On Exploration, Exploitation and Learning in Adaptive Importance Sampling

Xiaoyu Lu[1], Tom Rainforth[1], Yuan Zhou[1], Yee Whye Teh[1], Frank Wood[1], Hongseok Yang[2], Jan-Willem van de Meent[3]

[1]University of Oxford; [2]KAIST, South Korea; [3]Northeastern University

{*xiaoyu.lu, rainforth, y.w.teh*}*@stats.ox.ac.uk, yuan.zhou@cs.ox.ac.uk, fwood@robots.ox.ac.uk, hongseok.yang@kaist.ac.kr, j.vandemeent@northeastern.edu*

Adaptive Importance Sampling (AdaIS)

- Draw samples $x_t \sim q_t(x)$.

Adaptive Importance Sampling (AdaIS)

- Draw samples $x_t \sim q_t(x)$.
- Adapt the proposals $q_t(x)$ according to previous samples $x_1, \cdots x_{t-1}$.

Adaptive Importance Sampling (AdaIS)

- Draw samples $x_t \sim q_t(x)$.
- Adapt the proposals $q_t(x)$ according to previous samples $x_1, \cdots x_{t-1}$.
- Targer π is learnt through a sequence of queries at x_t.

Adaptive Importance Sampling (AdaIS)

- Draw samples $x_t \sim q_t(x)$.
- Adapt the proposals $q_t(x)$ according to previous samples $x_1, \cdots x_{t-1}$.
- Targer π is learnt through a sequence of queries at x_t.

$\rightarrow$ Trade off between **Exploitation** and **Exploration**.

AdaIS as a Bandit Problem

Partition the sample space $\mathcal{X}$ into K disjoint subspaces $\mathcal{X}_1, \cdots, \mathcal{X}_K$.
Proposal:

$$q_t(x) = \sum_{a=1}^{K} q_{at} g_a(x) \tag{1}$$

where q_{at} are the mixing proportions at time t, $g_a(x)$ is the importance proposal on $\mathcal{X}_a$.

- Arm: Subspace
- Reward: Importance weights
- Regret: $R(q_t) = KL(\pi \| q_t) - KL(\pi \| q^*) = \sum_a \pi_a \log \frac{\pi_a}{q_{at}}$

AdaIS as a Bandit Problem

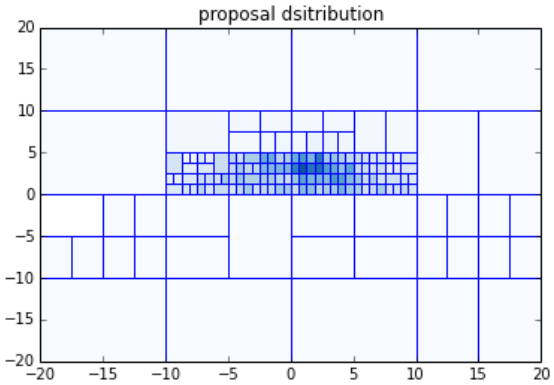
Partition the sample space $\mathcal{X}$ into K disjoint subspaces $\mathcal{X}_1, \cdots, \mathcal{X}_K$.
Proposal:

$$q_t(x) = \sum_{a=1}^{K} \underbrace{q_{at}}_{\substack{\text{boost when} \\ \text{under exploration}}} g_a(x) \tag{2}$$

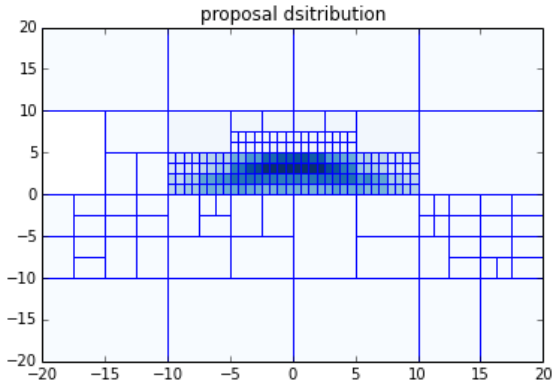
where q_{at} are the mixing proportions at time t, $g_a(x)$ is the importance proposal on $\mathcal{X}_a$.

- Arm: Subspace
- Reward: Importance weights
- Regret: $R(q_t) = KL(\pi \| q_t) - KL(\pi \| q^*) = \sum_a \pi_a \log \frac{\pi_a}{q_{at}}$

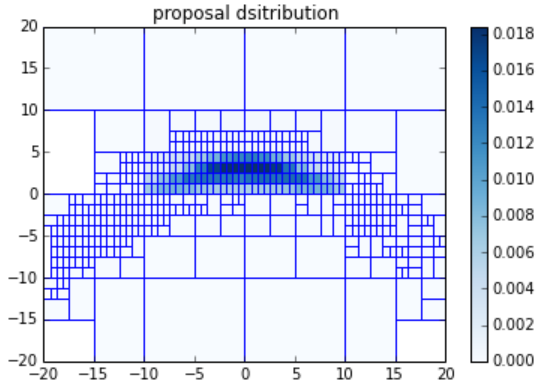
Banana-shaped Example



(a) 1000th iteration



(b) 10000th iteration



(c) 100000th iteration