Vol.2019-BIO-57 No.9
2019/3/9

Aggregated CMA-ES: An Effective and Stable Strategy for Neuron Model Optimization

Xu Han[1] Takahiro Shinozaki[1] Ryota Kobayashi[2]

概要：Covariance Matrix Adaptation Evolutionary Strategy (CMA-ES) is considered as one of the most effective method for black-box optimization issue. In this paper, we apply CMA-ES to the neuron model parameter optimization problem, and compare it with genetic algorithm (GA) and the Nelder-Mead method which are the widely used approaches. To enhanced robustness of CMA-ES, we extend it by making an aggregation of evolution. We analyze a public dataset recorded from a rat neocortical neuron, which shows that the proposed approach achieves higher performance than the conventional methods.

1. Introduction

In this paper we investigate the performance of CMA-ES by comparing it with conventional methods, including genetic algorithm (GA) and the Nelder-Mead method. We found that the performance of CMA-ES highly depends on the initial condition, i.e., CMA-ES performs better than GA on average, however it sometimes performs worse than GA. To overcome this drawback, we propose to make an aggregation of CMA-ES for the optimization.

2. CMA evolution strategy

2.1 Covariance matrix adaptation evolution strategy

Covariance matrix adaptation evolution strategy (CMA-ES) is an evolution strategy. It uses a multivariate Gaussian distribution $\mathcal{N}(\boldsymbol{x}|\boldsymbol{\theta})$ having a parameter set $\boldsymbol{\theta} = \{\boldsymbol{\mu}, \boldsymbol{\Sigma}\}$ to represent a gene distribution, where $\boldsymbol{x}$ is the real-valued vector representing a gene, $\boldsymbol{\mu}$ is a D-dimensional mean vector, $\boldsymbol{\Sigma}$ is a $D \times D$-dimensional covariance matrix, and D is the gene size. Instead of directly maximizing the fitness $f(\boldsymbol{x})$, CMA-ES maximizes an expected value of the fitness $\mathbb{E}[f(\boldsymbol{x})|\boldsymbol{\theta}]$ under the Gaussian distribution. Higher expectation means the Gaussian distribution generates good genes with high probability.

2.2 Aggregated CMA-ES

The results of CMA-ES largely vary for different trials, that is, the performance highly depends on the initialization. This is maybe because the landscape of the objective function in the gene space is fitted by a Gaussian distribution, which only describes a symmetric distribution with a single peak.

To improve CMA-ES for achieving near global optimum result, we propose an approach that runs CMA-ES for multiple trials with different initialization and finds the best individual among all the trials. We refer to this strategy as aggregated CMA-ES, and the number of trials as aggregation size. The original CMA-ES is a special case of the aggregated CMA-ES whose aggregation size is 1. In this way, we have multiple Gaussian distributions in an aggregated evolution process, and can represent complex divergence as a whole.

3. Neuron model

We considered the parameter optimization problem of Multiple-timescale Adaptive Threshold (MAT) model [1]. The neuron model generates spikes when the potential exceed the spike threshold. There are 5 threshold parameters that need a black-box optimization and we apply GA and CMA-ES.

We evaluate the model performance by the coincidence factor, which is defined by Equation (1).

$$\Gamma = \frac{N_c - 2f_m N_d \Delta}{N_d + N_m} \times \frac{2}{1 - 2f_m \Delta}, \tag{1}$$

where N_c is the number of coincident spikes with precision $\Delta = 4$ (ms), N_d (N_m) is the number of spikes of the real (model) neuron, and f_m is the spike frequency of the model neuron. The maximum value of $\Lambda = 1$ is achieved only if all the spikes coincide with precision Δ.

[1] Tokyo Institute of Technology, Japan
[2] National Institute of Informatics, Japan

表 1: Optimization by grid search and Nelder-Mead method.

Strategy	Grid Search	Nelder-Mead
Coincidence factor (Γ)	0.580	0.618

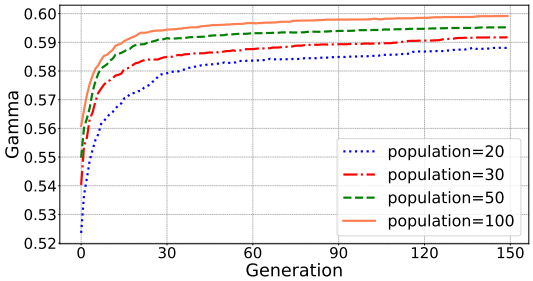

図 1: Results of GA using several population sizes.

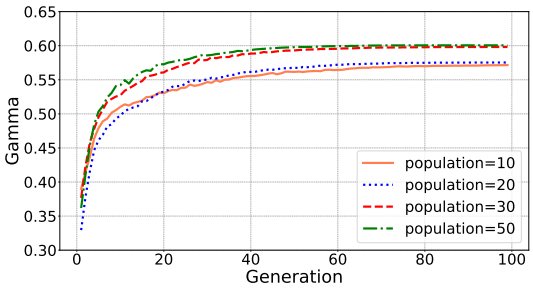

図 2: Results of CMA-ES using several population sizes.

4. Experiment and Results

We analyzed a public dataset from the International Competition on Quantitative Single-Neuron Modeling 2009 [2], [3] (Challenge A). The data consists of the stimulus and the voltage recorded from a rat neocortex neuron. Five threshold parameters are optimized by applying GA and CMA-ES to the dataset between 17.5 sec and 39 sec. We generated these initial genes by random sampling from a multidimensional uniform distribution over an interval. The fitness is evaluated by the model performance Γ. Because the results involve randomness, we repeated the experiments multiple times using independently sampled genes for the initialization.

For a comparison purpose, Table 1 shows results obtained by using the grid search and the Nelder-Mead method [4], which is one of the state-of-art optimization methods for neuron models. Figure 1 and Figure 2 shows the averaged performance of GA and CMA-ES with several population sizes. The horizontal axis is the generations and the vertical axis is the coincidence factor Γ. The experiments were repeated 50 times with different random initialization to obtain the averaged results. GA needs larger population size to obtain higher Γ. Γ improves with the generations until 150-th generation. When the population size was 100, Γ was 0.598 at 100-th generation and it was 0.599 at 150-th generation. CMA-ES can obtain higher Γ

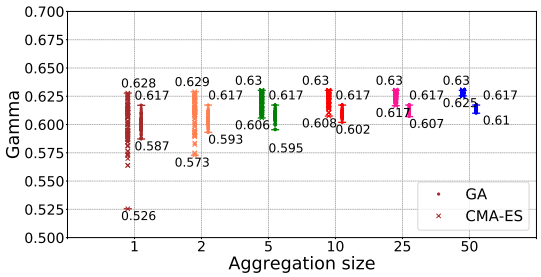

図 3: Distribution of Γ when aggregated GA and CMA-ES were used. The population size and the number of generations were 100 and 150 for GA, and 50 and 100 for CMA-ES.

than GA with smaller population size and converge faster (Before the 100-th generation, CMA-ES converged in most of the cases) At the 100-th generation, Γ by CMA-ES with 50 population size was 0.602.

Figure 3 shows the results of aggregated CMA-ES with different aggregation sizes and that of aggregated version of GA. As can be seen, aggregated CMA-ES provides higher Γ than the aggregated GA when we choose the aggregation size larger than 5. The best Γ was 0.617 by GA and 0.630 by CMA-ES when the aggregation size was larger than 5.

5. Conclusion

We have applied CMA-ES to a neuron model optimization problem for a public dataset recorded from rat neocortex, and have demonstrated that its averaged performance outperforms conventional optimization approaches based on GA and Nelder-Mead methods. The fitness obtained by CMA-ES largely depends on the initial condition. To address the problem, we tried a simple strategy that makes the aggregation of CMA-ES for the optimization, which we referred to as aggregated CMA-ES. Experimental results show that aggregated CMA-ES is robust against the initial condition, and stably achieves better performance than conventional methods.

参考文献

[1] Kobayashi, R., Tsubo, Y. and Shinomoto, S.: Made-to-Order Spiking Neuron Model Equipped with a Multi-Timescale Adaptive Threshold, *Frontiers in Computational Neuroscience*, Vol. 3,9 (2009).
[2] Gerstner, W. and Naud, R.: How good are neuron models?, *Science*, Vol. 326, No. 5951, pp. 379–380 (2009).
[3] Gerstner, W., Naud, R., Carandini, M., Sincich, L., Horton, J., Adams, D. and Economides, J.: Quantitative Single-Neuron Modeling Competition for year 2009, developed by researchers at EPFL; data from Wistar rat cortical neurons and rhesus monkey lateral geniculate nucleus and retinal ganglion cells, *CRCNS.org.* (2009).
[4] Nelder, J. A. and Mead, R.: A Simplex Method for Function Minimization, *The Computer Journal*, Vol. 7, No. 4, pp. 308–313 (1965).