Multi-Attribute Bayesian Optimization under Utility Uncertainty

Raul Astudillo
Cornell University
Ithaca, NY 14853
ra598@cornell.edu

Peter I. Frazier
Cornell University
Ithaca, NY 14853
pf98@cornell.edu

Abstract

We consider multi-attribute Bayesian optimization, where each design in an optimization problem's feasible space is associated with a vector of attributes that can be evaluated via a time-consuming computer code, and each vector of attributes is assigned a utility according to a decision-maker's utility function. A standard Bayesian optimization approach could be applied if the utility function were known to us: we would place a Bayesian prior distribution over the composition of the objective function, which returns a design's vector of attributes, and the utility function, which maps those attributes onto a utility. In contrast, we assume the utility function cannot be evaluated and is known implicitly only to the decision-maker.

We propose a Bayesian optimization algorithm that chooses the designs to evaluate, such that the expected utility of the design chosen by the decision-maker, according to our algorithm's estimate of the objective function, is large. In contrast with existing approaches for multi-attribute optimization that focus on estimating a Pareto frontier, our approach can take advantage of prior information about the decision-maker's utility, obtained from past experiences with the decision-maker or from a utility elicitation process.

1 Introduction

When making decisions using optimization methods, decision-makers often find it challenging to combine multiple competing attributes into a single objective function that quantifies desirability of an arbitrary design, and usually prefer instead to select from several concrete options [10]. This arises, for example, when trading different measures of quality in an information retrieval system [1], or speed against accuracy in the context of machine learning models [8].

One approach is to simply force the decision-maker to select a single-objective function, without any structure or support, and then optimize with respect to this objective. This is often problematic because the objective function chosen may be misaligned with the decision-maker's true preferences. A second approach is to use a utility elicitation technique [3, 5] to find the decision-makers' preferences over attributes. However, such elicitation processes are often limited by time or privacy constraints, and thus substantial uncertainty remains. When the decision-maker's preferences are known to be monotone with respect to the attributes, a third approach is to estimate the Pareto frontier [2, 11]. However, this approach ignores specific prior knowledge of the decision-maker's preferences, such as relative importance of the attributes. Furthermore, estimating the Pareto frontier requires a greater number of time-consuming function evaluations than does optimizing a single objective, especially if the number of attributes is large.

We present a novel approach to this problem by viewing it as a single-objective optimization problem determined by the partially hidden decision-maker's preferences. Specifically, we assume that a probability distribution over the decision-maker's preferences, obtained by a partial utility elicitation

31st Conference on Neural Information Processing Systems (NIPS 2017), Long Beach, CA, USA.

process or information from past experiences in similar contexts, is available. Given this probability distribution, the problem becomes to estimate the attributes for potential solutions in a way that, when they are made available to the decision-maker, the expected utility of the design chosen by the decision-maker is as large as possible.

Our approach is designed for multi-attribute optimization problems in which a design's attribute vector is obtained by a time-consuming computer code or an expensive physical process. We follow a Bayesian optimization framework [4], modeling this mapping from designs to attributes with a multi-output Gaussian process, and using our model of utility uncertainty and a knowledge-gradient approach [17, 7] to derive an acquisition function that we optimize to find each point to sample.

Recently, multi-attribute optimization has been studied by other researchers within the Bayesian optimization framework [13, 2, 11]. To the best of our knowledge, all of this work has focused on estimating the Pareto frontier of the objective function, typically by defining a single-valued function over the space of outputs that captures closeness to the Pareto front; this function, in turn, is used to construct an acquisition function that helps balancing exploration and exploitation.

Our work is also related to [6], who developed a knowledge-gradient method for the pure exploration multi-armed bandit problem with multiple attributes using a similar utility uncertainty approach.

2 Problem formulation

In this section we state our assumptions and formalize the problem to be solved.

We assume that the space of designs is a set $\mathcal{D} \subseteq \mathbb{R}^d$ and attributes are given by an expensive-to-evaluate black-box function $f : \mathcal{D} \to \mathbb{R}^m$. We also assume that there is a decision-maker whose preference over designs is characterized by a design's attributes $\mathbf{y}$ through a utility function, $U(\mathbf{y}) = \theta^\top \mathbf{y}$, i.e., the decision-maker prefers a design $\mathbf{x}$ over $\mathbf{x}'$ if and only if $\theta^\top f(\mathbf{x}) > \theta^\top f(\mathbf{x}')$. Thus, of all the designs, the decision-maker most prefers one in the set $\arg\max_{\mathbf{x}\in\mathcal{D}} \theta^\top f(\mathbf{x})$. If there is uncertainty about the value of $f(\mathbf{x})$ quantified by a Bayesian posterior distribution, then we suppose that the decision-maker selects according to the expected utility, given knowledge of her own utility function, most preferring $\mathrm{argmax}_{\mathbf{x}\in\mathcal{D}} \mathbb{E}\left[\theta^\top f(\mathbf{x}) \mid \theta\right]$, where the expectation is taken over $f(\mathbf{x})$.

If θ were known to us, we could apply a standard single-objective optimization technique to find one with largest $\theta^\top f(\mathbf{x})$. Instead, we assume that θ is unknown, and that the algorithm that samples f has access only to Bayesian prior probability distribution over θ, which may be obtained by a brief and incomplete utility elicitation exercise [3, 5], or by observing past decisions made by the decision-maker. We let $\mathbb{P}_\theta$ represent this prior probability distribution and assume it is concentrated on a set $\Theta \subset \mathbb{R}^m$. Thus, if the decision-maker were asked to choose the best among D designs with known attribute vectors $\mathbf{y}^{(1)}, \ldots, \mathbf{y}^{(D)}$, the expected utility generated would be $\mathbb{E}\left[\max_{i=1,\ldots,D} \theta^\top \mathbf{y}^{(i)}\right]$, where the expectation is over θ.

As is standard in Bayesian optimization [16], we place a Gaussian process prior on f. We use a multi-output Gaussian process $\mathcal{GP}(\mu, K)$ and assume that evaluations of f are either noise-free or of the form $\mathbf{y} = f(\mathbf{x}) + \epsilon(\mathbf{x})$, where $\epsilon(\mathbf{x}) \sim \mathcal{N}(0, \Sigma(\mathbf{x}))$ is independent across evaluations.

After observing N evaluations of f, $\mathbf{y}^{(1)}, \ldots, \mathbf{y}^{(N)}$, at points $\mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(N)}$, respectively, the estimates of the designs' attributes are given by the posterior distribution $\mathcal{GP}\left(\mu^{(N)}, K^{(N)}\right)$. We suppose that, after these N evaluations, the algorithm provides the posterior to the decision maker. As discussed above, the decision maker will then select a design in the set $\mathrm{argmax}_{\mathbf{x}\in\mathcal{D}} \mathbb{E}_N\left[\theta^\top f(\mathbf{x}) \mid \theta\right]$ where $\mathbb{E}_N$ indicates the expectation taken with respect to the posterior given N samples. The expected utility generated, given both θ and the N samples, is

$$\max_{\mathbf{x}\in\mathcal{D}} \mathbb{E}_N\left[\theta^\top f(\mathbf{x}) \mid \theta\right] = \max_{\mathbf{x}\in\mathcal{D}} \theta^\top \mu^{(N)}(\mathbf{x}), \tag{1}$$

Thus, our goal is to design a sampling policy that adaptively chooses points to sample $\mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(N)}$ to maximize the expected value of (1),

$$\mathbb{E}\left[\max_{\mathbf{x}\in\mathcal{D}} \theta^\top \mu^{(N)}(\mathbf{x})\right], \tag{2}$$

where this expectation is taken over both $\mu^{(N)}(\mathbf{x})$ and θ.

3 Sampling policy

While a sampling policy that maximizes (2) can in principle be computed by treating the problem as a partially observable Markov decision process [9, 14], we believe that doing so is computationally intractable in general.

Instead, we propose a sampling policy that approximately maximizes (2) using a one-step assumption, similar to that used to derive knowledge-gradient methods for single-objective Bayesian optimization problems [7]. This policy chooses the next point to sample that would be optimal if it were the last sample we were able to take. Thus, our work generalizes the knowledge-gradient to multi-attribute Bayesian optimization.

At period n of the sampling stage, this one-step optimal point $\mathbf{x}' \in \mathcal{D}$ is the one that maximizes (with ties broken arbitrarily) the expected value that would be received according to (1) if N were equal to $n+1$,

$$\mathbb{E}_n \left[\max_{\mathbf{x} \in \mathcal{D}} \theta^\top \mu^{(n+1)}(\mathbf{x}) \mid \mathbf{x}^{(n+1)} = \mathbf{x}' \right],$$

where the expectation is over θ and the randomness in $\mu^{(n+1)}(\mathbf{x})$ given the time-n posterior on f and our decision to sample $\mathbf{x}'$.

To make parallels to existing work on knowledge-gradient methods [7, 17], we may subtract from this the value that would be received if N were equal to n, to obtain a measure of improvement due to sampling. Since this quantity is the same for all sampling decisions $x^{(n+1)}$, our policy's sampling decision is also the point that maximizes this quantity. We call this the *multi-attribute knowledge-gradient (maKG)* factor,

$$\text{maKG}(\mathbf{x}') = \mathbb{E}_n \left[\max_{\mathbf{x} \in \mathcal{D}} \theta^\top \mu^{(n+1)}(\mathbf{x}) - \max_{\mathbf{x} \in \mathcal{D}} \theta^\top \mu^{(n)}(\mathbf{x}) \mid \mathbf{x}^{(n+1)} = \mathbf{x}' \right]. \tag{3}$$

Note that

$$\text{maKG}(\mathbf{x}') = \mathbb{E}\left[\text{KG}(\mathbf{x}'; \theta)\right], \tag{4}$$

where

$$\text{KG}(\mathbf{x}'; \theta) = \mathbb{E}_n \left[\max_{\mathbf{x} \in \mathcal{D}} \theta^\top \mu^{(n+1)}(\mathbf{x}) - \max_{\mathbf{x} \in \mathcal{D}} \theta^\top \mu^{(n)}(\mathbf{x}) \mid \mathbf{x}^{(n+1)} = \mathbf{x}', \theta \right]. \tag{5}$$

As the notation suggests, $\text{KG}(\mathbf{x}'; \theta)$ is simply the standard knowledge-gradient factor with respect to the function $\theta^\top f(x)$, which, by standard properties of the Gaussian distribution, at period n of the sampling stage is distributed according to a single-output Gaussian process $\mathcal{GP}\left(\theta^\top \mu^{(n)}, \theta^\top K^{(n)} \theta\right)$. In particular, $\text{KG}(\mathbf{x}'; \theta)$ can be computed as in the single-objective case; we refer the reader to [15] for a detailed description. Thus, in general we can compute $\text{maKG}(\mathbf{x}')$ using simple Monte Carlo by first sampling θ from $\mathbb{P}_\theta$ and then computing $\text{KG}(\mathbf{x}'; \theta)$. Furthermore, if the cardinality of Θ is small, $\text{maKG}(\mathbf{x}')$ can also be computed as

$$\text{maKG}(\mathbf{x}') = \sum_{\theta \in \Theta} \mathbb{P}_\theta(\theta) \text{KG}(\mathbf{x}'; \theta). \tag{6}$$

If the cardinality of $\mathcal{D}$ is small, maKG can be maximized through complete enumeration, otherwise, one may employ a stochastic gradient descent approach similar to [17].

We emphasize that a similar analysis allows to generalize other well known policies for single-objective Bayesian optimization, such as expected improvement or predictive entropy search, to the multi-attribute setting. In the next section we present numerical experiments analyzing the performance of our policy and the multi-attribute generalization of expected improvement, which we call *multi-attribute expected improvement (maEI)*.

4 Empirical results

To evaluate the performance of our policy, we perform numerical experiments on the following test problems. Our measure of performance is $\mathbb{E}\left[\max_{\mathbf{x} \in \mathcal{D}} \theta^\top f(\mathbf{x}) \mid f\right]$, where the expectation is over θ and f is the true underlying function.

In all cases we took the Gaussian process model to be independent across attributes, each with a squared exponential kernel. Also, independent Gaussian noise with variance $\sigma^2 = 1$ was added to each measurement for both attributes; the value of σ^2 was not known to our algorithm.

4.1 Synthetic test

In this problem, with results pictured below in Figure 1 (left), we consider the design space, $\mathcal{D} = [-5, 5]^2$, $f : \mathcal{D} \to \mathbb{R}^2$ given by $f_1 =$ Rosenbrock2 and $f_2(\mathbf{x}) = (x_1 - 1)^2 \sin^2(x_2)$, $\Theta = \{(\cos(\alpha_l), \sin(\alpha_l))\}_{l=1}^{10}$, where $\alpha_l = \frac{\pi}{2}\frac{l-1}{9}$, and $\mathbb{P}_\theta$ the uniform distribution over Θ.

4.2 The assemble-to-order benchmark

The assemble-to-order (ATO) benchmark is a reinforcement learning problem from a business application where the goal is to optimize a target level vector; see [12] for a detailed description of the problem. For our experiment we consider only the first two of items and take as attributes the profit and the total number of lost orders. Here, $\Theta = \{(1, -\frac{l}{1000})\}_{l=1}^{5}$, that is, the utility is given by the profit minus a linear penalization on the number of lost orders, and $\mathbb{P}_\theta$ is the uniform distribution over Θ. We picture the results below in Figure 1 (right).

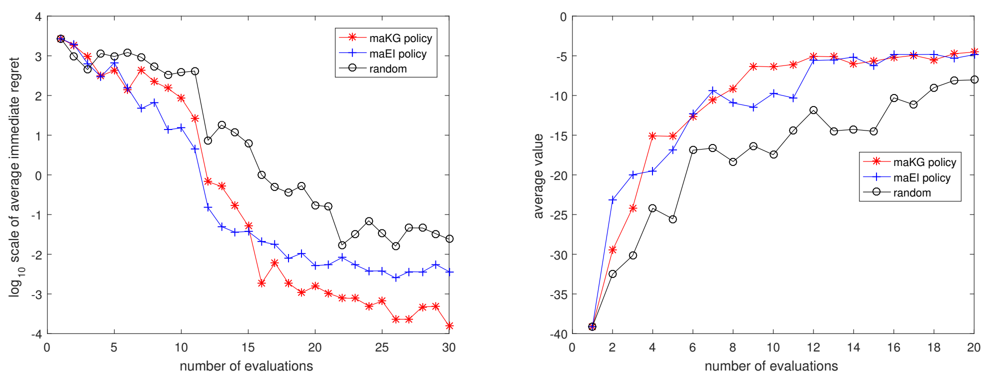

Figure 1: (left) Average performance over 50 replications on the synthetic problem of the maKG policy, the maEI policy and the policy that chooses the designs to sample uniformly at random. (right) Average performance over 25 replications of the aforementioned policies on the ATO benchmark.

5 Conclusion and future work

We introduced a new approach to multi-attribute Bayesian optimization in which the problem is seen as a single-objective optimization problem determined by the underlying decision-maker's preferences. Our formulation leverages prior information on these preferences to improve efficiency during the sampling stage.

One important direction of future work is to extend our analysis to a richer class of utility functions. However, this is challenging because the current scheme to compute the multi-attribute knowledge-gradient factor highly depends on the utility function being linear. Another direction of future work is to analyze the setting in which interaction with the decision-maker occurs during the sampling stage. Specifically, it would be interesting to understand the behavior of our policy under multiple updates of the distribution over the decision-maker's preferences. Finally, we currently assume that the decision-maker can select the best design according to the estimates offered, which is reasonable only if the space of designs is not too large. Thus, it would be relevant to understand how to offer only a small subset of designs such that the expected utility of the decision-maker's choice is still large.

References

[1] A. Al-Maskari, M. Sanderson, and P. Clough. The relationship between ir effectiveness measures and user satisfaction. In *Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval*, pages 773–774. ACM, 2007.

[2] D. C. Bautista. *A sequential design for approximating the pareto front using the expected pareto improvement function*. The Ohio State University, 2009.

[3] C. Boutilier. A pomdp formulation of preference elicitation problems. In *AAAI/IAAI*, pages 239–246, 2002.

[4] E. Brochu, V. M. Cora, and N. De Freitas. A tutorial on bayesian optimization of expensive cost functions, with application to active user modeling and hierarchical reinforcement learning. *arXiv preprint arXiv:1012.2599*, 2010.

[5] U. Chajewska, D. Koller, and R. Parr. Making rational decisions using adaptive utility elicitation. In *AAAI/IAAI*, pages 363–369, 2000.

[6] P. Frazier and A. M. Kazachkov. Guessing preferences: A new approach to multi-attribute ranking and selection. In *Proceedings of the 2011 Winter Simulation Conference (WSC)*, pages 4319–4331. IEEE, 2011.

[7] P. Frazier, W. Powell, and S. Dayanik. The knowledge-gradient policy for correlated normal beliefs. *INFORMS journal on Computing*, 21(4):599–613, 2009.

[8] M. A. Gelbart, J. Snoek, and R. P. Adams. Bayesian optimization with unknown constraints. *arXiv preprint arXiv:1403.5607*, 2014.

[9] D. Ginsbourger and R. Le Riche. Towards gaussian process-based optimization with finite time horizon. *mODa*, 9(96):89–96, 2010.

[10] S. Greco, J. Figueira, and M. Ehrgott. Multiple criteria decision analysis. *Springer's International series*, 2005.

[11] D. Hernandez-Lobato, J. Hernandez-Lobato, A. Shah, and R. Adams. Predictive entropy search for multi-objective bayesian optimization. In *Proceedings of The 33rd International Conference on Machine Learning*, pages 1492–1501, 2016.

[12] L. J. Hong, B. L. Nelson, S. G. Henderson, and J. Xie. `http://simopt.org/wiki/index.php?title=Assemble_to_order`, 2012. Accessed September 25, 2017.

[13] J. Knowles. Parego: A hybrid algorithm with on-line landscape approximation for expensive multiobjective optimization problems. *Transactions on Evolutionary Computation*, 10(1):50–66, 2006.

[14] G. E. Monahan. State of the art—a survey of partially observable markov decision processes: theory, models, and algorithms. *Management Science*, 28(1):1–16, 1982.

[15] W. Scott, P. Frazier, and W. Powell. The correlated knowledge gradient for simulation optimization of continuous parameters using gaussian process regression. *SIAM Journal on Optimization*, 21(3):996–1026, 2011.

[16] B. Shahriari, K. Swersky, Z. Wang, R. P. Adams, and N. de Freitas. Taking the human out of the loop: A review of bayesian optimization. *Proceedings of the IEEE*, 104(1):148–175, 2016.

[17] J. Wu and P. Frazier. The parallel knowledge gradient method for batch bayesian optimization. In *Advances in Neural Information Processing Systems*, pages 3126–3134. Curran Associates, Inc., 2016.