The central path as a set of analytic centers in semidefinite optimization.

Margaréta Halická

Institute of Applied Mathematics
Faculty of Mathematics and Physics
Comenius University
Bratislava, Slovakia
halicka@fmph.uniba.sk
http://pc2.iam.fmph.uniba.sk/institute/halicka/

Co-authors: E.de Klerk and K. Roos

OR seminar TU Delft
May 29, 2001

Outline

1. Preliminaries

1.1 Primal-dual pair of semidefinite programs

S^n – the space of real symmetric $n \times n$ matrices

$$M \bullet N = \text{trace}(M^T N) = \sum_{i,j} m_{ij} n_{ij}$$

$X \succeq 0$ (X symm. posit. semidefinite) $X \in S^n_+ \subset S^n$

$X \succ 0$ (X symm. posit. definite) $X \in S^n_{++} \subset S^n$

For given C, $A^i \in S^n$, $i = 1, ..., m$, and $b \in R^m$ we consider the primal-dual pair of semidefinite programs

$$\begin{aligned} \min_{X \in S^n} \quad & C \bullet X \\ \text{s.t.} \quad & A^i \bullet X = b^i, \quad i = 1, ..., m, \\ & X \succeq 0, \end{aligned} \tag{P}$$

$$\begin{aligned} \max_{y \in \mathbb{R}^m, Z \in S^n} \quad & b^T y \\ \text{s.t.} \quad & \sum_{i=1}^{m} A^i y^i + Z = C, \\ & Z \succeq 0. \end{aligned} \tag{D}$$

Feasible sets:

$\mathcal{P}$ - the set of feasible solution for (P)

$\mathcal{D}$ - the set of feasible solutions for (D), - i.e.

$$\mathcal{P} = \left\{ X \ : \ A^i \bullet X = b^i, \ i = 1, ..., m, \ X \succeq 0, \right\}$$
$$\mathcal{D} = \left\{ (y, Z) \ : \ \sum_{i=1}^{m} A^i y^i + Z = C, \ Z \succeq 0, \ y \in \mathbb{R}^n \right\}$$

Optimal sets:

$\mathcal{P}^*$ - the set of optimal solution for (P)

$\mathcal{D}^*$ - the set of optimal solutions for (D)

Duality gap:

$$X \in \mathcal{P}, \ (y, Z) \in \mathcal{D} \quad : \quad C \bullet X - b^T y = X \bullet Z \geq 0$$

Orthogonality property:

$$X^1, X^2 \in \mathcal{P}, \ Z^1, Z^2 \in \mathcal{D} \quad : \quad (X^1 - X^2) \bullet (Z^1 - Z^2) = 0$$

Assumption 1:

A^i, $i = 1, ..., m$, are linearly independent, i.e.,

$$\sum_{i=1}^{m} A^i u^i = 0 \quad \Rightarrow \quad u^i = 0, \quad i = 1, ..., m.$$

Assumption 2:

Both (P) and (D) have strictly feasible solutions, i.e.,

$$\exists X \in S^n \ : \qquad A^i \bullet X = b^i, \quad i = 1, ..., m, \quad X \succ 0;$$

$$\exists (y, Z) \in R^m \times S^n \ : \qquad \sum_{i=1}^{m} A^i y^i + Z = C, \quad Z \succ 0.$$

Facts:

(a) (P) and (D) have optimal solutions $\hat{X}$ and $(\hat{y}, \hat{Z})$;

(b) $C \bullet \hat{X} - b^T \hat{y} = \hat{X} \bullet \hat{Z} = 0$;

(c) $\hat{X}\hat{Z} = \hat{Z}\hat{X} = 0_{n\times n}$.

System of conditions for optimality:

primal feasibility	$A^i \bullet X = b^i, \ \forall i$	$X \succeq 0,$
dual feasibility	$\sum_{i=1}^{m} A^i y^i + Z = C,$	$Z \succeq 0,$
complementarity	$XZ = 0_{n\times n}.$	

Complementarity in SDO:

The property (c) ($XZ = ZX = 0_{n\times n}$) implies the complementarity of corresponding eigenvalues

Definition:

$(X, Z) \in \mathcal{P}^* \times \mathcal{D}^*$ is a **strictly complementary solution pair** iff rank (X)+rank $(Z) = n$.

Fact:

The existence of a strictly complementary solution pair is not guaranteed in SDO.

Definition:

$(X, Z) \in \mathcal{P}^* \times \mathcal{D}^*$ is a **maximally complementary solution pair** iff it maximizes rank (X)+rank (Z) over all optimal solution pairs.

1.2 Central path

For any $\mu > 0$ the **centering system** is defined as

$$
\begin{array}{lll}
\text{primal strict feas.} & A^i \bullet X = b^i, \ \forall i & X \succ 0, \\
\text{dual strict feas.} & \displaystyle\sum_{i=1}^{m} A^i y^i + Z = C, & Z \succ 0, \\
\text{perturbed compl.} & XZ = \mu I_{n\times n}. &
\end{array}
$$

Fact: For any $\mu > 0$ there exists a unique solution $(X(\mu), y(\mu), Z(\mu))$ to the centering system.

Central path: $\{ \ (X(\mu), y(\mu), Z(\mu)) \ \}_{\mu>0}$

Duality gap along the central path: $X(\mu) \bullet Z(\mu) = \mu n$

Facts:

(a) The central path restricted to $0 < \mu \leq \bar{\mu}$ $(\bar{\mu} > 0)$ is bounded.

(b) Each of the limit points of the central path for $\mu \downarrow 0$ is an optimal solution of (P), (D).

If the optimal set in not a singleton

(a) convergence of the central path

(b) characterization of the limit point

Fact: The central path converges as $\mu \downarrow 0$.

Proof by means of the theory of (semi)algebraic sets. Similar argumentation used by [Kojima et al : Lecture Notes in Computer Science 538,(1991)] in the context of LCP and SLCP.

Fact: Each of the limit points of the central path for $\mu \downarrow 0$ is a maximally complementary solution of (P), (D).

[De Klerk, Roos, Terlaky: Oper. Res. Lett. 20,(1997), 213-221]
[Goldfarb, Scheinberg: SIAM J. Opt. 8 (1998), 871-886]

Fact: The central path converges as $\mu \downarrow 0$ to the so-called analytic center of the optimal set, provided a strictly complementary solution exists.

[De Klerk, Roos, Terlaky: Topics in SD&IPM, 18 (1998) 215-236]
[Luo, Sturm, Zhang: SIAM J. Opt. 8, (1998), 59-81]

Fact ? : The central path converges as $\mu \downarrow 0$ to the so-called analytic center of the optimal set.

[Goldfarb, Scheinberg: SIAM J. Opt. 8 (1998), 871-886]

Idea of the proof: At any $\mu > 0$ the central path is the analytic center of a level set and keeps this property as $\mu \downarrow 0$. Gaps in the proof !

In what follows we restrict ourselves to the primal case.

2. Analytic centers

2.1 The central path as a set of analytic centers

For any $\mu > 0$ we define **the level set:**

$$M_\mu = \left\{ X : \begin{array}{c} A^i \bullet X = b_i \ (i = 1, \dots, m), \ X \in S^n_+, \\ C \bullet X - b^T y(\mu) = \mu n (= X \bullet Z(\mu)) \end{array} \right\}.$$

Since $X(\mu) \in M_\mu$ and $X(\mu) \succ 0$ we have that

$$\text{ri } M_\mu = \{X \in M_\mu \ : \ X \in S^n_{++}\} \neq \emptyset.$$

Fact: M_μ is convex, bounded.

For any $\mu > 0$ we define the following problem (C1):

$$\max_{X \in \text{ri} M_\mu} \ln \det X \qquad \text{(C1)}$$

Fact: There exists the unique solution to (C1) and it is called the **analytic center** of M_μ.

Proof:

Uniqueness: $\ln \det X$ is strictly concave.

Existence: The set $\{X \in \text{ri} M_\mu \ : \ \ln \det X \geq \ln \det X(\mu)\}$ is compact.

Fact. *At any $\mu > 0$, $X(\mu)$ is the analytic center of M_μ-i.e. $X(\mu)$ is the optimal solution to (C1).*

[E de Klerk, Ph.D. thesis (1997), Lemma 2.3.2] for the primal-dual case

Proof: Let $X \in \mathrm{ri}\ M_\mu$. Then

$$C \bullet X - b^T y(\mu) = X \bullet Z(\mu) = n\mu.$$

Dividing by $\mu > 0$ and substituting $\frac{Z(\mu)}{\mu} = X^{-1}(\mu)$ we obtain

$$X \bullet X^{-1}(\mu) = n$$

Applying the arithmetic-geometric inequality yields

$$\left[\det(X X^{-1}(\mu))\right]^{\frac{1}{n}} \leq \frac{1}{n}(X \bullet X^{-1}(\mu)) = 1$$

which implies the required result. $\square$

Let (y^*, Z^*) be a limit point of $(y(\mu), Z(\mu)$ as $\mu \downarrow 0$. Then

$$M^* = \left\{ X : \begin{array}{c} A^i \bullet X = b_i \ (\ i = 1, \ldots, m),\ X \in S^n_+, \\ C \bullet X - b^T y^* = 0 (= X \bullet Z^*) \end{array} \right\} = \mathcal{P}^*$$

Hence $\mathcal{P}^*$ is the limit set of M_μ as $\mu \downarrow 0$. This description of $\mathcal{P}^*$ does not allow to define the analytic center of $\mathcal{P}^*$.

2.2 Optimal partition

Let $(\bar{X}, \bar{Z})$ be a maximally complementary solution and

$$|B| := \text{rank } \bar{X}, \quad \text{and} \quad |N| := \text{rank } \bar{Z}$$

Obviously, $|B| + |N| \leq n$.

Since $\bar{X}\bar{Z} = \bar{Z}\bar{X}$, there exist common spectral decomposition of $\bar{X}$ and $\bar{Z}$. Therefore, we can assume without loss of generality (applying an orthonormal transformation of problem data, if necessary) that

$$\bar{X} = \begin{bmatrix} \bar{\Lambda}_B & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \bar{Z} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & \bar{\Omega}_N & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

where $\bar{\Lambda}_B \in S_{++}^{|B|}$ and $\bar{\Omega}_N \in S_{++}^{|N|}$ are diagonal matrices.

Fact: Each optimal solution pair $(\hat{X}, \hat{Z})$ is of the form

$$\hat{X} = \begin{bmatrix} \hat{X}_B & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \hat{Z} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & \hat{Z}_N & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

where $\hat{X}_B \in S_{+}^{|B|}$ and $\hat{Z}_N \in S_{+}^{|N|}$.

In what follows we consider the partition of any $M \in S^n$ corresponding to the above optimal partition so that

$$M = \begin{bmatrix} M_B & M_{BN} & M_{BT} \\ M_{NB} & M_N & M_{NT} \\ M_{TB} & M_{TN} & M_T \end{bmatrix}$$

Denote $\mathcal{I} = \{B, BN, BT, NB, N, NT, TB, TN, T\}$ the index set corresponding to the optimal partition. Referring to the all blocs of M except of M_B we will write

$$M_i \ (i \in \mathcal{I} - B).$$

By the above partition, the primal and dual feasibility equations can be rewritten into

$$\begin{aligned} \sum_{j \in \mathcal{I}} A^i_j \bullet X_j =& b_i \ (\ i = 1, \ldots, n), \\ \sum_{i=1}^{m} A^i_j y_i + Z_j =& C_j \ (j \in \mathcal{I}). \end{aligned}$$

2.3 Optimal sets

The optimal sets can be characterized as seen below. Here X and Z are considered in the block partition just mentioned above.

$$\mathcal{P}^* = \left\{ X : \begin{array}{l} A_B^i \bullet X_B = b_i \ (\ i = 1, \ldots, m), \ X_B \in S_+^{|B|}, \\ \\ X_k = 0 \ (k \in \mathcal{I} - B) \end{array} \right\},$$

$$\mathcal{D}^* = \left\{ (Z, y) : \begin{array}{l} \sum_{i=1}^m A_N^i y_i + Z_N = C_N, \ Z_N \in S_+^{|N|}, \\ \\ \sum_{i=1}^m A_k^i y_i = C_k, \ Z_k = 0 \ (k \in \mathcal{I} - N) \end{array} \right\}.$$

Since both $\mathcal{P}^*$ and $\mathcal{D}^*$ are non-empty, we have

Range property: $b \in \mathcal{R}(\mathcal{A}_B)$ and $C_B \in \mathcal{R}(\mathcal{A}_B^T)$

$\mathcal{R}(\mathcal{A}_B)$ is the range of the linear operator $\mathcal{A}_B$

($\mathcal{A}_B$ maps $X_B \in S^{|B|}$ to the m-vector

with components $A_B^i \bullet X_B$ ($i = 1, \ldots, m$)).

$\mathcal{R}(\mathcal{A}_B^T)$ is the range of the linear operator $\mathcal{A}_B^T$

($\mathcal{A}_B^T$ maps $y \in \mathbb{R}^m$ to $\sum_{i=1}^m A_B^i y_i$).

Fact: Both $\mathcal{P}^*$ and $\mathcal{D}^*$ are bounded.

2.4 Analytic center of the optimal set

$$\mathcal{P}^* = \left\{ X : \begin{array}{l} A_B^i \bullet X_B = b_i \ (\ i = 1, \ldots, m), \ X_B \in S_+^{|B|}, \\ X_k = 0 \ (k \in \mathcal{I} - B) \end{array} \right\},$$

$\mathcal{P}^*$ is bounded. Let X^* be a limit point of $X(\mu)$ as $\mu \downarrow 0$. Then $X^* \in \mathcal{P}^*$ and $X_B^* \in S_{++}^{|B|}$. Hence

$$\text{ri } \mathcal{P}^* = \left\{ X \in \mathcal{P}^* \ : \ X_B \in S_{++}^{|B|} \right\} \neq \emptyset$$

Fact: There exists the unique solution to the problem

$$\max_{X \in \text{ri } \mathcal{P}^*} \ln \det X_B \tag{C2}$$

and is called **the analytic center of $\mathcal{P}^*$**.

By using the Lagrange rule we obtain:

Lemma 1. *Let $X^a \in \mathcal{P}^*$, $X_B^a \succ 0$. Then the following claims are equivalent:*

(a) X^a is the analytic center of $\mathcal{P}^*$,

(b) $\exists \quad u^i (i = 1, ..., m) \quad : \quad (X_B^a)^{-1} - \sum_{i=1}^m u^i A_B^i = 0.$

(c) $(X_B^a)^{-1} \in \mathcal{R}(\mathcal{A}_B^T)$.

3. What's wrong with the proof by G&S?

3.1 Idea of the proof

Let X^* be an arbitrarily chosen limit point of the primal central path as $\mu \downarrow 0$. Since $X^* \in \mathcal{P}^*$, $X^*_B \succ 0$ it is a good candidate for the analytic center of $\mathcal{P}^*$.

Claim. *X^* is the analytic center of $\mathcal{P}^*$.*

Since the analytic center of $\mathcal{P}^*$ is defined uniquely we obtain

Corollary. *The central path $X(\mu)$ converges as $\mu \downarrow 0$, i.e. there exist finite $\lim_{\mu\downarrow 0} X(\mu)$.*

We have proved: $X(\mu)$ is the solution to (C1).

$X(\mu_t) \to X^*$ for some $\{\mu_t\}$.

We should prove: X^* is the optimal solution to (C2).

• Dimensions of the level sets in (C1) and (C2) are given by dimensions of the corresponding affine spaces.

• Dimension of the level sets (M_μ in (C1)) is changed at the limit ($\mathcal{P}^*$ in (C2)).

Idea: Restriction of (C1) to a problem of the same dimension as (C2). (The off diagonal blocs make trouble.)

3.2 Proof by Goldfarb and Scheinberg

Let the spectral decomposition of $X(\mu)$ is denoted as

$$X(\mu) = Q(\mu)\Lambda(\mu)Q(\mu)^T,\ Q(\mu)Q^T(\mu) = I,\ \Lambda \in D^n_{++}.$$

Without loss of generality we assume that $\{\mu_t\}$ is such that

$$X(\mu_t) \to X^*,\ Q(\mu_t) \to Q^*,\ \Lambda(\mu_t) \to \Lambda^*,$$

where

$$X^* = Q^*\Lambda^*Q^{*T}.$$

Any $X \in \mathrm{ri}\ M_\mu$ can be rewrite as $X = Q(\mu)\Lambda Q^T(\mu)$ for some $\Lambda \in S^n_{++}$. Hence (C1)(at $\mu > 0$) can be rewritten as

$$\begin{aligned} \max_{\Lambda \succ 0,} \quad & \ln\det\Lambda \\ \text{s.t.} \quad & Q^T(\mu)A^iQ(\mu) \bullet \Lambda = b_i\ (i = 1, ..., m), \\ & Q^T(\mu)CQ(\mu) \bullet \Lambda - b^Ty(\mu) = \mu n. \end{aligned}$$

Obviously $\Lambda(\mu)$ is the solution to this problem. We now assume that $Q(\mu) = (q_B(\mu), q_N(\mu), q_T(\mu))$ and we restrict our problem to matrices $\Lambda \in S^n_{++}$ which are block diagonal with matrices Λ_B, Λ_N and Λ_T on the diagonal.

Hence we obtain

$$
\begin{array}{ll}
\max\limits_{\Lambda \succ 0,} & \ln\det\Lambda \\
\text{s.t.} & q_B^T(\mu)A^i q_B(\mu)\bullet\Lambda_B + q_N^T(\mu)A^i q_N(\mu)\bullet\Lambda_N + \\
& \qquad q_T^T(\mu)A^i q_T(\mu)\bullet\Lambda_T = b_i \ (i=1,...,m), \\
& q_B^T(\mu)Cq_B(\mu)\bullet\Lambda_B + q_N^T(\mu)Cq_N(\mu)\bullet\Lambda_N + \\
& \qquad q_T^T(\mu)Cq_T(\mu)\bullet\Lambda_T - b^T y(\mu) = \mu n.
\end{array}
$$

Obviously, $\Lambda(\mu)$ is the optimal solution to the above problem. This problem can be further restricted to

$$
\begin{array}{ll}
\max\limits_{\Lambda_B \succ 0} & \ln\det\Lambda_B \\
\text{s.t.} & q_B^T(\mu)A^i q_B(\mu)\bullet\Lambda_B + q_N^T(\mu)A^i q_N(\mu)\bullet\Lambda_N(\mu) + \\
& \qquad q_T^T(\mu)A^i q_T(\mu)\bullet\Lambda_T(\mu) = b_i \ (i=1,...,m), \\
& q_B^T(\mu)Cq_B(\mu)\bullet\Lambda_B + q_N^T(\mu)Cq_N(\mu)\bullet\Lambda_N(\mu) + \\
& \qquad q_T^T(\mu)Cq_T(\mu)\bullet\Lambda_T(\mu) - b^T y(\mu) = \mu n.
\end{array}
$$

Since $\Lambda_B(\mu)$ is the optimal solution to this problem, by the Lagrange rule there exist u_μ^i $(i=1,...,m)$ and v_μ such that

$$
\Lambda_B^{-1}(\mu) - \sum_{i=1}^m u_\mu^i q_B^T(\mu)A^i q_B(\mu) - v_\mu q_B^T(\mu)Cq_B(\mu) = 0.
$$

Goldfarb and Sheinberg: *"by the limit approach we obtain the necessary and sufficient condition for the analytic center of the primal optimal solution set."*

$$\Lambda_B^{-1}(\mu) - \sum_{i=1}^{m} u_\mu^i q_B^T(\mu) A^i q_B(\mu) - v_\mu q_B^T(\mu) C q_B(\mu) = 0.$$

To claim this, **the boundedness of the Lagrange multipliers** u_μ^i $(i = 1, ..., m)$ and v_μ as $\mu \downarrow 0$ is needed!

Another approach can be used:

$$\Lambda_B^{-1}(\mu) \in \mathcal{R}(q_B^T(\mu) A^T q_B(\mu), q_B^T(\mu) C q_B(\mu))$$

for any $\mu > 0$. We also have that

$$\mathcal{R}(q_B^T(\mu_k) A^T q_B(\mu_k), q_B^T(\mu_k) C q_B(\mu_k)) \to \mathcal{R}(A_B^T, C_B)$$

and

$$\Lambda_B^{-1}(\mu_k) \to (\Lambda_B^*)^{-1} = (X_B^*)^{-1}.$$

However, merely from the above facts we are not allowed to claim that

$$(X_B^*)^{-1} \in \mathcal{R}(A_B^T, C_B)$$

in general.

4. Two proofs which fail for the general case

There are three proof techniques in **linear optimization**:

(1) The central path as a set of analytic centers of level sets; characterization of the analytic center.

(2) Orthogonality property and arithmetic-geometric inequality.

(3) Closedness of affine spaces and characterization of the analytic center

Contrary to (1), the proof techniques (2) and (3) use the inverze along the central path:

$$\frac{z_B(\mu)}{\mu} = x_B^{-1}(\mu) > 0$$

$$\frac{x_N(\mu)}{\mu} = z_N^{-1}(\mu) > 0$$

In **semidefinite optimization** analogous properties do not hold along the central path. They hold at the limit point, however, only if the strictly complementary solution exists.

4.1 Inverze along the central path

From the perturbed complementarity at any $\mu > 0$:

$$\frac{Z_B(\mu)}{\mu} = \left(X^{-1}(\mu)\right)_B \quad \text{and} \quad \frac{X_N(\mu)}{\mu} = \left(Z^{-1}(\mu)\right)_N .$$

Lemma 2. *Both* $\frac{Z_B(\mu_t)}{\mu_t}$ *and* $\frac{X_N(\mu_t)}{\mu_t}$ *converge as* $t \to \infty$ *(where, if necessary, we have taken a subsequence of* $\{\mu_t\}$*)to positive definite matrices. Moreover,*

(a) if $|B| + |N| = n$*, then following two formulas hold*

$$\lim_{t\to\infty} \frac{Z_B(\mu_t)}{\mu_t} = \lim_{t\to\infty} \left(X^{-1}(\mu_t)\right)_B = (X_B^*)^{-1},$$
$$\lim_{t\to\infty} \frac{X_N(\mu_t)}{\mu_t} = \lim_{t\to\infty} \left(Z^{-1}(\mu_t)\right)_N = (Z_N^*)^{-1}.$$

(b) If $|B|+|N| < n$*, then at least one of the above formulas does not hold, i.e.,*

$$\lim_{t\to\infty} \frac{Z_B(\mu_t)}{\mu_t} = \lim_{t\to\infty} \left(X^{-1}(\mu_t)\right)_B \neq (X_B^*)^{-1} \ \textit{or}$$
$$\lim_{t\to\infty} \frac{X_N(\mu_t)}{\mu_t} = \lim_{t\to\infty} \left(Z^{-1}(\mu_t)\right)_N \neq (Z_N^*)^{-1}$$

Proof of Lemma 2

Since $X(\mu)$ and $Z(\mu)$ commute, there exists an $n \times n$ matrix $Q(\mu)$ such that $Q(\mu)Q(\mu)^T = I$ and

$$X(\mu) = Q(\mu)\Lambda(\mu)Q(\mu)^T, \quad Z(\mu) = Q(\mu)\Omega(\mu)Q(\mu)^T$$

where $\Lambda(\mu)$ and $\Omega(\mu)$ are diagonal matrices with the eigenvalues of $X(\mu)$ and $Z(\mu)$ on the diagonal. Applying the optimal partition we obtain

$$Z_B(\mu) = Q_B\Omega_B Q_B^T + Q_{BN}\Omega_N Q_{BN}^T + Q_{BT}\Omega_T Q_{BT}^T$$
$$X_N(\mu) = Q_{NB}\Lambda_B Q_{NB}^T + Q_N\Lambda_N Q_N^T + Q_{NT}\Lambda_T Q_{NT}^T.$$

From the perturbed complementarity we have

$$\mu I = X(\mu)Z(\mu) = Q(\mu)\Lambda(\mu)\Omega(\mu)Q^T(\mu)$$
$$\mu I = \Lambda(\mu)\Omega(\mu)$$

and hence

$$\frac{\Omega_B(\mu)}{\mu} = \Lambda_B^{-1}(\mu) \text{ and } \frac{\Lambda_N(\mu)}{\mu} = \Omega_N^{-1}(\mu)$$

Without loss of generality we assume that the sequence $\{\mu_t\}$, for which the central path converges to (X^*, Z^*, y^*) has been chosen such that

$$Q(\mu_t) \to Q^*, \quad \Lambda(\mu_k) \to \Lambda^* \text{ and } \Omega(\mu_k) \to \Omega^*$$

where Q^*, Λ^* and Ω^* define the SF of X^* and Z^*. Hence

$$Q^* Q^{*T} = I,\ X^* = Q^* \Lambda^* Q^{*T} \text{ and } Z^* = Q^* \Omega^* Q^{*T}$$
$$\Lambda_N^* = 0,\ \Lambda_T^* = 0,\ \Omega_B^* = 0,\ \Omega_T^* = 0$$

and $\Lambda_B^* \succ 0$ and $\Omega_N^* \succ 0$ are diagonal matrices of positive eigenvalues of X^* and Z^*, respectively.

Moreover, we can assume that the orthonormal transformation of the problem data mentioned above was performed with respect to the maximally complementary pair (X^*, Z^*) and the orthonormal matrix Q^*. Hence now, as a consequence of the transformation, we have that

$$Q^* = I_n,\ X_B^* = \Lambda_B^* \text{ and } Z_N^* = \Omega_N^*.$$

Since $X^*, X(\mu) \in \mathcal{P}$ and $Z^*, Z(\mu) \in \mathcal{D}$ we have by the orthogonality property for any $\mu > 0$

$$(X^* - X(\mu)) \bullet (Z^* - Z(\mu)) = 0$$

and hence

$$X^* \bullet \frac{Z(\mu)}{\mu} + Z^* \bullet \frac{X(\mu)}{\mu} = X(\mu) \bullet Z(\mu) + X^* \bullet Z^*.$$

Substituting $X(\mu) \bullet Z(\mu) = \mu n$ and $X^* \bullet Z^* = 0$ and dividing the resulting equation by $\mu > 0$ we get

$$X^* \bullet \frac{Z(\mu)}{\mu} + Z^* \bullet \frac{X(\mu)}{\mu} = n.$$

Now we use the fact that $X_k^* = 0$ for all $k \in \mathcal{I} - B$ and $Z_k^* = 0$ for all $k \in \mathcal{I} - N$. Hence the last equality gives

$$X_B^* \bullet \frac{Z_B(\mu)}{\mu} + Z_N^* \bullet \frac{X_N(\mu)}{\mu} = n.$$

From the above formulas for $X_N(\mu)$ and $Z_B(\mu)$ we obtain

$$\frac{Z_B}{\mu} = Q_B \frac{\Omega_B}{\mu} Q_B^T + Q_{BN} \frac{\Omega_N}{\mu} Q_{BN}^T + Q_{BT} \frac{\Omega_T}{\mu} Q_{BT}^T$$

$$\frac{X_N}{\mu} = Q_{NB} \frac{\Lambda_B}{\mu} Q_{NB}^T + Q_N \frac{\Lambda_N}{\mu_k} Q_N^T + Q_{NT} \frac{\Lambda_T}{\mu_k} Q_{NT}^T.$$

Substituting these formulas we obtain the formula (R):

$$\Lambda_B^* \bullet \left(Q_B \frac{\Omega_B}{\mu} Q_B^T + Q_{BN} \frac{\Omega_N}{\mu} Q_{BN}^T + Q_{BT} \frac{\Omega_T}{\mu} Q_{BT}^T \right) +$$
$$\Omega_N^* \bullet \left(Q_{NB} \frac{\Lambda_B}{\mu} Q_{NB}^T + Q_N \frac{\Lambda_N}{\mu} Q_N^T + Q_{NT} \frac{\Lambda_T}{\mu} Q_{NT}^T \right) = n.$$

Consider (R) along the $\{\mu_t\}$. Since $Q(\mu_t) \to I_n$,

$$Q_B(\mu_t) \to I_{|B|} \text{ and } Q_N(\mu_t) \to I_{|N|}.$$

Moreover,

$$\begin{aligned} \frac{\Omega_B(\mu_t)}{\mu_t} &= \Lambda_B^{-1}(\mu_t) \to (\Lambda_B^*)^{-1} \succ 0 \\ \frac{\Lambda_N(\mu_t)}{\mu_t} &= \Omega_N^{-1}(\mu_t) \to (\Omega_N^*)^{-1} \succ 0 \end{aligned}$$

Thus we have that

$$\begin{aligned} \lim_{\mu_t \downarrow 0} Q_B(\mu_t) \frac{\Omega_B(\mu_t)}{\mu} Q_B^T(\mu_t) &= (\Lambda_B^*)^{-1} \\ \lim_{\mu_t \downarrow 0} Q_N(\mu_t) \frac{\Lambda_N(\mu_t)}{\mu_t} Q_N^T(\mu_t) &= (\Omega_N^*)^{-1} \end{aligned}$$

Since all items in (R) are non–negative and in sum they give n, each of them is bounded as $\{\mu_t\} \downarrow 0$. Thus, taking a subsequence of $\{\mu_t\}$, if necessary, we have a convergence of all items in (R).

Taking the limit in (R) we obtain

$$\begin{aligned}
&\Lambda_B^* \bullet (\Lambda_B^*)^{-1} + \Omega_N^* \bullet (\Omega_N^*)^{-1} + \\
&\Lambda_B^* \bullet \lim_{t\to\infty} \left(Q_{BN} \frac{\Omega_N(\mu_t)}{\mu_t} Q_{BN}^T + Q_{BT} \frac{\Omega_T(\mu_t)}{\mu_t} Q_{BT}^T \right) + \\
&\Omega_N^* \bullet \lim_{t\to\infty} \left(Q_{NB} \frac{\Lambda_B(\mu_t)}{\mu_t} Q_{NB}^T + Q_{NT}^t \frac{\Lambda_T(\mu_t)}{\mu_t} Q_{NT}^T \right) = n
\end{aligned}$$

It is easy to see that

$$\Lambda_B^* \bullet (\Lambda_B^*)^{-1} = |B|, \qquad \Omega_N^* \bullet (\Omega_N^*)^{-1} = |N|,$$

and

$$\Lambda_B^* \bullet \lim_{t\to\infty} \left(Q_{BN} \frac{\Omega_N(\mu_t)}{\mu_t} Q_{BN}^T + Q_{BT} \frac{\Omega_T(\mu_t)}{\mu_t} Q_{BT}^T \right) \geq 0$$

$$\Omega_N^* \bullet \lim_{t\to\infty} \left(Q_{NB} \frac{\Lambda_B(\mu_t)}{\mu_t} Q_{NB}^T + Q_{NT} \frac{\Lambda_T(\mu_t)}{\mu_t} Q_{NT}^T \right) \geq 0.$$

If $|B| + |N| = n$, then the two inequalities must hold as equalities and the statement (a) of the lemma is proved.

If $|B|+|N| < n$, then at least one of the inequalities holds as a strict inequality and thus the statement (b) of the lemma holds as well. $\square$

4.2 The first proof.

Let $\bar{X} \in \mathcal{P}^*$ and $\bar{Z} \in \mathcal{D}^*$. Since for any $\mu > 0$, $X(\mu) \in \mathcal{P}$ and $Z(\mu) \in \mathcal{P}$, the orthogonality property implies

$$(\bar{X} - X(\mu)) \bullet (\bar{Z} - Z(\mu)) = 0, \quad \text{for any } \mu > 0.$$

Hence

$$\bar{X} \bullet \frac{Z(\mu)}{\mu} + \bar{Z} \bullet \frac{X(\mu)}{\mu} = X(\mu) \bullet Z(\mu) + \bar{X} \bullet \bar{Z},$$

By $X(\mu)Z(\mu) = \mu I_n$ we have that $X(\mu) \bullet Z(\mu) = \mu n$. Moreover, $\bar{X} \bullet \bar{Z} = 0$. Therefore

$$\bar{X} \bullet \frac{Z(\mu)}{\mu} + \bar{Z} \bullet \frac{X(\mu)}{\mu} = n, \quad \text{for any } \mu > 0.$$

Now we use the fact that $\bar{X}_k = 0$ for all $k \in \mathcal{I} - B$ and $\bar{Z}_k = 0$ for all $k \in \mathcal{I} - N$. Hence for any $\mu > 0$

$$\bar{X}_B \bullet \frac{Z_B(\mu)}{\mu} + \bar{Z}_N \bullet \frac{X_N(\mu)}{\mu} = n.$$

Since by Lemma 2 $\frac{Z_B(\mu_t)}{\mu_t}$ and $\frac{X_N(\mu_t)}{\mu_t}$ converge we have

$$\bar{X}_B \bullet \lim_{t\to\infty} \frac{Z_B(\mu_t)}{\mu_t} + \bar{Z}_N \bullet \lim_{t\to\infty} \frac{X_N(\mu_t)}{\mu_t} = n.$$

If $|B| + |N| = n$, by Lemma 2(a) we get

$$\frac{1}{n}\left(\bar{X}_B \bullet (X_B^*)^{-1} + \bar{Z}_N \bullet (Z_N^*)^{-1}\right) = 1.$$

Using the algebraic geometric mean inequality we obtain

$$\det\left(\bar{X}_B (X_B^*)^{-1}\right) \det\left(\bar{Z}_N (Z_N^*)^{-1}\right) \leq 1.$$

Since $\bar{X}_B \in \mathcal{P}^*$ and $\bar{Z}_N \in \mathcal{D}^*$ were arbitrarily chosen we may substitute $\bar{Z}_N = Z_N^*$ into the last equation to obtain

$$\det\left(\bar{X}_B (X_B^*)^{-1}\right) \det\left(Z_N^* (Z_N^*)^{-1}\right) \leq 1$$

which means that

$$\det \bar{X}_B \leq \det X_B^* \quad \text{for all } \bar{X} \in \mathcal{P}^*$$

and hence X^* is the analytic center of $\mathcal{P}^*$.

If $|B| + |N| < n$, Lemma 2(b) does not allow to prove the required result.

4.3 The second proof.

For $Z_B(\mu)$ and the corresponding limit $Z_B^* = 0$ we have

$$\begin{aligned} \sum_{i=1}^{m} A_B^i y(\mu) + Z_B(\mu) &= C_B \\ \sum_{i=1}^{m} A_B^i y^* &= C_B. \end{aligned}$$

Subtracting the latter from the former and dividing by μ we obtain

$$\sum_{i=1}^{m} A_B^i \frac{y(\mu) - y^*}{\mu} + \frac{Z_B(\mu)}{\mu} = 0$$

and hence

$$\frac{Z_B(\mu)}{\mu} \in \mathcal{R}(\mathcal{A}_B^T), \quad \forall \mu > 0.$$

The space $\mathcal{R}(A_B^T)$ is closed, and since by Lemma 2 the sequence $\frac{Z_B(\mu_t)}{\mu_t}$ converges, we have

$$\lim_{t \to \infty} \frac{Z_B(\mu_t)}{\mu_t} \in \mathcal{R}(\mathcal{A}_B^T).$$

If $|B| + |N| = n$, we use Lemma 2(a). Accordingly

$$(X_B^*)^{-1} = \lim_{t \to \infty} \frac{Z_B(\mu_t)}{\mu_t} \in \mathcal{R}(\mathcal{A}_B^T)$$

and by Lemma 1 X^* is the analytic center of $\mathcal{P}^*$.

If $|B| + |N| < n$, the statement (b) of Lemma 2 does not allow to derive the required result.

5.Back to the analytic centers of level sets

5.1 Reformulation of the problem

Consider the following restriction of (C1) denoted as (C3):

$$\begin{aligned} \max_{X_{B,\mu} \in S^n_{++}} \quad & \ln\det X_{B,\mu} \\ \text{s.t.} \quad & A^i_B \bullet X_B + \sum_{j \in \mathcal{I}-B} A^i_j \bullet X_j(\mu) = b^i, \ \forall i, \\ & C_B \bullet X_B + \sum_{j \in \mathcal{I}-B} C_j \bullet X_j(\mu) - b^T y(\mu) = \mu n \end{aligned} \tag{C3}$$

where

$$X_{B,\mu} := \begin{bmatrix} X_B & X_{BN}(\mu) & X_{BT}(\mu) \\ X_{NB}(\mu) & X_N(\mu) & X_{NT}(\mu) \\ X_{TB}(\mu) & X_{TN}(\mu) & X_T(\mu) \end{bmatrix}$$

Obviously $X(\mu)$ is the optimal solution to (C3).

In what follows we rewrite (C3) into an equivalent problem.

It is easy to see that $X_{B,\mu}$ can be rewritten as

$$X_{B,\mu} = \begin{bmatrix} X_B & X_U(\mu) \\ X_U^T(\mu) & X_V(\mu) \end{bmatrix}$$

for some $X_U(\mu)$ and $X_V(\mu)$. Since $X_V(\mu) \in S^{n-|B|}_{++}$

$$X_{B,\mu} \in S^n_{++} \quad \Leftrightarrow \quad X_B - F(\mu) \in S^{|B|}_{++}$$

where

$$F(\mu) := X_U(\mu) X_V^{-1}(\mu) X_U^T(\mu) \in S^{|B|}_{++}.$$

By using the formula for the determinant of a block matrix:

$$\det X_{B,\mu} = \det X_V(\mu) \det \left(X_B - F(\mu) \right).$$

Hence (C3) can be rewritten as

$$\begin{aligned} \max_{X_B - F(\mu) \in S_{++}^{|B|}} \quad & \ln \det (X_B - F(\mu)) \\ \text{s.t.} \quad & A_B^i \bullet X_B + \sum_{j \in \mathcal{I} - B} A_j^i \bullet X_j(\mu) = b^i, \ \forall i, \\ & C_B \bullet X_B + \sum_{j \in \mathcal{I} - B} C_j \bullet X_j(\mu) - b^T y(\mu) = \mu n \end{aligned} \tag{C4}$$

Lemma 3. *For any $\mu > 0$, $X_B(\mu)$ is the optimal solution to (C4), and $(X_B(\mu) - F(\mu))^{-1} \in \mathcal{R}(\mathcal{A}^T)$.*

Proof: Since $X(\mu)$ is an optimal solution to (C3), from the above derivation of (C4) it follows that $X_B(\mu)$ is an optimal solution to (C4). By the Lagrange rule applied to (C4), there exist $u_\mu^i (i = 1, ..., m)$ and v_μ such that

$$(X_B(\mu) - F(\mu))^{-1} - \sum_{i=1}^{m} u_\mu^i A_B^i - v_\mu C_B = 0.$$

By the Range property $C_B \in \mathcal{R}(\mathcal{A}_B^T)$ and hence for any $\mu > 0$:

$$(X_B(\mu) - F(\mu))^{-1} \in \mathcal{R}(\mathcal{A}_B^T).$$

5.2 Shifted analytic center of the optimal set

We recall that

$$\mathcal{P}^* = \left\{ X : \begin{array}{l} A_B^i \bullet X_B = b_i \ (\ i = 1, \ldots, m), \ X_B \in S_+^{|B|}, \\ \\ X_k = 0 \ (k \in \mathcal{I} - B) \end{array} \right\},$$

$\mathcal{P}^*$ is bounded. Let F be such that

$$X_B - F \in S_{++}^{|B|}$$

for some $X \in \mathcal{P}^*$. Consider the following problem

$$\max_{\substack{X \in \mathcal{P}^* \\ X_B - F \succ 0}} \ \ln\det(X_B - F) \tag{C5}$$

Note: (C5) has the unique solution

Definition: The unique solution to (C5) is called

the F- shifted analytic center of $\mathcal{P}^*$.

By using the Lagrange rule we obtain:

Lemma 4. *Let $X^F \in \mathcal{P}^*$, $X_B^F - F \succ 0$. Then the following claims are equivalent:*

(a) X^F is the F- shifted analytic center of $\mathcal{P}^*$,

(b) $\exists \quad u^i(i = 1, ..., m) \quad : \quad (X_B^F - F)^{-1} - \sum_{i=1}^m u^i A_B^i = 0.$

(c) $(X_B^F - F)^{-1} \in \mathcal{R}(\mathcal{A}_B^T)$.

5. 3 Convergence to the shifted analytic center

L3: $(X_B(\mu) - F(\mu))^{-1} \in \mathcal{R}(\mathcal{A}_B^T)$ where $X_B(\mu) - F(\mu) \succ 0$.

L4: $X^F \in \mathcal{P}^*$ is the F- shifted analytic center of $\mathcal{P}^*$ if

$$(X_B^F - F)^{-1} \in \mathcal{R}(\mathcal{A}_B^T) \quad \text{where} \quad X_B^F - F \succ 0.$$

To the limit approach we need the convergence of $F(\mu_t)$:

Note: $X_B(\mu) - F(\mu)$ is the Shur complement to $X_V(\mu)$ in

$$X(\mu) = \begin{bmatrix} X_B(\mu) & X_U(\mu) \\ X_U^T(\mu) & X_V(\mu) \end{bmatrix}.$$

From the perturbed complementarity and the formula for the inverze of block matrices [Horn,Johnson] we have

$$\frac{Z_B(\mu)}{\mu} = \left(X^{-1}(\mu)\right)_B = (X_B(\mu) - F(\mu))^{-1}.$$

Let $X(\mu_t) \to X^*$ for some $\{\mu_t\} \downarrow 0$. Applying Lemma 2 to the above equality we obtain.

(1) $F(\mu_t) \to F^*$ (if necessary, we take a subsequence)

(2) $X_B^* - F^* \succ 0$

(3) If $|B| + |N| = n$ ($\Leftrightarrow$ strict compl.)$\Rightarrow F^* = 0$

(4) If $|B| + |N| < n$, then $F^* \neq 0$ (or the corresponding dual $G^* \neq 0$).

Result 1: X^* is the F^*- shifted analytic center of $\mathcal{P}^*$.

Since $\mathcal{R}(\mathcal{A}_B^T)$ is a closed space, Result 1 follows from Lemmas 3 and 4.

Result 2: If $|B|+|N|=n$, then X^* is the analytic center of $\mathcal{P}^*$ which implies the convergence of the central path.

Result 3: If $|B|+|N|<n$, then the convergence of the central path (as a consequence of this approach) is not clear since $(X(\mu), F(\mu))$ may have more limit points.

Note: The convergence of the central path follows from the theory of algebraic sets.

Denote:

$X(0) = \lim_{\mu \downarrow 0} X(\mu)$

X^a-the analytic center of $\mathcal{P}^*$

X^F - the F-shifted analytic center of $\mathcal{P}^*$

Result 4: $X(0) = X^F$ for all limit points F of $F(\mu)$.

Two questions:

(1) Does $X(0)$ differ from X^a for some cases?

(2) For which cases?

5.4 Counter-example

Example: $m = 3$, $n = 4$, no strict complementarity and dual degeneracy.

$$\min x_{44} \quad : \begin{bmatrix} 1-x_{22} & x_{12} & x_{13} & x_{14} \\ x_{12} & x_{22} & -0.5x_{44} & -0.5x_{33} \\ x_{13} & -0.5x_{44} & x_{33} & 0 \\ x_{14} & -0.5x_{33} & 0 & x_{44} \end{bmatrix} \succeq 0$$

$$\mathcal{P}^* = \left\{ \begin{bmatrix} 1-x_{22} & x_{12} & 0 & 0 \\ x_{12} & x_{22} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \succeq 0 \right\} \qquad (\Rightarrow x_{22} \in [0,1])$$

$$\max y_1 \quad : \begin{bmatrix} -y_1 & 0 & 0 & 0 \\ 0 & -y_1 & -y_3 & -y_2 \\ 0 & -y_3 & -y_2 & -y_4 \\ 0 & -y_2 & -y_4 & 1-y_3 \end{bmatrix} \succeq 0$$

$$\mathcal{D}^* = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

The analytic center of $\mathcal{P}^*$ is where $x_{22} = 0.5$ and $x_{12} = 0$.

The central path converges to $x_{22} = 0.6$ and $x_{12} = 0$.

Due to the structure of feasible $Z \in \mathcal{D}$, the primal central path $X(\mu)$ has the following structure

$$X(\mu) = \begin{bmatrix} 1 - X_{22}(\mu) & 0 & 0 & 0 \\ 0 & X_{22}(\mu) & -\frac{1}{2}X_{44}(\mu) & -\frac{1}{2}X_{33}(\mu) \\ 0 & -\frac{1}{2}X_{44}(\mu) & X_{33}(\mu) & 0 \\ 0 & -\frac{1}{2}X_{33}(\mu) & 0 & X_{44}(\mu) \end{bmatrix}$$

It is well-known that the primal central path is the analytic center of a level set. The level set is given by the primal feasibility and a level condition which is $x_{44} = x_{44}(\mu)$ in our case. This implies that $X(\mu)$ is the optimal solution to the problem:

$$\max \det \begin{bmatrix} 1 - X_{22} & 0 & 0 & 0 \\ 0 & X_{22} & -\frac{1}{2}X_{44} & -\frac{1}{2}X_{33} \\ 0 & -\frac{1}{2}X_{44} & X_{33} & 0 \\ 0 & -\frac{1}{2}X_{33} & 0 & X_{44} \end{bmatrix}$$

s. t.

$$X_{44} = X_{44}(\mu),$$

$$X_{22} \in (0, 1), \quad X_{33} > 0,$$

$$X_{22}X_{33}X_{44} - X_{33}^3 - X_{44}^3 > 0.$$

Setting the gradient (with respect to X_{22} and X_{33}) of the determinant to zero, we obtain the two equations:

$$X_{33}(\mu)X_{44}(\mu) - 2X_{22}(\mu)X_{33}(\mu)X_{44}(\mu) + \frac{1}{4}X_{44}(\mu)^3 + \frac{1}{4}X_{33}(\mu)^3 = 0 \quad (1)$$

$$(1 - X_{22}(\mu))\left(X_{22}(\mu)X_{44}(\mu) - \frac{3}{4}X_{33}(\mu)^2\right) = 0. \quad (2)$$

Since $X_{22}(\mu) \in (0,1)$, we obtain from (2):

$$X_{33}(\mu) = \frac{2}{\sqrt{3}}\sqrt{X_{22}(\mu)X_{44}(\mu)}.$$

Substituting this expression in (1) and simplifying, we obtain:

$$\frac{2}{\sqrt{3}}\sqrt{X_{22}(\mu)} - \frac{10}{3\sqrt{3}}X_{22}(\mu)^{3/2} + \frac{1}{4}X_{44}(\mu)^{3/2} = 0.$$

In the limit where $\mu \to 0$, we have $X_{44}(\mu) \to 0$. Moreover, $X_{22}(\mu)$ is positive in the limit, since the limit point of the central path is maximally complementary. Denoting $\lim_{\mu\to 0} X_{22}(\mu) := X_{22}(0) > 0$, we have:

$$\frac{2}{\sqrt{3}}\sqrt{X_{22}(0)} - \frac{10}{3\sqrt{3}}X_{22}(0)^{3/2} = 0,$$

which implies $X_{22}(0) = 0.6$.