Sándor Kisfaludi-Bak and Themis Gouleakis **Summer 2021**

Geometric algorithms with limited resources, Exercise Sheet 5

https://www.mpi-inf.mpg.de/departments/algorithms-complexity/teaching/summer21/geometric-algorithms-with-limited-resources

Total Points: 40 Due: 8am, Thursday, **July 8th**, 2021

You are allowed to collaborate on the exercise sheets, but you have to write down a solution on your own, ***using your own words****. Please indicate the names of your collaborators for each exercise you solve. Further, cite all external sources that you use (books, websites, research papers, etc.). You need to collect at least 50% of all points on exercise sheets to be admitted to the exam.*

Exercise 1 — **10** points

Let S be the unit sphere, i.e., the boundary of the unit ball in $\mathbb{R}^3$. A disk of radius r and center p in S is the set of points in S whose distance is at most r from p, where the distance is measured on the surface of S. Show that the VC-dimension of the set of all disks in S is $O(1)$.
Hint: Try to move the disks so that they are "canonical". How many canonical disks are there?

Exercise 2 — **10** points

Let P be a convex polytope such that $B_{1/2} \subset P \subset B_1$, where B_α is the ball of radius α centered at the origin. For a point $x \in S := \partial B_1$, let $n_P(x)$ denote the point of P nearest to x. Assume that each vertex of P has degree 3.

Let X be a set of points on the unit sphere such that any disk on S of radius ε contains some point of X. Let $X' = \{n_P(x) \mid x \in X\}$.

Suppose that F is a face of P that does not contain any point of X' (not even on the boundary of F). Show that the area of F is $O(\varepsilon)$.

Exercise 3 — **10** points

Suppose we changed the definition of PAC-learning to only require an algorithm succeed in finding a low-error hypothesis with probability at least $1/2$ (rather than with probability $1-\delta$). Show that this does not change what is learnable.
Specifically, show that if we had an algorithm $\mathcal{A}$ with at least a $1/2$ chance of producing a hypothesis of error at most $\epsilon/2$, we could convert it into an algorithm $\mathcal{B}$ that has at least a $1-\delta$ probability of producing a hypothesis of error at most ϵ. The reduction is that we first run $\mathcal{A}$ for $N = \log(2/\delta)$ times (so with probability at least $1-\delta/2$, at least one of the N hypotheses produced has error at most $\epsilon/2$), and we then test the N hypotheses produced on a new test set, choosing the one that performs best.

Hint: Use Chernoff bounds to analyze this second step and finish the argument. That is, assuming that at least one of N hypotheses has error at most $\epsilon/2$, give an asymptotic bound on a size for the test set that is sufficient so that with probability at least $1-\delta/2$, the hypothesis that performs best on the test set has error at most ϵ.

Exercise 4 — **4+6** points

Suppose that we modify the Perceptron algorithm so that it makes the update on positive examples even when it did not make a mistake.

(a) Show that, whenever the current hyperplane is not perpendicular to the true separating hyperplane, it is possible that an update moves the hyperplane in the wrong direction for sufficiently small margin γ. Specifically, show that the angle between the current hyperplane and the true separator hyperplane of the examples could increase.

(b) Show how this can cause the algorithm to make an unbounded number of mistakes even in the γ-margin case for any $\gamma < 1/2$.