Math-Net.Ru

Общероссийский математический портал

Ю. М. Нечепуренко, М. Садкане, О сходимости метода Ньютона–Канторовича для вычисления инвариантных подпространств, *Матем. заметки*, 2004, том 75, выпуск 1, 109–114

DOI: <https://doi.org/10.4213/mzm11>

Использование Общероссийского математического портала Math-Net.Ru подразумевает, что вы прочитали и согласны с пользовательским соглашением
<http://www.mathnet.ru/rus/agreement>

Параметры загрузки:

IP: 54.198.187.58

26 апреля 2023 г., 14:45:42

УДК 512.64

О СХОДИМОСТИ МЕТОДА НЬЮТОНА–КАНТОРОВИЧА ДЛЯ ВЫЧИСЛЕНИЯ ИНВАРИАНТНЫХ ПОДПРОСТРАНСТВ

Ю. М. Нечепуренко, М. Садкан

Предложен вариант метода Ньютона–Канторовича, позволяющий для заданной квадратной невырожденной матрицы и заданного числа ее минимальных по модулю собственных значений вычислять инвариантное подпространство, отвечающее этим собственным значениям. Получены оценки скорости сходимости через интегральный критерий качества круговой дихотомии.

Библиография: 4 названия.

1. Пусть $D \in \mathbb{C}^{n \times n}$ – невырожденная матрица и $\lambda_1, \dots, \lambda_n$ – ее собственные значения, упорядоченные так, что $|\lambda_1| \leq \dots \leq |\lambda_n|$. Пусть m – заданное целое число, удовлетворяющее следующим неравенствам: $1 \leq m < n$ и $|\lambda_m| < |\lambda_{m+1}|$; $\mathcal{Y}$ – инвариантное подпространство матрицы D , отвечающее $\lambda_1, \dots, \lambda_m$, и $Y \in \mathbb{C}^{n \times m}$ – матрица, столбцы которой образуют ортонормированный базис в подпространстве $\mathcal{Y}$: $\text{range } Y = \mathcal{Y}$, $Y^*Y = I'$.

Здесь и далее I' , I'' и I означают единичные матрицы порядков m , $n - m$ и n соответственно.

Матрица Y является решением системы матричных уравнений

$$D^{-1}Y - Y(Y^*D^{-1}Y) = 0, \quad Y^*Y = I'.$$

Для решения этой системы предлагается следующий вариант метода Ньютона–Канторовича:

$$\begin{aligned} B_j &= Y_j^* \Gamma_j, & \Delta_j &= \Gamma_j - Y_j B_j, \\ \Phi_j B_j - K(I - Y_j Y_j^*) \Phi_j &= K \Delta_j, & \Psi_{j+1} &= \Gamma_j + \Phi_j, \\ \Gamma_{j+1} &= \Psi_{j+1} (\Psi_{j+1}^* D^* \Psi_{j+1})^{-1/2}, & Y_{j+1} &= D \Gamma_{j+1}, \end{aligned} \quad (1)$$

возникший в наших частных беседах с С. К. Годуновым, как развитие техники вычисления инвариантных подпространств, предложенной в [1].

Здесь $\Gamma_0 \in \mathbb{C}^{n \times m}$ – заданная матрица такая, что матрица $Y_0 = D \Gamma_0$ унитарная прямоугольная, т.е. $Y_0^* Y_0 = I'$, и K – некоторая аппроксимация матрицы D^{-1} .

Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований, грант № 01-01-00269.

Используя подход, аналогичный предложенному в [2] для анализа методов ньютоновского типа, мы исследуем сходимость (1) и оцениваем тангенс

$$\tau_j = \operatorname{tg} \angle(\mathcal{Y}_j, \mathcal{Y}) \quad (2)$$

угла между искомым инвариантным подпространством $\mathcal{Y}$ и его аппроксимацией $\mathcal{Y}_j = \operatorname{range} Y_j$, где Y_j – матрица, вычисленная согласно (1). Итоговые оценки, полученные таким образом, включают только

$$\varepsilon = \|I - DK\|_2$$

и следующую характеристику матрицы D :

$$\eta(rD^{-1}) = r\|D^{-1}\|_2 \left\| \frac{1}{2\pi} \int_0^{2\pi} (re^{i\varphi}I - D)^{-1} * D^* D (re^{i\varphi}I - D)^{-1} d\varphi \right\|_2^2,$$

где r – любое положительное число между $|\lambda_m|$ и $|\lambda_{m+1}|$. Если D – дискретный аналог какого-либо дифференциального оператора $\mathcal{D}$, то эти оценки позволяют связать скорость сходимости метода (1) со свойствами оператора $\mathcal{D}$, которые обычно изучает теория дифференциальных операторов.

Следующая теорема показывает, что если подпространство $\mathcal{Y}_0$ достаточно близко к $\mathcal{Y}$, то сходимость линейная, если ε – достаточно малое положительное число, и квадратичная, если $\varepsilon = 0$.

ТЕОРЕМА 1. Пусть

$$11\eta(rD^{-1})(\varepsilon + \tau_0) \leq 1. \quad (3)$$

Тогда для любого целого $j \geq 0$ справедливы следующие неравенства:

$$\tau_{j+1} \leq 5, 44\eta(rD^{-1})(\varepsilon + \tau_j)\tau_j. \quad (4)$$

В этой теореме мы используем величину ε , как меру близости матрицы K к D^{-1} . В случае, когда эта величина не слишком мала, более точные оценки могут быть получены, если наряду с ней использовать величину

$$\varepsilon_0 = \|(I - YY^*)(I - DK)(I - YY^*)\|_2.$$

Следующая теорема устанавливает связь между тангенсом (2) и нормой спектральной невязки.

ТЕОРЕМА 2. В предположениях теоремы 1, следующие оценки справедливы при любом целом $j \geq 0$:

$$\frac{\tau_j}{(3, 47\eta(rD^{-1}))} \leq \frac{\|\Delta_j\|_2}{\|D^{-1}\|_2} \leq 2, 1\tau_j.$$

2. Рассмотрим случай $r = 1$. Будем предполагать, что единичная окружность не содержит собственных значений матрицы D и разбивает их множество на два непустых подмножества и m – сумма алгебраических кратностей собственных значений, лежащих в единичном диске. Для распространения полученных оценок на случай $r \neq 1$ достаточно заменить в них $\eta(D^{-1})$ на $\eta(rD^{-1})$.

Заметим прежде всего, что $\eta(G) = \|G\|_2 \kappa^2$, где κ означает интегральный критерий качества дихотомии спектра матрицы G единичной окружностью [3], [4]:

$$\kappa = \|H(G)\|_2, \quad H(G) = \frac{1}{2\pi} \int_0^{2\pi} (e^{i\varphi} I - G)^{-1*} (e^{i\varphi} I - G)^{-1} d\varphi.$$

Для получения оценок сходимости, основанных на интегральном критерии качества дихотомии, мы используем следующие две леммы, где G означает квадратную матрицу, не имеющую собственных значений на единичной окружности.

ЛЕММА 1. Пусть спектр матрицы G лежит либо внутри, либо вне единичного диска. Тогда для любого неотрицательного целого l выполняются соответственно следующие неравенства:

$$\|G^l\|_2 \leq \sqrt{\kappa} \left(1 - \frac{1}{\kappa}\right)^{l/2}, \quad \|G^{-l}\|_2 \leq \sqrt{\kappa + 1} \left(1 + \frac{1}{\kappa}\right)^{-l/2}.$$

ЛЕММА 2. Пусть E – квадратная матрица того же порядка, что и G , и спектр G лежит внутри (вне) единичного диска. Если

$$\left(\frac{\pi\kappa}{2} + \sqrt{\left(\frac{\pi\kappa}{2}\right)^2 + \kappa}\right) \|E\|_2 < 1,$$

то спектр суммы $G + E$ лежит строго внутри (вне) единичного диска.

Очевидно, достаточно доказать теоремы 1 и 2 для $j = 0$. Пусть U – унитарная матрица порядка n такая, что

$$\tilde{Y}_0 = U^* Y_0 = \begin{bmatrix} I' \\ 0 \end{bmatrix}.$$

Рассмотрим матрицу

$$\tilde{A} = U^* D^{-1} U = \begin{bmatrix} \tilde{A}_{11} & \tilde{A}_{12} \\ \tilde{A}_{21} & \tilde{A}_{22} \end{bmatrix},$$

где $\tilde{A}_{11}$, $\tilde{A}_{21}$, $\tilde{A}_{12}$ и $\tilde{A}_{22}$ – $m \times m$, $(n-m) \times m$, $m \times (n-m)$ и $(n-m) \times (n-m)$ матрицы соответственно.

ЛЕММА 3. Пусть $\tau_0 < \infty$. Тогда для матрицы $\tilde{A}$ справедливо следующее представление:

$$\tilde{A} = T C T^{-1}, \quad C = \begin{bmatrix} C_{11} & C_{12} \\ 0 & C_{22} \end{bmatrix}, \quad T = \begin{bmatrix} I' & 0 \\ X & I'' \end{bmatrix}, \quad (5)$$

где C_{11} и C_{22} – квадратные матрицы порядков m и $n - m$, спектры которых лежат строго вне и внутри единичного диска соответственно; C_{12} и X – $m \times (n - m)$ и $(n - m) \times m$ матрицы соответственно, и $\|X\|_2 = \tau_0$.

Введем следующие обозначения:

$$\tilde{K} = U^* K U = \begin{bmatrix} \tilde{K}_{11} & \tilde{K}_{12} \\ \tilde{K}_{21} & \tilde{K}_{22} \end{bmatrix}, \quad \tilde{E} = U^* (I - DK) U = \begin{bmatrix} \tilde{E}_{11} & \tilde{E}_{12} \\ \tilde{E}_{21} & \tilde{E}_{22} \end{bmatrix},$$

где разбиения на блоки такие же, как разбиение матрицы $\tilde{A}$.

Исключая Γ_j , B_j , Δ_j и Φ_j в (1) и полагая $j = 0$, получим следующее уравнение относительно Ψ_1 (уравнение Сильвестра):

$$\Psi_1 - K(I - Y_0 Y_0^*) \Psi_1 (Y_0^* D^{-1} Y_0)^{-1} = D^{-1} Y_0. \quad (6)$$

Используя леммы 2 и 3, можно показать, что при условии (3), спектры матриц $\tilde{A}_{11}$ и $\tilde{K}_{22}$ лежат строго вне и внутри единичного диска соответственно. Следовательно, спектры матриц $(Y_0^* D^{-1} Y_0)^{-1} = \tilde{A}_{11}^{-1}$ и

$$K(I - Y_0 Y_0^*) = U \begin{bmatrix} 0 & \tilde{K}_{12} \\ 0 & \tilde{K}_{22} \end{bmatrix} U^*$$

лежат строго внутри единичного диска. Таким образом уравнение (6) и, следовательно, система уравнений (1) однозначно разрешимы относительно Ψ_1 .

Умножая (6) слева на матрицу $U^* D$, мы получим уравнение, эквивалентное (6), которое можно записать следующим образом:

$$\tilde{\Psi}_1 - (I - \tilde{E})(I - \tilde{Y}_0 \tilde{Y}_0^*) \tilde{A} \tilde{\Psi}_1 (\tilde{Y}_0^* \tilde{A} \tilde{Y}_0)^{-1} = \tilde{Y}_0,$$

и, следовательно,

$$\begin{aligned} \tilde{\Psi}_{11} + \tilde{E}_{12}(\tilde{A}_{21} \tilde{\Psi}_{11} + \tilde{A}_{22} \tilde{\Psi}_{21}) \tilde{A}_{11}^{-1} &= I', \\ \tilde{\Psi}_{21} - (I'' - \tilde{E}_{22})(\tilde{A}_{21} \tilde{\Psi}_{11} + \tilde{A}_{22} \tilde{\Psi}_{21}) \tilde{A}_{11}^{-1} &= 0, \end{aligned} \quad (7)$$

где

$$\tilde{\Psi}_1 = U^* D \Psi_1 = \begin{bmatrix} \tilde{\Psi}_{11} \\ \tilde{\Psi}_{21} \end{bmatrix},$$

а $\tilde{\Psi}_{11}$ и $\tilde{\Psi}_{21} - m \times m$ и $(n - m) \times m$ матрицы соответственно.

Из неравенства (3) следует, что $\varepsilon < 1$. Следовательно, $\|\tilde{E}_{22}\|_2 < 1$. Учитывая, что матрица $I'' - \tilde{E}_{22}$ невырожденная, нетрудно показать, что матрица $\tilde{\Psi}_{11}$ невырожденная и $\tilde{\Psi}_{11}$ и $\tilde{\Psi}_{21} = Z \tilde{\Psi}_{11}$ удовлетворяют уравнению (7), если $\tilde{\Psi}_{11}$ и Z удовлетворяют уравнениям

$$(I' + \tilde{E}_{12}(I'' - \tilde{E}_{22})^{-1} Z) \tilde{\Psi}_{11} = I', \quad (8)$$

$$Z \tilde{A}_{11} - \tilde{A}_{22} Z = \tilde{A}_{21} - \tilde{E}_{22}(\tilde{A}_{21} + \tilde{A}_{22} Z) + Z(\tilde{A}_{11} - \tilde{\Psi}_{11} \tilde{A}_{11}^{-1} \tilde{\Psi}_{11}^{-1}). \quad (9)$$

Пусть справедливо равенство (8). Используя равенства

$$\begin{aligned}\tilde{A}_{12} &= C_{12}, \quad \tilde{A}_{11} = C_{11} - \tilde{A}_{12}X, \quad \tilde{A}_{22} = C_{22} + X\tilde{A}_{12}, \\ X\tilde{A}_{11} - \tilde{A}_{22}X + X\tilde{A}_{12}X &= \tilde{A}_{21},\end{aligned}\tag{10}$$

следующие из (5), мы можем показать, что следующее уравнение относительно матрицы $W = Z - X$ эквивалентно (9):

$$W - C_{22}WC_{11}^{-1} = FC_{11}^{-1} - \tilde{E}_{22}X,\tag{11}$$

где

$$\begin{aligned}F &= W\tilde{A}_{12}X + X\tilde{A}_{12}W + X\tilde{A}_{12}X - \tilde{E}_{22}\tilde{A}_{22}W \\ &\quad + (W + X)(I' + S)^{-1}(S\tilde{A}_{11} - \tilde{A}_{11}S), \\ S &= \tilde{E}_{12}(I'' - \tilde{E}_{22})^{-1}(W + X).\end{aligned}$$

Поскольку спектры матриц C_{11}^{-1} и C_{22} лежат строго внутри единичного диска, линейный оператор

$$\mathcal{L}(W) = W - C_{22}WC_{11}^{-1} : \mathbb{C}^{(n-m) \times m} \rightarrow \mathbb{C}^{(n-m) \times m}$$

невырожденный и

$$\mathcal{L}^{-1}(V) = \sum_{l=0}^{\infty} C_{22}^l V C_{11}^{-l}.$$

Поэтому, уравнение (11) и, следовательно, (9) эквивалентны уравнению

$$W = \mathcal{N}(W),\tag{12}$$

где

$$\mathcal{N}(W) = \sum_{l=0}^{\infty} C_{22}^l (FC_{11}^{-l-1} - \tilde{E}_{22}XC_{11}^{-l}).$$

Итак, при условии (3), имеем: (6) $\Leftrightarrow$ (7) $\Leftrightarrow$ (8) + (9) и (9) $\Leftrightarrow$ (11) $\Leftrightarrow$ (12). Пусть W является решением (12) таким, что $\|W\|_2 \leq \|X\|_2 = \tau_0$. Тогда $Z = W + X$ и $\tilde{\Psi}_{11} = (I' + S)^{-1}$ удовлетворяют (8) и (9) и, следовательно, матрица $\tilde{\Psi}_{11}$ невырождена и $\tilde{\Psi}_{11}$ и $\tilde{\Psi}_{21} = Z\tilde{\Psi}_{11}$ удовлетворяют (7). При условии (3) уравнение (7) однозначно разрешимо и, следовательно, матрица полного ранга

$$\Psi_1 = D^{-1}U \begin{bmatrix} \tilde{\Psi}_{11} \\ \tilde{\Psi}_{21} \end{bmatrix}$$

является единственным решением системы уравнений (1) при $j = 0$. Таким образом, система (1) при $j = 0$ однозначно разрешима относительно Γ_1 .

Из леммы 3 следует, что

$$U^* \mathcal{Y} = \text{range} \begin{bmatrix} I' \\ X \end{bmatrix}.$$

С другой стороны,

$$U^* \mathcal{Y}_1 = \text{range } \tilde{\Psi}_1 = \text{range } \begin{bmatrix} I' \\ Z \end{bmatrix}.$$

Поэтому [2]

$$\tau_1 = \text{tg } \angle(U^* \mathcal{Y}_1, U^* \mathcal{Y}) \leq \frac{\|W\|_2}{1 - \tau_0 \|W\|_2}.$$

Для получения оценки (4) при $j = 0$ и $r = 1$ осталось показать, используя принцип сжатых отображений, что (12) имеет решение W такое, что $\|W\|_2 \leq \|X\|_2 = \tau_0$, и оценить норму этого решения, используя уравнение (12), лемму 1 и следующие неравенства

$$\kappa(C_{ll}) \leq \left(1 + \frac{\tau_0^2}{2} + \sqrt{\left(1 + \frac{\tau_0^2}{2}\right)^2 - 1}\right)^2 \kappa(D^{-1}), \quad l = 1, 2,$$

которые следуют из леммы 3.

3. Учитывая, что

$$U^* \Delta_0 = \tilde{A} \tilde{Y}_0 - \tilde{Y}_0 \tilde{A}_{11} = \begin{bmatrix} 0 \\ \tilde{A}_{21} \end{bmatrix}$$

и, следовательно, $\|\Delta_0\|_2 = \|\tilde{A}_{21}\|_2$, и используя четвертое уравнение (10), можно получить верхнюю оценку из теоремы 2.

Используя второе и третье равенства (10) мы можем записать четвертое равенство (10) следующим образом:

$$X C_{11} - C_{22} X = \tilde{A}_{21} + X \tilde{A}_{12} X.$$

Поэтому,

$$X = \sum_{l=0}^{\infty} C_{22}^l (\tilde{A}_{21} + X \tilde{A}_{12} X) C_{11}^{-l-1},$$

откуда следует нижняя оценка.

Мы благодарны академику С.К. Годунову за очень полезное обсуждение результатов данной работы.

СПИСОК ЦИТИРОВАННОЙ ЛИТЕРАТУРЫ

- [1] Godunov S.K., Sadkane M., Robbe M. Smoothing techniques and computation of invariant subspaces of elliptic operators // Appl. Numer. Math. 2000. V. 33. P. 341–347.
- [2] Годунов С.К., Нечепуренко Ю.М. Оценки скорости сходимости метода Ньютона для вычисления инвариантных подпространств // ЖВМ и МФ. 2002. Т. 42. № 6. С. 771–779.
- [3] Булгаков А.Я., Годунов С.К. Круговая дихотомия матричного спектра // Сиб. матем. ж. 1988. Т. 29. № 5. С. 59–70.
- [4] Годунов С.К. Современные аспекты линейной алгебры. Новосибирск: Научная Книга, 1997.

(Ю.М. Нечепуренко) Институт вычислительной математики РАН, г. Москва
E-mail: yumn@inm.ras.ru
 (М. Садкан) Université de Bretagne Occidentale
E-mail: sadkane@univ-brest.fr

Поступило
 24.10 2002