



**Ордена Ленина
ИНСТИТУТ ПРИКЛАДНОЙ МАТЕМАТИКИ
имени М.В. Келдыша
Российской академии наук**

С.Ю. Медведев, Д.О. Чеботарев

**Метод конечных элементов с
базисными функциями
специального вида для задач
на собственные значения**

Препринт №

Москва

Российская Академия Наук
ОРДЕНА ЛЕНИНА ИНСТИТУТ ПРИКЛАДНОЙ МАТЕМАТИКИ
им. М.В. Келдыша

С.Ю. Медведев, Д.О. Чеботарев

МЕТОД КОНЕЧНЫХ ЭЛЕМЕНТОВ С БАЗИСНЫМИ ФУНКЦИЯМИ
СПЕЦИАЛЬНОГО ВИДА ДЛЯ ЗАДАЧ НА СОБСТВЕННЫЕ ЗНАЧЕНИЯ

Москва, 2005 г.

С.Ю. Медведев, Д.О. Чеботарев

Метод конечных элементов с базисными функциями специального вида для задач на собственные значения

Аннотация

В данной работе исследуется возможность применения метода residual-free bubbles (RFB) к решению задач на собственные значения. Рассматриваются задачи для эллиптического оператора в одномерной и двумерной области. Приводится доказательство того, что указанный метод эквивалентен методу конечных элементов с базисными функциями, которые выбраны специальным образом, т.е. можно выбрать базисные функции в методе конечных элементов таким образом, что получается та же разностная задача, что и в методе RFB. Рассматриваются методы решения матричных задач на собственные значения с нелинейным вхождением собственного числа, возникающих в результате применения метода RFB для задач на собственные значения.

S.Yu. Medvedev, D.O. Chebotarev

Finite elements method with special basic functions for eigenvalue problems

Abstract

In the present paper the application of the residual-free bubbles (RFB) method to eigenvalue problems is considered. Problems for the elliptic operator in one- and two-dimensional domains are considered. The proof that the RFB method is equivalent to the of finite elements method with special basic functions is presented, i.e. it is possible to choose basic functions in the finite elements method in such a manner that the same discrete problem appears, as in the RFB method. Methods of the solution of nonlinear matrix eigenvalue problems, which appear as a result of the RFB method application to eigenvalue problems are considered.

Оглавление:

I. ВВЕДЕНИЕ.	4
II. ПОСТАНОВКА ЗАДАЧИ И МЕТОД RFB ДЛЯ ЗАДАЧ НА СОБСТВЕННЫЕ ЗНАЧЕНИЯ.	4
III. БАЗИСНЫЕ ФУНКЦИИ СПЕЦИАЛЬНОГО ВИДА.	6
IV. ОДНОМЕРНЫЙ СЛУЧАЙ.	8
IV.1. Простейшая задача на собственные значения для оператора второй производной с постоянным коэффициентом.	8
IV.2. Задача на собственные значения для оператора второй производной с переменным коэффициентом.	12
V. ДВУМЕРНЫЙ СЛУЧАЙ.	17
V.1. Общая теория.	17
V.2. Задача для оператора Лапласа в квадрате	19
V.3. Задача для оператора Лапласа в L-образной области	21
VI. МЕТОДЫ РЕШЕНИЯ МАТРИЧНЫХ ЗАДАЧ НА СОБСТВЕННЫЕ ЗНАЧЕНИЯ С НЕЛИНЕЙНЫМ ВХОЖДЕНИЕМ СОБСТВЕННОГО ЧИСЛА.	23
VI.1. Метод обратных итераций для матричных задач на собственные значения с нелинейным входением собственного числа.	23
VI.2. Метод Дэвиса.	24
VI.3. Комбинация двух методов.	26
VII. ЗАКЛЮЧЕНИЕ.	27
СПИСОК ЛИТЕРАТУРЫ	28

I. Введение.

Целью данной работы является применение метода, известного в англоязычной литературе как RFB (residual-free bubbles)[1,4,5] к решению задач на собственные значения. Рассматриваются задачи для эллиптического оператора в одномерной и двумерной области. Суть метода RFB для задач на собственные значения излагается в главе 0.

Важной частью работы является доказательство того, что указанный метод эквивалентен методу конечных элементов с базисными функциями, которые выбраны специальным образом, т.е. можно выбрать базисные функции в методе конечных элементов таким образом, что получается та же разностная задача, что и в методе RFB. Именно этим фактом объясняется название работы. Этому вопросу посвящена глава III. В главе IV приводятся два примера применения описываемого метода в одномерном случае. И, наконец, в главе V рассматриваются двумерные задачи для оператора Лапласа.

Отметим, что в результате применения метода RFB для задач на собственные значения (в отличие от краевых задач) возникает матричная задача на собственные значения с нелинейным вхождением собственного числа. Поэтому встает вопрос о методах решения подобных задач. В качестве таких методов в данной работе выбраны модифицированный метод обратных итераций и метод Дэвиса, описываемые в главе VI. Кроме того, рассматривается вопрос применения комбинации указанных методов.

II. Постановка задачи и метод RFB для задач на собственные значения.

Рассмотрим абстрактную задачу на собственные значения в области Ω :

$$\begin{cases} Lu = \lambda u & \text{на } \Omega \\ u = 0 & \text{на } \partial\Omega \end{cases} \quad (1)$$

где L – некоторый эллиптический оператор, $u \in C^2(\Omega)$, $u \in C(\bar{\Omega})$, а Ω – двумерная или одномерная область.

Ей соответствует обобщенная задача:

$$\begin{aligned} &\text{Найти } u \in H_0^1(\Omega): \\ &a(u, v) = \lambda b(u, v), \forall v \in H_0^1(\Omega) \end{aligned} \quad (2)$$

Здесь $a(u, v)$ и $b(u, v)$ – билинейные формы. Заметим, что $b(u, v) = (u, v)$, а если u – классическое решение задачи (1), то билинейная форма $a(u, v)$ имеет вид:

$$a(u, v) = (Lu, v) \quad (3)$$

где $(\cdot, \cdot)$ – скалярное произведение в $L_2(\Omega)$.

Пусть $\{K\}$ – разбиение области Ω на подобласти.

Классический метод конечных элементов заключается в замене исходного Гильбертова пространства в обобщенной задаче на его конечномерное подпространство $V_h = V_1 \subset H_0^1(\Omega)$. В качестве пространства V_h обычно выбирается пространство

непрерывных кусочно-полиномиальных функций, обычно – кусочно-линейных. Тогда обобщенная задача (2) будет выглядеть следующим образом:

$$\begin{aligned} & \text{Найти } u_h \in V_h : \\ & a(u_h, v_h) = \lambda b(u_h, v_h), \forall v_h \in V_h \end{aligned} \quad (4)$$

При этом основная идея метода RFB состоит в расширении пространства кусочно-полиномиальных функций специальными функциями, которые обращаются в ноль на границе элементов сетки и построены таким образом, что получаемое решение удовлетворяет дифференциальной задаче внутри каждого элемента сетки (bubble functions). Идея подобного подхода тесно связана с так называемыми стабилизированными методами (stabilized methods). [1,4,5] Стабилизированные методы возникли в связи с тем, что применение традиционного метода конечных элементов к некоторым задачам, особенно задачам конвекции-диффузии не давало удовлетворительных результатов. При этом стабилизированные методы дают повышение точности в таких задачах за счет устранения осцилляций решения. Метод RFB возник в результате попытки подвести теоретическую основу под идеи стабилизированных методов и является одним из практических способов построения стабилизированных методов. Кроме того, существует другой способ построения стабилизированных методов, связанный с использованием функций Грина для дифференциальной задачи на каждом элементе сетки. Однако было доказано, что эти два способа получения стабилизированных методов эквивалентны[1].

В методе RFB [1,3,4,5] рассматривается пространство $V_h \subset H_0^1(\Omega)$, которое представляется как сумма подпространств $V_h = V_1 \oplus V_b$, где V_1 - пространство непрерывных кусочно-линейных функций, линейных на каждом элементе сетки, а V_b - некоторое, возможно бесконечномерное, замкнутое пространство, состоящее из функций, обращающихся в ноль на границе элементов сетки. Таким образом, по сравнению с традиционным методом производится расширение пространства, в котором ставится обобщенная задача. Здесь важно, что V_1 - пространство именно кусочно-линейных, а не кусочно-полиномиальных функций более высокого порядка. Многие факты, приводимые далее в общем случае для кусочно-полиномиальных функций неверны.

Примечание 1: В частности V_b может включать в себя все функции, обращающиеся в ноль на границе элементов сетки. При этом в одномерном случае $V_h = V_1 \oplus V_b = H_0^1(\Omega)$ (в пространствах большей размерности это неверно). Кроме того, в одномерном случае $u_h = u_1 + u_b \in V_1 \oplus V_b$ фактически будет представлять собой *точное* решение задачи (4), а u_1 будет линейным интерполянтом u_h . Однако в этом случае полученная задача (4) будет совпадать с исходной.

Из разложения $V_h = V_1 \oplus V_b$ следует, что любая функция u_h из V_h однозначно представляется в виде суммы:

$$u_h = u_1 + u_b, \quad u_1 \in V_1, \quad u_b \in V_b \quad (5)$$

Тогда задачу (4) можно переписать в следующем виде:

Найти $u_h = u_1 + u_b \in V_1 \oplus V_b$:

$$\begin{cases} a(u_1 + u_b, v_1) = \lambda b(u_1 + u_b, v_1), \forall v_1 \in V_1 \\ a(u_1 + u_b, v_b) = \lambda b(u_1 + u_b, v_b), \forall v_b \in V_b \end{cases} \quad (6)$$

Учитывая, что любая функция из V_b обращается в ноль на границе элементов сетки, V_b можно представить в виде:

$$V_b = \oplus_K V_b(K) \quad (7)$$

Здесь $\oplus_K$ прямая сумма по всем элементам сетки K , а $V_b(K)$ - это сужение пространства V_b на элементе K .

Тогда любая функция из V_b однозначно представляется в виде суммы:

$$u_b = \sum_K u_{b,K} \quad (8)$$

Следовательно, второе уравнение системы (6) можно записать отдельно для каждого элемента сетки, и система принимает вид:

$$\begin{aligned} & \text{Найти } u_h = u_1 + u_b = u_1 + \sum_K u_{b,K} \in V_1 \oplus V_b : \\ & \begin{cases} a(u_1 + u_b, v_1) = \lambda b(u_1 + u_b, v_1), \forall v_1 \in V_1 \\ a(u_1 + u_{b,K}, v_{b,K})_K = \lambda b(u_1 + u_{b,K}, v_{b,K})_K, \forall K, \forall v_{b,K} \in V_b(K) \end{cases} \end{aligned} \quad (9)$$

где индекс K у билинейных форм означает интегрирование по K -му элементу сетки.

Далее, согласно методу RFB, из второго уравнения системы (9) аналитически выражается $u_{b,K}$, при этом λ рассматривается как параметр, и фактически $u_{b,K}$ ищется отдельно на каждом элементе сетки K как решение *краевой* задачи, а не задачи на собственные значения, что значительно проще. Если же решить второе уравнение системы (9) не удастся, то можно рассмотреть более простую задачу, например полученную заменой переменных коэффициентов, входящих в оператор L их кусочно-постоянным приближением.

III. Базисные функции специального вида.

Обратимся ко второму уравнению системы (9):

$$a(u_1 + u_{b,K}, v_{b,K})_K = \lambda b(u_1 + u_{b,K}, v_{b,K})_K, \forall K, \forall v_{b,K} \in V_b(K) \quad (10)$$

Используя линейность форм $a(u, v)$ и $b(u, v)$ по первому аргументу, можно записать:

$$a(u_{b,K}, v_{b,K})_K - \lambda b(u_{b,K}, v_{b,K})_K = -(a(u_1, v_{b,K})_K - \lambda b(u_1, v_{b,K})_K), \quad \forall K, \forall v_{b,K} \in V_b(K) \quad (11)$$

Заметим, что функция u_1 является линейной на любом элементе сетки K . Следовательно, учитывая (3), а также то, что $v_{b,K} = 0$ на границе элемента, билинейную форму $a(u_1, v_{b,K})$ можно записать в виде $(Lu_1, v_{b,K})$. Далее, учитывая, что $b(u, v) = (u, v)$, получаем следующее уравнение:

$$a(u_{b,K}, v_{b,K}) - \lambda(u_{b,K}, v_{b,K}) = -(Lu_1 - \lambda u_1, v_{b,K}), \forall K \quad (12)$$

Теперь $u_{b,K}$ может быть представлено как решение этой задачи, зависящее от u_1 и λ , для любого $v_{b,K} \in V_b(K)$:

$$u_{b,K} = M_K(\lambda)u_1 \quad (13)$$

где $M_K(\lambda): V_1 \rightarrow V_b(K)$ – некоторый линейный ограниченный оператор, зависящий от параметра λ , действующий из пространства V_1 в пространство $V_b(K)$ и определенный для каждого элемента сетки $K[1]$.

Обратимся теперь к первому уравнению системы (9), учитывая (8):

$$a(u_1 + \sum_K u_{b,K}, v_1) = \lambda b(u_1 + \sum_K u_{b,K}, v_1), \forall v_1 \in V_1 \quad (14)$$

Далее уравнение (10) можно записать в следующем виде:

$$a(u_1 + \sum_K u_{b,K}, \sum_K v_{b,K}) = \lambda b(u_1 + \sum_K u_{b,K}, \sum_K v_{b,K}), \forall v_{b,K} \in V_b(K) \quad (15)$$

Уравнения (10) и (15) эквивалентны, так как в суммах $\sum_K u_{b,K}$ и $\sum_K v_{b,K}$ на каждом элементе сетки K , остается лишь одно слагаемое, соответствующее данному элементу сетки.

Складывая (14) и (15), получим:

$$\begin{aligned} a(u_1 + \sum_K u_{b,K}, v_1 + \sum_K v_{b,K}) &= \lambda b(u_1 + \sum_K u_{b,K}, v_1 + \sum_K v_{b,K}), \\ \forall v &= v_1 + \sum_K v_{b,K} \in V_1 \oplus V_b \end{aligned} \quad (16)$$

И, наконец, учтем равенства $u_{b,K} = M_K(\lambda)u_1$, а $v_{b,K} = M_K(\lambda)v_1$. Подставляя их в равенство (16), получим:

$$\begin{aligned} a(u_1 + \sum_K M_K(\lambda)u_1, v_1 + \sum_K M_K(\lambda)v_1) &= \\ = \lambda b(u_1 + \sum_K M_K(\lambda)u_1, v_1 + \sum_K M_K(\lambda)v_1), \forall v_1 \in V_1 \end{aligned} \quad (17)$$

Обозначим $\tilde{u}_h = u_1 + \sum_K u_{b,K} = u_1 + \sum_K M_K(\lambda)u_1$ для любого $u_1 \in V_1$. Обозначим пространство всех $\tilde{u}_h$ через $\tilde{V}_h$. Отметим, что пространство $\tilde{V}_h$ является линейным, так как пространства V_1 и V_b – линейны.

Используя новые обозначения, из уравнения (17) получаем следующую обобщенную задачу:

$$\begin{aligned} \text{Найти } \tilde{u}_h \in \tilde{V}_h : \\ a(\tilde{u}_h, \tilde{v}_h) = \lambda b(\tilde{u}_h, \tilde{v}_h), \forall \tilde{v}_h \in \tilde{V}_h \end{aligned} \quad (18)$$

Таким образом, мы получили обобщенную задачу, которая соответствует методу конечных элементов с модифицированными базисными функциями $\tilde{u}_h \in \tilde{V}_h$. Важно

отметить, что пространство $\tilde{V}_h$ конечномерно (это следует из конечномерности пространства V_1). Следовательно, задача (6), которая была получена из традиционного метода конечных элементов расширением исходного пространства пространством V_b (возможно бесконечномерным), эквивалентна задаче (18), полученной *модификацией* исходного пространства *без увеличения размерности*. Имеется в виду, что разностные задачи, получаемые указанными двумя способами совпадают.

Будем считать, что функции $u_{b,K}$ достаточно гладкие и на них можно подействовать оператором L . Тогда (12) можно записать в следующем виде:

$$(Lu_{b,K} - \lambda u_{b,K}, v_{b,K}) = -(Lu_1 - \lambda u_1, v_{b,K}), \forall K \quad (19)$$

Так как последнее равенство верно для любого $v_{b,K}$, то можно перейти к операторному уравнению:

$$Lu_{b,K} - \lambda u_{b,K} = -(Lu_1 - \lambda u_1) \quad (20)$$

Пусть размерность пространства V_1 равняется N . Введем в нем базис $\{\psi_i\}_{i=1}^N$. Тогда на каждом элементе сетки K базисные функции пространства $\tilde{V}_h$ представляются в виде $\tilde{\varphi}_i = \psi_i + M_K(\lambda)\psi_i$. Подставляя это разложение в (20), и учитывая что $M_K(\lambda)\psi_i$ обращается в ноль на границе элементов сетки, получаем следующую краевую задачу на каждом элементе сетки K :

$$\begin{cases} L\tilde{\varphi}_i = \lambda\tilde{\varphi}_i, & \text{на } K \\ \tilde{\varphi}_i = \psi_i, & \text{на } \partial K \end{cases} \quad (21)$$

Таким образом, базисные функции пространства $\tilde{V}_h$ в задаче (18) могут быть получены как решение этой краевой задачи. Это дает практический способ получения базисных функций $\tilde{\varphi}_i$. При этом в качестве кусочно-линейных базисных функций ψ_i выбираются так называемые “функции-домики”. Еще раз отметим, что описанный способ получения базисных функций $\tilde{\varphi}_i$ может использоваться, только если функции из $V_b(K)$ достаточно гладкие.

IV. Одномерный случай

IV.1. Простейшая задача на собственные значения для оператора второй производной с постоянным коэффициентом.

В качестве первого примера рассмотрим следующую простейшую краевую задачу на собственные значения:

$$\begin{cases} -u'' = \lambda u, x \in [0,1], \lambda > 0 \\ u(0) = u(1) = 0 \end{cases} \quad (22)$$

Введем равномерную сетку: $\omega_h = \{x_i = i/N, i = 0..N\}$. Определим финитные функции $\tilde{\varphi}_i(x)$, отличные от нуля на отрезке $[x_{i-1}, x_{i+1}]$, в соответствии с (21) как решение краевых задач на отрезках $[x_{i-1}, x_i]$ и $[x_i, x_{i+1}]$:

$$\begin{cases} -\tilde{\varphi}_i'' = \lambda \tilde{\varphi}_i, & x \in [x_{i-1}, x_i] \\ \tilde{\varphi}_i(x_i) = 1 \\ \tilde{\varphi}_i(x_{i-1}) = 0 \end{cases} \quad (23)$$

$$\begin{cases} -\tilde{\varphi}_i'' = \lambda \tilde{\varphi}_i, & x \in [x_i, x_{i+1}] \\ \tilde{\varphi}_i(x_i) = 1 \\ \tilde{\varphi}_i(x_{i+1}) = 0 \end{cases}$$

Нетрудно убедиться, что функции, удовлетворяющие системам (23) для данной задачи, будут иметь вид:

$$\tilde{\varphi}_i(x) = \begin{cases} -\frac{\sin(\sqrt{\lambda}(x_{i-1} - x))}{\sin(\sqrt{\lambda}h)}, & x \in [x_{i-1}, x_i] \\ \frac{\sin(\sqrt{\lambda}(x_{i+1} - x))}{\sin(\sqrt{\lambda}h)}, & x \in [x_i, x_{i+1}] \end{cases} \quad (24)$$

$$i = 1..N-1$$

На рис. 1 и 2 приведены графики функции $\tilde{\varphi}_i(x)$ при $N=10$ для собственных значений $\lambda = (3\pi)^2$ и $\lambda = (5\pi)^2$ соответственно.

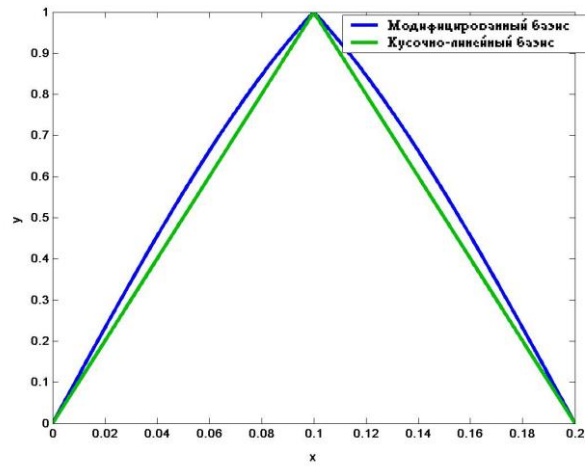


Рисунок 1

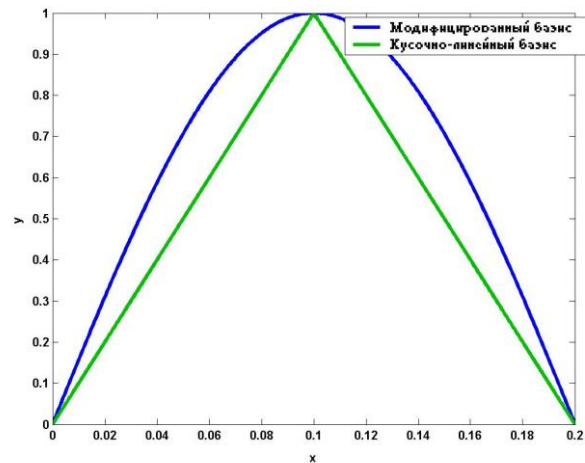


Рисунок 2

Рассмотрим теперь обобщенную постановку задачи (22). Пусть $\tilde{V}_h$ -конечномерное пространство функций с базисом $\tilde{\varphi}_i, i = 1..N-1$. Будем искать приближенное решение в виде разложения по базису $\tilde{V}_h$:

$$\tilde{u}_h(x) = \sum_{i=1}^{N-1} c_i \tilde{\varphi}_i(x) \quad (25)$$

Тогда обобщенная постановка задачи (22) выглядит следующим образом:

$$a(\tilde{u}_h, \tilde{v}_h) = \lambda b(\tilde{u}_h, \tilde{v}_h) \quad \forall \tilde{v}_h \in \tilde{V}_h \quad (26)$$

где билинейные формы $a(u, v)$ и $b(u, v)$ имеют вид:

$$\begin{aligned} a(u, v) &= \int_0^1 u'v' dx \\ b(u, v) &= \int_0^1 uv dx \end{aligned} \quad (27)$$

Подставляя (25) в (26), получим:

$$\sum_{i=1}^{N-1} c_i a(\tilde{\varphi}_i, \tilde{\varphi}_j) = \lambda \sum_{i=1}^{N-1} c_i b(\tilde{\varphi}_i, \tilde{\varphi}_j) \quad \forall j = 1..N-1 \quad (28)$$

Получаем матричную задачу на собственные значения с трехдиагональной матрицей. Учитывая, что базисные функции $\tilde{\varphi}_i$ по построению являются точным решением уравнения $-u'' = \lambda u$ на соответствующих отрезках сетки, можно найти коэффициенты системы следующим образом:

$$\begin{aligned} a(\tilde{\varphi}_i, \tilde{\varphi}_i) - \lambda b(\tilde{\varphi}_i, \tilde{\varphi}_i) &= \int_{x_{i-1}}^{x_i} (\tilde{\varphi}_i' \tilde{\varphi}_i' - \lambda \tilde{\varphi}_i \tilde{\varphi}_i) dx + \int_{x_i}^{x_{i+1}} (\tilde{\varphi}_i' \tilde{\varphi}_i' - \lambda \tilde{\varphi}_i \tilde{\varphi}_i) dx = \\ &= \tilde{\varphi}_i' \tilde{\varphi}_i \Big|_{x_{i-1}}^{x_i} + \tilde{\varphi}_i' \tilde{\varphi}_i \Big|_{x_i}^{x_{i+1}} = \tilde{\varphi}_i' \Big|_{x_i-0} - \tilde{\varphi}_i' \Big|_{x_i+0} \end{aligned} \quad (29)$$

Аналогично:

$$\begin{aligned} a(\tilde{\varphi}_{i+1}, \tilde{\varphi}_i) - \lambda b(\tilde{\varphi}_{i+1}, \tilde{\varphi}_i) &= -\tilde{\varphi}_{i+1}' \Big|_{x_i} \\ a(\tilde{\varphi}_{i-1}, \tilde{\varphi}_i) - \lambda b(\tilde{\varphi}_{i-1}, \tilde{\varphi}_i) &= \tilde{\varphi}_{i-1}' \Big|_{x_i} \end{aligned} \quad (30)$$

Таким образом, находя производные базисных функций $\tilde{\varphi}_i$:

$$\tilde{\varphi}_i'(x) = \begin{cases} \sqrt{\lambda} \frac{\cos(\sqrt{\lambda}(x_{i-1} - x))}{\sin(\sqrt{\lambda}h)}, & x \in [x_{i-1}, x_i] \\ -\sqrt{\lambda} \frac{\cos(\sqrt{\lambda}(x_{i+1} - x))}{\sin(\sqrt{\lambda}h)}, & x \in [x_i, x_{i+1}] \end{cases} \quad (31)$$

$$i = 1..N-1$$

мы получаем следующую нелинейную задачу на собственные значения:

$$A(\lambda)x = 0 \quad (32)$$

где матрица $A(\lambda)$ имеет вид:

$$A(\lambda) = \begin{pmatrix} \frac{2\sqrt{\lambda} \cos(\sqrt{\lambda}h)}{\sin(\sqrt{\lambda}h)} - \frac{\sqrt{\lambda}}{\sin(\sqrt{\lambda}h)} & & & \\ -\frac{\sqrt{\lambda}}{\sin(\sqrt{\lambda}h)} & \dots & & \\ & \dots & \dots & -\frac{\sqrt{\lambda}}{\sin(\sqrt{\lambda}h)} \\ & & -\frac{\sqrt{\lambda}}{\sin(\sqrt{\lambda}h)} & \frac{2\sqrt{\lambda} \cos(\sqrt{\lambda}h)}{\sin(\sqrt{\lambda}h)} \end{pmatrix} \quad (33)$$

Методы, позволяющие решить задачу (32), описаны в главе VI.

Нетрудно получить точное решение задачи (22):

$$\begin{aligned} \lambda &= (\pi n)^2, n \in \mathbb{Z} \\ u &= \sin \pi n x \end{aligned} \quad (34)$$

Интересно отметить, что исходная задача (22) имеет счетное множество решений (собственных значений и собственных функций). При этом получаемая разностная схема за счет специального выбора базисных функций является точной в точках сетки при заданном λ (см. Примечание 1 в главе 0). Отсюда следует, что получаемая с помощью метода конечных элементов с модифицированным базисом дискретная задача также имеет счетное множество решений. При этом, решая задачу (32), можно получить *любое* решение исходной задачи.

На рис. 3 и 4 представлены результаты расчетов для различных начальных приближений к собственным значениям ($n=5$ и $n=20$ в (34) соответственно). Из рис. 4 видно, что описанный метод применим даже для больших собственных значений и малого числа точек. Нужно отметить, что данный пример носит исключительно иллюстративный характер, и высокая точность объясняется использованием, информации о точном решении.

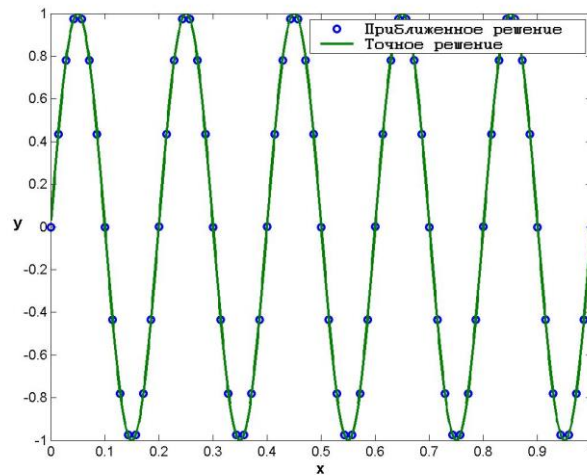


Рисунок 3

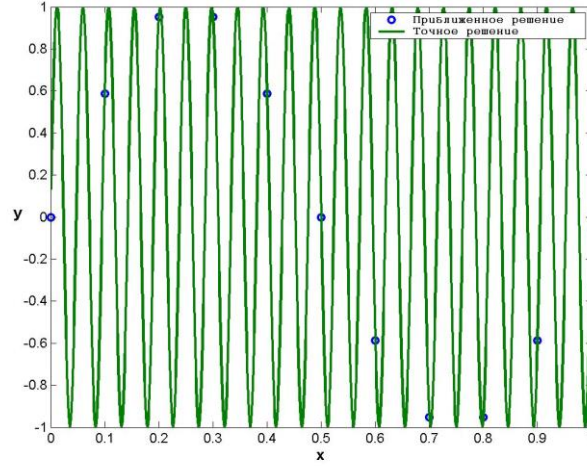


Рисунок 4

Замечание: сетка с постоянным шагом была выбрана в данном примере исключительно для простоты изложения. Описанный метод может быть применен и в случае сетки с переменным шагом.

IV.2. Задача на собственные значения для оператора второй производной с переменным коэффициентом

Рассмотрим следующую краевую задачу на собственные значения:

$$\begin{cases} -(x^2 u')' = \lambda u, x \in [0, 1], \lambda > 0 \\ x^2 u'|_{x=0} = 0 \\ u(1) = 0 \end{cases} \quad (35)$$

Известно, что данная задача имеет непрерывный спектр $[0.25, +\infty)$. Представляет интерес вопрос нахождения нижней границы спектра подобных задач.

Попытаемся найти точное решение уравнения $-(x^2 u')' = \lambda u$ на отрезке сетки. Для этого преобразуем левую часть уравнения по формуле дифференцирования произведения. В результате получим:

$$x^2 u'' + 2xu' + \lambda u = 0 \quad (36)$$

Полученное уравнение сводится к линейному дифференциальному уравнению с постоянными коэффициентами заменой переменных $x = e^t$.

Обозначим $q = \sqrt{1/4 - \lambda}$. Тогда общее решение уравнения (36) имеет вид:

$$u = c_1 x^{-1/2} \cos(q \ln x) + c_2 x^{-1/2} \sin(q \ln x) \quad (37)$$

Как видно полученная функция имеет особенность в нуле. Поэтому выбирать базис для метода конечных элементов в виде функций (37) нельзя. Чтобы избавиться от особенности в нуле сдвинем левую границу отрезка на ε и устремим затем ε к нулю. Таким образом, мы переходим к новой задаче на собственные значения:

$$\begin{cases} -(x^2 u')' = \lambda u, x \in [\varepsilon, 1], \lambda > 0 \\ x^2 u'|_{x=\varepsilon} = 0 \\ u(1) = 0 \end{cases} \quad (38)$$

Исследуем спектр этой задачи. Будем искать собственные функции в виде (37). Из граничных условий получаем:

$$\begin{cases} u(1) = c_1 = 0 \\ u'|_{\varepsilon} = -\frac{1}{2} c_2 \sin(q \ln \varepsilon) + q c_2 \cos(q \ln \varepsilon) = 0 \end{cases} \quad (39)$$

Из второго условия получаем уравнение относительно q :

$$\tan(q \ln \varepsilon) = 2q \quad (40)$$

Откуда приближенно (при $\varepsilon \rightarrow 0$):

$$q \approx \frac{\pi n}{2 - \ln \varepsilon}, n \in \mathbb{Z} \quad (41)$$

И следовательно:

$$\lambda \approx \frac{1}{4} + \left(\frac{\pi n}{2 - \ln \varepsilon} \right)^2 \quad (42)$$

Таким образом, задача (38) будет иметь дискретный спектр и нижняя граница спектра задачи стремится к 0.25 (т.е. к нижней границе спектра задачи (35)) при $\varepsilon \rightarrow 0$.

Оценка (42) может быть использована также как начальное приближение при численном решении задачи итерационным методом.

Введем равномерную сетку: $\omega_h = \{x_i = \varepsilon + \frac{(1-\varepsilon)i}{N}, i = 0..N\}$.

В соответствии с (21) будем искать базисные функции $\tilde{\varphi}_i$ как решение следующих краевых задач на отрезках $[x_{i-1}, x_i]$ и $[x_i, x_{i+1}]$:

$$\begin{cases} -(x^2 \tilde{\varphi}'_i)' = \lambda \tilde{\varphi}_i, & x \in [x_{i-1}, x_i] \\ \tilde{\varphi}_i(x_i) = 1 \\ \tilde{\varphi}_i(x_{i-1}) = 0 \end{cases} \quad (43)$$

$$\begin{cases} -(x^2 \tilde{\varphi}'_i)' = \lambda \tilde{\varphi}_i, & x \in [x_i, x_{i+1}] \\ \tilde{\varphi}_i(x_i) = 1 \\ \tilde{\varphi}_i(x_{i+1}) = 0 \end{cases}$$

Рассмотрим теперь обобщенную постановку задачи (38). Пусть, как и ранее, $\tilde{V}_h$ - конечномерное пространство функций с базисом $\tilde{\varphi}_i, i = 1..N-1$. Будем искать приближенное решение в виде разложения по базису $\tilde{V}_h$:

$$u(x) = \sum_{i=1}^{N-1} c_i \tilde{\varphi}_i(x) \quad (44)$$

Тогда в обобщенной постановке задача выглядит следующим образом:

$$a(u, v) = \lambda b(u, v) \quad \forall v \in \tilde{V}_h \quad (45)$$

где $a(u, v)$, $b(u, v)$ – билинейные формы, которые имеют вид:

$$\begin{aligned} a(u, v) &= \int_{\varepsilon}^1 x^2 u' v' dx \\ b(u, v) &= \int_{\varepsilon}^1 u v dx \end{aligned} \quad (46)$$

Зная точное решение (37), можно найти базисные функции, удовлетворяющие условиям (43):

$$\tilde{\varphi}_i = \begin{cases} c_{1,i}^- x^{-1/2} \cos(q \ln x) + c_{2,i}^- x^{-1/2} \sin(q \ln x), x \in [x_{i-1}, x_i] \\ c_{1,i}^+ x^{-1/2} \cos(q \ln x) + c_{2,i}^+ x^{-1/2} \sin(q \ln x), x \in [x_i, x_{i+1}] \end{cases} \quad (47)$$

где

$$\begin{aligned} c_{1,i}^- &= -x_i^{1/2} \frac{\sin(q \ln x_{i-1})}{\sin(q \ln \frac{x_i}{x_{i-1}})} \\ c_{2,i}^- &= x_i^{1/2} \frac{\cos(q \ln x_{i-1})}{\sin(q \ln \frac{x_i}{x_{i-1}})} \\ c_{1,i}^+ &= -x_i^{1/2} \frac{\sin(q \ln x_{i+1})}{\sin(q \ln \frac{x_i}{x_{i+1}})} \\ c_{2,i}^+ &= x_i^{1/2} \frac{\cos(q \ln x_{i+1})}{\sin(q \ln \frac{x_i}{x_{i+1}})} \end{aligned} \quad (48)$$

Важно отметить, что здесь, в отличие от предыдущего примера, функции $\tilde{\varphi}_i(x)$ различны на разных отрезках сетки. На рисунке 5 изображены графики функций $\tilde{\varphi}_i(x)$ при $N=10, i=1,2,3,4$.

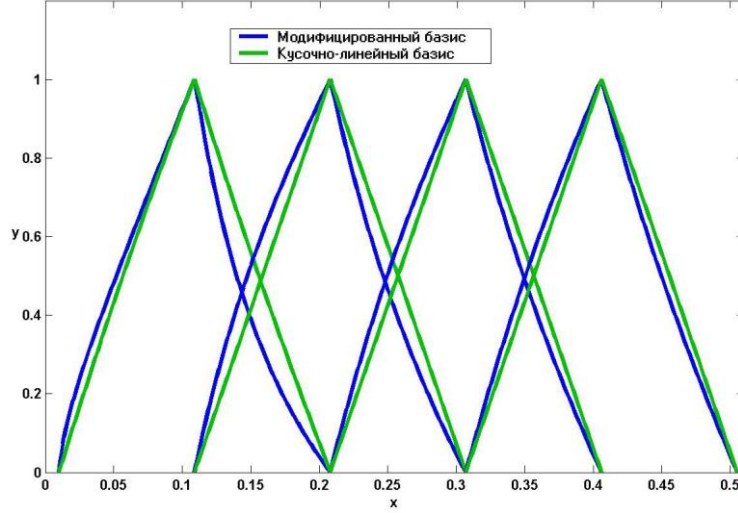


Рисунок 5

Аналогично предыдущей задаче вычисляем коэффициенты матрицы $A(\lambda)$:

$$\begin{aligned}
 a(\tilde{\varphi}_i, \tilde{\varphi}_i) - \lambda b(\tilde{\varphi}_i, \tilde{\varphi}_i) &= \tilde{\varphi}'_i|_{x_i-0} - \tilde{\varphi}'_i|_{x_i+0} \\
 a(\tilde{\varphi}_{i+1}, \tilde{\varphi}_i) - \lambda b(\tilde{\varphi}_{i+1}, \tilde{\varphi}_i) &= -\tilde{\varphi}'_{i+1}|_{x_i} \\
 a(\tilde{\varphi}_{i-1}, \tilde{\varphi}_i) - \lambda b(\tilde{\varphi}_{i-1}, \tilde{\varphi}_i) &= \tilde{\varphi}'_{i-1}|_{x_i}
 \end{aligned} \tag{49}$$

Для φ'_i получаем следующие выражения:

$$\tilde{\varphi}'_i = \begin{cases} c_{1,i}^- x^{-3/2} \left(-\frac{1}{2} \cos(q \ln x) - q \sin(q \ln x) \right) + \\ \quad c_{2,i}^- x^{-3/2} \left(-\frac{1}{2} \sin(q \ln x) + q \cos(q \ln x) \right), x \in [x_{i-1}, x_i] \\ c_{1,i}^+ x^{-3/2} \left(-\frac{1}{2} \cos(q \ln x) - q \sin(q \ln x) \right) + \\ \quad c_{2,i}^+ x^{-3/2} \left(-\frac{1}{2} \sin(q \ln x) + q \cos(q \ln x) \right), x \in [x_i, x_{i+1}] \end{cases} \tag{50}$$

На рис. 6 и 7 представлены результаты решения задачи при $\varepsilon = 10^{-3}$, $N = 50$ для первого и третьего собственных значений. Для сравнения на рис. 8 и 9 приведены результаты решения задачи с теми же параметрами методом конечных элементов с традиционным кусочно-линейным базисом.

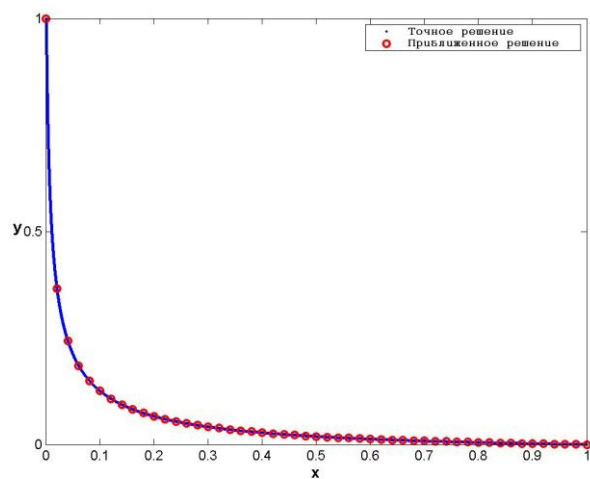


Рисунок 6

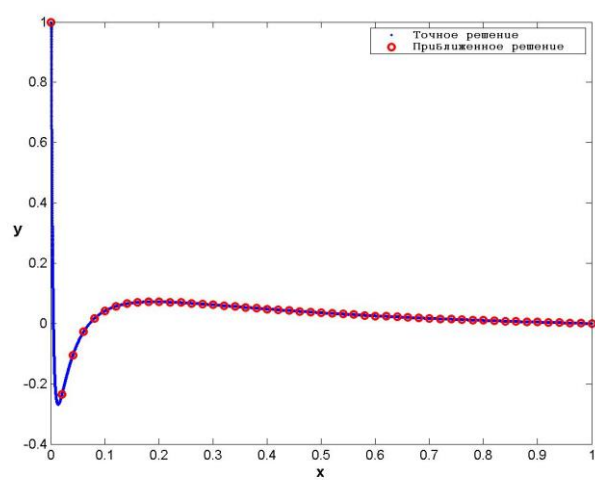


Рисунок 7

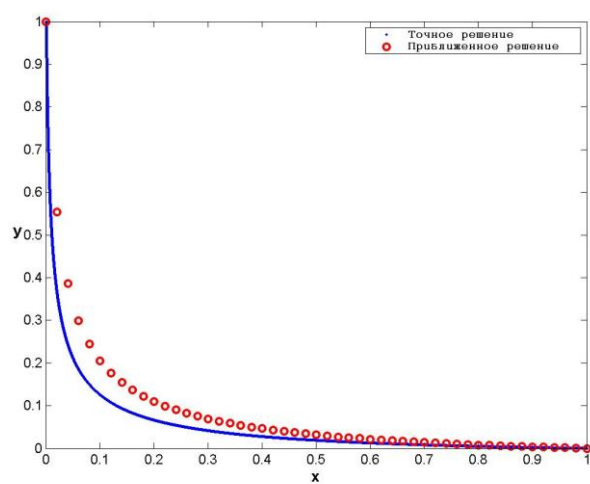


Рисунок 8

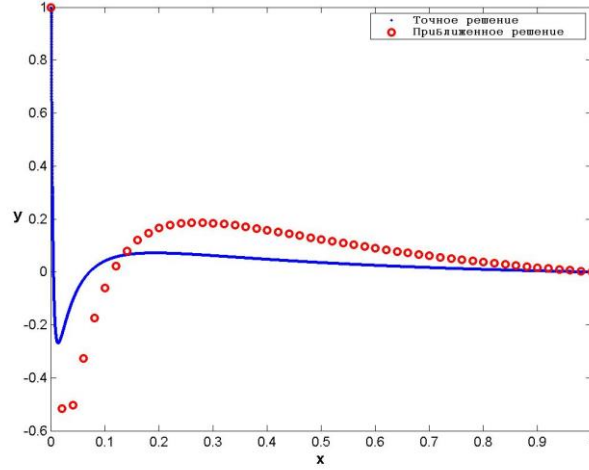


Рисунок 9

V. Двумерный случай.

V.1. Общая теория.

Рассмотрим задачу (1) в двумерной области Ω . Пусть $\{K\}$ - разбиение области Ω на треугольные подобласти. Согласно изложенному в главе III методу требуется решить краевую задачу (21), зависящую от параметра λ , для каждого элемента сетки K :

$$\begin{cases} L\tilde{\varphi}_i = \lambda\tilde{\varphi}_i, & \text{на } K \\ \tilde{\varphi}_i = \psi_i, & \text{на } \partial K \end{cases} \quad (51)$$

Здесь ψ_i - базисные функции пространства V_1 , линейные на каждом элементе сетки K .

Даже в простейшем случае, когда L – это оператор Лапласа с постоянными коэффициентами, аналитическое решение подобной задачи в треугольнике произвольной формы представляет сложность. Поэтому мы не можем выбрать в качестве пространства $V_b(K)$ на каждом элементе сетки K пространство *всех* функций из $H^1(K)$, обращающихся в ноль на границе элемента. Вместо этого введем на каждом элементе сетки K треугольную подсетку и в качестве пространства $V_b(K)$ рассмотрим пространство функций, которые являются кусочно-линейными внутри элемента сетки и линейными на каждом элементе подсетки.

Обозначим N – число внутренних точек сетки области Ω , T – число треугольников в триангуляции области Ω , M – число внутренних точек подсетки на каждом треугольнике глобальной сетки. Будем считать, что число внутренних точек на каждом треугольнике одинаково. Это делается исключительно для простоты дальнейшего изложения.

Пусть $\{\psi_i\}_{i=1}^N$ - базис пространства V_1 , а $\{\psi_{k,l}\}_{l=1}^M$ - базис пространства $V_b(K_k)$ для каждого треугольника сетки $k=1..T$. Обращаем внимание, что в данном случае пространство V_b будет конечномерным. Тогда любая функция $\tilde{u}_h$ из пространства $V_1 \oplus V_b$ раскладывается по базису:

$$\tilde{u}_h = \sum_{i=1}^N c_i \psi_i + \sum_{k=1}^T \sum_{l=1}^M c_{k,l} \psi_{k,l} \quad (52)$$

Подставляя разложение (52) в систему (9), получим следующую систему линейных алгебраических уравнений относительно коэффициентов c_i и $c_{k,l}$:

$$\left\{ \begin{array}{l} \sum_{i=1}^N c_i a(\psi_i, \psi_j) + \sum_{k=1}^T \sum_{l=1}^M c_{k,l} a(\psi_{k,l}, \psi_j) - \\ - \lambda \left[\sum_{i=1}^N c_i b(\psi_i, \psi_j) + \sum_{k=1}^T \sum_{l=1}^M c_{k,l} b(\psi_{k,l}, \psi_j) \right] = 0, \forall j = 1..N \\ \sum_{i=1}^N c_i a(\psi_i, \psi_{m,n}) + \sum_{k=1}^T \sum_{l=1}^M c_{k,l} a(\psi_{k,l}, \psi_{m,n}) - \\ - \lambda \left[\sum_{i=1}^N c_i b(\psi_i, \psi_{m,n}) + \sum_{k=1}^T \sum_{l=1}^M c_{k,l} b(\psi_{k,l}, \psi_{m,n}) \right] = 0, \forall m = 1..T, \forall n = 1..M \end{array} \right. \quad (53)$$

Следует отметить, что полученная система имеет большую размерность. Однако ее можно не решать непосредственно, а свести к системе относительно только коэффициентов c_i .

Изучим структуру матрицы системы (53). Размерность этой системы равна $N + TM$. Систему (53) можно записать в матричном виде:

$$\left\{ \begin{array}{l} Au^{(1)} + \sum_{k=1}^T B_k u_k^{(2)} = 0 \\ C_k u^{(1)} + D_k u_k^{(2)} = 0, k = 1..T \end{array} \right. \quad (54)$$

где $u^{(1)} \in \mathbf{C}^{N \times 1}$ - вектор-столбец неизвестных, соответствующих узлам глобальной сетки $c_i, i = 1..N$; $u_k^{(2)} \in \mathbf{C}^{M \times 1}$ - векторы-столбцы неизвестных, соответствующих узлам на подсетках $c_{k,l}, l = 1..M$; $A = A(\lambda, \psi_j, j = 1..N) \in \mathbf{C}^{N \times N}$ - квадратная матрица, зависящая от λ и от $\{\psi_j\}_{j=1}^N$; $B_k = B_k(\lambda, \psi_j, j = 1..N, \psi_{k,l}, l = 1..M) \in \mathbf{C}^{N \times M}$ и $C_k = C_k(\lambda, \psi_j, j = 1..N, \psi_{k,l}, l = 1..M) \in \mathbf{C}^{M \times N}$ - прямоугольные матрицы, зависящие от λ , $\{\psi_j\}_{j=1}^N$ и от $\{\psi_{k,l}\}_{l=1}^M$; $D_k = D_k(\lambda, \psi_{k,l}, l = 1..M) \in \mathbf{C}^{M \times M}$ - квадратные матрицы, зависящие от λ и от $\{\psi_{k,l}\}_{l=1}^M$.

Заметим, что, если оператор L – самосопряженный, то матрица системы (53) симметричная, и следовательно:

$$A = A^T, C_k = B_k^T, D_k = D_k^T, k = 1..T \quad (55)$$

Кроме того, нетрудно видеть, что матрицы B_k имеют только три ненулевых строки, соответствующие вершинам k -го треугольника (соответственно матрицы C_k имеют только три ненулевых столбца).

Покажем, что описываемый в данной главе метод соответствует общему методу, изложенному в главе III. Для этого достаточно определить вид оператора $M_k(\lambda)$ для любого элемента сетки K , действующего из пространства V_1 в пространство $V_b(K)$:

$$u_{b,K} = M_K(\lambda)u_1 \quad (56)$$

Из уравнений $C_k u^{(1)} + D_k u_k^{(2)} = 0, k = 1..T$ системы (54) $u_j^{(2)}$ формально выражается через $u^{(1)}$:

$$u_k^{(2)} = -D_k^{-1} C_k u^{(1)}, k = 1..T \quad (57)$$

Последнее соотношение фактически задает оператор $M_K(\lambda)$, действующий из пространства V_1 в пространство $V_b(K)$, и матрица этого оператора есть $[-D_k^{-1} C_k]$.

На практике наиболее эффективным методом решения системы (54) является исключение переменных из векторов $u_k^{(2)}$ из матричного уравнения $Au^{(1)} + \sum_{k=1}^T B_k u_k^{(2)} = 0$.

Причем это можно сделать независимо для каждого k . В результате получим систему:

$$\tilde{A}u^{(1)} = 0 \quad (58)$$

Где матрица $\tilde{A}$ представляется в виде $\tilde{A} = A - \sum_{k=1}^T B_k D_k^{-1} C_k$. При этом для самосопряженного оператора L из симметричности матрицы A и соотношений (55) следует, что матрица $\tilde{A}$ также является симметричной.

Таким образом, задача свелась к решению задачи на собственные значения с матрицей порядка N .

V.2. Задача для оператора Лапласа в квадрате

Рассмотрим сначала задачу для оператора Лапласа в квадрате:

$$\begin{cases} -\Delta u = \lambda u & \text{на } \Omega \\ u = 0 & \text{на } \partial\Omega \end{cases} \quad (59)$$

где область Ω - единичный квадрат $[0,1] \times [0,1]$.

В этом случае билинейные формы $a(u, v)$ и $b(u, v)$ представляются в виде:

$$\begin{aligned} a(u, v) &= \int_{\Omega} u'v' d\Omega \\ b(u, v) &= \int_{\Omega} uv d\Omega \end{aligned} \quad (60)$$

Тогда задача (9) принимает вид:

$$\begin{aligned} \text{Найти } u_h &= u_1 + u_b = u_1 + \sum_K u_{b,K} \in V_1 \oplus V_b : \\ \begin{cases} (u'_1 + u'_b, v'_1) = \lambda(u_1 + u_b, v_1), \forall v_1 \in V_1 \\ (u'_1 + u'_{b,K}, v'_{b,K})_K = \lambda(u_1 + u_{b,K}, v_{b,K})_K, \forall K, \forall v_{b,K} \in V_b(K) \end{cases} \end{aligned} \quad (61)$$

Известно точное решение задачи (59):

$$\begin{aligned} u &= \sin(\pi x) \sin(\pi y) \\ \lambda &= (\pi)^2 + (\pi n)^2, \quad m, n \in Z \end{aligned} \quad (62)$$

На рис. 10 и 11 приведены результаты решения задачи (зависимость погрешности нахождения собственного числа от $1/N$, где N – число точек сетки) для старшего собственного числа ($m = n = 1$) для стандартного метода конечных элементов с кусочно-линейным базисом и описанного выше модифицированного метода соответственно. При

этом в модифицированном методе число треугольников на подсетке было взято равным 64.

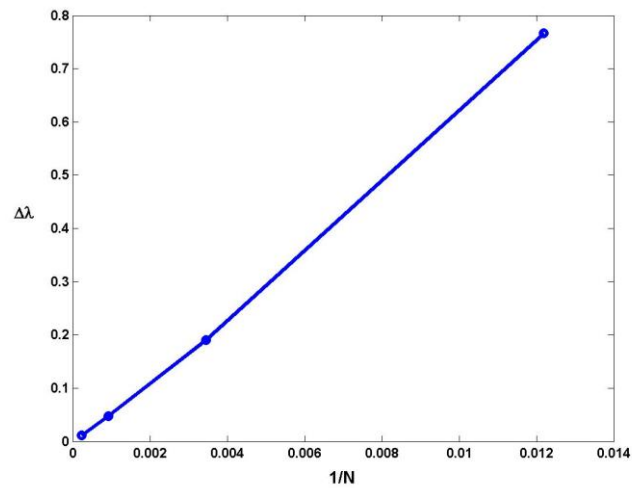


Рисунок 10

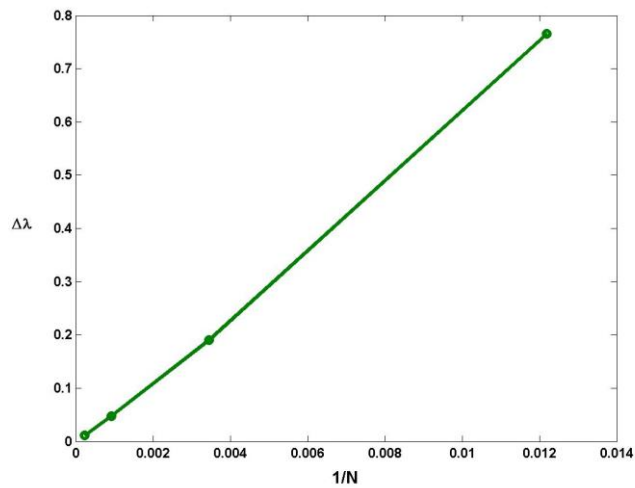


Рисунок 11

На рисунке 12 приведена зависимость $\lambda_{\kappa.э.} - \lambda_n$ от $1/M$, где $\lambda_{\kappa.э.}$ - собственное число, найденное с помощью обычного метода конечных элементов с кусочно-линейным базисом, λ_n - собственное число, найденное с помощью модифицированного метода на той же сетке(с введением подсетки), а M – число точек подсетки.

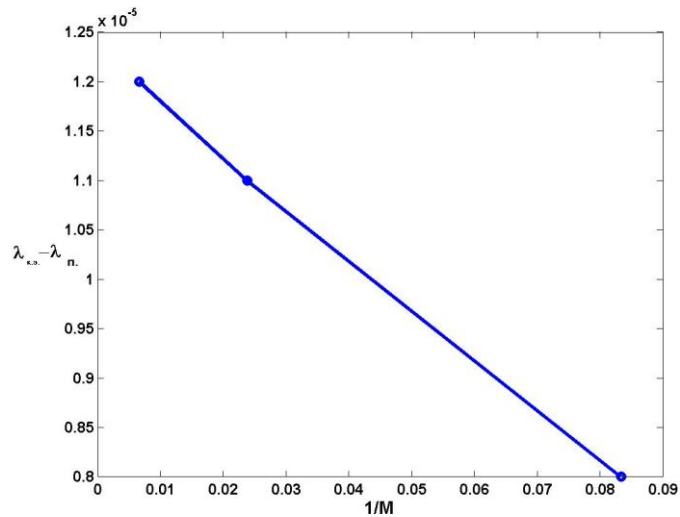


Рисунок 12

Как видно из рис. 10, 11 и 12 описанный метод, в отличие от одномерного случая, в двумерных задачах (в отличие от одномерных) хотя и дает повышение точности по сравнению со обычным методом, но незначительное. В чем же принципиальное отличие двумерного случая от одномерного для данного метода?

Напомним, что мы выбирали пространство, в котором ищется решение в виде $V_h = V_1 \oplus V_b$, где V_1 - пространство непрерывных кусочно-линейных функций, а V_b - некоторое, возможно бесконечномерное, замкнутое пространство функций, обращающихся в ноль на границе элементов сетки. В идеале V_b может включать в себя *все* функции, обращающиеся в ноль на границе элементов сетки. Тогда, как отмечено в Примечании 1 главы 0, в одномерном случае $V_h = V_1 \oplus V_b = H_0^1(\Omega)$, а $u_h = u_1 + u_b \in V_1 \oplus V_b$ фактически будет представлять собой *точное* решение. Конечно, в реальных задачах в качестве V_b выбирается некоторое подпространство пространства всех функций, обращающиеся в ноль на границе элементов сетки, однако при хорошем выборе можно получить решение близкое к точному. Принципиальное отличие двумерных задач в том, что даже если выбрать в качестве V_b пространство *всех* функции, обращающиеся в ноль на границе элементов сетки, мы не получим равенства $V_h = V_1 \oplus V_b = H_0^1(\Omega)$. Действительно, нетрудно видеть, что на ребрах треугольников сетки функции из V_h всегда будут линейными. Это является принципиальным отличием задач в областях более высокой размерности от одномерных. И, как показывает приведенный пример, это оказывается существенным. Следовательно, единственный возможный путь улучшения метода – это модификация базисных функций не только внутри элементов сетки, но и на их границе. Отметим, что вопрос о том, как это сделать, пока не решен.

V.3. Задача для оператора Лапласа в L-образной области

Рассмотрим сначала задачу для оператора Лапласа в L-образной области:

$$\begin{cases} -\Delta u = \lambda u & \text{на } \Omega \\ u = 0 & \text{на } \partial\Omega \end{cases} \quad (63)$$

где Ω - область, изображенная на рис. 13.

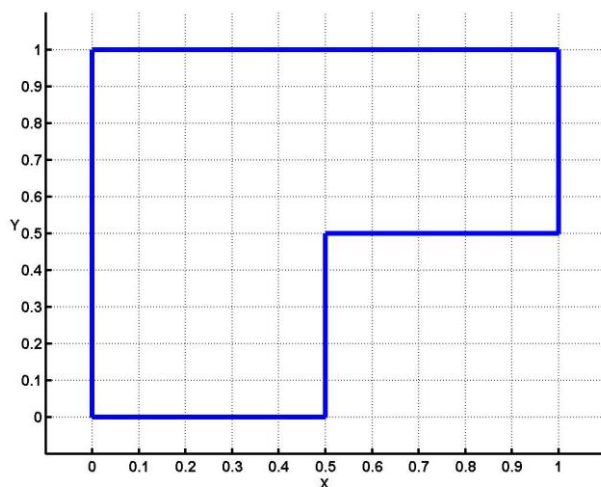


Рисунок 13

Покажем, что собственная функция, отвечающая старшему собственному числу, имеет особенность в точке $(0.5, 0.5)$. Представим собственную функцию в окрестности этой точки в полярных координатах:

$$u = \rho^\mu e^{in\varphi} \quad (64)$$

Здесь φ - это угол относительно оси x , а ρ - расстояние от точки $(0.5, 0.5)$. Тогда, записывая уравнение системы (63) в полярных координатах, получим:

$$\mu^2 \rho^{\mu-2} e^{in\varphi} - n^2 \rho^{\mu-2} e^{in\varphi} = \lambda \rho^\mu e^{in\varphi} \quad (65)$$

Из последнего равенства после очевидных преобразований получаем:

$$\mu^2 = -n^2 + \lambda \rho^2 \quad (66)$$

Из краевого условия системы (63) нетрудно получить, что $n = \frac{2}{3}$. Тогда учитывая, что в окрестности точки $(0.5, 0.5)$ ρ стремится к нулю, получаем $\mu \rightarrow \frac{2}{3}$. Отсюда и следует, что собственная функция будет иметь особенность в точке $(0.5, 0.5)$.

Введем билинейные формы $a(u, v)$ и $b(u, v)$ также, как и в предыдущей задаче, т.е. в виде (60) (но, естественно, на другой области Ω). Тогда обобщенная задача будет иметь вид (61).

На рис. 14 и 15 приведены результаты решения задачи описанным выше модифицированным методом для старшего собственного числа. При этом число треугольников на подсетке было взято равным 64. На рис. 14 приведена зависимость погрешности нахождения собственного числа от $1/N$, где N – число точек сетки (в качестве “точного” было взято значение $\lambda = 38.6048113$). На рис. 15 изображена собственная функция, соответствующая старшему собственному числу.

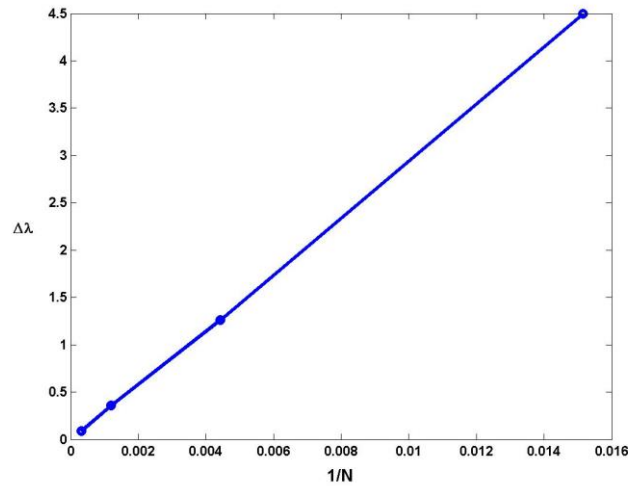


Рисунок 14

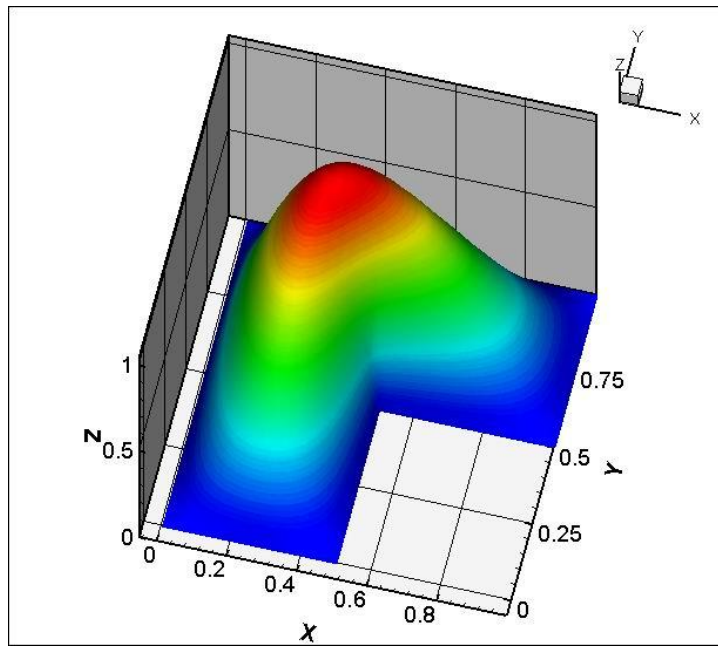


Рисунок 15

VI. Методы решения матричных задач на собственные значения с нелинейным вхождением собственного числа

VI.1. Метод обратных итераций для матричных задач на собственные значения с нелинейным вхождением собственного числа.

Рассмотрим следующую нелинейную задачу на собственные значения:

$$A(\lambda)x = 0 \quad (67)$$

где $A(\lambda) \in \mathbb{C}^{N \times N}$ - квадратная матрица с нелинейным вхождением λ .

Рассмотрим итерационный метод решения таких задач [6]. Пусть заданы два начальных приближения для собственного значения: λ^0 и λ^1 , а также начальное приближение для собственного вектора — x_0 . Два начальных приближения для собственного значения требуется для численного нахождения производной матрицы A .

Обозначим через $A'_n(\lambda)$ разностную производную матрицы A , вычисляемую на n -й итерации по формуле:

$$A'_n(\lambda) = \frac{A(\lambda^n) - A(\lambda^{n-1})}{\lambda^n - \lambda^{n-1}} \quad (68)$$

Это матричное выражение вычисляется поэлементно.

Метод состоит из двух вложенных циклов. Во внутреннем цикле проводятся итерации по k . На каждой итерации находится новое приближение к собственному вектору $-x_{k+1}$ и новое приближение к поправке на собственное значение - $\delta\lambda_{k+1}^n$ по формулам:

$$\begin{aligned}
A(\lambda^n)y_{k+1} &= -A'_n(\lambda^n)x_k \\
x_{k+1} &= \frac{y_{k+1}}{\|y_{k+1}\|} \\
\delta\lambda_{k+1}^n &= \frac{(A'_n(\lambda^n)x_k, (A'_n(\lambda^n)y_{k+1})^*)}{\|y_{k+1}\|^2}
\end{aligned} \tag{69}$$

где норма $\|\cdot\|$ определяется следующим образом: $\|\cdot\| = (A'_n(\lambda^n)\cdot, (A'_n(\lambda^n)\cdot)^*)^{1/2}$. В результате внутренних итераций по индексу k находится поправка на собственное значение - $\delta\lambda^n$. Во внешнем цикле проводятся итерации по индексу n , на каждой итерации после внутреннего цикла перевычисляется приближение к собственному значению по формуле:

$$\lambda^{n+1} = \lambda^n + \delta\lambda^n \tag{70}$$

Описанный метод фактически является комбинацией метода обратных итераций для линейных задач и метода Ньютона. Как известно, метод Ньютона сходится, если начальное приближение для λ достаточно близко к точному корню.

VI.2. Метод Дэвиса

Данный метод [2] предназначен для поиска *всех* нулей аналитической функции в заданной ограниченной области на комплексной плоскости.

Пусть $f(z)$ - аналитическая функция в области, ограниченной контуром C .

Тогда интегралы

$$S_n = \frac{1}{2\pi i} \int_C z^n \frac{f'(z)}{f(z)} dz, \quad n = 1, \dots, N \tag{71}$$

содержат в себе информацию обо всех нулях аналитической функции, лежащих в области, ограниченной контуром C . Здесь N – число корней в рассматриваемой области.

В качестве области будем рассматривать единичный круг с центром в начале координат. При этом фактически можно решать задачу в любой ограниченной области, которую можно конформно отобразить в единичный круг.

Первым шагом метода является определение числа корней. Число корней N в области определяется по принципу аргумента:

$$N = \frac{\Delta \text{Arg } f(z)}{2\pi} \tag{72}$$

где $\Delta \text{Arg } f(z)$ - изменение аргумента при обходе точкой z полного контура в положительном направлении.

Пусть $\{z_k\}_{k=1..m}$ - точки равномерно распределенные по контуру.

В работе [2] предлагаются следующие неравенства, которые гарантируют надежное определение числа нулей по принципу аргумента:

$$\begin{aligned}
\Phi_{\max} &< \frac{3\pi}{4} \\
\frac{1}{6.1} &< \left| \frac{f(z_k)}{f(z_{k+1})} \right| < 6.1, \quad k = 1..m
\end{aligned} \tag{73}$$

где $\Phi_{\max}$ - максимальное изменение аргумента от точки к точке.

Очевидно, что всегда можно добиться выполнения этих неравенств, увеличивая число точек m на контуре.

Вычисление интегралов (71) по квадратурным формулам непосредственно осложнено тем, что в выражение явно входит производная $f'(z)$. Для того, чтобы избавиться от явного вхождения производной производится интегрирование по частям:

$$\begin{aligned} S_n &= \frac{1}{2\pi i} \int_C z^n \frac{f'(z)}{f(z)} dz = \frac{1}{2\pi i} \int_C z^n \frac{f'(z)}{f(z)} dz - \frac{N}{2\pi i} \int_C z^{n-1} dz = \\ &= \frac{1}{2\pi i} \int_C z^n \frac{[z^{-N} f'(z) - N z^{-N-1} f(z)]}{z^{-N} f(z)} dz = \frac{-n}{2\pi i} \int_C z^{n-1} \ln[z^{-N} f(z)] dz, \quad n=1..N \end{aligned}$$

Так как рассматриваемая область является единичным кругом, удобно перейти к полярным координатам:

$$z = e^{i\theta} \quad (74)$$

Тогда интегралы S_n представляются в виде:

$$S_n = \frac{-n}{2\pi} \int_0^{2\pi} e^{in\theta} \ln[e^{-iN\theta} f(e^{i\theta})] d\theta \quad n=1, \dots, N \quad (75)$$

Полученные интегралы аппроксимируются квадратурными формулами в точках $\{z_k\}_{k=1..m}$ по методу трапеций:

$$S_n^m = \frac{-n}{m} \sum_{k=1}^m e^{\frac{2\pi i k n}{m}} \ln[e^{\frac{-2\pi i k N}{m}} f(e^{\frac{2\pi i k}{m}})], \quad n=1..N \quad (76)$$

Из формул (76) приближенно вычисляются значения S_n .

Следующий шаг – определение значений $\{a_i\}_{i=1..N}$ - нулей функции $f(z)$ в рассматриваемой области. Как известно, интегралы S_n обладают свойством:

$$S_n = \sum_{i=1}^N a_i^n, \quad n=1, \dots, N \quad (77)$$

При этом, если у функции $f(z)$ есть кратные нули, то они входят соответствующее число раз в сумму (77).

Будем искать нули $\{a_i\}_{i=1..N}$ как корни полинома:

$$P_N(z) = \prod_{i=1}^N (z - a_i) = \sum_{i=0}^N A_i z^{N-i} \quad (78)$$

Соотношение между коэффициентами A_i и суммами S_n известно из линейной алгебры:

$$\begin{aligned} S_1 + A_1 &= 0 \\ S_2 + A_1 S_1 + 2A_2 &= 0 \end{aligned} \quad (79)$$

$$S_k + A_1 S_{k-1} + A_2 S_{k-2} + \dots + kA_k = 0, \quad k=1..N$$

Отсюда, зная S_n , можно рекуррентно вычислить коэффициенты A_i . Таким образом, задача нахождения нулей $\{a_i\}_{i=1..N}$ свелась к задаче нахождения нулей полинома N -й степени. Кроме того, известно, что все нули этого полинома лежат внутри единичного круга. Задача нахождения этих нулей является простой, для этого может быть использован метод Ньютона.

В приведенных выше выкладках предполагалось, что в качестве рассматриваемой области выбран единичный круг с центром в начале координат. Однако фактически может использоваться любая ограниченная область, которую конформным отображением можно перевести в единичный круг. Действительно, пусть $f(w)$ - функция, заданная в некоторой области, а $w(z)$ - конформное отображение единичного круга в данную область, тогда описанный метод можно применить к функции, заданной в единичном круге $\varphi(z) = f(w(z))$.

В данной работе применялся вариант изложенного метода для круга произвольного радиуса с произвольным центром на комплексной плоскости. Метод использовался для решения матричной задачи на собственные значения с нелинейным вхождением собственного числа, для чего в качестве функции выбирался определитель матрицы: $f(\lambda) = \det A(\lambda)$.

В работе [2] приведена оценка погрешности вычисления сумм S_n . Пусть функция $f(z)$ аналитическая в круге $|z| < R$, где $R > 1$, $a_j, j = 1..N$ - нули функции внутри круга $|z| < 1$, $b_j, j = 1..N$ - нули функции в области $1 < |z| < R$. Тогда справедлива оценка:

$$|S_n - S_n^m| \leq \sum_{j=1}^N \frac{na_j^{n+m}}{n+m} + \sum_{j=1}^{N'} \frac{nb_j^{-m}}{m-n} + \frac{R^{1-m}}{1-R^{-m}} \sqrt{\frac{1}{2\pi i} \int_{|z|=R} |\Psi(z)|^2 \frac{dz}{z}} \quad (80)$$

где $\Psi(z)$ - некоторая аналитическая функция в рассматриваемой области.

Из приведенной оценки следует, что наибольшую погрешность вносят нули расположенные близко к контуру. Кроме того, погрешность зависит от числа нулей в выбранной области и снаружи вблизи границы области. Поэтому метод Дэвиса может использоваться для нахождения нулей с приемлемой точностью лишь при не очень большом их числе в выбранной области. При практическом использовании метода зависимость погрешности от числа точек интегрирования наблюдалась лишь при малом числе точек. Связано это, по всей видимости, с тем, что погрешность от приближенного вычисления интегралов мала по сравнению с погрешностью вычисления определителя матрицы (для вычисления определителя использовалось LU -разложение).

VI.3. Комбинация двух методов.

Достоинством метода Дэвиса является то, что он включает в себя способ определения числа корней в заданной области, что не всегда известно заранее. Недостатком метода является то, что он не позволяет определить собственные вектора и невысокая точность при большом числе корней и неудачном выборе области.

Используемый итерационный метод дает значительное более точное решение, чем метод Дэвиса и позволяет определить не только собственные числа, но и собственные векторы, однако требует достаточно хорошего начального приближения к собственному значению.

Исходя из изложенного, наиболее эффективной представляется комбинация двух методов, в которой сначала используется метод Дэвиса в рассматриваемой области для определения количества собственных значений и их приближенных значений, а затем полученные значения используются как начальные приближения для итерационного метода. В результате собственные значения определяются с большей точностью и имеется возможность определить собственные вектора.

VII. Заключение.

В данной работе рассмотрен вопрос применения метода RFB к решению задач на собственные значения. Указанный метод рассматривается для эллиптического оператора в одномерной или двумерной области. Кроме того, доказано, что указанный метод эквивалентен методу конечных элементов с базисными функциями, выбранными специальным образом. А именно, показано, что можно выбрать базисные функции в методе конечных элементов таким образом, что в результате получается та же разностная задача, что и в методе RFB.

Приведены примеры применения описываемого метода для двух одномерных задач и для двух двумерных. Приводятся результаты решения указанных задач и сравнение точности метода с обычным методом конечных элементов с кусочно-линейным базисом. При этом показывается, что в рассматриваемых одномерных задачах модифицированный метод дает более точное решение. Однако в двумерном случае в рассматриваемых задачах описываемый метод не дает существенного улучшения точности по сравнению со стандартным методом конечных элементов с кусочно-линейным базисом. Связано это с тем, что в двумерном случае, в отличие от одномерного, требуется модификация базисных функций не только внутри элементов сетки, но и на их границе. Отметим, что вопрос о том, как это сделать, пока не решен.

В результате применения метода RFB для задач на собственные значения возникает матричная задача на собственные значения с нелинейным вхождением собственного числа. Рассмотрены методы решения таких задач: модифицированный метод обратных итераций и метод Дэвиса. А также рассмотрен вопрос применения их в комбинации друг с другом.

Все программные модули, используемые для расчетов, в том числе работа с матрицами, работа с треугольными сетками, реализация изложенных выше методов решения нелинейных задач на собственные значения, реализация методов RFB и конечных элементов, выполнены на языке C++ в среде C++Builder 5.0. Для создания иллюстраций к работе были использованы средства Matlab 6.5 и Tecplot 9.0.

Список литературы

1. Brezzi F., Franca L.P., Hughes T.J.R., Russo A. $b = \int g$. Comput. Methods Appl. Mech. Engrg. 145, 329-339, 1997.
2. Davies B., Locating zeros of an analytic function, Journal of Computational Physics 66, 36, 1986.
3. Franca L.P., Nesliturk A., Stynes M., On the stability of residual-free bubbles for convection-diffusion problems and their approximation by a two-level finite element method. Comput. Methods Appl. Mech. Engrg. 166, 35-49, 1998.
4. Hughes T.J.R. Multiscale phenomena: Green's functions, The Dirichlet-to-Neumann formulation, subgrid scale models, bubbles, and the origins of stabilized methods. Comput. Methods Appl. Mech. Engrg. 127, 387-401, 1995.
5. Hughes T.J.R., Feijoo G.R., Mazzei L., Quincy J.B. The variational multiscale method – a paradigm for computational mechanics. Comput. Methods Appl. Mech. Engrg. 166, 3-24, 1998.
6. Дроздова О.М. О сходимости одного итерационного метода решения нелинейной спектральной задачи. Дифференциальные уравнения. Том XXII, № 7, 1263-1265, 1986.