Министерство общего и профессионального
образования Российской Федерации

Пензенский государственный технический университет

Ю.Г. Смирнов

Проекционные методы

Методические указания

Пенза 1997

Аннотация

УДК 519.6
С 50

Рассмотрены проекционные методы, используемые в качестве универсального инструмента для приближенного решения операторных уравнений как 2–го, так и 1–го рода. Изложены метод коллокации и метод Галеркина как частные случаи общей схемы проекционного метода. Дано описание числовых характеристик проекционных методов.

Методические указания подготовлены на кафедре "Высшая математика" и предназначены для студентов и аспирантов специальности "Прикладная математика".

Автор Ю.Г. Смирнов

Р е ц е н з е н т Ю.В. Шестопалов, д–р физ.–мат. наук, проф. кафедры "Математическая физика" Московского государственного университета им.М.В.Ломоносова.

1. Проекционные методы

Рассмотрим приближенное решение линейных операторных уравнений с помощью проектирования их на подпространства, которые для практических применений будем считать имеющими конечную размерность. Сначала напомним определение проекционного оператора. Ниже все операторы предполагаются линейными.

Определение 1. Пусть X — нормированное пространство и $U \subset X$ — нетривиальное подпространство. Ограниченный оператор $P : X \to U$ такой, что $P\varphi = \varphi$ для всех $\varphi \in U$ называется проекционным оператором (или проектором) X на U.

Теорема 1. Нетривиальный ограниченный оператор является проектором тогда и только тогда, когда он удовлетворяет условию $P^2 = P$. В этом случае $\|P\| \geq 1$.

Доказательство. Пусть $P : X \to U$ проектор. Тогда из условия $P\varphi \in U$ следует, что $P^2\varphi = P(P\varphi) = P\varphi$ для всех $\varphi \in X$.

Обратно, пусть $P^2 = P$ и положим $U := P(X)$. Тогда для всех $\varphi \in U$ можно записать $\varphi = P\psi$ с некоторым $\psi \in X$, и в результате получим $P\varphi = \varphi$.

Наконец, так как $P = P^2 = PP$, то $\|P\| \leq \|P\| \, \|P\|$, откуда $\|P\| \geq 1$. □

В гильбертовых пространствах важным примером проекционного оператора является так называемый *ортопроектор* на подпространство U, то есть оператор наилучшей аппроксимации фиксированного элемента из пространства X элементами из подпространства U.

Теорема 2. Пусть U — нетривиальное подпространство гильбертова пространства X. Тогда оператор P, переводящий элемент $\varphi \in X$ в элемент наилучшей аппроксимации в подпространстве U, есть проекционный оператор. Он называется ортопроектором на U и удовлетворяет условию $\|P\| = 1$.

Доказательство. Очевидно, $P\varphi = \varphi$ для всех $\varphi \in U$. Из условия ортогональности для элемента наилучшей аппроксимации в гильбертовых пространствах без труда проверяется, что оператор P линеен и что

$$\|\varphi\|^2 = \|P\varphi + (\varphi - P\varphi)\|^2 = \|P\varphi\|^2 + \|\varphi - P\varphi\|^2 \geq \|P\varphi\|^2$$

для всех $\varphi \in X$. Следовательно, $\|P\| \leq 1$ и в силу теоремы 1 $\|P\| = 1$. □

Вторым важным примером проекторов являются *операторы интерполяции.*

Задача 1. Докажите, что оператор интерполяции является проекционным оператором.

Заметим, что полиномиальная интерполяция или интерполяция тригонометрическими полиномами в общем случае не дает сходимости в норме C. В результате операторы интерполяции не являются равномерно ограниченными относительно нормы C [2].

Определение 2. Пусть X и Y — банаховы пространства и $A : X \to Y$ — инъективный оператор. Пусть $X_n \subset X$ и $Y_n \subset Y$ — две последовательности подпространств с условиями $\dim X_n = \dim Y_n = n$ и пусть $P_n : Y \to Y_n$ — проекционные операторы. Рассмотрим проекционный метод, образованный посредством X_n и P_n, который аппроксимирует уравнение

$$A\varphi = f$$

с помощью приближенного уравнения

$$P_n A\varphi_n = P_n f.$$

Этот проекционный метод называется сходящимся для оператора A, если существует число N такое, что для каждого $f \in \Im A$ приближенное уравнение $P_n A\varphi = P_n f$ имеет единственное решение $\varphi_n \in X_n$ для всех $n \geq N$, и если эти решения сходятся $\varphi_n \to \varphi$ при $n \to \infty$ к единственному решению φ уравнения $A\varphi = f$.

В терминах операторов сходимость проекционного метода означает, что для всех $n \geq N$ конечномерные операторы $A_n := P_n A : X_n \to Y_n$ являются обратимыми и что поточечная сходимость $A_n^{-1} P_n A \varphi \to \varphi$, $n \to \infty$ имеет место для всех $\varphi \in X$. (Напомним, что последовательность ограниченных операторов B_n *поточечно сходится* к ограниченному оператору B, $B_n \to B$, $n \to \infty$, если $\|B_n \varphi - B\varphi\| \to 0$ при $n \to \infty$ для всех $\varphi \in X$.)

В общем случае можно ожидать сходимость метода только тогда, когда подпространства X_n предельно плотны в X:

$$\inf_{\psi \in X_n} \|\psi - \varphi\| \to 0, \quad n \to \infty \tag{1}$$

для всех $\varphi \in X$. Эту оценку называют также *свойством аппроксимации* (произвольный элемент из X может быть аппроксимирован элементами из подпространств X_n с любой точностью в норме X). В последующем анализе будем всегда предполагать, что это условие выполняется.

Поскольку $A_n = P_n A$ — оператор, действующий в конечномерных пространствах, применение проекционного метода сводится к решению конечномерной системы линейных алгебраических уравнений. Ниже метод коллокации и метод Галеркина будут рассмотрены как проекционные методы, полученные с помощью оператора интерполяции и ортогонального проектора соответственно.

Первоначально докажем общий результат о сходимости проекционного метода и дадим соответствующую оценку для погрешности метода.

Теорема 3. Проекционный метод сходится тогда и только тогда, когда существуют число N и положительная константа M такие, что для всех $n \geq N$ конечномерные операторы $A_n := P_n A : X_n \to Y_n$ обратимы и операторы $A_n^{-1} P_n A : X \to X$ равномерно ограничены

$$\|A_n^{-1} P_n A\| \leq M. \tag{2}$$

В этом случае верна оценка для погрешности метода

$$\|\varphi_n - \varphi\| \leq (1 + M) \inf_{\psi \in X_n} \|\psi - \varphi\|. \tag{3}$$

Доказательство. Если проекционный метод сходится, то равномерная ограниченность норм операторов (2) есть следствие *принципа равномерной ограниченности* (см. [3]), согласно которому если последовательность ограниченных операторов поточечно сходится (к некоторому ограниченному оператору), то их нормы равномерно ограничены.

Обратно, если предположения теоремы выполняются, то можно записать

$$\varphi_n - \varphi = (A_n^{-1} P_n A - I)\varphi.$$

Так как для всех $\psi \in X_n$, то, очевидно, $A_n^{-1} P_n A \psi = \psi$, и тогда

$$\varphi_n - \varphi = (A_n^{-1} P_n A - I)(\varphi - \psi).$$

Следовательно, имеем оценку (3) для погрешности метода и в силу свойства аппроксимации (1) — сходимость проекционного метода. □

Оценка погрешности (3), приведенная в теореме 3, обычно называется *квазиоптимальной*. Она показывает, что ошибка в проекционном методе определяется тем, как хорошо точное решение может быть аппроксимировано с помощью элементов подпространства X_n (в норме пространства X).

Задача 2. Докажите, что если оператор K компактен, а последовательность операторов B_n поточечно сходится к оператору B, $B_n \to B$, $n \to \infty$, то $\|B_n K - B\,K\| \to 0$, $n \to \infty$.

Теорема 4. Предположим, что $S : X \to Y$ — есть ограниченный оператор, имеющий ограниченный обратный $S^{-1} : Y \to X$, и что проекционный метод является сходящимся для S. Пусть $K : X \to Y$ будет линейным ограниченным оператором, удовлетворяющим любому из двух условий:

$$a) \quad \|K\| \quad \text{достаточно мало}$$

или

$$b) \quad K \quad \text{компактен и} \quad S - K \quad \text{инъективен.}$$

Тогда проекционный метод также сходится для оператора $S - K$.

Доказательство. Оператор S удовлетворяет условиям теоремы 3, а именно, операторы $S_n := P_n S$ обратимы для всех достаточно больших n и удовлетворяют условию $\|S_n^{-1} P_n S\| \le M$ с некоторой константой M. Так как оператор S имеет ограниченный обратный, то поточечная сходимость $S_n^{-1} P_n S \to I$, $n \to \infty$ на X влечет поточечную сходимость $S_n^{-1} P_n \to S^{-1}$, $n \to \infty$ на Y. Покажем, что для достаточно больших n обратные операторы $I - S_n^{-1} P_n K : X \to X$ существуют и равномерно ограничены, если выполняются условия a) или b).

Пусть выполнено условие a). Применим принцип равномерной ограниченности к поточечно сходящейся последовательности $S_n^{-1} P_n$. Тогда обратные операторы $(I - S_n^{-1} P_n K)^{-1}$ существуют и равномерно ограничены для всех достаточно больших n, если норма оператора K удовлетворяет оценке

$$\sup_{n \in N} \|S_n^{-1} P_n\| \, \|K\| = q < 1.$$

Действительно, если $q < 1$, то

$$\|(I - S_n^{-1} P_n K)^{-1}\| \le \sum_{m=0}^{\infty} \|(S_n^{-1} P_n K)^m\| \le \sum_{m=0}^{\infty} q^m = 1/(1 - q).$$

Теперь пусть выполнено условие b). Поскольку оператор S имеет ограниченный обратный, то $S^{-1} K$ компактен, и, следовательно, оператор $I - S^{-1} K$ фредгольмов (см. [4]). Далее, так как оператор $S - K$ инъективен, то в силу альтернативы Фредгольма $I - S^{-1} K : X \to X$ имеет ограниченный обратный. Из поточечной сходимости последовательности $(S_n^{-1} P_n)$ и компактности K получаем сходимость по норме (задача 2)

$$\|S^{-1} K - S_n^{-1} P_n K\| \to 0, n \to \infty.$$

Тогда оператор

$$I - S_n^{-1} P_n K = (I - S^{-1} K) + (S^{-1} K - S_n^{-1} P_n K)$$

можно рассматривать как малое по норме (при достаточно больших n) возмущение ограниченного оператора, имеющего ограниченный обратный. Следовательно (см. [4]), обратные операторы $(I - S_n^{-1} P_n K)^{-1}$ существуют и равномерно ограничены для всех достаточно больших n.

Отметим, что $(I - S_n^{-1} P_n K)^{-1}$ отображает X_n на себя. Обозначим $\tilde{S} := S - K$, а $\tilde{S}_n := P_n \tilde{S}$. Тогда $\tilde{S}_n = S_n (I - S_n^{-1} P_n K) : X_n \to Y_n$ обратим для достаточно больших n с обратным оператором

$$\tilde{S}_n^{-1} = (I - S_n^{-1} P_n K)^{-1} S_n^{-1}.$$

Из формулы

$$\tilde{S}_n^{-1} P_n \tilde{S} = (I - S_n^{-1} P_n K)^{-1} \, S_n^{-1} P_n S (I - S^{-1} K)$$

получаем

$$\|\tilde{S}_n^{-1} P_n \tilde{S}\| \le \|(I - S_n^{-1} P_n K)^{-1}\| \, \|I - S^{-1} K\| \, M,$$

откуда следует, что условие (2) выполняется для $\tilde{S}$. Доказательство завершено. □

Для уравнения второго рода

$$\varphi - K\varphi = f \tag{4}$$

с линейным ограниченным оператором $K : X \to X$ необходимо только выбрать последовательность подпространств $X_n \subset X$ и проекционные операторы $P_n : X \to X_n$. Тогда проекционный метод запишется в форме

$$\varphi_n - P_n K\varphi_n = P_n f. \tag{5}$$

Заметим, что каждое решение $\varphi_n \in X$ уравнения (5) автоматически принадлежит X_n. Когда оператор K компактен, то по теореме 4 имеем следующее утверждение о сходимости.

Утверждение 1. Пусть оператор $K : X \to X$ компактен, $I - K$ инъективен и пусть проекторы $P_n : X \to X_n$ сходятся поточечно к единичному оператору, то есть $P_n\varphi \to \varphi$, $n \to \infty$ для всех $\varphi \in X$. Тогда проекционный метод для $I - K$ сходится.

Доказательство. Применим вторую часть теоремы 4 для $S = I$ и отождествим $X = Y$ и $X_n = Y_n$. □

Теперь дадим альтернативное доказательство последнего результата о сходимости. Это также даст нам возможность показать, что проекционный метод для уравнений второго рода может сходиться без условия поточечной сходимости проекционных операторов на всех элементах из X.

Теорема 5. Пусть $K : X \to X$ — компактный оператор, а $I - K$ инъективный. Предположим, что проекторы $P_n : X \to X_n$ удовлетворяют условию $\|P_n K - K\| \to 0$, $n \to \infty$. Тогда для достаточно больших n приближенные уравнения (5) однозначно разрешимы для всех $f \in X$ и справедливы оценки

$$\|\varphi_n - \varphi\| \leq M \, \|P_n\varphi - \varphi\| \tag{6}$$

с некоторой положительной константой M, зависящей от K.

Доказательство. Поскольку K — компактный оператор, $I - K$ — фредгольмов. Тогда из инъективности $I - K$ следует, что он имеет ограниченный обратный оператор $(I - K)^{-1} : X \to X$. Для достаточно больших n имеем оценки

$$\|(I - P_n K)^{-1}\| = \|((I - K) + (K - P_n K))^{-1}\| \leq 2\|(I - K)^{-1}\| = M,$$

если

$$\|(K - P_n K)\| \leq 1/(2\|(I - K)^{-1}\|).$$

Следовательно, операторы $(I - P_n K)^{-1}$ существуют и равномерно ограничены. Для проверки формулы (6) применим проектор P_n к левой и правой частям уравнения (4). В результате получим

$$\varphi - P_n K\varphi = P_n f + \varphi - P_n\varphi.$$

Вычитая это выражение из формулы (5), находим

$$(I - P_n K)(\varphi_n - \varphi) = P_n\varphi - \varphi,$$

откуда и следует формула (6). □

Отметим, что поточечная сходимость $P_n\varphi \to \varphi$, $n \to \infty$ для всех $\varphi \in X$ влечет $\|P_n K - K\| \to 0$, $n \to \infty$. Условие $\|P_n K - K\| \to 0$, $n \to \infty$ может выполняться и без поточечной сходимости последовательности проекторов P_n. Разумеется, сходимость проекционного метода может иметь место только в том случае, когда условие $P_n\,\varphi \to \varphi$, $n \to \infty$ выполняется для точного решения уравнения φ.

На практике обычно вместо точного уравнения (5) решается неточное уравнение

$$\varphi_n - P_n K_n \varphi_n = P_n f_n, \tag{7}$$

где операторы K_n аппроксимируют оператор K, а элементы f_n аппроксимируют правую часть f. Тогда мы имеем следующее предложение.

Утверждение 2. Пусть выполняется условие $\|P_n K_n - K\| \to 0$ при $n \to \infty$. Тогда для достаточно больших n уравнение (7) имеет единственное решение и справедливы оценки

$$\|\varphi_n - \varphi\| \leq M\{\|P_n\varphi - \varphi\| + \|P_n(K_n - K)\varphi\| + \|P_n(f_n - f)\|\}. \tag{8}$$

Теперь рассмотрим уравнение первого рода

$$S\varphi - K\varphi = f, \tag{9}$$

где $S : X \to Y$ ограниченный оператор, имеющий ограниченный обратный $S^{-1} : Y \to X$, $K : X \to Y$ компактный оператор и $S - K$ инъективный.

Тогда для проекционного метода

$$P_n(S - K)\varphi_n = P_n f \tag{10}$$

с выбором подпространств X_n и Y_n и проекторов $P_n : Y \to Y_n$ имеем следующее предложение.

Утверждение 3. Пусть $Y_n = S(X_n)$ и $\|P_n K - K\| \to 0$ при $n \to \infty$. Тогда для достаточно больших n уравнение (10) имеет единственное решение и справедливы оценки

$$\|\varphi_n - \varphi\| \leq M\|P_n S\varphi - S\varphi\|$$

с некоторой положительной константой M, зависящей от S и K.

Доказательство. Уравнение (10) эквивалентно следующему уравнению

$$S^{-1}P_n S(I - S^{-1}K)\varphi_n = S^{-1}P_n S S^{-1} f,$$

или, в другом виде,

$$Q_n(I - S^{-1}K)\varphi_n = Q_n S^{-1} f,$$

где $Q_n := S^{-1}P_n S : X \to X_n$, очевидно, проекционный оператор. Используя равенства $\|Q_n S^{-1}K - S^{-1}K\| = \|S^{-1}(P_n K - K)\|$ и $Q_n\varphi - \varphi = S^{-1}(P_n S\varphi - S\varphi)$ и применяя теорему 5, получаем требуемое утверждение. □

Суть утверждения 3 в том, что проекционный метод применяется непосредственно к уравнению первого рода, а результат о сходимости базируется на регуляризованном уравнении второго рода.

2. Метод коллокации

Метод коллокации для приближенного решения уравнения

$$A\varphi = f \tag{11}$$

состоит в нахождении приближенного решения из конечномерного подпространства посредством приравнивания значений функций в левой и правой частях уравнения (11) в конечном числе точек, называемых точками коллокации. Точнее, пусть $Y = C(G)$ и $A : X \to Y$ —

линейный ограниченный оператор. Пусть $X_n \subset X$ и $Y_n \subset Y$ — последовательности подпространств таких, что $\dim Y_n = \dim X_n = n$. Выберем n точек $x_1^{(n)}, \ldots, x_n^{(n)}$ в области G (мы также будем писать для упрощения $x_1, \ldots, x_n$ вместо $x_1^{(n)}, \ldots, x_n^{(n)}$) так, чтобы подпространство Y_n однозначно определялось по этим точкам. Тогда метод коллокации решения уравнения (11) состоит в нахождении приближенного решения $\varphi \in X_n$, удовлетворяющего уравнениям

$$(A\varphi_n)(x_j) = f(x_j), \quad j = 1, \ldots, n. \tag{12}$$

Пусть $X_n = \operatorname{span}\{u_1, \ldots, u_n\}$ — линейная оболочка элементов u_j (базисных функций). Выразим элемент φ_n в виде линейной комбинации

$$\varphi_n = \sum_{k=1}^{n} \gamma_k u_k.$$

Тогда метод коллокации (12) эквивалентен системе линейных алгебраических уравнений

$$\sum_{k=1}^{n} \gamma_k (Au_k)(x_j) = f(x_j), \quad j = 1, \ldots, n.$$

Метод коллокации можно рассматривать как вариант проекционного метода. Приравнивание левой и правой частей уравнения в конечном числе точек коллокации в форме (12) эквивалентно уравнению проекционного метода $P_n A\varphi_n = P_n f$ с некоторым оператором интерполяции P_n, являющимся проекционным оператором (задача 1). Действительно, если в качестве узлов интерполяции взять точки коллокации и выбрать, например, интерполяцию многочленами $n-1$ порядка,то в силу единственности интерполяционного многочлена, построенного по n различным узлам (многочлен Лагранжа), получаем вариант проекционного метода, где в качестве проектора P_n выбран оператор интерполяции многочленами $n-1$ порядка по узлам, совпадающими с точками коллокации. Отметим, что один и тот же метод коллокации можно представить в виде проекционного метода с различными проекторами P_n.

Рассмотрим интегральное уравнение 1-го рода с логарифмической особенностью ядра:

$$S\varphi := \frac{1}{2\pi} \int_0^{2\pi} \{\ln(4\sin^2(\frac{t-\tau}{2})) + K(t,\tau)\} \varphi(\tau)\, d\tau = f(t), \quad 0 \le t \le 2\pi, \tag{13}$$

где K — периодическая и аналитическая функция двух переменных.

Оператор S отображает пространство тригонометрических полиномов в себя. Выберем точки $t_j = j\pi/n$, $j = 0, \ldots, 2n-1$ в качестве точек коллокации.

Задача 3. Используя утверждение 3, доказать, что метод коллокации для интегрального уравнения (13) сходится экспоненциально, если K и f — аналитические функции.

Пример 1. Рассмотрим интегральное уравнение (13) с ядром

$$K(t,\tau) = \ln\{1 + b^2 - (1-b^2)\cos(t+\tau)\} + 1$$

и правой частью

$$f(t) = 2 - e^{\cos t}\, \cos(\sin t) - e^{c\cos t}\, \cos(c\sin t), \quad 0 \le t \le 2\pi,$$

где $c = (1-b)/(1+b)$, $0 < b \le 1$.

Это уравнение однозначно разрешимо. Точное решение имеет вид

$$\varphi(t) = e^{\cos t}\, \{\cos t\, \cos(\sin t) - \sin t\, \sin(\sin t)\}, \quad 0 \le t \le 2\pi.$$

Значения погрешности в трех точках при различных n для приближенного решения, найденного по методу коллокации представлены в таблице.

	n	$t=0$	$t=\pi/2$	$t=\pi$
$b=0.5$	4	-0.62927694	-0.19858702	0.20711238
	8	-0.02649085	0.00602690	0.01355414
	16	-0.00000724	0.00000110	0.00000457
	32	0.0	0.0	0.0
$b=0.2$	4	-0.77913195	-0.22694775	0.20990539
	8	-0.08896065	0.02823001	0.03008962
	16	-0.00167596	0.00010607	0.00000226
	32	-0.00000114	0.00000009	-0.00000001

3. Метод Галеркина

Для операторных уравнений в гильбертовых пространствах проекционный метод, строящийся с помощью ортопроекторов на конечномерные подпространства, приводит к *методу Галеркина.* Пусть X и Y — гильбертовы пространства и $A: X \to Y$ инъективный линейный ограниченный оператор. Пусть $X_n \subset X$ и $Y_n \subset Y$ последовательности подпространств таких, что $\dim X_n = \dim Y_n = n$, и пусть $P_n : Y \to Y_n$ последовательность ортопроекторов. Тогда $\varphi_n \in X_n$ будет приближенным решением уравнения $A\varphi = f$ с помощью проекционного метода, образованного посредством выбора подпространств X_n и проекторов P_n, тогда и только тогда, когда

$$(A\varphi_n, g) = (f, g) \tag{14}$$

для любого $g \in Y_n$. Действительно, выражение (14) эквивалентно уравнению $P_n(A\varphi_n - f)$. Уравнение (14) называется *уравнением Галеркина.*

В литературе общий случай метода Галеркина называют *методом Петрова – Галеркина,* а специальный случай, когда $X = Y$ и $X_n = Y_n$, называют *методом Бубнова – Галеркина.* Если оператор A самосопряжен и положительно определен, метод Бубнова – Галеркина совпадает с методом Релея – Ритца. Напомним, что линейный ограниченный оператор $A : X \to X$ самосопряжен, если $A = A^*$, и положительно определен, если, кроме того, справедливо неравенство

$$(A\varphi, \varphi) > 0 \tag{15}$$

для всех $\varphi \in X$, $\varphi \neq 0$. Положительно определенный оператор, очевидно, инъективен. Мы можем определить новое скалярное произведение в пространстве X посредством формулы

$$(\varphi, \psi)_A := (A\varphi, \psi) \tag{16}$$

с нормой

$$\|\varphi\|_A = \sqrt{(A\varphi, \varphi)_A},$$

называемой *энергетической нормой.* Рассмотрим так называемый функционал энергии, определенный на пространстве X:

$$E(\varphi) := (A\varphi, \varphi) - 2\Re\,(f, \varphi) \tag{17}$$

для любого $\varphi \in X$ и фиксированного элемента $f \in X$. Для $f \in A(X)$ мы можем записать

$$E(\varphi) - E(A^{-1}f) = \|\varphi - A^{-1}f\|_E^2.$$

Следовательно, решение уравнения $A\varphi = f$ эквивалентно минимизации функционала энергии на пространстве X. Впервые предложенный в работах Релея и Ритца метод Релея – Ритца состоит в минимизации функционала энергии E на конечномерном подпространстве X_n. В свою очередь он эквивалентен нахождению наилучшего приближения из подпространств X_n в энергетической норме к точному решению. Элемент наилучшего приближения существует и единственен и определяется соотношением

$$(\varphi_n, g)_A = (A^{-1}f, g)_A \quad \text{для всех} \quad g \in X_n.$$

Но это совпадает с уравнением Галеркина (14), то есть метод Релея – Ритца является частным случаем метода Галеркина. Отметим, что сходимость метода Релея – Ритца имеет место, если выполняется условие аппроксимации (1) относительно энергетической нормы.

Допустим, что подпространства X_n и Y_n являются линейными оболочками базисных и тестовых функций: $X_n = \operatorname{span}\{u_1, \ldots, u_n\}$, $Y_n = \operatorname{span}\{v_1, \ldots, v_n\}$. Представим φ_n в виде линейной комбинации базисных функций

$$\varphi_n = \sum_{k=1}^{n} \gamma_n u_n$$

и подставим это выражение в формулу (14). В результате получим эквивалентную систему линейных алгебраических уравнений порядка n

$$\sum_{k=1}^{n} \gamma_k (Au_k, v_j) = (f, v_j), \quad j = 1, \ldots, n \tag{18}$$

относительно неизвестных коэффициентов γ_n.

Таким образом, метод Галеркина может быть задан либо с помощью выбора подпространств X_n и проекторов P_n (или подпространств Y_n вместо проекторов P_n, что эквивалентно, поскольку проекторы ортогональные), либо посредством выбора базисных и тестовых функций. Важно отметить, что теоретически сходимость метода Галеркина не зависит от конкретного выбора базисных функций в подпространстве X_n и тестовых функций в подпространстве Y_n. Поэтому для теоретического исследования метода Галеркина фиксируют только подпространства X_n и проекторы P_n. На практике наоборот удобнее сразу выбрать базисные и тестовые функции, так как без них невозможно формирование матрицы коэффициентов по формуле (18).

Обычно скалярное произведение имеет форму интеграла. Следовательно, в случае одномерных интегральных уравнений система (18) требует вычисления двойного интеграла для каждого элемента матрицы, в отличие от метода коллокации, в котором требуется вычислять однократные интегралы. Это является наиболее существенным недостатком метода Галеркина, поскольку процедура вычисления двойного интеграла более трудоемка, чем однократного. В то же время метод Галеркина является более устойчивым к погрешностям вычислений.

Большие преимущества метода Галеркина связаны с использованием ортогональных проекторов и структуры гильбертова пространства. Например, для интегральных уравнений второго рода утверждение 1 можно всегда применить, поскольку условие аппроксимации (1) влечет поточечную сходимость ортогональных проекторов. Следовательно, имеет место следующее утверждение.

Теорема 6. Пусть $K : X \to Y$ компактный оператор и $I + K$ инъективный. Тогда метод Бубнова – Галеркина сходится для оператора $I + K$.

Для уравнений первого рода имеется несколько общих результатов.

Определение 3. Линейный ограниченный оператор $A : X \to X$ в гильбертовом пространстве X называется *коэрцетивным*, если существует константа c_0 такая, что

$$\Re\,(A\varphi, \varphi) \geq c_0\|\varphi\|^2 \tag{19}$$

для всех $\varphi \in X$.

Докажем следующее важное утверждение, которое носит название теоремы *Лакса – Мильграма.*

Теорема 7. В гильбертовом пространстве X коэрцетивный оператор $A : X \to X$ имеет ограниченный обратный оператор $A^{-1} : X \to X$.

Доказательство. Используя неравенство Коши – Буняковского, имеем следующие оценки:

$$\|A\varphi\|\|\varphi\| \geq \Re\,(A\varphi, \varphi) \geq c_0\|\varphi\|^2,$$

и, следовательно,

$$\|A\varphi\| \geq c_0\|\varphi\| \tag{20}$$

для всех $\varphi \in X$. Из уравнения (20) следует, что если $A\varphi = 0$, то $\varphi = 0$. Это означает, что — A инъективный оператор.

Покажем, что образ $A(X)$ замкнут. Пусть $\psi \in \overline{A(X)}$ и пусть $\{\psi_n\}$ последовательность из $A(X)$ такая, что $\psi_n \to \psi$, $n \to \infty$. Тогда мы можем написать $\psi_n = A\varphi_n$ с некоторыми элементами $\varphi_n \in X$ и из оценки (20) находим, что

$$c_0\|\varphi_n - \varphi_m\| \leq \|\psi_n - \psi_m\|$$

для всех натуральных чисел n, m. Следовательно, $\{\varphi_n\}$ фундаментальная последовательность в X и она сходится $\varphi_n \to \varphi$, $n \to \infty$ к некоторому элементу $\varphi \in X$. Таким образом, $\psi = A\varphi$ и равенство $A(X) = \overline{A(X)}$ доказано.

Поскольку теперь известно, что $A(X)$ замкнутое подпространство, обозначим через $P : A \to A(X)$ соответствующий ортопроектор. Возьмем произвольный элемент $f \in X$. Тогда $Pf - f \perp A(X)$. В частности, $(Pf - f, A(Pf - f)) = 0$. Из условия (19) получаем, что $f = Pf \in A(X)$. Следовательно, оператор A сюръективен и по теореме Банаха имеет ограниченный обратный оператор A^{-1}. Очевидно, что оценка для нормы обратного оператора $\|A^{-1}\| \leq 1/c_0$ есть прямое следствие оценки (20). □

Теорема 8. Пусть $A : X \to X$ коэрцетивный оператор. Тогда метод Бубнова – Галеркина сходится для оператора A.

Доказательство. Как и в предыдущем доказательстве, имеем оценки

$$\|P_n A\varphi\|\,\|\varphi\| \geq \Re\,(P_n A\varphi, \varphi) = \Re\,(A\varphi, \varphi) \geq c_0\,\|\varphi\|^2$$

для всех $\varphi \in X_n$, так как ортогональные проекторы — самосопряженные операторы (см. задачу 4). Следовательно,

$$\|P_n A\varphi\| \geq c_0\|\varphi\| \tag{21}$$

для всех $\varphi \in X_n$. Отсюда следует, что $A_n = P_n A : X_n \to X_n$ инъективны и, следовательно, обратимы (поскольку это конечномерные операторы). Для всех $\varphi \in X$, используя неравенство (21) и теорему 2, имеем оценки

$$c_0\|A_n^{-1}P_n A\varphi\| \leq \|P_n A A_n^{-1} P_n A\varphi\| = \|P_n A\varphi\| \leq \|A\|\|\varphi\|.$$

Тогда $\|A_n^{-1}P_n A\| \leq \|A\|/c_0$ для любого n, и утверждение теоремы следует из теоремы 3. □

Теоремы 7 и 8 могут быть распространены на коэрцетивные операторы, действующие из гильбертова пространства X в сопряженное пространство X^* (см. задачу 7).

Рассмотрим еще один специальный случай метода Галеркина.

Теорема 9. Пусть X и Y — гильбертовы пространства и $A : X \to Y$ — инъективный линейный ограниченный оператор. Пусть $X_n \subset X$ — конечномерное подпространство. Тогда для каждого $f \in Y$ существует единственный элемент $\varphi_n \in X_n$ такой, что

$$\|A\varphi_n - f\| = \inf_{\psi \in X_n} \|A\psi - f\|.$$

Такой способ нахождения приближенного решения φ_n для уравнения $A\varphi = f$ называется методом наименьших квадратов относительно подпространств X_n и совпадает с методом Петрова – Галеркина с выбором подпространств X_n и $Y_n := A(X_n)$.

Доказательство. Очевидно, φ_n будет приближенным решением для уравнения $A\varphi = f$, найденным по методу наименьших квадратов относительно подпространств X_n тогда и только тогда, когда элемент $A\varphi_n$ из Y_n будет наилучшим приближением для элемента f. Как известно (см. [4]) элемент наилучшего приближения существует и единствен. Инъективность оператора A влечет единственность решения, найденного по методу наименьших квадратов. Элемент наилучшего приближения $A\varphi_n$ однозначно определяется с помощью условия ортогональности $(A\varphi_n - f, g) = 0$ для всех $g \in Y_n$. Это эквивалентно уравнению Галеркина (14) для подпространств X_n и Y_n. $\square$

Отметим, что если решается интегральное уравнение методом наименьших квадратов, то для построения каждого элемента матрицы конечномерной системы (18) требуется вычисление тройного интеграла, а для вычисления каждого элемента правой части — двойного интеграла, так как $g_j = Au_j$, $j = 1, \ldots, n$. Таким образом, с точки зрения объема вычислений метод наименьших квадратов уступает по эффективности методу коллокации и методу Галеркина, если тестовые функции в последнем задаются непосредственно (без дополнительных вычислений).

Применим метод Галеркина для решения интегральных уравнений. Сначала рассмотрим метод Бубнова – Галеркина для решения уравнений второго рода

$$\varphi(x) - \int_G K(x,y)\,\varphi(x)\,\mathrm{d}y = f(x), \quad x \in G$$

с непрерывным или слабо сингулярным ядром K в гильбертовом пространстве $L^2(G)$. Пусть $X_n = \mathrm{span}\{u_1, \ldots, u_n\}$ — линейная оболочка базисных функций — n-мерное подпространство $L^2(G)$. Будем искать приближенное решение в виде

$$\varphi_n = \sum_{k=1}^{n} \gamma_k u_k.$$

Тогда уравнения, получаемые по методу Бубнова – Галеркина, будут иметь следующую форму

$$\sum_{k=1}^{n} \gamma_k \left\{ \int_G u_k(x)\overline{u_j(x)}\,\mathrm{d}x - \int_G \int_G K(x,y)\,u_k(y)\,\overline{u_j(x)}\,\mathrm{d}x\,\mathrm{d}y \right\}$$
$$= \int_G f(x)\overline{u_j(x)}\,\mathrm{d}x, \quad j = 1, \ldots, n. \tag{22}$$

При выборе простых базисных функций первое слагаемое в уравнении (22) вычисляется аналитически. Второе слагаемое в левой части уравнения (22) и элементы столбца правой части чаще всего вычисляются численно с помощью квадратурных формул. Если решается одномерное интегральное уравнение (например, G — интервал прямой), то при формировании элементов матрицы системы (22) требуется вычислять двойные интегралы. Если исходное уравнение двумерное (например, G — область на плоскости), то при формировании элементов матрицы системы требуется вычислять уже четырехкратные интегралы, при решении трехмерного интегрального уравнения — шестикратные интегралы. Таким образом, при переходе к задачам большей размерности вычислительные трудности резко возрастают.

Для интегральных уравнений, относительно периодических функций, использующих тригонометрические полиномы, отметим, что вместо рассмотрения метода Бубнова – Галеркина в $L^2(0, 2\pi)$ можно использовать пространства Соболева $H^s(0, 2\pi)$, $s \geq 0$. Поскольку тригонометрические базисные функции оргогональны в каждом из указанных пространств Соболева, ортогональные проекторы на подпространство тригонометрических полиномов всегда выражаются через отрезки рядов Фурье. Следовательно, уравнения метода Бубнова – Галеркина будут совпадать при всех $s \geq 0$. В частности, это означает, что мы автоматически получаем сходимость метода Бубнова – Галеркина, относительно всех соболевских норм, если ядро интегрального уравнения и правая часть — аналитические функции.

Задача 4. Покажите, что ортогональные проекторы являются самосопряженными операторами.

Задача 5. Пусть $A : X \to X$ — ограниченный положительный оператор в гильбертовом пространстве X. Выберем $u_0 \in X$ и определим $u_j = Au_{j-1}$, $j = 1, \ldots, n-1$. Рассмотрим уравнение $A\varphi = f$ и выберем подпространства $X_n = \operatorname{span}\{u_0, \ldots, u_{n-1}\}$. Покажите, что уравнения метода Бубнова – Галеркина однозначно разрешимы при всех n ($n > 0$). Более того, если f принадлежит замыканию линейной оболочки функций $\operatorname{span}\{A^j u_0 : j = 0, 1, \ldots\}$, то приближенные решения φ_n, найденные по методу Бубнова – Галеркина, сходятся к точному решению уравнения $A\varphi = f$. Покажите, что в специальном случае $u_0 = f$ приближенные решения φ_n могут быть вычислены итерационным способом с помощью формул $\varphi_0 = 0$, $p_0 = f$ и

$$\begin{aligned}
\varphi_{n+1} &= \varphi_n - \alpha_n p_n, \\
p_n &= r_n + \beta_{n-1} p_{n-1}, \\
r_n &= r_{n-1} - \alpha_{n-1} A p_{n-1}, \\
\alpha_{n-1} &= (r_{n-1}, p_{n-1})/(Ap_{n-1}, p_{n-1}), \\
\beta_{n-1} &= -(r_n, Ap_{n-1})/(Ap_{n-1}, p_{n-1}).
\end{aligned}$$

Здесь через r_n обозначена невязка $r_n = A\varphi_n - f$. Такой метод называется *методом сопряженных градиентов*.

Задача 6. Пусть выполнены условия теоремы 6 и φ_n — приближенное решение, найденное по методу Бубнова – Галеркина. Рассмотрим модифицированное приближенное решение $\tilde{\varphi}_n$, определенное посредством формул

$$\tilde{\varphi}_n := A\varphi_n + f.$$

Покажите, что

$$\|\tilde{\varphi}_n - \varphi\| \leq c_n \|\varphi_n - \varphi\|,$$

где $c_n \to 0$ при $n \to \infty$. Такая более быстрая сходимость модифицированного метода называется *суперсходимостью*.

Замечание. Покажите сначала, что $\tilde{\varphi}_n - AP_n\tilde{\varphi}_n = f$, что $\|AP_n - A\| \to 0$ при $n \to \infty$, и, наконец, что $\tilde{\varphi}_n - \varphi = (I - AP_n)^{-1}(AP_n - A)(\varphi - P_n\varphi)$.

Задача 7. Линейный ограниченный оператор $A : X \to X^*$, действующий из гильбертова пространства X в сопряженное пространство X^*, называется *коэрцетивным*, если существует постоянная $c_0 > 0$ такая, что

$$\Re\,(A\varphi)\,(\varphi) \geq c_0 \|\varphi\|^2$$

для всех $\varphi \in X$. Сформулируйте и докажите обобщения теорем 7 и 8 на этот случай.

Задача 8. Примените результаты задачи 7 к интегральному уравнению (13) в пространстве Соболева $H^{-1/2}(0, 2\pi)$.

Библиографический список

1. Kress R. Linear integral equations. Applied mathematical sciences. 82, 1989, Springer–Verlag New York Inc.

2. Бахвалов Н.С. Численные методы. – М.: Наука, 1975.

3. Садовничий В.А. Теория операторов. – М.: Наука, 1986.

4. Колмогоров А.Н., Фомин С.В. Элементы теории функций и функционального анализа. – М.: Наука, 1981.

5. Марчук Г.И., Агошков В.И. Введение в проекционно–сеточные методы. – М.: Наука, 1981.