Obsah

1 Optimalizační metody

Optimalizační metody slouží k hledání minima nebo maxima účelové funkce

$$\min f(x).$$

Tato kapitola představuje **přehledovou mapu témat**, nikoli detailní výklad. Slouží k rychlé orientaci a připomenutí hlavních oblastí předmětu.

Základní rozdělení metod

- **Metody bez derivací** → nevyžadují znalost derivací, pracují pouze s hodnotami funkce
  → typicky minimalizace jedné proměnné
  → unimodalita, interval a,b
- **Metody s derivacemi** → využívají první a často i druhou derivaci
  → hledání kořenů a minim pomocí iteračních metod

Minimalizace podle počtu proměnných

- **Jedna proměnná**
  → základní principy optimalizace
  → interval, přesnost, unimodalita
  → bezderivační a derivační metody
- **Více proměnných (bez omezení)**
  → proměnná $x \in \mathbb{R}^n$
  → gradient, Hessova matice
  → gradientní metody, Newtonovy metody

2 Metody s více proměnnými

Metody s více proměnnými se zabývají minimalizací funkcí, jejichž argumentem je vektor proměnných. Navazují na metody jedné proměnné, ale pracují s geometricky složitějšími objekty, jako jsou gradienty a matice druhých derivací.

2.0.1 Základní formulace úlohy

- obecná optimalizační úloha má tvar

$$\min f(x), \quad x \in \mathbb{R}^n$$

- proměnná x je vektor
- úloha **nemá omezení vazbami**
- hledáme lokální nebo globální minimum funkce

2.0.2 Minimum funkce více proměnných

- minimum je bod, kde má funkce nejmenší hodnotu v okolí
- nutná podmínka minima

$$\nabla f(x^*) = 0$$

- postačující podmínka (lokální minimum)
  - Hessova matice je **pozitivně definitní**

2.0.3 Gradient

- **gradient** je vektor prvních parciálních derivací

$$\nabla f(x) = \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{pmatrix}$$

- určuje směr **největšího růstu** funkce
- směr největšího poklesu je opačný směr gradientu

2.0.4 Hessova matice

- **Hessova matice** je matice druhých parciálních derivací

$$H(x) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{pmatrix}$$

- popisuje lokální zakřivení funkce
- používá se v Newtonových metodách

2.0.5 Základní skupiny metod

- **Gradientní metody**
  - využívají pouze gradient
  - jednoduché, ale pomalejší
  - viz Metoda největšího spádu
- **Newtonovy metody**
  - využívají gradient i Hessovu matici
  - velmi rychlé v okolí minima
  - viz Newton-Rhapsonova
- **Bezderivační metody**
  - nevyžadují derivace
  - pracují pouze s hodnotami funkce
  - vhodné pro malé dimenze
  - viz Nelder-Meadova

2.0.6 Charakteristika metod

- metody jsou většinou **iterační**
- vyžadují **počáteční bod** x_0
- konvergence závisí na:
  - vlastnostech funkce
  - volbě počátečního bodu
  - použité metodě

Optimalizace s omezením

- **Rovnostní omezení**
  → pohyb po omezené množině
  → kolmost gradientů
  → metoda Lagrangeových multiplikátorů

Matematické programování

- **Lineární programování (LP)**
  → lineární účelová funkce
  → lineární omezení
- **Kvadratické programování (QP)** → kvadratická účelová funkce
  → pozitivně definitivní matice

- **Nelineární programování (NLP)**
  → nelineární funkce a omezení
  → více lokálních extrémů

→ shrnuto v souboru Matematické programování

Tento soubor slouží jako **orientační uzel (mapa)**, ze kterého vedou odkazy na jednotlivá detailní témata.

3 Metody bez derivací

Metody bez derivací slouží k minimalizaci funkce jedné proměnné v situacích, kdy nejsou k dispozici derivace nebo je nechceme používat. Pracují výhradně s hodnotami funkce a jejich cílem je postupně zúžit interval, ve kterém se nachází minimum.

Základní úloha

- hledáme minimum funkce jedné proměnné
- obecný tvar úlohy

$$\min f(x), \quad x \in \langle a, b \rangle \text{ nebo } x \in \mathbb{R}$$

- výsledkem je hodnota x^*, pro kterou má funkce nejmenší hodnotu

Unimodalita funkce

- **unimodální funkce** má právě jedno minimum
- funkce:
  - nejprve klesá
  - poté roste
- unimodalita je **základní podmínkou** použití metod bez derivací
- minimum je bod, kde se chování funkce mění z klesajícího na rostoucí

Interval řešení

- **interval** $\langle a, b \rangle$ určuje oblast, ve které minimum hledáme
- funkce musí být:
  - spojitá
  - ideálně unimodální v celém intervalu
- metody bez derivací:
  - interval postupně zmenšují
  - minimum zůstává vždy uvnitř aktuálního intervalu

Kritérium ukončení

- Kritérium ukončení je dáno požadovanou přesností
- vyjadřuje, jak „dost dobrý“ výsledek požadujeme

Používaná kritéria:

- **Délka intervalu**

$$|a - b| < \epsilon$$

- **Velikost derivace**
  → používá se hlavně u metod s derivacemi

$$|f'(x)| < \epsilon$$

- **Změna mezi iteracemi**

$$|x_{k+1} - x_k| < \epsilon$$

Charakteristika metod bez derivací

- **nevyžadují derivace**
- pracují pouze s hodnotami $f(x)$
- porovnávají funkční hodnoty v různých bodech
- iterativně zužují interval s minimem
- jsou:
  - robustní
  - jednoduché na pochopení
- nevýhody:
  - pomalejší než derivační metody
  - použitelné především pro jednu proměnnou

Přehled základních metod

- **Trisekce** → rozdělení intervalu na třetiny
  → jednoduchá, ale pomalejší
- Zlatý řez → efektivní rozdělení intervalu pomocí konstantního poměru
  → méně výpočtů funkčních hodnot
- Fibonacciho metoda → optimalizovaná varianta zlatého řezu
  → vyšší přesnost
  → složitější implementace

3.1 Trisekce

Metoda trisekce je **bezderivační metoda** pro minimalizaci **unimodální funkce jedné proměnné**. Patří mezi nejjednodušší intervalové metody a slouží především k pochopení základního principu zúžování intervalu s minimem.

Tato metoda může být dále optimalizována viz metoda Zlatého řezu nebo Fibonacciho

Základní myšlenka

- minimum funkce leží v intervalu $\langle a, b \rangle$
- interval se v každém kroku rozdělí na **tři části**
- porovnáním funkčních hodnot se jedna třetina intervalu zahodí
- minimum zůstává vždy uvnitř nového, menšího intervalu

Rozdělení intervalu V každé iteraci se zvolí dva body:

$$u = a + \frac{b-a}{3}, \quad v = b - \frac{b-a}{3}$$

- platí $a < u < v < b$
- spočítají se hodnoty $f(u)$ a $f(v)$

Rozhodovací pravidlo

- pokud $f(u) < f(v)$
  $\to$ minimum leží v intervalu $\langle a, v \rangle$
- pokud $f(u) > f(v)$
  $\to$ minimum leží v intervalu $\langle u, b \rangle$
- v každém kroku se délka intervalu zmenší na $\frac{2}{3}$ původní délky

Podmínky použití

- funkce je:
  - **spojitá**
  - **unimodální** na intervalu $\langle a, b \rangle$
- derivace nejsou potřeba
- známe počáteční interval, ve kterém leží minimum

Přesnost a ukončení

- algoritmus končí, pokud:
$$|a - b| < \epsilon$$
- výsledkem je interval obsahující bod minima
- za aproximaci minima se často bere střed intervalu

Vlastnosti metody

- **výhody**
  - velmi jednoduchá
  - snadná implementace
  - robustní
- **nevýhody**
  - pomalejší než zlatý řez
  - v každé iteraci vyžaduje dva nové výpočty $f(x)$
  - nevyužívá dříve vypočtené hodnoty

3.2 Zlatý řez

Metoda zlatého řezu je **bezderivační intervalová metoda** pro minimalizaci **unimodální funkce jedné proměnné**. Oproti trisekci je efektivnější, protože **znovu využívá již vypočtené funkční hodnoty** a snižuje počet volání funkce.

Základní myšlenka

- minimum funkce leží v intervalu $\langle a, b\rangle$
- interval se zužuje pomocí **pevného poměru**
- v každé iteraci se zahazuje část intervalu, ve které minimum neleží
- jedna funkční hodnota se přenáší do další iterace

Zlatý poměr

- používá se konstanta

$$\varphi = \frac{\sqrt{5}-1}{2} \approx 0{,}618$$

- doplňkový poměr

$$1 - \varphi \approx 0{,}382$$

Volba bodů v intervalu V každé iteraci se volí body:

$$u = a + (1-\varphi)(b-a), \quad v = a + \varphi(b-a)$$

- platí $a < u < v < b$
- počítají se hodnoty $f(u)$ a $f(v)$
- při zúžení intervalu se **jeden z bodů znovu použije**

Rozhodovací pravidlo

- pokud $f(u) < f(v)$
  $\rightarrow$ minimum leží v intervalu $\langle a, v \rangle$
- pokud $f(u) > f(v)$
  $\rightarrow$ minimum leží v intervalu $\langle u, b \rangle$
- nový interval je kratší než původní v konstantním poměru

Podmínky použití

- funkce je:
  - spojitá
  - **unimodální** na intervalu $\langle a, b \rangle$
- derivace nejsou potřeba
- známe interval obsahující minimum

Přesnost a ukončení

- algoritmus se ukončí, pokud platí

$$|a - b| < \varepsilon$$

- výsledkem je interval obsahující minimum
- aproximací minima je obvykle střed intervalu

Vlastnosti metody

- **výhody**
  - méně výpočtů funkčních hodnot než trisekce
  - efektivní a stabilní
  - jednoduchá implementace
- **nevýhody**
  - pomalejší než derivační metody
  - použitelná pouze pro jednu proměnnou

Zařazení mezi metody

- metoda **bez derivací**
- minimalizace **jedné proměnné**
- intervalová metoda

3.3 Fibonacciho metoda

Fibonacciho metoda je **bezderivační intervalová metoda** pro minimalizaci **unimodální funkce jedné proměnné**. Představuje optimalizovanou variantu metody zlatého řezu, která využívá **Fibonacciho posloupnost** k přesně řízenému zmenšování intervalu.

Základní myšlenka

- minimum funkce leží v intervalu $\langle a, b \rangle$
- počet iterací je **předem dán**
- interval se zužuje pomocí poměrů vycházejících z Fibonacciho posloupnosti
- metoda minimalizuje počet výpočtů funkčních hodnot

Fibonacciho posloupnost

- posloupnost je definována vztahem

$$F_0 = 0, \quad F_1 = 1, \quad F_k = F_{k-1} + F_{k-2}$$

- poměry po sobě jdoucích členů se blíží zlatému poměru
- tyto poměry určují polohu testovacích bodů v intervalu

Volba bodů v intervalu

- pro zvolený počet kroků N se v každé iteraci určují body pomocí poměrů

$$u = a + \frac{F_{N-2}}{F_N}(b - a), \quad v = a + \frac{F_{N-1}}{F_N}(b - a)$$

- porovnávají se hodnoty $f(u)$ a $f(v)$
- interval se zúží podobně jako u zlatého řezu

Rozhodovací pravidlo

- pokud $f(u) < f(v)$
  $\rightarrow$ minimum leží v intervalu $\langle a, v \rangle$
- pokud $f(u) > f(v)$
  $\rightarrow$ minimum leží v intervalu $\langle u, b \rangle$
- jeden bod se přenáší do další iterace

Přesnost a ukončení

- počet iterací N se volí tak, aby výsledná délka intervalu splňovala

$$|a - b| < \varepsilon$$

- přesnost je tedy **řízena dopředu** volbou N
- výsledkem je interval obsahující bod minima

Podmínky použití

- funkce je:
  - spojitá
  - **unimodální** na intervalu $\langle a, b \rangle$
- derivace nejsou potřeba
- známe požadovanou přesnost ε

Vlastnosti metody

- **výhody**
  - velmi efektivní
  - minimální počet výpočtů $f(x)$
  - přesně řízená přesnost
- **nevýhody**
  - složitější implementace
  - nutnost znát požadovanou přesnost předem
  - méně flexibilní než zlatý řez

Zařazení mezi metody

- metoda **bez derivací**
- minimalizace **jedné proměnné**
- intervalová metoda

3.4 Nelder–Meadova metoda (simplex)

Nelder–Meadova metoda je **bezderivační metoda** pro minimalizaci **funkce více proměnných bez omezení**. Využívá geometrickou představu **simplexu** a pracuje pouze s hodnotami funkce, nikoli s jejími derivacemi.

Základní úloha

- minimalizace funkce tvaru

$$\min f(x), \quad x \in \mathbb{R}^n$$

- funkce nemá omezení vazbami
- derivace nejsou známy nebo je nechceme používat

Základní myšlenka

- metoda pracuje s **simplexem**:
  - v n-rozměrném prostoru má simplex $n + 1$ vrcholů
- v každém kroku se:
  - vyhodnotí funkční hodnoty ve vrcholech
  - porovnají se hodnoty
  - nejhorší bod je nahrazen lepším
- simplex se postupně posouvá směrem k minimu

Počáteční simplex

- volí se $n + 1$ bodů:
  - musí vyplnit okolí hledaného minima
- kvalita počátečního simplexu výrazně ovlivňuje chování metody

Hodnocení bodů

- ve vrcholech simplexu se spočítají hodnoty funkce
- body se seřadí podle velikosti funkčních hodnot:

$$f(x_{\text{nejlepší}}) < f(x_{\text{prostřední}}) < f(x_{\text{nejhorší}})$$

Základní operace se simplexem Reflexe

- **základní krok metody**
- nejhorší bod se odrazí přes těžiště zbylých bodů
- pokud je nový bod lepší než prostřední, je přijat

Expanze

- pokud je reflexní bod lepší než nejlepší
- simplex se zkusí rozšířit dále stejným směrem

- typický koeficient:

$$\alpha > 1 \quad (\text{často } \alpha = 2)$$

Kontrakce

- pokud reflexe nepřinese zlepšení
- simplex se zmenší směrem k lepším bodům
- typický koeficient:

$$\beta \in (0, 1) \quad (\text{často } \beta = 0{,}5)$$

Redukce

- použije se, pokud selžou předchozí kroky
- celý simplex se smrskne k nejlepšímu bodu
- používá se výjimečně

Podmínky použití

- funkce musí být **spočitatelná**
- derivace nejsou potřeba
- metoda je vhodná pro:
  - malé dimenze (typicky $n < 10$)

Přesnost a ukončení

- výpočet se ukončí, pokud platí

$$\max f(x_i) - \min f(x_i) < \varepsilon$$

- nebo po dosažení maximálního počtu iterací

Vlastnosti metody

- **výhody**
  - jednoduchá implementace
  - bez derivací
  - použitelná pro složité funkce
- **nevýhody**
  - pomalá
  - nespolehlivá ve vyšších dimenzích
  - nemá teoretické záruky konvergence

Zařazení mezi metody

- metoda **bez derivací**
- minimalizace **více proměnných**
- geometrická metoda (simplex)

$\rightarrow$ alternativa k derivačním metodám, zejména Metodě největšího spádu

4 Metody s derivacemi

Metody s derivacemi slouží k hledání **kořenů** a **minim funkcí** pomocí informací o derivacích účelové funkce. Oproti metodám bez derivací jsou zpravidla **rychlejší**, ale kladou vyšší nároky na hladkost funkce a volbu počátečního bodu.

Základní úloha

- minimalizace funkce jedné proměnné

$$\min f(x)$$

- nebo hledání kořene funkce

$$f(x) = 0$$

- metody jsou obvykle **iterační**
- vyžadují **počáteční bod** x_0

Kořen a minimum funkce

- **kořen funkce**
  - hodnota x, pro kterou platí

$$f(x) = 0$$

  - funkce protíná nebo se dotýká osy x
- **minimum funkce**
  - bod, kde má funkce nejmenší hodnotu
  - v minimu platí

$$f'(x^*) = 0, \quad f''(x^*) > 0$$

- hledání minima se často převádí na hledání kořene derivace

Charakter metod s derivacemi

- využívají:
  - **první derivaci**
  - často i **druhou derivaci**
- rozdíl mezi metodami není v principu, ale v cíli:
  - hledání kořene
  - hledání minima
- typickým zástupcem je **Newtonova metoda**

Počáteční bod

- **počáteční bod** x_0 je nutný pro iterační metody
- musí:
  - ležet v oblasti, kde existují derivace
  - být rozumně blízko řešení
- špatná volba může vést:
  - ke konvergenci k maximu
  - k inflexnímu bodu
  - k divergenci metody

Přesnost a ukončení

- výpočet se ukončuje při splnění některého z kritérií:
  - **velikost derivace**
    $$|f'(x_k)| < \varepsilon$$
  - **změna mezi iteracemi**
    $$|x_{k+1} - x_k| < \varepsilon$$
  - **maximální počet iterací**

Výpočet derivací

- **analytický**
  - derivace dány vzorcem
  - rychlý a přesný
- **numerický**
  - derivace jsou pouze odhadnuty
  - pomalejší
  - méně stabilní

Vlastnosti metod

- **výhody**
  - rychlá konvergence
  - vysoká přesnost
- **nevýhody**
  - nutnost znalosti derivací
  - citlivost na počáteční bod
  - vyšší nároky na hladkost funkce

Zařazení mezi metody

- metody **s derivacemi**
- minimalizace **jedné proměnné**
- iterační metody

$\rightarrow$ konkrétní postup je rozpracován v kapitole Newtonova metoda $\rightarrow$ případně pro více proměnných používáme metodu Newton-Rhapsonova

4.1 Newtonova metoda (1D)

Newtonova metoda je **iterační derivační metoda** používaná pro hledání **kořene funkce** nebo **minima funkce jedné proměnné**. Patří mezi nejdůležitější metody optimalizace díky své **rychlé konvergenci** v okolí řešení.

Základní princip

- metoda vychází z **lokální lineární aproximace** funkce
- v každém kroku se:
  - aproximuje funkce tečnou
  - průsečík tečny s osou x určí nový odhad řešení
- metoda vyžaduje **počáteční bod** x_0

Newtonova metoda pro hledání kořene

- hledáme bod, kde platí
$$f(x) = 0$$
- iterační vztah má tvar
$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}$$

Vlastnosti

- velmi rychlá konvergence v okolí kořene
- citlivá na volbu počátečního bodu
- může divergovat nebo konvergovat k jinému kořeni

Newtonova metoda pro hledání minima

- minimum funkce nastává v bodě, kde
$$f'(x^*) = 0$$
- úloha se převede na hledání kořene derivace
- iterační vztah má tvar
$$x_{k+1} = x_k - \frac{f'(x_k)}{f''(x_k)}$$

Podmínky použití

- funkce je:
  - diferencovatelná
  - ideálně **alespoň dvakrát diferencovatelná**
- v bodě minima platí

$$f''(x^*) > 0$$

- derivace existují v okolí řešení

Počáteční bod

- volba počátečního bodu x_0 je **kritická**
- špatná volba může vést:
  - ke konvergenci k maximu
  - k inflexnímu bodu
  - k divergenci metody
- počáteční bod volíme:
  - v blízkosti očekávaného minima
  - v oblasti, kde existují derivace

Přesnost a ukončení

- výpočet se ukončí, pokud platí alespoň jedna z podmínek:

$$|x_{k+1} - x_k| < \varepsilon$$

$$|f'(x_k)| < \varepsilon$$

- případně po dosažení maximálního počtu iterací

Výhody a nevýhody

- **výhody**
  - velmi rychlá konvergence
  - vysoká přesnost
- **nevýhody**
  - nutnost znalosti derivací
  - citlivost na počáteční bod
  - může selhat daleko od řešení

Zařazení mezi metody

- metoda **s derivacemi**
- minimalizace **jedné proměnné**
- iterační metoda

4.2 Newtonova–Rhapsonova metoda

Newtonova–Rhapsonova metoda je **iterační derivační metoda** pro minimalizaci **funkce více proměnných bez omezení**. Oproti gradientním metodám využívá kromě gradientu také **Hessovu matici**, což jí umožňuje velmi rychlou konvergenci v okolí minima.

Základní úloha

- minimalizace funkce tvaru

$$\min f(x), \quad x \in \mathbb{R}^n$$

- funkce nemá omezení vazbami
- gradient i Hessova matice jsou známy

Základní myšlenka

- funkce je v okolí aktuálního bodu **nahrazena kvadratickou aproximací**
- minimum této aproximace lze určit analyticky
- nový bod vznikne skokem přímo k minimu aproximující funkce

Iterační vztah

- obecný iterační krok má tvar

$$x_{k+1} = x_k - H^{-1}(x_k)\nabla f(x_k)$$

- $H(x_k)$ je Hessova matice funkce v bodě x_k
- metoda využívá informaci o zakřivení funkce

Podmínky použití

- funkce je:
  - diferencovatelná
  - alespoň dvakrát diferencovatelná v okolí řešení
- gradient existuje
- Hessova matice:
  - je invertibilní
  - je ideálně **pozitivně definitní**

Pozitivně definitní Hessova matice

- v bodě minima musí platit

$$v^T H v > 0, \quad v \neq 0$$

- funkce má v okolí bodu **miskovitý tvar**
- všechna vlastní čísla Hessovy matice jsou kladná
- v opačném případě může metoda směřovat:
  - k maximu
  - k sedlovému bodu

Počáteční bod

- volba počátečního bodu x_0 je **kritická**
- metoda:
  - je velmi rychlá blízko minima
  - může selhat daleko od minima
- špatný počátek může vést k divergenci

Přesnost a ukončení

- výpočet se ukončí při splnění některého z kritérií:

$$|\nabla f(x_k)| < \varepsilon$$

$$|x_{k+1} - x_k| < \varepsilon$$

- případně po dosažení maximálního počtu iterací

Vlastnosti metody

- **výhody**
  - velmi rychlá konvergence (kvadratická)
  - malý počet iterací
- **nevýhody**
  - drahý výpočet (Hessova matice, inverze)
  - citlivost na počáteční bod
  - může selhat mimo okolí minima

Zařazení mezi metody

- metoda **s derivacemi**
- minimalizace **více proměnných**
- Newtonova metoda

$\rightarrow$ bezderivační alternativa je Nelder–Meadova metoda

4.3 Metoda největšího spádu

Metoda největšího spádu je **iterační derivační metoda** pro minimalizaci **funkce více proměnných bez omezení**. Patří mezi základní gradientní metody a slouží jako výchozí nástroj pro pochopení chování optimalizačních algoritmů ve vícerozměrném prostoru.

Základní úloha

- minimalizace funkce tvaru

$$\min f(x), \quad x \in \mathbb{R}^n$$

- funkce nemá omezení vazbami
- gradient funkce je znám

Základní myšlenka

- **gradient** určuje směr největšího růstu funkce
- pro minimalizaci se postupuje **opačným směrem**
- v každém kroku se provede posun ve směru záporného gradientu

Iterační vztah

- obecný tvar kroku

$$x_{k+1} = x_k - \alpha_k \nabla f(x_k)$$

- α_k je **velikost kroku**
- volba kroku zásadně ovlivňuje chování metody

Volba velikosti kroku

- **konstantní krok**
  - jednoduchý
  - může vést k pomalé konvergenci nebo kmitání
- **proměnný krok**
  - volen např. minimalizací podél směru gradientu
  - stabilnější chování metody

Podmínky použití

- gradient existuje v celé oblasti řešení
- funkce je:
  - spojitá
  - diferencovatelná
- metoda konverguje zejména pro **konvexní funkce**

Vlastnosti metody

- **výhody**
  - jednoduchá
  - robustní
  - levná jedna iterace
- **nevýhody**
  - pomalá konvergence
  - klikatí se v úzkých údolích
  - silná závislost na volbě kroku

Přesnost a ukončení

- výpočet se ukončí, pokud platí

  $$|\nabla f(x_k)| < \varepsilon$$

  nebo

  $$|x_{k+1} - x_k| < \varepsilon$$
- případně po dosažení maximálního počtu iterací

Zařazení mezi metody

- metoda **s derivacemi**
- minimalizace **více proměnných**
- gradientní metoda

$\rightarrow$ rychlejší, ale náročnější alternativa je Newtonova–Rhapsonova metoda

5 Optimalizace s omezením

5.1 Lagrangeovy multiplikátory

Metoda Lagrangeových multiplikátorů slouží k hledání **extrémů funkce více proměnných s omezením**. Jedná se o **analytickou derivační metodu**, která umožňuje určit kandidáty minima nebo maxima na omezené množině.

Základní úloha

- hledáme extrém funkce

$$\min f(x)$$

- za rovnostního omezení

$$g(x) = 0$$

- proměnná $x \in \mathbb{R}^n$

Základní myšlenka

- pohybujeme se pouze po množině dané omezením
- v bodě extrému platí:
  - **gradient funkce je rovnoběžný s gradientem omezení**
- gradienty jsou tedy **lineárně závislé**

Nutná podmínka extrému

- existuje skalár λ tak, že

$$\nabla f(x) = \lambda \nabla g(x)$$

- λ je **Lagrangeův multiplikátor**
- multiplikátor vyjadřuje citlivost extrému na změnu omezení

Lagrangián

- zavede se pomocná funkce

$$L(x, \lambda) = f(x) + \lambda g(x)$$

- extrémy se hledají řešením soustavy rovnic

$$\nabla L(x, \lambda) = 0$$

Postup výpočtu

1) zavedení Lagrangiánu

2) výpočet parciálních derivací podle všech proměnných a multiplikátoru

3) řešení soustavy rovnic

4) získání **kandidátů extrému**

Podmínky použití

- funkce $f(x)$ i omezení $g(x)$ jsou diferencovatelné
- gradient omezení není nulový
- metoda řeší **pouze rovnostní omezení**

Vlastnosti metody

- **výhody**
  - elegantní
  - přesná
  - analytická
- **nevýhody**
  - složitější výpočty
  - neřeší nerovnostní omezení
  - poskytuje pouze kandidáty extrému

Použití

- optimalizace více proměnných
- úlohy s rovnostním omezením
- známé analytické derivace

Zařazení mezi metody

- metoda **s derivacemi**
- optimalizace **s omezením**
- analytická metoda

$\rightarrow$ obecnější rámec optimalizace s omezeními je uveden v kapitole Matematické programování

6 Matematické programování

Matematické programování představuje obecný rámec **optimalizace s omezeními**, ve kterém hledáme minimum nebo maximum účelové funkce při splnění rovnostních a nerovnostních vazeb. Důraz není kladen na konkrétní iterační postup, ale na **strukturu úlohy** a vlastnosti řešení.

6.1 Základní formulace úlohy

- obecná optimalizační úloha

$$\min f(x)$$

- za podmínek

$$g(x) = 0, \quad h(x) \leq 0$$

- proměnné jsou omezeny vazbami
- řešení leží v **přípustné oblasti**

6.2 Lineární omezení

- lineární omezení mají tvar

$$Ax \leq b$$

- vytvářejí oblast přípustných řešení
- geometrická interpretace:
  - v 2D mnohoúhelník
  - ve vyšších dimenzích polyedr
- optimalizace probíhá pouze uvnitř této oblasti

6.3 Lineární účelová funkce

- lineární účelová funkce:
  - přímka (2D)
  - rovina (3D)
- minimum nebo maximum vzniká na **hranici přípustné oblasti**
- vnitřek oblasti není překážkou

6.4 Lineární programování (LP)

- účelová funkce:
  - lineární
- omezení:

  - lineární
- proměnné:
  - spojité

Použití: - ekonomika - plánování výroby - alokace zdrojů
Vlastnosti: - úloha je **konvexní** - platí: - lokální minimum = globální minimum

6.5 Kvadratické programování (QP)

- účelová funkce:
  - kvadratická

$$\min \frac{1}{2}x^T Q x + c^T x$$

- omezení:
  - lineární

Použití: - kvadratické členy v zadání - aproximace nelineárních úloh
Vlastnosti: - pokud je matice Q pozitivně definitivní - úloha je konvexní - řešení je jednoznačné

6.6 Nelineární programování (NLP)

- účelová funkce:
  - nelineární
- omezení:
  - lineární i nelineární

Použití: - výskyt funkcí typu - sin, cos, exp - vyšší mocniny
Vlastnosti: - může existovat více lokálních minim - bez konvexity: - není zaručeno globální řešení - často se používají Lagrangeovy multiplikátory

6.7 Charakteristika matematického programování

- optimalizace s omezeními
- důraz na:
  - strukturu účelové funkce
  - typ omezení
- volba metody závisí na:
  - linearitě
  - konvexitě
  - rozměru úlohy