Нелинейные методы

(конспект Ивана Кобзаря, МФТИ)

1 Введение

Решается система уравнений вида,

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

где

$$f_i(x_1, ..., x_n) = 0, \quad i = 1, ..., n \tag{2}$$

В случае когда нет прямого метода для решения этой системы уравнений, нужно воспользоваться итеративным подходом.
Пусть x_k - текущее приближение; x - точное решение. Тогда определим понятия скорости сходимости p и фактора сходимости c, введя соотношение:

$$\frac{|x_{k+1} - x|}{|x_k - x|^p} \to c = const \tag{3}$$

Теперь можно определить понятия:
Линейные методы ($p = 1$) и **Квадратичные методы** ($p = 2$).
Также определим сходимость двух типов:
Локальная (методы сходятся при $x_0 \sim x$) и **Глобальная** (методы сходятся $\forall x_0$).

2 Метод Ньютона

Введем якобиан для преобразования (1):

$$J = \frac{df}{dx} = \begin{bmatrix} \frac{df_1}{dx_1} & \cdots & \frac{df_1}{dx_n} \\ \vdots & \ddots & \vdots \\ \frac{df_n}{dx_1} & \cdots & \frac{df_n}{dx_n} \end{bmatrix} \tag{4}$$

Определим также приращение

$$dx_k = x_{k+1} - x_k \tag{5}$$

И раздожение тейлора, где $J_k = J(x_k)$

$$f(x_k + dx_k) \approx f(x) = f(x_k) + J_k dx_k \tag{6}$$

Пользуясь (1), получаем следующую итерационную систему:

$$\begin{cases} J_k dx_k = -f(x_k) \\ x_{k+1} = dx_k = x_k \end{cases} \tag{7}$$

Предполагая отсутствие перегибов и непрерывность по Липшицу f, найдём фактор и скорость сходимости. Подставив разложением тейлора до второго порядка

$$f(x) = f(x_k) + (x - x_k)f^{'}(x_k) + \frac{1}{2}(x - x_k)^2 f^{''}(\xi), \quad \xi \in [x_0, x_1]. \tag{8}$$

в систему (7), получим равенство вида (3) для оценки скорости сходимости:

$$\frac{|x_{k+1}-x|}{|x_k-x|^2} = \left|\frac{f''(\xi)}{f'(x)}\right| \le M = \sup_{a,b\in[x-|x_0-x|,x+|x_0-x|]}\left|\frac{f''(a)}{f'(b)}\right|. \tag{9}$$

Таким образом, метод Ньютона - квадратичный, но локальный с условием сходимости

$$|x-x_0| < M^{-1} \tag{10}$$

3 Метод Халли

Введём гессиан функции:

$$H = \frac{\partial^2 f}{\partial x^2} \tag{11}$$

Рассмотрим двухшаговый метод:

$$\begin{cases} 1)\ J_k\widetilde{dx_k} = -f(x_k) \\ 2)\ (J_k + \frac{1}{2}H_k\widetilde{dx_k})dx_k = -f(x_k) \\ 3)\ x_{k+1} = x_k + dx_k \end{cases} \tag{12}$$

3.1 Поиск вдоль прямой

Пусть dx_k известно. Введём

$$\alpha = argmin\|f(x_k+\alpha dx_k)\| \tag{13}$$

Способы определения α:

1) Метод бисекций ($\alpha \in [0,1]$):

2) Метод золотого сечения:

3) Итерационный метод по правилам Армихо:

$$\begin{cases} 1)\ f(x_k+\alpha dx_k) \le f(x_k) + c_1\alpha_k dx_k^T J_k \\ 2) dx_k^T J(x_k+\alpha_k dx_k) \le -c_2 dx_k^T J_k \\ 3)\ 0 \le c_1 \le c_2 \le 1 \end{cases} \tag{14}$$

4 Метод ускорения сходимости Андерсена

Пусть $x_{k-m}, ..., x_k$ - m приближений к x, $f(x_{k-m}), ..., f(x_k)$ - m функций. Введём

$$\lambda_{k+1} = \sum_{j=k-m}^{k}\alpha_j x_j + Y\sum_{j=k-m}^{k}\alpha_j f(x_j) \tag{15}$$

Где Y- параметр, $\alpha : \|f(x_{k+1})\| \to min \quad$ при $\sum_{j=k-m}^{k}\alpha_j = 1$. Введём λ - множитель Лагранжа

$$\begin{bmatrix} RR^T & \beta\vec{1} \\ \beta\vec{1}^T & 0 \end{bmatrix}\begin{bmatrix} \alpha \\ \lambda \end{bmatrix} = \begin{bmatrix} 0 \\ \beta \end{bmatrix} \tag{16}$$

где

$$\beta = trace(RR^T), \tag{17}$$

$$R = \begin{bmatrix} f(x_{k-m}) \\ \vdots \\ f(x_k) \end{bmatrix} \tag{18}$$

Из (16) Вычисляем α и после этого находим x_{k+1}