Data-Driven Evolutionary Multi-Objective Optimization Based on Multiple-Gradient Descent for Disconnected Pareto Fronts*

Renzhi Chen[1] and Ke Li[2][0000−0001−7200−4244]

[1] National Innovation Institute of Defence Technology, Beijing, China
[2] Department of Computer Science, University of Exeter, Exeter, EX4 4QF, UK
k.li@exeter.ac.uk

1 Mathematical Proof

Proof. $\forall \mathbf{x} \in \mathbf{R}^d$, $\exists \epsilon > 0$, let $\mathbf{x}' = \mathbf{x} + \epsilon \mathbf{u}(\mathbf{x})$, according to Equation (5) and $\mathbf{u}(\mathbf{x}) \neq \mathbf{0}$, obviously we have:

$$\mathbf{u}(\mathbf{x}) \cdot \nabla f_i(\mathbf{x}) > 0$$

so for $i = 1, \ldots, m$:

$$\begin{aligned} f_i(\mathbf{x}') &= f_i(\mathbf{x} + \epsilon \mathbf{u}(\mathbf{x})) \\ &= f_i(\mathbf{x} + \frac{\mathbf{u}(\mathbf{x}) \cdot \nabla f_i(\mathbf{x})}{\|\nabla f_i(\mathbf{x})\|} \nabla f_i(\mathbf{x})) \\ &< f_i(\mathbf{x}) \end{aligned}$$

according to Definition 1, we have $\mathbf{x} \succ \mathbf{x}'$, $\mathbf{u}(\mathbf{x})$ is a convergent direction for objective functions $f_1(\mathbf{x}), \ldots, f_m(\mathbf{x})$.

Remark 1. The convergent speed for the MGD evolutionary search relate to the 'quality' of the initialization in Step 1 and the learning rate ϵ.

We assume that exists c satisfies:

$$f_j(\mathbf{x}) - f_j(\mathbf{x}^*) \geq c\mathcal{L}(\mathbf{x}), j = 1, \ldots, m \tag{1}$$

where $\mathcal{L}(\mathbf{x}) = \|\mathbf{x} - \mathbf{x}^*\|^2, \mathbf{x}^* = \mathrm{argmin}_{\mathbf{x}' \in PS}\{\|\mathbf{x} - \mathbf{x}'\|^2\}$ and the upper bound U for $\mathbf{u}^*$, $U = \max_{\mathbf{x} \in \mathbb{R}^d} \|\mathbf{u}^*(\mathbf{x})\|^2$, then we can give the bound of convergent speed as follow:

Lemma 1. *Sequence $\mathcal{S}_1, \ldots, \mathcal{S}_t$ is converge towards Pareto Set, as:*

$$\mathbb{E}[\mathcal{L}(\mathcal{S}_{t+1})] \leq \mathbb{E}[\mathcal{L}(\mathcal{S}_t)](1 - \epsilon_t c) + (\epsilon_t U)^2 \tag{2}$$

* This work was supported by UKRI Future Leaders Fellowship (MR/S017062/1), EPSRC (2404317), NSFC (62076056), Royal Society (IES/R2/212077) and Amazon Research Award.

2 Definitions of Benchmark Problems

Based on ZDT3, DTLZ7 and WFG2, we develop a series of problems (dubbed ZDT3$\star$, DTLZ7$\star$ and WFG2$\star$) with a controllable number of disconnected regions and imbalanced sizes. Their mathematical definitions are as follows.

– ZDT3$\star$

$$\begin{array}{ll} \text{minimize} & f_1(\mathbf{x}) = x_1 \\ \text{minimize} & f_2(\mathbf{x}) = g(\mathbf{x})(1 - \sqrt{x_1/g(\mathbf{x})} \\ & \qquad -x_1^{\alpha}/g(\mathbf{x}) \sin A\pi x_1^{\beta} \end{array}, \tag{3}$$

where

$$g(\mathbf{x}) = 1 + \frac{9}{n-1}\sum_{i=2}^{n} x_i. \tag{4}$$

– DTLZ7$\star$

$$\begin{array}{l} \text{minimize } f_1(\mathbf{x}) = x_1 \\ \qquad\qquad \vdots \\ \text{minimize } f_{m-1}(\mathbf{x}) = x_{M-1} \\ \text{minimize } f_m(\mathbf{x}) = (1 + g(\mathbf{x}_M))h(f_1, \cdots, f_{m-1}, g) \end{array}, \tag{5}$$

where

$$\begin{array}{ll} g(\mathbf{x}) & = 1 + \frac{9}{|\mathbf{x}_m|}\sum_{x_i \in \mathbf{x}_m} x_i \\ h(f_1, \cdots, f_{m-1}, g) & = m \\ & \quad - \sum_{i=1}^{m}[\frac{f_i}{1+g}(1 + f_i^{\alpha}\sin(A\pi f_i^{\beta}))] \end{array}, \tag{6}$$

where $x_i \in [0,1]$, $i \in \{1, \cdots, n\}$.

– WFG2$\star$

$$\begin{array}{l} \mathrm{t}^1 = \begin{cases} t^1_{i=1:k} & = x_i \\ t^1_{k+1:n} & = \texttt{s\_linear}(x_i, 0.35) \end{cases} \\ \mathrm{t}^2 = \begin{cases} t^2_{i=1:k} & = t^1_i \\ t^2_{k+1:n} & = \texttt{r\_nonsep}(\{t^1_{k+2(i-k)-1}, \\ & \qquad t^1_{k+2(i-k)}\}, 2) \end{cases} \\ \mathrm{t}^3 = \begin{cases} t^3_{i=1:m-1} & = \texttt{r\_sum}(\{t^2_{(i-1)k/(m-1)+1}, \\ & \qquad \cdots, t^2_{ik/(m-1)}\}, \{1, \cdots, 1\}) \\ t^3_m & = \texttt{r\_sum}(\{t^2_{k+1}, \cdots, t^2_{k+l/2}\}, \\ & \qquad \{1, \cdots, 1\}) \end{cases} \\ \text{shape} = \begin{cases} f_{1:m-1} & = \texttt{convex}_m \\ f_m & = \texttt{disc}_m(A, \alpha, \beta) \end{cases} \end{array}, \tag{7}$$

where the definitions of $\texttt{s\_linear}(\cdot)$, $\texttt{r\_nonsep}(\cdot)$, $\texttt{r\_sum}(\cdot)$, $\texttt{convex}_m$ and $\texttt{disc}_m(\cdot)$ can be found in [1].

Note that the A determines the number of disconnected regions of the PF. α controls the overall shape of the PF where $\alpha > 1$, $\alpha < 1$ and $\alpha = 1$ leads to a concave, a convex and a linear PF, respectively. β influences the location of the disconnected regions. In our experiments, we instantiate 7 test problem instances, the settings used in our experiments are given in Table **??**. Fig. **??** gives the illustrative examples of their PFs. The number of objectives is set to $m = 2$ for the ZDT and $m = 3$ for the DTLZ problems. As for the WFG problems, we consider both 2- and 3-objective cases. The number of variables is set as $n \in \{5, 10, 20, 30\}$ respectively for each benchmark test problem. In total, there are 136 test problem instances considered in our experiments.

References

1. S. Huband, P. Hingston, L. Barone, and R. L. While, "A review of multiobjective test problems and a scalable test problem toolkit," *IEEE Trans. Evol. Comput.*, vol. 10, no. 5, pp. 477–506, 2006.