Fourth Total Least Squares and Errors-in-Variables Modeling Workshop

Program and Abstracts

Leuven, Belgium, August 21–23, 2006

Organizing Committee

- Chairpersons: Sabine Van Huffel and Ivan Markovsky, Katholieke Universiteit Leuven, Belgium
- Secretariat: Ida Tassens (secretary), Ilse Pardon (accountant), and Bart Motmans (coordinator)

Scientific Committee

- Yasuo Amemiya, IBM Research Center, USA
- Chi-Lun Cheng, Academia Sinica, Taiwan
- Lieven De Lathauwer, CNRS–ETIS, France
- Bart De Moor, Katholieke Universiteit Leuven, Belgium
- Gene Golub, Stanford University, USA
- Helmut Kuchenhoff, University of Munich, Germany
- Alexander Kukush, Kiev National Taras Shevchenko University, Ukraine
- Nicola Mastronardi, CNR Bari, Italy
- Björn Ottersten, Royal Institute of Technology, Sweden
- Chris Paige, McGill University, Canada
- Rik Pintelon, Vrije Universiteit Brussel, Belgium
- Shalabh, Indian Institute of Technology Kanpur, India
- G.W. Stewart, University of Maryland, USA
- Richard Vaccaro, University of Rhode Island, USA
- Paul Van Dooren, Université Catholique de Louvain, Belgium

Sponsors

- Katholieke Universiteit Leuven
- Fund for Scientific Research Flanders (FWO)
- Scientific research network “Advanced numerical methods for mathematical modelling” of the Fund for Scientific Research Flanders
- The European Association for Signal Processing (EURASIP)

Contents

1 Welcome and general information

Dear participants,

it is our great pleasure to welcome all of you in the Arenberg castle, Leuven, Belgium, for the 4th edition of the workshop on Total Least Squares and Errors-In-Variables Modeling. Just like the three previous workshops, organized at the same place in August 1991, 1996 and 2001, this workshop attracts a highly interdisciplinary audience and therefore covers a broad scope of fields such as statistics, numerical analysis, system identification, signal processing, chemistry, computer vision, environmental sciences, etc., and subjects ranging from theory to practice. All lectures and poster presentations focus on the central question: “How to deal with measurement error?”. The invited lectures form the backbone of the workshop and aim to present a general overview of obtained results in a certain field over the past 5 years. Significant progress has been made in nonlinear measurement error and behavioral modeling, regularized and structured total least squares, numerical TLS algorithms, errors-in-variables system identification, geometric fitting, and applied fields such as signal processing, chemistry and environmental sciences. I strongly believe that bridging the gap between disciplines is a conditio sine qua non for tackling the scientific challenges in this area. It is our hope that this workshop will enhance this crossfertilization!

Enjoy your stay, the workshop, and the city of Leuven and its environments!

Sabine Van Huffel and Ivan Markovsky

Lunches

August 21, 12h00–13h30: ESAT (00.62 or 00.57)

August 22, 12h00–13h30: ESAT (00.62 or 00.57)

August 23, 12h00–13h30: ESAT (00.62 or 00.57)

Breakfast (for those staying in a dormitory room)

Every morning at ESAT (00.62 or 00.57).

Coffee breaks

In the Salons Arenberg Castle.

August 21, 10h30–11h00 and 15h00–15h30

August 22, 10h00–10h30 and 16h00–16h30

August 23, 10h15–10h30

Welcome drink and dinner on Monday

In the ”oude Kantien” at 19h00.

Banquet on Tuesday

In the faculty Club at 20h00.

Transportation

The Arenberg Castle is located approximately 3km. (30 minutes walking) from the center of Leuven. It is accessible by public transportation—bus number 2 direction Campus, stop De Oude Kantien. (In the opposite direction, bus number 2 goes to the station of Leuven.)

Computer facilities

Participants in the workshop will have an access to a computer lab in the building of the electrical engineering department (ESAT), see the map on page 77. Those who have laptops will be able to access Internet via a wireless connection within the ESAT building. Room 00.62 is reserved for the participants of the workshop.

Workshop address

Ida Tassens
Dept. of Electrical Engineering, ESAT-SCD (SISTA)
Katholieke Universiteit Leuven,
Kasteelpark Arenberg 10
B–3001 Leuven-Heverlee, Belgium

tel: 32/16/32.17.09, fax: 32/16/32.19.70
email: ida.tassens@esat.kuleuven.be

2 Proceedings

One special issue of *Signal Processing* and one special issue of *Computational Statistics and Data Analysis* will be published as proceedings of the workshop. However, the call for papers is also open to papers that are not presented during the workshop. We encourage authors, whose work is more statistically oriented to submit their manuscripts to Computational Statistics and Data Analysis and authors whose work is more applications oriented to submit their manuscripts for Signal Processing.

The scope of topics for the special issues overlaps with this of the workshop. The deadline for submissions for both special issues is October 1, 2006. Submissions for Computational Statistics and Data Analysis should be sent by e-mail to the workshop secretariat (`ida.tassens@esat.kuleuven.be`). Submissions for Signal Processing should be entered at `http://ees.elsevier.com/sigpro` selecting "TLS and EIV modeling" as the Article Type.

Guest editors for special issue of Signal Processing

- S. Van Huffel, K.U.Leuven, Dept. Elektrotechniek (ESAT), Kasteelpark Arenberg 10, B-3001 Leuven, Belgium
- I. Markovsky, K.U.Leuven, Dept. Elektrotechniek (ESAT), Kasteelpark Arenberg 10, B-3001 Leuven, Belgium
- R. Vaccaro, Department of Electrical & Computer Engineering, University of Rhode Island, 4 East Alumni Ave. Kingston, RI 02881, USA
- T. Söderström, Information Technology, Department of Systems and Control, Uppsala University, P O Box 337, SE-751 05, Uppsala, Sweden

Guest editors for special issue of Computational Statistics and Data Analysis

- S. Van Huffel, K.U.Leuven, Dept. Elektrotechniek (ESAT), Kasteelpark Arenberg 10, B-3001 Leuven, Belgium
- C.-L. Cheng, Institute of Statistical Science, Academia Sinica, Taipei, Taiwan, R.O.C
- N. Mastronardi, Istituto per le Applicazioni, del Calcolo "M.Picone" sez. Bari, National Council of Italy, via G. Amendola 122/D, I-70126 Bari, Italy
- C. Paige, McGill University, School of Computer Science, 3480 University Street, Montreal, PQ, Canada H3A 2A7
- A. Kukush, Kiev National Taras Shevchenko University, Volodymyrska st. 60, 01033, Kiev, Ukraine

3 Program

Monday, August 21

08h00–08h45	Registration in Arenberg Castle	
08h45–09h00	Opening and welcome by Sabine Van Huffel	

Session I: Regularized total least squares

Chair: S. Van Huffel

09h00–09h45	*G. Golub*, Matrices and moments: perturbation for least squares	11
09h45–10h30	*A. Beck*, The regularized total least squares problem: Theoretical properties and three globally convergent algorithms	11
10h30–11h00	*Break*	
11h00–11h30	*D. Sima*, Level choice in truncated total least squares	14
11h30–12h00	*A. Watson*, Robust counterparts of errors-in-variables problems	15
12h00–13h30	*Lunch*	

Session II: Nonlinear measurement error models

Chair: K. Kanatani

13h30–14h15	*C.-L. Cheng*, On the conditional score and corrected score estimation in nonlinear measurement error models	17
14h15–15h00	*A. Kukush and H. Schneeweiss*, Comparing the efficiency of structural and functional methods in measurement error models	18
15h00–15h30	*Break*	
15h30–16h00	*Shalabh* On the estimation of linear ultrastructural model when error variances are known	18
16h00–16h30	*G. Garg, Shalabh, N. Misra*, Consistent estimation of regression coefficients in measurement error model under exact linear restrictions	19

Poster Session (room 00.62 at ESAT)

17h00–18h30	*Poster Session* (for a list of posters see page 10)
19h00–22h00	Welcome drink and dinner in “De Oude Kantien”

Tuesday, August 22

Session III: Numerical methods for total least squares

Chair: G. Golub

Session IV: Geometric fitting

Chair: C.-L. Cheng

Session V: Total least squares applications in computer algebra

Chair: N. Mastronardi

Guided visit through Leuven

18h00–20h00	Guided visit through Leuven
20h00–22h00	*Banquet* in “The Faculty Club”

Wednesday, August 23

Session VI: Errors-in-variables system identification

Chair: I. Markovsky

Session VII: Total least squares applications in signal processing

Chair: R. Pintelon

List of posters

4 Abstracts

4.1 Matrices and moments: Perturbation for least squares

Gene H. Golub, Stanford University, golub@sccm.stanford.edu

Given a matrix A, $(m \times n)$ a vector b, and an approximate solution vector, we are interested in determining approximate error bounds induced by the approximate solution. We are able to obtain bounds for the perturbation using the Theory of Momnents. For an nxn symmetric, positive definite matrix A and a real vector u, we study a method to estimate and bound the quadratic form $u'F(A)u/u'u$ where F is a differentiable function. This problem arises in many applications in least squares theory eg computing a parameter in a least squares problem with a quadratic constraint, regularization and estimating backward perturbations of linear least squares problems. We describe a method based on the theory of moments and numerical quadrature for estimating the quadratic form. A basic tool is the Lanczos algorithm which can be used for computing the recursive relationship for orthogonal polynomials. We will present some numerical results showing the efficacy of our methods and will discuss various extensions of the method.

(Joint work with Zheng Su)

4.2 The regularized total least squares problem: Theoretical properties and three globally convergent algorithms

Amir Beck, Faculty of Industrial Engineering and Management, Technion - Israel Institute of Technology, Haifa, Israel, email: becka@ie.technion.ac.il

Total Least Squares (TLS) is a method for treating an overdetermined system of linear equations $Ax \approx b$, where both the matrix A and the vector b are contaminated by noise. In practical situations, the linear system is often ill-conditioned. For example, this happens when the system is obtained via discretization of ill-posed problems such as integral equations of the first kind (see e.g., [7] and references therein). In these cases the TLS solution can be physically meaningless and thus regularization is essential for stabilizing the solution.

Regularization of the TLS solution was addressed by several approaches such as truncation methods [6, 8] and Tikhonov regularization [1]. In this talk we will consider a third approach in which a quadratic constraint is introduced. It is well known [7, 11] that the quadratically constrained total least squares problem can be formulated as a problem of minimizing a ratio of two quadratic function subject to a quadratic constraint:

$$(RTLS) \quad \min_{x \in R^n} \left\{ \frac{\|Ax-b\|^2}{\|x\|^2+1} : \|Lx\|^2 \leq \rho \right\},$$

where $A \in R^{m \times n}, b \in R^m, \rho > 0$ and $L \in R^{k \times n} (k \leq n)$ is a matrix that defines a (semi)norm on the solution. The RTLS problem was extensively studied in recent years [2, 3, 7, 10, 11]. A key difficulty with this problem is its nonconvexity. As a result, several methods [7, 10] devised to solve it are not guaranteed to converge to a global optimum but rather to a point satisfying first order necessary optimality conditions.

We will present three globally and efficiently convergent algorithms, based on the algorithms proposed in [2, 3, 11], for solving the more general problem of minimizing a ratio of (possibly indefinite) quadratic functions subject to a

quadratic constraint:

$$(RQ) \quad \min_x \left\{ f(x) \equiv \frac{f_1(x)}{f_2(x)} : \|Lx\|^2 \le \rho \right\},$$

where

$$f_i(x) = x^T A_i x - 2b_i^T x + c_i, \quad i = 1,2,$$

$A_1, A_2 \in R^{n\times n}$ are symmetric matrices, $b_1, b_2 \in R^n$, $c_1, c_2 \in R$. We do not assume that A_1 and A_2 are positive semidefinite (as in the case of the RTLS problem). The only assumption made is that the problem is well defined. Surprisingly, at least with respect to the methodologies and techniques presented in the talk, there is no real advantage in dealing with the specific instance of the RTLS problem.

The procedure devised in [2] relies on the following key observation due to [5] for fractional programs:

Observation: given $\alpha \in R$, the following two statements are equivalent:
1. $\min_x \left\{ f_1(x)/f_2(x) : \|Lx\|^2 \le \rho \right\} \le \alpha$.
2. $\min_x \left\{ f_1(x) - \alpha f_2(x) : \|Lx\|^2 \le \rho \right\} \le 0$.

Based on the latter observation, we develop an efficient algorithm for finding the global optimal solution by converting the original problem into a sequence of simple optimization problems of the form

$$(GTRS) \quad \min_x \left\{ x^T Ax + 2b^T x + c : \|Lx\|^2 \le \rho \right\}$$

parameterized by a single parameter α. The optimal solution corresponds to a particular value of α, which can be found by a simple one-dimensional search. Problem (GTRS) is also known in the literature as the generalized trust region subproblem and, similarly to problem (RQ), is a nonconvex problem. Using the hidden convexity result of [4] we are able to convert the GTRS problem into a simple *convex* optimization problem that can be solved by finding the root of a one-dimensional secular equation. Overall, the algorithm finds an ε-optimal solution after solving $O(\log \varepsilon^{-1})$ GTRS problems. Practically, the numerical experiments in [2] show that a high-accuracy optimal solution is typically obtained after only few iterations.

The method devised in [11] was developed to solve the specific case of the RTLS problem. The starting point is the observation that x^* is an optimal solution of problem (RQ) if and only if

$$x^* \in \operatorname{argmin}_{y\in R^n} \{ f_2(y)(f(y) - f(x^*)) \ : \ \|Ly\|^2 \le \rho \},$$

(here $f_2(y) = \|y\|^2 + 1, f(y) = \|Ay - b\|^2/(\|y\|^2 + 1)$) which naturally leads to consider the following fixed point iterations:

$$x_{k+1} \in \operatorname{argmin}_{y\in R^n} \{ f_2(y)(f(y) - f(x_k)) : \|Ly\|^2 \le \rho \}.$$

The latter scheme, similarly to the one used in [2], also involves the solution of a GTRS problem at each iteration. A different method for solving the GTRS problem is discussed in [11]. Specifically, the GTRS is converted into an equivalent quadratic eigenvalue problem (QEP) for which efficient solvers are known to exist. The numerical results presented in [11] indicate that, similarly to the method proposed in [2], the method converges at a very fast rate and requires the solution of very few (up to 5) GTRS problems. The numerical results reported in [11] also indicate that the method produces a *global solution*. This fact was also validated empirically by comparing the two procedures in [2]. However, a proof of convergence to a global optimal solution of the RTLS was not given in [11].

The aforementioned results suggest that the problem (RQ) of minimizing a quadratically constrained ratio of two quadratic functions seems to share some kind of hidden convexity property, namely, it can be shown to be equivalent to some (tractable) convex optimization reformulation. In [3] we show that this is indeed the case. We obtain a simple condition in terms of the problem's data under which the attainment of the minimum in problem (RQ) is warranted. This condition allows us to derive an appropriate nonconvex reformulation of (RQ), and to apply an extension of the so-called S-Lemma for three quadratic homogeneous forms [9]. By so doing, we prove that problem (RQ) can be recast as a semidefinite programming problem for which efficient solution algorithms are known to exist (e.g., interior point methods). Based on the latter formulation, we propose a *third* globally and efficiently convergent algorithm for solving the RQ problem. Another byproduct of the aforementioned results is a superlinear convergence result for the iterative scheme suggested in [11], and which is extended for the *more general* class of problems (RQ). Moreover, it is shown that this algorithm produces an ε-global optimal solution in no more than $O(\sqrt{\log \varepsilon^{-1}})$ main loop iterations.

This result also provides a theoretical justification to the successful computational results reported in the context of (RTLS) in [11] and [3].

The talk is partially based on joint works with Aharon Ben-Tal and Marc Teboulle.

Bibliography

[1] A. Beck and A. Ben-Tal. On the Solution of the Tikhonov Regularization of the Total Least Squares Problem. *SIAM J. Optimization*, 17(1): 98–118, 2006.

[2] A. Beck, A. Ben-Tal and M. Teboulle. Finding a Global Optimal Solution for a Quadratically Constrained Fractional Quadratic Problem with Applications to the Regularized Total Least Squares. *SIAM J. Matrix Anal. Appl.* In print.

[3] A. Beck and M. Teboulle A Convex Optimization Approach for Minimizing the Ratio of Indefinite Quadratic Functions over an Ellipsoid. Submitted for publication.

[4] A. Ben-Tal and M. Teboulle. Hidden convexity in some nonconvex quadratically constrained quadratic programming. *Mathematical Programming*, 72(1):51–63, 1996.

[5] W. Dinkelbach. On nonlinear fractional programming. *Management Sciences*, 13:492–498, 1967.

[6] R. D. Fierro, G. H. Golub, P. C. Hansen and D. P. O'Leary. Regularization by truncated Total Least Squares. *SIAM J. Sci. Comp.*, 18(1): 1223–1241, 1997.

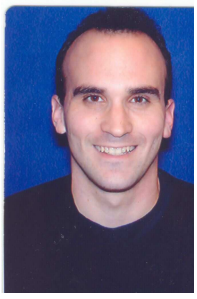
[7] G. H. Golub, P. C. Hansen and D. P. O'Leary. Tikhonov Regularization and Total Least Squares. *SIAM J. Matrix Anal. Appl.*, 21(2):185–194, 1999.

[8] P. C. Hansen. Regularization Tools, a Matlab package for analysis of discrete regularization problems. *Numerical Algorithms*, 6:1–35, 1994.

[9] B. T. Polyak. Convexity of quadratic transformations and its use in control and optimization. *J. Optim. Theory Appl.*, 99(3):553-583(1998).

[10] R. A. Renaut and H. Guo. Efficient algorithms for solution of regularized total least squares. *SIAM J. Matrix Anal. Appl.*, 26(2):457–476, 2005.

[11] D. Sima, S. Van Huffel and G. H. Golub. Regularized Total Least Squares Based on Quadratic Eigenvalue Problem Solvers, *BIT Numerical Mathematics*, 44(4):793–812, 2004.



About the author Amir Beck was born in Israel in 1975. He received the B.Sc. degree in pure mathematics (*Cum Laude*) in 1991, the M.sc. degree in operations research (*Suma Cum laude*) and the Ph.D. degree in operations research – all from Tel Aviv University (TAU), Tel Aviv, Israel. From 2003 to 2005 he was a Postdoctoral Fellow at the Minerva Optimization Center, Technion, Haifa, Israel.

He is currently a Senior Lecturer in the Department of Industrial Engineering at the Technion - Israel Institute of Technology, Haifa, Israel. His research interests are in continuous optimization, in particular, large-scale problems, conic and robust optimization as well as nonconvex quadratic optimization: theory, algorithms and applications in signal processing and communication.

4.3 Level choice in truncated total least squares

Diana Sima, Sabine Van Huffel, Katholieke Universiteit Leuven, Department of Electrical Engineering, ESAT-SCD, {diana.sima,sabine.vanhuffel}@esat.kuleuven.be

Introduction Ill-posed problems are problems where the solution does not depend continuously on the input data, where arbitrarily small perturbations in the input data produce arbitrarily large changes in the solution. Discrete ill-posed problems of the type $Ax \approx b$ might arise from the discretization of continuous problems, such as integral or differential models.

When the system $Ax \approx b$ is a *discrete ill-posed problem*, the least squares or total least squares methods yield unreliable solutions for x, dominated by noise in the data or by numerical approximation errors. This happens because discrete ill-posed problems have an intrinsic sensitivity, which is shown through the fact that the singular values of A decay without gap towards zero (or towards a "noise level").

Regularization techniques are used for many years as a way of stabilizing the computation of least squares solutions in discrete ill-posed problems. Truncated singular value decomposition and Tikhonov regularization are two of the most known methods.

Truncation methods for linear estimation The aim of regularization by truncation is to appropriately identify a good truncation level, and to construct a truncated solution that can capture the essential features of the unknown true solution, without explicit knowledge about the true solution, and even without *a priori* knowledge about the magnitude of the noise in the data.

A better understanding of truncation methods (such as truncated singular value decomposition (TSVD) and truncated total least squares (TTLS) [1]) is possible in view of the recent results on *core problems* of linear systems [4]. The core reduction of an incompatible linear system is a tool that is able to avoid the problems of nonuniqueness and nongenericity in the computation of the total least squares solution (and variations). We propose the use of *truncated core problems* in order to avoid close-to-nongenericity in ill-posed linear approximation problems.

If $A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^m$, then the TTLS core problem with truncation level k involves solving the core system

$$A_{11}^k x_1^k \approx b_1^k$$

in TLS sense, where $\begin{bmatrix} b_1^k & A_{11}^k \end{bmatrix}$ is a $(k+1)$-dimensional truncated core of $\begin{bmatrix} b & A \end{bmatrix}$, *e.g.*, a $(k+1)\times(k+1)$ partial bidiagonalization of $\begin{bmatrix} b & A \end{bmatrix}$. From the k-dimensional solution x_1^k, we can easily construct the n-dimensional TTLS solution $x_{\mathrm{TTLS},k}$, using, *e.g.*, the right orthogonal transformations associated with the partial bidiagonalization.

Choice of the truncation level Among model selection criteria that use the given data in order to select a good hyperparameter, we mention the *discrepancy principle*, the *L-curve*, *(generalized) cross validation* and *Akaike's information criterion*. Here, we focus on the generalized cross validation (GCV) [2, 5]. We make use of concepts from the field of regularization for nonlinear models [3], which are also linked to the interpretation of GCV as a rotation-invariant version of the ordinary cross-validation.

The GCV criterion can be written as

$$\min_k \frac{\|r_k\|_F^2}{\left(N - p_k^{\mathrm{eff}}\right)^2}, \tag{4.1}$$

where r_k denotes the residual between the corrected model $\begin{bmatrix} \widehat{b}_k & \widehat{A}_k \end{bmatrix}$ (reconstructed for a fixed value of k) and the given data $\begin{bmatrix} b & A \end{bmatrix}$; N is the number of "noisy" elements in the data, and the *effective number of parameters* p_k^{eff} is the trace of the so-called *generalized information matrix* [3]. The generalized information matrix is defined as the derivative of the reconstructed model with respect to the noisy data.

Since in the TLS case $A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^m$ are both considered noisy, $N = m(n+1)$. The residual error norm is given by

$$\|r_k\|_F^2 = \left\| \begin{bmatrix} b & A \end{bmatrix} - \begin{bmatrix} \widehat{b}_k & \widehat{A}_k \end{bmatrix} \right\|_F^2 = \frac{\|A\widehat{x}_{\mathrm{TTLS},k} - b\|_2^2}{\|\widehat{x}_{\mathrm{TTLS},k}\|_2^2 + 1} = \| \begin{bmatrix} b & A \end{bmatrix} \|_F^2 - \| \begin{bmatrix} b_1^k & A_{11}^k \end{bmatrix} \|_F^2 + (\sigma'')^2,$$

where σ'' is the smallest singular value of $\begin{bmatrix} b_1^k & A_{11}^k \end{bmatrix}$.

We prove that there exists a relatively easy-to-compute closed-form expression for the effective number of parameters in the truncated core TLS problem:

$$p_k^{\text{eff}} = \frac{1}{2}\text{Tr}\left\{ \left(\frac{A_{11}^{k\ \top} A_{11}^k}{(\sigma'')^2} - I_k + 8(v_1'')^2 x_1^k x_1^{k\top} \right)^{-1} \right\},$$

where (u'', σ'', v'') denotes the smallest singular triplet of the (bidiagonal) matrix $\begin{bmatrix} b_1^k & A_{11}^k \end{bmatrix}$. Note that if A_{11}^k is bidiagonal, the evaluation of p_k^{eff} involves inverting a (small) $k \times k$ tridiagonal plus rank-one matrix.

Plugging-in $\|r_k\|_F^2$ and p_k^{eff} into (4.1), we obtain a new closed-form expresion for GCV, specially adapted to the truncated TLS problem formulation. Its computation involves only the bidiagonal matrix obtained after k bidiagonalizations steps, and the smallest singular triplet of this bidiagonal matrix.

Conclusion As for truncated SVD, the truncated TLS problems admits closed-forms expressions for each of the classical model selection techniques for choosing truncation levels. We focused on generalized cross validation, which needs an important adjustment compared to the simple GCV criterion that applies to truncated SVD. However, the GCV function can still be efficiently computed during a partial bidiagonalization algorithm for truncated TLS.

Bibliography

[1] R. D. Fierro, G. H. Golub, P. C. Hansen, and D. P. O'Leary. Regularization by truncated total least squares. *SIAM Journal on Scientific Computing*, 18(1):1223–1241, 1997.

[2] G. H. Golub, M. Heath, and G. Wahba. Generalized cross-validation as a method for choosing a good ridge parameter. *Technometrics*, 21:215–223, 1979.

[3] J. E. Moody. The *effective* number of parameters: an analysis of generalization and regularization in nonlinear learning systems. In Moody, Hanson, and Lippmann, editors, *Advances in Neural Information Processing Systems*, volume 4, pages 847–854. Morgan Kaufmann, Palo Alto, 1992.

[4] C. C. Paige and Z. Strakoš. Core problems in linear algebraic systems. *SIAM Journal on Matrix Analysis and Applications*, 27(3):861–875, 2006.

[5] G. Wahba. *Spline Models for Observational Data*. CBMS-NSF Regional Conference Series in Applied Mathematics 59. SIAM, 1990.

About the author Diana M. Sima studied at the Mathematics Department, University of Bucharest, and obtained a Bachelor degree in *Computer Science* in 2000 and a Master in *Applied Statistics and Optimization*. She followed a PhD program at the Department of Electrical Engineering, K.U. Leuven, Belgium (2002-2006). Sabine Van Huffel is full professor at the Department of Electrical Engineering, K.U. Leuven, Belgium

4.4 Robust counterparts of errors-in-variables problems

G. Alistair Watson, Department of Mathematics,
University of Dundee, Dundee DD1 4HN, Scotland,
gawatson@maths.dundee.ac.uk

Let data points $(\mathbf{x}_i, y_i), i = 1, \ldots, m$, be given, where $\mathbf{x}_i \in \mathbb{R}^t$ and $y_i \in \mathbb{R}$, and all values contain errors. Let the points be related through a linear model, containing n parameters, so that we can write

$$y_i \approx \sum_{j=1}^{n} a_j \phi_j(\mathbf{x}_i),\ i = 1, \ldots, m, \tag{4.2}$$

where $\mathbf{a} \in I\!R^n$ is the vector of free parameters. Writing this in matrix/vector form as $\mathbf{y} \approx A\mathbf{a}$, then total least squares can be used to find values of the variables. In certain application areas, it may be more appropriate to solve a **robust counterpart** of this [1], which may be interpreted as the requirement to minimize $\|\tilde{\mathbf{y}} - \tilde{A}\mathbf{a}\|$ with respect to **a** over the **worst** of perturbations defined by $(\tilde{\mathbf{y}}, \tilde{A}) \in \mathscr{E}$, where $\mathscr{E}$ is an uncertainty set. For example, if

$$\mathscr{E} = \{(\mathbf{y} + \mathbf{r}, A + E) : \|\mathbf{r}\| \leq \rho_1, \|E\| \leq \rho_2\},$$

where ρ_1, ρ_2 are given, then the problem is equivalent to finding **a** to solve

$$\min_{\mathbf{a} \in R^n} \max_{\|\mathbf{r}\| \leq \rho_1, \|E\| \leq \rho_2} \|\mathbf{y} + \mathbf{r} - (A + E)\mathbf{a}\|,$$

and when the norms are least squares norms, good methods are available to compute solutions (for example, [3], [4]).

An alternative to total least squares is orthogonal distance regression [2], where direct perturbations of the variables themselves are considered. Let

$$\tilde{z}_i = \tilde{y}_i - \sum_{j=1}^{n} a_j \phi_j(\tilde{\mathbf{x}}_i), i = 1, \ldots, m.$$

Then robust counterparts of the basic problem correspond to the minimization of $\|\tilde{\mathbf{z}}\|$ with respect to **a** over the **worst** of all perturbations such that

$$(\tilde{y}_1, \ldots, \tilde{y}_m, \tilde{\mathbf{x}}_1, \ldots, \tilde{\mathbf{x}}_m) \in \mathscr{E},$$

where $\mathscr{E}$ is an uncertainty set. We consider such problems for different sets $\mathscr{E}$, and in particular for the set

$$\mathscr{E} = \{y_1 + r_1, \ldots, y_m + r_m, \mathbf{x}_1 + \mathbf{s}_1, \ldots, \mathbf{x}_m + \mathbf{s}_m : |r_i| \leq \rho_i, \|\mathbf{s}_i\|_A \leq \gamma_i, i = 1, \ldots, m\},$$

where $\|.\|_A$ is a norm on $I\!R^t$, and where $\rho_i, \gamma_i, i = 1, \ldots, m$ are given.

Because the orthogonal distance regression problem does not depend on the model being linear, we can also consider the treatment of nonlinear problems.

The intention is to use and build on ideas presented in [5], and the main focus is on algorithmic development.

Bibliography

[1] A. Ben-Tal and A. Nemirovski. Robust convex optimization. *Mathematics of Operations Research*, 23:769–803, 1998.

[2] P. T. Boggs, R. H. Byrd, and R. B. Schnabel. Stable and efficient algorithm for nonlinear orthogonal distance regression, *SIAMJ. Sci. Stat. Comput.*, 8:1052–1078, 1987.

[3] S. Chandrasekaran, G. H. Golub, M. Gu and A. H. Sayed. An efficient algorithm for a bounded errors-in-variables model. *SIAM J. Matrix Anal. Appl.*, 20:839–859, 1999.

[4] L. El Ghaoui and H. Lebret. Robust solutions to least squares problems with uncertain data. *SIAM J. Matrix Anal. Appl.*, 18:1035–1064, 1997.

[5] G. A. Watson. Robust solutions to a general class of approximation problems. *SIAM J. Sci. Stat. Comp.*, 25:1448–1460, 2004.

About the author Following BSc and MSc degrees at Edinburgh, and a PhD at the Australian National University in Canberra (supervised by Mike Osborne), Alistair Watson joined the University of Dundee in 1969. He was appointed to a Chair in Approximation Theory in 1988, and currently holds the Chair of Numerical Analysis (succeeding Ron Mitchell in 1993). He has been involved in the organisation of the Dundee Biennial Numerical Analysis conferences since 1971, and served as one of the main editors of the IMA Journal of Numerical Analysis from 1989 to 2004.

4.5 On the conditional score and corrected score estimation in nonlinear measurement error models

Chi-Lun Cheng, Institute of Statistical Science, Academia Sinica, Taiwan, R.O.C.
clcheng@stat.sinica.edu.tw

This paper reviews the conditional score and corrected score estimation of the unknown parameters in nonlinear measurement error (errors-in-variables) models. This includes the functional and structural models. The connection among these methodologies and total least squares (TLS) is also examined. A compendium of existing results as well as some possible extensions are discussed.

The ordinary regression models assume that the independent variables are measured without error. However, in many situations, the independent variables cannot measured precisely. When the measurement error is too large to ignore, the estimators for the regression parameters are biased and inconsistent. Measurement error models are important alternatives for ordinary regression models, in which we assume that relation between the independent variable y and independent variable ξ is known but one cannot observe ξ directly. Instead, one observes $\mathbf{x} = \xi + \delta$, where δ is independents of ξ and has mean zero.

The linear measurements error model has a long history and is dated back 1877 (Adcock, 1977), which has been well investigated. For a summary, see Fuller (1987) and Cheng and Van Ness (1999). For the past two decades, the researches on measurement error models are more focused on nonlinear measurement error models, see Carroll, Ruppert and Stefanski (1995) for a reference.

There are two general methodologies proposed in the literature to estimate the regression parameters in nonlinear measurement error models. The first one is the *conditional score* method that was proposed by Stefanski and Carroll (1987). The second one is called *corrected score* method, which was proposed by Stefanski (1989) and Nakamura (1990) independently.

In this paper, we will review these two methods. In our view, they have some fundamental difference in their assumptions that has been neglected in the literature. We will also bring some recent developments to attention and some possible extensions are discussed. Finally the connection between the conditional score method and TLS (total least squares) is addressed.

Bibliography

[1] R. J. Adcock. Note on the method of least squares. *Analyst*, 4: 183-184, 1877.

[2] R. J. Carroll, D. Ruppert, and L. A. Stefanski. Measurement Error in Nonlinear Models. *London: Chapman and Hall*, 1995.

[3] C. L. Cheng, J. W. Van Ness. Statistical Regression with Measurement Error *London: Chapman and Hall*, 1995.

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

[5] T. Nakamura. Corrected score functions for errors-in-variables models: methodology and application to generalized linear models. *Biometrika*, 77: 127-137, 1990.

[6] L. A. Stefanski. Unbiased estimation of a nonlinear function of a normal mean with application to measurement error models. *Comm. Statist. A*, 18: 4335-4358, 1989.

[7] L. A. Stefanski, and R. J. Carroll. Conditional scores and optimal scores in generalized linear measurement error models. *Biometrika*, 74: 703-716, 1987.

About the author Dr. Cheng received Ph.D. in 1987 at University of Texas (Dallas). He is Research Fellow, Institute of Statistical Science, Academia Sinica, Taiwan, R.O.C. His research interest is statistical inferences on measurement error (errors-in-variables) models.

4.6 Comparing the efficiency of structural and functional methods in measurement error models

Hans Schneeweiss, Department of Statistics, University of Munich, schneew@stat.uni-muenchen.de
Alexander Kukush, Department of Mechanics and Mathematics, Kiev National Taras Shevchenko University
alexander_kukush@univ.kiev.ua

The paper is a survey of recent investigations by the authors and others into the relative efficiencies of structural and functional estimators of the regression parameters in a measurement error model. While structural methods, in particular the quasi-score (QS) method, take advantage of the knowledge of the regressor distribution (if available), functional methods, in particular the corrected score (CS) method, discards such knowledge and works even if such knowledge is not available. Among other results, it has been shown that QS is more efficient than CS as long as the regressor distribution is completely known. However, if nuisance parameters in the regressor distribution have to be estimated, this is no more true in general. For small measurement errors, the efficiencies of QS and CS (and also of ML) become almost indistinguishable, whether nuisance parameters are present or not. QS is (asymptotically) biased if the regressor distribution has been misspecified, while CS is always consistent and thus more robust than QS.

4.7 On the estimation of linear ultrastructural model when error variances are known

Shalabh, Indian Institute of Technology Kanpur, India, shalab@iitk.ac.in, shalabh1@yahoo.com
H. Schneeweiss, University of Munich, Germany

In a linear measurement error model, the parameters can be estimated consistently only when some additional information besides the data set is available. There are various formulations that are commonly employed; e.g., [1] and [2]. Among them, an interesting formulation relates to the specification of the variances of the measurement errors associated with the variable in the linear relationship. Under such a specification, the slope parameter in a bivariate model is estimated by the technique of orthogonal regression in which the sum of squares of the perpendicular distances, rather than the horizontal and vertical distances, from the data points to the regression line is minimized. The resulting estimation turns out to be the maximum likelihood estimator if the errors are assumed to be normally distributed

Besides the techniques of orthogonal regression for the estimation of parameters, there are other alternative procedures but they have received far less attention in the literature of measurement error models. For instance, we may employ the technique of reduced major axis in which the slope parameter is estimated by the geometric mean of the two estimators arising from the direct and inverse regression; see, e.g., [3] for and expository account. Similarly, we may estimate the slope parameter by the arithmetic mean of the two estimators; see, e.g., [4]. Likewise the slope parameter may be estimated by the slope of the line that bisects the angle between the direct and inverse regression lines ; see, e.g., [5]. This paper considers all these techniques in the context of a linear ultrastructural model and discuss their asymptotic properties.

A simple question then arises that out of these suggested estimators, which estimation is better under what conditions. It can be well appreciated that the reliability ratios associated with study and explanatory variables are easily available or can be well estimated in measurement error models, see [6] and [7] for more details on this aspect. So an attempt is made in this paper to express the efficiency properties of all the estimators under consideration as a function of reliability ratios associated with study and explanatory variables only. This helps in obtaining the conditions for the superiority of one estimator over the other in terms of reliability ratios only.

Further, most of the literature associated with measurement error models generally assumes the normal distribution for the measurement errors. In practice, such an assumption may not always hold true. The distribution of measurement errors essentially depends on the nature of experiment. The specification of normally may thus sometime leads to invalid and erroneous statistical consequences. The effect of departure from normality is another aspect of study which is attempted in this paper.

The finite sample properties of the proposed estimators under different type of distributions of measurement errors is studied through a Monte-Carlo experiment. The validity of large sample approximations in small samples is also reported.

Bibliography

[1] C.L. Cheng and J.W. Van Ness. Statistical Regression with Measurement Errors, Arnold Publishers, 1999.

[2] W.A. Fuller. Measurement Error Models, John Wiley, 1987.

[3] R.R. Sokal and F.J. Rohalf. The Principal and Practice of Statistics in Biological Research, Second Edition, Freeman, 1981.

[4] M. Aaronsom, G. Bothum, J. Moultd, J. Huchra, R.A. Schommer and M.E. Cornell. A distance scale from the infrared magnitude/HI velority-width relation.5. distance moduli to 10 galaxy clusters, and positive detection of bulk supercluster motion toward the microwave anistrophy, Astrophysical Journal, 302, 536–563, 1986.

[5] M.J. Pierce and R.B. Tully. Distances to virgo and ursa major clusters and a determination of H_o, Astrophysical Journal, 330, 579–595, 1988.

[6] L.J. Gleser. The importance of assessing measurement reliability in multivariate regression, Journal of American Statistical Association, 87, 419, 696–407, 1992.

[7] L.J. Gleser. Estimation of slope in linear errors-in-variables regression models when the predictors have known reliability matrix, Statistics and Probability letters, 17 (2), 113–121, 1993.

Shalabh received his Master's and Ph.D. degrees in Mathematical Statistics from University of Lucknow, India in 1992 and 1996 respectively. He worked at University of Jammu (India), Panjab University (India) and then moved to Indian Institute of Technology Kanpur, India. He completed his Post-doc at University of Pittsburgh (U.S.A.) under Professor Leon J. Gleser. He has received Jan Tinbergen Award from International Statistical Institute, Young Scientist Award from Indian Science Congress Association, BOYSCAST fellowship from Government of India and Humboldt fellowship from Alexander von Humboldt Foundation, Germany. He is an Associate Editor of Journal of Multivariate Analysis.

4.8 Consistent estimation of regression coefficients in measurement error model under exact linear restrictions

Gaurav Garg, Shalabh, Neeraj Misra, Department of Mathematics & Statistics, Indian Institute of Technology, Kanpur-208016, India, ggarg@iitk.ac.in, shalab@iitk.ac.in, neeraj@iitk.ac.in

In linear regression models, the ordinary least squares estimator (OLSE) is inconsistent and biased when the observations on variables are observed with measurement errors. It is well known that in order to obtain the consistent estimators of regression coefficients, some additional information from outside the sample, e.g., measurement error variance, ratio of measurement error variances or reliability ratio etc. is required.

In many situations, some prior information on the regression coefficients is available which can be used to improve upon the OLSE. When such prior information can be expressed in the form of exact linear restrictions binding the regression coefficients and the data is observed without measurement errors, the restricted least squares estimator is commonly used. When there are no measurement error in the data, this estimator is consistent, satisfies the given linear restrictions on regression coefficients and has smaller variability than the OLSE. However the restricted least squares estimator becomes inconsistent and biased, when the observations are contaminated with measurement errors. So the problem of obtaining a consistent estimator, which also satisfies the restrictions is addressed in this paper.

The multivariate ultrastructural model is considered and no assumption is made about the distributional form of any of the measurement errors and random error component in the model. Only the existence and finiteness of first four moments of measurement errors and random error component are assumed. The additional knowledge of reliability matrix and covariance matrix of measurement errors associated with explanatory variables is used to obtain the consistent estimators which also satisfy the given restrictions.

The bias vectors and mean squared error matrices of the estimators are derived and studied using the large sample approximation theory. An inter-comparison of both the estimators is made and dominance conditions for the superiority of one estimator over the other are obtained under structural and functional forms of the measurement error models. The effect of departure from normal assumption is also studied. A Monte-Carlo simulation experiment is also conducted to study the behaviour of estimators in finite samples.

Bibliography

[1] A.K. Srivastava and Shalabh. Consistent Estimation for the non-normal ultrastructural model, *Stat. & Prob. Letters*, 34:67–73. 1997.

[2] A.K. Srivastava and Shalabh. Asymptotic efficiency properties of least squares in an ultrastructural model. *Sociedad Espanola de Estadistica e Investigacion Operativa*. 6(2):419–431. 1997.

[3] C.-L. Cheng and J.W. Van Ness. On the unreplicated ultrastructural model. *Biometrika*, 78:442–445, 1991.

[4] C.-L. Cheng and J.W. Van Ness. *Statistical Regression with Measurement Errors*. Arnold Publishers. 1999.

[5] C.R. Rao, H. Toutenberg. *Linear Models : Least Squares and Alternatives*, 2^{nd} *Ed.*. Springer. 1999.

[6] C.R. Rao and M.B. Rao. *Matrix Algebra and its Applications to Statistics and Econometrics*. World Scientific. 1998.

[7] E. Malinvaud. *Statistical Methods of Econometrics*. North Holland Publishing Company. 1966.

[8] H. Schneeweiss. Consistent Estimation of a Regression with Errors in the Variables. *Metrika*. 23:101–115. 1976.

[9] H. Schneeweiss. Note on a linear model with errors in the variables and with trend. *Statist. Papers*. 32:261-264. 1991.

[10] H. ToutenBurg. *Prior Information in Linear Models*. John Wiley. 1982.

[11] J.L. Gleser. A note on G.R. Dolby's ultrastructural model. *Biometrika*. 72:117–124. 1984.

[12] O. Rosen, J.L. Gleser and Shalabh. On the usefulness of the knowledge of error variances in the consistent estimation of an unreplicated ultrastructural model. *Journal of Statistical Computation & Simulation*. 74(6):391-417. 2004.

[13] S. Penev and T. Raykov. On choosing estimators in a simple linear error-in-variable model. *Communications in Statistics : Theory & Methods*. 22(9):2229-2444. 1993.

[14] W.A. Fuller. *Measurement Error Models*. John Willey. 1987.

Gaurav Garg completed his M.Sc. and M.Phil. degrees in Statistics from C.C.S. University, Meerut, India in 2000 and 2001 respectively. Currently he is working as a Ph.D. student at Indian Institute of Technology Kanpur, India. He is working in the area of measurement error models and linear regression models.

4.9 Bidiagonalization as a fundamental decomposition of data in linear approximation problems

Christopher C. Paige, McGill University, Montreal, Canada, chris@cs.mcgill.ca

Zdeněk Strakoš, Institute of Computer Science, Academy of Sciences of the Czech Republic, Prague, strakos@cs.cas.cz

Let A be a nonzero n by k real matrix, and b be a nonzero real n-vector. Consider estimating x from the linear approximation problem

$$Ax \approx b\,, \tag{4.3}$$

where the uninteresting case is for clarity of exposition excluded by the obvious assumption $b \perp \mathscr{R}(\mathscr{A})$, that is $A^T b \neq 0$. In a sequence of papers [1, 2, 3] it was proposed to orthogonally transform the the original data b, A into the form

$$P^T \left[\; b \;\middle\|\; AQ \;\right] = \left[\begin{array}{c|c||c} b_1 & A_{11} & 0 \\ \hline 0 & 0 & A_{22} \end{array}\right], \tag{4.4}$$

where $P^{-1} = P^T$, $Q^{-1} = Q^T$, $b_1 = \beta_1 e_1$, and A_{11} is a lower bidiagonal matrix with *nonzero bidiagonal elements*. The matrix A_{11} is either square, when (4.3) is compatible, or rectangular, when (4.3) is incompatible. The matrix A_{22}, and the corresponding block row and column in (4.4), can be nonexistent. The original problem is in this way decomposed into the approximation problem

$$A_{11} x_1 \approx b_1\,, \tag{4.5}$$

and the remaining part $A_{22} x_2 \approx 0$. It was proposed to find x_1 from (4.5), set $x_2 = 0$, and substitute for the solution of (4.3)

$$x \equiv Q \left[\begin{array}{c} x_1 \\ 0 \end{array}\right]. \tag{4.6}$$

The (partial) upper bidiagonalization of $[b, A]$ described above has remarkable properties, see [3, Theorems 2.2, 3.2 and 3.3].

- First, the lower bidiagonal matrix A_{11} with nonzero bidiagonal elements has full column rank and its singular values are simple. Consequently, any zero singular values or repeats that A has must appear in A_{22}.
- Second, A_{11} has minimal dimensions, and A_{22} has maximal dimensions, over all orthogonal transformations giving the block structure in (4.4), without any additional assumptions on the structure of A_{11} and b_1.
- Finally, all components of $b_1 = \beta_1 e_1$ in the left singular vector subspaces of A_{11}, that is, the first elements of all left singular vectors of A_{11} (multiplied by $\beta_1 \neq 0$), are nonzero.

In the approach represented by (4.3)–(4.6), the data b, A are fundamentally decomposed. The necessary and sufficient information for solving the problem (4.3) is given by b_1, A_{11}. All irrelevant and repeated information is filtered out to A_{22}. The problem (4.5) is therefore called a *core problem* within (4.3).

The core problem formulation can be used to solve least squares, scaled total least squares and data least squares problems. The core problem solutions are identical to the minimal 2-norm solutions of all formulations of the (scaled) total least squares problem with the single right hand side [4]. It gives the minimum norm solution determined by the algorithm of Golub and Van Loan [5], [4, Theorem 3.7, p. 58], if it exists. If such a solution does not exist, then the core problem approach gives the nongeneric minimum norm (scaled) total least squares solution described by Van Huffel and Vandewalle [4, Theorem 3.12, p. 72]. In this way, one simple and efficient approach can be applied to different classes of problems. The core problem formulation also offers a new theoretical insight into hybrid methods for solving noisy ill-posed problems [6, Section 6.6], and several authors have reported promising preliminary results in this direction.

In our contribution we will review the theory and recent applications of the core problem formulation, and describe the status of investigation of several open questions.

Bibliography

[1] C. C. Paige and Z. Strakoš. Scaled total least squares fundamentals. *Numer. Math.*, 91:117–146, (2002).

[2] C. C. Paige and Z. Strakoš. Unifying least squares, total least squares and data least squares, in "Total Least Squares and Errors-in-Variables Modeling", S. van Huffel and P. Lemmerling, editors, Kluwer Academic Publishers, Dordrecht, pp. 25–34, (2002).

[3] C. C. Paige and Z. Strakoš. Core problems in linear algebraic systems. *SIAM. J. Matrix Anal. Appl.*, to appear, (2006).

[4] S. VAN HUFFEL AND J. VANDEWALLE, *The Total Least Squares Problem: Computational Aspects and Analysis*, SIAM Publications, Philadelphia PA, 1991.

[5] G. H. Golub and C. F. Van Loan. An analysis of the total least squares problem. *SIAM J. Numer. Anal.*, 17:883–893, (1980).

[6] P. C. HANSEN, *Rank Deficient and Discrete Ill-Posed problems*, SIAM, Philadelphia PA, 1998.

About the authors Chris Paige graduated from the University of Sydney, Australia, and got his Ph.D. from the University of London, England, where he taught from 1967 to 1972. That year he moved to McGill University in Montreal, where he has been since, except for two years on leave at the University of Sydney in 1985 and 1986. In 1997 he retired from teaching because of poor health, and is now an Emeritus Professor at McGill University in the School of Computer Science. His primary research area is numerical linear algebra.

Zdeněk Strakoš graduated from the Czech Technical University in Prague, and got the Ph.D. and D.Sc. degrees from the Academy of Sciences of the Czech Republic. He currently works as a leading researcher at the Institute of Computer Science of the Academy, and teaches at several universities in the Czech Republic. He spent four years in the USA (IMA, University of Minnesota, and Emory University, Atlanta). His primary research area is numerical linear algebra; he actively collaborates with scientists throughout the world.

4.10 A band-Lanczos algorithm for least squares and total least squares problems

Åke Björck, Linköping University

Recently C. C. Paige and Zdenek Strakoš [7] have shown that the bidiagonaliztion

$$U^T \begin{pmatrix} b & AV \end{pmatrix} = \begin{pmatrix} \beta_1 e_1 & L \end{pmatrix},$$

where U and V are orthogonal and L lower bidiagonal, can be used to extract a regular core problem for the TLS and related linear approximation problems $Ax \approx b$. For computing this bidiagonal decomposition Golub and Kahan [3] gave two mathematically equivalent algorithms. The first uses Householder transformations applied alternately from left and right. It is a very stable algorithm and the method of choice for dense problems. The second algorithm uses a coupled two-term Lanczos recurrence, as in LSQR by Paige and Saunders [6]. It has the advantage that the matrix A is not explicitly transformed and therefore it is suitable for large scale problems with A sparse. An inherent drawback is that a loss of orthogonality will occur in the computed columns in U and V.

In this talk we develop a similar reduction for the multidimensional TLS problem

$$\min_{E,\,F} \| \begin{pmatrix} E & F \end{pmatrix} \|_F, \quad (A+E)X = B+F,$$

with d right-hand sides $B = (b_1, \ldots, b_d)$, For this the consistency relations can be written

$$\begin{pmatrix} B+F & A+E \end{pmatrix} \begin{pmatrix} -I_d \\ X \end{pmatrix} = 0,$$

and we now seek $(F,\ E)$ that reduces the rank of the matrix $(A,\ B)$ by d. The multidimensional TLS problem *cannot*, as the corresponding LS problem, be reduced to d separate problems.

We present an orthogonal decomposition of $\begin{pmatrix} B & A \end{pmatrix}$,

$$U^T \begin{pmatrix} B & AV \end{pmatrix} = \begin{pmatrix} R & L \end{pmatrix},$$

where R and L have band-width $d+1$, which generalizes the bidiagonal decomposition. As for $d = 1$, it can be realized by Householder reflections (dense case) or by a Lanczos-like recursion (sparse case). The Lanczos algorithm is new, but related to the symmetric band Lanczos algorithm of Axel Ruhe [8]. Ruhe's algorithm has been refined and generalized in several recent papers, notably in [1]. A survey of the symmetric case is given by R. W. Freund in [2].

With a single starting vector the algorithm terminates with a core problem if a zero element is encountered in one of the diagonals of L. When $d > 1$, the bandwidth can be reduced by one in the Householder algorithm, whenever a zero occurs in one of the outermost bands of L. When the bandwidth has been reduced d times the process will terminate with a separable problem.

The orthonormal columns in $U = (u_1, u_2, u_3, \dots)$ and $V = (v_1, v_2, v_3, \dots)$ form bases for the left right Krylov sequence

$$\begin{aligned} &b_1, b_2, (AA^T)b_1, (AA^T)b_2, (AA^T)^2 b_1, (AA^T)^2 b_2, \dots, \\ &A^T b_1, A^T b_2, (A^T A)A^T b_1, (A^T A)A^T b_2, (A^T A)^2 A^T b_1, \dots. \end{aligned}$$

The band Lanczos algorithm needs to include a deflation process in order to detect and remove linearly dependent vectors in these sequences. Suppose, for example, that the vector $(AA^T)b_2$ is linearly dependent upon previous vectors in the left Krylov sequence. Then this and all later vectors $(AA^T)^k b_2$, $k > 1$, must be removed. Since the left and right Krylov sequence are coupled, all vectors of the form $(A^T A)^k A^T b_2$ in the right Krylov sequence are must also be removed!

When deflation occurs the length of the recursions is reduced by one. When the bandwidth has been reduced d times the left and right Krylov subspaces have reached their maximal dimensions and the process terminates. Diana Sima [9] gives a rigorous implementation of this band Lanczos algorithm.

To solve the TLS problem the partial SVD of the reduced band matrix is needed. As a first step a reduction to bidiagonal form and performed. This can be done efficiently in $\approx 4n2(d-1)$ flops using the vectorized algorithm of Kaufman [5].

Bibliography

[1] J. I. Aliaga, D. L. Boley, R. W. Freund, and V. Hernández. A Lanczos-type method for multiple starting vectors. *Math. Comp..*, 69:1577–1601, 2000.

[2] Roland W. Freund. Band Lanczos method. In Zhaojun Bai, James W. Demmel, Jack J. Dongarra, Axel Ruhe, and Henk A. van der Vorst, editors, *Templates for the Solution of Algebraic Eigenvalue Problems: A Practical Guide*, pages 80–88. SIAM, Philadelphia, PA, 2000.

[3] G. H. Golub and W. Kahan. Calculating the singular values and pseudo-inverse of a matrix. *SIAM J. Numer. Anal. Ser. B*, 2:205–224, 1965.

[4] G. H. Golub, F. T. Luk, and M. L. Overton. A block Lanczos method for computing the singular values and corresponding singular vectors of a matrix. *ACM Trans. Math. Software*, 7:149–169, 1981.

[5] L. Kaufman. Banded eigenvalue solvers on vector machines. *ACM Trans. Math. Software*, 10:73–86, 1984.

[6] C. C. Paige and M. A. Saunders. LSQR. An algorithm for sparse linear equations and sparse least squares. *ACM Trans. Math. Software*, 8:43–71, 1982.

[7] C. C. Paige and Zdenek Strakoš. Core problems in linear algebraic systems. *SIAM J. Matrix. Anal. Appl.*, 2005.

[8] Axel Ruhe. Implementation aspects of band Lanczos algorithms for computation of eigenvalues of large sparse symmetric matrices. *Math. Comp.*, 33:680–687, 1979.

[9] Diana Maria Sima. *Regularization Techniques in Model Fitting and Parameter Estimation*. Ph.D. thesis, Katholieke Universiteit Leuven, Leuven, Belgium, 2006.

4.11 Minimal backward perturbations for data least squares problems

X.-W. Chang, School of Computer Science, McGill University, chang@cs.mcgill.ca
G.H. Golub, Department of Computer Science, Stanford University, golub@stanford.edu
C.C. Paige, School of Computer Science, McGill University, paige@cs.mcgill.ca

Let $A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^m$ with $m \geq n$. The data least squares (DLS) problem is defined as follows (see, e.g., [2], [4], [5]):

$$\min_{E,x} \|E\|_F^2, \text{ s.t. } (A+E)x = b. \tag{4.7}$$

It is known that the DLS problem is equivalent to $\min_x \|Ax-b\|_2^2/\|x\|_2^2$ (see, e.g., [5]).

Given a nonzero $y \in \mathbb{R}^n$, we might want to verify whether it is, for example, a backward stable solution to the DLS problem (4.7). So we would like to solve the following minimal backward perturbation problem:

$$\min_{\Delta A} \|\Delta A\|_F^2, \text{ s.t. } y = \arg\min_x \frac{\|(A+\Delta A)x-b\|_2^2}{\|x\|_2^2}. \tag{4.8}$$

From [5, (5.14)–(5.17)], y solves the perturbed DLS problem in (4.8) if and only if

$$(A+\Delta A)^T[(A+\Delta A)y-b] = y\sigma_M^2, \quad \sigma_M^2 \equiv \|(A+\Delta A)y-b\|_2^2/\|y\|_2^2 < \sigma_{min}^2(A+\Delta A). \tag{4.9}$$

Let $r = b-Ay$. Then from (4.9), we see that ΔA is a backward perturbation for our DLS problem if and only if it is in the following set

$$\begin{aligned}\mathscr{C} \equiv \big\{\Delta A : (y^Ty)(A+\Delta A)^T(\Delta Ay-r) - \big(\Delta Ay-r\big)^T(\Delta Ay-r)y = 0,\\ \|(A+\Delta A)y-b\|_2^2/\|y\|_2^2 < \sigma_{min}^2(A+\Delta A)\big\}.\end{aligned} \tag{4.10}$$

Since the inequality in (4.10) makes it difficult to derive a general expression for $\Delta A \in \mathscr{C}$, we consider a larger set

$$\mathscr{C}_+ \equiv \big\{\Delta A : \ (y^Ty)(A+\Delta A)^T(\Delta Ay-r) - \big(\Delta Ay-r\big)^T(\Delta Ay-r)y = 0\big\}. \tag{4.11}$$

In [1], a general expression for $\Delta A \in \mathscr{C}_+$ was derived as follows

Theorem 1 *The set $\mathscr{C}_+$ satisfies*

$$\mathscr{C}_+ = \{(I-vv^\dagger)ry^\dagger + vy^\dagger - vv^\dagger A + (I-vv^\dagger)Z(I-yy^\dagger) : v \in \mathbb{R}^m,\ Z \in \mathbb{R}^{m\times n},\ v^Tb = 0\}.$$

This theorem can also be derived by following the more complicated approach of [7]. Based on Theorem 1, we can obtain the following result.

Theorem 2 *Let $\lambda_* \equiv \lambda_{\min}\big((I-bb^\dagger)A(I-2yy^\dagger)A^T(I-bb^\dagger)\big)$. Then*

$$\begin{aligned}\mu_F(y) &\equiv \min_{\Delta A\in\mathscr{C}_+} \|\Delta A\|_F = \left(\frac{\|r\|_2^2}{\|y\|_2^2} + \lambda_*\right)^{1/2}\\ &= \sigma_{\min}\left(\left[(I-bb^\dagger)A(I-yy^\dagger), \frac{\|r\|_2}{\|y\|_2}(I-bb^\dagger)(I-rr^\dagger), \frac{\|r\|_2}{\|y\|_2}b_*\right]\right),\end{aligned}$$

where

$$b_* \equiv \begin{cases} b/\|b\|_2, & b \neq 0, \\ 0 & b = 0. \end{cases}$$

The following matrix is a solution to $\min_{\Delta A \in \mathscr{C}_+} \|\Delta A\|_F$*:*

$$\widehat{\Delta A} = \begin{cases} ry^\dagger, & \lambda_* = 0, \\ ry^\dagger - v_* v_*^T (A + ry^\dagger) - (v_*^T A y) v_* y^\dagger, & \lambda_* < 0, \end{cases}$$

where v_* *is a unit eigenvector of* $(I - bb^\dagger)A(I - 2yy^\dagger)A^T(I - bb^\dagger)$ *corresponding to* λ_**.*

Since $\mathscr{C} \subseteq \mathscr{C}_+$, $\mu_F(y) \leq \min_{\Delta A \in \mathscr{C}} \|A\|_F$, i.e., $\mu_F(y)$ is a lower bound on the optimal backward perturbation bound of DLS. However our computational experience so far indicates that when y is a reasonable approximation to the exact solution of the DLS problem (4.7), then y and the optimal $\widehat{\Delta A}$ satisfy the inequality in (4.9). Therefore, in such cases, $\mu_F(y)$ is the minimal backward perturbation bound of DLS.

Computing $\mu_F(y)$ directly is expensive. But we can derive an asymptotic estimate of $\mu_F(y)$ by using Corollary 3.4 of [3]. In fact, we can show that the quantity

$$\tilde{\mu}_F(y) \equiv \frac{1}{\|y\|_2} \left\| \left(A^T A + \frac{\|r\|_2^2}{\|y\|_2^2} I \right)^{-1/2} \left(A^T r + \frac{\|r\|_2^2}{\|y\|_2^2} y \right) \right\|_2$$

is an asymptotic estimate of $\mu_F(y)$, i.e., $\lim_{y \to x} \tilde{\mu}_F(y)/\mu_F(y) = 1$, where x is the exact solution to

$$(x^T x)A^T(b - Ax) + (b - Ax)^T(b - Ax)x = 0,$$

which is the equation in (4.11) with ΔA replaced by 0 and y by x. We can use two approaches to evaluating $\tilde{\mu}_F(y)$. One is to use the QR factorization of B, and the other is to use the moment method, see [6].

Bibliography

[1] X.-W. Chang, C. C. Paige and D. Titley-Peloquin, *Characterizing matrices that are consistent with given solutions*, in preparation.

[2] G. CIRRINCIONE, G. GANESAN, K.V.S. HARI, S. VAN HUFFEL, *Direct and neural techniques for the data least squares problem,* Proceedings of the 14th International Symposium on Mathematical Theory of Networks and Systems, Perpignan, France, June 19-23, 2000.

[3] J. F. Grcar, *Optimal sensitivity analysis of linear least squares*, Technical Report LBNL-52434. Lawrence Berkeley National Laboratory, 2003.

[4] R. D. D. Groat and E. M. Dowling, *The data least squares problem and channel equalization*, IEEE Trans. Signal Processing, 42 (1993), pp. 407–411.

[5] C.C. Paige and Z. Strakoš, *Scaled total least squares fundamentals*, Numerische Mathematik, 91:117–146 (2002).

[6] Z. Su, *Computational Methods for Least Squares Problems and Clinical Trials*. Ph.D Thesis, Scientific Computing & Computational Mathematics, Stanford University, 2005.

[7] B. WALDÉN, R. KARLSON AND J. SUN, *Optimal backward perturbation bounds for the linear least squares problem,* Numerical Linear Algebra with Applications, 2:271–286 (1995).

About the authors X.-W. Chang is Associate Professor of Computer Science at McGill University. His research interests are in the area of numerical linear algebra and its applications. Currently he is interested in numerical methods for estimation, such as real least squares estimation, integer least squares estimation, robust estimation, partially linear model estimation, etc, and their applications in global navigation satellite systems.

G.H. Golub is Fletcher Jones Professor of Computer Science at Stanford University. His work in matrix computations devises and analyzes algorithms for solving numerical problems arising in science and statistics. Specifically, he develops algorithms for solving linear systems with special structure, computes eigenvalues of sequences of matrices, and estimates functions of matrices.

C.C. Paige is Professor Emeritus of Computer Science at McGill University. His main research area is numerical linear algebra and its applications. Currently he is interested in large sparse matrix problems, Krylov subspace methods, parameter estimation, ill-posed problems and rounding error analyses.

4.12 Characterizing matrices consistent with given approximate solutions to LS, DLS, and scaled TLS problems

X.-W. Chang, School of Computer Science, McGill University, chang@cs.mcgill.ca
C.C. Paige, School of Computer Science, McGill University, paige@cs.mcgill.ca
D. Titley-Peloquin, School of Computer Science, McGill University, dtitle@cs.mcgill.ca

We describe a general constructive approach to characterizing sets of matrices consistent with given approximate solutions to various linear problems. This complements work done by Waldén, Karlson, and Sun (see [3]), who showed that given $A \in \mathscr{C}^{m\times n}$, $b \in \mathscr{C}^m$, and a nonzero approximate solution $y \in \mathscr{C}^n$ to the least squares problem $Ax \approx b$, then with $r \equiv b - Ay$, the set of matrices E for which y is the least squares solution to $(A+E)y \approx b$ is

$$\begin{aligned}\mathscr{E} &\equiv \{E \in \mathscr{C}^{m\times n} : (A+E)^H[(A+E)y - b] = 0\} \\ &= \{-\tilde{w}\tilde{w}^\dagger A + (I - \tilde{w}\tilde{w}^\dagger)[ry^\dagger + Z(I - yy^\dagger)] : \tilde{w} \in \mathscr{C}^m,\ Z \in \mathscr{C}^{m\times n}\}.\end{aligned}$$

The authors of [3] used this result to derive minimal perturbations in A and b for a given approximate solution y; that is, with A, b, y, and r as above, given a positive scalar γ, and defining $\mu \equiv \gamma^2/(1+\gamma^2\|y\|_2^2)$ and the minimum eigenvalue $\lambda_* \equiv \lambda_{min}(AA^H - \mu rr^H)$,

$$\min_{E,f}\{\|[E, \gamma f]\|_F^2 : (A+E)^H[(A+E)y - (b+f)] = 0\} = \mu\|r\|_2^2 + \min\{0,\ \lambda_*\}.$$

We give simple, constructive new proofs of some results in [3], and extend this approach to the data least squares, total least squares, and scaled total least squares problems.

Among other things, such results can be used to derive and analyze minimal backward errors for approximation problems, and develop stopping criteria for solution techniques. For example, Rigal and Gaches derived an optimal stopping criterion for the iterative solution of compatible systems (see [2] and, for example, [1, Thm. 7.1]) by showing that for given $A \in \mathscr{C}^{m\times n}$, $b \in \mathscr{C}^m$, $0 \neq y \in \mathscr{C}^n$, and positive scalars α and β,

$$\min_{\eta,E,f}\{\eta : (A+E)y = b+f,\ \|E\|_2 \le \eta\alpha,\ \|f\|_2 \le \eta\beta\} = \frac{\|b - Ay\|_2}{\beta + \alpha\|y\|_2}.$$

Our matrix characterization approach can be used to derive this and some new results, and might also be useful for deriving minimal backward errors for other problems, such as under-determined linear systems of equations, singular value problems, or eigenproblems.

Bibliography

[1] N.J. Higham. *Accuracy and Stability of Numerical Algorithms, 2nd Edn.*, SIAM, Philadelphia, PA, 2002.

[2] J.L. Rigal and J. Gaches. On the compatibility of a given solution with the data of a linear system. *JACM*, 14:543–548, 1967.

[3] B. Waldén, R. Karlson, and J. Sun. Optimal backward perturbation bounds for the linear least squares problem. *Numerical Linear Algebra with Applications*, 2:271–286, 1995.

About the authors X.-W. Chang is Associate Professor of Computer Science at McGill University. His research interests are in the area of numerical linear algebra and its applications. Currently he is interested in numerical methods for estimation, such as real least squares estimation, integer least squares estimation, robust estimation, partially linear model estimation, etc, and their applications in global navigation satellite systems.

C.C. Paige is Professor Emeritus of Computer Science at McGill University. His main research area is numerical linear algebra and its applications. Currently he is interested in large sparse matrix problems, Krylov subspace methods, parameter estimation, ill-posed problems and rounding error analyses.

D. Titley-Peloquin is a Ph.D. candidate at McGill University's School of Computer Science. He is co-supervised by X.-W. Chang and C.C. Paige.

4.13 On the equivalence between total least squares and maximum likelihood principal component analysis with applications in chemometrics

M. Schuermans, I. Markovsky, S. Van Huffel, KULeuven, ESAT-SCD, Kasteelpark Arenberg 10, B-3001 Leuven-Heverlee, mieke.schuermans,ivan.markovsky,sabine.vanhuffel@esat.kuleuven.be

The Maximum likelihood Principal Component Analysis (MLPCA) method has been devised in chemometrics as a generalization of the well-known PCA method in order to derive consistent estimators in the presence of errors with known error distribution. For similar reasons, the Total Least Squares (TLS) method has been generalized in the field of computational mathematics and engineering to maintain consistency of the parameter estimates in linear models with measurement errors of known distribution. The purpose of this talk is to explore the tight equivalences between MLPCA [2, 3] and element-wise weighted TLS (EW-TLS) [4, 1]. Moreover, an adapted version of the EW-TLS method is derived in order to make it useful for problems in chemometrics. We will present a computationally efficient algorithm and compare this algorithm with the standard EW-TLS algorithm and the MLPCA algorithm in computation time and convergence behaviour on chemical data.

Despite the seemingly different problem formulations of the MLPCA method and the TLS method, it is shown that both methods can be reduced to the same mathematical kernel problem, i.e. finding the closest (in a certain sense) weighted low rank matrix approximation where the weight is derived from the distribution of the errors in the data. Mathematically, we will consider the following weighted low rank matrix approximation problem:

$$\min_{\widehat{D}} \| D - \widehat{D} \|_W \quad \text{s.t. } \operatorname{rank}(\widehat{D}) \leq r, \tag{4.12}$$

with $D \in \mathbb{R}^{m \times n}$, the noisy data matrix, $\operatorname{rank}(D) = k$, $r < k$, $\widehat{\Delta D} = D - \widehat{D}$ the estimated measurement noise, W the covariance matrix of $vec(\widehat{\Delta D})$ where $vec(\widehat{\Delta D})$ stands for the vectorized form of $\widehat{\Delta D}$, i.e., a vector constructed by stacking the consecutive columns of $\widehat{\Delta D}$ in one vector and $\| \cdot \|_W = vec^\top(\cdot) W^{-1} vec(\cdot)$. When the measurement noise is centered, normal and independently and identically distributed, $W = I$, where I is the identity matrix, and the optimal closeness norm is the Frobenius norm, $\| \cdot \|_F$. This is used in the well-known TLS and the PCA methods. Nevertheless, when the measurement errors are not identically distributed the Frobenius norm is no longer optimal and a weighted norm is needed instead.

Different solution approaches, as used in MLPCA and EW-TLS, are discussed. These approaches differ in the representation of the rank constraint $\operatorname{rank}(\widehat{D}) \leq r$ in problem (4.12) and in the applied optimization technique in order to solve problem (4.12).

In the MLPCA approach, the rank constraint $\operatorname{rank}(\widehat{D}) \leq r$ is represented as

$$\widehat{D} = TP^\top,$$

with $T \in \mathbb{R}^{m\times r}$ and $P \in \mathbb{R}^{n\times r}$. So, problem (4.12) can be rewritten as follows:

$$\min_{T}\left(\min_{P,\widehat{D}} vec^{\top}(D-\widehat{D})W^{-1}vec(D-\widehat{D}) \quad \text{s.t. } \widehat{D}=TP^{\top}\right).$$

In the standard EW-TLS approach, the rank constraint is forced by rewriting $\text{rank}(\widehat{D}) \leq r$ as

$$\widehat{D}\begin{bmatrix} \widehat{B} \\ -I_{n-r} \end{bmatrix} = 0, \tag{4.13}$$

where $\widehat{B} \in \mathbb{R}^{r\times(n-r)}$. Moreover, the weighting matrix W is assumed to be block diagonal

$$W = \begin{bmatrix} W_1 & & \\ & \ddots & \\ & & W_m \end{bmatrix},$$

where each block W_i is the covariance matrix of the errors in the i-th row of the data matrix D. So, for the EW-TLS approach, problem (4.12) can be rewritten as

$$\min_{\widehat{B}}\left(\min_{\widehat{D}}\sum_{i=1}^{m}(d_i-\widehat{d_i})W_i^{-1}(d_i-\widehat{d_i})^{\top} \quad \text{s.t. } \widehat{D}\begin{bmatrix} \widehat{B} \\ -I_{n-r} \end{bmatrix} = 0\right), \tag{4.14}$$

with $d_i, \widehat{d_i} \in \mathbb{R}^n$ the i-th row of D and $\widehat{D}$, respectively, and W_i the i-th weighting matrix defined as the covariance matrix of the errors in d_i. Algorithms described in [1, 4], were designed to solve the standard EW-TLS problem (4.14) for the case when $m \geq n$ and when the measurement errors are only row-wise correlated. In chemometrics, however, the data matrix usually has size $m \times n$ with $m \leq n$, e.g., in problems of mixture analysis, curve resolution and data fusion. When the measurement errors are uncorrelated or column-wise correlated, the algorithms presented in [1, 4], can still be applied to the transposed data matrix. For other cases of measurement error correlation, the algorithms need to be optimized by considering the left kernel of $\widehat{D}$, i.e., the following modification of equation (4.13) should be used:

$$\begin{bmatrix} \widehat{B_2}^{\top} & -I_{m-r} \end{bmatrix}\widehat{D} = 0,$$

where $\widehat{B}_2 \in \mathbb{R}^{r\times(m-r)}$. By means of experiments on chemical data, we will show that the EW-TLS method certainly has potential for problems when the data matrix has size $m \times n$ with $m \geq n$ and only row-wise correlated measurement errors. It will also become clear that the standard EW-TLS approach is not the right method of choice for the case when $m \leq n$ and only row-wise correlated measurements and that an adapted version of the EW-TLS approach is needed for handling this case of row-wise correlated measurement errors in data sets where $m \leq n$. An algorithm will be derived to solve the following adapted version of the EW-TLS problem:

$$\min_{\widehat{B}_2}\left(\min_{\widehat{D}}\sum_{i=1}^{m}(d_i-\widehat{d_i})W_i^{-1}(d_i-\widehat{d_i})^{\top} \quad \text{s.t. } \begin{bmatrix} \widehat{B_2}^{\top} & -I_{m-r} \end{bmatrix}\widehat{D} = 0\right),$$

with $d_i, \widehat{d_i} \in \mathbb{R}^n$ the i-th row of D and $\widehat{D}$, respectively, and W_i the i-th weighting matrix defined as the covariance matrix of the errors in d_i. The developed algorithm will be compared with the standard EW-TLS algorithm and the MLPCA algorithm in computation time and convergence behaviour on chemical data.

Bibliography

[1] I. Markovsky, M. L. Rastello, A. Premoli, A. Kukush and S. Van Huffel S. The element-wise weighted total least squares problem. *Journal of Computational Statistics and Data Analysis*, 50(1): 181 – 209, 2005.

[2] P. D. Wentzell, D. T. Andrews, D. C. Hamilton, K. Faber and B. R. Kowalski. Maximum likelihood principal component analysis. *Journal of Chemometrics*, 11: 339 – 366, 1997.

[3] P. D. Wentzell and M. T. Lohnes. Maximum likelihood principal component analysis with correlated measurement errors: theoretical and practical considerations. *Chemometrics and Intelligent Laboratory Systems*, 45: 65 – 85, 1999.

[4] A. Premoli and M. L. Rastello. The parametric quadratic form method for solving problems with element-wise weighting. In S. Van Huffel and P. Lemmerling, editors, Total least squares and errors-in-variables modeling: Analysis, algorithms and applications, Kluwer: 67 – 76, 2002.

About the author Prof. dr. Sabine Van Huffel is a full professor, dr. Ivan Markovsky is a postdoctoral researcher and Mieke Schuermans is a research assistant at the Katholieke Universiteit Leuven, Belgium.

4.14 Hyperaccuracy for geometric fitting

Kenichi Kanatani, Department of Computer Science, Okayama University, 700-8530 Japan
kanatani@suri.it.okayama-u.ac.jp

Geometric Fitting *Geometric fitting* is to fit to noisy data a geometric model (a curve, a surface, or a relationship in high dimensions) expressed as an *implicit* equation that noise free data should satisfy [4, 7]. Its major differences from the traditional parameter estimation problem are:

- Unlike traditional statistical problems, there is no *explicit* model which explains observed data in terms of deterministic mechanisms and random errors. All descriptions are *implicit*.
- There are no inputs or outputs. No such concepts as causes and effects exist. No such distinctions as ordinates and abscissas exist.
- The underlying data space is homogeneous and isotropic with no inherent coordinate systems. Hence, the estimation process and the results should be *invariant* to changes of the coordinate system with respect to which the data are described.
- In many cases, the data are geometrically *constrained*. Typically, they are points on curves, surfaces, and hypersurfaces inherent in the data (e.g., unit vectors and matrices of determinant 0). Often, the parameters to be estimated are also similarly constrained. Hence, the *Gaussian distribution*, the most fundamental noise modeling, does not exist in its strict sense in such constrained spaces.

This type of problem plays a central role in computer vision applications. While in the traditional domain of statistics, the *total least-squares method* and the *errors-in-variable model* are rather abnormal concepts, to which attention is paid as a special research theme, the above mentioned properties are the *norm* in many computer vision problems.

Performance Evaluation Due to the above mentioned characteristics, performance evaluation of geometric fitting algorithms is very different from that in the traditional domain of statistics [4, 6, 7].

Traditional estimation It is customary to evaluate estimation methods by investigating the asymptotic performance as the number of data increases, and the *consistency* is one of the main concerns of statisticians. This is based on the tenet of statistics that random disturbances can be overcome by sampling many data, invoking the law of large numbers and the central limit theorem. This is reasonable in practice, too, since methods whose performance grows rapidly as the number of data increases are preferable; such methods can reach admissible accuracy with a fewer number of data than other methods.

Geometric fitting The data in computer vision applications are usually generated by a computer using image processing operations. They may contain pixel-level or subpixel-level errors, because image processing operations, often heuristically designed, are executed on digitized images and the results are not expected to be strictly correct. In such small noise domains, it is reasonable to focus on the asymptotic performance as the noise level approaches 0, since methods whose performance grows rapidly as the noise decreases can tolerate higher uncertainty in the data than others for admissible accuracy.

Hyperaccuracy In many computer vision applications, the constraint can be written as a linear form in the parameters to be estimated by (nonlinearly) changing variables and embedding them in a high dimensional space. For such linearized constraints, we can easily compute the ML (maximum likelihood) estimator. The best known are:

- The *renormalization* method of Kanatani [3, 4, 7].
- The *HEIV* (*heteroscedastic errors in variable*) method of Leedan and Meer [9].
- The *FNS* (*fundamental numerical scheme*) of Chojnacki et al. [2].

We call those estimation methods who have accuracy equivalent to ML *high accuracy* methods and those whose accuracy are lower *low accuracy methods*, typical examples being least squares and the method of Taubin et al. [10, 11]. In contrast, We call those methods who perform better than ML *hyperaccuracy methods*.

We demonstrate the existence of a hyperaccurate method [8]. Since the ML estimator achieves a theoretical accuracy bound, called the *KCR lower bound* [1, 5, 7], except for high order noise terms, the difference of hyperaccurate solution from the ML estimator ought to be in higher order noise terms and hence is necessarily very small. Nevertheless, the underlying principle for obtaining such a method is theoretically very important, illuminating the relationship between geometric ML and the KCR lower bound.

Bibliography

[1] N. Chernov and C. Lesort, Statistical efficiency of curve fitting algorithms, *Comput. Stat. Data Anal.*, **47**-4 (2004), 713–728.

[2] W. Chojnacki, M. J. Brooks, A. van den Hengel and D. Gawley, On the fitting of surfaces to data with covariances, *IEEE Trans. Patt. Anal. Mach. Intell.*, **22**-11 (2000), 1294–1303.

[3] K. Kanatani, Renormalization for unbiased estimation, *Proc. 4th Int. Conf. Comput. Vision (ICCV'93)*, May 1993, Berlin, Germany, pp. 599–606.

[4] K. Kanatani, *Statistical Optimization for Geometric Computation: Theory and Practice*, Elsevier, Amsterdam, The Netherlands, 1996; Dover, New York, U.S.A., 2005.

[5] K. Kanatani, Cramer-Rao lower bounds for curve fitting, *Graphical Models Image Processing*, **60**-2 (1998), 93–99.

[6] K. Kanatani, Uncertainty modeling and model selection for geometric inference, *IEEE Trans. Patt. Anal. Machine Intell.*, **26**-10 (2004), 1307–1319.

[7] K. Kanatani, Further improving geometric fitting, *Proc. 5th Int. Conf. 3-D Digital Imaging and Modeling (3DIM2005)*, June 2005, Ottawa, Canada, pp. 2–13.

[8] K. Kanatani, Ellipse fitting with hyperaccuracy, *Proc. 9th Euro. Conf. Comput. Vision (ECCV'06)*, May 2006, Graz, Austria, Vol. 1, pp. 484–495.

[9] Y. Leedan and P. Meer, Heteroscedastic regression in computer vision: Problems with bilinear constraint, *Int. J. Comput. Vision.*, **37**-2 (2000), 127–150.

[10] G. Taubin, Estimation of planar curves, surfaces, and non-planar space curves defined by implicit equations with applications to edge and rage image segmentation, *IEEE Trans. Patt. Anal. Mach. Intell.*, **13**-11 (1991), 1115–1138.

[11] G. Taubin, F. Cukierman, S. Sullivan, J. Ponce and D. J. Kriegman, Parameterized families of polynomials for bounded algebraic curve and surface fitting, *IEEE Trans. Patt. Anal. Mach. Intell.*, **16**-3 (1994), 287–303.

Kenichi Kanatani received his M.S. and Ph.D. in applied mathematics from the University of Tokyo in 1974 and 1979, respectively. After serving as Professor of computer science at Gunma University, Gunma, Japan, he is currently Professor of computer science at Okayama University, Okayama, Japan. He was a visiting researcher at the University of Maryland, U.S.A., the University of Copenhagen, Denmark, the University of Oxford, U.K., and INIRA at Rhone Alpes, France. He is the author of *Group-Theoretical Methods in Image Understanding* (Springer, 1990), *Geometric Computation for Machine Vision* (Oxford, 1993) and *Statistical Optimization for Geometric Computation: Theory and Practice* (Elsevier, 1996; Dover, 2005). He is an IEEE Fellow.

4.15 Low-rank approximation and its applications for data fitting

Ivan Markovsky, ESAT-SISTA, K.U.Leuven, `ivan.markovsky@esat.kuleuven.be`

Fitting models to data is an ubiquitous problem with long history and many ramifications, depending on the model class and the fitting criterion being used. A well known special case is the line fitting problem—find a line $\mathscr{B} \subset \mathbb{R}^2$ passing through the origin that best matches a set of given points $\mathscr{D}$,

$$\mathscr{D} := \{ d_1, \ldots, d_N \}, \qquad d_i =: \operatorname{col}(a_i, b_i).$$

The classical line fitting solution is given by the least squares method. In order to apply the least squares method, however, we have to represent the model $\mathscr{B}$ by equations. Two possible representations for a line in $\mathbb{R}^2$ passing through the origin are

$$\mathscr{B}_1(x) = \{ \operatorname{col}(a,b) \mid ax = b \} \quad \text{and} \quad \mathscr{B}_2(y) = \{ \operatorname{col}(a,b) \mid a = by \},$$

where x and y are parameter of the model in the two representations. The corresponding least squares problems are

$$\operatorname{col}(a_1, \ldots, a_N)x = \operatorname{col}(b_1, \ldots, b_N) \tag{SYS1}$$

and

$$\operatorname{col}(a_1, \ldots, a_N) = \operatorname{col}(b_1, \ldots, b_N)y. \tag{SYS2}$$

In general, the least squares fit is representation dependent, i.e., $\mathscr{B}_1(x_{\text{ls}}) \neq \mathscr{B}_2(y_{\text{ls}})$, where x_{ls} and y_{ls} are the least squares solutions of (SYS1) and (SYS2), respectively. From the data fitting point of view, this is an undesirable feature of the least squares method: the representation of the model is not part of the problem formulation and is therefore arbitrary. We would prefer a fitting method that does not depend on the choice of the representation.

An almost representation invariant line fitting method is the classical total least squares method. Generically $\mathscr{B}_1(x_{\text{tls}}) = \mathscr{B}_2(y_{\text{tls}})$, where x_{tls} and y_{ls} are the total least squares solutions of (SYS1) and (SYS2). There are, however, non-generic cases when x_{tls}, y_{ls}, or both fail to exist. For example, x_{tls} does not exist whenever the optimal fitting line $\mathscr{B}_{\text{tls}}$ is vertical. Note that in this case y_{tls} do exist.

Non-generic TLS problems occure because one of the variable (b in (SYS1) and a in (SYS2)) is required to be a function of the other variable. In system theory such representations are called input/output: the free variable is called input, and the bound (by the input and the model) variable is called output. By using (SYS1) and (SYS2), we fix an input/output partition of the variables prior to modeling the data. If the optimal model $\mathscr{B}_{\text{tls}}$ does not allow such an a priori fixed input/output partition, the classical TLS problem has no solution. In the context of data fitting, it is better to *deduce* the input/output partition from the model instead of *assuming* it in advance.

Representations of a line passing through the origin that does not force a given input/output partition are the image $\mathscr{B}(P) = \text{image}(P)$ and kernel $\mathscr{B}(R) = \ker(R)$ representations, where $P \in \mathbb{R}^{2\times 1}$ and $R \in \mathbb{R}^{1\times 2}$ are parameters of the model. The line fitting problem for the image and kernel representations leads to a low-rank approximation problem: approximate the data matrix $D := \begin{bmatrix} d_1 & \cdots & d_N \end{bmatrix}$ by a rank-one matrix. In fact, low-rank approximation is a representation free concept applying to general multivariable static and dynamic modeling problems.

The main difference between the static and the dynamic case is in the structure of the data matrix. In the context of linear time-invariant dynamic systems, the data matrix has structure, typically Hankel or Toeplitz, and the approximating matrix is searched in the class of matrices with the same structure. Such structured low-rank approximation problems are in general harder to solve than their unstructured counterparts. In this talk, we list applications of the structured low-rank approximation problem in system theory and signal processing, outline numerical solution methods, and show links to other problems.

Bibliography

[1] I. Markovsky and S. Van Huffel. Overview of total least squares methods. Technical Report 05–34, Dept. EE, K.U.Leuven, 2005.

About the author Ivan Markovsky obtained MS degree in Control and Systems Engineering in 1998 from the Technical University of Sofia and PhD degree in Electrical Engineering in 2005 from K.U.Leuven. Since February 2005 he is a postdoctoral research in the Electrical Engineering department of K.U.Leuven. His current research work is focused on identification methods in the behavioral setting and errors-in-variables estimation problems. He is an author of the book "Exact and Approximate Modeling of Linear Systems: A Behavioral Approach" (SIAM, Philadelphia, 2006).

4.16 Robust regression and ℓ_1 approximations for Toeplitz problems

Nicola Mastronardi , Istituto per le Applicazioni del Calcolo "M. Picone", Consiglio Nazionale delle Ricerche, Bari, Italy, n.mastronardi@ba.iac.cnr.it
Dianne P. O'Leary, Department of Computer Science and Institute for Advanced Computer Studies, University of Maryland, College Park, Maryland, USA, oleary@cs.umd.edu.

Consider the approximation problem

$$Ax \approx b$$

where $A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^m$ are given and $x \in \mathbb{R}^n$ is to be determined. We define the residual

$$r = b - Ax.$$

The usual approach to the problem is least squares, in which we minimize the 2-norm of the residual over all choices of x. This produces the minimum variance unbiased estimator of the solution when the errors in the observation b are independent and normally distributed with mean 0 and constant variance.

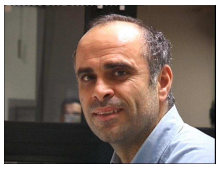
It is well known, however, that the least squares solution is not robust if outliers occur, i.e., if some of the components of b are contaminated by large error. In this case, alternate approaches have been proposed which judge the size of the residual in a way that is less sensitive to these components. These include the Huber M-function, the Talwar function, the logistic function, the Fair function, and the ℓ_1 norm [2]. In this work we consider how the solution to these problems can be computed efficiently, in particular when the matrix A has small displacement rank [1]. Matrices with small displacement rank include matrices that are Toeplitz, block-Toeplitz, block-Toeplitz with Toeplitz blocks

(BTTB), Toeplitz plus Hankel, and a variety of other forms. For exposition, we will focus on Toeplitz matrices, but the ideas apply to all matrices with small displacement rank. We also show how to compute the solution efficiently when we include a regularization term in case the matrix A is ill-conditioned or rank-deficient.

Bibliography

[1] T. Kailath and A.-H. Sayed. Displacement structure: theory and applications. *SIAM Review*, 37(3):297–386, 1995.

[2] D.-P. O'Leary, Robust regression computation using iteratively reweighted least squares. *SIAM J. Matrix Analysis Appl.*, 11:466–480, 1990.



About the author Nicola Mastronardi was born in Putignano, Italy, in 1962. He received the Ph.D. in Applied Sciences at the Department of Electrical Engineering, Katholieke Universiteit Leuven in May 2001. He is currently a Senior Researcher at Istituto per le Applicazioni del Calcolo "M. Picone", Consiglio Nazionale delle Ricerche, Bari, Italy. Nicola Mastronardi is an associate editor of SIAM Journal of Matrix Analysis and Applications.

4.17 A structured total least squares algorithm for approximate greatest common divisors of multivariate polynomials

Erich Kaltofen[1], Zhengfeng Yang[2] and Lihong Zhi[2],
[1]Dept. of Mathematics, Massachusetts Institute of Technology, Cambridge, Massachusetts, 02139-4307, USA and Dept. of Mathematics, North Carolina State University, Raleigh, North Carolina 27695-8205, USA,
[2]Key Laboratory of Mathematics Mechanization, Academy of Mathematical and System Sciences, Beijing 100080, China,
kaltofen@math.{mit, ncsu}.edu, {zyang, lzhi}@mmrc.iss.ac.cn

The approximate GCD problem has as inputs polynomials $f_1,\ldots,f_s$ in the variables $y_1,\ldots,y_r$ with real or complex coefficients. Let $d_i = \mathrm{tdeg}(f_i)$ be the total degree of f_i and $k \le d_i$ for all i with $1 \le i \le s$. The outputs are polynomials g and $f_1^*,\ldots,f_s^*$ with real or complex coefficients such that $\mathrm{tdeg}(g) \ge k$ and $\mathrm{tdeg}(\widetilde{f}_i) \le d_i$ for $\widetilde{f}_i = g f_i^*$, with $1 \le i \le s$, and such that $\|\triangle f_1\| + \cdots + \|\triangle f_s\|$ is minimized, where $\triangle f_i = \widetilde{f}_i - f_i$ for $1 \le i \le s$. Here $\widetilde{f}_i$ is the approximation to the input polynomials and $\triangle f_i$ is the applied change. We use Euclidean norm on the coefficient vector of multivariate polynomials.

Using a linear algebra formulation, we can apply a structure-preserving total least squares approach to our approximate GCD problem [2]. It has been proved that $\mathrm{tdeg}(\gcd(f_1,\ldots,f_s)) \ge k$ if and only if $S_k(f_1,\ldots,f_s)$ has rank deficiency at least one. Here the matrix $S_k(f_1,\ldots,f_s)$ is essentially a multi-polynomial generalized Sylvester matrix. The row and column dimensions of S_k are $\sum_{i=2}^{s}\binom{d_1+d_i-k+r}{r}$ and $\sum_{i=1}^{s}\binom{d_i-k+r}{r}$ respectively. Let $S_k(\zeta) = [A_1(\zeta) \mid b(\zeta) \mid A_2(\zeta)]$ and let $A(\zeta) = [A_1(\zeta) \mid A_2(\zeta)]$. The matrices $S_k(f_1,\ldots,f_s)$ and A and the vector b are parameterized via the vector ζ, which contains the coefficients of $f_1,\ldots,f_s$. The dimension of ζ is ν which is equal to $\sum_{i=1}^{s}\binom{d_i+r}{r}$. We wish to solve the two structure-preserving total least norm problems

$$\min_{\mathbf{z}\in\mathbb{R}^{\nu}} \|\mathbf{z}\| \text{ or } \min_{\mathbf{z}\in\mathbb{C}^{\nu}} \|\mathbf{z}\| \quad \text{with} \quad A(\mathbf{c}+\mathbf{z})\mathbf{x} = b(\mathbf{c}+\mathbf{z}) \text{ for some vector } \mathbf{x}, \tag{4.15}$$

where $\mathbf{c}$ is fixed to the initial coefficient vector. We choose the column corresponding to the absolutely largest component in the first singular vector of S_k [1].

We extend the structured total least squares (STLS) method in [3] to solve the minimization problems (4.15). If the optimization problem is over the complex numbers, real and complex parts need to be separated first. Suppose $\mathbf{z} = \mathbf{z}_R + \boldsymbol{i}\mathbf{z}_I$, $\mathbf{x} = \mathbf{x}_R + \boldsymbol{i}\mathbf{x}_I$ and $\lambda = \lambda_R + \boldsymbol{i}\lambda_I$. The problem (4.15) can be transformed into unconstrained optimization

by the Lagrangian:

$$\begin{aligned} L(\mathbf{z},\mathbf{x},\lambda) &= \tfrac{1}{2}\,\mathbf{z}^H\mathbf{z} - \lambda_R^{Tr}(b(\mathbf{c}_R+\mathbf{z}_R) - A(\mathbf{c}_R+\mathbf{z}_R)\mathbf{x}_R + A(\mathbf{c}_I+\mathbf{z}_I)\mathbf{x}_I) \\ &\quad - \lambda_I^{Tr}(b(\mathbf{c}_I+\mathbf{z}_I) - A(\mathbf{c}_R+\mathbf{z}_R)\mathbf{x}_I - A(\mathbf{c}_I+\mathbf{z}_I)\mathbf{x}_R) \\ &= \tfrac{1}{2}\,\mathbf{z}^H\mathbf{z} + \lambda_R^{Tr}\,\mathbf{r}_R(\mathbf{z},\mathbf{x}) + \lambda_I^{Tr}\,\mathbf{r}_I(\mathbf{z},\mathbf{x}). \end{aligned}$$

Before we apply the Newton method, two Sylvester-structured matrices H and Y need to be constructed, such that $H(\mathbf{x})\mathbf{z} = S_k(\mathbf{z})\mathbf{x}$ and $Y(\mathbf{x})\mathbf{z} = A(\mathbf{z})\mathbf{x}$. Applying the Newton method on the Lagrangian L yields:

$$\begin{bmatrix} W & J^{Tr} \\ J & \mathbf{0} \end{bmatrix} \begin{bmatrix} \triangle\mathbf{z}_R \\ \triangle\mathbf{z}_I \\ \triangle\mathbf{x}_R \\ \triangle\mathbf{x}_I \\ \triangle\lambda_R \\ \triangle\lambda_I \end{bmatrix} = - \begin{bmatrix} \mathbf{g} + J^{Tr} \begin{bmatrix} \lambda_R \\ \lambda_I \end{bmatrix} \\ \mathbf{r}_R(\mathbf{z},\mathbf{x}) \\ \mathbf{r}_I(\mathbf{z},\mathbf{x}) \end{bmatrix}, \quad \text{where } \mathbf{g} = W \begin{bmatrix} \mathbf{z}_R \\ \mathbf{z}_I \\ \mathbf{x}_R \\ \mathbf{x}_I \end{bmatrix} \tag{4.16}$$

and where

$$J = \begin{bmatrix} H(\mathbf{x}_R) & -Y(\mathbf{x}_I) & A(\mathbf{c}_R+\mathbf{z}_R) & -A(\mathbf{c}_I+\mathbf{z}_I) \\ Y(\mathbf{x}_I) & H(\mathbf{x}_R) & A(\mathbf{c}_I+\mathbf{z}_I) & A(\mathbf{c}_R+\mathbf{z}_R) \end{bmatrix},$$
$$W = \begin{bmatrix} I_{t_1\times t_1} & 0_{t_1\times t_2} \\ 0_{t_2\times t_1} & 0_{t_2\times t_2} \end{bmatrix}, t_1 = 2\nu, t_2 = -2 + 2\sum_{i=1}^{s} \tbinom{d_i-k+r}{r}.$$

We compute the solution of (4.16), and update $\mathbf{z} = \mathbf{z} + \triangle\mathbf{z}, \mathbf{x} = \mathbf{x} + \triangle\mathbf{x}, \lambda = \lambda + \triangle\lambda$ until $\|\triangle\mathbf{z}\|_2 \leq \text{tol}$, for a given tolerance.

We have implemented the above method in Maple and compared it with the STLN-based algorithm in [2]. Both algorithms can be applied to solve the approximate GCD problem and achieve globally optimal backward errors.

Bibliography

[1] Erich Kaltofen, Zhengfeng Yang, and Lihong Zhi. Structured low rank approximation of a generalized Sylvester matrix. Manuscript, 2005. 14 pages. Abstract at the Asian Symposium on Computer Mathematics ASCM 2005, 4 pages.

[2] Erich Kaltofen, Zhengfeng Yang, and Lihong Zhi. Approximate greatest common divisors of several polynomials with linearly constrained coefficients and singular polynomials. In Jean-Guillaume Dumas, editor, *ISSAC'06 Proc. 2006 Internat. Symp. Symbolic Algebraic Comput.*, New York, N. Y., 2006. ACM Press. To appear.

[3] Phillipe Lemmerling, Nicola Mastronardi, and Sabine Van Huffel. Fast algorithm for solving the Hankel/Toeplitz Structured Total Least Squares problem. *Numerical Algorithms*, 23:371–392, 2000.

Erich Kaltofen received both his M.S. degree in Computer Science in 1979 and his Ph.D. degree in Computer Science in 1982 from Rensselaer Polytechnic Institute. He was an Assistant Professor of Computer Science at the University of Toronto and an Assistant, Associate, and full Professor at Rensselaer Polytechnic Institute. Since 1996 he is a Professor of Mathematics at North Carolina State University. Kaltofen was the Chair of ACM's Special Interest Group on Symbolic & Algebraic Manipulation 1993–95. He has edited the Computer Algebra Handbook in 2002. Kaltofen's current interests in the symbolic computation discipline are hybrid symbolic/numeric algorithms, efficient algorithms for exact linear and polynomial algebra, generic programming techniques for algorithm implementation, and algebraic complexity theory.

Lihong Zhi received the B.S. degree in Mathematics from Peking University, Beijing, China, and the Dr.Sci. degree in Computer Algebra from the Institute of Systems Science, Academia Sinica, Beijing, China. She is an Associate Professor with the Institute of Systems Science, Academia Sinica. From 1998 to 2001, she was an Assistant

Professor with the Department of Computer Science, Ehime University, Matsuyama, Japan. From 2001 to 2002, she did post-doctoral research at the Ontario Research Center for Computer Algebra, University of Western Ontario, London, Canada. Zhi's current research interests are fast and reliable algorithms and their implementation for hybrid numeric/symbolic computation.

Zhengfeng Yang received the B. S. degree in Applied Mathematics from Changchun Normal College, China, in July 2001. Since then, He has become a Ph.D. student in Applied Mathematics in Key Lab of Mathematics Mechanization, Academy of Mathematics and System Science, China. At the 2004 International Symposium for Symbolic and Algebraic Computation Yang has received an ACM Distinguished Student Author Award. The subject of his Ph.D. thesis are hybrid numeric/symbolic algorithms and software for approximate polynomial GCD and factorization.

4.18 Structured matrix methods for the computation of a rank reduced Sylvester matrix

Joab R. Winkler, John D. Allan, Department of Computer Science, The University of Sheffield, United Kingdom
j.winkler@dcs.shef.ac.uk, j.allan@dcs.shef.ac.uk

The Sylvester resultant matrix $S(p,q)$ is a structured matrix that can be used to determine if two polynomials $p = p(y)$ and $q = q(y)$ are, or are not, coprime, and if they are not coprime, it allows their greatest common divisor (GCD) to be computed. In particular, the rank loss of $S(p,q)$ is equal to the degree of the GCD of $p(y)$ and $q(y)$, and the GCD can be obtained by reducing $S(p,q)$ to row echelon form.

The computation of the GCD of two polynomials arises in many applications, including computer graphics, control theory and geometric modelling. Experimental errors imply that the data consists of noisy realisations of the exact polynomials $p(y)$ and $q(y)$, and thus even if $p(y)$ and $q(y)$ have a non-constant GCD, their noisy realisations, $f(y)$ and $g(y)$ respectively, are coprime. It is therefore only possible to compute an *approximate GCD*, that is, a GCD of the polynomials $\tilde{f}(y)$ and $\tilde{g}(y)$ that are obtained by small perturbations of $f(y)$ and $g(y)$. Different perturbations of $f(y)$ and $g(y)$ yield different approximate GCDs, all of which are legitimate if the magnitude of these perturbations is smaller than the noise in the coefficients. It follows that $\tilde{f}(y)$ and $\tilde{g}(y)$ have a non-constant GCD, and thus the Sylvester resultant matrix $S(\tilde{f},\tilde{g})$ is a low rank approximation of the Sylvester matrix $S(f,g)$.

In this paper, the method of structured total least norm is used to compute the rank reduced Sylvester resultant matrix $S(\tilde{f},\tilde{g})$, given inexact polynomials $f(y)$ and $g(y)$ [3, 4]. Although this problem has been considered previously [1, 2], it is shown that there exist several issues that have not been addressed, and that these issues have a considerable effect on the degree of the computed approximate GCD.

Let the inexact polynomials $f(y)$ and $g(y)$ be given by

$$f(y) = \sum_{i=0}^{m} a_i y^{m-i} \qquad \text{and} \qquad g(y) = \sum_{i=0}^{n} b_i y^{n-i}, \qquad a_m, b_n \neq 0,$$

and let z_i be the perturbation of the coefficient a_i of $f(y)$, and z_{m+1+i} be the perturbation of the coefficient b_i of $g(y)$, that are required to perturb $S(f,g)$ into $S(\tilde{f},\tilde{g})$, which is a structured low rank approximation of $S(f,g)$. It is therefore required to minimise $\|z\|_2$, where

$$z = \begin{bmatrix} z_0 & \cdots & z_m & z_{m+1} & \cdots & z_{m+n+1} \end{bmatrix},$$

subject to the constraint that $S(\tilde{f},\tilde{g}) = S(f,g) + B(z)$, where the error matrix $B(z)$ has the same structure as $S(f,g)$, and $S(\tilde{f},\tilde{g})$ is a Sylvester resultant matrix of lower rank than $S(f,g)$.

A sequence of matrices A_k and a sequence of vectors c_k, $k = 1,\ldots,\min(m,n)$, are formed from the Sylvester matrix $S(f,g)$ of the inexact polynomials $f(y)$ and $g(y)$, and a sequence of matrices E_k and a sequence of vectors h_k, $k = 1,\ldots,\min(m,n)$, are formed from the error matrix $B(z)$. It is shown in [1] that the computation of $S(\tilde{f},\tilde{g})$ requires the solution of the least squares equality (LSE) problem,

$$\min_{\tilde{r}=0;z,x} \left\| \begin{matrix} \tilde{r} \\ z \end{matrix} \right\|_2, \qquad \tilde{r} = \tilde{r}(z,x) = (c_k + h_k) - (A_k + E_k)x, \qquad h_k = h_k(z), \qquad E_k = E_k(z),$$

for several values of k, until an acceptable solution is obtained. This optimisation yields the corrected polynomials $\tilde{f}(y)$ and $\tilde{g}(y)$ that have a non-constant GCD, and this enables the Sylvester matrix $S(\tilde{f},\tilde{g})$ to be constructed.

The GCD of $f(y)$ and $g(y)$ is equal (up to a scalar multiplier) to the GCD of $f(y)$ and $\alpha g(y)$, where α is an arbitrary constant, and it will be shown that α has a significant effect on the computed results. In particular, it will be shown that an incorrect value of α leads to unsatisfactory numerical answers, and methods for the determination of its optimal value will be considered. It will also be shown that a termination criterion of the optimisation algorithm that is based on a small normalised residual may lead to incorrect results, and that it is also necessary to monitor the singular values of $S(\tilde{f},\tilde{g})$ in order to achieve good results. Several non-trivial examples will be used to illustrate the importance of α, and the effectiveness of a termination criterion that is based on the normalised residual and the singular values of $S(\tilde{f},\tilde{g})$.

The dependence of the computed solution on the value of α has implications for the method that is used for the solution of the LSE problem. In particular, this problem is usually solved by the penalty method (method of weights), which requires that the value of the weight be set, but its value is defined heuristically, that is, it is independent of the data (the coefficients of the polynomials). As noted above, the value of the parameter α is crucial to the success or failure of the computed solution, and thus the presence of a parameter that is defined heuristically is not satisfactory. The *QR* decomposition, which does not suffer from this disadvantage, is therefore used to solve the problem described in this abstract.

Bibliography

[1] E. Kaltofen, Z. Yang, and L. Zhi. Structured low rank approximation of a Sylvester matrix, 2005. Preprint.

[2] B. Li, Z. Yang, and L. Zhi. Fast low rank approximation of a Sylvester matrix by structured total least norm. *J. Japan Soc. Symbolic and Algebraic Comp.*, 11:165–174, 2005.

[3] J. Ben Rosen, H. Park, and J. Glick. Total least norm formulation and solution for structured problems. *SIAM J. Mat. Anal. Appl.*, 17(1):110–128, 1996.

[4] J. Ben Rosen, H. Park, and J. Glick. Structured total least norm for nonlinear problems. *SIAM J. Mat. Anal. Appl.*, 20(1):14–30, 1998.

Joab Winkler obtained his undergraduate and doctorate degrees from Imperial College London and University College London, respectively. He worked for several years in industry before returning to academia, and he is now a Senior Lecturer in The Department of Computer Science at The University of Sheffield. His main research interests are algebraic and numerical problems in geometric modelling and computer aided geometric design, but he is also interested in the application of computational linear algebra to, for example, bioinformatics and natural language processing.

John Allan obtained his undergraduate degree in computer science from The University of Sheffield in 2004, where he is currently studying for a PhD. He is conducting research into accurate and computationally robust methods for computing high order multiple roots of polynomials that are expressed in the Bernstein basis.

4.19 Errors-in-variables methods in system identification

Torsten Söderström, Department of Information Technology, Uppsala University, `ts@it.uu.se`

The lecture gives a survey of errors-in-variables methods in system identification. Background and motivation are given, and examples illustrate why the identification problem can be difficult. Under general weak assumptions, the systems are not identifiable, but can be parameterized using one degree of freedom.

Examples where identifiability is achieved under additional assumptions are also provided. Such examples include modeling the noise-free input and the measurement noises as ARMA processes. Another possibility is to use multiple experiments, where some conditions on the different experiments have to be imposed.

It will be described how an Cramer-Rao lower bound for the parameter estimates can be computed, and how different estimators may be classified.

A number of approaches for parameter estimation of errors-in-variables models are presented. The underlying assumptions and principles for each approach are highlighted. Approaches covered include the instrumental variable estimator (based on Yule-Walker type of equations, or more sophisticated versions); various bias-compensating methods, where the linear least squares normal equations are complemented with a few more equations to handle the noise contributions; the so called Frisch scheme, applied for identifying a dynamical system; total least squares approaches; prediction error and maximum likelihood methods; and methods designed for using periodic data.

The lecture is primarily based on [1].

Bibliography

[1] T. Söderström. Errors-in-variables methods in system identification. Plenary paper. *14th IFAC Symposium on System Identification*, Newcastle, Australia, March 29-31, 2006.

Torsten Söderström was born in Malmö, Sweden, in 1945. He received his MSc degree ("civilingenjör") in engineering physics in 1969 and his PhD in automatic control in 1973, both from Lund Institute of Technology, Lund, Sweden. He is a Fellow of IEEE.

During the period 1967–1974 he held various teaching positions at the Lund Institute of Technology. Since 1974, he has been at Uppsala University, Uppsala, Sweden, where he is a professor of automatic control. During 1975-1998 he was the head of the Systems and Control Group, which now is a part of Department of Information Technology.

Dr Söderström is the author or co-author of many technical papers. His main research interests are in the fields of system identification, signal processing, and control of mechanical systems. He is the author or co-author of four books: "Theory and Practice of Recursive Identification", MIT Press, 1983 (with L Ljung), "The Instrumental Variable Methods for System Identification", Springer Verlag, 1983 (with P Stoica), "System Identification", Prentice–Hall, 1989 (with P Stoica) and "Discrete-Time Stochastic Systems", Prentice–Hall, 1994; 2nd edition, Springer-Verlag, 2002. In 1981 he was given (along with his co-authors) an Automatica Paper Prize Award. He is an Automatica editor for the area of System Parameter Estimation since 1992.

He has held a number of positions within IFAC, International Federation of Automatic Control, during 1993–2002. He is currently a council member of EUCA, European Union Control Association.

4.20 Some issues on errors-in-variables identification

Roberto Guidorzi, Roberto Diversi and Umberto Soverini, Department of Electronics, Computer Science and Systems, University of Bologna {rguidorzi,rdiversi,usoverini}@deis.unibo.it

Topics of the paper.

The EIV context The Errors-in-Variables context is a challenging environment well known from many years that has seen an increasing amount of research and, consequently, of new results, only in relatively recent times. One of the appealing features of EIV models consists in their intrinsic capability of describing real processes and in relying only on limited sets of *a–priori* assumptions [1, 2]. These features suggest the use of EIV models in all applications like, for instance, diagnosis, where the interest is focused on a realistic description of a process more than on other features like prediction.

The Frisch scheme The scheme proposed by the Nobel prize Ragnar Frisch in 1934 [3] is an interesting compromise between the great generality of the EIV environment and the possibility of real applications. Moreover, the Frisch

scheme encompasses some other important schemes like Least Squares and the Eigenvector Method and plays, consequently, a role of paramount importance also from a conceptual point of view. The compatibility of the Frisch scheme with a whole family of solutions has diverted the attention towards simpler schemes leading to single solutions.

Loci of solutions and their properties in the algebraic case Any analysis of the Frisch scheme cannot ignore the existence of two separate loci of solutions, one in the parameter space and the other in the space of the variances of the noise affecting the considered variables and, even more important, the maps between these loci. Some fundamental results [4, 5] describe these maps as well as the shape of the loci in the parameter space under specific conditions (the inverse of the covariance matrix of the noisy data must be Frobenius–like) [6]. Unfortunately the locus of solutions in the parameter space can be easily defined only when the data are compatible with a single linear relation; in all other cases the reference to the parameter space does not lead to significant results. The investigation of the properties of the locus of solutions in the noise space has offered the key for a deeper analysis that shows how this locus (a convex hypersurface lying in the first orthant), differently from what happens in the parameter space, does not degenerate in any case and enjoys some other important properties [7, 8].

The maximum corank problem One of the problems considered of great importance in the econometric field consists in determining the maximal number of linear relations compatible with a given set of noisy data. The importance attributed to this problem is due to the fact that econometricians consider the solution of this problem as linked to the extraction of the whole information contained in the data [9]. The solution of this problem in the context of the Frisch scheme has been possible only making reference to the properties of the locus of solutions in the noise space [10]; other approaches have led to determine an upper bound to this number [11].

Relations between algebraic and dynamic contexts When the data are generated by a dynamic system and the Frisch context is used for its identification, it is necessary to consider the properties of the loci of solutions under the constraints imposed by the shift properties of dynamic systems [12]. It can a bit surprising to discover that, in this respect, the dynamic case can be seen as a subcase of the algebraic one and that the previously mentioned shift properties lead (in general) to a unique solution [13, 14, 15]. It can also be surprising to show how this solution is linked to the solution of the maximal corank problem in the algebraic case.

EIV schemes and real data All previous considerations refer to an exact fullfillment of the assumptions behind the Frisch scheme (noise whiteness etc.) that could be satisfied, at best, only in asymptotic conditions. In all practical cases this cannot be achieved not even asymptotically because of a whole series of violations due to non linearity, non stationarity etc. The development of Frisch identification procedures requires thus the introduction of suitable criteria leading to the selection of a single model [16, 17, 18].

Bias Eliminated Least Squares and Instrumental Variable methods The Frisch scheme in the identification of dynamic processes enjoys some peculiarities like the congruence between the model and the estimated noise variances but does not constitute the only practical way for solving this problem. Another appealing method is based on the elimination of the bias that would be present by applying least squares. BELS methods constitute a large family of fast algorithms that, even if affected by convergence problems, can give good results [19, 20, 21]. The very general applicability context of IV methods allows also their use in the solution of EIV identification problems [22, 23, 24]. Despite their simple implementation and stimulating asymptotic properties these methods are affected by large estimation covariance with limited sets of data [25, 26].

Maximum Likelihood approaches When the ratio of the noise variances is *a–priori* known or when the input can be described by means of an ARMA process, it is possible, when the noise distribution is known, to apply a ML approach [27, 28, 29]. These approaches lead, in previous contexts, to the best achievable accuracy but can be affected by problems of convergence to local minima.

Bibliography

[1] S. Van Huffel (editor). *Recent Advances in Total Least Squares Techniques and Errors–in–Variables Modelling*. SIAM, Philadelphia, USA, 1997.

[2] S. Van Huffel and P. Lemmerling (editors). *Total Least Squares and Errors–in–Variables Modelling: Analysis, Algorithms and Applications*. Kluwer Academic Publishers, Dordrecht, The Netherlands, 2002.

[3] R. Frisch. Statistical confluence analysis by means of complete regression systems. Pub. No. 5, Economic Institute, Oslo University, 1934.

[4] R. E. Kalman. Identification from real data. In: M. Hazewinkel, H.G. Rinnooy Kan and D. Reidel (Eds.), *Current developments in the interface: Economics, Econometrics, Mathematics*, Dordrecht, The Netherlands, pp. 161–196, 1982.

[5] R. E. Kalman. *Nine lectures on identification*. Lecture notes on Economics and Mathematical Systems, Springer–Verlag, Berlin (to appear).

[6] B. De Moor and J. Vandewalle. The uniqueness versus the non–uniqueness principle in the identification of linear relations from noisy data. *Proc. of the 25th IEEE Conference on Decision and Control*, Athens, Greece, pp. 1663–1665, 1986.

[7] R. Guidorzi. Certain models from uncertain data: the algebraic case. *Systems & Control Letters*, 17:415–424, 1991.

[8] R. Guidorzi and M. Pierantoni. A new parametrization of Frisch scheme solutions. *Proc. of the 12th International Conference on Systems Science*, Wroclaw, Poland, pp. 114–120, 1995.

[9] E. Malinvaud. *Méthodes statistiques de l'économétrie*, 3rd edition, Dunod, Paris, 1980.

[10] R. Guidorzi. Identification of the maximal number of linear relations from noisy data. *Systems & Control Letters*, 24:159–166, 1995.

[11] K. G. Woodgate. An upper bound on the number of linear relations identified from noisy data by the Frisch scheme. *Systems & Control Letters*, 24:153–158, 1995.

[12] S. Beghelli, R. Guidorzi and U. Soverini. The Frisch scheme in dynamic system identification. *Automatica*, 26:171–176, 1990.

[13] M. Deistler. Linear errors–in–variables models. In: S. Bittanti (Ed.), *Time series and linear systems*, Lecture notes in Control and Information Sciences, Springer–Verlag, Berlin, pp. 37–67, 1986.

[14] B. D. O. Anderson and M. Deistler. Identifiability of dynamic errors–in–variables models. *Journal of Time Series Analysis*, 5:1–13, 1984.

[15] P. Stoica and A. Nehorai. On the uniqueness of prediction error models for systems with noisy input–output data. *Automatica*, 23:541–543, 1987.

[16] S. Beghelli, P. Castaldi, R. Guidorzi and U. Soverini. A robust criterion for model selection in identification from noisy data. *Proc. of the 9th International Conference on Systems Engineering*, Las Vegas, Nevada, pp. 480–484, 1993.

[17] R. Diversi, R. Guidorzi and U. Soverini. A new criterion in EIV identification and filtering applications. *Preprints of the 13th IFAC Symposium on System Identification*, Rotterdam, The Netherlands, pp. 1993–1998, 2003.

[18] R. Diversi, R. Guidorzi and U. Soverini. Frisch scheme–based algorithms for EIV identification. *Proc. of the 12th IEEE Mediterranean Conference on Control and Automation*, Kusadasi, Turkey, 2004.

[19] S. Sagara and K. Wada. On–line modified least–squares parameter estimation of linear discrete dynamic systems. *International Journal of Control*, 25:329–343, 1977.

[20] W. X. Zheng. Transfer function estimation from noisy input and output data. *International Journal of Adaptive Control and Signal Processing*, 12:365–380, 1998.

[21] W. X. Zheng. On unbiased parameter estimation of linear systems using noisy measurements. *Cybernetics and Systems*, 34:59–70, 2003.

[22] T. Söderström. Identification of stochastic linear systems in presence of input noise. *Automatica*, 17:713–725, 1981.

[23] S. Van Huffel and J. Vandewalle. Comparison of total least squares and instrumental variable methods for parameter estimation of transfer function models. *International Journal of Control*, 50:1039–1056, 1989.

[24] T. Söderström and K. Mahata. On instrumental variable and total least squares approaches for identification of noisy systems. *International Journal of Control*, 75:381–389,2002.

[25] U. Soverini and T. Söderström. Identification methods of dynamic systems in presence of input noise. *Preprints of the 12th IFAC Symposium on System Identification*, Santa Barbara, CA, USA, 2000.

[26] T. Söderström, U. Soverini and K. Mahata. Perspectives on errors–in–variables estimation for dynamic systems. *Signal Processing*, 82:1139–1154, 2002.

[27] M. J. Levin. Estimation of a system pulse transfer function in the presence of noise. *IEEE Transactions on Automatic Control*, 9:229–235, 1964.

[28] M. Aoki and P. C. Yue. On a priori error estimates of some identification methods. *IEEE Transactions on Automatic Control*, 15:541–548, 1970.

[29] K. V. Fernando and H. Nicholson. Identification of linear systems with input and output noise: the Koopmans–Levin method. *IEE Proceedings*, 132:30–36, 1985.

About the authors.

Roberto Guidorzi holds the chair of System Theory at the University of Bologna since 1980. He has, besides, been Director of the Computer Centre of the Engineering School of Bologna University from 1987 to 1992 and is, from January 1997, Director of CITAM (Interfaculty Center for Advanced and Multimedia Technologies). He has been visiting professor and invited speaker in European and American universities and has collaborated with several industries in the development of advanced projects, among them management systems for gas pipeline networks, satellite navigation systems, computerised injection systems, software for the identification of natural gas reservoirs, tracking and data fusion, early diagnosis in railway systems. He is the author of some 200 publications dealing with subjects of a methodological nature as well as applications of system theory methodologies. His present research interests concern errors-in-variables identification and filtering, blind channel equalisation and the development of e-learning environments.

Roberto Diversi was born in Faenza, Italy in 1970. He received the “Laurea” degree in Electronic Engineering in 1996 and the Ph.D. degree in System Engineering in 2000, both from the University of Bologna. Currently, he is an associate researcher at the Department of Electronics, Computer Science and Systems, University of Bologna. His research interests include system identification, optimal filtering, signal processing, fault detection.

Umberto Soverini was born in Bologna, Italy in 1959. He received the “Laurea” degree in Electronic Engineering from the University of Bologna in 1985 and the Ph.D. in System Engineering in 1990. In 1992 he has been appointed Research Associate at the Department of Electronics, Computer Science and Systems, University of Bologna where he is currently Associate Professor. In 1999 he has been a visiting researcher at the Department of Systems and Control, Uppsala University. His research interests include stochastic realization theory, signal processing, system identification and related applications to natural and industrial processes.

4.21 Model-based control in the errors-in-variables framework

Jens G. Linden, Benoit Vinsonneau, Keith J. Burnham
Control Theory and Applications Centre, Coventry University, UK, lindenj@coventry.ac.uk

Errors-in-variables (EIV) modelling techniques have received significant interest in recent years and their performance improvement, with respect to standard approaches, has been illustrated for certain cases [7]. In contrast,

however, comparably little work has been done in order to investigate the potential benefits of applying EIV techniques for the purpose of control. To address this deficiency, the aim here is to combine EIV filtering with existing control techniques in order to investigate the potential for improved closed-loop performance in comparison to a conventional approach. As a first step towards realising such an improvement on an practical application, a simulation study, which illustrates benefits of model-based control in the EIV framework, is presented.

In order to cope with the properties of closed-loop control systems, the conventional EIV setup (see [5] for instance) is modified accordingly: the additive noise on the inputs is then considered to model 'unobserved inputs' or uncertainties with respect to the true inputs of the system (e.g. actuator model mismatch) rather than measurement errors. This leads to a setup which is considered to be applicable to a wide range of industrial processes, whilst opening up opportunities to adopt a behavioural model structure [4].

A single-input single-output discrete-time nonlinear system, exhibiting both, bilinear terms as well as a hammerstein nonlinearity is simulated. An EIV-extended Kalman filter (EIV-EKF) [6], which is an extension of the EIV-Kalman filter (EIV-KF), developed in [2, 3], is applied in order to filter input and output noise components from the closed-loop system. In contrast to the EIV-KF, the EIF-EKF is able to deal with linear time-varying (LTV) systems and/or model mismatch, based on a linear default parameter set, the noisy measurements and the covariance matrix. The nonlinear system is to be controlled using a model-based controller of the incremental minimum variance type [1], which balances between the tracking of the reference signal and a smooth control action. The controller is based on an identified LTV model of the nonlinear system, which is evaluated by a separate adaptive online estimation algorithm. Moreover, due to the nature of the *k*-step ahead prediction scheme, the controller utilises the available/measured input and output signals. The estimation and control performance are assessed with and without the usage of the EIV-EKF. Preliminary results are encouraging, indicating improvement if the filtered signals are utilised. A Monte Carlo simulation is employed to show the consistency of the approach, hence demonstrating the benefits of EIV techniques in model based control.

Bibliography

[1] D. W. Clarke and P. J. Gawthrop. Self-tuning Control. *IEE Proceedings Part D: Control Theory and Applications*, 126(6):633–640, 1979.

[2] R. Diversi, R. Guidorzi, and U. Soverini. Algorithms for optimal errors-in-variables filtering. *Systems & Control Letters*, 48:1–13, 2003.

[3] R. Guidorzi, R. Diversi, and U. Soverini. Optimal errors-in-variables filtering. *Automatica*, 39:281–289, 2003.

[4] J. W. Polderman and J. C. Willems. *Introduction to mathematical systems theory: a behavioral approach*. Texts in Applied Mathematics. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 26 edition, 1998.

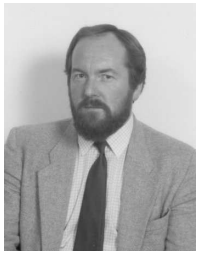
[5] T. Söderström, U. Soverini, and K. Mahata. Perspectives on errors-in-variables estimation for dynamic systems. *Signal Processing*, 82(8):1139–1154, 2002.

[6] B. Vinsonneau, D. P. Goodall, and K. J. Burnham. Errors-in-variables extended Kalman filter. In *Proc. IAR & ACD Conf.*, pages 217–222, Mulhouse, France, November 2005.

[7] Weihua Li and S. Joe Qin. Consistent dynamic PCA based on errors-in-variables subspace identification. *Journal of Process Control*, 11:661–678, 2001.

Jens G. Linden was born in Düren, Germany in 1980. In 2005, he received the Diploma in Technical Mathematics from Aachen University of Applied Sciences, Germany, and the MSc Degree (Distinction) in Control Engineering from Coventry University, UK. Currently, he is a PhD student in Control Engineering and based within the Control Theory and Applications Centre (CTAC), Faculty of Engineering and Computing, Coventry University, UK. His main research interests are in the fields of system identification, control engineering and related applications.

Benoit Vinsonneau was born in Angers, France, in 1981. He received the Degrees of BEng (Hons.) in Engineering in 2001 and MSc (Distinction) in Informatics and Control in 2002, from Coventry University, UK, and a DEA in Control, Signal Processing and Communications in 2003 from Henri Poincaré University Nancy 1, France. Currently, he is a PhD student in Control Engineering and based within the Control Theory and Applications Centre (CTAC), Faculty of Engineering and Computing, Coventry University, UK. His main research interests are in the fields of signal processing, system identification, control engineering and related applications to socio-economic, environmental and industrial processes. He is a student member of the Institute of Electronics and Electrical Engineering (IEEE).



Keith J. Burnham has been Professor of Industrial Control Systems, and Director of the Control Theory and Applications Centre (CTAC), Faculty of Engineering and Computing, Coventry University, UK, since 1999. Keith Burnham obtained his BSc (Mathematics), MSc (Control Engineering) and PhD (Adaptive Control) at Coventry University in 1981, 1984 and 1991, respectively. His main research interests are in the fields of adaptive control for nonlinear industrial systems and bilinear systems modelling and control. He is regularly consulted by industrial organisations to provide advice in areas of advanced algorithm development for control and condition monitoring. He is currently a member of the Editorial Board of the Transactions of the Institute of Measurement and Control (InstMC), and a member of the InstMC Systems Control Technology Panel, a member of the Informatics and Control Working Group of the Institution of Mechanical Engineers (IMechE), a member of the Control and Automation Professional Network Advisory Board of the Institution of Electrical Engineers (IEE), a member of the Automotive Control Technical Committee of the International Federation of Automatic Control (IFAC). He is also a corporate member of the IEE, InstMC, and Institute of Mathematics and its Applications (IMA).

4.22 Frequency domain maximum likelihood estimation of linear dynamic errors-in-variables models

R. Pintelon, and J. Schoukens, Vrije Universiteit Brussel, dept. ELEC, Pleinlaan 2, 1050 Brussels, Belgium
E-mail: Rik.Pintelon@vub.ac.be

This paper studies the linear dynamic errors-in-variables problem in the frequency domain. First the identifiability is shown under relaxed conditions. Next a frequency domain Gaussian maximum likelihood (ML) estimator is constructed that can handle discrete-time as well as continuous-time models on (a) part(s) of the unit circle or imaginary axis. The ML estimates are calculated via a computational simple and numerical stable Newton-Gauss minimization scheme. Finally the Cramr-Rao lower bound is derived.

Linear dynamic errors-in-variables (EIV) modelling is important in those applications where one is looking after a better understanding of the underlying input-output relation of a process rather than making an output prediction from noisy observations. One can distinguish between two cases: either the excitation of the process can freely be chosen, or one has to live with the operational (natural) perturbations. If the excitation can freely be chosen then it is strongly recommended to use periodic excitation signals because it simplifies significantly the identification problem: (i) nonparametric estimates of the disturbing noise (co-)variances are obtained in a preprocessing step, and (ii) since mutually correlated, coloured input/output errors are allowed, identification in feedback is just a special case of the general framework (see Pintelon and Schoukens, 2001). In the second case the excitation is often random and parts of it may even be unmeasurable. This paper handles the second case, assuming that the excitation is a stochastic process with rational power spectrum. As will be shown in the sequel of the paper the second case is much more complicated than the first: besides the plant model one should also identify simultaneously the signal, and the input/output noise models.

Identifiability is a first key issue in EIV modelling: under which conditions on the excitation, the input/ output

errors, and the process is the EIV problem uniquely solvable? This question has been studied in detail in econometrics and an extensive literature is available (see Sderstrm, 2006 for an exhaustive overview). For example, Anderson and Deistler (1984) handles the identifiability of scalar EIV problems with coloured input/output errors, while Nowak (1993) covers the multivariable case. A second key issue is the numerical calculation of the EIV estimates. Several algorithms have been proposed, each of them having their specific advantages and disadvantages (see Sderstrm, 2006 for an exhaustive overview). For example, spectral factorization is the computational bottle neck of the statistically efficient time domain maximum likelihood method (Sderstrm and Stoica, 1989), while the computational simple instrumental variable methods have low statistical accuracy (Sderstrm, 2006). Except for Mahata and Garnier (2005), all methods handle the discrete-time case and no algorithms for direct continuous-time EIV modelling are available. In Mahata and Garnier (2005) a method is presented for identifying continuous-time models from non-uniformly sampled data in the presence of white input/output errors.

The contributions of this paper are:

1. the identifiability of general linear dynamic EIV models is shown under relaxed conditions,
2. a (computational simple) frequency domain Gaussian maximum likelihood (ML) estimator is developed for the general case of coloured and mutually independent input/output errors,
3. the ML estimator can handle discrete-time as well as continuous-time modelling on (a) part(s) of the unit circle or imaginary axis,
4. a numerical stable Newton-Gauss minimization scheme of the ML cost function is derived,
5. easy calculation of the Cramr-Rao lower bound.

Bibliography

[1] Agüero, J. C., and Graham, G. C. (2006). Identifiability of errors in variables dynamic systems, 14th IFAC Symposium on System Identification (Sysid 06), New Castle (Australia), March 29-31.

[2] Anderson, B. D. O. (1985). Identification of scalar errors-in-variables models with dynamics, Automatica, vol. 21, no. 6, pp. 709-716.

[3] Anderson, B. D. O. and Deistler, M. (1984). Identifiability in dynamic errors-in-variables models, Journal of Times Series Analysis, vol. 5, no. 1, pp. 1-13.

[4] Brillinger, D. R. (1981). Time Series: Data Analysis and Theory. McGraw-Hill, New York.

[5] Caines, P. E. (1988). Linear Stochastic Systems. Wiley, New York.

[6] Deistler, M., and Anderson, B. D. O. (1989). Linear dynamic errors-in-variables models. Some structure theory, Journal of Econometrics, vol. 41, pp. 39-63.

[7] Golub, G. H. and Van Loan, C. F. (1996). Matrix Computations (3rd ed.). John Hopkins University Press, Baltimore.

[8] Karlsson, E., Söderström, T., and Stoica P. (2000). Computing the Cramr-Rao lower bound for noisy input output systems, Signal Processing, vol. 80, no. 11, pp. 2421-2447.

[9] Mahata, K., and Garnier, H. (2005). Direct identification of continuous-time errors-in-variables models, Preprints of the 16th World Congress of the International Fedeartion of Automatic Control, July 3-8, Prague (Czech Republic).

[10] Middleton, R. H., and G. C. Goodwin (1990). Digital Control and Estimation. Prentice-Hall, London.

[11] Nowak, E. (1993). The identification of multivariate linear dynamic errors-in-variables models, Journal of Econometrics, vol. 59, no. 3, pp. 213-227.

[12] Pintelon, R. (2006). A frequency domain bias compensated least squares algorithm for linear dynamic errors-in-variables models, Internal report no. RP1.01.06, dept. ELEC, Vrije Universiteit Brussel, Pleinlaan 2, 1050 Brussel (Belgium).

[13] Pintelon, R. and I. Kollr (2005). On the frequency scaling in continuous-time modeling. IEEE Trans. Instrum. Meas., vol. 53, no. 5, pp. 318-321.

[14] Pintelon, R. and Schoukens, J. (2001). System Identification: A Frequency Domain Approach, IEEE Press, New York (USA).

[15] Pintelon, R., and Schoukens, J. (2006). Box-Jenkins identification revisited - Part I: theory, Automatica, vol. 42, no. 1, pp. 63-75.

[16] Pintelon, R., Y. Rolain, and Schoukens, J. (2006). Box-Jenkins identification revisited - Part II: applications, Automatica, vol. 42, no. 1, pp. 77-84.

[17] Pintelon R., J. Schoukens, Y. Rolain, B. Cauberghe, E. Parloo, P. Guillaume (2005). Identification of continuous-time noise models. Proceedings of the 16th IFAC World Congress, Prague (Czec Republic), July 4-8.

[18] Söderström, T. (2006). Errors-in-variables methods in system identification, 14th IFAC Symposium on System Identification (Sysid 06), New Castle (Australia), March 29-31.

[19] Söderström, T. and P. Stoica (1989). System Identification. Prentice-Hall, Englewood Cliffs.

[20] Söderström, T., Soverini, U., and Mahata, K. (2002). Perspectives on errors-in-variables estimation for dynamic systems, Signal Processing, vol. 82, pp. 1139-1154.

[21] Söderström, T., Hong, M., and Zheng, W.X. (2005). Convergence properties of bias-eliminating algorithms for errors-in-variables identification, International Journal of Adaptive Contr. and Signal. Processing, vol. 19, no. 9, pp. 703-722.

[22] Zheng, W.X. (2002). A bias corrected method for identification of linear dynamic errors-in-variables models. IEEE Trans. Autom. Control, vol. 47, no. 7, pp. 1142-1147.

4.23 Identifiability analysis for errors-in-variables problems

J. Schoukens and R. Pintelon, Vrije Universiteit Brussel

In this paper, an analysis is made of the identifiability problem of the errors-in-variables identification as a function of the added prior information. The analysis is valid for continuous or discrete time systems. The measured input and output are disturbed by zero mean Gaussian distributed noise, that are in general colored and mutually correlated.

In this paper we study the identifiability using only 2nd order moments of the input and output. The analysis is done in the frequency domain, but the results are also valid for the time domain. In the frequency domain, it is known that the covariance matrix is asymptotically block diagonal, with one block per frequency. For Gaussian signals and noise the covariance matrices present all information that can be extracted from the data.

Without making any prior assumption on the signals, plant model, or the disturbing noise, the EIV problem is not identifiable. It will be shown that by adding additional assumptions like: white noise/parametric plant or noise model/no mutual correlation ..., identifiability can eventually be obtained. The basic idea is to analyse the number of independent constraints that can be extracted from the covariance matrices as a function of the added constraints, and these are compared to the number of unknown parameters. This is done for different combinations: parametric/ nonparametric plant/noise model; white/colored noise; mutual correlated/uncorrelated input/output noise.

Acknowledgement: This research is supported by the FWO-Vlaanderen, the Flemish (GOA-ILiNoS) and the Belgian government (IUAP-V22).

4.24 Optimal Parameter Estimation from Shift-Invariant Subspaces

Richard J. Vaccaro, Department of Electrical Engineering, University of Rhode Island, Kingston, RI, USA,
vaccaro@ele.uri.edu

Consider the following parameter estimation problems: (1) estimating frequencies, damping factors, amplitudes, and phases from data consisting of a sum of real or complex exponentially damped sinusoids; (2) estimating state-space models from matrix sequences (MIMO impulse responses); (3) estimating directions of arrival using a uniform linear array of sensors. All of these problems share the following two properties: (1) a certain matrix formed from noise-free data has low rank, and (2) at least one subspace of this matrix is shift invariant. When processing noisy data by so-called "subspace-based" methods, the first property is exploited by using a rank-revealing factorization such as the singular value decomposition (SVD) to get a subspace estimate. Because of noise, the second property is not satisfied, and the equation expressing shift invariance has no exact solution. The usual approach is to solve the shift equation in a least squares (LS) or total least squares (TLS) sense. For the three problems listed above, the LS and TLS solutions are not statistically optimal, and the variances of estimates obtained from these solutions do not achieve the CR bound.

In this paper we use a first-order subspace perturbation expansion to obtain an expression for the noisy subspace projection matrix estimates obtained from an SVD of the data matrix, in terms of the noise-free subspace projection matrix plus an additive perturbation term. This expression is used to calculate the covariance matrix of the additive perturbation analytically, and the result is used to solve a single weighted least squares problem for the underlying shift-invariant subspace. We call this procedure Optimal Subspace Estimation (OSE). Signal parameters are then obtained by solving the shift eqation using the OSE subspace estimate. The variances of the estimated parameters achieve the CR lower bounds for the problem over a range of signal-to-noise ratios.

This paper is related to previous work that has derived statistically optimal subspace methods for DOA estimation problem (problem 3 mentioned above). [1, 2] derive an optimally weighted ESPRIT algorithm, which improves the performance of the original, least-squares ESPRIT algorithm, so that the variances of the estimates reach the CR bound. Another subspace-based approach for estimating DOAs, derived by Viberg and Ottersten, is weighted subspace fitting [3]. Yet another approach for optimal DOA estimation is the MODE algorithm of Stoica and Sharman [4]. The common ingredient in all of the optimal algorithms cited here is a description of how the left singular vectors of a data matrix (or equivalently, the eigenvectors of a covariance matrix), are perturbed by additive noise. All of these algorithms were derived using an asymptotic (data length goes to infinity) first-order expansion for the perturbed singular vectors. This expansion appears in [5] based on results from Brillinger [6]. The expansion used in these papers is valid when the noise in each column of the data matrix be statistically independent. This is *not* the case in Problems 1 and 2 mentioned above. There is related work [7], which considers noise-free matrices that are not necessarily rank deficient and derives first and second-order perturbation formulas. For the rank deficient case considered here, such formulas have been derived in [8]. The biggest difference between these approaches is that [7] derives a perturbation expansion for individual basis vectors, while the SPE [8] expands an entire subspace.

The advantage of the non-iterative OSE algorithm, based on the subspace perturbation expansion (SPE), is that it can be used for all three problems mentioned above. The application of the SPE to the DOA estimation problem has been done in [9]. In this paper, we show how the SPE is used to derive statistically optimal algorithms for problems 1 and 2. In addition, we use a different approach for solving the resulting weighted least squares problems that has advantages over the approach used in [9].

Bibliography

[1] A. Eriksson and P. Stoica, "Optimally weighted ESPRIT for direction estimation," *Signal Processing*, vol. 38, pp. 223–229, 1994.

[2] P. Stoica and M. Viberg, "Weighted LS and TLS approache yield asymptotically equivalent results," *Signal Processing*, vol. 45, pp. 255–259, 1995.

[3] M. Viberg and B. Ottersten, "Sensor array processing based on subspace fitting," *IEEE Trans. Signal Processing*, vol. 39, pp. 1110–1121, 1991.

[4] P. Stoica and K. Sharman, "Maximum likelihood methods for direction-of-arrival estimation," *IEEE Trans. Acoustics, Speech, and Signal Processing*, vol. 38, pp. 1132–1143, 1990.

[5] M. Kaveh and A. Barabell, "The statistical performance of the MUSIC and the minimum-norm algorithms in resolving plane waves in noise," *IEEE Trans. Acoustics, Speech, and Signal Processing*, vol. 34, pp. 331–341, 1986.

[6] D. Brillinger, *Time Series: Data Analysis and Theory*. New York: Holt, Rinehart, and Winston, 1975.

[7] Y. Xu, "Perturbation analysis for subspace decomposition with application in subspace-based algorithms," *IEEE Trans. Signal Processing*, vol. 50, pp. 2820–2830, 2002.

[8] R. Vaccaro, "A second-order perturbation expansion for the SVD," *SIAM Journal on Matrix Analysis and Applications*, vol. 15, pp. 661–771, 1994.

[9] R. Vaccaro and Y. Ding, "A new state-space approach for direction finding," *IEEE Trans. Acoustics, Speech, and Signal Processing*, vol. 42, pp. 3234–3237, 1994.

About the author Richard J. Vaccaro is a Professor of Electrical Engineering at University of Rhode Island, where he has been a faculty member since 1983. His research interests include subspace methods for parameter estimation and signal modeling; direction-of-arrival estimation using sensor arrays; passive and active sonar signal processing.

4.25 On the role of constraints in system identification

Arie Yeredor, School of Electrical Engineering, Tel-Aviv University, `arie@eng.tau.ac.il`

The general framework System identification is concerned with the estimation of parameters characterizing an unknown system. The estimation is usually based on observations of the system's (possibly noisy) input(s) and output(s). In the so-called "blind" system identification scenario, the estimation is based on the observed output(s) only, aided by some general knowledge about statistical properties of the input(s), rather than by actual observations thereof.

Quite commonly, the discussion is limited to discrete-time systems, assumed to be linear and time-invariant (LTI), stable and causal. As such, their input-output relation can always[1] be described as

$$y[t] = \sum_{\ell=0}^{\infty} h[\ell]u[t-\ell] \quad \forall t \in \mathbb{Z}, \tag{4.17}$$

where $u[t]$ is the input, $y[t]$ is the output and $h[\ell]$ is the system's impulse response. In the Multiple-Inputs, Multiple-Outputs (MIMO) case, the inputs and outputs may assume a vector form, but the basic form of the convolutive relation remains the same:

$$y[t] = \sum_{\ell=0}^{\infty} H[\ell]u[t-\ell] \quad \forall t \in \mathbb{Z}, \tag{4.18}$$

where $u[t]$ and $y[t]$ denote (respectively) M-dimensional and L-dimensional input and output vectors, and $H[\ell]$ denote the $L \times M$ impulse response matrices.

Although such systems are fully described by their (generally infinite) impulse response, some prior knowledge pertaining to their structure often allows to assume that they can also be described by a reduced (finite) set of parameters. For example, the Single-Input Single-Output (SISO) model (4.17) is often also modeled by a difference equation,

[1] barring the usually uninteresting possibility of an additive constant.

$$a_0 y[t] = -\sum_{k=1}^{N_p} a_k y[t-k] + \sum_{k=1}^{N_z} b_k u[t-k], \tag{4.19}$$

where N_p and N_z are (respectively) the number of poles and zeros in this model and $\theta \stackrel{\triangle}{=} [a_0\, a_1\, a_2\, \cdots\, a_{N_p}\, b_0\, b_1\, \cdots\, b_{N_z}]^T$ is the finite vector of unknown system's parameters.

Likewise, the MIMO model (4.18) is often described using a State-Space model,

$$\begin{aligned} x[t+1] &= Ax[t] + Bu[t] \\ y[t] &= Cx[t] + Du[t], \end{aligned} \tag{4.20}$$

where $x[t]$ is an "internal" (unobserved) N-dimensional state-vector, and A, B, C and D are matrices of the appropriate dimensions, which together comprise the finite set of unknown model parameters θ.

Similarly, description of the MIMO system with a (matrix) difference equation model, or of the SISO system with a state-space model are also possible. We shall regard the SISO case as a particular case of the MIMO case, except where the distinction is necessary.

Identification via constrained optimization Assume that an observation interval of length T is available. Typical system identification approaches seek to minimize (or to maximize) some criterion, which generally involves all of the available output/input observations $Y \stackrel{\triangle}{=} [y[1]\, y[2]\, \cdots\, y[T]]$ and $U \stackrel{\triangle}{=} [u[1]\, u[2]\, \cdots\, u[T]]$ (which is absent in the blind scenario), the unknown system parameters θ, and possibly some additional "nuisance parameters" ϕ, often representing some underlying, unobserved signals. Optimization of the criterion is sought with respect to θ and ϕ, yielding in turn the estimates of these parameters:

$$\min_{\theta,\phi} C(Y,U;\theta,\phi) \quad \Rightarrow \quad \hat{\theta},\hat{\phi}. \tag{4.21}$$

Often, however, some constraints on either θ, ϕ or both are introduced into the optimization (4.21). The motivation for incorporating these constraints can come from a surprisingly large variety of perspectives on the problem. The main goal of this paper is to review the different approaches that lead to different types of constraints, each with the associated motivations, and to outline the resulting optimization and estimation approaches, providing some comparative study of the results.

Following are a few examples of useful constraints, with brief description of the motivation and frameworks by which they are applied.

- Constraints aimed at avoiding a trivial minimizer of the criterion. This is usually the basic motivation for adding constraints, where the associated optimization problem cannot yield a useful solution without excluding trivial solutions from the feasibility set.
- Constraints aimed at incorporating prior knowledge about the system, such as the locations of some of its poles or zeros (in a fashion similar to [4], [2]), so as to improve the resulting estimation accuracy by effectively reducing the number of degrees of freedom.
- Constraints aimed at imposing certain "natural" structures on some of the signals involves. This type of constraints usually involves the nuisance parameters ϕ, rather than the parameters of interest θ. For example, in the Structured Total Least Squares (STLS, e.g., [8]), ϕ can consist of the estimated noiseless signal matrix, whereas the constraints confine the elements of that matrix to obey a certain structure (Hankel, Toeplitz, etc.).
- Constraints aimed at mitigating the bias induced by additive output noise (e.g., [9], [7], [16]) or by the use of an inconsistent criterion [18].
- Constraints aimed at guaranteeing the stability of the resulting estimated system [5], [9].

Most of these constraints would take the form of "equality constraints", namely $f(\theta,\phi)=0$ (where $f(\cdot)$ is an associated vector function), but some may also involve forms of inequalities. The basic approaches for optimizing the criteria under the associated types of constraints (Lagrange multipliers, Method of multipliers, Successive projections, Linear programming) will also be reviewed and graphically illustrated, where applicable.

Bibliography

[1] D. P. Bertsekas. *Constrained optimization and Lagrange multiplier methods*, Academic Press, 1982.

[2] H. Chen, S. Van Huffel, J. Vandewalle. "Improved methods for exponential parameter estimation in the presence of known poles and noise," *IEEE Trans. Signal Processing*, vol. 45, pp.1390-1393, 1997.

[3] B. De Moor, M. Gevers, G. Goodwin. "L_2-Overbiased, L_2-underbiased and L_2-unbiased estimation of transfer functions," *Automatica*, vol. 30, pp. 893-898, 2004.

[4] E. M. Dowling, R. D. DeGroat, D. A. Linebarger. "Total least squares with linear constraints," em Proc. ICASSP'92, vol. V, pp.341-344, 1992.

[5] S. L. Lacy, D. S. Bernstein. "Subspace identification with guaranteed stability using constrained optimization," *IEEE Trans. Auto. Control*, vol. 48, pp.1259-1263, 2003.

[6] P. Lemmerling, B. De Moor. "Misfit versus latency," *Automatica* vol.37 pp.2057-2067, 2001.

[7] R. Lopez-Valcarce. "An algorithm for unit-norm equation error system identification based on the method of multipliers," *IEEE Trans. Signal Processing*, vol. 51, pp.3080-3085, 2003.

[8] I. Markovsky, J. C. Willem, S. Van Huffel, B. De Moor, R. Pintelon. "Application of structured total least squares for system identification and model reduction," *IEEE Trans. Auto. Control*, vol. 50, pp.1490-1500, 2005.

[9] P. A. Regalia. "An unbiased equation error identifier and reduced-order approximations," *IEEE Trans. Signal processing*, vol. 42, pp.1397-1411, 1994.

[10] H. C. So, Y. T. Chan. "Analysis of an LMS algorithm for unbiased impulse response estimation," *IEEE Trans. Signal Processing*, vol. 51, pp.2008-2013, 2003.

[11] T. Söderström, P. Stoica. *Instrumental variable methods for system identification*, Springer-Verlag, New-York, 1983.

[12] T. Söderström, P. Stoica. *System identification*, Prentice-Hall, Englewood Cliffs, NJ, 1989.

[13] P. Stoica, T. Söderström. "Bias correction in least squares identification," *Int. J. Control.*, vol.35, pp.449-457, 1982.

[14] K. T. Tsakalis, M. Deisher, A. Spanias. "System identification based on bounded error constraints," *IEEE Trans. Signal Processing*, vol. 43, pp.3071-3075, 1995.

[15] R. J. Vanderbei. *Linear programming: foundations and extensions*, Kluwer Academic, 2001.

[16] T. H. Van Pelt, D.S. Bernstein. "Quadratically constrained least squares identification," *Proc. American Control Conference, Arlington, VA, USA, 2001*, pp.3684-3689.

[17] M. Viberg. "Subspace-based methods for the identification of linear time-invariant systems," *Automatica*, vol. 31, pp.1835-1851, 1995.

[18] A. Yeredor, B. De Moor. "On homogeneous least-squares problems and the inconsistency introduced by misconstraining," em Computational Statistics and Data Analysis, vol. 47, pp.455-465, 2004.

About the author Arie Yeredor received the B.Sc. *(summa cum laude)* and Ph.D. degrees in electrical engineering from Tel-Aviv University (TAU), Tel Aviv, Israel, in 1984 and 1997, respectively. He is currently a Senior Lecturer with the Department of Electrical Engineering - Systems, at TAU's School of Electrical Engineering, where he teaches courses in statistical and digital signal processing, and has been awarded the yearly Best Lecturer of the Faculty of Engineering Award five times. He also holds a consulting position with NICE Systems Inc., Raanana, Israel, in the fields of speech and audio processing, video processing, and emitter location algorithms. Dr. Yeredor is an Associate Editor for IEEE SIGNAL PROCESSING LETTERS and for IEEE TRANSACTIONS ON CIRCUITS AND SYSTEM II, and is a member of the IEEE Signal Processing Society's Signal Processing Theory and Methods (SPTM) Technical Committee.

4.26 Principal component, independent component and parallel factor analysis

Lieven De Lathauwer, Lab. ETIS, CNRS/ENSEA/UCP, Cergy-Pontoise, France,
`Lieven.DeLathauwer@esat.kuleuven.be`

This talk is an introduction to Independent Component Analysis (ICA) and Parallel Factor Analysis (PARAFAC), the way they are related and their links with Principal Component Analysis (PCA). PCA is now a standard technique for the analysis of two-way multivariate data, i.e., data available in matrix format. However, principal components are subject to rotational invariance. By imposing statistical independence rather than uncorrelatedness, the solution becomes unique. This is ICA. On the other hand, PARAFAC is a technique for multiway data analysis, based on the decomposition of the data tensor in rank-1 terms. PARAFAC is unique under mild conditions on the factors. ICA decomposes a higher-order cumulant tensor in rank-1 terms. Hence, ICA uniqueness stems from PARAFAC uniqueness. PCA is often used as preprocessing, leading to PARAFAC with orthogonality constraints.

4.27 Applications of TLS and related methods in the environmental sciences

José A. Ramos, Department of Electrical and Computer Engineering, Indiana University Purdue University - Indianapolis, `jaramos@iupui.edu`

Rainfall-Runoff and Signal Separation Problems: Converting rainfall into runoff is a highly nonlinear process due to the soil-water interaction that starts when rainfall reaches the ground. Additional variables to consider are evaporation, transpiration, losses due to vegetation and land use, and the different flow processes that take place in a watershed. For instance, baseflow is a much slower process than groundwater and surface flow. Given records of rainfall and runoff data, one can build an accurate state-space model such as

$$\begin{aligned} x_{k+1} &= Ax_k + Bu_k + w_k \\ y_k &= Cx_k + Du_k + v_k, \end{aligned}$$

where at time k, u_k, y_k, and x_k are, respectively, the rainfall, runoff, and the state of the system. Such models have been used in real-time forecasting scenarios for flood control purposes [4]. However, the above model does not take into account the nonlinearities of the rainfall-runoff process. Most lumped rainfall-runoff models separate the baseflow and groundwater components from the measured runoff hydrograph in an attempt to model these as linear hydrologic reservoir units. Similarly, rainfall losses due to infiltration as well as other abstractions are separated from the measured rainfall hyetograph, which are then used as inputs to the linear hydrologic reservoir units. This data preprocessing is in essence a nonlinear signal separation problem where rainfall is separated into infiltration and excess rainfall, and the measured hydrograph into surface flow and groundwater flow as shown in Figure 1. These are then used to build separate linear models such as

$$\begin{aligned} x^g_{k+1} &= A_g x^g_k + B_g u^g_k \\ y^g_k &= C_g x^g_k + D_g u^g_k, \end{aligned}$$

$$\begin{aligned} x^s_{k+1} &= A_s x^s_k + B_s u^s_k \\ y^s_k &= C_s x^s_k + D_s u^s_k, \end{aligned}$$

where

$$\begin{array}{l} u_k = u^g_k + u^s_k \\ y_k = y^g_k + y^s_k \end{array}, \qquad \sum_{k=0}^{N-1} u_k = \sum_{k=0}^{N-1} y_k.$$

In the separation process, a TLS approach is used since the infiltration process is an exponential signal. Thus, the classical NMR fitting techniques [2, 3, 5] are used.

Physical Parameter Extraction Problems: When modeling physical processes such as the rainfall-runoff interaction, where water flows into different compartments, one is faced with a physical parameter extraction problem. This is quite evident in black-box system identification where an unknown similarity transformation matrix destroys the physical meaning of the problem. Here we show that such similarity trsnsformation can be recovered as a post identification TLS problem. That is, suppose the identified state-space system matrices are $\{\bar{A}, \bar{B}, \bar{C}, \bar{D}\}$, while the physical parameter matrices are those of a system with three compartments as shown in Figure 2. The the parameter matrices for both the physical and identified models are shown in Table 1.

The two systems are related by a similarity transformation T, i.e., $T\bar{A}T^{-1} = A$, $T\bar{B} = B$, and $\bar{C}T^{-1} = C$. As one can see, this system of equations is nonlinear, but if we rewrite these as $T\bar{A} = AT$, $T\bar{B} = B$, and $\bar{C} = CT$, then we convert the problem into a linear one. It turns out that the solution can be framed as an orthogonal complement problem of the form $x^T \mathscr{A} = 0_{p\times q}$, where p and q are problem dependent dimensions, x contains all the elements of the T matrix plus some of the parameters from the A matrix, and $\mathscr{A}$ is a matrix obtained from the identified model parameters, i.e., $\{\bar{A}, \bar{B}, \bar{C}\}$. We will generalize the above results and show an example of a two-tank reservoir model.

Other Applications and Related Methods: We will also discuss applications of TLS in hyperspectral analysis, variogram fitting of spatial processes, and Chemometrics applications in the environemental sciences.

Table 1. Physical (unknown) and identified model parameters.

Physical Model	Identified Model
$A = \begin{bmatrix} -k_{21} & k_{12} & 0 \\ k_{21} & -k_{12}-k_{32} & k_{23} \\ 0 & k_{32} & -k_{23}-k_{03} \end{bmatrix}$	$\bar{A}_c = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22}a_{23} & \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$
$B = \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix}$	$\bar{B}_c = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \end{bmatrix}$
$C = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$\bar{C}_c = \begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \end{bmatrix}$
$D = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$	$\bar{D}_c = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$

Bibliography

[1] Cadzow, J., "Signal enhancement: A composite property mapping algorithm," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 36, pp. 49-62, 1988.

[2] Chen, H., "Subspace based parameter estimation of exponentially damped sinusoids with application to nuclear magnetic resonance spectroscopy data," Ph.D. dissertation, Department of Electrical Engineering, Katholieke Universiteit Leuven, Belgium, 1996.

[3] De Moor, B. and J. Vandewalle, “A numerically reliable algorithm for fitting a sum of exponentials or sinusoids to noisy data,” Proceedings 3rd IFAC/IFIP International Symposium on Computer Aided Design in Control and Engineering System, Lyngby, Copenhagen, Denmark, 1985.

[4] Ramos, J. A., Mallants, D., and J. Feyen, “State-space identification of linear deterministic rainfall-runoff models,” *Water Resources Research*, vol. 31 (6), pp. 1519-1532, 1995.

[5] Van Huffel, S., “Enhanced resolution based on minimum variance estimation and exponential modeling,” *Signal Processing*, vol. 33, no. 3, pp. 333-335, 1992.

[6] Wackernagel, H., *Multivariate Geostatistics: An Introduction with Applications*, Springer-Verlag, Berlin, 3^{rd} edition, 2003.

[7] Zieger, H. P. and A. J. McEwen, “Approximate linear realization of given dimension via Ho’s algorithm,” *IEEE Trans. Automat. Contr.*, AC-19, pp. 153, 1974.

About the author

José Ramos obtained his BSCE from the University of Puerto Rico at Mayaguez in 1979. He then went to Georgia Institute of Technology and completed his MS and Ph.D. degrees in 1979 and 1985, respectively. From 1985 to 1990 he worked at United Technologies Optical Systems in West Palm Beach, Florida, where he developed Kalman filtering and tracking algorithms as part of a military program on adaptive optics for space applications. That same year joined the ESAT goup at the Katholieke Universiteit Leuven as a post doctoral fellow, working on subspace system identification algorithms for linear, bilinear, and 2-D systems. From 1991 to 1993 he was a post doctoral fellow at the Institute for Land and Water Management, Katholieke Universiteit Leuven. He has been at Indiana University Purdue University - Indianapolis since 1997, where he teaches courses in stochastic processes, instrumentation, modern control, circuit theory, system identification, and system theory. His current research interests are in the areas of system identification, multivariate data analysis, optimization, and applied numerical linear algebra. His applications are in the areas of water resources, biomedical signal processing, image registration, and time series analysis. He has collaborated with The University of Montpellier II on the use of splines in nonlinear system identification, and with the University of Oporto on various iterative subspace system identification algorithms.

4.28 Recursive total least squares not using SVD

Dan Pachner, Jirka Roubal, Pavel Trnka, Department of Control Engineering , Faculty of Electrical Engineering, Czech Technical University, Technicka 2, 166 27 Prague 6, Czech Republic,
{pachner, roubalj, trnkap}@control.felk.cvut.cz

Introduction The standard approach to the Total Least Squares TLS problem solution is the SVD de-composition algorithm [2]. The SVD solution is well known and reliable numerical tool nicely implemented in Matlab. Nonethe-less, for some applications like adaptive control, which may also be based on a TLS model, the SVD matrix decomposition can be unnecessarily complex. In our contribution, we propose a simplified TLS problem solution approach, which replaces the SVD by the QR decomposition.

Recursive Total Least Squares in a Rotating System of Co-Ordinates In our contribution, we propose an iterative TLS problem solution. At every iteration stage, an OLS (ordinary least squares) problem is solved. At the next iteration, the dependant variable is rotated to the direction perpendicular to the last stage model hyper-plane. Whereas with the ordinary least squares the error distances are measured along a defined direction, the TLS solution minimizes the squared distances along the direction perpendicular to the model hyper-plane. We propose the iterative scheme based on that direction update. We will show the sum of squared distances to be minimized by both OLS and TLS models are the same only the constraint is different: either the length of the vector of model parameters has to be one

or the last vector co-ordinate has to be one. There is one interesting consequence of this observation: the two OLS and TLS optimization problems would yield the same vector of parameters if it would have happened the OLS optimization problem would result in a vector of the last co-ordinate equal to one. In this case all model parameters would be zero except of the last one. In other words, if one calculates the OLS model parameters and the result happens to be such vector of all zeros except of the last one, there is no need to calculate the TLS model parameters because the result would have been exactly the same. The reason is the vector of parameters satisfies both constraints at the same time. Our algorithm is based on this idea. The geometric interpretation of the data rotation is the following one: At each iteration, the OLS problem supposes a special dependant variable *"y* as a linear combination of the TLS problem data *"x* such that this linear combination is perpendicular to the model hyper-plane from the previous iteration. It also supposes the independent variables to be an ortho-normal basis of the previous iteration hyper-plane. Thus, the distance, which has to be measured along the direction perpendicular to the model hyper-plane, is measured along the direction perpendicular to the last known OLS model hyper-plane. Both OLS problem solution and the rotation matrix update can be achieved via the QR matrix decomposition. Thus, our TLS problem solution approach replaces the SVD matrix decomposition with a series of the QR decompositions. We will demonstrate that our iterative algorithm can easily be generalized to the mixed TLS problems. A proof of convergence will be provided.

Adaptive TLS and State Space Models In our contribution will show this TLS problem solution is convenient for the adaptive TLS models and the recursive subspace system identification. The reason is the dependant variable direction update requires only a low number of iterations for slowly varying model parameters. Thus, exactly one matrix QR decomposition has to be performed instead of the SVD decomposition to track slowly varying TLS model parameters. We will show this algorithm can easily be combined with the standard forgetting techniques known from the adaptive control theory [3]. To demonstrate this, we will show how a rotating line which rotates round the clock like a clockwork hand, can be tracked by the algorithm, see the Figure 1. The clockwork hand position is measured under noisy conditions and the rotation pivot position is unknown. Next we will show that our algorithm can be used with a subspace identification approach related state space model identification recursive algorithm [1] and with an ARX model structure related multi-step predictor identification.

Bibliography

[1] PACHNER, D., ROUBAL, J., and TRNKA, P. On a Recursive State Space Model Identification Method. In *4th International Workshop on Total Least Squares and Errors-in-Variables Modeling*, Leuven, Belgium, 2006. ⟨http://homes.esat.kuleuven.be/∼imarkovs/workshop.html⟩.

[2] GOLUB, G.H, HOFFMAN, A., and STEWART, G.W., B. A generalization of the Eckart-Young-Mirsky matrix approximation theorem. Linear Algebra and Its Applications 88/89, pages 317-327, 1987

[3] PETERKA, V., Bayesian System Identification. P. EYKHOFF (Ed.). Pergamon Press, Oxford, UK, 1981, pp. 239-304.

4.29 On a recursive state space model identification method

Dan Pachner, Jirka Roubal, Pavel Trnka
Department of Control Engineering, Faculty of Electrical Engineering, Czech
Technical University, Technicka 2, 166 27 Prague 6, Czech Republic,
{pachner, roubalj, trnkap}@control.felk.cvut.cz

Introduction Most of our linear model identification applications were based on the Bayesian paradigm. The Bayesian statistical approach is based on the subjective belief measure [3]. Thanks to this statistical background, the Bayesian identification methods can easily combine the experimental data with the prior information about the process model. This technique proved to be very rewarding in all our applications. In many cases, the experimental data do not bring

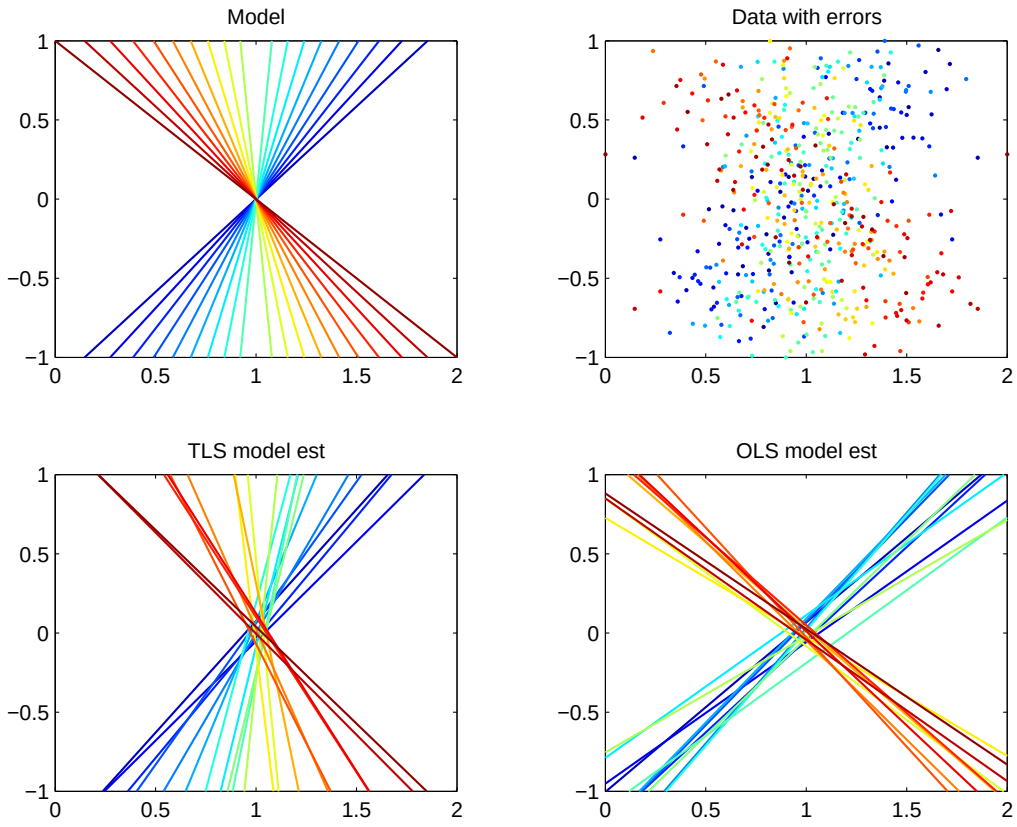


Figure 4.1: The clockwork hand tracked by the adaptive TLS and OLS models.

enough information to build a reasonable process model [5]. In practice, there can be found many prior information examples. For instance the process settling time is between 1 and 5 minutes, the process DC gain is exactly 1 (for a conveyer) etc. Quite often the low frequency information about the process model can be drawn from a first principles considerations (laws of physics, chemical laws), whereas the high frequency behaviour must be identified experimentally. Because for most processes, the fast transient effects are usually governed by too complicated theoretical laws.

Since the subspace identification methods have emerged in 1990's [1] there has been an obvious gap between Bayesian approach and the new methods. Whereas the new subspace identification proved to be very robust and reliable, the Bayesian methods could easily provide the process parameter uncertainty in terms of their conditional probability density function, and to combine the prior information with the data. the other Bayesian approach advantage has been it naturally provides recursive system identification methods, which are linked to the sufficient statistics calculations.

Since then, it has been our goal to understand the subspace identification methods in view of the Bayesian approach. We set the goal to find the new method Bayesian interpretation. We hope this will furnish the subspace identification methods with the standard Bayesian advantages as mentioned.

State Space Model Identification in the Bayesian Paradigm In our contribution, we will demonstrate how the state space linear system model parameters can be estimated within the Bayesian paradigm. It will be shown the state space model is naturally derived if the multi-step prediction conditional probability is considered instead of the one step-ahead prediction as usual. Thus, the state space model parameters A, B, C, D can be understood as the multistep prediction probability parameters. In our recursive identification method, we consider the following conditional covariance matrix

$$P_F = \mathrm{cov}\left(y_{t+T_F}^{t+1} | u_{t+T_F}^{t+1}\right), \tag{4.22}$$

where y is the process output and u is the process input on the future horizon T_F. It will be shown that the covariance matrix P_F eigenstructure is related to the process model order. Defining a matrix N composed row-wise from the co-

variance matrix eigenvectors related to the non-zero (statistically) eigenvalues, the future y values are decomposed to a sum of initial condition x response and the future controls forced response

$$Ny_{t+T_F}^{t+1} = N \begin{pmatrix} CA \\ CA^2 \\ \vdots \\ CA^{T_F} \end{pmatrix} x(t) + N \begin{pmatrix} D & & & \\ CB & D & & \\ \vdots & & \ddots & \\ CA^{T_F-1}B & \cdots & CB & D \end{pmatrix} u_{t+T_F}^{t+1}. \tag{4.23}$$

In our algorithm, we use the equation (4.23) to define process state function $NOx(t)$, which can picked to be our model representation of the process state. Here O is the system observability matrix as in (4.23). In our contribution, it will be shown the conditional probability distribution function related to the covariance matrix in (4.22) has to be modified to force the model causality. In other words the process outputs $y(k)$ have to be independent on the future controls $u(k+j)$, $j > 0$. Techniques known from the Bayesian information fusion can be used to force the process model causality.

The state representation $NOx(t)$ is calculated using the following conditional probability

$$p\left(y_{t+T_F}^{t+1} | u_{t+T_F}^{t+1}, u_t^{t-T_P}, y_t^{t-T_P}\right), \tag{4.24}$$

where T_P means the past horizon. Once the state representation is evaluated, the model matrices A, B, C, D are found to be parameters of the following conditional probabilities.

$$p\left(NOx(t) | NOx(t-1), u(t-1)\right), \qquad p\left(y(t) | NOx(t), u(t)\right). \tag{4.25}$$

It will be shown all the conditional probability functions evaluations can be performed recursively using a numerically robust orthonormal matrix triangularizations, known from recursive least squares methods. The matrix N calculation can either be based on the SVD matrix decomposition algorithm, or a modified least squares algorithm [2].

Advanced distributed parameters process control [4] require a low dimensional model that can be provided by our algorithm. the use of our algorithm will be demonstrated on a glass furnace data. This example involves the prior information incorporation to the state space model parameters.

Bibliography

[1] OVERSCHEE, P. and MOOR, B. *Subspace Identification for Linear Systems*. Kluwer Academic Publishers, 1996.

[2] PACHNER, D., ROUBAL, J., and TRNKA, P. Total Least Squares not Using SVD. In *4th International Workshop on Total Least Squares and Errors-in-Variables Modeling*, Leuven, Belgium, 2006. ⟨http://homes.esat.kuleuven.be/∼imarkovs/workshop.html⟩.

[3] PETERKA, V. *Bayesian System Identification In: Trends and Progress in System Identification*. P. Eykhoff Ed., Pergamon Press. Oxford, UK, 1981.

[4] ROUBAL, J. and HAVLENA, V. Range Control MPC Approach for Two-Dimensional System. In *Proceedings of 16th IFAC World Congress*, Prague, Czech Republic, 2005. (on CD).

[5] TULLEKEN, J. A. F. Grey-box Modelling and Identification Using Physical Knowledge and Bayesian Techniques. *Automatica*, 29(2):285–308, 1993.

About the authors Dan Pachner was born in Prague, Czechoslovakia in 1972. He received the Ph.D. degree in Control Engineering at Czech Technical University in Prague, Faculty of Electrical Engineering in February 2003. He is currently an assistant professor at the Department of Control Engineering, Czech Technical University in Prague. His research deals with the fault detection adaptive control and system identification.

Jirka Roubal was born in Louny, Czechoslovakia in 1978. He received the Ing. (MCs.) degree in Technical Cybernetics at Czech Technical University in Prague, Faculty of Electrical Engineering in February 2002. He is currently

a Ph.D. student at the Department of Control Engineering, Czech Technical University in Prague. His research deals with the predictive control of the distributed parameter processes described by the lumped inputs and the distributed output models.

Pavel Trnka was born in Psek, Czechoslovakia in 1980. He received the Ing. (MCs.) degree in Technical Cybernetics at Czech Technical University in Prague, Faculty of Electrical Engineering in February 2004. He is currently a Ph.D. student at the Department of Control Engineering, Czech Technical University in Prague. His research deals with the subspace identification methods.

4.30 Comparison of the joint output method and the sample maximum likelihood method in errors-in-variables identification

Mei Hong, Torsten Söderström, Johan Schoukens and Rik Pintelon,
mhong@it.uu.se, ts@it.uu.se, {Johan.Schoukens,Rik.Pintelon}@vub.ac.be

Introduction In errors-in-variables (EIV) identification, both the Joint Output (JO) method [1] and the Sampled Maximum Likelihood (SML) method [2] are attractive estimators because they can handle general noise conditions and give high accuracy. In the JO approach, the EIV system is regarded as a multivariable system with a two–dimensional output vector and three mutually uncorrelated white noise sources. By converting the model into the innovations form, a maximum likelihood JO(ML) method can be applied to give consistent parameter estimates. The SML method is a frequency-domain identification approach where the exact covariance matrices of the disturbing noise is replaced by the sampled covariance matrices, which are calculated from a small number (M) of independent, repeated experiments. Compared to the frequency domain Maximum Likelihood method assuming known noise variances, the loss in efficiency of SML is $(M-2)/(M-3)$, which is not large even for small M.

An essential assumption for the JO(ML) method is that the noise free input signal is stationary with rational spectrum, so that it can be described as an ARMA process. In the JO(ML) method, the input and output noises are also assumed to be described as ARMA processes with a small set of parameters. The SML method works under arbitrary true input signals and noise conditions, but with another important assumption: the noise-free signal is periodic. In general, both the JO(ML) method and the SML method can give good estimates but work under different experimental situations. If there is a condition suitable for both approaches, which method can give better estimation accuracy? It is of interest to investigate the relation between JO(ML) and SML methods.

In this paper, we compare these two methods by simulation under different cases, such as, assorted dynamic systems with different orders, different input signals, white or colored input output noises, varied signal-to-noise ratios (SNR) etc. Based on these results, more complicated theoretical comparisons might be attempted in the future.

Results and discussion Assume NM periodic data are available, where M is the number of periods and N is the number of data points in each period. Also assume that in each period the noise-free input signal is a realization of a stationary process.

The JO(ML) method uses all data points and assumes that the input signal is an ARMA process but does not exploit that the data are periodic. However, the SML method uses the periodic information but disregards that the input signal is an ARMA process. For comparison, we also give the Cramer-Rao lower bound based on known input-output noise variances and the period information but no assumptions on the input signals. See [2].

During the comparisons, the standard deviation (std) of each method are calculated from their theoretical covariance matrices of the estimation parameters, which have been proved to be well meet their relevant Monte-Carlo simulations. Details on these formulas can be found in [1], [4], [3] and [2].

Firstly, we analyze the effect of different order systems with white input and white output noise. Comparison results shows that for low order systems, the estimation accuracy of SML and JO(ML) method are quite similar. See Figure 1. While for high order systems, for those regions where the signal-to-noise ratios (SNR) is poor, the std of JO(ML) method is larger than that of the SML method. The smaller the SNR, the more distinct this phenomenon becomes. New comparison results under the same condition are shown except adding the periodic information to JO(ML) by

simple averaging the data over the M periods. The difference of JO(ML) and SML estimation results in the low SNR area has disappeared. According to these results, it seems that using the periodic information is important for high dynamic system especially when SNR is low. Furthermore, different input signals have been tried for various systems. The results show that for high order systems the input signal is more important than for low order systems.

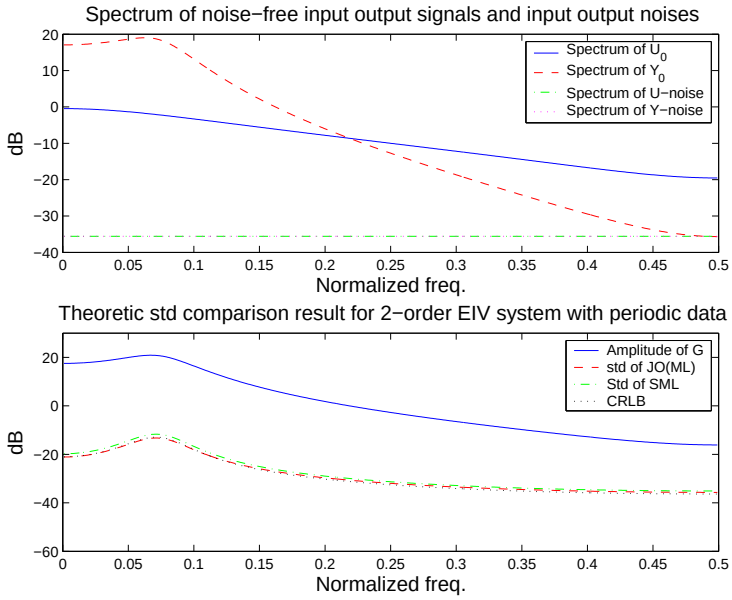


Figure 4.2: Spectrum of noise free signals and noises (up), and comparison of standard deviation of JO(ML) and SML (down) for a second order system white measurement noise.

Several examples with colored measurement noises were also tried. They give similar results as for white noise cases, i.e., if the periodic information has not been considered, the std of the JO(ML) method is similar to that of the SML method except at the very low SNR regions, where the estimation uncertainty of JO(ML) is larger than that of SML. This phenomenon is more pronounced for high order dynamic systems.

Bibliography

[1] T. Söderström. Identification of stochastic linear systems in presence of input noise. *Automatica*, 17:713–725, 1981.

[2] J. Schoukens, R. Pintelon, G. Vandersteen and P. Guillaume. Frequency domain system identification using non-parametric noise models estimated from a small number of data sets. *Automatica*, 33:1073–1086, 1997.

[3] R. Pintelon and J. Schoukens. System Identification. A Frequency Domain Approach. *IEEE Press*, New York, NY, USA, 2001.

[4] T. Söderström. On computing the Cramer-Rao bound and covariance matrices for PEM estimates in linear state space models. *Proc. 14th IFAC Symposium on System Identification*, Newcastle, Australia, March, 29-31, 2006.

About the author Mei Hong received her B.Sc. in Electronic Engineering from Xidian University, China in 1994 and M.Sc. from Chalmers University, Sweden in 2003. Now she is pursuing her Ph.D degree in Division of Systems and Control, Department of Information Technology, Uppsala University. Her current research interests are in system identification especially for errors-in-variables systems.

4.31 A fast algorithm for solving the Sylvester structured total least squares problem

Bingyu Li and Lihong Zhi, Key Lab of Mathematics Mechanization, AMSS, Beijing 100080 China,
{liby,lzhi}@mmrc.iss.ac.cn

Given two univariate polynomials f, $g \in \mathbb{R}[x]$ with $\deg(f) = m$ and $\deg(g) = n$. For a positive integer k with $k \leq \min(m,n)$, we wish to find perturbations $\triangle f$, $\triangle g$ with $\deg(\triangle f) \leq m$ and $\deg(\triangle g) \leq n$, which solve the following approximate GCD problem:

$$\min_{\triangle f, \triangle g} \|\triangle f\|_2^2 + \|\triangle g\|_2^2 \quad \text{such that} \quad \deg(\gcd(f+\triangle f, g+\triangle g)) \geq k. \tag{4.26}$$

Suppose $S(f,g)$ is the Sylvester matrix of f and g, the k-th Sylvester matrix S_k is a submatrix of S obtained by deleting the last $k-1$ rows of S and the last $k-1$ columns of coefficients of f and g separately in S. We know that $\deg(\gcd(f,g)) \geq k$ if and only if dim Nullspace$(S_k) \geq 1$ [3]. In [9, 8, 5, 6, 1, 10] structured total least squares(STLS) algorithms have been applied to solve the approximate GCD problem (4.26). The implementation of STLS which does not exploit the displacement structure of Sylvester matrix has its complexity of cubic in the degrees of the input polynomials. In [8, 9], a fast implementation based on Structured Total Least Norm (STLN) [12, 11] for constructing a Sylvester matrix of given lower rank was proposed. It has a quadratic amount of complexity in the degrees of the input polynomials. However, due to the large penalty used in STLN, we have to apply the generalized Schur algorithm in [4, 2] to operate on an ill-conditioned matrix. In the following, we describe a new fast algorithm which generalizes the method in [7] to solve the Sylvester structured total least squares problem.

Suppose $S_k = [\mathbf{b}_k, A_k]$, where $\mathbf{b}_k$ is the first column of S_k and A_k is a matrix consisting of the remaining columns of S_k. The Sylvester-structure preserving perturbation $[\mathbf{h}_k, E_k]$ of S_k can be represented by a vector $\mathbf{z} \in \mathbb{R}^{m+n+2}$ which contains the unknown coefficients of polynomials of degrees m and n. We solve the equality-constrained least squares problem:

$$\min_{\mathbf{z},\mathbf{x}} \|\mathbf{z}\|_2, \text{ subject to } \mathbf{r}(\mathbf{z},\mathbf{x}) = \mathbf{b}_k + \mathbf{h}_k - (A_k + E_k)\mathbf{x} = 0. \tag{4.27}$$

The above minimization problem can be transformed into unconstrained optimization by the Lagrangian:

$$L(\mathbf{z},\mathbf{x},\lambda) = 0.5\mathbf{z}^T\mathbf{z} - \lambda^T(\mathbf{b}_k + \mathbf{h}_k - A_k\mathbf{x} - E_k\mathbf{x}). \tag{4.28}$$

Construct a Sylvester structured matrix X_k such that

$$X_k\mathbf{z} = E_k\mathbf{x} - \mathbf{h}_k = [\mathbf{h}_k, E_k]\begin{bmatrix} -1 \\ \mathbf{x} \end{bmatrix}. \tag{4.29}$$

Apply the Newton method on the Lagrangian L and yields:

$$M\begin{bmatrix} \triangle\mathbf{z} \\ \triangle\mathbf{x} \\ \triangle\lambda \end{bmatrix} = \begin{bmatrix} I_{t_1\times t_1} & 0_{t_1\times t_2} & X_k^T \\ 0_{t_2\times t_1} & 0_{t_2\times t_2} & A_k^T + E_k^T \\ X_k & A_k + E_k & 0_{s\times s} \end{bmatrix}\begin{bmatrix} \triangle\mathbf{z} \\ \triangle\mathbf{x} \\ \triangle\lambda \end{bmatrix} = -\begin{bmatrix} \mathbf{g} + J^T\lambda \\ \mathbf{r}(\mathbf{z},\mathbf{x}) \end{bmatrix}, \tag{4.30}$$

where $t_1 = m+n+2, t_2 = m+n-2k+1, s = m+n-k+1, J = [X_k, A_k+E_k]$, and $\mathbf{g}$ is the gradient of the objective function in (4.27).

Since M is not strongly regular [7], a permutation matrix P is considered to transform M as:

$$\overline{M} = PMP^T = \begin{bmatrix} I_{t_1\times t_1} & X_k^T & 0_{t_1\times t_2} \\ X_k & 0_{s\times s} & A_k + E_k \\ 0_{t_2\times t_1} & A_k^T + E_k^T & 0_{t_2\times t_2} \end{bmatrix}. \tag{4.31}$$

The Schur complement of $\overline{M}$ w.r.t. the block $I_{t_1\times t_1}$, is:

$$\widehat{M} = \begin{bmatrix} -X_kX_k^T & A_k + E_k \\ A_k^T + E_k^T & 0_{t_2\times t_2} \end{bmatrix}. \tag{4.32}$$

Theorem 3 *$\widehat{M}$ is a strongly regular matrix of displacement rank at most* 8 *with respect to the displacement defined by $\widehat{M} - F\widehat{M}F^T$, where*

$$F = \begin{bmatrix} Z_{m+n-k+1} & \\ & Z_{m+n-2k+1} \end{bmatrix},$$

here Z_i is a lower shift matrix of order i.

By applying the generalized Schur algorithm with: $m+n-k+1$ negative steps and $m+n-2k+1$ positive steps, we obtain the LDL^T factorization of $\widehat{M}$ stably. Consequently, the solution to (4.30) can be obtained with $O((2m+2n-3k+2)^2)$ flops. We update $\mathbf{z} = \mathbf{z} + \triangle\mathbf{z}$, $\mathbf{x} = \mathbf{x} + \triangle\mathbf{x}$, $\lambda = \lambda + \triangle\lambda$ until $\|\triangle\mathbf{z}\|_2 \leq tol$, which is a given tolerance.

We have implemented the fast algorithm in Maple and applied it to compute the approximate GCD of univariate polynomials. The experiments show the efficiency and stability of the new fast algorithm compared with that of the STLN based fast algorithm [8, 9].

Bibliography

[1] Botting, Brad, Giesbrecht, Mark, and May, John. Using Riemannian SVD for problems in approximate algebra. In [13], pages 209–219.

[2] Chandrasekaran, S. and Sayed, A.H. A fast stable solver for nonsymmetric toeplitz and quasi-toeplitz systems of linear equations. *SIMAX*, 19(1):107–139, 1998.

[3] Gao, Shuhong, Kaltofen, Erich, May, John P., Yang, Zhengfeng, and Zhi, Lihong. Approximate factorization of multivariate polynomials via differential equations. In Gutierrez, Jaime, editor, *ISSAC 2004 Proc. 2004 Internat. Symp. Symbolic Algebraic Comput.*, pages 167–174, New York, N. Y., 2004. ACM Press. ISBN 1-58113-827-X.

[4] Kailath, Thomas and Sayed, Ali H. Displacement structure: theory and applications. *SIAM Review*, 37(3):297–386, 1995.

[5] Kaltofen, Erich, Yang, Zhengfeng, and Zhi, Lihong. Structured low rank approximation of a Sylvester matrix. In [13], pages 188–201.

[6] Kaltofen, Erich, Yang, Zhengfeng, and Zhi, Lihong. Structured low rank approximation of a generalized Sylvester matrix. Manuscript, 2005b. 14 pages. Abstract at the Asian Symposium on Computer Mathematics ASCM 2005, 4 pages.

[7] Lemmerling, Phillipe, Mastronardi, Nicola, and Van Huffel, Sabine. Fast algorithm for solving the Hankel/Toeplitz Structured Total Least Squares problem. *Numerical Algorithms*, 23:371–392, 2000.

[8] Li, Bingyu, Liu, Zhuojun, and Zhi, Lihong. Fast low rank approximation of a Sylvester matrix. In [13], pages 202–208.

[9] Li, Bingyu, Yang, Zhengfeng, and Zhi, Lihong. Fast low rank approximation of a Sylvester matrix by structured total least norm. *J. JSSAC (Japan Society for Symbolic and Algebraic Computation)*, 11(3,4):165–174, 2005b.

[10] Markovsky, Ivan and Huffel, Sabine Van. An algorithm for approximate common divisor computation. Internal Report 05-248, ESAT-SISTA, K.U.Leuven (Leuven, Belgium), 2005.

[11] Park, H., Zhang, L., and Rosen, J. B. Low rank approximation of a Hankel matrix by structured total least norm. *BIT*, 39(4):757–779, 1999.

[12] Rosen, J. B., Park, H., and Glick, J. Total least norm formulation and solution for structured problems. *SIAM J. Matrix Anal. Applic.*, 17(1):110–128, 1996.

[13] Wang, Dongming and Zhi, Lihong, editors. *Proc. 2005 International Workshop on Symbolic-Numeric*, July 2005. Distributed at the Workshop in Xi'an, China.

Bingyu Li received the M. Sc. degree in Applied Mathematics from Northeast Normal University, China, in 2003. Since then, she has become a Ph.D. student in Computer Algebra in Key Lab of Mathematics Mechanization, AMSS, China. The subject of her Ph.D. thesis is some fast algorithms related to approximate polynomial GCD.

Lihong Zhi received the B.S. degree in Mathematics from Peking University, Beijing, China, and the Dr.Sci. degree in Computer Algebra from the Institute of Systems Science, Academia Sinica, Beijing, China. She is an Associate Professor with the Institute of Systems Science, Academia Sinica. From 1998 to 2001, she was an Assistant Professor with the Department of Computer Science, Ehime University, Matsuyama, Japan. From 2001 to 2002, she did post-doctoral research at the Ontario Research Center for Computer Algebra, University of Western Ontario, London, Canada. Zhi's current research interests are fast and reliable algorithms and their implementation for hybrid numeric/symbolic computation.

4.32 Computational aspects of the geographically weighted regression

Andrzej Maćkiewicz and Barbara Maćkiewicz, Institute of Mathematics, Technical University of Poznań, Poland; mackiewi@sol.put.poznan.pl; Institute of Socio-Economic Geography and Spatial Management, Adam Mickiewicz University in Poznań, Poland; basic@amu.edu.pl

Geographically Weighted Regression (*GWR*) is a new and powerful method for the analysis of spatially varying relationships [2]. At the beginning of such an analysis it is assumed that we have defined m different *data points* in a given region of the country. Each i-th data point is uniquely defined by a pair of coordinates (u_i, v_i). We give this location a value y_i of the dependent variable y and values x_{ik} $(1 \le k \le n-1)$ of independent variables x_i.

Let us consider a *global regression model* written as $y_i = \beta_0 + \sum_{k=1}^{n-1} \beta_k x_{ik}$, where the dependent variable y is regressed over a set of independent variables x_i. Then, the classical regression equation can be written in a matrix form as an overdetermined (and usually not compatible) system of linear equations

$$y = X\beta, \tag{4.33}$$

where $X \in \mathbb{R}^{m\times n}$ is not necessarily sparse, $y \in \mathbb{R}^m$ and $\beta = (\beta_0, ..., \beta_{n-1})$ is the vector of parameters (constant over space) to be estimated. *GWR* extends this traditional regression framework by allowing local rather than global parameters to be estimated so that the *local model* for each individual i-th location (called the *i-th regression point*) is rewritten as $y_i = \beta_0(u_i, v_i) + \sum_{k=1}^{n-1} \beta_k(u_i, v_i) x_{ik}$, where $\beta_k(u_i, v_i)$ is the value of a continuous function $\beta_k(u, v)$ at this location. So, in this method we want to determine a matrix $B = (\beta_k(u_i, v_i)) \in \mathbb{R}^{m\times n}$ which consists of m sets of local parameters. The parameters in each i-th row of the above matrix B are determined from a weighted, overdetermined linear system

$$W_i y = W_i X \beta, \tag{4.34}$$

where a matrix $W_i = \mathrm{diag}(w_{i1}, w_{i2}, ..., w_{ik}, ..., w_{im})$ is an $m \times m$ spatial weighting matrix and $1 \ge w_{ik} > 0$ is the weight given to data point k in the calibration of the model for regression point i. The weight w_{ij} is usually calculated from the formula

$$w_{ij} = \exp\left(-\frac{1}{2}\left(\frac{d_{ij}}{b}\right)^2\right),$$

where d_{ij} is the distance between regression point i and data point j, and the parameter b (called *bandwidth*) is chosen so that weighting matrices W_i are well conditioned. Hence, data points closer to the regression point are weighted more heavily in the local regression than are data points farther away. If two regression points are close to each other, then the two corresponding weighted systems (4.34) are almost identical and should have very similar solutions. The spatial variation of each particular parameter estimate is analyzed graphically at the end of the process and compared with the behavior of the global solution of the system (4.33). In many applications this approach is much better suited to reality than an ordinary linear regession (related to the system 4.33).

Theoretically the system (4.33) and each of the systems (4.34) can be solved separately as an *LS* problem. As these systems are usually very big, this approach is hardly acceptable from the computational point of view. Moreover, the

matrices involved are often ill-conditioned and some kind of regularization must be performed in order to assure continuity of each parameter estimate over the region considered. To save the computation time the authors of the *GWR* method had recommend (in [2]) a normal equations approach as a method of choice for solving the systems (4.34) involved. This paper presents an alternative approach, still fast but more stable numerically. The new algorithm uses (only once) properly chosen Householder [3] postmultiplications in order to *"tri-orthogonalize"* [5] the coefficient matrix X. It is proved (on the basis of the Implicit Q Theorem and Krylov subspace methods) that these transformations provide an elegant way of extracting a well-conditioned *general core* subproblem both for the problem (4.33), and the correlated weighted problems (4.34). Next, a modified version of the *GMRES* algorithm is used to solve all of the weighted problems in the special *"dendrite like"* order (determined by the spatial distribution of the data points). As the special and effective preconditioning technique is implemented, and the solution of one system is used to start the next iteration process, the whole matrix B is determined quickly and fulfils the desired expectations. Sensitivity analysis is included.

This approach has been used successfully used for land price analysis in Poland. Possible new applications to signal and image processing (in a TLS approach) are mentioned. Computer programs (written in Fortran 95 and Matlab) will be delivered on request by the authors.

Bibliography

[1] Å. Björck, Bidiagonal Decomposition and Least Squares Problems. *BIT*, 44(2): 001–012, 2004.

[2] S. Fotheringham, Ch. Brunsdon and M. Charlton. Geographically Weighted Regression. *J. Wiley*, 2002.

[3] G,H. Golub and C.F. Van Loan. Matrix Computations. Third Edition. *Johns Hopkins University Press*, Baltimore, 1996.

[4] R. J. Hanson, and D. R. Kincaid. Notes on GMRES Algorithm Organization. Preprint submitted to Elsevier Science. 2005.

[5] A. Maćkiewicz and R. Rhalia. An efficient orthogonalization method for accurate computation of singular values. in: M. Doblare and J.M. Correas (Eds.) Metodos Numericos en Ingenieria, Vol.2, *SEMNI*, (Barcelona-Spain) : 1371–1382, 1996.

[6] Y. Saad. Iterative Methods for Sparse Linear Systems. *SIAM*, 2003.

About the authors Andrzej Mackiewicz was born in Czaplinek, Poland, on October 13, 1947. He married on September 25, 1974, with Janina Wrblewska (Matematician and Computer Programmer) and has 3 children: Barbara, Mikolaj (died on March 11, 2006) and Kasia. Doctoral degree in Numerical and Functional Analysis from Adam Mickiewicz University (1975). The author of three books (printed in Polish). Since 1987 Professor of Computational Mathematics, Technical University of Poznan; Poland. Visiting Professor at UBI (Portugal) 1992-1994 and UPC (Spain) 1995-1997. Research Interests: Numerical Analysis and Optimization, Operations Research, Numerical Linear Algebra, Algorithms, Transportation Research, Spatial Management.

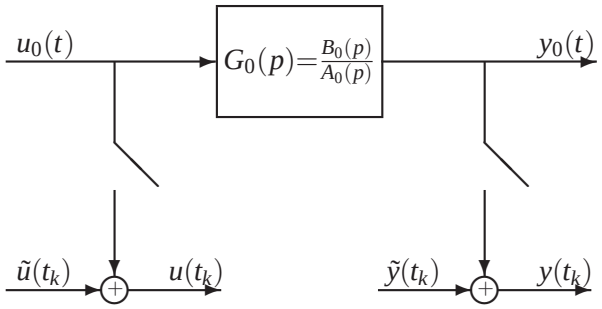
Barbara Mackiewicz was born in Poznan, Poland, on October 04, 1975. Doctoral degree in Socio-Economic Geography from Adam Mickiewicz University (2005). Assistant Professor since 2005. Interested in Statistics, Economics and Spatial Management.

4.33 A cumulant statistic–based method for continuous–time errors–in–variables model identification

Stéphane Thil, Hugues Garnier, Marion Gilson
Centre de Recherche en Automatique de Nancy – CRAN UMR 7039 CNRS-UHP-INPL,
Université Henri Poincaré, Nancy 1, BP 239 – 54506 Vandœuvre-lès-Nancy Cedex – France
firstname.lastname@cran.uhp-nancy.fr

Continuous-time errors-in-variables system identification

In this paper the continuous-time errors-in-variables model depicted in the opposite Figure is considered.



Errors-in-variables (EIV) models, where uncertainties or measurement noises are present on both input and output observations, play an important role when the identification purpose is the determination of the inner laws that describe the process, rather than the prediction of its future behavior. Numerous scientific disciplines use such EIV models, including time series modeling, array
signal processing for direction-of-arrival estimation, blind channel equalization, multivariate calibration in analytical chemistry, image processing, or environmental modeling [9].

Furthermore, in many areas of science and engineering, the identified dynamic models should be physically meaningful. As a result, there is a need for modeling approaches that are able to yield directly from the sampled data efficiently parameterized (parsimonious) continuous-time models that have clear physical interpretations. The attention in the system identification community was almost completely focused on the discrete-time model identification techniques until recently. The last decade has indeed witnessed considerable development in continuous-time approaches to system identification from sampled data (see [5] and [8, 3] for more recent references).

The goal of this paper is to present an approach for continuous-time modeling that can take into account colored measurement noise in both input and output observations. Many methods have been proposed to solve the EIV problem in discrete-time, whereas in continuous-time it is relatively unexplored. An overview of the main discrete-time methods can be found in [6]. Regarding the continuous-time, an approach has been recently proposed in [4], assuming the noises contaminating the data to be white.

Unless we impose certain assumptions on the signal and noise models, it is well-known that the general EIV model is not uniquely identifiable from second order statistics [1]. Although that problem can be overcome by adding supplementary conditions, EIV models suffer from this lack of identifiability. This motivates the approaches based on higher-order statistics.

Higher-order statistics The proposed methods are based upon the third-order cumulants; their main properties are quickly recalled. Some statistical assumptions on the noise-free input signal and on the noises are necessary: the probability density function of the input signal is assumed to be non-symmetric, whereas the noises are assumed to be symmetrically distributed. The differential equation of the system is then satisfied by the third-order cumulants [2]

$$C_{uyu}(\tau_1,\tau_2) = G(p,\theta)C_{uuu}(\tau_1,\tau_2) = \frac{B(p)}{A(p)}C_{uuu}(\tau_1,\tau_2) \tag{4.35}$$

where C_{uuu}, C_{uyu} are the third-order (cross-)cumulants and $G(p,\theta)$ is the parametrization of the real system. The noise-cancellation property of the third-order cumulants implies that equation (4.35) is (asymptotically) noise-free, consequently the simple least-squares method gives consistant estimates. However, when only a finite data record is available, errors appear in both left- and right-hand side of equation (4.35).
To obtain estimates of the parameter vector, two possibilities are then considered.

Linear regression. To estimate the parameter vector θ, the linear regression theory can be applied to equation (4.35). Minimizing the following equation error

$$e_1(\tau_1,\tau_2) = A(p)C_{uyu}(\tau_1,\tau_2) - B(p)C_{uuu}(\tau_1,\tau_2) \tag{4.36}$$

two criterion-based estimators are derived: the simple LS estimator and the TLS estimator.

Non-linear regression: the Steiglitz-McBride algorithm. From equation (4.35), θ can also be derived by minimizing the following output error

$$e_2(\tau_1,\tau_2) = C_{uyu}(\tau_1,\tau_2) - \frac{B(p)}{A(p)}C_{uuu}(\tau_1,\tau_2) \tag{4.37}$$

This output error is non-linear in the parameters. To avoid the recourse to non-linear optimization, following the work of J.M.M. Anderson in discrete-time [2], the Steiglitz-McBride [7] algorithm is used. An equation error is consequently defined, converging towards the output error (4.37) in an iterative fashion. Another criterion-based estimator is then defined.

The state variable filter. One of the key points in continuous-time system identification is how to handle time-derivation. Here the cumulants time-derivatives are needed and to estimate them the state variable filter [10] is utilized: in a first step the derivatives of the input/output signals are estimated, then the cumulants derivatives are computed from these estimates.

Bibliography

[1] B.D.O. Anderson and M. Deistler. Identifiability in dynamic errors-in-variables models. *Journal of Time Series Analysis*, 5(1):1–13, 1984.

[2] J.M.M. Anderson and G.B. Giannakis. Noisy input/output system identification using cumulants and the Steiglitz-McBride algorithm. *IEEE Transactions on Signal Processing*, 44(2):407–409, May 1996.

[3] H. Garnier, M. Mensler, and A. Richard. Continuous-time model identification from sampled data. Implementation issues and performance evaluation. *International Journal of Control*, 76(13):1337–1357, 2003.

[4] K. Mahata and H. Garnier. Direct identification of continous-time errors-in-variables models. In *16th World Congress of the IFAC*, Prague, Czech Republic, July 2005.

[5] N.K. Sinha and G.P. Rao. *Identification of continuous-time systems. Methodology and computer implementation*. Kluwer Academic Publishers, Dordrecht, 1991.

[6] T. Söderström, U. Soverini, and K. Mahata. Perspectives on errors-in-variables estimation for dynamic systems. *Signal Processing*, 82:1139–1154, 2002.

[7] K. Steiglitz and L.E. McBride. A technique for the identification of linear systems. *IEEE Transactions on Signal Processing*, 43(10):2386–2397, October 1965.

[8] H. Unbehauen and G.P. Rao. A review of identification in continuous-time systems. *Annual reviews in Control*, 22:145–171, 1998.

[9] S. Van Huffel. *Total least squares and errors-in-variables modeling*. Kluwer Academic Publishers, 2002.

[10] P.C. Young. Parameter estimation for continuous–time models - a survey. *Automatica*, 17(1):23–39, 1981.

About the author Thil Stéphane received his M.S. degree in Mathematics from Metz University (France) in 2003, and his Master of Research degree in Automatic Control and Signal Processing from Henri Poincaré University (France) in 2004. He is currently a Ph.D. student in Henri Poincaré University. His research interests focus on system identification (errors-in-variables, discrete- and continuous-time, Bayesian methods, closed-loop).

4.34 New errors-in-variables unscented Kalman filter algorithms with enhancements

Benoit Vinsonneau, David P. Goodall, Keith J. Burnham
Control Theory and Applications Centre, Coventry University, UK, b.vinsonneau@coventry.ac.uk

The errors-in-variables (EIV) filtering problem has only recently been formulated and solved [4] for systems that are assumed to be linear and time invariant (LTI). EIV algorithms derived for single input single output (SISO) systems are described in [1]. Extensions to the multiple input multiple output (MIMO) case have been proposed in both the stochastic [2, 3] and deterministic [9] contexts. Based on one of the proposed algorithms in [1], termed here the EIV Kalman filter (EIV-KF), a parallel development proposed in [10] has extended the approach to encompass linear time varying (LTV) systems. This is achieved with a new EIV extended Kalman filter (EIV-EKF), which estimates a set of compensating parameters within an extended state. As well as handling the LTV case, the EIV-EKF algorithm has been shown to provide a superior performance over the EIV-KF for the LTI case when the system parameters are not known exactly.

Unfortunatelly, the performance of the EIV-EKF decreases when applied to a system which exhibits severe nonlinearity, due probably to the limitations imposed when utilising the linearisation step in the prediction stage of both state and error covariance; noting that only the first order terms of the Taylor series are described accurately. A solution to such a problem, that has been proposed in 'classical' Kalman filtering, is to make use of an unscented transformation (UT) [5, 6, 7, 8]. Essentially, the UT provides a method of capturing the statistics of a random vector which undergoes a nonlinear transformation [8]. For non-Gaussian inputs, approximations are accurate to at least the second order, with the accuracy of the third and higher moments being determined by the choice of hyperparameters.

Prompted by the potential advantages of adopting a UT approach, the poster presents the results of a recent study utilising a new EIV-UT-KF algorithm. Whilst the work presented may be best described as 'work-in-progress', the novelty of the results and timeliness of the work is considered to be ideal for this Workshop. When applied to a nonlinear system having an assumed structure and known parameters, preliminary results of a Monte Carlo analysis have shown that a developed EIV-UT-KF algorithm provides improvement over the existing EIV-KF and EIV-EKF algorithms.

Bibliography

[1] R. Diversi, R. Guidorzi, and U. Soverini. Algorithms for optimal errors-in-variables filtering. *Systems & Control Letters*, 48:1–13, 2003.

[2] R. Diversi, R. Guidorzi, and U. Soverini. Kalman fitlering in symmetrical noise environments. In *Proc.* 11^{th} *IEEE Mediterranean Conf. on Control and Automation*, Rhodes, Greece, 2003.

[3] R. Diversi, R. Guidorzi, and U. Soverini. Optimal errors-in-variables filtering in the MIMO case. In *Proc. DVD-ROM* 16^{th} *IFAC World Congress*, Prague, Czech Republic, 2005.

[4] R. Guidorzi, R. Diversi, and U. Soverini. Optimal errors-in-variables filtering. *Automatica*, 39:281–289, 2003.

[5] S. Haykin. *Kalman filtering and neural networks*. ISBN 0-471-36998-5. John Wiley & Sons, New York, 2001.

[6] S. J. Julier and J. K. Uhlmann. A new extension of the Kalman filter to nonlinear systems. In *Int. Symp. Aerospace/Defense Sensing, Simulation and Controls*, Orlando, USA, 1997.

[7] S. J. Julier and J. K. Uhlmann. Reduced sigma point filters for the propagation of means and covariances through nonlinear transformations. In *Proc. American Control Conf.*, pages 887–892, Anchorage, Alaska, USA, 2002.

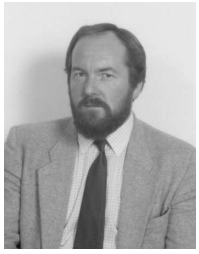
[8] S. J. Julier and J. K. Uhlmann. Unscented filtering and nonlinear estimation. *IEEE Review*, 92(3), 2004.

[9] I. Markovsky and B. De Moor. Linear dynamic filtering with noisy input and output. In *Proc.* 13^{th} *IFAC Symp. on System Identification*, pages 1749–1754, Rotterdam, The Netherlands, 2003.

[10] B. Vinsonneau, D. P. Goodall, and K. J. Burnham. Errors-in-variables extended Kalman filter. In *Proc. IAR & ACD Conf.*, pages 217–222, Mulhouse, France, 2005.

Benoit Vinsonneau was born in Angers, France, in 1981. He received the Degrees of BEng (Hons.) in Engineering in 2001 and MSc (Distinction) in Informatics and Control in 2002, from Coventry University, UK, and a DEA in Control, Signal Processing and Communications in 2003 from Henri Poincaré University Nancy 1, France. Currently, he is a PhD student in Control Engineering and based within the Control Theory and Applications Centre (CTAC), Faculty of Engineering and Computing, Coventry University, UK. His main research interests are in the fields of signal processing, system identification, control engineering and related applications to socio-economic, environmental and industrial processes. He is a student member of the Institute of Electronics and Electrical Engineering (IEEE).

David P. Goodall obtained a BSc (Hons.) degree in Mathematics from Hatfield Polytechnic, UK, in 1970, and MSc degree in Applied Functional Analysis from the University of Dundee, UK, in 1975, and a PhD degree from the University of Bath, UK, in 1989. His main research interests are in the fields of deterministic feedback stabilization of uncertain nonlinear control systems, stability of functional dynamical systems and identification of nonlinear systems. He has published widely in a number of journals and is a co-author of a book in the area of Signal Processing. Currently, he is a Reader in Dynamical Systems and Control based within the Faculty of Engineering and Computing at Coventry University, UK and, in addition, is an Associate Editor for the International Journal of Systems Science.



Keith J. Burnham has been Professor of Industrial Control Systems, and Director of the Control Theory and Applications Centre (CTAC), Faculty of Engineering and Computing, Coventry University, UK, since 1999. Keith Burnham obtained his BSc (Mathematics), MSc (Control Engineering) and PhD (Adaptive Control) at Coventry University in 1981, 1984 and 1991, respectively. His main research interests are in the fields of adaptive control for nonlinear industrial systems and bilinear systems modelling and control. He is regularly consulted by industrial organisations to provide advice in areas of advanced algorithm development for control and condition monitoring. He is currently a member of the Editorial Board of the Transactions of the Institute of Measurement and Control (InstMC), and a member of the InstMC Systems Control Technology Panel, a member of the Informatics and Control Working Group of the Institution of Mechanical Engineers (IMechE), a member of the Control and Automation Professional Network Advisory Board of the Institution of Electrical Engineers (IEE), a member of the Automotive Control Technical Committee of the International Federation of Automatic Control (IFAC). He is also a corporate member of the IEE, InstMC, and Institute of Mathematics and its Applications (IMA).

4.35 Algorithms for data fitting in error-in-variables models with linear and quadratic constraints

Burkhard Schaffrin, Department of Geological Science, The Ohio State University, Columbus, OH 43210-1275
Schaffrin.1@osu.edu

Yaron A. Felus, Surveying Engineering Department, Ferris State University, Big-Rapids, MI, 49307-2291,
felusy@ferris.edu

Fitting a surface to a given set of measurements is an essential function for engineers and geodesists, also known as trend analysis [7],[1],[4]. This technique uses Least-Squares (LS) adjustment to estimate the parameters (ξ) of a polynomial surface within a linear model ($y = A\xi + e$) that includes the vector of observed attribute values (y), a vector of normally distributed errors e, and a matrix of variables A, constructed from the geographical locations. However,

in this linear model the matrix of variables A is considered as fixed or error-free. This is not the case in many physical situations where errors exist both in the vector of attributes (y) and in the geographical locations matrix (A). The Total Least-Squares (TLS) approach as applied to the Error-in-Variables model is the proper method to treat problems where all the data are affected by random errors. The traditional approach to solve the TLS problem [2],[6] utilizes the Eckart-Young-Mirsky theorem to derive the best estimate $[\hat{A},\hat{y}]$ for the augmented matrix by setting the minimum eigenvalue of the Singular Value Decomposition of $[A; y]$ to zero.

A different approach to solve the TLS problem is based on the Euler-Lagrange theorem [4],[3] and follows an optimization process that minimizes a target function. In this contribution, the Euler-Lagrange approach is first used to solve an enhanced TLS problem that includes linear constraints as follows:

$$\begin{array}{lll} y-e=(A-E)\cdot\xi, & & rk[A]=m<n, \\ \kappa_0=K\cdot\xi, & & rk[K]=l\le m, \\ E\{[E,e]\}=0, & D\{vec[E,e]\}=\Sigma_0\otimes I_n, & C\{E,e\}=0, \end{array} \tag{4.38}$$

where E is a $n\times m$ random error matrix, $\Sigma_0=\sigma_0^2\cdot I_{m+1}$ is a *(m+1)×(m+1)* matrix with an unknown variance component (σ_0^2), K and κ_0 are $l\times m$ matrix, resp. $l\times 1$ vector of known values describing the constraints, m is the number of parameters, n the number of independent observations, and l the number of constraints. Denote $e_A=vec(E)\sim(0,\ \sigma_0^2\cdot I_m\otimes I_n)$; then the TLS Lagrange target function for model (4.38) is expressed by:

$$\Phi(e,e_A,\lambda,\mu,\xi):=e^Te+e_A^Te_A+2\lambda^T(y-A\xi-e+E\xi)-2\mu^T(\kappa_0-K\xi)=stationary \tag{4.39}$$

where λ *and* μ denote the $n\times 1$, *resp.* $l\times 1$ vectors of Lagrange multipliers. Using this target function, the following nonlinear normal equations are obtained:

$$\begin{bmatrix} A^TA & A^Ty & K^T \\ y^TA & y^Ty & \kappa_0^T \\ K & \kappa_0 & \hat{\nu}\cdot I \end{bmatrix}\cdot\begin{bmatrix} \hat{\xi} \\ -1 \\ \overline{\mu} \end{bmatrix}=\begin{bmatrix} \hat{\xi} \\ -1 \\ \overline{\mu} \end{bmatrix}\cdot\hat{\nu} \tag{4.40}$$

where $\hat{\nu}=(y-A\cdot\hat{\xi})^T\cdot(y-A\cdot\hat{\xi})/(1+\hat{\xi}^T\cdot\hat{\xi})$; all estimated parameters are marked with a bar $(\bar{\cdot})$ or a hat $(\hat{\cdot})$. Four algorithms have been developed to solve this TLS problem with linear constraints: One which is slow but guaranteed to converge, two which are fast but may only converge with good starting values, and a hybrid approach. A numerical example of fitting a surface to a set of surveyed points is used to demonstrate the computational efficiency of the algorithms and the accuracy improvement over the traditional model.

In some physical cases, the problem has quadratic constraints in addition to the linear constraints, for example when fitting an ellipsoidal surface through a list of measured points. The mathematical model in this case is:

$$\begin{array}{lll} y-e=(A-E)\cdot\xi, & & \\ \kappa_0=K\cdot\xi, & & \xi^T\cdot M\cdot\xi=\alpha^2, \\ E\{[E,e]\}=0, & D\{vec[E,e]\}=\Sigma_0\otimes I_n, & C\{E,e\}=0, \end{array} \tag{4.41}$$

using the same notations as before. Here M is a given $m\times m$ non-negative definite, symmetric matrix, and α^2 is a given constant. The TLS Lagrange target function for model (4.41) is:

$$\Phi(e,e_A,\lambda,\mu_1,\mu_2,\xi):=e^Te+e_A^Te_A+2\lambda^T(y-A\xi-e+E\xi)-2\mu_1^T(\kappa_0-K\xi)-\mu_2\left(\alpha^2-\xi^TM\xi\right)=stationary \tag{4.42}$$

leading to the following nonlinear normal equations:

$$\begin{bmatrix} A^TA+\overline{\mu}_2M & A^Ty & K^T \\ y^TA & y^Ty-\alpha^2\overline{\mu}_2 & \kappa_0^T \\ K & \kappa_0 & \hat{\nu}\cdot I_l \end{bmatrix}\cdot\begin{bmatrix} \hat{\xi} \\ -1 \\ \overline{\mu}_1 \end{bmatrix}=\begin{bmatrix} \hat{\xi} \\ -1 \\ \overline{\mu}_1 \end{bmatrix}\cdot\hat{\nu} \tag{4.43}$$

in conjunction with the modified secular equation:

$$\alpha^2=(A^Ty-K^T\overline{\mu}_1)^T(A^TA+\overline{\mu}_2M-\hat{\nu}\cdot I_m)^{-1}\cdot M\cdot(A^TA+\overline{\mu}_2M-\hat{\nu}\cdot I_m)^{-1}(A^Ty-K^T\overline{\mu}_1) \tag{4.44}$$

Using equations (4.43) and (4.44), an algorithm to solve the TLS problem with linear and quadratic constraints is derived and tested on a numerical example. The accuracy and computational efficiency of the newly developed algorithm is described, as well as some other open questions related to the current investigation.

Bibliography

[1] A.Y. Felus, and B. Schaffrin. A Total-Least Squares approach for Semivariogram modeling of aeromagnetic data. In "GIS and Spatial Analysis, IAMG'05", Q. Cheng and G. Bonham-Carter, Editors., China University of Geosciences Printing House: Beijing; pp. 215-221, 2005.

[2] H.G. Golub and F.C. van Loan. An analysis of the total least-squares problem, *SIAM Journal on Numerical Analysis*, 17 (6): 883-893, 1980.

[3] K. R. Koch. *Parameter Estimation and Hypothesis Testing in Linear Models* (second edition). Springer-Verlag: Berlin, 1999.

[4] B. Schaffrin, I. Lee, Y. Felus, and Y. Choi. Total Least-Squares (TLS) for Geodetic Straight-Line and Plane Adjustment. Submitted to the *Journal of Geodesy*, 2006.

[5] B. Schaffrin and Y. Felus. On Total Least-Squares Adjustment with Constraints. In "A Window on the Future of Geodesy ", F. Sansó, Editor. Series International Association of Geodesy symposia, Vol. 128, Springer–Verlag: Berlin. pp. 417-421, 2005.

[6] S. van Huffel and J. Vandewalle. *The Total Least-Squares Problem. Computational Aspects and Analysis*, SIAM Publication, Philadelphia/PA, 1991.

[7] G. S. Watson. Trend Surface Analysis. *Mathematical Geology*, 3(3):215-226, 1971.

Dr. Burkhard Schaffrin is an Associate Professor in the Geodetic Science Section of the Dept. of Geological Science at the Ohio State University. He holds a B.S. from Free University, Berlin (West), a M.S. from the University of Bonn, Germany (Mathematics), a M.S. from the University of Bonn, Germany (Geodetic Science & Surveying), a Ph.D. from the University of Bonn, Germany, and a Dr.-Ing.habil. from the University of Stuttgart, Germany (Theoretical Geodesy).

Dr. Yaron Felus is an Associate Professor in the Surveying Engineering Department at Ferris State University. He holds a B.S. from Ben-Gurion University, Israel (Electrical Engineering), a M.S. from the International Institute for Geo-Information Science and Earth Observation, Netherlands (Geomatics), and both M.S. and Ph.D. from the Ohio State University (Geodetic Science).

4.36 Inverse iteration for total least squares interpretted as combination of least squares problem and rotation of coordinates

Otakar Šprdlík, Department of Control Engineering , Faculty of Electrical Engineering, Czech Technical University in Prague, sprdlo1@control.felk.cvut.cz
Zdeněk Hurák, Center for Applied Cybernetics, Prague, z.hurak@c-a-k.cz

Inverse iteration method is one of the most powerful methods for computing selected eigenvectors (see for instance [1], section 7.6.1). It can also be used for finding numerical solution to total least squares problem (TLS) (see for instance [2], section 5.3). Recently, some new algorithms appeared in literature (see for instance [3] being presented at this workshop) that approach the numerical computation of TLS problem via iterative algorithm in which at eac

The statement of the total least squares problem is: For a given overdetermined system of linear equations $Ax \approx b$ where $A \in \mathbb{R}^{m\times n}$ and $m > n$, find $x \in \mathbb{R}^n$ satisfying $(A+\Delta A)x = b+\Delta b$ and minimizing $\|[\Delta A;\Delta b]\|_F$. It is well known

that under certain technical assumptions this minimizing vector $\hat{x}$ can be found in the direction of the right singular vector of $D = [\Delta A; \Delta b]$ corresponding to the least singular value.

Singular vector od D corresponding to the least singular value can be computed iteratively finding the eigenvector of $D^T D$ corresponding to its eigenvalue closest to zero by the inverse iteration method. Usage of QR factorization of D enables recursification of the algorithm. The following iteration can be followed

$$R_D^T R_D \theta^{(k)} = D^T D \theta^{(k)} = \begin{bmatrix} \hat{x}^{(k-1)} \\ -1 \end{bmatrix}, \qquad \begin{bmatrix} \hat{x}^{(k)} \\ -1 \end{bmatrix} = \frac{-\theta^{(k)}}{\theta_{n+1}^{(k)}},$$

where R_D denotes the triangular factor of D and $\hat{x}^{(k)}$ denotes TLS solution estimate at kth iteration. Further, it can be derived from R_D^{-1} viewed as an inverse of a block matrix that

$$\hat{x}^{(k+1)} = \tilde{x} + \frac{R_{Dn+1,n+1}}{1 + \tilde{x}^T \cdot \hat{x}^{(k)}} \cdot R_A^{-1} R_A^{-\mathrm{T}} \hat{x}^{(k)} \tag{4.45}$$

where $\tilde{x}$ denotes solution of least squares and R_A is the triangular QR factor of A. If the algorithm is initiated by $\hat{x}^{(0)} = 0$, result of the first iteration $\hat{x}^{(1)}$ is equal to the solution $\tilde{x}$ of ordinary least squares problem.

The core idea of TLS solved iteratively with the help of LS followed by rotation of coordinates is that under certain circumstances solutions to LS and TLS coincide. Aligning the coordinate systems can enable use of the LS algorithms for solving the TLS task. An intuitive idea is to rotate the original coordinate system at each step such that the new system is aligned with the solution of LS problem from the previous step.

Recall that the size of the augmented vector of TLS solution estimate $\hat{\theta}^{(k)}$ in kth iteration is $(n+1)$. In the first iteration the ordinary LS solution is assigned to $\hat{\theta}^{(1)}$, next an orthonormal basis of n-dimensional subspace orthogonal to the solution $\hat{\theta}^{(1)}$ is derived. The $(n+1)$th basis vector of the new coordinate system (= $(n+1)$th column of rotation matrix $Q^{(1)}$) is colinear with $\hat{\theta}^{(1)}$.

In each other iteration the data rotation is accomplished by mutliplication of D by $Q^{(k)}$ from the right. Standard LS problem is then solved for $\tilde{\theta}^{(k+1)}$ and estimate to the TLS problem is then obtained by rotation and normalization with respect to the last component

$$\underbrace{DQ^{(k)}}_{D^{(k+1)}} \tilde{\theta}^{(k+1)} \approx 0, \qquad \hat{\theta}^{(k+1)} = Q^{(k)} \tilde{\theta}^{(k+1)}, \qquad \begin{bmatrix} \hat{x}^{(k+1)} \\ -1 \end{bmatrix} = \frac{-\hat{\theta}^{(k+1)}}{\hat{\theta}_{n+1}^{(k+1)}},$$

where $Q^{(k)}$ is derived from $\hat{\theta}^{(k)}$ in the same way as $Q^{(1)}$ from $\hat{\theta}^{(1)}$. The normalization of $\hat{\theta}^{(k)}$ should be incorporated in the iterative scheme to provide numerical stability.

Considering $\hat{x}^{(0)}$ equals zero vector in the inverse iteration scheme, both algorithms start with $\hat{x}^{(1)} = \tilde{x}$. At the k-th step, the inverse iteration algorithm gives (with nonsingular $D^T D$)

$$\begin{bmatrix} \hat{x}^{(k)} \\ -1 \end{bmatrix} \propto (D^T D)^{-1} \begin{bmatrix} \hat{x}^{(k-1)} \\ -1 \end{bmatrix}, \text{ while the combination of LS and rotation yields } \begin{bmatrix} \hat{x}^{(k)} \\ -1 \end{bmatrix} \propto Q^{(k-1)} \begin{bmatrix} \tilde{x}^{(k)} \\ -1 \end{bmatrix}$$

where columns of $Q^{(k-1)}$ form an orthonormal basis of $\mathbb{R}^{(n+1)}$, last column of $Q^{(k-1)}$ is colinear with $[\hat{x}^{(k-1)}; -1]^T$, and $\tilde{x}^{(k)}$ is the solution of LS solved in the new coordinate frame $Q^{(k-1)}$.

With solution of LS in the new coordidate frame expressed using the fact explained below (4.45) the LS/rotation iteration became

$$\begin{bmatrix} \hat{x}^{(k)} \\ -1 \end{bmatrix} \propto \left(D^T D\right)^{-1} Q^{(k-1)} \begin{bmatrix} \mathbf{0} \\ -1 \end{bmatrix}$$

Vector $Q^{(k-1)}[\mathbf{0}; -1]^T$ is taken as the negative of the last column of $Q^{(k-1)}$, which is colinear with $[\hat{x}^{(k-1)}; -1]^T$ due to the construction of Q. Inverse iteration and combination of least squares and rotation of coordinate frame are thus equivalent.

We have shown inverse iteration scheme produces solutions of LS problem solved in a coordinate frame with the last basis vector colinear with the result of previous iteration and transformed back to the original coordinate frame. Results of iterations of the algorithm presented in [3] are the same as results of inverse iteration scheme initiated with all-zero $x^{(0)}$. The proven equilavence of the two algotithms – inverse iteration and combined LS and rotation – can perhaps bring some

Bibliography

[1] G. H. Golub and C. F. Van Loan. *Matrix Computations*. The John Hopkins University Press, 2nd edition, 1996.

[2] S. Van Huffel and J. Vandewalle. *The total least squares problem: Computational aspects and analysis*. SIAM, 1991.

[3] D. Pachner and J. Roubal and P. Trnka. Recursive total least squares not using SVD. In *The 4th International Workshop on Total Least Squares and Errors-in-Variables Modeling*, Leuven, Belgium, 2006. Accepted for presentation.

About the authors Otakar Šprdlík was born in 1980 in Jilemnice, Czech Republic. He received his Ing. (MSc.) degree in Technical Cybernetics at Czech Technical University in Prague, Faculty of Electrical Engineering in February 2005. He is currently a Ph.D. student at the Department of Control Engineering, Czech Technical University in Prague. His research deals with estimation and detection of human movement by means of inertial sensors.

Zdeněk Hurák was born in 1974 in Krnov, Czech Republic. He received his Ing. (MSc.) degree in Aviation Electrical Engineering at Air Force Faculty, Military Academy in Brno, Czech Republic in 1997 and Ph.D. degree in Control Engineering and Robotics at Czech Technical University in Prague, Faculty of Electrical Engineering in 2004. He is currently a research fellow at the Center for Applied Cybernetics. His research interests are robust control, convex optimization in control, numerical algorithm

4.37 Estimation in a multivariate errors-in-variables model with unknown noise variance ratio

Alexander Kukush, Kiev National Taras Shevchenko University,
E-mail: alexander_kukush@univ.kiev.ua

Consider a model $AX = B$, where $A \in \mathbb{R}^{m\times n}$, $X \in \mathbb{R}^{n\times p}$, and $B \in \mathbb{R}^{m\times p}$. The model means the following. For the true values, we have $\overline{A}X = \overline{B}$, where $\overline{A}, \overline{B}, X$ are non-random matrices. We observe $A = \overline{A} + \widetilde{A}$ and $B = \overline{B} + \widetilde{B}$, where $\widetilde{A}, \widetilde{B}$ are random matrices. We want to estimate X with fixed n and p and increasing m. This general model includes static models if the rows of $[\widetilde{A}, \widetilde{B}]$ are independent, and dynamic models if the matrices $[A, B]$ and $[\widetilde{A}, \widetilde{B}]$ are structured. Common assumption is that the covariance structure of $[\widetilde{A}, \widetilde{B}]$ is known up to a scalar factor, which leads to Elementwise–Weighted TLS Problem in the static case and to Structured TLS Problem in the dynamic case, see Kukush and Van Huffel (2004) and Kukush et al. (2005b).

In the present paper we assume that $[\widetilde{A}, \widetilde{B}]$ is partitioned into two uncorrelated blocks $[\widetilde{D_1}, \widetilde{D_2}]$ of sizes $m \times n_1$ and $m \times n_2$, and $E\widetilde{D}_k^T \widetilde{D}_k = \lambda_k^0 W_k,\ k = 1,2$, where W_k are known positive semidefinite matrices and λ_k^0 are two unknown positive scalars. In dynamic setting this means that both input and output noise variances are unknown.

To construct a consistent estimator $\widehat{X}$ we use a clustering assumption. The idea is due to Wald (1940), who studied a scalar model. We suppose that there are two copies of initial model $A(k)X = B(k),\ k = 1,2$, where $A(k) \in \mathbb{R}^{m_k\times n}, X \in \mathbb{R}^{n\times p}$, and $B(k) \in \mathbb{R}^{m_k\times p}$. Denote $\widetilde{D}_k = [\widetilde{A}(k), \widetilde{B}(k)]$ and let $\widetilde{d}_{ij}(k),\ 1 \le i \le m_k, 1 \le j \le n+p$ be the entries of $\widetilde{D}_k$, and $\widetilde{D}_k^T = [\widetilde{d}_1(k), \ldots, \widetilde{d}_m(k)]$. We list the conditions for consistency.

(i) $E\widetilde{d}_{ij}(k) = 0$, for all i, j, k.

(ii) $\exists\ \delta > 0\ \forall\ i, j, k : E|\widetilde{d}_{ij}(k)|^{4+\delta} \le const$.

(iii) Each of the sequences $\{\widetilde{d}_i(1), i \ge 1\}$ and $\{\widetilde{d}_i(2), i \ge 1\}$ is finite dependent.

(iv) $\exists\ n_1, 1 \le n_1 \le n+p-1\ \forall\ i \ge 1,\ 1 \le j \le n_1,\ n_1+1 \le l \le n+p,\ k = 1,2 : E\widetilde{d}_{ij}(k)\widetilde{d}_{il}(k) = 0$.

(v) $\widetilde{D}(k) = [\widetilde{D}_1(k), \widetilde{D}_2(k)], \widetilde{D}_1(k)$ of size $m \times n_1$, with $E\widetilde{D}_1^T(k)\widetilde{D}_2(k) = 0$, and $\widetilde{D}_j^T(k)\widetilde{D}_j(k) = \lambda_j^0 W_j(k),\ j = 1,2,\ k = 1,2$, where $W_j(k)$ are known positive semidefinite matices, and λ_j^0 are two unknown positive scalars.

(vi) For true matrices, $\|\overline{A}(k)\|_F/m_k \le const$, for all $m_k \ge 1,\ k = 1,2$.

(vii) $\liminf_{m_1,m_2\to\infty} \sigma_1(\overline{A}^T(1)\overline{A}(1)/m1 - \overline{A}^T(2)\overline{A}(2)/m2) > 0$, where $\sigma_1(c)$ is the smallest singular value of C.

Let $X_{ext} := [X^T, -I_p]^T = [X_1^T, X_2^T]^T, X_1 \in \mathbb{R}^{n_1\times p}$.

(viii) $\liminf_{m_1\to\infty} tr(X_j^T(W_j(1)/m_1)X_j) > 0$, for $j = 1,2$.

(ix) $W_j(1)/m_1 - W_j(2)/m_2 \to 0$, as $m_1, m_2 \to \infty$, $j = 1,2$.

(x) $\liminf_{m_k\to\infty} \lambda_{min}(\overline{A}^T(k)\overline{A}(k)/m_k) > 0$, for $k = 1,2$.

First we define the estimate of $\lambda^0 := (\lambda_1^0, \lambda_2^0)$. For $\lambda := (\lambda_1, \lambda_2) \in \mathbb{R}_+ \times \mathbb{R}_+$ introduce

$$\Psi^{(k)}(\lambda) = D^T(k)D(k) - \begin{bmatrix} \lambda_1 W_1(k) & 0 \\ 0 & \lambda_2 W_2(k) \end{bmatrix},$$

$$D(k) = [A(k), B(k)],$$

and let $\mu_{1k}(\lambda) \le \mu_{2k}(\lambda) \le \ldots \le \mu_{pk}(\lambda)$ be the p smallest eigenvalues of $\Psi^{(k)}(\lambda)$ with the corresponding orthonormal eigenvectors $f_{1k}(\lambda), \ldots, f_{pk}(\lambda)$, $L_{pk}(\lambda)$ be the span of $f_{1k}(\lambda), \ldots, f_{pk}(\lambda)$. The objective function is

$$Q(\lambda) = \sum_{k=1}^{2}\sum_{i=1}^{2} \mu_{ik}^2(\lambda) + c\|\sin\Theta(\lambda)\|^2.$$

Here $c > 0$ is a fixed constant and $\Theta(\lambda)$ is a diagonal matrix of canonical angles between $L_{p1}(\lambda), L_{p2}(\lambda)$, and $\sin\Theta(\lambda)$ is the diagonal matrix with diagonal elements the sines of these angles, see Steward and Sun (1990). For a fixed positive sequence $\{\varepsilon_q\}$, such that $\varepsilon_q \to 0$, as $q \to \infty$, an estimator $\widehat{\lambda} = (\widehat{\lambda}_1, \widehat{\lambda}_2) = \widehat{\lambda}(m_1, m_2)$ satisfies the inequality $Q(\widehat{\lambda}) \le \inf_{\lambda_1,\lambda_2\ge 0} Q(\lambda) + \varepsilon_q,\ q := min(m_1, m_2)$.

Now introduce a compound matrix $D_c = [D^T(1), D^T(2)]^T, W_{cj} := W_j(1) + W_j(2), j = 1,2$, and

$$\widehat{H} := D_c^T D_c - \begin{bmatrix} \widehat{\lambda}_1 W_{c1} & 0 \\ 0 & \widehat{\lambda}_2 W_{c2} \end{bmatrix}.$$

Let $L_p(\widehat{H})$ be the subspace spanned by the first p eigenvalues of $\widehat{H}$ corresponding to the smallest eigenvalues. An estimator $\widehat{X}$ is defined by the equality $L_p(\widehat{H}) = span(\widehat{z}_1, \ldots, \widehat{z}_p)$, where $[\widehat{X}^T - I_p]^T = [\widehat{z}_1, \ldots, \widehat{z}_p]$.

Theorem 4 *Under the conditions (i) to (x),* $\widehat{\lambda} \to \lambda^0$ *and* $\|\widehat{X} - X\|_F \to 0$, *as* $m_1, m_2 \to \infty$, *a.s.*

The results are applicable to system identification, with a turnover point in the input data. Simulation examples are discussed. Another estimator is proposed in Markovsky et al. (2006), that estimator is easier to compute, but its asymptotic properties are unclear. In case where nothing is known about covariance structure of the error matrix $[\widetilde{A}, \widetilde{B}]$, at least t clusters, $t = np$, are needed to estimate X consistently. The most of results are joint with Prof. S. Van Huffel and Dr. I. Markovsky (K.U.Leuven), see Kukush et al. (2005a) and Markovsky et al. (2006).

Bibliography

[1] A. Kukush and S. Van Huffel. Consistency of elementwise-weighted total least squares estimator in multivariate errors-in-variables model $AX = B$. Metrika, 59:75–97, 2004.

[2] A. Kukush, I. Markovsky, and S. Van Huffel. Estimation in linear multivariate measurement error model with clustering in the regression. Technical Report 05-170, Dept.EE, K.U.Leuven, 2005a.

[3] A. Kukush, I. Markovsky, and S. Van Huffel. Consistency of the structured total least squares estimator in a multivariate errors-in-variables model. Journal of Statistical Planning and Inference, 133:315–358, 2005b.

[4] I. Markovsky, A. Kukush, and S. Van Huffel. On errors-in-variables estimation with unknown noise variance ratio. 14th IFAC Symposium on Systen Identification, Newcastle, Australia, 2006.

[5] G. Steward and J. Sun. Matrix Perturbation Theory. Academic Press, Boston, 1990.

[6] A. Wald. The fitting of straight lines if both variables are subject to error. Annals of Mathematical Statistics, 11:284–300, 1940.

About the author A. Kukush is a full professor at the Dep. of Mechanics and Mathematics of Kiev National Taras Shevchenko University. Elected member of International Statistical Institute. In 2006 he got Shevchenko Award from Kiev National Taras Shevchenko University for a cycle of papers on errors-in-variables models. Supported by a senior postdoctoral fellowship from the Dep. of Applied Economics of the K.U.Leuven.

4.38 Rank-(R_1, R_2, R_3) reduction of tensors based on the Riemannian trust-region scheme

M. Ishteva[1,*], **L. De Lathauwer**[2,1], **P.-A. Absil**[3], **R. Sepulchre**[4], **S. Van Huffel**[1],
[1]ESAT/SCD, Katholieke Universiteit Leuven, Belgium, [2]CNRS, ETIS, Gergy-Pontoise, France,
[3]INMA, Université catholique de Louvain, Belgium, [4]Montefiore Institute, Université de Liège, Belgium,
[*]Mariya.Ishteva@esat.kuleuven.be

Introduction. Higher-order tensors are generalizations of vectors (order 1) and matrices (order 2) to three or more dimensions. They have various application areas, such as higher-order statistics, independent component analysis, biomedical engineering, and wireless communications. The mode-n ($n = 1, 2, \ldots$) vectors of a tensor are its columns, rows, etc. The dimension of the vector space spanned by the mode-n vectors is called the mode-n rank. This is a generalization of the column and row rank of a matrix. Contrary to the case of matrices, different mode-n ranks are not necessarily equal to each other.

We look for the best rank-(R_1, R_2, R_3) approximation of third-order tensors. In the matrix case, the best low-rank approximation can be obtained from the truncated Singular value decomposition (SVD). However, in the tensor case, the truncated Higher-order SVD (HOSVD) [2] gives a suboptimal low-rank approximation of the tensor, which can only be used as a starting value for iterative algorithms.

Problem formulation. For a real third-order tensor $\mathscr{A} \in \mathbb{R}^{I_1 \times I_2 \times I_3}$, find a tensor $\hat{\mathscr{A}} \in \mathbb{R}^{I_1 \times I_2 \times I_3}$ that minimizes the least-squares cost function $f(\hat{\mathscr{A}}) = \|\mathscr{A} - \hat{\mathscr{A}}\|^2$ under the constrains $\text{rank}_1(\hat{\mathscr{A}}) \leq R_1$, $\text{rank}_2(\hat{\mathscr{A}}) \leq R_2$, $\text{rank}_3(\hat{\mathscr{A}}) \leq R_3$. This minimization problem is equivalent to (see [3]) the maximization of

$$\overline{g}(\mathbf{U}, \mathbf{V}, \mathbf{W}) = \|\mathscr{A} \times_1 \mathbf{U}^T \times_2 \mathbf{V}^T \times_3 \mathbf{W}^T\|^2 \tag{4.46}$$

over the orthonormal matrices $\mathbf{U} \in \mathbb{R}^{I_1 \times R_1}$, $\mathbf{V} \in \mathbb{R}^{I_2 \times R_2}$, $\mathbf{W} \in \mathbb{R}^{I_3 \times R_3}$.

Riemannian trust-region scheme. This is an iterative algorithm for minimizing a cost function, consisting of the following steps (see [1]):

- compute an update η to be applied to the current iterate x; η solves a trust-region subproblem;
- evaluate the quality of the model;
- accept or reject the new iterate;
- update the trust-region radius.

Riemannian trust-region based rank-(R_1, R_2, R_3) approximation of a tensor. The function $\overline{g}$ has the invariance property:

$$\overline{g}(\mathbf{U}, \mathbf{V}, \mathbf{W}) = \overline{g}(\mathbf{U}\mathbf{Q}^{(1)}, \mathbf{V}\mathbf{Q}^{(2)}, \mathbf{W}\mathbf{Q}^{(3)}),$$

where $\mathbf{Q}^{(i)}$ are orthogonal matrices. Thus, $\overline{g}$ has a unique projection g:

$$g(\mathbf{U}O_{R_1}, \mathbf{V}O_{R_2}, \mathbf{W}O_{R_3}) := \overline{g}(\mathbf{U}, \mathbf{V}, \mathbf{W}).$$

In our algorithm, we express the tensor approximation problem as minimizing the cost function $-g$ on a proper manifold (a product of three Grassmann manifolds). We apply the Riemannian trust-region scheme, using the truncated conjugate-gradient method for solving the trust-region subproblem. Making use of second order information about the cost function, superlinear convergence is achieved. We provide some simulation results concerning the stability of the algorithm.

Bibliography

[1] P.-A. Absil, C.G. Baker, K.A. Gallivan. Trust-region methods on Riemannian Manifolds. *Accepted for publication in Found. Comput. Math., doi:10.1007/s10208-005-0179-9*, 2006.

[2] L. De Lathauwer, B. De Moor, J. Vandewalle. A Multilinear Singular Value Decomposition. *SIAM J. Matrix Anal. Appl.*, 21(4):1253–1278, 2000.

[3] L. De Lathauwer, B. De Moor, J. Vandewalle. On the Best Rank-1 and Rank-$(R_1, R_2, \ldots, R_N)$ Approximation of Higher-Order Tensors. *SIAM J. Matrix Anal. Appl.*, 21(4):1324–1342, 2000.

About the authors Mariya Ishteva is a Ph.D. student at the KULeuven, Belgium. Lieven De Lathauwer has a permanent research position with the French Centre National de la Recherche Scientifique, affiliated with the ETIS Lab and a honorary research position at the KULeuven, Belgium. Pierre-Antoine Absil is an associate professor at the Université catholique de Louvain, Belgium. Rodolphe Sepulchre is professor at the Université de Liège, Belgium. Sabine Van Huffel is a full professor at the KULeuven, Belgium.

4.39 On core problem formulation in linear approximation problems with multiple right-hand sides

Iveta Hnětynková, Martin Plešinger, Zdeněk Strakoš.
Institute of Computer Science, Academy of Sciences of the Czech Republic, Pod Vodárenskou věží 2, 182 07 Praha 8, Czech Republic
hnetynkova@cs.cas.cz, mata@cs.cas.cz, strakos@cs.cas.cz

Consider an orthogonally invariant linear approximation problem $Ax \approx b$. In [10] it is proved that the partial upper bidiagonalization of the extended matrix $[b,A]$ determines a *core approximation problem* $A_{11}x_1 \approx b_1$, with the necessary and sufficient information for solving the original problem given by b_1 and A_{11}. The transformed data b_1 and A_{11} can be computed either directly, using Householder orthogonal transformations, or iteratively, using the Golub-Kahan bidiagonalization [4]. The bidiagonalization is stopped at the first zero bidiagonal element. It is shown how the core problem can be used in a simple and efficient way for solving different formulations of the original approximation problems. The proof in [10] is based on the singular value decomposition of the matrix A. In [5], [6], the core problem formulation is derived from the relationship between the Golub-Kahan bidiagonalization and the Lanczos tridiagonalization [7], see [1], and from the well-known properties of Jacobi matrices. For a rewiev see the lecture [9] presented at this workshop.

In this contribution we concentrate on extension of the idea of core problem formulation to linear approximation problems with multiple right-hand sides. Here a concept of a (minimally dimensioned) approximation problem containing the necessary and sufficient information for solving the original problem seems more complicated. The analysis should start from the singular value decomposition, and computation should be based on the block (banded) Lanczos bidiagonalization, see [8, Section 2.3], [2], [3].

We will discuss several examples illustrating difficulties which have to be resolved in order to get a general multiple right-hand sides core problem formulation.

Acknowledgments. This work has been supported by the National Program of Research "Information Society" under project 1ET400300415, and by the Institutional Research Plan AV0Z10300504.

Bibliography

[1] A. Björck, A bidiagonalization algorithm for solving large and sparse ill-posed systems of linear equations, BIT **28**, pp. 659–670 (1988).

[2] A. Björck, A band-Lanczos generalization of bidiagonal decomposition, http://www.mai.liu.se/˜akbjo/kth06.pdf (2006).

[3] A. Björck, Bidiagonal decomposition and least squares, (2005). http://wwwmaths.anu.edu.au/events/sy2005/odatalks/canb05.pdf

[4] G. H. Golub, W. Kahan, Calculating the singular values and pseudo-inverse of a matrix, SIAM J. Numer. Anal. Ser. B **2**, pp. 205–224 (1965).

[5] I. Hnětynková, Z. Strakoš, Lanczos tridiagonalization and core problems, to appear in Lin. Alg. Appl.

[6] I. Hnětynková, M. Plešinger, Z. Strakoš, Lanczos tridiagonalization, Golub-Kahan bidiagonalization and core problem, to appear in PAMM.

[7] C. Lanczos, An iteration method for the solution of the eigenvalue problem of linear differential and integral operators, J. Res. Nat. Bur. Standards **45**, pp. 255–282 (1950).

[8] D. Sima, Regularization techniques in model fitting and parameter estimation, PhD thesis, Faculty of Engineering, K.U.Leuven, Belgium (2006).

[9] C. C. Paige, Z. Strakoš, Bidiagonalization as a fundamental decomposition of data in linear approximation problems, lecture at 4th Workshop on TLS and Errors-in-Variables Modeling, Leuven, Belgium (2006).

[10] C. C. Paige, Z. Strakoš, Core problems in linear algebraic systems, SIAM J. Matrix Anal. Appl. **27**, pp. 861–875 (2006).

5 List of participants

John Allan
Univ. Shetfield, Dept. Computer Science
211 Portobello Street
Sheffield South York S1 4DP, England
Tel. 44/114/222 18 34, Fax. 44/114/222 1810
Email john@dcs.shef.ac.uk

Amir Beck
Israel Inst. Technology
Fac. Industrial Eng. & Management
Technion City, Haifa 32000, Israel
Tel. 972/4/829 44 30, Fax. 972/4/829 5688
Email becka@ie.technion.ac.il

Åke Björck
Linköping Univ., Dept. Mathematics
SE-581 83 Linkping, Sweden
Tel. 46/13/28 14 00, Fax. 46/13/10 07 46
Email akbjo@mai.liu.se

Keith Burnham
Coventry University, Systems Engineering
Priory Street, CV1 5FB Coventry
United Kingdom
Email : csx167@coventry.ac.uk

Xiao-Wen Chang
McGill Univ., School Computer Science
3480 University Street
Montreal, Quebec H3A 2A7, USA
Tel. 1/514/398 82 59, Fax. 1/514/398 38 83
Email chang@cs.mcgill.ca

Chi-Lun Cheng
Inst. Statistical Science
Academia Sinica, Tapei, Taipei 115, Taiwan, R.O.C.
Tel. 886/2/2783 5611 420, Fax. 886/2/2783 1523
Email chcheng@stat.sinica.edu.tw

Lieven De Lathauwer
CNRS-Lab ETIS
6, avenue du Ponceau, BP 77
95014 Cergy, Pontoise Cedex, France
Tel. 33/1/30 73 66 27, Fax. 33/1/30 73 66 27
Email delathau@ensea.fr

Maarten De Vos
K.U.Leuven, Dept. ESAT/SCD (SISTA)
Kasteelpark Arenberg 10
3001 Leuven, Belgium
Tel. 32/16/32 18 84, Fax. 32/16/32 19 70
Email maarten.devos@esat.kuleuven

Roberto Diversi
Univ. Bologna
Dept. Electronics, Computer
Science and Systems
Viale del Risorgimento 2
40136 Bologna, Italy
Tel. +39/051/209 3770, Fax. +39/051/290 3073
Email rdiversi@deis.unibo.it

Giannone Domenico
ULB, ECARES
CP 114, Ave F.D. Roosevelt 50
1050 Brussels, Belgium
Tel. 32/2/650 4221, Fax. 32/2/650 4475
Email dgiannon@ulb.ac.be

Jurjen Duintjer Tebbens
Inst. Computer Science
Academy of Sciences
Pod Vodarenskou Vezi 2
182 07 Prague 8, Czech Republic
Tel. 420/26605 2182, Fax. 420/28658 5789
Email tebbens@cs.cas.cz

Yaron A. Felus
Ferris State Univ.
Dept. Surveying Engineering
915 Campus Dr., SWAN 314, Big Rapids
Michigan 49307, USA
Tel. 1/231/591 22 21, Fax. 1/231/591 2082
Email felusv@ferris.edu

Gaurav Garg
Indian Institute of Technology Kampur
Dept. Mathematics & Statistics
208 016 Uttar Pradesh, India
Tel. 91/9358408117
Email ggarg@iitk.ac.in

Gennaro Gelao
Politecnico di Bari, Dipt. Elettronica
Via Orabona 4, Bari 70125, Italy
Tel. 39/320/431 61 42, Fax. 39/080/596 34 10
Email g.gelao@poliba.it

Gene H. Golub
Stanford University
Computer Science, Gates 2B, 280
Stanford, CA 94305, USA
Tel. 1/650/723,3124, Fax. 1/650/618,2767
Email golub@sccm.stanford.edu

Roberto Guidorzi
Univ. Bologna, Dept. Electronics, Computer
Science and Systems
Viale des Risorgimento 2
40136 Bologna, Italy
Tel. +39/051/209 3770, Fax. +39/051/290 3073
Email rguidorzi@deis.unibo.it

Iveta Hnetynkova
Inst. Computer Science, Academy of Sciences
Pod Vodarenskou Vezi 2
182 07 Prague 8, Czech Republic
Tel. 420/26605 3787, Fax. 420/28658 5789
Email hnetynkova@cs.cas.cz

Mariya Ishteva K.U.Leuven, Dept. ESAT/SCD (SISTA)
Kasteelpark Arenberg 10
3001 Leuven, Belgium
Tel. +32/16/32 10 67, Fax. 32/16/32 19 70
Email mariya.ishteva@esat.kuleuven.be

Kenichi Kanatani
Okayama Univ., Dept. Computer Science
Okayama 700-8530, Japan
Tel. +81/86/251 8173, Fax. +81/86/251 8173
Email kanatani@suri.it.okayama-u.ac.jp

Alexander Kukush
Kyiv Nat. Taras Schevchenko Univ.
Reitarskaya st. 21 25, 01034 Kiev
Tel. 38/44/279 26 34, Fax. 38/44/266 15 58
Email alexander_kukush@univ.kiev.ua

Jens Linden
Coventry University, Systems Engineering
Priory Street, CV1 5FB Coventry, United Kingdom
Email :lindenj@coventry.ac.uk

Andrzej Mackiewicz
Technical Univ. Poznan
Inst. of Mathematics
ul. Piotrowo 3A, Poznan, Poland
Tel. 48/061/665 2320, Fax. 48/061/665 2348
Email andrzej.mackiewicz3@neostrada.pl

Ivan Markovsky
K.U.Leuven, Dept. ESAT/SCD (SISTA)
Kasteelpark Arenberg 10
3001 Leuven, Belgium
Tel. +32/16/32 11 437, Fax. 32/16/32 19 70
Email ivan.markovsky@esat.kuleuven.be

Nicola Mastronardi
National Res. Council Italy
Ist Applicazioni del Calcolo
”M. Picone” Bari, Via G. Amendola 122/D
70126 Bari, Italy
Tel. 39/080/59 29 740, Fax. 39/080/59 29 770
Email nicola.mastronardi

Mei Mei Honge
Uppsala Universityey Div. Systems & Control, Dept.
Flogstavägen 63B 233
752 72 Uppsala, Sweden
Tel. 46/18/471 33 98, Fax. 46/18/511 925
Email mei.hong@it.uu.se

Francesca Monti
ULB, ECARES
CP 114, Ave F.D. Roosevelt 50, 1050 Brussels
Tel. 32/2/650 4221, Fax. 32/2/650 4475
Email fmonti@ulb.ac.be

Daniel Pachner
Czech Technical University in Prague
Department of Control Engineering
Technicia 2, 166 27 Prague 6
Czech Republic, Fax. 420/224/913 646
Email pachner@control.felk.cvut.cz

Michal Pesta
Inst. Computer Science
Academy of Sciences
Pod Vodarenskou Vezi 2
182 07 Prague 8, Czech Republic
Tel. 421/903/043793
Email michal.pesta@matfyz.cz

Rik Pintelon
Vrije Universiteit Brussel
ELEC, Pleinlaan 2
1050 Brussel, Belgium
Tel. 32/2/629 29 44, Fax. 32/2/620 28 50
Email rik.pintelon@vub.ac.be

Martin Plesinger
Inst. Computer Science
Academy of Sciences
Pod Vodarenskou Vezi 2
182 07 Prague 8, Czech Republic
Tel. 420/26605 2182, Fax. 420/28658 5789
Email martin.plesinger@seznam.cz

José A. Ramos
Indiana Purdue Univ., Dept. Electrical & Comp. Eng.
IN 46112 Indianapolis, USA
Tel. 1/317/214 0044, Fax. 1/317/214 4493
Email jaramos@iupui.edu

Jiri Roubal
Czech Technical University in Prague
Department of Control Engineering
Technicia 2, 166 27 Prague 6, Czech Republic
Tel. +420/732/544 725, Fax. 420/224/918 646
Email roubalj@control.felk.cvut.cz

Burkhard Schaffrin
School of Earth Sciences
The Ohio State Univ., 275 Mendenhall Lab.
125 South Oval Mall, Columbus, OH 43210, USA
Tel. 1/614/292 0502, Fax. 1/614/292 7688
Email schaffrin.1@osu.edu

Hans Schneeweiss
Institute of Statistics, Univ. Munich, Akademiestrasse 1
80799 Munich, Germany
Tel. 49/89/2180 3224, Fax. 49/89/2180 5044
Email scheew@stat.uni-muenchen.de

Johan Schoukens
Vrije Universiteit Brussel
ELEC, Pleinlaan 2, 1050 Brussel, Belgium
Tel. 32/2/629 29 44, Fax. 32/2/620 28 50
Email johan.schoukens@vub.ac.be

Mieke Schuermans
K.U.Leuven, Dept. ESAT/SCD (SISTA)
Kasteelpark Arenberg 10
3001 Leuven, Belgium
Tel. +32/16/32 11 43, Fax. 32/16/32 19 70
Email mieke.schuermans@esat.kuleuven.be

Prof. Shalabh
Indian Inst. Technology, Dept. Mathematics & Statistics
Kanpur 208016, India
Tel. 91/512/259 7905, Fax. 91/512/2597500
Email shalab@iitk.ac.in

Diana Sima
K.U.Leuven, Dept. ESAT/SCD (SISTA)
Kasteelpark Arenberg 10
3001 Leuven, Belgium
Tel. +32/16/32 11 43, Fax. 32/16/32 19 70
Email diana.sima@esat.kuleuven.be

Torsten Söderström
Uppsala University, Div. Systems & Control, Dept.
PO Box 337, 75105 Uppsala, Sweden
Tel. 46/18/471 30 75
Email ts@it.uu.se

Umberto Soverini
Univ. Bologna
Dept. Electronics, Computer
Science and Systems
Viale del Risorgimento 2
40136 Bologna, Italy
Tel. +39/051/209 3026, Fax. +39/051/209 3073
Email usoverini@deis.unibo.it

Ctakar Sprdlik
Czech Techn. Univ.
Fac. Electrical Eng., Dept. Control Eng.
Technicka 2, 166 27 Praha, Czech Republic
Tel. +420/224/357 682
Email sprdlc1@control.felk.cvut.cz

Zdenek Strakos
Inst. Computer Science
Academy of Sciences, Pod Vodarenskou Vezi 2
182 07 Prague 8, Czech Republic
Tel. 420/26605 3290, Fax. 420/28658 5789
Email strakos@cs.cas.cz

Hideyuki Tanaka
Kyoto University
Dept. Applied Mathematics & Physics
Grad. School Informatics
Yoshida-honmachi, Sakyo ward, Kyoto city, Japan
Tel. +81-75-753-4754, Fax. +81-75-753-5507
Email htanaka@amp.i.kyoto-u.ac.jp

Stéphane Thil
Centre de Recherce en Automatique
de Nancy BP 239
54506 Vandoevre-les-Nancy Cedex, France
Tel. 33/383/684 461, Fax. 33/383/ 684 462
Email stephane.thil@ecran.uhp-nancy.fr

David Titley-Peloquin
McGill University
School Computer Science
3480 University Street
McConnell Eng. Bld., Room 318
Montreal QC H3A 2A7, Canada
Tel. 1/514/807 1915. Fax. 1/514/398 3883
Email dtitie@cs.mcgill.ca

Pavel Trnka
Czech Technical University in Prague
Department of Control Engineering
Technicia 2
166 27 Prague 6
Czech Republic
Tel. +420/732/544 725, Fax. 420/224/918646
Email trnkap@control.felk.cvut.cz

Richard Vaccaro
University of Rhode Island
Dept. Electrical Engineering
4, East Alumni Drive
Kingston, RI 02881, USA
Tel. 1/401/874 5816, Fax. 1/401/782 6422
Email vaccaro@ele.uri.edu

Sabine Van Huffel
K.U.Leuven, Dept. ESAT/SCD (SISTA)
Kasteelpark Arenberg 10
3001 Leuven, Belgium
Tel. 32/1632 17 03, Fax. 32/16/32 19 70
Email sabine.vanhuffel@esat.kuleuven.be

Joos Vandewalle
K.U.Leuven, Dept. ESAT/SCD (SISTA)
Kasteelpark Arenberg 10
3001 Leuven, Belgium
Tel. 32/16/32 10 52, Fax. 32/16/32 19 70
Email joos.vandewalle@esat.kuleuven.be

Benoit Vinsonneau
Coventry University, Systems Engineering
Priory Street, CV1 5FB Coventry, United Kingdom
Email : csy150@coventry.ac.uk

Alistair Watson
University of Dundee
Dept. Mathematics, Dundee DD1 4HN, Scotland
Tel. 44/1382/24 44 72, Fax. 44/1382/34 55 16
Email gawatson@maths.dundee.ee.uk

Jan C. Willems
K.U.Leuven, Dept. ESAT/SCD (SISTA)
Kasteelpark Arenberg 10
3001 Leuven
Belgium
Tel. +32/16/32 96 07, Fax. 32/16/32 19 70
Email jan.willems@esat.kuleuven.be

J. Winkler
Univ. Shetfield, Dept.Computer Science
211 Portobello Street
Sheffield South York S1 4DP, United Kingdom
Tel. 44/114/222 18 34, Fax. 44/114/222 1810
Email j.winkler@dcs.shef.ac.uk

Arie Yeredor
Tel-Aviv University
School Electrical Engineering
PO Box 39040, Tel Aviv 69978, Israel
Tel. 972/3/640 5314, Fax. 972/3/640 7095
Email arie@eng.tau.ac.il

Lihong Zhi
Academy Mathematics & Systems Scinece
55, East Road, Zhongguancun,
Beijing 100080, China
Tel. 8610/626 513 69, Fax. 8610/6263 0706
Email lzhi@mmrc.iss.ac.cn

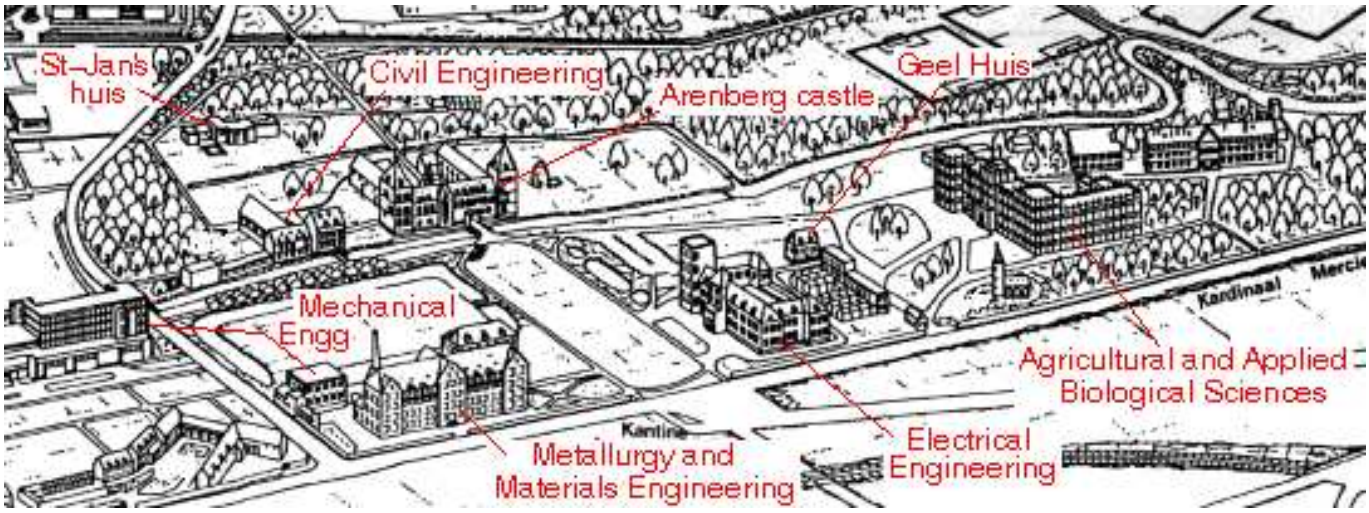
Jan Zitko
Charles Univ., Fac. Mathematics & Physics
Katedra numericke matematiky
Sokolovska 83, 186 75 Praha 8, Czech Republic
Tel. 11 191 33 62
Email zito@karlin.mff.cuni.cz

6 Location

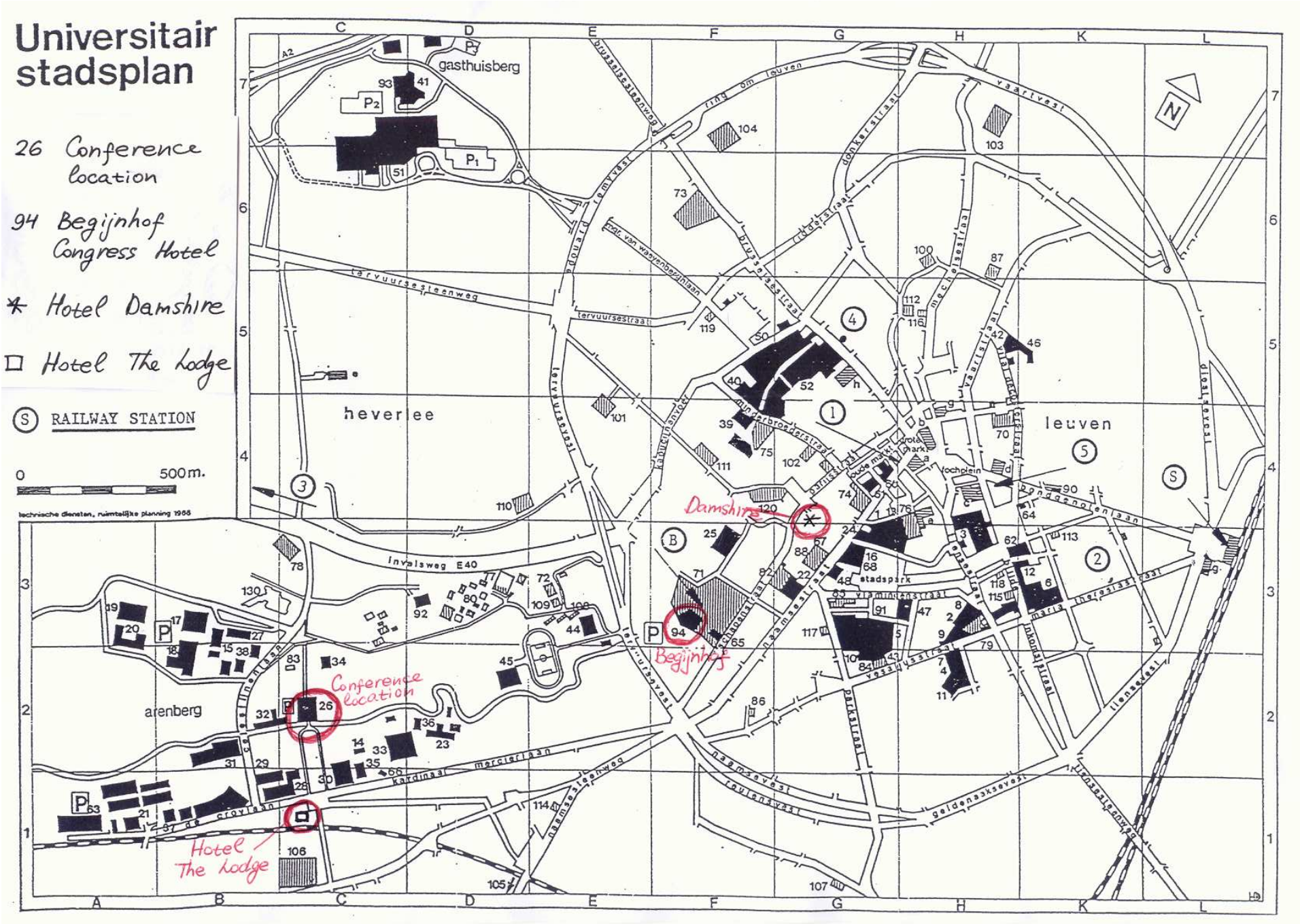
All sessions with oral presentations will take place in the **auditorium Arenberg**, Arenberg Castle, Kasteelpark Arenberg 1, 3001 Heverlee. (There are no parallel tracks for this workshop.) The poster session is scheduled for room 00.62 in the department of electrical engineering (ESAT). See the maps below and the route description on

`http://www.esat.kuleuven.ac.be/info/route.en.php`

The region of the Arenberg castle.



Universitair stadsplan



Monday, August 21

Session I: Regularized total least squares

09h00–09h45 *G. Golub*, Matrices and moments: perturbation for least
09h45–10h30 *A. Beck*, The regularized total least squares problem
11h00–11h30 *D. Sima*, Level choice in truncated total least squares
11h30–12h00 *A. Watson*, Robust counterparts of errors-in-variables problems

Session II: Nonlinear measurement error models

13h30–16h15 *C.-L. Cheng*, On the conditional score and corrected score estimation in nonlinear ME models
16h15–15h00 *H. Schneeweiss*, Comparing the efficiency of structural and functional methods in ME models
15h30–16h00 *Shalabh*, On the estimation of linear ultrastructural model when error variances are known
16h00–16h30 *G. Garg*, Consistent estimation of regression coefficients in ME model under exact linear restrictions

Tuesday, August 22

Session III: Numerical methods for total least squares

08h30–09h15 *Z. Strakoš*, Bidiagonalization as a fundamental decomposition of data in linear approximation problems
09h15–10h00 *Å. Björck*, A band Lanczos algorithm for least squares and total least squares problems
10h30–11h00 *X.-W. Chang*, Minimal backward perturbations for data least squares problems
11h00–11h30 *D. Titley-Peloquin*, Characterizing matrices consistent with given approximate solutions
11h30–12h00 *M. Schuermans*, On the equivalence between TLS and maximum likelihood principal component analysis

Session IV: Geometric fitting

13h30–14h30 *K. Kanatani*, Hyperaccuracy for geometric fitting
14h30–15h15 *I. Markovsky*, Low-rank approximation and its applications for data fitting
15h15–16h00 *A. Kukush*, Estimation in a multivariate errors-in-variables model with unknown noise variance ratio

Session V: Total least squares applications in computer algebra

16h30–17h00 *L. Zhi*, An STLS algorithm for approximate greatest common divisors of multivariate polynomials
17h00–17h30 *J. Winkler*, Structured matrix methods for the computation of rank reduced Sylvester matrix

Wednesday, August 23

Session VI: Errors-in-variables system identification

08h30–09h30 *T. Söderström*, Errors-in-variables methods in system identification
09h30–10h15 *R. Guidorzi*, Some issues on errors-in-variables identification
10h30–11h00 *J. Linden*, Model-based control in the errors-in-variables framework
11h00–11h30 *R. Pintelon*, Frequency domain maximum likelihood estimation of linear dynamic EIV models
11h30–12h00 *J. Schoukens*, Identifiability analysis for errors-in-variables problems

Session VII: Total least squares applications in signal processing

13h30–16h15 *A. Yeredor*, On the role of constraints in system identification
16h15–15h00 *R. Vaccaro*, Optimal parameter estimation from shift-invariant subspaces
15h00–15h45 *L. De Lathauwer*, Principal component, independent component and parallel factor analysis
15h45–16h30 *J. Ramos*, Applications of TLS and related methods in the environmental sciences