3/9/2015 Russell's course lec 3. Linial, Mansour, Nisan

Algorithmic uses of switching lemma:

learning depth d fns using Fourier representation [LMN]

Exercise: show that a rand rest. setting 2/3 of inputs makes bottom fan-in $O(\log S)$ whp.

Consequences of switching lemma:

let C be depth d fn of size S, $D > \log S$, $p = \Omega\left(\frac{1}{D^{d-1}}\right)$.

ρ a rand restr leaving pn vars unset. $\text{Prob}[DT[C|_\rho] \geq D] \leq 2^{-D}$

PAC learning C on unif distr.

Black-box access to C. Training: obtain $(x_1, C(x_1)) \ldots (x_T, y_T)$

where x_i are random n-bit inputs.

Generate hypothesis $h: \{0,1\}^n \to \{0,1\}$, w prob of success $\geq 1-\varepsilon$

- know C is depth d and size S, n inputs

Fourier representation:

~ represent true by -1, and false by 1.

Over $0,1$, $\sum x_i \bmod 2$ is parity. Over $\{-1,1\}$, parity is $\prod x_i$

~ function as a vector: $F(x) = \langle F(0..0), F(0..01), F(0..010) \ldots F(1..1) \rangle$. This is 2^n dim vector, embedded in 2^n-dim space. $\|\vec{F}\| = \sqrt{\sum_x (F(x))^2} = 2^{n/2}$.

Normalizing const $1/2^{n/2}$ 1 if $=$, -1 if $\neq$.

$\vec{F}, \vec{G}$. $\vec{F} \cdot \vec{G} = \frac{1}{2^{n/2}} \frac{1}{2^{n/2}} \sum_x F(x) \cdot G(x) =$

$= \frac{1}{2^n} \cdot ([\# x, F(x) = G(x)] - \# x, F(x) \neq G(x))) =$

$= \text{Prob}[F(x) = G(x)] - \text{Prob}[F(x) \neq G(x)] \in [-1, 1]$,

measure of correlation between fns

Geometrically, look under diff. bases. One basis $(1,0..0)$..

Orthonormal basis: set of Bool fns $F_1 \ldots F_{2^n}$ s.t. $\forall i \neq j$, $\langle F_i, F_j \rangle = 0$. So $\text{Prob}[F_i = F_j] = \text{Prob}[F_i \neq F_j]$.

Ideal subsets: parities of subsets. $S \subseteq [1..n]$. $\chi_S = \prod_{i \in S} x_i$ $(x_i \in \{-1,1\})$

$\chi_\emptyset = 1$. Let $S, T \subseteq [n]$, $S \neq T$. when does $\chi_S(x) = \chi_T(x)$?

$\chi_{S \Delta T}(x) = 1$ $\text{Prob}_x(\chi_{S \Delta T}(x) = 1) = 1/2$

χ_S are 2^n orthogonal vectors in 2^n dim space: basis.

Any Bool fn $\vec{F} = \sum \hat{\alpha}_S \vec{\chi}_S$

Fourier coeffs.

$\langle \vec{F}, \vec{\chi}_T \rangle = \sum_S \alpha_S \langle \chi_S, \chi_T \rangle = \alpha_T = \mathrm{Prob}_x [F(x) = \chi_T] - \mathrm{Prob}_x [F(x) \neq \chi_T]$

$\frac{1}{2^{n/2}} F(x) = \sum_S \hat{\alpha}_S \frac{1}{2^{n/2}} \chi_S(x) = \sum_S \hat{\alpha}_S \chi_S(x) = \sum_S \hat{\alpha}_S \prod_{i \in S} x_i$

– multilinear poly.

Dec. tree for F: x_1; $F|_{x_1=-1}$, $F|_{x_1=1}$

$F(x) = \frac{(1-x_1)}{2} \cdot F|_{x_1=-1} + \frac{(1+x_1)}{2} F|_{x_1=1}$ – harmonic, – multilinear rep. of deg d

$\Rightarrow$ any Fourier coef. of F of size $> d$ is 0.

Idea [LMN] this means large Fourier coeffs of F are small. ($|S| \geq \mathrm{polylog}(\mathrm{size}(ckt))$)

Parseval's identity: $\forall$ Bool fn, $\sum_S (\hat{\alpha}_S)^2 = 1$

$\|\vec{F}\| = 1$. $\hat{F} = \langle \alpha_\phi, \alpha_{\{x_1\}} \ldots \alpha_{[n]} \rangle$. $\|\vec{F}\| = \sqrt{1} = 1$

LMN thm: For F computable by size s, depth d ckt, then $\sum_{T, |T| \geq O(D^d)} (\hat{\alpha}_T)^2 \leq \exp(-D)$, $D \geq \log s$.

$F = \sum \hat{\alpha}_S \prod_{i \in S} x_i$. $F|_{x_1=1} = \sum_{S \subseteq \{2..n\}} (\hat{\alpha}_S + \hat{\alpha}_{S \cup \{1\}}) \prod_{i \in S} x_i$

$F|_{x_1=-1} = \sum_{S \subseteq \{2,..n\}} (\hat{\alpha}_S - \hat{\alpha}_{S \cup \{1\}}) \prod_{i \in S} x_i$

$\mathop{E}_{x_1 \in_R \{\pm 1\}} ((\hat{\alpha}_S)^2) = \frac{1}{2}(\hat{\alpha}_S + \hat{\alpha}_{S \cup \{1\}})^2 + \frac{1}{2}(\hat{\alpha}_S - \hat{\alpha}_{S \cup \{1\}})^2 = \hat{\alpha}_S^2 + \hat{\alpha}_{S \cup \{1\}}^2$

Say we set all vars in J at random ρ. Let $F'|_\rho = F'$. Let S be subset of target vars. Then $E(\hat{\alpha}'_S)^2 = \sum_{T' \subseteq T} (\hat{\alpha}_{S \cup T'})^2$

Lemma if ρ is rand. restr. leaving p frac. of vars unset, then

$\mathop{E}_\rho \left(\sum_{S \text{ unset}, \geq t} (\hat{\alpha}'_S)^2 \right) \geq c_0 \sum_{u, |u| \geq 3t/p} (\hat{\alpha}_u)^2$

Cond. on the set T of ρ-set vars, u contributes expected $\hat{\alpha}_u^2$ to $(\hat{\alpha}')^2_{u-T}$.

Part of sum if $|u-T| \geq t$. $E[|u-T|] = p|u| \geq 3t$. W. const prob $> t$.