Nelineárne optimalizačné modely a metódy

Téma prednášky č. 6

Klasifikácia a všeobecná charakteristika metód pre riešenie úloh nelineárneho programovania

Prof. Ing. Michal Fendek, CSc.

Katedra operačného výskumu a ekonometrie
Ekonomická univerzita
Dolnozemská 1
852 35 Bratislava

Pojem algoritmu

Pojem algoritmu je jedným zo základných pojmov aj v matematickom programovaní. Pod algoritmom rozumieme *súbor presne definovaných pravidiel určujúcich poradie vykonania istého konečného systému operácií, ktorý zabezpečuje riešenie všetkých úloh daného typu*.

Skúmajme úlohu nelineárneho programovania

$$f(\mathbf{x}) \to \min$$

pri ohraničeniach (6.1)

$$\mathbf{x} \in D$$

kde

f($\mathbf{x}$) je účelová funkcia,

D je množina prípustných riešení úlohy.

Algoritmus pre riešenie úlohy (6.1) možno formulovať v tvare iteračnej schémy, ktorá generuje postupnosť bodov $\mathbf{x}^k$ pre $k \in \{1,2,...\}$ v súlade s definovaným súborom pravidiel, ktorého súčasťou je aj pravidlo pre ukončenie iteračného procesu.

Najvýznamnejšou vlastnosťou algoritmov pre riešenie úloh triedy (6.1) je garancia konvergencie nimi generovanej postupnosti bodov x^k ku globálnemu optimálnemu riešeniu úlohy. Toto je *ideálna predstava o algoritmoch*.

Vo väčšine prípadov však musíme byť skromnejší a uspokojiť sa s podstatne menej reprezentatívnymi výsledkami. V praktických situáciách si nekonvexnosť funkcií, veľké rozmery úlohy a iné ťažkosti vynucujú ukončiť procedúru riešenia už vtedy, ak sme získali bod, ktorý patrí do nejakej nami akceptovateľnej množiny ψ, ktorú budeme nazývať *množinou riešení* a ktorá má určité postačujúce vlastnosti. Uveďme niekoľko typických množín riešení úlohy (6.1):

Klasifikácia metód NP

1. $\psi = \{ x^o \mid x^o$ je bodom lokálneho optima úlohy $\}$,

2. $\psi = \{ x^o \mid x^o \in D, f(x^o) < b \}$, kde b je nejaká ešte prijateľná hodnota účelovej funkcie,

3. $\psi = \{ x^o \mid'' x^o \in D, f(x^o) < LB + \varepsilon \}$, kde LB je dolná hranica hodnôt účelovej funkcie a e>0 je prípustná chyba,

4. $\psi = \{ x^o \mid x^o \in D, f(x^o) - f(x^*) < \varepsilon \}$, kde $f(x^*)$ je globálne minimum a $\varepsilon > 0$ prípustná chyba,

5. $\psi = \{ x^o \mid x^o$ spĺňa podmienky optimálnosti *Kuhna-Tuckera* $\}$,

6. $\psi = \{ x^o \mid x^o$ spĺňa podmienky optimálnosti *F. Johna* $\}$.

6.2 Klasifikácia metód pre riešenie úloh nelineárneho programovania

6.2.1 Metódy pre riešenie úloh na voľný extrém

Úloha voľnej optimalizácie spočíva v hľadaní minima alebo maxima funkcie bez akýchkoľvek vedľajších podmienok, resp. ohraničení. Bez ohľadu na to, že väčšina praktických optimalizačných úloh obsahuje ohraničujúce podmienky, skúmanie metód voľnej optimalizácie si zaslúži našu pozornosť, a to z viacerých dôvodov.

Niektoré algoritmy pre riešenie úloh na viazaný extrém totiž predpokladajú transformáciu úlohy na viazaný extrém na úlohu na voľný extrém, či už je to klasická *metóda Lagrangeových mukltiplikátorov*, alebo metódy založené na myšlienke *barierových*, resp. *penalizačných funkcií*.

Iná trieda metód je založená na princípe hľadania vhodného smeru presunu od aktuálneho prípustného bodu k nasledujúcemu a v druhej etape na určení optimálnej dĺžky kroku posunu v tomto smere, čo sa realizuje na báze riešenia pomocnej úlohy lineárneho hľadania. Táto úloha je ekvivalentná s minimalizáciou funkcie jednej premenej bez vedľajších podmienok, resp. s triviálnymi podmienkami ohraničení premenných.

a) metódy minimalizácie funkcie jednej premennej bez využívania derivácií

Tieto metódy umožňujú nájsť minimum kvázikonvexnej funkcie jednej premenej, pričom predpoklad kvázikonvexnosti je v niektorých prípadoch nahradený predpokladom unimodálnosti funkcie. Metódy využívajú len výpočty hodnôt funkcie a hľadajú jej minimum na uzatvorenom a ohraničenom intervale hodnôt premennej. Do tejto triedy metód napríklad patria:

- ***metóda dichotomického hľadania,***
- ***metóda zlatého rezu,***
- ***Fibonacciho metóda.***

b) metódy minimalizácie funkcie jednej premennej s využitím derivácií

Tieto metódy hľadajú minimum funkcie jednej premennej pri platnosti určitých predpokladov o jej diferencovateľnosti. Patria sem napríklad:

- *metóda polovinného delenia*, ktorá hľadá minimum pseudokonvexnej, a teda diferencovateľnej funkcie jednej premennej na uzavretom intervale,

- *Newtonova metóda*, ktorá hľadá minimum dvakrát diferencovateľnej funkcie jednej premennej pri zadanej východiskovej aproximácii minima.

Metódy minimalzácie funkcie *n* premenných bez využitia derivácií

Tieto metódy hľadajú minimum funkcie *n* premenných, t.j. riešia úlohy

$$min\ \{\ f(x)/x \in R^n\ \}.$$

Jeden z možných prístupov pre riešenie tejto úlohy je nasledovný. Pre zadaný východiskový bod $\mathbf{x}^1$ sa definuje smerový vektor **d** a potom, vychádzajúc z bodu $\mathbf{x}^1$ sa hľadá minimum funkcie v smere vektora **d.** Rieši sa teda úloha lineárneho hľadania

$$min\ \{\ f(x + \lambda d)\ /\ \lambda \in L\ \},$$

kde *L* je definovaný interval neurčitosti. Na riešenie tejto úlohy potom použijeme niektorú z hore uvedených metód lineárneho hľadania. Do tejto triedy metód patria:

- *metóda cyklickej posúradnicovej optimalizácie,*
- *Rosenbrockova metóda,*
- *metóda Hooka a Jeevensa*.

Metódy minimalzácie funkcie *n* premenných s využitím derivácií

Tieto metódy hľadajú minimum funkcie n premenných pri platnosti určitých predpokladov o diferencovateľnosti tejto funkcie. Medzi najznámejšie procedúry triedy metód minimalizácie diferencovateľnej fukcie n premenných patria:ň

- *gradientná metóda*, ktorá predstavuje základnú procedúru pre riešenie úloh tejto triedy,

- *Newtonova metóda* minimalizácie dvakrát diferencovateľnej funkcie.

Metódy pre riešenie úloh na viazaný extrém

V tejto časti uvádzame klasifikáciu a základné charakteristiky metód pre riešenie niektorých tried úloh nelineárneho programovania v tvare

$$min\ \{ f(x)\ /x \in D \subset R^n \}$$

Práve úlohy na viazaný extrém predstavujú metodologický aparát pre riešenie väčšiny aplikačných ekonomických optimalizačných úloh, ktorých typové príklady boli uvedené v 1. kapitole.

Klasifikácia metód NP

Metódy využívajúce transformáciu úlohy na viazaný extrém na úlohu na voľný extrém

Metódy tejto triedy využívajú transformáciu východiskovej úlohy nelineárneho programovania s ohraničeniami v tvare nerovníc, resp. rovníc na ekvivalentnú úlohu voľnej optimalizácie, alebo úlohu s jednoduchými ohraničeniami premenných. Táto úloha sa potom rieši niektorou s metód charakterizovaných v časti 6.2.1. Medzi metódy tejto triedy patria:

- *metóda Lagrangeových multipikátorov*, ktorá patrí medzi klasické metódy, jej základný variant je určený pre riešenie úloh s ohraničeniami v tvare rovníc,

- *metóda penalizačných funkcií*, v ktorej sa k účelovej funkcii pôvodnej úlohy pridáva funkcia, ktorá je interpretovaná ako pokuta, resp. *penále* za porušenie ktoréhokoľvek ohraničenia úlohy,

- *metóda barierových funkcií*, v ktorej sa k účelovej funkcii pôvodnej úlohy pridáva tzv. *barierový člen*, ktorý garantuje, že algoritmom generované body na jednotlivých iteráciách neprekročia hranice množiny prípustných riešení.

Klasifikácia metód NP

Metódy prípustných smerov

Táto trieda metód riešenia úloh nelineárneho programovania s ohraničujúcimi podmienkami je založená na postupnom presune od aktuálneho prípustného bodu k nasledujúcemu prípustnému bodu tak, že sa zlepší hodnota účelovej funkcie. Typická stratégia uplatňovaná v algoritmoch prípustných smerov je nasledovná:

Fáza I. Pre daný prípustný bod $\mathbf{x}^{\mathbf{k}}$ sa nájde taký *smerový vektor* $\mathbf{d}^{\mathbf{k}}$, že pre dostatočne malé $\lambda > 0$ sú splnené nasledovné podmienky:

a) bod $\mathbf{x}^{\mathbf{k}} + \lambda\mathbf{d}^{\mathbf{k}}$ je prípustným riešením úlohy,

b) hodnota účelovej funkcie v bode $\mathbf{x}^{\mathbf{k}} + \lambda\mathbf{d}^{\mathbf{k}}$ je *"lepšia"* ako hodnota účelovej funkcie v bode $\mathbf{x}^{\mathbf{k}}$.

Fáza II. Po nájdení takéhoto smerového vektora $\mathbf{d}^{\mathbf{k}}$ sa rieši úloha jednorozmernej minimalizácie, ktorej riešením sa určí optimálna dĺžka kroku λ^k v smere $\mathbf{d}^{\mathbf{k}}$. Výsledkom je nový bod $\mathbf{x}^{\mathbf{k+1}}$. Proces sa opakuje.

Do tejto triedy metód patria:

- *Zoutendijkova metóda,* v ktorej sa prípustný progresívny smer konštruuje prostredníctvom riešenia pomocnej úlohy, ktorou je obvykle úloha lineárneho programovania,

- *Rosenova metóda projekcie gradientu*, ktorá konštruuje progresívny smer pre úlohy s lineárnymi ohraničeniami takou projekciou gradientu účelovej funkcie, ktorá zabezpečuje zlepšenie hodnoty účelovej funkcie pri zachovaní prípustnosti jednotlivých bodov trajektórie,

Kvadratické programovanie

Úlohy kvadratického programovania predstavujú špeciálnu triedu úloh nelineárneho programovania, v ktorých je účelová funkcia kvadratická a sústava ohraničení je lineárna. Úloha kvadratického programoania je úlohou konvexného programovania a táto vlastnosť umožňuje nájsť jej riešenie prostredníctvom analýzy podmienok optimálnosti Kuhna-Tuckera.

Existuje celá trieda algoritmov pre riešenie úloh kvadratického programovania:

- ***Hildrethova matóda,***
- ***Wolfeho metóda,***
- ***Shettyho-Lemkeho metóda.***

Zlomkové programovanie

Úlohy zlomkového programovania predstavujú špeciálnu triedu úloh nelineárneho programovania, ktorých sústava ohraničení je tvorená lineárnymi funkciami a účelová funkcia je podielom dvoch lineárnych funkcií. Skutočnosť, že sústava ohraničení úlohy je lineárna spolu s vlastnosťami účelovej funkcie umožňujú pre riešenie úlohy využiť algoritmy založené na postupnej analýze krajných bodov množiny prípustných riešení.

Z metód pre riešenie úloh zlomkového programovania uvedieme nasledovné:

- *Metóda Gilmora a Gomoryho,* ktorá pri vyhodnocovaní kritéria optimálnosti a pri analýze extremálnych bodov úlohy využíva gradient účelovej funkcie,

- *Metóda Charnesa a Coopera*, ktorá využíva vlastnosti zlomkovej účelovej funkcie úlohy a na základe určitej substitúcie transformuje východiskovú úlohu nelineárneho programovania na ekvivalentnú úlohu lineárneho programovania, ktorá sa rieše štandardným algoritmom simplexovej metódy.

Separovateľné programovanie

Pod úlohou separovateľného programovania rozumieme takú úlohu nelineárneho programovania, v ktorej ako účelová funkcia, tak i funkcie sústavy ohraničení sa dajú vyjadriť v tvare súčtu funkcií jednej premennej. Funkcie s touto vlastnosťou nazývame *aditívne funkcie* a riešenie úloh tohto typu je založené na princípe, že každá z nelineárnych funkcií figurujúcich v úlohe je aproximivaná po častiach lineárnou funkciou a riešenie takejto aproximujúcej úlohy sa použije určitá modifikácia algoritmu simplexovej metódy.

Do tejto triedy metód patria:

- ***metóda aproximácie nelineárnych funkcií úlohy funkciami lineárnym po častiach,***

- ***Hadleyho metóda aproximácie.***