Quantum PAC - learning model

Bshouty , Jackson 1999

OUTLINE

1. QTM
2. Fourier transform over Boolean Cube
3. Quantum EXAMPLES ORACLE
4. learning DNF with QTM

AROOSA
27-12-2021

PTM

Q – set of states

Σ – input alphabet

Γ – tape alphabet

δ – $c_i \rightarrow c_j$ (transition matrix)

where c (state, content, head)

QTM

Transition prob —— complex numbers / amplitudes → interference

observation on $\sum_i a_i |c_i\rangle$

Example: 1 Qubit

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$$

($|0\rangle = \begin{pmatrix}1\\0\end{pmatrix}$, $|1\rangle = \begin{pmatrix}0\\1\end{pmatrix}$)

$\alpha, \beta \in \mathbb{C}$, $|\alpha|^2 + |\beta|^2 = 1$

↓ observe/measure

- w.p $|\alpha_0|^2$, the outcome is 0 and state collapses to $|0\rangle$
- w.p $|\beta|^2$, the outcome is 1 and state collapses to $|1\rangle$

Fourier transform over Boolean Cube

For any function $f : \{0,1\}^n \to \mathbb{R}$,

① inner product

$$\langle f, g \rangle = \frac{1}{2^n} \sum_{x \in \{0,1\}^n} f(x) \cdot g(x) = \mathbb{E}[f \cdot g]$$

↑ uniform dist over instance space

② $$\|f\|_2 = \sqrt{\langle f, f \rangle} = \sqrt{\mathbb{E}[f^2]}$$

Fourier Basis

define $\chi_a(x) = (-1)^{a \cdot x}$ where $a \cdot x = \sum_{i=1}^{n} a_i x_i \mod 2$

$\chi_a : \{0,1\}^n \to \{-1, +1\}$

$\langle \chi_a, \chi_b \rangle = \delta_{ab} \Longrightarrow \chi_a$ an orthonormal basis for $\mathbb{Z}_2^n$

Fourier spectrum of a function

$$f(x) = \sum_{a \in \{0,1\}^n} \hat{f}(a)\, \chi_a(x)$$

$\hat{f}(a)$: Fourier Coefficient of f at a

$$\hat{f}(a) = \langle f, \chi_a \rangle$$

$$= \mathbb{E}[f \cdot \chi_a]$$

Note If $f : \{0,1\}^n \to \{+1,-1\}$

$$\sum_a \hat{f}(a)^2 = \langle f, f \rangle = \mathbb{E}[f^2] = 1$$

Fourier Spectrum adds upto 1

$\Downarrow$

$\hat{f}(a)^2 \Longrightarrow$ Probabilities

Learning Models

distribution-independent

for all distributions in distribution class

class $\mathcal{F}$ can be

→ M- learnable using membership queries

→ weakly M- learnable

→ M-learnable by $\mathcal{H}$; $h \in \mathcal{H}$ ←

→ Properly M- learnable ←

→ M-learnable w.r.t $\mathcal{D}$ ← $\bigcup_n \mathcal{D}_n$

→ Quantum- learnable

Strong

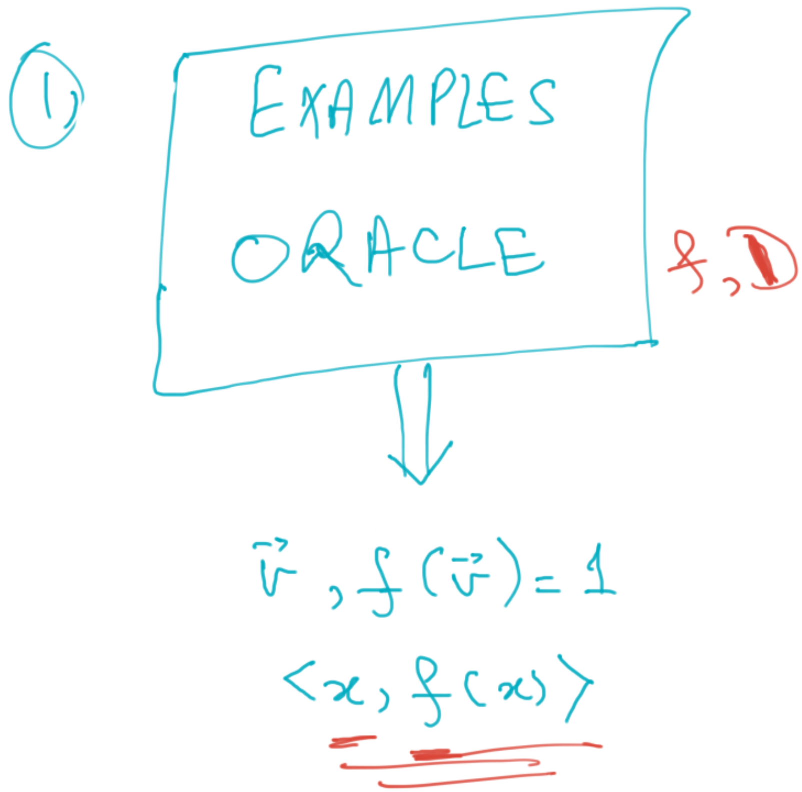
ε- approximator

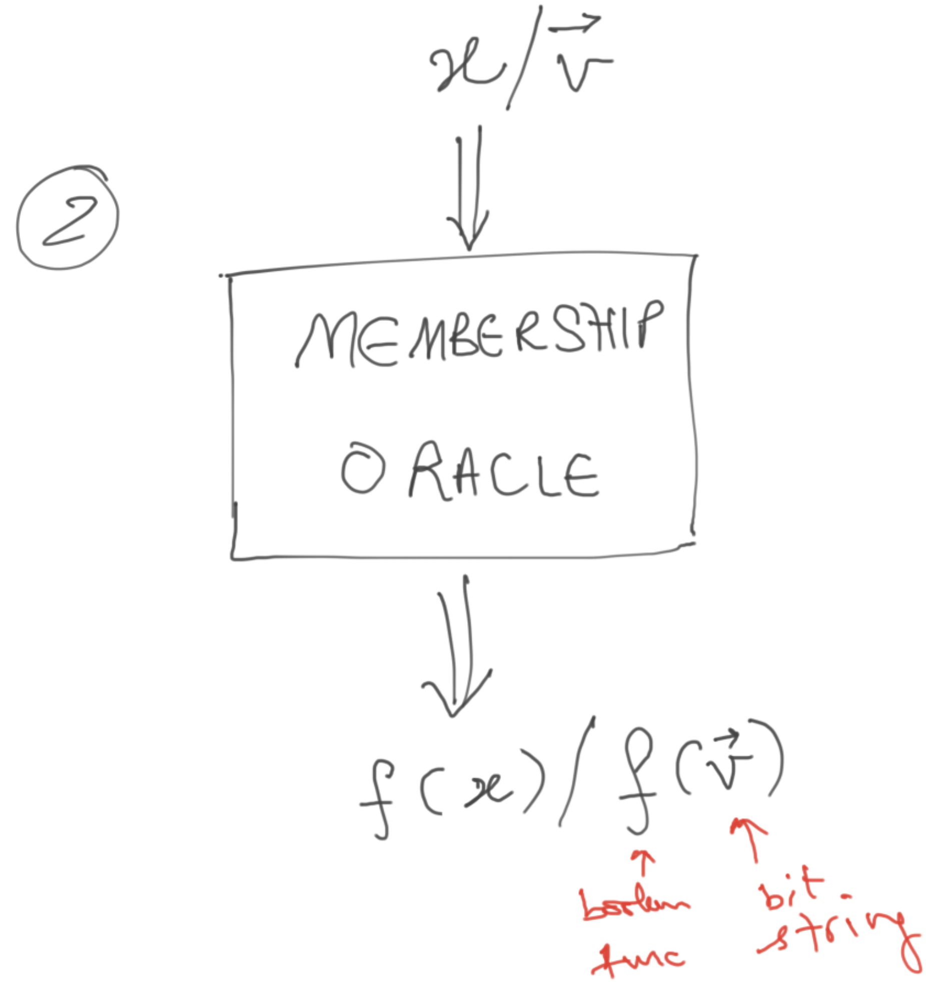
$f \in \mathcal{F}$

proper

$h \in \mathcal{F}$

ORACLES


1
EXAMPLES ORACLE
f, D
$\vec{v}, f(\vec{v}) = 1$
$\langle x, f(x) \rangle$


2
$x / \vec{v}$
MEMBERSHIP ORACLE
$f(x) / f(\vec{v})$
boolean func
bit string

HISTORY of learning DNF

$(m_1 + m_2 + \cdots + m_r : m_i = q_1 + q_2 \cdots q_k)$

→ Valiant 1984 — MONOTONE DNF learnable using both oracles

KV, Mansour 1993

→ Mansour, Rudich 1994 — DNF weakly-learnable w.r.t $\mathcal{U}$ $(\bigcup_n \mathcal{U}_n)$ using membership oracle

↓

→ Jackson 1996 — DNF learnable w.r.t $\mathcal{U}$ $(\bigcup_n \mathcal{U}_n)$ using membership oracle

Q Bshouty, Jackson 1999

QTM M + EXAMPLES ORACLE

$\Downarrow$

Can M efficiently learn an ϵ-approximation to f?

QUANTUM EXAMPLES ORACLE

$$\sum_x D(x) \langle x, f(x) \rangle \quad \Longleftrightarrow \quad \sum_x \sqrt{D(x)} \, \langle x, f(x) \rangle$$

$\downarrow$ Sample and output from 2^n possible configurations

$\downarrow$ Sample according to probability amplitude and observe!

→ $\mathcal{F}$ is quantum-learnable if $\mathcal{F}$ is PAC-learnable by a QTM using a QEX.

→ Every PAC-learnable function class is quantum-learnable

TM   efficient simulation   QTM

EX   ←———————   QEX

KUSHILEVITZ, Mansour 1993

$c \times \frac{1}{2s+c}$

① $\forall$ f_{DNF}, $D \rightarrow \chi_a$ is a weak-approximator of f w.r.t D

② Gave efficient algo to find χ_a using polynomial number of membership queries when $D = U$

Jackson, Rudich 1994

① Use boosting to get a strong learner of DNF w.r.t U.

② Algo

find χ_a to maximize $E[f, \chi_a]$

- define D_i ✓
- get weak-learner h_i w.r.t D_i ← requires membership and distribution oracle
- compute error of h using EXAMPLES
- update D_{i+1}

Bshouty, Jackson 1999

(1) find χ_a using QEX

$$\hat{f}^2(a) = \mathbb{E}^2[f \chi_a]$$

Algo

- start with empty tape
- QEX(f, U)
- FT
- sample and repeat

$$\frac{1}{\sqrt{2^n}} \sum_x |x, f(x)\rangle$$

↓ FT

$$\sum_a \hat{f}(a) \, |y, \hat{f}'(y)\rangle \; \hat{f}(a')$$

- QEX can be used to efficiently weak-learn
- QEX can be used to get errors of weak-learner