Knowledge Gradient for Selection with Covariates: Consistency and Computation

SHEN Haihui（沈海辉）

Sino-US Global Logistics Institute
Shanghai Jiao Tong University

Joint work with Liang Ding (HKUST), Jeff Hong (Fudan), and Xiaowei Zhang (HKU)

@ 2019 POMS CHINA

June 22, 2019

上海交通大学 SHANGHAI JIAO TONG UNIVERSITY | 董浩云航运与物流研究院 CY TUNG Institute of Maritime and Logistics 中美物流研究院 Sino-US Global Logistics Institute

Contents

1. Introduction
2. Formulation
3. Asymptotics
4. Computation
5. Remarks

1 Introduction

2 Formulation

3 Asymptotics

4 Computation

5 Remarks

Selection of the Best

- Select the best from a finite set of alternatives, whose performances are unknown and can only be learned by sampling.
- The samples may come from computer simulation or real experiments.
- E.g., select the best medicine (treatment), advertisement (recommendation), production line, inventory management, etc.

Selection of the Best

- Select the best from a finite set of alternatives, whose performances are unknown and can only be learned by sampling.
- The samples may come from computer simulation or real experiments.
- E.g., select the best medicine (treatment), advertisement (recommendation), production line, inventory management, etc.
- Sampling may be expensive (in time and/or money), thereby budget-constrained.
- **Goal**: a sampling strategy to learn the performances and identify the best as efficiently as possible.

Selection with Covariates

- In many cases, "the best" is not universal but depends on the covariates (contextual information).
- In the example of personalized medicine, the covariates may be *gender*, *age*, *weight*, *medical history*, *drug reaction*, etc.
- In the example of customized advertisement, the covariates may be *gender*, *age*, *location*, *education*, *browsing history*, etc.

Selection with Covariates

- In many cases, "the best" is not universal but depends on the covariates (contextual information).
- In the example of personalized medicine, the covariates may be *gender*, *age*, *weight*, *medical history*, *drug reaction*, etc.
- In the example of customized advertisement, the covariates may be *gender*, *age*, *location*, *education*, *browsing history*, etc.

- **Goal**: a sampling strategy to learn the performance surfaces (functions) as efficiently as possible.
    - With the learned performance surfaces, we can identify the best given certain covariates.

Knowledge Gradient

- Knowledge Gradient (KG), introduced in Frazier et al. (2008), is a design principle under Bayesian perspective for developing sequential sampling strategy.

Knowledge Gradient

- Knowledge Gradient (KG), introduced in Frazier et al. (2008), is a design principle under Bayesian perspective for developing sequential sampling strategy.

- For selection of the best (*without* covariates):
    - KG-based sampling strategies are widely used;
    - the performance is often competitive with or outperforms other sampling strategies (Ryzhov 2016).

- For selection of the best (*with* covariates):
    - KG-based sampling strategies are emerging (Pearce and Branke 2017);
    - the theory is not completed yet, e.g., no theoretical analysis of the asymptotic behavior of such strategies.

Knowledge Gradient

- Knowledge Gradient (KG), introduced in Frazier et al. (2008), is a design principle under Bayesian perspective for developing sequential sampling strategy.
- For selection of the best (*without* covariates):
  - KG-based sampling strategies are widely used;
  - the performance is often competitive with or outperforms other sampling strategies (Ryzhov 2016).
- For selection of the best (*with* covariates):
  - KG-based sampling strategies are emerging (Pearce and Branke 2017);
  - the theory is not completed yet, e.g., no theoretical analysis of the asymptotic behavior of such strategies.

What We Did?

- In this research, we
  - propose a sampling strategy based on the integrated KG, which is suitable for more general situation;
  - provide a theoretical analysis of the asymptotic behavior of the sampling strategy;
  - propose a stochastic gradient ascent (SGA) algorithm to solve the sampling strategy.

What We Did?

- In this research, we
  - propose a sampling strategy based on the integrated KG, which is suitable for more general situation;
  - provide a theoretical analysis of the asymptotic behavior of the sampling strategy;
  - propose a stochastic gradient ascent (SGA) algorithm to solve the sampling strategy.

	Pearce and Branke (2017)	Our Work
Sampling Noise	homoscedastic	can be heteroscedastic
Sampling Cost	constant	can be different

What We Did?

- In this research, we
  - propose a sampling strategy based on the integrated KG, which is suitable for more general situation;
  - provide a theoretical analysis of the asymptotic behavior of the sampling strategy;
  - propose a stochastic gradient ascent (SGA) algorithm to solve the sampling strategy.

	Pearce and Branke (2017)	Our Work
Sampling Noise	homoscedastic	can be heteroscedastic
Sampling Cost	constant	can be different
Asymptotic Analysis	numerical	theoretical

What We Did?

- In this research, we
  - propose a sampling strategy based on the integrated KG, which is suitable for more general situation;
  - provide a theoretical analysis of the asymptotic behavior of the sampling strategy;
  - propose a stochastic gradient ascent (SGA) algorithm to solve the sampling strategy.

	Pearce and Branke (2017)	Our Work
Sampling Noise	homoscedastic	can be heteroscedastic
Sampling Cost	constant	can be different
Asymptotic Analysis	numerical	theoretical
To Solve	sample average approximation	SGA

Difference with Multi-armed Bandit (MAB)

- Selection of the best vs. MAB problem
- Selection of the best with Covariates vs. Contextual MAB problem

Difference with Multi-armed Bandit (MAB)

- Selection of the best vs. MAB problem
- Selection of the best with Covariates vs. Contextual MAB problem
- Similar settings with different focuses:
  - MAB focuses on minimizing the regret which is caused by choosing inferior alternatives and accumulated during the sampling process;
  - selection of the best focuses on identifying the best alternative eventually when the budget is exhausted.

1 Introduction

2 Formulation

3 Asymptotics

4 Computation

5 Remarks

Setting

- M competing alternatives with *unknown* performance surface $\theta_i(\boldsymbol{x})$, $i = 1, \ldots, M$.
- The covariates $\boldsymbol{x} = (x_1, \ldots, x_d)^\intercal \in \mathcal{X} \subset \mathbb{R}^d$ has density $\gamma(\boldsymbol{x})$.
- We want to learn *offline*: $\operatorname{argmax}_{1 \le i \le M} \theta_i(\boldsymbol{x})$, for $\boldsymbol{x} \in \mathcal{X}$.

Setting

- M competing alternatives with *unknown* performance surface $\theta_i(\boldsymbol{x})$, $i = 1, \ldots, M$.
- The covariates $\boldsymbol{x} = (x_1, \ldots, x_d)^\intercal \in \mathcal{X} \subset \mathbb{R}^d$ has density $\gamma(\boldsymbol{x})$.
- We want to learn *offline*: $\operatorname{argmax}_{1 \le i \le M} \theta_i(\boldsymbol{x})$, for $\boldsymbol{x} \in \mathcal{X}$.

- For simplification purpose, in this presentation we just consider the constant sampling cost ($\equiv 1$), which is not necessary.
- The budget is then N samples.
- Sample on alternative i at location $\boldsymbol{x}$ has independent normal distribution with *unknown* mean $\theta_i(\boldsymbol{x})$ and *known* variance $\lambda_i(\boldsymbol{x})$.

Setting

- M competing alternatives with *unknown* performance surface $\theta_i(\boldsymbol{x})$, $i = 1, \ldots, M$.
- The covariates $\boldsymbol{x} = (x_1, \ldots, x_d)^\intercal \in \mathcal{X} \subset \mathbb{R}^d$ has density $\gamma(\boldsymbol{x})$.
- We want to learn *offline*: $\operatorname{argmax}_{1 \le i \le M} \theta_i(\boldsymbol{x})$, for $\boldsymbol{x} \in \mathcal{X}$.

- For simplification purpose, in this presentation we just consider the constant sampling cost ($\equiv 1$), which is not necessary.
- The budget is then N samples.
- Sample on alternative i at location $\boldsymbol{x}$ has independent normal distribution with *unknown* mean $\theta_i(\boldsymbol{x})$ and *known* variance $\lambda_i(\boldsymbol{x})$.

- We need a good strategy to guide the sampling decision (on *which alternative* and at *what location*) until the N samples are taken.

Bayesian Perspective

- Assign prior for $\{\theta_1(\boldsymbol{x}), \ldots, \theta_M(\boldsymbol{x})\}$, under which $\theta_i(\boldsymbol{x})$'s are independent Gaussian processes with:
  - mean function $\mu_i^0(\boldsymbol{x}) := \mathbb{E}[\theta_i(\boldsymbol{x})|\mathcal{F}^0]$;
  - covariance function $k_i^0(\boldsymbol{x}, \boldsymbol{x}') := \mathrm{Cov}[\theta_i(\boldsymbol{x}), \theta_i(\boldsymbol{x}')|\mathcal{F}^0]$.

Bayesian Perspective

- Assign prior for $\{\theta_1(\boldsymbol{x}), \dots, \theta_M(\boldsymbol{x})\}$, under which $\theta_i(\boldsymbol{x})$'s are independent Gaussian processes with:
  - mean function $\mu_i^0(\boldsymbol{x}) := \mathbb{E}[\theta_i(\boldsymbol{x})|\mathcal{F}^0]$;
  - covariance function $k_i^0(\boldsymbol{x}, \boldsymbol{x}') := \mathrm{Cov}[\theta_i(\boldsymbol{x}), \theta_i(\boldsymbol{x}')|\mathcal{F}^0]$.

- After n samples, $\{\theta_1(\boldsymbol{x}), \dots, \theta_M(\boldsymbol{x})\}$ are still independent Gaussian processes under the posterior with:
  - mean function $\mu_i^n(\boldsymbol{x}) := \mathbb{E}[\theta_i(\boldsymbol{x})|\mathcal{F}^n]$;
  - covariance function $k_i^n(\boldsymbol{x}, \boldsymbol{x}') := \mathrm{Cov}[\theta_i(\boldsymbol{x}), \theta_i(\boldsymbol{x}')|\mathcal{F}^n]$.

Bayesian Perspective

- Assign prior for $\{\theta_1(\boldsymbol{x}), \ldots, \theta_M(\boldsymbol{x})\}$, under which $\theta_i(\boldsymbol{x})$'s are independent Gaussian processes with:
  - mean function $\mu_i^0(\boldsymbol{x}) := \mathbb{E}[\theta_i(\boldsymbol{x})|\mathcal{F}^0]$;
  - covariance function $k_i^0(\boldsymbol{x}, \boldsymbol{x}') := \mathrm{Cov}[\theta_i(\boldsymbol{x}), \theta_i(\boldsymbol{x}')|\mathcal{F}^0]$.

- After n samples, $\{\theta_1(\boldsymbol{x}), \ldots, \theta_M(\boldsymbol{x})\}$ are still independent Gaussian processes under the posterior with:
  - mean function $\mu_i^n(\boldsymbol{x}) := \mathbb{E}[\theta_i(\boldsymbol{x})|\mathcal{F}^n]$;
  - covariance function $k_i^n(\boldsymbol{x}, \boldsymbol{x}') := \mathrm{Cov}[\theta_i(\boldsymbol{x}), \theta_i(\boldsymbol{x}')|\mathcal{F}^n]$.
- $\mu_i^n(\boldsymbol{x})$ is used as our estimator (or predictor) of $\theta_i(\boldsymbol{x})$, and $k_i^n(\boldsymbol{x}, \boldsymbol{x})$ characterizes the uncertainty at $\theta_i(\boldsymbol{x})$.

Bayesian Perspective

- Assign prior for $\{\theta_1(\boldsymbol{x}), \ldots, \theta_M(\boldsymbol{x})\}$, under which $\theta_i(\boldsymbol{x})$'s are independent Gaussian processes with:
  - mean function $\mu_i^0(\boldsymbol{x}) := \mathbb{E}[\theta_i(\boldsymbol{x})|\mathcal{F}^0]$;
  - covariance function $k_i^0(\boldsymbol{x}, \boldsymbol{x}') := \mathrm{Cov}[\theta_i(\boldsymbol{x}), \theta_i(\boldsymbol{x}')|\mathcal{F}^0]$.

- After n samples, $\{\theta_1(\boldsymbol{x}), \ldots, \theta_M(\boldsymbol{x})\}$ are still independent Gaussian processes under the posterior with:
  - mean function $\mu_i^n(\boldsymbol{x}) := \mathbb{E}[\theta_i(\boldsymbol{x})|\mathcal{F}^n]$;
  - covariance function $k_i^n(\boldsymbol{x}, \boldsymbol{x}') := \mathrm{Cov}[\theta_i(\boldsymbol{x}), \theta_i(\boldsymbol{x}')|\mathcal{F}^n]$.
- $\mu_i^n(\boldsymbol{x})$ is used as our estimator (or predictor) of $\theta_i(\boldsymbol{x})$, and $k_i^n(\boldsymbol{x}, \boldsymbol{x})$ characterizes the uncertainty at $\theta_i(\boldsymbol{x})$.

- Updating Equation: if the n-th sample y is taken on i at $\boldsymbol{v}$, then

$$\mu_i^n(\boldsymbol{x}) = \mu_i^{n-1}(\boldsymbol{x}) + k_i^{n-1}(\boldsymbol{x}, \boldsymbol{v})[k_i^{n-1}(\boldsymbol{v}, \boldsymbol{v}) + \lambda_i(\boldsymbol{v})]^{-1}[y - \mu_i^{n-1}(\boldsymbol{v})],$$

$$k_i^n(\boldsymbol{x}, \boldsymbol{x}') = k_i^{n-1}(\boldsymbol{x}, \boldsymbol{x}') - k_i^{n-1}(\boldsymbol{x}, \boldsymbol{v})[k_i^{n-1}(\boldsymbol{v}, \boldsymbol{v}) + \lambda_i(\boldsymbol{v})]^{-1}k_i^{n-1}(\boldsymbol{v}, \boldsymbol{x}').$$

Objective of Sampling Strategy

- After N samples, we will estimate $\text{argmax}_{1\le i\le M}\,\theta_i(\boldsymbol{x})$ via $\text{argmax}_{1\le i\le M}\,\mu_i^N(\boldsymbol{x})$.

Objective of Sampling Strategy

- After N samples, we will estimate $\text{argmax}_{1\le i\le M}\,\theta_i(\boldsymbol{x})$ via $\text{argmax}_{1\le i\le M}\,\mu_i^N(\boldsymbol{x})$.
- View $\max_i \mu_i^N(\boldsymbol{x})$ as a terminal "reward":
  - its expected value depends on the sampling strategy π;
  - maximize the reward $\iff$ minimize the "opportunity cost" $\max_i \theta_i(\boldsymbol{x}) - \max_i \mu_i^N(\boldsymbol{x})$.

Objective of Sampling Strategy

- After N samples, we will estimate $\text{argmax}_{1\le i\le M}\,\theta_i(\boldsymbol{x})$ via $\text{argmax}_{1\le i\le M}\,\mu_i^N(\boldsymbol{x})$.
- View $\max_i \mu_i^N(\boldsymbol{x})$ as a terminal "reward":
  - its expected value depends on the sampling strategy π;
  - maximize the reward $\iff$ minimize the "opportunity cost" $\max_i \theta_i(\boldsymbol{x}) - \max_i \mu_i^N(\boldsymbol{x})$.
- The objective becomes

$$\max_{\pi} \int_{\mathcal{X}} \mathbb{E}^{\pi}\left[\max_{1\le i\le M} \mu_i^N(\boldsymbol{x})\right]\gamma(\boldsymbol{x})\mathrm{d}\boldsymbol{x}.$$

Integrated Knowledge Gradient

- Let $(a^n, \boldsymbol{v}^n)$ denote the n-th sampling decision, i.e., on alternative a^n at location $\boldsymbol{v}^n$, and $S^n := (\mu_1^n, \ldots, \mu_M^n, k_1^n, \ldots, k_M^n)$ the random *state* after the n-th sample.

Integrated Knowledge Gradient

- Let $(a^n, \boldsymbol{v}^n)$ denote the n-th sampling decision, i.e., on alternative a^n at location $\boldsymbol{v}^n$, and $S^n := (\mu_1^n, \ldots, \mu_M^n, k_1^n, \ldots, k_M^n)$ the random *state* after the n-th sample.

- If $N = 1$, the optimal strategy is

$$\underset{1 \le i \le M, \boldsymbol{x} \in \mathcal{X}}{\operatorname{argmax}} \int_{\mathcal{X}} \mathbb{E}\left[\max_{1 \le a \le M} \mu_a^1(\boldsymbol{v}) \,\middle|\, S^0, a^1 = i, \boldsymbol{v}^1 = \boldsymbol{x} \right] \gamma(\boldsymbol{v}) \mathrm{d}\boldsymbol{v}.$$

Integrated Knowledge Gradient

- Let $(a^n, \boldsymbol{v}^n)$ denote the n-th sampling decision, i.e., on alternative a^n at location $\boldsymbol{v}^n$, and $S^n := (\mu_1^n, \ldots, \mu_M^n, k_1^n, \ldots, k_M^n)$ the random *state* after the n-th sample.

- If $N = 1$, the optimal strategy is

$$\underset{1\le i\le M, \boldsymbol{x}\in\mathcal{X}}{\operatorname{argmax}} \int_{\mathcal{X}} \mathbb{E}\left[\max_{1\le a\le M} \mu_a^1(\boldsymbol{v}) \,\middle|\, S^0, a^1 = i, \boldsymbol{v}^1 = \boldsymbol{x} \right] \gamma(\boldsymbol{v}) \mathrm{d}\boldsymbol{v}.$$

- Myopic Strategy: treat each time as if there were only one sample left, and allocate the n-th sample according to

$$\underset{1\le i\le M, \boldsymbol{x}\in\mathcal{X}}{\operatorname{argmax}} \int_{\mathcal{X}} \mathbb{E}\left[\max_{1\le a\le M} \mu_a^n(\boldsymbol{v}) \,\middle|\, S^{n-1}, a^n = i, \boldsymbol{v}^n = \boldsymbol{x} \right] \gamma(\boldsymbol{v}) \mathrm{d}\boldsymbol{v}.$$

Integrated Knowledge Gradient

- The previous myopic strategy is equivalent to maximizing

$$\int_{\mathcal{X}} \mathbb{E}\left[\max_{1\le a\le M} \mu_a^n(\boldsymbol{v}) - \underbrace{\max_{1\le a\le M} \mu_a^{n-1}(\boldsymbol{v})}_{\text{independent of } (i,\, \boldsymbol{x})} \,\Big|\, S^{n-1}, a^n = i, \boldsymbol{v}^n = \boldsymbol{x} \right] \gamma(\boldsymbol{v}) \mathrm{d}\boldsymbol{v}.$$

Integrated Knowledge Gradient

- The previous myopic strategy is equivalent to maximizing

$$\int_{\mathcal{X}} \mathbb{E}\left[\max_{1\le a\le M} \mu_a^n(\boldsymbol{v}) - \underbrace{\max_{1\le a\le M} \mu_a^{n-1}(\boldsymbol{v})}_{\text{independent of } (i,\boldsymbol{x})} \,\Big|\, S^{n-1}, a^n = i, \boldsymbol{v}^n = \boldsymbol{x} \right] \gamma(\boldsymbol{v}) \mathrm{d}\boldsymbol{v}.$$

- It is the integrated expected value of information gained by sampling $(i, \boldsymbol{x})$.

Integrated Knowledge Gradient

- The previous myopic strategy is equivalent to maximizing

$$\int_{\mathcal{X}} \mathbb{E}\left[\max_{1\le a\le M} \mu_a^n(\boldsymbol{v}) - \underbrace{\max_{1\le a\le M} \mu_a^{n-1}(\boldsymbol{v})}_{\text{independent of } (i,\boldsymbol{x})} \,\middle|\, S^{n-1}, a^n = i, \boldsymbol{v}^n = \boldsymbol{x} \right] \gamma(\boldsymbol{v}) \mathrm{d}\boldsymbol{v}.$$

- It is the integrated expected value of information gained by sampling $(i, \boldsymbol{x})$.

- We always search for $(i, \boldsymbol{x})$ that maximizes such integrated expected information gain, thus refer it as "Integrated Knowledge Gradient" (IKG) sampling strategy.

1 Introduction

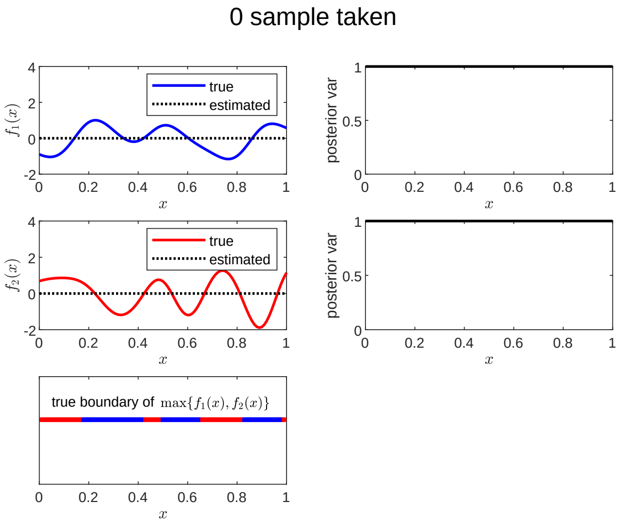
2 Formulation

3 Asymptotics

4 Computation

5 Remarks

Numerical Illustration



Theoretical Result

Theorem 1

Under some mild assumptions, the IKG sampling strategy is consistent, that is, as $N \to \infty$, for all $\boldsymbol{x} \in \mathcal{X}$,

(i) $k_i^N(\boldsymbol{x}, \boldsymbol{x}) \to 0$ *a.s. for* $i = 1, \ldots, M$*;*

(ii) $\mu_i^N(\boldsymbol{x}) \to \theta_i(\boldsymbol{x})$ *a.s. for* $i = 1, \ldots, M$*;*

(iii) $\operatorname{argmax}_i \mu_i^N(\boldsymbol{x}) \to \operatorname{argmax}_i \theta_i(\boldsymbol{x})$ *a.s.*

1. Introduction
2. Formulation
3. Asymptotics
4. Computation
5. Remarks

Compuation

- Recall that to decide the n-th sample, we need to solve

$$\underset{1\le i\le M, \boldsymbol{x}\in\mathcal{X}}{\operatorname{argmax}} \int_{\mathcal{X}} \mathbb{E}\left[\max_{1\le a\le M} \mu_a^n(\boldsymbol{v}) \,\middle|\, S^{n-1}, a^n = i, \boldsymbol{v}^n = \boldsymbol{x}\right] \gamma(\boldsymbol{v}) \mathrm{d}\boldsymbol{v}.$$

Compuation

- Recall that to decide the n-th sample, we need to solve

$$\underset{1\leq i\leq M,\boldsymbol{x}\in\mathcal{X}}{\operatorname{argmax}} \int_{\mathcal{X}} \mathbb{E}\left[\max_{1\leq a\leq M}\mu_a^n(\boldsymbol{v})\,\middle|\, S^{n-1}, a^n=i, \boldsymbol{v}^n=\boldsymbol{x}\right]\gamma(\boldsymbol{v})\mathrm{d}\boldsymbol{v}.$$

- So, we actually need to, for each i, solve

$$\underset{\boldsymbol{x}\in\mathcal{X}}{\operatorname{argmax}} \int_{\mathcal{X}} \mathbb{E}\left[\max_{1\leq a\leq M}\mu_a^n(\boldsymbol{v})\,\middle|\, S^{n-1}, a^n=i, \boldsymbol{v}^n=\boldsymbol{x}\right]\gamma(\boldsymbol{v})\mathrm{d}\boldsymbol{v}.$$

Compuation

- Recall that to decide the n-th sample, we need to solve

$$\underset{1\le i\le M, \boldsymbol{x}\in\mathcal{X}}{\operatorname{argmax}} \int_{\mathcal{X}} \mathbb{E}\left[\max_{1\le a\le M} \mu_a^n(\boldsymbol{v}) \,\middle|\, S^{n-1}, a^n = i, \boldsymbol{v}^n = \boldsymbol{x} \right] \gamma(\boldsymbol{v}) \mathrm{d}\boldsymbol{v}.$$

- So, we actually need to, for each i, solve

$$\underset{\boldsymbol{x}\in\mathcal{X}}{\operatorname{argmax}} \int_{\mathcal{X}} \underbrace{\mathbb{E}\left[\max_{1\le a\le M} \mu_a^n(\boldsymbol{v}) \,\middle|\, S^{n-1}, a^n = i, \boldsymbol{v}^n = \boldsymbol{x} \right]}_{\text{denoted as } h_i^n(\boldsymbol{v}, \boldsymbol{x})} \gamma(\boldsymbol{v}) \mathrm{d}\boldsymbol{v}.$$

Compuation

- Recall that to decide the n-th sample, we need to solve

$$\underset{1\le i\le M, \boldsymbol{x}\in\mathcal{X}}{\operatorname{argmax}} \int_{\mathcal{X}} \mathbb{E}\left[\max_{1\le a\le M} \mu_a^n(\boldsymbol{v}) \,\middle|\, S^{n-1}, a^n = i, \boldsymbol{v}^n = \boldsymbol{x}\right] \gamma(\boldsymbol{v}) \mathrm{d}\boldsymbol{v}.$$

- So, we actually need to, for each i, solve

$$\underset{\boldsymbol{x}\in\mathcal{X}}{\operatorname{argmax}} \int_{\mathcal{X}} \underbrace{\mathbb{E}\left[\max_{1\le a\le M} \mu_a^n(\boldsymbol{v}) \,\middle|\, S^{n-1}, a^n = i, \boldsymbol{v}^n = \boldsymbol{x}\right]}_{\text{denoted as } h_i^n(\boldsymbol{v}, \boldsymbol{x})} \gamma(\boldsymbol{v}) \mathrm{d}\boldsymbol{v}.$$

- $h_i^n(\boldsymbol{v}, \boldsymbol{x})$ can be computed explicitly.
- The computational challenge lies in the numerical integration.

Stochastic Gradient Ascent

- We need to, for each i, solve

$$\operatorname*{argmax}_{\boldsymbol{x}\in\mathcal{X}} \int_{\mathcal{X}} h_i^n(\boldsymbol{v},\boldsymbol{x})\gamma(\boldsymbol{v})\mathrm{d}\boldsymbol{v} = \operatorname*{argmax}_{\boldsymbol{x}\in\mathcal{X}} \mathbb{E}[h_i^n(\boldsymbol{\xi},\boldsymbol{x})],$$

where $\boldsymbol{\xi}$ is a $\mathcal{X}$-valued random variable with density $\gamma(\cdot)$.

Stochastic Gradient Ascent

- We need to, for each i, solve

$$\operatorname*{argmax}_{\boldsymbol{x}\in\mathcal{X}} \int_{\mathcal{X}} h_i^n(\boldsymbol{v},\boldsymbol{x})\gamma(\boldsymbol{v})\mathrm{d}\boldsymbol{v} = \operatorname*{argmax}_{\boldsymbol{x}\in\mathcal{X}} \mathbb{E}[h_i^n(\boldsymbol{\xi},\boldsymbol{x})],$$

where $\boldsymbol{\xi}$ is a $\mathcal{X}$-valued random variable with density $\gamma(\cdot)$.

- To solve such a stochastic optimization problem, the sample average approximation would be computationally prohibitive if $\mathcal{X}$ is high-dimensional.

Stochastic Gradient Ascent

- We need to, for each i, solve

$$\operatorname*{argmax}_{\boldsymbol{x}\in\mathcal{X}} \int_{\mathcal{X}} h_i^n(\boldsymbol{v},\boldsymbol{x})\gamma(\boldsymbol{v})\mathrm{d}\boldsymbol{v} = \operatorname*{argmax}_{\boldsymbol{x}\in\mathcal{X}} \mathbb{E}[h_i^n(\boldsymbol{\xi},\boldsymbol{x})],$$

  where $\boldsymbol{\xi}$ is a $\mathcal{X}$-valued random variable with density $\gamma(\cdot)$.

- To solve such a stochastic optimization problem, the sample average approximation would be computationally prohibitive if $\mathcal{X}$ is high-dimensional.

- Note that $\frac{\partial}{\partial \boldsymbol{x}} h_i^n(\boldsymbol{\xi},\boldsymbol{x})$ can also be computed explicitly here, which is an unbiased estimator of $\nabla_{\boldsymbol{x}} \mathbb{E}[h_i^n(\boldsymbol{\xi},\boldsymbol{x})]$ under mild regularity conditions (L'Ecuyer 1995).

- So we propose to use the stochastic gradient ascent to solve the above stochastic optimization problem.

Algorithm

Algorithm 1 Computing $\log \widehat{\text{IKG}}^n(i, \boldsymbol{x})$.

Inputs: $\mu_1^n, \ldots, \mu_M^n, k_1^n, \ldots, k_M^n, \lambda_1, \ldots, \lambda_M, \boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_J, i, \boldsymbol{x}, c_i(\boldsymbol{x})$
Outputs: *log_IKG*

$\mathcal{J} \leftarrow \emptyset$, *log_IKG* $\leftarrow -\infty$
for $j = 1$ to J **do**
  if $|\tilde{\sigma}_i^n(\boldsymbol{\xi}_j, \boldsymbol{x})| > 0$ **then**
    $u \leftarrow |\Delta_i^n(\boldsymbol{\xi}_j)| / |\tilde{\sigma}_i^n(\boldsymbol{\xi}_j, \boldsymbol{x})|$
    if $u < 20$ **then**
      $r \leftarrow \Phi(-u)/\phi(u)$
    else
      $r \leftarrow u/(u^2 + 1)$
    end if
    $g_j \leftarrow \log\left(\frac{|\tilde{\sigma}_i^n(\boldsymbol{\xi}_j, \boldsymbol{x})|}{\sqrt{2\pi} J}\right) - \frac{1}{2}u^2 + \texttt{log1p}(-ur)$ ▷ $\texttt{log1p}(x) = \log(1 + x)$.
    $\mathcal{J} \leftarrow \{\mathcal{J}, j\}$
  end if
end for
if $\mathcal{J} \neq \emptyset$ **then**
  $g^* \leftarrow \max_{j \in \mathcal{J}} g_j$
  log_IKG $\leftarrow g^* + \log \sum_{j \in \mathcal{J}} e^{g_j - g^*} - \log(c_i(\boldsymbol{x}))$
end if

Algorithm 2 Approximately Computing $(a^n, \boldsymbol{v}^n)$ Using SGA.

Inputs: $\mu_1^n, \ldots, \mu_M^n, k_1^n, \ldots, k_M^n, \lambda_1, \ldots, \lambda_M, \boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_J, c_1, \ldots, c_M$
Outputs: $\hat{a}^n, \hat{\boldsymbol{v}}^n$

for $i = 1$ to M **do**
  $\boldsymbol{x}_1 \leftarrow$ initial value
  for $k = 1$ to K **do**
    Generate independent sample $\{\boldsymbol{\xi}_{k1}, \ldots, \boldsymbol{\xi}_{km}\}$ from density $\gamma(\cdot)$
    $\boldsymbol{x}_{k+1} \leftarrow \Pi_{\mathcal{X}}[\boldsymbol{x}_k + b_k \bar{g}_i^n(\boldsymbol{\xi}_{k1}, \ldots, \boldsymbol{\xi}_{km}, \boldsymbol{x}_k)]$ ▷ Mini-batch SGA
  end for
  $\hat{\boldsymbol{v}}_i^n \leftarrow \frac{1}{K+2-K_0} \sum_{k=K_0}^{K+1} \boldsymbol{x}_k$ ▷ Polyak-Ruppert averaging
  log_IKG$_i \leftarrow \log \widehat{\text{IKG}}^n(i, \hat{\boldsymbol{v}}_i^n)$ ▷ Call Algorithm 1
end for
$\hat{a}^n \leftarrow \arg\max_i$ *log_IKG*$_i$
$\hat{\boldsymbol{v}}^n \leftarrow \hat{\boldsymbol{v}}_{\hat{a}^n}^n$

1. Introduction
2. Formulation
3. Asymptotics
4. Computation
5. Remarks

Concluding Remarks

- We propose an IKG sampling strategy, which is suitable for more general situation.
- We provide a theoretical analysis of the asymptotic behavior of the sampling strategy.
- We propose a SGA algorithm to solve the sampling strategy.

References

- Frazier, P. I., W. Powell, and S. Dayanik (2008). A knowledge gradient policy for sequential information collection. *SIAM J. Control Optim.* 47(5), 2410-2439.
- L'Ecuyer, P. (1995). Note: On the interchange of derivative and expectation for likelihood ratio derivative estimators. *Manag. Sci.* 41(4), 738-747.
- Pearce, M. and J. Branke (2017). Efficient expected improvement estimation for continuous multiple ranking and selection. In *Proc. 2017 Winter Simulation Conf.*, 2161-2172.
- Ryzhov, I. O. (2016). On the convergence rates of expected improvement methods. *Oper. Res.* 64(6), 1515-1528.

Thank you for your attention!

SHEN Haihui
shenhaihui@sjtu.edu.cn

June 22, 2019