MMSE ESTIMATION IN A LINEAR SIGNAL MODEL WITH ELLIPSOIDAL CONSTRAINTS

Stefan Uhlich and Bin Yang

Chair of System Theory and Signal Processing
Universität Stuttgart, Germany

ABSTRACT

The estimation of an unknown parameter vector in a Gaussian linear model is studied in this paper. Two different cases are analyzed: the parameter vector is assumed to lie either in or on a given ellipsoid. The best estimator in terms of the mean squared error is derived. The performance of this estimator is analyzed and compared with the ordinary least squares, the constrained least squares and the linear minimax approach.

Index Terms— Parameter estimation, Minimum mean squared error estimation, Restricted parameter space

1. INTRODUCTION

In this paper we consider the problem of estimating an unknown parameter vector $\underline{\theta} \in \mathbb{R}^M$ from noisy measurements where it is a priori known that the parameter vector is restricted to lie in a subset of $\mathbb{R}^M$. The two subset cases that are investigated are: the set of all points inside an ellipsoid and the set of all points on the surface of an ellipsoid. The first case has attracted a lot of interest in the past, see e.g. [1–5] where the linear minimax estimator is sought which minimizes the worst case mean squared error. It is well known that the ordinary least squares estimator is outperformed by this estimator. The second case is interesting for DOA estimation problems, see e.g. [6] for an example where a parameter vector has to be estimated that lies on a 3-D sphere which is a special case of the problem we consider.

The contribution of this paper is twofold: First, we derive in Sec. 3 the minimum mean squared error (MMSE) estimator for an unknown parameter vector in a linear model with ellipsoidal constraints. We show that the two M-dimensional integrations can be replaced by M one-dimensional integrations. Second, we analyze in Sec. 5 its performance with respect to different error measures and compare it with the ordinary least squares approach, the linear minimax estimator from [2, 7], the constrained least squares estimator and modifications of them. The simulation results show that the MMSE estimator outperforms the other approaches at the expense of an increased computational complexity.

Following notations are used throughout this paper: $\underline{x}$ denotes a vector, $\mathbf{X}$ a matrix, and $\mathbf{I}$ the identity matrix. Furthermore, the error function $\text{erf}(x)$ is defined as $\text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt$.

2. SIGNAL MODEL

We consider the linear signal model

$$\underline{x} = \mathbf{H}\underline{\theta} + \underline{z} \tag{1}$$

where $\underline{x} \in \mathbb{R}^N$ are observations of the unknown parameter vector $\underline{\theta} \in \mathbb{R}^M$ with $N \geq M$. $\mathbf{H} \in \mathbb{R}^{N \times M}$ is the known model matrix with a full column rank M and $\underline{z} \in \mathbb{R}^N$ the observation noise

which is Gaussian with $\underline{z} \sim \mathcal{N}(\underline{0}, \mathbf{C})$. It is a priori known that the parameter vector $\underline{\theta}$ satisfies one of the following two constraints:

Case A: $\underline{\theta}$ lies inside an ellipsoid $\mathcal{R}^*$, i.e. $\underline{\theta} \in \mathcal{R}^*$. The ellipsoid is given by

$$\mathcal{R}^* = \{\underline{\theta} : (\underline{\theta} - \underline{\theta}_0)^T \mathbf{\Theta} (\underline{\theta} - \underline{\theta}_0) \leq R^2\} \tag{2a}$$

or

Case B: $\underline{\theta}$ lies on an ellipsoid $\bar{\mathcal{R}}^*$, i.e. $\underline{\theta} \in \bar{\mathcal{R}}^*$, with

$$\bar{\mathcal{R}}^* = \{\underline{\theta} : (\underline{\theta} - \underline{\theta}_0)^T \mathbf{\Theta} (\underline{\theta} - \underline{\theta}_0) = R^2\}. \tag{2b}$$

The problem of finding a suitable estimator for $\underline{\theta}$ is a well known problem in signal processing. In [1] the linear minimax estimator for $\mathbf{H} = \mathbf{I}$ and $\mathbf{C} = \sigma^2 \mathbf{I}$ is derived if $\underline{\theta}$ is restricted to lie in an ellipsoid. This result was later extended in [2, 7].

In this paper, we will derive the MMSE estimator for this problem by modeling $\underline{\theta}$ as uniformly distributed in $\mathcal{R}^*$ or on $\bar{\mathcal{R}}^*$, respectively. This is the fundamental difference of our approach to the previous works which always assume $\underline{\theta}$ to be deterministic. The uniform distribution of the MMSE estimator has the interpretation that we only know that $\underline{\theta}$ is restricted to lie in $\mathcal{R}^*$ or on $\bar{\mathcal{R}}^*$, but no other a priori information is available. Thus, our results are directly comparable to those which were derived in a deterministic framework.

3. CASE A: MMSE ESTIMATOR FOR $\underline{\theta} \in \mathcal{R}^*$

We now derive the estimator for $\underline{\theta}$ in (1) which has the minimum mean squared error among all estimators. It is well known that the MMSE estimator is the mean of the a posteriori distribution [8]:

$$\hat{\underline{\theta}} = \text{E}[\underline{\theta}|\underline{x}] = \frac{\int_{\mathbb{R}^M} \underline{\theta}\, p_{x|\theta}(\underline{x}|\underline{\theta}) p_\theta(\underline{\theta}) d\underline{\theta}}{\int_{\mathbb{R}^M} p_{x|\theta}(\underline{x}|\underline{\theta}) p_\theta(\underline{\theta}) d\underline{\theta}}. \tag{3}$$

It requires in general two M-dimensional integrations which can often not be solved analytically. In our case, however, we can simplify these to M one-dimensional numerical integrations which are feasible. The derivation of the estimator is divided into three steps:

A. Transform the problem to one of estimating a vector in colored noise. The ellipsoid in which $\underline{\theta}$ lies is arbitrarily oriented.

B. Transform the problem to one of estimating a vector in white noise where the axes of the ellipsoid are parallel to the coordinate axes.

C. Give formulas for the remaining integrals.

Step A. As we assume $\underline{\theta} \in \mathcal{R}^*$ and $\underline{\theta}$ is a priori uniformly distributed, it has the probability density function (pdf)

$$p_\theta(\underline{\theta}) = \begin{cases} \text{const} & \underline{\theta} \in \mathcal{R}^* \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

Furthermore, $\underline{z} \sim \mathcal{N}(\underline{0}, \mathbf{C})$ and therefore we can write (3) as

$$\hat{\underline{\theta}} = \frac{\int_{\mathcal{R}^*} \underline{\theta} \exp\{-\frac{1}{2}(\underline{x} - \mathbf{H}\underline{\theta})^T \mathbf{C}^{-1}(\underline{x} - \mathbf{H}\underline{\theta})\} d\underline{\theta}}{\int_{\mathcal{R}^*} \exp\{-\frac{1}{2}(\underline{x} - \mathbf{H}\underline{\theta})^T \mathbf{C}^{-1}(\underline{x} - \mathbf{H}\underline{\theta})\} d\underline{\theta}}. \qquad (5)$$

Using the identity

$$(\underline{x} - \mathbf{H}\underline{\theta})^T \mathbf{C}^{-1}(\underline{x} - \mathbf{H}\underline{\theta}) = (\underline{\theta} - \underline{\tilde{x}})^T \tilde{\mathbf{C}}^{-1}(\underline{\theta} - \underline{\tilde{x}}) + \underline{x}^T \mathbf{C}^{-1}(\underline{x} - \mathbf{H}\underline{\tilde{x}}) \qquad (6)$$

with $\tilde{\mathbf{C}} = (\mathbf{H}^T \mathbf{C}^{-1} \mathbf{H})^{-1}$ and $\underline{\tilde{x}} = (\mathbf{H}^T \mathbf{C}^{-1} \mathbf{H})^{-1} \mathbf{H}^T \mathbf{C}^{-1} \underline{x}$, we can further write

$$\hat{\underline{\theta}} = \frac{\int_{\mathcal{R}^*} \underline{\theta} \exp\{-\frac{1}{2}(\underline{\theta} - \underline{\tilde{x}})^T \tilde{\mathbf{C}}^{-1}(\underline{\theta} - \underline{\tilde{x}})\} d\underline{\theta}}{\int_{\mathcal{R}^*} \exp\{-\frac{1}{2}(\underline{\theta} - \underline{\tilde{x}})^T \tilde{\mathbf{C}}^{-1}(\underline{\theta} - \underline{\tilde{x}})\} d\underline{\theta}}. \qquad (7)$$

The term $\underline{x}^T \mathbf{C}^{-1}(\underline{x} - \mathbf{H}\underline{\tilde{x}})$ in (6) is independent of $\underline{\theta}$ and is cancelled. Thus, the problem in (1) can be transformed into a simpler one by left-multiplying (1) with $(\mathbf{H}^T \mathbf{C}^{-1} \mathbf{H})^{-1} \mathbf{H}^T \mathbf{C}^{-1}$ which results in the signal model $\underline{\tilde{x}} = \underline{\theta} + \underline{\tilde{z}}$. The problem is now to estimate $\underline{\theta}$ which is observed in colored noise $\underline{\tilde{z}} \sim \mathcal{N}(\underline{0}, \tilde{\mathbf{C}})$. Note, that $\underline{\tilde{x}}$ represents the ordinary least squares estimate for $\underline{\theta}$. The MMSE estimator improves this estimate further by exploiting the restriction $\underline{\theta} \in \mathcal{R}^*$ and therefore shows a better performance. Thus, we can conclude that the projection into the column space of $\mathbf{H}$ as done by the least squares estimate $\underline{\tilde{x}}$ does not neglect information which we need for the MMSE estimator.

Step B. We define $\mathbf{U}$ as one matrix square root of $\tilde{\mathbf{C}}$, i.e. $\tilde{\mathbf{C}} = \mathbf{U}^T \mathbf{U}$. One possibility is the Cholesky factor of $\tilde{\mathbf{C}}$. Furthermore, let $\mathbf{V}$ and $\mathbf{D} = \mathrm{diag}(d_1, \ldots, d_M)$ be the matrix of eigenvectors and eigenvalues of $\mathbf{U}\mathbf{\Theta}\mathbf{U}^T$. Using the substitution $\underline{\theta} - \underline{\tilde{x}} = \mathbf{U}^T \mathbf{V} \underline{y}$, (7) reads

$$\hat{\underline{\theta}} = \frac{\int_{\mathcal{R}^{**}} (\underline{\tilde{x}} + \mathbf{U}^T \mathbf{V} \underline{y}) \exp\{-\frac{1}{2}\|\underline{y}\|^2\} d\underline{y}}{\int_{\mathcal{R}^{**}} \exp\{-\frac{1}{2}\|\underline{y}\|^2\} d\underline{y}}$$

$$= \underline{\tilde{x}} + \mathbf{U}^T \mathbf{V} \frac{\int_{\mathcal{R}^{**}} \underline{y} \exp\{-\frac{1}{2}\|\underline{y}\|^2\} d\underline{y}}{\int_{\mathcal{R}^{**}} \exp\{-\frac{1}{2}\|\underline{y}\|^2\} d\underline{y}} =: \underline{\tilde{x}} + \mathbf{U}^T \mathbf{V} \frac{\underline{i}_1}{i_2} \qquad (8)$$

where $\mathcal{R}^{**}$ is the ellipsoid $\mathcal{R}^{**} = \{\underline{y} : (\underline{y} - \underline{y}_0)^T \mathbf{D}(\underline{y} - \underline{y}_0) \le R^2\}$ with $\underline{y}_0 = (\mathbf{U}^T \mathbf{V})^{-1}(\underline{\theta}_0 - \underline{\tilde{x}})$.

Step C. The MMSE estimator requires in general a multidimensional integration which is very computational demanding. However, in this particular case the numerical integration can be substantially simplified. In [9, 10] it is shown that i_2

can be expressed as an infinite linear combination of χ^2 distributions, i.e. $i_2 = (2\pi)^{M/2} \sum_{k=0}^{\infty} a_k G(R^2/\beta; M + 2k)$ where $G(R^2/\beta; M + 2k)$ is the central χ^2 cumulated distribution function (cdf) with $M + 2k$ degrees of freedom evaluated at position R^2/β and β is an arbitrary constant. The coefficients a_k can be found by the recursive rule

$$a_0 = \prod_{m=0}^{M} \sqrt{\frac{\beta}{d_m}} e^{-y_{0m}^2/2}, \qquad a_k = \frac{1}{k} \sum_{l=0}^{k-1} b_{k-l} a_l$$

where $b_k = \frac{k}{2} \sum_{m=1}^{M} y_{0m}^2 \gamma_m^{k-1} + \frac{1}{2} \sum_{m=1}^{M} (1 - k y_{0m}^2) \gamma_m^k$, $\gamma_m = 1 - \frac{\beta}{d_m}$ and y_{0m} is the mth component of $\underline{y}_0$. The calculation of i_2 is therefore straightforward where the summation is stopped if the relative error is small enough. An estimate of the truncation error is derived in [9] and was used as stopping criterion. It therefore remains to calculate $\underline{i}_1$. Its mth element reads

$$i_{1m} = \int_{\mathcal{R}^{**}} y_m e^{-\frac{1}{2}\|\underline{y}\|^2} d\underline{y}$$

$$= \int_{y_{0m} - \frac{R}{\sqrt{d_m}}}^{y_{0m} + \frac{R}{\sqrt{d_m}}} y_m e^{-\frac{1}{2} y_m^2} \left(\int_{\mathcal{R}_m^{**}} e^{-\frac{1}{2}\|\underline{\tilde{y}}_m\|^2} d\underline{\tilde{y}}_m \right) dy_m \qquad (9)$$

where $\underline{\tilde{y}}_m = \mathbf{P}_m \underline{y}$ is identical to $\underline{y}$ except for the removed mth element. The selection matrix $\mathbf{P}_m \in \mathbb{R}^{(M-1) \times M}$ is the identity matrix with the mth row erased. Note, that the inner integral is of the same type as i_2 and can thus also be expressed as an infinite series of χ^2 cumulated density functions. The integration area $\mathcal{R}_m^{**}$ is given by $\mathcal{R}_m^{**} = \{\underline{\tilde{y}}_m : (\underline{\tilde{y}}_m - \mathbf{P}_m \underline{y}_0)^T \mathbf{P}_m \mathbf{D} \mathbf{P}_m^T (\underline{\tilde{y}}_m - \mathbf{P}_m \underline{y}_0) \le R^2 - d_m(y_m - y_{0m})^2\}$ and is a function of y_m. Eq. (9) can be efficiently solved by an one-dimensional numerical integration.

For the special case $M = 1$, i.e. $\underline{\theta} = \theta$ is a scalar, we can simplify (8) further and give a closed-form solution for $\hat{\theta}$. Using $\tilde{x} = \underline{h}^T \mathbf{C}^{-1} \underline{x}/(\underline{h}^T \mathbf{C}^{-1} \underline{h})$, $U = (\underline{h}^T \mathbf{C}^{-1} \underline{h})^{-1/2}$, $V = 1$, $D = d = \Theta/(\underline{h}^T \mathbf{C}^{-1} \underline{h})$ and $y_0 = (\underline{h}^T \mathbf{C}^{-1} \underline{h})^{1/2}(\theta_0 - \tilde{x})$, we obtain (10) at the bottom of this page.

4. CASE B: MMSE ESTIMATOR FOR $\underline{\theta} \in \bar{\mathcal{R}}^*$

For the case B that $\underline{\theta}$ lies on the surface of an ellipsoid, i.e. $\underline{\theta} \in \bar{\mathcal{R}}^*$, we can use the estimator (8) found in Sec. 3 where all integrations are now with respect to $\bar{\mathcal{R}}^*$ from (2b) and $\bar{\mathcal{R}}^{**} = \{\underline{y} : (\underline{y} - \underline{y}_0)^T \mathbf{D}(\underline{y} - \underline{y}_0) = R^2\}$ opposed to $\mathcal{R}^*$ and $\mathcal{R}^{**}$. Only the

$$\hat{\theta} = \frac{\underline{h}^T \mathbf{C}^{-1} \underline{x}}{\underline{h}^T \mathbf{C}^{-1} \underline{h}} + \frac{1}{(\underline{h}^T \mathbf{C}^{-1} \underline{h})^{1/2}} \frac{\int_{y_0 - \frac{R}{\sqrt{d}}}^{y_0 + \frac{R}{\sqrt{d}}} y e^{-\frac{1}{2} y^2} dy}{\int_{y_0 - \frac{R}{\sqrt{d}}}^{y_0 + \frac{R}{\sqrt{d}}} e^{-\frac{1}{2} y^2} dy} = \frac{\underline{h}^T \mathbf{C}^{-1} \underline{x}}{\underline{h}^T \mathbf{C}^{-1} \underline{h}} + \frac{1}{(\underline{h}^T \mathbf{C}^{-1} \underline{h})^{1/2}} \frac{2 e^{-\frac{1}{2}(y_0^2 + \frac{R^2}{d})} \sinh\left(y_0 \frac{R}{\sqrt{d}}\right)}{\sqrt{\frac{\pi}{2}} \left(\mathrm{erf}\left(\frac{y_0 + \frac{R}{\sqrt{d}}}{\sqrt{2}}\right) - \mathrm{erf}\left(\frac{y_0 - \frac{R}{\sqrt{d}}}{\sqrt{2}}\right) \right)} \qquad (10)$$

third step where the integrals are solved has to be changed. It holds

$$i_{1m} = \int\limits_{\mathcal{R}^{**}} y_m e^{-\frac{1}{2}\|\underline{y}\|^2} d\underline{y}$$

$$= \int\limits_{y_{0m}-\frac{R}{\sqrt{d_m}}}^{y_{0m}+\frac{R}{\sqrt{d_m}}} y_m e^{-\frac{1}{2}y_m^2} \left(\int\limits_{\tilde{\mathcal{R}}_m^{**}} e^{-\frac{1}{2}\|\underline{\tilde{y}}_m\|^2} d\underline{\tilde{y}}_m \right) dy_m \quad (11a)$$

$$i_2 = (2\pi)^{M/2} \frac{2}{\beta} \sum_{k=0}^{\infty} a_k g(R^2/\beta; M+2k) \quad (11b)$$

where $g(R^2/\beta; M+2k)$ is the central χ^2 probability density function (pdf) with $M+2k$ degrees of freedom evaluated at position R^2/β. Similar to Sec. 3, we can express i_2 as an infinite series of χ^2 densities and $\underline{i}_1$ is found again by M one-dimensional numerical integrations.

5. SIMULATION RESULTS

In the following we show some simulation results for case A and B. In both cases, $M = 3$ unknowns have to be estimated from $N = 10$ observations. The model matrix $\mathbf{H}$ is randomly chosen for each trial and in total we average over 10 000 trials, except for Fig. 2 where we average over 100 000 trials. The noise $\underline{z}$ has the distribution $\mathcal{N}(\underline{0}, \sigma^2 \mathbf{I})$ and the signal-to-noise ratio (SNR) is defined as SNR $= 10 \log_{10} \|\mathbf{H}\underline{\theta}\|^2/(N\sigma^2)$. For case A, we assume that the parameter vector $\underline{\theta}$ is restricted to lie in the unit sphere $\mathcal{R}^* = \{\underline{\theta} : \|\underline{\theta}\|^2 \leq 1\}$ and is uniformly distributed for the simulation. For case B, we assume that $\underline{\theta}$ is uniformly distributed on the unit sphere $\bar{\mathcal{R}}^* = \{\underline{\theta} : \|\underline{\theta}\|^2 = 1\}$.

5.1. Case A

We compare the derived MMSE estimator with the following four estimators for a spherical constraint:

- Ordinary LS estimator [8]

$$\hat{\underline{\theta}}_{\text{LS}} = (\mathbf{H}^T\mathbf{C}^{-1}\mathbf{H})^{-1}\mathbf{H}^T\mathbf{C}^{-1}\underline{x} = \underline{\tilde{x}} \quad (12)$$

- Modified ordinary LS estimator, inspired by [5]

$$\hat{\underline{\theta}}_{\text{MLS}} = \begin{cases} \hat{\underline{\theta}}_{\text{LS}} & \|\hat{\underline{\theta}}_{\text{LS}}\| \leq R \\ \frac{R}{\|\hat{\underline{\theta}}_{\text{LS}}\|}\hat{\underline{\theta}}_{\text{LS}} & \|\hat{\underline{\theta}}_{\text{LS}}\| > R \end{cases} \quad (13)$$

- Linear minimax estimator [2, 7]

$$\hat{\underline{\theta}}_{\text{MX}} = \frac{R^2}{R^2 + \text{tr}\{(\mathbf{H}^T\mathbf{C}^{-1}\mathbf{H})^{-1}\}}\hat{\underline{\theta}}_{\text{LS}} \quad (14)$$

- Modified minimax estimator [5]

$$\hat{\underline{\theta}}_{\text{MMX}} = \begin{cases} \hat{\underline{\theta}}_{\text{MX}} & \|\hat{\underline{\theta}}_{\text{MX}}\| \leq R \\ \frac{R}{\|\hat{\underline{\theta}}_{\text{MX}}\|}\hat{\underline{\theta}}_{\text{MX}} & \|\hat{\underline{\theta}}_{\text{MX}}\| > R \end{cases} \quad (15)$$

Fig. 1 shows the simulation results for the squared error $\|\underline{\theta} - \hat{\underline{\theta}}\|^2$ averaged over all $\underline{\theta} \in \mathcal{R}^*$. The derived MMSE estimator has clearly the minimal averaged squared error and is therefore superior to the other estimators as expected. It has, however, a higher computational complexity because of the M one-dimensional numerical integrations.

The second comparison is in terms of the risk of the linear minimax and the MMSE estimator. The risk of an estimator that corresponds to a quadratic loss function is [8]

$$R(\underline{\theta}, \hat{\underline{\theta}}) = \mathrm{E}_{\underline{x}|\underline{\theta}}\|\underline{\theta} - \hat{\underline{\theta}}(\underline{x})\|^2 = \int \|\underline{\theta} - \hat{\underline{\theta}}(\underline{x})\|^2 p(\underline{x}|\underline{\theta})d\underline{x}. \quad (16)$$

It is the quadratic loss $\|\underline{\theta} - \hat{\underline{\theta}}\|^2$ averaged over the distribution of the measurements with $\underline{\theta}$ fixed. This comparison is interesting as the linear minimax estimator minimizes the worst case mean squared error for each deterministic $\underline{\theta} \in \mathcal{R}^*$ opposed to the MMSE estimator which only considers the overall mean squared error. Note, that the estimator risk is rotational invariant in the parameter space of $\underline{\theta}$ as the model matrix $\mathbf{H}$ is chosen randomly. Thus, it is sufficient to plot the estimator risk as a function of the norm of $\underline{\theta}$ only and the risk in (16) is calculated by averaging the squared error over all $\underline{\theta}$ with $\|\underline{\theta}\| = \text{const}$. Fig. 2 depicts the simulation results for a SNR of 0dB. It shows that the MMSE estimator is, except for a small region around $\|\underline{\theta}\| = 0.4$, better than the linear minimax estimator. Thus, although we derived the estimator which minimizes the mean squared error averaged over all possible parameter vectors in $\mathcal{R}^*$, it is still almost always better in terms of the risk than the linear minimax estimator which was designed to minimize the worst case error for the deterministic case.

5.2. Case B

We compare the derived MMSE estimator with the ordinary least squares estimator (12) and the following three estimators:

- Scaled ordinary LS estimator, inspired by [5]

$$\hat{\underline{\theta}}_{\text{ScLS}} = \frac{R}{\|\hat{\underline{\theta}}_{\text{LS}}\|}\hat{\underline{\theta}}_{\text{LS}} \quad (17)$$

- Spherical LS estimator, corresponds to the class of constrained LS estimators [11]

$$\hat{\underline{\theta}}_{\text{SLS}} = \min_{\underline{\theta}}\|\underline{x} - \mathbf{H}\underline{\theta}\|^2 \text{ s.t. } \|\underline{\theta}\| = R \quad (18)$$

which we solved by expressing $\underline{\theta}$ in spherical coordinates and using a nonlinear least squares optimization procedure.

- Linear minimax estimator, which is for case B equal to the estimator given in (14) as the worst case MSE to be minimized is always located on the boundary of the ellipsoid. Hence, calculating the linear minimax estimator with respect to $\bar{\mathcal{R}}^*$ is equal to calculating it with respect to $\mathcal{R}^*$.

Fig. 3 and 4 show simulation results for case B with respect to the averaged squared error $\|\underline{\theta} - \hat{\underline{\theta}}\|^2$ and the averaged angle between $\underline{\theta}$ and $\hat{\underline{\theta}}$ in degrees $\cos^{-1}\left(\frac{\underline{\theta}^T\hat{\underline{\theta}}}{\|\underline{\theta}\|\|\hat{\underline{\theta}}\|}\right)$. The derived MMSE estimator is superior to the other estimators for both error measures. Especially the averaged angle error is interesting for DOA applications as in [6] where it is important that $\underline{\theta}$ and $\hat{\underline{\theta}}$ point to the same direction.

6. CONCLUSIONS

In this paper we derived the MMSE estimator for a linear Gaussian model under ellipsoidal constraints. It outperforms all other estimators we know at the expense of a higher computational complexity. A Matlab implementation of the MMSE estimator and the other considered estimators can be downloaded from [12].
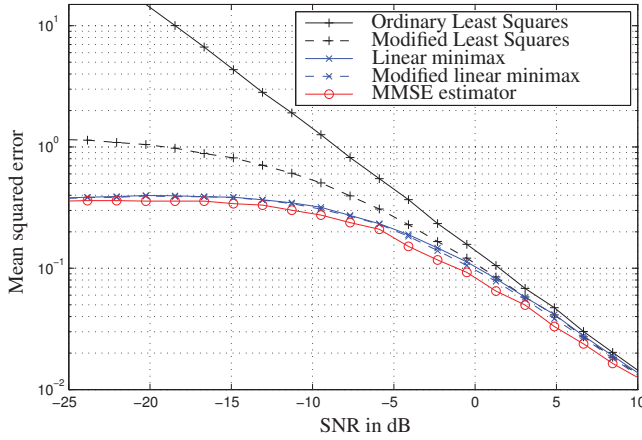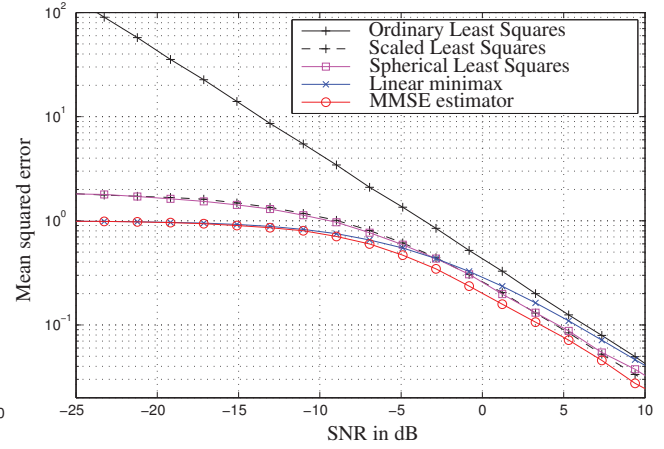
Fig. 1. Case A: Comparison of averaged squared error

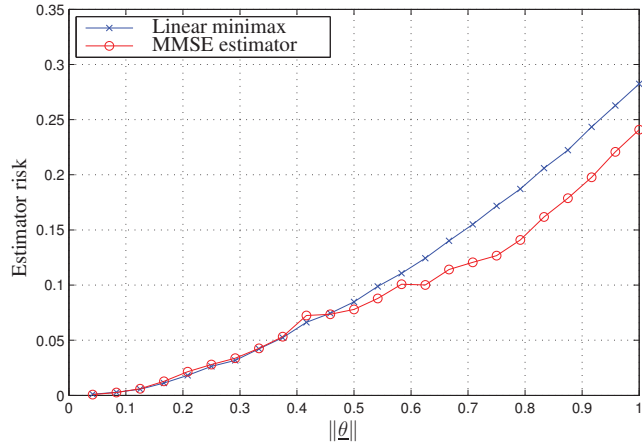**Fig. 3**. Case B: Comparison of averaged squared error

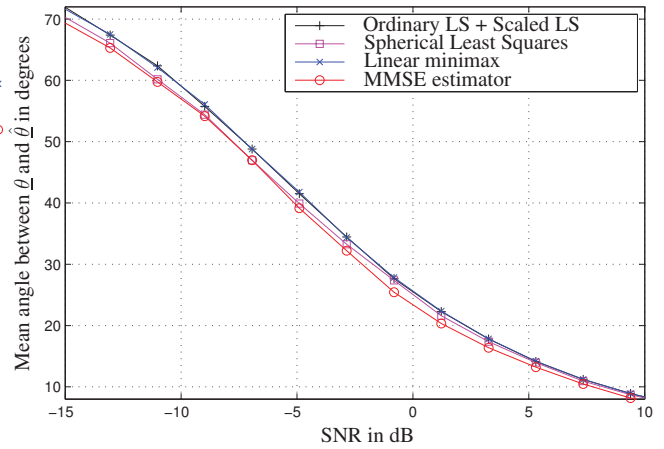**Fig. 2**. Case A: Comparison of estimator risk

Fig. 4. Case B: Comparison of averaged angle in degrees

7. REFERENCES

[1] M. S. Pinsker, "Optimal filtration of square-integrable signals in Gaussian noise," *Problems of Information Transmission*, vol. 16, pp. 120–133, 1980.

[2] J. Pilz, "Minimax linear regression estimation with symmetric parameter restrictions," *Journal of Statistical Planning and Inference*, vol. 13, pp. 297–318, 1986.

[3] Y. C. Eldar, "Uniformly improving the Cramer-Rao Bound and Maximum-Likelihood estimation," *IEEE Trans. Signal Processing*, vol. 54, no. 8, pp. 2943–2956, 2006.

[4] Y. C. Eldar, "Universal weighted MSE improvement of the Least-Squares estimator," *IEEE Trans. Signal Processing*, vol. 56, no. 5, pp. 1788–1800, 2008.

[5] P. Stoica and G. Ganesan, "Linear regression constrained to a ball," *Digital Signal Processing*, vol. 11, no. 1, pp. 80–90, 2000.

[6] R. Mukai, H. Sawada, S. Araki, and S. Makino, "Real-time blind source separation and DOA estimation using small 3-D microphone array," *Proc. International Workshop for Acoustic Echo and Noise Control (IWAENC)*, pp. 45–48, 2005.

[7] Y. C. Eldar, A. Ben-Tal, and A. Nemirovski, "Robust mean-squared error estimation in the presence of model uncertainties," *IEEE Trans. Signal Processing*, vol. 53, no. 1, pp. 168–181, 2005.

[8] L. L. Scharf, *Statistical Signal Processing: Detection, Estimation and Time Series Analysis*, Addison-Wesley, 1990.

[9] H. Ruben, "Probability content of regions under spherical normal distributions, IV: The distribution of homogeneous and non-homogeneous quadratic functions of normal variables," *Ann. of Math. Statist.*, vol. 33, no. 2, pp. 542–570, 1962.

[10] N. L. Johnson and S. Kotz, *Distributions in Statistics: Continuous Univariate Distributions 2*, John Wiley & Sons, 1970.

[11] T. K. Moon and W. C. Stirling, *Mathematical Methods and Algorithms for Signal Processing*, Prentice-Hall, 2000.

[12] S. Uhlich, "Matlab implementation of the MMSE estimator for a restricted parameter vector in a linear Gaussian model," Available: http://www.lss.uni-stuttgart.de/mitarbeiter/uhlich.en.html.