

Using Selective Orthonormalization to Update the Analytic Center After the Addition of Multiple Cuts¹

John E. Mitchell²

Srinivasan Ramaswamy³

October 18, 2002

Abstract

We study the issue of updating the analytic center after multiple cutting planes have been added through the analytic center of the current polytope. This is an important issue that arises at every ‘stage’ in a cutting plane algorithm. If $q \leq n$ cuts are to be added, we show that we can use a ‘Selective Orthonormalization’ procedure to modify the cuts before adding them — it is then easy to identify a direction for an affine step into the interior of the new polytope, and the next analytic center is then found in $O(q \log q)$ Newton steps. Further, we show that multiple cut variants with selective orthonormalization of standard interior point cutting plane algorithms have the same complexity as the original algorithms.

Key words: Cutting plane, analytic center, Selective Orthonormalization, linear programming, convex programming.

¹AMS 1991 subject classification. Primary 90C05. Secondary 90C25, 65K05.

²Department of Mathematical Sciences, Rensselaer Polytechnic Institute, Troy, NY 12180. Research partially supported by ONR Grant number N00014-94-1-0391, by NSF Grant number CCR-9901822, and by a grant from the Dutch NWO and Delft University of Technology for the 1997-98 academic year, while visiting ITS/TWI/SSOR at Delft University of Technology.

³Derivatives Strategy, JPMorgan, 270 Park Ave., New York, NY

1 Introduction

In this paper, we consider the problem of updating the analytic center when multiple cutting planes have been added. This is an issue that arises in recent attempts to solve the *convex feasibility problem* (see, e.g., Ref 1). This is the problem of finding a point in a given convex set $S \subseteq \mathbb{R}^n$, or establishing that the set must be empty. A related problem is the *separation problem* (Ref. 1) — given $x \in \mathbb{R}^n$, verify that $x \in S$ or find $a \in \mathbb{R}^n$ such that $a^T y \geq a^T x$ for every $x \in S$. If the convex set is such that the separation problem is solvable, then an *oracle* is said to exist for the feasibility problem. Many oracle based cutting plane methods have been designed for the feasibility problem, including the ellipsoid algorithm (Refs. 1–2).

We are interested in interior point oracle based cutting plane methods, a subject of much recent research interest. Roughly speaking, the philosophy of these methods is as follows: we begin with a polytope that is large enough to contain the set S . At each iteration, a certain well-defined point in the interior of the polytope is chosen as a test point. If this point is not in S , then the oracle returns a cut that is then added to the definition of the polytope. This is continued until a feasible point is found, or until it can be established that the set S is empty. Algorithms of this type include those in Refs. 3, 4, and 5. The test point commonly used in interior point approaches is the analytic center (Refs. 3, 4), initially proposed in Ref. 6, although the volumetric center is also an important point from a polynomial complexity point of view. It is interesting to note that the only polynomial algorithm based on analytic centers originates in Ref. 4, which adds as well as deletes cuts. The algorithm studied by Goffin *et al.*, which has been the subject of much study lately (Refs. 7, 8, 9, 10), is only fully polynomial, with the proofs of complexity using an inequality from Ref. 11. In Ref. 12, the authors have developed a long-step polynomial interior point method for solving convex optimization problems by extending Ref. 4. The analysis of interior point cutting plane methods has been extended to homogenized self-dual and homogeneous cutting plane methods in Refs. 13, 14, 15, and 16. For recent surveys of theoretical results in interior point cutting plane methods, see Refs. 17 and 18.

To further emphasize the importance of cutting plane methods, we mention applications to linear programming, convex programming and combinatorial optimization. Interior point cutting plane algorithms have been used to solve stochastic programming problems (Ref. 19), multicommodity network flow problems (Ref. 20), and integer programming problems (Ref. 21), as well as other forms of convex optimization problems (Refs. 22, 23, 24, 25, and 26). For some classes of linear ordering problems, a cutting plane scheme that combines an interior point method and a simplex

method has been shown to be up to ten times faster than one that just uses either of the methods individually (Ref. 27). For some max cut problems, an interior point cutting plane method has been shown to considerably outperform a simplex cutting plane method (Ref. 28).

From the viewpoint of practical efficiency, it is necessary to do two things — (i) add the cut as strongly as possible (i.e., through the current point, if not deeper), and (ii) add multiple cuts simultaneously (it is assumed that the oracle is capable of returning multiple cuts). In Ref. 3, it was shown that it is possible to add a cut right through the current point, and regain the analytic center in $O(1)$ iterations. Ref. 7 analyzed the column generation algorithm of Ref. 3 with multiple cuts, and showed that the total number of Newton steps required was of roughly the same order as the number of cuts needed. Goffin and Vial (Ref. 10) have proposed a method for finding a new analytic center after the addition of multiple cuts that requires the solution of a nonlinear programming problem, of the size of the number of added cuts. This approach is able to recover the new center in $O(q \log(q))$ Newton steps, where q denotes the number of added cuts. For homogeneous problems, the situation where all the normals of the added constraints make pairwise acute angles in a certain metric was considered in Ref. 16. This condition is similar to the one we analyze in §4.

In this paper, we propose a procedure that also finds a new approximate analytic center in $O(q \log(q))$ Newton steps, but it does not require the solution of a nonlinear programming problem as a subproblem. It does have the drawback of perhaps weakening the cutting planes. The modification of the cuts is to make them orthogonal in some metric, if necessary. In interior point cutting plane algorithms for combinatorial optimization problems, it appears useful to add orthogonal constraints in many situations (Ref. 21). Typically, these constraints are sparse, and distinct constraints correspond to different subsets of the variables. This has one immediate advantage: it helps numerically by reducing the density of the product $A^T A$.

The organization of this paper is as follows. In §2 we introduce preliminaries. If the cuts to be added are linearly independent, then there exists a feasible direction that is a positive combination of the normals (after scaling by the inverse of the Hessian). Since, however, the weights for each normal may be hard to find, in §3 we provide a ‘Selective Gram-Schmidt Orthonormalization’ procedure — we take the cuts returned by the oracle and generate an equal number of new valid cuts with certain nice properties, particularly that any positive combination is a feasible direction. In §4 we provide a technique for updating the analytic center of the new polytope, and discuss its complexity. The implications of our results for the algorithms of Refs.

7 and 4 are discussed in §5. In particular, we show that a multiple cut variant of the algorithm of Ref. 4 has the same complexity as the original algorithm, provided selective orthonormalization is used. We offer our conclusions in §6.

2 Preliminaries

First we introduce some notation. Let the current polytope be defined by

$$(P) \quad Ax \geq b$$

where $x \in \mathbb{R}^n$, $A \in \mathbb{R}^{m \times n}$, and $m \geq n$. Define $s := Ax - b$. Clearly we want $s \geq 0$. We use P to refer to the polytope defined by the system of inequalities (P). The dual problem associated with (P) is

$$\begin{aligned} \min \quad & b^T y \\ \text{subject to} \quad & A^T y = 0 \\ & y \geq 0. \end{aligned} \quad (D)$$

Our analysis will use the *primal-dual potential function*:

$$\Phi(y, s) := y^T s - m - \sum_{i=1}^m \ln(s_i y_i). \quad (1)$$

We can regard this as a function of x , since $s = Ax - b$. At the analytic center, $\Phi(y, s) = 0$, and this is the minimum value for the potential function. For a fixed y , the gradient and Hessian (denoted by H) of Φ with respect to x are given by

$$\nabla \Phi_x(y, s) = y - \sum_{i=1}^m \frac{a_i}{s_i} = -A^T S^{-1} e \quad (2)$$

and

$$H(x) = \nabla_{xx}^2 \Phi(y, s) = \sum_{i=1}^m \frac{a_i a_i^T}{s_i^2} = A^T S^{-2} A \quad (3)$$

respectively. (Throughout this paper, e will denote the vector of all ones of appropriate dimension. Also, if a symbol in lower case denotes a vector, then the upper case symbol will denote the diagonal matrix with the vector along the diagonal.)

We define a primal-dual measure of centrality:

$$\delta(y, s) := \|e - Sy\|. \quad (4)$$

An iterate is *approximately centered* if $\delta(y, s) < \theta$, where we will specify θ later. It will satisfy $0 < \theta < 1$. If an iterate is approximately centered, then the terms $y_i s_i$ can immediately be bounded:

$$1 - \theta \leq y_i s_i \leq 1 + \theta \quad (5)$$

for each component i .

A primal-dual interior point method can be used to find an approximate analytic center for any $\theta > 0$. Each iteration requires the calculation of a primal-dual Newton step. The performance of a primal-dual method with an appropriate step length can be described using the following three theorems. The result are standard and proofs can be found in, for example, Ref. 29.

Theorem 2.1 *Let (x, s, y) be a strictly feasible primal-dual pair. Assume $\delta(y, s) \geq \theta > 0$. Let (d_x, d_s, d_y) be the primal-dual Newton direction. Let $x(\alpha) := x + \alpha d_x$, $s(\alpha) := s + \alpha d_s$, and $y(\alpha) := y + \alpha d_y$. Then there exists a step length α such that*

$$\Phi(y(\alpha), s(\alpha)) \leq \phi(y, s) - \zeta,$$

where $\zeta = \frac{\theta}{2(1+\theta)} - \log(1 + \frac{\theta}{2(1+\theta)})$.

Theorem 2.2 *Let (x, s, y) be a strictly feasible primal-dual pair with $\delta(y, s) \leq \theta < 1$. Then $\Phi(y, s) \leq \frac{2\theta^2}{1-\theta^2}$.*

Theorem 2.3 *Let (x^0, s^0, y^0) be a strictly feasible primal-dual pair. A potential decrease algorithm using damped Newton steps can be designed to give an iterate (x, y, s) satisfying $\delta(y, s) \leq \theta < 1$ in at most*

$$O\left(\frac{\Phi(y^0, s^0)}{\zeta}\right)$$

iterations.

We state one other technical lemma that will prove useful in the analysis.

Lemma 2.1 (Ref. 29, Lemma C.1.) *Let $z \in \Re^n$ be such that $\|z\| < 1$. Then*

$$\|z\| - \log(1 + \|z\|) \leq e^T z - \sum_{j=1}^n \log(1 + z_j) \leq -\|z\| - \log(1 - \|z\|).$$

The cutting plane algorithm finds an approximate analytic center (x_0, s_0, y_0) for the current primal-dual system. An oracle is called to determine whether x_0 is feasible in the original convex set. If it is not, the oracle returns a number of separating hyperplanes. Let q denote the number of cuts to be added, and let these cuts be given by

$$w_i^T x \geq w_i^T x_0, \quad i = 1 \text{ to } q. \quad (6)$$

The new system then is

$$(\hat{P}) \quad Ax \geq b, \quad w_i^T x \geq w_i^T x_0 =: b_{m+i}, \quad i = 1 \text{ to } q.$$

Notice that the cuts are shifted to pass through the current center x_0 . The algorithm needs to now find a new approximate analytic center; finding such a point efficiently is the subject of this paper. We define $W := [w_1, w_2, \dots, w_q]$. We assume, without loss of generality, that $\|w_i\| = 1$ for each cut. (Throughout, $\|\cdot\|$ denotes the l_2 norm.)

Our aim is to extend the approach of Ref. 30 as it was used in Ref. 3. Thus, we will need to take an affine step to increase the slacks of the newly added constraints. If just one constraint $w_1^T x \geq b_{m+1}$ is added, then updating x to $x + \alpha w_1$ will ensure that the new constraint is satisfied strictly, for an appropriate positive step length α . In order to use a large step length, it is useful to multiply this direction by a positive definite matrix. We will use a primal-dual algorithm to retrieve a new analytic center after the addition of cutting planes, so we will scale using a diagonal matrix

$$D := Y^{0.5} S^{-0.5} \tag{7}$$

instead of S^{-1} . If the iterate is an exact analytic center, then these two scaling matrices are identical, of course. With this scaling matrix, the Dikin ellipsoid in the primal space can be defined as

$$E_D := \{x : (x - x_0)^T H_0 (x - x_0) \leq 1\}, \tag{8}$$

where

$$H_0 := A^T D_0^2 A \quad \text{and} \quad D_0 := Y_0^{0.5} S_0^{-0.5}. \tag{9}$$

It is then reasonable to choose a direction d_x that will find the largest value for the slack of the new constraint, subject to $x \in E_D$. This gives $d_x = H^{-1} w_1$.

We assume that the cuts returned by the oracle are linearly independent. Consequently, we assume that $q \leq n$. The simplest feasible direction would be one that is a weighted (positive) combination of the normals. However, finding appropriate weights is a non-trivial task. What we seek is an interior feasible solution to

$$W^T (H_0)^{-1} W \gamma \geq 0 \quad ; \quad \gamma \geq 0. \tag{10}$$

Such a solution exists as long as S has a nonempty interior. If we add in a homogenizing constraint of the form $\|\gamma\|_1 = 1$, for example, it is clear that this may be accomplished in $O(\sqrt{qt})$ iterations if we use a short step or a potential reduction algorithm. (All arithmetic operations here and in the ‘parent’ cutting plane algorithm are assumed to be on t bit numbers.) If the interior of the feasible set for the above problem is non-empty, we know that an interior point algorithm will converge to a point in the interior (see e.g., Ref. 31). Once a point in the interior has been found, it is then necessary to find a new approximate analytic center.

Ref. 10 proposed solving a nonlinear programming problem to find appropriate weights γ , and showed that the resulting weights enable a new analytic center to be recovered in $O(q \log(q))$ Newton steps.

3 Selective Gram-Schmidt Orthonormalization

Intuitively, it would be pleasing to set all the weights γ to 1 in (10), after normalizing the cuts in some way. Each of the new cuts would then have the same weight in determining the new iterate. However, this need not give us a feasible direction. Nevertheless, let us consider this: suppose the w_i 's were to satisfy $w_i^T H_0^{-1} w_j \geq 0$ for every i and j . Then it is clear that $c = We$ is a strictly feasible direction. In fact, any positive combination of the normals is now a strictly feasible direction. However, it is unreasonable to expect the oracle to return cuts with this property. So the question becomes, can we generate q cuts based on the cuts returned by the oracle, such that these cuts are valid and satisfy this property. It turns out that we can do this, by a scheme we call Selective Gram-Schmidt Orthonormalization.

Throughout this section, we will frequently use the inner product defined by H_0 , namely $w^T H_0 v$ for two vectors w and v , so we write $\langle w, v \rangle_{H_0} := w^T H_0^{-1} v$. We will also write $\|v\|_{H_0} := \sqrt{v^T H_0^{-1} v}$. Note that we will still use $\|v\|$ to denote the usual 2-norm, so $\|v\|^2 = v^T v$.

Selective Orthonormalization Procedure

Step 0: Define $v_1 := \frac{w_1}{\|w_1\|_{H_0}}$. Set $k = 2$.

Step k: Define $J(k) = \{i \in \{1, 2, \dots, k-1\} | \langle w_k, v_i \rangle_{H_0} < 0\}$.
Set $g_k := w_k - \sum_{i \in J(k)} \langle w_k, v_i \rangle_{H_0} v_i$.
Set $v_k := \frac{g_k}{\|g_k\|_{H_0}}$.

Step Check: If $k = q$, STOP. Else set $k = k + 1$, goto step k.

We are now in a position to state a theorem about the new cuts generated by the above procedure.

Theorem 3.1 *The cuts $v_i^T x \geq v_i^T x_0$ for $i = 1$ to q satisfy the following properties.*

- (i) *For any $i, j \in \{1, 2, \dots, q\}$, $\langle v_i, v_j \rangle_{H_0} \geq 0$.*
- (ii) *If $y \in \mathbb{R}^n$ is any point such that $W^T y \geq W^T x_0$, then $v_j^T y \geq v_j^T x_0$ for any j from 1 to q , so each v_j defines a valid inequality.*

Proof: Induction is used in the proof of each part.

(i) First, we observe that $\langle v_1, v_2 \rangle_{H_0} \geq 0$. This is evident if $\langle w_2, v_1 \rangle_{H_0} \geq 0$, since $v_2 = \frac{w_2}{\|w_2\|_{H_0}}$. However, if $\langle w_2, v_1 \rangle_{H_0} < 0$, then we have

$$g_2 = w_2 - \langle w_2, v_1 \rangle_{H_0} v_1.$$

Therefore,

$$\langle g_2, v_1 \rangle_{H_0} = \langle w_2, v_1 \rangle_{H_0} - \langle w_2, v_1 \rangle_{H_0} \langle v_1, v_1 \rangle_{H_0} = 0$$

since $\langle v_1, v_1 \rangle_{H_0} = 1$. Since g_2 is just a positive multiple of v_2 , it follows that we have $\langle v_2, v_1 \rangle_{H_0} \geq 0$.

Now we do the inductive step. Suppose that it is true for $k-1 \leq q-1$ that $\langle v_{k-1}, v_j \rangle_{H_0} \geq 0$ for $j < k-1$. We will show that $\langle v_k, v_j \rangle_{H_0} \geq 0$ for $j < k$. Recall that

$$g_k = w_k - \sum_{j \in J(k)} \langle w_k, v_j \rangle_{H_0} v_j.$$

Also recall that for all $j \in J(k)$, we have $\langle w_k, v_j \rangle_{H_0} < 0$. Now, for some index $i < k$, we look at $\langle g_k, v_i \rangle_{H_0}$. We do this in cases.

Case 1: $i \in J(k)$.

Now, we have

$$\begin{aligned} \langle g_k, v_i \rangle_{H_0} &= \langle w_k, v_i \rangle_{H_0} - \sum_{j \in J(k)} \langle w_k, v_j \rangle_{H_0} \langle v_j, v_i \rangle_{H_0} \\ &= \langle w_k, v_i \rangle_{H_0} - \sum_{j \in J(k), j \neq i} \langle w_k, v_j \rangle_{H_0} \langle v_j, v_i \rangle_{H_0} - \langle w_k, v_i \rangle_{H_0} \langle v_i, v_i \rangle_{H_0} \\ &= \langle w_k, v_i \rangle_{H_0} - \sum_{j \in J(k), j \neq i} \langle w_k, v_j \rangle_{H_0} \langle v_j, v_i \rangle_{H_0} - \langle w_k, v_i \rangle_{H_0} \\ &= - \sum_{j \in J(k), j \neq i} \langle w_k, v_j \rangle_{H_0} \langle v_j, v_i \rangle_{H_0} \\ &\geq 0 \end{aligned}$$

where that last part follows because $\langle w_k, v_j \rangle_{H_0} < 0$ for $j \in J(k)$, and because $\langle v_j, v_i \rangle_{H_0} \geq 0$ by the inductive hypothesis.

Case 2: i is not in $J(k)$.

Here, we know that $\langle w_k, v_i \rangle_{H_0} \geq 0$. Thus,

$$\begin{aligned} \langle g_k, v_i \rangle_{H_0} &= \langle w_k, v_i \rangle_{H_0} - \sum_{j \in J(k)} \langle w_k, v_j \rangle_{H_0} \langle v_j, v_i \rangle_{H_0} \\ &\geq - \sum_{j \in J(k)} \langle w_k, v_j \rangle_{H_0} \langle v_j, v_i \rangle_{H_0} \\ &\geq 0 \end{aligned}$$

since, as before, $\langle w_k, v_j \rangle_{H_0} < 0$ for $j \in J(k)$, and because $\langle v_j, v_i \rangle_{H_0} \geq 0$ by the inductive hypothesis.

Finally, since g_k is just a positive multiple of v_k these results hold for v_k as well. Thus, we have established that $\langle v_i, v_j \rangle_{H_0} \geq 0$ for any i, j from 1 to q .

(ii) Let x be any point in $\mathbb{R}^n \cap P$ that satisfies $w_i^T x \geq w_i^T x_0$ for all $i = 1$ to q . We will show that x also satisfies $v_i^T x \geq v_i^T x_0$ for all $i = 1$ to q — i.e., we will show that the v_i s are valid cuts. Clearly, v_1 defines a valid cut. We will first show that v_2 defines a cut too.

Case 1: $\langle w_2, v_1 \rangle_{H_0} \geq 0$.

Now we have

$$g_2^T x = w_2^T x \geq w_2^T x_0 = g_2^T x_0.$$

It follows that v_2 defines a valid inequality.

Case 2: $\langle w_2, v_1 \rangle_{H_0} < 0$.

Here, we have

$$\begin{aligned} g_2^T x &= w_2^T x - \langle w_2, v_1 \rangle_{H_0} v_1^T x \\ &\geq w_2^T x_0 - \langle w_2, v_1 \rangle_{H_0} v_1^T x_0 \\ &= g_2^T x_0 \end{aligned}$$

where the inequality follows because $w_2^T x \geq w_2^T x_0$, $v_1^T x \geq v_1^T x_0$ and $\langle w_2, v_1 \rangle_{H_0} < 0$.

Now, again, we use induction. Suppose it to be true that x satisfies $v_i^T x \geq v_i^T x_0$ for $i = 1$ to $k - 1$, where $k - 1 \leq q - 1$. We will show that v_k then must define a valid inequality. We have

$$\begin{aligned} g_k^T x &= w_k^T x - \sum_{j \in J(k)} \langle w_k, v_j \rangle_{H_0} v_j^T x \\ &\geq w_k^T x_0 - \sum_{j \in J(k)} \langle w_k, v_j \rangle_{H_0} v_j^T x_0 \\ &= g_k^T x_0 \end{aligned}$$

where the inequality in the second step follows because $w_k^T x \geq w_k^T x_0$, $v_j^T x \geq v_j^T x_0$ for all $j < k$, and because $\langle w_k, v_j \rangle_{H_0} < 0$ for all $j \in J(k)$. Thus, it is established that each v_i defines a valid cut, if we assume that the oracle returned valid cuts. \square

Finally, we conclude this section by stating some properties of these cuts. Let

$$V := [v_1, \dots, v_q] \tag{11}$$

$$\bar{c} := Ve = \sum_{i=1}^q v_i. \tag{12}$$

Lemma 3.1 *For any index i from 1 to q , we have the following:*

- (i) $\|v_i\|_{H_0} = 1$.
- (ii) $\langle v_i, \bar{c} \rangle_{H_0} \geq 1$. As a consequence, \bar{c} is a strictly feasible direction.
- (iii) $\langle \bar{c}, \bar{c} \rangle_{H_0} \leq q^2$.

Proof:

- (i) This is obvious since $v_i := \frac{g_i}{\|g_i\|_{H_0}}$.
- (ii) We have for every i

$$\begin{aligned}
 \langle v_i, \bar{c} \rangle_{H_0} &= \langle v_i, \sum_{j=1}^q v_j \rangle_{H_0} \\
 &= \sum_{j=1}^q \langle v_i, v_j \rangle_{H_0} \\
 &= \langle v_i, v_i \rangle_{H_0} + \sum_{j \neq i} \langle v_i, v_j \rangle_{H_0} \\
 &= 1 + \sum_{j \neq i} \langle v_i, v_j \rangle_{H_0} \\
 &\geq 1
 \end{aligned}$$

since for any i and j we have $\langle v_i, v_j \rangle_{H_0} \geq 0$.

- (iii) We have

$$\begin{aligned}
 \langle \bar{c}, \bar{c} \rangle_{H_0} &= \|\bar{c}\|_{H_0}^2 \\
 &= \left\| \sum_{j=1}^q v_j \right\|_{H_0}^2 \\
 &\leq \left(\sum_{j=1}^q \|v_j\|_{H_0} \right)^2 \\
 &= q^2.
 \end{aligned}$$

□

We note that what this procedure does is to weaken the cuts produced by the oracle. In the scaled space, if two cuts enclose an acute angle (i.e., their normals form an obtuse angle), one of the cuts is weakened so that the cuts enclose a right angle. Theoretically, this weakening is okay, from a complexity point of view, since in any case we cannot assume very much about the cuts returned by the oracle. From a practical viewpoint, however, this is a drawback. It is the price we pay in order to update the analytic center in $O(\sqrt{q} \log q)$ iterations, as will be seen shortly. However, as noted in the Introduction, in some situations it is computationally advantageous to design an oracle that returns orthogonal cuts.

4 Updating the Analytic Center

In this section, we discuss the procedure to update the analytic center. The new system is the following: (\hat{P}) :

$$(\bar{P}) \quad Ax \geq b, \quad v_i^T x \geq v_i^T x_0 = b_{m+i}, \quad i = 1 \text{ to } q.$$

where the vectors v_i satisfy the conclusion of Lemma 3.1, namely:

$$\|v_i\|_{H_0} = 1, \quad \langle v_i, \bar{c} \rangle_{H_0} \geq 1, \quad \text{and} \quad \langle \bar{c}, \bar{c} \rangle_{H_0} \leq q^2, \quad \text{where } \bar{c} = V\gamma \text{ and } \gamma = e. \quad (13)$$

Such vectors can be found using selective orthonormalization, if necessary, in which case (\bar{P}) is a weakening of the previously encountered (\hat{P}) .

We assume that we have an approximate analytic (x_0, s_0, y_0) center for (P) . In order to obtain an approximate analytic center for (\bar{P}) , we first take an affine step to a specific point in the interior of the modified polytope. An analysis of the primal-dual potential function at this point shows that it is possible to obtain a new approximate analytic center in $O(q \log(q))$ Newton steps. First, we define the updated iterate, and derive some properties of this update.

We define

$$d_x := H_0^{-1} \bar{c} \quad (14)$$

$$d_s := Ad_x \quad (15)$$

$$d_y := -D^2 Ad_x \quad (16)$$

Note that the directions d_s and d_y refer to only the first m components — we deal with the others explicitly. Thus, after this affine step, we have

$$x^+ := x_0 + d_x = x_0 + \alpha_P H_0^{-1} \bar{c} \quad (17)$$

$$s^+ := s_0 + d_s = s_0 + \alpha_P A H_0^{-1} \bar{c} \quad (18)$$

$$y^+ := y_0 + d_y = y_0 - \alpha_D D_0^2 A H_0^{-1} \bar{c} \quad (19)$$

for some primal and dual step lengths α_P and α_D . Again, s^+ and y^+ refer to only the first m components. The augmented slack and dual vectors in \mathbb{R}^{m+q} will be denoted by \bar{s}^+ and \bar{y}^+ respectively.

Define

$$\Delta_{PD} := \sqrt{\bar{c}^T (A^T D_0^2 A)^{-1} \bar{c}} = \sqrt{\bar{c}^T H_0^{-1} \bar{c}} = \|\bar{c}\|_{H_0}. \quad (20)$$

The following lemma can be verified by explicit calculation.

Lemma 4.1 *We have $d_y^T d_s = -\Delta_{PD}^2$.*

The $(m+i)^{th}$ components are given by

$$s_{m+i}^+ = v_i^T d_x = \alpha_P v_i^T H_0^{-1} \bar{c} \quad (21)$$

$$y_{m+i}^+ := \alpha_D \gamma_i = \alpha_D \quad (22)$$

for $i = 1$ to q . If $\alpha_P > 0$ then $s_{m+i} > 0$ from (13). To ensure that the other components of s remain positive, it is necessary to keep $s + \alpha_P d_s > 0$.

Lemma 4.2 *The following step length choices ensure positive iterates.*

1. If $0 < \alpha_P < \sqrt{1-\theta}/\Delta_{PD}$ then $s + \alpha_P d_s > 0$.
2. If $0 < \alpha_D < \sqrt{1-\theta}/\Delta_{PD}$ then $y + \alpha_D d_y > 0$.

Proof: To prove part 1, we have

$$\begin{aligned} s + \alpha_P d_s &= S e + \alpha_P A H_0^{-1} \bar{c} \quad \text{from (15)} \\ &= D_0^{-1} (Y^{0.5} S^{0.5} e + \alpha_P D_0 A^T H_0^{-1} \bar{c}) \\ &\geq D_0^{-1} (\sqrt{1-\theta} - \alpha_P \Delta_{PD}) e \end{aligned}$$

where the last line follows from (5) and (20). Similarly, to show part 2, we have

$$\begin{aligned} y + \alpha_D d_y &= Y e - \alpha_D D_0^2 A H_0^{-1} \bar{c} \quad \text{from (16)} \\ &= D_0 (Y^{0.5} S^{0.5} e - \alpha_P D_0 A H_0^{-1} \bar{c}) \\ &\geq D (\sqrt{1-\theta} - \alpha_P \Delta_{PD}) e. \end{aligned}$$

The result follows. □

It will prove useful to use the same step length in the primal and dual problems, particularly in Lemma 4.4. Therefore, we take

$$\alpha_P = \alpha_D = \alpha = \beta / \Delta_{PD} \quad (23)$$

for an absolute constant β to be specified later, with $0 < \beta < \sqrt{1-\theta}$.

Lemma 4.3 *After taking the specified step, the complementary slackness terms $y_{m+i} s_{m+i}$ satisfy*

$$\beta^2 / q^2 \leq y_{m+i} s_{m+i} \leq \beta^2$$

for $i = 1, \dots, q$.

Proof: From equations (21), (22), and (23), we have

$$\begin{aligned} y_{m+i}s_{m+i} &= \alpha^2 v_i^T H_0^{-1} \bar{c} \\ &= \beta^2 v_i^T H_0^{-1} \bar{c} / \Delta_{PD}^2. \end{aligned}$$

The result follows from (20) and (13). \square

Let us define the vector

$$\tilde{p} := D_0 A H_0^{-1} \bar{c}. \quad (24)$$

Notice that

$$\tilde{p}^T \tilde{p} = \Delta_{PD}^2 \quad (25)$$

and that

$$d_y = -D_0 \tilde{p} \text{ and } d_s = D_0^{-1} \tilde{p}. \quad (26)$$

Lemma 4.4 *The i th component of the complementary slackness is reduced to $y_i^+ s_i^+ = y_i s_i - \alpha^2 \tilde{p}_i^2$ for $i = 1, \dots, m$.*

Proof: From (15) and (16), $Y d_s = -S d_y$. It follows that

$$y_i^+ s_i^+ = (y_i + \alpha(d_y)_i)(s_i + \alpha(d_s)_i) = y_i s_i - \alpha^2 \tilde{p}_i^2.$$

\square

One consequence of Lemma 4.4 is that the objective function does not change after the step to recover an interior point, as shown in the following theorem.

Theorem 4.1 *The duality gap after taking the given step is the same as before taking the step.*

Proof: After taking the step, the duality gap is

$$\begin{aligned} \sum_{i=1}^q \bar{y}_i \bar{s}_i &= \sum_{i=1}^m y_i^+ s_i^+ + \sum_{i=1}^q y_{m+i} s_{m+i} \\ &= \sum_{i=1}^m (y_i s_i - \alpha^2 \tilde{p}_i^2) + \sum_{i=1}^q \beta^2 v_i^T H_0^{-1} \bar{c} / \Delta_{PD}^2 \\ &\quad \text{from Lemma 4.4 and the proof of Lemma 4.3} \\ &= y^T s - \beta^2 + \beta^2 \bar{c}^T H_0^{-1} \bar{c} / \Delta_{PD}^2 \\ &\quad \text{from (25), (23) and (13)} \\ &= y^T s - \beta^2 + \beta^2 \quad \text{from (20)}. \end{aligned}$$

The result follows. \square

We examine the potential function in order to obtain a complexity bound. Define $h \in \mathbb{R}^m$ by

$$h_i := \tilde{p}_i^2 / (y_i s_i). \quad (27)$$

It is useful to bound the norm of this vector.

Lemma 4.5 *The norm of the vector h is bounded:*

$$\alpha^2 \|h\| \leq \alpha^2 \Delta_{PD}^2 / (1 - \theta) = \beta^2 / (1 - \theta).$$

Proof: We have

$$\begin{aligned} \|h\|^2 &= \sum_{i=1}^m \tilde{p}_i^4 / (y_i s_i)^2 \\ &\leq \left[\sum_{i=1}^m \tilde{p}_i^2 \right]^2 / (1 - \theta)^2 \quad \text{from (5)}. \end{aligned}$$

The result follows from (25) and (23). \square

Notice that $\alpha^2 \|h\| < 1$ provided $\beta^2 < 1 - \theta$. Lemma 4.5 can be used to obtain a bound on $\sum_{i=1}^m \log(y_i^+ s_i^+)$.

Lemma 4.6 *The following bound holds for the new iterate, provided $\beta^2 < 1 - \theta$.*

$$\sum_{i=1}^m \log(y_i^+ s_i^+) \geq \sum_{i=1}^m \log(y_i s_i) - \beta^2 / (1 - \theta) + \log(1 - \beta^2 / (1 - \theta)).$$

Proof: We have

$$y_i^+ s_i^+ = y_i s_i - \alpha^2 \tilde{p}_i^2 = y_i s_i (1 - \alpha^2 \tilde{p}_i^2 / (y_i s_i))$$

from Lemma 4.4. Now, from Lemma 2.1, we have

$$\begin{aligned} \sum_{i=1}^m \log(1 - \alpha^2 \tilde{p}_i^2 / (y_i s_i)) &\geq -\alpha^2 \sum_{i=1}^m \tilde{p}_i^2 / (y_i s_i) + \alpha^2 \|h\| + \log(1 - \alpha^2 \|h\|) \\ &\geq -\beta^2 / (1 - \theta) + \log(1 - \beta^2 / (1 - \theta)) \\ &\quad \text{from (25), (5), and Lemma 4.5,} \end{aligned}$$

giving the required bound. \square

It is now possible to bound the value of the potential function Φ at $(\bar{x}^+, \bar{y}^+, \bar{s}^+)$.

Lemma 4.7 *Let $0 < \theta < 1$ and $0 < \eta < 1$ be fixed. Let $\beta = \sqrt{\eta(1 - \theta)}$. The potential function at the new iterate has value $O(q \log(q))$.*

Proof: The potential function at the new iterate $(\bar{x}^+, \bar{y}^+, \bar{s}^+)$ is

$$\begin{aligned}
\Phi(\bar{y}^+, \bar{s}^+) &= \sum_{i=1}^{m+q} \bar{y}_i^+ \bar{s}_i^+ - (m+q) - \sum_{i=1}^{m+q} \log(\bar{y}_i^+ \bar{s}_i^+) \\
&\leq \sum_{i=1}^m y_i s_i - (m+q) - \sum_{i=1}^m \log(y_i s_i) + \beta^2/(1-\theta) \\
&\quad - \log(1 - \beta^2/(1-\theta)) - q \log(\beta^2/q^2) \\
&\quad \text{by Theorem 4.1 and Lemmas 4.6 and 4.3} \\
&\leq \Phi(y, s) + \eta - \log(1 - \eta) + q(2 \log(q) - 1 - \log(\eta) - \log(1 - \theta)) \\
&\quad \text{from the choice of } \beta \\
&\leq \frac{2\theta^2}{1 - \theta^2} + \eta - \log(1 - \eta) + q(2 \log(q) - 1 - \log(\eta) - \log(1 - \theta)) \\
&\quad \text{from Theorem 2.2}
\end{aligned}$$

This gives the stated result. □

The following theorem is then an immediate consequence of Theorem 2.3 and Lemma 4.7.

Theorem 4.2 *The number of Newton steps to regain a new approximate analytic center is bounded by $O(q \log(q))$.*

5 Implications For Cutting Plane Algorithms

In this section, we discuss implications of this procedure to update the analytic center for two cutting plane methods that use analytic centers — namely, the algorithms of Refs. 32 and 4. For the sake of brevity here, we assume familiarity with both these algorithms.

The convergence proofs of these algorithms assume that the added constraints have norm one. After the Selective Orthonormalization Procedure, the added columns no longer have norm 1, they just have norm 1 in the metric defined by the Hessian. However, it is straightforward to ensure that the assumption is satisfied: after finding a new approximate analytic center, the new rows of A can then be scaled appropriately. The primal variables x can be left unchanged but s and y need to be rescaled also. Note that the scale factors used for y_{m+i} and s_{m+i} are reciprocals of one another, so

the product $y_{m+i}s_{m+i}$ is unchanged and therefore the new approximate center is still an approximate center even after this rescaling.

5.1 Ye's Algorithm

In Ref. 7, Ye analyzed the complexity of the cutting plane algorithm of Ref. 32 with multiple cuts, and showed that the total number of cuts needed is no more than $O(\frac{n^2q^2}{\epsilon^2})$, where $q \leq n$ is an upper bound on the number of cuts added at each stage, and ϵ is a small number such that the convex set (defining the convex feasibility problem) contains a ball of radius at least ϵ . We observe that this result is true even if we add the weaker cuts generated by selective orthonormalization. Thus we have the following result.

Theorem 5.1 *Let the algorithm of Ref. 7 terminate after k stages. At stage j , $1 \leq j \leq k$, let $q_j \leq q \leq n$ selectively orthonormalized cuts be added. The total number of Newton steps required is bounded by $O(\frac{n^2q^2}{\epsilon^2} \log(q))$.*

5.2 Atkinson and Vaidya's Algorithm

Here we visualize the algorithm of Ref. 4 with the only difference being that we allow the addition of multiple cuts through the current point in case 2 (see Ref. 4) of their algorithm. The reader familiar with Ref. 4 will note that almost all their results continue to hold. The only results that need to be established are:

1. If at some instance of case 2 of their algorithm, we add $q_i \leq q$ selectively orthonormalized cuts through the current analytic center, we need to show that their potential function grows by at least a constant.
2. We need to show that their lower bound on the determinant of the Hessian continues to hold.

It turns out that these results do hold. The proofs are exactly along the lines of the proofs in Ref. 4, and we don't include them here. The interested reader can find the proofs in Ref. 33. Thus it is true that in $O(nL^2)$ stages the algorithm terminates. We now establish a bound on the number of Newton steps needed.

Theorem 5.2 *Suppose that $q_i \leq q \leq n$ selectively orthonormalized cuts are added through the current iterate at the i^{th} occurrence of case 2 (i.e., the case where cuts are added) of the algorithm described in Ref. 4, and let this be the only modification*

to their algorithm. We then have that the total number of Newton steps is bounded by $O(nL^2 \log(q))$.

Proof: We observe that the number of occurrences of case 1.1 (the case where a cut is dropped) is bounded by $O(nL^2)$. This is because their potential function is guaranteed to increase by a constant in all cases of the algorithm, and if the function grows to $O(nL^2)$ the algorithm terminates. This observation implies that the total number of cuts added in case 2 throughout the course of the algorithm is also bounded by $O(nL^2)$, because the algorithm never keeps more than $O(nL)$ cuts at any time. Since every cut dropped in case 1.1 had to be added previously in an occurrence of case 2, the bound of $O(nL^2)$ on the number of cuts added cannot be improved further. Thus, if we let C_2 denote the number of occurrences of case 2, we have

$$\sum_{i=1}^{C_2} q_i = O(nL^2).$$

Since these cuts are selectively orthonormalized, it follows that occurrence i of case 2 requires no more than $O(q_i \log q_i)$ Newton steps. Thus, the total number of Newton steps needed in case 2 is bounded by $\sum_{i=1}^{C_2} O(q_i \log(q_i)) = O(nL^2 \log(q))$. Since the Newton steps required in other cases is also bounded by $O(nL^2)$, the result follows. \square

6 Conclusions

The selective orthonormalization procedure gives an efficient way to restart when adding multiple cuts in an interior point method. The number of interior point steps to recover a new approximate analytic center is $O(q \log(q))$ when q cuts are added, independent of the dimension of the problem or the number of constraints in the current formulation. If the interior point method was instead restarted from scratch, the number of iterations to obtain a new approximate analytic center would be a function of the size of the whole linear program. The approach we have described is an alternative to that of Ref. 10, which requires solving a nonlinear programming problem in order to determine an initial direction when restarting.

References

- [1] GRÖTSCHEL, M., LOVASZ, L., and SCHRIJVER, A., *Geometric Algorithms and Combinatorial Optimization*, Springer-Verlag, Berlin, Germany, 1988.

- [2] KHACHIYAN, L. G., *A polynomial algorithm in linear programming*, Doklady Akademiia Nauk SSSR, Vol. 224, pp. 1093–1096, 1979. English Translation: Soviet Mathematics Doklady, Vol. 20, pp. 1093–1096.
- [3] GOFFIN, J.-L., LUO, Z.-Q., and YE, Y., *Complexity analysis of an interior cutting plane method for convex feasibility problems*, SIAM Journal on Optimization, Vol. 6, pp. 638–652, 1996.
- [4] ATKINSON, D. S., and VAIDYA, P. M., *A cutting plane algorithm for convex programming that uses analytic centers*, Mathematical Programming, Vol. 69, pp. 1–43, 1995.
- [5] VAIDYA, P. M., *A new algorithm for minimizing convex functions over convex sets*, Mathematical Programming, Vol. 73, pp. 291–341, 1996.
- [6] SONNEVEND, G., *An “analytic center” for polyhedrons and new classes of global algorithms for linear (smooth, convex) programming*, Proceedings of the 12th IFIP Conference on Systems Modeling and Optimization, Budapest, 1985, Edited by A. Prekopa, J. Szelezsan, and B. Strazicky, Springer Verlag, Berlin, West Germany, Vol. 84 of Lecture Notes in Control and Information Sciences, pp. 866–876, 1986.
- [7] YE, Y., *Complexity analysis of the analytic center cutting plane method that uses multiple cuts*, Mathematical Programming, Vol. 78, pp. 85–104, 1997.
- [8] LUO, Z.-Q., ROOS, C., and TERLAKY, T., *Complexity analysis of a logarithmic barrier decomposition method for semi-infinite linear programming*, Applied Numerical Mathematics (Special issue dedicated to the HPOPT-I workshop held in Delft, September 19–20, 1996), Vol. 29, No. 3, pp. 379–395, 1999.
- [9] GOFFIN, J.-L., and VIAL, J.-P., *A two-cut approach in the analytic center cutting plane method*, Mathematical Methods of Operations Research, Vol. 49, No. 1, pp. 149–169, 1999.
- [10] GOFFIN, J.-L., and VIAL, J.-P., *Multiple cuts in the analytic center cutting plane method*, SIAM Journal on Optimization, Vol. 11, No. 1, pp. 266–288, 2001.
- [11] NESTEROV, Y., *Complexity estimates of some cutting plane methods based on the analytic center*, Mathematical Programming, Vol. 69, pp. 177–204, 1995.

- [12] MITCHELL, J. E., and RAMASWAMY, S., *A long-step, cutting plane algorithm for linear and convex programming*, Annals of Operations Research, Vol. 99, pp. 95–122, 2000.
- [13] ANDERSEN, E. D., MITCHELL, J. E., ROOS, C., and TERLAKY, T., *A homogenized cutting plane method to solve the convex feasibility problem*, Optimization Methods and Applications, Edited by X. Q. Yang *et al.*, Kluwer Academic Publishers, Dordrecht, The Netherlands, Chapter 10, 2001.
- [14] NESTEROV, Y., and VIAL, J.-P., *Homogeneous analytic center cutting plane methods for convex problems and variational inequalities*, SIAM Journal on Optimization, Vol. 9, No. 3, pp. 707–728, 1999.
- [15] NESTEROV, Y., PÉTON, O., and VIAL, J.-P., *Homogeneous analytic center cutting plane methods with approximate centers*, Optimization Methods and Software, Vol. 11, pp. 243–273, 1999.
- [16] PÉTON, O., and VIAL, J.-P., *Multiple cuts with a homogeneous analytic center cutting plane method*, Technical report, LOGILAB, Section of Management Studies, University of Geneva, 102 Bd Carl Vogt, CH-1211 Geneva 4, Switzerland, May 2001.
- [17] GOFFIN, J.-L., and VIAL, J.-P., *Convex nondifferentiable optimization: a survey focussed on the analytic center cutting plane method*, Technical Report 99.02, Logilab, Management Studies, University of Geneva, Geneva, Switzerland, February 1999.
- [18] MITCHELL, J. E., *Polynomial interior point cutting plane methods*, Technical report, Mathematical Sciences, Rensselaer Polytechnic Institute, Troy, NY 12180, July 2001.
- [19] BAHN, O., DU MERLE, O., GOFFIN, J.-L., and VIAL, J.-P., *A cutting plane method from analytic centers for stochastic programming*, Mathematical Programming, Vol. 69, pp. 45–73, 1995.
- [20] GOFFIN, J.-L., GONDZIO, J., SARKISSIAN, R., and VIAL, J.-P., *Solving nonlinear multicommodity network flow problems by the analytic center cutting plane method*, Mathematical Programming, Vol. 76, pp. 131–154, 1997.
- [21] MITCHELL, J. E., *Computational experience with an interior point cutting plane algorithm*, SIAM Journal on Optimization, Vol. 10, No. 4, pp. 1212–1227, 2000.

- [22] GONDZIO, J., *Warm start of the primal-dual method applied in the cutting plane scheme*, Mathematical Programming, Vol. 83, pp. 125–143, 1998.
- [23] GONDZIO, J., DU MERLE, O., SARKISSIAN, R., and VIAL, J.-P., *ACCPM — A library for convex optimization based on an analytic center cutting plane method*, European Journal of Operational Research, Vol. 94, pp. 206–211, 1996.
- [24] KRISHNAN, K., *Linear programming approaches to semidefinite programming problems*, PhD thesis, Mathematical Sciences, Rensselaer Polytechnic Institute, Troy, NY 12180, July 2002.
- [25] KRISHNAN, K., and MITCHELL, J. E., *A linear programming approach to semidefinite programming*, Technical report, Mathematical Sciences, Rensselaer Polytechnic Institute, Troy, NY 12180, May 2001.
- [26] OSKOOROUCHI, M. R., and GOFFIN, J.-L., *The analytic center cutting plane method with semidefinite cuts*, Technical Report G–2000–32, GERAD, McGill University, Montreal, Quebec, Canada, July 2000.
- [27] MITCHELL, J. E., and BORCHERS, B., *Solving linear ordering problems with a combined interior point/simplex cutting plane algorithm*, High Performance Optimization, Edited by H. L. Frenk *et al.*, Kluwer Academic Publishers, Dordrecht, The Netherlands, Chapter 14, pp. 349–366, 2000.
- [28] MITCHELL, J. E., *An interior point cutting plane algorithm for Ising spin glass problems*, Operations Research Proceedings, SOR 1997, Jena, Germany, Edited by P. Kischka and H.-W. Lorenz, Springer-Verlag, pp. 114–119, 1998.
- [29] ROOS, C., TERLAKY, T., and VIAL, J.-P., *Theory and Algorithms for Linear Optimization: An Interior Point Approach*. John Wiley, Chichester, 1997.
- [30] MITCHELL, J. E., and TODD, M. J., *Solving combinatorial optimization problems using Karmarkar’s algorithm*, Mathematical Programming, Vol. 56, pp. 245–284, 1992.
- [31] GÜLER, O., and YE, Y., *Convergence behaviour of interior point algorithms*, Mathematical Programming, Vol. 60, No. 2, pp. 215–228, 1993.
- [32] YE, Y., *A potential reduction algorithm allowing column generation*, SIAM Journal on Optimization, Vol. 2, pp. 7–20, 1992.

- [33] RAMASWAMY, S., *Cutting plane algorithms for linear and convex programming*, PhD thesis, Rensselaer Polytechnic Institute, DSES, RPI, Troy, NY, 1995.