ANNALES SOCIETATIS MATHEMATICAE POLONAE
Series I: COMMENTATIONES MATHEMATICAE XX (1977)
ROCZNIKI POLSKIEGO TOWARZYSTWA MATEMATYCZNEGO
Seria I: PRACE MATEMATYCZNE XX (1977)

JERZY K. BAKSALARY and RADOSŁAW KALA (Poznań)

A method of finding bases of a matrix

Abstract. The notion of a row-semireduced form of a matrix is introduced and an algorithm of transforming any given matrix to such a form is presented. A theorem is established which shows how, by a row-semireduction, bases of the row space and of the column space of any given matrix A can be found, such that consist of rows and columns of A, respectively. The proposed method requires a smaller number of operations than algorithms known in the literature.

1. Let $\mathscr{F}^{m\times n}$ be the class of $m\times n$ matrices over a field $\mathscr{F}$. For any given $\boldsymbol{A}\in\mathscr{F}^{m\times n}$ we denote by $\mathscr{R}(\boldsymbol{A})$ and $\mathscr{C}(\boldsymbol{A})$ the row space and the column space of $\boldsymbol{A}$, respectively.

The purpose of this paper is to find bases of the subspaces $\mathscr{R}(\boldsymbol{A})$ and $\mathscr{C}(\boldsymbol{A})$ consisting of rows and columns of the given $\boldsymbol{A}$, respectively. This question arises in such algebraical problems as solving a system of linear equations with rectangular or square singular matrix or finding a generalized inverse of a matrix. This question is also relevant to the theory of statistical linear models.

Several methods of determining the bases of $\mathscr{R}(\boldsymbol{A})$ and $\mathscr{C}(\boldsymbol{A})$ are already known (see, e.g., [1], p. 15, [2], p. 422, [4], p. 45–46) but not all of them give bases that consist of rows and columns of $\boldsymbol{A}$, respectively. In most cases, these methods are based on a transformation of a given matrix to the so-called row-echelon form. In this paper, the notion of a row-semireduced form of a matrix is introduced. It is shown that such a form is more convenient for the purpose of the present paper and, moreover, is obtainable by a smaller number of operations than the row-echelon form.

2. To emphasize the difference between the notions of a row-echelon form and a row-semireduced form of a matrix we recall the first of them in the version given in [3], p. 18, which is the weakest from various definitions available in the literature.

DEFINITION 1. A matrix $\boldsymbol{R}\in\mathscr{F}^{m\times n}$ is called a *row-echelon matrix* (also sometimes called a *row-reduced matrix*) if

(a) the leftmost non-zero element, called a *pivot*, in each non-zero row of $\boldsymbol{R}$ is 1,

(b) each column that contains a pivot of some non-zero row has 0 as each of its other elements.

Modifying condition (b) in the above definition, we get the following.

DEFINITION 2. A matrix $\boldsymbol{S}\in\mathscr{F}^{m\times n}$ is called a *row-semireduced matrix* if

(a) a pivot in each non-zero row of $\boldsymbol{S}$ is 1,

(b) each column containing a pivot of some non-zero row has 0 as each of its elements that appears below the pivot.

From the comparison of these two definitions it is evident that each row-echelon matrix is a row-semireduced matrix or, in other words, that the second notion is weaker than the first one.

3. We now show that any matrix $\boldsymbol{A}\in\mathscr{F}^{m\times n}$ can be transformed to a row-semireduced matrix $\boldsymbol{S}\in\mathscr{F}^{m\times n}$ by applying the following algorithm based on the elementary row operations (e_1) and (e_2), where the first consists in the multiplication of a given row by a non-zero scalar, and the second in the addition to a given row of a scalar multiple of another row.

Step 0. Place a given matrix $\boldsymbol{A}$ under $\boldsymbol{S}$. Go to Step 1 with the whole $\boldsymbol{S}$ as its submatrix.

Step 1. Find the uppermost non-zero row of a submatrix $\boldsymbol{S}$ (if there is none, $\boldsymbol{S}$ is in a row-semireduced form). Find a pivot of that row. If the pivot is $\neq 1$, apply a type (e_1) operation to change it to 1. Apply a sequence of type (e_2) operations to reduce to 0 all elements below the pivotal 1.

Step 2. Repeat Step 1 on the submatrix of $\boldsymbol{S}$ (in the form resulting from Step 1) that consists of all rows below that whose pivot has just been considered (if that row has been the last row of $\boldsymbol{S}$, $\boldsymbol{S}$ is in a row-semireduced form).

By the description of the algorithm it is clear that any $\boldsymbol{A}\in\mathscr{F}^{m\times n}$ can be transformed to a row-semireduced form by a finite sequence of elementary row operations. But, on the other hand, it is known (see, e.g., [3], p. 17–18) that each such operation can be viewed as the premultiplication of the transformed matrix by an elementary row matrix, which is non-singular and has only one row that is different from the corresponding row of the unit matrix $\boldsymbol{I}\in\mathscr{F}^{m\times m}$. Therefore, denoting by $\boldsymbol{E}$ the product of all elementary row matrices which are related to the elementary row operations performing the row-semireduction of $\boldsymbol{A}$, we get

$$\boldsymbol{EA} = \boldsymbol{S}, \tag{1}$$

where evidently $\boldsymbol{E}\in\mathscr{F}^{m\times m}$ is non-singular as a product of non-singular matrices.

4. Properties of a row-semireduced matrix relevant to the purpose of the present paper are given in the following

THEOREM 1. *Let* $\boldsymbol{S}$ *be any row-semireduced matrix. Then*

(i) *the non-zero rows of* $\boldsymbol{S}$ *form a basis for* $\mathscr{R}(\boldsymbol{S})$,

(ii) *the columns of* $\boldsymbol{S}$ *that contain the pivots form a basis for* $\mathscr{C}(\boldsymbol{S})$.

Proof. Suppose that $\boldsymbol{S} \in \mathscr{F}^{m\times n}$ has k non-zero rows and denote by $\boldsymbol{P}$, $\boldsymbol{P} \in \mathscr{F}^{m\times m}$, the permutation matrix such that

$$\boldsymbol{PS} = \begin{bmatrix} \boldsymbol{S}_{1\cdot} \\ \boldsymbol{0} \end{bmatrix}, \tag{2}$$

where $\boldsymbol{S}_{1\cdot} \in \mathscr{F}^{k\times n}$ consists of all non-zero rows of $\boldsymbol{S}$. From the algorithm it is clear that suitable permutations of rows and columns of $\boldsymbol{S}_{1\cdot}$ yield as its leftmost $k\times k$ minor an upper triangular matrix, $\boldsymbol{S}_{11}$ say, with 1 as each diagonal element. Hence it follows that $\boldsymbol{S}_{1\cdot}$ is of full row rank. Thus, the rows of $\boldsymbol{S}_{1\cdot}$ are linearly independent and, in view of the form (2), form a basis for $\mathscr{R}(\boldsymbol{PS})$. But $\boldsymbol{P}$ is non-singular, so they also form a basis for $\mathscr{R}(\boldsymbol{S})$. This completes the proof of (i).

Let now $\boldsymbol{Q}$, $\boldsymbol{Q} \in \mathscr{F}^{n\times n}$, be the permutation matrix such that

$$\boldsymbol{SQ} = [\boldsymbol{S}_{\cdot 1} \vdots \boldsymbol{S}_{\cdot 2}], \tag{3}$$

where $\boldsymbol{S}_{\cdot 1}$ consists of all the k columns that contain the pivots, while $\boldsymbol{S}_{\cdot 2}$ of all the other $n-k$ columns. It is seen that also $\boldsymbol{S}_{\cdot 1}$ has the minor $\boldsymbol{S}_{11} \in \mathscr{F}^{k\times k}$ specified in the first part of the proof, although, may be, in a permuted form. This implies that $\boldsymbol{S}_{\cdot 1}$ is of full column rank and so consists of linearly independent columns. On the other hand, it is known that $\dim \mathscr{C}(\boldsymbol{S}) = \dim \mathscr{R}(\boldsymbol{S})$. But $\dim \mathscr{R}(\boldsymbol{S}) = k$ and, in view of the non-singularity of $\boldsymbol{Q}$, $\mathscr{C}(\boldsymbol{S}) = \mathscr{C}(\boldsymbol{SQ})$. Thus, the k columns of $\boldsymbol{S}_{\cdot 1}$ are a basis of $\mathscr{C}(\boldsymbol{S})$ and the theorem is established.

5. The main result of this paper is given in

THEOREM 2. *Let* $\boldsymbol{S}$ *be a row-semireduced matrix obtained from a given matrix* $\boldsymbol{A} \in \mathscr{F}^{m\times n}$. *If the* r_1*-th*, r_2*-th*, ..., r_k*-th rows of* $\boldsymbol{S}$ *are all its non-zero rows and the* c_1*-th*, c_2*-th*, ..., c_k*-th columns of* $\boldsymbol{S}$ *are those containing the pivots, then*

(i) *the* r_1*-th*, r_2*-th*, ..., r_k*-th rows of* $\boldsymbol{A}$ *form a basis for* $\mathscr{R}(\boldsymbol{A})$,

(ii) *the* c_1*-th*, c_2*-th*, ..., c_k*-th columns of* $\boldsymbol{A}$ *form a basis for* $\mathscr{C}(\boldsymbol{A})$.

Proof. First we show part (ii). Postmultiplying both sides of (1) by the permutation matrix $\boldsymbol{Q}$ specified by (3), we get

$$\boldsymbol{EAQ} = \boldsymbol{SQ}.$$

Hence, by (3) and the non-singularity of $\boldsymbol{E}$,

$$[\boldsymbol{A}_{\cdot 1} \vdots \boldsymbol{A}_{\cdot 2}] = \boldsymbol{E}^{-1}[\boldsymbol{S}_{\cdot 1} \vdots \boldsymbol{S}_{\cdot 2}], \tag{4}$$

where $A_{\cdot 1}$ is the submatrix of A consisting of its c_1-th, c_2-th, ..., c_k-th columns. Theorem 1 states that the columns of $S_{\cdot 1}$ are a basis for $\mathscr{C}(S)$, so there exists a matrix $T \in \mathscr{F}^{k\times(n-k)}$ such that $S_{\cdot 2} = S_{\cdot 1}T$. Now, by (4),

$$A_{\cdot 1} = E^{-1}S_{\cdot 1} \tag{5}$$

and

$$A_{\cdot 2} = E^{-1}S_{\cdot 1}T. \tag{6}$$

From (5) it follows immediately that $A_{\cdot 1}$ is of full column rank, i.e., that its columns are linearly independent. On the other hand, using (5) in (6) gives $A_{\cdot 2} = A_{\cdot 1}T$, and so the columns of $A_{\cdot 2}$ are linear combinations of those in $A_{\cdot 1}$. Thus, the columns of $A_{\cdot 1}$ form a basis for $\mathscr{C}(A)$ and, therefore, (ii) is established.

To prove (i) we premultiply (1) by the permutation matrix P specified by (2). Since P is orthogonal, we can write

$$PEP'PA = PS. \tag{7}$$

The form of PS is given by (2), while PA can be expressed as

$$PA = \begin{bmatrix} A_{1\cdot} \\ A_{2\cdot} \end{bmatrix}, \tag{8}$$

where $A_{1\cdot} \in \mathscr{F}^{k\times n}$ consists of rows with the same indices as the rows of $S_{1\cdot}$, i.e., of the r_1-th, r_2-th, ..., r_k-th rows of A. Now let investigate the form of PEP'. It is known that E, defined as a product of elementary row matrices, can also be obtained by performing on $I \in \mathscr{F}^{m\times m}$ the same sequence of elementary row operations that transforms A to a row-semi-reduced form S. By the description of the algorithm it is seen that E differs from I in these columns only which correspond to non-zero rows of S. More precisely, all columns in E with indices different from $r_1, r_2, \ldots, r_k$ are some m-dimensional unit vectors. Therefore, since premultiplying E by P places the r_1-th, r_2-th, ..., r_k-th rows of E on the first k positions and postmultiplying PE by P' performs the analogous permutation of columns, the matrix PEP' can be written as

$$PEP' = \begin{bmatrix} E_{11} & O \\ E_{21} & I \end{bmatrix}, \tag{9}$$

where $E_{11} \in \mathscr{F}^{k\times k}$ and $I \in \mathscr{F}^{(m-k)\times(m-k)}$.

Now, applying (2), (8) and (9) to (7), we obtain two matrix equalities, one of which is

$$E_{11}A_{1\cdot} = S_{1\cdot}.$$

Hence, by (9) and the non-singularity of PEP', E_{11} is non-singular and so $A_{1\cdot} = E_{11}^{-1}S_{1\cdot}$. But, by Theorem 1, the rows of $S_{1\cdot}$ are a basis for $\mathscr{R}(S)$.

Thus, also the rows of $\boldsymbol{A}_{1\cdot}$ form a basis for $\mathscr{R}(\boldsymbol{S})$. Furthermore, since $\boldsymbol{E}$ is non-singular, (1) implies $\mathscr{R}(\boldsymbol{S}) = \mathscr{R}(\boldsymbol{A})$ and therefore the rows of $\boldsymbol{A}_{1\cdot}$ are also a basis for $\mathscr{R}(\boldsymbol{A})$. This completes the proof.

6. The authors are very indebted to Professor T. Caliński for his kind remarks concerning the preparation of the paper.

References

[1] A. Ben-Israel and T. N. E. Greville, *Generalized inverses: theory and applications*, Wiley, New York 1974.

[2] L. D. Pyle, *Generalized inverse computations using the gradient projection method*, J. ACM 11 (1964), p. 422–428.

[3] C. R. Rao, *Linear statistical inference and its applications*, 2-nd ed., Wiley, New York 1973.

[4] R. R. Stoll and E. T. Wong, *Linear algebra*, Academic Press, New York 1968.

DEPARTMENT OF STATISTICS, ACADEMY OF AGRICULTURE, POZNAŃ