Name:

Short Quiz 4B

8 December 2025

Question 1: Prove that $\|x+y\|^2 = \|x\|^2 + 2\langle x, y\rangle + \|y\|^2$.

Answer:

$$\begin{aligned}\|x+y\|^2 &= \sum_{i=1}^{d}(x_i+y_i)^2 = \sum_{i=1}^{d}(x_i^2 + 2x_iy_i + y_i^2)\\ &= \sum_{i=1}^{d} x_i^2 + 2\sum_{i=1}^{d} x_iy_i + \sum_{i=1}^{d} y_i^2 = \|x\|^2 + 2\langle x, y\rangle + \|y\|^2.\end{aligned}$$

Question 2: Compute the gradient of $x \mapsto \frac{1}{2}\|x\|^2$ at x_0.

Answer:

$$\begin{aligned}\frac{1}{2}\|x_0+h\|^2 &= \frac{1}{2}(\|x_0\|^2 + 2\langle x_0, h\rangle + \|h\|^2)\\ &= \frac{1}{2}\|x_0\|^2 + \langle x_0, h\rangle + o_{h\to 0}(\|h\|).\end{aligned}$$

Therefore, the gradient is x_0.

Question 3: What is the idea of boosting?

Answer: The idea of boosting is to train weak learners $f_1, \ldots, f_B$ sequentially, with weights $\alpha_1, \ldots, \alpha_B$, so that the predictor f_{b+1} performs well on training samples where $\alpha_1 f_1 + \cdots + \alpha_b f_b$ performs poorly. To do that, f_{b+1} is trained on a reweighted dataset, where more importance is given to such training samples. The final decision function is $\alpha_1 f_1 + \cdots + \alpha_B f_B$.