Patterns in random point distributions in higher dimensional Cartesian spaces

Venkata Pagadala
June 2018

Abstract

This note discusses algebraic operations defined using the patterns observed in the uniform distributions of random points in a higher dimensional space, in the context of a square metric. Using this algebra, I analyze the random points as objects resulting from oscillations in a finite interval with no synchronization.

Keywords: High-dimensions, Random points, Cartesian space, Square metric

Introduction

The following results inspect random points defined using a process of combining random values as the N-dimensional coordinates of the points in an N-Dimensional space where N is considered large.

Definitions

1. *Square metric*: I define a metric called square metric for any two points in the Euclidean space as the square of the Euclidean distance between them. Example: The path length, using the square metric, between two diametrically opposite points on a circle is same via any third point on the circle, due to the Pythagorean identity.

2. A_{mr} is a random point in an $N-$ dimensional Euclidian space, with its i^{th} Cartesian coordinate as a uniform random value a_i in the interval $[m_i - r,\ m_i + r]$ where $m_i \in \mathbb{R}$ and $r \in \mathbb{R};\ 0 < r < \infty;\ i = 1,2,3,..,N$.

   A_{mr} can also be denoted as A_m or A, where r or both r and m are insignificant in the context.

3. The point M with coordinates $(m_1, m_2, .., m_N)$ is called the anchor point for A. A is a random point in the context of M.

4. $M \rightarrow A$ is the random point system where A is the random point in the context of M.

5. $M \rightarrow A \rightarrow B \rightarrow \ldots \rightarrow Y$ is the chain of random points $A, B, \ldots, Y$ where every point is the anchor for the next point.

6. The square metric $|P_1, P_2|$ of a pair of points P_1 and P_2, is the square of the Euclidean distance between the two points. R_r is the squared metric $|A_{mr}, M|$.

7. Q_{mr} is a randomly selected vertex of a $N -$ cube with $2r$ as its side length, and M as its center. The coordinates of Q_{mr} have an absolute magnitude of r and may differ in sign. All vertices of the cube have an equal probability of getting selected.

8. U_{mr} is a randomly selected vertex of a $N -$ octahedron with 2N vertices, 2^N faces of (N-1)-simplexes, each vertex is located r units away from the origin on a primary axis. All vertices of the polyhedron have an equal probability of getting selected.

9. A' is a point which has same coordinates as A, but in the sorted ascending order.

10. *Random radial*: A line joining A_m and its anchor point M

11. Two square metrics $|A, B|$ and $|B, C|$ are orthogonal if $|A, B| + |B, C| \approx |A, C|$. The angle $\angle ABC$ is nearly $\pi/2$.

12. Two square metrics $|A, B|$ and $|C, D|$ are considered skewed orthogonal if $|A, B| + |B, C| + |C, D| \approx |A, D|$. The angles $\angle ABC$ and $\angle BCD$ are nearly $\pi/2$.

13. A_0 is the normalized random point for A_m with coordinates of $\{a_i - m_i\}$

A few results and their proofs follow.

Result #1:

$$|A_{mr}\,, M| \approx R = Nr^2/3$$

Distance of a random point from its anchor is a function of N (number of dimensions) and r (size of the interval used for the random distribution)

Proof:

One can prove this result using the variance for a uniform random distribution. The following is an alternate proof, which sets a method for the subsequent proofs.

Let d be the difference between adjacent coordinates of A'

Let $m \approx N/2$

$d = r/m$

$$\sum_{i=1}^{N} a_i^2 \approx 2\sum_{i=1}^{m} i^2 d^2$$

$$\approx 2d^2 \frac{m(m+1)(2m+1)}{6}$$

$$\approx \frac{2r^2m}{3} + 1 + \frac{1}{3m}$$

$$\approx Nr^2/3$$

Result #2:

$$|A_{mr}| \approx |B_{mr}|$$

All random points are equi-distant from their common anchor.

Proof:

This follows from the Result #1 which shows that the distance depends only on N and r which are same for both points.

Result #3:

$$|A_m, P| \approx |A_m, M| \ + \ |M, P|$$

The line passing through a random point and its anchor is orthogonal to any other line passing through any other random or non-random point and the anchor.

This means all lines passing through a point M in $\mathbb{R}^N$ are orthogonal to all vectors of the random points having M as their common anchor.

Proof:
The proof uses two sub-results
1) Due to uniform distribution of a_i around m_i :

$$\sum_{i=1}^{N}(a_i - m_i) \approx 0$$

2) For a two-flip event of a coin, there is an equal probability of outcomes of the two flips being same or opposite. If the two-flip event represents selecting successive coordinates of two points, then nearly half of the coordinate combinations will have opposite signs while other half will have same signs.

$$\begin{aligned}|A_m, P| &\approx \sum_{i=1}^{N}(a_i - p_i)^2 \\ &\approx \sum_{i=1}^{N}((a_i - m_i) - (p_i - m_i))^2 \\ &\approx \sum_{i=1}^{N}(a_i - m_i)^2 + (p_i - m_i)^2 - 2(a_i - m_i)(p_i - m_i)\end{aligned}$$

Then, using sub-results #1 and #2, the third term reduces to sum of differences between pair-wise differences of coordinates of M and P, multiplied by the absolute value of $(a_i - m_i)$. This value becomes insignificant, resulting in:

$$|A_m, P| \approx |A_m, M| \ + \ |M, P|$$

Result #4:

$$|A_{mr}, B_{mr}| \approx |A_{mr}| + |B_{mr}| \approx 2R$$

Distance between two random points is $\sqrt{2}$ times their radial distance.

Proof:
Substituting B_{mr} for P in the Result #3:

$$|A_{mr}, B_{mr}| \approx |A_{mr}, M| + |M, B_{mr}|$$
$$\approx 2R$$

Result #5:

$$|A_{m_1 r_1}, B_{m_2 r_2}| \approx |A| + |M_1, M_2| + |B|$$

The squared distance between two random points A and B with different anchors and different spreads, is the sum of squared distances to their anchors and the squared distance between the anchors.

Proof:
Due to the Result #3, the radial $A_{m_1 r_1}$ and the line $M_1 M_2$ are orthogonal and the radial $B_{m_2 r_2}$ and the line $M_1 M_2$ are orthogonal.

So, it is enough to prove that A_0 and B_0 are orthogonal. The Result #4 can be applied after scaling B_0 such that $|A_0| \approx |B_0|$ which shows that A_0 and B_0 are orthogonal.

Result #6:

$$if\ M \rightarrow A \rightarrow B \rightarrow \ldots \rightarrow Y; then$$
$$|A| + |B| + .. + |Y| \approx |M, Y|$$

The squared distance between the anchor of first random point A and the last random point Y in a chain is, sum of squared lengths of radials of all the random points in the chain.

Proof:
Every adjacent set of three points in the chain involve a middle point being in the role of an anchor, resulting in orthogonality between the two lines produced by the three points. Also the skewed orthogonality is extended to all distant pairs of points, due to Result #5.

Then all segments build the distance between M and Y similar to coordinates resulting in the approximation.

Result #7:

$$|A_{mr}, Q_{mk}| \approx \left\{ \begin{array}{cl} 4R & : k = r \\ R + Nk^2 & : otherwise \end{array} \right\}$$

The distance between A and a random cube vertex Q is twice the length of radial of A.

Proof:
Due to the results of two-flip events of a coin as discussed in Result #3,
a half of the combinations will have opposite signs and the other half will have same signs.

$$|A_{mr}, Q_{mk}| \approx \sum_{i=1}^{N/2} (k - a_i)^2 + \sum_{i=1}^{N/2} (k + a_i)^2$$

$$\approx \sum_{i=1}^{N/2}(k^2 - 2ka_i + a_i{}^2) + \sum_{i=1}^{N/2}(k^2 + 2ka_i + a_i{}^2)$$

$$\approx 2\sum_{i=1}^{N/2} k^2 + 2\sum_{i=1}^{N/2} a_i{}^2$$

$$\approx R + Nk^2$$

$$\approx R(1 + 3k^2)$$

$$\approx 4R \quad : if\ k = r$$

Result #8:

$$|A_{mr}, U_{mk}| \approx R + k^2$$

The distance between A and a point k units away from M along a randomly selected axis is, $R + k^2$. This result follows from the Result #3 by substituting U for P, as a special case.

Result #9:

$$|A_{mr}, M| \approx R$$
$$|A_{mr}, B_{mr}| \approx 2R$$
$$|M, Q_{mr}| \approx 3R$$
$$|A_{mr}, Q_{mr}| \approx 4R$$

Square metrics of the radial, edge between A and B, cube half diagonal and distance between A to cube vertex, are in arithmetic progression.

Proof:

$$|M, Q_{mr}| = \sum_{i=1}^{N} r^2$$
$$= Nr^2$$
$$= 3R$$

And the Results #1, #4 and #7 provide the other approximations.

Result #10:

As with the other results, the following are approximations.

1. The diameter of the "sphere" of a random distribution of the points is $\sqrt{2}$ times the radius.
2. For a random point, the nearest other random point in the distribution is at the same distance as the one diametrically opposite to it, which is at $\sqrt{2}$ distance. In $3D$, this would mean equator and opposite pole being at same distance from a pole.
3. The random points in a distribution form a simplex with dimensions equal to the number of the random points, even if it is more than the dimensions of the space in which the points are created.
4. All points in the distribution appear to form the base "circle" of a cone, with the apex of cone being any point, outside or inside of the distribution. This is due to random points being equi-distant from any other point.

Result #11:

$$|A_{0r}\,.\;B_{0r}| \approx Rr^2/3$$

The square metric of a random point obtained by multiplying same coordinates of two normalized random points is $r^2/3$ times the square metric of the random points.

Discussion

The following are a few effects of the patterns observed in the above results.

Infeasibility of constructing randomness through coordinates

Several theorems and results were stated in the areas of data science and related mathematics describing the concentration of volume and surface area of a unit sphere near its surface and equator, in high-dimensional spaces. The statements included how the random points would follow these concentrations, resulting in orthogonality between any two random points etc.

However, as seen in the results of this note, the random points possess a significant amount of patterns. So I would hesitate to call them random points, though they were built using random coordinates.

One might argue that the random points are uniformly distributed without a pattern, but the space or the volume itself is non-uniformly distributed along lower-dimensional measures. This anomaly reflects as patterns in the random point distributions.

However, the presence of patterns, and not it's reasons, leads to questioning of the labeling these points as random points.

Unsynchronized oscillations producing a sphere

Two points moving on the two axes in a $2D$ space, in a synchronized oscillation produce coordinates for the locus of a circle. Three such points in a $3D$ space would produce a sphere.

In both cases of $2D$ and $3D$, the points are not independent because the freedom of the points is restricted by the synchronization. However, the freedom in $3D$ is more compared to $2D$, while meeting the requirements of synchronization.

In an $N-$ dimensional space with a large N, the constraint due to synchronization is insignificant or absent, such that, independent random coordinates of a random point can produce a $(N-1)-$ sphere through their unsynchronized oscillations.

Coordinate systems failing to preserve variety in distance

Variety in distance is the ability of a distance to be different from other distance. The Cartesian system builds distance between two points as a function of sum of squared differences of coordinate pairs across the two points. This sum is sensitive to the values of coordinates in lower dimensions. However, the sensitivity reduces as the dimensions increase, and the distance is no longer a function of coordinates in higher-dimensions. This is reflected in the results shown above, for the random points.

Looking through the lens of randomness

Randomness is defined as the absence of patterns or communications among a collection of things, numbers or objects. However, the very definition forbids the pure randomness, by enforcing a pattern through labelling the collection members as things, numbers or objects.

Most of the patterns or lack of randomness observed in the the results of this note, is due to the insignificance of number of random points compared to the number of quadrants in the higher dimensional Cartesian coordinate system. However, one could imagine non-random points with an unusual order or pattern and have everything back in order.