Ivo Marek

Kellogg's iterations with minimizing parameters

Commentationes Mathematicae Universitatis Carolinae, Vol. 4 (1963), No. 2, 53--64

Persistent URL: http://dml.cz/dmlcz/104930

Commentationes Mathematicae Universitatis Carolinae
4, 2 (1963)

KELLOGG'S ITERATIONS WITH MINIMIZING PARAMETERS

Ivo MAREK, Praha

In the Kolomý's papers [4a], [4b] a method for the construction of the eigenelements of the linear symmetrizable operator T in Hilbert space is given, which differs from the usual process of succesive approximations

(1) $$x_{n+1} = T\, x_n \, .$$

Kolomý makes use of

(2) $$y_{n+1} = \lambda_{(n)}\, T\, y_n \, ,$$

where $\lambda^{-1}_{(n)}$ realize minimum of the function of real variable

(3) $$\varphi_n\,(\nu) = \| T\, y_n - \nu y_n \|^2 \, .$$

We shall show that Kolomý's method can be generalized in the natural way also for linear unsymmetrizable operators. Also the assumption that the transformation T operates in a Hilbert space can be replaced by the assumption that the concerned transformation operates in a Banach space.

Let X be the complex Banach space, X' the adjoint space of continuous linear forms on X and $[X]$ the space of bounded linear transformations mapping X into itself. Let us assume that the forms $y'_n \in X'$, $y' \in X'$ have the following property

(4) $$\lim_{n\to\infty} y'_n(x) = y'(x)$$

for every vector $x \in X$.

Let μ_o be a simple dominant eigenvalue of the operator $T \in [X]$. According to [7]

(5) $$x_{(n)} \to x_o , \quad \mu_{(n)} \to \mu_o$$

holds, where

(6) $$x_{(n)} = \frac{T^n x^{(o)}}{y_n'(T^n x^{(o)})}, \quad \mu_{(n)} = \frac{y_n'(T^{n+1} x^{(o)})}{y_n'(T^n x^{(o)})} ,$$

$x^{(o)} \in X$ is a suitable vector and x_o is an eigenvector of the operator T corresponding to the eigenvalue μ_o.

Let be

(7) $$y_{n+1} = \lambda_{(n)} \; T y_n , \quad y_o = x^{(o)} ,$$

where

(8) $$\lambda_{(n)} = \frac{y_n'(y_n)}{y_n'(T y_n)} .$$

Then the relations

(9) $$y_n \to w_o = c_1 x_o , \quad |c_1| > o , \quad \lambda_{(n)} \to \lambda_o = \mu_o^{-1}$$

follow from the theorem 3 of [7].

The form of the functionals $y_n' \in X'$ in the formula (8) is almost arbitrary. Therefore we shall ask which functionals y_n' are the most effective ones. The conception of the effectiveness of the process of the concerned type we shall put some of its extremal properties.

Let T' be the adjoint operator of T. Let the real function of the real variable $\nu \in (-\infty, +\infty)$

(10) $$\psi_n(\nu) = \ell_n'(T y_n - \nu y_n) , \text{ where } \ell_n' = T' y_n' - \nu y_n'$$

be given. Let the sequence $y_n' \in X'$ be defined by the formula

(11) $$y_{n+1}' = \lambda_{(n)} T' y_n' , \quad y_o' = x'^{(o)} \in X' ,$$

where the parameters $\lambda_{(n)}$ are to be determined from the condition that the function ψ_n shall catch the minimal value. Then

(12) $$\lambda^{-1}_{(n)} = \frac{y'_n(T^{n+1}\ y_o)}{y'_n\ (T^n\ y_o)} = \frac{y'_o(T^{2n+1}\ y_o)}{y'_o(T^{2n}\ y_o)} ,$$

from which we can infer that the "effective" process gives the approximations of μ_o of degree n which are equal to the approximation of degree 2 n given by the usual iteration process (6) with $y'_n = y'_o$ for $n = 0, 1, \dots$. This property can be successfully used for the symmetric operators in Hilbert space. In this case the process (11) is identical to the (7), therefore half of the computations falls off. We have

$$y'_n\ (x) = (x,\ y'_n) = (x,\ y_n)$$

and

$$\psi_n(\nu) = (T\ y_n - \nu\, y_n\ ,\ T\ y_n - \nu\, y_n) = \psi_n(\nu)$$

and this is the process of Kolomý.

T h e o r e m 1. Let μ_o be a real simple dominant eigenvalue of the operator $T \in [X]$. Let $x^{(o)} \in X$, $x'^{(o)} \in X'$ be the initial vectors of iterations (7), (11) such that

$$y'_o = x'^{(o)}\ ,\quad y'_o\ (B_1\ x^{(o)}) \neq 0\ ,$$

where $B_1 = \lim_{n\to\infty} \mu_o^{-n}\ T^n$. Let $y'_n(y_n) > 0$ hold for $n = 0, 1, \dots$. The values $\lambda_{(n)}$ defined by (8) let be real. Then

(13) $$y_n \to w_o\ ,\quad y'_n \to w'_o\ ,$$

(14) $$\lambda_{(n)} \to \lambda_o = \mu_o^{-1}$$

hold, where w_o , w'_o are eigenvectors of the operators T, T' corresponding to the eigenvalue μ_o . The value $\lambda^{-1}_{(n)}$

realizes the minimum of the function (10) for $n = 0, 1, \ldots$.

P r o o f . The formula $B_1 = \lim_{n\to\infty} \mu_o^{-n} T^n$ follows from [7] theorem 1.

If we put

$$z_n'(x) = \frac{y_n'(x)}{y_n'(x^{(o)})} = \frac{y_o'(T^n x)}{y_o'(T^n x^{(o)})}, \quad z'(x) = \frac{y_o'(B_1 x)}{y_o'(B_1 x^{(o)})},$$

then we have

$$\lambda_{(n)} = \frac{z_n'(T^n x^{(o)})}{z_n'(T^{n+1} x^{(o)})},$$

$$|z_n'(x) - z'(x)| \leq \left| z_n'(x) - \frac{y_o'(B_1 x)}{y_o'(\mu_o^{-n} T^n x^{(o)})} \right| +$$

$$+ \left| \frac{y_o'(B_1 x)}{y_o'(\ {}_o^{-n} T^n x^{(o)})} - \frac{y_o'(B_1 x)}{y_o'(B_1 x^{(o)})} \right| .$$

According to [7] lemma 2 we have

$$(15) \quad \| \mu_o^{-n} T^n x - B_1 x \| \leq c_2 \left| \frac{\mu_1}{\mu_o} \right|^n . \| x \| , \mu_o > \mu_1 > \mu ,$$

where μ is the radius of the smallest circle with the centre in the origine in which the whole spectrum $\sigma(T)$ of the operator T lies with the exception the point μ_o.

Clearly the inequalities

$$\left| z_n'(x) - \frac{y_o'(B_1 x)}{y_o'(\mu_o^{-n} T^n x^{(o)})} \right| = \left| \frac{y_o'(\ {}_o^{-n} T^n x^{(o)})}{y_o'(\mu_o^{-n} T^n x^{(o)})} - \right.$$

$$(16) \quad \left. - \frac{y_o'(B_1 x)}{y_o'(\mu_o^{-n} T^n x^{(o)})} \right| \leq \frac{c_2}{c_3} \| y_o' \|_{X'} . \| x \|_X . \left(\frac{\mu_1}{\mu_o} \right)^n$$

hold, where $c_3 = \inf_n |y_o'(\mu_o^{-n} T^n x^{(o)})| > 0$.

Similarly

$$(17) \quad \left| \frac{y_o'(B_1 x)}{y_o'(\mu_o^{-n} T^n x^{(o)})} - z'(x) \right| \leq \frac{c_2}{c_3} . \frac{\| y_o' \|_{X'}}{|y_o'(B_1 x^{(o)})|} \| x \|_X \| B_1 \| . \left(\frac{\mu_1}{\mu_o} \right)^n .$$

According to (16), (17) and according to [7] theorem 3 we obtain the validity of (13) and (14).

We shall still prove that the value $\lambda_{(n)}^{-1}$ realizes the minimum of the function (10). This statement follows from the relations

$$\frac{d}{d\nu}\psi_n(\nu) = 2\nu\, y_n'(y_n) - 2\, y_n'(T\, y_n)\ ,$$

$$\frac{d^2}{d\nu^2}\psi_n(\nu) = 2\, y_n'(y_n)\ ,$$

$$\frac{d}{d\nu}\psi_n(\nu)\Big|_{\nu=\lambda_{(n)}^{-1}} = 0\ , \quad \frac{d^2}{d\nu^2}\psi_n(\nu) = 2\, y_n'(y_n) > 0$$

The iterations of the type (2) were introduced by I.A. Birger [2] and by J. Kolomý [4a], [4b], where the proof of the convergence of the mentioned processes was given for the compact symmetric and for the compact symmetrizable operators mapping a Hilbert space into itself. But the formula (8) is also used for the construction of the eigenvalues of the unsymmetric operators, often without any motivation of the convergence ([6] pp. 72-73). In the monography [6] it is expected on the ground of physical ideas that the iteration algorithm (8) has some advantages in comparison with the other ones.

The statement of the theorem 1 can be utilized without difficulty for the construction of the eigenelements of the equations of the type

(18) $\quad Lx = Bx + \lambda Cx\ , \quad L'x' = B'x' + \lambda C'x'\ ,$

where L, B, C are, in general, unbounded linear operators mapping the domains $\mathcal{D}(L)$, $\mathcal{D}(B)$, $\mathcal{D}(C)$ into X. We shall assume that $\mathcal{D}(L) \subset \mathcal{D}(B) \subset \mathcal{D}(C)$ and that $\mathcal{D}(L)$ is dense in X. Then there exist the adjoint operators L', B', C'.

We define the iteration processes

(19) $w^{(n)} = C w_{(n)}$, $L w_{n+1} = B w_{n+1} + w^{(n)}$, $w_{(n+1)} = \lambda_{(n)} w_{n+1}$,
$w_{(o)} = x^{(o)}$,

(20) $L' z'_{(n)} = B' z'_{(n)} + z'^{(n)}$, $z'_{n+1} = C' z'_{(n)}$, $z'^{(n+1)} =$
$= \lambda_{(n)} z'_{n+1}$, $z'^{(o)} = x'^{(o)}$

(21) $$\lambda_{(n)} = \frac{z'^{(n)}(w_{(n)})}{z'^{(n)}(w_{n+1})} ,$$

where the initial vectors $x^{(o)}$, $x'^{(o)} = z'^{(o)}$ are chosen so that

(22) $$z'^{(o)}(B_1 x^{(o)}) \neq 0 ,$$

where $B_1 = \lim_{n\to\infty} \mu_o^{-n} T^n$, $T = (L - B)^{-1} C$. The existence of the operator $(L - B)^{-1} \in [X]$ is assumed.

T h e o r e m 2 . Let μ_o be a real simple dominant eigenvalue of the operator $T = (L - B)^{-1} C$, $T \in [X]$. Let the initial vectors $x^{(o)}$, $x'^{(o)}$ be given so that (22) holds. Let

(23) $$z'^{(n)}(w_{(n)}) > 0 \quad \text{for} \quad n = 0, 1, \ldots$$

be fulfilled. Then we have

(24) $$w_{(n)} \to x_o , \quad z'_{(n)} \to x'_o$$
$$\lambda_{(n)} \to \lambda_o = \mu_o^{-1} ,$$

where x_o, x'_o are the eigenvectors of the equation (18) corresponding to the characteristic value λ_o. The value $\lambda_{(n)}^{-1}$ realizes the minimum of the function

(25) $$\zeta_n(\nu) = z'_{n+1}(w_{n+1} - \nu w_{(n)}) - \nu z'_n(w_{n+1} - \nu w_{(n)}) .$$

E x a m p l e s .

Example 1. Let X be a complex Hilbert space, let T be a nonnegative linear operator, i.e. $(Tx, x) \geqq 0$ for $x \in X$. Let us put

$$y_n'(x) = (x, y_n) ,$$

where y_n, $n = 0, 1, \ldots$ are defined in (2) and

$$\varphi_n(\nu) = (T\ y_n - \nu y_n\ ,\ T\ y_n - \nu y_n) = \psi_n(\nu) .$$

Let the inequality

$$(x^{(o)},\ B_1\ x^{(o)}) \neq 0$$

be satisfied for the initial vector $x^{(o)} \in X$ where $B_1 = \lim_{n\to\infty} \mu_o^{-n}\ T^n$. Let $\|T\| = \mu_o < +\infty$ be simple isolated eigenvalue of the operator T. Thus all the assumptions of the theorem 1 are fulfilled.

Example 2. Let Y be a real Banach space with the cone K of the positive elements. Let X be the complex extension of Y ([5]) i.e. the space of the pairs $(x, y) = z$, $x \in Y$, $y \in Y$ with the norm

$$\|z\|_X = \sup_{0 \leqq \vartheta \leqq 2\pi} \|x \cos \vartheta + y \sin \vartheta\|_Y$$

or with some equivalent norm.

We assume that K is volume-type productive cone ([5]) i.e. the interior int K is nonempty and to every $y \in Y$ there exist sequences $\{y_{1\ n}\}$, $\{y_{2\ n}\}$, $y_{1\ n} \in K$, $y_{2\ n} \in K$ so that $y = \lim_{n\to\infty} (y_{1n} - y_{2n})$. Let W, $W \in [Y]$, be a compact operator such that $T = W + \tau I$ is strongly K-positive operator for some p (I denotes identical operator), i.e. for $x \in K$, $x \neq \emptyset$, $T^p \in$ int K for some natural p. About these assumptions there exists a real simple eigenvalue ν_o of the operator W having the maximal real part between the all points of the spectrum of the operator W .

Let $K' \subset Y'$ be the adjoint cone of K . If we put the initial vectors $x^{(o)} \in K$, $\alpha^{1(o)} \in K'$ we obtain according to theorem 1 the relations

$$y_n \to x_o , \quad y_n' \to x_o' , \quad \lambda_{(n)}^{-1} \to \nu_o + \tau ,$$

where $Wx_o = \nu_o x_o$, $W'x_o' = \nu_o x_o'$ and y_n, y_n' , $\lambda_{(n)}$ are defined by (7), (11), (8). The value $\lambda_{(n)}^{-1}$ realizes the minimum of the function (10).

Proof. Let us put

$$B_T = \frac{1}{2\pi i} \int_{C_{\mu_o}} R(\lambda, T)\, d\lambda , \quad B_W = \frac{1}{2\pi i} \int_{C_{\nu_o}} R(\lambda, W)\, d\lambda ,$$

where C_{μ_o} is the circle with the centre μ_o such that $\overline{\text{int } C_{\mu_o}} \cap \sigma(T) = \{\mu_o\}$ and similarly $\overline{\text{int } C_{\nu_o}} \cap \sigma(W) = \{\nu_o\}$. It can be easily shown that $B_T = B_W$. Thus $B_W x \in \text{int } K$ for $x \in K$, $x \neq \emptyset$.

It is well known ([5]) that there exists a positive simple dominant eigenvalue μ_o of the operator T and one and only one eigenvector $x_o \in K$, $\| x_o \| = 1$ corresponding to μ_o . Thus $\nu_o = \mu_o - \tau$ is real. The inequality

$$\text{Re } \nu \geqq \text{Re } \nu_o , \quad \nu \in \sigma(W)$$

is impossible for

$$| \nu + \tau | < \nu_o + \tau$$

and therefore Re $\nu_o >$ Re ν for all $\nu \in \sigma(W)$, $\nu \neq \nu_o$.

Remark. To the category of linear operators described in the example 2 belong the linear operators which operate in finite dimensional spaces and which are represented by the matrices $T = (t_{jk})$ with $t_{jk} > 0$ for $j \neq k$ and with arbitrary real t_{jj} ([1]).

E xample 3. Let G be a connected region in the three-dimensional Euclidean space R_3. Let Ω be the set of vectors of R_3 with the unit norm and let be $\mathcal{O} = G \times \Omega$. Let $X = L_2(\mathcal{O}) \times \dots \times L_2(\mathcal{O})$ be the cartesian product of m spaces $L_2(\mathcal{O})$, where $L_2(\mathcal{O})$ consists of the functions integrable with the square on $\mathcal{O}$. The scalar product and the norm are defined in X as usually by the formulae

$$(x,y) = \sum_{j=1}^{m} \int_{\mathcal{O}} x_j\ (\vec{r}, \vec{\Omega})\ \overline{y_j\ (\vec{r}, \vec{\Omega})}\ dr\ d\,\Omega\ ,$$

$$\|x\| = (x,\ x)^{\frac{1}{2}}$$

($\bar{\alpha}$ denotes the complex conjugate number of α), $x = (x_1, \dots, x_m)$, $y = (y_1, \dots, y_m)$.

Let $K \subset X$ be the set of the vector-functions nonnegative almost everywhere in $\mathcal{O}$. The set K is evidently a productive cone in X.

We shall investigate the so-called fundamental equations of the reactor-physics in the multigroup - energy approximation of the neutron transport theory [6] p. 285. The mentioned system of equations can be symbolically written as ([8])

(26) $$Lx = Bx + \lambda Cx\ , \qquad L'x' = B'x' + \lambda C'x'\ ,$$

where x denotes the neutron-flux, x' denotes the "neutron-price" and L, B, C denote the operators describing diffusion, absorbtion, scattering and fission in the given medium.

According to [8] $T = (L - B)^{-1} C$ is compact absolutely K-positive operator. Therefore there exists a positive simple dominant eigenvalue μ_0 of T and T' and to this eigenvalue correspond the eigenvectors x_o, x'_o of the operators T, $T = C'(L' - B')^{-1}$. These vectors are positive almost everywhere in $\mathcal{O}\!\mathcal{L}$. The mentioned eigenelements can be constructed by means of the iterations (19), (20), (21), where

$$\lambda_{(n)} = \frac{(w_{(n)}, z'^{(n)})}{(w_{n+1}, z'^{(n)})} , \tag{27}$$

$$\zeta_n(\nu) = (w_{n+1}, - \nu w_{(n)}, z'_{n+1} - \nu z'^{(n)}) . \tag{28}$$

Let be $x^{(o)} \in K$, $x'^{(o)} \in K$. It follows from the absolute K-positivity of T that the assumptions of the theorem 2 are fulfilled. Especially we have

$$(w_{(n)}, z'^{(n)}) > 0 ,$$

$$\frac{d}{d\nu} \zeta_n(\nu)\Big|_{\nu=\lambda_{(n)}^{-1}} = 0 , \quad \frac{d^2}{d\nu^2} \zeta_n(\nu) = 2 (w_{(n)}, z'^{(n)})$$

for $n = 0, 1, \ldots$.

Example 4. The necessary properties required in the theorem 2 has also the system of multigroup diffusion equations [6] pp. 113-135. The mentioned system can be obtained by substituting the operators L, B, C ; L', B', C' in (26) for other operators which describe the physically simpler approximation.

In the Hilbert spaces we can choose the values $\lambda_{(n)}$, $n = 0, 1, \ldots$ so that the functions

$$\eta_n(\nu) = \| \nu T y_n - y_n \|^2$$

will catch the minimal values. We obtain the formula

$$\lambda_{(n)} = \frac{\mathrm{Re}\ (y_n, Ty_n)}{(Ty_n, Ty_n)} . \tag{29}$$

The calculations show that the process (7), (8), (11) is more effective than the process (7), (8), (29) for the unsymmetrizable T .

In real Hilbert spaces H. Bueckner [3] used the formula (29) also to the construction of eigenvalues of non-linear equations.

References

[1] G. BIRKHOFF, R.S. VARGA. Reactor criticality and nonnegative matrices. Journ.Soc.Industr.Appl. Math. Vol.6 (1958), 354-377.

[2] I.A. BIRGER. Někotoryje matěmatičeskije rešenija inženěrnych zadač. Oborongiz, Moskva 1956.

[3] H.F. BUECKNER. An iterative method for solving non-linear integral equations. PICC Symposium, Rome 1960, 614-643.

[4a] J. KOLOMÝ. On convergence of the iterative methods. Comment.Math.Univ.Carol. 1,3 (1960),18-24.

[4b] J. KOLOMÝ. On the solution of homogeneous functional equations in Hilbert space. Comment.Math. Univ.Carol. 3,4 (1962),36-46.

[5] M.G.KREJN, M.A.RUTMAN. Linějnyje operatory ostavljajuščije invariantnym konus v prostranstvě Banacha. Usp.mat.nauk III (1948),3-97 .

[6] G.I.MARČUK. Čislennyje mětody rasčota jaděrnych reaktorov. Atomizdat, Moskva 1958.

[7] I. MAREK. Iterations of linear bounded operators and Kellogg s iterations in non-self adjoint eigenvalue problems. Czech.Math.Journ.12 (1962), 536-554.

[8] I. MAREK. Neutron transport theory in multigroup energetical approximation. Comment.Math.Univ.Carol. 3,3 (1962), 3-13.