Two-Thirds is Sharp for Affine Scaling

Leslie A. Hall Robert J. Vanderbei

Program in Statistics & Operations Research
Princeton University
Princeton, NJ 08544

September 1992

Technical Report SOR-92-9

Abstract

Tsuchiya and Muramatsu recently proved that the affine-scaling algorithm for linear programming generates convergent sequences of primal and dual variables whose limits are optimal for the corresponding primal and dual problems as long as the step size is no more than two-thirds of the distance to the nearest face of the polytope. An important feature of this result is that it does not require any nondegeneracy assumptions. In this paper we show that Tsuchiya and Muramatsu's result is sharp by providing a simple linear programming problem for which the sequence of dual variables fails to converge for every step size greater than two-thirds.

Key words: linear programming theory, linear programming algorithms, affine-scaling algorithms, fixed points

1 Introduction.

Consider the following primal-dual pair of linear programs

$$\begin{aligned} \text{Minimize} \quad & c^T x \\ \text{subject to} \quad & Ax = b \\ & x \geq 0, \end{aligned} \tag{P}$$

and

$$\begin{aligned} \text{Maximize} \quad & b^T y \\ \text{subject to} \quad & A^T y \leq c. \end{aligned} \tag{D}$$

The affine-scaling algorithm is defined as follows. Let

$$y(x) = (AD_x^2 A^T)^{-1} AD_x^2 c,$$

$$r(x) = c - A^T y(x),$$

and

$$T(x) = x - \gamma \frac{D_x^2 r(x)}{\max_j \{x_j r_j(x)\}},$$

where $D_x = \text{diag}(x)$ and γ is a fixed number between zero and one. The vector-valued function y is called the vector of *dual estimates*, r is called the vector of *reduced costs*, and T is the new primal solution. The algorithm starts with an arbitrary interior feasible point x^0 and generates a sequence of primal feasible solutions by repeated application of T:

$$x^{k+1} = T(x^k).$$

Associated with each iterate x^k of the primal solution is a corresponding dual estimate $y^k := y(x^k)$.

In a recent paper [2], Tsuchiya and Muramatsu study this algorithm under very weak assumptions. Namely, they assume that the primal and the dual are both feasible, that there is a strictly interior primal feasible solution, and that the constraint matrix A has full row rank. They then show that if $\gamma \leq 2/3$, x^k converges to a point in the relative interior of the primal optimal face and y^k converges to the analytic center of the dual optimal face. The importance of this result is that it makes no assumptions about primal and dual nondegeneracy. In a recent talk [1], Tsuchiya conjectured that $\gamma = 2/3$ is sharp. We prove this conjecture by exhibiting a linear program for which the sequence of dual estimates fails to converge for every $\gamma > 2/3$. (Tsuchiya and Muramatsu [2] give an example where the duals fail to converge to the analytic center of the dual optimal face. However, in contrast to our example, their sequence of duals is convergent.)

It is interesting to contrast this result with the analogous result obtained when primal and dual nondegeneracy are assumed. In that case it is known [3] that the primal and dual variables are convergent for all $\gamma < 1$.

2 The Example.

Consider the following linear program

$$\begin{array}{rl} \text{Minimize} & x_1 + x_2 + x_3 \\ \text{subject to} & x_1 + x_2 - x_3 - x_4 = 0 \\ & x \geq 0. \end{array}$$

The dual linear program is

$$\begin{array}{rl} \text{Maximize} & 0 \\ \text{subject to} & 0 \leq y \leq 1. \end{array}$$

Clearly, the optimal solution to the primal linear program is $x^* = 0$, and every feasible solution to the dual is optimal.

For this problem,

$$y(x) = \frac{x_1^2 + x_2^2 - x_3^2}{x_1^2 + x_2^2 + x_3^2 + x_4^2}$$

and

$$r(x) = \frac{1}{\sum_{i=1}^{4} x_i^2} \begin{bmatrix} 2x_3^2 + x_4^2 \\ 2x_3^2 + x_4^2 \\ 2x_1^2 + 2x_2^2 + x_4^2 \\ x_1^2 + x_2^2 - x_3^2 \end{bmatrix}.$$

We restrict our attention to the plane $x_1 = x_2$. Symmetry dictates and algebra proves that if we start on this plane, all iterates will lie on this plane. Let $z(x) = x_3/x_1$. For notational convenience, we generally omit explicit indication that z depends on x. Note that for feasible primal solutions, $0 \leq z \leq 2$. Simple calculation shows that

$$y(x) = \frac{2 - z^2}{2 + z^2 + (2 - z)^2} \tag{1}$$

and

$$\frac{D_x^2 r(x)}{\max_j\{x_j r_j(x)\}} = x_1 d(z)/\alpha(z),$$

where

$$d(z) = \begin{bmatrix} 3z^2 - 4z + 4 \\ 3z^2 - 4z + 4 \\ z^2(z^2 - 4z + 8) \\ (2 - z)^2(2 - z^2) \end{bmatrix}$$

and

$$\alpha(z) = \max\{3z^2 - 4z + 4,\ z(z^2 - 4z + 8),\ (2 - z)(2 - z^2)\}. \tag{2}$$

Let $a(z)$, $b(z)$, and $c(z)$ denote the three polynomial expressions in (2) and let $z_c = (5 - \sqrt{17})/2 \approx 0.4385$. The following assertions are easy to check:

$$a(z_c) = b(z_c) = c(z_c),$$

and

$$\alpha(z) = \begin{cases} b(z), & z_c < z \leq 2, \\ c(z), & 0 \leq z < z_c. \end{cases}$$

Application of the mapping T generates a new x-vector lying on the plane $x_1 = x_2$ and a corresponding new z-value, which can be computed directly from the old value as follows:

$$z^{k+1} = F(z^k),$$

where

$$F(z) = \begin{cases} z\,\dfrac{1-\gamma}{1-\gamma a(z)/b(z)}, & z \geq z_c, \\ \\ z\,\dfrac{1-\gamma b(z)/c(z)}{1-\gamma a(z)/c(z)}, & z \leq z_c. \end{cases}$$

The function F has the following properties:

- F is a map from $[0, 2]$ onto itself,
- F is C^1,
- $F(z_c) = z_c$,
- F is C^∞ and $F'' < 0$ on $\{z : 0 < z < z_c\}$,
- F is C^∞ and $F'' > 0$ on $\{z : z_c < z < 2\}$.

The sequence of iterates of such a function converges (to z_c) if and only if $F'(z_c) \geq -1$. The "only if" direction is easy; a simple geometric argument can be used to prove the converse. Indeed, if $F'(z) \geq 0$ for all $z \in (0, 2)$, the result is straightforward to prove. Otherwise, consider the largest interval I of the form $(z_c - \delta, z_c + \delta)$ with the property that $F'(z) < 0$ for all $z \in I$. Observe that F restricted to I is a strict contraction and hence iteration sequences starting in I will converge to z_c. By the smoothness of F, either $F'(z_c - \delta) = 0$ or $F'(z_c + \delta) = 0$; for the sake of the argument, we assume the former. It is now an easy exercise to show that:

- iteration sequences starting in $(0, z_c - \delta]$ increase monotonically while in $(0, z_c - \delta]$ and eventually enter I;

- iteration sequences starting in $[z_c+\delta, 2)$ decrease monotonically while in $[z_c+\delta, 2)$ and eventually enter either I or $(0, z_c - \delta]$, in which case we are reduced to the previous situation.

Proving these assertions suffices to prove the result.

Using the fact that $a(z_c) = b(z_c) = c(z_c)$, we find that

$$F'(z_c) = \frac{1 - \gamma + \gamma z_c \dfrac{a'(z_c) - b'(z_c)}{a(z_c)}}{1 - \gamma}.$$

Now, from the fact that

$$z_c \frac{a'(z_c) - b'(z_c)}{a(z_c)} = -1,$$

it follows that

$$F'(z_c) = \frac{1 - 2\gamma}{1 - \gamma}$$

and thus

$$F'(z_c) \geq -1 \text{ if and only if } \gamma \leq \frac{2}{3}.$$

Finally, we need to show that the nonconvergence of z^k implies the nonconvergence of y^k. We see from (1) that the dual variable y is only a function of z. If this function were one-to-one, then we would be done. But in fact it is not. Indeed, it is easy to check that y thought of as a function of z has a single critical point in the interval $[0, 2]$, which is a maximum, at z_c. Hence, there are at most two values of z that map to a given y, one of them less than z_c and the other larger than z_c. The only way that y^k could converge is if z^k had at most two limit points and these limit points mapped to the same y-value. We show that this does not happen as follows. First, we pick an arbitrary value of y for which there are two preimages z_+ and z_-. Explicit formulas for z_+ and z_- are

$$z_+(y) = \frac{2y + R(y)}{1 + 2y}$$

and

$$z_-(y) = \frac{2y - R(y)}{1 + 2y}$$

where $R(y) = \sqrt{2(1 - y - 4y^2)}$ and $1/3 \le y \le (\sqrt{17} - 1)/8$. Convergence of the y^k to y could happen only if $F(z_+(y)) = z_-(y)$ and $F(z_-(y)) = z_+(y)$. Using the definition of F, these two conditions can be rewritten as follows:

$$\gamma = \frac{z_+ - z_-}{z_+ - z_- \frac{a(z_+)}{b(z_+)}}$$

and

$$\gamma = \frac{z_+ - z_-}{z_+ \frac{a(z_-)}{c(z_-)} - z_- \frac{b(z_-)}{c(z_-)}}$$

where we have dropped explicit mention of the dependence of z_+ and z_- on y. Therefore, convergence to y can happen only if these two expressions for γ agree. However, simple but tedious calculations show that

$$z_+ - z_- \frac{a(z_+)}{b(z_+)} < z_+ \frac{a(z_-)}{c(z_-)} - z_- \frac{b(z_-)}{c(z_-)}$$

for all $y \in [1/3, (\sqrt{17} - 1)/8]$. Hence, it is impossible for y^k to converge to any of its possible limits.

As a final remark, we note that z^k does indeed have exactly two limit points for all $\gamma > 2/3$. Numerical experiments show that, for $\gamma < 0.995$, both limit points correspond to feasible duals, but for $\gamma > 0.995$, one corresponds to a feasible dual and the other to an infeasible one.

References

[1] T. Tsuchiya. Recent development on the global convergence analysis of long-step affine scaling algorithms. Talk given at AT&T Bell Labs, Murray Hill, August 1992.

[2] T. Tsuchiya and M. Muramatsu. Global convergence of a long-step affine scaling algorithm for degenerate linear programming problems. Technical Report 423, The Institute of Statistical Mathematics, 4-6-7 Minami-Azabu, Minato-ku, Tokyo, 106, Japan, January 1992. Revised September 1992.

[3] R. Vanderbei, M. Meketon, and B. Freedman. A modification of Karmarkar's linear programming algorithm. *Algorithmica*, 1:395–407, 1986.