Изв. АН Эстонии. Физ. Матем., 1989, 38, № 3, 268—276

УДК 517.98 https://doi.org/10.3176/phys.math.1989.3.05

Анна МИНЦ

О КРИТЕРИИ ОСТАНОВА МЕТОДА ОПЕРАТОРНЫХ ИТЕРАЦИЙ

(Представил Г. Вайникко)

В работе указывается критерий останова метода операторных итераций и приводится алгоритм применения метода.

Постановка задачи

Квадратная матрица $A=(a_{ij})^N_{i,j=1}$, где $A=A^*>0$, $\|A\|\leqslant 1$, обращается методом операторных итераций

$$B_k=B_{k-1}(2I-AB_{k-1}), \qquad B_0=g(A). \tag{1}$$

Здесь $g(\lambda)$ непрерывна на $[p, \|A\|]$ (p — минимальное собственное число A), $p>0$, B_k — очередное приближение к A^{-1},

$$0<g(\lambda)<2/\lambda. \tag{1'}$$

Предполагается, что на каждом шаге итераций (1) делается ошибка C_k, так что

$$\tilde{B}_k=\tilde{B}_{k-1}(2I-A\tilde{B}_{k-1})+C_k, \qquad \tilde{B}_0=B_0+C_0.$$

Здесь $\tilde{B}_k$ — приближенное значение B_k,

$$\|C_k\|\leqslant\varepsilon \qquad (k=0,\ 1,\ \ldots).$$

В дальнейшем указывается, какое число итераций (в зависимости от ε, $g(\lambda)$ и матрицы A) достаточно сделать для достижения разумной точности, а также даются рекомендации о наилучшем выборе $g(\lambda)$.

Предварительные обозначения и результаты

Обозначим $a_1\leqslant a_2\leqslant\ldots\leqslant a_N$ — собственные числа A с учетом кратности

$$q=\max_{\lambda\in\sigma(A)}|1-\lambda g(\lambda)|,$$
$$G_k(\lambda)=(1/\lambda)\cdot(1-\lambda g(\lambda))^{2^k}(1-(1-\lambda g(\lambda))^{2^k}),$$
$$F(\lambda)=|1-\lambda g(\lambda)|,$$
$$F_k(\lambda)=(1-\lambda g(\lambda))^{2^k},$$
$$E_k=\tilde{B}_k-B_k, \quad \|E_k\|=\varepsilon_k.$$

В [1] было показано, что если $\varepsilon<\sqrt{q}-q$, то последовательность ε_k мажорируется сходящейся последовательностью δ_k, причем

$$\delta_{k+1}=2q^{2^k}\delta_k+\delta_k^2+\varepsilon, \qquad \delta_0=\varepsilon,$$
$$\lim_{k\to\infty}\delta_k=(1-\sqrt{1-4\varepsilon})/2. \tag{2}$$

Заметим, что при достаточно малых ε эта величина эквивалентна ε.

Рассмотрим вопрос о поведении нормы $B_{k+1}-B_k$.

Поскольку $A=A^*$, то ее можно представить в виде $A=$

$=U\,\mathrm{diag}(a_1,\ \dots,\ a_N)\,U^{-1}$, где U — унитарная матрица, а так как (см. [2]) $B_k=g_k(A)$, то и B_k можно представить в виде

$$B_k=U\,\mathrm{diag}\,(b_1^{(k)},\ \dots,\ b_N^{(k)})\,U^{-1},$$

где $b_1^{(k)},\ \dots,\ b_N^{(k)}$ — собственные числа B_k с учетом кратности, и, как видно из (1), $b_i^{(k)}=b_i^{(k-1)}(2-a_ib_i^{(k-1)})$ $(i=1,\ \dots,\ N;$ $k=1,\ 2,\ \dots)$. Далее, $B_{k+1}-B_k=U\,\mathrm{diag}\,(\beta_1^{(k)},\ \dots,\ \beta_N^{(k)})\,U^{-1}$, где $\beta_1^{(k)},\ \dots,\ \beta_N^{(k)}$ — собственные числа $B_{k+1}-B_k$ с учетом кратности, причем $\beta_i^{(k)}=b_i^{(k)}(1-a_ig(a_i))^{2^k}$. Известно (см. [2]), что $I-AB_k=(I-AB_0)^{2^k}$, и таким образом, $b_i^{(k)}=(1/a_i)\cdot(1-(1-a_ib_i^{(0)})^{2^k})$, а поскольку $b_i^{(0)}=g(a_i)$, имеем

$$\beta_i^{(k)}=(1/a_i)\cdot(1-a_ig(a_i))^{2^k}(1-(1-a_ig(a_i))^{2^k}).$$

Отсюда следует, что

$$\|B_{k+1}-B_k\|=\max_{\lambda\in\sigma(A)}G_k(\lambda). \tag{3}$$

Поведение $\tilde{B}_{k+1}-\tilde{B}_k$

Поскольку мы предполагаем, что вычисления проводятся неточно, нас прежде всего должен интересовать вопрос о поведении $\tilde{B}_{k+1}-\tilde{B}_k$. Используя введенные ранее обозначения, запишем

$$\tilde{B}_{k+1}-\tilde{B}_k=B_{k+1}-B_k-B_kAE_k+E_k(I-AB_k)-E_kAE_k+C_{k+1}.$$

Отсюда

$$\|\tilde{B}_{k+1}-\tilde{B}_k\|\leqslant\|B_{k+1}-B_k\|+\|E_k\|\cdot(1+q^{2^k})+\varepsilon+\|E_k\|^2. \tag{4}$$

Докажем следующую вспомогательную лемму:

Лемма 1. *Пусть*

$$\gamma_k=\varepsilon^{1/2^k}\left(1-\frac{\varepsilon^{1/2}}{2^{k-1}}-\frac{\varepsilon^{3/4}}{2^{k-2}}-\dots-\frac{\varepsilon^{1-1/2^{k-1}}}{2}\right)-\underbrace{\sqrt{-\varepsilon+\sqrt{-\varepsilon+\dots+\sqrt{\varepsilon}}}}_{k}.$$

Тогда

$$\gamma_k=\varepsilon^{1/2^k}\left(\frac{\varepsilon\,(2^{k-1}-1)}{2^{2k-1}}+\frac{\varepsilon^{5/4}(2^{k-2}-1)}{2^{2k-3}}\right)+o\,(\varepsilon^{5/4}), \tag{5}$$

$$k=2,\ 3,\ \dots$$

Доказательство проведем по индукции, используя известное разложение

$$\sqrt{1-\alpha}=1-\frac{\alpha}{2}-\frac{\alpha^2}{8}-\frac{\alpha^3}{16}-\dots$$

и отбрасывая степени ε, большие 5/4. Используем знак $\simeq$ для обозначения равенства с точностью до $o(\varepsilon^{5/4})$. Имеем:

$$\sqrt{-\varepsilon+\sqrt{\varepsilon}}=\sqrt[4]{\varepsilon}\sqrt{1-\varepsilon^{1/2}}\simeq\sqrt[4]{\varepsilon}\left(1-\frac{\varepsilon^{1/2}}{2}-\frac{\varepsilon}{8}\right)=\sqrt[4]{\varepsilon}\left(1-\frac{\varepsilon^{1/2}}{2}\right)-\frac{\varepsilon^{5/4}}{8},$$

т. е. $\gamma_2\simeq\varepsilon^{1/4}\cdot\varepsilon/8$;

$$\sqrt{-\varepsilon+\sqrt{-\varepsilon+\sqrt{\varepsilon}}}=\varepsilon^{1/8}\cdot\sqrt{1-\varepsilon^{1/2}/2-\varepsilon^{3/4}-\varepsilon/8}\simeq\varepsilon^{1/8}\left(1-\frac{\varepsilon^{1/2}}{4}-\frac{\varepsilon^{3/4}}{2}-\right.$$

$$\left.-\frac{\varepsilon}{16}-\frac{1}{8}\left(\frac{\varepsilon^{1/2}}{2}+\varepsilon^{3/4}+\varepsilon/8\right)^2\right)\simeq\varepsilon^{1/8}\left(1-\frac{\varepsilon^{1/2}}{4}-\frac{\varepsilon^{3/4}}{2}\right)-\varepsilon^{1/8}\left(\frac{3\varepsilon}{32}+\frac{\varepsilon^{5/4}}{8}\right),$$

т. е.

$$\gamma_3 \asymp \varepsilon^{1/8}\cdot\left(\frac{3\varepsilon}{32}+\frac{\varepsilon^{5/4}}{8}\right).$$

Предположим, что (5) выполнено для некоторого k. Тогда

$$\underbrace{\sqrt{-\varepsilon+\sqrt{-\varepsilon+\ldots+\sqrt{\varepsilon}}}}_{k+1}=\sqrt{-\varepsilon+\varepsilon^{1/2^k}\left(1-\frac{\varepsilon^{1/2}}{2^{k-1}}-\ldots-\frac{\varepsilon^{1-1/2^{k-1}}}{2}-\gamma_k\right.}\asymp$$

$$\asymp \varepsilon^{1/2^{k+1}}\cdot\sqrt{1-\frac{\varepsilon^{1/2}}{2^{k-1}}-\ldots-\varepsilon^{1-1/2^k}-\frac{\varepsilon(2^{k-1}-1)}{2^{2k-1}}-\frac{\varepsilon^{5/4}(2^{k-2}-1)}{2^{2k-3}}}\asymp$$

$$\asymp \varepsilon^{1/2^{k+1}}\left(1-\frac{\varepsilon^{1/2}}{2^k}-\ldots-\frac{\varepsilon^{1-1/2^k}}{2}-\frac{\varepsilon(2^{k-1}-1)}{2^{2k}}-\frac{\varepsilon^{5/4}(2^{k-2}-1)}{2^{2k-2}}\right.$$

$$\left.-\frac{1}{8}\left(\frac{\varepsilon^{1/2}}{2^{k-1}}+\ldots+\varepsilon^{1-1/2^k}+\frac{\varepsilon(2^{k-1}-1)}{2^{2k-1}}+\frac{\varepsilon^{5/4}(2^{k-2}-1)}{2^{2k-3}}\right)^2\right)\asymp$$

$$\asymp \varepsilon^{1/2^{k+1}}\left(1-\frac{\varepsilon^{1/2}}{2^k}-\ldots-\frac{\varepsilon^{1-1/2^k}}{2}\right)-\varepsilon^{1/2^{k+1}}\left(\frac{\varepsilon(2^k-1)}{2^{2k+1}}+\frac{\varepsilon^{5/4}(2^{k-1}-1)}{2^{2k-1}}\right),$$

т. е.

$$\gamma_{k+1}\asymp\varepsilon^{1/2^{k+1}}\left(\frac{\varepsilon(2^k-1)}{2^{2k+1}}+\frac{\varepsilon^{5/4}(2^{k-1}-1)}{2^{2k-1}}\right).$$

Следовательно, (5) выполнено и для $k+1$, и лемма 1 доказана.

Из (5) следует, что

$$\gamma_k\leqslant\varepsilon^{1/2^k}\cdot\frac{\varepsilon+2\varepsilon^{5/4}}{2^k}+o(\varepsilon^{5/4}),$$

и можно показать, что при $k\geqslant\dfrac{(1+\varepsilon)\log_2 1/\varepsilon}{4}$ имеем $\gamma_k=o(\varepsilon^{5/4})$.

Докажем теперь следующую лемму:

Лемма 2. *Пусть* $\varepsilon<(1-q)/4$; $k_1=\min\left\{k:\varepsilon^{1/2^k}\geqslant\dfrac{q+\varepsilon}{1-3\varepsilon}\right\}$,

$$k_2=\left[\frac{1}{2}\log_2(1/\varepsilon)\right]+3,$$

$$k\geqslant\max(11,\,k_1,\,k_2).$$

Тогда

$$\varepsilon_k\leqslant 2\varepsilon+o(\varepsilon^{5/4}). \tag{6}$$

Доказательство. Рассмотрим последовательность δ_k, задаваемую формулами (2). Как уже отмечалось, для любого k выполнено неравенство $\varepsilon_k\leqslant\delta_k$. Поэтому, если удастся доказать, что в условиях леммы 2 будет выполнено неравенство $\delta_k\leqslant 2\varepsilon+o(\varepsilon^{5/4})$, то (6) получится автоматически. Используя (2), получим цепочку равносильных неравенств

$$\delta_k\leqslant 2\varepsilon\Leftrightarrow\delta_{k-1}\leqslant-q^{2^{k-1}}+\sqrt{\varepsilon+q^{2^k}}\Leftrightarrow$$

$$\Leftrightarrow\delta_{k-2}\leqslant-q^{2^{k-2}}+\sqrt{-\varepsilon+\sqrt{\varepsilon+q^{2^k}}}\Leftrightarrow$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\Leftrightarrow\delta_0=\varepsilon\leqslant-q+\sqrt{-\varepsilon+\sqrt{-\varepsilon+\ldots+\sqrt{\varepsilon+q^{2^k}}}}. \tag{7}$$

Вместо (7) рассмотрим неравенство

$$\varepsilon \leqslant -q+\sqrt{-\varepsilon+\sqrt{-\varepsilon+\ldots+\sqrt{\varepsilon}}}, \tag{7'}$$

очевидно, более сильное, чем (7).

Используем следующие приближенные формулы:

$$\sqrt{-\varepsilon+\sqrt{\varepsilon}}=\sqrt[4]{\varepsilon}\,(1-\sqrt{\varepsilon}/2)+o(\varepsilon^{5/4}),$$

. . . .

$$\sqrt{-\varepsilon+\sqrt{-\varepsilon+\ldots+\sqrt{\varepsilon}}}=\varepsilon^{1/2^k}\left(1-\frac{\varepsilon^{1/2}}{2^{k-1}}-\ldots-\frac{\varepsilon^{1-1/2^{k-1}}}{2}\right)+o(\varepsilon^{5/4}).$$

Как видно из леммы 1, неравенство (7′) можно при $k \geqslant k_2$ с точностью до $o(\varepsilon^{5/4})$ заменить неравенством

$$\varepsilon \leqslant -q+\varepsilon^{1/2^k}\left(1-\frac{\varepsilon^{1/2}}{2^{k-1}}-\ldots-\frac{\varepsilon^{1-1/2^{k-1}}}{2}\right). \tag{8}$$

Обозначим

$$S_k=\frac{\varepsilon^{1/2}}{2^{k-1}}+\ldots+\frac{\varepsilon^{1-1/2^{k-1}}}{2},$$

т. е.

$$S_k=\varepsilon\sum_{i=1}^{k-1}\frac{\varepsilon^{-1/2^i}}{2^{k-i}}=\varepsilon\sum_{\substack{m+n=k\\ m,n\geqslant 1}}2^{-m+\log_2\left(\frac{1}{\varepsilon}\right)/2^n}=$$

$$=\varepsilon\sum_{\substack{m-n=k\\ m,n\geqslant 1}}2^{-m+L/2^n}=\varepsilon\,(2^{-1+L/2^{k-1}}+2^{-2+L/2^{k-2}}+\ldots+2^{-(k-1)+L/2}),$$

где $L=\log_2 1/\varepsilon$.

Рассмотрим функцию $h(x)=-x+L/2^{-(k-x)}$. Она имеет единственный минимум при $x=k-\log_2(L\ln 2)$;

$$h(1)=-1+L/2^{k-1},\quad h(k-1)=-(k-1)+L/2.$$

Потребуем, чтобы $h(i)$ была отрицательна для любого $1 \leqslant i \leqslant k-1$. Это выполнено, если

$$\left.\begin{array}{r}-1+L/2^{k-1}<0\\ -(k-1)+L/2<0\end{array}\right\},\ \text{или}\ \left\{\begin{array}{l}L<2^{k-1}\\ L<2(k-1).\end{array}\right.$$

При $k>L/2+2$ оба эти неравенства выполнены, и

$$h(i)<0\quad (i=1,\ \ldots,\ k-1).$$

Так как $L<2(k-1)$, имеем

$$S_k \leqslant \varepsilon\cdot(2^{-1+(k-1)/2^{k-2}}+\ldots+2^{-(k-1)+(k-1)}).$$

Рассмотрим теперь функцию $t(x)=-x+\dfrac{k-1}{2^{k-x-1}}$ и выясним, когда

$$t(i+1)-t(i)<-1/2. \tag{9}$$

Так как $t(i+1)-t(i)=-1+(k-1)\left(\dfrac{1}{2^{k-i-2}}-\dfrac{1}{2^{k-i-1}}\right)=-1+\dfrac{k-1}{2^{k-i-1}}$, то (9) выполнено, если $(k-1)/2^{k-i-1}<1/2$, т. е. если

$$i<-1+(k-1)-\log_2(k-1). \tag{10}$$

При $k>10$ имеем

$$[k/2]<-1+(k-1)-\log_2(k-1),$$

и, таким образом, часть суммы

$$2^{-1+(k-1)/2^{k-2}}+\ldots+2^{-i+(k-1)/2^{k-i-1}}+\ldots+2^{-(k-1)+(k-1)} \tag{11}$$

при $i\leqslant[k/2]$ можно оценить суммой вида

$$\sum_{i=1}^{[k/2]}{}' 2^{-1+(k-1)/2^{k-2}-i/2},$$

представляющей собой геометрическую прогрессию со знаменателем $1/\sqrt{2}$.

И, наконец, рассмотрим функцию $r(x)=2^{-x+(k-1)/2^{k-x-1}}$; $r(k-1)=1$. Зададимся вопросом, при каких i имеем $r(x)\leqslant r(k-2)$. Это неравенство выполняется при i, удовлетворяющих неравенству $2^{-i+(k-1)/2^{k-i-1}}\leqslant$ $\leqslant 2^{-k/2+3/2}$.

При $k>10$ этому неравенству удовлетворяют

$$[k/2]+1\leqslant i\leqslant k-2,$$

и, таким образом, часть суммы (11) при таких i можно оценить суммой вида

$$\sum_{i=[k/2]+1}^{k-2} 2^{-k/2+3/2}+1=(k-2-[k/2])\cdot 2^{-k/2+3/2}+1,$$

и, учитывая оценку, полученную выше, оценим сумму (11) суммой вида

$$\sum_{i=1}^{[k/2]}{}' 2^{-1+(k-1)/2^{k-2}-i/2}+\sum_{[k/2]+1}^{k-2} 2^{-k/2+3/2}+1\leqslant$$

$$\leqslant\frac{2^{-3/2+(k-1)/2^{k-2}}\left(1-(1/\sqrt{2})^{[k/2]-1}\right)}{1-1/\sqrt{2}}+([k/2]-1)\cdot 2^{-k/2+3/2}+1.$$

Итак,

$$S_k\leqslant\varepsilon\cdot\left(\frac{2^{-3/2+(k-1)/2^{k-2}}\left(1-(1/\sqrt{2})^{[k/2]-1}\right)}{1-1/\sqrt{2}}+([k/2]-1)\cdot 2^{-k/2+3/2}+1\right)$$

Покажем, что при $k\geqslant 11$ выполнено неравенство

$$2^{(k-1)/2^{k-2}}\left(1-(1/\sqrt{2})^{[k/2]-1}\right)\leqslant 1, \tag{11'}$$

т. е.

$$2^{-(k-1)/2^{k-2}}+2^{1/2}/2^{[k/2]/2}\geqslant 1.$$

Используя тот факт, что $[k/2]/2\leqslant k/4$, рассмотрим более сильное неравенство

$$2^{-(k-1)/2^{k-2}}+2^{1/2}/2^{k/4}\geqslant 1.$$

Заметим, что при $k=11$ это неравенство выполнено.

Исследуем теперь поведение функции

$$\varkappa(x)=2^{-(x-1)/2^{x-2}}+2^{1/2}/2^{x/4}.$$

Производная этой функции

$$\varkappa'(x)\leqslant(4\ln 2/2^x)\left((x-1)\ln 2-1-2^{x/4}/2^{9/2}\right),$$

и видно, что при $x\geqslant 11$ имеем $\varkappa'(x)<0$. Таким образом, при $x\geqslant 11$

$\varkappa(x)$ убывает, $\varkappa(11)>1$, а $\lim\limits_{x\to\infty}\varkappa(x)=1$. Из этого следует, что при $x\geqslant 11$ имеем $\varkappa(x)\geqslant 1$, откуда и получаем (11′).

В свою очередь, из (11′) следует, что

$$(2^{-3/2+(k-1)/2^{k-2}}(1-(1/\sqrt{2})^{[k/2]-1})/(1-1/\sqrt{2})\leqslant 2^{-3/2}/(1-1/\sqrt{2}).$$

С другой стороны, при $k>10$ имеем

$$([k/2]-1)\cdot 2^{-k/2+3/2}\leqslant 2-2^{-3/2}/(1-1/\sqrt{2}).$$

Итак, при $k>10: S_k\leqslant 3\varepsilon$, и соответственно, $\varepsilon^{1/2^k}(1-S_k)\geqslant(1-3\varepsilon)\varepsilon^{1/2^k}$, а поскольку, как было показано ранее, (7′) можно с точностью до $o(\varepsilon^{5/4})$ заменить на (8), получаем, что (7′) выполнено, если

$$\varepsilon\leqslant -q+(1-3\varepsilon)\cdot\varepsilon^{1/2^k}$$

или

$$\varepsilon^{1/2^k}\geqslant(q+\varepsilon)/(1-3\varepsilon).$$

Так как по условию леммы $\varepsilon<(1-q)/4$, то $(q+\varepsilon)/(1-3\varepsilon)<1$, и существует $k_1=\min\{k:\varepsilon^{1/2^k}\geqslant(q+\varepsilon)/(1-3\varepsilon)\}$.

Таким образом, получаем, что в условиях леммы 2 выполнено (7) с точностью до $o(\varepsilon^{5/4})$, и следовательно, лемма доказана.

Пусть теперь

$$k_3=\min\{k:(q^{2^k}/p)(1-q^{2^k})\leqslant\varepsilon\}.$$

В обозначениях леммы 2

$$k_1=\min\{k:\varepsilon^{1/2^k}\geqslant(q+\varepsilon)/(1-3\varepsilon)\},$$

$$k_2=\left[\frac{1}{2}\log_2(1/\varepsilon)\right]+3.$$

Сформулируем главный результат данного пункта.

Теорема. *Пусть $g(\lambda)$ такая, что*

$$\max_{\lambda\in\sigma(A)}F(\lambda)=F(p), \tag{12}$$

$$\max_{\lambda\in\sigma(A)}G_k(\lambda)=G_k(p)\qquad(k\geqslant k_*,\ k_*=\max(11,k_1,k_2,k_3)), \tag{13}$$

$$\varepsilon<(1-q)/4.$$

Тогда

$$\|\tilde{B}_{k+1}-\tilde{B}_k\|\leqslant 4\varepsilon+o(\varepsilon^{5/4}).$$

Доказательство. Как было показано (см. (3)),

$$\|B_{k+1}-B_k\|=\max_{\lambda\in\sigma(A)}G_k(\lambda),$$

и из (13) получаем

$$\|B_{k+1}-B_k\|=1/pq^{2^k}(1-q^{2^k}).$$

Если $k\geqslant k_3$, то $\|B_{k+1}-B_k\|\leqslant\varepsilon$.

Далее, из леммы 2 следует, что в условиях теоремы

$$\|E_k\|\leqslant 2\varepsilon+o(\varepsilon^{5/4}).$$

Наконец, используя (4), получим

$$\|\tilde{B}_{k+1}-\tilde{B}_k\|\leqslant\varepsilon+2\varepsilon(1+q^{2^k})+\varepsilon+o(\varepsilon^{5/4}),$$

но при $k\geqslant k_3$ имеем $\varepsilon q^{2^k}=o(\varepsilon^{5/4})$, и следовательно,

$$\|\tilde{B}_{k+1}-\tilde{B}_k\|\leqslant 4\varepsilon+o(\varepsilon^{5/4}).$$

Теорема доказана.

Таким образом, если известны ε и p и надлежащим образом подобрана $g(\lambda)$, то еще до начала вычислений можно указать число итераций, достаточное для того, чтобы $\|\bar{B}_{k+1}-\bar{B}_k\|$ имела порядок погрешности округления ε. Следует заметить, что этого числа итераций достаточно и для того, чтобы $\|\bar{B}_k-B_k\|$ также имела порядок ε.

В следующем пункте будет исследован вопрос о наилучшем выборе $g(\lambda)$.

Выбор $g(\lambda)$

Функция $g(\lambda)$ должна удовлетворять условиям (1′) и (12)—(13). Желательно также, чтобы q было как можно меньше. Это связано с тем, что норма невязки $\|I-AB_k\|=q^{2^k}$, и следовательно, чем меньше q, тем быстрее сходимость. Кроме того, поскольку придется вычислять $g(A)$, необходимо, чтобы это вычисление было как можно проще.

На первый взгляд, очень удобной является $g_1(\lambda)=1/(1+\lambda)$. Она удовлетворяет (1′) и (12)—(13) для любых p и $\|A\|$, а вычисление $g_1(A)$ теоретически не представляет трудности, так как матрица $I+A$ хорошо обусловлена. Однако, как было показано в [3], при реальных вычислениях для больших размерностей A на подсчет $g_1(A)$ уходит примерно столько же времени, сколько на выполнение всех итераций, и, кроме того, погрешность C_0 получается весьма большой. Поэтому лучше использовать в качестве $g(\lambda)$ полином, а чтобы минимизировать C_0 и время вычисления $g(A)$, полином этот будем считать полиномом первой степени.

Заметим, что для того, чтобы q было минимально, следует выбирать убывающую функцию $g(\lambda)$, в противном случае при малых p величина $|1-pg(p)|$ будет очень близка к 1. Поэтому рассмотрим функцию вида

$$g_2(\lambda)=c_0-c_1\lambda,$$

где c_0, $c_1>0$.
Какими же должны быть c_0, c_1, чтобы выполнялись условия (1′) и (12)—(13)?

Обеспечим сначала выполнение условия (12). Для выполнения этого условия достаточно, чтобы $(F(\lambda))'$ было неположительно на $[p;\ \|A\|]$, т. е. $-c_0+2\lambda c_1\leqslant 0$ на $[p;\ \|A\|]$, что гарантируется выполнением условия $-c_0+2\|A\|c_1\leqslant 0$, или

$$c_1\leqslant c_0/(2\|A\|). \tag{14}$$

Потребуем дополнительно, чтобы $1-\lambda g(\lambda)$ было неотрицательно. Это условие равносильно условию $1-c_0\lambda+c_1\lambda^2\geqslant 0$, или, так как $1-\lambda g(\lambda)$ не возрастает, $1-c_0\|A\|+c_1\|A\|^2\geqslant 0$, т. е.

$$c_1\geqslant(\|A\|c_0-1)/\|A\|^2. \tag{15}$$

Для существования c_1, удовлетворяющего (14) и (15), необходимо, чтобы выполнялось $(\|A\|c_0-1)/\|A\|^2\leqslant c_0/(2\|A\|)$, откуда следует

$$c_0\leqslant 2/\|A\|. \tag{16}$$

Итак, при c_0, c_1, удовлетворяющих (14)—(16), условие (12) выполнено. Рассмотрим теперь условие (13). Продифференцируем $G_k(\lambda)$

$$(G_k(\lambda))'=\left(\frac{F_k(\lambda)(1-F_k(\lambda)}{\lambda}\right)=$$

$$=\frac{\lambda(F'_k(\lambda)-2F_k(\lambda)F'_k(\lambda)-F_k(\lambda)+F^2_k(\lambda))}{\lambda^2}.$$

Если удастся добиться, чтобы $G'_k(\lambda)$ была неположительна, то (13) будет выполнено. Воспользовавшись тем фактом, что

$$F'_k(\lambda) = -2^k F_k(\lambda) \cdot (\lambda g(\lambda))' / (1 - \lambda g(\lambda)),$$

получим

$$G'_k(\lambda) \leqslant 0 \Leftrightarrow F_k(\lambda) - 1 - \lambda 2 F'_k(\lambda) - \frac{\lambda(\lambda g(\lambda))' 2^k}{1 - \lambda g(\lambda)} \leqslant 0.$$

Заметим, что $F_k(\lambda) - 1$ всегда неположительно. Потребуем, чтобы

$$2F'_k(\lambda) + \frac{(\lambda g(\lambda))' \cdot 2^k}{1 - \lambda g(\lambda)} \geqslant 0,$$

тогда получим $G'_k(\lambda) \leqslant 0$. Поскольку $(1 - \lambda g(\lambda))' \leqslant 0$, то $(\lambda g(\lambda))' \geqslant 0$, и

$$2F'_k(\lambda) + \frac{(\lambda g(\lambda))' 2^k}{1 - \lambda g(\lambda)} \geqslant 0 \Leftrightarrow -2^k (1 - \lambda g(\lambda))^{2^k - 1} (\lambda g(\lambda))' +$$

$$+ 2^{k-1} \frac{(\lambda g(\lambda))'}{1 - \lambda g(\lambda)} \geqslant 0 \Leftrightarrow 1/(1 - \lambda g(\lambda)) - 2(1 - \lambda g(\lambda))^{2^{k-1}} \geqslant 0. \quad (17)$$

Если $q^{2^k} \leqslant 1/2$, то (17) выполнено, а это последнее условие, очевидно, выполнено при $k \geqslant k_3$.

Итак, мы молучили, что для выполнения условия (13), казавшегося наиболее сложным, не требуется никаких новых ограничений на c_0, c_1.

Рассмотрим теперь условие (1′):

$$0 < g(\lambda) < 2/\lambda.$$

То, что $g(\lambda) = c_0 - c_1(\lambda) > 0$, гарантируется условием (14), так как, если $c_1 \leqslant c_0/(2\|A\|)$, то $c_1 < c_0/\|A\|$, т. е. $c_0 - c_1\|A\| > 0$. Условие $\lambda g(\lambda) < 2$ гарантируется условием $1 - \lambda g(\lambda) > 0$, и таким образом, (1′) тоже не требует новых ограничений на c_0, c_1.

Какими же, наконец, должны быть c_0, c_1, чтобы q было минимально? В наших предположениях

$$q = 1 - c_0 p + c_1 p^2,$$

и следовательно, для минимизации q необходимо взять наибольшее возможное c_0 и наименьшее возможное c_1. Как видно из (14)—(16),

$$\begin{aligned} c_0 &= 2/\|A\|, \\ c_1 &= 1/\|A\|^2. \end{aligned} \quad (18)$$

Заметим, что при таком выборе c_0, c_1 и $\|A\| = 1$ получаем, что q, соответствующее

$$g_2(\lambda) = 1 - \lambda/2,$$

меньше q, соответствующего

$$g_1(\lambda) = 1/(1+\lambda).$$

Алгоритм применения метода и некоторые численные результаты

Теперь можно сформулировать алгоритм применения метода операторных итераций.

I. Определить $\|A\|$ и найти c_0, c_1 из (18).

II. Определить $q = \|I - c_0 A + c_1 A^2\|$ и, используя тот факт, что $q = 1 - c_0 p + c_1 p^2$, найти p.

III. Зная p, ε, q, найти k_1, k_2, k_3 и определить $k_* = \max(11, k_1, k_2, k_3)$.

IV. Применять (1) для $k = 0, 1, \dots, k_*$.

В таблице приведены значения k_* для различных p и ε в предположении, что $\|A\| = 1$, $g(\lambda) = 2 - \lambda$ (q в этом случае равно $1 - 2p + p^2 \approx 1 - 2p$).

p	ε							
	10^{-4}	10^{-5}	10^{-6}	10^{-7}	10^{-8}	10^{-9}	10^{-10}	10^{-11}
10^{-3}	14	14	15	15	16	17	19	21
10^{-4}	—	18	18	18	19	19	19	21
10^{-5}	—	—	22	22	22	22	22	22
10^{-6}	—	—	—	25	25	26	26	26
10^{-7}	—	—	—	—	29	29	29	29
10^{-8}	—	—	—	—	—	32	32	32
10^{-9}	—	—	—	—	—	—	36	36
10^{-10}	—	—	—	—	—	—	—	39

Примечание. Прочерки в графах таблицы показывают, что при данных p нельзя рассматривать такие ε, так как $\varepsilon > (1-q)/4$.

Автор выражает глубокую благодарность Г. Вайникко за помощь в работе и А. Вольбергу за ценное обсуждение.

ЛИТЕРАТУРА

1. *Минц А.* // Тез. докл. конф. «Теоретические и прикладные вопросы математики II». Тарту, ТГУ, 1985, 95—97.
2. *Вайникко Г.* Методы решения линейных некорректно поставленных задач в гильбертовых пространствах. Тарту, ТГУ, 1982.
3. *Минц А.* // Уч. зап. Тартуск. ун-та, 1987, вып. 762, 40—46.

Институт кибернетики
Академии наук Эстонской ССР

Поступила в редакцию
5/V 1988

Anna MINTS

ÜHEST KRITEERIUMIST OPERAATOR-ITERATSIOONIMEETODI PEATAMISEKS

On esitatud kriteerium operaator-iteratsioonimeetodi peatamiseks sümmeetrilise maatriksi pööramisega, võttes arvesse ümardamisvigade võimalikku kuhjumist, ja antud mõningad soovitused alglähendi valikuks.

Anna MINTS

ON THE CRITERION FOR STOPPING THE OPERATOR ITERATIONS METHOD

In this work a criterion for stopping the operator iterations method is obtained, and the corresponding algorithm for the application of this method is given.