On the Kellogg method and its variants for finding of eigenvalues and eigenfunctions of linear self-adjoint operators

J. Kolomý

(Prof. Dr. F. A. Willers zu seinem 100. Geburtstag gewidmet)

The Kellogg method, Birger’s method and the modified Birger procedure for determination of eigenvalues and eigenvectors of linear nonnegative and self-adjoint operators are investigated.

1. Introduction

In this note we show that the methods mentioned above converge even in the case, when the starting approximation is only different from zero. We also give the estimate for the distance of eigenvalues of two different operators. Moreover, some conditions for the starting approximation equivalent to that of Theorem 1 are established.

Let us recall that Birger [2] has proposed his method without any mathematical justification. However, he has found (on engineering problems) that his method has some advantages in comparison with the other ones, see also Marchuk [15], where a similar observation has been done on the ground of physical ideas. The convergence proofs of the Birger and the modified Birger methods for compact symmetrizable operators were given in [7, 8] under the condition that the starting approximation is not orthogonal to annihilator of the certain eigenspace. Later Buckner [3] independently proposed the same method as Birger and proved its convergence for the class of linear and nonlinear compact operators having some decomposition properties. Further results in these topics have been obtained under various hypotheses by Marek [14], Petryshyn [16] and the author [9—12]. We refer the reader also for instance to [1, 4, 13, 16] for some further and related methods. Let us note that the Birger method and the Buckner [3] results have been applied for instance by Birger [2], Marchuk [15], Conway and Thomas [5] and Thomas [18].
2. Convergence theorems

Let $X$ be a real Hilbert space with the scalar product $\langle \cdot , \cdot \rangle$, $A : X \rightarrow X$ a linear non-negative and self-adjoint operator defined on $X$. By saying that $A$ is non-negative we mean that $\langle Au, u \rangle \geq 0$ for each $u \in X$. Since $A$ is self-adjoint and is defined on all of $X$, $A$ is bounded by the closed graph theorem. The spectrum $\sigma (A)$ of $A$ lies in the segment $[m, \lambda_1]$, where $m = \inf \langle Au, u \rangle: \|u\| = 1$, $m \neq 0$, and $\lambda_1 = \sup \langle Au, u \rangle: \|u\| = 1$. The symbol $\{E_i\}$ stands for the spectral resolution of identity corresponding to the self-adjoint operator $A$.

Under the assumptions stated above the following approximate methods for determination of the eigenvalues and eigenvectors will be considered:

(i) the Kellogg method:

$$\alpha_{n+1} = \|Au_n\|, \quad u_{n+1} = \alpha_{n+1}^{-1} Au_n \quad (n = 0, 1, 2, \ldots),$$

where the starting approximation $u_0 \in X$ is such that $u_0 \notin \ker A, \|u_0\| = 1$. According to our assumptions $\alpha_n > 0$ and $\|u_n\| = 1$ for each $n$;

(ii) the modified Birger method:

$$\mu_{n+1} = \langle Av_n, v_n \rangle: \|v_n\|^2, \quad v_{n+1} = \mu_{n+1}^{-1} Av_n,$$

where the initial approximation $v_0 \in X$ satisfies the conditions $v_0 \notin \ker A, \|v_0\| = 1$;

(iii) the Birger method:

$$q_{n+1} = \langle Ay_n, y_n \rangle: \|Ay_n\|^2, \quad y_{n+1} = q_{n+1}^{-1} Ay_n,$$

where the starting approximation $y_0 \in X$ is such that $y_0 \notin \ker A, \|y_0\| = 1$.

**Lemma 1:** Let $A : X \rightarrow X$ be a linear non-negative and self-adjoint operator, $u_0, v_0, y_0 \in X$ the starting approximations of the methods (1) -- (3), respectively, such that $u_0 = v_0 = y_0, u_0 \notin \ker A, \|u_0\| = 1$. Then:

(i) $\mu_n > 0, \quad q_n > 0, \quad v_n \neq 0, \quad y_n \neq 0$;

(ii) $\mu_n \leq \alpha_n \leq q_n^{-1} \leq \mu_{n+1} \leq \alpha_{n+1} \leq q_n^{-1}$

(iii) $v_{n+1} = \left( \prod_{i=1}^{n+1} (\alpha_i q_i^{-1}) \right) u_{n+1}$

(iv) $y_{n+1} = \left( \prod_{i=1}^{n+1} \alpha_i \right) u_{n+1}$

for each $n \geq 1$.

**Proof:** Since $A$ is non-negative and self-adjoint, $\langle Au, v \rangle^2 \leq \langle Au, u \rangle \langle Av, v \rangle$ for each $u, v \in X$. According to our assumption $u_0 \notin \ker A$ we obtain

$$0 < \alpha_1^4 = \|Au_0\|^4 = \langle Au_0, Au_0 \rangle^2 \leq \langle Au_0, u_0 \rangle \langle A^2 u_0, Au_0 \rangle = \mu_1 \langle A^2 u_0, Au_0 \rangle.$$

Hence $\mu_1 > 0$ and $v_1 = \mu_1^{-1} Au_0 = \mu_1^{-1} Au_0 = \alpha_1 \mu_1^{-1} u_0 \neq 0$. Now assume that $\mu_k > 0, v_k \neq 0$ for each $k$ ($k = 1, 2, \ldots, n$). We are going to show that $\mu_{n+1} > 0$ and $v_{n+1} \neq 0$. From

$$0 < \alpha_{n+1}^4 = \langle Au_{n+1}, Au_{n+1} \rangle^2 \leq \mu_{n+1} \langle A^2 u_{n+1}, Au_{n+1} \rangle$$
we get that $\mu_{i+1} > 0$. Moreover, $u_n = v_n \|v_n\|^{-1}$ and

$$v_{n+1} = \mu_{n+1}^{-1} A v_n = \mu_{n+1}^{-1} \|v_n\| \cdot A \left( \frac{v_n}{\|v_n\|} \right) = \mu_{n+1}^{-1} \|v_n\| \cdot A u_n = \alpha_{n+1} \mu_{n+1}^{-1} \|v_n\| \cdot u_{n+1}$$

$$= \alpha_{n+1} \mu_{n+1}^{-1} \alpha_n u_{n+1} = \cdots = \left( \prod_{k=1}^{n+1} (\alpha_k \mu_k^{-1}) \right) u_{n+1} \neq 0.$$  

The proofs of the equality (iv) and the facts that $q_n > 0, y_n \neq 0$ are quite similar.

(ii) $\mu = \langle A v_{n-1}, v_{n-1} \rangle \|v_{n-1}\|^2 = \langle A u_{n-1}, u_{n-1} \rangle \leq \|A u_{n-1}\| \cdot \|u_{n-1}\| = \|A u_{n-1}\| = \alpha_n$,

$$q_n^{-1} = \langle A y_{n-1}, y_{n-1} \rangle^{-1} \|A y_{n-1}\|^2 = \|A u_{n-1}\| \langle A u_{n-1}, u_{n-1} \rangle^{-1} = \alpha_n (\alpha_n u_{n-1}) \geq \alpha_n.$$  

Furthermore, according to the generalized Schwarz inequality

$$q_n^{-1} = \langle A u_{n-1}, u_{n-1} \rangle^{-1} \|A u_{n-1}\|^2 \leq \langle A u_{n-1}, A u_{n-1} \rangle \cdot \|A u_{n-1}\|^{-2} \cdot \|A u_{n-1}\|^2$$

$$= \langle A^2 u_{n-1}, A u_{n-1} \rangle \|A u_{n-1}\|^{-2} = \langle A u_{n-1}, u_{n-1} \rangle = \mu_{n+1}.$$  

This completes the proof Lemma 1.

Lemma 1 enables a slight extension of the corresponding results of [11, 12] for non-negative operators.

**Theorem 1:** Let $A : X \to X$ be a linear non-negative self-adjoint operator, $u_0, v_0, y_0 \in X$ the starting approximation of the methods (1) — (3), respectively, such that $u_0 = v_0 = y_0$, $u_0 \notin \ker A$, $\|u_0\| = 1$.

Then the following conclusions are valid:

(i) If $u_0$ is such that $E u_0 = u_0$ for each $\lambda < \lambda_1$, then the sequences $(\alpha_n), (\mu_n), (q_n^{-1})$ are monotone increasing and converge to $\lambda_1$.

(ii) If $\lambda_1$ (not necessarily an isolated point of $\sigma(A)$ with finite multiplicity) is an eigenvalue of $A$ and $u_0$ is such that $u_0 \notin \ker (A - \lambda_1 I)^k$, then $(\alpha_n), (\mu_n), (q_n^{-1}) \to \lambda_1$ as $n \to \infty$.

(iii) If $\lambda_1$ is an isolated point of $\sigma(A)$ and $u_0 \in \ker (A - \lambda_1 I)^k$, then each of the following sequences $(u_n), (v_n), (y_n)$ converge to one of the eigenvectors of $A$ corresponding to $\lambda_1$.

Under the additional hypothesis that $A$ is compact we shall prove that the methods (1) to (3) are convergent even in the case when the starting approximations $u_0, v_0, y_0$ are only different from zero.

We shall use the following

**Lemma 2:** Under the assumptions of Lemma 1 the sequences $(\|v_n\|), (\|y_n\|)$ are monotone increasing and decreasing, respectively, and convergent.

Proof: It relies on Lemma 1 and the arguments of the proof of Lemma 2 [9] and [8], where the assumptions of positivity of $A$ is superfluous.

**Theorem 2:** Let $A : X \to X$ be a linear non-negative self-adjoint operator such that $A^n$ is compact for some positive integer $n_0$. Assume that the starting approximations $u_0, v_0, y_0$ are such that $u_0 = v_0 = y_0$ and $\|u_0\| = 1$.

Then there is valid:

(i) The sequences $(\alpha_n), (\mu_n), (q_n^{-1})$ are monotone increasing and converge to some positive eigenvalue $\lambda^*$ of $A$.

(ii) The sequences $(u_n), (v_n), (y_n)$ converge to one of the eigenvectors corresponding to $\lambda^*$.
Proof: According to Lemma 1 we have that \((x_n), (\mu_n), (q^{-1}_n) \times \lambda^*\) as \(n \to \infty\) and \(0 < \lambda^* \leq \lambda_1 = ||A||\). From (2) it follows that

\[
||v_{n+1} - v_n||^2 = ||\mu_n^{-1}Av_n - v_n||^2 = ||v_{n+1}||^2 - ||v_n||^2
\]

(4)

for each \(n\). Put \(z_n = v_{n+1}\) for each \(n \geq 1\). By Lemma 2 the sequence \((z_n - v_n)\) converges to 0 as \(n \to \infty\). Moreover, the sequence \((z_n)\) is bounded and our hypotheses imply that \(A\) is compact [20: ch. 12]. Hence \((z_n)\) contains a convergent subsequence \((z_{n_k})\). Denote \(v^* = \lim_{k \to \infty} z_{n_k}\). In view of (4) we conclude that \(v_{n_k} \to v^*\). Since \(||v_n||\) is monotone increasing and \(v_0 \neq 0\) we get that \(v^* \neq 0\). But continuity of \(A\) implies that \(A0^* = \lambda^*v^*\) and \(\mu_n \to \lambda^* = (Au^*, u^*) \cdot ||u^*||^{-2}\). From (4) it follows that \((v_n)\) is a Cauchy sequence. Since it contains a converging subsequence \((v_{n_k})\), the whole sequence \((v_n)\) converges to \(v^*\).

By Lemma 1

\[
v_n = \left( \prod_{i=1}^{n} (x_i\mu_i^{-1}) \right) u_n, \quad y_n = \left( \prod_{i=1}^{n} (x_iq_i) \right) u_n = \left( \prod_{i=1}^{n} (\mu_i\mu_i^{-1}) \right) u_n.
\]

Since \(\lim_{n \to \infty} ||v_n|| = \sup_{n=1,2,\ldots} ||v_n|| = \infty (\alpha \mu_i^{-1})\) is finite and positive, there exist \(\lim_{n \to \infty} u_n = u^*\), \(\lim_{n \to \infty} y_n = y^*\) and

\[
\prod_{i=1}^{\infty} (x_i\mu_i^{-1}) u^* = y^*, \quad \prod_{i=1}^{\infty} (x_iq_i) u^* = y^*.
\]

Hence \(u^* \in \ker (A - \lambda^*I), y^* \in \ker (A - \lambda^*I)\) and \(u^* \neq 0, y^* \neq 0\). Theorem 2 is proved.

Using arguments similar to that of [8], one can extend the result of Theorem 2 to symmetrizable compact operators.

Let \(B : X \to X, C : X \to X\) be linear self-adjoint operators, \(\lambda_0\) an eigenvalue of \(B\) such that \(\lambda_0 \in \sigma(C), \lambda^*\) an eigenvalue of \(C\). We shall say that \(\lambda^*\) is nearest to \(\lambda_0\) from the both sides if there exists a positive number \(e\) such that \(\lambda^* \in J = (\lambda_0 - e, \lambda_0 + e)\) and \(\lambda \in \sigma_p(C), \lambda = \lambda^*\) imply that \(\lambda \notin J\), where \(\sigma_p(C)\) denotes the point spectrum of \(C\).

The following result is an extension of Theorem 4 of [12].

**Theorem 3:** Let \(B : X \to X, C : X \to X\) be linear self-adjoint operators. Suppose that \(\lambda_0\) is an eigenvalue of \(B\), \(e_0 \in \ker (B - \lambda_0I), ||e_0|| = 1\) and that \(\lambda_0 \in \sigma(C)\). Let \(\lambda^*\) be an eigenvalue of \(C\) such that \(\lambda^*\) is nearest to \(\lambda_0\) from the both sides. If \(\lambda_0 \notin \ker (C - \lambda^*I)^1\), then

\[
||\lambda^* - \lambda_0|| \leq ||(C - \lambda_0I) u_0|| \leq \cdots \leq ||(C - \lambda_0I)^2 u_0|| \leq ||B - C||,
\]

where \((u_n)\) is defined by the Kellogg method (1) with \(A = \alpha I - (C - \lambda_0I)^2, u_0 = e_0\) and \(\alpha\) is an arbitrary constant such that \(\alpha > ||(C - \lambda_0I)^2||\).

Proof: The operators \(B, C\) are both bounded by the closed-graph theorem. Put

\[A = \alpha I - (C - \lambda_0I)^2, \quad \alpha > ||(C - \lambda_0I)^2||.\]

Then \(A\) is self-adjoint bounded and positive definite with the greatest isolated point \(\lambda_1\) of \(\sigma(A)\), where \(\lambda_1 = \alpha - (\lambda^* - \lambda_0)^2\). Put \(C_1 = C - \lambda_1I, \lambda^* = \lambda^* - \lambda_0, C_2 = C - \lambda^*I\). We show that \(\ker (C_1 - 2\lambda^*I) \supset \ker C_2\). Suppose that \(u \in \ker C_2\), this condition is equivalent to \(C_1u = \lambda^* u\). Then \(C_1^2u = C_1(\lambda^* u) = \lambda^* u\). Hence \(u \in \ker (C_1^2 - \lambda^*I)\) and we get that \(\ker (A - \lambda_1I) \supset \ker (C - \lambda^*I).\) Therefore \(\ker (A - \lambda_1I)^1 \subset \ker (C - \lambda^*I)^1\) and hence \(u_0 \notin \ker (A - \lambda_1I)^1\). According to Theorem 3 of [11] \(\mu_{n+1} = \sigma(A_{n+1}, u_0) > \lambda_1 = \alpha - (\lambda^* - \lambda_0)^2\).
Hence \( \langle (C - \lambda_0 I) u_n, u_n \rangle \sim (\lambda^* - \lambda_0)^2 \) as \( n \to \infty \). This conclusion implies that

\[
|\lambda^* - \lambda_0| \leq \langle (C - \lambda_0 I)^2 u_n, u_n \rangle^{1/2} = \langle (C - \lambda_0 I) u_{n-1}, u_{n-1} \rangle^{1/2} = \langle (C - \lambda_0 I) u_{n-1} \rangle \leq \cdots \leq \langle (C - \lambda_0 I) e_0 \rangle = \langle (C - B) e_0 \rangle \leq \|C - B\|
\]

because \( e_0 \in \ker (B - \lambda_0 I) \), \( \|e_0\| = 1 \) and \( u_0 = e_0 \), which concludes the proof.

We shall consider the condition concerning the starting approximations of the methods (1)–(3) which occurs in Theorem 1.

**Theorem 4:** Let \( A : X \to X \) be a linear non-negative self-adjoint operator, \( u_0 \in X \). Then the following assertions are equivalent:

(i) \( e_1 u_0 \not= u_0 \) for each \( \lambda < \lambda_1 \);

(ii) \( \|A^n u_0\|^{1/n} \to \lambda_1 \) as \( n \to \infty \);

(iii) \( \langle A^n u_0, u_0 \rangle^{1/n} \to \lambda_1 \) as \( n \to \infty \).

**Proof:** We prove that (i) \( \Rightarrow \) (ii) \( \Rightarrow \) (iii) \( \Rightarrow \) (i).

(i) \( \Rightarrow \) (ii): Assume that \( u_0 \in X \) is such that \( E_1 u_0 \not= u_0 \) for each \( \lambda < \lambda_1 \). Let \( \varepsilon \) be an arbitrary number such that \( 0 < \varepsilon < \lambda_1 \), where \( \lambda_1 \) denotes the lower upper bound of \( \sigma(A) \). Since \( \|A\| = \lambda_1 \), the spectral theorem implies that

\[
\lambda_1 \|u_0\|^2 \geq \|A^n u_0\|^2 \geq \|A^2 u_0\|^2 \geq \langle A^2 u_0, u_0 \rangle = \int_0^{\lambda_1} \lambda^2 d \|E_1 u_0\|^2 \geq \int_{\lambda_1 - \varepsilon}^{\lambda_1} \lambda^2 d \|E_1 u_0\|^2 \geq (\lambda_1 - \varepsilon)^2 \|u_0 - E_1 u_0\|^2.
\]

Hence

\[
\lambda_1 - \varepsilon \leq \liminf \|A^n u_0\|^{1/n} \leq \limsup \|A^n u_0\|^{1/n} \leq \lambda_1.
\]

Therefore \( \lim_{n \to \infty} \|A^n u_0\|^{1/n} = \lambda_1 \).

(ii) \( \Rightarrow \) (iii): Assume that \( \lim_{n \to \infty} \|A^n u_0\|^{1/n} = \lambda_1 \). We have that \( \langle A^n u_0, u_0 \rangle^{1/n} \leq \|A^n u_0\|^{1/n} \times \|u_0\|^{1/n} \). On the other hand, the Reid inequality [17] implies that

\[
\|A^n u_0\|^2 \leq \|A^n\| \langle A^n u_0, u_0 \rangle \leq \|A\|^n \langle A^n u_0, u_0 \rangle \leq \lambda_1^n \langle A^n u_0, u_0 \rangle.
\]

Therefore \( \langle A^n u_0, u_0 \rangle^{1/n} \geq \lambda_1^{-1} \|A^n u_0\|^{2/n} \). Now our conclusion follows at once from the above inequalities.

(iii) \( \Rightarrow \) (i): Assume that \( \lim_{n \to \infty} \langle A^n u_0, u_0 \rangle^{1/n} = \lambda_1 \) for some \( u_0 \in X \), \( u_0 \not= 0 \). We prove that \( E_2 u_0 \not= u_0 \) for each \( \lambda < \lambda_1 \). Suppose conversely, then there exists \( \lambda_0 \) such that \( \lambda_0 < \lambda_1 \) and \( E_2 u_0 = u_0 \). Then we have \( E_2 u_0 = E_1 E_2 u_0 = E_2 u_0 = u_0 \) for each \( \lambda \geq \lambda_0 \). Therefore

\[
0 \leq \langle A^n u_0, u_0 \rangle = \int_0^{\lambda_1} \lambda^n d \|E_2 u_0\|^2 \leq \lambda_0^n \int_0^{\lambda_0} d \|E_2 u_0\|^2 = \lambda_0^n \|E_2 u_0\|^2 = \lambda_0^n \|u_0\|^2.
\]

Hence

\[
0 \leq \langle A^n u_0, u_0 \rangle^{1/n} \leq \lambda_0 \|u_0\|^{2/n}
\]

and

\[
\lim_{n \to \infty} \langle A^n u_0, u_0 \rangle^{1/n} = \lambda_1 \leq \lambda_0 < \lambda_1,
\]

a contradiction. Hence \( E_2 u_0 \not= u_0 \) for each \( \lambda < \lambda_1 \). This finishes the proof.
Remark 1: The assertion (i) \( \Rightarrow \) (ii) of Theorem 4 is stated in [13: chapt. 2]. We gave the proof of this result here only for the sake of completeness. Let us note that the conclusion (ii) of Lemma 1 was observed firstly by Danes [6]. But our proof relies on quite another and much more simpler arguments than in [6].

Proposition 1: Under the assumptions of Theorem 4 suppose that \( \ker (A - \mu I) \neq (0) \) for some \( \mu \in (0, \lambda_1) \) and that \( u_0 \neq 0 \).

Then \( u_0 \in \ker (A - \mu I) \) if and only if there exist constants \( \lambda_0, \lambda^*, \lambda_0 \geq \mu, \lambda^* < \mu \) such that \( E_{\lambda_0}u_0 = u_0 \) and \( E_{\lambda^*}(u_0) = 0 \).

Proof: It is sufficient to prove the converse assertion: \( u_0 \in \ker (A - \mu I) \) if and only if \( E_{\lambda_0}u_0 = u_0 \) for each \( \lambda \geq \mu \) and \( E_{\lambda}u_0 = 0 \) for each \( \lambda < \mu \).

Assume that \( E_{\lambda_0}u_0 = u_0 \) for each \( \lambda \geq \mu \) and \( E_{\lambda}u_0 = 0 \) for all \( \lambda < \mu \). Let \( \epsilon \) be such that \( 0 < \epsilon < \min (\mu, \lambda_1 - \mu) \). Then

\[
\|(A - \mu I)u_0\|^2 = \int_0^{\lambda_1} (\lambda - \mu)^2 d \|E_{\lambda}u_0\|^2 = \lim_{\epsilon \downarrow 0} \int_0^{\lambda_1} (\lambda - \mu)^2 d \|E_{\lambda}u_0\|^2 \\
+ \lim_{\epsilon \downarrow 0} \int_0^{\lambda_1} (\lambda - \mu)^2 d \|E_{\lambda}u_0\|^2 = 0.
\]

Hence \( u_0 \in \ker (A - \mu I) \). Assume now that \( u_0 \in \ker (A - \mu I) \). Then

\[
0 = \|(A - \mu I)u_0\|^2 = \int_0^{\lambda_1} (\lambda - \mu)^2 d \|E_{\lambda}u_0\|^2 \\
+ \int_{\mu - \epsilon}^{\mu + \epsilon} (\lambda - \mu)^2 d \|E_{\lambda}u_0\|^2 = \int_{\mu - \epsilon}^{\mu + \epsilon} (\lambda - \mu)^2 d \|E_{\lambda}u_0\|^2.
\]

From this equality it follows that \( E_{\lambda}u_0 \) is constant on the intervals \( (0, \mu - \epsilon], [\mu + \epsilon, \lambda_1) \) for each \( \epsilon > 0 \). The properties of the spectral family \( \{E_{\lambda}\} \) imply that \( E_{\lambda}u_0 = \lim_{\lambda \downarrow 0} E_{\lambda} \times u_0 = 0 \) for each \( \lambda \in (0, \mu) \) and \( E_{\lambda}u_0 = \lim_{\lambda \uparrow \lambda_1} E_{\lambda}u_0 = u_0 \) for each \( \lambda \in [\mu, \lambda_1] \) as desired \( \blacksquare \)

REFERENCES


Manuskripteingang: 15. 04. 1982

VERFASSER:

Prof. Dr. Josef Kolomy
Matematický Ustav University Karlovy
18600 Praha 8 - Karlin, Sokolovska 83