

A Structured Data Least Squares Algorithm and its Application in Digital Filtering*

Hongbin Guo and Rosemary Renaut

*Arizona State University, Department of Mathematics and Statistics,
Tempe, AZ 85287-1804*

E-mail: hb_guo@asu.edu; renaut@asu.edu

Numerical methods for the solution of the structured data least squares problem with special application in digital filtering are investigated. While the minimum mean-square error, i.e. ordinary least squares formulation, solves the linear system of equations for the case of noise in the right hand side, data least squares is formulated for the problem with noise in the coefficient matrix. For the solution of the channel equalization problem of a linear time invariant channel, the coefficient matrix, and hence the error in the coefficient matrix, possesses Hankel structure. Experimental verification demonstrates that the imposition of the structure of the error in the formulation, structured data least squares, generates more accurate solutions than are achieved by either standard ordinary least squares or data least squares, for signals with high signal to noise ratios.

Keywords: zero-forcing channel equalization; data least squares; structured data least squares.

1. Introduction

In this paper we investigate numerical methods for the solution of the data least squares (DLS) problem,¹ in which the underlying coefficient matrix possesses Hankel structure. The formulation is motivated by the application of DLS for the solution of the channel equalization problem.² Digital data transmission through a linear communication channel is basically limited by *Intersymbol interference (ISI)*, i.e. overlap of successive pulses, and *additive thermal noise*.² Thus, the channel is modeled as a linear time-invariant filter

*Supported by NSF grants DMS 0513214 and DMS 0421846.

2

with impulse response $\{h_l\}$

$$y_k = \sum_{l=-\infty}^{\infty} h_l s_{k-l} + \alpha_k, \quad (1)$$

where s_k and y_k are the transmitted and received signals, respectively, and α_k represents the additive white Gaussian noise contamination of the signal. Generally, the channel $\{h_l\}$ is unknown.

The purpose of *channel equalization* is the recovery of the original signal s_k from measurements y_k by use of an appropriate equalizer $\{x_l\}$, at the front-end of the receiver. Here as usual we assume that the equalizer is a linear finite-impulse-response (FIR) filter.

$$s_k \approx \sum_{l=0}^{n-1} y_{k-l} x_l.$$

The common strategy is to obtain an equalizer from training pairs $\{s_k\}, k = n, \dots, n+m-1$, and $\{y_k\}, k = 1, \dots, n+m-1$. One of two classical linear equalization methods is *zero-forcing*,^{3,4} in which the equalizer $\{x_l\}_{l=0}^{n-1}$ is constructed so as to completely eliminate the ISI caused by $\{h_l\}$. The second method is designed to obtain the *minimum mean-square error* (MMSE),³ which removes partial ISI effects and partial noise $\{\alpha_k\}$. This leads to the ordinary least squares problem (OLS)

$$\min_{\mathbf{x}} \sum_{k=n}^{m+n-1} \left[\sum_{l=0}^{n-1} y_{k-l} x_l - s_k \right]^2 = \min_{\mathbf{x}} \|\mathbf{A}\mathbf{x} - \mathbf{s}\|_2^2.$$

Here $\mathbf{x}^T = [x_{n-1}, \dots, x_0]$ is the *equalizer* of order n , $\mathbf{s}^T = [s_n, \dots, s_{n+m-1}]$ is the *probing* or *training* signal and A is a rectangular Hankel matrix, $m > n$,

$$A = \begin{bmatrix} y_1 & y_2 & \cdots & y_n \\ y_2 & y_3 & \cdots & y_{n+1} \\ \vdots & \vdots & \ddots & \vdots \\ y_m & y_{m+1} & \cdots & y_{m+n-1} \end{bmatrix}. \quad (2)$$

In the rest of this paper we use the name OLS instead of MMSE, which is mostly used in communication. Our focus is the development of a new zero-forcing method for directly calculating the equalizer without estimating the channel $\{h_l\}$. Suppose that $\{x_l\}_{l=0}^{n-1}$ completely removes the ISI, then

$$\sum_{l=0}^{n-1} [y_{k-l} - \alpha_{k-l}] x_l = s_k, \quad k = n, n+1, \dots, m+n-1.$$

Let $-E$ be the noise portion of the matrix A , then the equalizer may be found by solving the DLS problem:

$$\min_{\mathbf{x}, E} \|E\|_F^2 \quad \text{subject to} \quad (A + E)\mathbf{x} = \mathbf{s}, \quad (3)$$

where $\|E\|_F = \sqrt{\text{trace}(E^H E)}$ is the Frobenius norm of matrix E . The solution of this standard DLS problem is given in Appendix A. Notice, however, that the data error matrix E has the same structure as matrix A . Thus it is more appropriate to modify the DLS formulation to explicitly take account of the Hankel structure in the matrix E . Specifically, the structured DLS (SDLS) algorithm should seek to solve (3) with the Hankel structure of A imposed on E , see (4), hence reducing the number of independent variables in E to $m + n - 1$ from mn .

The idea of imposing structure on the underlying model is not novel. The OLS and DLS problems are both special cases of the total least squares (TLS) problem, in which noise is assumed in both A and \mathbf{s} .⁵⁻⁸ The initial studies,⁹⁻¹² which include a structure constraint in the TLS model, namely structured TLS (STLS), have been shown to outperform algorithms which ignore the structure. Recently, numerical stability and efficiency have been taken into account for STLS algorithm development.^{13,14} Rosen *et al* extended the STLS formulation for solution in a general p -norm, including developing a structured total least norm (STLN) algorithm in which the p -norm of the linearized objective is iteratively minimized using a Gauss-Newton method.¹⁵ An elegant fast algorithm for STLS was developed by Mastronardi *et al*¹³ to solve the subproblem in the iteration steps of the STLN for the $p = 2$ case, i.e. STLS, by exploiting the Toeplitz structure. Another approach uses a BFGS quasi-Newton method for more general data block-Toeplitz and block-Hankel structures.¹⁶ Although, SDLS may be considered as a special case of the STLN problem, the STLN algorithm can not be used directly for SDLS. By careful consideration of the properties of the SDLS, a more efficient algorithm is presented. This algorithm is based on sequential quadratical programming (SQP),¹⁷ and uses the data structure to accelerate the speed of the linear system solutions at each iteration. The presented numerical experiments contrast the performance of OLS, DLS and SDLS for generating zero-forcing equalizers.

The rest of this paper is organized as follows. The solution of SDLS using SQP is introduced in Section 2. Numerical experiments are reported in Section 3, and conclusions of the analysis in Section 4. We conclude that the best method, of those methods considered here, for zero-forcing equalization of signals with high signal to noise ratio is indeed the SDLS.

2. An Algorithm for Structured Data Least Squares

We assume $A \in \mathcal{R}^{m \times n}$ and $\mathbf{s} \in \mathcal{R}^m$, the analysis and development for the complex case follows similarly.

2.1. Properties of the SDLS problem

As shown in Section 1 matrix E is used to account for the errors in coefficient matrix A due to the additive noise α_k in y_k and has Hankel structure, consistent with that of A ,

$$E = \begin{bmatrix} \alpha_1 & \alpha_2 & \cdots & \alpha_n \\ \alpha_2 & \alpha_3 & \cdots & \alpha_{n+1} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_m & \alpha_{m+1} & \cdots & \alpha_{m+n-1} \end{bmatrix}. \quad (4)$$

Imposition of this structure on DLS problem (3), yields the SDLS model:

$$\min_{\mathbf{x}, \alpha} \|\alpha\|_2^2 \quad \text{subject to} \quad (A + E)\mathbf{x} = \mathbf{s}, \quad (5)$$

where $\alpha^T = [\alpha_1, \dots, \alpha_{m+n-1}] \in \mathcal{R}^{m+n-1}$. Defining matrix $X \in \mathcal{R}^{m \times (m+n-1)}$ through

$$X = \begin{bmatrix} x_{n-1} & \cdots & x_0 & & 0 \\ & x_{n-1} & \cdots & x_0 & \\ & & \ddots & \ddots & \ddots \\ 0 & & & x_{n-1} & \cdots & x_0 \end{bmatrix}, \quad (6)$$

it is easy to see that $X\alpha = E\mathbf{x}$. Moreover, without loss of generality, assuming that \mathbf{x} is a non-zero vector, the solution of the SDLS problem can be characterized.

Theorem 2.1. *The SDLS problem (5) is equivalent to*

$$\min_{\mathbf{x}} \|X^\dagger(\mathbf{s} - A\mathbf{x})\|, \quad (7)$$

where $X^\dagger = X^T(XX^T)^{-1}$ is the pseudoinverse of X .

Proof. Using $X\alpha = E\mathbf{x}$, it is immediate that $(A + E)\mathbf{x} = \mathbf{s}$ is expressible as $X\alpha = \mathbf{s} - A\mathbf{x}$. Then SLDS problem (5) is replaced by

$$\min_{\mathbf{x}, \alpha} \|\alpha\|_2 \quad \text{subject to} \quad X\alpha = \mathbf{s} - A\mathbf{x}. \quad (8)$$

For any given \mathbf{x} the minimal norm solution of the constraint equation is $\alpha = X^\dagger(\mathbf{s} - A\mathbf{x})$, and formulation (7) follows immediately. \square

Theorem 2.2. *If $\{\alpha_k\}$ are samples from a Gaussian random variable g with zero mean, the solution of the SDLS problem (5) is equivalent to the maximum likelihood solution.*

Proof. Let $Z(\mathbf{x}) \in R^{(m+n-1) \times (n-1)}$ denote the complement of the row space of X . Thus $XZ(\mathbf{x}) = 0$. If \mathbf{x} is a zero-forcing equalizer, i.e. $(A+E)\mathbf{x} = \mathbf{s}$, the samples $\{\alpha_k\}$ satisfy $X\alpha = \mathbf{s} - A\mathbf{x}$. Thus α can be expressed as

$$\alpha = X^\dagger(\mathbf{s} - A\mathbf{x}) + Z(\mathbf{x})\mathbf{w},$$

for some \mathbf{w} . Let σ be the standard deviation of g , then the log likelihood function for the Gaussian samples α is

$$\begin{aligned} F(\mathbf{x}, \mathbf{w}) &= -\frac{1}{\sigma^2} \|X^\dagger(\mathbf{s} - A\mathbf{x}) + Z(\mathbf{x})\mathbf{w}\|^2 \\ &= -\frac{1}{\sigma^2} (\|X^\dagger(\mathbf{s} - A\mathbf{x})\|^2 + \|Z(\mathbf{x})\mathbf{w}\|^2), \end{aligned}$$

where the orthogonality between X and $Z(\mathbf{x})$ is used.

Maximizing $F(\mathbf{x}, \mathbf{w})$ is equivalent to minimizing

$$-\sigma^2 F(\mathbf{x}, \mathbf{w}) = \|X^\dagger(\mathbf{s} - A\mathbf{x})\|^2 + \|Z(\mathbf{x})\mathbf{w}\|^2.$$

The minimum is reached at $\mathbf{w} = 0$, i.e.

$$\min_{\mathbf{x}, \mathbf{w}} (-\sigma^2 F(\mathbf{x}, \mathbf{w})) = \min_{\mathbf{x}} \|X^\dagger(\mathbf{s} - A\mathbf{x})\|^2,$$

which is equivalent to (7) in the SDLS Theorem 2.1. \square

Rather than using (7) for the characterization of the SDLS solution, an efficient algorithm which takes advantage of the underlying structure in A is presented.

2.2. Algorithm Description

The Lagrangian of (5) is given by

$$\mathcal{L}(\alpha, \mathbf{x}, \lambda) = \frac{1}{2} \|\alpha\|_2^2 - \lambda^T (\mathbf{s} - (A+E)\mathbf{x}), \quad (9)$$

where $\lambda \in R^m$ is the vector of Lagrange multipliers. The first order Karush-Kuhn-Tucker (KKT) conditions for (9) are

$$\begin{aligned} X^T \lambda + \alpha &= 0, \\ (A+E)^T \lambda &= 0, \\ (A+E)\mathbf{x} - \mathbf{s} &= 0. \end{aligned} \quad (10)$$

The SQP solves the system of equations, $F(\alpha^T, \mathbf{x}^T, \lambda^T) = 0$, using Newton's method. In forming the Jacobian $K = \nabla F$ for (10) it is convenient to introduce the matrix $\Lambda \in R^{n \times (m+n-1)}$, defined as follows

$$\Lambda = \begin{bmatrix} \lambda_1 & \cdots & \lambda_m & & 0 \\ & \lambda_1 & \cdots & \lambda_m & \\ & & \ddots & \ddots & \ddots \\ 0 & & & \lambda_1 & \cdots & \lambda_m \end{bmatrix}. \quad (11)$$

Then using $X^T \lambda = \Lambda^T \mathbf{x}$ and $E^T \lambda = \Lambda \alpha$, the variable update in the Newton step is given by solving

$$\begin{bmatrix} I & \Lambda^T & X^T \\ \Lambda & 0 & (A+E)^T \\ X & A+E & 0 \end{bmatrix} \begin{bmatrix} \delta \alpha \\ \delta \mathbf{x} \\ \delta \lambda \end{bmatrix} = - \begin{bmatrix} \alpha + X^T \lambda \\ (A+E)^T \lambda \\ (A+E)\mathbf{x} - \mathbf{s} \end{bmatrix}. \quad (12)$$

This yields the following algorithm for solving the SDLS problem by SQP, where 0_n is the zero vector of length n .

Algorithm 2.1. SQP for solving SDLS.

Initialize the variables :

$$\begin{aligned} \mathbf{x} &= \operatorname{argmin}_{\mathbf{x}} \|\mathbf{s} - A\mathbf{x}\|, \\ \alpha &= \operatorname{argmin}_{\alpha} \|\mathbf{s} - A\mathbf{x} - X\alpha\| \\ \lambda &= \operatorname{argmin}_{\lambda} \left\| \begin{bmatrix} \alpha \\ 0_n \end{bmatrix} + [X, A+E]^T \lambda \right\| \end{aligned}$$

Do Until Convergence

- (1) Solve (12)
- (2) Form updates $\alpha^{\text{new}} = \alpha + \delta \alpha$, $\mathbf{x}^{\text{new}} = \mathbf{x} + \delta \mathbf{x}$ and $\lambda^{\text{new}} = \lambda + \delta \lambda$.
- (3) Test for convergence.

End (Do)

2.3. Numerical approaches for solving the iteration equation

The analysis and practical implementation of (12) is facilitated by using the block notations

$$K = \begin{bmatrix} B_1 & B_2^T \\ B_2 & 0 \end{bmatrix}, \quad B_1 = \begin{bmatrix} I_{m+n-1} & \Lambda^T \\ & \Lambda & 0 \end{bmatrix}, \quad B_2 = [X, A+E],$$

and introducing

$$\delta \mathbf{z} = \begin{bmatrix} \delta \alpha \\ \delta \mathbf{x} \end{bmatrix}, \quad \mathbf{f} = - \begin{bmatrix} \alpha + X^T \lambda \\ (A + E)^T \lambda \end{bmatrix}, \quad \mathbf{r} = \mathbf{s} - (A + E)\mathbf{x}.$$

Then (12) can be expressed in the block form

$$\begin{bmatrix} B_1 & B_2^T \\ B_2 & 0 \end{bmatrix} \begin{bmatrix} \delta \mathbf{z} \\ \delta \lambda \end{bmatrix} = \begin{bmatrix} \mathbf{f} \\ \mathbf{r} \end{bmatrix},$$

which implies

$$\begin{bmatrix} B_1 & B_2^T \\ 0 & B_2 B_1^{-1} B_2^T \end{bmatrix} \begin{bmatrix} \delta \mathbf{z} \\ \delta \lambda \end{bmatrix} = \begin{bmatrix} \mathbf{f} \\ B_2 B_1^{-1} \mathbf{f} - \mathbf{r} \end{bmatrix}. \quad (13)$$

Here the block Gaussian elimination requires that B_1 is nonsingular, which follows from the nonsingularity of Λ . Without loss of generality, we can assume that λ is a nonzero vector, so that immediately by its structure it can be seen that Λ is of full row rank. Under this assumption, the following theorem demonstrates the non-singularity of B_1 , and presents details about the spectrum of B_1 .

Theorem 2.3. *For λ a non-zero vector, B_1 is non-singular. Moreover, the spectrum of B_1 contains 1 with multiplicity $m - 1$, and $1/2(1 \pm \sqrt{1 + 4\eta_k^2})$, $k = 1, \dots, n$, where η_k are the singular values of Λ .*

Proof.

The proof of the eigenvalue distribution can be found in the paper of Fischer *et al.*¹⁸ We thus conclude that B_1 is nonsingular because $\eta_i \neq 0$ when Λ is of full row rank. \square

Under the additional assumption that \mathbf{x} is nonzero, the following theorem shows that the coefficient matrix K in (12) is nonsingular.

Theorem 2.4. *For λ and \mathbf{x} nonzero vectors, matrix K is nonsingular.*

Proof. By the assumption that λ and \mathbf{x} are nonzero vectors, both Λ and X are full row rank. Thus B_2 has full row rank and B_1 is nonsingular. Now from

$$\begin{aligned} \det(K) &= \det(B_1) \cdot \det(0 - B_2 B_1^{-1} B_2^T) \\ &= (-1)^m \det(B_1) \cdot \det(B_2 B_1^{-1} B_2^T), \end{aligned}$$

we can immediately conclude that $\det(K) \neq 0$, i.e. K is nonsingular. \square

Because both B_1 and K are nonsingular (Theorem 2.3 and 2.4) the Schur complement matrix $B_2B_1^{-1}B_2^T$ is nonsingular. The unknowns in (13) can be obtained by solving

$$B_2B_1^{-1}B_2^T\delta\lambda = B_2B_1^{-1}\mathbf{f} - \mathbf{r}. \quad (14)$$

$$B_1\delta\mathbf{z} = \mathbf{f} - B_2^T\delta\lambda. \quad (15)$$

It is convenient to first present the direct solution for (15), followed by the iterative approach for (14).

2.3.1. solution of $B_1\delta\mathbf{z} = \mathbf{f} - B_2^T\delta\lambda$:

The right hand side can be expanded as

$$\mathbf{f} - B_2^T\delta\lambda = \begin{bmatrix} -\alpha - X^T\lambda \\ -(A+E)^T\lambda \end{bmatrix} - \begin{bmatrix} X^T\delta\lambda \\ (A+E)^T\delta\lambda \end{bmatrix} = \begin{bmatrix} -\alpha - X^T\lambda^{new} \\ -(A+E)^T\lambda^{new} \end{bmatrix},$$

where $\lambda^{new} = \lambda + \delta\lambda$ is the updated solution of λ after solving (14). Fast and stable algorithms exist for triangularization of Toeplitz matrices, $\Lambda = [R, 0]Q$, see e.g. the paper by Park and Eldèn.¹⁹ In contrast to the case of general Toeplitz matrices, the displacement rank of $\Lambda \in R^{n \times (m+n-1)}$ is just 2 because the southeast and northwest corners are zero triangles.¹⁹ The calculation of the generator (i.e. first row of R), and the hyperbolic transformation, cost $\mathcal{O}((m+n)\log_2(m+n))$ and $3n^2$, respectively. Note, in this operation count, and throughout, we assume the use of the FFT for any Toeplitz/Hankel matrix-vector multiplication or convolution. From (15) it is easy to obtain

$$\Lambda\Lambda^T\delta\mathbf{x} = \Lambda(-\alpha - X^T\lambda^{new}) + (A+E)^T\lambda^{new}.$$

Because $\Lambda\Lambda^T = RR^T$, the cost for forming the righthand side and solving two triangular systems is $\mathcal{O}((m+n)\log_2(m+n)) + 2n^2$. Then $\delta\alpha = -\alpha - X^T\lambda^{new} - \Lambda^T\delta\mathbf{x}$ is calculated directly with additional cost $(m+n)\log_2n$.

2.3.2. Solution of $B_2B_1^{-1}B_2^T\delta\lambda = B_2B_1^{-1}\mathbf{f} - \mathbf{r}$

Forming $B_2B_1^{-1}B_2^T$ is prohibitive, but Krylov subspace methods using only matrix-vector multiplications can be utilized. As shown above the multiplication involving B_1^{-1} can be performed efficiently. But from Theorem 2.3 B_1^{-1} is not positive definite, thus in general $B_2B_1^{-1}B_2^T$ is a symmetric indefinite matrix. In such a case the MINRES and SYMMLQ algorithms can be used.²⁰ Both algorithms are based on three term recurrence relations for

construction of an orthogonal basis for the Krylov subspace and only need matrix-vector multiplications at cost $\mathcal{O}((m+n)\log_2(m+n) + n^2)$.

Thus the overall cost, $\mathcal{O}((m+n)\log_2(m+n) + n^2)$, is much cheaper than the general method which does not consider the structure and costs $\mathcal{O}((m+n)^3)$. Moreover the procedure is numerically stable. The stability analysis for the triangularization of Toeplitz matrices was analyzed by Park and Eldèn¹⁹ and the convergence behaviors of MINRES and SYMMLQ were analyzed by Paige *et al.*²¹

3. Experiments

Experiments are performed using MATLAB 7.0. The original signal $\{s_k\}$ is generated as a pseudorandom sequence uniformly distributed on $[0, 1]$. $m = 200$ and $n = 21$. For each example additive Gaussian noise of differing signal-to-noise ratios (SNR) is added. The SNR in unit *dB* is defined as follows

$$\text{SNR} = 10 \log_{10} \frac{\mathbf{y}^T \mathbf{y}}{\alpha^T \alpha}.$$

At each noise level, one hundred simulations are performed and average results generated. In each simulation the ideal zero-forcing equalizer, i.e. the inverse of the original channel impulse function, is known and denoted by $\{h_k^{-1}\}$. We use two quantities to evaluate our results, the ISI and the relative error of the estimated equalizer to the ideal zero-forcing equalizer $\{h_k^{-1}\}$. They are defined as follows

$$\text{ISI} = \frac{\sum_{k=0}^{n-1} c_k^2}{\max_k c_k^2} - 1,$$

where $c_k = \sum_{l=0}^{n-1} h_{k-l} x_l$, and

$$\text{Relerr} = \frac{\sum_{k=0}^{n-1} (x_k - h_k^{-1})^2}{\sum_{k=0}^{n-1} (h_k^{-1})^2}.$$

In communication applications the ISI is of interest, while the relative error of the solution is relevant mathematically. As seen in the definitions above, the sequences for calculating these two quantities are truncated if the impulse responses are infinite sequences. The discarded portion ($l > 20$) can be ignored.

Experiment 1. For the finite impulse response(FIR) channel,

$$y_k = s_k + 0.7s_{k-1} + \alpha_k,$$

10

the ideal zero-forcing equalizer should have an infinite impulse response (IIR) in order to completely remove the ISI,

$$\{h_l^{-1}\} = \{1, -0.7, (-0.7)^2, \dots, (-0.7)^k, \dots\}.$$

Experiment 2. For the IIR channel,

$$y_k = s_k + 0.7y_{k-1} + \alpha_k,$$

the ideal zero-forcing equalizer is a FIR sequence

$$\{h_l^{-1}\} = \{1, -0.7, 0, \dots, 0\}.$$

Experiment 3. For the complex channel

$$h_k = \sqrt{2 + 2i}e^{-k-ki},$$

the inverse channel is given by

$$\{h_l^{-1}\} = \left(\frac{1}{\sqrt{2 + 2i}}, -\frac{e^{-1-i}}{\sqrt{2 + 2i}}, 0, 0, \dots \right).$$

Results are detailed in Figures 1(a) to 1(f). In order to focus on the most interesting results, we only illustrate curves in the range for $\text{SNR} > 12$. If $\text{SNR} > 18$, the SDLS always outperforms DLS and OLS. Moreover, DLS outperforms OLS when $\text{SNR} > 22$. These are not surprising results. Indeed DLS, and in particular SDLS, is a better model than OLS for the zero-forcing channel equalization. In contrast to the situation with high SNR, for high noise levels, $\text{SNR} < 10$, the simple OLS models the problem better. Again, as expected in all cases SDLS performs better than DLS.

4. Conclusion

A numerically stable and efficient algorithm based on sequential quadratic programming has been developed for the SDLS problem. We conclude that SDLS is the best model for the zero-forcing equalization problem when the signal to noise ratio is reasonably large as compared to OLS and DLS.

Appendix A. Algorithm for Data Least Squares

The derivation of the DLS algorithm is similar to that for the LS-TLS mixed problem,²² in which some columns of the coefficient matrix are error free. Both are based on the generalized Eckart-Young-Mirsky matrix approximation theorem.²³ Although the derivation is straightforward, we have not

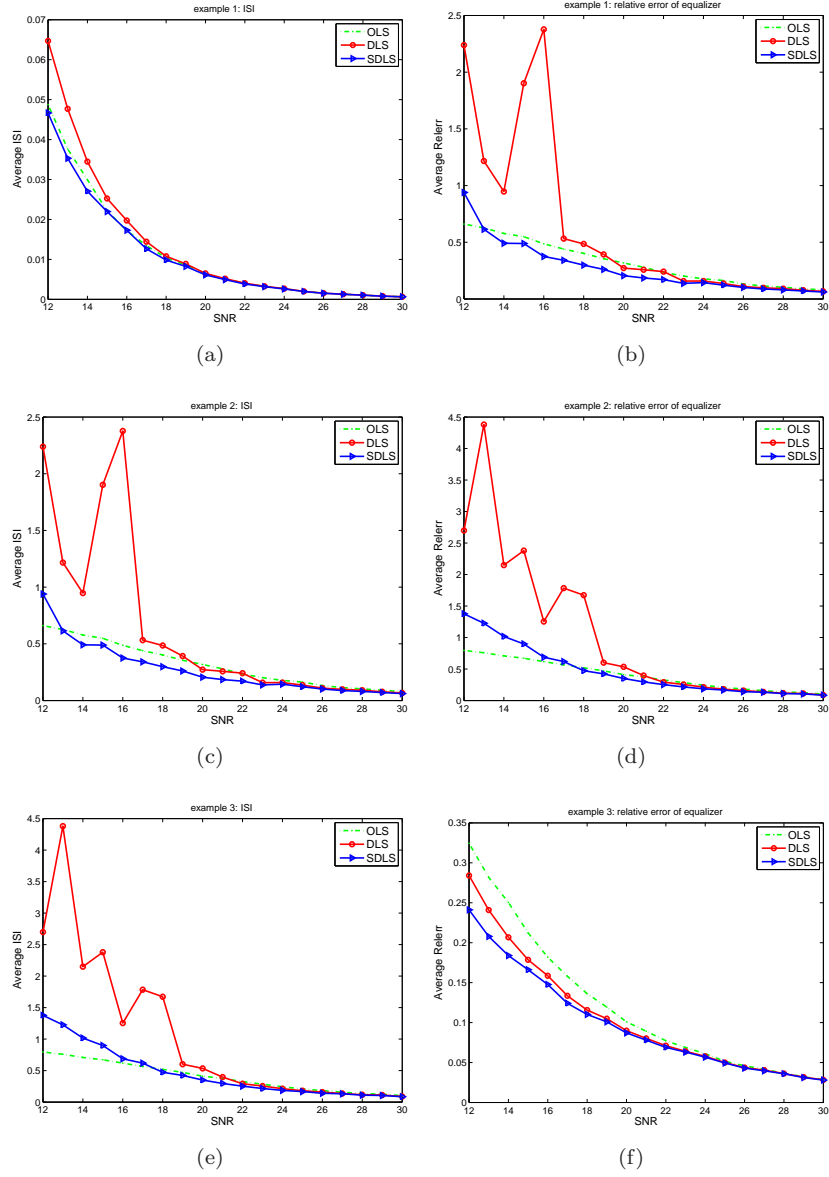


Fig. 1. Average ISI, left column, and average of Relerr, right column, for experiments 1 to 3, comparing methods MMSE-OLS, DLS and SDLS.

seen any literature which provides a clear exposition of the DLS algorithm. Let us rewrite the DLS problem as follows

$$\min \|E\|_F \quad \text{s.t.} \quad (\mathbf{s}, A + E) \begin{pmatrix} -1 \\ \mathbf{x} \end{pmatrix} = 0.$$

We perform Householder transformation such that $P\mathbf{s} = \|\mathbf{s}\|\mathbf{e}_1$ and denote PA by \bar{A} . Because the Frobenius norm is invariant under orthogonal transformation, the DLS problem is now

$$\min \|E\|_F \quad \text{s.t.} \quad (\|\mathbf{s}\|\mathbf{e}_1, \bar{A} + E) \begin{pmatrix} -1 \\ \mathbf{x} \end{pmatrix} = 0.$$

Rewriting the constraints in block style

$$\begin{pmatrix} \|\mathbf{s}\|, \bar{A}_1 + E_1 \\ 0, \bar{A}_2 + E_2 \end{pmatrix} \begin{pmatrix} -1 \\ \mathbf{x} \end{pmatrix} = 0, \quad \bar{A}_1 \in R^{1 \times n}, \quad \bar{A}_2 \in R^{(m-1) \times n},$$

shows that block (2, 2) is independent and associated with subproblem

$$\min \|E_2\|_F \quad \text{s.t.} \quad (\bar{A}_2 + E_2)\mathbf{x} = 0.$$

This is a regular TLS problem, the solution is given by

$$\mathbf{x} = k\mathbf{v}_n, \quad E_2 = -\gamma_n \mathbf{u}_n \mathbf{v}_n^T,$$

where $(\gamma_n, \mathbf{u}_n, \mathbf{v}_n)$ is the triple associated with the smallest singular value of \bar{A}_2 and scalar k is any real number. Any algorithm for TLS can be used directly to solve this subproblem.

To determine the scalar k we minimize $\|E_1\|$ subject to the first constraint equation and $\mathbf{x} = k\mathbf{v}_n$,

$$\min \|E_1\|_2 \quad \text{s.t.} \quad -\|\mathbf{s}\| + \bar{A}_1 k \mathbf{v}_n + E_1 k \mathbf{v}_n = 0.$$

Immediately we see $E_1 = 0$ and $k = \frac{\mathbf{s}}{\bar{A}_1 \mathbf{v}_n}$ if $\bar{A}_1 \mathbf{v}_n \neq 0$. If the original error matrix E is of interest, the calculated \bar{E} has to be transformed back by the Householder matrix P , namely form PE .

References

1. C. C. Paige and Z. Strakos, *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. 2534, (2002).
2. J. G. Proakis, *Digital Communications*. McGraw-Hill, 4th ed., New York, NY, 2001.
3. D. Tufts, *Nyquist's problem: The joint optimization of transmitter and receiver in pulse amplitude modulation*, Proc. IEEE, **53**, Mar. 1965. 248–259.

4. J. W. Smith, *The joint optimization of transmitted signal and receiving filter for data transmission system*, Bell Syst. Tech. J., **44**, 1965, 2363–2392.
5. G. H. Golub and C. Van Loan, *An analysis of the total least squares problem*, SIAM J. Numer. Anal., **17**, 1980, 883–893.
6. S. Van Huffel and J. Vandewalle, *The total least squares problem: computational aspects and analysis*, SIAM, Philadelphia, 1991.
7. H. Guo and R. A. Renaut, *A Regularized total least squares algorithm* in Total Least Squares and Errors-in-Variables Modeling, S. van Huffel and P. Lemmerling, editors, Kluwer Academic Publishers, Dordrecht, 57–66 (2002).
8. R. A. Renaut and H. Guo, *Efficient Algorithms for Solution of Regularized Total Least Squares*, SIAM J. Mat. Anal. Appl., **26**, 2005, 457–476.
9. M. Aoki and P. C. Yue, *On a priori error estimates of some identification methods*, IEEE Trans. Aut. Control, **15**, 1970, 541–548.
10. J. A. Cadzow, *Signal enhancement: A composite property mapping algorithm*, IEEE Trans. Acoust., Speech, Signal Processing, ASSP-**36**, 1988, 39–62.
11. T. Abatzoglou, J. Mendel and G. Harada, *The constrained total least squares technique and its application to harmonic super resolution*, IEEE Trans. Signal Proc., **39**, 1991, 1070–1087.
12. B. De Moor, *Structured total least squares and L_2 approximation problems*, Lin. Alg. and Its Appl, **188-189**, 1993, 163–207.
13. N. Mastronardi, P. Lemmerling, and S. Van Huffel, *Fast Structured Total Least Squares Algorithm for Solving the Basic Deconvolution Problem*, SIAM J. Mat. Anal. Appl., **22(2)**, 2000, 533–553.
14. A. Beck and A. Ben-Tal *A Global Solution for the Structured Total Least Squares Problem with Block Circulant Matrices*, SIAM J. Mat. Anal. Appl., **27(1)**, 2005, 238–255.
15. J.B. Rosen and H. Park and J. Glick, *Total least norm formulation and solution for structured problems*, SIAM J. Mat. Anal. Appl., **17(1)**, 1996, 110–128.
16. I. Markovsky and S. Van Huffel, *High-performance numerical algorithms and software for structured total least squares*, J. Comp. Appl. Math., **180(2)**, 2005, 311–331.
17. J. Nocedal and S. Wright, *Numerical Optimization*, Springer, New York, 1999.
18. B. Fischer and A. Ramage and D. J. Silvester and A. J. Wathen, *Minimum residual methods for Augmented Systems* BIT, 1998, **38(3)**, 527–543.
19. Haesun Park and Lars Eldèn, *Stability Analysis and Fast Algorithms for Triangularization of Toeplitz Matrices*, Numerische Mathematik, **76**, 1997, 383–402.
20. C. Paige and M. Saunders, *Solution of sparse indefinite systems of linear equations*, SIAM J. Numer. Anal., **12**, 1975, 617–629.
21. C. Paige and B. Parlett and H. Van Der Vorst, *Approximate solutions and eigenvalue bounds from Krylov subspaces*, Numer. Lin. Alg. Appls., **29**, 1995, 115–134.
22. S. Van Huffel and J. Vandewalle, *Analysis and properties of the generalized total least squares problem $AX \approx B$ when some or all columns in A are subject to error*, SIAM J. Mat. Anal. Appl., **10(3)**, 1989, 294–315.

23. G. H. Golub and A. Hoffman and G. W. Stewart, *A generalization of the Eckart-Young-Mirsky matrix approximation theorem*, Linear Algebra Appl. **88/89**, 1987, 317–327.