Solution to Homework 08

Name: Chen Shen

NetID: cs5236

1.

(a) With the dimensions of W, $0 \leq k_1, k_2 < 2$.

(b) The size ill be

$$(6-2+1) \times (5-2+1) = 5 \times 4$$

(c) Notice that

$$Z[i,j] = X[i,j] + X[i+1,j] - X[i,j+1] - X[i+1,j+1]$$

So $Z[i,j]$ will reach the largest positive value when the change between two column is largest positive. For example, $(i,j) = (1,3)$.

$$Z[1,3] = X[1,3] + X[2,3] - X[1,4] - X[2,4] = 3 + 3 - 0 - 0 = 6$$

(d) Notice that

$$Z[i,j] = X[i,j] + X[i+1,j] - X[i,j+1] - X[i+1,j+1]$$

So $Z[i,j]$ will reach the largest negative value when the change between two column is largest negative. For example, $(i,j) = (1,0)$.

$$Z[1,0] = X[1,0] + X[2,0] - X[1,1] - X[2,1] = 0 + 0 - 3 - 3 = -6$$

(e) For example, $(i,j) = (1,1)$.

$$Z[1,1] = X[1,1] + X[2,1] - X[1,2] - X[2,2] = 3 + 3 - 3 - 3 = 0$$

2.

(a) Since W is 3×3 and there are 20 channels, the shape of Z and U should be

$$(48-3+1) \times (64-3+1) \times 20 = 46 \times 62 \times 20$$

(b) Since W has shape $(3,3,10,20)$, there are 10 input channels and 20 output channels.

(c) The number of mutiplications are

$$46 \cdot 62 \cdot 20 \cdot 3 \cdot 3 \cdot 10 = 51333000$$

(d) Since W has shape $(3,3,10,20)$ and b has shape (20), the total number of parameters should be

$$3 \cdot 3 \cdot 10 \cdot 20 + 20 = 1820$$

3.

(a) By chain rule,

$$\frac{\partial J}{\partial Z[i,j,m]} = \frac{\partial J}{\partial U[i,j,m]}\frac{\partial U[i,j,m]}{\partial Z[i,j,m]} = \frac{\partial J}{\partial U[i,j,m]}\frac{e^{-Z[i,j,m]}}{(1+e^{-Z[i,j,m]})^2} = \frac{\partial J}{\partial U[i,j,m]}U[i,j,m](1-U[i,j,m])$$

(b)

$$\frac{\partial J}{\partial W[k_1,k_2,n,m]} = \sum_{i,j}\frac{\partial J}{\partial Z[i,j,m]}\frac{\partial Z[i,j,m]}{\partial W[k_1,k_2,n,m]} = \sum_{i,j}\frac{\partial J}{\partial Z[i,j,m]}X[i+k_1,j+k_2,n]$$

(c) First, let

$$i' = i-k_1, j' = j-k_2$$

So

$$Z[i',j',m] = \sum_{k_1}\sum_{k_2}\sum_{n} W[i-i',j-j',n,m]X[i,j,n] + b[m]$$

Thus

$$\frac{\partial Z[i',j',m]}{\partial X[i,j,n]} = W[i-i',j-j',n,m]$$

By chain rule,

$$\frac{\partial J}{\partial X[i,j,n]} = \sum_{i'}\sum_{j'}\sum_{m}\frac{\partial J}{\partial Z[i',j',m]}\frac{\partial Z[i',j',m]}{\partial X[i,j,n]} = \sum_{i'}\sum_{j'}\sum_{m}\frac{\partial J}{\partial Z[i',j',m]}W[i-i',j-j',n,m]$$

Replace i', j', with $i-k_1$, $j-k_2$, then we have

$$\frac{\partial J}{\partial X[i,j,n]} = \sum_{k_1}\sum_{k_2}\sum_{m}\frac{\partial J}{\partial Z[i-k_1,j-k_2,m]}W[k_1,k_2,n,m]$$