IV. Very important also is the study of the technical possibilities and technical facilities for effectively performing operations over data arrays, and also of efficient forms of organization of the transmission and processing of these arrays.

Thus, it is necessary to describe the technical possibilities of the computers in such a way that not only can the requirements towards them be specified, but we can also take into account the capabilities of computers based on technology, and what can be easily and fairly simply realized. In analyzing the algorithms, and in estimating and finding possible methods of solution, it is necessary to organize the algorithms in such a way that these real technical possibilities are best used.

V. Ascertaining the principles of construction of a high-level language that makes it possible to conveniently and effectively describe the above-mentioned mass methods; this language must be also oriented towards the corresponding technical facilities and take into account how they can be used.

In other words, it is important to construct a language that would represent the peculiar structural features of large-scale problems and would adequately describe them, while at the same time effectively helping to realize the corresponding capabilities of computer technology.

For example, microprogramming which was available in many computers has often remained practically unutilized. It was not clear where and how it could be used. The programming system and algorithmic languages were not suitable for this purpose. It is necessary to ascertain the principles of compilation and computer interpretation of the problems described in such a language, with efficient utilization of combined information units and operations realized by available technical facilities.

These are the problems to be discussed. We must find out what has been done already in this field, and what are the future prospects. This will make it possible to elaborate more efficient methods of development of computer technology and of using it.

THE SOLUTION OF THE ADDITIVE EIGENVALUE PROBLEM OF MATRICES

V. N. Kublanovskaya

UDC 518.512

It is required to determine a diagonal matrix X such that the symmetric matrix $A+X$ has preassigned eigenvalues (the additive problem). An iteration process is described that is based on the use of two-dimensional rotations, and convergence problems are analyzed.

1°. Statement of the Problem. The solving of the inverse eigenvalue problem of a matrix is considered in the following formulation.

For a given symmetric matrix A with constant elements it is required to find a real diagonal matrix X such that the matrix $A+X$ has preassigned eigenvalues $\mu_1, \ldots, \mu_n$. (Here $\mu_1, \ldots, \mu_n$ are real numbers.) This formulation is usually called the additive problem. This problem can be solved, but not for just any value of μ_i. In the literature we can find sufficient conditions of existence of a matrix X that solve the additive problem (see, e.g., [1-4]). We shall assume that for a certain numeration of the values μ_i we have the conditions

$$|\mu_i-\mu_j| \geqslant 2(A_i+A_j) \tag{1}$$

for $i,j=1,\ldots,n$. Here $A_k=\sum_{\substack{s=1\\ s\neq k}}^{n}|a_{sk}|$, n is the order of the matrix A.

Translated from Zapiski Nauchnykh Seminarov Leningradskogo Otdeleniya Matematicheskogo Instituta im. V. A. Steklova AN SSSR, Vol. 48, pp. 12-17, 1974.

0096-4104/78/1005-0647$07.50

Without loss of generality, we shall assume that the diagonal elements of the matrix A are equal to zero. Otherwise we can take the sought matrix in the form $X + \operatorname{diag} A$.

In this article we describe an algorithm of solution of the additive problem and we examine convergence problems. The algorithm is based on the use of two-dimensional rotations and it is a modification of the Jacobi process [5] adapted to the solution of the additive problem.

2°. Description of Algorithm. Let $\mathcal{D} = [\mu_1, \dots, \mu_n]$ be a diagonal matrix. By virtue of the conditions (1) the sought matrix X exists [3]. Hence there evidently exists an orthogonal matrix Q such that $Q(A+X)Q^T = \mathcal{D}$, and

$$X = Q^T \mathcal{D} Q - A . \tag{2}$$

A matrix Q that satisfies (2) will be constructed with the aid of a chain of two-dimensional rotations. For this purpose we shall construct a sequence of similar matrices $M_0, M_1, \dots$ according to the following prescription

$$M_0 = \mathcal{D}, \quad M_{k+1} = Q_{ij}^T M_k Q_{ij}, \quad k = 0, 1, \dots \tag{3}$$

Here $Q_{ij} = (q_{rs})$ is an elementary orthogonal matrix of two-dimensional rotations, i.e., a matrix that differs from the unit matrix by four elements $q_{ii} = q_{jj} = c$, $q_{ij} = -q_{ji} = s$, $c^2 + s^2 = 1$. The coefficients c and s are selected in such a way that the element standing in the position (i,j) of the matrix $M_{k+1} = (m_{rs}^{(k+1)})$ coincides with the corresponding element of the original matrix $A = (a_{rs})$. Since

$$m_{ij}^{(k+1)} = (c^2 - s^2) m_{ij}^{(k)} + cs(m_{jj}^{(k)} - m_{ii}^{(k)}) \tag{4}$$

we obtain for c and s the system

$$\left.\begin{aligned} (c^2 - s^2) m_{ij}^{(k)} + cs(m_{jj}^{(k)} - m_{ii}^{(k)}) &= a_{ij} \\ c^2 + s^2 &= 1 \end{aligned}\right\} . \tag{5}$$

Hence we obtain

$$c = \sqrt{\frac{d_k + 4a_{ij} m_{ij}^{(k)} + b_k}{2d_k}}, \quad s = \operatorname{sign} s \cdot |s|, \tag{6}$$

where

$$|s| = \sqrt{\frac{d_k - 4a_{ij} m_{ij}^{(k)} - b_k}{2d_k}};$$

$$\operatorname{sign} s = \operatorname{sign}\left\{(m_{jj}^{(k)} - m_{ii}^{(k)})[a_{ij} - (c^2 - s^2) m_{ij}^{(k)}]\right\},$$

$$d_k = (m_{ii}^{(k)} - m_{jj}^{(k)})^2 + 4m_{ij}^{(k)2},$$

$$b_k = \sqrt{d_k^2 - 4d_k(a_{ij}^2 + m_{ij}^{(k)2}) + 16 a_{ij}^2 m_{ij}^{(k)2}} .$$

The index pair (i,j) can be selected in the same way as in Jacobi's method for the diagonalization of a symmetric matrix. We shall distinguish between three methods of selection of (i,j) :1) $|m_{ij}^{(k)} - a_{ij}| = \max_{r \neq s} |m_{rs}^{(k)} - a_{rs}|$, which is a counterpart of the classical method of Jacobi. In the following we shall call it the modified classical method. 2) $|m_{ij}^{(k)} - a_{ij}| >$

$\Pi_s \quad s = 1, 2, \ldots, t$. Here $\Pi_1 > \Pi_2 > \ldots > \Pi_t$ are assigned obstructions. This is a counterpart of Jacobi's method with obstructions. 3) A cyclic process, i.e., the pairs (i, j) are selected in cyclic order $(i,j) = (1,2), (1,3), \ldots, (1,n), (2,3), \ldots, (2,n), \ldots, (n-1,n)$. As a measure of relaxation of the process we shall take the quantity

$$t^2(M_k - A) = \sum_{\substack{r,s=1 \\ r \neq s}}^{n} (m_{rs}^{(k)} - a_{rs})^2 .$$

The process of construction of the matrices M_k, and hence of the approximations $X_k = M_k - A$ of the diagonal matrix X will be continued until the quantity $t^2(M_k - A)$ is a small quantity with the prescribed accuracy or the effective accuracy. For a sufficiently large k it is possible to take the diagonal matrix $M_k - A$ as the sought matrix X.

3°. Convergence of the Algorithm. Let us prove the relaxation algorithm (3), i.e., let us ascertain the conditions under which

$$t^2(M_{k+1} - A) \leq t^2(M_k - A) \tag{7}$$

for $k = 0, 1, \ldots$. For simplicity we shall denote $M_{k+1} = \bar{M} = (\bar{m}_{rs})$, $M_k = M = (m_{rs})$. We have

$$t^2(\bar{M} - A) - t^2(M - A) = t^2(\bar{M}) + t^2(A) - 2\sum_{r \neq s} \bar{m}_{rs} a_{rs} - t^2(M) - t^2(A) + 2\sum_{r \neq s} m_{rs} a_{rs} .$$

Since

$$t^2(\bar{M}) - t^2(M) = 2(\bar{m}_{ij}^2 - m_{ij}^2) = 2(a_{ij}^2 - m_{ij}^2)$$

$\bar{m}_{rs} a_{rs} - m_{rs} a_{rs} = 0$ for any r and s that do not coincide with i or j, we obtain

$$t^2(\bar{M} - A) - t^2(M - A) = -2(a_{ij} - m_{ij})^2 + q , \tag{8}$$

where

$$q = 4 \sum_{\substack{p=1 \\ p \neq i, p \neq j}}^{n} [a_{pi}(m_{pi} - \bar{m}_{pi}) + a_{pj}(m_{pj} - \bar{m}_{pj})] .$$

For the modified classical process we have $(m_{ij} - a_{ij})^2 \geq \dfrac{t^2(M-A)}{n(n-1)}$, so that (8) can be written in the form

$$t^2(\bar{M} - A) \leq \left(1 - \frac{2}{n(n-1)}\right) t^2(M - A) + q . \tag{9}$$

Let us find an upper bound for q. From the transformation $\bar{M} = Q_{ij}^T M Q_{ij}$ we obtain

$$\bar{m}_{pi} = \bar{m}_{ip} = c m_{pi} + s m_{pj} , \quad \bar{m}_{pj} = \bar{m}_{jp} = -s m_{pi} + c m_{pj} ,$$

so that

$$q = 4 \sum_{\substack{p \neq i \\ p \neq j}} \left\{ a_{pi}[(1-c) m_{pi} - s m_{pj}] + a_{pj}[(1-c) m_{pj} + s m_{pi}] \right\}$$

Hence

$$|q| \leq 8 \max_{p \neq q} |m_{pq}| \left(\sum_{p \neq i} |a_{pi}| + \sum_{p \neq j} |a_{pj}| \right) = 8 \max_{p \neq q} |m_{pq}| (A_i + A_j) . \tag{10}$$

It is well known [6] that the range $\rho(M)$ of the matrix M satisfies the inequality

$$\rho(M) \geq 2 \max_{p \neq q} |m_{pq}|$$

On the other hand, the matrix $M \equiv M_k$ is similar to the matrix $\mathfrak{D}$ for any $k = 1, 2, \ldots$, so that

$$\rho \equiv \rho(M) \equiv \max_{i \neq j} |\mu_i - \mu_j| \geq 2 \max_{p \neq q} |m_{pq}| .$$

Moreover, by virtue of the fulfillment of condition (1) we have

$$2(A_i + A_j) \leq |\mu_i - \mu_j| ,$$

and hence

$$|q| \leq 2 \max_{i \neq j} |\mu_i - \mu_j|^2 = 2\rho^2 \tag{11}$$

Next let us find a lower bound for $t^2(M - A)$. Let us denote by $n^2(B) = \sum_{i,j=1}^{n} b_{ij}^2$ the sum of the squares of all the elements of the matrix B. Since $a_{pp} = 0$ for $p = 1, \ldots, n$, we hence obtain

$$t^2(M-A) = n^2(M-A) - \sum_{p=1}^{n} m_{pp}^2 = n^2(M) - \sum_{p=1}^{n} m_{pp}^2 +$$
$$+ n^2(A) - 2 \sum_{\substack{p,q=1 \\ p \neq q}}^{n} m_{pq} a_{pq} = t^2(M) + n^2(A) - 2 \sum_{\substack{p,q=1 \\ p \neq q}}^{n} m_{pq} a_{pq} \geq$$
$$\geq n^2(A) - 2 \max_{p \neq q} |m_{pq}| \cdot \sum_{\substack{p,q=1 \\ p \neq q}}^{n} |a_{pq}| \tag{12}$$

Next, since $2 \max_{p \neq q} |m_{pq}| \leq \rho \equiv \max_{i \neq j} |\mu_i - \mu_j|$, it follows from (12) that

$$t^2(M-A) \geq n^2(A) - \rho \sum_{\substack{p,q=1 \\ p \neq q}}^{n} |a_{pq}| . \tag{13}$$

We shall require that

$$2\rho^2 \leq \frac{2}{n(n-1)} \Big[n^2(A) - \rho \sum_{\substack{p,q=1 \\ p \neq q}}^{n} |a_{pq}| \Big] \tag{14}$$

Since $\rho > 0$, it follows from (14) that

$$0 < \rho \leq \beta , \tag{15}$$

where

$$\beta = \frac{-\sigma + \sqrt{\sigma^2 + 4n(n-1)\, n^2(A)}}{2n(n-1)} , \quad \sigma \equiv \sum_{\substack{p,q=1 \\ p \neq q}}^{n} |a_{pq}|$$

By virtue of (11) and (13) we hence conclude that when ρ varies in the interval $(0, \beta]$ we have the inequality

$$|q| \leq 2\rho^2 \leq \frac{2}{n(n-1)} \Big[n^2(A) - \rho\sigma \Big] \leq \frac{2}{n(n-1)} t^2(M-A),$$

so that by virtue of (9) we obtain for $0 < \rho \leq \beta$ the formula

$$t^2(\bar{M}-A) \leq \gamma\, t^2(M-A) \tag{16}$$

where $\gamma < 1$.

Thus, we have shown that under condition (15) the process (3) is a relaxation process:

$$t^2(M_{k+1} - A) \leq \gamma\, t^2(M_k - A) \leq \ldots \leq \gamma^{k+1} t^2(M_0 - A) ,$$

so that for $\kappa \to \infty$ we have $t^2(M_\kappa - A) \to 0$, and hence $m_{pq}^{(\kappa)} \to a_{pq}$ for $k \to \infty$ for any $p, q = 1, \ldots, n$, $p \neq q$.

Thus we have proved the convergence of the modified classical Jacobi process.

From the foregoing we obtain a theorem:

THEOREM. For the modified classical Jacobi process for the additive problem to be convergent, it suffices that

$$\alpha \leq \max_{i \neq j} |\mu_i - \mu_j| \leq \beta,$$

where

$$\alpha = 2 \max_{i \neq j} (A_i + A_j), \quad \beta = \frac{-\sigma + \sqrt{\sigma^2 + 4n(n-1)\, n^2(A)}}{2n(n-1)}.$$

n being the order of A, $A_\tau = \sum_{\kappa \neq \tau} |a_{\kappa\tau}|$.

LITERATURE CITED

1. K. P. Hadeler, "Ein inverses Eigenwertproblem," Linear Algebra and Its Applications, 1, 83-101 (1968).
2. Francoise Laborde, "Sur un probleme inverse d'un probleme de valeurs propres," C. R. Acad. Sci. Paris, 268, 153-156 (1969).
3. G. H. De Oliveira, "Note on an inverse characteristic value problem," Num. Math., 15, 345-347 (1970).
4. G. N. De Oliveira, "Note on the additive inverse eigenvalue problem," Rev. Fac. C. Lisboa, 13, 21-26 (1970).
5. D. K. Faddeev and V. N. Faddeeva, Computational Methods of Linear Algebra, W. H. Freeman, San Francisco (1963).
6. M. Marcus and H. Minc, Survey of Matrix Theory and Matrix Inequalities [Russian translation], Moscow (1972).

ORGANIZATION OF PSEUDOARRAYS IN REFAL

M. M. Lebedinskii

UDC 681.3.06:51

A pseudoarray in REFAL is described that makes it possible to speed up filing and the introduction of elements in an array. The corresponding programs are presented.

In computational problems it is often convenient to use an array structure of data. An array in the REFAL language is a sequence of terms; the index of an element (term) is its ordinal number. The principal operation in working with arrays, especially in REFAL, is to seek an element of an array according to its index. For REFAL the most natural organization of this process is to successively file through the array until the desired element is reached. For convenience we shall call this method simple filing. The average time of simple filing evidently has the same order of magnitude as the length of the array. Therefore it is desirable to have a better method of filing. In such a method the average filing time is approximately $\log_2 n$, where n is the array length. However, in view of the limitations of the parameters of a computer (e.g., the memory), the practical usefulness of such a reduction is limited, and it can be estimated only by experiment. The idea of such a reduction consists in preprocessing the array (i.e., prior to operating with its elements) for the purpose of adapting to REFAL the well-known method of seeking an element on the basis of a feature in an array that has been ordered according to this feature. Thus, we have a dichotomy, i.e.,

Translated from Zapiski Nauchnykh Seminarov Leningradskogo Otdeleniya Matematicheskogo Instituta im. V. A. Steklova AN SSSR, Vol. 48, pp. 18-22, 1974.