Czechoslovak Mathematical Journal

Vlastimil Dlab

A note on pseudocongruent matrices

Czechoslovak Mathematical Journal, Vol. 12 (1962), No. 1, 104–109

Persistent URL: http://dml.cz/dmlcz/100499

Terms of use:

© Institute of Mathematics AS CR, 1962

Institute of Mathematics of the Czech Academy of Sciences provides access to digitized documents strictly for personal use. Each copy of any part of this document must contain these *Terms of use*.

This document has been digitized, optimized for electronic delivery and stamped with digital signature within the project *DML-CZ: The Czech Digital Mathematics Library* http://dml.cz

Чехословацкий математический журнал т. 12 (87) 1962, Прага

A NOTE ON PSEUDOCONGRUENT MATRICES

VLASTIMIL DLAB, Khartoum (Sudan)

(Received May 20, 1960)

In the present paper, a generalization of the concept of congruent matrices is introduced and the Sylvester "Law of inertia" is derived.

1. The importance of the transformation $A \to P^T AP$ in the ring of the matrices over a field (where P is a regular matrix and P^T its transpose), in many branches of mathematics, is well-known. By the equivalence $A \sim P^T AP$, all matrices are distributed into classes of congruent matrices; every class with a symmetric matrix contains a diagonal matrix. There are some relations among the diagonal matrices of the same class depending on the given field. Particularly, these matrices are in the case of an ordered field connected by the Sylvester "Law of inertia".

In the present paper, some generalizations of these relations to matrices over a ring are studied. Throughout the paper, capital letters (excepting R and V) denote square matrices over a given ring. In particular, the scalar matrix with an element d in the diagonal is denoted by $D = [d]$. By A^T, we shall denote the transpose of a matrix A.

2. Let R be a ring. For a given natural number k, denote by $V_k(R)$ the (right) module (over R) of all finite sequences of k elements from R: $\mathscr{A} = (a_1, a_2, \ldots, a_k)$. An element $\mathscr{A}_0 \in V_k(R)$ is said to be *linear dependent* on $\mathscr{A}_1, \mathscr{A}_2, \ldots, \mathscr{A}_m$ ($\mathscr{A}_i \in V_k(R)$ for $i = 1, 2, \ldots, m$) if a relation

$$\mathscr{A}_0 s_0 + \mathscr{A}_1 s_1 + \ldots + \mathscr{A}_m s_m = 0 \quad \text{with} \quad s_i \in R \quad \text{and} \quad s_0 \neq 0 \tag{2.1}$$

holds. In an obvious way, we shall define *linearly independent* and *dependent* sets of elements from $V_k(R)$.

By a *V-ring* we understand such a ring R for which every maximal independent set of $V_k(R)$ has precisely k elements (for every natural k). Thus, every V-ring has the following *property P*: For every pair r, s of elements of R, there exists a non-trivial solution of the equation $rx - sy = 0$.[1])

On the other hand, we are going to prove

Theorem 1. *A ring R without divisors of zero having the property P is a V-ring.*

[1]) Thus, every commutative ring has the property P in a trivial way.

Proof. Our assertion is a consequence of three "Basic Theorems" of B. L. VAN DER WAERDEN (see [1], p. 100). The first and the second of these theorems are obvious from our definition of linear dependence in $V_k(R)$. To prove the third one, let us suppose that $\mathscr{A}_0 \in V_k(R)$ is linearly dependent on $\mathscr{A}_1, \mathscr{A}_2, \ldots, \mathscr{A}_m$ and that each of $\mathscr{A}_i$ $(i = 1, 2, \ldots, m)$ is linearly dependent on $\mathscr{B}_1, \mathscr{B}_2, \ldots, \mathscr{B}_n$. Hence, we have, besides (2.1), the following relation for suitable elements t_{ij} $(i = 1, 2, \ldots, m;$ $j = 0, 1, \ldots, n)$ from R:

$$(2.2)_i \qquad \mathscr{A}_i t_{i0} + \mathscr{B}_1 t_{i1} + \ldots + \mathscr{B}_n t_{in} = 0 \quad \text{with} \quad t_{i0} \neq 0 .$$

Clearly, we may assume $s_i \neq 0$ for all $i = 1, 2, \ldots, m$. By our hypothesis, there exists a non-trivial solution u_1, v_1 of the equation $s_1 x - t_{10} y = 0$ in R: $s_1 u_1 - t_{10} v_1 = 0$. Since R is without divisors of zero, both elements are, obviously, non-zero. Multiplying the relation (2.1) by u_1 and substituting for $\mathscr{A}_1 s_1 u_1$ from $(2.2)_1$, we obtain

$$\mathscr{A}_0 s_0 u_1 + \mathscr{B}_1 t_{11} v_1 + \ldots + \mathscr{B}_n t_{1n} v_1 + \mathscr{A}_2 s_2 u_1 + \ldots + \mathscr{A}_m s_m u_1 = 0 .$$

Now, $s_2 u_1 \neq 0$ and we can proceed by induction: In general, denote by u_l, v_l a non-trivial solution of the equation $s_l u_1 u_2 \ldots u_{l-1} x - t_{l0} y = 0$ $(l = 2, 3, \ldots, m)$. It is easy to see that $u_l \neq 0$, $v_l \neq 0$ for each l. Then, we obtain finally the relation

$$\mathscr{A}_0 s_0 u_1 \ldots u_m + \mathscr{B}_1 (t_{11} v_1 u_2 \ldots u_m + t_{21} v_2 u_3 \ldots u_m + \ldots + t_{m1} v_m) + \ldots +$$
$$+ \mathscr{B}_n (t_{1n} v_1 u_2 \ldots u_m + t_{2n} v_2 u_3 \ldots u_m + \ldots + t_{mn} v_m) = 0 \quad \text{with} \quad s_0 u_1 \ldots u_m \neq 0.$$

This completes our proof.

Remark 1. The assumption of the absence of divisors of zero in Theorem 1 is quite natural. Furthermore, an example of a free ring shows that the other assumption of the property P cannot be also omitted.

3. Let $\mathfrak{M}_n(R)$ be the system of all matrices of order n over a ring R. Two matrices A and B from $\mathfrak{M}_n(R)$ are said to be *congruent* if $P, Q \in \mathfrak{M}_n(R)$ exist such that $P^T A P = B$ and $Q^T B Q = A$. A and B are said to be *pseudocongruent*, if $U, V \in \mathfrak{M}_n(R)$ exist such that $U^T A U = rBs$ and $V^T B V = uAv$ for suitable $r, s, u, v \in R$. Thus, all matrices of $\mathfrak{M}_n(R)$ are divided into classes of congruent matrices, and, if R is commutative, also into classes of pseudocongruent matrices.

We shall say that an operation on a matrix is *elementary* if it can be performed by the following simple operations:

(i) by the multiplication of a column by a non-zero element of R;
(ii) by the addition of a multiple of a column to another column;
(iii) by a transposition of two columns.

A matrix is said to be *D-elementary*, if it can be obtained from a non-zero scalar matrix by an elementary operation.[2])

[2]) It can be proved that in the case of a field the concepts of D-elementary and regular matrices coincide.

To the end of this section, R denotes a commutative ring satisfying $r + r \neq 0$ for every non-zero element $r \in R$.

Lemma 1. *To every symmetric matrix $A \in \mathfrak{M}_n(R)$, there exists a D-elementary matrix P such that $P^T AP$ is diagonal.*

Proof. Let $A = (a_{ij})$. Take a non-zero element $r \in R$ and consider the scalar matrix $D_0 = [r]$. Now, if the matrix $A^{(1)} = D_0^T A D_0 = (a_{ij}^{(1)})$ is diagonal, the proof is completed. Otherwise, let l be the natural number such that

$$a_{ij}^{(1)} = 0 \text{ for } i > j,\ j < l \text{ and } a_{i_0 l}^{(1)} \neq 0 \text{ at least for one } i_0 > l\,.$$

Denote by $A^{(2)} = (a_{ij}^{(2)})$ the matrix $A^{(1)}$, if $a_{ll}^{(1)} \neq 0$, or, when this is not the case, the matrix $D_1^T A^{(1)} D_1$, where D_1 is the matrix obtained from D_0 by the transposition of i_0-nd and l-th columns (if $a_{i_0 i_0}^{(1)} \neq 0$) or by adding the i_0-th column to the l-th one (if $a_{i_0 i_0}^{(1)} = 0$). Clearly, $a_{ij}^{(2)} = 0$ for $i > j$, $j < l$ and $a_{l}^{(2)} \neq 0$ (for, $a_{ll}^{(2)}$ is either $a_{ll}^{(1)}$ or $a_{i_0 i_0}^{(1)}$ or $a_{i_0 l}^{(1)} + a_{l i_0}^{(1)}$).

Consider further the matrix D_2 originated from D_1 by multiplying the $(l + 1)$-th, $(l + 2)$-th, ..., n-th columns by the same element $a_{ll}^{(2)}$ and, finally, the matrix D_3 which we obtain from D_2 by the addition of $(-a_{ml}^{(2)})$-multiple of the l-th column to the m-th one (for $m = l + 1, l + 2, \ldots, n$). D_3 is, obviously, a D-elementary matrix again. Then, as one can immediately see, the matrix $A^{(3)} = D_3^T A D_3 = (a_{ij}^{(3)})$ satisfies the following conditions: $a_{ij}^{(3)} = 0$ for $i > j$, $j < l + 1$. Now, we may complete the proof of our lemma very easily by induction.

Remark 2. The proof of Lemma 1 gives, simultaneously, a very advantageous method for the numerical calculation of a diagonal form of a symmetric matrix and of the matrix of the corresponding transformation. In the case of a field, this method may be arranged in an obvious way into a much more suitable form.

Lemma 2. *Let $P^T AP = B$ with a D-elementary matrix P. Then a D-elementary matrix Q exists such that $Q^T BQ = r^2 A$ for a suitable $r \in R$.*

Proof. First, transform the relation $P^T AP = B$ by means of a scalar matrix $D = [d]$, $d \neq 0$: $d^2 P^T AP = D^T BD$. It is obvious that to prove our assertion it suffices only to prove the following proposition:

If $s^2 K^T AK = M^T BM$, where K is a matrix obtained from a matrix L by a simple operation of the form (i) or (ii) or (iii), M a D-elementary matrix and $s \in R$, then there exists a D-elementary matrix N such that $t^2 L^T AL = N^T BN$ for a suitable $t \in R$.

Suppose that K is obtained from L by an operation of the type (i), i. e., say, by the multiplication of the i_0-th column of L by $u \in R$. Then, multiplying all columns, with the exception of the i_0-th, of M by the same element u and denoting the D-elementary matrix thus obtained by N, we have readily obtained $s^2 u^2 L^T AL = N^T BN$.

In case that K is obtained from L by an operation of the type (ii) or (iii), L can be, on the contrary, obtained from K by the same operation. This operation, applied to the matrix M, gives us the desired matrix N.

The proof of our proposition, and, thus, also of Lemma 2, is completed.

Lemmas 1 and 2 then immediately imply

Theorem 2. *In each class of pseudocongruent matrices over a commutative ring satisfying $r + r \neq 0$ for every its non-zero element r which contains a symmetric matrix there exists a diagonal matrix.*

4. Now, we are going to give a proof of the Sylvester "Law of inertia". First, prove the following

Lemma 3. *Let R_0 be an ordered V-ring and A, B two diagonal matrices from $\mathfrak{M}_n(R)$ such that*

$$C^T AC = rBs \text{ for suitable } r, s \in R_0 \text{ and } C = (c_{ij}) \text{ from } \mathfrak{M}_n(R_0)\,. \tag{4.1}$$

Denote by k_1, k_2 the numbers of positive, and by l_1, l_2, the numbers of negative elements of A and B, respectively. If rs is positive, then $k_1 \geqq k_2$ and $l_1 \geqq l_2$. If rs is negative, then $k_1 \geqq l_2$ and $l_1 \geqq k_2$.

Proof. Let rs be positive. We are going to prove the inequality $k_1 \geqq k_2$. The other statements can be proved in a similar way. It is easy to see that we can assume (without loss of generality) that a_{pp} for $p = 1, 2, \ldots, k_1$ and b_{qq} for $q = 1, 2, \ldots, k_2$ are all positive elements of A and B, respectively.

Suppose that $k_1 < k_2$. Consider the following vectors of $V_k(R_0)$, the components of which are the elements of the matrix C: $\mathscr{V}_m = (c_{1m}, c_{2m}, \ldots, c_{k_1 m})$ for $m = 1, 2, \ldots, \ldots, k_1 + 1$. Then, by our hypothesis, there exist suitable elements $t_m \in R_0$ $(m = 1, 2, \ldots, k_1 + 1)$ such that at least one of them is non-zero and

$$\sum_{m=1}^{k_1+1} c_{pm} t_m = 0 \quad \text{for} \quad p = 1, 2, \ldots, k_1\,. \tag{4.2}$$

Denote by $F = (f_{ij})$ the matrix defined as follows:

$$f_{m,k_1+1} = t_m \text{ for } m = 1, 2, \ldots, k_1 + 1 \text{ and } f_{ij} = 0 \text{ otherwise.} \tag{4.3}$$

Further, denote by $G = (g_{ij})$ the product

$$G = CF\,. \tag{4.4}$$

Then, from (4.1), $G^T AG = F^T rBsF$. By (4.4), (4.3) and (4.2) we have

$$g_{p,k_1+1} = \sum_{m=1}^{n} c_{pm} f_{m,k_1+1} = \sum_{m=1}^{k_1+1} c_{pm} t_m = 0 \quad \text{for each} \quad p = 1, 2, \ldots, k_1.$$

Hence, at the position $(k_1 + 1, k_1 + 1)$ of the matrix $G^T AG$ we obtain the element

$$\sum_{i=1}^{n} g_{i,k_1+1} a_{ii} g_{i,k_1+1} = \sum_{i=1}^{n} \Big(\sum_{m=1}^{k_1+1} c_{im} t_m\Big) a_{ii} \Big(\sum_{m=1}^{k_1+1} c_{im} t_m\Big) = \sum_{i=k_1+1}^{n} \Big(\sum_{m=1}^{k_1+1} c_{im} t_m\Big) a_{ii} \Big(\sum_{m=1}^{k_1+1} c_{im} t_m\Big),$$

which is, obviously, non-positive.

On the other hand, in the same $(k_1 + 1, k_1 + 1)$ position of the matrix $F^T rBsF$ we obtain

$$\sum_{m=1}^{n} f_{m,k_1+1} r b_{mm} s f_{m,k_1+1} = \sum_{m=1}^{k_1+1} t_m r b_{mm} s t_m ;$$

that is, clearly, a positive element. This contradiction concludes the proof of our lemma.

From Lemma 3, we deduce immediately

Theorem 3. *The numbers of positive, resp. negative elements of two pseudocongruent diagonal matrices over an ordered V-ring are equal.*

Finally, using both Theorems 1 and 2, we may formulate

Corollary 1. *Let R_0 be an ordered ring with the property P. Then the numbers of positive, resp. negative elements of two pseudocongruent diagonal matrices of $\mathfrak{M}_n(R_0)$ are equal.*

Corollary 2. *Let R_0 be a commutative ordered ring. Then, in every class of pseudocongruent matrices of $\mathfrak{M}_n(R_0)$ which contains a symmetric matrix there exist diagonal matrices. The numbers of positive, resp. negative elements of all these matrices are equal.*

Bibliography

[1] *B. L. van der Waerden:* Modern Algebra I., New York 1949.

Резюме

ЗАМЕТКА О ПСЕВДОКОНГРУЭНТНЫХ МАТРИЦАХ

ВЛАСТИМИЛ ДЛАБ (Vlastimil Dlab), Хартум (Судан)

Пусть дано кольцо R. Обозначим через $V_k(R)$, где k — данное натуральное число, (правый) модуль (над R) всех конечных последовательностей k элементов из R: $\mathscr{A} = (a_1, a_2, \ldots, a_k)$. Мы скажем, что элемент $\mathscr{A}_0 \in V_k(R)$ *линейно зависит* от $\mathscr{A}_1, \mathscr{A}_2, \ldots, \mathscr{A}_m$ ($\mathscr{A}_i \in V_k(R)$ для $i = 1, 2, \ldots, m$), если существует соотношение

$$\mathscr{A}_0 s_0 + \mathscr{A}_1 s_1 + \ldots + \mathscr{A}_m s_m = 0 \quad \text{с} \quad s_i \in R \quad \text{и} \quad s_0 \neq 0 .$$

Очевидным образом мы определим *линейно независимые и зависимые множества* элементов из $V_k(R)$.

Кольцо R, для которого справедливо утверждение, что каждое максимальное независимое множество из $V_k(R)$ имеет в точности k элементов (для каждого натурального k), назовем *V-кольцом*. Теорема 1 дает достаточные условия для того, чтобы кольцо было V-кольцом:

Теорема 1. *Кольцо R без делителей нуля, в котором каждое уравнение $rx - sy = 0$ $(r,s \in R)$ имеет нетривиальное решение x, y, является V-кольцом.*

Пусть $\mathfrak{M}_n(R)$ означает систему всех (квадратных) матриц порядка n над R. Две матрицы A, B из $\mathfrak{M}_n(R)$ мы назовем *псевдоконгруэнтными*, если в $\mathfrak{M}_n(R)$ существуют матрицы U, V так, что $U^TAU = rBs$ и $V^TBV = uAv$ для подходящих r, s, u, $v \in R$.[1]) Итак, если R коммутативно, то матрицы из $\mathfrak{M}_n(R)$ распадаются на классы взаимно псевдоконгруэнтных матриц; кроме того справедлива

Теорема 2. *Пусть класс псевдоконгруэнтных матриц над коммутативным кольцом, удовлетворяющим неравенству $r + r \neq 0$ для любого ненулевого элемента r, содержит какую-либо симметрическую матрицу. Тогда в каждом таком классе имеется диагональная матрица.*

Теорема 3 выражает в обобщенном виде „закон инерции" Сильвестра:

Теорема 3. *Число положительных, соотв., отрицательных элементов двух псевдоконгруэнтных диагональных матриц над упорядоченным V-кольцом одинаково.*

Из теорем 1 – 3 тогда непосредственно получим

Следствие 1. *Пусть R_0 – упорядоченное кольцо, в котором каждое уравнение $rx - sy = 0$ $(r, s \in R)$ обладает нетривиальным решением x, y. Тогда число положительных, соотв., отрицательных элементов двух псевдоконгруэнтных диагональных матриц из $\mathfrak{M}_n(R_0)$ одинаково.*

Следствие 2. *Пусть R_0 – коммутативное упорядоченное кольцо. Тогда в каждом классе взаимно псевдоконгруэнтных матриц из $\mathfrak{M}_n(R)$, содержащем какую-либо симметрическую матрицу, существуют диагональные матрицы. Число положительных, соотв., отрицательных элементов в каждой из этих матриц одинаково.*

---

[1]) Символ P^T означает матрицу, транспонированную по отношению к матрице P.