1 Основные понятия теории оптимизации

Определение 1 (Окрестность). *Множество $U_\epsilon(X)$ называется окрестностью точки $X \in \mathbb{R}^N$, если существует $\epsilon > 0$ для любых $X' \in U_\epsilon(X)$ справедливо неравенство*

$$0 \leq ||X - X'|| = ||\delta X|| = \sqrt{\sum_{i=1}^{N}(x_i - x_i')^2} \tag{1}$$

$$\delta X = (x_1 - x_1', \ldots, x_n - x_n') \tag{2}$$

Определение 2 (Локалный эстремум). *Точка $X^* \in D$ называется точкой локального экстремума (локального минимума или максимума), если существует $U_\epsilon(X^*)$ такая что в которой выполняется неравенство*

$$f(X^*) \leq f(X) \text{ или } f(X^*) \geq f(X) \tag{3}$$

2 Постановка задачи оптимизации

Пусть даны:

1. функция $f(X)$ с областью определения D_f, $X \in D_f \subseteq \mathbb{R}^N$
2. множество D

Требуется найти $X^* \in D$ в которой функция достигает экстремального значения.

$$X^* : f(X^*) = \text{extr}_{(X \in D_f) \cap (X \in D)} f(X) \tag{4}$$

1. Если $D_f = D = \mathbb{R}^N$ такая задача называется задачей на безусловный экстремум.

2. Если $D_f \cap D \neq \emptyset$ и $D \subset \mathbb{R}^N$ или $D_f \subset \mathbb{R}^N$ то это задача на условный экстремум

3. Если $D_f \cap D = \emptyset$ экстремум не существует

Обычно задается не множесто D ,а система ограничений, его задающая

$$\psi_j(X) \leq, =, \geq 0, j = 1 \dots M \tag{5}$$

$$D = \{X \in \mathbb{R}^N \mid \psi_j(X) \leq, =, \geq 0, j = 1 \dots M\} \tag{6}$$

3 Условия существования безусловного экстремума

1. **Необходимые условия** – условия, которые вытекают из того факта, что рассматри точка есть экстремум

2. **Достаточные условия** – условия, из которых следует, что рассматриваемая точка экстремум

Теорема 1 (Ферма одномерный случай)**.** *Пусть функция f задана на $\mathbb{R}$, и в некоторой точке дифференцируема. Если эта точка имеет в точке x^* локальныйэкстремум, то $f'(x^*) = 0$*

Proof. Раздлжим в ряд Тейлора в форме Пеано

$$f(x^* + \delta x) = f(x^*) + f'(x^*)\delta x + o(\delta x) \tag{7}$$

$$f(x^* - \delta x) = f(x^*) - f'(x^*)\delta x + o(\delta x) \tag{8}$$

$$\begin{cases} f'(x^*)\delta x = f(x^* + \delta x) - f(x^*) \\ -f'(x^*)\delta x = f(x^* - \delta x) - f(x^*) \end{cases} \tag{9}$$

□

Теорема 2 (Необходимое услове экстремума первого порядка)**.** *Пусть функция $f(X)$ задана на $\mathbb{R}^N$, $X \in \mathbb{R}^N$,дифференцирума , $f(X^*) = \operatorname{extr}_{X\in\mathbb{R}^N} f(X)$, то* ***grad*** $f = 0$

Proof. Допусти, что точка X^* точка локального минимума. Для произвольной переменной x_i^* рассмотрим f как $f(x_1^*, x_2^*, \ldots, x_i, \ldots, x_N^*)$ может рассматриваться как в окрестности $U_\epsilon(X^*)$ как функция одной переменной с минимумом в точке x_i^*, тогда $f'_{x_i}(x_i^*) = 0$ по теореме 1 □

Определение 3. *X стационарная точка, если* ***grad*** $f(X) = 0$

Точки локального экстремума нужно искать среди стационарных точек

Пример 1. *Найти стационарные точки функции*

$$f(X) = \sum_{i=1}^{N} a_i x_i^2 \tag{10}$$

$$f'_{x_i} = 2a_i x_i = 0 \tag{11}$$

Такая система уравнений, имеет одно решение $X^* = (0, 0, \ldots, 0)$

1. $\forall a_i \geq 0$ X^**точка минимума*
2. $\forall a_i \leq 0$ X^* *точка максимума*

Пример 2.

$$f(X) = 5x_1^2 - 6x_2^2 \tag{12}$$

$$\begin{cases} f'_{x_1} = 10x_1 = 0 \\ f'_{x_2} = -12x_2 = 0 \end{cases} \tag{13}$$

Стационарная точка $(0, 0)$

$$f(x_1, 0) = 5x^2 \tag{14}$$

Возрастает

$$f(0, x_2) = -6x^2 \tag{15}$$

Убывает.

X^* – седловая точка

Теорема 3 (Необходимое условие экстремума второго порядка)**.** *Пусть функция f задана на $\mathbb{R}$, дважды непрерывно дифференцируема, тогда*

1. $f(x^*) = \min f(x) \implies f''(x^*) \geq 0$

2. $f(x^*) = \min f(x) \implies f''(x^*) \leq 0$

Proof. Разложим функцию в ряд Тейлора с остаточным членом в форме Пеано

$$f(x^* + \delta x) = f(x^*) + f'(x^*)\delta x + \frac{1}{2}f''(x^*)(\delta x)^2 + o((\delta x)^2) \tag{16}$$

Так как x^* точка локального минимума (максимума)

$$f(x^* + \delta x) = f(x^*) + \frac{1}{2}f''(x^*)(\delta x)^2 \tag{17}$$

$$\frac{1}{2}f''(x^*)(\delta x)^2 = f(x^* + \delta x) - f(x^*) \tag{18}$$

1. $f(x^*)\min \implies f(x^* + \delta x) - f(x^*) \geq 0$

2. $f(x^*)\max \implies f(x^* + \delta x) - f(x^*) \leq 0$

□

Определение 4 (матрица Гессе)**.**

$$H(X) = (f''_{x_i x_j})_{ij} \tag{19}$$

Это квадратная симметричная матрица

Определение 5. *Квадратная матрица называется*

1. *Положительно определенной если* $Q(X) = XAX^T > 0$

2. *Положительно полуопределенной если* $Q(X) = XAX^T \geq 0$

3. *Отрицательно определенной если* $Q(X) = XAX^T < 0$

4. *Отрицательно полуопределенной если* $Q(X) = XAX^T \leq 0$ *Для любого* $X = (x_1, \dots, x_N)$

Определение 6 (Угловые (ведущие) миноры)**.** $M_i(A)$ *определители* i *порядка располож вдоль главной диагонали* A *в первых* i *строках и столбцах с одинаковыми ноерами*

Определение 7 (Главные миноры)**.** m_i *матрицы A называют определители* i *порядка ,расположенные вдоль главной диагонали в* i *строках и столбцах, полученных при вычеркивании из* A *строк и столбцов с одними номерами*

Определение 8 (Условия Сильвестра)**.** 1. *Матрица A положительно определенная когда все угловые миноры положительны*

2. *Матрица A является отрицательно определнной, когда* $\operatorname{sgn} M_K(A) = (-1)^k$

3. *Матрица A является положительно полуопределенной, когда она вырожденная и ее главные миноры* ≥ 0

4. *Матрица A является отрицательно полуопределенной, когда значение* $m_k(A)$ *либо равно 0 либо* $\operatorname{sgn}(m_k) = (-1)^k$

Теорема 4 (Необходимое условие экстремума второго порядка)**.** *Пусть функция* f *задана на* $\mathbb{R}^N$ *и дважды непрерывно дифференцируема в некой окрестности* $U_\epsilon(X^*)$ *если в точке* X^* f *имеет локальный минимум (максимум), то Вычисленная в этой точке матрица Гессе, неотрицательно(неположительно) определена*

Proof. Разложим функцию в ряд тейлора в $U_\epsilon(X^*)$

$$f(X^* + \delta X) = f(X^*) + \mathbf{grad}\, f(X)\delta x + \frac{1}{2}\delta X H(X^*)\delta X^T + o(||\delta X||^2) \tag{20}$$

Так как X^* точка локального экстремума

$$\frac{1}{2}\delta X H(X^*)\delta X^T = f(X^* + \delta X) - f(X^*) \tag{21}$$

□

Теорема 5. *Рассмотрим* $f(x)$, x^* *стационарная точка, в окрестности точки существует непрерывная вторая производная*

Если $f''(x) > (<)0$ *то* x^* *точка локального минимума (максимума)*

Proof. Разложим f в $U_\epsilon(x^*)$ в ряд Тейлора

$$f(x^* + \delta x) = f(x^*) + f'(x^*)\delta x + \frac{1}{2}f''(x^*)(\delta x)^2 + o((\delta x)^2) \tag{22}$$

Так как x^* стационарная точка

$$\frac{1}{2}f''(x^*)(\delta x)^2 = f(x^* + \delta x) - f(x^*) \tag{23}$$

□

Теорема 6 (Достаточное условие экстремума)**.** *f задана на $\mathbb{R}^N$ и имеет стационарную точку, **grad** $f(X^*) = 0$, в окрестности которой все вторые производные существуют и непрерывны. Если матрица Гессе $H(X^*)$ положительно (отрицательно) определена, то X^* локального минимума (максимума) функции f. Если $H(X^*)$ не является знако определенной, то экстремума в точке X^* нет*

Proof.

$$f(X^* + \delta X) = f(X^*) + \textbf{grad}\, f(X^*) * \delta X + \frac{1}{2}\delta X H(X^*)\delta X^T + o(||\delta X||^2) \tag{24}$$

$$\frac{1}{2}\delta X H(X^*)\delta X^T = f(X^* + \delta X) - f(X^*) \tag{25}$$

□

Теорема 7 (Частный случай прошлой теоремы для $N = 2$)**.** *$f(x_1, x_2)$ задана на $\mathbb{R}^2$, $X^* = (x_1^*, x_2^*)$ стационарная точка, в которой вторые производные существуют и непрерывны*

1. *Если в точке X^**

$$f''_{x_1} f(X^*) * f''_{x_2}(X^*) - (f''_{x_1 x_2}(X^*)) > 0 \tag{26}$$

то при

(a) $f''_{x_1}(X^*) > 0$, X^* *точка локального минимума*

(b) $f''_{x_2}(X^*) < 0, X^*$ *точка локального максиумам*

2. *Если это не выполняется то локальього экстремума в точке* X^* *нет*

Определение 9. *Для функции* f *определенной на* $\mathbb{R}^N$ *вектор единичной длины задает направление убывания* $w^{\downarrow} \in \mathbb{R}^N$ *или возрастания* $w^{\uparrow} \in \mathbb{R}^N$ *в точке* X' *если при всех достаточно малых* α *выполяняется*

$$f(X' + \alpha w^{\downarrow}) < f(X') \tag{27}$$

$$f(X' + \alpha w^{\uparrow}) > f(X') \tag{28}$$

Определение 10. *Направления убывания (возрастания) функции* f *в точке* X' *образуют множество направлений убывания (возрастания)*

$$W^{\downarrow}_{X'} \subseteq \mathbb{R}^N \tag{29}$$

$$W^{\uparrow}_{X'} \subseteq \mathbb{R}^N \tag{30}$$

Точка X^* является точко минимума функции $f(X)$ если существует $U_\epsilon(X^*)$ целиком содержащаяся в $W^{\uparrow}_{X^*}$

Теорема 8. *Пусть* $f(X)$ *дифференцируема в точке* $X' \in \mathbb{R}^N$

1. *если вектор* $\mathbf{w} \in \mathbb{R}^N$ *удовлетворяет условию*

$$\boldsymbol{grad}\, f(X')\mathbf{w} < (>)0 \tag{31}$$

то $\mathbf{w} \in W^{\downarrow}_{X'}(W^{\uparrow}_{X'})$

2. *если* $\mathbf{w}^{\uparrow(\downarrow)} \in W^{\downarrow(\uparrow)}$, *то*

$$\boldsymbol{grad}\, f(X')\mathbf{w}^{\downarrow(\uparrow)} \leq (\geq)0 \tag{32}$$

Proof. 1.

$$f(X' + \alpha\mathbf{w}) = f(X') + \mathbf{grad}\, f(X')\alpha\mathbf{w} + o(\alpha) \tag{33}$$

$$\mathbf{grad}\, f X' \alpha \mathbf{w} = f(X' + \alpha \mathbf{w}) - f(X') \tag{34}$$

Эквивалентно $\mathbf{w}^{\downarrow(\uparrow)} \in W^{\downarrow(\uparrow)}$

2. Пусть $\mathbf{w}^{\downarrow(\uparrow)} \in W^{\downarrow(\uparrow)}$ то

$$f(X' + \alpha \mathbf{w}^{\downarrow(\uparrow)}) - f(X') = \mathbf{grad}\, f(X') \alpha \mathbf{w}^{\downarrow(\uparrow)} \tag{35}$$

□

4 Методы нелинейной оптимизации

4.1 Классическая задача условной оптимизации

Определение 11 (Формулировка задачи условной оптимизации общем виде)**.**

$$X^* : f(X^*) = \mathrm{extr}_{X \in D}\, f(X) \tag{36}$$

$$D = \{X \in \mathbb{R}^N \mid \psi_j \leq, (=, \geq) 0, j = 1 \dots M, X \in \mathcal{P}\} \tag{37}$$

Определение 12. *Вектор* $\mathbf{v} \in \mathbb{R}^N$ *задает возможное направление в точке* $X \in D$ *на множестве допустимых значений* D *, есл при всех достаточно малых* $\alpha > 0$ *,* $X' = X + \alpha \mathbf{v} \in D$

Все возможные направления образуют в точке x *функции* $f(X)$ *образуют множество возможных направлений* $V_X \subseteq \mathbb{R}^N$

Теорема 9. *Если точка* X^* *локальный минимум (максимум) задачи 36, то*

$$W_{X*}^{\downarrow(\uparrow)} \cap V_{X^*} = \emptyset \tag{38}$$

Proof. Предположим обратное тогда существует $\mathbf{r} \in W^{\downarrow(\uparrow)}$ Такой что при всех достаточно малых α

$$f(X^* + \alpha \mathbf{r}) < (>) f(X^*) \tag{39}$$

$$X^* + \alpha \mathbf{r} \in D \tag{40}$$

Следовательно X^* не является локальным минимумом (максимумом) f □

Теорема 10 (Вейрштрасса)**.** *Пусть $D \subset \mathbb{R}^N$ замкнутое ограниченное множество $f(X)$ непрерывная функция определенная на D. Тогда на D существуют точки глобального минимума и максимума функции f*

Определение 13. *Классическая задача оптимизации – это задача оптимизации, в которой множество допустимых значений D определяется системой функциональных равенств*

$$\psi_j(X) = 0, j = 1 \dots M, M < N \tag{41}$$

$$D = \{X \in \mathbb{R}^N \mid \psi_j(X) = 0, j = 1 \dots M\} \subset \mathbb{R}^n \tag{42}$$

Подразумевается что функции f, ψ_j определены на D

$$X^* : f(X^*) = \text{extr}_{X \in D} f(X) \tag{43}$$

Лемма 1. *Если область допустимых значений содержит точку X' и ее окрестность $U_\epsilon(X')$, то*

$$M < N \tag{44}$$

При этом предполагаем , что $\psi_j(X)$ дифференцируемые функции

Proof. разложим $\psi_j(X)$ в ряд тейлора в окрестности точки X'

$$\psi_j(X' + \delta X) = \psi_j(X') + \mathbf{grad}\, \psi_j(X') \cdot \delta X + o(||\delta X||) \tag{45}$$

$$\psi_j(X') = \psi_j(X' + \delta X) \tag{46}$$

Получили систему M линейных уравнений

$$\mathbf{grad}\, \psi_j \delta X = 0, j = 1 \dots M \tag{47}$$

$$\sum_{i=1}^{N} \frac{\partial \psi_i(X')}{\partial x_i}, j = 1 \dots M \tag{48}$$

1. $M > N$ система является переопределнной либо не содержит решений, либо содердит линейно зависимые ограничения, которые могут быть отброшены

2. $M = N$ система имеет тривиальное решение,либо содержит линейно зависимые ограниения

3. $M < N$ система имеет бесконечно много решений, которые определяют допустимую окрестность

$\square$

Лемма 2. *Предположим, что область допустимых значений D содержит хотя бы одну точку $X' \in D$. Если $M < N$ и якобиан $\mathbf{J} = (\frac{\partial \psi_j(X)}{\partial x_i})$ имеет в этой точке ранг равный M, то $U_\epsilon(X') \subseteq D$*

Proof. Пусть $X' + \delta X \notin D$, тогда

$$\psi_j(X' + \delta X) = \mathbf{grad}\, \psi_j(X') \cdot \delta X + o(||\delta X||) \neq 0 \tag{49}$$

Так как $M < N$ можно выбрать $P = N - M$ независимых переменных, обозначим их как $Z = (z_1, \ldots, z_P)$. Оставшиеся зависиимые $S = (s_1, \ldots, S_M)$, $X = (Z, S)$

$$\Psi(X' + \delta X) = \mathbf{J}_0(X')\delta S^T + \mathbf{C}(X')\delta Z^T \neq 0 \tag{50}$$

$$\delta X = (\delta S, \delta Z) \tag{51}$$

$$\Psi = (\psi_1, \ldots, \psi_M)^T \tag{52}$$

$$\mathbf{J}_0(X') = (\frac{\partial \psi_j(X')}{\partial s_l})_{j=1\ldots M, l=1\ldots M} \tag{53}$$

$$\mathbf{C}(X') = (\frac{\partial \psi_j(X')}{\partial z_k})_{j=1\ldots M, k=1\ldots P} \tag{54}$$

Так как ранг якобиана $\mathbf{J}$ равен M всегда можно сделать разделение что $\mathbf{J}_0$ будет невырождена

$$\delta S^T = -\mathbf{J}^{-1}(X')\mathbf{C}(X')\delta Z^t \tag{55}$$

где $\mathbf{C}(X')$ называется матрицей управления

$$\mathbf{J}_0 \delta S^T + \mathbf{C}(X')\delta Z^T = 0 \tag{56}$$

$$\square(X' + \delta X) = 0 \tag{57}$$

Получили что окрестность $U_\epsilon(X')$ допустимой точки, содержащая отличные точки от X', $X' + \delta X \in U_\epsilon(X') \subseteq D$ $\square$

5 Метод множителей Лагранжа

Метод множителей Лагранжа сводит классическую задачу на условный экстремум к задаче на безусловный экстремум

Определение 14 (Функция Лагранжа)**.**

$$L(\Lambda, \lambda_0, X) = \lambda_0 f(X) + \sum_{j}^{M} \lambda_j \psi_j(X) \tag{58}$$

Данная функция является функцией N переменных X, и $N+1$ параметров λ

$$\frac{\partial L(\Lambda, \lambda_0, X)}{\partial x_i} = \lambda_0 \frac{\partial f(X)}{\partial x_i} + \sum_{j=1}^{M} \lambda_j \frac{\partial \psi_j(X)}{\partial x_i} \tag{59}$$

Составим из частных производных вектор

$$\mathbf{L}'_X = (\lambda_0 \frac{\partial f(X)}{\partial x_i} + \sum_{j=1}^{M} \lambda_j \frac{\partial \psi_j(X)}{\partial x_i})_{i=1\dots N} \tag{60}$$

Частные производные по координатам λ_j

$$\frac{\partial L(\Lambda, \lambda_0, X)}{\partial \lambda_j} = \psi_j(X), j = 1 \dots M \tag{61}$$

Если точка X^* является точкой локального условного экстремума, то

$$\psi_j(X^*) = 0, L(\Lambda, \lambda_0, X^*) = \lambda_0 f(X^*) \tag{62}$$

L имеет безусловный локальный экстремум при $\lambda_0 \neq 0, X^* \in D$

Теорема 11 (Правило множителей Лагранжа)**.** *Пусть в $U_\epsilon(X^*) \subset D$*

1. *f , $\psi_j(X)$ непрерывно дифференцируемы*

2. *ранг матрицы Якоби $\mathbf{J}(X) = (\frac{\partial\psi_j}{\partial x_i})_{i=1\ldots N, j=1\ldots M}$ равен M*

Если $f(X^*) = \text{extr}_{X\in D} f(X)$ то существуют $\lambda_0^* \neq_0, \Lambda^* \neq 0$,что точка $(\lambda^*, 1, X^*)$ является стационарной точкой задачи на езусловный экстрмум функции Лагранжа

$$\textbf{\textit{grad}}\, f(\Lambda^*, 1, X^*) = 0 \tag{63}$$

Proof. Пусть точка X^* точка экстремума задачи на условный. Чтоб точка была стационарной точкой функции загранжа

$$\lambda_0 \frac{\partial f(X)}{\partial x_i} + \sum_{j=1}^{M} \lambda_j \frac{\partial \psi_j(X)}{\partial x_i} = 0, i = 1, \ldots N \tag{64}$$

$$\frac{\partial L(\Lambda, \lambda_0, X)}{\partial \lambda_j} = \psi_j(X) = 0, j = 1, \ldots M \tag{65}$$

$N + M$ уравнений, $N + M + 1$ переменная, если $\lambda_0 \neq 0$, то можно положить его равным $\lambda_0 = 1$, два случая

1. $\lambda_0 = 1$, система 64 линейна относительно Λ, ранг $\mathbf{J}(X^*) = M$, $(X^0, 1, \Lambda)$ стационарная точка

2. $\lambda_0 = 0$, система 64 имеет нулевое решение, строить функцию лагранжа нет смысла

□

$$L(\lambda, X) = f(X) \pm \sum_{j=1}^{M} \lambda_j \psi_j(X) \tag{66}$$

регуляная функция Лагранжа

Нам нужна будет Матрица Гессе для функции Лагранжа

$$\mathbf{H}(\Lambda, X) = \begin{pmatrix} \mathbf{0} & \mathbf{J}(X) \\ \mathbf{J}^T(X) & \mathbf{L}''_{XX}(\Lambda, X) \end{pmatrix} \tag{67}$$

$$\mathbf{L}''_{XX}(\Lambda, X) = (\frac{\partial L(\Lambda, X)}{\partial x_i x_j})_{i=1\ldots N, j=1\ldots N} \tag{68}$$

$$\mathbf{J}(X) = (\frac{\partial \psi_j(X)}{\partial x_i})_{j=1\ldots M, i=1\ldots N} \tag{69}$$

Теорема 12 (Необходимое условие второго порядка)**.** *Пусть в окрестности* $U_\epsilon(X^*) \subset D$ *точки* $X \in \mathbb{R}^N$

1. *функции* f, ψ_j *дважды непрерывно диффеенцируемы*

2. Rank $\mathbf{J}(X^*) = M$

Пусть точка X^* *точка локального минимума (максимума) классической задачи на условный экстремум тогда для любых отклонений* δX *от точки* X^* *удовлетворяющих условию*

$$\boldsymbol{grad}\, \psi_j(X^*) \cdot \delta X = 0 \tag{70}$$

или

$$\mathbf{J}(X^*)\delta X^T = 0 \tag{71}$$

в стационарной точке (Λ^*, X^*) *функции Лагранжа выполняется*

$$\delta X \mathbf{L}''_{XX}(\Lambda^*, X^*)\delta X^T \geq (\leq) 0 \tag{72}$$

Proof. Разложим функцию Лагранжа в ряд Тейлора в окрестности X^* с остатком в форме Пеано

$$L(\Lambda, X) = L(\Lambda^*, X^*) + \mathbf{grad}\, L(\Lambda^*, X^*)(\delta\Lambda, \Delta X) + \frac{1}{2}(\delta\Lambda, \delta X) H(\Lambda^*, X^*) \begin{pmatrix} \delta\Lambda^T \\ \delta X^T \end{pmatrix} + o(||\delta\Lambda, \delta X||^2) \tag{73}$$

$$f(X) = f(X^*) + \frac{1}{2}(\delta\Lambda, \delta X) \begin{pmatrix} \mathbf{0} & \mathbf{J}(X^*) \\ \mathbf{J}^T(X^*) & \mathbf{L}''_{XX}(\Lambda, X) \end{pmatrix} \begin{pmatrix} \delta\Lambda^T \\ \delta X^T \end{pmatrix} \tag{74}$$

$$(\delta\Lambda, \delta X) \begin{pmatrix} \mathbf{0} & \mathbf{J}(X^*) \\ \mathbf{J}^T(X^*) & \mathbf{L}''_{XX}(\Lambda, X) \end{pmatrix} \begin{pmatrix} \delta\Lambda^T \\ \delta X^T \end{pmatrix} = \delta X \mathbf{L}''_{XX}(\Lambda^*, X^*)\delta X^T \tag{75}$$

в силу 71

$$f(X) - f(X^*) = \delta X \mathbf{L}''_{XX}(\Lambda^*, X^*) \geq (\leq) 0 \tag{76}$$

□

Теорема 13 (Достаточное условие экстремума). *Дана классическая задача оптимизаци*

Предположим, что в окрестности допустимой точки X^*

1. $f(X), \psi_j(X)$ *дважды дифферринцируемы*

2. $\mathrm{Rank}\,\mathbf{J}(X^*) = M$

3. *существуют не равные нулю* (Λ^*, λ^*) *такие что* ***grad*** $L(\Lambda^*, X^*) = 0$

4. *существуют не нулевые отклонения* δX *от точки* X^**, удовлетворяющие условиям*

$$\boldsymbol{grad}\, \psi_j(X^*) \cdot \delta X = 0 \tag{77}$$

Для которых выполняется

$$\delta X \mathbf{L}''_{XX}(\Lambda^*, X^*) \delta X^T > (<) 0 \tag{78}$$

То X^* *локального минимума (максимума)* $f(X)$

Теорема 14 (Достаточное условие в терминах матрицы Гессе функции Лагранжа). *Дана классическая задача оптимизации. Допустим в точке* X^**,* f, ψ_j *дважды непрерывно дифференцируемы. Если существуют ненулевые значения множителей Лагранжа для которых* (X^*, Λ^*) *является стационарной точкой регулярной функции Лагранж*

$$L(\Lambda, X) = f(X) + \sum_{j=1}^{M} \lambda_j^M \lambda_j \psi_j \tag{79}$$

То X^* *является:*

1. *точкой минимума* $f(X)$

$$\mathrm{sgn}(M_{2M+k}(\mathbf{H})) = \mathrm{sgn}((-1)^M), k = 1 \ldots N - M \tag{80}$$

2. точкой максимума $f(X)$

$$\text{sgn}(M_{2M+k}(\mathbf{H})) = \text{sgn}((-1)^{M+1}) \tag{81}$$

6 Интерпретация множителей Лагранжа

$$X^* : f(X^*) = \text{extr}_{X \in D} f(X) \tag{82}$$

$$D = \{X \in \mathbb{R}^N \mid \phi_j(X) = b_j\} \subset \mathbb{R}^N \tag{83}$$

$$L(\Lambda, X) = f(X) - \sum_{j=1}^{M} \lambda_j(\phi_j(X) - b_j) \tag{84}$$

В стационарной точке (Λ^*, X^*)

$$\frac{\partial f}{\partial x_i} = \frac{\partial f(X)}{\partial x_i} - \sum_{j=1}^{M} \lambda_j \frac{\partial \phi_j}{\partial x_i}, i = 1 \dots N \tag{85}$$

$$\left.\frac{\partial f}{\partial x_1}\right|_{X^*} = \sum_{j=1}^{M} \lambda_j \left.\frac{\partial \phi_j}{\partial x_1}\right|_{X^*} \tag{86}$$

и так далее

Допустим , что X^* является точкой условного экстремума. Предположим что величины b_j варьируются , точки экстремума,значения функицй f, ψ_j могут рассматриваться как функции параметров b_j

$$x_i^* = x_i(b_1, \dots, b_M), i = 1 \dots N \tag{87}$$

$$f(X^*) = f(X^*(b_1, \dots b_M)) \tag{88}$$

$$\psi_j = \psi_j(X^*(b_1, \dots, b_M) \tag{89}$$

$$\left.\frac{\partial f}{\partial b_k}\right|_{X^*} = \sum_{i=1}^{N} \left.\frac{\partial f}{\partial x_i}\right|_{X^*} \frac{dx_i}{b_k} \tag{90}$$

7 Метод Якоби

$$f(X^*) = \text{extr}_{X \in D} f(X) \tag{91}$$

можно свести к задаче безусловного экстеремума, если Rank $\mathbf{J} = M$

Ограничения разложим в ряд Тейлора в окрестности $U_\epsilon^D(X)$ с остаточным членом в форме Пеано

$$\psi_j(X') = \psi_j(X + \delta X) = \psi_j(X) + \mathbf{grad}\, \psi_j(X) \cdot \delta X + o(|\delta X|), j = 1 \dots M \tag{92}$$

при малых δX

$$\delta\psi_j(X) = \psi_j(X') - \psi_j(X) = \mathbf{grad}\, \psi_j(X) \cdot \delta X, j = 1 \dots M \tag{93}$$

В допустимой окрестности имеем $U_\epsilon^D(X)$ имеем $\psi_j(X') = \psi_j(X) = 0$ поэтому

$$\mathbf{grad}\, \psi_j(X) \cdot \delta X = 0 j = 1 \dots M \tag{94}$$

Полученна система M линейных уравнений, относительно N отклонений δX от точки X. Выберем M зависимых переменных, обозначим их как $\delta S = (\delta s_1 \dots \delta s_M)$, оставшиеся $\delta Z = (\delta z_1, \dots, \delta z_{N-M})$, $\delta X = (\delta S, \delta Z)$

$$\mathbf{grad}\, \psi_j = (\mathbf{grad}_S\, \psi_j, \mathbf{grad}_Z\, \psi_j) \tag{95}$$

где

$$\mathbf{grad}_S\, \psi_j = \left(\frac{\partial \psi_j}{\partial s_i}\right)_{i=1\dots M} \tag{96}$$

$$\mathbf{grad}_Z\, \psi_j = \left(\frac{\partial \psi_j}{\partial z_i}\right)_{i=1\dots N-M} \tag{97}$$

$$\mathbf{grad}\, \psi_j(X) \cdot \delta X = (\mathbf{grad}_S\, \psi_j, \mathbf{grad}_Z\, \psi_j) \cdot (\delta S, \delta Z) = \mathbf{grad}_S\, \psi_j \cdot \delta S + \mathbf{grad}_Z \cdot \delta Z = 0 \quad j = 1 \dots M \tag{98}$$

$$\mathbf{C}(X) = \begin{pmatrix} \mathbf{grad}_Z\,\psi_1 \\ \vdots \\ \mathbf{grad}_Z\,\psi_M \end{pmatrix}_{M\times(N-M)} \tag{99}$$

Матрица управления

$$\mathbf{J}_0(X) = \begin{pmatrix} \mathbf{grad}_S\,\psi_1 \\ \vdots \\ \mathbf{grad}_S\,\psi_M \end{pmatrix}_{M\times M} \tag{100}$$

Так как ранг матрицы **J** равен M , то всегда можно выбрать координаты δS чтобы матрица $\mathbf{J}_0$ была невырожденной

$$\mathbf{J}(0)\delta S^T + \mathbf{C}(X)\delta Z^T = 0 \tag{101}$$

$$\delta S^T = -\mathbf{J}_0^{-1}(X)\mathbf{C}(X)\delta Z^T(X) \tag{102}$$

$$f(X') = f(X+\delta X) = f(X) + \mathbf{grad}\,f(X)\delta X + o(|\delta X|) \tag{103}$$

При малых δX

$$\delta f(X) = f(X+\delta X) - f(X) = \mathbf{grad}\,f(X)\cdot\delta X \tag{104}$$

$$\delta f(X) = \mathbf{grad}_S\,f(X)\delta S + \mathbf{grad}_Z\,f(X)\delta Z \tag{105}$$

$$\delta f(X) = (\mathbf{grad}_Z\,f(X) - \mathbf{grad}_S\,f(X)\mathbf{J}_0^{-1}(X)\mathbf{C}(X))\delta Z \tag{106}$$

переходим к пределу $Z \to 0$

$$\frac{\partial f(X)}{\partial Z} = \mathbf{grad}_*\,f(X) = \mathbf{grad}_Z\,f(X) - \mathbf{grad}_S\,f(X)\mathbf{J}_0^{-1}(X)\mathbf{C}(X) \tag{107}$$

$$\delta f(X) = \mathbf{grad}_*\,f(X)\cdot\delta Z \tag{108}$$

$\mathbf{grad}_*\,f(X)$ приведенный (условный) градиент