CS 747, Autumn 2020: Week 3, Q&A

Shivaram Kalyanakrishnan

Department of Computer Science and Engineering
Indian Institute of Technology Bombay

Autumn 2020

Question 1

- For $x, y \in [0, 1]$, show that $KL(x, y) \geq 2(x - y)^2$.

Question 1

- For $x, y \in [0, 1]$, show that $KL(x, y) \geq 2(x - y)^2$.

- For $x, y \in [0, 1]$ define

$$f_x(y) \stackrel{\text{def}}{=} KL(x, y) - 2(x - y)^2$$
$$= x \ln \frac{x}{y} + (1 - x) \ln \frac{1 - x}{1 - y} - 2(x - y)^2.$$

Question 1

- For $x, y \in [0, 1]$, show that $KL(x, y) \geq 2(x - y)^2$.

- For $x, y \in [0, 1]$ define

$$f_x(y) \stackrel{\text{def}}{=} KL(x, y) - 2(x - y)^2$$
$$= x \ln \frac{x}{y} + (1 - x) \ln \frac{1 - x}{1 - y} - 2(x - y)^2.$$

- We need to show for $x, y \in [0, 1]$: $f_x(y) \geq 0$.

Question 1

- For $x, y \in [0, 1]$, show that $KL(x, y) \geq 2(x - y)^2$.

- For $x, y \in [0, 1]$ define

$$f_x(y) \stackrel{\text{def}}{=} KL(x, y) - 2(x - y)^2$$
$$= x \ln \frac{x}{y} + (1 - x) \ln \frac{1 - x}{1 - y} - 2(x - y)^2.$$

- We need to show for $x, y \in [0, 1]$: $f_x(y) \geq 0$.
- Observe that if $y = x$, then $f_x(y) = 0$; otherwise if y is either 0 or 1, $f_x(y) = \infty$. The result seems to be correct at the "extremes"; what if $y \in (0, 1)$?

Question 1

- For $x, y \in [0, 1]$, show that $KL(x, y) \geq 2(x - y)^2$.

- You can work out that

$$\frac{df_x(y)}{dy} = (x - y)\left(4 - \frac{1}{y(1-y)}\right).$$

Question 1

- For $x, y \in [0, 1]$, show that $KL(x, y) \geq 2(x - y)^2$.

- You can work out that

$$\frac{df_x(y)}{dy} = (x - y)\left(4 - \frac{1}{y(1 - y)}\right).$$

- Since $y \in (0, 1)$, the second factor is never positive (it is 0 at $y = \frac{1}{2}$).

Question 1

- For $x, y \in [0, 1]$, show that $KL(x, y) \geq 2(x - y)^2$.

- You can work out that
$$\frac{df_x(y)}{dy} = (x - y)\left(4 - \frac{1}{y(1-y)}\right).$$

- Since $y \in (0, 1)$, the second factor is never positive (it is 0 at $y = \frac{1}{2}$).
- If $y \leq x$, $\frac{df_x(y)}{dy} \leq 0$; if $y \geq x$, $\frac{df_x(y)}{dy} \geq 0$.

Question 1

- For $x, y \in [0, 1]$, show that $KL(x, y) \geq 2(x - y)^2$.

- You can work out that

$$\frac{df_x(y)}{dy} = (x - y)\left(4 - \frac{1}{y(1-y)}\right).$$

- Since $y \in (0, 1)$, the second factor is never positive (it is 0 at $y = \frac{1}{2}$).
- If $y \leq x$, $\frac{df_x(y)}{dy} \leq 0$; if $y \geq x$, $\frac{df_x(y)}{dy} \geq 0$.

Hence, $f_x(y)$ is minimum at $y = x$, where it has value 0.

Question 2

- In our discussions on bandits, where all did we need/use the assumption of Bernoulli rewards?

Question 2

- In our discussions on bandits, where all did we need/use the assumption of Bernoulli rewards?

- **GLIE conditions** $C1$**,** $C2$: Although we wrote down the proof assuming Bernoulli rewards (implying a finite branching factor of 2 for histories), the proof can be redone for any bounded rewards. We would need integrals in place of sums if the rewards came from a continuous distribution.

Question 2

- In our discussions on bandits, where all did we need/use the assumption of Bernoulli rewards?

- **GLIE conditions** $C1$**,** $C2$: Although we wrote down the proof assuming Bernoulli rewards (implying a finite branching factor of 2 for histories), the proof can be redone for any bounded rewards. We would need integrals in place of sums if the rewards came from a continuous distribution.
- ϵ**-first,** ϵ**-greedy strategies**: These algorithms only compare the arms' empirical means, and don't depend on the Bernoulli assumption.

Question 2

- In our discussions on bandits, where all did we need/use the assumption of Bernoulli rewards?

- **GLIE conditions** $C1$**,** $C2$: Although we wrote down the proof assuming Bernoulli rewards (implying a finite branching factor of 2 for histories), the proof can be redone for any bounded rewards. We would need integrals in place of sums if the rewards came from a continuous distribution.
- ϵ**-first,** ϵ**-greedy strategies**: These algorithms only compare the arms' empirical means, and don't depend on the Bernoulli assumption.
- **UCB, KL-UCB**: Hoeffding's Inequality and the KL inequality hold for all bounded rewards. Hence UCB and KL-UCB work for all bounded rewards. The pseudocode given in class can be applied as is if rewards come from $[0, 1]$; otherwise they must first be scaled to lie within $[0, 1]$.

Question 2

- In our discussions on bandits, where all did we need/use the assumption of Bernoulli rewards?

- **Thompson Sampling**: We presented Thompson Sampling for Bernoulli rewards—for which the Beta distribution is the correct conjugate prior. If rewards come from some other parametric family, the form of the belief distribution would have to change. Note that any bounded rewards can be transformed into Bernoulli rewards (Agrawal and Goyal, 2012; see Algorithm 2).

Question 2

- In our discussions on bandits, where all did we need/use the assumption of Bernoulli rewards?

- **Thompson Sampling**: We presented Thompson Sampling for Bernoulli rewards—for which the Beta distribution is the correct conjugate prior. If rewards come from some other parametric family, the form of the belief distribution would have to change. Note that any bounded rewards can be transformed into Bernoulli rewards (Agrawal and Goyal, 2012; see Algorithm 2).
- **Lower bound**: Lai and Robbins' lower bound holds for arbitrary distributions (not necessary to even be bounded, but subject to certain technical conditions).

Question 2

- In our discussions on bandits, where all did we need/use the assumption of Bernoulli rewards?

- **Thompson Sampling**: We presented Thompson Sampling for Bernoulli rewards—for which the Beta distribution is the correct conjugate prior. If rewards come from some other parametric family, the form of the belief distribution would have to change. Note that any bounded rewards can be transformed into Bernoulli rewards (Agrawal and Goyal, 2012; see Algorithm 2).
- **Lower bound**: Lai and Robbins' lower bound holds for arbitrary distributions (not necessary to even be bounded, but subject to certain technical conditions).

In summary: the Bernoulli assumption in not critical.