88-040

25. *Computational methods in aeronautical fluid dynamics*
Edited by P. Stow
26. *Mathematics in signal processing II*
Edited by J. G. McWhirter

# Mathematics in Signal Processing II

Based on the proceedings of a conference organized by the Institute of Mathematics and its Applications on Mathematics in Signal Processing held at the University of Warwick in December 1988

Edited by

J. G. McWHIRTER
*Royal Signals and Radar Establishment*

CLARENDON PRESS · OXFORD · 1990

# RELIABLE AND EFFICIENT TECHNIQUES BASED ON TOTAL LEAST SQUARES FOR COMPUTING CONSISTENT ESTIMATORS IN MODELS WITH ERRORS IN THE VARIABLES

S. Van Huffel* and J. Vandewalle
*(ESAT Laboratory, Department of Electrical Engineering, K.U.Leuven, Heverlee, Belgium)*

ABSTRACT

The Total Least Squares (TLS) method has been devised as a more general fitting technique than the ordinary least squares technique for solving overdetermined sets of linear equations $AX \approx B$ when errors occur in all data. If the errors in the measurements A and B are uncorrelated with zero mean and equal variance, TLS is able to compute a strongly consistent estimate of the true solution of the corresponding unperturbed set $A_oX = B_o$. These coefficients are called the parameters of a classical errors-in-variables model.

In this paper, the TLS problem, as well as the TLS computations, are generalized in order to maintain consistency of the parameter estimates in a general errors-in-variables model, i.e. some of the columns of A may be known exactly and the errors in the remaining data matrix may be correlated and not equally sized. For this problem, a computationally efficient and numerically reliable Generalized TLS algorithm GTLS, based on the Generalized SVD (GSVD), is developed. Additionally, the equivalence between the GTLS solution and alternative expressions of consistent estimators, commonly used in linear regression analysis and system identification, is pointed out. These relations allow one to deduce the main statistical properties of the GTLS solution.

---

* Senior Research Assistant of the Belgian N.F.W.O. (National Fund of Scientific Research)

## 1 INTRODUCTION

Every linear parameter estimation problem gives rise to an overdetermined set of linear equations $AX \approx B$. Whenever both A and B are subject to errors, the Total Least Squares (TLS) method can be used for solving this set. Much of the literature concerns the classical TLS problem $AX \approx B$ in which all columns of A are subject to errors, see e.g. [7], [19-21]. If the errors in the measurements A and B are uncorrelated with zero mean and equal variance, then under mild conditions this TLS solution $\hat{X}$ is a strongly consistent estimate of the true solution X of the corresponding unperturbed set $A_o X = B_o$, i.e. $\hat{X}$ converges to X with probability one as the number of equations tends to infinity. However, in many linear parameter estimation problems some columns of A may be error-free. Moreover, the errors on the remaining data may be correlated and not equally sized. In order to maintain consistency of the result when solving these problems, the classical TLS formulation can be generalized as follows ($M^{-T}$ denotes the transposed inverse of matrix M):

Generalized TLS formulation: Given a set of m linear equations in n x d unknowns X:

$$AX \approx B \qquad A \in R^{mxn},\ B \in R^{mxd} \text{ and } X \in R^{nxd} \tag{1}$$

$$\text{Partition } A = [A_1;A_2] \qquad A_1 \in R^{mxn_1},\ A_2 \in R^{mxn_2} \text{ and } n=n_1+n_2 \tag{2}$$

$$X = [X_1^T;X_2^T]^T \qquad X_1 \in R^{n_1xd} \text{ and } X_2 \in R^{n_2xd} \tag{3}$$

and assume that the columns of $A_1$ are error-free and that nonsingular error equilibration matrices $R_D \in R^{mxm}$ and $R_C \in R^{(n_2+d)x(n_2+d)}$ are given such that the errors on $R_D^{-T}[A_2;B]R_C^{-1}$ are equilibrated, i.e. uncorrelated with zero mean and same variance.

Then, a GTLS solution of eqn (1) is any solution of the set

$$\hat{A}X = A_1X_1 + \hat{A}_2X_2 = \hat{B} \tag{4}$$

where $\hat{A} = [A_1;\hat{A}_2]$ and $\hat{B}$ are determined such that

$$\text{Range } (\hat{B}) \subseteq \text{Range}(\hat{A}) \tag{5}$$



$$\|R_D^{-T}[\Delta\hat{A}_2;\Delta\hat{B}]R_C^{-1}\|_F = \|R_D^{-T}[A_2 - \hat{A}_2;B - \hat{B}]R_C^{-1}\|_F \text{ is minimal} \tag{6}$$

The problem of finding $[\Delta\hat{A}_2;\Delta\hat{B}]$ such that eqn (5-6) are satisfied, is referred to as the GTLS problem. Whenever the solution is not unique, GTLS singles out the minimum norm solution denoted by $X = [\hat{X}_1^T;\hat{X}_2^T]^T$.

By varying $n_1$ from n to zero, this formulation can handle any LS ($n_1 = n; R_C, R_D \sim I$) and generalized LS problem ($n_1 = n, R_C \sim I$), as well as every TLS ($R_C, R_D \sim I$) and GTLS problem. The error equilibration matrices $R_D$ and $R_C$ are the square root of the error covariance matrices $C = E(\Delta^T D^{-1}\Delta)$ and $D = E(\Delta C^{-1}\Delta^T)$ respectively where E denotes the expected value operator and $\Delta_{mx(n_2+d)}$ represents the errors on the noisy data $[A_2;B]$. Often, only C and D are known: in these cases the matrices $R_C$ and $R_D$ are simply obtained from their Cholesky decomposition, i.e. $C = R_C^T R_C$ and $D = R_D^T R_D$, $R_C$ and $R_D$ upper triangular.

Although the name "total least squares" appeared only recently in the literature [7], this method of fitting is certainly not new and has a long history in the statistical literature where the method is known as orthogonal or errors-in-variables regression [5-6]. More recently, the TLS approach to fitting has also attracted interest outside of statistics. In numerical analysis, Golub and Van Loan [7] first studied this problem and presented an algorithm, based on the Singular Value Decomposition (SVD). The efficiency of their computations can be improved with a factor 2 or 3 by computing the SVD only "partially" [20], or iteratively, if a priori information about the solution is available. Here, the inverse iteration method is recommended [22]. We furthermore generalized the algorithm of Golub and Van Loan [7] to all cases in which their algorithm fails to produce a solution, described the properties of these so-called nongeneric TLS problems and proved that the proposed generalization still satisfies the TLS criteria eqn (5-6) if additional constraints are imposed on the solution space [21]. Also in experimental modal analysis, the TLS technique (sometimes called the $H_v$ technique), was studied recently [14], as well as in system identification, where TLS is more commonly known as the eigenvector method [15] or Koopmans-Levin method [3]. If some columns of the data matrix A in the set $AX \approx B$ are error-free, the classical TLS algorithms can be generalized in order to compute the more general TLS estimate

$\hat{X} = [\hat{X}_1^T;\hat{X}_2^T]^T$ satisfying the TLS criteria eqn (5-6) with $R_C \sim I$ and $R_D \sim I$ [19],[23] (see section 2). In particular, this algorithm is able to compute the Compensated Least Squares (CLS) estimate as derived by Guidorzi [10] and Stoica and Söderström [18]. When the only disturbance of the input-output sequences is given by zero-mean white noise sequences of equal variance, the CLS, GTLS and eigenvector methods all give the same estimate. For a detailed appraisal of the TLS method and its generalizations, see [19],[24].

## 2 THE GENERALIZED TLS ALGORITHM GTLS

As outlined below, the GTLS algorithm is based on an implicit GSVD method [2], which computes an SVD of a triple matrix product $E^{-1}FG^{-1}$ without explicitly forming the products and without inverting E or G. This guarantees its better numerical performance. Moreover, by first performing a QR factorization [8], only the GSVD of a smaller submatrix is required which makes the GTLS algorithm computationally more efficient than methods described in [3],[18].

Given: An m x d matrix B and an m x n matrix $A = [A_1;A_2]$ whose first $n_1$ columns $A_1$ have full column rank and are error-free, $n = n_1 + n_2$ and $m \geq n + d$.

The error equilibration matrices $(R_D)_{m\times m}$ and $(R_C)_{(n_2+d)\times(n_2+d)}$, as defined in the GTLS formulation.

Step 1: QR and QL factorizations

1.a. Compute the QR factorization of $[A_1;A_2;B]$:

$$[A_1;A_2;B] = Q_{AB}\begin{bmatrix} R_{AB} \\ 0 \end{bmatrix} \text{ with } R_{AB} = \begin{array}{c} \begin{bmatrix} R_{11} & R_{12} \\ 0 & R_{22} \end{bmatrix} \begin{array}{l} n_1 \\ n_2+d \end{array} \\ \begin{array}{cc} n_1 & n_2+d \end{array} \end{array} \text{ upper triangular}$$

1.b. If $R_D \sim I$ then $E_{22} \leftarrow I$ else compute the QL factorization of $R_D Q_{AB}$:



$$R_D Q_{AB} = Q_{\hat{D}} L_{\hat{D}} \text{ with } L^T_{\hat{D}} = \begin{array}{c} \begin{bmatrix} E_{11} & E_{12} & E_{13} \\ 0 & E_{22} & E_{23} \\ 0 & 0 & E_{33} \end{bmatrix} \begin{array}{l} n_1 \\ n_2+d \\ m-n-d \end{array} \\ \begin{array}{ccc} n_1 & n_2+d & m-n-d \end{array} \end{array} \text{ upper triangular}$$

If $n_I = n$ then begin $Z_2 \leftarrow -I_d$; go to step 2.d. end

1.c. If $R_C$ upper triangular then $R_{\hat{C}} \leftarrow R_C$ else compute the QR factorization of $R_C$:

$$R_C = Q_{\hat{C}} R_{\hat{C}} \qquad (R_{\hat{C}})_{(n_2+d)\times(n_2+d)} \text{ upper triangular}$$

Step 2: GSVD

2.a. Compute the implicit GSVD

$$U^T E_{22}^{-1} R_{22} R_{\hat{C}}^{-1} V = \mathrm{diag}(\sigma_1,\ldots,\sigma_{n_2+d}) \quad \sigma_{i-1} \geq \sigma_i \quad i=2,\ldots,n_2+d \tag{7}$$

with U and $V = [v_1,\ldots,v_{n_2+d}]$ orthonormal and $\sigma_i$ the generalized singular values.

2.b. If not user determined, compute the rank $r(\leq n_2)$ by means of a user-defined rank determinator $R_O$:

$$\sigma_1 \geq \ldots \geq \sigma_r > R_O \geq \sigma_{r+1} \geq \ldots \geq \sigma_{n_2+d}$$

2.c. If $R_C \sim I$ then $Z_2 \leftarrow [v_{r+1},\ldots,v_{n_2+d}]$ else solve $R_{\hat{C}} Z_2 = [v_{r+1},\ldots,v_{n_2+d}]$ by back substitution.

2.d. If $R_D \sim I$ then $\hat{Z} = 0$ else solve $E_{22}\hat{Z} = R_{22}Z_2$ by back substitution.

2.e. Solve $R_{11}Z_1 = E_{12}\hat{Z} - R_{12}Z_2$ by back substitution.

If $n_1 = n$ then begin $\hat{X} \leftarrow Z_1$;stop end.

Step 3: GTLS solution $\hat{X} = [\hat{X}_1^T; \hat{X}_2^T]^T$

3.a. If $R_C \neq I_{n+d}$, $d > 1$ and $r < n_2$, orthonormalize $\begin{bmatrix} Z_1 \\ Z_2 \end{bmatrix}$ using a QR factorization:

$$\begin{bmatrix} Z_1 \\ Z_2 \end{bmatrix} = Q_z R_z; \quad \begin{bmatrix} Z_1 \\ Z_2 \end{bmatrix} \leftarrow Q_z$$

3.b. Perform Householder transformations $Q$ such that

$$\begin{bmatrix} Z_1 \\ Z_2 \end{bmatrix} Q = \begin{bmatrix} W & Y \\ O & \Gamma \end{bmatrix} \begin{matrix} n \\ d \end{matrix} \text{ and } \Gamma \text{ upper triangular} \quad (8)$$
$$\qquad n_2-r \quad d$$

If $\Gamma$ nonsingular then solve $\hat{X}\Gamma = -Y$

else lower the rank r with the multiplicity of $\sigma_r$

go back to step 2.c.

END

The following comments are in order:

- Step 1 of the GTLS algorithm, based on the canonical correlation computation procedure of [2], reduces all 3 matrices involved in the GSVD to upper triangular form of equal dimension. In step 2.a. the algorithm PSVD-2 of [2] can readily be applied to find the implicit GSVD. For more details and an analysis of the computational complexity, see [2]. The special case where $R_D \sim I_m$ reduces PSVD-2 to the well-known GSVD algorithms [2-17] for computing the SVD of the product $FG^{-1}$ implicitly. In this case, the GTLS algorithm reduces to the algorithm described in [24]. These algorithms are all based on an implicit Kogbetliantz approach and are suitable for parallel implementation.

- If $R_C$ or $R_D$ are singular, the GSVD algorithm PSVD-2 can be adapted by using matrix adjoints (cf. [17]). Note that in this case the GSVD is not necessarily given by the SVD of $E_{22}^{\dagger} R_{22} R_C^{\dagger}$ ($\dagger$ denotes the pseudo-inverse). Similarly to [17], the GSVD algorithm could be adapted for the case that $m < n + d$ by adding zero rows and columns, if necessary, to give square matrices of equal dimension. These extensions are not yet fully analyzed.



- If $R_C \sim I_{n+d}$ and $R_D \sim I_m$, the GSVD in step 2 is simply the ordinary SVD of $R_{22}$ so that in this case the GTLS algorithm reduces to the classical TLS algorithms [7-8], [19, sec.1.8.1], [21] for the case that $n_1 = 0$ and [23] for the case that $n_1 \neq 0$. Observe also that the GTLS algorithm solves the ordinary LS problem, using a QR factorization, if all columns of A are error-free ($n_1 = n$), as well as the generalized LS problem described in [16].

- If $\Gamma$ in eqn (8) is nonsingular (resp., singular), the GTLS solution is called generic (resp., nongeneric). For more details, see [19], [21].

## 3. PROPERTIES OF THE GENERALIZED TLS SOLUTION

Based on the following important theorem, one can derive the main statistical properties of the GTLS solution:

Theorem 1 Consider eqn (1-2). Denote by $\sigma'$ (resp., $\sigma$) the minimal generalized singular value of the matrix pair $(R_D^{-T} A, R_{Ca}^*)$ (resp., $(R_D^{-T}[A;B], R_C^*)$) where $R_C^* =$

$$\begin{bmatrix} O & O \\ O & R_C \end{bmatrix} \begin{matrix} n_1 \\ n_2+d \end{matrix} = \begin{bmatrix} R_{Ca}^* & R_{Cab}^* \\ O & R_{Cb}^* \end{bmatrix} \begin{matrix} n \\ d \end{matrix} \text{ is upper triangular.}$$
$$\quad n_1 \quad n_2+d \qquad\qquad\qquad\quad n \quad d$$

Let $\sigma$ have multiplicity d and denote $D = R_D^T R_D$, $C_a^* = R_{Ca}^{*T} R_{Ca}^*$ and $C_{ab}^* = R_{Ca}^{*T} R_{Cab}^*$. If $\sigma' > \sigma$, the GTLS solution is given by

$$\hat{X} = (A^T D^{-1} A - \sigma^2 C_a^*)^{-1} (A^T D^{-1} B - \sigma^2 C_{ab}^*) \quad (9)$$

Proof: see [24, Theorem 4].

If $D = I$, eqn (9) is a well-known expression in linear regression analysis and statistics. Its consistency and other statistical properties have been investigated by Gallo [5]. Gleser [6] studied the special case that $R_C = I$ and $n_1 = 0$, corresponding to the classical TLS problem. Using their results and assuming that $\lim_{m\to\infty} \frac{1}{m} A_O^T A_O$ (defined below) exists and is positive definite, it can be concluded that the GTLS solution is a strongly consistent estimate of the true parameters X of the general errors-in-variables model, defined as:

$$B_O=(A_O)_{mxn}X_{nxd}=A_1X_1+(A_2)_OX_2; \quad A_2=(A_2)_O+\Delta A_2 \text{ and } B=B_O+\Delta B \qquad (10)$$

$X_1$ and $X_2$ are the true but unknown parameters to be estimated, $A_1$ and $(A_2)_O$ are of full column rank. They consist of constants as well as $B_O$. $A_1$ is known but $(A_2)_O$ and $B_O$ not. The observations $A_2$ and B of the unknown values $(A_2)_O$ and $B_O$ contain measurement errors $\Delta A_2$ and $\Delta B$ such that the rows of $[\Delta A_2; \Delta B]$ are independently and identically distributed (i.i.d.) with zero mean and known positive definite covariance matrix $C = R_C^T R_C$, up to a factor of proportionality.

Statistical properties of the GTLS solution for the case that the true values $A_O$ and $B_O$ in eqn (10) are random variables, have been proven by Kelly [12].

Not only in statistics but also in system identification, eqn (9) (with $D = I$) is well-known, see e.g. [18], [11], [10]. In particular, eqn (9) arises as a consistent estimator in ARMA modelling.

These models are given by:

$$y(t)+a_1y(t-1)+\ldots+a_{n_a}y(t-n_a)=b_1u(t-1)+\ldots+b_{n_b}u(t-n_b) \qquad (11)$$

where the $\{u(t)\}$ and $\{y(t)\}$ are the input and output sequences respectively and $\{a_j\}$ and $\{b_j\}$ are the unknown constant parameters of the system. If sufficient observations are taken, eqn (11) gives rise to an overdetermined Toeplitz-like set of equations. Assume that the only disturbances in the observed outputs and inputs (if they can't be measured exactly) are given by mutually independent zero-mean white noise sequences of equal variance, i.e. $R_C = I$ and $n_1 = 0$ or $n_b$. Then, because of the correspondence between the GTLS algorithm and the eigenvector method and compensated LS method, strong consistency of the GTLS solution of this set is proven under the same conditions as described in [1] and [18]. Several authors extended the eigenvector method, described by Levin [15], to multi-input multi-output systems in which the disturbances are not necessarily white provided the covariance matrix $C = R_C^T R_C$ of the correlated noise on the input-output data is known, up to a factor of proportionality. Again, the correspondence between GTLS and these extensions proves that the GTLS solution remains consistent in these cases under the same conditions as described in [13], [9], [11], [4].



REFERENCES

[1] Aoki, M., and Yue, P.C., (1970) On a priori error estimates of some identification methods, IEEE Transactions on Automatic Control, 15, 541-548, 1970.

[2] Ewerbring, L.M., and Luk, F.T., (1989) Canonical Correlations and Generalized SVD: Applications and New Algorithms, Proceedings SPIE 977, Real Time Signal Processing XI, paper 23°, 1988.

[3] Fernando, K.V., and Nicholson, H., (1985) Identification of linear systems with input and output noise: the Koopmans-Levin method, IEE Proceedings, part D, 132, 30-36, 1985.

[4] Furuta, K., and Paquet, J.G., (1970) On the identification of time-invariant discrete processes, IEEE Transactions on Automatic Control 15, 153-155, 1970.

[5] Gallo, P.P., (1982) Consistency of regression estimates when some variables are subject to error, Communications in Statistical Theoretical Methods, 11, 973-983, 1982.

[6] Gleser, L.J., (1981) Estimation in a multivariate "errors in variables" regression model: large sample results, Annals of Statistics, 9, 24-44, 1981.

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

[8] Golub, G.H., and Van Loan, C.F., (1983) Matrix Computations, The Johns Hopkins Univ. Press, Baltimore, Maryland, 1983.

[9] Grosjean, A., and Fouland, C., (1976) Extensions of the Levin's method (or eigenvector method) for the identification of discrete, linear, multivariable, stochastic, time invariant, dynamic systems, Proceedings 4th IFAC Symposium on Identification and System Parameter Estimation, Tbilisi, USSR, 2003-2010, 1976.

[10] Guidorzi, R.P., (1975) Canonical structures in the identification of multivariable systems, Automatica, 11, 361-374, 1975.

[11] James, P.N., Souter, P., and Dixon., D.C., (1972) Suboptimal estimation of the parameters of discrete systems in the presence of correlated noise, Electronics Letters, 8, 411-412, 1972.

[12] Kelly, G., (1984) The influence function in the errors in variables problem, Annals of Statistics, 12, 87-100, 1984.

[13] Kotta, Ü., (1979) Structure and parameter estimation of multivariable systems using the eigenvector method, Proceedings 5th IFAC Symposium on Identification and System Parameter Estimation, Darmstadt, FRG, 453-458, 1979.

[14] Leuridan, J., De Vis, D., Van der Auweraer, H., and Lembregts, F., (1986) A comparison of some frequency response function measurement techniques, Proceedings 4th International Modal Analysis Conference, Los Angeles, Feb. 3-6, 908-918, 1986.

[15] Levin, M.J., (1964) Estimation of a system pulse transfer function in the presence of noise. IEEE Transactions on Automatic Control, 9, 229-235, 1964.

[16] Paige, C.C., (1985) The general linear model and the generalized singular value decomposition, Linear Algebra and its Applications, 70, 269-284, 1985.

[17] Paige, C.C., (1986) Computing the generalized singular value decomposition, SIAM Journal on Scientific and Statistical Computing, 7, 1126-1146, 1986.

[18] Stoica, P., and Söderström, T., (1982) Bias correction in least squares identification, International Journal of Control, 35, 449-457, 1982.



[19] Van Huffel, S., Analysis of the total least squares problem and its use in parameter estimation, Doct. Dissertation, Dept. Electr. Eng., K.U. Leuven, Belgium, June 1987.

[20] Van Huffel, S., and Vandewalle, J., (1988) The partial total least squares algorithm, Journal of Computational and Applied Mathematics, 21, 333-341, 1988.

[21] Van Huffel, S., and Vandewalle, J., (1988) Analysis and solution of the nongeneric total least squares problem, SIAM Journal on Matrix Analysis and Applications, 9, 360-372, 1988.

[22] Van Huffel, S., and Vandewalle, J., (1988) Iterative speed improvement for solving slowly varying total least squares problems, Mechanical Systems and Signal Processing, 2(4), 327-348, 1988.

[23] Van Huffel, S., and Vandewalle, J., (1989) Comparison of total least squares and instrumental variable methods for parameter estimation of transfer function models, International Journal of Control, 50(4), 1039-1056.

[24] Van Huffel, S., and Vandewalle, J., (1989) Analysis and properties of the generalized total least squares problem AX ≈ B when some or all columns in A are subject to error, SIAM Journal on Matrix Analysis and Applications, 10 (3), 294-315.