A Polymorphical Mutation Operator for Evolution Strategies

Stefan Berlik
Department of Computer Science
University of Dortmund
44221 Dortmund, Germany
stefan.berlik@uni-dortmund.de

Abstract

Since the advent of evolution strategies many attempts have been done to improve the mutation operator [7]. The first enhancement could be seen in the introduction of a flexible step size. As a logical consequence later n step sizes, one per variable, were used. Highest flexibility is achieved with the correlated mutation but also quadratic growth of the number of strategy parameters [8]. A newer approach introduces asymmetry to the mutation operator getting by with linear growth of the strategy parameters [3]. The Polymorphical Mutation Operator, a new asymmetrical mutation operator, will be presented in the following. It overcomes some of the disadvantages of the one mentioned before.

Keywords: Polymorphical Mutation Operator, Polymorphical Mutation, Asymmetrical Mutation, Evolution Strategy, Optimization.

1 Asymmetrical Mutation

The main idea of the asymmetrical mutation is to focus on mutating into the most beneficial direction. In contrast to the classical mutation operator the possibility to get a positive or a negative random number is not equal any longer. The optimization strategy can adopt the most promising direction over the generations. Therefore the mutation operator needs a customizable asymmetrical distribution where the grade of asymmetry can be adjusted via a parameter.

For the asymmetrical mutation described in [3] an additive approach has been chosen. The density function is defined in sections and made up of two parts, one function for the negative domain and another function for the positive domain. One of these functions is exactly the density of the standard normal distribution. The other one is a compressed or expanded density function of the normal distribution. This construct has to be multiplied by a factor to normalize the integral value to one. The complete definition splits into four cases and can be seen below.

$$f_{AM}(x,c) = \begin{cases} \dfrac{\sqrt{2}}{\sqrt{\pi}(1+\sqrt{1-c})} e^{-\frac{1}{2}x^2} & \text{for } c < 0, x < 0 \\ \dfrac{\sqrt{2}}{\sqrt{\pi}(1+\sqrt{1-c})} e^{-\frac{x^2}{2(1-c)}} & \text{for } c < 0, x \geq 0 \\ \dfrac{\sqrt{2}}{\sqrt{\pi}(1+\sqrt{1+c})} e^{-\frac{1}{2}x^2} & \text{for } c \geq 0, x < 0 \\ \dfrac{\sqrt{2}}{\sqrt{\pi}(1+\sqrt{1+c})} e^{-\frac{x^2}{2(1+c)}} & \text{for } c \geq 0, x \geq 0 \end{cases} \quad (1)$$

As has been shown in [3] for some test problems, this mutation operator leads in many cases to faster convergence towards the optimum than the classical variants.

Nevertheless, some disadvantages are associated with this approach

- early stagnation occurs during the optimization of real world problems,
- expected values diverge for increasing asymmetry parameters and
- variances diverge.

The following figure 1 shows the graphs of the expected value and variance of the asymmetrical mutation.

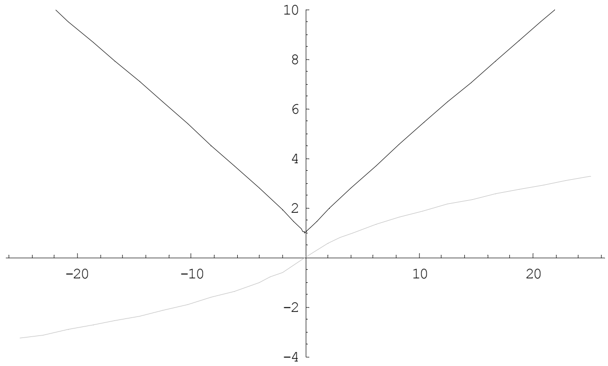

Figure 1: Expected value (depicted in gray) and variance of the Asymmetrical Mutation Operator for $c \in [-25, 25]$

The manner of these two parameters bares the reason for the stagnation during the optimizations. One can see that the asymmetry parameters have an objectionable impact on the variances and thus the step sizes. By favoring one direction the step size is enlarged implicitly. During the optimization process this has different effects.

At the beginning, after promising manipulation directions for the step sizes an the directions have been learned, the implicit step size manipulation (i.e. magnification) by the asymmetry parameters yields accelerated optimization progress.

After the strategy parameters have been adjusted to mutate toward the next local optimum the asymmetrical mutation strategy has problems in readjusting the parameters. Because of the interference of the different parameters new directions are hardly to be learned, thus the method is of limited applicability in high dimensional real world multimodal optimization problems.

2 Development of a Polymorphical Mutation Operator

The construction method for the polymorphical mutation differs significantly from that used to build up the asymmetrical mutation.

To overcome the drawbacks and to be able to express the density function in a closed form, for the polymorphical mutation a multiplicative approach has been chosen.

$$f_{PM}(x,a) = \sqrt{\frac{2}{\pi}} \frac{1}{1+e^{-a\,x}} e^{-\frac{x^2}{2}} \tag{2}$$

Formula (2) consists mainly of three parts

- the well known exp-term of the normal distribution, on the right side,
- a term introducing the asymmetry in the middle
- and a normalization factor on the left.

Most interesting of course is the term in the middle. It is a sigmoid function that enforces one side and attenuates at the same time the other one. In fact, this is a distribution itself, namely the Fermi-Dirac distribution function. By the real-valued parameter *a* the intensity of the asymmetry can be adjusted. For some values of *a* the polymorphical density functions are plotted in figure 2. One can see, that in contrast to the asymmetrical mutation there really occurs a change of the shape. The density function of the normal distribution is contained as a marginal in the case of $a=0$, meaning symmetry.

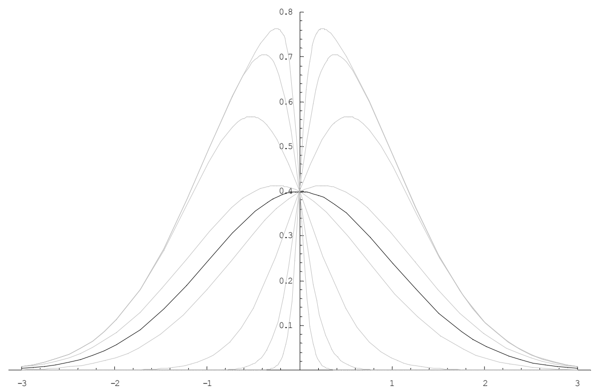

Figure 2: Probability density functions of the Polymorphical Mutation Operator for $a \in [-25, 25]$

The corresponding distributions are shown in figure 3. Again, the normal distribution is depicted in black.

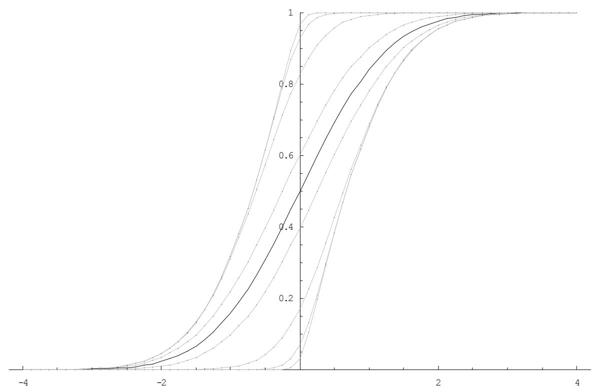

Figure 3: Distribution functions of the Polymorphical Mutation Operator for $a \in [-25, 25]$

3 Expected Values and Variances

In general, to generate asymmetrical random numbers, one has to introduce expected values unequal to zero. Of course this means that the mutation operator is not compliant to standard evolution strategy any longer postulating an expected value of zero, thus mutating around the origin. On the other hand it has to be ensured that the expected values converge, thereby forcing the mutation operator to continue mutating near by the origin. This demand is violated by the Asymmetrical Mutation Operator. Here mutating can lead to 'jumping' due to diverging expected values caused by increasing asymmetry parameters.

Of even greater interest are the variances. They can be seen as measures for the step sizes of the mutations and are strategy parameters in most evolution strategies. Because of this they should not be modified by the asymmetry parameters. In the ideal case the variance is a constant, independent of the asymmetry parameter. At least convergence is necessary and a small spread between minimal and maximal value is desired to limit the impact of the asymmetry on the step sizes. Again, the Asymmetrical Mutation Operator violates this demand.

The mentioned requirements should be fulfilled by any asymmetrical mutation operator. To see that this is the case for the Polymorphical Mutation Operator the next figure is dedicated to its expected value and variance.

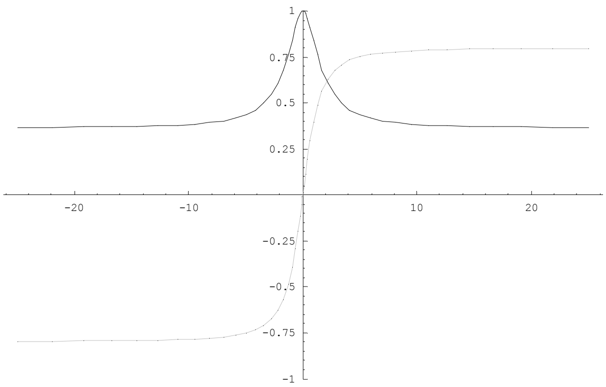

Figure 4: Expected value (depicted in gray) and variance of the Polymorphical Mutation Operator for $a \in [-25, 25]$

One can see, that the expected value is as desired, having a spread of about 1.6. Also the variance complies with the minimal demands, it spreads about 0.6.

4 Results

Some tests have been done using the Polymorphical Mutation Operator to optimize the scoop area of a twin-screw compressor. Details about twin-screw compressors, more results concerning this and other objectives can be found in [4] and [5]. Also multiobjective optimizations have been done using a NSGA-II like evolution strategy implemented with the EO-framework[1], refer to [2] resp. [6].

The dimensionality of this highly constrained example problem is 30 and the evaluation of the objective is expensive. Because of the hard constraints it has been made use of the extended creation scheme [1]. All runs have been done using an (8*30) evolution strategy, where the * indicates the use of extended creation scheme which has an influence of on the selection. The results can be seen in the following two diagrams, the first one for the polymorphical mutation, the second one for the asymmetrical mutation.

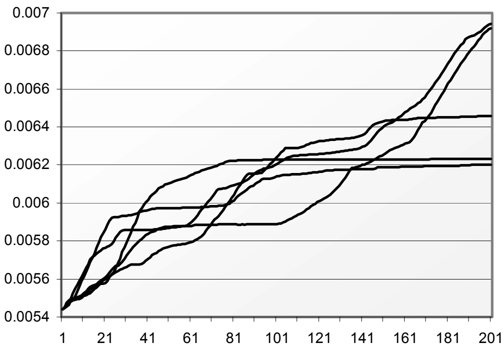

Figure 5: Maximization of the scoop-area using the polymorphical mutation

[1]EO (abbr. for 'Evolving Objects') is a templates-based, ANSI-C++ compliant evolutionary computation library. http://eodev.sourceforge.net/

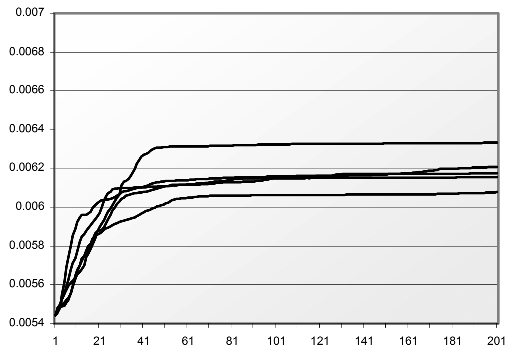

Figure 6: Maximization of the scoop-area using the asymmetrical mutation

As can be seen from the two diagrams the problem of stagnation is not that distinct using the Polymorphical Mutation Operator. Further it yields better overall results. What also can be seen is that the asymmetrical mutation is faster at the beginning of the optimizations, as explained before.

5 Conclusion and Outlook

The Polymorphical Mutation Operator is an asymmetrical mutation operator that overcomes the main disadvantages of the previous one, diverging expected value and variance. Besides it can be expressed in closed form. It has shown its capability and robustness in high dimensional, constrained, and multimodal real world optimization problems, for single objective optimizations as well as for multi-objective optimizations.

Further work has to be done in the field of symbolic integration. The integral of the polymorphical density function is still not known. All results have been obtained by numerical methods. But it has already been proven that no elementary integral can exist. Even integrating in terms of elementary functions and error functions might not be possible. Another topic that has to be studied in detail is the convergence speed.

6 References

[1] S. Berlik, M. Helpertz (2002). Optimizing Screw Machines using an Evolution Strategy with Extended Creation Scheme. In *IEEE International Conference on Systems, Man and Cybernetics (SMC 2002).*

[2] K. Deb (2001), *Multi-Objective Optimization using Evolutionary Algorithms*. John Wiley & Sons, Chichester, 2001.

[3] L. Hildebrand (2002). *Asymmetrische Evolutionsstrategien*. PhD thesis, Department of Computer Science, University of Dortmund, 2002.

[4] K. Kauder, B. Reusch, S. Berlik, M. Helpertz (2001). Automated Geometry Optimization of Rotors of Twin-Screw Compressors – Part 1. In K. Kauder, editor, *Schraubenmaschinen*, volume 9, pages 27-46, Dortmund, 2001.

[5] K. Kauder, B. Reusch, S. Berlik, M. Helpertz (2002). Automated Geometry Optimization of Rotors of Twin-Screw Compressors – Part 2. In K. Kauder, editor, *Schraubenmaschinen*, volume 10, pages 17-34, Dortmund, 2002.

[6] K. Kauder, B. Reusch, S. Berlik, M. Helpertz (2002). Optimization Methods for Rotors of Twin-Screw Compressors. In *Schraubenmaschinen 2002*, VDI-Berichte 1715, pages 29-50. VDI-Verlag, Düsseldorf, 2002.

[7] I. Rechenberg (1994). *Evolutionsstrategie `94*. frommann-holzboog, Stuttgart, 1994.

[8] H.-P. Schwefel (1994). *Evolution and Optimum Seeking*. John Wiley & Sons, New York, 1994.