ROCZNIKI POLSKIEGO TOWARZYSTWA MATEMATYCZNEGO
SERIA III: MATEMATYKA STOSOWANA IV (1975)

G. HOBOT i M. STEFAŃCZYK (Lublin)

Praktyka oszacowań błędów wartości własnych operatorów różniczkowych znajdowanych metodą różnicową sprzężoną z QR-algorytmem

1. Postawienie zadania. Będziemy rozpatrywać operator liniowy, samosprzężony, rzędu 2-go, typu eliptycznego postaci:

$$Lu = \sum_{i=1}^{m} \frac{\partial}{\partial x_i}\left(a_i \frac{\partial u}{\partial x_i}\right) - au, \tag{1.1}$$

gdzie $a_i(x)$, $a(x)$ – dane funkcje, $a_i(x) > 0$, $a(x) > 0$, $x = (x_1, \ldots, x_m)$, $u = u(x_1, \ldots, x_m)$.

Dla operatora L stawiamy następujące zadanie: szukamy liczb λ takich, że istnieje niezerowe rozwiązanie równania:

$$Lu(x) + \lambda\rho(x)\, u(x) = 0, \qquad x \in D, \tag{1.2}$$

$$lu(x) = 0, \qquad x \in \partial D, \tag{1.3}$$

gdzie D – m-wymiarowy obszar a ∂D oznacza brzeg tego obszaru, $\rho(x) > 0$ – zadana funkcja, l oznacza operator warunków brzegowych jednego z trzech typów

$$l_1 u = u, \qquad l_2 u = \frac{\partial u}{\partial \vec{n}}, \qquad l_3 u = \frac{\partial u}{\partial \vec{n}} - ku, \tag{1.4}$$

gdzie $\vec{n}$ – wektor normalny do ∂D skierowany do wnętrza obszaru, $k(x) > 0$, z góry zadana funkcja.

Tak postawiony problem nazywamy problemem własnym dla operatora L, a λ – wartością własną zadania (1.2)–(1.3).

Zadania na wartości własne dla operatorów typu (1.1) można rozwiązywać dokładnie tylko w nielicznych przypadkach. Z metod przybliżonych często stosowana jest metoda różnicowa ze względu na swoją prostotę.

2. Zadanie różnicowe sprzężone z (1.2)–(1.3). Wprowadzamy w obszarze D siatkę D_h o brzegu ∂D_h. Zastępujemy operatory L, l występujące w (1.2)–(1.3) odpowiednio operatorami różnicowymi L_h, l_h i przechodzimy do następującego zadania różnicowego:

$$L_h u^h(P) + \lambda^h \rho^h(P)\, u^h(P) = 0, \qquad P \in D_h, \tag{2.1}$$

(2.2) $$l_h u_h(P) = 0, \qquad P \in \partial D_h.$$

Operatory L_h i l_h otrzymamy z L i l zastępując np. wyrażenia różniczkowe odpowiednimi wyrażeniami różnicowymi.

Niech D_h zawiera N_h węzłów. Węzły te oznaczmy kolejno przez $P_1, P_2, \ldots, P_{N_h}$. Wtedy dowolna funkcja siatkowa $U(P)$ zdefiniowana dla każdego $P \in D_h$ może być przedstawiona jako wektor kolumnowy o wymiarze N_h, tzn.

(2.3) $$U = (u^h(P_1), u^h(P_2), \ldots, u^h(P_{N_h}))^T.$$

Po wprowadzeniu tego oznaczenia, można problem (2.1)–(2.2) zapisać inaczej w postaci

(2.4) $$(A - \lambda^h B)\, U = 0,$$

gdzie A jest macierzą rzędu N_h a B jest macierzą diagonalną o elementach $\rho(P_i)$. Macierz A jest symetryczna lub nie w zależności od sposobu aproksymacji L przez L_h i od rodzaju warunków brzegowych. Zilustrujemy to w przypadku $m = 1$.

Dane jest równanie

(2.5) $$\frac{d}{dx}\left(a_1(x)\frac{du}{dx}\right) - a(x)u = \lambda u$$

z warunkami

(2.6) $$u(0) = u(l) = 0.$$

Wprowadzamy siatkę $S_h = \{x_k\colon\ x_k = kh,\ h = l/n,\ k = 1\,(1)\,n\}$, n–ilość podprzedziałów na które został podzielony przedział $\langle 0,\ 1\rangle$. Jeśli zastąpimy (2.5)–(2.6) schematem postaci:

(2.7) $$\frac{1}{h^2}\{a_1(x_k + 0.5h)u(x_k + h) - [a_1(x_k + 0.5h) + a_1(x_k - 0.5h)]\ u(x_k) +$$
$$+ a_1(x_k - 0.5h)u(x_k - h)\} - a(x_k)u(x_k) = \lambda^h u(x_k)$$
$$\text{dla} \qquad k = 1\,(1)\,n{-}1,$$

(2.8) $$u(x_0) = u(x_n) = 0,$$

wówczas macierz A w zadaniu (2.4) będzie symetryczna a B będzie macierzą jednostkową.

Macierz A nie będzie symetryczna jeśli zastąpimy (2.5)–(2.6) schematem:

(2.9) $$\frac{1}{h^2}\{a_1(x_k + h)u(x_k + h) - [a_1(x_k + h) + a_1(x_k)]\ u(x_k) +$$
$$+ a_1(x_k)u(x_k - h)\} - a(x_k)u(x_k) = \lambda^h u(x_k) \qquad \text{dla} \quad k = 1\,(1)\,n{-}1.$$

(2.10) $$u(x_0) = u(x_n) = 0.$$

Inne sposoby przejścia od zadania (1.2)–(1.3) do (2.1)–(2.2) są opisane np. w [1], str. 359–400.

3. Oszacowanie błędu metody. Wprowadzamy oznaczenia:

λ_k – k-ta wartość własna zadania (1.2)–(1.3),

u_k – k-ta funkcja własna odpowiadająca wartości własnej λ_k,

λ_k^h – k-ta wartość własna zadania różnicowego (2.1)–(2.2),

u_k^h – k-ta funkcja własna problemu (2.1)–(2.2) odpowiadająca wartości własnej λ_k^h, której wartości określamy na dyskretnym zbiorze punktów.

Różnicę $\lambda_k - \lambda_k^h$ nazywamy błędem metody.

Dodatkowo załóżmy, że w zadaniu (1.2)–(1.3) dla każdego x $\rho(x) \equiv 1$ i $lu(x) = u(x)$.

Lokalny błąd obcięcia aproksymacji L przez L_h dla dostatecznie regularnej funkcji $\varphi(x)$ definiujemy następująco:

$$R[\varphi(P)] = L\varphi(P) - L_h\varphi(P), \qquad P \in D_h. \tag{3.1}$$

Keller w [3] wykazał następujące twierdzenie: Jeśli $D_h + \partial D_h$ i operator L_h postaci

$$L_h u^h(P) = \sum_{Q \in D_h + \partial D_h} a(P, Q)\, u^h(Q) \tag{3.2}$$

spełniają warunki:

$$D_h \subset D, \qquad \partial D_h \subset \partial D, \tag{3.3}$$

$$a(P, Q) = a(Q, P) \qquad \text{dla każdego } P, Q \in D_h, \tag{3.4}$$

to dla każdej wartości własnej λ_k problemu (1.2)–(1.3) istnieje wartość własna λ_k^h zagadnienia (2.1)–(2.2) taka, że

$$|\lambda_k - \lambda_k^h| \leqslant \frac{\|R[u_k]\|_2}{\|u_k\|_2}, \tag{3.5}$$

gdzie $\|u_k\|_2 = \sqrt{\sum_{P \in D_h} u_k^2(P)}$.

Oszacowanie (3.5) jest mało przydatne w praktyce ze względu na to, że tylko nieliczne zadania różnicowe (2.1)–(2.2) prowadzą do macierzy symetrycznej $A = (a_{ij})$ w (2.4). Ponadto zostało ono otrzymane dla bardzo szczególnego zadania, w którym rozpatruje się warunki brzegowe typu zerowania się rozwiązania na brzegu.

Wad tych nie ma oszacowanie błędu wartości własnych podane przez Ljaszenko w [4].

Rozpatrywał on problem własny (1.2)–(1.3).

Niech $\partial\bar{D}_h$ oznacza zbiór węzłów D_h takich, że operator L_h w tych punktach obejmuje węzły z ∂D_h. S_h – operator różnicowy, którego postać zależy od l_h. Wówczas

$$\lambda_k - \lambda_k^h = \|u_k^h\|^{-2} \Big[\sum_{P \in D_h} R[u_k(P)]\, u_k^h(P) + \sum_{P \in \partial\bar{D}_h} S_h[u_k(P)]\, u_k^h(P)\Big], \tag{3.6}$$

gdzie

$$\|u^h\|^2 = (\rho\, u^h,\, u^h) = \sum_{P \in D_h} \rho\, u^h(P)\, u^h(P),$$

ρ – tzw. funkcja wagowa.

W szczególności, jeżeli warunki brzegowe są 1-szego rodzaju, tzn. $u(P) = 0$ dla $P \in \partial D$, wtedy po prawej stronie (3.6) nie występuje 2-gi składnik.

Oszacowanie (3.6) jest wygodne, bo obejmuje przypadek aproksymacji obszaru D nie tylko siatkami prostokątnymi ale także sześciokątnymi i trójkątnymi, Ponadto, w rozpatrywanym zadaniu własnym, zostały uwzględnione warunki brzegowe 2-go i 3-go rodzaju.

4. Metoda rozwiązywania zadania algebraicznego (2.4) za pomocą QR-algorytmu i błąd wytworzony. Zadanie (2.1)–(2.2), sprowadzone do postaci (2.4), jest zagadnieniem z algebry liniowej obliczania wartości i wektorów własnych macierzy. Do rozwiązania tego zadania można stosować QR-algorytm ([5], [6], [7]) dla dowolnych macierzy rzeczywistych. W pracach tych podane są gotowe procedury w języku Algol-60. QR-algorytm oparty jest na rozkładzie ortogonalno-trójkątnym z przesunięciami zastosowanym do macierzy Hessenberga. Wyznaczanie wartości własnych metodą QR a następnie odpowiadających im wektorów własnych jest bardzo dobrze numerycznie stabilne. Z tego względu algorytm ten polecany jest do rozwiązywania zadań typu (2.4).

Niech $\tilde{\lambda}_k^h$ oznacza k-tą wartość własną otrzymaną za pomocą QR-algorytmu. Na ogół $\tilde{\lambda}_k^h \neq \lambda_k^h$.

Różnicę $\tilde{\lambda}_k^h - \lambda_k^h$ nazywamy błędem wytworzonym przez dany algorytm znajdowania wartości własnych, np. przez QR-algorytm.

Oszacowanie błędu wytworzonego można podać np. w przypadku, gdy macierz A w zadaniu (2.4) ma pojedyncze wartości własne $\lambda_1, \lambda_2, \ldots, \lambda_n$ (patrz [8], str. 528). Jeśli $x_1, x_2, \ldots, x_n$ oznaczają wektory własne należące odpowiednio do wartości własnych $\lambda_1, \lambda_2, \ldots, \lambda_n$ a $y_1, y_2, \ldots, y_n$ – lewe wektory własne, tzn. rozwiązania układu równań

$$y^T A = \lambda y^T, \tag{4.1}$$

a ponadto E jest macierzą zaburzeń dla A, to przy założeniu, że pomijamy wszystkie iloczyny zaburzeń, mamy

$$|\tilde{\lambda}_k^h - \lambda_k^h| \leqslant \frac{\|y_k^T\|_2\, \|E\|_2\, \|x_k\|_2}{|y_k^T x_k|}, \tag{4.2}$$

gdzie $\|\cdot\|_2$ – norma Euklidesowa.

Dla QR-algorytmu, w pracy [5] podane jest oszacowanie dla $\|E\|_2$, postaci

$$\|E\|_2 \leqslant K \text{ macheps } np \|A\|_2, \tag{4.3}$$

gdzie K – stała rzędu jedności, macheps – najmniejsza liczba dodatnia w maszynie taka, że macheps + 1 $\neq$ 1, p – ilość iteracji w QR-algorytmie, n – wymiar macierzy A.

W praktyce p nie przekracza $2n$ i dlatego prawa strona w (4.3) jest wielkością rzędu 2macheps $n^2 \|A\|_2$. Dla maszyny cyfrowej Odra 1204 macheps = $1.46 \cdot 10^{-12}$.

5. Przykłady numeryczne. Oszacowania podane w § 3 zilustrowane zostaną w przypadku $m = 1$. Do eksperymentów użyto zagadnień własnych rzędu 2-go postaci:

$$-u'' + a_1(x)\,u = \lambda\, u, \tag{5.1}$$

$$u(0) = u(l) = 0, \tag{5.2}$$

przy czym funkcja $a_1(x)$ była tak dobrana, aby znane były rozwiązania dokładne tych zagadnień, tzn. wartości i funkcje własne.

W przedziale $\langle 0,\ l \rangle$ wprowadzamy siatkę

$$S_h = \{x_i\colon\ x_i = ih,\quad i = 0(1)n,\ h = l/n\}. \tag{5.3}$$

Zadanie różniczkowe (5.1)–(5.2) sprowadzane było do zadania różnicowego jednej z dwóch postaci:

$$\frac{-u^h(x_{i+1}) + 2u^h(x_i) - u^h(x_{i-1})}{h^2} + a_1(x_i)\,u^h(x_i) = \lambda^h u^h(x_i) \qquad \text{dla} \quad i = 1(1)n-1, \tag{5.4}$$

$$u^h(x_0) = u^h(x_n) = 0 \tag{5.5}$$

lub

$$\frac{-u^h(x_2) + 2u^h(x_1) - u^h(x_0)}{h^2} + a_1(x_1)\,u^h(x_1) = \lambda^h u^h(x_1),$$

$$\frac{u^h(x_{i+2}) - 16u^h(x_{i+1}) + 30u^h(x_i) - 16u^h(x_{i-1}) + u^h(x_{i-2})}{12h^2} + \tag{5.6}$$

$$+ a_1(x_i)\,u^h(x_i) = \lambda^h u^h(x_i) \qquad \text{dla} \quad i = 2(1)n-2,$$

$$\frac{-u^h(x_n) + 2u^h(x_{n-1}) - u^h(x_{n-2})}{h^2} + a_1(x_{n-1})\,u^h(x_{n-1}) = \lambda^h u^h(x_{n-1}),$$

$$u^h(x_0) = u^h(x_n) = 0. \tag{5.7}$$

Dla zadania różnicowego (5.4)–(5.5), mamy

$$R[u(x_i)] = \frac{h^2}{12}\, u^{\mathrm{IV}}(x_i) + O(h^4) = \tag{5.8}$$

$$= \frac{h^2}{12} \{[a_1''(x_i) + (a_1(x_i) - \lambda)^2]\, u(x_i) + 2a_1'(x_i)\, u'(x_i)\} + O(h^4),$$

a dla (5.6)–(5.7)

$$(5.9) \qquad R[u(x_i)] = \frac{h^2}{12} u^{IV}(x_i) + O(h^4) =$$

$$= \frac{h^2}{12}\{[a_1''(x_i) + (a_1(x_i) - \lambda)^2]\,u(x_i) + 2a_1'(x_i)u'(x_i)\} + O(h^4)$$

$$\text{dla} \qquad i = 1,\, n-1,$$

$$R[u(x_i)] = \frac{h^4}{90}\{u^{VI}(x_i)\} + O(h^6) =$$

$$= \frac{h^4}{90}\{[a_1^{IV}(x_i) + 7a_1''(x_i)(a_1(x_i) - \lambda) + (a_1(x_i) - \lambda)^3 + 4a_1'^2(x_i)]\,u(x_i) +$$

$$+ [4a_1'''(x_i) + 6a_1'(x_i)(a_1(x_i) - \lambda)]\,u'(x_i)\} + O(h^6) \qquad \text{dla} \qquad i = 2(1)n{-}2.$$

Przy eksperymentach numerycznych odrzucamy w $R[u(x_i)]$ składniki $O(h^4)$ i wyższe. W wyrażeniach występujących po prawych stronach oszacowań podanych w § 3 λ i u zastępujemy wartościami λ i u otrzymanymi z QR-algorytmu, równocześnie przyjmując

$$(5.10) \qquad u'(x_i) \approx \frac{u^h(x_{i+1}) - u^h(x_{i-1})}{2h} \qquad \text{dla} \qquad i = 1(1)n{-}1.$$

Prawe strony w oszacowaniach z § 3 można przedstawić jako $h^p\,\psi(\lambda_k,\, u_k,\, u_k',\, u_k^h)$.

W opisanych metodach dla wartości własnych mamy oszacowanie $\lambda_k - \lambda_k^h = O(h^p)$ i podobnie dla funkcji własnych $u_k - u_k^h = O(h^p)$. Wtedy

$$u_k'(x_i) = \frac{u_k^h(x_{i+1}) - u_k^h(x_{i-1})}{2h} + O(h^{p-1})$$

i

$$(5.11) \qquad h^p\psi(\lambda_k,\, u_k,\, u_k',\, u_k^h) = h^p\psi\left(\lambda_k^h,\, u_k^h,\, \frac{u_k^h(x_{i+1}) - u_k^h(x_{i-1})}{2h},\, u_k^h\right) + O(h^{p-1}).$$

Obliczenia przeprowadzono na emc Odra-1204.

Do rozwiązania zadania (5.4)–(5.5) i (5.6)–(5.7) użyto procedury QR-algorytmu podanej w [2], która została oparta na procedurach z prac [5]–[7]. Dla (5.4)–(5.5) macierz A w (2.4) jest symetryczna a dla (5.6)–(5.7) – niesymetryczna, dlatego w 1-szym przypadku było stosowane oszacowanie (3.5) i (3.6) a w 2-gim (3.6).

$$\text{I} \qquad -u'' = \lambda u,$$

$$u(0) = u(1) = 0.$$

Rozwiązanie dokładne: $\lambda_k = (k\pi)^2$, $k = 1, 2, \ldots$

$$u_k(x) = \sin k\pi x.$$

Wyniki obliczeń dla różnych wartości n i dla najmniejszej wartości własnej $\lambda_1 = 9.8696044$ zostaną podane w tabelach poniżej.

S c h e m a t (5.4)–(5.5). W tym przypadku oszacowanie (3.5) jest identyczne z (3.6).

TABELA 5.1

n	oszacowanie (3.5)	błąd wytworzony z (4.2)	$\lambda_1 - \lambda_1^h$
8	$0.122987_{10}+0$	$0.1026660_{10}-5$	$0.1262_{10}+0$
10	$0.795887_{10}-1$	$0.2690805_{10}-5$	$0.8090_{10}-1$
16	$0.314652_{10}-1$	$0.2094520_{10}-4$	$0.3167_{10}-1$
18	$0.201937_{10}-1$	$0.3516317_{10}-4$	$0.2028_{10}-1$

S c h e m a t (5.6)–(5.7).

TABELA 5.2

n	oszacowanie (3.6)	błąd wytworzony z (4.2)	$\lambda_1 - \lambda_1^h$
8	$-0.137620_{10}-1$	$0.119756_{10}-6$	$0.1318_{10}-1$
10	$-0.463961_{10}-2$	$0.292187_{10}-6$	$0.4623_{10}-2$
15	$-0.567871_{10}-3$	$0.147852_{10}-5$	$0.6910_{10}-3$

II
$$-u'' + \left(\frac{0.75}{\sin^2 x} - 0.25\right)u = \lambda u,$$
$$u(0) = u(\pi) = 0.$$

Rozwiązanie dokładne: $\lambda_k = k(k+1), \quad k = 1, 2, \ldots$

$$u_k(x) = \sqrt{\sin x}\, P_k^{(1)}(\cos x),$$

gdzie $P_k^1(x)$ – tzw. dołączone funkcje Legendre'a 1-szego rodzaju. Wyniki obliczeń zostaną podane dla najmniejszej wartości własnej $\lambda_1 = 2$.

S c h e m a t r ó ż n i c o w y (5.4)–(5.5).

TABELA 5.3

n	oszacowanie (3.6)	oszacowanie (3.5)	błąd wytworzony z (4.2)	$\lambda_1 - \lambda_1^h$
8	$0.654258_{10}-1$	$0.235515_{10}+0$	$0.855726_{10}-7$	$0.6052_{10}-1$
10	$0.433140_{10}-1$	$0.113801_{10}+0$	$0.233157_{10}-6$	$0.4049_{10}-1$
16	$0.180737_{10}-1$	$0.100425_{10}+0$	$0.192406_{10}-5$	$0.1720_{10}-1$

S c h e m a t r ó ż n i c o w y (5.6)–(5.7).

TABELA 5.4

n	oszacowanie (3.6)	błąd wytworzony z (4.2)	$\lambda_1 - \lambda_1^h$
8	$0.213014_{10}-1$	$0.166801_{10}-7$	$0.1575_{10}-1$
10	$0.138736_{10}-1$	$0.406973_{10}-7$	$0.9899_{10}-2$
16	$0.561949_{10}-2$	$0.205961_{10}-6$	$0.3941_{10}-2$

W podanych przykładach obliczenia były przeprowadzone dla najmniejszych wartości własnych ze względu na to, że w metodzie różnicowej są one obliczane najdokładniej. W przypadku zagadnienia własnego I te uproszczone oszacowania (odrzucamy $O(h^4)$ lub $O(h^6)$ we wzorach na $R[u_k]$) dają dla większych n prawie dokładnie różnicę $\lambda_1 - \lambda_1^h$ (tabele 5.1 i 5.2).

Z tabeli 5.3 widać wyraźnie, że oszacowanie (3.6) jest dokładniejsze od (3.5). Ponadto jest ono wygodniejsze także dlatego, że macierz A w (2.4) nie musi być symetryczna i daje ono większą swobodę jeśli chodzi o postać warunków brzegowych.

Przy wzroście n błąd metody maleje a błąd wytworzony rośnie. Dla zadań, które można rozwiązywać na emc Odra 1204 (bez pamięci zewnętrznej), n jest rzędu 30 i błąd wytworzony dla takich n stosunkowo mało wpływa na całkowite oszacowanie.

Bibliografia

[1] Л. Колляц, *Задачи на собственные значения,* Москва 1968.

[2] M. Jankowski, *Numeryczne rozwiązywanie zagadnienia własnego,* Sprawozdanie Inst. Mat. i Zakładu Obl. Num. Uniwersytetu Warszawskiego nr 32.

[3] H. B. Keller, *On the accuracy of finite difference approximations to the eigenvalues of differential and integral operators,* Num. Math. 7(5) (1965), str. 412–419.

[4] Н.Н. Ляшенко, *К оцене прогрешности при определении собственных значений методом конечных разностей,* Вычислительная и прикладная математика 9, 1969, str. 79–87.

[5] S. R. Martin, G. Peters and J. H. Wilkinson, *The QR algorithm for real Hessenberg matrices,* Num. Math. 14(3) (1970), str. 219–231.

[6] R. S. Martin and J. H. Wilkinson, *Similarity reduction of general matrix to Hessenberg form,* Num. Math. 12(5) (1968), str. 349–368.

[7] B. N. Parlett and C. Reinsch, *Balancing a matrix for calculation of eigenvalues and eigenvectors,* Num. Math. 13(4) (1969), str. 293–304.

[8] A. Ralston, *Wstęp do analizy numerycznej,* Warszawa 1971.