Gowers Norm, Function Limits, and Parameter Estimation

Yuichi Yoshida

National Institute of Informatics, and
Preferred Infrastructure, Inc.

January 12, 2016

Affine-invariant Parameter

Definition

A *parameter* π maps a function $f : \mathbb{F}_2^n \to \{0, 1\}$ to a value in $[0, 1]$.

Affine-invariant Parameter

Definition

A *parameter* π maps a function $f : \mathbb{F}_2^n \to \{0,1\}$ to a value in $[0,1]$.

Definition

A parameter π is *affine-invariant* if $\pi(f) = \pi(f \circ A)$ for any bijective affine transformation $A : \mathbb{F}_2^n \to \mathbb{F}_2^n$.

E.g.

- $\#$ of ones divided by 2^n.
- Minimum Hamming distance to a linear function / 2^n.
- Spectral norm ($=$ the sum of absolute Fourier coefficients) / 2^n.

Parameter Estimation

Definition

An algorithm is an *estimator* for a parameter π if, given

- $n \in \mathbb{N}$,
- a query access to $f : \mathbb{F}_2^n \to \{0, 1\}$, and
- an error parameter $\epsilon > 0$,

it approximates $\pi(f)$ to within ϵ with probability at least $2/3$.

Parameter Estimation

Definition

An algorithm is an *estimator* for a parameter π if, given

- $n \in \mathbb{N}$,
- a query access to $f : \mathbb{F}_2^n \to \{0, 1\}$, and
- an error parameter $\epsilon > 0$,

it approximates $\pi(f)$ to within ϵ with probability at least $2/3$.

Definition

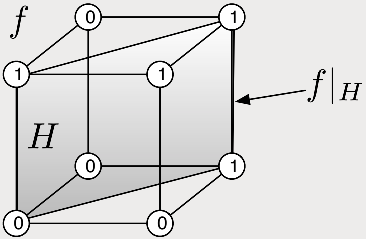
π is *constant-query estimable* if there is an estimator with query complexity that depends only on ϵ (and not on n).

Oblivious Estimator

Definition

A (constant-query) *oblivious estimator*

- Samples a random affine subspace H of dimension $h(\epsilon)$.
- Determines its output based only on the restriction $f|_H$.

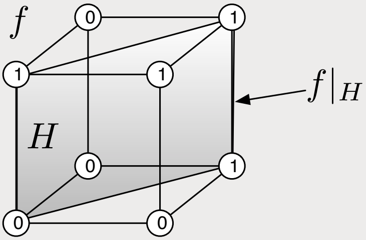


Oblivious Estimator

Definition

A (constant-query) *oblivious estimator*

- Samples a random affine subspace H of dimension $h(\epsilon)$.
- Determines its output based only on the restriction $f|_H$.



- Avoid "unnatural" parameters such as $\pi(f) = n \bmod 2$.
- For natural parameters, a constant-query estimator implies an oblivious constant-query estimator.

Main Result

Theorem (Informal)

An affine-invariant parameter π is (obliviously) constant-query estimable

$$\Updownarrow$$

For any function sequence $(f_i : \mathbb{F}_2^i \to \{0,1\})_{i \in \mathbb{N}}$ that "converges" in a certain metric, the sequence $\pi(f_i)$ converges.

Related work:

- A similar characterization for (dense) graphs [LS06].

Applications: Property testing

Definition

$f : \mathbb{F}_2^n \to \{0,1\}$ is *ϵ-far* from $\mathcal{P}$ if,

$$d_{\mathcal{P}}(f) := \min_{g \in \mathcal{P}} \#\{x \in \mathbb{F}_2^n \mid f(x) \neq g(x)\}/2^n \geq \epsilon.$$

Applications: Property testing

Definition

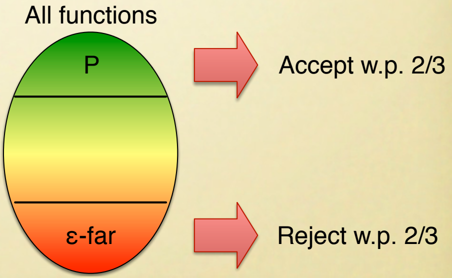
$f : \mathbb{F}_2^n \to \{0,1\}$ is *ϵ-far* from $\mathcal{P}$ if,

$$d_{\mathcal{P}}(f) := \min_{g \in \mathcal{P}} \#\{x \in \mathbb{F}_2^n \mid f(x) \neq g(x)\}/2^n \geq \epsilon.$$

A *tester* for a property $\mathcal{P}$:

Given

- $n \in \mathbb{N}$,
- a query access to $f : \mathbb{F}_2^n \to \{0,1\}$, and
- an error parameter $\epsilon > 0$,



Property Testing: Characterization

Corollary (Informal)

An affine-invariant property $\mathcal{P}$ is constant-query testable

$$\Updownarrow$$

For any function sequence $(f_i : \mathbb{F}_2^i \to \{0,1\})_{i \in \mathbb{N}}$ that "converges" in a certain metric, the sequence $d_{\mathcal{P}}(f_i)$ converges.

Simplified a previous characterization [Yos14], which involves many quantifiers and objects with seven parameters (regularity-instances).

Property Testing: Specific Properties

Corollary (Informal)

Suppose that a property $\mathcal{P}$ satisfies:

- *Any $f \in \mathcal{P}$ is of the form*

$$f(x) = \Gamma(P_1(x), \ldots, P_c(x), Q_1(x), \ldots, Q_{c'}(x)),$$

  where P_i's are low-degree polynomials, Q_i's are products of linear functions, $c + c' = O(1)$, $\Gamma : \mathbb{F}_2^{c+c'} \to \{0,1\}$.
- *(A minor condition)*

Then, $\mathcal{P}$ is obliviously constant-query testable.

Includes low-degree polynomials and having small spectral norm.

"Convergence" in a Certain Metric

"For any function sequence $(f_i : \mathbb{F}_2^{n_i} \to \{0,1\})$ *that converges in a certain metric, the sequence* $\pi(f_i)$ *converges."*

We have two issues:

- Metric?

"Convergence" in a Certain Metric

"For any function sequence $(f_i : \mathbb{F}_2^{n_i} \to \{0,1\})$ *that converges in a certain metric, the sequence* $\pi(f_i)$ *converges."*

We have two issues:

- Metric? $\Rightarrow$ *Gowers norm*

"Convergence" in a Certain Metric

"For any function sequence $(f_i : \mathbb{F}_2^{n_i} \to \{0,1\})$ *that converges in a certain metric, the sequence* $\pi(f_i)$ *converges."*

We have two issues:

- Metric? $\Rightarrow$ *Gowers norm*
- Convergence of functions on different domains?

"Convergence" in a Certain Metric

"For any function sequence $(f_i : \mathbb{F}_2^{n_i} \to \{0,1\})$ *that converges in a certain metric, the sequence* $\pi(f_i)$ *converges."*

We have two issues:

- Metric? $\Rightarrow$ *Gowers norm*
- Convergence of functions on different domains? $\Rightarrow$ *Non-standard analysis*

Gowers Norm

Definition

Let $f : \mathbb{F}_2^n \to \mathbb{R}$. The *Gowers norm of order d* for f is

$$\|f\|_{U^d} := \left(\mathop{\mathbf{E}}_{x, y_1, \ldots, y_d} \prod_{I \subseteq \{1, \ldots, d\}} f(x + \sum_{i \in I} y_i) \right)^{1/2^d}.$$

- $\| \cdot \|_{U^d}$ measures correlation with "polynomials" of degree $< d$.

A Metric for Functions on an Identical Domain

$\mu_{f,h}$: distribution of f restricted to an affine subspace of dimension h.

Fact

$\|f - g \circ A\|_{U^d}$ *is small (for large* d*)* $\Rightarrow \mu_{f,h} \approx \mu_{g,h}$.

A Metric for Functions on an Identical Domain

$\mu_{f,h}$: distribution of f restricted to an affine subspace of dimension h.

Fact

$\|f - g \circ A\|_{U^d}$ *is small (for large* d*)* $\Rightarrow \mu_{f,h} \approx \mu_{g,h}$.

Define

$$\upsilon^d(f, g) := \min_{A:\text{affine bijection}} \|f - g \circ A\|_{U^d}$$

Fact

$\upsilon^d(f, g)$ *is small* $\Leftrightarrow \mu_{f,h} \approx \mu_{g,h}$.

A Metric for Functions on an Identical Domain

Observation

Constant-query estimability $\Leftrightarrow$ *small* $\upsilon^d(f, g)$ *implies* $\pi(f) \approx \pi(g)$.

Proof sketch.

π is constant-query estimable.

$\Leftrightarrow$ If f and g are indistinguishable by a constant-query estimator (i.e., $\mu_{f,h} \approx \mu_{g,h}$), then $\pi(f) \approx \pi(g)$.

$\Leftrightarrow$ Small $\upsilon^d(f, g)$ implies $\pi(f) \approx \pi(g)$. $\square$

Convergence of a Function Sequence

If υ^d were a metric defined over functions on different domains, then

"small $\upsilon^d(f, g)$ implies $\pi(f) \approx \pi(g)$"

can be rephrased as

"If a function sequence $(f_i : \mathbb{F}_2^i \to \{0, 1\})_{i \in \mathbb{N}}$ converges in the υ^d-metric, then $\pi(f_i)$ converges."

Convergence of a Function Sequence

If υ^d were a metric defined over functions on different domains, then

> *"small $\upsilon^d(f,g)$ implies $\pi(f) \approx \pi(g)$"*

can be rephrased as

> *"If a function sequence $(f_i : \mathbb{F}_2^i \to \{0,1\})_{i\in\mathbb{N}}$ converges in the υ^d-metric, then $\pi(f_i)$ converges."*

To make this statement meaningful, we extend υ^d using *non-standard analysis*.

Brief Introduction to Non-standard Analysis

Non-standard analysis allows us to syntactically define a limit of any sequence (even if there's no metric).

Brief Introduction to Non-standard Analysis

- ω: a "nice" family of subsets of $\mathbb{N}$. (*non-principal ultrafilter*)

Brief Introduction to Non-standard Analysis

- ω: a "nice" family of subsets of $\mathbb{N}$. (*non-principal ultrafilter*)
- Introduce an equivalence relation $\sim$ on number sequences, where $(a_i)_{i \in \mathbb{N}} \sim (b_i)_{i \in \mathbb{N}}$ iff $\{i \in \mathbb{N} \mid a_i = b_i\} \in \omega$.

Brief Introduction to Non-standard Analysis

- ω: a "nice" family of subsets of $\mathbb{N}$. (*non-principal ultrafilter*)
- Introduce an equivalence relation $\sim$ on number sequences, where $(a_i)_{i \in \mathbb{N}} \sim (b_i)_{i \in \mathbb{N}}$ iff $\{i \in \mathbb{N} \mid a_i = b_i\} \in \omega$.
- The *ultralimit* of a sequence $(a_i)_{i \in \mathbb{N}}$, denoted by $\lim_{i \to \omega} a_i$, is the equivalence class it belongs to.

Brief Introduction to Non-standard Analysis

- ω: a "nice" family of subsets of $\mathbb{N}$. (*non-principal ultrafilter*)
- Introduce an equivalence relation $\sim$ on number sequences, where $(a_i)_{i \in \mathbb{N}} \sim (b_i)_{i \in \mathbb{N}}$ iff $\{i \in \mathbb{N} \mid a_i = b_i\} \in \omega$.
- The *ultralimit* of a sequence $(a_i)_{i \in \mathbb{N}}$, denoted by $\lim_{i \to \omega} a_i$, is the equivalence class it belongs to.
- Most operations can be naturally lifted to ultralimits.
  - E.g. $\lim_{i \to \omega} a_i + \lim_{i \to \omega} b_i = \lim_{i \to \omega} (a_i + b_i)$.

Brief Introduction to Non-standard Analysis

- ω: a "nice" family of subsets of $\mathbb{N}$. (*non-principal ultrafilter*)
- Introduce an equivalence relation $\sim$ on number sequences, where $(a_i)_{i\in\mathbb{N}} \sim (b_i)_{i\in\mathbb{N}}$ iff $\{i \in \mathbb{N} \mid a_i = b_i\} \in \omega$.
- The *ultralimit* of a sequence $(a_i)_{i\in\mathbb{N}}$, denoted by $\lim_{i\to\omega} a_i$, is the equivalence class it belongs to.
- Most operations can be naturally lifted to ultralimits.
  - E.g. $\lim_{i\to\omega} a_i + \lim_{i\to\omega} b_i = \lim_{i\to\omega}(a_i + b_i)$.
- A first order sentence ϕ is true in the ultralimit world $\Leftrightarrow$ ϕ is true for ω-many i's. (*Łoś' theorem*)
  - E.g. $\lim_{i\to\omega} a_i + \lim_{i\to\omega} b_i = \lim_{i\to\omega} c_i \Leftrightarrow \{i \mid a_i + b_i = c_i\} \in \omega$.

υ^d-Metric over Function Limits

The *function limit* $\mathbf{f}$ of a function sequence $(f_i : \mathbb{F}_2^i \to \{0,1\})$ is defined as

$$\mathbf{f}(\lim_{i \to \omega} x_i) = \lim_{i \to \omega} f_i(x_i).$$

(Formally, we take the standard part)

υ^d-Metric over Function Limits

The *function limit* $\mathbf{f}$ of a function sequence $(f_i : \mathbb{F}_2^i \to \{0,1\})$ is defined as

$$\mathbf{f}(\lim_{i\to\omega} x_i) = \lim_{i\to\omega} f_i(x_i).$$

(Formally, we take the standard part)

Definition

For two function limits $\mathbf{f}, \mathbf{g}$, we define

$$\upsilon^d(\mathbf{f}, \mathbf{g}) := \inf_{\mathbf{A}} \|\mathbf{f} - \mathbf{g} \circ \mathbf{A}\|_{U^d},$$

where $\mathbf{A}$ is over ultralimits of sequences of affine bijections.

Non-standard Analysis

Definition

For a function $f : \mathbb{F}_2^n \to \{0,1\}$, let

$$^*f = \text{the function limit of the sequence } (f \circ A_i)_{i \in \mathbb{N}},$$

where $A_i : \mathbb{F}_2^i \to \mathbb{F}_2^n$ is an arbitrary full-rank affine transformation.

Non-standard Analysis

Definition

For a function $f : \mathbb{F}_2^n \to \{0, 1\}$, let

$$^*f = \text{the function limit of the sequence } (f \circ A_i)_{i \in \mathbb{N}},$$

where $A_i : \mathbb{F}_2^i \to \mathbb{F}_2^n$ is an arbitrary full-rank affine transformation.

Definition

(f_i) is *υ^d-convergent* if the sequence $({}^*f_i)$ converges in the υ^d-metric.

The choice of A_i's is not important when discussing υ^d-convergence.

Main Result

Using the same idea as the identical domain case, we obtain:

Theorem

An affine-invariant parameter π is (obliviously) constant-query estimable

$$\Updownarrow$$

If a function sequence $(f_i : \mathbb{F}_2^i \to \{0,1\})_{i \in \mathbb{N}}$ is υ^d-convergent for any $d \in \mathbb{N}$, then the sequence $\pi(f_i)$ converges.

Proof ingredients:

- Tools from higher order Fourier analysis: non-classical polynomials, decomposition theorem.
- Another notion of convergence.

Summary and Open Problems

- Defined v^d-metric over function limits and obtained a concise characterization of constant-query estimable affine-invariant parameters.
- $\mathbb{F}_2$ can be generalized to $\mathbb{F}_p$ for any prime p, and for any prime power using recent techniques [BL15, BB15].

Summary and Open Problems

- Defined v^d-metric over function limits and obtained a concise characterization of constant-query estimable affine-invariant parameters.
- $\mathbb{F}_2$ can be generalized to $\mathbb{F}_p$ for any prime p, and for any prime power using recent techniques [BL15, BB15].

- Can we use our characterization to show other specific parameters are constant-query estimable?
- Can we characterize properties that are constant-query testable with one-sided error using function limits?