IDL Exercise 2.

Daniel Ruderman, David Ponarovsky

April 2025

1 Theoretical Questions:

1.1 LTI.

Show that a convolution with respect to any filter h is time/space invariant.

Solution. Let's denote by $\mathcal{L}$: funcs $\to$ funcs the convolution by the filter h, namely $\mathcal{L}[f] = f * h = \sum_x h(x)f(y-x)$. To show that $\mathcal{L}$ is an LTI operation, we have to show that $\mathcal{L}[f(x-t)](y) = \mathcal{L}[f](y-t)$. (Note that the linearity is obtained for free by the linearity of convolution, so only the time-invariant part is left). So:

$$\begin{aligned}\mathcal{L}[f(x-t)](y) &= f(x-t) * h = \sum_x h(x)f(\overbrace{y-t}^{\tau}-x) = \sum_x h(x)f(\tau - x) \\ &= \mathcal{L}[f](\tau) = \mathcal{L}[f](y-t)\end{aligned}$$

And we got that the convolution is an LTI operation.

1.2 TI.

Explain whether each of the following layers are time/space invariant or not:

1. Additive constant.

   Solution. TRUE.

   $$\mathcal{L}[f] = f(x) + c \Rightarrow \mathcal{L}[f(x-t)] = f(x-t) + c = \mathcal{L}[f](x-t)$$

2. Pointwise nonlinearity (such as ReLU)

   Solution. TRUE. Since the operator acts pointwise, we can denote it by $\mathcal{L}[f](x) = g(f(x))$. Here we think of the input x as the coordinate (that's consistent with the definition of the convolution from the lecture). Now:

   $$\Rightarrow \mathcal{L}[f(x-t)] = g(f(x-t)) = \mathcal{L}[f](x-t)$$

3. Strided pooling by a factor > 1

   Solution. **FALSE.** Consider the 2-factor pooling, namely $\mathcal{L}[f](x) = f(2x)$, (equivalent to taking only the even coordinates). Now consider the action of shifting the function f by a single time unit, The pooling gives only the odd coordinates, namely:

$$\mathcal{L}[f(x+1)] = f(2x+1) \neq \mathcal{L}[f](x+1) = f(2x+2)$$

4. As a result, is a CNN composed of all these operators (+convolution) time invariant?

   Solution. **FALSE.** The pooling is not TI and therfore a CNN isn't a TI.

1.3 Layers' Jacobian.

Calculate the Jacobian matrix of the following layers:

Recall the Jacobian defitnion:

$$J = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \frac{\partial f_m}{\partial x_2} & \cdots & \frac{\partial f_m}{\partial x_n} \end{bmatrix}$$

1. Additive bias vector

   Solution. In that case $f_i(x) = x_i + b_i$ and therefore:

$$\frac{\partial f_i}{\partial x_j} = \frac{\partial}{\partial x_j}(x_i + b_i) = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{else} \end{cases}$$

2. General Matrix multiplication

   Solution. In that case $f_i(x) = \sum_k A_{ik} x_k$ and therefore:

$$\frac{\partial f_i}{\partial x_j} = \frac{\partial}{\partial x_j}\left(\sum_k A_{ik} x_k\right) = A_{ij} \Rightarrow J = A$$

3. Convolution layer

   Solution. By the previous section and since convolution is just a multiplication by all the cyclic shifts of a vector (namely a matrix whose first row is h, the second is h shifted to the right, and so on), we get that the Jacobian will be the convolution matrix. In other words, J_{ij} would be the $(i-j)$th value of the filter. Namely, if the filter is h, then $J_{ij} = h_{i-j}$.