A Step Size Preserving Directed Mutation Operator

Stefan Berlik

Universität Dortmund, Computer Science, Chair I,
44221 Dortmund, Germany
Stefan.Berlik@Uni-Dortmund.de
http://ls1-www.cs.uni-dortmund.de/

Abstract. Using a directed mutation can improve the efficiency of processing many optimization problems. The first mutation operators of this kind proposed by Hildebrand [1], however, suffer from the asymmetry parameter influencing the mutation step size. Extreme asymmetry can lead to infinite step size. The operator presented here overcomes this drawback and preserves the step size.

The main idea of the directed mutation is to focus on mutating into the most beneficial direction by using a customizable asymmetrical distribution. In this way the optimization strategy can adopt the most promising mutation direction over the generations. It thus becomes nearly as flexible as with Schwefel's correlated mutation [2] but causes only linear growth of the strategy parameters instead of quadratic growth.

A normalization function is introduced to decouple asymmetry from the variance, i.e. the step size. By incorporating the normalization function the variance becomes independent of the asymmetry parameter. Given below are the definitions of the density function for the normalized directed mutation and its normalization function:

$$f_{\sigma,a}(x) = \begin{cases} \sqrt{\frac{2}{\pi}} \frac{\sqrt{1-a}}{\left(1+\sqrt{1-a}\right)\, \sigma_{\text{norm}}(a)\sigma} e^{-\frac{x^2}{2(\sigma_{\text{norm}}(a)\sigma)^2}} & \text{for } a \le 0, x \le 0 \\ \sqrt{\frac{2}{\pi}} \frac{\sqrt{1-a}}{\left(1+\sqrt{1-a}\right)\, \sigma_{\text{norm}}(a)\sigma} e^{-\frac{(1-a)x^2}{2(\sigma_{\text{norm}}(a)\sigma)^2}} & \text{for } a \le 0, x > 0 \\ \sqrt{\frac{2}{\pi}} \frac{\sqrt{1+a}}{\left(1+\sqrt{1+a}\right)\, \sigma_{\text{norm}}(a)\sigma} e^{-\frac{(1+a)x^2}{2(\sigma_{\text{norm}}(a)\sigma)^2}} & \text{for } a > 0, x \le 0 \\ \sqrt{\frac{2}{\pi}} \frac{\sqrt{1+a}}{\left(1+\sqrt{1+a}\right)\, \sigma_{\text{norm}}(a)\sigma} e^{-\frac{x^2}{2(\sigma_{\text{norm}}(a)\sigma)^2}} & \text{for } a > 0, x > 0 \end{cases} \tag{1}$$

$$\sigma_{\text{norm}}(a) = \sqrt{\frac{\pi\left(1+|a|\right)}{4\left(\sqrt{1+|a|}-1\right) + |a|(\pi-2) + \pi\left(2-\sqrt{1+|a|}\right)}}\,. \tag{2}$$

Formulas for the expected value and variance of a random variable X distributed according to the normalized asymmetrical distribution take the following form:

K. Deb et al. (Eds.): GECCO 2004, LNCS 3102, pp. 786–787, 2004.

$$\mathrm{E}(X)=\sqrt{\frac{2}{\pi}}\frac{a\sigma_{\text{norm}}(a)\,\sigma}{1+|a|+\sqrt{1+|a|}},\qquad \mathrm{V}(X)=\sigma^2. \tag{3}$$

To get a notion of this distribution the following figure shows some graphs of the density function and distribution for different asymmetry settings.

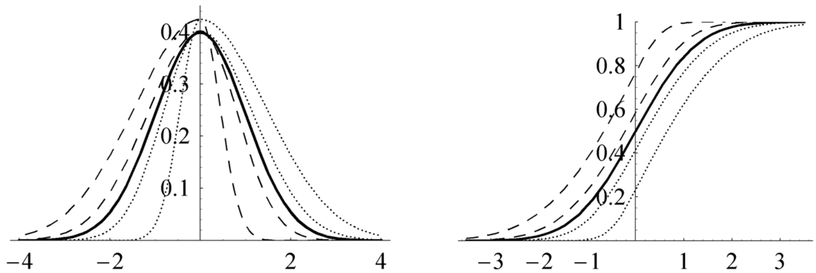

Fig. 1. Density function and distribution of the *normalized asymmetrical mutation* for $\sigma = 1$. Asymmetry parameters: *a = -10, a = -1 (dashed); a = 0 (solid); a = 1, a = 10 (dotted).*

Random numbers distributed according to the normalized asymmetrical distribution can be generated by multiplying its inverse function with uniformly distributed random numbers. The inverse function is defined by

$$\bar{\mathrm{F}}_{\sigma,a}(y)=\begin{cases}\sqrt{2}\,\sigma_{\text{norm}}(a)\,\sigma\,\text{inverf}\left(y\left(1+\frac{1}{\sqrt{1-a}}\right)-1\right) & \text{for } a\le 0,\ y\le\frac{\sqrt{1-a}}{1+\sqrt{1-a}}\\ \frac{\sqrt{2}\,\sigma_{\text{norm}}(a)\,\sigma}{\sqrt{1-a}}\text{inverf}\left(y\left(1+\sqrt{1-a}\right)-\sqrt{1-a}\right) & \text{for } a\le 0,\ y>\frac{\sqrt{1-a}}{1+\sqrt{1-a}}\\ \frac{\sqrt{2}\,\sigma_{\text{norm}}(a)\,\sigma}{\sqrt{1+a}}\text{inverf}\left(y\left(1+\sqrt{1+a}\right)-1\right) & \text{for } a>0,\ y\le\frac{1}{1+\sqrt{1+a}}\\ \sqrt{2}\,\sigma_{\text{norm}}(a)\,\sigma\,\text{inverf}\left(y\left(1+\frac{1}{\sqrt{1+a}}\right)-\frac{1}{\sqrt{1+a}}\right) & \text{for } a>0,\ y>\frac{1}{1+\sqrt{1+a}}.\end{cases} \tag{4}$$

Using the normalized directed mutation has shown to be very effective in optimizing test functions, as well as real world problems [3]. Taking into account that the application of the operator itself is quite fast, e.g. compared to the correlated mutation, the use of the directed mutation might be quite beneficial for many problems.

References

1. Hildebrand, L.: Asymmetrische Evolutionsstrategien. PhD thesis, Department of Computer Science, Universität Dortmund (2002)
2. Schwefel, H.-P.: Evolution and Optimum Seeking. John Wiley & Sons, New York (1994)
3. Berlik, S.: A Polymorphic Mutation Operator for Evolution Strategies. In: Proc. of the 3rd Int. Conf. in Fuzzy Logic and Technology, EUSFLAT'03, Zittau, Germany (2003)