Alexandria Journal of Mathematics (ISSN 2090-4320)
Volume1 -Number1 - June 2010

On the Characteristic Polynomial of Regular Linear Matrix Pencil

Yan Wu and Phillip D. Lorren
Department of Mathematical Sciences
Georgia Southern University
Statesboro, GA 30460 USA

Abstract
Linear matrix pencil, denoted by (***A***,***B***), plays an important role in control systems and numerical linear algebra. The problem of finding the eigenvalues of (***A***,***B***) is often solved numerically by using the well-known QZ method. Another approach for exploring the eigenvalues of (***A***,***B***) is by way of its characteristic polynomial, $P(\lambda)=|A-\lambda B|$. There are other applications of working directly with the characteristic polynomial, for instance, using Routh-Hurwitz analysis to count the stable roots of $P(\lambda)$ and transfer function representation of control systems governed by differential-algebraic equations. In this paper, we present an algorithm for algebraic construction of the characteristic polynomial of a regular linear pencil. The main theorem reveals a connection between the coefficients of $P(\lambda)$ and a lexicographic combination of the rows between matrices ***A*** and ***B***.

Keywords: Regular matrix pencil, characteristic polynomial, generalized eigenvalue problem, choose function, combinatorics, lexicographic order.

1. Introduction

A matrix pencil of degree n is defined as an n_{th} degree polynomial with matrix-coefficients as follows,

$$\Gamma(\lambda)=A_n\lambda^n+A_{n-1}\lambda^{n-1}+...+A_1\lambda+A_0$$

where $\lambda\in C$, $A_i\in C^{n\times n}, i=1,2,...,n$, and $A_n\neq \underline{0}$, a zero matrix. A particular case known as linear matrix pencil, denoted as $(A,B)=A-\lambda B$, has been studied extensively with various applications. A matrix pencil $A-\lambda B$ is said to be regular if

$$|A-\lambda B|\neq 0 \text{ for some } \lambda\in C. \tag{1.1}$$

A regular matrix pencil basically excludes the case that all complex numbers are (generalized) eigenvalues for the pencil. In other words, there exists a non-trivial characteristic polynomial for (***A***,***B***).

Matrix pencils, particularly linear matrix pencils, play important roles in numerical linear algebra as well as applications in control systems and signal processing. Ahmad and Byers [1] revealed the relation between the critical points of approximating the eigenvalues of matrix pencils and the pseudospectra of perturbed pencils. Perturbation of the spectra of diagonalizable matrix pencils is studied in [2] via unitary matrices as generalized commutators. Matrix pencils are used extensively in the studies of control systems with linear descriptors, such as

$$E\,\dot{x} = Ax + Bu\,, \tag{1.2}$$

where x is the state vector, u is the control input, and E, A, and B are matrices with appropriate dimensions. System (1.2) is known as a linear time-invariant descriptor system. Condition (1.1) serves as the necessary and sufficient condition for the existence and uniqueness of the solution of (1.2), see [3]. Fundamental control theories for the discrete analog of (1.2) can be found in [4]. Matrix pencils are also important tools widely adopted in the area of signal processing, see [5-7]. Generalized eigenvalue problems of matrix pencils have drawn great interests for decades from both mathematicians and engineers. Qualitative and quantitative analysis of generalized eigenvalue problems are essential to the studies and applications of matrix pencils. A relatively thorough survey on the spectra of regular matrix pencils is available in [8-9]. Various numerical algorithms were developed for solving the generalized eigenvalue problems of matrix pencils, see [10] and the references therein. However, the most popular algorithm for such a task is the so-called QZ-algorithm, which is due to the following theorem [11]:

Theorem 1.1 If A and B are in $C^{n\times n}$, then there exist unitary Q and Z such that $Q^H AZ = T$ and $Q^H BZ = S$ are upper triangular. If for some k, t_{kk} and s_{kk} are both zero, then $\lambda(A,B) = C$. Otherwise, $\lambda(A,B) = \{t_{ii}/s_{ii}\,, s_{ii} \neq 0\}$, where t_{ii} and s_{ii} are on the main diagonal of T and S, respectively.

Since we are focusing on regular matrix pencils in this paper, the case of $\lambda(A,B) = C$ is excluded. The QZ-algorithm is an implicit form of the well-known QR-algorithm for solving the eigenvalue problem $B^{-1}Av = \lambda v$ without explicitly formulating $B^{-1}A$ because, in most cases, the matrix B is not invertible. Moreover, the pencil (A,B) has the same number of eigenvalues at infinity as the number of zero eigenvalues from B.

Another useful approach for finding the eigenvalues of a matrix pencil is from its associated characteristic polynomial,

$$P(\lambda) = |A - \lambda B|, \tag{1.3}$$

due to the fact that there are a number of algorithms available for finding the roots of polynomials [12-14], and it is still an ongoing, active research area in numerical analysis. Polynomial (1.3) can be written explicitly in the following form,

$$P(\lambda) = a_r\lambda^r + a_{r-1}\lambda^{r-1} + ... + a_1\lambda + a_0\,,\ a_r \neq 0\,,\ 1 \le r \le n\,. \tag{1.4}$$

It can be easily seen from (1.4) that the pencil (A,B) has $n-r$ eigenvalues at infinity. It is preferable to obtain exact characteristic polynomial (1.4) from (1.3), i.e. no roundoff errors in the coefficients of $P(\lambda)$, if one chooses to find $\lambda(A,B)$ from (1.4). The characteristic polynomial (1.4) has other applications, such as in the extension of classical Cayley-Hamilton theorem to regular matrix pencils [15]. It is the purpose of this paper to propose a numerical approach for symbolically constructing $P(\lambda)$ from (1.3).

2. The Main Result

It is indeed possible to construct exact characteristic polynomial of a pencil (A,B) from (1.3) because the expansion of a determinant only requires additions and multiplications. No divisions are involved in the arithmetic operations. The algorithm is designed for numerical software such as MatLab instead of Maple, which already has the built-in symbolic functionalities.

The coefficients of the characteristic polynomial (1.4) are related to (1.3) via Taylor polynomial,

$$a_k = \frac{P^{(k)}(0)}{k!} = \frac{1}{k!}\frac{d^k}{d\lambda^k}|A-\lambda B|_{\lambda=0},\ k=0,1,\ldots,r. \tag{2.1}$$

Formula (2.1) reveals that higher-order derivatives of the determinant of the pencil are needed for computing the coefficients. Since the pencil is linear, it will be shown that $\frac{d^k}{d\lambda^k}|A-\lambda B|_{\lambda=0}$ can be carried out by forming new matrices from interchanging the rows between A and B and calculating the determinant of the resultant matrices.

To set the stage, let $F(y)=(f_{ij}(y))$ be an n by n matrix function and $F_i(y)$ represents the i_{th} row of the matrix, i.e. $F_i(y)=[f_{i1}(y)\quad f_{i2}(y)\quad \ldots\quad f_{in}(y)]$. It is understood that

$$\frac{d}{dy}F_i(y)=[\frac{d}{dy}f_{i1}(y)\quad \frac{d}{dy}f_{i2}(y)\quad \ldots\quad \frac{d}{dy}f_{in}(y)]$$

It is also understood that the notation $\left|\frac{d}{dy}f_{k1}(y)\ldots\frac{d}{dy}f_{kn}(y)\right|$ means the determinant is taken on $F(y)$ with its k_{th} row being replaced by $\frac{d}{dy}F_k(y)$. The following theorem provides a differentiation rule for a functional determinant, which will be used in (2.1).

Theorem 2.1 Suppose $F(y)=(f_{ij}(y))\in C^{n\times n}$ and $f_{ij}(y)\in C^1_{(a,b)}$. Then,

$$\frac{d}{dy}|F(y)| = \begin{vmatrix} \frac{d}{dy}F_1(y) \\ F_2(y) \\ \vdots \\ F_n(y)\end{vmatrix} + \begin{vmatrix} F_1(y) \\ \frac{d}{dy}F_2(y) \\ \vdots \\ F_n(y)\end{vmatrix} + \ldots + \begin{vmatrix} F_1(y) \\ F_2(y) \\ \vdots \\ \frac{d}{dy}F_n(y)\end{vmatrix} = \sum_{k=1}^{n}\left|\frac{d}{dy}f_{k1}(y)\ldots\frac{d}{dy}f_{kn}(y)\right| \tag{2.2}$$

where $|.|$ represents determinant.

Proof: This is done by induction on *n*. First of all, it is easy to verify that, if $F \in C^{2\times 2}$,

$$\frac{d}{dy}|F(y)| = \begin{vmatrix} \frac{d}{dy}F_1(y) \\ F_2(y) \end{vmatrix} + \begin{vmatrix} F_1(y) \\ \frac{d}{dy}F_2(y) \end{vmatrix}$$

based on the product rule for scalar functions. Now, assume (2.2) is true for all *k* by *k* matrix functions. Let $F \in C^{(k+1)\times(k+1)}$, use cofactor expansion along the first row of $F(y)$, let C_{1j} be the cofactor of f_{1j},

$$\frac{d}{dy}|F(y)| = \sum_{j=1}^{k+1}\frac{d}{dy}(f_{1j}C_{1j}) = \sum_{j=1}^{k+1}[C_{1j}\frac{d}{dy}f_{1j} + f_{1j}\frac{d}{dy}C_{1j}]$$

$$= \begin{vmatrix} \frac{d}{dy}f_{11} & \frac{d}{dy}f_{12} & \cdots & \frac{d}{dy}f_{1(k+1)} \end{vmatrix} + f_{11}\sum_{i=2}^{k+1}\begin{vmatrix} \frac{d}{dy}f_{i2}(y) \ldots \frac{d}{dy}f_{i(k+1)}(y) \end{vmatrix} +$$

$$\sum_{j=2}^{k+1}(-1)^{j-1}f_{1j}\sum_{l=2}^{k+1}\begin{vmatrix} \frac{d}{dy}f_{l1}(y) & \ldots & \frac{d}{dy}f_{l(j-1)}(y) & \frac{d}{dy}f_{l(j+1)}(y) & \ldots & \frac{d}{dy}f_{l(k+1)}(y) \end{vmatrix}$$

$$= \begin{vmatrix} \frac{d}{dy}f_{11} & \frac{d}{dy}f_{12} & \cdots & \frac{d}{dy}f_{1(k+1)} \end{vmatrix} + \sum_{j=2}^{k+1}\begin{vmatrix} \frac{d}{dy}f_{j1}(y) \ldots \frac{d}{dy}f_{j(k+1)}(y) \end{vmatrix}$$

$$= \sum_{j=1}^{k+1}\begin{vmatrix} \frac{d}{dy}f_{j1}(y) \ldots \frac{d}{dy}f_{j(k+1)}(y) \end{vmatrix} \qquad \square$$

Theorem 2.1 is useful for deriving a closed form formula for $\frac{d^k}{d\lambda^k}|A-\lambda B|_{\lambda=0}$. We will use a similar notation to that in Theorem 2.1 for the determinant of a matrix. Let A_i and B_i represent the rows of matrices *A* and *B*, respectively, and, for the sake of argument, the determinant of *A* is written as $|A| = \begin{vmatrix} A_1 \\ A_2 \\ \vdots \\ A_n \end{vmatrix}$. We begin with the first derivative,

$$\frac{d}{d\lambda}|A-\lambda B| = \begin{vmatrix} -B_1 \\ A_2-\lambda B_2 \\ \vdots \\ A_n-\lambda B_n \end{vmatrix} + \begin{vmatrix} A_1-\lambda B_1 \\ -B_2 \\ \vdots \\ A_n-\lambda B_n \end{vmatrix} + \ldots + \begin{vmatrix} A_1-\lambda B_1 \\ A_2-\lambda B_2 \\ \vdots \\ -B_n \end{vmatrix} \tag{2.3}$$

There are *n* determinants in (2.3). It is easy to see that, if one differentiates the right side of (2.3) again, each determinant would produce $(n-1)$ non-trivial (in the sense of noneexistence of zero rows in the matrix) determinants. For instance, the first determinant obtained from differentiating $\begin{vmatrix} -B_1 \\ A_2-\lambda B_2 \\ \vdots \\ A_n-\lambda B_n \end{vmatrix}$ in (2.3) yields a zero row on the first row. Therefore, that determinant is dropped because it equals zero. As a result, there are $n(n-1)$ non-trivial determinants in $\frac{d^2}{d\lambda^2}|A-\lambda B|$. Continuing this process, it can be induced that there are exactly $n(n-1)\cdots(n-k+1)$ non-trivial determinants in $\frac{d^k}{d\lambda^k}|A-\lambda B|$. It is also noticed that, in order to obtain $\frac{d^k}{d\lambda^k}|A-\lambda B|_{\lambda=0}$, all one has to do is to replace *k* rows in $A=\begin{bmatrix} A_1 \\ A_2 \\ \vdots \\ A_n \end{bmatrix}$ by the corresponding *k* rows from *B* (multiplied by -1) in a lexicographic order, which will be defined below, and evaluate the determinant for each new matrix as a result of combining the rows between *A* and *B*. This operation is easily observed from (2.3) for $k=1$ after substituting $\lambda=0$.

Definition 2.1 *A set of m-tuples, denoted by* $R^m_{i_1,i_2,\ldots i_n}$, $0<i_1<i_2<\ldots<i_n$, $m\le n$, *is said to be lexicographic if* $R^m_{i_1,i_2,\ldots i_n}$ *is defined as*

$$R^m_{i_1,i_2,\ldots i_n}=\left\{(j_1 \quad j_2 \quad \ldots \quad j_m),\ j_k<j_l \text{ if } k<l,\ j_k\in\{i_1 i_2\ldots i_n\},\ k,l=1,2,\ldots,m\right\} \tag{2.4}$$

Obviously, there are $\begin{pmatrix} n \\ m \end{pmatrix}$ (choose function, also denoted as C^m_n) elements in $R^m_{i_1,i_2,\ldots i_n}$. The set $R^m_{i_1,i_2,\ldots i_n}$ can be easily constructed since (2.4) turns out to be a simple combinatorial problem. Before stating the main theorem, we introduce a notation for row substitutions between two *n* by *n* matrices, i.e. $\langle A_x | B_x \rangle$, which means replacing *m* rows of elements in *A* with the corresponding *m* rows from *B*. The locations of the rows are determined by the *m*-tuple *x*. There is no replacement if $m=0$, i.e. $\langle A_x | B_x \rangle = A$; on the other hand, $\langle A_x | B_x \rangle = B$ if $m=n$. Other cases can be done accordingly. For example, let $x=(1,3,4)$, then, $\langle A_x | B_x \rangle$ represents a new matrix generated from *A* by replacing its 1_{st}, 3_{rd}, and 4_{th} rows with the corresponding rows in *B*, respectively.

Theorem 2.2 Let $A,B\in C^{n\times n}$ and $A-\lambda B$ be a regular pencil. Suppose its characteristic polynomial is given by (1.4). Then, the coefficients of the polynomial satisfy

$$a_k = \sum_{x \in R^k_{1,2,\ldots,n}} \left|\left\langle A_x \middle| -B_x \right\rangle\right|,\ k = 0,1,\ldots,r. \tag{2.5}$$

where $|.|$ represents determinant.

Proof: The coefficients a_k are related to the Taylor polynomial according to (2.1). Hence, it is sufficient to find an alternate expression for $\frac{d^k}{d\lambda^k}|A-\lambda B|_{\lambda=0}$. According to Theorem 2.1, there are $n(n-1)\cdots(n-k+1)$ non-trivial determinants in $\frac{d^k}{d\lambda^k}|A-\lambda B|$. After substituting $\lambda = 0$, the matrix in each determinant of $\frac{d^k}{d\lambda^k}|A-\lambda B|_{\lambda=0}$ is obtained from replacing certain k rows of A by the same number of corresponding rows from $-B$. It is readily seen from the proof of Theorem 2.1 that only product rule for differentiation is used in deriving the expression for the derivative of a functional determinant. Due to the product rule, each function in a product is differentiated exactly once. Therefore, there exists parity among the determinants in $\frac{d^k}{d\lambda^k}|A-\lambda B|_{\lambda=0}$ for the locations of those k rows of replacement, i.e. there are exactly $\binom{n}{k}$ ways of combinations of those k rows in matrix A. The locations of those k rows thereby follow the lexicographic order (2.4). This also shows that there are $\binom{n}{k}$ distinct determinants in $\frac{d^k}{d\lambda^k}|A-\lambda B|_{\lambda=0}$. However, $\binom{n}{k}$ is no greater than $n(n-1)\cdots(n-k+1)$, which implies that there are repeated determinants in $\frac{d^k}{d\lambda^k}|A-\lambda B|_{\lambda=0}$. Again, due to the parity among the determinants in $\frac{d^k}{d\lambda^k}|A-\lambda B|_{\lambda=0}$, each distinct determinant in the expansion of $\frac{d^k}{d\lambda^k}|A-\lambda B|_{\lambda=0}$ must be repeated the same number of times, and it is $k!$ to be exact. This is because of the following identity

$$\binom{n}{k} \cdot k! = n(n-1)\cdots(n-k+1),$$

which agrees with the total number of determinants in $\frac{d^k}{d\lambda^k}|A-\lambda B|_{\lambda=0}$ before combining the like terms (determinants). It is also easy to see that each distinct determinant in the expansion of $\frac{d^k}{d\lambda^k}|A-\lambda B|_{\lambda=0}$ can be written as $\left|\left\langle A_x \middle| -B_x \right\rangle\right|$, $x \in R^k_{1,2,\ldots,n}$. Therefore,

$$\frac{d^k}{d\lambda^k}\left|A-\lambda B\right|_{\lambda=0} = k!\sum_{x\in R^k_{1,2,\ldots,n}} \left|\langle A_x \,|\, -B_x\rangle\right|,\ k=0,1,\ldots,r\,. \qquad (2.6)$$

□

Notice that, because of the $k!$ in (2.6) that is eventually canceled by $\frac{1}{k!}$ in (2.1), calculating a_k using (2.5) only requires additions and multiplications as a result of evaluating the determinants. Hence, there are no roundoff errors in computing the coefficients numerically. This is particularly significant when the matrices in the pencil are integral matrices, namely the entries are integers. With the proposed algorithm, the coefficients of the characteristic polynomial will be integers as computed.

We conclude with an example to illustrate the algorithm proposed in this paper for symbolically constructing the characteristic polynomial of a regular linear matrix pencil. Suppose A and B are 4 by 4 matrices, and the characteristic polynomial of $A-\lambda B$ is written as

$$P(\lambda) = a_4\lambda^4 + a_3\lambda^3 + a_2\lambda^2 + a_1\lambda + a_0\,.$$

According to (2.6), the coefficients are

$$a_0 = |A|, \qquad a_4 = |-B|,$$

$$a_1 = \begin{vmatrix} -B_1 \\ A_2 \\ A_3 \\ A_4 \end{vmatrix} + \begin{vmatrix} A_1 \\ -B_2 \\ A_3 \\ A_4 \end{vmatrix} + \begin{vmatrix} A_1 \\ A_2 \\ -B_3 \\ A_4 \end{vmatrix} + \begin{vmatrix} A_1 \\ A_2 \\ A_3 \\ -B_4 \end{vmatrix},$$

$$a_2 = \begin{vmatrix} -B_1 \\ -B_2 \\ A_3 \\ A_4 \end{vmatrix} + \begin{vmatrix} -B_1 \\ A_2 \\ -B_3 \\ A_4 \end{vmatrix} + \begin{vmatrix} -B_1 \\ A_2 \\ A_3 \\ -B_4 \end{vmatrix} + \begin{vmatrix} A_1 \\ -B_2 \\ -B_3 \\ A_4 \end{vmatrix} + \begin{vmatrix} A_1 \\ -B_2 \\ A_3 \\ -B_4 \end{vmatrix} + \begin{vmatrix} A_1 \\ A_2 \\ -B_3 \\ -B_4 \end{vmatrix},$$

and

$$a_3 = \begin{vmatrix} -B_1 \\ -B_2 \\ -B_3 \\ A_4 \end{vmatrix} + \begin{vmatrix} -B_1 \\ -B_2 \\ A_3 \\ -B_4 \end{vmatrix} + \begin{vmatrix} -B_1 \\ A_2 \\ -B_3 \\ -B_4 \end{vmatrix} + \begin{vmatrix} A_1 \\ -B_2 \\ -B_3 \\ -B_4 \end{vmatrix}.$$

References:

[1] SK. S. Ahmad, R. Alam, and R. Byers, On Pseudospectra, Critical Points and Multiple Eigenvalues of Matrix pencils, *SIAM J. Matrix Anal. & Appl*., v. 31 (4), 1915-1933, 2010.

[2] R. C. Li, On Perturbations of Matrix Pencils with Real Spectra, A Revisit, *Math. Comp*. v. 72 (242), 715-728, 2003.

[3] E. L. Yip and R. F. Sincovec, Solvability, Controllability, and Observability of Continuous Descriptor Systems, *IEEE Trans. Automat. Contr*., v 26, 702-707, 1981.

[4] F. Lewis, Fundamental, Reachability, and Observability Matrices for Discrete Descriptor Systems, *IEEE Trans. Automat. Contr*., v 30 (5), 502-505, 1985.

[5] C. Chang, Z. Ding, S. F. Yau, and F. H. Y. Chan, A Mztrix-Pencil Approach to Blind Separation of Colored Nonstationary Signals, *IEEE Trans. Signal Process*., v. 48 (3), 900-907, 2000.

[6] S. Wang, X. Guan, D. Wang, X. Ma, and Y. Su, Application of Matrix Pencil Method for Estimating natural Resonances of Scatterers, *Electronic Letters*, v. 43 (1), 3-5, 2007.

[7] G. Nico and J. Fortuny, Using the Matrix Pencil Method to Solve Phase Unwrapping, *IEEE Trans. Signal Process*., v. 51 (3), 886-888, 2003.

[8] F. R. Gantmacher, *The Theory of Matrices*, Chelsea Publishing Company, New York, 1959.

[9] J. H. Wilkinson, The Algebraic Eigenvalue Problems, Clarendon, Oxford, 1965.

[10] G. H. Golub and Q. Ye, An Inverse Free Preconditioned Krylov Subspace Method for Symmetric Generalized Eigenvalue Problems, *SIAM J. Sci. Comput*. v. 24 (1), 312-334, 2002.

[11] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 2nd Ed., The Johns Hopkins University Press, Baltimore, Maryland, 1989.

[12] M. H. Kim and S. Sutherland, Polynomial root-finding algorithms and branched covers, *SIAM J. Comput*., v. 23 (2), 415-436, 1994.

[13] Z. Zeng, Algorithm 835: Multroot-a MatLab package for computing polynomial roots and multiplicities, *ACM Trans. Math. Software*, v. 30 (2), 218-236, 2004.

[14] L. Brugnano and D. Trigiante, Polynomial Roots: The Ultimate Answer?, *Linear Alg. and Its Appl.*, v. 225, 207-219, 1995.

[15] F. R. Chang and H. C. Chen, The generalized Cayley-Hamilton theorem for standard pencils, *Syst. Control Lett.*, v. 18, 179-182, 1992.