Digital Signal Processing 36 (2015) 82–92

Contents lists available at ScienceDirect

Digital Signal Processing

www.elsevier.com/locate/dsp

Robust least squares methods under bounded data uncertainties

N. Denizcan Vanli [a,*], Mehmet A. Donmez [b], Suleyman S. Kozat [a]

[a] Department of Electrical and Electronics Engineering, Bilkent University, Ankara, Turkey
[b] Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign, IL, USA

ARTICLE INFO

Article history:
Available online 24 October 2014

Keywords:
Data estimation
Least squares
Robust
Minimax
Regret

ABSTRACT

We study the problem of estimating an unknown deterministic signal that is observed through an unknown deterministic data matrix under additive noise. In particular, we present a minimax optimization framework to the least squares problems, where the estimator has imperfect data matrix and output vector information. We define the performance of an estimator relative to the performance of the optimal least squares (LS) estimator tuned to the underlying unknown data matrix and output vector, which is defined as the regret of the estimator. We then introduce an efficient robust LS estimation approach that minimizes this regret for the worst possible data matrix and output vector, where we refrain from any structural assumptions on the data. We demonstrate that minimizing this worst-case regret can be cast as a semi-definite programming (SDP) problem. We then consider the regularized and structured LS problems and present novel robust estimation methods by demonstrating that these problems can also be cast as SDP problems. We illustrate the merits of the proposed algorithms with respect to the well-known alternatives in the literature through our simulations.

1. Introduction

In this paper, we investigate estimation of an unknown deterministic signal that is observed through a deterministic data matrix under additive noise, which models a wide range of problems in signal processing applications [1–14]. In this framework, the data matrix and the output vector are not exactly known, however, estimates for both of them as well as uncertainty bounds on the estimates are given [2,8,15–19]. Since the model parameters are not known exactly, the performances of the classical LS estimators may significantly degrade, especially when the perturbations on the data matrix and the output vector are relatively high [9,15,16,20–22]. Hence, robust estimation algorithms are needed to obtain a satisfactory performance under such perturbations. This generic framework models several real-life applications, which require estimation of a signal observed through a linear model [9,16]. As an example, this setup models realistic channel equalization scenarios, where the data matrix represents a communication channel and the data vector is the transmitted information. The channel is usually unknown, especially for wireless communications applications, and possibly can be time-varying. Hence, in practical applications, the communication channel is estimated, where this estimate is usually subject to distortions [9,16]. Under such possible perturbations, robust equalization methods can be used to obtain a more consistent and acceptable performance compared to the LS (or MMSE) equalizer. In this sense, this formulation is comprehensive and can be used in other applications such as in feedback control systems to estimate a desired data under imperfect system knowledge.

A prevalent approach to find robust solutions to such estimation problems is the robust minimax LS method [8,9,16,23–27], in which the uncertainties in the data matrix and the output vector are incorporated into optimization framework via a minimax residual formulation and a worst-case optimization within the uncertainty bounds is performed. Although the robust LS methods are able to minimize the LS error for the worst-case perturbations, they usually provide unsatisfactory results on the average [15,23–27] due to their conservative nature. This issue is significantly exacerbated especially when the actual perturbations do not result in significant performance degradation. Another well-known approach to compensate for errors in the data matrix and the output vector is the total least squares method (TLS) [15], which may yield undesirable results since it employs a conservative approach due to data de-regularization. On the other hand, the data matrix usually has a known special structure, such as Toeplitz and Hankel, in many linear regression problems [9,15]. Hence, in [9,15], the authors illustrate that the performances of the estimators based on minimax approaches improve when such a prior knowledge on data matrix structure is integrated into the problem formulation.

* Corresponding author.
E-mail addresses: vanli@ee.bilkent.edu.tr (N.D. Vanli), donmez2@illinois.edu (M.A. Donmez), kozat@ee.bilkent.edu.tr (S.S. Kozat).

http://dx.doi.org/10.1016/j.dsp.2014.10.004

In all these methods, LS estimators under worst case perturbations are introduced to achieve robustness. However, due to this conservative problem formulation, in many practical applications, these approaches yield unsatisfactory performances [2,8,18,28–30].

In order to counterbalance this conservative nature of the robust LS methods [9], we propose a novel robust LS approach that minimizes a worst case "regret" that is defined as the difference between the squared residual error and the smallest attainable squared residual error with an LS estimator [2,8,18,28–30]. By this regret formulation, we seek a linear estimator whose performance is as close as possible to that of the optimal estimator for all possible perturbations on the data matrix and the output vector. Our main goal in proposing the minimax regret formulation is to provide a trade-off between the robust LS methods tuned to the worst possible data parameters (under the uncertainty bounds) and the optimal LS estimator tuned to the underlying unknown model parameters. Minimax regret approaches have been presented in signal processing literature to alleviate the pessimistic nature of the worst case optimization methods [2,8,18,28–30]. In [18,29], linear minimax regret estimators are introduced to minimize the mean squared error (MSE) under imperfect knowledge of channel statistics and true parameters, respectively. In [28], a minimum mean squared error (MMSE) estimation technique under imperfect channel and data knowledge is investigated. In [2], these robust estimation methods are extended to flat fading channels to perform channel equalization. These methods are shown to provide a better average performance compared to the minimax estimators, whereas under large perturbations the robustness of the minimax estimators are superior to these competitive methods. On the other hand, in this paper, the optimization frameworks investigated here are significantly different than [9,16,23–27], where the regret terms are directly adjoined in the cost functions. In particular, unlike [2,18,28,29], where the uncertainties are in the statistics of the transmitted signal or channel parameters, in this paper, the uncertainty is both on the data matrix and the output vector without any statistical assumptions. While in [8], the authors have considered a similar framework, the results of this paper build upon them and provide a complete solution to the regret based robust LS estimation methods unlike [8]. We emphasize that perturbation bounds on the data matrix and the output vector heavily depend on the estimation algorithms employed to obtain them. Since our methods are formulated for given perturbation bounds, different estimation algorithms can be readily incorporated into our framework with the corresponding perturbation bounds [16].

Our main contributions in this paper are as follows. *i*) We introduce a novel and efficient robust LS estimation method in which we find the transmitted signal by minimizing the worst-case regret, i.e., the worst-case difference between the residual error of the LS estimator and the residual error of the optimal LS estimator tuned to the underlying model. In this sense, we present a robust estimation method that achieves a tradeoff between the robust LS estimation methods and the direct LS estimation method tuned to the estimates of the data matrix and output vector. *ii*) We next propose a minimax regret formulation for the regularized LS estimation problem. *iii*) We then introduce a structured robust LS estimation method in which the data matrix is known to have a special structure such as Toeplitz or Hankel. *iv*) We demonstrate that the robust estimation methods we propose can be cast as SDP problems, hence our methods can be efficiently implemented in real-time [31]. *v*) In our simulations, we observe that our approaches provide better performance compared to the robust methods that are optimized with respect to the worst-case residual error [9,32], and the conventional methods that directly solve the estimation problem using the perturbed data.

The organization of the paper is as follows. An overview to the problem is provided in Section 2. In Section 3.1, we first introduce the LS estimation method based on our regret formulation, and then present the regularized LS estimation approach in Section 3.2. We then consider the structured LS approach in Section 3.3 and provide the explicit SDP formulations for all problems. The numerical examples are demonstrated in Section 4. Finally, the paper concludes with certain remarks in Section 5.

2. System overview

2.1. Notation

In this paper, all vectors are column vectors and represented by boldface lowercase letters. Matrices are represented by boldface uppercase letters. For a matrix $\mathbf{H}$, $\mathbf{H}^H$ is the conjugate transpose, $\|\mathbf{H}\|$ is the spectral norm, $\mathbf{H}^+$ is the pseudo-inverse, $\mathbf{H} > 0$ represents a positive definite matrix and $\mathbf{H} \geq 0$ represents a positive semi-definite matrix. For a square matrix $\mathbf{H}$, $\mathrm{Tr}(\mathbf{H})$ is the trace. Naturally, for a vector $\mathbf{x}$, $\|\mathbf{x}\| = \sqrt{\mathbf{x}^H\mathbf{x}}$ is the ℓ^2-norm. Here, $\mathbf{0}$ denotes a vector or matrix with all zero elements and the dimensions can be understood from the context. Similarly, $\mathbf{I}$ represents the appropriate sized identity matrix. The operator $\mathrm{vec}(\cdot)$ is the vectorization operator, i.e., it stacks the columns of a matrix of dimension $m \times n$ into an $mn \times 1$ column vector. Finally, the operator $\otimes$ is the Kronecker product [33].

2.2. Problem description

We investigate the problem of estimating an unknown deterministic vector $\mathbf{x} \in \mathbb{C}^n$ which is observed through a deterministic data matrix. However, instead of the actual data matrix and the output vector, their estimates $\mathbf{H} \in \mathbb{C}^{m\times n}$ and $\mathbf{y} \in \mathbb{C}^m$ and uncertainty bounds on these estimates are provided. In this sense, our aim is to find a solution to the following data estimation problem

$$\mathbf{y} \approx \mathbf{H}\mathbf{x},$$

such that

$$\mathbf{y} + \Delta\mathbf{y} = (\mathbf{H} + \Delta\mathbf{H})\mathbf{x},$$

for deterministic perturbations $\Delta\mathbf{H} \in \mathbb{C}^{m\times n}$, $\Delta\mathbf{y} \in \mathbb{C}^m$. Although these perturbations are unknown, a bound on each perturbation is provided, i.e.,

$$\|\Delta\mathbf{H}\| \leq \delta_H \quad \text{and} \quad \|\Delta\mathbf{y}\| \leq \delta_Y,$$

where $\delta_H, \delta_Y \geq 0$. In this sense, we refrain from any assumptions on the data matrix and the output vector, yet consider that the estimates $\mathbf{H}$ and $\mathbf{y}$ are at least accurate to "some degree" but their actual values under these uncertainties are completely unknown to the estimator.

Even in the presence of these uncertainties, the symbol vector $\mathbf{x}$ can be naively estimated by simply substituting the estimates $\mathbf{H}$ and $\mathbf{y}$ into the LS estimator [10]. For the LS estimator we have

$$\hat{\mathbf{x}} = \mathbf{H}^+\mathbf{y},$$

where $\mathbf{H}^+$ is the pseudo-inverse of $\mathbf{H}$ [33]. However, this approach yields unsatisfactory results, when the errors in the estimates of the data matrix and the output vector are relatively high [9,18,29,32]. A common approach to find a robust solution is to employ a worst-case residual minimization [9]

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}\in\mathbb{C}^n} \max_{\|\Delta\mathbf{H}\|\leq\delta_H, \|\Delta\mathbf{y}\|\leq\delta_Y} \left\|(\mathbf{y} + \Delta\mathbf{y}) - (\mathbf{H} + \Delta\mathbf{H})\mathbf{x}\right\|^2,$$

where $\mathbf{x}$ is chosen to minimize the worst-case residual error in the uncertainty region. However, since the solution is found with