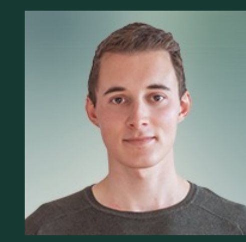
Meta-Learning Hypothesis Spaces for Sequential Decision-making

ETH zürich — Learning & Adaptive Systems

Parnian Kassraie, Jonas Rothfuss, Andreas Krause


Overview

- Obtaining reliable confidence sequences for unknown target functions is a central challenge in sequential decision-making tasks, e.g., Bayesian Optimization and model-based RL.
- These confidence sets are typically constructed by relying on oracle knowledge of the hypothesis space, e.g., a known RKHS.
- We propose any-time valid confidence sets that rely on a meta-learned hypothesis space, instead of assuming oracle knowledge.
- Applied to BO, our results imply a sublinear regret guarantee for the GP-UCB algorithm using our meta-learned kernel. This bound approaches that of the oracle algorithm as the amount of meta-data increases.
- Analysis of the regret that rely on RKHS confidence sets with a known kernel can be immediately extended to use our meta-learned confidence bounds, removing the dependency on the known kernel assumption.

Problem Setting

- Interacting with the environment

$$y_t = f^*(\boldsymbol{x}_t) + \varepsilon_t$$

$\boldsymbol{x}_t \in \mathcal{X}$, depends on the history — $f^*: \mathcal{X} \to R$, $f^* \in \mathcal{H}_{k^*}$, $\|f^*\|_{k^*} \le B$

$\mathcal{X} \subset \mathbb{R}^{d_0}$, compact — k^* unknown

ε_t: σ^2 sub-Gussian, i.i.d.

- Find $\hat{k}$ s.t. the confidence sets are valid

$$\mathbb{P}\left(\forall \boldsymbol{x} \in \mathcal{X},\, \forall t \ge 1:\; f^*(\boldsymbol{x}) \in \mathcal{C}_{t-1}(\hat{k}; \boldsymbol{x})\right) \ge 1 - \delta$$

$$\mathcal{C}_{t-1}(k; \boldsymbol{x}) = [\mu_{t-1}(k; \boldsymbol{x}) \pm \nu_t \sigma_{t-1}(k; \boldsymbol{x})] \qquad (1)$$

$$\mu_{t-1}(k;\boldsymbol{x}) = \boldsymbol{k}_{t-1}^T(\boldsymbol{x})(\boldsymbol{K}_{t-1} + \bar{\sigma}^2 \boldsymbol{I})^{-1}\boldsymbol{y}_{t-1}$$
$$\sigma_{t-1}^2(k;\boldsymbol{x}) = k(\boldsymbol{x},\boldsymbol{x}) - \boldsymbol{k}_{t-1}^T(\boldsymbol{x})(\boldsymbol{K}_{t-1} + \bar{\sigma}^2 \boldsymbol{I})^{-1}\boldsymbol{k}_{t-1}(\boldsymbol{x})$$

Meta-Learning Model

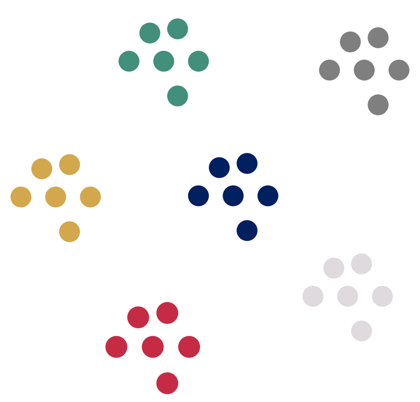
- Data from similar tasks is available (fixed design)

$$y_{s,i} = f_s(\boldsymbol{x}_{s,i}) + \varepsilon_{s,i}$$

$1 \le i \le n$ and $1 \le s \le m$

$\varepsilon_{s,i}$: also σ^2 sub-Gussian, i.i.d.

$f_s: \mathcal{X} \to R$, $f_s \in \mathcal{H}_{k^*}$, $\|f_s\|_{k^*} \le B$



- Assume that true kernel can be decomposed as

$$k^*(\boldsymbol{x}, \boldsymbol{x}') = \sum_{j=1}^{p} \eta_j^* k_j(\boldsymbol{x}, \boldsymbol{x}')$$

η_j^*: unknown, non-negative

k_j: known, finite-dimensional — exists $d_j < \infty$ where $k_j(\boldsymbol{x},\boldsymbol{x}') = \phi^T(\boldsymbol{x})\phi(\boldsymbol{x}')$ and $\phi \in \mathbb{R}^{d_j}$.

$p < \infty$

Meta-Learned Confidence Sets

- Let $\hat{k}$ be the minimizer of

$$\min_{\boldsymbol{\eta}, f_1, \dots, f_m} \frac{1}{m}\sum_{s=1}^{m}\left[\frac{1}{n}\sum_{i=1}^{n}(y_{s,i} - f_s(\boldsymbol{x}_{s,i}))^2\right] + \frac{\lambda}{2}\sum_{s=1}^{m}\|f_s\|_k^2 + \frac{\lambda}{2}\|\boldsymbol{\eta}\|_1$$

$$\text{s.t. } \forall s: f_s \in \mathcal{H}_k,\; k = \sum_{j=1}^{p} \eta_j k_j,\; 0 \le \boldsymbol{\eta} \qquad \text{(META-KEL)}$$

Proposition

Meta-KeL is convex, has a solution and optimizing it is as difficult as the Group Lasso.

$$\hat{k} = \sum_{j=1}^{p} \hat{\eta}_j k_j \qquad \hat{\eta}_j = \left\|\hat{\boldsymbol{\beta}}^{(j)}\right\|_2 \qquad \hat{\boldsymbol{\beta}}^{(j)} = \arg\min_{\boldsymbol{\beta}} \frac{1}{mn}\|\boldsymbol{y} - \boldsymbol{\Phi}\boldsymbol{\beta}\|_2^2 + \lambda \sum_{j=1}^{p}\left\|\boldsymbol{\beta}^{(j)}\right\|_2$$

- Construct the confidence sets (see Equation 1)

$$\mathcal{C}_{t-1}(\hat{k}; \boldsymbol{x}) = [\mu_{t-1}(\hat{k}; \boldsymbol{x}) \pm \nu_t \sigma_{t-1}(\hat{k}; \boldsymbol{x})]$$

$$\nu_t = B\left(1 + \epsilon(n,m)\right) + \sigma\sqrt{\hat{d}\log\left(1 + \bar{\sigma}^{-2}t\right) + 2 + 2\log(1/\delta)}$$

Theorem (Informal)

Under mild regularity assumptions on the meta-data, with probability greater than $1-\delta$,

- *$\hat{k}$ is sparse (in the sense of $\|\eta\|_1$)*
- *$\mathcal{H}_{k^*} \subseteq \mathcal{H}_{\hat{k}}$*
- *For $f \in \mathcal{H}_{k^*}$:*

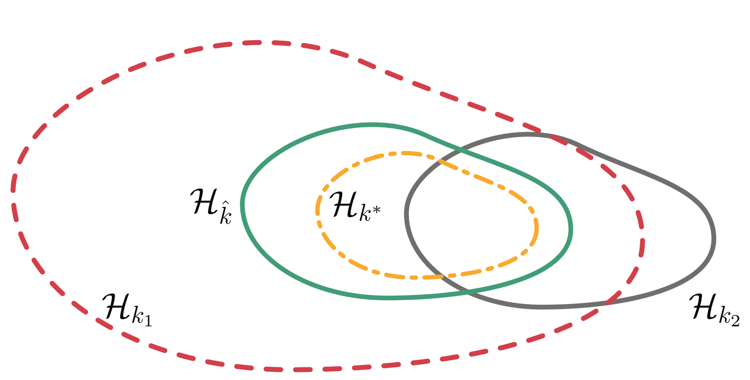
$$\mathbb{P}\left(\forall \mathbf{x} \in \mathcal{X},\, \forall t \ge 1:\; f(\mathbf{x}) \in \mathcal{C}_{t-1}(\hat{k}; \mathbf{x})\right) \ge 1 - \delta.$$

- This theorem implies, with probability greater than $1-\delta$

$$|\mu_{t-1}(\hat{k};\boldsymbol{x}) - f(\boldsymbol{x})| \le \sigma_{t-1}(\hat{k};\boldsymbol{x})\left(B + B\epsilon(n,m) + \sigma\sqrt{\hat{d}\log\left(1 + \bar{\sigma}^{-2}t\right) + 2 + 2\log(1/\delta)}\right)$$

As more meta-data is provided, i.e., as m and n grow, $\epsilon(n,m)$ vanishes and $\hat{d} \to d^*$, the dimension of k^*.

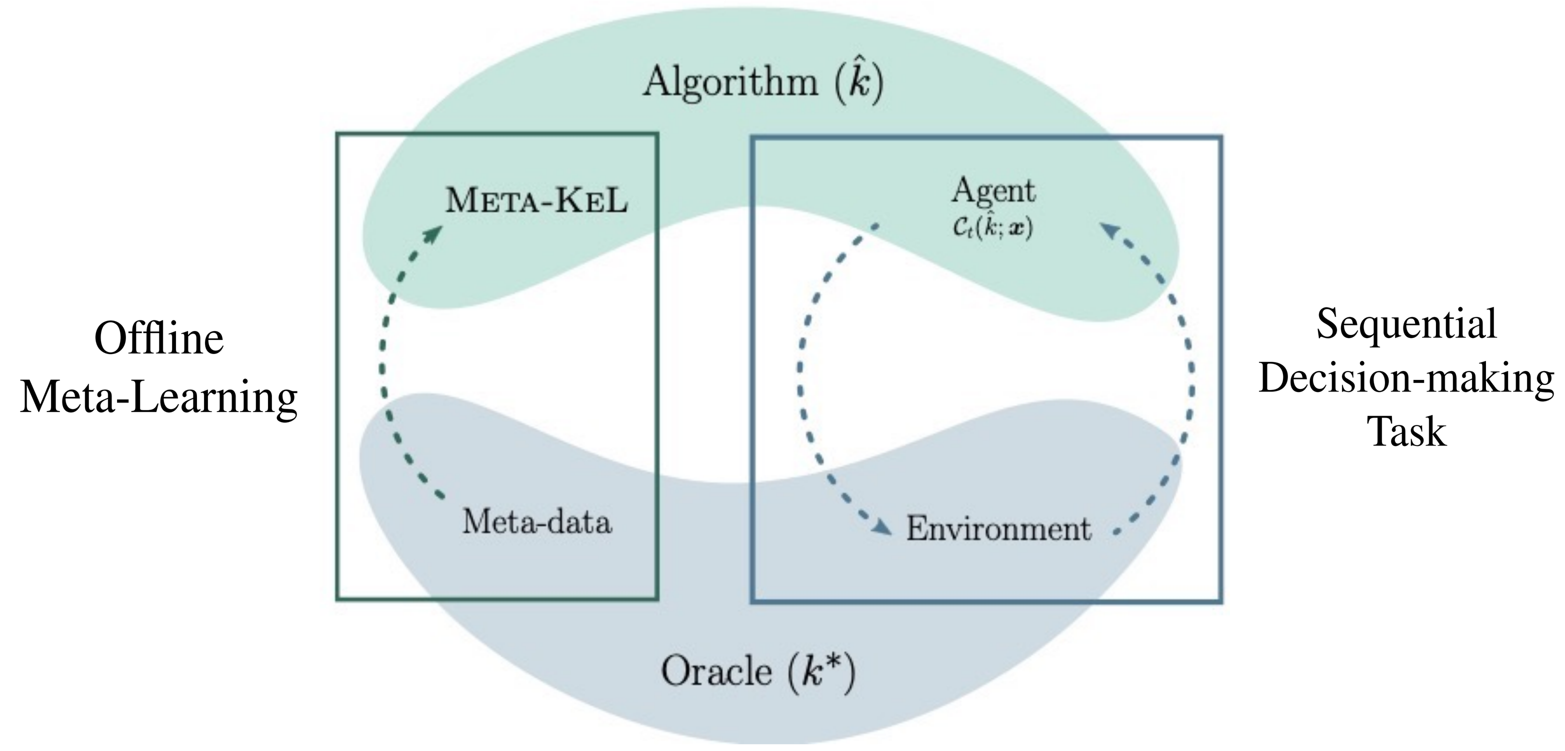
The meta-learned confidence bounds approach the oracle bounds, as amounts of offline data grows.



$\mathcal{C}_{t-1}(k_2; \boldsymbol{x})$	Invalid
$\mathcal{C}_{t-1}(k_1; \boldsymbol{x})$	Valid but too wide
$\mathcal{C}_{t-1}(\hat{k}; \boldsymbol{x})$	Valid and tight
$\mathcal{C}_{t-1}(k^*; \boldsymbol{x})$	True sets (valid)

Sequential Decision-making

- Plug and play



- Applications

Bandits	Safe BO
Bayesian Optimization	Model-Based RL

- Example: f is the objective function of a BO problem.

Regret	$R_T = \sum_{t=1}^{T}[f(\boldsymbol{x}^*) - f(\boldsymbol{x}_t)]$
Goal	$R_T/T \to 0 \quad T \to \infty$
Policy	$\boldsymbol{x}_t = \arg\max_{\boldsymbol{x}\in\mathcal{X}} \mathcal{C}_{t-1}(\hat{k}; \boldsymbol{x})$ [GP-UCB, Srinivas et al.]

Corollary

Provided that there is enough meta-data,

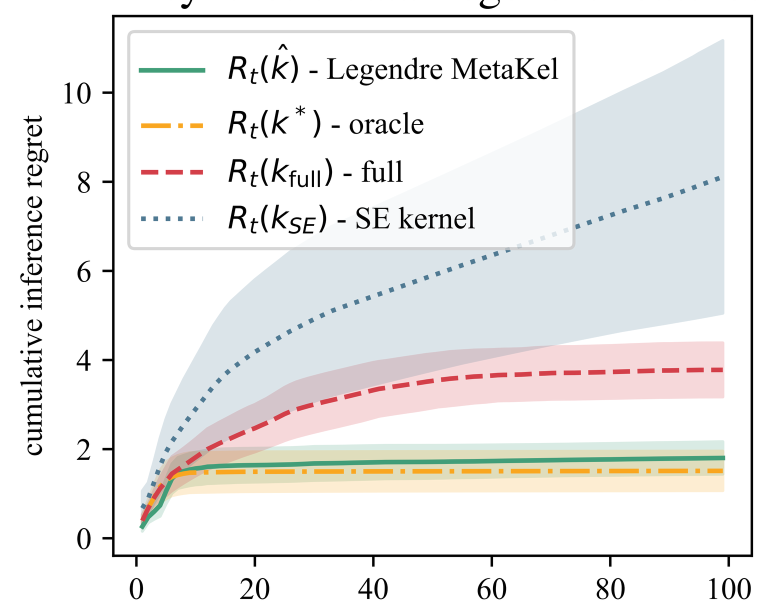
- *The learner achieves sublinear regret, w.h.p.*
- *This guarantee is tight compared to the one for the Oracle learner, and approaches it at a $\mathcal{O}(1/\sqrt{mn})$ rate.*

- This theorem implies, with probability greater than $1-\delta$

$$R_T = \mathcal{O}\left(\sqrt{\hat{d}T\log T}\left(B\left(1 + \epsilon(n,m)\right) + \sqrt{\hat{d}\log T + \log 1/\delta}\right)\right)$$

[Friedman et al 2010]

2D synthetic data Legendre features



GLMNET data + RFF