**A short study on minima distribution**

**Abstract**

|  |  |
| --- | --- |
|  |  |

**1. Introduction**

**2. Weak conditions of convergence and monotonicity**

Two most important properties of minima distribution are convergence and monotonicity. Stability and shrinkage are derived from the two properties. As a summary, Luo (Luo, 2019, p. 5) proposed following conditions for satisfying convergence and monotonicity of minima distribution:

* Function *τ*(*x*) is defined based on function *ρ* such that *τ*(*x*) = *ρ*(*f*(*x*)) where *ρ* is monotonically decreasing and *ρ* is positive, for instance, *ρ*(*y*) > 0 for all *x* belonging the domain *X* such that *y* = *f*(*x*).
* Function *τ*(*x*) is defined as *τ*(*x*) = exp(–*f*(*x*)) so as to easily obtain the monotonicity.

However, it is not necessary to strictly define *τ*(*x*) based on *ρ* and *f*(*x*) for convergence. Following the proofs of theorem 1 in (Luo, 2019, p. 7), a *weak convergence condition* is drawn that:

* Function *τ*(*x*)is positive for all *x* belonging to domain *X*.
* Given any *x* which is not a maximizer of *f*(*x*), there always exists an open set having nonzero Lebesgue measure such that *τ*(*t*) > *τ*(*x*) for all *t* belonging to this open set.

This condition is easily proved by proofs of theorem 1 (vi, vii) and theorem 2 (Luo, 2019, pp. 7-8). Of course, if *τ*(*x*) is defined as *τ*(*x*) = *ρ*(*f*(*x*)) where *ρ* is monotonically decreasing and *ρ* is positive, the weak convergence condition is satisfied. This condition is useful in practical cases that *τ*(*x*) is defined as a nonconstant, positive, and differential function which has no actual minimizer with note that some infima may not be actual minima.

Within the weak convergence condition, theorem 3 for nonnegativity (Luo, 2019, p. 8) is obtained because *τ*(*x*)is positive.

Monotonicity can be achieved if the first order derivative of with regard to *k* is greater than or equal to 0. Suppose *m*(*k*)(*x*) is differentiable, we have:

Here the positive function *τ*(*x*) is defined as *τ*(*x*) = *ρ*(*f*(*x*)) where *y* = *f*(*x*) such that:

|  |  |
| --- | --- |
|  | (1.1) |

Following the proof of theorem 1 (v) (Luo, 2019, p. 7), we have:

|  |  |
| --- | --- |
|  | (1.2) |

Where:

Therefore, we obtain (Luo, 2019, p. 9):

|  |  |
| --- | --- |
|  | (1.3) |

Where:

According to Gurland’s inequality (Luo, 2019, p. 6), if *g*(*y*) is nonincreasing with regard to *y* where *y* = *f*(*x*) then, *E*(*k*)(*yg*) ≤ (*E*(*k*)(*y*))(*E*(*k*)(*g*)). As a result, a *weak monotonicity condition* is that positive function *τ*(*x*) = *ρ*(*f*(*x*)) is defined by the way that equation 1.1 is satisfied in addition to that *g*(*y*) is nonincreasing regarding *y* where *y* = *f*(*x*). Note that *ρ* is not necessary to be monotonically decreasing and the weak monotonicity condition follows the weak convergence condition. Based on two weak conditions of convergence and monotonicity, convergence speed is proposed in the next section.

**2. Convergence speed**

For all real *k* and real Δ*k* > 0, two successive integrals within the weak convergence condition are determined as follows:

Convergence speed is defined as the absolute differential of two successive integrals as follows:

It is easy to recognize that the convergence speed is the absolute value of the first order derivative of specified by equation 1.2.

|  |  |
| --- | --- |
|  | (2.1) |

Note, *g*(*x*) is defined by equation 1.1. The magnitude of *cτ* is proportional to the difference between *y = f*(*x*) and *g*(*y*) which is determined by the following derivative:

Because 1 is constant, let *Qτ* be the metric that measures the convergence speed, which is defined as follows:

|  |  |
| --- | --- |
|  | (2.2) |

Actually, *Qτ* is absolutely slope of function *g*. The steeper *Qτ* is, the larger *cτ* is, which in turn, the faster convergence speed. For example, given *y* = *f*(*x*), if the definition of function *τ* is associated with normal distribution given mean 0 and variance 1, it becomes:

Therefore,

If the definition of function *τ* is associated with Gumbel distribution for extreme value given location parameter 0 and scale parameter 1, it becomes:

Therefore,

Obviously, convergence speed of *τG*(*x*) is faster than *τN*(*x*) because *gG*(*y*) is steeper than *gN*(*y*). In other words, *ρG*(*y*) is steeper than *ρN*(*y*). Given normal distribution with mean *μ* and variance *σ*2 then, *ρN*(*y*) is redefined as follows:

Variance *σ*2 reflects the steeper of *ρN*(*y*). Therefore, the smaller variance *σ*2 is, the steeper function *ρN*(*y*) is. The next section describes an experiment with varying variance *σ*2.

**3. Experiment with PSO algorithm**

In this research, convergence speed of minima distribution is calculated, which in turn derives the conclusion that function *τ*(*x*) = *ρ*(*f*(*x*)) should be more steeper regarding *f*(*x*) in order to improve the convergence speed. In particle swarm optimization (PSO) algorithm, movement of particles obeys normal distribution. Therefore, this section describes an experiment by varying variance of suchnormal distribution with note that *τ*(*x*) here is associated with the PSO normal distribution.

**4. Conclusions**

Some optimization algorithms like PSO take advantages of distribution of *x* instead of taking advantages of *f*(*x*). In other words, they define implicitly *τ*(*x*) as function of *x* instead of *τ*(*x*) = *ρ*(*f*(*x*)). Therefore, it is reasonable to assert their convergence by the weak convergence condition.

**References**

Luo, X. (2019, May 24). Minima distribution for global optimization. *arXiv preprint*. doi:10.48550/arXiv.1812.03457