Лекция 6

§ 4.5. Метод Ньютона

Знаменитый метод Ньютона является одним из наиболее эффективных методов решения самых разных нелинейных задач. Расчетную формулу метода можно получить, используя различные подходы. Рассмотрим два из них.

1.Метод касательных. Выведем расчетную формулу метода для решения нелинейного уравнения (4.1) из простых геометрических соображений. Соответствующая иллюстрация приведена на рис. 4.3.

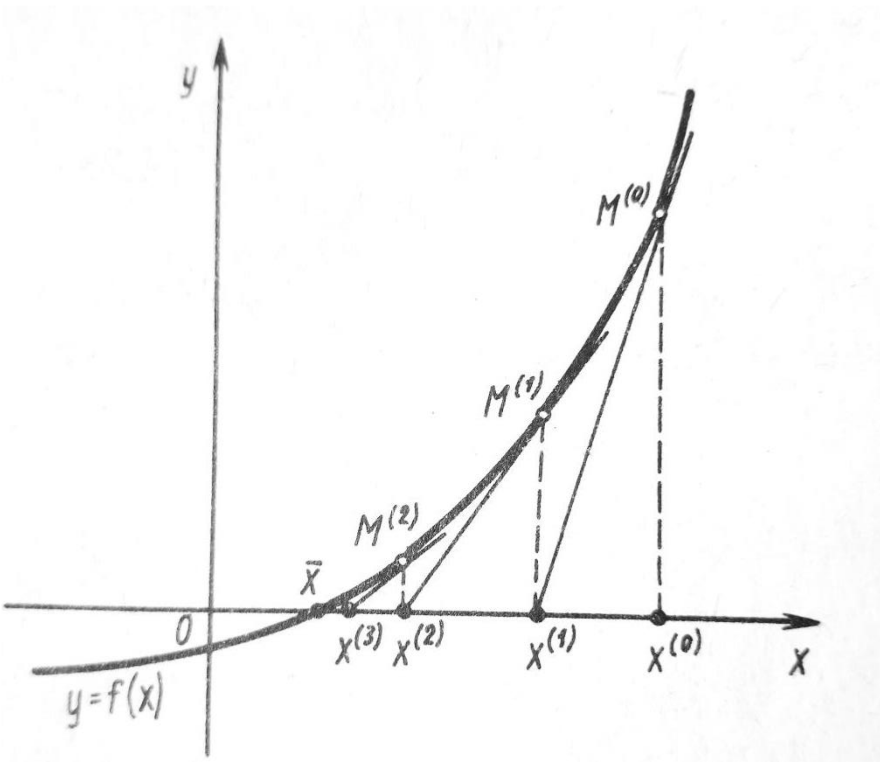

Рис. 4.3

Пусть $x^{(0)}$ – заданное начальное приближение к корню $\bar{x}$. В точке $M^{(0)}$ с координатами $(x^{(0)}, f(x^{(0)}))$ проведем касательную к графику функции $y = f(x)$ и за новое приближение $x^{(1)}$ примем абсциссу точки пересечения этой касательной с осью Ox. Аналогично за приближение $x^{(2)}$ примем абсциссу точки пересечения с осью Ox касательной, проведенной к графику в точке $M^{(1)}$ с координатами $(x^{(1)}, f(x^{(1)}))$.

Продолжая этот процесс далее, получим последовательность $x^{(0)}, x^{(1)}, x^{(2)}, ..., x^{(n)}, ...$ приближений к корню $\bar{x}$.

Напомним, что уравнение касательной, проведенной к графику функции $y = f(x)$ в точке $(x^{(n)}, f(x^{(n)}))$, имеет вид

$$y = f(x^{(n)}) + f'(x^{(n)})(x - x^{(n)}). \tag{4.27}$$

Полагая в равенстве (4.27) $y = 0$, замечаем, что при выполнении условия $f'(x^{(n)}) \neq 0$ абсцисса $x^{(n+1)}$ точки пересечения касательной с осью *Ox* удовлетворяет равенству

$$0 = f(x^{(n)}) + f'(x^{(n)})(x^{(n+1)} - x^{(n)}). \tag{4.28}$$

Выражая из него $x^{(n+1)}$, получаем расчетную формулу *метода Ньютона:*

$$x^{(n+1)} = x^{(n)} - \frac{f(x^{(n)})}{f'(x^{(n)})}, \; n \geq 0. \tag{4.29}$$

Благодаря такой геометрической интерпретации этот метод часто называют *методом касательных*.

2.Метод линеаризации. С более общих позиций метод Ньютона можно рассматривать как итерационный метод, использующий специальную линеаризацию задачи и позволяющий свести решение исходного нелинейного уравнения к решению последовательности линейных уравнений.

Пусть приближение $x^{(n)}$ уже получено. Представим функцию в окрестности точки $x^{(n)}$ по формуле Тейлора:

$$f(x) = f(x^{(n)}) + f'(x^{(n)})(x - x^{(n)}) + \frac{f''(\xi)}{2}(x - x^{(n)})^2. \tag{4.30}$$

Здесь ξ – некоторая точка, расположенная между x и $x^{(n)}$. Заменяя в уравнении $f(x) = 0$ функцию $f(x)$ главной линейной частью разложения (4.30), получим линейное уравнение

$$f(x^{(n)}) + f'(x^{(n)})(x - x^{(n)}) = 0. \tag{4.31}$$

Принимая решение уравнения (4.31) за новое приближение $x^{(n+1)}$, приходим к формуле (4.29).

3.Основная теорема о сходимости метода Ньютона.

Т е о р е м а 4.4. Пусть $\bar{x}$ *– простой корень уравнения f(x) = 0, в некоторой окрестности которого функция f дважды непрерывно дифференцируема. Тогда найдется такая малая ε-окрестность корня* $\bar{x}$*, что при произвольном выборе начального приближения* $x^{(0)}$ *из этой окрестности итерационная последовательность метода Ньютона не выходит за пределы этой окрестности и справедлива оценка*

$$|x^{(n+1)} - \bar{x}| \leq C|x^{(n)} - \bar{x}|^2, \; n \geq 0, \tag{4.32}$$

где C = σ^{-1}*, означающая, что метод сходится с квадратичной скоростью.*

Следствием оценки (4.32) является априорная оценка погрешности

$$|x^{(n)} - \bar{x}| \leq \sigma q^{2^n}, \; n \geq 0, \tag{4.33}$$

в которой $q = \sigma^{-1} \mid x^{(0)} - \bar{x}|$.

□ Так как $f'(\bar{x}) \neq 0$ (по определению простого корня), то в силу непрерывности функций f' и f'' найдется δ_0-окрестность корня, в которой при некоторых постоянных α и β выполнены неравенства
$0 < \alpha \leq |f'(x)|, |f''(x)| \leq \beta$.

Пусть $x^{(n)} \in (\bar{x} - \boldsymbol{\sigma}, \bar{x} + \boldsymbol{\sigma})$, где $\boldsymbol{\sigma} = min\, \{\delta_0, \frac{2\alpha}{\beta}\}$. Подставляя $x = \bar{x}$ в (4.30), получим равенство

$$0 = f(x^{(n)}) + f'(x^{(n)})(\bar{x} - x^{(n)}) + \frac{f''(\xi)}{2}(\bar{x} - x^{(n)})^2,$$ в котором

$\xi \in (\bar{x} - \boldsymbol{\sigma}, \bar{x} + \boldsymbol{\sigma})$. Вычитая из него равенство (4.28), имеем

$$f'(x^{(n)})(x^{(n+1)} - \bar{x}) = \frac{f''(\xi)}{2}(\bar{x} - x^{(n)})^2.$$

Тогда, приравняв модули обеих частей этого равенства и используя условия ограниченности $|f'(x)|$ и $|f''(x)|$, приходим к неравенству $\alpha|x^{(n+1)} - \bar{x} \mid \leq \frac{\beta}{2}(\bar{x} - x^{(n)})^2$, откуда следует справедливость оценки (4.32). Справедливость оценки (4.33) устанавливается методом индукции. ■

Таким образом, при выборе начального приближения из достаточно малой окрестности корня метод Ньютона сходится квадратично. Это означает, грубо говоря, что на каждой итерации число верных цифр приближения примерно удваивается.

Приведенные в теореме 4.4 оценки погрешности являются априорными и их использование в практике вычислений невозможно.

4.Критерий окончания. На практике предпочтительнее пользоваться простой апостериорной оценкой

$$|x^{(n)} - \bar{x}| \le |x^{(n)} - x^{(n-1)}|, \qquad (4.34)$$

справедливость которой обосновывается следующим утверждением.

Т е о р е м а 4.5. Пусть выполнены условия теоремы 4.4 и $x^{(0)} \in (\bar{x} - \boldsymbol{\sigma}/2, \bar{x} + \boldsymbol{\sigma}/2)$. *Тогда для всех* $n \ge 1$ *верна оценка (4.34).*

□ Из оценки (4.33) следует, что $(x^{(n-1)} - \bar{x}) \le \sigma q^{2^{n-1}} \le \boldsymbol{\sigma} q = |x^{(0)} - \bar{x}| < \boldsymbol{\sigma}/2$. Поэтому, применяя неравенство (4.32), получим цепочку неравенств $2\,|x^{(n)} - \bar{x}| \le 2\,\sigma^{-1}\,|x^{(n-1)} - \bar{x}|^2 \le |x^{(n-1)} - \bar{x}| \le |x^{(n-1)} - x^{(n)}| + |x^{(n)} - \bar{x}|$, из которой вытекает оценка (4.34). ■

Наличие оценки (4.34) позволяет сформулировать следующий практический критерий окончания итераций метода Ньютона. При заданной точности ε > 0 вычисления нужно вести до тех пор, пока не окажется выполненным неравенство

$$|x^{(n)} - x^{(n-1)}| < \varepsilon \qquad (4.35)$$

Теоремы 4.4 и 4.5 утверждают лишь о существовании малой окрестности корня, в которой имеет место сходимость метода Ньютона. Однако о критериях выбора такой окрестности речь в них не идет. Приведем без доказательства упрощенный вариант теоремы об условиях, предъявляемых к выбору отрезка локализации корня, поведению функции на нем и выбору начального приближения для гарантированной сходимости метода Ньютона:

Т е о р е м а 4.6. Пусть $[a, b]$ *- отрезок локализации простого корня* $\bar{x}$ *уравнения* $f(x) = 0$. *Предположим, что на этом отрезке функция* f *дважды непрерывно дифференцируема, а ее производные* $f'(x)$ и $f''(x)$ *знакопостоянны.*

Тогда начиная с n = 1 *итерационная последовательность метода Ньютона* $x^{(n)}$ *сходится к* $\bar{x}$ *монотонно с той стороны отрезка* $[a, b]$, *где* $f(x)\, f''(x) > 0$.

Иллюстрацией монотонного характера сходимости может служить рис. 4.3.

С л е д с т в и е. *Пусть* $\bar{x}$ *- корень уравнения* $f(x) = 0$, *функция* f *дважды непрерывно дифференцируема на всей числовой оси, а ее производные* $f'(x)$ и $f''(x)$ *знакопостоянны. Тогда метод Ньютона сходится при любом начальном приближении* $x^{(0)}$ *(т. е. является глобально сходящимся), причем начиная с* n=1 *итерационная последовательность* $x^{(n)}$ *сходится к* $\bar{x}$ *монотонно с той стороны от корня, где* $f(x)\, f''(x) > 0$.

Таким образом, метод Ньютона обладает в общем случае только *локальной сходимостью.* Это означает, что областью его сходимости является некоторая малая **σ**-окрестность корня, в которой функция должна быть гладкой (производные первого и второго порядка знакопостоянны), и для гарантии сходимости необходимо правильно выбирать начальное приближение.

Для практического применения метода Ньютона необходимо на каждом шаге вычислять производную $f'\left(x^{(n)}\right)$, однако часто бывает невозможно найти аналитическое выражение для f', а определить приближенное значение с высокой точностью либо очень трудно, либо очень трудоемко делать это на каждом шаге. В этих случаях приходится модифицировать метод, избегая непосредственного вычисления производной. Рассмотрим некоторые из таких модификаций.

§ 4.6. Модификации метода Ньютона

Рассматриваемые ниже итерационные методы решения нелинейного уравнения на каждой итерации используют некоторую процедуру его линеаризации, т.е. исходное нелинейное уравнение заменяется приближенно более простым линейным уравнением.

1.Упрощенный метод Ньютона. Если производная $f'(x)$ непрерывна, то ее значение вблизи простого корня $\bar{x}$ почти постоянно. Поэтому можно вычислить f' лишь однажды в точке $x^{(0)}$, а затем заменить в формуле (4.29) значение $f'\left(x^{(n)}\right)$ постоянной $f'\left(x^{(0)}\right)$. В результате получим расчетную формулу *упрощенного метода Ньютона:*

$$x^{(n+1)} = x^{(n)} - \frac{f(x^{(n)})}{f'(x^{(0)})}, \quad n \geq 0. \qquad (4.36)$$

Геометрическая иллюстрация метода приведена на рис. 4.4.

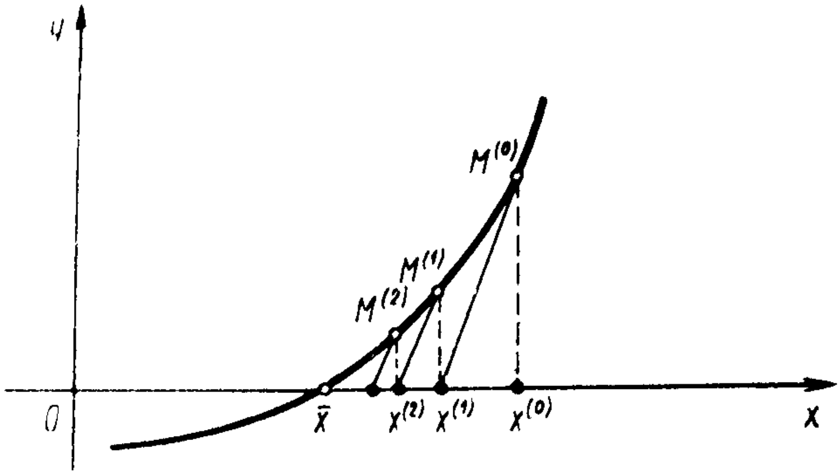

Рис. 4.4

В точке $(x^{(0)}, f(x^{(0)}))$ к графику функции $y = f(x)$ проводится касательная l_0 и за приближение $x^{(1)}$ принимается абсцисса точки пересечения этой касательной с осью *Ox* (как в методе Ньютона). Каждое следующее приближение $x^{(n+1)}$ получается здесь как абсцисса точки пересечения с осью *Ox* прямой, проходящей через точку M^n с координатами $(x^{(n)}, f(x^{(n)}))$ и параллельной касательной l_0.

Упрощение вычислений по сравнению с методом Ньютона достигается здесь ценой резкого падения скорости сходимости. Сходимость этого метода является уже не квадратичной, а линейной.

Метод (4.36) можно рассматривать как метод простой итерации с итерационной функцией $\varphi(x) = x - \frac{f(x)}{f'(x^{(0)})}$.

Так как $\varphi'(x) = 1 - \frac{f'(x)}{f'(x^{(0)})}$, то для знаменателя q соответствующей геометрической прогрессии имеем

$$q \approx \left| 1 - \frac{f'(\bar{x})}{f'(x^{(0)})} \right|.$$

Следовательно, скорость сходимости тем выше, чем ближе начальное приближение $x^{(0)}$ к решению $\bar{x}$.

2.Метод хорд. В основе этой и следующей модификаций лежит приближенное равенство

$$f'\left(x^{(n)}\right) \approx \frac{f\left(z^{(n)}\right)-f\left(x^{(n)}\right)}{z^{(n)}-x^{(n)}}. \qquad (4.37)$$

Оно верно при условии $z^{(n)} \approx x^{(n)}$ и следует из определения производной: $f'(x) = \lim_{z \to x} \frac{f(z)-f(x)}{z-x}$.

Пусть c – фиксированная точка, расположенная в окрестности простого корня $\bar{x}$. Заменим в расчетной формуле метода Ньютона производную $f'\left(x^{(n)}\right)$ правой частью приближенного равенства (4.37), полагая $z^{(n)} = c$. В результате придем к расчетной формуле *метода хорд:*

$$x^{(n+1)} = x^{(n)} - \frac{c - x^{(n)}}{f(c)-f(x^{(n)})} \cdot f(x^{(n)}, \ n \geq 0 \qquad (4.38)$$

Геометрическая иллюстрация метода приведена на рис. 4.5.

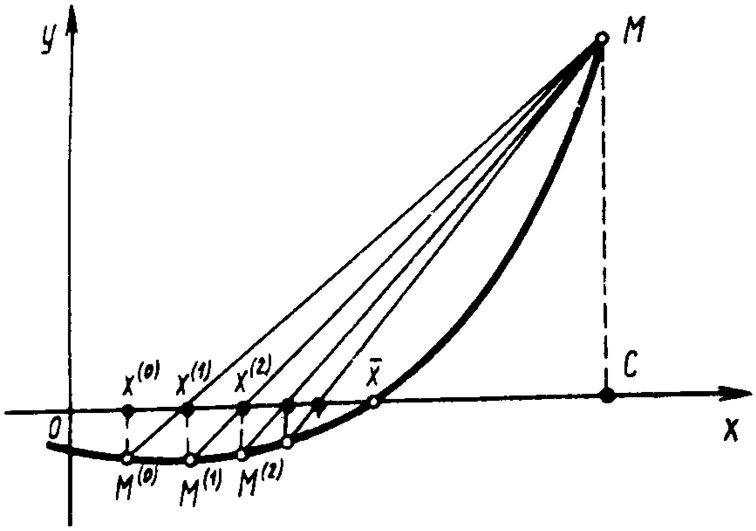

Рис. 4.5

Очередное приближение $x^{(n+1)}$ получается здесь как абсцисса точки пересечения с осью *Ox* прямой, проходящей через расположенные на графике функции *y = f(x)* точки *M* и $M^{(n)}$ с координатами *(c, f(c))* и $(x^{(n)}, f(x^{(n)}))$.

Метод (4.38) обладает только линейной сходимостью. Его можно рассматривать как метод простой итерации с итерационной функцией

$\varphi(x) = x - \frac{c-x}{f(c)-f(x)} f(x)$. Так как скорость сходимости определяется вблизи корня величиной $q \approx |\varphi'(x)| = |\, 1 - \frac{(c-\bar{x})f'(\bar{x})}{f(c)-f(\bar{x})} \,|$, то она тем выше, чем ближе окажется выбранная точка c к $\bar{x}$.

3.Метод секущих. Замена в формуле метода Ньютона производной $f'\left(x^{(n)}\right)$ приближением $\frac{f\left(x^{(n-1)}\right)-f\left(x^{(n)}\right)}{x^{(n-1)}-x^{(n)}}$ приводит к расчетной формуле *метода секущих*:

$$x^{(n+1)} = x^{(n)} - \frac{x^{(n-1)}-x^{(n)}}{f\left(x^{(n-1)}\right)-f\left(x^{(n)}\right)} f\left(x^{(n)}\right), \quad n \geq 1. \qquad (4.39)$$

Заметим, что этот метод двухшаговый, так как для нахождения очередного приближения $x^{(n+1)}$ требуется знание двух предыдущих приближений $x^{(n)}$ и $x^{(n-1)}$. В частности, для того чтобы начать вычисления, необходимо задать два начальных приближения $x^{(0)}$ и $x^{(1)}$.

На рис. 4.6 приведена геометрическая иллюстрация метода.

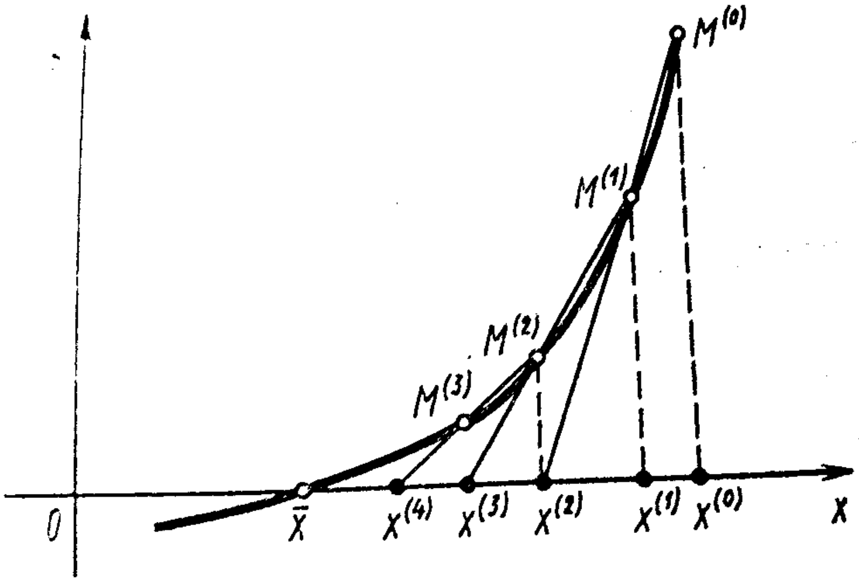

Рис. 4.6

Очередное приближение $x^{(n+1)}$ получается здесь как абсцисса точки пересечения с осью *Ox* секущей, соединяющей точки $M^{(n-1)}$ и $M^{(n)}$ графика функции *f(x)* с координатами $(x^{(n-1)}, f\left(x^{(n-1)}\right))$ и $(x^{(n)}, f\left(x^{(n)}\right))$.

Примечательно, что эта модификация метода Ньютона сохраняет сверхлинейную скорость сходимости, если вычисляется простой корень $\bar{x}$. Точнее, верно следующее утверждение.

Т е о р е м а 4.7. *Пусть $\bar{x}$ – простой корень уравнения f(x) = 0, в некоторой окрестности которого функция f дважды непрерывно дифференцируема, причем* $f''(\bar{x}) \neq 0$. *Тогда существует **σ**-окрестность корня $\bar{x}$ такая, что при произвольном выборе приближений* $x^{(0)}$ и $x^{(1)}$ *из этой **σ**-окрестности*

Метод сходится с порядком p = $\frac{1+\sqrt{5}}{2} \approx 1.618$, т.е. для n $\geq$ 1 справедлива оценка

$$|x^{(n+1)} - \bar{x}| \leq c\,|x^{(n)} - \bar{x}\,|^{p}, \quad p = \frac{1+\sqrt{5}}{2}.$$

Так как одна итерация метода секущих требует только одного нового вычисления значения функции f, а метод Ньютона – двух вычислений (f и f'), то трудоемкость двух итераций метода секущих приблизительно эквивалентна трудоемкости одной итерации метода Ньютона. Две итерации метода секущих дают порядок $p^2 \approx 2.618 > 2$, поэтому его можно расценивать как более быстрый по сравнению с методом Ньютона.

К сожалению, метод обладает только локальной сходимостью. Он требует выбора двух близких к корню начальных приближений $x^{(0)}$ и $x^{(1)}$. Если эти приближения выбраны неудачно, то метод расходится.