On the Equivalence of Two Singular Matrix Pencils

By J. WILLIAMSON, The Johns Hopkins University.

(*Received* 18*th October*, 1935. *Read* 1*st November*, 1935.)

In a recent paper[1] Turnbull, discussing a rational method for the reduction of a singular matrix pencil to canonical form, has shown how the lowest row, or column, minimal index may be determined directly without reducing the pencil to canonical form. It is the purpose of this note to show how all such indices may be determined, and at the same time to give conditions, somewhat simpler than the usual ones, for the equivalence of two matrix pencils.

For simplicity we adopt, as far as possible, Turnbull's notation and, for convenience of the reader, in § 1 give an abstract of the first part of his paper.

1. Let

(1) $$\Lambda = rA + sB = (ra_{ij} + sb_{ij}), \quad i=1, 2, .., n; j=1, 2, .., n'$$

be a matrix pencil, where the elements a_{ij}, b_{ij} all belong to the field K, while r and s are independent variables. We exclude the case in which A is a scalar multiple of B. If P and Q are two non-singular matrices, with elements in K, of orders n and n' respectively, such that

(2) $$P \Lambda Q = \Lambda_1 = rC + sD,$$

the pencils Λ and Λ_1 are said to be equivalent in K or briefly "to be equivalent."

Let ρ be the rank of Λ in r and s; *i.e.* every determinant of order $\rho + 1$ of Λ vanishes identically in r and s while at least one determinant of order ρ does not. If $\mu = n - \rho$ and $\mu' = n' - \rho$, then $\mu \geqq 0$ and $\mu' \geqq 0$. If one of the integers μ or μ', say $\mu > 0$, the pencil (1) is singular. By an application of Smith's Theorem[2],

[1] H. W. Turnbull, "On the reduction of singular matrix pencils," *Proc. Edin. Math. Soc.*, 4 (1935), 67.

[2] Turnbull & Aitken, *Canonical Matrices*, p. 23.

Turnbull shows that the rows of Λ satisfy exactly μ independent relations of the type

$$(3) \qquad \sum_{i=1}^{n} \phi_i \, \text{row}_i = 0,$$

where ϕ_i is a polynomial in r and s with coefficients in K.

The relation (3) may be written in the form of a vector equation,

$$(4) \qquad \theta \Lambda = [\phi_1, \phi_2, \ldots, \phi_n] (rA + sB) = 0,$$

identically in r and s. The μ vectors θ are independent in the sense that they satisfy no homogeneous linear equation with coefficients, which are polynomials in r and s. We may further suppose that the components ϕ_i of θ are homogeneous polynomials in r and s of the same degree and that they have no common factor. We thus have μ vectors θ_i, $i = 1, 2, \ldots, \mu$, which *annihilate* Λ, that is satisfy the equation

$$(5) \qquad \theta_i \Lambda = 0, \qquad i = 1, 2, \ldots, \mu.$$

We arrange the vectors θ_i in ascending degree so that, if θ_i is of degree m_i,

$$(6) \qquad 0 \leqq m_1 \leqq m_2 \leqq, \ldots, \leqq m_\mu.$$

This set of integers (6) is the set of Kronecker minimal indices of row dependence characterising the singular pencil Λ. A like set $\{m'_i\}$, μ' in number, refers to column dependence and is obtained by considering the column vectors ψ, which satisfy the equation $\Lambda \psi = 0$. These sets $\{m_i\}$ and $\{m'_i\}$ are invariant under transformations of type (2) and, together with the invariant factors, completely characterise the pencil Λ. Let

$$(7) \qquad M_1 = [A \quad B], \quad M_2 = \begin{bmatrix} A & B & 0 \\ 0 & A & B \end{bmatrix}, \quad M_3 = \begin{bmatrix} A & B & 0 & 0 \\ 0 & A & B & 0 \\ 0 & 0 & A & B \end{bmatrix}, \ldots.$$

$$N_1 = \begin{bmatrix} A \\ B \end{bmatrix}, \qquad N_2 = \begin{bmatrix} A & 0 \\ B & A \\ 0 & B \end{bmatrix}, \qquad N_3 = \begin{bmatrix} A & 0 & 0 \\ B & A & 0 \\ 0 & B & A \\ 0 & 0 & B \end{bmatrix}, \ldots.$$

so that M_i is a matrix of in rows and $(i+1)n'$ columns and N_i a matrix of $(i+1)n$ rows and in' columns. Let ρ_i and ρ'_i denote the ranks of M_i and N_i respectively. Then the integers

$$(8) \qquad \mu_i = in - \rho_i, \qquad \mu'_i = in' - \rho'_i$$

are all greater than or equal to zero.

2. Let V be a vector with elements in K which annihilates M_k, *i.e.* which satisfies the equation

$$VM_k = 0.$$

The vector V is of dimension nk and may be written in the form

$$V = (v_1, v_2, \ldots, v_k)$$

where each component vector v_i is of dimension n. It now follows that

(9) $$v_1 A = 0, \quad v_k B = 0, \quad v_i B + v_{i+1} A = 0, \quad i = 1, 2, \ldots, k-1.$$

As a consequence of (9) the vector

(10) $$\psi = v_1 r^{k-1} + v_2 r^{k-2} s + \ldots + v_k s^{k-1}$$

satisfies $\psi(rA + sB) = 0$ identically in r and s. We say that ψ is of *order* $k - 1$ and corresponds to the vector V. If ψ can be written as a linear combination of vectors of order $< k - 1$, we say that ψ is *reducible*. If ψ cannot be so expressed, we say that ψ is *irreducible*. We now prove the lemma:

LEMMA I. *The vector ψ in* (10) *corresponding to V is reducible, if and only if the first component vector v_1 of V is the first component of a vector U which annihilates M_{k-1}.*

Proof. Let the first component vector of U be v_1. Then, if 0 denotes the zero vector of dimension n, the vector $V - (U, 0) = (0, W)$ annihilates M_k, so that W annihilates M_{k-1}. If ψ_1 and ψ_2 are the vectors corresponding to U and W respectively,

(11) $$\psi = r\psi_1 + s\psi_2, \text{ and } \psi \text{ is reducible.}$$

Conversely, if ψ is reducible, ψ has the form (11), where ψ_1 and ψ_2 are of order $k - 2$. If U is the vector corresponding to ψ_1, which annihilates M_{k-1}, the first component of V is the same as the first component of U and our lemma is proved.

As in § 1, by an application of Smith's theorem, there exist exactly μ_k linearly independent linear relations connecting the rows of M_k. That is, there exist exactly μ_k linearly independent vectors V_a with elements in K, which annihilate M_k, *i.e.* which satisfy

(12) $$V_a M_k = 0, \qquad a = 1, 2, \ldots, \mu_k.$$

Moreover, every vector, which annihilates M_k, is a linear combination of the vectors V_a of (12), so that the vectors V_a form a basis for the

https://doi.org/10.1017/S0013091500027486 Published online by Cambridge University Press

set of all vectors annihilating M_k. Each vector V_α in (12) is of dimension kn and may be written as

$$V_\alpha = [v_{\alpha 1}, v_{\alpha 2}, \ldots, v_{\alpha k}],$$

where each component vector $v_{\alpha j}$ is of dimension n.

Let the vectors V_α be so arranged that their first component vectors $v_{11}, v_{21}, v_{31}, \ldots, v_{\lambda_k 1}$, are linearly independent, but that

$$v_{j1} = \sum_{\alpha=1}^{\lambda_k} c_{j\alpha} v_{\alpha 1}, \qquad j = \lambda_k + 1, \ldots, \mu_k.$$

Now the vectors

$$U_\alpha = V_\alpha, \qquad \alpha = 1, 2, \ldots, \lambda_k,$$

$$U_\alpha = V_\alpha - \sum_{\beta=1}^{\lambda_k} c_{\alpha\beta} V_\beta, \quad \alpha = \lambda_k + 1, \ldots, \mu_k,$$

are linearly independent and hence also form a basis for the set of vectors which annihilate M_k.

We may therefore suppose that the vectors V_α of (12) are such that, as regards their first component vectors,

(13) $\qquad v_{11}, v_{21}, \ldots, v_{\lambda_k 1}$ are independent, $v_{\alpha 1} = 0$, $\alpha > \lambda_k$.

If 0 denotes the zero vector of dimension n, each of the λ_k vectors W_β and the μ_k vectors X_α, defined by

$$(14) \qquad \begin{cases} W_\beta = (V_\beta, 0), & \beta = 1, 2, \ldots, \lambda_k, \\ X_\alpha = (0, V_\alpha), & \alpha = 1, 2, \ldots, \mu_k, \end{cases}$$

is of dimension $(k+1)n$ and annihilates M_{k+1}.

Lemma 2. *The $\mu_k + \lambda_k$ vectors* (14) *are linearly independent with respect to K.*

Proof. Let

$$\sum_{\beta=1}^{\lambda_k} c_\beta W_\beta + \sum_{\alpha=1}^{\mu_k} d_\alpha X_\alpha = 0;$$

then, since the first component vector of X_α is zero, $\sum_{\beta=1}^{\lambda_k} c_\beta v_{\beta 1} = 0$, and by (13) $c_\beta = 0, \beta = 1, 2, \ldots, \lambda_k$. Hence $\sum_{\alpha=1}^{\mu_k} d_\alpha X_\alpha = 0$ and consequently $\sum_{\alpha=1}^{\mu_k} d_\alpha V_\alpha = 0$. As the μ_k vectors V_α are linearly independent this last equation implies $d_\alpha = 0$, $\alpha = 1, 2, \ldots, \mu_k$. Consequently the lemma is proved.

But there exist μ_{k+1} linearly independent vectors which annihilate M_{k+1}. Accordingly a non-negative integer σ_k must exist such that

(15) $$\mu_{k+1} = \mu_k + \lambda_k + \sigma_k, \qquad \sigma_k \geqq 0.$$

Hence there also exist σ_k vectors

(16) $$Y_\gamma = (y_{\gamma 1}, y_{\gamma 2}, \ldots, y_{\gamma, k+1}), \qquad \gamma = 1, 2, \ldots, \sigma_k,$$

which annihilate M_{k+1} and such that the μ_{k+1} vectors W_β, X_α, Y_γ are linearly independent. (If $\sigma_k = 0$, no vector Y_γ appears).

LEMMA 3. *The first component vectors $v_{\beta 1}$, $y_{\gamma 1}$ of W_β and Y_γ respectively are linearly independent with respect to K, where $\beta = 1, 2, \ldots, \lambda_k$, $\gamma = 1, 2, \ldots, \sigma_k$.*

Proof. Let

$$\sum_{\beta=1}^{\lambda_k} c_\beta v_{\beta 1} + \sum_{\gamma=1}^{\sigma_k} d_\gamma y_{\gamma 1} = 0.$$

Then

(17) $$\sum_{\beta=1}^{\lambda_k} c_\beta W_\beta + \sum_{\gamma=1}^{\sigma_k} d_\gamma Y_\gamma = (0, G),$$

where, since the vector on the left of (17) annihilates M_{k+1}, the vector G annihilates M_k. By there mark that follows (12), such a vector must take the form $G = \sum_{\alpha=1}^{\mu_k} f_\alpha V_\alpha$, so that $(0, G) = \sum_{\alpha=1}^{\mu_k} f_\alpha X_\alpha$. Therefore, by (17),

$$\sum_{\beta=1}^{\lambda_k} c_\beta W_\beta + \sum_{\gamma=1}^{\sigma_k} d_\gamma Y_\gamma - \sum_{\alpha=1}^{\mu_k} f_\alpha X_\alpha = 0.$$

Since the vectors $W_\beta, X_\alpha, Y_\gamma$ are linearly independent, $c_\beta = d_\gamma = f_\alpha = 0$, $\beta = 1, 2, \ldots, \lambda_k$; $\gamma = 1, 2, \ldots, \sigma_k$; $\alpha = 1, 2, \ldots, \mu_k$. Accordingly the lemma is proved.

The formula corresponding to (15) when k is replaced by $k+1$ is

(18) $$\mu_{k+2} = \mu_{k+1} + \lambda_{k+1} + \sigma_{k+1}, \quad \sigma_{k+1} \geqq 0,$$

where by definition λ_{k+1} is the maximum number of vectors among those which annihilate M_{k+1}, whose first component vectors are linearly independent. But a vector U, which annihilates M_{k+1}, is a linear combination of the vectors W_β, X_α, Y_γ, and accordingly its first component vector u_1 is a linear combination of the vectors

$v_{\beta 1}$ and $y_{\gamma 1}$, where $\beta = 1, 2, \ldots, \lambda_k$, $\gamma = 1, 2, \ldots, \sigma_k$. Since, by lemma 3, these $\lambda_k + \sigma_k$ vectors are linearly independent, it follows that

$$\lambda_{k+1} = \lambda_k + \sigma_k;$$

and accordingly (18) becomes

$$\mu_{k+2} = \mu_{k+1} + \lambda_k + \sigma_k + \sigma_{k+1}; \tag{19}$$

so from (15) we have

$$\sigma_{k+1} = \mu_{k+2} + \mu_k - 2\mu_{k+1}, \qquad k = 1, 2, \ldots, . \tag{20}$$

It is easily seen that formula (20) also holds for $k = -1, 0$, if μ_{-1} and μ_0 are defined to be zero.

The vectors ψ_γ, $\gamma = 1, 2, \ldots, \sigma_k$, which correspond to the vectors Y_γ in (16), annihilate $rA + sB$ and are of order k. As a consequence of lemma 3, no linear combination of the vectors $y_{\gamma 1}$ is equal to a linear combination of the vectors $v_{\beta 1}$. Hence no linear combination of the first components of the vectors Y_γ is equal to the first component of a vector which annihilates M_k. Consequently, by lemma 1, all linear combinations, with constant coefficients, of the σ_k vectors ψ_γ are irreducible. Moreover, if ϕ is any vector of order k which annihilates Λ, and U is the corresponding vector which annihilates M_{k+1}, then

$$U = \sum_{\beta=1}^{\lambda_k} b_\beta W_\beta + \sum_{\alpha=1}^{\mu_k} c_\alpha X_\alpha + \sum_{\gamma=1}^{\sigma_k} d_\gamma Y_\gamma.$$

It follows by an argument similar to the above that the vector $\phi - \sum_{\gamma=1}^{\sigma_k} d_\gamma \psi_\gamma$ is reducible. Hence the σ_k vectors ψ_γ form a maximal set of vectors of order k, with the property that no linear combination of them, with constant coefficients, is reducible. Consequently, the σ_k vectors ψ_γ form a maximal set of independent vectors θ satisfying (5) of index k. Accordingly, we have proved

THEOREM I. *The number of minimal row indices* (6) *which have the value k is exactly $\mu_{k+1} + \mu_{k-1} - 2\mu_k$.*

COR. 1. *If μ_{k+1} is the first of the sequence $\mu_1, \mu_2, \mu_3, \ldots$ which is different from* 0, *then k is the value of the smallest minimal index of row dependence and μ_{k+1} is the number of such indices which are equal.*

This is Turnbull's Theorem 2.

COR. 2. *The number of minimal column indices* (6) *which have the value* k *is exactly*

$$\mu'_{k+1} + \mu'_{k-1} - 2\mu'_k.$$

By (20) and (8) we have

$$\begin{aligned}\sigma_{k+1} &= (k+2)\,n - \rho_{k+2} + kn - \rho_k - 2\,(k+1)\,n + 2\rho_{k+1} \\ &= 2\rho_{k+1} - \rho_{k+2} - \rho_k\end{aligned}$$

and similarly the analogous formula

$$\sigma'_{k+1} = 2\rho'_{k+1} - \rho'_{k+2} - \rho'_k.$$

Conversely, the ρ_k can now be expressed in terms of the σ_k. Accordingly we have proved

COR. 3. *The minimal indices* $\{m_i\}$ *and* $\{m'_i\}$ *are completely determined by the ranks of the matrices* M_k *and* N_k. *Conversely the ranks* ρ_k, ρ'_k *of the matrices* M_k *and* N_k *are completely determined by the minimal indices* $\{m_i\}$ *and* $\{m'_i\}$.

In order to determine the minimal indices it is sufficient to form the sequence of the positive or zero integers

(21) $$\mu_1, \quad \mu_2 - \mu_1, \quad \mu_3 - \mu_2, \quad \mu_4 - \mu_3, \; \ldots .$$

Since $$\mu_{k+1} - \mu_k - (\mu_k - \mu_{k+1}) = \sigma_k \geqq 0,$$

each term in the sequence is greater than or equal to its predecessor. The difference between the $(k+1)$th term and the kth term of the sequence (21) is the number of minimal row indices which have the value k. Since

$$\sigma_0 + \sigma_1 + \sigma_2 + \ldots + \sigma_k = \mu_{k+1} - \mu_k,$$

the $(k+1)$th term of (21) gives the total number of independent vectors of order $\leqq k$, which annihilate Λ. But the maximum number of such vectors is ρ, where ρ is the rank of Λ in r and s. Accordingly

$$\mu_{k+1} - \mu_k \leqq \rho.$$

Hence in forming the sequence (21), if the $(k+1)$th term has the value ρ all succeeding terms will have the value ρ and we need calculate no further.

By a proof exactly similar to that of Turnbull's Theorem 3, if $Y = (y_1, y_2, \ldots, y_{k+1})$ is one of the vectors (16), the component vectors $y_1, y_2, \ldots, y_{k+1}$ are linearly independent. Since each y_i is of dimension n, it follows that $k < n$ and consequently that

(22) $$\sigma_k = 0, \qquad k > n-1.$$

Hence in forming the sequence (21) we need never proceed beyond the nth term; and in corollary 3, k takes only the values 1, 2, ..., n and k' the values 1, 2, ..., n'.

Let $\Lambda_1 = rC + sD$ be a second pencil of the same order as Λ and let

$$I_k, \ k = 1, 2, \ldots, n; \quad J_{k'}, \ k' = 1, 2, \ldots, n',$$

be the matrices for this pencil corresponding to M_k and $N_{k'}$ respectively. Since a necessary and sufficient condition that the two pencils Λ and Λ_1 be equivalent in the sense of (2), is that Λ and Λ_1 have the same invariant factors and the same minimal indices[1], as a consequence of corollary 3 we have

THEOREM 2. *A necessary and sufficient condition that two pencils $rA + sB$ and $rC + sD$, with elements in a field K be equivalent in K is*:

(a) *that the invariant factors of $rA + sB$ be the same as the invariant factors of $rC + sD$,*

(b) *that the rank of M_k be the same as the rank of I_k, $k = 1, 2, \ldots, n$,*

(c) *that the rank of $N_{k'}$ be the same as the rank of $J_{k'}$, $k' = 1, 2, \ldots, n'$.*

[1] Turnbull and Aitken, *Op. cit.*, p. 129.