On Consistent Estimators in Linear and Bilinear Multivariate Errors-in-Variables Models [1]

Alexander Kukush, Ivan Markovsky, and Sabine Van Huffel[2]

August 2001

Published in
Total least squares and errors-in-variables modeling: Analysis, algorithms and applications
S. Van Huffel and Ph. Lemmerling editors
pp. 155–164, Kluver Academic Publishers

[1]This report is available by anonymous ftp from *ftp.esat.kuleuven.ac.be* in the directory *pub/sista/markovsky/reports/01-82.ps.gz*

[2]K.U.Leuven, Dept. of Electrical Engineering (ESAT),
Research group SCD (SISTA),
Kasteelpark Arenberg 10,
3001 Leuven-Heverlee, Belgium,
Tel. 32/16/32 17 09, Fax 32/16/32 19 70,
WWW: *http://www.esat.kuleuven.ac.be/sista.*
E-mail: *ivan.markovsky@esat.kuleuven.ac.be.*

This paper presents research results of the Belgian Programme on Interuniversity Poles of Attraction (IUAP V-10-29), initiated by the Belgian State, Prime Minister's Office - Federal Office for Scientific, Technical and Cultural Affairs, of the European Community TMR Programme, Networks, project CHRX-0160, of the Brite-Euram Programme, Thematic Network BRRT-CT97-5040 'Niconet', of the Concerted Research Action (GOA) projects of the Flemish Government MEFISTO-666 (Mathematical Engineering for Information and Communication Systems Technology), of the IDO/99/03 project (K.U.Leuven) "Predictive computer models for medical classification problems using patient data and expert knowledge", of the FWO projects G.0078.01, G.0200.00, and G0.0270.02.

Abstract

We consider three multivariate regression models related to the TLS problem. The errors are allowed to have unequal variances.

For the model $AX = B$, the elementwise-weighted TLS estimator is considered. The matrix $[A\ B]$ is observed with errors and has independent rows, but the errors in a row are correlated. In addition, the corresponding error covariance matrices may differ from row to row and some of the columns are allowed to be error-free. We give mild conditions for weak consistency of the estimator, when the number of rows in A increases. We derive the objective function for the estimator and propose an iterative procedure to compute the solution.

In a bilinear model $AXB = C$, where the data A, B, C are perturbed by errors, an adjusted least squares estimator is considered, which is consistent, i.e. converges to X, as the number m of rows in A and the number q of columns in B increase.

A similar approach is applied in a related model, arising in motion analysis. The model is $v^T F u = 0$, where the vectors u and v are homogeneous coordinates of the projections of the same rigid object point in two images, and F is a rank deficient matrix. Each pair (u, v) is observed with measurement errors. We construct a consistent estimator of F in three steps: a) estimate the measurement error variance, b) construct a preliminary matrix estimate, and c) project that estimate on the subspace of singular matrices.

A simulation study illustrates the theoretical results.

ON CONSISTENT ESTIMATORS IN LINEAR AND BILINEAR MULTIVARIATE ERRORS-IN-VARIABLES MODELS

Alexander Kukush, Ivan Markovsky and Sabine Van Huffel
ESAT-SISTA, K.U.Leuven
Kasteelpark 10, B-3001 Leuven-Heverlee, Belgium
{Alexander.Kukush, Ivan.Markovsky, Sabine.VanHuffel}@esat.kuleuven.ac.be

Abstract We consider three multivariate regression models related to the TLS problem. The errors are allowed to have unequal variances.

For the model $AX = B$, the elementwise-weighted TLS estimator is considered. The matrix $[A\ B]$ is observed with errors and has independent rows, but the errors in a row are correlated. In addition, the corresponding error covariance matrices may differ from row to row and some of the columns are allowed to be error-free. We give mild conditions for weak consistency of the estimator, when the number of rows in A increases. We derive the objective function for the estimator and propose an iterative procedure to compute the solution.

In a bilinear model $AXB = C$, where the data A, B, C are perturbed by errors, an adjusted least squares estimator is considered, which is consistent, i.e. converges to X, as the number m of rows in A and the number q of columns in B increase.

A similar approach is applied in a related model, arising in motion analysis. The model is $v^T F u = 0$, where the vectors u and v are homogeneous coordinates of the projections of the same rigid object point in two images, and F is a rank deficient matrix. Each pair (u, v) is observed with measurement errors. We construct a consistent estimator of F in three steps: a) estimate the measurement error variance, b) construct a preliminary matrix estimate, and c) project that estimate on the subspace of singular matrices.

A simulation study illustrates the theoretical results.

Keywords: linear and bilinear errors-in-variables models, elementwise-weighted total least squares, adjusted least squares, essential/fundamental matrix.

1. Introduction

We study three multivariate functional relation models. The errors are allowed to have unequal variances, which makes the asymptotic analysis of the estimators more complicated. Our goal is to derive consistent estimators, which admit a numerically reliable computation. *Consistent estimator* is an estimator that converges to the true value of the regression parameters, as the sample size tends to infinity.

2. Model $AX = B$

Let $A \in \mathbb{R}^{m\times n}$ and $B \in \mathbb{R}^{m\times d}$ be *observations*, and $X \in \mathbb{R}^{n\times d}$ be a *parameter* of interest. We suppose that

$$A = A_0 + \tilde{A}, \quad B = B_0 + \tilde{B}, \tag{1}$$

and that there exists $X_0 \in \mathbb{R}^{n\times d}$ such that

$$A_0 X_0 = B_0. \tag{2}$$

Here A_0 and B_0 are non-stochastic true values, and $\tilde{A}$ and $\tilde{B}$ are *errors*, and the matrix X_0 yields the true values of the parameters. The measurements are represented by the rows of A and B. We fix the dimension of X_0, and let the number of measurements m increase. For $i \geq 1$ let

$$\delta_{ij} := \tilde{a}_{ij}, \text{ if } j = 1, \ldots, n, \quad \text{and} \quad \delta_{i,n+k} := \tilde{b}_{ik}, \text{ if } k = 1, \ldots, d.$$

If all errors δ_{ij} are non-correlated and equally sized, the TLS technique is appropriate to solve an overdetermined set $AX \approx B$, see, e.g. [18]. In [17] the generalized TLS formulation is considered, where some of the columns of A are assumed to be known exactly, and the errors in the remaining data are allowed to be correlated and not equally sized, but they should be row-wise i.i.d. for consistency. In the present paper we confirm the claim from [17, p.313] about the consistency of the estimator in that model and further relax the conditions.

2.1 The elementwise-weighted TLS problem

Our main assumptions are as follows.

(i) The rows of $(\delta_{ij},\ i \geq 1,\ j = 1, \ldots, n + d)$ are independent with zero mean and finite second moments.

Let $\sigma_{ij}^2 := \mathbf{E}\delta_{ij}^2$. We allow some of σ_{ij}^2 to be equal to zero.

(ii) For a fixed $J \subset \{1, \ldots, n + d\}$, every $j \notin J$ and $i \geq 1$ satisfy $\sigma_{ij}^2 = 0$. For $i \geq 1$ denote $\Sigma_i := \mathrm{cov}(\delta_{ij}, j \in J)$.

(iii) There exists $\kappa > 0$ such that for every $i \geq 1$, $\lambda_{\min}(\Sigma_i) \geq \kappa^2$.

For instance we have ordinary regression if $J = \{n+1, \ldots, n+d\}$, i.e. if only B is observed with errors and A is error-free. We allow the Data Least Squares Problem, see [5], where $J \subset \{1, \ldots, n\}$, i.e. A is observed with errors and B is error-free.

Introduce two matrices by

$$X_{\text{ext}}^0 := \begin{bmatrix} X_0 \\ -I_d \end{bmatrix} = (x_{jk}^0,\ j = 1, \ldots, n+d,\ k = 1, \ldots, d)$$

with identity matrix $I_d \in \mathbb{R}^{d \times d}$, and $X_J^0 := (x_{jk}^0,\ j \in J,\ k = 1, \ldots, d)$.

We use similar notation X_{ext} and X_J for every $X \in \mathbb{R}^{n \times d}$.

(iv) $\operatorname{rank}\left(X_J^0\right) = d$.

Now, let $[\Delta A\ \Delta B] = (\Delta a_{ij},\ i = 1, \ldots, m,\ j = 1, \ldots, n+d)$, and column vectors $\Delta a_i^J := (\Delta a_{ij},\ j \in J) \in \mathbb{R}^{|J| \times 1}$.

The elementwise-weighted total least squares (EW-TLS) problem consists in finding the value $\hat{X}$ of unknown matrix X and the values $\Delta\hat{A}, \Delta\hat{B}$ of unknown error matrices $\tilde{A}, \tilde{B}$ minimizing the weighted sum of squared corrections

$$\min_{X, \Delta A, \Delta B} \sum_{i=1}^{m} ||\Sigma_i^{-\frac{1}{2}} \Delta a_i^J||^2 \tag{3}$$

subject to the constraints

$$\begin{array}{c} \operatorname{rank}(X_J) = d, \quad (A - \Delta A)X = B - \Delta B, \\ \text{and} \quad \Delta a_{ij} = 0, \text{ for } 1 = i, \ldots, m,\ j \notin J. \end{array} \tag{4}$$

The EW-TLS estimator $\hat{X}$ of X_0 in the model (1),(2) is a measurable value of X, which solves the optimization problem (3),(4).

Now, we derive an objective function for $\hat{X}$. Denote

$$(AX - B)^T := [r_1(X), \ldots, r_m(X)],$$

and

$$S_i := \operatorname{cov}(\delta_{ij},\ j = 1, \ldots, n+d), \quad i = 1, \ldots, m.$$

Theorem 1. *Let the conditions (i) to (iv) hold. The EW-TLS estimator $\hat{X}$ of X_0 in the model (1),(2) exists iff there exists an unconstrained minimum of the function*

$$Q(X) := \sum_{i=1}^{m} r_i(X)^T \left(X_J^T \Sigma_i X_J\right)^{-1} r_i(X), \quad \operatorname{rank}(X_J) = d. \tag{5}$$

Then $\hat{X}$ is a minimum point of $Q(X)$, and the values $\Delta\hat{A}$, $\Delta\hat{B}$ in the corresponding problem (3),(4) are given by

$$[\Delta\hat{A}\ \Delta\hat{B}] = \begin{bmatrix} r_1^T(\hat{X})\big(\hat{X}_J^T\Sigma_i\hat{X}_J\big)^{-1}\hat{X}_{\text{ext}}^T S_1 \\ \vdots \\ r_m^T(\hat{X})\big(\hat{X}_J^T\Sigma_i\hat{X}_J\big)^{-1}\hat{X}_{\text{ext}}^T S_m \end{bmatrix}, \text{ where } \hat{X}_{\text{ext}} := \begin{bmatrix} \hat{X} \\ -I_d \end{bmatrix}.$$

In the particular case $J = \{1, \ldots, n+d\}$ and $\Sigma_i = I_{n+d}$, $i = 1, \ldots, m$, the EW-TLS estimator coincides with the TLS-estimator. Then the function (5) equals

$$Q(X) = \operatorname{tr}\Big((AX - B)\big(X_{\text{ext}}^T X_{\text{ext}}\big)^{-1}(AX - B)^T\Big), \tag{6}$$

and the TLS estimator is known to minimize the function (6), see [16]. Thus (5) can be considered as a generalization of the objective function (6) for the case of unequally sized errors.

2.2 Consistency of the EW-TLS estimator

The next conditions guarantee existence of the estimator and its consistency. We denote $q := d(|J| - (d+1)/2)$, and let $\underline{\lambda}$ and $\bar{\lambda}$ be the minimal and maximal eigenvalues of $A_0^T A_0$, respectively.

(v) For fixed $r \geq 2$, $r > \frac{1}{2}\big(q + \max\{(n-1), q\}\big)$,

$$\sup_{i\geq 1,\ j\in J} \mathbf{E}|\delta_{ij}|^{2r} < \infty. \tag{7}$$

(vi) $\underline{\lambda}/\sqrt{m} \to \infty$, as $m \to \infty$.

(vii) $\underline{\lambda}^2/\bar{\lambda} \to \infty$, as $m \to \infty$.

The next two conditions are stronger versions of (vi), (vii).

(viii) For some $m \geq 1$, $\sum_{m=m_0}^{\infty}\big(\sqrt{m}/\underline{\lambda}\big)^r < \infty$.

(ix) For some $m \geq 1$, $\sum_{m=m_0}^{\infty}\big(\bar{\lambda}/\underline{\lambda}^2\big)^r < \infty$.

Theorem 2. *Suppose that the conditions (i) to (vii) hold. In the model (1),(2) we have:*

a) the EW-TLS estimator $\hat{X}$ exists with probability tending to one, as $m \to \infty$,

b) $\hat{X} \xrightarrow{\text{P}} X_0$, as $m \to \infty$,

c) instead of (vi),(vii), assume (viii),(ix), where r comes from (v). Then $\hat{X} \to X_0$ a.s., as $m \to \infty$.

Theorem 2b) extends the weak consistency result of Gallo [7], who considered a univariate homoscedastic model, i.e. B is a vector and the rows in $[\tilde{A}\ \tilde{B}]$ are i.i.d. The eigenvalue asymptotics (vi),(vii) are exactly his conditions, and he demanded (v), with $2r = 4$. In a homoscedastic case, we can weaken (v) and set $2r = 4$ in (7), thus obtaining a direct generalization of Gallo's result to the multivariate homoscedastic model.

In a homoscedastic multivariate model (1),(2) Gleser [8] required that $A_0^T A_0/m \to A_\infty > 0$, which implies that both $\underline{\lambda}$ and $\bar{\lambda}$ have the order m. He showed the strong consistency assuming that the fourth moments of the errors are limited. Theorem 2c) covers Gleser's result except for the condition that the moments of order $4+\delta$ be bounded (but we do not require that the errors are row-wise identically distributed).

In a homoscedastic model the asymptotic properties of $\hat{X}$ are derived from the explicit formula of the estimator. It is impossible to obtain such a formula in our case where all errors may be differently sized. Instead we show that the estimator is a solution of the unconstrained problem $\min_X Q(X)$, and analyze the asymptotic behavior of the objective function (5).

2.3 Computation of the EW-TLS estimator

We describe a numerical scheme to compute the estimator. Theorem 2a) proves that the minimum point $\hat{X}$ exists with probability tending to one, and that it is a root of the equation $Q'(X) := \mathcal{F}(X, Z) = 0$.

A direct computation gives

$$\mathcal{F}(X,Z) = \sum_{i=1}^{m} \Big(a_i r_i^T \big(Z^T S_i Z\big)^{-1} - P S_i^{\frac{1}{2}} Z \big(Z^T S_i Z\big)^{-1} r_i r_i^T \big(Z^T S_i Z\big)^{-1} \Big).$$

Here $r_i = r_i(X)$ and $Z = Z(X) := X_{\text{ext}}$ are functions of X, $P := [I_n\ 0_{n\times d}]$ and $A^T = [a_1 \cdots a_m]$. The proposed iterative algorithm for solving $\mathcal{F}(X, Z) = 0$ is as follows.

1) Compute an approximation $X_{(0)}$, using, e.g. weighted least squares.

2) Given an approximation $X_{(j)}$, we define $Z_{(j)} := \left[\begin{smallmatrix} X_{(j)} \\ -I_d \end{smallmatrix}\right]$ and $X_{(j+1)}$ is found from the optimization problem

$$\min_X ||\mathcal{F}(X, Z_{(j)})||_F^2,$$

which can be solved by Newton-type algorithms.

We mention the paper [15]. The authors consider a univariate problem (1),(2), where X and B are vectors, derive an objective function like (5) and propose an algorithm, which is related to the one presented above. They also solve the equation $\mathcal{F}(X, Z) = 0$, using a linear approximation with respect to X, while our approximation $\mathcal{F}(X, Z_{(j)})$ is quadratic with respect to X.

2.4 Simulation examples

In order to demonstrate the asymptotic properties of the EW-TLS estimator, we set up a simulation experiment. The true data $[A_0\ b_0]$ is a random matrix with independent elements, uniformly distributed in the interval $[0, 1]$. The true value X_0 of the parameters is known. The errors are selected in two different ways:

1) $\mathrm{cov}(\tilde{a}_i) = \sigma_i^2 I$ and $\mathrm{var}(\tilde{b}_i) = \sigma_i^2$, where σ_i, $i = 1, \ldots, m$ are uniformly distributed random variables in the interval $[0.02, 0.1]$, and

2) $\mathrm{cov}(\mathrm{vec}[\tilde{A}\ \tilde{B}]) = \sigma^2 I$, where $\sigma = 0.1$.

The plots on figures 1a and 1b show the relative error of estimation (averaged over 1000 noise realizations) as a function of the sample size m for the ordinary least squares (OLS) estimator, the weighted least-squares (WLS) estimator, the TLS estimator and the EW-TLS estimator. The OLS estimator uses a wrong assumption that A is noise free and the errors in b are equally distributed. The WLS uses the covariance information for $\tilde{b}$ but also assumes that A is noise free. The TLS estimator allows errors in both A and b but assumes that they are equally distributed. The EW-TLS estimator uses the full covariance information.

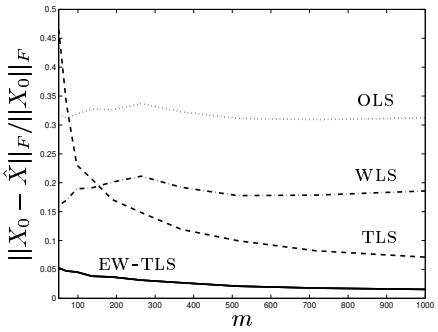

Figure 1a. Differently sized errors

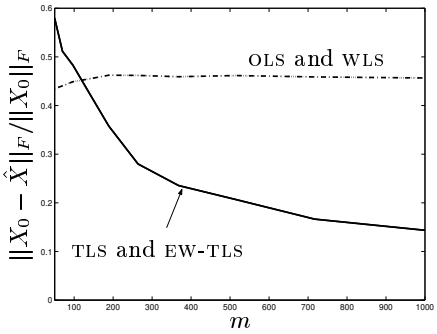

Figure 1b. Equally sized errors

3. Model $AXB = C$

Let $A \in \mathbb{R}^{m\times n}$, $B \in \mathbb{R}^{p\times q}$, $C \in \mathbb{R}^{m\times q}$ be *observations* and $X \in \mathbb{R}^{n\times p}$ be a *parameter* of interest. We suppose that

$$A = A_0 + \tilde{A}, \quad B = B_0 + \tilde{B}, \quad C = C_0 + \tilde{C}, \tag{8}$$

and that there exists $X_0 \in \mathbb{R}^{n\times p}$ such that

$$A_0 X_0 B_0 = C_0. \tag{9}$$

Here A_0, B_0, and C_0 are non-stochastic *true values* and $\tilde{A}$, $\tilde{B}$, and $\tilde{C}$ are *errors*. The matrix X_0 is the nominal or true value of the parameter. The measurements are represented by the rows of A, the columns of B and the elements of C. We fix the dimension of X_0, and let the numbers m and q increase.

We suppose that the errors $\{\tilde{a}_{ij},\ i \geq 1,\ 1 \leq j \leq n\}$, $\{\tilde{b}_{kl},\ 1 \leq k \leq p,\ l \geq 1\}$ and $\{\tilde{c}_{il},\ i \geq 1,\ l \geq 1\}$ form three independent arrays of random variables, which are centered and possess finite second order moments.

It is known that orthogonal regression is inconsistent for nonlinear measurement error models, [6, 13]. The orthogonal regression estimator in the bilinear model (8),(9) is a (weighted) TLS-estimator, see [18], therefore due to nonlinearity of the model, the TLS estimator is inconsistent in this case.

Let

$$V_{\tilde{A}} := \mathbf{E}\tilde{A}^T\tilde{A},\ V_{\tilde{B}} := \mathbf{E}\tilde{B}\tilde{B}^T.$$

The adjusted least squares (ALS) estimator is given by

$$\hat{X} := (A^TA - V_{\tilde{A}})^\dagger (A^TCB^T)(BB^T - V_{\tilde{B}})^\dagger. \tag{10}$$

It is similar to the ALS estimator proposed for a polynomial errors-in-variables model in [2]. The estimator (10) is consistent and asymptotically normal. Using the idea of [3] it is possible to construct a small sample modification of the estimator, which is more stable for small m and q and is asymptotically equivalent to the ALS estimator. Simulation results show that it reduces the error of estimation for small sample sizes [9].

The optimality of the estimator (10) is a topic of further research. For the model (1),(2) the ALS estimator is asymptotically efficient provided $V_{\tilde{A}}$ is known exactly and $\mathbf{E}\tilde{b}_{kl}^2$ are known up to a constant factor, see [11]. We conjecture that the estimator (10) is asymptotically efficient if $V_{\tilde{A}}$ and $V_{\tilde{B}}$ are known exactly and $\mathbf{E}\tilde{c}_{il}^2$ are known up to a constant factor.

4. Model $v^T F u = 0$

The next functional relation deals with the exploitation of the epipolar constraint information for the construction of the fundamental matrix for uncalibrated images, which once decomposed, solves the structure from motion problem, see [14, 4].

Let $u_i = [u_i(1)\ u_i(2)\ 1]^T \in \mathbb{R}^{3\times 1}$ and $v_i = [v_i(1)\ v_i(2)\ 1]^T \in \mathbb{R}^{3\times 1}$, $i = 1, \ldots, m$, represent the homogeneous pixel coordinates in the first and second image respectively. The model is

$$v_i^T F u_i = 0, \quad i = 1, \ldots, m,$$

where $F \in \mathbb{R}^{3\times 3}$ is the *fundamental matrix* with $\operatorname{rank}(F) = 2$, and F is a parameter of interest. We suppose that the vectors u_i, v_i are observed with errors,

$$u_i = u_{0,i} + \tilde{u}_i, \quad \text{and} \quad v_i = v_{0,i} + \tilde{v}_i, \quad i = 1, \ldots, m,$$

and that there exists $F_0 \in \mathbb{R}^{3\times 3}$, with $\operatorname{rank}(F_0) = 2$ and $||F_0||_F = 1$ such that

$$v_{0,i}^T F_0 u_{0,i} = 0, \quad i = 1, \ldots, m.$$

The matrix $F_0 \in \mathbb{R}^{3\times 3}$ is the true fundamental matrix F. The vectors $u_{0,i}$ and $v_{0,i}$ are the true values of the measurements u_i and v_i, respectively, and $\tilde{u}_i$, $\tilde{v}_i$ represent the measurement errors.

We derive a strongly consistent estimator for F_0 by taking more realistic assumptions than those in [14]. We suppose that the error terms $\{\tilde{u}_i,\ \tilde{v}_i,\ i \geq 1\}$ are independent and centered, and

$$\operatorname{cov}(\tilde{u}_i) = \operatorname{cov}(\tilde{v}_i) = \sigma_0^2 T, \quad i \geq 1$$

with $T := \operatorname{diag}(1, 1, 0)$ and fixed but unknown $\sigma_0^2 > 0$, $\sigma_0^2 \in (0, d^2]$.

The estimator of F_0 proposed below is the corrected minimum contrast estimator, considered in [12] in a more general context. It is related to the method of corrected score functions [1].

Let

$$S_m(\sigma^2) := \sum_{i=1}^{m} \left(u_i u_i^T - \sigma^2 T\right) \otimes \left(v_i v_i^T - \sigma^2 T\right),$$

$$Q_m(\sigma^2) := |\lambda_{\min}\left(S_m(\sigma^2)\right)|, \quad 0 \leq \sigma^2 \leq d^2.$$

A consistent estimator for σ_0^2 is given by

$$\hat{\sigma}^2 \in \arg \min_{0 \leq \sigma^2 \leq d^2} Q_m(\sigma^2).$$

Let $\hat{f}_1$ be a normalized eigenvector belonging to the minimal eigenvalue of $S_m(\hat{\sigma}^2)$. A preliminary estimator $\hat{F}_1$ is a matrix with $\mathrm{vec}(\hat{F}_1) = \hat{f}_1$. The final estimator $\hat{F}$ is obtained from $\hat{F}_1$ by projecting it to the set of rank deficient matrices, e.g. by computing the SVD of $\hat{F}_1$. Under appropriate conditions, $\mathrm{dist}(\hat{F}, \{\pm F_0\}) \to 0$ a.s., as $m \to \infty$. For details see [10].

We show the performance of the derived estimators $\hat{F}$ and $\hat{\sigma}^2$ by generating artificial data for which the true value F_0 is known. The fundamental matrix F_0 is a randomly chosen rank two matrix with unit Frobenius norm. The true coordinates $u_{0,i}$ and $v_{0,i}$ have third components equal to one, and the first two components are randomly chosen vectors in $\mathbb{R}^{2\times 1}$ with unit norm and random direction. The perturbations $\tilde{u}_i$ and $\tilde{v}_i$ are selected according to the assumptions stated in the paper. The estimation is repeated $N = 1000$ times with the same true data and different noise realizations. The presented results are the averaged ones.

We compare three estimators: the TLS estimator, the estimator $\hat{F}$ using the true noise variance σ_0^2, and the estimator $\hat{F}$ using the estimated noise variance $\hat{\sigma}^2$. Figure 2a shows the relative error of estimation as a function of the sample size m, and Figure 2b shows the convergence of the estimate $\hat{\sigma}^2$.

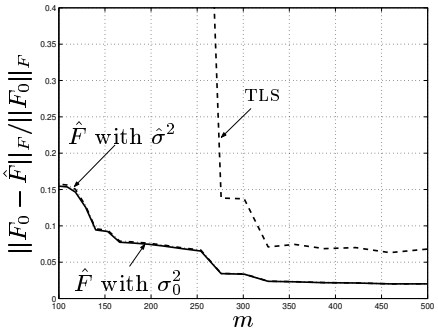

Figure 2a. Estimation of F_0

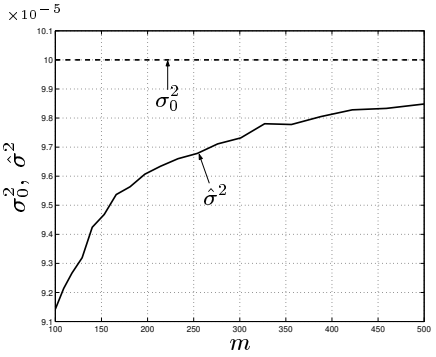

Figure 2b. Estimation of σ_0^2

Acknowledgments

A. Kukush is supported by a postdoctoral research fellowship of the Belgian office for Scientific, Technical and Cultural Affairs, promoting Scientific and Technical Collaboration with Central and Eastern Europe.

This work is supported by IUAP P4-02 and P4-24, the Concerted Research Action projects MEFISTO-666 and FWO project G.0200.00.

References

[1] R. J. Carroll, D. Ruppert, and L. A. Stefanski. *Measurement Error in Nonlinear Models*. Number 63 in Monographs on Statistics and Applied Probability. Chapman & Hall/CRC, 1995.

[2] C. Cheng and H. Schneeweiss. Polynomial regression with errors in the variables. *J. R. Statistical Society B*, (60):189–199, 1998.

[3] C. Cheng, H. Schneeweiss, and Markus Thamerus. A small sample estimator for a polynomial regression with errors in the variables. *J. R. Statistical Society*, (62):699–709, 2000.

[4] G. Cirrincione. TLS and constrained TLS neural networks for computer vision. In *this volume*.

[5] E. Dowling and R. DeGroat. The equivalence of the TLS and minimum norm methods. *IEEE Trans. on Signal Proc.*, 39:1891–1892, 1991.

[6] W. A. Fuller. *Measurement Error Models*. New York: Wiley, 1987.

[7] P. P. Gallo. Consistency of regression estimates when some variables are subject to error. *Comm. Stat. B-Theory Methods*, (11):973–893, 1982.

[8] L. J. Gleser. Estimation in a multivariate "errors in variables" regression model: Large sample results. *Ann. Stat.*, (9):24–44, 1981.

[9] A. Kukush, I. Markovsky, and S. Van Huffel. Consistent adjusted least squares estimator for errors-in-variables model $AXB = C$. Technical Report 01-72, K.U.Leuven, Dept. EE, K.U.Leuven, 2001. Submitted to Metrika.

[10] A. Kukush, I. Markovsky, and S. Van Huffel. Consistent fundamental matrix estimation in a quadratic measurement error model arising in motion analysis. Technical Report 01–64, ESAT-SISTA, Dept. EE, K.U.Leuven, 2001.

[11] A. Kukush and E.-O. Maschke. The efficiency of adjusted least squares in the linear functional relationship. Discussion Paper 208, SFB 386, Ludwig-Maximilians-University of Munich, 2000.

[12] A. Kukush and S. Zwanzig. On consistent estimators in nonlinear functional errors-in-variables models. In *this volume*.

[13] A. Kukush and S. Zwanzig. On inconsistency of the least squares estimator in nonlinear functional error-in-variables models. Preprint N96-12, Institut für Mathematische Stochastik, Universität Hamburg, 1996.

[14] M. Mühlich and R. Mester. The role of total least squares in motion analysis. In H. Burkhardt, editor, *Proc. European Conference on Computer Vision (ECCV'98)*, Springer Lecture Notes on Computer Science, pages 305–321. Springer-Verlag, June 1998.

[15] A. Premoli and M. L. Rastello. The parametric quadratic form method for solving total least squares problems with elementwise weighting. In *this volume*.

[16] P. Sprent. A generalized least-squares approach to linear functional relationships. *J. R. Statistical Society B*, (28):278–297, 1966.

[17] S. Van Huffel and J. Vandewalle. Analysis and properties of the generalized total least squares problem $AX \approx B$ when some or all columns in A are subject to error. *SIAM J. Matrix Anal.*, 10(3):294–315, 1989.

[18] S. Van Huffel and J. Vandewalle. *The Total Least Squares Problem: Computational Aspects and Analysis*. SIAM, Philadelphia, 1991.