9-2015

Conditional Dimension in Metric Spaces: A Natural Metric-Space Counterpart of Kolmogorov-Complexity-Based Mutual Dimension

Vladik Kreinovich
The University of Texas at El Paso, vladik@utep.edu

Luc Longpre
The University of Texas at El Paso, longpre@utep.edu

Olga Kosheleva
The University of Texas at El Paso, olgak@utep.edu

Comments:

Technical Report: UTEP-CS-15-69

To appear in *Mathematical Structures and Modeling*, 2016, Vol. 37.

Conditional Dimension in Metric Spaces: A Natural Metric-Space Counterpart of Kolmogorov-Complexity-Based Mutual Dimension

Vladik Kreinovich, Luc Longpré, and Olga Kosheleva
University of Texas at El Paso
500 W. University
El Paso, TX 79968, USA
vladik@utep.edu, longpre@utep.edu, olgak@utep.edu

Abstract

It is known that dimension of a set in a metric space can be characterized in information-related terms – in particular, in terms of Kolmogorov complexity of different points from this set. The notion of Kolmogorov complexity $K(x)$ – the shortest length of a program that generates a sequence x – can be naturally generalized to *conditional* Kolmogorov complexity $K(x : y)$ – the shortest length of a program that generates x by using y as an input. It is therefore reasonable to use conditional Kolmogorov complexity to formulate a conditional analogue of dimension. Such a generalization has indeed been proposed, under the name of *mutual dimension*. However, somewhat surprisingly, this notion was formulated in pure Kolmogorov-complexity terms, without any analysis of possible metric-space meaning. In this paper, we describe the corresponding metric-space notion of conditional dimension – a natural metric-space counterpart of the Kolmogorov-complexity-based mutual dimension.

1 Need for a Metric Analogue of Mutual Dimension: Formulation of a Problem

What is dimension: an informal idea. A straight line segment S_1 is a 1-dimensional set, meaning that to select a point on this segment, it is sufficient to describe the value of a *single* real-valued quantity.

Similarly, a planar area S_2 is a 2-dimensional set meaning that to select a point in this area, we need to describe the values of *two* real-valued quantities: namely, two coordinates of this point.

A spatial area S_3 is a 3-dimensional set meaning that to select a point in

this area, we need to describe the values of *three* real-valued quantities: namely, three coordinates of this point, etc.

Metric dimension as a formalization of this informal idea. In practice, we can only describe a real number with some accuracy, and thus, we can only describe a point with some accuracy $\varepsilon > 0$.

Let us start with a straight line segment S_1 of length L. On this segment, two points which are ε-close are, within this accuracy, indistinguishable. So, if we start with a point on this segment, the next ε-distinguishable point has to be at a distance $> \varepsilon$. Thus, the overall number of ε-distinguishable points (or, equivalently, ε-distinguishable real numbers) is

$$N_\varepsilon(S_1) \sim \frac{L}{\varepsilon}.$$

In a 2-D domain S_2 of area A, we can place

$$N_\varepsilon(S_2) \sim \frac{A}{\varepsilon^2}$$

ε-distinguishable points: e.g., we can place such points on a rectangular grid, with $\sim \dfrac{c}{\varepsilon}$ distinguishable values along each dimension. This number is asymptotically equal to the number $(N_\varepsilon(S_1))^2$ of pairs of ε-distinguishable real numbers – which is in perfect accordance with the fact that we need two real numbers to describe a point in the 2-D domain S_2.

Similarly, in a 3-D domain S_3 of volume V, we can place

$$N_\varepsilon(S_3) \sim \frac{V}{\varepsilon^3}$$

ε-distinguishable points: e.g., we can place such points on a rectangular grid. This number is asymptotically equal to the number $(N_\varepsilon(S_1))^3$ of triples of ε-distinguishable real numbers – which is also in perfect accordance with the fact that we need three real numbers to describe a point in the 3-D domain S_2.

For a general metric space, we arrive at the following natural definition.

Definition 1.

- *Let $\varepsilon > 0$ be a real number. We say that a finite set F is an* ε-net *for the metric space S if every points from S is ε-close to one of the points from the set F.*
- *For each set S in a metric space M, let $N_\varepsilon(S)$ denote the smallest possible number of points in an ε-net of S. We say that the set S has* dimension *α if $N_\varepsilon(S) \sim \varepsilon^{-\alpha}$ as $\varepsilon \to 0$.*

Comment. This definition is the main case of the so-called Hausdorff (metric) dimension. This definition goes beyond the usual 1-D, 2-D, and 3-D spaces, since it is also applicable to irregular sets called *fractals* [3]: e.g., a trajectory of a Brownian motion has dimension $\alpha = 1.5$.

Dimension and information. Dimension can be also described in information terms, namely, in terms of the number of bits (0s or 1s) which are needed to uniquely determine a point in a metric space S with a given accuracy ε.

Specifically, to describe a point with a given accuracy, we need to pinpoint one of the $N_\varepsilon(S)$ ε-close points. If we use b-bit binary strings, then we can identify no more than 2^b different points. Thus, the smallest number of bits b which is needed to identify $N_\varepsilon(S)$ different points is the smallest integer for which $2^n \ge N_\varepsilon(S)$, i.e., the value $b = \lceil \log_2(N_\varepsilon(S)) \rceil$. This logarithm is usually denoted by $H_\varepsilon(S)$ and is called an *ε-entropy* of the metric space.

In terms of the ε-entropy $H_\varepsilon(S) = \log_2(N_\varepsilon(S))$, the definition $N_\varepsilon(S) \sim \varepsilon^{-\alpha}$ takes the form

$$H_\varepsilon(S) \sim -\alpha \cdot \log_2(\varepsilon).$$

Comment. This asymptotic relation is used as an alternative definition of metric dimension.

Relation with Kolmogorov complexity. For subsets of a real line or subsets of an Euclidean space, we can have yet another reformulation of metric dimension: in terms of so-called Kolmogorov complexity $K(x)$. Kolmogorov complexity $K(x)$ of a string x is defined as the shortest length of a program (in some fixed programming language) that is needed to generate the string x; see, e.g., [2].

To generate all infinitely many bits of a well-defined sequence of bits such as 00...or 0101..., we can use a program of finite length. However, to generate n bits of a truly random sequence x, bits which do not follow any law, we need to actually list all these n bits in the description of the generating program, so we have $K(x) \le n$ (to be more precise, $K(x) \ge n - c$ for some constant c). This is exactly why Kolmogorov complexity was invented in the first place: to formally describe the meaning of a random sequence.

In the 1-D case, if we select a random infinite binary sequence

$$x = x_1 x_2 \ldots x_n \ldots$$

that describes a random point in a segment, then for its initial fragments $x_1 \ldots x_n$ that describe this point with accuracy 2^{-n}, we get $K(x_1 x_2 \ldots x_n) \sim n$, i.e.,

$$\lim_{n \to \infty} \frac{K(x_1 \ldots x_n)}{n} = 1.$$

For non-random points, we need fewer bits, so dimension 1 can be defined as the largest possible value of the above limit over all points

$$x = x_1 x_2 \ldots x_n \ldots$$

from this segment.

In the 2-D case, if we select a random point, i.e., a random pair of sequences

$$(x_1, x_2) = (x_{11} x_{12} \ldots x_{1n} \ldots, x_{21} x_{22} \ldots x_{2n} \ldots),$$

then, to describe this point with accuracy 2^{-n}, we need $K(x_{11}x_{12}\ldots x_n x_{21}x_{22}\ldots x_{2n}) \sim 2n$, i.e., we have

$$\lim_{n\to\infty} \frac{K(x_{11}x_{12}\ldots x_n x_{21}x_{22}\ldots x_{2n})}{n} = 2.$$

For non-random points, we need fewer bits, so dimension 2 can be defined as the largest possible value of the above limit over all points

$$(x_1, x_2) = (x_{11}x_{12}\ldots x_{1n}\ldots, x_{21}x_{22}\ldots x_{2n}\ldots)$$

from the corresponding 2-D domain.

In general, for any d-dimensional point

$$x = (x_1, \ldots, x_d) = (x_{11}x_{12}\ldots x_{1n}\ldots, \ldots, x_{d1}x_{d2}\ldots x_{dn}\ldots)$$

in an Euclidean space, we can define its dimension $\dim(x)$ as the limit

$$\dim(x) = \lim_{n\to\infty} \frac{K(x_{11}x_{12}\ldots x_{1n}x_{21}x_{22}\ldots x_{2n}\ldots x_{d1}x_{d2}\ldots x_{dn})}{n}.$$

Then, for many reasonable sets $S \subseteq \mathbb{R}^d$, the metric dimension is equal to the largest dimension of the corresponding points:

$$\dim(S) = \max_{x\in S} \dim(x).$$

Conditional Kolmogorov complexity and mutual dimension. The notion of Kolmogorov complexity $K(x)$ has been naturally extended to the notion of *conditional* Kolmogorov complexity $K(x : y)$ as the shortest length of a program that, given y as an input, generates x.

In [1], this notion was used to produce the corresponding analog of dimension – which the authors called *mutual dimension*. Specifically, let us consider a Euclidean space $M = \mathbb{R}^d$ which is a represented as a Cartesian product $M = X \times Y$, where $X = \mathbb{R}^{d_x}$ and $Y = \mathbb{R}^{d_y}$ with $d_x + d_y = d$. Then, for every point $(x, y) \in M$, we can defined *mutual dimension* $\dim(x : y)$ as

$$\dim(x : y) \stackrel{\text{def}}{=} \lim_{n\to\infty} \frac{K(x_{11}\ldots x_{1n}x_{21}\ldots x_{d_x 1}\ldots x_{d_x n} : y_{11}\ldots y_{d_y 1}\ldots x_{d_y n})}{n}.$$

Challenge. The problem with the above definition is that it is described in purely Kolmogorov-complexity terms, it is not clear what is its metric analogue of this definition. Such a metric analogue is described in this paper.

2 Conditional Dimension in Metric Spaces: A Natural Definition and its Relation to Kolmogorov-Complexity-Based Mutual Dimension

What is conditional dimension: an intuitive idea. As we have mentioned, the usual metric-space dimension describes how many bits of information we need to describe a point in the given set $S \subseteq M$ with a given accuracy $\varepsilon > 0$.

In this case, a natural interpretation is that points from the set S (or, more generally, from the metric space M) represent physical objects: a point can be an actual point in space, or it can be a point that characterizes a physical object. From the practical viewpoint, this interpretation makes perfect sense:

- all we know about each object is the results of measuring different quantities related to this object,
- so it is natural to represent the object as a tuple consisting of the values of these measurement results.

For example, the state of a point-wise mechanical object can be characterized, at each moment of time, by a tuple consisting of 6 real numbers: 3 spatial coordinates and three components of the velocity vector.

To describe the state of a possibly rotating solid body (e.g., a planet or an asteroid), we need to supplement these 6 numbers with 2 angles describing this body's current orientation and 3 numbers describing its angular velocity.

In many practical situations, we have a system consisting of two interacting subsystems. In such situations, to describe the state of a system, we need to describe the pair (x, y) of states: the state x of the first subsystem and the state y of the second subsystem. In mathematical terms, the set S of all possible states of the system as a whole is thus a subset of the space $M = X \times Y$ of all possible pairs of states, where:

- X is the set of all possible states of the first subsystem and
- Y is the set of all possible states of the second subsystem.

It is natural to ask a question: if we know the state y of the second subsystem, how many bits do we need to describe the state x of the first subsystem?

If the states x and y were unrelated, then of course, the knowledge of y would be of no help. But when they are related, we expect that the knowledge of y can help us find x.

Towards a formal definition. Once we know y, we thus know that the state of possible values of x is limited to the set $\{x : (x, y) \in S\}$. Let us denote this set by S_y. For each y, it is thus reasonable to describe the corresponding number of bits as $\dim(S_y)$.

This number may depend on the choice of y. We want to make sure that the corresponding bound on the number of bits holds for all possible y, so we

should consider the largest of the corresponding values $\dim(S_y)$ as the proper description of the "conditional" dimension.

Thus, we arrive at the following definition.

Definition 2. *Let X and Y be metric spaces, and let S is a subset of the set $X \times Y$ of all pairs (x,y). By the* conditional dimension $\dim_{X:Y}(S)$, *we mean*

$$\dim_{X:Y}(S) \stackrel{\text{def}}{=} \max_{y\in Y} \dim(S_y),$$

where $S_y \stackrel{\text{def}}{=} \{x : (x,y) \in S\}$.

Examples. Let us consider examples in which X and Y are straight line segments and S is a straight line in the rectangle $X \times Y$.

If this straight line is neither parallel to X nor to Y, then the value x is uniquely determined by the value y, i.e., S is a graph of a linear function $x = f(y)$: $S = \{(f(y), y) : y \in Y\}$. In this case, as expected, the conditional dimension is equal to 0: $\dim_{X:Y}(S) = 0 < \dim(\pi_X(S)) = 1$, where $\pi_X(S) \stackrel{\text{def}}{=} \{x : (x,y) \in S\}$ is the set of all possible x-values, i.e., in mathematical terms, a projection of the set S on X.

If X and Y are straight line segments and S is a straight line in the rectangle $X \times Y$ which is parallel to Y, then knowing y does not provide us any information about x, so in this case, $\dim_{X:Y}(S) = \dim(\pi_X(S)) = 1$.

If S is the graph of the Brownian motion, i.e., y is time and x is the value of the Brownian motion at time y, then:

- knowing time y, we can uniquely determine the value x, so $\dim_{X:Y}(S) = 0$;
- on the other hand, when we know x, we can only determine y with uncertainty; the corresponding set has dimension 0.5, so $\dim_{Y:X}(S) = 0.5$.

In many of these cases, we have

$$\dim(\pi_X(S)) = \dim_{X:Y}(S) + \dim(\pi_Y(S)),$$

where $\pi_Y(S) \stackrel{\text{def}}{=} \{y : (x,y) \in S\}$ is the projection of the set S on Y. However, sometimes,

$$\dim(\pi_X(S)) < \dim_{X:Y}(S) + \dim(\pi_Y(S)).$$

For example, if S consists of two straight-line segments, one parallel to X and one parallel to Y, then $\dim_{X:Y}(S) = 1$ and $\dim(\pi_Y(S)) = 1$, but

$$\dim(\pi_X(S)) = 1 < \dim_{X:Y}(S) + \dim(\pi_Y(S)) = 1 + 1 = 2.$$

Relation to Kolmogorov complexity. In many of the above examples, we have $\dim_{X:Y}(S) = \max\limits_{(x,y)\in S} \dim(x : y)$.

Acknowledgments. This work was supported in part by the National Science Foundation grants HRD-0734825 and HRD-1242122 (Cyber-ShARE Center of Excellence) and DUE-0926721.

The authors are thankful to all the participants of the 2015 North American Annual Meeting of the Association for Symbolic Logic (Urbana, Illinois, March 25–28, 2015), especially to Jack Lutz, for valuable discussions.

References

[1] A. Case and J. H. Lutz, "Mutual dimension", *Electronic Colloquium on Computational Complexity*, 2014, Report No. 133.

[2] M. Li and P. M. B. Vitányi, *An Introduction to Kolmogorov Complexity*, Springer-Verlag, Berlin, Heidelberg, New York, 2008.

[3] B. Mandelbrot, *The Fractal Geometry of Nature*, Freeman, San Francisco, California, 1983.