Lecture 11

- finish monotonicity of high dimensional Boolean fctns (slides)
- testing via learning (slides)
- testing monotonicity under L_1-distance

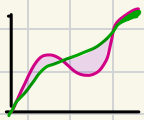
Testing monotonicity in L_1-distance:

Problem Given $f: [n] \to [0,1]$

$\sum_{i=1}^{n} |f(i) - g(i)|$

Pass if f monotone

Fail if f ε-far in L_1-distance from any monotone fctn



How does it compare to Hamming distance?

when $f: [n] \to \{0,1\}$, they are equal

for $f: [n] \to [0,1]$, $HD \geq L_1\text{-dist}$

for $f: [n] \to [0, \ldots, d]$, $HD \cdot d \geq L_1\text{-dist}$

Thm Can test if f monotone wrt L_1-dist in $O(1/\varepsilon)$ queries

↑ need $\Omega(\log n)$ for Hamming distance!

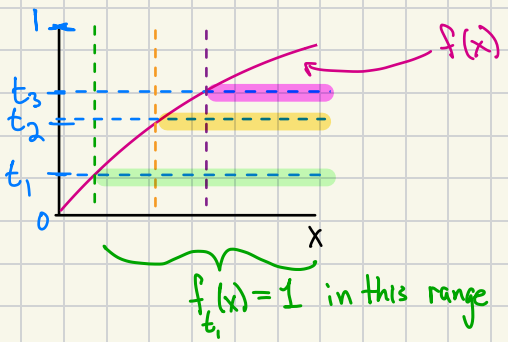
Proof

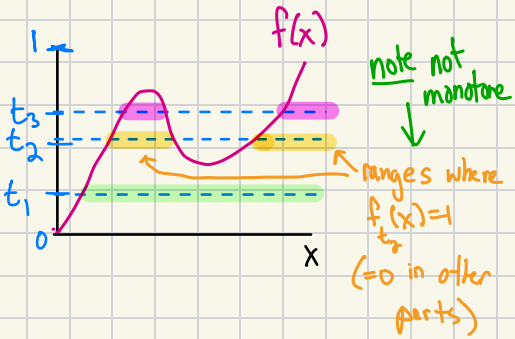
Idea: reduce to Boolean monotonicity testing

def for $t \in [0,1]$

threshold fctn $f_{(t)} : [n] \to \{0,1\}$

$$f_{(t)}(x) = \begin{cases} 1 & \text{if } f(x) \geq t \\ 0 & \text{o.w.} \end{cases}$$





Express f as sum of fctns mapping to $\{0,1\}$:

For any f s.t. $f(0) = 0$:

(Can be non-monotone)

$$f(x) = \int_0^{f(x)} 1 \cdot dt = \int_0^1 f_{(t)}(x)\, dt$$

$$= \int_0^{f(x)} f_{(t)}(x)\, dt + \int_{f(x)}^1 f_t(x)\, dt$$

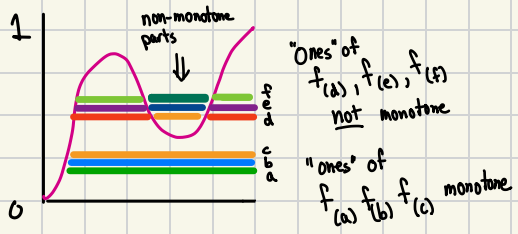
$f_{(t)}(x) = 1$ in this range

$f_t(x) = 0$

Let $L_1(f, M) \equiv L_1$ dist of f to closest monotone fctn

Lemma $L_1(f, M) = \int_0^1 L_1(f_{(t)}, M)\, dt$

ie. can express dist to monotonicity in terms of distances of threshold fctns.



Pf idea I. $L_1(f, M) \geq \int_0^1 L_1(f_{(t)}, M)\, dt$

To change f into monotone, must make each "row" monotone ← row is 0,1 fctn

need to change at least $L_1(f_{(t)}, M)$ in each row

$\Rightarrow L_1(f, M) \geq \int_0^1 L_1(f_{(t)}, M)\, dt$

Pf idea cont.

II. $L_1(f, M) \leq \int_0^1 L_1(f_{(t)}, M)\, dt$

When make each row monotone (which costs $L_1(f_{(t)}, M)$)
let g_t be resulting fctn
(closest monotone fctn to $f_{(t)}$)

let $g \equiv \int_0^1 g_t\, dt$

g monotone since sum of monotone fctns.

$$\Rightarrow L_1(f, M) \leq \|f - g\|_1$$
$$= \int_0^1 L_1(f_{(t)}, M)\, dt$$

□

Why is characterization helpful?

makes queries, passes iff answers are monotone

Lemma If T is a nonadaptive, 1 sided error tester for $f: D \to \{0,1\}$ w.r.t. Hamming dist

$\Rightarrow$ T " " " " " " " " "

exactly same test!

" " $f: D \to [0,1]$ w.r.t. L_1-dist

Can reduce to testing over binary domains in Hamming distance!

(and we know how to do this)

↖ was our algorithm nonadaptive?

Pf if f monotone:

T queries inputs in Q & passes iff f monotone on Q

$\Rightarrow$ T accepts (uses 1-sided error of T)

if f s.t. $L_1(f, M) \geq \varepsilon n$:

Lemma $\Rightarrow$ $L_1(f, M) = \int_0^1 L_1(f_{(t)}, M)\, dt$ = average value of $L_1(f_{(t)}, M)$

$\Rightarrow$ $\exists\, t^*$ s.t.

$$L_1(f_{(t^*)}, M) \geq \varepsilon n$$

since $f_{(t^*)}$ is Boolean,

Hamming & L_1-dist are same wrt dist to monotone.

$\Rightarrow$ T fails $f_{(t^*)}$ for $\geq 3/4$ choices of queries Q

$\Rightarrow$ with prob $\geq 3/4$, Q contains $x < y$ s.t. $f_{(t^*)}(x) > f_{(t^*)}(y)$

$\Downarrow$

$f(x) \geq t^* > y$

so Q outputs FAIL!

