Centroid Approximation for Bootstrap

Improving Particle Quality at Inference

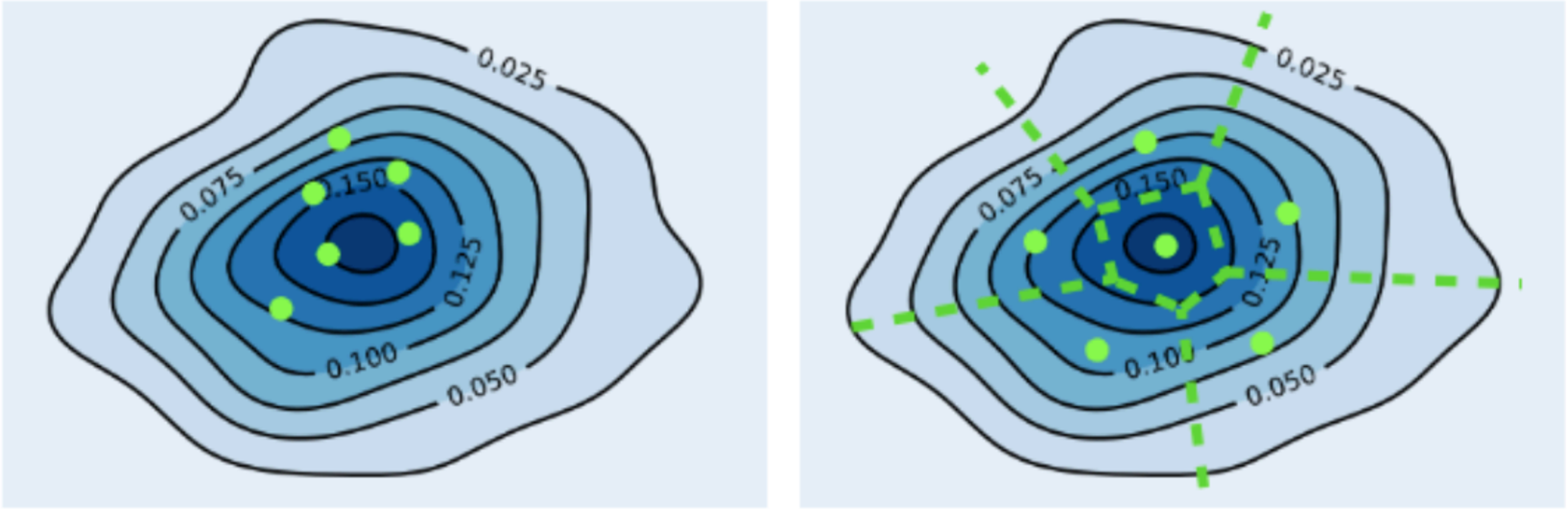
Mao Ye and Qiang Liu

Problem

- Bootstrap is a non-parametric and general method to inference model's data uncertainty.
  - Obtain a sample/parameter from bootstrap distribution by 1. randomly perturb the data weight and 2. estimate the parameter.
  - $\mathcal{L}_{\boldsymbol{w}}(\theta) = \sum_{i=1}^{n} w_i \ell(x_i, f_\theta)/n, \quad \hat{\theta}_{\boldsymbol{w}} = \arg\min_{\theta \in \Theta} \mathcal{L}_{\boldsymbol{w}}(\theta).$
- But requires a lot of i.i.d. samples to approximate the bootstrap uncertainty distribution.
- Hard to be applied to deep learning: too costly at inference time.

Solution

- Actively optimize Bootstrap particle instead of random sample.



- Minimizing Wasserstein distance between the true and particle distribution?
  - Intractable!

Solution

- Minimizing Wasserstein distance between the true and particle distribution.
  - Intractable!

Approximate Wasserstein distance with a tractable loss!

$$\{\theta_j^*\}_{j=1}^m = \arg\min_{\theta_1,\ldots,\theta_m \in \Theta} \mathbb{E}_{\boldsymbol{w}\sim\pi}\left[\min_{j\in[m]} \mathcal{L}_{\boldsymbol{w}}(\theta_j)\right].$$

Asymptotically efficient approximation with $O(\log n/n^{3/2})$ rate.

Experiment

- Improving confidence interval estimation.

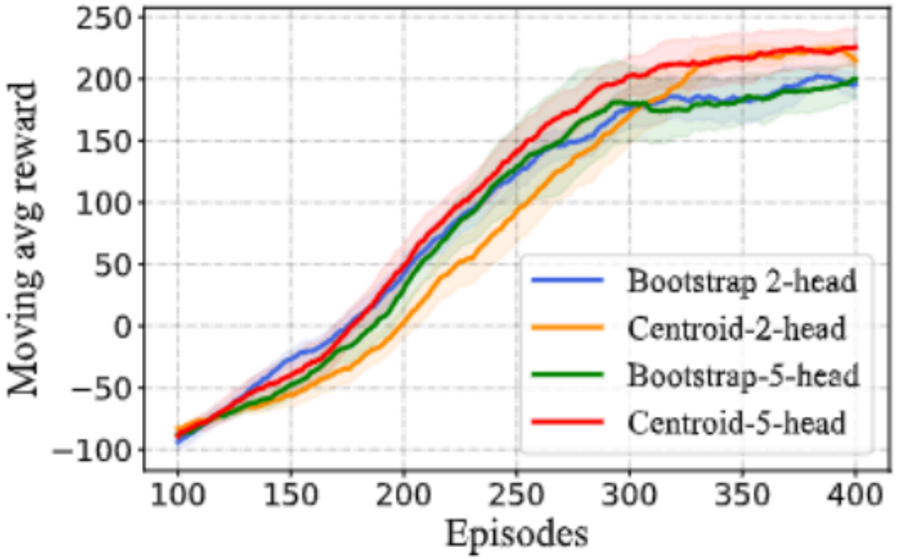
			$m = 20$	$m = 50$	$m = 100$	$m = 200$
$\alpha = 0.9$	Normal	Bootstrap	0.029 ± 0.010	0.031 ± 0.011	0.021 ± 0.010	0.017 ± 0.010
		Centroid	$\mathbf{0.027 \pm 0.010}$	$\mathbf{0.001 \pm 0.009}$	$\mathbf{0.012 \pm 0.010}$	$\mathbf{0.016 \pm 0.010}$
	Percentile	Bootstrap	0.101 ± 0.013	0.036 ± 0.011	0.021 ± 0.010	$\mathbf{0.014 \pm 0.010}$
		Centroid	$\mathbf{0.081 \pm 0.012}$	$\mathbf{0.021 \pm 0.010}$	$\mathbf{0.020 \pm 0.010}$	0.015 ± 0.010
	Pivotal	Bootstrap	0.106 ± 0.013	0.045 ± 0.011	0.025 ± 0.010	0.023 ± 0.010
		Centroid	$\mathbf{0.046 \pm 0.011}$	$\mathbf{0.013 \pm 0.009}$	$\mathbf{0.011 \pm 0.010}$	$\mathbf{0.020 \pm 0.010}$

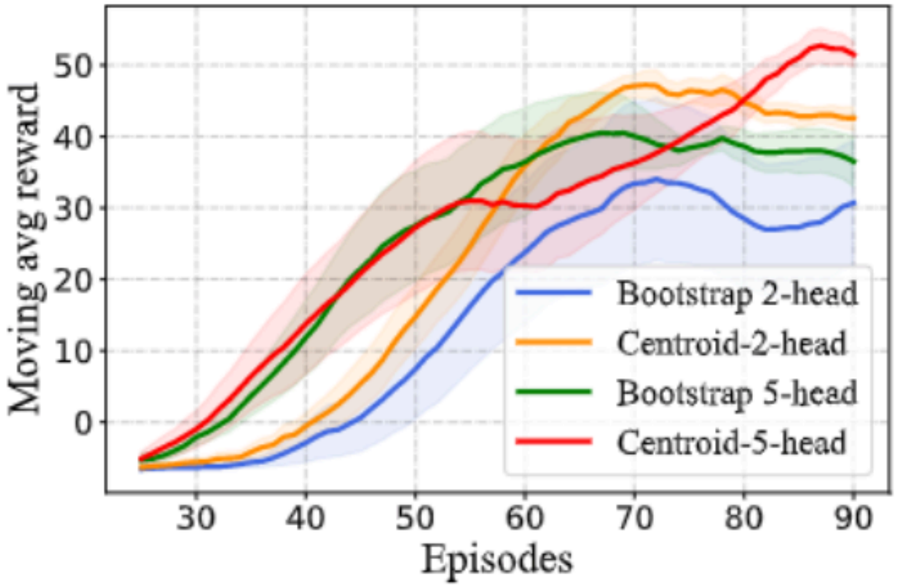
- Improving bootstrap method for contextual bandit.

		$m = 3$	$m = 4$	$m = 5$	$m = 10$
Mushroom	Bootstrap	3282.1 ± 72.8	3307.9 ± 69.2	3311.6 ± 79.3	3397.4 ± 51.4
	Centroid	$\mathbf{3702.7 \pm 89.8}$	$\mathbf{3723.1 \pm 78.7}$	$\mathbf{3799.6 \pm 84.2}$	$\mathbf{3796.9 \pm 36.1}$
Statlog	Bootstrap	1864.3 ± 6.4	1869.2 ± 5.2	1877.2 ± 4.1	1877.0 ± 2.7
	Centroid	$\mathbf{1893.6 \pm 6.0}$	$\mathbf{1892.6 \pm 3.6}$	$\mathbf{1891.3 \pm 3.5}$	$\mathbf{1892.6 \pm 2.8}$
Financial	Bootstrap	2255.8 ± 58.4	2265.4 ± 58.2	2269.3 ± 56.4	2281.4 ± 56.6
	Centroid	$\mathbf{2313.3 \pm 56.4}$	$\mathbf{2315.3 \pm 56.7}$	$\mathbf{2323.9 \pm 56.7}$	$\mathbf{2325.5 \pm 56.0}$

Experiment

- Improving Bootstrap DQN.





Thanks!