Nelineární optimalizace s omezeními
- obecně

Většinu spojitých optimalizačních problémů lze vyjádřit následujícím způsobem:

minimalizuj:

$f(x), x \in R^n$

za předpokladu:

$c_i(x) = 0, i \in E$

$c_i(x) \geq 0, i \in I$

Nelineární optimalizace s omezeními
- obecně II

Účelová funkce f, stejně jako funkce c_i, určující omezení, jsou zobrazení z R^n do R.

E množina omezení ve formě rovnic

I množina omezení ve formě nerovnic

Dále budeme kvůli stručnosti označovat gradienty ∇c_i jako a_i.

Nelineární optimalizace s omezeními - příklad

Potravinářský velkoobchod má k dispozici sklad na 100 tun obilovin a jeho manažer se rozhoduje, jaké plodiny nakoupit:

- hrách se prodává za 10 Kč/kg a jeho nákupní cena je 8 Kč/kg
- mouka se prodává za 8Kč/kg, nákupní cena malého množství je 7Kč/kg, při každé odebrané tuně se cena snižuje o 0,02 Kč
- čočka se prodává za 12 Kč/kg a její nákupní cena je 9 Kč/kg, speciálním potřebám pro skladování této plodiny však vyhovuje pouze jedna část skladu, která je schopná pojmout 50 t.

Kolik nakoupit hrachu, kolik mouky a kolik čočky?

Nelineární optimalizace s omezeními
- příklad II

Proměnné:

x množství hrachu

y množství mouky

z množství čočky

Omezení skladovací kapacity:

$x + y + z \leq 100$

$z \leq 50$

Nelineární optimalizace s omezeními
- příklad III

Zisk z jedné tuny:

hrách	2000 Kč
čočka	3000 Kč
mouka	(1000 + 20y) Kč

Poznámka: Cena za kg: 7 - 0,02y => zisk z tuny tedy bude 1000.(8 - (7 - 0,02y)) = 1000 + 20y

Účelová funkce má tedy tvar:

$$2000x + (1000 + 20y)y + 3000z =$$

$$= 2000x + 1000y + 20y^2 + 3000z$$

Nelineární optimalizace s omezeními
- definice

Množina přípustných bodů:

$S = \{\mathbf{x} \in R^n \mid c_i(\mathbf{x}) = 0,\ i \in E;\ c_i(\mathbf{x}) \geq 0,\ i \in I\}$

Přípustný bod $x^* \in E$, splňující omezující podmínky, nazveme **omezeným lokálním minimem** problému, pokud existuje okolí $\Omega(x^*)$ bodu x^* tak, že pro všechna $x \in \Omega(x^*)$, která splňují omezení, platí $f(x^*) \leq f(x)$.

Nelineární optimalizace s omezeními - definice II

Aktivní omezení v bodě $\mathbf{x}$ je indexová množina A, pro kterou platí:

$A(\mathbf{x}) = \{i; c_i(\mathbf{x}) = 0\}$

Je-li tedy $i \in A(\mathbf{x})$, pak $\mathbf{x}$ je na hranici oblasti definované i-tým omezením.

Velmi důležitá je množina $A^* = A(\mathbf{x}^*)$, jestliže je tato množina známa, pak můžeme ostatní omezení ignorovat a řešit úlohu jako problém s omezeními ve formě rovnic $E = A^*$.

Nelineární optimalizace s omezeními
- definice III

Volné extrémy:

Jsou to lokální extrémy účelové funkce, nezávisí na žádných dalších omezeních.

Vázané extrémy

Body účelové funkce, které jsou minimální (maximální) v oblasti, omezené jistými podmínkami.

Omezení v podobě rovnic
- obecně

Minimalizuj:

$f(x), \quad x \in R^n$

za předpokladu:

$c_i(x) = 0, i \in E$

$c_i(x) \geq 0, i \in I$

kde

$I = \emptyset$

Omezení v podobě rovnic
- eliminace proměnných

Jestliže jsou omezení tvořena pouze soustavou m rovnic, můžeme z nich vyjádřit m proměnných pomocí zbývajících n - m proměnných. Za těchto m proměnných pak můžeme dosadit v účelové funkci a řešit namísto původního problému o n proměnných s omezeními nepodmíněný minimalizační problém o n-m proměnných.

Omezení v podobě rovnic
- eliminace proměnných II

Pokud řešení nepodmíněného problému splňuje omezení původní úlohy, dospěli jsme k řešení.

Takto vytvořený nepodmíněný problém ale obecně nemusí být ekvivalentní původnímu problému, a proto je třeba při eliminaci proměnných zachovávat jistou opatrnost a konfrontovat nalezené řešení s původním zadáním.

Omezení v podobě rovnic

- eliminace proměnných III

Příklad:

Minimalizuj: $f(x) = -x_1 - x_2$

Za předpokladu: $x_1^2 + x_2^2 = 1$

Řešení:

$$x_2 = \pm\sqrt{1 - x_1^2}$$

x_2 dosadíme do rovnice pro f(x), odderivujeme a řešíme $f'(x) = 0$

globální minimum je řešením rovnice:

$$\min \quad -x_1 - \sqrt{1 - x_1^2}$$

Omezení v podobě rovnic

- metoda Lagrangeových činitelů

Cílem metody je najít vektory x* a λ* splňující rovnice:

$$g(x) = \sum_{i \in E} a_i(x).\lambda_i$$

$$c_i(x) = 0, \quad i \in E$$

Kde: $g(x)$ gradient $f(x)$

$a_i(x)$ $\nabla c_i(x)$

Omezení v podobě rovnic

- metoda Lagrangeových činitelů II

Pokud máme m omezení ($|E| = m$), pak dostáváme n+m rovnic o n+m neznámých.

Nicméně jde o systém nelineárních rovnic, takže nalezení řešení obecně nemusí být jednoduché. Navíc řešením mohou být kromě lokálních minim také lokální maxima a sedlové body.

Omezení v podobě rovnic

- metoda Lagrangeových činitelů III

Příklad:

Minimalizuj $x_1 + x_2$

Za předpokladu: $x_1^2 - x_2 = 0$

Sestavíme rovnice:

$$\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2x_1 \\ -1 \end{pmatrix} \lambda$$

$$x_1^2 - x_2 = 0$$

Odtud dostáváme:

$$\lambda^* = -1,\ x_1^* = -1/2 \text{ a } x_2^* = 1/4$$

Omezení v podobě nerovnic

- obecně

Nyní opět rozšíříme své úvahy na obecnou nelineární optimalizační úlohu:

Minimalizuj:

$f(x), \quad x \in R^n$

za předpokladu:

$c_i(x) = 0, i \in E$

$c_i(x) \geq 0, i \in I$

Omezení v podobě nerovnic

- obecně II

Nechť opět x* je řešením úlohy. V okolí bodu x* jsou pak důležité jen ty nerovnice, které představují aktivní omezení v x*. Označíme je jako I*, kde $I^* = A^* \cap I$.

Pokud x* je lokálním minimem, pak musí splňovat následující podmínky:

$$g(x^*) = \sum_{i \in A^*} a_i(x^*).\lambda_i^*$$

a zároveň $\lambda_i^* \geq 0,\ i \in I^*$

Omezení v podobě nerovnic
- příklad

Ukázka použití podmínek z minulého slidu pro řešení příkladu:

Minimalizuj: $f(x) = - x_1 - x_2$

Za předpokladů: $c_1(x) = x_2 - x_1^2 \geq 0$

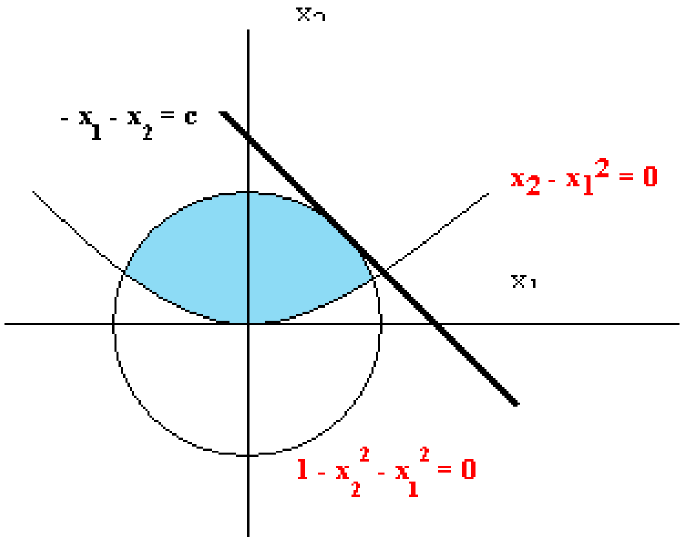
$c_2(x) = 1 - x_1^2 - x_2^2 \geq 0$

Nemáme žádný obecný návod, jak zjistit, které z omezení je aktivní v x*.

V tomto dvourozměrném případě lze nalézt aktivní omezení pomocí obrázku.

Omezení v podobě nerovnic
– příklad II

Nalezení aktivního omezení pomocí obrázku:



=> Aktivní omezení je pouze omezení 2.

Omezení v podobě nerovnic
– příklad III

Rovnice:

$$g(x^*) = \sum_{i \in A^*} a_i(x^*).\lambda_i^*$$

kde: $\lambda_i^* \geq 0,\ i \in I^*$

Tedy budou pro náš příklad vypadat následovně:

$$-1 = -2\lambda_2 x_1$$

$$-1 = -2\lambda_2 x_2$$

$$1 - x_1^2 - x_2^2 = 0$$

Uvedená soustava má dvě řešení, ale jen u jednoho z nich je $\lambda_2 \geq 0$: $\lambda_2 = x_1 = x_2 = \frac{1}{\sqrt{2}}$

Omezení v podobě nerovnic
– příklad IV

Při hledání aktivního omezení můžeme také postupovat tak, že systematicky otestujeme všechny kombinace daných omezení.

V našem případě by to tedy byly tyto kombinace:

a) $A^* = \emptyset$ Je zřejmé, že toto neplatí.

b) $A^* = \{1\}$ Neexistuje řešení, pro které by bylo $\lambda_1 \geq 0$.

c) $A^* = \{1,2\}$ Neexistuje řešení, pro které by bylo zároveň $\lambda_1 \geq 0$ a $\lambda_2 \geq 0$.

d) $A^* = \{2\}$ Tady řešení existuje :-).

Omezení v podobě nerovnic

- podmínky pro vázané extrémy

Na předchozích slidech jsme si ukázali, jakým způsobem lze odvodit nutné podmínky pro lokální minima problému. Dosažené výsledky nyní shrneme do jednoho tvrzení.

Omezení v podobě nerovnic

- podmínky pro vázané extrémy II

Předtím zavedeme následující termíny:

Podmínky regularity:

- vektory a_i*jsou lineárně nezávislé, nebo
- všechna omezení jsou lineární

Lagrangeova funkci (lagrangeián):

$$L(x,\lambda) = f(x) - \sum_i \lambda_i . c_i(x)$$

Omezení v podobě nerovnic
- podmínky pro vázané extrémy III

A nyní nutné podmínky pro lokální minimum problému (Kuhn-Tuckerovy podmínky).

Je-li x* lokálním minimem problému a pokud v x* platí podmínky regularity, pak existují Lagrangeovy činitele λ^* tak, že x* a λ^* splňují následující soustavu omezení:

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

$c_i(x) = 0, \qquad i \in E$

$c_i(x) \geq 0, \qquad i \in I$

$\lambda_i \geq 0, \qquad i \in I$

$\lambda_i \, c_i(x) = 0, \qquad \forall i$

Omezení v podobě nerovnic
- podmínky pro vázané extrémy IV

Dále zavedeme dostatečné podmínky pro lokální minimum problému: Nechť pro x* existuje λ^* tak, že platí Kuhn-Tuckerovy podmínky a platí:

$s^T.\nabla_x^2 L(x^*,\lambda^*).s > 0$

pro všechna $s \neq 0$ taková, že

$s^T.a_i^* = 0$ pro $i \in E \cup I_+^*$

$s^T.a_i^* \geq 0$ pro $i \in I_0^*$

kde

$I_+^* = \{i; i \in I^*, \lambda_i^* > 0\}$

$I_0^* = \{i; i \in I^*, \lambda_i^* = 0\}$

Pak x* je ostré lokální minimum úlohy.

Omezení v podobě nerovnic
- řešení problému

Využití klasických optimalizačních metod:

Simplexová metoda

Metody první derivace (spádové metody, konjugované gradienty)

Metody druhé derivace (Newtonovské metody atd.)

+ respektování omezení při výpočtu kroku $\delta^{(k)}$

+ testování podmínek, uvedených na předchozích slidech

Speciální případy NP
- obecně

Kvadratické programování:

kvadratická účelová funkce, lineární omezení

Konvexní programování:

účelová funkce i omezení jsou konvexní funkce

Separabilní programování:

bez smíšených členů $(2.x_1.x_2)$

Lomené programování:

účelová funkce a omezení jsou podíly lineárních funkcí

Celočíselné programování:

všechny nebo některé proměnné musí být celá čísla

Kvadratické programování

- obecně

Přirozeným zobecněním lineárního programování je úloha kvdratického programování:

Minimalizuj:

$$q(x) = \tfrac{1}{2}.x^T G x + g^T x$$

Za předpokladů:

$$a_i^T x = b_i,\ i \in E$$

$$a_i^T x \geq b_i,\ i \in I$$

Kvadratické programování

- obecně II

Tato úloha má význam nejen sama o sobě, ale také proto, že některé metody pro nelineární programování s omezeními převádějí obecný problém na posloupnost kvadratických problémů.

Pokud je Hessova matice G kladně semidefinitní, pak je funkce q konvexní a problém je speciálním případem úlohy konvexního programování a tedy má je jediné minimum.

Kvadratické programování
- obecně III

Naopak úloha

$\min \{-x^T x; -1 \le x_i \le 1, i = 1, \ldots, n\}$

má lokální minimum ve všech vrcholech intervalu $[-1, 1]^n$, tj. celkem 2^n lokálních minim.

Lze dokázat, že úloha kvadratického programování je NP-těžká (vzhledem k n).

Kvadratické programování

- metody řešení

Omezení v podobě rovnic:

- Přímá eliminace proměnných
- Metoda nulového prostoru

Omezení v podobě nerovnic (obecné):

- Metoda aktivní množiny

Celočíselné programování

- obecně

Řeší úlohy typu:

minimalizuj:

$f(x)$

za předpokladu:

$c_i(x) = 0, i \in E$

$c_i(x) \geq 0, i \in I$

přičemž pro všechny složky vektoru řešení x platí: $x_i \in Z$

Celočíselné programování

- příklad problému

Problém batohu:

Mějme n předmětů, pro něž jsou zadány veličiny:

a_j hmotnost j-tého předmětu

c_j cena j-tého předmětu

Batoh je třeba naplnit tak, abychom maximalizovali celkovou cenu nákladu a nepřekročíli přitom limit b jeho hmotnosti.

Celočíselné programování

- příklad problému II

Zavedeme proměnné x_j ($j = 1, 2, ..., n$):

$x_j = 1$ jestli naložíme j-tý předmět

$x_j = 0$ jinak

Matematická formulace úlohy:

$$\max f(x) = \sum_{j=1}^{n} c_j x_j$$

Za předpokladu:

$$\sum_{j=1}^{n} a_j x_j \leq b \qquad x_j \in \{0, 1\};\ j = 1, 2, \ldots, n$$

Celočíselné programování

- metody řešení

Metody řezných (sečných) nadrovin

Gomoryho algoritmus

Metoda větví a mezí

Speciální metody (maďarská metoda atd.)

Celočíselné programování
- metoda větví a mezí

V současné době nejpoužívanější metoda celočíselného programování.

Metoda je dostatečně obecná, takže ji lze použít na mnoho úloh.

Využívá se pro řešení celočíselných úloh LP i v rámci profesionálních programových systémů.

Celočíselné programování

- metoda větví a mezí II

Předpokládejme, že jsme nalezli optimální řešení, jehož některé (nebo všechny) složky x_i nesplňují podmínku celočíselnosti. Postupujeme následovně:

1)Vybereme takové $\mathbf{x_i}$ z optimálního řešení, které není celočíselné. (Při více takových x_i nezáleží na tom, které vybereme.)

Zvolíme interval: $\mathbf{p} \leq \mathbf{x_i} \leq \mathbf{q}$,

kde **p** je nejbližší menší celé číslo a **q** nejbližší větší celé číslo vzhledem k x_i.

Vzhledem k tomu, že celočíselné řešení není uvnitř intervalu (p,q), hledáme ho vně.

Celočíselné programování

- metoda větví a mezí III

2) Přidáme k původní úloze podmínku ve tvaru $\mathbf{x_i} \le \mathbf{p}$, tím vytvoříme **úlohu 2a**, a najdeme optimální řešení (nemusí být celočíselné).

Poté přidáme k původní úloze (tzn. úloze bez $\mathbf{x_i} \le \mathbf{p}$) podmínku ve tvaru $\mathbf{x_i} \ge \mathbf{q}$, tím vytvoříme **úlohu 2b**, a najdeme optimální řešení (nemusí být celočíselné).

Celočíselné programování

- metoda větví a mezí IV

3) Z optimálních řešení úloh 2a a 2b vybereme tu nadějnější větev (s vyšší hodnotou účelové funkce). V případě, že optimální řešení nadějnější větve ještě není celočíselné, pokračujeme stejně jako v bodě 2.

Opakujeme body 2 a 3, dokud nedospějeme k *celočíselnému optimálnímu řešení*.

Celočíselné programování

- metoda větví a mezí - příklad

Minimalizuj:

$f(x) = 2x_1 + 3x_2$

Za předpokladu:

$4x_1 + 3x_2 \leq 13$

$x_1 + 2x_2 \leq 6$

$x_1, x_2 \geq 0$

x_1, x_2 - celé

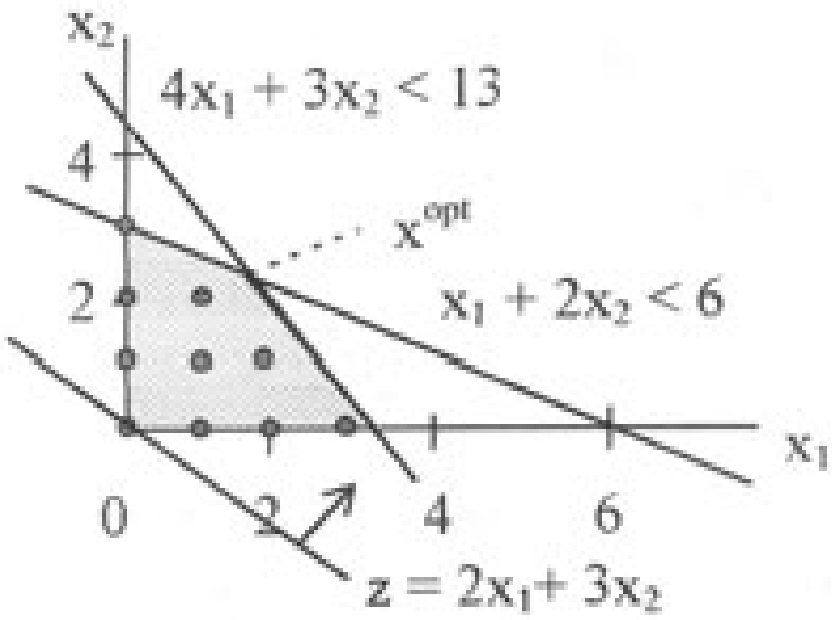
Celočíselné programování

- metoda větví a mezí – příklad II

1) Nalezení optimálního řešení (x) v R^2.

Graficky, výsledek:

$$x = (8/5,\ 11/5),\ f(x) = 49/5$$



Celočíselné programování

- metoda větví a mezí – příklad III

2) Vybereme x_1 a interval $\mathbf{p \leq x_1 \leq q}$,

tedy je: $\mathbf{1 \leq (x_1 = 8/5) \leq 2}$

Úloha 2a – 1. větev:

Přidáme k zadání podmínku: $\mathbf{2 \leq x_1}$

Řešení: $\mathbf{x' = (2, 5/3), f(x')= 9}$

Úloha 2b – 2. větev:

Přidáme k zadání podmínku: $\mathbf{x_1 \leq 1}$

Řešení: $\mathbf{x' = (1, 5/2), f(x')= 19/2}$

Celočíselné programování

- metoda větví a mezí – příklad IV

3) Nadějnější větev je větev 2 => **x = (1, 5/2)**

2) Vybereme x_2 a interval $\mathbf{p \leq x_2 \leq q}$,

tedy je: $\mathbf{2 \leq (x_2 = 5/2) \leq 3}$

Úloha 2a – 1. větev:

Přidáme k zadání podmínku: $\mathbf{x_2 \leq 2}$

Řešení: $\mathbf{x' = (1, 2), f(x') = 8}$ **STOP**

Úloha 2b – 2. větev:

Přidáme k zadání podmínku: $\mathbf{3 \leq x_2}$

Řešení: $\mathbf{x' = (0, 3), f(x') = 9}$ **STOP**

=> Celočíselným optimem je bod **x = (0,3)**

Celočíselné programování

- metoda větví a mezí – příklad V

Znázornění výpočtu pomocí stromu:

$x = (8/5,\ 11/5),\ f(x) = 49/5$

$x_1,\ x_1 \geq 2$

$x = (2,\ 5/3),\ f(x) = 9$

$x_1,\ x_1 \leq 1$

$x = (1,\ 5/2),\ f(x) = 19/2$

$x_2,\ x_2 \leq 2$

$x = (1,\ 2),\ f(x) = 8$

$x_2,\ x_2 \geq 3$

$x = (0,\ 3),\ f(x) = 9$

Cvičení

- metoda Lagrangeových činitelů

Příklad:

Minimalizuj $-x_1^2 - 4x_2^2$

Za předpokladu: $x_1 + 2x_2 = 6$

Sestavíme rovnice:

$$\begin{pmatrix} -2x_1 \\ 8x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \lambda$$

$$x_1 + 2x_2 = 6$$

Odtud dostáváme:

$$\lambda^* = -6,\ x_1^* = 3 \text{ a } x_2^* = 3/2$$

Cvičení

- celočíselné programování

Maximalizuj:

$f(x) = 2x_1 + x_2$

Za předpokladu:

$x_1 + x_2 \leq 5$

$-x_1 + 2x_2 \leq 0$

$6x_1 + 2x_2 \leq 21$

$x_1, x_2 \geq 0$

x_1, x_2 - celé