

On the Reduction of Singular Matrix Pencils

By H. W. TURNBULL.

(Received 13th March, 1934. Read 7th December, 1934.)

Introduction.

The following rational method of dealing with the reduction of a singular matrix pencil to canonical form has certain advantages. It is based on the principle of vector chains, the length of the chain determining a minimal index. This treatment is analogous to that employed by Dr A. C. Aitken and the author in *Canonical Matrices* (1932) 45-57, for the nonsingular case. In Theorems 1 and 2 tests are explicitly given for determining the minimal indices. Theorem 2 gives a method of discovering the lowest row (or column) minimal index. Theoretically it should be possible to state a corresponding theorem for each of these indices, not necessarily the lowest, and prior to any reduction of the pencil. This extension still awaits solution.

Theorem 3 is logically equivalent to the arguments used by Kronecker (who was the first to discuss the singular case, *Berlin Sitzungsberichte* (1890), 1375 and (1891), 9, 33) and subsequently by Dickson (*Trans. American Math. Soc.*, 29 (1927), 239-253). For a geometrical treatment see Segre, *Atti Acc. Torino*, 19 (1884), 878.

§ 1. Let

$$\Lambda = rA + sB = [ra_{ij} + sb_{ij}] \quad (1)$$

be a matrix pencil, where the elements a_{ij} and b_{ij} all belong to a field \mathcal{F} , while r, s are independent variables. Each matrix A, B, Λ is assumed to have n rows and n' columns, while neither A nor B is a scalar multiple of the other. It is proposed to reduce Λ to a canonical form

$$P \wedge Q, \quad |P| \neq 0, \quad |Q| \neq 0, \quad (2)$$

where P and Q are nonsingular constant matrices with elements in \mathcal{F} , P having n rows and columns, and Q having n' .

Let ρ be the rank of Λ in r and s : that is, let ρ be the highest order among the minor determinants of Λ which do not vanish

identically for all r and s . Then obviously there exist nonnegative integers μ and μ' such that

$$\mu = n - \rho \geq 0, \quad \mu' = n' - \rho \geq 0. \quad (3)$$

Two cases arise, the singular and the nonsingular. In the latter both μ and μ' are zero: in the former at least one, μ say, is nonzero. It will be proved that in the singular case μ is the number of ways, linearly independent in \mathcal{F} , in which the rows of Λ are linearly related, while μ' similarly relates the columns.

Such a row relation $\sum_{i=1}^n \theta_i \text{row}_i = 0$ can be conveniently written as a matrix product

$$\theta\Lambda \equiv [\theta_1, \theta_2, \dots, \theta_n](rA + sB) = 0 \quad (4)$$

as appears at once when written out in full. The coefficients θ_i of this relation here appear as the n components of a row vector θ which is said to annihilate Λ .

By Smith's Theorem¹ the matrix Λ can be reduced in \mathcal{F} to a diagonal form D such that

$$H\Lambda K = \text{diag}(E_1, E_2, \dots, E_\rho, 0, \dots, 0) = D \quad (5)$$

where H and K are nonsingular matrices each of whose elements are homogeneous polynomials in r, s , divided possibly by a power of s , whereas the determinants $|H|$ and $|K|$ are independent of r . The ρ nonzero elements E are the invariant factors. This allows relation (4) to take a simpler form: thus $0 = \theta\Lambda = \theta H^{-1} DK^{-1}$. Hence

$$\phi D = 0, \quad \text{where } \phi = \theta H^{-1}. \quad (6)$$

This ϕ , so found, is also a row vector: and clearly it can annihilate D if and only if its first ρ components are zero. Thus

$$\phi = [0, 0, \dots, 0, \phi_{\rho+1}, \phi_{\rho+2}, \dots, \phi_n] \quad (7)$$

where the last $\mu (= n - \rho)$ components are arbitrary functions of r and s . Let the unit matrix of n rows and columns be written

$$I = \{i_1, i_2, \dots, i_n\} \quad (8)$$

where $i_1 = [1, 0, \dots, 0]$, $i_2 = [0, 1, \dots, 0]$, \dots , $i_n = [0, \dots, 0, 1]$. Then (7) can be written

$$\phi = \phi_{\rho+1} i_{\rho+1} + \phi_{\rho+2} i_{\rho+2} + \dots + \phi_n i_n \quad (9)$$

¹ Turnbull and Aitken: *Canonical Matrices* (1932), 23.

which shows that the most general condition (4) is a consequence of μ linearly independent conditions

$$i_{\rho+1}D = 0, \quad i_{\rho+2}D = 0, \quad \dots, \quad i_n D = 0. \tag{10}$$

Each of these conditions is of the form $i_n H \Lambda K = 0$, from which the nonsingular K and any scalar common factor of the components can be deleted and the powers of s in the denominators cleared, the result being called $\theta \Lambda = 0$. This proves the following theorem.

THEOREM I. *If the pencil $\Lambda = rA + sB$ possesses row dependence, there are exactly μ distinct conditions*

$$\theta \Lambda = 0$$

where θ is a row-vector whose components are homogeneous polynomials in r and s with coefficients in \mathcal{F} , $(n - \mu)$ being the rank of Λ in r and s .

Correlatively: there are exactly μ' distinct relations of column dependence

$$\Lambda \theta' = 0$$

where θ' is a column vector homogeneous in r and s .

There is no necessary connection between θ and θ' .

§ 2. Let the relations just found be arranged in ascending degree in r and s , as

$$\theta_1 \Lambda = 0, \quad \theta_2 \Lambda = 0, \quad \dots, \quad \theta_\mu \Lambda = 0, \tag{1}$$

where the degree of the vector θ_i is m_i , so that

$$0 \leq m_1 \leq m_2 \leq \dots \leq m_\mu. \tag{2}$$

These are in fact the Kronecker minimal indices of row dependence characterising a singular pencil. A like set $[m_i']$, μ' in number, refers to column dependence. These sets $[m_i]$ and $[m_i']$ are numerical invariants under nonsingular transformation from Λ to the type $P \Lambda Q$ above, and also under nonsingular linear transformation from r, s to r', s' . The proof of these statements is immediate: in either case we have only to suppose the contrary and then obtain an identity of less than minimal order by applying the reciprocal transformation; which involves an absurdity.

These minimal indices together with the set of invariant factors E_1, E_2, \dots, E_ρ of § 1 (5) completely characterise the pencil Λ under such transformations, as Kronecker originally proved. A direct method will now be given for finding these minimal indices.

§ 3. Consider the following matrices

$$M_1 = [A, B], M_2 = \begin{bmatrix} A & B & \cdot \\ \cdot & A & B \end{bmatrix}, M_3 = \begin{bmatrix} A & B & \cdot & \cdot \\ \cdot & A & B & \cdot \\ \cdot & \cdot & A & B \end{bmatrix}, \dots,$$

$$N_1 = \begin{bmatrix} A \\ B \end{bmatrix}, N_2 = \begin{bmatrix} A & \cdot \\ B & A \\ \cdot & B \end{bmatrix}, N_3 = \begin{bmatrix} A & \cdot & \cdot \\ B & A & \cdot \\ \cdot & B & A \\ \cdot & \cdot & B \end{bmatrix}, \dots, \tag{1}$$

the M consisting of $n, 2n, 3n, \dots$ rows, and the N of $n', 2n', 3n', \dots$ columns respectively. Let ρ_i, ρ_j' denote their respective ranks. Then, if $\mu_i = in - \rho_i, \mu_j' = jn' - \rho_j'$, we have

$$\mu_1 = n - \rho_1 \geq 0, \mu_2 = 2n - \rho_2 \geq 0, \mu_3 = 3n - \rho_3 \geq 0, \dots, \tag{2}$$

and

$$\mu_1' = n' - \rho_1' \geq 0, \mu_j' \geq 0. \tag{3}$$

THEOREM 2. *If μ_{m+1} is the first nonzero integer in the sequence μ_1, μ_2, \dots , then m is the value of the smallest minimal index of row dependence, while μ_{m+1} is the number of such indices which are equal. Column dependence is given similarly by $\mu'_{m'+1}$.*

Proof. By Smith's Theorem, if $\mu_1 > 0$, exactly μ_1 distinct relations $\sum_i \lambda_i \text{row}_i = 0$ exist between the rows of M_1 , where the λ_i are $2n'$ constants which are not all zero in \mathcal{F} . On introducing the row vector

$$u = [\lambda_1, \lambda_2, \dots, \lambda_n] \neq 0$$

we may put such a relation in the form of a matrix product

$$u[A, B] = 0, \tag{4}$$

that is $uA = 0, uB = 0$: so that $u[rA + sB] = 0$ for all r, s . But u is a nonzero constant vector in \mathcal{F} . We have therefore secured a minimal index $m_1 = 0$: and the number of such is $\mu_1 (\neq 0)$.

Next if $\mu_1 = 0, \mu_2 > 0$, then a row vector consisting of $2n$ components exists such that

$$[u_1, u_2] \begin{bmatrix} A & B & \cdot \\ \cdot & A & B \end{bmatrix} = 0, \quad [u_1, u_2] \neq 0. \tag{5}$$

Here u_1 is a set of n components, u_2 is a further set, and in all there are $2n$ components. Hence

$$u_1 A = 0, u_1 B + u_2 A = 0, u_2 B = 0,$$

whence

$$[u_1 r + u_2 s] [rA + sB] = 0, \tag{6}$$

for all values of r, s . But this is explicitly a minimal relation $\theta\Lambda = 0$, where $\theta = ru_1 + su_2$ is a vector of index *unity*. There are μ_2 such distinct relations, while there are none of the zero index type, since $u[A, B] \neq 0$ if $\mu_1 = 0$, for all nonzero constant vectors u .

Next if $\mu_1 = 0, \mu_2 = 0, \mu_3 > 0$, then (4) and (5) are impossible, but three vectors u_1, u_2, u_3 each of n components exist such that

$$[u_1, u_2, u_3] M_3 = 0, \quad [u_1, u_2, u_3] \neq 0. \tag{7}$$

Hence

$$u_1 A = 0, \quad u_1 B + u_2 A = 0, \quad u_2 B + u_3 A = 0, \quad u_3 B = 0;$$

that is

$$[u_1 r^2 + u_2 rs + u_3 s^2] [rA + sB] = 0 \tag{8}$$

for all r, s . This gives μ_3 distinct relations of index 2. The general case is now evident: it also applies to columns by means of the

expressions $\begin{bmatrix} A \\ B \end{bmatrix} \{u_1'\}, \begin{bmatrix} A & . \\ B & A \\ . & B \end{bmatrix} \{u_1', u_2'\}$ etc., where $\{u_1', u_2'\}$ denotes

a column of $2n$ elements. This proves the theorem.

It should be remarked that the matrices N are not the transposed of the M : the elements within A (and B) maintain their same relative positions. Also, while the method discovers the initial index m_1 or m_1' it does not at once discover higher indices, if any.

For example:

$$\Lambda = \begin{bmatrix} r & . & . & . \\ s & r & . & . \\ . & s & r & . \\ . & . & s & . \\ s & . & . & r \\ . & . & . & s \end{bmatrix}, \quad A = \begin{bmatrix} 1 & . & . & . \\ . & 1 & . & . \\ . & . & 1 & . \\ . & . & . & . \\ . & . & . & 1 \\ . & . & . & . \end{bmatrix}, \quad B = \begin{bmatrix} . & . & . & . \\ 1 & . & . & . \\ . & 1 & . & . \\ . & . & 1 & . \\ 1 & . & . & . \\ . & . & . & 1 \end{bmatrix}. \tag{9}$$

Here the ranks of M_1, M_2, M_3, M_4 are 6, 12, 16, 20 respectively so that $\mu_3 = 2$ is the first nonzero μ . This implies two minimal indices each equal to $3 - 1 = 2$.

§ 4. In the singular case let m denote the smallest minimal index, so that the corresponding minimal relation $\theta\Lambda = 0$ can be written more explicitly as

$$[u_0 s^m - u_1 s^{m-1} r + u_2 s^{m-2} r^2 - \dots + (-)^m u_m r^m] (rA + sB) = 0 \tag{1}$$

where each of the $(m + 1)$ coefficients u_i is a row vector of n constant components. Since this is identically true for all r, s the coefficients

For the second part of the theorem let $m > 0$. If $u_q A$ is the first of the sequence $u_0 A, u_1 A, \dots$ to be linearly dependent upon its predecessors, let

$$u_q A + \beta_1 u_{q-1} A + \dots + \beta_q u_0 A = 0 \tag{6}$$

or $w_q A = 0$, where $w_q = u_q + \beta_1 u_{q-1} + \dots + \beta_q u_0$. By constructing w_0, w_1, \dots, w_{q-1} analogously to the v_i in (4) it again follows that a chain $0 = w_0 B, \dots, w_q A = 0$ exists, which in turn cannot be shorter than the chain (2). Hence $q \geq m$: and this proves the theorem.

Reduction to Canonical Form.

§ 5. Consider the following matrix relation

$$\begin{bmatrix} u_0 \\ u_1 \\ \vdots \\ u_m \\ P_0 \end{bmatrix} (rA + sB) = \begin{bmatrix} r & & & & \\ s & r & & & \\ & s & & & \\ & & \ddots & & \\ & & & r & \\ & & & s & \\ \hline rA_0 + sB_0 & rA_1 + sB_1 & & & \end{bmatrix} \begin{bmatrix} u_1 B \\ u_2 B \\ \vdots \\ u_m B \\ C_0 \end{bmatrix}, \quad m > 0, \tag{1}$$

or $PA = XQ^{-1}$, where u_0, u_1, \dots, u_m are the first $(m + 1)$ rows of P . Since by Theorem 3 these are linearly independent, P itself may be made a nonsingular constant matrix by choosing a suitable submatrix P_0 for its remaining $(n - m - 1)$ rows. If $m = 0$ then PA has a zero top row and we pass on to consider lower rows. If $m > 0$, then the m rows $u_i B$ are also linearly independent, so that a choice of a further submatrix C_0 is possible, the whole being nonsingular and written Q^{-1} . Owing to the chain of relations $u_i A = u_{i+1} B$ the first $(m + 1)$ rows of the product PA agree with the corresponding rows of XQ^{-1} . For example the i^{th} row gives

$$ru_{i-1} A + su_{i-1} B = su_{i-1} B + ru_i B.$$

Hence (1) is identically true provided that the remaining $(n - m - 1)$ rows of X are identical with those of PAQ . A canonical minimal submatrix L_m , of X has now been semi-isolated, such that

$$X = \begin{bmatrix} L_m & & \cdot \\ rA_0 + sB_0 & rA_1 + sB_1 & \end{bmatrix}, \quad m = m_1, \tag{2}$$

where, for example,

$$L_0 = 0, L_1 = \begin{bmatrix} r \\ s \end{bmatrix}, L_2 = \begin{bmatrix} r & \cdot \\ s & r \\ \cdot & s \end{bmatrix}, \dots, \tag{3}$$

m_1 being the lowest index of row dependence. If X contains a second such index m_2 , then m_2 will in fact be lowest row-index in the sub-matrix $rA_1 + sB_1$; but it will emerge more directly by selecting a new solution v_0 of the equation $uB = 0$, and forming a new chain (since $\mu > 1$)

$$v_0 B = 0, v_1 B = v_0 A, \dots, v_{m_2} A = 0, \quad m_2 \geq m_1. \tag{4}$$

THEOREM 4. *The $(m_2 + 1)$ vectors v are linearly independent of themselves and of the u vectors. Also all the vectors $u_i B, v_i B$ ($i > 0$) are linearly independent.*

Proof. Let v_p be the first such vector which is linearly dependent upon its predecessors u or v . (i) If no vector u with suffix higher than p enters, let the relation be

$$0 = \sum_{r=0}^p \alpha_r v_{p-r} + \sum_{r=0}^p \beta_r u_{p-r}, \quad \alpha_0 = 1.$$

Construct

$$\begin{aligned} w_0 &= \alpha_0 v_0 + \beta_0 u_0, \\ w_1 &= \alpha_0 v_1 + \alpha_1 v_0 + \beta_0 u_1 + \beta_1 u_0, \text{ etc.} \end{aligned} \tag{5}$$

exactly as in Theorem 3. Then the w vectors will form a chain, independent of the u vectors, such that

$$w_0 \neq 0, w_0 B = 0, w_1 B = w_0 A, \dots, w_h A = 0, \tag{6}$$

where $h = p - 1 < m_2$. This contradicts the assumption. The proof that the $u_i B, v_i B$ are unrelated is analogous to that in Theorem 3.

(ii) If however terms u_q ($q > p$) enter the relation, write it as

$$\sum_{r=0}^p \alpha_r v_{p-r} + \sum_{r=0}^p \beta_r u_{p-r} = \gamma_0 u_q + \gamma_1 u_{q+1} + \dots + \gamma_{m_1-q} u_{m_1}, \tag{7}$$

where $\alpha_0 = 1, \gamma_0 \neq 0, q > p$. Let w_p denote either side of this equality and let h be defined by

$$p \leq h = p + m_1 - q < m_1. \tag{8}$$

From $w_p = \gamma_0 u_q + \dots + \gamma_{m_1-1} u_{m_1}$, further vectors $w_{p+1}, w_{p+2}, \dots, w_h$ may be derived by successively adding unity to each suffix of w and u , and deleting terms of suffix exceeding m_1 . The concluding vector is then $w_h = \gamma_0 u_{m_1}$. With those defined by (5) the whole set w_0, w_1, \dots, w_h is then a chain of index less than m_1 , which again involves a contradiction. The proof that the $u_i B, v_i B$ are unrelated is analogous, starting with an identity such as (7) but with A appearing as final factor of each term. Again a chain w_0, \dots, w_h would exist, where w_0, \dots, w_p are defined by (5), and w_{p+1}, \dots, w_h by the rule just given. This proves the theorem.

This theorem allows us to take the v_i to be the first $(m_2 + 1)$ rows in P_0 , and the $v_{i+1}B$ the first m_2 rows in C_0 . The result is

$$\begin{bmatrix} u_i \\ v_i \\ \cdot \\ \cdot \\ \cdot \end{bmatrix} (rA + sB) = \begin{bmatrix} L_{m_1} & \cdot & \cdot \\ \cdot & L_{m_2} & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \begin{bmatrix} u_{i+1} B \\ v_{i+1} B \\ \cdot \\ \cdot \\ \cdot \end{bmatrix}$$

where two canonical minimal submatrices have now been semi-isolated. No new feature arises in further steps until all μ submatrices L_m have been semi-isolated. Among themselves they are completely isolated in the form

$$L = \text{diag} (L_{m_1}, L_{m_2}, \dots, L_{m_\mu}).$$

This exhausts all possible row dependence. Column dependence is then sought in μ' possible ways, but owing to the isolation of each L_{m_i} in its own row, such column dependence is independent of columns occupied by L . The result is a submatrix

$$L' = \text{diag} (L'_{m'_1}, \dots, L'_{m'_\mu}),$$

and any further submatrix X_0 not lying in the rows or columns of L and L' must be nonsingular. This can be reduced to rational or classical canonical form S say, and finally all remaining nonzero elements other than those of L , L' and S can be removed by the methods earlier explained.¹

§ 6. Also, for directly obtaining the rational form of the nonsingular portions of the pencil, vector chains of the same general type $u_{i+1}A = u_iB$ may be formed but for which $u_0 \neq 0$, $u_0A \neq 0$, $|A| \neq 0$. They must then be examined in descending order of their length, as in the rational case² for the collineatory group. The method is sufficiently illustrated by the following example:

$$\begin{bmatrix} u_0 \\ u_1 \\ u_2 \end{bmatrix} (rA + sB) = \begin{bmatrix} r & s & \\ & r & s \\ \alpha_0 s & \alpha_1 s & r + \alpha_2 s \end{bmatrix} \begin{bmatrix} u_0 A \\ u_1 A \\ u_2 A \end{bmatrix}.$$

In this example $u_3 = \alpha_0 u_0 + \alpha_1 u_1 + \alpha_2 u_2$ is the first of such a chain to be related to its predecessors.

¹ *Canonical Matrices* (1932), 127-8.

² *Canonical Matrices*, 49.

It is to be noted that in the example of §3(9) the chain appearing in the first four rows of Λ is not a true minimal. The failure is due to the presence of the lower element s in the first column. Every vector satisfying $u_0 B = 0$ must be of the form

$$[\alpha, \beta, 0, 0, -\beta, 0],$$

where α, β are arbitrary constants. Taking $\alpha = 1, \beta = 0$, the shortest chain is obtained as

$$u_0 = [1, 0, 0, 0, 0, 0], \quad u_1 = [0, 0, 0, 0, 1, 0], \quad u_2 = [0, 0, 0, 0, 0, 1]$$

where $u_0 B = u_0 A - u_1 B = u_1 A - u_2 B = u_2 A = 0$.

It may also be noted that the same method will furnish every submatrix of type

$$R_e = \begin{bmatrix} r & & & & & \\ s & r & & & & \\ & & \ddots & & & \\ & & & s & r & \end{bmatrix}, \quad |R_e| = r^e$$

due to a zero latent root, and belonging to the nonsingular core. All such are found according to ascending value of e by use of every vector u_0 for which $u_0 B = 0$ but which does not lead to a minimal chain. A modified chain now appears, following the same law except that it terminates abruptly with u_{e-1} at a point where it is impossible to satisfy the equation $u_e B = u_{e-1} A$ by any vector u_e .

