

СПЕКТРАЛЬНЫЕ МЕТОДЫ ПОЛИНОМИАЛЬНОЙ ИНТЕРПОЛЯЦИИ И АППРОКСИМАЦИИ

© 2025 г. В. П. Варин^{1,*}¹ 125047 Москва, Миусская пл., 4, Институт прикладной математики им. М.В. Келдыша РАН, Россия

*e-mail: varin@keldysh.ru

Поступила в редакцию 22.05.2024 г.

Переработанный вариант 18.10.2024 г.

Принята к публикации 08.11.2024 г.

Классическая задача интерполяции и аппроксимации функций полиномами здесь рассматривается как частный случай спектрального представления функций. Этот подход был ранее развит нами для ортогональных полиномов Лежандра и Чебышёва. Здесь в качестве базисных функций мы используем фундаментальные полиномы Ньютона. Показано, что спектральный подход имеет вычислительные преимущества по сравнению с методом разделенных разностей. В ряде задач интерполяции Ньютона и Эрмита неразличимы при нашем подходе и вычисляются по одним и тем же формулам. Также вычислительные алгоритмы, предложенные нами ранее с использованием ортогональных полиномов, переносятся без изменений на полиномы Ньютона и Эрмита. Библ. 13.

Ключевые слова: спектральные методы, полиномы Ньютона и Эрмита, интерполяция и аппроксимация.

DOI: 10.31857/S0044466925020027, **EDN:** CCAYAD

1. ВВЕДЕНИЕ

Термины “интерполяция” и “аппроксимация” в численном анализе имеют настолько близкий смысл, что часто встречаются вместе (см., например, [1]).

Интерполяция подразумевает моделирование некоторой сложной или эмпирической функции с помощью достаточно простого выражения (полинома или рациональной функции), по которому эта функция может быть вычислена (интерполирована) с достаточной точностью в некоторой ограниченной области. Специальные (и даже элементарные) функции часто имеют именно такое внутреннее представление в различных языках программирования и системах компьютерной алгебры (CAS).

С другой стороны, полученное выражение аппроксимирует нужную функцию в данной области, так что это и задача аппроксимации.

Аппроксимация также понимается как построение приближенного решения некоторой задачи, которое априори, разумеется, неизвестно.

Хотя эти две задачи весьма близки с точки зрения конечного результата, однако их решение может весьма различаться.

В первом случае нам известны точные (или полагаемые таковыми) значения функции в выбранных узлах. Задача аппроксимации здесь — это построение (как правило, с помощью конечных разностей) нужной формулы, т.е. *интерполанта*, аппроксимирующего данную функцию и совпадающего с ней в этих узлах.

Во втором случае, даже если используется та же формула для аппроксимации, задача состоит в вычислении коэффициентов некоторого разложения функции, по которому функция может быть (приближенно) вычислена в выбранных узлах и/или в нужной области.

На самом деле, интерполяция и аппроксимация функций всегда содержат два аспекта одной и той же задачи. Дело в том, что функция в численном анализе — это всегда некоторая таблица числовых значений, которая кодирует функцию в выбранной области.

В случае если функция представлена таблицей своих значений в выбранных узлах, будем говорить, что это *коллокационное представление* функции. А в случае, если функция представлена таблицей коэффициентов разложения функции по некоторому набору базисных функций, будем говорить, что это *спектральное представление* функции.

Оба этих представления используются для аппроксимации периодических функций тригонометрическими полиномами и связаны с классическим дискретным анализом Фурье. При этом прямое преобразование Фурье

переводит таблицу значений функции в (равноотстоящих) узлах в таблицу ее коэффициентов Фурье, а обратное преобразование Фурье — наоборот.

Как показано в [2] и [3], этот формализм переносится без существенных изменений на практически произвольные наборы базисных функций, которые должны лишь удовлетворять некоторым необременительным условиям. В частности, это справедливо для всех типов (классических) ортогональных полиномов.

Эта статья является продолжением работ [2], [3], где был предложен спектральный метод решения дифференциальных и функциональных уравнений. Здесь мы рассматриваем одно из возможных направлений развития этой темы: интерполяцию и аппроксимацию регулярных функций на конечном интервале (но не только) с помощью обычных полиномов.

Разумеется, это классическая задача, которую мы в некоторых ее аспектах уже решили. Так, в [2] мы аппроксимировали решения голономных ОДУ с помощью полиномов Лежандра, а в [3] мы решали краевые задачи для дифференциальных и функциональных уравнений с помощью полиномов Чебышёва.

Однако в этой статье мы рассматриваем “еще более классическую” постановку этой задачи, т.е. задачу интерполяции и аппроксимации функций полиномами Ньютона.

Один из мотиваторов этого исследования — это демонстрация настолько простого и понятного спектрального метода решения подобных задач, насколько это, на наш взгляд, вообще возможно.

Личный опыт автора (одно время работавшего в КБ инженером) показывает, что инженеры редко используют ортогональные полиномы при решении практических задач, а предпочитают использовать обычные полиномы, а иногда и просто отрезки степенных рядов. В то же время полиномы Ньютона и связанные с ними конечные и разделенные разности естественным образом появляются в процессе приближения функций полиномами, и поэтому они часто встречаются в инженерных расчетах.

Другим мотиватором было желание показать, что наш “спектральный подход” к этой классической задаче (интерполяции и аппроксимации полиномами) обладает в ряде случаев явными вычислительными преимуществами по сравнению с методом разделенных разностей и может давать больше и с меньшими затратами.

Эта последняя цель представляется недостижимой на фоне той роли, которую разделенные разности традиционно играют в этих задачах. Так, в рецензии [4] на книгу Стеффенсена [5] отмечается, что “Dr. Steffensen makes the fullest use of Newton’s Divided Difference Formula, the most powerful tool ever forged for use in Interpolation”.

Рассмотрим функцию $y(x)$, заданную пока что в некоторых *различных* точках, (как правило) принадлежащих некоторому интервалу:

$$x_1, x_2, \dots, x_N \in [a, b] \subset \mathbb{R}. \quad (1)$$

То есть пусть задана таблица $\{y_1 = y(x_1), \dots, y_N = y(x_N)\}$.

Напомним, что *разделенные разности*, обозначаемые $y[\dots]$ и вычисляемые с помощью конечных разностей, являются функциями таблиц значений функции в выбранных узлах и определяются рекурсивным образом:

$$y[x_k] = y_k, \quad y[x_1, x_2, \dots, x_k] = \frac{y[x_1, x_2, \dots, x_{k-1}] - y[x_2, \dots, x_k]}{x_1 - x_k}.$$

Тогда *интерполяционный полином Ньютона* этой функции дается формулой

$$y_N(x) = y[x_1] p_0(x) + y[x_1, x_2] p_1(x) + \dots + y[x_1, x_2, \dots, x_N] p_{N-1}(x), \quad (2)$$

где $p_n(x)$ — это *фундаментальные полиномы Ньютона*,

$$p_n(x) = \prod_{k=1}^n (x - x_k), \quad (3)$$

т.е. $p_0(x) = 1$, $p_1(x) = x - x_1$, $p_2(x) = (x - x_1)(x - x_2)$ и т.д.

Хотя формула (2) явно зависит от порядка следования точек (1), результат интерполяции, полином $y_N(x)$, как известно, зависит только от самого набора точек (1), но не от порядка их следования (с учетом того, что полиномы Ньютона (3) зависят от порядка точек (1)).

Таким образом, полином $y_N(x) \in \mathcal{P}_N$ (где \mathcal{P}_N — это множество полиномов степени меньше N) совпадает с полиномом лагранжевой интерполяции, инвариантность которого по отношению к порядку следования точек очевидна.

Одно из часто отмечаемых достоинств интерполяционного полинома Ньютона — это возможность добавлять новые узлы интерполяции к уже имеющимся, не пересчитывая при этом уже найденные коэффициенты.

Однако в этом способе вычисления интерполяционного полинома присутствуют и некоторые недостатки. Так, результат интерполяции регулярной на интервале $[a, b]$ функции $y(x)$ — это полином с коэффициентами,

которые регулярным образом зависят от выбранных узлов. Например, для функции $y(x) = 1/(1+x)$ и $N = 3$ получим

$$y_N(x) = \frac{1}{1+x_1} - \frac{x-x_1}{(1+x_1)(1+x_2)} + \frac{(x-x_1)(x-x_2)}{(1+x_1)(1+x_2)(1+x_3)}.$$

Последняя формула не только не зависит от порядка следования узлов, но и не требует их различия. Как нетрудно догадаться (устремив, например, $x_1 \rightarrow x_2$), в случае совпадения некоторых узлов мы получим частные случаи эрмитовой интерполяции. Например, при $x_1 = x_2 = x_3 = 0$ получим отрезок ряда Тейлора в нуле, а при $x_1 = x_2 = x_3 = 1$ — отрезок ряда Тейлора в единице, и т.д.

Однако разделенные разности нельзя вычислить при совпадении узлов даже в рациональной арифметике. А в случае, когда некоторые узлы близки, но различны, этот алгоритм становится плохо обусловленным при вычислениях в плавающей арифметике.

И, наконец, интерполяция Эрмита не может быть вычислена по формуле (2), хотя сам интерполяционный полином ее неявно содержит. Интерполяция Эрмита обычно вычисляется с помощью т.н. *конфлюентных разделенных разностей* (см. [6, р. 13]), т.е. неопределенности $0/0$, возникающие при совпадении узлов, заменяются значениями производных функции в выбранных узлах.

Как отмечалось в [1], “The Newton representation has a permanence property, and this is characteristic of Fourier series and other orthogonal and biorthogonal expansions.”

Это наблюдение как нельзя лучше соответствует нашим выводам в [3] о том, что полноценный “анализ Фурье” (т.е. совместное использование коллокационного и спектрального представления функций) можно построить для весьма общих наборов базисных функций. В этой статье мы строим анализ Фурье для полиномов Ньютона в качестве такого набора.

В разд. 2 мы приводим формулы для всех операторов, необходимых для аппроксимации функций полиномами Ньютона, пригодные также для эрмитовой интерполяции. В нашем формализме эти два подхода (почти) неразличимы.

Таким образом, все задачи, которые мы решали в [2] и [3] с помощью ортогональных полиномов Лежандра и Чебышёва, могут решаться (и по тем же формулам) с помощью полиномов Ньютона или Эрмита (не путать с ортогональными полиномами Эрмита).

В следующих разделах мы приводим ряд примеров решения задач этим способом.

2. СПЕКТРАЛЬНОЕ ПРЕДСТАВЛЕНИЕ РАЗДЕЛЕННЫХ РАЗНОСТЕЙ

Узлы в ньютоновой и эрмитовой интерполяциях могут выбираться без каких-либо ограничений. Но для удобства сравнения данного спектрального представления функций с введенными нами ранее (см. [2, 3]), здесь мы рассматриваем то же линейное функциональное пространство.

Любой интервал $[a, b] \subset \mathbb{R}$ можно без ограничения общности свести к интервалу $[0, 1]$. Поэтому в этой статье мы рассматриваем функции, регулярные на интервале $(0, 1)$, но могущие иметь особенности по краям интервала (пространство функций \mathcal{H}).

Это последнее допущение, казалось бы, противоречит заявленной цели: построению полиномиальных аппроксимаций на интервале. Однако, во-первых, такой выбор позволяет решать задачи, имеющие особенности по краям интервала, как мы это делали в [2] и [3]. Во-вторых, полиномиальные аппроксимации при увеличении размерности не обязаны сходиться на выбранном интервале к данной функции даже поточечно, а могут расходиться именно по краям интервала (феномен Рунге, см. ниже).

Спектральное представление функции $y(x) \in \mathcal{H}$ — это сопоставление этой функции формального ряда типа ряда Ньютона (2), отрезок которого имеет вид

$$y_N(x) = \sum_{n=0}^{N-1} a_n p_n(x), \quad (4)$$

где $p_n(x)$ — это полиномы Ньютона (3), что предполагает наличие некоторого фиксированного набора точек $\{x_n \in \mathbb{C}, n \in \mathbb{N}\}$. Кроме того, что точки фиксированы и упорядочены (возможно, с повторением), на этот набор более не накладывается никаких ограничений.

В случае, когда все узлы различны, полином (4) совпадает (по определению) с полиномом Ньютона (2), посчитанным с помощью разделенных разностей.

Назовем \mathcal{A}_N пространство коэффициентов разложений (4), т.е.

$$\langle a_0, a_1, \dots, a_{N-1} \rangle^t \in \mathcal{A}_N. \quad (5)$$

Умножение функции $y(x) \in \mathcal{H}$ на x дает функцию $xy(x) \in \mathcal{H}$, и это линейное отображение индуцирует линейное отображение X в пространстве \mathcal{A}_N . Иными словами,

$$X: \{a_n\} \rightarrow \{b_n\}, \quad xy(x) = \sum_{n=0}^{N-1} b_n p_n(x),$$

где равенство понимается в проективном смысле. Иными словами, *две функции равны в проективном смысле, если они имеют одинаковый набор коэффициентов Фурье в \mathcal{A}_N* .

Как показано в [2] и [3], отображение X (наряду с отображением D , см. ниже) играет ключевую роль в аппроксимации дифференциальных и разностных операторов. Поэтому укажем общий способ вычисления матрицы этого отображения для произвольных наборов базисных функций.

В случае, когда наборы базисных функций — это (классические) ортогональные полиномы, матрица отображения X — это всегда транспонированная матрица Якоби, ассоциированная с данной системой полиномов (см. [2, 7]). На самом деле, это всегда так, просто в общем случае эти матрицы (насколько нам известно) ранее не вводились.

Введем вектор $p = \langle p_0(x), p_1(x), \dots, p_{N-1}(x) \rangle^t$ и найдем матрицу J (которую назовем *матрицей Якоби*), такую, что выполняется *тождество Якоби*

$$xp = Jp + g \langle 0, 0, \dots, p_N(x) \rangle^t, \quad (6)$$

где g — это некоторая константа, зависящая от выбранного набора базисных функций и, возможно, от N .

В нашем случае из тождества (6) следует, что спектр матрицы J — это точки x_1, \dots, x_N , т.е. корни полинома $p_N(x)$. Поэтому матрица J аннулируется полиномом $p_N(x)$ (теорема Кэли—Гамильтона).

Транспонируя тождество (6) и умножая его скалярно на вектор (5), получаем, что $X = J^t$, т.е. X — это всегда транспонированная матрица J , найденная из тождества Якоби.

В нашем случае это следует из (очевидного) факта, что две функции, $u(x)$ и $v(x)$, имеют один и тот же интерполяционный полином Ньютона (2) (или более общий (4)) тогда и только тогда, когда функция $(u(x) - v(x))/p_N(x)$ является регулярной в точках x_1, \dots, x_N .

Это последнее свойство полиномов Ньютона является аналогом свойства ортогональности для классических ортогональных полиномов и дает оператор проектирования в пространстве \mathcal{H} .

Матрица J для полиномов Ньютона (и Эрмита) находится из очевидного тождества

$$xp_{n-1}(x) = x_n p_{n-1}(x) + p_n(x), \quad n \in \mathbb{N}.$$

Причем всегда $g = 1$ для любого набора узлов. Структура матрицы J очевидна и не требует формализации. Приведем матрицу $X = J^t$ для $N = 4$:

$$X = \begin{bmatrix} x_1 & 0 & 0 & 0 \\ 1 & x_2 & 0 & 0 \\ 0 & 1 & x_3 & 0 \\ 0 & 0 & 1 & x_4 \end{bmatrix}.$$

Любопытно, что в статье в Википедии о разделенных разностях приведена матрица J (заимствованная из [8]) именно в этом обозначении (т.е. обозначена буквой J), но без привязки к тождеству Якоби. В [8] эта матрица вводилась другим способом и использовалась для вычисления разделенных разностей от полиномиальных и рациональных функций (Opitz' formula).

Матрица X для узлов $x_1 = x_2 = \dots = 0$ уже использовалась нами в [3] для вычисления формальных степенных разложений решений голономных ОДУ. Разумеется, это обстоятельство не является случайным, так как при этих узлах полином Ньютона (4) превращается в отрезок степенного ряда, что можно также трактовать как весьма частный случай полиномов Эрмита.

Вообще, разложение (4) объединяет в себе обычный степенной ряд в любой точке $t_0 = x_1 = x_2 = \dots$; ряд Ньютона в случае различных узлов; и интерполяции Эрмита произвольного вида, в случае совпадения групп узлов в данном наборе.

Это последнее обстоятельство следует из симметрии разложения (4) по отношению к порядку следования узлов. Собрав совпадающие узлы в начале разложения, получим отрезок ряда Тейлора в этой точке. И так для любой группы узлов.

Пусть $\Lambda = \{x_1, x_2, \dots, x_N\}$ — это совокупность выбранных узлов (спектр матрицы X). И пусть $e = \langle 1, 0, \dots, 0 \rangle^t$ — это единичный вектор в \mathcal{A}_N . Справедливо следующее

Предложение 1 (см. также [3]). Для любой функции $y(x)$, регулярной в узлах Λ , все коэффициенты ее разложения (4) даются вектором $y(X)e$.

Доказательство. По условию матрица $y(X)$ определена, так как по формуле Лагранжа–Сильвестра (см. [9]) функция от матрицы определяется значениями функции $y(x)$ и необходимого количества ее производных на спектре матрицы X .

Вектор Xe , очевидно, соответствует функции x . Это же справедливо для любой степени X^n , $n \leq N-1$, т.е. вектор $X^n e$ соответствует функции x^n . Например, для $n = 3$ получаем

$$x^3 = x_1^3 p_0(x) + (x_1^2 + x_1 x_2 + x_2^2) p_1(x) + (x_1 x_2 + x_3) p_2(x) + p_3(x),$$

где вместо x , очевидно, можно подставить матрицу X .

Но для общей функции от матрицы $y(X)$ всегда существует полином $q(x)$ степени $d \leq N-1$, такой, что $y(X) = q(X)$ (см. [9]). В данном случае это следует из того, что полином $p_N(x)$ аннулирует матрицу X .

Поэтому вектор $y(X)e = q(X)e$ дает спектральное разложение (4) в силу линейности проделанных операций. Ч.т.д.

На самом деле, мы получили несколько больше, чем заявлено в предложении 1. Полином $q(x)$ по построению — это не что иное, как спектральное разложение (4). Поэтому матрица $y_N(X) = y(X)$ по формуле (4) — это не что иное, как формула Лагранжа–Сильвестра (см. [9]), записанная в форме общего полинома Ньютона (или Эрмита, т.е. для произвольных узлов).

Иными словами, вычислив ньютонову (или эрмитову) интерполяцию функции $y(x)$, мы получаем также функцию от матрицы $y(X)$, и наоборот.

Пример. Рассмотрим рациональную функцию, определенную в узлах Λ . Пусть для простоты $y(x) = 1/(1+x)$. Найдем полином $y_N(x)$, совпадающий с $y(x)$ в точках $\{0, 1/2, 1\}$, и потребуем также совпадения первых производных по краям интервала $[0, 1]$, а в середине интервала потребуем совпадения двух производных. Тогда имеем всего семь условий, наложенных в трех точках, поэтому $N = 7$ и $\Lambda = \{0, 0, 1/2, 1/2, 1/2, 1, 1\}$, причем узлы можно брать в произвольном порядке.

Для рациональных функций, как известно (см. [9]), функция от матрицы получается формальной подстановкой матрицы вместо независимой переменной. Поэтому найдем вектор $s = (E+X)^{-1}e$, где E — это единичная матрица, и составим полином (4), где $a_{n-1} = s_n$, $n = 1, \dots, N$, т.е. компоненты вектора s . Получаем

$$y_N(x) = 1 - x + \frac{107}{108}x^2 - \frac{11}{12}x^3 + \frac{37}{54}x^4 - \frac{1}{3}x^5 + \frac{2}{27}x^6, \quad (7)$$

что дает $(E+X)^{-1} = y_N(X)$, а также нетипичную интерполяцию Эрмита функции $y(x) = 1/(1+x)$.

Заметим, что мы не вычисляли предварительно значения функции $y(x)$ и ее производных в узлах. Полученный интерполянт $y_N(x)$ автоматически обладает нужными свойствами. Причем так будет всегда, когда матрица $y(X)$ известна в явном виде.

Таким образом, вычисление общей функции от матрицы $y(X)$ эквивалентно вычислению лагранжевой, ньютоновой или эрмитовой интерполяции этой функции в выбранных узлах. Например, просуммировав матричный ряд Тейлора для $\exp(X)$, получим общую эрмитову интерполяцию функции $\exp(x)$. Однако есть и более короткий путь.

Пусть пока что все точки x_n в спектре Λ различны. Мы вспоминаем, что для нас функция $y(x)$ — это таблица числовых значений, т.е.

$$y(x) = \langle a_0, a_1, \dots, a_{N-1} \rangle^t \in \mathcal{A}_N \text{ или } y(x) = \langle y_1, y_2, \dots, y_N \rangle^t \in \mathcal{H}_N, \quad y_n = y(x_n),$$

где \mathcal{H}_N обозначает пространство конечномерных аппроксимаций функций $y(x) \in \mathcal{H}$ в виде таблиц их (приближенных) значений в узлах x_n .

Тогда интерполяционная формула (4) на спектре Λ дает аналог обратного (дискретного) преобразования Фурье (которое мы обозначим F^{-1}):

$$F^{-1} = [p_{k-1}(x_j)]_{1 \leq j, k \leq N}, \quad F^{-1}: \mathcal{A}_N \rightarrow \mathcal{H}_N, \quad (8)$$

так как это линейный оператор, действующий в пространстве “коэффициентов Фурье” \mathcal{A}_N (далее кавычки опускаем) и восстанавливающий функцию $y(x)$, как таблицу ее (приближенных) значений в выбранных узлах.

Таким образом, матрица преобразования F^{-1} (8) всегда известна в явном виде для любого набора (необязательно различных) узлов Λ . Приведем эту матрицу для размерности $N = 4$:

$$F^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & (x_2 - x_1) & 0 & 0 \\ 1 & (x_3 - x_1) & (x_3 - x_2)(x_3 - x_1) & 0 \\ 1 & (x_4 - x_1) & (x_4 - x_2)(x_4 - x_1) & (x_4 - x_3)(x_4 - x_2)(x_4 - x_1) \end{bmatrix}.$$

Общий член этой нижнетреугольной матрицы размерности N имеет вид:

$$(F^{-1})_{m,n} = \prod_{k=1}^{n-1} (x_m - x_{n-k}).$$

Для попарно различных узлов x_n (и только для них) матрица F^{-1} всегда обратима, и что весьма существенно, в явном виде. Получим матрицу $F = (F^{-1})^{-1}$ — аналог обычного преобразования Фурье, которое переводит таблицу значений функции $y(x) \in \mathcal{H}_N$ в таблицу ее значений $y(x) \in \mathcal{A}_N$. Приведем матрицу F для размерности $N = 3$:

$$F = \begin{bmatrix} 1 & 0 & 0 \\ \frac{1}{(x_1 - x_2)} & \frac{1}{(x_2 - x_1)} & 0 \\ \frac{1}{(x_1 - x_2)(x_1 - x_3)} & \frac{1}{(x_2 - x_1)(x_2 - x_3)} & \frac{1}{(x_3 - x_1)(x_3 - x_2)} \end{bmatrix}.$$

Общий член этой нижнетреугольной матрицы размерности N имеет вид:

$$F_{m,n} = \prod_{k \neq n}^m (x_n - x_k)^{-1}.$$

Таким образом, для попарно различных узлов x_n коэффициенты a_n интерполяционной формулы Ньютона (4) функции $y(x)$ получаются просто:

$$\langle a_0, a_1, \dots, a_{N-1} \rangle^t = F \langle y_1, y_2, \dots, y_N \rangle^t, \quad y_n = y(x_n). \quad (9)$$

Причем этот способ вычисления полинома $y_N(x)$ обладает всеми преимуществами метода разделенных разностей, т.е. не требует пересчета полученных коэффициентов при добавлении новых точек интерполяции.

Явный вид матрицы F , на самом деле, следует из явного вида разделенных разностей, который давно известен (см., например, [10]). Наш вклад здесь состоит в привязке этих классических результатов к анализу Фурье и использовании этих результатов для аппроксимации дифференциальных операторов (см. ниже).

Теперь функция от матрицы $y(X)$ представима как

$$y(X) = F \text{Diag} [y(x_1), \dots, y(x_N)] F^{-1}, \quad (10)$$

где $\text{Diag}[]$ обозначает диагональную матрицу.

Таким образом, аппроксимация функций с помощью полиномов Ньютона формально обладает всеми преимуществами аппроксимаций полиномами Чебышёва. То есть спектр матрицы X , а также матрицы F , F^{-1} , известны в явном виде. Напомним, что среди ортогональных полиномов только полиномы Чебышёва обладают этим свойством (см. [3]).

В частности, матрица $y(X)$ (для различных узлов) всегда диагонализуема. Ее собственные значения — это величины $y(x_1), \dots, y(x_N)$, а ее собственные векторы — это столбцы матрицы F .

До сих пор то, что мы излагали, можно интерпретировать, как переформулировку классических результатов в “спектральной форме”. Однако для решения краевых задач для линейных ОДУ этого недостаточно.

Дифференцирование функции $y(x) \in \mathcal{H}$ по x дает функцию $y'(x) \in \mathcal{H}$, и это линейное отображение индуцирует линейное отображение D в пространстве \mathcal{A}_N :

$$D: \{a_n\} \rightarrow \{c_n\}, \quad y'(x) = \sum_{n=0}^{N-1} c_n p_n(x).$$

Отображение D (насколько нам известно) ранее не приводилось, так как дифференцирование полинома Ньютона дает весьма громоздкую формулу с неясной (пока) структурой. Тем не менее матрица оператора D может быть представлена в (почти) явном виде. Приведем эту матрицу для размерности $N = 5$:

$$D = \begin{bmatrix} 0 & 1 & (x_1 - x_2) & (x_1 - x_3)(x_1 - x_2) & (x_1 - x_4)(x_1 - x_3)(x_1 - x_2) \\ 0 & 0 & 2 & x_1 + x_2 - 2x_3 & (x_1 - x_3)(x_1 - x_2) + (x_2 - x_4)(x_1 + x_2 - 2x_3) \\ 0 & 0 & 0 & 3 & x_1 + x_2 + x_3 - 3x_4 \\ 0 & 0 & 0 & 0 & 4 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

В общем виде (опуская весьма громоздкие промежуточные выкладки) эта верхнетреугольная матрица размера N дается следующими формулами:

$$D_{n,n} = 0, \quad n = 1, \dots, N; \quad D_{n,n+1} = n, \quad n = 1, \dots, N-1.$$

Остальные элементы матрицы D заполняются по диагоналям с помощью рекуррентной формулы:

$$D_{k,k+m} = \sum_{j=1}^k (x_j - x_{j+m-1}) D_{j,j+m-1}, \quad m = 2, \dots, N-1, \quad k = 1, \dots, N-m.$$

Матрица D для узлов $x_1 = x_2 = \dots = 0$ уже использовалась нами в [3] для вычисления формальных степенных разложений решений голономных ОДУ.

Приведенный формализм можно использовать в точности так же и для решения того же класса задач, которые мы решали в [2] и [3]. Отличие состоит в том, что априорная оценка погрешности аппроксимации решения полиномами Ньютона неизвестна. Весьма общая оценка остаточного члена полинома Ньютона, которую можно найти в учебниках, здесь не поможет, так как свойства полиномов Ньютона в общем случае (т.е. для произвольных узлов) заранее неизвестны. Поэтому нужные оценки получаются апостериори и каждый раз в индивидуальном порядке.

3. РЕШЕНИЕ ГОЛОНОМНЫХ ОДУ И ФЕНОМЕН РУНГЕ

Решение краевых задач для голономных ОДУ с помощью полиномов Ньютона в нашем спектральном представлении формально ничем не отличается от решения этих задач с помощью ортогональных полиномов (см. [2, 3]). Однако последние имеют явные вычислительные преимущества по сравнению с полиномами Ньютона.

Дело в том, что полиномы Ньютона (или Эрмита) — это всего лишь форма записи обычных полиномов, у которых априори не предполагается никаких свойств, кроме их привязки к заданному набору узлов. В то время как, например, для рядов по полиномам Чебышёва заранее известны их свойства сходимости в зависимости от свойств аппроксимируемой функции. Например, для функций, голоморфных в некотором эллипсе, содержащем отрезок, коэффициенты разложения убывают экспоненциально быстро.

Рассмотрим, например, простейшую краевую задачу (задачу Коши):

$$y'(x) = f(x), \quad x \in [0, 1], \quad y(0) = b, \quad (11)$$

где $f(x)$ — это полином степени $N-2$. Ясно, что решение $y(x) \in \mathcal{P}_N$ находится однозначно, причем независимо от того, какие полиномы использовать в виде базисных.

Таким образом, для *точного* решения задачи (11) спектральным методом размерности N необходимо и достаточно, чтобы правая часть $f(x) \in \mathcal{P}_{N-1}$. Здесь термин “точно” понимается в смысле “сколь угодно точно” или с машинной точностью. Эта оговорка необходима потому, что вычисления проводятся на реальном, а не на идеальном компьютере с неограниченной разрядной сеткой.

По формулам разд. 2 решение может быть найдено следующим образом.

Правую часть задачи (11) выражаем в спектральном виде по формуле (9) (для различных узлов) или с помощью конфлюентных разделенных разностей (см. разд. 4). Тогда задача (11) принимает вид

$$Dy = \langle b_0, \dots, b_{N-2}, 0 \rangle^t \in \mathcal{A}_N, \quad (12)$$

где $y = \langle a_0, \dots, a_{N-1} \rangle^t$ — это коэффициенты разложения (4).

Ясно, что матрица D в (12) вырождена и имеет ранг $N - 1$. При этом у нее есть нулевая последняя строка, которую можно использовать. Для этого заменим ее на вектор-строку

$$\langle p_0(0), \dots, p_{N-1}(0) \rangle,$$

т.е. на значения полиномов Ньютона в нуле. Затем заменим нуль у вектора в правой части (12) на краевое (в данном случае начальное) значение b . Тогда полученная конечномерная аппроксимация задачи невырождена и дает нужное решение.

Последнее утверждение очевидно и доказывается так же, как в предложении 2 в [2] для полиномов Лежандра.

В случае, когда $f(x) \notin \mathcal{P}_{N-1}$ в правой части (11), это означает, что в спектральном представлении $f(x)$ величины $b_m \neq 0$, $m \geq N - 1$ (вообще говоря). Но величиной b_{N-1} придется пожертвовать, т.е. заменить ее на величину начального значения b .

Для спектральных разложений по полиномам Чебышёва или Лежандра эта операция не несет в себе риска потери точности решения, так как величина b_{N-1} будет весьма мала для гладкой функции $f(x)$ и достаточно большого N . Поэтому полученные приближенные решения весьма быстро стремятся к точному решению при увеличении N . Кроме того, точность полученного решения можно оценить по его коэффициентам Фурье—(Чебышёва)—(Лежандра) (см. [3]).

Все сказанное выше о решении задачи (11) спектральным методом также относится к любой краевой задаче для любого голономного дифференциального оператора, причем независимо от набора базисных полиномов. Примеры можно найти в [2] и [3].

Однако для аппроксимаций полиномами Ньютона, как было отмечено, не существует априорных оценок точности. Более того, метод вообще не обязан сходиться к нужной функции на выбранном интервале даже при аппроксимации функции, аналитической в окрестности отрезка.

Речь идет, разумеется, о феномене Рунге, который обнаружил это явление при интерполяции функции $1/(1+x^2)$ полиномами на сетке с равноотстоящими узлами на интервале $[-5, 5]$. Подробный анализ этого явления с оценками скорости расходимости можно найти в [5, р. 35].

Масштабируя пример Рунге на интервал $[0, 1]$, получим функцию

$$y(x) = (h - x + x^2)^{-1}, \quad h = \frac{13}{50}.$$

Будем приближать $y(x)$ полиномами Ньютона с узлами

$$x_n = \frac{n-1}{N-1}, \quad n = 1, \dots, N.$$

Для этого достаточно вычислить вектор

$$s = (hE - X + X^2)^{-1} e,$$

а затем составить полином (4), где $a_{n-1} = s_n$, $n = 1, \dots, N$.

Можно проверить, что внутри отрезка, т.е. для $x \in [\varepsilon, 1 - \varepsilon]$, наблюдается экспоненциально быстрая сходимость $y_N(x) \rightarrow y(x)$ для не слишком малых ε . Однако при $N = 10$ максимальное отклонение полинома $y_N(x)$, совпадающего с $y(x)$ в данных узлах, равно ≈ 30 . При $N = 20$ максимальное отклонение будет уже ≈ 860 и т.д.

Процесс расходимости весьма напоминает явление Гиббса. На графиках (которые мы опускаем) видно, что расходимость концентрируется по краям интервала в виде двух острых пиков.

Анализ феномена Рунге, данный в [5], не называет причины, по которой он вообще появляется. Причина, разумеется, в особенностях этой рациональной функции, расположенных в комплексной области слишком близко к вещественной оси.

Проблема в том, что для общего расположения узлов, которые даже не обязаны принадлежать отрезку, оценка того, что означает “слишком близко к вещественной оси” может быть только экспериментальной.

Если в данном примере взять $h = 16/50$, то феномен Рунге по-прежнему присутствует, но только в виде явления Гиббса. То есть $y_N(x)$ сходится при $N \rightarrow \infty$ на всем интервале $[0, 1]$ к $y(x)$ равномерно, хотя и медленно.

Заметим также, что заключение о сходимости/расходимости в данном случае нельзя сделать на основании оценки величины последнего коэффициента. Для $N = 40$ ($h = 16/50$) последний ненулевой коэффициент в разложении (4), $a_{N-2} \approx -1.30 \times 10^{17}$, поэтому последний член $a_{N-2} p_{N-2}(x)$ не мал вблизи конца отрезка $x = 1$ ($a_{N-2} p_{N-2}(1) \approx -921.47$). При этом максимальное отклонение $\max |y_N(x) - y(x)|$ на отрезке $[0, 1]$ не превосходит 0.16.

Недостатки полиномов Ньютона как базисных функций в спектральном методе по сравнению с ортогональными полиномами этим не исчерпываются. Это видно, если сравнить числа обусловленности матриц преобразования Фурье F в этих методах. Для полиномов Ньютона число обусловленности матрицы F стремится к бесконечности при $N \rightarrow \infty$, а для полиномов Чебышёва оно ограничено (и мало, см. [3]). Поэтому выбор узлов Чебышёва для полиномов Ньютона, вообще говоря, не решает всех проблем, возникающих при решении задач аппроксимации этим методом (см. [11], где полагаются на чебышёвские узлы).

Однако полиномы Ньютона обладают также и рядом несомненных достоинств, как то простота использования и гибкость в выборе узлов, которые вообще могут быть произвольными.

Рассмотрим, например, сингулярную краевую задачу, решение которой дает константу Эйлера–Гомпертца:

$$x^2 y'(x) + y(x) = x, \quad x \in [0, 1], \quad y(0) = 0, \quad y(1) = \delta,$$

где $\delta = \exp(1) \operatorname{Ei}(1, 1)$ и $\operatorname{Ei}()$ — это интегральная экспонента.

Это классическая задача, впервые решенная Эйлером. Мы ее также неоднократно рассматривали как модельную задачу (см. [3]).

Единственное ограниченное в нуле решение этой задачи имеет начальное значение $y(0) = 0$. Поэтому начальное значение можно не учитывать. Тогда конечномерная аппроксимация этой задачи имеет вид

$$A y = (X^2 D + E) y = X e = x = \langle x_1, 1, 0, \dots, 0 \rangle^t. \quad (13)$$

Если взять узлы $\Lambda = \{0, \dots, 0\}$, то (см. [3]) вектор $A^{-1} x$ дает отрезок всюду расходящегося ряда

$$y_N(x) = \sum_{n=1}^{N-1} (-1)^{n-1} (n-1)! x^n,$$

который совершенно бесполезен для аппроксимации функции.

Возьмем теперь узлы $\Lambda = \{1, \dots, 1\}$ и краевое значение $y(1) = \delta$, которое пока что неизвестно (т.е. просто символ). Для учета краевого значения последнюю строку матрицы A надо заменить на строку

$$\langle p_0(1), \dots, p_{N-1}(1) \rangle,$$

т.е. на значения полиномов Ньютона в единице. Затем заменяем последний элемент вектора x в правой части (13) на символ δ . Получим матрицу \tilde{A} и вектор \tilde{x} . Тогда полученная конечномерная аппроксимация задачи невырожденна и дает решение $y = \tilde{A}^{-1} \tilde{x}$, т.е. коэффициенты разложения (4). Получаем

$$y_N(x) = \delta + (1 - \delta)(x - 1) + \left(\frac{3}{2}\delta - 1\right)(x - 1)^2 + \left(\frac{4}{3} - \frac{13}{6}\delta\right)(x - 1)^3 + \\ + \left(\frac{73}{24}\delta - \frac{11}{6}\right)(x - 1)^4 + \left(\frac{5}{2} - \frac{167}{40}\delta\right)(x - 1)^5 + \dots,$$

что совпадает с тейлоровским разложением точного решения в точке $x = 1$, $y(x) = \exp(1/x) \operatorname{Ei}(1, 1/x)$, до $(N-1)$ -го члена включительно.

Отметим, что общий вид (или формула) коэффициентов последнего разложения неизвестен. Так что мы выразили эти коэффициенты с помощью линейных операций с известными матрицами.

Если теперь вспомнить, что $y(0) = 0$, то из последнего разложения получаем уравнение для определения δ . Например, для $N = 10$ получаем

$$\delta \approx \frac{748420}{1255151}$$

с погрешностью менее 6.850×10^{-5} . Если распределить узлы равномерно по интервалу $[0, 1]$, то точность аппроксимации можно значительно увеличить.

Также полиномы Ньютона весьма удобны, когда значения функции в некоторых точках известны. Например, вычислим функцию $y(x)$, где

$$y(1/m) = \sum_{n=1}^m \frac{1}{n^2},$$

и найдем значение $y(0) = \pi^2/6$, т.е. найдем сумму этого ряда при $m \rightarrow \infty$.

Для этого вычислим вектор

$$s = F \langle y(1), y(1/2), \dots, y(1/N) \rangle^t,$$

затем составим полином (4), $y_N(x)$, где $a_{n-1} = s_n$, $n = 1, \dots, N$. Тогда $y_N(0) \approx \pi^2/6$. Для $N = 10$ получаем

$$y_N(0) = \frac{758128119023}{460886630400}, \quad \left| y_N(0) - \frac{\pi^2}{6} \right| < 6 \times 10^{-10}.$$

Нетрудно проверить, что этот способ вычисления суммы ряда совпадает с экстраполяцией Ричардсона (см. [12]).

Феномен Рунге также объясняет, почему метод Ричардсона столь чувствителен к выбору вспомогательной последовательности. В примере выше вспомогательная последовательность — это $x_n = 1/n$, т.е. узлы полинома Ньютона. Но при очевидном выборе узлов можно случайно попасть на пик явления Гиббса на краю интервала. В то же время можно, в принципе, удачно выбрать узлы и улучшить результат экстраполяции на несколько порядков. Примеры см. в [12].

4. КОНФЛЮЕНТНЫЕ РАЗДЕЛЕННЫЕ РАЗНОСТИ

Для попарно различных узлов, которые в остальном могут быть произвольными, спектральное представление разделенных разностей, как мы показали, обладает рядом существенных преимуществ по сравнению с их классическим представлением.

Но в случае, когда ставится стандартная задача интерполяции известной либо эмпирической функции общими полиномами Эрмита, спектральное представление не дает существенных преимуществ по сравнению с конфлюентными разделенными разностями, так как матрица преобразования Фурье F в этом случае не существует.

Правда, можно обойти эту проблему с помощью символьных вычислений в CAS. Для этого узлы берутся сперва различными, затем вычисляются все нужные матрицы и коэффициенты полинома Ньютона, а затем нужные группы узлов полагаются равными в интерполянте $y_N(x)$.

Но эта последняя операция эквивалентна вычислению конфлюентных разделенных разностей, только на более поздней стадии. Поэтому уместно привести процедуру вычисления общей интерполяции Эрмита, основанную на классических правилах.

То, как это изложено в учебниках (и в Википедии), предполагает построение нужных таблиц разделенных разностей вручную, что нас не устраивает. Поэтому приведем готовую процедуру в CAS Maple, а затем дадим необходимые пояснения.

Процедура для интерполяции известной функции $y(x)$ имеет вид:

```
dd:=proc() option remember: global xp,y: local n,k:
  n:=nargs:
  if nops({args})=1 and member(args[1],xp,'k') then
    return eval(diff(y(x),[seq(x,i=1..n-1)]),x=xp[k])/(n-1)!
  fi:
  normal((dd(args[1..n-1])-dd(args[2..n]))/(args[1]-args[n]))
end:
```

Эрмитов интерполянт вычисляется так:

$$y_N(x) = \sum_{n=1}^N dd(x_1, \dots, x_n) p_{n-1}(x).$$

Проще всего работу процедуры объяснить на примере из разд. 2, где мы интерполировали функцию $y(x) = 1/(1+x)$. Напомним, что мы ищем полином $y_N(x)$, совпадающий с $y(x)$ в точках $x_p = [0, 1/2, 1]$ и имеющий те же первые производные по краям интервала $[0, 1]$, а также две общие производные в середине интервала. Поэтому здесь $N = 7$.

Массив x_p в процедуре `dd` — это все *различные* точки, в которых задана функция и ее производные, а массив узлов — это, как и ранее, $\Lambda = [0, 0, 1/2, 1/2, 1/2, 1, 1]$, где каждая точка из x_p берется столько раз, сколько условий наложено на интерполянт в этой точке. Причем на этот раз, в отличие от примера из разд. 2, все одинаковые точки группируются строго вместе.

Это последнее требование необходимо потому, что процедура `dd` для различных аргументов дает обычную разделенную разность, определенную в разд. 1, но для n одинаковых аргументов (и только для них):

$$dd(x, \dots, x) = \frac{1}{(n-1)!} \frac{d^{n-1}y}{dx^{n-1}}(x), \quad n \in \mathbb{N},$$

согласно определению конфлюэнтных разделенных разностей.

Отметим, что процедура рекурсивна и запоминает ранее вычисленные разности, поэтому никаких лишних вычислений не производится. Можно проверить, что этим способом вычисляется то же самое, что мы получили в разд. 2, т.е. полином (7).

В случае, когда функция задана таблицей своих значений и производных в наборе точек, строку в процедуре dd, которая начинается оператором “return”, следует заменить на строку

`return y[n-1,k]/(n-1)!`

где y теперь обозначает двумерный массив значений функции и ее производных. Структура этого массива должна соответствовать тому, как функция и ее производные вычислялись ранее. Для приведенного примера в нулевой строке массива y стоят значения функции в точках x_r . В первой строке — значения первых производных и т.д.

В точках, где производные не заданы, таблицу y можно не заполнять, т.е. процедура не обращается к этим значениям.

Согласно определению эрмита интерполянта в таблице значений функции y столбцы не могут иметь незаполненные ячейки. То есть если, например, в точке x_m задана k -я производная функции $y(x)$, то должны быть также заданы все значения

$$y(x_m), \frac{dy}{dx}(x_m), \dots, \frac{d^k y}{dx^k}(x_m).$$

Однако процедура dd также может быть использована для вычисления (в принципе) любого полиномиального интерполянта. Для этого пустые места в таблице y (т.е. где функция либо ее производные не заданы) заполняются неопределенными параметрами, которые определяются из условий обнуления такого же количества последних коэффициентов в разложении (4).

Правда, интерполянт нужной степени (т.е. количество наложенных условий минус единица) может и не существовать. Например, можно проверить, что в приведенном примере нельзя выкинуть одно условие на первую производную в точке $x = 1/2$ и получить при этом полином 5-й степени.

В заключение этого раздела (и статьи) несколько слов о терминологии.

То, что мы называем (общей) эрмитовой интерполяцией является общепринятым термином. Однако в некоторых современных учебниках такие эрмитовы интерполянты называют *osculating polynomials* (см. [13]). Эрмитова интерполяция в [13] — это полиномы, совпадающие с функцией в выбранных узлах вместе со значениями первой производной (что, вероятно, исторически правильно). Однако в русскоязычной литературе оскулирующими бывают орбиты, а не полиномы, поэтому мы придерживаемся общепринятой терминологии.

СПИСОК ЛИТЕРАТУРЫ

1. *Davis P.J.* Interpolation and Approximation. New-York: Dover. 1975.
2. *Варин В.П.* Аппроксимация дифференциальных операторов с учетом граничных условий // Ж. вычисл. матем. и матем. физ. 2023. Т. 63. № 8. С. 1251–1271.
3. *Варин В.П.* Спектральные методы решения дифференциальных и функциональных уравнений // Ж. вычисл. матем. и матем. физ. 2024. Т. 64. № 5. С. 713–728.
4. *P. M. M.* “Interpolation.” By J. F. STEFFENSEN, SC.D., Professor of Actuarial Science at the University of Copenhagen // REVIEWS. P. 325–332. (1927). <https://www.cambridge.org/core>
5. *Steffensen J.F.* Interpolation. Baltimore: The Williams and Wilkins Co. 1927.
6. *Milne-Thomson L.M.* The Calculus of Finite Differences. London: Macmillan and Co. 1933.
7. *Wilf H.S.* Mathematics for the physical sciences. NewYork: Wiley. 1962.
8. *Opitz G.* Steigungsmatrizen // Z. Angew. Math. Mech. 1964. V. 44. T52–T54.
9. *Gantmacher F.R.* Application of the Theory of Matrices. New-York: Chelsea Press. 1960.
10. *Markoff A.A.* Differenzenrechnung. Leipzig: Teubner. 1896.
11. *Trefethen L.N.* Approximation Theory and Approximation Practice. SIAM. 2013.
12. *Варин В.П.* Преобразование последовательностей в доказательствах иррациональности некоторых фундаментальных констант // Ж. вычисл. матем. и матем. физ. 2022. Т. 62. № 10. С. 1587–1614.

13. Burden R.L., Faires J.D. Numerical analysis. 9th ed. Boston: Brooks/Cole. 2010.

SPECTRAL METHODS OF POLYNOMIAL INTERPOLATION AND APPROXIMATION

V. P. Varin^{a,*}

^a*Keldysh Institute of Applied Mathematics RAS, Miusskaya pl. 4, 125047 Moscow, Russia*

^{*}*e-mail: varin@keldysh.ru*

Received May 22, 2024

Revised October 18, 2024

Accepted November 8, 2024

Abstract. The classical problem of interpolation and approximation of functions by polynomials is considered here as a special case of spectral representation of functions. This approach was previously developed by us for the orthogonal Legendre and Chebyshev polynomials. Here, we use fundamental Newton polynomials as basis functions. It is shown that the spectral approach has computational advantages over the divided difference method. In a number of problems, Newton and Hermite interpolations are indistinguishable with our approach and are calculated using the same formulas. Also, the computational algorithms that we proposed earlier using orthogonal polynomials are transferred without changes to Newton and Hermite polynomials.

Keywords: spectral methods, Newton and Hermite polynomials, interpolation and approximation