PORTUGALIAE MATHEMATICA
Vol. 54 Fasc. 3 – 1997

AXIOMS FOR INVARIANT FACTORS*

João Filipe Queiró

Abstract: We show that the invariant factors of matrices over certain types of rings are characterized by a short list of very simple properties.

1 – Introduction

An integral domain R is called an *elementary divisor domain* [3] if every matrix over R is equivalent to a "Smith normal form", that is, there exist U and V invertible over R such that

$$UAV = \begin{bmatrix} s_1(A) & & 0 \\ & s_2(A) & \\ 0 & & \ddots \end{bmatrix}$$

where $s_1(A) \mid s_2(A) \mid \ldots$. The elements $s_1(A)$, $s_2(A)$, ... are the *invariant factors* of A and are uniquely determined (apart from units) by the matrix, as follows from the characterization

$$s_k(A) = \frac{d_k(A)}{d_{k-1}(A)}, \qquad k = 1, ..., \operatorname{rank}(A)$$

($d_k(A)$ — the k-th *determinantal divisor* of A — is the g.c.d. of all $k \times k$ minors of A, $d_0 \equiv 1$). For convenience, we add a chain of 0's to the list of invariant factors.

Examples of elementary divisor domains are Euclidean domains (like $\mathbb{Z}$ and the rings $\mathbb{F}[\lambda]$, $\mathbb{F}$ a field) and, more generally, principal ideal domains. One example of an elementary divisor domain which is not a principal ideal domain

Received: January 22, 1996.
* Research supported by CMUC/JNICT and project PBIC/C/CEN/1129/92 "Representação de Grupos e Combinatória".

is the ring $H(\Omega)$ of all complex functions holomorphic in an open connected set $\Omega \subseteq \mathbb{C}$ [2].

The determinantal divisors (and hence also the invariant factors) are invariant under equivalence. Therefore, over an elementary divisor domain, the invariant factors completely determine the equivalence orbits: two matrices are equivalent if and only if they have the same invariant factors.

The following properties of invariant factors are very simple to establish:

(I) $s_1(cA) = c.s_1(A)$ for all $c \in R$;

(II) $\gcd(s_k(A), s_1(B)) \mid s_k(A+B)$ for all k, whenever $A+B$ exists;

(III) $s_k(A) \mid s_k(PAQ)$ for all k, whenever PAQ exists;

(IV) $k > \operatorname{rank}(A) \Rightarrow s_k(A) = 0$;

(V) $k \leq n \Rightarrow s_k(cI_n) \mid c$ for all $c \in R$.

Our main purpose in the present note is to show that this list of properties actually characterizes the chain of invariant factors.

2 – The main result

Theorem. *Let R be an elementary divisor domain. Suppose that to every matrix A over R we associate a sequence $h_1(A) \mid h_2(A) \mid ...$ of elements of R so that the following properties are satisfied:*

(I) $h_1(cA) = c.h_1(A)$ *for all* $c \in R$;

(II) $gcd(h_k(A), h_1(B)) \mid h_k(A+B)$ *for all k, whenever $A+B$ exists;*

(III) $h_k(A) \mid h_k(PAQ)$ *for all k, whenever PAQ exists;*

(IV) $k > \operatorname{rank}(A) \Rightarrow h_k(A) = 0$;

(V) $k \leq n \Rightarrow h_k(cI_n) \mid c$ *for all* $c \in R$.

Then $h_k(A) = s_k(A)$ for all A and k.

Remark. In condition (I) and in the conclusion of the theorem, equality means "apart from units" (and likewise in similar situations).

The proof of the theorem consists of a sequence of claims.

Claim 1. *If B is a submatrix of A, then $h_k(A) \mid h_k(B)$ for all k.*

Proof: There exist P and Q such that $B = PAQ$ whence (III) gives the result. ∎

Claim 2. *If A and B are equivalent then $h_k(A) = h_k(B)$ for all k.*

Proof: $B = UAV$ gives $h_k(A) \mid h_k(B)$. $A = U^{-1}BV^{-1}$ gives $h_k(B) \mid h_k(A)$. ∎

Claim 3. $h_1(A) = s_1(A)$.

Proof: A is equivalent to its Smith normal form

$$\Sigma = \begin{bmatrix} s_1(A) & & 0 \\ & s_2(A) & \\ 0 & & \ddots \end{bmatrix} = s_1(A).D \ ,$$

where

$$D = \begin{bmatrix} 1 & & & 0 \\ & \dfrac{s_2(A)}{s_1(A)} & & \\ & & \dfrac{s_3(A)}{s_1(A)} & \\ 0 & & & \ddots \end{bmatrix} .$$

Since 1 is a submatrix of D, we have $h_1(D) \mid h_1(1) = 1$, whence $h_1(D) = 1$. Therefore, $h_1(A) = h_1(\Sigma) = h_1(s_1(A).D) = s_1(A)\, h_1(D) = s_1(A)$. ∎

Claim 4. $s_k(A) \mid h_k(A)$ *for all k.*

Proof: By claim 2, we may assume A is in Smith normal form. Let

$$X = \begin{bmatrix} s_1(A) & & & & 0 \\ & \ddots & & & \\ & & s_{k-1}(A) & & \\ & & & 0 & \\ 0 & & & & \ddots \end{bmatrix} \quad \text{(with the size of } A) \ .$$

Clearly $\operatorname{rank}(X) < k$, whence $h_k(X) = 0$. We have

$$\begin{aligned} s_1(A-X) = h_1(A-X) &= \gcd(0, h_1(A-X)) \\ &= \gcd(h_k(X), h_1(A-X)) \mid h_k(X+A-X) = h_k(A) \ , \end{aligned}$$

where we have used claim 3 and (II). But obviously $s_1(A-X)=s_k(A)$. ∎

Claim 5. *$h_k(A) \mid s_k(A)$ for all k.*

Proof: It is enough to consider $k \leq \operatorname{rank}(A)$. Clearly there exist P and Q such that

$$PAQ = \begin{bmatrix} s_1(A) & & 0 \\ & \ddots & \\ 0 & & s_k(A) \end{bmatrix}.$$

Put

$$E = \begin{bmatrix} \dfrac{s_k(A)}{s_1(A)} & & & 0 \\ & \ddots & & \\ & & \dfrac{s_k(A)}{s_{k-1}(A)} & \\ 0 & & & 1 \end{bmatrix}$$

and write $Q' = QE$. Then $PAQ' = s_k(A).I_k$ whence, by (III) and (V),

$$h_k(A) \mid h_k(PAQ') = h_k(s_k(A).I_k) \mid s_k(A)\ . \blacksquare$$

Claims 4 and 5 prove the theorem.

Remark 1. The inspiration for this theorem came from a characterization of singular values by Pietsch [6]. In the language of matrices his result reads as follows: Suppose that to every matrix A over $\mathbb{C}$ we associate a sequence $h_1(A) \geq h_2(A) \geq ...$ of nonnegative numbers so that the following properties are satisfied:

(**I**) $h_1(A) = \|A\|$;

(**II**) $h_k(A) + h_1(B) \geq h_k(A+B)$ for all k, whenever $A+B$ exists;

(**III**) $h_1(P)\, h_k(A)\, h_1(Q) \geq h_k(PAQ)$ for all k, whenever PAQ exists;

(**IV**) $k > \operatorname{rank}(A) \Rightarrow h_k(A) = 0$;

(**V**) $k \leq n \Rightarrow h_k(I_n) = 1$.

Then, for all A, $h_1(A), h_2(A), ...$ are the singular values of A.

Remark 2. Condition (V) in our theorem cannot be replaced by $k \leq n \Rightarrow h_k(I_n) = 1$, as the sequence of determinantal divisors would also satisfy the new axiom list.

3 – Applications

The theorem can be applied to obtain alternative characterizations of invariant factors. Denote s_1 ($=$ gcd) by μ. Let $A \in R^{m\times n}$. Then, for all k, we have, among others, the following characterizations:

$$s_k(A) = \operatorname{lcm}\Big\{\mu(A-X)\colon\ \operatorname{rank}(X) < k\Big\}\ , \tag{1}$$

$$s_k(A) = \gcd\Big\{c \in R\colon\ PAQ = cI_k,\ P \in R^{k\times m},\ Q \in R^{n\times k}\Big\}\ , \tag{2}$$

$$s_k(A) = \operatorname*{lcm}_{\substack{E\le R^n,\\ \dim E=n-k+1}}\ \gcd_{\substack{x\in E,\\ \mu(x)=1}}\ \mu(Ax)\ , \tag{3}$$

$$s_k(A) = \gcd_{\substack{E\le R^n,\\ \dim E=k}}\ \operatorname*{lcm}_{\substack{x\in E,\\ \mu(x)=1}}\ \mu(Ax)\ . \tag{4}$$

The first of these four characterizations appeared in [7]. The third and fourth appeared in [1]. All can be proved very easily by showing that the right-hand side satisfies properties (I)–(V) of the theorem. For this, one must assume beforehand that the indicated lcm's and gcd's actually exist. This is automatic if R is, for example, a principal ideal domain (which was the situation considered in [1] and [7]).

As noted in [1] with respect to (3) and (4), each of these characterizations provides a new proof of the fact that the invariant factors are uniquely determined by the matrix, a matter approached in [5] in a different way.

The alternative characterizations can in turn be used to obtain easy proofs of known results about invariant factors. We list some of these.

Interlacing of invariant factors. If A' $m'\times n'$ is a submatrix of A $m\times n$, then, for all k,

$$s_k(A) \mid s_k(A') \mid s_{k+(m-m')+(n-n')}(A)$$

(the so-called interlacing "inequalities"). The proof is trivial using characterization (1) [7]. (The original proof can be found in [8], [9].)

Invariant factors of sums. We have

$$\gcd\Big(s_i(A), s_j(B)\Big) \mid s_{i+j-1}(A+B)$$

for all i, j. Again the proof is trivial using characterization (1) [7]. (For the original proof, valid only for principal ideal domains, see [10].)

Invariant factors of products. This is an extensively studied problem. For $n \times n$ A and B, known relations have the form

$$\text{(P)} \qquad s_{i_1}(A) \cdots s_{i_t}(A)\, s_{j_1}(B) \cdots s_{j_t}(B) \mid s_{k_1}(AB) \cdots s_{k_t}(AB) ~,$$

where $1 \le t \le n$, $1 \le i_1 < ... < i_t \le n$, $1 \le j_1 < ... < j_t \le n$, $1 \le k_1 < ... < k_t \le n$. The problem is to find all the "right" sequences $\boldsymbol{i} = (i_1, ..., i_t)$, $\boldsymbol{j} = (j_1, ..., j_t)$, $\boldsymbol{k} = (k_1, ..., k_t)$. A very general description of sequences $\boldsymbol{i}$, $\boldsymbol{j}$, $\boldsymbol{k}$ for which (P) holds is due to R.C. Thompson [11]. In that work the ring must be a principal ideal domain.

An important corollary of Thompson's work is that (P) holds when $k = i_u + j_u - u$, $1 \le u \le t$:

$$s_{i_1}(A) \cdots s_{i_t}(A)\, s_{j_1}(B) \cdots s_{j_t}(B) \mid s_{i_1+j_1-1}(AB) \cdots s_{i_t+j_t-t}(AB)$$

(the "standard" inequalities). For $t = 1$, this gives the well-known relation

$$s_i(A)\, s_j(B) \mid s_{i+j-1}(AB) ~.$$

The standard inequalities can also be proved using the Carlson–Sá characterizations (3)–(4) [4]. So they hold for matrices over elementary divisor domains.

REFERENCES

[1] Carlson, D. and Marques de Sá, E. – Generalized minimax and interlacing theorems, *Linear and Multilinear Algebra,* 15 (1984), 77–103.

[2] Helmer, O. – The elementary divisor theorem for certain rings without chain condition, *Bulletin AMS,* 49 (1943), 225–236.

[3] Kaplansky, I. – Elementary divisors and modules, *Transactions AMS,* 66 (1949), 464–491.

[4] Martins, R.A. and Queiró, J.F. – Spectral inequalities for generalized Rayleigh quotients, *Linear Algebra and its Applications,* 220 (1995), 311–320.

[5] Newman, M. – Equivalence without determinantal divisors, *Linear and Multilinear Algebra,* 7 (1979), 107–109.

[6] Pietsch, A. – s-Numbers of operators in Banach spaces, *Studia Mathematica,* 51 (1974), 201–223.

[7] Queiró, J.F. – Invariant factors as approximation numbers, *Linear Algebra and its Applications,* 49 (1983), 131–136.

[8] Marques de Sá, E. – Imbedding conditions for λ-matrices, *Linear Algebra and its Applications,* 24 (1979), 33–50.

[9] Thompson, R.C. – Interlacing inequalities for invariant factors, *Linear Algebra and its Applications,* 24 (1979), 1–32.

[10] Thompson, R.C. – The Smith invariants of a matrix sum, *Proceedings AMS,* 78 (1980), 162–164.

[11] Thompson, R.C. – Smith invariants of a product of integral matrices, in "Linear Algebra and its Role in Systems Theory", *Contemporary Mathematics,* 47 (1985), 401–435.

João Filipe Queiró,
Departamento de Matemática, Universidade de Coimbra,
Apartado 3008, 3000 Coimbra – PORTUGAL