CSCI B609 – Foundations of Data Science **Fall 2016**

Homework 2: October 04

Name: YOUR NAME HERE *Due: Monday, October 17, 11:59pm EST*

Problem 2.1 (Exercise 3.1, text modified) *Given a set of points* $\{(x_i, y_i)|1 \leq i \leq n\}$ *give formulas for parameters* m *and* b *of the line of the form* $y = mx + b$ *that minimizes squared vertical distance between the points and the line (rather than the distance to the closest point on the line as we did in class). Formally if* (x_i, y_i) *is a data point and* (x_i, y) *is a point on the line then the vertical distance equals* $|y_i - y|$.

Problem 2.2 (Problem 3.23, text modified)

1. *For any matrix* A *show that* $\sigma_k \leq \frac{\|A\|_F}{\sqrt{k}}$.
2. *Prove that there exists a matrix* B *of rank at most* k *such that* $\|A - B\|_2 \leq \frac{\|A\|_F}{\sqrt{k}}$.
3. *Does there exist a matrix* B *of rank at most* k *such that* $\|A - B\|_F \leq \frac{\|A\|_F}{\sqrt{k}}$*? If yes, construct* B*, if no then give a counterexample.*

Problem 2.3 (Excercise 6.11) *What is the VC-dimension of the class* $\mathcal{H}$ *of axis-parallel boxes in* $\mathbb{R}^d$. *That is* $\mathcal{H} = \{h_{a,b} : a, b \in \mathbb{R}^d\}$ *where* $h_{a,b}(x) = 1$ *is* $a_i \leq x \leq b_i$ *for all* $i = 1, \ldots, d$ *and* $h_{a,b}(x) = 0$ *otherwise.*

1. *Prove that the VC-dimension is at least your chosen* V *by given a set of* V *points that is shattered by the class (and explaining why it is shattered).*
2. *Prove taht the VC-dimension is at most your chosen* V *by proving that no set of* $V + 1$ *points can be shattered.*

Problem 2.4 (Excercise 6.12) *Recall that the margin of a linear separator* w^* *is defined as* $\gamma = 1/\|w^*\|_2$. *Say that a set of points is shattered by linear separators of margin* γ *if every laeling of the points in* S *is achievable by a linear separator of margin at least* γ. *Prove that no set of* $1/\gamma^2 + 1$ *points in the unit ball is shattered by linear separators of margin* γ.

Hint: *think about the Perceptron algorithm and try a proof by contradiction.*