ЛЕКЦИЯ № 2

1. Лагранжева теория двойственности

Линейное программирование

1. Базисные допустимые решения

2. Критерий разрешимости

ЛАГРАНЖЕВА ТЕОРИЯ ДВОЙСТВЕННОСТИ

Рассмотрим задачу $\boldsymbol{P}$ с произвольными функциями $\boldsymbol{f}$ и $\boldsymbol{\varphi_i}$:

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

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

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

Пусть

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

Тогда

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

Лагранжева теория двойственности

Задача $\boldsymbol{P}$ эквивалентна следующей:

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

Пусть

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

Рассмотрим задачу $(\boldsymbol{D})$:

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

$(\boldsymbol{D})$ – задача двойственная к прямой (или исходной) задаче $\boldsymbol{P}$. $\lambda_1, \ldots, \lambda_m$ – 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).

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

Док-во: $\boldsymbol{h}(\boldsymbol{\lambda}) = \inf_{\tilde{\boldsymbol{x}} \in \boldsymbol{R}^n} \boldsymbol{L}(\tilde{\boldsymbol{x}}, \boldsymbol{\lambda}) \leq \boldsymbol{L}(\boldsymbol{x}, \boldsymbol{\lambda}) \leq \sup_{\tilde{\lambda} \geq 0} \boldsymbol{L}(\boldsymbol{x}, \tilde{\boldsymbol{\lambda}}) = \boldsymbol{f}(\boldsymbol{x})$.

■

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

Определение 2. Пара $(\boldsymbol{x}^*, \boldsymbol{\lambda}^*)$, $\boldsymbol{\lambda}^* \geq 0$, называется седловой точкой функции Лагранжа, если

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

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

Док-во:

Лагранжева теория двойственности

1) НЕОБХОДИМОСТЬ.

Пусть $\overline{\boldsymbol{x}}, \overline{\boldsymbol{\lambda}}$ – оптимальные решения прямой и двойственной задачи. Тогда:

$$\boldsymbol{f}(\overline{\boldsymbol{x}}) = \sup_{\boldsymbol{\lambda} \geq \mathbf{0}} \boldsymbol{L}(\overline{\boldsymbol{x}}, \boldsymbol{\lambda}) \geq \boldsymbol{L}(\overline{\boldsymbol{x}}, \overline{\boldsymbol{\lambda}}) \geq \inf_{\boldsymbol{x} \in \boldsymbol{R}^n} \boldsymbol{L}(\boldsymbol{x}, \overline{\boldsymbol{\lambda}}) = \boldsymbol{h}(\overline{\boldsymbol{\lambda}}).$$

Но $\boldsymbol{f}(\overline{\boldsymbol{x}}) = \boldsymbol{h}(\overline{\boldsymbol{\lambda}})$. Возьмем произвольные $\boldsymbol{x} \in \boldsymbol{R}^n$, $\boldsymbol{\lambda} \geq \mathbf{0}$. Получается, что

$$\boldsymbol{L}(\overline{\boldsymbol{x}}, \boldsymbol{\lambda}) \leq \sup_{\boldsymbol{\lambda} \geq \mathbf{0}} \boldsymbol{L}(\overline{\boldsymbol{x}}, \boldsymbol{\lambda}) = \boldsymbol{L}(\overline{\boldsymbol{x}}, \overline{\boldsymbol{\lambda}}) = \inf_{\boldsymbol{x} \in \boldsymbol{R}^n} \boldsymbol{L}(\boldsymbol{x}, \overline{\boldsymbol{\lambda}}) \leq \boldsymbol{L}(\boldsymbol{x}, \overline{\boldsymbol{\lambda}}).$$

Из последнего следуют (3) и (4).

Лагранжева теория двойственности

2) ДОСТАТОЧНОСТЬ.

Пусть $(\overline{x}, \overline{\lambda})$ – седловая точка функции Лагранжа. Тогда из (3) следует:

$$\sum_{i=1}^{m}(\lambda_i - \overline{\lambda}_i)\varphi_i(\overline{x}) \leq 0, \forall \lambda \geq 0.$$

Предположим, что $\overline{x} \notin Q$. Т.е. $\exists i : \varphi_i(\overline{x}) > 0$. Тогда для достаточно большого $\lambda_i > 0$:

$$\sum_{i=1}^{m}(\lambda_i - \overline{\lambda}_i)\varphi_i(\overline{x}) > 0.$$

Противоречие. Следовательно, $\overline{x} \in Q$.

Лагранжева теория двойственности

При $\boldsymbol{\lambda} = \mathbf{0}$ имеем:

$$\sum_{i=1}^{m} \overline{\boldsymbol{\lambda}}_i \boldsymbol{\varphi}_i(\overline{\boldsymbol{x}}) \geq \mathbf{0},$$

откуда:

$$\overline{\boldsymbol{\lambda}}_i \boldsymbol{\varphi}_i(\overline{\boldsymbol{x}}) = \mathbf{0}, \forall \boldsymbol{i} = \overline{\mathbf{1}, \boldsymbol{m}}.$$

Следовательно, $\boldsymbol{f}(\overline{\boldsymbol{x}}) = \boldsymbol{L}(\overline{\boldsymbol{x}}, \overline{\boldsymbol{\lambda}})$. Из (4) следует, что:

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

По Лемме 2, $\overline{\boldsymbol{x}}$ и $\overline{\boldsymbol{\lambda}}$ – оптимальные решения прямой и двойственной задачи, соответственно. ■

Лагранжева теория двойственности

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

1. Пара $(\overline{x}, \overline{\lambda})$ — седловая точка функции Лагранжа.
2. $f(\overline{x}) = h(\overline{\lambda})$.
3. $\min\limits_{x} \sup\limits_{\lambda \geq 0} L(x, \lambda) = \max\limits_{\lambda \geq 0} \inf\limits_{x} L(x, \lambda)$.

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:

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

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

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

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

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

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

Линейное программирование (ЛП)

Произвольная задача линейного программирования:

$$w(x) = (c, x) \longrightarrow \min(\max)$$

$$(a_i, x) \le b_i, i \in I_1;$$

$$(a_i, x) = b_i, i \in I_2;$$

$$x_j \ge 0, j \in J_1;$$

$$x_j \text{ – свободное}, j \in J_2.$$

Замечание 2: любую задачу ЛП можно свести к эквивалентной задаче ЛП в канонической форме:

1)$(a_i, x) \le b_i \Leftrightarrow (a_i, x) + y = b_i, i \in I_1; y \ge 0,$

2)x_j – свободное $\Leftrightarrow x_j + y_1 - y_2 = 0; y_1, y_2 \ge 0,$

3)$(c, x) \longrightarrow \max \Leftrightarrow -(c, x) \longrightarrow \min$

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

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

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

Обозначения:

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

где $\boldsymbol{N} = [\boldsymbol{A}_j]_{j \in S'}$, $\boldsymbol{x} = (\boldsymbol{x}_B, \boldsymbol{x}_N)$, $\boldsymbol{x}_B = (\boldsymbol{x}_{\sigma(1)}, \ldots, \boldsymbol{x}_{\sigma(m)})$ — базисные, а $\boldsymbol{x}_N = (\boldsymbol{x}_j)_{j \in S'}$ — небазисные переменные.

$$\boldsymbol{E}\, \boldsymbol{x}_B + \boldsymbol{B}^{-1} \boldsymbol{N} \boldsymbol{x}_N = \boldsymbol{B}^{-1} \boldsymbol{b} \qquad (6')$$

Определение 3. Решение $(\boldsymbol{x}_B, \boldsymbol{x}_N) = (\boldsymbol{B}^{-1}\boldsymbol{b}, 0)$ системы уравнений (6) назовем базисным (соответствующим базису $\boldsymbol{B}$).

ЛП: понятие б.д.р.

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

Док-во: Дополним столбцы с индексами из $S(x)$ до базиса. ■

Определение 4. **Базисным допустимым решением (б.д.р.)** называется любой элемент множества Q, являющийся базисным решением системы уравнений (6).

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

ЛП: понятие б.д.р.

Определение 5. Вектор $\boldsymbol{x} \in \boldsymbol{Q}$ — **крайняя точка или вершина** множества $\boldsymbol{Q}$, если не существует допустимых решений $\boldsymbol{x}^1 \neq \boldsymbol{x}^2$ из $\boldsymbol{Q}$ таких, что $\boldsymbol{x} = \boldsymbol{\alpha x}^1 + (1 - \boldsymbol{\alpha})\boldsymbol{x}^2$, где $\boldsymbol{0 < \alpha < 1}$.

Теорема 2. **Вектор $\boldsymbol{x}$ — б.д.р. тогда и только тогда, когда $\boldsymbol{x}$ — крайняя точка множества $\boldsymbol{Q}$.**

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

$$\exists\, \boldsymbol{y} \neq \boldsymbol{0} : \ \boldsymbol{Ay} = \boldsymbol{0}, \qquad (*)$$

при этом $(\boldsymbol{x_j} = \boldsymbol{0} \Longrightarrow \boldsymbol{y_j} = \boldsymbol{0}) \Longrightarrow$

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

ЛП: понятие б.д.р.

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

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

тогда из $(**) \Longrightarrow$

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

$\Longrightarrow$

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

$\Longrightarrow$

$x^1 \neq x^2$ и $x = \frac{1}{2}x^1 + \frac{1}{2}x^2 \Rightarrow x$ не крайняя точка — противоречие $\Rightarrow$

$$x \text{ — б.д.р.}$$

ЛП: понятие б.д.р.

Пусть $\boldsymbol{x}$ — б.д.р., но

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

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

$$A(x^1 - x^2) = 0 \Longleftrightarrow \{A_j | x_j^1 \neq x_j^2\} \text{ — линейно зависимо } \Rightarrow$$

$$\{A_j | \alpha x_j^1 + (1-\alpha)x_j^2 > 0\} \text{ — линейно зависимо } \Rightarrow$$

$\{A_j | x_j > 0\}$ — линейно зависимо $\Rightarrow$ $\boldsymbol{x}$ не б.д.р. — противоречие,

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

ЛП: Критерий разрешимости

Теорема 3 (Критерий разрешимости). Задача линейного программирования (5)-(7) разрешима тогда и только тогда, когда множество допустимых решений не пусто и целевая функция ограничена на нем.

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

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

Определение грани множества Q

Пусть $\boldsymbol{S'} \cup \boldsymbol{S} = \{\mathbf{1}, \ldots, \boldsymbol{n}\}$, $\boldsymbol{S'} \cap \boldsymbol{S} = \emptyset$. Множество решений системы уравнений

$$\boldsymbol{Ax} = \boldsymbol{b}, \boldsymbol{x_j} = \mathbf{0}, \boldsymbol{j} \in \boldsymbol{S'}, \boldsymbol{x_j} \geq \mathbf{0}, \boldsymbol{j} \in \boldsymbol{S},$$

называется гранью множества допустимых решений (6)-(7).

Определение грани множества Q

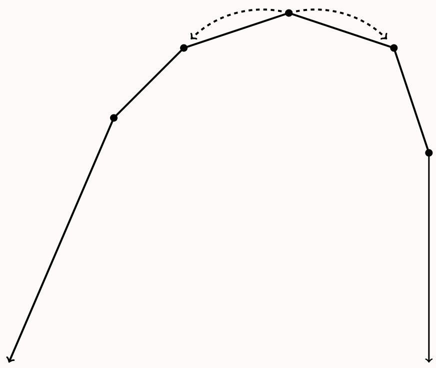
Величина $\boldsymbol{n-m-|S'|}$ — размерность данной грани (здесь $\boldsymbol{m+|S'|}$ – ранг системы уравнений).

Так как $\boldsymbol{\overline{x}}$ – б.д.р., то $\boldsymbol{|S'|=n-m}$, следовательно, $\boldsymbol{\overline{x}}$ – грань размерности 0.

Если $\boldsymbol{|S'|=n-m-1}$, то получим грань размерности 1, т.е. ребро: ограниченное или неограниченное.

Идея симплекс-метода

1) Разумно упорядоченный перебор вершин множества $\boldsymbol{Q}$ (б.д.р.);
2) Переход от одной вершины к другой по ребру.



Идея симплекс-метода

Пусть $\overline{\boldsymbol{x}}$ — б.д.р., $\boldsymbol{B} = (\boldsymbol{A}_{\boldsymbol{\sigma}(1)}, \ldots, \boldsymbol{A}_{\boldsymbol{\sigma}(m)})$ — базисная матрица. Тогда

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

следовательно

$$\boldsymbol{E x_B = B^{-1} b - B^{-1} N x_N}, \text{ и}$$

$$\boldsymbol{w(x) = c_B B^{-1} b + (c_N - c_B B^{-1} N) x_N}, \qquad (\mathbf{5}')$$

Идея симплекс-метода

$$z_{0j} = c_j - c_B B^{-1} A_j,\ j = 1, \ldots, n,$$

Величины z_{0j} называются оценками замещения. Знак оценки замещения определяет характер изменения целевой функции при движении по ребру.

$$(z_{10}, \ldots, z_{m0})^\top = B^{-1} b,$$

$$(z_{1j}, \ldots, z_{mj})^\top = B^{-1} A_j,\ \ j = 1, \ldots, n.$$

Элементарное преобразование б.д.р.

$x(t), t \geq 0$:

$$\begin{aligned} x_{\sigma(i)}(t) &= \overline{x}_{\sigma(i)} - z_{is}t, \\ x_s(t) &= t, \\ x_j(t) &= 0, j \in S' \backslash s \end{aligned} \tag{8}$$

Лемма 4 (признак оптимальности). Если оценки замещения неотрицательны, то текущее базисное допустимое решение $\overline{x}$ является оптимальным решением задачи (5)–(7).

Пусть $x \in Q$. Так как $z_{0j} \geq 0$ и $x_j \geq 0$, $j \in S'$, то из $(5')$ следует, что

$$w(x) = c_B B^{-1} b + \sum_{j \in S'} z_{0j} x_j \geq c_B B^{-1} b =$$

$$= w(\overline{x}) \blacksquare$$

Идея симплекс-метода

Лемма 5 (о неразрешимости). Если для номера s оценка замещения $z_{0s} < 0$ и для всех индексов i коэффициенты замещения z_{is} неположительны, то в задаче (5)–(7) не существует оптимального решения.

Идея симплекс-метода

Лемма 6 (о существовании лучшей вершины). Если оценка замещения $z_{0s} < 0$ и существуют базисные переменные с коэффициентами замещения $z_{is} > 0$, то элементарное преобразование приведёт либо в вершину с меньшим значением целевой функции, либо вершина останется прежней, но изменится её базис.

Идея симплекс-метода

Пусть

$$\bar{t} = \frac{\bar{x}_{\sigma(r)}}{z_{rs}} = \frac{z_{r0}}{z_{rs}} = \min_{z_{is}>0,\ i\geq 1} \frac{z_{i0}}{z_{is}}.$$

Предположим, что $\bar{t} > 0$. Тогда

$$\forall i \forall t < \bar{t}:\ x_{\sigma(i)}(t) = \bar{x}_{\sigma(i)} - z_{is}t > 0,$$

$x_{\sigma(r)}(\bar{t}) = 0$ и для $\forall t > \bar{t}$ имеем $x_{\sigma(r)}(\bar{t}) < 0$ $\Longrightarrow$

Идея симплекс-метода

Семейство векторов $\boldsymbol{x(t), 0 \le t \le \bar{t}}$ – ограниченное ребро множества $\boldsymbol{Q}$.

Вектор $\boldsymbol{x(\bar{t})}$ – б.д.р.:

$$\boldsymbol{(z_{1s}, \ldots, z_{ms})^\top = B^{-1} A_s \Longleftrightarrow}$$

$$\boldsymbol{A_s = B(z_{1s}, \ldots, z_{ms})^\top \Longleftrightarrow}$$

$\boldsymbol{A_s = \sum_{i=1}^{m} z_{is} A_{\sigma(i)}}$ (отсюда и условия) $\boldsymbol{z_{rs} > 0}$

следует, что матрица $\boldsymbol{B'}$ со столбцами

Идея симплекс-метода

$[A_{\sigma(1)}, \dots, A_{\sigma(r-1)}, A_s, A_{\sigma(r+1)}, \dots, A_{\sigma(m)}]$ невырождена.

Следовательно B' – базис б.д.р. $x(\bar{t})$. Т.к.

$$\forall t, 0 < t \le \bar{t}, w(x(t)) = -z_{00} + z_{0s} t < -z_{00},$$

то $x(\bar{t})$ – искомое б.д.р. с меньшим значением целевой функции.

Идея симплекс-метода

Пусть $\bar{\boldsymbol{t}} = \mathbf{0}$. В силу выбора $\boldsymbol{r} : \boldsymbol{x}_{\boldsymbol{\sigma}(\boldsymbol{r})}(\bar{\boldsymbol{t}}) =$
$= \boldsymbol{x}_{\boldsymbol{\sigma}(\boldsymbol{r})}(\mathbf{0}) = \overline{\boldsymbol{x}}_{\boldsymbol{\sigma}(\boldsymbol{i})} = \boldsymbol{z}_{\boldsymbol{r}\mathbf{0}} = \mathbf{0}$.

Т.к. $\boldsymbol{z}_{\boldsymbol{rs}} > \mathbf{0}$, то матрица $\boldsymbol{B}'$ со столбцами

$$[\boldsymbol{A}_{\boldsymbol{\sigma}(\mathbf{1})}, \ldots, \boldsymbol{A}_{\boldsymbol{\sigma}(\boldsymbol{r}-\mathbf{1})}, \boldsymbol{A}_{\boldsymbol{s}}, \boldsymbol{A}_{\boldsymbol{\sigma}(\boldsymbol{r}+\mathbf{1})}, \ldots, \boldsymbol{A}_{\boldsymbol{\sigma}(\boldsymbol{m})}]$$

снова невырождена. Следовательно $\boldsymbol{B}'$ – другой базис вершины $\overline{\boldsymbol{x}}$. ■