AN INEQUALITY FOR INVARIANT FACTORS

ROBERT C. THOMPSON

Abstract. A divisibility relation is proved connecting the invariant factors of integral matrices $A$, $B$, and $C$ when $C = AB$.

Let $A$, $B$, and $C$ be $n \times n$ matrices with entries in a principal ideal domain $\mathcal{R}$, and with $C = AB$. In a recent note [3] on the multiplicative property of the Smith normal form, Morris Newman observed the fact: if $d_i(A)$ denotes the $i$th determinantal divisor of $A$, then $d_i(A)d_i(B) \mid d_i(C)$, where $\mid$ denotes divisibility. The objective of this paper is to prove the following divisibility property of invariant factors, a property containing Newman’s observation as a special case.

Notation. $\alpha_1 \mid \alpha_2 \mid \cdots \mid \alpha_n$, $\beta_1 \mid \beta_2 \mid \cdots \mid \beta_n$, $\gamma_1 \mid \gamma_2 \mid \cdots \mid \gamma_n$ are the invariant factors of $A$, $B$, and $C$, respectively. See [4] for all properties of invariant factors used here.

Theorem. We have

$$\alpha_{i_1} \alpha_{i_2} \cdots \alpha_{i_m} \beta_{j_1} \beta_{j_2} \cdots \beta_{j_m} \mid \gamma_{i_1+j_1-1} \gamma_{i_2+j_2-2} \cdots \gamma_{i_m+j_m-m}$$

whenever the integer subscripts satisfy

$$1 \leq i_1 < i_2 < \cdots < i_m, \quad 1 \leq j_1 < j_2 < \cdots < j_m, \quad i_m + j_m \leq m + n.$$

Proof. Let $p$ be a fixed prime in $\mathcal{R}$, and let $\mathcal{R}_p$ be the ring of all fractions $a/b$, where $a$, $b$ lie in $\mathcal{R}$ and $p$ does not divide $b$. Ring $\mathcal{R}_p$ is a principal ideal ring, with every nontrivial ideal a power of the principal ideal generated by $p$. Observe that (1) holds when the $\alpha$’s, $\beta$’s, and $\gamma$’s are invariant factors if and only if it holds when the $\alpha$’s, $\beta$’s, and $\gamma$’s are the elementary divisors belonging to the prime $p$, for each choice of $p$ dividing $\det C$. And these elementary divisors are the invariant factors of $A$, $B$, $C$ when the matrices are regarded as having elements in the local ring $\mathcal{R}_p$, an observation due some years ago to L. Gerstein [1]. So (1) will be proved if it can be proved when $A$, $B$, and $C$ are matrices over $\mathcal{R}_p$.

The proof will easily be completed once the following lemma is established.

Lemma. Over $\mathcal{R}_p$, we may assume that:

(i) $B$ is diagonal, $B = \text{diag}(\beta_1, \beta_2, \ldots, \beta_n)$;
(ii) $A$ is triangular, $A = [a_{ij}]$ with $a_{ij} = 0$ for $i > j$;
(iii) $C$ is triangular, $C = [c_{ij}]$ with $c_{ij} = 0$ for $i > j$, and $c_{ii} = \gamma_i$, $c_{ij} \mid c_{ij}$ for $j > i$, $1 \leq i \leq n$.

Received by the editors September 15, 1981.
1980 Mathematics Subject Classification. Primary 15A36; Secondary 15A39.
Proof of Lemma. From \( C = AB \) we get \( UCV = (UAW^{-1})(WBV) \), where \( U, V, W \) are unimodular. First, choose \( W \) and \( V \) to put \( B \) into its Smith form: \( WBV = \text{diag}(\beta_1, \beta_2, \ldots, \beta_n) \). Then choose \( U \) to put \( AW^{-1} \) into Hermite (triangular) form. Since we are only interested in invariant factors, which these transformations preserve, we may henceforth assume that \( B \) is diagonal, \( A \) and \( C \) are triangular. All of this holds over \( \mathbb{R} \) as well as over \( \mathbb{R}_p \).

If \( C = 0 \), the claims are trivially correct, so suppose that \( C \neq 0 \).

Always \( \gamma_i \) is the greatest common divisor of the \( c_{ij} \), but over \( \mathbb{R}_p \) \( \gamma_i \) is the power of \( p \) exactly present in those \( c_{ij} \) exhibiting the lowest exponent on \( p \). Among these minimal \( c_{ij} \), select one with \( i \) least. If \( i > 1 \) we may left multiply \( C \) and \( A \) by a unimodular \( U \) that adds row \( i \) to row 1. For the \( C = AB \) now at hand, the minimal \( p \) power exactly dividing an element of \( C \) appears in a first row element. So we may suppose that \( i = 1 \). If the minimal \( c_{1j} \) for which \( j \) is least has \( j > 1 \), proceed as follows: Choose unimodular \( V \) such that in \( CV \) column \( j \) of \( C \) is added to column 1, and let \( W \) be unimodular so that in \( WB \) the \( \beta_j/\beta_1 \) multiple of row 1 is subtracted from row \( j \). Then in \( CV = (AW^{-1})(WBV) \), we have \( WVB = \text{diag}(\beta_1, \ldots, \beta_n) \) and \( \gamma_1 \) (to within a unit) is the \((1,1)\) element of \( CV \). But \( CV \) and \( AW^{-1} \) are no longer triangular.

Rename the matrices at hand as \( C = AB \), with \( c_{11} = \gamma_1 \) and \( B = \text{diag}(\beta_1, \ldots, \beta_n) \). Since \( c_{1i} = a_{i1}\beta_1 \) and \( c_{11} \mid c_{1i} \), evidently \( a_{11} \mid a_{1i} \). Elementary row operations on \( A \) (therefore on \( C \) also) now make \( a_{21} = \cdots = a_{n1} = 0 \), whence \( c_{21} = \cdots = c_{n1} = 0 \).

Thus \( A \) and \( C \) are block triangular, and \( c_{11} = \gamma_1 \) divides each \( c_{ij} \). Since \( C \) is unimodularly equivalent to the direct sum of \( c_{11} \) and \( [c_{ij}]_{2 \leq i, j \leq n} \), evidently the trailing \((n-1)\)-square block in \( C \) has invariant factors \( \gamma_2, \ldots, \gamma_n \). And by an obvious left multiplication by a unimodular \( U \), the trailing blocks in \( A \) and \( C \) may be assumed triangular.

We now repeat this procedure on the last \( n-1 \) rows and columns if the trailing block in \( C \) is nonzero, there being nothing further to prove if it is zero. Continuing in this manner, the lemma is established.

Proof of Theorem concluded. We proceed by induction on \( n \). The initial value is \( n = m \), in which case (1) merely asserts that \( \det A \det B \mid \det C \), trivially true. So suppose \( n > m \).

We adapt a trick used by M. F. Smiley [5] in quite another context. Define integers \( u \) and \( v \) by

\[
i_1 = 1, \ldots, i_u = u, \quad u = m \quad \text{or} \quad i_{u+1} > u + 1, \\
j_1 = 1, \ldots, j_v = v, \quad v = m \quad \text{or} \quad j_{v+1} > v + 1.
\]

One of \( u, v \) is the smaller, and by transposing \( C \) if necessary we may assume that \( v \leq u \). Now apply the lemma and so have \( B = \text{diag}(\beta_1, \ldots, \beta_n) \), \( A \) and \( C \) triangular, \( c_{ii} = \gamma_i \) and \( c_{ii} \mid c_{ij} \) for all \( i \leq j \).

Let \( C' \) be the matrix gotten from \( C \) by deleting row \( v + 1 \) and column \( v + 1 \); similarly for \( A' \), \( B' \) from \( A \) and \( B \), respectively. The diagonal form of \( B \) then implies that \( C' = A'B' \), and that the invariant factors of \( B' \) are

\[
\beta'_i = \beta_1, \ldots, \beta'_v = \beta_v, \quad \beta'_{v+1} = \beta_{v+2}, \ldots, \beta'_{n-1} = \beta_n.
\]
The invariant factors \( a'_1 | \cdots | a'_{n-1} \) of \( A' \) are known \([6]\) to satisfy

\[
\alpha_1 | \alpha'_1, \ldots, \alpha'_{n-1} | a'_{n-1}.
\]

And the special structure of \( C \) shows that the invariant factors of \( C' \) are

\[
\gamma'_1 = \gamma_1, \ldots, \gamma'_v = \gamma_v, \quad \gamma'_{v+1} = \gamma_{v+2}, \ldots, \gamma'_{n-1} = \gamma_n.
\]

Let

\[
I_1 = i_1, \ldots, I_m = i_m,
\]

\[
J_1 = j_1, \ldots, J_v = j_v, \quad J_{v+1} = J_{v+1} - 1, \ldots, J_m = J_m - 1.
\]

By induction on \( n \), the inequality

\[
(2) \quad \alpha'_1 \cdots \alpha'_m | y'_{i_1+j_1-1} \cdots y'_{i_m+j_m-m}
\]

holds. However, the relations written down in the last several lines show that (2) implies (1).

**Comments.** The inequality just proved is one of a large family in which the indices are Littlewood-Richardson sequences. The entire family was established some years ago by the present author, but never published, using a method based on results of Klein \([2]\). Since the inequality (1) is so clean, and its proof so elementary, it seems worthwhile to publish it separately.

The special case \( \alpha_i \beta_j | y_{i+j-1} \) was used in \([7]\) and proved there by a method somewhat similar to the one used above. This special case can be shown to imply the multiplicative property of the Smith form, that is, \( S(AB) = S(A)S(B) \) when \( A \) and \( B \) have relatively prime determinants; \( S(A) \) is the Smith form of \( A \). See \([7]\) for details.

The role of Littlewood-Richardson sequences in some of the classical eigenvalue problems of linear algebra was described in \([8]\).

**Acknowledgement.** The preparation of this paper was partly supported by AFOSR Grant 79-0127.

**References**

7. ______, Left multiples and right divisors of integral matrices (submitted).

**Algebra Institute and Department of Mathematics, University of California at Santa Barbara, Santa Barbara, California 93106**