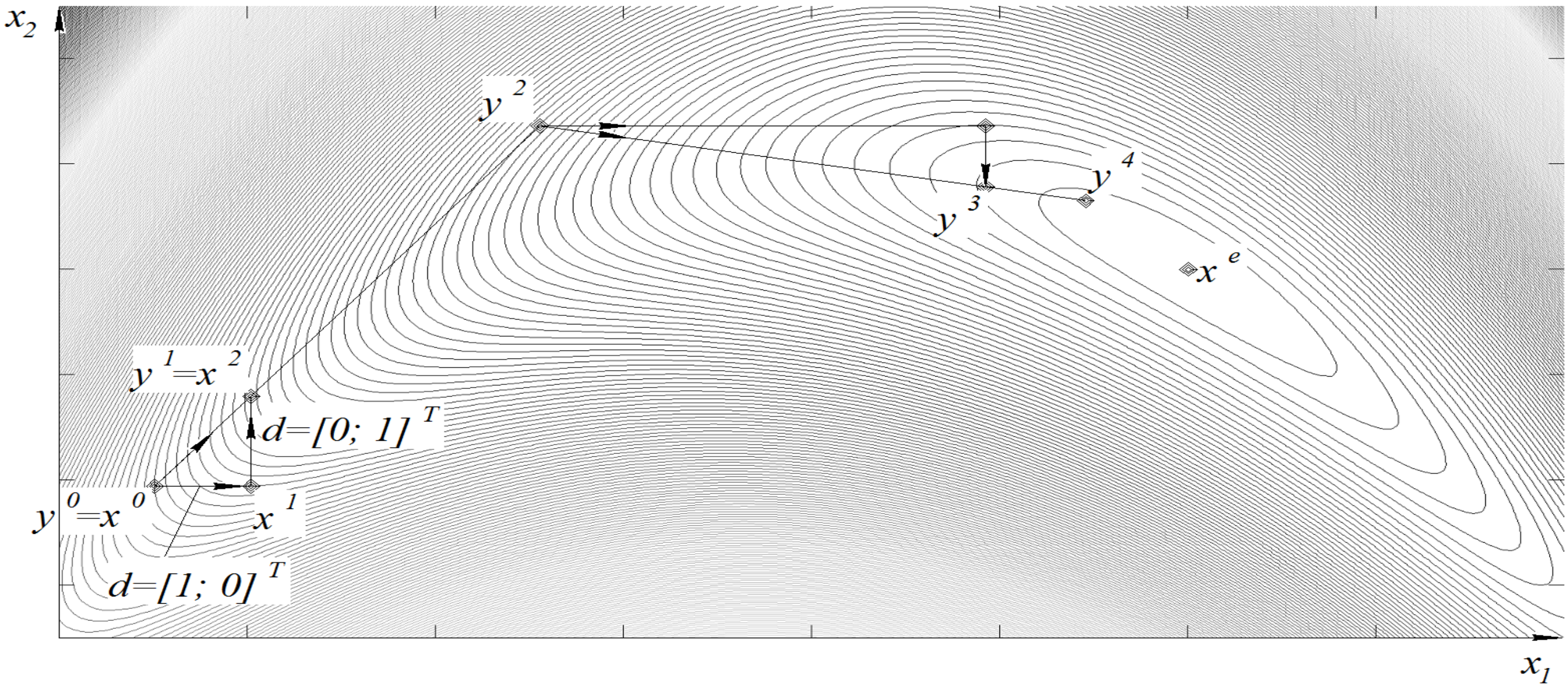
ЛЕКЦИЯ № 8

Методы многомерной оптимизации

Методы многомерной оптимизации без ограничений.

Метод Хука-Дживса (Х-Д)



Геометрическая интерпретация метода Хука-Дживса в случае двухпараметрической задачи

Метод Хука-Дживса (Х-Д)

Алгоритм:

Ш.0. Выбор: $y^0 \in X \in E^n$ (стартовая точка), ε, δ , $l = 0$ (счётчик точек).

Ш.1. Выполняется покоординатный спуск – однократный проход: от точки $x^0 \equiv y^l$ поочередно по каждой координате $x_k \left(k = \overline{1, n}\right)$ к точке x^n

Ш.2. Осуществляется движение «по образцу»: $y^{(p+1)} = y^{(p)} + \Delta y^{(p)}; \Delta y^{(p)} = \alpha_{II}^{*(p)} d^{(p)}$

$\alpha_{II}^{*(p)} = Arg \min \varphi(y^{(p)} + \alpha^{(p)} d^{(p)})$

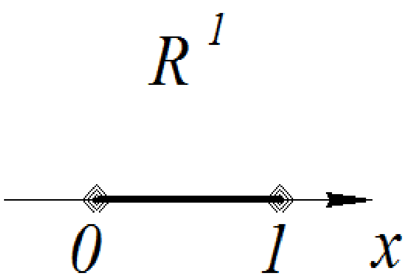
Ш3. Проверка сходимости:

Если $\left[\left(\left\| y^{(p+1)} - y^{(p)} \right\| \le \delta\right) \wedge \left(\left| \varphi(y^{p+1}) - \varphi(y^{(p)}) \right| \le \varepsilon\right)\right]$, то $\left\{x^e = y^{(p+1)}; \text{конец}\right\}$

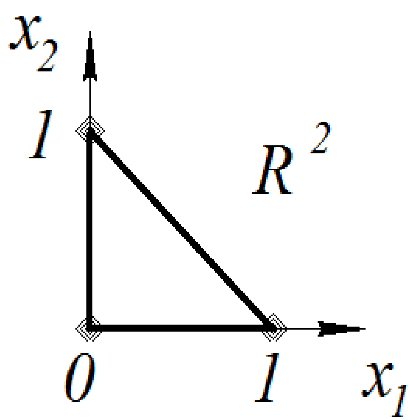
иначе $\left\{x^0 = y^{(p+1)}; k = 0; p = p + 1; \text{переход на Ш1.}\right\}$

.

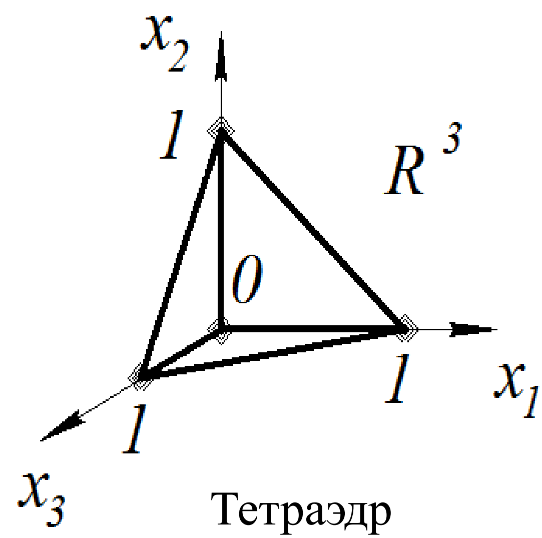
Метод Нелдера-Мида (деформируемых симплексов)



Отрезок [0 1].



Треугольник



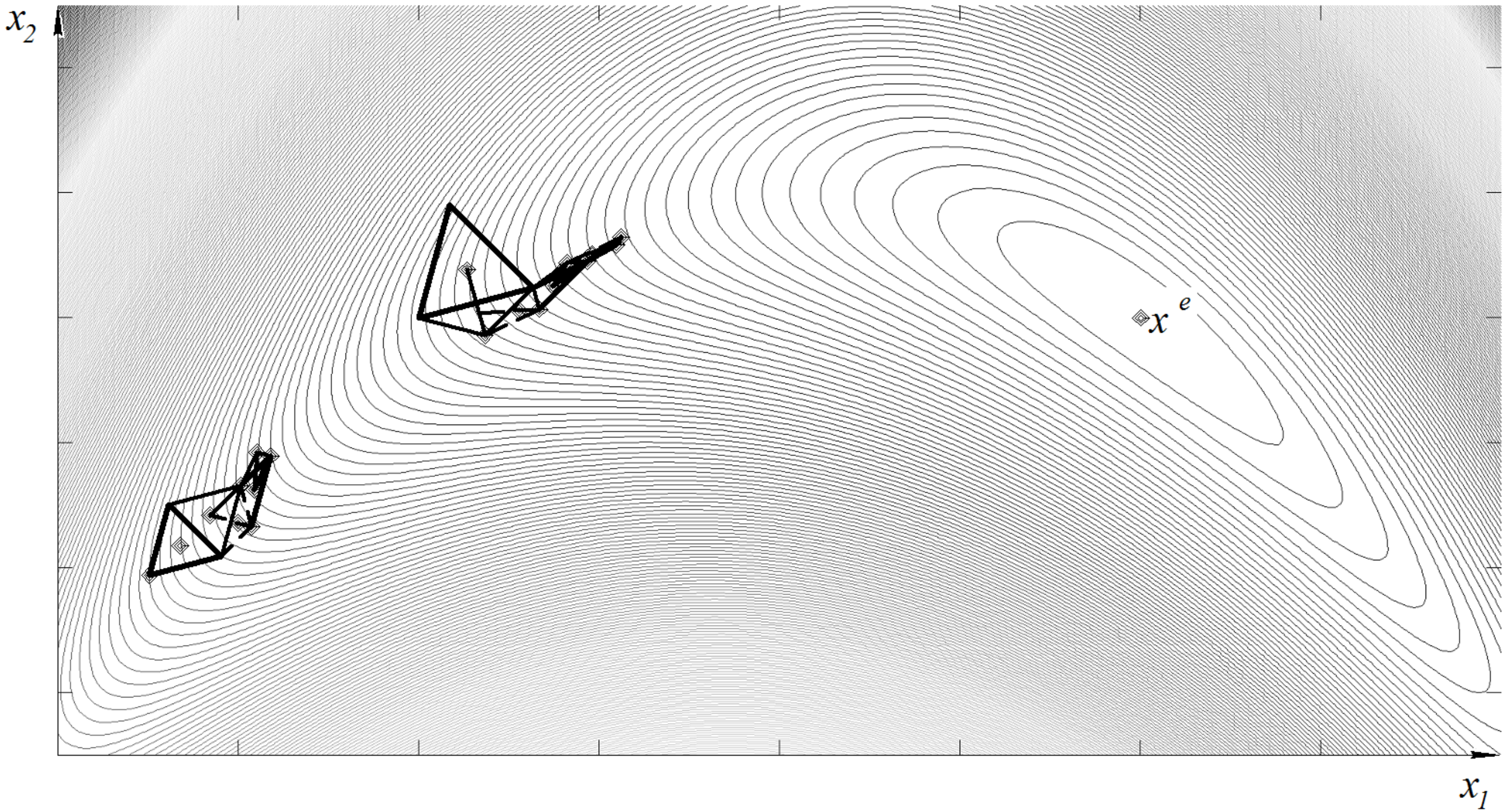
Тетраэдр

Метод Нелдера-Мида (деформируемых симплексов)

Идея

- Пусть целевая функция – овражная. Число граничных точек – N+1, где N – размерность пространства варьируемых переменных. Строим симплекс из начальной точки. Ищем его центр тяжести и из наихудшей точки через центр тяжести проводим прямую и строим новый симплекс. Симплексы постепенно уменьшаются при подходе к экстремуму.
- Метод деформируемых симплексов не имеет ничего общего с симлекс-методом линейного программирования.

Метод Нелдера-Мида (деформируемых симплексов)



Геометрическая интерпретация метода Нелдера-Мида в случае двухпараметрической задачи

Метод Нелдера-Мида (деформируемых симплексов)

<u>Алгоритм</u>

<u>Ш0.</u> Задание: 1) начального шага ***S***; 2) точности ε 3) начальной точки $X^{(0)}$

коэффициенты $\begin{cases} \alpha = 1 \text{ - коэф.отражения} \\ \beta = 0{,}5 \text{ - коэф. сжатия} \\ \gamma = 2 \text{ - коэф. растяжения} \end{cases}$

<u>Ш1.</u> Построение исходного симплекса (т.е. нахождение n + 1 граничных точек)

$$x_i^{(0)} = x_1^{(0)} + D_i \text{, где } D_i = \begin{bmatrix} d_1 \\ d_2 \\ \dots \\ d_2 \end{bmatrix},\ D_2 = \begin{bmatrix} d_2 \\ d_1 \\ \dots \\ d_2 \end{bmatrix}, \dots, D_n = \begin{bmatrix} d_2 \\ d_2 \\ \dots \\ d_1 \end{bmatrix}$$

$$d_1 = \frac{s}{n\sqrt{2}}(\sqrt{n+1}+n-1) \qquad d_2 = \frac{s}{n\sqrt{2}}(\sqrt{n+1}-1)$$

n – размерность пространства, d1 – смещается в соответствии с номером точки.

Метод Нелдера-Мида (деформируемых симплексов)

Ш2. Вычислить $x_l^{(k)} = Arg\min\left\{\varphi(x_i^{(k)})\right\}$

$x_h^{(k)} = Arg\max\left\{\varphi(x_i^{(k)})\right\}$

Ш3. Вычислить и.т. симплекса . $x_{n+2}^{(k)} = \frac{1}{n}\left[\sum_{i=1}^{n+1} x_i^{(k)} - x_h^{(k)}\right]$

Ш4. Отражение. $x_{n+3}^{(k)} = x_{n+2}^{(k)} + \alpha(x_{n+2}^{(k)} - x_h^{(k)})$

.

Ш5. 1) Если $\left[\varphi(x_{n+3}^{(k)}) \le \varphi(x_l^{(k)})\right]$, то $x_{n+4}^{(k)} = x_{n+2}^{(k)} + \gamma(x_{n+3}^{(k)} - x_{n+2}^{(k)})$

2) Если $\left[\varphi(x_{n+3}^{(k)}) > \varphi(x_i^{(k)}), \forall i \ne h\right]$, то $x_{n+5}^{(k)} = x_{n+2}^{(k)} + \beta(x_h^{(k)} - x_{n+2}^{(k)})$

3) Если $\left[\varphi(x_{n+3}^{(k)}) > \varphi(x_h^{(k)})\right]$, то $x_i^{(k)} = x_i^{(k)} + 0.5(x_i^{(k)} - x_l^{(k)})$

– редукция всего симплекса в сторону наилучшего решения.

Ш6. Если $\left[\frac{1}{n+1}\sum_{i=1}^{n+1}\left[\varphi(x_i^{(k)}) - \varphi(x_l^{(k)})\right]^2\right]^{1/2} \le \varepsilon$, то $\left\{x^e = x_l; \text{конец}\right\}$

иначе $\left\{k = k+1; \text{переход на Ш2.}\right\}$

Методы, использующие производные.

Как мы уже видели, простейшие методы I порядка (метод наискорейшего спуска) и II порядка (метод Ньютона) в чистом виде не очень удобны для вычислений. В связи с этим были разработаны методы, которые позволяют преодолеть их недостатки. Наиболее распространенными в настоящее время методами являются метод сопряженных градиентов (метод Флетчера-Ривза), который является усовершенствованием метода наискорейшего спуска и квазиньютоновские методы, которые представляют собой развитие метода Ньютона.

Метод Флетчера-Ривза

Метод сопряженных градиентов состоит в построении матрицы сопряжения Q(k) такой, что $d^{(k+1)T}Q^{(k)}d^{(k)}=0$. Флетчером-Ривзом была предложена процедура , состоящая в следующем:

Строится последовательность точек: $x^{(k+1)}=x^{(k)}+\alpha^{*(k)}d^{(k)}$

$$d^{(k)}=-\nabla\varphi(x^{(k)})+\omega^{(k)}d^{(k-1)}$$

$$\omega^{(k)}=-\frac{\nabla\varphi(x^{(k)})^T\nabla\varphi(x^{(k)})}{\nabla\varphi(x^{(k-1)})^T\nabla\varphi(x^{(k-1)})}=-\frac{\left\|\nabla\varphi(x^{(k)})\right\|^2}{\left\|\nabla\varphi(x^{(k-1)})\right\|^2}$$

– поправка, формирование матрицы сопряжения Q(k).

Алгоритм метода Флетчера-Ривза.

Ш0. Выбор параметров точности $\delta, \varepsilon, \mathrm{k} = 0.$

Ш1. Выбор начальной точки $x^{(0)}$. Вычислить $\varphi(x^{(0)}), \nabla\varphi(x^{(0)}), d^{(0)} = -\nabla\varphi(x^{(0)})$

Ш2. Вычислить $\alpha^{*(k)} = Arg\min\varphi(x^{(k)} + \alpha^{(k)} d^{(k)}); x^{(k+1)} = x^{(k)} + \alpha^{*(k)} d^{(k)}$

Ш3. Вычислить $\omega^{(k+1)} = -\dfrac{\left\|\nabla\varphi(x^{(k+1)})\right\|^2}{\left\|\nabla\varphi(x^{(k)})\right\|^2}; d^{(k+1)} = -\nabla\varphi(x^{(k+1)}) + \omega^{(k+1)} d^{(k)}.$

Ш4. Проверка сходимости:

Если $(\left\|d^{(k+1)}\right\| \le \varepsilon) \vee (\left\|x^{(k+1)} - x^{(k)}\right\| \le \delta)$, то $\left\{x^e = x^{(k+1)}; \text{конец}\right\}$

Иначе {если $[k+1 < n]$, то $\left\{x^{(k)} = x^{(k+1)}; k = k+1; \text{переход на Ш2}\right\}$

если [k=n], то $\left\{x^{(0)} = x^{(k+1)}; k = 0; \text{переход на Ш1}\right\}$

$\bar{x}$

Квазиньютоновские методы

(Методы переменной метрики).

Идея метода:

Продифферинцируем разложение в ряд Тейлора функцию φ(x) в окрестности т. $\bar{x}$ т.е. $U_\varepsilon(\bar{x})$.

$$\varphi(\bar{x}+\Delta x)=\varphi(\bar{x})+\nabla\varphi(\bar{x})^T\Delta x+0(\|\Delta x\|^2)$$

Обозначим:

$$x^{(k)}=\bar{x}$$

$$\Delta x^{(k)}=\Delta x$$

$$\nabla\varphi(x^{(k)})=\nabla\varphi(\bar{x})$$

$$\nabla(\varphi(x^{(k+1)}))=\nabla\varphi(x^{(k)}+\Delta x^{(k)})$$

$$\nabla\varphi(\bar{x}+\Delta x)=\nabla\varphi(\bar{x})+H(\bar{x})^T\Delta x+\ldots$$

Тогда

$$\Delta x^{(k)}=x^{(k+1)}-x^{(k)}$$

$$\Delta g^{(k)}=\nabla\varphi(x^{(k+1)})-\nabla\varphi(x^{(k)})$$

$$\Delta g^{(k)}=H(x^{(k)})\Delta x^{(k)}$$

откуда $\Delta x^{(k)}=H^{-1}\Delta g^{(k)}$,

$H^{-1(k)}$ – обратная матрица Гессе, требующая вычисления 2-х частных производных.

Идея квазиньютоновских методов состоит в аппроксимации матрицы Гессе матрицами, корректирующими на каждой итерации исходную матрицу так, чтобы через n-итераций получить матрицу близкую к матрице Гессе.

Квазиньютоновские методы (Методы переменной метрики).

Обозначим $\Delta x^{(k)} = \widetilde{H}^{(k+1)}\Delta g^{(k)}$, где $\widetilde{H}^{(k+1)}$ - аппроксимация обратной матрицы Гессе.

Пусть $\widetilde{H}^{(k+1)} = \widetilde{H}^{(k)} + \Delta\widetilde{H}^{(k)}$, тогда $\Delta x^{(k)} = (\widetilde{H}^{(k)} + \Delta\widetilde{H}^{(k)})\Delta g^{(k)}$

или $\Delta x^{(k)} = \widetilde{H}^{(k)}\Delta g^{(k)} + \Delta\widetilde{H}^{(k)}\Delta g^{(k)}$. Выразим $\Delta\widetilde{H}^{(k)}$ через $\widetilde{H}^{(k)}, \Delta g^{(k)}$ и вспомогательные векторы y и z. Тогда: $\Delta\widetilde{H}^{(k)} = \dfrac{\Delta x^{(k)} y^T}{y^T \Delta g^{(k)}} - \dfrac{\widetilde{H}^{(k)}\Delta g^{(k)} z^T}{z^T \Delta g^{(k)}}$,

то есть $\Delta\widetilde{H}^{(k)} = A^{(k)} - B^{(k)}$. Необходимо, чтобы векторы y и z имели размерность $\Delta x, \Delta g$.

В зависимости от подбора векторов y и z получим различные варианты метода переменной метрики.

$$A^{(k)} = \frac{\Delta x^{(k)} y^T}{y^T g^{(k)}}, \quad B^{(k)} = \frac{\widetilde{H}^{(k)}\Delta g^{(k)} z^T}{z^T \Delta g^{(k)}}.$$

Квазиньютоновские методы
(Методы переменной метрики).

1. Метод Девидона – Флетчера – Пауэлла (1964г. DFP).

$$y^{(k)} = \Delta x^{(k)}$$

$$z^{(k)} = \widetilde{H}^{(k)} \Delta g^{(k)}$$

$$\Delta \widetilde{H}^{(k)} = \frac{\Delta x^{(k)} \Delta x^{(k)T}}{\Delta x^{(k)T} \Delta g^{(k)}} - \frac{\widetilde{H}^{(k)} \Delta g^{(k)} \Delta g^{(k)T} \widetilde{H}^{(k)T}}{\Delta g^{(k)T} \widetilde{H}^{(k)T} \Delta g^{(k)}}$$

Квазиньютоновские методы (Методы переменной метрики).

2.Метод Бройдена (B).

$$y^{(k)} = z^{(k)} = (\Delta x^{(k)} - \widetilde{H}^{(k)} \Delta g^{(k)})$$

$$\Delta \widetilde{H}^{(k)} = \frac{\Delta x^{(k)} (\Delta x^{(k)} - \widetilde{H}^{(k)} \Delta g^{(k)})^T}{(\Delta x^{(k)} - \widetilde{H}^{(k)} \Delta g^{(k)})^T \Delta g^{(k)}} - \frac{\widetilde{H}^{(k)} \Delta g^{(k)} (\Delta x^{(k)} - \widetilde{H}^{(k)} \Delta g^{(k)})^T}{(\Delta x^{(k)} - H^{(k)} \Delta g^{(k)})^T \Delta g^{(k)}}.$$

Здесь так же как и в методе сопряженных градиентов на каждой итерации формируется новая матрица сопряжения $Q^{(k)}$, где

$$Q^{(k)} = \widetilde{H}^{-1(k)}; \; d^{(k+1)T} Q^{(k)} d^{(k)} = 0 \;;$$

$$\mathrm{x}^{(\mathrm{k}+1)} = x^{(k)} + \alpha^{*(k)} d^{(k)};$$

При этом: $$d^{(k)} = -(\widetilde{H}^{(k)} + \Delta \widetilde{H}^{(k)})^T \nabla \varphi(x^{(k)}) \;;$$

$$\alpha^{*(k)} = Arg \min \varphi(x^{(k)} + \alpha^{(k)} d^{(k)}) \,.$$

Квазиньютоновские методы
(Методы переменной метрики).

Алгоритм квазиньютоновских методов.

Ш0. Выбор параметров точности ε,δ, k=0. Выбор нач. точки $x^{(0)}$.

Ш1. Вычислить: $\left\{\nabla\varphi(x^{(0)}); \tilde{H}^{(0)} = I; d^{(0)} = -\tilde{H}^{(0)}\nabla\varphi(x^{(0)});\right\}$

Ш2. Вычислить: $\left\{\alpha^{*(k)} = Arg\min\varphi(x^{(k)} + \alpha^{*(k)}d^{(k)}); x^{(k+1)} = x^{(k)} + \alpha^{*(k)}d^{(k)}\right\}$

Ш3. Вычислить: $\left\{\nabla\varphi(x^{(k+1)}); \Delta g^{(k)} = \nabla\varphi^{(k+1)} - \nabla\varphi^{(k)};\right\}$

Ш4. Вычислить: $\left\{\Delta\tilde{H}^{(k)} = A^{(k)} - B^{(k)}; \tilde{H}^{(k+1)} = \tilde{H}^{(k)} + \Delta\tilde{H}^{(k)}; d^{(k+1)} = -\tilde{H}^{(k+1)}\nabla\varphi\left(x^{(k+1)}\right)\right\}$

Ш5. Если $\left[\left(\left\|d^{(k+1)}\right\| \le \varepsilon\right) \vee \left(\left|x^{(k+1)} - x^{(k)}\right| \le \delta\right)\right]$, то $\left\{x^{e} = x^{(k+1)}; КОНЕЦ\right\}$

Иначе $\left\{x^{(k)} = x^{(k+1)}; k = k+1; \text{переход на Ш2.}\right\}$