How much do we need to know to recognise a process?

Nishant Chandgotia[1] Benjamin Weiss[2]

[1]Tata Institute of Fundamental Research, Bangalore

[2]Hebrew University of Jerusalem

October, Expanding Dynamics

The Central Question

The context for this presentation is that of “single orbit dynamics”:

The Central Question

The context for this presentation is that of “single orbit dynamics”: What can a sample reveal about a process?

The Central Question

The context for this presentation is that of "single orbit dynamics": What can a sample reveal about a process?

In this talk all processes, denoted by $X_i; i \in \mathbb{Z}$, will be stationary, Radon, ergodic and conservative processes defined on spaces of infinite measure and take value in a countable discrete set A.

The Central Question

The context for this presentation is that of "single orbit dynamics": What can a sample reveal about a process?

In this talk all processes, denoted by $X_i; i \in \mathbb{Z}$, will be stationary, Radon, ergodic and conservative processes defined on spaces of infinite measure and take value in a countable discrete set A.

Question

For which sets $P \subset \mathbb{Z}$, can we identify the process given its sample on the set P, that is, $X_i; i \in P$?

Some notation: Cylinder Sets

Given a finite set $B \subset \mathbb{Z}$ and $w \in A^B$, we write

$$[w] := \{x \in A^{\mathbb{Z}} \; : \; x|_B = w\}.$$

Such sets are called cylinder sets.

Some notation: Cylinder Sets

Given a finite set $B \subset \mathbb{Z}$ and $w \in A^B$, we write

$$[w] := \{x \in A^{\mathbb{Z}} \ : \ x|_B = w\}.$$

Such sets are called cylinder sets.

We will frequently use the correspondence between processes and infinite measure preserving transformations: The process will be denoted by $X_i; i \in \mathbb{Z}$ while the corresponding measure preserving action will be denoted by $(A^{\mathbb{Z}}, \mu, \sigma)$ where σ is the shift map.

What if we know everything?

Hopf's ratio ergodic theorem says that if $f, g \in L^1(\mu); g \geq 0$ then for almost every $x \in A^{\mathbb{Z}}$,

$$\lim_{n\to\infty} \frac{\sum_{i=1}^{n} f(\sigma^i(x))}{\sum_{i=1}^{n} g(\sigma^i(x))} = \frac{\int f \ d\mu}{\int g \ d\mu}.$$

What if we know everything?

Hopf's ratio ergodic theorem says that if $f, g \in L^1(\mu); g \geq 0$ then for almost every $x \in A^{\mathbb{Z}}$,

$$\lim_{n\to\infty} \frac{\sum_{i=1}^{n} f(\sigma^i(x))}{\sum_{i=1}^{n} g(\sigma^i(x))} = \frac{\int f \ d\mu}{\int g \ d\mu}.$$

Thus if a sample of a process is given on $\mathbb{N}$ then (up to measure zero), we can apply Hopf's ratio ergodic theorem to find the ratio of $\mu([u])$ and $\mu([v])$ for all cylinder sets $[u], [v]$ enabling us to recognise the underlying process.

What if we just know the process only on the even integers?

Take two processes: One which alternates between the symbols 1 and 2 and another which alternates between 1 and 3.

What if we just know the process only on the even integers?

Take two processes: One which alternates between the symbols 1 and 2 and another which alternates between 1 and 3.

Then given a sample of either process, we will not be able to identify it with probability $1/2$.

These two examples motivate the following definition.

Recognition sets

Let $Meas(A^{\mathbb{Z}})$ denote the space of all ergodic, conservative (under the shift action), Radon measures.

Recognition sets

Let $Meas(A^{\mathbb{Z}})$ denote the space of all ergodic, conservative (under the shift action), Radon measures.

A set $P \subset \mathbb{Z}$ is called a recognition set if there is a measurable function

$$f : A^P \to Meas(A^{\mathbb{Z}})$$

such that for all processes $X_i; i \in \mathbb{Z}$ corresponding to measure μ,

$$f(X_i; i \in P) = k\mu$$

for some $k > 0$ (up to samples of measure zero).

Recognition sets

Let $Meas(A^{\mathbb{Z}})$ denote the space of all ergodic, conservative (under the shift action), Radon measures.

A set $P \subset \mathbb{Z}$ is called a recognition set if there is a measurable function

$$f : A^P \to Meas(A^{\mathbb{Z}})$$

such that for all processes $X_i; i \in \mathbb{Z}$ corresponding to measure μ,

$$f(X_i; i \in P) = k\mu$$

for some $k > 0$ (up to samples of measure zero).

$\mathbb{N}$ is a recognition set while $2\mathbb{N}$ is not.

Question

Is there a nice characterisation of recognition sets?

Question

Is there a nice characterisation of recognition sets?

While we have not found one, we do have a partial understanding.

Recognition sets for processes on probability spaces

Let $\mathcal{P}(A^{\mathbb{Z}})$ denote the space of all ergodic (under the shift action) probability measures.

Recognition sets for processes on probability spaces

Let $\mathcal{P}(A^{\mathbb{Z}})$ denote the space of all ergodic (under the shift action) probability measures.

$P \subset \mathbb{Z}$ is called a recognition set for processes on probability spaces if there is a measurable function

$$f : A^P \to \mathcal{P}(A^{\mathbb{Z}})$$

such that for all processes $X_i; i \in \mathbb{Z}$ corresponding to a probability measure μ,

$$f(X_i; i \in P) = \mu.$$

(up to sets of measure zero).

Recognition sets for processes on probability spaces

Let $\mathcal{P}(A^{\mathbb{Z}})$ denote the space of all ergodic (under the shift action) probability measures.

$P \subset \mathbb{Z}$ is called a recognition set for processes on probability spaces if there is a measurable function

$$f : A^P \to \mathcal{P}(A^{\mathbb{Z}})$$

such that for all processes $X_i; i \in \mathbb{Z}$ corresponding to a probability measure μ,

$$f(X_i; i \in P) = \mu.$$

(up to sets of measure zero).

Theorem (Chandgotia, Weiss)

P is a recognition set for processes on probability spaces if and only if it is thick, that is, it contains intervals of arbitrary size.

There are thick sets which are not recognition sets

A set $P \subset \mathbb{Z}$ is called thick if it contains intervals of arbitrary size.

We mentioned that this characterises recognition sets for processes on probability spaces in the previous slide. This is not true in more generality.

There are thick sets which are not recognition sets

A set $P \subset \mathbb{Z}$ is called thick if it contains intervals of arbitrary size.

We mentioned that this characterises recognition sets for processes on probability spaces in the previous slide. This is not true in more generality.

Theorem (Chandgotia, Weiss)

If $P \subset \mathbb{N}$ is recognition set then it must have bounded gaps.

There are thick sets which are not recognition sets

A set $P \subset \mathbb{Z}$ is called thick if it contains intervals of arbitrary size.

We mentioned that this characterises recognition sets for processes on probability spaces in the previous slide. This is not true in more generality.

Theorem (Chandgotia, Weiss)

If $P \subset \mathbb{N}$ is recognition set then it must have bounded gaps.

In particular, there are thick sets which do not have bounded gaps so there are recognition sets for processes on probability spaces which are not recognition sets (on spaces of infinite measure).

IP* sets

Theorem (Chandgotia, Weiss)

If $P \subset \mathbb{N}$ is a recognition set (for all processes) then it must have bounded gaps.

IP* sets

Theorem (Chandgotia, Weiss)

If $P \subset \mathbb{N}$ is a recognition set (for all processes) then it must have bounded gaps.

This is a consequence of more general theorem which we now state. Given a sequence of positive integers $i_1, i_2, \ldots$, we write

$$IP(i_1, i_2, \ldots) := \{ \sum_{t \in \mathbb{N}} \epsilon_t i_t \ : \ \epsilon_t \text{ is 0 or 1} \}.$$

IP* sets

Theorem (Chandgotia, Weiss)

If $P \subset \mathbb{N}$ is a recognition set (for all processes) then it must have bounded gaps.

This is a consequence of more general theorem which we now state. Given a sequence of positive integers $i_1, i_2, \ldots$, we write

$$IP(i_1, i_2, \ldots) := \{\sum_{t \in \mathbb{N}} \epsilon_t i_t \ : \ \epsilon_t \text{ is 0 or 1}\}.$$

A set $P \subset \mathbb{N}$ is called IP* if it intersects all IP sets.

IP* sets

Theorem (Chandgotia, Weiss)

If $P \subset \mathbb{N}$ is a recognition set (for all processes) then it must have bounded gaps.

This is a consequence of more general theorem which we now state. Given a sequence of positive integers $i_1, i_2, \ldots$, we write

$$IP(i_1, i_2, \ldots) := \{ \sum_{t \in \mathbb{N}} \epsilon_t i_t \ : \ \epsilon_t \text{ is 0 or } 1\}.$$

A set $P \subset \mathbb{N}$ is called IP* if it intersects all IP sets.

IP* sets have bounded gaps.

Maharam sets

IP* sets have bounded gaps.

A set $P \subset \mathbb{N}$ is called a Maharam set if for all conservative measure preserving actions, (X, μ, T) and sets $A \subset X$ of positive measure, there exists $n \in P$ such that

$$\mu(A \cap T^{-n}A) > 0.$$

Maharam sets

IP* sets have bounded gaps.

A set $P \subset \mathbb{N}$ is called a Maharam set if for all conservative measure preserving actions, (X, μ, T) and sets $A \subset X$ of positive measure, there exists $n \in P$ such that

$$\mu(A \cap T^{-n}A) > 0.$$

Theorem (Chandgotia, Weiss)

A set $P \subset \mathbb{N}$ is a Maharam set if and only if it is IP. In particular it must have bounded gaps.*

A recognition set has to be a Maharam set and hence must have bounded gaps.

Examples: Removal of a sparse sequence

A sequence of natural numbers n_k is called sparse if $n_{k+1} - n_k$ is an increasing sequence.

Examples: Removal of a sparse sequence

A sequence of natural numbers n_k is called sparse if $n_{k+1} - n_k$ is an increasing sequence.

Theorem (Chandgotia, Weiss)

Let (X, μ, T) be a conservative, ergodic measures preserving action and $f \in L^1(\mu); f \geq 0$. Then for almost every x,

$$\lim_{k\to\infty} \frac{f(T^{n_1}x) + f(T^{n_2}x) + \ldots + f(T^{n_k}x)}{f(Tx) + f(T^2x) + \ldots + f(T^{n_k}x)} \to 0.$$

Examples: Removal of a sparse sequence

A sequence of natural numbers n_k is called sparse if $n_{k+1} - n_k$ is an increasing sequence.

Theorem (Chandgotia, Weiss)

Let (X, μ, T) be a conservative, ergodic measures preserving action and $f \in L^1(\mu); f \geq 0$. Then for almost every x,

$$\lim_{k\to\infty} \frac{f(T^{n_1}x) + f(T^{n_2}x) + \ldots + f(T^{n_k}x)}{f(Tx) + f(T^2x) + \ldots + f(T^{n_k}x)} \to 0.$$

It follows that $\mathbb{N} \setminus \{n_k \ : \ k \in \mathbb{N}\}$ is a recognition set.

Examples: Removal of a sparse sequence

A sequence of natural numbers n_k is called sparse if $n_{k+1} - n_k$ is an increasing sequence.

Theorem (Chandgotia, Weiss)

Let (X, μ, T) be a conservative, ergodic measures preserving action and $f \in L^1(\mu); f \geq 0$. Then for almost every x,

$$\lim_{k \to \infty} \frac{f(T^{n_1}x) + f(T^{n_2}x) + \ldots + f(T^{n_k}x)}{f(Tx) + f(T^2x) + \ldots + f(T^{n_k}x)} \to 0.$$

It follows that $\mathbb{N} \setminus \{n_k \ : \ k \in \mathbb{N}\}$ is a recognition set.

Question

Do all recognition sets have density one?

Summary

- We are interested in conservative, infinite, ergodic, invariant Radon measures on $\mathbb{N}^{\mathbb{Z}}$.
- A set $P \subset \mathbb{Z}$ is called recognition set if processes can be recognised by their samples restricted to P.
- Recognition sets in the natural numbers have bounded gaps.
- If n_k forms a sparse sequence then $\mathbb{N} \setminus \{n_k \ : \ k \in \mathbb{N}\}$ is a recognition set.

Questions:

- Is there a nice characterisation of recognition sets?
- Do they necessarily have density one?