ЛЕКЦИЯ № 7

Методы одномерной оптимизации

Методы одномерной оптимизации

После того, как выбрано направление поиска $d^{(k)}$, определяется величина шага

$\alpha^{*(k)}$ – минимизирующая функцию $\varphi(\alpha^{(k)})$ в этом направлении.

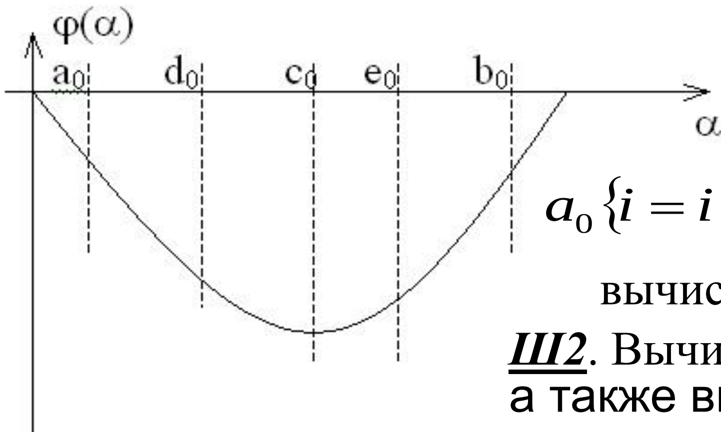
Таким образом отрабатывается композиция 2-х алгоритмов. Мы рассмотрим работу второго алгоритма – выбор величины шага в заданном направлении:

$$\alpha^{*(k)} = Arg\min_{\alpha^{k}\in E^1} \varphi(x^{(k)} + \alpha^{(k)} d^{(k)}) = Arg\min_{\alpha^{(k)}\in E^1} \varphi(\alpha^{(k)})$$

В этом случае также могут быть использованы методы, основанные на вычислении производных. Однако в реальных алгоритмах они используются редко. Вместо них широкое применение нашли методы одномерного поиска и методы интерполяции.

Методы одномерного поиска (методы стягивающихся отрезков).

1. Метод половинного деления (дихотомия).



Алгоритм:

Ш1. Задание δ, ε, j=0, N–число шагов для определения b0. I=0, α_0 , $h=\dfrac{\left\|x^{*(k)}-x^{*}\right\|}{N}$ – величина шага для определения b0.

$$a_0\{i=i+1;\alpha_i=\alpha_0+ih;b_0=\alpha_i\}WHILE(\varphi(\alpha_i)<\varphi(\alpha_0))$$

вычисление $\varphi(a_0),\varphi(b_0),a_0=\alpha_i,b_0=\alpha_{i+1}$

Ш2. Вычисление точек c_0 , d_0, e_0, $c_0=\dfrac{a_0+b_0}{2},d_0=\dfrac{a_0+c_0}{2},e_0=\dfrac{c_0+b_0}{2}$. а также вычисление функций $\varphi(c_0),\varphi(d_0),\varphi(e_0)$.

Ш3. 1)Если $\left|(\varphi(d_j)<\varphi(c_j))\wedge(\varphi(c_j)\le\varphi(e_j))\right|$, то

$$\left\{a_{j+1}=a_j,b_{j+1}=e_j,c_{j+1}=\frac{a_j+e_j}{2};d_{j+1}=\frac{a_{j+1}+c_{j+1}}{2};e_{j+1}=\frac{c_{j+1}+b_{j+1}}{2};вычс\,\varphi(c_{j+1}),\varphi(d_{j+1}),\varphi(e_{j+1})\right\}$$

2) Если $\left|(\varphi(d_j)\ge\varphi(c_j))\wedge(\varphi(c_j)>\varphi(e_j))\right|$, то

$$\left\{a_{j+1}=d_j,b_{j+1}=b_j,c_{j+1}=\frac{a_j+b_j}{2};d_{j+1}=\frac{a_{j+1}+c_{j+1}}{2};e_{j+1}=\frac{c_{j+1}+b_{j+1}}{2};вычс\,\varphi(c_{j+1}),\varphi(d_{j+1}),\varphi(e_{j+1})\right\}$$

Методы одномерного поиска (методы стягивающихся отрезков).

3) Если $\left[(\varphi(d_j) \ge \varphi(c_j)) \wedge (\varphi(c_j) \le \varphi(e_j))\right]$

$$\left\{a_{j+1} = d_j, b_{j+1} = e_j, c_{j+1} = \frac{a_{j+1} + b_{j+1}}{2} = c_j; d_{j+1} = \frac{a_{j+1} + c_{j+1}}{2}; e_{j+1} = \frac{c_{j+1} + b_{j+1}}{2}; вычс\,\varphi(d_{j+1}), \varphi(e_{j+1})\right\}$$

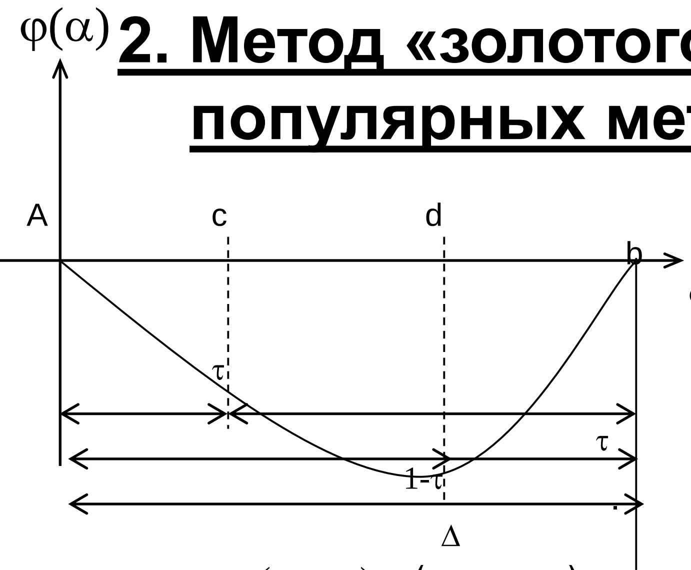
Ш4. Если $\left[\left(\frac{d_{j+1} - e_{j+1}}{2} < \delta\right) \vee \left(\varphi(c_j) - \varphi(c_{j+1})\right) < \varepsilon\right]$, то $\left\{\alpha^* = c_{j+1}, конец\right\}$

иначе $\left\{j = j + 1; переход\,на\,Ш3\right\}$

На каждой итерации требуется 3-х кратное вычисление функции $\varphi(\alpha)$

Методы одномерного поиска (методы стягивающихся отрезков).

2. Метод «золотого сечения» (один из наиболее популярных методов).



$$c_j = a_j + (1-\tau)*(b_j - a_j)$$
$$d_j = a_j + \tau(b_j - a_j)$$

$$\frac{1}{\tau} = \frac{\tau}{1-\tau} \Rightarrow \tau^2 + \tau - 1 = 0 \Rightarrow \tau_{1,2} = \frac{-1 \pm \sqrt{5}}{2}$$

Используем положительный корень τ_1

$$\tau := \tau_1 = \frac{-1+\sqrt{5}}{2} = 0{,}618; \quad 1-\tau = 0{,}382.$$

Методы одномерного поиска (методы стягивающихся отрезков).

2. Метод «золотого сечения»

Алгоритм:

1)Алгоритм обладает сублинейной скоростью сходимости.

2)На каждой итерации требуется 1-кратное вычисление функции $\varphi(\alpha)\cdot$

Ш1. Задание $\delta, \varepsilon, j = 0$. Выбор длины интервала $\Delta_0 = b_0 - a_0$.

$\{$ Вычислить $d_0 = a_0 + \tau(b_0 - a_0)$; $c_0 = a_0 + (1-\tau)(b_0 - a_0)$.

Вычислить $\varphi(a_0), \varphi(b_0), \varphi(c_0), \varphi(d_0.)$ $\}$

Ш2. Если

$$\left[\varphi(c_j) \geq \varphi(d_j)\right], \text{то } \{a_{j+1} = c_{j:}; c_{j+1} = d_j; b_{j+1} = b_j; d_{j+1} = a_{j+1} + \tau(b_{j+1} - a_{j+1}); \text{вычислить } \varphi(d_{o+1})\}$$

Ш3. Если $\left[\left(\left|\frac{c_{j+1} d_{j+1}}{2} \langle \delta\right|\right) \vee \left(\left|\frac{\varphi(c_{j+1} - \varphi(d_{j+1}))}{2}\right| \langle \varepsilon\right)\right]$, то

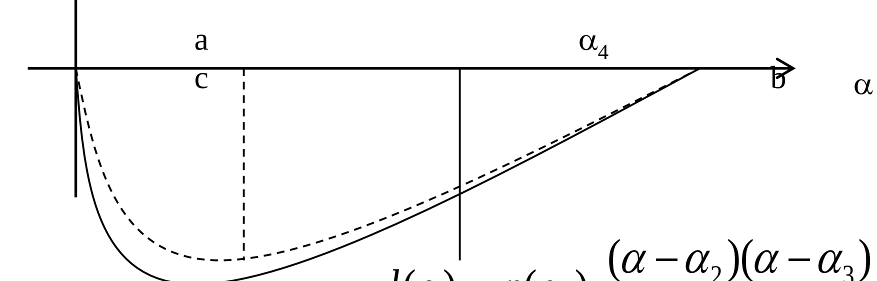
$$\left\{\alpha^* = \frac{c_{j+1} - d_{j+1}}{2}; \text{конец}\right\}, \textit{иначе}\{j = j+1; \text{переход на Ш2.}\}$$

Методы интерполяции.

Обладают большей скоростью сходимости, но используются только для $\varphi(\alpha) \in C^1(\mathrm{A})$.

Метод квадратичной интерполяции.

Пусть имеем 3 точки a,b,c: причем $\varphi(a) \rangle \varphi(c) \langle \varphi(b)$



$$l(\alpha) = \sum_{i=1}^{3} \varphi(\alpha_i) \frac{\prod\limits_{i \neq j}^{3} (\alpha - \alpha_i)}{\prod\limits_{i \neq j}^{3} (\alpha_i - \alpha_j)} \quad т.е.$$

$$l(\alpha) = \varphi(\alpha_1)\frac{(\alpha-\alpha_2)(\alpha-\alpha_3)}{(\alpha_1-\alpha_2)(\alpha_1-\alpha_3)} + \varphi(\alpha_2)\frac{(\alpha-\alpha_1)(\alpha-\alpha_3)}{(\alpha_2-\alpha_1)(\alpha_2-\alpha_3)} + \varphi(\alpha_3)\frac{(\alpha-\alpha_1)(\alpha-\alpha_2)}{(\alpha_3-\alpha_1)(\alpha_3-\alpha_2)}$$

Продифференцировав по α и приравняв 0, получим:

$$\alpha_4 = \frac{1}{2}\frac{\varphi(\alpha_1)(\alpha_2^2-\alpha_3^2)+\varphi(\alpha_2)(\alpha_3^2-\alpha_1^2)+\varphi(\alpha_3)(\alpha_1^2-\alpha_2^2)}{\varphi(\alpha_1)(\alpha_2-\alpha_3)+\varphi(\alpha_2)(\alpha_3-\alpha_1)+\varphi(\alpha_3)(\alpha_1-\alpha_2)}$$ так как $\alpha_4 \in [\alpha_1, \alpha_3]$,

используем процедуру стягивающихся отрезков.

Упрощенные методы.

Эти методы предполагают дифференцируемость функции $\varphi(x)$ в точке x^k,

т.е. $\varphi(x) \in C^1(X)$. Обычно они используются для решения оптимизационных задач методами первого и второго порядков.

1. Правило Армихо.

Можно зафиксировать число $\hat{\alpha} > 0, \varepsilon, \theta \in (0,1).$ (Обычно $\varepsilon = 0.7; \theta = 0.3.$)

Будем полагать, что $\alpha = \hat{\alpha}.$

Алгоритм:

Ш1. Проверяется выполнение неравенства $\varphi(x^k + \alpha d^k) \le \varphi(x^k) + \varepsilon\alpha \langle \varphi'(x^k), d^k \rangle.$

Ш2 . **Если** это неравенство не выполнено, то α заменяется на $\theta\alpha$ и переход на ***Ш1.***

Иначе полагаем $\alpha^k = \alpha$. ***Конец.***

Упрощенные методы.

2. Правило Голдстейна.

Алгоритм:

Ш1. Проверяется неравенство $\varepsilon_1 \le \dfrac{\varphi(x^k + \alpha d) - \varphi(x^k)}{\alpha \left\langle \varphi'(x^k), d \right\rangle} \le \varepsilon_2$

Ш2 . Если это неравенство не выполнено, то α заменяется на $\theta\alpha$ и переход на ***Ш1.*** ***Иначе*** полагаем $\alpha^k = \alpha$. ***Конец.***

При этом обычно $\varepsilon_1 = 0.2; \varepsilon_2 = 0.8$.

Упрощенные методы.

3. Правило Вульфе-Пауэлла.

Алгоритм:

Ш1. Проверяются неравенства

$$\varphi(x^k+\alpha d^k)\le\varphi(x^k)+\varepsilon_1\alpha\left\langle\varphi'(x^k),d\right\rangle,$$

$$\left\langle\varphi'(x^k+\alpha d^k),d^k\right\rangle\ge\varepsilon_2\left\langle\varphi'(x^k),d\right\rangle.$$

Ш2 . Если это неравенство не выполнено, то α заменяется на $\theta\alpha$

и переход на ***Ш1.*** ***Иначе*** полагаем $\alpha^k=\alpha$. ***Конец.***

При этом, как и в предыдущем алгоритме $\varepsilon_1=0.2;\varepsilon_2=0.8$.

Наиболее часто эти методы используются совместно с квазиньютоновскими методами.