A low-poly, geometric illustration of a dog, possibly a Corgi, in shades of brown, tan, and white. The dog is sitting and facing forward. To its left is a large, yellow, low-poly rock or object. The background is a simple gradient from light gray to white.

# Двойственность в линейном программировании

Даниил Меркулов

Методы оптимизации. МФТИ

## Двойственность в линейном программировании

# Двойственность

Прямая задача:

$$\begin{aligned} & \min_{x \in \mathbb{R}^n} c^\top x \\ \text{s.t. } & Ax = b \\ & x_i \geq 0, \quad i = 1, \dots, n \end{aligned} \quad (1)$$

# Двойственность

Прямая задача:

$$\begin{aligned} \min_{x \in \mathbb{R}^n} c^\top x \\ \text{s.t. } Ax = b \\ x_i \geq 0, \quad i = 1, \dots, n \end{aligned} \quad (1)$$

ККТ для оптимальных  $x^*, \nu^*, \lambda^*$ :

$$L(x, \nu, \lambda) = c^\top x + \nu^\top (Ax - b) - \lambda^\top x$$

$$-A^\top \nu^* + \lambda^* = c$$

$$Ax^* = b$$

$$x^* \succeq 0$$

$$\lambda^* \succeq 0$$

$$\lambda_i^* x_i^* = 0$$

# Двойственность

Прямая задача:

$$\begin{aligned} \min_{x \in \mathbb{R}^n} c^\top x \\ \text{s.t. } Ax = b \\ x_i \geq 0, \quad i = 1, \dots, n \end{aligned}$$

ККТ для оптимальных  $x^*, \nu^*, \lambda^*$ :

$$\begin{aligned} L(x, \nu, \lambda) &= c^\top x + \nu^\top (Ax - b) - \lambda^\top x \\ &\quad - A^\top \nu^* + \lambda^* = c \\ Ax^* &= b \\ x^* &\succeq 0 \\ \lambda^* &\succeq 0 \\ \lambda_i^* x_i^* &= 0 \end{aligned}$$

Имеет следующую двойственную:

$$\begin{aligned} \max_{\nu \in \mathbb{R}^m} -b^\top \nu \\ \text{s.t. } -A^\top \nu \preceq c \end{aligned} \tag{2}$$

Найдите двойственную задачу к задаче выше (она должна быть исходной задаче ЛП). Также запишите условия ККТ для двойственной задачи, чтобы убедиться, что они идентичны условиям ККТ для прямой задачи.

## Сильная двойственность в линейном программировании

- (i) Если любая из задач Equation 1 или Equation 2 имеет (конечное) решение, то и другая имеет, и значения целевых функций равны.

## Сильная двойственность в линейном программировании

- (i) Если любая из задач Equation 1 или Equation 2 имеет (конечное) решение, то и другая имеет, и значения целевых функций равны.
- (ii) Если любая из задач Equation 1 или Equation 2 неограничена, то другая задача недопустима.

## Сильная двойственность в линейном программировании

- (i) Если любая из задач Equation 1 или Equation 2 имеет (конечное) решение, то и другая имеет, и значения целевых функций равны.
- (ii) Если любая из задач Equation 1 или Equation 2 неограничена, то другая задача недопустима.

## Сильная двойственность в линейном программировании

- (i) Если любая из задач Equation 1 или Equation 2 имеет (конечное) решение, то и другая имеет, и значения целевых функций равны.
- (ii) Если любая из задач Equation 1 или Equation 2 неограничена, то другая задача недопустима.

**ДОК-ВО.** Для (i) предположим, что Equation 1 имеет конечное оптимальное решение  $x^*$ . Из ККТ следует, что существуют оптимальные векторы  $\lambda^*$  и  $\nu^*$  такие, что  $(x^*, \nu^*, \lambda^*)$  удовлетворяет ККТ. Мы отметили выше, что ККТ для Equation 1 и Equation 2 эквивалентны. Более того, как утверждалось,  $c^T x^* = (-A^T \nu^* + \lambda^*)^T x^* = -(\nu^*)^T A x^* = -b^T \nu^*$ . Симметричный аргумент справедлив, если мы начнем с предположения, что двойственная задача Equation 2 имеет решение.

## Сильная двойственность в линейном программировании

- (i) Если любая из задач Equation 1 или Equation 2 имеет (конечное) решение, то и другая имеет, и значения целевых функций равны.
- (ii) Если любая из задач Equation 1 или Equation 2 неограничена, то другая задача недопустима.

**ДОК-ВО.** Для (i) предположим, что Equation 1 имеет конечное оптимальное решение  $x^*$ . Из ККТ следует, что существуют оптимальные векторы  $\lambda^*$  и  $\nu^*$  такие, что  $(x^*, \nu^*, \lambda^*)$  удовлетворяет ККТ. Мы отметили выше, что ККТ для Equation 1 и Equation 2 эквивалентны. Более того, как утверждалось,  $c^T x^* = (-A^T \nu^* + \lambda^*)^T x^* = -(\nu^*)^T A x^* = -b^T \nu^*$ . Симметричный аргумент справедлив, если мы начнем с предположения, что двойственная задача Equation 2 имеет решение.

Чтобы доказать (ii), предположим, что прямая задача неограничена, то есть существует последовательность точек  $x_k$ ,  $k = 1, 2, 3, \dots$  такая, что

$$c^T x_k \downarrow -\infty, \quad A x_k = b, \quad x_k \geq 0.$$

## Сильная двойственность в линейном программировании

- (i) Если любая из задач Equation 1 или Equation 2 имеет (конечное) решение, то и другая имеет, и значения целевых функций равны.
- (ii) Если любая из задач Equation 1 или Equation 2 неограничена, то другая задача недопустима.

**ДОК-ВО.** Для (i) предположим, что Equation 1 имеет конечное оптимальное решение  $x^*$ . Из ККТ следует, что существуют оптимальные векторы  $\lambda^*$  и  $\nu^*$  такие, что  $(x^*, \nu^*, \lambda^*)$  удовлетворяет ККТ. Мы отметили выше, что ККТ для Equation 1 и Equation 2 эквивалентны. Более того, как утверждалось,  $c^T x^* = (-A^T \nu^* + \lambda^*)^T x^* = -(\nu^*)^T A x^* = -b^T \nu^*$ . Симметричный аргумент справедлив, если мы начнем с предположения, что двойственная задача Equation 2 имеет решение.

Чтобы доказать (ii), предположим, что прямая задача неограничена, то есть существует последовательность точек  $x_k$ ,  $k = 1, 2, 3, \dots$  такая, что

$$c^T x_k \downarrow -\infty, \quad A x_k = b, \quad x_k \geq 0.$$

Предположим также, что двойственная задача Equation 2 допустима, то есть существует вектор  $\bar{\nu}$  такой, что  $-A^T \bar{\nu} \leq c$ . Из последнего неравенства вместе с  $x_k \geq 0$  имеем, что  $-\bar{\nu}^T A x_k \leq c^T x_k$ , и поэтому

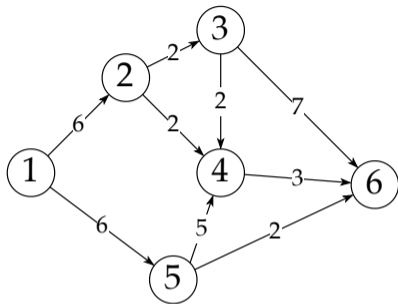
$$-\bar{\nu}^T b = -\bar{\nu}^T A x_k \leq c^T x_k \downarrow -\infty,$$

что приводит к противоречию. Следовательно, двойственная задача должна быть недопустимой.

Аналогичный аргумент можно использовать, чтобы показать, что неограниченность двойственной задачи влечет недопустимость прямой.

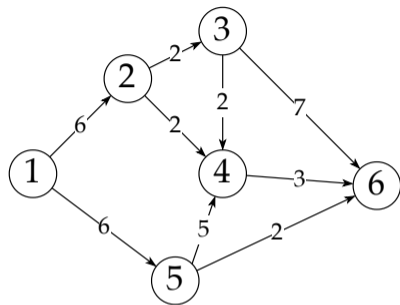
## Максимальный поток - минимальный разрез

## Задача максимального потока



Узлы — это маршрутизаторы, рёбра — это каналы связи; с каждым узлом связана пропускная способность — узел 1 может передавать узлу 2 до 6 Мбит/с, узел 2 может передавать узлу 4 до 2 Мбит/с и т.д.

# Задача максимального потока

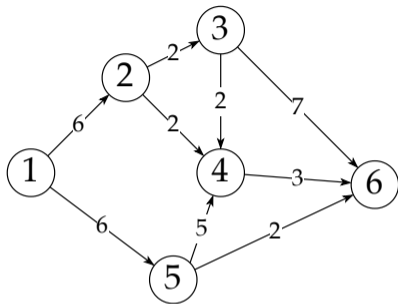


## Вопрос:

- Сеть узлов и рёбер представляет каналы связи, каждый с указанной пропускной способностью.

Узлы — это маршрутизаторы, рёбра — это каналы связи; с каждым узлом связана пропускная способность — узел 1 может передавать узлу 2 до 6 Мбит/с, узел 2 может передавать узлу 4 до 2 Мбит/с и т.д.

# Задача максимального потока

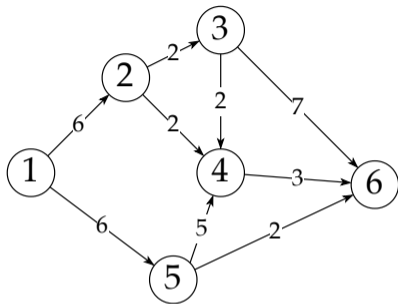


## Вопрос:

- Сеть узлов и рёбер представляет каналы связи, каждый с указанной пропускной способностью.
- Пример: Может ли узел 1 (источник) передавать узлу 6 (сток) со скоростью 6 Мбит/с? 12 Мбит/с? Какова максимальная скорость?

Узлы — это маршрутизаторы, рёбра — это каналы связи; с каждым узлом связана пропускная способность — узел 1 может передавать узлу 2 до 6 Мбит/с, узел 2 может передавать узлу 4 до 2 Мбит/с и т.д.

# Задача максимального потока

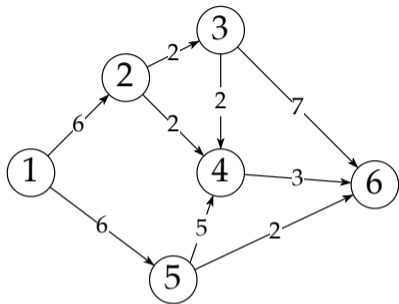


## Вопрос:

- Сеть узлов и рёбер представляет каналы связи, каждый с указанной пропускной способностью.
- Пример: Может ли узел 1 (источник) передавать узлу 6 (сток) со скоростью 6 Мбит/с? 12 Мбит/с? Какова максимальная скорость?

Узлы — это маршрутизаторы, рёбра — это каналы связи; с каждым узлом связана пропускная способность — узел 1 может передавать узлу 2 до 6 Мбит/с, узел 2 может передавать узлу 4 до 2 Мбит/с и т.д.

# Задача максимального потока



Узлы — это маршрутизаторы, рёбра — это каналы связи; с каждым узлом связана пропускная способность — узел 1 может передавать узлу 2 до 6 Мбит/с, узел 2 может передавать узлу 4 до 2 Мбит/с и т.д.

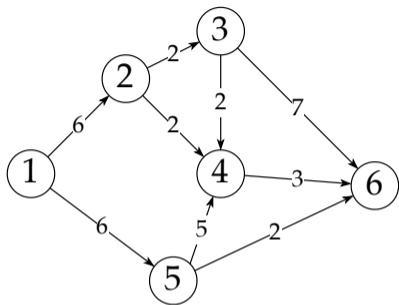
## Вопрос:

- Сеть узлов и рёбер представляет каналы связи, каждый с указанной пропускной способностью.
- Пример: Может ли узел 1 (источник) передавать узлу 6 (сток) со скоростью 6 Мбит/с? 12 Мбит/с? Какова максимальная скорость?

## Матрица пропускной способности:

$$C = \begin{bmatrix} 0 & 6 & 0 & 0 & 6 & 0 \\ 0 & 0 & 2 & 2 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 & 7 \\ 0 & 0 & 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 5 & 0 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

# Задача максимального потока



Узлы — это маршрутизаторы, рёбра — это каналы связи; с каждым узлом связана пропускная способность — узел 1 может передавать узлу 2 до 6 Мбит/с, узел 2 может передавать узлу 4 до 2 Мбит/с и т.д.

## Вопрос:

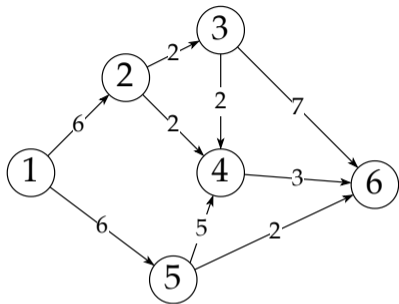
- Сеть узлов и рёбер представляет каналы связи, каждый с указанной пропускной способностью.
- Пример: Может ли узел 1 (источник) передавать узлу 6 (сток) со скоростью 6 Мбит/с? 12 Мбит/с? Какова максимальная скорость?

## Матрица пропускной способности:

$$C = \begin{bmatrix} 0 & 6 & 0 & 0 & 6 & 0 \\ 0 & 0 & 2 & 2 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 & 7 \\ 0 & 0 & 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 5 & 0 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

**Матрица потока:**  $X[i, j]$  представляет поток от узла  $i$  к узлу  $j$ .

# Задача максимального потока



## Вопрос:

- Сеть узлов и рёбер представляет каналы связи, каждый с указанной пропускной способностью.
- Пример: Может ли узел 1 (источник) передавать узлу 6 (сток) со скоростью 6 Мбит/с? 12 Мбит/с? Какова максимальная скорость?

## Матрица пропускной способности:

$$C = \begin{bmatrix} 0 & 6 & 0 & 0 & 6 & 0 \\ 0 & 0 & 2 & 2 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 & 7 \\ 0 & 0 & 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 5 & 0 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Узлы — это маршрутизаторы, рёбра — это каналы связи; с каждым узлом связана пропускная способность — узел 1 может передавать узлу 2 до 6 Мбит/с, узел 2 может передавать узлу 4 до 2 Мбит/с и т.д.

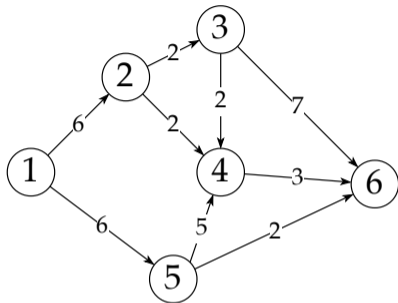
**Матрица потока:**  $X[i, j]$  представляет поток от узла  $i$  к узлу  $j$ .

## Ограничения:

$$0 \leq X \leq C$$

Сохранение потока: 
$$\sum_{j=2}^N X(i, j) = \sum_{k=1}^{N-1} X(k, i), \quad i = 2, \dots, N-1$$

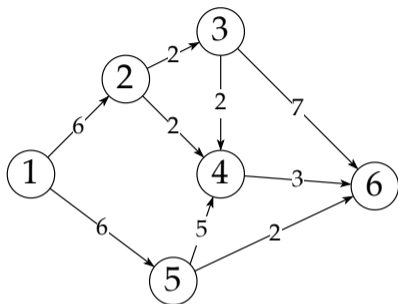
## Задача максимального потока



При данной настройке, когда всё, что производится источником, пойдёт к стоку, поток сети — это просто сумма всего, что выходит из источника:

$$\sum_{i=2}^N X(1, i) \quad (\text{Поток})$$

# Задача максимального потока



$$\begin{aligned}
 &\max \langle X, S \rangle \\
 &\text{s.t. } -X \leq 0 \\
 &\quad X \leq C \\
 &\langle X, L_n \rangle = 0, \quad n = 2, \dots, N-1,
 \end{aligned}
 \quad (\text{Задача максимального потока})$$

$L_n$  составляется как вычитание двух матриц. Обе матрицы почти полностью состоят из нулей. Только в первой матрице  $n$ -ый столбец состоит из единиц (кроме последней строки). А во второй матрице только  $n$ -ая строка состоит из единиц (кроме первого столбца).

При данной настройке, когда всё, что производится источником, пойдёт к стоку, поток сети — это просто сумма всего, что выходит из источника:

$$\sum_{i=2}^N X(1, i) \quad (\text{Поток})$$

$$S = \begin{bmatrix} 0 & 1 & \dots & 1 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{bmatrix}, \quad L_2 = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & -1 & \dots & -1 \\ 0 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 \end{bmatrix}.$$

# Вывод двойственной задачи к максимальному потоку

## Вывод двойственной задачи к максимальному потоку

$$\begin{aligned} & \min \langle \Lambda, C \rangle \\ & \Lambda, \nu \\ \text{s.t. } & \Lambda + Q \succeq S \\ & \Lambda \succeq 0 \end{aligned} \quad \text{(Двойственная задача максимального потока)}$$

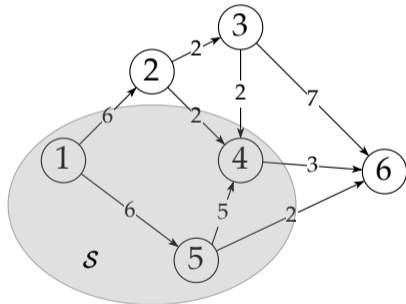
где

$$Q = \begin{bmatrix} 0 & \nu_2 & \nu_3 & \cdots & \nu_{N-1} & 0 \\ 0 & 0 & \nu_3 - \nu_2 & \cdots & \nu_{N-1} - \nu_2 & -\nu_2 \\ 0 & \nu_2 - \nu_3 & 0 & \cdots & \nu_{N-1} - \nu_3 & -\nu_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & \nu_2 - \nu_{N-1} & \nu_3 - \nu_{N-1} & \cdots & 0 & -\nu_{N-1} \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{bmatrix}.$$

## Задача минимального разреза

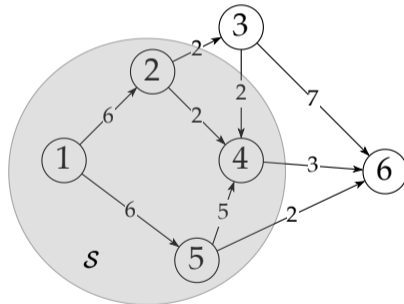
Разрез сети разделяет вершины на два множества: одно содержит источник (мы называем это множество  $\mathcal{S}$ ), а другое содержит сток. Пропускная способность разреза — это общее значение рёбер, выходящих из  $\mathcal{S}$  — мы разделяем множества, “отрезая поток” вдоль этих рёбер.

$$\mathcal{S} = \{1, 4, 5\}$$



Рёбра в разрезе:  $1 \rightarrow 2$ ,  $4 \rightarrow 6$ , и  $5 \rightarrow 6$ , пропускная способность этого разреза равна  $6 + 3 + 2 = 11$ .

$$\mathcal{S} = \{1, 2, 4, 5\}$$



Рёбра в разрезе:  $2 \rightarrow 3$ ,  $4 \rightarrow 6$ , и  $5 \rightarrow 6$ , пропускная способность этого разреза равна  $2 + 3 + 2 = 7$ .

## Минимальный разрез является двойственным к максимальному потоку

Каково минимальное значение наименьшего разреза? Мы докажем, что оно равно оптимальному значению решения  $d^*$  двойственной программы (Двойственная задача максимального потока).

## Минимальный разрез является двойственным к максимальному потоку

Каково минимальное значение наименьшего разреза? Мы докажем, что оно равно оптимальному значению решения  $d^*$  двойственной программы (Двойственная задача максимального потока).

Сначала предположим, что  $\mathcal{S}$  — допустимый разрез. Из  $\mathcal{S}$  мы можем легко найти двойственную допустимую точку, которая соответствует его пропускной способности: для  $n = 1, \dots, N$  возьмём

$$\nu_n = \begin{cases} 1, & n \in \mathcal{S}, \\ 0, & n \notin \mathcal{S}, \end{cases} \quad \text{и} \quad \lambda_{i,j} = \begin{cases} \max(\nu_i - \nu_j, 0), & i \neq 1, j \neq N, \\ 1 - \nu_j, & i = 1, \\ \nu_i, & j = N. \end{cases}$$

## Минимальный разрез является двойственным к максимальному потоку

Каково минимальное значение наименьшего разреза? Мы докажем, что оно равно оптимальному значению решения  $d^*$  двойственной программы (Двойственная задача максимального потока).

Сначала предположим, что  $\mathcal{S}$  — допустимый разрез. Из  $\mathcal{S}$  мы можем легко найти двойственную допустимую точку, которая соответствует его пропускной способности: для  $n = 1, \dots, N$  возьмём

$$\nu_n = \begin{cases} 1, & n \in \mathcal{S}, \\ 0, & n \notin \mathcal{S}, \end{cases} \quad \text{и} \quad \lambda_{i,j} = \begin{cases} \max(\nu_i - \nu_j, 0), & i \neq 1, j \neq N, \\ 1 - \nu_j, & i = 1, \\ \nu_i, & j = N. \end{cases}$$

Заметим, что эти выборы подчиняются ограничениям в двойственной задаче, и что  $\lambda_{i,j}$  будет равен 1, если  $i \rightarrow j$  разрезан, и 0 в противном случае, поэтому

$$\text{capacity}(\mathcal{S}) = \sum_{i,j} \lambda_{i,j} C_{i,j}.$$

## Минимальный разрез является двойственным к максимальному потоку

Каково минимальное значение наименьшего разреза? Мы докажем, что оно равно оптимальному значению решения  $d^*$  двойственной программы (Двойственная задача максимального потока).

Сначала предположим, что  $\mathcal{S}$  — допустимый разрез. Из  $\mathcal{S}$  мы можем легко найти двойственную допустимую точку, которая соответствует его пропускной способности: для  $n = 1, \dots, N$  возьмём

$$\nu_n = \begin{cases} 1, & n \in \mathcal{S}, \\ 0, & n \notin \mathcal{S}, \end{cases} \quad \text{и} \quad \lambda_{i,j} = \begin{cases} \max(\nu_i - \nu_j, 0), & i \neq 1, j \neq N, \\ 1 - \nu_j, & i = 1, \\ \nu_i, & j = N. \end{cases}$$

Заметим, что эти выборы подчиняются ограничениям в двойственной задаче, и что  $\lambda_{i,j}$  будет равен 1, если  $i \rightarrow j$  разрезан, и 0 в противном случае, поэтому

$$\text{capacity}(\mathcal{S}) = \sum_{i,j} \lambda_{i,j} C_{i,j}.$$

Каждый разрез допустим, поэтому

$$d^* \leq \text{MINCUT}.$$

## Минимальный разрез является двойственным к максимальному потоку

Теперь мы покажем, что для каждого решения  $\nu^*, \lambda^*$  двойственной задачи существует разрез с пропускной способностью не более  $d^*$ . Мы генерируем разрез *случайно*, а затем показываем, что ожидаемое значение пропускной способности разреза меньше  $d^*$  — это означает, что должен существовать хотя бы один с пропускной способностью  $d^*$  или меньше.

## Минимальный разрез является двойственным к максимальному потоку

Теперь мы покажем, что для каждого решения  $\nu^*, \lambda^*$  двойственной задачи существует разрез с пропускной способностью не более  $d^*$ . Мы генерируем разрез *случайно*, а затем показываем, что ожидаемое значение пропускной способности разреза меньше  $d^*$  — это означает, что должен существовать хотя бы один с пропускной способностью  $d^*$  или меньше.

Пусть  $Z$  — равномерная случайная величина на  $[0, 1]$ . Вместе с  $\lambda^*, \nu_2^*, \dots, \nu_{N-1}^*$ , полученными решением (Двойственная задача максимального потока), возьмём  $\nu_1 = 1$  и  $\nu_N = 0$ . Создадим разрез  $\mathcal{S}$  по правилу:

если  $\nu_n^* > Z$ , то возьмём  $n \in \mathcal{S}$ .

. . . Вероятность того, что конкретное ребро  $i \rightarrow j$  находится в этом разрезе, равна

$$\begin{aligned} P(i \in \mathcal{S}, j \notin \mathcal{S}) &= P(\nu_j^* \leq Z \leq \nu_i^*) \\ &\leq \begin{cases} \max(\nu_i^* - \nu_j^*, 0), & 2 \leq i, j \leq N-1, \\ 1 - \nu_j^*, & i = 1; j = 2, \dots, N-1, \\ \nu_i^*, & i = 2, \dots, N-1; j = N, \\ 1, & i = 1; j = N. \end{cases} \\ &\leq \lambda_{i,j}^*, \end{aligned}$$

## Минимальный разрез является двойственным к максимальному потоку

Последнее неравенство следует просто из ограничений в двойственной программе (Двойственная задача максимального потока). Этот разрез случайный, поэтому его пропускная способность — случайная величина, и её ожидание равно

$$\begin{aligned}\mathbb{E}[\text{capacity}(\mathcal{S})] &= \sum_{i,j} C_{i,j} P(i \in \mathcal{S}, j \notin \mathcal{S}) \\ &\leq \sum_{i,j} C_{i,j} \lambda_{i,j}^* = d^*.\end{aligned}$$

## Минимальный разрез является двойственным к максимальному потоку

Последнее неравенство следует просто из ограничений в двойственной программе (Двойственная задача максимального потока). Этот разрез случайный, поэтому его пропускная способность — случайная величина, и её ожидание равно

$$\begin{aligned}\mathbb{E}[\text{capacity}(\mathcal{S})] &= \sum_{i,j} C_{i,j} P(i \in \mathcal{S}, j \notin \mathcal{S}) \\ &\leq \sum_{i,j} C_{i,j} \lambda_{i,j}^* = d^*.\end{aligned}$$

Таким образом, должен существовать разрез, пропускная способность которого не более  $d^*$ . Это устанавливает, что

$$\text{MINCUT} \leq d^*.$$

## Минимальный разрез является двойственным к максимальному потоку

Последнее неравенство следует просто из ограничений в двойственной программе (Двойственная задача максимального потока). Этот разрез случайный, поэтому его пропускная способность — случайная величина, и её ожидание равно

$$\begin{aligned}\mathbb{E}[\text{capacity}(\mathcal{S})] &= \sum_{i,j} C_{i,j} P(i \in \mathcal{S}, j \notin \mathcal{S}) \\ &\leq \sum_{i,j} C_{i,j} \lambda_{i,j}^* = d^*.\end{aligned}$$

Таким образом, должен существовать разрез, пропускная способность которого не более  $d^*$ . Это устанавливает, что

$$\text{MINCUT} \leq d^*.$$

Объединяя эти два факта, конечно, получаем, что

$$d^* = \text{MINCUT} = \text{MAXFLOW} = p^*,$$

где  $p^*$  — решение прямой задачи, и равенство следует из сильной двойственности для линейного программирования.

## Минимальный разрез является двойственным к максимальному потоку

Последнее неравенство следует просто из ограничений в двойственной программе (Двойственная задача максимального потока). Этот разрез случайный, поэтому его пропускная способность — случайная величина, и её ожидание равно

$$\begin{aligned}\mathbb{E}[\text{capacity}(\mathcal{S})] &= \sum_{i,j} C_{i,j} P(i \in \mathcal{S}, j \notin \mathcal{S}) \\ &\leq \sum_{i,j} C_{i,j} \lambda_{i,j}^* = d^*.\end{aligned}$$

Таким образом, должен существовать разрез, пропускная способность которого не более  $d^*$ . Это устанавливает, что

$$\text{MINCUT} \leq d^*.$$

Объединяя эти два факта, конечно, получаем, что

$$d^* = \text{MINCUT} = \text{MAXFLOW} = p^*,$$

где  $p^*$  — решение прямой задачи, и равенство следует из сильной двойственности для линейного программирования.

**i** Теорема максимального потока — минимального разреза.

Максимальное значение s-t потока равно минимальной пропускной способности среди всех s-t разрезов.

## Анализ чувствительности

# Анализ чувствительности

Давайте перейдём от исходной задачи оптимизации

$$\begin{aligned} f_0(x) &\rightarrow \min_{x \in \mathbb{R}^n} \\ \text{s.t. } f_i(x) &\leq 0, \quad i = 1, \dots, m \\ h_i(x) &= 0, \quad i = 1, \dots, p \end{aligned} \quad (\text{P})$$

# Анализ чувствительности

Давайте перейдём от исходной задачи оптимизации

К возмущённой версии:

$$\begin{aligned} f_0(x) &\rightarrow \min_{x \in \mathbb{R}^n} \\ \text{s.t. } f_i(x) &\leq 0, \quad i = 1, \dots, m \\ h_i(x) &= 0, \quad i = 1, \dots, p \end{aligned} \quad (\text{P})$$

$$\begin{aligned} f_0(x) &\rightarrow \min_{x \in \mathbb{R}^n} \\ \text{s.t. } f_i(x) &\leq u_i, \quad i = 1, \dots, m \\ h_i(x) &= v_i, \quad i = 1, \dots, p \end{aligned} \quad (\text{Per})$$

## Анализ чувствительности

Давайте перейдём от исходной задачи оптимизации

К возмущённой версии:

$$\begin{aligned} f_0(x) &\rightarrow \min_{x \in \mathbb{R}^n} \\ \text{s.t. } f_i(x) &\leq 0, \quad i = 1, \dots, m \\ h_i(x) &= 0, \quad i = 1, \dots, p \end{aligned} \quad (\text{P})$$

$$\begin{aligned} f_0(x) &\rightarrow \min_{x \in \mathbb{R}^n} \\ \text{s.t. } f_i(x) &\leq u_i, \quad i = 1, \dots, m \\ h_i(x) &= v_i, \quad i = 1, \dots, p \end{aligned} \quad (\text{Per})$$

Заметим, что у нас по-прежнему есть только переменная  $x \in \mathbb{R}^n$ , в то время как  $u \in \mathbb{R}^m, v \in \mathbb{R}^p$  мы рассматриваем как параметры. Очевидно, что  $\text{Per}(u, v) \rightarrow P$ , если  $u = 0, v = 0$ . Мы будем обозначать оптимальное значение  $\text{Per}$  как  $p^*(u, v)$ , в то время как оптимальное значение исходной задачи  $P$  — просто  $p^*$ . Можно сразу сказать, что  $p^*(u, v) = p^*$ .

## Анализ чувствительности

Давайте перейдём от исходной задачи оптимизации      К возмущённой версии:

$$\begin{aligned} f_0(x) &\rightarrow \min_{x \in \mathbb{R}^n} \\ \text{s.t. } f_i(x) &\leq 0, \quad i = 1, \dots, m \\ h_i(x) &= 0, \quad i = 1, \dots, p \end{aligned} \quad (\text{P})$$

$$\begin{aligned} f_0(x) &\rightarrow \min_{x \in \mathbb{R}^n} \\ \text{s.t. } f_i(x) &\leq u_i, \quad i = 1, \dots, m \\ h_i(x) &= v_i, \quad i = 1, \dots, p \end{aligned} \quad (\text{Per})$$

Заметим, что у нас по-прежнему есть только переменная  $x \in \mathbb{R}^n$ , в то время как  $u \in \mathbb{R}^m, v \in \mathbb{R}^p$  мы рассматриваем как параметры. Очевидно, что  $\text{Per}(u, v) \rightarrow P$ , если  $u = 0, v = 0$ . Мы будем обозначать оптимальное значение  $\text{Per}$  как  $p^*(u, v)$ , в то время как оптимальное значение исходной задачи  $P$  — просто  $p^*$ . Можно сразу сказать, что  $p^*(u, v) = p^*$ .

Говоря о значении некоторого  $i$ -го ограничения, можно сказать, что

- $u_i = 0$  оставляет исходную задачу

## Анализ чувствительности

Давайте перейдём от исходной задачи оптимизации      К возмущённой версии:

$$\begin{aligned} f_0(x) &\rightarrow \min_{x \in \mathbb{R}^n} \\ \text{s.t. } f_i(x) &\leq 0, \quad i = 1, \dots, m \\ h_i(x) &= 0, \quad i = 1, \dots, p \end{aligned} \quad (\text{P})$$

$$\begin{aligned} f_0(x) &\rightarrow \min_{x \in \mathbb{R}^n} \\ \text{s.t. } f_i(x) &\leq u_i, \quad i = 1, \dots, m \\ h_i(x) &= v_i, \quad i = 1, \dots, p \end{aligned} \quad (\text{Per})$$

Заметим, что у нас по-прежнему есть только переменная  $x \in \mathbb{R}^n$ , в то время как  $u \in \mathbb{R}^m, v \in \mathbb{R}^p$  мы рассматриваем как параметры. Очевидно, что  $\text{Per}(u, v) \rightarrow P$ , если  $u = 0, v = 0$ . Мы будем обозначать оптимальное значение  $\text{Per}$  как  $p^*(u, v)$ , в то время как оптимальное значение исходной задачи  $P$  — просто  $p^*$ . Можно сразу сказать, что  $p^*(u, v) = p^*$ .

Говоря о значении некоторого  $i$ -го ограничения, можно сказать, что

- $u_i = 0$  оставляет исходную задачу
- $u_i > 0$  означает, что мы ослабили неравенство

## Анализ чувствительности

Давайте перейдём от исходной задачи оптимизации

К возмущённой версии:

$$\begin{aligned} f_0(x) &\rightarrow \min_{x \in \mathbb{R}^n} \\ \text{s.t. } f_i(x) &\leq 0, \quad i = 1, \dots, m \\ h_i(x) &= 0, \quad i = 1, \dots, p \end{aligned} \quad (\text{P})$$

$$\begin{aligned} f_0(x) &\rightarrow \min_{x \in \mathbb{R}^n} \\ \text{s.t. } f_i(x) &\leq u_i, \quad i = 1, \dots, m \\ h_i(x) &= v_i, \quad i = 1, \dots, p \end{aligned} \quad (\text{Per})$$

Заметим, что у нас по-прежнему есть только переменная  $x \in \mathbb{R}^n$ , в то время как  $u \in \mathbb{R}^m, v \in \mathbb{R}^p$  мы рассматриваем как параметры. Очевидно, что  $\text{Per}(u, v) \rightarrow P$ , если  $u = 0, v = 0$ . Мы будем обозначать оптимальное значение  $\text{Per}$  как  $p^*(u, v)$ , в то время как оптимальное значение исходной задачи  $P$  — просто  $p^*$ . Можно сразу сказать, что  $p^*(u, v) = p^*$ .

Говоря о значении некоторого  $i$ -го ограничения, можно сказать, что

- $u_i = 0$  оставляет исходную задачу
- $u_i > 0$  означает, что мы ослабили неравенство
- $u_i < 0$  означает, что мы ужесточили ограничение

## Анализ чувствительности

Давайте перейдём от исходной задачи оптимизации

К возмущённой версии:

$$\begin{aligned} f_0(x) &\rightarrow \min_{x \in \mathbb{R}^n} \\ \text{s.t. } f_i(x) &\leq 0, \quad i = 1, \dots, m \\ h_i(x) &= 0, \quad i = 1, \dots, p \end{aligned} \quad (\text{P})$$

$$\begin{aligned} f_0(x) &\rightarrow \min_{x \in \mathbb{R}^n} \\ \text{s.t. } f_i(x) &\leq u_i, \quad i = 1, \dots, m \\ h_i(x) &= v_i, \quad i = 1, \dots, p \end{aligned} \quad (\text{Per})$$

Заметим, что у нас по-прежнему есть только переменная  $x \in \mathbb{R}^n$ , в то время как  $u \in \mathbb{R}^m, v \in \mathbb{R}^p$  мы рассматриваем как параметры. Очевидно, что  $\text{Per}(u, v) \rightarrow P$ , если  $u = 0, v = 0$ . Мы будем обозначать оптимальное значение  $\text{Per}$  как  $p^*(u, v)$ , в то время как оптимальное значение исходной задачи  $P$  — просто  $p^*$ . Можно сразу сказать, что  $p^*(u, v) = p^*$ .

Говоря о значении некоторого  $i$ -го ограничения, можно сказать, что

- $u_i = 0$  оставляет исходную задачу
- $u_i > 0$  означает, что мы ослабили неравенство
- $u_i < 0$  означает, что мы ужесточили ограничение

## Анализ чувствительности

Давайте перейдём от исходной задачи оптимизации      К возмущённой версии:

$$\begin{aligned} f_0(x) &\rightarrow \min_{x \in \mathbb{R}^n} \\ \text{s.t. } f_i(x) &\leq 0, \quad i = 1, \dots, m \\ h_i(x) &= 0, \quad i = 1, \dots, p \end{aligned} \quad (\text{P})$$

$$\begin{aligned} f_0(x) &\rightarrow \min_{x \in \mathbb{R}^n} \\ \text{s.t. } f_i(x) &\leq u_i, \quad i = 1, \dots, m \\ h_i(x) &= v_i, \quad i = 1, \dots, p \end{aligned} \quad (\text{Per})$$

Заметим, что у нас по-прежнему есть только переменная  $x \in \mathbb{R}^n$ , в то время как  $u \in \mathbb{R}^m, v \in \mathbb{R}^p$  мы рассматриваем как параметры. Очевидно, что  $\text{Per}(u, v) \rightarrow P$ , если  $u = 0, v = 0$ . Мы будем обозначать оптимальное значение  $\text{Per}$  как  $p^*(u, v)$ , в то время как оптимальное значение исходной задачи  $P$  — просто  $p^*$ . Можно сразу сказать, что  $p^*(u, v) = p^*$ .

Говоря о значении некоторого  $i$ -го ограничения, можно сказать, что

- $u_i = 0$  оставляет исходную задачу
- $u_i > 0$  означает, что мы ослабили неравенство
- $u_i < 0$  означает, что мы ужесточили ограничение

Можно даже показать, что когда  $P$  является задачей выпуклой оптимизации,  $p^*(u, v)$  является выпуклой функцией.

## Анализ чувствительности

Предположим, что сильная двойственность выполняется для исходной задачи, и предположим, что  $x$  — любая допустимая точка для возмущённой задачи:

$$p^*(0, 0) = p^* = d^* = g(\lambda^*, \nu^*) \leq$$

## Анализ чувствительности

Предположим, что сильная двойственность выполняется для исходной задачи, и предположим, что  $x$  — любая допустимая точка для возмущённой задачи:

$$\begin{aligned} p^*(0, 0) = p^* = d^* = g(\lambda^*, \nu^*) &\leq \\ &\leq L(x, \lambda^*, \nu^*) = \end{aligned}$$

## Анализ чувствительности

Предположим, что сильная двойственность выполняется для исходной задачи, и предположим, что  $x$  — любая допустимая точка для возмущённой задачи:

$$\begin{aligned} p^*(0, 0) &= p^* = d^* = g(\lambda^*, \nu^*) \leq \\ &\leq L(x, \lambda^*, \nu^*) = \\ &= f_0(x) + \sum_{i=1}^m \lambda_i^* f_i(x) + \sum_{i=1}^p \nu_i^* h_i(x) \leq \end{aligned}$$

## Анализ чувствительности

Предположим, что сильная двойственность выполняется для исходной задачи, и предположим, что  $x$  — любая допустимая точка для возмущённой задачи:

$$\begin{aligned} p^*(0, 0) &= p^* = d^* = g(\lambda^*, \nu^*) \leq \\ &\leq L(x, \lambda^*, \nu^*) = \\ &= f_0(x) + \sum_{i=1}^m \lambda_i^* f_i(x) + \sum_{i=1}^p \nu_i^* h_i(x) \leq \\ &\leq f_0(x) + \sum_{i=1}^m \lambda_i^* u_i + \sum_{i=1}^p \nu_i^* v_i \end{aligned}$$

## Анализ чувствительности

Предположим, что сильная двойственность выполняется для исходной задачи, и предположим, что  $x$  — любая допустимая точка для возмущённой задачи:

$$\begin{aligned} p^*(0, 0) &= p^* = d^* = g(\lambda^*, \nu^*) \leq \\ &\leq L(x, \lambda^*, \nu^*) = \\ &= f_0(x) + \sum_{i=1}^m \lambda_i^* f_i(x) + \sum_{i=1}^p \nu_i^* h_i(x) \leq \\ &\leq f_0(x) + \sum_{i=1}^m \lambda_i^* u_i + \sum_{i=1}^p \nu_i^* v_i \end{aligned}$$

## Анализ чувствительности

Предположим, что сильная двойственность выполняется для исходной задачи, и предположим, что  $x$  — любая допустимая точка для возмущённой задачи:

$$\begin{aligned} p^*(0, 0) &= p^* = d^* = g(\lambda^*, \nu^*) \leq \\ &\leq L(x, \lambda^*, \nu^*) = \\ &= f_0(x) + \sum_{i=1}^m \lambda_i^* f_i(x) + \sum_{i=1}^p \nu_i^* h_i(x) \leq \\ &\leq f_0(x) + \sum_{i=1}^m \lambda_i^* u_i + \sum_{i=1}^p \nu_i^* v_i \end{aligned}$$

Что означает

$$f_0(x) \geq p^*(0, 0) - \lambda^{*T} u - \nu^{*T} v$$

## Анализ чувствительности

Предположим, что сильная двойственность выполняется для исходной задачи, и предположим, что  $x$  — любая допустимая точка для возмущённой задачи:

$$\begin{aligned} p^*(0, 0) &= p^* = d^* = g(\lambda^*, \nu^*) \leq \\ &\leq L(x, \lambda^*, \nu^*) = \\ &= f_0(x) + \sum_{i=1}^m \lambda_i^* f_i(x) + \sum_{i=1}^p \nu_i^* h_i(x) \leq \\ &\leq f_0(x) + \sum_{i=1}^m \lambda_i^* u_i + \sum_{i=1}^p \nu_i^* v_i \end{aligned}$$

Что означает

$$f_0(x) \geq p^*(0, 0) - \lambda^{*T} u - \nu^{*T} v$$

И взяв оптимальную  $x$  для возмущённой задачи, имеем:

$$p^*(u, v) \geq p^*(0, 0) - \lambda^{*T} u - \nu^{*T} v \quad (3)$$

## Анализ чувствительности

В сценариях, где выполняется сильная двойственность, мы можем сделать несколько выводов о чувствительности оптимальных решений по отношению к множителям Лагранжа.

- **Влияние ужесточения ограничения (большое  $\lambda_i^*$ ):**

Когда множитель Лагранжа  $i$ -го ограничения,  $\lambda_i^*$ , имеет существенное значение, и если это ограничение ужесточается (выбирается  $u_i < 0$ ), гарантируется, что оптимальное значение, обозначенное  $p^*(u, v)$ , значительно увеличится.

## Анализ чувствительности

В сценариях, где выполняется сильная двойственность, мы можем сделать несколько выводов о чувствительности оптимальных решений по отношению к множителям Лагранжа.

- **Влияние ужесточения ограничения (большое  $\lambda_i^*$ ):**

Когда множитель Лагранжа  $i$ -го ограничения,  $\lambda_i^*$ , имеет существенное значение, и если это ограничение ужесточается (выбирается  $u_i < 0$ ), гарантируется, что оптимальное значение, обозначенное  $p^*(u, v)$ , значительно увеличится.

- **Эффект корректировки ограничений с большими положительными или отрицательными  $\nu_i^*$ :**

Если  $\nu_i^*$  велик и положителен, и выбирается  $v_i < 0$ , или если  $\nu_i^*$  велик и отрицателен, и выбирается  $v_i > 0$ , то в любом сценарии ожидается, что оптимальное значение  $p^*(u, v)$  значительно увеличится.

## Анализ чувствительности

В сценариях, где выполняется сильная двойственность, мы можем сделать несколько выводов о чувствительности оптимальных решений по отношению к множителям Лагранжа.

- **Влияние ужесточения ограничения (большое  $\lambda_i^*$ ):**

Когда множитель Лагранжа  $i$ -го ограничения,  $\lambda_i^*$ , имеет существенное значение, и если это ограничение ужесточается (выбирается  $u_i < 0$ ), гарантируется, что оптимальное значение, обозначенное  $p^*(u, v)$ , значительно увеличится.

- **Эффект корректировки ограничений с большими положительными или отрицательными  $\nu_i^*$ :**

Если  $\nu_i^*$  велик и положителен, и выбирается  $v_i < 0$ , или если  $\nu_i^*$  велик и отрицателен, и выбирается  $v_i > 0$ , то в любом сценарии ожидается, что оптимальное значение  $p^*(u, v)$  значительно увеличится.

- **Последствия ослабления ограничения (малое  $\lambda_i^*$ ):**

Если множитель Лагранжа  $\lambda_i^*$  для  $i$ -го ограничения относительно мал, и ограничение ослабляется (выбирается  $u_i > 0$ ), ожидается, что оптимальное значение  $p^*(u, v)$  не значительно уменьшится.

## Анализ чувствительности

В сценариях, где выполняется сильная двойственность, мы можем сделать несколько выводов о чувствительности оптимальных решений по отношению к множителям Лагранжа.

- **Влияние ужесточения ограничения (большое  $\lambda_i^*$ ):**

Когда множитель Лагранжа  $i$ -го ограничения,  $\lambda_i^*$ , имеет существенное значение, и если это ограничение ужесточается (выбирается  $u_i < 0$ ), гарантируется, что оптимальное значение, обозначенное  $p^*(u, v)$ , значительно увеличится.

- **Эффект корректировки ограничений с большими положительными или отрицательными  $\nu_i^*$ :**

Если  $\nu_i^*$  велик и положителен, и выбирается  $v_i < 0$ , или если  $\nu_i^*$  велик и отрицателен, и выбирается  $v_i > 0$ , то в любом сценарии ожидается, что оптимальное значение  $p^*(u, v)$  значительно увеличится.

- **Последствия ослабления ограничения (малое  $\lambda_i^*$ ):**

Если множитель Лагранжа  $\lambda_i^*$  для  $i$ -го ограничения относительно мал, и ограничение ослабляется (выбирается  $u_i > 0$ ), ожидается, что оптимальное значение  $p^*(u, v)$  не значительно уменьшится.

- **Результаты крошечных корректировок в ограничениях с малыми  $\nu_i^*$ :**

Когда  $\nu_i^*$  мал и положителен, и выбирается  $v_i > 0$ , или когда  $\nu_i^*$  мал и отрицателен, и выбирается  $v_i < 0$ , в обоих случаях оптимальное значение  $p^*(u, v)$  незначительно уменьшится.

## Анализ чувствительности

В сценариях, где выполняется сильная двойственность, мы можем сделать несколько выводов о чувствительности оптимальных решений по отношению к множителям Лагранжа.

- **Влияние ужесточения ограничения (большое  $\lambda_i^*$ ):**

Когда множитель Лагранжа  $i$ -го ограничения,  $\lambda_i^*$ , имеет существенное значение, и если это ограничение ужесточается (выбирается  $u_i < 0$ ), гарантируется, что оптимальное значение, обозначенное  $p^*(u, v)$ , значительно увеличится.

- **Эффект корректировки ограничений с большими положительными или отрицательными  $\nu_i^*$ :**

Если  $\nu_i^*$  велик и положителен, и выбирается  $v_i < 0$ , или если  $\nu_i^*$  велик и отрицателен, и выбирается  $v_i > 0$ , то в любом сценарии ожидается, что оптимальное значение  $p^*(u, v)$  значительно увеличится.

- **Последствия ослабления ограничения (малое  $\lambda_i^*$ ):**

Если множитель Лагранжа  $\lambda_i^*$  для  $i$ -го ограничения относительно мал, и ограничение ослабляется (выбирается  $u_i > 0$ ), ожидается, что оптимальное значение  $p^*(u, v)$  не значительно уменьшится.

- **Результаты крошечных корректировок в ограничениях с малыми  $\nu_i^*$ :**

Когда  $\nu_i^*$  мал и положителен, и выбирается  $v_i > 0$ , или когда  $\nu_i^*$  мал и отрицателен, и выбирается  $v_i < 0$ , в обоих случаях оптимальное значение  $p^*(u, v)$  незначительно уменьшится.

## Анализ чувствительности

В сценариях, где выполняется сильная двойственность, мы можем сделать несколько выводов о чувствительности оптимальных решений по отношению к множителям Лагранжа.

- **Влияние ужесточения ограничения (большое  $\lambda_i^*$ ):**

Когда множитель Лагранжа  $i$ -го ограничения,  $\lambda_i^*$ , имеет существенное значение, и если это ограничение ужесточается (выбирается  $u_i < 0$ ), гарантируется, что оптимальное значение, обозначенное  $p^*(u, v)$ , значительно увеличится.

- **Эффект корректировки ограничений с большими положительными или отрицательными  $\nu_i^*$ :**

Если  $\nu_i^*$  велик и положителен, и выбирается  $v_i < 0$ , или если  $\nu_i^*$  велик и отрицателен, и выбирается  $v_i > 0$ , то в любом сценарии ожидается, что оптимальное значение  $p^*(u, v)$  значительно увеличится.

- **Последствия ослабления ограничения (малое  $\lambda_i^*$ ):**

Если множитель Лагранжа  $\lambda_i^*$  для  $i$ -го ограничения относительно мал, и ограничение ослабляется (выбирается  $u_i > 0$ ), ожидается, что оптимальное значение  $p^*(u, v)$  не значительно уменьшится.

- **Результаты крошечных корректировок в ограничениях с малыми  $\nu_i^*$ :**

Когда  $\nu_i^*$  мал и положителен, и выбирается  $v_i > 0$ , или когда  $\nu_i^*$  мал и отрицателен, и выбирается  $v_i < 0$ , в обоих случаях оптимальное значение  $p^*(u, v)$  незначительно уменьшится.

Эти интерпретации предоставляют основу для понимания того, как изменения в ограничениях, отражённые через соответствующие множители Лагранжа, влияют на оптимальное решение в задачах, где выполняется сильная двойственность.

## Локальная чувствительность

Предположим теперь, что  $p^*(u, v)$  дифференцируема в точке  $u = 0, v = 0$ .

## Локальная чувствительность

Предположим теперь, что  $p^*(u, v)$  дифференцируема в точке  $u = 0, v = 0$ .

$$\lambda_i^* = -\frac{\partial p^*(0, 0)}{\partial u_i} \quad \nu_i^* = -\frac{\partial p^*(0, 0)}{\partial v_i} \quad (4)$$

## Локальная чувствительность

Предположим теперь, что  $p^*(u, v)$  дифференцируема в точке  $u = 0, v = 0$ .

$$\lambda_i^* = -\frac{\partial p^*(0, 0)}{\partial u_i} \quad \nu_i^* = -\frac{\partial p^*(0, 0)}{\partial v_i} \quad (4)$$

Чтобы показать этот результат, мы рассматриваем направленную производную  $p^*(u, v)$  вдоль направления некоторого  $i$ -го базисного вектора  $e_i$ :

## Локальная чувствительность

Предположим теперь, что  $p^*(u, v)$  дифференцируема в точке  $u = 0, v = 0$ .

$$\lambda_i^* = -\frac{\partial p^*(0, 0)}{\partial u_i} \quad \nu_i^* = -\frac{\partial p^*(0, 0)}{\partial v_i} \quad (4)$$

Чтобы показать этот результат, мы рассматриваем направленную производную  $p^*(u, v)$  вдоль направления некоторого  $i$ -го базисного вектора  $e_i$ :

$$\lim_{t \rightarrow 0} \frac{p^*(te_i, 0) - p^*(0, 0)}{t} = \frac{\partial p^*(0, 0)}{\partial u_i}$$

## Локальная чувствительность

Предположим теперь, что  $p^*(u, v)$  дифференцируема в точке  $u = 0, v = 0$ .

$$\lambda_i^* = -\frac{\partial p^*(0, 0)}{\partial u_i} \quad \nu_i^* = -\frac{\partial p^*(0, 0)}{\partial v_i} \quad (4)$$

Чтобы показать этот результат, мы рассматриваем направленную производную  $p^*(u, v)$  вдоль направления некоторого  $i$ -го базисного вектора  $e_i$ :

$$\lim_{t \rightarrow 0} \frac{p^*(te_i, 0) - p^*(0, 0)}{t} = \frac{\partial p^*(0, 0)}{\partial u_i}$$

Из неравенства Equation 3 и взяв предел  $t \rightarrow 0$  с  $t > 0$ , имеем

## Локальная чувствительность

Предположим теперь, что  $p^*(u, v)$  дифференцируема в точке  $u = 0, v = 0$ .

$$\lambda_i^* = -\frac{\partial p^*(0, 0)}{\partial u_i} \quad \nu_i^* = -\frac{\partial p^*(0, 0)}{\partial v_i} \quad (4)$$

Чтобы показать этот результат, мы рассматриваем направленную производную  $p^*(u, v)$  вдоль направления некоторого  $i$ -го базисного вектора  $e_i$ :

$$\lim_{t \rightarrow 0} \frac{p^*(te_i, 0) - p^*(0, 0)}{t} = \frac{\partial p^*(0, 0)}{\partial u_i}$$

Из неравенства Equation 3 и взяв предел  $t \rightarrow 0$  с  $t > 0$ , имеем

$$\frac{p^*(te_i, 0) - p^*}{t} \geq -\lambda_i^* \rightarrow \frac{\partial p^*(0, 0)}{\partial u_i} \geq -\lambda_i^*$$

## Локальная чувствительность

Предположим теперь, что  $p^*(u, v)$  дифференцируема в точке  $u = 0, v = 0$ .

$$\lambda_i^* = -\frac{\partial p^*(0, 0)}{\partial u_i} \quad \nu_i^* = -\frac{\partial p^*(0, 0)}{\partial v_i} \quad (4)$$

Чтобы показать этот результат, мы рассматриваем направленную производную  $p^*(u, v)$  вдоль направления некоторого  $i$ -го базисного вектора  $e_i$ :

$$\lim_{t \rightarrow 0} \frac{p^*(te_i, 0) - p^*(0, 0)}{t} = \frac{\partial p^*(0, 0)}{\partial u_i}$$

Из неравенства Equation 3 и взяв предел  $t \rightarrow 0$  с  $t > 0$ , имеем

$$\frac{p^*(te_i, 0) - p^*}{t} \geq -\lambda_i^* \rightarrow \frac{\partial p^*(0, 0)}{\partial u_i} \geq -\lambda_i^*$$

Для отрицательного  $t < 0$  имеем:

## Локальная чувствительность

Предположим теперь, что  $p^*(u, v)$  дифференцируема в точке  $u = 0, v = 0$ .

$$\lambda_i^* = -\frac{\partial p^*(0, 0)}{\partial u_i} \quad \nu_i^* = -\frac{\partial p^*(0, 0)}{\partial v_i} \quad (4)$$

Чтобы показать этот результат, мы рассматриваем направленную производную  $p^*(u, v)$  вдоль направления некоторого  $i$ -го базисного вектора  $e_i$ :

$$\lim_{t \rightarrow 0} \frac{p^*(te_i, 0) - p^*(0, 0)}{t} = \frac{\partial p^*(0, 0)}{\partial u_i}$$

Из неравенства Equation 3 и взяв предел  $t \rightarrow 0$  с  $t > 0$ , имеем

$$\frac{p^*(te_i, 0) - p^*}{t} \geq -\lambda_i^* \rightarrow \frac{\partial p^*(0, 0)}{\partial u_i} \geq -\lambda_i^*$$

Для отрицательного  $t < 0$  имеем:

$$\frac{p^*(te_i, 0) - p^*}{t} \leq -\lambda_i^* \rightarrow \frac{\partial p^*(0, 0)}{\partial u_i} \leq -\lambda_i^*$$

## Локальная чувствительность

Предположим теперь, что  $p^*(u, v)$  дифференцируема в точке  $u = 0, v = 0$ .

Та же идея может быть использована для установления факта о  $v_i$ .

$$\lambda_i^* = -\frac{\partial p^*(0, 0)}{\partial u_i} \quad \nu_i^* = -\frac{\partial p^*(0, 0)}{\partial v_i} \quad (4)$$

Чтобы показать этот результат, мы рассматриваем направленную производную  $p^*(u, v)$  вдоль направления некоторого  $i$ -го базисного вектора  $e_i$ :

$$\lim_{t \rightarrow 0} \frac{p^*(te_i, 0) - p^*(0, 0)}{t} = \frac{\partial p^*(0, 0)}{\partial u_i}$$

Из неравенства Equation 3 и взяв предел  $t \rightarrow 0$  с  $t > 0$ , имеем

$$\frac{p^*(te_i, 0) - p^*}{t} \geq -\lambda_i^* \rightarrow \frac{\partial p^*(0, 0)}{\partial u_i} \geq -\lambda_i^*$$

Для отрицательного  $t < 0$  имеем:

$$\frac{p^*(te_i, 0) - p^*}{t} \leq -\lambda_i^* \rightarrow \frac{\partial p^*(0, 0)}{\partial u_i} \leq -\lambda_i^*$$

## Локальная чувствительность

Предположим теперь, что  $p^*(u, v)$  дифференцируема в точке  $u = 0, v = 0$ .

$$\lambda_i^* = -\frac{\partial p^*(0, 0)}{\partial u_i} \quad \nu_i^* = -\frac{\partial p^*(0, 0)}{\partial v_i} \quad (4)$$

Чтобы показать этот результат, мы рассматриваем направленную производную  $p^*(u, v)$  вдоль направления некоторого  $i$ -го базисного вектора  $e_i$ :

$$\lim_{t \rightarrow 0} \frac{p^*(te_i, 0) - p^*(0, 0)}{t} = \frac{\partial p^*(0, 0)}{\partial u_i}$$

Из неравенства Equation 3 и взяв предел  $t \rightarrow 0$  с  $t > 0$ , имеем

$$\frac{p^*(te_i, 0) - p^*}{t} \geq -\lambda_i^* \rightarrow \frac{\partial p^*(0, 0)}{\partial u_i} \geq -\lambda_i^*$$

Для отрицательного  $t < 0$  имеем:

$$\frac{p^*(te_i, 0) - p^*}{t} \leq -\lambda_i^* \rightarrow \frac{\partial p^*(0, 0)}{\partial u_i} \leq -\lambda_i^*$$

Та же идея может быть использована для установления факта о  $v_i$ .

Результат локальной чувствительности Equation 4 предоставляет способ понять влияние ограничений на оптимальное решение  $x^*$  задачи оптимизации. Если ограничение  $f_i(x^*)$  отрицательно в  $x^*$ , оно не влияет на оптимальное решение, что означает, что небольшие изменения в этом ограничении не изменят оптимальное значение. В этом случае соответствующий оптимальный множитель Лагранжа будет равен нулю, согласно принципу дополнительной нежёсткости.

## Локальная чувствительность

Предположим теперь, что  $p^*(u, v)$  дифференцируема в точке  $u = 0, v = 0$ .

$$\lambda_i^* = -\frac{\partial p^*(0, 0)}{\partial u_i} \quad \nu_i^* = -\frac{\partial p^*(0, 0)}{\partial v_i} \quad (4)$$

Чтобы показать этот результат, мы рассматриваем направленную производную  $p^*(u, v)$  вдоль направления некоторого  $i$ -го базисного вектора  $e_i$ :

$$\lim_{t \rightarrow 0} \frac{p^*(te_i, 0) - p^*(0, 0)}{t} = \frac{\partial p^*(0, 0)}{\partial u_i}$$

Из неравенства Equation 3 и взяв предел  $t \rightarrow 0$  с  $t > 0$ , имеем

$$\frac{p^*(te_i, 0) - p^*}{t} \geq -\lambda_i^* \rightarrow \frac{\partial p^*(0, 0)}{\partial u_i} \geq -\lambda_i^*$$

Для отрицательного  $t < 0$  имеем:

$$\frac{p^*(te_i, 0) - p^*}{t} \leq -\lambda_i^* \rightarrow \frac{\partial p^*(0, 0)}{\partial u_i} \leq -\lambda_i^*$$

Та же идея может быть использована для установления факта о  $v_i$ .

Результат локальной чувствительности Equation 4 предоставляет способ понять влияние ограничений на оптимальное решение  $x^*$  задачи оптимизации. Если ограничение  $f_i(x^*)$  отрицательно в  $x^*$ , оно не влияет на оптимальное решение, что означает, что небольшие изменения в этом ограничении не изменят оптимальное значение. В этом случае соответствующий оптимальный множитель Лагранжа будет равен нулю, согласно принципу дополнительной нежёсткости.

Однако, если  $f_i(x^*) = 0$ , что означает, что ограничение точно выполняется в оптимуме, то ситуация иная.

Значение  $i$ -го оптимального множителя Лагранжа,  $\lambda_i^*$ , даёт нам понимание того, насколько “чувствительно” или “активно” это ограничение. Малое  $\lambda_i^*$  указывает на то, что небольшие корректировки ограничения не будут значительно влиять на оптимальное значение.

Наоборот, большое  $\lambda_i^*$  подразумевает, что даже незначительные изменения ограничения могут иметь значительное влияние на оптимальное решение.

- Теория оптимизации (MATH4230) курс @ CUHK, professor Tieyong Zeng