Mathematica Bohemica

Fernando Marcos; Edgar Pereira

A fixed point method to compute solvents of matrix polynomials

Mathematica Bohemica, Vol. 135 (2010), No. 4, 355–362

Persistent URL: `http://dml.cz/dmlcz/140826`

135 (2010) MATHEMATICA BOHEMICA No. 4, 355–362

A FIXED POINT METHOD TO COMPUTE SOLVENTS OF MATRIX POLYNOMIALS

Fernando Marcos, Guarda, Edgar Pereira, Covilhã

(Received October 15, 2009)

Dedicated to José Vitória on the occasion of his 70th birthday

Abstract. Matrix polynomials play an important role in the theory of matrix differential equations. We develop a fixed point method to compute solutions of matrix polynomials equations, where the matricial elements of the matrix polynomial are considered separately as complex polynomials. Numerical examples illustrate the method presented.

Keywords: fixed point method, matrix polynomial, matrix differential equation

MSC 2010: 34M99, 65H10

1. Introduction

Let

$$P\Big(\frac{\mathrm{d}}{\mathrm{d}t}\Big)x = A_0x^{(m)}(t) + A_1x^{(m-1)}(t) + \ldots + A_mx(t) = 0 \tag{1.1}$$

be a homogeneous ordinary differential equation of order m, where

$$x^{(i)}(t) = \frac{\mathrm{d}^i x}{\mathrm{d}t^i} \tag{1.2}$$

and $A_0, A_1, \ldots, A_m$ are constant complex matrices of order n. This equation is linked to several applications [5], in particular to vibrating systems [10]. Usually, A_0 is supposed to be the identity matrix. This can also be achieved when A_0 is nonsingular; such case is called monic. Assuming $m = 2$ leads us to a very important second-order differential equation, which appears in many engineering applications, such as mechanical and electrical oscillation [16].

Associated to equation (1.1) is the matrix polynomial

$$P(X) = A_0 X^m + A_1 X^{m-1} + \ldots + A_m \tag{1.3}$$

of degree m in the unknown $n \times n$ matrix X.

A matrix S such that

$$P(S) = 0 \tag{1.4}$$

is a solvent of $P(X)$.

An important approach in searching numerical solutions to equation (1.1) is through the computation of solvents of the associated matrix polynomial ([12], p. 525). The first work we know in numerical analysis dealing with matrix polynomials is [2], which gave origin to [3] and [4], where an algebraic theory was developed and some algorithms were presented.

We cite from [3] two iterative methods to find solvents.

The first is a generalization of Traub's scalar polynomial algorithm and its purpose is the computation of a dominant solvent, that is, a solvent with the eigenvalues greater, in modulus, than the eigenvalues of any other solvent.

The second algorithm is a matrix version of Bernoulli's algorithm, which is essentially a block matrix power method applied to a block companion matrix of $P(X)$. Since [4], several works have been considering this method ([11], [17], [7], [13]). The classical Newton's method also had been generalized to matrix polynomials, first to the quadratic equation [1] and then to a general degree m [9]. Also, this method has been studied in the past years ([6], [14]).

All methods mentioned above are based on matrix arithmetics, that is, they solve the equation $P(X) = 0_n$ in $\mathbb{C}^{n \times n}$. Here we will develop a fixed point method considering the matrix elementwise, so the computations will be done at the scalar level; our attempt is to avoid the complications of matrix manipulations especially when dealing with the inverse of a matrix.

Next, in Section 2 we develop the theory of the fixed point method and in Section 3 we present some numerical experiments.

2. The method

The method is based on the construction of a function in which the existence of a fixed point implies a solvent of $P(X)$. This fixed point is computed by an iterative algorithm.

First we consider the elements of the matrix polynomial $P(X)$

$$p(x)_{ij} = p(x_{11}, x_{12}, \ldots, x_{ij}, \ldots, x_{nn})_{ij} \tag{2.1}$$

for $i, j = 1, 2, \ldots n$, where each of these elements is a multivariate complex polynomial.

Secondly we stack the elements of the matrix variable X in a vector; for this let

$$\mathrm{vec}(X) = (x_1^T x_2^T \ldots x_n^T)^T, \tag{2.2}$$

where $x_1, x_2, \ldots, x_n \in \mathbb{C}^n$ are the columns of X.

Then we define the function

$$\begin{aligned} f\colon\ \mathbb{C}^{n^2} &\longrightarrow \mathbb{C}^{n^2}, \\ x = \mathrm{vec}(X) &\mapsto f(x) \end{aligned} \tag{2.3}$$

where the coordinates $f_l(x)$, $l = 1, 2, \ldots, n^2$, $l = (j-1)n + i$, are obtained by solving the respective polynomials $p(x)_{ij}$ with respect to the leading term in x_{ij}, that is; the term in x_{ij} with the highest power. If $i = j$ the power of the leading term is m, which is the degree of the polynomial, otherwise it is

$$\begin{cases} \frac{1}{2}m & \text{for } m \text{ even}, \\ \frac{1}{2}(m+1) & \text{for } m \text{ odd}. \end{cases} \tag{2.4}$$

We explain the construction of the function f by considering $n = 2$ and $m = 2$, that is, the two-dimensional quadratic equation

$$P(X) = A_0 X^2 + A_1 X + A_2.$$

We write

$$X = \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix}, \quad A_k = \begin{bmatrix} a_{k_{11}} & a_{k_{12}} \\ a_{k_{21}} & a_{k_{22}} \end{bmatrix} \quad \text{for } k = 0, 1, 2,$$

and

$$P(X) = \begin{bmatrix} p(x)_{11} & p(x)_{12} \\ p(x)_{21} & p(x)_{22} \end{bmatrix}.$$

Thus, we consider

$$\begin{aligned} \mathrm{vec}(P(X)) &= \begin{bmatrix} p(x)_{11} \\ p(x)_{21} \\ p(x)_{12} \\ p(x)_{22} \end{bmatrix} \\ &= \begin{bmatrix} a_{0_{11}} x_{11}^2 + a_{0_{11}} x_{12} x_{21} + a_{0_{12}}(x_{11} x_{21} + x_{22} x_{21}) + a_{1_{11}} x_{11} + a_{1_{12}} x_{21} + a_{2_{11}} \\ a_{0_{21}} x_{11}^2 + a_{0_{21}} x_{12} x_{21} + a_{0_{22}}(x_{11} x_{21} + x_{22} x_{21}) + a_{1_{21}} x_{11} + a_{1_{22}} x_{21} + a_{2_{21}} \\ a_{0_{12}} x_{22}^2 + a_{0_{12}} x_{12} x_{21} + a_{0_{11}}(x_{11} x_{12} + x_{22} x_{12}) + a_{1_{11}} x_{12} + a_{1_{12}} x_{22} + a_{2_{12}} \\ a_{0_{22}} x_{22}^2 + a_{0_{22}} x_{12} x_{21} + a_{0_{21}}(x_{11} x_{12} + x_{22} x_{12}) + a_{1_{21}} x_{12} + a_{1_{22}} x_{22} + a_{2_{22}} \end{bmatrix}. \end{aligned}$$

Then, solving the polynomials $p(x)_{11}$, $p(x)_{21}$, $p(x)_{12}$, $p(x)_{22}$ with respect to the leading terms in x_{11}, x_{21}, x_{12}, x_{22}, respectively, we obtain

$$\begin{bmatrix} x_{11} \\ x_{21} \\ x_{12} \\ x_{22} \end{bmatrix} = \begin{bmatrix} \Big(-\dfrac{a_{0_{11}}x_{12}x_{21} + a_{0_{12}}(x_{11}x_{21} + x_{22}x_{21}) + a_{1_{11}}x_{11} + a_{1_{12}}x_{21} + a_{2_{11}}}{a_{0_{11}}}\Big)^{1/2} \\ -\dfrac{a_{0_{21}}x_{11}^2 + a_{1_{21}}x_{11} + a_{2_{21}}}{a_{0_{21}}x_{12} + a_{0_{22}}(x_{11} + x_{22}) + a_{1_{22}}} \\ -\dfrac{a_{0_{12}}x_{22}^2 + a_{1_{12}}x_{22} + a_{2_{12}}}{a_{0_{12}}x_{21} + a_{0_{11}}(x_{11} + x_{22}) + a_{1_{11}}} \\ \Big(-\dfrac{a_{0_{22}}x_{12}x_{21} + a_{0_{21}}(x_{11}x_{12} + x_{22}x_{12}) + a_{1_{21}}x_{12} + a_{1_{22}}x_{22} + a_{2_{22}}}{a_{0_{22}}}\Big)^{1/2} \end{bmatrix}$$

$$\stackrel{\text{def}}{=} \begin{bmatrix} f_1(x) \\ f_2(x) \\ f_3(x) \\ f_4(x) \end{bmatrix} = f(x).$$

As a direct consequence of the construction of the function f we have

Theorem 2.1. *If s is a fixed point of f, then $S = \operatorname{vec}^{-1}(s)$ is a solvent of $P(X)$.*

The algorithm for computing the fixed point is stated next.

Algorithm 2.1. *Let $x^{(0)} = \operatorname{vec}(X_0)$, where X_0 is the initial guest matrix, and define vectors $x^{(k)}$ by*

$$x^{(k+1)} = f(x^{(k)}) \tag{2.5}$$

for $k = 1, 2, \ldots$

The convergence of Algorithm 2.1 is based on the Schauder fixed point theorem (Theorem 2.2) and the respective asymptotic stability (Theorem 2.3), also known as the conjecture of Belitskii and Lyubich [15].

Theorem 2.2. *Let E be a Banach space, and let $K \subset E$ be a non-empty bounded convex open set.*

Then given any compact continuous mapping $g\colon \overline{K} \to \overline{K}$, there exists $\hat{x} \in \overline{K}$ such that $g(\hat{x}) = \hat{x}$.

Theorem 2.3. *Let E be a Banach space, let Ω be an open subset of E and let $g\colon \Omega \subseteq X \to X$ be compact and continuously Fréchet differentiable in Ω. Suppose $D \subset \Omega$ is a non-empty bounded convex open subset of E such that $g(\overline{D}) \subset \overline{D}$ and $\sup\limits_{x \in \overline{D}} \varrho(g'(x)) < 1$, where $\varrho(\cdot)$ is the spectral radius.*

Then there exists a unique $\hat{x} \in \overline{D}$ *such that* $g(\hat{x}) = \hat{x}$. *In addition, if* E *is a complex Banach space,* $\hat{x}$ *is globally asymptotically stable, i.e. the sequence* $g^{(k)}(x)$ *of iterates converges to* $\hat{x}$ *for any* $x \in \overline{D}$.

Furthermore, we will need the following theorem [8].

Theorem 2.4. *Let* A *be a bounded linear operator on a normed space* E.

Then for each $\delta > 0$ *there is a norm* $\|\cdot\|_\delta$ *on* E *equivalent to the original norm such that* $\|A\|_\delta \leqslant \varrho(A) + \delta$.

We present now conditions for the convergence of Algorithm 2.1.

Theorem 2.5. *Let* f *be the function defined as above. If there exists a non-empty bounded convex open set* U*, where* $f\colon U \subseteq \mathbb{C}^{n^2} \to \mathbb{C}^{n^2}$ *is continuously differentiable and*

i) *there exists an* $s \in U$ *such that* $s = f(s)$,

ii) *there exists a* $V_\varepsilon(s) = \{x \in \mathbb{C}^{n^2} \colon \|x - s\| < \varepsilon\}$, *with* $\overline{V_\varepsilon(s)} \subseteq U$ *such that* $\sup\limits_{x \in V_\varepsilon(s)} \varrho(f'(x)) < 1$.

Then for any $x^{(0)} \in V_\varepsilon(s)$ *the sequence* $x^{(k+1)} = f(x^{(k)})$, $k = 1, 2, \ldots$, *is in* $V_\varepsilon(s)$ *and converges to* s*, the unique solution in* $V_\varepsilon(s)$.

Proof. Let $r = \sup\limits_{x \in V_\varepsilon(s)} \varrho(f'(x))$ and set $\delta < 1 - r$, then by Theorem 2.4 for each x in $V_\varepsilon(s)$ there is $\|.\|_\delta$ such that $\|f'(x)\|_\delta \leqslant r + \delta < r + 1 - r = 1$.

Therefore, if we take an $\hat{x} \in \overline{V_\varepsilon(s)}$, then $\|f(\hat{x}) - s\|_\delta < r\|\hat{x} - s\|_\delta \leqslant \varepsilon$, and so $f(\hat{x}) \in \overline{V_\varepsilon(s)}$.

Thus, $f(\overline{V_\varepsilon(s)}) \subset \overline{V_\varepsilon(s)}$ and we can consider $f\colon \overline{V_\varepsilon(s)} \longrightarrow \overline{V_\varepsilon(s)}$, where $\overline{V_\varepsilon(s)}$ is a non-empty, compact and convex subset of the Banach space $\mathbb{C}^{n^2}$.

Then, by the Schauder fixed point theorem (Theorem 2.2), there exists an $\hat{x} \in \overline{V_\varepsilon(s)}$ such that $f(\hat{x}) = \hat{x}$.

Furthermore, this fixed point is unique by Theorem 2.3, hence $\hat{x} = s$. It also follows that for any $x^{(0)} \in V_\varepsilon(s)$, the sequence $x^{(k+1)} = f(x^{(k)})$ is in $V_\varepsilon(s)$ and

$$\|x^{(k)} - s\|_\delta < \|f(x^{(k-1)}) - f(s)\|_\delta < r\|x^{(k-1)} - s\|_\delta < r^k\|x^{(0)} - s\|_\delta < r^k\varepsilon \underset{k\to\infty}{\longrightarrow} 0.$$

Hence we have that $x^{(k)} \underset{k\to\infty}{\longrightarrow} s$. □

Corollary 2.1. *The sequence* $\operatorname{vec}^{-1}(x^{(k)})$ *converges to a solvent of* $P(X)$.

3. Numerical examples

The implementation of Algorithm 2.1 poses several challenges. First, for degrees $m > 2$ the expressions for the function f become very extensive, so we only carry out examples with $m = 2$. Second, considering that matrix polynomials with no solvents is a common situation, Theorem 2.5 must fail in at least one condition. Usually we perform the test for the spectral radius of the Jacobian of the function f evaluated for $x^{(0)}$, where $x^{(0)} = \mathrm{vec}(X_0)$ and X_0 is the initial guest.

We made various experiments with matrices of orders $n = 2$ and $n = 3$ to verify the usefulness of the method. Next we present two examples for the two-dimensional case ($n = 2$).

Example 3.1. We consider the quadratic matrix polynomial

$$P(X) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} X^2 + \begin{bmatrix} -5 & 0 \\ -34.667 & -4 \end{bmatrix} X + \begin{bmatrix} 4 & 0 \\ 34.667 & 104 \end{bmatrix}.$$

It can be verified that the equation $P(X) = 0_n$ has 5 solvents, namely

$$S_1 = \begin{bmatrix} 1 & 0 \\ 0 & 2+10\mathrm{i} \end{bmatrix}, \quad S_2 = \begin{bmatrix} 1 & 0 \\ 0 & 2-10\mathrm{i} \end{bmatrix}, \quad S_3 = \begin{bmatrix} 1 & 3 \\ 0 & 4 \end{bmatrix},$$
$$S_4 = \begin{bmatrix} 4 & 0 \\ 2-10\mathrm{i} & 2+10\mathrm{i} \end{bmatrix}, \quad S_5 = \begin{bmatrix} 4 & 0 \\ 2+10\mathrm{i} & 2-10\mathrm{i} \end{bmatrix}.$$

None of these solvents is a dominant one, so in this case Traub's method does not converge.

We take as the initial guest $X_0 = S_5 + 0.2I_n$, $\|P(X^{(0)})\|_2 = 0.91307$.

We have $x^{(0)} = \mathrm{vec}(X_0)$. Computing the spectral radius we obtain $\varrho(f'(x^{(0)})) = 0.61758$.

The algorithm gives, after 21 iterations,

$$\begin{aligned} X_{21} &= \mathrm{vec}^{-1}(x^{(21)}) \\ &= \begin{bmatrix} 4 + 2.05755 \times 10^{-6}\mathrm{i} & 0 \\ 2.00001 - 10\mathrm{i} & 2+10\mathrm{i} \end{bmatrix}, \quad \text{with } \|P(X_{21})\|_2 = 0.000045. \end{aligned}$$

Now, with the initial guest $X_0 = -A_1$, we have $\varrho(f'(x^{(0)})) = 0.54554$.

We also get after 23 iterations an approximation to S_5, that is,

$$X_0 = \begin{bmatrix} 5 & 0 \\ 34.667 & 4 \end{bmatrix} \quad \text{with } \|P(X_0)\|_2 = 109.663$$

and so

$$X_{23} = \begin{bmatrix} 4.00002 & 0 \\ 2.00004 - 10.0001\mathrm{i} & 2 + 10\mathrm{i} \end{bmatrix} \quad \text{with } \|P(X_{23})\|_2 = 0.000349.$$

E x a m p l e 3.2. We consider another quadratic matrix polynomial

$$P(X) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} X^2 + \begin{bmatrix} -0.15 & -0.075 \\ 0.01 & -0.355 \end{bmatrix} X + \begin{bmatrix} 6.1333 & -9.46667 \\ -2.7333 & 33.0333 \end{bmatrix}.$$

We choose

$$X_0 = \begin{bmatrix} -36.1317 & 0 \\ 0 & -36.1317 \end{bmatrix} \quad \text{with } \varrho(f'(x^{(0)})) = 0.78877.$$

After 6 iterations we get the following approximation for a solvent:

$$X_6 = \begin{bmatrix} 0.075918 + 2.39468\mathrm{i} & 0.03808 - 1.166698\mathrm{i} \\ 0.007214 - 0.33701\mathrm{i} & 0.17669 + 5.71041\mathrm{i} \end{bmatrix} \quad \text{with} \quad \|P(X_6)\|_2 = 0.00128.$$

In this example, our implementation for Newton's method does not converge.

The main limitations that we experimented are common to all known methods for computing solvents for matrix polynomials, there is no sure condition for the choice of the initial guest and the result is in general very unpredictable, even when it converges. But we believe that the method we present could be useful mainly for solving problems related to second-order matrix differential equations.

The authors wish to thank the anonymous Referee and the Executive Editor for providing helpful comments, suggestions and corrections that improved this paper.

References

[1] *G. J. Davis*: Numerical solution of a quadratic matrix equation. SIAM J. Scient. Computing *2* (1981), 164–175.

[2] *E. Dennis, J. F. Traub, R. P. Weber*: On the Matrix Polynomial, Lambda-Matrix and Block Eigenvalue Problems. Computer Science Department, Technical Report, Cornell University, Ithaca, New York and Carnegie-Mellon University, Pittsburgh, Pennsylvania, 1971.

[3] *J. E. Dennis, J. F. Traub, R. P. Weber*: The algebraic theory of matrix polynomials. SIAM J. Numer. Anal. *13* (1976), 831–845.

[4] *J. E. Dennis, J. F. Traub, R. P. Weber*: Algorithms for solvents of matrix polynomials. SIAM J. Numer. Anal. *15* (1978), 523–533.

[5] *I. Gohberg, P. Lancaster, L. Rodman*: Matrix Polynomials. Academic Press, New York, 1982.

[6] *N. J. Higham, H. M. Kim*: Solving a quadratic matrix equation by Newton's method with exact line searchers. SIAM J. Matrix Anal. Appl. *23* (2001), 303–316.

[7] *N. J. Higham, H. M. Kim*: Numerical analysis of a quadratic matrix equation. IMA J. Numer. Anal. *20* (2000), 499–519.

[8] *R. B. Holmes*: A formula for the spectral radius of an operator. Am. Math. Mon. *75* (1968), 163–166.
[9] *W. Kratz, E. Stickel*: Numerical solution of matrix polynomial equations by Newton's method. IMA J. Numer. Anal. *7* (1987), 355–369.
[10] *P. Lancaster*: Lambda-Matrices and Vibrating Systems. Pergamon Press, New York, 1966.
[11] *P. Lancaster*: A fundamental theorem on lambda matrices with applications II. Difference equations with constant coefficients. Linear Algebra Appl. *18* (1977), 213–222.
[12] *P. Lancaster, M. Tismenetsky*: The Theory of Matrices, 2nd edition. Academic Press, New York, 1985.
[13] *E. Pereira, J. Vitória*: Deflation of block eigenvalues of block partitioned matrices with an application to matrix polynomials of commuting matrices. Comput. Math. Appl. *42* (2001), 1177–1188.
[14] *E. Pereira, R. Serodio, J. Vitória*: Newton's method for matrix polynomials. Int. J. Math. Game Theory Algebra *17* (2008), 183–188.
[15] *M. Shih, J. Wu*: Asymptotic stability in the Schauder fixed point theorem. Stud. Math. *2* (1998), 143–148.
[16] *F. Tisseur, K. Meerbergen*: The quadratic eigenvalue problem. SIAM Rev. *43* (2001), 235–286.
[17] *J. S. H. Tsai, L. S. Shieh, T. T. C. Shen*: Block power method for computing solvents and spectral factors of matrix polynomials. Comput. Math. Appl. *16* (1988), 683–699.

Authors' addresses: *Fernando Marcos*, DM, Instituto Politécnico da Guarda, Guarda, Portugal, e-mail: `marcos@ipg.pt`; *Edgar Pereira* (corresponding author), IT-DI, Universidade da Beira Interior, 6201-001 Covilhã, Portugal, e-mail: `edgar@di.ubi.pt`.