Úloha **NLP**

Uvažujme

$$\min\{f(\vec{x}) : g(\vec{x}) \leqq \vec{0}, h(\vec{x}) = \vec{0}, \vec{x} \in X\} \qquad \text{(NLP)}$$

kde

- $g(\vec{x}) = (g_1(\vec{x}), \ldots, g_m(\vec{x}))^\top$ je vektorová funkcia nerovností
- $h(\vec{x}) = (h_1(\vec{x}), \ldots, h_p(\vec{x}))^\top$ je vektorová funkcia rovností
- množina X predstavuje napr. nezápornosť niektorých premenných
- množinu prípustných riešení úlohy **NLP** označujeme

$$M = \{\vec{x} \in X : g(\vec{x}) \leqq \vec{0}, h(\vec{x}) = \vec{0}\}$$

- nelinearita úlohy je podmienená nelineárnosťou aspoň jednej z funkcií f, g_i a h_k pre $i \in \{1, \ldots, m\}$ a $k \in \{1, \ldots, p\}$

- špeciálnym prípadom úlohy **NLP** je úloha **KP** t.j. účelová funkcia je kvadratická a podmienky sú lineárne

$$f(\vec{x}\,) = \vec{x}^{\top} \cdot \boldsymbol{C} \cdot \vec{x} + \vec{p}^{\top} \cdot \vec{x} \longrightarrow \min \tag{KP}$$

pri ohraničeniach $\vec{x} \in D$, kde

- n je počet rozhodovacích premenných úlohy
- $\vec{x} = (x_1, \ldots, x_n)^T \in \mathbb{R}^n$ je vektor rozhodovacích premenných
- $\boldsymbol{C} \in \mathbb{R}^{n \times n}$ je symetrická pozitívne semidefinitná matica n-tého stupňa
- $\vec{p} \in \mathbb{R}^n$ je vektor koeficientov lineárnej časti funkcie f
- $D \subset \mathbb{R}^n$ je konvexná polyedrická množina v $\mathbb{R}^n$

Tri špeciálne prípady polyedrickej množiny D

$$D_1 = \{\vec{x} : \boldsymbol{A} \cdot \vec{x} \leqq \vec{b}, \vec{x} \geqq \vec{0}\}$$

$$D_2 = \{\vec{x} : \boldsymbol{A} \cdot \vec{x} = \vec{b}, \vec{x} \geqq \vec{0}\}$$

$$D_3 = \{\vec{x} : \boldsymbol{A} \cdot \vec{x} = \vec{b}\}$$

kde $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ a $\vec{b} \in \mathbb{R}^m$

- prípady množín D_2 a D_3 vieme vhodnými dodatočnými podmienkami previesť na problém pre prípad D_1

$$\boldsymbol{A} \cdot \vec{x} = \vec{b} \Longleftrightarrow (\boldsymbol{A} \cdot \vec{x} \leqq \vec{b} \wedge \boldsymbol{A} \cdot \vec{x} \geqq \vec{b})$$

$$\boldsymbol{A} \cdot \vec{x} = \vec{b} \Longleftrightarrow (\boldsymbol{A} \cdot \vec{x} \leqq \vec{b} \wedge -\boldsymbol{A} \cdot \vec{x} \leqq -\vec{b})$$

- stanovenie podmienok optimálnosti pre sústavu ohraničení v tvare D_1

Základný tvar úlohy **KP**

$$\begin{aligned} f(\vec{x}) = \vec{x}^\top \cdot \boldsymbol{C} \cdot \vec{x} + \vec{p}^\top \cdot \vec{x} &\longrightarrow \min \\ \boldsymbol{A} \cdot \vec{x} &\leqq \vec{b} \\ \vec{x} &\geqq \vec{0} \end{aligned}$$

kde $\boldsymbol{C} \in \mathbb{R}^{n\times n}$ je symetrická pozitívne semidefinitná matica, $\vec{p} \in \mathbb{R}^n$, $\boldsymbol{A} \in \mathbb{R}^{m\times n}$, $\vec{b} \in \mathbb{R}^m$ a $\vec{x} = (x_1, \dots, x_n)^T \in \mathbb{R}^n$

- podmienky optimálnosti Kuhna-Tuckera nám pomôžu pri hľadaní riešenia úlohy **KP**

- pri hľadaní viazaného extrému použijeme Lagrangeovu funkciu s multiplikátormi $\vec{u} \in \mathbb{R}^m$ a $\vec{w} \in \mathbb{R}^n$

$$L(\vec{x}, \vec{u}, \vec{w}) = \vec{x}^\top \cdot \boldsymbol{C} \cdot \vec{x} + \vec{p}^\top \cdot \vec{x} + \vec{u}^\top \cdot (\boldsymbol{A} \cdot \vec{x} - \vec{b}) + \vec{w}^\top \cdot (-\vec{x})$$

- nakoľko rozhodovacie premenné sú nezáporné, môžeme použiť zovšeobecnenú Lagrangeovu funkciu

$$L(\vec{x}, \vec{u}) = \vec{x}^\top \cdot \boldsymbol{C} \cdot \vec{x} + \vec{p}^\top \cdot \vec{x} + \vec{u}^\top \cdot (\boldsymbol{A} \cdot \vec{x} - \vec{b}) \qquad \text{(L)}$$

Postup na získanie Kuhnových-Tuckerových podmienok:

- na stanovenie Kuhnových-Tuckerových podmienok sú potrebné parciálne derivácie Lagrangeovej funkcie $L(\vec{x}, \vec{u})$

$$\frac{\partial L}{\partial u_i} = A_i \cdot \vec{x} - b_i \qquad \text{pre} \quad \forall i \in \{1, 2, \ldots, m\}$$

$$\frac{\partial L}{\partial x_j} = 2 \cdot C_j \cdot \vec{x} + p_j + A_j^\top \cdot \vec{u} \quad \text{pre} \quad \forall j \in \{1, 2, \ldots, n\}$$

kde A_i zodpovedá i-temu riadku matice $\boldsymbol{A}$; C_j zodpovedá j-temu riadku matice $\boldsymbol{C}$ a $A_j^\top$ zodpovedá j-temu riadku matice $\boldsymbol{A}^\top$

- pre vypočítané derivácie majú platiť podmienky komplementárnej rovnováhy

$$\frac{\partial L}{\partial x_j} \geqq 0; \quad x_j \frac{\partial L}{\partial x_j} = 0; \quad x_j \geqq 0 \quad \text{pre} \quad \forall j \in \{1, 2, \ldots, n\}$$

$$\frac{\partial L}{\partial u_i} \leqq 0; \quad u_i \frac{\partial L}{\partial u_i} = 0; \quad u_i \geqq 0 \quad \text{pre} \quad \forall i \in \{1, 2, \ldots, m\}$$

- z vypočítaných parciálnych derivácií dostávame

$$
\begin{aligned}
2 \cdot \boldsymbol{C} \cdot \vec{x} + \vec{p} + \boldsymbol{A}^\top \cdot \vec{u} &\geqq \vec{0} \\
\vec{x}^\top \cdot (2 \cdot \boldsymbol{C} \cdot \vec{x} + \vec{p} + \boldsymbol{A}^\top \cdot \vec{u}) &= \vec{0} \\
\boldsymbol{A} \cdot \vec{x} - \vec{b} &\leqq \vec{0} \\
\vec{u}^\top \cdot (\boldsymbol{A} \cdot \vec{x} - \vec{b}) &= \vec{0} \\
\vec{x} &\geqq \vec{0} \\
\vec{u} &\geqq \vec{0}
\end{aligned}
$$

- zavedením vhodných substitúcií, t.j. nových premenných $\vec{v} \in \mathbb{R}^n$ a $\vec{y} \in \mathbb{R}^m$ v tvare

$$\begin{aligned} \vec{v} &= 2 \cdot \boldsymbol{C} \cdot \vec{x} + \vec{p} + \boldsymbol{A}^\top \cdot \vec{u} \\ \vec{y} &= -\boldsymbol{A} \cdot \vec{x} + \vec{b} \end{aligned}$$

a následnom dosadení dostávame

$$\begin{aligned} 2 \cdot \boldsymbol{C} \cdot \vec{x} + \vec{p} + \boldsymbol{A}^\top \cdot \vec{u} &= \vec{v} \geqq \vec{0} \\ x_j \cdot v_j &= 0 \qquad \text{pre} \quad \forall j \in \{1, 2, \dots, n\} \\ \vec{x} &\geqq \vec{0} \\ \boldsymbol{A} \cdot \vec{x} - \vec{b} &= -\vec{y} \leqq \vec{0} \\ u_i \cdot y_i &= 0 \qquad \text{pre} \quad \forall i \in \{1, 2, \dots, m\} \\ \vec{u} &\geqq \vec{0} \end{aligned}$$

- a po malých úpravách sústavu rovníc, ktorú nazývame Kuhnove-Tuckerove podmienky

Kuhnove-Tuckerove podmienky pre úlohu KP v tvare D_1

$$
\begin{aligned}
-2\cdot \boldsymbol{C}\cdot \vec{x} - \boldsymbol{A}^\top \cdot \vec{u} + \vec{v} &= \vec{p} \\
\boldsymbol{A}\cdot \vec{x} + \vec{y} &= \vec{b} \\
x_j \cdot v_j &= 0 \quad \text{pre} \quad \forall j \in \{1, 2, \ldots, n\} \\
u_i \cdot y_i &= 0 \quad \text{pre} \quad \forall i \in \{1, 2, \ldots, m\} \\
\vec{x}, \vec{y}, \vec{u}, \vec{v} &\geqq \vec{0}
\end{aligned}
$$

- autormi sú Mike C. Shetty a Carlton E. Lemke (1962)
- algoritmus (Shetty-Lemke) pozostáva z troch krokov:

- K0: (Inicializačná fáza) - podmienky optimálnosti Kuhna-Tuckera upravíme o parameter μ s príslušnými koeficientami

$$
\begin{aligned}
-2 \cdot \boldsymbol{C} \cdot \vec{x} - \boldsymbol{A}^\top \cdot \vec{u} + \vec{v} + \vec{q} \cdot \mu &= \vec{p} \\
\boldsymbol{A} \cdot \vec{x} + \vec{y} + \vec{t} \cdot \mu &= \vec{b} \\
\vec{x}^\top \cdot \vec{v} &= 0 \\
\vec{y}^\top \cdot \vec{u} &= 0 \\
\vec{x}, \vec{y}, \vec{u}, \vec{v} &\geqq \vec{0}
\end{aligned}
$$

kde vektory $\vec{q} \in \mathbb{R}^n$ a $\vec{t} \in \mathbb{R}^m$ sú definované nasledovne:

$$q_j = \begin{cases} -1 & \text{ak } p_j < 0 \\ 0 & \text{ak } p_j \geqq 0 \end{cases} \qquad (\vec{q})$$

$$t_i = \begin{cases} -1 & \text{ak } b_i < 0 \\ 0 & \text{ak } b_i \geqq 0 \end{cases} \qquad (\vec{t})$$

- K1: (Fáza riešenia) - riešeniu danej sústavy rovníc zodpovedá simplexová tabuľka

B	$\vec{p} \mid \vec{b}$	$\vec{x}$	$\vec{u}$	$\vec{v}$	$\vec{y}$	μ
$\vec{v}$	$\vec{p}$	$-2 \cdot C$	$-A^\top$	E	0	$\vec{q}$
$\vec{y}$	$\vec{b}$	A	0	0	E	$\vec{t}$

- východiskové bázické riešenie je v tvare

$$\vec{x} = \vec{0}, \; \vec{u} = \vec{0}$$

$$\vec{v} = \vec{p}, \; \vec{y} = \vec{b}, \; \mu = 0$$

Následne overíme:

(1) ak $\vec{v} \geqq \vec{0}$ a $\vec{y} \geqq \vec{0}$, tak vektor $\vec{x} = \vec{0}$ je optimálnym riešením úlohy

(2) inak realizujeme elementárnu zmenu bázy, pri ktorej sa vedúcim stĺpcom stáva stĺpec reprezentovaný premennou μ a vedúci riadok je určený minimálnym prvkom pravej strany ohraničení; pokračujeme pokiaľ premenná μ nevystúpi z bázy

- nezápornosť vektorov $\vec{x}$, $\vec{y}$, $\vec{u}$ a $\vec{v}$ garantuje kritérium prípustnosti riešenia pri primárnom riešení simplexovej metódy

- K2: (Pravidlo výpočtu) - úlohu riešime pomocou simplexového algoritmu s dodatočnými pravidlami pre vstup premenných do bázy
  1. ak z bázy vystúpila v $(k-1)$-ej iterácii premenná x_j, tak v aktuálnej k-ej iterácii vstúpi do bázy premenná v_j
  2. ak z bázy vystúpila v $(k-1)$-ej iterácii premenná v_j, tak v aktuálnej k-ej iterácii vstúpi do bázy premenná x_j
  3. ak z bázy vystúpila v $(k-1)$-ej iterácii premenná y_i, tak v aktuálnej k-ej iterácii vstúpi do bázy premenná u_i
  4. ak z bázy vystúpila v $(k-1)$-ej iterácii premenná u_i, tak v aktuálnej k-ej iterácii vstúpi do bázy premenná y_i

- autorom je Philip Wolfe (1956) a metóda je rozpracovaná v závislosti od vlastností účelovej funkcie v dvoch variantoch:

(1) - špeciálny prípad, tzv. krátky tvar Wolfeho metódy

- tento variant umožňuje riešiť úlohy, v ktorých matica $\boldsymbol{C}$ je pozitívne definitná alebo vektor $\vec{p}$ je nulový
- v praxi platí, že krátky tvar Wolfeho metódy konverguje v mnohých prípadoch, aj keď matica $\boldsymbol{C}$ je len pozitívne semidefinitná a vektor $\vec{p}$ je nenulový; teoreticky je však možné konvergenciu zaručiť len pri splnení vyššie uvedených predpokladov

(2) - všeobecný prípad, tzv. dlhý tvar Wolfeho metódy

- tento tvar je možné použiť bez vyššie uvedených obmedzení

- algoritmus (Wolfe v krátkom tvare) pozostáva z piatich krokov:

- K0: (Inicializačná fáza) - nájdeme prípustnú bázu $\boldsymbol{B} \subset \{x_1, \ldots, x_n, y_1, \ldots, y_m\}$ pre nezáporné riešenie sústavy $\boldsymbol{A} \cdot \vec{x} + \vec{y} = \vec{b}$; ak takáto primárne prípustná báza neexistuje, tak choď na krok K3
- K1: (Fáza riešenia) - zostavíme pomocnú úlohu pre vektor parametrov $\vec{w} = (w_1, \ldots, w_n)^\top$

$$\begin{aligned} \sum_{i=1}^{n} w_i &\longrightarrow \min \\ \boldsymbol{A} \cdot \vec{x} + \vec{y} &= \vec{b} \\ 2 \cdot \boldsymbol{C} \cdot \vec{x} + \boldsymbol{A}^\top \cdot \vec{u} - \vec{v} + \boldsymbol{D} \cdot \vec{w} &= -\vec{p} \\ \vec{x}, \vec{y}, \vec{u}, \vec{v} &\geqq \vec{0} \end{aligned}$$

kde $\boldsymbol{D} = (d_{i,j})_{n,n}$ je diagonálna matica definovaná nasledovne:

$$d_{k,k} = \begin{cases} -1 & \text{ak } 2\sum_{j=1}^{n} c_{kj} \cdot x_j(\boldsymbol{B}) > -p_k \\ 1 & \text{ak } 2\sum_{j=1}^{n} c_{kj} \cdot x_j(\boldsymbol{B}) \leqq -p_k \end{cases} \qquad (\boldsymbol{D})$$

- položíme $L_0 = \boldsymbol{B} \cup \{w_1, \ldots, w_n\}$ a L_0 bude prípustnou bázou takto zostavenej úlohy; pokračujeme pokiaľ všetky premenné vektora $\vec{w}$ nevystúpia z bázy

- K2: (Pravidlo výpočtu) - úlohu riešime pomocou simplexového algoritmu s dodatočnými pravidlami pre vstup premenných do bázy
  (1) ak $z_{x_j} - c_{x_j} > 0$ a v_j nie je bázická premenná, tak x_j môže vstupovať do bázy
  (2) ak $z_{v_j} - c_{v_j} > 0$ a x_j nie je bázická premenná, tak v_j môže vstupovať do bázy
  (3) ak $z_{y_i} - c_{y_i} > 0$ a u_i nie je bázická premenná, tak y_i môže vstupovať do bázy
  (4) ak $z_{u_i} - c_{u_i} > 0$ a y_i nie je bázická premenná, tak u_i môže vstupovať do bázy

- za počiatočnú bázu simplexového algorimu vezmeme L_0 a takto upravený algoritmus skončí po konečnom počte krokov; označme konečnú nájdenú bázu L_{kon}

(1) ak $\sum_{i=1}^{m} u_j(L_{\text{kon}}) > 0$, tak prejdeme na krok K4

(2) ak $\sum_{i=1}^{m} u_j(L_{\text{kon}}) = 0$, tak prejdeme na krok K5

- K3: (Ukončenie algoritmu) - úloha nemá prípustné riešenie
- K4: (Ukončenie algoritmu) - komponenty bázického riešenia $x(L_{\text{kon}})$ sú optimálne riešenia úlohy
- K5: (Ukončenie algoritmu) - algoritmus nenašiel optimálne riešenie úlohy