# MINIMAX REGRET ESTIMATION IN LINEAR MODELS

*Yonina C. Eldar, Aharon Ben-Tal and Arkadi Nemirovski*

Technion–Israel Institute of Technology
Haifa, 32000, Israel

**ABSTRACT**

We develop a new linear estimator for estimating an unknown vector $\mathbf{x}$ in a linear model, in the presence of bounded data uncertainties. The estimator is designed to minimize the worst-case *regret* across all bounded data vectors, namely the worst-case difference between the MSE attainable using a linear estimator that does not know the true parameters $\mathbf{x}$, and the optimal MSE attained using a linear estimator that knows $\mathbf{x}$. We demonstrate through several examples that the minimax regret estimator can significantly increase the performance over the conventional least-squares estimator, as well as several other least-squares alternatives.

## 1. INTRODUCTION

We consider the generic problem in which we seek to estimate an unknown deterministic parameter vector $\mathbf{x}$ in the linear model

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{w}, \tag{1}$$

where $\mathbf{H}$ is a known matrix and $\mathbf{w}$ is a noise vector. Estimation problems of this form arise in many different fields in science and engineering, and consequently have attracted much attention in the estimation literature.

Since the MSE of a linear estimator $\hat{\mathbf{x}}$ of $\mathbf{x}$ will in general depend on the unknown vector $\mathbf{x}$, we cannot design a linear estimator to minimize the MSE. Instead, a common approach is to seek linear estimators that minimize some function of the data error $\hat{\mathbf{y}} - \mathbf{y}$, where $\hat{\mathbf{y}} = \mathbf{H}\hat{\mathbf{x}}$ is the estimated data vector. The celebrated least-squares (LS) estimator seeks the estimator $\hat{\mathbf{x}}$ of $\mathbf{x}$ that minimizes the squared-norm of the data error $\|\hat{\mathbf{y}} - \mathbf{y}\|^2$. It is well known that the LS estimate has the smallest variance among all linear *unbiased* estimators. On the negative side, an unbiased estimator does not necessarily lead to a small MSE. In fact, in many cases the LS estimator can result in a large MSE. Various modifications of the LS estimator have been proposed, including Tikhonov regularization [1], the linear shrunken estimator [2], and the covariance shaping LS estimator [3]. In general, these LS alternatives attempt to reduce the MSE in estimating $\mathbf{x}$ by allowing for a bias. Each of the estimators above can be shown to be a solution to an optimization problem which involves minimizing some function that depends on the *data error*.

In an estimation context, we typically would like to minimize the *estimation error*, rather than the data error. To this end we assume that $\mathbf{x}$ is known to satisfy a (possibly weighted) norm constraint, and then seek a robust estimator whose performance is reasonably good across all possible choices of the parameters $\mathbf{x}$, in the region of uncertainty. The most common approach for designing robust estimators is the minimax MSE approach, in which we seek the estimator that minimizes the worst-case MSE in the region of uncertainty. The minimax approach, in which the goal is to optimize the worst-case performance, is one of the major techniques for designing robust systems with respect to modelling uncertainties, and has been applied to many problems in detection and estimation [4, 5].

Following the popular minimax approach, we may seek the linear estimator that minimizes the worst-case MSE over all possible values of $\mathbf{x}$ that satisfy a weighted norm constraint. The minimax estimator of this form for arbitrary matrices $\mathbf{H}$ and arbitrary noise vectors $\mathbf{w}$, is developed in [6], in which the case of uncertainties in the model matrix $\mathbf{H}$ is also considered.

Although the minimax approach has enjoyed widespread use in the design of robust methods for signal processing and communication, its performance is often unsatisfactory. The main limitation of this approach is that it tends to be overly conservative since it optimizes the performance for the worst possible choice of unknowns. As we show in the context of a concrete example in Section 4, this can often lead to degraded performance.

To improve the performance of the minimax MSE estimator, we propose, in Section 2, a new approach to linear estimation, in which we seek a linear estimator whose performance is as close as possible to that of the optimal linear estimator, *i.e.,* the one minimizing the MSE when $\mathbf{x}$ is assumed to be known. Specifically, we seek the estimator that minimizes the worst-case *regret*, which is the difference between the MSE of the linear estimator which does not know $\mathbf{x}$, and the smallest attainable MSE with a linear estimator that knows $\mathbf{x}$. Note that as we show in Section 2, since we are restricting ourselves to linear estimators, we cannot achieve zero MSE even in the case in which the parameters $\mathbf{x}$ are known. By considering the *difference* between the MSE and the optimal MSE rather than the MSE directly, we can counterbalance the conservative character of the minimax approach, as is evident in the examples we consider in Section 4. Some special cases of the minimax regret estimator for different choices of the weighting matrix are considered in Section 3.

Proofs of theorems, which are omitted here for brevity, can be found in [8].

## 2. MINIMAX REGRET ESTIMATOR

We denote vectors in $\mathbb{C}^m$ by boldface lowercase letters and matrices in $\mathbb{C}^{n\times m}$ by boldface uppercase letters. The Hermitian conjugate of a matrix is denoted by $(\cdot)^*$.

Consider the problem of estimating the unknown deterministic parameter vector $\mathbf{x}$ in the linear model (1), where $\mathbf{H}$ is a known $n \times m$ matrix with full rank $m$, and $\mathbf{w}$ is a zero-mean random vector with covariance $\mathbf{C}_w$. We assume that $\mathbf{x}$ is known to satisfy the weighted norm constraint $\|\mathbf{x}\|_{\mathbf{T}} \leq L$ for some positive definite covariance $\mathbf{T}$ and scalar $L > 0$, where $\|\mathbf{x}\|^2_{\mathbf{T}} = \mathbf{x}^*\mathbf{T}\mathbf{x}$.

We estimate $\mathbf{x}$ using a linear estimator so that $\hat{\mathbf{x}} = \mathbf{G}\mathbf{y}$ for some $m \times n$ matrix $\mathbf{G}$. The MSE of $\hat{\mathbf{x}}$ is given by

$$E(\|\hat{\mathbf{x}} - \mathbf{x}\|^2) = \text{Tr}(\mathbf{G}\mathbf{C}_w\mathbf{G}^*) + \mathbf{x}^*(\mathbf{I} - \mathbf{G}\mathbf{H})^*(\mathbf{I} - \mathbf{G}\mathbf{H})\mathbf{x}. \quad (2)$$

Since the MSE depends on the unknown parameter vector $\mathbf{x}$, we cannot choose an estimate to directly minimize the MSE (2).

Instead, we seek the linear estimator that minimizes the worst-case regret $\mathcal{R}(\mathbf{x}, \mathbf{G})$, which is the difference between the MSE of an estimator $\hat{\mathbf{x}} = \mathbf{G}\mathbf{y}$ of $\mathbf{x}$, and the best possible MSE attainable using any estimator of the form $\hat{\mathbf{x}} = \mathbf{G}(\mathbf{x})\mathbf{y}$ where $\mathbf{x}$ is assumed to be known, so that $\mathbf{G}$ can depend explicitly on $\mathbf{x}$. As we now show, since we are restricting ourselves to linear estimators of the form $\hat{\mathbf{x}} = \mathbf{G}\mathbf{y}$, even in the case in which the parameters $\mathbf{x}$ are known we cannot achieve zero MSE. The best possible MSE is illustrated schematically in Fig. 1. Instead of seeking an estimator to minimize the worst-case MSE, we therefore propose seeking an estimator to minimize the worst-case difference between its MSE and the best possible MSE, as illustrated in Fig. 1.

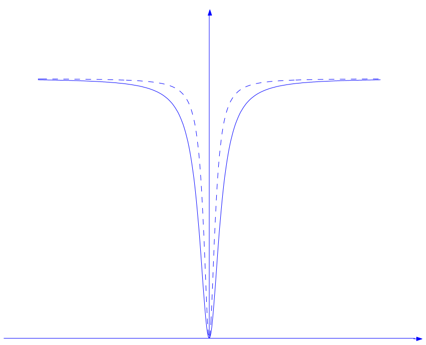

**Fig. 1**. The line represents the best attainable MSE as a function of $\mathbf{x}$ when $\mathbf{x}$ is known, and the dashed line represents a desirable graph of MSE with small regret as a function of $\mathbf{x}$, using some linear estimator that does not depend on $\mathbf{x}$.

To develop the minimax regret estimator, we first determine the best attainable MSE, which we denote by $\text{MSE}^o$. To this end, we seek the estimator $\hat{\mathbf{x}} = \mathbf{G}(\mathbf{x})\mathbf{y}$ that minimizes the MSE when $\mathbf{x}$ is known. Differentiating the MSE of (2) with respect to $\mathbf{G}$ and equating to 0, the optimal $\mathbf{G}(\mathbf{x})$ is

$$\begin{aligned} \mathbf{G}(\mathbf{x}) &= \mathbf{x}\mathbf{x}^*\mathbf{H}^*(\mathbf{C}_w + \mathbf{H}\mathbf{x}\mathbf{x}^*\mathbf{H}^*)^{-1} \\ &= \frac{1}{1 + \mathbf{x}^*\mathbf{H}^*\mathbf{C}_w^{-1}\mathbf{H}\mathbf{x}}\mathbf{x}\mathbf{x}^*\mathbf{H}^*\mathbf{C}_w^{-1}, \end{aligned} \quad (3)$$

where we used the Matrix Inversion Lemma [7]. Substituting $\mathbf{G}(\mathbf{x})$ back into (2), $\text{MSE}^o$ is given by

$$\text{MSE}^o = \frac{\mathbf{x}^*\mathbf{x}}{1 + \mathbf{x}^*\mathbf{H}^*\mathbf{C}_w^{-1}\mathbf{H}\mathbf{x}}. \quad (4)$$

Since $\mathbf{x}$ is unknown, we cannot implement the optimal estimator (3). Instead we seek the estimator $\hat{\mathbf{x}} = \mathbf{G}\mathbf{y}$ that minimizes the worst-case regret $\mathcal{R}(\mathbf{x}, \mathbf{G})$ subject to the constraint $\|\mathbf{x}\|_{\mathbf{T}} \leq L$. Thus, we seek the matrix $\mathbf{G}$ that is the solution to the problem

$$\min_{\mathbf{G}} \max_{\mathbf{x}^*\mathbf{T}\mathbf{x} \leq L^2} \mathcal{R}(\mathbf{x}, \mathbf{G}), \quad (5)$$

where

$$\begin{aligned} \mathcal{R}(\mathbf{x}, \mathbf{G}) &= E(\|\mathbf{G}\mathbf{y} - \mathbf{x}\|^2) - \text{MSE}^o = \\ &= \text{Tr}(\mathbf{G}\mathbf{C}_w\mathbf{G}^*) + \mathbf{x}^*(\mathbf{I} - \mathbf{G}\mathbf{H})^*(\mathbf{I} - \mathbf{G}\mathbf{H})\mathbf{x} - \\ &\quad - \frac{\mathbf{x}^*\mathbf{x}}{1 + \mathbf{x}^*\mathbf{H}^*\mathbf{C}_w^{-1}\mathbf{H}\mathbf{x}}. \end{aligned} \quad (6)$$

For analytical tractability, we restrict our attention to weighting matrices $\mathbf{T}$ such that $\mathbf{T}$ and $\mathbf{H}^*\mathbf{C}_w^{-1}\mathbf{H}$ have the same eigenvector matrix. Thus, if $\mathbf{H}^*\mathbf{C}_w^{-1}\mathbf{H}$ has an eigendecomposition $\mathbf{H}^*\mathbf{C}_w^{-1}\mathbf{H} = \mathbf{V}\Sigma\mathbf{V}^*$ where $\mathbf{V}$ is a unitary matrix and $\Sigma$ is a diagonal matrix, then $\mathbf{T} = \mathbf{V}\Lambda\mathbf{V}^*$ for some diagonal matrix $\Lambda$. Under this assumption, the form of the minimax regret estimator is given by the following theorem [8].

**Theorem 1** *Let $\mathbf{x}$ denote the unknown deterministic parameter vector in the model $\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{w}$, where $\mathbf{H}$ is a known $n \times m$ matrix with rank $m$, and $\mathbf{w}$ is a zero-mean random vector with covariance $\mathbf{C}_w$. Let $\mathbf{H}^*\mathbf{C}_w^{-1}\mathbf{H} = \mathbf{V}\Sigma\mathbf{V}^*$ where $\mathbf{V}$ is a unitary matrix and $\Sigma$ is an $m \times m$ diagonal matrix with diagonal elements $\sigma_i > 0$ and let $\mathbf{T} = \mathbf{V}\Lambda\mathbf{V}^*$ where $\Lambda$ is an $m \times m$ diagonal matrix with diagonal elements $\lambda_i > 0$. Then the solution to the problem*

$$\min_{\hat{\mathbf{x}} = \mathbf{G}\mathbf{y}} \max_{\|\mathbf{x}\|_{\mathbf{T}} \leq L} \left\{ E(\|\hat{\mathbf{x}} - \mathbf{x}\|^2) - \min_{\hat{\mathbf{x}} = \mathbf{G}(\mathbf{x})\mathbf{y}} E(\|\hat{\mathbf{x}} - \mathbf{x}\|^2) \right\}$$

*has the form*

$$\hat{\mathbf{x}} = \mathbf{V}\mathbf{D}\mathbf{V}^*(\mathbf{H}^*\mathbf{C}_w^{-1}\mathbf{H})^{-1}\mathbf{H}^*\mathbf{C}_w^{-1}\mathbf{y},$$

*where $\mathbf{D}$ is an $m \times m$ diagonal matrix with diagonal elements $d_i$ which are the solution to the convex optimization problem*

$$\min_{\tau, d_i} \tau$$

*subject to*

$$\begin{aligned} \textstyle\sum_{i=1}^m \frac{d_i^2}{\sigma_i} &\leq \tau \\ \Phi(\{d_i\}) + \textstyle\sum_{i=1}^m \frac{d_i^2}{\sigma_i} &\leq \tau \end{aligned}$$

*where*

$$\Phi(\{d_i\}) = \max_{s_i \geq 0, \sum_i \lambda_i s_i = L^2} \left\{ \sum_{i=1}^m (1 - d_i)^2 s_i - \frac{\sum_{i=1}^m s_i}{1 + \sum_{i=1}^m \sigma_i s_i} \right\}.$$

Theorem 1 reduces the problem of minimizing the regret to a simpler convex optimization problem in $m$ unknowns. In the next section we show that for certain choices of the weighting matrix $\mathbf{T}$, the problem can be further simplified, and in some cases a closed form solution for the minimax regret estimator exists.

## 3. MINIMAX REGRET ESTIMATOR FOR SPECIAL CHOICES OF T

We now consider two choices of $\mathbf{T}$ which simplify the optimization problem of Theorem 1.

Suppose first that $\mathbf{T} = \mathbf{H}^*\mathbf{C}_w^{-1}\mathbf{H}$. This case may be of interest, for example, when an unknown signal $\mathbf{x}$ is sent through a known channel, and the output signal-to-noise ratio is bounded. As we show in Theorem 2 below, when $L$ is large enough with respect to $m$, the optimal minimax regret estimator in this case is a shrunken estimator proposed by Mayer and Willke [2], which is simply a scaled version of the LS estimator, with a specific choice of shrinkage factor. For small values of $L$, the optimal estimator is given in terms of a single parameter, which is the solution to a nonlinear equation.

**Theorem 2** *Let* $\mathbf{x}$ *denote the unknown deterministic parameter vector in the model* $\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{w}$*, where* $\mathbf{H}$ *is a known* $n \times m$ *matrix with rank* $m$*, and* $\mathbf{w}$ *is a zero-mean random vector with covariance* $\mathbf{C}_w$*. Let* $\mathbf{H}^*\mathbf{C}_w^{-1}\mathbf{H} = \mathbf{V}\Sigma\mathbf{V}^*$ *where* $\mathbf{V}$ *is a unitary matrix and* $\Sigma$ *is an* $m \times m$ *diagonal matrix with diagonal elements* $\sigma_1 \geq \ldots \geq \sigma_m > 0$*. Then the minimax regret estimator with* $\mathbf{T} = \mathbf{H}^*\mathbf{C}_w^{-1}\mathbf{H}$ *is given by*

$$\hat{\mathbf{x}} = \begin{cases} \left(1 - \sqrt{\frac{1}{1+L^2}}\right)(\mathbf{H}^*\mathbf{C}_w^{-1}\mathbf{H})^{-1}\mathbf{H}^*\mathbf{C}_w^{-1}\mathbf{y}, & L^2 \geq \alpha; \\ \mathbf{V}\mathbf{D}\mathbf{V}^*(\mathbf{H}^*\mathbf{C}_w^{-1}\mathbf{H})^{-1}\mathbf{H}^*\mathbf{C}_w^{-1}\mathbf{y}, & L^2 < \alpha, \end{cases}$$

*where* $\alpha = (m-1)^2 - 1$*, and* $\mathbf{D}$ *is an* $m \times m$ *diagonal matrix with diagonal elements* $d_i$ *that are given by*

$$d_i = \begin{cases} 0, & i \leq k(\rho_*); \\ 1 - \sqrt{\rho_*\sigma_i + 1/(L^2+1)}, & i \geq k(\rho_*) + 1. \end{cases}$$

*Here* $\rho_*$ *is the unique positive root of*

$$\phi(\rho) = \sum_{i=1}^{m} \sigma_i \gamma_i(\rho) - L^2,$$

*where* $\gamma_i(\rho)$ *is defined by*

$$\gamma_i(\rho) = \max\left[\frac{1}{\sigma_i}\left(\frac{1}{\sqrt{\rho\sigma_i + 1/(L^2+1)}} - 1\right), 0\right],$$

*and*

$$k(\rho_*) = \begin{cases} 0, & \gamma_1(\rho_*) > 0; \\ \max\{i : \gamma_i(\rho_*) = 0\}, & \textit{otherwise}. \end{cases}$$

As we expect intuitively, when $L \to \infty$, the minimax regret estimator $\hat{\mathbf{x}}$ of Theorem 2 reduces to the LS estimator. Indeed, when the weighted norm of $\mathbf{x}$ can be made arbitrarily large, the MSE, and therefore the regret, will also be arbitrarily large unless the bias is equal to zero. Therefore, in this limit, the worst-case regret is minimized by choosing an estimator with zero bias that minimizes the variance, which leads to the LS estimator.

We next consider the case in which $\mathbf{T} = \mathbf{I}$, which may be of interest when an unknown signal $\mathbf{x}$ is sent through a known channel, and the power of $\mathbf{x}$ is bounded. Theorem 3 below shows that in this case the minimax regret estimator can be determined by solving $m$ convex optimization problems, each in 3 unknowns.

**Theorem 3** *Let* $\mathbf{x}$ *denote the unknown deterministic parameter vector in the model* $\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{w}$*, where* $\mathbf{H}$ *is a known* $n \times m$ *matrix with rank* $m$*, and* $\mathbf{w}$ *is a zero-mean random vector with covariance* $\mathbf{C}_w$*. Let* $\mathbf{H}^*\mathbf{C}_w^{-1}\mathbf{H} = \mathbf{V}\Sigma\mathbf{V}^*$ *where* $\mathbf{V}$ *is a unitary matrix and* $\Sigma$ *is an* $m \times m$ *diagonal matrix with diagonal elements* $\sigma_1 \geq \ldots \geq \sigma_m > 0$*. Then the minimax regret estimator with* $\mathbf{T} = \mathbf{I}$ *has the form*

$$\hat{\mathbf{x}} = \mathbf{V}\mathbf{D}\mathbf{V}^*(\mathbf{H}^*\mathbf{C}_w^{-1}\mathbf{H})^{-1}\mathbf{H}^*\mathbf{C}_w^{-1}\mathbf{y},$$

*where* $\mathbf{D}$ *is an* $m \times m$ *diagonal matrix with diagonal elements* $d_i$ *that are given by*

$$d_i = \begin{cases} 1 - \sqrt{\lambda - \sigma_i\mu}, & i \leq k; \\ 0, & i \geq k+1, \end{cases}$$

*with* $k = \arg\min \tau_\ell$*,* $\mu = \mu_k$ *and* $\lambda = \lambda_k$*. Here* $\tau_\ell$*,* $\mu_\ell$ *and* $\lambda_\ell$ *for* $1 \leq \ell \leq m$ *are the optimal solutions to*

$$\min_{\tau,\mu,\lambda} \tau$$

*subject to*

$$\begin{array}{c} \sum_{i=1}^{\ell} \frac{\left(1-\sqrt{\lambda-\sigma_i\mu}\right)^2}{\sigma_i} \leq \tau \\ -2L\sqrt{\mu} + \mu + L^2\lambda + \sum_{i=1}^{\ell} \frac{\left(1-\sqrt{\lambda-\sigma_i\mu}\right)^2}{\sigma_i} \leq \tau \\ \sigma_1\mu \leq \lambda \leq 1 + \sigma_\ell\mu \\ \mu \geq 0 \\ \textit{if } \ell < m \textit{ then } \lambda \geq 1 + \sigma_{\ell+1}\mu. \end{array}$$

As before, we can readily show that when $L \to \infty$, the minimax regret estimator of Theorem 3 reduces to the LS estimator.

| $(k_{\ell,1}, k_{\ell,2})$ | $a_\ell$ | $\phi_\ell$ |
|---|---|---|
| (1,1) | 1.0681 | 2.1438 |
| (2,1) | 0.8704 | 3.3557 |
| (1,2) | 1.2027 | 4.5686 |
| (2,2) | 1.0466 | 1.9433 |
| (3,2) | 0.9449 | 5.2684 |

**Table 1**. Simulation parameters.

## 4. EXAMPLE

We now present an example, illustrating the performance advantage of the minimax regret estimator.

We consider the problem of estimating a 2D image from noisy observations, which are obtained by blurring the image with a 2D filter, and adding random Gaussian noise. The image $x(z_1, z_2)$ is chosen as a sum of $m$ harmonic oscillations:

$$x(z_1, z_2) = \sum_{\ell=1}^{m} a_\ell \cos(\omega_{\ell,1}z_1 + \omega_{\ell,2}z_2 + \phi_\ell), \tag{7}$$

where $\omega_{\ell,i} = 2\pi k_{\ell,i}/n$, and $k_{\ell,i} \in \mathbb{Z}^2$ are given parameters. Clearly, the image $x(z_1, z_2)$ is periodic with period $n$. Therefore, we can represent the image by a length-$n^2$ vector $\mathbf{x}$, with components $\{x(z_1, z_2) : 0 \leq z_1, z_2 \leq n-1\}$.

The observed image $y(z_1, z_2)$ is given by

$$y(z_1, z_2) = \sum_{\tau_1,\tau_2} H(\tau_1, \tau_2)x(z_1-\tau_1-d_1, z_2-\tau_2-d_2) + \sigma w(z_1, z_2),$$

where $H(z_1, z_2)$ is a blurring filter defined by

$$H(z_1, z_2) = \max\left(1 - \frac{\sqrt{z_1^2 + z_2^2}}{\rho}, 0\right), \tag{8}$$

for some parameter $\rho$, $d_1$ and $d_2$ are randomly chosen shifts, and $w(z_1, z_2)$ is an independent, zero-mean, Gaussian noise process so that for each $z_1$ and $z_2$, $w(z_1, z_2)$ is $\mathcal{N}(0, 1)$. By defining the vectors $\mathbf{y}$ and $\mathbf{w}$ with components $y(z_1, z_2)$ and $w(z_1, z_2)$, respectively, and defining a matrix $\mathbf{H}$ with the appropriate elements $H(z_1, z_2)$, the observations $\mathbf{y}$ can be expressed in the form (1).

In Fig. 2 we consider the case in which $m = 5$, $n = 128$, $\sigma = 0.5$, $L = \|\mathbf{x}\|$, and $n\rho = \sqrt{2}$. The values of $k_{\ell,i}$, $a_\ell$ and $\phi_\ell$ are given in Table 1. To estimate the image $x(z_1, z_2)$ from the noisy observations $y(z_1, z_2)$ we consider 4 different estimators: The LS estimator, the minimax regret (RGR) estimator of Theorem 3, and

| Estimator | Relative Error |
|---|---|
| LS | 5.0e8 |
| MMX | 1.00 |
| WNR | 6.17 |
| RGR | 0.843 |

**Table 2**. Relative error for the data of Table 1.

two other estimators, the deterministic Wiener estimator (WNR), and the minimax estimator (MMX), which we now describe.

The LS estimator does not incorporate the knowledge on $\sigma$ and $L = \|\mathbf{x}\|$. To develop an estimator that incorporates this knowledge, we may assume that $\mathbf{x}$ is a random vector with covariance $L^2\mathbf{I}$ independent of $\mathbf{w}$, and design a minimum MSE Wiener estimator matched to this covariance. The resulting estimator is

$$\hat{\mathbf{x}} = \left(\mathbf{H}^*\mathbf{H} + \frac{\sigma^2}{L^2}\mathbf{I}\right)^{-1}\mathbf{H}^*\mathbf{y}. \tag{9}$$

The minimax estimator minimizes the worst-case MSE over all $\mathbf{x}^*\mathbf{x} \le L^2$, and is given by [6]

$$\hat{\mathbf{x}} = \frac{L^2}{L^2+\gamma_0}(\mathbf{H}^*\mathbf{C}_w^{-1}\mathbf{H}^*)^{-1}\mathbf{H}^*\mathbf{C}_w^{-1}\mathbf{y}, \tag{10}$$

where $\gamma_0 = \mathrm{Tr}\left(\left(\mathbf{H}^*\mathbf{C}_w^{-1}\mathbf{H}\right)^{-1}\right)$.

In Table 2 we report the relative error $\epsilon = \|\hat{\mathbf{x}} - \mathbf{x}\|/\|\mathbf{x}\|$ corresponding to the $4$ estimators. The performance of the LS, minimax, and Wiener estimators are severely degraded. The surprising result is that even though in this example $\mathbf{H}$ is ill-conditioned, the minimax regret estimator works pretty well, as can be seen from the results of Table 2, as well as in Fig. 2. Since the error in the LS estimate is so large, we do not show the resulting image.

## 5. REFERENCES


[1] A. N. Tikhonov and V. Y. Arsenin, *Solution of Ill-Posed Problems*, Washington, DC: V.H. Winston, 1977.

[2] L. S. Mayer and T. A. Willke, "On biased estimation in linear models," *Technometrics*, vol. 15, pp. 497–508, Aug. 1973.

[3] Y. C. Eldar and A. V. Oppenheim, "Covariance shaping least-squares estimation," *IEEE Trans. Signal Processing*, vol. 51, pp. 686–697, Mar. 2003.

[4] S. A. Kassam and H. V. Poor, "Robust techniques for signal processing: A survey," *IEEE Proc.*, vol. 73, pp. 433–481, Mar. 1985.

[5] S. Verdú and H. V. Poor, "On minimax robustness: A general approach and applications," *IEEE Trans. Inform. Theory*, vol. IT-30, pp. 328–340, Mar. 1984.

[6] Y. C. Eldar, A. Ben-Tal, and A. Nemirovski, "Robust mean-squared error estimation in the presence of model uncertainties," *IEEE Trans. Signal Processing*, to appear.

[7] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge, UK: Cambridge Univ. Press, 1985.

[8] Y. C. Eldar, A. Ben-Tal, and A. Nemirovski, "Linear minimax regret estimation with bounded data uncertainties," Tech. Rep. CCIT 437, Technion—Israel Institute of Technology, Haifa, Israel, July 2003, *IEEE Trans. Signal Processing*, to appear.


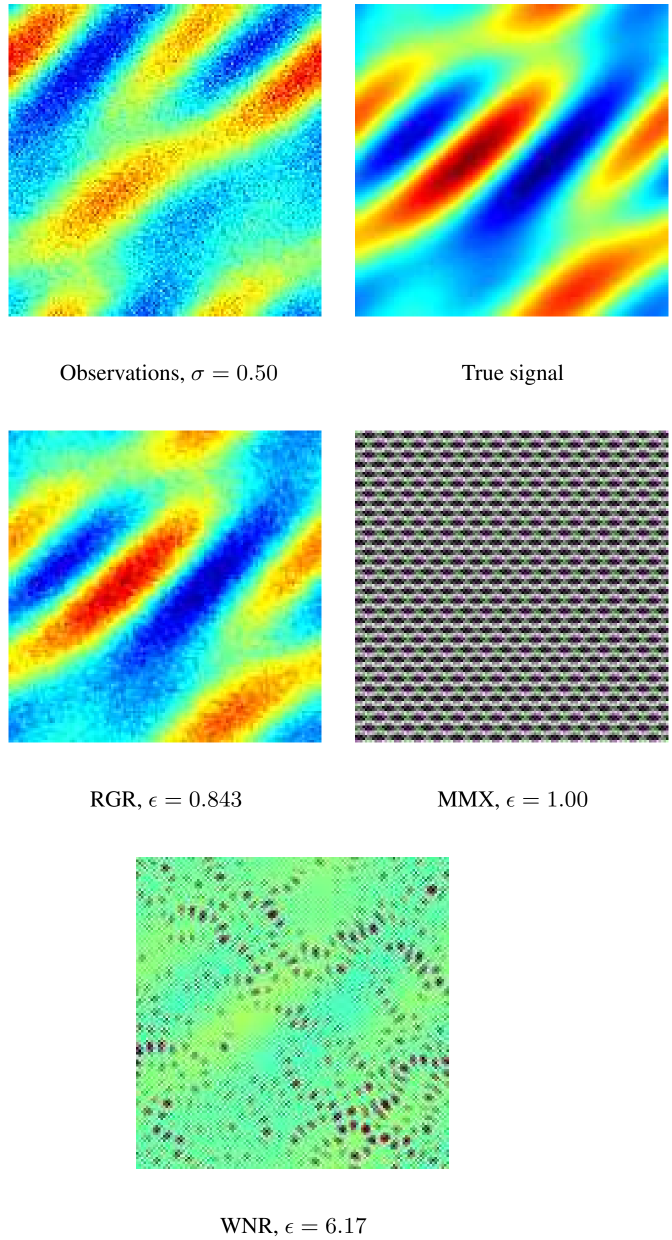


**Fig. 2**. Comparison of the minimax regret estimator of Theorem 3, and the least-squares, Wiener, and minimax estimators.