Learning Parity Fctns

PAC Setting:

from which distribution?

Given samples $x, f(x)$

Find χ_s s.t. χ_s + f agree a lot ← large Fourier Coeffs.

Thought to be hard:

if x from arbitrary distribution then NP-hard

"Maximum likelihood decoding of linear codes"

if x from uniform dist. then ~~still~~ thought to be hard

" hardness of parity with noise"

" hardness of decoding linear codes"

used as hardness assumption eg. in Crypto

if noise random:

"hardness of decoding random linear codes"

"noisy parity"

[A. Blum Kalai Wasserman]: Can solve in $2^{O(n/\log n)}$

used to determine lattice vector & length,
Crytoanalysis
+ other learning problems

What if given query access to f for arbitrary inputs??

Learning Parities with Queries

$$x \longrightarrow \boxed{f} \longrightarrow f(x)$$

Given f, θ

1) Output all coeffs S s.t. $|\hat{f}(S)| \geq \theta$ (get all "close" fctns)

2) Only output coeffs S s.t. $|\hat{f}(S)| \geq \frac{\theta}{2}$ (no real junk)

(Using Boolean Parseval's: $\sum \hat{f}^2(S) = 1$ only $O(1/\theta^2)$ such coeffs)

recall $\Pr_x[f(x) = \chi_S(x)] = \frac{1}{2} + \frac{\hat{f}(S)}{2}$

so case 1 $\Rightarrow$ $\Pr_x[f(x) = \chi_S(x)] \geq \frac{1}{2} + \frac{\theta}{2}$

2 $\Rightarrow$ $\leq \frac{1}{2} + \frac{\theta}{4}$

Warmup #0:

poly queries, unbnded time $\Big\}$ find all f that agree enough

Warmup #1: (poly queries, poly time) from now on

Suppose f agrees with χ_S everywhere for some S

(i.e. 0-error case) only one S s.t. $\chi_S \neq 0$

Algorithm 1: equation solving for coeffs

Algorithm 2: $\forall\, i \in [n]$ put i in S if $f(1111) \neq f(\underbrace{111(-1)111}_{e_i})$ (ith spot ↓)

Outputs

Note if $i \in S$ $\chi_S(u) \cdot \chi_S(ue_i) = -1$

Warmup #2

(∃ S s.t. agreement with $\chi_S \approx 1$ + all other χ_T's is ≈ 0)

Suppose f agrees with χ_S $\underbrace{\text{"almost"}}$ everywhere $\left(\ll \frac{1}{\text{poly}(n)} \text{-negligible fraction of inputs}\right)$ for some S

Note: can't use previous algorithm since error might be on (1111⋯1)

Algorithm:

```
choose $r \in \{\pm 1\}^n$
$\forall i \in [n]$
    put $i$ in $S$ if $f(r) \neq f(r \odot e_i)$
                                  ↑ coordinatewise multiplication
Output S
```

Why? (sketch)

f(r), f(r⊙e_i) agree with $\chi_S(r)$ $\chi_S(r \cdot e_i)$ for almost all r

unif dist

so $\Pr[S \text{ not correct}] \leq 2n \cdot \underbrace{\text{negligible}}_{\text{union bnd}}$

Warmup #3

Suppose f agrees with χ_S on $3/4 + \varepsilon$ for some S

$\uparrow$
$\geq \frac{1}{\text{poly}(n)}$

(here get better result on # solns than Boolean Parsevals: BP ⇒ ≤ 3 but actually there is only unique soln.

Algorithm:

```
choose $r_1 \cdots r_t \in \{\pm 1\}^n$
$\forall i \in [n]$
    put $i$ in $S$ if $\underbrace{\text{majority}}_{t \text{ samples}}$ of $f(r_j) \neq f(r_j \odot e_i)$
Outputs S
```

(warmup ~~3~~ 3 cont)

why?

$(-1)^{1_{i\in S}}$

///

$Pr[$"wrong" answer for r_j on $i]$

$$= Pr[f(r_j)\cdot f(r_j \oplus e_i)\cdot(-1)^{1_{i\in S}} \neq 1]$$

↑ "right" should be different if $i \in S$ same if $i \notin S$

$$\leq Pr[f(r_j) \neq \chi_S(r_j)] + Pr[f(r_j \oplus e_j) \neq \chi_S(r_j \oplus e_j)]$$

uniformly distributed

Union bound on two bad events **BUT** we are doing union bound on same $f(r_j)$ event over + over + over!!!

$$\leq \left(\tfrac{1}{4}-\varepsilon\right) + \left(\tfrac{1}{4}-\varepsilon\right) = \tfrac{1}{2} - 2\varepsilon$$

∴ get correct answer with prob slightly $> 1/2$

∴ for i, most r_j are right with prob $> 1 - \delta/n$ — Chernoff: picking $t = \Theta(\log n)$

for all i, most r_j are right with prob $> 1-\delta$

Warmup ~~4~~ 4

output all S s.t. f agrees with χ_S on $\geq \frac{1}{2}+\varepsilon$ fraction of inputs

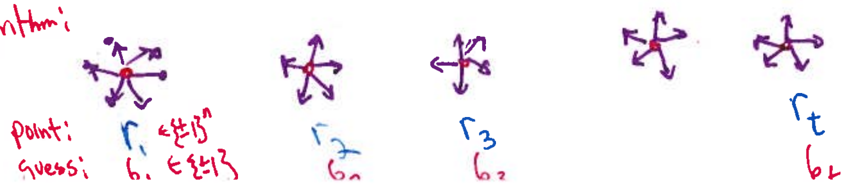
↑ constant

idea 1 guess answers to $f(r_j)$'s

Since only $O(\log n)$, can run over all possible guesses } saves half the union bound error!!!

Idea 2 Can test Candidates + rule out junk

Picture of Algorithm: Pick $r_1 \dots r_t$ uniformly + compare to all nbrs



Algorithm

- Choose $r_1 \cdots r_t \in \{\pm 1\}^n$ $\quad t = O(\log n)$
- For all possible settings of $\sigma_1 \cdots \sigma_t$
  $\{$"guesses" to values of $\chi_S(r_i)$'s$\}$
  - $\forall i \in [n]$ put i in $S_{\sigma_1 \cdots \sigma_t}$ if

i.e. by testing if
$$f(r_j) \neq f(r_j \odot e_j)$$
$$\Updownarrow$$
$$\sigma_j \neq f(r_j \odot e_j)$$
$\longrightarrow$ majority of (over $j \in [t]$) $\sigma_j \neq f(r_j \odot e_i)$ $\}$ generate a candidate for S

  - Sample to see if $\chi_{S_{\sigma_1 \cdots \sigma_t}}$ agrees
    with f on $\geq \frac{1}{2} + \frac{3}{8}\theta$ inputs
    if yes, output $\chi_{S_{\sigma_1 \cdots \sigma_t}}$

$\}$ test candidate + weed out junk

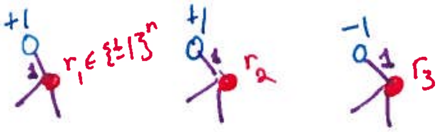
Note: many settings of $\sigma_1 \cdots \sigma_t$ could give good answer since could have lots of linear fctns agreeing with f on enough inputs

Why?

For each S that should be output
consider $\sigma_1 \cdots \sigma_t$ s.t. $\sigma_i = \chi_S(r_i)$

For this setting
(see next page)

Example of what happens with $i=1$ for all guesses of σ_i's:



σ_1	σ_2	σ_3	$f(r_1 \odot 1(i)) = +1$	$f(r_2 \odot 1(i)) = +1$	$f(r_3 \odot 1(i)) = -1$	$1 \in S$?
+	+	+	+ vs. +1	+ vs +	+ vs −	no
+	+	−	+ vs +	+ vs +	− vs −	no
+	−	+	+ vs +	− vs +	+ vs −	yes
+	−	−	+ vs F	− vs +	− vs −	no
−	+	+	− vs +	+ vs +	+ vs −	yes
−	+	−	− vs +	+ vs +	− vs −	no
−	−	+	− vs +	− vs +	+ vs −	yes
−	−	−	− vs +	− vs +	− vs −	yes

For this setting:

$$\Pr[\text{wrong answer for } r_j \text{ on } i]$$

$$= \Pr[\sigma_j \cdot f(r_j \odot e_i) \cdot (-1)^{1_{i \in S}} = -1]$$

$$\text{assumption} \Rightarrow \overset{\sigma_j}{\underset{\chi_S(r_j)}{\|}} \cdot \overset{\|?}{\chi_S(r_j \odot e_i)} \cdot (-1)^{1 \in S} = -1$$

$$\leq \Pr[f(r_j \odot e_i) \neq \chi_S(r_j \odot e_i)]$$

$$\leq \frac{1}{2} - \varepsilon$$

Chernoff bnds + $O(\log n)$ r_j's $\Rightarrow$ $\Pr[\text{wrong answer on } i] \leq 1/2n$

+ union bnd $\Rightarrow$ $\Pr[\text{wrong answer on } \underline{\text{any}}\ i] \leq 1/2$

$\therefore$ S is output with prob $\geq 1/2$

for each S that <u>should not</u> be output:

$$\Pr[\text{output } S] \leq \Pr[S \text{ passes testing phase}]$$

runtime:

since $t \approx \theta(\log n)$, need $2^{\theta(\log n)}$ iterations $\Rightarrow$ poly(n)

Learning Parity Functions

General Case

Output all S st f agrees with χ_S on $\geq \frac{1}{2} + \varepsilon$ fraction of inputs

↑ can be $\frac{1}{\text{poly}(n)}$

Show that not too many such S

idea

in earlier warmup, if ε small $(\approx \frac{1}{\text{poly}(n)})$ need more samples for Chernoff to kick in - ie. if need poly(n) samples then need $2^{\text{poly}(n)}$ guesses!

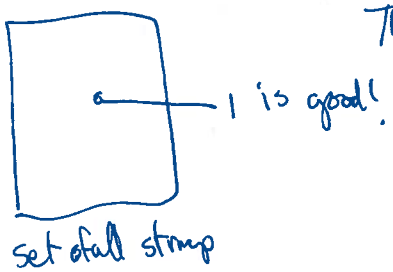
Fix

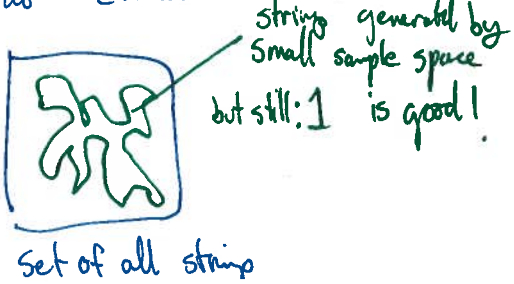
Choose many more $r_1 \ldots r_t$ but not independently

ie. choose them pairwise independently

that is - find sample space of poly size (ie. $2^{\underbrace{O(\log n)}_{\#\text{p.i. bits needed}}}$) which behaves in the same way as iid vars.

Then do exhaustive search on sample space!





Algorithm

- choose $s_1 \cdots s_K \in \{\pm 1\}^n$ $\quad K = \log(t+1)$ $\quad$ # guesses

  $t = \Theta(n/\varepsilon^2)$ $\quad$ # r_i's generated

  $\geq \frac{2n}{\varepsilon^2}$

- For all possible settings of $\sigma_1 \cdots \sigma_K \in \{\pm 1\}^K$: $\quad$ {all "guesses" for values of $\chi_S(s_i)$'s}

  {generate <u>a lot</u> ($2^K \approx n/\varepsilon^2$) of labelled samples}

  - For every $W \subseteq \{1..K\}$ $\quad W \neq \emptyset$

    set $r_W \leftarrow \bigoplus_{j \in W} s_j$ $\quad$ ← pairwise random bits

    $\rho_W \leftarrow \prod_{j \in W} \sigma_j$ $\quad$ if initial guesses of σ_i's "correct" according to χ_S then $\rho_W = \chi_S(r_W)$

  - $\forall\, i \in [n]$ put i in $S_{\sigma_1 \cdots \sigma_K}$ if majority of $\rho_W \neq f(r_W \oplus e_i)$ $\quad$ ← creates $S_{\sigma_1 \cdots \sigma_K}$

  - Test $S_{\sigma_1 \cdots \sigma_K}$ to see if agrees enough with f

    $\geq \frac{1}{2} + \frac{3}{4}\varepsilon$ fraction

    if yes, output it

Behavior

For s' s.t. f agrees with $\chi_{s'}$ on $\geq \frac{1}{2}+\varepsilon$ of inputs:

1) if setting of σ_i's agrees with χ_s

i.e. $\forall i \quad \sigma_i = \chi_s(s_i)$

then $\forall w$

$$\left.\begin{aligned} p_w &= \prod_{j\in w} \chi_{s'}(s_j) && \text{def of } p_w \\ &= \chi_{s'}\Big(\bigoplus_{j\in w} s_j\Big) \\ &= \chi_{s'}(r_w) && \text{def of } r_w \end{aligned}\right\} \text{so all } p_w\text{'s are consistent with } \sigma$$

From now on, assume this setting of σ_i's . . .

2) r_w's are pairwise independent [in fact, generated via a known construction]

i.e. $\Pr[r_w = b_1 \;\&\; r_{w'} = b_2] = \Pr[r_w = b_1] \cdot \Pr[r_{w'} = b_2]$

also $r_w \odot e_i$'s are p.i.

3) $\Pr[$ Algorithm generates s' when considering $S_{\sigma_1 \cdots \sigma_k}]$:

$\Pr[$it get s' right on index $i]$

$$= \Pr\Big[\underbrace{p_w \cdot f(r_w \odot e_i) \cdot (-1)^{1_{i\in s}} = 1}_{\text{indicator } X_w = \begin{cases} 1 & \text{if holds} \\ 0 & \text{o.w.}\end{cases}}\Big]$$

Note: if $f(r_w \odot e_i) = \chi_s(r_w \odot e_i)$ ← ??

& $p_w = \chi_s(r_w)$ ← assumption

then $X_w = 1$

$E[X_w] \geq \frac{1}{2} + \varepsilon$ since $r_w \oplus e_i$ uniform dist

Variance $\sigma_w^2 = E[X_w^2] - E[X_w]^2$

$$\geq \frac{1}{2} + \varepsilon - \left(\frac{1}{2} + \varepsilon\right)^2 = \frac{1}{4} - \varepsilon^2$$

$$E\left[\sum_{w \in [K]} X_w\right] \geq t\left(\frac{1}{2} + \varepsilon\right)$$

$$\Pr\left[\sum_w X_w < \frac{t}{2}\right] \leq \frac{\left(\frac{1}{2}\right)^2 - \varepsilon^2}{t\ \varepsilon^2} \leq \frac{1}{t\varepsilon^2} \leq \frac{1}{2n}$$

union bnd: $\Pr[\$ \text{ not output}] \leq \frac{1}{2}$

Chebyshev:
$X_1 \cdot X_n$ p.i.
$E[X_i] = \mu$
$Var[X_i] = \sigma^2$
$\Pr\left[\left|\frac{\sum X_i}{n} - \mu\right| > \varepsilon\right] \leq \frac{\sigma^2}{\varepsilon^2 n}$

Also shows:

#parity fctns agreeing with f on $\geq \frac{1}{2} + \varepsilon$ is $O\left(\frac{n}{\varepsilon^2}\right)$