

RESEARCH

Open Access



ℓ_1 -regularized recursive total least squares based sparse system identification for the error-in-variables

Jun-seok Lim* and Hee-Suk Pang

*Correspondence:
jslim@sejong.ac.kr
Department of Electronics
Engineering, Sejong
University, Kunja, Kwangjin,
98, 143-747 Seoul, Korea

Abstract

In this paper an ℓ_1 -regularized recursive total least squares (RTLS) algorithm is considered for the sparse system identification. Although recursive least squares (RLS) has been successfully applied in sparse system identification, the estimation performance in RLS based algorithms becomes worse, when both input and output are contaminated by noise (the error-in-variables problem). We proposed an algorithm to handle the error-in-variables problem. The proposed ℓ_1 -RTLS algorithm is an RLS like iteration using the ℓ_1 regularization. The proposed algorithm not only gives excellent performance but also reduces the required complexity through the effective inversion matrix handling. Simulations demonstrate the superiority of the proposed ℓ_1 -regularized RTLS for the sparse system identification setting.

Keywords: Adaptive filter, TLS, RLS, Convex regularization, Sparsity, ℓ_1 -norm

Background

There has been a recent interest in adaptive algorithms to handle sparsity in various signals and systems (Gu et al. 2009; Chen et al. 2009; Babadi et al. 2010; Angelosante et al. 2010; Eksioglu 2011; Eksioglu and Tanc 2011; Kalouptsidis et al. 2011). The idea is to exploit a priori knowledge about sparsity in a signal that needs to be processed for system identification. Several algorithms based on the least-mean square (LMS) (Gu et al. 2009; Chen et al. 2009) and the recursive least squares (RLS) (Babadi et al. 2010; Angelosante et al. 2010; Eksioglu 2011; Eksioglu and Tanc 2011) techniques have been reported with different penalty or shrinkage functions. In a broad range of signal processing applications, not only the system output is corrupted by measurement noise, but also the measured input signal may often be corrupted by the additive noise due to such as sampling error, quantization error and wide-band channel noise. However, the algorithms for sparsity can handle only the corrupted output case. It is necessary to derive an algorithm to handle a noisy case of both noisy input and noisy output (i.e. the error-in-variables problem).

One of the potential counterparts to handle the error-in-variables problem is the total-least-squares estimator (TLS) that seeks to minimize the sum of squares of residuals on all of the variables in the equation instead of minimizing the sum of squares of residuals

on only the response variable. Golub and Van Loan introduced the TLS problem to the field of numerical analysis (Golub and Loan 1980); consequently, other researchers have developed and analyzed adaptive algorithms that employ the TLS formulation and its extensions (Dunne and Williamson 2000, 2004; Feng et al. 1998; Arablouei et al. 2014; Davila 1994; Arablouei et al. 2015). Recursive based algorithms were also studied along with adaptive TLS estimators (Soijer 2004; Choi et al. 2005). The algorithms were denoted as recursive total least squares (RTLS) or sequential total least squares (STLS). These algorithms recursively calculate and track the eigenvector corresponding to the minimum eigenvalue (the TLS solution) from the inverse covariance matrix of the augmented sample matrix. Some TLS based algorithms were proposed for the sparse signal processing (Tanc 2015; Arablouei 2016; Zhu et al. 2011; Dumitrescu 2013; Lim and Pang 2016). The algorithms in Tanc (2015), Arablouei (2016) utilized the gradient based method. The algorithms in Zhu et al. (2011), Dumitrescu (2013) were based on the block coordinate descent method. In Lim and Pang (2016), the TLS method was applied to handle the group sparsity problem.

In this paper, we consider the ℓ_1 regularization for the RTLS cost function, in which the recursive procedure is derived from the generalized eigendecomposition method in Davila (1994) and Choi et al. (2005), and the regularization approach outlined in Eksioglu and Tanc (2011) is used in order to handle the sparsity. We develop the update algorithm for the ℓ_1 -regularized RTLS using results from subgradient calculus. As a result, we propose the algorithm superior to the algorithm of Eksioglu and Tanc (2011) in the error-in-variables. We also reduce the total complexity by utilizing the inverse matrix update effectively. The proposed algorithm improves the sparse system estimation performance in the error-in-variables with only a little additional complexity. We provide simulation results to examine the performance of the proposed algorithm in comparison with the algorithm of Eksioglu and Tanc (2011).

Sparse system identification problem

In the sparse system identification problem of interest, the system observes a signal represented by an $M \times 1$ vector $\mathbf{x}(k) = [x_1(k), \dots, x_M(k)]^T$ at time instant k , performs filtering and obtains the output $y(k) = \mathbf{x}^T(k)\mathbf{w}_o(k)$, where $\mathbf{w}_o(k) = [w_1(k), \dots, w_M(k)]^T$ is an M -length finite-impulse-response (FIR) system that represents the actual system. For system identification, an adaptive filter with M coefficients $\hat{\mathbf{w}}(k)$ is employed in such a way that observes $\mathbf{x}(k)$ and produces an estimate $\hat{y}(k) = \mathbf{x}^T(k)\hat{\mathbf{w}}(k)$. The system identification scheme then compares the output of the actual system $y(k)$ and the adaptive filter $\hat{y}(k)$, resulting in an error signal $e(k) = y(k) + n(k) - \hat{y}(k) = \tilde{y}(k) - \hat{y}(k)$, where $n(k)$ is the measurement noise. In this context, the goal of an adaptive algorithm is to identify the system by minimizing the cost function defined by

$$\hat{\mathbf{w}} = \arg \min_{\hat{\mathbf{w}}} \frac{1}{2} \sum_{m=0}^k \lambda^{k-m} (e(m))^2. \quad (1)$$

The gradient based minimization derives the following equation.

$$\Phi(k)\hat{\mathbf{w}}(k) = \mathbf{r}(k), \quad (2)$$

where $\Phi(k) = \sum_{m=0}^k \lambda^{k-m} \mathbf{x}(m) \mathbf{x}^T(m)$ and $\mathbf{r}(k) = \sum_{m=0}^k \lambda^{k-m} \tilde{y}(m) \mathbf{x}(m)$. This equation is the matrix form of the normal equations for least squares solution.

Especially, we call M -th order $\mathbf{w}_o(k)$ sparse system, when the number of nonzero coefficients $K \ll M$. In order to estimate the M -th order sparse system, most estimation algorithms exploit non-zero coefficients of the system to obtain performance benefits and/or a computational complexity reduction (Gu et al. 2009; Chen et al. 2009; Babadi et al. 2010; Angelosante et al. 2010; Eksioglu 2011; Eksioglu and Tanc 2011).

ℓ_1 -regularized RTLS (recursive total least squares)

In TLS, we assume a given unknown system that both the input and output are corrupted by noise. The system should be estimated from the noisy observation of the input and the output (Fig. 1). In this system output is given by:

$$\tilde{y}(k) = \tilde{\mathbf{x}}^T(k) \mathbf{w}_o + n_o(k), \quad (3)$$

where the output noise $n_o(k)$ is the Gaussian white noise with variance σ_o^2 and independent of the input signal. The noisy input of the system is given by:

$$\tilde{x}(k) = x(k) + n_i(k), \quad (4)$$

where the input noise $n_i(k)$ is the Gaussian white noise with variance σ_i^2 .

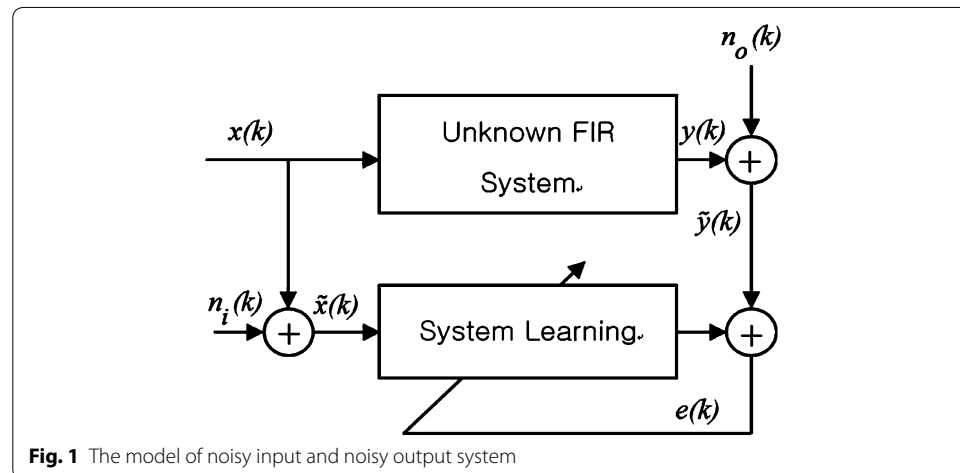
For TLS solution, the augmented data vector is considered as:

$$\bar{\mathbf{x}}(k) = [\tilde{\mathbf{x}}^T(k), \tilde{y}(k)]^T \in \mathbb{R}^{(M+1) \times 1}. \quad (5)$$

And its covariance matrix has the following structure:

$$\bar{\mathbf{R}} = \begin{bmatrix} \tilde{\mathbf{R}} & \mathbf{p} \\ \mathbf{p}^T & c \end{bmatrix}, \quad (6)$$

where $\mathbf{p} = E\{\tilde{\mathbf{x}}(k)y(k)\}$ and $c = E\{y(k)y(k)\}$, $\tilde{\mathbf{R}} = E\{\tilde{\mathbf{x}}(k)\tilde{\mathbf{x}}^T(k)\} = \mathbf{R} + \sigma_i^2 \mathbf{I}$, $\mathbf{R} = E\{\mathbf{x}(k)\mathbf{x}^T(k)\}$. In Davila (1994) and Choi et al. (2005), the TLS problem is thus reduced to finding the eigenvector that is associated with the smallest eigenvalue of $\bar{\mathbf{R}}$. The



following equation is the simplified cost function to find the eigenvector that is associated with the smallest eigenvalue of $\bar{\mathbf{R}}$.

$$\tilde{J}(\mathbf{w}) = \frac{\tilde{\mathbf{w}}^T \bar{\mathbf{R}} \tilde{\mathbf{w}}}{\tilde{\mathbf{w}}^T \bar{\mathbf{D}} \tilde{\mathbf{w}}} = \frac{[\mathbf{w}^T, -1] \bar{\mathbf{R}} [\mathbf{w}^T, -1]^T}{[\mathbf{w}^T, -1] \bar{\mathbf{D}} [\mathbf{w}^T, -1]^T}, \quad (7)$$

where $\bar{\mathbf{D}} = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \gamma \end{bmatrix}$ with $\gamma = \frac{\sigma_o^2}{\sigma_i^2}$. Minimum value of (7) is recognized as the smallest generalized eigenvalue of $\bar{\mathbf{R}}$ (Dunne and Williamson 2004). Therefore, we can find the eigenvector associated with the smallest eigenvalue. The smallest eigenvector can also be derived from the maximization of (8).

$$\tilde{J}(\mathbf{w}) = \frac{\tilde{\mathbf{w}}^H \bar{\mathbf{D}} \tilde{\mathbf{w}}}{\tilde{\mathbf{w}}^H \bar{\mathbf{R}} \tilde{\mathbf{w}}} = \frac{[\mathbf{w}^H, -1] \bar{\mathbf{D}} [\mathbf{w}^T, -1]^T}{[\mathbf{w}^H, -1] \bar{\mathbf{R}} [\mathbf{w}^T, -1]^T}. \quad (8)$$

We adopt the modified cost function by the addition of a penalty function. This penalty function can be chosen to reflect a priori knowledge about the true sparsity system.

$$\tilde{J}(\mathbf{w}) = \frac{[\mathbf{w}^H, -1] \bar{\mathbf{D}} [\mathbf{w}^T, -1]^T}{[\mathbf{w}^H, -1] \bar{\mathbf{R}} [\mathbf{w}^T, -1]^T} + \gamma f(\tilde{\mathbf{w}}) = \frac{\tilde{\mathbf{w}}^H \bar{\mathbf{D}} \tilde{\mathbf{w}}}{\tilde{\mathbf{w}}^H \bar{\mathbf{R}} \tilde{\mathbf{w}}} + \gamma f(\tilde{\mathbf{w}}), \quad (9)$$

where γ is the regularized parameter in Eksioglu and Tanc (2011). We adopt the ℓ_1 penalty function as follows:

$$f(\tilde{\mathbf{w}}) = \|\tilde{\mathbf{w}}\|_1 = \sum_{i=0}^{M-1} |\tilde{w}_i|. \quad (10)$$

We solve the equations by $\nabla_{\tilde{\mathbf{w}}} J(k) = 0$,

$$\begin{aligned} \nabla_{\tilde{\mathbf{w}}} J(k) = 0 : & \frac{\bar{\mathbf{D}} \tilde{\mathbf{w}}(k) (\tilde{\mathbf{w}}^H(k) \bar{\mathbf{R}}(k) \tilde{\mathbf{w}}(k))}{(\tilde{\mathbf{w}}^H(k) \bar{\mathbf{R}}(k) \tilde{\mathbf{w}}(k))^2} - \frac{\tilde{\mathbf{w}}^H(k) \bar{\mathbf{D}} \tilde{\mathbf{w}}(k) (\bar{\mathbf{R}}(k) \tilde{\mathbf{w}}(k))}{(\tilde{\mathbf{w}}^H(k) \bar{\mathbf{R}}(k) \tilde{\mathbf{w}}(k))^2} \\ & + \gamma_k \nabla f(\tilde{\mathbf{w}}(k)) = 0. \end{aligned} \quad (11)$$

where $\bar{\mathbf{R}}(k) = \lambda \bar{\mathbf{R}}(k-1) + \bar{\mathbf{x}}(k) \bar{\mathbf{x}}^H(k)$, $\tilde{\mathbf{w}}(k) = [\hat{\mathbf{w}}^T(k), -1]^T$ and $\hat{\mathbf{w}}(k)$ is the estimation result for the unknown system at k -th time step. The subgradient of $f(\tilde{\mathbf{w}})$ is $\nabla^s \|\tilde{\mathbf{w}}\|_1 = \text{sgn}(\tilde{\mathbf{w}})$ (Babadi et al. 2010; Kalouptsidis et al. 2011), and $\text{sgn}(\cdot)$ is the component-wise sign function. In (11), the regularized parameter, γ_k , is time-varying, which governs a tradeoff between the approximation error and the penalty function. From (11), we obtain

$$\begin{aligned} \tilde{\mathbf{w}}(k) = & \frac{\tilde{\mathbf{w}}^H(k) \bar{\mathbf{R}}(k) \tilde{\mathbf{w}}(k)}{\|\tilde{\mathbf{w}}(k)\|^2} \left(\bar{\mathbf{R}}^{-1}(k) \tilde{\mathbf{w}}(k) \right) \\ & + \gamma_k \frac{(\tilde{\mathbf{w}}^H(k) \bar{\mathbf{R}}(k) \tilde{\mathbf{w}}(k))^2}{\|\tilde{\mathbf{w}}(k)\|^2} \bar{\mathbf{R}}^{-1}(k) \nabla f(\tilde{\mathbf{w}}(k)). \end{aligned} \quad (12)$$

And we obtain the estimated parameter of the unknown system as

$$\hat{\mathbf{w}}(k) = -\tilde{\mathbf{w}}_{1:M}(k) / \tilde{w}_{M+1}(k). \quad (13)$$

As the estimated unknown system can be derived from the ratio between $\tilde{w}_{M+1}(k)$ and the elements in $\tilde{\mathbf{w}}_{1:M}(k)$, the normalization of $\tilde{\mathbf{w}}(k)$ as $\tilde{\mathbf{w}}(k) = \tilde{\mathbf{w}}(k)/\|\tilde{\mathbf{w}}(k)\|$ keeps the same solution in (13) as well as numerical stability in the iteration. By applying the normalization of (12) becomes,

$$\begin{aligned}\tilde{\mathbf{w}}(k) &= (\tilde{\mathbf{w}}^H(k)\bar{\mathbf{R}}(k)\tilde{\mathbf{w}}(k)) \\ &\times \left(\bar{\mathbf{R}}^{-1}(k)\tilde{\mathbf{w}}(k) + \gamma_k(\tilde{\mathbf{w}}^H(k)\bar{\mathbf{R}}(k)\tilde{\mathbf{w}}(k))\bar{\mathbf{R}}^{-1}(k)\nabla f(\tilde{\mathbf{w}}(k)) \right).\end{aligned}\quad (14)$$

In addition, we can approximate (14) as follows.

$$\begin{aligned}\tilde{\mathbf{w}}(k) &\approx \left(\bar{\mathbf{R}}^{-1}(k)\tilde{\mathbf{w}}(k-1) \right) \\ &+ \gamma_k(\tilde{\mathbf{w}}^H(k-1)\bar{\mathbf{R}}(k-1)\tilde{\mathbf{w}}(k-1))\bar{\mathbf{R}}^{-1}(k)\nabla f(\tilde{\mathbf{w}}(k-1)),\end{aligned}\quad (15)$$

where $\bar{\mathbf{R}}(k) = \lambda\bar{\mathbf{R}}(k-1) + \bar{\mathbf{x}}(k)\bar{\mathbf{x}}^H(k)$. In (15),

$$\begin{aligned}\tilde{\mathbf{w}}^H(k)\bar{\mathbf{R}}(k)\tilde{\mathbf{w}}(k) &= \lambda\tilde{\mathbf{w}}^H(k)\bar{\mathbf{R}}(k-1)\tilde{\mathbf{w}}(k) + \tilde{\mathbf{w}}^H(k)\bar{\mathbf{x}}(k)\bar{\mathbf{x}}^H(k)\tilde{\mathbf{w}}(k) \\ &\approx \lambda\tilde{\mathbf{w}}^H(k-1)\bar{\mathbf{R}}(k-1)\tilde{\mathbf{w}}(k-1) + \tilde{\mathbf{w}}^H(k)\bar{\mathbf{x}}(k)\bar{\mathbf{x}}^H(k)\tilde{\mathbf{w}}(k) \\ &= \lambda\tilde{\mathbf{w}}^H(k-1)\bar{\mathbf{R}}(k-1)\tilde{\mathbf{w}}(k-1) + \bar{y}(k)^2.\end{aligned}\quad (16)$$

In (11), we apply the same γ_k in Eksioglu and Tanc (2011).

$$\gamma_k = \frac{2 \frac{\text{tr}(\bar{\mathbf{R}}^{-1}(k))}{M} (f(\hat{\mathbf{w}}_{aug}(k)) - \rho) + \nabla^s f(\hat{\mathbf{w}}_{aug}(k))\bar{\mathbf{R}}^{-1}(k)\varepsilon(k)}{\|\bar{\mathbf{R}}^{-1}(k)\nabla^s f(\hat{\mathbf{w}}_{aug}(k))\|_2^2}, \quad (17)$$

where $\hat{\mathbf{w}}_{aug}(k) = [\hat{\mathbf{w}}^T(k), -1]^T$, $\hat{\mathbf{w}}_{aug,RLS}(k) = [\hat{\mathbf{w}}_{RLS}^T(k), -1]^T$, $\varepsilon(k) = \hat{\mathbf{w}}_{aug}(k) - \hat{\mathbf{w}}_{aug,RLS}(k)$, and $\hat{\mathbf{w}}_{RLS}(k)$ is the parameter estimated by recursive least squares (RLS).

A simplified way to solve ℓ_1 -regularized RTLS

The proposed algorithm needs a solution in (15) as well as RLS solution for $\hat{\mathbf{w}}_{RLS}(k)$ in γ_k . However, this makes the algorithm complex and we find a less complex way from block matrix inversion lemma in Moon and Stirling (2000). The required calculation complex-

ity can be simplified if we use the following matrix manipulation: $\bar{\mathbf{X}}(k) = \left[\tilde{\mathbf{X}}(k)^T; \tilde{\mathbf{y}}(k) \right]^T$

and $\bar{\mathbf{X}}(k)\bar{\mathbf{X}}^T(k) = \begin{bmatrix} \tilde{\mathbf{X}}(k)\tilde{\mathbf{X}}^T(k) & \tilde{\mathbf{X}}(k)\tilde{\mathbf{y}}(k) \\ \tilde{\mathbf{y}}^T(k)\tilde{\mathbf{X}}^T(k) & \tilde{\mathbf{y}}^T(k)\tilde{\mathbf{y}}(k) \end{bmatrix} = \begin{bmatrix} \tilde{\mathbf{X}}(k)\tilde{\mathbf{X}}^T(k) & \mathbf{a} \\ \mathbf{a}^T & c \end{bmatrix}$, then:

$$\begin{aligned}\bar{\mathbf{R}}^{-1}(k) &= \left(\bar{\mathbf{X}}(k)\bar{\mathbf{X}}^T(k) \right)^{-1} \\ &= \begin{bmatrix} \mathbf{P}_{xx}(k) + \beta\mathbf{P}_{xx}(k)\mathbf{a}\mathbf{a}^T\mathbf{P}_{xx}(k) & -\beta\mathbf{P}_{xx}(k)\mathbf{a} \\ -\beta\mathbf{a}^T\mathbf{P}_{xx}(k) & \beta \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix},\end{aligned}\quad (18)$$

where $\tilde{\mathbf{X}}(k) = \left[\tilde{\mathbf{x}}(k), \sqrt{\lambda}\tilde{\mathbf{x}}(k-1), \dots, \left(\sqrt{\lambda} \right)^k \tilde{\mathbf{x}}(0) \right]$, $\beta = \left(c - \mathbf{a}^H \left(\tilde{\mathbf{X}}(k)\tilde{\mathbf{X}}^T(k) \right)^{-1} \mathbf{a} \right)^{-1}$.

$\mathbf{A}_{12} = -\beta \left(\tilde{\mathbf{X}}(k)\tilde{\mathbf{X}}^T(k) \right)^{-1} \mathbf{a}$ in (18) includes

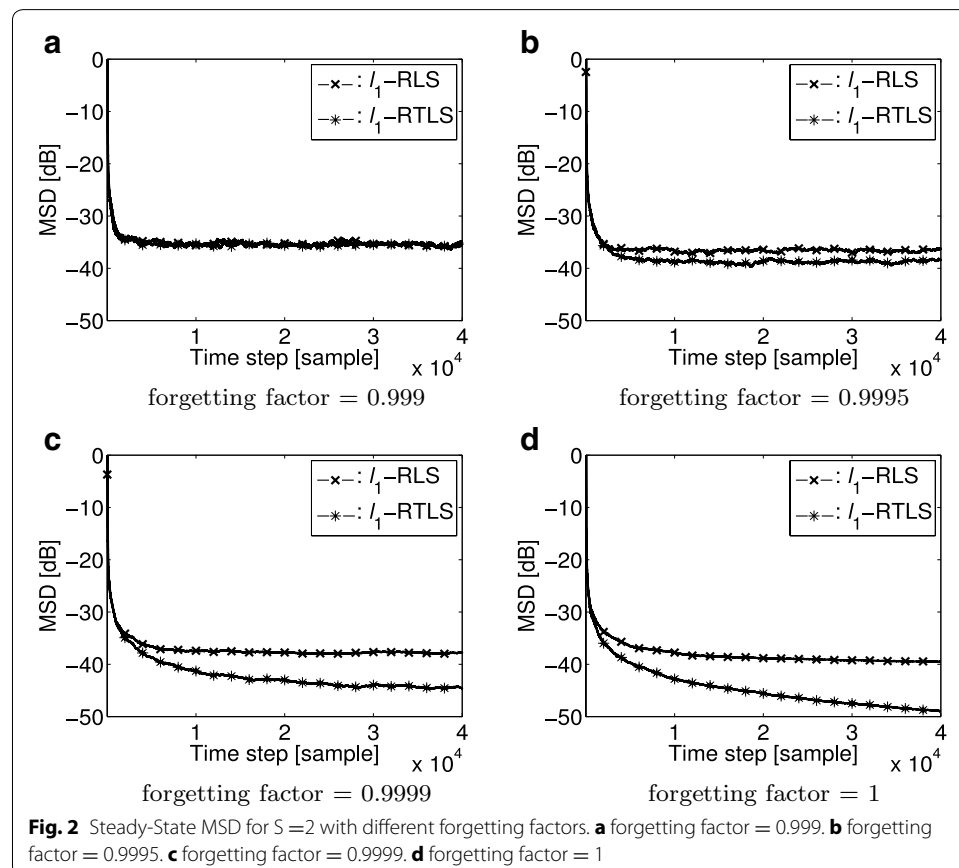
$$\left(\tilde{\mathbf{X}}(k)\tilde{\mathbf{X}}(k)^T\right)^{-1}\mathbf{a}=\left(\tilde{\mathbf{X}}(k)\tilde{\mathbf{X}}(k)^T\right)^{-1}\tilde{\mathbf{X}}(k)\tilde{\mathbf{y}}(k), \quad (19)$$

from which the RLS solution, $\hat{\mathbf{w}}_{RLS}(k)$ can be derived by dividing \mathbf{A}_{12} with the constant $-\beta$. When we solve the proposed ℓ_1 -RTLS in (15–17), γ_k in (17) needs to solve the RLS for $\varepsilon(k)$. Therefore, all procedures in (15–17) need $2M^2 + 2M$ additional multiplications. However, the simplification in this section needs only an additional division.

Simulation results

In this experiment, we follow the experiment scenario in Eksioglu and Tanc (2011). The true system function \mathbf{w} has a total of $N = 64$ taps, where only S of them are nonzero. The nonzero coefficients are positioned randomly and take their values from an $N(0, 1/S)$ distribution. The input signal is $x_k \sim N(0, 1)$. Noise is added to both the input and the output, and the additive input and output noises in this paper are $n_{in,k} \sim N(0, \sigma_{in}^2)$ and $n_{out,k} \sim N(0, \sigma_{out}^2)$, respectively. These additional noises are necessary to experiment the errors-in-variables problem. The proposed ℓ_1 -RTLS algorithm is realized with the automatic γ_k using (17). The ρ value in (17) is taken to be the true value of $f(\mathbf{w})$ as in Eksioglu and Tanc (2011), that is $\rho = \|\mathbf{w}_o\|_1$ for the ℓ_1 -RTLS. We also compare the ordinary RLS and the ℓ_1 -RLS of Eksioglu and Tanc (2011) with the proposed ℓ_1 -RTLS.

The proposed algorithm needs the inversion of the covariance matrix as (15) in order to derive the eigenvector. The better covariance matrix is needed for the better eigenvector. Therefore, the forgetting factor is needed close to 1. Figure 2 compares the



estimation performance results between the ℓ_1 -RLS and the ℓ_1 -RTLS in mean square deviation (MSD) with $S = 4$ and the different forgetting factors. We add the noise at both input and output with $\sigma_{in} = \sigma_{out} = 0.1$. For this comparison, we set the forgetting factor to 0.999, 0.9995, 0.9999 and 1, respectively. Figure 2 shows that the performance becomes better as the forgetting factor goes to 1 although the proposed ℓ_1 -RTLS becomes better when the forgetting factor is greater than 0.999.

In Fig. 3 we simulate the algorithms with $\sigma_{in} = \sigma_{out} = 0.1$ and for $S = 4, 8, 16$ and 64, where $S = 64$ corresponds to a completely non-sparse system. In this simulation we set the forgetting factor to 0.9999. Figure 3a–d plot the MSD curves of the proposed ℓ_1 -RTLS, the ℓ_1 -RLS and the ordinary RLS with the different S values. Figure 3 includes the MSD curves from the ℓ_1 -RLS and the ordinary RLS with the contaminated output only, and shows the estimation performance of the ℓ_1 -RLS and the ordinary RLS significantly deteriorates when input and output are contaminated with noise. The proposed ℓ_1 -RTLS, however, outperforms the ℓ_1 -RLS and the ordinary RLS when input and output are contaminated with noise.

Table 1 summarizes the steady-state MSD values at the end of 500 independent trials for the algorithms. In this simulation, we set the forgetting factor to 0.999, 0.9995, 0.9999 and 1 and vary the sparsity S to 4, 8, 16 and 64, respectively. Table 1 shows the performance of the ℓ_1 -RLS is almost the same as that of the ordinary RLS. This means the ℓ_1 -RLS cannot improve the estimation performance when both input and output are

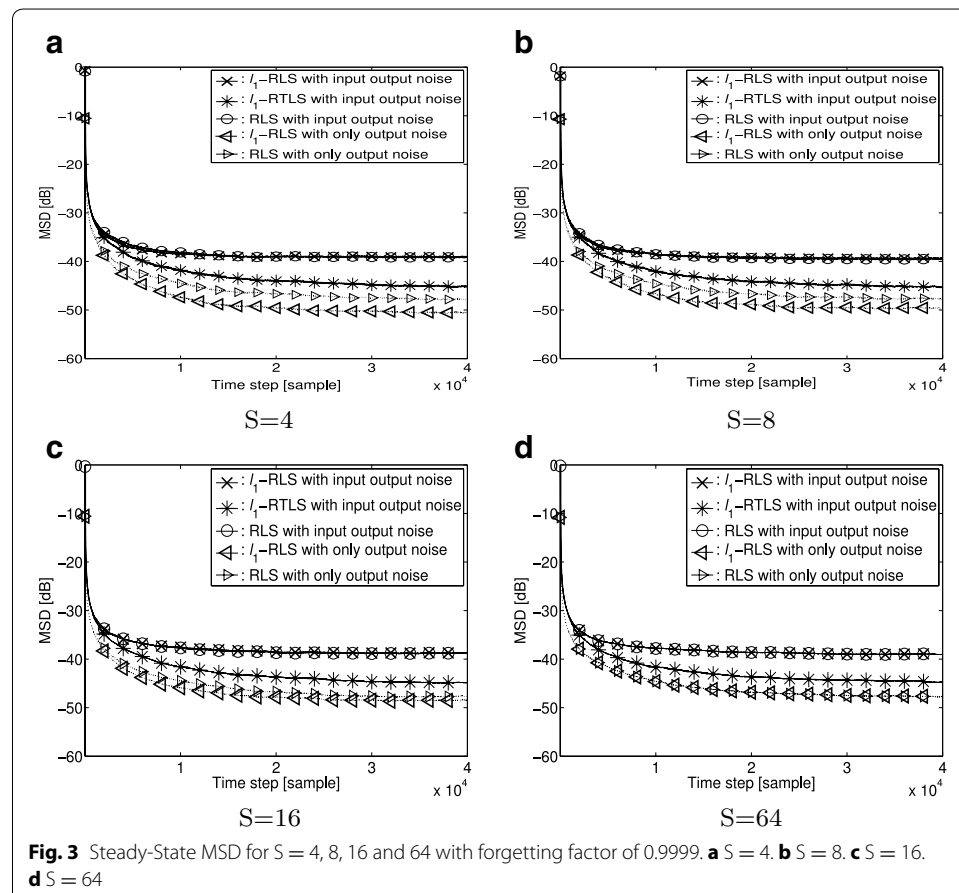


Table 1 Comparison of MSD in case of noisy input and noisy output

Forgetting factor	Algorithm	S = 4 (dB)	S = 8 (dB)	S = 16 (dB)	S = 64 (dB)
0.999	RLS	−34.0	−33.5	−33.8	−33.7
	ℓ_1 -RLS	−34.8	−33.6	−33.5	−33.7
	ℓ_1 -RTLS	−35.4	−35.0	−34.9	−34.8
	$(\ell_1\text{-RLS})-(\ell_1\text{-RTLS})$	0.6	1.4	1.4	1.1
0.9995	RLS	−35.7	−35.9	−36.1	−35.9
	ℓ_1 -RLS	−36.0	−35.8	−35.8	−35.9
	ℓ_1 -RTLS	−38.1	−38.0	−38.0	−38.0
	$(\ell_1\text{-RLS})-(\ell_1\text{-RTLS})$	2.1	2.2	2.2	2.1
0.9999	RLS	−39.0	−39.3	−38.6	−38.6
	ℓ_1 -RLS	−38.9	−39.1	−38.5	−38.6
	ℓ_1 -RTLS	−44.0	−44.1	−43.6	−43.6
	$(\ell_1\text{-RLS})-(\ell_1\text{-RTLS})$	5.1	5.0	5.1	5.0
1	RLS	−39.0	−38.5	−38.6	−38.9
	ℓ_1 -RLS	−39.0	−38.5	−38.6	−38.9
	ℓ_1 -RTLS	−45.0	−44.6	−44.8	−44.7
	$(\ell_1\text{-RLS})-(\ell_1\text{-RTLS})$	6.0	6.1	6.2	5.8

* ℓ_1 -RLS is the algorithm in Eksioglu and Tanc (2011)

contaminated by noise. However, the proposed ℓ_1 -RTLS outperforms the other algorithms. The improvement goes better as the forgetting factor gets close to 1.

Conclusion

In this paper, we propose an ℓ_1 -regularized RTLS for sparse system identification. The proposed algorithm keeps good performance in case of both noisy input and noisy output. We develop the recursive procedure for total least squares solution with an ℓ_1 -regularized cost function. We also present a simplified solution requiring only a little additional complexity in order to integrate the regularization factor. Simulations show that the introduced ℓ_1 -regularized RTLS algorithm shows better performance than RLS and ℓ_1 -regularized RLS in the sparse system with noisy input and noisy output.

Authors' contributions

This paper considers the sparse system identification. Especially, authors contribute to propose a new recursive sparse system identification algorithm, when both input and output are contaminated by noise. The algorithm not only gives excellent performance but also reduces the required complexity. Both authors read and approved the final manuscript.

Acknowledgements

This paper was supported by Agency for Defense Development (ADD) in Korean (15-106-104-027).

Competing interests

The authors declare that they have no competing interests.

Received: 21 September 2015 Accepted: 22 August 2016

Published online: 31 August 2016

References

- Angelosante D, Bazerque J, Giannakis G (2010) Online adaptive estimation of sparse signals: where RLS meets the ℓ_1 -norm. *IEEE Trans Signal Process* 58:3436–3447
- Arablouei R, Werner S, Dogancay K (2014) Analysis of the gradient-decent total least-squares algorithm. *IEEE Trans Signal Process* 62:1256–1264

- Arablouei R, Dogancay K, Werner S (2015) Recursive total least-squares algorithm based on inverse power method and dichotomous coordinate-descent iterations. *IEEE Trans Signal Process* 63:1941–1949
- Arablouei R (2016) Fast reconstruction algorithm for perturbed compressive sensing based on total least-squares and proximal splitting. *Signal Process*. doi:10.1016/j.sigpro.2016.06.009
- Babadi B, Kalouptsidis N, Tarokh V (2010) SPARLS: the sparse RLS algorithm. *IEEE Trans Signal Process* 58:4013–4025
- Chen Y, Gu Y, Hero A (2009) Sparse LMS for system identification. In: Paper presented at IEEE international conference on acoustics, speech and signal processing, Taiwan, 19–24 April, 2009
- Choi N, Lim J, Sung K (2005) An efficient recursive total least squares algorithm for training multilayer feedforward neural networks. *LNCS* 3496:558–565
- Davila C (1994) An efficient recursive total least squares algorithm for FIR adaptive filtering. *IEEE Trans Signal Process* 42:268–280
- Dumitrescu B (2013) Sparse total least squares: analysis and greedy algorithms. *Linear Algebra Appl* 438:2661–2674
- Dunne B, Williamson G (2000) Stable simplified gradient algorithms for total least squares filtering. In: Paper presented at the 32nd annual asilomar conference on signals, systems, and computers, Pacific Grove, Oct 29–Nov 1 2000
- Dunne B, Williamson G (2004) Analysis of gradient algorithms for TLS-based adaptive IIR filters. *IEEE Trans Signal Process* 52:3345–3356
- Eksioglu E (2011) Sparsity regularized RLS adaptive filtering. *IET Signal Process* 5:480–487
- Eksioglu E, Tanc A (2011) RLS algorithm with convex regularization. *IEEE Signal Process Lett* 18:470–473
- Feng D, Bao Z, Jiao L (1998) Total least mean squares algorithm. *IEEE Trans Signal Process* 46:2122–2130
- Golub G, Van Loan C (1980) An analysis of the total least squares problem. *SIAM J Numer Anal* 17:883–893
- Gu Y, Jin J, Mei S (2009) Norm constraint LMS algorithm for sparse system identification. *IEEE Signal Process Lett* 16:774–777
- Kalouptsidis N, Mileounis G, Babadi B, Tarokh V (2011) Adaptive algorithms for sparse system identification. *Signal Process* 91:1910–1919
- Lim J, Pang H (2016) Mixed norm regularized recursive total least squares for group sparse system identification. *Int J Adapt Control Signal Process* 30:664–673
- Moon T, Stirling W (2000) Mathematical methods and algorithm for signal processing. Prentice Hall, New Jersey
- Soijer MW (2004) Sequential computation of total least-squares parameter estimates. *J Guidance* 27:501–503
- Tanc A (2015) Sparsity regularized recursive total least-squares. *Digital Signal Process* 40:176–180
- Zhu H, Leus G, Giannakis G (2011) Sparsity-cognizant total least-squares for perturbed compressive sampling. *IEEE Trans Signal Process* 59:2002–2016

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Immediate publication on acceptance
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► springeropen.com
