Метод Ньютона

Решаем уравнение $f(x)=0$

Пусть известно, что на [a, b] есть корень x_* уравнения $f(x)=0$, причем $f'(x)$ и $f''(x)$ непрерывны и сохраняют определенные знаки на [a, b].

Пусть x_k – некоторое приближенное значение корня. Можно записать $x_* = x_k + h$, где h – малая величина.

Применяем разложение в ряд Тейлора:

$$f(x_*) = f(x_k + h) \approx f(x_k) + h f'(x_k) = 0,$$

Метод Ньютона

$$h \approx -f(x_k)/f'(x_k)$$

$$x_* \approx x_k - f(x_k)/f'(x_k)$$

Таким образом, следующее приближение к корню будет

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}$$

Это итерационная схема носит название метода Ньютона (или метода *касательных*)

Метод Ньютона

Уравнение касательной:

$$y = y_0 + f'(x_0)(x - x_0)$$

Для y=0 значение x будет равно x_1:

$$0 = y_0 + f'(x_0)(x_1 - x_0)$$

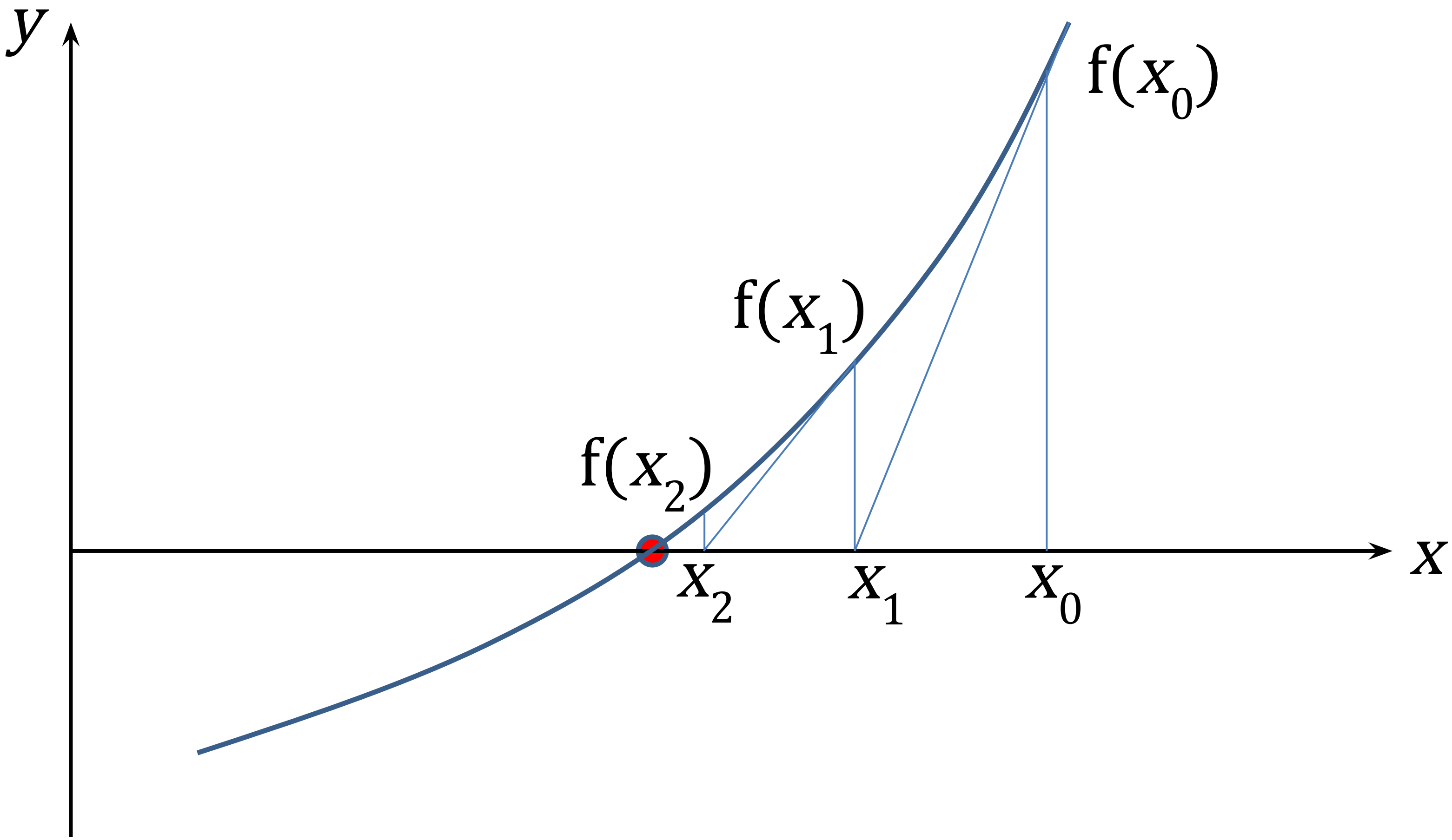
Находим x_1:

$$x_1 = x_0 - \frac{y_0}{f'(x_0)}$$

Т.к. $y_0 \equiv f(x_0)$, то

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}$$

Метод Ньютона



Метод Ньютона

Условие сходимости метода Ньютона

Т.к. $$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}$$

а в методе простых итераций $x_{k+1} = \varphi(x_k)$, то в методе Ньютона можно принять

$$\varphi(x) \equiv x - \frac{f(x)}{f'(x)}$$

Условие сходимости $|\varphi'(x)| \le q < 1$ дает

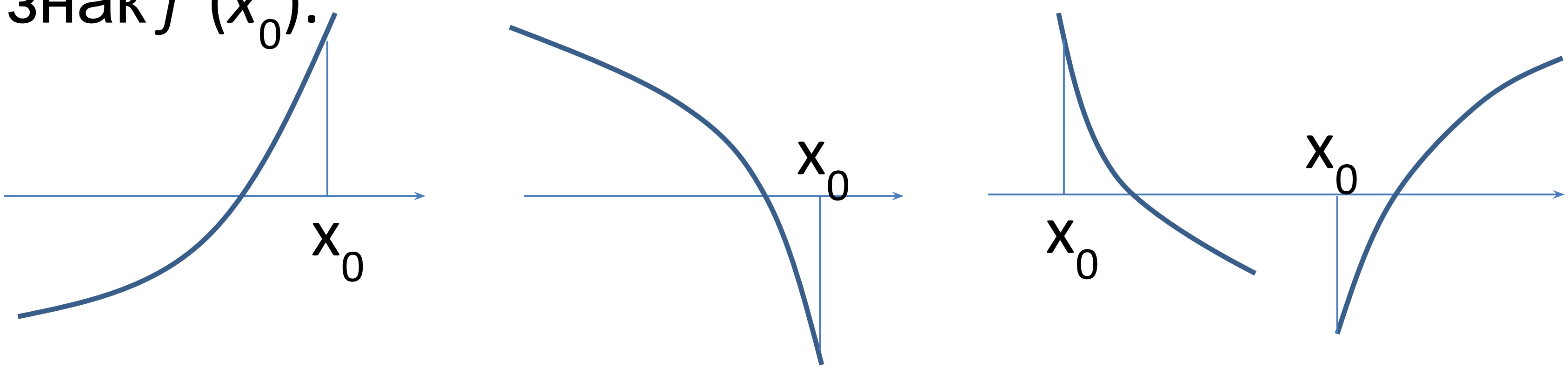
$$|ff''/(f')^2| \le q < 1$$

Метод Ньютона

«Хорошим» начальным приближением x_0 является то, для которого выполняется неравенство

$$f(x_0)\, f''(x_0)>0.$$

Таким образом в качестве исходной точки x_0 можно выбрать тот конец интервала [*a, b*], которому отвечает ордината того же знака, что и знак $f''(x_0)$.



Метод Ньютона

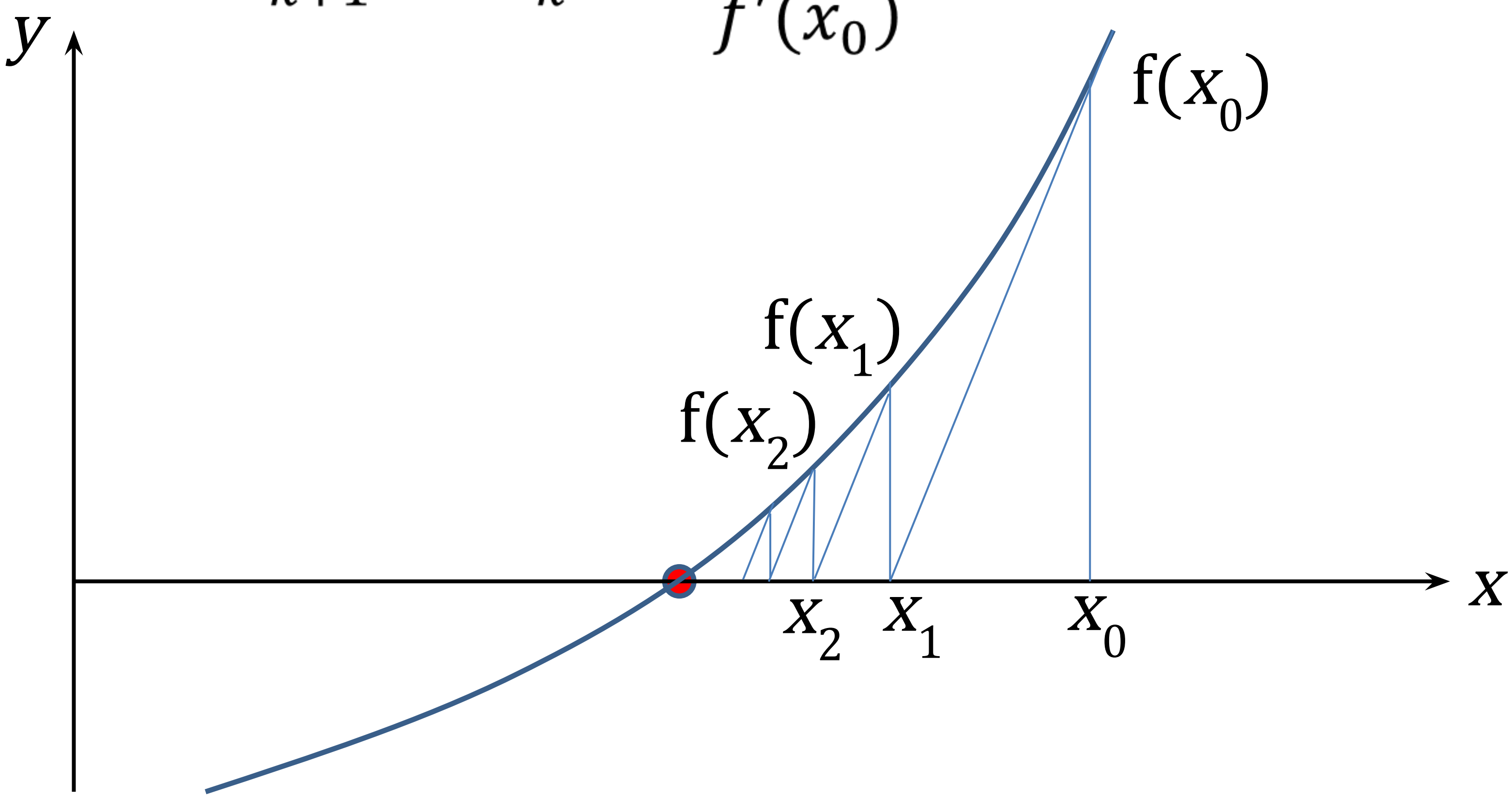
Примеры «плохого» поведения метода
Н

Метод Ньютона

Модификации метода Ньютона

1) Метод «замороженной» производной

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_0)}$$

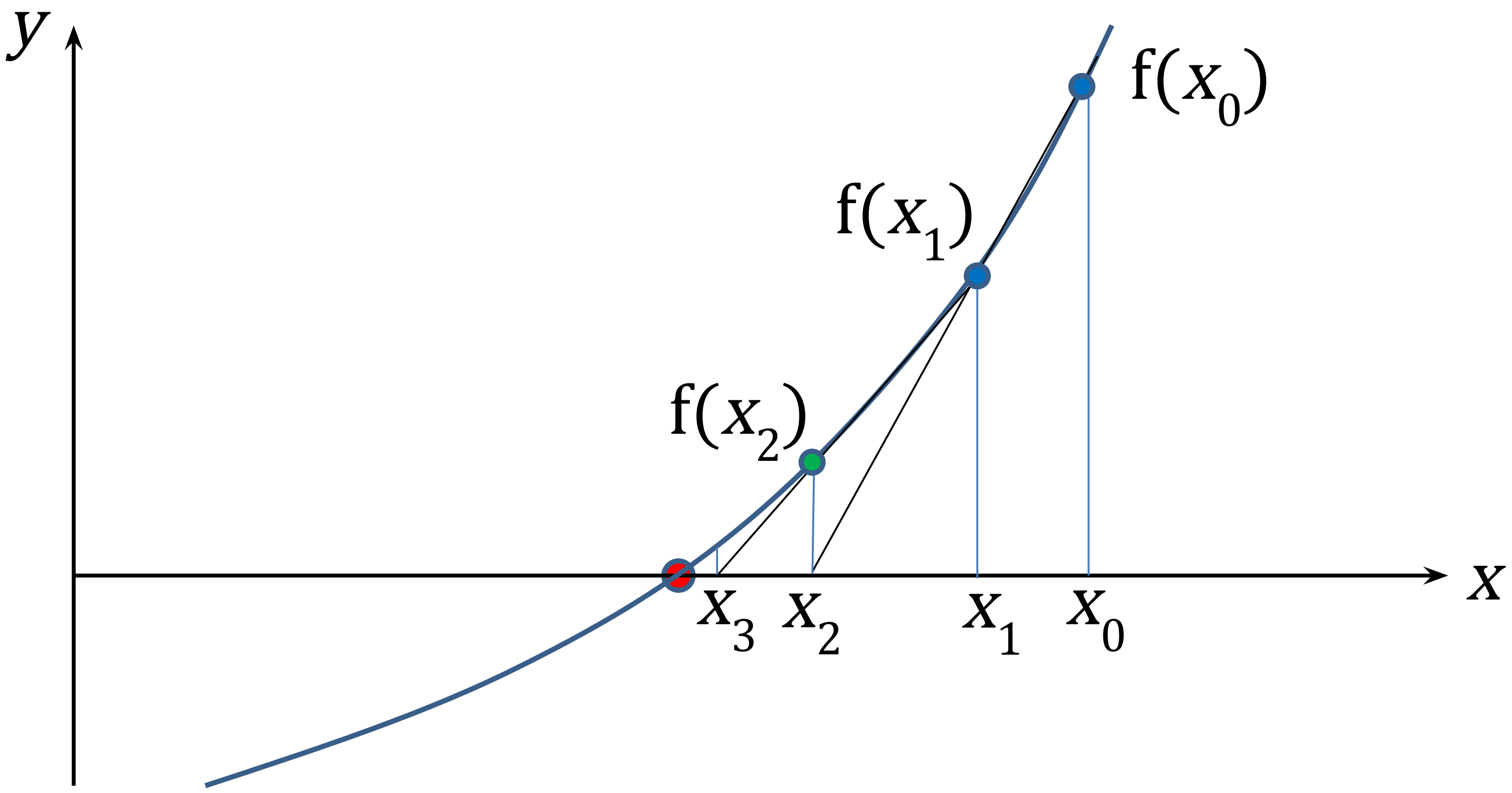


Метод Ньютона

Модификации метода Ньютона

2) Метод секущих

$$x_{k+1} = x_k - \frac{x_k - x_{k-1}}{f(x_k) - f(x_{k-1})} f(x_k)$$



Метод Чебышева построения итераций высших порядков

Пусть уравнение $y = f(x) = 0$ на $[a, b]$ имеет корень x_*.

Пусть $f'(x) \neq 0$. Функция $y = f(x)$ имеет обратную функцию $x = F(y)$ и $x \equiv F(f(x))$.

Тогда

$$x_* = F(0) = F(y) + \frac{F'(y)}{1!}(0-y)^1 + \frac{F''(y)}{2!}(0-y)^2 + \cdots \approx$$

$$\approx F(y) + \sum_{j=1}^{n} (-1)^j \frac{F^{(j)}(y)}{j!} f^j(x)$$

Метод Чебышева построения итераций высших порядков

Можно записать, что

$$x_* \approx x + \sum_{j=1}^{n} (-1)^j \frac{F^{(j)}(y)}{j!} f^j(x)$$

Следовательно, итерационная схема будет иметь вид

$$x_{k+1} = x_k + \sum_{j=1}^{n} (-1)^j \frac{F^{(j)}(f(x_k))}{j!} f^j(x_k)$$

Найдем производные от обратной функции $F^{(j)}(f(x_k))$

Метод Чебышева построения итераций высших порядков

Дифференцируем соотношение $x = F(y) \equiv F(f(x))$

$$1 = F'(y)f'(x) \quad \rightarrow \quad \boxed{F'(y) = 1/f'(x)}$$

Дифференцируем соотношение еще раз

$$0 = F''(y)f'(x)f'(x) + F'(y)f''(x) =$$

$$= F''(y)[f'(x)]^2 + f''(x)/f'(x)$$

Метод Чебышева построения итераций высших порядков

$$\boxed{F''(y) = -f''(x)/[f'(x)]^3}$$

И т.д.

Частные случаи

j=1 $$x_{k+1} = x_k + (-1)^1 \frac{F^{(1)}(f(x_k))}{1!} f^1(x_k)$$

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}$$ – формула Ньютона

Метод Чебышева построения итераций высших порядков

j=2

$$x_{k+1} = x_k + \sum_{j=1}^{2} (-1)^j \frac{F^{(j)}(f(x_k))}{j!} f^j(x_k) =$$

$$= x_k + (-1)^1 \frac{F^{(1)}(f(x_k))}{1!} f^1(x_k) + (-1)^2 \frac{F^{(2)}(f(x_k))}{2!} f^2(x_k)$$

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)} - \frac{f''(x_k) \cdot f^2(x_k)}{2[f'(x)]^3}$$

Скорость сходимости методов

$$\left|X_{k+1} - X_*\right| = q\left|X_k - X_*\right|^{\alpha}$$

Название метода	q	α
Метод половинного деления	0.5	1
Метод хорд	$1-\dfrac{m_1}{M_1}$	1
Метод простых итераций	$\left|\varphi'(x)\right|$	1
Метод Ньютона	$\dfrac{M_2}{2m_1}$	2
Метод Чебышева	$\dfrac{Q_3}{3!}M_1^3$	3

$$m_1 = \min\left|f'(x)\right|$$

$$M_1 = \max\left|f'(x)\right|$$

$$M_2 = \max\left|f''(x)\right|$$

$$Q_3 = \max\left|F'''(y)\right|$$

Нахождение корней полиномов

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

Рассмотренные выше методы решения нелинейных трансцендентных уравнений пригодны и для алгебраических уравнений.

Особенности нахождения корней полиномов.

1) Неустойчивость корней некоторых полиномов как функций от их коэффициентов, что предъявляет особые требования к точности вычислений.

Нахождение корней полиномов

2) Полином степени n имеет ровно n корней, действительных или комплексных, при условии, что каждый корень считается столько раз, какова его кратность.

3) Комплексные корни образуют комплексно-сопряженные пары, т.е. каждому корню $x = c + id$ соответствует $x = c - id$.

4) Известны оценки границ корней полинома:

$$r < |x_k| < R$$

Нахождение корней полиномов

$$R = 1 + \frac{max\{|a_{n-1}|, |a_{n-2}|, \dots, |a_0|\}}{|a_n|}$$

$$r = \frac{1}{1 + \frac{max\{|a_n|, |a_{n-1}|, \dots, |a_1|\}}{|a_0|}}$$

1) Это завышенные оценки.
2) В связи с малостью *r* на практике пользуются оценкой $|x_k| < R$.
3) Эти оценки справедливы и для комплексных корней.

Нахождение корней полиномов

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

Теорема Лагранжа.

Пусть $a_n > 0$ и $0 \le k \le n-1$ – номер первого отрицательного коэффициента полинома $P_n(x)$, если просматривать их в порядке убывания степени x.

Тогда за верхнюю границу положительных корней уравнения $P_n(x)=0$ может быть взято число $R = 1 + \sqrt[n-k]{\frac{C}{a_n}}$.

C – наибольшая из абсолютных величин отрицательных коэффициентов полинома $P_n(x)$.

Если отрицательных коэффициентов нет, то $P_n(x)$ не имеет положительных корней.

Нахождение корней полиномов

Как быть, если рассматриваемое уравнение имеет ***комплексные корни***?

1) Если все значения функции вещественны и в качестве начального приближения x_0 также взято ***вещественное*** число, то и все последующие значения x_k также будут вещественны.

2) Однако если взять в качестве x_0 ***комплексное*** число, то последующие x_k ***могут*** оказаться комплексными.

3) Если найден один комплексный корень $x_1 = c + id$, то всегда будет существовать сопряженный ему корень!

Нахождение корней полиномов

Если многочлен $P_n(x)$ имеет пару комплексно-сопряженных корней $x_{1,2} = c \pm id$, то его можно записать в следующем виде:

$P_n(x) = (x - x_1)(x - x_2)\, P_{n-2}(x)$

или $P_n(x) = ((x - c)^2 + d^2) P_{n-2}(x)$

Дальше можно решать уравнение $P_{n-2}(x) = 0$ вместо $P_n(x) = 0$.

Решение систем нелинейных уравнений

В координатном виде система нелинейных уравнений имеет вид:

$$\begin{cases} f_1(x_1, x_2, \dots, x_n) = 0 \\ f_2(x_1, x_2, \dots, x_n) = 0 \\ \dots\dots\dots\dots\dots\dots \\ f_n(x_1, x_2, \dots, x_n) = 0 \end{cases}$$

или

$$f_i(x_1, x_2, \dots, x_n) = 0, \qquad i = 1, \dots n$$

В векторном виде $\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{0}$,

$$\boldsymbol{f} = \{f_1, f_2, \dots, f_n\}^T; \qquad \boldsymbol{x} = \{x_1, x_2, \dots, x_n\}$$

Метод простых итераций

Представим исходную систему уравнений в виде:

$$\begin{cases} x_1 = \varphi_1(x_1, x_2, \dots, x_n) \\ x_2 = \varphi_2(x_1, x_2, \dots, x_n) \\ \dots\dots\dots\dots\dots\dots \\ x_n = \varphi_n(x_1, x_2, \dots, x_n) \end{cases}$$

В векторном виде: $\boldsymbol{x} = \varphi(\boldsymbol{x})$

Метод простых итераций записывается в виде:

$$\boldsymbol{x}^{(k+1)} = \varphi\left(\boldsymbol{x}^{(k)}\right)$$

$\boldsymbol{x}^{(0)}$ – вектор начального приближения

Достаточное условие сходимости: $\|\varphi'(\boldsymbol{x})\| \leq \mathrm{q} < 1$

Метод простых итераций

Матрица Якоби:

$$\varphi'(\boldsymbol{x}) = \left\{\frac{\partial\varphi_i}{\partial x_j}\right\} = \begin{pmatrix} \frac{\partial\varphi_1}{\partial x_1} & \cdots & \frac{\partial\varphi_1}{\partial x_n} \\ \cdots & \cdots & \cdots \\ \frac{\partial\varphi_n}{\partial x_1} & \cdots & \frac{\partial\varphi_n}{\partial x_n} \end{pmatrix}$$

В норме Чебышева:

$$\|\varphi'(\boldsymbol{x})\| = \max_{1\le i\le n} \sum_{j=1}^{n} \left|\frac{\partial\varphi_i}{\partial x_j}\right|$$

$$\|\boldsymbol{x}\| = \max_{1\le i\le n} |x_i|$$

Метод простых итераций

В норме Евклида:

$$\|\varphi'(\boldsymbol{x})\| = \sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n}\left(\frac{\partial\varphi_i}{\partial x_j}\right)^2} \qquad \|\boldsymbol{x}\| = \sqrt{\sum_{i=1}^{n}(x_i)^2}$$

Оценка погрешности метода (используемая для прекращения процесса итерации) имеет вид:

$$\left\|\boldsymbol{x}^{(k)} - \boldsymbol{x}^{(k-1)}\right\| \le \frac{1-q}{q}\varepsilon$$

Модифицированный метод простых итераций

$$
\begin{cases}
x_1^{(k+1)} = \varphi_1\left(x_1^{(k)}, x_2^{(k)}, \ldots, x_{n-1}^{(k)}, x_n^{(k)}\right) \\
x_2^{(k+1)} = \varphi_2\left(x_1^{(k+1)}, x_2^{(k)}, \ldots, x_{n-1}^{(k)}, x_n^{(k)}\right) \\
\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
x_n^{(k+1)} = \varphi_n\left(x_1^{(k+1)}, x_2^{(k+1)}, \ldots, x_{n-1}^{(k+1)}, x_n^{(k)}\right)
\end{cases}
$$

Скорость сходимости выше, чем в методе простых итераций.

Метод Ньютона

Этот метод обладает гораздо ***более быстрой*** сходимостью, чем метод простых итераций.

Недостатки:

1) Очень важен удачный выбор начального приближения для обеспечения хорошей сходимости;
2) Сходимость ухудшается с увеличением числа уравнений системы.

Метод Ньютона

Решаем систему уравнений:

$$\begin{cases} f_1(x_1, x_2, \dots, x_n) = 0 \\ f_2(x_1, x_2, \dots, x_n) = 0 \\ \dots\dots\dots\dots\dots\dots \\ f_n(x_1, x_2, \dots, x_n) = 0 \end{cases}$$

или

в векторном виде $\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{0}$,

Метод Ньютона

Метод Ньютона можно применить, если могут быть вычислены все частные производные функций f_i по переменным x_j: $\frac{\partial f_i}{\partial x_j}$

Процесс начинается с задания произвольного начального приближения $\boldsymbol{x}^{(0)}$.

Далее разлагаем функцию f в ряд Тейлора в окрестности точки $\boldsymbol{x}^{(0)}$:

$$\boldsymbol{f}(\boldsymbol{x})=\boldsymbol{f}\left(\boldsymbol{x}^{(0)}\right)+J\left(\boldsymbol{x}^{(0)}\right)\left(\boldsymbol{x}-\boldsymbol{x}^{(\mathbf{0})}\right)+\cdots$$

Метод Ньютона

$J(\boldsymbol{x})$ – матрица Якоби, составленная из частных производных:

$$J(\boldsymbol{x}) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_n} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial f_n}{\partial x_1} & \frac{\partial f_n}{\partial x_2} & \cdots & \frac{\partial f_n}{\partial x_n} \end{pmatrix}$$

Метод Ньютона

Ограничиваясь линейным приближением, имеем: $\boldsymbol{f}(\boldsymbol{x}) \approx \boldsymbol{f}\left(\boldsymbol{x}^{(0)}\right) + J\left(\boldsymbol{x}^{(0)}\right)\left(\boldsymbol{x} - \boldsymbol{x}^{(0)}\right)$

Предположим, что $\boldsymbol{f}(\boldsymbol{x})=0$, тогда

$$\boldsymbol{f}\left(\boldsymbol{x}^{(0)}\right) + J\left(\boldsymbol{x}^{(0)}\right)\left(\boldsymbol{x} - \boldsymbol{x}^{(0)}\right) \approx \mathbf{0}$$

Чтобы найти следующее приближение $\boldsymbol{x}^{(1)}$ к решению системы $\boldsymbol{f}(\boldsymbol{x})=\mathbf{0}$, решаем систему уравнений:

$$\boldsymbol{f}\left(\boldsymbol{x}^{(0)}\right) + J\left(\boldsymbol{x}^{(0)}\right)\left(\boldsymbol{x}^{(1)} - \boldsymbol{x}^{(0)}\right) = \mathbf{0}$$

Метод Ньютона

Для нахождения приближения $\boldsymbol{x}^{(k)}$, решаем систему уравнений:

$$\boldsymbol{f}\left(\boldsymbol{x}^{(k)}\right) + J\left(\boldsymbol{x}^{(k)}\right)\left(\boldsymbol{x}^{(k+1)} - \boldsymbol{x}^{(k)}\right) = \boldsymbol{0}$$

Эту систему можно переписать для поправки $\Delta\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k+1)} - \boldsymbol{x}^{(k)}$

$$\boxed{\begin{gathered} J\left(\boldsymbol{x}^{(k)}\right)\Delta\boldsymbol{x}^{(k+1)} = -\boldsymbol{f}\left(\boldsymbol{x}^{(k)}\right) \\[2em] \boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} + \Delta\boldsymbol{x}^{(k+1)} \end{gathered}}$$

Метод Ньютона

Решение можно записать и в другом виде – непосредственно для $\boldsymbol{x}^{(k+1)}$:

$$\boxed{\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} - \left[J\left(\boldsymbol{x}^{(k)}\right)\right]^{-1} \boldsymbol{f}\left(\boldsymbol{x}^{(k)}\right)}$$

Систему линейных уравнений, как правило, решать проще, чем искать обратную матрицу

$$\left[J\left(\boldsymbol{x}^{(k)}\right)\right]^{-\mathbf{1}}.$$

Еще одно ограничение, накладываемое на метод Ньютона: ***определитель*** матрицы Якоби должен быть ***отличен от нуля***.

Метод Ньютона для решения системы 2-х уравнений

Рассматривается система уравнений:

$$\begin{cases} f_1(x_1, x_2) = 0 \\ f_2(x_1, x_2) = 0 \end{cases}$$

Матрица Якоби:

$$J(\boldsymbol{x}) = \begin{pmatrix} \dfrac{\partial f_1}{\partial x_1} & \dfrac{\partial f_1}{\partial x_2} \\ \dfrac{\partial f_2}{\partial x_1} & \dfrac{\partial f_2}{\partial x_2} \end{pmatrix}$$

Обратная ей матрица: $\left[J\left(\boldsymbol{x}^{(k)}\right)\right]^{-\mathbf{1}} = \dfrac{1}{|J|}\tilde{A}$

Метод Ньютона для решения системы 2-х уравнений

$\tilde{A}$ – присоединенная матрица, составленная из алгебраических дополнений матрицы $J(\boldsymbol{x})$

$$\tilde{A} = \begin{pmatrix} \dfrac{\partial f_2}{\partial x_2} & -\dfrac{\partial f_1}{\partial x_2} \\ -\dfrac{\partial f_2}{\partial x_1} & \dfrac{\partial f_1}{\partial x_1} \end{pmatrix}$$

$$|J| = \frac{\partial f_1}{\partial x_1}\frac{\partial f_2}{\partial x_2} - \frac{\partial f_1}{\partial x_2}\frac{\partial f_2}{\partial x_1}$$

Метод Ньютона для решения системы 2-х уравнений

Уравнение

$$\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} - \left[J\left(\boldsymbol{x}^{(k)}\right)\right]^{-1} \boldsymbol{f}\left(\boldsymbol{x}^{(k)}\right)$$

записывается в виде:

$$\binom{x_1}{x_2}^{(k+1)} = \binom{x_1}{x_2}^{(k)} - \frac{1}{|J^{(k)}|} \begin{pmatrix} \dfrac{\partial f_2}{\partial x_2} & -\dfrac{\partial f_1}{\partial x_2} \\ -\dfrac{\partial f_2}{\partial x_1} & \dfrac{\partial f_1}{\partial x_1} \end{pmatrix}^{(k)} \binom{f_1}{f_2}^{(k)}$$

Метод Ньютона для решения системы 2-х уравнений

Для каждой переменной уравнения записываются в виде:

$$x_1{}^{(k+1)} = x_1{}^{(k)} - \frac{1}{|J^{(k)}|}\left(f_1\frac{\partial f_2}{\partial x_2} - f_2\frac{\partial f_1}{\partial x_2}\right)^{(k)}$$

$$x_2{}^{(k+1)} = x_2{}^{(k)} - \frac{1}{|J^{(k)}|}\left(f_2\frac{\partial f_1}{\partial x_1} - f_1\frac{\partial f_2}{\partial x_1}\right)^{(k)}$$