УДК 519.6   ББК 22.18   22.193

ОБ ЭФФЕКТИВНОМ ПОИСКЕ БЕЗУСЛОВНОГО ЭКСТРЕМУМА ГЛАДКИХ ФУНКЦИОНАЛОВ В КОНЕЧНОМЕРНЫХ ЗАДАЧАХ

Голубева О.В., Ехилевский С.Г., Пастухов Ю.Ф., Пастухов Д.Ф.

Аннотация: Получена теорема с условиями для эффективной сходимости итерации 3-го порядка гладкости функционалов, представленных в виде композиции элементарных функций, даются формулы необходимого числа итераций для поиска экстремума конечномерного функционала.
Приведена разностная формула для эффективного поиска экстремума. Вторая теорема с разностной формулой простой итерации рассматривает априорно гладкие функционалы, не представимые в виде композиции элементарных функций, показана эквивалентность точности итерационных формул в теоремах, определена верхняя граница оптимального шага.
Приведены программы и примеры, подтверждающие эффективность доказанных теорем.
Ключевые слова: Гладкий функционал ,диагональное преобладание элементов матрицы Гесса, равномерная невырожденность,центральная разность первого порядка, оптимальный шаг итерационной формулы.
UDC 519.6 BBC 22.18 22.193
ABOUT EFFICIENT SEARCHING for of the UNCONDITIONAL EXTREMUM SMOOTH FUNKCIONALOV In KONECHNOMERNYH PROBLEM Golubeva O.V., Ehilevskiy S.G., Pastuhov YU.F., Pastuhov D.F.
The Abstract: is Received theorem with condition for efficient convergence of the iterations 3 orders to smoothness, presented in the manner of compositions elementary function, are given formulas of the necessary number iteration for searching for of the optimum certainly measured function.
The numerical formula is Brought for efficient searching for of the optimum. The Second theorem with the numerical formula iteration idle time considers a priori smooth to functions, not presented in the manner of compositions elementary function, is shown equivalence to accuracy iteration - molded the theorems, is determined upper border of the optimum step.
The Broughted programs and examples, confirming efficiency of the proved theorems.
The Keywords: Smooth function ,diagonal prevalence element matrixes Gessa, even absence of degeneracy,central first-order difference, optimum step to sequences.

Введение

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

Градиентные методы являются не достаточно эффективными с точки зрения точности. Действительно, пусть необходимо найти решение с абсолютной точностью 10^{-3}, а значение функционала в экстремальной точке с точностью порядка $\left(10^{-3}\right)^2 = 10^{-6}$ - 6 значащими цифрами. Зададим шаг итерации $h = 0.001$ с начальным расстоянием между первым приближением и стационарной точкой на 100 единиц, потребуется число $10^2 / 10^{-3} = 10^5$ итераций, в то время как можно получить решение и экстремальное значение гладкого функционала с абсолютной точностью 10^{-15} всего за 30 шагов итерации с тем же начальным удалении от стационарной точки. Число арифметических операций и ошибка округления пропорциональны размерности цикла. Одним из эффективных методов численного решения системы уравнений является метод Зейделя[2]. Даже более эффективным по сравнению с формулой касательных Ньютона (в случае если метод Зейделя и метод Ньютона применимы одновременно). В матричном методе Ньютона нужно вычислить n^2 элементов матрицы Якоби, погрешность дополнительно увеличивается при отыскании обратной матрицы к матрице Якоби:
$\overline{x^{m+1}} = \overline{x^m} - \left(F'\left(\overline{x^m}\right)\right)^{-1} * F\left(\overline{x^m}\right)$, где $\left(F'\left(\overline{x^m}\right)\right)^{-1}$ - матрица обратная к матрице Якоби. Метод Зейделя использует только диагональные элементы матрицы Якоби, т.е. n элементов. Поэтому при одинаковом числе итераций в методе Зейделя меньше элементарных операций и, следовательно, ошибка округления.

Определение 1. Среди двух методов, решающих одну и ту же задачу с одинаковой точностью при одинаковых начальных условиях, более эффективным назовём тот, который использует минимальное число элементарных операций(+,-,/,*).

Постановка задачи

Рассмотрим n - мерную $x = \{x_1, x_2, ..., x_n\} \in A \subset R^n$ задачу на безусловный экстремум достаточно гладкого функционала $f(x) \in C^3(A)$, т.е. множество функционалов трижды непрерывно дифференцируемых в открытой области A, имеющих непрерывные частные производные до 3 – го порядка включительно[1].

$$\begin{cases} f(x_1, x_2, ..., x_n) \to extr \\ (x_1, x_2, ..., x_n) \in A \subset R^n,\ f(x) \in C^3(A) \end{cases} \qquad (1)$$

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

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

где $(\overline{x_1}, \overline{x_2}, ..., \overline{x_n})$ - стационарная точка, т.е. решение системы уравнений (2).

Будем решать систему уравнений (2**) численно** методом простой итерации.

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

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

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

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

Обозначим вектор $\overline{\delta x^m} = \left(\delta x_1{}^m, \delta x_2{}^m, ..., \delta x_n{}^m\right) = \left(x_1^m - \overline{x_1}, x_2^m - \overline{x_2}, ..., x_n^m - \overline{x_n}\right)$, аналогично: $\overline{\delta x^{m+1}} = \left(\delta x_1{}^{m+1}, \delta x_2{}^{m+1}, ..., \delta x_n{}^{m+1}\right) = \left(x_1^{m+1} - \overline{x_1}, x_2^{m+1} - \overline{x_2}, ..., x_n^{m+1} - \overline{x_n}\right)$,где вектор $(\overline{x_1}, \overline{x_2}, ..., \overline{x_n})$ - представляет стационарную искомую точку решения системы (2) $\Leftrightarrow$ (5), т.е.:

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

Перейдём к старой переменной – функции $f(x_1, x_2, ..., x_n)$, учитывая обозначения (5)

$f'_{x_i}\left(x_1^{m+1}, x_2^{m+1}, ..., x_i^{m+1}, ..., x_n^m\right) = F_i\left(x_1^{m+1}, x_2^{m+1}, ..., x_i^{m+1}, ..., x_n^m\right), i = \overline{1, n}$.

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, в частных прозводных $f_{x_i}^{'}$, $f_{x_i x_i}^{''}$ переменные с порядком m и $m+1$ выразим через переменные δx_k^m, δx_k^{m+1}:

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

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

Сходимость метода

Теорема 1. (условия сходимости итерации (6)).
Пусть открытая область $A \subset R^n$ содержит начальную итерацию $\left(x_1^0, x_2^0, ..., x_n^0\right) \in A$ и стационарную точку $\left(\overline{x_1}, \overline{x_2}, ..., \overline{x_n}\right) \in A$ (решение системы уравнений (2)). Функция $f(x), x \in A$ конечного числа n переменных:

1) Трижды непрерывно дифференцируема $f(x) \in C^3(A)$
2) Матрица вторых частных производных (матрица Гесса) обладает диагональным преобладанием

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

3) Диагональные элементы матрицы Гесса равномерно не вырождены:

$$\left| f_{x_i x_i}^{''}(x) \right| \geq M_{2,i} > 0, i = \overline{1,n}, \forall x \in A \text{ где } M_{1,i}, M_{2,i}, i = \overline{1,n} \text{ - некоторые константы.}$$

$$\frac{M_{1,i}}{M_{2,i}} = q_i < 1, i = \overline{1,n} \text{ .Обозначим } q = \max_{i=\overline{1,n}} q_i$$

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

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

.

Доказательство проведём по индукции (достаточность).

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

1) $i=1$**:**

$$f'_{x_1}(x_1^m,x_2^m,...,x_n^m)=f'_{x_1}(\overline{x_1},\overline{x_2},...,\overline{x_n})+\sum_{j=1}^{n}f''_{x_1x_j}(\bar{x})\delta x_j^m+O\left(\alpha_1\left|\delta x^m\right|^2\right)=\sum_{j=1}^{n}f''_{x_1x_j}(\bar{x})\delta x_j^m+O\left(\alpha_1\left|\delta x^m\right|^2\right)$$ в силу справедливости (5). Где $\left|\delta x^m\right|=\max_{j=\overline{1,n}}\left|\delta x_j^m\right|,\ \left|\delta x_j^m\right|\le\left|\delta x^m\right|,\forall j=\overline{1,n},\forall m=0,1,2,...$. (7)

$\lim\limits_{\left|\delta x^m\right|\to 0}\alpha_i(\delta x^m)=C_i\ne 0, i=\overline{1,n}$.

В силу условий (5),(6),(7) и условия 1) теоремы:

$$\delta x_1^{m+1}=\delta x_1^m-\frac{f'_{x_1}(\overline{x_1}+\delta x_1^m,\overline{x_2}+\delta x_2^m,...,\overline{x_n}+\delta x_n^m)}{f''_{x_1x_1}(\overline{x_1}+\delta x_1^m,\overline{x_2}+\delta x_2^m,...,\overline{x_n}+\delta x_n^m)}=\delta x_1^m-\frac{f'_{x_1}\left(\bar{x}\right)+\sum\limits_{j=1}^{n}f''_{x_1x_j}(\bar{x})\delta x_j^m+O\left(\alpha_1\left|\delta x^m\right|^2\right)}{f''_{x_1x_1}(\overline{x_1}+\delta x_1^m,\overline{x_2}+\delta x_2^m,...,\overline{x_n}+\delta x_n^m)}=$$

$$\delta x_1^m-\frac{\sum\limits_{j=1}^{n}f''_{x_1x_j}(\bar{x})\delta x_j^m+O\left(\alpha_1\left|\delta x^m\right|^2\right)}{f''_{x_1x_1}(\overline{x_1}+\delta x_1^m,\overline{x_2}+\delta x_2^m,...,\overline{x_n}+\delta x_n^m)}=\delta x_1^m-\frac{\delta x_1^m}{\dfrac{f''_{x_1x_1}(\overline{x_1}+\delta x_1^m,\overline{x_2}+\delta x_2^m,...,\overline{x_n}+\delta x_n^m)}{f''_{x_1x_1}(\overline{x_1},\overline{x_2},...,\overline{x_n})}}-$$

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

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

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

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

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

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

Используя условие равномерной невырожденности 3) и условие диаганального преобладания 2) Т.1:

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

$$\left|\delta x_1^{m+1}\right|\le\left|\delta x^m\right|\frac{M_{1,1}}{M_{2,1}}=\left|\delta x^m\right|q_1\le\left|\delta x^m\right|q<\left|\delta x^m\right| \qquad (8.1)$$

3) Для произвольного $i-$го уравнение системы (6) продолжим.

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:

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

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

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

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

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

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

Сокращая промежуточные записи, получим:

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

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

Обозначение последней формулы отличается от общей нумерации в виду её важности.

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

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

$$\left|\delta x_i^{m+1}\right| \le \left|\delta x^m\right| \frac{M_{1,i}}{M_{2,i}} = \left|\delta x^m\right| q_i \le \left|\delta x^m\right| q < \left|\delta x^m\right| \qquad (8.i)$$

Индуктивно доказано справедливость $\left|\delta x_i^{m+1}\right| \le \left|\delta x^m\right| q < \left|\delta x^m\right|, i = \overline{1, n}$, поэтому

$$\left|\delta x^{m+1}\right| = \max_{i=\overline{1,n}} \left|\delta x_i^{m+1}\right| \le \left|\delta x^m\right| q < \left|\delta x^m\right|.$$

Таким образом, сходимость при выполнениии условий теоремы доказана.

4) Оценим погрешность метода.

Пусть $\left|\delta x^{m+1}\right| \le \left|\delta x^m\right| q \le \left|\delta x^{m-1}\right| q^2 \le ... \le \left|\delta x^1\right| q^m \le \left|\delta x^0\right| q^{m+1} \underset{m\to\infty}{\to} 0$, $\qquad (9)$

где $\left|\delta x^0\right| = \max_{i=\overline{1,n}} \left|\delta x_i^0\right| = \max_{i=\overline{1,n}} \left|x_i^0 - \overline{x_i}\right|$ начальное приближение стационарной точки, начальная точка $x^0 = \left(x_1^0, x_2^0, ..., x_n^0\right) \in A$.

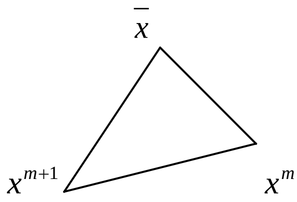

Рис.1

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:

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

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

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

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

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

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

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

Окончательно оценка погрешности имеет вид:

$$\left|\delta x^{m}\right| = \left|x^{m} - \overline{x}\right| \le \frac{(1+q)\left|\delta x^{0}\right| q^{m}}{1-q} \le \frac{(1+q)q^{m}}{(1-q)^{2}} l_0. \qquad (10)$$

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

$\left(\overline{x_1}, \overline{x_2}\right) = (0{,}0)$. $\quad f_{x_1x_1}\left(\overline{x_1}, \overline{x_2}\right) = f_{x_2x_2}\left(\overline{x_1}, \overline{x_2}\right) = 2,\ f_{x_1x_2}\left(\overline{x_1}, \overline{x_2}\right) = f_{x_2x_1}\left(\overline{x_1}, \overline{x_2}\right) = 3, q = \frac{3}{2} > 1$.

Согласно(*):

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

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

Откуда видно, что во всех чётных и нечётных итерациях каждый раз удаляются от стационарной точки, т.е. итерация расходится.

Замечание 2. Формула (*) 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 $\left(\overline{x_1}, \overline{x_2}, \ldots, \overline{x_n}\right) \in A$.

Замечание 3. Определение равномерной невырожденности матрицы приведено в[3]. Оно обеспечивает взимнооднозначное отображение окрестности стационарной точки и поля градиента (стационарная точка –

особая точка поля градиента) в каждой итерации (6) – теорема об обратной функции[1]. Если итерация задаётся формулой $\delta x^{m+1} = A\delta x^{m}$ - аналогом формулы (*), 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) и 3) Т.1) , то по теореме о неподвижной точке в метрических пространствах[4] сжимающее отображение имеет единственное решение. Таким образом, единственность решения итерации(6) доказана.

Замечание 4. Сходимость (6)выполнена даже при разных значениях $\frac{M_{1,i}}{M_{2,i}} = q_i < 1, i = \overline{1,n}$. Если по всем переменным $i = \overline{1,n}$ в итерации (6) недиагональные элементы матрицы Гесса $\frac{\partial^2 f(x)}{\partial x_i \partial x_j} \equiv 0, j,i = \overline{1,n}, j \ne i$.

В этом случае скорость сходимости (6) не линейная, а квадратичная.
В качестве экстремальной задачи, 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 $f(x_1,x_2,x_3)$:

$$\nabla f(x_1,x_2,x_3) = (2x_1 - x_2 + 1,\ 2x_2 - x_1, 2x_3 - 2).$$

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

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

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

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

Запишем итерацию по формуле (6):

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

Из (10) необходимое число итераций $N \ge \ln\left(\frac{\left|\delta x^m\right|}{\left|\delta x^0\right|}\left(\frac{1-q}{1+q}\right)\right) \Big/ \ln q = \ln\left(\frac{\left|\delta x^0\right|}{\left|\delta x^m\right|}\left(\frac{1+q}{1-q}\right)\right) \Big/ \ln 1/q$ (11)

$$N \ge \ln\left(\frac{\left|\delta x^m\right|}{l_0}\left(\frac{(1-q)^2}{1+q}\right)\right) \Big/ \ln q = \ln\left(\frac{l_0}{\left|\delta x^m\right|}\left(\frac{1+q}{(1-q)^2}\right)\right) \Big/ \ln 1/q \qquad (12)$$

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

Составим на языке си 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;
        printf("Galeev - Tihomirov (primer p.169)\n");
        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*x1+x2*x2 +x3*x3-x1*x2+x1-2*x3\n");
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 Гесса 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 условия теоремы 1. Тогда итерационные формулы (6) и (13) сравнимы по точности с порядком $O(h^2) = O\left(\left|\delta x^m\right|^2\right)$ и справедливы результаты теоремы 1. Оценка верхней границы для оптимального шага определяется формулами:

$$\left(h_i^m\right)_{apper}{}^2=\left|\frac{240 f_{x_i}^{(3)}\left(\bar{x}\right) f_{x_i}^{''}\left(\bar{x}\right)}{20 f_{x_i}^{(3)}\left(\bar{x}\right) f_{x_i}^{(4)}\left(\bar{x}\right)-3 f_{x_i}^{(5)}\left(\bar{x}\right) f_{x_i}^{''}\left(\bar{x}\right)}\right|, i=\overline{1,n}\ \left(\left(h_i^m\right)_{apper}{}^2=\left|\frac{12 f_{x_i}^{''}\left(\bar{x}\right)}{f_{x_i}^{(4)}\left(\bar{x}\right)}\right|, i=\overline{1,n}\right).$$

Доказательство:
Разложим в ряд Тейлора числитель и знаменатель формул (6) и учтём, что компилятор автоматически присваивает значениям переменных $\delta x_k^{m+1} \to \delta x_k{}^m, k=\overline{1,i-1}$ в формулах, обозначим $x^m=\left(x_1^m,\ldots,x_n^m\right)$:

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

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

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

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

$$\begin{cases} x_1^{m+1}=x_1^m-\dfrac{\dfrac{f\left(x_1^m+\dfrac{h}{2},x_2^m,\ldots,x_n^m\right)-f\left(x_1^m-\dfrac{h}{2},x_2^m,\ldots,x_n^m\right)}{h}+O\left(h^2\right)}{\dfrac{f\left(x_1^m+h,x_2^m,\ldots,x_n^m\right)+f\left(x_1^m-h,x_2^m,\ldots,x_n^m\right)-2 f\left(x_1^m,x_2^m,\ldots,x_n^m\right)}{h^2}+O\left(h^2\right)}, \\ \ldots. \\ x_i^{m+1}=x_i^m-\dfrac{\dfrac{f\left(x_1^m,\ldots,x_i^m+\dfrac{h}{2},\ldots,x_n^m\right)-f\left(x_1^m,\ldots,x_i^m-\dfrac{h}{2},\ldots,x_n^m\right)}{h}+O\left(h^2\right)}{\dfrac{f\left(x_1^m,\ldots,x_i^m+h,\ldots,x_n^m\right)+f\left(x_1^m,\ldots,x_i^m-h,\ldots,x_n^m\right)-2 f\left(x_1^m,x_2^m,\ldots,x_n^m\right)}{h^2}O\left(h^2\right)} \\ \ldots. \\ x_i^{m+1}=x_i^m-\dfrac{\dfrac{f\left(x_1^m,x_2^m,\ldots,x_n^m+\dfrac{h}{2}\right)-f\left(x_1^m,x_2^m,\ldots,x_n^m-\dfrac{h}{2}\right)}{h}+O\left(h^2\right)}{\dfrac{f\left(x_1^m,x_2^m,\ldots,x_n^m+h\right)+f\left(x_1^m,x_2^m,\ldots,x_n^m-h\right)-2 f\left(x_1^m,x_2^m,\ldots,x_n^m\right)}{h^2}+O\left(h^2\right)} \end{cases} \qquad (13)$$

Откуда видно, что (6) отличается от (13) с точностью до $O\left(h^2\right)$. Заменим h на $\left|x_i{}^{m+1}-x_i{}^m\right|=h_i^m$ - расстояние между соседними узлами.

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

$\dfrac{h_i^m}{(1+q)}=\dfrac{\left|x_i{}^{m+1}-x_i{}^m\right|}{(1+q)} \le \left|\delta x_i^m\right| \le \dfrac{\left|x_i{}^{m+1}-x_i{}^m\right|}{(1-q)}=\dfrac{h_i^m}{(1-q)}$, $(1-q)\left|\delta x_i^m\right| \le h_i^m \le (1+q)\left|\delta x_i^m\right|$. Значит, шаг схемы с хорошей точностью можно заменить на $\left|\delta x_i^m\right|=h_i^m$ с сохранением точности

$O\left(h_i^{m^2}\right)=O\left(\left|\delta x_i^m\right|^2\right)$.Поэтому схемы (6) и (13) эквивалентны с одинаковой точностью. Если в теореме 2 выполнены условия теоремы 1, то существует единственное решение итерации (13), сходящееся к решению системы(2), число итераций определяется, по – прежнему, формулами(11),(12). Теорема доказана.

Кроме того, систему уравнений (13) удобнее преобразовать к форме(14). На первый взгляд, формулы (14) от шага h_i^m не звисят. Действительно, числитель и знаменатель дроби прямопропорциональны $\left(h_i^m\right)^2$ и сокращаются на $\left(h_i^m\right)^2$. Поэтому, для оптимального шага с наименьшей погрешностью нужно учесть члены более высокого порядка $O\left(h_i^m\right)^2$.

$$\begin{cases} x_1^{m+1}=x_1^m-h_1^m\dfrac{f\left(x_1^m+\dfrac{h_1^m}{2},x_2^m,...,x_n^m\right)-f\left(x_1^m-\dfrac{h_1^m}{2},x_2^m,...,x_n^m\right)}{f\left(x_1^m+h_1^m,x_2^m,...,x_n^m\right)+f\left(x_1^m-h_1^m,x_2^m,...,x_n^m\right)-2f\left(x_1^m,x_2^m,...,x_n^m\right)}+O\left(\left(h_1^m\right)^2\right)\\ \quad .... \\ x_i^{m+1}=x_i^m-h_i^m\dfrac{f\left(x_1^m,...,x_i^m+\dfrac{h_i^m}{2},...,x_n^m\right)-f\left(x_1^m,...,x_i^m-\dfrac{h_i^m}{2},...,x_n^m\right)}{f\left(x_1^m,...,x_i^m+h_i^m,...,x_n^m\right)+f\left(x_1^m,...,x_i^m-h_i^m,...,x_n^m\right)-2f\left(x_1^m,x_2^m,...,x_n^m\right)}+O\left(\left(h_i^m\right)^2\right)\\ \quad .... \\ x_n^{m+1}=x_n^m-h_n^m\dfrac{f\left(x_1^m,x_2^m,...,x_n^m+\dfrac{h_n^m}{2}\right)-f\left(x_1^m,x_2^m,...,x_n^m-\dfrac{h_n^m}{2}\right)}{f\left(x_1^m,x_2^m,...,x_n^m+h_n^m\right)+f\left(x_1^m,x_2^m,...,x_n^m-h_n^m\right)-2f\left(x_1^m,x_2^m,...,x_n^m\right)}+O\left(\left(h_n^m\right)^2\right) \end{cases} \quad (14)$$

Получим, преобразуя i - е уравнение (14):

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

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

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

$$\left.+\left(h_i^m\right)^4\left(\frac{1}{144}\left(\frac{f^{(4)}_{x_i}\left(x^m\right)}{f''_{x_i}\left(x^m\right)}\right)^2+\frac{f^{(5)}_{x_i}\left(x^m\right)}{1920f'_{x_i}\left(x^m\right)}-\frac{f^{(6)}_{x_i}\left(x^m\right)}{360f''_{x_i}\left(x^m\right)}-\frac{f^{(3)}_{x_i}\left(x^m\right)}{288f'_{x_i}\left(x^m\right)}\frac{f^{(4)}_{x_i}\left(x^m\right)}{f''_{x_i}\left(x^m\right)}\right)+O\left(\left(h_i^m\right)^6\right)\right),i=\overline{1,n} \quad (15)$$

Чтобы формула(15) имела порядок точности $O\left(\left(h_i^m\right)^6\right)$ необходимо:

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

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

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

Учитывая систему уравнений (2) $f_{x_i}^{'}\left(x^m\right)\approx f_{x_i}^{'}\left(\overline{x}\right)=0$, упростим (16):

$$\left(h_i^m\right)^2=\left|\frac{240f_{x_i}^{(3)}\left(x^m\right)f_{x_i}^{''}\left(x^m\right)}{20f_{x_i}^{(3)}\left(x^m\right)f_{x_i}^{(4)}\left(x^m\right)-3f_{x_i}^{(5)}\left(x^m\right)f_{x_i}^{''}\left(x^m\right)}\right|\approx\left|\frac{240f_{x_i}^{(3)}\left(\overline{x}\right)f_{x_i}^{''}\left(\overline{x}\right)}{20f_{x_i}^{(3)}\left(\overline{x}\right)f_{x_i}^{(4)}\left(\overline{x}\right)-3f_{x_i}^{(5)}\left(\overline{x}\right)f_{x_i}^{''}\left(\overline{x}\right)}\right|,i=\overline{1,n} \quad (17)$$

Для полиномов n переменных не выше четвёртой степени: $\left(h_i^m\right)^2=\left|\dfrac{12f_{x_i}^{''}\left(\overline{x}\right)}{f_{x_i}^{(4)}\left(\overline{x}\right)}\right|,i=\overline{1,n}$ (18)

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

$f_{x_1}^{'}\left(\overline{x}\right)=4\overline{x_1}^3-\overline{x_2}+2\overline{x_1}=0$, $f_{x_2}^{'}\left(\overline{x}\right)=4\overline{x_2}^3-\overline{x_1}+2\overline{x_2}=0\Rightarrow\left(\overline{x_1}-\overline{x_2}\right)\left(4\left(\overline{x_1}^2+\overline{x_2}^2+\overline{x_1x_2}\right)+3\right)=0$

$\Leftrightarrow\overline{x_1}=\overline{x_2}\Leftrightarrow4\overline{x_1}^3+\overline{x_1}=0\Leftrightarrow\overline{x_1}=\overline{x_2}=0$. $\dfrac{\partial^2f\left(\overline{x_1},\overline{x_2}\right)}{\partial x_i\partial x_j}=\begin{vmatrix}12\overline{x_1}^2+2 & -1\\ -1 & 12\overline{x_2}^2+2\end{vmatrix}=\begin{vmatrix}2 & -1\\ -1 & 2\end{vmatrix}$

По теореме 1 матрица Гесса имеет диаганальное преобладание и равномерно невырождена. Оптимальный шаг найдём по формуле (18) $\left(h_1^m\right)^2=\left(h_2^m\right)^2=\dfrac{12*2}{24}=1,h_1^m=h_2^m=1$.

Замечание1 (Т2). Как видно из приведенного примера $\left(h_i^m\right)^2$ велико, (должно быть $\left(h_i^m\right)^2<<1$ для выполнения асимптотики $O\left(\left(h_i^m\right)^6\right)$), что явно не выполняется. Поэтому формулы(17),(18) можно рассматривать как верхнюю оценку величины h_i^m ,т.е. $0<h_i^m\le\left(h_i^m\right)_{apper}$. Можно предложить следующий алгоритм выбора $\left(h_i^m\right)=\left\{\left(h_i^m\right)_{apper},10^{-1}\left(h_i^m\right)_{apper},10^{-2}\left(h_i^m\right)_{apper}\right\}$, Оптимальным шагом считаем тот для которого значение функции наименьшее (наибольшее), рассчитанное программой. Существование оптимального шага $0<h_i^m\le\left(h_i^m\right)_{apper}$ подтверждает численный эксперимент, когда погрешность вычислений экстремума 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 си (в данной программе $h_1^m=\left(h_1^m\right)_{apper}=h_2^m=\left(h_2^m\right)_{apper}=1$, при этом мы 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,delta1,delta2;
double h1,h2;
n=60;
        a1=10.0;
```

```
            a2=10.0;
            h1=1e0;
            h2=1e0;
            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=%.20lf,x2=%.20lf,extr=%.20lf\n",b1,b2,fx(b1,b2));
printf(" h1=%.16lf,h2=%.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

$$x_1 = 0.00000000000000111, x_2 = 0.00000000000000117, extr = 1.00000000000000000000$$

Выводы

1) В первой теореме получены условия для эффективной сходимости итерации 3-го порядка гладкости функционалов представленных в виде композиции элементарных функций, даются формулы необходимого числа итераций для поиска экстремума конечномерного функционала.
2) Вторая теорема доказана для априорно гладких функционалов, 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 теорем.

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.