MASARYKOVA UNIVERZITA

Prírodovedecká fakulta

Metódy kvadratického programovania

Bakalárska práca

Brno, 2009 Branislav Karniš

Poďakovanie

Ďakujem vedúcemu bakalárskej práce prof. RNDr. Ondřejovi Došlému, DrSc. za cenné rady a pomoc pri vedení bakalárskej práce. Tiež by som sa chcel poďakovať rodine, za podporu pri štúdiu.

Prehlásenie

Prehlasujem, že som túto bakalársku prácu vypracoval samostatne za použitia uvedenej literatúry.

....

Obsah

Úvod

Predkladaná bakalárska práca sa zaoberá problematikou matematického programovania, resp. úlohou kvadratického programovania, ktorá je jeho špecifickým prípadom. Optimalizovať úlohu kvadratického programovania znamená hľadať extrém konvexnej kvadratickej funkcie pri lineárnych obmedzujúcich podmienkach. Cieľom práce je teda popísať základné metódy riešenia takýchto úloh, pričom text obsahuje aj základy teórie matematického programovania.

Obsahovo je práca rozdelená na štyri kapitoly. Prvá kapitola je úvodná a objasňuje základné používané pojmy, predovšetkým je v nej formulovaný pojem úloha matematického programovania. Obsahom druhej kapitoly sú najdôležitejšie pojmy z teoretického jadra matematického programovania, predovšetkým pojem Kuhn-Tuckerove podmienky a teória duality v matematickom programovaní. Poznatky z tejto kapitoly sú potom aplikované na úlohu kvadratického programovania v tretej kapitole. V poslednej najobsiahlejšej kapitole sú popísané samotné metódy riešenia úloh kvadratického programovania (Metóda Hildrethova a d'Esopova, Metóda Theila a van de Panneho a Wolfeho metóda.) Pre názornosť je každá z popisovaných metód doplnená ilustračným príkladom.

Hlavným zdrojom poznatkov pri písaní tejto práce bola publikácia [1]. Takisto používané značenie je zhodné so značením v tejto knihe. Príslušné tvrdenia v texte sú uvádzané bez dôkazov, pričom v poznámke pod čiarou je vždy uvedený odkaz na literatúru, v ktorej možno daný dôkaz nájsť. Predpokladom k štúdiu textu sú základné znalosti matematickej analýzy a lineárnej algebry. Určitou výhodou je znalosť teórie lineárneho programovania predovšetkým pre časť venovanej Wolfeho metóde, ktorá pri riešení využíva upravený simplexový algoritmus lineárneho programovania.

Kapitola 1

Základné pojmy

1.1.Základné pojmy matematického programovania

Úlohu matematického programovania by sme mohli formulovať ako úlohu určiť maximum, (resp. minimum) nejakej funkcie *n*-premenných $f_0(x_1, x_2, \dots, x_n)$ vzhľadom na podmienky $f_j(x_1, x_2, \dots, x_n) \le b_j$, $j = 1, 2, \dots, m$. Ide teda o úlohu určiť *n*-čísel $x_1^0, x_2^0, \dots, x_n^0$ tak , aby hodnota $f_0(x_1^0, x_2^0, \dots, x_n^0)$ bola čo najväčšia (resp. čo najmenšia). Ak sú funkcie f_j, $j = 0,1,2, \dots, m$ lineárne hovoríme o úlohe *lineárneho programovania.* V prípade , že aspoň jedna z uvedených funkcií nie je lineárna, hovoríme o úlohe *nelineárneho programovania.*

Uvažujme funkciu $f : X \to \mathbb{R}$, kde $X \subset \mathbb{R}^n$. Ak funkcia *f(x)* má v bode $x^* \in X$ minimum , resp. v bode $x^{**} \in X$ maximum , potom funkcia *–f(x)* má v bode $x^* \in X$ maximum , resp. v bode $x^{**} \in X$ minimum , pričom platí :

$$-\min_{x \in X} f(x) = \max_{x \in X}\{-f(x)\} \qquad \text{resp.} \qquad -\max_{x \in X} f(x) = \min_{x \in X}\{-f(x)\}.$$

Uvedená vlastnosť nám umožňuje pri riešení extrémov funkcie sa formálne obmedziť napr. len na minimum funkcie. Tomu prispôsobíme nasledujúce úvahy. Základnú minimalizačnú úlohu budeme zapisovať v tvare

$$f(x) \to \min, \quad x \in X. \tag{1.1}$$

Funkciu f nazývame *cieľovou* (*účelovou*) *funkciou*, množinu X nazývame *prípustnou množinou,* body $x \in X$ budeme nazývať *prípustnými bodmi (prípustnými riešeniami*) a číslo $f^* = \inf_{x \in X} f(x)$ nazývame *hodnotou úlohy* (1.1).

Poznamenajme, že v prípade $X = \emptyset$, kladieme $f^* = \infty$

DEFINÍCIA 1.1. Bod $x^* \in X$ nazveme bodom *globálneho minima* funkcie f na X alebo tiež *riešením úlohy* (1.1), ak $f(x^*) \le f(x)$ pre všetky $x \in X$. Bod $x^* \in X$ nazveme bodom *lokálneho minima* alebo tiež *lokálnym riešením úlohy* (1.1), ak existuje okolie $O_\varepsilon(x^*) = \left\{ x \in \mathbb{R}^n : \left\| x - x^* \right\| < \varepsilon \right\}$ bodu x^* také, že $f(x^*) \le f(x)$ pre každé $x \in X \cap O_\varepsilon(x^*)$. V prípade ostrosti vyššie uvedených nerovností pre $x \ne x^*$ hovoríme o *ostrých* globálnych, resp. lokálnych minimách.

1.2.Klasifikácia extremálnych úloh

DEFINÍCIA 1.2. Konečno-rozmernú extremálnu úlohu $f(x) \to \min, \quad x \in X$ nazývame:

a) úlohou na voľný extrém, ak $X = \mathbb{R}^n$,

b) úlohou na viazaný extrém v užšom zmysle, resp. *klasickou úlohou na viazaný extrém,* ak množina X je určená sústavou m rovníc o n neznámych ($m < n$):

$$\begin{aligned} f_1(x_1, x_2, \ldots, x_n) &= b_1 \\ f_2(x_1, x_2, \ldots, x_n) &= b_2 \\ &\ldots\ldots\ldots\ldots \\ f_m(x_1, x_2, \ldots, x_n) &= b_m \end{aligned}$$

Skrátene zapísané $X = \left\{ x \in \mathbb{R}^n, f_j(x) = b_j, j = 1, 2, \ldots, m \right\}$,

c) úlohou matematického programovania v užšom zmysle, ak množina X je určená sústavou m nerovníc, t.j.

$$X = \left\{ x \in \mathbb{R}^n, f_j(x) \le b_j, j = 1, 2, \ldots, m \right\}, \tag{1.2}$$

d) úlohou matematického programovania, resp. *úlohou matematického programovania v širšom zmysle,* ak množina X je určená nejakou sústavou rovníc a nerovníc, t.j.

$$X = \left\{ x \in \mathbb{R}^n, \quad \begin{matrix} f_j(x) \le b_j, j = 1, 2, \ldots, r \\ f_k(x) = b_k, k = r+1, \ldots, m \end{matrix} \right\}.$$

Podľa vlastností funkcií $f_j(x), j = 1, 2, \ldots, m,$ možno úlohy matematického programovania klasifikovať podrobnejšie. Základné členenie spočíva v rozdelení na dve skupiny :

I. *Lineárne úlohy* , ak všetky funkcie $f(x), f_j(x), j = 1, 2, \ldots, m$ sú lineárne.

II. *Nelineárne úlohy,* ak aspoň jedna z funkcií $f(x), f_j(x), j = 1, 2, \ldots, m$ nie je lineárna.

Nelineárne úlohy možno ďalej rozdeliť na:

1.*Konvexné úlohy* , ak všetky funkcie $f(x), f_j(x), j = 1, 2, \ldots, m,$ v (1.2) sú konvexné. Do tejto oblasti patria aj úlohy lineárneho programovania, keďže lineárna funkcia je konvexná.

2.*Nekonvexné úlohy* , ak aspoň jedna z funkcií $f(x), f_j(x), j = 1, 2, \ldots, m,$ nie je konvexná.

Kapitola 2

Nutné a dostatočné podmienky optimality

Uvažujme úlohu matematického programovania

$$f(x) \to \min, \quad x \in X, \tag{2.1}$$

kde prípustná množina X je zadaná systémom rovností a nerovností

$$X = \{x \in P \subseteq \mathbb{R}^n : g_i(x) \le 0, i = 1,\dots,k, \tag{2.2}$$

$$g_j(x) = 0,\ j = k+1,\dots,m\}.$$

Obmedzenie $x \in P$ sa nazýva *priame obmedzenie* a obmedzenia $g_i(x) \le 0, i = 1,\dots,k$, $g_j(x) = 0,\ j = k+1,\dots,m$ sa nazývajú *funkcionálne obmedzenia.* $f^* = \inf\limits_{x \in X} f(x)$ je *hodnota úlohy* (2.1) , (2.2).

2.1.Lagrangeov princíp a Kuhn-Tuckerove podmienky

Lagrangeovou funkciou úlohy (2.1) , (2.1) nazývame funkciu $L : P \times \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}$, ktorá je definovaná predpisom

$$L(x, y_0, y) = y_0 f(x) + \sum_{i=1}^{m} y_i g_i(x) \tag{2.3}$$

Ďalej označme:

$$Q = \{y = (y_1,\dots y_m) \in \mathbb{R}^m : y_1,\dots y_k \ge 0\}$$

a pre $x^* \in X$

$$I(x^*) = \{i = \{1,\dots,k\}, g_i(x^*) = 0\}.$$

$I(x^*)$ označuje tzv. *množinu aktívnych nerovností v bode* x^*.Obsahuje indexy tých nerovností $g_i(x) \le 0, i = 1,\dots,k$ definujúcich množinu X, ktoré sa v bode x^* realizujú ako rovnosti. Číslo y_0 a prvky y_i vektora y sa nazývajú *Lagrangeove multiplikátory.*

Nasledujúca veta (Lagrangeov princíp) udáva *nutnú podmienku* pre lokálny extrém úlohy (2.1) , (2.2).

VETA 2.1. (Lagrangeov princíp)[1] Nech množina $P \subseteq \mathbb{R}^n$ je konvexná, funkcie $f,\ g_i,\ i = 1,\dots,k$ sú diferencovateľné v bode $x^* \in X$ a $g_j,\ j = k+1,\dots,m$, sú spojito diferencovateľné v nejakom okolí bodu x^*. Ak je bod x^* lokálnym extrémom úlohy (2.1) , (2.2) potom existujú Lagrangeove multiplikátory $y_0^*, y_1^*,\dots,y_k^* \geq 0,\ y_{k+1}^*,\dots,y_m^* \in \mathbb{R}$ tak, že nie všetky $y_0^*, y^* = (y_1^*,\dots,y_m^*)$ sú nulové a

$$\left\langle L'_x\left(x^*, y_0^*, y^*\right), x - x^* \right\rangle \geq 0, \quad \forall x \in P, \tag{2.4}$$

$$y_i^* g_i(x^*) = 0, \quad i = 1,\dots,m \tag{2.5}$$

Podmienka (2.4) znamená, že bod x^* je *stacionárny bod* príslušnej Lagrangeovej funkcie[2]. Vzťahy (2.5) nazývame *podmienky komplementarity.* Znamenajú, že Lagrangeove multiplikátory prislúchajúce pasívnym nerovnostiam v bode x^* sú nulové. Súhrne podmienky (2.4) a (2.5) označujeme ako *Kuhn-Tuckerove podmienky* úlohy (2.1) , (2.2) .

Pridaním dodatočných podmienok konvexností funkcií $f,\ g_1,\dots,g_k$ a afinity funkcií $g_{k+1},\dots,g_m$, sa Kuhn-Tuckerove podmienky (2.4) a (2.5) stávajú *dostatočnými* pre existenciu globálneho riešenia úlohy (2.1) , (2.2)

VETA 2.2.[3] Nech množina P je konvexná, funkcie $f,\ g_1,\dots,g_k$ sú konvexné a diferencovateľné na P , funkcie $g_{k+1},\dots,g_m$ sú afinné a nech platia podmienky (3.4) a (3.5) s $y_0^* = 1$. Potom x^* je globálnym riešením úlohy (2.1)

VETA 2.3. (Kuhna-Tuckera v diferenciálnom tvare)[4] Nech P je konvexná, funkcie $f, g_1,\dots,g_k$ sú konvexné a diferencovateľné, $g_{k+1},\dots,g_m$ sú afinné a nech platí aspoň jedna z podmienok regularity:

(i) *Slaterova:* Funkcionálne obmedzenia sú len v tvare nerovností, t.j. $k = m$, a existuje $\overline{x} \in P$ také, že $g_i(\overline{x}) < 0, i = 1,\dots,k$;

(ii) *Lineárna:* Množina P je polyéder, funkcie $g_1,\dots,g_k$ sú afinné;

[1] Dôkaz: viď. [1] str. 65
[2] Definícia: viď. [1] str. 62
[3] Dôkaz: viď. [1] str.66
[4] Dôkaz: viď. [1] str.66

(iii) *Modifikovaná Slaterova:* Existuje $l \in \{0,\dots,k\}$ také, že funkcie $g_{l+1},\dots,g_k$ sú afinné, $\exists \overline{x} \in \operatorname{ri} P$[5] tak, že $g_i(\overline{x}) < 0, i = 1,\dots,l$.

Potom x^* je riešenie (2.1) , (2.2), práve vtedy keď existuje $y^* \in Q$ tak, že platia podmienky (2.4) a (2.5) s $y_0^* = 1$.

2.2.Dualita v úlohách matematického programovania

DEFINÍCIA 2.4. Vektor $y \in Q$ sa nazýva *vektor Kuhna-Tuckera* úlohy (2.1) , (2,2) ak $f^* \le f(x) + \sum_{i=1}^{m} y_i g_i(x) = L(x,y)$ pre všetky $x \in P$.

Postačujúce podmienky existencie vektora Kuhna-Tuckera udáva nasledujúca veta:

VETA 2.5.[6] Nech P je konvexná množina, $f, g_1,\dots,g_k$ sú konvexné a $g_{k+1},\dots,g_m$ sú afinné a platí aspoň jedna z podmienok regularity:

(i) *Slaterova:* $k = m$ a existuje $\overline{x} \in P$ tak, že $g_i(\overline{x}) < 0,\ i = 1,\dots,k$.

(ii) *Lineárna: P* je polyéder, funkcie $f, g_1,\dots,g_k$ sú afinné a $X \neq \emptyset$.

(iii) *Modifikovaná Slaterova: P* je polyéder, existuje index $l \in \{0,\dots,k\}$ taký, že funkcie f a $g_1,\dots,g_l$ sú konvexné na relatívne otvorenej množine $U \supset P$, funkcie $g_{l+1},\dots,g_k$ sú lineárne a existuje $\overline{x} \in \operatorname{ri} P$, také, že $g_i(\overline{x}) < 0$, $i = 1,\dots,l$.

Potom existuje vektor Kuhna-Tuckera úlohy (2.1) , (2.2).

Prejdime k samotnej definícii duálnej úlohy matematického programovania.

DEFINÍCIA2.6. Nech $y \in Q$. Definujme funkciu

$$\varphi(y) = \inf_{x \in P} L(x,y) = \inf_{x \in P}\left[f(x) + \sum_{i=1}^{m} y_i g_i(x) \right].$$

Označme ďalej $Y = \{ y \in Q : \varphi(y) > -\infty \}$. *Duálnou úlohou k úlohe* (2.1), (2.2) nazveme úlohu

$$\varphi(y) \to \max, \qquad y \in Y, \tag{2.6}$$

a veličinu

---

5 Množina **ri *X*** sa nazýva *relatívny vnútrajšok množiny X.* Predstavuje množinu všetkých relatívne vnútorných bodov množiny X. Bod $x \in X$ sa nazýva *relatívne vnútorným bodom* množiny X, ak existuje okolie $O(x)$ bodu x tak, že $O(x) \cap \operatorname{aff} X \subseteq X$, kde aff *X* je afinný obal množiny *X*.

6 Dôkaz: viď. [1] str.70

$$\varphi^* = \sup_{y \in Y} \varphi(y)$$

nazveme *hodnotou duálnej úlohy* (2.6).

POZNÁMKA 2.7. Úloha (2.1) , (2.2) sa v tomto kontexte označuje ako *primárna úloha.*

Ústredným problémom teórie duality v úlohách matematického programovania je dokázať tzv. *vzťah duality*:

$$\inf_{x \in X} f(x) = f^* = \varphi^* = \sup_{y \in Y} \varphi(y). \tag{2.7}$$

Hodnota primárnej úlohy sa teda má rovnať hodnote duálnej úlohy. Nasledujúca veta udáva postačujúcu podmienku pre vzťah duality.

VETA 2.8.[7] Nech (2.1) , (2.2) je *regulárna úloha konvexného programovania*, t.j. P je konvexná množina, $f, g_1, \ldots, g_k$ sú konvexné, $g_{k+1}, \ldots, g_m$ sú afinné a platí aspoň jedna z *podmienok regularity* (i)-(iii) uvedených vo Vete 2.5. Ak $f^* > -\infty$, potom platí *vzťah duality* (2.7), pričom množina riešení úlohy (2.6) je rovná množine všetkých vektorov Kuhna-Tuckera úlohy (2.1) , (2.2).

Priamym dôsledkom tejto vety je nasledujúce tvrdenie:

DÔSLEDOK 2.9. Nech (2.1) , (2.2) je regulárna úloha konvexného programovania a Y je prípustná množina úlohy (2.6). Potom platí:

(i) Ak je množina Y neprázdna, potom je duálna úloha riešiteľná a $f^* > -\infty$.

(ii) Ak je množina $Y = \emptyset$, potom je $f^* = -\infty$.

Na záver tejto kapitoly ešte uvedieme vetu, ktorá býva označovaná ako Kuhnova-Tuckerova veta v nediferenciálnom tvare.

VETA 2.9.[8] Nech (2.1) , (2.2) je regulárna úloha konvexného programovania. Potom $x^* \in X$ je riešenie tejto úlohy práve vtedy, keď platí aspoň jedna z nasledujúcich podmienok:

(i) Existuje $y^* \in Q$ také, že $\varphi(y^*) = f(x^*)$;

(ii) Existuje $y^* \in Q$ také, že $L(x^*, y^*) = \min_{x \in P} L(x, y^*)$ a

$$y_i^* g_i^*(x^*) = 0, \quad i = 1, \ldots, k.$$

[7] Dôkaz: viď. [1] str.74

[8] Dôkaz: viď. [1] str.75

Kapitola 3

Úloha kvadratického programovania

Obsahom tejto kapitoly bude aplikácia teoretických záverov z predchádzajúcej kapitoly na úlohu kvadratického programovania. Najprv teda uvedieme Kuhn-Tuckerove podmienky pre úlohu kvadratického programovania, v ďalšej časti prejdeme k duálnej úlohe.

3.1. Kuhn-Tuckerove podmienky pre úlohu kvadratického programovania

Úloha kvadratického programovania predstavuje úlohu konvexného programovania optimalizovať kvadratickú účelovú funkciu s lineárnymi ohraničeniami.

DEFINÍCIA 3.1. Nech matica $C \in \mathbb{R}^n$ je symetrická a pozitívne semidefinitná a nech $d,\, a_i \in \mathbb{R}^n,\, b_i \in \mathbb{R},\, i = 1,\dots,m$. Označme:

$$P = \left\{ x \in \mathbb{R}^n,\ x_1,\dots,x_s \geq 0,\ s \in \left\{0,\dots,n\right\} \right\},$$

$$X = \left\{ x \in P;\ \left\langle a_i, x \right\rangle \leq b_i,\ i = 1,\dots,k;\ \left\langle a_i, x \right\rangle = b_i,\ i = k+1,\dots,m \right\}.$$

Potom *úlohou kvadratického programovania* nazývame úlohu

$$f(x) = \frac{1}{2}\left\langle Cx, x \right\rangle + \left\langle d, x \right\rangle \to \min, \quad x \in X \tag{3.1}$$

Ďalej budeme požívať označenie: $A = \left[a_1,\dots,a_m\right]^T$, kde $a_1,\dots,a_m \in \mathbb{R}^n$ sú riadky matice A, $b = \left[b_1,\dots,b_m\right]^T$. V súlade s definíciou z Kapitoly 2 môžeme Lagrangeovu funkciu úlohy (3.1) vyjadriť v tvare

$$L(x,y) = \frac{1}{2}\left\langle Cx, x \right\rangle + \left\langle d, x \right\rangle + \left\langle y, Ax - b \right\rangle = \frac{1}{2}\left\langle Cx, x \right\rangle + \left\langle A^T y + d, x \right\rangle - \left\langle y, b \right\rangle,$$

a teda

$$L'_x(x,y) = Cx + A^T y + d.$$

Poznamenajme, že $y \in Q = \{ y = (y_1, \dots, y_m), y_1, \dots, y_k \geq 0 \}$.

Využitím poznatkov z Kapitoly 2 , hlavne Vety 2.3. môžeme odvodiť nasledujúce tvrdenie:

VETA 3.2. Bod x^* je riešením úlohy (3.1) práve vtedy, keď existuje $y^* \in Q$ tak, že sú splnené Kuhn-Tuckerove podmienky, t.j.

$$\left\langle Cx^* + A^T y^* + d, x - x^* \right\rangle \geq 0, \quad \forall x \in P \tag{3.2}$$

$$y_i^* (Ax^* - b)_i = 0, \quad i = 1, \dots, m \tag{3.3}$$

Z teoretického i praktického hľadiska stačí ak sa pri skúmaní úlohy kvadratického programovania obmedzíme na nasledujúce (najčastejšie sa vyskytujúce) typy množiny *X:*

(i) $X_I = \{ x \mid Ax \leq b, x \geq 0 \}$

(ii) $X_{II} = \{ x \mid Ax = b, x \geq 0 \}$

(iii) $X_{III} = \{ x \mid Ax \leq b \}$

Pre tieto tri typy úloh s rôznymi obmedzeniami teraz odvodíme príslušné Kuhn-Tuckerove podmienky.

(i) $\frac{1}{2}\langle Cx, x \rangle + \langle d, x \rangle \to \min, \quad Ax \leq b, x \geq 0:$

Podmienku (3.2) možno prepísať na tvar

$$Cx^* + A^T y^* + d \geq 0,$$

$$\left\langle x^*, Cx^* + A^T y^* + d \right\rangle = 0.$$

Podmienka (3.3) je ekvivalentná podmienke $\left\langle y^*, b - Ax^* \right\rangle = 0$, čím sa hľadanie riešenia úlohy (3.1) stáva ekvivalentné riešeniu sústavy:

$$\begin{aligned} Cx + A^T y + d &\geq 0, \\ \langle x, Cx + A^T y + d \rangle &= 0, \\ \langle y, b - Ax \rangle &= 0, \\ Ax \leq b, \quad x, y &\geq 0, \end{aligned} \quad \Longleftrightarrow \quad \begin{aligned} Cx + A^T y - u &= -d \\ Ax + v &= b, \\ \langle x, u \rangle + \langle y, v \rangle &= 0, \\ x, y, u, v &\geq 0. \end{aligned}$$

pri zavedení označenia $u = Cx + A^T y + d$ a $v = b - Ax$.

(ii) $\frac{1}{2}\langle Cx, x \rangle + \langle d, x \rangle \to \min, \quad Ax = b, x \geq 0:$

Podmienka (3.2) je rovnaká ako v predošlom prípade a podmienka (3.3) je splnená vždy, lebo $Ax = b$. Hľadaná sústava má teda podobu:

$$\begin{aligned} Cx + A^T y + d &\geq 0 \\ \left\langle x, Cx + A^T y + d \right\rangle &= 0\,, \\ Ax &= b\,, \\ x &\geq 0\,, \end{aligned} \quad \Leftrightarrow \quad \begin{aligned} Cx + A^T y - u &= -d\,, \\ Ax &= b\,, \\ \left\langle x, u \right\rangle &= 0\,, \\ x, u &\geq 0\,. \end{aligned}$$

(iii) $\frac{1}{2}\left\langle Cx, x \right\rangle + \left\langle d, x \right\rangle \to \min, \quad Ax \leq b$:

Pretože $P = \mathbb{R}^n$, môžeme namiesto podmienky (3.2) písať $Cx^* + A^T y^* + d = 0$, podmienka (3.3) je rovnaká ako v prvom prípade. Hľadáme teda riešenie systému:

$$\begin{aligned} Cx + A^T y + d &= 0\,, \\ \left\langle y, b - Ax \right\rangle &= 0\,, \\ Ax &\leq b\,, \\ y &\geq 0\,, \end{aligned} \quad \Leftrightarrow \quad \begin{aligned} Cx + A^T y &= -d\,, \\ Ax + v &= b\,, \\ \left\langle y, v \right\rangle &= 0\,, \\ y, v &\geq 0\,. \end{aligned}$$

Množiny X_I, X_{II}, X_{III} sú ekvivalentné v tom zmysle, že formálnymi úpravami (predovšetkým zavedením doplnkových premenných , zmenou rovnosti na dve nerovnosti apod.) môžeme prejsť z jedného spôsobu zápisu na druhý. Prípad, v ktorom je množina obmedzení v tvare $Ax = b$ možno riešiť metódou Lagrangeových multiplikátorov ako úlohu na viazaný extrém, alebo prevedením úlohy na úlohu duálnu, ktorá je v tomto prípade (ako ďalej uvedieme) bez akýchkoľvek obmedzení.

3.2.Dualita v úlohách kvadratického programovania

Využijeme výsledky získané v predošlej kapitole. Uvažujme úlohu kvadratického programovania

$$\frac{1}{2}\left\langle Cx, x \right\rangle + \left\langle d, x \right\rangle \to \min, \quad Ax \leq b.$$

Duálnu úlohu vyjadríme v tvare

$$\varphi(y) \to \max, \quad y \geq 0, \text{ kde } \varphi(y) = \inf_{x \in \mathbb{R}^n} L(x, y)\,.$$

Je to teda úloha:

$$\frac{1}{2}\left\langle Cx, x \right\rangle + \left\langle A^T y + d, x \right\rangle - \left\langle y, b \right\rangle \to \max, \quad Cx + A^T y + d = 0, y \geq 0,$$

t.j.

$$-\frac{1}{2}\left\langle Cx, x \right\rangle - \left\langle b, y \right\rangle \to \max, \quad Cx + A^T y = -d, \, y \geq 0\,.$$

Aplikovaním všeobecných tvrdení z Kapitoly 2 na primárnu a duálnu úlohu kvadratického programovania získame nasledujúce tvrdenie:

VETA 3.3.

(i) Ak je prípustný obor jednej z úloh prázdny, tak u druhej je prípustný obor prázdny alebo účelová funkcia nie je nad prípustným oborom zdola (zhora) ohraničená.

(ii) Ak nie je u jednej z úloh účelová funkcia nad prípustným oborom zdola (zhora) ohraničená, tak u druhej je prípustný obor prázdny.

(iii) Ak má jedná z úloh riešenie, má riešenie aj druhá a hodnoty oboch úloh sú rovnaké.

V prípade, že matica C je pozitívne definitná, tak existuje inverzná matica C^{-1} a premennú x môžeme z duálnej úlohy vylúčiť.

$$L'_x(x,y) = Cx + A^T y + d = 0 \;\Rightarrow\; x = -C^{-1}(A^T y + d).$$

Úpravami dostaneme

$$\varphi(y) = L(-C^{-1}(A^T y + d), y) = \frac{1}{2}\left\langle A^T y + d, -C^{-1}(A^T y + d)\right\rangle - \left\langle y, b\right\rangle =$$

$$-\frac{1}{2}\left\langle AC^{-1}A^T y, y\right\rangle - \left\langle AC^{-1}d + b, y\right\rangle - \frac{1}{2}\left\langle C^{-1}d, d\right\rangle.$$

V prípade, že obmedzenia úlohy sú v tvare $Ax = b$, tak duálna úloha je úloha nájsť maximum kvadratickej funkcie s negatívne semidefinitnou maticou *bez obmedzenia* ,t.j.

$$\varphi(y) \to \max, y \in \mathbb{R}^m \quad \Leftrightarrow \quad -\varphi(y) \to \min, y \in \mathbb{R}^m.$$

Ak je y^* jej riešením , potom $x^* = -C^{-1}(A^T y^* + d)$ je riešením úlohy $f(x) \to \min, Ax = b.$

Kapitola 4

Metódy kvadratického programovania

V tejto kapitole sa dostaneme k samotným metódam riešenia úloh kvadratického programovania. Keďže v súčasnosti existuje veľké množstvo metód obmedzíme sa len základné využívané metódy, konkrétne metódu Hildrethovu a d´Esopovu, metódu Theila a van de Panneho a Wolfeho metódu.

4.1.Metóda Hildrethova a d´Esopova

Táto metóda riešenia úloh kvadratického programovania je založená na využívaní duálnej úlohy a jej vlastností. Metódu je vhodné použiť v prípadoch , keď počet premenných prevyšuje počet ohraničení (prípadne podmienok nezápornosti). Uvažujme úlohu zadanú v tvare:

$$\frac{1}{2}\langle Cx, x\rangle + \langle d, x\rangle \to \min, \quad Ax \le b, \tag{4.1}$$

kde matica C je pozitívne definitná.

Za týchto predpokladov môžeme duálnu úlohu k tejto úlohe vyjadriť ako úlohu

$$\phi(y) = \frac{1}{2}\langle Gy, y\rangle + \langle h, y\rangle \to \min, \quad y \ge 0, \tag{4.2}$$

pri zavedení označenia $G = AC^{-1}A^T$ a $h = AC^{-1}d + b$. Funkcia $\phi(y)$ odpovedá funkcii $-\varphi(y)$ až na konštantu $\frac{1}{2}\langle C^{-1}d, d\rangle$, ktorá však nemá na riešenie vplyv. Poznamenajme ešte, že matica G je pozitívne semidefinitná s kladnými diagonálnymi prvkami $g_{ii} = \langle C^{-1}a_i, a_i\rangle, i = 1,\dots,m.$

Iteratívny algoritmus metódy najprv nájde riešenie y^* duálnej úlohy a pomocou neho potom vypočítame riešenie pôvodnej primárnej úlohy x^*, pre ktoré platí $x^* = -C^{-1}(A^T y^* + d)$.

K predpokladom ešte musíme zaradiť predpoklad zaisťujúci konvergenciu metódy úlohy (4.1). Musia existovať také prípustné body, splňujú podmienky:

$$\langle a_i, x \rangle \le b_i - \varepsilon, \quad i = 1, \dots, m,$$

Kde $\varepsilon > 0$ je ľubovoľne malé, a_i sú riadky matice A a b_i sú zložky vektora b. Pri splnení týchto predpokladov o účelovej funkcii a obmedzeniach úlohy má úloha (4.1) jednoznačné riešenie x^*, pretože C je pozitívne definitná, účelová funkcia je teda silne konvexná. Ak je matica G pozitívne definitná, tak je aj y^* určené jednoznačne.

4.1.1.Algoritmus

Postup začína v ľubovoľnom prípustnom bode $y_0 \ge 0$, pričom vhodné je voliť $y^0 = (0, \dots, 0)^T$ vzhľadom k tomu, že v konečnom riešení y^* sú spravidla niektoré zložky nulové. Zložky ďalšieho bodu $y^1 = (y_1^1, \dots, y_m^1)$ určíme postupne ako výsledok m-násobnej minimalizácie funkcie $\phi(y)$ vzhľadom k jednotlivým komponentom y_j za podmienky $y_j \ge 0$. Ostatné zložky sú pritom fixované hodnotami, ktoré nadobudli pri poslednej minimalizácii. To znamená, že pre $i > j$ má i-tá zložka hodnotu y_i^0 a pre $i < j$ hodnotu y_i^1. Obdobným spôsobom sa získa y^2 z y^1 atd. Predpokladajme, že sme už zostrojili $y^0, y^1, \dots, y^{k-1}$.

1.krok: Riešime úlohu

$$\phi(t, y_2^{k-1}, \dots, y_m^{k-1}) \to \min, \quad t \ge 0$$

Keďže funkcia ϕ je konvexná vzhľadom k t, nadobúda svoje minimum pre $t \ge 0$ buď v $t = 0$ alebo v stacionárnom bode. Minimum t_1^k určíme pomocou prvej derivácie. Riešime teda rovnicu $\frac{\partial \phi}{\partial y_1}(t, y_2^{k-1}, \dots, y_m^{k-1}) = 0$. Vzhľadom na to, že minimum funkcie ϕ môže byť alebo vo vnútri prípustnej oblasti , alebo na hranici prípustnej oblasti, musíme rozlišovať dva prípady:

a) ak $t_1^k \ge 0$, vtedy $y_1^k = t_1^k$ (minimum vo vnútri)

b) ak $t_1^k < 0$, vtedy $y_1^k = 0$ (minimum na hranici)

Ak je teda t_1^k riešením rovnice $\frac{\partial \phi}{\partial y_1}(t, y_2^{k-1}, \dots, y_m^{k-1}) = 0$ tak potom $y_1^k = \max(0, t_1^k)$.

***j*-ty krok:** Riešime úlohu

$$\phi(y_1^k, \dots, y_{j-1}^k, t, y_{j+1}^{k-1}, \dots, y_m^{k-1}) \to \min, \quad t \ge 0.$$

Podobne ako v prvom kroku, ak je t_j^k riešením rovnice $\frac{\partial \phi}{\partial y_j}(y_1^k, \dots, y_{j-1}^k, t, y_{j+1}^{k-1}, \dots, y_m^{k-1}) = 0$, tak potom $y_j^k = \max(0, t_j^k)$.

m-tý krok: Riešime úlohu

$$\phi(y_1^k, y_2^k, \dots, y_{m-1}^k, t) \to \min, \quad t \ge 0.$$

Ak je t_m^k riešením rovnice $\frac{\partial \phi}{\partial y_m}(y_1^k, y_2^k, \ldots, y_{m-1}^k, t) = 0$ tak potom platí $y_m^k = \max(0, t_m^k)$.

Pritom v j-tom kroku je stacionárny bod v tvare:

$$t_j^k = -\frac{1}{g_{jj}}\left(\sum_{i=1}^{j-1} g_{ji} y_i^k + h_j + \sum_{i=j+1}^{m} g_{ji} y_i^{k-1}\right), \tag{4.3}$$

o čom sa možno presvedčiť priamym výpočtom derivovaním $\frac{\partial \phi}{\partial y_j} = \sum_{i=1}^{m} g_{ji} y_i + h_j$ a riešením rovnice $\frac{\partial \phi}{\partial y_j}(y_1^k, \ldots, y_{j-1}^k, t, y_{j+1}^{k-1}, \ldots, y_m^{k-1}) = 0$.

Zrejme platí $\phi(y^k) \le \phi(y^{k-1})$ teda postupnosť $\phi(y^k)$ je klesajúca. V iteráciách pokračujeme až dovtedy, kým sa y^k nemení v medziach presnosti výpočtu. Pritom v jednotlivých iteráciách je možné zameniť poradie výpočtu komponentov y^k. To má vplyv len na rýchlosť konvergencie, keďže množinu indexov $\{1, \ldots, m\}$ možno ľubovoľne permutovať pričom princíp metódy sa nemení.

VETA 4.1.[9] Nech $y^0 \ge 0$ je ľubovoľné a postupnosť $\phi(y^k)$ je zostrojená vyššie popísaným spôsobom. Potom $\phi(y^k)$ konverguje k minimálnej hodnote $\phi^* = \phi(y^*)$, y^k konverguje k niektorému riešeniu y^* a $x(y^k) := -C^{-1}(A^T y^k + d)$ konverguje k riešeniu $x^* = x(y^*)$.

4.1.2. Ilustračný príklad[10]

Riešte Hildrethovou a d´Esopovou metódou úlohu:

$$f(x) = 2x_1^2 - 6x_1x_2 + 5x_2^2 - 16x_1 + 40x_2 \to \min, \quad \text{pri obmedzeniach} \quad \begin{aligned} x_1 + x_2 &\ge 3, \\ 2x_1 - x_2 &\ge 2, \\ x_1, x_2 &\ge 0. \end{aligned}$$

Riešenie: Úlohu najprv vyjadríme tvare $\frac{1}{2}\langle Cx, x\rangle + \langle d, x\rangle \to \min, \quad Ax \ge b,$ kde

$$C = \begin{pmatrix} 4 & -6 \\ -6 & 10 \end{pmatrix}, \quad d = \begin{pmatrix} -16 \\ 40 \end{pmatrix}, \quad A = \begin{pmatrix} -1 & -1 \\ -2 & 1 \\ -1 & 0 \\ 0 & -1 \end{pmatrix}, \quad b = \begin{pmatrix} -3 \\ -2 \\ 0 \\ 0 \end{pmatrix}, \quad x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

Ďalším krokom bude vyjadrenie duálnej úlohy, ktorá má podľa (4.2) tvar $\phi(y) = \frac{1}{2}\langle Gy, y\rangle + \langle h, y\rangle \to \min, \quad y \ge 0$, kde

[9] Dôkaz: viď.[1] str.135
[10] [2] str.137

$$G = AC^{-1}A^T = \begin{pmatrix} 6,5 & 5,5 & 4 & 2,5 \\ 5,5 & 5 & 3,5 & 2 \\ 4 & 3,5 & 2,5 & 1,5 \\ 2,5 & 2 & 1,5 & 1 \end{pmatrix}, \quad h = AC^{-1}d + b = \begin{pmatrix} -39 \\ -26 \\ -20 \\ -16 \end{pmatrix}, \quad y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix}.$$

Zvolíme $y^0 = (0,0,0,0)^T$ a zahájime výpočet podľa vyššie uvedeného algoritmu a s využitím vzťahu (4.3).

Vypočítame prvú iteráciu y^1:

$$t_1^1 = -\tfrac{1}{6,5}\left(-39 + 5,5\cdot 0 + 4\cdot 0 + 2,5\cdot 0\right) = 6 \geq 0 \quad \Rightarrow y_1^1 = 6$$
$$t_2^1 = -\tfrac{1}{5}\left(5,5\cdot 6 - 26 + 3,5\cdot 0 + 2\cdot 0\right) = -1,4 < 0 \quad \Rightarrow y_2^1 = 0$$
$$t_3^1 = -\tfrac{1}{2,5}\left(4\cdot 6 + 3,5\cdot 0 - 20 + 1,5\cdot 0\right) = -1,6 < 0 \quad \Rightarrow y_3^1 = 0$$
$$t_4^1 = -\left(2,5\cdot 6 + 2\cdot 0 + 1,5\cdot 0 - 16\right) = 1 \geq 0 \quad \Rightarrow y_4^1 = 1$$

Teda $y^1 = (6,0,0,1)^T$. Pokračujeme výpočtom y^2:

$$t_1^2 = -\tfrac{1}{6,5}\left(-39 + 5,5\cdot 0 + 4\cdot 0 + 2,5\cdot 1\right) = 5,6 \geq 0 \quad \Rightarrow y_1^2 = 5,6$$
$$t_2^2 = -\tfrac{1}{5}\left(5,5\cdot 5,6 - 26 + 3,5\cdot 0 + 2\cdot 0\right) = -1,36 < 0 \quad \Rightarrow y_2^2 = 0$$
$$t_3^2 = -\tfrac{1}{2,5}\left(4\cdot 5,6 + 3,5\cdot 0 - 20 + 1,5\cdot 1\right) = -1,56 < 0 \quad \Rightarrow y_3^2 = 0$$
$$t_4^1 = -\left(2,5\cdot 5,6 + 2\cdot 0 + 1,5\cdot 0 - 16\right) = 2 \geq 0 \quad \Rightarrow y_4^1 = 2$$

Odtiaľ teda dostávame $y^2 = (5,6\,;0;0;2)^T$. Vyrátame tretiu iteráciu y^3:

$$t_1^3 = -\tfrac{1}{6,5}\left(-39 + 5,5\cdot 0 + 4\cdot 0 + 2,5\cdot 2\right) = 5,2 \geq 0 \quad \Rightarrow y_1^3 = 5,2$$
$$t_2^3 = -\tfrac{1}{5}\left(5,5\cdot 5,2 - 26 + 3,5\cdot 0 + 2\cdot 2\right) = -1,32 < 0 \quad \Rightarrow y_2^3 = 0$$
$$t_3^3 = -\tfrac{1}{2,5}\left(4\cdot 5,2 - 3,5\cdot 0 - 20 + 1,5\cdot 2\right) = -1,52 < 0 \quad \Rightarrow y_3^3 = 0$$
$$t_4^1 = -\left(2,5\cdot 5,2 + 2\cdot 0 + 1,5\cdot 0 - 16\right) = 3 \geq 0 \quad \Rightarrow y_4^3 = 3$$

Dostávame $y^3 = (5,2;0;0;3)$.Princíp výpočtu je zrejme jasný, preto ďalšie iterácie nebudeme takto podrobne rozpisovať, ale zobrazíme ich v tabuľke:

y^0	0	0	0	0
y^1	6	0	0	1
y^2	5,6	0	0	2
y^3	5,2	0	0	3
y^4	4,8	0	0	4
y^5	4,4	0	0	5
y^6	4	0	0	6
y^7	3,7	0	0	6,8
y^8	3,4	0	0	7,5
y^9	3,1	0	0	8,3
y^{10}	2,8	0	0	9
y^{11}	2,5	0	0	9,8
y^{12}	2,2	0	0	10,5
y^{13}	2	0	0	11
y^{14}	1,8	0	0	11,5
y^{15}	1,6	0	0	12
y^{16}	1,4	0	0	12,5
y^{17}	1,2	0	0	13
y^{18}	1	0	0	13,5
y^{19}	0,8	0	0	14
y^{20}	0,6	0	0	14,5
y^{21}	0,4	0	0	15
y^{22}	0,2	0	0	15,5
y^{23}	0	0	0	16
y^{24}	0	0	0	16

Tabuľka teda zobrazuje celkom 24 iterácií vyrátaných s presnosťou 10^{-1}, pričom iterácia y^{24} je vyrátaná presne. Keďže $y^{23} = y^{24}$, iterácia y^{24} určuje optimálne riešenie duálnej úlohy, platí teda:

$$y^{24} = y^* = (0,0,0,16).$$

Riešenie pôvodnej úlohy získame dosadením do vzťahu:

$$x^* = -C^{-1}(A^T y^* + d) = (4,0)^T.$$

Ostáva ešte určiť minimálnu hodnotu funkcie $f(x)$. Jednoduchým dosadením zistíme, že hodnota úlohy je $f^* = -32$.

Skutočnosť, že zmena poradia výpočtu komponentov jednotlivých iterácií y^k nemá vplyv na riešenie zobrazuje nasledujúca tabuľka, v ktorej sú zobrazené iterácie s tým istým východiskovým riešením y^0, ale rozdielnym poradím výpočtu zložiek:

Por.	4	1	2	3
y^0	0	0	0	0
y^1	0	5,2	0,7	4,8
y^2	0	2,9	1,2	8,4
y^3	0	1	1,6	11,6
y^4	0	0	1	14,5
y^5	0	0	0	16
y^6	0	0	0	16

4.2.Metóda Theila a Van de Panneho

Táto ďalšia z metód kvadratického programovania sa používa na riešenie úloh zadaných v tvare:

$$\frac{1}{2}\langle Cx, x\rangle + \langle d, x\rangle \to \min, \quad Ax \le b, \tag{4.4}$$

kde C je pozitívne definitná matica s rozmermi $n \times n$, A je matica s rozmermi $m \times n$ a b je vektor z $\mathbb{R}^m$. Dôležitým pojmom pri popise Metódy Theila a Van de Panneho je pojem singulárneho bodu úlohy (4.4). Preto uvedieme jeho definíciu a jeho vlastnosti vo vzťahu k riešeniu úlohy.

DEFINÍCIA 4.2. Povieme, že prípustný bod $\bar{x} \in \mathbb{R}^n$ je *singulárnym bodom úlohy* (4.3) , ak existuje $J \subseteq \{1,\ldots,m\}$ (nevylučuje sa prípad keď $J = \emptyset$) taká, že $\bar{x}$ je riešením úlohy

$$f(x) \to \min, \quad \langle a_i, x \rangle = b_i, \; i \in J. \tag{4.5}$$

Pričom $a_i, i = 1,\ldots,m$ predstavujú riadky matice A a $b_i, i = 1,\ldots,m$ sú zložky vektora b .

VETA 4.3.[11] Ak je x^* riešením úlohy (4.3), tak je x^* jej singulárnym bodom.

Keďže matica C je pozitívne definitná, úloha (4.5) má jediné riešenie. Počet podmnožín množiny $\{1,\ldots,m\}$ je 2^m, existuje teda najviac 2^m singulárnych bodov. Úlohu teda môžeme riešiť tak , že najprv určíme voľné minimu. Ak neleží v prípustnej množine, tak pre riešenie x^* musí byť jedna alebo viac nerovníc slnených ako rovnosti. Funkciu teda môžeme minimalizovať pri splnení požiadavky, že len určité podmienky sú splnené, a to v tvare rovníc. Ak tento postup aplikujeme na všetky možné podmnožiny obmedzujúcich podmienok, dostaneme určitý počet bodov, z ktorých potom vylúčime tie, ktoré neležia v prípustnej množine. U ostatných porovnáme funkčné hodnoty a zistíme, ktorý z nich predstavuje riešenie úlohy.

Na tomto princípe je založená aj metóda Theila a van de Panneho, s tým rozdielom, že výber podmnožín nie je „náhodný" , ale prebieha spôsobom, pri ktorom nie sú „zbytočné" množiny skúmané. Ďalší rozdiel spočíva v tom, že sme schopní rozpoznať či sme už našli riešenie.
Uvažujeme úlohu v tvare (4.4), kde C je pozitívne definitná matica. Označme x^0 minimum funkcie f na $\mathbb{R}^n$, platí teda $x^0 = -C^{-1}d$, ďalej označme x^J riešenie úlohy (4.5). Poznamenajme ešte, že obe riešenia existujú. Ďalší dôležitým predpokladom je tzv. *predpoklad nedegenerovanosti,* t.j. pre každú množinu $J = \{1,\ldots,m\}, J \neq \emptyset$, platí $x^J \neq x^{J\setminus\{i\}}$ pre každé $i \in J$. Teraz uvedieme niekoľko tvrdení, na ktorých je metóda založená:

VETA 4.4. [12] Nech x^* je riešenie úlohy (4.4) a nech

$$J^* = \left\{ i, i \in \{1,\ldots,m\}, \langle a_i, x^* \rangle = b_i \right\}.$$

Potom platí , že $x^* = x^{J^*}$.

VETA 4.5.[13] Nech pre nejakú množinu $J^* \subseteq \{1,\ldots,m\}, J^* \neq \emptyset$, platí $x^* = x^{J^*}$, kde x^* je riešenie úlohy (4.4). Potom pre každú vlastnú podmnožinu $J \subset J^*$ existuje aspoň jeden index $i \in J^* \setminus J$ taký, že $\langle a_i, x^J \rangle > b_i$.

[11] Dôkaz: viď.[1] str.138
[12] Dôkaz: viď.[1] str.139
[13] Dôkaz: viď.[1] str.140

DÔSLEDOK 4.6. Ak má množina J^* k prvkov $i_1,\dots,i_k$, potom existuje usporiadanie prvkov $i_{v_1},\dots,i_{v_l}$ také, že $\left\langle a_{i_{v_{l+1}}}, x^{J_l}\right\rangle > b_{i_{v_{l+1}}}$ pre $l=0,1,\dots,k-1$, pričom $J_l=\left\{i_{v_1},\dots,i_{v_l}\right\}, J_0=\varnothing$.

VETA 4.7.[14] Nech pre nejakú množinu $J\subseteq\{1,\dots,m\}$ platí $x^J\in X$ a nech pre každé $j\in J$ platí $\left\langle a_j, x^{J\setminus\{j\}}\right\rangle > b_j$. Potom $x^{J^*}=x^*$ je riešením úlohy (4.4)

Z Vety 4.4 je vidieť, že stačí nájsť množinu J^*, pretože potom nájdením x^{J^*} určíme súčasne x^*. Využitím Viet 4.5 až 4.7 odvodíme pravidlá, pomocou ktorých možno nájsť túto množinu:

(1) Ak porušuje x^0 určité obmedzujúce podmienky, tak potom x^* musí aspoň jednu z týchto podmienok splňovať ako rovnosť (plynie z V.4.5 pri $J=\varnothing$).

(2) Ak porušuje x^J určité obmedzujúce podmienky, tak potom musíme k J pripojiť aspoň jednu z týchto podmienok, aby sme dostali J^* (plynie z V.4.5).

(3) Určité x^J, ktoré neporušuje žiadne obmedzujúce podmienky je riešením x^* práve vtedy, keď $x^{J\setminus\{i\}}$ porušuje i-tu (vynechanú) obmedzujúcu podmienku (plynie z V.4.5 pri $J=J^*\setminus\{i\}$ a z V.4.7).

4.2.1.Postup výpočtu

0. Vypočítame x^0 zo vzťahu $x^0=-C^{-1}d$. Ak $x^0\in X=\left\{x\in\mathbb{R}^n:\left\langle a_i,x\right\rangle\le b_i,\right.$ $\left.i=1,\dots,m\right\}$ tak potom $x^0=x^*$. Ak $x^0\notin X$, nasleduje prvý krok.
1. Označme $G_1=\left\{i,\left\langle a_i,x^0\right\rangle>b_i\right\}$ a vyrátame $x^{\{i\}}$ pre $i\in G_1$. V prípade, že pre nejaký index i platí $x^{\{i\}}\in X$ tak potom $x^{\{i\}}=x^*$. Ak $x^{\{i\}}\notin X$ pre každé $i\in G_1$ pokračujeme druhým krokom.
2. Označme $G_2=\left\{\{j,i\}, j\in G_1,\left\langle a_i,x^j\right\rangle>b_i\right\}$. Pre každú dvojprvkovú množinu $\{j,i\}$, pre ktorú platí $x^{\{j,i\}}\in X$ vyšetríme podľa pravidla (3) či $x^{\{j,i\}}=x^*$ je riešením a v prípade, že tomu tak nie je vylúčime príslušné $\{j,i\}$ z ďalšieho výpočtu. Podobne skúmame tri a viac prvkové množiny.
3. Postupne sme teda preskúmali množiny $G_1,\dots,G_{k-1}$ zložené z jednoprvkových až $(k-1)$-prvkových podmnožín množiny $\{1,\dots,m\}$. Označme ďalej $G_k=\left\{J_k:=\{J_{k-1},i\}:J_{k-1}\in G_{k-1},\left\langle a_i,x^{J_{k-1}}\right\rangle>b_i\right\}$. Ak neexistuje žiadne i, pre ktoré $\left\langle a_i,x^{J_k}\right\rangle>b_i$ potom $x^{J_k}\in X$ a podľa pravidla (3) skúmame

[14] Dôkaz: viď.[1] str.141

či $x^{J_k} = x^*$. V prípade, že nenájdeme riešenie x^* ako riešenie niektorej z úloh $f(x) \to \min$, $\langle a_i, x \rangle = b_i, i \in J_k, J_k \in G_k$ zostrojíme množinu G_{k+1} tvorenú $(k+1)$-ticami indexov a pokračujeme $(k+1)$-ým krokom.

V prípade, že má množina J^* k-prvkov, nájdeme riešenie s istotou v k-tom kroku, kedy skúmame práve tie množiny J_k, ktoré vyhovujú nutným podmienkam pre J^* daným Dôsledkom 4.6. Postup preto končí najviac v m-tom kroku.

4.2.2. Ilustračný príklad[15]

Riešte metódou Theila a Van de Panneho úlohu:

$$f(x) = x_1^2 + x_2^2 + x_3^2 - x_1x_2 + x_1x_3 + 2x_1 - 2x_2 - x_3 \to \min$$

za podmienok

$$2x_1 - x_2 + x_3 \leq -3,$$
$$x_1 + 2x_2 + 3x_3 \leq 2,$$
$$x_1 + x_2 + x_3 \leq -1.$$

Riešenie: Riešime podľa vyššie uvedeného postupu. Najprv vyrátame $x^0 = -C^{-1}d$, kde

$$C = \begin{pmatrix} 2 & -1 & 1 \\ -1 & 2 & 0 \\ 1 & 0 & 2 \end{pmatrix}, d = \begin{pmatrix} 2 \\ -2 \\ -1 \end{pmatrix}.$$

Teda $x^0 = \left(-\frac{3}{2}, \frac{1}{4}, \frac{5}{4} \right)^T$. Keďže x^0 nesplňuje žiadnu z obmedzujúcich podmienok, zostrojíme množinu $G_1 = \{\{1\}, \{2\}, \{3\}\}$ a hľadáme riešenie medzi $x^{\{1\}}$, $x^{\{2\}}$, $x^{\{3\}}$.

$x^{\{1\}}$:

Riešime úlohu $f(x) \to \min, 2x_1 - x_2 + x_3 = 3$. Na to využijeme metódu Lagrangeových multiplikátorov. Príslušná Lagrangeová funkcia má podobu:
$L(x, \lambda) = x_1^2 + x_2^2 + x_3^2 - x_1x_2 + x_1x_3 + 2x_1 - 2x_2 - x_3 + \lambda(2x_1 - x_2 + x_3 + 3).$
Derivovaním podľa jednotlivých premenných získame sústavu:

$$2x_1 - x_2 + x_3 + 2 + 2\lambda = 0,$$
$$2x_2 - x_1 - 2 - \lambda = 0,$$
$$2x_3 + x_1 - 1 + \lambda = 0,$$
$$2x_1 - x_2 + x_3 + 3 = 0.$$

[15] [1] str.142

Riešením tejto sústavy je $x^{\{1\}} = \left(-2, \frac{1}{4}, \frac{5}{4}\right)^T$. Po dosadení do podmienok pôvodnej úlohy zistíme, že $x^{\{1\}}$ nevyhovuje druhej a tretej podmienke.

$x^{\{2\}}$:

Riešime úlohu $f(x) \to \min, x_1 + 2x_2 + 3x_3 = 2.$ Postupujeme podobne ako v predošlom prípade.

$L(x, y) = x_1^2 + x_2^2 + x_3^2 - x_1x_2 + x_1x_3 + 2x_1 - 2x_2 - x_3 + \lambda(x_1 + 2x_2 + 3x_3 - 2)$.

Derivovaním dostávame:

$$
\begin{aligned}
2x_1 - x_2 + x_3 + 2 + \lambda &= 0,\\
2x_2 - x_1 - 2 + 2\lambda &= 0,\\
2x_3 + x_1 - 1 + 3\lambda &= 0,\\
x_1 + 2x_2 + 3x_3 - 2 &= 0.
\end{aligned}
$$

Riešenie $x^{\{2\}} = \left(-\frac{14}{9}, \frac{1}{9}, \frac{10}{9}\right)^T$ nevyhovuje prvej a tretej podmienke pôvodnej úlohy.

$x^{\{3\}}$:

Riešime úlohu $f(x) \to \min, x_1 + x_2 + x_3 = -1.$ Vyjadríme Lagrangeovu funkciu

$L(x, y) = x_1^2 + x_2^2 + x_3^2 - x_1x_2 + x_1x_3 + 2x_1 - 2x_2 - x_3 + \lambda(x_1 + x_2 + x_3 + 1),$

derivovaním ktorej získame sústavu:

$$
\begin{aligned}
2x_1 - x_2 + x_3 + 2 + \lambda &= 0,\\
2x_2 - x_1 - 2 + \lambda &= 0,\\
2x_3 + x_1 - 1 + \lambda &= 0,\\
x_1 + x_2 + x_3 + 1 &= 0.
\end{aligned}
$$

Riešenie tejto sústavy je $x^{\{3\}} = \left(-2, -\frac{1}{4}, \frac{5}{4}\right)^T$. Toto riešenie nevyhovuje prvej podmienke, druhá podmienka je splnená.

Preskúmali sme teda jednoprvkové podmnožiny. Riešenie sme nenašli a tak v súlade s postupom riešenia pristúpime k vyšetrovaniu dvojprvkových podmnožín. Z výpočtu $x^{\{1\}}$ dostávame množiny $\{1,2\}$ a $\{1,3\}$, z výpočtu $x^{\{2\}}$ množiny $\{2,1\}$ a $\{2,3\}$, a z výpočtu $x^{\{3\}}$ dostávame množiny $\{3,1\}$. Celkovo teda dostávame množinu $G_3 = \{\{1,2\},\{1,3\},\{2,3\}\}$, ktorú budeme vyšetrovať v druhom kroku výpočtu.

$x^{\{1,2\}}$:

Riešime úlohu $f(x) \to \min, 2x_1 - x_2 + x_3 = -3, x_1 + 2x_2 + 3x_3 = 2.$ Lagrangeova funkcia má tvar

$$L(x,\lambda) = x_1^2 + x_2^2 + x_3^2 - x_1x_2 + x_1x_3 + 2x_1 - 2x_2 - x_3 +$$
$$+\lambda_1(2x_1 - x_2 + x_3 + 3) + \lambda_2(x_1 + 2x_2 + 3x_3 - 2).$$

Derivovaním získame systém

$$\begin{aligned} 2x_1 - x_2 + x_3 + 2 + 2\lambda_1 + \lambda_2 &= 0, \\ 2x_2 - x_1 - 2 - \lambda_1 + 2\lambda_2 &= 0, \\ 2x_3 + x_1 - 1 + \lambda_1 + 3\lambda_2 &= 0, \\ 2x_1 - x_2 + x_3 + 3 &= 0, \\ x_1 + 2x_2 + 3x_3 - 2 &= 0, \end{aligned}$$

riešením ktorého je $x^{\{1,2\}} = \left(-2, \frac{1}{5}, \frac{6}{5}\right)^T$. Toto riešenie nevyhovuje prvej podmienke.

$x^{\{1,3\}}$:

Riešime úlohu $f(x) \to \min, 2x_1 - x_2 + x_3 = -3, x_1 + x_2 + x_3 = -1$.

$$L(x,\lambda) = x_1^2 + x_2^2 + x_3^2 - x_1x_2 + x_1x_3 + 2x_1 - 2x_2 - x_3 +$$
$$+\lambda_1(2x_1 - x_2 + x_3 + 3) + \lambda_2(x_1 + x_2 + x_3 + 1).$$

Riešením sústavy

$$\begin{aligned} 2x_1 - x_2 + x_3 + 2 + 2\lambda_1 + \lambda_2 &= 0, \\ 2x_2 - x_1 - 2 - \lambda_1 + \lambda_2 &= 0, \\ 2x_3 + x_1 - 1 + \lambda_1 + \lambda_2 &= 0, \\ 2x_1 - x_2 + x_3 + 3 &= 0, \\ x_1 + x_2 + x_3 + 1 &= 0, \end{aligned}$$

je $x^{\{1,3\}} = \left(-\frac{13}{6}, -\frac{1}{12}, \frac{5}{4}\right)^T$. Riešenie $x^{\{1,3\}}$ vyhovuje všetkým podmienkam a k tomu ešte platí, že $x^{\{1\}}$ porušuje tretiu podmienku a $x^{\{3\}}$ porušuje prvú podmienku našli sme riešenie pôvodnej úlohy:

$$x^{\{1,3\}} = x^* = \left(-\frac{13}{6}, -\frac{1}{12}, \frac{5}{4}\right)^T.$$

Minimalizovaná funkcia v tomto bode nadobúda hodnotu $f(x^*) = -\frac{49}{24}$.

4.3. Wolfeho metóda

Táto metóda riešenia úloh kvadratického programovania je založená na Kuhn-Tuckerových podmienkach a na ich riešenie Wolfe rozpracoval upravený simplexový algoritmus. Pôvodne bola táto metóda rozpracovaná v závislosti od vlastností účelovej funkcie v dvoch variantoch a to:

(i) Špeciálny prípad, tzv. *krátky tvar* Wolfeho metódy: týmto variantom možno riešiť úlohy, v ktorých matica C kvadratickej formy účelovej funkcie je pozitívne definitná, alebo vektor d lineárnej formy je nulový. Prakticky krátky tvar Wolfeho metódy konverguje v mnohých prípadoch keď matica C je len pozitívne semidefinitná a vektor d nenulový, teoreticky je však možné konvergenciu zaučiť len pri splnení vyššie uvedených predpokladov.

(ii) Všeobecný prípad, tzv. *dlhý tvar* Wolfeho metódy: tento tvar je možné použiť bez vyššie uvedených obmedzení.

Uvažujme teda úlohu kvadratického programovania zadanú v tvare:

$$\frac{1}{2}\langle Cx, x\rangle + \langle dx, x\rangle \to \min, \quad Ax = b, \quad x \geq 0, \tag{4.5}$$

kde C je pozitívne semidefinitná matica. Ďalej predpokladáme, že $b \geq 0$ a hodnosť matice A je $h(A) = m$. Pre úlohu v tvare (4.5) sme Kuhn-Tuckerove podmienky odvodili v Kapitole 3.1. Bod x^* splňujúci $Ax^* = b, x^* \geq 0$ je teda riešením úlohy (4.5) práve vtedy keď existuje vektor $y^* \in \mathbb{R}^m$ taký, že sú splnené podmienky:

$$Cx^* + A^T y^* + d \geq 0,$$
$$\langle x^*, Cx^* + A^T y^* + d\rangle = 0.$$

Zavedením substitúcie $u = Cx + A^T y + d$ sa problém nájsť riešenie úlohy (4.5) zmení na problém vyriešiť nasledujúcu sústavu:

$$\begin{aligned} Cx + A^T y - u &= -d, \\ Ax &= b, \\ x, u &\geq 0 \end{aligned} \tag{4.6}$$

pričom zároveň musí byť splnená podmienka (tzv. podmienka ortogonality)

$$\langle x, u\rangle = 0. \tag{4.7}$$

Aby táto podmienka mohla byť splnená, musí byť x_i alebo u_i rovné nule pre každé $i = 1, \ldots, n$. Do úvahy teda prichádzajú len také riešenia sústavy (4.6) , v ktorých je z $2n + m$ premenných nenulových najviac $n + m$, teda toľko, koľko má sústava rovníc. Táto podmienka je splnená práve bázickými riešeniami. Princíp metódy teda spočíva v tom, že spomedzi prípustných bázických riešení sústavy (4.6) vyberieme tie, ktoré zároveň splňujú podmienku (4.7). Na to použijeme obmenený simplexový algoritmus. Zavedením pomocných (resp. umelých) premenných sa sústava (4.6) rozširi na tvar, v ktorom možno bezprostredne určiť počiatočné bázické riešenie vyhovujúce podmienke (4.7). Potom simplexovým algoritmom sa pomocné premenné anulujú, pričom simplexový algoritmus sa upraví tak, aby v každej iterácii bola splnená aj podmienka (4.7).

4.3.1.Krátky tvar Wolfeho metódy

Uvažujme úlohu zadanú v tvare (4.5) a hľadajme riešenie sústavy (4.6) pri dodržaní podmienky (4.7). Zavedieme pomocné premenné:

$$w \in \mathbb{R}^m, z^1, z^2 \in \mathbb{R}^n, w, z^1, z^2 \geq 0 .$$

Systém (4.6) tak prejde v rozšírený systém:

$$Cx + A^T y - u + z^1 + z^2 = -d,$$
$$Ax + w = b, \tag{4.8}$$
$$x, u, z^1, z^2, w \geq 0,$$
Pričom stále predpokladáme splnenie podmienky (4.7).
Počiatočné prípustné bázické riešenie určíme tak, aby nenulových ostalo najviac $n+m$ premenných. Teda $x=0$, $y=0$, $w=b$ a pre z^1, z^2 musí platiť

$$z_i^1 = \begin{cases} -d_i, & d_i \leq 0, \\ 0, & d_i \geq 0, \end{cases} \qquad z_i^2 = \begin{cases} d_i, & d_i \geq 0, \\ 0, & d_i \leq 0, \end{cases}$$

Báza potom obsahuje premenné $w_j = b_j$ pre $j = 1, \ldots, m$ a jedna z premenných z_i^1, z_i^2. V prípade $d_i = 0$ je jedno či bude v báze z_i^1, alebo z_i^2.
Ďalej nasledujú dva kroky, počas ktorých eliminujeme pomocné premenné na nulu a navrátime sa tak k pôvodnému systému (4.6).

PRVÝ KROK: Vychádzame z počiatočného bázického riešenia a riešime úlohu :

$$\sum_{j=1}^{m} w_j \to \min \tag{4.9}$$

za podmienok (4.8) a pri $y = 0, u = 0$. Celkovo tak obmedzenia pre úlohu (4.9) môžeme vyjadriť nasledovne:

$$Cx + z^1 - z^2 = -d,$$
$$Ax + w = b,$$
$$x, w, z^1, z^2 \geq 0.$$

Zároveň je splnená podmienka (4.7) pretože y a u nevstupujú v celom prvom kroku do bázy. Úlohu (4.9) riešime simplexovou metódou lineárneho programovania.[16] Výpočet prevádzame dovtedy kým sa v báze nenachádza žiadna zložka w. Ak počas výpočtu nedôjde k degenerácií, tak na konci prvého kroku sú v báze premenné x a jedna z premenných z_i^1, z_i^2 pre $i = 1, \ldots, n$. Ak zapíšeme predchádzajúcu sústavu maticovým spôsobom, teda v tvare:

$$\begin{pmatrix} C & I & -I & 0 \\ A & 0 & 0 & I \end{pmatrix} \begin{pmatrix} x \\ z^1 \\ z^2 \\ w \end{pmatrix} = \begin{pmatrix} -d \\ b \end{pmatrix},$$

vidíme, že premenné z^1 a z^2 nemôžu byť súčasne v báze. Ak by sa tak stalo, boli im prislúchajúce stĺpce lineárne závislé čo je v rozpore s definíciou bázického riešenia.

[16] viď. napr. J.PLESNÍK, J.DUPAĆOVÁ, M.VLACH, Lineárne programovanie, Alfa, Bratislava, 1990

Označme $\hat{x}, \hat{z}^1, \hat{z}^2$ a $\hat{w}$ zložky riešenia úlohy (4.9). V prípade $\hat{w} \neq 0$ úloha (4.5) nemá riešenie, pretože má prázdnu prípustnú množinu. V prípade, že $\hat{w} = 0$, položíme pre $i = 1, \ldots, n$

$$z_i = \begin{cases} \hat{z}_i^1, & \hat{z}_i^1 > 0, \\ \hat{z}_i^2, & \text{ak je } \hat{z}_i^2 > 0, \\ 0, & \hat{z}_i^1 = \hat{z}_i^2 = 0. \end{cases}$$

Vektor z je teda zložený z tých komponentov vektorov z^1, z^2, ktoré sa na konci prvého kroku vyskytovali v báze.

DRUHÝ KROK: V tomto kroku riešime úlohu

$$\sum_{i=1}^{n} z_i \to \min, \tag{4.10}$$

pričom za počiatočné bázické riešenie považujeme $y = 0, u = 0, x = \hat{x}$ a vektor z tak, ako je definovaný vyššie. Obmedzujúce podmienky sú v tvare:

$$\begin{aligned} Cx + A^T y - u + Dz &= -d, \\ Ax &= b, \\ x, u, z &\geq 0, \end{aligned}$$

pričom súčasne musí byť splnená podmienka (4.7). Matica D je diagonálna a na diagonále má prvky:

$$d_{ii} = \begin{cases} -1, & \text{ak je } z_i = \hat{z}_i^1, \\ 1, & \phantom{\text{ak je }} z_i = \hat{z}_i^2, \end{cases} \quad i = 1, \ldots, n.$$

Túto úlohu lineárneho programovania opäť riešime simplexovou metódou. Pretože však nie každé optimálne riešenie musí vyhovovať podmienke (4.7), doplníme algoritmus o nasledujúce pravidlo (tzv. Wolfeho pravidlo):

„Ak x_i je bázickou premennou, tak pri prechode k ďalšiemu bázickému riešeniu nesmie do bázy vstúpiť u_i, a naopak, ak je v báze u_i, tak do bázy nesmie vstúpiť x_i."

Vzniká otázka, či Wolfeho pravidlom upravený simplexový algoritmus konverguje, t.j. či sa nemôže stať, že $\Sigma z_i > 0$ a nebude možné pokračovať v simplexových iteráciách. Odpoveď na túto otázku dáva nasledujúca veta:

VETA 4.8.[17] Ak matica C pozitívne definitná , alebo ak $d = 0$, tak potom úloha (4.10) je riešiteľná simplexovým algoritmom pri rešpektovaní vyššie uvedeného Wolfeho pravidla, t.j. po konečnom počte iterácií dostaneme $\Sigma z_i = 0$.

[17] Dôkaz: viď.[2] str.144

Predpokladajme teda, že sú splnené predpoklady Vety 4.8. a označme $\bar{x}, \bar{y}, \bar{u}, \bar{z}$ riešenie úlohy (4.10). Ak je $\bar{z} = 0$ tak $\bar{x}, \bar{y}, \bar{u}$ sú riešením pôvodného systému (4.6) a $x^* = \bar{x}$ je riešenie úlohy (4.5).

4.3.2.Dlhý tvar Wolfeho metódy

Dlhý tvar Wolfeho metódy sa používa v prípadoch kedy nie sú splnené predpoklady Vety 4.8 . Pozostáva z troch fáz, pričom prvé dve sa v podstate zhodujú s krátkym tvarom metódy. Hlavný rozdiel spočíva v zavedení novej premennej $\mu \geq 0$, pomocou ktorej parametrizujeme lineárnu formu účelovej funkcie. Prvá rovnica systému (4.8) sa po zavedení tejto premennej zmení na rovnicu v tvare:

$$Cx + A^T y - u + z^1 + z^2 + \mu d = 0.$$

Úlohu začneme riešiť krátkym tvarom Wolfeho metódy pričom v prvých dvoch krokoch premenná μ nevstupuje do bázy a platí $\mu = 0$. Sú teda splnené predpoklady Vety 4.8, pretože vektor d je nahradený nulovým vektorom. Predpokladajme, že sme uskutočnili prvé dva kroky krátkeho tvaru a máme teda riešenie $\bar{x}, \bar{y}, \bar{u}, (\bar{z} = 0)$ systému:

$$\begin{aligned} Cx + A^T y - u + \mu d &= 0, \\ Ax &= b, \\ x, u, \mu &\geq 0. \end{aligned} \tag{4.11}$$

Je zároveň splnená podmienka (4.7) a $\mu = 0$. Pretože optimálne riešenie pôvodnej úlohy môžeme interpretovať ako optimálne riešenie parametrickej úlohy s hodnotou parametra $\mu = 1$, v ďalšom (treťom) kroku budeme zväčšovať hodnotu parametra μ.

TRETÍ KROK: V tomto kroku budeme vychádzať z bázického riešenia, ktoré sme získali v druhom kroku, teda $x = \bar{x},\ y = \bar{y},\ u = \bar{u},\ \mu = 0$. Opäť využijeme simplexovú metódu na riešenie úloh lineárneho programovania a riešime úlohu:

$$-\mu \to \min \tag{4.12}$$

za obmedzujúcich podmienok (4.11) a pri rešpektovaní Wolfeho pravidla. Pri riešení tejto úlohy môžu nastať dva prípady. Prvý prípad predstavuje situáciu kedy hodnoty $-\mu$ nie sú zdola ohraničené, t.j. $\mu \to \infty$. V druhom prípade proces riešenia úlohy (4.12) končí konečnou hodnotou. Túto situáciu popisuje nasledujúce tvrdenie:

VETA 4.9. Ak riešenie úlohy (4.12) má konečnú hodnotu, tak potom je táto hodnota $\mu^* = 0$ a hodnota účelovej funkcie f nie je na prípustnej množine zdola obmedzená, teda úloha (4.5) nemá riešenie.

Podrobnejšie sa teda budeme zaoberať prípadom kedy hodnota $-\mu$ nie je zdola obmedzená. Pri riešení úlohy (4.12) simplexovou metódou najprv dostaneme konečnú postupnosť bázických riešení sústavy (4.11)

$$(x^j, y^j, u^j, \mu^j), \quad j = 1, \ldots, k$$

a potom polpriamku

$$(x^k, y^k, u^k, \mu^k) + \lambda(x^{k+1}, y^{k+1}, u^{k+1}, \mu^{k+1}), \quad 0 \le \lambda < \infty,$$

na ktorom μ nadobúda ľubovoľne veľké hodnoty, t.j. $(-\mu) \to -\infty$. Táto polpriamka vznikne tým, že jedna z nebázických premenných môže prebehnúť kladné hodnoty, zatiaľ čo ostatné nebázické premenné zostávajú nulové, bez toho aby sme vylúčili niektorú z bázických premenných.

Wolfeho pravidlo zaručuje splnenie podmienky (4.7) nielen v jednotlivých bázických riešeniach, t.j.

$$\left\langle x^j, u^j \right\rangle = 0, \quad j = 1, \ldots, k+1,$$

ale takisto v susedných bázických riešeniach, t.j.

$$\left\langle x^j, u^{j+1} \right\rangle = 0, \quad \left\langle x^{j+1}, u^j \right\rangle = 0, \quad j = 1, \ldots, k.$$

Z uvedeného vyplýva, že podmienka (4.7) je splnená pre ľubovoľnú konvexnú kombináciu dvoch po sebe nasledujúcich bázických riešeniach a pre všetky body uvedenej polpriamky pretože

$$\left\langle \alpha x^j + \beta x^{j+1}, \alpha u^j + \beta u^{j+1} \right\rangle = \alpha^2 \left\langle x^j, u^j \right\rangle + \beta^2 \left\langle x^{j+1}, u^{j+1} \right\rangle +$$
$$+ \alpha\beta \left\langle x^j, u^{j+1} \right\rangle + \alpha\beta \left\langle x^{j+1}, u^j \right\rangle = 0,$$
$$\alpha, \beta \ge 0, \alpha + \beta = 1, \ j = 1, \ldots, k-1,$$

a tiež

$$\left\langle x^k + \lambda x^{k+1}, u^k + \lambda u^{k+1} \right\rangle = \left\langle x^k, u^k \right\rangle + \lambda^2 \left\langle x^{k+1}, u^{k+1} \right\rangle +$$
$$+ \lambda \left\langle x^k, u^{k+1} \right\rangle + \lambda \left\langle x^{k+1}, u^k \right\rangle = 0.$$

V prípade nedegenerovanosti sústavy (4.11) platí pre postupnosť bázických riešení $0 = \mu^1 < \mu^2 < \ldots < \mu^k$. V ďalšom musíme rozlišovať dve možnosti, ktoré môžu nastať:

(i) V prípade, že $\mu^k \ge 1$, vytvoríme konvexnú kombináciu dvoch po sebe idúcich bázických riešení, pre ktoré $\mu^j < 1 \le \mu^{j+1}$, pričom koeficient α v konvexnej kombinácii vezmeme taký, aby platilo: $\alpha\mu^j + (1-\alpha)\mu^{j+1} = 1$, t.j. $\alpha = \dfrac{\mu^{j+1} - 1}{\mu^{j+1} - \mu^j}$. Teda

$$(\tilde{x}, \tilde{y}, \tilde{u}, \tilde{\mu}) = \frac{\mu^{j+1} - 1}{\mu^{j+1} - \mu^j}(x^j, y^j, u^j, \mu^j) + \frac{1 - \mu^j}{\mu^{j+1} - \mu^j}(x^{j+1}, y^{j+1}, u^{j+1}, \mu^{j+1}).$$

Keďže $\tilde{\mu} = 1$ je $(\tilde{x}, \tilde{y}, \tilde{u}, \tilde{\mu})$ riešením (4.6) a je splnená aj podmienka (4.7). Znamená to teda, že $x^* = \tilde{x}$ je riešením úlohy (4.5).

(ii) Ak nastane situácia $\mu^k < 1$, tak riešením (4.11) a (4.7) je

$$(\tilde{x}, \tilde{y}, \tilde{u}, \tilde{\mu}) = (x^k, y^k, u^k, \mu^k) + \frac{1 - \mu^k}{\mu^{k+1}}(x^{k+1}, y^{k+1}, u^{k+1}, \mu^{k+1}).$$

Opäť teda platí, že $\tilde{\mu} = 1$ a teda $x^* = \tilde{x}$ je riešením úlohy (4.5).

4.3.3.Ilustračný príklad[18]

Riešte Wolfeho metódou úlohu

$$f(x)=\frac{1}{2}(x_1^2+2x_2^2+2x_3^2)+x_1x_2+x_1x_3+x_1+2x_2\to\min$$

pri splnení podmienok

$$\begin{aligned} x_1+x_2+x_3&=5,\\ 2x_1+x_2-x_3&=4,\\ x_1,x_2,x_3&\geq 0. \end{aligned}$$

Riešenie: Zo zadania úlohy vyjadríme:

$$C=\begin{pmatrix}1&1&1\\1&2&0\\1&0&2\end{pmatrix},\quad d=\begin{pmatrix}1\\2\\0\end{pmatrix},\quad A=\begin{pmatrix}1&1&1\\2&1&-1\end{pmatrix},\quad b=\begin{pmatrix}5\\4\end{pmatrix}.$$

Vzhľadom na to, že matica C je len semidefinitná a vektor d nie je nenulový, nie sú splnené predpoklady pre použitie krátkeho tvaru Wolfeho metódy. Musíme teda použiť dlhý tvar k čomu zavedieme nasledujúce pomocné premenné: $w,z^1,z^2,\mu\geq 0$. Počiatočná tabuľka prvého kroku má nasledujúcu podobu:

	x_1	x_2	x_3	z_1^1	z_2^1	z_3^1	z_1^2	z_2^2	z_3^2	w_1	w_2	μ	
w_1	1	1	1	0	0	0	0	0	0	1	0	0	5
w_2	$\boxed{2}$	1	−1	0	0	0	0	0	0	0	1	0	4
z_1^2	−1	−1	−1	−1	0	0	1	0	0	0	0	−1	0
z_2^2	−1	−2	0	0	−1	0	0	1	0	0	0	−2	0
z_3^2	−1	0	−2	0	0	−1	0	0	1	0	0	0	0
$\sum w_j$	3	2	0	0	0	0	0	0	0	0	0	0	9

V ľavom stĺpci vyčítame bázické premenné a v pravom ich hodnoty. V báze sú teda premenné $w=b$ a $z^2=\mu d$. Posledný riadok obsahuje funkciu, ktorú v prvom kroku budeme minimalizovať. Tabuľka zatiaľ neobsahuje premenné y a u, keďže tieto premenné v prvok kroku nevstupujú do bázy.
Využijeme teda simplexovú metódu na odstránenie vektora w z bázy. V každej tabuľke je v rámčeku označený kľúčový prvok, podľa ktorého sa tabuľka upravuje. Je to vždy kladný prvok nachádzajúci sa v stĺpci, v ktorom je kladná hodnota v poslednom riadku. Výber riadku, v ktorom sa nachádza kľúčový prvok je určený tým, že podiel hodnoty príslušnej bázickej premennej a kľúčového prvku je minimálny. Nasledujúce dve tabuľky zobrazujú postup v prvom kroku výpočtu:

[18] [1] str. 152

	x_1	x_2	x_3	z_1^1	z_2^1	z_3^1	z_1^2	z_2^2	z_3^2	w_1	w_2	μ	
x_3	0	$\frac{1}{2}$	$\boxed{\frac{3}{2}}$	0	0	0	0	0	0	1	$-\frac{1}{2}$	0	3
w_2	1	$\frac{1}{2}$	$-\frac{1}{2}$	0	0	0	0	0	0	0	$\frac{1}{2}$	0	2
z_1^2	0	$-\frac{1}{2}$	$-\frac{3}{2}$	−1	0	0	1	0	0	0	$\frac{1}{2}$	−1	2
z_2^2	0	$-\frac{3}{2}$	$-\frac{1}{2}$	0	−1	0	0	1	0	0	$\frac{1}{2}$	−2	2
z_3^2	0	$\frac{1}{2}$	$-\frac{5}{2}$	0	0	−1	0	0	1	0	$\frac{1}{2}$	0	2
$\sum w_j$	0	$\frac{1}{2}$	$\frac{3}{2}$	0	0	0	0	0	0	0	$-\frac{3}{2}$	0	3

	x_1	x_2	x_3	z_1^1	z_2^1	z_3^1	z_1^2	z_2^2	z_3^2	w_1	w_2	μ	
x_3	0	$\frac{1}{3}$	1	0	0	0	0	0	0	$\frac{2}{3}$	$-\frac{1}{3}$	0	2
x_1	1	$\frac{2}{3}$	0	0	0	0	0	0	0	$\frac{1}{3}$	$\frac{1}{3}$	0	3
z_1^2	0	0	0	−1	0	0	1	0	0	1	0	−1	5
z_2^2	0	$-\frac{4}{3}$	0	0	−1	0	0	1	0	$\frac{1}{3}$	$\frac{1}{3}$	−2	3
z_3^2	0	$\frac{4}{3}$	0	0	0	−1	0	0	1	$\frac{5}{3}$	$-\frac{1}{3}$	0	7
$\sum w_j$	0	0	0	0	0	0	0	0	0	−1	−1	0	0

Premenné w_1 a w_2 sú z bázy vylúčené a môžeme pristúpiť k druhému kroku. Do tabuľky pridáme stĺpce odpovedajúce vektorom y a u a odstránime stĺpce, ktoré patrili k premennej w a z^1. Vektor z^2 ďalej budeme označovať len z a v ďalšom postupe ho budeme odstraňovať z bázy.
Počiatočná tabuľka druhého kroku má teda nasledujúcu podobu :

	x_1	x_2	x_3	y_1	y_2	u_1	u_2	u_3	z_1	z_2	z_3	μ	
x_3	0	$\frac{1}{3}$	1	0	0	0	0	0	0	0	0	0	2
x_1	1	$\frac{2}{3}$	0	0	0	0	0	0	0	0	0	0	3
z_1	0	0	0	−1	−2	1	0	0	1	0	0	−1	5
z_2	0	$-\frac{4}{3}$	0	$\boxed{-1}$	−1	0	1	0	0	1	0	−2	3
z_3	0	$-\frac{4}{3}$	0	−1	1	0	0	1	0	0	1	0	7
$\sum z_i$	0	0	0	−3	−2	1	1	1	0	0	0	−3	15

Opäť postupujeme simplexovou metódou, pri súčasnom dodržaní Wolfeho pravidla. Keďže na vektor y nie je kladené žiadne obmedzenie, platí pre neho, že ho môžeme vložiť do bázy aj v prípade keď je hodnota v poslednom riadku záporná a pritom vyberáme záporný kľúčový prvok. Ak je už y v báze, už z nej nevystupuje. Postup v druhom kroku zobrazujú nasledujúce tri tabuľky:

	x_1	x_2	x_3	y_1	y_2	u_1	u_2	u_3	z_1	z_2	z_3	μ	
x_3	0	$\frac{1}{3}$	1	0	0	0	0	0	0	0	0	0	2
x_1	1	$\frac{2}{3}$	0	0	0	0	0	0	0	0	0	0	3
z_1	0	$\frac{4}{3}$	0	0	−1	1	−1	0	1	−1	0	1	2
y_1	0	$\frac{4}{3}$	0	1	1	0	−1	0	0	−1	0	2	−3
z_3	0	$\boxed{\frac{8}{3}}$	0	0	2	0	−1	1	0	−1	1	2	4
$\sum z_i$	0	4	0	0	1	1	−2	1	0	−3	0	3	6

	x_1	x_2	x_3	y_1	y_2	u_1	u_2	u_3	z_1	z_2	z_3	μ	
x_3	0	0	1	0	$-\frac{1}{4}$	0	$\frac{1}{8}$	$-\frac{1}{8}$	0	$\frac{1}{8}$	$-\frac{1}{8}$	$-\frac{1}{4}$	$\frac{3}{2}$
x_1	1	0	0	0	$-\frac{1}{2}$	0	$\frac{1}{4}$	$-\frac{1}{4}$	0	$\frac{1}{4}$	$-\frac{1}{4}$	$-\frac{1}{2}$	2
z_1	0	0	0	0	$\boxed{-2}$	1	$-\frac{1}{2}$	$-\frac{1}{2}$	1	$-\frac{1}{2}$	$-\frac{1}{2}$	0	0
y_1	0	0	0	1	0	0	$-\frac{1}{2}$	$-\frac{1}{2}$	0	$-\frac{1}{2}$	$-\frac{1}{2}$	1	−5
x_2	0	1	0	0	$\frac{3}{4}$	0	$-\frac{3}{8}$	$\frac{3}{8}$	0	$-\frac{3}{8}$	$\frac{3}{8}$	$\frac{3}{4}$	$\frac{3}{2}$
$\sum z_i$	0	0	0	0	−2	1	$-\frac{1}{2}$	$-\frac{1}{2}$	0	$-\frac{3}{2}$	$-\frac{3}{2}$	0	0

	x_1	x_2	x_3	y_1	y_2	u_1	u_2	u_3	z_1	z_2	z_3	μ	
x_3	0	0	1	0	0	$-\frac{1}{8}$	$\frac{3}{16}$	$-\frac{1}{16}$	$-\frac{1}{8}$	$\frac{3}{16}$	$-\frac{1}{16}$	$-\frac{1}{4}$	$\frac{3}{2}$
x_1	1	0	0	0	0	$-\frac{1}{4}$	$\frac{3}{8}$	$-\frac{1}{8}$	$-\frac{1}{4}$	$\frac{3}{8}$	$-\frac{1}{8}$	$-\frac{1}{2}$	2
y_2	0	0	0	0	1	$-\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{4}$	$-\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{4}$	0	0
y_1	0	0	0	1	0	0	$-\frac{1}{2}$	$-\frac{1}{2}$	0	$-\frac{1}{2}$	$-\frac{1}{2}$	1	−5
x_2	0	1	0	0	0	$\frac{3}{8}$	$-\frac{9}{16}$	$\frac{3}{16}$	$\frac{3}{8}$	$-\frac{9}{16}$	$\frac{3}{16}$	$\frac{3}{4}$	$\frac{3}{2}$
$\sum z_i$	0	0	0	0	0	0	0	0	−1	−1	−1	0	0

Hodnota minimalizačnej funkcie je nulová a všetky premenné vektoru z už vypadli z bázy. Druhý krok teda končí, nasleduje krok tretí. Odstránime stĺpce prislúchajúce z. V treťom kroku budeme minimalizovať funkciu $-\mu$. Počiatočná tabuľka je nasledujúca:

	x_1	x_2	x_3	y_1	y_2	u_1	u_2	u_3	μ	
x_3	0	0	1	0	0	$-\frac{1}{8}$	$\frac{3}{16}$	$-\frac{1}{16}$	$-\frac{1}{4}$	$\frac{3}{2}$
x_1	1	0	0	0	0	$-\frac{1}{4}$	$\frac{3}{8}$	$-\frac{1}{8}$	$-\frac{1}{2}$	2
y_2	0	0	0	0	1	$-\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{4}$	0	0
y_1	0	0	0	1	0	0	$-\frac{1}{2}$	$-\frac{1}{2}$	1	-5
x_2	0	1	0	0	0	$\frac{3}{8}$	$-\frac{9}{16}$	$\frac{3}{16}$	$\boxed{\frac{3}{4}}$	$\frac{3}{2}$
$-\mu$	0	0	0	0	0	0	0	0	1	0

Začíname s hodnotou $\mu^1 = 0$ a postupujeme analogicky ako v predchádzajúcom kroku.

	x_1	x_2	x_3	y_1	y_2	u_1	u_2	u_3	μ	
x_3	0	$\frac{1}{3}$	1	0	0	0	0	0	0	2
x_1	1	$\frac{2}{3}$	0	0	0	0	0	0	0	3
y_2	0	0	0	0	1	$-\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{4}$	0	0
y_1	0	$-\frac{4}{3}$	0	1	0	$-\frac{1}{2}$	$\frac{1}{4}$	$-\frac{3}{4}$	0	-7
μ	0	$\frac{4}{3}$	0	0	0	$\frac{1}{2}$	$-\frac{3}{4}$	$\frac{1}{4}$	1	2
$-\mu$	0	$-\frac{4}{3}$	0	0	0	$-\frac{1}{2}$	$\frac{3}{4}$	$-\frac{1}{4}$	0	-2

Keďže hodnota $\mu^2 = 2 \geq 1$ výpočet v tabuľke končí a ostáva dorátať riešenie x^* ako konvexnú kombináciu bázických riešení posledných dvoch tabuliek. Teda

$$x^* = \alpha x^1 + (1-\alpha)x^2,$$

kde

$$x_1^1 = 2, x_2^1 = \frac{3}{2}, x_3^1 = \frac{3}{2}, \mu^1 = 0, \qquad x_1^2 = 3, x_2^2 = 0, x_3^2 = 2, \mu^2 = 2,$$

$$\alpha = \frac{\mu^2 - 1}{\mu^2 - \mu^1} = \frac{1}{2}.$$

Zložky vektora x^* teda určíme výpočtom:

$$x_1^* = \frac{1}{2}\cdot 2 + \frac{1}{2}\cdot 3 = \frac{5}{2}, \quad x_2^* = \frac{1}{2}\cdot\frac{3}{2} + \frac{1}{2}\cdot 0 = \frac{3}{4}, \quad x_3^* = \frac{1}{2}\cdot\frac{3}{2} + \frac{1}{2}\cdot 2 = \frac{7}{4}.$$

Dostávame teda riešenie úlohy:

$$x^* = \left(\frac{5}{2}, \frac{3}{4}, \frac{7}{4}\right)^T.$$

Dosadením tohto riešenia do minimalizovanej funkcie zistíme, že hodnota úlohy je 17.

Literatúra

[1] O. Došlý, Základy konvexní analýzy a optimalizace v $\mathbb{R}^n$, Masarykova univerzita, Brno, 2005, 1.vydanie.

[2] M. Hamala, Nelineárne programovanie, Vydavateľstvo technickej a ekonomickej literatúry, Bratislava, 2.vydanie.

[3] V. Jarník, Diferenciální počet II, Academia, Praha 1974.