# Konvexná optimalizácia

# Metódy vnútorného bodu

Uvažujme úlohu konvexného programovania

$$\begin{aligned} Min \quad & f_0(x) \\ & f_i(x) \le 0, \quad i = 1, \dots, m \qquad\qquad (P) \\ & Ax = b, \end{aligned}$$

**Predpoklady:**

- Funkcie $f_0, f_1, \dots, f_m$ sú dvakrát spojito diferencovatel'né.
- Úloha (P) má optimálne riešenie $x^*$.
- Platí Slaterova podmienka, t. j. existuje bod $x^0 \in \mathbb{R}^n$ taký, že

$$Ax^0 = b, \quad , f_i(x^0) < 0, \; i = 1, \dots, m.$$

Predpoklady zabezpečujú, že existuje optimálne riešenie $(u^*, v^*) \in \mathbb{R}^m \times \mathbb{R}^p$ duálnej úlohy, ktoré spolu s $x^*$ spĺňa systém KKT podmienok:

$$\boxed{\begin{aligned} A^*x = b,\ f_i(x^*) &\le 0, \quad i = 1, \ldots, m, \\ u^* &\ge 0, \\ u_i f_i(x^*) &= 0, \quad i = 1, \ldots, m \\ \nabla f_0(x^*) + \textstyle\sum_{i=1}^m u_i^* \nabla f_i(x^*) + A^T v^* &= 0. \end{aligned}}$$

Metódy vnútorného bodu riešia úlohu (P), resp. systém KKT aplikovaním Newtonovej metódy na

1. postupnosť' parametrizovaných úloh (neohraničených alebo s ohraničeniami typu rovností) $\longrightarrow$ **barérové metódy vnútorného bodu**,

resp.

2. postupnosť' perturbovaných (modifikovaných) systémov KKT podmienok $\longrightarrow$ **primárno-duálne metódy vnútorného bodu**

Zovšeobecnené úlohy konvexného programovania (SDP, SOCP) sa riešia zovšeobecnenými metédami vnútorného bodu.

## Logaritmická bariéra a centrálna trajektória

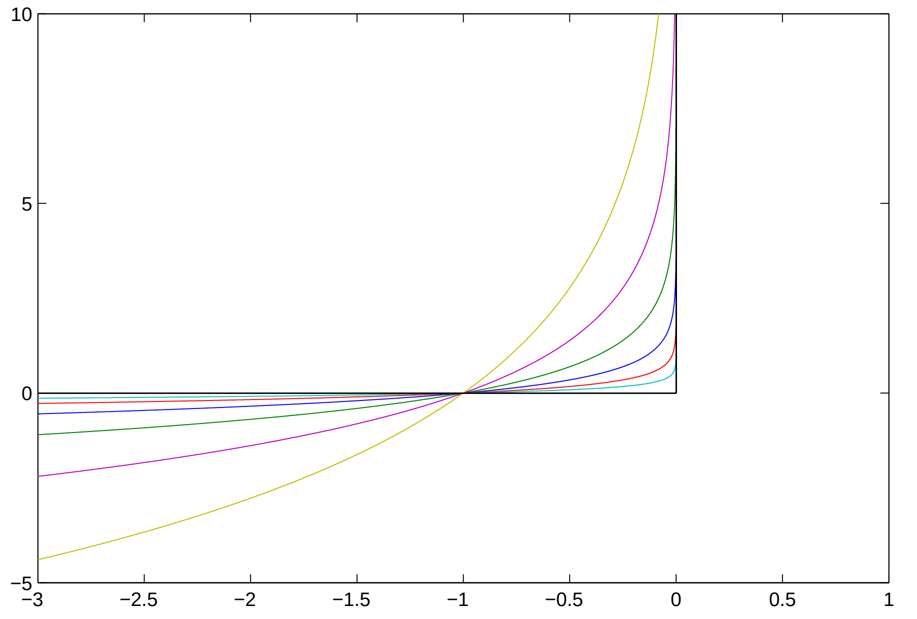
Ekvivalentná formulácia úlohy (P):

$$\begin{array}{ll} Min & f_0(x) + \sum_{i=1}^m I(f_i(x)) \\ & Ax = b, \end{array} \qquad (PI)$$

kde $I : \mathbb{R} \to \mathbb{R}$ je daná ako

$$I(t) = \begin{cases} 0 & t \leq 0 \\ \infty & t > 0 \end{cases}$$

- Úloha (PI) nemá diferencovatel'nú účelovú funkciu - Newtonova metóda sa nedá aplikovat'.
- Funkcia $I$ sa aproximuje diferencovatel'nou funkciou $B(t) = -\frac{1}{\mu} \ln(-t)$, kde $\mu > 0$ je parameter, ktorý určuje presnost' aproximácie.




Aproximácie funkcie $I$ pre hodnoty $\mu = \frac{1}{8}, \frac{1}{4}, \frac{1}{2}, 1, 2, 4$.

## Aproximácia úlohy (P):

$$\begin{array}{ll} Min & f_0(x) - \frac{1}{\mu}\sum_{i=1}^{m} \ln(-f_i(x)) \\ & Ax = b, \end{array} \qquad (PA)$$

Funkcia

$$B : \mathbb{R}_{--} \to \mathbb{R}, \qquad B(t) = -\frac{1}{\mu}\ln(-t)$$

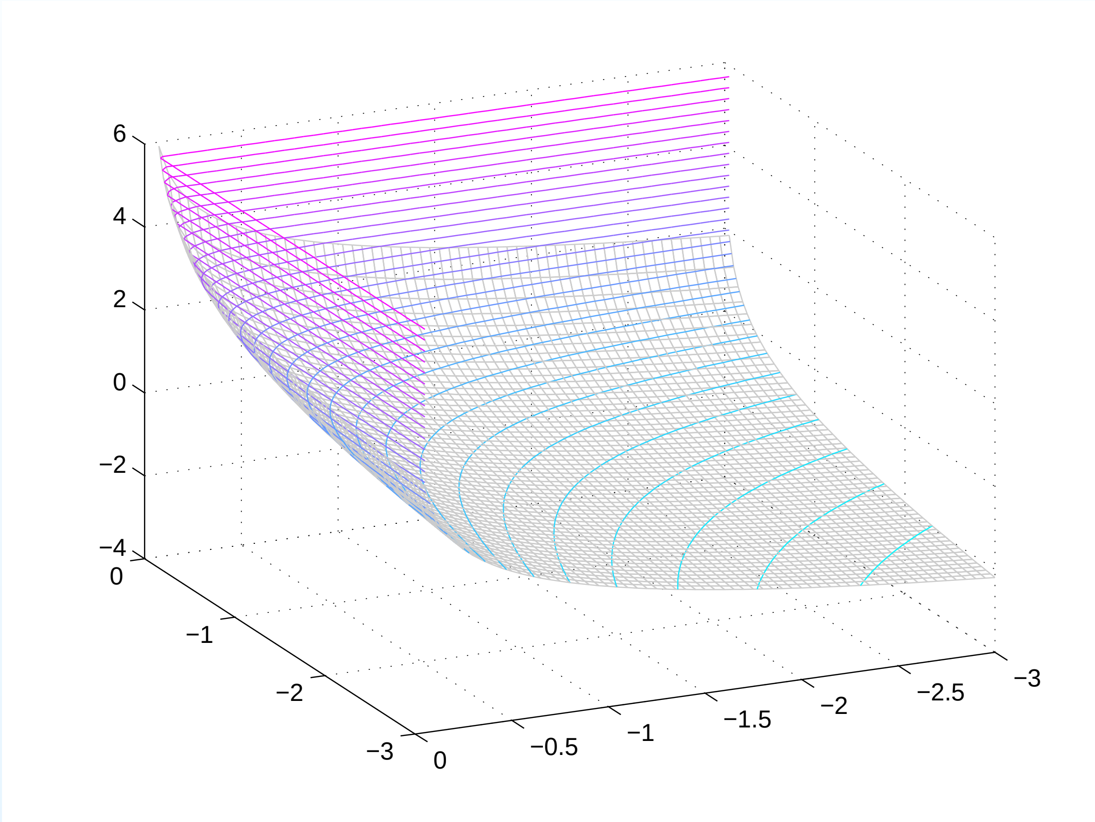
má vlastnosti:

- $\lim_{t\to 0-} B(t) = +\infty$
- $B'(t) > 0$ (rastúca)
- $B''(t) > 0$ (konvexná)

Funkcia

$$\phi(x) = -\sum_{i=1}^{m} \ln(-f_i(x))$$

s definičným oborom $\mathcal{D}(\phi) = \{x \mid f_i(x) < 0, i = 1, \ldots, m\}$ sa nazýva **logaritmická bariéra.**




Logaritmická bariéra pre ohraničenie $x \leq 0$.

- Bez ohl'adu na hodnotu parametra $\mu > 0$ hodnota logaritmickej bariéry prudko rastie pre $f_i(x) \to 0$.

- Úloha (PA) je len aproximáciou úlohy (PI), ale pre zväčšujúce sa hodnoty parametra $\mu$ sa úloha (PA) stále viac podobá na (PI).

- Pre velké hodnoty $\mu$ je však vel'mi náročné minimalizovat' funkciu $f_0(x) + \frac{1}{\mu}\phi(x)$ Newtonovou metódou - je potrebný dobrý štartovací bod.

- **Myšlienka MVB:** Rieši sa postupnost' parametrizovaných úloh (PA) pre postupne väčšie hodnoty parametra $\mu$, pričom štartovací bod pre Newtonovu metódu je optimálne riešenie úlohy pre predchádzajúcu hodnotu parametra $\mu$.

**Centrálna trajektória**

Riešme úlohu (PA) v tvare

$$\begin{aligned} Min \quad & \mu f_0(x) + \phi(x) \\ & Ax = b, \end{aligned} \qquad (PB)$$

Predpokladáme, že sa úloha dá riešit' Newtonovou metódou a že $\forall \mu > 0$ má jediné riešenie $x^*(\mu)$.
**Centrálna trajektória** prináležiaca úlohe (PB) je množina

$$c(\mu) = \{x^*(\mu) \mid \mu > 0\}.$$

Body $x^*(\mu)$ spĺňajú (nutné a postačujúce) **podmienky optimality**

- $Ax^*(\mu) = b, \; f_i(x^*(\mu)) < 0, i = 1, \ldots, m$
- $\exists \hat{v} \in \mathbb{R}^p : \; \mu \nabla f_0(x^*(\mu)) + \nabla \phi(x^*(\mu)) + A^T \hat{v} = 0.$

Každý bod centrálnej trajektórie určuje **duálne prípustný bod** (pôvodnej úlohy) a teda aj dolné ohraničenie na hodnotu $p^*$:

$$u_i^*(\mu) = -\frac{1}{\mu f_i(x^*(\mu))}, \quad i = 1, \dots, m, \qquad v^*(\mu) = \frac{\hat{v}}{\mu}$$

Zrejme $u^*(\mu) > 0$. Navyše z podmienok optimality (vyššie) vyplýva, že $x^*(\mu)$ je minimom Lagrangeovej funkcie $L(x, u, v)$ úlohy (P) pre $u = u^*(\mu), v = v^*(\mu)$.

**Duálna funkcia** spĺňa:

$$G(u^*(\mu), v^*(\mu)) =$$

$$f_0(x^*(\mu)) + \sum_{i=1}^m u_i^*(\mu) f_i(x^*(\mu)) + (v^*(\mu))^T (Ax^*(\mu) - b) = f_0(x^*(\mu)) - \frac{m}{\mu}$$

**Duálna medzera** pre $x^*(\mu), u^*(\mu), v^*(\mu)$ je $\frac{m}{\mu}$.

## Bariérová metóda

- Bod $x^*(\mu)$ je najviac $\frac{m}{\mu}$ suboptimálny.
- Teda $\varepsilon$-optimálne riešenie úlohy (P) by sme mohli nákst' vyriešením úlohy

$$\begin{array}{ll} Min & \frac{m}{\varepsilon} f_0(x) + \phi(x) \\ & Ax = b. \end{array}$$

- Táto metóda funguje len pre malé rozmery $m$, dobrý štartovací bod a nie vel'mi vysokú presnost'.
- **Jednoduché rozšírenie tejto metódy:** rieši sa postupnost' úloh typu (PB) pre rastúce hodnoty parametra $\mu$, pričom posledné nájdené optimálne riešenie slúži ako štartovací bod pre nasledujúcu úlohu.
- **Fiacco, McCormick** - 1960' - SUMT

## Algoritmus bariérovej metódy

Vstup: $x^0 \in \mathcal{P}$ - vnútorný bod, $\mu = \mu(0) > 0$,
$\beta > 1$ - koeficient rastu parametra $\mu$,
tolerancia $\varepsilon$.

Opakuj 1. (Centrujúci krok.) Rieš úlohu

$$\begin{aligned} Min \quad & \mu f_0(x) + \phi(x) \\ & Ax = b \end{aligned}$$

Newtonovou metódou so štartovacím bodom $x^0$ a nájdi optimálne riešenie $x^*(\mu)$.

2. $x^0 = x^*(\mu)$.
3. Ak $m/\mu < \varepsilon \longrightarrow$ KONIEC.
4. $\mu = \beta\mu$.

**Presnost' centrujúceho kroku**:

- Nie je nutné počí tat' presné riešenie $x^*(\mu)$.
- Príslušné $u^*(\mu), v^*(\mu)$ však potom nie sú duálne prípustné (rieši sa korekciou definujúcich vzt'ahov)
- Výpočet vel'mi preného riešenia vyžaduje len niekol'ko Newtonovych iterácií navyše - môžeme uvažovat' presné riešenia.

**Výber** $\beta$ (koeficient rastu parametra $\mu$)

- Ak je $\beta$ malé ($\approx 1$) - malý počet vnútorných (Newtonovych) iterácií, vel'ký počet vonkaších iterácií (centrujúcich krokov).
- Ak je $\beta$ vel'ké - vel'a vnútorných (Newtonovych) iterácií, malý počet vonkaších iterácií (centrujúcich krokov).
- V praxi $\beta \in [10, 20]$.

**Výber $\mu(0)$**

- Ak je známe nejaké duálne prípustné riešenie $(u^0, v^0)$ a $\nu = f_0(x^0) - G(u^0, v^0)$ je príslušná duálna medzera, tak môžeme vziať $\mu(0) = m/\nu$.
- Rieši sa minimalizačná úloha

$$Min_{\mu,v} \|\mu \nabla f_0(x^0) + \nabla \phi(x^0) + A^T v\|_2$$

**Analýza konvergencie**

- Riešime postupnosť úloh (PB) pre $\mu = \mu(0), \beta\mu(0), \beta^2\mu(0), \ldots$.
- Duálna medzera po $k$ centrujúcich krokoch je teda $\frac{m}{\beta^k \mu(0)}$.
- $\varepsilon$-presné riešenie teda dostaneme po

$$\left\lceil \frac{\ln \frac{m}{\varepsilon \mu(0)}}{\ln \beta} \right\rceil$$

krokoch.

**Newtonova metóda pre úlohy s ohraničeniami $Ax = b$**
Majme úlohu

$$\left.\begin{array}{ll} Min & F(x) \\ & Ax = b \end{array}\right\} \quad PO: Ax = b,\ \exists v: \nabla F(x) + A^T v = 0$$

V okolí $x$ funkciu $F$ aproximujeme **Taylorovym polynómom 2. rádu:**

$$\begin{array}{ll} Min & \hat{F}(x+s) = F(x) + \nabla F(x)^T s + \frac{1}{2} s^T \nabla^2 F(x) s \\ & A(x+s) = b \end{array}$$

**Podmienky optimality:**

$$\exists w: \begin{bmatrix} \nabla^2 F(x) & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} s \\ w \end{bmatrix} = \begin{bmatrix} -\nabla F(x) \\ 0 \end{bmatrix} \qquad (SNM)$$

## Algoritmus Newtonovej metódy

---

Vstup: Štartovací bod: $x: \; Ax = b$

Tolerancia $\varepsilon > 0$.

Opakuj 1. Vypočítaj gradient $\nabla F(x)$ a Hessovu maticu $\nabla^2 F(x)$

2. Ukončovacie kritérium: ak $\nu(x) < \varepsilon$ - KONIEC

3. Rieš systém podmienok optimality (SNM),

nájdi Newtonov krok $s = \triangle_{NK}$

4. Urči dĺžku kroku $\lambda$

5. $x = x + \lambda \triangle_{NK}$

---

## Ukončovacie kritérium

- $\nu(x) = \|\nabla F(x)\|_2$
- Alternatívne: $\nu(x) = \frac{1}{2}\triangle_{NK}^T(\nabla^2 F(x))^{-1}\triangle_{NK}$, v tomto prípade je $\nu(x)$ odhadom hodnoty $F(x) - p^*$ založenej na kvadratickej aproximácií $F$, t. j.

$$\nu(x) = F(x) - \inf\{\hat{F}(x+s) \mid A(x+s) = b\}.$$

## Dĺžka kroku

- Konštantný krok $\lambda = 1$
- Optimálny krok $\lambda = \arg\min_t F(x + t\triangle_{NM})$
- Približne optimálny krok.

Uvažujme opät' úlohu

$$\begin{array}{ll} Min & \mu f_0(x) + \phi(x) \\ & Ax = b, \end{array} \qquad (PB)$$

Systém (SNM) pre hl'adanie Newtonovského smeru má tvar

$$\begin{bmatrix} \mu\nabla^2 f_0(x) + \nabla^2\phi(x) & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} \triangle_{NM} \\ w \end{bmatrix} = \begin{bmatrix} -\mu\nabla f_0(x) + \nabla\phi(x) \\ 0 \end{bmatrix}$$

resp. v označení

$$g = \nabla f_0(x) + \frac{1}{\mu}\nabla\phi(x), \quad H = \nabla^2 f_0(x) + \frac{1}{\mu}\nabla^2\phi(x)$$

$$\begin{bmatrix} \mu H & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} \triangle_{NM} \\ w \end{bmatrix} = \begin{bmatrix} -\mu g \\ 0 \end{bmatrix}$$

Bod $x = x^*(\mu), u = u^*(\mu), v = v^*(\mu)$ spĺňa systém **modifikovaných KKT podmienok:**

- $Ax = b$
- $\nabla f_0(x) + \sum_{i=1}^{m} u_i \nabla f_i(x) + A^T v = 0$
- $-u_i f_i(x) = 1/\mu, \; i = 1, \ldots, m$

resp. po dosadení z poslednej podmienky:

$$\nabla f_0(x) + \sum_{i=1}^{m} \frac{1}{-\mu f_i(x)} \nabla f_i(x) + A^T v = \nabla f_0(x) + \frac{1}{\mu} \nabla \phi(x) + A^T v = 0, \; Ax = b$$

čo je systém $n + p$ rovníc s $n + p$ neznámymi.
Pripomíname:

- $\nabla \phi(x) = \sum_{i=1}^{m} \frac{1}{-f_i(x)} \nabla f_i(x)$,
- $\nabla^2 \phi(x) = \sum_{i=1}^{m} \frac{1}{f_i(x)^2} \nabla f_i(x) \nabla f_i(x)^T - \sum_{i=1}^{m} \frac{1}{f_i(x)} \nabla^2 f_i(x)$.

Riešme systém modifikovaných KKT Newtonovou metódou:

$$\nabla f_0(x+a) + \frac{1}{\mu}\nabla\phi(x+s) + A^T v \approx$$

$$\nabla f_0(x) + \frac{1}{\mu}\nabla\phi(x) + A^T v + \nabla^2 f_0(x)s + \frac{1}{\mu}\nabla^2\phi(x)s = 0,$$

$$A(x+s) - b = Ax - b + As = As = 0,$$

čo je ekvivalentné

$$Hs + A^T v = -g, \quad As = 0.$$

Newtonov smer $\triangle_{NM}$ a $w$ v centrujúcom kroku bariérovej metódy spĺňajú:

$$\mu H\triangle_{NM} + A^T w = -\mu g, \quad A\triangle_{NM} = 0.$$

Platí vzt'ah: $s = \triangle_{NM},\ v = w/\mu$.

## Primárno duálne metódy vnútorného bodu

- Na rozdiel od bariérových metód v algoritme nerozlišujeme vonkajšie a vnútorné iterácie.
- Newtonova metóda sa apolikuje na systém modifikovaných KKT podmienok.
- Iterácie nune negenerujú prípustné body.

Pre $\mu > 0$ definujme systém

$$r(x,u,v) = \begin{bmatrix} r_D \\ r_C \\ r_P \end{bmatrix} = \begin{bmatrix} \nabla f_0(x) + Df(x)^T u + A^T v \\ -diag(u)f(x) - \frac{1}{\mu}\mathbf{1} \\ Ax - b \end{bmatrix} = 0$$

kde $f(x) = \begin{bmatrix} f_1(x) \\ \vdots \\ f_m(x) \end{bmatrix}, \quad Df(x) = \begin{bmatrix} \nabla f_1(x)^T \\ \vdots \\ \nabla f_m(x)^T \end{bmatrix}$

Ak $(x, u, v)$ spĺňa systém $r(x, u, v) = 0$, tak
$x = x^*(\mu), u = u^*(\mu), v = v^*(\mu)$ a duálna medzera je rovná $m/\mu$.

Označme $y = (x, u, v)$ **a** $\triangle y = (\triangle x, \triangle u, \triangle v)$ a riešme systém
$r(y) = r(x, u, v) = 0$ **Newtonovou metódou:**

$$r(y + \triangle y) \approx r(y) + Dr(y)\triangle y = 0,$$

t.j.

$$\begin{bmatrix} \nabla^2 f_0(x) + \sum_{i=1}^{m} u_i \nabla^2 f_i(x) & Df(x)^T & A^T \\ -diag(u)Df(x) & -diag(f(x)) & 0 \\ A & 0 & 0 \end{bmatrix} \begin{bmatrix} \triangle x \\ \triangle u \\ \triangle v \end{bmatrix} = - \begin{bmatrix} r_D \\ r_C \\ r_P \end{bmatrix} \quad (PD)$$

Riešenie $\triangle y = (\triangle x, \triangle u, \triangle v)$ systému (PD) sa nazýva **primárno duálny smer**.

**Náhradná duálna medzera** $\eta = -f(x)^T u = m/\mu$.

## Algoritmus primárno-duálnej metódy vnútorného bodu

---

| | |
|---|---|
| Vstup: | $x^0 \in \mathcal{P}$ - vnútorný bod, $u > 0, v$, |
| | Tolerancie $\varepsilon_P > 0, \varepsilon_O > 0$, faktor redukcie $\beta > 0$. |
| Opakuj | 1. $\mu = \beta m/\eta$ |
| | 2. Rieš systém (PD) a nájdi $\triangle y$. |
| | 3. Urči dĺžku kroku $\lambda$. |
| | 4. $y = y + \lambda \triangle y$ |
| pokiaľ | $\|r_P\| \leq \varepsilon_P,\ \|r_D\| \leq \varepsilon_P,\ \eta \leq \varepsilon_O$. |

---

## Samokonkordantnosť

Predpoklady pre konvergenciu Newtonovej metódy z prednášky z Nelineárneho programovania:

- $F$ má spojité 2. parciálne derivácie
- Lipschitzovskosť Hessovej matice: $\exists L > 0 : \ \forall x, y$:

$$\|\nabla^2 F(x) - \nabla^2 F(y)\|_2 \leq L\|x - y\|_2$$

- Existuje ostré lokálne minimum $\hat{x}$ funkcie $F$.

Počet iterácií na dosiahnutie $\varepsilon$-presného riešenia závisí od neznámych konštánt $L, \lambda_{min}(\nabla^2(\hat{x}))$, ktoré navyše nie sú škálovo invariantné.

V r. 1994 **Nesterov a Nemirovski** vyriešili tento problém a "náhradný" predpoklad nazvali samokonkordantnosť:

**Definícia:** Konvexná funkcia $f : \mathbb{R} \to \mathbb{R}$ sa nazýva **samokonkordantná**, ak $\forall x \in \mathbb{R}$ platí

$$|f'''(x)| \leq 2f''(x)^{\frac{3}{2}}. \tag{SK}$$

Funkcia $g : \mathbb{R}^n \to \mathbb{R}$ sa nazýva samokonkordantná, ak každé jej zúženie na priamku je samokonkordantná funkcia, t. j. ak pre každé $x, y \in \mathbb{R}^n$ je funkcia

$$f : \mathbb{R} \to \mathbb{R}, \quad f(t) = g(x + ty)$$

samokonkordantná v zmysle (SK).