

# Dualita pri lineárnom programovaní

Mišo Forišek

<misof@mfnotes.ksp.sk>  
 Department of Computer Science  
 Faculty of Mathematics, Physics, and Informatics  
 Comenius University, Bratislava, Slovakia

júl 2009

Cieľom týchto poznámok je urobiť intuitívne jasno v dualite pri lineárnom programovaní.

## Obsah

<b>1</b>	<b>Dva diskkrétne problémy na úvod</b>	<b>1</b>
1.1	<i>Ešte raz v blede modrej množinovej forme</i>	2
1.2	<i>Prvý príklad duality</i>	2
<b>2</b>	<b>Toky a rezy</b>	<b>3</b>
2.1	<i>Dualita tokov a rezov</i>	4
2.2	<i>Hľadanie toku a rezu ako lineárne programy</i>	4
<b>3</b>	<b>Dualita medzi covering a packing LP</b>	<b>5</b>
3.1	<i>Príklad s cukrárňou</i>	5
3.2	<i>Príklad s príviesmi</i>	6
<b>4</b>	<b>Všeobecná dualita</b>	<b>7</b>
4.1	<i>Weak duality theorem</i>	8
4.2	<i>Strong duality theorem</i>	8
4.3	<i>Complementary slackness</i>	9
4.4	<i>Minimálny rez vracia úder</i>	10
<b>5</b>	<b>Užitočné linky</b>	<b>11</b>

## 1 Dva diskkrétne problémy na úvod

Spoločný setting: Máme  $N$  ľudí a  $M$  druhov jedla. O každom človeku vieme množinu jedál, ktoré je ochotný zjesť.

Problém #1: Cecilka organizuje párty pre všetkých  $N$  ľudí a chceme to mať s čo najmenšou námahou. Preto by chcela vybrať čo najmenej druhov jedla, ktoré pripraví. Samozrejme tak, aby si každý našiel aspoň niečo, čo je ochotný zjesť.

Problém #2: Paulína rada varí, a tak navarila všetkých  $M$  jedál. Teraz by rada pozvala čo najviac spomedzi dotýčnych  $N$  ľudí na návštevu. Nechce ale, aby sa niektorí dvaja pohádali, že obaja chcú to isté jedlo. Aby tomu naisto zabránila, potrebuje, aby všetci pozvaní mali po dvoch disjunktné množiny chutí.

## 1.1 Ešte raz v blede modrej množinovej forme

V oboch prípadoch ide o známe problémy.

Ten Cecilkin sa volá SETCOVER a v množinovej terminológii znie nasledovne: Máme množinu prvkov  $A = \{1, \dots, N\}$  a tiež postupnosť  $B$  tvorenú  $M$  podmnožinami množiny  $A$ . Z postupnosti  $B$  treba vybrať čo najmenej prvkov tak, aby ich zjednotením bola celá množina  $A$ .

Inými slovami, treba z  $B$  vybrať čo najmenej množín tak, aby dokopy „pokryli“ celú  $A$ .

Ten Paulínin sa volá SETPACKING. Tentoraz máme množinu prvkov  $B = \{1, \dots, M\}$  a postupnosť jej podmnožín  $A$ . Z postupnosti  $A$  chceme vybrať čo najviac prvkov tak, aby každé dva boli disjunktné.

Inými slovami, treba z  $A$  vybrať čo najviac množín, ktoré sa naraz „zmestia“ do  $B$ .

Oba tieto problémy, teda presnejšie ich rozhodovacie verzie, sú  $NP$ -úplné.

## 1.2 Prvý príklad duality

Naše dva problémy sú v istom zmysle duálne. Aby sme to lepšie videli, predstavme si maticu  $Z$  rozmerov  $N \times M$ , v ktorej v riadku  $i$  a stĺpci  $j$  je 0 alebo 1 podľa toho, či človek  $i$  zje jedlo  $j$ .

Na tejto matici teraz môžeme naše úlohy preformulovať nasledovne:

C: Vyber čo najmenej stĺpcov tak, aby v každom riadku bola vybratá aspoň jedna 1.

P: Vyber čo najviac riadkov tak, aby v každom stĺpci bola vybratá najviac jedna 1.

Vidíme teda, že zadania sa na seba podobajú. Ale súvisia spolu aj riešenia?

Súvisia. Presnejšie, platí nasledujúce tvrdenie: Nech  $s_1, \dots, s_c$  je nejaký (nie nutne optimálny) výber stĺpcov, ktorý spĺňa podmienku C, a nech  $r_1, \dots, r_p$  je nejaký (opäť, nie nutne optimálny) výber riadkov, ktorý spĺňa podmienku P. Potom  $c \geq p$ .

(To isté slovne: Ľubovoľné riešenie pre pakovanie nám dáva dolný odhad pre minimálne pokrytie. Ľubovoľné riešenie pre pokrytie nám dáva horný odhad pre pakovanie.)

Prečo? Pozrime sa na riadky  $r_1, \dots, r_p$ . Mohol nastať jeden patologický prípad: Niektorý z týchto riadkov môže obsahovať samé 0. Ale ak toto nastalo, tak neexistuje žiadne riešenie problému C – aj keby sme vybrali úplne všetky stĺpce, v tomto riadku 1 nebude obsahovať žiaden z nich.

Môžeme teda predpokladať, že každý z riadkov  $r_1, \dots, r_p$  obsahuje aspoň jednu 1. Čo vieme teraz povedať o riešení problému C? Potrebujeme vybrať takú množinu stĺpcov, aby sme v

každom riadku vybrali aspoň jednu 1. Špeciálne teda musíme vybrať aspoň jednu 1 v každom z riadkov  $r_1, \dots, r_p$ . Lenže keďže tieto riadky sú riešením problému P, ľubovoľný stĺpec má 1 nanajvyš v jednom z nich. Aby sme teda vybrali aspoň jednu 1 v každom z riadkov  $r_1, \dots, r_p$ , musíme vybrať aspoň  $p$  stĺpcov, q.e.d.

(Dôkaz sa samozrejme dal robiť aj opačným smerom. Majme nejaký výber stĺpcov  $s_1, \dots, s_c$  taký, že v každom riadku vyberieme aspoň jednu jednotku. Čo teraz vieme povedať o riešení problému P? Každý riadok, ktorý vyberieme, má 1 v aspoň jednom zo stĺpcov  $s_1, \dots, s_c$ . Ak by sme teda vybrali viac ako  $c$  riadkov, určite by sme (z Dirichletovho princípu) dostali nejaké dva, ktoré majú 1 v tom istom stĺpci – a to nesmieme. Preto môžeme vybrať najviac  $c$  riadkov, q.e.d.)

Poznámka na záver: Ani ak máme zaručené, že má problém C riešenie, nemusí nastať rovnosť medzi veľkosťou minimálneho pokrytia a veľkosťou maximálneho pakovania. Neskôr uvidíme, že rovnosť by v takomto prípade vždy nastala, ak by sme povolili „vyberať riadky/stĺpce čiastočne“, t. j. napríklad zobrať 0.7-krát prvý a druhý a 0.6-krát piaty stĺpec.

## 2 Toky a rezy

Iná známa dualita, dokázaná pánmi Fordom a Fulkersonom, je dualita medzi maximálnym tokom a minimálnym rezom v orientovanom grafe.

Inštanciou pre oba problémy je orientovaný ohodnotený graf s dvoma špeciálnymi vrcholmi  $s$  a  $t$  (source, target). Ohodnotenia hrán nech sú pre jednoduchosť nezáporné celé čísla. Pri toku ich voláme *capacity*, pri reze zväčša *ceny*.

Pri maximálnom toku sa na graf dívame ako na sústavu potrubí (s chlopňami, aby boli jednosmerné), pričom o každom vieme, akú má priepustnosť za sekundu. Cieľom je zistiť priepustnosť siete – koľko najviac vody môžeme každú sekundu pustiť do  $s$  tak, aby všetka stíhala odtekať z  $t$ ?

Formálnejšie, každej hrane  $e$  chceme priradiť nejaký tok  $f_e$ , ktorý bude nezáporný a nanajvyš rovný jej kapacite  $c_e$ . Toto priradenie musí byť také, že v každom vrchole okrem  $s$  a  $t$  spĺňa prvý Kirchhoffov zákon – teda súčet toku na prichádzajúcich hranách musí byť rovný súčtu na odchádzajúcich. Veľkosťou toku je potom rozdiel medzi súčtom odchádzajúceho a prichádzajúceho toku v  $s$ . (Ekvivalentne sa to dá definovať aj v  $t$ .)

Pri minimálnom  $s - t$  reze je našim cieľom „rozstrihnúť“ niektoré hrany tak, aby sa už z  $s$  do  $t$  nedalo dostať. Formálnejšie,  $s - t$  rez je partícia množiny vrcholov na dve podmnožiny  $P_1$  a  $P_2$  také, že  $s \in P_1$  a  $t \in P_2$ . Cena tohto rezu je rovná súčtu cien všetkých hrán, ktoré začínajú v  $P_1$  a končia v  $P_2$ . (Toto sú hrany, ktoré musíme rozstrihnúť, ak chceme zabezpečiť, aby sa z  $P_1$  nedalo dostať do  $P_2$ .)

## 2.1 Dualita tokov a rezov

Pre toky a rezy platí podobná dualita ako v predchádzajúcom príklade: veľkosť ľubovoľného platného toku je menšia alebo rovná ako veľkosť ľubovoľného rezu.

Dôkaz je zjavný: pozrime si ľubovoľný rez  $(P_1, P_2)$ . Z množiny  $P_1$  do množiny  $P_2$  vieme za každú sekundu pretlačiť nanajvýš toľko vody, koľko „pustí“ tento rez. A keďže  $s \in P_1$  a  $t \in P_2$ , tento rez tým pádom obmedzuje množstvo vody, ktoré vieme „pretlačiť“ z  $s$  do  $t$  pri ľubovoľnom toku.

Tentokrát však navyše platí aj tvrdenie o rovnosti: vždy existujú tok a jemu zodpovedajúci rez také, že ich veľkosti sa rovnajú.

Jeden možný dôkaz je založený na zlepšujúcich cestách. Majme ľubovoľný tok. Zostrojíme zvyškovú sieť nasledovne: pre každú hranu  $e = uv$  v pôvodnom grafe: ak je  $f_e < c_e$ , pridáme do zvyškovej siete hranu  $e$ , a tiež ak  $f_e > 0$ , pridáme do zvyškovej siete hranu  $\bar{e} = vu$ .

Ak by v tejto sieti existovala cesta z  $s$  do  $t$  (takúto voláme *zlepšujúca cesta*), znamená to, že pôvodný tok vieme zväčšiť – vieme zobrať jednotkové množstvo vody navyše a to sieťou s pôvodným tokom „pretlačiť“ po práve nájdenej ceste.

Keďže sú kapacity celé čísla, tento postup môžeme opakovať len konečne veľa krát. V konečnom čase teda dostaneme tok, pre ktorý už žiadna zlepšujúca cesta neexistuje. V tomto okamihu nech  $P_1$  je množina tých vrcholov, do ktorých existuje zlepšujúca cesta a  $P_2$  nech je množina zvyšných vrcholov. Takto sme zjavne dostali  $s - t$  rez.

Všimnime si teraz, ako vyzerá náš tok na rozhraní  $P_1$  a  $P_2$ . Nech  $e$  je ľubovoľná hrana z  $P_1$  do  $P_2$ . Potom musí byť  $f_e = c_e$ , inak by sme sa vo zvyškovej sieti vedeli dostať do vrcholu z  $P_2$ . Analogicky, nech  $\bar{e}$  je ľubovoľná hrana z  $P_2$  do  $P_1$ , potom musí byť  $f_{\bar{e}} = 0$ .

Dostávame teda, že celkový tok z  $P_1$  do  $P_2$  je presne rovný cene rezu  $(P_1, P_2)$ . Lenže tento tok z  $P_1$  do  $P_2$  je nutne rovný veľkosti nášho toku – je jedno, či ju meriame pri zdroji alebo hocikde inde. Preto sme našli tok a rez, ktoré majú rovnakú veľkosť, a teda sú oba optimálne.

## 2.2 Hľadanie toku a rezu ako lineárne programy

Úlohu o nájdení maximálneho toku vieme triviálne formulovať ako lineárny program. Budeme mať jednu premennú  $f_e$  pre každú hranu  $e$ . Obmedzenia pre tok na jednotlivých hranách sú zjavne OK. Takisto Kirchhoffov zákon pre vrchol je lineárny – konkrétna kombinácia premenných  $f$  s koeficientami 1 alebo  $-1$  má byť rovná nule. A do tretice, aj funkcia, ktorú maximalizujeme, je takouto kombináciou premenných pre hrany incidentné s vrcholom  $s$ .

Súvis tokov a rezov naznačuje, že aj úloha o minimálnom reze by sa mala dať formulovať ako lineárny program. Ako na to?

Jedna možnosť vyzerá nasledovne: Pre každý vrchol  $x$  budeme mať jednu premennú  $p_x$ , ktorú budeme volať jeho potenciál. Tieto potenciály musia byť z intervalu  $[0, 1]$ , pričom požadujeme  $p_s = 1$  a  $p_t = 0$ . Pre ľubovoľnú hranu  $e = uv$  nazvime hodnotu  $w_e = p_u - p_v$  jej váhou. Cieľom

bude minimalizovať hodnotu  $\sum_e c_e w_e$ . Toto je zjavne nejaká lineárna kombinácia premenných  $p_x$ , takže všetko je v poriadku a máme lineárny program.

Prečo ale tento program zodpovedá minimálnemu rezu? Všimnite si, že síce sme povolili neceločíselné potenciály (podobne ako pri maximálnom toku povoľujeme neceločíselné toky), nikdy sa nám ale neoplatia. Pre ľubovoľné riešenie totiž platí, že ak máme vrchol  $x$  taký, že  $0 < p_x < 1$ , tak aspoň jedna zo zmien „nastav  $p_x$  na 0“ a „nastav  $p_x$  na 1“ nám nezvýši celkovú cenu. Existuje teda optimálne riešenie, v ktorom je každé  $p_x$  rovné 0 alebo 1, a toto riešenie evidentne zodpovedá minimálnemu rezu.

Všimli ste si chybu, ktorú sme úmyselne ponechali v predchádzajúcej úvahe?

Tá je v tom, že tak, ako sme tento LP formulovali, sa nám spokojne môže stať, že niektoré hrany budú mať záporný rozdiel potenciálov. (Pri reze  $(P_1, P_2)$  by šlo o hrany z  $P_2$  do  $P_1$ .) A toto nie je to, čo chceme. My by sme tieto hrany chceli odignorovať (netreba ich rezať), lenže v našom lineárnom programe každá takáto hrana zníži hodnotu funkcie, ktorú minimalizujeme.

Uvažujme napríklad jednoduchý graf len s dvoma vrcholmi  $s$  a  $t$ , a s dvoma hranami  $st$  a  $ts$ , oboma s jednotkovou cenou. Ak by sme zostrojili LP podľa vyššie uvedeného postupu, dostali by sme LP, ktorý by mal optimálne riešenie 0 – pre ľubovoľné prípustné hodnoty premenných  $p_s$  a  $p_t$ .

Ako teda túto situáciu zachrániť? Pre každú hranu  $uv$  označme  $d_{uv}$  hodnotu, ktorou chceme aby prispela do riešenia. Ak je  $p_u - p_v$  kladné, malo by byť  $d_{uv} = p_u - p_v$ , inak by malo byť  $d_{uv} = 0$ . Toto vieme v reči LP zapísať pomocou podmienok  $d_{uv} \geq 0$  a  $d_{uv} \geq p_u - p_v$ . Funkcia, ktorú budeme minimalizovať, sa potom zmení na  $\sum_e d_e \cdot c_e$ .

(Všimnite si, že z nami uvedených podmienok vyplýva len  $d_{uv} \geq \max(0, p_u - p_v)$ , v optimálnom riešení však zjavne bude pre každú hranu platiť rovnosť.)

### 3 Dualita medzi covering a packing LP

Nech  $A$  je matica  $M \times N$  nezáporných reálnych čísel, nech  $b$  je vektor  $N$  nezáporných reálnych čísel, a nech  $c$  je vektor  $M$  nezáporných reálnych čísel.

Covering (pokrývací) LP je LP s premennými  $x = (x_1, \dots, x_N)$  v ktorom je cieľom minimalizovať  $\sum_j b_j x_j$  pri podmienkach  $\forall j : x_j \geq 0$  a  $\forall i : \sum_j A_{i,j} x_j \geq c_i$ .

Packing (pakovací) LP je LP s premennými  $y = (y_1, \dots, y_M)$  v ktorom je cieľom maximalizovať  $\sum_i c_i y_i$  pri podmienkach  $\forall i : y_i \geq 0$  a  $\forall j : \sum_i A_{i,j} y_i \leq b_j$ .

Oba LP si teraz ukážeme na príkladoch

#### 3.1 Príklad s cukrárňou

Cyril sa vybral do cukrárne. Má chuť na tri rôzne dobroty: čokoládu, karamel a cukor. Jeho cieľom je zjesť aspoň  $c_1$  jednotiek čokolády, aspoň  $c_2$  karamelu a aspoň  $c_3$  cukru.

V cukrárni však nepredávajú cukor, ale koláče. Menovite dva druhy: veterníky po  $b_1$  peniažkov a karamelové rezy po  $b_2$  peniažkov. Ak Cyril pekne poprosí, odkroja mu aj časť koláča a naučtujú mu za ňu príslušnú časť ceny.

Matica  $A$  bude popisovať zloženie koláčov. Koláč  $j$  obsahuje presne  $A_{i,j}$  jednotiek suroviny  $i$ .

Z Cyrilovho pohľadu dostávame nasledujúci lineárny program: Cieľom je minimalizovať hodnotu  $b_1x_1 + b_2x_2$ , t. j. sumu, ktorú za koláče zaplatí. Pre každú dobrotu dostávame jednu podmienku, ktorá hovorí, že kúpené koláče musia mať dotyčnej dobroty dosť. Teda napr. pre cukor dostávame podmienku  $A_{3,1}x_1 + A_{3,2}x_2 \geq c_3$ .

Toto je covering LP. Cyril musí vybrať koláče, ktoré kúpi, tak, aby pokryl všetky svoje potreby. Všimnite si, že covering LP je zovšeobecnením nášho úplne prvého problému SETCOVER. Cyril musí v matici  $A$  nastaviť váhy stĺpcov tak, aby v každom riadku dostal dostatočne veľký váhovaný súčet.

Pozrime sa teraz na situáciu v cukrárni z opačného uhlu pohľadu. Cukrár zistil, že nemá žiadne suroviny na to, aby Cyrilovi napiekol koláče, a tak odišiel do obchodu. Obchodník Pavol vie, že cukrár ide nakupovať pre Cyrila, pozná teda tiež hodnoty  $A$ ,  $b$  aj  $c$ . A keďže obchodník je líška mazaná, rozhodol sa upraviť ceny surovín (t. j. čokolády, karamelu a cukru) tak, aby jeho zaručený zisk bol čo najvyšší.

Pavlove premenné  $y_1$ ,  $y_2$ ,  $y_3$  teda predstavujú cenu jednotky čokolády, karamelu a cukru.

Podmienky, ktoré musí dodržať, vyplývajú z cien, za ktoré cukrár predáva hotové koláče. Ak by cukrára suroviny vyšli drahšie, nenakúpil by tu. Teda napr. pre veterník dostávame podmienku  $A_{1,1}y_1 + A_{2,1}y_2 + A_{3,1}y_3 \leq b_1$ .

Pri dodržaní podmienok má Pavol isté, že cukrár nakúpi dosť na to, aby splnil Cyrilovu požiadavku. Pavol má teda istý zisk  $c_1y_1 + c_2y_2 + c_3y_3$ , a jeho cieľom bude tento maximalizovať.

Toto je packing LP, zovšeobecnenie problému SETPACKING. Pavol určuje váhy riadkov matice tak, aby v žiadnom stĺpci neprekročil stanovený limit.

### 3.2 Príklad s príviesmi

Tentoraz začneme packing problémom. Máme firmu, ktorá vyrába tri typy príviesov za auto: obyčajný, predĺžený a krytý. Všetko, čo vyrobí, rovno predáva štátnym inštitúciám za vopred dohodnuté ceny. Zisk je  $c_1$  za každý obyčajný,  $c_2$  za každý predĺžený, a  $c_3$  za každý krytý prívies.

Problémom je, že na výrobu príviesu je treba dva druhy prác: opracovanie dreva a opracovanie kovov. Na obe práce má firma len obmedzenú kapacitu. (A kvôli tomu teda nestíha vyrobiť príviesov nekonečne veľa.) Rôzne príviesy si vyžadujú rôzne veľa práce.

Presnejšie,  $j$ -ty typ príviesu si vyžaduje  $A_{i,j}$  jednotiek  $i$ -teho typu práce. Firma za týždeň stihne spraviť nanajvýš  $b_i$  jednotiek  $i$ -teho typu práce.

Z pohľadu firmy tu máme očividný packing problém: Premenné  $y_i$ , ktoré volíme, sú počty príviesov každého typu, ktoré vyrobíme za týždeň. (Tieto počty môžu byť aj neceločíselné,

jednoducho potom ostane nejaký prívies rozrobený a v ďalšom týždni sa na ňom pokračuje.) Máme obmedzenia na výrobu, a pri ich dodržaní chceme maximalizovať zisk.

Výhodou tejto metafory je očitnosť nasledujúceho kroku. Položme si nasledujúcu otázku: Firma dostala možnosť prenajať si navyše nejaké výrobné haly a pracovníkov, čím jej stúpnu obmedzenia na oba typy práce. Koľko má nanaajvýš byť naša firma ochotná zaplatiť za každú jednotku prvého typu práce navyše? A koľko za jednotku druhého typu?

Tieto hodnoty sa zvyknú označovať tieňové ceny práce (*shadow prices*).

Označme si  $x_1$  a  $x_2$  tieňovú cenu jednotky práce prvého a druhého druhu. Čo o nich vieme povedať?

Keby nám niekto dal  $A_{1,1}$  jednotiek práce prvého a  $A_{2,1}$  jednotiek práce druhého druhu, môžeme si vyrobiť navyše jeden obyčajný prívies, a tým zarobiť  $c_1$ . Skutočná cena tejto práce (všetkej dokopy) je teda aspoň  $c_1$ . (Vieme zarobiť  $c_1$  a možno nejakým lepším spôsobom aj viac.) Dostávame teda nerovnosť  $A_{1,1}x_1 + A_{2,1}x_2 \geq c_1$ . Úvahu samozrejme zopakujeme aj pre ostatné druhy príviesov.

No a už sme takmer doma. Naša firma za týždeň vie spraviť  $b_i$  jednotiek práce  $i$ -teho druhu. Teda všetka práva, ktorú dokopy za týždeň spravíme, má cenu  $b_1x_1 + b_2x_2$ . A samozrejme reálnej cene našej práce zodpovedá najmenšia možná hodnota tohto výrazu pre prípustné hodnoty  $x_1, x_2$ .

Všimnite si, čo sme práve dostali. Rovnako ako v predchádzajúcom príklade sme pre dané dáta sformulovali dva rôzne lineárne programy. Tentokrát sme však ukázali aj súvis medzi nimi: začali sme s packing problémom a ukázali sme, že k nemu duálny covering problém v skutočnosti počíta údaje súvisiace s riešením pôvodného problému.

Ešte stále nevieme presne, aký je medzi našimi dvoma duálnymi LP vzťah, ale už sa tam dostávame.

## 4 Všeobecná dualita

Nech  $A$  je matica  $M \times N$  reálnych čísel, nech  $b$  je vektor  $N$  reálnych čísel, a nech  $c$  je vektor  $M$  reálnych čísel. (Všimnite si, že už nepožadujeme ich nezápornosť.)

Uvažujme LP s premennými  $y = (y_1, \dots, y_M)$  v ktorom je cieľom maximalizovať  $\sum_i c_i y_i$  pri podmienkach  $\forall i : y_i \geq 0$  a  $\forall j : \sum_i A_{i,j} y_i \leq b_j$ .

(Uvedomte si, že toto je normálny tvar pre lineárne programy, hocijaký LP vieme upraviť do tohoto tvaru. Rovnosť nahradíme dvoma nerovnosťami, nerovnosť so zlým znamienkom prenásobíme  $-1$ .

Jeho duálnym lineárnym programom budeme volať LP s premennými  $x = (x_1, \dots, x_N)$  v ktorom je cieľom minimalizovať  $\sum_j b_j x_j$  pri podmienkach  $\forall j : x_j \geq 0$  a  $\forall i : \sum_j A_{i,j} x_j \geq c_i$ .

Ak niekedy budeme potrebovať pomenovať prvý popísaný LP, budeme ho volať primárny.

Všimnite si, že premenné duálneho LP v istom zmysle zodpovedajú obmedzeniam primárneho LP. A naopak, obmedzenia v duálnom LP zodpovedajú premenným v primárnom.

Tiež si všimnite, že keby sme našli duálny LP k nášmu duálnemu LP, dostali by sme späť náš primárny LP.

## 4.1 Weak duality theorem

V tejto časti sa konečne dostávame k dôležitým poznatkom o dualite. Ich formulácia už iste nikoho neprekvapí, keďže ide o zovšeobecnenie tvrdení, ktoré sme už stretli na príkladoch.

Weak duality theorem hovorí nasledovne: Nech  $y_1, \dots, y_M$  je ľubovoľné prípustné (nie nutne optimálne) riešenie primárneho LP a nech  $x_1, \dots, x_N$  je ľubovoľné prípustné riešenie duálneho LP. Potom hodnota riešenia primárneho LP je menšia alebo rovná ako hodnota riešenia duálneho LP.

(Čo nám táto veta hovorí o okrajových prípadoch? V prvom rade to, že ak jeden je z našich LP neohraničený, tak ten druhý nemôže mať vôbec žiadne prípustné riešenie. A pozor, ešte je tu jeden špeciálny prípad: môže sa stať, že prípustné riešenie nemá ani jeden z nich.)

Dôkaz je triviálny, táto skutočnosť vyplýva priamo z prípustnosti dotyčných dvoch riešení.

Riešenie primárneho LP spĺňa podmienky  $\forall j : \sum_i A_{i,j}y_i \leq b_j$ , zatiaľ čo riešenie duálneho LP spĺňa podmienky  $\forall i : \sum_j A_{i,j}x_j \geq c_i$ .

Pozrime sa najskôr na primárny LP. Pre ľubovoľné  $j$  platí  $\sum_i A_{i,j}y_i \leq b_j$ . Zoberme hodnotu  $x_j$  z riešenia duálneho LP. O tej vieme, že je nezáporná. Preto platí aj  $\sum_i A_{i,j}x_jy_i \leq b_jx_j$ .

Pre každé  $j$  takto dostaneme jednu nerovnicu. Keď ich všetky sčítame, dostávame:

$$\sum_i \sum_j A_{i,j}x_jy_i \leq \sum_j b_jx_j$$

Keď spravíme analogické úpravy s podmienkami pre duálny LP, dostávame druhú nerovnosť:

$$\sum_i \sum_j A_{i,j}x_jy_i \geq \sum_i c_iy_i$$

Zložením týchto dvoch nerovností dostávame  $\sum_j b_jx_j \geq \sum_i c_iy_i$ , q.e.d.

## 4.2 Strong duality theorem

Strong duality theorem nás ničím novým neprekvapí. Znie: Ak má primárny LP konečné optimálne riešenie, tak aj duálny LP má konečné optimálne riešenie a ich hodnoty sa rovnajú.

Dôkaz neuvedieme, lebo ho nevieme podať formou, ktorá by čitateľovi čokoľvek ďalšie objasnila, či priniesla nejaký nadhľad. Proste sa pošrotia matice a vyjde to.

### 4.3 Complementary slackness

Už sme videli jeden normálny tvar pre LP. Teraz si ukážeme iný: Každý LP sa dá upraviť na ekvivalentný, kde máme pre každú premennú podmienku že je nezáporná, a všetky zvyšné podmienky sú rovnosti.

Ako na to? Trik je jednoduchý: keď máme napr. nerovnosť  $\sum_j A_{i,j}x_j \geq c_i$ , môžeme ju zmeniť na rovnosť tak, že zavedieme novú premennú  $x_{N+j}$  a pôvodnú podmienku zmeníme na  $\sum_j A_{i,j}x_j = c_i + x_{N+j}$ .

Takto zavedené nové premenné teda pre každú podmienku akoby explicitne pomenúvajú „voľnosť“, ktorú pri nej máme. V anglickej terminológii sa tejto hodnote hovorí slackness.

Pozrime sa teraz ešte raz na Weak duality theorem. Ukázali sme, že pre ľubovoľné prípustné riešenia primárneho a duálneho LP ( $y_1, \dots, y_M$  a  $x_1, \dots, x_N$ ) platí nerovnosť:

$$\sum_j b_j x_j \geq \sum_i \sum_j A_{i,j} x_j y_i \geq \sum_i c_i y_i$$

Zo Strong duality theorem vieme, že pre optimálne riešenia, ak obe existujú, nastáva medzi pravou a ľavou stranou rovnosť. Nech teda  $\hat{y}$  a  $\hat{x}$  sú optimálne riešenia. Pozrime sa teraz napr. na ľavú nerovnosť, teda v našom prípade rovnosť:  $\sum_j b_j \hat{x}_j = \sum_i \sum_j A_{i,j} \hat{x}_j \hat{y}_i$ .

Tú môžeme upraviť do nasledovného tvaru:

$$\sum_j \left( b_j - \sum_i A_{i,j} \hat{y}_i \right) \hat{x}_j = 0$$

To, čo sme dostali v zátvorke, silne pripomína jednu z podmienok nášho primárneho LP. Presnejšie, výraz v zátvorke je priamo slackness v dotyčnej nerovnosti. Pripomeňme si, že si pre slackness v primárnom LP môžeme zaviesť nové premenné  $y_{M+j} = b_j - \sum_i A_{i,j} y_i$ . Potom môžeme vyššie uvedenú rovnosť zjednodušiť na nasledovný tvar:  $\sum_j \hat{y}_{M+j} \cdot \hat{x}_j = 0$ .

Všetky premenné vystupujúce v tejto sume sú nezáporné. Aby teda vyšiel výsledok 0, znamená to, že v každej dvojici  $y_{M+j}$  a  $x_j$  musí byť aspoň jedna hodnota nulová.

Čo to pre nás znamená? Ak sa na primárny LP dívame v duchu príkladu s príviesmi ako na využívanie zdrojov, dostávame vzťah medzi dvoma hodnotami – jednou je slackness  $y_{M+j}$  hovoríaca, ktoré zdroje nie sú plne využívané, druhou sú riešenia duálneho problému, o ktorých vieme, že nám predstavujú tieňové ceny jednotlivých zdrojov. Slovné teda môžeme toto tvrdenie zhrnúť nasledovne: zdroj môže mať nenulovú cenu len vtedy, ak je plne využitý (jemu zodpovedajúca slackness je 0).

Práve odvodený vzťah (ktorému sa hovorí Complementary slackness) nám teda ukazuje, že jednotlivé premenné primárneho LP zodpovedajú jednotlivým obmedzeniam duálneho LP a naopak.

#### 4.4 Minimálny rez vracia úder

Keď už teda vieme, ako to vo všeobecnosti funguje s dualitou lineárnych programov, pozrime sa ešte raz poriadnejšie na pre nás programátorov zaujímavú tému: maximálny tok vs. minimálny rez.

Ukážeme teraz dvojicu LP, ktoré sú navzájom duálne, a navyše majú tvar, z ktorého je zjavné, ako zodpovedajú maximálnemu toku, resp. minimálnemu rezu.

Nech  $G = (V, E)$  je orientovaný graf, pričom pre každú hranu  $e \in E$  vieme jej ohodnotenie (kapacitu/cenu)  $c_e$ , a nech  $s, t \in V$ .

Zostrojme  $G' = (V, E')$ , kde  $E'$  vznikne z  $E$  pridaním hrany z  $t$  do  $s$ . (Táto hrana nebude mať obmedzenú kapacitu. Formálne stačí, aby jej  $c$  bolo väčšie ako súčet ostatných, ale nebudeme to potrebovať.)

Teraz pre každú hranu  $e \in E'$  zavedieme premennú  $f_e$  predstavujúcu tok po tejto hrane. Všetky tieto premenné musia samozrejme byť nezáporné, a pre každé  $e \in E$  musí byť  $f_e \leq c_e$ . Veľkosť toku, ktorú maximalizujeme, bude rovná  $f_{ts}$  – teda toku po novo pridanej hrane.

Ešte nám ostávajú podmienky pre Kirchhoffov zákon v jednotlivých vrcholoch. Tie sme síce pôvodne formulovali ako rovnosti, ale bohaté ich stačí formulovať ako nerovnosti hovoriace „v tomto vrchole žiaden tok nepribúda“:

$$\forall v \in V : \sum_{w:vw \in V} f_{vw} - \sum_{w:vw \in V} f_{wv} \leq 0$$

Tým sme zostrojili novú verziu LP pre maximálny tok.

Pozrime sa teraz, čo sa stane, keď k nej mechanicky zostrojíme duálny LP.

V primárnom LP máme  $|V| + |E|$  podmienok (okrem nezápornosti premenných), v duálnom teda bude  $|V| + |E|$  premenných. Podmienke pre hranu  $e$  bude v duálnom programe zodpovedať premenná  $d_e$ , podmienke pre vrchol  $v$  bude zodpovedať premenná  $p_v$ .

Ako budú vyzeráť obmedzenia v duálnom LP? Tie budú zodpovedať premenným v primárnom LP, teda hranám v  $E'$ .

Pre každú hranu  $e = uv \in E$  dostávame obmedzenie:  $p_v + d_{uv} - p_u \geq 0$ .

No a pre špeciálnu hranu  $ts$  dostávame obmedzenie:  $p_s - p_t \geq 1$ . (Tá 1 namiesto 0 sa tam zjavila preto, že v pôvodnej funkcii, ktorú sme maximalizovali, mal tok na tejto hrane koeficient 1, zatiaľ čo tok na ostatných hranách mal koeficient 0.

Posledná otázka: ako bude vyzeráť funkcia, ktorú minimalizujeme? V primárnom LP sú pravé strany nerovností pre vrcholy nulové, teda premenné  $p_v$  budú mať nulové koeficienty. No a každá premenná  $d_e$  bude mať koeficient  $c_e$ . Cieľom je teda minimalizovať  $\sum_{e \in E} d_e \cdot c_e$ .

Zdôrazňujeme, že predchádzajúci postup neobsahoval myslenie. Šlo len o mechanické použitie definície duálneho LP.

Pozrime sa teraz na LP, ktorý sme zostrojili. Ukážeme, že skutočne zodpovedá minimálnemu rezu.

Premenné  $p_v$ , podobne ako keď sme prvýkrát definovali LP pre minimálny rez, budú predstavovať potenciály vrcholov. Namiesto pôvodného  $p_s = 1, p_t = 0$  máme teraz  $p_s - p_t \geq 1$ , ale o chvíľu uvidíme, že to vôbec neprekáža.

Ostatné obmedzenia môžeme prepísať do formy  $d_{uv} \geq p_u - p_v$ , teda hodnota  $d_{uv}$  je aspoň zmena potenciálu na hrane  $uv$ .

Je zjavné, že v optimálnom riešení bude pre všetky hrany  $uv \in E$  platiť rovnosť – totiž každá hrana  $uv$  má kladnú cenu  $c_{uv}$ , a teda ak by sme zvolili  $d_{uv}$  väčšie ako treba, zbytočne tým zvyšujeme hodnotu funkcie, ktorú minimalizujeme.

Teda presnejšie, každé  $d_{uv}$  musí samozrejme byť nezáporné, preto bude platiť rovnosť  $d_{uv} = \max(0, p_u - p_v)$ .

V tomto okamihu už máme ten istý LP ako sme predtým zostrojili ručne. Jediný rozdiel je spomínaná nerovnosť  $p_s - p_t \geq 1$ . No a tu už je všetko jasné: v optimálnom riešení aj tak musí platiť  $p_s - p_t = 1$ . Ak totiž máme riešenie, kde  $p_s - p_t > 1$ , môžeme všetky potenciály preškalovať do intervalu  $[0, 1]$ , čím dostaneme iné platné riešenie s nižšou cenou.

Takže duálny LP, ktorý sme mechanicky zostrojili k primárnemu LP pre maximálny tok, skutočne zodpovedá minimálnemu rezu.

Z Strong duality theorem teraz priamo vyplýva rovnosť medzi veľkosťou maximálneho toku a minimálneho rezu. Z Complementary slackness navyše napr. dostávame pozorovanie, že tie hrany  $e$ , kde je v optimálnom riešení duálneho LP  $d_e > 0$ , majú všetky v optimálnom riešení primárneho LP nulovú slackness. Inými slovami, platí pre ne  $f_e = c_e$ . A teda rez zodpovedajúci optimálnemu riešeniu duálneho LP je v optimálnom riešení primárneho LP celý nasýtený nájdeným tokom.

Domáca úloha: sformulujte úlohu o najkratších cestách ako LP. Čomu zodpovedá duálny problém? Čo vieme zistiť z Complementary slackness?

## 5 Užitočné linky

- [people.cs.uchicago.edu/~brady/CSPP55005/dualexample.pdf](http://people.cs.uchicago.edu/~brady/CSPP55005/dualexample.pdf)
- <http://web.mit.edu/15.053/www/AMP-Chapter-04.pdf>

**Disclaimer.** Tieto poznámky môžete voľne používať na ľubovoľné nekomerčné účely. Na akékoľvek komerčné využitie je potrebný súhlas autora. Ak v mojich poznámkach objavíte nejakú chybu, prípadne ich nejakým spôsobom viete doplniť, budem rád, ak mi dáte vedieť.

Pre potreby prípadného citovania má tento kus poznámok evidenčné číslo MF-0015.