

Вычислительная математика

Лекция 2. Численные методы решения нелинейных уравнений

Исупов К. С.
isupov.k@gmail.com

28 сентября 2018 г.

- 1 Постановка задачи
- 2 Локализация корней
- 3 Уточнение корней
- 4 Методы уточнения корней
 - Метод половинного деления (бисекции, дихотомии)
 - Метод хорд
 - Метод касательных Ньютона
 - Модифицированный метод Ньютона
 - Метод секущих
 - Комбинированный метод хорд и касательных
 - Метод простых итераций (Fixed-point iteration)
 - Неподвижные точки функции
 - Алгоритм простых итераций
 - Условия сходимости
 - Выбор начального приближения и скорость сходимости
 - Преобразование уравнения к каноническому виду
 - Критерий останова
 - Наибольшая скорость сходимости

Постановка задачи

Нелинейное уравнение с одним неизвестным можно записать в виде:

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

где $f(x)$ — некоторая функция аргумента x .

Определение

1. Функция $f(x)$ имеет **корень (однократный корень, простой корень)** $x = r$, если $f(r) = 0$.
2. Если в точке $x = r$ наряду с функцией обращаются в ноль и ее производные до $(k - 1)$ -го порядка, то число r — **корень кратности k** .

Типы нелинейных уравнений

1. Алгебраические: $f(x) = a_0 + a_1x + a_2x^2 + \dots + a_nx^n = 0$.
2. Трансцендентные: $2^x - 2 \cos(x) = 0, \lg(x + 5) = \cos(x)$.

Методы решения нелинейных уравнений

1. Прямые (аналитические, точные). Позволяют записать решение в виде формулы.
2. Численные. Позволяют получить приближения корней с любой заданной точностью.

Этапы численного решения нелинейных уравнений

1. Локализация, т.е. отделение корней (Bracketing a root).
2. Уточнение корней, т.е. вычисление приближенных значений корней с заданной точностью.

Локализация корней

Задачи

1. Определить, что на заданном отрезке $[a, b]$ существуют корни.
2. Определить, что на заданном отрезке $[a, b]$ корень единственный.

Теорема о существовании корней

Если функция $f(x)$ непрерывна на отрезке $[a, b]$, а на концах отрезка ее значения имеют разные знаки, т.е. $f(a)f(b) < 0$, то на этом отрезке расположен, по крайней мере, один корень.

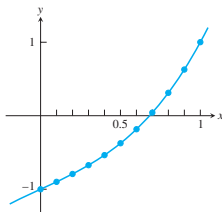


Рис. 1: График функции $f(x) = x^3 + x - 1$. $f(0)f(1) = -1 \cdot 1 < 0$, функция имеет корень на отрезке $[0, 1]$, т.е. найдется такое число r , что $f(r) = 0$

Теорема о единственности корня

Пусть дана функция $f(x)$, причем на отрезке $[a, b]$ она непрерывна и дифференцируема, и имеет разные знаки на концах, т.е. $f(a)f(b) < 0$. Тогда на этом отрезке содержится ровно один корень, если $f(x)$ монотонна, т.е. $f'(a)f'(b) > 0$.

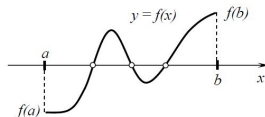


Рис. 2: Отделение корней. Функция $f(x)$ не монотонна на отрезке $[a, b]$

Способы отделения корней

1. Аналитический (нахождение интервалов монотонности функции).
2. Графический (построение графика функции $y = f(x)$ и поиск пересечения y с осью абсцисс, либо построение графика $f(x_1) = f(x_2)$ и поиск точки пересечения этих функций).
3. Табличный.

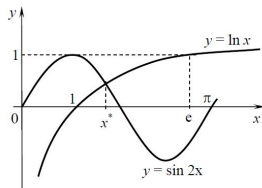


Рис. 3: Графическое отделение корней уравнения $\sin 2x - \ln x = 0$

Уточнение корней

На данном этапе получается приближенное значение корня, принадлежащего отрезку $[a, b]$, с заданной точностью ε . Т.е. вычисленное значение корня x должно отличаться от точного r не более, чем на величину ε :

$$|r - x| \leq \varepsilon.$$

Этапы уточнения корней

1. Выбор начального приближения к корню $x_0 \in [a, b]$.
2. Вычисление по некоторой формуле последующих приближений x_1, x_2, \dots (итерационный процесс).

Если значения x_k с ростом k стремятся к точному значению корня:

$$\lim_{k \rightarrow \infty} \{x_k\} = r,$$

то говорят, что итерационный процесс **сходится**.

Сходимость итерационного процесса: погрешность каждого последующего приближения должна быть меньше погрешности предыдущего приближения, т.е. погрешность приближенных значений с каждым шагом должна уменьшаться:

$$|r - x_{k+1}| < |r - x_k|.$$

В общем случае это неравенство можно представить в виде:

$$|r - x_{k+1}| < q|r - x_k|^\alpha,$$

где $q > 0$ и $\alpha \leq 1$ — числа, значения которых определяются методом уточнения корня. Главный показатель сходимости метода — значение α , называемое **порядком сходимости**. При $\alpha = 1$ сходимость линейная, при $\alpha > 1$ — сверхлинейная.

Метод половинного деления (бисекции, дихотомии)

Пусть дано нелинейное уравнение

$$f(x) = 0.$$

Считаем, что отделение корней проведено и на отрезке $[a, b]$ расположен один корень, который нужно уточнить с погрешностью ε .

Алгоритм

1. В качестве начального приближения корня примем середину отрезка: $c_0 = (a + b)/2$.
2. Исследуем значение функции $f(x)$ на концах отрезков $[a, c_0]$ и $[c_0, b]$: тот из отрезков, на концах которого $f(x)$ принимает значения разных знаков, содержит искомый корень, поэтому принимаем его в качестве нового отрезка $[a_1, b_1]$. Вторую же половину отрезка $[a, b]$, на которой $f(x)$ не меняет знак, отбрасываем.
3. В качестве следующего приближения корня принимаем середину нового отрезка: $c_1 = (a_1 + b_1)/2$, и повторяем пункт 2. Таким образом, k -е приближение вычисляется как:

$$c_k = \frac{a_k + b_k}{2}.$$

4. Условия прекращения итерационного процесса:

$$|r - c_k| < \varepsilon. \quad (2)$$

Так как

$$|r - c_k| < \frac{|b_k - a_k|}{2},$$

то (2) будет выполнено, если:

$$|b_k - a_k| < 2\varepsilon. \quad (3)$$

Метод половинного деления (бисекции, дихотомии)

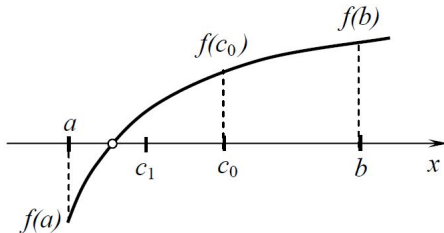


Рис. 4: Метод дихотомии

Преимущества

1. Безусловная сходимость.
2. Простота, применим для любых уравнений.

Недостатки

Медленный (линейная сходимость): с каждым шагом погрешность уменьшается в два раза:

$$|r - c_{k+1}| \leq \frac{1}{2} |r - c_k|.$$

Метод хорд

Очередное приближение в отличие от метода дихотомии берем не в середине отрезка, а в точке x_0 , где пересекает ось абсцисс прямая (хорда), проведенная через точки A и B :

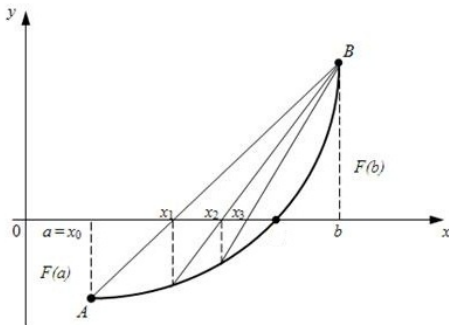


Рис. 5: Метод хорд

Уравнение прямой, проходящей через точки A и B :

$$\frac{x - a}{b - a} = \frac{y - f(a)}{f(b) - f(a)}. \quad (4)$$

Для точки пересечения с осью абсцисс ($x = x_0$, $y = 0$) получим уравнение:

$$x_0 = a - f(a) \frac{b - a}{f(b) - f(a)}. \quad (5)$$

Это и есть первое приближенное значение корня. В качестве нового интервала изоляции выбираем $[x_0, b]$, где функция меняет знак.

Метод хорд

Расчетные формулы

В зависимости от вида функции, неподвижной может быть точка a или b .

1. Когда неподвижен правый конец (b), а $x_0 = a$:

$$x_{k+1} = x_k - f(x_k) \frac{b - x_k}{f(b) - f(x_k)}. \quad (6)$$

2. Когда неподвижен левый конец (a), а $x_0 = b$:

$$x_{k+1} = x_k - f(x_k) \frac{x_k - a}{f(x_k) - f(a)}. \quad (7)$$

Неподвижным будет тот конец интервала изоляции, для которого знак функции совпадает со знаком ее второй производной $f''(x)$, а последовательные приближения лежат по другую сторону корня.

Оценка погрешности

$$m = \min_{[a,b]} |f'(x)|; \quad M = \max_{[a,b]} |f''(x)|.$$

Если не накладывается ограничений на длину интервала изоляции, то погрешность оценивается по формуле:

$$|r - x_k| \leq f(x_k)/m. \quad (8)$$

Чаще стремятся сузить интервал изоляции корня до выполнения неравенства: $M \leq 2m$. Тогда погрешность оценивают по формуле:

$$|r - x_k| \leq |x_k - x_{k-1}|. \quad (9)$$

Метод касательных Ньютона

Пусть известно начальное приближение x_0 . Проведем в этой точке касательную к кривой $y = f(x)$:

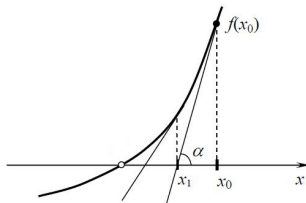


Рис. 6: Метод Ньютона

Эта касательная пересечет ось абсцисс в точке x_1 , которая рассматривается в качестве следующего приближения. Значение x_1 находится из уравнения:

$$\operatorname{tg} \alpha = \frac{f(x_0)}{x_0 - x_1} = f'(x_0). \quad (10)$$

Выражая x_1 , имеем

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}. \quad (11)$$

В общем случае:

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}, \quad k = 0, 1, \dots \quad (12)$$

Функция $f(x)$ должна быть дифференцируемой и $f'(x)$ в окрестности корня не должна менять знак.

Условия окончания итерационного процесса те же, что и для метода хорд.

Метод касательных Ньютона

Другой способ вывода метода

Пусть x_0 — некоторое начальное приближение корня. Заменим функцию $f(x)$ в окрестности точки x_0 отрезком ряда Тейлора:

$$f(x) \approx f(x_0) + (x - x_0)f'(x_0).$$

Вместо нелинейного уравнения $f(x) = 0$ решим линеаризованное уравнение относительно x_1 :

$$f(x_0) + (x - x_0)f'(x_0) = 0. \quad (13)$$

Будем рассматривать решение этого уравнения как следующее приближение к искомому значению корня:

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}. \quad (14)$$

Скорость сходимости

- Скорость сходимости метода определяется модулем первой производной функции в окрестности корня: чем больше крутизна графика, тем меньше поправка, которую надо прибавить к n -у приближению, чтобы получить $(n + 1)$ -е приближение.
- На каждом шаге погрешность пропорциональна квадрату погрешности на предыдущем шаге. Следовательно, сходимость квадратичная.

Выбор начального приближения

Если задан отрезок $[a, b]$, содержащий корень, и известно, что функция $f(x)$ монотонна на этом отрезке, то в качестве начального приближения x_0 следует выбрать ту границу отрезка $[a, b]$, где совпадают знаки функции $f(x)$ и ее второй производной $f''(x)$.

Модифицированный метод Ньютона

1. Зачастую нахождение производной может быть сопряжено с существенными затратами.
2. В модифицированном методе Ньютона производная вычисляется только в начальной точке (x_0) , а расчетные формулы имеют вид:

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

Метод секущих

Производную в окрестности точки x_k заменяют разностной аппроксимацией:

$$f'(x) \approx \frac{f(x_k) - f(x_{k-1})}{x_k - x_{k-1}}. \quad (15)$$

После подстановки в формулу Ньютона

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

имеем:

$$x_{k+1} = x_k - \frac{f(x_k)(x_k - x_{k-1})}{f(x_k) - f(x_{k-1})}, \quad k = 1, 2, \dots \quad (16)$$

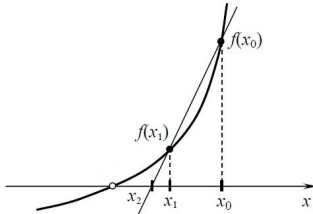


Рис. 7: Метод секущих: секущая, проведенная через точки $(x_0, f(x_0))$ и $(x_1, f(x_1))$ пересекает ось абсцисс в точке x_2 , значение которой определяется формулой (16)

Для начала итерационного процесса нужно задать два начальных приближения: x_0 и x_1 (можно выбрать границы интервала изоляции).

Комбинированный метод хорд и касательных

Суть комбинированного метода состоит в том, чтобы сузить интервал изоляции с двух сторон.

Расчетные формулы

В зависимости от неподвижного конца интервала изоляции используются две пары формул, и на каждом шаге вычисляются приближенное значение корня по недостатку x_{k+1} и по избытку \tilde{x}_{k+1} .

1. Если неподвижна точка a :

- значение по недостатку вычисляется методом касательных:

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

- значение по избытку вычисляется методом хорд:

$$\tilde{x}_{k+1} = \tilde{x}_k - f(\tilde{x}_k) \frac{\tilde{x}_k - x_k}{f(\tilde{x}_k) - f(x_k)}, \quad \tilde{x}_0 = b;$$

2. Если неподвижна точка b :

- значение по недостатку вычисляется методом хорд:

$$x_{k+1} = x_k - \frac{\tilde{x}_k - x_k}{f(\tilde{x}_k) - f(x_k)}, \quad x_0 = a;$$

- значение по избытку вычисляется методом касательных:

$$\tilde{x}_{k+1} = \tilde{x}_k - \frac{f(\tilde{x}_k)}{f'(\tilde{x}_k)}, \quad \tilde{x}_0 = b;$$

На каждом новом шаге расчета используют суженный с двух сторон интервал $[x_k, \tilde{x}_k]$.

Метод простых итераций. Неподвижные точки функции

Пример

1. Выберем произвольный аргумент, например $x = 1.85$.
2. Посчитаем значение функции $\cos x = \cos(1.85) \approx -0.275590$.
3. Посчитаем значение функции $\cos(-0.275590) \approx 0.962265$.
4. Посчитаем значение функции $\cos(0.962265) \approx 0.571663$.
5. ...
6. В результате приходим к числу 0.73908513322 То есть

$$\cos \cos \cos \cos \dots \cos x \rightarrow 0.73908513322, \quad \text{где } x \text{ — произвольное число.}$$

Таким образом, если $\xi = 0.73908513322$, то $\cos(\xi) = \xi$, и ξ — это *неподвижная точка* для \cos .

Определение

Вещественное число ξ называется неподвижной точкой функции ϕ , если $\phi(\xi) = \xi$.

Применение

Пусть, например, дано уравнение $\cos x - x = 0$. С точки зрения неподвижных точек это уравнение может быть записано в виде:

$$x = \cos x. \quad (17)$$

Очевидно, что если значение функции совпадает со значением аргумента, то есть ξ — неподвижная точка функции $\cos x$, то ξ и есть корень уравнения (17).

Метод простых итераций. Алгоритм

Пусть дано нелинейное уравнение

$$f(x) = 0.$$

Считаем, что отделение корней проведено и на отрезке $[a, b]$ расположен один корень ξ , который нужно уточнить с погрешностью ε .

Алгоритм

1. Преобразуем уравнение к каноническому виду:

$$f(x) = 0 \rightarrow x = \phi(x). \quad (18)$$

2. Выберем начальное приближение x_0 — любую точку из интервала $[a, b]$.
3. Итерационный процесс осуществляем в соответствии с рекуррентным соотношением:

$$x_{n+1} = \phi(x_n), \quad n = 1, 2, \dots \quad (19)$$

В развернутом виде:

$$\begin{aligned} x_1 &= \phi(x_0), \\ x_2 &= \phi(x_1), \\ x_3 &= \phi(x_2), \\ &\dots \text{profit!} \end{aligned}$$

Доказано, что при определенных свойствах функции $\phi(x)$ последовательность $\{x_n\}$ сходится к корню ξ уравнения $f(x) = 0$.

Главный вопрос: при каких условиях итерационный процесс (19) будет сходиться?

Метод простых итераций. Условия сходимости

Геометрическая интерпретация сходящегося процесса

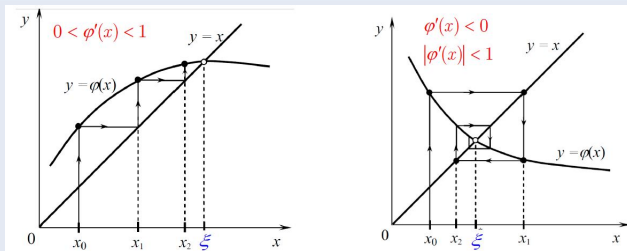


Рис. 8: Монотонно (слева) и двусторонне (справа) сходящиеся последовательности

Геометрическая интерпретация расходящегося процесса

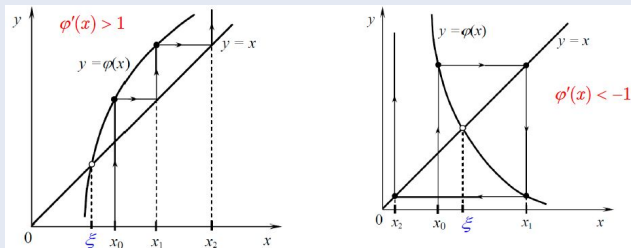


Рис. 9: Монотонно (слева) и двусторонне (справа) расходящиеся последовательности

Метод простых итераций. Условия сходимости

Теорема о сходимости итерационного процесса

Итерационный процесс сходится при условии, что первая производная итерационной функции $\phi(x)$ по модулю меньше единицы.

Доказательство.

1. Представим n -е и $(n + 1)$ -е приближения в форме

$$x_n = \xi + \varepsilon_n, \quad x_{n+1} = \xi + \varepsilon_{n+1}, \quad (20)$$

где $\varepsilon_n, \varepsilon_{n+1}$ — отклонения от корня.

2. Функцию $\phi(x)$ в окрестности точки ξ заменим двумя первыми членами ряда Тейлора:

$$\phi(x) \approx \phi(\xi) + \varepsilon \phi'(\xi). \quad (21)$$

3. Тогда итерационная формула $x_{n+1} = \phi(x_n)$ примет вид

$$\xi + \varepsilon_{n+1} = \phi(\xi) + \varepsilon_n \phi'(\xi). \quad (22)$$

4. Поскольку ξ — корень уравнения, то $\xi = \phi(\xi)$, откуда следует, что

$$\varepsilon_{n+1} = \varepsilon_n \phi'(\xi). \quad (23)$$

5. Для сходимости итерационного процесса необходимо и достаточно, чтобы

$$|\varepsilon_{n+1}| < |\varepsilon_n|, \quad (24)$$

следовательно, для любого x в окрестности точки ξ должно выполняться условие:

$|\phi'(x)| < 1$, что непосредственно следует из (23).

Метод простых итераций. Выбор начального приближения и скорость сходимости

1. Если условие

$$|\phi'(x)| < 1 \quad (25)$$

выполняется, то на отрезке $[a, b]$, на котором локализован корень, в качестве начального приближения можно взять **любую** точку $x_0 \in [a, b]$.

2. Скорость сходимости зависит от абсолютной величины производной: чем меньше $|\phi'(x)|$, тем быстрее сходится процесс, что непосредственно следует из формулы (23):

$$\varepsilon_{n+1} = \varepsilon_n \phi'(\xi). \quad (26)$$

Метод простых итераций. Преобразование уравнения к каноническому виду

Задача: для заданного уравнения $f(x) = 0$ построить функцию $\phi(x)$ так, чтобы выполнялось условие (25):

$$|\phi'(x)| < 1.$$

Пример

Пусть дано уравнение

$$x^3 - x - 1 = 0, \quad (27)$$

и известно, что один из корней расположен на $I = [1, 2]$.

1. Преобразуем уравнение следующим образом:

$$x = x^3 - 1. \quad (28)$$

Проверим условие сходимости для точки $x_0 = \text{mid}(I) = 1.5$:

$$\phi(x) = x^3 - 1, \quad \phi'(x) = 3x^2, \quad |\phi(1.5)| = 6.75 > 1.$$

2. Преобразуем уравнение другим способом:

$$x^3 = x + 1 \Rightarrow x = \sqrt[3]{x + 1}. \quad (29)$$

Проверим условие сходимости для точки $x_0 = \text{mid}(I) = 1.5$:

$$\phi(x) = \sqrt[3]{x + 1}, \quad \phi'(x) = \frac{1}{3(x + 1)^{2/3}}, \quad |\phi(1.5)| = 0.133 < 1.$$

Метод простых итераций. Преобразование уравнения к каноническому виду

Общий алгоритм преобразования уравнения к итерационному виду

1. Умножить левую и правую части уравнения $f(x) = 0$ на произвольную константу $k \neq 0$:

$$k \cdot f(x) = 0 \cdot k.$$

2. Добавить к обоим частям уравнения неизвестное x :

$$k \cdot f(x) + x = 0 \cdot k + x \Leftrightarrow x = x + k \cdot f(x).$$

3. Полученное уравнение, очевидно, эквивалентно уравнению

$$x = \phi(x)$$

с функцией $\phi(x) = x + k \cdot f(x)$. Произвольный выбор константы k позволяет обеспечить выполнение условия сходимости $|\phi'(x)| < 1$.

4. Поскольку $\phi'(x) = 1 + k \cdot f'(x)$, то значение k следует выбирать таким образом, чтобы в окрестности корня выполнялось условие:

$$|\phi'(x)| = |1 + k \cdot f'(x)| < 1. \quad (30)$$

Например, если

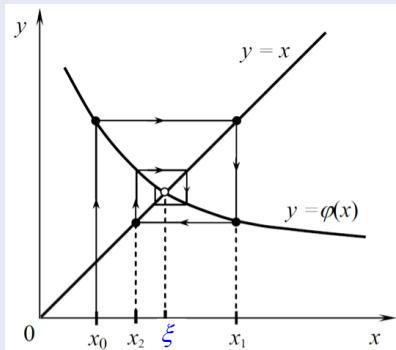
$$|k| < \frac{2}{\max |f'(x)|}, \quad \text{sgn}(k) = \text{sgn}(f'(x)), \quad (31)$$

то условие (30) выполняется.

Критерий останова

Двусторонняя сходимость

Если $\phi'(x) < 0$, то сходимость двусторонняя:



В этом случае

$$|x_{n+1} - x_n| > |x_{n+1} - \xi|, \quad (32)$$

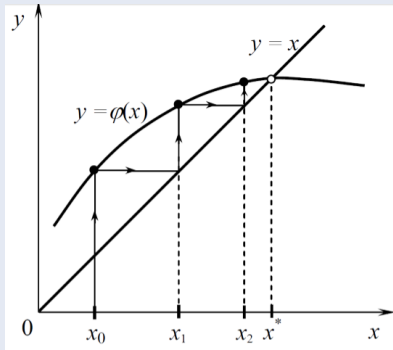
и критерий окончания итерационного процесса

$$|x_{n+1} - x_n| \leq \varepsilon \quad (33)$$

является объективным.

Монотонная сходимость

Если $\phi'(x) > 0$, то сходимость односторонняя:



В этом случае контроль достигнутой точности следует осуществлять по проверке неравенства:

$$\frac{1}{1-q} |x_{n+1} - x_n| \leq \varepsilon, \quad (34)$$

где

$$q = \max_{[a,b]} |\phi'(x)|.$$

Метод простых итераций. Наибольшая скорость сходимости

Наибольшая скорость сходимости итерационного процесса:

$$x_{n+1} = \phi(x_n), \quad n = 1, 2, \dots \quad (35)$$

достигается при

$$\phi'(x) = 0.$$

Этого можно добиться, если выбрать параметр k в уравнении

$$\phi(x) = x + k \cdot f(x) \quad (36)$$

зависящим от x в виде

$$k = -\frac{1}{f'(x)}. \quad (37)$$

При этом итерационная формула 35 переходит в формулу метода касательных Ньютона:

$$x_{n+1} = \phi(x_n) = x - \frac{f(x)}{f'(x)}, \quad n = 1, 2, \dots \quad (38)$$