Adaptation in Differential Evolution

Radka Poláková

May, 2012

ISCAMI, Malenovice

Introduction

- Global optimization problem
- Differential evolution algorithm (DE) and its parameters
- Adaptive approaches in DE (JADE, SADE, jDE, EPSDE, CDE, ...)
- Study of present adaptive approaches $\Rightarrow$ new adaptive algorithm(s)
- New adaptive approach in DE algorithm based on CDE and jDE is proposed here and tested on CEC2005 functions

Global optimization problem

- Objective function $f : S \to \mathcal{R}$, $S \subset \mathcal{R}^D$
- $\boldsymbol{x}^*$ is called global minimum point when

$$\boldsymbol{x}^* = arg\ min_{\boldsymbol{x} \in S}\ f(\boldsymbol{x})$$

- Search space S is closed compact set

$$S = \prod_{i=1}^{D} [a_i, b_i];\ a_i < b_i \quad i = 1, 2, \ldots, D$$

DE Algorithm

generate an initial generation $P, (\boldsymbol{x}_i, i = 1, 2, \ldots, NP)$
while stopping condition not achieved **do**
    for $i := 1$ **to** NP **do**
        generate a new trial vector $\boldsymbol{y}$ using P
        if $f(\boldsymbol{y}) \leq f(\boldsymbol{x}_i)$ **then** insert $\boldsymbol{y}$ into Q
        else insert $\boldsymbol{x}_i$ into Q
        end if
    end for
    $P := Q$
end while

Mutation types used in this study

- **rand/1** (the most popular strategy)

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

  $\boldsymbol{r}_1$, $\boldsymbol{r}_2$, $\boldsymbol{r}_3$ are mutually distinct points taken randomly from population P not coinciding with the current $\boldsymbol{x}_i$; $F > 0$ is input parameter

- **randrl/1** (random localization)

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

  for $\boldsymbol{r}_1$, $\boldsymbol{r}_2$, $\boldsymbol{r}_3$ and F hold the same statements like for strategy rand, but in addition $\boldsymbol{r}_1$ is the best one among $\boldsymbol{r}_1$, $\boldsymbol{r}_2$ and $\boldsymbol{r}_3$, therefore $\boldsymbol{r}_1 = arg\ min_{i \in 1,2,3}\ f(\boldsymbol{r}_i)$

- **current-to-rand/1**

$$\boldsymbol{y} = \boldsymbol{x}_i + rand(0,1) \times (\boldsymbol{r}_1 - \boldsymbol{x}_i) + F(\boldsymbol{r}_2 - \boldsymbol{r}_3)$$

  for $\boldsymbol{r}_1$, $\boldsymbol{r}_2$, $\boldsymbol{r}_3$ and F hold the same statements like for strategy rand, this mutation generates a trial point $\boldsymbol{y}$ directly, because it includes so called arithmetic crossover

Binomial crossover

$$y_j = \begin{cases} u_j & \text{if} \quad U_j \le CR \quad \text{or} \quad j = l \\ x_{ij} & \text{if} \quad U_j > CR \quad \text{and} \quad j \ne l, \end{cases}$$

- $\forall j \in \{1, 2, \ldots, D\}$
- l – randomly chosen number from $\{1, 2, \ldots, D\}$
- U_j, $\forall j \in \{1, 2, \ldots, D\}$ – independent random variable uniformly distributed in $[0, 1]$
- CR – input parameter influencing the number of elements to be exchanged by crossover
- at least one element of $\boldsymbol{x}_i$ is changed, even if $CR = 0$

Exponential crossover

- starting position of crossover is chosen randomly from $\{1, 2, \ldots, D\}$, then L consecutive elements (counted in circular manner) are taken from the mutant $\boldsymbol{u}$
- probability of replacing the k-th element in the sequence $1, 2, \ldots, L$ ($L \leq D$) decreases exponentially with increasing k
- relation between probability of crossover, p_m and CR for exponential crossover is not linear

$$CR^D - Dp_m CR + Dp_m - 1 = 0$$

Parameters of DE and their adaptation

- Parameters of DE – size of population *NP*, type of mutation, parameter of mutation *F*, type of crossover, parameter of crossover *CR*, stopping condition
- Setting of parameters can take a lot of time when particular problem is solved by DE
- Adaptation – effective tool to avoid time-consuming setting of parameters

Competitive DE (CDE), Tvrdik 2006

- let us have H settings
- we choose among them randomly with probabilities q_h, $h \in \{1, 2, \ldots, H\}$
- h-th setting is successful if it generates a trial point $\boldsymbol{y}$ better then $\boldsymbol{x}_i$, $f(\boldsymbol{y}) \leq f(\boldsymbol{x}_i)$
- probabilities q_h, $h \in \{1, 2, \ldots, H\}$, are adjusted according to success rate of settings in preceding steps

$$q_h = \frac{n_h + n_0}{\sum_{j=1}^{H} (n_j + n_0)}$$

  n_h is current count of successes of the h-th setting, $n_0 > 1$ prevents a dramatic change in q_h by random successful use of h-th setting
- to avoid degeneration of the search process, current values q_h are reset to their starting values ($q_h = 1/H$) if any $q_h < \delta$ (δ is input parameter, $\delta > 0$)

jDE, Brest 2006

- only strategy rand/1/bin is used in this adaptive approach
- each point $\boldsymbol{x}_i$ of population has its own values of F and CR
- they are randomly initialized and survive with point
- the values of F and CR can be randomly mutated with probabilities τ_1 and τ_2 before new trial point $\boldsymbol{y}$ is computed
- if the new pair of values generates successful trial point $\boldsymbol{y}$, this new pair survives with new point, else old pair of parameters survives with original point
- each generated value of parameter F is uniformly distributed in $[F_l, F_u]$ and the value of CR is also uniformly distributed, $CR \in [0, 1]$

New adaptive algorithm comp3jDE

comp3jDE

- it is competition with 3 settings
- the first setting is DE/rand/1/bin
- the second one is DE/randrl/1/exp
- the third one is DE/current-to-rand/1
- each of these three settings does not use only one value of F and one value of CR, but parameters are for each setting adapted by mechanism of jDE

Experiments

- 25 benchmark functions – CEC2005
- size of population $NP = 60$
- other parameters, for competition $H = 3$, $n_0 = 2$, $\delta = 1/15$, for mechanism of jDE $F_l = 0.1$, $F_u = 0.9$, $\tau_1 = 0.1$, $\tau_2 = 0.1$
- dimension of test problems, $D = 30$
- stopping condition was ($nfe \geq 300000$) (according to CEC competition)
- 25 independent runs for each function and each algorithm

CDE – b6e6rl algorithm

- CDE algorithm used for comparison is b6e6rl
- b6e6rl is competition of 12 settings of DE algorithm

Mutation	F	Crossover	CR
randrl	0.5	bin	0
			0.5
			1
	0.8	exp	CR1
			CR2
			CR3

Results

F	CDE - b6e6rl mean	std	jDE mean	std	comp3jDE mean	std
F1	0.00E+00	0.0E+00	0.00E+00	0.0E+00	3.17E-09	1.6E-08
F2	1.18E-13	6.3E-14	1.11E-06	2.0E-06	2.22E-07	8.6E-07
F3	9.09E+04	5.3E+04	1.98E+05	1.1E+05	9.15E+05	4.4E+05
F4	1.11E-13	7.2E-14	4.40E-02	1.3E-01	1.70E+02	4.1E+02
F5	5.44E+02	5.4E+02	5.11E+02	4.4E+02	2.89E+03	5.8E+02
F6	3.41E-14	2.8E-14	2.35E+01	2.5E+01	4.03E+01	4.5E+01
F7	6.30E-03	7.7E-03	1.18E-02	7.8E-03	2.89E-02	3.6E-02
F8	2.10E+01	5.7E-02	2.09E+01	4.9E-02	2.09E+01	4.3E-02
F9	0.00E+00	0.0E+00	0.00E+00	0.0E+00	3.98E-02	2.0E-01
F10	6.38E+01	1.0E+01	5.54E+01	8.5E+00	5.68E+01	2.2E+01
F11	2.66E+01	2.1E+00	2.79E+01	1.6E+00	1.75E+01	2.8E+00
F12	1.48E+04	6.3E+03	8.63E+03	8.3E+03	7.71E+03	8.1E+03
F13	1.42E+00	1.2E-01	1.66E+00	1.4E-01	1.58E+00	1.6E-01
F14	1.26E+01	2.3E-01	1.30E+01	2.0E-01	1.22E+01	3.8E-01
F15	3.64E+02	1.2E+02	3.77E+02	8.0E+01	4.08E+02	8.3E+01
F16	1.32E+02	1.0E+02	7.94E+01	3.0E+01	1.22E+02	1.2E+02
F17	1.61E+02	7.1E+01	1.37E+02	3.8E+01	1.06E+02	7.9E+01
F18	9.05E+02	1.2E+00	9.04E+02	1.1E+01	9.16E+02	3.9E+01
F19	9.06E+02	1.7E+00	9.04E+02	1.1E+00	9.06E+02	4.9E+01
F20	9.05E+02	1.0E+00	9.04E+02	1.1E+00	9.24E+02	1.5E+01
F21	5.00E+02	1.2E-13	5.00E+02	4.8E-13	5.76E+02	1.7E+02
F22	8.82E+02	1.9E+01	8.75E+02	1.9E+01	9.24E+02	3.1E+01
F23	5.34E+02	3.6E-04	5.34E+02	2.8E-04	7.68E+02	2.5E+02
F24	2.00E+02	6.0E-13	2.00E+02	2.9E-14	2.00E+02	3.1E-12
F25	2.11E+02	1.1E+00	2.11E+02	7.3E-01	2.14E+02	2.5E+00

Order of algorithms

Function	CDE - b6e6rl	jDE	comp3jDE
F1	1.5	1.5	3
F2	1	3	2
F3	1	2	3
F4	1	2	3
F5	2	1	3
F6	1	2	3
F7	1	2	3
F8	3	1	2
F9	1.5	1.5	3
F10	3	1	2
F11	2	3	1
F12	3	2	1
F13	1	3	2
F14	2	3	1
F15	1	2	3
F16	3	1	2
F17	3	2	1
F18	2	1	3
F19	2	1	3
F20	2	1	3
F21	2	1	3
F22	2	1	3
F23	2	1	3
F24	2	2	2
F25	2	1	3
Sum of orders	47	42	61
Average order	1.88	1.68	2.44

Conclusion

- new adaptive algorithm based on two of the state-of-the-art adaptive algorithms is proposed
- CDE allows to use several settings of DE, each setting is strategy (mutation and crossover) with one value of F and one value of CR
- jDE adapts values of F and CR for one strategy
- values of F and CR in our new algorithm are adapted for three strategies and these DE settings (setting – strategy with values of F and CR) are competing each other

Conclusion

Using several pairs of specific values of parameters F and CR for each strategy in CDE is better way of adaptation than to adapt these values for each competing strategy by mechanism of jDE.

Thank you for your kind attention

Conclusion

Using several pairs of specific values of parameters F and CR for each strategy in CDE is better way of adaptation than to adapt these values for each competing strategy by mechanism of jDE.

Thank you for your kind attention