MathDNN Homework 10

Department of Computer Science and Engineering
2021-16988 Jaewan Park

Problem 1

Expansion over the expectation operator gives the following.

$$\begin{aligned}
\nabla_\phi \mathbb{E}_{Z\sim q_\phi(z)}\left[\log\left(\frac{h(Z)}{q_\phi(Z)}\right)\right] &= \nabla_\phi \int_{\mathbb{R}^k} \log\left(\frac{h(z)}{q_\phi(z)}\right) q_\phi(z)dz \\
&= \int_{\mathbb{R}^k} \left(\nabla_\phi \log\left(\frac{h(z)}{q_\phi(z)}\right) q_\phi(z) + \log\left(\frac{h(z)}{q_\phi(z)}\right)\nabla_\phi q_\phi(z)\right)dz \\
&= \int_{\mathbb{R}^k} \left(-\frac{\nabla_\phi q_\phi(z)}{q_\phi(z)} q_\phi(z) + \log\left(\frac{h(z)}{q_\phi(z)}\right)\nabla_\phi q_\phi(z)\right)dz \\
&= \int_{\mathbb{R}^k} \log\left(\frac{h(z)}{q_\phi(z)}\right)\nabla_\phi q_\phi(z)dz = \int_{\mathbb{R}^k} \log\left(\frac{h(z)}{q_\phi(z)}\right)\frac{\nabla_\phi q_\phi(z)}{q_\phi(z)} q_\phi(z)dz \\
&= \mathbb{E}_{Z\sim q_\phi(z)}\left[(\nabla_\phi \log q_\phi(Z))\log\left(\frac{h(Z)}{q_\phi(Z)}\right)\right]
\end{aligned}$$

Problem 2

Since $C = \left\{x\in\mathbb{R}^2 \mid x_1 = a,\, 0\le x_2\le 1\right\}$, we obtain the following.

$$\begin{aligned}
\Pi_C(y) &= \underset{x\in C}{\operatorname{argmin}} \|x-y\|^2 = \underset{x\in C}{\operatorname{argmin}} \left\{(a-y_1)^2 + (x_2-y_2)^2\right\} = \underset{x\in C}{\operatorname{argmin}}\ (x_2-y_2)^2 \\
&= \begin{bmatrix} a \\ \min\{\max\{y_2, 0\}, 1\}\end{bmatrix} \quad (\because 0\le x_2\le 1)
\end{aligned}$$

Problem 4

(a) Since $f_1(x) = Ax = PL(U+\operatorname{diag}(s))x$, we obtain the following.

$$\begin{aligned}
\log\left|\frac{\partial f_1}{\partial x}\right| &= \log|A| = \log|PL(U+\operatorname{diag}(s))| = \log(|P||L||U+\operatorname{diag}(s)|) = \log|\operatorname{diag}(s)| \\
&= \sum_{i=1}^{C}\log|s_i|
\end{aligned}$$

(b) Different choices of reshape results differ by permutations, i.e., multiplying a permutation matrix to one result of reshape gives another possible result of reshape. Let $x, y\in\mathbb{R}^{abc}$, and $P\in\mathbb{R}^{abc}$ a permutation matrix. We obtain the following.

$$\left|\frac{\partial(Py)}{\partial(Px)}\right| = \left|\frac{\partial(Py)}{\partial y}\frac{\partial y}{\partial x}\frac{\partial x}{\partial(Px)}\right| = |P|\left|\frac{\partial y}{\partial x}\right|\frac{1}{|P|} = \left|\frac{\partial y}{\partial x}\right|$$

Therefore Jacobian determinants between two vectors or between their permutations give the same result. Therefore the choice of reshape does not matter in the given calculation.

(c) Select the reshape operator as reshaping $X \in \mathbb{R}^{C\times m\times n}$ into

$$(X_{1,1,1}, X_{2,1,1}, \cdots, X_{C,1,1}, X_{1,2,1}, X_{2,2,1}, \cdots, X_{C,2,1}, \cdots, X_{1,m,n}, X_{2,m,n}, \cdots, X_{C,m,n}).$$

Then we know

$$f_2(X \mid P, L, U, s).\text{reshape}(Cmn) = \begin{bmatrix} A & & 0 \\ & \ddots & \\ 0 & & A \end{bmatrix} X.\text{reshape}(Cmn)$$

where the As in the diagonal are repeated mn times. Therefore using the result of (a), we obtain the following.

$$\begin{aligned} \log\left|\frac{\partial f_2(X \mid P, L, U, s)}{\partial X}\right| &= \log\left|\frac{\partial f_2(X \mid P, L, U, s).\text{reshape}(Cmn)}{\partial X.\text{reshape}(Cmn)}\right| \\ &= \log\begin{vmatrix} A & & 0 \\ & \ddots & \\ 0 & & A \end{vmatrix} = \log|A|^{mn} = mn\log|A| = mn\sum_{i=1}^{C}\log|s_i| \end{aligned}$$

(d) Select the reshape operator as reshaping $X \in \mathbb{R}^{2C\times m\times n}$ into

$$(X_{1,1,1}, \cdots, X_{C,1,1}, \cdots, X_{1,m,n}, \cdots, X_{C,m,n}, X_{C+1,1,1}, \cdots, X_{2C,1,1}, \cdots, X_{C+1,m,n}, \cdots, X_{2C,m,n}).$$

Then we know

$$Z.\text{reshape}(2Cmn) = \begin{bmatrix} I & & & 0 \\ & A & & \\ & & \ddots & \\ 0 & & & A \end{bmatrix} X.\text{reshape}(2Cmn)$$

where $I \in \mathbb{R}^{Cmn\times Cmn}$ and the As in the diagonal are repeated mn times. Therefore using the result of (a), we obtain the following.

$$\begin{aligned} \log\left|\frac{\partial Z}{\partial X}\right| &= \log\left|\frac{\partial Z.\text{reshape}(2Cmn)}{\partial X.\text{reshape}(2Cmn)}\right| \\ &= \log\begin{vmatrix} I & & & 0 \\ & A & & \\ & & \ddots & \\ 0 & & & A \end{vmatrix} = \log|A|^{mn} = mn\log|A| = mn\sum_{i=1}^{C}\log|s_i| \end{aligned}$$