

Multi-Objective Bayesian Optimization with Active Preference Learning

Ryota Ozaki¹, Kazuki Ishikawa¹, Youhei Kanzaki¹,
Shion Takeno³, Ichiro Takeuchi^{2, 3}, Masayuki Karasuyama¹

¹Nagoya Institute of Technology

²Nagoya University

³RIKEN AIP

ozaki.ryota.mllab.nit@gmail.com, ishikawa.kazuki.mllab.nit@gmail.com, kanzaki.youhei.mllab.nit@gmail.com,
shion.takeno@riken.jp, takeuchi.ichiro@mae.nagoya-u.ac.jp, karasuyama@nitech.ac.jp

Abstract

There are a lot of real-world black-box optimization problems that need to optimize multiple criteria simultaneously. However, in a multi-objective optimization (MOO) problem, identifying the whole Pareto front requires the prohibitive search cost, while in many practical scenarios, the decision maker (DM) only needs a specific solution among the set of the Pareto optimal solutions. We propose a Bayesian optimization (BO) approach to identifying the most preferred solution in the MOO with expensive objective functions, in which a Bayesian preference model of the DM is adaptively estimated by an interactive manner based on the two types of supervisions called the pairwise preference and improvement request. To explore the most preferred solution, we define an acquisition function in which the uncertainty both in the objective function and the DM preference is incorporated. Further, to minimize the interaction cost with the DM, we also propose an active learning strategy for the preference estimation. We empirically demonstrate the effectiveness of our proposed method through the benchmark function optimization and the hyper-parameter optimization problems for machine learning models.

1 Introduction

In many real-world problems, simultaneously optimizing multiple expensive black-box functions $f_1(\mathbf{x}), \dots, f_L(\mathbf{x})$ are often required. For example, considering an experimental design for developing novel drugs, there can be several criteria to evaluate drug performance such as the efficacy of the drug and the strength of side effects. If we pursue high efficacy, the strength of side effects usually tends to deteriorate. Another example is in the hyper-parameter optimization of a machine learning model, in which multiple different criteria such as accuracy on different classes, computational efficiency, and social fairness should be simultaneously optimized.

In general, multi-objective optimization (MOO) problems have multiple optimal solutions (Figure 1 (a)), called Pareto optimal solutions, and MOO solvers typically aim to find all of the Pareto optimal solutions. However, the search cost of this approach often becomes prohibitive because the number of Pareto optimal solutions can be large even with small L .

On the other hand, in many practical scenarios, a decision maker (DM) only needs one of the optimal solutions that matches their demands (e.g., a drug developer may choose just one drug design considering the best balance between the efficacy and side effects). Therefore, if we can incorporate the DM's preference at the MOO stage, preferred solutions for the DM can be efficiently identified without enumerating all Pareto solutions. However, for the DM, directly defining their preference mathematically can be difficult.

We propose a Bayesian optimization (BO) method for the preference-based MOO, optimizing \mathbf{x} through an interactive (human-in-the-loop based) estimation of the preference based on weak-supervisions provided by the DM. Let $U : \mathbb{R}^L \rightarrow \mathbb{R}$ be a utility function that quantifies the DM preference. We assume that when the DM prefers $\mathbf{f} \in \mathbb{R}^L$ to $\mathbf{f}' \in \mathbb{R}^L$, the utility function values satisfy $U(\mathbf{f}) > U(\mathbf{f}')$. By using U , the optimization problem can be formulated as $\max_{\mathbf{x}} U(\mathbf{f}(\mathbf{x}))$, where $\mathbf{f}(\mathbf{x}) = (f_1(\mathbf{x}), \dots, f_L(\mathbf{x}))^\top$. Our proposed method is based on a Bayesian modeling of $\mathbf{f}(\mathbf{x})$ and U , by which uncertainty of both the objective functions and the DM preference can be incorporated. For $\mathbf{f}(\mathbf{x})$, we use the Gaussian process (GP) regression, following the standard convention of BO. For U , we employ a Chebyshev scalarization function (CSF) based parametrized utility function because of its simplicity and capability of identifying any Pareto optimal points depending on a setting of the preference parameter $\mathbf{w} \in \mathbb{R}^L$ (Figure 1 (b)).

To estimate utility function U (i.e., to estimate the parameter \mathbf{w}), we consider two types of weak supervisions provided by the interaction with the DM. First, we use the pairwise comparison (PC) over two vectors $\mathbf{f}, \mathbf{f}' \in \mathbb{R}^L$. In this supervision, the DM answers whether $U(\mathbf{f}) > U(\mathbf{f}')$ holds based on their preference. In the context of preference learning (e.g., Chu and Ghahramani 2005), this first type of supervision is widely known that the relative comparison is often much easier to be provided by the DM than the exact value of $U(\mathbf{f})$. As the second preference information, we propose to use improvement request (IR) for a given $\mathbf{f} \in \mathbb{R}^L$. The DM provides the index ℓ for which the DM hopes that the ℓ -th dimension of \mathbf{f} , i.e., f_ℓ , is required to improve most among all the L dimensions. To our knowledge, this way of supervision has never been studied in preference learning nevertheless it is obviously easy to provide and important information for the DM. We show that IR can be formulated

as a weak supervision of the gradient of the utility function.

We show that the well-known expected improvement (EI) acquisition function of BO can be defined for the optimization problem $\max_{\mathbf{x}} U(\mathbf{f}(\mathbf{x}))$. Here, the expectation is taken over the both of U and $\mathbf{f}(\mathbf{x})$, by which exploration is performed based on the current uncertainty for both of them. Further, to reduce querying cost to the DM, we also propose active learning that selects effective queries (PC or IR) to estimate the preference parameter \mathbf{w} . Our contributions can be summarized as follows:

- We propose a preference-based MOO algorithm by combining BO and preference learning in which both the MOO objective $\mathbf{f}(\mathbf{x})$ and the DM preference (utility function) are modeled by a Bayesian manner.
- Bayesian preference learning for the utility function is proposed based on two types of weak supervisions, i.e., PC and IR. In particular, IR is a novel paradigm of preference learning.
- Active learning for PC and IR is also proposed. Our approach is based on BALD (Bayesian Active Learning by Disagreement) (Houlsby et al. 2011), which uses mutual information to measure the benefit of querying.
- Numerical experiments on several benchmark functions and hyperparameter optimization of machine learning models show superior performance of our framework compared with baselines such as an MOO extension of BO without preference information.

2 Problem Setting

We consider a multi-objective optimization (MOO) problem that maximizes L objective functions. Let $f_\ell(\mathbf{x})$ for $\ell \in [L]$ be a set of objective functions and $\mathbf{f}(\mathbf{x}) = (f_1(\mathbf{x}), \dots, f_L(\mathbf{x}))^\top$ be their vector representation, where $\mathbf{x} \in \mathbb{R}^d$ is an input vector. In general, an MOO problem can have multiple optimal solutions, called the Pareto optimal points. For example, in Figure 1 (a), all the green points are optimal points. See MOO literature (e.g., Giagkiozis and Fleming 2015) for the detailed definition of the Pareto optimality.

Since the number of the Pareto optimal points can be large even with small L (e.g., Ishibuchi, Tsukamoto, and Nojima 2008), identifying all the Pareto optimal points can be computationally prohibitive. Instead, we consider optimizing a utility function $U : \mathbb{R}^L \rightarrow \mathbb{R}$ that satisfies $U(\mathbf{f}) > U(\mathbf{f}')$ when the DM prefers \mathbf{f} to \mathbf{f}' , where $\mathbf{f}, \mathbf{f}' \in \mathbb{R}^L$. The optimization problem that seeks the solution preferred by the DM is represented as

$$\max_{\mathbf{x}} U(\mathbf{f}(\mathbf{x})). \quad (1)$$

In our problem setting, both the functions $\mathbf{f}(\mathbf{x})$ and $U(\mathbf{f})$ are assumed to be unknown.

About the querying to the objective function $\mathbf{f}(\mathbf{x})$, we follow the standard setting of BO. Let $\mathbf{y}_i = \mathbf{f}(\mathbf{x}_i) + \varepsilon$ be a noisy observation of $\mathbf{f}(\mathbf{x}_i)$ where $\varepsilon \sim \mathcal{N}(0, \sigma^2 I)$ is an independent noise term with the variance σ^2 and the identity matrix I . Since observing \mathbf{y}_i requires high observation cost, a sample efficient optimization strategy is required.

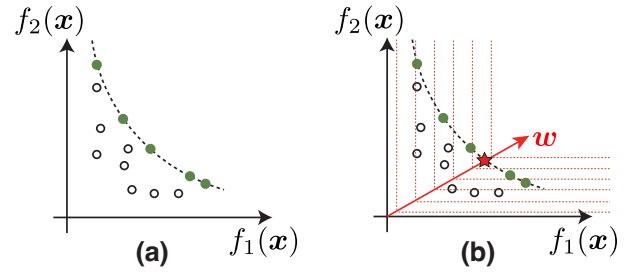


Figure 1: (a) Illustration of Pareto optimal points. The green points are Pareto optimal because, from each one of the green points, no points can improve both objectives simultaneously. The dashed line is an underlying set of all the Pareto optimal points, called the Pareto front. (b) Illustration of CSF (see § 3.2 for detail of CSF). The direction of the red arrow corresponds to \mathbf{w} . The red dashed lines are the contour lines of CSF, by which the red star becomes the optimal point.

On the other hand, directly observing the value of the utility function $U(\mathbf{f}(\mathbf{x}))$ is usually difficult because of difficulty in defining a numerical score of the DM preference. Instead, we assume that we can query to the DM about the following two types of weak supervisions:

Pairwise Comparison (PC): PC indicates the relative preference over given two \mathbf{f} and \mathbf{f}' , i.e., the DM provides whether \mathbf{f} is preferred than \mathbf{f}' or not.

Improvement Request (IR): IR indicates the dimension ℓ in a given \mathbf{f} that the DM considers improvement is required most among $\ell \in [L]$.

For the DM, these two types of information are much easier to provide than the exact value of $U(\mathbf{f})$ itself. PC is a well-known format of a supervision in the context of preference learning (Chu and Ghahramani 2005) and dueling bandit (Sui et al. 2018). On the other hand, it should be noted that IR has not been studied in these contexts to our knowledge, though it considers a practically quite common scenario.

3 Proposed Method

In this section, we first describe the modeling of the MOO objective functions $\mathbf{f}(\mathbf{x})$ in § 3.1. Next, in § 3.2, the modeling of the utility functions U that represents the DM preference is described. In § 3.3, we show the acquisition functions for BO of $U(\mathbf{f}(\mathbf{x}))$ and active learning of U . Then, the entire algorithm of the proposed method is shown in § 3.4. Finally, we discuss a variation of the utility function in § 3.5. Due to the space limitation, some technical details are omitted. See our extended pre-print (Ozaki et al. 2023) for further details.

3.1 Gaussian Process for Objective Functions

As a surrogate model of $\mathbf{f}(\mathbf{x})$, we employ the Gaussian process (GP) regression. For simplicity, the L dimensional-output of $\mathbf{f}(\mathbf{x})$ is modeled by the L independent GPs, each one of which has $k(\mathbf{x}, \mathbf{x}')$ as a kernel function. When we

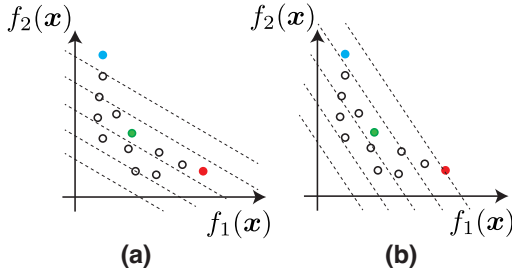


Figure 2: Illustrative examples of contour plots the linear utility function $U(\mathbf{f}) = \sum_{\ell=1}^L w_\ell f_\ell$, where $w_\ell \geq 0$. The plots (a) and (b) have different coefficients. In (a), the blue point is the optimal in a sense of U , while the red point is the optimal in (b). On the other hand, the green point can never be the optimal point by any selection of the coefficients.

observe a set of t observations $D_{\text{GP}} = \{(\mathbf{x}_i, \mathbf{y}_i)\}_{i=1}^t$, the predictive distribution can be obtained as the posterior $p(\mathbf{f}(\mathbf{x}) \mid D_{\text{GP}})$, for which the well-known analytical calculation is available (see e.g., Williams and Rasmussen 2006). Although we here use the independent setting for the L outputs, incorporating correlation among them is also possible by using the multi-output GP (Alvarez et al. 2012).

3.2 Bayesian Modeling for Utility Function

The simplest choice of the utility function is a linear function $U(\mathbf{f}) = \sum_{\ell=1}^L w_\ell f_\ell$ with a weight parameter $w_\ell \geq 0$. However, in a linear utility function, depending of the shape of the Pareto front, it is possible that: 1) there exists a Pareto optimal solution that cannot be identified by any weight parameters, and 2) only one objective function with the highest weight is exclusively optimized (Figure 2). To incorporate the DM preference, variety of scalarization functions have been studied (Bechikh et al. 2015). We here mainly consider the following Chebyshev scalarization function (CSF) (Giagkiozis and Fleming 2015):

$$U(\mathbf{f}(\mathbf{x})) = \min \left(\frac{f_1(\mathbf{x})}{w_1}, \dots, \frac{f_L(\mathbf{x})}{w_L} \right), \quad (2)$$

where $\mathbf{w} = (w_1, \dots, w_L)^\top$ is a preference vector that satisfies $\sum_{\ell \in [L]} w_\ell = 1$ and $w_\ell > 0$. Figure 1 (b) is an illustration of CSF. In the MOO literature (Giagkiozis and Fleming 2015), it is known that for any Pareto optimal solution \mathbf{f}^* , there exist a weighting vector \mathbf{w} under which the maximizer \mathbf{x} of (2) derives $\mathbf{f}^* = \mathbf{f}(\mathbf{x})$. For further discussion on the choice of the utility function, see § 3.5.

Likelihood for Pairwise Comparison We define the likelihood of PC, inspired by the existing work on preference learning of a GP (Chu and Ghahramani 2005). Suppose that the DM prefers $\mathbf{f}^{(i)}$ to $\mathbf{f}^{(i')}$, for which we write $\mathbf{f}^{(i)} \succ \mathbf{f}^{(i')}$. We assume that the preference observation is generated from the underlying true utility U contaminated with the Gaussian noise as follows:

$$\mathbf{f}^{(i)} \succ \mathbf{f}^{(i')} \Leftrightarrow U(\mathbf{f}^{(i)}) + \epsilon_i > U(\mathbf{f}^{(i')}) + \epsilon'_i,$$

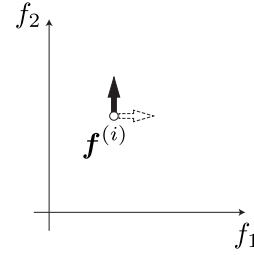


Figure 3: Assume that the DM requests that the direction of f_2 should be improved than f_1 at $\mathbf{f}^{(i)}$. This indicates that the utility function should have larger increase when $\mathbf{f}^{(i)}$ increases along with the axis of f_2 (the direction of the black arrow) compared with the axis of f_1 (the white arrow).

where $\epsilon_i, \epsilon'_i \sim N(0, \sigma_{\text{PC}}^2)$ is the Gaussian noise having the variance σ_{PC}^2 . For a given set of n preference observations $\{\mathbf{f}_i \succ \mathbf{f}'_i\}_{i=1}^n$, the likelihood is written as

$$p(\{\mathbf{f}^{(i)} \succ \mathbf{f}^{(i')}\}_{i=1}^n \mid \mathbf{w}) = \prod_{i=1}^n \Phi \left(\frac{U(\mathbf{f}^{(i)}) - U(\mathbf{f}^{(i')})}{\sqrt{2}\sigma_{\text{PC}}} \right), \quad (3)$$

where Φ is the cumulative distribution function of the standard Gaussian distribution.

Likelihood for Improvement Request For a given $\mathbf{f}^{(i)} = (f_1^{(i)}, \dots, f_L^{(i)})^\top$, if the DM considers $f_{\ell_i}^{(i)}$ has a higher priority to be improved more than other $f_{\ell'_i}^{(i)}$, we write $\ell_i \succ \ell'_i$. In IR, an observation is a dimension ℓ_i for which the DM requires improvement most strongly among L dimensions. This can be considered that we observe $L - 1$ relations $\ell_i \succ \ell'_i$ for $\ell'_i \in [L] \setminus \ell_i$.

The observation $\ell_i \succ \ell'_i$ for $\mathbf{f}^{(i)}$ can be interpreted that the gradient of $U(\mathbf{f}^{(i)})$ with respect to $f_{\ell_i}^{(i)}$ is larger than those of $f_{\ell'_i}^{(i)}$. In other words, the direction of $f_{\ell_i}^{(i)}$ should improve the preference $U(\mathbf{f}^{(i)})$ more rapidly than the direction of $f_{\ell'_i}^{(i)}$ (Figure 3). Let $g_\ell(\mathbf{f}) = \frac{\partial U(\mathbf{f})}{\partial f_\ell}$ be the ℓ -th dimension of the gradient of $U(\mathbf{f})$. Then, the event $\ell_i \succ \ell'_i$ is characterized through the underlying gradient $g_\ell(\mathbf{f})$ as follows:

$$\ell_i \succ \ell'_i \Leftrightarrow g_{\ell_i}(\mathbf{f}^{(i)}) - g_{\ell'_i}(\mathbf{f}^{(i)}) > e_i$$

where $e_i \sim N(0, \sigma_{\text{IR}}^2)$ is the observation noise with the variance σ_{IR}^2 . Suppose that we have m observations $\{\ell_i \succ \ell'_i\}_{i=1}^m$ in total. The likelihood is

$$p(\{\ell_i \succ \ell'_i\}_{i=1}^m \mid \mathbf{w}) = \prod_{i=1}^m \Phi \left(\frac{g_{\ell_i}(\mathbf{f}^{(i)}) - g_{\ell'_i}(\mathbf{f}^{(i)})}{\sqrt{2}\sigma_{\text{IR}}} \right). \quad (4)$$

Prior and Posterior We employ the Dirichlet distribution $p(\mathbf{w}) = \frac{1}{B(\boldsymbol{\alpha})} \prod_{i=1}^L w_i^{\alpha_i - 1}$ as a prior distribution of \mathbf{w} because it has the constraint $\|\mathbf{w}\|_1 = 1$, where B is the beta function and $\boldsymbol{\alpha} = [\alpha_1, \dots, \alpha_L]^\top$ is a parameter. Let D_{pre} be

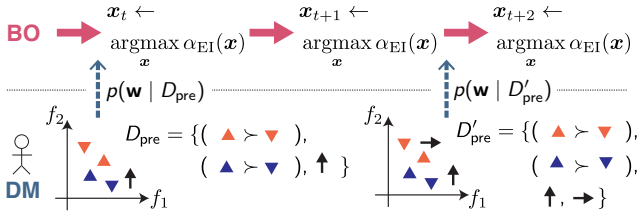


Figure 4: A schematic illustration of the procedure of the proposed framework. $x^{(t)}$ is the selected x at the t -th iteration of BO. The DM can add the preference information at any time point. A pair in a PC observation is represented by the two triangles with the same color (e.g., for the orange-colored \blacktriangle and \blacktriangledown , $\blacktriangle \succ \blacktriangledown$). Each black arrow represents the direction that the DM requests to improve. In D'_{pre} , the DM newly adds the request for the improvement on the horizontal direction.

a set of PC and IR observations. The posterior distribution of w is written as

$$p(w | D_{pre}) \propto p(w) \times p(\{f^{(i)} \succ f^{(i')}\}_{i=1}^n | w) p(\{\ell_j \succ \ell'_j\}_{j=1}^m | w). \quad (5)$$

3.3 Acquisition Functions

Bayesian Optimization of Utility Function Our purpose is to identify x that maximizes $U(f(x))$, for which we employ the well-known expected improvement (EI) criterion of Bayesian optimization. Let $x_{best} = \operatorname{argmax}_{x \in \Theta} U(f(x))$ be the best x in the set Θ consisting of x_i contained in the observed set D_{GP} . The acquisition function that selects the next x is defined by

$$\alpha_{EI}(x) = \mathbb{E}_{f(x), w} [\max \{U(f(x)) - U(f(x_{best})), 0\}]. \quad (6)$$

Unlike the standard EI in BO, the expectation is jointly taken over $f(x)$ and w , and the current best term $U(f(x_{best}))$ is also random variable. Therefore, the analytical calculation of (6) is difficult and we employ the Monte Carlo (MC) method to evaluate (6). The objective function $f(x)$ can be easily sampled because it is represented by the GPs. The parameter of utility function w is sampled from the posterior (5), for which we use the standard Markov chain MC (MCMC) sampling. Note that since both of f and U are represented as Bayesian models, uncertainty with respect to both the objective functions and the user preference are incorporated in our acquisition function.

Active Learning for Utility Function Estimation We also propose an active learning (AL) acquisition function for efficiently estimating the utility function U . Since PC- and IR- observations require interactions with the DM, accurate preference estimation with the minimum observations is desired. Our AL acquisition function is based on the Bayesian active learning framework called BALD (Bayesian Active Learning by Disagreement) (Houlsby et al. 2011). BALD is an information theoretic approach in which the next query

is selected by maximizing mutual information (MI) between an observation and a model parameter.

We here describe the case of PC only because for IR, almost the same procedure is derived. We need to select a pair f and f' that efficiently reduces the uncertainty of w . Let

$$z_{PC} = \begin{cases} 1 & \text{if } f \succ f', \\ 0 & \text{if } f' \succ f \end{cases}$$

be the indicator of the preference given by the DM. Our AL acquisition function is defined by the MI between z_{PC} and w :

$$MI(z_{PC}; w) = H[z_{PC}] - \mathbb{E}_w[H[z_{PC} | w]], \quad (7)$$

where H represents the entropy. Houlsby et al. (2011) clarify that this difference of the entropy representation results in a simpler computation than other equivalent representations of MI.

For the first term of (7), since z_{PC} follows the Bernoulli distribution, we have

$$H[z_{PC}] = \sum_{z_{PC} \in \{1,0\}} p(z_{PC}) \log p(z_{PC}).$$

Unfortunately, $p(z_{PC})$, in which w is marginalized, is difficult to evaluate analytically. On the other hand, the conditional distribution $p(z_{PC} | w)$ is easy to evaluate as shown in (3). Therefore, we employ a sampling based approximation

$$p(z_{PC}) \approx \sum_{w \in \mathcal{W}} p(z_{PC} | w) / |\mathcal{W}|,$$

where \mathcal{W} is a set of w generated from the posterior (5) (e.g., by using the MCMC sampling). For the second term of (7), the same sample set \mathcal{W} can be used as

$$\begin{aligned} & \mathbb{E}_w[H[z_{PC} | w]] \\ & \approx \sum_{w \in \mathcal{W}} \sum_{z_{PC} \in \{1,0\}} \frac{p(z_{PC} | w) \log p(z_{PC} | w)}{|\mathcal{W}|}. \end{aligned}$$

3.4 Algorithm

The procedure of the proposed framework is shown in Figure 4 and Algorithm 1 (MBO-APL: Multi-objective BO with Active Preference Learning). Note that although Algorithm 1 also only considers the case of PC, the procedure for IR is almost the same. BO and the preference learning can be in parallel because the training of the GPs and w are independent. When the DM adds the preference data (PC and/or IR) into D_{pre} , the posterior of w is updated. The updated posterior can be immediately used in the next acquisition function calculation of BO (in Figure 4, for example, $p(w | D'_{pre})$ can be used to determine $x^{(t+2)}$).

3.5 Selection on Utility Function

Although we mainly focus on (2) as a simple example of the utility function, different utility functions can be used in our framework. For example, in MOO literature, the following

Algorithm 1: Proposed Method

```

1: procedure MBO-APL
2:   Run ACTIVE-PREF-LEARNING in background
3:   for  $t = 1, \dots$  do
4:     Fit GPs to  $D_{\text{GP}} = \{(\mathbf{x}_i, \mathbf{y}_i)\}_{i=1}^t$ 
5:      $\mathbf{x}_{t+1} \leftarrow \operatorname{argmax}_{\mathbf{x}} \alpha_{\text{EI}}(\mathbf{x})$ 
       using current  $p(\mathbf{w} \mid D_{\text{pref}})$ 
6:     Observe  $(\mathbf{x}_{t+1}, \mathbf{y}_{t+1})$ 
       and  $D_{\text{GP}} \leftarrow D_{\text{GP}} \cup (\mathbf{x}_{t+1}, \mathbf{y}_{t+1})$ 
7:   end for
8: end procedure
9: procedure ACTIVE-PREF-LEARNING
10:  for  $t = 1, \dots$  do
11:    Update  $p(\mathbf{w} \mid D_{\text{pref}})$  with the current  $D_{\text{pref}}$ 
12:     $\mathbf{f}, \mathbf{f}' \leftarrow \operatorname{argmax}_{\mathbf{f}, \mathbf{f}'}$   $\text{MI}(z_{\text{PC}}; \mathbf{w})$ 
13:    Query  $z_{\text{PC}}$  to the DM and add the result to  $D_{\text{pref}}$ 
14:  end for
15: end procedure

```

augmented CSF is often used (Bechikh et al. 2015; Hakanen and Knowles 2017):

$$U(\mathbf{f}(\mathbf{x})) = \min_{\ell \in [L]} \frac{f_{\ell}(\mathbf{x}) - f_{\ell}^{\text{ref}}}{w_{\ell}} + \rho \sum_{\ell \in [L]} \frac{f_{\ell}(\mathbf{x}) - f_{\ell}^{\text{ref}}}{w_{\ell}},$$

where $\rho > 0$ is an augmentation coefficient (usually a small constant) and f_{ℓ}^{ref} is a reference point. If the DM can directly provide the reference point, f_{ℓ}^{ref} is a fixed constant, while it is also possible that we estimate f_{ℓ}^{ref} as a random variable from PC and IR. The second term avoids weakly Pareto optimal solutions. Using the augmented CSF in our framework is easy because the posterior of \mathbf{w} (and f_{ℓ}^{ref}) can be derived by the same manner described in § 3.2.

Instead of parametric models such as CSF, nonparametric approaches are also applicable to defining the utility function. Since our problem setting is the maximization of $\mathbf{f}(\mathbf{x})$, the utility function U should be monotonically increasing. Therefore, for example, the GP regression with the monotonicity constraint (Riihimäki and Vehtari 2010) can be used to build a utility function with high flexibility. However, because of its high flexibility, the GP model may require a larger number of observations to estimate the DM preference accurately. In this sense, the simple CSF-based and the GP-based approaches are expected to have trade-off relation about their flexibility and complexity of the estimation.

4 Related Work

In the MOO literature, incorporating the DM preferences into exploration algorithms have been studied. For example, a hyper-rectangle or a vector are often used to represent the preference of the DM in the objective function space (e.g., Hakanen and Knowles 2017; Palar et al. 2018; He et al. 2020; Paria, Kandasamy, and Póczos 2020). Another example is that Abdolshah et al. (2019) represent the DM preference by the order of importance for the objective functions. In particular, Paria, Kandasamy, and Póczos (2020) use a similar formulation to our approach in which CSF with

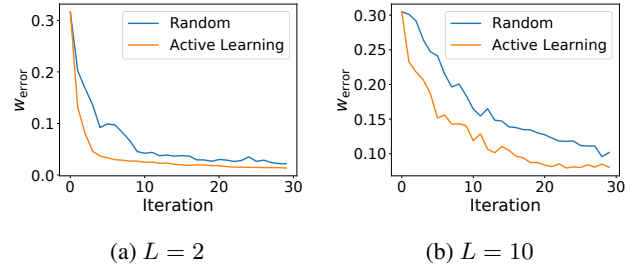


Figure 5: Estimation error of \mathbf{w} .

the random parameters can be used for the utility function. These approaches do not estimate a model of the DM preference, by which the DM needs to directly specify the detailed requirement of the performance though it is often difficult in practice. For example, in the case of the hyper-rectangle, there may not exist any solutions in the specified region by the DM.

The interactive preference estimation with MOO has also been studied mainly in the context of evolutionary algorithms (Hakanen et al. 2016). For example, Taylor et al. (2021) combine preference learning with the multi-objective evolutionary algorithm. On the other hand, to our knowledge, combining an interactively estimated Bayesian preference model and the multi-objective BO has not been studied, though multi-objective extension of BO has been widely studied (e.g., Emmerich 2005; Hernández-Lobato, Hoffman, and Ghahramani 2014). Astudillo and Frazier (2020) and Jerry Lin et al. (2022) consider a different preference based BO in which the DM can have an arbitrary preference over the multi-dimensional output space, meaning that the problem setting is not MOO anymore (in the case of MOO, the preference should be monotonically increasing). Further, these studies do not discuss the IR-type supervision.

5 Experiments

We perform two types of experiments. First, in § 5.1, we evaluate the performance of our MI-based active learning (described in § 3.3) that efficiently learns the preference model (2). Next, in § 5.2, we evaluate the performance of the entire framework of our proposed method, for which we used a benchmark function and two settings of hyper-parameter optimization of machine learning models. In all experiments, the true utility function (the underlying true DM preference) is defined by (2) with the parameter \mathbf{w}_{true} , determined through the sampling from the Dirichlet distribution ($\alpha = (2, \dots, 2)^{\top}$). GPs for $\mathbf{f}(\mathbf{x})$ employs the RBF kernel. Preference observations are generated with the noise variance $\sigma_{\text{PC}} = \sigma_{\text{IR}} = 0.1$. Due to the space limitation, some experimental details are omitted. See our extended pre-print (Ozaki et al. 2023) for further details and additional results.

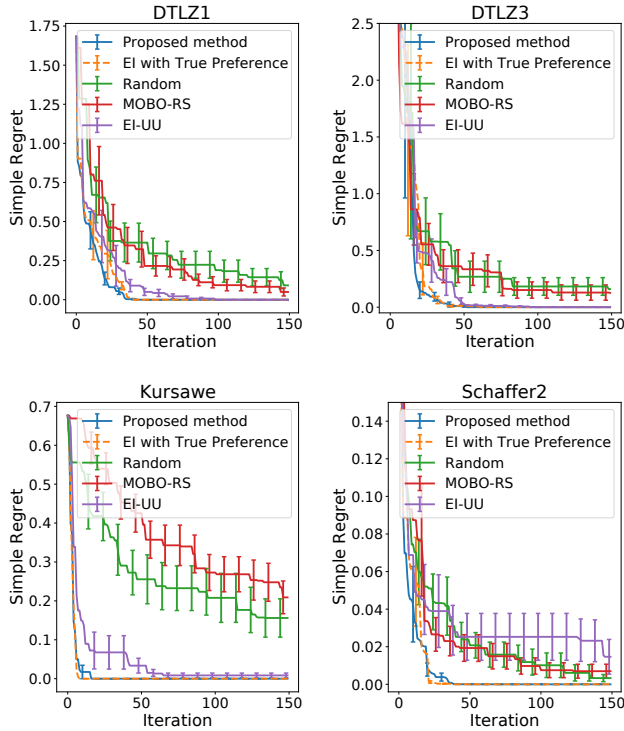


Figure 6: Simple regret on benchmark functions.

5.1 Preference Learning with Active Query Selection

We evaluate estimation accuracy of the preference parameter w . Our MI-based acquisition function and the random query selection were compared. We iteratively added the preference information of the DM. At each iteration, a PC observation and an IR observation are provided.

Figure 5 shows the results. In each plot, the horizontal axis is the iteration and the vertical axis is $w_{\text{error}} = \frac{1}{T} \sum_{t=1}^T \|w_{\text{true}} - w_t\|_2$, where w_1, \dots, w_T are sampled from the posterior ($T = 1000$). The results are the average of 10 runs. We can see that the error rapidly decreases, and further, active learning obviously improves the accuracy. Even when $L = 10$, w_{error} decreased to around 0.1 only with a few tens of iterations. Since w_{error} becomes 0 only when the posterior generates the exact w_{true} T times, we consider $w_{\text{error}} \approx 0.1$ is sufficiently small to accelerate our preference based BO (note that $\|w_t\|_1 = 1$).

5.2 Utility Function Optimization

We evaluate efficiency of the proposed method by evaluating the simple regret on the true utility function $U_{w_{\text{true}}}$:

$$\max_{\mathbf{x}} U_{w_{\text{true}}}(\mathbf{f}(\mathbf{x})) - \max_{\mathbf{x} \in \Theta_s} U_{w_{\text{true}}}(\mathbf{f}(\mathbf{x})),$$

where Θ_s is a set of \mathbf{x} already observed. This evaluates the difference of the utility function values between the true optimal (the first term) and the current best value (the second term). The true parameter w_{true} was determined through the

sampling from the Dirichlet distribution. We assume that a PC- and an IR- observation are obtained when the DM provides the preference information. The results are shown in the average and the standard error of 10 runs. The performance was compared with the three methods: random selection (Random), a random scalarization-based multi-objective BO, called MOBO-RS (Paria, Kandasamy, and Póczos 2020), and multi-attribute BO with preference-based expected improvement acquisition function, called EI-UU (Astudillo and Frazier 2020). In hyper-parameter optimization experiments, we also compare our proposed methods with two other multi-objective BO methods, Expected Hypervolume Improvement (EHI) (Emmerich and Klinkenberg 2008) and Max-value Entropy Search for Multi-objective Bayesian Optimization (MESMO) (Belakaria, Deshwal, and Doppa 2019). Further, we evaluate the performance of EI (6) with the fixed ground-truth preference vector w_{true} , referred to as EI with True Preference (EI-TP). We interpret EI-TP as a best possible baseline for our proposed method, because it has a complete DM preference from the beginning.

We used benchmark functions and two problem settings of hyper-parameter optimization for machine learning models (cost-sensitive learning and fairness-aware machine learning).

Benchmark Function We use well-known MOO benchmark functions, called DTLZ1 and DTLZ3 (Deb et al. 2005). In both the functions, the input and output dimensions are $d = 3$ and $L = 3$. In addition, we use benchmark functions called Kursawe (Kursawe 1990), and Schaffer2 (Schaffer et al. 1985), in which $(d, L) = (3, 2)$ and $(d, L) = (1, 2)$, respectively. We prepare 1000 input candidates by taking grid points in each input dimension. At every iterations, a PC- and an IR- observation are selected by active learning, and the posterior of w is updated.

The results are shown in Figure 6. We see that our proposed method can efficiently reduce simple regret compared with Random, MOBO-RS (which seeks the entire Pareto front), and EI-UU. On the other hand, the proposed method and EI-TP show similar performance, which indicates that the posterior of w provides sufficiently useful information for the efficient exploration.

We further provide results on other benchmark functions, sensitivity evaluation to the number of samplings in EI, and ablation study on PC and IR in Appendix E.2, F.3, and F.1, respectively, in Ozaki et al. (2023).

Hyper-parameter Optimization for Cost-sensitive Learning

In many real-world applications of multi-class classification, a DM may have different preferences over each type of misclassifications (Elkan 2001; Li et al. 2021). For example, in a disease diagnosis, the miss-classification cost for the disease class can be higher than those of the healthy class. In learning algorithms, the cost balance are often controlled by hyper-parameters. In this section, we consider a hyper-parameter optimization problem for L -class classification models considering the importance of each class.

For the hyper-parameter optimization of cost-sensitive learning, we used LightGBM (Ke et al. 2017) and neural net-

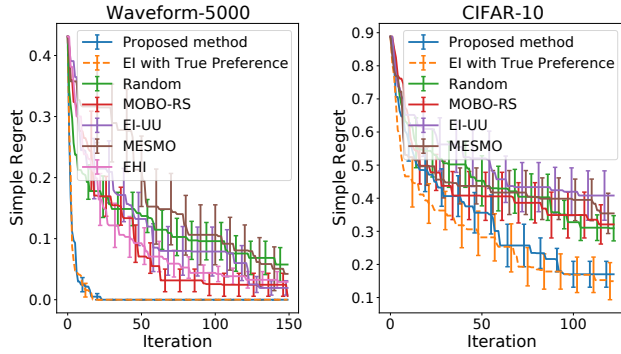


Figure 7: Simple regret on the hyper-parameter optimization for cost-sensitive learning.

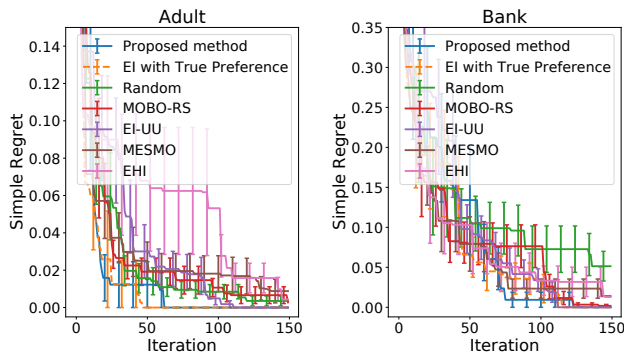


Figure 8: Simple regret on the hyper-parameter optimization for fairness-aware learning.

works. For LightGBM, \mathbf{x} is defined by the L -dimensional ‘class_weight’ parameters, which controls the importance of each class. In the case of neural networks, the weighted cross-entropy loss $\sum_{i=1}^L -\lambda_i y_i \log \hat{y}_i$ was used, where $\lambda_i > 0$ is the class weight, $y_i \in \{0, 1\}$ is one-hot encoding of the label and \hat{y}_i is the corresponding class predicted probability. In this case, \mathbf{x} consists of $\lambda_1, \dots, \lambda_L$. The objective functions $f(\mathbf{x})$ are defined by recall of each class on the validation set (e.g. f_1 represents recall of the first class, f_2 represents recall of the second class, ...). The datasets for the classifiers are Waveform-5000 ($L = 3$) and CIFAR-10 ($L = 10$) (Krizhevsky, Hinton et al. 2009). For Waveform-5000, we used LightGBM. For CIFAR-10, we used Resnet18 (He et al. 2016) pre-trained by Imagenet (Deng et al. 2009). The input dimension (dimension of hyper-parameters) equals to the number of classes, i.e., $d = L$.

Figure 7 shows the results. Overall, we see that the same tendency with the case of benchmark function optimization. The proposed method outperformed MOBO-RS, EHI, MESMO, and EI-UU, and was comparable with EI-TP. In the CIFAR-10 dataset, which has the highest output dimension (10), EI-TP was better than the proposed method. When the output dimension is high, the preference estimation can

be more difficult, but we still obviously see that the proposed method drastically accelerates the exploration compared with searching the entire Pareto front. Note that EHI is not applied to the CIFAR-10 dataset because it is difficult to calculate the acquisition function values due to the high output dimension.

Hyper-parameter Optimization for Fair Classification

As another example of preference-aware hyper-parameter optimization, we consider the fairness in classification problem. Although specific sensitive attributes (e.g. gender, race) should not affect the prediction results for fairness, generally high fairness degrades the classification accuracy (Zafar et al. 2017). Zafar et al. (2017) propose the classification method which maximizes fairness under accuracy constraints, and it has a hyper-parameter named γ to control trade-off between fairness and accuracy.

We consider the hyper-parameter optimization with 2-dimensional objective functions consisting of the level of fairness (p%-rule (Biddle 2006)) and accuracy of the classification. We use the logistic regression classifier proposed in Zafar et al. (2017) and the hyper-parameter γ which control trade-off between fairness and accuracy is equivalent to \mathbf{x} ($d = 1$). For two datasets, Adult (Becker and Kohavi 1996) and Bank (Moro, Rita, and Cortez 2012), we prepare 201 candidates of \mathbf{x} by taking grid points in $[0, 1]$. The data pre-processing and some other settings comply with Zafar et al. (2017).

Figure 8 shows the results. The proposed method finds a better or the best solution in short iteration compared with other methods, especially for the Adult dataset.

6 Conclusion

We proposed a multi-objective Bayesian optimization (BO) method in which the preference of the decision maker (DM) is adaptively estimated through a human-in-the-loop manner. The DM’s preference is represented by a Chebyshev scalarization based utility function, for which we assume that pairwise comparison (PC) and improvement request (IR) are provided as weak supervisions by a decision maker (DM). Our acquisition function is based on the well-known expected improvement by which uncertainty of both the original objective functions and the preference model can be incorporated. We further proposed a mutual information based active learning strategy that reduces the interaction cost with the DM. Empirical evaluation indicated that our proposed method accelerates the optimization on several benchmark functions, and applications to hyper-parameter optimization of machine learning models are also shown.

Acknowledgements

This work was supported by MEXT KAKENHI (20H00601, 21H03498, 22H00300, 23K17817), JST CREST (JPMJCR21D3, JPMJCR22N2), JST Moonshot R&D (JPMJMS2033-05), JST AIP Acceleration Research (JPMJCR21U2), NEDO (JPNP18002, JPNP20006), RIKEN Center for Advanced Intelligence Project, JSPS KAKENHI Grant Number JP21J14673, and JST ACT-X (JPMJAX23CD).

References

- Abdolshah, M.; Shilton, A.; Rana, S.; Gupta, S.; and Venkatesh, S. 2019. Multi-objective Bayesian optimisation with preferences over objectives. In Wallach, H.; Larochelle, H.; Beygelzimer, A.; d'Alché-Buc, F.; Fox, E.; and Garnett, R., eds., *Advances in Neural Information Processing Systems*, volume 32. Curran Associates, Inc.
- Alvarez, M. A.; Rosasco, L.; Lawrence, N. D.; et al. 2012. Kernels for vector-valued functions: A review. *Foundations and Trends® in Machine Learning*, 4(3): 195–266.
- Astudillo, R.; and Frazier, P. 2020. Multi-attribute Bayesian optimization with interactive preference learning. In *International Conference on Artificial Intelligence and Statistics*, 4496–4507. PMLR.
- Bechikh, S.; Kessentini, M.; Said, L. B.; and Gh-dira, K. 2015. Preference Incorporation in Evolutionary Multiobjective Optimization: A Survey of the State-of-the-Art. volume 98 of *Advances in Computers*, 141–207. Elsevier.
- Becker, B.; and Kohavi, R. 1996. Adult. UCI Machine Learning Repository. DOI: <https://doi.org/10.24432/C5XW20>.
- Belakaria, S.; Deshwal, A.; and Doppa, J. R. 2019. Max-value entropy search for multi-objective Bayesian optimization. *Advances in neural information processing systems*, 32.
- Biddle, D. 2006. *Adverse impact and test validation: A practitioner's guide to valid and defensible employment testing*. Gower Publishing, Ltd.
- Chu, W.; and Ghahramani, Z. 2005. Preference learning with Gaussian processes. In *Proceedings of the 22nd international conference on Machine learning*, 137–144.
- Deb, K.; Thiele, L.; Laumanns, M.; and Zitzler, E. 2005. *Scalable test problems for evolutionary multiobjective optimization*. Springer.
- Deng, J.; Dong, W.; Socher, R.; Li, L.-J.; Li, K.; and Fei-Fei, L. 2009. Imagenet: A large-scale hierarchical image database. In *2009 IEEE conference on computer vision and pattern recognition*, 248–255. Ieee.
- Elkan, C. 2001. The Foundations of Cost-Sensitive Learning. In *Proceedings of the 17th International Joint Conference on Artificial Intelligence - Volume 2, IJCAI'01*, 973–978. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc.
- Emmerich, M.; and Klinkenberg, J.-w. 2008. The computation of the expected improvement in dominated hypervolume of Pareto front approximations. *Rapport technique, Leiden University*, 34: 7–3.
- Emmerich, M. T. M. 2005. Single- and Multi-objective Evolutionary Design Optimization Assisted by Gaussian Random Field Metamodels. PhD thesis, FB Informatik, University of Dortmund.
- Giagkiozis, I.; and Fleming, P. J. 2015. Methods for multi-objective optimization: An analysis. *Information Sciences*, 293: 338–350.
- Hakanen, J.; Chugh, T.; Sindhya, K.; Jin, Y.; and Miettinen, K. 2016. Connections of reference vectors and different types of preference information in interactive multiobjective evolutionary algorithms. In *2016 IEEE Symposium Series on Computational Intelligence (SSCI)*, 1–8.
- Hakanen, J.; and Knowles, J. D. 2017. On Using Decision Maker Preferences with ParEGO. In Trautmann, H.; Rudolph, G.; Klamroth, K.; Schütze, O.; Wiecek, M.; Jin, Y.; and Grimme, C., eds., *Evolutionary Multi-Criterion Optimization*, 282–297. Cham: Springer International Publishing.
- He, K.; Zhang, X.; Ren, S.; and Sun, J. 2016. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 770–778.
- He, Y.; Sun, J.; Song, P.; Wang, X.; and Usmani, A. S. 2020. Preference-driven Kriging-based multiobjective optimization method with a novel multipoint infill criterion and application to airfoil shape design. *Aerospace Science and Technology*, 96: 105555.
- Hernández-Lobato, J. M.; Hoffman, M. W.; and Ghahramani, Z. 2014. Predictive Entropy Search for Efficient Global Optimization of Black-box Functions. In *Advances in Neural Information Processing Systems 27*, 918–926. Curran Associates, Inc.
- Houlsby, N.; Huszár, F.; Ghahramani, Z.; and Lengyel, M. 2011. Bayesian Active Learning for Classification and Preference Learning.
- Ishibuchi, H.; Tsukamoto, N.; and Nojima, Y. 2008. Evolutionary many-objective optimization: A short review. In *2008 IEEE Congress on Evolutionary Computation (IEEE World Congress on Computational Intelligence)*, 2419–2426.
- Jerry Lin, Z.; Astudillo, R.; Frazier, P.; and Bakshy, E. 2022. Preference Exploration for Efficient Bayesian Optimization with Multiple Outcomes. In Camps-Valls, G.; Ruiz, F. J. R.; and Valera, I., eds., *Proceedings of The 25th International Conference on Artificial Intelligence and Statistics*, volume 151 of *Proceedings of Machine Learning Research*, 4235–4258. PMLR.
- Ke, G.; Meng, Q.; Finley, T.; Wang, T.; Chen, W.; Ma, W.; Ye, Q.; and Liu, T.-Y. 2017. Lightgbm: A highly efficient gradient boosting decision tree. *Advances in neural information processing systems*, 30.
- Krizhevsky, A.; Hinton, G.; et al. 2009. Learning multiple layers of features from tiny images.
- Kursawe, F. 1990. A variant of evolution strategies for vector optimization. In *International conference on parallel problem solving from nature*, 193–197. Springer.
- Li, M.; Zhang, X.; Thrampoulidis, C.; Chen, J.; and Oymak, S. 2021. AutoBalance: Optimized Loss Functions for Imbalanced Data. *Advances in Neural Information Processing Systems*, 34: 3163–3177.
- Moro, S.; Rita, P.; and Cortez, P. 2012. Bank Marketing. UCI Machine Learning Repository. DOI: <https://doi.org/10.24432/C5K306>.

Ozaki, R.; Ishikawa, K.; Kanzaki, Y.; Suzuki, S.; Takeno, S.; Takeuchi, I.; and Karasuyama, M. 2023. Multi-Objective Bayesian Optimization with Active Preference Learning. arXiv:2311.13460.

Palar, P. S.; Yang, K.; Shimoyama, K.; Emmerich, M.; and Bäck, T. 2018. Multi-Objective Aerodynamic Design with User Preference Using Truncated Expected Hypervolume Improvement. In *Proceedings of the Genetic and Evolutionary Computation Conference, GECCO '18*, 1333–1340. New York, NY, USA: Association for Computing Machinery.

Paria, B.; Kandasamy, K.; and Póczos, B. 2020. A flexible framework for multi-objective bayesian optimization using random scalarizations. In *Uncertainty in Artificial Intelligence*, 766–776. PMLR.

Riihimäki, J.; and Vehtari, A. 2010. Gaussian processes with monotonicity information. In *Proceedings of the thirteenth international conference on artificial intelligence and statistics*, 645–652. JMLR Workshop and Conference Proceedings.

Schaffer, J. D.; et al. 1985. Multiple objective optimization with vector evaluated genetic algorithms. In *Proceedings of an international conference on genetic algorithms and their applications*, 93–100.

Sui, Y.; Zoghi, M.; Hofmann, K.; and Yue, Y. 2018. Advancements in Dueling Bandits. In *Proceedings of the 27th International Joint Conference on Artificial Intelligence*, 5502–5510. AAAI Press.

Taylor, K.; Ha, H.; Li, M.; Chan, J.; and Li, X. 2021. Bayesian Preference Learning for Interactive Multi-Objective Optimisation. In *Proceedings of the Genetic and Evolutionary Computation Conference, GECCO '21*, 466–475. New York, NY, USA: Association for Computing Machinery.

Williams, C. K.; and Rasmussen, C. E. 2006. *Gaussian processes for machine learning*, volume 2. MIT press Cambridge, MA.

Zafar, M. B.; Valera, I.; Ródriguez, M. G.; and Gummadi, K. P. 2017. Fairness constraints: Mechanisms for fair classification. In *Artificial intelligence and statistics*, 962–970. PMLR.