ИСПОЛЬЗОВАНИЕ АФФИННОЙ АРИФМЕТИКИ ДЛЯ ВНУТРЕННЕГО ОЦЕНИВАНИЯ ОБЛАСТЕЙ ЗНАЧЕНИЙ ФУНКЦИЙ

Р. Р. Ахмеров
Алтайский государственный университет, Россия

In this paper, a method for inner estimation of the ranges of values of functions is proposed. The method is based on using affine arithmetic, which was originally developed for outer estimation of the ranges of values. We describe the technique for the construction of inner estimates is described and discuss some its applications: tracing the accuracy of outer estimations, checking the solvability of the systems of nonlinear equations, taking into account constraints in global optimization problems.

Введение

Техника внешнего оценивания областей значений функций хорошо изучена и с успехом используется для решения широкого круга задач вычислительной математики. В качестве примера приведём две популярные задачи непрерывного анализа: глобальная оптимизация и решение систем нелинейных уравнений. До изобретения методов внешнего оценивания эти задачи считались практически нерешаемыми [2].

Одним из наиболее универсальных и простых методов построения внешних оценок является интервальный метод, основанный на применении интервальной арифметики и интервальных расширений элементарных функций [3, 4]. Этот метод позволяет построить внешнюю интервальную оценку области значений функции, если известны внешние интервальные оценки её аргументов.

Главным недостатком интервального оценивания является его тенденция давать слишком широкие оценивающие интервалы, которые часто оказываются практически безполезны (проблема зависимости, см. [4, 5]). В целях улучшения внешних интервальных оценок было разработано несколько альтернативных методов [3, 5], одним из которых является так называемая *аффинная арифметика.*

В данной работе представлена новая возможность применения аффинной арифметики — построение внутренних оценок областей значений функций.

1. Аффинная арифметика

В этом параграфе мы приводим краткое описание аффинной арифметики и техники аффинного оценивания. Подробности смотрите в [1, 5].

1.1. Аффинные формы

В аффинной арифметике частично неизвестная величина x представлена *аффинной формой* $\hat{x}$, которая является многочленом первой степени:

$$\hat{x} = x_0 + x_1\varepsilon_1 + x_2\varepsilon_2 + \cdots + x_n\varepsilon_n.$$

Коэффициенты x_i есть числа с плавающей точкой, а ε_i — символьные вещественные переменные, чьи значения не известны, но приняты лежащими в интервале $\mathbf{U} = [-1, 1]$. Число x_0 называют центром аффинной формы $\hat{x}$, коэффициенты x_i — частичными отклонениями, а ε_i — символами шума.

Аффинные формы неявно выражают частичные зависимости между величинами. Когда две аффинные формы совместно используют общие символы шума, то это означает, что величины, ими представляемые, по крайней мере, частично зависят одна от другой. Принятие таких зависимостей во внимание позволяет аффинной арифметике существенно повысить точность оценок по сравнению с чисто интервальной. Особенно сильно это проявляется в длинных вычислительных цепочках, когда результаты одних алгоритмов становятся входными параметрами для других. Конечно, аффинная модель более трудоемка, но многочисленные примеры показывают, что повышенная трудоемкость почти всегда окупается точностью получаемых оценок.

1.2. Применение для внешнего оценивания

Если нам дана функция $f : \mathbb{R}^n \to \mathbb{R}$ и брус $\mathbf{X} \subseteq \mathbb{R}^n$, то для поиска внешней оценки f на $\mathbf{X}$ аффинная арифметика используется следующим образом. Вначале все входные интервалы — компоненты бруса $\mathbf{X}$ — представляются в виде аффинных форм. Интервалу $[a, b]$ будет соответствовать аффинная форма вида

$$\frac{a+b}{2} + \frac{b-a}{2}\varepsilon_k,$$

где ε_k — некоторый вновь введённый символ шума. После этого f оценивается в аффинной арифметике. В результате мы получаем некоторую аффинную форму $\hat{y}$, которую в свою очередь оцениваем извне интервалом. Если $\hat{y} = y_0 + y_1\varepsilon_1 + y_2\varepsilon_2 + \cdots + y_s\varepsilon_s$, то для любых $\varepsilon_i \in [-1, 1]$ значения $\hat{y}(\varepsilon)$ будут принадлежать интервалу

$$\left[y_0 - \sum_{i=1}^{s} |y_i|\,,\ y_0 + \sum_{i=1}^{s} |y_i| \right],$$

а значит этот интервал будет внешней оценкой области значений f на $\mathbf{X}$.

При оценивании f в аффинной арифметике используется такая же методика, как и при оценивании f в интервальной арифметике. Для этого все элементарные операции и функции в записи f заменяются на соответствующие им операции и функции над аффинными формами, после чего проводятся вычисления, дающие в результате аффинную форму.

1.3. Операции над аффинными формами

Опишем общую идею, лежащую в основе вычислений над аффинными формами.

Пусть некоторые величины x и y представлены соответственно аффинными формами $\hat{x}$ и $\hat{y}$:

$$\begin{aligned} \hat{x} &= x_0 + x_1\varepsilon_1 + x_2\varepsilon_2 + \cdots + x_n\varepsilon_n, \\ \hat{y} &= y_0 + y_1\varepsilon_1 + y_2\varepsilon_2 + \cdots + y_n\varepsilon_n. \end{aligned}$$

Тогда если $z \leftarrow f(x, y)$, то z можно записать в виде :

$$z \;=\; f(x,y) \;=\; f\left(x_0 + \sum_i x_i\varepsilon_i, y_0 + \sum_i y_i\varepsilon_i \right) \;=\; f^*(\varepsilon_1, \ldots, \varepsilon_n),$$

где $f^*(\varepsilon_1, \ldots, \varepsilon_n)$ есть функция, действующая из $\mathbf{U}^n$ в $\mathbb{R}$. Если f^* нелинейна, то z не может быть выражена в виде линейной комбинации символов шума ε_i. В этом случае мы должны подобрать некоторую линейную функцию от ε_i,

$$f^a(\varepsilon_1, \ldots, \varepsilon_n) = z_0 + z_1\varepsilon_1 + \cdots + z_n\varepsilon_n$$

которая достаточно хорошо приближает $f^*(\varepsilon_1, \ldots, \varepsilon_n)$ на $\mathbf{U}^n$, и затем добавить дополнительный член $z_k\varepsilon_k$ для представления ошибки аппроксимации. Таким образом мы получим

$$\hat{z} = f^a(\varepsilon_1, \ldots, \varepsilon_n) + z_k\varepsilon_k = z_0 + z_1\varepsilon_1 + \cdots + z_n\varepsilon_n + z_k\varepsilon_k.$$

Символ шума ε_k должен отличаться от всех других символов шума, которые появлялись ранее в едином вычислительном процессе, а коэффициент z_k должен сверху оценивать максимальную ошибку аппроксимации функции $f^*(\varepsilon_1, \ldots, \varepsilon_n)$ функцией $f^a(\varepsilon_1, \ldots, \varepsilon_n)$.

Если операция линейная, то $\hat{z}$ можно вычислить легко:

$$\begin{aligned} \hat{x} + \hat{y} &= (x_0 + y_0) + (x_1 + y_1)\varepsilon_1 + \cdots + (x_n + y_n)\varepsilon_n, \\ \hat{x} - \hat{y} &= (x_0 - y_0) + (x_1 - y_1)\varepsilon_1 + \cdots + (x_n - y_n)\varepsilon_n, \\ \alpha\hat{x} &= (\alpha x_0) + (\alpha x_1)\varepsilon_1 + \cdots + (\alpha x_n)\varepsilon_n, \qquad \alpha \in \mathbb{R}. \end{aligned}$$

Заметим, что на практике даже при линейных операциях над аффинными формами приходится вводить новые символы шума, чтобы учесть ошибки округлений.

Для всех основных вещественных элементарных функций и операций были разработаны их аффинные эквиваленты, что дало возможность выполнять аффинное оценивание для широкого и практически важного класса функций.

2. Внутреннее оценивание

2.1. Структура аффинной формы

Предположим, что в результате аффинного оценивания некоторой функции $f : \mathbb{R}^n \to \mathbb{R}$ на брусе $\mathbf{X}$ была получена аффинная форма

$$\hat{x} = x_0 + x_1\varepsilon_1 + \cdots + x_l\varepsilon_l.$$

Заметим, что символы шума ε_i в записи $\hat{x}$ не равнозначны. Действительно, часть из них была введена для представления входных интервалов. Другая часть была добавлена на этапе вычисления операций над аффинными формами. Между этими двумя типами символов есть принципиальное различие. В чем оно состоит?

Будет считать, что символы $\varepsilon_1, \dots, \varepsilon_n$, $n < l$, представляют входные интервалы. Этого всегда можно добиться путем подходящего переобозначения.

Напомним определение символа шума: ε_i — это символьные вещественные переменные, чьи значения не известны точно, но приняты лежащими в интервалах $\mathbf{U} = [-1, 1]$, причём всегда полагается, что символы с разными индексами независят друг от друга. Но про каждый ε_i для $i = 1, \dots, n$ мы может сказать, что он не просто лежит в интервале $[-1, 1]$, но еще и *принимает все значения* из этого интервала, независимо от других символов. Иными словами, первые n символов шума представляют собой, по существу, не точечные величина, а множества. Можно ли то же самое сказать про символы с индексами большими n? Для ответа на этот вопрос обратимся к источнику возникновения новых символов.

Символ шума возникает при выполнении операции над аффинными формами, когда нам нужно учесть и оценить

- либо ошибку округления при выполнении операции с числами с плавающей точкой,
- либо ошибку аппроксимации при приближении нелинейной функции линейной.

Это означает, что каждый вновь введённый символ шума фактически *является функцией существующих символов*, а значит функцией входных символов $\varepsilon_1, \dots, \varepsilon_n$. Когда мы полагаем символы шума независимыми друг от друга, это просто дает нам удобный для внешнего оценивания областей значений способ оперирования ими, но всегда надо помнить об их возможной взаимосвязи. Таким образом, мы можем записать

$$\varepsilon_i = \varepsilon_i(\varepsilon_1, \dots, \varepsilon_n) \quad \text{при} \quad i > n.$$

Более того, если исходная функция f непрерывна на $\mathbf{X}$, то и все ε_i для $i > n$ непрерывны на $\mathbf{U}^n$.

Вышеизложенное даёт нам право утверждать, что аффинная форма $\hat{x}$ есть некий иной вид представления функции f на брусе $\mathbf{X}$, который заключается в выделении линейной части (линейного ядра) f и в структурировании и оценивании нелинейной.

2.2. Система обозначений

В предыдущем пункте мы показали, что символы шума в записи аффинной формы можно разбить на две группы:

1. Символы, представляющие независимые величины, принимающие все значения из интервала $[-1, 1]$.
2. Символы, являющиеся функциями величин, представляемых символами первого типа, про значения которых известно только то, что они принадлежат интервалу $[-1, 1]$.

Для удобства изложения дальнейшего материала и по причине принципиального различия этих двух типов символов, мы будем называть символы первого типа *символами диапазона* и обозначать через α_i. Название и обозначение символов второго типа мы оставим прежним — символы шума, ε_i. Такой выбор нам представляется наиболее благоразумным, так как выражает суть этих понятий.

Тогда форму $\hat{x}$ можно записать в виде:

$$\hat{x} = x_0 + x_1\alpha_1 + \cdots + x_n\alpha_n + \tilde{x}_1\varepsilon_1 + \cdots + \tilde{x}_s\varepsilon_s = x_0 + \sum_{i=1}^{k} x_i\alpha_i + \sum_{i=1}^{s} \tilde{x}_i\varepsilon_i \tag{1}$$

или, введя обозначения $\alpha = (\alpha_1, \dots, \alpha_n)$, $\varepsilon = (\varepsilon_1, \dots, \varepsilon_s)$, $x = (x_1, \dots, x_n)$ и $\tilde{x} = (\tilde{x}_1, \dots, \tilde{x}_s)$, в виде:

$$\hat{x} = x_0 + \langle x, \alpha\rangle + \langle \tilde{x}, \varepsilon\rangle,$$

где острыми скобками обозначено скалярное произведение векторов.

2.3. Условие существования и метод вычисления внутренних оценок

Пусть нам дана аффинная форма $\hat{x}$ вида (1). Как уже отмечалось, $\hat{x}$ можно воспринимать как функцию $\hat{x}(\alpha)$, действующую из $\mathbf{U}^n$ в $\mathbb{R}$.

Обозначим через D область значений $\hat{x}(\alpha)$ на $\mathbf{U}^n$. Тогда внешней интервальной оценкой для D будет интервал

$$\mathbf{y} = \left[\, x_0 - \sum_{i=1}^{n} |x_i| - \sum_{i=1}^{s} |\tilde{x}_i| \,, x_0 + \sum_{i=1}^{n} |x_i| + \sum_{i=1}^{s} |\tilde{x}_i| \,\right],$$

причём это наилучший интервал, который можно получить при имеющейся у нас информации.

Обозначим $\mathbf{y}^* = [\,\inf\{d \mid d \in D\}, \sup\{d \mid d \in D\}\,]$ — интервальная оболочка D, наименьший по отношению включения интервал, содержащий D. Величина

$$e = \sup\{|\bar{\mathbf{y}} - \bar{\mathbf{y}}^*|\,, |\underline{\mathbf{y}} - \underline{\mathbf{y}}^*|\} = \sup\{\bar{\mathbf{y}} - \bar{\mathbf{y}}^*, \underline{\mathbf{y}}^* - \underline{\mathbf{y}}\}$$

есть ошибка оценки $\mathbf{y}$.

Утверждение 1.

$$e \le 2 \sum_{i=1}^{s} |\tilde{x}_i|\,.$$

Доказательство. Так как α_i могут принимать все значения из $[-1, 1]$, то, в частности, можно положить $\alpha_i = \operatorname{sgn} x_i$, $i = 1, 2, \ldots, n$. Тогда

$$\hat{x}(\alpha) = x_0 + \sum_i |x_i| + \sum_i \tilde{x}_i \varepsilon_i(\alpha) \;\in\; \left[x_0 + \sum_i |x_i| - \sum_i |\tilde{x}_i|\,,\; x_0 + \sum_i |x_i| + \sum_i |\tilde{x}_i| \right].$$

Так как $\hat{x}(\alpha) \in \mathbf{y}^*$, то

$$\bar{\mathbf{y}}^* \ge x_0 + \sum_i |x_i| - \sum_i |\tilde{x}_i| = \bar{\mathbf{y}} - 2 \sum_i |\tilde{x}_i|\,.$$

Следовательно $\bar{\mathbf{y}} - \bar{\mathbf{y}}^* \le 2 \sum_i |\tilde{x}_i|\,.$

Аналогично, взяв все $\alpha_i = -\operatorname{sgn}(x_i)$, получим $\underline{\mathbf{y}}^* - \underline{\mathbf{y}} \le 2 \sum_i |\tilde{x}_i|$. Тем самым *утверждение доказано.*

Утверждение 2. Если все ε_i, $i = 1, \ldots, s$, непрерывны на $\mathbf{U}^n$ и $\sum_i |\tilde{x}_i| \le \sum_i |x_i|$, то

$$\mathbf{y}_* = \left[\, x_0 - \sum_{i=1}^{n} |x_i| + \sum_{i=1}^{s} |\tilde{x}_i|\,,\; x_0 + \sum_{i=1}^{n} |x_i| - \sum_{i=1}^{s} |\tilde{x}_i| \right] \subseteq D,$$

т.е. является внутренней интервальной оценкой D.

Доказательство. Если ε_i непрерывны, то область значений аффинной формы связна и $D = \mathbf{y}^*$. Из доказательства Утверждения 1 следует, что концы $\mathbf{y}_*$ принадлежат D, а значит $\mathbf{y}_* \subseteq D$.

Подытоживая вышеизложенное можно сказать, что используя аффинную арифметику мы можем вычислить внешнюю интервальную оценку области значений вещественнозначной функции, оценить ошибку этой оценки и, в ряде важных случаев вычислить внутреннюю интервальную оценку.

2.4. Случай векторнозначной функции

Если функция f действует из $\mathbb{R}^n$ в $\mathbb{R}^m$, то ее область значений на брусе тоже можно оценить изнутри, используя аффинную арифметику. Этот случай значительно сложнее, так как здесь недостаточно построить интервальные внутренние оценки областей значений компонент-координат f. Здесь важно учитывать еще и взаимосвязи координат.

Автором разработан метод внешнего и внутреннего оценивания областей значений векторнозначных функций многогранниками специального вида. Но это тема отдельной статьи.

3. Приложения внутреннего оценивания

Ниже перечислены некоторые приложения внутреннего оценивания областей значений:

- Проверка существования решений систем нелинейных уравнений.

- Учет ограничений при решении задачи глобальной оптимизации методом ветвей и границ.
- Контроль за точностью внешних оценок и проверка критерия остановки в алгоритмах, использующих дробление области для повышения точности внешних оценок.

Надо заметить, что для всех приведенных выше задач разработано много методов [2, 4], не использующих внутреннего оценивания. Поэтому имеет смысл объединить эти методы с алгоритмами построения внутренних оценок, что вероятно должно увеличить эффективность получающихся новых методов.

Список литературы

[1] Comba J. L. D. and Stolfi J. Affine arithmetic and its applications to computer graphics// Proc. of VI SIBGRAPI, 1993. P. 9–18.

[2] Hansen E. R. Global Optimization Using Interval Analysis. N. Y., Marcel Dekker, 1992.

[3] Moore E. R. Methods and Applications of Interval Analysis. SIAM, Philadelpia, 1979.

[4] Neumaier A. Interval Methods for Systems of Equations. Cambridge, Cambridge University Press, 1990.

[5] Stolfi J. and de Figueiredo L. H. Self-Validated Numerical Methods and Applications. Notes of 21^{st} Brazilian Mathematics Colloquium, 1997.