**An extension of minima distribution**

**Abstract**

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

* Calculating the convergence speed
* Extending minima distribution with probabilistic distribution.
* Applying Gumbel distribution to Bayesian optimization
* Comparison of Gumbel distribution and normal distribution

**1. Introduction**

**2. An extension of minima distribution**

In theory, there are two conditions for *ρ*(*y*) such as 1) positiveness and 2) monotonicity but these conditions can be replaced by two weaker conditions as follows:

1. The limit of *ρk*(*y*) is existent when *k* approaches positive infinity.
2. If *y* = *f*(*x*) has a maximum *y*\* = *f*\* = *f*(*x*\*) then, *ρ*(*y\**) gets maximal.

For the weak condition 1, *ρk*(*y*) is actually a symbolic notation, which can be defined arbitrarily, for example, as *ρk*(*y*) = (*ρ*(*y*))*k*, *ρk*(*y*) = (*ρ*(*y*))*αk*, *ρk*(*y*) = , etc. where *α* is constant.

**2. Convergence speed**

For all real *k* and real Δ*k* > 0, corollary 1 (Luo, 2019, p. 10) specifies the following converged sequences:

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

Suppose *f*(*x*) and *τ*(*x*) are differentiable, we have

Because *f*(*x*) and *τ*(*x*) are differentiable, by applying L’ Hospital’s rule, we have:

According to theorem 1 (Luo, 2019, p. 6), we have:

Where,

Due to:

We obtain:

As a result, equation 2.1 specifies the convergence speed:

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

Where,

Suppose *τ*(*x*) is defined as exponential function of –*f*(*x*) such that:

We obtain

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

Where,

Suppose *E*(*k*)(*f*) is constant, let *g*(*x*) = (*f*(*x*) – *E*(*k*)(*f*))*f*(*x*), we have:

Because *g*(*x*) is proportional to *f*(*x*) and the sequences which are converged to the minimum *f*\* are decreased, convergence speed *cτ*,*f* gets larger when *g*(*x*) is inversely proportional to *m*(*k*)(*x*). In other words, the steeper function *τ*(*f*(*x*)) with regard to *f*(*x*) is, the faster convergence speed is. This conclusion is important in this research. For example, if the definition of function *τ* is associated with normal distribution given mean 0 and variance 1, it becomes:

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

Obviously, convergence speed of *τG*(*x*) is faster than *τN*(*x*) because *τG*(*x*) is steeper than *τN*(*x*). In formal specification, how steep a function is depends on the magnitude of its derivative. Given normal distribution with mean *μ* and variance *σ*2 then, *τN*(*x*) is redefined as follows:

Variance *σ*2 reflects the steeper of *τN*(*x*). Therefore, the smaller variance *σ*2 is, the steeper function *τN*(*x*) 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*) 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**

**References**