Advanced Topics in Machine Learning 2020-2021

Yevgeny Seldin Christian Igel Chloé Rouyer Yi-Shan Wu

Home Assignment 1

Deadline: 22:00, Tuesday, 8 September 2020

The assignments must be answered individually - each student must write and submit his/her own solution. We encourage you to work on the assignments on your own, but we do not prevent you from discussing the questions in small groups. If you do so, you are requested to list the group partners in your individual submission.

Submission format: *Please, upload your answers in a single* `.pdf` *file and additional* `.zip` *file with all the code that you used to solve the assignment. (The* `.pdf` *should* ***not*** *be part of the* `.zip` *file.)*

IMPORTANT: *We are interested in how you solve the problems, not in the final answers. Please, write down the calculations and comment your solutions.*

Assignment structure: The assignment contains 5 mandatory questions and 2 optional questions. All questions are designed to help you understand the material, however, the optional questions are not for submission and you can get help with the optional questions at the TA session if needed.

1 Grid World (25 points)

The floorplan in Figure 1 represents a 3×4 grid-world. Each of the 12 rooms corresponds to one state. The agent can go from one room to another if the rooms are connected by a door. The actions are the movements $\{\mathrm{up}, \mathrm{down}, \mathrm{right}, \mathrm{left}\}$. Every movement (except in the terminal state) gives a negative reward of -1 (i.e., every movement costs 1). If you bump into a wall without a door, you stay in the same room, but the movement still has to be paid (i.e., it still gives a reward of -1). There are three rooms that give you an *additional* negative reward of -5 and -10, respectively, *when you enter them*, see figure. The room at the bottom right is a terminal state. An episode ends when this room is reached, which can be modelled by every action in this state leaving the state unchanged and having no cost (i.e., the reward is zero).

This exercise requires an implementation of the MDP. Note that all rewards and transitions are deterministic. They could be described by simple mappings $S \times A \to \mathbb{R}$ and $S \times A \to S$, respectively, which can be encoded by simple lookup tables. Consider a discount factor $\gamma = 1$.

1. Use an algorithm presented in the lecture to compute the value function V^{rand} of the random policy, that is, the policy that chooses an action uniformly at random in every state. If you use an iterative algorithm, initialize all V-values with zero. Report the 12 V^{rand} rand values. Provide an implementation of the algorithm.

2. Use an algorithm presented in the lecture to compute the optimal value function V^*. If you use an iterative algorithm, initialize all V-values with zero. Report the 12 V^* values. Provide an implementation of the algorithm.

2 Numerical comparison of kl inequality with its relaxations and with Hoeffding's inequality (25 points)

Let $X_1, \ldots, X_n$ be a sample of n independent Bernoulli random variables with bias $p = \mathbb{P}\{X = 1\}$. Let $\hat{p}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$ be the empirical average. In this question we make a numerical comparison of the relative power of various bounds on p we have studied. Specifically, we consider the following bounds:

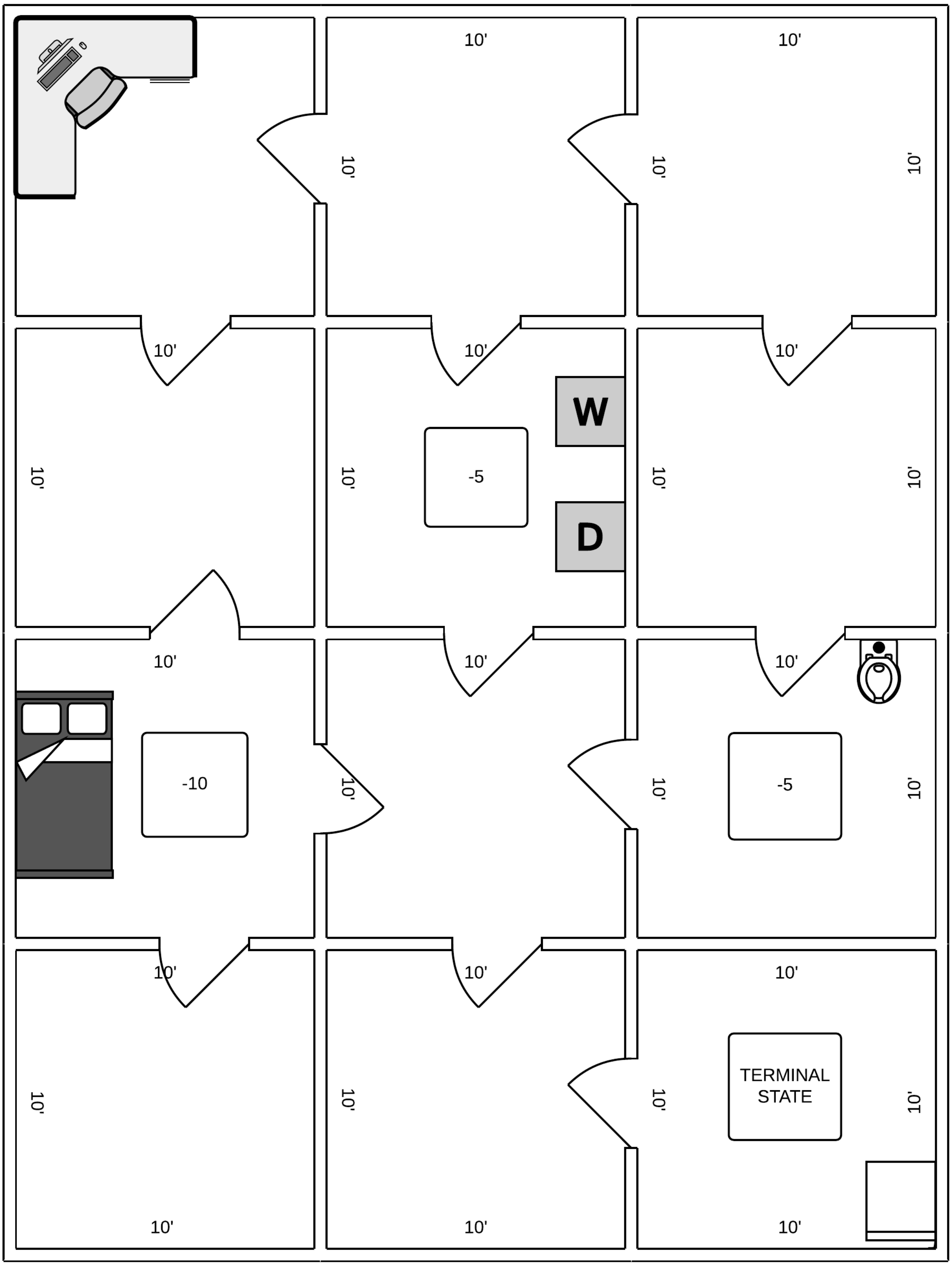

Figure 1: Floorplan defining an MDP. Every movement $\{\text{up}, \text{down}, \text{right}, \text{left}\}$ except in the terminal state in the bottom right gives a negative reward of -1, three rooms give an *additional* negative reward as indicated.

A. **Hoeffding's inequality**: by Hoeffding's inequality, with probability greater than $1-\delta$:

$$p \leq \hat{p}_n + \sqrt{\frac{\ln \frac{1}{\delta}}{2n}}.$$

(This is the bound we would like you to plot.)

B. kl **inequality**: again, you should take the bound in a form "with probability greater than $1-\delta$, $p \leq \ldots$". In the lecture notes we provide a bound on $\mathrm{kl}(\hat{p}_n \| p)$. The upper bound on p follows by taking the "upper inverse" of kl. Namely, we define $\mathrm{kl}^{-1^+}(\hat{p}_n, z) = \max\{p : \mathrm{kl}(\hat{p}_n \| p) \leq z\}$. We have that if $\mathrm{kl}(\hat{p}_n \| p) \leq z$ then $p \leq \mathrm{kl}^{-1^+}(\hat{p}_n, z)$.

We provide a MATLAB function for numerical computation of the inverse kl^{-1^+}.

C. **Pinsker's relaxation of the** kl **inequality**: the bound on p that follows from kl-inequality by Pinsker's inequality.

D. **Refined Pinsker's relaxation of the** kl **inequality**: the bound on p that follows from kl-inequality by refined Pinsker's inequality.

In this task you should do the following:

1. Write down explicitly the four bounds on p you are evaluating.

2. Plot the four bounds on p as a function of $\hat{p}_n$ for $\hat{p}_n \in [0,1]$, $n = 1000$, and $\delta = 0.01$. You should plot all four bounds in one figure, so that you can directly compare them. Clip all bounds at 1, because otherwise they are anyway meaningless and will only destroy the scale of the figure.

3. Generate a "zoom in" plot for $\hat{p}_n \in [0, 0.1]$.

4. Compare Hoeffding's lower bound on p with kl lower bound on p for the same values of $\hat{p}_n, n, \delta$ in a separate figure (no need to consider the relaxations of kl). The kl lower bound follows from the "lower inverse" of kl defined as $\mathrm{kl}^{-1^-}(\hat{p}_n, z) = \min\{p : \mathrm{kl}(\hat{p}_n \| p) \leq z\}$.

5. Write down your conclusions from the experiment. For what values of $\hat{p}_n$ which bounds are tighter and is the difference significant?

6. [Optional, not for submission.] You are welcome to experiment with other values of n and δ.

We provide a MATLAB function for inverting the kl-divergence with respect to its second argument. The function computes the "upper inverse" $\mathrm{kl}^{-1^+}(\hat{p}, \varepsilon) = \max\{p : \mathrm{kl}(\hat{p} \| p) \leq \varepsilon\}$. The inversion is computed by binary search. You are not obliged to use the function and can write your own if you like using any programming language you like. For the "lower inverse" $\mathrm{kl}^{-1^-}(\hat{p}, \varepsilon) = \min\{p : \mathrm{kl}(\hat{p} \| p) \leq \varepsilon\}$ you can either adapt the "upper inverse" function (and we leave it to you to think how to do this) or write your own function. Whatever way you chose you should explain in your main `.pdf` submission file how you computed the upper and the lower bound. Please, attach all code that you used for solving the assignment in a separate `.zip` file.

3 Refined Pinsker's Lower Bound (20 points)

Prove that if $\mathrm{kl}(p \| q) \leq \varepsilon$ then $q \geq p - \sqrt{2p\varepsilon}$. You are allowed to use Refined Pinsker's inequality in Lemma 2.18 in Yevgeny's lecture notes.

4 Occam's razor with kl inequality (20 points)

Prove the following theorem.

Theorem 1. *Let S be an i.i.d. sample of n points, let ℓ be the zero-one loss, let $\mathcal{H}$ be countable, and let $\pi(h)$ be such that it is independent of the sample S and satisfies $\pi(h) \geq 0$ for all h and $\sum_{h \in \mathcal{H}} \pi(h) \leq 1$. Let $\delta \in (0,1)$. Then with probability greater than $1-\delta$, for all $h \in \mathcal{H}$ simultaneously:*

$$\mathrm{kl}(\hat{L}(h,S) \| L(h)) \leq \frac{\ln \frac{n+1}{\pi(h)\delta}}{n}.$$

Briefly emphasize in your proof where you are using the assumption that $\pi(h)$ is independent of S and why it is necessary.

5 PAC-Bayes vs. Occam (10 points)

The change of measure inequality that is at the basis of PAC-Bayesian analysis can be seen as a replacement of the union bound, which is at the basis of Occam's razor. In this question we compare the tightness of the two approaches.

1. Prove the following theorem.

   Theorem 2. *Let S be an i.i.d. sample of n points, let ℓ be the zero-one loss, let $\mathcal{H}$ be countable, and let $\pi(h)$ be such that it is independent of the sample S and satisfies $\sum_{h \in \mathcal{H}} \pi(h) \leq 1$. Let $\delta \in (0,1)$. Then with probability greater than $1-\delta$, for all distributions ρ over $\mathcal{H}$ simultaneously:*

   $$\mathrm{kl}\left(\mathbb{E}_\rho\left[\hat{L}(h,S)\right] \middle\| \mathbb{E}_\rho\left[L(h)\right]\right) \leq \frac{\mathbb{E}_\rho\left[\ln \frac{1}{\pi(h)}\right] + \ln \frac{n+1}{\delta}}{n}. \tag{1}$$

   You can use the result from Theorem 1 to prove Theorem 2.

2. Recall that by PAC-Bayes-kl inequality, under the conditions of Theorem 2 we have that with probability greater than $1-\delta$, for all distributions ρ over $\mathcal{H}$ simultaneously:

   $$\mathrm{kl}\left(\mathbb{E}_\rho\left[\hat{L}(h,S)\right] \middle\| \mathbb{E}_\rho\left[L(h)\right]\right) \leq \frac{\mathrm{KL}(\rho \| \pi) + \ln \frac{n+1}{\delta}}{n}. \tag{2}$$

   Show that the PAC-Bayes-kl inequality in equation (2) is always at least as tight as the Occam's razor bound with kl in equation (1). Hint: the entropy of a discrete distribution is always non-negative, $\mathrm{H}(\rho) \geq 0$.

---

6 [Optional, not for submission] Asymmetry of kl divergence

Prove that kl is asymmetric in its arguments by providing an example of p and q for which $\mathrm{kl}(p\|q) \neq \mathrm{kl}(q\|p)$.

7 [Optional, not for submission] Fast convergence rates when the empirical loss is zero

In this question we provide a simple and intuitive explanation on why faster convergence rates are possible when the empirical loss is zero. The kl inequality provides a continuous interpolation between fast convergence rates (of order $\frac{1}{n}$) when the empirical loss is zero and slow convergence rates (of order $\sqrt{\frac{1}{n}}$) when it is close to 1/2.

1. Let $X_1, \ldots, X_n$ be a sample of n independent Bernoulli random variables with bias $p = \mathbb{P}\{X = 1\}$. Let $\hat{p}_n = \frac{1}{n}\sum_{i=1}^n X_i$ be the empirical average. Recall that by Hoeffding's inequality

$$\mathbb{P}\left\{p \geq \hat{p}_n + \sqrt{\frac{\ln \frac{1}{\delta}}{2n}}\right\} \leq \delta.$$

   Prove that if $p \geq \varepsilon$ then $\mathbb{P}\{\hat{p}_n = 0\} \leq e^{-n\varepsilon}$. In other words, if $p \geq \frac{\ln \frac{1}{\delta}}{n}$ then $\mathbb{P}\{\hat{p}_n = 0\} \leq \delta$.

   The result means that the probability of observing a sample with $\hat{p}_n = 0$ that is non-representative (diverges from the true mean p) by more than $\frac{\ln \frac{1}{\delta}}{n}$ is bounded by δ. (Note that for a sample with $\hat{p}_n = \frac{1}{2}$ we can only bound divergence from the true mean by $\sqrt{\frac{\ln \frac{1}{\delta}}{2n}}$ with the same probability δ.)

   Hint for the proof: what is the probability that we make n independent flips of a coin with bias p and get all zeros? The inequality $1 + x \leq e^x$ is helpful for the proof.

2. Let S be an i.i.d. sample of n points. Let $\mathcal{H}$ be countable and let $\pi(h)$ be such that it is independent of S and $\sum_{h \in \mathcal{H}} \pi(h) \leq 1$. Assume that for all $h \in \mathcal{H}$ we have $L(h) \geq \frac{\ln \frac{1}{\pi(h)\delta}}{n}$. Show that

$$\mathbb{P}\Big\{\exists h \in \mathcal{H} : \hat{L}(h, S) = 0\Big\} \leq \delta.$$

   The result means that the probability of observing an empirical loss $\hat{L}(h^*, S) = 0$ that is non-representative by more than $\frac{\ln \frac{1}{\pi(h^*)\delta}}{n}$ is bounded by δ.

Good luck!
Yevgeny, Christian, Chloé, and Yi-Shan