Inner Ellipsoidal Approximation of Membership Set: A Fast Recursive Algorithm[†]

F. Dabbene*, P. Gay** and B. T. Polyak***

*DAUIN – Politecnico di Torino – Italy
**DEIAFA – Università degli Studi di Torino – Italy
***Institute for Control Science – Moscow – Russia.

Abstract

In robust system identification the measurement errors are usually assumed to be unknown but bounded, leading to a description of the model parameters in terms of membership set. Due to the difficulty to obtain an exact description of this set, in the past several methods have been proposed for its approximation. In this paper a fast recursive algorithm for the determination of an inner ellipsoidal approximation of the membership set is presented.

1 Introduction

A quite standard assumption in the set membership approach to system identification is to consider the measurement error (that includes measurement noise and modeling error) to be unknown but bounded. This leads to the determination of a membership set of the model parameters that, in the case of linear in the parameters models, is a convex polytope whose elements are compatible with the measurements, the assumed model structure and the a priori error bounds. One of the main drawbacks of this approach to identification is given by the difficulty to describe in an exact way the set of feasible parameters, especially when recursive algorithms are needed. To this extent, several algorithms have been proposed in the past years based on the idea of approximating the convex polytope of feasible parameters with different low complexity sets. However, as the number of measurements increases, the computational burden of computing off-line solutions of such approximation problems easily becomes unaffordable. Hence the interest for recursive algorithms which can update parameter estimates after each measurement while requiring a limited amount of memory and computing time. Examples of such technique can be found in e.g. [6] (ellipsoidal approximation), [2] (orthotopes) and [11] (low complexity polytopes).

In this paper a new efficient recursive algorithm for constructing an inner ellipsoidal approximation of the membership set is proposed.

[†]This work was supported by CNR grant 99.01960.CT07. Part of this work was carried on while B. T. Polyak was visiting IRITI-CNR supported by a NATO-CNR Guest Fellowship.

2 Preliminaries and Notation

We consider the single-output linear regression model

$$y = a^T\theta + e \tag{1}$$

where $\theta \in \mathbb{R}^p$ is the (true) parameter vector and $a^T \in \mathbb{R}^p$ is the regressor. The measurement error e is assumed to be unknown but bounded, that is $|e| \leq \bar{e}$. The error bound $\bar{e}$ is assumed to be known a priori.

The experiment condition of a single measurement are summarized by the variable Δ that we refer to as *experiment data*. We define then the set $\boldsymbol{\Delta}$ of all possible experiment outcomes, that describes the set of all regressors and measurement outputs that can possibly occur performing the measurement. This set $\boldsymbol{\Delta}$ can be a closed set, or simply it can be a finite (numerable) set, i.e. $\boldsymbol{\Delta} = \{\Delta_1, \ldots, \Delta_N\}$. The experiment Δ may have either a stochastic or a deterministic nature over the set $\boldsymbol{\Delta}$. In the first case, we define a probability density function (pdf) $f_\Delta(\Delta)$ with support $\boldsymbol{\Delta}$. In other words, in this case we associate to every single measurement a *probability of outcome*.

Each inequality of the type (1) defines a strip $\mathcal{S}(\Delta)$ in the parameter space

$$\mathcal{S}(\Delta) \doteq \left\{\theta : y(\Delta) - \bar{e} \leq a^T(\Delta)\theta \leq y(\Delta) + \bar{e}\right\}, \tag{2}$$

where the notation $y(\Delta), a^T(\Delta)$ is used to indicate that the measurement output and the regressor are relative to the experiment Δ.
The *feasible set* $\mathcal{D}$ is then defined as the set of parameters θ that are consistent with the acquired data, and is represented by the intersection of the strips

$$\mathcal{D} \doteq \left\{\theta : |y(\Delta) - a^T(\Delta)\theta| \leq \bar{e},\ \Delta \in \boldsymbol{\Delta}\right\} = \bigcap_{\Delta\in\boldsymbol{\Delta}} \mathcal{S}(\Delta).$$

In order to approximate the feasible set, we consider ellipsoids of the form

$$\mathcal{E} \doteq \left\{\theta : \|P^{-1}(\theta - c)\| \leq 1\right\}, \tag{3}$$

where $c \in \mathbb{R}^p$ and the positive definite matrix $P = P^T > 0$ represent the ellipsoid center and the *shape* matrix, respectively. These information are collected in a single parameter $x \doteq (c, P)$.

3 Recursive algorithm

We first derive the conditions for the ellipsoid $\mathcal{E}$ to belong to the strip $\mathcal{S}(\Delta)$.

Lemma 1 *The ellipsoid $\mathcal{E}$ defined in (3) belongs to the strip $\mathcal{S}(\Delta)$ given in (2) if and only if the following condition holds*

$$\varphi(x,\Delta) \leq 0, \tag{4}$$

where $\varphi(x,\Delta) \doteq |y(\Delta) - a^T(\Delta)c| + \|Pa(\Delta)\| - \bar{e}$.

The proof of the lemma, in slightly different form, can be found in [4, 10]. We remark that, for fixed Δ, $\varphi(x,\Delta)$ is a scalar convex function of x. The above result can be applied to write the conditions for the ellipsoid $\mathcal{E}$ to belong to the convex set $\mathcal{D}$ as a set of N scalar convex inequalities, where N is the cardinality of $\mathbf{\Delta}$. In contrast with the above mentioned papers, we do not convert the conditions into LMIs, but we deal with these convex inequalities directly.

The problem of finding the largest ellipsoid inscribed in $\mathcal{D}$ (provided that this set is nonempty) can therefore be formulated as a convex optimization problem

$$\max_x f(x), \quad f(x) = \mathrm{Tr}P \tag{5}$$

$$s.t. \quad P > 0, \quad \varphi(x,\Delta) \leq 0 \quad \forall \Delta \in \mathbf{\Delta}.$$

The choice of the cost function $\mathrm{Tr}P$ corresponds to the maximization of the sum of the ellipsoid semiaxes. However, the results derived in this paper can be easily extended to the case of the maximum volume ellipsoid simply choosing the cost function $f(x) = \log\det\ P$. The function $\mathrm{Tr}P$ has some additional advantages, because it is a linear function of P.

If the set $\mathbf{\Delta}$ is finite, this problem reduces to a convex optimization on a set of convex constraints and can be immediately formulated as an LMI optimization problem [4]. However, in practice the number of collected measurements can be quite large, leading to LMIs of large dimensions that are not affordable by the standard tools available. This consideration motivates the need of a recursive algorithm.

In order to find a recursive solution of the above problem, we first concentrate on a feasibility problem. That is, we fix some level μ for the cost function and consider the general problem of iteratively finding a feasible solution to the set of scalar inequalities

$$\varphi(x,\Delta) \leq 0 \quad \forall \Delta \in \mathbf{\Delta}, \quad x \in \mathcal{T},$$

where $\mathcal{T}$ is the closed convex set $\mathcal{T} = \{P \geq 0, TrP \geq \mu\}$. The only assumption needed is a condition of *strong feasibility*, i.e.

$$\exists x^* \in \mathcal{T}, \alpha > 0 : \varphi(x,\Delta) \leq 0$$

for all $\Delta \in \mathbf{\Delta}$ and $x \in \mathcal{T}, ||x - x^*|| \leq \alpha$.

Roughly speaking, we require that the feasible set is not empty and doesn't reduce to a singleton.

Considering a series of experiments $\Delta_1, \ldots, \Delta_k$, under the above assumptions we propose the following recursive algorithm

Algorithm 1 *Let $x_0 = (c_0, P_0) \in \mathcal{T}$ be an initial estimate, construct the following recursion*

$$\begin{cases} x_{k+1} = \mathcal{P}_{\mathcal{T}}(x_k - \lambda_k \partial\varphi(x_k,\Delta_k)) & \text{if } \varphi(x_k,\Delta_k) > 0 \\ x_{k+1} = x_k & \text{otherwise;} \end{cases} \tag{6}$$

where

$$\lambda_k = \frac{\varphi(x_k,\Delta_k) + \alpha||\partial\varphi(x_k,\Delta_k))||}{||\partial\varphi(x_k,\Delta_k))||^2} \tag{7}$$

and $\mathcal{P}_{\mathcal{T}}$ stands for projection on the set $\mathcal{T}$. The subgradients involved in (6) and (7) are given by

$$\partial\varphi(x,\Delta) = \begin{pmatrix} a^T(\Delta)\mathrm{sgn}(a^T(\Delta)c - y(\Delta)) \\ \frac{a(\Delta)a^T(\Delta)P + Pa(\Delta)a^T(\Delta)}{2\|Pa(\Delta)\|} \end{pmatrix}.$$

This leads to $||\partial\varphi(x_k,\Delta_k))||^2 = 2\|a\|^2$.

Consider first the case when the set $\mathbf{\Delta}$ is finite, then with respect to Algorithm 1, the following result holds.

Theorem 1 *Consider the problem (5) with $\mathbf{\Delta}$ finite. Under the assumption of strong feasibility, let*

$$M = \frac{\|x^0 - x^*\|^2}{\alpha^2}$$

Then, the following results hold

1. *if the Δ_k's are processed cyclically then Algorithm 1 terminates after a number of cycles less or equal than M.*
2. *if the Δ_k's are chosen such that $\varphi(x_k,\Delta_k) = \max_{\Delta\in\mathbf{\Delta}} \varphi(x_k,\Delta)$, then the Algorithm 1 terminates after a number of iterations less or equal than M.*

The proof of this result is omitted for brevity and can be found in [5]. Similar results, applied to probabilistic robust design and LMI optimization problems, can be found in [3, 9].

In the case of $\mathbf{\Delta}$ not finite, we consider experiments Δ with associated pdf $f_\Delta(\Delta)$ such that

$$\forall x \notin \mathcal{D} \qquad \Pr\{\varphi(x,\Delta) > 0\} > 0\ , \tag{8}$$

that is equivalent to a persistent excitation condition in the standard framework of identification.
Under such assumptions the following theorem holds.

Theorem 2 *Under the assumption of strong feasibility, if condition (8) holds then the proposed algorithm terminates at a finite number of iterations with probability one.*

The proof, omitted here for brevity, can be found in [5]. If we choose the initial ellipsoid $x_0 = (c_0, P_0)$ with $P_0 > 0$ sufficiently large and apply the method with $\mathcal{T} = \{P \geq 0\}$, paying no attention to the cost function $\mathrm{Tr}P$, Algorithm 1 terminates in a finite number of steps. Due to the choice of P_0 large we expect that the obtained ellipsoid, parameterized by (c^*, P^*), is a good approximation to the solution of the optimization problem (5). To refine this result, some heuristic considerations can be introduced. The idea is to apply again the above result with $\mu > \mathrm{Tr}P^*$. Heuristic rules to adjust μ in accordance with the results of calculations can be easily derived. However, we can also provide some rigorous results for algorithms explicitely tailored for the optimization problem (5). Below is the simplest result of this sort. We suppose that finding the most violated constraint is not a difficult task.

Algorithm 2 *Let $x_0 = (c_0, P_0) \in \mathcal{T}$, $\mathcal{T} = \{P \geq 0\}$, be an initial estimate. For a point $x_k = (c_k, P_k), P_k \geq 0$ calculate the most violated constraint:*

$$v_k = \varphi(x_k, \Delta_k) = \max_{\Delta \in \mathbf{\Delta}} \varphi(x_k, \Delta)$$

and construct the following recursion

$$\begin{cases} x_{k+1} = \mathcal{P}_{\mathcal{T}}(x_k - \lambda_k \partial\varphi(x_k, \Delta_k)) & \text{if } v_k > 0 \\ x_{k+1} = P_{\mathcal{T}}(x_k - \gamma_k \nabla f(x_k)) & \text{if } v_k \leq 0 \end{cases},$$

where λ_k is defined as in Algorithm 1, with $\alpha = 0$, $P_{\mathcal{T}}$ is the projection on the positive definite matrices cone $\mathcal{T}$,

$$\nabla_P f(x) = I, \quad \nabla_c f(x) = 0,$$

and $\gamma_k > 0$ is a scalar sequence such that

$$\gamma_k \to 0, \quad \sum \gamma_k = \infty.$$

With respect to Algorithm 2, the following result holds.

Theorem 3 *If problem (5) is feasible, then Algorithm 2 converges for $k \to \infty$ to a solution x^* of the problem.*

The proof, omitted here for brevity, can be found in [5]. We remark that Algorithm 2 has some common features with the method proposed in [8]. The need to determine at each step the most violated constraint can be computationally cumbersome. However, we are currently validating other methods, with Δ_k processed cyclically or randomly, that allow to avoid this step.
We remark that, for $\mu = 0$, the center of the ellipsoid obtained by Algorithms 1 and 2 provide an interpolatory estimate in a finite number of iterations. Moreover, an important information about the reliability of the estimate is given by the size (and the shape) of the ellipsoid. This result extends previous works on successive projections methods for set membership identification (see e.g. [1]).

4 Numerical Example

For visualization purposes, we provide here an example with $p = 2$. A larger example, showing in more details the performances of the algorithm, can be found in [5]. We considered a set of regressors parameterized as follows

$$a^T(\Delta) = [\ 1\ \Delta\ ] \tag{9}$$

where Δ is uniformly distributed over the domain $\mathbf{\Delta} = [-1, 1]$. A set of $N = 200$ measurements errors have been generated in the interval $[-\bar{e},\ \bar{e}]$, $\bar{e} = 1$, and the corresponding measurement outputs $y(\Delta)$ have been computed. The results of Algorithm 1 are reported in Figure 1. The dotted lines represent ten iteration steps of Algorithm 1, while the solid lines represent the ellipsoids $\mathcal{E}$ obtained running ten times the algorithm with increasing values of μ. The feasible parameter set $\mathcal{D}$, derived applying the Norton-Mo algorithm for the exact determination of feasible set [7], is also reported.

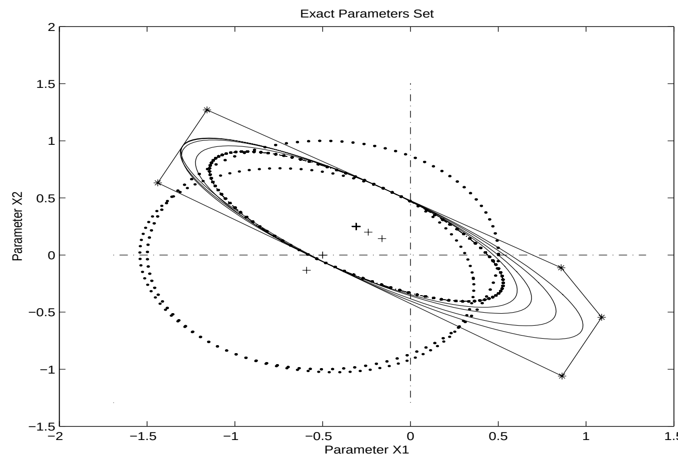

Figure 1: Application of Algorithm 1

References

[1] E. W. Bai, H. Cho, R. Tempo (1998). Convergence Properties of the Membership Set and Noise Model Selection, *Automatica*, Vol. 34, pp. 1245-1249

[2] G. Belforte, T. T. Tay (1993). Two New Estimation Algorithms for Linear Models with Unknown but Bounded Measurement Noise, *IEEE Transactions on Automatic Control*, Vol. 38, No. 8, pp. 1273-1279.

[3] G. Calafiore and B. T. Polyak (2000). Fast algorithms for exact and approximate feasibility of robust LMIs, Proc. of IEEE Conference of Decision and Control, Sydney.

[4] S. P. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan (1994). *Linear Matrix Inequalities in Systems and Control Theory.* Philadelphia: SIAM Publications.

[5] F. Dabbene, P. Gay, B.T. Polyak (2000). Almost Guaranteed Membership Set Identification in the Presence of Outliers, Internal Report DGP01-2000, Politecnico di Torino.

[6] E. Fogel, Y. F. Huang (1982). On the Value of Information in System Identification - Bounded Noise Case, *Automatica*, Vol. 18(2), PP. 229-238.

[7] S. H. Mo, J. P. Norton (1998). Recursive Parameter Bounding Algorithms which Compute Polytope Bounds, *IMACS Congress on Scientific Computation,*, Vol. 2, pp. 477-480.

[8] B. T. Polyak, A General Method for Solving Extremum Problems (1967). *Soviet. Math. Doklady,* Vol. 3, pp. 593-597.

[9] B. T. Polyak, R. Tempo (2000). Probabilistic Robust Design with Linear Quadratic Regulators, Proc. of IEEE Conference of Decision and Control, Sydney.

[10] S. Tarasov, L. Khachiayn, I. Erlikh (1988). The Method of Inscribed Ellipsoids, *Soviet Math. Doklady* Vol. 37, pp. 226-230.

[11] A. Vicino, G. Zappa (1996). Sequential Approximation of Feasible Parameter Sets for Identification with Set Membership Uncertanty, *IEEE Transactions on Automatic Control*, Vol. 41, No. 6, pp. 774-785.