Ivo Babuška; Sergei Lvovich Sobolew

Оптимизация численных методов

Aplikace matematiky, Vol. 10 (1965), No. 2, 96–129

Persistent URL: http://dml.cz/dmlcz/102941

ОПТИМИЗАЦИЯ ЧИСЛЕННЫХ МЕТОДОВ

И. БАБУШКА, С. Л. СОБОЛЕВ (I. BABUŠKA, S. L. SOBOLEV)

0. ВВЕДЕНИЕ

Вычислительная математика и в частности численные методы играют сейчас все большую роль.

Огромные возможности быстродействующих машин привлекли к себе внимание математиков, физиков, инженеров и др.

При этом методы вычислений, разработанные в прошлом оказались не всегда достаточным и часто не соответствуют новым техническим средствам. По этой причине произошла перестройка этой части математики. Если в прошлом численные методы представляли собой множество более или менее искусственных приемов, то сейчас они приводятся в систему и объединяются в стройное целое. Этим новым идеям наряду с их конкретным осуществлением мы и посвятили наш доклад. Важной практической целью вычислительной математики является отыскание наилучших, т. е. кратчайших, наиболее быстрых и дешевых способов решения математических задач, т. е. отпимизация вычислительных алгорифмов. Это ворпос не простой.

Решение нужной задачи обычно происходит с использованием многичисленных данных не всегда указанных прямо в условиях задачи. Эти данные могут быть самой различной природы. Например:

а) Мы можем воспользоваться при решении таблицами специальных функций или известными нам константами.

б) В процессе решения можно использовать некоторые готовые стандартные приемы или стандартные схемы, как например, разные схемы решения задач линейной алгебры или какую-либо схему решения системы обыкновенных дифференциальных уравнений.

в) Разными могут быть технические средства, применяемые в процессе решения. Мы можем воспользоваться машинами разных марок с различным быстродействием.

г) Четвертый тип дополнительных данных — данные, обычно указаываемые в задаче, например, ограничения на дифференцируемость или рост производных неизвестной функции.

В зависимости от всех этих данных будет находиться оптимальный, т. е. наиболее подходящий алгорифм решения данной задачи.

Задача об оптимизации в общем виде состоит в том, чтобы найти условный экстремум для вычислительной работы среди всех мыслимых алгорифмов, учитывая данные, указанные в а), б), в) и г).

Использование всех средств, указанных в а), б), в) и г) есть использование различной информации, как прямо так и косвенно относящейся к данной задаче. Некоторый минимум такой используемой информации всегда существует, хотя бы в виде таблицы умножения или сложения первого десятка чисел при ручном счете. Максимально возможная информация это наличие переписанного готового ответа на том месте, где он нам нужен, без необходимости поиска в таблицах.

При имеющейся дополнительной информации в виде а), б), в) или г) суживается класс возможных ответов, т. е. тем самым суживается класс рассматриваемых задач.

Очень важно следующее: во всяком решении, достаточно сложной задачи, алгорифм не бывает сразу до конца определен и не только число действий неизвестно, но и сами эти действия вырабатываются в процессе решения задачи. Благодаря этому, например, удается решить при помощи программы заложенной в машину, состоящей из нескольких тысяч операций, такую задачу, для которой общее число операций, выполняемых машиной составляет десятки и сотни миллионов. Машина сама детализирует вычисления, следуя правилам заложенным в программе.

Поиск путей решения в процессе самого решения, выполняет не только машина. Занимаясь задачей, мы должны часто принимать дополнительные решения, осуществлять добавочный выбор.

Например, используя разные формулы для механических кубатур (квадратур функций многих переменных), мы должны при этом иногда находить и оптимальную систему узлов и коэффициенты такой формулы. Подсчет коэффициентов является таким образом частью работы, выполняемой при решении задачи.

Конечно, такое положение дел в каждом-то смысле подчас делает задачу оптимизации неопределенной.

Отметим, еще одно обстоятельство; если один и тот же подсчет, например составление таблиц, сделан для многих задач, то имеет смысл затратить на него сравнительно большее время и силы. Этот труд окупится при пользовании этими таблицами. В обратном случае это уже не так. Очевидно, далее что заранее предугадать весь возможный объем использования тех или иных таблиц, стандартных программ и т. п., в большом числе случаев просто нельзя.

Существенную и притом никак не выражаемую явно роль играет и то, что создаваемое вспомогательное средство заготовки и т. д. используются не только один раз в настоящем, но как правило служат долгое время (таблицы логарифмов до создания арифмометров). Поэтому правильно распределить долю затрат, приходящихся на каждый случай их использования, нельзя из-за отсутствия данных. Это использование носит динамический характер.

Каждое множество результатов вычислений, полученных определенным методом при условных весах использованных средств имеет в свою очередь определенный вес измеренный в затратах вычислительного труда. Эти затраты зависят от избранного метода. Достижение оптимизации вычислений неотделимо от систематизации наших знаний об этих затратах для всех мыслимых вычислительных задач.

Однако, такая полная систематизация, если только она вообще возможна — дело будущего. Сейчас делаются в этом направлении лишь некоторые первые шаги, в выяснении некоторых важных свойств вычислительных задач.

Как видно отсюда, изучая проблемы оптимизации решения при данных условиях, т. е. оптимальные методы и способы их нахождения, мы должны одновременно с помощью общих математических теорем изучать и классифицировать задачу, оценивать a priori каковы могут быть в принципе минимальные затраты труда на их решение.

Как мы видели, рассмотрение одной изолированной математической задачи оптимизации большей частью не решает никакого практического вопроса. Однако, умея находить условный экстремум, т. е. наилучший способ решения при заданных возможностях а), б), в) и г) каждой локальной задачи, мы тем самым даем ключ к решению общей проблемы.

Положение здесь в точности таково, как при выводе например, уравнений Эйлера в вариационной задаче. Общий экстремум должен обязательно являться одновременно и локальным экстремумом на любом участке при соответствующих граничных условиях согласования с соседними участниками. Локальные экстремальные задачи уже доступны для решения.

Поскольку в проблемах оптимизации речь идет о наилучшем использовании ресурсов для решения сразу многих задач — эти проблемы имеют также много общего с задачей математической экономики и линейного программирования.

Наш доклад будет состоять из четырех разделов.

Изложим вкратце их содержание.

1. Оптимизация конкретных задач

В разных задачах поиски наилучших решений производились уже давно, причем разумеется получить окончательные результаты даже в условной постановке удается далеко не всегда. Этот раздел нашего доклада будет повсящен тем задачам, где это удается довести до конца. Мы приведем, разумеется, не все, но лишь некоторую часть результатов на наш взгляд наиболее интересных.

2. Асимптотически оптимальные результаты для вычислений линейных функционалов

В некоторых задачах, несмотря на отсутствие практических путей для нахождения оптимальных способов расчета, удается построить алгофирфмы близкие к оптимальным, оценив уклонения их от оптимальных. Таково, например, положение в некоторых постановках задачи о численном интегрировании, т. е. в теории кубатурных формул.

Уместно заметить еще раз, что с практической точки зрения при учете оптиальности метода, труд поиска экстремальных алгорифмов нужно включить в общие затраты труда на решение задачи. При этом в ряде случаев, оптимальный в каких-либо заданных условиях алгорифм, вообще может оказаться практически и теоретически мало пригодным из-за трудности нахождения самого этого алгорифма.

3. Оптимизация линейных задач

Для класса задач, связанных с обращением дифференциальных операторов, как правило имеющих вполне непрерывный обратный, можно бывает иногда дать довольно близкую оценку числа необходимых действий (из заданного набора первичных, которые мы считаем осуществленными, снабженных определенным весом). Эта оценка, основанная на общих понятиях теории множеств, связанная с т. н. поперечниками множеств введенными А. Н. Колмогоровым позволяют найти алгорифмы близкие к оптимальным по числу действий.

4. Общие идеи и оценки

В общем случае, оценку числа необходимых действий для решения той или иной задачи можно бывает получит асимптотически исходя из другого важного понятия теории множеств также введенного Колмогоровым именно из понятия ε-энтропии компактных множеств.

Современному состоянию этих идей мы отведем последний заключительный раздел нашего доклада.

Прежде чем перейти к самому исложению мы должны будем еще несколько уточнить предмет, которым будем заниматься.

В современном представлении вычислительная математика состоит как бы из двух основных частей. Ее классическая часть — это теория приближенных операций в бесконечных множествах, конечно-мерных и бесконечно-мерных функциональных пространствах. Другая часть это — математика конечного, математика дискретного.

Дискретная математика входит в нее двумя сторонами. Прежде всего — это новый круг задач, новые вопросы поставленные в последнее время жизнью.

С другой стороны сами математические методы теории вычисления близко соприкасаются с математической логикой, теорией оптимизации дискретных вычислений.

В нашем докладе мы не будем затрагивать проблем дискретной математики так как это было бы, слишко обширной темой и не уложилось бы в поставленные нам рамки.

Переходим к основному предмету.

1. ОПТИМИЗАЦИЯ КОНКРЕТНЫХ ЗАДАЧ

Задача о сравнении алгорифмов, выходящих из равной общей информации и дающих равные результаты не так проста. Сравнение всегда относительное в зависимости, от того, какие средства (технические и общие) мы имеем в своем распоряжении от того, какие требования накладываются на процесс вычисления результатов.

Сравнивая разные алгорифмы с точки зрения их удобства при ручном счете, счете на арифмометре или счете на машине, при условии паличия таблиц или при отсуствии их, при различной требуемой точности мы как правило придем к разным результатам. Ясно, что имеет смысл сравнивать только такие алгорифмы, которые исходя их равных начальных данных дадут равные результаты.

Возникает вопрос, что такое алгорифм. И когда два алгорифма сравнимы. Дадим определение.

Определение 1. *Пусть дана последовательность метрических пространств*

$$X_i\,, \quad i = -p, \dots, 0, 1, \dots, N\,, \quad p \geqq 0\,, \quad N \leqq \infty$$

и последовательность операторов A_i, $i = 0, 1, 2, \dots, N-1$, *которые отображают картезианское прямое произведение*

$$X_{-p} \times X_{-p+1} \times \dots \times X_i \quad \text{в} \quad X_{i+1}\,.$$

Пусть далее мы имеем последовательность чисел

$$0 < n_k \leqq N\,, \quad k = 1, 2, \dots, q\,.$$

Последовательность уравнений

$$x_{i+1} = A_i(x_{-p}, \dots, x_i)\,, \quad x_i \in X_i\,, \qquad i = 0, 1, 2, \dots, N-1 \tag{1}$$

мы вудем называть численным алгорифмом. Если заданы

$$x_{-p}, \dots, x_0\,, \quad x_i \in X_i\,,$$

то последовательность элементов, определенных уравнениями (1), *мы называем решением. Элементы* $x_{n_k} \in X_{n_k}$, $k = 1, \ldots, q$ *называем результатами и элементы* $x_{-p}, \ldots, x_0$ *начальными данными.*

Далее мы увидим, что весьма существенно, будут ли пространства X_i, $i = -p, \ldots, 0$ пространства только с одним элементом (или конечным числом элементов) или бесконечным числом элементов.

В случае, если эти пространства (или только некоторые) одноточечные мы говорим о стандартных алгорифмах, в случае если эти пространства имеют бесконечное множество элементов мы говорим об общих алгорифмах. Определение 1 включает реальные алгорифмы, которые употребляются на машинах при научно-исследовательских и технических расчетах но не при вопросах обработки данных.

Пространства X_i, $i = 1, 2, \ldots, N - 1$ – это в действительности (т. е. при работе машины) пространства действительных (или комплексных) чисел; X_i, $i = -p, \ldots, 0$ некоторые множества действительных (комплексных) чисел и операторы A_i представляют элементарные операции $\pm$, $\times$, : и операцию $\leqq 0$.

Это значит, что A_i бинарная зависящая может быть от sign остальных переменных.

Ясно, что ограничения структуры операторов A_i условное. В действительной машине имеется ряд (100 и больше) инструкций, т. е. в некотором смысле элементарных операций. Если X_i, $i = 1, 2, \ldots, N - 1$ это пространство действительных или комплексных чисел, мы будем говорить о непрерывной модели над телом действительных или комплексных чисел.

Важное значение имеет еще дискретная δ-модель, когда

$$(2) \qquad X_i = E[x \in R\,, \quad x = k\delta\,, \quad k - \text{целое}]\,.$$

Операции $\pm$, $\times$, : это будут операции $\pm$, $\times$ и : с округлением.

Ясно, что это тоже идеализация, которая должна отражать вычисление в фиксированной запятой. Аналогичную модель можно построить и для вычислений с плавающей запятой. Заметим, что наша цель получить искомое решение с точностью до ε. С точки зрения дискретной модели необходимо брать различные δ и с изменением δ изменяется фактическая длина алгорифма.

Определение 2. *Два численные алгорифма определенные через пространства*

$$(3) \qquad X^{(j)}_{-p}, \ldots, X^{(j)}_{N_j}\,, \quad j = 1, 2$$

и операторы

$$(4) \qquad A^{(j)}_i\,, \quad i = 0, 1, 2, \ldots, N_j - 1; \quad j = 1, 2$$

мы будем называть ε-*сравнимыми* ($\varepsilon \geqq 0$) *относительно* $\{m^{(j)}_k\}$, $k = 1, \ldots, q$, *если*

(5) $$X^{(1)}_{-l} = X^{(2)}_{-l}, \quad l = 0, 1, 2, \ldots, p, \quad X^{(1)}_{m_k^{(1)}} = X^{(2)}_{m_k^{(2)}}, \quad k = 1, \ldots, q,$$

а если из

(6) $$x^{(1)}_{-l} = x^{(2)}_{-l}, \quad x^{(1)}_{-l} \in X^{(1)}_{-l}, \quad x^{(2)}_{-l} \in X^{(2)}_{-l}, \quad l = 0, \ldots, p,$$

следует

(7) $$\varrho\left(x^{(1)}_{m_k^{(1)}},\ x^{(2)}_{m_k^{(2)}}\right) \leqq \varepsilon, \quad k = 1, 2, \ldots, q,$$

где ϱ *метрика в данных пространствах.*

Если $\varepsilon = 0$, *то алгорифмы чисто сравнимы.*

Подбор метода — это оптимальный подбор в классе ε-сравнимых алгорифмов. Важен критерий оптимального подбора. В качестве такого критерия можно взять например:

а) длину алгорифма,

б) объем памяти (количество необходимых ячеек).

Возможны еще сложные критерии, которые в конечном итоге связаны с этими двумя.

В качестве примера можно указать устойчивость счета. Для слабоустойчивых алгорифмов при желании получить высокую точность, длина алгорифма вместе с числом необходимых знаков в каждом действии очень сильно увеличивается. Поскольку очень часто алгорифм используется для решения сразу нескольких задач или целого класса, или для получения результатов с различной точностью, этот критерий может оказаться стол же важным как и ряд других. Еще один пример это возможность стандартизации и т. п.

Полезно заметить, что длина алгорифма зависит от начальных данных или другими словами, от того, какую память мы используем. Если у нас достаточная память, то может быть лучше выбирать например значение $\sin x$ из таблицы, чем делать вычисление по апроксимации Чебышева.

Длину алгорифма проще всего характеризовать числом элементарных операций $\pm$, $\times$, $:$. Эти операции (например $+$ и $\times$) вообще говоря неравноценны, так как время на большинстве машин для $+$ и $\times$ разное иногда они могут оказаться равноценными, так как на некоторых машинах они требуют равное время. Поэтому операции, входящие в его состав, нужно снабдить весом. Вернемся к примеру.

Предположим, что a_i, $i = 0, \ldots, n$ и x-любые действительные числа. Речь идет в этом случае об общем алгорифме для вычисления значения полинома $g(x) = \sum_{i=0}^{n} a_i x^i$.

Предположим, что мы имеем непрерывную модель над телом действительных (или соответственно комплексных) чисел и поставим ворпос: какое минимальное число операций $\times$, : и $\pm$ необходимо для вычисления значения в точке x некоторого многочлена $g(x) = \sum_{i=0}^{n} a_i x^i$ заданного своими коэффициентами.

Вопрос этого рода предложил в 1954 Ostrowski [1954] и доказал, что для многочленов не выше 4-той степени необходимо $2n$ операций $\pm$, $\times$, :. Пан [1962] доказал для общего многочлена n-той степени, что необходимо n операций $\times$, : и n операций $\pm$. Поскольку схема Горнера имеет n операций $\pm$ и n операций $\times$, то из теоремы Пана выходит, что в смысле длины, лучшего алгорифма чем алгорифм Горнера не существует.

Посмотрим, как изменится эта оценка, если мы выходим из данных a_i и x_j, $j = 1, 2, \ldots, s$ и требуемые результаты будут $g(x_j)$, $j = 1, 2, \ldots, s$.

Для больших s вопрос возможно ставить в следующей форме. Как можно наилучшим образом характеризовать однозначно многочлен n-той степени, чтобы для его вычисления было необходимо минимальное число операций. Motzkin [1955] показал, что для $n = 6$ можно многочлен характеризовать такими параметрами b_i, $i = 0, \ldots, 6$ при задании которых, для вычисления его значения при каждом x хватило бы менее чем 6 операций $\times$, :.

Многочлен $g(x)$ задается формулами $p_1(x) = x(x + b_1)$, $p_2(x) = (p_1 + x + b_2)(p_1 + b_3)$, $p_3(x) = (p_2 + b_4)(p_1 + b_5)$, $g(x) = b_0(p_3 + b_6)$.

Параметры $b_0, \ldots, b_6$ должны быть найдены (например из коэффициентов a_i) один раз для всех значений x.

Как видно здесь только 4 операции $\times$, : и 7 операций $\pm$. Белага [1958], [1961] показал в общем случае, что минимальное число операций $\times$, : соотв. $\pm$ при соответствующем задании любого многочлена составляет $[(n + 1)/2]$ и Пан [1959] уточнил эту оценку так, что необходимо $[(n + 3)/2]$ операций $\times$, : и n операций $\pm$.

Белага [1958] дал тоже конструкцию (над телом комплексных чисел), которая требует $[(n + 3)/2]$ операций $\times$, : и $(n + 1)$ операцию $\pm$.

Пан [1959], [1961] получил такой же результат над телом действительных чисел. Eve (1964) дал конструкцию которая требует $[(n + 4)/2]$ операций $\times$, : и n операций $\pm$.

Вычислительные алгорифмы этого вида называются методами с предварительной обработкой данных.

Мы занимались вопросом общих алгорифмов. Ясно, что в стандартных для конкретных случаев может хватить меньшего числа операций. Например, если

(8) $$g(x) = \sum_{n=0}^{31} x^n,$$

то

$$g(x) = \frac{x^{32} - 1}{x - 1}$$

и для вычисления достаточно произвести 6 операций $\times$, : и 2 операции $\pm$.

Вопрос о необходимом числе операций для конкретного многочлена $g(x) = x^p$ решил Р. Э. Вальский [1959] и в некоторой другой форме Brauer [1939]. Он показал что при правильном выборе алгорифма хватит $(1 + o(1)) \cdot [\lg_2 p]$ действий. Оценка снизу очевидно $[\lg_2 p]$.

Вопросы методов с предварительной обработкой данных имеют важное значение, на практике при конструкции стандартных программ например для специальных функций (т. е. апроксимации).

В связи с тем, что мы сейчас сказали наметим еще один вопрос. Задача по оптимальной апроксимации функций многочленом n-той степени выражает — в некотором смысле — требование наилучщей апроксимации при данной вычислительной работе, характеризуемой числом n, т. е. степенью многочлена. Вопрос можно ставить по другому: найти наилучший апроксимирующий многочлен, который требует n действий (например умножений или умножений и сложений).

Вопросы оптимальных алгорифмов — это почти нерешенные вопросы. Вычисление многочленов изучено в последнее время, как начало этого важного направления.

Важной задачей во многих вопросах теории вычислений является задача о решении системы n линейных алгебраических уравнений. В случае, когда матрица системы содержит только три диагонали с элементами неравными нулю, можно показать, что для решения необходимо в случае общего алгорифма для любых коэффициентов $(4n - 3)$ операции.

Метод исключения имеет $7(n - 1)$ операции. Видно, что метод исключения по порядку оптимальный. Когда мы имеем общую (т. е. полную) матрицу то, как можно просто видеть, необходимо $0(n^2)$ операций. Все до сих пор известные методы имеют $0(n^3)$ операций. Неизвестно, существует ли алгорифм решения системы n линейных уравнений, которая имеет $o(n^3)$ операций.

Очень интересные результаты получены А. Карацубой и А. Тоомом.[1] Ими рассматривалась задача об оценке минимального количества логических действий при умножении n-значных двоичных чисел.

Оказалось, что для перемножения таких чисел требуется при правильном выборе алгорифма не более $0(n^{1+o(1)})$ действий.

Таким образом известные с древности способы умножения чисел столбиком не являются наилучшими при умножении чисел большой значности.

Мы изучали некоторые вопросы оптимизации алгорифмов в некоторых случаях в предположении, что мы имеем в виду непрерывную модель (над телом действительных или комплексных чисел). В действительности машина не работает как непрерывная модель и вопросы становятся более сложными. Это вопросы численной устойчивости.

Покажем простой пример.

---

[1]) До сих пор неопубликовано.

Вычислим методом интегрирования по частям интеграл

(9) $$I_n = \frac{1}{e}\int_0^1 x^n e^x \,\mathrm{d}x\,.$$

Получим рекуррентную формулу

(10) $$I_k = 1 - kI_{k-1}\,, \quad I_0 = 1 - \frac{1}{e}\,.$$

Эта формула представляет собою уравнение в конечных разностях, в котором известны как I_0 так и I_∞, очевидно, равное нулю и его можно решать с двух сторон.

Мы получим так два ε-сравнимые алгорифма для вычисления I_n.

а) I. алгорифм:

(11) $$I_0 = 1 - \frac{1}{e}\,, \quad I_k = 1 - kI_{k-1}\,, \quad k = 1, 2, \ldots, n\,.$$

Легко видеть, что влияние малой ошибки в I_0 на I_n очень велико при больших n, причем ошибки для растущего n накапливаются и растут.

б) II. алгорифм:

(12) $$I_{[(\alpha+1)n]} = 0\,, \quad \alpha \geqq \left|\frac{\lg \varepsilon}{n \lg n}\right|,$$

$$I_k = \frac{1}{k+1}\left(1 - I_{k+1}\right), \quad k = [(\alpha + 1)\, n] - 1, \ldots, n\,.$$

Этот алгорифм будет, как легко видеть, очень устойчивым. Влияние ошибок округления будет маленькое.

Эти два алгорифма ε-сравнимые над телом действительных чисел.

Приведем результат действительных вычислений на машине:

(13) На машине ZUSE 23 получилось ($n = 13$, $\alpha = 1, 2$, $\varepsilon = 10^{-15}$)

	I. алгорифм	II. алгорифм
I_{13}	$-0{,}95098662$	$+0{,}0669477025$

Таким образом невозможно в этом примере практически применить алгорифм I. непрерывной модели, если хотим иметь реальный алгоригм; применение алгорифма II. дает очень хороший результат.

Из этого примера видно, что если бы мы взяли дискретную модель и хотели бы получить I_{13}, то для того чтобы эти два алгорифма были ε-сравнимы, было бы необходимо брать очень маленькое δ, или если мы хотим получить I_{13} с точностью до ε, то нам необходимо брать для первого алгорифма существенно меньшие δ, чем для алгорифма II.

Оттуда следует что с точки зрения непрерывной модели алгорифм I. лучше, чем алгорифм II., а с точки зрения дискретной модели, наоборот.

2. АСИМПТОТИЧЕСКИЕ ОПТИМАЛЬНЫЕ РЕЗУЛЬТАТЫ ДЛЯ ВЫЧИСЛЕНИЙ ЛИНЕЙНЫХ ФУНКЦИОНАЛОВ

Точный оптимум вычислений, как мы говорили, удается получить очень редко. Для более широкого класса задач удается построить асимптотически оптимальные формулы. Очень характерным является здесь пример теории формул приближенного интегрирования: квадратных и кубатурных формул.

Квадратурной формулой называется приближенная формула, выражающая значение интеграла

$$\int_\Omega f(x)\,\mathrm{d}x$$

в виде линейной комбинации значений подинтегральной функции в точках называемых узлами

$$\int_\Omega f\,\mathrm{d}x \cong \sum_{k=1}^{N} C_k\, f(x^{(k)})\,.$$

Не следует смешивать теорию кубатурных формул с более общей теорией приближенного вычисления интеграла. Легко указать пример, когда вычисление это, при помощи кубатурных формул, практически плохо пригодно. Например, если фунция $f(x)$ одной переменной дана в виде сходящегося степенного ряда

$$f(x) = \sum_{k=0}^{\infty} a_k x^k\,,$$

то может оказаться самым выгодным почленное интегрирование этого ряда.

Интеграл – это специальный линейный функционал, определяемый для непрерывных функций через пределы некоторых сумм (мы будем рассматривать здесь почти исключительно непрерывные функции).

С точки зрения анализа, исследование квадратных или кубатурных формул, есть исследование характера этого предельного перехода.

Однако, мы можем поставить в связи с этим более общий вопрос. Пусть мы имеем некоторое линейное нормированное пространство H. Предположим далее, что мы имеем конус $\mathfrak{M}$ в пространстве функционалов H'. Конусом мы называем такое подмножество $\mathfrak{M}$ пространства H', что если $\varphi \in \mathfrak{M}$ то и $a\varphi \in \mathfrak{M}$ для любого числа a.

Пусть далее мы имеем $F \in H'$, $F \notin \mathfrak{M}$, пусть $f_1, f_2, \ldots, f_N \in \mathfrak{M}$ и определим

$$(14) \qquad \inf_{c_j} \left\| F - \sum_{n=1}^{N} C_j f_j \right\| = \varrho(F, f_1, \ldots, f_N),$$

$$(15) \qquad a_N^{\mathfrak{M},F} = \inf_{f_j \in \mathfrak{M}} \varrho(F, f_1, \ldots, f_N) = \inf_{f_j \in \mathfrak{M}} \left\| F - \sum_{j=1}^{N} f_j \right\|.$$

Ясно что важное значение имеет вопрос имеетли место условие $a_N^{\mathfrak{M},F} \to 0$ для $N \to \infty$. Если существуют $\hat{f}_1, \hat{f}_2, \ldots, \hat{f}_N \in \mathfrak{M}$ так, что

$$\left\| F - \sum_{j=1}^{N} \hat{f}_j \right\| = a_N^{\mathfrak{M},F},$$

то мы будем говорить об апроксимации оптимальной в $\mathfrak{M}$ или оптимальных функционалах $\hat{f}_i$. Функционалы g_i, $i = 1, \ldots, N$ такие, что $\left\| F - \sum_{j=1}^{N} g_j \right\| = \varrho(F, g_1, \ldots, g_N)$ мы назовем локально оптимальными. В случае квадратурных формул задача об отыскании локально оптимальных функционалов это задача об отыскании коэффициентов при заданных узлах этой формулы.

В чем суть дела с практической точки зрения? Дело в том, что F – невычислимый функционал и $\mathfrak{M}$ – множество функционалов вычислимых. В качестве F можно например взять интеграл, а за $\mathfrak{M}$ множество функционалов типа

$$f(\varphi) = \varphi(x_i).$$

Ясно, что „работу“ на вычисление функционалов из $\mathfrak{M}$ мы должны определить вначеле. Мы предполагали, что все алгорифмы вычисления всех функционалов из $\mathfrak{M}$ равноценны. Но очевидно, можно обобщить задачу. Для каждого $f \in \mathfrak{M}$ определим некоторый вес как функцию x, определенную на единичном шаре (и вес f это значение $x(f/\|f\|)$). Этим мы не будем заниматься.

Очевидно $a_{n+1}^{\mathfrak{M},F} \leqq a_n^{\mathfrak{M},F}$.

Последовательность чисел $a_n^{\mathfrak{M},F}$, $n = 1, 2, \ldots$ – это характеристика апроксимирующих свойств множества $\mathfrak{M}$ по отношению к F.

Пусть $f_i^{(j)} \in \mathfrak{M}$, $i = 1, \ldots, j$, $j = 1, 2, \ldots$ некоторая последовательность апроксимирующих функционалов.

Если

$$\limsup_{j \to \infty} \frac{\left\| F - \sum_{i=1}^{j} f_i^{(j)} \right\|}{a_j^{\mathfrak{M},F}} < \infty,$$

то будем говорить, что $f_i^{(j)}$, $i = 1, 2, \ldots, j$ по порядку оптимальные апроксимирующие функционалы.

Если

$$\lim_{j \to \infty} \frac{\left\| F - \sum_{i=1}^{j} f_i^{(j)} \right\|}{a_j^{\mathfrak{M},F}} = 1 \,,$$

то мы говорим о асимптотически оптимальных апроксимирующих функционалах. Подобно этому будем говорить о асимптотически локально оптимальных по порядку или локально оптимальных асимптотически функционалах.

С точки зрения вычислительной математики важны

1. оценки $a_N^{\mathfrak{M},F}$;
2. конструкция оптимальных или оптимальных по порядку или оптимальных асимптотически апроксимирующих функционалов.

Различные постановки этой задачи для функций одной переменной рассматривали Никольский [1950], Sard [1949], Meyer, Sard [1950], Шайдаева [1959] и также Davis [1959]. Некоторыми общими вопросами подобного рода занимались Golomb, Weinberger [1959]. Для классов $W_p^{(l)}(\Omega)$ при некоторых p и l ими были найдены в точности оптимальные формулы, то-есть наилучшая апроксимация функционала в пространствах $W_p^{(l)}$. Для других p и $l \geqq 1$ найдены формулы, наилучшие по порядку.

Н. М. Коробов [1957], [1959] предложил способ интегрирования функций многих переменных с применением некоторых результатов теории чисел и оценил этот способ для классов $E_s^{(\alpha)}$ функций s переменных, у которых ограничены все смешанные производные, где порядок дифференцирования по каждой переменной в отдельности превосходит α.

Н. С. Бахвалов [1959], [1960] дал оценки числа необходимых действий в этих классах снизу, которые по порядку совпали с оценками сверху данными Н. М. Коробовым и дал целую серию оценок по вероятности, т. е. оценок числа действий необходимых для получения результатов с заданной вероятностью. На других классах ряд оценок сверху получили Бахвалов [1964], Соболь [1960], Смоляк [1960], Солодов [1959].

При существовании лишь обобщенных производных для пространств $L_2^{(m)}$ оценки погрешности получаются при $m > n/2$, т. е. при таких m и n, которые в силу теории вложения обеспечивали линейность оператора вычисления интеграла.

Укажем более конкретно относящиеся сюда результаты, как по оценкам, так и по оптимизации формул.

Прежде всего, укажем некоторые самые общие оценки порядка необходимого числа действий, т. е. необходимого числа N точек $x^{(j)}$ и характер их расположения.

Поместим в данной области $\Omega(E_n)$ некоторое количество кубов Q_j с длиной ребер h. Внутри каждого такого куба содержится некоторое количество узлов кубатурной формулы.

Пусть сумма коэффициентов кубатурной формулы $\sum_{x^{(k)} \in Q_j} C_k = \Delta_j h^n$. Будем рассматривать множество S непересекающихся кубов Q_j, для которых это отношение сторого отличается от единицы: например,

$$\Delta_j < q < 1 .$$

Верна следующая теорема:

Теорема 1. *Справедлива следующая оценка снизу для ошибки кубатурной формулы*

$$\|l\|_{L_2^{(m)}(\Omega)} \geqq k \left(\frac{\sum_s |Q_j|}{|\Omega|} \right)^{\frac{1}{2}} (1 - q)\, h^m , \tag{16}$$

где $|Q_j|$ *объем* Q_j.

Отсюда следует сразу например, что ни при каких условиях мы не можем получить погрешность кубатурной формулы меньше, чем $kN^{-m/n}$. Действительно: покрыв область $\Omega \subset E_n$ сеткой кубов со стороной $(\frac{1}{2}N)^{-1/n}$, мы получим, что в половине таких кубов не будет содержаться вообще ни одной точки решетки. При этом оценка (16) дает нам желаемый результат.

Подобного рода оценка получена Н. С. Бахваловым для несколько более широкого класса функций.

Теорема 1′. [Бахвалов [1963]] *Пусть* $C_{r_1, r_2, \ldots, r_s}(A_1, \ldots, A_s)$ *класс функций определенных в* s*-мерном кубе и удовлетворящих условиям*

$$\left| \frac{\partial^r f}{\partial x_k^r} \right| \leqq A_k \quad \text{при} \quad 1 \leqq r \leqq r_k . \tag{17}$$

Тогда для любой кубатурной формулы с N *узлами найдется функция из рассмотриваемого класса, для которой ошибка интегрирования не меньше*

$$\min_k (r_k)\, C(r_1, \ldots, r_s) \min_i (A_i) \,.\, N^{-r} ,$$

$$r^{-1} = r_1^{-1} + \ldots + r_s^{-1} .$$

Приведенные оценки очевидно сильно зависят от класса рассматримых функций. Если пространство будет более, специальное, напр. если мы добавим условие что все функции гармонические, то мы получим лучшие оценки. Теоремы этого характера служит средством для оценок по порядку числа действий нужных для получения необходимой точности в ряде задач вычислительной математики к этому вопросу мы еще вернемся в разделе 4.

Из оценки, приведенной нами, следует также, что точки в кубатурной формуле должны быть распределены по возможности равномерно. Теоремы 1 и 1′ доказываются непосредственно при помощи конструкции функции из $L_2^{(m)}$ соотв.

$C_{r_1,\ldots,r_s}(A_1, \ldots, A_s)$, на которой функционал ошибки (l, φ) достигает самого большого значения.

Следующий ряд результатов, которые мы приведем также без доказательства, основан на теории уравнений в частных производных и вариационном исчислении.

Основой его служит уравнение Эйлера для экстремального элемента из $L_2^{(m)}(E_n)$, т. е. того элемента, на котором функционал ошибки (l, φ) достигает своего максимума. Это уравнение в терминологии обобщенных функций будет записываться в виде

$$\Delta^m u = (-1)^{m-1}\, l(x)\,, \tag{17}$$

где

$$l(x) = \mathscr{E}_\Omega(x) - \sum_k C_k\, \delta(x - x^{(k)})\,. \tag{18}$$

Через $\mathscr{E}_\Omega(x)$ обозначена характеристическая функция области интегрирования, а $\delta(x)$ так называемая функция Дирака.

При нашем выборе пространства $L_2^m(E_n)$ никаких граничных условий ненужно. Заметим еще, что выбор нормы в пространстве функций конечно не однозначен. Мы ограничиваемся нормой $L_2^{(m)}(E_n)$ потому, что с одной стороны такой выбор дает возможность получить довольно простые результаты а с другой стороны такая норма не меняется при любих ортогональных преобразованиях пространства. Потому если нет оснований предпочитать одно направление другим, то такой выбор естествен.

Решение уравнения (17) выражается явно при помощи т. н. функции Грина $G(x, x^{(0)})$,

$$u(x^{(0)}) = \int G(x, x^{(0)})\, l(x)\, \mathrm{d}x \tag{19}$$

для разных случаев. Анализ этого решения позволил получить ряд результатов, о которых пойдет речь ниже.

Оценка снизу, даваемая теоремой 1 по порядку является точной.

Разобъем область Ω на элементарные неперекрывающиеся области Ω_j с объемом kh^n, где k лежит в ограниченных пределах и составим множество кубатурных формул

$$\int_{\Omega_j} \varphi(x)\, \mathrm{d}x \cong \sum_{d(x^{(k)},\, \Omega_j) \leqq Lh} C_k^{(j)}\, \varphi(x^{(k)})\,, \quad L > 0\,, \tag{20}$$

где через d обозначено расстояние от точки $x^{(k)}$ до области Ω_j.

Допустим, что все формулы для Ω_j справедливы для любых многочленов степени $m - 1$, а коэффициенты этих формул не превосходят Ah^n.

Суммируя все такие формулы, получим

$$(21)\qquad \int_\Omega f\,\mathrm{d}x \cong \sum_k\Big(\sum_j C_k^{(j)}\,\varphi(x^{(k)})\Big) = \sum_k C_k\,\varphi(x^{(k)})\,.$$

Эту формулу мы будем называть формулой с равномерным распределением узлов и коэффициентов. Справедлива теорема

Теорема 2. *Норма функционала погрешности l кубатурной формулы с равномерным распределением узлов и коэффициентов удовлетворяет неравенству*

$$(22)\qquad \|l\|_{L_2^{(m)}} \leqq kN^{-m/n}\,,$$

где постоянная k зависит только от чисел L, A и не зависит от N.

Из теоремы 1 и 2 вытекает, что формулы с равномерным распределением узлов и коэффициентов являются по порядку оптимальными.

Более тонким является вопрос о наилучшем значении постоянной k в оценке (22). В общем виде этот вопрос представляет значительные трудности. Несколько ниже мы укажем его решение лишь для некоторых частных случаев. Пока что рассмотрим один специальный случай кубатурных формул.

Рассмотрим кубатурные формулы для периодических функций обладающих свойством

$$(23)\qquad \varphi(x + H\beta) = \varphi(x)\,,$$

где H-матрица периодов, а x и β векторы столбцы, причем x какой угодно вектор, а β вектор с целочисленными компонентами. Разумеется, простым афинным преобразованием $x = Hy$ эти функции могут быть переведены в функции с периодом 1 по каждому переменному. Пока речь идет о оптимальных формулах по порядке то такое преобразование не играет очевидно роль.

Функции, удовлетворящие условию периодичности (23) достаточно задать на некоторой „фундаментальной" области Ω_0. Их значения во всех остальных точках определяются с помощью этого условия. Легко убедиться в том, что эта область при отождествлении ее противоположеных границ изоморфна некоторому тору.

Мы будем сейчас рассматривать различные кубатурные формулы, узлы которых образуют правильную решетку с матрицей периодов

$$(24)\qquad H\,.\begin{pmatrix} 1/k_1, & 0, & 0, & \dots, & 0 \\ 0, & 1/k_2, & 0, & \dots, & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0, & 0, & 0, & \dots, & 1/k_n \end{pmatrix},$$

где $k_1, \dots, k_n$ — целые числа, т.е. такую, периоды которой укладываются в периодах H целое число раз. Для удобства предположим, что объем фундаментальной области Ω_0 равен единице. Простейший случай такой решетки, это решетка со сторонами hH, где $h = N^{-1/n}$ причем $N^{1/n}$ целое число.

Нетрудно заметить, что всякая финитная функция в E_n простым изменением масштаба приводится в область Ω_0 и может в силу этого, также, рассматриваться как периодическая, если изучать только тор.

Постоянная в оценке (22) существенным образом зависит от выбора решетки H.

Теорема 3. *Справедливо равенство*

$$\|l\|_{L_2^{(m)}(\Omega_0)} = h^m \sqrt{B_{n,m}(H')}\,, \tag{25}$$

где

$$B_{n,m}(H) = (2\pi)^{-n/2} \sum_{\gamma \neq 0} \frac{1}{(\gamma H^{-1} H^{-1*} \gamma^*)^m} \tag{26}$$

является обобщенным числом Бернули для решетки H, γ пробегает всевозможные целочисленные вектор строки и H^{-1} матрица обратная матрице периодов H. (H^{-1*} соотв. γ^* это транспонированная к H^{-1} матрица соотв. вектор столбец.)

При больших значениях m

$$B_{n,m} = 0(r_{\min}^{-2m})\,, \tag{27}$$

где $r_{\min}$ есть кратчайшее расстояние между узлами решетки с периодом $2\pi H^{-1}$, двойственной к решетке H. Теорема 3 позволяет поставить интересный с практической точки зрения вопрос.

Пусть $\varphi(x)$ задана в Ω_0 и бесконечно дифференцируема в этой области. Функционал ошибки l некоторой формулы является общим для последовательности пространств $L_2^{(m)}$, $m = 1, 2, \ldots, \infty$, $m > n$ и мы будем иметь для него ряд норм, определяемых формулой (25).

Данная функция $\varphi(x)$ также является элементом всех пространств $L_2^{(m)}$. Поэтому для ошибки конкретного интегрирования именно этой функции мы получим ряд оценок

$$|(l, \varphi)| \leqq h^m \sqrt{B_{n,m}(H)}\, \|\varphi\|_{L_2^{(m)}}(\Omega)_0\,. \tag{28}$$

Какая из этих оценок будет наилучшей при любом заданном φ? Каков порядок приближения в зависимости от h? Ответ на этот вопрос можно дать для классов функций с данным порядком и типом роста производных. Будем, как это принято сейчас, обозначать производные от φ символом D_φ^α, где α обозначает вектор $(\alpha_1, \ldots, \alpha_n)$ и пусть

$$|\alpha| = \alpha_1 + \alpha_2 + \ldots + \alpha_n\,, \quad \alpha_j \geqq 0\,.$$

Пусть еще

$$\alpha! = \alpha_1!\,\alpha_2! \dots \alpha_n!\,.$$

Пусть можно указать такие два числа $1 > \beta > 0$ и $1 > A > 0$ что для всех α выполняется при некотором k неравенство

$$|\alpha|^{(1-\beta)|\alpha|} \left| \frac{D^\alpha \varphi}{\alpha!} \right| < k(A + \varepsilon)^{|\alpha|} \tag{29}$$

и для бесконечно большого набора α и k

$$|\alpha|^{(1-\beta)|\alpha|} \left| \frac{D^\alpha \varphi}{\alpha!} \right| > k(A - \varepsilon)^{|\alpha|}\,. \tag{30}$$

Тогда число β будем называть порядком роста производных от φ, а число A типом этого роста. При разных β и A мы получим разные классы функций, иногда квазианалитические, иногда аналитические при всех вещественных x, иногда целые, а иногда просто тригонометрические многочлены.

Справедлива теорема

Теорема 4. *Для любой φ имеет место оценка*

$$|(l, \varphi)| \leqq k e^{-\frac{\beta}{e}\left(\frac{Ah}{2\pi e r_{\min}}\right)^{-1/\beta}}\,. \tag{31}$$

Эта оценка показывает, что порядок сходимости кубатурных формул для бесконечно дифференцируемых функций с ростом производных порядка β типа A гораздо более сильной, чем степенной. Оценка (31) получается из сравнения оценок (28) для разных m с выражением (29). Для оптимального значения m мы получим формулу

$$m = \left[\frac{1}{e}\left(\frac{Ah}{2\pi e r_{\min}}\right)^{-1/\beta}\right]. \tag{32}$$

Мы указали каким образом оценивается ошибка кубатурных формул для финитных и бесконечно дифференцируемых функций. Как мы видим оптимальной решеткой будет такая, в которой при единичном объеме фундаментальной области расстояние $r_{\min}$ будет наибольшим. Это расстояние очевидно равно радиусу шариков с центрами в узлах решетки, которые можно поместить в области так, чтобы они касались друг друга и не пересекались. Та решетка, для которой эта величина наибольшая, дает возможность занять шарами наибольший объем. Эта упаковка называется наиплотнейшей упаковкой. Мы получим теорему

Теорема 5. *При больших m наивыгоднейшей будет такая решетка периодов H, для которой взаимная с ней служит решеткой наиплотнейшей упаковки шаров.*

Задача об отыскании наиплотнейшей упаковки шаров решена Вороным, удобный алгорифм для небольших n предложен недавно Н. К. Игнатьевым [1964].

До сих пор не доказано, что наилучшая упаковка всегда найдется среди правильных решеток.

Для функций $\varphi(x)$, заданных в ограниченной области мы не можем указать оптимальных формул, т. к. уже самый выбор узлов такой формулы представляет собою задачу огромной трудности.

Теоретическое ее решение для любой области сводится к исследованию решений некоторых полигармонических уравнений в этой области, которое практически неосуществимо при достаточно больших N из-за очень больших вычислительных трудностей.

Однако, если ограничиться правильными решетками узлов, то эта задача асимптотически разрешима.

Как мы видели, очень важно равномерное распределение этих узлов и поэтому ограничение правильности решеток является естественным.

Покажем как это сделать.

Справедлива следующая теорема

Теорема 6. *Норма ошибки кубатурной формулы*

$$(l, \varphi) = \int_{\Omega} \varphi \, dx - \sum_{k=1}^{N} C_k \, \varphi(x^{(k)}) \cong 0 \tag{33}$$

при фиксированных узлах и переменных коэффициентах достигает наименьшего значения для такой системы значений C_k, при которых решения уравнения $u \in L_2^{(m)}(E_n)$

$$\Delta^m u = \mathscr{E}_{\Omega}(x) - \sum_{k=1}^{N} C_k \, \delta(x - x^{(k)}) \tag{34}$$

допускает решение обращающееся в нуль во всех точках $x^{(k)}$.

Таким образом отыскание оптимальных формул оказывается задачей, сводящейся к решению системы уравнений вида

$$\sum_{k=1}^{N} C_k \, G(x^{(k)}, x^{(l)}) = f^{(l)} \, . \tag{35}$$

Непосредственное ее решение в случае, если число независимых переменных n больше единицы, очень трудно т. к. эта система содержит как правило очень много неизвестных.

Однако, можно построит формулы, хотя и не являющиеся оптимальными, но асимптотически оптимальные.

Такими асимптотически оптимальными формулами служат формулы с регулярными пограничным слоем.

Для функций одной переменной такие формулы впервые были построены Gregori.

Их структура следующая. Область Ω покрывается правильной сеткой с матрицей периодов hH. Все точки этой сетки

$$x^{(\gamma)} = hH\gamma\,, \tag{36}$$

где γ-целочисленный вектор — столбец, разбиватся на два множества: B_2-состоящее из тех внутренних точек Ω, расстояние которых до границы области Ω не превосходит Lh, где L некоторая постоянная, зависящая от m и n но не от h, и B_1, в которое входят все внутренние точки Ω расстояние от которых до границы больше или равно Lh.

Кубатурная формула с регулярным пограничным слоем это формула, представимая в том же виде, как в условиях теоремы 2 и притом такая, что все коэффициенты C_γ для точке $x^{(\gamma)} \in B_1$ равны просто h^n.

Справедлива теорема

Теорема 7. *Формулы с регулярным пограничным слоем асимптотически локально оптимальны.*

Мы изложили некоторые результаты для частных случаев пространств H и множеств $\mathfrak{M}$. Важны и другие пространства, другие множества $\mathfrak{M}$ и другие функционали, чем простая интеграция. Покажем один такой частный случай.

Положим, что φ-функция одного переменного

$$F(\varphi) = \frac{1}{2\pi}\int_{-\pi}^{+\pi} e^{ipx}\,\varphi(x)\,\mathrm{d}x\,. \tag{37}$$

Пусть H — пространство всех 2π — периодических функций из $L_2^{(l)}$. Норму мы возьмем в $L_2^{(l)}$. Пусть $\mathfrak{M}$ конус как раньше т. е. $\mathfrak{M} \subset H'$ состоит из функционалов вида $f(\varphi) \in C\,\varphi(x_i)$. Можно показать, что локально оптимальная апроксимация

$$F(\varphi) \sim C(p, n, l)\frac{1}{n}\sum_{j=1}^{n} l^{ip2\pi j/n)}\,\varphi\left(\frac{j2\pi}{n}\right) \tag{38}$$

и

$$\frac{1}{C(p, n, l)} = \frac{1}{(2l-1)!}\frac{p^{2l}}{n^{2l}}\left(\frac{\partial^{2(l-1)}}{\partial z^{2(l-1)}}\left(\frac{\pi}{\sin \pi z}\right)^2\right)_{z=p/n} \tag{39}$$

и

$$a_n^{\mathfrak{M},F} = \frac{1}{p^l}\frac{1}{\sqrt{2\pi}}\left(1 - C(p, n, l)\right)^{\frac{1}{2}}\,. \tag{40}$$

Если $p = 0$, то $C(p, n, l) = 1$ для всех p, n, l (формулу для $a_n^{\mathfrak{M},F}$ в этом случае нужно изменить).

Мы видим, что существует случай, когда для всех l одновременно мы получаем оптимальные формулы.

Имеет практическое значение — строить там, где это возможно для $n \to \infty$ такие формулы, которые оптимальные одновременно для всех l.

К сожалению это возможно только в исключительных случаях. Если не достаточно ясно в которое пространство H вложить интегрируемую функцию (например которое l выбрать) то лучше всего использовать последовательности формул асимптотически оптимальных по возможности для самого широкого класса пространств H. Например если везмем $C(p, n, l) = 1$ то для всех l (для данного p) получим асимптотически локально оптимальную последовательность формул. Заметим что требование асимптотической оптимальности для самого широкого класса пространств очень важно с практической точки зрения и можно ее в большинстве случаев выполнить. Этим вопросом мы небудем здесь заниматься.

Как мы уже сказали, важное значение имеет вопрос вложения данной интегрируемой функции в самое хорошее H характеризованное например числом l. Некоторые соображения по этому поводу были приведены нами выше. Но и в том случае когда l мы знаем, для оценки ошибок нам необходимо еще знать $\|f\|_{L_2^{(l)}}$. Оценить величину $\|f\|_{L_2^{(l)}}$ можно иногда через разностные отношения. Такие оценки не являются строгими но могут принести пользу.

Как мы видели, вопрос о выборе формулы решается в зависимости от очень многих обстоятельств. Оптимизация для каждого конкретного случая формулы будет зависеть не только от выбора шага но и от того например сколько верных знаков может сохранить машина на которой мы считаем. В частности например формулы Котеса с большими знакотременными коэффициентами будут плохи при машинах данной точности и могут оказаться харошии для машин с большим числом знаков.

Хочется остановиться еще на одном вопросе: как можно наилучшим образом использовать постепенно вычисляющиеся значения функций. Другими словами, как использовать новую информацию, которую мы получаем во время вычисления. Этот вопрос мы не будем изучать. Как иллюстрацию приведем только один пример задачи аналогичной задаче наискорейшего спуска. Вопрос в построении последовательности элементов $f_1, f_2, \ldots, f_N \in \mathfrak{M}$ и чисел $C_i^{(n)}$, $i = 1, 2, \ldots, n$, $n = 1, 2, \ldots$ таких что

$$\inf_{C_i} \left\| F - \sum_{i=1}^{n} C_i f_i - f_{n+1} \right\| = \inf_{\substack{C_i \\ f \in \mathfrak{M}}} \left\| F - \sum_{i=1}^{n} C_i f_i - f \right\| . \tag{41}$$

Покажем пример того, как нужно в этих целях выбирать узлы в пространствах $L_2^{(1)}$ и $L_2^{(2)}$ на $\langle 0, 1\rangle$ когда даны $f_1(\varphi) = \varphi(0)$ и $f_2(\varphi) = \varphi(1)$. Получим таблицу

(42)

	$L_2^{(1)}$	$L_2^{(2)}$
1	$\varphi(0)$	$\varphi(0)$
2	$\varphi(1)$	$\varphi(1)$
3	$\varphi(\frac{1}{2})$	$\varphi(\frac{1}{2})$
4	$\varphi(\frac{1}{4})$	$\varphi(0{,}15407)$
5	$\varphi(\frac{3}{4})$	$\varphi(0{,}78750)$

Как видно в этом случае узлы уже не будут распределены равномерно.

До сих пор мы занимались некоторым классом формул, которые обыкновенно называются детерминированными.

С другой стороны очень важны и т. н. недетерминированные способы, где конкретный способ интегрирования выбирается случайным образом из некоторого множества способов интегрирования. Типический способ этого рода — методы Монте-Карло. Методы этого рода очень важны на практике, где достаточно пости всегда иметь пределы результата с большой вероятностью. Требовать, чтобы ошибки была, например, гораздо меньше вероятности возникновения мировой катострофы в ближайшие полчаса, явно не имеет смысла. И здесь имеются вопросы оптимизации, но мы не будем ими заниматься. Ср. Бахвалов [1964].

До сих пор мы предполагали, что имеем непрерывную модель над телом действительных чисел.

Но в действительности мы работаем с дискретной моделью.

Вопросы, которые возникают, на практике это вопросы численной устойчивости. Мы не будем заниматься подробно этим вопросом, потому-что он обсуждается на следующий день конференции. Мы коснемся этого с точки зрения оптимизации. Предположим, что мы имеем некоторую последовательность $f_1^{(j)}, f_2^{(j)}, \ldots, f_j^{(j)} \in \mathfrak{M}$ апроксимирующих функционалов. Очевидно, что

(43) $$b^{(j)} = \|F - \sum_{i=1}^{j} f_i^{(j)}\| \geqq a_j^{\mathfrak{M},F}.$$

Предположим, что мы работаем с расчетной δ-моделью.

Предположим, что $f_i^{(j)} = C_i^{(j)} f_i^{(j)*}$ и что $\sup\limits_{\|\varphi\| \leqq 1} \|f_i^{*(j)}(\varphi) - \tilde{f}_i^{*(j)}(\varphi)\| \leqq \delta$, $\|f_i^{*(j)}\| < k$ (мы обозначили $\tilde{f}_i^{(j)}(\varphi)$ и соответственно $f_i^{(j)}(\varphi)$ значение данного функционала в дискретной или соответственно в непрерывной модели; ясно что $\tilde{f}_i^{(j)}$ является нелинейным функционалом).

Через D будем обозначать конкретную δ-модель, т. е. и способ округления всех операций участвующих в вычислениях. Положим

$$\sup_{\|\varphi\|=1} \left\| \sum_{i=1}^{j} f_i^{(j)}(\varphi) - \sum_{i=1}^{j} \widetilde{\tilde{C}_i^{(j)} \tilde{f}_i^{j*}(\varphi)} \right\| = C_j^{[D,\delta]} . \tag{44}$$

Далее пусть

$$\sup_{D} C_j^{[D,\delta]} = C_j^{\delta} .$$

Пусть $B(\varepsilon)$ обратная фунция к $b^{(j)}$, принимающая очевидно лишь целые значения. Мы будем говорить, что способ интегрирования по порядку α устойчив, если (для малых ε)

$$C_{B(\varepsilon)}^{\delta} \leqq \delta K \varepsilon^{-\alpha} , \quad \text{где} \quad \alpha \geqq 0 . \tag{45}$$

Суть дела здесь в том, что для $\varepsilon \to 0$ число операций повышается и в большинстве способов $C_{B(\varepsilon)}^{\delta} \to \infty$ для $\varepsilon \to 0$.

Дело в том, чтобы затрата точности в связи с тем, что употребляется дискретная модель не была бы слишком большой.

Возможна тоже оптимизация с учетом влияния устойчивости. Но мы не будем этим заниматься.

3. ЛИНЕЙНЫЕ ЗАДАЧИ СО ВПОЛНЕ НЕПРЕРЫВНЫМИ ОПЕРАТОРАМИ

Предложим, что даны два линейные нормированные пространства

$$H \quad \text{и} \quad K$$

а пусть дан оператор A линейный, отображающий плотное подмножество $H_1 \subset H$ в K. Предположим, что наша задача состоит в решении уравнения

$$Ay = f , \quad y \in H , \quad f \in K \tag{46}$$

и предположим, что существует абсолютно непрерывный оператор A^{-1} (отображающий K в H).

Примером задач этого типа служат краевые задачи для дифференциальных уравнений.

Ряд методов для решения этих задач состоит

а) в приближенной замене пространств H и K n-мерным пространством H_n и K_n;

б) в приближенной замене линейного оператора A^{-1} линейным оператором A_n^{-1} отображающим K_n в H_n.

Приближенное решение y_n мы получим в форме

$$y_n = Q_n A_n^{-1} P_n f, \tag{47}$$

где P_n соотв. Q_n это отображение K на K_n соотв. H_n в H. Оператор ошибки L приближенного решения будет

$$L_n = A^{-1} - Q_n A_n^{-1} P_n. \tag{48}$$

Возникает аналогичный вопрос как в разделе 2 вопрос об минимальном возможном значении нормы оператора L в зависимости от n. Число n это — в некотором смысле — характеристика количества вычислительной работы.

Если по предположению A^{-1} абсолютно непрерывный, то отоборажение единичного шара из K в H будет компакт и вопрос сводится к вопросу наилучших приближений в компактах. Этот вопрос связан с теорией поперечников.

Изложим это понятие.

В произвольном линейном нормированном пространстве R расстоянием точки x от множества $F \subset R$ называется нижняя грань

$$\varrho(x, F) = \inf_{y \in F} \varrho(x, y), \tag{49}$$

$\varrho(x, y)$ это расстояние точек x, y, т. е. $\varrho(x, y) = \|x - y\|$. Отклонением $F \subset R$ от $G \subset R$ называется верхняя грань

$$\delta(F, G) = \sup_{x \in F} \varrho(x, G). \tag{50}$$

Пусть $S \subset R$ компакт и пусть $\mathfrak{M}$ некоторый конус из R.

Обозначим

$$d_n^{\mathfrak{M}}(S) = \inf_{L_n} \delta(S, L_n), \tag{51}$$

где L_n это n-размерный линеал, натянутый на n элементов из конуса $\mathfrak{M}$. $d_n^{\mathfrak{M}}$ будем называть относительным n-тым поперечником S по отношении к $\mathfrak{M}$.

Очевидно $d_{n+1}^{\mathfrak{M}} \leqq d_n^{\mathfrak{M}}$ и если $\mathfrak{M}$ плотное в S, то $d_n^{\mathfrak{M}} \to 0$ для $n \to \infty$. Выбранные n элементов из $\mathfrak{M}$ (или линеал $\hat{L}_n$) мы называем экстремальными, если

$$d_n^{\mathfrak{M}}(S) = \delta(S, \hat{L}_n). \tag{52}$$

На эти понятия впервые обратил внимание А. Н. Колмогоров [1936].

Будем называть последовательность пространств L_n, $n = 1, 2, \ldots$ по порядку экстремальной, если

$$\limsup_{n \to \infty} \frac{\delta(S, L_n)}{d_n^{\mathfrak{M}}(S)} < \infty. \tag{53}$$

Последовательность пространств L_n, $n = 1, 2, \dots$ асимптотически экстремальной, если

$$\frac{\delta(S, L_n)}{d_n^{\mathfrak{M}}(S)} \to 1 \quad \text{при} \quad n \to \infty . \tag{54}$$

По отношению к нашей задаче будем рассматривать $S = A^{-1}(T)$, T-единичный шар в K и $\mathfrak{M}$ пространство решений задачи.

Очевидно $\|L_n\| \geqq d_n^H(S)$.

Приведем пример.

Пусть мы имеем краевую задачу $Ay = f$ для эллиптического самосопряженного дифференциального положительно определенного оператора, $2l$-того порядка на $\Omega \in E_n$ для данных однородных краевых условий.

Пусть $K \equiv L_2$ и H пространство функций с конечной энергией, т. е.

$$\|y\|_H^2 = (Ay, y) \quad \text{и} \quad \mathfrak{M} = \hat{H} , \tag{55}$$

где $\hat{H}$ подпространство характеризирующее пространство решений (в зависимости от краевых условий); тогда верна следующая теорема:

Теорема 8. *Справедлива формула*

$$d_n^{\mathfrak{M}}(S) = \frac{1}{\sqrt{\lambda_{n+1}}} ,$$

где λ_n, это собственное число задачи

$$Ay - \lambda y = 0$$

и линеал L_n, натянутый на первые n собственных функций — это эксремальный линеал.

Отметим, что экстремальный линеал вообще говоря не определен однозначно.

Если мы имеем ввиду метод Ритца, то как известно, при данных координатных функциях $y_1, y_2, \dots, y_n$ решение задачи $Ay = f$ получается в форме $\sum_{i=1}^{n} C_i y_i$, где C_i зависит от f и определяется как решение системы линейных алгебраических уравнений. Это значит, что для данных y_i метод Ритца дает приближенный оператор $A^{-1}_{[y_1, \dots, y_n]}$ отображающий K в H. Из того, что мы говорили, следует, что норма оператора ошибки

$$\|A^{-1} - A^{-1}_{[y_1, \dots, y_n]}\| \tag{56}$$

не может быть меньше, чем $d_n^{\mathfrak{M}}(S)$.

Если $y_1, y_2, \dots, y_n$ таковы, что натянутое на них подпространство совпадает с подпространством натянутым на первые собственные функции, мы получим минимально возможное значение нормы оператора ошибки L_n.

Мы видим, что поперечники $d_n^{\mathfrak{M}}$ это максимально возможно получаемая точность. Таким образом поперечник может быть масштабом эффективности метода. Оценки поперечников и конструкция экстремальных линеалов важные теоретические задачи. Существует ряд работ, изучающих эти вопросы. Напомним работы Колмогорова, Тихомирова и других.

С практической точки зрения нахождение экстремального линеала, например ситемы собственных функций может оказаться сложнее, чем решение нашей задачи. Это значит, что нахождение экстремального линеала практически неудобно.

Но нахождение экстремальных по порядку линеалов или аисмптотически экстремальных линеалов в ряде случаев просто. Этот нахождение основано на методе подобных задач.

Изложим главную мысль в случае по порядку экстремальных линеалов.

Предположим, что мы имеем две самосопряженные положительно определенные задачи для дифференциального оператора $2l$-того порядка для достаточно гладких областей и однородных краевых условий

$$Ay = f \quad \text{и} \quad A^*y = f^* \tag{57}$$

и пусть A и A^* имеют равные области определения и

$$C_2\|y\|^2_{W_2^{(l)}} \leqq (Ay, y) \leqq C_1\|y\|^2_{W_2^{(l)}}\,, \tag{58}$$

$$C_2^*\|y\|^2_{W_2^{(l)}} \leqq (A^*y, y) \leqq C_1^*\|y\|^2_{W_2^{(l)}}\,. \tag{59}$$

Предположим еще что A^*A^{-1} и AA^{*-1} линейные в K операторы. Такие две задачи будем называть подобными. Верна теорема

Теорема 9. *Экстремальные по порядку линеалы для одной из двух подобных задач являются по порядку экстремальными и для другой задачи.*

Приведем пример. Пусть мы имеем задачу

$$-(ay')' + by = f\,, \quad y(0) = y(1) = 0\,, \tag{60}$$

и задачу

$$-y'' = f\,, \quad y(0) = y(1) = 0\,. \tag{61}$$

Предположим $a > 0$, $b \geqq 0$, что a' ограниченная функция.

Эти две задачи подобны.

Экстремальные линеалы для второй задачи получим, если применим функцию $y_j = \sin \pi jx$. Из того, что мы сказали следует, что эти функции, будут создавать по порядку экстремальные линеалы и для задачи первой.

Экстремальные линеалы не определены однозначно. Например, для задачи (61) система функций $\varphi_j(x)$

$$\varphi_i(x) = \begin{cases} \left(x - \dfrac{i-1}{N}\right) N\,; & \dfrac{i-1}{N} \leqq x \leqq \dfrac{i}{N}\,, \\ 1 - \left(x - \dfrac{i}{N}\right) N\,; & \dfrac{i}{N} \leqq x \leqq \dfrac{i+1}{N}\,, \\ 0; \quad 0 < x < \dfrac{i-1}{N}\,, & 1 > x > \dfrac{i+1}{N}\,, \end{cases} \tag{62}$$

$$i = 1, \ldots, N-1$$

составляет экстремальный линеал.

Как уже мы сказали приближенное решение, это линейное отображение пространства K в n-размерное подпространство пространства H или другими словами

$$y_n = \sum_{i=1}^{n} C_i^{(n)}(f)\, u_i^{(n)}\,, \tag{63}$$

где $C_i^{(n)}(f)$ некоторые функционалы из K'.

До сих пор мы занимались вопросом оптимальнлго выбора если $C_i^{(n)}(f)$ любые функционалы. В некоторых задачах, как это было в частности в теории кубатурных формул, возникает вопрос о том, как выбирать $u_i^{(n)}$, если $C_i^{(n)}$ могут быть только функционалами из некоторого конуса допустимых функционалов $\mathfrak{M}$, например типа

$$C_i(f) = f(x_i)\,.$$

Итак пусть $\hat{H}$ линейное пространсто, натянутое на компакт S, где $S = A^{-1}(T)$, T – единичный шар в K.

Пусть даны линейные функционалы $C_i^{(n)}(f) \in K'$, $i = 1, 2, \ldots, n$ и обозначим $B_{C_i^{(n)}}^{u_i^{(n)}}$ линейное отображение определенное равенством

$$B_{C_i^{(n)}}^{u_i^{(n)}}(f) = \sum_{i=1}^{n} C_i^{(n)}(f)\, u_i^{(n)} \tag{64}$$

и

$$\sigma_{C_i^{(n)}} = \inf_{u_i^{(n)} \in \hat{H}} \left\| B_{C_i^{(n)}}^{u_i^{(n)}} - A^{-1} \right\|.$$

Функции $u_i^{(n)}$ называем локально экстремальными, если $\left\| B_{C_i^{(n)}}^{u_i^{(n)}} - A^{-1} \right\| = \sigma_{C_i^{(n)}}$ Далее пусть

$$b_n^{\mathfrak{M}} = \inf_{C_i^{(n)} \in \mathfrak{M}} \sigma_{C_i^{(n)}}\,. \tag{65}$$

Функционалы $\hat{C}_i^{(n)}$ и функции $\hat{u}_i^{(n)}$ такие, что $b_n^{\mathfrak{M}} = \left\| B_{\hat{C}_i^{(n)}}^{\hat{u}_i^{(n)}} - A^{-1} \right\|$ называем экстремальными.

Большое значение имеют оценки $\sigma_{C_i^{(n)}}$ и $b_n^{\mathfrak{M}}$ и конструкция локально экстремальных и экстремальных функционалов и функций. Ясно, тоже, каково будет определение по порядку или асимптотически экстремальных функционалов и функций.

Пусть K пространство Гильберта и пусть даны функционалы C_i, $i = 1, 2, \ldots, n$. Обозначим Q подпространство K такое, что, если $f \in Q$, то $C_i(f) = 0$, $i = 1, 2, \ldots, n$.

Пусть $Q^\perp$ ортогональное дополнение Q в K. Очевидно $Q^\perp n$ - мерное подпространство. Если сейчас взять в качестве координатных функций $u_i^{(n)} = A^{-1} f_i^{(n)}$, где $f_i^{(n)} \in Q^\perp$, $C_i(f_j^{(n)}) = \delta_i^{(j)}$ ($\delta_i^{(j)}$ символ Кронеккера), то $u_i^{(n)}$ будут локально экстремальными и для них можно получить некоторые важные соотношения.

Покажем пример.

Пусть имеем краевую задачу

$$y''(x) = f(x), \quad y(0) = y(1) = 0 . \tag{66}$$

Пусть

$$K = L_2^{(l)}, \quad l \geqq 1, \quad H = C .$$

Пусть

$$C_j^{(n)}(f) = f(x_j), \quad x_j = \frac{j}{N}, \quad j = 0, 1, \ldots, N \tag{67}$$

и пусть $\hat{u}^{(n)}$ взяты так, как мы указали выше. Обозначим

$$^{(i)}y_{n+1} = y_{n+1}(x_i) = \sum_{j=0}^{n} C_j^{(n)}(f)\, \hat{u}_j^{(n)}(x_i), \quad f(x_i) = f_i ; \tag{68}$$

тогда получим для y_{N+1} систему уравнений

(69) 1) $l = 1$,

$$^{(0)}y = {}^{(N)}y = 0,$$

$$^{(i-1)}y - 2\,^{(i)}y + {}^{(i+1)}y = \frac{h^2}{6}\left[f_{i-1} + 4f_i + f_{i+1}\right] ;$$

(70) 2) $l = 2$, $^{(0)}y = {}^{(N)}y = 0$, $^{(-1)}y = -\,^{(1)}y + h^2 f_0$, $^{(N+1)}y = -\,^{(N-1)}y + h^2 f_N$,

$$f_{-1} = 2f_0 - f_1, \; f_{N+1} = 2f_N - f_{N-1},$$

$$^{(i-2)}y + 2\,^{(i-1)}y - 6^{(2)}y + 2\,^{(i+1)}y + {}^{(i+2)}y = \frac{h^2}{20}\left[f_{i-2} + 26f_{i-1} + 66f_i + 26f_{i+1} + f_{i+2}\right] .$$

Все рассмотрения этого раздела были сделаны нами с точки зрения непрерывной модели. В дискретной модели, как мы уже говорили ранее, играет очень важную роль численная устойчивость и мы должны каждый раз более конкретно изучать алгорифмы в целом.

Покажем это на простом примере метода Ритца. Экстремальное множество L_n мы можем получить через разные функции. С точки зрения результата в случае непрерывной модели все эти функции равноценны. Но с точки зрения дискретной модели, т. е. с точки зрения устойчивости, это не так. Очевидно матрица Грамма, которую мы должны обратить при решении методом Ритца (или решить данную систему уравнений) зависит от выбора координатных функций. Очевидно, если функции, создающие базис будут ,,почти" линейно зависимы, то при решении данной системы возникает большая ошибка за счет того, что мы работаем с дискретной моделью, другими словами мы получим неустойчивый алгорифм. Но этими вопросами мы уже заниматься не будем.

4. ОБЩИЕ ИДЕИ И ОЦЕНКИ

Оценки, характерные для линейных задач, связанные с поперечником множества непосредственно не применимы к задачам нелинейным, так как в них не определяются линейные многообразия состоящие из решений.

А. Н. Колмогоровым [1956] и его ученикам введены еще несколько важных понятий более общего характера, позволяющих найти границу для необходимого числа действий при решении вычислительных задач. Значение оценок, получаемых таким образом особенно важно потому, что оно дает возможность планомерного поиска в случаях когда есть несовпадения асимптотики, то есть когда оценки сверху и снизу расходятся.

Решение каждой задачи состоит в выборе одного из ответов среди множества $\mathfrak{M}$ всех возможных. Имея в руках идеальный метод распознания удачного ответа от неудачного и отбрасывания подмножества не содержащего верный ответ мы должны будем совершать каждый раз в процессе решения несколько двоичных выборов, для того, чтобы его индивидуализировать. Число таких двоичных выборов будет равно числу двоичных разрядов, нужных для того, чтобы переменить все элементы множества, т. е. двоичному логарифму числа элементов, среди которых нужно сделать выбор.

Пусть нам нужна точность ε. Тогда мы должны исследуемое нами множество ответов заменить некоторым конечным числом $N(\varepsilon)$ элементов $u_\varepsilon^{(j)}$, $j = 1, 2, \ldots, N(\varepsilon)$ таким, что любой из элементов множества $\mathfrak{M}$ будет находиться в ε-близости от какого-либо $u_\varepsilon^{(j)}$. Говорят, что $u_\varepsilon^{(j)}$ образуют конечную ε-сеть.

Существование конечной ε-сети для любого ε равносильно требованию компактности множества $\mathfrak{M}$. Число

$$H(\varepsilon) = \inf \lg_2 N(\varepsilon)\,, \tag{71}$$

называемое ε-энтропией множества, характеризует работу, нужную для выделения индивидуального элемента ε-сети и поэтому оно может служить мерой трудности ε-приближения к элементам $\mathfrak{M}$. Рост при $\varepsilon \to 0$ является весьма важной характеристикой компакта. Эта характеристика предложена А. Н. Колмогоровым [1956].

Решая задачу вычислительной математики очень интересно с теоретической и практической точки зрения сравнить между собой труд T_1, необходимый для решения ее с тем трудом $T_2 = H(\varepsilon)$, который нужно затратить, чтобы характеризовать конкретно ту функцию от которой зависит ответ в классе возможных функций. Конечно мы должны возможно ограничить заранее этот класс. Интуитивно можно ожидать что $T_2(\varepsilon)$ минимальный необходимый труд для решения задач вычислительной математики если класс возможных ответов характеризован достаточно точно.

Оказывается, что при достаточно малых ε для многих задач существуют методы при использовании которых эти величины будут одного порядка. Поскольку при постепенном уменьшении ε дальнейший процесс распознания всегда ведется уже в малом участке исследуемого компакта, значит порядок роста ε-энтропии определяется локальной топологией того пространства, где мы ищем ответ. Таким образом оказалось, что труд по разысканию ответа задач вычислительной техники во многих случаях по порядку совпадает с ε энтропией множества элементов от которых этот ответ зависит.

В различных работах, принадлежащих советским и другим исследователям, даны оценки числа действий необходимых для решения многих вычислительных задач с нужной точностью через ε-энтропию. Эти оценки дают возможность оценить число действий не только сверху, но и снизу. При условии совпадения по порядку числа действий с нижней оценкой мы будем говорить о неулучшаемом по порядку методе.

Приведем некоторые относящиеся сюда результаты. А. Н. Колмогоров [1956], А. Витушкин [1959], В. Ерохин, В. Тихомиров [1960] и И. В. Арнольд получили оценки ε-энтропии для широкого круга классов аналитических функций и дифференцируемых функций конечной гладкости.

Для класса функций s переменных $C_s^{(r)}(A)$ у которых все производные до порядка r не превосходит A

$$H(\varepsilon) \asymp \left(\frac{A}{\varepsilon}\right)^{s/r}. \tag{72}$$

Для классов аналитических функций $f(z_1, \ldots, z_s)$ s-переменных ограниченных некоторой постоянной A на некотором множестве значений переменных $z_1, \ldots,$ $\ldots, z_s$, введем норму как максимум функции f на некотором подмножестве. Мы получим типичную оценку ε-энтропии

$$H(\varepsilon) \asymp \ln^{s+1}\left(\frac{1}{\varepsilon}\right). \tag{73}$$

В работе Витушкина [1959] даны оценки снизу объема вычислительной работы нужной для восстановления функций этих классов по заданной информации. Эти оценки связаны именно с ε-энтропией.

Коснемся еще некоторых других вопросов и мыслей в связи с числом действий для решения задач вычислительной математики.

При решении задачи Коши для обыкновенного уравнения 1-го порядка

$$y' = f(x, y)$$

в классе функций, где $|y^{(p+1)}| \leqq A$, $|f_y| < L$, классические методы типа Адамса обеспечивают оценку сверху погрешности как $0(N^{-p})$. Можно показать, что лучшей оценки получить нельзя, так как этот класс содержит в себе уравнения вида $y' = f(x)$ для которого оценка снизу, вытекающая из теоремы 1′ по порядку совпадает с указанной.

Изучим интегральное уравнение

$$\varphi(x) = \int_0^1 K(x, y)\,\varphi(y)\,\mathrm{d}y + f(x)\,, \tag{74}$$

$$K(x, y) \in C_{r,r}(A)\,, \quad f(x) \in C_2(A)\,.$$

Если $\|K\| \geqq 1$, то преобразуем уравнение так, чтобы выполнялось условие $\|K\| < 1$.

Для этого можно применить к уравнению (74) оператор типа $(E - K^*)^{-1}$, где K^* вырожденный оператор и $\|K - K^*\|$ мала. Затем заменяем интегралы суммами и решаем системы методом итераций.

Асимптотика числа действий $\varepsilon^{-2/r} \ln 1/\varepsilon$. Оценка снизу дает величину $\varepsilon^{-2/r}$.

Для ядер из классов типа $H_s^\alpha(A)$ рассмотрена задача в работах Коробова [1959a], Шахова [1961] и др.

Для некоторых типов классов и уравнений оценки сверху и снизу сходятся.

Для классов параболических и гиперболических уравнений с переменными коэффициентами из какого-то класса $C_{r_0,r_1,\ldots,r_s}(A)$ наилучшими, повидмному, будут методы переменных направлений с соотношением шагов $h_0^{r_0} \approx h_1^{r_1} \approx$ $\approx h_2^{r_2} \approx \ldots \approx h_s^{r_s}$. Оценки сверху и снизу сходятся лищь для некоторых частных случаев.

В случае, когда коэффициенты относятся к разным классам, например уравнение

$$u_t = u_{xx} + au + f\,,$$

$$a \in C_{r_1,r_2}\,, \quad f \in C_{r_1,r_2}\,,$$

задача, не решена в простейших случаях.

Возможно, что на пути решения задачи оптимизации на таких классах возникнут новые методы решения и новые методы оценок снизу.

Коснемся некоторых вопросов когда решение более гладкое в некоторой части области, где можно поэтому применять более редкую сетку метода конечных разностей.

а) Рассиотрим уравнение теплопроводности.

Пусть

$$u_t = u_{xx}\,, \quad u(t, 0) = u(t, 1) = 0\,, \quad u(0, x) = u_0(x)\,.$$

Предположим дальше, что после продолжения нечетным образом в точке 0 и 1 функции $u_0(x)$, получим на любом отрезке длины 1 функцию из класса $L_2^{(r)}$.

Оценка снизу количества значений $u_0(x)$, которое нужно для получения значения $u(t_0, x_0)$ решения с точностью ε, имеет порядок $\varepsilon^{-1/r}$.

Построена сетка, содержащая $\approx \varepsilon^{-1/r}$ узлов, при интегрировании, по которой производятся $\varepsilon^{-1/r}$ действий, причем удается получить таблицу решений при $t \leqq T_0$. С таблицы решение интерполируется в любую точку за конечное число действий.

б) Уравнение Лапласа (Бахвалов [1961]).

Пусть

$$\Delta u = 0\,, \quad u|_\Gamma = \varphi\,, \quad \varphi \in C_r(A)\,, \quad \Gamma \in \Omega_r(A)\,, \quad r \leqq 3$$

(гладкость границы — r-производных).

Оценка снизу количества действий имеем порядок $\varepsilon^{-1/2}$. Аналогично тому как в прошлом случае, с помощью специально построенно сетки можно получить оценку сверху порядка $\varepsilon^{-1/r} \lg_2 1/\varepsilon$.

Если $\Gamma \in \Omega_{r+\delta}(A)$ $\delta > 0$, то для получения решения во внутренней подобласти будем имет оценку сверху и снизу порядка $\varepsilon^{-1/r}$.

В случае более сложных эллиптических уравнений и эллиптических систем часто не проведены оценки сверху и снизу. Поэтому там стоит сравнивать оценки объема труда при решении сеточных уравнений.

1) Уравнение Пуассона в параллелепипеде размерности s. Получается система порядка h^{-s} уравнений. Метод который предложили Peaceman, Rachford [1955] и Douglas, Rachford [1956] (ср. также Varga [1962], стр. 209) требует $0(h^{-s} \lg 1/h \lg 1/\varepsilon)$ действий.

2) Эллиптическое уравнение

$$\sum_{\substack{i=1\\j=1}}^{s} a_{ij} u_{x_i x_j} + 2 \sum_{i=1}^{s} a_i u_{x_i} + a_0 u = f\,.$$

Получается система порядка h^{-s} уравнений. Метод Younga [1954] требует при $a_0 < 0$, $0(h^{-s-1} \lg 1/\varepsilon)$ действий.

Для эллиптических систем хороших оценок сверху нет.

Для бигармонического уравнения уже в прямоугольнике при граничных условиях $u|_\Gamma = \varphi$, $\partial u/\partial^n|_\Gamma = \psi$ неизвесто оценок сверху лучших h^{-4}.

Авторы доклада выражают благодарность Н. С. Бахвалову оказавшему им большую помощь.

Бахвалов Н. С. [1959]: О приближенном вычислении кратных интегралов. Вест. МГУ 3 (1959).

Бахвалов Н. С. [1959a]: О численном решении задачи Дирихле для уравнения Лапласа. Вестн. Моск. ун-та. Сер. мат. физ. 5, 171—195 (1959).

Бахвалов Н. С. [1960]: Оценка в среднем остаточного члена квадратурных формул. Журн. вычисл. матем. и матем. физ. 1, 64 (1960).

Бахвалов Н. С. [1963]: Об оценке количества вычислительнои работы необходимой при приближнном решении задач. Доп. к книги Годунов-Рябеyький: Введенне в теорию разностных сеем. Москва (1962).

Бахвалов Н. С. [1964]: Об оптимальных оценках сходимости квадратурных процессов и методов интегрирования типа Монте-карло на классах функций. Численные методы решения дифф. и инт. уравнений и квадратурные формулы. Москва (1964).

Белага Э. Г. [1958]: Некоторые вопросы вычисления многочленов. ДАН СССР 123, 775 (1958).

Белага Э. Г. [1961]: О вычислении значений многочленов от одного переменного с предварительной обработкой коэффициентов. Проблемы кибернетики 5, 7 (1961).

Brauer A. [1939]: On addition chains. BAMS 45, 736— 739 (1939).

Чжан-Гуань-Цзюань [1962]: О минимальном числе узлов при численном интегрировании уавнения теплопроводности. Журн. вычисл. матем. и матем. физ. 2, 1. 80—88 (1962).

Davis P. J. [1959]: On the Numerical Integration of Periodic Analytic Functions. Proc. of Symp., Wisconsin Press (1959).

Douglas J. J., Rachford H. H. Jr. [1956]: On the Numerical Solution of Heat Conduction Problems in Two or Three Space Variables. Trans. Amer. Math. Soc. 82, 421 (1956).

Eve J. [1964]: The Evaluation of Polynomials. Numerische Mathematik 6, 17 (1964).

Golomb M., Weinberger H. F. [1959]: Optimal Approximation and Errors Bounds. Proceedings of a Symposium Conducted by the Mathematics Research Center, United States Army at the University of Wisconsin, Madison, Apr. 21—23, 1958; The University of Wisconsin Press (1959).

Игнатыв Н. К. [1964]: Об одном практическом приеме нахождения плотных упаковок и мерных сфер. Сибир. матем. журн. V, 815 (1964).

Kolmogoroff A. [1936]: Über die beste Annäherung von Funktionen einen gegebenen Functionenklassen. Ann. of Math. 37, 107 (1936).

Колмогоров А. [1956]: О некоторых аисмптотических характеристиках вполне ограниченных метрических пространств. ДАН СССР 108, 3 (1956).

Коробов Н. М. [1957]: Приближенное вычисление кратных интегралов с помощью метода теории чисел. ДАН СССР 115, 1062 (1957).

Коробов Н. М. [1959]: О приближенном вычислении кратных интегралов. ДАН СССР 124, 1207 (1959).

Коробов Н. М. [1959]: Вычисление кратных интегралов методом оптимальных коэффициентов Вест. МГУ 4, 19 (1959).

Коробов Н. М. [1959a]: О приближенном решении интегральных уравнений. ДАН СССР 128, 235 (1959).

Motzkin T. S. [1955]: Evaluation of Polinomials. BAMS 61, 2 (1955).

Никольский С. М. [1950]: К вопросу об оценках приближений квадратурными формулами. УМН V, 2, 165 (1950).

Ostrowski A. M. [1954]: Studies in Mathematics and Mechanics 44. (Presented to R. von Mises New York.) Academic Press (1954).

Пан В. Я. [1959]: Некоторые схемы для вычисления многочленов с вещественными коэффициентами. ДАН СССР 127, 266 (1959).

Пан В. Я. [1961]: Некоторые схемы для вычисления значений полиномов с вещественными коэффициентами. Проблемы кибернетики 5, 17 (1961).

Пан В. Я. [1962]: О некоторых способах вычислений значений многочленов. Проблемы кибернетики 7, 21 (1962).
Peaceman D. W., Rachford H. H. Jr. [1955]: The Numerical Solution of Parabolic and Elliptic Differential Equations. J. of Soc. Industr. Appl. Math. 3, 28 (1955).
Sard A. [1949]: Best Approximate Integration Formulas; Best Approximation Formulas. Amer. J. Math. 71, 80 (1949).
Sard A., Meyers L. F. [1950]: Best Approximate Integration Formulas. J. Math. Phys. 29, 118 (1950).
Смоляк С. А. [1960]: Интерполяционные и квадратурные формулы на классах W_s^α и E_s^α. ДАН СССР 131, 1028 (1960).
Соболь И. М. [1960]: Точная оценка погрешности многомерных квадратурных формул для функций класса S_p. ДАН СССР 132, 1041 (1960).
Солодов В. М. [1959]: О вычислении кратных интегралов. ДАН СССР 127, 753 (1959).
Шайдаева Т. А. [1959]: Квадратурные формулы с наименшей оценкой остатка для некоторых классов функций. Труды Мат. Инст. им. Стеклова 53, 313 (1959).
Шахов Ю. Н. [1961]: О приближенном решении уравнений Вольтерра II. рода методом итераций. ДАН СССР 136, 1302 (1961).
Тихомиров В. М. [1960]: Поперечники множеств в функциональных пространствах и теория найлучших приближений. УМН 15, 3, 81 (1960).
Тихомиров В. М., Колмогоров А. Н. [1959]: ε-энтропия и э-емкость множеств в функциональных пространствах.
Вальский Р. Э. [1959]: О наименшем числе умножений для возведения в данную степень. Проблемы кибернетики 2, 73 (1959).
Varga R. S. [1962]: Matrix Iterative Analysis. Prentice Hall Inc. Englewood cliffs, New Yersey (1962).
Витушкин А. Г. [1959]: Оценка сложности задачи табулирования. ГИФМЛ, Москва (1959).
Young D. [1954]: Iterative Methods for Solving Partial Difference Equations of Elliptic Type. Trans. Amer. Math. Soc. 76, 92 (1954).

Dr. Sc. *Ivo Babuška,* Matematický ústav ČSAV, Praha 1, Žitná 25, ČSSR.

Академик *С. Л. Соболев,* Сибирское отделение АН СССР, Новосибирск 72, СССР.