Equivalence of Partitioned Matrices

Robert B. Feinberg

Institute for Basic Standards, National Bureau of Standards, Washington, D.C. 20234

(November 18, 1975)

It is shown that if \( M = \begin{bmatrix} A & T \\ 0 & B \end{bmatrix} \) is a partitioned matrix over a principal ideal domain \( R \) such that the matrices \( A \) and \( B \) are both square, then \( M \) is equivalent to \( A + B \) (\( = \)) the matrix equation \( T = AY + XB \) is solvable. The result is generalized to treat the case when

\[
M = \begin{bmatrix}
M_{11} & M_{12} & \ldots & M_{1t} \\
0 & M_{22} & \ldots & \cdot \\
\cdot & \cdot & \ddots & \cdot \\
0 & \ldots & 0 & M_{tt}
\end{bmatrix},
\]

where each \( M_{ij} \) is square.

Key words: Determinantal divisors; equivalence; matrix equation; partitioned matrix; Smith normal form.

Let \( R \) be a principal ideal domain and let \( R_{mn} \) denote the collection of \( m \times n \) matrices over \( R \). According to Theorem 2 of [2], if \( A \in R_{rr} \), \( B \in R_{ss} \), and \( (\det A, \det B) = 1 \), then for any \( T \in R_{rs} \),

\[
S \begin{bmatrix} A & T \\ 0 & B \end{bmatrix} = S \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix},
\]

where \( S(M) \) denotes the Smith normal form of a matrix \( M \). The proof consists essentially of establishing two elementary propositions:

(i) For arbitrary \( A \in R_{rr} \), \( B \in R_{ss} \), and \( T \in R_{rs} \), if the matrix equation (*) \( T = AY + XB \) has a solution \( X \), \( Y \in R_{rs} \), then

\[
\begin{bmatrix} A & T \\ 0 & B \end{bmatrix} \tilde{E} \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix},
\]

where \( \tilde{E} \) denotes equivalence of matrices (ii). In the case when \( (\det A, \det B) = 1 \); (*) is always solvable.

The central result of this note (Theorem 1) provides a converse to (i), namely that if

\[
\begin{bmatrix} A & T \\ 0 & B \end{bmatrix} \tilde{E} \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix},
\]

*This paper was prepared while the author was a National Academy of Sciences-National Research Council Postdoctoral Research Associate at the National Bureau of Standards, Washington, D.C. 20234. He would like to thank Dr. Morris Newman for several helpful conversations.

**Present address: Clarkson College of Technology, Potsdam, New York 13676.
then (*) must be solvable. We generalize this result to the case when

\[
M = \begin{bmatrix}
M_{11} & M_{12} & \cdots & M_{1t} \\
0 & M_{22} & \cdots & \\
\vdots & \vdots & \ddots & \\
0 & 0 & \cdots & M_{tt}
\end{bmatrix},
\]

where each \(M_{ii}\) is square, and also derive some corollaries.

Subsequent to completion of this work, the author discovered that Theorem 1 had been established in [3] in the case when \(R\) is the domain of polynomials over a field. The proof there carries over immediately to the case when \(R\) is an arbitrary P.I.D., and is similar to the proof of Theorem 1 presented here. The generalization of Theorem 1 is not developed there, however.

In the sequel \(R'_s\) will denote the group of unimodular \(n \times n\) matrices over \(R\), \(I_n\) will denote the identity matrix of order \(n\), \(I\) will denote an identity matrix of unspecified order, \(0_{mn}\) will denote the 0 matrix of order \(m \times n\), \(0_{m}\) will denote \(0_{mm}\), and \(d_k[M]\) will denote the \(k\)th determinantal divisor of the matrix \(M\).

See [1] for a good general reference on matrices over a P.I.D.

**Theorem 1:** Let \(R\) be a P.I.D., \(A \in R_{rs}, B \in R_{ss},\) and \(T \in R_{rs}\). Then

\[
\begin{bmatrix}
A & T \\
0 & B
\end{bmatrix} E \begin{bmatrix}
A & 0 \\
0 & B
\end{bmatrix} = T = AY + XB, \text{ for suitable } X, Y \in R_{rs}.
\]

**Proof:**\((=)\) Note that

\[
\begin{bmatrix}
I_r & X \\
0 & I_s
\end{bmatrix} \begin{bmatrix}
A & 0 \\
0 & I_s
\end{bmatrix} \in R_{rs}, \text{ and that }
\begin{bmatrix}
I_r & X \\
0 & I_s
\end{bmatrix} \begin{bmatrix}
A & 0 \\
0 & B
\end{bmatrix} \begin{bmatrix}
I_r & Y \\
0 & I_s
\end{bmatrix} = \begin{bmatrix}
A & T \\
0 & B
\end{bmatrix}.
\]

Hence

\[
\begin{bmatrix}
A & T \\
0 & B
\end{bmatrix} E \begin{bmatrix}
A & 0 \\
0 & B
\end{bmatrix}.
\]

\((=)\) Let \(\varphi \begin{bmatrix}
A & T \\
0 & B
\end{bmatrix}\) be the statement we wish to prove, namely

\[
\begin{bmatrix}
A & T \\
0 & B
\end{bmatrix} E \begin{bmatrix}
A & 0 \\
0 & B
\end{bmatrix} = \exists X, Y \in R_{rs}
\]
such that \(T = AY + XB\).

We will begin with four reduction steps

(i) We may assume w.l.o.g. (without loss of generality) that \(A = S(A), B = S(B)\). Justification: Choose \(U, U^* \in R'_r; V, V^* \in R'_{ss}\) such that \(UAV^* = S(A), VAV^* = S(B)\). Note that

\[
\begin{bmatrix}
U & 0 \\
0 & V
\end{bmatrix} \begin{bmatrix}
A & 0 \\
0 & B
\end{bmatrix} \begin{bmatrix}
U^* & 0 \\
0 & V^*
\end{bmatrix} = \begin{bmatrix}
S(A) & 0 \\
0 & S(B)
\end{bmatrix}
\]

and that

\[
\begin{bmatrix}
U & 0 \\
0 & V
\end{bmatrix} \begin{bmatrix}
A & T \\
0 & B
\end{bmatrix} \begin{bmatrix}
U^* & 0 \\
0 & V^*
\end{bmatrix} = \begin{bmatrix}
S(A) & \tilde{T} \\
0 & S(B)
\end{bmatrix},
\]

where \(\tilde{T} = UAV^*\).

Hence

\[
\begin{bmatrix}
A & 0 \\
0 & B
\end{bmatrix} E \begin{bmatrix}
S(A) & 0 \\
0 & S(B)
\end{bmatrix} \text{ and } \begin{bmatrix}
A & T \\
0 & B
\end{bmatrix} E \begin{bmatrix}
S(A) & \tilde{T} \\
0 & S(B)
\end{bmatrix}.
\]

90
Thus \[
\begin{bmatrix}
A & T \\
0 & B
\end{bmatrix} \tilde{E} \begin{bmatrix}
A & 0 \\
0 & B
\end{bmatrix} \leq > \begin{bmatrix}
S(A) & T \\
0 & S(B)
\end{bmatrix} \tilde{E} \begin{bmatrix}
S(A) & 0 \\
0 & S(B)
\end{bmatrix}.
\]

Note also that \( T = AY + XB \leq > UTV^* = UAYV^* + UXBV^* \leq > T = UAU^* \left[ (U^*)^{-1} YV^* \right] + UXV^{-1} \left[ YV^* \right] \leq > T = S(A)Y + XS(B) \), where \( X = UXV^{-1} \), \( Y = (U^*)^{-1} YV^* \). It follows that

\[
\varphi \begin{bmatrix}
A & T \\
0 & B
\end{bmatrix} \leq > \varphi \begin{bmatrix}
S(A) & T \\
0 & S(B)
\end{bmatrix}.
\]

Hence setting \( T = \bar{T} \), we may assume w.l.o.g. that \( A = S(A), B = S(B) \).

(ii) Let \( r' = \text{rank } A, s' = \text{rank } B \). We may assume w.l.o.g. that \( T = (t_i, r_{i,j}) \leq < i < r, j \leq s \), where

\[
t_{i, r_{i,j}} = 0 \text{ for } (i, j) \text{ such that } r' < i \leq r \text{ or } s' < i \leq s.
\]

Justification: Let \( A = S(A) = \text{diag}(\alpha_1, \ldots, \alpha_r, 0, \ldots, 0), B = S(B) = \text{diag}(\beta_1, \ldots, \beta_s, 0, \ldots, 0) \), where \( \alpha_1 \leq \alpha_2 \leq \cdots \leq \alpha_r \) and \( \beta_1 \leq \beta_2 \leq \cdots \leq \beta_s \). Assume first that \( r' < r \) and \( s' < s \). If \( (i, j) \in (r', r] \times (s', s] \) such that \( t_{i, r_{i,j}} \neq 0 \), then it would follow that \( \text{rank} \begin{bmatrix}
A & T \\
0 & B
\end{bmatrix} > r' + s' = \text{rank} \begin{bmatrix}
A & 0 \\
0 & B
\end{bmatrix} \), a contradiction. Hence \( t_{i, r_{i,j}} = 0 \) for \( (i, j) \in (r', r] \times (s', s] \).

Assume now that \( r' < r \) and choose \( (i, j) \in (r', r] \times [1, s'] \). Then it is easily seen that \( d_{r', s'} \begin{bmatrix}
A & 0 \\
0 & B
\end{bmatrix} = \alpha_1 \ldots \alpha_r \beta_1 \ldots \beta_s, \) and that

\[
\delta = \begin{cases}
\alpha_1 \ldots \alpha_r \beta_1 \ldots \beta_j t_{i, r_{i,j}}, & j < s' \\
\alpha_1 \ldots \alpha_r \beta_1 \ldots \beta_j t_{i, r_{i,j}}, & j = s'
\end{cases}
\]

is an \((r'+s') \times (r'+s')\) determinantal divisor of \( \begin{bmatrix}
A & T \\
0 & B
\end{bmatrix} \). Since \( \begin{bmatrix}
A & T \\
0 & B
\end{bmatrix} \tilde{E} \begin{bmatrix}
A & 0 \\
0 & B
\end{bmatrix} \), it follows that \( \alpha_1 \ldots \alpha_r \beta_1 \ldots \beta_i \delta \), from which we deduce that \( \beta_i | t_{i, r_{i,j}} \). Hence we may choose \( w_{i, r_{i,j}} \in R \) such that \( t_{i, r_{i,j}} = w_{i, r_{i,j}} \beta_i \).

Assume finally that \( s' < s \) and choose \( (i, j) \in [1, r'] \times (s', s] \). Then it is easily seen that

\[
\eta = \begin{cases}
\alpha_1 \ldots \alpha_{i-1} \alpha_i t_{i, r_{i,j}}, & i < r' \\
\alpha_1 \ldots \alpha_{i-1} \beta_i t_{i, r_{i,j}}, & i = r'
\end{cases}
\]

is an \((r'+s') \times (r'+s')\) determinantal divisor of \( \begin{bmatrix}
A & T \\
0 & B
\end{bmatrix} \). It follows that \( \alpha_1 \ldots \alpha_r \beta_1 \ldots \beta_s \eta \), from which we deduce that \( \alpha_i | t_{i, r_{i,j}} \). Hence we may choose \( z_{i, r_{i,j}} \in R \) such that \( t_{i, r_{i,j}} = \alpha_i z_{i, r_{i,j}} \).

Now for \( 1 \leq i \leq r, 1 \leq j \leq s \) set

\[
\tilde{w}_{i, r_{i,j}} = \begin{cases}
w_{i, r_{i,j}}, & \text{if } r' < i \leq r, 1 \leq j \leq s', t_{i, r_{i,j}} \neq 0 \\
0, & \text{otherwise}
\end{cases}
\]

and set

\[
\tilde{z}_{i, r_{i,j}} = \begin{cases}
z_{i, r_{i,j}}, & \text{if } 1 \leq i \leq r', s' < j \leq s, t_{i, r_{i,j}} \neq 0 \\
0, & \text{otherwise}.
\end{cases}
\]

Let \( \tilde{W} = (\tilde{w}_{i, r_{i,j}})_{1 \leq i \leq r, 1 \leq j \leq s}, \tilde{Z} = (\tilde{z}_{i, r_{i,j}})_{1 \leq i \leq r, 1 \leq j \leq s}, \).

91
Then \[
\begin{bmatrix}
I_r - W & A \\
0 & I_s
\end{bmatrix}
\begin{bmatrix}
A & T \\
0 & B
\end{bmatrix}
\begin{bmatrix}
I_r - Z \\
0 & I_s
\end{bmatrix} =
\begin{bmatrix}
A & \tilde{T} \\
0 & B
\end{bmatrix},
\] where \(\tilde{T} = T - AZ - WB\). Note that
\[
(\tilde{T})_{i,r+j}, 1 \leq i \leq r' \text{ or } 1 \leq j \leq s'
\]
It follows that
\[
\begin{bmatrix}
A & T \\
0 & B
\end{bmatrix}
E
\begin{bmatrix}
A & 0 \\
0 & B
\end{bmatrix}
\Rightarrow
\begin{bmatrix}
A & \tilde{T} \\
0 & B
\end{bmatrix}
E
\begin{bmatrix}
A & 0 \\
0 & B
\end{bmatrix}.
\]
It is also immediate that \(T = AY + XB\) \(\Rightarrow\) \(\tilde{T} = \tilde{A}Y + \tilde{X}B\), where \(\tilde{X} = X - W, \tilde{Y} = Y - Z\). Thus
\[
\varphi \begin{bmatrix}
A & T \\
0 & B
\end{bmatrix} \Rightarrow \varphi \begin{bmatrix}
\tilde{A} & \tilde{T} \\
0 & B
\end{bmatrix}.
\]
Hence we may assume w.l.o.g. that \(T = \tilde{T}\), i.e., that \(T\) is of the form specified above.

(iii) We may assume w.l.o.g. that \(A = r, \text{ rank } B = s\). Justification: We have from (i) and (ii) that we may assume that \(A = \tilde{A} + 0_{r-r'}, \text{ where } \tilde{A} = \text{diag} (\alpha_1, \ldots, \alpha_{r'})\), \(B = \tilde{B} + 0_{s-s'}, \text{ where } \tilde{B} = \text{diag} (\beta_1, \ldots, \beta_{s'})\), and that \(T = \tilde{T} + 0_{r-r', s-s'}\), where \(\tilde{T} \in \mathbb{R}_{r,s}'\). It is not difficult to show that
\[
\begin{bmatrix}
A & T \\
0 & B
\end{bmatrix}
E
\begin{bmatrix}
A & 0 \\
0 & B
\end{bmatrix} = \begin{bmatrix}
\tilde{A} & \tilde{T} \\
0 & B
\end{bmatrix}.
\]
also that \(T = AY + XB\) for some \(X, Y \in \mathbb{R}_{rs}\) \(\Rightarrow\) \(\tilde{T} = \tilde{A}Y + \tilde{X}B\) for some \(\tilde{X}, \tilde{Y} \in \mathbb{R}_{rs}'\). Thus
\[
\varphi \begin{bmatrix}
A & T \\
0 & B
\end{bmatrix} \Rightarrow \varphi \begin{bmatrix}
\tilde{A} & \tilde{T} \\
0 & B
\end{bmatrix}.
\]
Hence we may assume w.l.o.g. that \(A = \tilde{A}, B = \tilde{B}\), i.e., that rank \(A = r, \text{ rank } B = s\).

(iv) We may assume w.l.o.g. that \(r = s\). Justification: Assume \(r < s\). Let
\[
\tilde{A} = I_{s-r} \hat{A} + A, T = \begin{bmatrix} 0_{s-r,s} \end{bmatrix}.
\]
It is an easy consequence of [1, Ch. 2, ex. 1] that
\[
\begin{bmatrix}
A & T \\
0 & B
\end{bmatrix}
E
\begin{bmatrix}
A & 0 \\
0 & B
\end{bmatrix} \Rightarrow \begin{bmatrix}
\tilde{A} & \tilde{T} \\
0 & B
\end{bmatrix}.
\]
It is also not difficult to show that \(T = AY + XB\) for some \(X, Y \in \mathbb{R}_{rs}\) \(\Rightarrow\) \(\tilde{T} = \tilde{A}Y + \tilde{X}B\) for some \(\tilde{X}, \tilde{Y} \in \mathbb{R}_{rs}'\). Thus
\[
\varphi \begin{bmatrix}
A & T \\
0 & B
\end{bmatrix} \Rightarrow \varphi \begin{bmatrix}
\tilde{A} & \tilde{T} \\
0 & B
\end{bmatrix}.
\]
Assume now that \(s < r\). Let \(\tilde{B} = I_{r-s} \hat{B} + B, \tilde{T} = [0_{r-r-s}, T]\). Proceeding as above, we can show that
\[
\varphi \begin{bmatrix}
A & T \\
0 & B
\end{bmatrix} \Rightarrow \varphi \begin{bmatrix}
A & \tilde{T} \\
0 & B
\end{bmatrix}.
\]
It then follows from this and the above case that we may assume w.l.o.g. that \(r = s\).

We now complete the proof of the theorem. By (i) – (iv), we may assume w.l.o.g. that \(A = S(A) = \text{diag} (\alpha_1, \ldots, \alpha_r), B = S(B) = \text{diag} (\beta_1, \ldots, \beta_r)\), where \(\alpha_i, \beta_j \neq 0, 1 \leq i, j \leq r\). Note that
\[
\begin{bmatrix}
A & T \\
0 & B
\end{bmatrix}
E
\begin{bmatrix}
A & 0 \\
0 & B
\end{bmatrix} \Rightarrow \forall k \leq 2r, d_k \begin{bmatrix} A & 0 \\
0 & B \end{bmatrix}\text{ every } k \times k \text{ determinantal divisor of } \begin{bmatrix} A & T \\
0 & B \end{bmatrix}. Note
also that \( T = \mathbf{A}Y + \mathbf{X}B \) \( \implies \) \( (\alpha_i, \beta_j) \mid t_{i,r+j}, 1 \leq i, j \leq r \). Now

\[
d_k \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} = (\alpha_1 \ldots \alpha_k, \beta_1 \ldots \beta_k, \{\alpha_{r-t-1} \ldots \alpha_t \beta_1 \ldots \beta_{k-1-i}\}_1 \leq t < k-2),
\]

as is easily calculated. We consider two cases.

a. \( i \geq j \). Let \( \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}_{r+j,i} \) denote the matrix obtained by deleting row \( r+j \) and column \( i \) from \( \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} \). It is easily seen that if \( \rho \) is any \( (i-1) \times (i-1) \) determinantal minor of \( \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}_{r+j,i} \)

then \( \rho t_{i,r+j} \) is an \( i \times i \) determinantal minor of \( \begin{bmatrix} A & T \\ 0 & B \end{bmatrix} \). It then follows that

\[
d_i \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}_{r+j,i} d_{i-1} \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}_{r+j,i} t_{i,r+j},
\]

Note that

\[
d_{i-1} \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}_{r+j,i} = (\alpha_1 \ldots \alpha_{i-1}, \{\beta_1 \ldots \beta_{j-1} \beta_{j+1} \ldots \beta_i, \text{if} \ j < i \}
\{
\beta_1 \ldots \beta_{i-1}, \text{if} \ j = i
\}
\{\alpha_1 \ldots \alpha_{i-u} \beta_1 \ldots \beta_{u-1} \} 2 \leq u \leq j,
\{\alpha_1 \ldots \alpha_{i-u} \beta_1 \ldots \beta_{j-1} \beta_{j+1} \ldots \beta_{r} \} j+1 \leq r \leq i-1 \} (\text{for} \ j \leq i-2).
\]

From this it may be verified that

\[
(\alpha_i, \beta_j) d_{i-1} \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}_{r+j,i} d_i \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}_{r+j,i}.
\]

It follows that \( (\alpha_i, \beta_j) \mid t_{i,r+j} \), as was to be shown.

b. \( j > i \). We proceed as in (a). If \( \sigma \) is any \( (j-1) \times (j-1) \) determinantal minor of \( \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}_{r+j,i} \)

then \( \sigma t_{i,r+j} \) is a \( j \times j \) determinantal minor of \( \begin{bmatrix} A & T \\ 0 & B \end{bmatrix} \). It then follows that

\[
d_j \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} d_{j-1} \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}_{j,r+j,i} t_{i,r+j}.
\]

Note that

\[
d_{j-1} \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}_{r+j,i} = (\alpha_1 \ldots \alpha_{i-1} \alpha_{i+1} \ldots \alpha_j,
\beta_1 \ldots \beta_{j-1}, \{\alpha_1 \ldots \alpha_{u-1} \beta_1 \ldots \beta_{j-1} \beta_{j+1} \ldots \beta_{r} \} j+1 \leq r \leq \alpha_{i-1} \alpha_{i+1} \ldots \alpha_j \beta_1 \ldots \beta_{j-1} \beta_{j+1} \ldots \beta_{r} \}(j \leq i-2).
\]

From this it may be verified that \( (\alpha_i, \beta_j) d_{j-1} \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}_{r+j,i} d_j \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}_{r+j,i}.
\]

It follows that \( (\alpha_i, \beta_j) \mid t_{i,r+j} \), as was to be shown.

With this we have completed proof of the theorem. Q.E.D.

**Corollary 1.1:** Let \( A, B, \) and \( R \) be as before, and suppose \( P, Q \in \mathbb{R}_{rs} \).
Then \[ \begin{bmatrix} A & Q-P \\ 0 & B \end{bmatrix} \bar{E} \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} \Rightarrow \begin{bmatrix} A & P \\ 0 & B \end{bmatrix} \bar{E} \begin{bmatrix} A & Q \\ 0 & B \end{bmatrix} \].

PROOF: By Theorem 1, \( Q - P = AY + XB \), for some \( X, Y \in \mathbb{R}_r \). Note that
\[
\begin{bmatrix} I & X \\ 0 & 1 \end{bmatrix} \begin{bmatrix} A & P \\ 0 & B \end{bmatrix} \begin{bmatrix} I & Y \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} A & Q \\ 0 & B \end{bmatrix}.
\]
Hence \[ \begin{bmatrix} A & P \\ 0 & B \end{bmatrix} \bar{E} \begin{bmatrix} A & Q \\ 0 & B \end{bmatrix} \].

Q.E.D.

NOTE: The converse to Corollary 1.2 fails. For example, \[ \begin{bmatrix} 6 & 4 \\ 0 & 9 \end{bmatrix} \bar{E} \begin{bmatrix} 6 & 2 \\ 0 & 9 \end{bmatrix} \], as may be verified by considering determinantal divisors, but \[ \begin{bmatrix} 6 & 4-2 \\ 0 & 9 \end{bmatrix} \] is not equivalent to \[ \begin{bmatrix} 6 & 0 \\ 0 & 9 \end{bmatrix} \].

We now generalize Theorem 1 as follows:

**THEOREM 2:** Let \( M \) be a matrix over \( \mathbb{R} \), and suppose that \( M \) may be partitioned as
\[
M = \begin{bmatrix} M_{11} & M_{12} & \ldots & M_{1t} \\ 0 & M_{22} & \ldots & M_{2t} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \ldots & 0 & M_{tt} \end{bmatrix}
\]
where each \( M_{ii} \) is square, of order \( r_i \). Then \( M \bar{E} \) \( \text{diag}[M_{11}, \ldots, M_{tt}] \) \((=)\) for \( 1 \leq i < j \leq t \) \( \exists \) \( X_{ij} \), \( Y_{ij} \in \mathbb{R}_{r_i r_j} \) such that \( M_{ij} = M_{ii} Y_{ij} + \sum_{k=i+1}^{t} X_{ik} M_{kj} \).

PROOF \((\Leftarrow):\) Let
\[
U = \begin{bmatrix} 1 & X_{12} & \ldots & X_{1t} \\ 0 & 1 & \ldots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \ldots & 0 & 1 \end{bmatrix}
\]
and
\[
V = \begin{bmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \ldots & 0 & 1 \end{bmatrix}
\]

Then \( U, V \) are invertible and \( U \) \( \text{diag}[M_{11}, \ldots, M_{tt}] \) \( V = M \), as may be verified. Hence \( M \bar{E} \) \( \text{diag} [M_{11}, \ldots, M_{tt}] \).
(⇒) Let $A_1 = M_{11}$, $T_1 = [M_{12}, \ldots, M_{1t}]$ and

$$B_1 = \begin{bmatrix} M_{22} & \cdots & \cdots & M_{2t} \\ 0 & \ddots & & \\ \vdots & \ddots & \ddots & \\ 0 & \cdots & \cdots & M_{tt} \end{bmatrix}$$

so that $M = \begin{bmatrix} A_1 & T_1 \\ 0 & B_1 \end{bmatrix}$. Note that to obtain a minor of $\begin{bmatrix} A_1 & 0 \\ 0 & B_1 \end{bmatrix}$ with nonzero determinant, it is necessary that the number of rows deleted which pass through the block $A_1$ equal the number of columns deleted which pass through this block. It follows from this that every determinantal divisor of $\begin{bmatrix} A_1 & 0 \\ 0 & B_1 \end{bmatrix}$ is also a determinantal divisor of $M$. Since $M \tilde{E} \text{diag}[M_{11}, \ldots, M_{tt}]$, it follows that $\begin{bmatrix} A_1 & 0 \\ 0 & B_1 \end{bmatrix} \tilde{E} \text{diag}[M_{11}, \ldots, M_{tt}]$ as well, so that $M \tilde{E} \begin{bmatrix} A_1 & 0 \\ 0 & B_1 \end{bmatrix}$. Hence by Theorem 1 there are matrices $X_1, Y_1$ of appropriate size such that $T_1 = A_1Y_1 + X_1B$. We may write $X_1 = [X_{12}, \ldots, X_{1t}]$, $Y_1 = [Y_{12}, \ldots, Y_{1t}]$, where $X_{1j}, Y_{1j} \in R_{r_{1j}}, 2 \leq j \leq t$, from which it follows that $M_{1j} = M_{11}Y_{1j} + \sum_{k=2}^{t} X_{1k}M_{kj}$.

Now let $A_2 = M_{22}$, $T_2 = [M_{23}, \ldots, M_{2t}]$,

$$B_2 = \begin{bmatrix} M_{33} & \cdots & \cdots & M_{3t} \\ 0 & \ddots & & \\ \vdots & \ddots & \ddots & \\ 0 & \cdots & \cdots & M_{tt} \end{bmatrix}$$

Then $M \tilde{E} \begin{bmatrix} A_2 & T_2 \\ 0 & B_2 \end{bmatrix} \tilde{E} \begin{bmatrix} A_2 & 0 \\ 0 & B_2 \end{bmatrix}$ as before, so that $M \tilde{E} \begin{bmatrix} A_2 & T_2 \\ 0 & B_2 \end{bmatrix} \tilde{E} \begin{bmatrix} A_2 & 0 \\ 0 & B_2 \end{bmatrix}$. Proceeding as above we obtain that $\exists X_2 = [X_{23}, \ldots, X_{2t}]$, $Y_2 = [Y_{23}, \ldots, Y_{2t}]$ such that

$$M_{2j} = M_{22}Y_{2j} + \sum_{k=3}^{t} X_{2k}M_{kj}, 3 \leq j \leq t.$$ 

Continuing in this manner we obtain the desired linear recurrence relation.

**COROLLARY 2.1:** Let $R$ and $\{M_{ij}\}_{i, j}$ be as in Theorem 2, and

$$M = \begin{bmatrix} M_{11} & M_{12} & \cdots & M_{1t} \\ 0 & M_{22} & \cdots & M_{2t} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & \cdots & M_{tt} \end{bmatrix}$$

Then $M \tilde{E} \text{diag}[M_{11}, M_{22}, \ldots, M_{tt}] \forall \{M_{ij}\}_{1 \leq i \leq j \leq t}$, where $M_{ij} \in R_{r_{ij}}, r_{ij}$ (\(=\))

(i) $(\det M_{ij}, \det M_{jj}) = 1, 1 \leq i \leq j \leq t$ or

(ii) $\exists j \in [1, n]$ such that $\det M_{jj} = 0$ and such that for all $k \neq j$, $M_{kk}$ is unimodular.
PROOF: \((\Leftarrow)\) (i) This is essentially Theorem 3 of [1]. (ii) Consider first the equation

\[
M_{t-1,t} = M_{t-1,t-1}Y_{t-1,t} + X_{t-1,t}M_{t,t}.
\]

By hypothesis, at least one of \(M_{t-1,t-1}\), \(M_{t,t}\) is unimodular. If \(M_{t-1,t-1}\) is unimodular, we may let \(X_{t-1,t} = 0\) and \(Y_{t-1,t} = M_{t-1,t-1}^{-1} M_{t-1,t}\) to obtain a solution to \((*)\). If \(M_{t,t}\) is unimodular, it is again easy to solve \((*)\).

Consider now the equations

\[
M_{t-2,t-1} = M_{t-2,t}Y_{t-2,t-1} + X_{t-2,t}M_{t-1,t-1}
\]

and

\[
M_{t-2,t} = M_{t-2,t}Y_{t-2,t} + X_{t-2,t-1}M_{t-1,t} + X_{t-2,t}M_{t,t}
\]

Proceeding as above, it is again easy to solve \((**)*\), this time for \(X_{t-2,t-1}, Y_{t-2,t-1}\). Now rewrite \((**)*\) as

\[
M_{t-2,t} = M_{t-2,t-1}M_{t-1,t} = M_{t-2,t-2}Y_{t-2,t-1} + X_{t-2,t-1}M_{t-1,t} + X_{t-2,t}M_{t,t}
\]

and note that the matrices on the left-hand side have all been given or determined previously. Again, it is easy to solve \((**)*\), for \(X_{t-2,t}, Y_{t-2,t}\).

Proceeding in this manner, for \(1 \leq i < j \leq t\) we may find \(X_{i,j}, Y_{i,j} \in \mathbb{R}^{i,j}\) such that \(M_{i,j} = M_{i,j}Y_{i,j} + \sum_{k=1}^{i} X_{i,k}M_{k,j}\). Hence by Theorem 2, \(M \in \mathbb{R}^{i,j}\) is unimodular. If we choose any \(k < \ell \leq t\) and \(r' < r_j\), we may obtain \(\tilde{Y}_{i,j}\) at rank \(M_{ij} = \det M_{ij} = 0\) or \(v \neq j\), we would obtain \(\tilde{Y}_{i,j}\) at rank \(M > \det M_{ij} = 0\), contradicting hypothesis. Hence there is at most one \(j \in [1,n]\) such that \(\det M_{i,j} = 0\).

Suppose first that there is such an \(i\). We will show that \((i)\) holds in this case. Choose any \(j > i\) (if such exist) and let \(M_{i,j} = M_{i,j}Y_{i,j} + \sum_{k=1}^{j} X_{i,k}M_{k,j}\). Hence by Theorem 2,

\[
M_{i,j} = M_{i,j}Y_{i,j} + \sum_{k=1}^{j} X_{i,k}M_{k,j} = M_{i,j}Y_{i,j} + X_{i,j}M_{j,j}.
\]

Considering the \((k, l)\) component of this matrix equation, for \(r' < k \leq r_i\) and \(l \leq r_j\) we obtain that

\[
1 = (X_{i,j})_{k,l}\alpha_{j,l}.
\]

This implies that \(\alpha_{j,l}\) is a unit, \(1 \leq l \leq r_j\). Since \(\det M_{j,j} = \prod_{i=1}^{r_j} \alpha_{j,l}\), we obtain that \(M_{j,j}\) is unimodular. If we choose any \(j < i\), we may obtain by a similar argument that \(\det M_{i,j}\) is a unit, so that \(M_{i,j}\) is unimodular in this case as well. We can thus conclude that \((i)\) holds when \(\det M_{i,j} = 0\) for some \(i\).

Now suppose that \(\det M_{i,j} \neq 0\), \(1 \leq i \leq t\). We will show that \((ii)\) holds in this case. Choose \(i \neq j \in [1, t]\) and let \(M_{i,j} = M_{i,j}Y_{i,j} + \sum_{u \neq i} X_{i,j}M_{u,j}\) be defined as before. Again, \(M_{i,j} = M_{i,j}Y_{i,j} + X_{i,j}M_{j,j}\). Considering the \((k, l)\) component of this matrix equation, for \(1 \leq k \leq r_i\), \(1 \leq l \leq r_j\) we obtain that

\[
1 = (X_{i,j})_{k,l} + (X_{i,j})_{k,l}\alpha_{j,l}.
\]
It follows that \((\alpha_{ik}, \alpha_{jl}) = 1\). Since \(\det M_{ii} = \prod_{k=1}^{r_i} \alpha_{ik}\) and \(\det M_{jj} = \prod_{l=1}^{r_j} \alpha_{jl}\), we obtain that \((\det M_{ii}, \det M_{ij}) = 1\). We can thus conclude that (ii) holds when \(\det M_{ii} \neq 0\), all \(i\).

This completes the proof of the corollary.

Q.E.D.

References


(Paper 80B1–435)