arXiv:2303.09478v1 [cs.LG] 16 Mar 2023

通过简单的基于种群的进化实现任意阶元学习

Chris Lu、Sebastian Towers 和 Jakob Foerster

工程科学系，牛津大学，英国牛津 christopher.lu@exeter.ox.ac.uk

摘要

元学习，即学习如何学习，使学习系统能够快速、灵活地解决新任务。这通常涉及定义一组外循环元参数，随后利用它们更新一组内循环参数。大多数元学习方法都采用复杂且计算开销巨大的双层优化方案来更新这些元参数。理想情况下，系统应执行多阶元学习，即学习如何学习如何学习，依此类推，从而加速自身的学习。不幸的是，标准元学习技术往往不适用于这些高阶元参数，因为元优化过程会变得过于复杂或不稳定。受到我们在现实世界中的进化中观察到的高阶元学习的启发，我们表明，简单的基于种群的进化能够隐式地优化任意高阶的元参数。首先，我们从理论上证明并通过实证表明，在一个简单的设定中，基于种群的进化能够隐式优化任意高阶的元参数。接着，我们引入一种最简的自指参数化方式，原则上可实现任意阶元学习。最后，我们表明，高阶元学习能够提升时间序列预测任务的表现。

引言

自然界包含多个层次的元进化与适应（Vanchurin 等人，2022年）。例如，DNA 不仅进化出了生物体，也进化出了可进化性（Zheng 等人，2020年；Woods 等人，2011年；Metzgar 和 Wills，2000年）。换言之，DNA 经历了进化，使得基因型中的随机突变常常能令由此产生的生物体表型发生有益或适应性的变化。此外，DNA 的进化创造了能够在其生命周期内适应的生物体，其中一种形式是进行强化学习的生物体（Bateson，1984）。这些学习型生物体还会通过社会互动和文化进一步影响自身的学习（Henrich，2015；Heyes，2018）。

然而，现有的大多数研究仅探讨单阶元学习，例如进化强化学习算法（Lu 等人，2022a）。这些方法通常采用计算成本高昂的双层优化方案，而这类方案很快就会变得不稳定，或在计算上变得难以处理，尤其是在应用于更高阶元学习时（Metz 等人，2021b）。

以往研究通过实证表明，基于种群的进化会隐式地选择单阶元学习，通常是通过同时进化突变率来实现的（Frans 和 Witkowski，2021；B¨ack 等人，1992；Smith，1998）。其他研究考察了多个层次的元学习，但其背景分别是基于梯度的优化（Chandra 等人，2019）和多智能体学习（Willi 等人，2022）。最后，Kirsch 和 Schmidhuber （ 2022 年）、 Lange 等人（ 2022 年）以及 Metz 等人（2021a）通过实证研究考察了在自指系统中使用类进化算法来实现自指元学习；这一思想最早由 Schmidhuber （ 1987 年）提出。然而，这些研究并未从理论上证明其实现了高阶元学习。我们将这些研究联系起来，从理论上证明并通过实证表明，在某些情况下，简单的基于种群的进化会选择任意高阶元学习；原则上，这使自指系统能够实现任意阶的自我改进。

数值适应度世界

我们通过在每一世代选择并变异适应度最高的前 k 个个体来进行基于种群的进化。与以往的工作不同，我们在包含多个阶元参数的基因组上进行这一操作。具体而言，对于具有 n 阶元参数的情况，我们将世代 t 的基因组 x 表示为一个包含 n 个参数的向量， $x_t = \{x_t^0, x_t^1, \cdots, x_t^n\}$，其中 x^i 表示第 i 阶元参数。我们考虑 Frans 和 Witkowski（2021）提出的“数值适应度世界”设定，其中适应度(x_t) $=x_t^0$。

我们变异 x_t 时使用以下更新规则：

$$x_{t+1}^i = x_t^i + x_t^{i+1} + B_t^i, 0 \leq i < n, 0 < t \quad (1)$$

$$x_{t+1}^n = x_t^n + B_t^n, 0 < t \quad (2)$$

$$B_t^i \sim \mathcal{N}(0, \beta), i.i.d, \forall t, i \quad (3)$$

换句话说，我们通过加上第 $i+1$ 个参数以及从正态分布中采样的噪声，来更新第 i 阶元参数。我们更新最后一个元参数

（x_t^n）时只加入噪声。若要创建一种自指式参数化，则通过加入它自身和噪声来更新最后一个元参数。这种确切形式的自指在大多数情况下可能并不合适，但对于其他参数化方式（例如神经网络）或许是合理的（Irie 等人，2022年）。

理论结果

我们证明，在此设定下，前 k 名选择当且仅当满足 $k > 1$时，才会选择适应度更高的高阶元参数。

令 P 定义由个体成员构成的种群，定义如上。令 x 为 P中的一个特定成员。令 P定义一个与 P 完全相同的种群，但x的第 n 个参数除外。更具体地说，$\bar{x}^n - x^n = \delta$，$\delta > 0$。

令 $F(P, B, t)$ 和 $F^{-1}(P, B, t)$ 分别表示种群 P 中 x 经过 t 世代的选择和突变向量 B 后，子代和非子代的适应度集合。注意，$|F(P, B, t+1)|$ 因此将是 x 在第 t 世代后的子代数量。

首先，我们证明，前 1 名（单基因组）选择不会选择高阶元参数。

定理 1. $\mathbb{E}[|F(P, B, n+1)|] = \mathbb{E}[|F(P, B, n+1)|]$在前 *1* 名选择下满足 $n > 1$。

证明。第 $t > 1$ 世代的子代数量完全由第一次选择决定。第 n 世代的种群成员要么全部 是 x 的子代，要么一个也不是。由于 x^n $(n > 1)$ 不影响第一次选择，因此它与子代数量无关。 □

接下来，我们证明前 k 名选择会为 $k > 1$ 选择高阶元参数。

引理 2。 $|F(P, B, n+1|B = b)| \geq |F(P, B, n+1|B = b)|$，对任意突变向量 b 均成立。

证明。注意，对于 $t < n$，$F(P, B, t|B = b) = F(P, B, t|B = b)$，由于适应度均不受 x^n 影响，而 x^n 是这两个种群之间唯一不同的值。

$F(P, B, n|B = b) = F(P, B, n|B = b) \oplus \delta$，其中 $\oplus$ 表示分布式加法。

$F^{-1}(\bar{P}, B, n|B = b) = F^{-1}(P, B, n|B = b)$ 因为x^n 无法在第 $n+1$ 世代之前影响其他成员的选择或适应度。

因此，在下一世代的前 k 名中，$\bar{x}$ 的子代数量不会少于 x 的子代数量。 □

定理 3。 $\mathbb{E}[|F(P, B, n+1)|] > \mathbb{E}[|F(P, B, n+1)|]$

证明。根据引理 2，$|F(P, B, n+1)| \geq |F(P, B, n+1)|$。因此，只需证明 $\mathbb{P}(|F(P, B, n+1)| > |F(P, B, n+1)|) > 0$ 即可得出我们的结论。具体而言，我们证明 $\mathbb{P}(|F(P, B, n+1)| = 0 \cap |F(P, B, n+1)| = 1) > 0$。在 B 上存在一个区间集合，使得 $F^{-1}(P, B, n)$ 中恰有 k 个成员落在范围$[\max F(P, B, n), \max F(P, B, n) + \delta]$内，其余成员均小于 $\max F(P, B, n)$。因此，经过选择后，$\bar{x}$恰有一个子代，而 x没有子代。 □

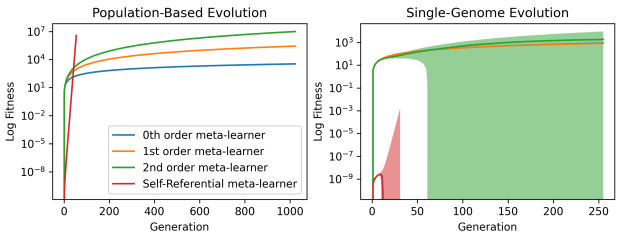

图 1：种群规模为 2048 时，采用不同元学习顺序的基于种群的进化（前 2 名）与单基因组进化。阴影区域表示 1024 个随机种子所得结果的均值标准误。

	元学习顺序			
$f(t)$	0th	1st	2nd	3rd
t	1.0	**3.7e-4**	9.4e-3	5.0e-2
t^2	1.3e7	7.5	0.77	**0.56**
$\sin(t)$	6.6e-2	1.0e-3	**9.4e-4**	1.2e-2
$\sin(t\sin(t))$	2.4	0.67	0.31	**0.16**

表1：种群规模为16384、采用前1024名选择并使用64个随机种子的情况下，经过4096代进化后的平均预测误差。对于每个实验，我们都调整了$\beta \in \{1.0, 0.5, 0.1, 0.05, 0.01\}$

.

实证结果

我们使用 JAX（Bradbury等人，2018）模拟进化，并在图 1 中展示结果。我们观察到，适应度的渐近增长大致为x^n ，其中 n 为元参数的数量。此外，自指元学习器的适应度呈指数级增长。因此，在我们的基于种群的设置中，更高阶的元参数能够提升适应度。相比之下，在单基因组选择中，适应度的期望值基本不受元参数数量影响，这表明单基因组选择无法进行元优化。

时间序列预测

接下来，我们考虑一个时间序列预测任务，其目标是预测某个函数 $f(t)$ 的下一个值。个体 x_t 的适应度由 适应度$(x_t) = -|f(t/100) - x_t^0|$ 决定。我们在表1中报告了若干函数上的结果。在许多此类设置中，更高阶的元参数可以提升性能。

未来工作

可以研究多智能体系统（Lu 等人，2022b,c）或人工生命（Langton，1997）中高阶元学习的涌现。未来的工作还应从理论上分析这些系统的长期性质，并在更实际的时间序列预测任务中评估其他参数化和选择方案。

参考文献

B¨ack, T. 等 (1992). 遗传算法中的自适应. 载于第一届欧洲人工生命会议论文集，第263–271页。剑桥麻省理工学院出版社。

Bateson, P. (1984). 基因、学习、生物学与学习. 载于《学习生物学与学习研讨会学习生物学报告》，柏林，*1983*年，*10*月*23*日至*28*日，第75–88页。施普林格。

Bradbury, J., Frostig, R., Hawkins, P., Johnson, M., Leary, C., Maclaurin, D., Necula, G., Paszke, A., VanderPlas, J., Wanderman-Milne, S. 和 Zhang, Q.（2018）。JAX：可组合变换的 Python+NumPy 程序。

Chandra, K., Meijer, E., Andow, S., Arroyo-Fang, E., Dea, I., George, J., Grueter, M., Hosmer, B., Stumpos, S., Tempest, A., 等（2019）。梯度下降：终极优化器。*arXiv* 预印本 *arXiv:1909.13371*。

Frans, K. 和 Wilkinson, D. (2022). 基于种群的进化优化元学习目标。*arXiv* 预印本 *arXiv:2103.06435*。

Henrich, J. (2015). 我们成功的秘密。普林斯顿大学出版社。

Heyes, C. (2018). 认知工具：思维的文化演化。哈佛大学出版社。

Irie, K., Schlag, I., Csord´as, R., 以及 Schmidhuber, J. (2022年). 一种能够学习如何修改自身的现代自指权重矩阵。发表于国际机器学习大会，第9660–9677页。PMLR。

Kirsch, L. 和 Schmidhuber, J. (2022年). 通过自指元学习消除元优化。*arXiv* 预印本 *arXiv:2212.14392*。

Lange, R. T., Schaul, T., Chen, Y., Zahavy, T., Dallibard, V., Lu, C., Singh, S., 和 Flennerhag, S.（2022年）。通过元黑盒优化发现进化策略。*arXiv* 预印本 *arXiv:2211.11260*。

Langton, C. G. (1997). 人工生命：概述. w.

Lu, C., Kuba, J., Letcher, A., Metz, L., de Witt, C. S., 以及 Foerster, J. (2022a). 发现式策略优化。*arXiv* 预印本 *arXiv:2210.05639*。

Lu, C., Willi, T., de Witt, C. A. S., and Foerster, J.（2022b）。模型无关的对手塑造。载于国际机器学习会议，第14398–14411页。PMLR。Lu, C., Willi, T., Letcher, A., and Foerster, J.（2022c）。对抗性廉价交谈。*arXiv*预印本 *arXiv:2211.11030*。Metz, L., Freeman, C. D., Maheswaranathan, N., and Sohl-Dickstein, J.（2021a）。使用随机初始化的学习型优化器训练学习型优化器。*arXiv*预印本 *arXiv:2101.07367*。Metz, L., Freeman, C. D., Schoenholz, S. S., and Kachman, T.（2021b）。梯度并非你所需的一切。*arXiv*预印本 *arXiv:2111.05803*。Metzgar, D. and Wills, C.（2000）。突变率适应性进化的证据。《细胞》，101(6):581–584。

Schmidhuber, J.（1987）。自指学习中的演化原理，或关于如何学习学习：元元……钩子。博士论文，慕尼黑工业大学。Smith, J.（1998）。演化算法中的自适应。博士论文，Citeseer。Vanchurin, V., Wolf, Y. I., Katsnelson, M. I., and Koonin, E. V.（2022）。迈向多层次学习的演化理论。《美国国家科学院院刊》，119(6):e2120037119。Willi, T., Letcher, A. H., Treutlein, J., and Foerster, J.（2022）。Cola：具有对手学习感知的一致性学习。载于国际机器学习会议，第23804–23831页。PMLR。Woods, R. J., Barrick, J. E., Cooper, T. F., Shrestha, U., Kauth, M. R., and Lenski, R. E.（2011）。大型大肠杆菌种群中促进可演化性的二阶选择。《科学》，331(6023):1433–1436。Zheng, J., Guo, N., and Wagner, A.（2020）。选择通过提高突变稳健性和可折叠性增强蛋白质可演化性。《科学》，370(6521):eabb5962。