STRUCTURED LEAST SQUARES WITH BOUNDED DATA UNCERTAINTIES

M. Pilanci[1], *O. Arikan*[1], *B. Oguz*[2], *M.C. Pinar*[3]

[1]Department of Electrical and Electronics Engineering, Bilkent University, Ankara, Turkey
[2]Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, USA
[3]Department of Industrial Engineering, Bilkent University, Ankara, Turkey

ABSTRACT

In many signal processing applications the core problem reduces to a linear system of equations. Coefficient matrix uncertainties create a significant challenge in obtaining reliable solutions. In this paper, we present a novel formulation for solving a system of noise contaminated linear equations while preserving the structure of the coefficient matrix. The proposed method has advantages over the known Structured Total Least Squares (STLS) techniques in utilizing additional information about the uncertainties and robustness in ill-posed problems. Numerical comparisons are given to illustrate these advantages in two applications: signal restoration problem with an uncertain model and frequency estimation of multiple sinusoids embedded in white noise.

***Index Terms*—** total least squares, robust solutions, inverse problems, structured perturbations, bounded data uncertainties

1. INTRODUCTION

In various signal processing applications such as deconvolution, signal modeling, frequency estimation and system identification, it is important to produce robust estimates for an unknown vector $\hat{\mathbf{x}}$ from a set of measurements $\mathbf{y}$. Typically, a linear model is used to relate the unknowns to the available measurements: $\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{w}$, where the matrix $\mathbf{H} \in \mathbb{R}^{m \times n}$ describes the linear relationship and $\mathbf{w}$ is an additive noise vector. There are many well known approaches to provide estimates $\hat{\mathbf{x}}$. For instance, if $\mathbf{x}$ is a random vector with known first and second order statistics, the Wiener estimator, which minimizes the mean-squared error (MSE) over all linear estimators, is a reasonable choice. In the absence of such a statistical information on $\mathbf{x}$, least squares techniques are commonly used.

In many applications the elements of matrix $\mathbf{H}$ are also subject to errors since they are results of some other measurements or an imprecise model. It has been shown that if the errors in $\mathbf{H}$ and $\mathbf{w}$ are both independent identically distributed Gaussian noise, the Maximum Likelihood (ML) estimate for $\mathbf{x}$ is provided by the Total Least Squares (TLS) technique, which ”corrects” the system with minimum perturbation so that it is consistent [1]. However, in many applications $\mathbf{H}$ has a certain structure, such as Toeplitz and Structured Total Least Squares (STLS) techniques have been developed to perform structured perturbations [2].

A major drawback of both the TLS and the STLS techniques is that, in trying to reach to a consistent system, they can produce unacceptably large perturbations on $\mathbf{H}$ and $\mathbf{y}$. Another significant problem of TLS arises in nonzero residual problems in which the original system is inconsistent, may be due to lower order linear modeling or actual nonlinear relationship between the unknowns and the measurement. In these cases the TLS solution may be more sensitive than the LS solution and it is necessary to relax the consistency condition, and incorporate perturbation bounds [1]. For this purpose, two alternative formulations have been proposed. In Min-Max formulation, which is also referred to as Bounded Data Uncertainties (BDU) or Robust Least Squares (RLS) [3], $\hat{\mathbf{x}}$ is chosen as a minimizer of the maximum error over the set of allowed perturbations. In Min-Min formulation, which is referred to as Bounded Errors-in-Variables Model [4], $\hat{\mathbf{x}}$ is chosen as a minimizer of the minimum error over the set of allowed perturbations. Therefore, Min-Max approach provides more conservative estimates than the estimates obtained by the Min-Min approach.

In this paper, we formulate a new Min-Min type approach, the Structured Least Squares with Bounded Data Uncertainties (SLS-BDU), to overcome the sensitivity problems in STLS methods. In the SLS-BDU approach the residual norm $\|(\mathbf{H} + \mathbf{\Delta H})\mathbf{x} - (\mathbf{y} + \mathbf{\Delta y})\|$ subject to bounded and structured perturbations is minimized with respect to $\mathbf{x}$ as well as the perturbations $\mathbf{\Delta H}$ and $\mathbf{\Delta y}$. Hence, the consistency is not forced, and the sensitivity of the solution is kept under control with the perturbation bounds. Before proceeding with the details of the proposed approach, we first present a review on TLS, STLS, Min-Min and Min-Max approaches. Then, on two different applications, we report results of a comparison study. Finally, the drawn conclusions are presented.

2. REVIEW: TOTAL LEAST SQUARES AND THE STRUCTURED TOTAL LEAST SQUARES

Given the overdetermined linear system of equations, $\mathbf{Hx} \approx \mathbf{y}$, where both $\mathbf{H}$ and $\mathbf{y}$ may have imprecisions, TLS produces $\mathbf{x}$ as the minimum norm solution to $(\mathbf{H} + \mathbf{\Delta H})\mathbf{x} = (\mathbf{y} + \mathbf{\Delta y})$ where $[\mathbf{\Delta H}\, \mathbf{\Delta y}]$ is chosen to be minimum norm perturbation on the original system which results in a consistent system. The TLS problem can be solved using the Singular Value Decomposition (SVD) as [1]:

$$\mathbf{x}_{TLS} = (\mathbf{H}^{\mathrm{T}}\mathbf{H} - \sigma_{n+1}^2 \mathbf{I})^{-1}\mathbf{H}^{\mathrm{T}}\mathbf{y} \ , \tag{1}$$

where σ_{n+1} is the smallest singular value of $[\mathbf{H}\ \mathbf{y}]$ and subtracted to remove the bias introduced by the error in $\mathbf{H}$. However, the subtraction of $\sigma_{n+1}^2\mathbf{I}$ from the diagonal of $\mathbf{H}^{\mathrm{T}}\mathbf{H}$ deregulates the inverse operation, hence results in sensitivity issues.

In the Structured Total Least Squares (STLS) formulation the problem becomes,
$\min_{\mathbf{\Delta H}, \mathbf{\Delta y}, \mathbf{x}} \|\mathbf{\Delta H}\ \mathbf{\Delta y}\|_F$, $s.t.(\mathbf{H}+\mathbf{\Delta H})\mathbf{x} = (\mathbf{y}+\mathbf{\Delta y})$ and $[\mathbf{\Delta H}\ \mathbf{\Delta y}]$ has the same structure as $[\mathbf{A}\ \mathbf{b}]$.

This problem is non-convex and known to be NP-hard and developed solution methods are based on local optimization. When the matrices are ill conditioned the solution has a huge norm and variance since STLS introduce deregularization similar to TLS.

3. REVIEW: MIN-MAX AND MIN-MIN METHODOLOGY

3.1. Robust Least Squares

One of the Min-Max techniques is known as the Robust Least Squares (RLS) which generates estimate to $\mathbf{x}$ as a solution to the following optimization problem:

$$\min_{\mathbf{x}} \max_{\|[\mathbf{\Delta H}\ \mathbf{\Delta y}]\|_{\mathrm{F}} \leq \rho} \|(\mathbf{H} + \mathbf{\Delta H})\mathbf{x} - (\mathbf{y} + \mathbf{\Delta y})\| \ . \tag{2}$$

RLS minimizes the worst case residual over a set of perturbations with bounded Frobenius norm. As the bound ρ gets larger, the obtained solutions deviate more from the least squares solution. Hence, the RLS approach trades accuracy for robustness.

SRLS is the structured version of RLS with $\mathbf{\Delta H} = \sum_{i=1}^{p} \delta_i \mathbf{H}_i$ and solutions to both the RLS and the SRLS problems can be obtained using convex, second-order cone programming [3].

3.2. Bounded Errors-in-Variables Model

One of the Min-Min techniques is known as the Bounded Errors-in-Variables Model, where the inner maximization of the RLS cost function is replaced with a minimization over the allowed perturbations:

$$\min_{\mathbf{x}} \min_{\substack{\|[\mathbf{\Delta H}]\|_{\mathrm{F}} \leq \eta_H \\ \|[\mathbf{\Delta y}]\|_2 \leq \eta_y}} \|(\mathbf{H} + \mathbf{\Delta H})\mathbf{x} - (\mathbf{y} + \mathbf{\Delta y})\| \ .$$

As opposed to the cautious approach in the Min-Max techniques, this technique has an optimistic approach and searches for the most favorable perturbation in the allowed set of perturbations. In this sense it is closer to the TLS technique, but more robust since it does not pursue the consistency as in TLS resulting in sensitivity issues. However, unlike the Min-Max case, the Min-Min approach may be degenerate if the residual becomes zero [4]. The nondegenerate and unstructured case has the same form of the TLS solution

$$\mathbf{x}_{Min-Min} = (\mathbf{H}^{\mathrm{T}}\mathbf{H} - \gamma \mathbf{I})^{-1}\mathbf{H}^{\mathrm{T}}\mathbf{y},$$

for some positive valued γ which depends on the perturbation bounds. For small enough bounds on the perturbations, it can be shown that the value of γ is less than that of σ_{n+1}^2 in the TLS solution given in Eqn. 1. [4]. Thus, the deregularization of the Min-Min solution is less than that of the TLS, resulting in more robust solutions.

4. PROPOSED STRUCTURED LEAST SQUARES WITH BOUNDED DATA UNCERTAINTIES APPROACH

The SLS-BDU approach is a structured Min-Min approach, that is developed to provide more robust solutions than the STLS technique. Although the STLS utilizes structured perturbations, because it seeks consistency, the perturbations can be unreasonably large even if a penalty on $\|\mathbf{x}\|$ is added to the objective. In many signal processing applications perturbations beyond some bounds cannot be justified. Therefore in our proposed approach, we want to consider perturbations that are within a given tolerable bound only. The following cases illustrate the need for the bounded perturbations:

1. The given linear equations may be inadequate to represent the observed phenomenon, e.g., wrong model, nonlinear data, where seeking consistency of equations is not appropriate.

2. Some elements of the matrix may be exactly known or given with confidence intervals, e.g., econometric or mechanical models.

3. Forcing the consistency in ill-posed problems may result a very sensitive estimator and the mean-squared error is not desirable as it will be shown in numerical examples.

In SLS-BDU approach, we propose to use the following linearly structured version of the Bounded Errors-in-Variables optimization:

$$\min_{\mathbf{x}} \min_{\|\mathbf{W}\alpha\| \leq \rho} \left\| (\mathbf{H} + \sum_{i=1}^{p} \alpha_i \mathbf{H}_i)\mathbf{x} - (\mathbf{y} + \sum_{i=1}^{p} \alpha_i \mathbf{y}_i) \right\| \quad . \quad (3)$$

Similar to the SRLS formulation, the structure is encoded to $\mathbf{H}_i$ and $\mathbf{y}_i$ with α_i's determining the amount of perturbation. The SLS-BDU formulation allows bounds defined over any convex set. Here, for the sake of simplicity in the presentation, we only consider a weighted norm bound on the α with a positive definite weighting matrix $\mathbf{W}$.

The SLS-BDU optimization given in Eqn. 3. is nonconvex. However, as we will show next, an iterative algorithm can be used to a find a local minimum of it. For this purpose, first define:

$$\mathbf{H}(\alpha) = \mathbf{H} + \sum_{i=1}^{p} \alpha_i \mathbf{H}_i, \quad \mathbf{y}(\alpha) = \mathbf{y} + \sum_{i=1}^{p} \alpha_i \mathbf{y}_i \ , \alpha = [\alpha_1 ... \alpha_p]^{\mathrm{T}} \ . \quad (4)$$

Then, simplify the SLS-BDU optimization given in Eqn. 3. as:

$$\min_{\mathbf{x}} \min_{\|\mathbf{W}\alpha\| \leq \rho} J(\mathbf{x}, \alpha) \ , \quad (5)$$

where $J(\mathbf{x}, \alpha)$ is defined as $\|\mathbf{H}(\alpha)\mathbf{x} - \mathbf{y}(\alpha)\|$. For a fixed α, minimization of $J(\mathbf{x}, \alpha)$ with respect to $\mathbf{x}$ becomes a convex ordinary least squares problem which can be solved easily. Now we will show that for a fixed $\mathbf{x}$ minimization of $J(\mathbf{x}, \alpha)$ with respect to α is also a convex optimization problem.

$$\min_{\|\mathbf{W}\alpha\| \leq \rho} J(\mathbf{x}, \alpha) = \min_{\|\mathbf{W}\alpha\| \leq \rho} \|\epsilon(\mathbf{x}) + [(\mathbf{h}_1 - \mathbf{y}_1) \dots (\mathbf{h}_p - \mathbf{y}_p)]\alpha\|$$

where $\epsilon(\mathbf{x}) = \mathbf{H}\mathbf{x} - \mathbf{y}$, $\mathbf{h}_i = \mathbf{H}_i\mathbf{x}$. Hence, for a fixed $\mathbf{x}$ minimization of $J(\mathbf{x}, \alpha)$ with respect to α becomes:

$$\min_{\|\mathbf{W}\alpha\| \leq \rho} \|\epsilon(\mathbf{x}) + \mathbf{U}\alpha\| \ , \quad (6)$$

where $\mathbf{U} = [(\mathbf{h}_1 - \mathbf{y}_1) \dots (\mathbf{h}_p - \mathbf{y}_p)]$. This final form is a Constrained Least Squares problem which can be solved by using the method of Lagrange multipliers [5].

The above derived convexity results enables us to use the following iterative optimization algorithm to converge to a local minimum of the SLS-BDU optimization given in Eqn. 3.:

Step 1 **Set** $\hat{\alpha}_0 = \mathbf{0}$, and $\hat{\mathbf{x}}_0 = (\mathbf{H}^{\mathrm{T}}\mathbf{H})^{-1}\mathbf{H}^{\mathrm{T}}\mathbf{y}$, $\hat{\alpha}_0 = 0$.

Step 2 **For** $k \geq 1$, by using the method of Lagrange multipliers update $\hat{\alpha}_{k+1}$ as the solution to (6).

Step 3 **Set** $\hat{\mathbf{x}}_{k+1} = (\mathbf{H}(\hat{\alpha}_k)^{\mathrm{T}}\mathbf{H}(\hat{\alpha}_k))^{-1}\mathbf{H}^{\mathrm{T}}\mathbf{y}(\hat{\alpha}_k)$ where $\mathbf{H}(\alpha)$ and $\mathbf{y}(\alpha)$ are defined in Eqn.4.

Step 4 **Repeat** steps 2 and 3, until $\|\hat{\mathbf{x}}_k - \hat{\mathbf{x}}_{k-1}\| \leq \varepsilon$, where ε is a user defined threshold of convergence. If problems are encountered in evaluating $\hat{\mathbf{x}}_k$, one can use QR decomposition or Tikhonov regularization.

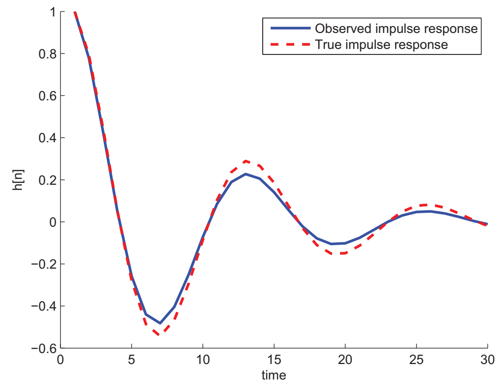

Fig. 1. Nominal and actual impulse responses are shown in solid and dashed lines respectively.

ϵ_b / b_{true}	0.2	0.6
$\|\mathbf{x}_{true} - \mathbf{x}_{LS}\| \,/\, \|\mathbf{x}_{true}\|$	0.0820	0.2123
$\|\mathbf{x}_{true} - \mathbf{x}_{SLS-BDU}\| \,/\, \|\mathbf{x}_{true}\|$	0.0274	0.1279
$\|\mathbf{H}_{true} - \mathbf{H}\|_{\mathrm{F}} \,/\, \|\mathbf{H}_{true}\|_{\mathrm{F}}$	0.1072	0.2589
$\|\mathbf{H}_{true} - \mathbf{H}_{SLS-BDU}\|_{\mathrm{F}} \,/\, \|\mathbf{H}_{true}\|_{\mathrm{F}}$	0.0655	0.1284

Table 1. $\mathbf{x}_{true}$, $\mathbf{x}_{LS}$ and $\mathbf{x}_{SLS-BDU}$ correspond to actual signal and estimates, $\mathbf{H}_{true}$, $\mathbf{H}$, $\mathbf{H}_{SLS-BDU}$ correspond to actual, nominal and corrected matrices respectively.

5. NUMERICAL EXAMPLES

5.1. Signal Restoration with an Uncertain Kernel

Suppose that the observed signal is $y[n] = \sum_{k=0}^{L-1} x[n-k]h[k] + w[n]$, $n = 0, ..N - 1$ where

$$h[n] = \sum_{i=1}^{Np} (a_i + \delta a_i) e^{-(b_i + \delta b_i)n} \cos(w_i n + \phi_i)$$

is the kernel of convolution with bounded data uncertainties on amplitudes $| \ \delta a_i \ | \leq \epsilon_{a_i}$ and dampings $| \ \delta b_i \ | \leq \epsilon_{b_i}$, $i = 1, ..., N_p$. $x[n]$ is the signal to be estimated and $w[n]$ is white Gaussian noise. The uncertainties in b_i's can be linearized by a first order approximation, $e^{-(b_i+\delta b_i)n} \approx e^{-b_i n}(1 - \delta b_i n)$, and the uncertain matrix representation becomes,

$$\mathbf{y} = (\mathbf{H} + \sum_{i=1}^{Np} \alpha_i \mathbf{H}_i)\mathbf{x} + \mathbf{w} \ ,$$

with the constraint $\|\mathbf{W}\alpha\|_{\infty} \leq \epsilon$, where $\mathbf{H}_i$ are fixed Toeplitz matrices.

Suppose that we observe the nominal impulse response shown in Fig. 1. and have a priori bounds on the uncertainty. Structured Least Squares with Bounded Data Uncertainties

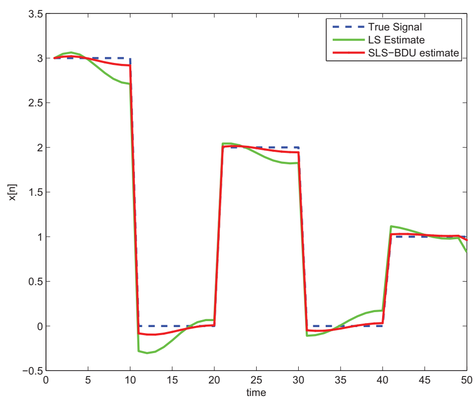

Fig. 2. Actual and restored signals are shown in dashed and solid lines respectively.

	$\min_i E_i$	$\max_i E_i$	mean(E_i)
LS	0.8653	1.0044	0.9345
STLS	8.6991e-7	7.6497e+9	9.1162e+7
SLS-BDU	0.0187	1.0889	0.6396

Table 2. Minimum, Maximum and Mean Relative Errors for LS, STLS and SLS-BDU

corrects the system in given perturbation bounds and restores the original signal with better accuracy as shown in Fig. 2. and Table 2. Note that if the uncertainty is not bounded as in STLS, the approximation may not be valid and the corresponding estimator is not desirable.

5.2. Frequency Estimation of Multiple Sinusoids

Linear prediction equations can be solved to estimate the parameters of multiple sinusoids and it is shown that STLS estimator corresponds to the ML estimator when noise is normally distributed [6]. Consider the case where parameters of two sinusoids which are close in frequency need to be estimated with frequencies $f_1 = 0.32$ Hz and $f_2 = 0.30$ Hz in white noise w_n:

$$x(n) = cos(2\pi f_1 n)+cos(2\pi f_2 n)+w_n, n = 0, 1, \ldots, 99.$$

We set the constraint on the perturbations as $\|\alpha\| \leq \delta$ such that there exists an energy bound on the observed signal. The relative estimation error $E_i \triangleq \frac{\|\mathbf{x}_{true}-\mathbf{x}^{[i]}\|}{\|\mathbf{x}_{true}\|}$ of LS, STLS [2] and the proposed SLS-BDU estimators are evaluated in independent trials at 23 dB SNR and plotted in Fig. 3. As it can be seen in Table 2 when the consistency condition is relaxed as in SLS-BDU, the sensitivity problem of STLS is avoided significantly without adding a regularization term and therefore preserving details in the signals which can be resolved.

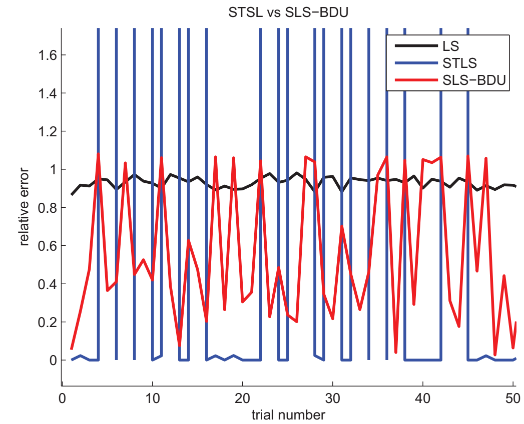

Fig. 3. Relative Estimation Error of LS, STLS and SLS-BDU in 50 independent trials. Frequently $\|\mathbf{x}_{STLS}\|$ attains huge values because of ill conditioning.

6. CONCLUSIONS

A new robust estimation technique is proposed for the solution of structured linear system of equations with bounded data uncertainties. Numerical examples showed that the proposed SLS-BDU technique achieves better mean-squared error and utilizes additional information about the uncertainties. An iterative algorithm to compute the proposed estimator is shown to be accurate and efficient. Our formulation can be used to obtain robust and accurate results in many other signal processing applications, especially in commonly occurring ill-posed problems with significant sensitivity issues.

7. REFERENCES

[1] S. Van Huffel and P. Lemmerling, "Total Least Squares and Errors-in-Variables Modeling Analysis, Algorithms and Applications", Kluwer Academic, 2002.

[2] I. Markovsky, S. Van Huffel, and R. Pintelon, "Block-Toeplitz/Hankel Structured Total Least Squares", Tech. Rep. 03–135,2003.

[3] L. El-Ghaoui, H. Lebret. "Robust solutions to least-square problems to uncertain data matrices", SIAM J. Matrix Anal. Appl. 18, 1035–1064 (1997).

[4] S. Chandrasekaran, M. Gu, A. H. Sayed, and K. E. Schubert, "The degenerate bounded error-in-variables model", SIAM J. Matrix Anal. Appl., vol. 23, pp. 138–166, 2001.

[5] G. H. Golub and U. von Matt, "Quadratically constrained least squares and quadratic problems", Numer. Math., 59 (1991), pp. 561–580.

[6] T. Abatzoglou, J. Mendel, and G. Harada, "The constrained total least squares technique and its application to harmonic superresolution", IEEE Trans. Signal Process., 39 (1991), pp. 1070–1087.