Redundancy of information: Lowering effective dimension

Jun Le Goh
joint with Joseph Miller, Mariya Soskova, Linda Westrick

University of Wisconsin-Madison

19th International Conference on Computability and Complexity in Analysis
May 2022, online

Summary

We study the interaction between effective Hausdorff dimension

$$\dim(X) = \liminf_{n\to\infty} \frac{K(X \upharpoonright n)}{n} \in [0,1]$$

and Besicovitch pseudo-distance

$$d(X,Y) = \limsup_{n\to\infty} \frac{|(X \upharpoonright n)\Delta(Y \upharpoonright n)|}{n} \in [0,1]$$

of binary sequences. Specifically, fix $t < s$ in $[0,1]$.

- Given X with $\dim(X) = t$, how close to X can we find Y with $\dim(Y) = s$?
- Given Y with $\dim(Y) = s$, how close to Y can we find X with $\dim(X) = t$?

This line of inquiry was initiated by Greenberg, Miller, Shen, Westrick 2018 (henceforth GrMShW).

Kolmogorov complexity of strings

The Kolmogorov complexity $K(\sigma)$ of a finite binary string σ is the length of the shortest description of σ, where descriptions are given by a fixed universal Turing machine.

We are concerned with the asymptotics of $\frac{K(\sigma)}{|\sigma|}$ (where σ is an initial segment of some $X \in 2^\omega$), so it does not matter which universal Turing machine we fix.

Nor does it matter whether we use plain Kolmogorov complexity or prefix-free Kolmogorov complexity.

The entropy function $H : [0,1] \to [0,1]$

Given a string σ of length n, here is a way to describe it:

(1) specify the number of 1s and 0s in σ (say pn and $(1-p)n$ respectively), and

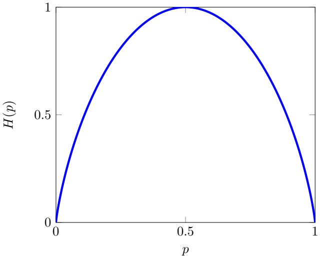
(2) specify σ among the strings of length n with pn many 1s.

(1) can be done with $O(\log n)$ bits.

(2) can be done with $H(p)n$ bits, where

$$H(p) = -p\log(p) - (1-p)\log(1-p)$$

is the entropy function.



Effective Hausdorff dimension of sequences

Definition (Lutz; Mayordomo)

The (effective Hausdorff) dimension of a sequence $X \in 2^\omega$ is

$$\dim(X) = \liminf_{n \to \infty} \frac{K(X \upharpoonright n)}{n} \in [0, 1].$$

Observations:

- Computable sequences have dimension 0.
- Martin-Löf random sequences have dimension 1.
- Flipping every bit in a sequence does not change its dimension.

Upper density and dimension

If a sequence X has upper density p, i.e.,

$$\limsup_{n\to\infty} \frac{|\{i < n : X(i) = 1\}|}{n} = p,$$

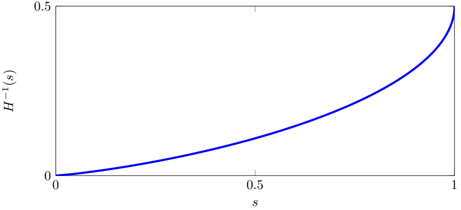
then we can bound the dimension of X in terms of p:

Proposition

A sequence with upper density p has dimension $\leq H(p)$.

Corollary

If a sequence has dimension s, then its upper density is at least $H^{-1}(s)$. (We use the branch $H^{-1} : [0,1] \to [0,1/2]$.)



Hamming distance and Besicovitch pseudo-distance

The Hamming distance $\Delta(\sigma, \tau)$ between strings $\sigma, \tau \in 2^n$ is the number of bits where they differ.

Definition

The (Besicovitch pseudo-)distance between sequences $X, Y \in 2^\omega$ is

$$d(X, Y) = \limsup_{n \to \infty} \frac{\Delta(X \upharpoonright n, Y \upharpoonright n)}{n} \in [0, 1].$$

Observations:

- The distance between X and $00\cdots$ is the upper density of X.
- If we modify X on a set of positions of upper density 0, then the result Y satisfies $d(X, Y) = 0$.

Distance versus dimension

Proposition (GrMShW)

If $\dim(X) = t$ and $\dim(Y) = s$, then $|s - t| \leq H(d(X, Y))$.

In particular:

1. The previous proposition is the special case where Y is $00\cdots$.
2. If $d(X, Y) = 0$, then X and Y have the same dimension.

Proof idea: We can describe an initial segment of X by describing the corresponding initial segment of Y, as well as their differences. This shows that

$$t \leq s + H(d(X, Y)).$$

Distance versus dimension

Proposition (GrMShW)

If $\dim(X) = t$ and $\dim(Y) = s$, then $|s - t| \leq H(d(X, Y))$, i.e.,

$$d(X, Y) \geq H^{-1}(|s - t|).$$

Motivating Question

Is this the best possible bound?

In a weak sense, yes:

Proposition (GrMShW)

For every $t < s$, there are X and Y with $\dim(X) = t$, $\dim(Y) = s$, and $d(X, Y) \leq H^{-1}(s - t)$ (hence $d(X, Y) = H^{-1}(s - t)$).

However, it is not the case that for every X of dimension t, there is some Y of dimension s such that $d(X, Y) \leq H^{-1}(s - t)$.

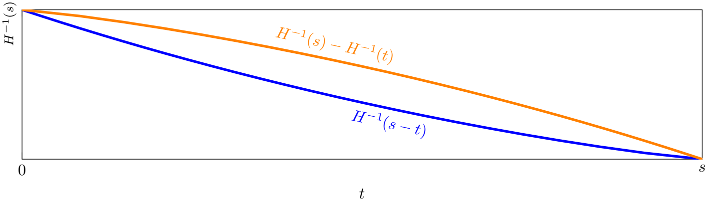
Raising dimension from t to s

Observation (GrMShW)

Suppose $t < s$. There is some X of dimension t such that for every Y of dimension s, $d(X, Y) > H^{-1}(s - t)$.

To see this, fix X with dimension t and density $H^{-1}(t)$. For every Y with dimension s, the density of Y is at least $H^{-1}(s)$, so

$$d(X, Y) \geq H^{-1}(s) - H^{-1}(t) > H^{-1}(s - t).$$



Raising dimension from t to s

Observation (GrMShW)

Fix X with dimension t and density $H^{-1}(t)$. For every Y with dimension s, the density of Y is at least $H^{-1}(s)$, so

$$d(X, Y) \geq H^{-1}(s) - H^{-1}(t).$$

The above is the worst that could happen when trying to increase the dimension of a given sequence X:

Theorem (GrMShW)

Suppose $t < s$. For every X of dimension t, there is some Y of dimension s such that $d(X, Y) \leq H^{-1}(s) - H^{-1}(t)$.

Lowering dimension from s to t

Given Y of dimension s, how close to Y can we find some X of dimension t?

$H^{-1}(s-t)$ is the closest that we can hope for, but this is not always attainable.

The information in Y may be stored redundantly, i.e., it may be hard to erase.

Lowering dimension from s to t: Redundancy in Y

(GrMShW) Take Y to be $Z \oplus Z$, where Z is a random.

Imagine you're trying to flip bits of Y in order to obtain an X of lower dimension.

In order for you to succeed, it must be hard to recover Y from X.

X can detect (for free) its inconsistencies, i.e., the i such that $X(2i) \neq X(2i+1)$. It is relatively cheap to fix all inconsistencies. Example:

X	$000011\mathbf{01}00\mathbf{10}11\mathbf{01}\cdots$
Extra info	$001\cdots$
$\tilde{X}$	$000011\mathbf{00}00\mathbf{00}11\mathbf{11}\cdots$

If, in addition to the above, we specify the set of i such that $X(2i) = X(2i+1) \neq Z(i)$, then we can recover all of Y.

Lowering dimension from s to t: Earlier results

Theorem (GrMShW)

For each Y of dimension s and each $t < s$, there is some X of dimension t with $d(X, Y) \leq H^{-1}(1 - t)$.

This was proved using the corresponding result for finite strings:

Proposition (GrMShW)

For each $\sigma \in 2^n$ and $t \in [0, 1]$, there is some $\tau \in 2^n$ such that

$$\frac{K(\tau)}{n} \lesssim t$$
$$\frac{\Delta(\sigma, \tau)}{n} \lesssim H^{-1}(1 - t).$$

If $s = 1$, the above theorem is optimal. How about if $s < 1$?

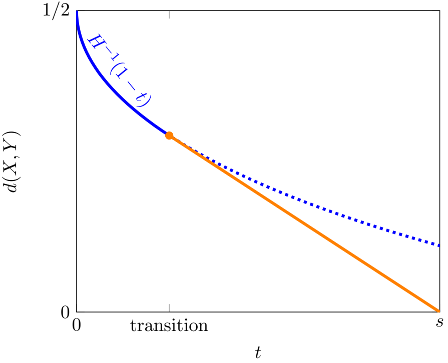
Lowering dimension from s to t: Another strategy

If $s < 1$, there is another strategy for finding a nearby X of dimension t.

The previous theorem was proved by applying the previous proposition to each interval in Y to obtain X. Instead:

- We leave some intervals in Y unchanged, and
- apply the previous proposition to the other intervals to obtain strings of dimension $< t$.

If t is sufficiently close to s, then this strategy is better.

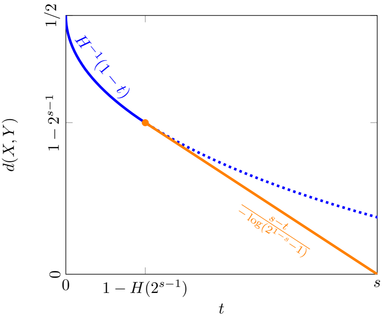


Lowering dimension from s to t: Our main theorem

Theorem (GoMSoW)

For each Y of dimension s and each $t < s$, there is some X of dimension t such that

$$d(X,Y) \leq \begin{cases} H^{-1}(1-t) & \text{if } t \leq 1 - H(2^{s-1}) \\ \frac{s-t}{-\log(2^{1-s}-1)} & \text{otherwise} \end{cases}.$$



For $s = 1$, this specializes to the previous theorem of GrMShW.

The above piecewise function is continuous, and even differentiable.

Lowering dimension from s to t: Optimality

Theorem (GoMSoW)

For each s, there is some Y_s of dimension s such that the previous bounds are optimal, i.e., for each $t < s$ and each X of dimension t,

$$d(X, Y_s) = \begin{cases} H^{-1}(1-t) & \text{if } t \leq 1 - H(2^{s-1}) \\ \frac{s-t}{-\log(2^{1-s}-1)} & \text{otherwise} \end{cases}.$$

Such Y_s are constructed in order to maximize the redundancy of their information.

Notice Y_s does not depend on t.

Lowering dimension from s to t: Constructing optimal Y_s

For integers $r \leq n$, a set $C \subseteq 2^n$ is an r-covering code if every string of length n lies at most distance r from C.

"Small" r-covering codes exist by a probabilistic argument (Delsarte, Piret 1986).

We show that "small" r-covering codes which are "well-distributed" exist, i.e., no string of length n is "close" to "too many" elements of the covering code.

Once we know that such covering codes exist, we can compute them by exhaustive search.

Henceforth, for each $r \leq n$, fix such a covering code C_r^n.

Witnesses of optimality: s-codewords

For each n, let I_n denote the interval of integers $\left[\sum_{i<n} i, \sum_{i<n+1} i\right)$. Note $|I_n| = n$.

Definition (GoMSoW)

For $s \in [0,1]$, we say that $Y_s \in 2^\omega$ is an s-codeword if there is some Y of dimension 1 and integers $\langle r_n \rangle_{n\in\omega}$ such that:

- $r_n \approx H^{-1}(1-s)n$
- for each n, $\Delta(Y \restriction I_n, Y_s \restriction I_n) \leq r_n$
- $Y_s \restriction I_n \in C^n_{r_n}$.

Facts:

1. Given Y of dimension 1, we can construct an s-codeword Y_s.
2. $\dim(Y_s) \geq s$ because $d(Y, Y_s) \leq H^{-1}(1-s)$.
3. $\dim(Y_s) \leq s$ because the covering codes $C^n_{r_n}$ are "small".

s-codewords are far from sequences of lower dimension

Suppose $t < s$, Y_s is an s-codeword, and $\dim(X) = t$.

Recall that Y_s restricted to the interval I_n lies in the r_n-covering code $C^n_{r_n}$.

We can describe $Y_s \upharpoonright I_n$ by:

(1) describing $X \upharpoonright I_n$
(2) specifying the distance q between $X \upharpoonright I_n$ and $Y_s \upharpoonright I_n$
(3) specifying $Y_s \upharpoonright I_n$ among the strings in $C^n_{r_n}$ which lie within distance q of $X \upharpoonright I_n$.

(3) can't be too short, otherwise we could give a simple description of $Y_s \upharpoonright I_n$.

We can obtain a lower bound for q by giving an upper bound for the number of bits needed for (3) in terms of q. Here we use the fact that $C^n_{r_n}$ is "well-distributed".

For each $t > s$, s-codewords are close to some sequence of dimension t

Proposition (GoMSoW)

Suppose $t > s$. For every s-codeword Y_s, there is some X of dimension t such that $d(X, Y_s)$ is as small as possible, i.e., $H^{-1}(t - s)$.

Thanks!