Вычислительная математика. Методы решения нелинейных уравнений.

МФТИ

Основная задача

Найти изолированный корень уравнения $f(x) = 0$

Основная задача

Найти изолированный корень уравнения $f(x) = 0$

- Если f - непрерывная функция и z - простой корень, то z принадлежит отрезку, на концах которого $f(x)$ разного знака

Основная задача

Найти изолированный корень уравнения $f(x) = 0$

- Если f - непрерывная функция и z - простой корень, то z принадлежит отрезку, на концах которого $f(x)$ разного знака
- Простейшие методы: деление пополам, метод золотого сечения

Основная задача

Найти изолированный корень уравнения $f(x) = 0$

- Если f - непрерывная функция и z - простой корень, то z принадлежит отрезку, на концах которого $f(x)$ разного знака
- Простейшие методы: деление пополам, метод золотого сечения
- Итерационные методы: используют значения функции и производных в предыдущих приближениях

Критерий остановки

Критерий остановки

Теорема 14.1

Если f' непрерывна и $\forall y \in [x_k - \epsilon, x_k + \epsilon]$ верно:

$$|f(x_k)/f'(y)| \leq \epsilon$$

то $f(z) = 0$ для некоторого $z \in [x_k - \epsilon, x_k + \epsilon]$

Критерий остановки

Теорема 14.1

Если f' непрерывна и $\forall y \in [x_k - \epsilon, x_k + \epsilon]$ верно:

$$|f(x_k)/f'(y)| \leq \epsilon$$

то $f(z) = 0$ для некоторого $z \in [x_k - \epsilon, x_k + \epsilon]$

- По теореме Лагранжа $f(x_k + t) = f(x_k) + f'(y(t))t$

Критерий остановки

Теорема 14.1

Если f' непрерывна и $\forall y \in [x_k - \epsilon, x_k + \epsilon]$ верно:

$$|f(x_k)/f'(y)| \leq \epsilon$$

то $f(z) = 0$ для некоторого $z \in [x_k - \epsilon, x_k + \epsilon]$

- По теореме Лагранжа $f(x_k + t) = f(x_k) + f'(y(t))t$
- Следовательно

$$\frac{f(x_k + t)}{f'(y)} = \frac{f(x_k)}{f'(y)} + t \geq 0 \text{ при } t = \epsilon, \leq 0 \text{ при } t = -\epsilon$$

Критерий остановки

Теорема 14.1

Если f' непрерывна и $\forall y \in [x_k - \epsilon, x_k + \epsilon]$ верно:

$$|f(x_k)/f'(y)| \le \epsilon$$

то $f(z) = 0$ для некоторого $z \in [x_k - \epsilon, x_k + \epsilon]$

- По теореме Лагранжа $f(x_k + t) = f(x_k) + f'(y(t))t$
- Следовательно

$$\frac{f(x_k + t)}{f'(y)} = \frac{f(x_k)}{f'(y)} + t \ge 0 \text{ при } t = \epsilon, \le 0 \text{ при } t = -\epsilon$$

- Производная не меняет знак $\Rightarrow f(x_k + \epsilon) f(x_k - \epsilon) \le 0 \quad \square.$

Метод простой итерации

Метод простой итерации

- $f(x) = 0 \iff x = F(x)$: $x_{k+1} = F(x_k), k = 0, 1, \ldots$

Метод простой итерации

- $f(x) = 0 \iff x = F(x)$: $x_{k+1} = F(x_k), k = 0, 1, \ldots$
- Решение z уравнения $z = F(z)$ называется *неподвижной точкой отображения* F.

Метод простой итерации

- $f(x) = 0 \iff x = F(x)$: $x_{k+1} = F(x_k), k = 0, 1, \ldots$
- Решение z уравнения $z = F(z)$ называется *неподвижной точкой отображения* F.

Теорема 14.2 (о сжимающих отображениях)

Пусть M полное метрическое пространство с расстоянием $\rho(\cdot, \cdot)$. $F : M \to M$ – сжимающее:

$$\rho(F(x), F(y)) \leq q\rho(x, y), 0 < q < 1 \quad \forall x, y \in M$$

Тогда $x = F(x)$ имеет единственное решение z и $\forall x_0$ МПИ сходится к z:

$$\rho(x_k, z) \leq \frac{q^k}{1-q}\rho(x_1, x_0)$$

Доказательство теоремы

Доказательство теоремы

- При $m \geq k$

$$\rho(x_m, x_k) \leq \sum_{i=k}^{m-1} \rho(x_{i+1}, x_i) \leq$$

$$\sum_{i=k}^{m-1} q^i \rho(x_1, x_0) \leq \frac{q^k}{1-q}\rho(x_1, x_0)$$

Доказательство теоремы

- При $m \geq k$

$$\rho(x_m, x_k) \leq \sum_{i=k}^{m-1} \rho(x_{i+1}, x_i) \leq$$

$$\sum_{i=k}^{m-1} q^i \rho(x_1, x_0) \leq \frac{q^k}{1-q} \rho(x_1, x_0)$$

- $\Rightarrow \{x_k\}$ - фундаментальная $\Rightarrow$ в силу полноты M она сходится к некоторому $z \in M$.

Доказательство теоремы

- При $m \geq k$

$$\rho(x_m, x_k) \leq \sum_{i=k}^{m-1} \rho(x_{i+1}, x_i) \leq$$

$$\sum_{i=k}^{m-1} q^i \rho(x_1, x_0) \leq \frac{q^k}{1-q} \rho(x_1, x_0)$$

- $\Rightarrow \{x_k\}$ - фундаментальная $\Rightarrow$ в силу полноты M она сходится к некоторому $z \in M$.
- $0 \leq \rho(x^{k+1}, F(z)) = \rho(F(x^k), F(z)) \leq q\rho(x^k, z) \Rightarrow z = F(z)$

Доказательство теоремы

- При $m \geq k$

$$\rho(x_m, x_k) \leq \sum_{i=k}^{m-1} \rho(x_{i+1}, x_i) \leq$$

$$\sum_{i=k}^{m-1} q^i \rho(x_1, x_0) \leq \frac{q^k}{1-q} \rho(x_1, x_0)$$

- $\Rightarrow \{x_k\}$ - фундаментальная $\Rightarrow$ в силу полноты M она сходится к некоторому $z \in M$.
- $0 \leq \rho(x^{k+1}, F(z)) = \rho(F(x^k), F(z)) \leq q\rho(x^k, z) \Rightarrow z = F(z)$
- Если $z_1 = F(z_1) : \rho(z_1, z) = \rho(F(z_1), F(z)) < \rho(z_1, z) \Rightarrow z_1 = z$

Доказательство теоремы

- При $m \geq k$

$$\rho(x_m, x_k) \leq \sum_{i=k}^{m-1} \rho(x_{i+1}, x_i) \leq$$

$$\sum_{i=k}^{m-1} q^i \rho(x_1, x_0) \leq \frac{q^k}{1-q} \rho(x_1, x_0)$$

- $\Rightarrow \{x_k\}$ - фундаментальная $\Rightarrow$ в силу полноты M она сходится к некоторому $z \in M$.
- $0 \leq \rho(x^{k+1}, F(z)) = \rho(F(x^k), F(z)) \leq q\rho(x^k, z) \Rightarrow z = F(z)$
- Если $z_1 = F(z_1) : \rho(z_1, z) = \rho(F(z_1), F(z)) < \rho(z_1, z) \Rightarrow z_1 = z$
- Переход к $m \to \infty$ даёт оценку ошибки $\square$.

Условие достижения заданной точности

$$\rho(x_n, x^\star) \leqslant \rho(x_n, x_{n+1}) + \rho(x_{n+1}, x^\star) \leqslant \rho(x_n, x_{n+1}) + q\rho(x_n, x^\star)$$

$$\rho(x_n, x^\star) \leqslant \frac{\rho(x_{n+1}, x_n)}{1-q} \leqslant \frac{q^n \rho(x_1, x_0)}{1-q} \leqslant \varepsilon$$

$$|x_n - x^\star| \leqslant \frac{|x_{n+1} - x_n|}{1-q} \leqslant \frac{q^n |x_1 - x_0|}{1-q}$$

МПИ в $\mathbb{R}^1$

МПИ в $\mathbb{R}^1$

- Пусть $F \in C^1[z-\delta, z+\delta]$, z – единственная неподвижная точка

МПИ в $\mathbb{R}^1$

- Пусть $F \in C^1[z-\delta, z+\delta]$, z – единственная неподвижная точка
- Если $|F'(z)| < 1$, то в некоторой δ-окрестности

$$q = \max_{|x-z|\le\delta} |F'(x)| < 1 \;\Rightarrow |F(x) - F(y)| \le q|x-y| \;\forall x, y$$

МПИ в $\mathbb{R}^1$

- Пусть $F \in C^1[z-\delta, z+\delta]$, z – единственная неподвижная точка
- Если $|F'(z)| < 1$, то в некоторой δ-окрестности

$$q = \max_{|x-z| \le \delta} |F'(x)| < 1 \ \Rightarrow |F(x) - F(y)| \le q|x-y| \ \forall x, y$$

- Следовательно, F - сжимающее отображение на $M = [z-\delta, z+\delta]$, МПИ сходится $\forall x_0 \in M$

МПИ в $\mathbb{R}^n$

$$F'(x) = \begin{bmatrix} \frac{\partial F_1(x)}{\partial x_1} & \dots & \frac{\partial F_1(x)}{\partial x_n} \\ \dots & \dots & \dots \\ \frac{\partial F_n(x)}{\partial x_1} & \dots & \frac{\partial F_n(x)}{\partial x_n} \end{bmatrix}$$

МПИ в $\mathbb{R}^n$

$$F'(x) = \begin{bmatrix} \frac{\partial F_1(x)}{\partial x_1} & \dots & \frac{\partial F_1(x)}{\partial x_n} \\ \dots & \dots & \dots \\ \frac{\partial F_n(x)}{\partial x_1} & \dots & \frac{\partial F_n(x)}{\partial x_n} \end{bmatrix}$$

Теорема 14.3 (достаточное условие сходимости)

Пусть отображение $F : \mathbb{R}^n \to \mathbb{R}^n$ имеет единственную неподвижную точку $z = F(z)$ и непрерывно дифференцируемо в некоторой её окрестности.
Тогда, если $\rho(F'(z)) < 1$, то $\forall x_0$ из некоторой окрестности точки z метод простой итерации сходится к z.

Метод простой итерации с параметром

Метод простой итерации с параметром

- $F(x) = x - \alpha(x)f(x)$

Метод простой итерации с параметром

- $F(x) = x - \alpha(x) f(x)$
- Если z изолированный *простой* корень ($f'(z) \neq 0$), то α всегда можно выбрать так, чтобы выполнялось $|F'(z)| < 1$.

Метод простой итерации с параметром

- $F(x) = x - \alpha(x) f(x)$
- Если z изолированный *простой* корень $(f'(z) \neq 0)$, то α всегда можно выбрать так, чтобы выполнялось $|F'(z)| < 1$.
- Если известны оценки $|f'(x)|$, то можно выбрать *оптимальное* значение α (как в м-де Ричардсона).

Метод простой итерации с параметром

- $F(x) = x - \alpha(x) f(x)$
- Если z изолированный *простой* корень ($f'(z) \neq 0$), то α всегда можно выбрать так, чтобы выполнялось $|F'(z)| < 1$.
- Если известны оценки $|f'(x)|$, то можно выбрать *оптимальное* значение α (как в м-де Ричардсона).
- Нулевое значение $|F'(z)|$ будет при $\alpha = 1/f'(z)$. Можно взять приближение, тогда получим метод Ньютона:

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

Метод Ньютона

Метод Ньютона

- f приближается многочленом 1-й степени (мн-н Тейлора или интерполяционный мн-н Эрмита)

$$f(x) \approx H(x) = f(x_k) + f'(x_k)(x - x_k)$$

Метод Ньютона

- f приближается многочленом 1-й степени (мн-н Тейлора или интерполяционный мн-н Эрмита)

$$f(x) \approx H(x) = f(x_k) + f'(x_k)(x - x_k)$$

- Пусть $f \in C^2$, $f'(z) \neq 0$. Тогда

$$f(z) - H(z) = \frac{f''(\xi_k)}{2}(z - x_k)^2$$
$$H(x_{k+1}) - H(z) = f'(x_k)(x_{k+1} - z)$$

Метод Ньютона

- f приближается многочленом 1-й степени (мн-н Тейлора или интерполяционный мн-н Эрмита)

$$f(x) \approx H(x) = f(x_k) + f'(x_k)(x - x_k)$$

- Пусть $f \in C^2$, $f'(z) \neq 0$. Тогда

$$f(z) - H(z) = \frac{f''(\xi_k)}{2}(z - x_k)^2$$
$$H(x_{k+1}) - H(z) = f'(x_k)(x_{k+1} - z)$$

- Т.к. $f(z) = 0, H(x_{k+1}) = 0$, получаем оценку ошибки:

$$\boxed{e_{k+1} = -\frac{f''(\xi_k)}{2f'(x_k)}e_k^2}, \; e_k = z - x_k$$

Метод Ньютона, свойства

Метод Ньютона, свойства

- Метод Ньютона в лучшем случае сходится *квадратично*.

Метод Ньютона, свойства

- Метод Ньютона в лучшем случае сходится *квадратично*.
- По определению последовательность x_k сходится к z с порядком p, если

$$\lim_{k\to\infty} \sup \left|\frac{e_{k+1}}{e_k^p}\right| \leq c \leq \infty$$

Метод Ньютона, свойства

- Метод Ньютона в лучшем случае сходится *квадратично*.
- По определению последовательность x_k сходится к z с порядком p, если

$$\lim_{k\to\infty} \sup \left|\frac{e_{k+1}}{e_k^p}\right| \leq c \leq \infty$$

- Условие $f'(z) \neq 0$ означает, что корень является простым. В общем случае, z - корень кратности m, если $f^{(j)}(z) = 0$ при $0 \leq j \leq m-1$ и $f^{(m)}(z) \neq 0$.

Метод Ньютона, свойства

- Метод Ньютона в лучшем случае сходится *квадратично*.
- По определению последовательность x_k сходится к z с порядком p, если

$$\lim_{k\to\infty} \sup \left|\frac{e_{k+1}}{e_k^p}\right| \leq c \leq \infty$$

- Условие $f'(z) \neq 0$ означает, что корень является простым. В общем случае, z - корень кратности m, если $f^{(j)}(z) = 0$ при $0 \leq j \leq m-1$ и $f^{(m)}(z) \neq 0$.
- Метод Ньютона может сходиться и для кратного корня, но сходимость может не быть квадратичной. Для $f(x) = x^2$, $e_{k+1} = e_k/2$.

Сходимость метода Ньютона

Теорема 14.4

Пусть z - простой корень уравнения $f(x) = 0$ и

Сходимость метода Ньютона

Теорема 14.4

Пусть z - простой корень уравнения $f(x) = 0$ и

1. $f \in C^2[z - \delta, z + \delta]$

Сходимость метода Ньютона

Теорема 14.4

Пусть z - простой корень уравнения $f(x) = 0$ и

1. $f \in C^2[z-\delta, z+\delta]$
2. $f'(x) \neq 0$ при $x \in [z-\delta, z+\delta]$

Сходимость метода Ньютона

Теорема 14.4

Пусть z - простой корень уравнения $f(x) = 0$ и

1. $f \in C^2[z-\delta, z+\delta]$
2. $f'(x) \neq 0$ при $x \in [z-\delta, z+\delta]$
3. $\gamma \equiv \dfrac{\max\limits_{|x-z|\leq\delta} \left|f''(x)\right|}{2 \min\limits_{|x-z|\leq\delta} \left|f'(x)\right|} \neq 0$

Сходимость метода Ньютона

Теорема 14.4

Пусть z - простой корень уравнения $f(x) = 0$ и

1. $f \in C^2[z-\delta, z+\delta]$
2. $f'(x) \neq 0$ при $x \in [z-\delta, z+\delta]$
3. $\gamma \equiv \dfrac{\max\limits_{|x-z|\leq\delta} \left|f''(x)\right|}{2\min\limits_{|x-z|\leq\delta} \left|f'(x)\right|} \neq 0$

Тогда $\forall\epsilon : \ 0 < \epsilon < \min\{\delta, \gamma^{-1}\}$ метод Ньютона сходится при любом начальном приближении $x_0 \in [z-\epsilon, z+\epsilon]$ и верно:

Сходимость метода Ньютона

Теорема 14.4

Пусть z - простой корень уравнения $f(x) = 0$ и

1. $f \in C^2[z-\delta, z+\delta]$
2. $f'(x) \neq 0$ при $x \in [z-\delta, z+\delta]$
3. $\gamma \equiv \dfrac{\max\limits_{|x-z|\leq\delta} |f''(x)|}{2\min\limits_{|x-z|\leq\delta} |f'(x)|} \neq 0$

Тогда $\forall\epsilon :\ 0 < \epsilon < \min\{\delta, \gamma^{-1}\}$ метод Ньютона сходится при любом начальном приближении $x_0 \in [z-\epsilon, z+\epsilon]$ и верно:

1. $|e_{k+1}| \leq \gamma |e_k|^2$

Сходимость метода Ньютона

Теорема 14.4

Пусть z - простой корень уравнения $f(x) = 0$ и

1. $f \in C^2[z-\delta, z+\delta]$
2. $f'(x) \neq 0$ при $x \in [z-\delta, z+\delta]$
3. $\gamma \equiv \dfrac{\max\limits_{|x-z|\le\delta} |f''(x)|}{2 \min\limits_{|x-z|\le\delta} |f'(x)|} \neq 0$

Тогда $\forall \epsilon :\ 0 < \epsilon < \min\{\delta, \gamma^{-1}\}$ метод Ньютона сходится при любом начальном приближении $x_0 \in [z-\epsilon, z+\epsilon]$ и верно:

1. $|e_{k+1}| \le \gamma |e_k|^2$
2. $|e_k| \le \gamma^{-1}(\gamma|e_0|)^{2^k}$

Доказательство

Доказательство

- Пусть $x_k \in [z-\epsilon, z+\epsilon]$. Т.к. $\epsilon < \gamma^{-1}$:

$$e_{k+1} = -\frac{f''(\xi_k)}{2f'(x_k)} e_k^2 \Rightarrow$$
$$|e_{k+1}| \leq \gamma |e_k|^2 \leq (\gamma\epsilon)\epsilon \leq \epsilon \Rightarrow x_{k+1} \in [z-\epsilon, z+\epsilon]$$

Доказательство

- Пусть $x_k \in [z - \epsilon, z + \epsilon]$. Т.к. $\epsilon < \gamma^{-1}$:

$$e_{k+1} = -\frac{f''(\xi_k)}{2f'(x_k)} e_k^2 \Rightarrow$$
$$|e_{k+1}| \leq \gamma |e_k|^2 \leq (\gamma\epsilon)\epsilon \leq \epsilon \Rightarrow x_{k+1} \in [z - \epsilon, z + \epsilon]$$

- Умножим обе части на γ и обозначим $d_k = \gamma|e_k|$, тогда

$$d_{k+1} \leq d_k^2 \Rightarrow d_k \leq d_0^{2^k}$$

Доказательство

- Пусть $x_k \in [z - \epsilon, z + \epsilon]$. Т.к. $\epsilon < \gamma^{-1}$:

$$e_{k+1} = -\frac{f''(\xi_k)}{2f'(x_k)} e_k^2 \Rightarrow$$
$$|e_{k+1}| \le \gamma |e_k|^2 \le (\gamma\epsilon)\epsilon \le \epsilon \Rightarrow x_{k+1} \in [z - \epsilon, z + \epsilon]$$

- Умножим обе части на γ и обозначим $d_k = \gamma |e_k|$, тогда

$$d_{k+1} \le d_k^2 \Rightarrow d_k \le d_0^{2^k}$$

- Из условий на начальное приближение ($|x - z| < \gamma^{-1}$) следует, что $d_0 < 1 \Rightarrow d_k \to 0$ $\square$

Доказательство

- Пусть $x_k \in [z-\epsilon, z+\epsilon]$. Т.к. $\epsilon < \gamma^{-1}$:

$$e_{k+1} = -\frac{f''(\xi_k)}{2f'(x_k)} e_k^2 \Rightarrow$$
$$|e_{k+1}| \le \gamma |e_k|^2 \le (\gamma\epsilon)\epsilon \le \epsilon \Rightarrow x_{k+1} \in [z-\epsilon, z+\epsilon]$$

- Умножим обе части на γ и обозначим $d_k = \gamma|e_k|$, тогда

$$d_{k+1} \le d_k^2 \Rightarrow d_k \le d_0^{2^k}$$

- Из условий на начальное приближение ($|x-z| < \gamma^{-1}$) следует, что $d_0 < 1 \Rightarrow d_k \to 0$ $\square$
- Следствие: $\lim\limits_{k\to\infty} \frac{e_{k+1}}{e_k^2} = -\frac{f''(z)}{2f'(z)}$

Достижение заданной точности

$$|x_n - x^\star| \leqslant M_2 |x_n - x_{n-1}|^2/(2m_1) < \varepsilon$$

Пример: квадратный корень из числа

Метод Ньютона в $\mathbb{R}^n$

$$x_{k+1} = x_k - \left[f'(x_k)\right]^{-1} f(x_k)$$

Метод Ньютона в $\mathbb{R}^n$

$$x_{k+1} = x_k - [f'(x_k)]^{-1} f(x_k)$$

Теорема 14.5

$f : \mathbb{R}^n \to \mathbb{R}^n$, $f(z) = 0$ и в $B_\delta(z) = \{x : \|x - z\|_\infty \leq \delta\}$ f - непрерывно дифференцируемо, якобиан f существует и невырожден, и $\forall x, y \in B_\delta$:

$$\|f'(x) - f'(y)\|_\infty \leq c\|x - y\|_\infty,\ c > 0$$

Пусть $\gamma = c \max\limits_{\|z-x\|_\infty \leq \delta} \|[f'(x)^{-1}]\|_\infty$, $0 < \epsilon < \min\{\delta, \gamma^{-1}\}$. Тогда $\forall x_0 \in B_\epsilon(z)$ метод Ньютона сходится и

$$\|e_{k+1}\|_\infty \leq \gamma \|e_k\|_\infty^2, \quad \|e_k\|_\infty \leq \gamma^{-1}(\gamma\|e_0\|_\infty)^{2^k}$$

Доказательство

Доказательство

- По теореме Лагранжа найдутся точки $\xi_1, \dots, \xi_n$:

$$f(x) - f(y) = J_k(x - y), \quad J_k = \begin{bmatrix} \frac{\partial f_1(\xi_1)}{\partial x_1} & \dots & \frac{\partial f_1(\xi_1)}{\partial x_n} \\ \dots & \dots & \dots \\ \frac{\partial f_n(\xi_n)}{\partial x_1} & \dots & \frac{\partial f_n(\xi_n)}{\partial x_n} \end{bmatrix}$$

Доказательство

- По теореме Лагранжа найдутся точки $\xi_1, \ldots, \xi_n$:

$$f(x) - f(y) = J_k(x-y), \quad J_k = \begin{bmatrix} \frac{\partial f_1(\xi_1)}{\partial x_1} & \ldots & \frac{\partial f_1(\xi_1)}{\partial x_n} \\ \ldots & \ldots & \ldots \\ \frac{\partial f_n(\xi_n)}{\partial x_1} & \ldots & \frac{\partial f_n(\xi_n)}{\partial x_n} \end{bmatrix}$$

- $e_{k+1} = e_k - [f'(x_k)]^{-1}(f(z) - f(x_k)) = e_k - [f'(x_k)]^{-1} J_k e_k = [f'(x_k)]^{-1}(f'(x_k) - J_k)e_k$

Доказательство

- По теореме Лагранжа найдутся точки $\xi_1, \dots, \xi_n$:

$$f(x) - f(y) = J_k(x-y), \quad J_k = \begin{bmatrix} \frac{\partial f_1(\xi_1)}{\partial x_1} & \dots & \frac{\partial f_1(\xi_1)}{\partial x_n} \\ \dots & \dots & \dots \\ \frac{\partial f_n(\xi_n)}{\partial x_1} & \dots & \frac{\partial f_n(\xi_n)}{\partial x_n} \end{bmatrix}$$

- $e_{k+1} = e_k - [f'(x_k)]^{-1}(f(z) - f(x_k)) = e_k - [f'(x_k)]^{-1} J_k e_k = [f'(x_k)]^{-1}(f'(x_k) - J_k)e_k$
- По условию Липшица

$$\|f'(x_k) - J_k\|_\infty \le c \max_{1 \le j \le n} \|x_k - \xi_j\|_\infty \le c\|x_k - z\|_\infty$$

Доказательство

- По теореме Лагранжа найдутся точки $\xi_1, \dots, \xi_n$:

$$f(x) - f(y) = J_k(x - y), \quad J_k = \begin{bmatrix} \frac{\partial f_1(\xi_1)}{\partial x_1} & \dots & \frac{\partial f_1(\xi_1)}{\partial x_n} \\ \dots & \dots & \dots \\ \frac{\partial f_n(\xi_n)}{\partial x_1} & \dots & \frac{\partial f_n(\xi_n)}{\partial x_n} \end{bmatrix}$$

- $e_{k+1} = e_k - [f'(x_k)]^{-1}(f(z) - f(x_k)) = e_k - [f'(x_k)]^{-1} J_k e_k = [f'(x_k)]^{-1}(f'(x_k) - J_k)e_k$
- По условию Липшица

$$\|f'(x_k) - J_k\|_\infty \le c \max_{1 \le j \le n} \|x_k - \xi_j\|_\infty \le c\|x_k - z\|_\infty$$

- $\|e_{k+1}\|_\infty = \|[f'(x_k)]^{-1}(f'(x_k) - J_k)e_k\|_\infty \le \gamma \|e_k\|_\infty^2$

Метод Ньютона, решение линейной системы

Метод Ньютона, решение линейной системы

- $x_{k+1} = x_k - \left[f'(x_k)\right]^{-1} f(x_k)$

Метод Ньютона, решение линейной системы

- $x_{k+1} = x_k - \left[f'(x_k)\right]^{-1} f(x_k)$
- $\left[f'(x_k)\right] \Delta x_{k+1} = -f(x_k), \quad \Delta x_{k+1} = x_{k+1} - x_k$

Метод Ньютона, решение линейной системы

- $x_{k+1} = x_k - \left[f'(x_k)\right]^{-1} f(x_k)$
- $\left[f'(x_k)\right] \Delta x_{k+1} = -f(x_k), \quad \Delta x_{k+1} = x_{k+1} - x_k$
- При больших n – большая вычислительная сложность

Метод Ньютона, решение линейной системы

- $x_{k+1} = x_k - [f'(x_k)]^{-1} f(x_k)$
- $[f'(x_k)]\,\Delta x_{k+1} = -f(x_k), \quad \Delta x_{k+1} = x_{k+1} - x_k$
- При больших n – большая вычислительная сложность
- $[f'(x_k)]$ можно обновлять не на каждой итерации:

$$[f'(x_0)]\,\Delta x_{k+1} = -f(x_k)$$

Метод Ньютона, решение линейной системы

- $x_{k+1} = x_k - \left[f'(x_k)\right]^{-1} f(x_k)$
- $\left[f'(x_k)\right] \Delta x_{k+1} = -f(x_k), \quad \Delta x_{k+1} = x_{k+1} - x_k$
- При больших n – большая вычислительная сложность
- $\left[f'(x_k)\right]$ можно обновлять не на каждой итерации:

$$\left[f'(x_0)\right] \Delta x_{k+1} = -f(x_k)$$

- Имеет смысл использовать итерационные методы:

Метод Ньютона, решение линейной системы

- $x_{k+1} = x_k - \left[f'(x_k)\right]^{-1} f(x_k)$
- $\left[f'(x_k)\right] \Delta x_{k+1} = -f(x_k), \quad \Delta x_{k+1} = x_{k+1} - x_k$
- При больших n – большая вычислительная сложность
- $\left[f'(x_k)\right]$ можно обновлять не на каждой итерации:

$$\left[f'(x_0)\right] \Delta x_{k+1} = -f(x_k)$$

- Имеет смысл использовать итерационные методы:
  - $\Delta x_{k+1} \approx 0$ – хорошее начальное приближение

Метод Ньютона, решение линейной системы

- $x_{k+1} = x_k - [f'(x_k)]^{-1} f(x_k)$
- $[f'(x_k)]\,\Delta x_{k+1} = -f(x_k), \quad \Delta x_{k+1} = x_{k+1} - x_k$
- При больших n – большая вычислительная сложность
- $[f'(x_k)]$ можно обновлять не на каждой итерации:

$$[f'(x_0)]\,\Delta x_{k+1} = -f(x_k)$$

- Имеет смысл использовать итерационные методы:
  - $\Delta x_{k+1} \approx 0$ – хорошее начальное приближение
  - Не нужно решать систему точно

Метод Ньютона, решение линейной системы

- $x_{k+1} = x_k - [f'(x_k)]^{-1} f(x_k)$
- $[f'(x_k)]\, \Delta x_{k+1} = -f(x_k), \quad \Delta x_{k+1} = x_{k+1} - x_k$
- При больших n – большая вычислительная сложность
- $[f'(x_k)]$ можно обновлять не на каждой итерации:

$$[f'(x_0)]\, \Delta x_{k+1} = -f(x_k)$$

- Имеет смысл использовать итерационные методы:
  - $\Delta x_{k+1} \approx 0$ – хорошее начальное приближение
  - Не нужно решать систему точно
  - $[f'(x_k)]\, \Delta x \approx f(x_k + \Delta x) - f(x_k)$ – можно не хранить матрицу Якоби

Метод Ньютона и методы оптимизации

Метод Ньютона и методы оптимизации

- Рассмотрим функционал $f : \mathbb{R}^n \to \mathbb{R}$

Метод Ньютона и методы оптимизации

- Рассмотрим функционал $f:\mathbb{R}^n \to \mathbb{R}$
- Если $f \in C^2$ имеет единственную точку минимума – градиент в этой точке равен нулю:

$$\operatorname{grad} f(x) = \begin{bmatrix} \frac{\partial f(x)}{\partial x_1} \\ \vdots \\ \frac{\partial f(x)}{\partial x_n} \end{bmatrix} = 0$$

Метод Ньютона и методы оптимизации

- Рассмотрим функционал $f : \mathbb{R}^n \to \mathbb{R}$
- Если $f \in C^2$ имеет единственную точку минимума – градиент в этой точке равен нулю:

$$\operatorname{grad} f(x) = \begin{bmatrix} \frac{\partial f(x)}{\partial x_1} \\ \vdots \\ \frac{\partial f(x)}{\partial x_n} \end{bmatrix} = 0$$

- Для решения можно применить метод Ньютона:

$$x_{k+1} = x_k - \left[f''(x_k)\right]^{-1} \operatorname{grad} f(x_k)$$

Метод Ньютона и методы оптимизации

- Рассмотрим функционал $f : \mathbb{R}^n \to \mathbb{R}$
- Если $f \in C^2$ имеет единственную точку минимума – градиент в этой точке равен нулю:

$$\operatorname{grad} f(x) = \begin{bmatrix} \frac{\partial f(x)}{\partial x_1} \\ \vdots \\ \frac{\partial f(x)}{\partial x_n} \end{bmatrix} = 0$$

- Для решения можно применить метод Ньютона:

$$x_{k+1} = x_k - \left[f''(x_k)\right]^{-1} \operatorname{grad} f(x_k)$$

- $f''(x) = H(x) = [\operatorname{grad} f(x_k)]' = \left[\frac{\partial f}{\partial x_i \partial x_j}\right]$ - гессиан

Метод Ньютона и методы оптимизации

- Рассмотрим функционал $f : \mathbb{R}^n \to \mathbb{R}$
- Если $f \in C^2$ имеет единственную точку минимума – градиент в этой точке равен нулю:

$$\operatorname{grad} f(x) = \begin{bmatrix} \frac{\partial f(x)}{\partial x_1} \\ \vdots \\ \frac{\partial f(x)}{\partial x_n} \end{bmatrix} = 0$$

- Для решения можно применить метод Ньютона:

$$x_{k+1} = x_k - [f''(x_k)]^{-1} \operatorname{grad} f(x_k)$$

- $f''(x) = H(x) = [\operatorname{grad} f(x_k)]' = \left[\frac{\partial f}{\partial x_i \partial x_j}\right]$ - гессиан
- $H(x) = H(x)^T$, в точке минимума $H(z) > 0$

Метод секущих

Идея методов высокого порядка

Локально приблизим f или f^{-1} многочленом, и для многочлена решим задачу точно.

Метод секущих

Идея методов высокого порядка

Локально приблизим f или f^{-1} многочленом, и для многочлена решим задачу точно.

- По двум последним приближениям x_{k-1}, x_k построим линейный интерполянт:

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

Метод секущих

Идея методов высокого порядка

Локально приблизим f или f^{-1} многочленом, и для многочлена решим задачу точно.

- По двум последним приближениям x_{k-1}, x_k построим линейный интерполянт:

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

- Корень многочлена:

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

Скорость сходимости метода секущих

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

Скорость сходимости метода секущих

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

- Пусть z - корень, $e_k \equiv z - x_k$, $f \in C^2$, $f'(z) \neq 0$

Скорость сходимости метода секущих

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

- Пусть z - корень, $e_k \equiv z - x_k$, $f \in C^2$, $f'(z) \neq 0$
-

$$\underbrace{f(z)}_{0} - L(z) = \frac{f''(\xi_k)}{2}(z - x_k)(z - x_{k-1})$$

$$\underbrace{L(x_{k+1})}_{0} - L(z) = \frac{f(x_k) - f(x_{k-1})}{x_k - x_{k-1}}(x_{k+1} - z) =$$

$$= f'(\zeta_k)(x_{k+1} - z)$$

Скорость сходимости метода секущих

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

- Пусть z - корень, $e_k \equiv z - x_k$, $f \in C^2$, $f'(z) \neq 0$
- $$\underbrace{f(z)}_{0} - L(z) = \frac{f''(\xi_k)}{2}(z - x_k)(z - x_{k-1})$$
  $$\underbrace{L(x_{k+1})}_{0} - L(z) = \frac{f(x_k) - f(x_{k-1})}{x_k - x_{k-1}}(x_{k+1} - z) =$$
  $$= f'(\zeta_k)(x_{k+1} - z)$$
- $e_{k+1} = -\dfrac{f''(\xi_k)}{2f'(\zeta_k)}e_k e_{k-1}$

Скорость сходимости м-да секущих(2)

$$e_{k+1} = -\frac{f''(\xi_k)}{2f'(\zeta_k)} e_k e_{k-1}$$

Скорость сходимости м-да секущих(2)

$$e_{k+1} = -\frac{f''(\xi_k)}{2f'(\zeta_k)} e_k e_{k-1}$$

- Предположим, что метод сходится

Скорость сходимости м-да секущих(2)

$$e_{k+1} = -\frac{f''(\xi_k)}{2f'(\zeta_k)} e_k e_{k-1}$$

- Предположим, что метод сходится
- для некоторого $\gamma > 0$: $|e_{k+1}| \leq \gamma |e_k||e_{k-1}|$

Скорость сходимости м-да секущих(2)

$$e_{k+1} = -\frac{f''(\xi_k)}{2f'(\zeta_k)} e_k e_{k-1}$$

- Предположим, что метод сходится
- для некоторого $\gamma > 0$: $|e_{k+1}| \leq \gamma |e_k||e_{k-1}|$
- Введем величины $d_k = \gamma |e_k|$ и предположим, что $d_0 \leq d < 1$, $d_1 \leq d < 1$

Скорость сходимости м-да секущих(2)

$$e_{k+1} = -\frac{f''(\xi_k)}{2f'(\zeta_k)} e_k e_{k-1}$$

- Предположим, что метод сходится
- для некоторого $\gamma > 0$: $|e_{k+1}| \leq \gamma |e_k||e_{k-1}|$
- Введем величины $d_k = \gamma |e_k|$ и предположим, что $d_0 \leq d < 1$, $d_1 \leq d < 1$
- $d_2 \leq d_1 d_0 \leq d^2$

Скорость сходимости м-да секущих(2)

$$e_{k+1} = -\frac{f''(\xi_k)}{2f'(\zeta_k)} e_k e_{k-1}$$

- Предположим, что метод сходится
- для некоторого $\gamma > 0$: $|e_{k+1}| \leq \gamma |e_k||e_{k-1}|$
- Введем величины $d_k = \gamma |e_k|$ и предположим, что $d_0 \leq d < 1$, $d_1 \leq d < 1$
- $d_2 \leq d_1 d_0 \leq d^2$
- $d_3 \leq d_2 d_1 \leq d^3 \ldots$

Скорость сходимости м-да секущих(2)

$$e_{k+1} = -\frac{f''(\xi_k)}{2f'(\zeta_k)} e_k e_{k-1}$$

- Предположим, что метод сходится
- для некоторого $\gamma > 0$: $|e_{k+1}| \leq \gamma |e_k||e_{k-1}|$
- Введем величины $d_k = \gamma |e_k|$ и предположим, что $d_0 \leq d < 1$, $d_1 \leq d < 1$
- $d_2 \leq d_1 d_0 \leq d^2$
- $d_3 \leq d_2 d_1 \leq d^3 \ldots$
- $d_k \leq d^{\phi_k}$, $\phi_0 = \phi_1 = 1$, $\phi_k = \phi_{k-1} + \phi_{k-2}$, $k = 2, \ldots$

Скорость сходимости м-да секущих(2)

$$e_{k+1} = -\frac{f''(\xi_k)}{2f'(\zeta_k)} e_k e_{k-1}$$

- Предположим, что метод сходится
- для некоторого $\gamma > 0$: $|e_{k+1}| \leq \gamma |e_k||e_{k-1}|$
- Введем величины $d_k = \gamma |e_k|$ и предположим, что $d_0 \leq d < 1$, $d_1 \leq d < 1$
- $d_2 \leq d_1 d_0 \leq d^2$
- $d_3 \leq d_2 d_1 \leq d^3 \ldots$
- $d_k \leq d^{\phi_k}$, $\phi_0 = \phi_1 = 1$, $\phi_k = \phi_{k-1} + \phi_{k-2}$, $k = 2, \ldots$
- $\phi_k = \frac{1}{\sqrt{5}} \left(\left(\frac{1+\sqrt{5}}{2} \right)^{k+1} - \left(\frac{1-\sqrt{5}}{2} \right)^{k+1} \right) = O(1.618^k)$

Скорость сходимости м-да секущих(2)

$$e_{k+1} = -\frac{f''(\xi_k)}{2f'(\zeta_k)} e_k e_{k-1}$$

- Предположим, что метод сходится
- для некоторого $\gamma > 0$: $|e_{k+1}| \leq \gamma |e_k||e_{k-1}|$
- Введем величины $d_k = \gamma |e_k|$ и предположим, что $d_0 \leq d < 1$, $d_1 \leq d < 1$
- $d_2 \leq d_1 d_0 \leq d^2$
- $d_3 \leq d_2 d_1 \leq d^3 \dots$
- $d_k \leq d^{\phi_k}$, $\phi_0 = \phi_1 = 1$, $\phi_k = \phi_{k-1} + \phi_{k-2}$, $k = 2, \dots$
- $\phi_k = \frac{1}{\sqrt{5}} \left(\left(\frac{1+\sqrt{5}}{2} \right)^{k+1} - \left(\frac{1-\sqrt{5}}{2} \right)^{k+1} \right) = O(1.618^k)$
- Для метода Ньютона: $d_k \leq d^{2^k}$

Методы высоких порядков

- Метод третьего порядка

$$x_{n+1} = x_n - \frac{f_n}{(f')_n} - \frac{(f'')_n}{2(f')_n^3} f_n^2$$

- Метод четвертого порядка

$$x_{n+1} = x_n - \frac{f_n}{(f')_n} - \frac{(f'')_n}{2(f')_n^3} f_n^2 - \frac{(f'')_n^2}{2(f')_n^5} f_n^3 + \frac{(f''')_n^2}{6(f')_n^4} f_n^3$$

Заключение

Заключение

- Метод простой итерации

Заключение

- Метод простой итерации
- Принцип сжимающих отображений

Заключение

- Метод простой итерации
- Принцип сжимающих отображений
- Метод Ньютона

Заключение

- Метод простой итерации
- Принцип сжимающих отображений
- Метод Ньютона
- Идея построение методов высокого порядка

Заключение

- Метод простой итерации
- Принцип сжимающих отображений
- Метод Ньютона
- Идея построение методов высокого порядка
- Метод секущих