Section 7: A lower bound algorithm

Sylwester Zając

17.06.2014

As previously, we consider a block structure

$$\mathit{\Delta} = \left\{ \mathrm{diag}\left(\delta_1 I_{r_1}, \ldots, \delta_n I_{r_n}, \Delta_{S+1}, \ldots, \Delta_{S+F}\right) : \delta_j \in \mathbb{C}, \Delta_{S+j} \in \mathbb{C}^{m_j \times m_j} \right\},$$

where

$$r_1 + \ldots, +r_S + m_1 + \ldots + m_F = n.$$

Recall that for $M \in \mathbb{C}^{n \times n}$

$$\mu_{\mathit{\Delta}}(M) = \max_{\Delta \in \mathbf{B}_{\mathit{\Delta}}} \rho(\Delta M) = \max_{Q \in \mathbf{Q}} \rho(QM)$$

and

$$\begin{aligned} \mathbf{B}_{\mathit{\Delta}} &= \{\Delta \in \mathit{\Delta} : \bar{\sigma}(\Delta) \leq 1\}, \\ \mathbf{Q} &= \{Q \in \mathit{\Delta} : Q \text{ is unitary}\}, \\ \mathbf{D} &= \{\mathrm{diag}\left(D_1, \ldots, D_S, d_{S+1} I_{m_1}, \ldots, d_{S+F} I_{m_F}\right) : \\ &\qquad D_j \in \mathbb{C}^{r_j \times r_j}, D_j > 0, d_{S+j} \in \mathbb{R}_{>0}\}. \end{aligned}$$

Note that every $D \in \mathbf{D}$ and $\Delta \in \mathit{\Delta}$ commute.

Observation

Let $M \in \mathbb{C}^{n\times n}$, $\beta > 0$. *If there exist* $Q \in \mathbf{Q}$, $D \in \mathbf{D}$, $\xi \in \mathbb{C}^n$, $\|\xi\| = 1$ *such that*

$$QD^{\frac{1}{2}}MD^{-\frac{1}{2}}\xi = \beta\xi, \quad (1)$$

$$D^{-\frac{1}{2}}M^*D^{\frac{1}{2}}Q^*\xi = \beta\xi, \quad (2)$$

then

$$\beta \le \mu_\Delta(M).$$

Putting $\eta = \|D^{-\frac{1}{2}}\xi\|^{-1}D^{-\frac{1}{2}}\xi$ we obtain that (1) and (2) are equivalent to

$$\exists\, \eta \in \mathbb{C}^n, \|\eta\| = 1 : \; QM\eta = \beta\eta, \quad (QM)^*D\eta = \beta D\eta. \quad (3)$$

Proof of Observation: β is an eigenvalue of QM, so $\beta \le \rho(QM) \le \mu_\Delta(M)$ (actually, here we use only (1)).

Example

$S = 1,\ F = 0,\ r_1 = 2,\ \Delta = \{\mathrm{diag}\,(\delta, \delta) : \delta \in \mathbb{C}\}.$

$$M := \begin{bmatrix} 3 & -1 \\ 1 & 1 \end{bmatrix}.$$

We have $\det(qM - \lambda) = (2q - \lambda)^2$ and hence

$$\mu_\Delta(M) = \max_{|q|=1} \rho(QM) = 2,$$

but there is no $\beta > 0$ which satisfies (1) and (2) (although there is much β's satisfying (1) with some Q, D, ξ).

Example

$S = 1,\ F = 0,\ r_1 = 2,\ \Delta = \{\mathrm{diag}\,(\delta, \delta) : \delta \in \mathbb{C}\}$.

$$M := \begin{bmatrix} 3 & -1 \\ 1 & 1 \end{bmatrix}.$$

We have $\det(qM - \lambda) = (2q - \lambda)^2$ and hence

$$\mu_\Delta(M) = \max_{|q|=1} \rho(QM) = 2,$$

but there is no $\beta > 0$ which satisfies (1) and (2) (although there is much β's satisfying (1) with some Q, D, ξ).

Indeed, (1) and (2) hold iff (3) hold with some $Q \in \mathbf{Q}$, $D \in \mathbf{D}$, $\|\eta\| = 1$. This gives $Q = I_2$, $\beta = 2$ and

$$\eta \in \mathbb{R} \cdot (1, 1), \quad D\eta \in \mathbb{R} \cdot (1, -1).$$

Hence $0 = \langle D\eta, \eta \rangle > 0$, a contradiction.

Example

$S = 1,\ F = 0,\ r_1 = 3,\ \Delta = \{\mathrm{diag}\,(\delta, \delta, \delta) : \delta \in \mathbb{C}\}.$

$$M := \begin{bmatrix} 3 & -1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

We have $\det(qM - \lambda) = (2q - \lambda)^2(q - \lambda)$ and hence

$$\mu_\Delta(M) = \max_{|q|=1} \rho(QM) = 2,$$

but the only $\beta > 0$ which satisfies (1) and (2) is $\beta = 1$.

Let $x, y \in \mathbb{C}^n$. Write them compatibly to Δ:

$$\begin{aligned} x &= (x_1, \ldots, x_S, x_{S+1}, \ldots, x_{S+F}) \in \mathbb{C}^{r_1} \times \ldots \times \mathbb{C}^{r_S} \times \mathbb{C}^{m_1} \times \ldots \times \mathbb{C}^{m_F} = \mathbb{C}^n, \\ y &= (y_1, \ldots, y_S, y_{S+1}, \ldots, y_{S+F}) \in \mathbb{C}^{r_1} \times \ldots \times \mathbb{C}^{r_S} \times \mathbb{C}^{m_1} \times \ldots \times \mathbb{C}^{m_F} = \mathbb{C}^n. \end{aligned}$$

We say that the pair (x, y) is *non-degenerate* (with respect to Δ), if

$$\langle x_1, y_1 \rangle, \ldots, \langle x_S, y_S \rangle \neq 0 \text{ and } x_{S+1}, y_{S+1} \neq 0, \ldots, x_{S+F}, y_{S+F} \neq 0.$$

Theorem

Let $M \in \mathbb{C}^{n \times n}$. Assume that there exists $Q_0 \in \mathbf{Q}$ such that:

- $\rho(Q_0 M) = \max_{Q \in \mathbf{Q}} \rho(QM) > 0$,
- *$\rho(Q_0 M)$ is a distinct eigenvalue of $Q_0 M$,*
- *there exist a non-degenerate pair $(x, y) \in \mathbb{C}^n \times \mathbb{C}^n$ such that*

$$Q_0 M x = \rho(Q_0 M) x, \quad (Q_0 M)^* y = \rho(Q_0 M) y.$$

Then there exists $D \in \mathbf{D}$ and $\xi \in \mathbb{C}^n$, $\|\xi\| = 1$ such that

$$\begin{aligned} Q_0 D^{\frac{1}{2}} M D^{-\frac{1}{2}} \xi &= \mu_\Delta(M) \xi, \\ D^{-\frac{1}{2}} M^* D^{\frac{1}{2}} Q_0^* \xi &= \mu_\Delta(M) \xi. \end{aligned}$$

Lemma

Let $x, y \in \mathbb{C}^n$ be non-degenerate vectors. Then the following are equivalent:

1. *there exists $D \in \mathbf{D}$ such that $y = Dx$.*
2. *for every $G \in \Delta$ such that $G + G^* \leq 0$, $GG^* = G^*G$ there is $\mathrm{Re}\,\langle Gx, y\rangle \leq 0$.*

Proof of the lemma. (1) $\Rightarrow$ (2): D and G commute and $y = Dx$, so

$$\langle Gx, y\rangle = \langle Gx, D^{\frac{1}{2}}D^{\frac{1}{2}}x\rangle = \langle D^{\frac{1}{2}}Gx, D^{\frac{1}{2}}x\rangle = \langle GD^{\frac{1}{2}}x, D^{\frac{1}{2}}x\rangle,$$

$$\overline{\langle Gx, y\rangle} = \langle D^{\frac{1}{2}}x, GD^{\frac{1}{2}}x\rangle = \langle G^*D^{\frac{1}{2}}x, D^{\frac{1}{2}}x\rangle.$$

Hence

$$2\mathrm{Re}\,\langle Gx, y\rangle = \langle (G + G^*)D^{\frac{1}{2}}x, D^{\frac{1}{2}}x\rangle \leq 0.$$

(2) $\Rightarrow$ (1): It suffices to prove it for each block separately, so in fact we need to consider only two cases:

- $\Delta = \mathbb{C}^{n\times n}$,
- $\Delta = \{\mathrm{diag}\,(\delta, \ldots, \delta) : \delta \in \mathbb{C}\}$.

The case $\Delta = \mathbb{C}^{n\times n}$. Since

$$\mathbf{D} = \{\mathrm{diag}\,(d, \dots, d) : d \in \mathbb{R}_{>0}\},$$

we have to prove that $y = dx$ for some $d > 0$. We may assume that

$$\|x\| = \|y\| = 1.$$

There exists an unitary matrix $U \in \mathbb{C}^{n\times n}$ such that

$$Ux = y.$$

We have

$$U = P^* JP$$

for some unitary matrices $P, J \in \mathbb{C}^{n\times n}$, $J = \mathrm{diag}\,(\zeta_1, \dots, \zeta_n)$, $\zeta_j \in \mathbb{T}$. Write

$$Px = (a_1, \dots, a_n).$$

There is

$$Py = JPx = (\zeta_1 a_1, \dots, \zeta_n a_n).$$

It suffices to show that $a_j\zeta_j = a_j$ for every j.

Without loss of generality: $j = 1$.
Suppose that $a_1\zeta_1 \neq a_1$, i.e. $a_1 \neq 0$ and $\zeta_1 \neq 1$. There exists $\eta_1 \in \mathbb{T}$ such that

$$\mathrm{Re}\,\eta_1 \leq 0,\ \mathrm{Re}\,(\eta_1\bar{\zeta}_1) > 0.$$

Let

$$G := P^*\mathrm{diag}\,(\eta_1, 0, \ldots, 0)P$$

We have

$$\begin{aligned} 0 \geq \mathrm{Re}\,\langle Gx, y\rangle &= \langle \mathrm{diag}\,(\eta_1, 0, \ldots, 0)Px, Py\rangle \\ &= \langle(\eta_1 a_1, 0, \ldots, 0), (a_1\zeta_1, \ldots, a_n\zeta_n)\rangle \\ &= |a_1|^2\mathrm{Re}\,(\eta_1\bar{\zeta}_1) > 0, \end{aligned}$$

a contradiction.

The case $\Delta = \{\mathrm{diag}\,(\delta, \ldots, \delta) : \delta \in \mathbb{C}\}$. We have

$$\mathbf{D} = \{D \in \mathbb{C}^{n\times n} : D > 0\},$$

Existence of D such that $Dx = y$ follows from the fact that $\langle x, y\rangle > 0$.

Proof of the theorem. Step 1: If $Q_0 = I$, then there exists $D \in \mathbf{D}$ such that

$$y = Dx.$$

Proof of step 1: Let $G \in \Delta$, $G + G^* \leq 0$, $GG^* = G^*G$. It suffices to show that

$$\operatorname{Re} \langle Gx, y \rangle \leq 0.$$

Define

$$W(t) := e^{tG} M, \; t \geq 0.$$

$\rho(M)$ is a distinct eigenvalue of M, so $\langle x, y \rangle \neq 0$ and we may assume $\langle x, y \rangle = 1$.
There are $\epsilon > 0$ and $\mathbb{R}$-analytic maps $X, Y : (-\epsilon, \epsilon) \to \mathbb{C}^n$, $\lambda : (-\epsilon, \epsilon) \to \mathbb{C}$ s.t.

- $\lambda(0) = \rho(W(0)) = \rho(M)$, $X(0) = x$, $Y(0) = y$,
- $\lambda(t)$ is a distinct eigenvalue of $W(t)$,
- $|\lambda(t)| = \rho(W(t))$,
- $W(t)X(t) = \lambda(t)X(t)$, $W(t)^* Y(t) = \overline{\lambda(t)} Y(t)$,
- $\langle X(t), Y(t) \rangle = 1$.

Since $GG^* = G^*G$, we have

$$e^{tG} \in \mathbf{B}_\Delta, \ t \geq 0,$$

so the function $[0,\epsilon) \ni t \mapsto \mathrm{Re}\,\lambda(t)$ has a maximum at $t = 0$. There holds

$$\lambda'(0) = \langle W'(0)x, y\rangle.$$

Hence

$$0 \geq \mathrm{Re}\,\lambda'(0) = \mathrm{Re}\,\langle W'(0)x, y\rangle = \mathrm{Re}\,\langle GMx, y\rangle = \rho(M)\mathrm{Re}\,\langle Gx, y\rangle.$$

Step 2: We prove the conclusion.
By **step 1** for $\widetilde{M} := Q_0 M$ we have $y = Dx$ for some $D \in \mathbf{D}$. Put

$$\xi = D^{\frac{1}{2}}x = D^{-\frac{1}{2}}y.$$

As Q_0 and $D^{\frac{1}{2}}$ commute, we have

$$Q_0 D^{\frac{1}{2}} M D^{-\frac{1}{2}}\xi = D^{\frac{1}{2}} Q_0 M D^{-\frac{1}{2}} D^{\frac{1}{2}} x = \rho(Q_0 M) D^{\frac{1}{2}} x = \mu_\Delta(M)\xi$$

and similarly we obtain the condition (2).

From now, for simplicity: $S = F = 1$, $r + m = n$, i.e.

$$\begin{aligned} \varDelta &= \left\{ \operatorname{diag}\left(\delta_1 I_r, \Delta_2\right) : \delta_1 \in \mathbb{C}, \Delta_2 \in \mathbb{C}^{m\times m} \right\}, \\ \mathbf{D} &= \left\{ \operatorname{diag}\left(D_1, d_2 I_m\right) : D_1 \in \mathbb{C}^{r\times r}, D_1 > 0, d_2 \in \mathbb{R}_{>0} \right\}. \end{aligned}$$

Proposition

Let $M \in \mathbb{C}^{n\times n}$, $\beta > 0$. Then there exist $Q \in \mathbf{Q}$, $D \in \mathbf{D}$ and $\xi = (\xi_1, \xi_2) \in \mathbb{C}^n$, $\|\xi\| = 1$ such that

$$\begin{aligned} QD^{\frac{1}{2}} M D^{-\frac{1}{2}} \xi &= \beta \xi, \\ D^{-\frac{1}{2}} M^* D^{\frac{1}{2}} Q^* \xi &= \beta \xi, \end{aligned}$$

iff there exist $a = (a_1, a_2), w = (w_1, w_2), b, z \in \mathbb{C}^r \times \mathbb{C}^m$ such that (a, w) is non-degenerate and

$$\begin{aligned} \beta a = Mb, &\quad z = \left(\frac{\langle a_1, w_1 \rangle}{|\langle a_1, w_1 \rangle|} w_1, \frac{\|w_2\|}{\|a_2\|} a_2 \right), \\ \beta w = M^* z, &\quad b = \left(\frac{\langle w_1, a_1 \rangle}{|\langle w_1, a_1 \rangle|} a_1, \frac{\|a_2\|}{\|w_2\|} w_2 \right). \end{aligned}$$

Proof of the proposition. ($\Rightarrow$) We have

$$M\left(D^{-\frac{1}{2}}\xi\right) = \beta\left(D^{-\frac{1}{2}}Q^*\xi\right), \quad M^*\left(D^{\frac{1}{2}}Q^*\xi\right) = \beta\left(D^{\frac{1}{2}}\xi\right).$$

Set

$$b := D^{-\frac{1}{2}}\xi, \quad a := D^{-\frac{1}{2}}Q^*\xi, \quad z := D^{\frac{1}{2}}Q^*\xi, \quad w := D^{\frac{1}{2}}\xi.$$

($\Leftarrow$) Set

$$Q = \operatorname{diag}\left(\frac{\langle w_1, a_1\rangle}{|\langle w_1, a_1\rangle|}I_r, Q_2\right),$$

where $Q_2 \in \mathbb{C}^{m\times m}$ is unitary and such that $Q_2 z_2 = w_2$,

$$D = \operatorname{diag}\left(D_1, \frac{\|w_2\|}{\|a_2\|}I_m\right),$$

where $D_1 \in \mathbb{C}^{r\times r}$ is positive and such that $D_1 a_1 = z_1$ (it exists, because $\langle a_1, z_1\rangle > 0$), and

$$\xi := D^{\frac{1}{2}}b.$$

An iterative algorithm for a lower bound for $\mu_\Delta(M)$:

Take $a^1 = (a_1^1, a_2^1), w^1 = (w_1^1, w_2^1), b^1, z^1 \in (\mathbb{C}^r)_* \times (\mathbb{C}^m)_*$ such that

$$\langle a_1^1, w_1 \rangle \neq 0, \ \|a^1\| = \|b^1\| = \|z^1\| = \|w^1\| = 1.$$

Define (assuming that the following definitions are proper) $\widetilde{\beta}_k, \widehat{\beta}_k \in \mathbb{R}_{>0}$, $a^k = (a_1^k, a_2^k), w^k = (w_1^k, w_2^k), b^k, z^k \in (\mathbb{C}^r)_* \times (\mathbb{C}^m)_*$ as

$$\begin{aligned}
\widetilde{\beta}_{k+1} a^{k+1} &:= Mb^k \text{ with } \|a^{k+1}\| = 1 \\
z^{k+1} &:= \left(\frac{\langle a_1^{k+1}, w_1^k \rangle}{|\langle a_1^{k+1}, w_1^k \rangle|} w_1^k, \frac{\|w_2^k\|}{\|a_2^{k+1}\|} a_2^{k+1} \right), \\
\widehat{\beta}_{k+1} w^{k+1} &:= M^* z^{k+1} \text{ with } \|w^{k+1}\| = 1 \\
b^{k+1} &:= \left(\frac{\langle w_1^{k+1}, a_1^{k+1} \rangle}{|\langle w_1^{k+1}, a_1^{k+1} \rangle|} a_1^{k+1}, \frac{\|a_2^{k+1}\|}{\|w_2^{k+1}\|} w_2^{k+1} \right).
\end{aligned}$$

We have

$$\|a^k\| = \|b^k\| = \|w^k\| = \|z^k\| = 1.$$

If $\widehat{\beta}_k$, $\widetilde{\beta}_k$, a_k, b_k, w_k, z_k converge (respectively) to some $\widehat{\beta}, \widetilde{\beta} \in \mathbb{R}_{>0}$, $a = (a_1, a_2), w = (w_1, w_2), b, z \in \mathbb{C}^r \times \mathbb{C}^m$ with (a, w) being non-degenerate, then

$$\widehat{\beta} = \widetilde{\beta} \leq \mu_\Delta(M).$$

Proof: We have

$$\widetilde{\beta}a = Mb, \quad z = \left(\frac{\langle a_1, w_1 \rangle}{|\langle a_1, w_1 \rangle|} w_1, \frac{\|w_2\|}{\|a_2\|} a_2 \right),$$
$$\widehat{\beta}w = M^* z, \quad b = \left(\frac{\langle w_1, a_1 \rangle}{|\langle w_1, a_1 \rangle|} a_1, \frac{\|a_2\|}{\|w_2\|} w_2 \right).$$

Thus

$$\widetilde{\beta}\langle a, z \rangle = \langle Mb, z \rangle = \langle b, M^* z \rangle = \widehat{\beta}\langle b, w \rangle = \widehat{\beta}\langle a, z \rangle,$$

so $\widehat{\beta} = \widetilde{\beta}$. Now apply the previous proposition.

In practice:

- If the algorithm seems to converge, i.e. if $\widehat{\beta}_k \approx \widetilde{\beta}_k$, then $\widehat{\beta}_k$ and $\widetilde{\beta}_k$ approximate some lower bound of $\mu_\Delta(M)$.

In practice:

- If the algorithm seems to converge, i.e. if $\widehat{\beta}_k \approx \widetilde{\beta}_k$, then $\widehat{\beta}_k$ and $\widetilde{\beta}_k$ approximate some lower bound of $\mu_\Delta(M)$.
- In the opposite case we can just restart the algorithm with different initial points.

In practice:

- If the algorithm seems to converge, i.e. if $\widehat{\beta}_k \approx \widetilde{\beta}_k$, then $\widehat{\beta}_k$ and $\widetilde{\beta}_k$ approximate some lower bound of $\mu_\Delta(M)$.
- In the opposite case we can just restart the algorithm with different initial points.
- After some modifications of the algorithm it is possible to avoid the problems with definition of the sequences or with non-convergence.

Thank you for attention.