SET ESTIMATION VIA ELLIPSOIDAL APPROXIMATIONS

Ashutosh Sabharwal and Lee C. Potter

Department of Electrical Engineering
The Ohio State University
Columbus OH 43210

ABSTRACT

In most estimation and design problems, there exists more than one solution that satisfies all constraints. In this paper, we address the problem of estimating the *complete set of feasible solutions*. Multiple feasible solutions are frequently encountered in signal restoration, image reconstruction, array processing, system identification and filter design. An estimate of the size of the feasibility set can be utilized to quantitatively evaluate inclusion and effectiveness of added constraints. Further, set estimation can be used to determine a null feasibility set. We compute ellipsoidal approximations to the set of feasible solutions using a new ellipsoid algorithm and the method of analytic centers. Both algorithms admit multiple convex constraint sets with ease. Also, the algorithms provide a solution which is guaranteed to be in the interior of the feasibility set.

1. Motivation

Most estimation and design algorithms can be classified as point estimation schemes. Typically, the algorithms involve optimizing a criterion with or without constraints. Maximum entropy (MEM), minimum mean-square error (MMSE) and maximum likelihood (ML) are examples of the optimality criterion frequently employed in estimating solutions. Choice of the optimality criterion, in many cases, is highly subjective and can be greatly influenced by the tractability of the resulting optimization problem.

In [1], Youla and Webb introduced the estimation problem in a set-theoretic framework. The underlying principle of the set theoretic formulation [2, 3] is to express each desired property of the solution as a set in the appropriate parameter space. The intersection of these property sets is the set of feasible solutions. This formulation discards the notion of a unique optimal solution.

Apart from avoiding an optimality criterion, the set theoretic formulation offers other advantages. The change in set sizes provides a measure of the impact of adding or altering a constraint. Further, a small set of feasible solutions implies that all point estimation schemes will be comparable in their performance under any criterion. Similarly, in design applications, smaller sets of feasible solutions mean tighter designs. Importantly, set estimation can be effectively used to determine null feasibility sets.

Here, we seek to compute a measure of the *uncertainty* inherent in a given feasibility problem. In most practical applications, exact estimation of the feasibility set is an NP-hard problem. However, ellipsoids are cheaply computed and can provide bounds on the feasibility set. The approximation gives a "trust region" for a design or estimate. Further, we obtain, as a byproduct, an estimate of a center for the feasibility set, in the spirit of Bayesian or minimax estimation.

2. Problem Formulation

Let $\mathcal{K}_1, \ldots, \mathcal{K}_M$ represent closed convex subsets of $\mathbb{R}^n$. We assume that set $\mathcal{K}_1$ is bounded. The intersection $\cap_{i=1}^M \mathcal{K}_i$ represents the set of feasible solutions and is denoted by $\mathcal{F}$. The feasibility set is closed and convex since $\mathcal{K}_i$, $i = \{1, \ldots, M\}$ is closed and convex. Additionally, $\mathcal{F}$ is bounded since $\mathcal{K}_1$ is assumed to be bounded. Thus, there exists a unique ellipsoid of minimal volume which contains $\mathcal{F}$ [4]. We seek to estimate the minimum volume ellipsoid which bounds the feasibility set $\mathcal{F}$.

3. Ellipsoid Algorithm

In this section, we present an ellipsoid algorithm to estimate feasibility sets. The ellipsoid method was used by Khachiyan [5] to prove the polynomial solvability of linear programming problems; other applications are developed in [6]. We derive optimal cutting halfspaces. Polynomial time convergence is proved and a characterization of the limit ellipsoid is given.

The idea of the ellipsoid algorithm is simple. The algorithm initializes with a ball around the bounded constraint set, $\mathcal{K}_1$. At every step, a cut is computed and used to obtain a reduced volume ellipsoidal approximation. The process is repeated until no further reduction in volume is possible. The idea of computing a cut is similar to having a separation oracle [4]. Both the separation oracle and the optimal cut provide a halfspace which serves to reduce the volume of the ellipsoid bounding the feasible set. A separation oracle provides a separating hyperplane for the feasibility set. On the other hand, the cut computed in the proposed ellipsoidal algorithm is not only a separating halfspace but is also a support halfspace of one of the constituent constraint sets, $\mathcal{K}_1, \ldots, \mathcal{K}_M$.

3.1. Optimal Cuts

Definition 1 Let $\mathcal{K}$ be a closed convex subset of $\mathbb{R}^n$. Also, let $\mathcal{E} = \{x : (x-c)^T P^{-1} (x-c) \leq 1\}$ denote an ellipsoid in $\mathbb{R}^n$. A support halfspace, $H = \{x : a^T x \leq \beta\}$, of $\mathcal{K}$ is called the *optimal cut* for $\mathcal{E}$ if the volume in the intersection, $\mathcal{E} \cap H$, is minimized among all support halfspaces of $\mathcal{K}$.

The following propositions allow the computation of the optimal cut provided by a closed convex set $\mathcal{K}$, with respect

 0-7803-2431-5/95 $4.00

to an ellipsoid $\mathcal{E} \subset \mathbb{R}^n$. The basic idea of finding an optimal cut is as follows. Consider the affine transformation $\hat{x} = \mathcal{J}^{-1}(x - c)$ where $P = \mathcal{J}\mathcal{J}^T$. In the tranformed coordinate system, the quantity $\alpha = (a^T c - \beta)/\sqrt{a^T P a}$ represents the algebraic distance of a cutting halfspace, $\mathcal{J}^{-1}(H - c)$, from the origin. And the cut, $H = \{x : a^t x \leq \beta\}$, for which α is maximized, provides minimal volume in the intersection $\mathcal{E} \cap H$ [7].

Proposition 2 *Let $\mathcal{E}$ be an ellipsoid with center c and defined by the matrix $P > 0$. Also, assume that $\mathcal{K}$ is a closed convex set. If $c \notin \mathcal{K}$, then the optimal cut, H, provided by the set $\mathcal{K}$ is given by*

$$\begin{aligned} H &= \{x : a^T x \leq \beta\} \\ a &= \mathcal{J}^{-T}(-a_p) \\ \beta &= -\|a_p\|^2 + a^T c \end{aligned}$$

and

$$a_p = \Pi_{\mathcal{J}^{-1}(\mathcal{K}-c)}(0) \text{ and } P = \mathcal{J}\mathcal{J}^T$$

where $\Pi_{\mathcal{J}^{-1}(\mathcal{K}-c)}$ represents the nearest point projection onto the closed convex set $\mathcal{J}^{-1}(\mathcal{K} - c)$.

For proofs of the results, see [8].

For the case when the the center of the ellipsoid lies in the convex set $\mathcal{K}$, we introduce the following definition. Assume that $\mathcal{M}$ is a closed convex set such that the origin belongs to the set. We define $q(\mathcal{M})$ as a minimum norm element in the boundary of the set $\mathcal{M}$. In other words, $q(\mathcal{M})$ can be viewed as an output of the following constrained program

$$q(\mathcal{M}) = \arg \inf_{x \in \partial\mathcal{M}} \|x\|$$

We call q the *minimum-norm boundary function*.

Proposition 3 *Let $\mathcal{E}$ be an ellipsoid with center c and defined by the matrix $P > 0$. Also, assume that $\mathcal{K}$ is a closed convex set. If $c \in \mathring{\mathcal{K}}$, then an optimal cut, H, provided by the set $\mathcal{K}$ is given by*

$$\begin{aligned} H &= \{x : a^T x \leq \beta\} \\ a &= \mathcal{J}^{-T}(a_p) \\ \beta &= \|a_p\|^2 + a^T c \end{aligned}$$

where

$$a_p = q(\mathcal{J}^{-1}(\mathcal{K} - c)) \text{ and } P = \mathcal{J}\mathcal{J}^T$$

If $c \in \partial\mathcal{K}$, then $a = \phi(c)$ and $\beta = a^T c$, where $\phi(x)$ is an outer normal to $\mathcal{K}$ at $x \in \partial\mathcal{K}$.

3.2. Ellipsoid Algorithm

In this section, we obtain an ellipsoid algorithm using the optimal cuts from Section 3.1. The closed convex constraints $\mathcal{K}_1, \ldots, \mathcal{K}_M$ are known via their nearest-point projection and it is assumed that the minimum-norm boundary function is computable for each constraint set. The ellipsoid algorithm initiates with a ball containing the bounded set $\mathcal{K}_1$. At every step, M optimal cuts are computed, one for each of the constraint sets $\mathcal{K}_1, \ldots, \mathcal{K}_M$. The cut which offers the maximal volume reduction is chosen from the calculated optimal cuts. Using the chosen cut, an updated ellipsoidal approximation is obtained. The algorithm terminates once the update results in no further reduction in the volume of the ellipsoid. The algorithm can be written as follows.

Ellipsoid Algorithm :

1. Initialize the algorithm with an ellipsoid $\mathcal{E}_0$ with center $c_0 = 0$ and $\mathcal{K}_1 \subset \mathcal{E}_0$. Set $k = 0$.
2. Find the optimal cuts $H_{\mathcal{K}_1}, H_{\mathcal{K}_2}, \ldots, H_{\mathcal{K}_M}$ for the sets $\mathcal{K}_1, \ldots, \mathcal{K}_M$ with respect to the ellipsoid $\mathcal{E}_k$, using Propositions (2) and (3). Choose H as the cut providing the maximum volume reduction.
3. If *termination condition* met then STOP else goto step (4).
4. Compute $\mathcal{E}_{k+1}$ as the minimum volume ellipsoid bounding the intersection $\mathcal{E}_k \cap H$ (the closed form expression is given in [7, 8]).
5. $k = k + 1$. Goto step (2).

The *termination condition* at step (3) is evaluated as follows. Let $H = \{x : a^T x \leq \beta\}$ be the halfspace chosen at step (2). The termination condition is met if

$$\frac{a^T c_k - \beta}{\sqrt{a^T P_k a}} < -\frac{1}{n} + \gamma$$

where the constant γ satisfies $0 < \gamma < 1/n$. The vector c_k represents the center of the ellipsoid $\mathcal{E}_k$ defined by the positive-definite matrix P_k. Figure 1 illustrates a typical step in the proposed ellipsoid algorithm.

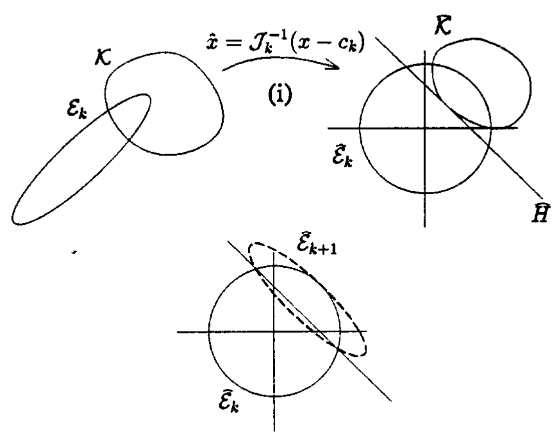

Figure 1: Typical steps in the ellipsoid algorithm : (i) Transform current approximation $\mathcal{E}_k$ into a unit hypersphere $\widehat{\mathcal{E}}_{k+1}$ and calculate optimal cut for $\widehat{\mathcal{K}}$ and (ii) calculate new approximation $\widehat{\mathcal{E}}_{k+1}$, assuming that $\widehat{H}$ was chosen in step (2).

4. Analysis of Ellipsoid Algorithm

Lemma 4 *For the ellipsoid algorithm of Section 3.2*

1. *If $V_k = vol(\mathcal{E}_k)$, then $V_k < V_{k-1}$.*
2. *$\mathcal{F} \subset \mathcal{E}_k$, where $\mathcal{F}$ is the bounded feasibility set.*

Since $\text{vol}(\mathcal{E}_k) \geq 0$ and the volumes of the ellipsoids $\text{vol}(\mathcal{E}_k)$ form a strictly decreasing sequence, hence the volumes converge to some real number $V \geq 0$. Thus, the proposed ellipsoid algorithm provides a strictly decreasing sequence of ellipsoidal approximations containing the feasibility set. In other words, the quadratic approximations generated in the form of ellipsoids at every step improve as the algorithm progresses in time.

Theorem 5 *Let $\mathcal{K}_1, \ldots, \mathcal{K}_M$ be closed and bounded convex subsets of $\mathbb{R}^n$. Also, assume that the set $\mathcal{K}_1$ is bounded. Let $\mathcal{F} = \cap_{i=1}^{M} \mathcal{K}_i$ denote the bounded, closed and convex feasibility set. Then the sequence of the ellipsoids $\mathcal{E}_k$, generated by the ellipsoid algorithm, terminates in a finite number of steps to an ellipsoid $\mathcal{E}$ such that $\mathcal{F} \subset \mathcal{E}$. The number of steps is bounded by $\log(V_m/V_0)\log(\delta)^{-1}$, where $V_0 = vol(\mathcal{E}_0)$ and V_m represents the volume of the unique minimum volume ellipsoid bounding $\mathcal{F}$. Also, the center of the ellipsoid lies in the interior of the set $\mathcal{F}$.*

It is important to be able to estimate the size of the ellipsoid computed using the proposed ellipsoid algorithm. To this end, we modify a result provided in [4].

Theorem 6 *Let $-1/n + \gamma = -1/n_\gamma$ where $0 < \gamma < 1/n$, and $\mathcal{E}$ is the limit ellipsoid. Then the following holds*

$$\mathcal{E}_{in} = \frac{1}{\sqrt{n}(n_\gamma + 1)} \mathcal{E} \subseteq \mathcal{F} \subseteq \mathcal{E}$$

where $\mathcal{F} \subset \mathbb{R}^n$ is the feasibility set.

Note that the proposed ellipsoid algorithm does not compute an optimal cut for the set $\mathcal{F}$; instead, the cut is chosen from the cuts computed for the individual sets $\mathcal{K}_1, \ldots, \mathcal{K}_M$. If, instead, an optimal cut for $\mathcal{F}$ is computed at step (2) of the algorithm, we refer to the resulting algorithm as the *optimal algorithm.* To obtain an optimal cut for $\mathcal{F}$ requires computation of $\Pi_{\mathcal{F}}(0)$ or $q(\mathcal{F})$ which, in practice, may not be easily computable. On the positive side, in Corollary (7), we demonstrate that the proposed algorithm performs as well as the optimal algorithm, in the sense described below.

Theorem (6) provides a criterion to evaluate the performance of the proposed ellipsoid algorithm. Let X and Y be two algorithms which provide the two-sided approximations $(r_x\mathcal{E}_x, \mathcal{E}_x)$ and $(r_y\mathcal{E}_y, \mathcal{E}_y)$, respectively, of a bounded and closed convex set. We say algorithm X performs *better* than algorithm Y if $r_x < r_y$. Based on this criterion, we obtain the following result.

Corollary 7 *Let $\mathcal{F} = \cap_{i=1}^{M} \mathcal{K}_i \subset \mathbb{R}^n$ be a bounded and closed convex set where $\mathcal{K}_i \subset \mathbb{R}^n$, $i = 1, \ldots, M$ is closed and convex. Also, $\mathcal{K}_1$ is assumed to be bounded. Then the optimal algorithm performs no better than the ellipsoid algorithm proposed in Section 3.2, in the sense defined above.*

5. Method of Analytic Centers

The ellipsoid algorithm proposed in the previous section seeks to approximate the feasibility set $\mathcal{F}$ from outside. As a byproduct, both an inner approximation to $\mathcal{F}$ and a point in $\mathcal{F}$ are obtained. In contrast, in this section, we proceed by estimating the *analytic center* of the set $\mathcal{F}$. Both inner and outer ellipsoidal approximation to $\mathcal{F}$ are obtained as a consequence. We call this a method of analytic centers (MAC).

We use the widely studied [9, 10] concept of the analytic center of a linear matrix inequality (LMI). An LMI is defined as

$$F(x) \triangleq F_0 + \sum_{i=1}^{n} x_i F_i > 0 \tag{1}$$

where $x \in \mathbb{R}^n$ is the variable and the symmetric matrices $F_i = F_i^T \in \mathbb{R}^{p \times p}$, $i = 0, \ldots, n$, represent convex constraints.

Many commonly encountered convex constraints can be represented as an LMI [10]. The feasibility set $\mathcal{F}$ is given by $\{x \in \mathbb{R}^n : F(x) > 0\}$. The barrier function for $\mathcal{F}$ is defined as [10]

$$\phi(x) \triangleq \begin{cases} \log\det F(x)^{-1} & x \in \mathcal{F} \\ \infty & x \notin \mathcal{F} \end{cases}$$

and has a unique minimizer x^* referred to as the *analytic center* of the affine matrix inequality $F(x) > 0$. Analytic center provides inner and outer ellipsoidal approximations. As proved in [9], there exists a pair of inner and outer ellipsoids centered at x^* and with shape determined by $H(x^*)$, the Hessian of ϕ at x^*. The following approximations hold for $\mathcal{F}$

$$\frac{1}{n(n-1)} \mathcal{E} \subseteq \mathcal{F} \subseteq \mathcal{E}$$

where $\mathcal{E} \triangleq \left\{x \in \mathbb{R}^n : (x - x^*)^T H(x^*)(x - x^*) < 1\right\}$.

Finally, to compute the the analytic center, x^*, a modified Newton-Raphson iteration scheme is given in [9].

6. Comparisons

The two algorithms differ in their basic philosophy. While the ellipsoid algorithm (EA) proceeds by estimating the set and provides a point estimate as a consequence, the method of analytic centers (MAC) estimates a point in the set and provides set estimates as a consequence.

Both the algorithms provide a pair of ellipsoids $(r\mathcal{E}, \mathcal{E})$ such that

$$r\mathcal{E} \subseteq \mathcal{F} \subseteq \mathcal{E}$$

where

$$r = r_{EA} = \frac{1}{\sqrt{n}(n_\gamma + 1)}, \text{ for an ellipsoid algorithm}$$

$$r = r_{MAC} = \frac{1}{n(n-1)}, \text{ for the MAC}$$

For small γ, $r_{EA} < r_{MAC}$. Thus, the EA performs better than the MAC in the sense described in Section 4.

Both the algorithms are computationally intensive. The EA suffers from slow convergence, if the size of the feasibility set is small and the number of dimensions is large. Also, computation of the projection operator and the minimum-norm boundary function is expensive for quadratic convex sets. On the other hand, the MAC requires computation of the Hessian, gradient and the inverse of the Hessian. Additionally, the sizes of the matrices involved grows linearly with the number of constraints.

7. NMR Porosimetry

Finally, to illustrate application of the two set estimation techniques, we consider a linear inverse problem arising in nuclear magnetic resonance (NMR) porosity measurements. The NMR spin-lattice relaxation technique recently has been applied to the study of porous media, such as oil-bearing rock formations [11] and microfiltration type membrane filters [12]. The measured magnetization signal can be described as a Fredholm integral equation of the first kind,

$$M(t) = M_0 \int_0^\infty \left(1 - 2e^{-t/T_1}\right) f(T_1) dT_1$$

where $f(T_1)$ is the pore volume distribution function. Four properties of the solution are expressed as convex constraint sets. Specifically, we consider candidate distributions which are nonnegative, are zero outside the interval $[10^{-3}, 10]$, and have a fixed L_1 norm. The fourth constraint set is consistency with the magnetization measurements: *i.e.*, $\|Af - b\| \leq 0.01$, where b is the sampled data vector consisting of 32 samples of $M(t)$, logarithmically spaced from 0.5 msec to 15 sec.

We compute the constrained least-squares solution [3], the regularized minimum-norm solution and the point estimates provided by EA and MAC. The result is plotted on a logarithmic scale in Figure 2. It is evident that incorporation of the constraints greatly improves the reconstructed pore size distribution. Also, the inner ellipsoid has nonzero volume, indicating that there exist distributions consistent with all constraints. Moreover, the outer approximating ellipsoid has a very small volume, implying that all constrained point estimators will provide similar performance for this data.

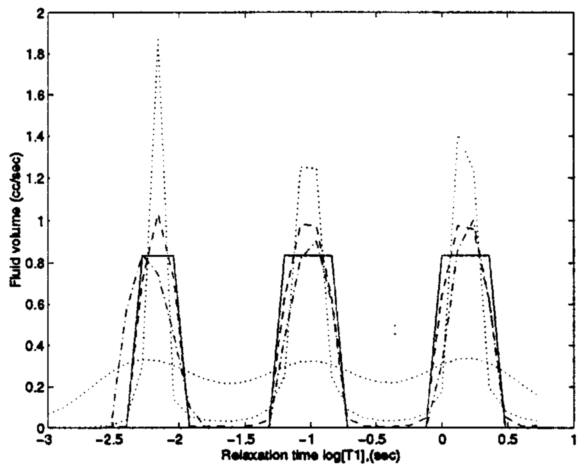

Figure 2: NMR recovery of pore size distribution. — True distribution, - - - ellipsoid algorithm, · · ·(upper) MAC solution, —·— constrained minimum norm least squares solution and · · ·(lower) regularized minimum-norm solution

8. References

[1] D. C. Youla and H. Webb, "Image restoration by the method of convex projections : Part 1 – Theory," *IEEE Trans. Med. Imag.*, vol. MI-1, pp. 81–94, Oct. 1982.

[2] P. L. Combettes, "Foundations of set estimation," *Proceedings of the IEEE*, vol. 81, no. 2, pp. 182–208, 1993.

[3] L. C. Potter and K. S. Arun, "A dual approach to linear inverse problems with convex constraints," *SIAM J. Control and Optimization*, vol. 31, pp. 1080–1092, July 1993.

[4] M. Grótschela, L. Lovász, and A. Schrijver, *Geometric Algorithms and Combinatorial Optimization.* Springer–Verlag, 1988.

[5] H. Konig and D. Pallashke, "On Khachian's algorithm and minimal ellipsoids," *Numer. Math.*, vol. 36, pp. 211–223, 1981.

[6] J. R. Deller and F. S. Odeh, "Adaptive set-membership identification in O(m) time for linear-in-parameters models," *IEEE Transactions on Signal Processing*, pp. 1906–1924, May 1993.

[7] D. Goldfarb and M. J. Todd, "Modifications and implementation of the ellipsoid algorithm for linear programming," *Mathematical Programming*, vol. 23, pp. 1–19, 1982.

[8] A. Sabharwal, "Constrained signal reconstruction," Master's thesis, The Ohio State University, 1994.

[9] Y. Nesterov and A. Nemirovsky, *Interior point polynomial methods in convex programming : Theory and applications.* SIAM, 1993.

[10] S. Boyd and L. E. Ghaoui, "Method of centers for minimizing generalized eigenvalues," *Linear Algebra and Applications, special issue on Numerical Linear Algebra Methods in Control, Signals and Systems*, vol. 188, pp. 63–111, July 1993.

[11] J. Howard, W. Kenyon, and C. Straley, "Proton magnetic resonance and pore size variations in reservoir sandstones," in *65th Annual Technical Conference, SPE*, (New Orleans), 1990.

[12] C. L. Glaves and D. M. Smith, "Membrane pore structure analysis via NMR spin-lattice relaxation experiments," *J. Membrane Science*, vol. 46, pp. 167–184, 1989.