MATEMATICKO-FYZIKÁLNÍ FAKULTA
Univerzita Karlova

BAKALÁŘSKÁ PRÁCE

Michaela Minaříková

Bendersova dekompozice v optimalizaci

Katedra pravděpodobnosti a matematické statistiky

Vedoucí bakalářské práce:	RNDr. Martin Branda, Ph.D.
Studijní program:	matematika
Studijní obor:	obecná matematika

Praha 2018

Prohlašuji, že jsem tuto bakalářskou práci vypracoval(a) samostatně a výhradně s použitím citovaných pramenů, literatury a dalších odborných zdrojů.

Beru na vědomí, že se na moji práci vztahují práva a povinnosti vyplývající ze zákona č. 121/2000 Sb., autorského zákona v platném znění, zejména skutečnost, že Univerzita Karlova má právo na uzavření licenční smlouvy o užití této práce jako školního díla podle §60 odst. 1 autorského zákona.

V .... dne .... Podpis autora

Ráda bych na tomto místě poděkovala RNDr. Martinu Brandovi, Ph.D. za vedení mé práce, za jeho rady, připomínky a čas, což mi velmi pomohlo k dosažení výsledné úrovně práce.

Název práce: Bendersova dekompozice v optimalizaci

Autor: Michaela Minaříková

Katedra: Katedra pravděpodobnosti a matematické statistiky

Vedoucí bakalářské práce: RNDr. Martin Branda, Ph.D., Katedra pravděpodobnosti a matematické statistiky

Abstrakt: Práce pojednává o Bendersově dekompozici v optimalizaci, konkrétně ve stochastickém lineárním programování. Čtenář je nejdříve seznámen s důležitými pojmy používanými v dekompozičním algoritmu. Následně je vysvětleno, jak lze úlohu stochastického lineárního programování přeformulovat na tvar vhodný pro Bendersův algoritmus. V třetí kapitole je dekompoziční algoritmus, založený na řezech přípustnosti a optimality, vysvětlen včetně podmínek konvergence algoritmu. Pro dvoustupňové stochastické lineární programování je uvedena modifikace algoritmu. V průběhu práce je Bendersův algoritmus ilustrován na dvou menších příkladech.

Klíčová slova: Bendersova dekompozice, stochastické lineární programování, řez přípustnosti, řez optimality

Title: Benders decomposition in optimization

Author: Michaela Minaříková

Department: Department of Probability and Mathematical Statistics

Supervisor: RNDr. Martin Branda, Ph.D., Department of Probability and Mathematical Statistics

Abstract: The Bachelor thesis is dealing with Benders decomposition in optimization, especially in stochastic linear programming. In the begining the reader will be introduced to the important terms used in the decomposition algorithm. Consequently it is demonstrated how to reformulate the problem of stochastic linear programming to a special structure suitable for Benders decomposition. In the third chapter, the decomposition algorithm, using the feasibility and optimality cuts, is explained including conditions of convergence of the algorithm. There follows modification of algorithm for two stage stochastic linear programming. Finally, we illustrate Benders algorithm on two smaller problems.

Keywords: Benders decomposition, stochastic linear programming, feasibility cut, optimality cut

Obsah

Úvod

Bendersova dekompozice je jedním ze stále se rozvíjejících postupů v moderní optimalizaci. Myšlenku poprvé uvedl J.F. Benders v článku [1] v roce 1962, od té doby se tento algoritmus modifikuje pro různé optimalizační úlohy. Benders původně algoritmus zamýšlel pro smíšené celočíselné programování. I když tak, jak ho uvedl, má mnohem obecnější tvar s jasnými podmínkami pro nalezení řešení. Bendersova dekompozice totiž využívá speciální struktury a urychluje výpočet pro větší úlohy. V této práci si uvedeme modifikaci algoritmu uzpůsobenou především pro dvoustupňové stochastické lineární programování.

Myšlenka algoritmu spočívá v rozdělení problému na dva podproblémy a jejich iterační řešení. Postupně se přidávají další podmínky na řešení pomocí řezů optimality a přípustnosti. Benders svou úlohu rozdělil na podproblémy klasického linárního programování a celočíselného programování.

V první kapitole bakalářské práce uvedeme důležité definice a věty z optimalizace, jejichž znalost je pro pochopení algoritmu nutná. V druhé kapitole vysvětlíme, jak je třeba upravit formulaci stochastického lineárního programování pro použití Bendersovy dekompozice.

Ve třetí kapitole bude vysvětlen algoritmus včetně podmínek konvergence. Jelikož práce je zaměřená především na stochastické lineární programování, tak algoritmus Benderse neboli duální algoritmus, bude uveden jako speciální případ. Na něm budou i demonstrovány podmínky konvergence. Obecným případem je myšleno použití pro dvoustupňové stochastické lineární programování.

Poslední kapitola bude ještě doplněna dvěma jednoduchými příklady. První příklad bude jen pro ukázku algoritmu a optimálního řezu. Duální algoritmus použijeme na lineární programování, které má strukturu vhodnou pro tento algoritmus, což uvádí i Benders ve svém článku [1] jako jednu z možných aplikací. Druhý příklad bude stochastický. Ukážeme si na něm přeformulování stochastického lineárního programování podle druhé kapitoly a následné řešení pomocí obecného případu algoritmu z kapitoly třetí.

Cílem práce je seznámit čtenáře, znalého základních poznatků z optimalizace, s Bendersovou dekompozicí a její možností aplikace.

1. Lineární programování

V této kapitole si shrneme důležité definice a věty, týkající se lineárního programování (zkráceně LP), které jsou uvedeny v knize Úvod do optimalizace [4], jež budeme dále intenzivně využívat.

1.1 Formulace LP

Lineární programování je matematický problém optimalizace, který se snaží najít minimální (maximální) hodnotu lineární účelové funkce za daných podmínek určených lineárními funkcemi.
Problém LP se nejčastěji formuluje ve standardním tvaru:

$$\begin{aligned} \min\ c^\top x & \qquad (1.1)\\ \text{za podm.: } Ax &= b \qquad (1.2)\\ x &\geq 0, \qquad (1.3) \end{aligned}$$

kde $A \in \mathbb{R}^{m\times n}$; c,$x \in \mathbb{R}^n$ a $b \in \mathbb{R}^m$. Ať už dostaneme zadání LP jakékoliv, vždy se dá přeformulovat do standardního tvaru pomocí několika triků:

$$\begin{aligned} \max\ c^\top x &= -\min\ -c^\top x\\ x &= x^+ - x^-\\ x^+, x^- &\geq 0\\ \text{pomocná skluzová proměnná: } y &\geq 0, \end{aligned}$$

kde $y \in \mathbb{R}^n$. Pomocná skluzová proměnná se využívá u podmínek, které jsou zadány s nerovností, přičtením/odečtením této proměnné můžeme nerovnici transformovat do rovnice.

1.2 Vlastnosti úlohy LP ve standardním tvaru

Nechť máme úlohu LP ve standardním tvaru 1.3. Z lineární algebry víme, že soustava $Ax = b$ je řešitelné právě tehdy, když rank(A,b) =rank(A). Množinou přípustných řešení úlohy LP ve standardním tvaru rozumíme

$$M = \{x \in \mathbb{R}^n; Ax = b, x \geq 0\}$$

Nyní si uvedeme několik pojmů z optimalizace vhodných pro popis množiny přípustných řešení a konstrukci algoritmu.

Definice 1 (Uzavřený poloprostor)**.** *Uzavřeným poloprostorem rozumíme množinu* $\{x \in \mathbb{R}^n : a \neq 0,\ a^\top x \geq c\}$

Definice 2 (Konvexní polyedrická množina)**.** *Je-li množina průnikem konečně mnoha uzavřených poloprostorů, pak ji nazýváme konvexní polyedrickou množinou.*

Definice 3 (Konvexní polyedr)**.** *Je-li množina konvexní polyedrická a omezená, pak ji nazveme konvexním polyedrem.*

Definice 4 (Kužel)**.** *Kuželem rozumíme množinu $K \subset \mathbb{R}^n$, která obsahuje počátek a nezáporné násobky svých prvků. Tedy $\forall x \in K\ \forall \alpha \geq 0 : \alpha x \in K$.*

Definice 5 (Konvexní kužel)**.** *Konvexním kuželem rozumíme množinu $K \subset \mathbb{R}^n$ která obsahuje počátek a s každými dvěma body obsahuje také všechny jejich pozitivní lineární kombinace. Tedy $\forall x, y \in K\ \forall \alpha \geq 0\ \forall \beta \geq 0 : \alpha x + \beta y \in K$.*

Definice 6 (Konvexní polyedrický kužel)**.** *Množina, která je konvexní kužel a současně i konvexní polyedrická množina, se nazývá konvexní polyedrický kužel.*

V dekompozici využíváme vlastnost tvaru množiny přípustných řešení úlohy LP. Následující věty budou důležité i pro některé části důkazu a budeme se na ně odvolávat.

Věta 1. *Množina přípustných řešení úlohy LP ve standardním tvaru je konvexní polyedrická množina.*

Důkaz. V knize [4] na straně 21.

□

Věta 2. *Množina přípustných řešení úlohy LP ve standardním tvaru je buď prázdná nebo je tvaru algebraického součtu $M = \mathcal{P} + \mathcal{K}$, kde $\mathcal{P}$ je konvexní polyedr a množina $\mathcal{K} = \{y \in \mathbb{R}^n : Ay = 0, y \geq 0\}$ je konvexní polyedrický kužel.*

Důkaz. V knize [2] na straně 58.

□

Pro algoritmus je klíčová následující ekvivalence o optimálním řešení.

Věta 3 (Existence optimálního řešení)**.** *Úloha LP ve standardním tvaru má optimální řešení právě tehdy, když*
1. $M \neq \emptyset$,
2. $\forall y \in \{x \in \mathbb{R}^n : Ax = 0,\ x \geq 0\} :\ c^\top y \geq 0$.

Důkaz. V knize [4] na straně 23.

□

V řešení některých kroků algoritmu budeme využívat dualitu, jelikož vše bude uvedeno ve standardním tvaru, tak nám stačí zadefinovat duální úlohu pouze pro standardní tvar.

Definice 7 (Duální úloha LP ve standardním tvaru)**.** *Nechť máme úlohu* $\min\ \{c^\top x : Ax = b, x \geq 0\}$, *pak úloha* $\max\ \{b^\top y : A^\top y \leq c\}$ *se nazývá duální úlohou LP ve standardním tvaru.*

Pro dvojici duálních úloh platí silná věta o dualitě, kterou využijeme při vyhodnocování rekurzivní funkce v algoritmu.

Věta 4 (Silná věta o dualitě)**.** *Úloha* $\min\ \{c^\top x : Ax = b, x \geq 0\}$ *má optimální řešení právě tehdy, když má úloha* $\max\ \{b^\top y : A^\top y \leq c\}$ *optimální řešení. V takovém případě platí, že optimální hodnoty se rovnají.*

Důkaz. V knize [4] na straně 34.

□

Ke konstrukci řezů budeme využívat vztahy mezi množinami přípustných řešení duálních úloh.

Věta 5. *Nechť máme úlohu LP* $\min \{c^\top x : Ax = b, x \geq 0\}$ *s množinou řešení* $\mathcal{M}$ *a její duální úlohu* $\max \{b^\top y : A^\top y \leq c\}$ *s množinou přípustných řešení* $\mathcal{N}$*, pak nastává jedna ze čtyř možností:*
(i) Žádná z těchto úloh nemá přípustné řešení, tj. $\mathcal{M} = \emptyset$, $\mathcal{N} = \emptyset$.
(ii) $\mathcal{M} = \emptyset$, $\mathcal{N} \neq \emptyset$*, ale úloha je neomezená (tj.* $\max b^\top y = \infty$*, tedy neexistuje optimální řešení).*
(iii) $\mathcal{N} = \emptyset$, $\mathcal{M} \neq \emptyset$*, ale úloha je neomezená (tj.* $\min c^\top x = -\infty$*, tedy neexistuje optimální řešení).*
(iv) Obě úlohy mají optimální řešení a platí $\min c^\top x = \max b^\top y$, $x \in \mathcal{M}$, $y \in \mathcal{N}$.

Důkaz. V knize [4] na straně 35.

□

2. Stochastické lineární programování

V optimalizaci můžeme řešit i problémy, které jsou určené chováním náhodných matic, vektorů. Úloha LP s náhodnými omezeními může být:

$$
\begin{aligned}
\min\ & c^\top x \\
\text{za podm.: } Ax &= b \\
Tx &= h \\
x &\geq 0,
\end{aligned}
\tag{2.1}
$$

kde proměnnou T označujeme náhodnou matici a h je náhodný vektor definované na pravděpodobnostním prostoru $(\Omega, \mathcal{A}, P)$. Takto zadaná úloha může být pro některé realizace T a h přeurčená.
V následujícím textu uvedeme formulaci dvoustupňového stochastického programování a jeho klasifikaci. Ukážeme, jak se dá úloha 2.1 přeformulovat na dvoustupňové stochastické lineární programování. Právě úlohy se strukturou, kterou má dvoustupňové SLP, je vhodné řešit algoritmem, který bude v následující kapitole.

2.1 Dvoustupňové stochastické programování

Uvažujme dvoustupňové stochastické lineární programování ve tvaru:

$$
\begin{aligned}
\min\ & c^\top x + \mathbf{E}\, Q(x;\xi) \\
\text{za podm.: } & Ax = b \\
& x \geq 0,
\end{aligned}
\tag{2.2}
$$

kde $Q(x,\xi)$ je optimální hodnota problému druhého stupně:

$$
\begin{aligned}
Q(x;\xi) &:= \min\ q^\top y \\
\text{za podm.: } Wy &= h - Tx \\
y &\geq 0,
\end{aligned}
$$

písmenem ξ označujeme realizaci (q,h,T,W) a funkci $Q(x,\xi)$ nazýváme rekurzivní funkcí. Prvky ξ jsou náhodné vektory/matice (nemusí být všechny prvky náhodné).
Pokud je matice W deterministicky určená, pak říkáme, že máme problém s fixní rekurzí. Speciální případem volby ξ pro fixní rekurzi je $W = (I, -I)$, vektor $q \geq 0$ je deterministický, T a h jsou náhodné, pak hovoříme o jednoduché kompenzaci.
V případě, že $\forall x \in \{x : Ax = b, x \geq 0\}$ je množina přípustných řešení funkce $Q(x,\xi)$ neprázdná pro skoro všechna ξ, mluvíme o relativně úplné kompenzaci.

2.2 Struktura SLP pro dekompozici

Na začátku kapitoly jsme si uvedli úlohu SLP 2.1. Předpokládejme že T a h mají diskrétní rozdělení. Nejdříve je potřeba vyřešit omezení určená maticí A

a vektorem b, což nazýváme podmínkami prvního stupně. Pro všechny realizace náhodné matice T a náhodného vektoru h nemusí být splněna rovnost $Tx = h$, neboť problém bude přeurčený. Jelikož jde o SLP, tak využijeme poznatků z předchozí podkapitoly a úlohu přeformulujeme abychom minimalizovali chybu $h - Tx$. Máme tedy fixní rekurzivní funkci:

$$
\begin{aligned}
Q(x,T,h) &:= \min\, q^\top y \\
\text{za podm.: } Wy &= h - Tx \\
y &\geq 0,
\end{aligned}
$$

a úlohu uvedeme do tvaru 2.2.

Pokud náhodné proměnné T a h mají diskrétní rozdělení, definovaná hodnotami (T^j, h^j) s pravděpodobnostmi p_j, kde $j = 1, \ldots, S$ ($p_j > 0$ a $\sum_{j=1}^S p_j = 1$), problém zmíněný výše si tak ekvivalentně zapíšeme:

$$
\begin{array}{llcll}
 & \min\, c^\top x \; + \; \sum_{j=1}^S p_j q^\top y^j & & & \\
\text{za podm.} & Ax & = & b & \\
 & T^j x \quad + W y^j & = & h^j, & j = 1, \ldots, S \\
 & x & \geq & 0 & \\
 & \qquad\qquad y^j & \geq & 0, &
\end{array}
\tag{2.3}
$$

tím pádem jsme opět získali problém lineárního programování se speciální strukturou vhodnou pro duální dekompozici, což uvidíme v následující kapitole. I když úloha je řešitelná přímo algoritmy LP, pro větší dimenze se vyplatí využít struktury a výpočet urychlit Bendersovou dekompozicí.

3. Duální dekompozice

V předchozí kapitole jsme odvodili, jak se dá dvoustupňové SLP s rekurzí vyjádřit jako LP problém se speciální strukturou. Takto strukturované LP se dá vyřešit pomocí algoritmu poprvé prezentovaném Bendersem v článku [1].

3.1 Speciální případ

Nejdříve pro S=1. Předpokládejme, že pro LP problém 2.3 existuje optimální řešení a množina přípustných řešení $\{x; Ax = b, x \geq 0\}$ je omezená. Podle Věty 3 řešitelnost soustavy implikuje, že

$$\{(x,y) : Ax = b,\, Tx + Wy = h, x \geq 0, y \geq 0\} \neq \emptyset$$

a

$$c^\top \zeta + q^\top \eta \geq 0 \;\forall\; (\zeta,\eta) \in \{(\zeta, \eta) : A\zeta = 0,\, T\zeta + W\eta = 0,\, \zeta \geq 0,\, \eta \geq 0\},$$

speciálně pro volbu $\zeta = 0$:

$$q^\top \eta \geq 0 \;\forall \eta \in \{\eta : W\eta = 0,\, \eta \geq 0\}.$$

Definujme rekurzivní funkci

$$f(x) := \min\, \{q^\top y : \; Wy = h - Tx,\, y \geq 0\},$$

která pro každé x najde minimální hodnotu $q^\top y$ za daných omezení. Takto definovaná funkce je konečná, pokud jsou omezení funkce neprázdná. V opačném případě definujeme $f(x) = \infty$.
Poté můžeme náš problém přepsat do tvaru NLP (nelineární programování):

$$\begin{aligned} \min\; & c^\top x + f(x) \\ \text{za podm.}\; & Ax = b \\ & x \geq 0. \end{aligned}$$

Stejně tak můžeme tento problém zapsat jako:

$$\min\; c^\top x + \theta \tag{3.1}$$
$$\text{za podm.}\; Ax = b \tag{3.2}$$
$$\theta - f(x) \geq 0 \tag{3.3}$$
$$x \geq 0. \tag{3.4}$$

Tento problém se dá řešit algoritmem sepsaným v [3] v podkapitole 2.6, který uvedeme v následující podkapitole.

3.1.1 Dekompoziční algoritmus

Bendersova dekompozice je založena na myšlence rozdělení problému na dva podproblémy. V každém kroku, kdy nenajde optimální řešení, tak přidává řezy, které zmenšují množinu přípustných řešení. Jsou dva typy řezu - optimality a přípustnosti. V některých krocích se pro výpočty využívá vlastností duálních úloh, které jsme uvedli v první kapitole.

1. Inicializace
Nejdříve najdeme optimální hodnotu:

$$\min\{q^\top y : Ax = b, Tx + Wy = h, x \geq 0, y \geq 0\},$$

kterou označíme θ_0 a vyřešíme LP úlohu

$$\min\{c^\top x + \theta : Ax = b, x \geq 0, \theta \geq \theta_0\},$$

čímž máme řešení ($\hat{x}$,$\hat{\theta}$). Definujme

$$\mathcal{B}_0 = \{(x,\theta) : Ax = b, x \geq 0, \theta \in \mathbb{R}\},$$

$$\mathcal{B}_1 = \{\mathbb{R}^n \times \{\theta\} : \ \theta \geq \theta_0\}.$$

2. Vyhodnocení rekurzivní funkce
Najdeme hodnotu

$$f(\hat{x}) = \min\{q^\top y : Wy = h - T\hat{x}, y \geq 0\}$$

pomocí duální úlohy LP:

$$f(\hat{x}) = \max\{(h - T\hat{x})^\top u : W^\top u \leq q\}.$$

Pokud $f(\hat{x}) = \infty$, nastal případ (ii) z věty 5 a jdeme na krok 3, jinak jdeme na krok 4.

3. Řez přípustnosti
Jestliže $\hat{x}$ není přípustné řešení pro 2.3, tak podle věty 3 existuje vektor $\bar{u}$ takový, že $W^\top \bar{u} \leq 0$ a $(h - T\hat{x})^\top \bar{u} > 0$, zatímco pro x, které je přípustným řešením 2.3 platí, že existuje $y \geq 0$ takové, že $Wy = h - Tx$. Vynásobením $\bar{u}^\top$ dostaneme nerovnost

$$\bar{u}^\top (h - Tx) = \bar{u}^\top Wy \leq 0,$$

která musí platit pro kterékoli přípustné x, ale neplatí pro $\hat{x}$. Proto předefinujeme

$$\mathcal{B}_1 := \mathcal{B}_1 \cap \{(x,\theta) : \bar{u}^\top (h - Tx) \leq 0\}.$$

Tím pádem jsme "odsekli" $\hat{x}$, protože $\hat{x} \notin \mathcal{B}_1$, takže jdeme na krok 5.

4. Řez optimality
Jelikož $f(\hat{x})$ je konečná hodnota, pak podle věty 2.3 z knihy [3] na straně 16

existuje řešení duální úlohy $\hat{u}$ přímo vypočtené v kroku 2 takové, že $f(\hat{x}) = (h - T\hat{x})^\top \hat{u}$. Pro každé x platí, že

$$f(x) = \sup\{(h - Tx)^\top u : W^\top u \leq q\} \geq (h - Tx)^\top \hat{u},$$

kde nerovnost plyne z vlastnosti suprema a faktu, že $\hat{u} \in \{W^\top u \leq q\}$.
Nerovnost $\theta - f(x) \geq 0$ v 3.3 implikuje podmínku

$$\theta \geq (h - Tx)^\top \hat{u}.$$

Pokud pro $(\hat{x},\hat{\theta})$ je nerovnost splněna, neboli $f(\hat{x}) \leq \hat{\theta}$, pak zastavíme proceduru a řešením prvního stupně je $\hat{x}$. V opačném případě předefinujeme množinu přípustných řešení:

$$\mathcal{B}_1 := \mathcal{B}_1 \cap \{(x,\theta) : \theta \geq (h - Tx)^\top \hat{u}\}$$

a jdeme na krok 5.

5. Přepočítání LP
Vyřešíme novou úlohu LP:

$$\min \{c^\top x + \theta : (x,\theta) \in \mathcal{B}_0 \cap \mathcal{B}_1\}$$

s novým řešením $(\hat{x},\hat{\theta})$ a vrátíme se na krok 2.

3.1.2 Příklad LP

Uvedeme si jednoduchý příklad LP na ukázku algoritmu. Mějme optimalizační úlohu zadanou:

$$\min 4x_1 + 2x_2 + 5y$$

$$\begin{array}{lllll} \text{za podm.} & 3x_1 & + x_2 & & = 6 \\ & 2x_1 & + 2x_2 & + y & = 10 \\ & x_1 & & & \geq 0 \\ & & x_2 & & \geq 0 \\ & & & y & \geq 0 \end{array} \tag{3.5}$$

1. Krok
Definujme: $\theta_0 = 0$, neb dolní mez je 0.
Řešíme LP:

$$\min\{4x_1 + 2x_2 + \theta : 3x_1 + x_2 = 6 \ x_1,x_2 \geq 0, \ \theta \geq \theta_0\},$$

což nám dává řešení $\hat{x} = (2,0)^\top$ a $\hat{\theta} = 0$.

2. Rekurzivní funkce:
$f(x) = \min\{5y : y = 10 - 2x_1 - 2x_2, \ y \geq 0\}$
Řešíme duální úlohu pro $\hat{x} = (2,0)^\top$:
$f(\hat{x}) = \max\{6u : u \leq 5\}$
$f(\hat{x}) = 30$

3. Optimální řez
Naše $\hat{\theta} = 0$, což je menší než hodnota $f(\hat{x})$, takže předefinujeme množinu řešení podmínkou:

$$\theta \geq 50 - 10x_1 - 10x_2.$$

4. Přepočítání LP
Řešíme úlohu LP:

$$\min\{4x_1 + 2x_2 + \theta : 3x_1 + x_2 = 6,\ 50 - 10x_1 - 10x_2 \leq \theta\ x_1,x_2 \geq 0\},$$

což nám dává řešení $(\frac{1}{2},\frac{9}{2},0)$.

5. Rekurzivní funkce:
Nové řešení z předchozího kroku je $\hat{x} = (\frac{1}{2},\frac{9}{2})$, pro které $f(\hat{x}) = 0$, což se rovná $\hat{\theta}$ a tedy je splněna nerovnost ze 4. kroku algoritmu a nalezli jsme optimální řešení.

3.2 Podmínky konvergence

V následujících větách stále předpokládáme S=1, pokud mluvíme o úloze 2.3.

Věta 6. *Rekurzivní funkce $f(x)$ definovaná na omezené množině $\{x : Ax = b,\ Tx + Wy = h, x \geq 0\} \neq \emptyset$, je po částech lineární, konvexní a omezená.*

Důkaz. Podle našich předpokladů množina $\mathcal{B}_1 := \{x : Ax = b,\ x \geq 0\}$ je omezená. Pak množina

$$\mathcal{B} := \mathcal{B}_1 \cap \{x : \exists y \geq 0 : Wy = h - Tx\} \neq \emptyset$$

je omezená.
Podle věty 1 je $\mathcal{B}_1$ konvexní polyedr. Tedy $\mathcal{B}$ je průnikem konvexní polyedrické množiny a konvexního polyedru, tudíž je také konvexní polyedr.
Nechť je $x \in \mathcal{B}$ pevné, pak

$$f(x) = q_B^\top W_B^{-1}(h - Tx),$$

kde W_B je příslušná optimální báze W vybraná z konečně mnoha přípustných bází W a q_B jsou příslušné složky vektoru q. Proto $f(x)$ je po částech lineární a ze spodu omezená pro $x \in \mathcal{B}$.
Nechť $x^1 \in \mathcal{B}$, $x^2 \in \mathcal{B}$ takové, že $f(x^i)$, $i = 1{,}2$ jsou konečné hodnoty s odpovídajícími y^1, y^2:

$$f(x^i) = q^\top y^i,\ i = 1{,}2 \text{ a } Wy^i = h - Tx^i,\ y^i \geq 0, i = 1{,}2.$$

Nechť je $\lambda \in (0{,}1)$ libovolné a nechť $\tilde{x} = \lambda x^1 + (1 - \lambda)x^2$, pak

$$\lambda y^1 + (1 - \lambda)y^2 \in \{y : Wy = h - T\tilde{x},\ y \geq 0\},$$

neboť

$$h - T\tilde{x} = \lambda(h - Tx^1) + (1 - \lambda)(h - Tx^2) =$$
$$\lambda Wy^1 + (1 - \lambda)Wy^2 = W(\lambda y^1 + (1 - \lambda)y^2),$$

z linearity a z definice dvojic (x^1,y^1), (x^2,y^2).
Potom

$$f(\tilde{x}) = \min\{q^\top y : Wy = h - T\tilde{x}, y \geq 0\} \leq$$
$$q^\top(\lambda y^1 + (1-\lambda)y^2) = \lambda f(x^1) + (1-\lambda)f(x^2),$$

jelikož z předchozího víme, že vektor $\lambda y^1 + (1-\lambda)y^2$ patří do množiny, přes kterou minimalizujeme, a z vlastnosti minima plyne nerovnost. Tím je konvexita funkce f(x) na množině $\mathcal{B}$ dokázána.

□

Věta 7. *Nechť jsou splněny všechny předpoklady uvedené výše, pak algoritmus duální dekompozice nám dá optimální řešení po konečně mnoha krocích.*

Důkaz. Vzhledem k větě 6, tak nejnižší mez θ_0 z prvního kroku existuje. Vzhledem k řešitelnosti naší úlohy 2.3, tak podmínky duální úlohy ($W^\top u \leq q$) jsou přípustné a nezávislé na x. Proto platí, že duální reprezentace rekurzivní funkce $f(\hat{x})$ v 2. kroce je vždy řešitelná nebo je nekonečno. Neomezenost implikuje nepřípustnost primární úlohy (podle bodu (ii) ve větě 5).
Jestliže $f(\hat{x}) = \infty$, pak $\hat{x}$ není přípustné řešení pro naši úlohu 2.3. Vzhledem k větě 3 existuje

$$\tilde{u} : W^\top \tilde{u} \leq 0 \text{ a } (h - T\hat{x})^\top \tilde{u} > 0.$$

Můžeme uvažovat $\tilde{u}$, které je jedním z prvků generující kužel $\{u : W^\top u \leq 0\}$. Jelikož kužel je konečně generovaná množina, pak můžeme přidat nejvýše konečně mnoho omezení

$$(h - Tx)^\top \tilde{u} \leq 0$$

dokud nedostaneme konečnou hodnotu rekurzivní funkce ve všech následujících krocích.
Jestliže $f(\hat{x}) = (h - Tx)^\top \hat{u}$ je konečná hodnota, pak $\hat{u}$ je optimálním řešením duální úlohy, neboli je bazické přípustné řešení. Je jen konečně mnoho přípustných bazických řešení, tedy jen konečně mnoho podmínek:

$$\theta \geq (h - Tx)^\top \hat{u},$$

které přidáváme v kroce 4 (řez optimality) duálního algoritmu. Tím pádem v konečně mnoha krocích najdeme v 5. kroku (přepočítání LP) duálního algoritmu řešení, které splňuje všechny podmínky dané právě těmito přípustnými bazickými řešeními $\hat{u}$:

$$\hat{\theta} \geq (h - Tx)^\top \hat{u} = f(\hat{x}).$$

Vzhledem k výše uvedenému, tak
a) množina přípustných řešení úlohy 3.2 je obsažena v množině přípustných řešení $\mathcal{B}_0 \cap \mathcal{B}_1$ v 5. kroce duálního algoritmu, který nám dá řešení ($\hat{x}$,$\hat{\theta}$).
b) Tohle řešení je také přípustné pro úlohu 3.2.
Nechť máme optimální řešení (x^*, θ^*) úlohy 3.2, pak

$$c^\top x^* + \theta^* = c^\top x^* + f(x^*)$$

$$\geq c^\top \hat{x} + \hat{\theta},$$

protože dle a) je množina přípustných řešení obsažena v množině, která dá optimální řešení $(\hat{x},\hat{\theta})$, tedy řešení úlohy 3.2 nemůže být menší. Dále

$$c^\top \hat{x} + \hat{\theta} \geq c^\top x^* + \theta^*,$$

protože $(\hat{x},\hat{\theta})$ je přípustné pro úlohu 3.2 dle b), tak optimální řešení této úlohy nemůže být větší než jiné přípustné řešení. Máme tedy (x^*,θ^*) na obou stranách nerovností, takže vše platí s rovností, takže algoritmus skutečně nalezl optimální řešení. Tedy algoritmus nám nalezne optimální řešení $\hat{x}$ prvního stupně úlohy 2.3 v konečně mnoha krocích.

□

3.3 Obecný případ

Nechť S>1, pak máme dvoustupňové SLP 2.3. Analogicky, jako pro S=1 můžeme zadefinovat rekurzivní funkce pro každý scénář (T^j,h^j):

$$f_j(x) = \min\{q^\top y^j : Wy^j = h^j - T^j x,\, y^j \geq 0\}, j = 1, \ldots, S.$$

Uvedeme si dvě možnosti jak modifikovat algoritmus z podkapitoly 3.1.

3.3.1 Varianta s více řezy

Ve druhé kapitole jsme si odvodili strukturu 2.3. Jednou z možností modifikace algoritmu z podkapitoly 3.1 je varianta s více řezy. Úlohu si, podobně jako tomu bylo ve speciálním případě, přepíšeme do tvaru:

$$\begin{array}{lrcll} & \min \{c^\top x \; + \; \sum_{j=1}^S p_j\theta_j\} & & & \\ \text{za podm.} & Ax & = & b & \\ & \theta_j \; - \; f_j(x) & \geq & 0, & j = 1, \ldots, S \\ & x & \geq & 0 & \end{array} \tag{3.6}$$

Postupujeme dle algoritmu z podkapitoly 3.1 tak, že každý krok, kde řešíme podmínky pro rekurzivní funkci, musíme mít splněné dané podmínky pro všechny rekurzivní funkce. S tím, že v druhém kroku algoritmu se definuje množina

$$J = \{j : \; f_j(\hat{x}) = \infty\}.$$

Pokud platí $J \neq \emptyset$, pak je potřeba řez přípustnosti. Tento řez se odvodí jen pro j-té omezení, $j \in J$. V opačném případě se provádí řez optimality.
Pokud $\forall j$ je splněna podmínka $f_j(\hat{x}) \leq \hat{\theta}_j$, pak jsme nalezli přípustné řešení. Pokud existuje j, že není podmínka splněna, pak pro dané j provedeme řez optimality a pokračujeme dalšími kroky algoritmu.

3.3.2 Varianta s jedním řezem

Druhým přístupem k řešení úlohy 2.3 je verze algoritmu s jedním řezem. Na rozdíl od předchozí kapitoly zavedeme pouze jednorozměrné θ a to následujícím způsobem:

$$\begin{array}{lrcll} & \min \{c^\top x \ +\theta\} & & & \\ \text{za podm.} & Ax & = & b & \\ & \theta \ - \ \sum_{j=1}^{S} p_j f_j(x) & \geq & 0, & j = 1, \dots, S \\ & x & \geq & 0 & \end{array} \tag{3.7}$$

Postupujeme podobně jako u verze s více řezy, tedy řešíme podmínky pro všechny rekurzivní funkce $f_j(x)$. Řez přípustnosti je potřeba dokud existuje j takové, že $f_j(\hat{x}) = \infty$.
Pokud ve čtvrtém kroku, tedy v řezu optimality, nastane

$$\hat{\theta} < \sum_{j=1}^{S} p_j \hat{u_j}^\top (h^j - T^j \hat{x}),$$

pak přidáme podmínku tvaru:

$$\theta \geq \sum_{j=1}^{S} p_j \hat{u_j}^\top (h^j - T^j x).$$

Následně pokračujeme podle algoritmu, tedy jdeme přepočítat úlohu LP.

3.3.3 Stochastický příklad

V této části si uvedeme jednoduchý příklad stochastického lineárního programovaní s náhodnými proměnnými diskrétního rozdělení. Pro nalezení optimálního řešení použijeme verzi s více řezy.
Nechť máme zadanou úlohu ve tvaru:

$$\min 2x_1 + x_2$$

$$\text{za podm.: } x_1 + x_2 = 9$$

$$T^j x = h^j, \ j = 1{,}2,$$

kde $(T^1, h^1) = ((1, 4){,}30)$ s pravděpodobností $\frac{2}{3}$ a $(T^2, h^2) = ((3{,}1){,}12)$ s pravděpodobností $\frac{1}{3}$. Pak podle kapitoly 2 si převedeme úlohu na tvar:

$$\begin{array}{lrcll} & \min 2x_1 + x_2 + \ \sum_{j=1}^{2} p_j (y_1^j + y_2^j) & & & \\ \text{za podm.} & x_1 + x_2 & = & 9 & \\ & T^j x \ + y_1^j - y_2^j & = & h^j, & j = 1{,}2 \\ & x & \geq & 0 & \\ & y^j & \geq & 0, & \end{array} \tag{3.8}$$

matici W jsme volili jako (1, -1) a vektor $q = (1,\ 1)^\top$, tedy jsme zvolili jednoduchou kompenzaci.
1. Krok

Definujeme dolní meze $\tilde{\theta}_j = 0$ pro $j = 1{,}2$
Řešíme úlohu LP:

$$\min\{2x_1 + x_2 + \frac{2}{3}\theta_1 + \frac{1}{3}\theta_2 :\ x_1 + x_2 = 9, x \geq 0, \theta \geq \tilde{\theta}\},$$

což dává řešení $\hat{x} = (0,\ 9)^\top$ a $\hat{\theta} = (0,\ 0)^\top$

2. Krok - vyhodnocení rekurzivních funkcí
Nejdříve

$$f_1(\hat{x}) = \min\{y_1^1 + y_2^1 : y_1^1 - y_2^1 = 30 - \hat{x}_1 - 4\hat{x}_2,\ y^1 \geq 0\},$$

kterou vyřešíme duální úlohou a dostaneme, že $f_1(\hat{x}) = 6$, čímž není splněna podmínka $\theta_1 - f_1(\hat{x}) \geq 0$. Proto přidáme podmínku:

$$\theta_1 \geq -30 + x_1 + 4x_2$$

Vyhodnotíme i funkci

$$f_2(\hat{x}) = \min\{y_1^2 + y_2^2 : y_1^2 - y_2^2 = 12 - 3\hat{x}_1 - \hat{x}_2,\ y^1 \geq 0\},$$

hodnota této funkce v bodě $\hat{x}$ je rovna 3, tedy není splněna podmínka a je potřeba řez optimality. Tímto řezem získáváme novou podmínku

$$\theta_2 \geq 12 - 3x_1 - x_2$$

3. Krok - přepočítání úlohy
Řešíme úlohu z kroku 1 rozšířenou o nové podmínky. Získáváme řešení

$$\hat{x} = (1.5,\ 7.5)^\top \text{a } \hat{\theta} = (1.5, 0)^\top$$

s hodnotou účelové funkce 11.

4.Krok - vyhodnocení rekurzivních funkcí
$f_1(\hat{x}) = 1{,}5$, splňuje podmínku optimality, neb $\hat{\theta} - f_1(\hat{x}) \geq 0$. A $f_2(\hat{x}) = 0$, což také splňuje podmínku optimality. Tedy jsme nalezli optimální řešení.

Závěr

Jsme na konci práce, ve které jsme se seznámili s Bendersovou dekompozicí a ilustrovali ji na konkrétních optimalizačních problémech. Zejména jsme se věnovali použití algoritmu na dvoustupňové stochastické lineární programování, i když tohle využití Benders v článku [1] původně nezamýšlel. Dekompoziční algoritmus byl vysvětlen ve tvaru, který je uveden v knize [3]. Následně byla ukázána modifikace, která se používá pro stochastické lineární programování, což bylo i ukázáno na konkrétním příkladu. Práce tak nejen vysvětlila myšlenku algoritmu v používání řezů přípustnosti a optimality, ale ukázala i další využití daného algoritmu.

Ve třetí kapitole jsme si mimo jiné ukázali vlastnosti rekurzivní funkce, používané v algoritmu. Následně byla dokázána konečnost algoritmu. V důkazech jsme používali poznatky z optimalizace, které jsme si, pro přehlednost práce, uvedli v první kapitole.

Práce by se dala více rozvíjet na další využití Bendersovy dekompozice. Jednou z možností by bylo řešení problému na smíšené lineární programování podle Bendersova původního záměru. Další aplikace by také mohla být na nelineární programování. V neposlední řadě by bylo zajímavé zkoumat efektivnost algoritmu pro klasické lineární programování, kde matice omezení má speciální blokovou strukturu, ve srovnání s jinými algoritmy typickými pro řešení lineárního programování.

Seznam použité literatury

[1] BENDERS, J. F. (1962). Partitioning procedures for solving mixed-variables programming problems. *Numerische Mathematik 4*, pages 238–252.

[2] DUPAČOVÁ, J. (1982). *Lineární programování.* Státní pedagogické nakladatelství.

[3] KALL, P. a MAYER, J. (2005). *Stochastic Linear Programming.* Springer. ISBN 0-387-23385-7.

[4] LACHOUT, P. a DUPAČOVÁ, J. (2011). *Úvod do optimalizace.* Matfyzpress. ISBN 978-80-7378-176-7.