

1 Problem mnożników Lagrange'a

$$\begin{aligned} z^* &= \min \mathbf{c}^T \mathbf{x} \\ \mathbf{Ax} &\leq \mathbf{b} \\ \mathbf{x} &\in \mathbb{X} \end{aligned}$$

$$\begin{aligned} L(\boldsymbol{\mu}) &= \min \mathbf{c}^T \mathbf{x} + \boldsymbol{\mu}^T (\mathbf{Ax} - \mathbf{b}) \\ \mathbf{x} &\in \mathbb{X} \end{aligned}$$

Problem mnożników Lagrange'a (najlepsze dolne ograniczenie)

$$L^* = \max_{\boldsymbol{\mu} \geq \mathbf{0}} L(\boldsymbol{\mu}).$$

W przypadku $\mathbf{Ax} = \mathbf{b}$ problem mnożników Lagrange'a ma postać

$$L^* = \max_{\boldsymbol{\mu}} L(\boldsymbol{\mu}).$$

Maksymalizujemy $\boldsymbol{\mu}$ minimalizując $\mathbf{x}$.

Problem najkrótszej ścieżki z ograniczeniem czasowym (Constrained Shortest Path) CSP

$$\begin{aligned} z^* &= \min \sum_{(i,j) \in A} c_{ij} x_{ij} \\ \sum_{j \in S(i)} x_{ij} - \sum_{j \in P(i)} x_{ji} &= \begin{cases} 1 & i = s, \\ -1 & i = t, \\ 0 & \text{w przeciwnym przypadku,} \end{cases} \\ \sum_{(i,j) \in A} t_{ij} x_{ij} &\leq T \\ x_{ij} &\in \{0, 1\} \quad (i, j) \in A. \end{aligned}$$

Relaksacja CSP

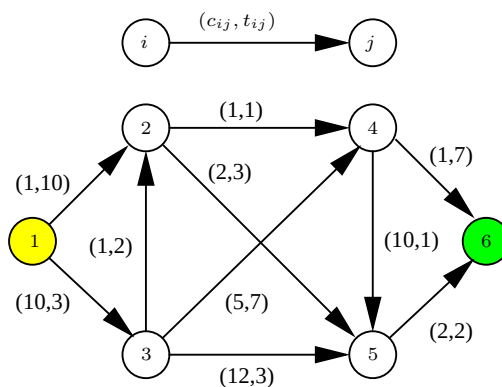
$$\begin{aligned} L(\mu) &= \min \sum_{(i,j) \in A} c_{ij} x_{ij} + \mu \left(\sum_{(i,j) \in A} t_{ij} x_{ij} - T \right) \\ \sum_{j \in S(i)} x_{ij} - \sum_{j \in P(i)} x_{ji} &= \begin{cases} 1 & i = s, \\ -1 & i = t, \\ 0 & \text{w przeciwnym przypadku,} \end{cases} \\ x_{ij} &\in \{0, 1\} \quad (i, j) \in A. \end{aligned}$$

Inne sformułowanie (ścieżkowe)

$$L(\mu) = \min_{p \in \mathcal{P}_{st}} \sum_{(i,j) \in p} c_{ij} + \mu \left(\sum_{(i,j) \in p} t_{ij} - T \right) = \min_{p \in \mathcal{P}_{st}} c_p + \mu(t_p - T)$$

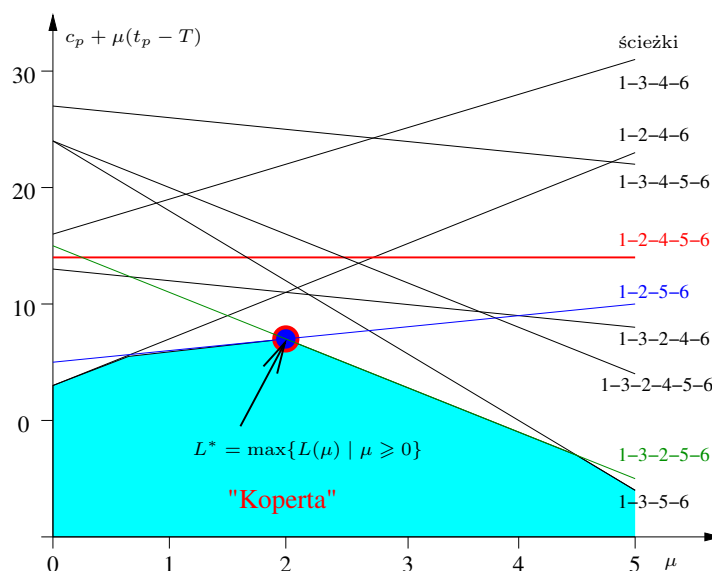
gdzie $\mathcal{P}_{st}$ jest zbiorem ścieżek od s do t .

Należy znaleźć najtańszą ścieżkę od wierzchołka 1 do 6, której czas przejazdu nie przekracza 14 ($T = 14$).



Jaka jest natura problemu mnożników Lagrange'a?

ścieżka p	koszt c_p	czas t_p	$c_p + \mu(t_p - T)$
1-2-4-6	3	18	$3 + 4\mu$
1-2-5-6	5	15	$5 + \mu$
1-2-4-5-6	14	14	14
1-3-2-4-6	13	13	$13 - \mu$
1-3-2-5-6	15	10	$15 - 4\mu$
$\vdots$	$\vdots$	$\vdots$	$\vdots$



Rozważmy uogólnienie problemu

$$\begin{aligned} \min \mathbf{c}^T \mathbf{x} \\ \mathbf{Ax} = \mathbf{b} \\ \mathbf{x} \in \mathbb{X} \end{aligned}$$

gdzie $\mathbb{X}$ jest zbiorem skończonym $\mathbb{X} = \{\mathbf{x}^1, \mathbf{x}^2, \dots, \mathbf{x}^K\}$. Dokonując relaksacji Lagrange'a otrzymujemy

$$L(\boldsymbol{\mu}) = \min_{\mathbf{x} \in \mathbb{X}} \mathbf{c}^T \mathbf{x} + \boldsymbol{\mu}^T (\mathbf{Ax} - \mathbf{b})$$

Z definicji

$$L(\boldsymbol{\mu}) \leq \mathbf{c}^T \mathbf{x}^k + \boldsymbol{\mu}^T (\mathbf{Ax}^k - \mathbf{b}) \text{ dla } k = 1, \dots, K.$$

$\mathbf{c}^T \mathbf{x}^k + \boldsymbol{\mu}^T (\mathbf{Ax}^k - \mathbf{b})$ jest hiperpłaszczyzną.

Hiperpłaszczyzny $\mathbf{c}^T \mathbf{x}^k + \boldsymbol{\mu}^T (\mathbf{Ax}^k - \mathbf{b})$ dla $k = 1, \dots, K$ tworzą "kopertę".

W problemie mnożników Lagrange'a chcemy wyznaczyć najwyższy punkt tej "koperty". Co sprowadza się do rozwiązania następującego problemu

$$\begin{aligned} \max w \\ w \leq \mathbf{c}^T \mathbf{x}^k + \boldsymbol{\mu}^T (\mathbf{Ax}^k - \mathbf{b}) \text{ dla } k = 1, \dots, K \\ \boldsymbol{\mu} \text{ dowolna.} \end{aligned}$$

Rozwiązanie powyższego problemu sprowadza się do rozwiązania szeregu zadań LP.

Metoda gradientów

Niech funkcja celu f będzie różniczkowalna i wklęsła. Gradient funkcji

$$\nabla f(\mathbf{x}^*) = \left(\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n} \right),$$

o ile nie jest równy zero, wskazuje kierunek wzrostu wartości funkcji w pobliżu punktu, w którym jest wyliczony. Jeżeli wybierzemy taki kierunek $\mathbf{d}$, że $\nabla f(\mathbf{x})^T \mathbf{d} > 0$, wówczas niewielkie przesunięcie o θ ,

$$\mathbf{x} + \mathbf{d}\theta$$

w tym kierunku spowoduje wzrost wartości funkcji

$$f(\mathbf{x} + \mathbf{d}\theta) \geq f(\mathbf{x}).$$

Kierunek gradientu jest kierunkiem "w górę stoku".

$$L(\boldsymbol{\mu}) = \min_{\mathbf{x} \in \mathbb{X}} \mathbf{c}^T \mathbf{x} + \boldsymbol{\mu}^T (\mathbf{Ax} - \mathbf{b})$$

- Startujemy z $\boldsymbol{\mu}^0$.
- dla $k := 0, 1, \dots$

Wyznaczamy $\mathbf{x}^k$ rozwiązując $L(\boldsymbol{\mu}^k)$

$$\boldsymbol{\mu}^{k+1} := \boldsymbol{\mu}^k + \theta^k (\mathbf{Ax}^k - \mathbf{b}).$$

Problem z dobozem θ^k - duże, małe

Założmy, że $\mathbf{x}^k$ rozwiązuje $L(\boldsymbol{\mu})$ dla $\boldsymbol{\mu} = \boldsymbol{\mu}^k$. Wówczas jest rozwiązaniem w otoczeniu $\boldsymbol{\mu}$. Aproxymujemy liniowo $L(\boldsymbol{\mu})$ za pomocą

$$r(\boldsymbol{\mu}) = \mathbf{c}^T \mathbf{x}^k + \boldsymbol{\mu}^T (\mathbf{A}\mathbf{x}^k - \mathbf{b}).$$

Założmy, że znamy L^* , wówczas możemy tak wykonać ruch (dobrać θ^k), że osiągniemy L^* .

$$\begin{aligned} r(\boldsymbol{\mu}^{k+1}) &= \mathbf{c}^T \mathbf{x}^k + \boldsymbol{\mu}^{k+1T} (\mathbf{A}\mathbf{x}^k - \mathbf{b}) = L^* \\ \boldsymbol{\mu}^{k+1} &= \boldsymbol{\mu}^k + \theta^k (\mathbf{A}\mathbf{x}^k - \mathbf{b}) \\ r(\boldsymbol{\mu}^{k+1}) &= \mathbf{c}^T \mathbf{x}^k + (\boldsymbol{\mu}^k + \theta^k (\mathbf{A}\mathbf{x}^k - \mathbf{b}))^T (\mathbf{A}\mathbf{x}^k - \mathbf{b}) = L^*. \end{aligned}$$

Z równań wyznaczamy θ^k

$$\theta^k = \frac{L^* - L(\boldsymbol{\mu}^k)}{\|\mathbf{A}\mathbf{x}^k - \mathbf{b}\|^2}.$$

Ponieważ L^* nie jest znane, możemy użyć górnego ograniczenia UB na wartość funkcji celu z^*

$$\theta^k = \frac{\lambda_k (UB - L(\boldsymbol{\mu}^k))}{\|\mathbf{A}\mathbf{x}^k - \mathbf{b}\|^2},$$

gdzie $\lambda_k \in [0, 2]$.

2 Związek relaksacji Lagrange'a z LP

Twierdzenie 1. *Jeżeli zastosujemy relaksację Lagrange'a do następującego zadania LP*

$$\begin{aligned} \min \quad & \mathbf{c}^T \mathbf{x} \\ \text{s.t.} \quad & \mathbf{A}\mathbf{x} = \mathbf{b} \\ & \mathbf{D}\mathbf{x} \leq \mathbf{q} \\ & \mathbf{x} \geq \mathbf{0} \end{aligned} \tag{1}$$

dokonując relaksacji ograniczeń $\mathbf{A}\mathbf{x} = \mathbf{b}$, wówczas optymalna wartość funkcji Lagrange'a L^ w problemie mnożników równa jest optymalnej wartości funkcji celu problemu (1).*

Dowód. (zob. dodatek A). □

- Relaksacja Lagrange'a może być alternatywną sposobem rozwiązywania problemów LP.
- W niektórych przypadkach zastosowanie relaksacji Lagrange'a do LP może prowadzić do łatwiejszego rozwiązania problemu.

- Zastosowanie relaksacji może prowadzić do podziału problemu na mniejsze podproblemy (dekompozycje).

Rozważmy problem optymalizacji dyskretnej.

$$\begin{aligned} z^* &= \min \mathbf{c}^T \mathbf{x} \\ \mathbf{A}\mathbf{x} &= \mathbf{b} \\ \mathbf{x} &\in \mathbb{X} \end{aligned} \quad (2)$$

gdzie $\mathbb{X}$ jest zbiorem dyskretnym $\mathbb{X} = \{\mathbf{x} \mid \mathbf{D}\mathbf{x} \leq \mathbf{q}, \mathbf{x} \geq \mathbf{0} \text{ i całkowity}\}$, $\mathbf{D}$ jest macierzą, której elementy są całkowite, a $\mathbf{q}$ jest wektorem z całkowitymi elementami.

$$\begin{aligned} z^* &= \min \mathbf{c}^T \mathbf{x} \\ \mathbf{A}\mathbf{x} &= \mathbf{b} \\ \mathbf{D}\mathbf{x} &\leq \mathbf{q} \\ \mathbf{x} &\geq \mathbf{0} \text{ i całkowity} \end{aligned} \quad (3)$$

Niech

$$\begin{aligned} z^o &= \min \mathbf{c}^T \mathbf{x} \\ \mathbf{A}\mathbf{x} &= \mathbf{b} \\ \mathbf{D}\mathbf{x} &\leq \mathbf{q} \\ \mathbf{x} &\geq \mathbf{0} \end{aligned}$$

będzie liniową relaksacją problemu (3). Wiemy, że

$$z^o \leq z^*.$$

oraz

$$L^* \leq z^*.$$

Pokażemy, że wartość L^* jest najlepszym dolnym ograniczeniem na z^* , czyli

$$\mathbf{z}^o \leq \mathbf{L}^*.$$

Założmy, że zbiór $\mathbb{X} = \{\mathbf{x}^1, \mathbf{x}^2, \dots, \mathbf{x}^K\}$ jest skończony.

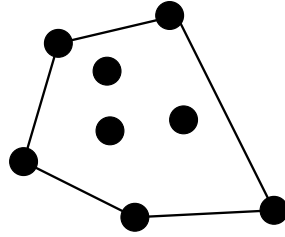
Rozwiązanie $\mathbf{x}$ jest wypukłą kombinacją rozwiązań $\mathbf{x}^1, \mathbf{x}^2, \dots, \mathbf{x}^K$, jeżeli

$$\begin{aligned} \mathbf{x} &= \sum_{k=1}^K \lambda_k \mathbf{x}^k \\ \sum_{k=1}^K \lambda_k &= 1 \\ \lambda_k &\geq 0 \quad k = 1, \dots, K. \end{aligned}$$

$\mathcal{H}(\mathbb{X})$ jest otoczką wypukłą $\mathbb{X}$, jeżeli

$$\mathcal{H}(\mathbb{X}) = \{\mathbf{x} \mid \mathbf{x} \text{ jest wypukłą kombinacją punktów } \mathbb{X}\}.$$

Otoczka wypukła w $\mathbb{R}^2$



Otoczka wypukła $\mathcal{H}(\mathbb{X})$ jest najmniejszym wielościannem wypukłym zawierającym wszystkie punkty zbioru $\mathbb{X}$.

Własność 1.

- a) $\mathcal{H}(\mathbb{X})$ jest wielościannem, który może być wyrażony za pomocą $\mathcal{H}(\mathbb{X}) = \{\mathbf{x} \mid \mathbf{C}\mathbf{x} \leq \mathbf{d}\}$, gdzie $\mathbf{C}$ jest macierzą, a $\mathbf{d}$ jest wektorem.
- b) Każdy punkt ekstremalny (wierzchołek) $\mathcal{H}(\mathbb{X})$ należy do $\mathbb{X}$.
Jeżeli rozwiązujemy zadanie $\min\{\mathbf{c}^T \mathbf{x} \mid \mathbf{x} \in \mathcal{H}(\mathbb{X})\}$, to rozwiązanie optymalne należy do $\mathbb{X}$.
- c) Jeżeli $\mathbb{X} \subseteq \mathbb{Y} = \{\mathbf{x} \mid \mathbf{C}\mathbf{x} \leq \mathbf{d}, \mathbf{x} \geq \mathbf{0}\}$, to $\mathcal{H}(\mathbb{X}) \subseteq \mathbb{Y}$.

Dowód.

Ad. a) Jest to znany fakt z algebry liniowej.

Ad. b) Pierwsza część wynika z faktu, że punkty należące do $\mathcal{H}(\mathbb{X})$ a nie należące do $\mathbb{X}$ są wypukłą kombinacją z dodatnimi wagami co najmniej dwóch punktów z $\mathbb{X}$. Zatem nie są ekstremalne (nie są wierzchołkami $\mathbb{X}$).

Druga część wynika z faktu, że optimum w zadaniach **LP** jest osiągane zawsze w co najmniej jednym punkcie ekstremalnym (wierzchołku).

Ad. c) Ponieważ każdy element zbioru $\mathbb{X}$ należy do wypukłego zbioru $\mathbb{Y}$. Stąd każda wypukła kombinacja elementów zbioru $\mathbb{X}$ ($\mathcal{H}(\mathbb{X})$) również należy do $\mathbb{Y}$. $\square$

Twierdzenie 2. Optymalna wartość funkcji Lagrange'a L^* w problemie mnożników Lagrange'a jest równa optymalnej wartości funkcji celu następującego zadania **LP**

$$\begin{aligned} z_{LP} &= \min \mathbf{c}^T \mathbf{x} \\ \mathbf{A}\mathbf{x} &= \mathbf{b} \\ \mathbf{x} &\in \mathcal{H}(\mathbb{X}) \end{aligned} \tag{4}$$

Dowód. Rozważmy relaksację Lagrange'a problemu (2) dla pewnego $\boldsymbol{\mu}$.

$$L(\boldsymbol{\mu}) = \min_{\mathbf{x} \in \mathbb{X}} \mathbf{c}^T \mathbf{x} + \boldsymbol{\mu}^T (\mathbf{A}\mathbf{x} - \mathbf{b})$$

która równoważna jest (własność 1 b) następującemu zadaniu (może ono być wyrażone jako **LP** – własność 1 a)

$$L(\boldsymbol{\mu}) = \min_{\mathbf{x} \in \mathcal{H}(\mathbb{X})} \mathbf{c}^T \mathbf{x} + \boldsymbol{\mu}^T (\mathbf{A}\mathbf{x} - \mathbf{b}) \quad (5)$$

Problem (5) jest relaksacją Lagrange’a (relaksacja ograniczeń $\mathbf{A}\mathbf{x} = \mathbf{b}$) problemu (4). Zatem z twierdzenia 1 optymalna wartość L^* problemu mnożników jest równa optymalnej wartości z_{LP} problemu (4). $\square$

Twierdzenie 3. *Jeżeli zastosujemy relaksację Lagrange’a, i rozwiążemy problem mnożników, dla problemu*

$$\begin{aligned} & \min \mathbf{c}^T \mathbf{x} \\ & \mathbf{A}\mathbf{x} = \mathbf{b} \\ & \mathbf{D}\mathbf{x} \leq \mathbf{q} \\ & \mathbf{x} \geq \mathbf{0} \text{ i całkowity.} \end{aligned}$$

Wówczas $z^o \leq L^*$, gdzie

$$\begin{aligned} z^o &= \min \mathbf{c}^T \mathbf{x} \\ & \mathbf{A}\mathbf{x} = \mathbf{b} \\ & \mathbf{D}\mathbf{x} \leq \mathbf{q} \\ & \mathbf{x} \geq \mathbf{0}. \end{aligned}$$

Dowód. Z twierdzenia 2 mamy

$$\begin{aligned} L^* &= \min \mathbf{c}^T \mathbf{x} \\ & \mathbf{A}\mathbf{x} = \mathbf{b} \\ & \mathbf{x} \in \mathcal{H}(\mathbb{X}). \end{aligned}$$

Z własności 1 c dostajemy

$$\mathcal{H}(\mathbb{X}) = \mathcal{H}(\{\mathbf{x} \mid \mathbf{D}\mathbf{x} \leq \mathbf{q}, \mathbf{x} \geq \mathbf{0} \text{ i całkowity}\}) \subseteq \{\mathbf{x} \mid \mathbf{D}\mathbf{x} \leq \mathbf{q}, \mathbf{x} \geq \mathbf{0}\}.$$

Zatem $z^o \leq L^*$. Powyższe rozumowanie ilustruje rysunek 1. $\square$

Kiedy $z^o = L^*$?

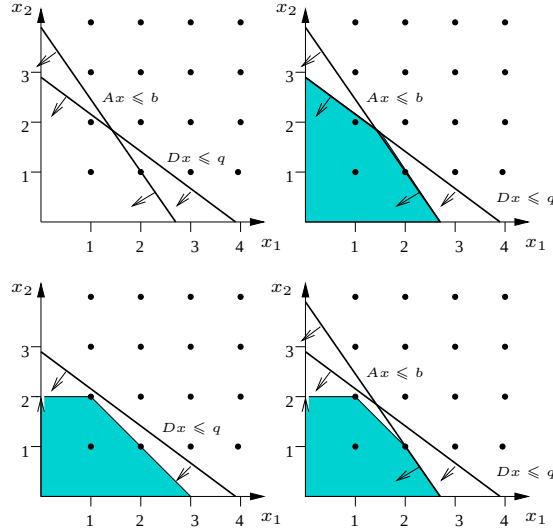
Niech $\mathbb{X} = \{\mathbf{x} \mid \mathbf{D}\mathbf{x} \leq \mathbf{q}, \mathbf{x} \geq \mathbf{0} \text{ i całkowity}\}$.

Definicja 1. $\mathbb{X}$ ma własność całkowitoliczbowości jeżeli następujące zadanie **LP** ma rozwiązania całkowitoliczbowe dla każdego $\mathbf{d}$

$$\begin{aligned} & \min \mathbf{d}^T \mathbf{x} \\ & \mathbf{D}\mathbf{x} \leq \mathbf{q} \\ & \mathbf{x} \geq \mathbf{0}. \end{aligned}$$

$$\{x \mid Ax \leq b, Dx \leq q, x \geq 0 \text{ i całkowity}\}$$

$$\{x \mid Ax \leq b, Dx \leq q, x \geq 0\}$$



Otoczka wypukła $\mathcal{H}(\mathbb{X})$

$$\{x \mid Ax \leq b, x \in \mathcal{H}(\mathbb{X})\}$$

Rysunek 1: Ilustracja dowodu twierdzenia 3.

$\mathbb{X}$ ma **własność całkowitoliczbowości**, jeśli np. D jest całkowicie uni-modularna i q jest całkowity.

Fakt 1. Jeżeli $\mathbb{X}$ ma własność całkowitoliczbowości, to

$$\begin{aligned} \min\{d^T x \mid Dx \leq q, x \geq 0\} &= \min\{d^T x \mid x \in \mathbb{X}\} \\ &= \min\{d^T x \mid x \in \mathcal{H}(\mathbb{X})\}. \end{aligned}$$

Twierdzenie 4. Jeżeli podproblem Lagrange'a po relaksacji problemu

$$\begin{aligned} \min c^T x \\ Ax &= b \\ Dx &\leq q \\ x &\geq 0 \text{ i całkowity.} \end{aligned}$$

ma własność całkowitoliczbowości, to $L^* = z^o$, gdzie

$$\begin{aligned} z^o &= \min c^T x \\ Ax &= b \\ Dx &\leq q \\ x &\geq 0. \end{aligned}$$

Dowód. Z faktu 1. □

Uwagi na temat treści wykładu

Treść wykładu w całości została przygotowana na podstawie książki [1].

Literatura

- [1] Ravindra K. Ahuja, Thomas L. Magnanti, James B. Orlin. *Network Flows: theory, algorithms, and applications*. Prentice Hall, Englewood Cliffs, New Jersey, 1993.

A Dodatek

Dowód twierdzenia 1. Skorzystamy z dualności zadań **LP**. Załóżmy, że $\mathbf{x}^*$ jest rozwiązaniem optymalnym problemu

$$\begin{aligned} \min \mathbf{c}^T \mathbf{x} \\ \mathbf{Ax} = \mathbf{b} \\ \mathbf{Dx} \leq \mathbf{q} \\ \mathbf{x} \geq \mathbf{0} \end{aligned} \tag{6}$$

a $(\boldsymbol{\pi}^*, \boldsymbol{\gamma}^*)$ jest rozwiązaniem optymalnym odpowiadającego zadania dualnego. Muszą więc spełniać warunki dopuszczalności i warunki komplementarności (twierdzenie o różnicach dopełniających):

$$\begin{aligned} \mathbf{c}^T + \boldsymbol{\pi}^{*T} \mathbf{A} + \boldsymbol{\gamma}^{*T} \mathbf{D} &\geq \mathbf{0} \\ (\mathbf{c}^T + \boldsymbol{\pi}^{*T} \mathbf{A} + \boldsymbol{\gamma}^{*T} \mathbf{D}) \mathbf{x}^* &= 0 \\ \boldsymbol{\pi}^{*T} (\mathbf{Ax}^* - \mathbf{b}) &= 0 \\ \boldsymbol{\gamma}^{*T} (\mathbf{Dx}^* - \mathbf{q}) &= 0 \end{aligned}$$

Niech $L(\boldsymbol{\mu}) = \min\{\mathbf{c}^T \mathbf{x} + \boldsymbol{\mu}^T (\mathbf{Ax} - \mathbf{b}) \mid \mathbf{Dx} \leq \mathbf{q}, \mathbf{x} \geq \mathbf{0}\}$ będzie relaksacją problemu (6). Przyjmijmy za $\boldsymbol{\mu} = \boldsymbol{\pi}^*$. Zatem $L(\boldsymbol{\pi}^*) = \min\{\mathbf{c}^T \mathbf{x} + \boldsymbol{\pi}^{*T} (\mathbf{Ax} - \mathbf{b}) \mid \mathbf{Dx} \leq \mathbf{q}, \mathbf{x} \geq \mathbf{0}\}$. Rozwiązanie $\mathbf{x}^*$ jest rozwiązaniem dopuszczalnym dla tego problemu, ponieważ jest rozwiązaniem dopuszczalnym dla (6). Po przyjęciu $\boldsymbol{\mu} = \boldsymbol{\pi}^*$ warunki komplementarności dla problemu Lagrange'a są takie same jak warunki powyżej. Zatem $\mathbf{x}^*$ rozwiązuje problem Lagrange'a dla $\boldsymbol{\mu} = \boldsymbol{\pi}^*$. Ponieważ $\boldsymbol{\pi}^{*T} (\mathbf{Ax}^* - \mathbf{b}) = 0$, $L(\boldsymbol{\pi}^*) = \mathbf{c}^T \mathbf{x}^*$. Stąd $L^* = L(\boldsymbol{\pi}^*) = \mathbf{c}^T \mathbf{x}^*$. $\square$