

I. I. Dikin's Convergence Result for the Affine-Scaling Algorithm

R. J. VANDERBEI AND J. C. LAGARIAS

ABSTRACT. The affine-scaling algorithm is an analogue of Karmarkar's linear programming algorithm that uses affine transformations instead of projective transformations. Although this variant lacks some of the nice properties of Karmarkar's algorithm (for example, it is probably not a polynomial-time algorithm), it nevertheless performs well in computer implementations. It has recently come to the attention of the western mathematical programming community that a Soviet mathematician, I. I. Dikin, proposed the basic affine-scaling algorithm in 1967 and published a proof of convergence in 1974. Dikin's convergence proof assumes only primal nondegeneracy, while all other known proofs require both primal and dual nondegeneracy. Our aim in this paper is to give a clear presentation of Dikin's ideas.

1. Introduction

In 1984, N. K. Karmarkar [7] discovered a polynomial-time algorithm for linear programming, which, unlike earlier such algorithms, was said to perform better than the traditional simplex method. This pioneering work inspired a huge amount of research on interior methods for linear programming. Today, these methods are divided into two basic classes. Algorithms in the first class are called projective-scaling algorithms (see [10] for a list of references). This class includes the original algorithm studied by Karmarkar. Roughly speaking, projective-scaling algorithms are hard to describe but easy to analyze. That is, the description of these algorithms involves technicalities such as logarithmic barrier functions, assumptions that the optimal objective function value is zero, etc., which although they originally seemed rather obscure are now seen as forcing the sequence of points generated by the algorithm to home in on something called the 'central trajectory.' Once these technicalities are understood, the analysis proceeds to a proof of polynomial-time convergence.

Algorithms in the second class are called affine-scaling algorithms (see, e.g.,

1980 *Mathematics Subject Classification* (1985 Revision). Primary 90C05.

©1990 American Mathematical Society
0271-4132/90 \$1.00 + \$.25 per page

[4], [13]). In comparison, these algorithms are easy to describe but hard to analyze. That is, affine-scaling algorithms have very simple, geometrically intuitive descriptions, but proving convergence is a difficult and interesting mathematical problem. Until recently, the only known proofs of convergence (in the west) involved assuming that the linear program was primal and dual nondegenerate. (Empirical evidence suggests that neither of these assumptions is necessary.) Furthermore, there is strong evidence [8] to support the belief that these algorithms are in fact exponential in the worst case.

This segregation into two classes is not intended to imply that projective-scaling algorithms are 'first-class' algorithms whereas the affine-scaling algorithms are 'second class.' In fact, all serious large scale implementations currently being pursued use affine-scaling methods ([1], [2], [9], [11], [12]). The reason is that the affine-scaling methods have certain advantages over the projective-scaling algorithms. For example, they apply directly to problems presented in standard form, and on the average they are computationally more efficient. Ironically, this brings us back to the old situation where what is theoretically best (worst case) is not practically best (average case).

A further point worth mentioning is that the distinction between these two classes of algorithms is being blurred as people discover polynomial-time algorithms that look more and more like affine-scaling algorithms. The most notable example of this is the recent paper by Monteiro et al. [10], where it is shown that the affine-scaling algorithm applied to the primal-dual problem is in fact polynomial if the initial point is close to the center of the polytope and the step size is not too big.

It has recently come to the attention of the western scientific community that a Soviet mathematician, I. I. Dikin, proposed the basic affine-scaling algorithm in the Soviet Mathematics Doklady in 1967 [5]. He published a proof of convergence in 1974 [6]. It turns out that, not only did Dikin predate the west by almost 20 years, but also his proof of convergence does not require the dual nondegeneracy assumption. The purpose of this paper is to give a clear presentation of Dikin's methods.

Perhaps the most interesting open problem in this area is to prove that the affine-scaling algorithm converges even if the problem is primal degenerate. Essentially, all real-world problems are both primal and dual degenerate, and yet practical experience shows that this does not present any difficulty (except that the code has to be able to solve a consistent system of equations even when there are dependent or almost dependent rows). In a recent paper [3] by Adler and Monteiro, it was shown that the continuous trajectories associated with the affine-scaling algorithm do indeed converge even when the problem is primal and/or dual degenerate. Hence the problem is the discreteness of the affine-scaling algorithm. Dikin chose a step size that is smaller than the one in [13] and was able to remove the dual nondegeneracy assumption. Perhaps by taking an even smaller (but noninfinitesimal) step size, it might be possible to remove the primal nondegeneracy assumption as well.

Acknowledgment. We would like to thank Mike Todd for several stimulating discussions. In particular, he pointed out the appropriate argument to show that $(AD_x^2 A^T)^{-1}$ is bounded.

2. The main convergence result

In this section, we describe the basic affine-scaling algorithm studied by Dikin. The *primal linear program* is

$$(P) \quad \begin{aligned} \min c \cdot x \\ Ax = b \\ x \geq 0. \end{aligned}$$

The associated *dual linear program* is

$$(D) \quad \begin{aligned} \max b \cdot w \\ A^T w \leq c. \end{aligned}$$

We begin by introducing some notation. Let $\Omega = \{x \in R^n: Ax = b, x \geq 0\}$ denote the polytope for the primal problem; let $\Omega^0 = \{x \in \Omega: x > 0\}$ denote the relative interior of Ω ; and let $\partial\Omega = \Omega - \Omega^0$ denote the boundary of Ω . Finally, given a vector x , let D_x denote the diagonal matrix having the components of x along its diagonal.

The motivation behind the affine-scaling algorithm can be found in many papers (see, e.g., [13]). Therefore, in this paper we assume that the reader has seen the motivation and we go straight to the definition of the algorithm. For this, we need to introduce three important functions. The first function $w: \Omega^0 \rightarrow R^m$ associates with each $x \in \Omega^0$ a vector of *dual variables*:

$$w(x) = (AD_x^2 A^T)^{-1} AD_x^2 c.$$

The second function $r: \Omega^0 \rightarrow R^n$ measures the slackness in the inequality constraints of the dual problem:

$$r(x) = c - A^T w(x).$$

Note that $w(x)$ is dual feasible if and only if $r(x) \geq 0$. The vector $r(x)$ is called the vector of *reduced costs*. The third function $y: \Omega^0 \rightarrow \Omega$ is given by

$$(1) \quad y(x) = x - \frac{D_x^2 r(x)}{|D_x r(x)|}.$$

The affine-scaling algorithm is defined in terms of the function y :

$$x^{k+1} = \begin{cases} y(x^k), & x^k \in \Omega^0 \\ x^k, & x^k \in \partial\Omega. \end{cases}$$

The algorithm also generates a sequence of dual variables

$$w^k = \begin{cases} w(x^k), & x^k \in \Omega^0 \\ w^{k-1}, & x^k \in \partial\Omega \end{cases}$$

and a sequence of reduced costs

$$r^k = c - A^T w^k.$$

For notational convenience, put $D_k = D_{x^k}$. Note that if the sequence x^k hits the boundary of the polytope, then it becomes fixed at the point where it first hits. In contrast, the dual variables become fixed at the values associated with the last interior point before the boundary was hit.

This affine-scaling algorithm differs slightly from the ones studied in [4] and [13]. The difference is in the step size. In words, the algorithm studied here steps 100 percent of the way to the surface of the inscribed ellipsoid (see [4] for the definition of this ellipsoid), whereas the algorithm in [4] steps only a certain fraction α of the way to the surface. Hence, while the algorithm presented here can stop in a finite number of steps, the one in [4] always involves an asymptotic approach to the optimal solution. In contrast, the algorithm presented in [13] steps a certain fraction of the way to the nearest face. This means that the mapping $\gamma(x)$ defined by (1) has to be changed to

$$\gamma^{[13]}(x) = \frac{x - \alpha D_x^2 r(x)}{\gamma(x)},$$

where

$$\gamma(x) = \max_j x_j r_j(x).$$

It is easy to see that

$$\gamma(x) \leq |D_x r(x)|_\infty \leq |D_x r(x)|$$

(assumption (1) below implies that $\gamma(x) > 0$). Hence, the step length chosen in [13] is the longest, followed by the L^∞ norm, followed by the conservative L^2 step length studied here. It should be noted that all implementations use the step length described in [13], and so, in a sense, convergence proofs for that case are the most interesting. We do not study that case here, but it is easy to peruse the proof given here for the L^2 case and see that it can be easily modified to cover the L^∞ case as long as we also introduce a contraction factor $\alpha < 1$.

The following assumptions are made:

1. $-\infty < \min_\Omega c \cdot x < \max_\Omega c \cdot x$,
2. Ω^0 is nonempty; $x^0 \in \Omega^0$ is given.
3. A has full-row rank.
4. $AD_x^2 A^T$ is invertible for all $x \in \Omega$.

Assumption (4) is called *primal nondegeneracy*. Note that assumption (3) implies that $AD_x^2 A^T$ is invertible for all $x \in \Omega^0$. Hence, the primal nondegeneracy assumption is really an assumption about the boundary of Ω . Also note that the primal nondegeneracy assumption implies that the domain of the function w can be extended to all of Ω . Also, assumption (1) implies

that $|D_x r(x)| \neq 0$ for all $x \in \Omega^0$, which in turn implies that $y(x)$ is well defined.

The main result is

THEOREM. *The sequence x^k converges to a primal feasible point \bar{x} . The sequence w^k converges to a dual feasible point \bar{w} . The pair consisting of the limiting primal variables \bar{x} and the limiting reduced costs $\bar{r} = c - A^T \bar{w}$ satisfy strong complementarity:*

$$(2) \quad \bar{x}_j = 0 \quad \text{if and only if} \quad \bar{r}_j > 0.$$

In the next section we prove this theorem assuming the polytope is bounded. Then, in Section 4, we remove the boundedness assumption.

3. Proof assuming compactness

In this section, we assume that Ω is bounded. For this case, we break the proof up into a series of steps.

Step 1. Primal feasibility is preserved. It is easy to check that $AD_x^2 r(x) = 0$, and therefore, since $Ax = b$, we see that

$$Ay(x) = Ax - \frac{AD_x^2 r(x)}{|D_x r(x)|} = b.$$

From the definition of $y(x)$, we see that the j th component is given by

$$(3) \quad y_j(x) = x_j \left(1 - \frac{x_j r_j(x)}{|D_x r(x)|} \right).$$

The subtracted term above must lie between -1 and 1 , and hence

$$(4) \quad 0 \leq y_j(x) \leq 2x_j.$$

Step 2. Hitting a face implies optimality. Suppose that $x^k \in \Omega^0$ and that $x_j^{k+1} = 0$. Then (3) implies that $x_j^k r_j^k = |D_x r^k|$. Hence, since $x_i^k > 0$ for all i , we see that $r_j^k > 0$ and $r_i^k = 0$ for all $i \neq j$. Since $\bar{x} = x^{k+1}$ and $\bar{w} = w^k$, we see now that (2) holds. Henceforth, we assume that $x^k \in \Omega^0$ for all k .

Step 3. The objective function decreases. We begin with the obvious:

$$c \cdot x - c \cdot y(x) = \frac{c \cdot D_x^2 r(x)}{|D_x r(x)|}.$$

Next, note that $D_x r(x) = P_x D_x c$, where $P_x = I - D_x A^T (AD_x A^T)^{-1} AD_x$ denotes the projection onto the null space of AD_x . Hence,

$$c \cdot x - c \cdot y(x) = \frac{c \cdot D_x P_x D_x c}{|D_x r(x)|}$$

and, since P_x is a projection matrix, it is idempotent and symmetric, and so

$$c \cdot x - c \cdot y(x) = [P_x D_x c] = |D_x r(x)|.$$

Therefore,

$$c \cdot x^k - c \cdot x^{k+1} = |D_k r^k|.$$

Step 4. Complementary slackness holds in the limit. Since $c \cdot x^k$ converges (it is decreasing and bounded below), it follows from the previous step that

$$|D_k r^k| \rightarrow 0.$$

Step 5. The dual variables are bounded. Assumption (4) implies that $w(x)$ is a continuous function on the compact set Ω ; hence $w(x)$ is bounded.

We now introduce some auxiliary notation. Let \bar{w} be a limit point of w^k . Let k_p , $p = 1, 2, \dots$, denote a subsequence along which w^k converges to \bar{w} . Put $\bar{r} = c - A^T \bar{w}$, and let

$$B = \{j: \bar{r}_j = 0\}, \quad N = \{j: \bar{r}_j \neq 0\}.$$

The indices in B are called the *basic* indices and those in N are called *nonbasic*. Abusing notation, we use the same letters to denote the partition of A into its basic and nonbasic parts:

$$A = [B \mid N].$$

We denote the corresponding partitioning of n -vectors using subscripts:

$$x = \begin{bmatrix} x_B \\ x_N \end{bmatrix}.$$

Step 6. $\lim_{x_N \rightarrow 0} r(x) = \bar{r}$. First note that

$$\begin{aligned} r(x) &= [I - A^T (AD_x^2 A^T)^{-1} AD_x^2] c \\ &= [I - A^T (AD_x^2 A^T)^{-1} AD_x^2] \bar{r}, \end{aligned}$$

where the second equality follows from the definition of \bar{r} and the fact that the bracketed matrix annihilates A^T . Rearranging the last equation and using the fact that the basic components of \bar{r} vanish, we get

$$\bar{r} - r(x) = A^T (AD_x^2 A^T)^{-1} N D_{x_N}^2 \bar{r}_N.$$

Now, since $(AD_x^2 A^T)^{-1}$ exists and is continuous throughout the compact set Ω , it follows that it is bounded. Hence as x_N tends to zero the $D_{x_N}^2$ factor dominates and drives the difference to zero.

Step 7. The nonbasic components of x convergence to zero. Fix $j \in N$. By Step 4, we know that $x_j^k r_j^k$ tends to zero. We also know that r_j^k tends along the subsequence k_p to \bar{r}_j , which is a nonzero number. Hence, x_j^k must

tend to zero along the subsequence k_p . Since this is true for any nonbasic index j , we see that

$$x_N^{k_p} \rightarrow 0.$$

Now suppose that x_N^k does not tend to zero along the entire sequence; i.e., there exists a $\delta > 0$ such that $x_N^k \in C_\delta^c$ infinitely often, where $C_\delta = \{x_N: 0 \leq x_i < \delta \text{ for all } i \in N\}$. (Given such a δ , any smaller value will also work.) Fix δ small enough that C_δ^c is visited infinitely often and

$$(5) \quad |r_j^{k'_p}| \geq \frac{1}{2} |\bar{r}_j|$$

for all j and for all x such that $x_N \in C_\delta$. This, of course, is possible because of the previous step. Let $K = \{k: x_N^k \in C_\delta \text{ and } x_N^{k+1} \leq C_\delta^c\}$. Then clearly K is an infinite set. Now let $K_j = \{k \in K: x_j^{k+1} \geq \delta\}$. Then $K = \bigcup_{j \in N} K_j$. Since N is finite, there exists a j such that K_j is infinite. Let k'_p be the subsequence that enumerates K_j . Then

$$x_N^{k'_p} < \delta \quad \text{and} \quad x_j^{k'_p+1} \geq \delta.$$

Using formula (4), the second statement above implies that $x_j^{k'_p} \geq \delta/2$. From (5) we see that

$$|x_j^{k'_p} r_j^{k'_p}| \geq \frac{1}{4} \delta |\bar{r}_j| > 0$$

for all p . This contradicts Step 4 and so we are led to conclude that x_N^k actually tends to zero.

Step 8. The dual variables converge to \bar{w} . As in Step 6, we start by noting that

$$\begin{aligned} w^k - \bar{w} &= (AD_k^2 A^T)^{-1} AD_k^2 (c - A^T \bar{w}) \\ &= (AD_k^2 A^T)^{-1} AD_k^2 \bar{r} \\ &= (AD_k^2 A^T)^{-1} ND_{x_N^k}^2 \bar{r}_N. \end{aligned}$$

Again, we use the fact that $(AD_x^2 A^T)^{-1}$ is bounded to conclude that the difference converges to zero, since x_N^k is now known to converge to zero.

Step 9. The limiting nonbasic reduced costs are positive. Suppose there exists a $j \in N$ such that $\bar{r}_j < 0$. Since $w^k \rightarrow \bar{w}$, there exists a K such that $r_j^k < 0$ for all $k \geq K$. Hence, we see that

$$x_j^{k+1} = x_j^k \left(1 - \frac{x_j^k r_j^k}{|D_k r^k|} \right) > x_j^k \quad \text{for all } k \geq K.$$

This contradicts Step 7.

Step 10. The basic components of x^k converge (say to \bar{x}_B). From the definition of the algorithm, we have

$$(6) \quad x_j^{k+1} - x_j^k = -\frac{(x_j^k)^2 r_j^k}{|D_k r^k|}.$$

Therefore, x_j^k converges if

$$\sum_{k=0}^{\infty} \frac{(x_j^k)^2 |r_j^k|}{|D_k r^k|} < \infty.$$

For $j \in B$,

$$r_j(x) = r_j(x) - \bar{r}_j = e_j^T A^T (AD_x^2 A^T)^{-1} N D_{x_N}^2 \bar{r}_N.$$

Hence,

$$x_j^2 r_j = \sum_{l \in N} \sigma_{jl}(x) x_l^2 \bar{r}_l,$$

where

$$\sigma_{jl}(x) = e_j^T D_x^2 A^T (AD_x^2 A^T)^{-1} N e_l.$$

Continuity and compactness now imply that there are bounding constants $\bar{\sigma}_{jl}$: $|\sigma_{jl}(x)| \leq \bar{\sigma}_{jl}$ for all $x \in \Omega$. Therefore,

$$(7) \quad \sum_{k=0}^{\infty} \frac{(x_j^k)^2 |r_j^k|}{|D_k r^k|} \leq \sum_{l \in N} \bar{\sigma}_{jl} \sum_{k=0}^{\infty} \frac{(x_l^k)^2 \bar{r}_l}{|D_k r^k|}.$$

Step 7 and equation (6) imply that, for any $l \in N$,

$$\sum_{k=0}^{\infty} \frac{(x_l^k)^2 r_l^k}{|D_k r^k|}$$

converges. Since $r_l^k \rightarrow \bar{r}_l$, we see that the right-hand side in (7) is finite.

Step 11. Strong complementarity holds. We only need to show that $\bar{x}_B > 0$. Fix $j \in B$. To show that $\bar{x}_j > 0$, it suffices to show that

$$\sum_{k=0}^{\infty} \frac{\log x_j^{k+1}}{x_j^k}$$

converges absolutely. Clearly, this sum converges absolutely if and only if

$$\sum_{k=0}^{\infty} \frac{x_j^{k+1} - x_j^k}{x_j^k}$$

converges absolutely. Using (3), we see that

$$\sum_{k=0}^{\infty} \frac{x_j^{k+1} - x_j^k}{x_j^k} = \sum_{k=0}^{\infty} \frac{x_j^k r_j^k}{|D_k r^k|}.$$

To see that this last sum converges absolutely, we use exactly the same argument as in the previous step, except that instead of σ_{jl} we get something similar.

$$\rho_{jl}(x) = e_j^T D_x A^T (AD_x^2 A^T)^{-1} N e_l.$$

In the previous section, we found a bound for $\sigma_{jl}(x)$ that was valid throughout Ω . Here we settle for slightly less: we only need to show that $\rho_{jl}(x)$ is bounded along the sequence x^k . This follows from the fact that x_j^k converges and $(AD_x^2 A^T)^{-1}$ is bounded. This completes the proof for the case where Ω is assumed to be bounded.

4. Proof without compactness

The boundedness assumption (in conjunction with primal nondegeneracy) was used in three places:

1. In Step 5, showing that the dual variables function $w(x)$ is bounded.
2. In Steps 6, 8, and 11, showing that $(AD_x^2 A^T)^{-1}$ is bounded.
3. In Step 10, showing that the function $\sigma_{jl}(x)$ is bounded.

We will now show that the boundedness assumption is not necessary for any of these three statements to hold and that primal nondegeneracy is not necessary for the first and third.

We begin with the claim that $(AD_x^2 A^T)^{-1}$ is bounded. Let $\widehat{\Omega}$ denote the compact set consisting of the convex hull of the extreme points of Ω . Then any point x in Ω can be written as $x = \widehat{x} + \widehat{t}$, where \widehat{x} belongs to $\widehat{\Omega}$ and $\widehat{t} \geq 0$, $A\widehat{t} = 0$. Hence, $AD_x^2 A^T \geq AD_{\widehat{x}}^2 A^T$ in the sense that the difference is positive semidefinite. Therefore, $\lambda_1(x) \geq \lambda_1(\widehat{x})$, where $\lambda_1(x)$ denotes the smallest eigenvalue of $AD_x^2 A^T$. Primal nondegeneracy and compactness of $\widehat{\Omega}$ imply that $\lambda_1(\widehat{x}) \geq \bar{\lambda}_1 > 0$ for all $\widehat{x} \in \widehat{\Omega}$. Since the L^2 norm of $(AD_x^2 A^T)^{-1}$ is exactly $1/\lambda_1(x)$, we see that

$$\|(AD_x^2 A^T)^{-1}\|_2 = \frac{1}{\lambda_1(x)} \leq \frac{1}{\lambda_1(\widehat{x})} \leq \frac{1}{\bar{\lambda}_1} < \infty.$$

Note that even though we have removed the boundedness assumption on Ω , we still have used primal nondegeneracy. This is essential, as the following example shows:

$$\begin{aligned} &\min x_1 \\ &x_1 - x_2 = 0 \\ &x_1 \geq 0, \quad x_2 \geq 0. \end{aligned}$$

Clearly, $(x_1, x_2) = (0, 0)$ is a point in the polytope at which $(AD_x^2 A^T)^{-1} = 1/(x_1^2 + x_2^2)$ blows up.

The other two claims follow immediately from the following statement:

LEMMA. For any n -vector γ , the function $w^\gamma(x) = (AD_x^2 A^T)^{-1} AD_x^2 \gamma$ is bounded.

To prove this lemma, we use Cramer's rule and the Cauchy-Binet theorem to write the i th component as follows:

$$w_i^\gamma = \frac{\sum_{1 \leq j_1 < \dots < j_m \leq n} (x_{j_1} \dots x_{j_m})^2 \det_{j_1, \dots, j_m}(a_1, \dots, a_m) \det_{j_1, \dots, j_m}(a_1, \dots, \overset{\text{ith position}}{\downarrow} \gamma, \dots, a_m)}{\sum_{1 \leq j_1 < \dots < j_m \leq n} (x_{j_1} \dots x_{j_m})^2 [\det_{j_1, \dots, j_m}(a_1, \dots, a_m)]^2},$$

where $\det_{j_1, \dots, j_m}(a_1, \dots, a_m)$ denotes the determinant of the $m \times m$ matrix obtained by selecting columns j_1, \dots, j_m from the $m \times n$ matrix whose rows are the n -vectors $\alpha_1, \dots, \alpha_m$ and where a_i denotes the i th row of A . To bound this expression, we use the following simple inequality:

$$\frac{\left| \sum_{i=1}^N r_i \right|}{\sum_{i=1}^N s_i} \leq \max_i \frac{|r_i|}{s_i},$$

which is valid for any pair of sequences with $s_i > 0$. Since terms in the numerator in (8) vanish whenever terms in the denominator vanish, we can use this simple inequality to obtain the following bound:

$$|w_i^\gamma(x)| \leq \max_{\substack{1 \leq j_1 < \dots < j_m \leq n \\ \text{denom} \neq 0}} \left| \frac{\det_{j_1, \dots, j_m}(a_1, \dots, \gamma, \dots, a_m)}{\det_{j_1, \dots, j_m}(a_1, \dots, a_m)} \right|.$$

The right-hand side is independent of x and hence furnishes a bound.

5. Dikin's proof

The ideas, methods, and results presented in this paper stem directly from Dikin's work. However, there are a few minor gaps in Dikin's paper that have been rectified here. The first is that he claimed that the limiting nonbasic reduced costs are positive (Step 9) before he showed that the dual variables converge (Step 8). His argument is extremely terse and combined with his proof that the nonbasic components of x converge to zero (Step 7), so it seems that at the very least he was leaving a lot to the reader. In any case, the proof of Step 9 here certainly depends on the result of Step 8. Mike Todd has a proof for Step 9 that does not depend on Step 8, but the proof given here is easier.

The other difficulty is that in Steps 6, 8, and 11, where the boundedness of $(AD_x^2 A^T)^{-1}$ is needed, Dikin argues that the desired result follows from Cramer's rule and the Cauchy-Binet theorem. We do not see how it is possible to establish the boundedness of this matrix using that approach.

REFERENCES

1. I. Adler, N. K. Karmarkar, and M. G. C. Resende, *An implementation of Karmarkar's algorithm for linear programming*, Math. Programming **44**, (1989), 297-335..
2. I. Adler, N. K. Karmarkar, M. G. C. Resende, and G. Veiga, *Data structures and programming techniques for the implementation of Karmarkar's algorithm*, ORSA J. Comp. **1** (1989), 84-106.
3. I. Adler and R. D. C. Monteiro, *Limiting behavior of the affine scaling continuous trajectories for linear programming problems*, this volume, pp. 189-211.
4. E. R. Barnes, *A variation on Karmarkar's algorithm for solving linear programming problems*, Math. Programming **36** (1986), 174-182.
5. I. I. Dikin, *Iterative solution of problems of linear and quadratic programming*, Soviet Math. Dokl. **8** (1967), 674-675.
6. —, *On the speed of an iterative process*, Upravlyaemye Sistemi **12** (1974), 54-60.
7. N. K. Karmarkar, *A new polynomial-time algorithm for linear programming*, Combinatorica **4** (1984), 373-395.
8. N. Megiddo and M. Shub, *Boundary behavior of interior-point algorithms in linear programming*, Math. Oper. Res. **14** (1989), 97-146.
9. C. L. Monma and A. J. Morton, *Computational experiments with a dual affine variant of Karmarkar's method for linear programming*, Oper. Res. Lett. **6** (1987), 261-267.
10. R. D. C. Monteiro, I. Adler, and M. G. C. Resende, *A polynomial-time primal-dual affine scaling algorithm for linear and convex quadratic programming and its power series extension*, Technical Report, Department of Industrial Engineering and Operations Research, University of California, Berkeley, Calif., 1988.
11. J. A. Tomlin, *An experimental approach to Karmarkar's projective method for linear programming*, Technical Report, Ketron, Inc., Mountain View, CA, 1985.
12. J. A. Tomlin and J. S. Welch, *Implementing an interior-point method in a mathematical programming system*, Technical Report, Ketron, Inc., Mountain View, Calif., 1986.
13. R. J. Vanderbei, M. S. Meketon, and B. A. Freedman, *A modification of Karmarkar's linear programming algorithm*, Algorithmica **1** (1986), 395-407.

AT&T BELL LABORATORIES, MURRAY HILL, NEW JERSEY 07974

E-mail address, R. J. Vanderbei: rvdb@research.att.com

E-mail address, J. C. Lagarias: jcl@research.att.com