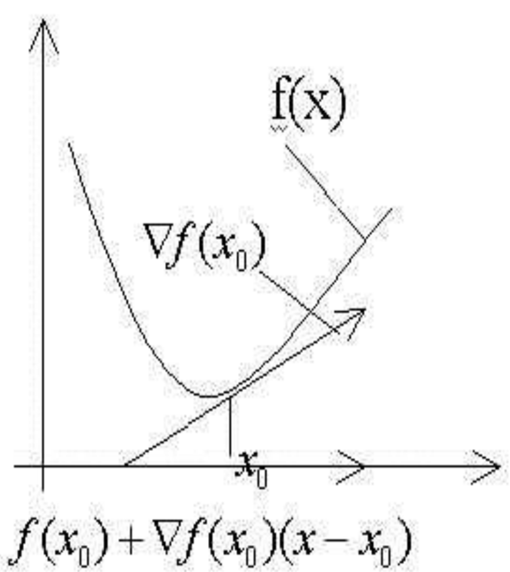
ЛЕКЦИЯ № 4

Математические основы оптимизации

1. Дифференцирование функций.

1.1. Градиент функции.

Пусть f: $X \to Y; X \subset E^n, Y \subset E^1$ f(x) – всюду гладкая функция.



$$gradf(x_0) = \nabla_x f(x_0) = \left\{\frac{df}{dx_i}\right\}_{i=1,n}; \nabla_x f(x) = \left[\frac{df}{dx_1}, \frac{df}{dx_2}, \cdots \frac{df}{dx_n}\right]^T$$

для выпуклой функции:

$$f(x) \geq f(x_0) + \nabla_x^T f(x_0)(x - x_0)$$

фундаментальное неравенство выпуклых функций.

1. Дифференцирование функций.

1.2. Субградиент и его субдифференциал.

Определение 1. Субградиентом выпуклой не всюду гладкой функции f в точке Xo называется вектор $\gamma = [\gamma_1, \gamma_2, \gamma_3 \dots \gamma_n]^T$, удовлетворяющий условию $f(x) \geq f(x_0) + \gamma^T (x - x_0)$. $X \subset R^n, f(x) \in R^1$

В пространстве $E^{n+1} \subset R^{n+1}$ гиперплоскость: $z = f(x_0) + \gamma^T (x - x_0)$ касающаяся функции f(x) в т. Xo , и лежащая всюду не выше этой функции называется опорной гиперплоскостью функции f(x) в точке Xo.

Определение 2. Субдифференциалом выпуклой необязательно гладкой функции f в т.Xo называется множество всех субградиентов функции f(x) в т. Xo

$$df(x_0) = \{\gamma_j\}^k{}_{j=1}\Big|_{x_0}$$

Обобщение: дифференциал Фреше - главная линейная часть функции f(x):

$$Adx = [\gamma_j^T] dx$$

1. Дифференцирование функций.

Теорема: Для любой точки $x_0 (\forall x_0 \in I(S_\alpha))$ элемента внутренности множества сечения уровня $df(x_0)$ - компактное множество.

1.3. Производная функции по направлению.

Определение 3. Пусть f: $X \to Y; X \subset E^n, Y \subset E^1,$ тогда производной по направлению d функции f(x) в т. Xo называется предел:

$$f'_d(x_0) = \lim_{\alpha \to 0} \frac{f(x_0 + \alpha d) - f(x_0)}{\alpha} \geq \gamma^T d\, x\Big|_{x_0}$$

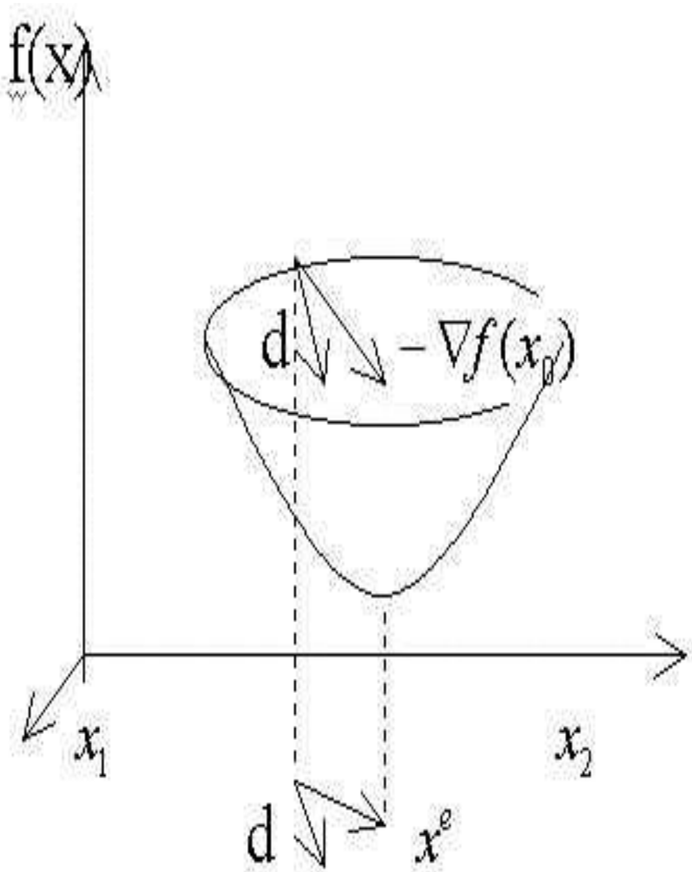
или

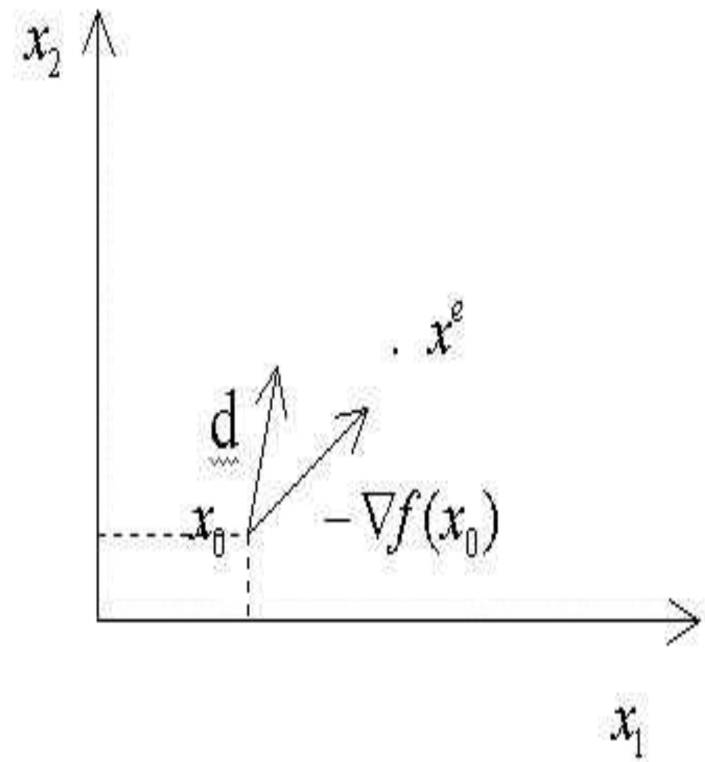
$$\left.\frac{df(x)}{dx}\right|_d = \lim_{\alpha \to 0} \frac{f(x_0 + \alpha d) - f(x_0)}{\alpha} \geq \gamma^T d$$

Обобщение: Производная Гато: $\left.\frac{df(x_0)}{d\alpha}\right|_d = \max \gamma^T d$

Из существования производной Гато не следует существование дифференциала Фреше.

1. Дифференцирование функций.





1. Дифференцирование функций.

1.4. Разложение функции в ряд Тейлора.

В случае n-мерного пространства переменных - разложение в ряд Тейлора.

Пусть f : $X \to Y; X \subset E^n, Y \subset E^1,$ тогда , если f(x) n раз дифференцируема в т. x0, то $f(x_0 + \Delta x) = f(x_0) + \nabla^T f(x_0)\Delta x + \frac{1}{2}\Delta x^T H(x_0)\Delta x + ... + O(\|\Delta x\|^T);$

где O - остаточный член в форме Пеано;

$\nabla^T f(x_0) = gradf(x_0) = \left\{\frac{\partial f(x_0)}{\partial x_i}\right\}_{i=1,n}$ - градиент функции f(x) в т. x0;

$H(x_0) = \left\{\frac{\partial^2 f(x_0)}{\partial x_i \partial x_j}\right\}_{\substack{i=1,n \\ j=1,n}}$ – матрица Гессе 2-х частных производных в т. x0.

Если $\det H(x_0) \ge 0$, функция f(x) в т. x0 выпукла (если > – строго).

Если $\det H(x_0) \le 0$, функция f(x) в т. x_0 вогнута (если < – строго).

Если $\det H(x_0) = 0$, функция f(x) в т. x_0 имеет перегиб.

1. Дифференцирование функций.

1.5. Матрица Якоби. Якобиан.

Пусть g - операторное отображение: g: $X \to Z, X \subset E^n, Z \subset E^m,$

Тогда $grad\ g(x_0) = J(x_0)$ – матрица Якоби, где $J(x_0) = \left[\dfrac{\partial g_j(x_0)'}{\partial x_i} \right]_{\substack{i=1,n \\ j=1,m}}$

а det(J(x0)) – Якобиан. det(J(x0)) = |J(x0)|.

Если $\det J(x) \neq 0,$ то система уравнений gj(x) = C имеет решение.

$\forall x \in X$. В этом случае существует обратное отображение

$g^{-1}(x_0) \to x_0, [m.e. g^{-1} : Z \to X].$

Это имеет большое значение при решении задач с неявно заданными функциями.

1. Дифференцирование функций.

1.6. Классы дифференцируемых функций.

Пусть f: $X \to Y; X \subset E^n, Y \subset E^1,$ тогда

$f(x) \in C^0(X)$ – класс непрерывных функций;

$f(x) \in C^1(X)$ – класс 1 раз дифференцируемых функций;

$f(x) \in C^2(X)$ – класс 2 раза дифференцируемых функций;

……………………………………………………………………….

$f(x) \in C^n(X)$ – класс n раз дифференцируемых функций.

2. Необходимые и достаточные условия существования экстремума функции.

2.1. Условия существования экстремума без учета ограничений.

2.1.1. Необходимые и достаточные условия выпуклости функций.

Теорема1. (Необходимые условия).

Пусть f: $X \to Y; X \subset E^n, Y \subset E^1$ $\|x - \bar{x}\| \leq \varepsilon$

$y = f(x); \forall x, x \in X, y \in Y, f(x) \in C^1(X)$, тогда для того, чтобы f(x)

была выпуклой в Uε(x), необходимо, чтобы

$f(x) \geq f(\bar{x}) + \nabla_x^T f(x) U_\varepsilon(\bar{x})$, где $|x - \bar{x}| \in U_\varepsilon(\bar{x})$.

Теорема2. (достаточные условия).

Пусть f: $X \to Y; X \subset E^n, Y \subset E^1$ и $\exists \varepsilon$, для которой $\|x - \bar{x}\| \leq \varepsilon, U_\varepsilon(\bar{x})$,

$\forall x, \bar{x} \in X, y \in Y, f(x) \in C^2(x)$, тогда для того, чтобы f(x) , была

выпуклой в U(x) достаточно, чтобы $\det H(\bar{x}) > 0$.

2.Необходимые и достаточные условия существования экстремума функции.

2.1.2.Необходимые и достаточные условия существования экстремума функции без ограничений.

Теорема3. Пусть f: $X \to Y; X \subset E^n, Y \subset E^1; \; y = f(x); \forall x \in X, y \in Y, f(x) \in C^2(X).$

Тогда, для того чтобы $x^e = Argextr_{x \in X} f(x)$ необходимо и достаточно, чтобы

были выполнены условия: $\begin{cases} \nabla f(x^e) = 0 \\ \det H(x^e) \neq 0 \end{cases}$

При этом: $\begin{cases} \nabla f(x^e) = 0 \\ \det H(x^e) > 0 \end{cases}$ приводит к $x^e = Arg\min_{x \in X} f(x)$

$\begin{cases} \nabla f(x^e) = 0 \\ \det H(x^e) \leq 0 \end{cases}$ приводит к $x^e = Arg\max_{x \in X} f(x)$

2.Необходимые и достаточные условия существования экстремума функции.

2.2. Условия существования экстремума с учетом ограничений.

2.2.1.Постановка задачи:

Пусть f: $X \to Y; X \subset E^n, Y \subset E^1$ y=f(x) – выпуклая функция;

$x \in X, y \in Y; f(x) \in C^1(X)$ и пусть g: $X \to Z; X \subset E^n, Z \subset E^m$

$f(x) \in C^1(X)$

$g_i(x) \in C^1(X), j = 1..m.$ – вогнутые или выпуклые функции,

$g = [g_1(x), g_2(x), ...g_m(x)]^T$ а также $\forall g_j(x) = 0$ – ограничения типа равенства.

Тогда задачу поиска экстремума функции при наличии ограничений можно записать:

$$\begin{cases} x^e = Arg\min f(x) \\ x_i \geq 0, i = 1..n \\ g_j(x) = 0, j = 1..m \end{cases}$$

2.Необходимые и достаточные условия существования экстремума функции.

2.2.2.Метод множителей Лагранжа. (1788г.”Аналитическая механика”)

Функция вида: $L(x,\lambda) = f(x) + \sum_{j=1}^{m} \lambda_j g_j(x).$

Если взять вариацию от функции L(x,λ) в U(x):

$$\delta L(x,\lambda) = \sum_{i=1}^{n} \frac{\partial f(x)}{\partial x_i}\delta x_i + \sum_{i=1}^{n}\sum_{j=1}^{m} \lambda_j \frac{\partial g_j}{\partial x_i}\delta x_i = 0 \text{ при } x = x^e$$

$$\delta L(x,\lambda) = \sum_{i=1}^{n} \left[\frac{\partial f}{\partial x_i} + \sum_{j=1}^{m} \lambda_j \frac{\partial g_j}{\partial x_i} \right]\Bigg|_{x=x^e} \delta x_i = 0$$

Таким образом получаем систему уравнений : $\left[\frac{\partial f}{\partial x_i} + \sum_{j=1}^{m} \lambda_j \frac{\partial g_j}{\partial x_i} = 0 \right]\Bigg|_{x=x^e}$

Подбором определенных значений λ_j можно добиться того, чтобы эти уравнения удовлетворялись при $\forall i = \overline{1..n}.$ Именно в точке x^e достигается стационарность функции L(x,λ), т.е. экстремум функции f(x) при заданных ограничениях $g_j(x)=0$.

2.Необходимые и достаточные условия существования экстремума функции.

2.2.2.Метод множителей Лагранжа.

Задачу можно переформулировать следующим образом.

Требуется найти: $\begin{cases} x^e = Arg\min_{x\in X} L(x,\lambda) \\ \lambda^e = Arg\max_{\lambda\in\Lambda} L(x,\lambda) \end{cases}$

Тогда решение задачи состоит в совместном решении уравнений:

$$\begin{cases} \dfrac{\partial L(x,\lambda)}{\partial x_i} = 0, i = 1..n; \\ \dfrac{\partial L(x,\lambda)}{\partial \lambda_j} = 0,\ j = 1..m. \end{cases}$$

Составляются и решаются 2 системы уравнений:

$$\begin{cases} \dfrac{\partial f}{\partial x_i} + \sum_{j=1}^{m} \lambda_j \dfrac{\partial g}{\partial x_i} = 0, -n\_\textit{уравнений} \\ g_j(x) = 0, -m\_\textit{уравнений} \end{cases}$$

2.Необходимые и достаточные условия существования экстремума функции.

2.2.4.Область применения классического метода множителей Лагранжа.

Задача поиска экстремума методами классического математического анализа связана с удовлетворением следующих условий, накладываемых на функции:

1.Целевая функция f(x) должна быть выпуклой (в случае поиска минимума).

2.Функции ограничений gj(x) должны быть вогнутыми или выпуклыми функциями.

3.Ограничения gj(x) = 0 только типа равенств.

Последнее условие связано с тем, что если gj(x) < 0, то классический метод множителей Лагранжа не дает единственного решения $x^e = Arg\min f(x)$

и, следовательно, не работает.