Learning & Adaptive Systems

Meta-Learning Hypothesis Spaces for Sequential Decision-making

Parnian Kassraie, Jonas Rothfuss, Andreas Krause

ETH zürich

Overview

- Sequential decision-making describes a class of problems where a learner sequentially interacts with an unknown stochastic environment, with the goal of maximizing a reward.
- 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. This is a strong assumption.
- We propose any-time valid confidence sets that rely on a meta-learned hypothesis space, instead of assuming oracle knowledge.

Problem Setting

- Interacting with the environment

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

$\boldsymbol{x}_t \in \mathcal{X}$, depends on the history $(\boldsymbol{x}_{1:t-1}, y_{1:t-1})$

$\mathcal{X} \subset \mathbb{R}^{d_0}$, compact

ε_t: σ^2 sub-Gaussian, i.i.d.

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

k^* unknown

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

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

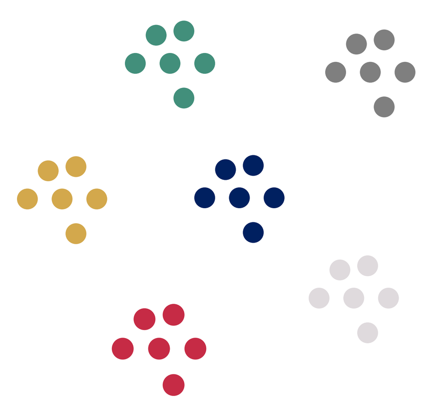
$$\mathcal{C}_{t-1}(k; \boldsymbol{x}) = [\mu_{t-1}(k; \boldsymbol{x}) \pm \nu_t \sigma_{t-1}(k; \boldsymbol{x})] \qquad (1)$$

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

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

- Data from similar tasks is available (fixed design)

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



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

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

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

References

[1] Lounici, K., Pontil, M., Van De Geer, S., and Tsybakov, A. B. Oracle inequalities and optimal inference under group sparsity. *The annals of statistics*, 39(4):2164–2204, 2011.

[2] Chowdhury, S. R. and Gopalan, A. On kernelized multi-armed bandits. In *International Conference on Machine Learning*, pp. 844–853. PMLR, 2017.

Model

- Assume that true kernel can be decomposed as

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

$p < \infty$

η_j^*: unknown, non-negative — Better choice is $\eta_j^* \in \{0, 1\}$.

k_j: known, finite-dimensional — Is extendable to infinite-dimensional kernels.

$\|\boldsymbol{\eta}^*\|_1 \leq 1$ and $k_j(\boldsymbol{x}, \boldsymbol{x}') \leq 1$ — exists $d_j < \infty$ where $k_j(\boldsymbol{x}, \boldsymbol{x}') = \phi^T(\boldsymbol{x})\phi(\boldsymbol{x}')$ and $\phi \in \mathbb{R}^{d_j}$.

Meta-KeL

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

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

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

Proposition

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

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

Theorem (Hypothesis Space Recovery, Informal)

Let $0 < \delta < 1$. Assume meta-data satisfies the Restricted Eigenvalue Assumption and (a relaxed) Beta-min Condition. For the appropriate choice of λ, with probability greater than $1 - \delta$,

$$\mathcal{H}_{k^*} \subseteq \mathcal{H}_{\hat{k}},$$

$$\|f\|_{\hat{k}} \leq \|f\|_{k^*}\Big(1 + \epsilon(n, m) + o(\epsilon(n, m))\Big),$$

where $\epsilon(n, m) = O\left(1/\sqrt{mn}\left(\log p/\delta + \sqrt{md_{\max} \log p/\delta}\right)\right)$.

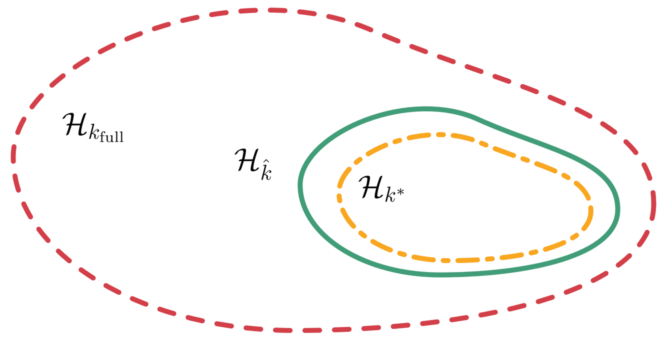
Proposition (Sparsity of $\hat{k}$, Informal)

Let $0 < \delta < 1$. Assume $\boldsymbol{\eta}^*$ is s-sparse. Under assumptions of the theorem above, and for mn large enough, $\hat{\boldsymbol{\eta}}$ is also s-sparse with probability greater than $1 - \delta$.

i.e., the sparsity patterns match with high probability.

$$k_{\text{full}}(\boldsymbol{x}, \boldsymbol{x}') = \frac{1}{p} \sum_{j=1}^{p} k_j(\boldsymbol{x}, \boldsymbol{x}')$$

$$\mathcal{H}_{k^*} \overset{\text{w.h.p.}}{\subseteq} \mathcal{H}_{\hat{k}} \overset{\text{w.h.p.}}{\subsetneq} \mathcal{H}_{k_{\text{full}}}$$



Confidence Sets

- Construct the confidence sets (see Equation 1)

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

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

where $\hat{d}$ is the dimension of $\hat{k}$. i.e., exists $\hat{d} < \infty$ where $\hat{k}(\boldsymbol{x}, \boldsymbol{x}') = \boldsymbol{\phi}^T(\boldsymbol{x})\boldsymbol{\phi}(\boldsymbol{x}')$ where $\boldsymbol{\phi} \in \mathbb{R}^{\hat{d}}$

For calculating μ_{t-1} and σ_{t-1}, set $\bar{\sigma} = 1 + 2/t$.

Theorem (Confidence Bounds with Meta-KeL)

Let $f \in \mathcal{H}_{\boxed{k^*}}$ with $\|f\|_{k^*} \leq B$, where k^* is unknown. Under the assumptions of the previous theorem, with probability greater than $1 - \delta$,

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

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

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

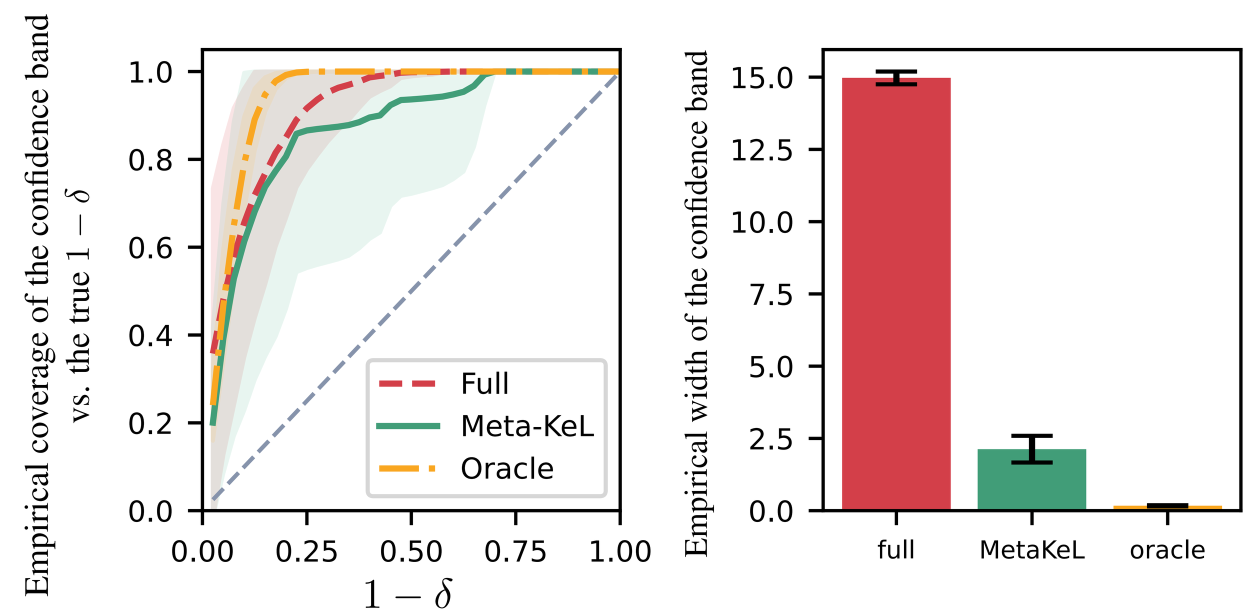
As more meta-data is provided, i.e., as m and n grow, $\epsilon(n, m)$ vanishes and $\hat{d} \to d^*$, the dimension of k^*.

- With oracle knowledge of k^*,

$$|\mu_{t-1}(k^*; \boldsymbol{x}) - f(\boldsymbol{x})| \leq \sigma_{t-1}(k^*; \boldsymbol{x})\Big(B + \sigma\sqrt{\boxed{d^*} \log(1 + \bar{\sigma}^{-2} t) + 2 + 2\log(1/\delta)}\Big)$$

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

- We let $k^*(x, x') = \frac{1}{5} \sum_{j \in J^*} P_j(x) P_j(x')$, where P_j is a degree-j Legendre polynomial and J^* is a random subset of size 5 from $\{1, \cdots, 20\}$. Using this kernel, we generate a random meta-data set of size $m = n = 50$. On a test task, with a fresh new function f^*, we assess the confidence sets at $t = 4$. We check 1) if the sets are valid and 2) if they are tight.



Both quantities are averages over a sample of size 1000 from $\mathcal{X}$. Error bars show standard error for 50 runs of the problem.

- If applied to Bandit optimization, the sets imply a sublinear regret guarantee for the GP-UCB algorithm using $\hat{k}$. This regret bound approaches that of the oracle algorithm.