УДК 519.6

ОБ ЭФФЕКТИВНОМ ПОИСКЕ БЕЗУСЛОВНОГО ЭКСТРЕМУМА ГЛАДКИХ ФУНКЦИОНАЛОВ В КОНЕЧНОМЕРНЫХ ЗАДАЧАХ

Пастухов Д.Ф., Пастухов Ю.Ф.

Полоцкий государственный университет

Аннотация: *Получены условия быстрой сходимости простой итерации в конечномерной задаче на экстремум для функционала, имеющего третий порядок гладкости. Приведены формулы необходимого числа итераций для достижения заданной наибольшей точности координат стационарной точки и значения функционала в стационарной точке.Во второй теореме выведена разностная итерационная формула, рассматривающая априорно гладкие функционалы, возможно не представимые в виде композиции элементарных функций. Показана эквивалентность порядка точности итерационных формул в обеих теоремах, определена верхняя граница оптимального шага. Приведены программы и примеры, подтверждающие эффективность методов для поиска точек экстремума и точек перевала.*

Ключевые слова: Гладкий функционал, строгое диагональное преобладание элементов матрицы Гессе, центральная разность первого порядка, оптимальный шаг итерационной формулы.

About efficient searching for of the unconditional extremum smooth function in certainly measured problem

Pastuhov D.F., Pastuhov YU.F.

(Polotsk state university)

The Abstract: are Received condition to quick convergence iteration idle time in certainly measured problem on extremum for function, having third order to smoothness. The Broughted formulas of the necessary number iteration for achievement given most accuracy of the coordinates of the stationary point and importances of the function in stationary point.In the second theorem is received iteration , considering a priori smooth to functions, possible not presented in the manner of compositions elementary function. It Is Shown equivalence of the order to accuracy molded in both theorem, is determined upper border of the optimum step.The Broughted programs and examples, confirming efficiency of the methods for searching for point extremum and point of the mouting pass.

The Keywords: Smooth function, diagonal prevalence of element matrixes Gesse, central first-order difference, optimum step of iteration idle time.

Введение

В данной работе рассматриваются эффективные методы поиска безусловного экстремума гладких функционалов конечного числа переменных.

Градиентные методы поиска экстремума являются более распространёнными в численных методах, так как решение принадлежит широкому классу непрерывно дифференцируемых функций. Однако, градиентные методы не достаточно эффективны с точки зрения точности, на практике фактически невозможно получить градиентными методами решение с двойной точностью, на что и указывает В.М. Тихомиров[1,2,3]. Возможно, потому, что итерационные точки ложатся по разные стороны от стационарной точки, в то время как в формуле касательных Ньютона итерации находятся по одну сторону. Как показано в данной работе для случая трижды непрерывно дифференцируемых функционалов и строгим диагональным преобладанием матрицы Гессе можно получить решение и экстремальное значение с относительной точностью 10^{-16} (двойной точностью double для чисел с плавающей запятой)всего за 60 итераций при начальным удалении от стационарной точки в100 единиц. Наиболее быстрые итерационные формулы удаётся построить при совмещении матричного метода Ньютона и метода Зейделя [2].В матричном методе Ньютона нужно вычислить n^2 элементов матрицы Якоби, погрешность дополнительно увеличивается при отыскании обратной матрицы к матрице Якоби

$x^{m+1} = x^m - \left(F'\left(x^m\right)\right)^{-1} * F\left(x^m\right)$, где $\left(F'\left(x^m\right)\right)^{-1}$ - матрица обратная к матрице Якоби.

Метод Зейделя использует только диагональные элементы матрицы Якоби, т.е. n элементов. Поэтому при одинаковом числе итераций в формуле Зейделя - Ньютона меньше элементарных операций, чем в маричной формуле Ньютона и, следовательно, ошибка округления.

Определение 1. Среди двух методов, решающих одну задачу при одинаковых условиях, назовём ***более эффективным*** методом метод, дающий большую точность. При одинакой достигнутой точности решения назовём ***более эффективным*** метод с меньшим числом элементарных арифметических операций (сложением, вычитанием, умножением, делением).

Постановка задачи

Пусть в R^n задана открытая область $A \subset R^n$, вектор $x = \{x_1, x_2, \ldots, x_n\} \in A$. Рассмотрим задачу на безусловный экстремум достаточно гладкого функционала $f : R^n \to R$, а именно множество функционалов трижды непрерывно дифференцируемых в открытой области A ,т.е. имеющих непрерывные частные производные до 3 – го порядка включительно[1].

$$\begin{cases} f(x_1, x_2, \ldots, x_n) \to extr \\ f(x) \in C^3(A) \end{cases} \quad (1)$$

Для поиска экстремальных точек задачи (1) можно использовать 2 подхода. Первый заключается в прямом исследовании основного функционала, например, градиентными методами. Другой подход заключается в использовании необходимых условий экстремума функции нескольких переменных[1]:

$$\begin{cases} f'_{x_1}(x_1, x_2, \ldots, x_n) \equiv F_1(x_1, x_2, \ldots, x_n) = 0 \\ f'_{x_2}(x_1, x_2, \ldots, x_n) \equiv F_2(x_1, x_2, \ldots, x_n) = 0 \\ \qquad \ldots \\ f'_{x_n}(x_1, x_2, \ldots, x_n) \equiv F_n(x_1, x_2, \ldots, x_n) = 0 \end{cases} \quad (2)$$

где $\overline{x} = (\overline{x_1}, \overline{x_2}, \ldots, \overline{x_n}) \in A \subset R^n$ - стационарная точка, т.е. решение системы уравнений (2).

Обозначим $x_i^m, i = \overline{1,n}$ координату итерационной точки, у которой нижний индекс i (номер координаты) и с номером итерации $m = 0,1,2,\ldots$ (верхний индекс).Будем решать систему уравнений (2**)** **численно** методом простой итерации[2].

Метод Зейделя для системы уравнений заданных в неявном виде:[2]:

$$\begin{cases} F_1\left(x_1^{m+1}, x_2^m, \ldots, x_n^m\right) = 0, \\ F_2\left(x_1^{m+1}, x_2^{m+1}, \ldots, x_n^m\right) = 0, \\ \qquad \ldots \\ F_n\left(x_1^{m+1}, x_2^{m+1}, \ldots, x_{n-1}^{m+1}, x_n^{m+1}\right) = 0 \end{cases} \quad (3)$$

То есть исходная система (2) сводится к последовательному решению n уравнений системы (3), каждое из которых $F_i\left(x_1^{m+1}, x_2^{m+1}, \ldots, x_i^{m+1}, \ldots, x_n^m\right) = 0$, $i = \overline{1,n}$ представляет уравнение с одной неизвестной x_i^{m+1} переменной, а все остальные переменные при фиксированной строке остаются “замороженными”, т.е. постоянными. В этом случае для нахождения i - ой переменной можно использовать i - ое уравнение с явным видом итерации – формулу касательных Ньютона для уравнения с одной неизвестной переменной:

$$\begin{cases} x_1^{m+1} = x_1^m - \dfrac{F_1(x_1^m, x_2^m, \ldots, x_n^m)}{F'_{1x_1}(x_1^m, x_2^m, \ldots, x_n^m)}, \\ x_2^{m+1} = x_2^m - \dfrac{F_2(x_1^{m+1}, x_2^m, \ldots, x_n^m)}{F'_{2x_2}(x_1^{m+1}, x_2^m, \ldots, x_n^m)}, \\ \qquad \ldots. \\ x_i^{m+1} = x_i^m - \dfrac{F_i(x_1^{m+1}, x_2^{m+1}, \ldots, x_{i-1}^{m+1}, x_i^m, \ldots, x_n^m)}{F'_{ix_i}(x_1^{m+1}, x_2^{m+1}, \ldots, x_{i-1}^{m+1}, x_i^m, \ldots, x_n^m)}, \\ \qquad \ldots. \\ x_n^{m+1} = x_n^m - \dfrac{F_n(x_1^{m+1}, x_2^{m+1}, \ldots, x_{i-1}^{m+1}, x_i^{m+1}, \ldots, x_{n-1}^{m+1}, x_n^m)}{F'_{nx_n}(x_1^{m+1}, x_2^{m+1}, \ldots, x_{i-1}^{m+1}, x_i^{m+1}, \ldots, x_{n-1}^{m+1}, x_n^m)} \end{cases} \quad (4)$$

Обозначим вектор (невязку, погрешность) $\overline{\delta x^m} = \left(\delta x_1{}^m, \delta x_2{}^m, \ldots, \delta x_n{}^m\right) = \left(x_1^m - \overline{x_1}, x_2^m - \overline{x_2}, \ldots, x_n^m - \overline{x_n}\right)$- разность между итерационной точкой x^m на шаге итерации с номером m и стационарной точкой $\overline{x}$, аналогично:

$\delta x^{m+1} = \left(\delta x_1{}^{m+1}, \delta x_2{}^{m+1}, \ldots, \delta x_n{}^{m+1}\right) = \left(x_1^{m+1} - \overline{x_1}, x_2^{m+1} - \overline{x_2}, \ldots, x_n^{m+1} - \overline{x_n}\right)$ невязка итерации на шаге итерации с номером $m+1$.

Пусть система итерационных уравнений (4) имеет предельную точку $\overline{x}$,в силу (2) и (4), получим

$$\begin{cases} \lim\limits_{m\to\infty} x_1^{m+1} = \lim\limits_{m\to\infty} x_1^m = \overline{x_1} \Leftrightarrow f'_{x_1}(\overline{x_1}, \overline{x_2}, \ldots, \overline{x_n}) = 0, \ldots, \lim\limits_{m\to\infty} x_n^{m+1} = \lim\limits_{m\to\infty} x_n^m = \overline{x_n} \Leftrightarrow f'_{x_n}(\overline{x_1}, \overline{x_2}, \ldots, \overline{x_n}) = 0 \\ f''_{x_i x_i}(\overline{x_1}, \overline{x_2}, \ldots, \overline{x_n}) \neq 0, i = \overline{1,n} \end{cases} \quad (5)$$

Вычтем из правой и из левой части (4) каждого i - го уравнения $\overline{x_i}$, $\delta x_i^{m+1} = x_i^{m+1} - \overline{x_i}$, $\delta x_i^m = x_i^m - \overline{x_i}$. Кроме того, аргументы частных производных $f_{x_i}^{'}$, $f_{x_i x_i}^{''}$ с порядками m и $m+1$ выразим через невязки δx_k^m , δx_k^{m+1} :

$$x_k^{m+1} = \overline{x_k} + \delta x_k^{m+1}, x_k^m = \overline{x_k} + \delta x_k^m, k = \overline{1,n}$$

$$\begin{cases} \delta x_1^{m+1} = \delta x_1^m - \dfrac{f_{x_1}^{'}(\overline{x_1} + \delta x_1^m, \overline{x_2} + \delta x_2^m, \ldots, \overline{x_n} + \delta x_n^m)}{f_{x_1 x_1}^{''}(\overline{x_1} + \delta x_1^m, \overline{x_2} + \delta x_2^m, \ldots, \overline{x_n} + \delta x_n^m)}, \\ \delta x_2^{m+1} = \delta x_2^m - \dfrac{f_{x_2}^{'}(\overline{x_1} + \delta x_1^{m+1}, \overline{x_2} + \delta x_2^m, \ldots, \overline{x_n} + \delta x_n^m)}{f_{x_2 x_2}^{''}(\overline{x_1} + \delta x_1^{m+1}, \overline{x_2} + \delta x_2^m, \ldots, \overline{x_n} + \delta x_n^m)}, \\ \quad \ldots. \\ \delta x_i^{m+1} = \delta x_i^m - \dfrac{f_{x_i}^{'}(\overline{x_1} + \delta x_1^{m+1}, \overline{x_2} + \delta x_2^{m+1}, \ldots, \overline{x_{i-1}} + \delta x_{i-1}^{m+1}, \overline{x_i} + \delta x_i^m, \ldots, \overline{x_n} + \delta x_n^m)}{f_{x_i x_i}^{''}(\overline{x_1} + \delta x_1^{m+1}, \overline{x_2} + \delta x_2^{m+1}, \ldots, \overline{x_{i-1}} + \delta x_{i-1}^{m+1}, \overline{x_i} + \delta x_i^m, \ldots, \overline{x_n} + \delta x_n^m)}, \\ \quad \ldots. \\ \delta x_n^{m+1} = \delta x_n^m - \dfrac{f_{x_n}^{'}(\overline{x_1} + \delta x_1^{m+1}, \overline{x_2} + \delta x_2^{m+1}, \ldots, \overline{x_{n-1}} + \delta x_{n-1}^{m+1}, \overline{x_n} + \delta x_n^m)}{f_{x_n x_n}^{''}(\overline{x_1} + \delta x_1^{m+1}, \overline{x_2} + \delta x_2^{m+1}, \ldots, \overline{x_{n-1}} + \delta x_{n-1}^{m+1}, \overline{x_n} + \delta x_n^m)} \end{cases} \quad (6)$$

Система итерационных последовательностей (4) сходится к стационарной точке $\overline{x}$ если и только если невязки в системе итерационных уравнений (6) сходятся к нулю $\lim_{m\to\infty} \delta x_i^{m+1} = \lim_{m\to\infty} \delta x_i^m = 0, i = \overline{1,n}$. То есть системы уравнений (4) и (6)эквивалентны.

Определение 2. Говорят[2], что числовая итерационная последовательность x^{m+1} сходится к предельной точке $\overline{x}$ с порядком скорости p , если $\exists\, C, p > 0 : \lim_{m\to\infty} \left(\sup_{p>0} \dfrac{\left|\delta x^{m+1}\right|}{\left|\delta x^m\right|^p} \le C \right)$.

Сходимость метода

Теорема 1. (условия сходимости итерации (6)).

Пусть открытая область $A \subset R^n$ содержит начальную итерацию $\left(x_1^0, x_2^0, \ldots, x_n^0\right) \in A$ и стационарную точку $\overline{x} = \left(\overline{x_1}, \overline{x_2}, \ldots, \overline{x_n}\right) \in A$ (решение системы уравнений (2)). Функция $f(x), x \in A \subset R^n$ конечного числа n переменных:

1) Трижды непрерывно дифференцируема $f(x) \in C^3(A)$

2) Пусть матрица вторых частных производных (матрица Гессе) с элементами $H_{i,j} = f_{x_i x_j}^{''}(x)$ обладает строгим диагональным преобладанием:

$$\left| f_{x_i x_i}^{''}(x) \right| > \sum_{j=1, j\ne i}^{n} \left| f_{x_i x_j}^{''}(x) \right|, \forall i = \overline{1,n}, \forall x \in A$$

Обозначим $q_i = \sup_{x\in A} \dfrac{\sum_{j=1, j\ne i}^{n} \left| f_{x_i x_j}^{''}(x) \right|}{\left| f_{x_i x_i}^{''}(x) \right|} < 1, i = \overline{1,n}$ в силу диагонального преобладания

Обозначим $q = \max_{i=\overline{1,n}} q_i < 1$

Тогда система уравнений (6) сходится к единственной стационарной точке(5), по крайней мере, с первым порядком скорости ($C = q, p = 1$) и имеет место оценка погрешности после m итераций:

$\left|\delta x^m\right| = \left|x^m - \overline{x}\right| \le \dfrac{(1+q)\left|\delta x^0\right| q^m}{1-q} \le \dfrac{(1+q)q^m}{\left(1-q\right)^2} l_0$, где $l_0 = \left|x^1 - x^0\right|$ - расстояние между начальными итерациями x^0, x^1 , $\delta x^0 = \left(x_1^0 - \overline{x_1}, x_2^0 - \overline{x_2}, \ldots, x_n^0 - \overline{x_n}\right)$.

Доказательство проведём по индукции (достаточность).

Разложим последовательно первую частную производную $f'_{x_i}(\overline{x_1}+\delta x_1^{m+1},\overline{x_2}+\delta x_2^{m+1},\dots,\overline{x_{i-1}}+\delta x_{i-1}^{m+1},\overline{x_i}+\delta x_i^{m},\dots,\overline{x_n}+\delta x_n^{m}),\ i=\overline{1,n}$, входящую в каждое уравнение системы (6) в ряд Тейлора с центром в стационарной точке, для первого уравнения имеем:

1) $i=1$:

$f'_{x_1}(x_1^m,x_2^m,\dots,x_n^m)=f'_{x_1}(\overline{x_1},\overline{x_2},\dots,\overline{x_n})+\sum_{j=1}^{n}f''_{x_1x_j}(\overline{x})\delta x_j^m+O\left(\left|\delta x^m\right|^2\right)=\sum_{j=1}^{n}f''_{x_1x_j}(\overline{x})\delta x_j^m+O\left(\left|\delta x^m\right|^2\right)$ в силу формулы (5). Где

$$\left|\delta x^m\right|=\max_{j=\overline{1,n}}\left|\delta x_j^m\right|,\quad \left|\delta x_j^m\right|\le\left|\delta x^m\right|,\forall j=\overline{1,n},\forall m=0,1,2,\dots. \tag{7}$$

В силу условий (5),(6),(7) и условия 1) теоремы преобразуем дробь в правой части первого уравнения (6)

$$\delta x_1^{m+1}=\delta x_1^m-\frac{f'_{x_1}(\overline{x_1}+\delta x_1^m,\overline{x_2}+\delta x_2^m,\dots,\overline{x_n}+\delta x_n^m)}{f''_{x_1x_1}(\overline{x_1}+\delta x_1^m,\overline{x_2}+\delta x_2^m,\dots,\overline{x_n}+\delta x_n^m)}=\delta x_1^m-\frac{f'_{x_1}\left(\overline{x}\right)+\sum_{j=1}^{n}f''_{x_1x_j}(\overline{x})\delta x_j^m+O\left(\left|\delta x^m\right|^2\right)}{f''_{x_1x_1}(\overline{x_1}+\delta x_1^m,\overline{x_2}+\delta x_2^m,\dots,\overline{x_n}+\delta x_n^m)}=$$

$$\delta x_1^m-\frac{\sum_{j=1}^{n}f''_{x_1x_j}(\overline{x})\delta x_j^m+O\left(\left|\delta x^m\right|^2\right)}{f''_{x_1x_1}(\overline{x_1}+\delta x_1^m,\overline{x_2}+\delta x_2^m,\dots,\overline{x_n}+\delta x_n^m)}=\delta x_1^m-\frac{\delta x_1^m}{\dfrac{f''_{x_1x_1}(\overline{x_1}+\delta x_1^m,\overline{x_2}+\delta x_2^m,\dots,\overline{x_n}+\delta x_n^m)}{f''_{x_1x_1}(\overline{x_1},\overline{x_2},\dots,\overline{x_n})}}-$$

(В силу условия 1) теоремы) $f(x)\in C^3(A)$:

$$\frac{\sum_{j=2}^{n}f''_{x_1x_j}(\overline{x})\delta x_j^m}{f''_{x_1x_1}(\overline{x_1}+\delta x_1^m,\overline{x_2}+\delta x_2^m,\dots,\overline{x_n}+\delta x_n^m)}+O\left(\left|\delta x^m\right|^2\right)=\delta x_1^m-\frac{\delta x_1^m}{1+\dfrac{\sum_{j=1}^{n}f^{(3)}_{x_1x_1x_j}(\overline{x_1},\overline{x_2},\dots,\overline{x_n})\delta x_j^m}{f''_{x_1x_1}(\overline{x_1},\overline{x_2},\dots,\overline{x_n})}}+O\left(\left|\delta x^m\right|^2\right)-$$

$$-\frac{\sum_{j=2}^{n}f''_{x_1x_j}(\overline{x})\delta x_j^m}{f''_{x_1x_1}(\overline{x_1}+\delta x_1^m,\overline{x_2}+\delta x_2^m,\dots,\overline{x_n}+\delta x_n^m)}=\delta x_1^m-\delta x_1^m-\frac{\sum_{j=2}^{n}f''_{x_1x_j}(\overline{x})\delta x_j^m}{f''_{x_1x_1}(\overline{x_1}+\delta x_1^m,\overline{x_2}+\delta x_2^m,\dots,\overline{x_n}+\delta x_n^m)}+O\left(\left|\delta x^m\right|^2\right)$$

$$\delta x_1^{m+1}=-\frac{\sum_{j=2}^{n}f''_{x_1x_j}(\overline{x})\delta x_j^m}{f''_{x_1x_1}(\overline{x_1}+\delta x_1^m,\overline{x_2}+\delta x_2^m,\dots,\overline{x_n}+\delta x_n^m)}+O\left(\left|\delta x^m\right|^2\right).$$

Оценим последнее выражение по модулю (в следующей оценке использовано неравенство треугольника для модуля суммы величин и неравенство (7)):

$$\left|\delta x_1^{m+1}\right|\le\frac{\sum_{j=2}^{n}\left|f''_{x_1x_j}(\overline{x})\right|\left|\delta x_j^m\right|}{\left|f''_{x_1x_1}(\overline{x_1}+\delta x_1^m,\overline{x_2}+\delta x_2^m,\dots,\overline{x_n}+\delta x_n^m)\right|}+O\left(\left|\delta x^m\right|^2\right)\le\frac{\left|\delta x^m\right|\sum_{j=2}^{n}\left|f''_{x_1x_j}(\overline{x})\right|}{\left|f''_{x_1x_1}(\overline{x_1}+\delta x_1^m,\overline{x_2}+\delta x_2^m,\dots,\overline{x_n}+\delta x_n^m)\right|}+O\left(\left|\delta x^m\right|^2\right)$$

Используя условие 2) теоремы 1:

$$\left|f''_{x_1x_1}(\overline{x_1}+\delta x_1^m,\overline{x_2}+\delta x_2^m,\dots,\overline{x_n}+\delta x_n^m)\right|>\sum_{j=1,\,j\ne1}^{n}\left|f''_{x_1x_j}(x)\right|=\sum_{j=2}^{n}\left|f''_{x_1x_j}(x)\right|\forall x\in A,$$

$$\left|\delta x_1^{m+1}\right|\le\left|\delta x^m\right|q_1+O\left(\left|\delta x^m\right|^2\right)\le\left|\delta x^m\right|q+O\left(\left|\delta x^m\right|^2\right)<\left|\delta x^m\right|+O\left(\left|\delta x^m\right|^2\right) \tag{8}$$

3) Используя разложение в ряд Тейлора, преобразуем произвольное $i-$е уравнение системы (6) в окрестности стационарной точки.

По индукции предположим выполнение неравенств $\left|\delta x_k^{m+1}\right|\le\left|\delta x^m\right|,k=\overline{1,i-1}$, тогда повторяя преобразование с $i-$м уравнением системы

$$f'_{x_i}(x_1^{m+1},x_2^{m+1},\dots,x_i^m,\dots,x_n^m)=f'_{x_i}(\overline{x_1},\overline{x_2},\dots,\overline{x_n})+\sum_{j=1}^{i-1}f''_{x_ix_j}(\overline{x})\delta x_j^{m+1}+\sum_{j=i}^{n}f''_{x_ix_j}(\overline{x})\delta x_j^m+O\left(\left|\delta x^m\right|^2\right)=$$

$$=\sum_{j=1}^{i-1}f''_{x_ix_j}(\overline{x})\delta x_j^{m+1}+\sum_{j=i}^{n}f''_{x_ix_j}(\overline{x})\delta x_j^m+O\left(\left|\delta x^m\right|^2\right)$$

$$\delta x_i^{m+1}-\delta x_i^m=-\frac{f'_{x_i}(\overline{x_1}+\delta x_1^{m+1},\ldots,\overline{x_{i-1}}+\delta x_{i-1}^{m+1},\overline{x_i}+\delta x_i^m,\ldots,\overline{x_n}+\delta x_n^m)}{f''_{x_ix_i}(\overline{x_1}+\delta x_1^{m+1},\ldots,\overline{x_{i-1}}+\delta x_{i-1}^{m+1},\overline{x_i}+\delta x_i^m,\ldots,\overline{x_n}+\delta x_n^m)}=$$

$$-\frac{\sum\limits_{j=1}^{i-1}f''_{x_ix_j}(\overline{x})\delta x_j^{m+1}+\sum\limits_{j=i}^{n}f''_{x_ix_j}(\overline{x})\delta x_j^m+O\left(\left|\delta x^m\right|^2\right)}{f''_{x_ix_i}(\overline{x_1}+\delta x_1^{m+1},\ldots,\overline{x_{i-1}}+\delta x_{i-1}^{m+1},\overline{x_i}+\delta x_i^m,\ldots,\overline{x_n}+\delta x_n^m)}=-\frac{\delta x_i^m}{\dfrac{f''_{x_ix_i}(\overline{x_1}+\delta x_1^{m+1},\ldots,\overline{x_n}+\delta x_n^m)}{f''_{x_ix_i}(\overline{x_1},\overline{x_2},\ldots,\overline{x_n})}}-$$

$$-\frac{\sum\limits_{j=1}^{i-1}f''_{x_ix_j}(\overline{x})\delta x_j^{m+1}+\sum\limits_{j=i+1}^{n}f''_{x_ix_j}(\overline{x})\delta x_j^m}{f''_{x_ix_i}(\overline{x_1}+\delta x_1^{m+1},\ldots,\overline{x_n}+\delta x_n^m)}+O\left(\left|\delta x^m\right|^2\right)=-\frac{\delta x_i^m}{1+\dfrac{\sum\limits_{j=1}^{i-1}f^{(3)}_{x_ix_ix_j}(\overline{x})\delta x_j^{m+1}+\sum\limits_{j=i}^{n}f^{(3)}_{x_ix_ix_j}(\overline{x})\delta x_j^m}{f''_{x_ix_i}(\overline{x_1},\overline{x_2},\ldots,\overline{x_n})}}+O\left(\left|\delta x^m\right|^2\right)-$$

$$-\frac{\sum\limits_{j=1}^{i-1}f''_{x_ix_j}(\overline{x})\delta x_j^{m+1}+\sum\limits_{j=i+1}^{n}f''_{x_ix_j}(\overline{x})\delta x_j^m}{f''_{x_ix_i}(\overline{x_1}+\delta x_1^{m+1},\ldots,\overline{x_n}+\delta x_n^m)}=-\delta x_i^m-\frac{\sum\limits_{j=1}^{i-1}f''_{x_ix_j}(\overline{x})\delta x_j^{m+1}+\sum\limits_{j=i+1}^{n}f''_{x_ix_j}(\overline{x})\delta x_j^m}{f''_{x_ix_i}(\overline{x_1}+\delta x_1^{m+1},\ldots,\overline{x_n}+\delta x_n^m)}+O\left(\left|\delta x^m\right|^2\right)$$

Сокращая промежуточные записи, получим:

$$\delta x_i^{m+1}-\delta x_i^m=-\delta x_i^m-\frac{\sum\limits_{j=1}^{i-1}f''_{x_ix_j}(\overline{x})\delta x_j^{m+1}+\sum\limits_{j=i+1}^{n}f''_{x_ix_j}(\overline{x})\delta x_j^m}{f''_{x_ix_i}(\overline{x_1}+\delta x_1^{m+1},\ldots,\overline{x_n}+\delta x_n^m)}+O\left(\left|\delta x^m\right|^2\right),\text{или:}$$

$$\delta x_i^{m+1}=-\frac{\sum\limits_{j=1}^{i-1}f''_{x_ix_j}(\overline{x})\delta x_j^{m+1}+\sum\limits_{j=i+1}^{n}f''_{x_ix_j}(\overline{x})\delta x_j^m}{f''_{x_ix_i}(\overline{x_1}+\delta x_1^{m+1},\ldots,\overline{x_n}+\delta x_n^m)}+O\left(\left|\delta x^m\right|^2\right) \qquad (9)$$

Учитывая индуктивное предположение $\left|\delta x_k^{m+1}\right|\le\left|\delta x^m\right|, k=\overline{1,i-1}$, получим:

$$\left|\delta x_i^{m+1}\right|\le\frac{\sum\limits_{j=1}^{i-1}\left|f''_{x_ix_j}(\overline{x})\right|\left|\delta x_j^{m+1}\right|+\sum\limits_{j=i+1}^{n}\left|f''_{x_ix_j}(\overline{x})\right|\left|\delta x_j^m\right|}{\left|f''_{x_ix_i}(\overline{x_1}+\delta x_1^{m+1},\ldots,\overline{x_n}+\delta x_n^m)\right|}+O\left(\left|\delta x^m\right|^2\right)\le\frac{\sum\limits_{j=1,j\ne i}^{n}\left|f''_{x_ix_j}(\overline{x})\right|\left|\delta x_j^m\right|}{\left|f''_{x_ix_i}(\overline{x_1}+\delta x_1^{m+1},\ldots,\overline{x_n}+\delta x_n^m)\right|}+O\left(\left|\delta x^m\right|^2\right)$$

$$\left|\delta x_i^{m+1}\right|\le\left|\delta x^m\right|q_i=\left|\delta x^m\right|q_i\le\left|\delta x^m\right|q<\left|\delta x^m\right| \qquad (10)$$

Индуктивно доказано, что $\left|\delta x_k^{m+1}\right|\le\left|\delta x^m\right|q<\left|\delta x^m\right| \text{ при } k=i$, поэтому в силу доказанного по индукции верно

$$\left|\delta x_k^{m+1}\right|\le\left|\delta x^m\right|q<\left|\delta x^m\right|,\ \forall k=\overline{1,n}, \text{ следовательно, } \left|\delta x^{m+1}\right|=\max_{k=\overline{1,n}}\left|\delta x_k^{m+1}\right|\le\left|\delta x^m\right|q<\left|\delta x^m\right|.$$

Таким образом, сходимость систем итераций (4) и (6) при выполнении условий теоремы доказана. Оценим погрешность метода.

Пусть $\left|\delta x^{m+1}\right|\le\left|\delta x^m\right|q\le\left|\delta x^{m-1}\right|q^2\le\ldots\le\left|\delta x^1\right|q^m\le\left|\delta x^0\right|q^{m+1}\underset{m\to\infty}{\longrightarrow}0$, (11)

где $\left|\delta x^0\right|=\max\limits_{i=\overline{1,n}}\left|\delta x_i^0\right|=\max\limits_{i=\overline{1,n}}\left|x_i^0-\overline{x_i}\right|$ начальное приближение стационарной точки, начальная точка $x^0=\left(x_1^0,x_2^0,\ldots,x_n^0\right)\in A$. Первый порядок скорости сходимости следует из формулы (11) и определения 2 $\left(C=q, p=1\right)$.

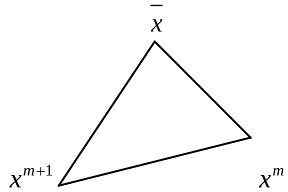

Рис.1.

На практике известны не величины $\left|\delta x^{m+1}\right|$ и $\left|\delta x^m\right|$, а расстояния между последовательными итерациями x^m, x^{m+1} (рис.1). Из неравенства треугольника получим

$$\left|\delta x^m\right| = \left|x^{m+1} - x^m\right| = \left|x^{m+1} - \bar{x} - \left(x^m - \bar{x}\right)\right| = \left|\delta x^{m+1} - \delta x^m\right| \le \left|\delta x^m\right| + \left|\delta x^{m+1}\right| \le (1+q)\left|\delta x^m\right|,$$

учитывая доказанное неравенство $\left|\delta x^{m+1}\right| \le q\left|\delta x^m\right|$ и используя неравенство треугольника, получим после m итераций:

$$\left|x^m - \bar{x}\right| \le \left|x^m - x^{m+1}\right| + \left|x^{m+1} - x^{m+2}\right| + \left|x^{m+2} - x^{m+3}\right| + \ldots + \left|x^{m+n} - x^{m+n+1}\right| + \left|x^{m+n+1} - \bar{x}\right| + \ldots =$$

$$= \sum_{i=m}^{\infty}\left|x^{i+1} - x^i\right| \le (1+q)\left|\delta x^m\right| + (1+q)\left|\delta x^{m+1}\right| + (1+q)\left|\delta x^{m+2}\right| + \ldots = (1+q)(\left|\delta x^m\right| + \left|\delta x^{m+1}\right| + \left|\delta x^{m+2}\right| + \ldots) \le$$

$$\le (1+q)\left|\delta x^0\right| q^m \sum_{i=0}^{\infty} q^i = \frac{(1+q)\left|\delta x^0\right| q^m}{1-q}.$$

Поскольку $\left|\delta x^1\right| \le q\left|\delta x^0\right|$, запишем неравенство треугольника с избытком и с недостатком, получим:

$$(1-q)\left|\delta x^0\right| = \left|\delta x^0\right| - \left|\delta x^1\right| \le l_0 = \left|x^1 - x^0\right| \le \left|\delta x^0\right| + \left|\delta x^1\right| = (1+q)\left|\delta x^0\right|.$$

Окончательно оценка погрешности(невязки) имеет вид:

$$\left|\delta x^m\right| = \left|x^m - \bar{x}\right| \le \frac{(1+q)\left|\delta x^0\right| q^m}{1-q} \le \frac{(1+q)q^m}{(1-q)^2} l_0. \qquad (12)$$

Замечание 1 (необходимость). Условие 2) - диагонального преобладания матрицы Гесса является также и необходимым условием сходимости. Достаточно привести 1 пример с условием $q > 1$, в котором итерация (6) расходится

$f(x_1, x_2) = x_1^2 + x_2^2 + 3x_1x_2$, $f_{x_1}(x_1, x_2) = 2x_1 + 3x_2$, $f_{x_2}(x_1, x_2) = 2x_2 + 3x_1$. Стационарная точка

$\left(\bar{x}_1, \bar{x}_2\right) = (0{,}0)$. $f_{x_1x_1}\left(\bar{x}_1, \bar{x}_2\right) = f_{x_2x_2}\left(\bar{x}_1, \bar{x}_2\right) = 2$, $f_{x_1x_2}\left(\bar{x}_1, \bar{x}_2\right) = f_{x_2x_1}\left(\bar{x}_1, \bar{x}_2\right) = 3$, $q = \dfrac{3}{2} > 1$.

Согласно(9):

$$\begin{cases} \delta x_1^{m+1} = -\dfrac{f_{x_1x_2}\left(\bar{x}_1, \bar{x}_2\right)\delta x_2^m}{f_{x_1x_1}\left(\bar{x}_1 + \delta x_1^m, \bar{x}_2 + \delta x_2^m\right)} + O\left(\left|\delta x^m\right|^2\right) = -\dfrac{3}{2}\delta x_2^m \\ \delta x_2^{m+1} = -\dfrac{f_{x_2x_1}\left(\bar{x}_1, \bar{x}_2\right)\delta x_1^m}{f_{x_2x_2}\left(\bar{x}_1 + \delta x_1^{m+1}, \bar{x}_2 + \delta x_2^m\right)} + O\left(\left|\delta x^m\right|^2\right) = -\dfrac{3}{2}\delta x_1^m \end{cases} \Leftrightarrow \begin{cases} \delta x_1^{m+1} = \dfrac{9}{4}\delta x_1^{m-1} + O\left(\left|\delta x^m\right|^2\right) \\ \delta x_2^{m+1} = \dfrac{9}{4}\delta x_2^{m-1} + O\left(\left|\delta x^m\right|^2\right) \end{cases}$$

$$\begin{cases} \left|\delta x_1^{m+1}\right| = \dfrac{9}{4}\left|\delta x_1^{m-1}\right| + O\left(\left|\delta x^m\right|^2\right) \\ \left|\delta x_2^{m+1}\right| = \dfrac{9}{4}\left|\delta x_2^{m-1}\right| + O\left(\left|\delta x^m\right|^2\right) \end{cases}, m = 2,4,6,\ldots \begin{cases} \left|\delta x_1^{m}\right| = \dfrac{9}{4}\left|\delta x_1^{m-2}\right| + O\left(\left|\delta x^{m-1}\right|^2\right) \\ \left|\delta x_2^{m}\right| = \dfrac{9}{4}\left|\delta x_2^{m-2}\right| + O\left(\left|\delta x^{m-1}\right|^2\right) \end{cases}, m = 2,4,6,\ldots$$

Откуда видно, что во всех чётных и нечётных итерациях невязка растёт, а итерация удаляется от стационарной точки.

Замечание 2. Формула (9) выполняется локально, т.е. условия Т.1 должны выполняться обязательно в окрестности стационарной точки. Так как условия Т.1 выполняются абсолютно во всей области $A \subset R^n$, то они выполнены и локально в точке $\left(\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_n\right) \in A$.

Замечание 3. Если итерация для невязки задаётся формулой $\delta x^{m+1} = A\delta x^m$ - аналогом формулы (9), где A линейный оператор, и A- сжимающее отображение, т.е. $\left|\delta x^{m+1}\right| \le q\left|\delta x^m\right| < \left|\delta x^m\right|$ (что обеспечивается условиями 2) Т.1) , то по теореме о неподвижной точке в метрических пространствах[4] сжимающее отображение имеет единственное решение. Таким образом, единственность и существование решения итерации(6) доказана.

Замечание 4. Сходимость (6)выполняется при разных значениях $q_i < 1, i = \overline{1,n}$. Если по всем переменным $i = \overline{1,n}$ в итерации (6) недиагональные элементы матрицы Гесса $\dfrac{\partial^2 f(x)}{\partial x_i \partial x_j} \equiv 0, j, i = \overline{1,n}, j \ne i$. В этом случае скорость сходимости (6) не линейная, а квадратичная.

Замечание 5. Отметим, что с помощью первой и второй теорем можно находить не только минимумы и максимумы, но и другие стационарные точки, в том числе и седловые точки в теории игр и матричных играх. Например, $f(x_1, x_2) = x_1^2 - x_2^2, (\bar{x}_1, \bar{x}_2) = (0{,}0)$.

В качестве примера задачи на экстремум, решённой численно, рассмотрим[1]:

$f(x_1, x_2, x_3) = x_1^2 + x_2^2 + x_3^2 - x_1x_2 + x_1 - 2x_3 \to extr$.

Запишем градиент и матрицу Гесса для функции $f(x_1, x_2, x_3)$:
$\nabla f(x_1, x_2, x_3) = (2x_1 - x_2 + 1,\ 2x_2 - x_1, 2x_3 - 2)$.

$H_{i,j}(x_1, x_2, x_3) = \dfrac{\partial^2 f(\bar{x})}{\partial x_i \partial x_j} = \begin{vmatrix} 2 & -1 & 0 \\ -1 & 2 & 0 \\ 0 & 0 & 2 \end{vmatrix}$. Условия теоремы 1 выполнены:

$|H_{1,1}| = 2 > |H_{1,2}| + |H_{1,3}| = 1 + 0 = 1$, $M_{11} = 1, M_{21} = 2, q_1 = \dfrac{1}{2} < 1$

$|H_{2,2}| = 2 > |H_{2,1}| + |H_{2,3}| = 1 + 0 = 1$, $M_{12} = 1, M_{22} = 2, q_2 = \dfrac{1}{2} < 1$,

$|H_{3,3}| = 2 > |H_{3,1}| + |H_{3,2}| = 0 + 0 = 0$. $M_{13} = 0, M_{23} = 2, q_2 = 0 < 1, q = \max\left\{\dfrac{1}{2}, \dfrac{1}{2}, 0\right\} = \dfrac{1}{2}$.

Запишем итерацию по формуле (6):

$$\begin{cases} x_1^{m+1} = x_1^m - \dfrac{(2x_1^m - x_2^m + 1)}{2}, \\ x_2^{m+1} = x_2^m - \dfrac{(2x_2^m - x_1^{m+1})}{2} \\ x_3^{m+1} = x_3^m - \dfrac{2x_3^m - 2}{2} = 1 \end{cases}$$

Из (12) необходимое число итераций $N \geq \ln\left(\dfrac{|\delta x^m|}{|\delta x^0|}\left(\dfrac{1-q}{1+q}\right)\right) \Big/ \ln q = \ln\left(\dfrac{|\delta x^0|}{|\delta x^m|}\left(\dfrac{1+q}{1-q}\right)\right) \Big/ \ln 1/q$ (13)

$$N \geq \ln\left(\frac{|\delta x^m|}{l_0}\left(\frac{(1-q)^2}{1+q}\right)\right) \Big/ \ln q = \ln\left(\frac{l_0}{|\delta x^m|}\left(\frac{1+q}{(1-q)^2}\right)\right) \Big/ \ln 1/q \qquad (14)$$

Выберем $|\delta x^m| = 10^{-15}, |\delta x^0| = 10^2$, получим необходимое число $N = \ln(10^{17} * 3) \Big/ \ln 2 = 58$ итераций.

Составим на языке с программу:

```
#include<stdio.h>
#include<math.h>
double fx1(double x1,double x2, double x3);
double fx2(double x1,double x2, double x3);
double fx3(double x1, double x2, double x3);
double fxx1(double x1, double x2, double x3);
double fxx2(double x1,double x2, double x3);
double fxx3(double x1, double x2, double x3);
int main()
{
int n,i;
double x1,x2,x3;
n=60;
x1=-100.0;
x2=100.0;
x3=100.0;
for(i=1;i<=n ;i++)
{
x1=x1-fx1(x1,x2,x3)/fxx1(x1,x2,x3);
x2=x2-fx2(x1,x2,x3)/fxx2(x1,x2,x3);
x3=x3-fx3(x1,x2,x3)/fxx3(x1,x2,x3);
}
printf("x1=%.16lf,x2=%.16lf,x3=%.16lf,extr=%.16lf\n",x1,x2,x3,x1*x1+x2*x2+x3*x3-x1*x2+x1-2.0*x3);
}
double fx1(double x1,double x2,double x3)
{
return 2.0*x1-x2+1.0;
```

```
}
 double fx2(double x1,double x2, double x3)
{
return 2.0*x2-x1;
}
double fx3(double x1,double x2,double x3)
{
return 2.0*x3-2.0;
}
double fxx1(double x1, double x2, double x3)
{
return 2.0;
}
double fxx2(double x1,double x2,double x3)
{
return 2.0;
}
double fxx3(double x1,double x2,double x3)
{
return 2.0;
}
```

Программа возвращает решение задачи и значение функционала:

$x_1 = -0.6666666666666666, x_2 = 0.3333333333333333, x_3 = 1.0000000000000000$

$extr = -1.3333333333333330$. Матрица Гессе положительно определена и , следовательно, точке $\bar{x} = \left(-\frac{2}{3}, \frac{1}{3}, 1 \right)$ локальный минимум функции**.**

Точное решение есть: $\overline{x_1} = -\frac{2}{3}, \overline{x_2} = \frac{1}{3}, \overline{x_3} = 1,\ f\left(\overline{x_1}, \overline{x_2}, \overline{x_3}\right) = -\frac{4}{3}$.

Рассмотрим разностную формулу, полученную из (6), в которой первая производная заменена центральной разностью с шагом $h/2$

Теорема 2. Пусть выполнены все условие 2) первой теоремы 1 и функционал $f(x) \in C^6(A)$ имеет гладкость не хуже шестого порядка. Тогда итерационные формулы (6) и (13) сравнимы по точности со вторым порядком $O\left(h^2\right) = O\left(\left|\delta x^m\right|^2\right)$ и справедливы результаты теоремы 1. Оценка верхней границы для оптимального шага h_i^m определяется формулами:

$$h_i^{apper} = \min\left\{ \sqrt{\frac{12\left| f_{x_i}^{"}\left(\bar{x}\right)\right|}{\left| f_{x_i}^{(4)}\left(\bar{x}\right)\right|}}, \sqrt{\left(\frac{240 f_{x_i}^{(3)}\left(\bar{x}\right) f_{x_i}^{"}\left(\bar{x}\right)}{20 f_{x_i}^{(3)}\left(\bar{x}\right) f_{x_i}^{(4)}\left(\bar{x}\right) - 3 f_{x_i}^{(5)}\left(\bar{x}\right) f_{x_i}^{"}\left(\bar{x}\right)} \right)} \right\}$$

Доказательство:

Разложим в рял Тейлора числитель и знаменатель формул (6) , обозначим $x^m = \left(x_1^m, \ldots, x_n^m\right)$:

$$f\left(x_1^m, .., x_i^m \pm \frac{h}{2}, \ldots, x_i^m\right) = f\left(x^m\right) \pm \frac{h}{2} f_{x_i}^{'}\left(x^m\right) + \frac{h^2}{8} f_{x_i}^{"}\left(x^m\right) \pm \frac{h^3}{48} f_{x_i}^{(3)}\left(x^m\right) + \frac{h^4}{384} f_{x_i}^{(4)}\left(x^m\right) \pm \frac{h^5}{3840} f_{x_i}^{(5)}\left(x^m\right)$$

$$\frac{f\left(x_1^m, \ldots, x_i^m + \frac{h}{2}, \ldots, x_n^m\right) - f\left(x_1^m, \ldots, x_i^m - \frac{h}{2}, \ldots, x_n^m\right)}{h} = f_{x_i}^{'}\left(x^m\right) + \frac{h^2}{24} f_{x_i}^{(3)}\left(x^m\right) + \frac{h^4}{1920} f_{x_i}^{(5)}\left(x^m\right) + O\left(h^6\right)$$

$$f\left(x_1^m,\ldots,x_i^m \pm h,\ldots,x_n^m\right)= f\left(x^m\right)\pm hf_{x_i}^{'}\left(x^m\right)+\frac{h^2}{2}f_{x_i}^{''}\left(x^m\right)\pm\frac{h^3}{6}f_{x_i}^{(3)}\left(x^m\right)+\frac{h^4}{24}f_{x_i}^{(4)}\left(x^m\right)\pm\frac{h^5}{120}f_{x_i}^{(5)}\left(x^m\right)+\frac{h^6}{720}f_{x_i}^{(6)}\left(x^m\right)$$

$$\frac{f\left(x_1^m,\ldots,x_i^m+h,\ldots,x_n^m\right)+f\left(x_1^m,\ldots,x_i^m-h,\ldots,x_n^m\right)-2f\left(x_1^m,x_2^m,\ldots,x_n^m\right)}{h^2}=f_{x_i}^{''}\left(x^m\right)+\frac{h^2}{12}f_{x_i}^{(4)}\left(x^m\right)+\frac{h^4}{360}f_{x_i}^{(6)}\left(x^m\right)+O\left(h^6\right)$$

$$\begin{cases} x_1^{m+1}=x_1^m-\dfrac{\dfrac{f\left(x_1^m+\frac{h}{2},x_2^m,\ldots,x_n^m\right)-f\left(x_1^m-\frac{h}{2},x_2^m,\ldots,x_n^m\right)}{h}+O\left(h^2\right)}{\dfrac{f\left(x_1^m+h,x_2^m,\ldots,x_n^m\right)+f\left(x_1^m-h,x_2^m,\ldots,x_n^m\right)-2f\left(x_1^m,x_2^m,\ldots,x_n^m\right)}{h^2}+O\left(h^2\right)}, \\ \ldots. \\ x_i^{m+1}=x_i^m-\dfrac{\dfrac{f\left(x_1^m,\ldots,x_i^m+\frac{h}{2},\ldots,x_n^m\right)-f\left(x_1^m,\ldots,x_i^m-\frac{h}{2},\ldots,x_n^m\right)}{h}+O\left(h^2\right)}{\dfrac{f\left(x_1^m,\ldots,x_i^m+h,\ldots,x_n^m\right)+f\left(x_1^m,\ldots,x_i^m-h,\ldots,x_n^m\right)-2f\left(x_1^m,x_2^m,\ldots,x_n^m\right)}{h^2}+O\left(h^2\right)} \\ \ldots. \\ x_n^{m+1}=x_n^m-\dfrac{\dfrac{f\left(x_1^m,x_2^m,\ldots,x_n^m+\frac{h}{2}\right)-f\left(x_1^m,x_2^m,\ldots,x_n^m-\frac{h}{2}\right)}{h}+O\left(h^2\right)}{\dfrac{f\left(x_1^m,x_2^m,\ldots,x_n^m+h\right)+f\left(x_1^m,x_2^m,\ldots,x_n^m-h\right)-2f\left(x_1^m,x_2^m,\ldots,x_n^m\right)}{h^2}+O\left(h^2\right)} \end{cases} \quad (15)$$

Откуда видно, что (6) отличается от (15) с точностью $O\left(h^2\right)$. Заменим h на $\left|x_i^{m+1}-x_i^m\right|=h_i^m$, где h_i^m - абсолютная разность i - ой координаты между соседними итерациями с номерами $m,m+1$.

Но $\left|\delta x_i^m\right|(1-q)=\left|\delta x_i^m\right|-\left|\delta x_i^{m+1}\right|\le\left|x_i^{m+1}-x_i^m\right|\le\left|\delta x_i^m\right|+\left|\delta x_i^{m+1}\right|=\left|\delta x_i^m\right|(1+q)$

$\dfrac{h_i^m}{(1+q)}=\dfrac{\left|x_i^{m+1}-x_i^m\right|}{(1+q)}\le\left|\delta x_i^m\right|\le\dfrac{\left|x_i^{m+1}-x_i^m\right|}{(1-q)}=\dfrac{h_i^m}{(1-q)}$, $(1-q)\left|\delta x_i^m\right|\le h_i^m\le(1+q)\left|\delta x_i^m\right|$.

По определению символики "О большое" (Б.П. Демидович Сборник задач и упражнений по математическому анализу):

Запись $\varphi(x)=O(\psi(x))$ при $x\in X$ означает, что существует постоянная $A>0$ такая, что

$|\varphi(x)|\le A|\psi(x)|$ для $x\in X$

Для дискретного варианта задания функции аналогично мы можем записать

$\varphi_m=O(\psi_m)$ при $m=0,1,2,\ldots$, если существует постоянная А , что

$|\varphi_m|\le A|\psi_m|$ для $m=0,1,2,\ldots$

Поскольку для функций h_i и δx_i раньше получено неравенство $0\le h_i^m\le(1+q)\left|\delta x_i^m\right|$, т.е. существует постоянная $A=1+q$, такая что $h_i^m\le A\left|\delta x_i^m\right|,\forall m=0,1,2,\ldots$ Следовательно, $h_i=O(\delta x_i), i=\overline{1,n}$.

С другой стороны $\left|\delta x_i^m\right|\le\dfrac{h_i^m}{1-q},\forall m=0,1,2,\ldots$ Тогда $\delta x_i=O(h_i), i=\overline{1,n}$ с постоянной $A=\dfrac{1}{1-q}$.

Имеем $O\left(h_i^2\right)=O(O(\delta x_i)*O(\delta x_i))=O\left(O\left(\delta x_i^2\right)\right)=O\left(\delta x_i^2\right)$.

Мы использовали 2 известных свойства символики О большое:

$O\left(x^n\right)O\left(x^m\right)=O\left(x^{n+m}\right); O(O(f(x)))=O(f(x))$

Аналогично, получаем, что $O\left(\delta x_i^2\right)=O(O(h_i)*O(h_i))=O\left(O\left(h_i^2\right)\right)=O\left(h_i^2\right)$.

Поэтому для любого $m=0,1,2,\ldots O\left(\left(h_i^m\right)^2\right)=O\left(\left|\delta x_i^m\right|^2\right)$ и системы уравнений (6) и (15) справедливы одновременно со вторым порядком точности. Если в теореме 2 выполнены условия теоремы 1, то существует единственное решение итерации (15), сходящееся к решению системы(2), число итераций определяется, по – прежнему, формулами(13),(14). Теорема доказана.

Кроме того, систему уравнений (15) удобнее преобразовать к форме(16). На первый взгляд, формулы (16) от шага h_i^m не зависят. Действительно, числитель и знаменатель дроби прямо пропорциональны в первом приближении $\left(h_i^m\right)^2$ и сокращаются на $\left(h_i^m\right)^2$. Поэтому для поиска оптимального шага с большей точностью чем $O\left(h_i^m\right)^2$ нужно учесть члены более высокого порядка.

$$\begin{cases} x_1^{m+1} = x_1^m - h_1^m \dfrac{f\left(x_1^m + \dfrac{h_1^m}{2}, x_2^m, \ldots, x_n^m\right) - f\left(x_1^m - \dfrac{h_1^m}{2}, x_2^m, \ldots, x_n^m\right)}{f\left(x_1^m + h_1^m, x_2^m, \ldots, x_n^m\right) + f\left(x_1^m - h_1^m, x_2^m, \ldots, x_n^m\right) - 2f\left(x_1^m, x_2^m, \ldots, x_n^m\right)} + O\left(\left(h_1^m\right)^2\right) \\ \qquad \ldots \\ x_i^{m+1} = x_i^m - h_i^m \dfrac{f\left(x_1^m, \ldots, x_i^m + \dfrac{h_i^m}{2}, \ldots, x_n^m\right) - f\left(x_1^m, \ldots, x_i^m - \dfrac{h_i^m}{2}, \ldots, x_n^m\right)}{f\left(x_1^m, \ldots, x_i^m + h_i^m, \ldots, x_n^m\right) + f\left(x_1^m, \ldots, x_i^m - h_i^m, \ldots, x_n^m\right) - 2f\left(x_1^m, x_2^m, \ldots, x_n^m\right)} + O\left(\left(h_i^m\right)^2\right) \\ \qquad \ldots \\ x_n^{m+1} = x_n^m - h_n^m \dfrac{f\left(x_1^m, x_2^m, \ldots, x_n^m + \dfrac{h_n^m}{2}\right) - f\left(x_1^m, x_2^m, \ldots, x_n^m - \dfrac{h_n^m}{2}\right)}{f\left(x_1^m, x_2^m, \ldots, x_n^m + h_n^m\right) + f\left(x_1^m, x_2^m, \ldots, x_n^m - h_n^m\right) - 2f\left(x_1^m, x_2^m, \ldots, x_n^m\right)} + O\left(\left(h_n^m\right)^2\right) \end{cases} \tag{16}$$

Получим, преобразуя i - е уравнение (16):

$$x_i^{m+1} = x_i^m - h_i^m \frac{f\left(x_1^m, \ldots, x_i^m + \dfrac{h_i^m}{2}, \ldots, x_n^m\right) - f\left(x_1^m, \ldots, x_i^m - \dfrac{h_i^m}{2}, \ldots, x_n^m\right)}{f\left(x_1^m, \ldots, x_i^m + h_i^m, \ldots, x_n^m\right) + f\left(x_1^m, \ldots, x_i^m - h_i^m, \ldots, x_n^m\right) - 2f\left(x_1^m, x_2^m, \ldots, x_n^m\right)} =$$

$$x_i^m - \frac{f'_{x_i}\left(x^m\right) + \dfrac{\left(h_i^m\right)^2}{24} f^{(3)}_{x_i}\left(x^m\right) + \dfrac{\left(h_i^m\right)^4}{1920} f^{(5)}_{x_i}\left(x^m\right) + O\left(\left(h_i^m\right)^6\right)}{f''_{x_i}\left(x^m\right) + \dfrac{\left(h_i^m\right)^2}{12} f^{(4)}_{x_i}\left(x^m\right) + \dfrac{\left(h_i^m\right)^4}{360} f^{(6)}_{x_i}\left(x^m\right) + O\left(\left(h_i^m\right)^6\right)} = x_i^m - \frac{f'_{x_i}\left(x^m\right)}{f''_{x_i}\left(x^m\right)}\left(1 + \frac{\left(h_i^m\right)^2}{24}\frac{f^{(3)}_{x_i}\left(x^m\right)}{f'_{x_i}\left(x^m\right)} + \frac{\left(h_i^m\right)^4}{1920}\frac{f^{(5)}_{x_i}\left(x^m\right)}{f'_{x_i}\left(x^m\right)} + O\left(\left(h_i^m\right)^6\right)\right) *$$

$$\left(1 - \frac{\left(h_i^m\right)^2}{12}\frac{f^{(4)}_{x_i}\left(x^m\right)}{f''_{x_i}\left(x^m\right)} - \frac{\left(h_i^m\right)^4}{360}\frac{f^{(6)}_{x_i}\left(x^m\right)}{f''_{x_i}\left(x^m\right)} + \left(\frac{\left(h_i^m\right)^2}{12}\frac{f^{(4)}_{x_i}\left(x^m\right)}{f''_{x_i}\left(x^m\right)}\right)^2 + O\left(\left(h_i^m\right)^6\right)\right) = x_i^m - \frac{f'_{x_i}\left(x^m\right)}{f''_{x_i}\left(x^m\right)}\left(1 + \frac{\left(h_i^m\right)^2}{24}\left(\frac{f^{(3)}_{x_i}\left(x^m\right)}{f'_{x_i}\left(x^m\right)} - \frac{2f^{(4)}_{x_i}\left(x^m\right)}{f''_{x_i}\left(x^m\right)}\right) + \right.$$

$$\left. + \left(h_i^m\right)^4\left(\frac{1}{144}\left(\frac{f^{(4)}_{x_i}\left(x^m\right)}{f''_{x_i}\left(x^m\right)}\right)^2 + \frac{f^{(5)}_{x_i}\left(x^m\right)}{1920 f'_{x_i}\left(x^m\right)} - \frac{f^{(6)}_{x_i}\left(x^m\right)}{360 f''_{x_i}\left(x^m\right)} - \frac{f^{(3)}_{x_i}\left(x^m\right)}{288 f'_{x_i}\left(x^m\right)}\frac{f^{(4)}_{x_i}\left(x^m\right)}{f''_{x_i}\left(x^m\right)}\right) + O\left(\left(h_i^m\right)^6\right)\right), i = \overline{1, n} \tag{17}$$

В последнем разложении используется формула $\dfrac{1}{1+\alpha} = \sum_{k=0}^{\infty}\left(-\alpha\right)^k$ и для её применимости достаточно $\left|\alpha\right| < 1, где \left|\alpha\right| = \dfrac{1}{\left|f''_{x_i}\left(x^m\right)\right|}\left|\dfrac{\left(h_i^m\right)^2}{12} f^{(4)}_{x_i}\left(x^m\right) + \dfrac{\left(h_i^m\right)^4}{360} f^{(6)}_{x_i}\left(x^m\right)\right|$ (18)

$$\frac{1}{\left|f''_{x_i}\left(x^m\right)\right|}\left|\frac{\left(h_i^m\right)^2}{12} f^{(4)}_{x_i}\left(x^m\right) + \frac{\left(h_i^m\right)^4}{360} f^{(6)}_{x_i}\left(x^m\right)\right| \le \frac{1}{\left|f''_{x_i}\left(x^m\right)\right|}\frac{\left(h_i^m\right)^2}{12}\left|f^{(4)}_{x_i}\left(x^m\right)\right| + \frac{1}{\left|f''_{x_i}\left(x^m\right)\right|}\frac{\left(h_i^m\right)^4}{360}\left|f^{(6)}_{x_i}\left(x^m\right)\right| < 1$$

Обозначим $A = \dfrac{\left|f^{(4)}_{x_i}\left(x^m\right)\right|}{12\left|f''_{x_i}\left(x^m\right)\right|} \ge 0, B = \dfrac{\left|f^{(6)}_{x_i}\left(x^m\right)\right|}{360\left|f''_{x_i}\left(x^m\right)\right|} \ge 0, y = \left(h_i^m\right)^2$, последнее неравенство перейдёт в

$Ay + By^2 - 1 < 0, y \ge 0$, решая два неравенства, оценим приближённо верхнюю границу шага итерации $0 \le y < \dfrac{-A + \sqrt{A^2 + 4B}}{2B} = \dfrac{4B}{2B\left(\sqrt{A^2 + 4B} + A\right)} = \dfrac{2}{\sqrt{A^2 + 4B} + A} \le \dfrac{2}{2A} = \dfrac{1}{A} = \dfrac{12\left|f''_{x_i}\left(x^m\right)\right|}{\left|f^{(4)}_{x_i}\left(x^m\right)\right|}$ (19)

Чтобы формула(17) имела точность $O\left(\left(h_i^m\right)^6\right)$ необходимо:

$$\left(h_i^m\right)^2=\frac{\dfrac{f_{x_i}^{(3)}\left(x^m\right)}{f_{x_i}^{'}\left(x^m\right)}-\dfrac{2f_{x_i}^{(4)}\left(x^m\right)}{f_{x_i}^{''}\left(x^m\right)}}{-\dfrac{1}{6}\left(\dfrac{f_{x_i}^{(4)}\left(x^m\right)}{f_{x_i}^{''}\left(x^m\right)}\right)^2-\dfrac{f_{x_i}^{(5)}\left(x^m\right)}{80f_{x_i}^{'}\left(x^m\right)}+\dfrac{f_{x_i}^{(6)}\left(x^m\right)}{15f_{x_i}^{''}\left(x^m\right)}+\dfrac{f_{x_i}^{(3)}\left(x^m\right)}{12f_{x_i}^{'}\left(x^m\right)}\dfrac{f_{x_i}^{(4)}\left(x^m\right)}{f_{x_i}^{''}\left(x^m\right)}}=$$

$$\frac{240\left(f_{x_i}^{(3)}\left(x^m\right)f_{x_i}^{''}\left(x^m\right)-2f_{x_i}^{(4)}\left(x^m\right)f_{x_i}^{'}\left(x^m\right)\right)}{16f_{x_i}^{(6)}\left(x^m\right)f_{x_i}^{'}\left(x^m\right)+20f_{x_i}^{(3)}\left(x^m\right)f_{x_i}^{(4)}\left(x^m\right)-3f_{x_i}^{(5)}\left(x^m\right)f_{x_i}^{''}\left(x^m\right)-40\dfrac{f_{x_i}^{'}\left(x^m\right)\left(f_{x_i}^{(4)}\left(x^m\right)\right)^2}{f_{x_i}^{''}\left(x^m\right)}} \qquad (20)$$

$i=\overline{1,n}$.

Учитывая систему уравнений (2) $f_{x_i}^{'}\left(x^m\right)\approx f_{x_i}^{'}\left(\bar{x}\right)=0$, упростим (20):

$$\left(h_i^m\right)^2=\left(\frac{240f_{x_i}^{(3)}\left(x^m\right)f_{x_i}^{''}\left(x^m\right)}{20f_{x_i}^{(3)}\left(x^m\right)f_{x_i}^{(4)}\left(x^m\right)-3f_{x_i}^{(5)}\left(x^m\right)f_{x_i}^{''}\left(x^m\right)}\right)\approx\left(\frac{240f_{x_i}^{(3)}\left(\bar{x}\right)f_{x_i}^{''}\left(\bar{x}\right)}{20f_{x_i}^{(3)}\left(\bar{x}\right)f_{x_i}^{(4)}\left(\bar{x}\right)-3f_{x_i}^{(5)}\left(\bar{x}\right)f_{x_i}^{''}\left(\bar{x}\right)}\right), i=\overline{1,n} \qquad (21)$$

Необходимо выполнить условия (19) и (21)одновременно, в результате выбираем наименьшее значение

$$h_i^{apper}=\min\left\{\sqrt{\frac{12\left|f_{x_i}^{''}\left(x^m\right)\right|}{\left|f_{x_i}^{(4)}\left(x^m\right)\right|}},\sqrt{\left(\frac{240f_{x_i}^{(3)}\left(\bar{x}\right)f_{x_i}^{''}\left(\bar{x}\right)}{20f_{x_i}^{(3)}\left(\bar{x}\right)f_{x_i}^{(4)}\left(\bar{x}\right)-3f_{x_i}^{(5)}\left(\bar{x}\right)f_{x_i}^{''}\left(\bar{x}\right)}\right)}\right\} \qquad (22)$$

Для полиномов n переменных не выше четвёртой степени из (22) получим $h_i^{apper}=\sqrt{12\dfrac{\left|f_{x_i}^{''}\left(\bar{x}\right)\right|}{\left|f_{x_i}^{(4)}\left(\bar{x}\right)\right|}}$ (23)

, т.к. (19) и (21) дают равные значения. Рассмотрим пример

$f\left(x_1,x_2\right)=1+x_1^{\ 4}+x_2^{\ 4}-x_1x_2+x_1^{\ 2}+x_2^{\ 2}$,

$f_{x_1}^{'}\left(\bar{x}\right)=4\bar{x}_1^3-\bar{x}_2+2\bar{x}_1=0$, $f_{x_2}^{'}\left(\bar{x}\right)=4\bar{x}_2^3-\bar{x}_1+2\bar{x}_2=0\Rightarrow\left(\bar{x}_1-\bar{x}_2\right)\left(4\left(\bar{x}_1^2+\bar{x}_2^2+\bar{x}_1\bar{x}_2\right)+3\right)=0$

$\Leftrightarrow\bar{x}_1=\bar{x}_2\Leftrightarrow4\bar{x}_1^3+\bar{x}_1=0\Leftrightarrow\bar{x}_1=\bar{x}_2=0$. $\dfrac{\partial^2f\left(\bar{x}_1,\bar{x}_2\right)}{\partial x_i\partial x_j}=\begin{vmatrix}12\bar{x}_1^2+2 & -1\\ -1 & 12\bar{x}_2^2+2\end{vmatrix}=\begin{vmatrix}2 & -1\\ -1 & 2\end{vmatrix}$

По теореме 1 матрица Гесса имеет диагональное преобладание. Оптимальный шаг найдём по формуле (23) $\left(h_1^m\right)^2=\left(h_2^m\right)^2=\dfrac{12*2}{24}=1, h_1^m=h_2^m=1$.

Замечание1 (Существование оптимального шага $0<h_i^m\le\left(h_i^m\right)_{apper}$ подтверждает численный эксперимент, поскольку погрешность вычислений стационарной точки увеличивается как при $h_i^m<\left(h_i^m\right)_{apper}$, так и при $h_i^m>\left(h_i^m\right)_{apper}$.Примеры оценки оптимального шага можно встретить в [5].

Замечание 2. Несмотря на априорную гладкость функционала в (14), оценка оптимального шага требует как минимум гладкости четвёртого порядка (в общем случае частных производных 5 порядка).

Запишем программу на языке с (в данном примере мы используем $h_1^m=h_2^m=1.0$ - получаем решение с максимально возможной точностью – 16 значащих цифр):

```
#include<stdio.h>
#include<math.h>
double fx(double x1,double x2);
int main()
{
int n,i;
double a1,b1,a2,b2,a3,b3,delta1,delta2,delta3;
double h1,h2,h3,h;
n=60;
a1=100.0;
a2=100.0;
h1=1.0;
h2=1.0;
for(i=1;i<=n;i++)
{
b1=a1-h1*(fx(a1+(h1/2.0),a2)-fx(a1-(h1/2.0),a2))/(fx(a1+h1,a2)+fx(a1-h1,a2)-2.0*fx(a1,a2));
```

```
delta1=sqrt((b1-a1)*(b1-a1));
a1=b1;
b2=a2-h2*(fx(a1,a2+(h2/2.0))-fx(a1,a2-(h2/2.0)))/(fx(a1,a2+h2)+fx(a1,a2-h2)-2.0*fx(a1,a2));
delta2=sqrt((b2-a2)*(b2-a2));
a2=b2;
}
printf("x1=%.30lf,x2=%.30lf,extr=%.30lf\n",b1,b2,fx(b1,b2));

printf(" h1=%.16lf,h2=%.16lf,h3=%.16lf\n",h1,h2);
}
double fx(double x1, double x2)
{
return    1.0+x1*x1*x1*x1+x2*x2*x2*x2+x1*x1-x1*x2+x2*x2;
}
```

Программа возвращает для $h_1 = 1.0, h_2 = 1.0$

$x_1 = 0.00000000000000011, x_2 = 0.00000000000000011, extr = 1.00000000000000000000$

Выводы

1) В первой теореме получены условия для эффективной сходимости итерации поиска стационарной точки для функционалов 3-го порядка гладкости, даются формулы необходимого числа итераций для поиска экстремума с заданной точностью в конечномерной задаче.
2) Вторая теорема доказана для априорно гладких функционалов, возможно не представимых в виде композиции элементарных функций, показана эквивалентность точности итерационных формул в теоремах, определена верхняя граница оптимального шага.
3) Приведены программы и примеры, подтверждающие эффективность доказанных методов. Используя полученные теоремы можно находить как экстремумы, так и другие стационарные точки, например, точки перевала.

Литература

1) Галеев Э.М., Тихомиров В.М. Краткий курс теории экстремальных задач. – Изд. – во Моск. Ун – та, 1989. – 204с.: ил.

2)Н.С.Бахвалов, Н.П.Жидков, Г.М.Кобельков. Численные методы. – 7 – е изд.: БИНОМ. Лаборатория знаний,2011. – 636 с. – (Классический университетский учебник).

3)Федоренко Р.П. Введение в вычислительную физику: Учебное пособие для вузов. – Долгопрудный: Издательский дом “Интеллект”, 2008. – 504с.

4) Колмогоров А.Н., Фомин С.В.Элементы теории функции и функционального анализа. – М.:1989 – 450 с.5) Бахвалов Н.С., Лапин А.В., Чижонков Е.В. Численные методы в задачах и упражнениях: учебное пособие. – М.:БИНОМ. Лаборатория знаний,2010. – 240 с.