Statistical Machine Learning
Homework 4

Instructor: Prof. Jun Zhu

May 11, 2021

Requirements:

- We recommend that you typeset your homework using appropriate software such as LaTeX. If you submit your handwritten version, please make sure it is cleanly written up and legible. The TAs will not invest undue effort to decrypt bad handwritings.
- There are optional problems in the assignments. We will give bonus points to those who succeed in solving theses problems.
- Please finish your homework independently. In addition, **you should write in your homework the set of people with whom you collaborated**.

1 Generative Models

Problem 1.1 (IWAE, see P29 in the DGM slides; 2pt)**.** The variational lower bound in IWAE is

$$\mathrm{IWAE}(K) = \mathbb{E}_{q(\mathbf{z}|\mathbf{x};\boldsymbol{\phi})}\left[\log\left(\frac{1}{K}\sum_{k=1}^{K}\frac{p(\mathbf{z}^{(k)},\mathbf{x};\boldsymbol{\theta})}{q(\mathbf{z}^{(k)}|\mathbf{x};\boldsymbol{\phi})}\right)\right].$$

When $K = 1$, it recovers the lower bound of VAE. Please show that

1. $\mathrm{IWAE}(1) \leq \mathrm{IWAE}(K) \leq \log p(\mathbf{x};\boldsymbol{\theta})$.
2. $\lim_{K\to\infty}\mathrm{IWAE}(K) = \log p(\mathbf{x};\boldsymbol{\theta})$. (You only need to give a intuitive explanation instead of a complete and rigorous proof. The latter may require $\log\frac{p(\cdot,\mathbf{x};\boldsymbol{\theta})}{q(\cdot|\mathbf{x};\boldsymbol{\phi})}$ to be bounded.)
3. (Bonus, 0.1pt) If $K < L$, $\mathrm{IWAE}(K) \leq \mathrm{IWAE}(L)$.

Remark 1.1. The above shows that $\mathrm{IWAE}(K)$ is a tighter lower bound of $\log p(\mathbf{x};\boldsymbol{\theta})$ than the VAE lower bound. If K is large, then the lower bound could be rather tight even if the variational family q cannot approximate the true posterior well. However, there are other issues in IWAE. Interested readers can refer to the paper *Tighter Variational Bounds are Not Necessarily Better*.

Problem 1.2 (Affine VAE and probabilistic PCA; 1pt)**.** Consider the following affine VAE: the prior $p(\mathbf{z})$ is the κ-dimensional standard normal distribution, the decoder is $p(\mathbf{x}|\mathbf{z}) = \mathcal{N}(\boldsymbol{W}\mathbf{z}+\boldsymbol{b},\lambda\boldsymbol{I})$. where $(\boldsymbol{W},\boldsymbol{b},\lambda)$ are parameters with appropriate dimensions, and the encoder is $q(\mathbf{z}|\mathbf{x}) = \mathcal{N}(\boldsymbol{\mu}(\mathbf{x}),\boldsymbol{S}(\mathbf{x})\boldsymbol{S}(\mathbf{x})^\top)$ where $\boldsymbol{\mu},\boldsymbol{S}$ are arbitrarily flexible functions.

1. Write down the ELBO given observations $\{\mathbf{x}^{(1)},\ldots,\mathbf{x}^{(n)}\}$. Derive the optimal $\boldsymbol{S},\boldsymbol{\mu}$, and show that for any encoder parameters $(\boldsymbol{W},\boldsymbol{b},\lambda)$, the "collapsed objective", $\mathrm{ELBO}(\boldsymbol{W},\boldsymbol{b},\lambda,\boldsymbol{\mu}^*,\boldsymbol{S}^*)$, is equivalent to

$$-\sum_i(\mathbf{x}^{(i)}-\boldsymbol{b})^\top(\boldsymbol{W}\boldsymbol{W}^\top+\lambda\boldsymbol{I})^{-1}(\mathbf{x}^{(i)}-\boldsymbol{b}) - n\log|\boldsymbol{W}\boldsymbol{W}^\top+\lambda\boldsymbol{I}|, \tag{1}$$

modulo constants. Eq.(1) is the probabilistic PCA objective. (Hint: You may use the fact that for any $\boldsymbol{A}\in\mathbb{R}^{d\times d}$, $\log|\boldsymbol{A}\boldsymbol{A}^\top| = \inf_{\boldsymbol{\Gamma}\succ 0}(\mathrm{Tr}(\boldsymbol{A}\boldsymbol{A}^\top\boldsymbol{\Gamma}^{-1})+\log|\boldsymbol{\Gamma}|-d)$, where $\boldsymbol{\Gamma}\succ 0$ means $\boldsymbol{\Gamma}$ is positive definite. Alternatively there may be a simpler proof.)

2. (Bonus, 0pt) Eq.(1) has no local optima which is unsurprising. If we further restrict $\boldsymbol{S}$ to be diagonal, will the new collapsed objective have local optimas?

Problem 1.3 (GAN, see P42 in the DGM slides; 1pt)**.** The minimax objective function in GAN is

$$\min_G \max_D \left(\mathbb{E}_{p_{\text{data}}(\mathbf{x})}[\log D(\mathbf{x})] + \mathbb{E}_{p(\mathbf{z})}[\log(1 - D(G(\mathbf{z}))]\right).$$

Prove that the optimal solution of D is

$$D(\mathbf{x}) = \frac{p_{\text{data}}(\mathbf{x})}{p_{\text{data}}(\mathbf{x}) + p_{\text{model}}(\mathbf{x})},$$

where p_{model} is the distribution of $G(\mathbf{z})$ when $\mathbf{z} \sim p(\mathbf{z})$.

2 Reinforcement Learning

Problem 2.1 (Policy Improvement Theorem, 1pt)**.** In a Markov Decision Process, suppose we currently have a policy π. In policy iteration, we update π to π' such that

$$\forall s \in \mathcal{S}, \pi'(s) = \operatorname*{argmax}_a q^{\pi}(s, a).$$

Please prove that

$$\forall s \in \mathcal{S}, v^{\pi'}(s) \geq v^{\pi}(s).$$