Апостериорные интервально-аффинные вычисления

Ахмеров Р.Р.

Алтайский государственный университет

Аффинная арифметика

Оперирует аффинными формами вида:

$$\hat{\mathsf{x}} = a_0 + a_1\varepsilon_1 + a_2\varepsilon_2 + \cdots + a_s\varepsilon_s, \tag{1}$$

где $a_i \in \mathbb{R}$ и ε_i — независимые переменные, такие что $\operatorname{ran}\varepsilon_i = [-1, 1] \subset \mathbb{R}$ для любого i, и $a_i \neq 0$ при $i > 0$.
Область значений аффинной формы $\hat{\mathsf{x}}$ вида (1) это интервал

$$\operatorname{ran}\hat{\mathsf{x}} = [a_0 - \sum_{i=1}^{s} |a_i|, a_0 + \sum_{i=1}^{s} |a_i|].$$

Любому интервалу $\boldsymbol{y} \in \mathbb{IR}$ можно сопоставить аффинную форму $\hat{\mathsf{y}} = \operatorname{mid}(\boldsymbol{y}) + \operatorname{rad}(\boldsymbol{y})\varepsilon_k$. И наоборот, любой аффинной форме $\hat{\mathsf{z}}$ можно сопоставить интервал, равный области ее значений $\operatorname{ran}\hat{\mathsf{z}}$.

Оценивание областей значений производных

Пусть $f(x) = f(x_1, x_2, \ldots, x_n) \colon \mathbb{R}^n \to \mathbb{R}$ и f определена и дифференцируема на брусе $\boldsymbol{x} \in \mathbb{IR}^n$. Обозначим

$$\nabla f = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \ldots, \frac{\partial f}{\partial x_n} \right)$$

— градиент f. Во многих случаях нужно строить внешнюю интервальную оценку для множества значений ∇f на брусе $\boldsymbol{x}$, т.е. такой брус $\boldsymbol{d} \in \mathbb{IR}^n$, что выполняется

$$\{\, \nabla f(x) \mid x \in \boldsymbol{x} \,\} \subset \boldsymbol{d}$$

Примеры:

- Среднезначная форма (Mean value form)
- Тесты в интервальных методах глобальной оптимизации
- Решение систем нелинейных уравнений

Сложность построения оценок “напрямую”

Функция f от n переменных записана в виде программы, состоящей из μ операций. Для оценки области значений f в интервальной арифметике нужно выполнить $O(\mu)$ операций. В аффинной арифметике $O(\mu n)$ операций. Если производные оценивать напрямую по их аналитической записи независимо друг от друга, то для оценки ∇f в интервальной арифметике потребуется $O(\mu n)$ операций, в аффинной арифметике $O(\mu n^2)$ операций.

	f	∇f	Ошибка оценки
Интервальная арифметика	$\sim \mu$	$\sim \mu n$	$O(\operatorname{diam} \boldsymbol{x})$
Аффинная арифметика	$\sim \mu n$	$\sim \mu n^2$	$O((\operatorname{diam} \boldsymbol{x})^2)$

Апостериорный метод оценивания производных

Прямой проход:

Строим оценку f, сохраняя выполненные операции.

$$\longrightarrow \quad \longrightarrow \quad \longrightarrow \quad \longrightarrow \quad \longrightarrow \quad \longrightarrow \quad \longrightarrow \quad \longrightarrow$$

Обратный проход:

По программе вычисления f строим оценки для $\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n}$.

$$\longleftarrow \quad \longleftarrow \quad \longleftarrow \quad \longleftarrow \quad \longleftarrow \quad \longleftarrow \quad \longleftarrow \quad \longleftarrow$$

Матиясевич, Ю.В.: Вещественные числа и ЭВМ, Кибернетика и вычислительная техника, 2 (1986), стр.104–133.

	f	∇f
Интервальная арифметика	$\sim \mu$	$\sim 6\mu$
Аффинная арифметика	$\sim \mu n$	$\sim 6\mu n$

Численные эксперименты

Решались задачи глобальной оптимизации различными модификациями метода ветвей и границ:

- Int — На базе чистой интервальной арифметики.
- Aff — На базе чистой интервально-аффинной арифметики.
- MGInt — На базе интервальной арифметики с использованием среднезначной формы и теста градиента.
- MGAff — На базе интервально-аффинной арифметики с использованием среднезначной формы и теста градиента.

Численные эксперименты

Сравнивались параметры работы алгоритмов:

OptEncl — оценка для оптимума функции,

wEncl — ширина оценки для оптимума функции,

NBoxes — число брусов в рабочем списке после завершения алгоритма,

BoxWid — ширина максимального бруса из рабочего списка,

NBis — число выполненных бисекций,

Time — время работы алгоритма в секундах.

Критерий остановки алгоритмов: ширина максимального бруса в рабочем списке меньше $\varepsilon = 1e - 07$ или число брусов в рабочем списке превысило допустимый предел (2000).

Задача eg1

Задача eg1 (сборник задач А. Ноймайера). Найти минимум функции на брусе:

$$f(x, y, z) = x^2 + (yz)^4 + xz + y\sin(x + z) + y,$$

$$-10 \leq x \leq 10,$$

$$-1 \leq y \leq 1,$$

$$1 \leq z \leq 2.$$

Метод	OptEncl	wEncl	NBoxes	BoxWid	NBis	Time
Int	[-1.461, -1.426]	0.0352	2000	0.625	2733	6
Aff	[-1.429.., -1.429..]	2.3e-08	2000	0.00048	4783	104
MGInt	[-1.429.., -1.429..]	2.3e-08	1	7.4e-08	279	< 1
MGAff	[-1.429.., -1.429..]	2.3e-08	1	7.4e-08	234	1

Задача Colville4

Задача Colville4 (Iwaarden phd). Найти минимум функции на брусе:

$$f(x_1,x_2,x_3,x_4) = 100(x_2 - x_2^2)^2 + (1 - x_1)^2 + 90(x_4 - x_3^2)^2 + (1 - x_3)^2 + 10.1((x_2 - 1)^2 + (x_4 - 1)^2) + 19.8(x_2 - 1)(x_4 - 1),$$

$$-10 \leq x_i \leq 10.$$

Метод	OptEncl	wEncl	NBoxes	BoxWid	NBis	Time
Int	[-10.4.., 1.44..]	11.89	2000	5	4535	9
Aff	[-5.6.., 1.49..]	7.15	2000	5	4703	127
MGInt	[-3.8e-13, 5.1e-14]	4.4e-13	39	7.4e-08	3978	16
MGAff	[-1.03e-13, 5.1e-14]	1.5e-13	11	7.4e-08	1250	8

Задача Six Hump Camel Back

Задача Six Hump Camel Back. Найти минимум функции на брусе:

$$f(x_1, x_2) = 4x_1^2 - 2.1x_1^4 + x_1^6/3 + x_1x_2 - 4x_2^2 + 4x_2^4,$$

$$-5 \leq x_i \leq 5.$$

Метод	OptEncl	wEncl	NBoxes	BoxWid	NBis	Time
Int	[-1.041.., -1.031..]	0.0103	2000	0.156	3657	5
Aff	[-1.031.., -1.031..]	2.66e-14	20	7.4e-08	557	1
MGInt	[-1.031.., -1.031..]	6.3e-14	2	7.4e-08	355	1
MGAff	[-1.031.., -1.031..]	2.35e-14	2	7.4e-08	147	1

Модифицированный Six Hump Camel Back

Модифицированный Six Hump Camel Back. Найти минимум функции на брусе:

$$f(x_1, x_2) = 4x_1^2 - 2.1x_1^4 + x_1^6/3 + x_1x_2 - 4x_2^2 + 4x_2^4 + \\ - \frac{0.0001x_1}{x_1^2 + x_2^2 + 2x_1x_2\sin(x_1 - x_2) + 0.01},$$

$$-5 \leq x_i \leq 5.$$

Метод	OptEncl	wEncl	NBoxes	BoxWid	NBis	Time
Int	$[-\infty$, -0.988..]	∞				∞
Aff	[-1.031.., -1.031..]	2.5e-14	6	7.45e-08	662	6
MGInt	$[-\infty$, -0.988..]	∞				∞
MGAff	[-1.031.., -1.031..]	2.5e-14	1	7.45e-08	429	11

Заключение

Апостериорный метод построения внешних оценок для областей значений производных позволяет уменьшить трудоемкость вычислений на один порядок. Это особенно важно для задач высокой размерности. Для аффинного метода оценивания апостериорный подход также дает повышение точности оценок.