Technion – Israel Institute of Technology

HW3

Autonomous Navigation and Perception

086762

Alon Spinner	305184335	alonspinner@gmail.com
Dan Hazzan	308553601	danhazzan@campus.technion.ac.il

April 26, 2022

Theoretical questions

Q1:

Consider two distributions $p(X,Y)$ & $q(X,Y)$ over random variables X & Y.

Prove that the KL divergence between these two distributions can be factorized as:

$KL[p(X,Y)||q(X,Y)] = KL[p(Y)||q(Y)] + KL[p(X|Y)||q(X|Y)]$

<u>Proof:</u>

$$KL[p(X,Y)||q(X,Y)] \doteq \Sigma_X \Sigma_Y p(X,Y) \log\left(\frac{p(X,Y)}{q(X,Y)}\right) =$$

$$\underbrace{=}_{CR} \Sigma_X \Sigma_Y p(X,Y) \log\left(\frac{p(Y)p(X|Y)}{q(Y)q(X|Y)}\right) =$$

$$\underbrace{=}_{\log rules} \Sigma_X \Sigma_Y p(X,Y) \left[log\left(\frac{p(Y)}{q(Y)}\right) + log\left(\frac{p(X|Y)}{q(X|Y)}\right)\right] =$$

$$= \Sigma_Y \Sigma_X p(X,Y) log\left(\frac{p(Y)}{q(Y)}\right) + \Sigma_Y \Sigma_X p(X,Y) log\left(\frac{p(X|Y)}{q(X|Y)}\right) =$$

$$= \Sigma_Y \big(\Sigma_X p(X,Y)\big) log\left(\frac{p(Y)}{q(Y)}\right) + \Sigma_Y \Sigma_X p(X|Y)p(Y) log\left(\frac{p(X|Y)}{q(X|Y)}\right) =$$

$$\underbrace{=}_{MR+CR} \Sigma_Y p(Y) log\left(\frac{p(Y)}{q(Y)}\right) + \Sigma_Y p(Y)\left(\Sigma_X p(X|Y) log\left(\frac{p(X|Y)}{q(X|Y)}\right)\right) \doteq$$

$$\doteq KL[p(Y)||q(Y)] + KL[p(X|Y)||q(X|Y)]$$

Q2:

a. $J(b_k, a_k) = \underbrace{\mathbb{E}}_{z_{k+1}} [r(b_k, a_k) + r(b_{k+1})]$

$$= \Sigma_{z_{k+1} \epsilon Z} p(z_{k+1}|b_k, a_k)(r(b_k, a_k) + r_\tau(b_{k+1}|b_k, a_k, z_{k+1})) =$$

$$= r(b_k, a_k) + \Sigma_{z_{k+1} \epsilon Z} p(z_{k+1}|b_k, a_k) r_\tau(b_{k+1}|b_k, a_k, z_{k+1})$$

Steps:

1. Calculate $r(b_k, a_k)$
2. sample $N = |Z|$ observations z_{k+1}^i from probability distribution
3. From for each observation z_{k+1}^i update belief b_{k+1}^i
4. Calculate expectancy $\Sigma_{z_{k+1} \epsilon Z} p(z_{k+1}|b_k, a_k) r_\tau(b_{k+1}|b_k, a_k, z_{k+1})$
5. Add calculated $r(b_k, a_k)$ to the expectancy calculated in step 4.

b.

$$p(z_{k+1}|b_k, a_k) \underbrace{=}_{marginalization} \Sigma_{b_{k+1}} p(z_{k+1}, b_{k+1}|, b_k, a_k)$$

$$\underbrace{=}_{chain\ rule} \Sigma_{b_{k+1}} p(z_{k+1}|b_{k+1}, b_k, a_k)p(b_{k+1}|b_k, a_k)$$

$$\underbrace{=}_{markov} \Sigma_{b_{k+1}} p(z_{k+1}|b_{k+1})p(b_{k+1}|b_k, a_k)$$

c.

Following the solutions above, we wrote an algorithm to compute $p(z_{k+1}|b_k, a_k)$, given that $b_k \doteq b[x_k]$, which results in $O_z = O(|X|^2)$

The algorithm for (b):

init $p(z_{k+1}=z^*|b_k, a_k) = 0$

for each x_i in X do:

  init $p(b_{k+1}=x_i|b_k, a_k) = 0$

  for each x_j in X do:

    $p(b_{k+1}=x_i|b_k, a_k) += p(b_{k+1}=x_i|b_k=x_j, a_k) \cdot p(b_k=x_j)$

  $p(z_{k+1}=z^*|b_k, a_k) += p(z_{k+1}=z^*|b_{k+1}=x_i) \cdot p(b_{k+1}=x_i|b_k, a_k)$

$\Rightarrow O_z = O(|X|^2)$

d.

$$b_{k+1} = p(x_{k+1}|b_k, a_k, z_{k+1}) \underbrace{=}_{bayes} \frac{p(z_{k+1}|b_k, a_k, x_{k+1})p(x_{k+1}|b_k, a_k)}{p(z_{k+1}|b_k, a_k)}$$

$$\underbrace{=}_{markov\ \&\ same\ as\ 2(b)} \frac{p(z_{k+1}|x_{k+1})p(x_{k+1}|b_k, a_k)}{\Sigma_{b_{k+1}} p(z_{k+1}|b_{k+1})p(b_{k+1}|b_k, a_k)}$$

e.

<u>Numerator:</u>

Computing $p(z_{k+1}|x_{k+1}) \rightarrow O(1)$

Computing $p(x_{k+1}|b_k, a_k) \rightarrow O(|X|)$

<u>Denominator:</u>

Computing the denominator has the same computational complexity as in (2c)

$$\rightarrow O(|X|^2)$$

Hence:

$$O_b = O(1) + O(|X|) + O(|X|^2) = O(|X|^2)$$

f.

$$r(b[x], a) = H(b[x]) = -\Sigma_x b[x] \log\left(b[x]\right)$$

Since the algorithm runs over all the states, $O_r = O(|x|)$

g.

Calculating the objective, we have:

$$J(b_k, a_k) = \Sigma_{z_{k+1} \epsilon Z} p(z_{k+1}|b_k, a_k)\big(r(b_k, a_k) + r(b_{k+1})\big) =$$

$$= r(b_k, a_k) + \Sigma_{z_{k+1} \epsilon Z} p(z_{k+1}|b_k, a_k) r_\tau(b_{k+1}|b_k, a_k, z_{k+1})$$

Running over all observations: $|Z|$

$$\Rightarrow O\big(J(b_k, a_k)\big) = O_r + \ |Z| \cdot \big(O_z + (O_r + O_b)\big) = O(|Z| \cdot |X|^2)$$