An inexact interior-point method for semi-definite programming, a description and convergence proof

Janne Harju, Anders Hansson
Division of Automatic Control
E-mail: harju@isy.liu.se, hansson@isy.liu.se

14th September 2007

Report no.: LiTH-ISY-R-2819

Address:
Department of Electrical Engineering
Linköpings universitet
SE-581 83 Linköping, Sweden

WWW: http://www.control.isy.liu.se
Abstract
In this report we investigate convergence for an infeasible interior-point method for semidefinite programming.

Keywords: LMI optimization; Interior-point methods; Iterative computation; convergence.
An inexact interior-point method for semi-definite programming, a description and convergence proof

Janne Harju
2007-09-14

Inexact search directions in interior-point methods for semi-definite programming has been considered in [3] where also a proof of convergence is given. However, because the method considered was a feasible method the inexact search direction had to be projected onto the feasible space at a high computational cost. In this report we instead investigate convergence for an infeasible interior-point method for semidefinite programming.

1 Optimization problem

Let $\mathcal{X}$ be a finite-dimensional real vector space with an inner product $\langle \cdot, \cdot \rangle_{\mathcal{X}} : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ and define the linear mappings

$$A : \mathcal{X} \to S^n$$
$$A^* : S^n \to \mathcal{X}$$

where $S^n$ denotes the space of symmetric matrices of size $n \times n$ and where $A^*$ is the adjoint of $A$. The space $S^n$ has the inner product $\langle X, Y \rangle_{S^n} = \text{Tr}(XY^T)$. We will with abuse of notation use the same notation $\langle \cdot, \cdot \rangle$ for inner products defined on different spaces when the inner product used is clear from context.

Now consider the primal and dual optimization problems

$$\begin{align*}
\min & \quad \langle c, x \rangle \\
\text{s.t} & \quad A(x) + M_0 = S \\
& \quad S \succeq 0
\end{align*}$$

$$\begin{align*}
\max & \quad -\langle M_0, Z \rangle \\
\text{s.t} & \quad A^*(Z) = c \\
& \quad Z \succeq 0
\end{align*}$$

where $c, x \in \mathcal{X}$ and $S, Z \in S^n$. Additionally define $z = (x, S, Z)$ and the corresponding finite-dimensional vector space $Z = \mathcal{X} \times S^n \times S^n$ with its inner product $\langle \cdot, \cdot \rangle_Z$. We define the corresponding 2-norm $\| \cdot \|_{2} : Z \to \mathbb{R}$ by $\|z\|_{2}^{2} = \langle z, z \rangle$. We notice that the 2-norm of a matrix with this definition will be the Frobenius norm and not the induced 2-norm.
1.1 Optimality conditions

If strong duality holds then the Karush-Kuhn-Tucker conditions defines the solution to the primal and dual optimization problems, page 244 in [1]. The Karush-Kuhn-Tucker conditions for the optimization problems defined in the previous section are

\[ A(x) + M_0 = S \]  \hspace{1cm} (9)
\[ A^*(Z) = c \]  \hspace{1cm} (10)
\[ ZS = 0 \]  \hspace{1cm} (11)
\[ S \succeq 0, \ Z \succeq 0 \]  \hspace{1cm} (12)

It is assumed that the mapping \( A \) has full rank.

**Definition 1.1** The complementary slackness \( \nu \) is defined as

\[ \nu = \frac{(Z, S)}{n} \]  \hspace{1cm} (13)

**Definition 1.2** Define the central-path as the solution points for

\[ A(x) + M_0 = S \]  \hspace{1cm} (14)
\[ A^*(Z) = c \]  \hspace{1cm} (15)
\[ ZS = \nu I \]  \hspace{1cm} (16)
\[ S \succeq 0, \ Z \succeq 0 \]  \hspace{1cm} (17)

where \( \nu \geq 0 \). Note that the central-path converges to a solution of the Karush-Kuhn-Tucker conditions when \( \nu \) tends to zero.

2 Interior-point method

For a thorough description of algorithms and theory within the area of interior-point methods see [12] for linear programming while [11] gives an extensive overview of semidefinite programming. Here follows a brief discussion on what relaxed KKT conditions are to be solved when using an infeasible interior-point method. Then an inexact infeasible predictor-corrector method for semidefinite programming is described in detail.

2.1 Infeasible interior-point method

In this section an infeasible interior-point method is discussed. Such a method is initiated with an infeasible or a feasible point \( z \) and then its iterates tend toward feasibility and optimality by computing a sequence of search directions and taking steps in these directions. To derive equations for the search directions the next iterate \( z^+ = z + \Delta z \) is introduced and inserted into (14)-(17). This gives a nonlinear system of equations for \( \Delta z \). Even after linearization the variables \( \Delta S \) and \( \Delta Z \) of the solution \( \Delta z \) to these equations are not guaranteed to be symmetric since this requirement is only implicit. A solution to this remedy is to introduce the symmetry transformation \( \mathcal{H} : \mathbb{R}^{n \times n} \rightarrow \mathbb{S}^n \) that is defined by

\[ \mathcal{H}(X) = \frac{1}{2} \left( R^{-1} X R + (R^{-1} X R)^T \right) \]  \hspace{1cm} (18)
where $R \in \mathbb{R}^{n \times n}$ is the so called scaling matrix. For a thorough description of scaling matrices, see [11] and [13]. In [13] it is shown that the relaxed complementary slackness condition $ZS = \nu I$ is equivalent with

$$\mathcal{H}(ZS) = \nu I$$

(19)

for any nonsingular matrix $R$. Hence we may replace (16) with (19). Replacing $z$ with the next iterate $z^* + \Delta z$ in (14), (15), (17) and (19) results in

$$A(\Delta x) - \Delta S = -(A(x) + M_0 - S)$$

(20)

$$A^*(\Delta Z) = (c - A^*(Z))$$

(21)

$$\mathcal{H}(\Delta ZS + Z\Delta S) = \nu I - \mathcal{H}(ZS) - \mathcal{H}(\Delta Z \Delta S)$$

(22)

$$S + \Delta S \succeq 0, Z + \Delta Z \succeq 0$$

(23)

If the nonlinear term in (22) is ignored $\Delta S$ and $\Delta Z$ will be symmetric. Several approaches to handling the nonlinear term in the complementary slackness equation have been presented in the literature. A direct solution is to ignore the higher order term which gives a linear system of equations. Another approach is presented in [6].

2.2 Inexact predictor-corrector method

Here an inexact infeasible predictor-corrector method is presented. The idea of using inexact search directions has been applied to model predictive control applications in [4] and to monotone variational inequality problems in [8]. Inexact search directions have also been applied for a potential reduction method in [10] and [3].

In a predictor-corrector method alternating steps are taken. There are two separate objectives in the strategy. The predictor step decreases the duality gap while the corrector step moves the iterate towards the central-path. To this end the parameter $\nu$ in (22) is replaced with $\sigma \nu$. Then for small values of $\sigma$ a step is taken to reduce the complementary slackness $\nu$, and for values of $\sigma$ close to 1 a step to find an iterate close to the central path is taken. Then the linear system of equations to be solved for the search directions is

$$A(\Delta x) - \Delta S = -(A(x) + M_0 - S)$$

(24)

$$A^*(\Delta Z) = (c - A^*(Z))$$

(25)

$$\mathcal{H}(\Delta ZS + Z\Delta S) = \sigma \nu I - \mathcal{H}(ZS)$$

(26)

**Lemma 2.1** If the operator $A$ has full rank, i.e. $A(x) = 0$ implies that $x = 0$, and if $Z \succ 0$ and $S \succ 0$, then the linear system of equations in (24)-(26) has a unique solution.

**Proof** See Theorem 10.2.2 in [11].

For later use and to obtain an easier notation define $F : \mathcal{Z} \to S^n \times S^n \times S^n$ as

$$F(z) = \begin{bmatrix} F_p(z) \\ F_d(z) \\ F_c(z) \end{bmatrix} = \begin{bmatrix} A(x) + M_0 - S \\ A^*(Z) - c \\ \mathcal{H}(ZS) \end{bmatrix}$$

(27)
where $z = (x, S, Z)$. Also define

$$
\mathbf{r} = \begin{bmatrix}
\mathbf{r}_p \\
\mathbf{r}_d \\
\mathbf{r}_c
\end{bmatrix} = \begin{bmatrix}
-(A(x) + M_0 - S) \\
(c - A^*(Z)) \\
\sigma \nu \mathbf{I} - \mathcal{H}(ZS)
\end{bmatrix}
$$  \hspace{1cm} (28)

Now define the set $\Omega$ as

$$
\Omega = \{ z = (x, S, Z) | S \succeq 0, Z \succeq 0, \\
\| F_p(z) \|_2 \leq \beta \nu, \| F_d(z) \|_2 \leq \beta \nu, \\
\gamma \nu \mathbf{I} \succeq F_c(z) \succeq \eta \nu \mathbf{I} \}
$$  \hspace{1cm} (29)

where the scalars $\beta$, $\gamma$ and $\eta$ will be defined later on. Then define the set $\Omega_+$ as

$$
\Omega_+ = \{ z \in \Omega | S > 0, Z > 0 \}
$$  \hspace{1cm} (30)

Finally define the set $\mathcal{S}$ for which the Karush-Kuhn-Tucker conditions (9)-(12) are fulfilled.

$$
\mathcal{S} = \{ z | F_p(z) = 0, F_d(z) = 0, F_c(z) = 0, S \succeq 0, Z \succeq 0 \}
$$  \hspace{1cm} (31)

2.2.1 Algorithm

Below the overall algorithm is summarized, which is taken from [8] and adapted to semidefinite programming.

1. Initialize the counter $j = 1$ and choose $0 < \eta < \eta_{\text{max}} < 1$, $\gamma \geq n$, $\beta > 0$, $\kappa \in (0, 1)$, $0 < \sigma_{\text{min}} < \sigma_{\text{max}} < 1/2$, $\epsilon > 0$, $0 < \chi < 1$ and $z^0 \in \Omega$.
2. Evaluate stopping criteria. If fulfilled, terminate the algorithm.
3. Choose $\sigma \in (\sigma_{\text{min}}, \sigma_{\text{max}})$.
4. Compute the scaling matrix $R$.
5. Solve (24)-(26) for search direction $\Delta z^j$ with a residual tolerance $\epsilon \sigma \beta \nu / 2$.
6. Choose a step length $\alpha^j$ as the first element in the sequence $\{1, \chi, \chi^2, \ldots\}$ such that $z^{j+1} = z^j + \alpha^j \Delta z^j \in \Omega$ and such that $\nu^{j+1} \leq (1 - \alpha \kappa(1 - \sigma)) \nu^j$.
7. Update the variables, $z^{j+1} = z^j + \alpha^j \Delta z^j$ and the counter $j := j + 1$.
8. Return to step 1.

Note that any iterate generated by the algorithm is in $\Omega$, which is a closed set, since it is defined as an intersection of closed sets, see Section 3.3 for details.

3 Convergence

A global proof of convergence will be presented. This proof of convergence is due to [7]. It has been extended in [8], and inexact solutions of the equations for the search direction were considered in [4] for the application to model predictive control and in [2] for variational inequalities. The convergence result is that
either the sequence generated by the algorithm terminates at a solution to the Karush-Kuhn-Tucker conditions in a finite number of iterations, or all limit points, if any exist, are solutions to the Karush-Kuhn-Tucker conditions defined in (9)-(12). Here the proof is extended to the case of semidefinite programming.

3.1 Convergence of inexact interior-point method

In order to prove convergence some preliminary results are presented in the following lemmas.

**Lemma 3.1** Any iterate generated by the algorithm in Section 2.2.1 is in \( \bar{\Omega} \), which is a closed set. If \( z \notin \mathcal{S} \) and \( z \in \Omega \) then \( z \in \Omega_+ \).

**Proof** Any iterate generated by the algorithm is in \( \Omega \) from the definition of the algorithm. The set \( \Omega \) is closed, since it is defined as an intersection of closed sets. For a detailed proof, see Section 3.3.

The rest of the proof follows by contradiction. Assume that \( z \notin \mathcal{S} \), \( z \in \Omega \) and \( z \notin \Omega_+ \). Now study the two cases \( \nu = 0 \) and \( \nu > 0 \) separately. Note that \( \nu \geq 0 \) by definition.

First assume that \( \nu = 0 \). Since \( \nu = 0 \) and \( z \in \Omega \) it follows that \( \mathcal{F}_c(z) = 0 \), \( \| \mathcal{F}_p(z) \|_2 = 0 \) and \( \| \mathcal{F}_d(z) \|_2 = 0 \). This implies that \( \mathcal{F}_p(s) = \mathcal{F}_d(s) = 0 \). Note that \( z \in \Omega \) also implies that \( Z \succeq 0 \) and \( S \succeq 0 \). Combining these conclusions gives that \( z \in \mathcal{S} \), which is a contradiction.

Now assume that \( \nu > 0 \). Since \( \nu > 0 \) and \( z \in \Omega \) it follows that \( \mathcal{F}_c(z) = \mathcal{H}(ZS) \succ 0 \). To complete the proof two inequalities are needed. First note that \( \det(\mathcal{H}(ZS)) > 0 \). To find the second inequality the Ostrowski-Taussky inequality, see page 56 in [5],

\[
\det \left( \frac{X + X^T}{2} \right) \leq | \det(X) | \tag{32}
\]

is applied to (18). This gives

\[
\det(\mathcal{H}(ZS)) \leq | \det(R^{-1}ZSR) | = | \det(Z) | \cdot | \det(S) | = \det(Z) \cdot \det(S)
\]

where the last equality follows from \( z \in \Omega \). Combining the two inequalities gives

\[
0 < \det(\mathcal{H}(ZS)) \leq \det(Z) \cdot \det(S) \tag{33}
\]

Since the determinant is the product of all eigenvalues and \( z \in \Omega \), (33) shows that the eigenvalues are nonzero and therefore \( Z \succ 0 \) and \( S \succ 0 \). This implies that \( z \in \Omega_+ \), which is a contradiction.

The linear system of equations in (24)-(26) for the step direction is now rewritten as

\[
\frac{\partial \text{vec}(\mathcal{F}(z))}{\partial \text{vec}(z)} \text{vec}(\Delta z) = \text{vec}(r) \tag{34}
\]

Note that the vectorization is used for the theoretical proof of convergence. In practice solving the equations is preferably made with a solver based on the operator formalism.
Lemma 3.2  Assume that $A$ has full rank. Let $\hat{z} \in \Omega_+$, and let $\epsilon \in (0,1)$. Then there exist scalars $\hat{\delta} > 0$ and $\hat{\alpha} \in (0,1]$ such that if

$$\left\| \frac{\partial \text{vec}(F(z))}{\partial \text{vec}(z)} \text{vec}(\Delta z) - \text{vec}(r) \right\|_2 \leq \frac{\epsilon}{2} \sigma \beta \nu, \quad (35)$$

and if the algorithm takes a step from any point in

$$B = \{z \mid \|z - \hat{z}\|_2 \leq \hat{\delta}\} \quad (37)$$

then the calculated step length $\alpha$ will satisfy $\alpha \geq \hat{\alpha}$.

Proof  See Section 3.2.

Now the global convergence proof is presented.

Theorem 3.3  Assume that $A$ has full rank. Then for the iterates generated by the interior-point algorithm either

- $z^j \in S$ for some finite $j$.

- all limit points of $\{z^j\}$ belongs to $S$.

Remark  Note that nothing is said about the existence of a limit point. It is only stated that if a convergent subsequence exists, then its limit point is in $S$. A sufficient condition for the existence of a limit point is that $\{z^j\}$ is uniformly bounded, [7].

Proof  Suppose that the sequence $\{z^j\}$ is infinite and it has a subsequence which converges to $\hat{z} \notin S$. Denote the corresponding subsequence $\{j_i\}$ with $K$. Then for all $\delta' > 0$, there exist a $k$ such that for $j \geq k$ and $j \in K$ it holds that

$$\|z^j - \hat{z}\|_2 \leq \delta' \quad (38)$$

Since $\hat{z} \notin S$ it holds by Lemma 3.1 that $\hat{z} \in \Omega_+$ and hence from Lemma 3.2 and (13) that there exist a $\delta > 0$ and $\hat{\alpha} \in (0,1]$ such that for all $z^j$, $j \geq k$ such that $\|z^j - \hat{z}\|_2 \leq \delta$ it holds that

$$\nu^j - \nu \leq -\hat{\alpha}(1 - \sigma_{\text{max}})\nu^j < -\hat{\alpha}(1 - \sigma_{\text{max}})\delta^2 < 0 \quad (39)$$

Now take $\delta' = \delta$. Then for two consecutive points $z^j$, $z^j+1$ of the subsequence with $j \leq k$ and $j \in K$ it holds that

$$\nu^j - \nu \leq (\nu^j - \nu^{j-1}) + (\nu^{j-1} - \nu^{j-2} + \cdots) + (\nu^{j+1} - \nu^j) < \nu^{j+1} - \nu^j \leq -\hat{\alpha}(1 - \sigma_{\text{max}})\delta^2 < 0 \quad (40)$$

Since $K$ is infinite it holds that $\nu^j$ diverges to $-\infty$. The assumption $\hat{z} \in \Omega_+$ gives that $\nu > 0$, which is a contradiction.

Lemma 3.4  Assume $z \in B$ as defined in (37). Then $Z \succeq \hat{Z} - \hat{\delta}I$ and $S \succeq \hat{S} - \hat{\delta}$.
Proof We have \( \|z - \hat{z}\|_2 \leq \hat{\delta} \). Hence \( \|z - \hat{z}\|_2^2 \leq \hat{\delta}^2 \) and
\[
\|S - \hat{S}\|_2^2 + \|Z - \hat{Z}\|_2^2 + \|x - \hat{x}\|_2^2 \leq \hat{\delta}^2.
\]
Therefore \( \|S - \hat{S}\|_2^2 \leq \hat{\delta}^2 \) and \( \|Z - \hat{Z}\|_2^2 \leq \hat{\delta}^2 \). Hence by Lemma 3.5 below
\(-\delta I \leq S - \hat{S} \leq \delta I \) and \( -\delta I \leq Z - \hat{Z} \leq \delta I \).

Lemma 3.5 Assume \( A \) is symmetric. If \( \|A\|_2 \leq a \), then \(-aI \preceq A \preceq aI\).

Proof Denote the \( i^{th} \) largest eigenvalue of a matrix \( A \) with \( \lambda_i(A) \). Then the norm defined in Section 1 fulfills
\[
\|A\|_2^2 = \sum_{i=1}^n \lambda_i^2(A)
\]
see page 647 in [1]. Then
\[
\|A\|_2 \leq a \iff \sum_{i=1}^n \lambda_i^2(A) \leq a^2 \Rightarrow \lambda_i^2(A) \leq a^2 \Rightarrow |\lambda_i(A)| \leq a \iff \quad (43)
\]
and hence \(-aA \preceq A \preceq aA\).
\[ C_0 > 0. \] Introduce \( \delta \Delta z = \Delta z - \Delta z_0. \) Then from the bound on the residual in the assumptions of this lemma it follows that \( \| \partial \text{vec}(F(z))/\partial \text{vec}(z))\text{vec}(\delta \Delta z) \|_2 \) is bounded. Hence there must be a constant \( \delta C > 0 \) such that \( \| \delta \Delta z \|_2 \leq \delta C. \) Let \( C = C_0 + \delta C > 0. \) It now holds that \( \| \Delta z \|_2 \leq C \) for all \( z \in B, \sigma \in (\sigma_{\text{min}}, 1/2]. \) Notice that it also holds that \( \| \Delta Z \|_2 \leq C \) and \( \| \Delta S \|_2 \leq C. \) Define \( \hat{g}^{(1)} = \delta/2C. \) Then for all \( \alpha \in (0, \hat{\alpha}^{(1)}) \) it holds that

\[
Z(\alpha) = Z + \alpha \Delta Z \succeq \hat{Z} - \delta I + \alpha \Delta Z \succeq 2\delta I - \delta I - \frac{\delta}{2} I > 0 \tag{49}
\]

where the first inequality follows by Lemma 3.4 and \( z \in B, \) and where the second inequality follows from (47). The proof for \( S(\alpha) > 0 \) is analogous. Hence is is possible to take a positive step without violating the constraint \( S(\alpha) > 0, Z(\alpha) > 0. \)

Now we prove that

\[
F_c(z) = \mathcal{H}(Z(\alpha)S(\alpha)) \succeq \eta \nu(\alpha) I \tag{50}
\]

for some positive \( \alpha. \) This follows from two inequalities that utilizes the fact that

\[
\frac{\partial \text{vec}(F_c(z))}{\partial \text{vec}(z)} \Delta z - \text{vec}(r_c) = 0 \iff \mathcal{H}(\Delta ZS + Z\Delta S) + \mathcal{H}(ZS) - \sigma \nu I = 0 \tag{51}
\]

and the fact that the previous iterate \( z \) is in the set \( \Omega. \) Note that \( \mathcal{H}(\cdot) \) is a linear operator. Hence

\[
\mathcal{H}(Z(\alpha)S(\alpha)) = \mathcal{H}(ZS + \alpha(\Delta ZS + Z\Delta S) + \alpha^2 \Delta Z \Delta S)
\]

\[
= (1 - \alpha)\mathcal{H}(ZS) + \alpha(\mathcal{H}(ZS) + \mathcal{H}(\Delta ZS + Z\Delta S)) + \alpha^2 \mathcal{H}(\Delta Z \Delta S)
\]

\[
\succeq (1 - \alpha)\eta \nu I + \alpha \sigma \nu I + \alpha^2 \mathcal{H}(\Delta Z \Delta S)
\]

\[
\succeq (\{(1 - \alpha)\eta + \alpha \sigma \nu - \alpha^2 \lambda_{\text{min}}^2\}) I
\]

where \( \lambda_{\text{min}}^2 = \min, \lambda_i(\mathcal{H}(\Delta Z \Delta S)). \) Note that \( \lambda_{\text{min}}^2 \) is bounded. To show this we note that in each iterate in the algorithm \( Z \) and \( S \) are bounded and hence is \( R \) bounded since it its calculated from \( Z \) and \( S. \) Hence is \( \mathcal{H}(\Delta Z \Delta S) \) is bounded since \( ||\Delta Z||_2 < C \) and \( ||\Delta S||_2 < C. \) In each iterate in the algorithm \( Z \) and \( S \) are bounded and hence is \( R \) bounded since it its calculated from \( Z \) and \( S. \)

Moreover

\[
n \nu(\alpha) = \langle Z(\alpha), S(\alpha) \rangle = \langle \mathcal{H}(Z(\alpha)S(\alpha)), I \rangle
\]

\[
= \langle (1 - \alpha)\mathcal{H}(ZS) + \alpha \sigma \nu I + \alpha^2 \mathcal{H}(\Delta Z \Delta S), I \rangle
\]

\[
= (1 - \alpha) ||Z, S|| + \alpha \sigma \nu I + \alpha^2 ||\Delta Z, \Delta S||
\]

\[
\leq (1 - \alpha)n \nu + \alpha \sigma \nu I + \alpha^2 C^2
\]

The last inequality follows from \( ||\Delta Z, \Delta S|| \leq ||\Delta Z||_2 ||\Delta S||_2 \leq C^2. \) Rewriting (53) gives that

\[
\eta \nu(\alpha) \leq \left( (1 - \alpha + \alpha \sigma) \nu + \frac{\alpha^2 C^2}{n} \right) \eta \tag{54}
\]
Clearly
\[
\left(\{(1-\alpha)\eta + \alpha \sigma |\nu - \alpha^2 |\lambda^{\Delta}_{\min}\}\right) I \succeq \left(\{(1-\alpha)\eta + \alpha \sigma |\nu + \frac{\alpha^2 C^2 \eta}{n}\}\right) I
\] (55)
implies (50), which (assuming that \(\alpha > 0\)) is fulfilled if
\[
\alpha \leq \frac{\sigma \nu (1-\eta)}{C^2 \eta / n + |\lambda^{\Delta}_{\min}|}
\] (56)
Recall that \(\sigma \geq \sigma_{\min} > 0\), \(\eta < \eta_{\max} < 1\) by assumption and that \(0 < \hat{\delta}^2 \leq \nu\) by (48). Hence with
\[
\hat{\alpha}^{(2)} = \min \left(\hat{\alpha}^{(1)}, \frac{\sigma_{\min} \hat{\delta}^2 (1-\eta_{\max})}{C^2 \eta_{\max} / n + |\lambda^{\Delta}_{\min}|}\right)
\] (57)
(50) is satisfied for all \(\alpha \in (0, \hat{\alpha}^{(2)}]\).

We now show that
\[
\gamma \nu (\alpha) I \succeq H(Z(\alpha) S(\alpha))
\] (58)
First note that \(\gamma \geq n\). Then
\[
\frac{\gamma}{n} \sum_i \lambda_i (H(Z(\alpha) S(\alpha))) \geq \lambda^{\max} (H(Z(\alpha) S(\alpha))) \Leftrightarrow
\] (59)
\[
\frac{\gamma}{n} \sum_i \lambda_i (H(Z(\alpha) S(\alpha))) I \succeq \lambda^{\max} (H(Z(\alpha) S(\alpha))) I \Leftrightarrow
\] (60)
\[
\frac{\gamma}{n} \text{Tr} (H(Z(\alpha) S(\alpha))) I \succeq \lambda^{\max} (H(Z(\alpha) S(\alpha))) I
\] (61)
where the first expression is fulfilled by definition. The second equivalence follows from a property of the trace of a matrix, see page 41 in [5]. From the definition of complementary slackness in (13) it now follows that (58) holds.

Now we prove that (45) is satisfied. Let
\[
\hat{\alpha}^{(3)} = \min \left(\hat{\alpha}^{(2)}, \frac{n \hat{\delta}^2 (1-\kappa)}{2 C^2}\right)
\] (62)
Then for all \(\alpha \in [0, \hat{\alpha}^{(3)}]\) it holds that
\[
\alpha^2 C^2 \leq \alpha n (1-\kappa) \hat{\delta}^2 / 2 \leq \alpha n (1-\kappa) \nu / 2 \leq \alpha n (1-\kappa) (1-\sigma) \nu
\] (63)
where the second inequality follows from (48) and the third inequality follows from the assumption \(\sigma \leq 1/2\). This inequality together with (53) implies that
\[
\nu (\alpha) \leq \left(\{(1-\alpha) + \alpha \sigma \nu + \frac{\alpha^2 C^2}{2}\}\right) \leq \left(\{(1-\alpha) + \alpha \sigma \nu + \alpha (1-\kappa) (1-\kappa) (1-\kappa) (1-\sigma)\}\right)
\] (64)
\[
\leq \left(\{(1-\alpha) + \alpha \sigma \nu + \alpha (1-\kappa) (1-\sigma)\}\right)
\] (65)
where the second inequality is due to \(\sigma < 1/2\) and hence (45) is satisfied.

It now remains to prove that
\[
\|F_p (z(\alpha))\|_2 \leq \beta \nu (\alpha)
\] (66)
\[
\|F_d (z(\alpha))\|_2 \leq \beta \nu (\alpha)
\] (67)
Since the proofs are similar, it will only be proven that (66) holds true. Use Taylor’s theorem to write

\[ \text{vec}(\mathcal{F}_p(z(\alpha))) = \text{vec}(\mathcal{F}_p(z)) + \alpha \frac{\partial \text{vec}(\mathcal{F}_p(z))}{\partial \text{vec}(z)} \text{vec}(\Delta z) + \alpha R \]  

(68)

\[ R = \int_0^1 \left( \frac{\partial \text{vec}(\mathcal{F}_p(z(\theta \alpha)))}{\partial \text{vec}(z)} - \frac{\partial \text{vec}(\mathcal{F}_p(z))}{\partial \text{vec}(z)} \right) \text{vec}(\Delta z) \, d\theta \]  

(69)

Using \( \Delta z = \Delta z_0 + \delta \Delta z \) and the fact that \( \left( \frac{\partial \text{vec}(\mathcal{F}_p(z(\theta \alpha)))}{\partial \text{vec}(z)} - \frac{\partial \text{vec}(\mathcal{F}_p(z))}{\partial \text{vec}(z)} \right) \text{vec}(\Delta z_0) = -\text{vec}(\mathcal{F}_p(z)) \), it follows that

\[ \text{vec}(\mathcal{F}_p(z(\alpha))) = (1 - \alpha) \text{vec}(\mathcal{F}_p(z)) + \alpha \frac{\partial \text{vec}(\mathcal{F}_p(z))}{\partial \text{vec}(z)} \text{vec}(\delta \Delta z) + \alpha R \]  

(70)

Furthermore

\[ \|R\|_2 \leq \max_{\theta \in (0, 1)} \left\| \frac{\partial \text{vec}(\mathcal{F}_p(z(\theta \alpha)))}{\partial \text{vec}(z)} - \frac{\text{vec}(\mathcal{F}_p(z))}{\partial \text{vec}(z)} \right\|_2 \cdot \|\text{vec}(\Delta z)\|_2 \]  

(71)

Since \( \|\Delta z\|_2 \leq C \) and since \( \frac{\partial \text{vec}(\mathcal{F}_p(z))}{\partial \text{vec}(z)} \) is continuous, there exists an \( \hat{\alpha}^{(4)} > 0 \) such that for \( \alpha \in (0, \hat{\alpha}^{(4)}) \), \( \epsilon \in (0, 1) \) and \( z \in B \) it holds that

\[ \|R\|_2 < \frac{1 - \epsilon}{2} \sigma \beta \nu \]  

(72)

Using the fact that \( \|\mathcal{F}_p(z)\|_2 \leq \beta \nu \), it now follows that

\[ \|\mathcal{F}_p(z(\alpha))\|_2 \leq (1 - \alpha) \beta \nu + \alpha \frac{\sigma}{2} \beta \nu \]  

(73)

for all \( \alpha \in [0, \hat{\alpha}^{(4)}] \). By reducing \( \hat{\alpha}^{(4)} \), if necessary, it follows that

\[ \alpha C^2 < \frac{\sigma}{2} n \nu \]  

(74)

for all \( \alpha \in [0, \hat{\alpha}^{(4)}] \). Similar (53), but bounding below instead, it holds that

\[ n \nu(\alpha) \geq (1 - \alpha(1 - \sigma)) n \nu - \alpha^2 C^2 \iff \]  

\[ (1 - \alpha) \nu \leq \nu(\alpha) - \alpha \sigma \nu - \frac{\alpha^2 C^2}{n} \]  

(75)

Hence

\[ \|\mathcal{F}_p(z(\alpha))\|_2 \leq \beta \left( \nu(\alpha) - \alpha \sigma \nu - \frac{\alpha^2 C^2}{n} \right) + \alpha \frac{\sigma}{2} \beta \nu \]  

(76)

\[ = \beta \nu(\alpha) - \alpha \beta \left( \sigma \nu - \frac{\alpha C^2}{n} - \frac{\sigma \nu}{2} \right) \leq \beta \nu(\alpha) \]  

(77)

where the last inequality follows from (74). The proof for \( \mathcal{F}_d(z(\alpha)) \leq \beta \nu(\alpha) \) is done analogously, which gives an \( \hat{\alpha}^{(5)} > 0 \).

### 3.3 Proof of closed set

Here the proof of that the set \( \Omega \) defined in (29) defines a closed set is presented.
Definition 3.6 Let $X$ and $Y$ denote two metric spaces and define the mapping $f : X \to Y$. Let $C \subseteq Y$. Then the inverse image $f^{-1}(C)$ of $C$ is the set \( \{ x \mid f(x) \in C, x \in X \} \).

Lemma 3.7 A mapping of a metric space $X$ into a metric space $Y$ is continuous if and only if $f^{-1}(C)$ is closed in $X$ for every closed set $C \subseteq Y$.

Proof See page 81 in [9].

Lemma 3.8 The set \( \{ z \mid \|F_p(z)\|_2 \leq \beta \nu, z \in \mathcal{Z} \} \) is a closed set.

Proof Consider the mapping $F_p(z) : \mathcal{Z} \to \mathbb{S}^n$ and the set $C = \{ C \mid \|C\|_2 \leq \beta \nu, C \in \mathbb{S}^n \}$ that is closed since the norm defines a closed set, see page 634 in [1]. Now note that the mapping $F_p(z)$ is continuous. Then the inverse image \( \{ z \mid F_p(z) \in C, z \in \mathcal{Z} \} = \{ z \mid \|F_p(z)\|_2 \leq \beta \nu, z \in \mathcal{Z} \} \) is a closed set.

Lemma 3.9 The set \( \{ z \mid \|F_d(z)\|_2 \leq \beta \nu, z \in \mathcal{Z} \} \) is a closed set.

Proof Analogous with the proof of Lemma 3.8.

Lemma 3.10 The set \( \{ z \mid \gamma \nu I \succeq H(ZS) \succeq \eta \nu I, z \in \mathcal{Z} \} \) is a closed set.

Proof Define the mapping $h(z) = H(ZS) : \mathcal{Z} \to \mathbb{S}^n$ and the set $C_1 = \{ C_1 \mid C_1 \succeq \eta \nu I, C_1 \in \mathbb{S}^n \}$ that is closed, page 43 in [1]. Since the mapping is continuous the inverse image \( \{ z \mid H(ZS) \succeq \eta \nu I, z \in \mathcal{Z} \} \) is a closed set. Now define the set $C_2 = \{ C_2 \mid C_2 \preceq \gamma \nu I, C_2 \in \mathbb{S}^n \}$. Using continuity again gives that \( \{ z \mid H(ZS) \preceq \gamma \nu I, z \in \mathcal{Z} \} \) is a closed set. Since the intersection of closed sets is a closed set, Theorem 2.24 in [9], \( \{ z \mid \gamma \nu I \succeq H(ZS) \succeq \eta \nu I, z \in \mathcal{Z} \} \) is a closed set.

Lemma 3.11 The set $\Omega$ is a closed set.

Proof Note that $(\mathcal{Z}, \rho)$ is a metric space with distance $\rho(u, v) = \|u - v\|_2$. Hence is $\mathcal{Z}$ a closed set. Lemma 3.8, 3.9 and 3.10 gives that each additional constraint in $\Omega$ defines a closed set. Since the intersection of closed sets is a closed set, Theorem 2.24 in [9], $\Omega$ is a closed set.
References


Titel: An inexact interior-point method for semi-definite programming, a description and convergence proof

Författare: Janne Harju, Anders Hansson

Sammanfattning: In this report we investigate convergence for an infeasible interior-point method for semi-definite programming.

Nyckelord: LMI optimization; Interior-point methods; Iterative computation; convergence.