Прямо-Двойственные методы в задачах условной оптимизации

в т.ч. применительно к выводу в марковских случайных полях

Александр Темлянцев[1]

[1]факультет Вычислительной Математики и Кибернетики
МГУ

24.11.2010

План работы

1. Прямая и двойственная задача условной оптимизации
   - Базовая конструкция
   - Геометрическая интерпретация
2. Прямо-двойственный алгоритм для задач линейного программирования
   - Общая схема
   - Примеры
3. Прямо-двойственный алгоритм для задачи вывода в MRF
   - Общая схема
   - алгоритм PD-1

Outline

Постановка задачи

$$\begin{cases} f_0(x) \to \min\limits_x \\ f_i(x) \leq 0 \\ h_i(x) = 0 \end{cases}$$

Двойственная функция

$$\mathcal{L}(x, \lambda, \nu) = f_0(x) + \sum_{i=1:m} \lambda_i f_i(x) + \sum_{i=1:p} \nu_i h_i(x)$$

$$g(\lambda, \nu) = \inf_{x \in \Omega} (\mathcal{L}(x, \lambda, \nu))$$

Двойственная функция. Пример

$$\begin{cases} c^T x \to \min\limits_x \\ Ax = b \\ x \geq 0 \end{cases}$$

$$\mathcal{L}(x, \lambda, \nu) = c^T x - \lambda^T (Ax - b) - \sum_{i=1:p} \nu_i x = \lambda^T b + (c^T + \lambda^T A - \nu^T) x \quad (1)$$

$$g(\lambda, \nu) = \begin{cases} \lambda^T b, & \text{если } c^T + \lambda^T A - \nu^T = 0; \\ -\infty, & \text{иначе.} \end{cases}$$

Двойственная задача

$$\begin{cases} d(\lambda, \nu) \to \max\limits_{\lambda,\nu} \\ \nu \geq 0 \end{cases}$$

Двойственная задача. Пример

$$\begin{cases} \lambda^T b \to \max\limits_{\lambda,\nu} \\ c^T + \lambda^T A - \nu^T = 0 \\ \nu \geq 0 \end{cases}$$

$$\begin{cases} \lambda^T b \to \max\limits_{\lambda} \\ A^T \lambda \leq c \end{cases}$$

Outline

Геометрическая постановка

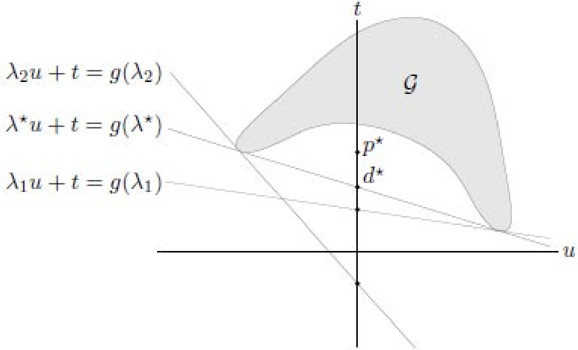
$$\mathcal{G} = \{(u, v, t) \in \mathbf{R}^m \times \mathbf{R}^p \times \mathbf{R} | u_i = f_i(x), v_j = h_j(x), t = f_0(x), x \in \Omega\}$$

$$p^* = inf\{t | (\exists u \preceq 0)((u, 0, t) \in \mathcal{G})\}$$

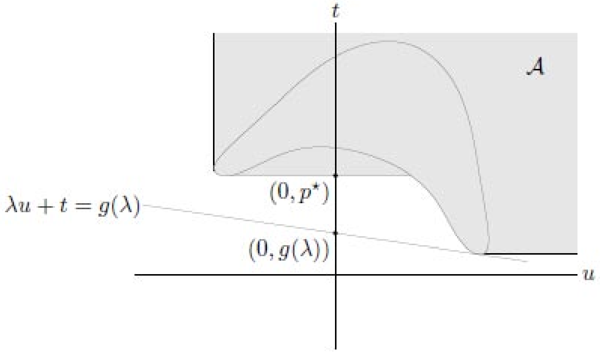
$$g(\lambda, \nu) = inf\{(\lambda, \nu, 1)^T (u, v, t) | (u, v, t) \in \mathcal{G}\}$$

$$\mathcal{A} = \mathcal{G} + \mathbf{R}_+^m \times \mathbf{0}^p \times \mathbf{R}_+\}$$

Интерпретация слабой двойственности



Интерпретация duality gap



$$g(\lambda, \nu) = inf\{(\lambda, \nu, 1)^T(u, v, t) | (u, v, t) \in \mathcal{A}\}(0, 0, p^*) \in \mathcal{A}; \quad (2)$$

$$p^* = (\lambda, \nu, 1)^T(0, 0, p^*) \geq g(\lambda, \nu); \quad (3)$$

Сильная двойственность в выпуклых задачах оптимизации

Условия Слейтера: Пусть функции $f_i(x), i = 0 : m$ выпуклы и

$$(\exists x \in int(\Omega))(f_i(x) < 0, i = \overline{1..m})$$

тогда задача

$$\begin{cases} f_0(x) \\ f_i(x) \leq 0 \\ Ax = b \end{cases}$$

обладает свойством сильной двойственности

Сильная двойственность в выпуклых задачах оптимизации

Условия Слейтера: Пусть функции $f_i(x), i = 0 : m$ аффинные и

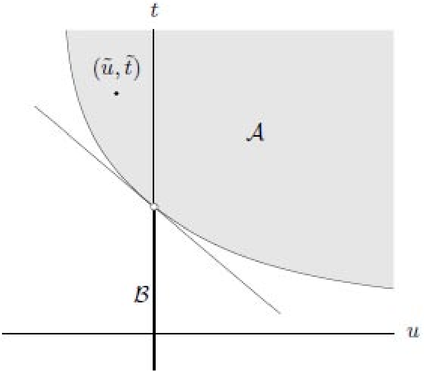
$$(\exists x \in int(\Omega))(f_i(x) \leq 0, i = \overline{1..m})$$

тогда задача

$$\begin{cases} f_0(x) \\ f_i(x) \leq 0 \\ Ax = b \end{cases}$$

обладает свойством сильной двойственности

Сильная двойственность в выпуклых задачах оптимизации



Условия дополняющей нежесткости

Пусть $p^* = d^*$, тогда

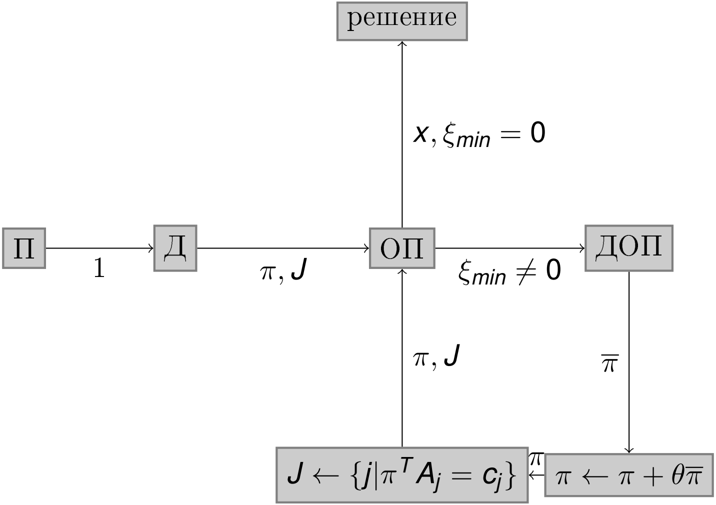
$$f_0(x^*) = d(\lambda^*, \nu^*) \leq \mathcal{L}(x^*, \lambda^*, \nu^*) \leq f_0(x^*)$$

$$f_0(x^*) = f_0(x^*) + \sum_{i=1:m} \lambda_i^* f_i(x^*)$$

$$\sum_{i=1:m} \lambda_i^* f_i(x^*) = 0$$

Outline

Общая схема

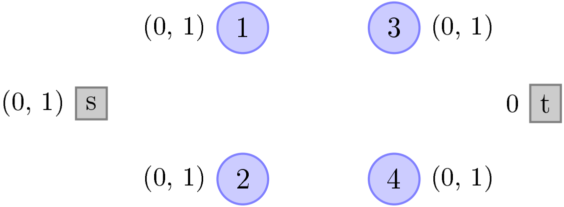

решение
$x, \xi_{min} = 0$
П
1
Д
π, J
ОП
$\xi_{min} \neq 0$
ДОП
π, J
$\overline{\pi}$
$J \leftarrow \{j | \pi^T A_j = c_j\}$
π
$\pi \leftarrow \pi + \theta\overline{\pi}$

Outline

Задача о кратчайшем пути в графе

$$\theta = \max_{\substack{(i,j)\in E: \\ \overline{\pi_i} \neq \overline{\pi_j}}} (c_{i,j} - (\pi_i - \pi_j)); \pi \leftarrow \pi + \theta * \overline{\pi}$$

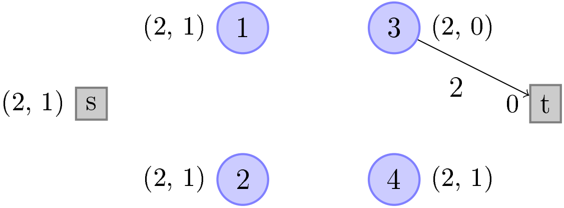
$$\pi = 0; J = \emptyset;$$



Задача о кратчайшем пути в графе

$$\theta = \max_{\substack{(i,j)\in E: \\ \overline{\pi_i}\neq\overline{\pi_j}}} (c_{i,j} - (\pi_i - \pi_j)); \pi \leftarrow \pi + \theta * \overline{\pi}$$

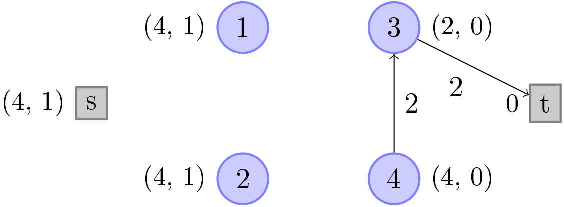
$$\pi_3 - \pi_t = 2 \Rightarrow (3, t) \in J$$



Задача о кратчайшем пути в графе

$$\theta = \max_{\substack{(i,j)\in E:\\ \overline{\pi_i}\neq\overline{\pi_j}}} (c_{i,j} - (\pi_i - \pi_j)); \pi \leftarrow \pi + \theta * \overline{\pi}$$

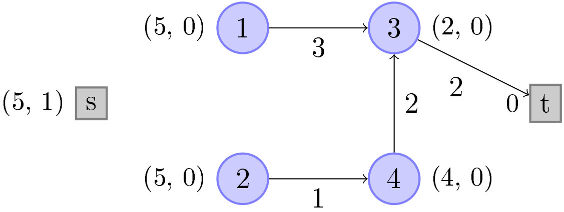
$$\pi_4 - \pi_3 = 2 \Rightarrow (4,3) \in J$$



Задача о кратчайшем пути в графе

$$\theta = \max_{\substack{(i,j)\in E:\\ \overline{\pi_i}\neq\overline{\pi_j}}} (c_{i,j} - (\pi_i - \pi_j)); \pi \leftarrow \pi + \theta * \overline{\pi}$$

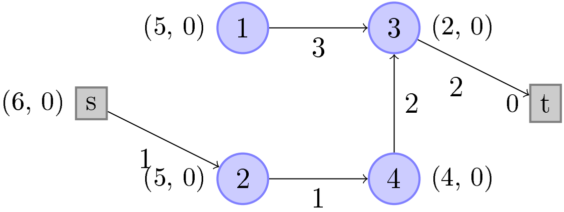
$$\pi_2 - \pi_4 = 1 \Rightarrow (2,4) \in J$$



Задача о кратчайшем пути в графе

$$\theta = \max_{\substack{(i,j)\in E:\\ \overline{\pi_i}\neq\overline{\pi_j}}} (c_{i,j} - (\pi_i - \pi_j)); \pi \leftarrow \pi + \theta * \overline{\pi}$$

$$\pi_2 - \pi_4 = 1 \Rightarrow (2,4) \in J$$



Outline

1. Прямая и двойственная задача условной оптимизации
   - Базовая конструкция
   - Геометрическая интерпретация
2. Прямо-двойственный алгоритм для задач линейного программирования
   - Общая схема
   - Примеры
3. Прямо-двойственный алгоритм для задачи вывода в MRF
   - Общая схема
   - алгоритм PD-1

Постановка

$$
\begin{cases}
\displaystyle\sum_{p\in V}\sum_{a\in L} c_p^a x_p^a + \sum_{(p,q)\in E} \omega_{p,q} \sum_{a,b\in L} d^{a,b} x_{p,q}^{a,b} \to \min_x \\
\displaystyle\sum_{a\in L} x_p^a = 1 \\
\displaystyle\sum_{b\in L} x_{p,q}^{a,b} = x_p^a \\
\displaystyle\sum_{a\in L} x_{p,q}^{a,b} = x_q^b \\
x_p^a \geq 0 \\
x_{p,q}^{a,b} \geq 0
\end{cases}
$$

Постановка

$$\sum_{p\in V}\sum_{a\in L} c_p^a x_p^a + \sum_{(p,q)\in E} \omega_{p,q} \sum_{a,b\in L} d^{a,b} x_{p,q}^{a,b} \to \min_x$$

$$\begin{cases} \displaystyle\sum_{p\in V} y_p \to \max_y \\ \displaystyle -y_p + c_p^a + \sum_q y_{p,q}^a = \lambda_p \\ -y_{p,q}^a - y_{q,p}^b + w_{p,q} d^{a,b} = \lambda_{p,q} \\ \lambda_p^a, \lambda_{p,q}^{a,b} \geq 0 \end{cases}$$

Постановка

$$\sum_{p\in V}\sum_{a\in L} c_p^a x_p^a + \sum_{(p,q)\in E} \omega_{p,q} \sum_{a,b\in L} d^{a,b} x_{p,q}^{a,b} \to \min_x$$

$$\begin{cases} \displaystyle\sum_{p\in V} y_p \to \max_y \\ y_p \le c_p^a + \displaystyle\sum_q y_{p,q}^a = h_p^a \\ y_{p,q}^a + y_{q,p}^b \le w_{p,q} d^{a,b} \end{cases}$$

Общая схема

$$c^T x \le (1+\varepsilon) b^T y \Rightarrow c^T x \le (1+\varepsilon) c^T x^*$$

$$(1+\varepsilon)\sum_i a_{i,j} y_i \ge c_j \Rightarrow$$

$$(1+\varepsilon)\sum_j \sum_i a_{i,j} x_j y_i \ge c_j x_j \Rightarrow$$

$$(1+\varepsilon)\sum_i b_i y_i \ge c_j x_j \quad (4)$$

Общая схема

$$\begin{cases} x_{p,a} > 0 \Rightarrow y_p \geq c_p^a / f_1 + \sum\limits_{q \sim p} y_{p,q}^a \\ x_{p,a}^{q,b} > 0 \Rightarrow y_{p,q}^a + y_{q,p}^b \geq \omega_{p,q} d^{a,b} / f_2 \end{cases}$$

$$\begin{cases} y_p \geq c_p^{x_p} / f_1 + \sum\limits_{q \sim p} y_{p,q}^{x_p} \\ y_{p,q}^{x_p} + y_{q,p}^{x_q} \geq \omega_{p,q} d^{x_p, x_q} / f_2 \end{cases}$$

Общая схема

$$\begin{cases} y_p \geq c_p^{x_p}/f_1 + \sum_{q \sim p} y_{p,q}^{x_p} \\ y_{p,q}^{x_p} + y_{q,p}^{x_q} \geq \omega_{p,q} d^{x_p,x_q}/f_2 \end{cases}$$

$$y_p = \min_a h_p^a \tag{4}$$

$$h_p^{x_p} = \min_a h_p^a \tag{5}$$

$$y_{p,q}^{x_p} + y_{q,p}^{x_q} \geq \omega_{p,q} d^{x_p,x_q}/f_2 \tag{6}$$

$$y_{p,q}^a \leq \omega_{p,q} d_{min}/2 \tag{7}$$

Outline

пересчет двойственных переменных

пересчет базовых переменных