KYBERNETIKA ČÍSLO 3, ROČNÍK 4/1968

Optimalizácia n-rozmernej spojitej funkcie

FRIDRICH SLOBODA

V článku je popísaná nová metóda optimalizácie n-rozmernej spojitej funkcie, ktorá v danej oblasti nadobúda práve jeden extrém. Algoritmus predstavuje sekvenčný proces optimalizácie funkcií jedného parametru.

1. ÚVOD

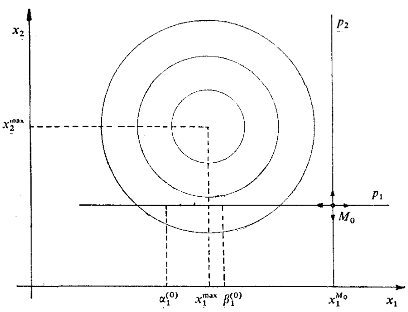
Nech $F(x_1, \ldots, x_n)$ je stacionárna spojitá funkcia n-premenných, ktorá v danej oblasti D_n nadobúda práve jeden extrém. Predpokladajme, že analytické vyjadrenie tejto funkcie nepoznáme, a máme možnosť iba zisťovať jej hodnoty v danej oblasti. Úlohou je nájsť extrém funkcie $F(x_1, \ldots, x_n)$ v danej oblasti D_n. Popíšeme metódu, ktorá nepoužíva skúšobné kroky (až na počiatočné, čo nie je nutnosť), a má premenlivý krok.

2. POPIS ALGORITMU

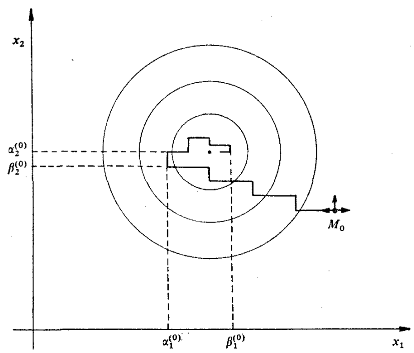
Obmedzíme sa na hľadanie maxima funkcie $F(x_1, \ldots, x_n)$. Nech M_0 je ľubovoľný bod z danej množiny D_n, $M_0 \in D_n$. Pre názornosť uvažujeme dvojrozmernú funkciu $F(x_1, x_2)$, ktorá má v oblasti D_2 určenej úsečkami rovnobežnými s osami vrstevnice, ktoré tvoria sústredné kružnice, obr. 1. V bode M_0 sa z dvoch možných smerov pre každý parameter x_i, $i = 1, \ldots, n$, vyberie ten, ktorý pre daný skúšobný krok dáva väčší prírastok funkcie. Veďme rovnobežky p_1, p_2 v bode M_0 s osami x_1, x_2 a zisťujme hodnoty funkcie $F(x_1, x_2)$ na zvolenej priamke p_1, keď parameter x_1 sa mení daným krokom Δx_1. Označme $\max F(x_1, x_2) = \max F(x) = F(x_1^{\max}, x_2^{\max})$.

Postupnosť hodnôt funkcie $F(x_1, x_2)$ je rastúca pre $x_1 \in \langle x_1^{M_0}, x_1^{\max} \rangle$ v danom smere na priamke p_1 i na každej priamke rovnobežnej s p_1. Pretože funkcia $F(x_1, x_2)$ nadobúda v danej oblasti extrém, existuje také $x_1 = \alpha_1^{(0)}$, že hodnota funkcie je už menšia ako v predchádzajúcom kroku. Ak tento postup prevedieme pre $x_1 \in$

$\in \langle \alpha_1^{(0)}, x_1^{M_0} \rangle$ krokom $\Delta x_2 < \Delta x_1$, analogicky existuje $\beta_1^{(0)}$, že hodnota funkcie $F(\beta_1^{(0)}, x_2)$ je už menšia ako hodnota funkcie v predchádzajúcom kroku. Teda sme dostali interval $\langle \alpha_1^{(0)}, \beta_1^{(0)} \rangle$ taký, že $x_1^{\max} \in \langle \alpha_1^{(0)}, \beta_1^{(0)} \rangle$.



Obr. 1.



Obr. 2.

Analogicky definujeme intervaly pre $x_i^{\max}$, ak $i = 2, 3, \ldots, n$. Ako z predchádzajúceho vidieť, kroky rovnobežné sa súradnými osami sa môžu striedať. Krok pre daný parameter sa robí na základe predchádzajúceho kroku pre daný parameter. Ak funkcia $F(x_1, \ldots, x_n)$ je funkciou n-premenných, potom nasledujúci krok pre daný parameter sa prevedie po $n - 1$ krokoch pre ostatné parametre. Ak v predchádzajúcom kroku pre daný parameter x_i, $i = 1, 2, \ldots, n$, bol prírastok funkcie kladný, resp. nulový, potom nasledujúci krok pre daný parameter x_i $i = 1, 2, \ldots, n$, bude v tom istom smere a tak veľký ako v predchádzajúcom kroku pre daný parameter.

Ak v predchádzajúcom kroku pre daný parameter bol prírastok funkcie záporný, v nasledujúcom kroku pre daný parameter x_i, $i = 1, 2, \ldots, n$, bude krok opačným smerok ako v predchádzajúcom kroku pre daný parameter a hodnota kroku bude menšia. Takýmto spôsobom dostaneme pre funkciu znázornenú na obr. 1. pre každý parameter x_1, x_2, obecne pre x_i, $i = 1, 2, \ldots, n$, systém do seba zapadajúcich intervalov $\langle_i^{(l)}, \beta_i^{(l)}\rangle$ pre $i = 1, 2, \ldots, n$ a $l = 0, 1, 2, \ldots$, v ktorých sa nachádzajú extremálne súradnice hľadaného extrému. Celkový pohyb v danej oblasti môže byť znázornený pre danú funkciu (na obr. 1) ako na obr. 2.

Je zrejmé, že pre zložitejšie funkcie hodnota kroku pre daný parameter môže nadobudnúť veľmi malú hodnotu a extrém nemusí byť dosiahnutý. Funkciu budeme nazývať *zložitú* vzhľadom na vychodzí bod M_0, ak sa aspoň jedna hodnota $x_i^{\max}$ nenachádza v prvovytvorenom intervale $\langle\alpha_i^{(0)}, \beta_i^{(0)}\rangle$ pre $i = 1, \ldots, n$. V ďalšom uvedieme spôsob zväčšovania hodnoty kroku a rozšírime našu predchádzajúcu úvahu.

3. KONVERGENCIA

Uvažujme kartézsky pravouhlý systém a zložky vektora pohybu rovnobežné so súradnými osami tohto systému. Takto uvažovaný vektor pohybu vyhovuje pre jednoduché funkcie (viď definíciu zložitej funkcie) a pre určitú podmnožinu zložitých funkcií. Nevyhovuje pre tú triedu zložitých funkcií, ktoré predstavujú telesá, ktoré majú ostrú hranu sklonenú k súradným osiam. Aj je táto hrana zhodná s niektorým smerom súradnej osi, alebo ak odklon tejto hrany od niektorého smeru súradnej osi je malý, potom algoritmus si práceschopnosť zachováva pre ľubovolný východzí bod M_0. V prípade, ak teleso má hranu sklonenú k súradným osiam, algoritmus pre daný vektor pohybu sa ukončí nájdením niektorého bodu na hrane telesa. Ak predpokladáme, že hranu (resp. hrebeň funkcie) možno po čiastkach aproximovať priamkami resp. priamkou, potom ďalší pohyb po hrebeni bude možný ak pôvodné zložky vektora pohybu otočíme (pri zachovaní pravouhlosti) tak, aby jedna zložka vektoru pohybu bola zhodná so smerom priamky aproximujúcu hranu, alebo aby odklon od nej bol malý. V ďalšom prevedieme dôkaz konvergencie pre jednoduché funkcie a pre tú časť zložitých funkcií (v zmysle definície), u ktorých si algoritmus zachováva práceschopnosť pri danom vektore pohybu.

Nech $F(x_1, \ldots, x_n)$ je spojitá funkcia n-premenných, ktorá v danej oblasti D_n nadobúda práve jeden extrém a jej analytické vyjadrenie nepoznáme. Ukážeme konvergenciu algoritmom získanej postupnosti funkčných hodnôt k extrému funkcie z udanej triedy, podľa nižše uvedeného algoritmu.

Označme hodnoty počiatočných krokov daným smerom (určeným skúšobnými krokmi vo východzom bode) $\Delta x_i^{(0)}$ pre $i = 1, 2, \ldots, n$. Kroky sú rovnobežné so súradnými osami a sa striedajú. Nech $F(x_1^{(0)}, x_2^{(0)}, \ldots, x_n^{(0)})$ je hodnota funkcie vo východzom bode M_0 a nech $\{1/k^m\}$ je klesajúca postupnosť pre ktorú platí, že $\lim_{m\to\infty} 1/k^m = 0$ pre reálne $k > 1$, $m = 1, 2, 3, \ldots$,

$$x_i^{(j)} = x_i^{(j-1)} + \Delta x_i^{(j-1)} \quad \text{pre} \quad i = 1, \ldots, n;\ j = 1, 2, \ldots ;$$

$$\Delta F_i^{(j)} = F(x_1^{(j)}, x_2^{(j)}, \ldots, x_i^{(j)}, x_{i+1}^{(j-1)}, \ldots, x_{i+n}^{(j-1)}) - F(x_1^{(j)}, x_2^{(j)}, \ldots, x_{i-1}^{(j)}, x_i^{(j-1)}, \ldots, x_n^{(j-1)}),$$

$$\Delta_i^{(j+1)} = \Delta x_i^{(j)} \quad \text{ak} \quad \Delta F_i^{(j)} \geqq 0 \quad \text{pre} \quad i = 1, 2, \ldots, n;\ j = 1, 2, \ldots ;$$

$$\Delta_i^{(j+1)} = -\frac{\Delta x_i^{(j)}}{k} \quad \text{ak} \quad \Delta F_i^{(j)} < 0 \quad \text{pre} \quad i = 1, 2, \ldots, n;\ j = 1, 2, \ldots$$

Postupnosti hodnôt $x_i^{(j)}$, pre ktoré $\Delta F_i^{(j)} < 0$ budeme priradovať označenie z postupnosti $\{\alpha_i^{(0)}, \beta_i^{(0)}, \alpha_i^{(1)}, \beta_i^{(1)}, \alpha_i^{(2)}, \beta_i^{(2)}, \ldots, \alpha_i^{(l)}, \beta_i^{(l)}, \ldots\}$ pre $i = 1, 2, \ldots, n$ a $l = 0, 1, \ldots$ pričom $x_i^{(j)} = \alpha_i^{(0)}$ odpovedá prvej hodnote $x_i^{(j)}$, pre ktorú $\Delta F_i^{(j)} < 0$.

$\Delta x_i^{(j+1)} = \Delta x_i^{(j)} \cdot k^2$ ak algoritmom získaná hodnota $x_i^{(j)} \notin \langle \alpha_i^{(l)}, \beta_i^{(l)} \rangle$ (resp. ak $x_i^{(j)} \notin \langle \beta_i^{(l)}, \alpha_i^{(l)} \rangle$, čo závisí od východzieho bodu M_0, pričom hodnota funkcie pre $x_i^{(j)}$ ešte nebola zistená) pričom interval $\langle \alpha_i^{(l)}, \beta_i^{(l)} \rangle$ sa ďalej nebude uvažovať pre pôvodné hodnoty $x_i^{(j)}$, pre ktoré $\Delta F_i^{(j)} < 0$, ale pre nové hodnoty $x_i^{(j)}$ získané novou hodnotou kroku, pre ktoré $\Delta F_i^{(j)} < 0$. Nová hodnota kroku odpovedá hodnote kroku v najmenšom nadintervale pôvodného intervalu $\langle \alpha_i^{(l)}, \beta_i^{(l)} \zeta$, tj. v intervale $\langle \alpha_i^{(l-1)}, \beta_i^{(l-1)} \rangle$ daným smerom.

Tvrdenie 1. Ak $x_i^{\max} \in \langle \alpha_i^{(0)}, \beta_i^{(0)} \rangle$, potom horeuvedeným algoritmom sa pre danú presnosť konečným počtom krokov vytvorí systém do seba zapadajúcich intervalov $\langle \alpha_i^{(l)}, \beta_i^{(l)} \rangle$ pre každé $i = 1, 2, \ldots, n$ pričom $l = 0, 1, 2, \ldots$ taký, že

$$x_i^{\max} \in \langle \alpha_i^{(l)}, \beta_i^{(l)} \rangle$$

a

$$\lim_{l \to \infty} \alpha_i^{(l)} = \lim_{l \to \infty} \beta_i^{(l)} = x_i^{\max} .$$

Pretože interval $\langle \alpha_i^{(0)}, \beta_i^{(0)} \rangle$ je konečný, na konečný počet zväčšení hodnoty kroku, sa vytvorí systém do seba zapadajúcich intervalov pričom pre hodnotu kroku platí, že

$$\lim_{m \to \infty} \frac{\Delta x_i^{(0)}}{k^m} = 0$$

pre $m = 1, 2, \ldots$ a reálné $k > 1$,
kde $\Delta x_i^{(0)}$ je hodnota počiatočného kroku.

Pretože funkcia $F(x_1, \ldots, x_n)$ je spojitá a v danej oblasti nadobúda práve jeden extrém, $x_i^{\max}$ pre $i = 1, 2, \ldots, n$ nám udávajú súradnice hľadaného extrému.

Ak $x_i^{\max} \notin \langle \alpha_i^{(0)}, \beta_i^{(0)} \rangle$, potom na konečný počet krokov algoritmom získaná hodnota $x_i^{(j)} \notin \langle \alpha_i^{(0)}, \beta_i^{(0)} \rangle$ (pričom hodnota funkcie pre $x_i^{(j)}$ ešte nebola zistená). Potom hodnota kroku sa určí v danom smere vzhľadom na polohu $x_i^{(j)}$ v intervale $\langle a_i, b_i \rangle$ (viď nasledujúci odsek). Novou hodnotou kroku sa vytvorí nový interval $\langle \alpha_i^{(0)}, \beta_i^{(0)} \rangle$, pričom pôvodný sa dalej neuvažuje. Na konečný počet krokov sa takto vytvorí taký interval $\langle\alpha_i^{(0)}, \beta_i^{(0)}\rangle$, že $x_i^{\max} \in \langle\alpha_i^{(0)}, \beta_i^{(0)}\rangle$ pre každé $i = 1, 2, \ldots, n$, pretože D_n je konečná oblasť, v ktorej existuje práve jeden extrém spojitej funkcie $F(x_1, \ldots, x_n)$. Z predchádzajúceho vyplýva nasledujúce tvrdenie:

Tvrdenie 2. Na konečný počet krokov sa algoritmom vytvorí taký interval $\langle\alpha_i^{(0)}, \beta_i^{(0)}\rangle$, pre ktorý $x_i^{\max} \in \langle\alpha_i^{(0)}, \beta_i^{(0)}\rangle$ pre každé $i = 1, 2, \ldots, n$, a teda podľa predchádzajúceho tvrdenia, v ktorom sa vytvorí systém do seba zapadajúcich intervalov $\langle\alpha_i^{(l)}, \beta_i^{(l)}\rangle$ $i = 1, 2, \ldots, n$, $l = 0, 1, 2, \ldots$, že

$$x_i^{\max} \in \langle\alpha_i^{(l)}, \beta_i^{(l)}\rangle\,,$$

$$\lim_{l\to\infty} \alpha_i^{(l)} = \lim_{l\to\infty} \beta_i^{(l)} = x_i^{\max}\,.$$

Poznámka. Je zrejmé, že miesto postupnosti $\{1/k^m\}$ sa môže použiť ľubovoľná klesajúca postupnosť $\{g_m\}$, pre ktorú

$$\lim_{m\to\infty} g_m = 0$$

pričom $g_m < 1$ pre $m = 1, 2, \ldots,$

4. OPTIMÁLNA VOĽBA KROKU PRE POSTUPNOSŤ $\{1/k^m\}$

Predpokladajme, že obor definície funkcie je určený úsečkami rovnobežnými s osami, $D_n \equiv \{a_1 \leqq x_1 \leqq b_1, \ldots, a_n \leqq x_n \leqq b_n\}$. Úlohou je určiť koeficient k klesajúcej postupnosti $\{1/k^m\}$ a počiatočnú hodnotu kroku tak, aby počet krokov na dosiahnutie extrému s danou presnosťou bol minimálny, za predpokladu, že každý bod má rovnakú pravdepodobnosť byť hodnotou extremálnou. Najvýhodnejšie je voliť východzí bod stred danej úsečky pre každý parameter, pretože disperzia tohto bodu vzhľadom na ostatné body intervalu je minimálna. Skúšobný krok nám udá smer prvého kroku pre každý parameter. Hodnotu počiatočného kroku pre každý parameter je najvýhodnejšie voliť tak ako v predchádzajúcom prípade vzhľadom na stred udanej poloviny úsečky. Východzí bod a hodnota počiatočného kroku sa určí pre každé x_i, $i = 1, \ldots, n$. Chceme najsť optimálnu hodnotu k klesajúcej postupnosti $\{1/k^m\}$ pre reálne $k > 1$, pre algoritmom získaný systém do seba zapadajúcich intervalov.

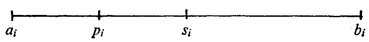
Nech $x_i \in \langle a_i, b_i\rangle$ pre $i = 1, \ldots, n$. Východzí bod označme s_i a stred udanej poloviny úsečky p_i pre $i = 1, \ldots, n$ (obr. 3), kde

$$s_i = \frac{|a_i - b_i|}{2}\,.$$

Nech k je koeficient klesajúcej postupnosti $\{1/k^m\}$. Nech hodnota východzieho kroku je $|s_i - p_i|/k$.

Potom $k + 1$ nám udáva počet krokov, ktoré treba previesť, aby nám klesla hodnota funkcie, ak predpokladaný extremálny parameter $x_i^{\max} \equiv p_i$. Hodnota kroku po $m - 1$ klesnutí hodnoty funkcie $F(x_1, \ldots, x_n)$ bude $|s_i - p_i|/k^m$, kde $m = 1, 2, \ldots$

Obr. 3.



Treba najsť minimum funkcie $G(k)$,

$$G(k) = m(k + 1) \tag{1}$$

pre reálne $k > 1$, ktorá udáva počet krokov pri podmienke, že hodnota p_i bude dosiahnutá s udanou presnosťou $c_1 < |s_i - p_i|$ tj.

$$\frac{|s_i - p_i|}{k^m} = c_1 \tag{2}$$

alebo

$$k^m = c\,, \tag{2a}$$

kde

$$c = \frac{|s_i - p_i|}{c_1}\,.$$

Po úprave rovnice (2a) dostávame

$$m \log k = \log c\,,$$

$$m = \frac{\log c}{\log k}\,. \tag{3}$$

Dosaďme (3) do (1) a najdime minimum $G(k)$ pre reálne $k > 1$:

$$G(k) = \frac{\log c}{\log k}(k + 1) = \log c\left(\frac{k}{\log k} + \frac{1}{\log k}\right), \tag{4}$$

$$\frac{\mathrm{d}G}{\mathrm{d}k} = \log c\,\frac{\log k - 1}{(\log k)^2} + \log c\left(-\frac{1/k}{(\log k)^2}\right) = \log c\left(\frac{\log k - 1 - 1/k}{(\log k)^2}\right), \tag{5}$$

$$\frac{\mathrm{d}G}{\mathrm{d}k} = 0 \quad \text{ak} \quad \begin{cases} 1.\ \log c = 0 \Rightarrow c = 1\,, \\ 2.\ \log k - 1 - 1/k = 0\,. \end{cases}$$

Riešiť budeme druhú rovnicu, pretože prvá rovnica nevyhovuje našej podmienke (2). Riešením (iteratívnym) druhej rovnice je $k = 3{,}591\,121$. Pretože druhá derivácia $\mathrm{d}^2G/\mathrm{d}k^2$ je kladná pre reálne $k > 1$, funkcia $G(k)$ pre $k = 3.591\,121$ dosahuje minimum.

Získaná hodnota k je optimálny koeficient postupnosti $\{1/k^m\}$ pre ľubovolnú presnosť, ako vidieť z rov. (5). Hodnoty počiatočných krokov sú:

$$\Delta x_i^{(0)} = \frac{|s_i - p_i|}{k},$$

kde $k = 3{,}591\,121$ pre $i = 1, \ldots, n$.

V treťom odseku spomínané určenie kroku vzhľadom na polohu $x_i^{(j)}$ v intervalu $\langle a_i, b_i \rangle$ sa prevedie tak, že posledná hodnota $x_i^{(j)}$ sa považuje za bod s_i a hodnota kroku v udanom smere sa prevedie vzhľadom na stred vzdialenosti bodu $x_i^{(j)}$ od koncového bodu intervalu $\langle a_i, b_i \rangle$.

Ak extrém funkcie sa v danej oblasti nachádza blízko hranice oblasti, alebo na hranici danej oblasti, potom sa môže stať, že alkgoritmom získaná hodnota $x_i^{(j)}$ nebude z intervalu $\langle a_i, b_i \rangle$. V tom prípade

$$x_i^{(j)} = \begin{cases} a_i \\ b_i \end{cases},$$

a nasledujúci krok je do vnútra intervalu $\langle a_i, b_i \rangle$, pričom hodnota kroku je

$$\Delta x_i^{(j)} = \frac{\Delta k_i^{(j-1)}}{k}$$

pre $i = 1, 2, \ldots, n$.

5. ZÁVER

Uvedená metóda nepoužíva skúšobné kroky, ako ich používajú gradientove metódy, má premenlivý krok a možno ju použiť i pri riešení úloh lineárneho programovania. Algoritmus tvorí sekvenčný proces optimalizácie funkcií jedného parametru. Z tohoto hľadiska, pretože teória dynamickej optimalizácie pre jednoparametrové funkcie je dostatočne rozpracovaná napr. v práci [1], možno túto metódu využiť pre dynamickú optimalizáciu viacparametrových sústav. Metódu možno tiež využiť v extremalnych adaptívnych systémoch. Rýchlosť uvedenej metódy pre jednoduchšie funkcie sa dá usúdiť z rovníc (1) a (2) v štvrtom odseku. Niektoré výsledky boli overené na systéme GIER.

(Došlo dňa 11. októbra 1967.)

[1] И. И. Перельман: Один способ самонастройки шаговых поисковых систем. Автоматика и телемеханика (1967), 4, 80—93.

SUMMARY

Optimization of n-dimensional Continuous Function

FRIDRICH SLOBODA

The article describes a new method of optimization of an n-dimensional continuous function, that has only one optimum in the given domain. The method does not use test steps (except the initial ones, which, however, are not necessary) and has variable step. The algorithm is presented as a sequential process of optimization functions with one parameter.

Further in the article there are derived the relations for the proof of convergence and for the optimal value of coefficient forstep change.

The convergence of this method is better than that of the gradient methods.

Fridrich Sloboda, prom. mat., Ústav technickej kybernetiky SAV, Dúbravská cesta 1, Bratislava.