Лекция 4

Тема 4. Методы решения нелинейных уравнений

Рассмотрим задачу отыскания корней нелинейных уравнений и основные методы ее решения. Решение нелинейных уравнений представляет собой редкий пример задачи, которая может быть сравнительно полно исследована элементарными средствами. В то же время многие проблемы, возникающие при отыскании корней нелинейных уравнений, типичны, а некоторые методы их решения (в особенности метод Ньютона и метод простой итерации) допускают широкие обобщения и играют в вычислительной математике фундаментальную роль.

§ 4.1. Постановка задачи. Основные этапы решения

Задача отыскания корней нелинейного уравнения с одним неизвестным вида

$$f(x) = 0 \tag{4.1}$$

имеет многовековую историю, но не потеряла свою актуальность и в наши дни, так как она часто возникает как элементарный шаг при решении различных научных и технических проблем.

Напомним, что *корнем* (или *решением*) уравнения (4.1) называется значение $\bar{x}$, при котором $f(\bar{x}) = 0$.

Будем далее предполагать, что в окрестности каждого из искомых корней функция $f(x)$ дважды непрерывно дифференцируема.

Корень $\bar{x}$ уравнения (4.1) называется *простым*, если $f'(\bar{x}) \neq 0$. В противном случае (т.е. в случае $f'(\bar{x}) = 0$) корень $\bar{x}$ называется *кратным*. Число m назовем кратностью корня $\bar{x}$, если $f^{(k)}(\bar{x}) = 0$ для $k = 1, 2, \ldots, m\text{-}1$ и $f^{(m)}(\bar{x}) \neq 0$.

Геометрически корень $\bar{x}$ соответствует точке пересечения графика функции $y = f(x)$ с осью Ox. Корень $\bar{x}$ является простым, если график пересекает ось Ox под ненулевым углом, и кратным, если пересечение происходит под нулевым углом. Задача отыскания простых корней является значительно более простой (и чаще встречающейся), чем задача отыскания кратных корней. Большинство методов решения уравнения (4.1) ориентировано именно на вычисление простых корней.

В подавляющем большинстве случаев представить решение уравнения (4.1) в виде конечной формулы оказывается невозможным. Даже для простейшего алгебраического уравнения n-й степени

$$a_n x^n + a_{n-1} x^{n-1} + \ldots + a_1 x + a_0 = 0 \qquad (4.2)$$

явные формулы найдены лишь при n = 2, 3, 4. Однако уже для уравнений пятой и более высоких степеней таких формул не существует (теорема Абеля). При этом нужно иметь в виду, что даже точное решение, если оно существует, оказывается выраженным достаточно громоздкой формулой, вычисление по которой все равно дает приближенное значение, так как содержит вычислительную погрешность. Поэтому в дальнейшем мы откажемся от попыток найти точное решение и сосредоточимся на методах приближенного вычисления корней с заданной точностью ε.

Решение задачи вычисления корней нелинейного уравнения, как правило, осуществляется в два этапа. Первый этап называется *этапом локализации (*или *отделения)* корней, второй – *этапом итерационного уточнения корней.*

Л о к а л и з а ц и я к о р н е й. Отрезок [a, b], содержащий только один корень $\bar{x}$ уравнения (4.1), называют отрезком *локализации* корня $\bar{x}$. Цель этапа локализации считается достигнутой, если для каждого из подлежащих локализации корней удалось указать отрезок локализации (его длину стараются по возможности сделать минимальной).

Способы локализации корней многообразны, и указать универсальный метод не представляется возможным. Иногда отрезок локализации известен заранее или он определяется из физических соображений. В простых случаях хороший результат может дать графический метод. Широко применяют также построение таблиц значений функции f вида $y_i = f(x_i)$, $i = 1, 2, \ldots, n$. При этом способе локализации о наличии на отрезке $[x_{i-1}, x_i]$ корня судят по перемене знака функции на концах отрезка. Основанием для применения указанного способа служит известная теорема математического анализа:

Т е о р е м а 4.1. *Пусть функция f непрерывна на отрезке $[a, b]$ и принимает на его концах значения разных знаков, т.е. $f(a) \cdot f(b) < 0$. Тогда отрезок $[a, b]$ содержит по крайней мере один корень уравнения $f(x) = 0$.*

И т е р а ц и о н н о е у т о ч н е н и е к о р н е й. На этом этапе решения задачи для вычисления каждого из корней с точностью ε > 0 используют тот или иной итерационный метод, позволяющий построить последовательность $x^{(0)}, x^{(1)}, \dots, x^{(n)}, \dots$ приближений к корню $\bar{x}$.

Общие представления об итерационных методах и основные определения были даны в разделе 3. Введем дополнительно некоторые определения.

Итерационный метод называется *одношаговым*, если для вычисления очередного приближения $x^{(n+1)}$ используется только одно предыдущее приближение $x^{(n)}$ и *k-шаговым*, если для вычисления $x^{(n+1)}$ используются k предыдущих приближений $x^{(n-k+1)}$, $x^{(n-k+2)}, \dots, x^{(n)}$. Заметим, что для построения итерационной последовательности одношаговым методом требуется задание только одного начального приближения, в то время как при использовании k-шагового метода – k начальных приближений $x^{(0)}, x^{(1)}, \dots, x^{(k-1)}$.

Скорость сходимости – одна из важнейших характеристик итерационных методов. Говорят, что *метод сходится со скоростью геометрической прогрессии,* знаменатель которой $q < 1$, если для всех n справедлива следующая оценка:

$$|x^{(n)} - \bar{x}| \le c_0 q^n \tag{4.5}$$

Как нетрудно видеть, из оценки (4.5) действительно вытекает сходимость метода.

Пусть одношаговый итерационный метод обладает следующим свойством: существует б-окрестность корня $\bar{x}$ такая, что если приближение $x^{(n)}$ принадлежит этой окрестности, то справедлива оценка

$$|x^{(n+1)} - \bar{x}| \le C\, |x^{(n)} - \bar{x}|^{\,p}, \tag{4.6}$$

где C > 0 и $p \ge 1$ – постоянные. Если $p = 1$ и C < 1, то говорят, что метод обладает *линейной скоростью сходимости* в указанной б-окрестности корня. Если *$p > 1$,* то принято говорить о *сверхлинейной скорости сходимости.* При $p = 2$ скорость сходимости называют *квадратичной,* а при $p = 3$ – *кубической*.

Л е м м а 4.1. *Пусть одношаговый итерационный метод обладает линейной скоростью сходимости в некоторой б-окрестности корня $\bar{x}$. Тогда при любом выборе начального приближения $x^{(0)}$ из* *σ-окрестности корня $\bar{x}$ итерационная последовательность $x^{(n)}$ не выходит за пределы этой окрестности, метод сходится со скоростью геометрической прогрессии со знаменателем q = C и имеет место следующая оценка погрешности:*

$$|x^{(n)} - \bar{x}| \leq q^n \mid x^{(0)} - \bar{x} \mid, \quad n \geq 0. \tag{4.7}$$

□ Заметим, что принадлежность $x^{(n)}$ окрестности ($\bar{x} - \sigma$, $\bar{x} + \sigma$) является следствием неравенства (4.7). В самом деле, так как $q < 1$, то $|x^{(n)} - \bar{x}| \leq \mid x^{(0)} - \bar{x} \mid < \sigma$. Сходимость $x^{(n)}$ к $\bar{x}$ также вытекает из (4.7).

Справедливость (4.7) установим методом индукции. При $n = 0$ оно переходит в очевидное: $\mid x^{(0)} - \bar{x} \mid \leq \mid x^{(0)} - \bar{x} \mid$. Пусть неравенство (4.7) выполнено при $n = m - 1$. Тогда

$$\mid x^{(m)} - \bar{x} \mid \leq q \mid x^{m-1)} - \bar{x} \mid \leq q^m \mid x^{(0)} - \bar{x} \mid,$$

т.е. неравенство выполнено при $n = m$. ■

Можно доказать аналогичное утверждение и для многошагового итерационного метода. Мы этого делать не будем.

С помощью указанной леммы исследование сходимости итерационных методов сводится только к получению оценки (4.6).

§ 4.2. Обусловленность задачи вычисления корня

Пусть $\bar{x}$ – корень уравнения (4.1), подлежащий определению. Будем считать, что входными данными для задачи вычисления корня $\bar{x}$ являются значения функции *f(x)* в малой окрестности корня. Так как значения *f(x)* будут вычисляться на ЭВМ по некоторой программе, то в действительности задаваемые значения являются приближенными и их следует обозначать через $f^*(x)$. Ошибки в $f^*(x)$ могут быть связаны не только с неизбежными ошибками округления, но и с использованием приближенных методов. К сожалению, нельзя ожидать, что в окрестности корня относительная погрешность $\delta(f^*)$ окажется малой. Достаточно обратиться к примеру 3.7, чтобы установить, что для достаточно простой функции $y = \sin x$ в окрестности корней $x = \pi \cdot к$, *к = 0, ±1, ±2, …* значения этой функции не могут быть найдены с малой относительной погрешностью. Реально рассчитывать можно лишь на то, что малой окажется абсолютная погрешность вычисления значений функции.

Будем предполагать, что в достаточно малой окрестности корня выполняется неравенство $| f(x) - f^*(x)| < \bar{\Delta}$, где $\bar{\Delta} = \bar{\Delta}(f^*)$ – граница абсолютной погрешности.

Если функция f непрерывна, то найдется такая малая окрестность $(\bar{x} - \bar{\varepsilon}, \bar{x} + \bar{\varepsilon})$ корня $\bar{x}$, имеющая радиус $\bar{\varepsilon} > 0$, в которой выполняется неравенство

$$| f(x) | < \bar{\Delta}. \tag{4.9}$$

Для $x \in (\bar{x} - \bar{\varepsilon}, \bar{x} + \bar{\varepsilon})$ знак вычисленного значения $f^*(x)$, вообще говоря, не обязан совпадать со знаком $f(x)$ и, следовательно, становится невозможным определить, какое именно значение x из интервала $(\bar{x} - \bar{\varepsilon}, \bar{x} + \bar{\varepsilon})$ обращает функцию $f(x)$ в нуль. Будем называть этот интервал интервалом неопределенности корня $\bar{x}$.

Найдем оценку величины $\bar{\varepsilon}$. Пусть корень $\bar{x}$ – простой. Для близких к $\bar{x}$ значений x справедливо приближенное равенство

$$f(x) \approx f(\bar{x}) + f'(\bar{x})(x - \bar{x}) = f'(\bar{x})(x - \bar{x})$$

Поэтому неравенство (4.9) примет вид $|f'(\bar{x})(x - \bar{x})| < \bar{\Delta}$, откуда получаем

$$\bar{x} - \frac{\bar{\Delta}}{|f'(\bar{x})|} < x < \bar{x} + \frac{\bar{\Delta}}{|f'(\bar{x})|}.$$

Следовательно,

$$\bar{\varepsilon} \approx \nu_\Delta \, \bar{\Delta}(f^*), \tag{4.10}$$

здесь $\nu_\Delta = \frac{1}{|f'(\bar{x})|}$ – число, которое в рассматриваемой задаче играет роль абсолютного числа обусловленности. Действительно, если $\bar{x}^*$ – корень уравнения $f^*(x) = 0$, то $|f(\bar{x}^*)| < \bar{\Delta}$ и тогда выполнено неравенство

$$| \bar{x} - \bar{x}^*| \le \bar{\Delta}(\bar{x}^*) \le \bar{\varepsilon} \approx \nu_\Delta \, \bar{\Delta}(f^*). \tag{4.11}$$

Заметим, что радиус интервала неопределенности прямо пропорционален погрешности $\bar{\Delta}$ вычисления значения f. Кроме того, $\bar{\varepsilon}$ возрастает (обусловленность задачи ухудшается) с уменьшением

$|f'(\bar{x})|$, т.е. с уменьшением модуля тангенса угла наклона, под которым график функции пересекает ось *Ox*.

Если же $f'(\bar{x}) = 0$ (т.е. корень $\bar{x}$ – кратный), то формула (4.10) уже не верна. Пусть кратность корня равна m. Тогда в силу формулы Тейлора справедливо приближенное равенство

$$f(x) \approx f(\bar{x}) + f'(\bar{x})(x - \bar{x}) + \frac{f''(\bar{x})}{2!}(x - \bar{x})^2 + \ldots + \frac{f^{(m)}(\bar{x})}{m!}(x - \bar{x})^m,$$

в правой части которого все слагаемые, кроме последнего, равны нулю. Следовательно, неравенство (4.9) имеет вид

$$\left| \frac{f^{(m)}(\bar{x})}{m!}(x - \bar{x})^m \right| < \bar{\Delta}.$$

Решая его, получаем аналогично (4.10) оценку радиуса интервала неопределенности:

$$\bar{\varepsilon} \approx \left| \frac{m!}{f^{(m)}(\bar{x})} \right|^{1/m} \bar{\Delta}^{1/m}.$$

Эта оценка означает, что для корня кратности m радиус интервала неопределенности пропорционален $\bar{\Delta}^{1/m}$, что свидетельствует о плохой обусловленности задачи вычисления кратных корней.

В реальных условиях оценить величину и даже порядок радиуса интервала неопределенности довольно сложно. Однако знать о его существовании нужно, по крайней мере, по двум причинам. Во-первых, не имеет смысла ставить задачу о вычислении корня $\bar{x}$ с точностью $\varepsilon < \bar{\varepsilon}$. В условиях неопределенности, вызванных приближенным заданием функции, любое значение $\bar{x}^* \in (\bar{x} - \bar{\varepsilon}, \bar{x} + \bar{\varepsilon})$ может быть с одной и той же степенью достоверности принято за решение уравнения. Во-вторых, нельзя требовать от алгоритмов отыскания корня получения достоверных результатов после того, как очередное приближение попало в интервал неопределенности или оказалось очень близко от него; в этой ситуации вычисления следует прекратить и считать, что получен максимум действительно возможного.

Для большинства итерационных методов определить этот момент можно, поскольку, начиная с него поведение приближений $x^{(n)}$

становится крайне нерегулярным. Если вдали от интервала неопределенности величина

$$q^{(n)} = |\, x^{(n)} - x^{(n-1)}| \,/\, |\, x^{(n-1)} - x^{(n-2)}| \tag{4.12}$$

обычно бывает меньше единицы (т.е. $|\, x^{(n)} - x^{(n-1)}| < |\, x^{(n-1)} - x^{(n-2)}|$), то появление при некотором при некотором n значений $q^{(n)} > 1$ свидетельствует о начале «разболтки» $-$ хаотического поведения итерационной последовательности. В этой ситуации вычисления следует прекратить и принять правильное решение. Лучшим из последовательностей приближений к решению следует считать, конечно, $x^{(n-1)}$. Использование для контроля вычислений величины (4.12) называют *правилом Гарвика*.