ЛЕКЦИЯ № 2

Лагранжева теория двойственности

1. Определения

2. Теорема о седловой точке

3. Линейное программирование

4. Теория двойственности линейного программирования

Лагранжева теория двойственности

Рассмотрим задачу P с произвольными функциями f и φ_i:

$$f(x) \longrightarrow \min \tag{1}$$

$$\varphi_i(x) \le 0, i = \overline{1, m}. \tag{2}$$

Определение 1. Функцию

$$L(x, \lambda) = f(x) + \sum_{i=1}^{m} \lambda_i \varphi_i(x),$$

определенную при всех x и λ, назовем функцией Лагранжа для задачи (1), (2).

Лагранжева теория двойственности

Определение 2. **Пара (x^*, λ^*) называется седловой точкой функции Лагранжа, если**

$$L(x^*, \lambda) \overset{(3)}{\leq} L(x^*, \lambda^*) \overset{(4)}{\leq} L(x, \lambda^*) \ \forall x \in R^n, \forall \lambda \geq 0.$$

Пусть

$$g(x) = \sup_{\lambda \geq 0} L(x, \lambda)$$

Тогда

$$g(x) = \begin{cases} f(x), \ x \in Q, \\ +\infty, \ x \notin Q. \end{cases} \Longrightarrow$$

Лагранжева теория двойственности

Задача P эквивалентна следующей:

$$g(x) \longrightarrow \min$$

Пусть

$$h(\lambda) = \inf_{x \in R^n} L(x, \lambda).$$

Рассмотрим задачу (D):

$$h(\lambda) \longrightarrow \max_{\lambda \geq 0}.$$

(D) – задача двойственная к прямой (или исходной) задаче P. $\lambda_1, \ldots, \lambda_m$ – двойственные переменные, а $x_1, \ldots, x_n$ — прямые переменные.

Лагранжева теория двойственности

Если $\boldsymbol{x} \in \boldsymbol{Q}, \boldsymbol{\lambda} \geq \mathbf{0}$, то $\boldsymbol{x}$ — допустимое решение прямой задачи, а $\boldsymbol{\lambda}$ — допустимое решение двойственной задачи.

Лемма 1. (Слабая теорема двойственности).

$$\forall \boldsymbol{x} \in \boldsymbol{Q} \; \forall \boldsymbol{\lambda} \geq \mathbf{0} \; (\boldsymbol{h}(\boldsymbol{\lambda}) \leq \boldsymbol{f}(\boldsymbol{x})).$$

Лемма 2. Если $\overline{\boldsymbol{x}} \in \boldsymbol{Q}$ и $\overline{\boldsymbol{\lambda}} \geq \mathbf{0}$ и $\boldsymbol{f}(\overline{\boldsymbol{x}}) = \boldsymbol{h}(\overline{\boldsymbol{\lambda}})$, то $\overline{\boldsymbol{x}}$ и $\overline{\boldsymbol{\lambda}}$ — оптимальные решения задачи $\boldsymbol{P}$ и $\boldsymbol{D}$, соответственно.

Лагранжева теория двойственности

Теорема 1. Вектора $\overline{\boldsymbol{x}}, \overline{\boldsymbol{\lambda}}$ — оптимальные решения прямой и двойственной задачи и $\boldsymbol{f}(\overline{\boldsymbol{x}}) = \boldsymbol{h}(\overline{\boldsymbol{\lambda}})$ тогда и только тогда, когда пара $(\overline{\boldsymbol{x}}, \overline{\boldsymbol{\lambda}})$ — седловая точка функции Лагранжа. При этом $\boldsymbol{L}(\overline{\boldsymbol{x}}, \overline{\boldsymbol{\lambda}}) = \boldsymbol{f}(\overline{\boldsymbol{x}}) = \boldsymbol{h}(\overline{\boldsymbol{\lambda}})$.

Следствие 1. Пусть $\overline{\boldsymbol{x}} \in \boldsymbol{Q}(\boldsymbol{P}), \overline{\boldsymbol{\lambda}} \geq 0$. Следующие утверждения эквивалентны:

1. Пара $(\overline{\boldsymbol{x}}, \overline{\boldsymbol{\lambda}})$ — седловая точка функции Лагранжа.

2. $\boldsymbol{f}(\overline{\boldsymbol{x}}) = \boldsymbol{h}(\overline{\boldsymbol{\lambda}})$.

3. $\min\limits_{\boldsymbol{x}} \sup\limits_{\boldsymbol{\lambda} \geq 0} \boldsymbol{L}(\boldsymbol{x}, \boldsymbol{\lambda}) = \max\limits_{\boldsymbol{\lambda} \geq 0} \inf\limits_{\boldsymbol{x}} \boldsymbol{L}(\boldsymbol{x}, \boldsymbol{\lambda})$.

Лагранжева теория двойственности

Следствие 2. **Пусть $x^*, \overline{x} \in Q$, $\lambda^*, \overline{\lambda} \geq 0$.**

Если пары $(\overline{x}, \overline{\lambda})$ и (x^*, λ^*) — седловые точки функции Лагранжа, то пары $(\overline{x}, \lambda^*)$ и $(x^*, \overline{\lambda})$ — также седловые точки функции Лагранжа, причем

$$L(x^*, \overline{\lambda}) = L(\overline{x}, \lambda^*) = L(\overline{x}, \overline{\lambda}) = L(x^*, \lambda^*).$$

Линейное программирование (ЛП)

Задача линейного программирования (ЛП) в канонической форме:

$$\boldsymbol{w(x) = (c, x) \longrightarrow \min} \tag{5}$$

$$\boldsymbol{Ax = b,} \tag{6}$$

$$\boldsymbol{x \geq 0,} \tag{7}$$

где $\boldsymbol{c = (c_j), x = (x_j) \in R^n}$, $\boldsymbol{A = (a_{ij}) - (m \times n)}$ матрица, $\boldsymbol{b = (b_i) \in R^m}$, $\boldsymbol{m \leq n}$, $\boldsymbol{rang(A) = m}$.

$$\boldsymbol{Ax = b \equiv (a_i, x) = b_i, i = \overline{1, m},}$$

$$\boldsymbol{Ax = b \equiv \sum_{j=1}^{n} A_j x_j = b}$$

ЛП: двойственная задача

Задача $(5)-(7)$ эквивалентна задаче

$$(c, x) \longrightarrow \min$$

$$\begin{aligned} (a_i, x) - b_i &\le 0, & \lambda_i^1 \ge 0 \\ -(a_i, x) + b_i &\le 0, & \lambda_i^2 \ge 0 \\ -x_j &\le 0. & \mu_j \ge 0 \end{aligned}$$

Ее функция Лагранжа:

$$L(x, \lambda^1, \lambda^2, \mu) = (c, x) + (\lambda^1, Ax - b) + (\lambda^2, -Ax + b) + (\mu, -x) =$$

$$= \Big(c + (\lambda^1 - \lambda^2)A - \mu, x\Big) - (\lambda^1 - \lambda^2, b).$$

Следовательно, целевая функция двойственной задачи имеет вид:

ЛП: двойственная задача

$$h(\boldsymbol{\lambda}^1, \boldsymbol{\lambda}^2, \boldsymbol{\mu}) = \inf_{\boldsymbol{x}} L(\boldsymbol{x}, \boldsymbol{\lambda}^1, \boldsymbol{\lambda}^2, \boldsymbol{\mu}) =$$

$$= \begin{cases} -(\boldsymbol{b}, \boldsymbol{\lambda}^1 - \boldsymbol{\lambda}^2), & \text{если } \boldsymbol{c} + (\boldsymbol{\lambda}^1 - \boldsymbol{\lambda}^2)\boldsymbol{A} - \boldsymbol{\mu} = \boldsymbol{0}, \\ -\infty, & \text{иначе.} \end{cases} \Longrightarrow$$

Двойственная задача

$$h(\boldsymbol{\lambda}^1, \boldsymbol{\lambda}^2, \boldsymbol{\mu}) \longrightarrow \sup_{\boldsymbol{\lambda}^1 \geq 0, \boldsymbol{\lambda}^2 \geq 0, \boldsymbol{\mu} \geq 0},$$

эквивалентна задаче ЛП:

$$-(\boldsymbol{b}, \boldsymbol{\lambda}^1 - \boldsymbol{\lambda}^2) \longrightarrow \max$$

$$\boldsymbol{c} + (\boldsymbol{\lambda}^1 - \boldsymbol{\lambda}^2)\boldsymbol{A} - \boldsymbol{\mu} = \boldsymbol{0} \equiv \boldsymbol{c} + (\boldsymbol{\lambda}^1 - \boldsymbol{\lambda}^2)\boldsymbol{A} \geq \boldsymbol{0}.$$

Умножим ограничения на $-\boldsymbol{1}$, обозначим $\boldsymbol{y} = -(\boldsymbol{\lambda}^1 - \boldsymbol{\lambda}^2)$

ЛП: двойственная задача

Получим

$$(b, y) \longrightarrow \max \tag{8}$$

$$yA \le c. \tag{9}$$

Замечание 2. Для задач (5)-(7) и (8)-(9) выполняются все утверждения: л. 1, л. 2, теор. 1, следствия 1 — 2.

Теорема 2. Задача двойственная к задаче (8)-(9) совпадает с исходной задачей (5)-(7).

Доказательство. Задача (8)-(9) эквивалентна задаче

$$-(b, y) \longrightarrow \min$$

$$yA \le c.$$

ЛП: двойственная задача

$$yA \le c \equiv \text{системе неравенств } (y, A_j) - c_j \le 0, \qquad x_j \ge 0$$

(сопоставили каждому ограничению двойственную переменную (множитель Лагранжа))

Функция Лагранжа:

$$L(y, x) = -(b, y) + (x, yA - c) =$$

$$-(b, y) + (Ax, y) - (x, c) = (Ax - b, y) - (c, x).$$

Целевая функция двойственной задачи

$$h(x) = \inf_y L(y, x) =$$

$$\begin{cases} -(c, x), & \text{если } Ax = b, \\ -\infty, & \text{иначе.} \end{cases} \Longrightarrow$$

ЛП: двойственная задача

Задача

$$\max_{x \geq 0} h(x)$$

эквивалентна задаче ЛП:

$$-(c, x) \longrightarrow \max$$

$$Ax = b,$$

$$x \geq 0$$

или

$$\min(c, x)$$

$$Ax = b,$$

$$x \geq 0. \qquad \blacksquare$$

ЛП: двойственная задача

Прямая задача		Двойственная задача
$w(x)=\sum_{j=1}^{n} c_j x_j \to \min$		$z(y)=\sum_{i=1}^{m} b_i y_i \to \max$
$a_i x \geq b_i$	$i \in I_1$	$y_i \geq 0$
$a_i x = b_i$	$i \in I_2$	y_i — своб.
$x_j \geq 0$	$j \in J_1$	$y A_j \leq c_j$
x_j — своб.	$j \in J_2$	$y A_j = c_j$.

Упражнение.

Техника получения этой схемы: либо повторить выкладки, приведшие к задаче (8), (9), либо воспользоваться сводимостью общей задачи ЛП к задаче ЛП в канонической форме и применить готовый рецепт (задача (8), (9)).

ЛП: понятие базисного допустимого решения (б.д.р.).

Базис — любой набор $A_{\sigma(1)}, \ldots, A_{\sigma(m)}$ из m линейно независимых столбцов матрицы системы ограничений A.

Матрица $B = [A_{\sigma(1)}, \ldots, A_{\sigma(m)}]$ называется базисной.

Обозначения:

$$S = \{\sigma(1), \ldots, \sigma(m)\}, S' = \{1, \ldots, n\} \setminus S, A = [B, N],$$

где $N = [A_j]_{j \in S'}$, $x = (x_B, x_N)$, $x_B = (x_{\sigma(1)}, \ldots, x_{\sigma(m)})$ — базисные, а $x_N = (x_j)_{j \in S'}$ — небазисные переменные.

$$E\, x_B + B^{-1} N x_N = B^{-1} b \qquad (6')$$

Определение 3. **Решение $(x_B, x_N) = (B^{-1} b, 0)$ системы уравнений (6) назовем базисным (соответствующим базису B).**

ЛП: понятие б.д.р.

Лемма 3. **Вектор $\boldsymbol{x}$ — базисное решение системы (6) тогда и только тогда, когда множество столбцов с индексами из множества $\boldsymbol{S(x) = \{j | x_j \neq 0\}}$ — линейно независимо.**

Определение 3. **Базисным допустимым решением (б.д.р.)** называется любой элемент множества $\boldsymbol{Q}$, являющийся базисным решением системы уравнений $\mathbf{(6)}$.

Замечание 3. Решение соответствующее базису $\boldsymbol{B}$ — б.д.р. $\Longleftrightarrow$ $\boldsymbol{B^{-1}b \geq 0}$.

ЛП: понятие б.д.р.

Определение 4. Вектор $x \in Q$ — **крайняя точка или вершина** множества Q, если не существует допустимых решений $x^1 \neq x^2$ из Q таких, что $x = \alpha x^1 + (1-\alpha)x^2$, где $0 < \alpha < 1$.

Теорема 3. **Вектор x — б.д.р. тогда и только тогда, когда x — крайняя точка множества Q.**

Доказательство. Пусть x — крайняя точка, но не б.д.р. Из леммы 3 $\Longrightarrow$

$$\exists\, y \neq 0 : \; Ay = 0. \qquad (*)$$

ЛП: понятие б.д.р.

При этом можем считать, что $(x_j = 0 \Longrightarrow y_j = 0) \Longrightarrow$

$$\{j | y_j \neq 0\} \subseteq \{j | x_j > 0\}. \qquad (**)$$

Т.к. $x \in Q$, то из $(*) \Longrightarrow$

$$\forall\, t \in R : z(t) = x + t\, y - \text{решение (6)},$$

тогда из $(**) \Longrightarrow$

$$\forall \text{ малого } t \in R : \ z(t) \in Q.$$

$\Longrightarrow$

$$\exists\, \varepsilon > 0 : \ x^1 = x + \varepsilon y \in Q, x^2 = x - \varepsilon y \in Q,$$

$\Longrightarrow$

$$x^1 \neq x^2 \text{ и } x = \frac{1}{2}x^1 + \frac{1}{2}x^2 \Rightarrow\rightarrow\leftarrow\Rightarrow x - \text{б.д.р.}.$$

ЛП: понятие б.д.р.

Пусть $\boldsymbol{x}$ — б.д.р., но

$$\exists\, \mathbf{0} < \boldsymbol{\alpha} < \mathbf{1}, \exists\, \boldsymbol{x}^1 \neq \boldsymbol{x}^2, \boldsymbol{x}^1, \boldsymbol{x}^2 \in \boldsymbol{Q} : \boldsymbol{x} = \boldsymbol{\alpha x}^1 + (\mathbf{1} - \boldsymbol{\alpha})\boldsymbol{x}^2.$$

По условию $\boldsymbol{Ax}^1 = \boldsymbol{Ax}^2(= \boldsymbol{b}) \Longrightarrow \boldsymbol{A}(\boldsymbol{x}^1 - \boldsymbol{x}^2) = \mathbf{0}$, но

$\boldsymbol{A}(\boldsymbol{x}^1 - \boldsymbol{x}^2) = \mathbf{0} \Longleftrightarrow \{\boldsymbol{A}_j | \boldsymbol{x}_j^1 \neq \boldsymbol{x}_j^2\}$ — линейно зависимо $\Rightarrow$

$\{\boldsymbol{A}_j | \boldsymbol{\alpha x}_j^1 + (\mathbf{1} - \boldsymbol{\alpha})\boldsymbol{x}_j^2 > \mathbf{0}\}$ — линейно зависимо $\Rightarrow$

$\{\boldsymbol{A}_j | \boldsymbol{x}_j > \mathbf{0}\}$ — линейно зависимо $\Rightarrow \rightarrow\leftarrow$,

т.е. если x — б.д.р., то $\boldsymbol{x}$ — крайняя точка множества Q. ■

ЛП: Критерий разрешимости

Теорема 4 (Критерий разрешимости). **Задача линейного программирования (5)-(7) разрешима тогда и только тогда, когда множество допустимых решений не пусто и целевая функция ограничена на нем.**

Доказательство. Необходимость очевидна. Для доказательства достаточн покажем, что

$$\forall\, \boldsymbol{x}^0 \in \boldsymbol{Q}\ \exists\ \text{б.д.р. } \overline{\boldsymbol{x}} : \boldsymbol{w}(\overline{\boldsymbol{x}}) \leq \boldsymbol{w}(\boldsymbol{x}^0).$$

Пусть

$$\overline{\boldsymbol{x}} \in \boldsymbol{Q}^0 = \{\boldsymbol{x} \in \boldsymbol{Q} | \boldsymbol{w}(\boldsymbol{x}) \leq \boldsymbol{w}(\boldsymbol{x}^0)\} \neq \emptyset$$

и имеет минимальное число ненулевых компонент $(\boldsymbol{supp}(\overline{\boldsymbol{x}}))$.

Доказем, что $\overline{\boldsymbol{x}}$ — б.д.р. Допустим противное. $\Longrightarrow$

ЛП: Критерий разрешимости

множество

$$\{A_j | \overline{x}_j > 0\} \text{ линейно зависимо } \Longrightarrow$$

$$\exists\, y \neq 0 : Ay = 0 \text{ и если } \overline{x}_j = 0, \text{ то } y_j = 0.$$

Пусть $w(y) \leq 0$ (если необходимо, то возьмем $-y$). Положим $x(t) = \overline{x} + ty$. Выполняется следующее свойство

$$\forall \text{ малого } t \in R : \ x(t) \in Q.$$

1). Пусть $\forall j : y_j \geq 0 \Rightarrow \forall t \geq 0\ x(t) \geq 0 \Rightarrow \forall t \geq 0\ x(t) \in Q$

отсюда и неравенства $w(x(t)) = w(\overline{x}) + tw(y) \geq const$(по условию)

учитывая, что $w(y) \leq 0$ и $t \geq 0$ — произвольно,

имеем: $w(y) = 0$ и, следовательно, $w(x(t)) = w(\overline{x})\ \forall t.$

ЛП: Критерий разрешимости

Т.к. из условия $\boldsymbol{y_j > 0 \Rightarrow \overline{x}_j > 0}$, то

$$\forall \text{ малого по абсолютной величине } \boldsymbol{t < 0 : \; x(t) \in Q.}$$

Найдем такое $\boldsymbol{\bar{t}}$ наибольшее по абсолютной величине из условия:

$$\boldsymbol{\forall j (y_j > 0 \Rightarrow \overline{x}_j + \bar{t} y_j \geq 0)}$$

которое эквивалентно условию:

$$\boldsymbol{\forall j (y_j > 0 \Rightarrow \overline{x}_j \geq (-\bar{t}) y_j) \iff}$$

$$\boldsymbol{(-\bar{t}) = \min_{y_j > 0} \frac{\overline{x}_j}{y_j} \iff \bar{t} = - \min_{y_j > 0} \frac{\overline{x}_j}{y_j}.}$$

Итак $\boldsymbol{x(\bar{t}) \in Q}$ и $\boldsymbol{w(x(\bar{t})) = w(\overline{x}) \leq w(x^0) \Rightarrow x(\bar{t}) \in Q^0}$.

Получили противоречие. Т.к. $\boldsymbol{supp(x(\bar{t})) = supp(\overline{x}) - 1}$.

ЛП: Критерий разрешимости

2). Пусть $\exists \boldsymbol{j} : \boldsymbol{y_j} < \mathbf{0}$. Тогда

$$\forall \text{ достаточно малых } \boldsymbol{t} \geq \mathbf{0} : \ \boldsymbol{x}(\boldsymbol{t}) \in \boldsymbol{Q}.$$

Найдем наибольшее такое $\bar{\boldsymbol{t}}$ из условия:

$$\forall \boldsymbol{j}(\boldsymbol{y_j} < \mathbf{0} \Rightarrow \overline{\boldsymbol{x}}_{\boldsymbol{j}} + \bar{\boldsymbol{t}}\boldsymbol{y_j} \geq \mathbf{0})$$

которое эквивалентно условию:

$$\forall \boldsymbol{j}(\boldsymbol{y_j} < \mathbf{0} \Rightarrow \overline{\boldsymbol{x}}_{\boldsymbol{j}} \geq \bar{\boldsymbol{t}}(-\boldsymbol{y_j})) \iff$$

$$\iff \bar{\boldsymbol{t}} = \min_{\boldsymbol{y_j} < 0} \frac{\overline{\boldsymbol{x}}_{\boldsymbol{j}}}{-\boldsymbol{y_j}}.$$

Итак $\boldsymbol{x}(\bar{\boldsymbol{t}}) \in \boldsymbol{Q}$ и т.к. $\bar{\boldsymbol{t}} > \mathbf{0}, \boldsymbol{d} \leq \mathbf{0}$, то $\boldsymbol{w}(\boldsymbol{x}(\bar{\boldsymbol{t}})) = \boldsymbol{w}(\overline{\boldsymbol{x}}) + \boldsymbol{d}\bar{\boldsymbol{t}} \leq$ $\leq \boldsymbol{w}(\boldsymbol{x}^{\mathbf{0}}) \Rightarrow \boldsymbol{x}(\bar{\boldsymbol{t}}) \in \boldsymbol{Q}^{\mathbf{0}}$.

Получили противоречие. Т.к. $\boldsymbol{supp}(\boldsymbol{x}(\bar{\boldsymbol{t}})) = \boldsymbol{supp}(\overline{\boldsymbol{x}}) - \mathbf{1}$.

ЛП: Критерий разрешимости

Т.к. по условию $\boldsymbol{Q} \neq \varnothing$, то множество базисных допустимых решений задачи не пусто. Т.к. оно конечно, то

$$\exists\ \boldsymbol{x}^* \text{ — б.д.р.: } \boldsymbol{w}(\boldsymbol{x}^*) \leq \boldsymbol{w}(\boldsymbol{x})\ \forall \text{ б.д.р. } \boldsymbol{x}.$$

Из ранее доказанного следует, что $\boldsymbol{x}^*$ — оптимальное решение. ■

Следствие 3. Если множество допустимых решений задачи ЛП не пусто, то существуют базисные допустимые решения.

Следует из доказательства теоремы 4 (взять $\boldsymbol{w}(\boldsymbol{x}) \equiv \boldsymbol{0}$).

Следствие 4. Если задача ЛП разрешима, то существует оптимальное базисное решение.