VISUAL GUIDE OF F AND CR PARAMETERS INFLUENCE ON DIFFERENTIAL EVOLUTION SOLUTION QUALITY

M.Červenka[*], H. Boudná[**]

Abstract: *Choice of configuration parameters of an evolutionary algorithm is not straightforward. This contribution studies impact of two Differential Evolution configuration parameters - the mutation factor F and the crossover rate CR - on the actual solution quality the optimisation process can yield. A subset of visual maps displaying this relationship for ten variants of Differential Evolution benchmarked on six standard test functions is presented.*

Keywords: Differential Evolution, evolutionary algorithms, black-box optimisation, parameter tuning.

1. Introduction

The Differential Evolution (DE) as proposed by (Storn & Price, 1997) is a powerful optimisation technique designed for global optimisation. It can be used as a general optimiser yet its qualities emerge in assignments where conventional optimisation techniques fail. DE does not explicitly rely on gradient of the optimised problem, it works with multiple agents collaborating in a direct-search manner, treating the optimisation problem as a black-box, which merely emits a measure of fitness for candidate solutions. DE then creates new candidate solutions by combining solutions according a variant-specific chain of rules from its current set of agents (*population*), accepting new solutions in case of fitness improvement.

The main advantage of DE lies with its conceptual simplicity, ease of use and low number of control parameters. However, like any other evolutionary algorithm, the success of DE is also very sensitive on setting of its control parameters. There are three DE's control parameters: (1) the population size NP, (2) the mutation factor F (a real-value factor that controls amplification of differential variations) and (3) the crossover factor CR (also a real value, controlling the crossover operation).

2. DE control parameters

One of the main problems in evolution strategies of DE is to choose controls parameters such that they stimulate good behaviour – i.e. the algorithm does not prematurely converge to a point that is not globally optimal or stagnate and has an acceptable rate of convergence towards the global optimum. Premature convergence may occur under different situations: the population has converged to local optimum of the objective function or the population has lost its diversity or the search algorithm proceeds slowly or does not proceed at all. It has been observed that DE sometimes stops proceeding towards a global optimum and stagnation occurs. Stagnation may occur under various situations: the population has not converged to a local optimum or any other point or the population is still retaining some diversity and occasionally even of new individuals enter or are injected into the population. Then the search algorithm does not proceed towards any better solution.

For successful run of DE, proper choice of control parameters is crucial. Best configuration depends on the problem in hand and limitation of computation time and accuracy ((Liu & Lampinen, 2002)). As discussed

---

[*] Ing. Miroslav Červenka, Ph.D.: Institute of Aerospace Engineering, Faculty of Mechanical Engineering, Brno University of Technology, Technická 2896/2, 616 69 Brno; CZ, cervenka@fme.vutbr.cz
[**] Ing. Hana Boudná: Department of Active Safety Systems, Airborne Instruments Ltd., Cambridge; UK

Tab. 1: Analysed variants of Differential Evolution

	DE Variant	Formula
1	DE/best/1	$v = x^G_{best,j} + F \cdot (x^G_{r_1,j} - x^G_{r_2,j})$
2	DE/rand/1	$v = x^G_{r_1,j} + F \cdot (x^G_{r_2,j} - x^G_{r_3,j})$
3	DE/rand-to-best/1	$v = x^G_{i,j} + \lambda \cdot (x^G_{best,j} - x^G_{i,j}) + F \cdot (x^G_{r_1,j} - x^G_{r_2,j})$
4	DE/best/2	$v = x^G_{best,j} + F \cdot (x^G_{r_1,j} + x^G_{r_2,j} - x^G_{r_3,j} - x^G_{r_4,j})$
5	DE/rand/2	$v = x^G_{r_5,j} + F \cdot (x^G_{r_1,j} + x^G_{r_2,j} - x^G_{r_3,j} - x^G_{r_4,j})$

in (Storn & Price, 1997), selection of DE's control parameters is not easy there is no simple rule of thumb for choosing a good ones. On the other hand, it is also reported that choosing proper control parameters for DE can be mode difficult than expected. It is important to select optimal parameters for each problem separately and carefully to avoid premature convergence or even stagnation. (Brest et al., 2006) assessed the selection of control parameters and reported that efficiency and robustness of DE are much more sensitive to setting of mutation factor F and crossover ration CR than to the value of population size NP. (Zaharie, 2007) discussed the relationship between control parameters of DE and the evolution of population variance and reported critical interval for control parameters of DE. (Teo, 2005) proposed a method of self adapting population size in addition to self adapting mutation and crossover factor. There are also other papers indicating that the self adaptation and self evolution in DE stand for an interesting issue to look onto (Teo, 2005). Some other attempts to DE control parameters adaptation are summarized by (Liu & Lampinen, 2002).

Run of DE (and evolutionary algorithms in general) is indeed a dynamic and adaptive process. Usage of rigid parameters that do not change their values is thus in contrast to this spirit. Additionally, it is intuitively obvious, and it has been empirically and theoretically demonstrated, that different values of control parameters might be optimal at different stages of the evolutionary process. Therefore (Eiben et al., 1999) claim that use of static configuration does not bring any benefits to the optimisation process.

Even though using fixed configuration parameters for DE optimisation might look obsolete, intention of our paper is to look visually on the influence of mutation factor F and crossover factor CR on overall quality of final solution.

2.1. Choice of the control parameters

In (Storn & Price, 1997), the authors claim that it is not difficult to choose NP, F and CR in order to obtain good results. According to their experience, a reasonable choice for N is between $5 \cdot Dim$ and $10 \cdot Dim$ but NP must be at least 4 to ensure that DE will have enough mutually independent vectors to work with. As for F, $F = 0.5$ is usually supposed to be a good initial choice. If the population converged prematurely, then F and/or NP should be increased. Values of F smaller than 0.4, similarly as those greater than 1, are only occasionally effective. A good first choice for CR is 0.1, but since a large CR often speeds the convergence up, trying first $CR = 0.9$ or $CR = 1.0$ is appropriate in order to see if a quick solution is possible. Eventually, (Storn & Price, 1997) advise setting up values of F and CR depending on the results of preliminary tuning in the region $0.5 \leq F \leq 1$ and $0 \leq CR \leq 1$. There are also other papers dealing with the problem of setting DE's control parameters. (Ali & Törn, 2004) suggests simple expression how to adapt the value of the scaling factor F within the search process. (Zaharie, 2007) derivescritical intervals for the control parameters. Tvrdik (2007) suggests good initial settings to be $F = 0.8$ and $CR = 0.5$. (Babu & Jehan, 2003) deal with the F factor up to value of 1.2 and in one case even suggests using $F \leq 2$.

In overall, in the literature there are no rigid guidelines one can follow when starting with DE optimisation. It is known that quality or the optimisation process strongly depends on the F and CR parameters. There are also some recommended intervals of their initial values. But the relation between F, CR and the solution quality has still not been thoroughly mapped. Therefore, a set of experiments has been designed to shed some light into this issue and to clearly visualise this relationship.

Tab. 2: Best DE configurations by optimised function

TestFunction	DE Variant	F	CR	BestCostValue
AckleyTest	DE/RAND_1/BIN	0.05	0.60	−289.4354
DeJong1	DE/BEST_2/BIN	0.60	0.35	0.0007
Griewangk	DE/BEST_2/BIN	0.70	0.30	1.0000
Rana	DE/BEST_2/BIN	1.00	0.30	−37989.5305
Rastrigin6	DE/RAND_1/BIN	0.00	0.20	−969.6812
Schwefel	DE/RAND_1/BIN	0.05	0.65	−41894.2869

3. The experiments

3.1. Experiments settings

For a thorough exploration we used the following configuration of DE: $NP = 400$, $generations = 200$, $minDiv = 0$ (every run of DE is forced to perform 200 generations). Over the solution quality plot we laid a mesh of F and CR parameters in the range $F \in < 0; 0.05; 2 >$ and $CR \in < 0; 0.05; 1 >$ resulting in 861 F-CR combinations, where every setting was evaluated 100 times. The presented experiment is a compillation if $6 \times 10 \times 861 \times 100 = 5.166.000$ DE optimisations with total of 413.280.000.000 cost function evaluations.

3.2. Experimental results and findings

In general, Differential Evolution can provide very good results when you choose good configuration parameters for optimisation of a particular function. In case you pick improper values, the outcoming result can be fairly poor. When not using a self-adaptive or self-tuning version of DE algorithm, manual setting of DE configuration parameters can become rather tricky process. In the stage of conducting preliminary trial experiments, the user can easily find a locally-optimal pair of (F,CR) producing fairly good results and without finding any better configuration, this settings will be further used for the actual optimisation.

For a DE operator it is good to have an idea, to know which pairs of (F, CR) yield useful results while avoiding combinations resulting in insufficient of even bad results. Therefore, (F, CR) pairs for combinations of several standard test functions (AckleyTest, DeJong1, Griewangk, Rana, Rastrigin6 and Schwefel) were created and selected versions of Differential Evolution as in Table 1 were used. The maps (see Figure 1) indicated regions which should be rather avoided (lighter colours) and areas which can bring good results even if the configuration parameters are not exactly at their optimal values (darker colours).

The experimental results show that the relationship between (F, CR) and solution quality strongly depends on both, the optimised function and DE variant used. As stated by (Storn & Price, 1997), there is no need to use the mutation factor $F > 1$, alghough statement 'solutions tend to be more time consuming and less reliable than if $F < 1$' is somehow arguable. One of the best solutions found was in cofiguration with $F = 1$ ($DE/BEST\_2/BIN$ on Rana's function, see Table 3.2.) and the heat map clearly indicated that region of the most suitable solutions extends just behind $F \geq 1$. On the other hand, this is a solitary example, in most of the other cases the value of F was below 1. Range from which to choose the crossover factor is difficult to narrow down. The interval from the literature $0 \leq CR \leq 1$ can be cropped by the the values of $CR \leq 0.1$ as DE with the crossover factor set below this value never performed good optimisation.

4. Conclusions

The aim of this work was to create maps visualising the relationship between the DE's configuration parameters F, CR and the solution quality. Presented maps show dependency of final solution quality not only on the DE's configuration parameters, but also on DE variant and mainly on particular optimised function. There are some similarities betweeen the maps for various cost functions and DE variants, but no solid association can be determined. Nonetheless, it is obvious the generally recommended values of F, CR are not very useful. Every particular problem needs to have these values mapped to find the ones giving best performance and result every or most of the time.
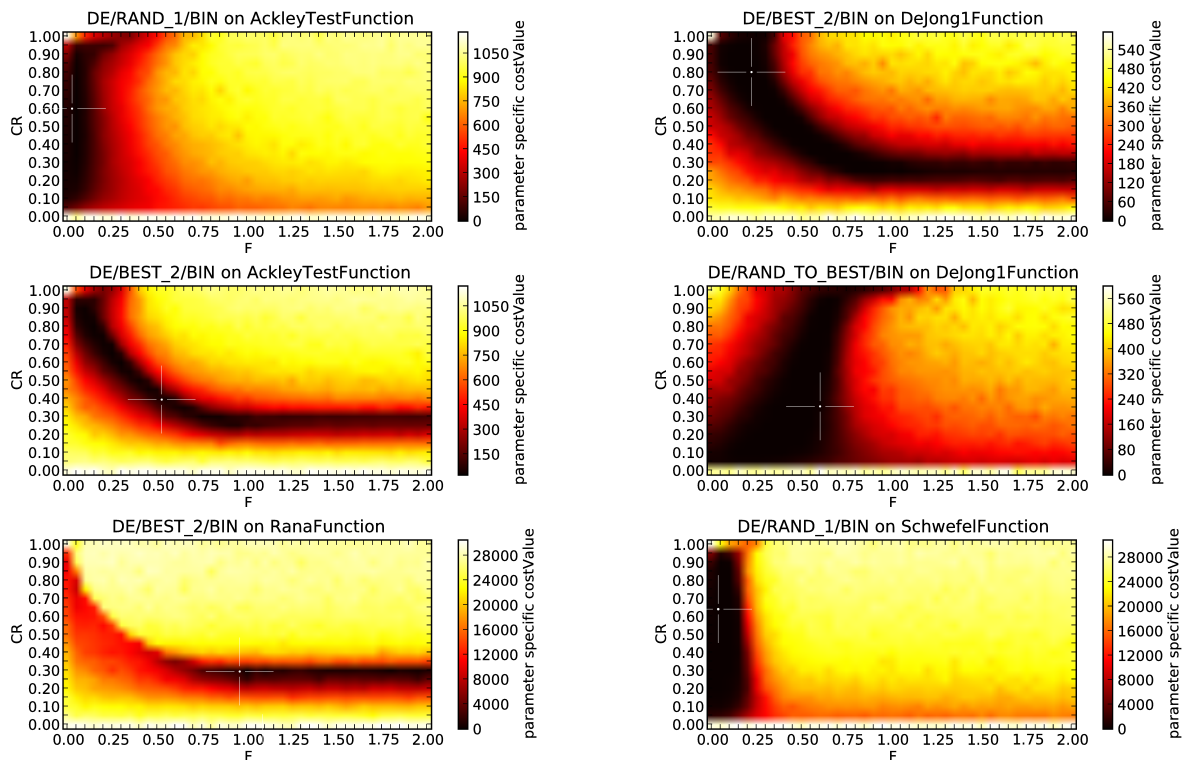
Fig. 1: DE solution quality as a function of F and CR control parameters. Note how much the areas of best restults differs between different variants of DE and given cost functions. Darker colour represents better solutions yielded for particular pair of (F, CR). White crosshair indicates best set of (F, CR) for specified experiment configuration.

The targeted outcome and the main reason for this preparative work is to define, design and validate a tool for efficient optimisation for acoustic emmision sensors placement on a wing which is currently being designed at the department with cooperation with external partners.

Acknowledgements

References

Ali, M. M., Törn, A. (2004), Population Set-Based Global Optimization Algorithms: Some Modifications and Numerical Studies In: *Computers & Operation Research 31*, pp 1703-1725.

Babu, B. V., Jehan, M. M. L. (2003), Differential Evolution for Multi-Objective Optimization In: *Proc. of the 2003 Congress on Evolutionary Computation (CEC'2003), Vol. 4*, pp 2696-2703.

Brest, J.,Zumer, V., Maucec, M. S. (2006), Self-Adaptive Differential Evolution Algorithm in Constrained Real-Parameter Optimization In: *Proc. of the 2006 IEEE Congress on Evolutionary Computation*, pp 215-222.

Eiben, A. E.,Hinterding, R., Michalewicz, Z. (1999), Parameter Control in Evolutionary Algorithms In: *Proc. of IEEE-EC*, pp 124-141.

Liu, J., Lampinen, J. (2002), On the Setting of Control Parameter of the Differential Evolution Method In: *Proc. of MENDEL*, pp 11-18.

Storn, R., Price, K. (1997), Differential Evolution – A Simple and Efficient Heuristic for global Optimization over Continuous Spaces In: *Journal of Global Optimization*, pp 341-359.

Teo, J. (2005), Differential Evolution with Self-adaptive Populations In: *KES, Lecture Notes in Computer Science, Vol. 3681*, pp 1284-1290.

Zaharie, D. (2007), A Comparative Analysis of Crossover Variants in Differential Evolution In: *Proc. of the International Multiconference on Computer Science and Information Technology*, pp 171-181.