Generalized Eigenvalue Problems with Specified Eigenvalues

DANIEL KRESSNER†
SB–MATHICSE–ANCHP, EPF Lausanne,
Station 8, CH-1015 Lausanne, Switzerland

EMRE MENGİ‡
Department of Mathematics, Koç University,
Rumelifeneri Yolu, 34450 Sarıyer-Istanbul, Turkey

IVICA NAKIĆ§
Department of Mathematics, University of Zagreb,
Bijenička 30, 10000 Zagreb, Croatia

NINOSLAV TRUHAR¶
Department of Mathematics, University of Osijek,
Trg Ljudevita Gaja 6, HR-31 000 Osijek, Croatia

[Received on 14 May 2012]

We consider the distance from a (square or rectangular) matrix pencil to the nearest matrix pencil in 2-norm that has a set of specified eigenvalues. We derive a singular value optimization characterization for this problem and illustrate its usefulness for two applications. First, the characterization yields a singular value formula for determining the nearest pencil whose eigenvalues lie in a specified region in the complex plane. For instance, this enables the numerical computation of the nearest stable descriptor system in control theory. Second, the characterization partially solves the problem posed in [Boutry et al. 2005] regarding the distance from a general rectangular pencil to the nearest pencil with a complete set of eigenvalues. The involved singular value optimization problems are solved by means of BFGS and Lipschitz-based global optimization algorithms.

Keywords: Matrix pencils, eigenvalues, optimization of singular values, inverse eigenvalue problems, Lipschitz continuity, Sylvester equation.

1. Introduction

Consider a matrix pencil $A - \lambda B$ where $A, B \in \mathbb{C}^{n \times m}$ with $n \geq m$. Then a scalar $\rho \in \mathbb{C}$ is called an eigenvalue of the pencil if there exists a nonzero vector $v \in \mathbb{C}^m$ such that

$$(A - \rho B)v = 0. \quad (1.1)$$

The vector $v$ is said to be a (right) eigenvector associated with $\rho$ and the pair $(\rho, v)$ is said to be an eigenpair of the pencil.
In the square case $m = n$, the eigenvalues are simply given by the roots of the characteristic polynomial $\det(A - \lambda B)$ and there are usually $n$ eigenvalues, counting multiplicities. The situation is quite the opposite for $n > m$. Generically, a rectangular pencil $A - \lambda B$ has no eigenvalues at all. To see this, notice that a necessary condition for the satisfaction of (1.1) is that $n!/(n - m)!m!$ polynomials, each corresponding to the determinant of a pencil obtained by choosing $m$ rows of $A - \lambda B$ out of $n$ rows, must have a common root. Also, the generic Kronecker canonical form of a rectangular matrix pencil only consists of singular blocks (see Demmel & Edelman (1995)). Hence, (1.1) is an ill-posed problem and requires reformulation before admitting numerical treatment.

To motivate our reformulation of (1.1), we describe a typical situation giving rise to rectangular matrix pencils. Let $M \in \mathbb{C}^{n \times n}$ and suppose that the columns of $U \in \mathbb{C}^{n \times m}$ form an orthonormal basis for a subspace $\mathcal{W} \subset \mathbb{C}^n$ known to contain approximations to some eigenvectors of $M$. Then it is quite natural to consider the $n \times m$ matrix pencil

$$A - \lambda B := MU - \lambda U.$$  

(1.2)

The approximations contained in $\mathcal{W}$ and the approximate eigenpairs of $A - \lambda B$ are closely connected to each other. In one direction, suppose that $(\rho, x)$ with $x \in \mathcal{W}$ satisfies

$$(M + \Delta M - \rho I)x = 0$$  

(1.3)

for some (small) perturbation $\Delta M$. Then there is $v \in \mathbb{C}^n$ such that $x = Uv$. Moreover, we have

$$(A + \Delta A - \rho B)v = 0$$  

(1.4)

with $\Delta A := \Delta M \cdot U$ satisfying $\|\Delta A\|_2 \leq \|\Delta M\|_2$. In the other direction, the relation (1.4) with an arbitrary $\Delta A$ implies (1.3) with $\Delta M = \Delta A \cdot U^*$ satisfying $\|\Delta M\|_2 = \|\Delta A\|_2$. Unless $M$ is normal, the first part of this equivalence between approximate eigenpairs of $M$ and $A - \lambda B$ does not hold when the latter is replaced by the more common compression $U^*MU$. This observation has led to the use of rectangular matrix pencils in, e.g., large-scale pseudospectra computation (see Toh & Trefethen (1996)) and Ritz vector extraction (see Jia & Stewart (2001)).

This paper is concerned with determining the 2-norm distance from the pencil $A - \lambda B$ to the nearest pencil $(A + \Delta A) - \lambda B$ with a subset of specified eigenvalues. To be precise, let $\mathcal{S} = \{\lambda_1, \ldots, \lambda_k\}$ be a set of distinct complex numbers and let $r$ be a positive integer. Let $m_j(A + \Delta A, B)$ denote the (possibly zero) algebraic multiplicity\(^1\) of $\lambda_j$ as an eigenvalue of $(A + \Delta A) - \lambda B$. Then we consider the distance

$$\tau_r(\mathcal{S}) := \inf \left\{ \|\Delta A\|_2 : \sum_{j=1}^k m_j(A + \Delta A, B) \geq r \right\}.  

(1.5)$$

We allow $B$ to be rank-deficient. However, we require that rank$(B) \geq r$. Otherwise, if rank$(B) < r$, the pencil $(A + \Delta A) - \lambda B$ has fewer than $r$ finite eigenvalues for all $\Delta A$ and consequently the distance $\tau_r(\mathcal{S})$ is ill-posed.

For $k = r = 1$, it is relatively easy to see that

$$\tau_1(\{\lambda_1\}) = \sigma_m(A - \lambda_1 B),$$

\(^1\)For a rectangular matrix pencil, the algebraic multiplicity of $\lambda_j$ is defined as the sum of the sizes of associated regular Jordan blocks in the Kronecker canonical form, see also Section 2. By definition, this number is zero if $\lambda_j$ is actually not an eigenvalue of the pencil.
where, here and in the following, $\sigma_k$ denotes the $k$th largest singular value of a matrix. (The particular form of this problem with $k = r = 1$, and when $A$ and $B$ are perturbed simultaneously, is also studied for instance in Byers & Nichols (1993).) One of the main contributions of this paper is a derivation of a similar singular value optimization characterization for general $k$ and $r$, which facilitates the computation of $\tau_r(\mathbb{S})$. Very little seems to be known in this direction. Existing results concern the square matrix case ($m = n$ and $B = I$); see the works by Malyshev (1999) for $k = 1$ and $r = 2$ as well as Lippert (2005) for $k = 2$ and $r = 3$, Ikramov & Nazari (2003) for $k = 1$ and $r = 3$, and Mengi (2011a) for $k = 1$ and arbitrary $r$. Some attempts have also been made by Lippert (2010) for arbitrary $k$ and $r$ and for the square matrix case, and by Papathanasiou & Psarrakos (2008) for $k = 1$ and $r = 2$ and for the square matrix polynomial case.

Another class of applications arises in (robust) control theory, where a number of tasks require the determination of a (minimal) perturbation that moves some or all eigenvalues into a certain region in the complex plane. With the region of interest denoted by $\Omega \subseteq \mathbb{C}$, the results in this paper are an important step towards rendering the numerical computation of the distance

$$
\tau_r(\Omega) := \inf \{ \| \Delta A \|_2 : (A + \Delta A) - \lambda B \text{ has } r \text{ finite eigenvalues in } \Omega \}
$$

feasible. Here and in the following, multiple eigenvalues are counted according to their algebraic multiplicities. For $r = 1$ and $\Omega$ equal to $\mathbb{C}^+$ (right-half complex plane), the quantity $\tau_1(\mathbb{C}^+)$ amounts to the distance to instability, also called stability radius. In Van Loan (1984), a singular value characterization of $\tau_1(\mathbb{C}^+)$ was provided, forming the basis of a number of algorithms for computing $\tau_1(\mathbb{C}^+)$, see, e.g., Boyd & Balakrishnan (1990); Byers (1988). In our more general setting, we can also address the converse question: Given an unstable matrix pencil $A - \lambda B$, determine the closest stable pencil. Notice that this problem is intrinsically harder than the distance to instability. For the distance to instability it suffices to perturb the system so that one of the eigenvalues is in the undesired region. On the other hand to make an unstable system stable one needs to perturb the system so that all eigenvalues lie in the region of stability.

An important special case, $\Omega = \mathbb{C}$ leads to

$$
\tau_r(\mathbb{C}) := \inf \{ \| \Delta A \|_2 : (A + \Delta A) - \lambda B \text{ has } r \text{ finite eigenvalues } \}
$$

$$
= \inf_{\mathbb{S} \subseteq \mathbb{C}} \tau_r(\mathbb{S}).
$$

For $r = 1$ and particular choices of rectangular $A$ and $B$, the distance $\tau_1(\mathbb{C})$ corresponds to the distance to uncontrollability for a matrix pair (see Burke et al. (2005); Eising (1984)). For general $r$, a variant of this distance was suggested in Boutry et al. (2005) to solve an inverse signal processing problem approximately. More specifically, this problem is concerned with the identification of the shape of a region in the complex plane given the moments over the region. If the region is assumed to be a polygon, then its vertices can be posed as the eigenvalues of a rectangular pencil $A - \lambda B$, where $A$ and $B$ are not exact due to measurement errors, causing the pencil to have no eigenvalues (see Elad et al. (2004) for details). Then the authors attempt to locate nearby pencils with a complete set of eigenvalues. In this work we allow perturbations to $A$ only, but not to $B$. This restriction is only justified if the absolute value of $\lambda$ does not become too small. We consider our results and technique as significant steps towards the complete solution of the problem posed in Elad et al. (2004).

The outline of this paper is as follows. In the next section, we review the Kronecker canonical form for the pencil $A - \lambda B$. In §3, we derive a rank characterization for the condition $\sum_{j=1}^k m_j(A,B) \geq r$. 
This is a crucial prerequisite for deriving the singular value characterizations of $\tau_r(S)$ in §4. We discuss several corollaries of the singular value characterizations for $\tau_r(S)$, in particular for $\tau_r(\Omega)$ and $\tau_r(C)$, in §5. The singular value characterizations are deduced under certain mild multiplicity and linear independence assumptions. Although we expect these assumptions to be satisfied for examples of practical interest, they may fail to hold as demonstrated by an academic example in §6. Interestingly, the singular value characterization remains true for this example despite the fact that our derivation no longer applies. Finally, a numerical approach to solving the involved singular value optimization problems is briefly outlined in §7 and applied to a number of settings in §8. The main point of the developed numerical method and the experiments is to demonstrate that the singular value characterizations facilitate the computation of $\tau_r(S)$, $\tau_r(\Omega)$ and $\tau_r(C)$. We do not claim that the method outlined here is as efficient as it could be, neither do we claim that it is reliable.

2. Kronecker Canonical Form

Given a matrix pencil $A - \lambda B \in \mathbb{C}^{n \times m}$, the Kronecker canonical form (KCF), see Gantmacher (1959), states the existence of invertible matrices $P \in \mathbb{C}^{n \times n}$ and $Q \in \mathbb{C}^{m \times m}$ such that the transformed pencil $P(A - \lambda B)Q$ is block diagonal with each diagonal block taking the form $J_p(\alpha) - \lambda I_p$ or $I_p - \lambda J_p(0)$ or $F_p - \lambda G_p$ or $F_p^T - \lambda G_p^T$, where

$$J_p(\alpha) = \begin{bmatrix} \alpha & 1 & \cdots & \alpha \\ & \ddots & \ddots & \ddots \\ & & \alpha & 1 \\ & & & \alpha \end{bmatrix}_{p \times p}, \quad F_p = \begin{bmatrix} 1 & 0 & \cdots & 1 \\ & \ddots & \ddots & \ddots \\ & & 1 & 0 \\ & & & 1 \end{bmatrix}_{p \times (p+1)}, \quad G_p = \begin{bmatrix} 0 & 1 & \cdots & 0 \\ & \ddots & \ddots & \ddots \\ & & 0 & 1 \end{bmatrix}_{p \times (p+1)}$$

for some $\alpha \in \mathbb{C}$. Regular blocks take the form $J_p(\alpha) - \lambda I_p$ or $I_p - \lambda J_p(0)$, with $p \geq 1$, corresponding to finite or infinite eigenvalues, respectively. The blocks $F_p - \lambda G_p$ and $F_p^T - \lambda G_p^T$ are called right and left singular blocks, respectively, with $p \geq 0$ corresponding to a so called Kronecker index.

In large parts of this paper, indeed until the main singular value optimization characterization, we will assume that $A - \lambda B$ has no right singular blocks $F_p - \lambda G_p$. Eventually, we will remove this assumption by treating the occurrence of such blocks separately in Section 4.3.

3. Rank Characterization for Pencils with Specified Eigenvalues

In this section we derive a rank characterization for the satisfaction of the condition

$$\sum_{j=1}^k m_j(A, B) \geq r,$$

where $m_j(A, B)$ denotes the algebraic multiplicity of the eigenvalue $\lambda_j$. The following classical result (Gantmacher, 1959, Theorem 1, p. 219) concerning the dimension of the solution space for a Sylvester equation will play a central role.

**Theorem 3.1** Let $F \in \mathbb{C}^{m \times m}$ and $G \in \mathbb{C}^{r \times r}$. Then the dimension of the solution space for the Sylvester equation

$$FX - XG = 0$$
only depends on the Jordan canonical forms of the matrices $F$ and $G$. Specifically, suppose that $\mu_1, \ldots, \mu_\ell$ are the common eigenvalues of $F$ and $G$. Let $c_{j,i}, \ldots, c_{j,i_j}$ and $p_{j,1}, \ldots, p_{j,\tilde{l}_j}$ denote the sizes of the Jordan blocks of $F$ and $G$ associated with the eigenvalue $\mu_j$, respectively. Then

$$\dim \{ X \in \mathbb{C}^{m \times r} : FX - XG = 0 \} = \sum_{j=1}^{\ell} \sum_{i=1}^{c_{j,i}} \sum_{q=1}^{p_{j,q}} \min(c_{j,i}, p_{j,q}).$$

For our purposes, we need to extend the result of Theorem 3.1 to a generalized Sylvester equation of the form

$$AX - BXC = 0,$$  \hspace{1cm} (3.2)

where $C$ is a matrix with the desired set of eigenvalues $S$ and with correct algebraic multiplicities. For this type of generalized Sylvester equation, the extension is straightforward.\footnote{Košir (1996) provides an extension of Theorem 3.1 to a more general setting.} To see this, let us partition the KCF

$$P(A - \lambda B)Q = \text{diag}(A_F - \lambda I, I - \lambda A_I, A_S - \lambda B_S),$$  \hspace{1cm} (3.3)

such that

- $A_F - \lambda I$ contains all regular blocks corresponding to finite eigenvalues;
- $I - \lambda A_I$ contains all regular blocks corresponding to infinite eigenvalues;
- $A_S - \lambda B_S$ contains all left singular blocks of the form $F_T^T - \lambda G_T^T$.

As explained in Section 2, we exclude the occurrence of right singular blocks for the moment. Note that the finite eigenvalues of $A - \lambda B$ are equal to the eigenvalues of $A_F$ with the same algebraic and geometric multiplicities.

Using (3.3), $X$ is a solution of the generalized Sylvester equation (3.2) if and only if

$$(PAQ)(Q^{-1}X) - (PBQ)(Q^{-1}X)C = 0 \iff \text{diag}(A_F, I, A_S)Y - \text{diag}(I, A_I, B_S)YC = 0$$

where $Y = Q^{-1}X$. Consequently, the dimension of the solution space for (3.2) is the sum of the solution space dimensions of the equations

$$A_FY_1 - Y_1 C = 0 \quad \text{and} \quad Y_2 - A_IY_2 C = 0 \quad \text{and} \quad A_SY_3 - B_SY_3 C = 0.$$ 

Results by Demmel & Edelman (1995) show that the last two equations only admit the trivial solutions $Y_2 = 0$ and $Y_3 = 0$. To summarize: the solution spaces of the generalized Sylvester equation (3.2) and the (standard) Sylvester equation

$$A_FX - XC = 0$$

have the same dimension. Applying Theorem 3.1 we therefore obtain the following result.

**Theorem 3.2** Let $A, B \in \mathbb{C}^{n \times m}$ with $n \geq m$ be such that the KCF of $A - \lambda B$ does not contain right singular blocks. Then the dimension of the solution space for the generalized Sylvester equation

$$AX - BXC = 0$$
only depends on the Kronecker canonical form of \( A - \lambda B \) and the Jordan canonical form of \( C \in \mathbb{C}^{r \times r} \).

Specifically suppose that \( \mu_1, \ldots, \mu_r \) are the common eigenvalues of \( A - \lambda B \) and \( C \). Let \( c_{j,1}, \ldots, c_{j,\ell_j} \) and \( p_{j,1}, \ldots, p_{j,\tilde{\ell}_j} \) denote the sizes of the Jordan blocks of \( A - \lambda B \) and \( C \) associated with the eigenvalue \( \mu_j \), respectively. Then

\[
\dim \{ X \in \mathbb{C}^{m \times r} : AX - BXC = 0 \} = \sum_{j=1}^{r} \sum_{q=1}^{\tilde{\ell}_j} \min(c_{j,q}, p_{j,q}).
\]

We now apply the result of Theorem 3.2 to the generalized Sylvester equation

\[
AX - BXC(\mu, \Gamma) = 0,
\]

where \( C(\mu, \Gamma) \) takes the form

\[
C(\mu, \Gamma) = \begin{bmatrix}
\mu_1 & -\gamma_1 & \cdots & -\gamma_{r-1} \\
0 & \mu_2 & \cdots & \vdots \\
& & \ddots & \vdots \\
0 & & & \mu_r
\end{bmatrix},
\]

with

\[
\mu = [ \mu_1 \quad \mu_2 \quad \cdots \quad \mu_r ]^T \in \mathbb{S}, \quad \Gamma = [ \gamma_1 \quad \gamma_2 \quad \cdots \quad \gamma_{r-1} ]^T \in \mathbb{C}^{(r-1)/2}.
\]

As explained in the introduction, the set \( \mathbb{S} = \{ \lambda_1, \ldots, \lambda_k \} \) contains the desired approximate eigenvalues. Suppose that \( \lambda_j \) occurs \( p_j \) times in \( \mu \). Furthermore, as in Theorem 3.2, denote the sizes of the Jordan blocks of \( A - \lambda B \) and \( C(\mu, \Gamma) \) associated with the scalar \( \lambda_j \) by \( c_{j,1}, \ldots, c_{j,\ell_j} \) and \( p_{j,1}, \ldots, p_{j,\tilde{\ell}_j} \), respectively. Note that \( p_j = \sum_{q=1}^{\tilde{\ell}_j} p_{j,q} \). In fact, for generic values of \( \Gamma \) the matrix \( C(\mu, \Gamma) \) has at most one Jordan block of size \( p_j \) associated with \( \lambda_j \) for \( j = 1, \ldots, k \), see Demmel & Edelman (1995). In the following, we denote this set of generic values for \( \Gamma \) by \( \mathcal{G}(\mu) \). By definition, this set depends on \( \mu \) but not on \( A - \lambda B \).

First, suppose that inequality (3.1) holds. If we choose \( \mu \) such that \( \sum_{j=1}^{k} p_j = r \) and \( p_j \leq m_j(A, B) = \sum_{i=1}^{\ell_j} c_{j,i} \), then Theorem 3.2 implies that the dimension of the solution space for the generalized Sylvester equation (3.4) is

\[
\sum_{j=1}^{k} \sum_{i=1}^{\ell_j} \min(c_{j,i}, p_{j,q}) \geq \sum_{j=1}^{k} \sum_{i=1}^{\ell_j} \min(c_{j,i}, p_j) \geq \sum_{j=1}^{k} \min(m_j(A, B), p_j) = \sum_{j=1}^{k} p_j = r.
\]

In other words, there exists a vector \( \mu \) with components from \( \mathbb{S} \) such that the dimension of the solution space of the Sylvester equation (3.4) is at least \( r \).

Now, on the contrary, suppose that inequality (3.1) does not hold. Then for generic values \( \Gamma \in \mathcal{G}(\mu) \), the solution space dimension of (3.4) is

\[
\sum_{j=1}^{k} \sum_{i=1}^{\ell_j} \min(c_{j,i}, p_j) \leq \sum_{j=1}^{k} \sum_{i=1}^{\ell_j} c_{j,i} = \sum_{j=1}^{k} m_j(A, B) < r.
\]

In other words, no matter how \( \mu \) is formed from \( \mathbb{S} \), the dimension is always less than \( r \) for \( \Gamma \in \mathcal{G}(\mu) \). This shows the following result.
THEOREM 3.3 Let $A, B \in \mathbb{C}^{n \times m}$ with $n \geq m$ be such that the KCF of $A - \lambda B$ does not contain right singular blocks. Consider a set $S = \{\lambda_1, \ldots, \lambda_k\}$ of distinct complex scalars, and a positive integer $r$. Then the following two statements are equivalent.

(1) $\sum_{j=1}^k m_j(A, B) \geq r$, where $m_j(A, B)$ is the algebraic multiplicity of $\lambda_j$ as an eigenvalue of $A - \lambda B$.

(2) There exists $\mu \in S^r$ such that

$$\dim \{X \in \mathbb{C}^{m \times r} : AX - BXC(\mu, \Gamma) = 0\} \geq r$$

for all $\Gamma \in \mathscr{G}(\mu)$, where $C(\mu, \Gamma)$ is defined as in (3.5).

To obtain a matrix formulation of Theorem 3.3, we use the Kronecker product $\otimes$ to vectorize the generalized Sylvester equation (3.4) and obtain

$$\left( (I \otimes (C^T(\mu, \Gamma) \otimes B)) \right) \text{vec}(X) = \mathscr{L}(\mu, \Gamma, A, B) \text{vec}(X) = 0,$$

with the lower block triangular matrix

$$\mathscr{L}(\mu, \Gamma, A, B) := \begin{bmatrix}
A - \mu_1 B & \gamma_2 B & \cdots & A - \mu_2 B \\
\gamma_1 B & A - \mu_3 B & \cdots & \gamma_3 B \\
\vdots & \ddots & \ddots & \ddots \\
\gamma_k B & \cdots & A - \mu_{r-1} B & A - \mu_r B
\end{bmatrix}. \quad (3.6)$$

The operator vec stacks the columns of a matrix into one long vector. Clearly, the solution space of the generalized Sylvester equation and the null space of $\mathscr{L}(\mu, \Gamma, A, B)$ have the same dimension. Consequently, Theorem 3.3 can be rephrased as follows.

COROLLARY 3.1 Under the assumptions of Theorem 3.3, the following two statements are equivalent.

(1) $\sum_{j=1}^k m_j(A, B) \geq r$.

(2) There exists $\mu \in S^r$ such that $\text{rank} \left( \mathscr{L}(\mu, \Gamma, A, B) \right) \leq mr - r$ for all $\Gamma \in \mathscr{G}(\mu)$.

4. A singular value characterization for the nearest pencil with specified eigenvalues

As before, let $S = \{\lambda_1, \ldots, \lambda_k\}$ be a set of distinct complex scalars and let $r$ be a positive integer. The purpose of this section is to derive a singular value optimization characterization for the distance $\tau_r(S)$ defined in (1.5). Our technique is highly inspired by the techniques in Mengi (2011a,b) and in fact the main result of this section generalizes the singular value optimization characterizations from these works. We start by applying the following elementary result (Golub & Van Loan, 1996, Theorem 2.5.3, p.72) to the rank characterization derived in the previous section.

LEMMA 4.1 Consider $C \in \mathbb{C}^{\ell \times q}$ and a positive integer $p < \min(\ell, q)$. Then

$$\inf \left\{ \|\Delta C\|_2 : \text{rank}(C + \Delta C) \leq p \right\} = \sigma_{p+1}(C).$$
Defining
\[ \mathcal{P}_r(\mu) := \inf \left\{ \| \Delta A \|_2 : \text{rank}(\mathcal{L}(\mu, \Gamma, A + \Delta A, B)) \leq mr - r \right\} \quad (4.1) \]
for some \( \Gamma \in \mathcal{G}(\mu) \), Corollary 3.1 implies
\[ \tau_s(S) = \inf_{\mu \in S'} \mathcal{P}_r(\mu), \]
independent of the choice of \( \Gamma \). By Lemma 4.1, it holds that
\[ \mathcal{P}_r(\mu) = \inf\{\| A \|_2 : \text{rank}(\mathcal{L}(\mu, \Gamma, A + \Delta A, B)) \leq mr - r \} \]
\[ \geq \sigma_{mr-r+1}(\mathcal{L}(\mu, \Gamma, A, B)), \]
using the fact that \( A \) enters \( \mathcal{L} \) linearly. Note that this inequality in general is not an equality due to the fact that the allowable perturbations to \( \mathcal{L}(\mu, \Gamma, A, B) \) in the definition of \( \mathcal{P}_r(\mu) \) are not arbitrary. On the other hand, the inequality holds for all \( \Gamma \in \mathcal{G}(\mu) \) and hence – by continuity of the singular value \( \sigma_{mr-r+1}(\cdot) \) with respect to \( \Gamma \) – we obtain the lower bound
\[ \mathcal{P}_r(\mu) \geq \sup_{\Gamma \in \mathcal{G}(\mu)} \sigma_{mr-r+1}(\mathcal{L}(\mu, \Gamma, A, B)) =: \kappa_r(\mu). \quad (4.2) \]
For \( m = n \), it can be shown that \( \sigma_{mr-r+1}(\mathcal{L}(\mu, \Gamma, A, B)) \) tends to zero as \( \| \Gamma \| := \sum |\gamma_{ij}|^2 \to \infty \) provided that \( \text{rank}(B) \geq r \); see Appendix A for details. From this fact and the continuity of singular values, it follows that the supremum is attained at some \( \Gamma_0 \) in the square case:
\[ \kappa_r(\mu) = \sigma_{mr-r+1}(\mathcal{L}(\mu, \Gamma_0, A, B)). \]
In the rectangular case, numerical experiments indicate that the supremum is still attained if \( \text{rank}(B) \geq r \), but a formal proof does not appear to be easy. Moreover, it is not clear whether the supremum is attained at a unique \( \Gamma_0 \) or not. However, as we will show in the subsequent two subsections, any local extremum of the singular value function is a global maximizer under mild assumptions. (To be precise, the satisfaction of the multiplicity and linear independence qualifications at a local extremum guarantees that the local extremum is a global maximizer; see Definitions 4.1 and 4.2 below for multiplicity and linear independence qualifications.)

Throughout the rest of this section we assume that the supremum is attained at some \( \Gamma_0 \) and that \( \Gamma_0 \in \mathcal{G}(\mu) \). The latter assumption will be removed later, in Section 4.3.

We will establish the reverse inequality \( \mathcal{P}_r(\mu) \leq \kappa_r(\mu) \) by constructing an optimal perturbation \( \Delta A_s \) such that
\[ (i) \quad \| \Delta A_s \|_2 = \kappa_r(\mu), \quad \text{and} \]
\[ (ii) \quad \text{rank}(\mathcal{L}(\mu, \Gamma_0, A + \Delta A_s, B)) \leq mr - r. \]
Let us consider the left and right singular vectors \( U \in \mathbb{C}^n \) and \( V \in \mathbb{C}^m \) satisfying the relations
\[ \mathcal{L}(\mu, \Gamma_0, A, B) V = \kappa_r(\mu) U, \quad U^* \mathcal{L}(\mu, \Gamma_0, A, B) = V^* \kappa_r(\mu), \quad \| U \|_2 = \| V \|_2 = 1. \quad (4.3) \]
The aim of the next two subsections is to show that the perturbation
\[ \Delta A_s := -\kappa_r(\mu) \mathcal{Y}^+ \quad (4.4) \]
with \( \mathcal{Y} \in \mathbb{C}^{n \times r} \) and \( \mathcal{Y}^+ \in \mathbb{C}^{m \times r} \) such that \( \text{vec}(\mathcal{Y}) = U \) and \( \text{vec}(\mathcal{Y}) = V \) satisfies properties (i) and (ii). Here, \( \mathcal{Y}^+ \) denotes the Moore-Penrose pseudoinverse of \( \mathcal{Y} \). The optimality of \( \Delta A_s \) will be established under the following additional assumptions.
\textbf{Definition 4.1 (Multiplicity Qualification)} We say that the multiplicity qualification holds at \((\mu, \Gamma)\) for the pencil \(A - \lambda B\) if the multiplicity of the singular value \(\sigma_{\Gamma - \lambda + 1}((\mathcal{L}(\mu, \Gamma, A, B))\) is one.

\textbf{Definition 4.2 (Linear Independence Qualification)} We say that the linear independence qualification holds at \((\mu, \Gamma)\) for the pencil \(A - \lambda B\) if there is a right singular vector \(V\) associated with \(\sigma_{\Gamma - \lambda + 1}((\mathcal{L}(\mu, \Gamma, A, B))\) such that \(V \in \mathbb{C}^{m \times r}\), with \(\text{vec}(V) = V\), has full column rank.

### 4.1 The 2-norm of the optimal perturbation

Throughout this section we assume that the multiplicity qualification holds at the optimal \((\mu, \Gamma)\) for the pencil \(A - \lambda B\). Moreover, we can restrict ourselves to the case \(\kappa_r(\mu) \neq 0\), as the optimal perturbation is trivially given by \(\Delta A = 0\) when \(\kappa_r(\mu) = 0\).

Let \(\mathcal{A}(\gamma)\) be a matrix-valued function depending analytically on a parameter \(\gamma \in \mathbb{R}\). If the multiplicity of \(\sigma_j(\mathcal{A}(\gamma))\) is one and \(\sigma_j(\mathcal{A}(\gamma)) \neq 0\), then \(\sigma_j(\mathcal{A}(\gamma))\) is analytic at \(\gamma = \gamma_s\), with the derivative

\[
\frac{\partial \sigma_j(\mathcal{A}(\gamma_s))}{\partial \gamma} = \text{Re} \left( u_j \frac{\partial \mathcal{A}(\gamma_s)}{\partial \gamma} v_j^* \right),
\]

where \(u_j\) and \(v_j\) denote a consistent pair of unit left and right singular vectors associated with \(\sigma_j(\mathcal{A}(\gamma_s))\), see, e.g., Bunse-Gerstner \textit{et al.} (1991); Malyshev (1999); Rellich (1936).

Let us now define

\[
f(\Gamma) := \sigma_{\Gamma - \lambda + 1}((\mathcal{L}(\mu, \Gamma, A, B)),
\]

where we view \(f\) as a mapping \(\mathbb{R}^{r(r-1)} \rightarrow \mathbb{R}\) by decomposing each complex parameter \(\gamma_{j\ell}\) contained in \(\Gamma\) into its real and imaginary parts \(\Re \gamma_{j\ell}\) and \(\Im \gamma_{j\ell}\). By (4.5), we have

\[
\frac{\partial f(\Gamma)}{\partial \Re \gamma_{j\ell}} = \text{Re}(U_j^*BV_{\ell}), \quad \frac{\partial f(\Gamma)}{\partial \Im \gamma_{j\ell}} = \text{Re}(iU_j^*BV_{\ell}) = -\text{Im}(U_j^*BV_{\ell}),
\]

where \(U_j \in \mathbb{C}^n\) and \(V_{\ell} \in \mathbb{C}^m\) denote the \(j\)th and \(\ell\)th block components of \(U\) and \(V\), respectively. Furthermore, the fact that \(\Gamma_s\) is a global maximizer of \(f\) implies that both derivatives are zero. Consequently we obtain the following result.

**Lemma 4.2** Suppose that the multiplicity qualification holds at \((\mu, \Gamma)\) for the pencil \(A - \lambda B\) and \(\kappa_r(\mu) \neq 0\). Then \(U_j^*BV_{\ell} = 0\) for all \(j = 2, \ldots, r\) and \(\ell = 1, \ldots, j - 1\).

Now by exploiting Lemma 4.2 we show \(\mathcal{U}^*\mathcal{U} = \mathcal{V}^*\mathcal{V}\). Geometrically this means that the angle between \(U_i\) and \(U_j\) is identical with the angle between \(V_i\) and \(V_j\).

**Lemma 4.3** Under the assumptions of Lemma 4.2 it holds that \(\mathcal{U}^*\mathcal{U} = \mathcal{V}^*\mathcal{V}\).

**Proof** Expressing the first two equalities in the singular value characterization (4.3) in matrix form yields the generalized Sylvester equations

\[
A\mathcal{V} - B\mathcal{V}C(\mu, \Gamma) = \kappa_r(\mu)\mathcal{U}
\]

and

\[
\mathcal{U}^*A - C(\mu, \Gamma)\mathcal{U}^*B = \kappa_r(\mu)\mathcal{V}^*.
\]

By multiplying the first equation with \(\mathcal{U}^*\) from the left-hand side, multiplying the second equation with \(\mathcal{V}^*\) from the right-hand side, and then subtracting the second equation from the first we obtain

\[
\kappa_r(\mu)(\mathcal{U}^*\mathcal{U} - \mathcal{V}^*\mathcal{V}) = C(\mu, \Gamma)\mathcal{U}^*B\mathcal{V} - \mathcal{U}^*B\mathcal{V}C(\mu, \Gamma).
\]

(4.6)
Lemma 4.2 implies that $\mathcal{W}^*B\mathcal{W}$ is upper triangular. Since $C(\mu, \Gamma_\ast)$ is also upper triangular, the right-hand side in (4.6) is strictly upper triangular. But the left-hand side in (4.6) is Hermitian, implying that the right-hand side is indeed zero, which – together with $\kappa_\nu(\mu) \neq 0$ – completes the proof. □

The result of Lemma 4.3 implies $\|\mathcal{W}\mathcal{Y}^+\|_2 = 1$. A formal proof of this implication can be found in (Malyshev, 1999, Lemma 2) and (Mengi, 2011a, Theorem 2.5). Indeed, the equality $\|\mathcal{W}\mathcal{Y}^+\|_2 = 1$ can be directly deduced from $\|\mathcal{W}\mathcal{Y}^+x\|_2 = \|\mathcal{Y}^+x\|_2$ for every $x$ (implied by Lemma 4.3), and $\|VV^+\|_2 = 1$ (since $VV^+$ is an orthogonal projector).

**Theorem 4.3** Suppose that the multiplicity qualification holds at $(\mu, \Gamma_\ast)$ for the pencil $A - \lambda B$. Then the perturbation $\Delta A_\ast$ defined in (4.4) satisfies $\|\Delta A_\ast\|_2 = \kappa_\nu(\mu)$.

4.2 **Satisfaction of the rank condition by the optimally perturbed pencil**

Now we assume that the linear independence qualification (Definition 4.2) holds at $(\mu, \Gamma_\ast)$ for the pencil $A - \lambda B$. In particular we assume we can choose a right singular “vector” $\text{vec}(\mathcal{Y})$ so that $\mathcal{Y}$ has full column rank. We will establish that

$$\text{rank} (\mathcal{L}(\mu, \Gamma_\ast, A + \Delta A_\ast, B)) \leq mr - r$$

(4.7)

for $\Delta A_\ast$ defined as in (4.4).

Writing the first part of the singular vector characterization (4.3) in matrix form leads to the generalized Sylvester equation

$$A\mathcal{Y} - B\mathcal{Y}C(\mu, \Gamma_\ast) = \kappa_\nu(\mu)\mathcal{W}.$$  

The fact that $\mathcal{Y}$ has full column rank implies $\mathcal{Y}^+\mathcal{Y} = I$ and hence

$$A\mathcal{Y} - B\mathcal{Y}C(\mu, \Gamma_\ast) = \kappa_\nu(\mu)\mathcal{W} \mathcal{Y}^+\mathcal{Y}.$$  

$$\implies (A - \kappa_\nu(\mu)\mathcal{W}\mathcal{Y}^+)\mathcal{Y} - B\mathcal{Y}C(\mu, \Gamma_\ast) = 0$$  

$$\implies (A + \Delta A_\ast)\mathcal{Y} - B\mathcal{Y}C(\mu, \Gamma_\ast) = 0.$$  

Let us consider $\mathcal{M} = \{D \in \mathbb{C}^{r \times r} : C(\mu, \Gamma_\ast)D - DC(\mu, \Gamma_\ast) = 0\}$, the subspace of all $r \times r$ matrices commuting with $C(\mu, \Gamma_\ast)$. By Theorem 3.1, $\mathcal{M}$ is a subspace of dimension at least $r$. Clearly for all $D \in \mathcal{M}$, we have

$$0 = (A + \Delta A_\ast)\mathcal{Y}D - B\mathcal{Y}C(\mu, \Gamma_\ast)D = (A + \Delta A_\ast)(\mathcal{Y}D) - B(\mathcal{Y}D)C(\mu, \Gamma_\ast).$$

In other words, $\{\mathcal{Y}D : D \in \mathcal{M} \}$ has dimension at least $r$ (using the fact that $\mathcal{Y}$ has full column rank) and represents a subspace of solutions to the generalized Sylvester equation

$$(A + \Delta A_\ast)X - BXC(\mu, \Gamma_\ast) = 0.$$  

Reinterpreting this result in terms of the matrix representation, the desired rank estimate (4.7) follows. This completes the derivation of $\mathcal{P}_\nu(\mu) \leq \kappa_\nu(\mu)$ under the stated multiplicity and linear independence assumptions.

4.3 **Main Result**

To summarize the discussion above, we have obtained the singular value characterization

$$\tau_\nu(\mathcal{S}) = \inf_{\mu \in \mathcal{S}} \sup_{\Gamma} \sigma_{mr - r+1} (\mathcal{L}(\mu, \Gamma, A, B)).$$  

(4.8)
Among our assumptions, we have

(i) the KCF of \( A - \lambda B \) has no right singular blocks and (ii) \( \Gamma_s \in \mathscr{G}(\mu) \). \hspace{1cm} (4.9)

In this section, we show that these two assumptions can be dropped. We still require that \( \text{rank}(B) \geq r \). As explained in the introduction, the distance problem becomes ill-posed otherwise.

(i) Suppose that the KCF of \( A - \lambda B \) contains a right singular block \( F_p - \lambda G_p \in \mathbb{R}^{p \times (p+1)} \) for some \( p > 0 \). By (Košir, 1996, Sec. 4), the generalized Sylvester equation \( F_p Y - G_p XC(\mu, \Gamma) = 0 \) has a solution space of dimension \( r \). This implies that also the solution space of \( AX - BXC(\mu, \Gamma) = 0 \) has dimension at least \( r \), and consequently \( \sigma_{mr-r+1}(\mathscr{L}(\mu, \Gamma, A, B)) \) is always zero. On the other hand, the presence of a right singular block implies that for any \( \varepsilon > 0 \) and \( \mu_1, \ldots, \mu_r \in \mathbb{C} \) with \( r \leq \text{rank}(B) \) there is a perturbation \( \Delta A \) such that \( ||\Delta A||_2 \leq \varepsilon \) and \( (A + \Delta A) - \lambda B \) has eigenvalues \( \mu_1, \ldots, \mu_r \), see De Terán & Kressner (2012). This shows \( \tau_r(S) = 0 \) and hence both sides of (4.8) are equal to zero.

In summary, we can replace the assumption (i) by the weaker assumption \( \text{rank}(B) \geq r \).

(ii) To address (ii), we first note that both \( \mathscr{P}_r(\mu) \) and \( \kappa_r(\mu) \), defined in (4.1) and (4.2), change continuously with respect to \( \mu \). Suppose that \( \mu \) has repeating elements, which allows for the possibility that \( \Gamma_s \notin \mathscr{G}(\mu) \). But for all \( \tilde{\mu} \) with distinct elements, we necessarily have \( \mathscr{G}^{*}(\tilde{\mu}) = \mathbb{C}^{(r-1)/2} \). Moreover, when \( \tilde{\mu} \) is sufficiently close to \( \mu \) then \( \mathscr{P}_r(\tilde{\mu}) = \kappa_r(\tilde{\mu}) \), provided that the multiplicity and linear independence assumptions hold at \( (\mu, \Gamma_s) \) (implying the satisfaction of these two assumptions for \( \tilde{\mu} \) also). Then the equality \( \mathscr{P}_r(\mu) = \kappa_r(\mu) \) follows from continuity. Consequently, the assumption (ii) in (4.9) is also not needed for the singular value characterization.

We conclude this section by stating the main result of this paper.

**Theorem 4.4 (Nearest Pencils with Specified Eigenvalues)** Let \( A - \lambda B \) be an \( n \times m \) pencil with \( n \geq m \), and let \( r \) be a positive integer such that \( r \leq \text{rank}(B) \) and let \( S = \{\lambda_1, \ldots, \lambda_k\} \) be a set of distinct complex scalars.

(i) Then

\[
\tau_r(S) = \inf_{\mu \in \mathbb{S}^r} \sup_{\Gamma_s} \sigma_{mr-r+1}(\mathscr{L}(\mu, \Gamma, A, B))
\]

holds, provided that the optimization problem on the right is attained at some \( (\mu_s, \Gamma_s) \) for which \( \Gamma_s \) is finite and the multiplicity as well as the linear independence qualifications hold.

(ii) A minimal perturbation \( \Delta A \) such that \( \sum_{j=1}^{k} m(A + \Delta A_{s}, B) \geq r \) is given by (4.4), with \( \mu \) replaced by \( \mu_s \).

5. Corollaries of Theorem 4.4

As discussed in the introduction one potential application of Theorem 4.4 is in control theory, to ensure that the eigenvalues lie in a particular region in the complex plane. Thus let \( \Omega \) be a subset of the complex plane. Then, provided that the assumptions of Theorem 4.4 hold, we have the following singular value
characterization for the distance to the nearest pencil with \( r \) eigenvalues in \( \Omega \):

\[
\tau_r(\Omega) := \inf_{S \subseteq \Omega} \tau_r(S) \\
= \inf_{S \subseteq \Omega} \inf_{\mu \in S} \sup_{\Gamma} \sigma_{m_r-\Gamma+1}(\mathcal{L}(\mu, \Gamma, A, B)) \\
= \inf_{\mu \in \Omega} \sup_{\Gamma} \sigma_{m_r-\Gamma+1}(\mathcal{L}(\mu, \Gamma, A, B)),
\]

(5.1)

where \( \Omega^r \) denotes the set of vectors of length \( r \) with all entries in \( \Omega \).

When the pencil \( A - \lambda B \) is rectangular, that is \( n > m \), the pencil has generically no eigenvalues. Then the distance to the nearest rectangular pencil with \( r \) eigenvalues is of interest. In this case, the singular value characterization takes the following form:

\[
\tau_r(C) = \inf_{\mu \in C^r} \sup_{\Gamma} \sigma_{m_r-\Gamma+1}(\mathcal{L}(\mu, \Gamma, A, B)).
\]

(5.2)

The optimal perturbations \( \Delta A^* \) such that the pencil \( (A + \Delta A^*) - \lambda B \) has eigenvalues (in \( C \) and \( \Omega \)) are given by (4.4), with \( \mu \) replaced by the minimizing \( \mu \) values in (5.2) and (5.1), respectively.

6. Multiplicity and linear independence qualifications

The results in this paper are proved under the assumptions of multiplicity and linear independence qualifications. This section provides an example for which the multiplicity and linear independence qualifications are not satisfied for the optimal value of \( \Gamma \). Note that this does not mean that these assumptions are necessary to prove the results from this paper. In fact, numerical experiments suggest that our results may hold even if these assumptions are not satisfied.

Consider the pencil

\[
\begin{bmatrix}
-1 & 0 & 0 \\
0 & 5 & 0 \\
0 & 0 & 2
\end{bmatrix}
- \lambda
\begin{bmatrix}
0 & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & 1
\end{bmatrix}.
\]

Let \( \mu = [5 \ 1]^T \), that is, the target eigenvalues are 5 and 1. Then it is easy to see that the optimal perturbation is given by

\[
\Delta A^* = \begin{bmatrix}
0 & 0 & 0 \\
0 & 0 & 0 \\
0 & 0 & -1
\end{bmatrix}.
\]

The singular values of the matrix \( \mathcal{L}(\mu, \gamma, A, B) \) are

\[
0, 1, \sqrt{16 + |\gamma|^2}, \sqrt{5 + \frac{1}{2} |\gamma|^2 \pm \frac{1}{2} \sqrt{|\gamma|^4 + 20|\gamma|^2 + 64}}
\]

where the multiplicity of the singular value 1 is two. Hence

\[
\sigma_5(\mathcal{L}(\mu, \gamma, A, B)) = \sqrt{5 + \frac{1}{2} |\gamma|^2 - \frac{1}{2} \sqrt{|\gamma|^4 + 20|\gamma|^2 + 64}}.
\]

Clearly the supremum is attained for \( \gamma = 0 \) and \( \sigma_5(\mathcal{L}(\mu, 0, A, B)) = 1 \). Hence the multiplicity condition at the optimal \( \gamma \) is violated. All three pairs of singular vectors corresponding to the singular value 1 at the optimal \( \gamma \) violate the linear independence condition, but one pair does lead to the optimal perturbation \( \Delta A^* \).
7. Computational issues

A numerical technique that can be used to compute \( \tau_r(\Omega) \) and \( \tau_r(C) \) based on the singular value characterizations was already described in Mengi (2011a,b). For completeness, we briefly recall this technique in the following. The distances of interest can be characterized as

\[
\tau_r(\Omega) = \inf_{\mu \in \Omega^r} g(\mu) \quad \text{and} \quad \tau_r(C) = \inf_{\mu \in C^r} g(\mu),
\]

where \( g : \mathbb{C}^r \to \mathbb{R} \) is defined by

\[
g(\mu) := \sup_{\Gamma \in C^{(r-1)/2}} \sigma_{mr-r+1}(L(\mu, \Gamma, A, B)).
\]

The inner maximization problems are solved by BFGS, even though \( \sigma_{mr-r+1}(\cdot) \) is not differentiable at multiple singular values. In practice this is not a major issue for BFGS as long as a proper line search (e.g., a line search respecting weak Wolfe conditions) is used, as the multiplicity of the \( r \)th smallest singular value is one generically with respect to \( \Gamma \) for any given \( \mu \); see the discussions in Lewis & Overton (2012). If the multiplicity and linear independence qualifications hold at a local maximizer \( \Gamma^* \), then \( \Gamma^* \) is in fact a global maximizer and hence \( g(\mu) \) is retrieved. If, on the other hand, BFGS converges to a point where one of these qualifications is violated, it needs to be restarted with a different initial guess. In practice we have almost always observed convergence to a global maximizer immediately, without the need for such a restart.

Although the function \( g(\mu) \) is in general non-convex, it is Lipschitz continuous:

\[
|g(\mu + \delta\mu) - g(\mu)| \leq \|\delta\mu\|_2 \cdot \|B\|_2.
\]

There are various Lipschitz-based global optimization algorithms in the literature stemming mainly from ideas due to Piyavskii and Shubert (see Piyavskii (1972); Shubert (1972)). The Piyavskii-Shubert algorithm is based on the idea of constructing a piecewise linear approximation lying beneath the Lipschitz function. We used DIRECT (see Jones et al. (1993)), a sophisticated variant of the Piyavskii-Shubert algorithm. DIRECT attempts to estimate the Lipschitz constant locally, which can possibly speed up convergence.

The main computational cost involved in the numerical optimization of singular values is the retrieval of the \( r \)th smallest singular value of \( L(\mu, \Gamma, A, B) \) at various values of \( \mu \) and \( \Gamma \). As we only experimented with small pencils, we used direct solvers for this purpose. For medium to large scale pencils, iterative algorithms such as the Lanczos method (see Golub & Van Loan (1996)) are more appropriate.

8. Numerical Experiments

Our algorithm is implemented in Fortran, calling routines from LAPACK for singular value computations, the limited memory BFGS routine written by J. Nocedal (discussed in Liu & Nocedal (1989)) for inner maximization problems, and an implementation of the DIRECT algorithm by Gablonsky (described in Gablonsky (2001)) for outer Lipschitz-based minimization. A mex interface provides convenient access via MATLAB.

The current implementation is not very reliable, which appears to be related to the numerical solution of the outer Lipschitz minimization problem, in particular the DIRECT algorithm and its termination criteria. We rarely obtain results that are less accurate than the prescribed accuracy. The multiplicity and linear independence qualifications usually hold in practice and don’t appear to affect the numerical accuracy. For the moment, the implementation is intended for small pencils (e.g., \( n, m < 100 \)).
8.1 Nearest Pencils with Multiple Eigenvalues

As a corollary of Theorem 4.4 it follows that for a square pencil $A - \lambda B$ the nearest pencil having $S = \{\mu\}$ as a multiple eigenvalue is given by

$$\tau_2(S) = \sup_{\gamma} \left( \begin{bmatrix} A - \mu B & 0 \\ \gamma B & A - \mu B \end{bmatrix} \right)$$

provided that the multiplicity and linear independence qualifications are satisfied at the optimal $(\mu, \gamma^*)$. Therefore, for the distance from $A - \lambda B$ to the nearest square pencil with a multiple eigenvalue the singular value characterization takes the form

$$\inf_{\mu \in \mathbb{C}} \sup_{\gamma} \sigma_{2n-1} \left( \begin{bmatrix} A - \mu B & 0 \\ \gamma B & A - \mu B \end{bmatrix} \right).$$

(8.1)

Specifically, we consider the pencil

$$A - \lambda B = \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix} - \lambda \begin{bmatrix} -1 & 2 & 3 \\ 2 & -1 & 2 \\ 4 & 2 & -1 \end{bmatrix}.$$  (8.2)

Solving the above singular value optimization problem results in a distance of 0.59299 to the nearest pencil with a multiple eigenvalue. By (4.4), a nearest pencil turns out to be

$$\begin{bmatrix} 1.91465 & -0.57896 & -1.21173 \\ -1.32160 & 1.93256 & -0.57897 \\ -0.72082 & -1.32160 & 1.91466 \end{bmatrix} - \lambda \begin{bmatrix} -1 & 2 & 3 \\ 2 & -1 & 2 \\ 4 & 2 & -1 \end{bmatrix},$$

with the double eigenvalue $\lambda^* = -0.85488$. The optimal maximizing $\gamma$ turns out to be zero, which means neither the multiplicity nor the linear independence qualifications hold. (This is the non-generic case; had we attempted to calculate the distance to the nearest pencil with $\mu$ as a multiple eigenvalue for a given $\mu$, optimal $\gamma$ appears to be non-zero for generic values of $\mu$.) Nevertheless, the singular value characterization (8.1) remains to be true for the distance as discussed next.

The $\epsilon$-pseudospectrum of $A - \lambda B$ (subject to perturbations in $A$ only) is the set $\Lambda_{\epsilon}(A,B)$ containing the eigenvalues of all pencils $(A + \Delta A) - \lambda B$ such that $\|\Delta A\|_2 \leq \epsilon$. Equivalently,

$$\Lambda_{\epsilon}(A,B) = \{ \lambda \in \mathbb{C} : \sigma_{\text{min}}(A - \lambda B) \leq \epsilon \}.$$

It is well known that the smallest $\epsilon$ such that two components of $\Lambda_{\epsilon}(A,B)$ coalesce equals the distance to the nearest pencil with multiple eigenvalues. (See Alam & Bora (2005) for the case $B = I$, but the result easily extends to arbitrary invertible $B$.) Figure 1 displays the pseudospectra of the pencil in (8.2) for various levels of $\epsilon$. Indeed, two components of the $\epsilon$-pseudospectrum coalesce for $\epsilon = 0.59299$, confirming our result.

8.2 Nearest Rectangular Pencils with at least Two Eigenvalues

As an example for a rectangular pencil, let us consider the $4 \times 3$ pencil

$$A - \lambda B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0.1 & 0 \\ 0 & 2 & 0.3 \\ 0 & 1 & 2 \end{bmatrix} - \lambda \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$
The KCF of this pencil contains a $4 \times 3$ singular block and therefore the pencil has no eigenvalues. However, if the entry $a_{22}$ is set to zero, the KCF of the resulting pencil contains a $2 \times 1$ singular block and a $2 \times 2$ regular block corresponding to finite eigenvalues. Hence, a perturbation with 2-norm 0.1 is sufficient to have two eigenvalues.

According to the corollaries in Section 5 the distance to the nearest $4 \times 3$ pencil with at least two eigenvalues has the characterization

$$
\tau_2(C) = \inf_{\mu \in \mathbb{C}} \sup_{\gamma} \sigma_{2m-1} \left( \begin{bmatrix} A - \mu_1 B & 0 \\ \gamma B & A - \mu_2 B \end{bmatrix} \right)
$$

for $m = 3$. Our implementation returns $\tau_2(C) = 0.03927$. The corresponding nearest pencil (4.4) is given by

$$
\begin{bmatrix}
0.99847 & -0.03697 & -0.01283 \\
0 & 0.08698 & 0.03689 \\
0 & 2.00172 & 0.30078 \\
0.00007 & 1.00095 & 2.00376
\end{bmatrix} - \lambda
\begin{bmatrix}
0 & 0 & 0 \\
1 & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & 1
\end{bmatrix}
$$

and has eigenvalues at $\mu_1 = 2.55144$ and $\mu_2 = 1.45405$. This result is confirmed by Figure 2, which illustrates the level sets of the function $g(\mu)$ defined in (8.3) over $\mathbb{R}^2$.

For this example the optimal $\gamma$ is 2.0086. The smallest three singular values of the matrix in (8.3) are 1.4832, 0.0393 and 0.0062 for these optimal values of $\mu$ and $\gamma$. The linear independence qualification also holds.
Fig. 2. Level sets over $\mathbb{R}^2$ of the function $g(\mu)$ defined in (8.3). The asterisk marks the numerically computed global minimizer of $g$, which corresponds to the eigenvalues of a nearest pencil with two eigenvalues.

8.3 Nearest Stable Pencils

As a last example, suppose that $Bx'(t) = Ax(t)$ with $A, B \in \mathbb{C}^{n \times n}$ is an unstable descriptor system. The distance to a nearest stable descriptor system is a special case of $\tau_n(\Omega)$, with $\Omega = \mathbb{C}^-$, the open left-half of the complex plane. A singular value characterization is given by

$$
\tau_n(\mathbb{C}^-) = \inf_{\lambda_j \in \mathbb{C}^-} \sup_{\gamma_{ik} \in \mathbb{C}} \sigma_{n^2-n+1}\left(\begin{array}{ccc}
A - \lambda_1 B & 0 & 0 \\
\gamma_{21} B & A - \lambda_2 B & 0 \\
\gamma_{n1} B & \gamma_{n2} B & A - \lambda_n B \\
\end{array}\right).
$$

Specifically, we consider a system with $B = I_2$ and

$$
A = \left[\begin{array}{cc}
0.6 - \frac{i}{2} & -0.2 + \frac{3}{2}i \\
-0.1 + \frac{3}{2}i & 0.5 + \frac{1}{2}i
\end{array}\right].
$$

(8.4)

Both eigenvalues $\lambda_1 = 0.7 - i$ and $\lambda_2 = 0.4 + i$ are in the right-half plane. Based on the singular value characterization, we have computed the distance to a nearest stable system $x'(t) = (A + \Delta A)x(t)$ as 0.6610. The corresponding perturbed matrix

$$
A + \Delta A_s = \left[\begin{array}{cc}
0.0681 - 0.3064i & -0.4629 + 1.2524i \\
0.2047 + 0.5858i & -0.1573 + 0.3064i
\end{array}\right]
$$
at a distance of 0.6610 has one eigenvalue $(\lambda_*)_1 = -0.0885 + 0.9547i$ in the left-half plane and the other $(\lambda_*)_2 = -0.9547i$ on the imaginary axis. The $\epsilon$-pseudospectrum of $A$ is depicted in Figure 3. For $\epsilon = 0.6610$, one component of the $\epsilon$-pseudospectrum crosses the imaginary axis, while the other component touches the imaginary axis.

9. Concluding Remarks

In this work a singular value characterization has been derived for the 2-norm of a smallest perturbation to a square or a rectangular pencil $A - \lambda B$ such that the perturbed pencil has a desired set of eigenvalues. The immediate corollaries of this main result are

(i) a singular value characterization for the 2-norm of the smallest perturbation so that the perturbed pencil has a specified number of its eigenvalues in a desired region in the complex plane, and

(ii) a singular value characterization for the 2-norm of the smallest perturbation to a rectangular pencil so that it has a specified number of eigenvalues.

Partly motivated by an application explained in the introduction, we allow perturbations to $A$ only. The extension of our results to the case of simultaneously perturbed $A$ and $B$ remains open.

The development of efficient and reliable computational techniques for the solution of the derived singular value optimization problems is still in progress. As of now the optimization problems can be
solved numerically only for small pencils with small number of desired eigenvalues. The main task that needs to be addressed from a computational point of view is a reliable and efficient implementation of the DIRECT algorithm for Lipschitz-based optimization. For large pencils it is necessary to develop Lipschitz-based algorithms converging asymptotically faster than the algorithms (such as the DIRECT algorithm) stemming from the Piyavskii-Shubert algorithm. The derivatives from Section 4.1 might constitute a first step in this direction.

Acknowledgments We are grateful to two anonymous referees for their valuable comments. The research of the second author is supported in part by the European Commission grant PIRG-GA-268355 and the TÜBİTAK (the scientific and technological research council of Turkey) carrier grant 109T660.

A. Proof that \( \sigma_{m-r+1}(\mathcal{L}(\mu, \Gamma, A, B)) \to 0 \) as \( \Gamma \to \infty \)

We prove that the \( r \) smallest singular values of \( \mathcal{L}(\mu, \Gamma, A, B) \) decay to zero as soon as at least one entry of \( \Gamma \) tends to infinity, provided that \( n = m \). In the rectangular case, \( n > m \), these singular values generally do not decay to zero.

We start by additionally assuming that \( A - \mu_i B \) are non–singular matrices for all \( i = 1, \ldots, r \). We will first prove the result under this assumption, and then we will drop it. Our approach is a generalization of the procedure from (Ikramov & Nazari, 2005, §5), which in turn is a generalization of (Malyshev, 1999, Lemma 2).

Under our assumptions the matrix \( \mathcal{L}(\mu, \Gamma, A, B) \) is non–singular, and one can explicitly calculate the inverse. It is easy to see that the matrix \( \mathcal{L}^{-1}(\mu, \Gamma, A, B) \) has the form

\[
\begin{pmatrix}
(A - \mu_1 B)^{-1} & 0 & \cdots & 0 \\
X_{21} & (A - \mu_2 B)^{-1} & \cdots & 0 \\
X_{31} & X_{32} & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
X_{r1} & X_{r2} & \cdots & (A - \mu_r B)^{-1}
\end{pmatrix}
\]

We will use the well–known relations

\[
\sigma_{m-r+1}(\mathcal{L}(\mu, \Gamma, A, B)) = \sigma_r(\mathcal{L}(\mu, \Gamma, A, B)^{-1}) \leq \sigma_r(X_{ij})^{-1}. \tag{A.1}
\]

We first compute the matrices \( X_{21}, \ldots, X_{r,r-1} \) which lie on the first sub–diagonal. By a straightforward computation we obtain

\[
X_{i+1,i} = -\gamma_{i+1,i}(A - \mu_{i+1} B)^{-1}B(A - \mu_i B)^{-1}.
\]

If \( \sigma_r((A - \mu_{i+1} B)^{-1}B(A - \mu_i B)^{-1}) > 0 \), then from (A.1) it follows that if any of \( |\gamma_{i+1,i}| \) tends to infinity, we obtain the desired result. But \( \sigma_r((A - \mu_{i+1} B)^{-1}B(A - \mu_i B)^{-1}) > 0 \) easily follows from the assumption \( \text{rank}(B) \geq r \).

If this is not the case, meaning \( \max_i |\gamma_{i+1,i}| \) is bounded, then we use the entries on the next sub–diagonal \( X_{i+2,i} \). Again by straightforward computation we obtain

\[
X_{i+2,i} = -\gamma_{i+2,i}(A - \mu_{i+2} B)^{-1}B(A - \mu_i B)^{-1} + \gamma_{i+2,i+1}\gamma_{i+1,i}(A - \mu_{i+2} B)^{-1}B(A - \mu_{i+1} B)^{-1}B(A - \mu_i B)^{-1}.
\]

Because again \( \text{rank}(B) \geq r \) implies \( \sigma_r((A - \mu_{i+2} B)^{-1}B(A - \mu_i B)^{-1}) > 0 \), it follows that if any of \( |\gamma_{i+2,i}| \)
we obtain the inequality

\[ \sigma \]

From the arguments above, it follows that there exists \( \gamma \)\(|A| < 1\) for all \( i = 1, \ldots, r \). Thus, we have the recursive formula

\[ X_{i+j,i} = -\gamma_{i+j,i}(A - \mu_i B)^{-1}B(A - \mu_i B)^{-1} - \sum_{k=1}^{j-1} \gamma_{i+j,i+k}(A - \mu_i B)^{-1}BX_{i+k,i}. \]

Applying the same procedure as above, we conclude the proof in this case.

To remove the assumption that the matrices \( A - \mu_i B \) are non–singular, we fix any \( \varepsilon > 0 \). Let us choose a matrix \( A_\varepsilon \) such that \( \|A_\varepsilon - A\| < \varepsilon \) and that the matrices \( A_\varepsilon - \mu_i B \) are non–singular for all \( i = 1, \ldots, r \). From the arguments above, if follows that there exists \( \gamma_\varepsilon > 0 \) such that \( \sigma_{nr-r+1}(\varphi(\mu; \mu_i; A_\varepsilon; B)) < \varepsilon \), when \( \|\Gamma\| > \gamma_\varepsilon \).

Since

\[ \sigma_{nr-r+1}(\varphi(\mu; \Gamma; A, B)) \leq \sigma_{nr-r+1}(\varphi(\mu; \Gamma; A_\varepsilon, B)) + \varepsilon, \]

we obtain the inequality \( \sigma_{nr-r+1}(\varphi(\mu; \Gamma; A, B)) < 2\varepsilon \), when \( \|\Gamma\| > \gamma_\varepsilon \).

**References**


