数理解析研究所講究録
1004 巻 1997 年 128-137

Affine Scaling Algorithm Fails For Semidefinite Programming

Masakazu MURAMATSU[1]
Sophia University

Abstract

In this paper, we introduce an affine scaling algorithm for semidefinite programming, and give an example of a semidefinite program such that the affine scaling algorithm converges to a non-optimal point. Both our program and its dual have interior feasible solutions, and unique optimal solutions which satisfy strict complementarity, and they are nondegenerate everywhere.

1 Introduction

When both primal and dual problems have interior feasible solutions, semidefinite programming (SDP) has remarkable resemblance with linear programming (LP), e.g., both problems have optimal solutions and satisfy strong duality. In this case, it is known that several interior point methods for LP and their polynomial convergence analysis can be naturally extended to SDP (Alizadeh [1], Alizadeh, Haeberly and Overton [2], Helmberg, Rendl, Vanderbei and Wolkowicz [9], Jarre [10], Kojima, Shindoh, and Hara [12], Kojima, Shida and Shindoh [13], Lin and Saigal [14], Lou, Sturm and Zhang [15], Monteiro [16], Monteiro and Zhang [18], Nesterov and Nemirovskii [19, 20], Nesterov and Todd [21, 22], Potra and Sheng [23], Sturm and Zhang [24], Vandenberghe and Boyd [28], and Zhang [30]).

The affine scaling algorithm for LP is originally proposed by Dikin [5], and rediscovered by Barnes [4], and Vanderbei, Meketon and Freedman [29] after Karmarkar [11] proposed the first polynomial-time interior point method . The affine scaling algorithm has been widely implemented and extensively studied, but unlike many other interior point methods, the question of whether the affine scaling algorithm is polynomially convergent is still an open problem. The strongest convergence result so far is due to Tsuchiya and Muramatsu [27], who establish global convergence of the affine scaling algorithm where the step is taken as a fixed fraction less than 2/3 of the whole step to the boundary of the feasible region. Simpler proofs of this global convergence results can also be found in Monteiro, Tsuchiya and Wang [17], and Saigal [25].

The affine scaling algorithm can also be naturally extended to SDP. Faybusovich [6] investigated the affine scaling vector field for SDP, Faybusovich [7] proposed the discrete version of the affine sclaing algorithm, and Goldfarb and Scheinberg [8] proved the global convergence of the associated continuous trajectories. So far, there exists no convergence analysis for the discrete version of the affine scaling algorithm for SDP.

In this paper, we give an instance of SDP such that the affine scaling algorithm converges to a non-optimal point. We prove that for both the short and the long step version of the affine scaling algorithm, there exists a region of starting points such that the generated sequence converges to a non-optimal point. Our program and its dual have interior feasible solutions, unique optimal solutions which satisfy strict complementarity, and both programs are nondegenerate everywhere. (For degeneracy in SDP, see Alizadeh, Haeberly, and Overton [3].)

This paper is organized as follows. In Section 2, we introduce the affine scaling algorithm for SDP. In Section 3, we give an instance of SDP and prove that the affine scaling algorithm converges to a non-optimal point. In Section 4, we give some concluding remarks.

2 The Affine Scaling Algorithm for SDP

Let $\mathcal{S}(n)$ denote the set of $n \times n$ real symmetric matrices. Consider the SDP problem

$$\min C \bullet X \text{ subject to } A_i \bullet X = b_i, \quad i = 1, \ldots, m, \quad X \succeq 0, \tag{1}$$

[1] 7-1 Kioi-cho, Chiyoda-ku, Tokyo 102 Japan. muramatume.sophia.ac.jp

where $C, X, A_i \in \mathcal{S}(n)$, the operator $\bullet$ denotes the standard inner product in $\mathcal{S}(n)$, i.e., $C \bullet X \triangleq \mathrm{tr}(CX) = \sum_{i,j} C_{ij}X_{ij}$, and $X \succeq 0$ means that X is positive semidefinite. The induced norm $\|X\|_F \triangleq \sqrt{X \bullet X} = \sqrt{\mathrm{tr}(X^tX)}$ is called the *Frobenius norm*. The dual of (1) is

$$\max b^t u \text{ subject to } S + \sum_{i=1}^{m} u_i A_i = C, \quad S \succeq 0. \tag{2}$$

We assume that an interior (i.e., positive definite) feasible point of the primal problem (1) exists, and the matrices $A_i, i = 1, \ldots, m$ are linearly independent.

There are several equivalent characterization of the affine scaling direction for LP. Similarly, we can define the affine scaling direction for SDP in different ways. Following the original derivation proposed by Dikin [5], here we define the affine scaling direction for SDP by first defining the associated dual estimate. For a detailed motivation of the affine scaling algorithm for LP, we recommend the textbook by Saigal [26] which deals with this algorithm extensively.

Given an interior feasible solution $X \succ 0$, we define the dual estimate $(u(X), S(X))$ as the unique solution of the following optimization problem:

$$\min \|X^{1/2}SX^{1/2}\|_F^2 \text{ subject to } S + \sum_i u_i A_i = C. \tag{3}$$

Solving the Karush-Kuhn-Tucker condition, we have the explicit formula

$$u(X) = G(X)^{-1}p(X), \quad S(X) = C - \sum_i u_i(X)A_i, \tag{4}$$

where $G(X) \in \mathcal{S}(m)$ and $p(X) \in R^m$ are such that $G_{ij}(X) = \mathrm{tr}(A_iXA_jX)$ and $p_j(X) = \mathrm{tr}(A_jXCX)$ for all $i, j = 1, \ldots, m$, respectively. Here, the linear independence of the A_j's ensures that $G(X)$ is invertible. In fact, we have the following lemma.

Proposition 1 *If $X \succ 0$, then $G(X) \succ 0$.*

Proof : First, we note that for any matrices M_1, M_2, M_3 which have appropriate sizes, $\mathrm{tr}(M_1M_2M_3) = \mathrm{tr}(M_2M_3M_1) = \mathrm{tr}(M_3M_1M_2)$ holds.

For any $v \in R^m$, we have

$$\begin{aligned} v^tGv &= \sum_{i,j} G_{ij}v_iv_j \\ &= \sum_{i,j} \mathrm{tr}(A_iXA_jX)v_iv_j \\ &= \mathrm{tr}((\sum_i v_iA_i)X(\sum_j v_jA_j)X) \\ &= \mathrm{tr}(X^{1/2}(\sum_i v_iA_i)X(\sum_j v_jA_j)X^{1/2}) \\ &= \|X^{1/2}(\sum_i v_iA_i)X^{1/2}\|_F^2 \geq 0, \end{aligned} \tag{5}$$

in which the equality holds if and only if $\sum_i v_iA_i = 0$. The linear independence of A_i's implies $v = 0$. □

In (3), any decomposition of $X = RR^t$ can be used instead of $X^{1/2}$, because $\|R^tSR\|_F^2 = \mathrm{tr}(SXSX) = \|X^{1/2}SX^{1/2}\|_F^2$. Namely, the dual estimate is independent of the choice of decomposition. We however use $X^{1/2}$ for notational simplicity.

The affine scaling direction $D(X)$ is defined as

$$D(X) \triangleq XS(X)X = X(C - \sum_i u_i(X)A_i)X. \tag{6}$$

The following proposition gives some properties of $D(X)$.

Proposition 2 *We have*

$$A_i \bullet D(X) = 0 \tag{7}$$

for all $i = 1, \ldots, m$, *and*

$$C \bullet D(X) = \|X^{1/2} S X^{1/2}\|_F^2. \tag{8}$$

Proof: The first equation follows from

$$A_j \bullet D(X) = \mathrm{tr}(A_j X(C - \sum_{i=1}^{m} u_i(X) A_i) X) = p_j - \sum_{i=1}^{m} G_{ji} u_i(X) = 0. \tag{9}$$

Hence,

$$\mathrm{tr}(X^{1/2}(C - \sum_i u_i(X) A_i) X A_j X^{1/2}) = 0 \tag{10}$$

holds for all j, and we have

$$\begin{aligned} C \bullet D(X) &= \mathrm{tr}(CX(C - \sum_i u_i A_i) X) \\ &= \mathrm{tr}(X^{1/2}(C - \sum_i u_i A_i) X C X^{1/2}) \\ &= \mathrm{tr}(X^{1/2}(C - \sum_i u_i A_i) X (C - \sum_j u_j A_j) X^{1/2}) \\ &= \|X^{1/2}(C - \sum_i u_i A_i) X^{1/2}\|_F^2. \quad \square \end{aligned} \tag{11}$$

For the affine scaling algorithm, there are two major strategies to define the stepsize: short step strategy and long step strategy.

In the short step strategy, the stepsize parameter is determined with respect to the Dikin's ellipsoid. Specifically, the iteration of the short step affine scaling algorithm can be written as

$$X_{k+1} = X_k - \lambda \frac{D(X_k)}{\|X_k^{1/2} S_k X_k^{1/2}\|_F} = X_k - \lambda \frac{D(X_k)}{\sqrt{C \bullet D(X_k)}} \tag{12}$$

where $\lambda \leq 1$ is a stepsize parameter. It is easily seen that the condition $\lambda \leq 1$ ensures that the next point X_{k+1} is feasible.

The long step strategy, which is widely used in practice, means that the next iterate is chosen by moving a fixed ratio of the whole step to the boundary of the feasible region in the direction of affine scaling direction, namely,

$$X_{k+1} = X_k - \mu \rho(X_k) D(X_k), \tag{13}$$

where

$$\rho(X) = \sup\{\, \rho > 0 \mid X - \rho D(X) \succ 0 \,\}, \tag{14}$$

and $\mu < 1$ is the ratio.

In the following section, we give an instance of (1) for which both the long step and the short step affine scaling algorithm fail to converge to an optimal solution.

3 An Example

Consider the following SDP problem,

$$\min \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \bullet X \text{ subject to } \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \bullet X = 2, \quad X \succeq 0, \tag{15}$$

and its dual

$$\max 2u \text{ subject to } S + u\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad S \succeq 0, \tag{16}$$

where $X, S \in \mathcal{S}(2)$ and $u \in R$. We note that both the primal and dual have interior feasible solutions. By defining

$$X = \begin{pmatrix} x & 1 \\ 1 & y \end{pmatrix}, \tag{17}$$

it is easy to see that the problem (15) is equivalent to

$$\min x + y \text{ subject to } x \geq 0, \quad y \geq 0, \quad xy \geq 1, \tag{18}$$

whose optimal solution is $(x, y) = (1, 1)$. Hence, it can also be easily verified that

$$X^* = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \quad \text{and } (u^*, S^*) = \left(1, \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}\right) \tag{19}$$

are the unique optimal solutions of the primal (15) and the dual (16), respectively, that the optimal values coincide, and that X^* and (u^*, S^*) satisfy strict complementarity, namely, we can decompose X^* and (u^*, S^*) as

$$X^* = Q\begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}Q^t, \text{ and } S^* = Q\begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix}Q^t, \tag{20}$$

where

$$Q = \begin{pmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} \tag{21}$$

is the orthogonal matrix whose columns are the eigenvectors of X^* and S^* corresponding to each nonzero eigenvalues.

We can calculate the dual estimates by (4) as follows.

$$\begin{aligned} G(X) &= 2(xy+1), && (22) \\ p(X) &= 2(x+y), && (23) \\ u(X) &= G(X)^{-1}p(X) = \frac{x+y}{xy+1}, && (24) \\ S(X) &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - u(X)\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & -u(X) \\ -u(X) & 1 \end{pmatrix}. && (25) \end{aligned}$$

By a straightforward calculation, we have the affine scaling direction

$$D(X) = \frac{xy-1}{xy+1}\begin{pmatrix} x^2-1 & 0 \\ 0 & y^2-1 \end{pmatrix}. \tag{26}$$

Hence an iteration of the general affine scaling algorithm can be written as

$$\begin{pmatrix} x^{k+1} & 1 \\ 1 & y^{k+1} \end{pmatrix} = \begin{pmatrix} x^k & 1 \\ 1 & y^k \end{pmatrix} - \alpha^k\begin{pmatrix} (x^k)^2-1 & 0 \\ 0 & (y^k)^2-1 \end{pmatrix} \tag{27}$$

for some scaler $\alpha^k > 0$. Thus we have

$$\begin{aligned} x^{k+1} &= x^k - \alpha^k((x^k)^2 - 1), && (28) \\ y^{k+1} &= y^k - \alpha^k((y^k)^2 - 1), && (29) \end{aligned}$$

which is the expression of the affine scaling algorithm in the (x, y)-space. In other words, the sequence $\{X_k\}$ generated by the affine scaling algorithm for the SDP (15) corresponds to the sequence $\{(x^k, y^k)\}$ generated by (28) and (29).

Let us define $\mathcal{F} \triangleq \{(x, y) \mid xy > 1, x < 1, y > 1\}$. We have the following proposition.

Proposition 3 *Assume that we do one iteration* (28) *and* (29) *of the affine scaling algorithm from* $(x,y) \in \mathcal{F}$ *to produce the next iterate* (x^+,y^+). *Then we have* $x^+ \geq x$, $y^+ \leq y$, *and* $(x^+,y^+) \in \mathcal{F}$.

Proof : The former relations are obvious from (28), (29) and the definition of $\mathcal{F}$. Letting

$$x(\alpha) = x - \alpha(x^2 - 1), \tag{30}$$
$$y(\alpha) = y - \alpha(y^2 - 1), \tag{31}$$

below we prove the last one by showing that if $y(\alpha) = 1$ then $x(\alpha) < 1$, which implies that $x(\alpha) < 1$ at the boundary.

From $y(\alpha) = 1$, $\alpha = 1/(y+1)$ follows. Substituting this relation into (30), we have

$$x(\alpha) - 1 = x - 1 - \frac{(x^2-1)}{y+1}, \tag{32}$$

and

$$\frac{1-x(\alpha)}{1-x} = 1 - \frac{(x+1)}{y+1} = \frac{y-x}{y+1} > 0. \tag{33}$$

Therefore, the proposition follows. □

Corollary 4 *If the initial point* (x^0,y^0) *is contained in* $\mathcal{F}$, *the sequence generated by affine scaling algorithm converges regardless of stepsize.*

Proof : $\{x^k\}$ is monotonically increasing and bounded above by 1, thus has a limit point. Similarly, $\{y^k\}$ is monotonically decreasing and bounded below by 1. Thus $\{y^k\}$ also has a limit point, and $\{(x^k,y^k)\}$ converges. □

Proposition 5 *Let* $(x,y) \in \mathcal{F}$ *and* (x^+,y^+) *be the next iterate of the affine scaling algorithm with* $\alpha^k = \alpha$. *Then we have*

$$\frac{x^+y^+ - 1}{xy - 1} = 1 - \alpha(x+y) - \alpha^2\frac{(1-x^2)(y^2-1)}{xy-1} < 1. \tag{34}$$

This relation follows from (28) and (29) by a straightforward calculation, thus we omit the proof.

Now we consider the long step affine scaling algorithm. Note that if the left hand side of (34) is 0, then (x^+,y^+) is on the boundary of the feasible region. Let $\sigma(x,y)$ be the positive solution of

$$1 - \sigma(x+y) - \sigma^2\frac{(1-x^2)(y^2-1)}{xy-1} = 0. \tag{35}$$

Noting that

$$\sigma(x,y)\begin{pmatrix} x^2-1 \\ y^2-1 \end{pmatrix} \tag{36}$$

is the whole step of the way to the boundary of the feasible region in the (x,y)-space, we see that the iteration of the long step affine scaling algorithm can be written as follows:

$$x^{k+1} = x^k - \mu\sigma(x^k,y^k)((x^k)^2 - 1), \tag{37}$$
$$y^{k+1} = y^k - \mu\sigma(x^k,y^k)((y^k)^2 - 1), \tag{38}$$

where $\mu < 1$ is a stepsize parameter. We denote $\sigma(x^k,y^k)$ by σ^k in what follows.

Theorem 6 *Assume that we use the long step affine scaling algorithm with stepsize* $\mu < 1$ *for* (15). *Fix any positive number* ϵ, *and choose the initial point* $(x^0,y^0) \in \mathcal{F}$ *to satisfy*

$$\sqrt{x^0y^0 - 1} < \frac{\mu}{2}\sqrt{\frac{1-(x^0)^2}{(y^0)^2-1}}(y^0 - 1 - \epsilon). \tag{39}$$

Then the sequence $\{(x^k,y^k)\}$ *converges to a non-optimal point.*

Note that the inequality (39) is always possible since for fixed ϵ, μ, and y^0, the left hand side goes to 0 as $x \downarrow 1/y^0$, while the right hand side remains positive.

Proof: Assume by contradiction that $(x^k, y^k) \to (1,1)$, which is the optimal solution. Then $y^k < 1+\epsilon$ for sufficiently large k. Let L be the iteration number such that $y^L \geq 1+\epsilon$ and $y^{L+1} < 1+\epsilon$. Since y^k is monotonically decreasing, such L is unique.

In view of (34) with $\alpha = \mu\sigma^k$, and (35), we have

$$\begin{aligned} \frac{x^{k+1}y^{k+1}-1}{x^k y^k - 1} &= 1 - \mu\sigma^k(x^k+y^k) - \mu^2(\sigma^k)^2\frac{(1-(x^k)^2)((y^k)^2-1)}{x^k y^k - 1} \\ &= 1 - \mu\sigma^k(x^k+y^k) - \mu^2(1-(x^k+y^k)\sigma^k) \\ &= 1 - \mu^2 - (x^k+y^k)\mu\sigma^k(1-\mu) \\ &< 1-\mu^2. \end{aligned} \tag{40}$$

Furthermore from (35), it follows that

$$(\sigma^k)^2 = \frac{x^k y^k - 1}{(1-(x^k)^2)((y^k)^2-1)}(1-(x^k+y^k)\sigma^k) \leq \frac{x^k y^k - 1}{(1-(x^k)^2)((y^k)^2-1)}, \tag{41}$$

and

$$\sigma^k \leq \sqrt{\frac{x^k y^k - 1}{(1-(x^k)^2)((y^k)^2-1)}}. \tag{42}$$

Now we have

$$\begin{aligned} y^0 - y^{L+1} &= \sum_{k=0}^{L}(y^k - y^{k+1}) \\ &= \mu\sum_{k=0}^{L}\sigma^k((y^k)^2-1) \\ &\leq \mu\sum_{k=0}^{L}\sqrt{\frac{(x^k y^k-1)((y^k)^2-1)}{1-(x^k)^2}} \\ &\leq \mu\sqrt{\frac{(y^0)^2-1}{1-(x^0)^2}}\sum_{k=0}^{L}\sqrt{x^k y^k - 1} \\ &\leq \mu\sqrt{\frac{((y^0)^2-1)(x^0y^0-1)}{1-(x^0)^2}}\sum_{k=0}^{L}\sqrt{(1-\mu^2)}^k \\ &\leq \mu\sqrt{\frac{((y^0)^2-1)(x^0y^0-1)}{1-(x^0)^2}}\sum_{k=0}^{L}(1-\frac{\mu^2}{2})^k \\ &\leq \frac{2}{\mu}\sqrt{\frac{((y^0)^2-1)(x^0y^0-1)}{1-(x^0)^2}}. \end{aligned} \tag{43}$$

Due to the choice (39) of the initial point, we have

$$y^0 - y^{L+1} \leq y^0 - 1 - \epsilon, \tag{44}$$

which contradicts the relation that $y^0 - y^{L+1} > y^0 - 1 - \epsilon$. Therefore, the limit point is not optimal. □

Next we turn to the short step affine scaling algorithm. From (12) and (26), the iteration of the short step affine scaling algorithm can be written as follows:

$$x^{k+1} = x^k - \lambda\theta(x^k,y^k)((x^k)^2-1), \tag{45}$$
$$y^{k+1} = y^k - \lambda\theta(x^k,y^k)((y^k)^2-1), \tag{46}$$

where

$$\theta(x,y) = \sqrt{\frac{xy-1}{(xy+1)(x^2+y^2-2)}}. \tag{47}$$

We denote $\theta(x^k, y^k)$ by θ^k in what follows.

Theorem 7 *Assume that we use the short step affine scaling algorithm with stepsize $\lambda \le 1$ for (15). Let for $\eta > 4$,*

$$\mathcal{G}(\eta) \triangleq \{ (x,y) \in \mathcal{F} \,|\, xy \le 2, \quad \eta \ge y \ge 4 \}. \tag{48}$$

Choose the initial point $(x^0, y^0) \in \mathcal{G}(\eta)$ to satisfy

$$\sqrt{x^0y^0-1} < \frac{15\sqrt{7}\lambda}{4\eta^4}(y^0-4). \tag{49}$$

Then the sequence $\{(x^k, y^k)\}$ converges to a non-optimal point.

Note that the inequality (49) is always possible since for fixed η, λ, and y^0, we can arbitrarily reduce the left hand side by choosing appropriate x^0.

Proof: Assume by contradiction that $(x^k, y^k) \to (1,1)$, which is the optimal solution. Then by Propositions 3 and 5, the sequence must be in $\mathcal{H} \triangleq \{ (x,y) \in \mathcal{F} \,|\, xy \le 2, \quad y < 4 \}$ in finite number of iterations. Let L be the iteration number such that $(x^L, y^L) \in \mathcal{G}(\eta)$ and $(x^{L+1}, y^{L+1}) \in \mathcal{H}$. Since y^k is monotonically decreasing, and $\mathcal{G}(\eta)$ and $\mathcal{H}$ are disjoint, such L is unique.

From Proposition 5, substituting $\alpha = \lambda\theta^k$, we have

$$\begin{aligned} \frac{x^{k+1}y^{k+1}-1}{x^ky^k-1} &= 1 - \lambda\theta^k(x^k+y^k) - \lambda^2(\theta^k)^2\frac{(1-(x^k)^2)((y^k)^2-1)}{x^ky^k-1} \\ &\le 1 - \lambda^2\frac{(1-(x^k)^2)((y^k)^2-1)}{(x^ky^k+1)((x^k)^2+(y^k)^2-2)}. \end{aligned} \tag{50}$$

When $(x^k, y^k) \in \mathcal{G}(\eta)$, we see that

$$\begin{aligned} &(y^k)^2 - 1 \ge 15, \qquad 1-(x^k)^2 \ge 3/4, \\ &x^ky^k+1 \le 3, \qquad (x^k)^2+(y^k)^2-2 \le \eta^2-1 \le \eta^2. \end{aligned} \tag{51}$$

Substituting these relations into (50), we have for $k \le L$,

$$\frac{x^{k+1}y^{k+1}-1}{x^ky^k-1} \le 1 - \lambda^2\frac{45/4}{3\eta^2} = 1 - \frac{15\lambda^2}{4\eta^2}. \tag{52}$$

Now we have

$$\begin{aligned} y^0 - y^{L+1} &= \sum_{k=0}^{L}(y^k - y^{k+1}) \\ &= \lambda\sum_{k=0}^{L}\sqrt{\frac{x^ky^k-1}{(x^ky^k+1)((x^k)^2+(y^k)^2-2)}}((y^k)^2-1) \\ &\le \lambda\sum_{k=0}^{L}\frac{\sqrt{x^ky^k-1}}{\sqrt{28}}\eta^2 \quad \text{(Because } (x^k,y^k) \in \mathcal{G}(\eta) \subset \mathcal{F}.\text{)} \\ &\le \frac{\lambda\eta^2\sqrt{x^0y^0-1}}{2\sqrt{7}}\sum_{k=0}^{L}\left(\sqrt{1-\frac{15\lambda^2}{4\eta^2}}\right)^k \quad \text{(Use (52))} \\ &\le \frac{\lambda\eta^2\sqrt{x^0y^0-1}}{2\sqrt{7}}\sum_{k=0}^{L}\left(1-\frac{15\lambda^2}{8\eta^2}\right)^k \end{aligned}$$

$$\leq \frac{\lambda\eta^2\sqrt{x^0y^0-1}}{2\sqrt{7}}\frac{8\eta^2}{15\lambda^2}$$
$$= \frac{4\eta^4\sqrt{x^0y^0-1}}{15\sqrt{7}\lambda}. \tag{53}$$

Due to the choice (49) of the initial point, we have

$$y^0 - y^{L+1} < y^0 - 4, \tag{54}$$

which contradicts the relation that $y^0 - y^{L+1} \geq y^0 - 4$ because $y^{L+1} \in \mathcal{H}$. Therefore, the limit point is not optimal. □

4 Concluding Remarks

In our example, the dual estimates also converge to an infeasible point. In fact, we have from (24),

$$\lim_{k\to\infty} u(x^k, y^k) = \lim_{k\to\infty} \frac{x^k + y^k}{x^ky^k + 1} = \frac{x^\infty + y^\infty}{x^\infty y^\infty + 1}, \tag{55}$$

and by noting that $x^\infty y^\infty = 1$, we have

$$\frac{x^\infty + y^\infty}{x^\infty y^\infty + 1} = \frac{x^\infty + 1/x^\infty}{2} > 1, \tag{56}$$

where the last inequality follows from $x^\infty < 1$. Hence, we have $u(x^\infty, y^\infty) > 1$, which implies that the limit of the dual estimates is infeasible. Note that the dual estimate is continuous everywhere on the feasible region. In LP, nondegeneracy assumption implies this condition, which produces a much simpler proof of the global convergence of the affine scaling algorithm. Our example however, shows that this condition is not effective in proving the global convergence of the affine scaling algorithm for SDP.

By choosing an initial point to satisfy (39) of Theorem 6 in the long step strategy or (49) of Theorem 7 in the short step strategy, we can easily do a numerical experience such that the sequences of the affine scaling algorithm converge to a non-optimal point. For example, in the short step affine scaling algorithm with stepsize $\lambda = 0.5$, the sequence generated from $(x^0, y^0) = (0.20001, 5.0)$ converges to approximately $(0.2078065619, 4.8121675791)$, a non-optimal point.

Even though our example shows that the global convergence from arbitrary starting point is impossible, it may still be possible to prove global convergence from well-chosen starting points, or by allowing variable stepsizes.

Acknowledgements

The author wishes to thank Prof. Takashi Tsuchiya, Prof. Renato Monteiro, and Prof. Shinji Mizuno of the Institute of Statistical Mathematics for having many helpful conversations with the author, and for pointing out several improvements of the first draft of this paper.

This work was partially supported by the Grant-in-Aid for Encouragement of Young Scientists of Ministry of Education, Science, Sports and Culture of Japan, and the Grant-in-Aid for Developmental Scientific Research (B) of Ministry of Education, Science, Sports and Culture of Japan.

References

[1] F. Alizadeh, Interior point methods in semidefinite programming with application to combinatorial optimization, SIAM J. Optim. 5, (1995) 13-51.

[2] F. Alizadeh, J. P. A. Haeberly, and M. L. Overton, Primal-dual interior-point methods for semidefinite programming, Technical report, (1994).

[3] F. Alizadeh, J. P. A. Haeberly, and M. L. Overton, Complementarity and nondegeneracy in semidefinite programming, Technical report, (1995).

[4] E. R. Barnes : A Variation on Karmarkar's Algorithm for Solving Linear Programming Problems. *Mathematical Programming,* Vol. 36, (1986) 174-182.

[5] I. I. Dikin, Iterative solution of problems of linear and quadratic programming, Sov. Math. Doklady 8, (1967) 674-675.

[6] L. Faybusovich, On a matrix generalization of affine-scaling vector fields, SIAM J. Mat. Anal. Appl. 16, No. 3, (1995) 886-897.

[7] L. Faybusovich, Dikin's algorithm for matrix linear programming problems, Lecture notes in control and information sciences, J. Henry and J. Pavon editors, Springer-Verlag, Vol. 197, (1994) 237-247.

[8] D. Goldfarb and K. Scheinberg, Interior point trajectories in semidefinite programming, Technical report, (1996).

[9] C. Helmberg, F. Rendl, R. J. Vanderbei, and H. Wolkowicz, An interior-point method for semidefinite programming, SIAM J. Optim. 6, (1996) 342-361.

[10] F. Jarre, An interior-point method for minimizing the maximum eigenvalue of a linear combination of matrices. SIAM J. Cont. and Optim. 31, (1993) 1360-1377.

[11] N. Karmarkar, A new polynomial-time algorithm for linear programming, Combinatorica 4, No. 4, (1984) 373-395.

[12] M. Kojima, S. Shindoh and S. Hara, Interior point methods for the monotone semidefinite linear complementarity problems, Research Report #282, Dept. of Mathematical and Computing Sciences, Tokyo Institute of Technology, Oh-Okayama, Meguro, Tokyo 152, Japan, (1994, Revised 1995).

[13] M. Kojima, M. Shida and S. Shindoh, Local convergence of predictor-corrector infeasible-interior-point algorithms for SDPs and SDLCPs, Research Report #B-306, Dept. of Mathematical and Computing Sciences, Tokyo Institute of Technology 2-12-1 Oh-Okayama, Meguro-ku, Tokyo 152, Japan, (1995).

[14] C-J. Lin and R. Saigal, A predictor-corrector method for semi-definite programming, Working paper, Dept. of Industrial and Operations Engineering, The University of Michigan, Ann Arbor, Michigan 48109-2177, (1995).

[15] Z-Q. Lou, J. F. Sturm and S. Zhang, Superlinear convergence of a symmetric primal-dual path-following algorithms for semidefinite programming, Working paper, (1996).

[16] R. D. C. Monteiro, Primal-dual path-following algorithms for semidefinite programming, Working paper, (1995). To appear in *SIAM J. Optim.*

[17] R. D. C. Monteiro, T. Tsuchiya and Y. Wang, A simplified global convergence proof of the affine scaling algorithm, Ann. Oper. Res. 47, (1993) 443-482.

[18] R. D. C. Monteiro, and Y. Zhang, A unified analysis for a class of path-following primal-dual interior-point algorithms for semidefinite programming, Working Paper, (1996).

[19] Y. E. Nesterov and A. S. Nemirovskii. Optimization over positive semidefinite matrices: Mathematical background and user's manual. Technical report, Central Economic & Mathematical Institute, USSR Acad. Sci. Moscow, USSR, (1990).

[20] Yu. E. Nesterov and A. Nemirovskii, Interior-point Polynomial Algorithms in Convex Programming, SIAM, Philadelphia, PA, USA.

[21] Y. E. Nesterov and M. J. Todd, Self-scaled barriers and interior-point methods for convex programming, Technical Report 1091, School of Operations Research and Industrial Engineering, Cornell University, Ithaca, New York, 14853-3801, (1995).

[22] Y. E. Nesterov and M. J. Todd, Primal-dual interior-point methods for self-scaled cones, Technical Report 1125, School of Operations Research and Industrial Engineering, Cornell University, Ithaca, New York, 14853-3801, (1995).

[23] F. A. Potra and R. Sheng, A superlinearly convergent primal-dual infeasible-interior-point algorithm for semidefinite programming. Reports on Computational Mathematics, No.78, Department of Mathematics, The University of Iowa, (1995).

[24] J. F. Sturm and S. Zhang, Symmetric primal-dual path-following algorithms for semidefinite programming, Report 9554/A, Ecometric Institute, Erasmus University Rotterdam, The Netherlands, (1995).

[25] R. Saigal, A simple proof of primal affine scaling method, Technical Report, Department of Industrial and Operations Engineering, University of Michigan, Ann Arbor, MI, USA, (1993). (To appear in Ann. Oper. Res.)

[26] R. Saigal, Linear programming: A modern integrated analysis, Kluwer Academic Publishers, USA, (1996).

[27] T. Tsuchiya and M. Muramatsu, Global convergence of a long-step affine scaling algorithm for degenerate linear programming problems, SIAM J. Optim. 5, No. 3, (1995).

[28] L. Vandenberghe and S. Boyd, A primal-dual potential reduction method for problems involving matrix inequalities, *Mathematical Programming* Vol.69, (1995) 205-236.

[29] R. J. Vanderbei, M. S. Meketon and B. A. Freedman, A Modification of Karmarkar's Linear Programming Algorithm. *Algorithmica*, Vol. 1, (1986) 395-407.

[30] Y. Zhang, On extending primal-dual interior-Point algorithms from linear programming to semidefinite programming, Technical Report TR95-20, Dept. of Math/Stat, University of Maryland Baltimore County, Baltimore, MD 21228, USA, (1995).