Proceedings of the American Control Conference
Albuquerque, New Mexico June 1997
0-7803-3832-4/97/$10.00 © 1997 AACC

WORST-CASE PARAMETER ESTIMATION WITH BOUNDED MODEL UNCERTAINTIES*

S. CHANDRASEKARAN, GENE H. GOLUB, MING GU, AND ALI H. SAYED

Abstract

We formulate and solve a new parameter estimation problem in the presence of bounded model uncertainties. The new method is suitable when a-priori bounds on the uncertain data are available, and its solution guarantees that the effect of the uncertainties will never be unnecessarily over-estimated beyond what is reasonably assumed by the a-priori bounds. This is in contrast to other methods, such as total least-squares and robust estimation, that do not incorporate explicit bounds on the size of the uncertainties. A geometric interpretation of the solution of the new problem is provided, along with a closed form expression for it. We also consider the case in which only selected columns of the coefficient matrix are subject to perturbations.

1. INTRODUCTION

The central problem in estimation is to recover, to good accuracy, a set of unobservable parameters from corrupted data. Several optimization criteria have been used for estimation purposes over the years, but the most important, at least in the sense of having had the most applications, are criteria that are based on quadratic cost functions. The most striking among these is the linear least-squares criterion, which was perhaps first developed by Gauss (ca. 1795) in his work on celestial mechanics. Since then, it has enjoyed widespread popularity in many diverse areas as a result of its attractive computational and statistical properties.

Alternative optimization criteria have also been proposed over the years including, among others, regularized least-squares, ridge regression, total least-squares, and robust (or H_∞) estimation (see, e.g., [1]-[5]). These different formulations allow, in one way or another, incorporation of further a priori information about the unknown parameter into the problem statement. They are also more effective in the presence of data errors and incomplete statistical information about the exogenous signals (or measurement errors).

*S. Chandrasekaran is with Dept. of ECE, UCSB, CA 93106. G. H. Golub is with Dept. of CS, Stanford University, CA 94305, M. Gu is with Dept. of Math., UCLA, CA 90095, and A. H. Sayed is with Dept. of EE, UCLA, CA 90095. The work of A. H. Sayed was in part supported by the National Science Foundation under Award No. MIP-9409319.

1.1. The TLS Method

A notable variation is the total least-squares (TLS) method, also known as orthogonal regression or errors-in-variables method in statistics and system identification [6]. In contrast to the standard least-squares problem, the TLS formulation allows for errors in the data matrix. But it still exhibits certain drawbacks that degrade its performance in some practical situations. In particular, it may unnecessarily over-emphasize the effect of noise and uncertainties and can, therefore, lead to overly conservative results.

To clarify this remark, assume $A \in \mathbf{R}^{m \times n}$ is a given full rank matrix with $m \geq n$, and $b \in \mathbf{R}^m$ is a given vector. Consider the problem of solving the inconsistent linear system $A\hat{x} \approx b$ in the least-squares sense. The TLS solution assumes data uncertainties in A and proceeds to correct A and b by replacing them by their projections, $\hat{A}$ and $\hat{b}$, onto a specific subspace and by solving the now consistent linear system of equations $\hat{A}\hat{x} = \hat{b}$. The spectral norm of the correction $(A - \hat{A})$ in the TLS solution is bounded by the smallest singular value of $\begin{bmatrix} A & b \end{bmatrix}$. While this norm might be small for vectors b that are close enough to the range space of A, it need not always be so. In other words, the TLS solution may lead to situations in which the correction term is unnecessarily large.

Consider, for example, a situation in which the uncertainties in A are very small, say A is almost known exactly. Assume further that b is far from the column space of A. In this case, it is not difficult to visualize that the TLS solution will need to rotate (A, b) into $(\hat{A}, \hat{b})$ and may therefore end up with an overly corrected approximant for A, despite the fact that A is almost exact.

1.2. Motivation

These facts motivate us to introduce a parameter estimation formulation that imposes bounds on the size of the allowable corrections to the data. More specifically, we formulate and solve a new estimation problem that is more suitable for scenarios in which a-priori bounds on the uncertain data are known. In this case, the solution will guarantee that the effect of the uncertainties will never be unnecessarily over-estimated, beyond what is reasonably assumed by the a-priori bounds.

We note that while preparing this paper, the related work [7] has come to our attention, where the authors have independently formulated and solved a similar estimation problem by using convex semidefinite programming techniques and interior-point methods. The resulting computational complexity of the proposed solution is $O(nm^2 + m^{3.5})$, where m is the larger matrix dimension.

The solution proposed in this paper proceeds by first providing a geometric formulation of the problem, followed by an algebraic derivation that establishes that the optimal solution can in fact be obtained by solving a related regularized problem. The regression parameter of the regularization step is further shown to be obtained as the unique positive root of a secular equation and as a function of the given data. In this sense, the new formulation turns out to provide automatic regularization and, hence, has some useful regularization properties: the regularization parameter is not selected by the user but rather determined by the algorithm. Our solution involves an SVD step and its computational complexity amounts to $O(mn^2 + n^3)$, where n is the smaller matrix dimension. In the companion paper [8] we study an alternative problem formulation that involves a non-convex cost function, with further interesting connections with TLS and H_∞ methods.

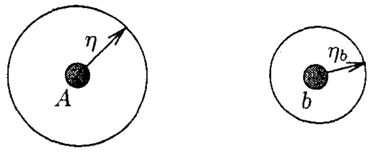

Figure 1: *Bounded data uncertainties.*

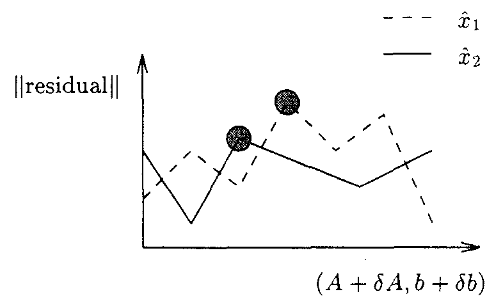

Figure 2: *Two illustrative residual-norm curves.*

2. PROBLEM FORMULATION

Let $A \in \mathbf{R}^{m\times n}$ be a given full rank matrix with $m \geq n$ and let $b \in \mathbf{R}^m$ be a given vector. The quantities (A, b) are assumed to be linearly related via an unknown vector of parameters $x \in \mathbf{R}^n$, $b = A \cdot x + v$, where $v \in \mathbf{R}^m$ explains the mismatch between $A \cdot x$ and b.

We assume that the "true" coefficient matrix is $A+\delta A$, and that we only know an upper bound on the perturbation δA, say $\|\delta A\|_2 \leq \eta$. Likewise, we assume that the "true" observation vector is $b + \delta b$, and that we know an upper bound η_b on the perturbation δb, say $\|\delta b\|_2 \leq \eta_b$. The notation $\|\cdot\|_2$ denotes either the 2−induced norm of its matrix argument or the Euclidean norm of its vector argument.

We pose the problem of finding an estimate $\hat{x}$ that performs "well" for any possible perturbation $(\delta A, \delta b)$. That is, we would like to determine, if possible, an $\hat{x}$ that solves

$$\min_{\hat{x}} \left(\max_{\|\delta A\|_2 \leq \eta,\ \|\delta b\|_2 \leq \eta_b} \| (A + \delta A) \cdot \hat{x} - (b + \delta b)\|_2 \right) . \quad (1)$$

Any value that we pick for $\hat{x}$ would lead to many residuals norms, $\| (A + \delta A)\cdot\hat{x} - (b+\delta b)\|_2$, one for each possible choice of A in the disc $(A + \delta A)$ and b in the disc $(b + \delta b)$. We want to determine the particular value(s) for $\hat{x}$ whose maximum residual is the least possible. The situation is depicted in Figs. 1 and 2. It turns out that this problem always has a unique solution except in a special degenerate case in which the solution is nonunique.

We note that if $\eta = 0 = \eta_b$, then problem (1) reduces to a standard least squares problem. Therefore, we shall assume throughout that $\eta > 0$. [It will turn out that the solution to the above constrained min-max problem is independent of η_b].

2.1. Geometric Interpretation

The problem also admits an interesting geometric formulation. For this purpose, and for the sake of illustration, assume we have a unit-norm vector b, $\|b\|_2 = 1$, with no uncertainties in it ($\eta_b = 0$; it turns out that the solution does not depend on η_b). Assume further that A is simply a column vector, say a, with $\eta \neq 0$, and consider (1) in this setting:

$$\min_{\hat{x}} \left(\max_{\|\delta a\|_2 \leq \eta} \| (a + \delta a) \cdot \hat{x} - b\|_2 \right) .$$

The situation is depicted in Fig. 3. The vectors a and b are indicated in thick black lines. The vector a is shown in the horizontal direction and a circle of radius η around its vertex indicates the set of all possible vertices for $a + \delta a$.

For any $\hat{x}$ that we pick, the set $\{(a + \delta a)\hat{x}\}$ describes a disc of center $a\hat{x}$ and radius $\eta\hat{x}$. This is indicated in the figure by the largest rightmost circle, which corresponds to a choice of a positive $\hat{x}$ that is larger than one. The vector in $\{(a+\delta a)\hat{x}\}$ that is furthest away from b is the one obtained by drawing a line from b through the center of the rightmost circle. The intersection of this line with the circle defines a residual vector r_3 whose norm is the largest among all possible residual vectors in the set $\{(a + \delta a)\hat{x}\}$.

Likewise, if we draw a line from b that passes through the vertex of a, it will intersect the circle at a point that defines a residual vector r_2. This residual will have the largest norm among all residuals that correspond to the particular choice $\hat{x} = 1$.

More generally, any $\hat{x}$ that we pick will determine a circle, and the corresponding largest residual is obtained by finding the furthest point on the circle from b. This is the point where the line that passes through b and the

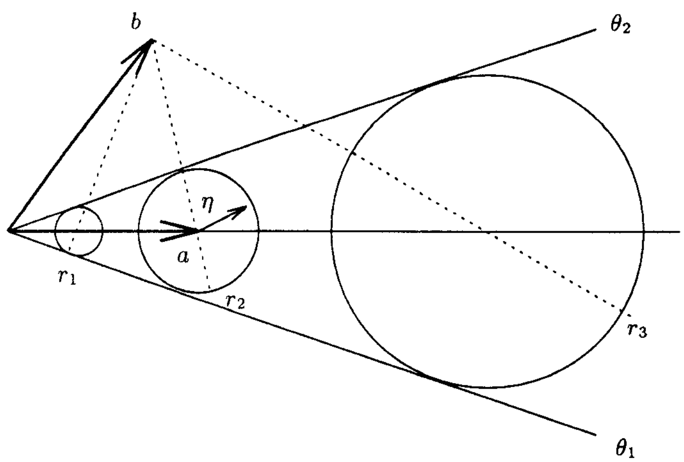

Figure 3: *Geometric construction in a simplified scenario.*

center of the circle intersects the circle on the other side of b.

We need to pick an $\hat{x}$ that minimizes the largest residual. For example, it is clear from the figure that the norm of r_3 is larger than the norm of r_2. The claim is that in order to minimize the largest residual we need to proceed as follows: we drop a perpendicular from b to the lower tangent line denoted by θ_1. This perpendicular intersects the horizontal line in a point where we draw a new circle (the leftmost circle) that is tangent to both θ_1 and θ_2. This circle corresponds to a choice of $\hat{x}$ such that the furthest point on it from b is the foot of the perpendicular from b to θ_1. The residual indicated by r_1 corresponds to the desired solution (it has the minimum norm among the largest residuals). The radius of this circle will be $\eta\hat{x}$, where $\hat{x}$ is the optimal solution. Also, the foot of the perpendicular on θ_1 will be the optimal $\hat{b}$.

The projection $\hat{b}$ (and consequently the solution $\hat{x}$) will be nonzero as long as b is not orthogonal to the direction θ_1. This imposes a condition on η. Indeed, the direction θ_1 will be orthogonal to b only when η is large enough. This requires that the circle centered around a has radius a^Tb, which is the length of the projection of a onto the unit norm vector b. This is depicted in Fig. 4. Hence, the largest value that can be allowed for η in order to have a nonzero solution $\hat{x}$ must be smaller than $\eta < |a^Tb|$.

For a non-unity b, the upper bound on η would take the form $\eta < \frac{|a^Tb|}{\|b\|_2}$. We shall see that in the general case a similar bound holds, for nonzero solutions, and is given by $\eta < \frac{\|A^Tb\|_2}{\|b\|_2}$. We now proceed to an algebraic solution of the constrained min-max problem.

3. ALGEBRAIC SOLUTION

We start by showing how to reduce the constrained min-max problem (1) to a standard minimization problem. To begin with, we note the upper bound:

$$\|(A+\delta A)\,\hat{x} - (b+\delta b)\|_2 \quad\leq\quad \|A\hat{x}-b\|_2 + \eta\|\hat{x}\|_2 \;+\; \eta_b\,,$$

which is in fact achievable for the choices:

$$\delta A^o = \frac{(A\hat{x}-b)}{\|A\hat{x}-b\|_2}\,\frac{\hat{x}^T}{\|\hat{x}\|_2}\eta\,,\quad \delta b^o = -\frac{(A\hat{x}-b)}{\|A\hat{x}-b\|_2}\,\eta_b\,.$$

In this case, the quantities $\{(A\hat{x}-b), \delta A^o\hat{x}, \delta b^o\}$ will be collinear vectors that point in the same direction and it will follow that

$$\|(A+\delta A^o)\,\hat{x} - (b+\delta b^o)\|_2 \quad=\quad \|A\hat{x}-b\|_2 + \eta\|\hat{x}\|_2 + \eta_b\,.$$

This argument establishes that

$$\max_{\|\delta A\|_2\leq\eta\,,\|\delta b\|_2\leq\eta_b} \|(A+\delta A)\,\hat{x} - (b+\delta b)\|_2 =$$

$$\|A\hat{x}-b\|_2 + \eta\|\hat{x}\|_2 + \eta_b\,,$$

which in turn reduces the constrained min-max problem (1) to an unconstrained minimization problem of the form:

$$\min_{\hat{x}} \;(\|A\hat{x}-b\|_2 + \eta\|\hat{x}\|_2 + \eta_b)\;. \tag{2}$$

We should note that this problem formulation is significantly distinct from a regularized least-squares formulation, where the *squared* Euclidean norms $\{\|A\hat{x}-b\|_2^2, \|\hat{x}\|_2^2\}$ are used rather than the norms themselves.

3.1. Solving the Minimization Problem

The cost function $\mathcal{L}(\hat{x}) = \|A\hat{x}-b\|_2 + \eta\|\hat{x}\|_2 + \eta_b$ is convex and continuous in $\hat{x}$. Hence any local minimum of $\mathcal{L}(\hat{x})$ is also a global minimum. But at any local minimum of $\mathcal{L}(\hat{x})$, it either holds that $\mathcal{L}(\hat{x})$ is not differentiable or its gradient $\nabla\mathcal{L}(\hat{x})$ is 0. In particular, note that $\mathcal{L}(\hat{x})$ is not differentiable only at $\hat{x}=0$ and at any $\hat{x}$ that satisfies $A\hat{x}-b=0$.

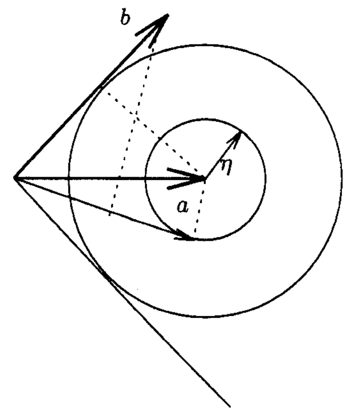

Figure 4: *Geometric condition for a nonzero solution.*

We first consider the case in which $\mathcal{L}(\hat{x})$ is differentiable and, hence, the gradient of $\mathcal{L}(\hat{x})$ exists and is given by

$$\nabla\mathcal{L}(\hat{x}) = \frac{1}{||A\hat{x}-b||_2}\left(\left(A^TA+\alpha I\right)\hat{x}-A^Tb\right),$$

where we have introduced the *positive* real number

$$\alpha = \frac{\eta||A\hat{x}-b||_2}{||\hat{x}||_2}. \qquad (3)$$

By setting $\nabla\mathcal{L}(\hat{x}) = 0$ we obtain that any stationary solution $\hat{x}$ of $\mathcal{L}(\hat{x})$ is given by

$$\hat{x} = \left(A^TA+\alpha I\right)^{-1}A^Tb. \qquad (4)$$

Expressions (3)-(4) define a system of equations with two unknowns $\{\hat{x},\alpha\}$. If we replace (4) into (3) we obtain a nonlinear equation in α, which will further lead to what we shall refer to as the *secular equation*. We shall show later that the secular equation has only one positive root. Hence, determining its positive root uniquely determines α, which in turn uniquely determines $\hat{x}$.

Since we are also interested in the numerical reliability of the resulting computational procedure, we find it useful to perform these substitutions and calculations by invoking the SVD of A, say

$$A = U\begin{bmatrix}\Sigma\\0\end{bmatrix}V^T, \qquad (5)$$

where $U \in \mathbf{R}^{m\times m}$ and $V \in \mathbf{R}^{n\times n}$ are orthogonal, and $\Sigma = \text{diag}(\sigma_1,\ldots,\sigma_n)$ is diagonal, with $\sigma_1 \geq \cdots \geq \sigma_n \geq 0$, being the singular values of A. We further partition the vector U^Tb into $\begin{bmatrix}b_1\\b_2\end{bmatrix} = U^Tb$, where $b_1 \in \mathbf{R}^n$ and $b_2 \in \mathbf{R}^{m-n}$. In this case, the expression (3) can be seen to collapse to

$$\alpha = \frac{\eta\sqrt{||b_2||_2^2+\alpha^2||\left(\Sigma^2+\alpha I\right)^{-1}b_1||_2^2}}{||\Sigma\left(\Sigma^2+\alpha I\right)^{-1}b_1||_2}. \qquad (6)$$

Note that only the norm of b_2, and not b_2 itself, is needed in the above expression.

Remark. We have assumed in the derivation so far that A is full rank. If this were not the case, i.e., if A (and hence Σ) were singular, then equation (6) can be reduced to an equation of the same form but with a non-singular Σ of smaller dimension. Indeed, if we partition

$$\Sigma = \begin{bmatrix}\hat{\Sigma} & 0\\0 & 0\end{bmatrix},$$

where $\hat{\Sigma} \in \mathbf{R}^{k\times k}$ is non-singular, and let $\hat{b}_1 \in \mathbf{R}^k$ be the first k components of b_1; $\tilde{b}_1 \in \mathbf{R}^{n-k}$ be the last $n-k$ components of b_1; and let

$$||\hat{b}_2||_2^2 = ||b_2||_2^2 + ||\tilde{b}_1||_2^2.$$

Then equation (6) reduces to

$$\alpha = \frac{\eta\sqrt{||\hat{b}_2||_2^2+\alpha^2||\left(\hat{\Sigma}^2+\alpha I\right)^{-1}\hat{b}_1||_2^2}}{||\hat{\Sigma}\left(\hat{\Sigma}^2+\alpha I\right)^{-1}\hat{b}_1||_2}, \qquad (7)$$

the same form as (6). From now on, we assume that A is full rank and, hence, Σ is invertible.

3.2. The Secular Equation

Define the nonlinear function in α,

$$\mathcal{G}(\alpha) = b_1^T\left(\Sigma^2-\eta^2 I\right)\left(\Sigma^2+\alpha I\right)^{-2}b_1 - \frac{\eta^2}{\alpha^2}||b_2||_2^2. \qquad (8)$$

It is clear that α is a positive solution to (6) if, and only if, it is a positive root of $\mathcal{G}(\alpha)$. Following [1][p. 564], we refer to the equation $\mathcal{G}(\alpha) = 0$ as a *secular* equation.

The function $\mathcal{G}(\alpha)$ has several useful properties that will allow us to provide conditions for the existence of a unique positive root α (proofs are omitted for brevity - see [9]). First note that a necessary and sufficient condition for b to belong to the column span of A is $b_2 = 0$.

Lemma. *The function $\mathcal{G}(\alpha)$ satisfies the following:*

1. *It can have at most one positive root. In addition, if $\hat{\alpha} > 0$ is a root of $\mathcal{G}(\alpha)$, then $\hat{\alpha}$ is a simple root.*

2. *Assume $b_2 \neq 0$, i.e., b does not belong to the column span of A. Then $\mathcal{G}(\alpha)$ has a unique positive root if, and only if,*

$$0 < \eta < \frac{||A^Tb||_2}{||b||_2}.$$

3. *Assume $b_2 = 0$, i.e., b belongs to the column span of A (this case arises, for example, when A is square and invertible). Define*

$$\tau_1 = \frac{||\Sigma^{-1}b_1||_2}{||\Sigma^{-2}b_1||_2} \quad \textit{and} \quad \tau_2 = \frac{||\Sigma b_1||_2}{||b_1||_2}. \qquad (9)$$

Then $\mathcal{G}(\alpha)$ has a unique positive root if, and only, if $\tau_1 < \eta < \tau_2$.

4. *Whenever $\mathcal{G}(\alpha)$ has a positive root $\hat{\alpha}$, the corresponding vector $\hat{x}$ in (4) must be the global minimizer of $\mathcal{L}(\hat{x})$ (its Hessian matrix is positive-definite).*

We still need to consider the points at which $\mathcal{L}(\hat{x})$ is not differentiable. These include $\hat{x} = 0$ and any solution of $A\hat{x} = b$. We omit the details and state the final result.

3.3. Statement of Solution

The solution of the constrained min-max problem proceeds as follows.

Theorem. *Given $A \in \mathbf{R}^{m\times n}$, with $m \geq n$ and A full rank, $b \in \mathbf{R}^m$, and nonnegative real numbers (η, η_b). The optimization problem (1) always has a solution $\hat{x}$. The solution(s) can be constructed as follows.*

- *Introduce the SVD of A as in (5).*
- *Partition the vector $U^T b$ into $col\{b_1, b_2\}$, where $b_1 \in \mathbf{R}^n$ and $b_2 \in \mathbf{R}^{m-n}$.*
- *Introduce the secular function $\mathcal{G}(\alpha)$ as in (8).*
- *Define $\{\tau_1, \tau_2\}$ as in (9).*

First case: b does not belong to the column span of A.

1. *If $\eta \geq \tau_2$ then the unique solution is $\hat{x} = 0$.*
2. *If $\eta < \tau_2$ then the unique solution is $\hat{x} = (A^TA + \hat{\alpha}I)^{-1}A^Tb$, where $\hat{\alpha}$ is the unique positive root of the secular equation $\mathcal{G}(\alpha) = 0$.*

Second case: b belongs to the column span of A.

1. *If $\eta \geq \tau_2$ then the unique solution is $\hat{x} = 0$.*
2. *If $\tau_1 < \eta < \tau_2$ then the unique solution is $\hat{x} = (A^TA + \hat{\alpha}I)^{-1}A^Tb$, where $\hat{\alpha}$ is the unique positive root of the secular equation $\mathcal{G}(\alpha) = 0$.*
3. *If $\eta \leq \tau_1$ then the unique solution is $\hat{x} = V\Sigma^{-1}b_1 = A^\dagger b$.*
4. *If $\eta = \tau_1 = \tau_2$ then there are infinitely many solutions that are given by $\hat{x} = \beta V\Sigma^{-1}b_1 = \beta A^\dagger b$, for any $0 \leq \beta \leq 1$.*

3.4. Automatic Regularization

Note that the expression for the unique solution $\hat{x}$ has the form

$$\hat{x} = (A^TA + \hat{\alpha}I)^{-1}A^Tb \,.$$

This can be regarded as the exact solution of a regularized least-squares problem of the form:

$$\min_{\hat{x}} \ \left(\hat{\alpha}\|\hat{x}\|_2^2 \ + \ \|A\hat{x} - b\|_2^2\right)$$

with squared Euclidean distances. In this sense, the solution to the original problem (2) (with norms only rather than squared norms) can be seen to lead to automatic regularization. That is, the solution first determines a regularization parameter $\hat{\alpha}$ and then uses it to solve a regularized least-squares problem of the above form.

The scalar $\hat{\alpha}$ can be determined by employing a bisection-type algorithm to solve the secular equation, thus requiring $O\left(n \log \frac{\hat{\alpha}}{\epsilon}\right)$, where ϵ is the desired precision.

4. RESTRICTED PERTURBATIONS

We have so far considered the case in which all the columns of the A matrix are subject to perturbations. It may happen in practice, however, that only selected columns are uncertain, while the remaining columns are known precisely. This situation can be handled by the approach of this paper. The details can be found in [9]. We only state the problem here.

Given $A \in \mathbf{R}^{m\times n}$, we partition it into block columns, $A = \left[\begin{array}{cc} A_1 & A_2 \end{array}\right]$, and assume, without loss of generality, that only the columns of A_2 are subject to perturbations while the columns of A_1 are known exactly. We can then pose the following problem. Determine $\hat{x}$ such that

$$\min_{\hat{x}} \ \max_{\|\delta A_2\|_2 \leq \eta_2, \ \|\delta b\|_2 \leq \eta_b} \left\{ \left\| \left[\begin{array}{cc} A_1 & A_2 + \delta A_2 \end{array}\right] \hat{x} - (b + \delta b) \right\|_2 \right\} .$$

Its solution can be found in [9].

5. FURTHER REMARKS

The solution proposed herein requires the computation of the SVD of the data matrix and the determination of the unique positive root of the nonlinear secular equation. In the companion paper [10], we establish the existence of a fundamental contraction mapping and use this observation to propose an approximate recursive algorithm that avoids the need for explicit SVDs and for the solution of the nonlinear equation.

Moreover, in the second companion paper [8] we consider an alternative formulation of the estimation problem that has some interesting ties with TLS and H_∞ methods.

References

[1] G. H. Golub and C. F. Van Loan, *Matrix Computations*, The Johns Hopkins University Press, Baltimore, third ed., 1997.

[2] S. V. Huffel and J. Vandewalle, *The Total Least Squares Problem: Computational Aspects and Analysis*, SIAM, Philadelphia, 1991.

[3] B. Hassibi, A. H. Sayed, and T. Kailath, *Linear estimation in Krein spaces - Part I: Theory*, IEEE Trans. Automatic Control, 41 (1996), pp. 18–33.

[4] P. Khargonekar and K. M. Nagpal, *Filtering and smoothing in an $H^\infty-$ setting*, IEEE Trans. on Automatic Control, AC-36 (1991), pp. 151–166.

[5] U. Shaked and Y. Theodor, *$H^\infty-$optimal estimation: A tutorial*, in Proc. IEEE Conference on Decision and Control, Tucson, AZ, Dec. 1992, pp. 2278–2286.

[6] L. Ljung and T. Söderström, *Theory and Practice of Recursive Identification*, MIT Press, Cambridge, MA, 1983.

[7] L. E. Ghaoui and H. Lebret, *Robust solutions to least-squares problems with uncertain data*, preprint.

[8] S. Chandrasekaran, G. H. Golub, M. Gu, and A. H. Sayed, *Best-fit parameter estimation for a bounded errors-in-variables model*, in Proceedings of this Conference (ACC'97).

[9] S. Chandrasekaran, G. H. Golub, M. Gu, and A. H. Sayed, *Parameter estimation in the presence of bounded data uncertainties*, SIAM J. Matrix Anal. Appl., to appear.

[10] A. H. Sayed, A. Garulli, and S. Chandrasekaran, *A fast iterative solution for worst-case parameter estimation with bounded model uncertainties*, in Proceedings of this Conference (ACC'97).