2020270026 应用统计硕士 王姿文

---

1 Generative Models

Problem 1.1

IWAE

1

$By\ Jensen's\ Inequality,\ f(E[X]) \leq E[f(x)], \forall\ f(.)\ is\ a\ convex\ function$
$\because\ log(.)\ isn't\ a\ convex\ function$
$\Rightarrow IWAE(K) = E_{q(z|x,\phi)}[log(\frac{1}{K}\sum_{k=1}^{K}\frac{p(z^{(k)},x;\theta)}{q(z^{(k)}|x;\phi)})]$
$\leq log(\frac{1}{K}\sum_{k=1}^{K}E_{q(z|x,\phi)}[\frac{p(z^{(k)},x;\theta)}{q(z^{(k)}|x;\phi)}]) = logp(x;\theta)$
$\Rightarrow IWAE(K) \leq logp(x;\theta) - (1)$
又$\frac{1}{K}\sum_{k=1}^{K}log(\frac{p(z^{(k)},x;\theta)}{q(z^{(k)}|x,\phi)}) \leq log(\frac{1}{K}\sum_{k=1}^{K}\frac{p(z^{(k)},x;\theta)}{q(z^{(k)}|x,\phi)})$
$\Rightarrow \frac{1}{K}\sum_{k=1}^{K}E_{q(z|x,\phi)}[log(\frac{p(z^{(k)},x;\theta)}{q(z^{(k)}|x,\phi)})] \leq E_{q(z|x,\phi)}[log(\frac{1}{K}\sum_{k=1}^{K}\frac{p(z^{(k)},x;\theta)}{q(z^{(k)}|x,\phi)})]$
$\Rightarrow E_{q(z|x,\phi)}[log(\frac{p(z,x;\theta)}{q(z|x,\phi)})] \leq E_{q(z|x,\phi)}[log(\frac{1}{K}\sum_{k=1}^{K}\frac{p(z^{(k)},x;\theta)}{q(z^{(k)}|x,\phi)})]$
$\Rightarrow IWAE(1) \leq IWAE(K) - (2)$
$By\ (1),\ (2)$
$\Rightarrow IWAE(1) \leq IWAE(K) \leq logp(x;\theta)$得证

2

首先得$lim_{k\rightarrow+\infty}\frac{1}{K}\sum_{k=1}^{K}\frac{p(z^{(k)},x;\theta)}{q(z^{(k)}})) = E_{q(z|x,\phi)}[\frac{p(z,x;\theta)}{q(z|x,\phi)}] = p(z,x;\theta)$
$\therefore lim_{k\rightarrow+\infty}IWAE(K) = lim_{k\rightarrow+\infty}E_{q(z|x,\phi)}[log(\frac{1}{K}\sum_{k=1}^{K}\frac{p(z^{(k)},x;\theta)}{q(z^{(k)}|x;\phi)})]$
$= E_{q(z|x,\phi)}[log(lim_{k\rightarrow+\infty}\frac{1}{K}\sum_{k=1}^{K}\frac{p(z^{(k)},x;\theta)}{q(z^{(k)}|x;\phi)})]$
$= E_{q(z|x,\phi)}[logp(z,x;\theta)]$
$= logp(z,x;\theta)$
$\Rightarrow lim_{k\rightarrow+\infty}IWAE(K) = logp(z,x;\theta)$得证

3

$\because\ log(.)\ is\ a\ monotonic\ function$
$\therefore K < L, log(\frac{1}{K}\sum_{k=1}^{K}\frac{p(z^{(k)},x;\theta)}{q(z^{(k)}|x;\phi)}) \leq log(\frac{1}{L}\sum_{k=1}^{L}\frac{p(z^{(k)},x;\theta)}{q(z^{(k)}|x;\phi)}) \Rightarrow IWAE(K) \leq IWAE(L)$得证

Problem 1.2

Affine VAE and probabilistic PCA

$By\ equation\ in\ slide,\ logp(x^{(1)}, x^{(2)}, \ldots, x^{(n)}) = log \prod_{i=1}^{n} p(x^{(i)})$

$= \sum_{i=1}^{n} \int q(z|x^{(i)}) log \frac{p(x^{(i)},z)}{q(z|x^{(i)})} dz + \sum_{i=1}^{n} KL(q(z|x^{(i)})||p(z|x^{(i)}))$

$= ELBO + \sum_{i=1}^{n} KL(q(z|x^{(i)})||p(z|x^{(i)}))$

$\Rightarrow\ ELBO = \sum_{i=1}^{n} \int q(z|x^{(i)}) log \frac{p(x^{(i)},z)}{q(z|x^{(i)})} dz$

如果$x\ is\ constant$，则$logp(x)\ is\ constant$
已知$\mu^*,\ S^* \Rightarrow q(z|x^{(i)}) = p(z|x^{(i)})$
　$\Rightarrow p(z|x) \sim Multivariant\ Guassian\ Distribution$
　$\Rightarrow max ELBO = logp(x)$

故下求$logp(x)$：
$p(x) = \int p(x|z)p(z)dz = N(b, \lambda I + WW^T)$

$q(z|x) = p(z|x) = N((\lambda I + W^T W)^{-1} W^T (x-b), \lambda(\lambda I + W^T W)^{-1}) = N(\mu^*,\ (S^*)^2)$

$\Rightarrow\ ELBO = logp(x) = \sum_{i=1}^{n} logp(x^i) = -\sum_i (x^{(i)} - b)^T (\lambda I + WW^T)^{-1} (x^{(i)} - b) - nlog|\lambda I + WW^T|$

Problem 1.3

GAN

$By\ min_G max_D (E_{p_{data(x)}}[logD(X)] + E_{p(x)}[log(l - D(G(Z)))])$
$D(x) = argmax_D E_{p_{data(x)}}[logD(X)] + E_{p(x)}[log(l - D(G(Z)))]$

$= argmax_D \int p_{data(x)}[logD(x)]dx + \int p(x)[log(l - D(G(z)))]dz$

$= argmax_D \int p_{data(x)}[logD(x)]dx + \int p(G^{-1}(x))[log(l - D(G(x)))] \frac{1}{(G^{-1}(x))'} dx$

$= argmax_D \int p_{data(x)}[logD(x)]dx + p(G^{-1}(x))[log(l - D(G(x)))] \frac{1}{(G^{-1}(x))'} dx$

$= argmax_D D(x)^{p_{data(x)}} (1 - D(x))^{p_{model(x)}}$

$\Rightarrow\ D(x) = \frac{p_{data(x)}}{p_{data(x)} + p_{model(x)}}$得证

2 Reinforcement Learning

Problem 2.1

Policy Improvement Theorem

在$Markov\ Decision\ Process$，有$policy\ \pi$，可用下述方式将$\pi$更新至$\pi^{'}$：
$\forall s \in S,\ \pi^{'}(s) = argmax_a q^{\pi}(s,a)$
欲证$\forall s \in S,\ v^{\pi^{'}}(s) \geq v^{\pi}(s)$

下证：
$By\ definition\ of\ state-value\ funciton\ and\ action\ function,\ v^{\pi}(s) = \sum_a \pi(a|s)q^{\pi}(s,a) \leq q^{\pi}(s,\pi^{'}(s$
$\Rightarrow v^{\pi}(s) \leq \pi^{'}(s)$
且$q^{\pi}(s,\pi^{'}(s)) = E_{\pi}[R_{t+1} + \ldots | S_t = s, A_t = \pi^{'}(s]$
故证$q^{\pi}(s,\pi^{'}(s)) = \pi^{'}(s) \leq v^{\pi}(s)$，则$v^{\pi^{'}}(s) \geq v^{\pi}(s)$即可得证

现令$\pi^{'}(s) = f(1),\ v^{\pi}(s) = f(\infty)$，并求证$f(k) \leq f(k+1)$
$f(k) = \sum_{i=0}^{k-1} \gamma^i E_{\pi'}[R_t + i + 1 | S_t = s] + \sum_{s_1} \ldots \sum_{s_k} \sum_a P_{ss_1}^{\pi^{'}(s)} P_{s_1 s_2}^{\pi^{'}(s_1)} \ldots P_{s_{k-1} s_k}^{\pi^{'}(s_{k-1})} \pi(a|s_k) E_{\pi}[\sum_{j=k}^{\infty} \gamma^j R_{t+}$
$= \sum_{i=0}^{k-1} \gamma^k E_{\pi'}[R_t + i + 1 | S_t = s] + \sum_{s_1} \ldots \sum_{s_k} \sum_a P_{ss_1}^{\pi^{'}(s)} P_{s_1 s_2}^{\pi^{'}(s_1)} \ldots P_{s_{k-1} s_k}^{\pi^{'}(s_{k-1})} \pi(a|s_k) \gamma^k q^{\pi}(s_k, a)$
$\leq \sum_{i=0}^{k-1} \gamma^k E_{\pi'}[R_t + i + 1 | S_t = s] + \sum_{s_1} \ldots \sum_{s_k} \sum_a P_{ss_1}^{\pi^{'}(s)} P_{s_1 s_2}^{\pi^{'}(s_1)} \ldots P_{s_{k-1} s_k}^{\pi^{'}(s_{k-1})} \gamma^k q^{\pi}(s_k, \pi^{'}(s_k)) = f(k+1$
$\Rightarrow f(1) \leq f(\infty)$
$\Rightarrow \pi^{'}(s) \leq v^{\pi}(s)$
$\Rightarrow v^{\pi}(s) \leq \pi^{'}(s) \leq v^{\pi}(s)$
$\Rightarrow v^{\pi}(s) \leq v^{\pi}(s)$得证