \mathrm{E} q(Z_i, \beta) \right|^{2k} \right)^{1/k} \le C^{1/k}$;

(iii) there exists a random variable $\mathrm{M}_{(n)}$, constant C, and real number $k \ge 1$ such that, for all n and $\beta, \beta' \in B$, we have

$$\frac{1}{n} \sum_{i=1}^{n} | q(Z_i, \beta) - \mathrm{E} q(Z_i, \beta) - q(Z_i, \beta') + \mathrm{E} q(Z_i, \beta') |^2 \le \mathrm{M}_{(n)} \| \beta - \beta' \|^2$$

and $\mathrm{E}\mathrm{M}_{(n)}^k \le C$.

Theorem 1. *Suppose that condition (i) and conditions (ii) and (iii), where $k \ge 1$ is fixed, are satisfied. Then the following assertions are true:*

(a) if $k > p/2$, then

$$\forall \tau > 0 \ \ \forall n \ge 1 \colon \mathrm{P}_{\beta_0}(\| \beta_{\mathrm{ad}} - \beta_0 \| > \tau) \le \mathrm{const} \cdot \rho^{-2k}(\tau) \cdot n^{-k + p/2} ;$$

(b) if $k > p/2 + 1$, then $\beta_{\mathrm{ad}} \to \beta_0$ a.s.

For the functions q_μ from (7), we consider the following analogs of conditions (ii) and (iii):

(iv) $\exists c > 0\ \exists k \in [1, \infty)\ \exists \gamma_1 = \gamma_1(k) > 0\ \forall \mu \le \mu_0\ \forall \beta \in B$:

$$\frac{1}{n}\sum_{i=1}^{n}\left(\mathrm{E}\left|q_\mu(Z_i,\beta) - \mathrm{E}\,q_\mu(Z_i,\beta)\right|^{2k}\right)^{1/k} \le \frac{c^{1/k}}{\mu^{\gamma_1(k)/k}};$$

(v) there exists a random variable $\mathrm{M}_{(n,\mu)}$, constant c, real number $k \ge 1$, and $\gamma_2 = \gamma_2(k) > 0$ such that, for all n and $\mu \le \mu_0$ and every $\beta, \beta' \in B$, we have

$$\frac{1}{n}\sum_{i=1}^{n}\left|q_\mu(Z_i,\beta) - \mathrm{E}\,q_\mu(Z_i,\beta) - q_\mu(Z_i,\beta') + \mathrm{E}\,q_\mu(Z_i,\beta')\right|^2 \le \mathrm{M}_{(n,\mu)}\cdot\|\beta - \beta'\|^2$$

and $\mathrm{EM}^k_{(n,\mu)} \le c/\mu^{\gamma_2(k)}$.

Theorem 2. *Suppose that* $\mu = \mu(n) = an^{-r}$, $r > 0$, $a > 0$, *in relation (8). Also assume that condition (i) and conditions (iv) and (v), where* k *is fixed and such that* $k \ge 1$ *and* $k = p/2 + \delta$, $\delta > 0$, *are satisfied. We set*

$$r = \frac{2k(\delta - \chi)}{\gamma_1(k)(2k - p) + \gamma_2(k)\,p}.$$

Then the following assertions are true:

(a) if $\chi > 0$, *then*

$$\forall \tau > 0\ \forall n > \left(\frac{2a}{\rho(\tau)}\right)^{1/r}: \quad \mathrm{P}_{\beta_0}\left(\left\|\beta_{\mu(n)} - \beta_0\right\| > \tau\right) \le \mathrm{const}\cdot\left(\rho(\tau) - an^{-r}\right)^{-2k} n^{-\chi};$$

(b) if $\chi > 1$, *then* $\beta_{\mu(n)} \to \beta_0$ *a.s.*

3. Examples

3.1. Adaptive Least-Squares Estimator in an Explicit Model. For model (3), we set

$$c(\zeta_i,\beta) = G^2(\zeta_i,\beta) = (g(\xi_i,\beta_0) - g(\zeta_i,\beta))^2.$$

Assume that the separability condition (i), where

$$\rho(t) = at^2, \quad t \ge 0, \quad a > 0, \tag{9}$$

is satisfied. An adaptive estimator of the least contrast β_{lse} is a measurable solution of the optimization problem

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

where f satisfies the convolution equation

$$\forall \xi_1\ \forall \beta \in B:\quad \mathrm{E}_{\xi_1} f(x_1,\beta) = g(\xi_1,\beta) \tag{10}$$

and h satisfies a similar equation with the function $g^2(\xi_1,\beta)$ on the right-hand side. This estimator was studied in [7].

3.2. Approximate Adaptive Estimator of the Least Contrast in an Explicit Model of a Special Form. We give an example of a model where the approximate solution (7) naturally appears due to the fact that the regression function is not smooth. Assume that $s = 2$ in the explicit model (3) and, for fixed $A > 0$ and $d > 0$, we have

$$g(\xi_i,\beta_0) = \beta_0|\xi_i|,\qquad \beta_0 \in B = [-d, d],\quad |\xi_i| \le A,\quad i \ge 1. \tag{11}$$

Also assume that $\{\varepsilon_{1i}\}$ and $\{\varepsilon_{2i}\}$ are independent, $\mathrm{E}\varepsilon_{1i} = 0$, $\mathrm{E}\varepsilon_{1i}^2 = \sigma^2 < \infty$, and ε_{2i} have the canonical Laplace distribution with density $p(u) = e^{-|u|}/2$. We set

$$c(\zeta_i,\beta_0) = (g(\xi_i,\beta_0) - g(\xi_i,\beta))^2 + \sigma^2 = \xi_i^2(\beta_0-\beta)^2 + \sigma^2. \tag{12}$$

To construct an adaptive estimator of the least contrast, it is necessary to solve the convolution equations

$$\mathrm{E} f(\xi+\varepsilon_{2i},\beta) = \beta|\xi|,\qquad \beta,\xi \in \boldsymbol{R}, \tag{13}$$

$$\mathrm{E} h(\xi+\varepsilon_{2i},\beta) = \beta^2\xi^2,\qquad \beta,\xi \in \boldsymbol{R}. \tag{14}$$

Since $D\varepsilon_{2i} = 2$, the function $h(x,\beta) = \beta^2(x^2-2)$ satisfies Eq. (14). A formal solution of Eq. (13) is expressed in terms of the Dirac δ-function as follows:

$$f(x,\beta) = \beta(|x| - 2\delta(x));$$

indeed, integrating by parts, one can verify that

$$\int |\xi+t|p(t)dt - 2p(\xi) = |\xi|,\qquad \xi \in \boldsymbol{R}. \tag{16}$$

We approximate the δ-function by the δ-shaped family $\delta_\mu(t) = \mu^{-1}\omega(\mu^{-1}t)$, $t \in \boldsymbol{R}$, $\mu > 0$, where ω is a continuous probability density with the bounded support $\operatorname{supp}\omega = [-1, 1]$. Then the approximation test involves the following approximation of function (15):

$$f_\mu(x,\beta) = \beta\left(|x| - 2\delta_\mu(x)\right). \tag{17}$$

We set

$$q_\mu(y_i, x_i, \beta) = \left(y_i - f_\mu(x_i, \beta)\right)^2 + h(x_i, \beta) - f_\mu^2(x_i, \beta) = y_i^2 - 2y_i f_\mu(x_i, \beta) + \beta^2(x_i^2 - 2). \tag{18}$$

Let us show that q_μ satisfies (7) up to a constant factor.

Lemma 1. *For all* $\mu > 0$, *the following relation is true:*

$$\sup_{|\beta| \le d} \sup_{|\xi| \le A} \left| \mathrm{E} q_\mu\left(\beta_0 |\xi| + \varepsilon_{1i}, \ \xi + \varepsilon_{2i}\right) - \xi^2(\beta_0 - \beta) - \sigma^2 \right| \le \mathrm{const} \cdot \mu. \tag{19}$$

Proof. Let $\tilde{y}_i = \beta_0|\xi| + \varepsilon_{1i}$ and $\tilde{x}_i = \xi + \varepsilon_{2i}$. It follows from (18) that

$$\mathrm{E} q_\mu(\tilde{y}_i, \tilde{x}_i, \beta) = \xi^2(\beta_0 - \beta)^2 + \sigma^2 - 2\beta_0|\xi|\left(\mathrm{E} f_\mu(\tilde{x}_i, \beta) - \beta|\xi|\right). \tag{20}$$

Further, taking (16) into account, for $|\beta| \le d$ we get

$$\left| \mathrm{E} f_\mu(\tilde{x}_i, \beta) - \beta|\xi| \right| = 2d \left| \mathrm{E} \delta_\mu(\xi + \varepsilon_{2i}) - p(\xi) \right| \le d\mu. \tag{21}$$

The required inequality (19) follows from relations (20) and (21).

For the special model considered, we now define an approximate adaptive estimator of the least contrast β_μ as a measurable solution of the optimization problem (8), where $Z_i = (y_i, x_i)$ and the function q_μ is defined by equality (18). The direct application of Theorem 2 enables us to establish conditions for the consistency of the estimator β_μ.

Theorem 3. *Suppose that relations (11) are satisfied in model (3) and the quantities* ε_{2i} *have the canonical Laplace distribution. Also assume that the following conditions are satisfied:*

(A) $\liminf_{n \to \infty} n^{-1} \sum_{i=1}^{n} \xi_i^2 > 0$;

(B) for a certain fixed real $k \ge 1$, *we have* $\mathrm{E}|\varepsilon_{1i}|^{4k} < \infty$.

For $\chi \in (0, \ k - 1/2)$, *we set*

$$r = \frac{1}{2} - \frac{1 + 2\chi}{4k}.$$

Let $\mu = \mu(n) = an^{-r}$, $a > 0$, *in the definition of approximate adaptive estimator of the least contrast. Then the following assertions are true:*

(a) $\beta_{\mu(n)} \to \beta_0$ *in probability;*

(b) for $k > 3/2$ *and* $\chi \in (1, \ k - 1/2)$, *we have* $\beta_{\mu(n)} \to \beta_0$ *a.s.*

Proof. According to Lemma 1, functions (18) and (12) satisfy relation (7) up to a constant factor on the right-hand side. This enables us to apply Theorem 2 to the estimator $\beta_{\mu(n)}$. It is necessary to verify condition (i) and, for $p = 1$, conditions (iv) and (v).

According to condition (A), for $n \ge n_0$ and $|\beta| \le d$ we have

$$\frac{1}{n}\sum_{i=1}^{n}[c(\zeta_i,\beta)-c(\zeta_i,\beta_0)] = (\beta-\beta_0)^2\cdot\frac{1}{n}\sum_{i=1}^{n}\xi_i^2 \ge \text{const}\cdot(\beta-\beta_0)^2,$$

where const > 0, and the separability condition (i), where $\rho(t) = \text{const}\cdot t^2$, is satisfied.

Further, we verify condition (iv) for $\gamma_1(k) = 2k$. To this end, it suffices to show that, for $\mu \le \mu_0$, $|\beta| \le d$, and $|\xi| \le A$, we have

$$\mathrm{E}\left|q_\mu(\tilde{y}_i,\tilde{x}_i,\beta)\right|^{2k} \le \frac{\text{const}}{\mu^{2k}}, \tag{22}$$

where $\tilde{y}_i$ and $\tilde{x}_i$ are defined in the proof of Lemma 1. To verify inequality (22), we use the inequality $\delta_\mu(x) \le \text{const}\cdot\mu^{-1}$ and the fact that $\mathrm{E}\,(\tilde{y}_i^2)^{2k} \le \text{const}$ by virtue of condition (B) [see the definition of q_μ in (18)].

Finally, to verify condition (v) for $\gamma_2(k) = 2k$, it suffices to show that, for $\mu \le \mu_0$ and $|\xi| \le A$, we have

$$\mathrm{E}\sup_{|\beta|\le d}\left|\frac{\partial q_\mu(\tilde{y}_i,\tilde{x}_i,\beta)}{\partial\beta}\right|^{2k} \le \frac{\text{const}}{\mu^{2k}}. \tag{23}$$

Since

$$\frac{\partial q_\mu(\tilde{y}_i,\tilde{x}_i,\beta)}{\partial\beta} = -2\tilde{y}_i\big(|\tilde{x}_i|-2\delta_\mu(\tilde{x}_i)\big) + 2\beta(\tilde{x}_i^2-2),$$

we can prove inequality (23) by analogy with (22).

Thus, all conditions of Theorem 2 are satisfied. Then, in the considered case $\delta = k-1/2$, we have

$$r = \frac{k-1/2-\chi}{2k} = \frac{1}{2}-\frac{1+2\chi}{4k}$$

in Theorem 2.

The assertions of Theorem 3 follow from assertions (a) and (b) of Theorem 2. Theorem 3 is proved.

3.3. Adaptive Estimator of the Least Contrast in an Explicit Model with Regression Function Represented in the Form of a Ratio. Assume that $s = 2$ in model (3) and $g(\xi_i,\beta_0) = g_1(\xi_i,\beta_0)/g_2(\xi_i,\beta_0)$. Let us transform this model into the implicit model (1) with the function $G(\zeta_i,\beta) = g_2(\zeta_{2i},\beta)\cdot\zeta_{1i} - g_1(\zeta_{2i},\beta)$. We set $c(\zeta_i,\beta) = G^2(\zeta_i,\beta)$ and construct an adaptive estimator of the least contrast according to relation (8). We can solve Eq. (7) in the case where, e.g., g_2 is a polynomial and g_1 is either a polynomial or a function exponential with respect to ξ_i. If $g_2(\xi_1, \beta)$ is separated from zero, then condition (i) with function (9) is satisfied if

$$\forall n \geq 1: \frac{1}{n}\sum_{i=1}^{n}\left(g(\xi_i, \beta_0) - g(\xi_i, \beta)\right)^2 \geq L \cdot \|\beta - \beta_0\|^2, \quad L > 0.$$

3.4. Approximate Adaptive Estimator of the Least Contrast of the L_1 Type. Consider model (3) for $s = 2$. Within the framework of the L_1-approach, two cases of approximate adaptation are possible.

Let ε_{2i} have the canonical Laplace distribution. The first estimator is based on the consistent L_1-estimation in the classical nonlinear regression. The function of contrast has the form $c_0(\xi_i, \beta) = \mathrm{E}_{\beta_0}|y_i - g(\xi_i, \beta)|$ and the convolution equation is as follows:

$$\mathrm{E}_{\zeta_i} q(\zeta_{1i}, x_i, \beta) = |\zeta_{1i} - g(\xi_i, \beta)|.$$

The function c_0 generates a contrast C_n satisfying condition (i), where $\rho(t) = at^2$, $a > 0$ [9].

Now let $\{\varepsilon_{1i}, \varepsilon_{2i}\}$ be independent and identically distributed according to the canonical Laplace distribution. The second estimator is based on the function of contrast $c(\xi_i, \beta) = |g(\xi_i, \beta) - g(\xi_i, \beta_0)|$. The convolution equation has the form $\mathrm{E}_{\zeta_i, \beta_0} q(x_i, y_i, \beta) = c(\xi_i, \beta)$. In the separability condition (i), we set $\rho(t) = at$, $a > 0$.

For both estimators, the convolution equations are only approximately solvable in the sense of (7).

The proofs of Theorems 1 and 2 are based on the Whittle inequality [10] and can be found in [11], where Examples 3.1, 3.3, and 3.4 are investigated in more detail.

This work was supported by grant No. DFG 436UKR 17/10/96 and grant No. 436UKR 17/29/96.

REFERENCES

1. J. Neyman and E. L. Scott, "Consistent estimates based on partially consistent observations," *Econometrica,* **16**, No. 1, 1–32 (1948).
2. F. Leise and I. Vajda, "Necessary and sufficient conditions for consistency of generalized *M*-estimates," *Metrica,* **42**, 291–324 (1995).
3. L. A. Stefanski, "Correcting data for measurement error in generalized linear models," *Commun. Statist. Theory Meth.,* **18**, No. 5, 1715–1733 (1989).
4. T. Nakamura, "Corrected score function for error-in-variables models: Methodology and applications to generalized linear models," *Biometrica,* **77**, No. 1, 127–137 (1990).
5. J. S. Buzas and L. A. Stefanski, "A note on corrected-score mean," *Statist. Prob. Lett.,* **28**, 1–8 (1996).
6. J. J. Hanfelt and Kuung-Yee Liang, "Approximate likelihoods for generalized linear error-in-variables models," *J. R. Statist. Soc. B,* **59**, No. 3, 627–637 (1997).
7. I. Fazekas and A. Kukush, "Asymptotic properties of an estimator in nonlinear functional errors-in-variables models," *Computers Math. Appl.,* **34**, No. 10, 23–39 (1997).
8. L. A. Stefanski and R. J. Carroll, "Deconvoluting kernel density estimators," *Statistics,* **21**, 169–184 (1990).
9. S. Zwanzig, "On L_1-norm estimators in nonlinear regression and in nonlinear error-in-variables models," *Lect. Notes. Monograph Ser.,* **31**, 101–118 (1997).
10. P. Whittle, "Bounds for the moments of linear and quadratic forms in independent variables," *Theory Probab. Appl.,* **5**, 302–305 (1960).
11. A. Kukush and S. Zwanzig, *On a Corrected Contrast Estimator in the Implicit Nonlinear Functional Relation Model,* Technical Report 97–12, Inst. Math. Stochast. Univ. Hamburg, Hamburg (1997).