

Численные методы

Конспект по 3 курсу специальности «прикладная
математика»
(лектор А. М. Будник)

Оглавление

1	Методы решения нелинейных уравнений.	2
1.1	Постановка задачи.	2
1.2	Метод простой итерации решения нелинейного уравнения.	3
1.3	Метод Ньютона решения нелинейного уравнения.	8
1.4	Видоизменения метода Ньютона и метода простой итерации.	13
1.4.1	Модификации метода Ньютона.	13
1.4.2	Модификации метода простой итерации.	15
1.5	Метод Лобачевского.	17
1.6	Методы решения систем нелинейных уравнений (СНУ).	20
1.6.1	Метод простой итерации (МПИ).	20
1.6.2	Видоизменения метода простой итерации.	22
1.6.3	Метод Ньютона.	23
1.6.4	Видоизменения метода Ньютона.	25
1.6.5	Метод градиентного спуска.	26
2	Приближение функций.	27
2.1	Общие положения проблемы приближения функций.	27

Глава 1

Методы решения нелинейных уравнений.

В данной главе будут рассмотрены некоторые методы решения нелинейных уравнений и систем уравнений. Рассмотрим случай одного нелинейного уравнения.

1.1 Постановка задачи.

Пусть задана функция $f(x)$ действительного переменного $x \in \mathbb{R}$. Требуется найти корни уравнения

$$f(x) = 0, \quad (1)$$

или, что то же самое, нули функции $f(x)$. Выясним, является ли задача корректно поставленной. Для ответа на вопрос существования и единственности решения введем теорему из математического анализа.

Теорема. *Если функция $f(x)$ непрерывна на отрезке $[a, b]$ и принимает на его концах значения разных знаков, то на этом отрезке существует по крайней мере один корень уравнения $f(x) = 0$. Если при этом функция $f(x)$ будет монотонной на отрезке $[a, b]$, то она может иметь только один корень.*

◆ Без доказательства. ☒

Выясним условие устойчивости для рассматриваемой задачи. Как правило, в качестве входных данных мы имеем функцию $f(x)$, заданную в виде функциональной формы. Поэтому понятие устойчивости здесь отпадает.

Нелинейное уравнение в зависимости от вида функции $f(x)$ можно разделить на два класса:

1. алгебраические;
2. трансцендентные.

В первом классе функция $f(x)$ содержит только алгебраические функции. Например, полином $P_n(x)$ является целой алгебраической функцией. Ко второму классу относятся все другие функции, которые содержат выражения тригонометрические, показательные, логарифмические и так далее.

Методы решения нелинейных уравнений делятся на прямые и итерационные. Мы будем

рассматривать лишь итерационные методы.

Задача нахождения корней уравнения (1) обычно решается в два этапа:

1. отделение корней;

На этом этапе изучается расположение корней в общем случае на комплексной плоскости, проводится их разделение, т.е. выделяются области, содержащие только один корень. Кроме того изучается вопрос о кратности корней. Находятся некоторые начальные приближения x^0 для точного решения.

2. построение метода.

На этом этапе, используя заданное начальное приближение, строится итерационный процесс, позволяющий уточнить значение отыскиваемого корня до некоторой заданной точности ε . Т.е., зная x^0 , строим последовательность $x^k \xrightarrow[k \rightarrow \infty]{\varepsilon} x^*$.

В заключение этого параграфа запишем несколько соображений, касающихся первого этапа. По отделению корней мы можем выделить несколько способов нахождения начального приближения:

- из физических соображений;
- графический способ;
- построение таблицы значений функции $f(x)$ на заданной сетке узлов;
- метод деления отрезка пополам (метод дихотомии, метод бисекции).

Метод деления отрезка пополам заключается в том, что мы берем отрезок $[a, b]$ и смотрим, чтобы на этом отрезке функция меняла знак. Затем делим отрезок пополам, берем точку $c : a < c < b$ и в зависимости от того, где меняется знак, переходим к следующему отрезку и так далее. В итоге мы приходим к тому, что отрезок получится меньше ε , то есть мы и получим искомый корень. Число делений отрезка пополам $N \geq \log_2 \frac{b-a}{\varepsilon}$.

1.2 Метод простой итерации решения нелинейного уравнения.

Применение метода требует предварительного приведения уравнения $f(x) = 0$ к каноническому виду

$$x = \varphi(x), \quad (1)$$

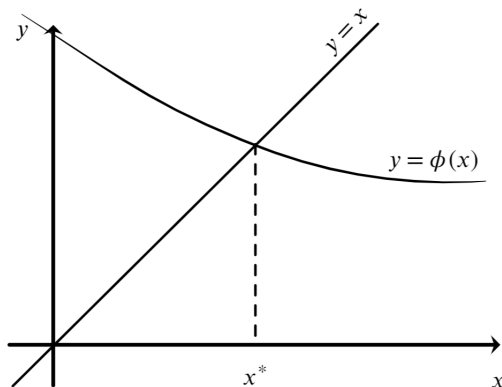
где $\varphi(x)$ — это заданная функция. Метод простой итерации будет иметь следующий вид:

$$x^{k+1} = \varphi(x^k), \quad k = 0, 1, 2, \dots, \quad (2)$$

где x^k — последовательность, начинающаяся с x^0 , которая должна сходиться к точному решению. Область изменения аргумента x на числовой оси обозначим через X , а через Y обозначим область значений функции $y = \varphi(x)$. Тогда функцию $\varphi(x)$ можно рассматривать как оператор, преобразующий X в Y :

$$\varphi : X \rightarrow Y.$$

Таким образом, нам нужно найти такие точки области X , которые при преобразовании оператором φ переходят сами в себя, то есть точки остающиеся неподвижными при преобразовании X в Y . Значит решения уравнения (1) — это точки, остающиеся неподвижными при преобразовании X в Y . Геометрически это можно изобразить следующим образом:



Таким образом, после процедуры отделения корней мы находим начальное приближение x^0 в окрестности корня x^* . И по найденному начальному приближению по формуле (2) строится итерационная последовательность, которая и называется **методом простой итерации**.

Мы должны обеспечить сходимость этого итерационного процесса. Сформулируем и докажем для этого теорему.

Теорема (о сходимости метода простой итерации). Пусть выполняются следующие условия:

1. функция $\varphi(x)$ определена на отрезке

$$|x - x^0| \leq \delta, \quad (3)$$

непрерывна на нем и удовлетворяет условию Липшица с постоянным коэффициентом меньше единицы, то есть $\forall x, \tilde{x}$

$$|\varphi(x) - \varphi(\tilde{x})| \leq q|x - \tilde{x}|, \quad 0 \leq q < 1; \quad (4)$$

2. для начального приближения x^0 верно неравенство

$$|x^0 - \varphi(x^0)| \leq m;$$

3. числа δ, q, m удовлетворяют условию

$$\frac{m}{1 - q} \leq \delta. \quad (5)$$

Тогда

1. уравнение (1) в области (3) имеет решение;
2. последовательность x^k построенная по правилу (2) принадлежит отрезку $[x^0 - \delta, x^0 + \delta]$, является сходящейся и ее предел удовлетворяет уравнению (1):

$$x^k \xrightarrow[k \rightarrow \infty]{} x^*;$$

3. скорость сходимости x^k к x^* оценивается неравенством

$$|x^* - x^k| \leq \frac{m}{1-q} q^k, \quad k = 1, 2, \dots \quad (6)$$

Также эта теорема может называться **методом сжимающих отображений**.

◆ Докажем второй пункт, т.е. принадлежность. Методом математической индукции покажем, что при всех значениях $k = 1, 2, \dots$ приближения $x^k \in [x^0 - \delta, x^0 + \delta]$ и для них верно неравенство

$$|x^{k+1} - x^k| \leq m q^k. \quad (7)$$

При $k = 0$ имеем $x^1 = \varphi(x^0)$, а x^1 всегда может быть найден, поскольку φ определена в x^0 . Кроме того

$$|x^1 - x^0| = |\varphi(x^0) - x^0| \leq m,$$

т.е. формула (7) справедлива. Докажем, что x^1 находится не дальше, чем m от x^0 :

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

отсюда следует, что $x^1 \in [x^0 - \delta, x^0 + \delta]$.

Пусть данное предположение справедливо при $x^0, x^1, \dots, x^k \in [x^0 - \delta, x^0 + \delta]$ и

$$|x^{n+1} - x^n| \leq m q^n, \quad n = 0, 1, \dots, k-1.$$

По предположению $x^k \in [x^0 - \delta, x^0 + \delta]$, следовательно, $x^{k+1} = \varphi(x^k)$ может быть вычислено. По сделанному допущению справедливо

$$|x^k - x^{k-1}| \leq m q^{k-1}.$$

Теперь рассмотрим неравенство для $k+1$ -ой итерации:

$$|x^{k+1} - x^k| = |\varphi(x^k) - \varphi(x^{k-1})| \leq q |x^k - x^{k-1}| \leq m q^k.$$

Осталось проверить $x^k \in [x^0 - \delta, x^0 + \delta]$. Рассмотрим разность

$$|x^{k+1} - x^0| = \left| (x^{k+1} - x^k) + (x^k - x^{k-1}) + \dots + (x^1 - x^0) \right| \leq m q^k + m q^{k-1} + \dots + m.$$

Легко видеть, что эта сумма легко подсчитывается, как сумма геометрической прогрессии, и равна

$$\frac{m - m q^{k+1}}{1 - q} < \frac{m}{1 - q} \leq \delta.$$

Итак, мы доказали, что x^{k+1} принадлежит отрезку (3).

Докажем сходимость последовательности. Для этого покажем, что для последовательности выполняется условие Больцано-Коши

$$|x^{k+p} - x^k| = \left| (x^{k+p} - x^{k+p-1}) + (x^{k+p-1} - x^{k+p-2}) + \dots + (x^{k+1} - x^k) \right| \leq \frac{m}{1-q} q^k.$$

Так как оценка не зависит от p и учитывая то, что $0 \leq q < 1$, можно утверждать, что признак сходимости для последовательности x^k выполняется, а значит существует предел этой последовательности

$$\exists \lim_{k \rightarrow \infty} x^k = x^*.$$

Нужно доказать, что $x^* \in [x^0 - \delta; x^0 + \delta]$ и x^* удовлетворяет формуле (1). Это следует из того, что все x^k принадлежат этому отрезку, то есть и предел находится в этом отрезке. Для доказательства второго в формуле (2) устремим $k \rightarrow \infty$:

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

Ввиду непрерывности функции, x^* является решением искомого уравнения, т.е. уравнение (1) превращается в тождество.

Последнее, что нужно доказать, — оценка из пункта 3. Для получения неравенства (6) достаточно в соотношении

$$|x^{k+p} - x^k| \leq \frac{m}{1-q} q^k.$$

устремить $p \rightarrow \infty$. То есть

$$|x^* - x^k| \leq \frac{m}{1-q} q^k,$$

что и является искомой оценкой. □

Замечания.

1. На всяком множестве точек, где для функции $\varphi(x)$ выполняется условие

$$|\varphi(x) - \varphi(y)| < |x - y|, \quad x \neq y$$

уравнение (1) может иметь не более одного решения.

2. Пользуясь оценкой (6), можно получить априорное количество итераций, необходимое для получения приближенного решения с заданной точностью

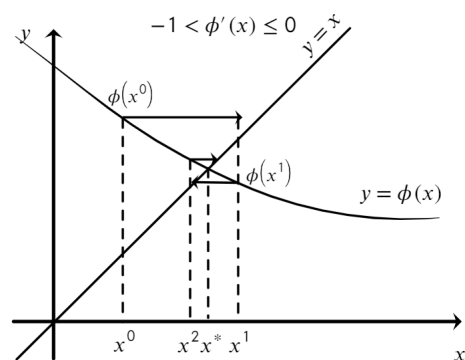
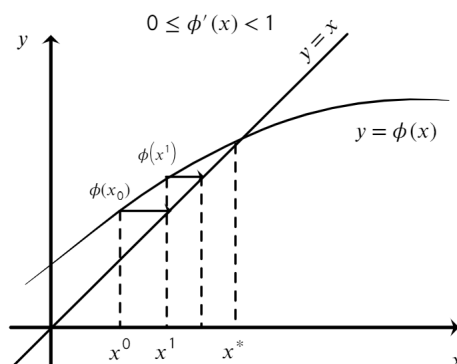
$$k \geq \frac{\lg \frac{\varepsilon(1-q)}{m}}{\lg q}.$$

3. Для построения сходящегося метода простой итерации в практических вычислениях условие 1 теоремы о сходимости метода простой итерации обычно заменяется более строгим требованием, а именно для всех x из отрезка $|x - x^0| \leq \delta$ функция $\varphi(x)$ имеет непрерывную первую производную $\varphi'(x)$ такую, что

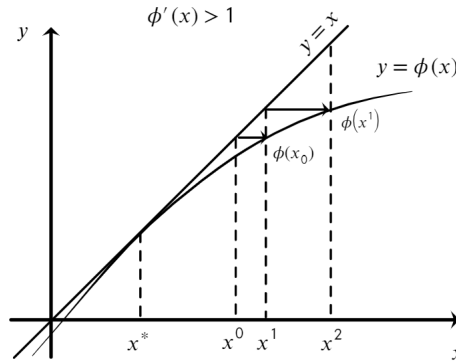
$$|\varphi'(x)| < 1 \quad \forall x \in [x_0 - \delta; x_0 + \delta].$$

Более того, если $0 \leq \varphi'(x) < 1$, то поведение последовательных приближений будет монотонным. Если $-1 < \varphi'(x) \leq 0$, то поведение итерационной последовательности будет колебательным.

Геометрический смысл метода простой итерации продемонстрируем на графике:



В свою очередь, при $\varphi'(x) > 1$ процесс расходится, это можно увидеть из графика



4. Так как сходимость метода простых итераций возможна при сжимающем отображении, то условие $|\varphi'(x)| < 1$ является определяющим при приведении исходного уравнения к каноническому виду.

Наиболее универсальным способом приведения к каноническому виду является преобразование

$$x = \underbrace{x + f(x)}_{\varphi(x)},$$

но нам необходимо выполнение условия $|\varphi'(x)| < 1$. Поэтому мы вводим параметр $\psi(x)$, выбираемый таким образом, чтобы обеспечить сходимость:

$$x = \underbrace{x + \psi(x)f(x)}_{\varphi(x)},$$

Параметр $\psi(x)$ должен быть непрерывным и $\psi(x^*) \neq 0$. Самый простой вариант — взять постоянную функцию $\psi(x) = \text{const}$ и подобрать эту константу из условия $|\varphi'(x)| < 1$.

5. Поведение последовательности приближений мы будем исследовать, изучая величину

$$\varepsilon_k = x^* - x^k$$

— **погрешность приближенного решения на k -ой итерации**. Из этого соотношения легко увидеть, что

$$x^k = x^* - \varepsilon_k$$

и подставим это в формулу (2). Тогда

$$x^* - \varepsilon_{k+1} = \varphi(x^* - \varepsilon_k).$$

Предполагая, что функция $\varphi(x)$ имеет непрерывную производную в окрестности точек x_k и x_{k+1} , разложим правую часть в ряд Тейлора в окрестности x^* :

$$x^* - \varepsilon_{k+1} = \varphi(x^*) - \varphi'(x^*)\varepsilon_k + O(\varepsilon_k^2).$$

Такое разложение возможно при условии, что функция $\varphi(x)$ дифференцируема и при предположении достаточной малости ε_k , чтобы мы могли отбросить остальные члены. Учитывая $x^* = \varphi(x^*)$ и отбрасывая достаточно малые слагаемые $O(\varepsilon_k^2)$, получим

$$\varepsilon_{k+1} \approx \varphi'(x^*)\varepsilon_k. \quad (8)$$

Формула (8) дает ответ о скорости сходимости метода простой итерации. То есть погрешность на каждой итерации уменьшается по сравнению с предыдущей в величину $\varphi'(x^*)$. Таким образом,

- (а) нам нужно обеспечить $|\varphi'| < 1$, чтобы $\varepsilon_{k+1} < \varepsilon_k$;
- (б) сходимость метода осуществляется по закону геометрической прогрессии со знаменателем $q = \varphi'$.

1.3 Метод Ньютона решения нелинейного уравнения.

Рассмотрим уравнение

$$f(x) = 0, \quad (1)$$

где $f(x)$ достаточно гладкая функция вещественного переменного. Предположим, что для точного решения x^* каким-либо образом задано начальное приближение x^0 . Для построения метода рассмотрим погрешность $\varepsilon_0 = x^* - x^0$. В предположении, что ε_0 достаточно малая по модулю величина, подставим в уравнение (1) x^* вместо x , тогда

$$f(x^0 + \varepsilon_0) = 0.$$

Разложим это выражение в ряд Тейлора в окрестности точки x^0 :

$$f(x^0 + \varepsilon_0) = f(x^0) + \varepsilon_0 f'(x^0) + O(\varepsilon_0^2) = 0.$$

Теперь отбросим слагаемое $O(\varepsilon_0^2)$ и получим в рамках отброшенной величины получим приближенное уравнение

$$f(x^0) + \varepsilon_0 f'(x^0) \approx 0.$$

Решая это уравнение относительно ε_0 , получим

$$\varepsilon_0 \approx -\frac{f(x^0)}{f'(x^0)}.$$

Тогда выразим из $x^* = x^0 + \varepsilon_0$ и учитывая, что равенство приближенное, получим

$$x^* \approx x^0 - \frac{f(x^0)}{f'(x^0)}.$$

В итоге, повторяя описанную процедуру, мы можем построить итерационную формулу, которая носит название **метода Ньютона**

$$x^{k+1} = x^k - \frac{f(x^k)}{f'(x^k)}, \quad k = 0, 1, \dots; \quad x_0 \quad (2)$$

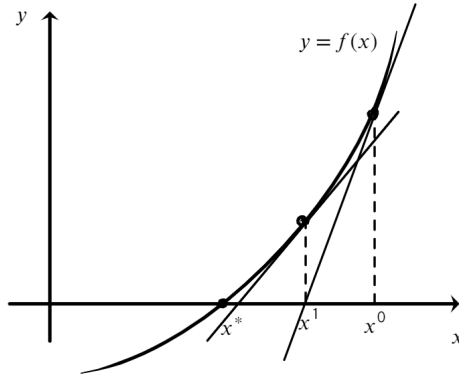
(добавка x_0 означает, что начальное приближение задано). Иногда этот метод называют **методом касательных**. Это название следует из геометрического смысла. Если рассмотреть уравнение кривой $y = f(x)$, то в точке x^k касательная к ней задается уравнением

$$y - f(x^k) = f'(x^k)(x - x^k).$$

Находим точку пересечения касательной с осью Ox , полагая $y = 0$, и тогда

$$x = x^k - \frac{f(x^k)}{f'(x^k)}.$$

Таким образом строим приближение x^{k+1} и так далее:



То есть мы приближаемся к корню по последовательности касательных прямых.

Выясним, какова скорость сходимости этого метода. С помощью подстановки получим формулу для скорости сходимости

$$\varepsilon_{k+1} = \frac{\varepsilon_k f'(x^* - \varepsilon_k) + f(x^* - \varepsilon_k)}{f'(x^* - \varepsilon_k)}.$$

Для того, чтобы получить ответ на вопрос, какова скорость сходимости, необходимо сделать несколько преобразований данного выражения. Воспользуемся тем, что мы можем разложить функции в ряд Тейлора в окрестности x^* :

$$f(x^* - \varepsilon_k) = f(x^*) - \varepsilon_k f'(x^*) + \frac{1}{2} \varepsilon_k^2 f''(x^*) + O(\varepsilon_k^3),$$

$$f'(x^* - \varepsilon_k) = f'(x^*) - \varepsilon_k f''(x^*) + \frac{1}{2} \varepsilon_k^2 f'''(x^*) + O(\varepsilon_k^3).$$

В итоге после подстановки мы получим формулу

$$\varepsilon_{k+1} = -\frac{1}{2} \frac{f''(x^*)}{f'(x^*)} \varepsilon_k^2 + O(\varepsilon_k^3).$$

Отбросив величину более высокого порядка, чем ε_k^2 , мы получим приближенное равенство

$$\varepsilon_{k+1} \approx -\frac{1}{2} \frac{f''(x^*)}{f'(x^*)} \varepsilon_k^2 = \alpha \varepsilon_k^2. \quad (3)$$

Формула (3) доказывает, что при $|\alpha| < 1$ последовательность x^k построенная по формуле (2) обладает квадратичной сходимостью.

Теорема (о сходимости метода Ньютона). Пусть выполняются следующие условия:

1. Функция $f(x)$ определена и дважды непрерывно дифференцируема на отрезке

$$s_0 = [x^0; x^0 + 2h_0], \quad h_0 = -\frac{f(x^0)}{f'(x^0)}.$$

При этом на концах отрезка $f(x)f'(x) \neq 0$.

2. Для начального приближения x^0 выполняется неравенство

$$2|h_0|M \leq |f'(x_0)|, \quad M = \max_{x \in s_0} |f''(x)|.$$

Тогда справедливы следующие утверждения:

1. Внутри отрезка s_0 уравнение $f(x) = 0$ имеет корень x^* и при этом этот корень единственный.
2. Последовательность приближений x^k , $k = 1, 2, \dots$ может быть построена по формуле (2) с заданным приближением x^0 .
3. Последовательность x^k сходится к корню x^* , то есть $x^k \xrightarrow[k \rightarrow \infty]{} x^*$.
4. Скорость сходимости характеризуется неравенством

$$|x^* - x^{k+1}| \leq |x^{k+1} - x^k| \leq \frac{M}{2|f'(x^*)|} \cdot (x^k - x^{k-1})^2, \quad k = 0, 1, 2, \dots \quad (4)$$

◆ Сначала докажем утверждение 2, т.е., что последовательность приближений x^k может быть построена. Будем доказывать по индукции. По условию 1 теоремы первый член последовательности (2) можно построить

$$x^1 = x^0 - \frac{f(x_0)}{f'(x_0)}, \quad f'(x_0) \neq 0.$$

Чтобы доказать возможность построения x^2 , докажем, что $x^1 \in s_0$ и $f'(x^1) \neq 0$. Учитывая тот факт, что

$$x_1 = x_0 + h_0,$$

получим тот факт, что x^1 является серединой отрезка s_0 . Далее рассмотрим следующее выражение, пользуясь вторым условием, теоремы

$$|f'(x^1) - f'(x^0)| = \left| \int_{x^0}^{x^1} f''(x) dx \right| \leq M|x^1 - x^0| = M|h_0| \leq \frac{|f'(x^0)|}{2}.$$

Теперь рассмотрим

$$|f'(x^1)| = |f'(x^0) - (f'(x^0) - f'(x^1))| \geq |f'(x^0)| - |f'(x^0) - f'(x^1)| \geq |f'(x^0)| - \frac{|f'(x^0)|}{2} = \frac{|f'(x^0)|}{2} \neq 0.$$

Таким образом, $f'(x^1) \neq 0$, а значит x^2 может быть построено. Тогда

$$x^2 = x^1 + h_1, \quad h_1 = -\frac{f(x^1)}{f'(x^1)}.$$

И так далее все x^k могут быть вычислены.

Рассмотрим, как себя ведут отрезки для того, чтобы доказать сходимость итерационного процесса. Наряду с отрезком s_0 рассмотрим отрезок

$$s_1 = [x^1; x^1 + 2h_1].$$

Середина этого отрезка — это x^2 . Покажем, что $s_1 \subset s_0$. Для этого нам нужно показать, что $h_1 < h_0$. Оценим величину h_1 . Для этого используем разложение в ряд Тейлора:

$$|f(x^1)| = \left| f(x^0) + h_0 f'(x^0) + \frac{h_0^2}{2} f''(x^0 + \theta h_0) \right| = \left| \frac{h_0^2}{2} f''(x^0 + \theta h_0) \right| \leq \frac{h_0^2}{2} M.$$

$$|h_1| = \left| -\frac{f(x^1)}{f'(x^1)} \right| \leq \frac{h_0^2}{2} \frac{M}{|f'(x^1)|} \leq \frac{h_0^2}{2} \frac{2M}{|f'(x^0)|} = h_0 \frac{M}{|f'(x^0)|} \leq \frac{|h_0|}{2}.$$

Итак $2|h_1| \leq |h_0|$, следовательно,

$$x^1 + 2h_1 = x^0 + h_0 + 2h_1 \leq x^0 + 2h_0 \in s_0.$$

Отсюда следует, что $s_1 \subset s_0$.

Далее мы можем показать по индуктивному предположению, что на отрезке s_1 итерация x^1 будет удовлетворять условиям 1 и 2 теоремы. Обе части неравенства $|h_1| \leq \frac{|h_0|}{2}$

умножим на $\frac{2M}{|f'(x^1)|}$, тогда

$$\frac{2M}{|f'(x^1)|} |h_1| \leq \frac{2|h_0|M}{2|f'(x^1)|}$$

Воспользуемся ранее произведенными оценками:

$$2|h_0|M \leq |f'(x^0)|, \quad 2|f'(x^1)| \geq \frac{1}{2}|f'(x^0)|$$

Тогда

$$\frac{2|h_0|M}{2|f'(x^1)|} \leq 1 \Rightarrow 2|h_1|M \leq |f'(x^1)|.$$

Таким образом, на отрезке s_1 функция $f(x)$ удовлетворяет условиям теоремы 1 и 2. Теперь по индукции очевидна возможность построения последовательности x^{k+1} по формуле (2). При этом x^{k+1} является серединой отрезка

$$s_k = [x^k; x^k + 2h_k], \quad h_k = -\frac{f(x^k)}{f'(x^k)}.$$

А отрезок $s_k \subset s_{k-1}$ и не превосходит половины длины s_{k-1} . Кроме того, выполняется неравенство, являющегося оценкой половины длины отрезка

$$|h_k| \leq \frac{h_{k-1}^2 M}{2|f'(x^k)|}.$$

То есть мы доказали утверждение 2.

Докажем утверждения 3 и 1. Так как мы построили последовательность вложенных отрезков

$$s_k \subset s_{k-1} \subset \dots \subset s_1 \subset s_0,$$

длины которых с ростом k стремятся к нулю, то, таким образом, эти отрезки стягиваются в точку. А следовательно последовательность x^{k+1} , элементы которой являются серединами этих отрезков, также является сходящейся к некоторому значению x^* . Отсюда

$$x^{k+1} \xrightarrow[k \rightarrow \infty]{} x^*,$$

но существование предела еще не означает, что это нужный нам предел. Покажем, что x^* – это корень уравнения (1). Для этого в формуле (2) перейдем к пределу при $k \rightarrow \infty$:

$$x^* = x^* - \frac{f(x^*)}{f'(x^*)},$$

но дробь нужно рассмотреть отдельно. Для того, чтобы перейти к пределу в $f(x^k)$, мы должны доказать, что

$$\lim_{k \rightarrow \infty} f(x^k) = f(\lim_{k \rightarrow \infty} x^k),$$

этот переход возможен в силу непрерывности функции f и в силу того, что $f'(x^k) \neq 0 \forall k$. Тогда записанная нами формула будет верна. А из этой формулы можно сделать вывод, что

$$f(x^*) = 0.$$

Теперь докажем единственность этого корня x^* . Для этого предположим, что $M > 0$ (случай $M = 0$ мы не рассматриваем, иначе функция будет линейной, а в таком случае на первой же итерации мы получим точное решение). По условию теоремы

$$f'(x^0) \neq 0, \quad f'(x^0 + 2h_0) \neq 0.$$

Учитывая этот факт, мы можем утверждать, что

$$f'(x) \neq 0, \quad \forall x \in s_0,$$

действительно докажем это. Для этого рассмотрим любую точку отрезка $x \in s_0$:

$$|f'(x) - f'(x^0)| = \left| \int_{x^0}^x f''(t) dt \right| \leq M|x - x^0| < M \cdot 2|h_0| \leq |f'(x^0)|.$$

Теперь мы можем оценить величину $\forall x \in s_0$

$$|f'(x)| = |f'(x^0) - (f'(x^0) - f'(x))| \geq |f'(x^0)| - |f'(x^0) - f'(x)| > |f'(x^0)| - |f'(x_0)| = 0.$$

То есть $f'(x) \neq 0$ в любой точке отрезка s_0 . Этот факт говорит о том, что $f(x)$ строго монотонна на s_0 . Следовательно, уравнение (1) имеет не более одного корня.

Докажем утверждение 4. По доказанным ранее утверждениям x^{k+1} — это середина отрезка s_k длиной $2|h_k|$ и $x^* \in s_k$. Тогда можно рассмотреть

$$|x^* - x^{k+1}| \leq |h_k| \leq \frac{h_{k-1}^k M}{2|f'(x^1)|}, \quad k = 0, 1, \dots$$

Отсюда и следует формула (4). □

Замечания.

1. Из оценки (4) можно получить априорную оценку количества итераций, необходимых для достижения заданной точности ε (доказать самостоятельно)

$$k \geq \log_2 \frac{\ln(\alpha\varepsilon)}{\ln(\alpha|x^1 - x^0|)}, \quad \alpha = \max_{x \in s_0} \left| \frac{f''(x)}{2f'(x)} \right|.$$

2. Если в окрестности корня производная $f'(x)$ сохраняет знак и монотонна, то приближение x_k построенное по формуле (2) сходится с одной стороны.

1.4 Видоизменения метода Ньютона и метода простой итерации.

1.4.1 Модификации метода Ньютона.

Все видоизменения связаны с тем, что мы хотим упростить формулу метода Ньютона и уменьшить количество арифметических операций, а для этого будем пытаться заменить вычисление производной вычислением другой более простой функции.

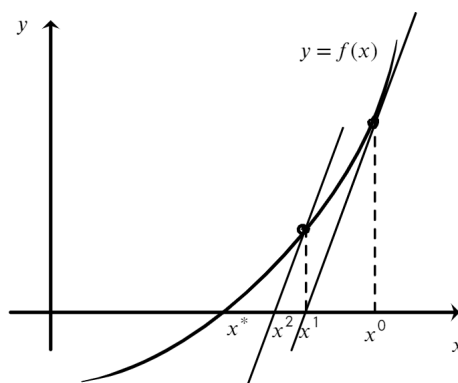
Метод Ньютона с постоянной производной.

Формула этого метода имеет следующий вид

$$x^{k+1} = x^k - \frac{f(x^k)}{f'(x^0)}, \quad k = 0, 1, \dots, \quad x^0.$$

Это видоизменение напрямую связано с уменьшением количества арифметических операций, поскольку мы отказываемся от вычисления последовательности $f'(x^k)$. Таким образом, с точки зрения количества операций метод простой итерации и метод Ньютона становятся сравнимы между собой.

Геометрически это означает, что, выбрав x^0 , мы движемся по касательной. Найдя x^1 , мы будем двигаться из точки x^1 по той же касательной, т.е. все касательные будут параллельны касательной в точке, которая является начальным приближением к корню.



Но скорость сходимости данного метода ухудшится. Легко видеть, что погрешность на каждой итерации будет меняться по следующему закону

$$\varepsilon_{k+1} = \varepsilon_k - \frac{f(x^* - \varepsilon_k)}{f'(x^0)}.$$

Проделав необходимые вычисления, связанные с разложением функции в окрестности x^* , можно получить

$$\varepsilon_{k+1} \approx \left(1 - \frac{f'(x^*)}{f'(x^0)}\right) \varepsilon_k. \quad (2)$$

Исходя из вида формулы (2), мы можем утверждать, что такая модификация имеет линейную скорость сходимости.

Метод секущих.

Возьмем за основу формулу производной

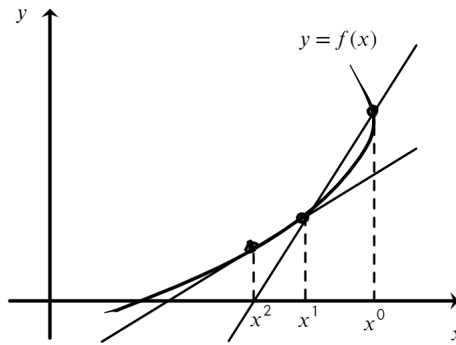
$$f'(x^k) \approx \frac{f(x^k) - f(x^{k-1})}{x^k - x^{k-1}}, \quad k = 1, 2, \dots$$

И, подставляя в формулу Ньютона, мы получим следующую формулу

$$x^{k+1} = x^k - f(x^k) \frac{x^k - x^{k-1}}{f(x^k) - f(x^{k-1})}, \quad k = 1, 2, \dots; \quad x^0 \quad (3)$$

Однако мы должны знать не только x^0 , но и x^1 , поэтому метод секущих двухшаговый.

Геометрически мы выбираем два приближения x^0 и x^1 и через две эти точки мы проводим прямую, и она является не касательной, а секущей. Таким образом, при пересечении секущей с осью Ox мы получаем точку x^2 . Проводим через x^1 и x^2 следующую секущую, получаем точку x^3 и так далее.



Количество операций в этом случае сравнимо с количеством операций метода Ньютона с постоянной производной. Но при этом мы выигрываем в скорости, покажем это. Мы имеем следующее уравнение для погрешности:

$$\varepsilon_{k+1} = \varepsilon_k - \frac{(\varepsilon_k - \varepsilon_{k+1})f(x^* - \varepsilon_k)}{f(x^* - \varepsilon_k) - f(x^* - \varepsilon_{k-1})}.$$

После выделения главной части из формулы и приведения подобных слагаемых, мы получим соотношение между погрешностями

$$\varepsilon_{k+1} \approx -\frac{1}{2} \frac{f''(x^*)}{f'(x^*)} \varepsilon_k \varepsilon_{k-1} \quad (4)$$

Таким образом, она выше чем линейная, но ниже, чем квадратичная. Для уточнения необходимо преобразовать данную величину. Соотношение на $k + 1$ и k итерациях может быть оценено как

$$\varepsilon_{k+1} \approx C \varepsilon_k^\alpha, \quad \alpha = \frac{1 + \sqrt{5}}{2}.$$

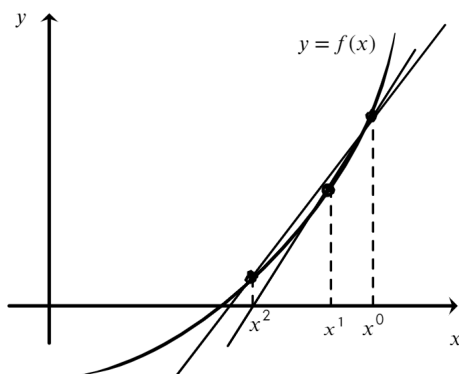
Метод хорд.

Формула метода хорд имеет вид

$$x^{k+1} = x^k - f(x^k) \frac{x^k - x^0}{f(x^k) - f(x^0)}, \quad k = 1, 2, \dots; \quad x^0, x^1 \quad (5)$$

Для подсчетов нам нужно два приближения, но сам метод одношаговый.

Геометрически мы строим хорды, проходящие через точку $f(x^0)$ и $f(x^k)$ на каждой итерации. Точка пересечения этой хорды с осью Ox приводит нас к новому приближению x^{k+1} .



В количестве операций мы не выигрываем. Можно показать, что погрешность в данном случае будет иметь вид

$$\varepsilon_{k+1} \approx -\frac{1}{2} \frac{f''(x^*)}{f'(x^*)} \varepsilon_0 \varepsilon_k \quad (6)$$

Отсюда можно сделать вывод, что метод хорд сходится по закону геометрической прогрессии, а значит по линейному закону, но знаменатель прогрессии будет зависеть от ε_0 . При достаточно хорошем начальном приближении этот метод может сходиться быстрее, чем остальные методы. Практически обычно метод хорд используется для того, чтобы сузить область, где находится корень.

1.4.2 Модификации метода простой итерации.

Все модификации сводятся к тому, что мы хотим повысить скорость сходимости метода.

Метод Стеффенсена.

Метод Стеффенсена основывается на том, что мы укажем способ вычисления x^{k+1} через x^k таким образом, чтобы обеспечить квадратичную скорость сходимости. Для увеличения скорости сходимости в данном методе используется преобразование Эйткена. Суть его состоит в том, что, если имеется сходящаяся последовательность чисел $s_0, s_1, \dots, s_n, \dots$, которая сходится к числу s , и при этом мы знаем, что характер сходимости носит вид

$$s_n = s + Aq^n, \quad A = \text{const}, q < 1,$$

то есть сходимость по закону геометрической прогрессии со знаменателем q . Тогда закон Эйткена позволяет сразу получить значение искомого предела по формуле Эйткена, построив последовательность

$$\sigma_0, \sigma_1, \dots, \sigma_n, \quad \sigma_n = s = \frac{s_{n+1}s_{n-1} - s_n^2}{s_{n+1} - 2s_n + s_{n-1}} = \lim_{n \rightarrow \infty} s_n. \quad (7)$$

Мы будем использовать эту формулу для того, чтобы сразу найти нужный нам предел в методе простой итерации.

Пусть мы имеем x^0 . Берем приближения

$$x^1 = \varphi(x^0), \quad x^2 = \varphi(x^1) = \varphi(\varphi(x^0)).$$

Тогда, используя формулу (7), мы можем при $n = 1$ получить

$$\sigma_1 = \frac{x^0 x^2 - (x^1)^2}{x^2 - 2x^1 + x^0} = \frac{x^0 \varphi(\varphi(x^0)) - (\varphi(x^0))^2}{\varphi(\varphi(x^0)) - 2\varphi(x^0) + x^0}.$$

Заменяем в этой формуле соответствующим образом индексы. В итоге получается итерационная формула, которая получила название **метод Стеффенсена**

$$x^{k+1} = \frac{x^k \varphi(\varphi(x^k)) - (\varphi(x^k))^2}{\varphi(\varphi(x^k)) - 2\varphi(x^k) + x^k}, \quad k = 0, 1, \dots; \quad x^0. \quad (8)$$

Метод Стеффенсена можно трактовать как метод простой итерации примененный к уравнению вида

$$x = \Phi(x),$$

где

$$\Phi(x) = \frac{x \varphi(\varphi(x)) - \varphi^2(x)}{\varphi(\varphi(x)) - 2\varphi(x) + x}.$$

Возникает вопрос сходимости метода. Можно доказать, что функция $\Phi(x)$ вместе со своей производной будет непрерывна в окрестности точки x^* , причем

$$\lim_{x \rightarrow x^*} \Phi(x) = x^*.$$

Если предположить, что функция $\Phi(x^*) = x^*$ (то есть мы доопределяем ее), то $\Phi(x)$ будет непрерывна в точке x^* . Кроме того можно утверждать, что

$$\Phi'(x^*) = \lim_{x \rightarrow x^*} \frac{\Phi(x) - \Phi(x^*)}{x - x^*} = 0.$$

Таким образом, можно утверждать, что сходимость метода Стеффенсена будет квадратичной. То есть мы построили метод, обладающий повышенной скоростью сходимости. Но в 2 раза увеличивается объем вычислений, из-за того, что нужно вычислять функцию $\varphi(\varphi(x))$.

Метод Чебышева.

Идея метода базируется на способе построения итерационного процесса таким образом, чтобы обеспечить обращение в ноль производных от функции $\varphi(x)$ в точке x^* , то есть мы берем уравнение

$$x = \varphi(x)$$

и стараемся построить метод, у которого максимальное количество производных обращается в ноль в точке x^* . Для этого функцию $\varphi(x)$ запишем в виде

$$\varphi(x) = x + \psi_1(x)f(x) + \psi_2(x)f^2(x) + \dots + \psi_{n-1}(x)f^{n-1}(x), \quad (9)$$

где $f(x)$ — это исходная функция, для которой мы ищем корни. Требуется выбрать функции $\psi_1(x), \dots, \psi_{n-1}(x)$ так, чтобы

$$\varphi^{(j)}(x) \Big|_{f(x)=0} = 0, \quad j = 1, 2, \dots, n-1 \quad (10)$$

Рассмотрим условие на первую производную

$$\begin{aligned} \varphi'(x) \Big|_{f(x)=0} &= 1 + \psi_1'(x)f(x) + \psi_1(x)f'(x) + \psi_2'(x)f^2(x) + 2\psi_2(x)f(x)f'(x) + \dots \Big|_{f(x)=0} = \\ &= 1 + \psi_1(x)f'(x) \Big|_{f(x)=0} = 0. \end{aligned}$$

Аналогичным образом мы можем записать вторую производную

$$\varphi''(x) \Big|_{f(x)=0} = 2\psi_1' f'(x) + \psi_1(x) f''(x) + 2\psi_2(x) (f'(x))^2 \Big|_{f(x)=0} = 0.$$

Из условия $\varphi'(x) \Big|_{f(x)=0} = 0$ следует, что функция

$$\psi_1(x) = -\frac{1}{f'(x)}.$$

Отсюда

$$\varphi(x) = x + \left(-\frac{1}{f'(x)}\right) f(x),$$

то есть мы пришли к методу Ньютона, итерационному процессу второго порядка. Из условия, что $\varphi''(x) \Big|_{f(x)=0} = 0$, применяя простые арифметические действия, мы можем получить

$$\psi_2(x) = -\frac{f''(x)}{2(f'(x))^3}.$$

Учитывая выражения для ψ_1 и ψ_2 мы можем построить итерационный процесс третьего порядка с кубической скоростью сходимости

$$x^{k+1} = x^k - \frac{f(x^k)}{f'(x^k)} - \frac{f^2(x^k) f''(x^k)}{2(f'(x^k))^3} \quad (11)$$

и будем называть эту формулу **методом Чебышева**. В этом методе мы также увеличиваем количество операций, так как необходимо вычислять значения $f(x)$, $f'(x)$, $f''(x)$.

1.5 Метод Лобачевского.

Метод Лобачевского является методом отыскания корней алгебраического уравнения. Данный метод не требует предварительного задания начального приближения для корней и кроме того позволяет найти сразу все корни.

Рассмотрим алгебраическое уравнение следующего вида

$$P(x) = a_0 x^n + a_1 x^{n-1} + \dots + a_{n-1} x + a_n = 0, \quad a_0 \neq 0. \quad (1)$$

Пусть $a_i \in \mathbb{R}$ и предположим, что все корни x_i простые, вещественные и удовлетворяют соотношению

$$|x_1| \gg |x_2| \gg \dots \gg |x_n|, \quad (2)$$

то есть отношение эквивалентно

$$\frac{|x_{k+1}|}{|x_k|} \ll 1, \quad k = \overline{1, n-1}.$$

Если выполняется условие (2), то говорят, что **корни сильно разделены** в смысле отношения $(k+1)$ -го корня к k -ому.

Из алгебры известно, что соотношение Виета, связывающее корни многочлена с его коэффициентами, является системой вида

[illegible]

В случае выполнения соотношения (2) в левых частях соотношений (3) главными членами будут все первые слагаемые. Тогда вместо точных равенств можно записать приближенные:

$$\left\{ \begin{array}{l} x_1 \approx -\frac{a_1}{a_0}, \\ x_1x_2 \approx \frac{a_2}{a_0}, \\ x_1x_2x_3 \approx -\frac{a_3}{a_0}, \\ \dots\dots\dots \\ x_1x_2\dots x_n \approx (-1)^n \frac{a_n}{a_0}. \end{array} \right.$$

Отсюда можно найти приближенные значения корней

$$x_i \approx -\frac{a_i}{a_{i-1}}, i = \overline{1, n}. \quad (4)$$

Если требование сильной разделенности корней не выполняется, то мы будем строить новое уравнение, корни которого будут высокими степенями корней исходного уравнения. При этом можно надеяться получить уравнение с сильно разделенными корнями. Метод Лобачевского основан на построении уравнения, корни которого являются квадратами корней исходного.

Запишем уравнение $P(x) = 0$ в виде

$$P(x) = a_0(x - x_1)(x - x_2) \dots (x - x_n) = 0.$$

Также рассмотрим полином

$$P^*(x) = a_0(x + x_1)(x + x_2) \dots (x + x_n) = 0.$$

Корни полинома $P^*(x)$ отличаются от корней исходного уравнения только знаком. По многочленам $P(x)$ и $P^*(x)$ построим многочлен $P_1(y)$ такой, что корнями этого полинома будут являться значения $y_i = x_i^2$, $i = \overline{1, n}$. Для этого перемножим

$$P(x)P^*(x) = a_0^2(x^2 - x_1^2)(x^2 - x_2^2) \dots (x^2 - x_n^2).$$

Сделав замену $y_i = x_i^2$, мы получим полином $P_1(y)$. Обозначим коэффициенты полинома $P_1(y)$ через $a_i^{(1)}$ и формально перемножим $P(x)$ с $P^*(x)$:

$$P(x)P^*(x) = (a_0x^n + a_1x^{n-1} + \dots + a_n)(a_0x^n - a_1x^{n-1} + \dots + (-1)^na_n).$$

И так далее. При достаточно больших k с требуемой точностью эти величины будут равны. Значит мы достигли требуемой степени разреженности корней. Таким образом, условием того, что достигнута требуемая разделимость корней является следующая связь между коэффициентами многочлена

$$a_i^{(k+1)} \approx (-1)^i (a_i^{(k)})^2. \quad (6)$$

Тогда, если это выполнено, x_i могут быть посчитаны по формулам

$$|x_i| = 2^{k+1} \sqrt{-\frac{a_i^{(k+1)}}{a_{i-1}^{(k+1)}}} \quad (7)$$

Таким образом, алгоритм метода Лобачевского определен.

1.6 Методы решения систем нелинейных уравнений (СНУ).

В общем виде СНУ можно записать как в координатном виде

$$f_i(x_1, \dots, x_n) = 0, \quad i = \overline{1, n}, \quad (1)$$

так и в векторном виде

$$f(x) = 0, \quad f = (f_1, \dots, f_n)^T, \quad x = (x_1, \dots, x_n) \quad (1)$$

Частные случаи при $n = 1$ – одно нелинейное уравнение. Если $n > 1$, но функции f_i линейные, то получим СЛАУ. Поэтому рассмотрим случай, когда функции нелинейные и $n \geq 2$. Основные этапы решения СНУ:

1. отделение корня;
2. построение последовательности приближений;
3. контроль сходимости.

Следует иметь ввиду, что проблема отделения корня в общем случае не имеет решений.

1.6.1 Метод простой итерации (МПИ).

Применение метода простой итерации требует приведения исходной системы (1) к виду удобному для итераций, т.е. канонической форме,

$$\begin{cases} x_1 = \varphi_1(x_1, \dots, x_n), \\ \text{\tiny}, \\ x_n = \varphi_n(x_1, \dots, x_n); \end{cases} \quad (2)$$

или в векторной форме

$$x = \varphi(x), \quad \varphi = (\varphi_1, \dots, \varphi_n)^T, \quad x = (x_1, \dots, x_n).$$

Мы будем предполагать, что $x^* = (x_1^*, \dots, x_n^*)$ — точное решение, а $x^k = (x_1^k, \dots, x_n^k)$ — итерационное приближение. Если выбрано начальное приближение x_0 , то все последующие приближения находятся по формуле

$$x_i^{k+1} = \varphi_i(x_1^k, \dots, x_n^k), \quad i = \overline{1, n}, \quad k = 0, 1, \dots$$

или в векторной форме

$$x^{k+1} = \varphi(x^k) \quad (3)$$

Будем считать функции $\varphi_i(x)$ непрерывно дифференцируемыми в общей области их задания. Будем далее предполагать, что решение x^* так же, как и все приближения x^k , лежат внутри этой области.

Выясним поведение вектора погрешности $\varepsilon^k = (\varepsilon_1^k, \dots, \varepsilon_n^k)$

$$\varepsilon^k = x^* - x^k = (x_1^* - x_1^k, \dots, x_n^* - x_n^k)$$

Посмотрим, как будет вести себя погрешность с увеличением количества итераций. Для этого подставим в формулу (3) выражение x^k через ε^k и получим следующее выражение

$$x_i^* - \varepsilon_i^{k+1} = \varphi_i(x_1^* - \varepsilon_1^k, \dots, x_n^* - \varepsilon_n^k).$$

Разложим правую часть последних равенств по степеням ε^k в окрестности точки x^* и выделим главную часть, учитывая, что $x^* = \varphi(x^*)$. Тогда получим

$$\varepsilon_i^{k+1} = \sum_{j=1}^n \frac{\partial}{\partial x_j} \varphi_i(x_1^*, \dots, x_n^*) \varepsilon_j^k + O\left(\max_j (\varepsilon_j^k)^2\right), \quad i = \overline{1, n}.$$

Если мы предположим, что ε_j^k достаточно малые, то можем отбросить достаточно малые слагаемые и записать приближенно отношения между k -ой и $(k+1)$ -ой итерациями:

$$\varepsilon^{k+1} \approx A \varepsilon^k \quad (4)$$

где A – это матрица Якоби построенная по системе функций φ_i :

$$A = \begin{pmatrix} \frac{\partial \varphi_1(x^*)}{\partial x_1} & \cdots & \frac{\partial \varphi_1(x^*)}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial \varphi_n(x^*)}{\partial x_1} & \cdots & \frac{\partial \varphi_n(x^*)}{\partial x_n} \end{pmatrix}$$

Таким образом, видно что вектор погрешности на одном шаге испытывает линейное преобразование. Для того, чтобы сделать вывод об условиях сходимости, преобразуем формулу (4). Запишем спектральное разложение матрицы A . Предположим, что все элементарные делители матрицы A являются простыми. Тогда эту матрицу можно записать в виде спектрального разложения

$$A = S^{-1} \Lambda S, \quad \Lambda = \text{diag}\{\lambda_1, \dots, \lambda_n\}$$

Тогда соотношение (4) можно записать в виде

$$\varepsilon^{k+1} \approx S^{-1} \Lambda S \varepsilon^k.$$

Обозначим

$$r^k = S \varepsilon^k.$$

Домножим обе части слева на S , тогда

$$r^{k+1} \approx \Lambda r^k$$

или по координатам

$$r_i^{k+1} = \lambda_i r_i^k, \quad i = \overline{1, n}$$

Если все $|\lambda_i| < 1 \ \forall i$, то

$$r_i^k \xrightarrow[k \rightarrow \infty]{} 0.$$

По условию $\varepsilon^k = S^{-1}r^k$, а значит

$$\varepsilon^k \xrightarrow[k \rightarrow \infty]{} 0.$$

Тогда можно утверждать, что

$$x^k \xrightarrow[k \rightarrow \infty]{} x^*.$$

Таким образом, необходимым и достаточным условием сходимости является условие $|\lambda_i| < 1 \forall i$ (по аналогии с МПИ для СЛАУ). Но в практических целях проверка этого условия достаточно затруднительна. Сформулируем теорему о достаточных условиях сходимости МПИ в случае СЧУ.

Теорема (о сходимости МПИ в случае СЧУ). Пусть выполняются условия

1. функции $\varphi_i(x_1, \dots, x_n)$ определены и непрерывно дифференцируемы в области

$$S_\delta = \max_{1 \leq i \leq n} |x_i - x_i^0| \leq \delta;$$

2. функции φ_i удовлетворяют на S_δ неравенству

$$\max_{1 \leq i \leq n} \max_{x \in S_\delta} \sum_{j=1}^n \left| \frac{\partial \varphi_i(x)}{\partial x_j} \right| \leq q < 1;$$

3. для начального приближения x^0 выполняется условие

$$\max_{1 \leq i \leq n} |x_i^0 - \varphi_i(x_1^0, \dots, x_n^0)| \leq m;$$

4. для чисел m, δ, q выполняется неравенство

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

Тогда исходная система (2) в области S_δ имеет решение x^* , к которому сходится итерационная последовательность x^k , вычисляемая по правилу (3). Кроме того скорость сходимости $x^k \rightarrow x^* \in S_\delta$ определяется неравенством

$$\max_{1 \leq i \leq n} |x_i^* - x_i^k| \leq \frac{m}{1-q} q^k.$$

♦ Аналогично подобной теореме для одномерного случая.

1.6.2 Видоизменения метода простой итерации.

Метод Зейделя.

Аналогично методу Зейделя для СЛАУ, уточненную координату x_i мы будем использовать при уточнении следующей координаты:

[illegible]

– **метод Зейделя** для системы (2). Скорость сходимости метода Зейделя линейная, как и у МПИ. Достаточное условие сходимости аналогично достаточному условию сходимости для МПИ, за исключением того, что в условии 2 теперь

$$\left\| \frac{\partial \varphi(x)}{\partial x} \right\|_1 < 1, \quad x \in S_\delta, \quad \|x - x^0\|_1 \leq \delta. \quad (6)$$

В отличие от метода Зейделя это видоизменение не требует предварительного приведения приведения системы (1) к каноническому виду. Итерационный процесс будет выглядеть следующим образом

Нахождение каждого нового значения x_i^{k+1} требует решения нелинейного уравнения вида

где все $x_j^k, j > i$ известны как значения с предыдущей итерации, а значения $x_j^{k+1}, j < i$ также известны как координаты точек, которые мы ранее вычислили. Тогда уравнение является уравнением от одной переменной x_i^{k+1} . Тогда для решения этого уравнения мы можем применять методы, используемые для одного уравнения.

В качестве примера запишем метод Гаусса-Зейделя с организацией внутреннего итерационного процесса по методу Ньютона (индекс k – внешний итерационный процесс, а s – внутренний)

В качестве начального приближения возьмем значение соответствующей компоненты, полученное на предыдущей внешней итерации

Построение итерационной последовательности аналогично случаю одного уравнения. Обозначим погрешность

23

Подставим x^* в уравнение и получим

$$f(x^k + \varepsilon^k) = 0.$$

Разложим в ряд Тейлора левую часть по степеням ε^k в окрестности x^* , отбросив достаточно малые слагаемые:

$$f(x^k) + \frac{\partial f(x^k)}{\partial x} \varepsilon^k \approx 0.$$

Обозначим

$$\Delta x^k \approx \varepsilon^k,$$

тогда получим следующее выражение

$$\frac{\partial f(x^k)}{\partial x} \Delta x^k = -f(x^k). \quad (9)$$

Таким образом, если матрица Якоби $\frac{\partial f(x^k)}{\partial x}$ будет невырожденной, то из последнего равенства можно единственным образом найти вектор поправок Δx^k . Тогда, зная Δx^k , мы можем построить итерационный процесс

$$x^{k+1} = x^k + \Delta x^k. \quad (10)$$

Если из (9) выразить Δx^k и подставить в формулу (10), то можно получить объединенную формулу

$$x^{k+1} = x^k - \left(\frac{\partial f(x^k)}{\partial x} \right)^{-1} f(x^k), \quad k = 0, 1, \dots \quad (11)$$

— **метод Ньютона** решения СНУ.

Если мы будем решать систему линейных алгебраических уравнений (9) итерационными методами, то мы снова получим вложенные итерационные процессы.

На качественных примерах можно показать, что этот метод будет обладать более высокой скоростью сходимости ежели метод простой итерации.

Сформулируем теорему о сходимости метода Ньютона для СНУ.

Теорема (о сходимости метода Ньютона для СНУ). *Пусть в области*

$$\Omega(x^*, \delta) = \{x \in \mathbb{R} : \|x^* - x\| \leq \delta\}$$

при некоторых значениях $\delta > 0$ и значениях констант $0 \leq a_1, a_2 < \infty$ выполнены условия

1.

$$\left\| \left(\frac{\partial f(x^k)}{\partial x} \right)^{-1} \right\| \leq a_1 \quad \forall x \in \Omega(x^*, \delta);$$

2.

$$\left\| f(x') - f(x'') - \frac{\partial f(x'')}{\partial x} (x' - x'') \right\| \leq a_2 \|x' - x''\|^2, \quad \forall x', x'' \in \Omega(x^*, \delta);$$

3.

$$x^0 \in \Omega(x^*, b), \quad b = \min\{\delta, c^{-1}\}, \quad c = a_1 \cdot a_2.$$

Тогда метод Ньютона (11) сходится в области $\Omega(x^*, b)$ и имеет место оценка погрешности

$$\|x^k - x^*\| \leq \frac{1}{c} \left(c \|x^0 - x^*\| \right)^{2^k}.$$

Замечания.

1. Из оценки погрешности следует квадратичная сходимость метода Ньютона.
2. В отличие от одномерного случая существования решения x^* предполагается, что существует обратная матрица $\left(\frac{\partial f}{\partial x}\right)^{-1}$.
3. Условие 2 теоремы выполняется автоматически, если функция $f \in C^2(\Omega(x^*, \delta))$.

1.6.4 Видоизменения метода Ньютона.

Метод Ньютона с постоянной матрицей Якоби.

Запишем формулы метода

$$\frac{\partial f(x^0)}{\partial x} \Delta x^k = -f(x^k), \quad x^{k+1} = x^k + \Delta x^k, \quad k = 0, 1, \dots \quad (12)$$

Данное видоизменение характеризуется тем, что на каждой итерации необходимо решать СЛАУ с одной и той же матрицей. Это позволяет уменьшить объем вычислений на каждом k -ом шаге.

Дискретный метод Ньютона.

Задаем некоторый векторный параметр $h^k \in \mathbb{R}^n$, все компоненты которого достаточно малы (близки к нулю) для того, чтобы обеспечить требуемую точность. Тогда частные производные, которые выходят в матрицу Якоби, мы приближенно заменяем на разность:

$$\frac{\partial f_i(x_1, \dots, x_n)}{\partial x_j} \approx \frac{f_i(x_1, \dots, x_j, \dots, x_n) - f_i(x_1, \dots, x_j - h_j^k, \dots, x_n)}{h_j^k}.$$

Тогда матрица Якоби в методе Ньютона заменяется некоторой матрицей:

$$\left(\frac{\partial f(x^k)}{\partial x}\right) \sim J(x^k, h^k),$$

элементами которой являются отношения, указанные выше. Тогда получится следующий метод:

$$J(x^k, h^k) \Delta x^k = -f(x^k), \quad x^{k+1} = x^k + \Delta x^k, \quad k = 0, 1, \dots \quad (13)$$

— дискретный метод Ньютона.

Метод секущих.

Это частный случай дискретного метода. В данном случае мы будем рассматривать разность

$$h^k = x^k - x^{k-1}.$$

Тогда

$$\frac{\partial f_i(x_1^k, \dots, x_n^k)}{\partial x_j} \approx \frac{f_i(x_1^k, \dots, x_j^k, \dots, x_n^k) - f_i(x_1^k, \dots, x_j^{k-1}, \dots, x_n^k)}{x_j^k - x_j^{k-1}}. \quad (14)$$

Таким образом, с учетом последней формулы мы получим метод вида (13), где элементы матрицы $J(x^k, h^k)$ определяются по формулам (14).

1.6.5 Метод градиентного спуска.

На основании исходной системы нелинейных уравнений системы (1) рассмотрим функцию

$$\Phi(x) = \sum_{i=1}^n f_i^2(x_1, \dots, x_n). \quad (15)$$

Функция $\Phi(x)$ неотрицательна и обращается в ноль тогда и только тогда, когда все $f_i \equiv 0$. Таким образом, решение исходной системы нелинейных уравнений будет одновременно нулевым минимумом скалярной функции многих переменных $\Phi(x)$.

По методу градиентного спуска итерационная последовательность к решению определяется формулой

$$x^{k+1} = x^k - t \operatorname{grad} \Phi(x^k), \quad k = 0, 1, \dots, \quad t \geq 0 \quad (16)$$

Параметр t выберем из условия минимума функции Φ в точке x^{k+1} :

$$\Phi(x^{k+1}) = \Phi(x^k - t \operatorname{grad} \Phi(x^k)) = \varphi(t).$$

И будем искать такую функцию, чтобы $\varphi(t)$ была минимальна.

В итоге, решая уравнение

$$\varphi'(t) = 0,$$

находим t на каждой итерации. Таким образом и строится формула (16). Для того, чтобы найти t из этого уравнения, можно использовать любой метод для нахождения корня нелинейной функции одной переменной.

Иногда функцию $\varphi(t)$ бывает сложно посчитать. Условие $\varphi'(t) = 0$ можно заменить эквивалентным ему. Выразим

$$\operatorname{grad} \Phi(x) = 2(f_i, \operatorname{grad} f_i).$$

И пользуясь выражением для градиента, мы можем записать уравнение $\varphi'(t) = 0$ в следующем виде:

$$\sum_{i=1}^n 2f_i(x^k - t \operatorname{grad} \Phi(x^k)) \cdot \frac{d}{dt} f_i(x^k - t \operatorname{grad} \Phi(x^k)) = 0 \quad (17)$$

Для определения t можно использовать один из методов решения нелинейных уравнений.

Замечания.

1. Если решить уравнение относительно t не представляется возможным, то требование $\varphi'(t) = 0$ заменяется на менее жесткое:

$$\Phi(x^{k+1}) < \Phi(x^k).$$

2. Методы спуска сходятся для гладких функций всегда, но зачастую довольно медленно. Но они могут использоваться для получения хорошего начального приближения к решению, чтобы использовать более быстро сходящиеся методы.

Глава 2

Приближение функций.

2.1 Общие положения проблемы приближения функций.

В самом широком смысле проблему приближения функций мы сформулируем следующим образом:

1. имея значения функции $f(x_i)$ в одних точках, найти значения функции в других точках;
2. имея некоторую функцию $f(x)$, которую трудно вычислить, мы будем заменять ее другой функцией $\varphi(x)$, которую легко вычислить, при этом задача связанная с приближением функций состоит в том, чтобы найти такую функцию $\varphi(x)$ и она приближала функцию $f(x)$ с некоторой точностью ε .

Примеры задач в рамках поставленной проблемы:

1. Задачи планирования экспериментов;
2. Задачи обработки данных экспериментов;
3. Задачи вычисления элементарных или специальных функций.