Proof. Let $\mathfrak{A}_H$ be the discretisation of the space H defined by a family of orthogonal projections $\{\pi_h^H\}_{h\in\Theta}$ onto subspaces V_h. Since the family $\{\pi_h^H\}_{h\in\Theta}$ is uniformly bounded, then from Lemma 3.2 it follows that $\mathfrak{A}_H$ approximates H. By Lemma 2.5 we see that π_h^H is entirely determined by the basis $\{\varphi_1^{Hh}, \ldots, \varphi_{d_h}^{Hh}\}$. It now suffices to apply Theorem 6.3. ■

The trivial example of a problem satisfying conditions (6.1) and (6.2) stated in this paragraph is the following

$$\Delta u = \lambda u \quad \text{on } \Omega,$$
$$u = 0 \quad \text{on } \partial\Omega,$$

where Ω is a bounded domain in R^d with regular boundary. In this case we can take $V = H_0^2(\Omega)$, $H = H^0(\Omega) = L_2(\Omega)$ and $A = \Delta$ (in generalized sense). In this case, by Rellich Theorem, the imbedding $J\colon V \to H$ is a compact linear map, and A and A^{-1} are bounded operators.

More advanced examples of elliptic problems are discussed also in [5] (see Section 4, as well as works quoted in this paper).

Remark. In the paper [5] there are given error bounds obtained in a similar way, in the case of a differential operator in the Sobolev space. The information about the rate of convergence is obtained when the spaces $V_h = p_h(X_h)$ are special finite-dimensional subspaces of piecewise polynomials.

References

[1] N. Dunford and J. T. Schwartz, *Linear operators*, Part I, Chapter VII, Interscience, New York 1958.
[2] J. P. Aubin, *Approximation of elliptic boundary-value problems*, Wiley–Interscience, New York 1972.
[3] J. L. Lions, *Analyse numerique*, Ecole Polytechnique 1971.
[4] F. Riesz et Sz.-Nagy, *Leçons d'analyse fonctionelle*, Budapest 1952.
[5] J. H. Bramble and J. E. Osborn, Mathematics of computations, 27 (123) (1973), pp. 525–540.

Presented to the Semester
Mathematical Models and Numerical Methods
(February 3–June 14, 1975)

BANACH CENTER PUBLICATIONS
VOLUME 3

ОБ ЭФФЕКТИВНЫХ МЕТОДАХ ВЫЧИСЛЕНИЯ СОБСТВЕННЫХ ЧИСЕЛ И СОБСТВЕННЫХ ВЕКТОРОВ С ПОМОЩЬЮ МЕТОДА ЯКОБИ И МЕТОДА ЛОЖНЫХ ВОЗМУЩЕНИЙ

Ф. КУНЕРТ

Технический ВУЗ, Карл-Маркс-Штадт, ГДР

В заметке содержится часть лекций, прочитанных автором во время работы семестра „Математические модели и численные методы" Международного математического центра им. Стефана Банаха в Варшаве. Описываются эффективные методы приближенного вычисления отдельного собственного числа и соответствующего ему собственного вектора симметрической матрицы, состоящие из двух этапов. На первом этапе применяется метод Якоби элементарных плоских вращений до тех пор, пока мы не сможем указать отрезок на вещественной оси, содержащий интересующее нас собственное число. На втором этапе применяется метод ложных возмущений для уточнения собственного числа и собственного вектора.

Метод Якоби, являющийся „глобальным" методом для полной проблемы собственных чисел, устанавливает при этом условия, достаточные для применимости и быстрой сходимости метода ложных возмущений. Метод ложных возмущений является „локальным" методом для частной проблемы собственных чисел.

Идея соединения этих двух методов содержится в работе [12] (см. также [13]).

1. Теоремы теории возмущения спектра

Приведем некоторые основные теоремы теории возмущения спектра самосопряженных операторов A в гильбертовом пространстве H со скалярным произведением $(\cdot,\cdot)$ (детальное изложение приведенных здесь фактов можно найти, например, в монографиях [1] и [8], см. также [9]).

ТЕОРЕМА 1. *Пусть A самосопряженный оператор в гильбертовом пространстве H, B ограниченный самосопряженный оператор и пусть E_t и E_t' обозначают* *спектральные разложения операторов A и $A' = A+B$. Обозначим через $\Delta = [a,b]$ и $\Delta' = [a',b'] = [a-\|B\|-\varepsilon,\ b+\|B\|+\varepsilon]$, $\varepsilon>0$ произвольно, конечные промежутки.*

Тогда имеем:

$$\dim G = \dim E(\Delta)H \leqslant \dim G' = \dim E'(\Delta')H,$$

где $E(\Delta) = E_b - E_a$, $E'(\Delta') = E'_{b'} - E'_{a'}$.

При помощи этой теоремы доказывается

Теорема 2. *Пусть A, B и A' те же, как в теореме 1. Пусть спектр оператора A в конечном промежутке $\Delta = [a,b]$ состоит из собственных чисел суммарной кратности $k<\infty$. Пусть a и b сами не являются точками спектра оператора A и пусть расстояние от a и b до ближайшей точки спектра оператора A больше, чем $2\|B\|$.*

Тогда спектр оператора A' в промежутке $\tilde{\Delta} = [a-\|B\|,\ b+\|B\|]$ состоит из собственных чисел суммарной кратности k.

Доказательство. Обозначим через $\sigma(A)$ спектр оператора A. По предположению имеем:

$$\min\{\inf_{\lambda\in\sigma(A)\setminus\Delta}|a-\lambda|,\ \inf_{\lambda\in\sigma(A)\setminus\Delta}|b-\lambda|\} = 2\|B\|+\delta,\qquad \delta>0.$$

Тогда для каждого промежутка $\Delta_\gamma = [a-2\|B\|-\gamma,\ b+2\|B\|+\gamma]$, $0\leqslant\gamma\leqslant\delta$ имеем

$$E(\Delta_\gamma) = E(\Delta). \tag{1.1}$$

Пусть $\varepsilon<\delta/4$. Обозначим $\Delta' = [a-\|B\|-\varepsilon,\ b+\|B\|+\varepsilon]$ и применим теорему 1:

$$k = \dim G = \dim E(\Delta)H \leqslant \dim E'(\Delta')H = \dim G'. \tag{1.2}$$

После введения обозначений

$$A'' = A+B+(-B) = A'+(-B),$$

$$\Delta'' = [a'-\|B\|-\varepsilon,\ b'+\|B\|+\varepsilon] = [a-2\|B\|-2\varepsilon,\ b+2\|B\|+2\varepsilon] = \Delta_{2\varepsilon},$$

мы применим теорему 1 вторично (при этом мы исходим из оператора A' и промежутка Δ'). Учитывая соотношение (1.1), мы получим неравенство

$$\dim G' \leqslant \dim E(\Delta'')H = \dim E(\Delta_{2\varepsilon})H = k,$$

которое вместе с (1.2) дает равенство

$$k = \dim G' = \dim E'(\Delta')H.$$

Отсюда получим ввиду произвольности $\varepsilon>0$ требуемое соотношение $k = \dim E'(\tilde{\Delta})H$. ■

Теорема 3. *Пусть B ограниченный самосопряженный оператор и пусть λ обозначает простое собственное число самосопряженного оператора A. Если выполнено условие*

$$\varrho = \inf_{\substack{\mu\in\sigma(A)\\ \mu\neq\lambda}}|\lambda-\mu| > 2\|B\|, \tag{1.3}$$

то спектр оператора $A' = A+B$ в промежутке $\Delta' = [\lambda-\|B\|,\ \lambda+\|B\|]$ состоит из единственного простого собственного числа λ':

$$|\lambda-\lambda'| \leqslant \|B\|. \tag{1.4}$$

Доказательство этой теоремы мы получим, если применим теорему 2 на случай промежутков $\Delta = [\lambda-\delta,\ \lambda+\delta]$ и $\tilde{\Delta} = [\lambda-\|B\|-\delta,\ \lambda+\|B\|+\delta]$ для достаточно малых $\delta>0$, а затем перейдем к пределу при $\delta\to0$.

Теорема 4. *Пусть выполнены условия теоремы 3. Тогда имеем для $\nu'\in\sigma(A')$, $\nu'\neq\lambda'$, оценку*

$$|\lambda-\nu'| \geqslant \varrho-\|B\|.$$

Доказательство этой теоремы можно найти в работе [3] (см. также [9]).

Предположим, что самосопряженный оператор A имеет простое собственное число λ. Пусть

$$A\varphi = \lambda\varphi,\qquad \|\varphi\| = 1,\qquad \varphi\in D(A). \tag{1.5}$$

Если мы обозначим $H_\varphi = \{u\in H\colon (u,\varphi)=0\}$, то оператор $A-\lambda I$ осуществляет взаимно однозначное отображение множества $H_\varphi\cap D(A)$ на подпространство H_φ.

Оператор, обратный к $A-\lambda I$ на подпространстве H_φ, обозначим через R. Он является самосопряженным и ввиду равенства $D(R) = H_\varphi$ ограниченным оператором. Полагая $R\varphi = 0$, мы расширим оператор R на все пространство H, так что R есть самосопряженный и ограниченный оператор, определенный на всем пространстве H.

Очевидно, оператор $(A-\lambda I)R$ есть ортопроектор P в пространстве H на подпространство H_φ, так что справедливо соотношение:

$$\begin{aligned}(A-\lambda I)Rf &= (A-\lambda I)R(f-(f,\varphi)\varphi)+(f,\varphi)(A-\lambda I)R\varphi\\ &= f-(f,\varphi)\varphi,\qquad f\in H.\end{aligned} \tag{1.6}$$

Аналогично, имеем для $f\in D(A)$ соотношение:

$$R(A-\lambda I)f = f-(f,\varphi)\varphi. \tag{1.7}$$

Таким образом, оператор $R(A-\lambda I)$ совпадает на множестве $D(A)$ с ортопроектором P и может быть расширен на все пространство H.

Для нормы оператора R имеем равенство:

$$\|R\| = \varrho^{-1}. \tag{1.8}$$

Теорема 5. *Пусть выполнены условия теоремы 3. Тогда вектор*

$$\varphi' = \{I-R(\mu I-B)\}^{-1}\varphi,\qquad \mu = \lambda'-\lambda, \tag{1.9}$$

является собственным вектором оператора $A' = A+B$ к собственному числу λ', причем справедливо $(\varphi,\varphi') = 1$.

Доказательство. Пусть $\varphi' \neq 0$ есть собственный вектор оператора A', соответствующий собственному числу λ'. Предположим, что $(\varphi, \varphi') = 0$. Тогда из уравнения $(A' - \lambda' I)\varphi' = 0$ получим

$$(A - \lambda I + B - \mu I)\varphi' = 0$$

или

$$(1.10) \qquad (A - \lambda I)\varphi' = (\mu I - B)\varphi'.$$

Применяя на обе стороны уравнения (1.10) оператор R и учитывая соотношение (1.7), получим

$$\varphi' = R(\mu I - B)\varphi'.$$

Отсюда следует в результате (1.4), (1.8) и предположения $2\|B\| < \varrho$ неравенство

$$\|\varphi'\| \leqslant \|R\| (|\mu| + \|B\|) \|\varphi'\| < \|\varphi'\|.$$

Итак, мы можем считать, что $(\varphi, \varphi') \neq 0$. Выбираем нормировку вектора φ' в соответствии с равенством $(\varphi, \varphi') = 1$. В этом случае мы получим в результате применения оператора R на обе стороны уравнения (1.10) соотношение

$$\varphi' - \varphi = R(\mu I - B)\varphi'$$

или

$$(1.11) \qquad \varphi' = \{I - R(\mu I - B)\}^{-1}\varphi.$$

Содержащийся в этой формуле обратный оператор существует, так как в силу наших предположений выполнено неравенство $\|R(\mu I - B)\| < 1$. ■

Если мы скалярно умножим обе стороны уравнения (1.10) на вектор φ, то мы получим в силу соотношения

$$((A - \lambda I)\varphi', \varphi) = (\varphi', (A - \lambda I)\varphi) = 0$$

следующее уравнение

$$((\mu I - B)\varphi', \varphi) = 0,$$

которое дает

$$(1.12) \qquad \mu = (B\varphi', \varphi) = (\varphi', B\varphi).$$

2. Метод ложных возмущений

В этом пункте мы опишем приближенный метод вычисления простого собственного числа и соответствующего ему собственного вектора самосопряженного оператора, который основывается на результатах теории возмущения спектра. Исходя из формул (1.11) и (1.12), разлагаем вектор φ' и число μ в бесконечные ряды, в которых сохраняем лишь некоторые первые члены. При этом оператор-возмущение B задается специальным образом. Получающийся численный метод описан для самосопряженных операторов в работе [4] и для несамосопряженных операторов в работе [10] (см. также [11]).

Пусть для самосопряженного оператора A требуется определить собственное число λ и соответствующий ему собственный вектор φ. Предположим, что задано некоторое приближение $\varphi_0 \in D(A)$ к вектору φ. Для определенности считаем, что $\|\varphi_0\| = 1$. Обозначим

$$\lambda_0 = (A\varphi_0, \varphi_0), \quad \sigma_0 = A\varphi_0 - \lambda_0\varphi_0.$$

Очевидно, справедливо равенство $(\varphi_0, \sigma_0) = 0$.

Сущность метода ложных возмущений состоит в построении оператора-возмущения B_0 такого, что для оператора $A_0 = A - B_0$ выполняется уравнение

$$(2.1) \qquad A_0\varphi_0 = \lambda_0\varphi_0.$$

Для того, чтобы применить выше приведенные теоремы теории возмущения спектра, оператор B_0 должен быть ограниченным и самосопряженным. Далее, из уравнения (2.1) вытекает, что этот оператор должен удовлетворить условию

$$(2.2) \qquad B_0\varphi_0 = \sigma_0.$$

Отсюда видно, что норма $\|B_0\|$ не меньше, чем норма $\|\sigma_0\|$: $\|B_0\| \geqslant \|\sigma_0\|$.

Оператор B_0, удовлетворяющий всем этим условиям, определен неоднозначно. В дальнейшем мы рассмотрим оператор B_0, действующий по формуле

$$(2.3) \qquad B_0 f = (f, \varphi_0)\sigma_0 + (f, \sigma_0)\varphi_0.$$

Этот оператор B_0 удовлетворяет всем выше перечисленным условиям, кроме того он вполне непрерывен и для его нормы мы имеем равенство $\|B_0\| = \|\sigma_0\|$.

Если λ_0 является простым собственным числом оператора A_0, то мы можем рассмотреть (по аналогии с оператором R) ограниченный и самосопряженный оператор R_0, обратный к оператору $A_0 - \lambda_0 I$ на подпространстве $H_{\varphi_0} = \{v \in H: (v, \varphi_0) = 0\}$. Положим $R_0\varphi_0 = 0$. Тогда имеем

$$\|R_0\| = \varrho_0^{-1}, \quad \varrho_0 = \inf_{\substack{\nu \in \sigma(A_0) \\ \nu \neq \lambda_0}} |\lambda_0 - \nu|.$$

Условиями, достаточными для простоты собственного числа λ_0 по отношению к оператору A_0, являются, например,

$$(2.4) \qquad \|\sigma_0\| < \varrho/3, \quad |\lambda - \lambda_0| < \varrho - \|\sigma_0\|$$

(см. [4]). При этих условиях справедлива оценка

$$\varrho_0 > \varrho/3.$$

Кроме того имеет место следующее разложение для собственного вектора φ оператора A:

$$(2.5) \quad \varphi = \{I - R_0(\mu_0 I - B_0)\}^{-1}\varphi_0 = \varphi_0 - \sum_{k=1}^{\infty} \mu_0^{k-1} R_0^k \sigma_0, \quad \mu_0 = \lambda - \lambda_0.$$

Здесь вектор φ нормирован в соответствии с условием $(\varphi, \varphi_0) = 1$.

Первый шаг итерационного метода ложных возмущений заключается в том, что в разложении (2.5) оставим лишь слагаемые, независимые от μ_0, т.е. мы считаем вектор

$$\varphi_1 = \frac{\varphi_0 - R_0\sigma_0}{\|\varphi_0 - R_0\sigma_0\|} = \frac{\varphi_0 - u_0}{\|\varphi_0 - u_0\|} \tag{2.6}$$

следующим за φ_0 приближением. При этом вектор u_0 является единственным решением системы

$$(A_0 - \lambda_0 I)x = \sigma_0, \quad (x, \varphi_0) = 0. \tag{2.7}$$

Решение $x = u_0 = R_0\sigma_0$ системы (2.7) обычно называется нормальным решением операторного уравнения $(A_0 - \lambda_0 I)x = \sigma_0$.

Отсюда получим следующее формальное описание итерационного метода ложных возмущений:

Пусть известны векторы $\varphi_0, \ldots, \varphi_n$. Положим

$$\lambda_n = (A\varphi_n, \varphi_n), \qquad \sigma_n = A\varphi_n - \lambda_n\varphi_n,$$

и рассмотрим оператор B_n, действующий по формуле

$$B_n f = (f, \varphi_n)\sigma_n + (f, \sigma_n)\varphi_n, \quad f \in H.$$

Обозначим $A_n = A - B_n$. Определим очередной вектор φ_{n+1} по формуле

$$\varphi_{n+1} = \frac{\varphi_n - u_n}{\|\varphi_n - u_n\|}, \tag{2.6$_n$}$$

где через u_n обозначено единственное решение системы

$$(A_n - \lambda_n I)x = \sigma_n, \quad (x, \varphi_n) = 0. \tag{2.7$_n$}$$

При этом мы конечно предполагаем, что λ_n является простым собственным числом оператора A_n. В работе [4] показано, что достаточным для этого является выполнение условий (2.4). Кроме того, выполнением неравенств (2.4) для начального вектора φ_0 гарантируется сходимость последовательностей σ_n, $\varphi - \varphi_n$, $\lambda - \lambda_n$ со скоростью q^{3^n}, $q = \text{const} < 1$.

Для описанного варианта метода ложных возмущений приближенного вычисления простого собственного числа λ самосопряженного оператора A и соответствующего собственного вектора φ требуется наличие начального приближения φ_0 и выполнение некоторых условий, достаточных для реализации расчетных формул и сходимости. В следующем пункте мы покажем, как в случае симметрических матриц можно определить начальный вектор φ_0, удовлетворяющий некоторым достаточным условиям сходимости метода ложных возмущений. С этой целью приведем ряд теорем.

Теорема 6. *Если простое собственное число λ оператора A является единственной точкой спектра этого оператора в промежутке $\Delta_0 = [\lambda_0 - \|\sigma_0\|, \lambda_0 + \|\sigma_0\|]$, то λ_0 является простым собственным числом оператора A_0 и изолированной точкой спектра этого оператора.*

Доказательство. Так как спектр самосопряженного оператора замкнут, то можно указать $\delta > 0$ такое, что спектр оператора A в промежутке $\Delta_\delta = [\lambda_0 - \|\sigma_0\| - \delta, \lambda_0 + \|\sigma_0\| + \delta]$ состоит из единственного числа λ. Из полной непрерывности оператора B_0 мы заключаем, что в промежутке Δ_δ не находится ни одна точка непрерывного спектра оператора $A_0 = A - B_0$. Тогда число λ_0 можно окружить промежутком $[\lambda_0 - \gamma, \lambda_0 + \gamma]$, $\gamma < \delta/2$, таким, что он кроме λ_0 не содержит никаких других точек спектра оператора A_0.

Пусть кратность k_0 собственного числа λ_0 оператора A_0 больше 1. Тогда мы получим из теоремы 1 в силу $\|B_0\| = \|\sigma_0\|$, что промежуток $\Delta_{\gamma+\varepsilon} = [\lambda_0 - \|\sigma_0\| - \gamma - \varepsilon, \lambda_0 + \|\sigma_0\| + \gamma + \varepsilon]$, $\varepsilon > 0$, содержит точки спектра оператора A, причем, справедливо соотношение

$$\dim E(\Delta_{\gamma+\varepsilon})H = k \geqslant k_0 > 1.$$

Но это противоречит при $\varepsilon < \delta/2$ предположению о том, что спектр оператора A в промежутке Δ_δ содержит лишь простое собственное число λ. ■

Теорема 7. *Если простое собственное число λ оператора A является единственной точкой спектра этого оператора в промежутке $\Delta_0 = [\lambda_0 - \|\sigma_0\|, \lambda_0 + \|\sigma_0\|]$ и если выполнено неравенство*

$$\|\sigma_0\| = \|B_0\| < \varrho_0 \tag{2.8}$$

то

$$\varphi = \varphi_0 - \sum_{k=1}^{\infty} \mu_0^{k-1} R_0^k \sigma_0, \qquad \mu_0 = \lambda - \lambda_0, \tag{2.9}$$

есть собственный вектор оператора A к собственному числу λ, причем, справедливо условие нормировки $(\varphi, \varphi_0) = 1$.

Доказательство. При выполнении условия (2.8) ряд справа в формуле (2.9) сходится и равняется вектору $\{I - R_0(\mu_0 I - B_0)\}^{-1}\varphi_0$, который является собственным для оператора A к собственному числу λ (см. формулу (2.5)). ■

Итак, если условия теоремы 7 выполнены, то определение вектора φ_1 по формуле (2.6) осмысленно, причем, $u_0 = R_0\sigma_0$ является единственным решением системы (2.7).

Умножая следующее очевидное равенство

$$\begin{aligned} A\varphi_1 - \lambda_0\varphi_1 &= (A - \lambda_0 I)\varphi_1 = (1 + \|u_0\|^2)^{-1/2}(A_0 + B_0 - \lambda_0 I)(\varphi_0 - u_0) = \\ &= -(1 + \|u_0\|^2)^{-1/2}(u_0, \sigma_0)\varphi_0 \end{aligned} \tag{2.10}$$

скалярно на вектор φ_1, мы получим при помощи обозначения $\lambda_1 = (A\varphi_1, \varphi_1)$ следующее соотношение

$$\lambda_1 - \lambda_0 = -(1 + \|u_0\|^2)^{-1/2}(u_0, \sigma_0)(\varphi_0, \varphi_1) = -(1 + \|u_0\|^2)(u_0, \sigma_0) \tag{2.11}$$

между приближенными числами λ_0 и λ_1. Отсюда вытекает оценка

$$(2.12)\qquad |\lambda_0-\lambda_1| \leqslant \frac{\|u_0\|}{1+\|u_0\|^2}\,\|\sigma_0\| < \tfrac{1}{2}\|\sigma_0\|\,.$$

Далее, из равенства $\sigma_1=(A-\lambda_1 I)\varphi_1$ мы получим в силу $\|\varphi_1\|=1$ и $(\sigma_1,\varphi_1)=0$ следующий результат:

$$\|\sigma_1\|^2 = \big((A-\lambda_1 I)\varphi_1,\sigma_1\big) = \big((A-\lambda_0 I)\varphi_1,\sigma_1\big) = \\ = \big((A-\lambda_0 I)\varphi_1,(A-\lambda_1 I)\varphi_1\big) = \|(A-\lambda_0 I)\varphi_1\|^2-(\lambda_0-\lambda_1)^2.$$

Подставляя сюда соотношения (2.10) и (2.11) мы имеем оценку

$$\|\sigma_1\|^2 = (1+\|u_0\|^2)^{-1}|(u_0,\sigma_0)|^2-(1+\|u_0\|^2)^{-2}|(u_0,\sigma_0)|^2 = \\ = \frac{\|u_0\|^2\,|(u_0,\sigma_0)|^2}{(1+\|u_0\|^2)^2} \leqslant \frac{\|u_0\|^4}{(1+\|u_0\|^2)^2}\,\|\sigma_0\|^2.$$

Если мы учитываем неравенство

$$\|u_0\| = \|R_0\sigma_0\| \leqslant \|R_0\|\,\|\sigma_0\| < 1,$$

являющееся следствием условий теоремы 7, то получим из последней оценки следующую оценку нормы $\|\sigma_1\|$ через норму $\|\sigma_0\|$:

$$(2.13)\qquad \|\sigma_1\| \leqslant \tfrac{1}{2}\|\sigma_0\|\,.$$

Оценки (2.12) и (2.13), справедливые при выполнении условий теоремы 7, показывают, что промежуток $\Delta_1=[\lambda_1-\|\sigma_1\|,\ \lambda_1+\|\sigma_1\|]$ содержится в промежутке $\Delta_0=[\lambda_0-\|\sigma_0\|,\ \lambda_0+\|\sigma_0\|]$: $\Delta_1\subset\Delta_0$.

Тогда справедлива

Теорема 8. *Если выполнены условия теоремы 7, то собственное число λ оператора A является единственной точкой спектра этого оператора в промежутке $\Delta_1=[\lambda_1-\|\sigma_1\|,\ \lambda_1+\|\sigma_1\|]$. Кроме того, λ_1 является простым собственным числом оператора $A_1=A-B_1$,*

$$B_1 f = (f,\varphi_1)\sigma_1+(f,\sigma_1)\varphi_1,\qquad f\in H,$$

и изолированной точкой спектра этого оператора.

Доказательство. Число λ_1 является, очевидно, собственным числом оператора A_1. Следовательно, оператор $A=A_1+B_1$, $\|B_1\|=\|\sigma_1\|$, имеет в промежутке $\Delta_1=[\lambda_1-\|\sigma_1\|,\lambda_1+\|\sigma_1\|]$ точки спектра (см. теорему 2). Но так как справедливо $\Delta_1\subset\Delta_0$ и спектр оператора A в промежутке Δ_0 состоит из одного лишь простого собственного числа λ, то необходимо $\lambda\in\Delta_1$. Простота собственного числа λ_1 оператора A_1 доказывается точно так же, как простота собственного числа λ_0 оператора A_0 в доказательстве теоремы 6. ∎

Для того, чтобы применить теорему 7 на ситуацию, полученную в результате проведения первого шага итерационного процесса, необходимо выполнение условия

$$\|\sigma_1\|<\varrho_1,\qquad \varrho_1=\inf_{\substack{\nu\in\sigma(A_1)\\ \nu\neq\lambda_1}}|\nu-\lambda_1|.$$

Однако, условия теоремы 7 не гарантируют, вообще говоря, выполнения этого неравенства.

Теорема 9. *Пусть простое собственное число λ оператора A является единственной точкой спектра как в промежутке $\Delta_0=[\lambda_0-\|\sigma_0\|,\ \lambda_0+\|\sigma_0\|]$ так и в промежутке $\Delta^0=[\lambda_0-2\|\sigma_0\|,\ \lambda_0+2\|\sigma_0\|]$ и пусть выполнено неравенство $\|\sigma_0\|<\varrho_0$.*

Тогда λ является единственной точкой спектра оператора A как в промежутке $\Delta_1=[\lambda_1-\|\sigma_1\|,\ \lambda_1+\|\sigma_1\|]$, так и в промежутке $\Delta^1=[\lambda_1-2\|\sigma_1\|,\ \lambda_1+2\|\sigma_1\|]$. Кроме того, справедлива оценка $\|\sigma_1\|<\varrho_1$.

Доказательство. Первая часть утверждения теоремы 9 содержится в теореме 8, или может быть доказано совершенно аналогично.

Так как имеем $\lambda\in\Delta_0$ и так как Δ^0 не содержит точек спектра оператора A кроме собственного числа λ, то разность между λ и остальными точками λ' спектра, $\lambda'\neq\lambda$, не меньше чем $\|\sigma_0\|$. С другой стороны, точки спектра $\lambda_1'\neq\lambda_1$ оператора $A_1=A-B_1$, $\|B_1\|=\|\sigma_1\|<\frac{1}{2}\|\sigma_0\|$, находятся в области (см. теорему 1)

$$\bigcup_{\substack{\lambda'\in\sigma(A)\\ \lambda'\neq\lambda}}[\lambda'-\|\sigma_1\|,\ \lambda'+\|\sigma_1\|]\,.$$

Но расстояние между этой областью и промежутком $\Delta_0=[\lambda_0-\|\sigma_0\|,\ \lambda_0+\|\sigma_0\|]$ больше чем δ,

$$\delta \geqslant \|\sigma_0\|-\|\sigma_1\| > \tfrac{1}{2}\|\sigma_0\| > \|\sigma_1\|\,.$$

Но так как $\lambda_1\in\Delta_1\subset\Delta_0$, то имеем для всех точек спектра λ_1' оператора A_1, $\lambda_1'\neq\lambda_1$, оценку

$$|\lambda_1-\lambda_1'| \geqslant \delta > \|\sigma_1\|\,.$$

Следовательно, имеем

$$\varrho_1 = \inf_{\substack{\lambda_1'\in\sigma(A_1)\\ \lambda_1'\neq\lambda_1}}|\lambda_1-\lambda_1'| \geqslant \delta > \|\sigma_1\|\,. \ ∎$$

Условия теоремы 9 таковы, что после проведения одного шага итерационного метода ложных возмущений, мы можем получить аналогичные условия и, тем самым, итерационный процесс осуществим неограничено. Кроме того, из оценок (2.12) и (2.13) видно, что после нескольких шагов выполняются неравенства, аналогичные (2.4). Этим доказывается сходимость итерационного метода ложных возмущений.

3. Метод Якоби

В этом пункте мы покажем, что для вещественных симметрических матриц применение метода Якоби приближенного вычисления собственных чисел является эффективным приемом для установления условий теоремы 9. При этом рассматривается частная задача вычисления собственных чисел, а именно приближенное вычисление некоторого одного собственного числа матрицы. Метод Якоби применяется до тех пор, пока не выполнены условия теоремы 9 (первый этап). После этого счет продолжается по формулам метода ложных возмущений, т.е. по формулам (2.6_n) и (2.7_n), где начальный вектор φ_0 определяется на этапе работы метода Якоби (второй этап).

Обозначим через $A = (a_{kj})$, $k, j = 1, \ldots, m$, вещественную симметрическую матрицу, а через $\boldsymbol{R}^m$ m-мерное линейное пространство векторов $x = (x_1, \ldots \ldots, x_m)^T$, $x_j \in R$ (R обозначает поле вещественных чисел). Введем в пространстве $\boldsymbol{R}^m$ скалярное произведение по формуле

$$(x, y) = \sum_{j=1}^{m} x_j y_j .$$

Предположим, что матрица A имеет m попарно различных собственных чисел $\Lambda_1, \ldots, \Lambda_m$, упорядоченных убывающим образом:

$$\Lambda_1 > \Lambda_2 > \ldots > \Lambda_m .$$

Через $\psi_1, \ldots, \psi_m$ обозначим ортонормированную систему собственных векторов матрицы A: $A\psi_j = \Lambda_j \psi_j$.

Ставится задача приближенного вычисления k-го собственного числа Λ_k и соответствующего ему собственного вектора ψ_k матрицы A.

Метод Якоби состоит в определении последовательности симметрических матриц $A^{(0)} = A$, $A^{(1)}$, $A^{(2)}, \ldots$, где матрицы

$$A^{(h)} = (a_{pq}^{(h)}) \quad \text{и} \quad A^{(h+1)} = (a_{pq}^{(h+1)})$$

связаны между собой соотношением

$$A^{(h+1)} = U_{i_h j_h}^T A^{(h)} U_{i_h j_h} .$$

Здесь $U_{i_h j_h}$ обозначает матрицу элементарного вращения в плоскости (i_h, j_h), причем ее коэффициенты определяются из условия $a_{i_h j_h}^{(h+1)} = 0$. Для разных способов выбора последовательности индексов (i_h, j_h) последовательность матриц $A^{(h)}$ сходится к диагональной матрице с собственными числами матрицы A на главной диагонали (см. например [14], [15], [17]).

Обозначим

(3.1) $$\tilde{A} = U_N^T A U_N, \quad \tilde{A} = (\tilde{a}_{pq}), \quad U_N = U_{i_0 j_0} U_{i_1 j_1} \ldots U_{i_{N-1} j_{N-1}},$$

(3.2) $$\tilde{p}_j = \Big(\sum_{\substack{i=1\\ i\neq j}}^{m} |\tilde{a}_{ji}|^2\Big)^{1/2}, \quad \tilde{q}_j = \sum_{\substack{i=1\\ i\neq j}}^{m} |\tilde{a}_{ji}| .$$

Пусть $\tilde{a}_{rr}$ обозначает k-ый коэффициент главной диагонали матрицы $\tilde{A}$, если они расположены в порядке убывания, т.е. если для коэффициентов $d_1, \ldots, d_m$ главной диагонали матрицы $\tilde{A}$ справедливо

$$d_1 \geqslant d_2 \geqslant \ldots \geqslant d_m ,$$

то пусть $d_k = \tilde{a}_{rr}$.

Далее, предположим, что

(1) промежуток $[\tilde{a}_{rr} - \tilde{q}_r, \tilde{a}_{rr} + \tilde{q}_r]$ не пересекается со всеми остальными промежутками $[\tilde{a}_{ii} - \tilde{q}_i, \tilde{a}_{ii} + \tilde{q}_i]$:

(3.3) $$[\tilde{a}_{rr} - \tilde{q}_r, \tilde{a}_{rr} + \tilde{q}_r] \cap [\tilde{a}_{ii} - \tilde{q}_i, \tilde{a}_{ii} + \tilde{q}_i] = \emptyset, \quad i \neq r,$$

(2) выполнено неравенство

(3.4) $$\tilde{p}_r < s_1 = \min_i \{|\tilde{a}_{rr} - \tilde{a}_{ii}| - \tilde{p}_r - \tilde{q}_i\} .$$

Обозначим $\varphi_0 = U_N e_r$, $e_r = (0, \ldots, \underset{r}{1}, \ldots, 0)^T \in \boldsymbol{R}^m$. Очевидно, мы имеем в силу ортогональности матрицы U_N равенство

$$\|\varphi_0\| = \|U_N e_r\| = \|e_r\| = 1 .$$

Далее, справедливо

$$\lambda_0 = (A\varphi_0, \varphi_0) = (U_N^T A U_N e_r, e_r) = (\tilde{A} e_r, e_r) = \tilde{a}_{rr} .$$

Тогда получим для нормы невязки $\sigma_0 = A\varphi_0 - \lambda_0 \varphi_0$ следующее соотношение:

$$\|\sigma_0\| = \|A\varphi_0 - \tilde{a}_{rr}\varphi_0\| = \|U_N^T(AU_N e_r - \tilde{a}_{rr} U_N e_r)\| = \|\tilde{A} e_r - \tilde{a}_{rr} e_r\| = \tilde{p}_r .$$

Наконец, рассмотрим матрицу $B_0 = \varphi_0 \sigma_0^T + \sigma_0 \varphi_0^T$ и $A_0 = \tilde{A} - B_0$.

В дальнейшем мы покажем, что при выполнении условий (3.3) и (3.4) имеют место предположения теоремы 9, если вектор φ_0 выбираем согласно формулы $\varphi_0 = U_{Nr} e_r$.

Теорема 10. *Пусть выполнены условия* (3.3) *и* (3.4). *Тогда спектр матрицы A как в промежутке $\Delta_0 = [\lambda_0 - \|\sigma_0\|, \lambda_0 + \|\sigma_0\|] = [\tilde{a}_{rr} - \tilde{p}_r, \tilde{a}_{rr} + \tilde{p}_r]$, так и в промежутке $\Delta^0 = [\lambda_0 - 2\|\sigma_0\|, \lambda_0 + 2\|\sigma_0\|] = [\tilde{a}_{rr} - 2\tilde{p}_r, \tilde{a}_{rr} + 2\tilde{p}_r]$ состоит из собственного числа Λ_k. Кроме того, имеет место неравенство*

$$\|\sigma_0\| < \varrho_0 = \min_{\substack{\nu \in \sigma(A_0)\\ \nu \neq \lambda_0}} |\nu - \lambda_0| .$$

Доказательство. Пусть через $d_1 \geqslant \ldots \geqslant d_{k-1} > d_k = \tilde{a}_{rr} > d_{k+1} \geqslant \ldots \geqslant d_m$ обозначены коэффициенты главной диагонали матрицы $\tilde{A}$, упорядоченные убывающим образом. Промежутки $[\tilde{a}_{ii} - \tilde{q}_i, \tilde{a}_{ii} + \tilde{q}_i]$ суть в силу определения величин $\tilde{q}_i$ (см. (3.2)) круги Гершгорина [5] для матрицы $\tilde{A}$ (см. например [15]). Совокупность кругов Гершгорина распадается в силу предположения (3.3) на три взаимно непересекающихся подобласти:

(1) сумма $k-1$ кругов, середины которых являются $d_1 \geqslant \ldots \geqslant d_{k-1}$, образует первую подобласть,

(2) промежуток $[\tilde{a}_{rr} - \tilde{q}_r, \tilde{a}_{rr} + \tilde{q}_r]$ образует вторую подобласть,

(3) сумма $m-k$ кругов, середины которых являются $d_{k+1} \geqslant \ldots \geqslant d_m$, образует третью подобласть.

Следовательно, в каждой подобласти находится ровно столько собственных чисел матрицы $\tilde{A}$, сколько кругов образуют данную подобласть (см. например [17]). Заметим, что собственные числа матриц A и $\tilde{A}$ тождествены. Каждое собственное число матрицы A, лежащее в первой подобласти, больше, чем то собственное число, которое находится во второй подобласти. Точно так же, каждое собственное число матрицы A, лежащее в третьей подобласти, меньше, чем то собственное число, которое лежит во второй подобласти. Так как первая подобласть содержит $k-1$ собственное число и третья подобласть $m-k$ собственных чисел матрицы A, то необходимо k-ое собственное число Λ_k лежит во второй подобласти и является единственной точкой спектра матрицы A в этой подобласти, т.е. имеем

$$\Lambda_k \in [\tilde{a}_{rr}-\tilde{q}_r, \tilde{a}_{rr}+\tilde{q}_r].$$

Далее, $\lambda_0 = \tilde{a}_{rr}$ является собственным числом матрицы A_0, так что матрица $\tilde{A} = A_0 + B_0$ имеет в промежутке

$$[\lambda_0 - \|B_0\|, \lambda_0 + \|B_0\|] = [\lambda_0 - \|\sigma_0\|, \lambda_0 + \|\sigma_0\|] = [\tilde{a}_{rr}-\tilde{p}_r, \tilde{a}_{rr}+\tilde{p}_r]$$

собственные числа. В силу очевидного неравенства $\tilde{p}_r \leqslant \tilde{q}_r$ отсюда вытекает, что Λ_k является единственной точкой спектра матрицы $\tilde{A}$ не только в промежутке $[\tilde{a}_{rr}-\tilde{q}_r, \tilde{a}_{rr}+\tilde{q}_r]$, но и в меньшем промежутке $[\tilde{a}_{rr}-\tilde{p}_r, \tilde{a}_{rr}+\tilde{p}_r] = \Delta_0$.

Тогда очевидно, что

$$\Lambda_k \in [\tilde{a}_{rr}-2\tilde{p}_r, \tilde{a}_{rr}+2\tilde{p}_r] = [\lambda_0 - 2\|\sigma_0\|, \lambda_0 + 2\|\sigma_0\|] = \Delta^0.$$

Кроме того, Λ_k является единственной точкой спектра матрицы A в промежутке Δ^0, так как в силу предположения (3.4) промежуток Δ^0 не пересекается ни с одним промежутком вида $[\tilde{a}_{ii}-\tilde{q}_i, \tilde{a}_{ii}+\tilde{q}_i]$, $i \neq r$, в которых лежат собственные числа Λ_j, $j \neq k$, матрицы A.

Наконец, оценим ϱ_0. Так как в сумме промежутков $[\tilde{a}_{ii}-\tilde{q}_i, \tilde{a}_{ii}+\tilde{q}_i]$, $i \neq r$, находится ровно $m-1$ собственное число матрицы $\tilde{A}$, то в сумме промежутков

$$\Delta^i = [\tilde{a}_{ii}-\tilde{q}_i - \|B_0\|, \tilde{a}_{ii}+\tilde{q}_i+\|B_0\|] = [\tilde{a}_{ii}-\tilde{q}_i-\tilde{p}_r, \tilde{a}_{ii}+\tilde{q}_i+\tilde{p}_r], \qquad i \neq r,$$

находится не меньше, чем $m-1$ собственное число матрицы $A_0 = \tilde{A} - B_0$ (см. теорему 1). Так как в силу предположения (3.4) расстояние от этих промежутков до собственного числа $\lambda_0 = \tilde{a}_{rr}$ матрицы A_0 больше, чем $\tilde{p}_r = \|\sigma_0\|$, то в сумме промежутков Δ^i, $i \neq r$, находится ровно $m-1$ собственное число матрицы A_0 и поэтому справедлива оценка

$$\varrho_0 = \min_{\substack{\nu \in \sigma(A_0) \\ \nu \neq \lambda_0}} |\lambda_0 - \nu| > \|\sigma_0\|. \ \blacksquare$$

Если рассматривается задача определения k-го собственного числа вещественной симметрической матрицы, то первый этап предлагаемого метода состоит в применении метода Якоби до тех пор, пока не выполнены условия (3.3) и (3.4). Эти условия легко проверяемы после каждого плоского вращения в методе Якоби. Однако, на практике оказалось, что, например, при применении циклического метода Якоби (см. [2], а также [14] и [17]) целесообразно

проверить условия (3.3) и (3.4) не после каждого элементарного вращения, а после прохождения одного полного цикла. В примерах, приведенных в работах [12] и [13], понадобилось не больше трех циклов. При этом оказалось, что условия (3.3) и (3.4) выполняются не одновременно, а — как правило — одно выполняется раньше другого. В работе [12] приведены наряду с условиями (3.3) и (3.4) еще другие, в которых входит вместе с постоянными $\tilde{p}_j$ и $\tilde{q}_j$ из формулы (3.2) еще постоянная

$$t(\tilde{A}) = \Big(\sum_{\substack{i,j=1 \\ i\neq j}}^{m} |\tilde{a}_{ij}|^2\Big)^{1/2}.$$

Заметим, что предположение $\Lambda_1 > \Lambda_2 > \ldots > \Lambda_m$ относительно собственных чисел Λ_j матрицы A можно ослабить в следующем виде:

$$\Lambda_1 \geqslant \ldots \geqslant \Lambda_{k-1} > \Lambda_k > \Lambda_{k+1} \geqslant \ldots \geqslant \Lambda_m,$$

т.е. искомое собственное число Λ_k должно быть простым и изолированным. Кроме того, общая кратность собственных чисел, больших Λ_k, должна быть равной $k-1$. Доказательство теоремы 10 при этом не изменится.

Наряду с методом, описанным выше, можно упомянуть метод Гивенса [6] и метод Хаусхолдера [7] для вычисления собственных чисел матрицы, в которых данная матрица переводится в трехдиагональную форму при помощи ортогональных матриц. После этого определяется характеристический полином этой трехдиагональной матрицы, который получается из конечной последовательности рекуррентных соотношений. Кроме того, из этих рекуррентных соотношений строится последовательность Штурма для характеристического полинома. После этого полувинием (бисекцией) промежутка легко можно получить произвольно малые промежутки, в которых находятся собственные числа матрицы.

Заметим, что методы Гивенса и Хаусхолдера связаны с вычислительными трудностями, если вместе с собственными числами требуется вычислить и собственные векторы. В этой связи часто ссылаются на метод Виландта [16], который похож на метод ложных возмущений, но который имеет меньшую скорость сходимости, чем метод ложных возмущений. В работе [11] обсуждается класс методов типа Ньютона для вычисления собственных чисел и собственных векторов симметрических матриц, к которому принадлежит описанный выше метод ложных возмущений. Доказывается, что метод ложных возмущений имеет наибольшую скорость сходимости среди всех методов этого класса. Приведены устойчивые схемы счета.

Литература

[1] H. Baumgärtel, *Endlichdimensionale analytische Störungstheorie*, Akademie-Verlag, Berlin 1972.

[2] G.E. Forsythe, P. Henrici, *The cyclic Jacobi method for computing the principal values of a complex matrix*, Trans. Amer. Math. Soc. 94 (1960), стр. 1–23.

[3] М. К. Гаврин, *О собственных числах операторов, зависящих от параметра*, Вестник Ленингр. ун-та, 9 (1952), стр. 77–95.
[4] —, *О методе ложных возмущений для разыскания собственных значений*, Журн. вычисл. матем. и матем. физики 1 (1961), стр. 757–770.
[5] S. Gerschgorin, *Abgrenzung der Eigenwerte einer Matrix*, Bull. Acad. Sci. USSR, Leningrad, classe math., 7 (1931), стр. 749–754.
[6] W. Givens, *Numerical computation of the characteristic values of a real symmetric matrix*, Oak Ridge Nat. Lab. Report ORNL-1574, 1954.
[7] A. S. Householder, *Principles of numerical analysis*, New York 1953.
[8] T. Kato, *Perturbations theory for linear operators*, Springer-Verlag, Berlin 1966.
[9] F. Kuhnert, *Die Pseudostöriteration zur Bestimmung von Eigenwerten und Eigenvektoren*, Deut. Verlag Wiss., Berlin 1971.
[10] —, *Die Pseudostöriteration bei abgeschlossenen Operatoren*, Rend. Circ. Mat. Palermo 16 (1967), стр. 5–32.
[11] —, *Über die Realisierung von iterativen Verfahren zur Eigenwertbestimmung*, Beitr. zur Numer. Math. 4 (1976), стр. 141–146.
[12] A. Meyer, *Die Konstruktion eines geeigneten Startvektors für die Pseudostöriteration mittels Jacobiverfahrens*, Diplomarbeit, Tech. Hochschule Karl-Marx-Stadt, 1974.
[13] A. Meyer, W. Lang, *Die Konstruktion eines geeigneten Startvektors für die Pseudostöriteration mit Hilfe des Jacobiverfahrens*, Beitr. zur Numer. Math. 5 (1976), стр. 119–128.
[14] H. R. Schwarz, H. Rutishauser, E. Stiefel, *Numerik symmetrischer Matrizen*, B. G. Teubner, Stuttgart 1968.
[15] G. W. Stewart, *Introduction to matrix computation*, Academic Press, New York–London 1973.
[16] H. Wielandt, *Bestimmung höherer Eigenwerte durch gebrochene Iteration*, Ber. d. aerodyn. Versuchsanstalt Göttingen 44/J/37, 1944.
[17] J. H. Wilkinson, *The algebraic eigenvalue problem*, Clarendon Press, Oxford 1965.

Presented to the Semester
Mathematical Models and Numerical Methods
(February 3–June 14, 1975)

BANACH CENTER PUBLICATIONS
VOLUME 3

МЕТОД ЛОЖНЫХ ВОЗМУЩЕНИЙ ДЛЯ НЕЛИНЕЙНЫХ ЗАДАЧ НА СОБСТВЕННЫЕ ЗНАЧЕНИЯ

РЕЙНАРД ЛЕМАНН

Технический ВУЗ, Карл-Маркс-Штадт, ГДР

1. Введение

Математические модели той или иной физической или технической проблемы нередко сводятся к нелинейной задаче на собственные значения (см. [3]), т.е. к задаче о нахождении такого параметра $\lambda = \Lambda$, для которого операторное уравнение

$$A(\lambda)x = 0 \tag{1}$$

в некотором гильбертовом пространстве H имеет хотя бы одно ненулевое решение $x = \varphi$. Здесь $A(\lambda)$ — линейный оператор в H, который нелинейно зависит от комплексного параметра $\lambda \in G \subseteq \boldsymbol{C}$. Всюду в дальнейшем предполагается, что эта зависимость либо имеет вид полинома относительно λ с операторными коэффициентами,

$$A(\lambda) \equiv \sum_{k=0}^{m} \lambda^{m-k} A_k,$$

либо операторная функция $A(\lambda)$ является дифференцируемой относительно λ достаточное число раз. Семейство операторов $A = \{A(\lambda), \lambda \in G \subseteq \boldsymbol{C}\}$ называется *операторным пучком*.

Число $\lambda = \Lambda$ называется *собственным значением* пучка A, если уравнение (1) при $\lambda = \Lambda$ имеет хотя бы одно ненулевое решение $x = \varphi$, называемое *собственным вектором* пучка A к собственному значению Λ, причём число $\dim\ker A(\Lambda)$ называется (геометрической) *кратностью* данного собственного значения Λ. Все остальные понятия спектральной теории линейных операторов, как спектр, резольвентное множество, резольвента пучка A, определяются, очевидно, аналогичным образом. В частности, оператор $R(\zeta)$ является резольвентой пучка A в точке ζ из резольвентного множества пучка A, если $R(\zeta)$ есть ограниченный обратный к оператору $A(\zeta)$.