Лекция 5

Необходимые и достаточные условия условного экстремума

ПОСТАНОВКА ЗАДАЧИ И ОСНОВНЫЕ ОПРЕДЕЛЕНИЯ

$$f(x^*) = \min_{x \in X} f(x); \quad f(x^*) = \max_{x \in X} f(x), \tag{1}$$

где $X = \left\{ x \,\middle|\, \begin{array}{l} g_j(x) = 0,\ j = 1,\ldots,m;\ m < n \\ g_j(x) \le 0,\ j = m+1,\ldots,p \end{array} \right\}$, m и p – числа; $f(x)$ – целевая функция, $g_j(x),\ j = 1,\ldots,p$, – функции, задающие ограничения (условия).

Будем считать функции $f(x)$; $g_j(x),\ j = 1,\ldots,p$, дважды непрерывно дифференцируемыми на множестве R^n, а функции $g_j(x)$, задающие ограничения, – называть для краткости просто ограничениями. При $p = m$ задача (1) со *смешанными ограничениями* преобразуется в задачу с *ограничениями типа равенств*, а при $m = 0$ в задачу с *ограничениями типа неравенств*.

Определение 1. Функция

$$L\left(x,\lambda_0,\lambda\right)=\lambda_0 f(x)+\sum_{j=1}^{p}\lambda_j g_j(x) \tag{2}$$

называется *обобщенной функцией Лагранжа*, числа $\lambda_0,\lambda_1,\ldots,\lambda_p$ – *множителями Лагранжа*, $\lambda=\left(\lambda_1,\ldots,\lambda_p\right)^T$. *Классической функцией Лагранжа* называется функция

$$L\left(x,\lambda\right)=f(x)+\sum_{j=1}^{p}\lambda_j g_j(x). \tag{3}$$

Определение 2. *Градиентом обобщенной* (*классической*) *функции Лагранжа по x* называется вектор-столбец, составленный из ее частных производных первого порядка по $x_i, i = 1, ..., n$:

$$\nabla_x L(x, \lambda_0, \lambda) = \begin{pmatrix} \dfrac{\partial L(x, \lambda_0, \lambda)}{\partial x_1} \\ \vdots \\ \dfrac{\partial L(x, \lambda_0, \lambda)}{\partial x_n} \end{pmatrix}, \qquad (4)$$

$$\left[\nabla_x L(x, \lambda) = \begin{pmatrix} \dfrac{\partial L(x, \lambda)}{\partial x_1} \\ \vdots \\ \dfrac{\partial L(x, \lambda)}{\partial x_n} \end{pmatrix} \right].$$

Определение 3. *Вторым дифференциалом обобщенной* (*классической*) *функции Лагранжа* $L(x,\lambda_0,\lambda)$ $\left[L(x,\lambda)\right]$ называется функция

$$d^2L(x,\lambda_0,\lambda)=\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{\partial^2 L(x,\lambda_0,\lambda)}{\partial x_i\,\partial x_j}\,dx_i\,dx_j, \qquad (5)$$

$$\left[\ d^2L(x,\lambda)=\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{\partial^2 L(x,\lambda)}{\partial x_i\,\partial x_j}\,dx_i\,dx_j\ \right].$$

Определение 4. *Первым дифференциалом ограничения* $g_j(x)$ называется функция

$$dg_j(x) = \sum_{i=1}^{n} \frac{\partial g_j(x)}{\partial x_i} dx_i, \quad j = 1, \ldots, p. \tag{6}$$

Определение 5. Ограничение $g_j(x) \le 0$ называется *активным* в точке x^*, если $g_j(x^*) = 0$. Если $g_j(x^*) < 0$, то ограничение называется *пассивным.*

Определение 6. Градиенты ограничений $g_1(x),\ldots,g_m(x)$ являются *линейно независимыми* в точке x^*, если равенство $\lambda_1\nabla g_1(x^*)+\lambda_2\nabla g_2(x^*)+\ldots+\lambda_m\nabla g_m(x^*)=0$ выполняется только при $\lambda_1=\lambda_2=\ldots=\lambda_m=0$. Если существуют числа $\lambda_1,\ldots,\lambda_m$, одновременно не равные нулю, для которых равенство выполняется, то градиенты *линейно зависимы*. В этом случае один из них есть линейная комбинация остальных. Один вектор $\nabla g_1(x^*)$ тоже образует систему векторов: при $\nabla g_1(x^*)\neq 0$ линейно независимую, а при $\nabla g_1(x^*)=0$ линейно зависимую.

Система векторов, содержащая нулевой вектор, всегда линейно зависима. Если $\mathrm{rang}A=\mathrm{rang}\left(\nabla g_1(x^*)\nabla g_2(x^*)\ldots\nabla g_m(x^*)\right)=m$, то система векторов линейно независима. Если $\mathrm{rang}A<m$, то система линейно зависима.

УСЛОВНЫЙ ЭКСТРЕМУМ ПРИ ОГРАНИЧЕНИЯХ ТИПА РАВЕНСТВ

Постановка задачи

Даны дважды непрерывно дифференцируемые целевая функция $f(x) = f\left(x_1, \ldots, x_n\right)$ и функции ограничений $g_j(x) = g_j\left(x_1, \ldots, x_n\right) = 0,\ j = 1, \ldots, m,$ определяющие множество допустимых решений X.

Требуется исследовать функцию $f(x)$ на экстремум, т.е. определить точки $x^* \in X$ ее локальных минимумов и максимумов на множестве X:

$$f(x^*) = \min_{x \in X} f(x), \quad f(x^*) = \max_{x \in X} f(x), \tag{7}$$

где $X = \left\{ x \,\middle|\, g_j(x) = 0, \quad j = 1, \ldots, m; \quad m < n \right\}.$

Утверждение 1 (необходимые условия экстремума первого порядка).

Пусть x^* *есть точка локального экстремума в задаче* (7). *Тогда найдутся числа* $\lambda_0^*, \lambda_1^*, \ldots, \lambda_m^*$ *, не равные одновременно нулю и такие, что выполняются следующие условия*:

- *условие стационарности обобщенной функции Лагранжа по x*:

$$\frac{\partial L(x^*, \lambda_0^*, \lambda^*)}{\partial x_i} = 0, \quad i = 1, \ldots, n; \qquad (8\text{ а})$$

- *условие допустимости решения*:

$$g_j(x^*) = 0, \quad j = 1, \ldots, m. \qquad (8\text{ б})$$

Если при этом градиенты $\nabla g_1(x^*), \ldots, \nabla g_m(x^*)$ *в точке* x^* *линейно независимы* (*выполняется условие регулярности*), *то* $\lambda_0^* \neq 0$.

Утверждение 2 (необходимые условия экстремума второго порядка).

Пусть x^* *– регулярная точка минимума* (*максимума*) *в задаче* (7) *и имеется решение* (x^*,λ^*) *системы* (9). *Тогда второй дифференциал классической функции Лагранжа, вычисленный в точке* (x^*,λ^*), *неотрицателен* (*неположителен*):

$$d^2L(x^*,\lambda^*) \ge 0 \quad \left(d^2L(x^*,\lambda^*) \le 0\right) \tag{10}$$

для всех $dx \in R^n$ *таких, что*

$$dg_j(x^*) = \sum_{i=1}^{n} \frac{\partial g_j(x^*)}{\partial x_i}\, dx_i = 0, \qquad j = 1,\ldots,m. \tag{11}$$

Утверждение 3 (достаточные условия экстремума).

Пусть имеется точка (x^*, λ^*)*, удовлетворяющая системе* (9). *Если в этой точке*

$$d^2L(x^*, \lambda^*) > 0 \quad \left(d^2L(x^*, \lambda^*) < 0\right) \tag{12}$$

для всех ненулевых $dx \in R^n$ *таких, что*

$$dg_j(x^*) = \sum_{i=1}^{n} \frac{\partial g_j(x^*)}{\partial x_i} dx_i = 0, \qquad j = 1, \ldots, m,$$

то точка x^* *является точкой локального минимума* (*максимума*) *в задаче* (7).

З а м е ч а н и я.

1. Точки x^*, удовлетворяющие системе при некоторых λ_0^*, λ^*, называются *условно-стационарными.*

2. При решении задач проверка условия регулярности затруднена, так как точка x^* заранее неизвестна. Поэтому, как правило, рассматриваются два случая: $\lambda_0^* = 0$ и $\lambda_0^* \neq 0$. Если $\lambda_0^* \neq 0$, в системе (8 а) полагают $\lambda_0^* = 1$. Это эквивалентно делению системы уравнений (8 а) на λ_0^* и замене $\dfrac{\lambda_j^*}{\lambda_0^*}$ на λ_j^*. При этом обобщенная функция Лагранжа становится классической, а сама система (8) имеет вид

$$\frac{\partial L(x^*,\lambda^*)}{\partial x_i} = 0, \quad i = 1,\dots,n; \tag{9 а}$$

$$g_j(x^*) = 0, \quad j = 1,\dots,m. \tag{9 б}$$

Здесь число уравнений равно числу неизвестных.

Точка экстремума, удовлетворяющая системе (8) при $\lambda_0^* \neq 0$, называется *регулярной*, а при $\lambda_0^* = 0$ – *нерегулярной.* Случай $\lambda_0^* = 0$ отражает вырожденность ограничений. При этом в обобщенной функции Лагранжа исчезает член, содержащий целевую функцию, а в необходимых условиях экстремума не используется информация, представляемая градиентом целевой функции.

З а м е ч а н и я.

1. Достаточные и необходимые условия экстремума второго порядка проверяются в условно-стационарных точках, которые удовлетворяют системе (8) при $\lambda_0^* \neq 0$ или системе (9), так как для практики, безусловно, представляет интерес случай, когда в функции Лагранжа присутствует целевая функция, экстремум которой ищется.

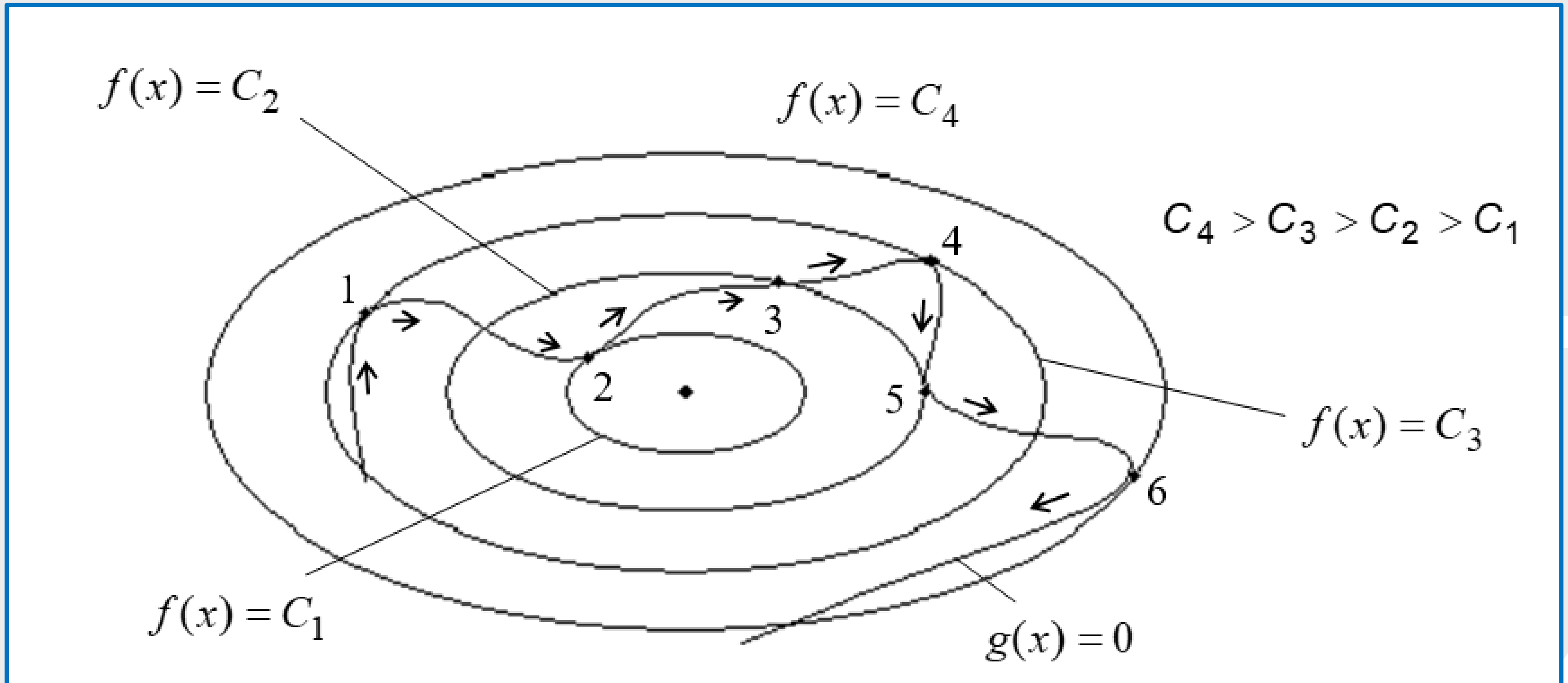
2. Иногда удается проверить условие линейной независимости градиентов ограничений на множестве X (см. определение 6.). Если оно выполняется, то на шаге 1 следует записать классическую функцию Лагранжа (3), на шаге 2 можно записывать сразу систему (9), а на шаге 3 отсутствует случай $\lambda_0^* = 0$.

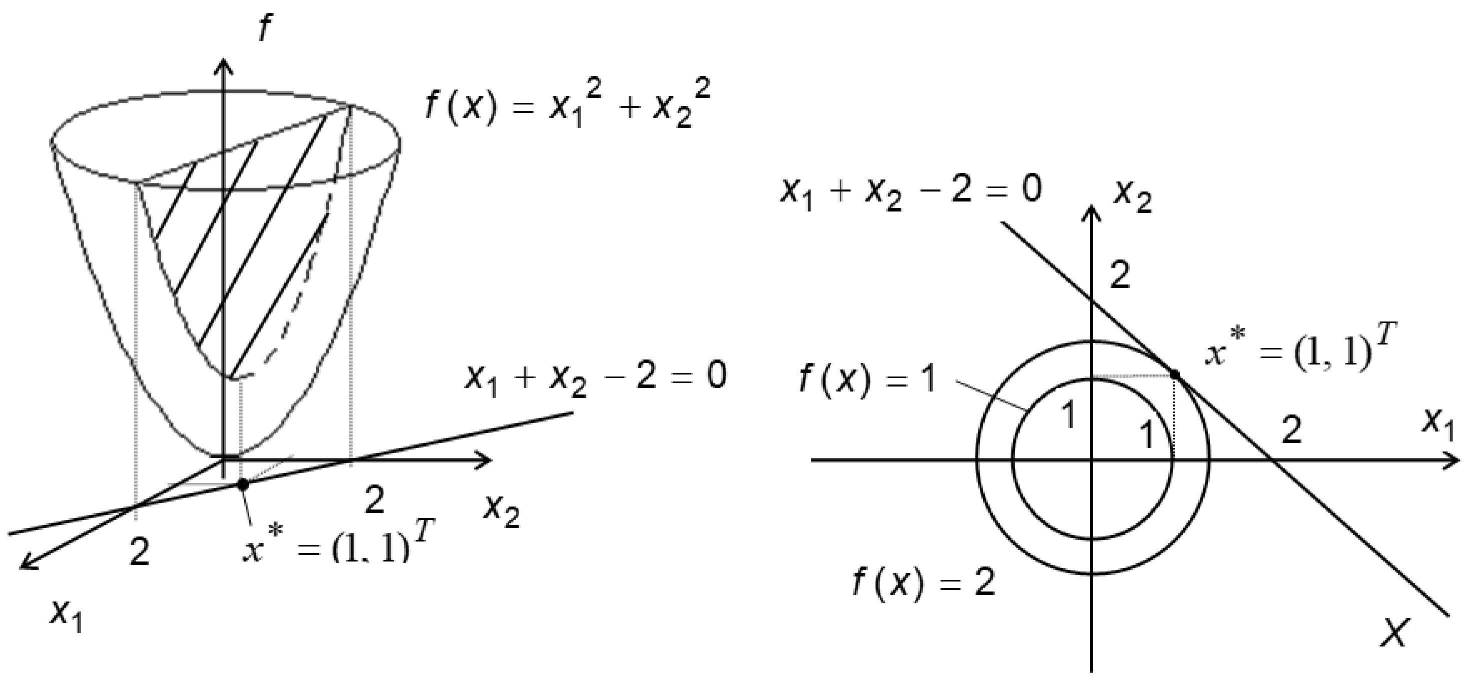
Для нахождения графического решения задачи (при $n=2, m=1$) следует:

а) построить множество допустимых решений X;

б) построить семейство линий уровня целевой функции и найти точки их касания с кривыми, описывающими ограничения. Эти точки являются «подозрительными» на условный экстремум;

в) исследовать поведение целевой функции при движении вдоль ограничения к исследуемой точке и от нее. Классифицировать точки, используя определение экстремума (см. определения 1 и 2).





f
f (x) = x1^2 + x2^2
x1 + x2 − 2 = 0
2
2
x2
x1
x* = (1, 1)^T
x1 + x2 − 2 = 0
x2
2
x* = (1, 1)^T
f (x) = 1
1
1
2
x1
f (x) = 2
X