习题 5-1

(1)

$$
\begin{aligned}
d^2x(t) &= d(dx(t)) \\
&\approx d(x(t+\Delta t) - x(t)) \\
&\approx x(t+\Delta t) - x(t) - (x(t) - x(t-\Delta t)) \\
&= x(t+\Delta) + x(t-\Delta t) - 2x(t) \\
&\approx x(t+1) + x(t-1) - 2x(t)
\end{aligned}
$$

(2)

$$
\begin{aligned}
\nabla^2 f(x,y) =& d_x^2 f(x,y) + d_y^2 f(x,y) \\
\approx& f(x+1,y) + f(x-1,y) - 2f(x,y) \\
&+ f(x,y+1) + f(x,y-1) - 2f(x,y)
\end{aligned}
$$

so,

$$
w = \begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix}
$$

习题 5-3

Efficiently change the dimensions of feature.

Ref: What does 1x1 convolution mean in a neural network?

习题 5-4

(1) 3x3

time complexity:

$$\begin{aligned}&(100*100)*(3*3)*256*256\\=&5.89824e9\end{aligned}$$

space complexity:

$$\begin{aligned}&100*100*256\\&+(3*3)*256*256\\&+100*100*256\\=&2560000+589824+2560000\\=&5.709824e6\end{aligned}$$

(2) 1x1 then 3x3

time complexity:

$$\begin{aligned}&(100*100)*(1*1)*256*64\\&+(100*100)*(3*3)*64*256\\=&1.6384e9\end{aligned}$$

space complexity:

$$\begin{aligned}&(100*100)*256\\&+(1*1)*256*64\\&+(100*100)*64\\&+(3*3)*64*256\\=&3.36384e6\end{aligned}$$

习题 5-5

let

$$x = \begin{bmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \end{bmatrix}$$

$$w = \begin{bmatrix} w_{11} & w_{12} \\ w_{21} & w_{22} \end{bmatrix}$$

$$x' = [x_{11}, x_{12}, x_{13}, x_{21}, x_{22}, x_{23}, x_{31}, x_{32}, x_{33}]^T$$

then,

$$\begin{aligned} z &= w \otimes x \\ &= \begin{bmatrix} w_{11} & x_{12} & 0 & w_{21} & w_{22} & 0 & 0 & 0 & 0 \\ 0 & w_{11} & x_{12} & 0 & w_{21} & w_{22} & 0 & 0 & 0 \\ 0 & 0 & 0 & w_{11} & x_{12} & 0 & w_{21} & w_{22} & 0 \\ 0 & 0 & 0 & 0 & w_{11} & x_{12} & 0 & w_{21} & w_{22} \end{bmatrix} x' \\ &= Cx' \end{aligned}$$

习题 5-6

(1)

let

$$x = [x_1, \cdots, x_D]^T$$

$$x_m = \max(x)$$

then,

$$\begin{aligned} \frac{\partial y}{\partial x} &= \frac{\partial x_m}{\partial x} \\ &= [0, \cdots, 1, \cdots, 0]^T \end{aligned}$$

(2)

This post gives some discussion: argmax differentiable?