EECS545 WN20 - Homework 6 solution

Due date: 11:55pm, Tuesday 4/19/2022

1 [35 points] Conditional Variational Autoencoders.

In this problem, you will implement a conditional variational autoencoder (CVAE) and train it on the MNIST dataset.

(a) (10 pts) Derive the variational lower bound of a conditional variational autoencoder. Show that:

$$\begin{aligned}\log p_\theta(\mathbf{x}|\mathbf{y}) &\geq \mathcal{L}(\theta, \phi; \mathbf{x}, \mathbf{y}) \\ &= \mathbb{E}_{q_\phi(\mathbf{z}|\mathbf{x},\mathbf{y})}\left[\log p_\theta(\mathbf{x}|\mathbf{z},\mathbf{y})\right] - D_{KL}\left(q_\phi(\mathbf{z}|\mathbf{x},\mathbf{y}) \| p_\theta(\mathbf{z}|\mathbf{y})\right), \qquad (1)\end{aligned}$$

Answer: From the second line in the following, we will denote the $\mathbb{E}_{\mathbf{z}\sim q(\mathbf{z}|\mathbf{x},\mathbf{y})}$ as $\mathbb{E}_{\mathbf{z}}$, q_ϕ as q, p_θ as p for notational simplicity.

$$\begin{aligned}\log p_\theta(\mathbf{x}|\mathbf{y}) &= \mathbb{E}_{\mathbf{z}\sim q(\mathbf{z}|\mathbf{x},\mathbf{y})}\left[\log p_\theta(\mathbf{x}|\mathbf{y})\right] & (2)\\ &= \mathbb{E}_{\mathbf{z}}\left[\log \frac{p(\mathbf{x}|\mathbf{y},\mathbf{z})p(\mathbf{z}|\mathbf{y})}{p(\mathbf{z}|\mathbf{x},\mathbf{y})}\right] & (3)\\ &= \mathbb{E}_{\mathbf{z}}\left[\log \frac{p(\mathbf{x}|\mathbf{y},\mathbf{z})p(\mathbf{z}|\mathbf{y})}{p(\mathbf{z}|\mathbf{x},\mathbf{y})}\frac{q(\mathbf{z}|\mathbf{x},\mathbf{y})}{q(\mathbf{z}|\mathbf{x},\mathbf{y})}\right] & (4)\\ &= \mathbb{E}_{\mathbf{z}}[\log p(\mathbf{x}|\mathbf{y},\mathbf{z})] - \mathbb{E}_{\mathbf{z}}\left[\log \frac{q(\mathbf{z}|\mathbf{x},\mathbf{y})}{p(\mathbf{z}|\mathbf{y})}\right] + \mathbb{E}_{\mathbf{z}}\left[\log \frac{q(\mathbf{z}|\mathbf{x},\mathbf{y})}{p(\mathbf{z}|\mathbf{x},\mathbf{y})}\right] & (5)\\ &= \mathbb{E}_{\mathbf{z}}[\log p(\mathbf{x}|\mathbf{y},\mathbf{z})] - D_{KL}(q(\mathbf{z}|\mathbf{x},\mathbf{y})\|p(\mathbf{z}|\mathbf{y})) + \underbrace{D_{KL}(q(\mathbf{z}|\mathbf{x},\mathbf{y})\|p(\mathbf{z}|\mathbf{x},\mathbf{y}))}_{\geq 0} & (6)\\ &\geq \mathbb{E}_{\mathbf{z}}[\log p(\mathbf{x}|\mathbf{y},\mathbf{z})] - D_{KL}(q(\mathbf{z}|\mathbf{x},\mathbf{y})\|p(\mathbf{z}|\mathbf{y})) & (7)\\ &= \mathcal{L}(\theta, \phi; \mathbf{x}, \mathbf{y}) & (8)\end{aligned}$$

(b) (10 pts) Derive the analytical solution to the KL-divergence between two Gaussian distributions $D_{KL}\left(q_\phi(\mathbf{z}|\mathbf{x},\mathbf{y}) \| p_\theta(\mathbf{z}|\mathbf{y})\right)$. Let us assume that $p_\theta(\mathbf{z}|\mathbf{y}) \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ and show that:

$$D_{KL}\left(q_\phi(\mathbf{z}|\mathbf{x},\mathbf{y}) \| p_\theta(\mathbf{z}|\mathbf{y})\right) = -\frac{1}{2}\sum_{j=1}^{m}\left(1 + \log\left(\sigma_j^2\right) - \mu_j^2 - \sigma_j^2\right), \qquad (9)$$

Answer: First, remember that all z_j are assumed to be independent, i.e. $p_\theta(\mathbf{z}|\mathbf{y}) = \Pi_j p_\theta(z_j|\mathbf{y})$. Therefore, we have

$$D_{KL}\left(q_\phi(\mathbf{z}|\mathbf{x},\mathbf{y})\|p_\theta(\mathbf{z}|\mathbf{y})\right) = \int_{\mathbf{z}} q_\phi(\mathbf{z}|\mathbf{x},\mathbf{y}) \log \frac{q_\phi(\mathbf{z}|\mathbf{x},\mathbf{y})}{p_\theta(\mathbf{z}|\mathbf{y})} d\mathbf{z} \tag{10}$$

$$= \int_{\mathbf{z}} q_\phi(\mathbf{z}|\mathbf{x},\mathbf{y}) \log q_\phi(\mathbf{z}|\mathbf{x},\mathbf{y}) d\mathbf{z} - \int_{\mathbf{z}} q_\phi(\mathbf{z}|\mathbf{x},\mathbf{y}) \log p_\theta(\mathbf{z}|\mathbf{y}) d\mathbf{z} \tag{11}$$

$$= \int_{\mathbf{z}} q_\phi(\mathbf{z}|\mathbf{x},\mathbf{y}) \log \prod_j q_\phi\left(z_j|\mathbf{x},\mathbf{y}\right) d\mathbf{z} - \int_{\mathbf{z}} q_\phi(\mathbf{z}|\mathbf{x},\mathbf{y}) \log \prod_j p_\theta\left(z_j|\mathbf{y}\right) d\mathbf{z} \tag{12}$$

$$= \int_{\mathbf{z}} q_\phi(\mathbf{z}|\mathbf{x},\mathbf{y}) \sum_j \log q_\phi\left(z_j|\mathbf{x},\mathbf{y}\right) d\mathbf{z} - \int_{\mathbf{z}} q_\phi(\mathbf{z}|\mathbf{x},\mathbf{y}) \sum_j \log p_\theta\left(z_j|\mathbf{y}\right) d\mathbf{z} \tag{13}$$

$$= \sum_j \int_{\mathbf{z}} q_\phi(\mathbf{z}|\mathbf{x},\mathbf{y}) \log q_\phi\left(z_j|\mathbf{x},\mathbf{y}\right) d\mathbf{z} - \sum_j \int_{\mathbf{z}} q_\phi(\mathbf{z}|\mathbf{x},\mathbf{y}) \log p_\theta\left(z_j|\mathbf{y}\right) d\mathbf{z} \tag{14}$$

Now, note that for each j in the sum, we have

$$\int_{\mathbf{z}} q_\phi(\mathbf{z}|\mathbf{x},\mathbf{y}) \log q_\phi\left(z_j|\mathbf{x},\mathbf{y}\right) d\mathbf{z} = \int_{\mathbf{z}} \prod_{j'} q_\phi\left(z_{j'}|\mathbf{x},\mathbf{y}\right) \log q_\phi\left(z_j|\mathbf{x},\mathbf{y}\right) d\mathbf{z} \tag{15}$$

$$= \int_{z_j} q_\phi\left(z_j|\mathbf{x},\mathbf{y}\right) \log q_\phi\left(z_j|\mathbf{x},\mathbf{y}\right) dz_j \prod_{j'\neq j} \int_{z_{j'}} q_\phi\left(z_{j'}|\mathbf{x},\mathbf{y}\right) dz_{j'} \tag{16}$$

$$= \int_{z_j} q_\phi\left(z_j|\mathbf{x},\mathbf{y}\right) \log q_\phi\left(z_j|\mathbf{x},\mathbf{y}\right) dz_j \tag{17}$$

Therefore, we have

$$D_{KL}\left(q_\phi(\mathbf{z}|\mathbf{x},\mathbf{y})\|p_\theta(\mathbf{z}|\mathbf{y})\right) = \sum_j \int_{z_j} q_\phi\left(z_j|\mathbf{x},\mathbf{y}\right) \log q_\phi\left(z_j|\mathbf{x},\mathbf{y}\right) dz_j - \sum_j \int_{z_j} q_\phi\left(z_j|\mathbf{x},\mathbf{y}\right) \log p_\theta\left(z_j|\mathbf{y}\right) dz_j \tag{18}$$

Now, let us plug in Gaussian distributions. Let's start with the left side:

$$\sum_j \int_{z_j} q_\phi\left(z_j|\mathbf{x},\mathbf{y}\right) \log q_\phi\left(z_j|\mathbf{x},\mathbf{y}\right) dz_j = \sum_j \int_{z_j} q_\phi\left(z_j|\mathbf{x},\mathbf{y}\right) \log\left(\frac{1}{\sqrt{2\pi\sigma_j^2}} e^{-\frac{(z_j-\mu_j)^2}{2\sigma_j^2}}\right) dz_j \tag{19}$$

$$= \sum_j \int_{z_j} q_\phi\left(z_j|\mathbf{x},\mathbf{y}\right) \left(\log \frac{1}{\sqrt{2\pi\sigma_j^2}} - \frac{\left(z_j-\mu_j\right)^2}{2\sigma_j^2}\right) dz_j \tag{20}$$

$$= \sum_j -\frac{1}{2}\log 2\pi - \frac{1}{2}\log\sigma_j^2 - \frac{1}{2\sigma_j^2}\int_{z_j} q_\phi\left(z_j|\mathbf{x},\mathbf{y}\right)\left[\left(z_j-\mu_j\right)^2\right] dz_j \tag{21}$$

$$= \sum_j -\frac{1}{2}\log 2\pi - \frac{1}{2}\log\sigma_j^2 - \frac{1}{2\sigma_j^2}\sigma_j^2 \tag{22}$$

$$= \sum_j -\frac{1}{2}\log 2\pi - \frac{1}{2}\log\sigma_j^2 - \frac{1}{2} \tag{23}$$

Now, let us solve the right side. Remember that we assume that $p_\theta(\mathbf{z}|\mathbf{y}) \sim \mathcal{N}(\mathbf{0},\mathbf{I})$. Therefore:

$$\sum_j \int_{z_j} q_\phi\left(z_j|\mathbf{x},\mathbf{y}\right)\log p_\theta\left(z_j|\mathbf{y}\right)dz_j = \sum_j \int_{z_j} q_\phi\left(z_j|\mathbf{x},\mathbf{y}\right)\log\left(\frac{1}{\sqrt{2\pi}}e^{-\frac{z^2}{2}}\right)dz_j \tag{24}$$

$$= \sum_j \int_{z_j} q_\phi\left(z_j|\mathbf{x},\mathbf{y}\right)\left(\log\frac{1}{\sqrt{2\pi}} - \frac{z_j^2}{2}\right)dz_j \tag{25}$$

$$= \sum_j -\frac{1}{2}\log 2\pi - \frac{1}{2}\int_{z_j} q_\phi\left(z_j|\mathbf{x},\mathbf{y}\right)z_j^2 dz_j \tag{26}$$

$$= \sum_j -\frac{1}{2}\log 2\pi - \frac{1}{2}\left(\sigma_j^2 + \mu_j^2\right) \tag{27}$$

Therefore, combining the solutions of the left and right side we have:

$$D_{KL}\left(q_\phi(\mathbf{z}|\mathbf{x},\mathbf{y})\|p_\theta(\mathbf{z}|\mathbf{y})\right) = \left(\sum_j -\frac{1}{2}\log 2\pi - \frac{1}{2}\log\sigma_j^2 - \frac{1}{2}\right) - \left(\sum_j -\frac{1}{2}\log 2\pi - \frac{1}{2}\left(\sigma_j^2 + \mu_j^2\right)\right) \tag{28}$$

$$= \sum_j -\frac{1}{2}\log\sigma_j^2 - \frac{1}{2} + \frac{1}{2}\left(\sigma_j^2 + \mu_j^2\right) \tag{29}$$

$$= -\frac{1}{2}\sum_j\left(1 + \log\sigma_j^2 - \sigma_j^2 - \mu_j^2\right) \tag{30}$$

Proof for

$$D_{KL}\left(q_\phi\left(\mathbf{z}|\mathbf{x},\mathbf{y}\right)\|p_\theta\left(\mathbf{z}|\mathbf{y}\right)\right) = \sum_{j=1}^m D_{KL}\left(q_\phi\left(z_j|\mathbf{x},\mathbf{y}\right)\|p_\theta\left(z_j|\mathbf{y}\right)\right).$$

$$\begin{aligned}
D_{KL}\left(q_\phi\left(\mathbf{z}|\mathbf{x},\mathbf{y}\right)\|p_\theta\left(\mathbf{z}|\mathbf{y}\right)\right) &= \int_{\mathbf{z}} q_\phi(\mathbf{z}|\mathbf{x},\mathbf{y})\log\frac{q_\phi(\mathbf{z}|\mathbf{x},\mathbf{y})}{p_\theta(\mathbf{z}|\mathbf{y})}d\mathbf{z} \\
&= \sum_{j=1}^m \int_{\mathbf{z}} q_\phi(\mathbf{z}|\mathbf{x},\mathbf{y})\log\frac{q_\phi(z_j|\mathbf{x},\mathbf{y})}{p_\theta(z_j|\mathbf{y})}d\mathbf{z} \\
&= \sum_{j=1}^m \int_{\mathbf{z}}\left(\prod_{k=1}^m q_\phi(z_k|\mathbf{x},\mathbf{y})\right)\log\frac{q_\phi(z_j|\mathbf{x},\mathbf{y})}{p_\theta(z_j|\mathbf{y})}d\mathbf{z} \\
&= \sum_{j=1}^m \int_{z_j} q_\phi(z_j|\mathbf{x},\mathbf{y})\log\frac{q_\phi(z_j|\mathbf{x},\mathbf{y})}{p_\theta(z_j|\mathbf{y})}dz_j\left(\prod_{k\neq j}\int_{z_k} q_\phi(z_k|\mathbf{x},\mathbf{y})dz_k\right) \\
&= \sum_{j=1}^m D_{KL}\left(q_\phi\left(z_j|\mathbf{x},\mathbf{y}\right)\|p_\theta\left(z_j|\mathbf{y}\right)\right)
\end{aligned}$$

(c) (15 pts) CVAE Implementation and results.

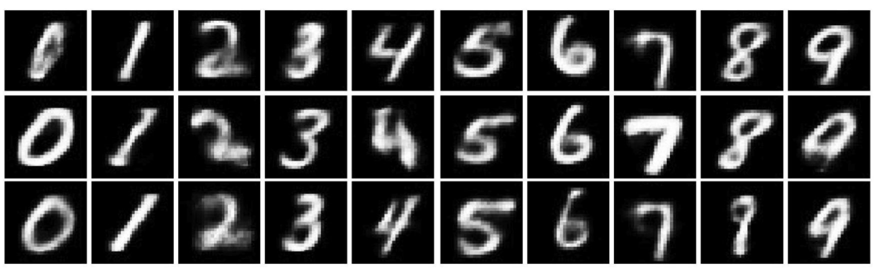


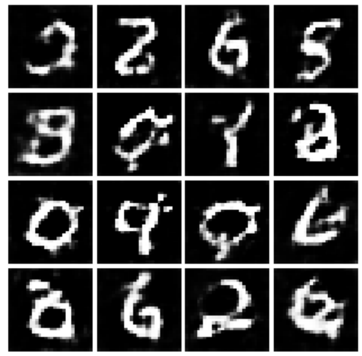
Figure 1: Example output of CVAE.

2 [30 points] Generative Adversarial Networks.

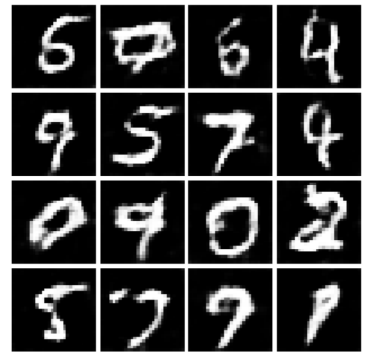
This problem asks you to implement generative adversarial networks and train it on MNIST dataset.

Solution: Please refer to the solution code for the detail.

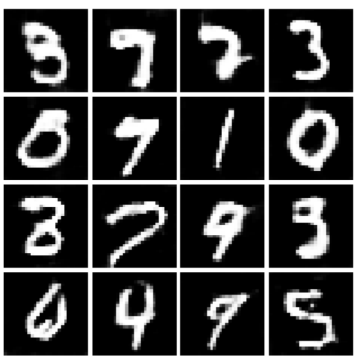
(a) After 0 Iterations



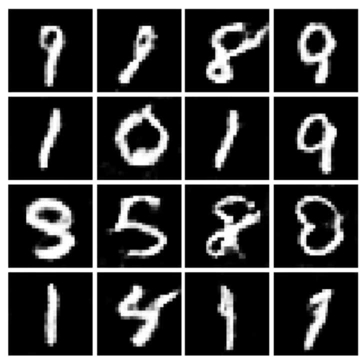
(b) After 250 Iterations



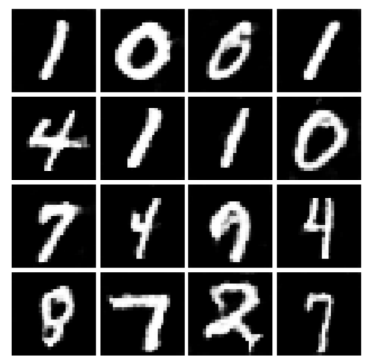
(c) After 500 Iterations



(d) After 1000 Iterations



(e) After 2000 Iterations



(f) After 3000 Iterations

Figure 2: Example outputs generated from DCGAN.

The result should look like a random noise after 0 iterations (right after initialization), but should converge in around 2000 iterations. It is expected that the generated image has not as good perceptual quality at 250-500 iterations.

3 [25 points] Deep Q-Network (DQN).

In this problem, you are asked to implement DQN algorithm.

Solution: Please refer to the solution code for the details. Below are some example outputs from the reference solution. It is OK to have a somewhat fluctating results (i.e. the performance drops due to the instability of DQN), but the peak average reward should reach higher than 300.

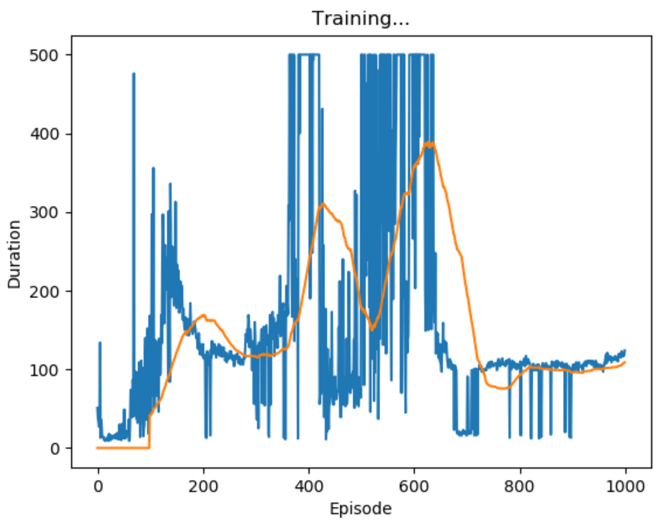


Figure 3: An example learning curve of DQN.

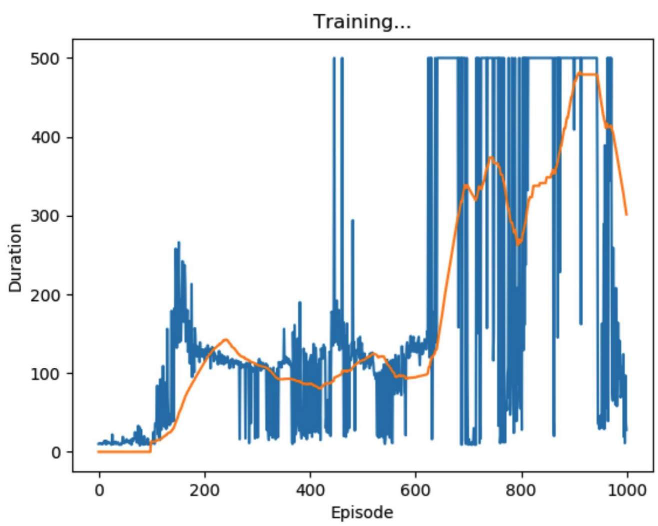


Figure 4: An example learning curve of DQN.

4 [10 points] Policy Gradients.

(a) Let $r(\tau) = \sum_{i=1}^{T} r_t$. In this case, the above gradient estimator becomes

$$\nabla_\theta J(\theta) \approx \frac{1}{N}\sum_{i=1}^{N}\sum_{t=1}^{T} \nabla_\theta \log\pi_\theta(a_t^i|s_t^i) \sum_{t'=1}^{T} r_{t'}^i$$

Show that the following is an unbiased estimate of $\nabla_\theta J(\theta)$.

$$\nabla_\theta J(\theta) \approx \frac{1}{N}\sum_{i=1}^{N}\sum_{t=1}^{T} \nabla_\theta \log\pi_\theta(a_t^i|s_t^i) \sum_{t'=t}^{T} r_{t'}^i$$

That is, we can omit the rewards collected in the past while keeping the estimator unbiased. The new estimator has the advantage of having lower variance than the original estimator.

Solution: According to the property of Monte Carlo method and derivation in part 1, we know that

$$\mathbb{E}\left[\frac{1}{N}\sum_{i=1}^{N}\sum_{t=1}^{T} \nabla_\theta \log\pi_\theta(a_t^i|s_t^i) \sum_{t'=1}^{T} r_{t'}^i\right] = \nabla_\theta J(\theta).$$

Now we decompose the $\sum_{t'=t}^{T}$ term into the difference between a summation $\sum_{t'=1}^{T}$ and another summation $\sum_{t'=1}^{t-1}$:

$$\mathbb{E}\left[\frac{1}{N}\sum_{i=1}^{N}\sum_{t=1}^{T} \nabla_\theta \log\pi_\theta(a_t^i|s_t^i) \sum_{t'=t}^{T} r_{t'}^i\right] \tag{31}$$

$$= \mathbb{E}\left[\frac{1}{N}\sum_{i=1}^{N}\sum_{t=1}^{T} \nabla_\theta \log\pi_\theta(a_t^i|s_t^i) \sum_{t'=1}^{T} r_{t'}^i\right] - \mathbb{E}\left[\frac{1}{N}\sum_{i=1}^{N}\sum_{t=1}^{T} \nabla_\theta \log\pi_\theta(a_t^i|s_t^i) \sum_{t'=1}^{t-1} r_{t'}^i\right] \tag{32}$$

$$= \nabla_\theta J(\theta) - \mathbb{E}\left[\frac{1}{N}\sum_{i=1}^{N}\sum_{t=1}^{T} \nabla_\theta \log\pi_\theta(a_t^i|s_t^i) \sum_{t'=1}^{t-1} r_{t'}^i\right] \tag{33}$$

To prove that $\sum_{t=1}^{T}(\sum_{t'=t}^{T} r_{t'}^i)\nabla_\theta \log\pi_\theta(a_t^i|s_t^i)$ is an unbiased estimator of $\nabla_\theta J(\theta)$, it suffices to show

$$\mathbb{E}\left[\sum_{t=1}^{T}(\sum_{t'=1}^{t-1} r_{t'}^i)\nabla_\theta \log\pi_\theta(a_t^i|s_t^i)\right] = 0.$$

If $t' < t$, then $r_{t'}$ does not depend on a_t because the current action wouldn't change the rewards in the previous timesteps (this should be mentioned clearly):

$$\mathbb{E}_{a_t\sim\pi_\theta(a_t|s_t)}[r_{t'}\nabla_\theta \log\pi_\theta(a_t|s_t)] \tag{34}$$

$$= r_{t'}\mathbb{E}_{a_t\sim\pi_\theta(a_t|s_t)}[\nabla_\theta \log\pi_\theta(a_t|s_t)] \tag{35}$$

$$= r_{t'}\int_a \pi_\theta(a|s_t)\nabla_\theta \log\pi_\theta(a|s_t)da \tag{36}$$

$$= r_{t'}\int_a \nabla_\theta \pi_\theta(a|s_t)da \tag{37}$$

$$= r_{t'}\nabla_\theta \int_a \pi_\theta(a|s_t)da \tag{38}$$

$$= r_{t'}\nabla_\theta 1 \tag{39}$$

$$= 0 \tag{40}$$

and we are done. $\square$

(b) [10 points] Show that adding a state dependent baseline does not introduce any bias in the estimator. i.e., Show that the following is an unbiased estimator of the gradient. Adding a baseline can further reduce the variance of the estimator.

$$\nabla_\theta J(\theta) \approx \frac{1}{N}\sum_{i=1}^{N}\sum_{t=1}^{T}\nabla_\theta \log\pi_\theta(a_t^i|s_t^i)\left(\left[\sum_{t'=t}^{T} r_{t'}^i\right] - b(s_t^i)\right)$$

Solution: So far, we have shown that

$$\mathbb{E}\left[\frac{1}{N}\sum_{i=1}^{N}\sum_{t=1}^{T}\nabla_\theta \log\pi_\theta(a_t^i|s_t^i)\sum_{t'=1}^{T} r_{t'}^i\right] = \nabla_\theta J(\theta). \tag{41}$$

So we just need to show

$$\mathbb{E}\left[\frac{1}{N}\sum_{i=1}^{N}\sum_{t=1}^{T}\nabla_\theta \log\pi_\theta(a_t^i|s_t^i) b(s_t^i)\right] = 0 \tag{42}$$

to prove that the estimator is unbiased. Indeed,

$$\begin{aligned}
&\mathbb{E}_{a_t\sim\pi_\theta(a_t|s_t)}[b(s_t)\nabla_\theta \log\pi_\theta(a_t|s_t)] && (43)\\
&= b(s_t)\mathbb{E}_{a_t\sim\pi_\theta(a_t|s_t)}[\nabla_\theta \log\pi_\theta(a_t|s_t)] && (44)\\
&= b(s_t)\int_a \pi_\theta(a|s_t)\nabla_\theta \log\pi_\theta(a|s_t)da && (45)\\
&= b(s_t)\int_a \nabla_\theta\pi_\theta(a|s_t)da && (46)\\
&= b(s_t)\nabla_\theta\int_a \pi_\theta(a|s_t)da && (47)\\
&= b(s_t)\nabla_\theta 1 && (48)\\
&= 0 && (49)\\
& && (50)
\end{aligned}$$

and we are done. □