УДК 519.6

Н. М. ШЕРЫХАЛИНА, А. А. ОШМАРИН

ЧИСЛЕННАЯ ФИЛЬТРАЦИЯ ДАННЫХ, ИСКАЖЕННЫХ НЕРЕГУЛЯРНОЙ ПОГРЕШНОСТЬЮ

В статье рассматривается применение численной фильтрации в качестве средства оценки погрешности, уточнения и проверки достоверности полученных результатов. Разрабатываются критерии надежности оценок. Приводятся конкретные примеры, на которых демонстрируется работоспособность предложенных методов фильтрации в случаях, когда обычные методы не дают желаемого результата. *Численная фильтрация, оценка погрешности, критерии надежности оценок, численные методы, экстраполяция*

ВВЕДЕНИЕ

В условиях отсутствия аналитических способов решения, разность между приближенным и точным результатом является самой надежной оценкой погрешности. Но такой метод может быть использован только для очень немногих примеров, которые имеют аналитическое решение. Для других примеров применение такой оценки очень ненадежно. Выход может быть найден, если вместо точного использовать приближенное, но более точное по сравнению с проверяемым, значение. Это более точное значение может быть всего в три раза точнее, чем проверяемое. Однако при этом возникают два вопроса: как получить это более точное значение и как проверить, что оно действительно точнее исходного.

И проверяемое, и более точное значение вычисляются одним и тем же способом. Но часто получить более точное значение невозможно, так как это приводит к дополнительным требованиям к ресурсам, которые могут оказаться невыполнимыми. Но существует и другой способ: использовать более грубые результаты (с меньшим числом узлов и временем счета). Если погрешность метода подчиняется некоторому закону, то, зная этот закон (в виде характера зависимости, например степенной, экспоненциальный и т. п.), можно по нескольким результатам провести идентификацию и экстраполяцию, и приближенно предсказать значение, соответствующее бесконечному числу узлов [1].

Ответить на второй вопрос можно с помощью повторной экстраполяции, т.е. экстраполяцией экстраполированных результатов, полученных для разных наборов исходных данных [2]. В этом случае получается оценка погрешности экстраполированных результатов (или размытость оценки погрешности). Цель достигнута, если эта оценка удовлетворяет требованиям: в три и более раз меньше оценки погрешности исходных данных (относительная размытость меньше 1/3). Если не удовлетворяет, то данный способ оценки в конкретном случае следует признать ненадежным. Подробнее критерии надежности будут рассмотрены ниже.

Кроме того, при хороших оценках результаты экстраполяции можно использовать вместо вычисленных данных как более точные. При этом, чтобы убедиться в надежности полученных таким способом результатов, необходима дополнительная экстраполяция. В некоторых случаях путем повторной экстраполяции можно получить результаты, на многие порядки более точные, чем рассчитанные непосредственно с помощью численного метода, и чего невозможно было бы добиться прямым расчетом в связи с огромными затратами времени, превышающими разумные пределы.

1. ПОСТАНОВКА ЗАДАЧИ

1. 1. Численная фильтрация

При экстраполяции требуется априорное знание характера зависимости результата расчетов от числа узлов (или математической модели погрешности), например следующим образом:

$$z_n = z + c_1 n^{-k_1} + c_2 n^{-k_2} + \ldots + c_L n^{-k_L} + \delta(n), \quad (1)$$

где z – точное значение; z_n – приближенный результат, полученный при числе узловых точек, равном n; c_j – коэффициенты, которые предполагаются независящими от n; $\delta(n)$ – величина, полагаемая малой по сравнению с $c_j n^{-k_j}$ при тех значениях n, которые использовались в данных конкретных расчетах, $k_1,\ldots, k_L$ – произвольные действительные числа (предполагается, что $k_j>0$, $k_1< k_2<\ldots< k_L$).

В математическом анализе обычно оценивается только первый член, поскольку остальные являются асимптотически (при $n\to\infty$) бесконечно малыми более высокого порядка. Однако для конечных n остальные слагаемые могут вносить существенный вклад и должны приниматься во внимание.

Если решение задачи не особенное, то можно допустить возможность его разложения по формуле Тейлора, тогда k_j – это часть ряда натуральных чисел.

Решение задачи численной фильтрации есть последовательное устранение степенных слагаемых суммы (1) с помощью вычисления линейных комбинаций результатов, полученных для различных n.

Пусть $q = n_{j-1}/n_j < 1$, $m>L$. В соответствии с (1) составим систему линейных алгебраических уравнений

$$z_{n_1} = z + c_1 + c_2 + \ldots + c_L + v_1,$$

$$z_{n_2} = z + c_1 q^{k_1} + c_2 q^{k_2} + \ldots + c_L q^{k_L} + v_2,$$

....

$$z_{n_m} = z + c_1 q^{(m-1)k_1} + c_2 q^{(m-1)k_2} + \ldots + c_L q^{(m-1)k_L} + v_m$$

или в матричном виде

$$y = Ax + v,\ x = \begin{pmatrix} z \\ c_1 \\ \vdots \\ c_L \end{pmatrix},\quad y = \begin{pmatrix} z_{n_1} \\ \vdots \\ z_{n_m} \end{pmatrix},$$

$$A = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 1 & q^{k_1} & \cdots & q^{k_L} \\ . & . & . & . \\ 1 & q^{(m-1)k_1} & \cdots & q^{(m-1)k_L} \end{pmatrix},\quad v = \begin{pmatrix} \delta(n_1) \\ \vdots \\ \delta(n_m) \end{pmatrix}. \qquad (2)$$

1. 2. Метод минимизации дисперсии ожидаемой погрешности

Задача ставится следующим образом. Рассмотрим систему m линейных алгебраических уравнений (2) с $n=L+1$ неизвестными, предполагая наличие погрешности в каждом уравнении.

Значения погрешностей неизвестны, поэтому система (2) всегда недоопределенная, независимо от соотношения m и n и от ранга матрицы A. В зависимости от подхода вектор погрешностей v характеризуется либо нормой (максимальным значением $|v_i|$, либо дисперсиями σ^2_i, либо матрицей ковариаций). Математические ожидания $M(v_j)$, как правило, принимают равными нулю.

Требуется найти оценку решения этой системы, причем такую, чтобы она была наилучшей в смысле определенного критерия. Кроме того, необходимо найти оценку погрешности результата.

Пусть для системы уравнений (2) матрица A постоянна и известна точно. Заданы области неопределенности ожидаемой изменчивости вектора погрешностей v и искомого вектора x. Квадратические размеры областей определяются ковариационными матрицами K_v и K_x

$$K_v = \begin{pmatrix} \sigma_1^2 & r_{12}\sigma_1\sigma_2 & \ldots & r_{1m}\sigma_1\sigma_m \\ r_{12}\sigma_1\sigma_2 & \sigma_2^2 & \ldots & r_{2m}\sigma_2\sigma_m \\ \ldots & \ldots & \ldots & \ldots \\ r_{1m}\sigma_1\sigma_m & r_{2m}\sigma_2\sigma_m & \ldots & \sigma_m^2 \end{pmatrix},$$

где σ_j^2 – дисперсии случайных величин v_j, r_{ij} – коэффициенты корреляции v_i и v_j, (матрица K_x определяется аналогично).

Необходимо найти оптимальный способ определения оценки x вектора искомых параметров x по результатам измерений или вычислений $y_1, \ldots, y_n$.

В качестве критерия оптимальности принимается минимум математического ожидания квадрата ошибки неизвестного x_1, т.е.

$$F_1 = M|x_1 - x_1|^2 \to \min. \qquad (3)$$

Оценку x будем искать в виде $x = By$, где B – матрица размерности $n\times m$ (т.е. ограничимся рассмотрением линейных операторов B).

Задача, аналогичная (3), с критерием $F_1 = M|x - x|^2 \to \min$ решена в [3]

$$B = \left(A^T K_v^{-1} A + K_x^{-1}\right)^{-1} A^T K_v^{-1}$$

Или если $K_x^{-1} = 0$

$$B = \left(A^T K_v^{-1} A\right)^{-1} A^T K_v^{-1} \qquad (4)$$

Как оказалось, решение задачи (3) совпадает с (4).

Тем самым задача определения неизвестного z решается методом наименьших квадратов.

Отметим что, при L=1, m=2 получается формула фильтрации, которая совпадает с Ричардсоновской экстраполяционной формулой

$$z_{n_2}^* = z_{n_2} + \frac{z_{n_2} - z_{n_1}}{Q^{k_j} - 1} \qquad (5)$$

Проводя последовательно экстраполяцию по всем парам соседних значений, получим отфильтрованную зависимость, не содержащую члена с n^{-k_j}

$$z_n^* = z + c_1^* n^{-k_1} + \ldots + c_{j-1}^* n^{-k_{j-1}} + c_{j+1}^* n^{-k_{j+1}} + $$
$$+ c_L^* n^{-k_L} + \delta^*(n), \qquad (6)$$

где

$$c_l^* = c_l \frac{Q^{k_j} - Q^{k_l}}{Q^{k_j} - 1} \qquad (7)$$

Заметим, что отфильтрованная последовательность $z_{n_j}^*$ содержит на один член меньше, чем исходная. Если она содержит больше одного члена, то ее также можно отфильтровать, устранив степенную составляющую с n^{-k_l}. Операции фильтрации можно повторять последовательно для $n^{-k_1}, \ldots, n^{-k_L}$, если исходная последовательность содержит достаточное количество членов. Результаты экстраполяций удобно представлять в виде треугольной матрицы

n	z_n	z_n^*	z_n^{**}	z_n^{***}	$\ldots$
10	z_{10}	–	–	–	
20	z_{20}	z_{20}^*	–	–	
40	z_{40}	z_{40}^*	z_{40}^{**}	–	
80	z_{80}	z_{80}^*	z_{80}^{**}	z_{80}^{***}	
…					

(8)

Применение повторной экстраполяции известно под названием метода Ромберга. Но при его применении возникает ряд ограничений, не позволяющих автоматически использовать полученные путем экстраполяции значения. Это требует разработки методов повышения достоверности полученных оценок.

2. СРЕДСТВА ПОВЫШЕНИЯ ДОСТОВЕРНОСТИ

2. 1. Критерий размытости оценки

Оценка погрешности по правилу Рунге сводится к сравнению значения z_n с экстраполированным значением z_n^*. Поскольку эта оценка справедлива при допущении, что величина z_n^* точнее, чем z_n, то необходима проверка справедливости этого допущения.

Это можно сделать следующим образом. Повторим процесс экстраполяции и получим значение z_n^{**}. Разность $\Delta_n = z_n - z_n^*$ представляет собой оценку погрешности приближенного значения z_n. Разность $\Delta_{\Delta n} = z_n^* - z_n^{**}$ является оценкой погрешности экстраполированного значения z_n^* или оценкой погрешности оценки погрешности (рис. 1). Отношение $\delta_n = |\Delta_{\Delta n}/\Delta_n|$ имеет смысл относительной размытости оценки погрешности.

Если $\delta_n << 1$, то это означает, что относительная размытость оценки Δ_n мала, и такой оценке можно доверять.

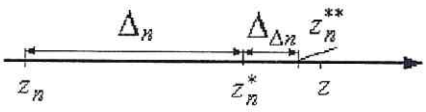

Рис. 1. Размытость оценки погрешности

Сначала постановим, что оценка погрешности представляется в виде интервала $z = z_n^* \pm \Delta_n$. Для определения порогового значения δ_n для принятия или отклонения полученной оценки желательно на основании имеющейся информации установить, не может ли при гипотетическом продолжении экстраполяций произойти переход получающихся значений левее $z_n = z_n^* - \Delta_n$ или правее $z_n^* + \Delta_n$. Для этого предположим, что при последующих гипотетических экстраполяциях значение δ_n как коэффициента уменьшения расстояния между соседними экстраполированными значениями будет сохраняться: $\delta_n = \delta$. Тогда максимальное удаление предельного значения от z_n^* определяется суммой геометрической прогрессии $\Delta_{\max} = \Delta_n^*/(1-\delta)$.

Отсюда следует неравенство

$$K\frac{\Delta_n^*}{1-\delta} \le \Delta_n,$$

где $K \ge 1$ – коэффициент «запаса» надежности оценки. Необходимость введения коэффициента K вызвано желанием получать достаточно надежные оценки в условиях неопределенности, вызванной влиянием нерегулярных составляющих погрешности. Тогда получим условие (критерий принятия оценки):

$$\delta \le \frac{1}{K+1}. \qquad (9)$$

Примем величину K=2. Тогда пороговое значение $\delta = 1/3$, при $\delta \le 1/3$ оценка принимается, а при $\delta > 1/3$ отвергается. Практика показывает [3], что этот критерий действительно «работает».

2. 2. Визуализация результатов экстраполяции

Результаты экстраполяции и оценки погрешности представляются в виде зависимости $-\lg \Delta$ (десятичного логарифма относительной погрешности) от $-\lg n$.

На рис. 2-4 кривые *0* соответствует погрешности расчетных данных, кривые *1-3* результатам первой, второй и третьей экстраполяции.

Визуализация позволяет проводить фильтрацию в интерактивном режиме и принимать решение о достоверности оценки на основе совместного анализа совокупности полученных путем экстраполяции данных.

3. ПРИМЕРЫ ПРИМЕНЕНИЯ ФИЛЬТРАЦИИ

Численная фильтрация применялась для обработки результатов, полученных различными численными методами. В качестве примера на рис. 2 приведены результаты вычисления производной по симметричной разностной формуле второго порядка, на рис. 3 результаты численного интегрирования методом правых прямоугольников. Экстраполяция проводилась методом Ромберга с визуальным контролем результатов. В результате анализа можно утверждать, что вычисленные значения интегралов имеют 1-5 точных десятичных знаков, результат первой экстраполяции 3-10 знаков при относительной размытости менее 0,1.

Ограничение на уровне 16 знака объясняется погрешностью округления (применялись числа с двойной точностью). При этом следует отметить увеличение погрешности округления при численном дифференцировании с ростом *n*. При численном интегрировании этот рост погрешности округления менее заметен.

Аналогично выглядят результаты обработки численных данных, полученных при решении дифференциальных уравнений (обыкновенных и с частными производными).

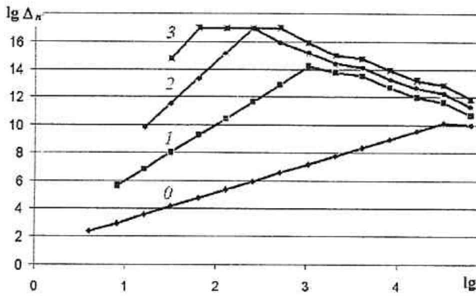

Рис. 2. Результаты экстраполяции для численного дифференцирования

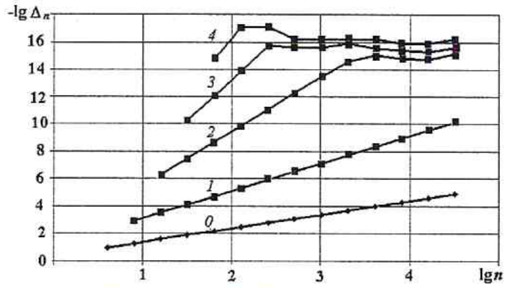

Рис. 3. Результаты экстраполяции для численного интегрирования

Экстраполяция часто осложняется необходимостью нахождения максимальных значений распределенных параметров. Различные положения экстремума относительно узлов сетки приводит к возникновению нерегулярной составляющей погрешности, затрудняющей анализ. В этом случае применение формулы (5) (рис. 4,*а*, кривая *2*) позволяет только оценить погрешность (около 4-х знаков) и размытость оценки $\delta<0{,}01$, так как последующие экстраполяции не повышают точность. Уточнить результат невозможно, так как соответствующая кривая искажена нерегулярной составляющей.

В результате применения общей формулы фильтрации (4) становится возможной повторная экстраполяция с отчетливо различимой линией *3* (рис.4,*б*). Поэтому можно оценить погрешность результата первой экстраполяции (более 6 знаков при размытости порядка 0,2), что существенно улучшает результат и повышает надежность оценки. Следует, однако, отметить, что получение диаграммы, приведенной на рис. 4,*а*, требуется около 30 вычисленных значений, а для рис. 4,*б* – в 25 раз больше.

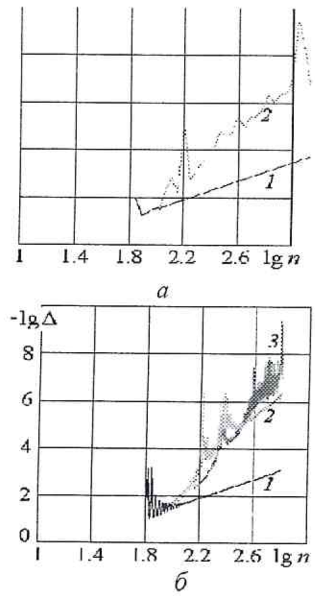

Рис. 4. Результаты экстраполяции при наличии нерегулярных составляющих погрешности

ЗАКЛЮЧЕНИЕ

Как было показано, применение численной фильтрации на основе экстраполяции позволяет существенно увеличить точность численных результатов с малыми затратами времени. Кроме того, анализ нескольких составляющих погрешности повышает достоверность оценок. Этому также способствует применение критерия размытости и визуализация результатов обработки в виде представленных графиков.

Многокомпонентный анализ помогает также определить несовершенства алгоритма, которые не заметны из-за наличия более значительных составляющих погрешности.

СПИСОК ЛИТЕРАТУРЫ

1. **Zhitnikov, V. P.** Application of extrapolation methods of numerical results for improvement of hydrodynamics problem solution / N. M. Sherykhalina, V. P. Zhitnikov // Computational Fluid Dynamics Journ. 2002. V. 11, No. 2. P. 155–160.

2. **Житников, В. П.** Методы верификации математических моделей в условиях неопределенности / В. П. Житников, Н. М. Шерыхалина // Вестник УГАТУ. 2000. № 2. С. 53–60.

3. **Житников, В. П.** Линейные некорректные задачи. Верификация численных результатов / В. П. Житников, Н. М. Шерыхалина, А. Р. Ураков. Уфа, 2002. 90 с.

4. **Житников, В. П.** Решение плоских и осесимметричных задач с помощью методов теории функции комплексного переменного : учеб. пособие / В. П. Житников. Уфа : УГАТУ, 1994. 106 с.

5. **Житников, В. П.** Применение экстраполяции для оценки погрешности и уточнения численного решения нестационарных задач электрохимического формообразования / В. П. Житников, О. Р. Зиннатуллина, Г. И. Федорова // Вычислительные технологии. 2006. Т. 11. С. 82–93.

6. **Житников, В. П.** Применение численной фильтрации как средства уточнения результатов вычисления и оценки погрешности / В. П. Житников, О. Р. Зиннатуллина, Н. М. Шерыхалина, А. А. Ошмарин // Гидродинамика больших скоростей и численное моделирование : матер. 3-й летн. науч. шк. Кемерово, 2006. С.73–77.