Применение интервального метода Ньютона и его модификации для решения задачи поиска глобального оптимума функций

М. А. Лядова
Новосибирский государственный университет
e-mail: tilvit-teg@rambler.ru

Н. В. Панов
Конструкторско-технологический институт вычислительной техники СО РАН

Интервальный метод Ньютона является одним из методов решения нелинейных уравнений. Он основан на последовательном применении сжимающего оператора Ньютона для уменьшения размера исходного интервала. Таким образом, интервальный метод Ньютона позволяет локализовать корни в том числе нелинейных уравнений.

Метод Ньютона – это итерационный численный метод решения нелинейных уравнений. Поиск решения осуществляется путём построения последовательных приближений. Метод обладает квадратичной сходимостью. Интервальный метод Ньютона позволяет локализовать корни, в том числе нелинейных уравнений. В задачах глобальной оптимизации это может быть использовано при распространении ограничений, а так же при прямом поиске нулей первой производной целевой функции. В точке экстремума производная должна обращаться в ноль, причем по каждой из координат. Таким образом, если на каком-либо подбрусе интервальное расширение первой производной целевой функции по какой-либо координате не содержит нуля (в многомерном случае требуется анализировать градиент функции), то, если это не граничный брус, то он гарантированно не содержит точек экстремума, а, следовательно, может быть удален из рассмотрения.

Рассмотрим одномерный случай для некоторого уравнения $f(x) = 0$. Можем разложить данную функцию в ряд Тейлора в окрестности некоторой точки x_0 $f(x) = f(x_0) + f'(x_0)(x - x_0) + O(x^2) = 0$, откуда следует $f(x) \cong f(x_0) + f'(x_0)(x - x_0) = 0$. При решении этого уравнения получаем лучшее приближение $x = \frac{x_0 - f(x_0)}{f'(x_0)}$, если $f'(x_0) \neq 0$. И так повторяем цикл и получаем следующий итерационный алгоритм:

$$\begin{cases} \text{Выбираем } x^{(0)} \\ \text{Итерируем:} \\ x^{(k+1)} \leftarrow x^{(k)} - \dfrac{f\left(x^{(k)}\right)}{f'\left(x^{(k)}\right)}, \quad k = 0, 1, 2, \dots \end{cases}$$

Получили одномерный метод Ньютона. Сходимость метода Ньютона зависит от выбора x_0. При неудачном начальном значении данный метод может, вообще говоря, и не сойтись.

Перейдём к интервальному случаю одномерного метода Ньютона.

Пусть $\mathbb{x}$ – интервал на нём рассматриваем уравнение $f(x) = 0$, где f - непрерывно дифференцируемая функция на $\mathbb{x}$. Если $\tilde{x}$ – некоторая точка из интервала $\mathbb{x}$, то в силу теоремы Лагранжа о среднем получаем:

$$(*)\ f(x) - f(\tilde{x}) = f'(\xi)(x - \tilde{x}), \text{где } \xi \in \mathbb{x}.$$

Если же $x = x^*$, где $f(x^*) = 0$, т.е. x^* - решение уравнения $f(x) = 0$, то путём простых преобразований из $(*)$ получаем $(**)\ x^* = \tilde{x} - \frac{f(\tilde{x})}{f'(\xi)}$.

Пусть $\mathbb{f}'$ - *интервальное расширение* f'. *Тогда, интервализуя по* $\xi \in \mathbb{x}$, *получим из* $(**)$ *следующее* $x^* \in \tilde{x} - \frac{f(\tilde{x})}{\mathbb{f}'(\boldsymbol{x})}$ *при условии, что данное выражение определено.*

Отображение $\mathbb{R} \times \mathbb{IR} \to \mathbb{IR}$, *определяемое как* $\mathcal{N}(\tilde{x},\ \mathbb{x}) = \tilde{x} - \frac{f(\tilde{x})}{\mathbb{f}'(\mathbb{x})}$ *для заданной функции* $f\colon \mathbb{R} \to \mathbb{R}$ *с интервальным расширением её первой производной* $\mathbb{f}'$, *есть (одномерный) интервальный оператор Ньютона.*

То есть, если заданы $f, \mathbb{x}$, то $\mathcal{N}(\tilde{x},\ \mathbb{x})$ определяет новый интервал, содержащий решение x^* уравнения $f(x) = 0$.

Следовательно, более точные границы для x^* можно взять в виде $\mathbb{x} \cap \mathcal{N}(\tilde{x},\ \mathbb{x})$.

Далее снова начинаем процесс. В итоге, можем определить итерации:

$$\begin{cases} \mathbb{x}^{(0)} \leftarrow \mathbb{x} \\ \mathbb{x}^{(k+1)} \leftarrow \mathbb{x}^{(k)} \cap \mathcal{N}\left(\tilde{x},\ \mathbb{x}^{(k)}\right) \quad k = 0, 1, 2, \ldots \end{cases}$$

Получили простейший одномерный случай интервального метода Ньютона. Часто определяют $\tilde{x}$ как $mid\mathbb{x}^{(k)}$, где $mid\mathbb{x}$ - это середина интервала $\mathbb{x}$.

Работая с этим методом, мы можем получить интервал, содержащий ноль. В этом случае в операторе Ньютона возникает деление на ноль, в результате чего мы получаем бесконечный интервал. К примеру, пусть у нас есть интервал $\mathbb{x} = [-a, a]$. Тогда $\frac{1}{\mathbb{x}} = \left[-\infty, -\frac{1}{a}\right] \cup \left[\frac{1}{a}, +\infty\right]$. Прибавляя сюда интервал, мы можем вообще получить всю числовую ось. В частности, чтобы избежать этого, мы и пересекаем новый интервал с предыдущим.

Реальные задачи в большинстве своём являются многомерными, и для их решения необходим многомерный случай метода Ньютона.

Рассмотрим систему уравнений $F(x) = 0$, где $F(x) = \big(F_1(x),\ F_2(x),\ \dots,\ F_n(x)\big)^T$. Нужно найти решение этой системы. Исходя из одномерного случая, получим следующее уравнение для некоторой точки x_0 из области определения функции:

$$F'(x_0)(x - x_0) = -F(x_0), \text{ где } F'(x_0) \text{ - матрица Якоби.}$$

Если $F'(x_0)$ не вырождена, то система решается. Путём простых преобразований получим:

$$x = x_0 - (F'(x_0))^{-1} F(x_0).$$

Получили лучшее приближение. Далее повторяем процесс:

$$\begin{cases} \text{Берём } x^{(0)} \\ \text{Итерируем:} \\ x^{(k+1)} = x^{(k)} - \left(F'\left(x^{(k)}\right)\right)^{-1} F\left(x^{(k)}\right), \ \ k = 0, 1, 2, \dots \end{cases}$$

Получили многомерный случай метода Ньютона.

Теперь мы можем рассмотреть интервальный случай многомерного метода Ньютона.

Пусть $\mathbb{x} \in \mathbb{IR}^n$ – многомерный интервал, на котором ищем решение, и пусть $x, \tilde{x} \in \mathbb{x}$. Аналогично одномерному интервальному случаю получим для многомерной функции $F(x)$, определённой на интервале $\mathbb{x}$:

$$F(x) - F(\tilde{x}) = F'(\xi)(x - \tilde{x}), \text{ где } \xi \in (x, \tilde{x}) \text{ в } \mathbb{R}^n.$$

Возьмём $x = x^*$, где x^* - решение системы, т. е. $F(x^*) = 0$. И пусть известно интервальное расширение $\mathbb{F}'(\mathbb{x})$ для $F'(\xi)$ на интервале $\mathbb{x}$. Тогда получим:

$$0 \in F(\tilde{x}) + \mathbb{F}'(\mathbb{x})(x^* - \tilde{x}).$$

Прибегнем к помощи одного известного факта из теории интервальных линейных систем, а именно характеризации Бека:

Пусть $x \in \Xi(\mathbb{A}, \mathbb{b})$ *- множество решений системы линейных алгебраических уравнений (ИСЛАУ)* $\mathbb{A}y = \mathbb{b}$. *Тогда* $\mathbb{A}x \cap \mathbb{b} \neq 0 \iff 0 \in \mathbb{A}x - \mathbb{b}$.

Отсюда следует, что $x^* - \tilde{x}$ принадлежит множеству решений интервальной системы

$$\mathbb{F}'(\mathbb{x}) \cdot y = -F(\tilde{x}).$$

А это может быть тогда и только тогда, когда $x^* \in \tilde{x} + \big(\text{множество решений ИСЛАУ } \mathbb{F}'(\mathbb{x}) \cdot y = -F(\tilde{x})\big)$.

И, следовательно,
$x^* \in \tilde{x} + $ (внешняя интервальная оценка решений ИСЛАУ $\mathbb{F}'(\mathbb{x}) \cdot y = -F(\tilde{x})$).
Или, другими словами, $x^* \in \tilde{x} + Encl\left(\Xi\left(\mathbb{F}'(\mathbb{x}), -F(\tilde{x})\right)\right)$.

В этом случае оператор Ньютона для многомерной функции $F(x)$ и интервального расширения её первой производной $\mathbb{F}'(\mathbb{x})$ будет выглядеть следующим образом:

$$\mathcal{N}(\tilde{x}, \mathbb{x}) = \tilde{x} + Encl\left(\Xi\left(\mathbb{F}'(\mathbb{x}), -F(\tilde{x})\right)\right).$$

Здесь, как и в одномерном случае интервального метода Ньютона, $\mathcal{N}(\tilde{x}, \mathbb{x})$ определяет новый интервал, содержащий решение x^* уравнения $F(x) = 0$. Точно так же более точные границы для x^* можно взять в виде $\mathbb{x} \cap \mathcal{N}(\tilde{x}, \mathbb{x})$. И, повторяя процесс итерирования, получим схожий алгоритм для многомерного случая интервального метода Ньютона:

$$\begin{cases} \mathbb{x}^{(0)} \leftarrow \mathbb{x} \\ \mathbb{x}^{(k+1)} \leftarrow \mathbb{x}^{(k)} \cap \mathcal{N}\left(\tilde{x}, \mathbb{x}^{(k)}\right) \quad k = 0, 1, 2, \dots \end{cases}$$

Так же можем брать $\tilde{x} = mid\mathbb{x}^{(k)}$. Здесь $\mathbb{x}$ - это интервальная матрица, а $mid\mathbb{x}$ - скалярная матрица, состоящая из середин компонент-интервалов матрицы $\mathbb{x}$. Важно помнить, что у нас теперь $\mathbb{x}$ - многомерный интервал и оператор Ньютона определяется по-новому.

В случае задачи оптимизации, когда нужно найти, например, минимум функции на заданном интервале, нужно использовать градиент этой функции и рассматривать уже систему уравнений $\nabla F(x) = 0$.

Многомерный случай интервального метода Ньютона имеет свои недостатки. В ходе работы с этим методом постоянно нужно решать задачу внешнего оценивания множества решений интервальных систем алгебраических уравнений, а эта задача, вообще говоря, является NP-трудной. Также интервальная линейная система, которую необходимо решать при реализации интервального метода Ньютона, может быть вырожденной, а тогда её множество решений будет неограничено.

Существуют различные альтернативные методы, преодолевающие эти недостатки, основанные на методе Ньютона, такие как метод Кравчика, метод Хансена-Сенгупты и другие.

Описанные методы и их модификации разрабатываются для использования в программе поиска глобального оптимума функций для нахождения нулей градиента. Что, как было указано выше, можно использовать для повышения эффективности решения задачи глобальной оптимизации.

Список литературы

[1] Шарый С.П. Конечномерный интервальный анализ. Электронная книга

http://www-sbras.nsc.ru/interval/Library/InteBooks/ SharyBook.pdf/