Exercises 11 & 12

Analysis of Boolean Functions, Autumn 2025, University of Chicago
Instructor: William Hoza (williamhoza@uchicago.edu)

Submission. Solutions are due **Friday, November 14** at 11:59pm Central time. Submit your solutions through Gradescope. You are encouraged, but not required, to typeset your solutions using a LaTeX editor such as Overleaf.

The policies below can also be found on the course webpage.

Collaboration. You are encouraged to collaborate with your classmates on exercises, but you must adhere to the following rules.

- Work on each exercise on your own for at least five minutes before discussing it with classmates.
- Feel free to explain your ideas to your classmates in person, and feel free to use whiteboards/chalkboards/etc. However, do not share any written/typeset solutions with your classmates for them to study on their own time. This includes partial solutions.
- Write your solutions on your own. While you are writing your solutions, do not consult any notes that you might have taken during discussions with classmates.
- In your write-up, list any classmates who helped you figure out the solution. The fact that student A contributed to student B's solution does not necessarily mean that student B contributed to student A's solution.

Permitted Resources for Full Credit. In addition to discussions with me and discussions with classmates as discussed above, you may also use the course textbook, any slides or notes posted in the "Course Timeline" section of the course webpage, and Wikipedia. If you wish to receive full credit on an exercise, you may not use any other resources.

Outside Resources for Partial Credit. If you wish, you may use outside resources (ChatGPT, Stack Exchange, etc.) to solve an exercise for partial credit. If you decide to go this route, you must make a note of which outside resources you used when you were working on each exercise. You must disclose using a resource even if it was ultimately unhelpful for solving the exercise. Furthermore, if you consult an outside resource while working on an exercise, then you must not discuss that exercise with your classmates.

Exercise 11 (10 points)**.**

(a) Let $f, g\colon \{\pm 1\}^n \to \mathbb{R}$ and $i, j \in [n]$, with $i \neq j$. Assume $f(x)$ does not depend on x_i and $g(x)$ does not depend on x_j. Prove that

$$\mathop{\mathbb{E}}_{x}[x_i \cdot x_j \cdot f(x) \cdot g(x)] = \langle D_j f, D_i g \rangle.$$

(b) Let $p\colon \{\pm 1\}^n \to \mathbb{R}$ and let $f(x) = \operatorname{sign}(p(x))$, using the convention $\operatorname{sign}(0) = +1$. Prove that

$$\mathrm{I}[f] \leq \sqrt{n + \sum_{i=1}^{n} \mathrm{I}[\operatorname{sign}(D_i p)]}.$$

Hints:

i. Prove that $(D_i f)(x) = \operatorname{sign}((D_i p)(x))$ whenever $(D_i f)(x) \neq 0$.

ii. Prove that $\mathrm{Inf}_i[f] = \mathbb{E}_x[f(x) \cdot x_i \cdot \operatorname{sign}((D_i p)(x))]$.

iii. Prove that $\mathrm{I}[f] \leq \mathbb{E}_x[|\sum_{i=1}^{n} x_i \cdot \operatorname{sign}((D_i p)(x))|]$.

iv. Apply Jensen's inequality, part (a), and the AM-GM inequality.

(c) Let $f\colon \{\pm 1\}^n \to \{\pm 1\}$ be a degree-k polynomial threshold function, i.e., $f(x) = \operatorname{sign}(p(x))$ where $\deg(p) = k$. Prove that $\mathrm{I}[f] \leq O(n^{1-2^{-k}})$.

Hint: Use part (b) and induction on k.

(d) Prove that for all constants $k \in \mathbb{N}$ and $\varepsilon \in (0, 1)$, degree-k polynomial threshold functions can be learned from random examples with error ε in polynomial time.

A *de Morgan formula* is a binary tree in which each internal node is labeled $\wedge$ or $\vee$ and each leaf is labeled with a constant or a literal. The *leafsize* of the formula is the number of leaves that are not labeled with constants. For a function $f\colon \{\pm 1\}^n \to \{\pm 1\}$, we define $L(f)$ to be the minimum leafsize of any de Morgan formula computing f. For this exercise, you may take for granted the following theorem regarding the effect of random restrictions on de Morgan formulas.

Theorem 1. *Let* $f\colon \{\pm 1\}^n \to \{\pm 1\}$ *and let* $p \in (0, 1)$. *Then*

$$\mathop{\mathbb{E}}_{\rho \sim \mathcal{R}_p}[L(f|_\rho)] = O\left(p^2 \cdot L(f) + p \cdot \sqrt{L(f)}\right).$$

Exercise 12 (5 points)**.** Prove that de Morgan formulas with leafsize at most m can be learned from random examples with error ε in time $n^{O(\sqrt{m}/\varepsilon)}$.