Homework Week 2

Mathematics of deep learning
MASH & IASD 2026

Lecturer: Bruno Loureiro, bruno.loureiro@di.ens.fr

Instructions: This homework is **due on Monday 02/02/2026**. Please upload your solutions in a PDF file named HW2_Nom_Prenom.pdf here. Formats accepted: PDF (LaTeX or a **readable** scan of handwritten solutions).

1 Exercises

Exercise 1.

Show that the $L^p(\mu)$ norm:

$$||f||_p := \left(\int \mu(\mathrm{d}\boldsymbol{x}) |f(\boldsymbol{x})|^p \right)^{1/p} \tag{1}$$

is an increasing function of $p \in [1, \infty]$:

$$||f||_{L^\infty(\mu)} \geq \cdots \geq ||f||_{L^2(\mu)} \geq ||f||_{L^1(\mu)} \tag{2}$$

where we recall $||f||_{L^\infty(\mu)} = \sup_{x \in \mathrm{supp}(\mu)} |f(\boldsymbol{x})|$. Conclude that we have the inclusion:

$$L^\infty(\mu) \subset \cdots \subset L^2(\mu) \subset L^1(\mu) \tag{3}$$

Exercise 2.

Let μ denote a probability measure in $\mathbb{R}$ and $[a, b] \subset \mathbb{R}$ a compact subset of your choice. Give examples of functions $f, g : \mathbb{R} \to \mathbb{R}$ such that:

$$||f||_{L^\infty(\mu)} \leq ||f||_{L^\infty([a,b])}, \qquad \text{and } ||g||_{L^\infty(\mu)} \geq ||g||_{L^\infty([a,b])} \tag{4}$$

Note: In the above, we want $[a, b]$ and μ to be the same in both inequalities.

Exercise 3.

Consider the following continuous function on $\mathbb{R}$:

$$g(x) = \begin{cases} 0 & x < -1/2 \\ x + 1/2 & x \in [-1/2, 1/2] \\ 1 & x > 1/2 \end{cases} \tag{5}$$

How many neurons p are needed to approximate g within a precision $\epsilon > 0$ on the compact set $[-1, 1] \subset \mathbb{R}$ using the two-layer neural network with step-size activation from Proposition 2 in the lectures? Show that we could do as well by using fewer neurons if we adapt the partition to the function.

Exercise 4.

Show that:

$$\inf_{f_\theta \in \mathcal{F}_{\mathrm{relu},1}} \sup_{x \in \mathbb{R}} |f_\theta(x) - \sin(x)| \geq 1 \tag{6}$$

where $\mathcal{F}_{\mathrm{relu},1}$, the class of two-layer neural networks over $\mathbb{R}$ with relu activation and unbounded width. Conclude that the compactness assumption in the definition of a universal approximator is crucial to define meaningful approximation results.