Малик Ігор Володимирович, доктор фіз.-мат. наук, професор

Літвінчук Юлія Анатоліївна

Чернівецький національний університет ім. Ю. Федьковича, Чернівці

https://orcid.org/0000-0002-5277-0109

МОДЕЛЮВАННЯ РОЗШИРЕНОГО АЛГОРИТМУ CMA-ES НА БАЗІ СУМIШЕЙ НОРМАЛЬНИХ РОЗПОДIЛIВ

Коваріаційна матриця стратегії еволюції адаптації (Covariance matrix adaptation, CMA-ES) один із кращих методів оптимізації гіперпараметрів нейронної мережі. CMA-ES є сучасним еволюційним алгоритмом, який досить добре працює для великої кількості параметрів [1]. Даний алгоритм являється більш потужнішим в порівнянні з Байєсiвською оптимiзацiєю [2], пошуком в сітці, випаковим пошуком та має деякi кориснi властивостi iнварiантностi. Суть даного методу полягає у перерахунку коварiацiйної матрицi розподiлiв гiперпараметрiв мiж епохами еволюцiйного алгоритму з подальшим вибором параметрiв та врахуванням даної матрицi.

Однак, недолiком методу CMA-ES є те, що припускається однопiковiсть щiльностi розподiлу гiперпараметрiв. Та на практицi у великій кількості оптимізаційних задач цiльова функцiя не є однопiковою, що приводить до збiльшення областi пошуку та включення в область пошуку генетичного алгоритму область зi значеннями, що значно вiдрiзняються вiд локальних екстремумів.

У зв’язку з цим запропоновано розширенний CMA-ES алгоритм, що використовує багатопiковi моделi на базі сумiшей нормальних розподiлiв [3, 4]

$p(x)=\sum\_{s=1}^{k}π\_{s}p(x|μ\_{s};θ\_{s})$ (1)

де $p(x|μ\_{s};θ\_{s})$ – щільність багатовимірного нормального розподілу з $R^{d}$ та параметрами ($μ\_{s};θ\_{s}$) [5].

 Нехай $P(θ; X\_{1:k}, y\_{1:k})$ – розподiл гiперпараметрiв нейронної мере-
жi на основi значень цiльової функцiї, отриманої на основi $k$епох, де
$X\_{k}$– значення гiперпараметрiв на $k$-му кроцi, $y\_{k} $– значення цiльової
функцiї на $k$-му кроцi. Розглянемо новий алгоритм еволюцiйної стратегiї на основi розширеного CMA- ES, який можна описати наступними кроками:

1. Визначення областi змiни гiперпараметрiв $(a\_{0})$, розмiрностi сумiшi $(n)$, кiлькостi генiв в генетичному алгоритмi $(N)$, точностi методу $(ε)$.
2. Задання випадковим чином $(π^{(0)}, µ^{(0)}, θ^{(0)})$.
3. Вибiр $N$ генiв $X\_{k}$ згiдно розподiлу (1) та обчислень значень цiльової функцiї $y\_{k}$.
4. Перерахунок параметрiв $(π^{(k+1)}, µ^{(k+1)}, θ^{(k+1)})$ на основi ЕМ-алгоритму [6].
5. Якщо задовольняється умова виходу $|L\_{k+1} - L\_{k}| < ε$ то перейти до виконання генетичного алгоритму на основi розподiлу гiперпараметрiв з розподiлом $p(θ)=P(θ; X\_{1:k}, y\_{1:k})$. Якщо, $|L\_{k+1} - L\_{k}| < ε$ то перейти до кроку 3.

На практиці розширений CMA-ES алгоритм для оцiнки гiперпара-
метрiв складних систем у порівнянні з генетичними алгоритмами [7] ABCO, GA, PSO показав високу продуктивність (рис. 1).



 а) б)

Рис. 1. Графік залежності продуктивності генетичних алгоритмів

**Література**

1. Ilya Loshchilov , Frank Hutter. CMA-ES for Hyperparameter Optimization of Deep Neural Networks. arXiv:1604.07269v1 [cs.NE] 25 Apr 2016. – 9p.
2. Snoek J., Rippel O., Swersky K., Kiros R., Satish N., Sundaram N., Patwary M., Ali M., Adams R., et al. Scalable bayesian optimization using deep neural networks. arXiv preprint arXiv:1502.05700, 2015
3. Benyamin Ghojogh, Aydin Ghojogh, Mark Crowley, Fakhri Karray. Fitting A Mixture Distribution to Data: Tutorial. arXiv:1901.06708v2 [stat.OT] 11 Oct 2020. – 12p.
4. Lee, Gyemin and Scott, Clayton. Em algorithms for multivariate gaussian mixture models with truncated and censored data. Computational Statistics & Data Analysis, 56(9):2816–2829, 2012. – 13p.
5. Litvinchuk, Y. і Malyk, I. 2023. Розширений алгоритм стратегiї еволюцiї адаптацiї коварiацiйної матрицї. Буковинський математичний журнал. 10, 2 (Січ 2023), 137-143. DOI:https://doi.org/10.31861/bmj2022.02.09.
6. Малик Ігор Володимирович, Літвінчук Юлія Анатоліївна. Побудова еволюційної стратегії на основі сумішей. Інформаційне суспільство: технологічні, економічні та технічні аспекти становлення (випуск 70): 22-23 вересня 2022 р., с. 55-57. http://www.konferenciaonline.org.ua/ua/article/id-656/
7. Venkatesan D., Kannan K., Saravanan R. A genetic algorithm-based artifiial neural network
model for the optimization of machining processes. Neural Computing and Applications. February. 2009. 7p.