tions in TLS and EIV modeling has increased exponentially, because of its emergence in new fields such as computer vision, image reconstruction, speech and audio processing, and its gain in popularity in fields as signal processing, modal and spectral analysis, system identification and astronomy. In the book, the use of TLS and errors-in-variables models in the most important application fields, such as signal processing and system identification, are overviewed and new algorithms that apply the TLS concept to the model characteristics used in that field are described. For example, in system identification the well-known Kalman filtering is extended to the errors-in-variables context in which noise on the inputs as well as on the outputs is taken into account thereby improving the filtering performance.

In response to this growing interest in TLS and EIV modeling by researchers and practitioners, I invited well-known experts from several disciplines to prepare a survey paper and present it at the third international workshop on TLS and EIV modeling held in Leuven, Belgium, August 27–29, 2001. These invited papers, representing two-thirds of the book, together with a selection of other presented contributions, are included in this book and yield a complete overview of the main scientific achievements since 1996 in TLS and errors-in-variables modeling. Not only computational issues, but also statistical, numerical, algebraic properties are described, as well as many new generalizations and applications. In this way, the topic is studied from different points of view as was also the case in the first two books. This *interdisciplinary* character points out the strongness of this book and enables the reader to get a complete overview of the state of the art of the research in TLS and errors-in-variables modeling as a whole.

All these methods and applications haven't been described in the first two books on TLS (SIAM, 1991 and 1997). Therefore, the new book nicely completes the earlier ones by updating the description of the state of the art on TLS since 1996. Being aware of the growing interest in these techniques by researchers and applied engineers, I strongly believe that this book will aid and stimulate users to apply the new techniques and models correctly to their own practical problems.

SABINE VAN HUFFEL

INTRODUCTION TO TOTAL LEAST SQUARES AND ERRORS-IN-VARIABLES MODELING

Bridging the Gap between Statistics, Computational Mathematics and Engineering

Sabine Van Huffel and Philippe Lemmerling
ESAT-SISTA, K.U.Leuven
Kasteelpark Arenberg 10
B-3001 Leuven-Heverlee, Belgium
{Sabine.Vanhuffel, Philippe.Lemmerling}@esat.kuleuven.ac.be

1. Introduction, formulation and state of the art until 1996

1.1 Introduction and problem formulation

The main purpose of this book is to present an overview of the progress in the last five years of a modeling technique which is known as Total Least Squares (TLS) in computational mathematics and engineering, and as Errors-In-Variables (EIV) modeling or orthogonal regression in the statistical community.

The basic motivation for TLS is the following. Let a set of multidimensional data points (vectors) be given. How can we obtain a linear model that explains the data? The idea is to modify all data points in such a way that some norm of the modification is minimized subject to the constraint that the modified vectors satisfy a linear relation. The origin of this basic insight can be traced back to the beginning of the previous century. It was rediscovered many times, often independently, mainly in the statistical and psychometric literature (where it is known as errors-in-variables or orthogonal regression and where its relation to principal component analysis and factor analysis was thoroughly investigated in the past half century). However, it is only in the last 20 years that the technique penetrated in scientific and engineering applications. One of the main reasons for its sudden popularity is the availability of efficient and numerically robust algorithms, in which the Singular Value

S. Van Huffel and P. Lemmerling (eds.), Total Least Squares and Errors-in-Variables Modeling, 1–11.
© 2002 *Kluwer Academic Publishers. Printed in the Netherlands.*

Decomposition (SVD) plays a prominent role. Another reason is the fact that TLS is really an applications oriented procedure. It is ideally suited for situations in which all data are corrupted by noise, which is almost always the case in engineering applications. In this sense it is a powerful extension of the classical Least Squares (LS) idea, which corresponds only to a partial modification of the data.

A comprehensive description of the state of the art on TLS from its conception up to the summer of 1990 and its use in parameter estimation has been presented in the monograph "*The total least squares problem: computational aspects and analysis*" by S. Van Huffel and J. Vandewalle (SIAM, 1991). While the latter book is entirely devoted to TLS, a second book, entitled "*Recent advances in total least squares techniques and errors-in-variables modeling*" (editor: S. Van Huffel, SIAM, 1997), presents the progress in TLS and in the broader field of errors-in-variables modeling from 1990 till 1996. Hence, this new volume nicely *completes* these first two books.

The problem of *linear parameter estimation* arises in a broad class of scientific disciplines such as signal processing, automatic control, system theory and in general engineering, statistics, physics, economics, biology, medicine, etc. It starts from a model described by a linear equation:

$$\alpha_1 x_1 + \cdots + \alpha_n x_n = \beta \tag{1}$$

where $\alpha_1, \ldots, \alpha_n$ and β denote the variables and $x = [x_1, \ldots, x_n]^T \in \mathbb{R}^n$ plays the role of a parameter vector that characterizes the specific system. A basic problem of applied mathematics is to determine an estimate of the *true* but *unknown* parameters from certain measurements of the variables. This gives rise to an *overdetermined* set of m linear equations $(m > n)$:

$$Ax \approx b \tag{2}$$

where the i-th row of the data matrix $A \in \mathbb{R}^{m \times n}$ and the vector of observations $b \in \mathbb{R}^m$ contain the measurements of the variables $\alpha_1, \ldots, \alpha_n$ and β respectively.

In the classical least squares approach the measurements A of the variables α_i (the lefthand side of (2)) are assumed to be free of error and hence, all errors are confined to the observation vector b (the righthand side of (2)). However, this assumption is frequently unrealistic: sampling errors, human errors, modeling errors and instrument errors may imply inaccuracies of the data matrix A as well. One way to take errors in A into account is to introduce *perturbations also in* A and consider the following TLS problem ($R(A)$ denotes the range of A):

Definition 1 (Total Least Squares problem). *Given an overdetermined set of m linear equations $Ax \approx b$ in n unknowns x. The total least squares problem seeks to*

$$\underset{[\widehat{A}\ \widehat{b}] \in \mathbb{R}^{m \times (n+1)}}{\text{minimize}} \quad \| [A\ b] - [\widehat{A}\ \widehat{b}] \|_F \tag{3}$$

$$\text{subject to} \quad \widehat{b} \in R(\widehat{A}) \tag{4}$$

Once a minimizing $[\widehat{A}\ \widehat{b}]$ is found, then any x satisfying

$$\widehat{A}x = \widehat{b} \tag{5}$$

is called a TLS solution.

1.2 Computation

In most problems, the TLS problem (3-5) has a unique solution which can be obtained from a simple scaling of the right singular vector of $[A\ b]$ corresponding to its smallest singular value. Extensions of this basic TLS problem to multidimensional TLS problems $AX \approx B$ having more than one righthand side vector, to problems in which the TLS solution is no longer unique or fails to have a solution altogether and to mixed LS-TLS problems that assume some of the columns of A to be error-free, are also considered in the first book on TLS (SIAM, 1991). In addition, it is shown how to speed up the TLS computations directly by computing the SVD only partially or iteratively if a good starting vector is available. More recent advances, e.g. recursive TLS algorithms, rank-revealing TLS algorithms, regularized TLS algorithms, TLS algorithms for large scale problems, etc., were presented during the second workshop on TLS and Errors-in-Variables Modeling (Leuven, August 21–24, 1996) and reviewed in the book, entitled "*Recent advances in total least squares techniques and errors-in-variables modeling*" (editor: S. Van Huffel, SIAM, 1997).

1.3 Properties

The TLS principle, as introduced in numerical analysis, is strongly related to the concept of EIV or measurement modeling in the field of statistics, studied by Adcock (1878) and Pearson (1902) already more than one century ago. These models with *only measurement error*, referred to as classical *errors-in-variables* (EIV) models, are characterized by the fact that the true values of the observed variables satisfy one or more unknown but exact linear relations of the form (1). In particular, in case of one underlying linear relation, we assume that the m

measurements in A, b are related to n unknown parameters x by :

$$A_0 x = b_0 \qquad A = A_0 + \Delta A \text{ and } b = b_0 + \Delta b \tag{6}$$

where $\Delta A, \Delta b$ represent the measurement errors and all rows of $[\Delta A; \Delta b]$ are independently and identically distributed with zero mean and covariance matrix C, known up to a scalar multiple σ_ν^2. If additionally $C = \sigma_\nu^2 I$ is assumed with I the identity matrix (i.e. Δa_{ij} and Δb_i are uncorrelated random variables with equal variance) and A_0 has full rank, then it can be proved that the TLS solution x_{TLS} of $Ax \approx b$ estimates the *true* parameter values x_0, given by $A_0^\dagger b_0$, *consistently*, i.e. x_{TLS} converges to x_0 as m tends to infinity. This property of TLS estimates does not depend on any assumed distributional form of the errors. Note also that the LS estimates are inconsistent in this case. In these cases, TLS gives better estimates than does LS, as confirmed by simulations. This situation may occur far more often in practice than is recognized. It is very common in agricultural, medical and economic science, in humanities, business and many other data analysis situations. Hence TLS should be a quite useful tool to data analysts. In fact, the keyrole and importance of LS in regression analysis is the same as that of TLS in EIV regression. Nevertheless, a lot of confusion exists in the fields of numerical analysis and statistics about the principle of TLS and its relation to EIV modeling. In particular, the name "Total Least Squares" is still largely unknown in the statistical community, while inversely the concept of EIV modeling did not penetrate sufficiently well in the field of computational mathematics and engineering. Roughly speaking, we can say that TLS is a special case of EIV variables estimation and, as such, TLS is reduced to a method in statistics but, on the other hand, TLS is used in other fields like signal processing, system identification, modal analysis, etc., where mainly the data modification idea is used and explained from a geometric point of view, independently from its statistical interpretation.

Let us now discuss some of the main properties of the TLS method. The properties of TLS are best understood by comparing them with those of LS. First of all, a lot of insight can be gained by comparing their analytical expressions, given by:

$$\text{LS:} \quad x_{LS} = (A^T A)^{-1} A^T b \tag{7}$$

$$\text{TLS:} \quad x_{TLS} = (A^T A - \sigma_{n+1}^2 I)^{-1} A^T b \tag{8}$$

with A of full rank and σ_{n+1} the smallest singular value of the augmented data matrix $[A\ b]$.

From a numerical analyst's point of view, these formulas tell us that the TLS solution is more ill-conditioned than the LS solution since it has a higher condition number. This implies that errors in the data more likely affect the TLS solution than the LS solution. This is particularly true under worst case perturbations. Hence, TLS can be considered as a kind of *de*regularizing procedure. However, from a statistical point of view, these formulas tell us that TLS is doing the right thing in the presence of independently and identically distributed (i.i.d.) equally sized errors since it removes (asymptotically) the bias by subtracting the error covariance matrix (estimated by $\sigma_{n+1}^2 I$) from the data covariance matrix $A^T A$.

Secondly, while LS minimizes a sum of squared residuals, TLS minimizes a sum of squared *normalized* residuals, expressed as follows:

$$\text{LS:} \quad \min_x \|Ax - b\|^2 \tag{9}$$

$$\text{TLS:} \quad \min_{z=\left[\begin{smallmatrix} x \\ -1 \end{smallmatrix}\right]} \frac{\|[A\ b]z\|^2}{\|z\|^2} = \min_x \frac{\|Ax - b\|^2}{\|x\|^2 + 1} \tag{10}$$

From a numerical analyst's point of view, we say that TLS minimizes the Rayleigh quotient. From a statistical point of view, we say that we normalize the residuals by multiplying them with the inverse of their covariance matrix (up to a scaling factor) in order to derive consistent estimates.

Other properties of TLS, which were studied in the field of numerical analysis, are its sensitivity in the presence of errors on all data. Differences between the LS and TLS solution are shown to increase when the ratio $\sigma_n([A\ b])/\sigma_n(A)$ (with $\sigma_n(X)$ the nth largest singular value of a matrix X) is growing. This is the case when the set of equations $Ax \approx b$ becomes less compatible, when the vector b is growing in length and when A tends to be rank-deficient. Assuming i.i.d. equally sized errors, the improved accuracy of the TLS solution compared to that of TLS is maximal when the orthogonal projection of b is parallel with the nth singular vector of A, corresponding to its smallest singular value $\sigma_n(A)$. Additional algebraic connections and sensitivity properties of the TLS and LS problem, as well as many more statistical properties of the TLS estimators, based on knowledge of the distribution of the errors in the data, are analyzed in the first two books on TLS (SIAM, 1991 and 1997).

1.4 Extensions

Furthermore, several extensions of the TLS problem have been investigated. In particular, the *mixed LS-TLS* problem formulation allows to extend consistency of the TLS estimator in EIV models, where some of the variables α_i are measured without error. The *data least squares* problem refers to the special case in which all variables except β are

measured with error and was introduced in the field of signal processing by Dowling and DeGroat in the mid nineties. Whenever the errors are independent but unequally sized, *weighted TLS* problems should be considered using appropriate diagonal scaling matrices in order to maintain consistency. If additionally, the errors are also correlated then the *generalized TLS* problem formulation allows to extend consistency of the TLS estimator in EIV models, provided the corresponding error covariance matrix is known up to a factor of proportionality. More general problem formulations, such as *restricted TLS*, which also allow the incorporation of equality constraints, have been proposed, as well as equivalent problem formulations using other L_p norms and resulting in the so-called *Total L_p approximations*. The latter problems proved to be useful in the presence of outliers. Robustness of the TLS solution is also improved by adding regularization, resulting in the *regularized TLS* methods, or by exploiting the matrix structure (e.g. Toeplitz, Hankel) in the problem formulation, leading to the so-called *structured TLS* problems. The latter problems are also studied in more detail in this volume and fast algorithms are presented, by exploiting the matrix structure in the computations.

Finally, we mention the important extension to *nonlinear EIV* models, nicely studied in the book of Caroll, Ruppert and Stefanski (1995). In these models, the relationship between the variables α_i and β is assumed to be nonlinear. It is important to notice here that the close relationship between nonlinear TLS and EIV stops to exist. Indeed, consider the bilinear EIV model $AXB \approx C$, in which A, B, and C are noisy. Applying the TLS principle to this model leads to the following bilinear TLS problem:

$$\min_{\Delta A, \Delta B, \Delta C, X} \quad \|\Delta A\|_F^2 + \|\Delta B\|_F^2 + \|\Delta C\|_F^2 \tag{11}$$

$$\text{subject to} \quad (A - \Delta A)\, X\, (B - \Delta B) = C - \Delta C \tag{12}$$

However, solving this problem yields inconsistent estimates of X. As recently proved by Kukush, a consistent estimate can be obtained using the adjusted LS estimator (the full rank case is considered here for reasons of simplicity):

$$X_{ALS} = (A^T A - V_A)^{-1} (A^T C B^T)(B B^T - V_B)^{-1} \tag{13}$$

with $V_A = \mathbf{E}\tilde{A}^T\tilde{A}$ and $V_B = \mathbf{E}\tilde{B}^T\tilde{B}$ where $\mathbf{E}$ denotes the expected value and $\tilde{A}$ and $\tilde{B}$ represent the errors on A and B respectively. Corrections for small samples are derived in this volume and shown to give superior performance for small sized problems.

2. New evolutions since 1996: outline of the book

Since the appearance of the first two books on TLS (SIAM, 1991 and 1997), many new TLS algorithms have been developed, such as rank-revealing, neural based and regularized TLS algorithms, as well as new algorithms for solving TLS problems with elementwise weighting. A lot of attention was focused to structured TLS problems: fast and robust algorithms have been described. In particular for Toeplitz- or Hankel-structured problems, arising in many signal processing and system identification applications, a significant reduction in computation time is achieved by exploiting the displacement structure. New properties of EIV models in the presence of colored and nongaussian noise distributions, outliers and latency errors have been investigated, and new generalizations have been presented, in particular extensions to nonlinear and polynomial EIV models and to Cox's proportional Hazards model. The statistical properties of the estimators of the unknown model parameters are discussed and algorithms are derived for computing these estimators in a reliable way. In addition, various types of bounded uncertainties have been proposed in order to improve robustness of the estimators under various noise conditions and algorithms are outlined. Another hot topic that uses the TLS idea is orthogonal curve fitting, where one aims to minimize the orthogonal distances from the measurement points to a prescribed geometry: line, hyperplane, etc. New and faster algorithms for computing these geometries have been developed and analyzed. And last but not least, the number of applications in TLS and EIV modeling has increased exponentially, because of its emergence in new fields such as computer vision, image reconstruction, speech and audio processing, and its gain in popularity in fields as signal processing, modal and spectral analysis, system identification and astronomy.

In this volume, the use of TLS and errors-in-variables models in the most important application fields, such as signal processing and system identification, are surveyed and new algorithms that apply the TLS concept to the model characteristics used in those fields are described. For example, in system identification the well-known Kalman filtering is extended to the errors-in-variables context in which noise on the inputs as well as on the outputs is taken into account thereby improving the filtering performance.

All these methods and applications haven't been described in the first two books on TLS (SIAM, 1991 and 1997). Therefore, the new volume nicely completes the earlier ones by updating the description of the state of the art on TLS since 1996. Moreover, the new volume broadens the

view on errors-in-variables modeling by also discussing other useful types of errors-in-variables models with bounded uncertainties, together with the algorithms to compute the corresponding model parameters, and by discussing new application fields such as modal analysis, speech processing, information retrieval, computer vision and galactic kinematics.

All contributions evolve around the following nine themes, corresponding to the nine parts of this book:

Part I:	Basic concepts and analysis in errors-in-variables modeling
Part II:	Total least squares algorithms
Part III:	Structured total least squares problems
Part IV:	Nonlinear errors-in-variables models and statistical estimators
Part V:	Errors-in-variables modeling with bounded uncertainties
Part VI:	Curve fitting
Part VII:	Errors-in-variables estimation in system identification
Part VIII:	Errors-in-variables estimation in signal processing
Part IX:	Errors-in-variables applications in other fields

In **Part I**, the TLS estimator is analyzed in the framework of information geometry of semiparametric statistical models and in a scaled TLS framework, unifying LS, TLS and data LS problems. **Part II** overviews recent progress in TLS algorithms, based on the implicitly restarted Lanczos and using the rank-revealing orthogonal decomposition. In addition, new algorithms are presented for solving ill-conditioned TLS problem with Tikhonov regularization and TLS problems with independent but elementwise differently sized errors. **Part III** focuses on structured TLS problems: different problem formulations are analyzed, various computational algorithms are outlined and shown to reduce the computation time by exploiting the displacement structure. In addition, it is shown how to extend the problem and solve it, if latency or modeling errors are included. In **Part IV**, various types of nonlinear EIV models, including bilinear, polynomial, nonlinear functional, semi-linear and Cox's proportional Hazards models, are considered and consistent estimators are derived. In addition, the advantages of using Bayesian approaches are investigated. **Part V** presents new EIV models with various types of bounded uncertainties, including ellipsoidal and parametric uncertainties, based on robust regularized LS formulations. It is shown how these models improve the robustness of the TLS estimator and how estimators can be computed efficiently. **Part VI** focuses on curve fitting, including orthogonal distance regression or geometric fitting, where we aim to minimize the sum of squared orthogonal distances from the data points to a prescribed geometry. Fast algorithms for orthogonal distance regression, making use of the sparsety structure in the QR factorization of the associated Jacobian, as well as new algorithms, based on neural computing, for minor component analysis (which leads directly to the

TLS solution) and for structural line segment fitting, using the method of moments estimator, are investigated. In addition, the properties of a new approach, called least volume model fitting, are investigated.

Finally, Parts VII, VIII and IX review TLS applications in various applied fields to illustrate the improvements obtained so far with TLS based estimators. First of all in **Part VII**, applications in the field of system identification and input-output system modeling are discussed. Various methods for identifying dynamic EIV systems are reviewed and it is shown how to extend the Kalman filter to an EIV context in case the inputs as well as the outputs are contaminated by noise. **Part VIII** focuses on applications in signal processing, in particular in-vivo magnetic resonance spectroscopy and audio coding, and presents new state-space based methods for spectral estimation, using decimation and a multichannel approach. In addition, it is shown how to extend the least mean squares (LMS) algorithm to the EIV context for use in adaptive signal processing and various noise environments. Finally, **part IX** discusses applications in other fields, including information retrieval, image reconstruction, astronomy, and computer vision. It is shown how the TLS methods and its generalizations, including structured, regularized and generalized TLS, can be successfully applied.

This list of applications of TLS and EIV modeling is certainly not exhaustive and clearly illustrates the increased interest over the past five years. It is hoped that this book will aid and stimulate the reader to apply TLS and EIV modeling in his or her own applications and problems.

3. Using the Book

This book presents a state of the art of the progress in TLS and EIV modeling over the last five years and consists of invited overview papers, as well as selected contributed papers, presented at the third international workshop on TLS and EIV modeling held in Leuven, Belgium, August 27–29, 2001. All contributions are self-contained and can be read separately. They are written for an interdisciplinary audience, thereby avoiding to be too technical or too specialized.

To understand the book, the reader is assumed to have basic knowledge of the TLS problem and EIV modeling. Therefore, the reader is encouraged to read first the book on TLS by S. Van Huffel and J. Vandewalle (SIAM, 1991), although this is not a strong requirement. In addition, the reader is assumed to have basic knowledge of linear algebra, matrix computations and elementary statistics.

All used notations are clearly defined in each contribution, except for the following very common notations in linear algebra which are globally defined here for the whole book:

- A matrix is always denoted by a capital letter, e.g. A. The corresponding lowercase letter with the subscript i and ij refers to the i-th column and (i,j)-th entry respectively, e.g. a_i, a_{ij}.
- A vector is represented by a lowercase letter, e.g. b. The individual components are denoted with single subscripts, e.g. b_i.
- The superscripts ^{T}, ^{H}, and † denote respectively the transpose, conjugate transpose, and Moore-Penrose pseudo-inverse of a vector or matrix.
- The 2-norm or Euclidean norm of a vector s is denoted by $\|s\|_2$ or, simply, by $\|s\|$, while $\|S\|_2$ or $\|S\|$ denotes the 2-norm of a matrix S. The Frobenius norm of a matrix S is denoted by $\|S\|_F$, while the L_p-norm with $1 \leq p \leq \infty$ is denoted by $\|S\|_p$.
- If S is an $m \times n$ matrix and diagonal, then write

$$S = \text{diag}(\sigma_1, \ldots, \sigma_p), \text{ or } S = \text{diag}(\sigma), \quad p = \min\{m, n\},$$

  i.e. $s_{ij} = 0$ whenever $i \neq j$ and $s_{ii} = \sigma_i$ for $i = 1, \ldots, p$, where $\sigma = [\sigma_1, \ldots, \sigma_p]^T$.
- The $m \times m$ identity matrix is denoted by I_m or, simply, by I.
- $\mathbf{E}x$ denotes the expected value of x.
- The kth unit vector, with kth component equal to one and zeroes elsewhere, is denoted by e_k.
- Denote the singular value decomposition (SVD) of an $m \times n$ matrix S by

$$S = U\Sigma V^T$$

$$\begin{aligned} U &= [u_1, \ldots, u_m], \quad u_i \in \mathbb{R}^m, \; U^T U = I_m, \\ V &= [v_1, \ldots, v_n], \quad v_i \in \mathbb{R}^n, \; V^T V = I_n, \\ \Sigma &= \text{diag}(\sigma_1, \ldots, \sigma_p), \; p = \min\{m, n\} \; \sigma_1 \geq \cdots \geq \sigma_p \geq 0. \end{aligned}$$

  The σ_i are the singular values of S. The vectors u_i and v_i are respectively the i-th left and right singular vector of S. If differentiation is made between the SVD of several matrices, accents or primes are added to each symbol in the SVD.

Acknowledgments

It is our pleasure to acknowledge the assistance of many well-known contributors responsible for the success of this book and the third international workshop on TLS and EIV modeling. To all these authors and to the workshop participants we express our warmest thanks. Furthermore, we express our feelings of thanks to our research group ESAT-SISTA for the help during the workshop. In particular, we are indebted to Ida Tassens for her invaluable assistance in organizing the workshop and to Ivan Markovsky for his careful help in editing this volume. In addition, we would like to thank the sponsors of our workshop at which all contributions in this book have been presented: the Fund for Scientific Research — Flanders (FWO) and its Scientific Research Community on numerical methods for mathematical modeling, the European Association of Signal Processing (EURASIP), and the Katholieke Universiteit Leuven. We hope that this forum of interdisciplinary communication has encouraged the participants to broaden their perception on TLS and EIV modeling.

Finally, we wish to thank Kluwer Academic Publishers, in particular the editing staff, for their assistance in the publication of this book.

Sabine Van Huffel and Philippe Lemmerling