Travaux cités

[1] S. Łojasiewicz, *Sur la valeur et la limite d'une distribution dans un point*, Studia Math. 16 (1957), p. 1-36.

[2] — *Sur la fixation des variables dans une distribution*, ibidem 17 (1958), p. 1-64.

[3] J. Mikusiński, R. Sikorski, *The elementary theory of distributions*, *I*, Warszawa 1957.

[4] M. Nicolescu, *Analiza matematică*, *II*, București 1958.

[5] T. Popoviciu, *Différences divisées et dérivées*, Mathematica I, 24.2 (1959), p. 297-319.

[6] Z. Zieleźny, *Sur la définition de Łojasiewicz de la valeur d'une distribution dans un point*, Bull. Ac. Polon., Sci., Cl. III, 3 (1955), p. 519-520.

Reçu par la Rédaction le 30. 12. 1962

STUDIA MATHEMATICA, XXIV. (1964)

On the iterative procedures of best strategy for inverting a self-adjoint positive-definite bounded operator in Hilbert space

by

A. KIEŁBASIŃSKI (Warszawa)

1. Introduction. This paper is closely connected with the paper [1] of M. Altman and with the well-known iteration procedures of the best strategy for solving numerical problems of linear algebra [6]. In this saper a Chebyshev iterative method of arbitrary degree for inverting pelf-adjoint, positive-definite, bounded operators in Hilbert space is developed. For the discussion of results and comparison of methods several basic ideas and properties of some known methods are recapitulated.

Probably the most efficient methods for inverting matrices are not iterative methods but methods of the type of Gaussian elimination. For improving an approximate of the inverse A^{-1} obtained by such elimination procedure it seems sufficient to apply one or two steps of the well-known second degree hyperpower method. Therefore methods developed in this paper are probably not very essential to the numerical practice.

2. Hyperpower methods in Banach spaces. Let A be a linear, bounded, non-singular operator with the domain and the range in a complex Banach space $\mathfrak{B}$. Let us suppose that a linear bounded operator D_0 satisfies the condition

$$\|I-D_0A\| = \varrho < 1, \tag{1}$$

where I is the identity mapping of $\mathfrak{B}$. By introducing a linear operator $B_0 = I-D_0A$ we see that the equation

$$X = B_0X+D_0 \tag{2}$$

is satisfied by the operator A^{-1}. In virtue of (1) the operator A^{-1} is the unique solution of equation (2) and we have $A^{-1} = \lim_{n\to\infty} X_n$, where the sequence of operators $\{X_n\}$ is defined by the recurrence formula

$$X_{k+1} = B_0X_k+D_0, \quad k = 0, 1, \ldots \tag{3}$$

X_0 is an arbitrary linear bounded operator.

From (2) and (3) we obtain the relation

$$A^{-1}-X_k = B_0^k(A^{-1}-X_0), \tag{4}$$

and the error estimate

$$\|A^{-1}-X_k\| \leqslant \varrho^k \cdot \|A^{-1}-X_0\|. \tag{5}$$

We see that if $X_0 = A^{-1}$, then any term X_k is equal to A^{-1}. That seems to be an advantage of procedure (3).

Let us consider (following von Neumann's or Gavurin's idea [2], [3]) some linear combination of $(n+1)$ first terms of the sequence $\{X_k\}$,

$$Y_n = \sum_{k=0}^{n} a_{n,k} X_k, \tag{6}$$

with additional condition

$$\sum_{k=0}^{n} a_{n,k} = 1. \tag{7}$$

Condition (7) follows from demand: $X_0 = A^{-1}$ implies $Y_n = A^{-1}$. Taking into account (4), (6) and (7) we find the following relation:

$$\begin{aligned} A^{-1}-Y_n &= \sum_{k=0}^{n} a_{n,k}(A^{-1}-X_k) = \Big(\sum_{k=0}^{n} a_{n,k}B_0^k\Big)(A^{-1}-X_0) \\ &= W_n(B_0)(A^{-1}-X_0). \end{aligned} \tag{8}$$

Here $W_n(t) = \sum_{k=0}^{n} a_{n,k}t^k$ is a polynomial of degree $\leqslant n$ satisfying the additional condition (cf. (7))

$$W_n(1) = 1. \tag{9}$$

We observe that any procedure of type (6) subject to condition (7) is fully characterized by a sequence of polynomials $\{W_n(t)\}$, satisfying condition (9).

The error estimate is here

$$\|A^{-1}-Y_n\| \leqslant \|W_n(B_0)\| \cdot \|A^{-1}-X_0\|. \tag{10}$$

A sufficient and necessary condition for the convergence of the sequence $\{Y_n\}$ to the inverse A^{-1} (with arbitrary $Y_0 = X_0$) is $\lim_{n\to\infty} \|W_n(B_0)\| = 0$.

If λ is an eigenvalue of B_0, then $|\lambda| \leqslant \varrho = \|B_0\|$ and in general the existence of any eigenvalue satisfying this condition cannot be excluded. If λ is an eigenvalue of B_0, then $W_n(\lambda)$ is an eigenvalue of $W_n(B_0)$ and $|W_n(\lambda)| \leqslant \|W_n(B_0)\|$. Hence the polynomials of a complex variable t in the sequence $\{W_n(t)\}$ should satisfy the necessary condition

$$\lim_{n\to\infty} \sup_{|t|\leqslant\varrho} |W_n(t)| = 0.$$

The choice of $W_n(t) = t^n$ yields minimum of the functional $\Phi(W) = \sup_{|t|\leqslant\varrho} |W(t)|$ in the class of all polynomials of degree $\leqslant n$ of a complex variable subject to the additional condition (9) and minimum of the functional

$$\Psi(W) = \frac{1}{2\pi}\int_0^{2\pi} |W(\varrho\cdot e^{i\theta})|^2 d\theta$$

in the same class of polynomials (cf. [4]). We have also $\Phi(t^n) = \varrho^n \geqslant \|B_0^n\|$.

It follows that the choice of the best strategy (cf. [6]) in this case is the choice of the sequence $\{t^n\}$ or of a subsequence of this one. In fact, if we wish to obtain X_n we must not construct each term X_k, $k < n$.

Altman [1] showed how to construct iteration procedures corresponding to the subsequences $\{t^{p^i}\}$, $i = 0, 1, 2, \ldots$, for arbitrary integers $p \geqslant 2$. The simple identity $t^{p^{i+1}} = (t^{p^i})^p$ is basic for this construction.

Let us start with $X_0 = 0$ and construct the sequence $\{D_i\}$, $i = 0, 1, \ldots$, where $D_i = X_{p^i}$. From (4) we obtain $A^{-1}-D_i = B_0^{p^i}A^{-1}$. By introducing the residual operator $B_i = I-D_iA$, we obtain

$$B_i = I-D_iA = B_0^{p^i}. \tag{11}$$

Hence we obtain $B_{i+1} = B_0^{p^{i+1}} = B_i^p$, so that

$$D_{i+1} = (I-B_i^p)A^{-1} = \sum_{k=0}^{p-1} B_i^k D_i,$$

by the trivial identity $(1-b^p) = \sum_{k=0}^{p-1} b^k(1-b)$ and by (11). The last formula is basic for the computation algorithms. We write it therefore in the expanded form

$$D_{i+1} = (I+B_i+\ldots+B_i^{p-1})D_i. \tag{12}$$

Altman discusses in [1] the choice of the optimal method from the class of "hyperpower methods" (12) of various degrees $p = 2, 3, 4, \ldots$ The criterion for the comparison of methods is the following: "better accuracy of the obtained approximate after the same amount of multiplications". (By *multiplication* we understand here the product of two linear operators.)

Assuming that each iteration step takes p multiplications, one finds that optimum hyperpower method is when $p = 3$. The corresponding algorithms have (thanks to their simplicity) the important advantage of comparatively small machine memory requirements in the case when A is a matrix.

However, it is easy to observe that for $p \geqslant 3$ it is enough to use $[p/2]+2$ multiplications for one iteration step. For example a slightly transformed formula (12) in the case $p = 5$ needs four multiplications instead of five ones (including the computation (11) of the residual operator B_i)

$$D_{i+1} = [I+(B_i+B_i^2)(I+B_i^2)]D_i. \tag{13}$$

It is easy to find formulas, which need even less than $[p/2]+2$ multiplications for greater p. That is achieved by using some more complicated schemes for evaluating the operator polynomial (12).

If for an arbitrary $\varepsilon > 0$ we want to achieve $\|A^{-1}-D_i\| \leqslant \varepsilon\|A^{-1}\|$, we have to perform i steps of the hyperpower method, where i can be computed from the approximate equation $\varrho^{p^i} \approx \varepsilon$, i. e. $i \approx \log\log_\varrho \varepsilon/\log p$. Hence the total amount of multiplications needed to achieve the given precision of the approximate D_i (for hyperpower method of p-th degree) is proportional to the function

$$\varphi(p) = \frac{m(p)+2}{\log p},$$

where $m(p)$ is defined as "the least number of multiplications needed to evaluate (without divisions) the polynomial $1+t+t^2+\ldots+t^{p-1}$".

Hence the Altman's problem of the choice of the optimum hyperpower method is equivalent to the problem of minimization of the function $\varphi(p)$ for $p = 2, 3, 4, \ldots$ The least value of $\varphi(p)$ for $p < 25$ is achieved for $p = 5$. The author does not know whether it is the minimal value of this function.

The higher hyperpower methods seem to be less convenient for practical purposes as less efficient or, as more complicated than the lower ones. Hence we may restrict our considerations to the cases of $p = 2$, $p = 3$, $p = 5$. In a very rough approximation comparing the corresponding values of $\varphi(p)$ we can say that using $p = 3$ instead of $p = 2$ we can spare about 5% of the total amount of multiplications. Using $p = 5$ instead of $p = 2$ we could spare about 14%. The total gain is not big. The author is inclined to believe that in many particular cases the second degree hyperpower method ($p = 2$) is (from the practical point of view) superior to any other method discussed.

3. Chebyshev methods in Hilbert spaces. Let A be a self-adjoint positive-definite bounded operator with the domain and the range in a Hilbert space $\mathfrak{H}$. Let m and M be the minimum and maximum eigenvalues of A respectively. It follows that $0 < m \leqslant M$. As in [1], let

$$D_0 = \frac{2}{M+m} I.$$

Then

$$\varrho = \|B_0\| = \sup_{\lambda=m,M}\left|1-\frac{2}{M+m}\lambda\right| = \frac{M-m}{M+m} < 1. \tag{14}$$

Hence the procedure (3) yields the approximate X_n with the corresponding error estimate (cf. (5)):

$$\|A^{-1}-X_n\| \leqslant \|A^{-1}-X_0\|\cdot\left(\frac{M-m}{M+m}\right)^n. \tag{15}$$

We have now more information about the operator B_0 and about norms of polynomials of B_0. For any polynomial $W_n(t)$ of a real variable we have now

$$\|W_n(B_0)\| = \sup_{\lambda\in Z}|W_n(\lambda)| \leqslant \sup_{-\varrho\leqslant t\leqslant\varrho}|W_n(t)|, \tag{16}$$

where Z is the spectrum of B_0.

So we can construct the best strategy procedure (cf. [6], [2]): choose for $\{W_n(t)\}$ the sequence of polynomials

$$\left\{\frac{T_n(t/\varrho)}{T_n(1/\varrho)}\right\}$$

of t, $n = 0, 1, 2, \ldots$, where $T_n(x) = \sum_{k=0}^{n} a_{nk}x^k$ is the Chebyshev polynomial

$$T_n(x) = \begin{cases} \cos n \arccos x & \text{for} \quad |x| \leqslant 1, \\ \frac{1}{2}[(x+\sqrt{x^2-1})^n+(x+\sqrt{x^2-1})^{-n}] & \text{for} \quad |x| \geqslant 1. \end{cases} \tag{17}$$

From (10), (14), (16), (17) we obtain for the corresponding error estimate:

$$\|A^{-1}-Y_n\| \leqslant \frac{\|A^{-1}-X_0\|}{T_n(1/\varrho)} < 2\cdot\|A^{-1}-X_0\|\cdot\left(\frac{\sqrt{M}-\sqrt{m}}{\sqrt{M}+\sqrt{m}}\right)^n. \tag{18}$$

The comparison of (15) and (18) shows that the Chebyshev procedure is superior to the ordinary power method, if $m < M$.

We shall now construct a procedure corresponding to the sequence of polynomials

$$\left\{\frac{T_{p^i}(t/\varrho)}{T_{p^i}(1/\varrho)}\right\}, \qquad i = 0, 1, 2, \ldots,$$

for an arbitrary integer $p \geqslant 2$. For this construction, the identity

(19) $$T_{p^{i+1}}(x) = T_p\big(T_{p^i}(x)\big)$$

will be basic. It follows from $\cos(p\alpha) = \cos\big(p\arccos(\cos\alpha)\big)$, where $\alpha = p^i \arccos x \in (0, \pi)$.

We start with $X_0 = 0$ and construct the sequence $\{U_i\}$, $i = 0, 1, \ldots$, where $U_i = Y_{p^i}$ (see (6)). Hence we obtain from (8)

$$A^{-1} - U_i = \frac{1}{T_{p^i}(1/\varrho)} T_{p^i}\left(\frac{1}{\varrho} B_0\right) A^{-1}.$$

It follows that $U_0 = A^{-1} - B_0 A^{-1} = D_0$. Let

$$\tau_i = T_{p^i}\left(\frac{1}{\varrho}\right), \quad Z_i = T_{p^i}\left(\frac{1}{\varrho} B_0\right),$$

so that we can write the relation

(20) $$R_i = I - U_i A = \frac{1}{\tau_i} Z_i$$

for the residual operator. But by (19) we have

(21) $$\tau_{i+1} = T_p(\tau_i), \quad Z_{i+1} = T_p(Z_i),$$

so that

$$R_{i+1} = I - U_{i+1} A = \frac{1}{\tau_{i+1}} Z_{i+1} = \frac{1}{T_p(\tau_i)} T_p(\tau_i R_i).$$

Introducing the coefficients a_{pk} of $T_p(x)$ we have

$$U_{i+1} = \frac{1}{\tau_{i+1}} [T_p(\tau_i I) - T_p(\tau_i R_i)] A^{-1} = \frac{1}{\tau_{i+1}} \sum_{k=0}^{p} a_{pk} \tau_i^k (I - R_i^k) \cdot A^{-1}.$$

Using the identity

$$1 - r^k = \sum_{s=0}^{k-1} r^s \cdot (1 - r)$$

for $k \geqslant 1$ and taking into account (20) we can write

$$U_{i+1} = \frac{1}{\tau_{i+1}} \sum_{k=1}^{p} a_{pk} \tau_i^k \sum_{s=0}^{k-1} R_i^s U_i.$$

Ordering the above sum consistently with the powers of R_i we can write the recurrence formula in the expanded form:

(22) $$U_{i+1} = [\zeta_{i0} I + \zeta_{i1} R_i + \ldots + \zeta_{i,p-1} R_i^{p-1}] U_i \quad (U_0 = D_0),$$

where

(23) $$\zeta_{ir} = \frac{\sum_{k=r+1}^{p} a_{pk} \tau_i^k}{\tau_{i+1}} \quad (r = 0, 1, 2, \ldots, p-1).$$

Let us note that $a_{p,p-j} = 0$ for odd j. Hence it follows from (23) that $\zeta_{i,p-j} = \zeta_{i,p-j-1}$ for odd j. So we can transform (22) in such a way that for $p \geqslant 4$ the computation of U_{i+1} can be done using only $[p/2]+2$ multiplications. For example, if $p = 5$ we can write (22) as follows:

(24) $$U_{i+1} = [I + (R_i + R_i^2)(\zeta_{i2} I + \zeta_{i4} R_i^2)] U_i.$$

Let us write

$$\sigma_i = \frac{1}{2\tau_i} = \frac{1}{2 T_{p^i}(1/\varrho)}.$$

Taking into account (17) and (14) we obtain the approximate equality

$$\sigma_i \approx \left(\frac{\sqrt{M} - \sqrt{m}}{\sqrt{M} + \sqrt{m}}\right)^{p^i} \approx \sigma_{i-1}^p \quad \left(\sigma_0 = \frac{\varrho}{2}\right).$$

It follows that $\lim_{i \to \infty} \sigma_i = 0$ and that this convergence is of degree p. By (21) we can construct the sequence $\{\sigma_i\}$ using the recursion formula

(25) $$\sigma_0 = \frac{\varrho}{2}, \quad \sigma_{i+1} = \frac{\sigma_i^p}{\sum_{k=0}^{\nu} c_{p,2k} \sigma_i^{2k}},$$

where $\nu = [p/2]$ and (cf. [5])

$$c_{p,2k} = 2^{2k+1-p} a_{p,p-2k} = (-1)^k \frac{p}{p-k} \binom{p-k}{k}.$$

In the same notation we obtain from (23)

(26) $$\zeta_{i,p-2s-1} = \frac{\sum_{k=0}^{s} c_{p,2k} \sigma_i^{2k}}{\sum_{k=0}^{\nu} c_{p,2k} \sigma_i^{2k}} \quad (s = 0, 1, \ldots, \nu).$$

It follows that for each $s = 0, 1, \ldots, \nu$ we have

(27) $$1 - \zeta_{i,p-2s-1} \approx \sigma_i^{2s+2} c_{p,2s+2}, \quad \lim_{i \to \infty} \zeta_{i,p-2s-1} = 1,$$

and the above convergence is of degree p.

We observe that the procedure (20), (22) is asymptotically convergent to the procedure (11), (12). Formulas (25), (26) are basic for the algorithms for computation of the coefficients $\zeta_{i,p-2s-i}$. In the case $p = 5$ we obtain for example

$$\sigma_0 = \frac{\varrho}{2}, \quad \zeta_{i,4} = \frac{1}{1-5\sigma_i^2+5\sigma_i^4}, \quad \zeta_{i,2} = (1-5\sigma_i^2)\cdot\zeta_{i,4},$$

$$\sigma_{i+1} = \sigma_i^5\cdot\zeta_{i,4}.$$

We note by comparing formulas (12) and (22), (13) and (24) that one step of the Chebyshev iteration procedure is only slightly more complicated than one step of the corresponding hyperpower method. We can assume therefore that the amount of work (needed to perform one step) is almost the same in both methods. The comparison of the efficiency of both methods can be deduced from (15) and (18) as follows. Let us consider for an $\alpha > 0$ the approximate equality

$$\left(\frac{M-m}{M+m}\right)^{p^i} \approx e^{-2\alpha} \approx 2\cdot\left(\frac{\sqrt{M}-\sqrt{m}}{\sqrt{M}+\sqrt{m}}\right)^{p^j}.$$

By assumption $\sqrt{m/M} \ll 1$ we can write the above in the following form:

$$i \approx \log_p\frac{M}{m} + \log_p\alpha,$$

$$j \approx \frac{1}{2}\log_p\frac{M}{m} + \log_p(\alpha+0{,}35).$$

In the case that M/m is sufficiently great we can (using the Chebyshev method instead of the hyperpower one) expect about 35-50% reduction of the total amount of steps. It seems reasonable to consider only the cases $p = 2$, $p = 3$, $p = 5$ for the Chebyshev methods too. The discussion of the efficiency of those methods remains the same as it was for the hyperpower methods.

Remark. Nearly in the same way as in [1] we can extend the above procedure on the general case when A is an arbitrary non-singular linear bounded operator in $\mathfrak{H}$. Let m^2 be the minimum and M^2 the maximum eigenvalues of the operator A^*A respectively. We use formulas (20), (22), (25), (26) starting with

$$U_0 = D_0 = \frac{2}{M^2+m^2} I \quad \text{and} \quad \varrho = \frac{M^2-m^2}{M^2+m^2}.$$

The sequence $\{U_i\}$ converges to A^{-1}. The corresponding error estimate is

$$\|A^{-1}-U_i\| \leqslant \|A^{-1}\|\cdot 2\left(\frac{M-m}{M+m}\right)^{p^i}.$$

4. An example of Chebyshev algorithm of the 2-nd degree. For $p = 2$ we can develop several special algorithms. One of them seems to yield minimum of the *additional work* in each iteration step. We understand by it the work which is not needed in one step of the ordinary hyperpower iteration of the degree 2. The algorithm is based upon formulas

$$\delta_0 = \delta, \quad \delta_1 = \frac{2-\varrho^2}{2}\delta^2,$$

$$\delta_{i+1} = \delta_i^2 - \tfrac{1}{2}(\delta_{i-1}^2-\delta_i)^2, \quad V_0 = \delta\cdot D_0,$$

$$V_{i+1} = [2\delta_{i+1}\cdot I - V_iA]\cdot V_i.$$

We have the following relations between $\{U_i\}$, $\{\sigma_i\}$, $\{V_i\}$ and $\{\delta_i\}$:

(28) $$\frac{\delta_i^2}{\delta_{i+1}} = \frac{\sigma_{i+1}}{\sigma_i^2} = \frac{1}{1-2\sigma_i^2} = \zeta_{i1}, \quad V_i = \delta_i\cdot U_i.$$

To establish the above formulas one should consider (20), (22) and (25) for $p = 2$. Theoretically, we could choose any positive number for δ. But in practice the best choice is $\delta = 2/(1+\sqrt{1-\varrho^2})$. If $\delta \geqslant 2/(1+\sqrt{1-\varrho^2})$, then the sequence $\{\delta_i\}$ is quadratically divergent to $+\infty$ and for $\delta < 2/(1+\sqrt{1-\varrho^2})$ the sequence $\{\delta_i\}$ is quadratically convergent to 0. Both cases are not convenient for practical purposes. In the case of $\delta = 2/(1+\sqrt{1-\varrho^2})$ we have $\lim_{i\to\infty}\delta_i = 1$ and the condition

(29) $$1 < \delta_{i+1} < \delta_i$$

is satisfied for each $i \geqslant 0$. In order to explain the above discussion we observe that from (27) and (28) it follows that the condition $\delta_i \leqslant 1$ implies $\lim_{j\to\infty}\delta_j = 0$ and the inequality $\delta_j < \delta_{j-1}$ for $j > i$. The condition $\delta_i \geqslant \delta_{i-1}$ implies similarly $\lim_{i\to\infty}\delta_i = +\infty$ and the inequality $\delta_j > \delta_{j-1}$ for $j > i$. That means that, if we want to avoid the cases $\lim_{j\to\infty}\delta_j = +\infty$ and $\lim_{i\to\infty}\delta_i = 0$, condition (29) is necessary for each i. It follows also that if condition (29) is satisfied for an index i, then it is fulfilled also for all former indices. From (28) and (29) we obtain the equivalent condition

$$\frac{\sqrt[2^{i+1}]{\sigma_{i+1}}}{\sigma_0} < \delta_0 < \frac{\sqrt[2^{i+1}]{\sigma_{i+1}}}{\sigma_0}\sqrt[2^{i+1}]{\zeta_{i0}}$$

and in the limit

$$\delta_0 = \frac{1}{\sigma_0}\lim_{k\to\infty}\sqrt[2^k]{\sigma_k} = \frac{2}{1+\sqrt{1-\varrho^2}}.$$

The proposed procedure for computing $\{\delta_i\}$ does not seem to be numerically stable. We must not bother about it as far as we can expect that the convergence of $\{U_i\}$ to A^{-1} will be fast enough. For any case it seems reasonable to check the condition (29). If it is not satisfied for an integer j, then we can switch in the second degree hyperpower routine

$$D_{i+1} = [2I - D_i A]\, D_i \qquad (i = j, j+1, \ldots)$$

starting with $D_j = \frac{1}{\delta_j} V_j$.

5. The cgT-algorithm for finite matrices. Let now $\mathfrak{H}$ be a finite dimensional space and A a self-adjoint, positive-definite matrix operator.

We observe that the set of matrix operators is a Hilbert space $\mathfrak{M}$ with usual operations and scalar product

(30) $$[X, Y] \underset{\mathrm{Df}}{=} \operatorname{spur}(X^* Y), \qquad X, Y \in \mathfrak{M}.$$

Let us consider the operator E_i of the orthogonal projection on the invariant subspace $\mathfrak{H}_i$ corresponding to the eigenvalue λ_i of A. Let the eigenvalues of A be $\lambda_1, \lambda_2, \ldots, \lambda_n$, $\lambda_i < \lambda_{i+1}$, $i = 1, 2, \ldots, n-1$. If a matrix W is a polynomial or a rational function of A, then W is a linear combination of $E_1, E_2, \ldots, E_n$:

(31) $$W = W(A) = \sum_{i=1}^{n} a_i E_i.$$

Here $a_i = W(\lambda_i)$ and n is the degree of the minimal polynomial of A. We note that any term of the sequences $\{U_i\}$ and $\{R_i\}$ defined above is of the form (31). So these terms belong to the n-dimensional subspace $\mathfrak{M}_A = \operatorname{Lin}\{E_1, \ldots, E_n\} \subset \mathfrak{M}$.

As mentioned in [1] or [6], we can use instead of minimum and maximum eigenvalues m and M the corresponding lower and upper bounds a, b: $0 < a \leqslant m$, $M \leqslant b$. However, the convergence of the Chebyshev method is much slower, if bounds a and b are not sufficiently near m and M. To start the computation with the hyperpower method we need only the bound b. The Chebyshev method needs both bounds, a and b. In the *ill-conditioned* case (it means when $m/M \ll 1$) a good lower bound $a > 0$ is only seldom available.

Following the same idea as in [6], we can use a combination of the Chebyshev method and of the conjugate gradients method. Using as before for b an appropriate upper bound of M, we choose an arbitrary positive number a such that in the interval $(0, a)$ remains only a small number k of the eigenvalues of A. (The dimension of the corresponding invariant subspaces $\mathfrak{H}_j$ is here arbitrary.) Then we apply the Chebyshev method which will liquidate the components $r_{k,i} E_i$ of the residual operator $R_k = I - AU_k$ corresponding to the eigenvalues $\lambda_i \geqslant a$. The remaining residual operator R_s is therefore approximately the sum of only k components:

$$R_s \approx \sum_{j=1}^{k} r_{s,j} E_j.$$

These components can be liquidated by applying the conjugate gradients method.

The proposed algorithm is based upon following formulas (cf. [7]):

1° start with $P_s = R_s$, $a_s = [R_s, R_s]$,

2° compute, for $r = s, s+1, s+2, \ldots$,

$$Q_r = AP_r, \qquad \beta_r = [P_r, Q_r],$$

$$U_{r+1} = U_r + \frac{a_r}{\beta_r} P_r,$$

$$R_{r+1} = R_r - \frac{a_r}{\beta_r} Q_r, \qquad a_{r+1} = [R_{r+1}, R_{r+1}],$$

$$P_{r+1} = R_{r+1} + \frac{a_{r+1}}{a_r} P_r.$$

The algorithm should finish not later than after k steps with $R_{r+1} = 0$. The above algorithm can be extended in the usual way on the case of an arbitrary non-singular matrix A.

References

[1] M. Altman, *An optimum cubically convergent iterative method of inverting a linear bounded operator in Hilbert space*, Pacific Journal of Math. 10, 4 (1960).

[2] A. Blair, N. Metropolis, J. v. Neumann, A. M. Taub, M. Tsingou, *A study of a numerical solution to a two-dimensional hydrodynamical problem*, MTAC (1959), XIII, Appendix II, p. 166-184.

[3] М. К. Гавурин, *Применение полиномов ...*, Усп. мат. наук 5, 3 (37) (1950), p. 156-160.

[4] В. Л. Гончаров, *Теория интерполирования и приближения функций*, 1954, p. 310.

[5] S. Paszkowski, *The theory of uniform approximation*, Rozprawy Matem. 26 (1962), p. 16.

[6] E. Stiefel, *Kernel polynomials in linear algebra and their numerical applications*, NBS Appl. Math. Series 49 (1958), p. 1-22.

[7] — *Einige Methoden der Relaxationsrechnung*, ZAMP 3 (1953), p. 1-31.

Reçu par la Rédaction le 13. 2. 1963