6.S183 (IAP 2025) Problem Set 1

due: 10 Jan 2025 23:59
updated: 06 Jan 2005

Collaboration Policy: In the case of written homework assignments, your assignment must represent your own individual work. Although you may discuss homework problems with other students, please write up your own solutions.
Submission: Gradescope (`https://www.gradescope.com/courses/931575`).

Question 1 (3 points)

a. Prove that for any distribution over random variables X and Y we have

$$\operatorname{argmin}_f \mathop{\mathbb{E}}_{x,y} \left[\|f(y) - x\|^2 \right] = \mathbb{E}[x \mid y].$$

Hint: Solve the optimization problem pointwise for fixed y.

b. Let μ be the density function of a data distribution, so that $\mu(x_0) \geq 0$ for all $x_0 \in \mathbb{R}^n$ and $\int_{\mathcal{K}} \mu(x_0) dx_0 = 1$. Consider the following loss function for fixed σ.

$$\mathcal{L}_\sigma(\epsilon_\sigma) = \mathop{\mathbb{E}}_{x_0 \sim \mu,\, \epsilon \sim N(0, I_n)} \left[\|\epsilon_\sigma(x_0 + \sigma\epsilon) - \epsilon\|^2 \right] \tag{1}$$

Write down the exact minimizer $\epsilon^*_\sigma(x_\sigma)$ of (1) in terms of the data distribution $\mu(x_0)$ and input x_σ. Hint: use part a, Bayes' rule, as well as the probability density function of $N(x_0, \sigma^2 I_n)$.

Question 2 (3 points)

We will write code to train a diffusion model for a toy 2D dataset. See the provided Jupyter notebook for instructions. You can open the notebook in Google Colab (for this assignment a CPU-only instance is sufficient) `https://colab.research.google.com/drive/1gNZkGePLZmH9uGdjhroxkYe_GXA5KDSy`. For submission, export the notebook as a PDF (ensure all plots and figures are visible!) and upload via gradescope.