Miroslav Fiedler; Vlastimil Pták

Some results on matrices of class K and their application to the convergence rate of iteration procedures

Czechoslovak Mathematical Journal, Vol. 16 (1966), No. 2, 260–273

Чехословацкий математический журнал, т. 16 (91) 1966, Прага

SOME RESULTS ON MATRICES OF CLASS K AND THEIR APPLICATION TO THE CONVERGENCE RATE OF ITERATION PROCEDURES

MIROSLAV FIEDLER and VLASTIMIL PTÁK, Praha

(Received May 12, 1965)

Introduction. The present paper represents a continuation of the authors' series of communications concerning matrices of type **K** and their applications to spectral problems. The paper is divided into three sections, the first section being devoted to a recapitulation of some definitions and terminological conventions. The new results on matrices of class **K** are collected in section two. Especially, we present improvements of two theorems of the first paper [2] of the series. Theorems (2,5) and (2,6) of the present paper constitute a quantitative sharpening of theorem (4,6) of [2]. Theorem (2,10) is a considerable improvement of theorem (6,7) of [2] in that it gives conditions under which the new matrix can be singular.

As an illustration, section 3 contains theorems which are closely connected with convergence theorems in relaxation methods. Theorem (3,3) recalls – under appropriate assumptions – the monotonous dependence of the convergence rate on the choice of the matrix B in the iteration formula $x_{n+1} = B^{-1}(B - A)\,x_n + B^{-1}b$ for the solution of $Ax = b$. This theorem was proved in [1] for A symmetric, R. S. VARGA [4] generalized this result for the non-symmetric case. Theorem (3,4) shows that analogous estimates to those obtained by Varga [5] are valid for a more general class of Gauss-Seidel procedures.

1. Definitions and notation. In the whole paper, n will be a fixed natural number. The set of all natural numbers $\leqq n$ will be denoted by N. A matrix is a real function on $N \times N$, the value of a matrix A at the point (i, k) being denoted by a_{ik}. A matrix A is said to be nonnegative if $a_{ik} \geqq 0$ for each i and k. In this case, we write simply $A \geqq 0$. The (unique) nonnegative proper value of a nonnegative matrix A which has the greatest modulus of all proper values of A will be called Perron root of A and denoted by $p(A)$.

A matrix A is said to be diagonal if $a_{ik} = 0$ for $i \neq k$. Such a matrix will be denoted by diag $(a_{11}, a_{22}, \ldots, a_{nn})$. A positive diagonal matrix is a diagonal matrix with $a_{ii} > 0$ for all i.

The spectral radius of a matrix A is the maximum of the moduli of the proper values of A and will be denoted by $|A|_\sigma$. In accordance with common usage we shall, sometimes, drop the unit matrix in expressions like $\lambda E - A$.

We shall denote by $\mathbf{Z}$ the class of all matrices A for which $a_{ik} \leqq 0$ for $i \neq k$. The subclass of $\mathbf{Z}$ consisting of all matrices $A \in \mathbf{Z}$ which have all principal minors positive will be called $\mathbf{K}$, the subclass of all matrices $A \in \mathbf{Z}$ which have all principal minors nonnegative will be denoted by $\mathbf{K}_0$. The matrices which belong to $\mathbf{K}$ are usually called M-matrices by various authors. The paper [2] presented by the authors is devoted to the study of both the important classes $\mathbf{K}$ and $\mathbf{K}_0$ and contains a whole series of equivalent characterizations of matrices in $\mathbf{K}$ or $\mathbf{K}_0$. Since we shall repeatedly use different results on matrices of these types contained in [2], it will be convenient to simplify references to this paper in using the symbol 2 to denote results of [2]. Thus, theorem (2; 2,3) will be theorem (2,3) of [2] whereas (2,3) is theorem (2,3) of the present paper.

Finally, we recall the following notation from [2]. If A is a matrix in $\mathbf{K}$ or $\mathbf{K}_0$, we denote by $q(A)$ the (unique) nonnegative proper value of A which has the smallest modulus of all proper values of A.

2. In this section, we shall prove some theorems on nonnegative matrices, and on matrices of classes $\mathbf{K}$ and $\mathbf{K}_0$.

(2,1) *A matrix A belongs to $\mathbf{K}$ if and only if it may be written in the form $A = \lambda - P$ where P is nonnegative and $\lambda > p(P)$. Similarly, A belongs to $\mathbf{K}_0$ if and only if it may be written in the form $A = \lambda - P$ where P is nonnegative and $\lambda \geqq p(P)$.*

Proof. Suppose that $A \in \mathbf{K}$. Clearly there exists a $\lambda > 0$ such that $P = \lambda - A \geqq 0$. The number $\lambda - p(P)$ is a real proper value of A whence $\lambda - p(P) > 0$ according to (2; 4,3). On the other hand, if a matrix $\tau - P$ is given where $P \geqq 0$ and $\tau > p(P)$, we have $\tau > |P|_\sigma$ so that $(\tau - P)^{-1} = E + P + P^2 + \ldots$ exists and is nonnegative. Hence $\tau - P$ belongs to $\mathbf{K}$ by (2;4,3). The statement about matrices of type $\mathbf{K}_0$ may be obtained in an analogous manner or follows directly from (2;5,1).

(2,2) *Let M and S be two nonnegative matrices such that $m_{ii} > 0$ and S is symmetric. Then $p(MS) = 0$ implies $S = 0$.*

Proof. The matrix $A = MS$ is nonnegative and $p(A) = 0$. It follows from the theory of nonnegative matrices that there exists a permutation matrix P such that $B = PAP^{-1}$ is a matrix with $b_{ik} = 0$ for $i \leqq k$. If $\tilde{M} = PMP^{-1}$ and $\tilde{S} = PSP^{-1}$, we have for $i \leqq k$

$$\tilde{m}_{ii}\tilde{s}_{ik} \leqq \sum_r \tilde{m}_{ir}\tilde{s}_{rk} = b_{ik} = 0$$

so that $\tilde{s}_{ik} = 0$, the number $\tilde{m}_{ii}$ being clearly positive. Since $\tilde{S}$ is symmetric, this means that $\tilde{S} = 0$ which implies $S = 0$.

(2,3) *Let* $0 \leqq A \leqq B$ *and suppose that* $p(A) = p(B)$. *If* A *is irreducible then* $A = B$.

Proof. Suppose that A is irreducible. If $n = 1$, the result is obvious. If $n \geqq 2$, B is irreducible as well, we have $p(A) > 0$ and there exist positive vectors x and y such that $Ax = p(A)\, x$ and $y'B = p(B)\, y'$. We have thus

$$p(A)\, y'x = y'Ax \leqq y'Bx = p(B)\, y'x = p(A)\, y'x$$

whence $y'Ax = y'Bx$. Both vectors y and x being positive, this implies $A = B$.

(2,4) *Let* $P \leqq Q$ *and suppose that both* P *and* Q *belong to* $\mathbf{K}_0$. *If* Q *is singular then so is* P. *Moreover, if* Q *is irreducible then* Q *singular implies* $P = Q$.

Proof. Suppose that $P \in \mathbf{K}_0$, $Q \in \mathbf{K}_0$ and $P \leqq Q$. If P is nonsingular, we have $P \in \mathbf{K}$ by (2;5,5) and it follows from (2;4,6) that $Q \in \mathbf{K}$ as well. This proves the first assertion. Suppose now that Q is singular. There exists an $\alpha > 0$ such that both matrices $A = \alpha E - Q$ and $B = \alpha E - P$ are nonnegative. It follows from (2;5,1) that $\alpha = p(A) = p(B)$.

We have thus $A \leqq B$ and $p(A) = p(B)$; if Q is irreducible then A is irreducible as well so that, by (2,3), we have $A = B$ whence $P = Q$.

(2,5) *Let* $A \in \mathbf{K}$. *If* $B \geqq A$ *and* $B \in \mathbf{Z}$ *then*

1° $B \in \mathbf{K}$,

2° $0 \leqq B^{-1} \leqq A^{-1}$,

3° $\det B \geqq \det A > 0$,

4° $A^{-1}B \geqq E$ *and* $BA^{-1} \geqq E$,

5° $E \geqq B^{-1}A$ *and* $E \geqq AB^{-1}$ *and both matrices* $B^{-1}A$ *and* AB^{-1} *belong to* $\mathbf{K}$,

6° $1 - p(E - B^{-1}A) = 1 - p(E - AB^{-1}) = \dfrac{1}{p(A^{-1}B)} = \dfrac{1}{p(BA^{-1})}$,

7° $q(B) \geqq q(A)$.

Proof. If $B \in \mathbf{Z}$ and $B \geqq A$, the matrix $\tau E - B$, and hence also $\tau E - A$, will be nonnegative for a suitable positive τ. Since $A = \tau E - (\tau E - A)$, the number $\tau - p(\tau E - A)$ is a proper value of A so that $\tau - p(\tau E - A)$ is positive by 7° of (2;4,3). We have $0 \leqq \tau E - B \leqq \tau E - A$ whence $p(\tau E - B) \leqq p(\tau E - A) < \tau$. It follows that both the series

$$E + \left(E - \frac{1}{\tau} B\right) + \left(E - \frac{1}{\tau} B\right)^2 + \dots,$$

$$E + \left(E - \frac{1}{\tau} A\right) + \left(E - \frac{1}{\tau} A\right)^2 + \dots$$

are convergent. The first series converges to $(1/\tau) \cdot B^{-1}$, the second series to $(1/\tau) \cdot A^{-1}$ It follows that $0 \leqq B^{-1} \leqq A^{-1}$. This proves 2°; further it follows from 11° of (2;4,3) that $B \in \mathbf{K}$. The inequalities in 4° and 5° may be obtained upon multiplying $B - A \geqq 0$ by the nonnegative matrices A^{-1} and B^{-1}. Since $E \geqq B^{-1}A$ and $E \geqq AB^{-1}$, we have $B^{-1}A \in \mathbf{Z}$ and $AB^{-1} \in \mathbf{Z}$. Further, these matrices have inverses $A^{-1}B$ and BA^{-1} which are nonnegative by 4°. It follows that both $B^{-1}A$ and AB^{-1} belong to $\mathbf{K}$. To prove 6°, let us note first that the matrices $B^{-1}A$ and AB^{-1} are similar so that it suffices to prove $1 - p(E - B^{-1}A) = 1/p(A^{-1}B)$. If we write λ for $p(E - B^{-1}A)$, it follows that $1 - \lambda$ is a proper value of $B^{-1}A$. Since $B^{-1}A \in \mathbf{K}$, the number $1 - \lambda$ is positive according to 7° of (2;4,3). We intend to show now that $1/(1 - \lambda)$ is the Perron root of $A^{-1}B$. Indeed, $1/(1 - \lambda)$ is a proper value of $A^{-1}B = (B^{-1}A)^{-1}$. If $\mu > 1/(1 - \lambda)$, we may write $\mu = 1/(1 - \sigma)$ for a suitable $\sigma > \lambda$. It follows that

$$\mu E - A^{-1}B = \frac{1}{1 - \sigma} E - A^{-1}B = \frac{1}{1 - \sigma} A^{-1}B(B^{-1}A - (1 - \sigma) E) =$$

$$= \frac{1}{1 - \sigma} A^{-1}B(\sigma E - (E - B^{-1}A))$$

and the last matrix is nonsingular since $\sigma > \lambda = p(E - B^{-1}A)$.

To prove 7°, it is sufficient to show that $\lambda E - B$ is nonsingular if $\lambda < q(A)$. But in this case $\alpha - \lambda \geqq q(A) - \lambda > 0$ for each real proper value α of A so that $A - \lambda E \in \in \mathbf{K}$ by 7° of (2;4,3). Since $B - \lambda E \geqq A - \lambda E$ and $B - \lambda E \in \mathbf{Z}$, $B - \lambda E \in \mathbf{K}$ and thus nonsingular. The proof is complete.

(2,6) *Let $M \in \mathbf{K}$. Suppose we are given two matrices B_1 and B_2 which satisfy*

$$B_2 \geqq B_1 \geqq M\,.$$

If $B_2 \in \mathbf{Z}$, then both B_2 and B_1 belong to $\mathbf{K}$. Further, both $B_2^{-1}M$ and $B_1^{-1}M$ belong to $\mathbf{K}$ and

$$0 \leqq p(B_1^{-1}(B_1 - M)) \leqq p(B_2^{-1}(B_2 - M)) < 1\,.$$

Proof. The inclusions $B_2 \in \mathbf{K}$ and $B_1 \in \mathbf{K}$, $B_2^{-1}M \in \mathbf{K}$ and $B_1^{-1}M \in \mathbf{K}$ follow immediately from the preceding theorem. Clearly it suffices to prove

$$0 < 1 - p(B_2^{-1}(B_2 - M)) \leqq 1 - p(B_1^{-1}(B_1 - M)) \leqq 1\,.$$

According to 6° of the preceding theorem, we have

$$1 - p(E - B_2^{-1}M) = \frac{1}{p(M^{-1}B_2)} \leqq \frac{1}{p(M^{-1}B_1)} = 1 - p(E - B_1^{-1}M)\,.$$

Together with the obvious facts $1/p(M^{-1}B_2) > 0$ and $p(E - B_1^{-1}M) \geqq 0$ this yields the desired inequalities.

(2,7) *Let* $A \in \mathbf{K}$, $B \in \mathbf{K}$ *and suppose that* $AB \in \mathbf{Z}$. *Then* $AB \in \mathbf{K}$.

Proof. We use condition 11° of (2;4,3). Since A and B belong to $\mathbf{K}$, they are both nonsingular and $A^{-1} \geqq 0$, $B^{-1} \geqq 0$. It follows that $(AB)^{-1}$ exists and $(AB)^{-1} = B^{-1}A^{-1} \geqq 0$ whence $AB \in \mathbf{K}$, taking into account the inclusion $AB \in \mathbf{Z}$.

(2,8) *Let* $A \in \mathbf{K}$, $B \in \mathbf{Z}$. *If* $AB \in \mathbf{K}$, *then* $B \in \mathbf{K}$. *If* $AB \in \mathbf{K}_0$ *and is irreducible, then* $B \in \mathbf{K}_0$.

Proof. By 2° of (2;4,3) there exists a vector $x > 0$ such that $ABx = y > 0$. Since $A \in \mathbf{K}$, it follows that $A^{-1} \geqq 0$ with all diagonal elements positive. Hence $Bx = A^{-1}y > 0$ and it follows from 2° of (2;4,3) that $B \in \mathbf{K}$.

Let now $AB \in \mathbf{K}_0$ and let AB be irreducible. It suffices to discuss only the case that AB is singular since otherwise $AB \in \mathbf{K}$ and $B \in \mathbf{K}$. In this case, there exists, by (2;5,6), a vector $x > 0$ such that $ABx = 0$. Thus we have $Bx \geqq 0$ and $B \in \mathbf{K}_0$ by (2;5,4). The proof is complete.

(2,9) *Let* $A \in \mathbf{K}_0$ *be singular and suppose* z *is a vector for which* $Az \geqq 0$. *If* A *is irreducible then* $Az = 0$.

Proof. According to (2;5,6) there exists a vector $y > 0$ such that $y'A = 0$. If $u = Az$, we have $y > 0$, $u \geqq 0$ and $y'u = y'Az = 0$ so that u must be the zero vector.

We shall need further a sharpening of theorem (2;6,7). For the sake of completeness we intend to give the entire proof although a part of the present result is already contained in (2;6,7). We introduce first a notation.

(2,10) *Let* A *and* B *be two matrices of type* (n, n) *and let* $0 < \alpha < 1$ *be given. We shall denote by* $g(A, B)$ *the matrix* G *where*

$$g_{ii} = |a_{ii}|^{\alpha} |b_{ii}|^{1-\alpha}, \quad g_{ik} = - |a_{ik}|^{\alpha} |b_{ik}|^{1-\alpha} \quad \text{for} \quad i \neq k .$$

(2,11) *Let* $0 < \alpha < 1$ *be given. Then the following implications hold*:

1° *If* $A \in \mathbf{K}$, $B \in \mathbf{K}$ *then* $g(A, B) \in \mathbf{K}$.

2° *If* $A \in \mathbf{K}_0$, $B \in \mathbf{K}_0$ *then* $g(A, B) \in \mathbf{K}_0$.

3° *Let* A *and* B *belong to* $\mathbf{K}_0$ *and let* $g(A, B)$ *be singular. Suppose further that* $g(A, B)$ *is irreducible. Then*

31° *both* A *and* B *are singular and there exist vectors* $x_0 > 0$ *and* $y_0 > 0$ *with* $Ax_0 = 0$ *and* $By_0 = 0$;

32° *if* $x > 0$, $y > 0$ *and* $Ax = 0$, $By = 0$ *then the vector* z *with coordinates* $z_i = x_i^{\alpha} y_i^{1-\alpha}$ *satisfies* $g(A, B)\, z = 0$;

33° *there exist positive diagonal matrices* P *and* Q *such that* $PA = BQ$;

34° *if* $x > 0$, $y > 0$ *satisfy* $Ax = 0$, $By = 0$ *and if* $X = \operatorname{diag}(x_1, \ldots, x_n)$, $Y = \operatorname{diag}(y_1, \ldots, y_n)$ *then there exists a positive diagonal matrix* D *such that* $AX = DBY$.

4° *Conversely, let* A *and* B *be matrices of type* (n, n) *and let* $A \in \mathbf{Z}$. *Let* X, Y, D *be positive diagonal matrices. Let* e *be the vector with* $e_i = 1$ *for every* i *and suppose that* $AXe = 0$. *Let* B *satisfy the relation* $AX = DBY$. *Then both* A *and* B *belong to* $\mathbf{K}_0$, $BYe = 0$ *and* $g(A, B)$ *is singular.*

Proof. We shall use the Hölder inequality in the following form: if a_i and b_i are nonnegative numbers then

$$\sum a_i^{\alpha} b_i^{1-\alpha} \leqq \left(\sum a_i\right)^{\alpha} \left(\sum b_i\right)^{1-\alpha}$$

and equality holds if and only if the vectors a and b are linearly dependent. Consider first the case $A, B \in \mathbf{K}$. According to 2° of (2;4,3), there exist positive vectors x and y such that $Ax > 0$ and $By > 0$. We are going to show that $g(A, B)\, z > 0$ where z is the vector with coordinates $z_i = x_i^{\alpha} y_i^{1-\alpha}$. Indeed, we have

$$\sum_{k \neq i} |a_{ik}|^{\alpha} |b_{ik}|^{1-\alpha} z_k = \sum_{k \neq i} \left(|a_{ik}|\, x_k\right)^{\alpha} \left(|b_{ik}|\, y_k\right)^{1-\alpha} \leqq$$
$$\left(\sum_{k \neq i} |a_{ik}|\, x_k\right)^{\alpha} \left(\sum_{k \neq i} |b_{ik}|\, y_k\right)^{1-\alpha} < \left(a_{ii} x_i\right)^{\alpha} \left(b_{ii} y_i\right)^{1-\alpha} = a_{ii}^{\alpha} b_{ii}^{1-\alpha} z_i\,.$$

This completes the proof of 1°. Suppose now that A and B belong to $\mathbf{K}_0$. We are going to show that $g(A, B) + \varepsilon E$ belongs to $\mathbf{K}$ for each positive ε. Clearly there exist positive numbers s_i and t_i such that

$$g_{ii} + \varepsilon = \left(a_{ii} + s_i\right)^{\alpha} \left(b_{ii} + t_i\right)^{1-\alpha}.$$

If S and T are diagonal matrices with s_i and t_i as diagonal elements, we have $A + S \in \mathbf{K}$ and $B + T \in \mathbf{K}$ by (2;5,11) and 3° of (2;5,1). Hence $g(A, B) + \varepsilon E = g(A + S, B + T) \in \mathbf{K}$ by the first assertion of the present theorem. It follows from (2;5,1) that $g(A, B) \in \mathbf{K}_0$.

To prove 3°, assume $A, B \in \mathbf{K}_0$ and suppose that $g(A, B)$ is singular and irreducible. According to 2°, we have $g(A, B) \in \mathbf{K}_0$. Since $g(A, B)$ is irreducible, both A and B are irreducible as well. Since both $A, B \in \mathbf{K}_0$ it follows from (2;5,8) that there exist vectors $x_0 > 0$ and $y_0 > 0$ for which $Ax_0 \geqq 0$ and $By_0 \geqq 0$. If z is the vector with coordinates $z_i = x_{0i}^{\alpha} y_{0i}^{1-\alpha}$, we obtain in the same manner as above $z > 0$ and $g(A, B)\, z \geqq 0$. Now it follows from (2,9) that $g(A, B)\, z = 0$; hence equality is attained in the inequalities

$$\sum_{k \neq i} |a_{ik}|^{\alpha} |b_{ik}|^{1-\alpha} z_k = \sum_{k \neq i} \left(|a_{ik}|\, x_{0k}\right)^{\alpha} \left(|b_{ik}|\, y_{0k}\right)^{1-\alpha} \leqq$$
$$\leqq \left(\sum_{k \neq i} |a_{ik}|\, x_{0k}\right)^{\alpha} \left(\sum_{k \neq i} |b_{ik}|\, y_{0k}\right)^{1-\alpha} \leqq \left(a_{ii} x_{0i}\right)^{\alpha} \left(b_{ii} y_{0i}\right)^{1-\alpha} = a_{ii}^{\alpha} b_{ii}^{1-\alpha} z_i$$

so that $Ax_0 = 0$ and $By_0 = 0$. This proves 31°.

To prove 32°, 34° and 33°, let $x > 0$, $y > 0$ be vectors for which $Ax = 0$, $By = 0$. Then, an analogous chain of inequalities as for x_0, y_0 is satisfied for x, y and for the vector z, $z_i = x_i^\alpha y_i^{1-\alpha}$. By (2,9), we have $g(A, B)\, z = 0$ which proves 32°. In these inequalities equality is attained. Hence, for each i, the vectors

$$u^{(i)} = \left(|a_{i1}|\, x_1, \ldots, |a_{i,i-1}|\, x_{i-1}, |a_{i,i+1}|\, x_{i+1}, \ldots, |a_{in}|\, x_n\right),$$
$$v^{(i)} = \left(|b_{i1}|\, y_1, \ldots, |b_{i,i-1}|\, y_{i-1}, |b_{i,i+1}|\, y_{i+1}, \ldots, |b_{in}|\, y_n\right)$$

are linearly dependent. Since $x > 0$, $y > 0$ and both A and B are irreducible, none of these is the zero vector so that there exists a $d_i > 0$ with $u^{(i)} = d_i v^{(i)}$. Since $a_{ii}x_i = \sum_{k \neq i} |a_{ik}|\, x_k = d_i \sum_{k \neq i} |b_{ik}|\, y_k = d_i b_{ii} y_i$ as well, we have proved the equation $AX = DBY$ where $D = \operatorname{diag}(d_1, \ldots, d_n)$. This proves 34°. Since there exist vectors x and y according to 31°, 33° is satisfied for $P = D^{-1}$, $Q = YX^{-1}$.

To prove 4°, let us write $x = Xe$, $y = Ye$ so that $x > 0$ and $y > 0$. We have $A \in \mathbf{Z}$, $x > 0$ and $Ax = AXe = 0$. It follows from (2;5,4) that $A \in \mathbf{K}_0$. Since $B = D^{-1}AXY^{-1}$, we have $B \in \mathbf{Z}$ and

$$By = BYe = D^{-1}AXe = 0\,.$$

Since $y > 0$, it follows from (2;5,4) that $B \in \mathbf{K}_0$. To see that $g(A, B)$ is singular, it suffices to take the vector z with coordinates $z_i = x_i^\alpha y_i^{1-\alpha}$ and show that $g(A, B)\, z = 0$. This follows from a direct computation.

The last theorem concerns matrices with all principal minors positive or nonnegative.

(2,12) *Let A be a real matrix such that $A + A^*$ is positive definite. Then, all principal minors of A are positive. If $A + A^*$ is nonnegative definite then all principal minors of A are nonnegative.*

Proof. The first part follows from (2;3,3) if we put $D_x = E$ for each x. To prove the second part, if suffices to consider the set of matrices $A + \varepsilon E$ for $\varepsilon > 0$ and apply the preceding result.

3. Some applications. As an illustration of the preceding results we shall prove here a theorem which generalizes some earlier results of R. S. Varga. In its formulation we shall need some notions concerning relations and their decompositions.

A relation on a set M is an arbitrary subset of $M \times M$. If R is a relation on M we shall write xRy for $(x, y) \in R$. A cycle in the relation R is a sequence $g_1, \ldots, g_m \in M$ such that

$$g_1 R g_2 R g_3 \ldots g_{m-1} R g_m R g_1\,.$$

A relation is said to be symmetric if aRb implies bRa. If R is a symmetric relation on M, we shall denote by R^e the relation defined as follows:

$aR^e c$ if and only if one of the following conditions is satisfied:

1° $a = c$,

2° aRc,

3° there exist elements $b_1, \ldots, b_k \in M$ such that $aRb_1Rb_2 \ldots b_kRc$.

Clearly R^e is the minimal equivalence containing R. We shall say that R is connected if $xR^e y$ for each x and y in M. (This is clearly in conformity with the terminology of the theory of graphs.)

Let us introduce now the following definition:

(3,1) Let R be a symmetric relation on M. We shall say that the three subsets S, P, P^* of R form a conservative decomposition of R if the following conditions are satisfied:

1° the sets S, P, P^* are pairwise disjoint;

2° iPk if and only if kP^*i;

3° for each cycle $g_1, \ldots, g_m$ in R

$$p(g_1, g_2) + p(g_2, g_3) + \ldots + p(g_{m-1}, g_m) + p(g_m, g_1) = 0$$

where

$$\begin{aligned} p(i, k) &= 0 \quad \text{for} \quad iSk\,, \\ p(i, k) &= -1 \quad \text{for} \quad iPk\,, \\ p(i, k) &= 1 \quad \text{for} \quad iP^*k\,. \end{aligned} \tag{1}$$

(3,2) *Let $S \cup P \cup P^*$ be a decomposition of a symmetric relation R satisfying 1° and 2° of (3,1). This decomposition is conservative if and only if there exists an integer-valued function V on M such that $i_0Ri_1R \ldots Ri_s$ implies*

$$V(i_s) - V(i_0) = \sum_{k=1}^{s} p(i_{k-1}, i_k)\,, \tag{2}$$

$p(i, k)$ being defined in (1).

Moreover, this function V is unique up to an additive constant if R is connected.

Proof. It is immediately seen that the condition (2) implies 3° of (3,1). Now, let the decomposition $S \cup P \cup P^*$ be conservative and let $M_1, \ldots, M_m$ be classes of equivalent elements in the equivalence R^e. Choose arbitrary elements $g_i \in M_i$, $i = 1, \ldots, m$ and put $V(g_i) = 0$. Let $h \in M$. If $h \in M_k$, we have one of the following three possibilities: either $g_k = h$ or g_kRh or there exists a sequence $a_1, \ldots, a_t$ such that

$$g_kRa_1Ra_2R \ldots Ra_tRh\,.$$

Let us form the sum

$$V(h) = p(g_k, a_1) + p(a_1, a_2) + \ldots + p(a_t, h).$$

To include all the three possibilities in the definition of $g_k R^e h$, let us agree that we take this sum to be empty if $g_k = h$ or has just one term if $g_k R h$.

Let us show that $V(h)$ is independent on the sequence from g_k to h. Indeed, let $g_k R b_1 R \ldots R b_n R h$ (in the same generalized sense) as well. Then,

$$g_k R a_1 R a_2 \ldots R a_t R h R b_n R \ldots R b_1 R g_k$$

is a cycle in R and from 3° it follows that

$$V(h) + p(h, b_n) + \ldots + p(b_1, g_k) = 0.$$

The independence follows immediately from the skew symmetry of p. We have thus obtained an integer-valued function on M. To prove the formula (2), let $a_0 R a_1 R \ldots$ $\ldots R a_s$ and let all these elements a_i belong to M_k. Hence there exist sequences $b_1, \ldots, b_v$ and $c_1, \ldots, c_w$ such that $g_k R b_1 R \ldots R b_v R a_0$, $a_s R c_1 R \ldots R c_w R g_k$, which complete the given sequence to a cycle. It follows in a similar manner as above that

$$V(a_0) + p(a_0, a_1) + \ldots + p(a_{s-1}, a_s) - V(a_s) = 0.$$

The formula is thus verified.

Let now R be connected (thus $m = 1$). If W is another function on M satisfying condition (2) then this formula yields

$$V(a) - V(b) = W(a) - W(b)$$

for all $a, b \in M$. It follows that

$$V(a) = W(a) + C$$

where C is independent on $a \in M$. The proof is complete.

In the sequel, we shall apply these notions to the case that the set M is the set of all natural numbers $\leq n$ and that $R = R(A)$ is the relation on M corresponding to a square n-rowed matrix A, i.e. $(i, k) \in R(A)$ if and only if $a_{ik} \neq 0$.

Let now A be a given matrix. Choose a nonsingular matrix B and consider the iteration procedure

(3) $$B x_{n+1} = (B - A) x_n + b;$$

if the sequence x_n converges, its limit x will be a solution of $Ax = b$. The preceding Gauss-Seidel procedure is clearly equivalent to the ordinary Ritz procedure

(4) $$x_{n+1} = B^{-1}(B - A) x_n + B^{-1} b.$$

It is therefore convenient to introduce the following abbreviation: given A, we shall denote by $\lambda(B)$ the spectral radius of $B^{-1}(B - A)$. The number $\lambda(B)$ may be considered as a measure of the convergence-rate of the procedure (3). The question of estimating $\lambda(B)$ as a function of B is of considerable practical importance.

Suppose now that $A \in \mathbf{K}$ and that we choose a matrix $B \in \mathbf{Z}$ and $B \geqq A$. According to (2,5) the matrix B belongs to $\mathbf{K}$ as well so that, in particular, B will be nonsingular. Further, $\lambda(B) = p(E - B^{-1}A)$ and we see from 6° of (2,5) that $\lambda(B) < 1$ so that the procedure (3) is convergent.

The following theorem on the monotonic dependence was proved for a symmetric matrix A in [1], for the general case in [3]:

(3,3) *Let $A \in \mathbf{K}$ and let B_1, B_2 be two matrices from $\mathbf{Z}$ such that $A \leqq B_1 \leqq B_2$. Then $\lambda(B_1) \leqq \lambda(B_2)$.*

The proof follows immediately from (2,6).

(3,4) *Let $A \in \mathbf{K}$ be symmetric. Suppose that $B \geqq A$. Put $D = B + B^* - A$ and suppose that $D \in \mathbf{Z}$. Then $B \in \mathbf{K}$, $D \in \mathbf{K}$ and D is symmetric.*

Proof. If $i \neq k$, we have $b_{ik} + b_{ki} \leqq a_{ik}$ since $D \in \mathbf{Z}$. Since $B \geqq A$, we have $-b_{ik} \leqq -a_{ik}$ which, together with the preceding inequality, yields $b_{ki} \leqq 0$. We have thus $B \in \mathbf{Z}$ so that $B \in \mathbf{K}$ by (2;4,6). Since $B \geqq A$, we have $D \geqq A$ as well and $D \in \mathbf{Z}$ by assumption. It follows that $D \in \mathbf{K}$.

In the sequel, the matrix B will be taken in the form $B = D - C^*$ where D is a symmetric matrix of class $\mathbf{K}$ and $C \geqq 0$ is such that $A = D - C - C^*$. In the following theorem estimates of $\lambda(B)$ will be given in terms of $\lambda(D)$ using the methods of section 2:

(3,5) **Theorem.** *Let A be a symmetric positive definite matrix and $A \in \mathbf{Z}$. Let $A = D - C - C^*$ where $D \in \mathbf{K}$ and $C \geqq 0$. Then, $B = D - C^*$ belongs to $\mathbf{K}$ and*

$$(\lambda(D))^2 \leqq \lambda(B) \leqq \frac{\lambda(D)}{2 - \lambda(D)} .$$

Suppose that A is irreducible. Then $\lambda(B) = (\lambda(D))^2$ if and only if $R(D) \cup R(C) \cup \cup R(C^*)$ is a conservative decomposition of $R(A)$.

Proof. Clearly $B \in \mathbf{Z}$ and $B = A + C \geqq A$. Since $A \in \mathbf{K}$ we have $B \in \mathbf{K}$ as well according to (2;4,6). Now let $\sigma > \lambda(B)$; since $\lambda(B) = p(B^{-1}(B - A)) = p(B^{-1}C)$, the matrix $\sigma - B^{-1}C$ belongs to $\mathbf{K}$ by (2,1). Further, $\sigma B - C \in \mathbf{Z}$ and $\sigma B - C = = B(\sigma - B^{-1}C)$ where both B and $\sigma - B^{-1}C$ belong to $\mathbf{K}$. It follows from (2,7) that $\sigma B - C \in \mathbf{K}$; clearly $\sigma B^* - C^* \in \mathbf{K}$ as well. Now take $\alpha = \frac{1}{2}$ and apply theorem (2,11) to the matrices $\sigma B - C$ and $\sigma B^* - C^*$. It follows that $g(\sigma B - C, \sigma B^* - C^*) \in$

$\in \boldsymbol{K}$. Denote by W the matrix $\sigma D - \sigma^{\frac{1}{2}} C - \sigma^{\frac{1}{2}} C^*$ so that $W \in \boldsymbol{Z}$. To show that $W \in \boldsymbol{K}$, it suffices, by (2;4,6), to show that $W \geqq g(\sigma B - C, \sigma B^* - C^*)$. Indeed,

$$w_{ii} = \sigma d_{ii} - 2\sigma^{\frac{1}{2}} c_{ii} \geqq \sigma(d_{ii} - c_{ii}) - c_{ii}\,; \tag{5}$$

for $i \neq k$

$$w_{ik} = \sigma d_{ik} - \sigma^{\frac{1}{2}}(c_{ik} + c_{ki}) \leqq 0$$

and

$$(\sigma d_{ik} - \sigma^{\frac{1}{2}}(c_{ik} + c_{ik}))^2 \leqq (\sigma(d_{ik} - c_{ik}) - c_{ki})(\sigma(d_{ik} - c_{ki}) - c_{ik})$$

since

$$0 \leqq -\sigma(1 - \sigma^{\frac{1}{2}})^2 d_{ik}(c_{ik} + c_{ki}) + (1 - \sigma)^2 c_{ik} c_{ki}\,. \tag{6}$$

We have thus shown that $W \in \boldsymbol{K}$. It follows that $\sigma^{\frac{1}{2}} D - C - C^* \in \boldsymbol{K}$ as well. Denote by F the matrix $\sigma^{\frac{1}{2}} - D^{-1}(C + C^*)$ so that $F \in \boldsymbol{Z}$. Since $DF = \sigma^{\frac{1}{2}} D - C - C^* \in \boldsymbol{K}$ and $D \in \boldsymbol{K}$, it follows from (2,8) that $F \in \boldsymbol{K}$ whence

$$\sigma^{\frac{1}{2}} > p(D^{-1}(C + C^*)) = p(D^{-1}(D - A)) = \lambda(D).$$

To prove the estimate of $\lambda(B)$ from above, we shall denote by M the matrix

$$\frac{p_2}{2 - p_2} E - (D - C^*)^{-1} C$$

where $p_2 = \lambda(D)$.

The matrix $(D - C^*)^{-1} C$ is nonnegative and $M \in \boldsymbol{Z}$. We know already that $B \in \boldsymbol{K}$. Les us consider the matrix

$$BM = \frac{p_2}{2 - p_2}(D - C^*) - C\,.$$

The matrix $p_2 D - (C + C^*)$ belongs to $\boldsymbol{K}_0$ by (2,1) and is, accordingly, nonnegative definite.

Since

$$BM + (BM)^* = \frac{2p_2}{2 - p_2}\left[D - \frac{1}{p_2}(C + C^*)\right]$$

is nonnegative definite as well and $BM \in \boldsymbol{Z}$, it follows from lemma (2,12) that $BM \in \boldsymbol{K}_0$. An application of (2,8) shows that $M \in \boldsymbol{K}_0$. It follows that $p_2/(2 - p_2) \geqq p[(D - C^*)^{-1} C] = \lambda(B)$.

Suppose now that $\lambda(B) = \lambda(D)^2$. We shall distinguish two cases.

If $\lambda(D) = 0$, we shall show that $C = 0$. Indeed, we have $p(D^{-1}(C + C^*)) = \lambda(D) = 0$ and D^{-1}, being inverse to a matrix of class $\boldsymbol{K}$, has positive diagonal elements. Since $C + C^*$ is nonnegative and symmetric, it follows from lemma (2,2) that $C + C^* = 0$. Since $C \geqq 0$, we have $C = 0$ as well. It is easy to see that, conversely, $C = 0$ implies $\lambda(B) = \lambda(D) = 0$. The assertion of the theorem is easily seen to be valid.

Suppose now that $\lambda(D) \neq 0$ and that A is irreducible. Write τ for $\lambda(B)$ and observe that $\tau B - C$ is singular. Further $g(\tau B - C, \tau B^* - C^*) \leqq \tau D - \tau^{\frac{1}{2}} C - \tau^{\frac{1}{2}} C^* = \tau^{\frac{1}{2}}(\tau^{\frac{1}{2}} D - C - C^*) = \tau^{\frac{1}{2}} (\lambda(D) D - C - C^*) = \tau^{\frac{1}{2}} D(\lambda(D) - D^{-1}(D - A))$ and this last matrix is singular. Since $\tau \neq 0$ and $D - C - C^* = A$ is irreducible, the matrix $\tau D - \tau^{\frac{1}{2}} C - \tau^{\frac{1}{2}} C^*$ is irreducible as well. By lemma (2,4) we have

$$g(\tau B - C, \tau B^* - C^*) = \tau D - \tau^{\frac{1}{2}} C - \tau^{\frac{1}{2}} C^* .$$

It follows that equality is attained both in (5) and (6) for $\sigma = \tau$. Equation (5) yields $c_{ii} = 0$ for all i. From (6) we obtain that for each $i, k, i \neq k$, at most one of the numbers d_{ik}, c_{ik}, c_{ki} is different from zero.

We know already that both matrices $\tau B - C$ and $\tau B^* - C^*$ are singular. Clearly they are irreducible as well so that there exist (essentially unique) vectors $x > 0$ and $y > 0$ for which $(\tau B - C) x = 0$ and $(\tau B^* - C^*) y = 0$. Further we have just seen that $g(\tau B - C, \tau B^* - C^*) = \tau D - \tau^{\frac{1}{2}} C - \tau^{\frac{1}{2}} C^*$ is singular and irreducible.

It follows from (2,11) that there exists a positive diagonal matrix H such that

$$(7) \qquad (\tau B - C) X = H(\tau B^* - C^*) Y$$

where $X = \operatorname{diag}(x_1, \ldots, x_n)$ and $Y = \operatorname{diag}(y_1, \ldots, y_n)$. On comparing the diagonal elements and taking into account the fact that $c_{ii} = 0$ we obtain for the diagonal elements $h(i)$ of H the equation

$$h(i) = \frac{x_i}{y_i} .$$

Now let $i \neq k$. If $c_{ik} \neq 0$, then $d_{ik} = 0$ and $c_{ki} = 0$ and it follows from (7) that $-c_{ik} x_k = -h(i) \tau c_{ik} y_k$, or, in other words,

$$(8) \qquad \tau h(i) = h(k) .$$

If $d_{ik} \neq 0$, we have $c_{ik} = c_{ki} = 0$ and it follows in the same way that

$$(9) \qquad h(i) = h(k) .$$

For $i \neq k$, let us define a number $p(i, k)$ in the following manner:

$$\begin{aligned} p(i, k) &= -1 \quad \text{if} \quad c_{ik} \neq 0 , \\ p(i, k) &= 1 \quad \text{if} \quad c_{ki} \neq 0 , \\ p(i, k) &= 0 \quad \text{otherwise.} \end{aligned}$$

This is possible since $c_{ik} c_{ki} = 0$ for all i, k.

Since $A = D - C - C^*$, we see that $a_{ik} \neq 0$ if and only if exactly one of the elements d_{ik}, c_{ik}, c_{ki} is different from zero. This enables us to replace (8) and (9) by a single formula

$$\frac{h(i)}{h(k)} = \tau^{p(k,i)}$$

whenever $a_{ik} \neq 0$.

Suppose now that $i_1, i_2, \ldots, i_m$ is a cycle in $R(A)$; in other words, all the elements $a_{i_1 i_2}, a_{i_2 i_3}, \ldots, a_{i_{m-1} i_m}, a_{i_m i_1}$ are different from zero. Clearly

$$\frac{h(i_1)}{h(i_2)} \frac{h(i_2)}{h(i_3)} \cdots \frac{h(i_{m-1})}{h(i_m)} \frac{h(i_m)}{h(i_1)} = 1$$

whence, τ being different from 1, $p(i_1, i_2) + p(i_2, i_3) + \ldots + p(i_{m-1}, i_m) + p(i_m, i_1) = 0$.

Thus, $R(D) \cup R(C) \cup R(C^*)$ is a conservative decomposition of $R(A)$.

Conversely, it is easily seen that if $R(D) \cup R(C) \cup R(C^*)$ is a conservative decomposition of $R(A)$ then $(\lambda(D))^2 = \lambda(B)$.

Bibliography

[1] *M. Fiedler, Vl. Pták:* Über die Konvergenz des verallgemeinerten Seidelschen Verfahrens zur Lösung von Systemen linearer Gleichungen. Math. Nachr. *15* (1956), 31—38.

[2] *M. Fiedler, Vl. Pták:* On matrices with non-positive off-diagonal elements and positive principal minors. Czech. Math. J. *12* (*87*) (1962), 382—400.

[3] *A. S. Householder:* The Theory of Matrices in Numerical Analysis. Blaisdell Publ. Co., N. York, Toronto, London 1964.

[4] *R. S. Varga:* A comparison of the successive overrelaxation method and semi-iterative methods using Chebyshev polynomials. J. Soc. Indust. Appl. Math. *5* (1957), 39—46.

[5] *R. S. Varga:* Orderings of the successive overrelaxation scheme. Pac. J. Math. *9* (1959), 925—939.

[6] *D. Young:* Iterative methods for solving partial difference equations of elliptic type. Trans. Amer. Math. Soc. *76* (1954), 92—111.

Author's addresses: Praha 1, Žitná 25, ČSSR (Matematický ústav ČSAV).

Резюме

НЕКОТОРЫЕ РЕЗУЛЬТАТЫ О МАТРИЦАХ КЛАССА *K* И ИХ ПРИМЕНЕНИЯ К СКОРОСТИ СХОДИМОСТИ ИТЕРАТИВНЫХ МЕТОДОВ

МИРОСЛАВ ФИДЛЕР, ВЛАСТИМИЛ ПТАК, (Miroslav Fiedler, Vlastimil Pták), Прага

Новые результаты и уточнения известных результатов о матрицах классов **K** и $\mathbf{K}_0$ применяются к изучению скорости сходимости обобщенных итерационных методов Гаусса-Зейделя. Основная теорема обобщает результати Р. С. Варги,

следовательно которому консервативные методы имеют наибольшую скорость сходимости среди циклических итеративных методов для матриц типа Янга A.

Если A данная матрица и B некоторая невырожденная матрица, потом скорость сходимости итеративного метода

$$Bx_{n+1} = (B - A)\,x_n + b$$

измеряется спектральным радиусом матрицы $B^{-1}(B - A)$, обозначаемым $\lambda(B)$.

В главной теореме 5,5 доказывается следующая оценка для $\lambda(B)$: *Если A симметрическая, положительно определенная матрица такая, что $a_{ik} \leqq 0$ для $i \neq k$, и если $A = D - C - C^*$ (C^* – транспонированная матрица к C), где $C \geqq 0$ и D положительно определенная матрица такая, что $d_{ik} \leqq 0$ для $i \neq k$, потом*

$$[\lambda(D)]^2 \leqq \lambda(B) \leqq \lambda(D)/(2 - \lambda(D))$$

Дается комбинаторная характеризация случая равенства в левом неравенстве.