Intro to ML

HW8

`Tianyuunyang Wang` `N14855366` `tw1719`

1.

1. Let X and W be arrays,

$$X = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 3 & 3 & 3 & 0 \\ 0 & 3 & 3 & 3 & 0 \\ 0 & 3 & 2 & 3 & 0 \\ 0 & 3 & 2 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad W = \begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix}.$$

Let Z be the 2D convolution (without reversal):

$$Z[i,j] = \sum_{k_1,k_2} W[k_1,k_2]X[i+k_1, j+k_2]. \tag{1}$$

Assume that the arrays are indexed starting at $(0,0)$.

(a) What are the limits of the summations over k_1 and k_2 in (1)?

(b) What is the size of the output $Z[i,j]$ if the convolution is computed only on the *valid* pixels (i.e. the pixel locations (i,j) where the summation in (1) does not exceed the boundaries of W or X).

(c) What is the largest positive value of $Z[i,j]$ and state one pixel location (i,j) where that value occurs.

(d) What is the largest negative value of $Z[i,j]$ and state one pixel location (i,j) where that value occurs.

(e) Find one pixel location where $Z[i,j] = 0$.

> - (a)
> $0 \le (k_1 + k_2) \le 2$
> - (b)
>   If the convolution is computed only on the valid pixels, the size of Z[i,j] will be (5,4)
> - (c)
>   The largest positive value of Z[i,j] is 6, you can find it at point (1,3), (2,3), or (3,3)
> - (d)
>   The largest negative value of Z[i,j] is 6, you can find it at point (1,0), (2,0), or (3,0)
> - (e)
>   Location (0,1), (0,2), (1,1), (1,2)

2. Suppose that a convolutional layer of a neural network has an input tensor $X[i,j,k]$ and computes an output via a convolution and ReLU activation,

$$Z[i,j,m] = \sum_{k_1}\sum_{k_2}\sum_{n} W[k_1,k_2,n,m]X[i+k_1,j+k_2,n] + b[m],$$
$$U[i,j,m] = \max\{0, Z[i,j,m]\}.$$

for some weight kernel $W[k_1,k_2,n,m]$ and bias $b[m]$. Suppose that X has shape (48,64,10) and W has shape (3,3,10,20). Assume the convolution is computed on the *valid* pixels.

(a) What are the shapes of Z and U?

(b) What are the number of input channels and output channels?

(c) How many multiplications must be performed to compute the convolution in that layer?

(d) If W and b are to be learned, what are the total number of trainable parameters in the layer?

- (a)

  Size of Z and U: (46,62,20)
- (b)

  input channel: 10, output channel: 20
- (c)

  One pixel needs $3*3$ times.

  Thus we need $3*3*46*62*20*10 = 5.13*10^6$ multiplications.
- (d)

  The total number of trainable parameters is : $3*3*10*20+20 = 1820$

3. Suppose that a convolutional layer in some neural network is described as a linear convolution followed by a sigmoid activation,

$$Z[i,j,m] = \sum_{k_1}\sum_{k_2}\sum_{n} W[k_1,k_2,n,m]X[i+k_1,j+k_2,n] + b[m],$$
$$U[i,j,m] = 1/(1+\exp(-Z[i,j,m])).$$

where $X[i,j,n]$ is the input of the layer and $U[i,j,m]$ is the output. Suppose that during back-propagation, we have computed the gradient $\partial J/\partial U$ for some loss function J. That is, we have computed the components $\partial J/\partial U[i,j,m]$. Show how to compute the following:

(a) The gradient components $\partial J/\partial Z[i,j,m]$.

(b) The gradient components $\partial J/\partial W[k_1,k_2,n,m]$.

(c) The gradient components $\partial J/\partial X[i,j,n]$.

- (a)

  $\frac{\alpha U}{\alpha Z} = \frac{1}{1+e^{-z}}' = \frac{e^{-z}}{(1+e^{-z})^2}$

  According to the chain rule, $\frac{\alpha J}{\alpha Z} = \frac{\alpha J}{\alpha U} * \frac{\alpha U}{\alpha Z} = \frac{\alpha J}{\alpha U} * \frac{e^{-z}}{(1+e^{-z})^2}$
- (b)

  $\frac{\alpha Z}{\alpha W} = \sum_i \sum_j X[i+k_1, j+k_2, n]$

  According to the chain rule,

  $\frac{\alpha J}{\alpha W} = \frac{\alpha J}{\alpha Z} * \frac{\alpha Z}{\alpha W} = \sum_i \sum_j X[i+k_1, j+k_2, n] * \frac{\alpha J}{\alpha U} * \frac{e^{-z}}{(1+e^{-z})^2}$

- (c)
  $\frac{\alpha J}{\alpha X} = \frac{\alpha J}{\alpha Z} * \frac{\alpha Z}{\alpha X}$
  Suppose we have two new variables l_1, l_2, where $k_1 = l_1 - i, k_2 = l_2 - j.$
  $Z[i,j,m] = \sum_{l_1-i} \sum_{l_2-j} \sum_n W[l_1 - i, l_2 - j, n, m]X[l_1, l_2, n] + b[m]$
  $\frac{\alpha Z}{\alpha X} = W[l_1 - i. l_2 - j, n, m]$
  $\frac{\alpha J}{\alpha X} = \sum_i \sum_j \frac{\alpha J}{\alpha U} * \frac{e^{-z}}{(1+e^{-z})^2} * W[l_1 - i. l_2 - j, n, m]$