On the long step path–following method for semidefinite programming

Jos F. Sturm *
Shuzhong Zhang †

January 8, 1998

Abstract

It has been shown in various recent research reports that the analysis of short step primal-dual path following algorithms for linear programming can be nicely generalized to semidefinite programming. However, the analysis of long step path-following algorithms for semidefinite programming appeared to be less straightforward. For such an algorithm, Monteiro [9] obtained an $O(n^{1.5}\log(1/\epsilon))$ iteration bound for obtaining an ϵ-optimal solution, where n is the order of the semidefinite decision variable. In this paper, we propose to use a different search direction, viz. the so-called V-space direction. It is shown that this modification reduces the iteration complexity to $O(n\log(1/\epsilon))$. Independently, Monteiro and Y. Zhang obtained a similar result using Nesterov-Todd directions.

Key words. semidefinite programming, long step path following, symmetric primal–dual transformation.

1. Introduction

A number of recent reports are dedicated to the generalization of primal-dual path-following algorithms from linear towards semidefinite programming. This research has yielded various ways of generating search directions. Alizadeh, Haeberly and Overton [1] proposed a method based on the so-called $XZ + ZX$ (or AHO) direction and Q–methods. Local quadratic convergence for the $XZ + ZX$–method was shown by Kojima, Shida and Shindoh [6], and Monteiro [10] recently obtained some polynomiality results for this method. Helmberg et al. [2], Nesterov and Todd [12, 13] and Monteiro [9] proposed other methods based on different search directions. These methods fit in the general frameworks of Kojima, Shindoh and Hara [5] and Y. Zhang [17]. For algorithms in which the central path is followed closely, an $O(\sqrt{n}\log(1/\epsilon))$ worst case iteration bound has been obtained for several of these search directions. The generalization of the *long step path-following* method of Kojima, Mizuno and Yoshise [3] appeared to be more difficult. A generalization was proposed by Monteiro [9], but he obtained an $O(n^{1.5}\log(1/\epsilon))$ iteration bound, therefore requiring an extra factor $\sqrt{n}$ compared to the well known result for linear programming [3]. Independently of our work, Monteiro and Y. Zhang [11] show that the iteration bound can be reduced to $O(n\log(1/\epsilon))$ if Nesterov-Todd directions [12, 13, 14, 16] are used.

We show in this short paper that we can also get rid of the afore mentioned extra factor $\sqrt{n}$ if we use the so-called *V-space direction* instead of Monteiro's directions. The V-space direction was recently proposed in our report [15]. We showed in [15] that this direction is well suited for approximating weighted centers that are located in a wide neighborhood of the central path. It is therefore a natural idea to use the V–space direction in a long step path–following method.

* Tinbergen Institute Rotterdam, The Netherlands, sturm@few.eur.nl.
† Erasmus University Rotterdam, The Netherlands, zhang@few.eur.nl.

There is a close relation between Nesterov–Todd directions and V–space directions. Namely, the Nesterov–Todd direction [12, 13] can be derived from the linearization of Λ_{XZ}, where X and Z are respectively the primal and the dual semidefinite matrix variables, and Λ_{XZ} is the diagonal matrix of the eigenvalues of XZ, see [14]. Similarly, the V–space direction is based on the linearization of $\Lambda_{XZ}^{1/2}$, see [15]. The matrix $\Lambda_{XZ}^{1/2}$ has an attractive interpretation. Namely, after applying a certain linear transformation to the semidefinite programming problem (the so–called symmetric primal–dual transformation [14]), both the primal variable X and the dual variable Z map onto the diagonal matrix $\Lambda_{XZ}^{1/2}$. The V–space is an extension of the v–space approach that was developed by Kojima et al. [4] for complementarity problems.

Notation.
Given X and Y in $\Re^{n\times n}$, the standard inner product is defined by $X \bullet Y = \mathrm{tr} X^{\mathrm{T}} Y$. The Euclidean norm and its associated operator norm, viz. the spectral norm, are both denoted by $\|\cdot\|$. The Frobenius norm of a matrix $X \in \Re^{n\times n}$ is $\|X\|_F = \sqrt{X \bullet X}$. The space of symmetric $n \times n$ matrices is denoted by $\mathcal{S}$. If $X \in \mathcal{S}$ is positive (semi-) definite, we write $X \succ 0$ ($X \succeq 0$). The cone of positive semi-definite matrices is denoted by $\mathcal{S}_+$ and the cone of positive definite matrices is $\mathcal{S}_{++}$. For $X \in \mathcal{S}_+$ we let $\lambda_{\min}(X)$ denote the smallest eigenvalue of X. The order n identity matrix is denoted by I.

1.1. Problem statement

In this paper, we study the standard primal semidefinite programming problem (SDP)

$$(P) \quad \inf\{C \bullet X \mid X \succeq 0,\ A_i \bullet X = b_i \text{ for } i = 1, \ldots, m\}$$

along with its dual

$$(D) \quad \sup\{b^{\mathrm{T}} y \mid Z \succeq 0,\ Z = C - \textstyle\sum_{i=1}^m y_i A_i\}.$$

The primal feasible set is denoted by

$$\mathcal{F}_P := \{X \in \mathcal{S}_+ \mid A_i \bullet X = b_i \text{ for } i = 1, \ldots, m\}$$

and the dual feasible set is

$$\mathcal{F}_D := \{Z \in \mathcal{S}_+ \mid Z = C - \sum_{i=1}^m y_i A_i,\ \exists y \in \Re^m\}.$$

We make the standard assumption that both $\mathcal{F}_P$ and $\mathcal{F}_D$ contain positive definite matrices. This assumption guarantees the existence of an optimal solution pair (X, Z) with $X \bullet Z = 0$, and we can replace 'inf' by 'min' and 'sup' by 'max' in problems (P) and (D) respectively.

1.2. Long step path-following

A primal-dual pair of feasible solutions $(X, Z) \in \mathcal{F}_P \times \mathcal{F}_D$ is said to be a *μ-center*, if

$$XZ = \mu I$$

for some $\mu \in \Re_{++}$. The long step primal-dual path-following method for semidefinite programming [9, 17] generates an iterative sequence $(X^{(k)}, Z^{(k)})$ in the *wide neighborhood* of the central path

$$\mathcal{N}(\theta) = \{(X, Z) \in \mathcal{F}_P \times \mathcal{F}_D \mid \lambda_{\min}(XZ) \geq \theta^2 \mu,\ \mu = \frac{X \bullet Z}{n}\}$$

with parameter $\theta \in (0,1)$. Given an iterate $(X^{(k)}, Z^{(k)}) \in \mathcal{N}(\theta)$, a new iterate

$$X^{(k+1)} = X^{(k)} + 2t_k \Delta X^{(k)}, \quad Z^{(k+1)} = Z^{(k)} + 2t_k \Delta Z^{(k)}$$

is determined for some step length t_k and search direction $(\Delta X^{(k)}, \Delta Z^{(k)})$. The step length t_k is chosen such that $(X^{(k+1)}, Z^{(k+1)}) \in \mathcal{N}(\theta)$. We let $t_k(\theta)$ denote the step length to the boundary of $\mathcal{N}(\theta)$, i.e.

$$t_k(\theta) = \max\{\bar{t} \mid (X^{(k)} + 2t\Delta X^{(k)}, Z^{(k)} + 2t\Delta Z^{(k)}) \in \mathcal{N}(\theta),\ \forall t \in [0, \bar{t}]\}. \tag{1}$$

The long step path-following algorithm is now described as follows:

Algorithm 1 (Long step path-following)**.**
Parameters $\theta \in (0,1)$, $\gamma \in (0,1)$ and $\epsilon > 0$. Initial solution $(X^{(0)}, Z^{(0)}) \in \mathcal{N}(\theta)$.

Step 0 Set $k = 0$.

Step 1 Stop if $X^{(k)} \bullet Z^{(k)} \leq \epsilon$.

Step 2 Compute a Newton direction $(\Delta X^{(k)}, \Delta Z^{(k)})$ towards the $(\gamma^2 \mu^{(k)})$-center, where

$$\mu^{(k)} := \frac{X^{(k)} \bullet Z^{(k)}}{n}.$$

Step 3 Choose a step length $t_k \geq t_k(\theta)/2$ such that

$$(X^{(k)} + 2t_k \Delta X^{(k)}, Z^{(k)} + 2t_k \Delta Z^{(k)}) \in \mathcal{N}(\theta).$$

Step 4 Let

$$X^{(k+1)} = X^{(k)} + 2t_k \Delta X^{(k)}, \quad Z^{(k+1)} = Z^{(k)} + 2t_k \Delta Z^{(k)},$$

replace k by $k+1$ and return to Step 1.

In Step 2 of the above algorithm, different ways of computing a Newton direction are possible. In fact, there is an ongoing investigation on the different ways of generating Newton directions for semidefinite programming. In this paper, we will use the V-space Newton directions as proposed in [15].

1.3. The V-space

In this section, we briefly review the V-space approach for semidefinite programming. Consider $X, Z \in \mathcal{S}_{++}$ with $X \in \mathcal{F}_P$ and $Z \in \mathcal{F}_D$. Let Λ_{XZ} denote a positive diagonal matrix whose diagonal entries are the eigenvalues of the matrix XZ. Let Q denote an arbitrary unitary matrix of order n (i.e. $QQ^{\mathrm{T}} = I$), and define

$$V := Q\Lambda_{XZ}^{1/2} Q^{\mathrm{T}}.$$

It is shown in [14, 15, 16] that there exists an invertible matrix L_d such that

$$L_d^{-1} X L_d^{-\mathrm{T}} = L_d^{\mathrm{T}} Z L_d = V.$$

In particular, a procedure for computing L_d is given in Section 2.4 of [15]. We see that V is both feasible for the transformed primal problem

$$\min\{L_d^{\mathrm{T}} C L_d \bullet \bar{X} \mid L_d \bar{X} L_d^{\mathrm{T}} \in \mathcal{F}_P\}$$

and for its dual

$$\max\{b^{\mathrm{T}} y \mid \bar{Z} \succeq 0,\ \bar{Z} = L_d^{\mathrm{T}}(C - \sum_{i=1}^m y_i A_i) L_d\}.$$

This explains why the above transformation is known as the *symmetric primal-dual transformation*. Note that the duality gap of the pair (X, Z) in the original SDP is the same as the duality gap of the pair (V, V) for the transformed SDP. It follows that the optimal solution set will be approached if $V \to 0$.

1.4. Primal-dual search directions

It is easily seen that given a symmetric matrix $D_v \in \mathcal{S}$ and an invertible matrix $L_d \in \Re^{n\times n}$, the relations

$$\left\{ \begin{array}{rcll} D_x + D_z &=& D_v, & \\ L_d^{\mathrm{T}} A_i L_d \bullet D_x &=& 0, & \text{for } i = 1, 2, \ldots, m, \\ D_z + \sum_{i=1}^m \Delta y_i L_d^{\mathrm{T}} A_i L_d &=& 0, & \\ \Delta X &=& L_d D_x L_d^{\mathrm{T}}, & \\ \Delta Z &=& L_d^{-\mathrm{T}} D_z L_d^{-1} & \end{array} \right. \tag{2}$$

uniquely determine the primal search direction ΔX and the dual search direction $(\Delta y, \Delta Z)$. Namely, the transformed search directions D_x amd D_z form an orthogonal decomposition of D_v. In the sequel, "the search direction" will refer to the matrix D_v, which combines the primal and the dual search directions.

Once the search direction D_v is given, we obtain a solution pair

$$X(t) := X + 2t\Delta X, \quad Z(t) := Z + 2t\Delta Z$$

by moving simultaneously along ΔX in the primal space and along ΔZ in the dual-space, with a step length $t \geq 0$.

It is shown in [15] that the choice

$$D_v = \gamma\sqrt{\mu} I - V \tag{3}$$

is a Newton direction for obtaining the $(\gamma^2\mu)$-center, which is by definition the solution pair $(\hat{X}, \hat{Z})$ for which all the eigenvalues of the matrix $\hat{X}\hat{Z}$ are equal to the quantity $\gamma^2\mu$.

The Nesterov–Todd direction can be interpreted using the V–space notation, see [14]. Indeed, the Nesterov–Todd direction is the solution of system (2) with $D_v = \gamma\mu V^{-1} - V$.

2. Worst-case analysis

For simplicity, we will restrict ourselves to parameter choices with $\gamma \leq 2\theta$ (a similar restriction on the parameters was made in Mizuno, Todd and Ye [8]). Consider a primal-dual pair $(X, Z) \in \mathcal{N}(\theta)$. We let

$$\phi := \frac{\mathrm{tr}(\Lambda_{XZ}^{1/2})}{n}.$$

From the definition of $\mathcal{N}(\theta)$, we obtain

$$\phi \geq \theta\sqrt{\mu}. \tag{4}$$

On the other hand, application of the Cauchy-Schwartz inequality yields

$$\phi \leq \frac{\left\|\Lambda_{XZ}^{1/2}\right\|_F}{\sqrt{n}} = \sqrt{\mu}. \tag{5}$$

Since $D_v = \gamma\sqrt{\mu}I - V$, we have

$$\|D_v\|_F^2 = \gamma^2 n\mu - 2\gamma n\sqrt{\mu}\phi + n\mu \leq n\mu, \tag{6}$$

where we used (4) and the fact that $\gamma \leq 2\theta$.
We let

$$\mu(t) := \frac{X(t) \bullet Z(t)}{n},$$

so that $n\mu(t)$ denotes the duality gap after taking a step of length t. An upper bound on $\mu(t)$ is derived below.

Lemma 2.1. *There holds*

$$\mu(t) \leq (1 - 2(1-\gamma)t)\mu.$$

Proof:
First remark that

$$\begin{aligned} n\mu(t) = \mathrm{tr}X(t)Z(t) &= \mathrm{tr}L_d^{-1}X(t)Z(t)L_d \\ &= \mathrm{tr}(V + 2tD_x)(V + 2tD_z) \\ &= V \bullet (V + 2tD_v) + 4t^2 D_x \bullet D_z. \end{aligned}$$

Now using the orthogonality $D_x \perp D_z$, we have

$$n\mu(t) = V \bullet (V + 2tD_v) = (1-2t)n\mu + 2nt\gamma\phi\sqrt{\mu}.$$

The lemma follows by applying (5) to the above relation.

□

The following lemma bounds the eigenvalues of $X(t)Z(t)$ from below. This result is important for estimating $t(\theta)$, the step length towards the boundary of $\mathcal{N}(\theta)$, see definition (1).

Lemma 2.2. *It holds that*

$$\lambda_{\min}(X(t)Z(t)) \geq \theta^2\mu(t) + 2\theta(1-\theta)\gamma\mu t - n\mu t^2.$$

Proof:
By definition, we have

$$L_d^{-1}X(t)L_d^{-\mathrm{T}} = V + 2tD_x = V + tD_v + t(D_x - D_z),$$

and

$$L_d^{\mathrm{T}}Z(t)L_d = V + 2tD_z = V + tD_v - t(D_x - D_z),$$

so that

$$\begin{aligned} L_d^{-1}X(t)Z(t)L_d &= (V+tD_v)^2 - t^2(D_x - D_z)^2 \\ &\quad + t[(D_x - D_z)(V+tD_v) - (V+tD_v)(D_x - D_z)]. \end{aligned}$$

It follows from the above identity that

$$\frac{L_d^{-1}X(t)Z(t)L_d + [L_d^{-1}X(t)Z(t)L_d]^{\mathrm{T}}}{2} = (V+tD_v)^2 - t^2(D_x - D_z)^2. \tag{7}$$

We remark that a lower bound on $\lambda_{\min}(X(t)Z(t))$ can easily be obtained from (7) by applying Lemma 3.3 of Monteiro [9] or Lemma 5.3 of Monteiro and Zhang [11]. Alternatively, we can establish such a bound by using the elementary properties of eigenvectors and symmetric matrices, as shown below.
Let $q(t)$ be an eigenvector corresponding to the smallest eigenvalue of $X(t)Z(t)$, such that

$$X(t)Z(t)q(t) = \lambda_{\min}(X(t)Z(t))q(t), \quad \left\| L_d^{-1}q(t) \right\|_2 = 1.$$

Then

$$\begin{aligned} \lambda_{\min}(X(t)Z(t)) &= q(t)^{\mathrm{T}}L_d^{-\mathrm{T}}L_d^{-1}X(t)Z(t)q(t) \\ &= \frac{1}{2}(L_d^{-1}q(t))^{\mathrm{T}}[L_d^{-1}X(t)Z(t)L_d]L_d^{-1}q(t) \\ &\quad + \frac{1}{2}(L_d^{-1}q(t))^{\mathrm{T}}[L_d^{-1}X(t)Z(t)L_d]^{\mathrm{T}}L_d^{-1}q(t) \\ &= (L_d^{-1}q(t))^{\mathrm{T}}[(V+tD_v)^2 - t^2(D_x - D_z)^2]L_d^{-1}q(t), \end{aligned} \tag{8}$$

where we used (7) in the last identity. However,

$$V + tD_v = (1-t)V + t\gamma\sqrt{\mu}I \succeq ((1-t)\theta + t\gamma)\sqrt{\mu}I,$$

where we used the fact that $(X,Z) \in \mathcal{N}(\theta)$. Together with (8) and the fact that $\left\| L_d^{-1}q(t) \right\|_2 = 1$, it follows that

$$\lambda_{\min}(X(t)Z(t)) \geq ((1-t)\theta + t\gamma)^2\mu - t^2 \left\| D_x - D_z \right\|_2^2. \tag{9}$$

Using Lemma 2.1, we have

$$\begin{aligned} ((1-t)\theta + t\gamma)^2 &= \theta^2 + 2t\theta(\gamma - \theta) + t^2(\gamma - \theta)^2 \\ &\geq \frac{\theta^2\mu(t)}{\mu} + 2t\theta(1-\theta)\gamma + t^2(\gamma-\theta)^2 \\ &\geq \frac{\theta^2\mu(t)}{\mu} + 2t\theta(1-\theta)\gamma. \end{aligned} \tag{10}$$

Recall that $D_x \perp D_z$, so that, using also (6),

$$\|D_x - D_z\|_2^2 \leq \|D_x - D_z\|_F^2 = \|D_x + D_z\|_F^2 = \|D_v\|_F^2 \leq n\mu. \tag{11}$$

Combining (9)–(11), it finally follows that

$$\lambda_{\min}(X(t)Z(t)) \geq \theta^2\mu(t) + 2\theta(1-\theta)\gamma\mu t - n\mu t^2.$$

□

Based on Lemma 2.2, it is straightforward to derive the main result of this short paper:

Theorem 2.1. *Consider the long step path following algorithm with parameters $\gamma \in (0,1)$, $\theta \in (0,1)$, $\gamma \leq 2\theta$, and the Newton direction defined by (3). Given a tolerance $\epsilon > 0$, the algorithm terminates with an iterate satisfying $X^{(k)} \bullet Z^{(k)} \leq \epsilon$ in*

$$O(n \log \frac{X^{(0)} \bullet Z^{(0)}}{\epsilon})$$

main iterations.

Proof:
Recall from (1) and the definition of $\mathcal{N}(\theta)$ that $t_k(\theta)$ is the largest step length in iteration $k \in \{0,1,\ldots\}$ satisfying

$$\lambda_{\min}((X^{(k)} + 2t\Delta X^{(k)})(Z^{(k)} + 2t\Delta Z^{(k)})) \geq \mu_k(t)\theta^2 \text{ for all } 0 \leq t \leq t_k(\theta).$$

It thus follows from Lemma 2.2 that

$$t_k(\theta) \geq \frac{2\gamma\theta(1-\theta)}{n} = \frac{1}{O(n)},$$

for all $k = 0, 1, \ldots$. Since $t_k \geq t_k(\theta)/2$, we also have $1/t_k = O(n)$. Applying Lemma 2.1, we thus obtain

$$X^{(k+1)} \bullet Z^{(k+1)} \leq (1 - \frac{1}{O(n)})X^{(k)} \bullet Z^{(k)}$$

for all $k = 0, 1, \ldots$. The theorem follows immediately from the above inequality.

□

References

[1] Alizadeh, F., Haeberley, J.-P.A. and Overton, M.L., "Primal-dual interior point methods for semidefinite programming: convergence rates, stability and numerical results," Report 721, New York University Computer Science Department, New York, NY, 1996.

[2] Helmberg, C., Rendl, F., Vanderbei, R.J. and Wolkowicz, H., "An interior point method for semidefinite programming," *SIAM Journal on Optimization* 6 (1996) 342–361.

[3] Kojima, M., Mizuno, S. and Yoshise, A., "A primal-dual interior point algorithm for linear programming," in *Progress in Mathematical Programming: Interior point and related methods* pp. 29–37, (ed. Megiddo, N.), Springer Verlag, New York, 1989.

[4] M. Kojima, N. Megiddo, T. Noma, and A. Yoshise. *A Unified Approach to Interior Point Algorithms for Linear Complementarity Problems.* Springer-Verlag, Berlin, 1991.

[5] Kojima, M., Shindoh, S. and Hara, S., "Interior-point methods for the monotone linear complementarity problem in symmetric matrices," *SIAM Journal on Optimization* 7 (1997) 1, 86–125.

[6] Kojima, M., Shida, M. and Shindoh, S., "A predictor-corrector interior-point algorithm for the semidefinite linear complementarity problem using the Alizadeh-Haeberley-Overton search direction," Research Reports on Information Sciences, B-311, Dept. of Information Sciences, Tokyo Institute of Technology, 2-12-1 Oh-Okayama, Meguro-ku, Tokyo 152, Japan, 1996.

[7] Kojima, M., Shida, M. and Shindoh, S., "A note on the Nesterov-Todd and the Kojima-Shindoh-Hara search directions in semidefinite programming," Research Reports on Information Sciences, B-313, Dept. of Information Sciences, Tokyo Institute of Technology, 2-12-1 Oh-Okayama, Meguro-ku, Tokyo 152, Japan, 1996.

[8] Mizuno, S., Todd, M.J. and Ye, Y., "On Adaptive- Step Primal-Dual Interior-Point Algorithms for Linear Programming," *Mathematics of Operations Research* 18 (1993) 4, 964-981.

[9] Monteiro, R.D.C., "Primal-dual path following algorithms for semidefinite programming," *SIAM Journal on Optimization* 7 (1997) 663-678.

[10] Monteiro, R.D.C., "Polynomial convergence of primal–dual algorithms for semidefinite programming based on Monteiro and Zhang family of search directions," Technical Report, School of Industrial and Systems Engineering, Georgia Tech, Atlanta, Georgia, U.S.A., 1996.

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

[12] Nesterov, Y. and Todd, M.J., "Self-scaled barriers and interior-point methods for convex programming," *Mathematics of Operations Research* 22 (1997) 1, 1–42.

[13] Nesterov, Y. and Todd, M.J., "Primal-dual interior-point methods for self-scaled cones," Technical Report 1125, School of Operations Research and Industrial Engineering, Cornell University, Ithaca, New York, 1995. To appear in *SIAM Journal on Optimization*.

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

[15] Sturm, J.F. and Zhang, S., "On weighted centers for semidefinite programming," Report 9636/A, Econometric Institute, Erasmus University Rotterdam, Rotterdam, The Netherlands, 1996.

[16] Todd, M.J., Toh, K.C. and Tütüncü, R.H., "On the Nesterov-Todd direction in semidefinite programming," Technical Report 1154, School of Operations Research and Industrial Engineering, Cornell University, Ithaca, New York, 1996.

[17] Zhang, Y., "On extending primal-dual interior-point algorithms from linear programming to semidefinite programming," Technical Report, Dept. of Mathematics and Statistics, University of Maryland Baltimore County, Baltimore, Maryland, U.S.A., 1995.