ЛЕКЦИЯ 12.1 ЗАДАЧА ЧИСЛЕННОГО РЕШЕНИЯ НЕЛИНЕЙНОГО УРАВНЕНИЯ. ЛОКАЛИЗАЦИЯ КОРНЯ. МЕТОДЫ ПОЛОВИННОГО ДЕЛЕНИЯ И ПРОСТОЙ ИТЕРАЦИИ

1. Нелинейные уравнения. Постановка задачи

Решение уравнений – это важная прикладная задача. Практически во всех инженерных, научных расчётах приходится решать уравнения или их системы. Для некоторых видов уравнений известны формулы точных корней (квадратные, кубические, некоторые тригонометрические и др.). Но, во-первых, они охватывают весьма узкий круг уравнений, в то время как в вычислительной практике приходится решать самые разнообразные уравнения. Во-вторых, формулы точного решения могут быть громоздки, а потому трудны для практического применения. Поэтому возникает задача численного, или приближённого, решения уравнений и систем. В этой лекции мы изучим методы решения нелинейных уравнений.

Уравнение

$$f(x) = 0 \tag{1}$$

называется *нелинейным*, если функция f нелинейна. *Корнем* уравнения называется такое число, что его подстановка вместо переменной обращает уравнение в тождество. Точное (теоретическое) значение корня будем обозначать так же, как переменную, т.е. x. Задача заключается в нахождении приближённого значения корня x^*, т.е. такого числа x^*, что

$$f(x^*) \approx 0.$$

Это приближённое равенство означает, что

$$|f(x^*)| < \varepsilon, \tag{2}$$

где $\varepsilon > 0$ - точность решения.

2. Локализация корней

Все методы численного решения нелинейных уравнений - итерационные (прямые методы разрабатываются в фундаментальной математике). Они состоят из двух этапов.

Первый, предварительный – это локализация (отделение) корня. Второй, основной, – это итерационное уточнение корня, т.е. собственно решение уравнения.

Локализация корня – это обязательный этап решения. Более того, он очень важен: от него во многом зависит успех решения, т.е. сходимость итерационного процесса и её скорость. Локализация, или отделение, корня - это определение промежутка числовой оси, содержащего ровно один корень.

Локализация осуществляется исследованием функции f. Для этого применяются самые разнообразные методы. Они сильно зависят от уравнения, поэтому здесь невозможно дать общий универсальный алгоритм. Но есть некоторые общие рекомендации.

Первая – использовать график. Можно построить график (или эскиз графика) функции f и по нему определить промежутки, на которых он пересекает ось OX. Абсциссы точек пересечения – это и есть корни уравнения $f(x) = 0$.

Вторая – преобразовать уравнение (1). Например, можно привести его к виду

$$\varphi(x) = \phi(x),$$

функции φ и ϕ выбираются так, чтобы это уравнение было проще для исследования, чем исходное. Положение корня можно определить опять же графически, построив графики функций $y = \varphi(x)$, $y = \phi(x)$. Корень уравнения - абсцисса точки пересечения этих графиков. Иногда полезно привести уравнение к виду

$$x = \varphi(x).$$

Тогда корень ищется как абсцисса точки пересечения графика функции $y = \varphi(x)$ с прямой $y = x$.

Наконец, можно построить таблицу значений функции и по ней определить промежутки, на которых функция f меняет знак. Если f непрерывна, то на таком промежутке будет хотя бы один корень. Понятно, что для лучшей локализации надо, чтобы узлы таблицы шли с небольшим шагом, а также следует применять другие методы исследования функций. В частности, определить промежутки монотонности функции.

При исследовании свойств уравнения можно применять следующие теоремы из курса математического анализа.

Теорема 1 (Больцано-Коши). Если непрерывная на отрезке $[a; b]$ функция f имеет на его концах противоположные знаки, т.е. $f(a)f(b) < 0$, то на интервале $(a; b)$ она хотя бы один раз обращается в нуль.

Теорема 2. Непрерывная строго монотонная функция имеет на отрезке единственный нуль тогда и только тогда, когда на его концах она принимает значения разных знаков.

Также при локализации применяются все методы исследования функции из математического анализа.

Пример. Рассмотрим уравнение

$$x \operatorname{tg} \frac{x}{3} - x - 1 = 0.$$

Функция непрерывна на интервале $\left(-\frac{3}{2}\pi; \frac{3}{2}\pi\right)$. Рассмотрим значения функции в нескольких точках:

$$f(-\pi) = \pi \operatorname{tg} \frac{\pi}{3} + \pi - 1 > 0,$$

$$f(0) = -1 < 0,$$

$$f(\pi) = \pi \operatorname{tg} \frac{\pi}{3} - \pi - 1 > 0.$$

Непрерывная на отрезке $[-\pi; \pi]$ функция f дважды меняет на нём знак. Значит, на этом отрезке функция имеет по крайней мере два нуля, и уравнение имеет хотя бы два корня. Докажем, что на отрезках $[-\pi; 0]$ и $[0; \pi]$ имеется ровно по одному корню.

Первая производная функции f

$$f'(x) = \operatorname{tg} \frac{x}{3} + \frac{x}{3\cos^2 \frac{x}{3}} - 1$$

на отрезке $[-\pi; \pi]$ монотонна. В этом можно убедиться с помощью второй производной

$$f''(x) = \left(1 + \frac{1}{3} \operatorname{tg} \frac{x}{3}\right) \frac{2}{3\cos^2 \frac{x}{3}},$$

которая положительна на всем рассматриваемом отрезке. Первая производная меняет знак на отрезке $[-\pi; \pi]$ (в этом можно убедиться, подставив $-\pi$, π в формулу производной). Как следствие, первая производная имеет одну точку экстремума на отрезке $\left[0; \frac{\pi}{2}\right]$, причем, это точка минимума функции: $f'(0) = -1$, $f'\left(\frac{\pi}{2}\right) > 0$.

Можно сделать следующие выводы: функция непрерывна, на концах отрезка $[-\pi;\ \pi]$ её значения положительны; имеется одна точка минимума, в которой значение функции отрицательно. Значит, на отрезке $[-\pi;\ \pi]$ имеется две точки, в которых значения функции равны нулю. Одна точка принадлежит отрезку $[-\pi; 0]$, другая точка отрезку $[0;\ \pi]$. Таким образом, удалось локализовать два корня исходного уравнения.

3. Метод половинного деления

После локализации корня вступает в действие его итерационное уточнение: выбирается начальная итерация $x^{(0)}$ из промежутка локализации и строится последовательность $x^{(0)}, x^{(1)}, x^{(2)}, \dots$, сходящаяся к точному решению x. Понятно, что вычисление продолжается не бесконечно, а останавливается при достижении заданной точности. Поэтому для полноты метода надо определить:

1. Расчётную формулу для произвольной k-й итерации;
2. Оценку погрешности итерации (для критерия останова);
3. Условия сходимости (понятно, что интервал локализации сам по себе не обеспечивает сходимости).

Всё перечисленное зависит от конкретного метода. Начнём с самого простого – метода половинного деления (бисекций).

Пусть $(a; b)$ – интервал локализации; функция f непрерывна на отрезке $[a; b]$ и на его концах принимает значения разных знаков:

$$f(a)f(b) < 0.$$

Делим отрезок $[a; b]$ пополам, c - его середина (рис. 1):

$$c = \frac{a+b}{2}. \tag{3}$$

Проверяем условие

$$f(c) = 0,$$

т.е. является ли c корнем уравнения. Если да, то решение найдено: c – корень. Если нет, то определяем, на каком из отрезков теперь корень, на $[a; c]$ или $[c; b]$. Для этого достаточно проверить знак произведения $f(a)f(c)$. Если оно отрицательно, то корень на отрезке $[a; c]$, если положительно, то корень на отрезке $[c; b]$. Отрезок, содержащий корень, уменьшился вдвое.

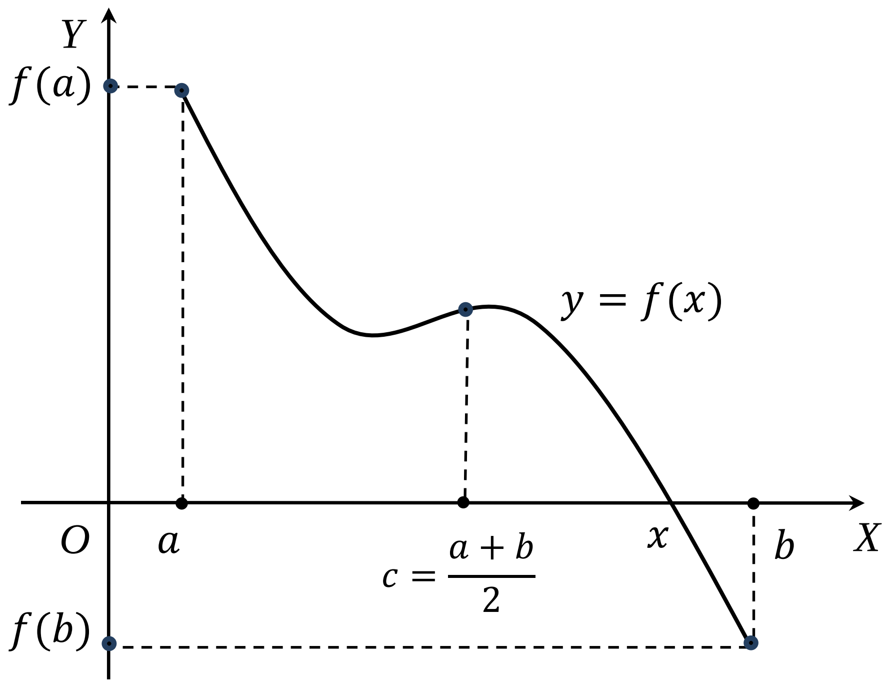

Рис. 1. Метод половинного деления

Далее в качестве нового отрезка $[a; b]$ берём тот, на котором находится корень:

$$[a; b] = \begin{cases} [a; c], \text{если } f(a)f(c) < 0, \\ [c; b], \text{если } f(a)f(c) > 0. \end{cases}$$

На новом отрезке $[a; b]$ повторяется та же процедура деления пополам, и так далее. Понятно, что выполнение точного равенства (3) маловероятно, а при приближённом вычислении f оно практически невозможно, поэтому на самом деле проверяют выполнение этого равенство с заданной точностью ε (см. (2)).

На каждом шаге длина отрезка, содержащего корень, уменьшается вдвое; и можно локализовать корень с какой угодно точностью. Процесс заканчивается при условии $b - a < \delta$ или при условии $|f(c)| < \varepsilon$, где δ, ε – заданные числа, связанные с погрешностью уточнения корня. Значение c выдаётся как приближение корня x^* (достигнуто условие (2)). Число δ будет при этом точностью корня, так как очевидно, что после завершения алгоритма

$$\Delta x^* = |x - x^*| = |x - c| < \delta.$$

На рис. 2 изображена блок-схема алгоритма.

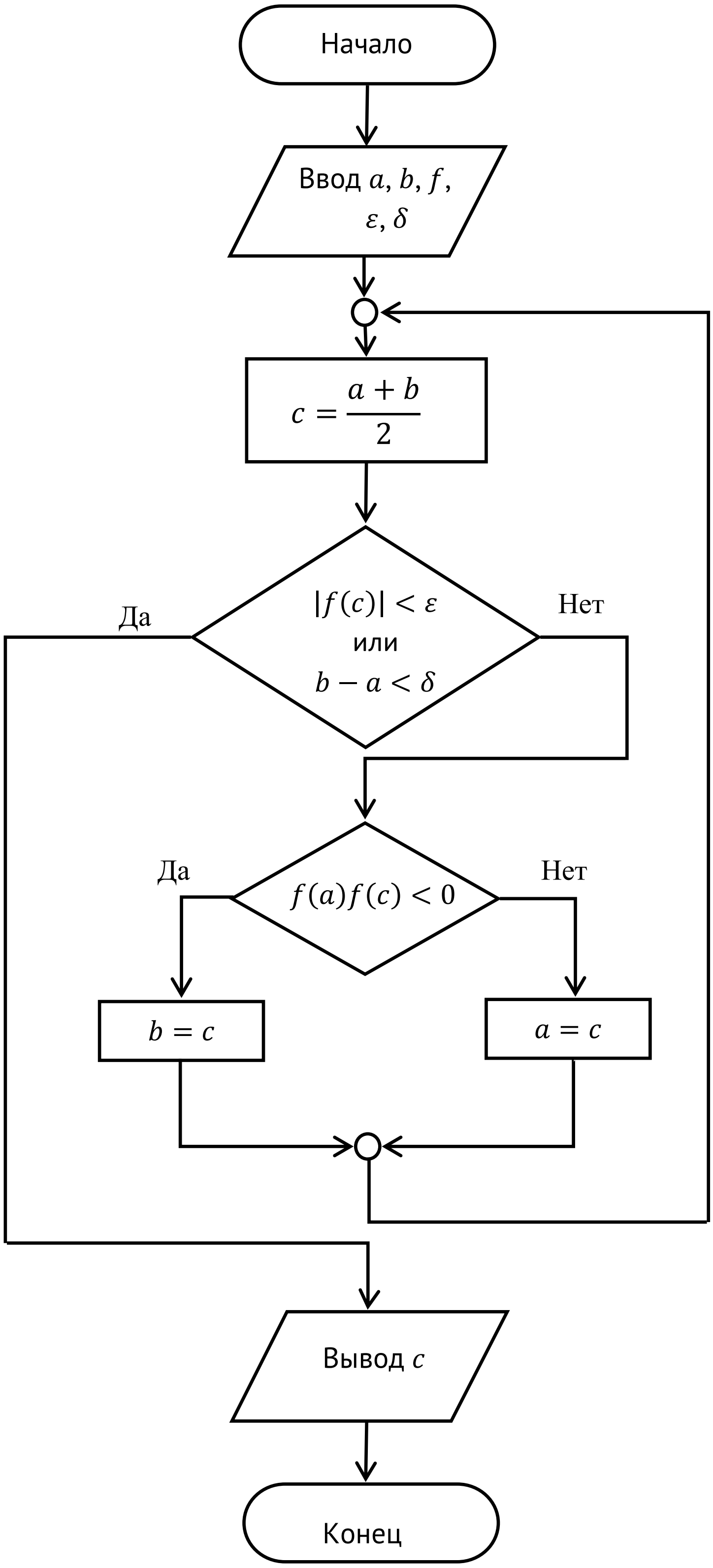

Рис. 2. Блок-схема алгоритма половинного деления

Если задаться точностью локализации корня δ, то можно определить число шагов, за которое она будет достигнута. Через k шагов деления отрезка пополам его длина станет равной

$$\frac{b-a}{2^k}.$$

Тогда для достижения точности корня δ достаточно выполнения неравенства

$$\frac{b-a}{2^k} < \delta,$$

из которого следует

$$2^k > \frac{b-a}{\delta} \Rightarrow k > \log_2 \frac{b-a}{\delta}.$$

Следовательно, после выполнения

$$k = \left[\log_2 \frac{b-a}{\delta}\right] + 1$$

делений длина промежутка локализации корня гарантированно станет меньше δ.

Очевидно, что рано или поздно нужная локализация будет достигнута при любой начальной итерации, поэтому метод деления отрезка пополам относится к гарантированно сходящимся. Но он имеет очевидный и существенный недостаток: малая скорость сходимости. На каждом шаге погрешность корня уменьшается всего в два раза, т.е. метод половинного деления имеет линейную скорость сходимости и сходится со скоростью геометрической прогрессии со знаменателем $^1/_2$.

Пример. Пусть надо решить уравнение $\sin x = 0$. Для него известны точные корни. Но на его примере можно рассмотреть приближенные методы.

Отправляясь от начального отрезка $[-0{,}5; 1{,}5]$, локализуем делением отрезка известный корень $x = 0$. В таблице 1 приведены результаты 4 шагов. Число c может приблизиться к корню, а на следующем шаге отдалиться, но длина отрезка, содержащего корень, с каждым шагом уменьшается.

Табл. 1. Итерации метода половинного деления в примере на с. 7

Шаг	1	2	3	3
a	$-0{,}5$	$-0{,}5$	$-0{,}083$	$-0{,}083$
b	$1{,}17$	$0{,}335$	$0{,}335$	$0{,}126$
c	$0{,}335$	$-0{,}083$	$0{,}126$	$0{,}022$
$f(a)$	$-0{,}479$	$-0{,}479$	$-0{,}082$	$-0{,}082$
$f(b)$	$0{,}921$	$0{,}329$	$0{,}329$	$0{,}126$
$f(c)$	$0{,}329$	$-0{,}082$	$0{,}126$	$0{,}022$
$b-a$	$1{,}67$	$0{,}835$	$0{,}418$	$0{,}209$

4. Метод простой итерации

4.1. Алгоритм метода

Метод простой итерации - это обобщение метода простой итерации для линейных систем на нелинейный случай. Исходное уравнение (1) приводится к эквивалентному уравнению

$$x = \varphi(x). \tag{4}$$

Этот переход можно сделать многими способами, и он важен, так как от функции φ зависит сходимость итерационной последовательности. Например, можно поступить совсем просто: умножить (1) на число $-c$ ($c \neq 0$) и прибавить к обеим частям x. Получим уравнение

$$x = x - cf(x). \tag{5}$$

Итерационная последовательность строится так же, как и в линейном случае. Берётся начальная итерация $x^{(0)}$ из промежутка локализации, подставляется в правую часть (4), полученное значение принимается за следующую итерацию $x^{(1)}$; $x^{(1)}$ опять подставляется в правую часть и так далее. Расчётная формула выглядит так:

$$x^{(k)} = \varphi\left(x^{(k-1)}\right), \tag{6}$$

$k = 1, 2, \dots$. Если итерационная последовательность $\{x^{(k)}\}$ имеет предел $\tilde{x}$ и функция φ непрерывна на промежутке локализации, то этот предел является корнем уравнения. Это доказывается предельным переходом в (6):

$$x^{(k)} = \varphi\big(x^{(k-1)}\big) \Rightarrow \lim_{k\to\infty} x^{(k)} = \lim_{k\to\infty} \varphi\big(x^{(k-1)}\big) = \varphi\left(\lim_{k\to\infty} x^{(k-1)}\right) \Rightarrow \tilde{x} = \varphi(\tilde{x}).$$

На рис. 3 показана графическая иллюстрация метода простой итерации. Корень уравнения – абсцисса точки пересечения прямой $y = x$ и кривой $y = \varphi(x)$. Для начальной точки $x^{(0)}$ находится точка $\left(x^{(0)}, \varphi\big(x^{(0)}\big)\right)$. Через неё проводится прямая параллельно оси OX до пересечения с прямой $y = x$. Абсцисса точки пересечения - новая итерация $x^{(1)}$. Для неё проводится то же построение. Последовательность итераций на рисунке сходится к точному значению корня: предел $\tilde{x}$ последовательности $\{x^{(k)}\}$ существует и совпадает с корнем.

Последовательность $\{x^{(k)}\}$ может расходиться, как показано на рис. 4. Это не значит, что уравнение не имеет корня. Просто, последовательность к нему не сходится.

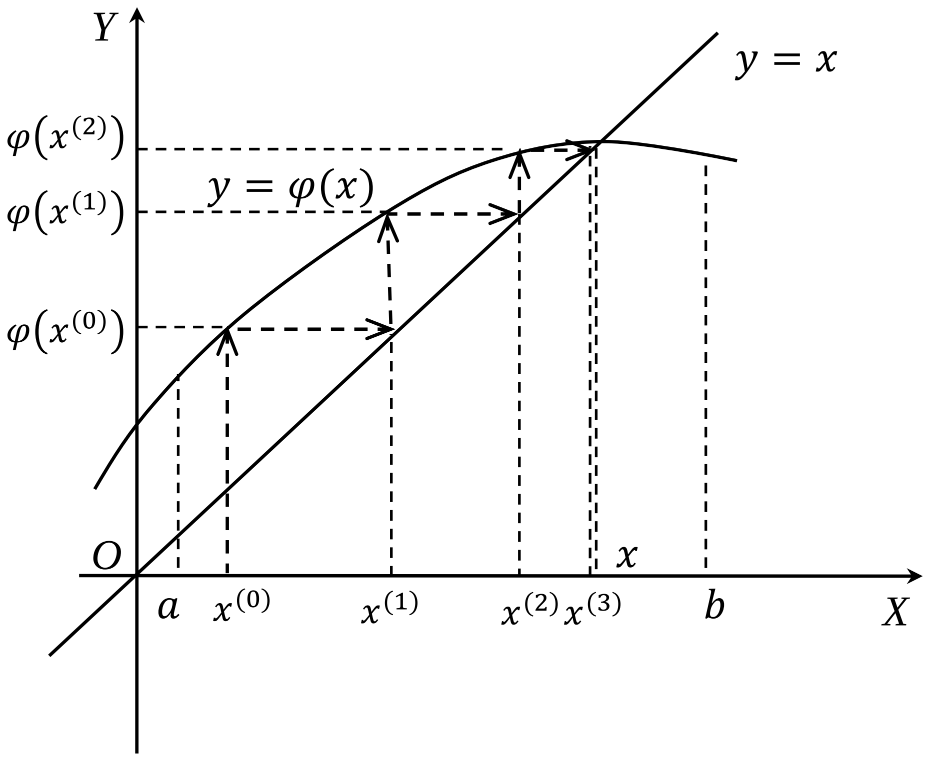

Рис. 3. Метод простой итерации

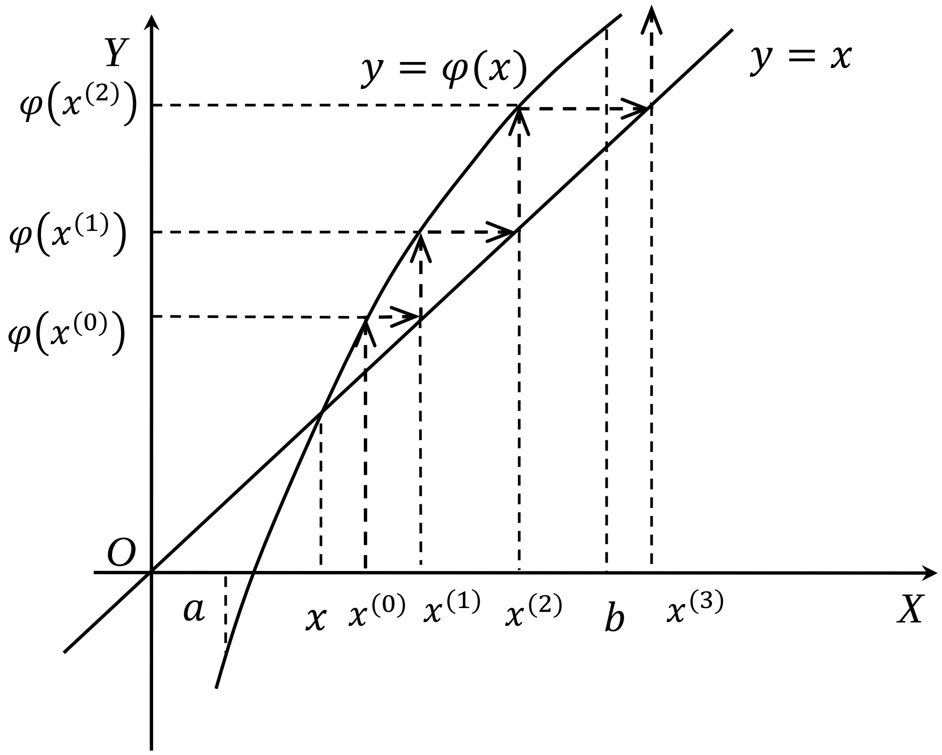

Рис. 4. Расходимость метода простой итерации

4.2. Сходимость метода

Естественно, возникает вопрос об условиях сходимости, а также оценке погрешности итерации для критерия останова. Это даёт следующая теорема.

Теорема 3. Пусть для функции φ и начального приближения $x^{(0)}$ выполнены условия:

1. Для любых x' и x'' из отрезка $\left|x - x^{(0)}\right| \leq \delta$ имеет место неравенство

$$|\varphi(x') - \varphi(x'')| \leq q|x' - x''|,$$

где $0 < q < 1$.

2. Справедливо неравенство

$$\frac{m}{1-q} \leq \delta,$$

где $m = \left|x^{(0)} - \varphi\left(x^{(0)}\right)\right|$.

Тогда уравнение имеет единственный корень x, к которому сходится последовательность $\left\{x^{(k)}\right\}$ метода простой итерации, и имеет место оценка погрешности

$$\left|x - x^{(k)}\right| \leq \frac{m}{1-q} q^k,$$

где $k = 0, 1, 2, \ldots$.

С помощью этой теоремы можно вести вычисления до достижения заданной точности: критерием останова будет выполнение неравенства

$$\frac{m}{1-q}q^k < \varepsilon.$$

Более того, можно найти число k необходимых шагов из этого же неравенства. Однако теорема требует знания q, что может быть не так просто (вторую константу m найти очень просто). Поэтому были найдены другие, более практичные, условия сходимости. Одно из них даёт теорема 4.

Теорема 4. Пусть на отрезке локализации $[a; b]$ функция φ определена, непрерывна и дифференцируема и выполняются условия:

1. $\varphi(x) \in [a; b]$;
2. $|\varphi'(x)| \le q < 1$.

Тогда для любого начального приближения $x^{(0)} \in [a; b]$ последовательность $\{x^{(k)}\}$ метода простой итерации сходится к корню уравнения, при этом справедлива оценка

$$\left|x - x^{(k)}\right| \le \frac{q}{1-q}\left|x^{(k)} - x^{(k-1)}\right|. \qquad (7)$$

Доказательство. Вычитая (4) из (6), получаем

$$x^{(k)} - x = \varphi\left(x^{(k-1)}\right) - \varphi(x). \qquad (8)$$

Применяем к (8) теорему Лагранжа (функция φ удовлетворяет её условиям):

$$x^{(k)} - x = \varphi\left(x^{(k-1)}\right) - \varphi(x) = \varphi'(\xi_k)\left(x^{(k-1)} - x\right),$$

где ξ_k – некоторая точка между $x^{(k-1)}$ и x. В силу первого условия теоремы и формулы (6) при любом начальном приближении $x^{(0)} \in [a; b]$ последовательность $\{x^{(k)}\}$ не выходит за пределы отрезка $[a; b]$. А значит, $\xi_k \in [a; b]$ при всех k. Отсюда в силу второго условия следует оценка:

$$\left|x^{(k)} - x\right| = \left|\varphi'(\xi_k)\left(x^{(k-1)} - x\right)\right| = |\varphi'(\xi_k)| \cdot \left|x^{(k-1)} - x\right| \le q\left|x^{(k-1)} - x\right|. \qquad (9)$$

Применяя последовательно (9), имеем

$$\left|x^{(k)} - x\right| \le q\left|x^{(k-1)} - x\right| \le q^2\left|x^{(k-2)} - x\right| \le \cdots \le q^k\left|x^{(0)} - x\right|.$$

А из этого неравенства следует, что

$$\lim_{k\to\infty}\left|x^{(k)} - x\right| = 0.$$

Сходимость доказана. Теперь надо получить оценку (7). В правой части (9) добавим и вычтем $x^{(k)}$:

$$\left|x^{(k)}-x\right| \le q\left|x^{(k-1)}-x^{(k)}+x^{(k)}-x\right|.$$

Теперь применим свойство модуля:

$$\left|x^{(k)}-x\right| \le q\left(\left|x^{(k-1)}-x^{(k)}\right|+\left|x^{(k)}-x\right|\right)=q\left|x^{(k-1)}-x^{(k)}\right|+q\left|x^{(k)}-x\right|,$$

а отсюда уже нетрудно вывести (7) (по условию (2) теоремы $1-q>0$). ■

Константу q в этой теореме найти уже проще. Если в условиях теоремы $q \ll 1$, то можно пользоваться следующим правилом. Вычисления надо вести до тех пор, пока два последовательных приближения не совпадут в пределах заданной точности:

$$\left|x^{(k)}-x^{(k-1)}\right| < \varepsilon.$$

В связи с рассмотрением вопроса о сходимости введём следующую терминологию. Пусть некоторый итерационный процесс генерирует последовательность $\left\{x^{(k)}\right\}_{k=0}^{\infty}$, имеющую пределом $\tilde{x}$. Сходимость последовательности итераций к $\tilde{x}$ называется *линейной*, если существуют такая постоянная $C \in (0;1)$ и такой номер K, что

$$\left|\tilde{x}-x^{(k+1)}\right| \le C\left|\tilde{x}-x^{(k)}\right|$$

при всех $k \ge K$. Сходимость называется *сверхлинейной*, если существуют такая положительная сходящаяся к нулю числовая последовательность $\left\{C^{(k)}\right\}_{k=0}^{\infty}$ и такой номер K, что

$$\left|\tilde{x}-x^{(k+1)}\right| \le C^{(k)}\left|\tilde{x}-x^{(k)}\right|$$

при всех $k \ge K$.

Последовательность $\left\{x^{(k)}\right\}_{k=0}^{\infty}$ сходится к $\tilde{x}$ по меньшей мере *с p-порядком*, если найдутся такие константы $C>0, p \ge 1, K>0$, что

$$\left|\tilde{x}-x^{(k+1)}\right| \le C\left|\tilde{x}-x^{(k)}\right|^{p}$$

при всех $k \ge K$. При $p=1$ получается линейная сходимость.

Из (9) следует, что метод простой итерации имеет линейную сходимость. В лекции 10.2 такая сходимость была названа сходимостью со скоростью геометрической прогрессии. Число q из теоремы 4 – это знаменатель прогрессии, характеризующий скорость сходимости: чем меньше q, тем она выше.

В заключение вернёмся к вопросу о переходе от уравнения (1) к (5). Из теоремы 4 следует, что число c надо подобрать так, чтобы производная $\varphi'(x)=1-cf'(x)$ была мала по модулю в нужной области. Для конкретного подбора c надо знать свойства производной f'. Если $0 < \alpha \le f'(x) \le \gamma < \infty$, то $1 - c\gamma \le \varphi'(x) \le 1 - c\alpha$. Значит, $|\varphi'(x)| \le$ $\le q(c)$, где $q(c) = \max\{|1 - c\alpha|; |1 - c\gamma|\}$. Если $c \in \left(0; \frac{\alpha}{\gamma}\right)$, то $q(c) < 1$. Оптимальное значение c, при котором q минимально, равно

$$c_0 = \frac{2}{\alpha + \gamma}.$$

Если же нет нужных оценок для f', то можно подобрать c так, чтобы

$$\varphi'\left(x^{(0)}\right) = 0. \tag{10}$$

Тогда в некоторой малой окрестности $x^{(0)}$ производная φ' будет мала. Из (10) выводится выражение для c:

$$1 - \ cf'\left(x^{(0)}\right) = 0 \Rightarrow \ c = \frac{1}{f'(x^{(0)})}.$$

С этим c расчётная формула простых итераций (6) принимает вид

$$x^{(k)} = x^{(k-1)} - \frac{f\left(x^{(k-1)}\right)}{f'(x^{(0)})}.$$

Эта формула представляет собой формулу модифицированного метода Ньютона, который будет изучен в следующей лекции.

Пример. Рассмотрим уравнение

$$x \operatorname{tg} \frac{x}{3} - x - 1 = 0.$$

Методом простой итерации уточним корни, которые имеются на отрезке $[-\pi; \ \pi]$: один на отрезке $[-\pi; \ 0]$, другой – на $[0; \ \pi]$. Приведём исходное уравнение к виду (5). Для каждого отрезка, содержащего по одному корню, подберем число c.

Уточнение корня на отрезке $[-\boldsymbol{\pi}; \ \mathbf{0}]$. На этом отрезке производная f' отрицательна. Надо подобрать c так, чтобы $\varphi'(x) = 1 - cf'(x)$ в окрестности корня была по модулю меньше единицы. Поскольку

$$f'(x) = \operatorname{tg} \frac{x}{3} + \frac{x}{3\cos^2 \frac{x}{3}} - 1,$$

$$f'(-\pi) = -\sqrt{3} - \frac{4\pi}{3} - 1 \approx -6{,}9,$$

$$f'(0) = -1,$$

то $-6{,}9 \le f'(x) \le -1$ (ранее мы выяснили, что f'' положительна на всём отрезке $[-\pi;\ \pi]$). Можно взять $c = -\frac{1}{4}$, для которого $|\varphi'(x)| < 1$, $x \in [-\pi;\ 0]$. Если взять начальное приближение к корню $x^{(0)} = -1{,}5$, то на 11-й итерации имеем $x^{(11)} = -0{,}792$, $f\left(x^{(11)}\right) = 0{,}0062$. Итерационная последовательность сходится медленно. Ситуацию можно улучшить, либо изменив начальное приближение, либо взяв другое число c. Для $c = -\frac{1}{2}$ при том же начальном приближении имеем $x^{(3)} = -0{,}791$, $f\left(x^{(3)}\right) = 0{,}004$.

Уточнение корня на отрезке $[\mathbf{0}; \boldsymbol{\pi}]$. На этом отрезке производная f' меняет знак, возрастая от -1 до $f'(\pi) = \sqrt{3} + \frac{4\pi}{3} - 1 \approx 5$. Для отрезка $[0; \pi]$ диапазон изменения производной такой, что невозможно подобрать c для уменьшения φ'. Уменьшим отрезок локализации положительного корня:

$$f'\left(\frac{\pi}{2}\right) = \frac{\sqrt{3}}{3} + \frac{2\pi}{9} - 1 \approx 0{,}3,$$

$$f\left(\frac{\pi}{2}\right) = \frac{\pi}{2} \cdot \frac{\sqrt{3}}{3} - \frac{\pi}{2} - 1 < 0.$$

Отрезок, содержащий корень, сужаем до $\left[\frac{\pi}{2}; \pi\right]$. На нём $0{,}3 \le f'(x) \le 5$. Для того, чтобы выполнялось условие $|\varphi'(x)| < 1$, можно взять $c = \frac{1}{3}$. Тогда, начиная с $x^{(0)} = 2$, получим

$$x^{(3)} = 2{,}804,$$

$$f\left(x^{(3)}\right) = -0{,}008.$$