ŘÍZENÉ NÁHODNÉ PROHLEDÁVÁNÍ S ALTERNUJÍCÍMI HEURISTIKAMI

Josef Tvrdík

Ostravská universita [1]

Abstract. *The paper deals with stochastic algorithms for global optimization over continuous space. Special attention is given to the controlled random search (CRS). A generalization of CRS is proposed where several heuristics generating a new trial point are changing at random. The condition for asymptotic convergence of the algorithm are briefly discussed. The generalized CRS algorithm has been implemented in Matlab and tested on several problems. The Matlab implementation of the CRS algorithm is available at the author of the paper.*

Klíčová slova: *globální optimalizace, stochastické algoritmy, konvergence, řízené náhodné prohledávání, heuristiky, Matlab.*

Problém globální optimalizace.

Úlohu nalezení globálního minima můžeme formulovat takto: Mějme funkci

$$f : D \to \mathcal{R}, D \subseteq \mathcal{R}^d$$

Pak

$$\mathbf{x}^{opt} = \arg\min_{\mathbf{x} \in D} f(\mathbf{x})$$

je globální minimum. Funkci f nazýváme cílovou funkcí (objective function). Nalézt obecné řešení takto jednoduše formulovaného problému je obtížné, zvláště když cílová funkce je multimodální, není diferencovatelná, případně má další nepříjemné vlastnosti. Nedávné teoretické výsledky indikují, že patrně neexistuje algoritmus řešící obecnou úlohu globální optimalizace (tj. nalezení dostatečně přesné aproximace $\mathbf{x}^{opt}$) v polynomiálním čase [2], str. 35–62. V posledních desetiletích se s poměrným úspěchem pro hledání globálního minima takových funkcí užívají stochastické algoritmy zejména evolučního typu, viz např. [2, 3, 4, 6, 8].

V tomto příspěvku se omezíme na úlohy, kdy hledáme globální minimum ve spojité oblasti $D = \prod_{i=1}^{d} < a_i, b_i >,\ a_i < b_i,\ i = 1, 2, \ldots, d$ (tzv. box constraints) a cílovou funkci $f(\mathbf{x})$ umíme vyhodnotit s požadovanou přesností v každém bodě $\mathbf{x} \in D$. Pro úlohy řešené numericky na počítači nepředstavuje box constraints žádné podstatné omezení, neboť hodnoty a_i, b_i jsou tak jako tak omezeny datovým typem užitým pro $\mathbf{x}$ a $f(\mathbf{x})$.

[1] Tato práce byla podporována z projektu institucionálního výzkumu CEZ: J09/98:179000002.

Řízené náhodné prohledávání.

Jedním z mnoha stochastických algoritmů hledání globálního minima je řízené náhodné prohledávání (controlled random search, CRS) [12]. Tento algoritmus můžeme zapsat

> **procedure** CRS;
> **begin**
>   generuj populaci P, tj. N bodů náhodně v D,
>   **repeat**
>   najdi $\mathbf{x}_{max} \in P$ takové, že $f(\mathbf{x}_{max}) \geq f(\mathbf{x}), \mathbf{x} \in P$
>   **repeat**
> $\mathbf{y} := \mathit{heuristika}(P),\ \mathbf{y} \in D$
>   **until** $f(\mathbf{y}) \leq f(\mathbf{x}_{max})$;
> $\mathbf{x}_{max} := \mathbf{y}$;
>   **until** podmínka ukončení;
> **end** {CRS};

Funkcí *heuristika*(P) v algoritmu CRS rozumíme nějaké nedeterministické pravidlo, které z bodů populace P vytvoří bod $\mathbf{y} \in D$. Pokud tento bod $\mathbf{y}$ „zlepšuje" populaci P, nahradí v ní nejhorší bod $\mathbf{x}_{max}$. Podmínka ukončení bývá většinou formulována tak, že populace nebo její část je dostatečně homogenní co do vzdálenosti jejích bodů nebo hodnot cílové funkce. V původním algoritmu CRS Price [12] jako heuristiku užíval reflexi známou ze simplexové metody [11] aplikovanou na náhodně vybraný simplex ($d+1$ bodů) z populace P. Jsou známy i jiné varianty tohoto algoritmu [1, 7].

Z popisu algoritmu CRS je zřejmé, že v něm užitá funkce *heuristika*(P) nemusí být založena na stejných pravidlech v průběhu celého prohledávání. Střídání více heuristik bylo užito v evolučním algoritmu se soutěžícími heuristikami [15], kde jsou i navržena pravidla pro změny hodnot pravděpodobnosti výběru heuristiky v závislosti na její dosavadní úspěšnosti v průběhu vyhledávání. V tomto příspěvku se bude zabývat algoritmem užívajícím jen jednoduché náhodné střídání h heuristik, kdy pravděpodobnost výběru $i-$té heuristiky je konstantní během celého procesu a je rovna $1/h$, tj. pravděpodobnost je rozdělena rovnoměrně. Aplikací tohoto pravidla dostaneme algoritmus CRS s alternujícími heuristikami.

Algoritmus CRS je speciálním případem dosti obecné třídy evolučních algoritmů [9], pro kterou byly zkoumány podmínky teoretické konvergence. Evoluční algoritmus je konvergentní tehdy, když limita pravděpodobnosti nalezení řešení velmi blízkého hodnotě $\mathbf{x}^{opt}$ je rovna jedné. Z výsledků Mišíka et al [9] vyplývá, že konvergence algoritmu CRS závisí na vlastnostech užité heuristiky. Algoritmus CRS je konvergentní, když v každém kroku je alespoň s pravděpodobností p_i užita heuristika zaručující, že pravděpodobnost vygenerování bodu $\mathbf{y}$ z libovolné otevřené množiny $S \subset D$ je kladná a řada $\sum_{i=1}^{\infty} p_i$ je divergentní. Asymptotická konvergence algoritmu CRS s alternujícími heuristikami je tedy zaručena. když alespoň jedna ze střídajících se heuristik „umí" s kladnou pravděpodobností vygenerovat bod $y \in S$.

Výsledky experimentální testování (rovněž v [9]) však ukázaly, že teoreticky konvergentní algoritmus nemusí být nutně úspěšnější při řešení konkrétního problému globální optimalizace (ukončeném po konečném počtu kroků) než nekonvergentní. Bohužel často je horší nejen v rychlosti konvergence, ale i ve spolehlivosti nalezení řešení blízkého globálnímu minimu.

Heuristiky.

Heuristik pro střídání v algoritmu CRS se nabízí nepřeberné množství a jejich výběrem můžeme ovlivnit vlastnosti algoritmu, jak teoretickou konvergenci, tak rychlost a spolehlivost prohledávání. Před návrhem množiny alternujících heuristik pro tuto implementaci algoritmu byly provedeny poměrně rozsáhlé testovací pokusy k získání přehledu o základních vlastnostech jednotlivých heuristik připadajících v úvahu pro použití v CRS. Pro toto testování byly zvoleny tři často užívané funkce [13, 14]:

- první de Jongova funkce (vícerozměrná koule), tj. jednomodální funkce považovaná za velmi snadnou úlohu, $< a_i, b_i >=< -5.12, 5.12 >$,
- druhá de Jongova funkce (Rosenbrockovo sedlo, známá i pod názvem banánové údolí), jednomodální funkce, středně obtížná pro algoritmy globální optimalizace, $< a_i, b_i >=< -2.048, 2.048 >$,
- Ackleyho funkce, multimodální s mnoha lokálními minimy, $< a_i, b_i >=< -30, 30 >$.

Na tyto úlohy byl opakovaně aplikován algoritmus CRS, vždy užívající jednu z 50 testovaných heuristik. Pro každou úlohu a heuristiku bylo provedeno 100 nezávislých opakování. Sledovanými veličinami byly spolehlivost nalezení globálního minima vyjádřená jako relativní četnost úspěšných výsledků a rychlost konvergence měřená počtem vyhodnocení cílové funkce potřebným k dosažení podmínky ukončení prohledávání. Kromě toho byly sledovány i další charakteristiky procesu prohledávání, jejichž analýza bude předmětem jiné práce.

Užité heuristiky byly inspirovány různými postupy užívanými ve stochastických algoritmech optimalizace:

- evoluční strategie (ES), [2, 3, 6], testováno 8 variant,
- reflexe simplexu, [12, 7], testováno 12 variant,
- diferenciální evoluce (DE), [13, 14], testováno 12 variant,
- breeder genetic algorithm (BGA), [5, 10], testováno 18 variant.

Heuristiky BGA byly nejméně úspěšné a v tomto článku se jimi nebudeme dále zabývat, ostatní heuristiky stručně popíšeme.

Heuristiky založené na ES generují nový bod $\mathbf{y}$ podle následujícího pravidla

$$y_i = x_i^{best} + U, \quad i = 1, 2, \ldots, d,$$

kde x_i^{best} je $i-$tá souřadnice bodu $\mathbf{x}^{best}$, tj. bodu s nejmenší funkční hodnotou v populaci P a U je náhodná veličina, $U \sim N(0, \sigma_i^2)$. Hodnota směrodatné odchylky σ_i je stanovována z celé populace (heuristika *esbest-pop*) jako

$$\sigma_i = k(\max x_i - \min x_i) + \varepsilon,$$

operátory $\max, \min$ znamenají největší, resp. nejmenší hodnotu i-té souřadnice v aktuální populaci P, ε (v implementaci $\varepsilon = 1 \times 10^{-4}$) zabezpečuje, že hodnota σ_i je kladná, k je vstupní řídící parametr heuristiky, $k > 0$. V heuristice *esbest-2pts* je směrodatná odchylka

σ_i určována ze dvou náhodně vybraných bodů $\mathbf{r}$, $\mathbf{s}$ z populace P při vynechání nejlepšího bodu $\mathbf{x}^{best}$

$$\sigma_i = k \ \mid r_i - s_i \mid + \varepsilon,$$

k je opět vstupní parametr heuristiky, ostatní symboly mají stejný význam jako výše.

Znáhodněná reflexe simplexu [7] náhodně vybraného z populace P je popsána vztahem

$$\mathbf{y} = \mathbf{g} + Z(\mathbf{g} - \mathbf{x}),$$

kde $\mathbf{x}$ je bod simplexu s nejvyšší funkční hodnotou (v heuristice *refl-worst*) nebo náhodný bod simplexu (v heuristice *refl-rand*), $\mathbf{g}$ je těžiště zbývajících d bodů simplexu a Z je spojitá náhodná veličina rovnoměrně rozdělená na intervalu $\langle 0, \alpha)$, α je vstupní parametr heuristiky.

Z diferenciální evoluce vycházejí dva typy heuristik. Heuristika *de-rand* generuje bod $\mathbf{u}$

$$\mathbf{u} = \mathbf{r}_1 + F(\mathbf{r}_2 - \mathbf{r}_3),$$

$\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3$ jsou navzájem různé body náhodně vybrané z populace P. F je vstupní parametr heuristiky, $F > 0$. Heuristika *de-best* je založena na modifikaci bodu $\mathbf{x}^{best}$ (s nejmenší funkční hodnotou a aktuální populaci) podle následujícího pravidla

$$\mathbf{u} = \mathbf{x}^{best} + F(\mathbf{r}_1 + \mathbf{r}_2 - \mathbf{r}_3 - \mathbf{r}_4),$$

$\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \mathbf{r}_4$ jsou opět navzájem různé body náhodně vybrané z populace P (kromě $\mathbf{x}^{best}$) a F je vstupní parametr.

Nový vektor $\mathbf{y}$ vznikne „křížením“ vektoru $\mathbf{u}$ a náhodně vybraného vektoru $\mathbf{x}$ tak, že každá složka x_i je přepsána hodnotou u_i s pravděpodobností C, $C \in \langle 0, 1\rangle$ je další vstupní parametr heuristiky. Pokud žádné x_i nebylo přepsáno hodnotou u_i nebo při volbě $C = 0$, přepisuje se jedna náhodně vybraná složka vektoru $\mathbf{x}$.

Žádná z uvedených heuristik nezaručuje, že jí vygenerovaný bod $\mathbf{y}$ nutně leží v D. Pokud nastane, že $\mathbf{y} \notin D$, je aplikována tzv. perturbace [6], která překlopí souřadnici y_i neležící v $\langle a_i, b_i\rangle$ dovnitř mnohorozměrného kvádru D zrcadlově vzhledem k příslušné stěně.

Výsledky algoritmu CRS s jednou heuristikou.

Pro všechny testovací úlohy byly voleny shodné vstupní parametry algoritmu CRS, a to

- velikost populace $N = 10d$
- podmínka ukončení $f_{(N/2)} - f_{(1)} \leq 1 \times 10^{-7}$, když populace je uspořádána podle funkčních hodnot, tj. $f_{(1)} \leq f_{(2)} \leq \ldots \leq f_{(N)}$

Úspěšnost vyhledávání byla hodnocena podle toho, zda dosažená hodnota $f_{(1)}$ byla dostatečně blízká hodnotě $f(\mathbf{x}^{opt})$, která je pro uvedené testované funkce rovna 0. Vyhledávání je považováno za úspěšné, pokud $f_{(1)} < 1 \times -3$ pro Ackleyho funkci a $f_{(1)} < 1n \times -6$ pro ostatní testované funkce. Vstupní parametry jednotlivých heuristik jsou spolu s výsledky uvedeny v tab.1. Ve sloupcích *NE* v této tabulce jsou průměrné hodnoty počtu vyhodnocení cílové funkce pro splnění podmínky ukončení u pokusů, které skočily úspěšným nalezením globálního minima. Spolehlivost vyhledávání R je relativní četnost úspěšného nalezení globálního minima, vyjádřená v procentech.

TABULKA 1. Spolehlivost a rychlost konvergence algoritmu CRS s jednou heuristikou

			funkce					
			deJong1 $d=3$		Rosenbrock $d=2$		Ackley $d=2$	
heuristika	p_1 (k, F, α)	p_2 (C)	R	NE	R	NE	R	NE
esbest-2pts	1	-	100	585	100	3380	0	-
esbest-pop	0.2	-	100	457	100	2603	0	-
de-best	0.5	0.5	100	702	100	1060	93	752
de-best	0.9	0.5	100	1403	100	2586	96	1266
de-rand	0.5	0.5	96	879	39	1393	60	850
de-rand	0.9	0.5	100	1174	82	2049	75	1038
refl-rand	2	-	100	892	100	745	99	1392
refl-rand	6	-	100	2364	100	1565	99	3556
refl-worst	2	-	100	875	100	576	99	1228
refl-worst	6	-	100	2463	100	1430	98	3413

Algoritmus CRS a alternujícími heuristikami.

Algoritmus CRS byl implementován v Matlabu se střídáním 10 heuristik uvedených v tab.1. Výsledky dosažené na testovacích funkcích jsou uvedeny v tab.2. Ve sloupci sm.odch. jsou směrodatné odchylky počtu vyhodnocení cílové funkce a ve sloupci $NE1$ průměrné hodnoty počtu vyhodnocení cílové funkce potřebné pro přiblížení nejlepšího bodu v populaci ($f_{(1)}$) k hledanému globálnímu minimu. Oproti tab.1 jsou uvedeny i výsledky pro Ackleyho funkci s $d = 10$ a pro Griewangkovu funkci [13], $< a_i, b_i >=< -400, 400 >$, $i = 1, 2, \ldots, 10$. Tato multimodální funkce je považována ze obtížnou úlohu globální optimalizace. Porovnáním výsledků v tabulkách 1 a 2 vidíme, že algoritmus s alternujícími heuristikami je spolehlivější než je průměrná spolehlivost algoritmů CRS s jednou heuristikou a také hodnoty NE jsou podstatně nižší než průměr z hodnot dosažených CRS s jednou heuristikou.

TABULKA 2. Algoritmus CRS s alternujícími heuristikami

funkce	d	R	NE	sm.odch.	$NE1$
DeJong1	3	100	858	55	647
Rosenbrock	2	100	1111	115	798
Ackley	2	95	1137	106	538
Ackley	10	99	11881	526	5746
Griewangk	10	62	10131	2727	9055

Zdrojový text implementace algoritmu CRS s alternujícími heuristikami gcrs_alth.m a další odtud volané funkce (esbest_pop.m, esbest_2pts.m, de_rand.m, de_best.m, refl_worst.m, refl_rand.m, zrcad.m) jsou k dispozici u autora příspěvku a budou dostupné i na webovské stránce www.osu.cz/~tvrdik.

Závěr.

CRS s alternujícími heuristikami je jednoduchý algoritmus, který může být snadno použit k hledání globální minima v situacích, kdy lokální optimalizátory nejsou úspěšné. Výhodou tohoto algoritmu je jednoduchost užití, neboť všechny řídící parametry algoritmu jsou nastaveny implicitně. Nevýhodou je poněkud větší časová náročnost výpočtu než u lokálních optimalizátorů a samozřejmě i společná vlastnost všech stochastických algoritmů, totiž to, že o výsledku dosaženém v konečném počtu kroků nikdy nevíme s jistotou, zda je skutečně blízký globálnímu minimu.

Literatura

[1] Ali, M.M., Törn, A. and Viitanen, S., A numerical comparison of some controlled random search algorithms, *J. Global Optimization*, **11** 377–385, (1997).

[2] Bäck, T., *Evolutionary Algorithms in Theory and Practice.* New York: Oxford University Press (1996).

[3] Bäck, T. and Schwefel, H.P., An Overview of Evolutionary Algorithms for Parameter Optimization, *Evolutionary Computation*, **1**, 1–23 (1993).

[4] Goldberg, D.E., *Genetic algorithms in search, optimization, and machine learning* Addison Wesley, Reading (1989).

[5] Haslinger, J., Jedelský, D., Kozubek, T. and Tvrdík, J., Genetic and random search methods in optimal shape design problems, *J. Global Optimization* **16**, 109–131 (2000).

[6] Kvasnička, V., Pospíchal, J. and Tiňo, P.. *Evolučné algoritmy*, STU Bratislava, (2000)

[7] Křivý, I. and Tvrdík, J.: The Controlled Random Search Algorithm in Optimizing Regression Models. *Comput. Statist. and Data Anal.*, 20, 229–234, (1995).

[8] Michalewicz, Z., *Genetic algorithms + data structures = evolutionary program*, Springer-Verlag, Berlin, (1992).

[9] Mišík, L., Tvrdík, J., Křivý, I., On Convergence of a Class of Stochastic Algorithms, In: *ROBUST'2000* (eds J. Antoch a G. Dohnal), 198–209, JČMF Praha, (2001).

[10] Mühlenbein, H., Schlierkamp-Voosen, D., Predictive models for breeder genetic algorithm, I. Continuous parameter optimization, *Evolutionary Computation*, **1**, 25–49, (1993)

[11] Nelder, J.A., Mead, R., A simplex method for function minimization, *Computer J.*, **7**(1) 308–313, (1964).

[12] Price, W.L., A controlled random search procedure for global optimization. *Computer J.*, **20**(4) 367-370, (1976).

[13] Storn, R. and Price, K., Differential evolution - a Simple and Efficient Heuristic for Global Optimization over Continuous Spaces, *J. Global Optimization* **11**, 341–359, 1997.

[14] Tvrdík, J. and Křivý, I., Simple Evolutionary Heuristics for Global Optimization. *Comput. Statist. and Data Anal.* **30**, 345–352, (1999).

[15] Tvrdík, J., Křivý, I., Mišík, L., Evolutionary algorithm with competing heuristics, In: *MENDEL'2001. 6th International Conference on Soft Computing*, 58–64, VUT Brno, (2001).

Adresa pracoviště autora:

Katedra infromatiky a počítačů
Přírodovědecká fakulta
Ostravská universita
30. dubna 22, 701 03 Ostrava
e-mail: tvrdik@osu.cz