A Randomized Method for Solving Semidefinite Programs

B. T. Polyak   P. S. Shcherbakov

Trapeznikov Institute of Control Science,
Russian Academy of Sciences, Moscow, Russia
(e-mail: {boris,sherba}@ipu.ru)

Abstract: Proposed is a novel iterative method for solving semidefinite programs. It exploits the ideology of cutting hyperplane through the center of mass of a convex body. To estimate the center of mass, we use a random walk technique known as the Hit-and-Run algorithm. The results of numerical simulations are compared to those obtained with presently available approaches. Robust versions of the method are considered, where the coefficient matrices contain norm-bounded uncertainties.

Keywords: linear matrix inequalities, optimization, random walk, center of mass.

1. INTRODUCTION AND OVERALL SCHEME OF THE METHOD

In this section we formulate the problem under consideration and present a brief schematic description of the main ideas underlying our method.

Considered is the standard semidefinite program (SDP) of the form

$$\min c^\mathsf{T} x \quad \text{s.t.} \quad A(x) \doteq A_0 + \sum_{i=1}^{n} x_i A_i \,\leq\, 0, \qquad (1)$$

where $c \in \mathbb{R}^n$ and $A_i \in \mathbb{R}^{m\times m}$, $i = 0,\dots,n$, are known symmetric matrices; the notation $A \leq 0$ stands for negative semidefiniteness of the matrix A. The constraint inequality in (1) is called a *linear matrix inequality* (LMI), and the convex set

$$D_{feas} = \{x \in \mathbb{R}^n : A(x) \leq 0\}$$

is referred to as the *feasible domain* of this LMI.

This problem is known to be one of the key problems in the theory of linear matrix inequalities Boyd et al. (1994). It has numerous applications in various fields of system theory and control, and at present there exist efficient solution techniques based on interior-point methods; e.g., see Nesterov and Nemirovskii (1994).

Inspired by the recent results in the rapidly developing area of randomized methods in control system analysis and design Tempo et al. (2005), we propose a novel approach to solving problem (1), which is based on totally different ideas. The iterative method that we developed leans on *random walks*, estimation of the *center of mass* of a convex set—the feasible domain D_{feas}, and uses the new notion of *boundary oracle* for LMIs.

The cornerstone components of our approach are

(a) the cutting hyperplane ideology, which is used to compute iteratively a sequence of embedded subdomains $D_{k+1} \subset D_k \subset D_{feas}$ and monotonically decrease the value of the objective function $f(x) = c^\mathsf{T} x$;

(b) the so-called Hit-and-Run (HR) algorithm for estimating the center of mass of D_k required in item **(a)** above;

(c) a boundary oracle which is needed for implementation of the HR-algorithm.

Specifically, let $D_k \subset D_{feas}$ be the domain obtained at the kth step of the iterative method under consideration. For simplicity, it is assumed that D_{feas} is bounded in order to guarantee the boundedness of D_k. Using HR-algorithm, we generate N_{hr} random points distributed approximately uniformly on D_k and adopt their average x^k as an estimate of the center of mass of D_k. The point x^k might as well be taken to compute the current estimate $f^k = c^\mathsf{T} x^k$ of the objective function. Next, the hyperplane

$$H_k = \{x \in \mathbb{R}^n\colon\ c^\mathsf{T}(x - x^k) = 0\}$$

is drawn to cut off the "idle" portion of D_k thus reducing it to

$$D_{k+1} = \{x \in \mathbb{R}^n\colon x \in D_k,\ c^\mathsf{T} x \leq c^\mathsf{T} x^k\} \subset D_k,$$

and the process is repeated with the set D_{k+1}. In other words, the convex set D_{k+1} is bounded by the LMI constraints in (1) and the half-space $\{x \in \mathbb{R}^n :\ c^\mathsf{T}(x - x^k) \leq 0\}$ defined by the hyperplane H_k. Schematically, the behavior of the method is represented in Fig. 1 for the two-dimensional case where the vector c is taken in the form $c = (0,\ 1)^\mathsf{T}$.

Under the assumption that x^k is a reasonably accurate estimate of the true center of mass, the lemma in Radon (1916) on the measures of symmetry of convex bodies is used to obtain an estimate on the *guaranteed rate* of decrease in the objective function. Thus, the method is expected to have a geometric convergence rate. Note that by no means do we intend to estimate the optimal point $x^* = \arg\min_{x\in D_{feas}} c^\mathsf{T} x$, but rather evaluate the optimal value f^* of the objective function.

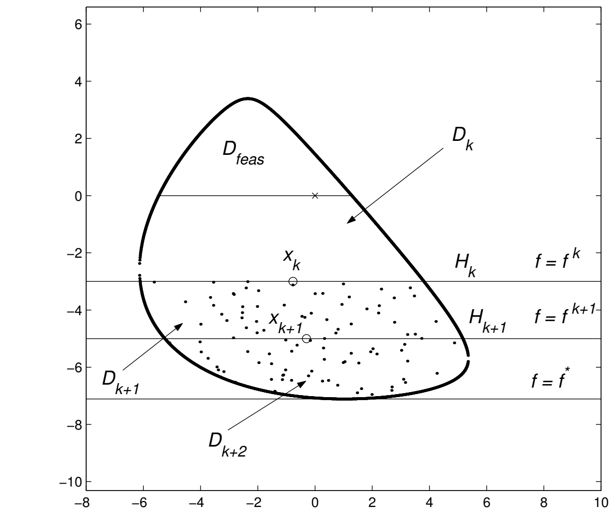

Fig. 1. Schematic representation of the method.

2. CORNERSTONES

We now present the techniques and results underlying each of the items **(a)**–**(c)** above.

2.1 Cutting hyperplane

The implementation of item **(a)** and the guaranteed reduction of the value of the objective function is based on the following lemma.

Lemma 1 Radon (1916). *Let $D \subset \mathbb{R}^n$ be a convex bounded body and $g \in D$ be its center of mass. Denote by H an arbitrary $(n-1)$-dimensional hyperplane through g, and let H_1 and H_2 be the two hyperplanes supporting to D and parallel to H. Denote by*

$$r(H) \doteq \frac{\min\{\mathbf{dist}(H,H_1),\, \mathbf{dist}(H,H_2)\}}{\max\{\mathbf{dist}(H,H_1),\, \mathbf{dist}(H,H_2)\}}$$

the ratio of the distances from H to H_1 and H_2, respectively. Then

$$\min_H r(H) \geq \frac{1}{n}.$$

As applied to the setup in this paper, let H_1 and H denote the two hyperplanes through the two successive iterations x^k and x^{k+1} of the method, and let H_2 be the supporting hyperplane through the optimal point x^*. Assuming that the *exact* value of the center of mass is known, the following estimate is readily available:

$$f^{k+1} - f^* \leq \varkappa(f^k - f^*), \quad \varkappa = \frac{n}{n+1},$$

where f^* is the optimal value of the objective function and $f^k = c^\mathsf{T} x^k$ is the estimate obtained at the kth iteration.

Notably, to the best of our knowledge, this result has never been used in optimization, in contrast to similar results on the guaranteed *volumetric* reduction, which are typical to various modifications of the ellipsoid method.

2.2 Hit-and-Run algorithm

We now describe the Hit-and-Run algorithm, an efficient randomized technique exploited in this paper in the estimation of the center of mass of the sets D_k. This random walk algorithm originally proposed in Smith (1984) is simple to describe. It applies to a bounded convex body $D \in \mathbb{R}^n$ and returns a random point z having approximately uniform distribution on D. The arithmetic mean of this distribution is then adopted as an estimate of the center of mass.

Specifically, an initial point z^0 in the interior of D is selected, and let z^j be the point obtained at the jth step of the algorithm. A random direction y is generated (say, in the form $\xi/\|\xi\|$, where ξ is a Gaussian random vector with zero mean and identity covariance matrix, and $\|\cdot\|$ is the euclidean vector norm). The 1D-line $z^j + \lambda y$ is considered and the points $\underline{z}^j$ and $\overline{z}^i$ of its intersection with the boundary of D are computed. The next-step point z^{j+1} is then generated randomly uniformly on the chord $[\underline{z}^j, \overline{z}^j]$.

In Smith (1984); Lovász (1999) it has been shown that the sequence of random vectors $\{z^i\}_1^{N_{hr}}$ generated in such a way forms a discrete Markov chain having the property of uniform ergodicity. In other words, the distribution of the random vector z^i tends to the limiting uniform distribution on D with geometric rate; this property is referred to as *fast mixing*. The mixing rate depends on the shape of the set D and on the "position" of the initial point. The best results are obtained if the distribution of z^0 is close to the uniform (the so-called *warm start* of the HR-algorithm), and the set D is *isotropic*, i.e., it has "equal dimensions" along all directions.

There exist other random walk techniques; the HR-algorithm is chosen here because it is simple to implement, produces practically reasonable approximations to the uniform distribution, requires minimum a priori information about the set D, and can also be applied to nonconvex sets under mild extra assumptions.

It should be noted that the HR-algorithm has been first applied to solving convex optimization problems in Bertsimas and Vempala (2004), however the overall method in that paper essentially differs from the one proposed here.

2.3 Boundary oracle

To perform the HR-algorithm over D_k, we need to efficiently compute the intersections of a 1D-line and the boundary of D_k, which is accomplished via use of the semidefinite boundary oracle developed in Polyak and Shcherbakov (2006). The core of this oracle is the lemma below.

Lemma 2 Polyak and Shcherbakov (2006). *Let $A < 0$ and $B = B^\mathsf{T}$, then the minimal and the maximal values of the parameter $\lambda \in \mathbb{R}$ retaining the negative definiteness of the matrix $A + \lambda B$ are given by*

$$\underline{\lambda} = \begin{cases} \max\limits_{\lambda_i<0} \lambda_i, & \\ -\infty, & \textit{if all} \quad \lambda_i > 0; \end{cases} \tag{2}$$

and

$$\overline{\lambda} = \begin{cases} \min\limits_{\lambda_i>0} \lambda_i, & \\ +\infty, & \textit{if all} \quad \lambda_i < 0, \end{cases} \tag{3}$$