An algorithm for linear programming which requires $O(((m+n)n^2+(m+n)^{1.5}n)L)$ arithmetic operations

Pravin M. Vaidya

AT&T Bell Laboratories
Murray Hill, New Jersey 07974

Abstract

We present an algorithm for linear programming which requires $O(((m+n)n^2+(m+n)^{1.5}n)L)$ arithmetic operations where m is the number of inequalities, and n is the number of variables. Each operation is performed to a precision of $O(L)$ bits. L is bounded by the number of bits in the input.

1. Introduction

We study the linear programming problem

$$\max c^T x$$
$$\text{s.t. } Ax \geq b$$

where $A \in R^{m \times n}$, $b \in R^m$, and $c \in R^n$. We assume that the polytope defined by $Ax \geq b$ is bounded and has a non-zero interior. As the polytope is bounded we can assume that $m \geq n$, and that the columns of A are linearly independent.

A polynomial time algorithm for the linear programming problem was first presented by Khachian [3] using the ellipsoid method. Khachian's algorithm requires $O(mn^3L)$ arithmetic operations in the worst case, and each operation is performed to a precision of $O(L)$ bits where

$$L = \log_2(\text{largest absolute value of the determinant of any square submatrix of } A) + \log_2(\max_i c_i) + \log_2(\max_i b_i) + \log_2(m+n).$$

In [2] Karmarkar presents an interior point algorithm which requires $((m^{1.5}n^2+m^2n)L)$ arithmetic operations, each operation being performed to a precision of $O(L)$ bits. We present an algorithm for the linear programming problem which requires $O((mn^2+m^{1.5}n)L)$ arithmetic operations in the worst case, and it is adequate to perform each arithmetic operation to a precision of $O(L)$ bits. The algorithm presented in this paper is thus faster than Karmarkar's algorithm [2] by a factor of $\sqrt{m}$ for all values of m and n. It is also faster than Khachian's ellipsoid algorithm by a factor of n for $m \leq n^2$, and continues to be faster than his algorithm for $m \leq n^4$. (Typically, m and n are of the same order).

In [5] Renegar gives an algorithm for linear programming which requires $O(m^{1.5}n^2L)$ arithmetic operations. We build on the ideas in [5], and obtain a faster algorithm. In our algorithm we construct a sequence $P^0, P^1, \cdots, P^k, \cdots$ of smaller and smaller polytopes which shrink towards the optimal vertex (facet). During the *kth* iteration, we move from the center of polytope P^{k-1} to the center of polytope P^k by performing a local optimization which consists of minimizing a linear function over an ellipsoid. The total number of iterations performed by the algorithm is $O(\sqrt{m}L)$, and on the average each iteration requires $O(\sqrt{m}n^2+mn)$ arithmetic operations where the average is taken over all the iterations. Thus the total number of arithmetic operations is $O((mn^2+m^{1.5}n)L)$. When the algorithm terminates we have a feasible point that is sufficiently close in objective function value to the optimum over the original polytope, and we may then jump to an optimal solution as described in [4] in $O(mn^2)$ arithmetic operations.

In section 2 we give an overview, and describe the relationship with related work [1, 5]. In section 3 we give the algorithm. In section 4 we give some properties of the potentials used to measure convergence, and in section 5 we describe the local optimizations. In section 6 we show how to amortize the number of arithmetic operations. In section 7 we show how to reduce any linear program to the format required by the algorithm, and in section 8 we show that $O(L)$ precision is adequate for arithmetic operations.

2. An Overview

Let P be the given polytope

$$P = \{x : Ax \geq b\}$$

and let $\beta^{\max}$ be the maximum value of the objective function c^Tx over P. Let $\beta^0, \beta^1, \cdots, \beta^k, \cdots$ be a strictly increasing sequence such that $\lim_{k \to \infty} \beta^k = \beta^{\max}$. Let π^k denote the set of linear inequalities $\{Ax \geq b,\ c^Tx \geq \beta^k\}$, and P^k denote the polytope $P^k = \{x : Ax \geq b,\ c^Tx \geq \beta^k\}$. Let a_i^T denote the *ith* row of the constraint matrix A. The center of the system of linear inequalities π^k is defined to be the unique point that maximizes the potential function

$$F^k(x) = \sum_{i=1}^{m} \ln(a_i^Tx - b_i) + m\ln(c^Tx - \beta^k)$$

over the interior of the polytope P^k. The center of π^k is unique since $F^k(x)$ is a strictly concave function. We let ω^k denote the center of the system π^k. Let $f^k(x)$ be the normalized potential function given by $f^k(x) = F^k(\omega^k) - F^k(x)$.

$f^k(x)$ is a strictly convex function. Furthermore, the value of $F^k(x)$ at ω^k is not required for evaluating the derivatives of $f^k(x)$, and for computing $f^k(y) - f^k(z)$ for a pair of points y and z.

The algorithm generates a sequence of points $x^0, x^1, \cdots, x^k, \cdots$ such that x^k is in the interior of polytope P^k, and x^k is a good approximation to the center ω^k of the system π^k. Specifically, each x^k satisfies the condition $f^k(x^k) \leq 0.04$. We shall assume that β^0 satisfies the condition $\beta^{\max} - \beta^0 = 2^{O(L)}$, and that we have a point x^0 such that $f^0(x^0) \leq 0.04$. In section 7 we show how to transform the given linear program so that a required starting point x^0 is available. In the *kth* iteration, β^k is computed as $\beta^k = \beta^{k-1} + \frac{\alpha}{\sqrt{m}}(c^T x^{k-1} - \beta^{k-1})$, for some constant α. As the bounding objective function hyperplane shifts due to the change in β, x^{k-1} may no longer be a good approximation to the new center ω^k, and so a local optimization is performed to decrease the potential $f^k(x)$ and obtain a better approximation x^k to the center ω^k. The local optimization consists of minimizing a linear function over an ellipsoid. The sequence $\beta^0, \beta^1, \cdots, \beta^k, \cdots$ satisfies the condition $\beta^{\max} - \beta^k \leq (1 - \frac{\alpha'}{\sqrt{m}})(\beta^{\max} - \beta^{k-1})$ for some constant α' dependent on α. So the corresponding sequence of systems $\pi^0, \pi^1, \cdots, \pi^k, \cdots$ induces a sequence of smaller and smaller polytopes $P^0, P^1, \cdots, P^k, \cdots$ which shrink geometrically towards the optimal vertex (facet). In $O(\sqrt{m}L)$ iterations we obtain a point where the objective function value is $2^{-O(L)}$ away from the optimal value over the original polytope P, and we may then isolate the set of constraints which define an optimal vertex of P [4].

We shall now describe how to find x^k from x^{k-1}. It may be shown that if $f^{k-1}(x^{k-1}) \leq 0.04$ and $\alpha \leq 1/30$ then $f^k(x^{k-1}) \leq 0.05$. Thus at the start of the *kth* iteration we have a point x^{k-1} satisfying the condition $f^k(x^{k-1}) \leq 0.05$, and we have to obtain a point x^k such that $f^k(x^k) \leq 0.04$. Let D be an $m \times m$ diagonal matrix such that the *ith* diagonal entry D_{ii} satisfies the condition $\frac{1}{1.1(a_i^T x^{k-1} - b_i)^2} \leq D_{ii} \leq \frac{1.1}{(a_i^T x^{k-1} - b_i)^2}$. Let $E^k(r)$ be the ellipsoid defined by

$$E^k(r) = \{x : (x - x^{k-1})^T G^k (x - x^{k-1}) \leq r^2 f^k(x^{k-1})\}$$

where

$$G^k = A^T D A + \frac{m}{(c^T x^{k-1} - \beta^k)^2} c c^T$$

and r is a suitable parameter between 0 and 1. We note that $E^k(r)$ is contained in the polytope P^k. Let η^k be the gradient of $f^k(x)$ evaluated at x^{k-1}. Consider the power series expansion of $f^k(x)$ at x^{k-1}. Within the ellipsoid $E^k(r)$, the magnitude of the second order term in this series is bounded by $0.55 r^2 f^k(x^{k-1})$, and the sum of magnitudes of the higher order terms is at most $\frac{0.1 r^3 f^k(x^{k-1})}{(1 - 0.27r)}$. Whereas the minimum value attained by the first order (linear) term within the ellipsoid is less than $-0.9 r f^k(x^{k-1})$. So minimizing $(\eta^k)^T (x - x^{k-1})$, the linear term in the power series, over the ellipsoid $E^k(r)$ will decrease $f^k(x)$ by approximately a factor of $(1 - 0.9r + 0.55r^2)$. Let $z^k(r)$ be the point that minimizes the linear function $(\eta^k)^T x$ over the ellipsoid $E^k(r)$. If $0.5 \leq r \leq 0.8$ then $z^k(r)$ reduces the potential $f^k(x)$ by at least 25% and $f^k(z^k(r)) \leq 0.75 f^k(x^{k-1})$.

x^k is computed as follows. From the theory of convex functions [6], $z^k(r) - x^{k-1}$ satisfies the system of linear equations

$$(A^T D A + \frac{m}{(c^T x^{k-1} - \beta^k)^2} c c^T)(z^k(r) - x^{k-1}) = -t(r)\,\eta^k$$

for some scalar $t(r) > 0$. We first compute ξ^k, a vector in the direction of $z^k(r) - x^{k-1}$, by solving the system of linear equations

$$(A^T D A + \frac{m}{(c^T x^{k-1} - \beta^k)^2} c c^T)\,\xi^k = -\eta^k$$

Next, we find a scalar $t^k > 0$ such that

$$0.018 \leq (t^k)^2 (\xi^k)^T (A^T D A + \frac{m}{(c^T x^{k-1} - \beta^k)^2} c c^T)\,\xi^k \leq 0.0196$$

If $0.04 \leq f^k(x^{k-1}) \leq 0.05$ then $x^{k-1} + t^k \xi^k$ minimizes the function $(\eta^k)^T x$ over the ellipsoid $E^k(r_0)$, for some r_0 in the range $[0.6, 0.7]$, and $f^k(x^{k-1} + t^k \xi^k) \leq 0.75 f^k(x^{k-1}) \leq 0.04$. Thus either $f^k(x^{k-1}) \leq 0.04$ or $f^k(x^{k-1} + t^k \xi^k) \leq 0.04$. It suffices to let x^k be that point where the potential $f^k(x)$ is lower among the two points x^{k-1} and $x^{k-1} + t^k \xi^k$.

An algorithm based on the idea of producing a sequence of shrinking polytopes together with a sequence of approximate centers was given by Renegar [5]. An approach based on centers is also suggested in [1, 9] but without any analysis. In developing our algorithm we follow Renegar's approach, but there are two critical differences which enable us to obtain a faster algorithm. First, closeness to the center of a polytope is measured in a different manner. Renegar [5] measures closeness to the center in terms of euclidean distance in a transformed domain, and shows that a local optimization decreases the distance to the current center. We measure closeness in terms of the potentials $f^k(x)$. Second, the local optimizations described in this paper are quite different from Renegar's [5], and can actually increase the distance metric used by him in [5] to measure closeness to the center. Finally, Renegar's algorithm requires $O(m^{1.5} n^2 L)$ arithmetic operations. Measuring local convergence in terms of the potentials $f^k(x)$ allows us to amortize the number of arithmetic operations, and obtain a bound of $O((mn^2 + m^{1.5} n)L)$ on the total number of arithmetic operations performed by our algorithm.

In [1] Bayer and Lagarias analyze an infinitesimal version of Karmarkar's algorithm, and study trajectories leading from each point to the optimum defined by taking infinitesimal steps. It is interesting to note that the points $\omega^0, \omega^1, \cdots, \omega^k, \cdots$ lie on such a trajectory that would be generated by starting the infintesimal version of Karmarkar's algorithm at ω^0. Thus our algorithm could also be viewed as efficiently following this trajectory to the optimum.

3. The Algorithm

In this section we give the actual algorithm. We assume that we are given a β^0 such that $\beta^{\max} - \beta^0 = 2^{O(L)}$, and an x^0 which is close enough to the center ω^0 of the system of linear inequalities $\pi^0 = \{Ax \geq b,\ c^T x \geq \beta^0\}$. Specifically, $f^0(x^0) \leq 0.04$.

At the beginning of the *kth* iteration we have a parameter β^{k-1}, and a feasible point x^{k-1} such that $c^T x^{k-1} > \beta^{k-1}$, and $f^{k-1}(x^{k-1}) \leq 0.04$. We also have a diagonal matrix D such that $\frac{1}{1.1(a_i^T x^{k-1} - b_i)^2} \leq D_{ii} \leq \frac{1.1}{(a_i^T x^{k-1} - b_i)^2}$, for $i = 1, \cdots, m$. During the *kth* iteration we perform the following computations in sequence.

1. $\beta^k := \beta^{k-1} + \frac{1}{30\sqrt{m}}(c^Tx^{k-1} - \beta^{k-1})$.

2. Determine a direction ξ^k by solving

$$(A^TDA + \frac{m}{(c^Tx^{k-1}-\beta^k)^2}cc^T)\,\xi^k = -\eta^k .$$

where η^k is the gradient of $f^k(x)$ evaluated at x^{k-1}.

3. Compute a scalar $t^k > 0$ such that

$$0.018 \le (t^k)^2(\xi^k)^T(A^TDA + \frac{m}{(c^Tx^{k-1}-\beta^k)^2}cc^T)\,\xi^k \le 0.0196 .$$

4. If $f^k(x^{k-1}+t^k\xi^k) < f^k(x^{k-1})$ then $x^k := x^{k-1}+t^k\xi^k$ else $x^k := x^{k-1}$.

5. For each i, $1 \le i \le m$,

$$\text{if } (D_{ii} < \frac{1}{1.1\,(a_i^Tx^k-b_i)^2}) \text{ or } (D_{ii} > \frac{1.1}{(a_i^Tx^k-b_i)^2})$$

$$\text{then } D_{ii} := \frac{1}{(a_i^Tx^k-b_i)^2} .$$

The algorithm halts when $c^Tx^k - \beta^k \le 2^{-\theta L}$ for some prespecified constant $\theta \ge 13$.

Each iterate x^k is a good approximation to the center ω^k, and satisfies the condition $f^k(x^k) \le 0.04$. It is this property of the algorithm which leads to a bound of $O(\sqrt{m}\,L)$ on the number of iterations. Using this property we can show the following Theorem.

Theorem 1. If $m \ge 16$ then for all k, $c^Tx^k - \beta^k \ge 0.4(\beta^{\max} - \beta^k)$.

From Theorem 1. it follows that, $\beta^{\max} - \beta^k \le (1 - \frac{0.4}{30\sqrt{m}})(\beta^{\max} - \beta^{k-1})$, and since $\beta^{\max} - \beta^0 = 2^{O(L)}$, $c^Tx^k - \beta^k$ must fall below $2^{-\theta L}$ in $O(\sqrt{m}\,L)$ iterations. Thus the algorithm halts in $O(\sqrt{m}\,L)$ iterations. From Theorem 1, it also follows that when the algorithm halts, $\beta^{\max} - c^Tx^k \le 1.25 \times 2^{-\theta L}$. Using the final point generated by the algorithm an exact optimum (an optimal vertex) may be computed in $O(mn^2)$ arithmetic operations [4].

The proof of Theorem 1 is based on the following Lemmas. For the lemmas below we assume that $m \ge 16$.

Lemma 1. For all k,

$$c^T\omega^k - \beta^k \ge 0.5(\beta^{\max} - \beta^k) ,$$

and if $f^k(x) \le 0.04$ then

$$\frac{|c^T\omega^k - c^Tx|}{c^T\omega^k - \beta^k} \le \frac{0.4}{\sqrt{2m}} .$$

Lemma 2. For all k, if $f^{k-1}(x^{k-1}) \le 0.04$ then $f^k(x^{k-1}) \le 0.05$.

Lemma 3. For all k, if $0.04 \le f^k(x^{k-1}) \le 0.05$ then $f^k(x^{k-1}+t^k\xi^k) \le 0.04$.

Lemma 1 follows from Lemmas 4 and 5 which are proved in section 4. Lemma 2 follows from Lemma 7 that is also proved in section 4. A proof of Lemma 3 is given in section 5. From Lemmas 2 and 3, it inductively follows that for all k, $f^k(x^k) \le 0.04$. Theorem 1 then follows from Lemma 1.

4. Potential Functions

We shall study some properties of the potential

$$F(x) = \sum_{i=1}^{m} \ln(a_i^Tx - b_i) + m\ln(c^Tx - \beta)$$

and the point ω that maximizes $F(x)$ over the polytope $P^\beta = \{x : Ax \ge b,\ c^Tx \ge \beta\}$. We note that ω is a unique point since $F(x)$ is strictly concave. Let $f(x)$ be the normalized potential given by $f(x) = F(\omega) - F(x)$.

Lemma 4. For any point x in the polytope P^β,

$$\sum_{i=1}^{m} \frac{(a_i^Tx - b_i)}{(a_i^T\omega - b_i)} + m\frac{c^Tx - \beta}{c^T\omega - \beta} = 2m ,$$

and

$$c^T\omega - \beta \ge 0.5(c^Tx - \beta) .$$

Proof. Since the gradient of $F(x)$ vanishes at ω, taking the dot product of the gradient at ω with $x - \omega$ gives

$$\sum_{i=1}^{m} \frac{a_i^T(x-\omega)}{a_i^T\omega - b_i} + m\frac{c^T(x-\omega)}{c^T\omega - \beta} = 0 .$$

Lemma 4 then follows. ■

The next Lemma states that if x and ω are close in potential then the value of a_i^Tx (as well as c^Tx) at x and at ω can differ only by a small amount.

Lemma 5. If $f(x) \le \frac{\delta^2}{2} - \frac{\delta^3}{3}$, where $0 \le \delta < 1$, then

$$\frac{|a_i^Tx - a_i^T\omega|}{a_i^T\omega - b_i} \le \delta,\ 1 \le i \le m ,$$

and

$$\frac{|c^T\omega - c^Tx|}{c^T\omega - \beta} \le \frac{\delta}{\sqrt{2m}} .$$

Proof. Suppose that $f(x) \le \frac{\delta^2}{2} - \frac{\delta^3}{3}$, where $0 \le \delta < 1$. Let $\psi_i(x) = \frac{a_i^Tx - a_i^T\omega}{a_i^T\omega - b_i}$, for $i = 1, 2, \cdots, m$, and let $\psi_i(x) = \frac{c^Tx - c^T\omega}{c^T\omega - \beta}$, for $i = m+1, \cdots, 2m$. First, we show that

$$\sum_{i=1}^{2m} \psi_i(x)^2 \le \delta^2 \qquad \cdots (5.1).$$

As $f(x)$ is strictly convex, the minimum value of $f(x)$ over the region $\{x : \sum_{i=1}^{2m}\psi_i(x)^2 \ge \delta^2,\ Ax \ge b,\ c^Tx \ge \beta\}$ occurs on the boundary of the ellipsoid $\Sigma(\delta) = \{x : \sum_{i=1}^{2m}\psi_i(x)^2 \le \delta^2\}$. We shall lower bound the value of $f(x)$ on the boundary of $\Sigma(\delta)$. We have that

$$f(x) = -\sum_{i=1}^{2m} \ln(1 + \psi_i(x))$$

Using the Taylor series expansion for $\ln(1+\psi_i(x))$, within the ellipsoid $\Sigma(\delta)$, $f(x)$ may be expressed as

$$f(x) = \sum_{i=1}^{2m}\sum_{j=1}^{\infty} \frac{(-1)^j\psi_i(x)^j}{j}$$

From Lemma 4, $\sum_{i=1}^{2m} \psi_i(x) = 0$. Also, on the boundary of the region $\Sigma(\delta)$, $\sum_{i=1}^{2m} \sum_{j=4}^{\infty} \frac{(-1)^j \psi_i(x)^j}{j} > 0$. Thus on the boundary of $\Sigma(\delta)$,

$$f(x) > \sum_{i=1}^{2m} \frac{1}{2}\psi_i(x)^2 - \frac{1}{3}\psi_i(x)^3$$

$$> \sum_{i=1}^{2m} \psi_i(x)^2 (\frac{1}{2} - \frac{\delta}{3}), \qquad \text{since} |\psi_i(x)| \le \delta .$$

$$> \delta^2 (\frac{1}{2} - \frac{\delta}{3}).$$

(5.1). above then follows. Then

$$\sum_{i=1}^{2m} |\psi_i(x)| = \delta\sqrt{2m}$$

and since $\sum_{i=1}^{2m} \psi_i(x) = 0$ we get that

$$\sum_{i=m+1}^{2m} |\psi_i(x)| = m \frac{|c^T x - c^T \omega|}{c^T \omega - \beta} \le \delta\sqrt{m/2} \quad \blacksquare$$

We note that Lemma 1 in section 3 follows from Lemmas 4 and 5 above.

Let $F'(x)$ be the potential defined as $F'(x) = \sum_{i=1}^{m} \ln(a_i^T x - b_i) + m \ln(c^T x - \beta')$, and let ω' be the point that maximizes $F'(x)$ over the region $\{x : Ax \ge b,\ c^T x \ge \beta'\}$. Let $f'(x) = F'(\omega') - F'(x)$ be the normalized potential corresponding to $F'(x)$. The next lemma bounds the change in objective function value at the center due to change in the parameter β.

Lemma 6. Let $\beta' \ge \beta$. Then $c^T \omega' \ge c^T \omega$ and $c^T \omega' - c^T \omega \le \beta' - \beta$.

Proof. Given in [5].

The next Lemma bounds the change in potential at a point due to change in the parameter β.

Lemma 7. Let x be a point in the interior of the polytope $P^\beta = \{x : Ax \ge b,\ c^T x \ge \beta\}$ such that $f(x) \le \frac{\delta^2}{2} - \frac{\delta^3}{3}$, where $0 \le \delta < 1$. Let $\beta' = \beta + \frac{\alpha}{\sqrt{m}}(c^T x - \beta)$, where $\alpha \ge 0$. Then

$$f'(x) \le f(x) + \frac{\alpha\delta}{\sqrt{2}(1 - \frac{\alpha}{\sqrt{m}})} + \frac{\alpha^2 (1 + \frac{\delta}{\sqrt{2m}})^2}{(1 - \frac{\alpha}{\sqrt{m}} - \frac{\alpha\delta}{\sqrt{2m}})}.$$

Proof. We may write $f'(x)$ as

$$f'(x) = f(x) + f'(\omega) + m \ln(\frac{(c^T x - \beta)(c^T \omega - \beta')}{(c^T x - \beta')(c^T \omega - \beta)})$$

We have

$$\frac{(c^T x - \beta)(c^T \omega - \beta')}{(c^T x - \beta')(c^T \omega - \beta)})$$

$$= 1 + \frac{(\beta' - \beta)(c^T \omega - c^T x)}{(c^T x - \beta')(c^T \omega - \beta)}$$

$$\le 1 + \frac{\alpha\delta}{m} \frac{1}{\sqrt{2}(1 - \frac{\alpha}{\sqrt{m}})}, \cdots \text{(by Lemma 5)}.$$

Thus

$$m \ln(\frac{(c^T x - \beta)(c^T \omega - \beta')}{(c^T x - \beta')(c^T \omega - \beta)}) \le \frac{\alpha\delta}{\sqrt{2}(1 - \frac{\alpha}{\sqrt{m}})}$$

Next,

$$\frac{c^T \omega' - \beta'}{c^T \omega - \beta'}$$

$$= \frac{c^T \omega' - \beta}{c^T \omega - \beta}(1 + \frac{(\beta' - \beta)(c^T \omega' - c^T \omega)}{(c^T \omega' - \beta)(c^T \omega - \beta')})$$

$$\le \frac{c^T \omega' - \beta}{c^T \omega - \beta} + \frac{(\beta' - \beta)^2}{(c^T \omega - \beta)(c^T \omega - \beta')} \cdots \text{(by Lemma 6)}$$

$$\le \frac{c^T \omega' - \beta}{c^T \omega - \beta} + \frac{\alpha^2 (1 + \frac{\delta}{\sqrt{2m}})^2}{m(1 - \frac{\alpha}{\sqrt{m}} - \frac{\alpha\delta}{\sqrt{2m}})} \cdots \text{(by Lemma 5)}$$

Then

$$f'(\omega)$$

$$= \sum_{i=1}^{m} \ln(\frac{a_i^T \omega' - b_i}{a_i^T \omega - b_i}) + m \ln(\frac{c^T \omega' - \beta'}{c^T \omega - \beta'})$$

$$\le \sum_{i=1}^{m} (\frac{a_i^T \omega' - b_i}{a_i^T \omega - b_i} - 1) + m(\frac{c^T \omega' - \beta'}{c^T \omega - \beta'} - 1)$$

$$\le \sum_{i=1}^{m} \frac{a_i^T \omega' - b_i}{a_i^T \omega - b_i} + m \frac{c^T \omega' - \beta}{c^T \omega - \beta} - 2m$$

$$+ \alpha^2 (1 + \frac{\delta}{\sqrt{2m}})^2 (1 - \frac{\alpha}{\sqrt{m}} - \frac{\alpha\delta}{\sqrt{2m}})^{-1}$$

From Lemma 4,

$$\sum_{i=1}^{m} \frac{a_i^T \omega' - b_i}{a_i^T \omega - b_i} + m \frac{c^T \omega' - \beta}{c^T \omega - \beta} - 2m = 0.$$

Thus

$$f'(w) \le \frac{\alpha^2 (1 + \frac{\delta}{\sqrt{2m}})^2}{(1 - \frac{\alpha}{\sqrt{m}} - \frac{\alpha\delta}{\sqrt{2m}})} \quad \blacksquare$$

We note that Lemma 2 in section 3 follows from Lemma 7 above.

5. Local Optimizations

In this section we describe the local optimizations, and show that x^k, the point computed during the kth iteration, sufficiently reduces the potential $f^k(x)$. We shall collect together a few definitions. Let $F^k(x)$ be the potential defined as

$$F^k(x) = \sum_{i=1}^{m} \ln(a_i^T x - b_i) + m\ln(c^T x - \beta^k)$$

and ω^k be the point that maximizes $F^k(x)$ over the region $\{x : Ax \geq b,\ c^T x \geq \beta^k\}$. Let $f^k(x) = F^k(\omega) - F^k(x)$ be the normalized potential corresponding to $F^k(x)$. Let x^{k-1} be the point at the beginning of the kth iteration. Let η^k be the gradient of $f^k(x)$ evaluated at x^{k-1}, and let D be a diagonal matrix such that $\frac{1}{1.1\,(a_i^T x^{k-1} - b_i)^2} \leq D_{ii} \leq \frac{1.1}{(a_i^T x^{k-1} - b_i)^2}$.

x^k is obtained from x^{k-1} as follows.

1. Determine a direction ξ^k by solving the system of linear equations

$$(A^T DA + \frac{m}{(c^T x^{k-1} - \beta^k)^2} cc^T)\, \xi^k = -\eta^k$$

2. Next, compute a scalar $t^k > 0$ satisfying the condition

$$0.018 \leq (t^k)^2 (\xi^k)^T (A^T DA + \frac{m}{(c^T x^{k-1} - \beta^k)^2} cc^T)\, \xi^k \leq 0.0196,$$

3. If $f^k(x^{k-1} + t^k \xi^k)) < f^k(x^{k-1})$ then $x^k := x^{k-1} + t^k \xi^k$ else $x^k := x^{k-1}$.

It is adequate to show that if $f^k(x^{k-1})$ exceeds 0.04 then $x^{k-1} + t^k \xi^k$ sufficiently reduces the potential $f^k(x)$. The following lemma was introduced in section 3.

Lemma 4. If $0.04 \leq f^k(x^{k-1}) \leq 0.05$ then $f^k(x^{k-1} + t^k \xi^k) \leq 0.75 f^k(x^{k-1}) \leq 0.04$.

In this section we shall prove an alternate Lemma, i.e. Lemma 8 below, and Lemma 4 will follow as a consequence of Lemma 8. We shall require some additional notation. Let $E^k(r)$ be the ellipsoid

$$E^k(r) = \{x : (x - x^{k-1})^T G^k (x - x^{k-1}) \leq r^2 f^k(x^{k-1})\}$$

where

$$G^k = A^T DA + \frac{m}{(c^T x^{k-1} - \beta^k)^2} cc^T$$

and r lies between 0 and 1. Let $z^k(r)$ be the point that minimizes the linear function $(\eta^k)^T x$ over the ellipsoid $E^k(r)$. Let H^k denote the Hessian of $f^k(x)$ evaluated at x^{k-1}. Note that H^k may be written as

$$H^k = \sum_{i=1}^{m} \frac{1}{(a_i^T x^{k-1} - b_i)^2} a_i a_i^T + \frac{m}{(c^T x^{k-1} - \beta^k)^2} cc^T.$$

Lemma 8. Suppose that $0 < f^k(x^{k-1}) \leq \frac{\delta^2}{2} - \frac{\delta^3}{3}$, where $0 \leq \delta < 1$. Then

$$f^k(z^k(r)) \leq (1 - \mu r + 0.55 r^2 + \nu r^3) f^k(x^{k-1}),$$

where

$$\mu = \frac{\sqrt{(1-\delta)}}{\left(1.1\,(\sum_{j=0}^{\infty} \frac{\delta^j}{(j+1)(j+2)})\right)^{\frac{1}{2}}},$$

and

$$\nu = \frac{(1.1)^{3/2}}{3} \frac{\sqrt{f^k(x^{k-1})}}{(1 - \sqrt{1.1 r^2 f^k(x^{k-1})})}.$$

Before proving Lemma 8, we shall show how Lemma 4 follows from Lemma 8. As $z^k(r)$ minimizes $(\eta^k)^T x$ over the ellipsoid $E^k(r)$, from the theory of convex functions [6], it follows that $z^k(r) - x^{k-1}$ satisfies the system of linear equations

$$(A^T DA + \frac{m}{(c^T x^{k-1} - \beta^k)^2} cc^T)\,(z^k(r) - x^{k-1}) = -t(r)\,\eta^k,$$

for some scalar $t(r) > 0$. So ξ^k, the direction computed during the kth iteration, and $z^k(r) - x^{k-1}$ are in the same direction. Furthermore, when $f^k(x^{k-1})$ is in the range [0.04, 0.05], $x^{k-1} + t^k \xi^k$ equals $z^k(r_0)$, for some r_0 in the range [0.6, 0.7]. So from Lemma 8 we may conclude that if $0.04 \leq f^k(x^{k-1}) \leq 0.05$ then $f^k(x^{k-1} + t^k \xi^k) \leq 0.75 f^k(x^{k-1}) \leq 0.04$.

We shall now give a proof of Lemma 8.

Proof of Lemma 8. Let $x = x^{k-1} + t\xi$. Using a power series expansion at x^{k-1}, $f^k(x^{k-1} + t\xi)$ may be written as

$$f^k(x^{k-1} + t\xi) = f^k(x^{k-1}) + t(\eta^k)^T \xi + \frac{t^2}{2} \xi^T H^k \xi + \sum_{j=3}^{\infty} \frac{(-1)^j t^j}{j} (\sum_{i=1}^{m} \frac{(a_i^T \xi)^j}{(a_i^T x^{k-1} - b_i)^j} + \frac{m\,(c^T \xi)^j}{(c^T x^{k-1} - \beta^k)^j})$$

Since $\frac{1}{1.1\,(a_i^T x^{k-1} - b_i)^2} \leq D_{ii} \leq \frac{1.1}{(a_i^T x^{k-1} - b_i)^2}$,

$$\frac{t^2}{2} \xi^T H^k \xi \leq \frac{1.1\, t^2}{2} \xi^T (A^T DA + \frac{m}{(c^T x^{k-1} - \beta^k)^2} cc^T)\xi,$$

and so within the ellipsoid $E^k(r)$ the magnitude of the second order term in the above series is upper bounded by $0.55 r^2 f^k(x^{k-1})$. Next, we bound the sum of the magnitudes of the third and higher order terms in the power series for an arbitrary point $x^{k-1} + t\xi$ in the ellipsoid $E^k(r)$.

$$|\sum_{j=3}^{\infty}\frac{(-1)^j t^j}{j}(\sum_{i=1}^{m}\frac{(a_i^T\xi)^j}{(a_i^Tx^{k-1}-b_i)^j}+\frac{m\,(c^T\xi)^j}{(c^Tx^{k-1}-\beta^k)^j})|$$

$$\le\sum_{j=3}^{\infty}\frac{1}{j}(1.1r^2f^k(x^{k-1}))^{j/2}$$

$$\le\frac{(1.1)^{3/2}}{3}\frac{r^3\sqrt{f^k(x^{k-1})}}{(1-\sqrt{1.1r^2f^k(x^{k-1})})}f^k(x^{k-1}).$$

Having upper bounded the sum of the magnitudes of the second and the higher order terms in the power series for points in $E^k(r)$, we shall lower bound the maximum change in the linear term that is possible within the ellipsoid $E^k(r)$. Such a bound is provided by Lemma 9. Let x' be the point where the line joining x^{k-1} and ω^k intersects the boundary of the ellipsoid $E^k(r)$. By Lemma 9,

$$(\eta^k)^T(x'-x^{k-1})\le-\mu\, r f^k(x^{k-1}).$$

Since $(\eta^k)^Tz^k(r)\le(\eta^k)^Tx'$, Lemma 8 follows from Lemma 9 and the upper bounds on sum of the higher order terms in the above power series. ■

Lemma 9. Let x' be the point where the line joining x^{k-1} and ω^k intersects the boundary of the ellipsoid $E^k(r)$. If $0<f^k(x^{k-1})\le\frac{\delta^2}{2}-\frac{\delta^3}{3}$, where $0\le\delta<1$, then

$$(\eta^k)^T(x'-x^{k-1})\le-\mu\, r f^k(x^{k-1})$$

where $\mu=\dfrac{\sqrt{(1-\delta)}}{\left(1.1(\sum_{j=0}^{\infty}\frac{\delta^j}{(j+1)(j+2)})\right)^{\frac{1}{2}}}$.

Proof. Let $x^{k-1}-x'=\lambda u$, where u is the unit vector in the direction of $x^{k-1}-x'$, and $\lambda=\|x^{k-1}-x'\|_2$. We have

$$\lambda^2u^T(A^TDA+\frac{m}{(c^Tx^{k-1}-\beta^k)^2}cc^T)u=r^2f^k(x^{k-1}).$$

Since $\dfrac{1}{1.1(a_i^Tx^{k-1}-b_i)^2}\le D_{ii}\le\dfrac{1.1}{(a_i^Tx^{k-1}-b_i)^2}$,

$$\lambda^2u^TH^ku\ge\frac{1}{1.1}r^2f^k(x^{k-1}).$$

Thus

$$\lambda\ge r\sqrt{(\frac{f^k(x^{k-1})}{1.1\,u^TH^ku})}.$$

Hence

$$(\eta^k)^T(x^{k-1}-x')$$

$$\ge\frac{r\sqrt{f^k(x^{k-1})}}{\sqrt{1.1}}\frac{(\eta^k)^Tu}{\sqrt{u^TH^ku}}$$

$$\ge\frac{rf^k(x^{k-1})}{\sqrt{1.1}}\frac{(\eta^k)^T(x^{k-1}-\omega^k)}{\sqrt{f^k(x^{k-1})}\sqrt{(x^{k-1}-\omega^k)^TH^k(x^{k-1}-\omega^k)}}\;..(9.1).$$

Let $y_i=\dfrac{a_i^Tx^{k-1}-a_i^T\omega^k}{a_i^T\omega^k-b_i}$, for $i=1,\cdots,m$, and let $y_i=\dfrac{c^Tx^{k-1}-c^T\omega^k}{c^T\omega^k-\beta^k}$ for $i=m+1,\cdots,2m$,. Then

$$(\eta^k)^T(x^{k-1}-\omega^k)=\sum_{i=1}^{2m}(\frac{1}{1+y_i}-1),$$

$$(x^{k-1}-\omega^k)^TH^k(x^{k-1}-\omega^k)=\sum_{i=1}^{2m}(\frac{1}{1+y_i}-1)^2,$$

and

$$f^k(x^{k-1})=\sum_{i=1}^{2m}\ln(\frac{1}{1+y_i}).$$

From Lemma 4 in section 4,

$$\sum_{i=1}^{2m}y_i=0.$$

Since $f^k(x^{k-1})\le\frac{\delta^2}{2}-\frac{\delta^3}{3}$, where $0\le\delta<1$, from Lemma 5 in section 4 we get that

$$|y_i|\le\delta,\ \ i=1,2,\cdots,2m.$$

We can thus apply Lemma 10 below, and from (9.1). above conclude that

$$(\eta^k)^T(x^{k-1}-x')\ge rf^k(x^{k-1})\frac{\sqrt{(1-\delta)}}{\left(1.1(\sum_{j=0}^{\infty}\frac{\delta^j}{(j+1)(j+2)})\right)^{\frac{1}{2}}}\ ■$$

Lemma 10. Suppose that $\sum_{i=1}^{2m}y_i==0$, $\sum_{i=1}^{2m}\ln(1+y_i)<0$, and $|y_i|\le\delta<1$, for $i=1,2,\cdots,2m$. Then

$$\sum_{i=1}^{2m}(\frac{1}{1+y_i}-1)\ge(1-\delta)\sum_{i=1}^{2m}(\frac{1}{1+y_i}-1)^2,$$

and

$$\sum_{i=1}^{2m}(\frac{1}{1+y_i}-1)\ge\frac{\sum_{i=1}^{2m}\ln(\frac{1}{1+y_i})}{\sum_{j=0}^{\infty}\frac{\delta^j}{(j+1)(j+2)}}.$$

Proof. Since $\sum_{i=1}^{2m}y_i=0$,

$$\sum_{i=1}^{2m}(\frac{1}{1+y_i}-1)=\sum_{i=1}^{2m}(\frac{1}{1+y_i}+y_i-1)=\sum_{i=1}^{2m}(\frac{y_i^2}{1+y_i})$$

So as $|y_i|\le\delta$,

$$\sum_{i=1}^{2m}(\frac{1}{1+y_i}-1)\ge(1-\delta)\sum_{i=1}^{2m}\frac{y_i^2}{(1+y_i)^2}$$

$$\ge(1-\delta)\sum_{i=1}^{2m}(\frac{1}{1+y_i}-1)^2$$

Next, as $\sum_{i=1}^{2m}y_i=0$,

$$\sum_{i=1}^{2m}\ln(\frac{1}{1+y_i})=\sum_{i=1}^{2m}(y_i-\ln(1+y_i))$$

$$=\sum_{y_i\ne0,\,1\le i\le m}\frac{y_i^2}{1+y_i}\frac{(1+y_i)(y_i-\ln(1+y_i))}{y_i^2}$$

Using the Taylor series expansion for $\ln(1+y_i)$, for $y_i \neq 0$ we get

$$\frac{(1+y_i)(y_i - \ln(1+y_i))}{y_i^2}$$

$$= \frac{1}{2} + \sum_{j=1}^{\infty} \frac{(-1)^{j+1} y_i^j}{(j+1)(j+2)}$$

$$\leq \sum_{j=0}^{\infty} \frac{\delta^j}{(j+1)(j+2)}, \text{ as } |y_i| \leq \delta .$$

Thus

$$\sum_{i=1}^{2m} \ln(\frac{1}{1+y_i}) \leq \left(\sum_{j=0}^{\infty} \frac{\delta^j}{(j+1)(j+2)} \right) \sum_{i=1}^{2m} (\frac{y_i^2}{1+y_i})$$

Lemma 10 then follows. ■

6. Amortizing the number of arithmetic operations

In this section we show that the total number of arithmetic operations performed by the algorithm is $O((mn^2 + m^{1.5}n)L)$. The amortization of the number of arithmetic operations is similar to the one in [2]. The total number of arithmetic operations is determined by the number of operations required for the following computations.

1. Solving systems of linear equations to determine the directions ξ^k.
2. Computing the gradients η^k of the potentials $f^k(x)$ and computing the scalars t^k.

In the kth iteration we determine a direction ξ^k by solving the system of linear of equations

$$(A^T DA + \frac{m}{(c^T x^{k-1} - \beta^k)^2} cc^T) \xi^k = -\eta^k$$

where η^k is the gradient of $f^k(x)$ evaluated at x^{k-1}, and we find a scalar t^k such that

$$0.018 \leq (t^k)^2 (\xi^k)^T (A^T DA + \frac{m}{(c^T x^{k-1} - \beta^k)^2} cc^T) \xi^k \leq 0.0196$$

The gradient η^k can be computed in $O(mn)$ operations, and once we have ξ^k, a required t^k may also be obtained in $O(mn)$ operations. The total number of iterations is $O(\sqrt{m}L)$. So the total number of operations required to compute η^k and scalar t^k over all the iterations is $O(m^{1.5}nL)$. We maintain $(A^T DA)^{-1}$ and update it whenever the matrix D changes. Once $(A^T DA)^{-1}$ is available, $(A^T DA + \frac{m}{(c^T x^{k-1} - \beta^k)^2} cc^T)^{-1}$ may be computed in $O(n^2)$ operations as the two matrices differ by a rank one matrix, and then ξ^k may be obtained in $O(n^2)$ extra operations. We shall show that the total number of operations required to maintain $(A^T DA)^{-1}$ during the entire execution of the algorithm is $O(mn^2 L)$, and then the desired bound on the total number of arithmetic operations performed by the algorithm would follow.

At the end of the kth iteration D is updated as follows. For $i = 1, \cdots, m$, the ith diagonal element D_{ii} is reset to $\frac{1}{(a_i^T x^k - b_i)^2}$ if $D_{ii} \notin [\frac{1}{1.1 (a_i^T x^k - b_i)^2}, \frac{1.1}{(a_i^T x^k - b_i)^2}]$.

Suppose that D' is the matrix obtained by changing the ith diagonal element of D to d'. Then

$$A^T D' A = A^T DA + (d' - D_{ii}) a_i a_i^T$$

Thus whenever an element of D is changed, $A^T DA$ changes by a rank one matrix, and hence $(A^T DA)^{-1}$ changes by a rank one matrix. Therefore when an element of D is changed, $(A^T DA)^{-1}$ may be updated in $O(n^2)$ operations, using th rank one update formula

$$(B + uv^T)^{-1} = B^{-1} - \frac{(B^{-1}u)(B^{-1}v)^T}{1 + v^T B^{-1} u}.$$

So to obtain a bound of $O(mn^2 L)$ on the total number of operations required to maintain $(A^T DA)^{-1}$, it is sufficient to show that the total number of changes to the matrix D during the entire execution of the algorithm is $O(mL)$.

Let

$$\phi_i^k = |\ln(\frac{a_i^T x^k - b_i}{a_i^T x^{k-1} - b_i})|.$$

Suppose D_{ii} was reset at the lth iteration and at the jth iteration but was not reset between the lth and jth iterations. Then $\sum_{k=l+1}^{j} \phi_i^k \geq \ln(1.1)$, and the total number of times D_{ii} is changed during the execution of the algorithm is $O(\sum_{k=1}^{I} \phi_i^k)$, where I is the number of iterations performed by the algorithm. Thus, the total number of changes to D is $O(\sum_{i=1}^{m} \sum_{k=1}^{I} \phi_i^k)$.

A bound on $\sum_{i=1}^{m} \sum_{k=1}^{I} \phi_i^k$ may be obtained as follows. x^k lies within an ellipsoid around x^{k-1}, and $x^k - x^{k-1}$ satisfies the condition

$$(x^k - x^{k-1})^T (A^T DA + \frac{m}{(c^T x^{k-1} - \beta^k)^2} cc^T)(x^k - x^{k-1}) \leq 0.0196$$

Since at the start of the kth iteration, for $i = 1, \cdots, m$, $D_{ii} \in [\frac{1}{1.1 (a_i^T x^{k-1} - b_i)^2}, \frac{1.1}{(a_i^T x^{k-1} - b_i)^2}]$, it follows that

$$\sum_{i=1}^{m} (\frac{a_i^T x^k - b_i}{a_i^T x^{k-1} - b_i} - 1)^2 \leq 1.1 \times 0.0196,$$

and thus

$$\sum_{i=1}^{m} |(\frac{a_i^T x^k - b_i}{a_i^T x^{k-1} - b_i} - 1)| \leq 1.1 \times 0.0196 \times \sqrt{m}.$$

Then using the taylor series expansion for the natural logarithm, it is easily shown that $\sum_{i=1}^{m} \phi_i^k = O(\sqrt{m})$, and since I, the number of iterations, is $O(\sqrt{m}L)$ we may conclude that $\sum_{i=1}^{m} \sum_{k=1}^{I} \phi_i^k = O(mL)$.

7. Recasting a linear program into the required format

In this section we show how to transform the given linear program so that a suitable starting point for the transformed program is available. There are several ways to carry out such a transformation. The one we give is similar to the one in [5]. The given linear program is

$$\max \quad p^T x$$
$$\text{s.t.} \quad Hz \ge q$$

where $z \in R^{n_1}$, $p \in R^{n_1}$, $q \in R^{m_1}$, and $H \in R^{m_1 \times n_1}$. We reserve A, x, b, and c to refer to a linear program that is already in the required format. Let

L_1 = $\log_2$(largest absolute value of the determinant of any square submatrix of H) $+ \log_2(\max_i p_i) + \log_2(\max_i q_i) + \log_2(m_1+n_1)$.

Note that

1. If the given linear program has an optimal solution then every optimal vertex z^{opt} satisfies the condition $\|z^{opt}\|_\infty \le 2^{L_1}$.
2. If the polytope $\{z : Hz \ge q\}$ is unbounded then there is a feasible solution z^f such that $\|z^f\|_\infty \le 2^{L_1}$.

Let $t \in R$, let $e \in R^{m_1}$ be a vector given by $e^T = [1, 1, \cdots, 1]$, let $\lambda = m_1 n_1 2^{2L_1}$, and let $\mu = 2^{30L_1}$. Let h_i^T denote the ith row of H. The transformed linear program is as follows.

$$\max \quad p^T z + \mu t$$
$$\text{s.t.} \quad h_i^T z - (\lambda + q_i)t \ge -\lambda, \; i = 1,2, \cdots, m_1.$$
$$-e^T H z \ge -\lambda$$
$$z_j \ge -\lambda, \; j = 1,2, \cdots, n_1.$$
$$-z_j \ge -\lambda, \; j = 1,2, \cdots, n_1.$$
$$-\lambda t \ge -\lambda$$
$$((m+1)\lambda + e^T q)t \ge -\lambda$$

Let $m = m_1 + 2n_1 + 3$, and let $n = n_1 + 1$. Let $A \in R^{m \times n}$ denote the constraint matrix, $b \in R^m$ denote the right hand side of the constraints, $c \in R^n$ denote the objective function vector, and $x \in R^n$ the variables in the transformed problem. Note that $c^T = [p^T, \mu]$, and $x^T = [z^T, t]$. The tranformed linear program may be written as

$$\max \quad c^T x$$
$$\text{s.t.} \quad Ax \ge b.$$

The transformed problem has the following properties.

1. Since $((m+1)\lambda + e^T b) > 0$, t is bounded, and thus the polytope $\{x : Ax \ge b\}$ is bounded.
2. Let

   L = $\log_2$(largest absolute value of the determinant of any square submatrix of A) $+ \log_2(\max_i c_i) + \log_2(\max_i b_i) + \log_2(m+n)$.

   Then $L \le 40L_1$. The bound on L follows from the observations that the largest absolute value of the determinant of any square submatrix of A is at most $(m_1+n_1)^6 2^{6L_1}$, and that $\max_i b_i \le m_1 n_1 2^{2L_1}$, and that $\max_i c_i \le 2^{30L_1}$.
3. 0 is feasible. Since the sum of the rows of A is the zero vector and all the coordinates of b have the same value, the gradient of the function $\sum_{i=1}^{m} \ln(a_i^T x - b_i)$ vanishes at 0. Furthermore, as the polytope $\{x : Ax \ge b\}$ is bounded, 0 is the unique point that maximizes the function $\sum_{i=1}^{m} \ln(a_i^T x - b_i)$.
4. A point $\begin{bmatrix} z \\ 1 \end{bmatrix}$ is feasible for the transformed linear program iff $Hz \ge q$, $e^T H z \le \lambda$, and $-\lambda \le z_j \le \lambda$, $j = 1, 2, \cdots, n_1$.
5. μ is large enough so that if there exists a feasible point with $t = 1$ then every optimal solution has $t = 1$. This is because the minimum vertex to vertex variation of the function μt exceeds the maximum change in the function $p^T z$ over the entire polytope $\{x : Ax \ge b\}$.

We shall now show that 0 is an adequate starting point for running the algorithm in section 3 on the transformed problem. Let $\beta^0 = -m^3 2^{2L}$. Let $F^0(x) = \sum_{i=1}^{m} \ln(a_i^T x - b_i) + m \ln(c^T x - \beta^0)$, and ω^0 be the point that maximizes $F^0(x)$. Since the magnitude of β^0 is large enough $|m \ln(\frac{c^T \omega^0 - \beta^0}{c^T x - \beta^0})| \le \frac{1}{m}$ for any point x in the transformed polytope. Thus, $F^0(\omega^0) - F^0(0) \le 0.04$ as required, for $m \ge 4$.

Finally, we have the following easily shown lemma.

Lemma 11. Let $\begin{bmatrix} z^{opt} \\ t^{opt} \end{bmatrix}$ be an optimal vertex in the transformed linear program.

1. If $t^{opt} < 1$ then the original linear program is infeasible.
2. If $t^{opt} = 1$, $e^T H z^{opt} < \lambda$, and $|z_j^{opt}| < \lambda$, $j = 1, 2, \cdots, n_1 + 1$, then z^{opt} is an optimal vertex in the original linear program.
3. If $t^{opt} = 1$, and either $|z_j^{opt}| = \lambda$, for some j, or $e^T H z^{opt} = \lambda$, then either the original problem is unbounded or z is an optimal solution. In this case the transformed problem is solved again with λ replaced by 2λ to obtain a new optimal point $\begin{bmatrix} z^* \\ t^* \end{bmatrix}$. If $p^T z^* > p^T z^{opt}$ then the original problem is unbounded, otherwise z^{opt} is an optimal solution. ■

8. Precision of arithmetic operations

The error in the solution of a system of linear equations is directly related to the conditon number of the matrix describing the system, and the precision used for arithmetic operations [7, 8]. In the first part of this section we show that the entries in the matrix D at each iteration are upper and lower bounded by 2^{18L} and 2^{-4L} respectively, and that the condition numbers of the matrices arising during local optimizations are upper bounded by 2^{30L}. Using these bounds on condition numbers we shall argue that it is adequate to maintain $(A^TDA)^{-1}$ to an accuracy of νL bits for some constant ν. Then in the second part of the section we describe how sufficient accuracy in $(A^TDA)^{-1}$ may be maintained during rank one changes using $O(L)$ bits of precision.

8.1. Condition number of local optimization matrices.

As before let x^{k-1} be the point at the beginning of the kth iteration. During the kth iteration we determine a direction ξ^k by solving the system of linear equations

$$(A^TDA + \frac{m}{(c^Tx^{k-1}-\beta^k)^2}cc^T)\xi^k = -\eta^k$$

where D is a diagonal matrix such that $D_{ii} \in [\frac{1}{1.1(a_i^Tx^{k-1}-b_i)}, \frac{1.1}{a_i^Tx^{k-1}-b_i}]$, and $-\eta^k = \sum_{i=1}^{m}\frac{1}{a_i^Tx^{k-1}-b_i}a_i + \frac{m}{c^Tx^{k-1}-b_i}c$ is the gradient of $f^k(x)$ evaluated at x^{k-1}. We maintain $(A^TDA)^{-1}$ by performing rank one changes, and compute $(A^TDA + \frac{m}{(c^Tx^{k-1}-\beta^k)^2}cc^T)^{-1}$ by a rank one change to $(A^TDA)^{-1}$.

We shall first bound the entries in D. Note that the absolute value of $a_i^Tx-b_i$ and c^Tx for all feasible x is upper bounded by 2^{4L}. Moreover, we may assume that $c^Tx^{k-1}-\beta^{k-1} > 2^{-13L}$ (since the algorithm halts when $c^Tx^k-\beta^k$ is less than 2^{-13L}). Next, we shall lower bound the value of $a_i^Tx^{k-1}-b_i$, for $1 \le i \le m$. Let $\overline{\beta^{k-1}}$ be obtained by rounding β^{k-1} to $15L$ bits. A vertex of the polytope $\{x : Ax \ge b\ ,\ c^Tx \ge \overline{\beta^{k-1}}\}$ has rational coordinates with a common denominator which is most 2^{16L}, and so the maximum change in the value of $a_i^Tx-b_i$ over this polytope is at least 2^{-16L}. Thus, the maximum value of $a_i^Tx-b_i$ over the polytope $\{Ax \ge b\ ,\ c^Tx \ge \beta^{k-1}\}$ is at least 2^{-16L}. Then from Lemma 4 in section 3,

$$a_i^T\omega^{k-1}-b_i \ge \frac{2^{-16L}}{2m},$$

where ω^{k-1} is the point that maximizes $F^{k-1}(x)$. Furthermore, as $F^{k-1}(\omega^{k-1})-F^{k-1}(x^{k-1}) \le 0.04$, from Lemma 5 in section 3 we get that

$$a_i^Tx^{k-1}-b_i \ge \frac{2^{-16L}}{4m},$$

for $m \ge 16$. Thus we may conclude that the entries in D are upper and lower bounded by 2^{18L} and 2^{-4L} respectively.

We shall now obtain bounds on condition numbers. Note that the condition number of a symmetric positive definite matrix is just the ratio of the largest to the the smallest eigenvalue, and that D, A^TDA, and $A^TDA + \frac{m}{(c^Tx^{k-1}-\beta^k)^2}cc^T$, are symmetric positive definite matrices. An elementary arguement using rayleigh quotients [7, 8] shows that the condition number of A^TDA is bounded by the product of the condition numbers of A^TA and D. Furthermore, a straightforward calculation shows that the condition number of A^TA is at most 2^{7L}. So from the above bounds for entries in D we may conclude that the condition number of A^TDA is at most 2^{29L}. A similar calculation gives a lower bound of 2^{-8L} on the smallest eigenvalue of A^TDA, and an upper bound of 2^{21L} on the largest eigenvalue of A^TDA. This in turn leads to a bound of 2^{30L} on the condition of $A^TDA + \frac{m}{(c^Tx^{k-1}-\beta^k)^2}cc^T$.

Let ν be fixed constant greater than 100. We maintain an approximate inverse $(A^TDA)_a^{-1}$ such that

$$(A^TDA)^{-1} = (A^TDA)_a^{-1} + \Delta$$

where $\|\Delta\|_2 = 2^{-\nu L}$. Then, from the above bounds on condition numbers,

$$(A^TDA)_a^{-1} - \frac{m(A^TDA)_a^{-1}cc^T(A^TDA)_a^{-1}}{1+mc^T(A^TDA)_a^{-1}c}$$
$$= (A^TDA + \frac{m}{(c^Tx^{k-1}-\beta^k)^2}cc^T)^{-1} + \Delta'$$

where $\|\Delta'\|_2 \le 2^{-(\nu-40)L}$. Once the approximate inverse $(A^TDA)_a^{-1}$ is available, $x^{k-1}+t^k\xi^k$ is obtained as follows. The gradient η^k, the approximation to $(A^TDA + \frac{m}{(c^Tx^{k-1}-\beta^k)^2}cc^T)^{-1}$ given by the above formula, the direction ξ^k, and the scalar t^k are all computed using $2\nu L$ bits of precision, and at the end each coordinate of the point $x^{k-1}+t^k\xi^k$ is rounded off to $40L$ bits.

Finally, as $\|\eta^k\|_2 \le 2^{22L}$, and $t^k \le 2^{2L}$, it is easily shown that the error in $x^{k-1}+t^k\xi^k$ is $O(2^{-40L})$. Then the potential difference between the computed and the exact value of $x^{k-1}+t^k\xi^k$ is neglible.

8.2. Maintaining accuracy in the inverse

Here we shall briefly describe how to maintain accuracy in $(A^TDA)^{-1}$ during rank one changes. Let B denote $(A^TDA)^{-1}$. Suppose we have an approximate inverse B' of B such that $BB' = I + E_1$, and $\|E_1\| \le 2^{-\gamma_1 L}$. Then $B' = B^{-1} + B^{-1}E_1$, and $\|B^{-1}E_1\|_2 \le 2^{-(\gamma_1-29)L}$ as the condition number of B is at most 2^{29L}. A good approximation to $(B + uv^T)^{-1}$ is computed as follows.

1. Compute an initial approximation $B'' = B' - \frac{(B'u)(B'v)^T}{1+v^TB'u}$.
2. $(B + uv^T)(B'') = I + E_1 + E_2$, where E_2 is a constant rank matrix computable in $O(n^2)$ operations. An adequate approximate inverse of B is obtained by rounding the entries in $B'' - B''E_2$ to multiples of $2^{-\gamma_2 L}$.

Let $B'' - B''E_2 + E_3$ be the computed approximate inverse. Then

$$(B + uv^T)(B'' - B''E_2 + E_3)$$

$$= I + E_1 - (E_1 + E_2)E_2 + (B + uv^T)E_3 .$$

We note that the condition number of $A^T DA$ is bounded by 2^{29L}. It is then easily shown that $\|E_2\|_2 \le 2^{60L}\|E_1\|_2$. So we may choose γ_1 and γ_2 large enough so that $\|E_2^2\|_2$, and $\|(B + uv^T)E_3\|_2$ are each less than $2^{-L}\|E_2\|_2$. Then

$$\|E_2(E_1 + E_2) + (B + uv^T)E_3\|_2 \le (1 + O(2^{-L}))\, 2^{-\gamma_1 L}$$

and the error in B^{-1} grows very slowly.

9. Conclusion

We have presented an algorithm for linear programming which requires $O(((m+n)n^2 + (m+n)^{1.5} n)L)$ arithmetic operations where m is the number of constraints, and n is the number of variables. Each operation is performed to a precision of $O(L)$ bits. L is bounded by the number of bits in the input.

Acknowledgements

The author would like to thank Sanjiv Kapoor, Narendra Karmarkar, and Jeff Lagarias for helpful discussions and suggestions.

References

1. D. A. Bayer, and J. C. Lagarias, The non-linear geometry of Linear Programming I. Affine and Projective scaling trajectories, *Trans. Amer. Math. Soc.*, (to appear).
2. N. Karmarkar, A new polynomial time algorithm for linear programming, *Combinatorica*, Vol. 4, 1984, pp. 373-395.
3. L. G. Khachian, Polnomial algorithms in linear programming, *Zhurnal Vychislitelnoi Mathematiki i Matematicheskoi Fiziki*, Vol. 20, 1980, pp. 53-72.
4. C. Papadimitriou, and K. Steiglitz, *Combinatorial Optimization: Algorithms Complexity*, Prentice-Hall, Inc., Englewood Cliffs, New Jersey.
5. J. Renegar, A polynomial-time algorithm, based on Newton's method, for linear programming, MSRI 07118-86, Mathematical Sciences Research Institute, Berkeley, California.
6. G. Zoutendijk, *Mathematical Programming Methods*, North-Holland, New York, 1976.
7. G. W. Stewart, *Introduction to matrix computations*, 1973, Academic Press, Inc., New York.
8. J. H. Wilkinson, *The Algebraic Eigenvalue Problem*, 1965, Oxford University Press (Clarendon), London and New York.
9. Gy. Sonnevand, An analytical center for polyhedrons and new classes of global algorithms for linear (smooth, convex) programming, preprint, Dept. of Numerical Analysis, Institute of Mathematics, Eotvos University, 1088, Budapest, Muzeum Korut, 6-8.