HW3

October 21, 2019

1 GR5242 HW3 Zihan Zhou (zz2573)

1.1 Problem 1

1.1.1 (a)

Each time we generate a random sample x_i from the density p, generate another random sample y_i from Uniform$(0, p(x_i))$. Then (x_i, y_i) is a point that is distributed uniformly on the area under the curve p.

1.1.2 (b)

If the training data is linear separable or a large proportion of the data is linear separable, the log-likelihood can be increased by scaling the length of $\mathbf{v}$ without moving the decision boundary. As the length of $\mathbf{v}$ gets larger, the sigmoid function becomes sharper and more similar to the indicator function. Therefore we might lose the smoothness of sigmoid and overfitting occurs.

1.1.3 (c)

X_n and X_2 are dependent if we do not know the information about $Z_1, \ldots, Z_n$. For example, in the dishonest casino example, if $X_2 = 6$, it might be more likely that the loaded dice is used at Z_2 and thus affect how X_3 is distributed. Once we observe what Z_3 is, X_3 can be sampled from $P(\cdot|Z_3)$ and is independent of X_2.

1.2 Problem 2

Here $\phi_1(x) = \phi_2(x) = x$, $\phi(x) = \mathbb{I}(x = 0)$.

1.3 Problem 3

The observed data consists of one data point in category 3, because the dirichlet posterior has a shifted mean and a larger concentration on the category 3 side.

1.4 Problem 4

This neural network represents $f(\mathbf{x}) = \mathbb{I}\{\mathbf{w}'\mathbf{x} - c \geq 0\}$. Thus $f(\mathbf{x}) = \mathbb{I}\{-\frac{3}{\sqrt{2}} - \frac{1}{2\sqrt{2}} \geq 0\} = -1$, $f(\mathbf{x}') = \mathbb{I}\{\frac{2}{2\sqrt{2}} - \frac{1}{2\sqrt{2}} \geq 0\} = 1$.

1.5 Problem 5

1.5.1 (a)

$q(x) = 1, \ \forall x \in [0,1]$. The optimal choice for M makes $\sup_{x\in[0,1]} \widetilde{p}(x) = \widetilde{p}(1) = \frac{2}{M} = 1$, i.e. $M = 2$.

1.5.2 (b)

The proposal distribution q forms a square with four vertices $(0,0), (0,1), (1,1), (1,0)$, and $\widetilde{p}$ is the diagonal segment from $(0,0)$ to $(1,1)$. The acceptance probability is the fraction of two areas, which is $\frac{1}{2}$. Therefore we need to sample $n = 2m$ times to get m valid samples on average.

1.5.3 (c)

The importance sampling estimate of the mean is $\hat{\mu} = \sum_{i=1}^{n} X_i \widetilde{p}(X_i) = \sum_{i=1}^{n} \frac{1}{2} X_i^2$, $X_1, \ldots, X_n \sim$ *iid* Uniform$(0,1)$. Therefore

$$\begin{aligned}
\mathrm{Var}(\hat{\mu}) &= \mathrm{Var}(\sum_{i=1}^{n} \frac{1}{2} X_i^2) \\
&= \frac{1}{4} \sum_{i=1}^{n} \mathrm{Var}(X_i^2) \\
&= \frac{n}{4}(E(X_1^4) - E^2(X_1^2)) \\
&= \frac{n}{4}(\frac{1}{5} - \frac{1}{9}) = \frac{n}{45}
\end{aligned}$$

1.6 Problem 6

From the formula we know that the normalizing constant is not needed, because if we use $\widetilde{p}(x) = \frac{1}{Z} p(x)$ instead of $p(x)$, the normalizing constant would cancel out in the fraction and still we get $\mathbb{P}(X_i \leq x | A) = \int_{-\infty}^{x} p(x_i) dx_i$.

1.6.1 (a)

$$\begin{aligned}
\mathbb{P}(X_i \leq x | R) = \mathbb{P}(X_i \leq x | U_i > \tfrac{p(X_i)}{kr(X_i)}) &= \frac{\mathbb{P}(X_i \leq x, U_i > \frac{p(X_i)}{kr(X_i)})}{\mathbb{P}(U_i > \frac{p(X_i)}{kr(X_i)})} \\
&= \frac{\int_{-\infty}^{x} \left[1 - \int_0^{p(x_i)/kr(x_i)} dy\right] r(x_i) dx_i}{\int_{-\infty}^{\infty} \left[1 - \int_0^{p(x_i)/kr(x_i)} dy\right] r(x_i) dx_i} \\
&= \frac{\int_{-\infty}^{x} r(x_i) dx_i - \frac{1}{k} \int_{-\infty}^{x} p(x_i) dx_i}{\int_{-\infty}^{\infty} r(x_i) dx_i - \frac{1}{k} \int_{-\infty}^{\infty} p(x_i) dx_i} \\
&= \frac{k}{k-1} \int_{-\infty}^{x} r(x_i) dx_i - \frac{1}{k-1} \int_{-\infty}^{x} p(x_i) dx_i
\end{aligned}$$

$\Rightarrow X_i | Z_i = R$ has density $\frac{k}{k-1} r(\cdot) - \frac{1}{k-1} p(\cdot)$.

1.6.2 (b)

An importance sampler is a weighted average of samples, whose importance weights are determined by the fraction of target density $p(\cdot)$ and proposal density $q(\cdot)$. As we proved above, in this case, if $Z_i = A$, $X_i \sim q(\cdot) = p(\cdot)$, the importance weight is 1; if $Z_i = R$, $X_i \sim q(\cdot) = \frac{kr(\cdot)-p(\cdot)}{k-1}$, the importance weight is $\frac{(k-1)p(X_i)}{kr(X_i)-p(X_i)}$.

[ ]:

```

```