On regular polynomial matrices I. A new criterion for a polynomial matrix to be simple

By ANNA LEE (Budapest)

1. Introduction

There is an extensive literature about polynomial matrices, beginning with works of Weierstrass and Kronecker and giving new results even in our days (cf. [1]—[9]). The first investigations were related to the problem of simultaneous transformation of two quadratic forms into sums of quadrats. The discussions of Weierstrass in this topic lead to his theory of elementary divisors. In the treatment of polynomial matrices of higher degree there is also made use of the right and left sided division of the polynomial matrix in question by another polynomial matrix ([3]—[7]). But even up to now the theory of elementary divisors proved to be the most effective tool in the investigations, giving a rather precise information about the structure of some polynomial matrices.

Polynomial matrices play an important role in the modern discussion of damped vibrating systems. The importance of such matrices — especially simple ones — is fairly demonstrated in LANCASTER's book [7]. Here there is also a bibliography of polynomial matrices.

The present paper is an introductory note on our investigations in this field. The method applied is different from both the theory of elementary divisors an the method of one-sided division as well. We will analyse the reduced adjoint of the polynomial matrix in question (i.e. the adjoint, divided by the gcd of its elements), and the derivatives of this matrix. This new aproach is useful for the very reason, since the comparison of statements obtained by the different methods yields sometimes interesting new results. (c. f. [11])

2. Perliminaries

We begin with some definitions, since the use of several notions concerning polynomial matrices is not uniform.

A polynomial matrix (or lambda-matrix) is a matrix with entries from $K[\lambda]$, the polynomial ring over the complex number field K. Any polynomial matrix of degre l can be represented as

(2.1) $$D_l(\lambda) = A_0\lambda^l + A_1\lambda^{l-1} + \dots + A_l \quad (A_0 \neq 0),$$

where the entries of the matrices A_i are from K.

Definition 1. The polynomial matrix (2.1) is *regular*, if

(i) all A_i are n-square matrices,

(ii) $\det D_l(\lambda) = |D_l(\lambda)| \not\equiv 0$.

This definition of the regular polynomial matrix does not exclude the case of $|A_0|=0$. Therefore it is more suitable instead of (2.1) the homogeneous form

$$D_l(\lambda, u) = A_0\lambda^l + A_1\lambda^{l-1}\mu + \dots + A_l u^l \quad (A_i \text{ are } n\text{-squares}). \tag{2.2}$$

The characteristic polynomial of the regular polynomial matrix (2.2) can be represented as

$$\Delta(\lambda, \mu) = |D_l(\lambda, \mu)| = c\mu^{\alpha_0} \prod_{k=1}^{s} (\lambda - \lambda_k\mu)^{\alpha_k} \tag{2.3}$$

$$c \neq 0; \quad \lambda_k \neq \lambda_i \Leftrightarrow k \neq i; \quad \sum_{k=1}^{s} \alpha_k = nl.$$

In virtue of (2.3) a regular polynomial matrix (2.2) has always exactly nl eigenvalues if they are counted according to their multiplicities. The usual (finite) eigenvalues λ_k of multiplicity α_k are given by (2.3), substituting $\mu=1$ $(k=1, \dots, s)$, and we will denote by λ_0 the "infinite eigenvalue" of multiplicity α_0 given by (2.3), if substituting $\mu=0$.

In the following a regular polynomial matrix $D_l(\lambda, \mu)$ will mean always a polynomial matrix represented by (2.2) and (2.3).

For arbitrary polynomial $f(\lambda, \mu)$ homogeneous in λ and μ we will use the short notation

$$f(\lambda_k) = \begin{cases} f(\lambda, \mu)|_{\lambda=\lambda_k, \mu=1} & k = 1, \dots, s, \\ f(\lambda, \mu)|_{\lambda=1, \mu=0} & k = 0, \end{cases} \tag{2.4}$$

$$f'(\lambda_k) = \begin{cases} \left.\dfrac{\partial f(\lambda, \mu)}{\partial \lambda}\right|_{\lambda=\lambda_k, \mu=1} & k = 1, \dots, s, \\ \left.\dfrac{\partial f(\lambda, \mu)}{\partial \mu}\right|_{\lambda=1, \mu=0} & k = 0, \end{cases}$$

where λ_k $(k=0, 1, \dots, s)$ are the roots of $\Delta(\lambda, \mu)$ in (2.3). (The roots of a polynomial $f(\lambda, \mu)$ will always mean the — finite and infinite — roots of the polynomial equation $f(\lambda, \mu)=0$).

Definition 2. A regular polynomial matrix $D_l(\lambda, \mu)$ is *simple*, if

$$\varrho[D_l(\lambda_k)] = n - \alpha_k, \quad k = 0, 1, \dots, s$$

where ϱ (X) denotes the rank of the matrix X.

This definition of the simple polynomial matrix is more general than that of Lancaster (cf. [7] pp. 42.), since it does not exclude the case $|A_0|=0$.

Let $\Delta_j(\lambda, \mu)$ denote the gcd of all jth order subdeterminants of $D_l(\lambda, \mu)$ ($j=1, \ldots, n$). The polynomials $\Delta_j(\lambda, \mu)$ are called the *determinantal divisors* of $D_l(\lambda, \mu)$ and we set $\Delta_0(\lambda, \mu)=1$ for notational convenience. The *invariant factors* of $D_l(\lambda, \mu)$ are defined by

$$i_j(\lambda, \mu) \doteq \frac{\Delta_j(\lambda, \mu)}{\Delta_{j-1}(\lambda, \mu)} \quad j = 1, \ldots, n; \; \Delta_n(\lambda, \mu) \equiv \frac{1}{c}\Delta(\lambda, \mu).$$

We will need only the reduced characteristic polynomial $\delta(\lambda, \mu)$, which is defined by

$$\delta(\lambda, \mu) \doteq \frac{\Delta(\lambda, \mu)}{\Delta_{n-1}(\lambda, \mu)} = c\mu^{\beta_0} \prod_{k=1}^{s} (\lambda - \lambda_k \mu)^{\beta_k} \tag{2.5}$$

and for which $\delta(\lambda, \mu) \equiv c i_n(\lambda, \mu)$.

Let us define the homogeneous Lagrangean-polynomials belonging to an arbitrary homogeneous polynomial

$$f(\lambda, \mu) = a\mu \prod_{k=1}^{s} (\lambda - \lambda_k \mu), \quad \lambda_k \neq \lambda_i \Leftrightarrow k \neq i$$

as follows:

$$L_k(\lambda, \mu) \doteq \frac{f(\lambda, \mu)}{f'(\lambda_k)(\lambda - \lambda_k \mu)}, \quad k = 1, \ldots, s \tag{2.6}$$

$$L_0(\lambda, \mu) \doteq \frac{f(\lambda, \mu)}{f'(\lambda_0)\mu}, \qquad k = 0.$$

Considering the notation (2.4), it is easy to see that the polynomials (2.6) fulfil the conditions

$$L_k(\lambda_i) = \delta_{ki} \quad k, i = 0, 1, \ldots, s. \tag{2.7}$$

The homogeneous polynomial $g(\lambda, \mu)$ of degree at most s, which satisfies the conditions

$$g(\lambda_k) = c_k \quad k = 0, 1, \ldots, s$$

can be constructed with help of the just defined Lagrangean polynomials as:

$$g(\lambda, \mu) = \sum_{k=0}^{s} c_k L_k(\lambda, \mu). \tag{2.8}$$

The reduced adjoint of $D_l(\lambda, \mu)$ is defined by

$$F(\lambda, \mu) \doteq \frac{\operatorname{adj} D_l(\lambda, \mu)}{\Delta_{n-1}(\lambda, \mu)}. \tag{2.9}$$

If $1 \leqq r \leqq n$, then $Q_{r,n}$ denote the totality of strictly increasing sequences of r integers chosen from $1, \ldots, n$. Let $X=(x_{ij})$ an n-square matrix. If $\omega=(i_1, \ldots, i_r)$ and $\tau=(j_1, \ldots, j_r)$ are sequences in $Q_{r,n}$, then $X[\omega|\tau]$ will denote the r-square *submatrix of* X *lying in rows* ω *and columns* τ. The notation $X(\omega|\tau)$ designate the $(n-r)$-square submatrix of X, whose rows and columns resp. are precisely those complementary to ω and τ resp.

We formulate with this notation a well-known determinantal indentity (cf. e.g. [10] pp. 66.) as follows:

$$|(\mathrm{adj}\, X)[\omega|\tau]| = (-1)^{\omega+\tau}|X|^{r-1}|X(\omega|\tau)|, \tag{2.10}$$

where

$$(-1)^{\omega+\tau} = (-1)^{\Sigma i_\nu + \Sigma j_\nu}, \quad i_\nu \in \omega, \quad j_\nu \in \tau.$$

3. Statements

Theorem 1. *If the reduced characteristic polynomial $\delta(\lambda, \mu)$ of the regular polynomial matrix $D_l(\lambda, \mu)$ has only simple roots, that is if in (2.5) the equality*

$$\beta_k = 1 \quad k = 0, 1, \dots, s \tag{3.1}$$

holds, then the matrices

$$P_k = \frac{1}{\delta'(\lambda_k)} F(\lambda_k) \quad k = 0, 1, \dots, s \tag{3.2}$$

satisfy the following conditions:

(i) $\varrho(P_k) = \varrho[D_l'(\lambda_k) P_k] = \varrho[P_k D_l'(\lambda_k)] = \alpha_k, \quad k = 0, 1, \dots, s$

(ii) $\left.\begin{matrix} [D_l'(\lambda_k) P_k]^2 = D_l'(\lambda_k) P_k \\ [P_k D_l'(\lambda_k)]^2 = P_k D_l'(\lambda_k) \end{matrix}\right\} \quad k = 0, 1, \dots, s.$

$\left.\begin{matrix} [D_l'(\lambda_k) P_k]^2 = D_l'(\lambda_k) P_k \\ [P_k D_l'(\lambda_k)]^2 = P_k D_l'(\lambda_k) \end{matrix}\right\}$

L trices (3.2) in basis factor, that is

$$P_k = U_k V_k^*, \quad \varrho(U_k) = \varrho(V_k) = \varrho(P_k) = \alpha_k \quad k = 0, 1, \dots, s \tag{3.3}$$

where U_k and V_k are matrices of type $n \times \alpha_k$ and asterisk on a matrix indicates the conjugate transpose. Then in virtue of

$$D_l(\lambda, \mu) F(\lambda, \mu) = \delta(\lambda, \mu) E \tag{3.4}$$

$$F(\lambda, \mu) D(\lambda, \mu) = \delta(\lambda, \mu) E$$

we get immediately

$$D_l(\lambda_k) U_k = 0; \quad V_k^* D_l(\lambda_k) = 0 \quad k = 0, 1, \dots, s. \tag{3.5}$$

On account of (3.3) and (3.5) there follows from theorem 1:

Corollary 1. According to the assumption of theorem 1 the columns of the matrix U_k give α_k linearly independent right eigenvectors and the rows of V_k^* give α_k linearly independent left eigenvectors belonging to the eigenvalue λ_k, which are biorthogonal with respect to $D_l'(\lambda_k)$ ($k = 0, 1, \dots, s$). Thus $D_l(\lambda, \mu)$ is a simple polynomial matrix.

By the assumption of theorem 1 the entries of the matrix $F(\lambda, \mu)$ defined by (2.9) are polynomials of degree at most s. Therefore (2.6)—(2.8) can be applied for $f(\lambda, \mu) = \delta(\lambda, \mu)$ to the polynomial matrix $F(\lambda, \mu)$ and hence we receive

$$F(\lambda, \mu) = \sum_{k=0}^{s} F(\lambda_k) \frac{\delta(\lambda, \mu)}{\delta'(\lambda_k)(\lambda - \lambda_\nu \mu)} \quad \lambda - \lambda_0 \mu \doteq \mu. \tag{3.6}$$

If $\lambda \neq \lambda_k$ $(k=0, 1, \dots, s)$, then $D_l(\lambda, \mu)$ and $F(\lambda, \mu)$ are invertable and from (3.4)

$$[D_l(\lambda, \mu)]^{-1} = \frac{1}{\delta(\lambda, \mu)} F(\lambda, \mu).$$

Thus due to (3.6) we obtain from theorem 1 also

Corollary 2. According to the assumption of theorem 1 the inverse of $D_l(\lambda, \mu)$ can be represented as

$$[D_l(\lambda, \mu)]^{-1} = \sum_{k=0}^{s} \frac{P_k}{(\lambda - \lambda_k \mu)} \quad \lambda \neq \lambda_k \quad k = 0, 1, \dots, s,$$

or — substituting the from (3.3) of P_k — as

$$(3.7) \qquad [D_l(\lambda, \mu)]^{-1} = U \left\langle \frac{E_{\alpha_0}}{\mu}, \frac{E_{\alpha_1}}{(\lambda - \lambda_1 \mu)}, \dots, \frac{E_{\alpha_s}}{(\lambda - \lambda_s \mu)} \right\rangle V^*,$$

where $\langle X_{\alpha_0}, X_{\alpha_1}, \dots, X_{\alpha_s} \rangle$ denotes a partitioned matrix having nonzero blocks in the main diagonal only, and the block standing in the kth position of the main diagonal is the α_k-square matrix X_{α_k} $(k=0, 1, \dots, s)$. The transformation matrices U and V^* are composed of the matrices (3.3) as

$$U = [U_0 U_1 \dots U_1] \quad V^* = [V_0 V_1 \dots V_s]^*.$$

It is easy to see that in the case $|A_0| = 0$ $(\alpha_0 = 0)$ (3.7) yields formula (4.4.11) of Lancaster's book (cf. [7] pp. 66.).

From our considerations discussed up to now it can be seen that the assumption of the polynomial $\delta(\lambda, \mu)$ having only simple roots is a *sufficient condition* for a regular polynomial matrix $D_l(\lambda, \mu)$ to be simple. This condition is necessary. There holds namely

Theorem 2. *If the regular polynomial matrix* $D_l(\lambda, \mu)$ *is not simple, then the polynomial* $\delta(\lambda, \mu)$ *has at least one multiple root, that is* $\beta_k > 1$ *holds in* (2.5) *for some index* k.

Theorem 2 means in other words: for a non-simple regular polynomial matrix alrady the n th invariant factor yields at least one non-linear elementary divisor.

The main result of the present paper is an immediate consequence of theorems 1and 2, namely

Theorem 3. *A regular polynomial matrix* $D_l(\lambda, \mu)$ *is simple if and only if the polynomial* $\delta(\lambda, \mu)$ *has only simple roots.*

This criterion can be formulated in a different way as well:

Theorem 3. *A regular polynomial matrix is simple if and only if its* nth *invariant factor yields only linear elementary divisors.*

From the latter formulation it is obvious that our criterion is based only on elementary divisors yielded by the nth invariant factor. Comparing this with the wellknown criterion of a *regular polynomial matrices being simple if and only if all its elementary divisors are linear**), we obtain

Theorem 4. *If the* nth *invariant factor of a regular polynomial matrix* $D_l(\lambda, \mu)$ *yields only linear elementary divisors, then all elementary divisors of* $D_l(\lambda, \mu)$ *are linear.*

4. Proofs

We fiirst show that

Lemma. *Let* $D_l(\lambda, \mu)$ *be a regular polynomial matrix,* $\Delta(\lambda, \mu)$ *its characteristic polynomial and* $\delta(\lambda, \mu)$ *the reduced characteristic polynomial. Each distinct root of* $\Delta(\lambda, \mu)$ *is a root of* $\delta(\lambda, \mu)$ *as well, that is in* (2.5) *there holds the inequality*

$$1 \leqq \beta_k \leqq \alpha_k \quad k = 0, 1, \dots, s. \tag{4.1}$$

Proof. Applying the identity (2.10) in the case $X = D_l(\lambda, \mu)$ and $r = n$ we obtain

$$|\mathrm{adj}\, D_l(\lambda, \mu)| = \Delta(\lambda, \mu)^{n-1}. \tag{4.2}$$

The multiplicity of the factor $(\lambda - \lambda_k \mu)$ in $\Delta_{n-1}(\lambda, \mu)$ is according to (2.5) equal to $\alpha_k - \beta_k$ $(k=0, 1, \dots, s)$. Thus we may see that the left side of (4.2) is dicisible by *at least* the $n(\alpha_k - \beta_k)$th power of $(\lambda - \lambda_k \mu)$. But the multiplicity of the factor $(\lambda - \lambda_k \mu)$ on the right side of (4.2) is exactly $(n-1)\alpha_k$. Therefore the inequality

$$n(\alpha_k - \beta_k) \leqq (n-1)\alpha_k \tag{4.3}$$

holds and as $\alpha_k > 0$ $(k=1, \dots, s)$ $\alpha_0 \geqq 0$ and $\beta_k \geqq 0$ $k=0, 1, \dots, s$ are integers, from (4.3) we have immediately

$$1 \leqq \beta_k \leqq \alpha_k, \quad k = 1, \dots, s$$

$$1 \leqq \beta_0 \leqq \alpha_0, \quad \text{if} \quad \alpha_0 > 0.$$

Proof *of theorem 1.* Differentiating (3.4) and evaluating at $\lambda = \lambda_k$ we obtain by (2.4)

$$D_l'(\lambda_k) F(\lambda_k) + D_l(\lambda_k) F'(\lambda_k) = \delta'(\lambda_k) E$$

$$F(\lambda_k) D_l'(\lambda_k) + F'(\lambda_k) D_l(\lambda_k) = \delta'(\delta_k) E.$$

According to the assumption of the theorem these yield simply

$$[D_l'(\lambda_k) F(\lambda_k)]^2 = \delta'(\lambda_k) D_l'(\lambda_k) F(\lambda_k)$$

$$[F(\lambda_k) D_l'(\lambda_k)]^2 = \delta'(\lambda_k) F(\lambda_k) D_l'(\lambda_k).$$

After dividing both sides by $\delta'(\lambda_k)^2 \neq 0$, we obtain assertion (ii) of the theorem.

*) This criterion is formulated only in the case $|A_0| \neq 0$ (cf. e.g. [7] pp. 46.), but it is obviously valid also for our definition of the simple polynomial matrix. We have only to extend the criterion to the elementary divisors belonging to the infinite eigenvalue λ_0.

In order to prove assertion (i) we apply again (2.10) to the matrix $D_l(\lambda, \mu)$ and divide both sides by $\Delta_{n-1}(\lambda, \mu)^r$, then

$$(4.4)\qquad \left|F(\lambda, \mu)[\omega|\tau]\right| = (-1)^{\omega+\tau} \frac{\delta(\lambda, \mu)^{r-1}}{\Delta_{n-1}(\lambda, \mu)} \left|D_l(\lambda, \mu)(\omega|\tau)\right|.$$

As $\delta(\lambda, \mu)$ has only simple roots, we have on the right side of (4.4)

$$\left.\frac{\delta(\lambda, \mu)^{r-1}}{\Delta_{n-1}(\lambda, \mu)}\right|_{\lambda=\lambda_k} \begin{cases} = 0 & \text{if} \quad r > \alpha_k \\ \neq 0 & \text{if} \quad r = \alpha_k. \end{cases}$$

But $\varrho[D_l(\lambda_k)] \geqq n - \alpha_k$, hence for suitable $D_l(\lambda_k)(\omega|\tau)$ the right side of (4.4) differs from zero. Thus

$$\varrho(P_k) = \varrho[F(\lambda_k)] = \alpha_k \quad k = 0, 1, \ldots, s$$

and by (ii) also (i) holds.

Proof *of theorem 2.* According to the assumption the matrix $D_l(\lambda, \mu)$ is not simple. Then we have for at least one eigenvalue λ_k

$$\varrho[D_l(\lambda_k)] > n - \alpha_k$$

and on account of (3.4)

$$(4.5)\qquad \varrho[F(\lambda_k)] < \alpha_k.$$

Let us consider now (4.4) for $r = \alpha_k$ and evaluate at $\lambda = \lambda_k$. The left side of (4.5) is due to (4.6) certainly equal to zero. Thus on the right side of (4.4)

$$\left.\frac{\delta(\lambda, \mu)^{\alpha_k - 1}}{\Delta_{n-1}(\lambda, \mu)}\right|_{\lambda=\lambda_k} = 0$$

must hold. Hence $(\alpha_k - 1)\beta_k > \alpha_k - \beta_k$, i.e.

$$\beta_k > 1,$$

which proves the theorem.

References

[1] K. Weierstrass, Zur Theorie der bilinearen und quadratischen Formen, *Monatsh. Akad. Wiss.*, Berlin, (1895), 310—338.

[2] L., Kronecker, Algebraische Reduktion der Scharen bilinearen Formen, *Sitz-Ber. Akad. Wiss., Phys-math. Kl.*, Berlin (1890), 763—775.
Über Scharen von Quadratischen und bilinearen Formen, *Monatsber. d. Königlich Preussischen Akad. Wiss.*, Berlin (1874), S: 59—76, 149—156, 206—232.

[3] M. M. Flood, Division by non-singular matrix polynomials, *Ann. of Math.*, (2) **36** (1935) 859—869.

[4] H. W. Turnbull—A. C. Aitken, An Introduction to the Theory of Canonical Matrices, *London*, 1932. Chapt. III.

[5] FRAZER R. A. — DUNCAN W. J. Elementary Matrices, *Cambridge*, 1938. Chapt. II.
[6] F. R. GANTMACHER, Matrizenrechnung, *Berlin* 1958. *Kapitel* VI., VII., XII.
[7] LANCASTER P., Lambda—Matrices and Vibrating Systems, *Oxford—London*, 1966.
[8] S. BARNETT, Regular polynomial matrices having relative prime determinants, *Proc. Cambridge Phil. Soc.*, **65** (1969), 585—590.
[9] S. BARNETT, Degrees of greatest common divisors of inveriant factors of two regular polynomial matrices, *Proc. Cambridge Phil. Soc.*, **66** (1969), 241—245.
[10] F. R. GANTMACHER—M. G. KREIN, Oszillationsmatrizen, Oszillationskerne und kleine Schwingungen mechanischer Systeme, *Berlin*, 1960.
[11] A. LEE, On regular polynomial matrices III. Nonsimple matrix pencils, *To appear*.

(Received December 28, 1969; in revised form July 21, 1974.)