NOPT048: Optimalizační metody - LS 2012 Jiří Sgall

Konvexní množiny

2. 3. 2012, 2. přednáška *Zapsal: Jan Roztočil, Tomáš Musil*
Poslední změna: 16. června 2012

1 Příklady

1.1 Vrcholové pokrytí

Je dán neorientovaný graf $G = (V, E)$, vrcholové pokrytí je podmnožina vrcholů $U \subseteq V$ taková, že každá hrana z E je incidentní aspoň jednomu vrcholu z U. Zavedeme proměnné pro vrcholy x_v, jejich hodnota je buď 1, pokud je vrchol v ve vrcholovém pokrytí, nebo 0, když ne. Celočíselný program:

$$\begin{aligned} \text{minimalizuj} \quad & \sum_{v \in V} x_v \\ \text{pro} \quad & \mathbf{x} \in \{0,1\}^n \\ \text{za podmínek} \quad & x_u + x_v \geq 1 \quad \forall \{u,v\} \in E \end{aligned}$$

Podmínky $x_u + x_v \geq 1$ nám zaručí, že každá hrana bude pokrytá vrcholem a výběr proměnných z $\{0,1\}$ zase celočíselnost. **LP relaxace** této úlohy:

$$\begin{aligned} \text{minimalizuj} \quad & \sum_{v \in V} x_v \\ \text{pro} \quad & 0 \leq x_v \leq 1 \\ \text{za podmínek} \quad & x_u + x_v \geq 1 \quad \forall \{u,v\} \in E \end{aligned}$$

Problém je, že optimální řešení této LP relaxace nemusí (a pravděpodobně také nebude) optimálním řešením původní úlohy celočíselného programování. Navíc proměnné budou reálná čísla mezi 0 a 1.

1.2 Nezávislá množina

Je dán neorientovaný graf $G = (V, E)$, nezávislá množina je podmnožina vrcholů $U \subseteq V$ taková, že Žádné dva vrcholy z U nejsou spojené hranou. V celočíselném programu opět použijeme proměnné pro vrcholy. x_v je rovno 1, pokud v patří do nezávislé množiny, 0 jinak.

$$\begin{aligned} \text{maximalizuj} \quad & \sum_{v \in V} \mathbf{x}_v \\ \text{pro} \quad & \mathbf{x} \in \{0,1\}^n \\ \text{za podmínek} \quad & x_u + x_v \leq 1 \quad \forall \{u,v\} \in E \end{aligned}$$

Podmínky $x_u + x_v \leq 1$ zaručují, že do vrcholového pokrytí se může dostat nejvýše jeden z vrcholů incidentních každé hraně. LP relaxace by se provedla jako u vrcholového pokrytí.

Podobně můžeme modelovat i další grafové problémy, lineární programování je vhodné zj. pro ty, kdy jsou omezující podmínky lokální (tedy ne např. TSP, kde je problém zajistit souvislost).

1.3 Minimální perfektní párování v bipartitním grafu

Na vstupu máme bipartitní graf $G = (U, V, E)$ $|U| = |V| = n$, $E \subseteq U \times V$, G má perfektní párování, a váhovou funkci $w : E \to \mathbb{R}_0^+$. A hledáme perfektní párovaní $M \subseteq E$ takové, že $\sum_{e \in M} w(e)$ je minimální.

V celočíselném programu tentokrát zvolíme proměnné pro hrany:

$$\begin{array}{rll} \text{minimalizuj} & \sum\limits_{e \in E} w(e) x_e & \\ \text{pro} & \mathbf{x} \in \{0,1\}^n & \\ \text{za podmínek} & \sum\limits_{v \in V} x_{uv} = 1 & \forall u \in U \\ & \sum\limits_{u \in U} x_{uv} = 1 & \forall v \in V \end{array}$$

Podmínky zajistí, že do každého vrcholu v U i V vchází právě jedna hrana tedy, že se jedná o perfektní párování. Přímo dostaneme i LP relaxaci této úlohy:

$$\begin{array}{rll} \text{minimalizuj} & \sum\limits_{e \in E} w(e) x_e & \\ \text{pro} & \mathbf{x} \geq 0 & \\ \text{za podmínek} & \sum\limits_{v \in V} x_{uv} = 1 & \forall u \in U \\ & \sum\limits_{u \in U} x_{uv} = 1 & \forall v \in V \end{array}$$

Tady už nám podmínky nezajistí perfektní párování. Může se stát, že např. dostaneme vrchol stupně 2 takový, že proměnné pro hrany jemu incidentní budou mít obě hodnotu 0.5 nebo dokonce 0.9 a 0.1. Naštěstí pro bipartitní grafy platí následující věta.

Věta 1. *Mějme vážený bipartitní graf $G = (V, E)$ a celočíselný program odpovídající hledání minimálního perfektního párování v G, pak existuje optimální řešení $\mathbf{x}$ LP relaxace, které zároveň splňuje $\mathbf{x} \in \{0,1\}^m$.*

Důkaz. Pokud graf nemá perfektní párování, pak neexistuje ani přípustné řešení, není tedy co dokazovat. Naopak když graf má perfektní párování, pak optimální řešení existuje, protože přípustná řešení tvoří kompaktní množinu (je omezená, protože $0 \leq \mathbf{x} \leq 1$). Vezmeme tedy $\mathbf{x}$ optimální řešení LP relaxace s minimálním počtem neceločíselných proměnných.

Vybereme podmnožinu hran $E' = \{e \in E \mid x_e \in (0,1)\}$. Tvrdíme, že graf G' indukovaný hranami z E' neobsahuje vrcholy stupně 1, kdyby totiž takový vrchol měl, tak ten by musel být celočíselný, protože platí podmínky z LP, tady spor s výběrem hran do E'

Dále tvrdíme, že v G' existuje cyklus $C \subseteq E'$, což plyne z předchozího. A protože je G bipartitní, tak $|C| = 2k$ (C je sudé délky). Hrany v cyklu očíslujeme $C = e_1, e_2, \ldots, e_{2k}$.

Zvolme jiné řešení $\mathbf{y}$ a nějaké malé reálné ε takto:

$$y_e = \begin{cases} x_e + \varepsilon & e_1, e_3, \ldots, e_{2k-1} \\ x_e - \varepsilon & e_2, e_4, \ldots, e_{2k} \\ x_e & \text{jinak} \end{cases}$$

Máme stále přípustné řešení? Protože ke každému vrcholu jednou ε přičteme a jednou od něj odečteme, tak suma zůstane stejná. A protože x_e pro hrany z C jsou z ostrého intervalu $(0, 1)$, tak pro dostatečně malé ε bude platit i podmínka $0 \leq x_e \leq 1$.

Jak to změní účelovou funkci?

$$\sum_{e \in E} w(e) y_e = \sum_{e \in E} w(e) x_e + \varepsilon \overbrace{\Big(\sum_{\text{sudé } e \in C} w(e) x_e - \sum_{\text{liché } e \in C} w(e) x_e \Big)}^{=0}$$

Protože $\mathbf{x}$ je optimální, musí být výraz ve svorce 0, jinak bychom mohli vhodnou volbou kladného nebo záporného ε zlepšit optimum.

Pro maximální ε takové, že $\mathbf{y}$ je přípustné řešení, platí pro některou hranu $e \in \{e_1, \ldots, e_{2k}\}$, že $x_e \in \{0, 1\}$, tedy $\mathbf{x}$ má méně neceločíselných složek. Postup opakujeme, dokud nejsou všechny x_e celočíselné.

$\square$

2 Opakování lineární algebry

2.1 Afinní prostory

Definice 2. A je **afinní prostor**, pokud $A = \mathbf{x} + L$ pro nějaké $\mathbf{x} \in \mathbb{R}^n$, L vektorový podprostor.

Definice 3. **Afinní obal** $X \subseteq \mathbb{R}^n$ je průnik všech afinních prostorů A, $X \subseteq A$.

Definice 4. **Afinní kombinace** bodů z X je bod (výraz) $\alpha_0 a_0 + \ldots + \alpha_k a_k$, kde $k \in \mathbb{N}, a_i \in X, \alpha_i \in \mathbb{R}, \sum \alpha_i = 1$

Věta 5. *Afinní obal X je roven $\{\alpha_0 a_0 + \ldots + \alpha_k a_k \mid k \in \mathbb{N}, a_i \in X, \alpha_i \in \mathbb{R}, \sum \alpha_i = 1\}$*

Definice 6. $a_0, \ldots, a_k$ jsou **afinně nezávislé**, pokud neexistují netriviální $\alpha_0, \ldots, \alpha_k \in \mathbb{R}$ (tj. ne všechny $= 0$) tak, že $\alpha_0 a_0 + \ldots + \alpha_k a_k = 0$ a $\sum \alpha_i = 0$.

Pozorování 7. *$a_0, \ldots, a_k$ afinně nezávisle právě tehdy, když $a_1 - a_0, a_2 - a_0, \ldots, a_k - a_0$ jsou lineárně nezávislé*

Definice 8. **Dimenze množiny** $X \subseteq \mathbb{R}^n$ je dimenze afinního obalu X.

Pozorování 9. *Dimenze množiny X je maximální d takové, že v X existují afinně nezávislé $a_0, \ldots, a_d$.*

2.2 Konvexní množiny

Definice 10. $X \subseteq \mathbb{R}^n$ je **konvexní**, pokud $\forall \mathbf{x}, \mathbf{y} \in X$ a $\forall t \in [0, 1]$ platí $t\mathbf{x} + (1 - t)\mathbf{y} \in X$ (s každými dvěma body je v X i úsečka mezi nimi).

Definice 11. Konvexní obal $X \subseteq \mathbb{R}^n$ je definován jako průnik všech konvexních množin obsahujících X:

$$\operatorname{conv}(X) = \bigcap_{\substack{C \text{ konvexní} \\ C \supseteq X}} C.$$

Definice 12. Konvexní kombinace bodů z X je (bod nebo výraz): $\alpha_0 a_0 + \ldots + \alpha_k a_k$, kde $k \in \mathbb{N}, a_i \in X, \alpha_i \geq 0, \sum \alpha_i = 1$.

Věta 13. *Konvexní obal X je roven* $\{\alpha_0 a_0 + \ldots + \alpha_k a_k \mid k \in \mathbb{N}, a_i \in X, \alpha_i \geq 0, \sum \alpha_i = 1\}$

Důkaz. TODO □

Věta 14 (Carathéodory)**.** *Nechť* $X \subseteq \mathbb{R}^n$*,* $\dim X = d$*. Pak* $\operatorname{conv}(X) = \{\alpha_0 a_0 + \ldots + \alpha_d a_d \mid a_i \in X, \alpha_i \geq 0, \sum \alpha_i = 1\}$

Důkaz. Nechť $\mathbf{x} \in \operatorname{conv}(X)$ lze zapsat jako konvexní kombinaci $\mathbf{x} = \alpha_0 a_0 + \ldots + \alpha_k a_k$, $\alpha_i \geq 0$, $\sum \alpha_i$ tak, že k je minimální.

Pokud $k \leq d$, pak je $\mathbf{x}$ zjevně i konvexní kombinace d prvků, máme hotovo.

Pokud $k > d$, pak $a_0, \ldots, a_k$ jsou afinně závislé, tedy existují $\beta_i \geq 0$ tak, že $\beta_0 a_0 + \ldots + \beta_k a_k = \mathbf{0}$, $\sum \beta_i = 0$, ale pro nějaké i, $\beta_i \neq 0$, mezi β_i jsou tedy kladná i záporná čísla. Dále vezmeme γ maximální takové, že $(\forall i)(\gamma\beta_i + \alpha_i \geq 0)$. Takové číslo bude existovat, protože množina γ, ze které vybíráme, je neprázdná (obsahuje $\gamma = 0$), omezená (protože některé β_i je záporné) a uzavřená (daná neostrými nerovnostmi)

$$\sum_{i=0}^{k} \underbrace{(\gamma\beta_i + \alpha_i)}_{\geq 0} a_i = \mathbf{x}$$

Vidíme, že je tato kombinace konvexní, nezápornost máme díky výběru γ a platí $\sum(\gamma\beta_i + \alpha_i) = 1$, to plyne z $\sum \alpha_i = 1$ a $\sum \beta_i = 0$.

Dále díky maximalitě γ bude existovat i tak, že $\gamma\beta_i + \alpha_i = 0$, tedy $\mathbf{x}$ lze vyjádřit jako konvexní kombinaci méně bodů.

□