ÖREBRO UNIVERSITET
Institutionen för naturvetenskap och teknik

Perturbations of symmetric matrix polynomials and their linearization

Edgar Skönnegård

Örebro universitet
Institutionen för naturvetenskap och teknik
Självständigt arbete för kandidatexamen i matematik, 15 hp

Perturbations of symmetric matrix polynomials and their linearization

Edgar Skönnegård

Maj 2020

Handledare: Andrii Dmytryshyn
Examinator: Jens Fjelstad

Abstract

The canonical stucture information, i.e. the elementary divisors and minimal indices of a matrix polynomial, is sensitive to perturbations of the matrix coefficients of the polynomial, e.g., the eigenvalues may change or disappear. Passing to a strong linearization is a way to solve a number of problems for matrix polynomials, the linearization then has the same finite and infinite elementary divisors and the change in minimal indices is known. However, when the linearization is perturbed by a full perturbation the correspondence between the linearization and matrix polynomial is lost, hence we seek a method to restore a matrix polynomial that corresponds to perturbed linearization. Therefore we present a numerical method for computing the perturbation of a matrix polynomial from a full perturbation of its linearization. Our method is iterative and requires of solving a system of coupled Sylvester equations. We limit the method to symmetric matrix polynomials.

Contents

Chapter 1

Introduction

1.1 Matrix polynomials and their linearization

A system of many polynomials can be represented as a matrix of polynomials or more commonly as a matrix polynomial, where the coefficients of the polynomial are matrices. Let

$$P(\lambda) = \lambda^d A_d + \cdots + \lambda^1 A_1 + A_0, \quad A_i \in \mathbb{C}^{n\times m} \quad \text{for} \quad i = 0, \ldots, d \tag{1.1}$$

be such a matrix polynomial, it is a matrix-valued function defined on the complex numbers. The grade of a matrix polynomial is defined as the highest power of λ. The degree of a matrix polynomial is defined as the highest power of λ with a nonzero coefficient matrix, i.e. we can then say that this matrix polynomial is of degree d, if $A_d \neq 0$ then the grade and the degree coincide. Hence grade $\geq$ degree. The rank of a matrix is given by the amount of linearly independent columns, the rank of a matrix polynomial $P(\lambda)$ is equal to $P(\lambda)$ at any point $\lambda \in \mathbb{C}$ which is not a zero of $P(\lambda)$.

Polynomial eigenvalue problems [8], the simplest and one of the most important of the nonlinear eigenvalue problems for matrix-valued functions, are most commonly solved by passing to a linearization that has the same eigenvalues. The eigenvalue problem for the linearization is then solved with general pencil algorithms [10] like the QZ algorithm. The eigenvalues of a matrix A are the solutions λ_0 of

$$Au = \lambda_0 u, \quad u \neq 0,$$

where u is a vector and referred to as the eigenvector for eigenvalue λ_0. As for matrix polynomials, λ_0 is an eigenvalue of $P(\lambda)$ if there exists a nonzero vector u such that

$$P(\lambda_0)u = 0.$$

The complete eigenstructure, comprised of the elementary divisors and minimal indices of a matrix polynomial gives an understanding of properties and behaviours of the underlying physical systems. For any eigenvalue λ_0 the invariant polynomials $d_i(\lambda)$ of a polynomial P with rank r can each be uniquely factored as

$$d_i(\lambda) = (\lambda - \lambda_0)^{\alpha_i} p_i(\lambda) \quad \text{for} \quad i = 1, \dots, r, \quad \text{with} \quad \alpha_i \geq 0, \;\; p_i(\lambda_0) \neq 0. \tag{1.2}$$

The elementary divisors, [4, 8], of P, with rank r and eigenvalue λ_0 , are the collection of factors $(\lambda - \lambda_0)^{\alpha_i}$. α_i comes from the the partial multiplicity sequence $(\alpha_1, \alpha_2, .., \alpha_r)$, of λ_0, which includes repetitions and where $\alpha_i \neq 0$. The partial multiplicity sequence is the sequence of exponents for a given λ_0 that satisfies the condition $0 \leq \alpha_1 \leq \alpha_2 \leq \cdots \leq \alpha_r$. For finite eigenvalues the elementary divisors may be referred to as finite elementary divisors, and infinite elementary divisors for infinite eigenvalues. Let us also define minimal indices, [4, 5], of P. The following definition is derived from [5]. We define the left and right null-spaces (1.3), respectively, for an $n \times m$ matrix polynomial $P(\lambda)$ over the field $\mathbb{C}(\lambda)$, of rational functions.

$$\begin{aligned} \mathcal{N}_{left}(P(\lambda)) &:= \{y(\lambda)^T \in \mathbb{C}(\lambda)^{1\times n} : y(\lambda)^T P(\lambda) = 0^T\} \quad \text{and} \\ \mathcal{N}_{right}(P(\lambda)) &:= \{x(\lambda) \in \mathbb{C}(\lambda)^{m\times 1} : P(\lambda)x(\lambda) = 0\}. \end{aligned} \tag{1.3}$$

Every subspace $\mathcal{W}$ of $\mathbb{C}(\lambda)^n$ has bases consisting entirely of vector polynomials, polynomials where the coefficients are vectors. A basis of $\mathcal{W}$ consisting of vector polynomials whose sum of degrees is minimal among all bases of $\mathcal{W}$ consisting of vector polynomials is a minimal basis of $\mathcal{W}$. The minimal indices of $\mathcal{W}$ are the degrees of the vector polynomials in a minimal basis of $\mathcal{W}$. Let the sets $\{y_1(\lambda)^T, \dots, y_{n-r}(\lambda)^T\}$ and $\{x_1(\lambda), \dots, x_{m-r}(\lambda)\}$, ordered by degree, be minimal bases of $\mathcal{N}_{left}(P)$ and $\mathcal{N}_{right}(P)$, respectively. Let $\eta_k = deg(y_k)$ for $k = 1, \dots, n-r$ and $\mu_k = deg(x_k)$ for $k = 1, \dots, m-r$, where deg is short for degree. Then the scalars $0 \leq \eta_1 \leq \cdots \leq \eta_{m-r}$ and $0 \leq \mu_1 \leq \cdots \leq \mu_{n-r}$ are, respectively, the left and right minimal indices of P.

A linearization of a matrix polynomial of degree d replaces the matrix polynomial with a matrix pencil, i.e. a matrix polynomial of degree 1. The most commonly used linearization has been the Frobenius companion forms, see [1, 8], although they do not preserve any of the most important matrix polynomial structures, being Hermitian, symmetric, alternating or palindromic. A linearization that preserves the underlying structure of the original polynomial is called structure preserving.

In this thesis we will focus on symmetric matrix polynomials of grade and degree d defined as

$$P(\lambda) = \lambda^d A_d + \cdots + \lambda^1 A_1 + A_0, \quad A_i \in \mathbb{C}^{n\times n}, A_i = A_i^T \quad \text{for} \quad i = 0, \dots, d. \tag{1.4}$$

A matrix pencil $\mathcal{L}$ is called a linearization of a matrix if they have the same finite elementary divisors and if rev$\mathcal{L}$ is a linearization of rev$P(\lambda) := \lambda^d P(1/\lambda)$ then $\mathcal{L}$ is called a strong linearization. We use a structure preserving strong linearization $\mathcal{L}_{P(\lambda)}$ of (1.4) from [4, 1]:

$$\mathcal{L}_{P(\lambda)} = \lambda \begin{bmatrix} A_d & & & & & \\ & 0 & & & & \\ & & \ddots & I_n & & \\ & & I_n & A_3 & & \\ & & & & 0 & I_n \\ & & & & I_n & A_1 \end{bmatrix} + \begin{bmatrix} A_{d-1} & I_n & & & & \\ I_n & 0 & & & & \\ & & \ddots & & & \\ & & & A_2 & I_n & \\ & & & I_n & 0 & \\ & & & & & A_0 \end{bmatrix} \tag{1.5}$$

where I_n is the $n \times n$ identity matrix and all non-specified blocks are zero. We restrict the linearization (1.5) to matrix polynomials of odd degrees.

1.2 Perturbations

A system may encounter shifts which change the underlying physical structure and thus behaviour, we call such a shift a perturbation. Let $\mathcal{F}$ be a structured perturbation of (1.5), defined as

$$\mathcal{F} = \lambda \begin{bmatrix} F_d & & & & & \\ & 0 & & & & \\ & & \ddots & & & \\ & & & F_3 & & \\ & & & & 0 & \\ & & & & & F_1 \end{bmatrix} + \begin{bmatrix} F_{d-1} & & & & & \\ & 0 & & & & \\ & & \ddots & & & \\ & & & F_2 & & \\ & & & & 0 & \\ & & & & & F_0 \end{bmatrix} \tag{1.6}$$

where $F_i = F_i^T$, and

$$\mathcal{E} = \lambda \begin{bmatrix} E_{11} & E_{12} & E_{13} & \cdots & E_{1d} \\ E_{21} & E_{22} & E_{23} & \cdots & E_{2d} \\ E_{31} & E_{32} & E_{33} & \cdots & E_{3d} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ E_{d1} & E_{d2} & E_{d3} & \cdots & E_{dd} \end{bmatrix} + \begin{bmatrix} \widetilde{E}_{11} & \widetilde{E}_{12} & \widetilde{E}_{13} & \cdots & \widetilde{E}_{1d} \\ \widetilde{E}_{21} & \widetilde{E}_{22} & \widetilde{E}_{23} & \cdots & \widetilde{E}_{2d} \\ \widetilde{E}_{31} & \widetilde{E}_{32} & \widetilde{E}_{33} & \cdots & \widetilde{E}_{3d} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \widetilde{E}_{d1} & \widetilde{E}_{d2} & \widetilde{E}_{d3} & \cdots & \widetilde{E}_{dd} \end{bmatrix} \tag{1.7}$$

be a full perturbation of (1.5) such that $E_{ij} = E_{ji}^T$ and $\widetilde{E}_{ij} = \widetilde{E}_{ji}^T$ for $i, j = 1, \cdots, d$. When we refer to an unstructured perturbation we mean the part of a full perturbation that is not the structured part, i.e.

$$\lambda \begin{bmatrix} 0 & E_{12} & E_{13} & \cdots & E_{1d} \\ E_{21} & E_{22} & E_{23} & \cdots & E_{2d} \\ E_{31} & E_{32} & 0 & \cdots & E_{3d} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ E_{d1} & E_{d2} & E_{d3} & \cdots & 0 \end{bmatrix} + \begin{bmatrix} 0 & \widetilde{E}_{12} & \widetilde{E}_{13} & \cdots & \widetilde{E}_{1d} \\ \widetilde{E}_{21} & \widetilde{E}_{22} & \widetilde{E}_{23} & \cdots & \widetilde{E}_{2d} \\ \widetilde{E}_{31} & \widetilde{E}_{32} & 0 & \cdots & \widetilde{E}_{3d} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \widetilde{E}_{d1} & \widetilde{E}_{d2} & \widetilde{E}_{d3} & \cdots & 0 \end{bmatrix}.$$

A perturbation is usually small i.e. the entries are very small numbers and often follow a certain distribution. We denote an unstructured perturbation with u, ($\mathcal{E}^{\mathrm{u}}$), and a structured perturbation with s, ($\mathcal{E}^{\mathrm{s}}$).

For future reference please note that the Frobenius norm for every matrix $X = \left[x_{ij}\right]$ defined as $\|X\| = \|X\|_F := \left(\sum_{ij} |x_{ij}|^2\right)^{1/2}$ is used hereafter unless otherwise stated. We also define the norm of a $n \times m$ polynomial $P(\lambda)$ of grade d as

$$\|P(\lambda)\| := \left(\sum_{k=0}^{d} \|A_k\|^2\right)^{1/2}.$$

Chapter 2

Obtaining a structured perturbation

We begin with a full perturbation (1.7) of the linearization of a matrix polynomial, $\mathcal{L}_{P(\lambda)}$, the objective is to perform a number of iterations to transform this full perturbation into a structured perturbation (1.6) such that the change in the eigenstructure is known. How we achieve this is by solving a system of equations (2.2) and minimizing the unstructured part through every iteration. The iterations can then be comprised into a transformation matrix V, [4], that transforms the full perturbation into a structured perturbation. The purpose of the method is to derive a perturbation of the original matrix polynomial $P(\lambda)$. The transformation matrix V, derived from [4] and defined as $V = (I + X_1) \cdots (I + X_k)$ where X_i is a solution of a system of coupled Sylvester equations (3.1), such that

$$V^T(\mathcal{L}_{P(\lambda)} + \mathcal{E}_1)V = \mathcal{L}_{P(\lambda)} + \mathcal{E}^{\mathrm{s}}. \tag{2.1}$$

We acknowledge Theorem 5.2 in [2] where this result for the transformation matrix V is stated, specifically, for symmetric polynomials.

2.1 Auxiliary results

The basis of the method is the solution of the coupled Sylvester equation

$$\begin{aligned} AX + YA &= M, \\ BX + YB &= N \end{aligned} \tag{2.2}$$

where A, B, M, N are symmetric $n \times n$ matrices. In this section we go through how it is solved and also how it is used.

Definition 2.1.1. Let $\mathrm{vec}(X)$ be the vectorization of X defined as the n^2-long ordered stack of columns of X from left to right.

Example 2.1.1. We provide an example of Definition 2.1.1

$$\text{for} \quad X = \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix} \quad \text{then} \quad \operatorname{vec}(X) = \begin{bmatrix} x_{11} \\ x_{21} \\ x_{12} \\ x_{22} \end{bmatrix}.$$

Definition 2.1.2. LetA, B be $n \times m$ matrices and let $A \otimes B$ be the Kronecker product defined as

$$A \otimes B = \begin{bmatrix} a_{11} \cdot B & \cdots & a_{1m} \cdot B \\ a_{21} \cdot B & \cdots & a_{2m} \cdot B \\ \vdots & \ddots & \vdots \\ a_{n1} \cdot B & \cdots & a_{nm} \cdot B \end{bmatrix}.$$

Using definitions 2.1.1 and 2.1.2 we can rewrite the first equation in (2.2)

$$\begin{aligned} &\big(A \otimes I_n\big) \operatorname{vec}(Y) + \big(I_n \otimes A\big) \operatorname{vec}(X) = \\ &\qquad \begin{bmatrix} A \otimes I_n & I_n \otimes A \end{bmatrix} \begin{bmatrix} \operatorname{vec}(Y) \\ \operatorname{vec}(X) \end{bmatrix} = \operatorname{vec}(M). \end{aligned}$$

To show that this is allowed we evaluate the computation of the first element $m_{1,1}$ in M. Originally we had $AX + YA = M$ giving us the equation

$$m_{11} = A(1,:) \cdot X(:,1) + Y(1,:) \cdot A(:,1). \tag{2.3}$$

From the rewritten system

$$\begin{aligned} &\begin{bmatrix} A \otimes I_n & I_n \otimes A \end{bmatrix} \begin{bmatrix} \operatorname{vec}(Y) \\ \operatorname{vec}(X) \end{bmatrix} \\ &= \begin{bmatrix} a_{1,1} \cdot I_n & \cdots & a_{1,n} \cdot I_n & 1 \cdot A & 0 \cdot A & \cdots & 0 \cdot A \\ a_{2,1} \cdot I_n & \cdots & a_{2,n} \cdot I_n & 0 \cdot A & 1 \cdot A & \cdots & 0 \cdot A \\ \vdots & \ddots & \vdots & 0 \cdot A & \vdots & \ddots & \vdots \\ a_{n,1} \cdot I_n & \cdots & a_{n,n} \cdot I_n & 0 \cdot A & 0 \cdot A & \cdots & 1 \cdot A \end{bmatrix} \begin{bmatrix} \operatorname{vec}(Y) \\ \operatorname{vec}(X) \end{bmatrix}, \end{aligned}$$

we get the equation

$$m_{1,1} = \begin{bmatrix} a_{1,1} \\ 0 \\ \vdots \\ 0 \\ a_{1,2} \\ 0 \\ \vdots \\ 0 \\ \vdots \\ a_{1,n} \\ 0 \\ \vdots \\ 0 \\ a_{1,1} \\ \vdots \\ a_{1,n} \\ 0 \\ \vdots \\ 0 \end{bmatrix}^T \begin{bmatrix} y_{1,1} \\ \vdots \\ y_{n,1} \\ \vdots \\ y_{1,n-1} \\ \vdots \\ y_{n,n} \\ x_{1,1} \\ \vdots \\ x_{n,1} \\ \vdots \\ x_{1,n-1} \\ \vdots \\ x_{n,n} \end{bmatrix}.$$

Which simplifies to

$$\begin{aligned} m_{1,1} &= a_{1,1}x_{1,1} + a_{1,2}x_{1,2} + \cdots + a_{1,n}x_{1,n} + a_{1,1}y_{1,1} + a_{1,2}y_{2,1} + \cdots + a_{1,n}y_{n,1} \\ &= A(1,:)Y(1,:) + A(1,:)X(:,1). \end{aligned} \tag{2.4}$$

Comparing the equations (2.3) and (2.4)

$$\begin{gathered} A(1,:) \cdot X(:,1) + Y(1,:) \cdot A(:,1), \\ A(1,:) \cdot Y(1,:) + A(1,:) \cdot X(:,1) \end{gathered}$$

Recalling that A is symmetric, i.e. $A(:,1) = A(1,:)$, we can say that these equations are equal.

Similarly, we can then rewrite the whole system (2.2) as

$$\begin{bmatrix} A \otimes I_n & I_n \otimes A \\ B \otimes I_n & I_n \otimes B \end{bmatrix} \begin{bmatrix} \mathrm{vec}(Y) \\ \mathrm{vec}(X) \end{bmatrix} = \begin{bmatrix} \mathrm{vec}(M) \\ \mathrm{vec}(N) \end{bmatrix}.$$

we add the condition $Y = X^T$ with the equation

$$\begin{bmatrix} I_{n^2} & -P \end{bmatrix} \begin{bmatrix} \mathrm{vec}(Y) \\ \mathrm{vec}(X) \end{bmatrix} = 0$$

where P is the permutation matrix consisting of ones and zeroes such that $\mathrm{vec}(X^T) = P \cdot \mathrm{vec}(X)$. Now we have rewritten the system of coupled Sylvester equations as a system of linear equations, $Tx = b$. Where

$$T = \begin{bmatrix} A \otimes I_n & I_n \otimes A \\ B \otimes I_n & I_n \otimes B \\ I_{n^2} & -P \end{bmatrix}, \quad x = \begin{bmatrix} \mathrm{vec}(Y) \\ \mathrm{vec}(X) \end{bmatrix}, \quad b = \begin{bmatrix} \mathrm{vec}(M) \\ \mathrm{vec}(N) \end{bmatrix}$$

$$M = -\begin{bmatrix} E_{11} & E_{12} & E_{13} & \cdots & E_{1d} \\ E_{21} & E_{22} & E_{23} & \cdots & E_{2d} \\ E_{31} & E_{32} & E_{33} & \cdots & E_{3d} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ E_{d1} & E_{d2} & E_{d3} & \cdots & E_{dd} \end{bmatrix}, \qquad N = -\begin{bmatrix} \widetilde{E}_{11} & \widetilde{E}_{12} & \widetilde{E}_{13} & \cdots & \widetilde{E}_{1d} \\ \widetilde{E}_{21} & \widetilde{E}_{22} & \widetilde{E}_{23} & \cdots & \widetilde{E}_{2d} \\ \widetilde{E}_{31} & \widetilde{E}_{32} & \widetilde{E}_{33} & \cdots & \widetilde{E}_{3d} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \widetilde{E}_{d1} & \widetilde{E}_{d2} & \widetilde{E}_{d3} & \cdots & \widetilde{E}_{dd} \end{bmatrix},$$

$$A = \begin{bmatrix} A_d & & & & & \\ & 0 & & & & \\ & & \ddots & I_n & & \\ & & I_n & A_3 & & \\ & & & & 0 & I_n \\ & & & & I_n & A_1 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} A_{d-1} & I_n & & & & \\ I_n & 0 & & & & \\ & & \ddots & & & \\ & & & A_2 & I_n & \\ & & & I_n & 0 & \\ & & & & & A_0 \end{bmatrix}.$$

To compute a solution from this system that will minimize the unstructured perturbation we seek the property

$$(Y(\mathcal{L}_{P(\lambda)} + \mathcal{E}_1) + (\mathcal{L}_{P(\lambda)} + \mathcal{E}_1)X)^{\mathrm{u}} = -\mathcal{E}_1^{\mathrm{u}}, \tag{2.5}$$

which in turn gives us the property (2.7). What we specifically do not want is $(Y(\mathcal{L}_{P(\lambda)} + \mathcal{E}_1) + (\mathcal{L}_{P(\lambda)} + \mathcal{E}_1)X)^{\mathrm{s}} = -\mathcal{E}_1^{\mathrm{s}}$, as it would mean that we are also minimizing the structured perturbation which is not our goal. Now, to achieve the property that we do want we need to exclude some equations in $Tx = b$. We update the right hand side accordingly

$$b \to b = \begin{bmatrix} \mathrm{vec}(M^{\mathrm{u}}) \\ \mathrm{vec}(N^{\mathrm{u}}) \end{bmatrix} = -\mathcal{E}_1^{\mathrm{u}}.$$

The equations we want to exclude are the ones related to $E_{11}, E_{33}, \cdots, E_{dd}$ and $\widetilde{E}_{11}, \widetilde{E}_{33}, \cdots, \widetilde{E}_{dd}$ i.e. the parts removed from the right hand side. By removing the rows of T corresponding to the excluded elements in the right hand side we obtain the property (2.5).
Then solving the updated system $Tx = b$ with the least squares method will give the solution

$$x = \begin{bmatrix} \mathrm{vec}(X_1^T) \\ \mathrm{vec}(X_1) \end{bmatrix}.$$

We use the solution x to update the perturbation like in (2.1),

$$\begin{aligned}\mathcal{L}_{P(\lambda)} + \mathcal{E}_2 &= (I + X_1^T)(\mathcal{L}_{P(\lambda)} + \mathcal{E}_1)(I + X_1) \\ &= (\mathcal{L}_{P(\lambda)} + \mathcal{E}_1) + X_1^T(\mathcal{L}_{P(\lambda)} + \mathcal{E}_1) + (\mathcal{L}_{P(\lambda)} + \mathcal{E}_1)X_1 \\ &+ X_1^T(\mathcal{L}_{P(\lambda)} + \mathcal{E}_1)X_1.\end{aligned} \tag{2.6}$$

Since $(X_1^T(\mathcal{L}_{P(\lambda)} + \mathcal{E}_1) + (\mathcal{L}_{P(\lambda)} + \mathcal{E}_1)X_1)^{\mathrm{u}} = -\mathcal{E}_1^{\mathrm{u}}$ we get that the new unstructured part of the perturbation is in the last term $X_1^T(\mathcal{L}_{P(\lambda)} + \mathcal{E}_1)X_1$, such that

$$\mathcal{E}_2^{\mathrm{u}} = (X_1^T(\mathcal{L}_{P(\lambda)} + \mathcal{E}_1)X_1)^{\mathrm{u}} \tag{2.7}$$

where $\|(X_1^T(\mathcal{L}_{P(\lambda)} + \mathcal{E}_1)X_1)^{\mathrm{u}}\|$ will be small, as shown in section 3.1. Specifically, it will be smaller than the previous unstructured perturbation

$$\|\mathcal{E}_2^{\mathrm{u}}\| < \|\mathcal{E}_1^{\mathrm{u}}\|.$$

We continue until $\|\mathcal{E}_k^{\mathrm{u}}\| <$ tol. Then

$$\begin{aligned}\mathcal{L}_{P(\lambda)} + \mathcal{E}_k &= (I + X_k^T)(\mathcal{L}_{P(\lambda)} + \mathcal{E}_{k-1})(I + X_k) = \cdots \\ &= (I + X_k^T)\cdots(I + X_1^T)(\mathcal{L}_{P(\lambda)} + \mathcal{E}_1)(I + X_1)\cdots(I + X_k) \\ &= V^T(\mathcal{L}_{P(\lambda)} + \mathcal{E}_1)V.\end{aligned}$$

2.2 Algorithm

In the previous section we went through one iteration of the method, now we present it explicitly in the form of an algorithm.

Algorithm 2.2.1. *Let $\mathcal{L}_{P(\lambda)}$ be a structure preserving linearization of a symmetric matrix polynomial $P(\lambda)$ and $\mathcal{E}_1$ be a symmetric full perturbation of $\mathcal{L}_{P(\lambda)}$.*

Input: Perturbed matrix pencil $\mathcal{L}_{P(\lambda)} + \mathcal{E}_1$ and the tolerance parameter tol;

Initiation: $V := I$, and $\mathcal{F} := -\mathcal{E}_1^{\mathrm{u}}$;

Computation: while $\|\mathcal{F}\| > tol$

- *solve the coupled Sylvester equations: $((\mathcal{L}_{P(\lambda)} + \mathcal{E}_i)X + X^T(\mathcal{L}_{P(\lambda)} + \mathcal{E}_i))^{\mathrm{u}} = \mathcal{F}$;*
- *update the perturbation of the linearization: $(I + X^T)(\mathcal{L}_{P(\lambda)} + \mathcal{E}_i)(I + X)$;*
- *update the transformation matrix: $V_{i+1} := V_i(I + X)$;*
- *extract the new unstructured perturbation to be eliminated: $\mathcal{F} := -\mathcal{E}_{i+1}^{\mathrm{u}}$;*

Output: Structurally perturbed linearization pencil $\mathcal{L}_{P(\lambda)+E(\lambda)} = \mathcal{L}_{P(\lambda)} + \mathcal{E}_k$, where $\mathcal{E}_k$ is a structured perturbation (since the norm of $\|\mathcal{E}_k^{\mathrm{u}}\| <$tol), and the transformation matrix V.

Chapter 3

Bounds for convergence

This chapter is dedicated to show that Algorithm 2.2.1 will converge. To show this we use bounds for the different parts of our algorithm, we show that each part converges and thus the algorithm converges.

3.1 Bound on the solution of the Sylvester equations

In this section we show that the solution of the Sylvester equation that we use to update the perturbation in Algorithm 2.2.1 is bounded.

Lemma 3.1.1. *Let A, B, M, N be symmetric $n \times n$ matrices and let X and Y be $n \times n$ matrices that are the smallest norm solution of the system of coupled Sylvester equations*

$$\begin{aligned} AX + YA &= M, \\ BX + YB &= N, \\ Y &= X^T. \end{aligned} \tag{3.1}$$

Then

$$\|X\| \leq \frac{k(T)}{\sqrt{2}\sqrt{(2n\|A\|^2 + 2n\|B\|^2 + 2n^2))}} \sqrt{(\|M\|^2 + \|N\|^2)} \tag{3.2}$$

where $T = \begin{bmatrix} A \otimes I_n & I_n \otimes A \\ B \otimes I_n & I_n \otimes B \\ I_{n^2} & -P \end{bmatrix}$ *is the Kronecker product matrix associated with the system* (3.1) *and* $k(T) := \|T\|\|T^\dagger\|$ *is the Frobenius condition number of* T. $T^\dagger$ *is the pseudoinverse of* T

Proof. Using the Kronecker product we can rewrite the system of coupled

Sylvester equations as a system of linear equations $Tx = b$ or explicitly

$$\begin{bmatrix} A \otimes I_n & I_n \otimes A \\ B \otimes I_n & I_n \otimes B \\ I_{n^2} & -P \end{bmatrix} \begin{bmatrix} \mathrm{vec}(Y) \\ \mathrm{vec}(X) \end{bmatrix} = \begin{bmatrix} \mathrm{vec}(M) \\ \mathrm{vec}(N) \end{bmatrix}. \tag{3.3}$$

The least-squares solution of the smallest norm of such a system can be written as $x = T^{\dagger}b$, implying $\|x\| \leq \|T^{\dagger}\|\|b\|$ or explicitly, and taking into account $\|x\| = \|X\|^2 + \|Y\|^2 = \|X\|^2 + \|X^T\|^2 = 2\|X\|^2$

$$\begin{aligned} 2\|X\|^2 &\leq \|T^{\dagger}\|^2(\|M\|^2 + \|N\|^2) = \frac{k(T)^2}{\|T\|^2}(\|M\|^2 + \|N\|^2) \\ &= \frac{k(T)^2}{n\|A\|^2 + n\|A\|^2 + n\|B\|^2 + n\|B\|^2 + \|I_{n^2}\|^2 + \|-P\|^2}(\|M\|^2 + \|N\|^2) \\ &= \frac{k(T)^2}{2n\|A\|^2 + 2n\|B\|^2 + n^2 + n^2}(\|M\|^2 + \|N\|^2). \end{aligned} \tag{3.4}$$

We obtain

$$\|X\| \leq \frac{k(T)}{\sqrt{2}\sqrt{(2n\|A\|^2 + 2n\|B\|^2 + 2n^2))}}\sqrt{(\|M\|^2 + \|N\|^2)}$$

□

Now we use the bound on X to show that the unstructured perturbation of the new perturbation, updated using solution X_i, tends to zero as the number of iterations grows.

Theorem 3.1.1. *Let $\mathcal{L}_{P(\lambda)} + \mathcal{E}_1$ be a perturbation of the linearization and let $\alpha\|\mathcal{E}_1\| < 1$, where $\alpha = \alpha(\mathcal{L}_{P(\lambda)} + \mathcal{E}_1)$ is defined in* (3.8)*, then Algorithm 2.2.1 converges to a structured perturbation of the linearization, i.e. $\|\mathcal{E}_i^u\| \to 0$ if $i \to \infty$.*

Proof. Using the norm of the unstructured part of a perturbation at the i-th step, of the algorithm, we prove a bound for the norm of the unstructured part of a perturbation at step $i+1$. Define $\mathcal{L}_{P(\lambda)} = \lambda W + \widetilde{W}$.

Following Algorithm 2.2.1 we solve the system of coupled Sylvester equations, at step i,

$$\begin{aligned} ((W + E_i)X_i + X_i^T(W + E_i))^{\mathrm{u}} &= -E_i^{\mathrm{u}}, \\ ((\widetilde{W} + \widetilde{E}_i)X_i + X_i^T(\widetilde{W} + \widetilde{E}_i))^{\mathrm{u}} &= -\widetilde{E}_i^{\mathrm{u}}. \end{aligned} \tag{3.5}$$

Using the solution X_i of the system (3.5) we update the perturbation by computing

$$\begin{aligned} W + E_{i+1} &= (I + X_i^T)(W + E_i)(I + X_i), \\ \widetilde{W} + \widetilde{E}_{i+1} &= (I + X_i^T)(\widetilde{W} + \widetilde{E}_i)(I + X_i). \end{aligned} \tag{3.6}$$

Now we find the unstructured perturbation of the updated perturbation using (3.5) and (3.6)

$$\begin{aligned} E_{i+1}^{\mathrm{u}} &= (W + E_{i+1})^{\mathrm{u}} \\ &= (W + E_i)^{\mathrm{u}} + ((W + E_i)X_i + X_i^T(W + E_i))^{\mathrm{u}} + (X_i^T(W + E_i)X_i)^{\mathrm{u}} \\ &= E_i^{\mathrm{u}} - E_i^{\mathrm{u}} + (X_i^T(W + E_i)X_i)^{\mathrm{u}} = (X_i^T(W + E_i)X_i)^{\mathrm{u}}, \text{ similarly} \end{aligned}$$

$$\widetilde{E}_{i+1}^{\mathrm{u}} = (X_i^T(\widetilde{W} + \widetilde{E}_i)X_i)^{\mathrm{u}}.$$

In general, E_{i+1}^{u} and $\widetilde{E}_{i+1}^{\mathrm{u}}$ are not zero matrices but we will show that they tend to zero (entry-wise) when $i \to \infty$. Using the bound (3.2) on the Frobenius norm of X we have

$$\|E_{i+1}^{\mathrm{u}}\| \le \|(X_i^T(W+E_i)X_i)^{\mathrm{u}}\| \le \|(X_i^T(W+E_i)X_i)\| \le \|X_i\|\cdot\|X_i^T\|\cdot\|W+E_i\|$$

$$= \|X_i\|^2 \cdot \|W + E_i\| \le \frac{k(T_i)^2\|W + E_i\|}{2(2n\|W + E_i\|^2 + 2n\|\widetilde{W} + \widetilde{E}_i\|^2 + 2n^2)}\|\mathcal{E}_i^{\mathrm{u}}\|^2,$$

similarly for $\|\widetilde{E}_{i+1}^{\mathrm{u}}\|$,

$$\|\widetilde{E}_{i+1}^{\mathrm{u}}\| \le \|(X_i^T(\widetilde{W}+\widetilde{E}_i)X_i)^{u}\| \le \|(X_i^T(\widetilde{W}+\widetilde{E}_i)X_i)\| \le \|X_i\|\cdot\|X_i^T\|\cdot\|\widetilde{W}+\widetilde{E}_i\|$$

$$= \|X_i\|^2 \cdot \|\widetilde{W} + \widetilde{E}_i\| \le \frac{k(T_i)^2\|W + \widetilde{E}_i\|}{2(2n\|W + E_i\|^2 + 2n\|\widetilde{W} + \widetilde{E}_i\|^2 + 2n^2)}\|\mathcal{E}_i^{\mathrm{u}}\|^2$$

Where

$$T_i = \begin{bmatrix} (W + E_i) \otimes I_n & I_n \otimes (W + E_i) \\ (\widetilde{W} + \widetilde{E}_i) \otimes I_n & I_n \otimes (\widetilde{W} + \widetilde{E}_i) \\ I_{n^2} & -P \end{bmatrix} \tag{3.7}$$

is the Kronecker product matrix associated with the system of coupled Sylvester equations (3.5).

Define α as follows

$$\alpha = \max\left\{ \frac{k(T_i)^2\|W+E_i\|}{\sqrt{2}\cdot(2n\|W+E_i\|^2+2n\|\widetilde{W}+\widetilde{E}_i\|^2+2n^2)},\ \frac{k(T_i)^2\|\widetilde{W}+\widetilde{E}_i\|}{\sqrt{2}\cdot(2n\|W+E_i\|^2+2n\|\widetilde{W}+\widetilde{E}_i\|^2+2n^2)} \right\}. \tag{3.8}$$

Using α we can write the bounds on the unstructured part of the perturbation for both of the matrices in the matrix pencil at step $i + 1$ as

$$\|E_{i+1}^{\mathrm{u}}\| \le \frac{\alpha}{\sqrt{2}}\|\mathcal{E}_i^{\mathrm{u}}\|^2 \quad \text{and} \quad \|\widetilde{E}_{i+1}^{\mathrm{u}}\| \le \frac{\alpha}{\sqrt{2}}\|\mathcal{E}_i^{\mathrm{u}}\|^2. \tag{3.9}$$

This results in the bound for the whole pencil

$$\|\mathcal{E}_{i+1}^{\mathrm{u}}\| = (\|E_{i+1}^{\mathrm{u}}\|^2 + \|\widetilde{E}_{i+1}^{\mathrm{u}}\|^2)^{1/2} \leq \alpha \|\mathcal{E}_i^{\mathrm{u}}\|^2 \tag{3.10}$$

Expanding (3.10) ρ steps we get

$$\begin{aligned} \|\mathcal{E}_{i+1}^{\mathrm{u}}\| &\leq \alpha \|\mathcal{E}_i^{\mathrm{u}}\|^2 \\ &\leq \alpha(\alpha\|\mathcal{E}_{i-1}^{\mathrm{u}}\|^2)^2 \leq \cdots \leq \alpha(\alpha(\cdots(\alpha\|\mathcal{E}_{i-\rho}^{\mathrm{u}}\|^2)^2\cdots)^2)^2 \\ &= \alpha^{2^{\rho-1}+1}\|\mathcal{E}_{i-\rho}^{\mathrm{u}}\|^{2^\rho} = \alpha^{2^\rho-1}\|\mathcal{E}_{i-\rho}^{\mathrm{u}}\|^{2^\rho}. \end{aligned} \tag{3.11}$$

Using (3.11), with $\rho = k-1$, we can write the norm of the unstructured perturbation at step k, explicitly, as

$$\|\mathcal{E}_k^{\mathrm{u}}\| \leq \alpha^{2^{k-1}-1}\|\mathcal{E}_1\|^{2^{k-1}}. \tag{3.12}$$

If $\alpha\|\mathcal{E}_1\| < 1$ then the norm of the unstructured part of the perturbation tends to zero as the iteration grows. □

3.2 Bound on the transformation matrix

In this section we show, by Theorem 3.2.1, that the transformation matrix V of Algorithm 2.2.1 as defined

$$V = \lim_{i\to\infty}(I_n + X_1)\cdots(I_n + X_i) \quad \text{and} \quad V^T = \lim_{i\to\infty}(I_n + X_i^T)\cdots(I_n + X_1^T)$$

converges to a nonsingular matrix.

Theorem 3.2.1. *$\mathcal{L}_{P(\lambda)} + \mathcal{E}_1$ be a perturbation of the linearization $\mathcal{L}_{P(\lambda)}$, and $\alpha \cdot \|\mathcal{E}_1\| < 1$, where $\alpha = \alpha\left(L_{P(\lambda)}, \mathcal{E}_1\right)$ is defined in* (3.8)*. Let X_i be a solution of* (3.5) *for corresponding index i and I_n be the $n \times n$ identity matrix. Then*

$$\lim_{i\to\infty}(I_n + X_i)\cdots(I_n + X_2)(I_n + X_1) \tag{3.13}$$

exist and is a nonsingular matrix.

Proof. By Theorem 4 in [9] the limit in (3.13) exist and is a nonsingular matrix if the sum

$$\|X_1\| + \|X_2\| + \|X_3\| + ... = \sum_{i=1}^{\infty}\|X_i\| \tag{3.14}$$

absolutely converges.
Using the bound (3.2) for a solution of coupled Sylvester equations and bounds (3.10) and (3.12) we have

$$\|X_{i+1}\|^2 = \frac{\alpha \|\mathcal{E}_i^{\mathrm{u}}\|^2}{\sqrt{2}\max\left\{\|W\| + \|E\|, \quad \|\widetilde{W}\| + \|\widetilde{E}\|\right\}} \leq \alpha^{2^{i-1}-1} \|\mathcal{E}_1\|^{2^{i-1}}. \tag{3.15}$$

Bound (3.15) allow us to conclude that equation (3.14) absolutely converges for $\alpha\mathcal{E}_1 < 1$. □

Corollary 3.2.1. *If Theorem 3.2.1 holds for X then it holds for X^T.*

With Theorem 3.2.1 and Corollary 3.2.1 we can state that V and V^T exist and are nonsingular.

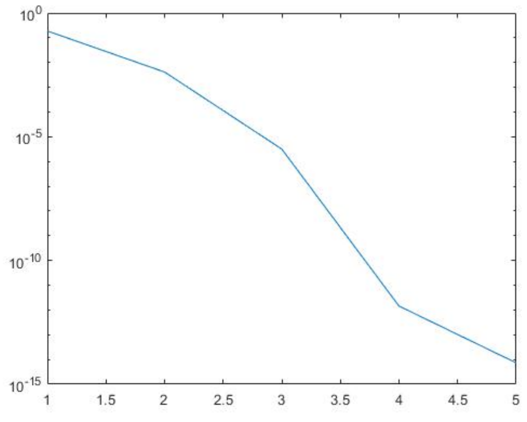
Chapter 4

Testing of the algorithm

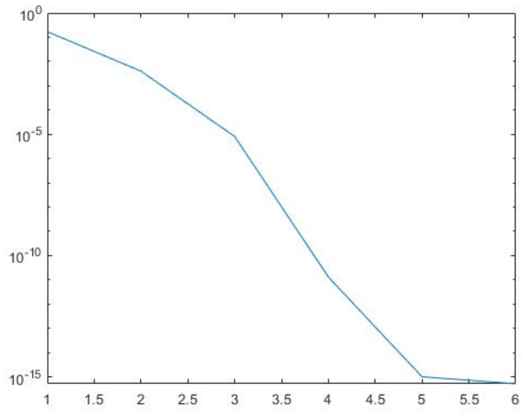
In this chapter we investigate the Algorithm 2.2.1 by performing a number of tests to see how well it converges. All of the examples are performed with randomly generated matrix polynomials.

4.1 Experiments

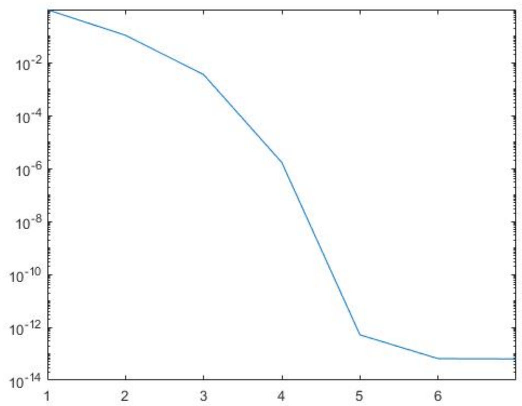
Example 4.1.1. Consider a random symmetric matrix polynomial of the size 5×5 and degree 5. The polynomial is normalized to have the Frobenius norm equal to 1 and is perturbed by adding a matrix polynomial (1.7), whose matrix coefficients have entries that are uniformly distributed numbers on the interval (0,0.1). The assigned tolerance is 10^{-14} and convergence occurs after six iterations as shown in the figure below.



Example 4.1.2. Consider a symmetric random polynomial as in Example 4.1.1 but of the size 8×8. Convergence occurs after six iterations.



Example 4.1.3. Consider a symmetric random polynomial as in Example 4.1.1 but of the size 10×10 and degree 7. After 6 iterations the norm of the unstructured part is $6.4061 \cdot 10^{-14}$.



Example 4.1.4. Consider a symmetric 3×3 cubic matrix polynomial, we scale the the matrix coefficients of this polynomial in an effort to observe the increase in the structured final perturbation and determine whether we still have convergence under assigned tolerance 10^{-15}. The norm of the initial perturbation is $\|\mathcal{E}_1\| = 0.8898$ and is the same for all scalings. We summarize the results in the table below.

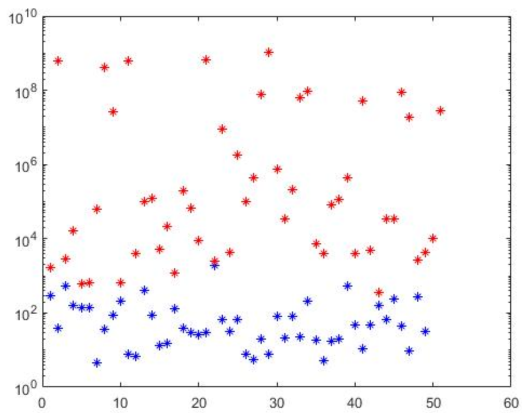
α_3	α_2	α_1	α_0	$\|\mathcal{E}^{\mathrm{s}}\|$	$\|\mathcal{E}^{\mathrm{s}}\|/\|\mathcal{E}_1\|$	$\|V\|$	conv.
$\frac{1}{\|Q(\lambda)\|}$	$\frac{1}{\|Q(\lambda)\|}$	$\frac{1}{\|Q(\lambda)\|}$	$\frac{1}{\|Q(\lambda)\|}$	0.1396	0.1569	1.055	yes
1	1	1	1	0.2026	0.2277	1.0808	yes
1	0.1	1	1	0.1911	0.2148	1.0732	yes
1	0.1	0.1	10	20.0435	22.5271	2.318	yes
1	0.1	0.1	10	9.72	10.9244	1.8913	yes
1	0.1	0.01	10	" - "	" - "	" - "	no
1	0.01	0.1	10	18,8964	21.2378	2.2591	yes
1	0.01	1	100	" - "	" - "	" - "	no
1	0.01	10	100	" - "	" - "	" - "	no

The tabel shows how the choice of scalars α_i, $i = 0, 1, 2, 3$ in the matrix polynomial $Q(\lambda) = \alpha_3 A_3 \lambda^3 + \alpha_2 A_2 \lambda^2 + \alpha_1 A_1 \lambda + \alpha_0 A_0$, changes the norm of the resulting structured perturbation $\mathcal{E}^{\mathrm{s}}$ and if the method converges. The initial perturbation $\mathcal{E}_1$ has entries equidistributed in (0,0.1).

Example 4.1.5. We consider 100 randomly generated 3×3 symmetric matrix polynomials of degree 3. As in the previous example we scale the polynomials with α_i presented in the table below, in an attempt to find the limit of the structured perturbation when the method no longer converges, assigned tolerance is 10^{-15}. The perturbations are randomly generated symmetric matrices with entries equidistributed in (0,0.1).

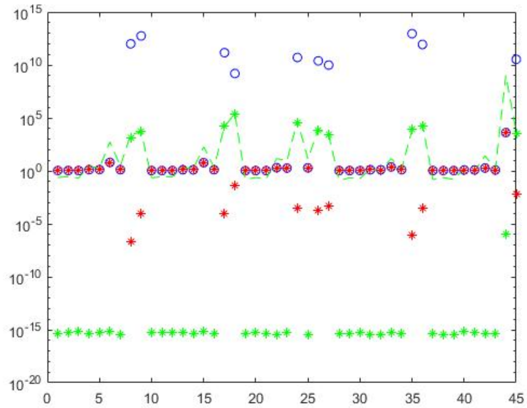
α_0	α_1	α_2	α_3
1	0.01	10	6

The result is presented in the figure below, where the red dots shows the norm of the perturbation of the structured part when the method does not converge and blue when convergence occurs.



We note that when the method no longer converges the norm of the structured perturbation appears to be about 10^3 or larger.

Example 4.1.6. In this example we investigate the transformation matrix. In the code we compute $V = (I+X_1)\cdots(I+X_k)$ and $V^T = (I+X_k^T)\cdots(I+X_1^T)$ from the solution of $Tx = b$, if the computed Y of the solution is not equal to X^T, or respectively for X, then the symmetry is broken and convergence will not occur. Consider 3 randomly generated symmetric 3×3 matrix polynomials of degree 3, we perturb each polynomial and scale them as in Example 4.1.4 with $\alpha_i, i = 0,1,2,3$. We compute the transformation matrices, V(red *) and V^T(blue o), and plot their norm along with the norm of the obtained unstructured perturbation("green *") and the structured perturbation ("green–") in the figure below. The perturbations are randomly generated symmetric matrices with entries equidistributed in (0,0.1).



We can see in the figure above that when the transformation matrix has norm much larger than 1 then we do not have convergence.

4.2 Conclusions

Throughout the performed testing we noted that the method worked good for well conditioned problems and for small entrywise perturbations. We also noted that when approximation errors and nonconvergent solutions of the least-square solution for $Tx = b$ occur then the lost symmetry in the matrix pencil makes the method diverge.

Bibliography

[1] E. Antoniou and S. Vologiannidis. A new family of companion forms of polynomial matrices. *Electron. J. Linear Algebra*, 11:78-87, 2004.

[2] F. De Terán, A. Dmytryshyn, F. M. Dopico. Generic symmetric matrix polynomials with bounded rank and fixed odd grade. https://arxiv.org/abs/1911.01408

[3] A. Dmytryshyn. Recovering a perturbation of a matrix polynomial from a perturbation of its linearization, *preprint*, 2020.

[4] A. Dmytryshyn. Structure preserving stratification of skew-symmetric matrix polynomials, *Linear Algebra Appl.*, 532:266-286, 2017.

[5] A. Dmytryshyn, S. Johansson, B. Kågström, and P. Van Dooren. Geometry of spaces for matrix polynomaial Fiedler linearizations. *Found. comput. Math.*, 20, 423-450, 2020.

[6] I. Gohberg, P. Lancaster, and L. Rodman, *Matrix polynomials*. SIAM Publications, 2009. Originally published: New York, Academic Press, 1982.

[7] A. Kaabi, An Arnoldi based algorithm for solving coupled Sylvester matrix equations, *J. Appl. Math. and Bioinformatics*, 2:151-163, 2012.

[8] D. S. Mackey, N. Mackey, and F. Tisseur. Polynomial Eigenvalue Problems: Theory, Computation, and Structure. In *Numerical Algebra, Matrix Theory, Differential-Algebraic Equations and Control Theory*, pages 319-348. Springer, 2015.

[9] W. F. Trench. Invertibly convergent infinite products of matrices. *J. Comput. Appl. Math.*, 101(1):255-263, 1999.

[10] P. Van Dooren and P. Dewilde. The eigenstructure of an arbitrary polynomial matrix: Computational aspects. *Linear Algebra Appl.*, 50:545-579, 1983.