The curse of dimensionality

Mario Köppen
Fraunhofer IPK Berlin
Pascalstr. 8-9, 10587 Berlin, Germany
E-Mail: *mario.koeppen@ipk.fhg.de*

Abstract

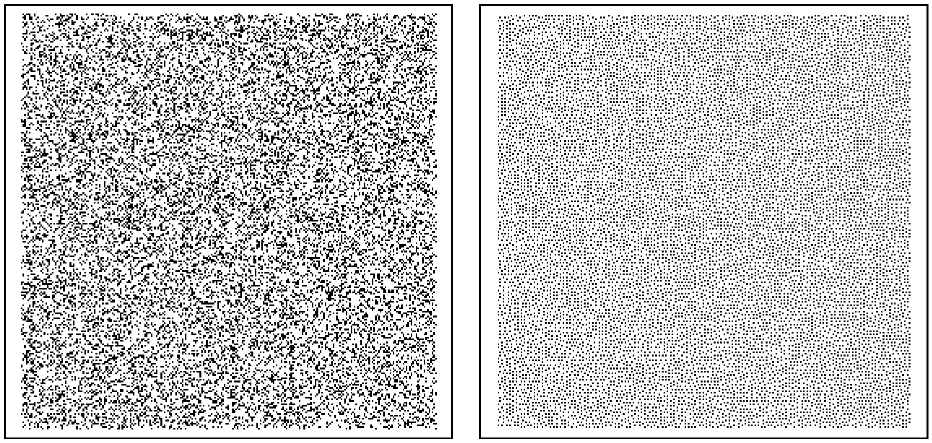
In this text, some question related to higher dimensional geometrical spaces will be discussed. The goal is to give the reader a feeling for geometric distortions related to the use of such spaces (e.g. as search spaces).

1 Introduction

In one of his essays “Glow, Big Glowworm”[3] Stephen Jay Gould reports on a visit of the Waitomo glowworm caves on Lake Roturura in New Zealand. He witnessed a magnificent natural spectacle, which motivated him for an interesting observation (idem, p. 260):

”‘Here, in utter silence, you glide by boat into a spectacular underground planetarium, an amphitheater lit with thousands of green dots — each the illuminated rear end of a fly larva (not a worm at all). (I was dazzled by the effect because I found it so unlike the heavens. Stars are arrayed in the sky at random with respect to the earth’s position. Hence, we view them as clumped into constellations. This may sound paradoxical, but my statement reflects a proper and unappreciated aspect of random distributions. Evenly spaced dots are well ordered for cause. Random arrays always include some clumping, just as we will flip several heads in a row quite often so long as we can make enough tosses — and our sky is not wanting for stars. The glowworms, on the other hand, are spaced more evenly because larvae compete with, and even eat, each other — and construct an exclusive territory. The glowworm grotto is an ordered heaven.)”’

By reading this, the american Nobel laureate Ed Purcell was inspired to perform a small experiment for the purpose of illustrating the *clumping* in random patterns, about which Gould spoke. The result was presented in the *Postscriptum* of the essay. Figure 1 shows the same experiment in a slightly modified version. Within an image of size 256×256 pixel, 20,000 positions were randomly selected. The difference between image 1 (a) and (b) is as follows: while the positions in figure (a) were selected with uniformly distributed x- and y-coordinates, for figure (b) only positions were allowed, which do not lie within the 8-neighborhood of an already selected position. Hence, figure (a) more ressembles the night sky case, while figure (b) more ressembles the glowworm case, for which each glowworm requires a minimum distance to each of its neighbors.



(a) (b)

Figure 1: Distribution of 20,000 randomly selected positions in a grid of size 256×256: without any further restrictions (a); and with the restriction that none of the selected positions may be direct neighbors (b).

The paradox given by Gould is related to the fact that the observer perceives features only in figure (a). We consider figure (b) a better representation of randomness due to the lack of positional differentiation. There seems to be no region in figure (b) distinguished from any other. But, for figure (a), the human cognition impresses several distinguishing features, despite of the fact that for the creation of figure (a) there was never a distinction among positions introduced. This was only the case for figure (b), for the creation of which positions were steadily prohibited!

In a letter to Gould (idem, p. 268), Ed Purcell wrote:

"'What interests me more in the random field of 'stars' is the overimposing impression of 'features' of one sort or another. It is hard to accept the fact that any perceived feature — be it string, clump, constellation, corridor, curved chain, lacuna — is a totally meaningless accident, having as its only cause the avidity for pattern of my eye and brain! Yet that is perfectly true in this case."'

The paradox is related to two interacting factors, a physiological one and a mathematical one.

1. Human physiology of visual perception, based on the model of receptive fields. It is widely accepted that the receptive fields in the neuronal layers of the visual cortex may be approximated by GABOR functions. Those functions are characterized by local frequency and exponential decay. During the so-called preattentive perception, local orientations are determined. Adjacent positions in the "star field" of figure 1 (a) more strongly guide to the perception of a local orientation than more distant ones. The further steps to the cognition of a local feature shall not be considered here.

2. The second factor is related to a mathematical issue. With the positions itself being equally distributed, the average distances of these positions are not equally distributed as well. To see this, we present a little numerical experiment. In a square of sidelength 1, two points are selected with a uniform distribution, and their distance is computed. This will be repeated for 100,000 times. The so-estimated distances will be divided among 100 "cups" according to their magnitudes. The first cup takes distances from 0 to $\sqrt{2}/100$, the last cup distances from $99/100\sqrt{2}$ to $\sqrt{2}$. This gives an estimate for the probability density of the distribution of the distances of two randomly selected points within the unit square. The computed frequencies, normalized to the total sum 1, are given in figure 2. As can be seen, the distances are not equally distributed! The distribution is a more left-oriented one. In combination with the exponential decay (see dotted line in figure 2) of receptive fields, we can understand why our perception of random patterns is so "clumpsy": since they are really present, and since they are strengthened by the functionality of our visual perception.

For now, we will not follow that *unappreciated aspect* of random patterns (see [5] for more examples) any further. But we will make a note on the following key aspect: the distribution of distances of uniformly distributed points in a search space by itself is not uniformly distributed.

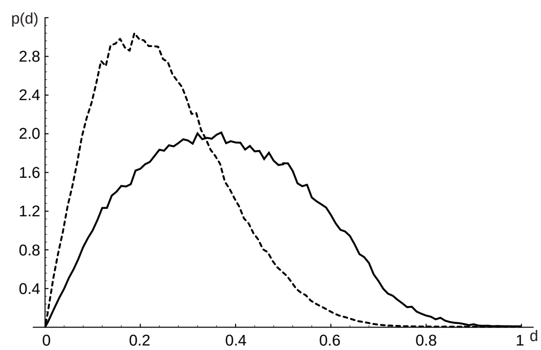


Figure 2: Estimation of the probability density $p(d)$ of the distance of two randomly selected points in the unit square (solid line) and after multiplication with $\frac{8}{\sqrt{\pi}}e^{-8d^2}$ for simulating the activation of a receptive field (dotted line).

Among the few researchers, which considered such phanomenons in more detail, is Ingo Rechenberg (the "father of evolutionary strategy"). In his book "Evolutionsstrategie'94" [7] there are many examples for such geometrical distortions in the context of higher dimensional spaces. Rechenberg even employs those issues for deriving a theoretical model for the progress of an evolutionary strategy, the so-called *sphere model.*

In the following sections, some of those aspects are presented. Section 2 gives some information about highdimensional regular bodies (hypercube, hypersimplex and hypersphere). Section 4 then reconsiders the distribution of distances in highdimensional spaces in a more accurate manner.

This will help to give the reader a feeling for what is reliable in higher dimensional spaces (e.g. search spaces) and what is not, and for which dimensions those effects become to be important.

Dimension	1	2	3	4	n
Count of corners	2	4	8	16	2^n
Count of edges	1	4	12	32	$n2^{n-1}$
Count of faces	0	1	6	24	$2^{n-2}\binom{n}{2}$
Count of bordering cubes	0	0	1	8	$2^{n-3}\binom{n}{3}$
Count of hypersurfaces	2	4	6	8	$2n$
Count of hypersurfaces per corner	1	2	3	4	n
Count of disjoint hypersurfaces per corner	1	2	3	4	n
Volume with sidelength l	l	l^2	l^3	l^4	l^n
Diagonallength with sidelength 1	1	$\sqrt{2}$	$\sqrt{3}$	2	$\sqrt{n}$

Table 1: Some counts and measures for the hypercube.

2 Geometrics of *n*-dimensional spaces

2.1 Hypercube

The most simple n-dimensional body is the hypercube. The corners of this geometrical body with sidelength l are all those points with coordinates being either 0 or l. Thus, the bordering faces of the hypercube are orthogonal to each other or parallel.

For human imagination, the hypercube is only accessible for the dimensions 2 and 3 by way of a square or cube respectively. For consistency, a line of length 1 is sometimes considered a unit hypercube of dimension 1. Figure 3 gives two projections of the 4-dimensional hypercube, the CAVALIER perspective and the central projection.

Table 1 gives some counts and measures for the hypercubes of dimensions 1, 2, 3 and 4 and for the general case as function of the dimension n. For deriving such values, the concept of a *scheme* can be employed. Base object are *bitstrings*, i.e. vectors with n components, all of which are either 0 or 1, and which may represent the corners of a hypercube. A scheme of size $m \leq n$ is a bitstring with m unspecified positions, as e.g. `10**01` being a scheme of size 2. The asterisk * stands for such an unspecified position (also known as wildcard symbol). The bitstrings `100101` or `101101` can be considered *realisations* of that scheme, but the bitstring `000101` not.

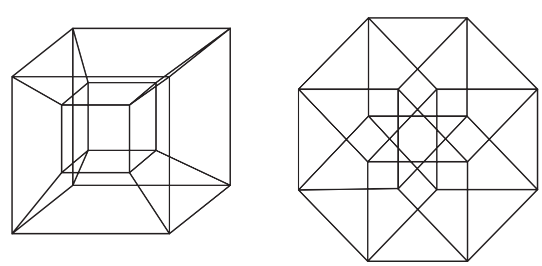


Figure 3: Two projections of the 4 dimensional hypercube (tesseract): central projection (left) and CAVALIER perspective (right).

For each scheme of size k there is exactly one hypercube of dimension k, which borders the hypercube of dimension n. So, an edge connects two corners of a hypercube. The connected corners only differ in one position of their coordinates. Therefore, an edge corresponds to a scheme of size 1. As an example, the scheme `00*` represents the edge going from corner $(0,0,0)$ to the corner $(0,0,1)$, or the scheme `**0` the lower face of a cube.

The count of schemes of size k equals the number of possibilities for selecting k wildcard positions out of n, multiplied with the number of possibilities to assign 0 or 1 to the remaining $n-k$ positions, i.e. 2^{n-k}. Therefrom it follows for the count N_k^n of k-dimensional hypercubes bordering the n-dimensional hypercube:

$$N_k^n = 2^{n-k}\binom{n}{k}. \tag{1}$$

Equation (1) was used to derive the values in table 1. The number of hypersurfaces, to which a corner belongs, can be found in a similar manner. It has to be counted, how many schemes of size $n-1$ are realized by a given bitstring. This are just n schemes, one for each bit position. The remaining $2n-n=n$ hypersurfaces of the hypercube are disjoint to that corner, and each connecting line from the corner to an inner point of one of those disjoint hypersurfaces completely lies within the hypercube.

All other entries in table 1 are obvious.

The volume of a n-dimensional unit hypercube is 1. For $n\to\infty$, the volume of a hypercube with $l>1$ goes to infinity, while for $l<1$ it goes to 0. Also, the length of the diagonal of a unit hypercube ($\sqrt{n}$) goes to infinity, the hypercube becomes more and more extended. According to [4], the hypercube can be imagined as a highly anisotropical body, more ressembling a spherical "hedgehog" than a convex body. The inner ball-like part with radius $1/2$ is covered with a large number (2^n) of "spikes" of length $\sqrt{n}/2$ (going to infinity for large n) (see figure 4). "'*The surfaces of cubes are so horribly jagged that they might even be thought of as being almost fractal.*"' ([4], p. 42).

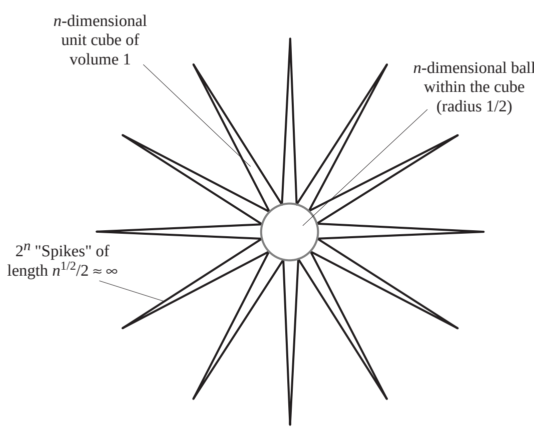


Figure 4: "'Spiking Hypercube"' [4].

Finally, a short computation will show the change of relative volumina within the hypercube, when problem dimension increases. On the main diagonal of a hypercube, a random point P is selected with coordinates $p_1=p_2=\ldots=p_n=p$ and $p>1/2$. This way, two subcubes of the hypercube are defined, with one including the point

$(0,0,\ldots,0)$ and a sidelength of p (A_1), and the other including the point $(1,1,\ldots,1)$ and a sidelength of $1-p$ (A_2). Now, P will be shifted slightly by the amount δ. The question is how to choose δ, given dimension n, in order to double the ratio r of the volumina of A_1 and A_2. The old and new value of r are given by

$$r_{old} = \left(\frac{p}{1-p}\right)^n \tag{2}$$

$$r_{new} = \left(\frac{p+\delta}{1-(p+\delta)}\right)^n. \tag{3}$$

By resolving $r_{new}/r_{old} = 2$ one gets:

$$\delta = \underbrace{\frac{p(1-p)}{1-p+\sqrt[n]{2}}}_{term1}\underbrace{(\sqrt[n]{2}-1)}_{term2}. \tag{4}$$

The interesting property of this result is that $\sqrt[n]{2}$ with $n \to \infty$ exponentially goes to 1. For large n, term 1 remains nearly constant, while term 2 rapidly goes to 0. This decay dominates term 1 for large n, too.

This means that the necessary shift of P becomes smaller and smaller, when dimension n increases. This decay is even exponential! And, of course, this holds for other ratios as 2 as well. Therefore, mental constructs as the segmentation of a hypercube into a system of subcubes, standing for classification boundaries or so, are very instable designs and opted to sudden breakdown when their carriers are slightly modified.

2.2 Hypersimplex

The hypersimplex is composed of $n+1$ equally distanced points in the n-dimensional space (there are maximal $n+1$ points in R_n with this property). Each two corners of the hypersimplex are connected by an edge of the hypersimplex, each three span a regular triangle, each four span a regular tetrahedron a.s.f. The n-dimensional hypersimplex is bounded by $\binom{n+1}{k+1}$ hypersimplexes of dimension k. In that sense, a hypersimplex is a very compact body.

In the appendix, some relations for the hypersimplex are derived. Therefrom, the volume of a hypersimplex with sidelength a and dimension n is given by (see equation (52))

$$V_n = \frac{1}{n!}\sqrt{\frac{n+1}{2^n}}\,a^n. \tag{5}$$

For large n, this expression is dominated by the decay of $1/n!$, and the volume goes rapidly to 0 (independently of the size of a).

For the height h_n of a hypersimplex, i.e. the length of the perpendicular from one corner to the hypersurface spanned by the remaining n points, one gets (see equation (49))

$$h_n = \sqrt{\frac{n+1}{2n}}\,a \tag{6}$$

and for the radius r_n of the surrounding ball, i.e. the hypersphere containing all $n+1$ corner points of the hypersimplex (see equation (50))

$$r_n = \sqrt{\frac{n}{2(n+1)}}a. \tag{7}$$

Therefore, for large n, h_n and r_n are approaching $1/\sqrt{2}$, and it holds

$$\lim_{n\to\infty} |h_n - r_n| = 0. \tag{8}$$

In other words, the center of the hypersimplex gets closer and closer to the centers of its outer surfaces, the hypersimplex collapses.

The inner angle of the hypersimplex, i.e. the angle between two lines connecting the center with two corners of the hypersimplex, is (see equation (55))

$$\theta = \arccos\left(-\frac{1}{n}\right). \tag{9}$$

For large n this goes to $\pi/2$, the hypersimplex becomes a more cube-like body.

Of course, all such relations are impossible in finite-dimensional spaces.

2.3 Hypersphere

Finally, the n-dimensional hypersphere will be considered. This is the geometrical place of all those points, which have a distance of maximal R from its center. Also, the surface of this body, with distance equal to R, is sometimes referred to as hypersphere. The 2-dimensional hypersphere is the common sphere, the 3-dimensional hypersphere the ball. Sometimes, the unit line is considered a 1-dimensional hypersphere.

For obtaining the volume of a hypersphere of radius R, it has to be distinguished whether n is odd or even. For $n = 2p$ even one gets (see appendix)

$$V_{2p} = \frac{R^{2p}\pi^p}{p!} \tag{10}$$

and for $n = 2p+1$ being odd

$$V_{2p+1} = \frac{R^{2p+1}2^{p+1}\pi^p}{1\cdot 3\cdot 5\cdot \ldots \cdot (2p+1)}. \tag{11}$$

With the following definition for $p!!$

$$p!! \quad = \quad 1 \text{ for } p < 2 \tag{12}$$
$$p!! \quad = \quad p\cdot(p-2)!! \text{ else} \tag{13}$$

this can be given in a more compact manner

$$V_n = \frac{2^{[(n+1)/2]}\pi^{[n/2]}}{n!!}R^n \tag{14}$$

with $[x]$ the largest integer smaller or equal to x.

n	V_n
1	2.00000
2	3.14159
3	4.18879
4	4.93480
5	5.26379
6	5.16771
7	4.72477
8	4.05871
9	3.29851
10	2.55016

Table 2: Volume of the unit cube for the dimensions 1 to 10.

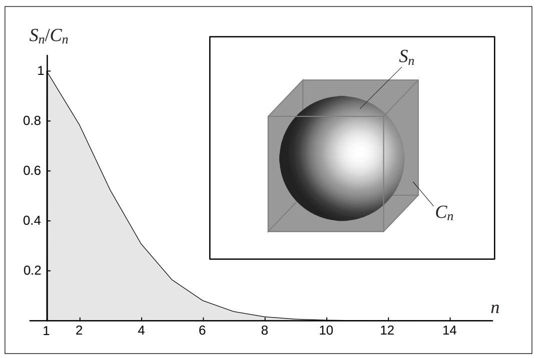


Figure 5: Ratio of the volumes of unit hypersphere and embedding hypercube of side-length 2 up to the dimension 14.

As it was the case for the hypersimplex, the volume of a hypersphere goes to 0, independent from the size of its radius. Table 2 gives the volumes of some unit cubes for $n = 1, \ldots, 10$. The hypersphere for $n = 5$ has the biggest volume, but this depends on R. Especially, for $R = 1/\sqrt{2}$, the hypersphere attains its maximum in "our" 3-dimensional world.

As an example consequence for search spaces, the ratio of volumes of unit hypersphere and embedding hypercube (with a sidelength of 2) will be considered (see figure 5). Despite of the fact that the volume of the hypercube goes to infinity, and the hypersphere touches all faces of the hypercube (i.e. at $2n$ points), the volume of the embedded hypersphere goes to 0! Moreover, for $n > 10$ we could neglect this volume part within the hypercube for all practical computations. This is an important distinction between search methods, which explores the hypercube, and search methods, which explores the hypersphere. The last ones will not "see" very much from their world.

Also, it has to be noted that the term "high dimension" may refer to values of n as small as 10 or so.

3 Short Break: The fourth spatial dimension

Geometrical relations in spaces with more than three spatial dimensions seems to be impossible to imagine. This is a remarkable property of the human cognition, the discussion of which will be shortly given in this section.

We will concentrate on the most simple question in this context about a fourth spatial dimension. No mathematical training seems to be suitable to access a mental state, which makes it possible to operate with the tesseract in the same way we are used to operate with objects in the 2- or 3-dimensional space. Therefore, it is no wonder that most efforts in this direction were made a little bit apart from traditional science.

Nevertheless, the choices offered by a possible access to a fourth spatial dimension are astonishing. Just to name a few:

- It would be possible to directly perceive the interior of a body.
- The content of a box could be taken without opening the box.
- A node could be removed from a string without moving the ends of the string.
- A body could be lifted without external forces.
- A body could be moved into its mirror form.

Of course, this reminds one more on the tricks of a magician than on real effects of physical interactions. So, first efforts on speculating about a fourth dimension can be found in some esoteric circles of the 19th century. The fourth dimension was regarded as the world of demons and ghosts, but even as the world where god and his angels used to live as well. One of the first more serious discussions of a fourth spatial dimension can be found in the essay "Flatland" of Edwin Abbott Abbott (1838–1926) from 1884 [1]. Besides of being a mathematical excurse into the questions of living in a 2-dimensional world, the essay also offers some critque on the victorian society. The main character of this essay, A Square, is sentenced to prison for teaching his "flat-minded" contemporaries the existence of a 3-dimensional world.

Another person of interest is Charles Howard Hinton (1853–1907), who used to live, in modern words, a wild life. He developped a system for memorizing an arrangement of $36 \times 36 \times 36 = 46656$ cubes, each of which carrying a unique latin name. So, he created something like a "threedimensional retina" and became able, after years of training, for mental images of spatial arrangements even in a 4-dimensional space. Also Henri Poincaré reports on such a successfull training in his book "Last Essays" [6]. In a more pessimistic mood we find Rudy Rucker ([8], p. 7): "*It is very hard to visualize such a dimension directly. Off and on for some fifteen years, I have tried to do so. In all this time I've enjoyed a grand total of perhaps fifteen minutes' worth of direct vision into four-dimensional space.*"

Interesting enough, Poincaré also considered the question about the three dimensional nature of our world, and he gives a surprisingly modern-sounding idea. According to him, a thinking entity, as the human beings are, could assign to their world either dimension they like, since all of them are mathematically equivalent. Thereby, the world appears to be an abstract premise for sensoric perception and muscular activity, and spatial dimensions are basically a mental product. The choice for three depends on the configuration of the human nervous system and the so-acquired evolutionary advantage. By using two 2-dimensional retinas, the movement of the two hands has to be monitored. A dimension of 2 would not suffice to perform such a control, and a dimension of 4 would allow for movements, during which the hands may even shortly be disconnected from the remaining body. So, a dimension of 3 seems to be a good compromise.

Basing on such arguments, mental operations with more than three dimensions seems to be possible, but dangerous for the being undergoing such experiments in our 3-dimensional *organized* world as well.

The question should not mistaken with the question for a physical fourth dimension. The limitation of feeling unable to imagine a tesseract is related to human cognition, and it hints on an evolutionary designed human being without any need for such abilities. Concepts of modern physics, as quantum theory, spin, quarks, black holes or string theory rely on abstractions beyond human cognition as well, and lead to many contradictions for our thinking, but this is related to the use of models to describe objective matter.

So, considering time as a fourth dimension is not the same as considering a fourth *spatial* dimension. While it is a good trick for describing relativity, it is not just more: we can not move forth and back in time. Also string theory, which offers even ten dimensions, six of which got lost after creation of the universe, has similar disadvantages.

Therefrom, this short consideration of a fourth spatial dimension should show that we could learn something about the "bootstrapping" of human cognition, or, in other words, the way that thinking is rooted into its material base, the central nervous system.

Whether there are any physical principles, which may temporarily cross a fourth dimension, remains an open question. Up to now, there is no experiment known which demonstrates the fact of a merely 3-dimensional world, and also there is no fact known for which a 3-dimensional world would be optimal. On the other hand, the practical possibilities offered by such an access would be astonishing.

Last, but not least, the question for a fourth dimension has inspired arts as well. In [2] there are lots of computer generated pictures on this subject. The 1954 Salvador Dalí painting "Corpus Hypercubicus" shows the famous religious motive using an unfolded tesseract. The Cubists also tried in their multi-perspective approach to capture the idea of a choice for simultaneous perception of different aspects of the same thing - as it would be a simple act for a fourdimensional (thinking) being, knowing about Charles Hintons' *Kata* perspective.

4 Equal distribution in the hypercube

4.1 Square

As it was discussed at the beginning of this text, equally distributed points in a square will not necessarily give equally distributed distances (pairwise distances or distance to one corner of the square). In the following, this will be derived mathematically.

For doing so, the probability density $p(R)$ of the distances R of equally dsitributed random points will be determined. This is given by the derivation of the probability function $P(R)$, which describes the probability of obtaining a distance value lower or

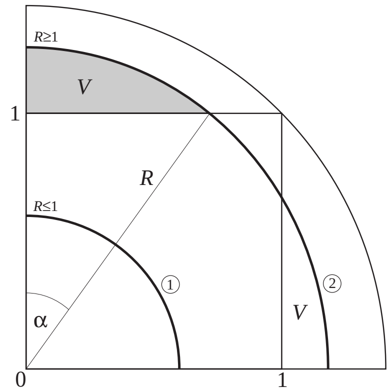


Figure 6: On the derivation of average distance of equally distributed points to a corner of a square.

equal to R, according to R. For $R \leq 0$, this probability is 0, and for $R \geq \sqrt{2}$ it is 1. For $0 \leq R \leq 1$ (case 1), $P(R)$ is the portion of a quarter circle with radius R of the whole square (see figure 6). The distance of all those points of the quarter circle is smaller or equal to R. Case 2 happens, when $1 \leq R \leq \sqrt{2}$. The set of all points with a distance to the origin smaller or equal to R is the intersection of the quarter circle with radius R and the square. Therefore, from the area of the quarter circle, the area of the overlapping circular segments V has to be subtracted two times. The computations for both cases give:

Case 1. The quarter circle has an area of $A = 1/4\pi R^2$, therefore the probability function is $p(R) = \pi/2R$.

Case 2. V is given as the difference of areas of a circular sector with angle α and a rectangular triangle with cathetus 1 and hypotenuse R. Then (see figure 6):

$$\cos\alpha = \frac{1}{R} \tag{15}$$

$$\begin{aligned} V &= \pi R^2 \frac{\arccos 1/R}{2\pi} - \frac{1}{2}\sqrt{R^2-1} \\ &= \frac{1}{2}\left(R^2 \arccos 1/R - \sqrt{R^2-1}\right) \end{aligned} \tag{16}$$

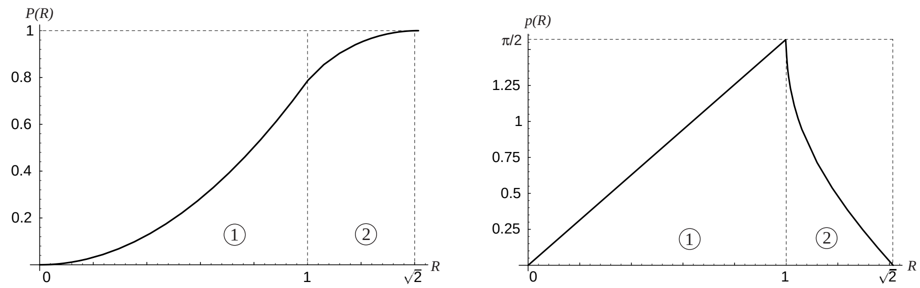
$$\begin{aligned} A &= \frac{1}{4}\pi R^2 - 2V = \frac{1}{4}\pi R^2 - R^2 \arccos 1/R + \sqrt{R^2-1} \\ &= R^2\left(\frac{\pi}{4} - \arccos\frac{1}{R}\right) + \sqrt{R^2-1}. \end{aligned} \tag{17}$$

By deriving $A = A(R)$ according to R, $p(R)$ follows as:

$$p(R) = \frac{R}{2}\left(\pi - 4\arccos\frac{1}{R}\right). \tag{18}$$

Since

$$\int_0^{\sqrt{2}} p(R)dR = 1 \tag{19}$$



(a) (b)

Figure 7: Probability function (a) and probability density (b) of the distances of equally distributed points to a corner within the unit square.

1	2	3	4	5	6	7	8	9	10
0.543	0.543	0.544	0.541	0.541	0.541	0.540	0.542	0.540	0.539

Table 3: Ten experimentally computed values for the average distance of 20,000 random points from a corner.

there is no need to normalize $p(R)$.

Figure 7(a) and (b) shows the plots of $P(R)$ and $p(R)$. As can be seen, there is no uniform distribution of the distances (which would be given by a flat line). To get the expectation value of the distance, the first moment has to be computed:

$$\int_0^{\sqrt{2}} x p(x) dx = \frac{\sqrt{2} + \ln(\sqrt{2}+1)}{3} \simeq 0.7652. \tag{20}$$

So we get for the expectation value of the distance of a randomly selected point within a square with diagonallength 1 to one corner

$$E[R/\sqrt{2}] \simeq 0.541. \tag{21}$$

Table 3 gives some experimentally computed values, using 20,000 random points in each case.

For the variancy one gets:

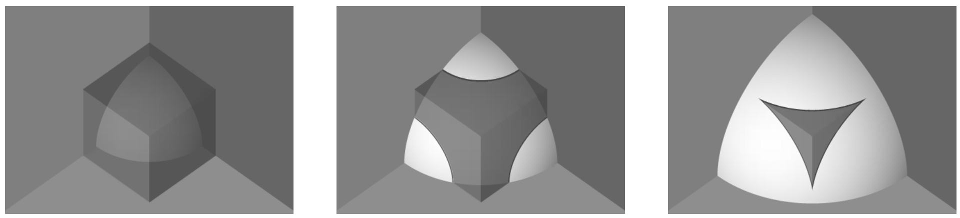
$$\begin{aligned} \sigma^2[R/\sqrt{2}] &= \frac{1}{2}\int_0^{\sqrt{2}} (x - E[R])^2 p(x) dx \\ &= \frac{4 - 2\sqrt{2}\log(1+\sqrt{2}) - \log(1+\sqrt{2})^2}{18} \\ &\simeq 0.0405 \end{aligned} \tag{22}$$

Table 4 gives ten experimentally computed values, using 10,000 random points in each case.

To summarize: equally distributed points in a unit square will have an average distance of 0.76 ± 0.28 to a corner of the square (with 0.28 given as $\sigma[R]$).

1	2	3	4	5	6	7	8	9	10
40.5	40.5	39.9	39.9	40.5	40.7	40.8	39.5	39.9	41.6

Table 4: Ten experimentally computed values for the variancy of 10,000 random points, multiplied with 1000.



$R \leq 1$ $\quad$ $1 \leq R \leq \sqrt{2}$ $\quad$ $\sqrt{2} \leq R \leq \sqrt{3}$

Figure 8: Three cases for the intersection of eighth ball and unit cube.

4.2 Cube

Now, the computations will be given for the case of a unit cube. As can be seen from figure 8, this time there are three cases to consider.
Case 1: $0 \leq R \leq 1$. The eighth ball lies completely within the cube, therefore

$$P(R) = \frac{1}{6}\pi R^3. \tag{23}$$

Case 2: $1 \leq R \leq \sqrt{2}$. Here, the eighth ball sticks out of the cube on its three faces. Each of these quarter ball caps has the height $R-1$. The volume of a ball cap κ with radius R and height h is $V_\kappa = 1/3\pi h^2(3R-h)$. Therefore

$$\begin{aligned} P(R) &= \frac{1}{6}\pi R^3 - \frac{3}{4}\cdot\frac{1}{3}\pi(R-1)^2(3R-R+1) \\ &= \frac{1}{6}\pi R^3 - \frac{1}{4}\pi(R-1)^2(2R+1). \end{aligned} \tag{24}$$

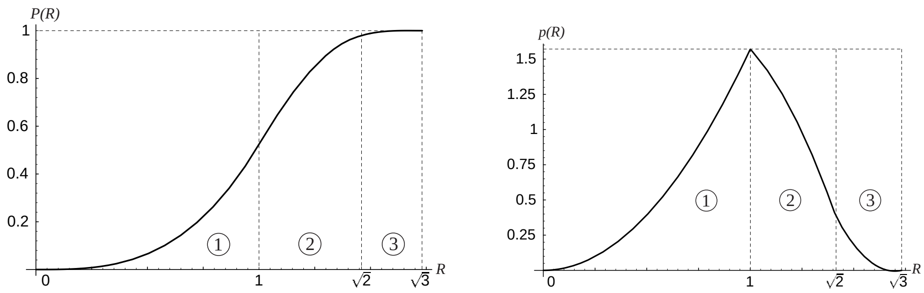
Case 3: $\sqrt{2} \leq R \leq \sqrt{3}$. The quarter ball caps start to overlap. The resulting body (a cube corner sticking out of the eighth ball) is hard to describe mathematically. Because only a small part of the functions $p(R)$ and $P(R)$ is covered by this case, instead of an exact computation a spline interpolation was used.

The details of the derivation of $P(R)$ will be omitted here. Figure 9 gives the plots of both functions. The skewness of the distribution, compared with the square case, has increased.

From the first moment the expectation value $E[R/\sqrt{3}]$ and the variancy $\sigma^2[E/\sqrt{3}]$ can be computed. It follows for the average distance in a cube of diagonallength 1 the value 0.55 and for the variancy 0.026. In the unit cube, equally distributed random points will have an average distance of 0.96 ± 0.28 to each corner.

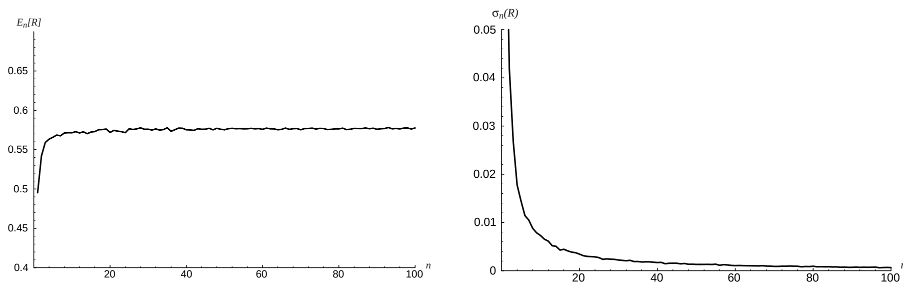
4.3 General case

For the general case $n > 3$, the formal computations become too complex. Thats why we restrict our attention on some numerical experiments.



(a) (b)

Figure 9: Probability function (a) and density (b) for the distances of random points within the unit cube to a corner.



(a) (b)

Figure 10: Expectation value and variancy of the distance of a random point of a hypercube with diagonallength 1 to one of its corners.

The simulation is quite simple: for given dimension n, n equally distributed random numbers x_i from $[0,1]$ are generated. Then, compute

$$R_n = \frac{1}{\sqrt{n}}\sqrt{x_1^2 + x_2^2 + \ldots + x_n^2}. \tag{25}$$

This is repeated for a given number of trials, and the expectation value and variancy of the values R_n is computed. To get the plot in figure 10 (a), for n going from 1 to 100, in each case 2000 trials were made, for figure 10(b) 1000 trials. Two things are obvious: the expectation values approaches a constant value (R_∞ of about 0.58); and the variancy decays to 0 with the order of about $1/\sqrt{n}$.

On a first glance, this may be a surprise. For higher dimensions, the reliability of the fact that a random point will have a fixed distance to a corner of a hypercube, rapidly goes to 1. The probability of obtaining a different distance becomes nearly 0. Therefore, it is impossible to stay nearby the corner region of a cube.

It has to be noted that the points itself remain random. They are still uniformly scattered over the whole volume of the hypercube. The convergence is only a statistical one. This can be understood in the sense that the majority of the hypercube volume is concentrated in the inner part, and the portion of the spikes (corner regions) goes down to 0.

Figure 11 illustrates this fact. By numerical simulation, the probability density functions for dimensions 2, 3, 4, 5 , 10, 30, 50 and 100 have been obtained. With increasing dimension one gets a more and more bell shaped curve.

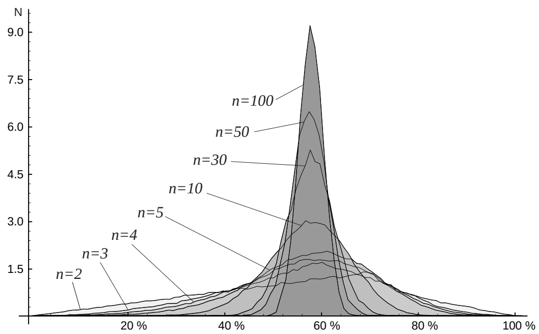


Figure 11: Probability density functions for the distance of a random point from the corner of a unit hypercube for dimensions up to 100.

What is the exact value of R_∞? For getting this value, equation (25) is re-considered under the viewpoint that R_n^2 is the expectation value of n squares of equally distributed random numbers from $[0,1]$ as well. Now, n is considered to be infinitesimal large, but all those random numbers are “rounded” up to the nearest out of a set of m equally-distanced values from $[0,1]$. So, R_∞ will be derived as a function of m with m going to infinity. To illustrate this, we consider the case $m=3$ with the three values 0, $1/2$ and 1. So, 0.321 will be rounded to 0.5, 0.872 to 1 and 0.123 to 0. There are three cases:

1. Random numbers from 0 to 0.25 will be rounded to 0. This happens in $1/4$ of all cases. The square will take the value 0 in 25% of all trials.

2. Random numbers from 0.25 to 0.75 will be rounded to 0.5. Hence, with a probability of $1/2$ the square will have the value 0.25.

3. Random numbers from 0.75 to 1 will be rounded to 1. This happens with probability $1/4$, and the square will have the value 1.

Thus follows for the expectation value of the square after rounding:

$$R_\infty(3) = \frac{1}{4}\cdot 0 + \frac{1}{2}\cdot\frac{1}{4} + \frac{1}{4}\cdot 1 = \frac{3}{8}. \tag{26}$$

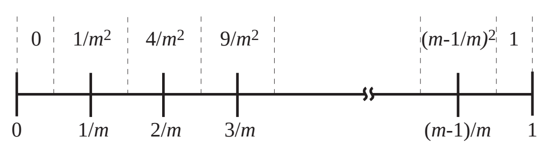


Figure 12: Computing the expectation value of the square of random numbers from $[0,1]$ by rounding them to m intervall values.

Now we consider the general case of $(m+1)$ equally distanced intervall values for rounding (see figure 12). The intervall values are $0, 1/m, 2/m, 3/m, \ldots, (m-1)/m, 1$. There are $(m-1)$ intervalls of length $1/m$, for which the squares take the values $1/m^2, 2/m^2, \ldots, (m-1)^2/m^2$, and two intervalls of length $1/2m$, with the square being

0 and 1. For $R_\infty(m)$, we obtain:

$$\begin{aligned} R_\infty(m) &= \frac{1}{m}\left(\frac{1}{m^2}+\frac{4}{m^2}+\frac{9}{m^2}+\ldots+\frac{(m-1)^2}{m^2}\right)+\frac{1}{2m} \\ &= \frac{1}{m^3}\left(1^2+2^2+3^2+\ldots+(m-1)^2\right)+\frac{1}{2m} \\ &= \frac{1}{m^3}\frac{(m-1)m(2m-1)}{6}+\frac{1}{2m} \\ &= \frac{1}{3}+\frac{1}{6m^2}. \end{aligned} \tag{27}$$

In this derivation, the formula for the sum of the first k square numbers $\sum_{i=1}^{k} i^2 = 1/6 \cdot k(k+1)(2k+1)$ was used. For the limes we get:

$$R_\infty^2 = \lim_{m\to\infty} R_\infty(m) = \frac{1}{3}. \tag{28}$$

Hence, the expectation value goes to $1/\sqrt{3} \simeq 0.57735$, which is quite similar to the estimate of figure 10.

The presented approach is not suitable for getting the decay of variancy as well. But, if the magnitude n of the dimension is considered to be a measure for the exactness of the computation of R_∞, from equation (28) it can be seen that variancy goes down with the order of about $1/\sqrt{n}$.

Finally we will mention the question for the *pairwise* distance of two randomly selected points:

$$D_\infty^2 = \lim_{n\to\infty} E_n[(x_i - x_j)^2]. \tag{29}$$

This can be derived in a manner similar to the derivation of R_∞. One gets for the expectation value of the product of two randomly selected points from $[0,1]$ the value $1/4$. For the average distance it follows:

$$D_\infty^2 = 2E[x_i^2] - 2E[x_i x_j] = \frac{2}{3} - 2\cdot\frac{1}{4} = \frac{1}{6}. \tag{30}$$

thus $D_\infty = 1/\sqrt{6} \simeq 0.40825$. In higher dimensions, we will not meet the "Gould-Effect" (see p. 2), since the pairwise distances of random points approach a constant value. This may give raise to the question whether pattern recognition is even possible for beings using a higherdimensional sensation space. We see more evidence for the fact that with increasing dimension probabilities become reliabilities.

5 Summary

Some geometrical and statistical issues of higherdimensional spaces, i.e. spaces with a dimension greater than 3, have been discussed. Despite of the fact that they are simply accessible from a formal point of view, their geometrical and statistical properties may give some problems. This is only partially related to the limitation of human cognition to perform mental operations on such objects as the 4-dimensional hypercube. The more important are geometric distortions of positional and volume relations, which finally repress randomness at all. In particular, evidence was given for the following claims:

- The ratio of the volume of a hypersphere to the volume of its embedding hypercube goes to 0 by the order of $n!$. This means for search methods that it makes a big difference, whether searchspace is explored spherically or "blockwise."

- Segmentations become unstable according to relative ratios of volumes. By slightly shifting one of the segmentation boundaries and for sufficient large n, the ratios of volumina could be changed to either value.

- The distance of a randomly selected point in a hypercube of diagonallength 1 to one corner gets closer and closer to $1/\sqrt{3}$, as n increases. Hence, for random search it is impossible to select positions nearby the corner of a hypercube.

- Two randomly selected points in a hypercube will have nearly the same distance for larger n. Therefrom, $(n+1)$ of them will give a hypersimplex, the volume of which rapidly goes to 0. An estimation of the volume occupied by a random set of points (basing e.g. the Monte Carlo method) is no more reliable.

- Features in random patterns (e.g. textures) are a unique property of lowdimensional spaces.

Therefore, using analogies from 2- or 3-dimensional spaces for higherdimensional spaces always has to be done with care.

A Hypersimplex and Hypersphere

A.1 The generalized Cavalier principle

A cone over an object (base) to a point A not belonging to this object is defined as the set of all lines connecting A with any of the interior points of the object. The distance of A from the object is defined as the height of the cone. The CAVALIER principle says that two cones have the same volume if their bases and heights are equal.

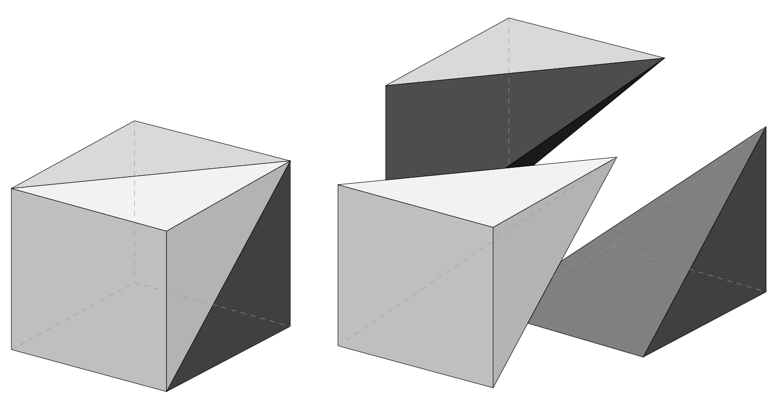


Figure 13: Splitting of a cube into three coni with equal volume. A corner V of the cube is connected to all points of a face, which does not touch the corner point.

Now consider the coni, which are generated by connecting a corner of a cube with all points of a face, which is disjoint with the corner (there are three such faces). Thus, the cube is splitted into three disjoint coni, which have the same volumes according to the CAVALIER principle (in all cases, base is a face of the cube, and height is the sidelength of the cube). Since a cube with sidelength R has a volume of R^3, each of the subconi has a volume of $R^3/3$. With $A = R^2$ being the area of the base of such a conus, and $h = R$ being its height, this can be written as $V_{conus} = hA/3$ as well.

When the base of a conus is not a square, the base is assumed to be covered by a sequence of smaller squares, which approximate the area of the base. For each conus over a subsquare, the equation for the volume holds, and in the limiting case, for the volume of a conus with base A and height h one gets:

$$V = \frac{1}{3}\, h \cdot A. \tag{31}$$

Note that for the validity of equation (31) the base needs not to be flat.

A similar relation can be given for the n-dimensional case as well. Then, the hypercube is split into n disjoint coni, for a corner of a hypercube being disjoint to exactly n hyperfaces (see p. 5). Therefore, a n-cone with base A and height h has a (hyper)volume of

$$V = \frac{1}{n}\, h \cdot A. \tag{32}$$

A.2 Volume of a hypersphere

The volume of a hypersphere could be derived by using multiple integration, but there is a more simple way [1].

A n-ball is the set of all points in n-dimensional space, which have the maximum distance of R from a given central point, i.e. a n-dimensional hypersphere. A $(n-1)$-sphere is the set of all points with its distance to the center being exactly R, i.e. the surface of the n-ball.

A n-ball is the cone over its surface as well. This cone has the base of $A_{n-1}(R)$ and the height R. Thus, due to the generalized CAVALIER principle, for volume $V_n(R)$ of a n-ball and its surface volume $A_{n-1}(R)$, the following relation holds:

$$V_n(R) = \frac{1}{n}\, R \cdot A_{n-1}(R). \tag{33}$$

Also $A_n(R)$ can be given as a function of $V_{n-1}(R)$. For doing so, consider the embedding of the ball into a tophat with the same radius.

This embedding can be seen in 14. There, the area of a stripe on the surface of the ball is equal to the area of its projection onto the surface of the tophat, which embeds the ball.

Figure 15 shows the geometric relations for the 3-dimensional case. The (infinitesimal small) ball stripe has radius r_s and width w_s. So, the area is $2\pi r_s w_s$. The projected tophat stripe has a width of w_c and radius r_c, which is equal to the radius of the ball. Its area is $2\pi r_c w_c$. Since the two triangles marked in gray are similar, $r_s/r_c = w_c/w_s$ holds. But then

$$2\pi r_s w_s = 2\pi r_c w_c. \tag{34}$$

[1] This follows an idea of Evelyn Sander and Bob Hesse, see http://freeabel.geom.umn.edu/docs/forum/ndvolumes/.

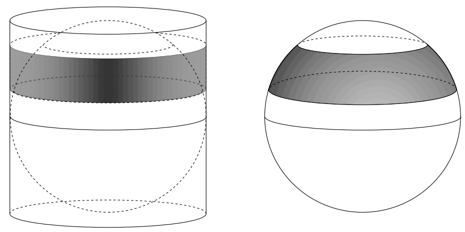


Figure 14: A stripe on the surface of the ball has the same area as its projection onto the surface of a tophat, which embeds the ball.

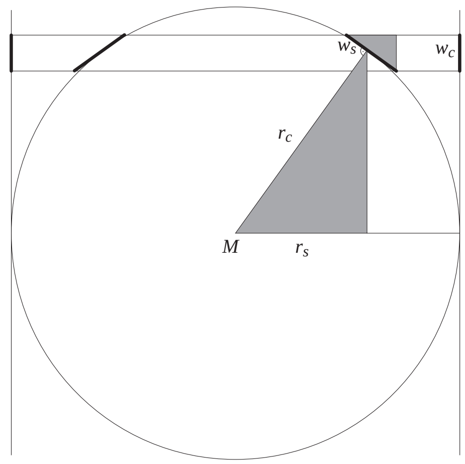


Figure 15: Similar triangles for the proof that ball and tophat stripes do have the same area. It is $r_s/r_c = w_c/w_s$.

This is also true in the general case. The $(n-1)$-sphere has the same volume as the crossproduct of circle and $(n-2)$-ball.

$$A_{n-1}(R) = 2\pi R \cdot V_{n-2}(R) \tag{35}$$

Together with equation (33) this gives a recursive formula for $V_n(R)$:

$$V_n(R) = \frac{2\pi}{n} R^2 V_{n-2}(R). \tag{36}$$

Using

$$V_1(R) = 2R \tag{37}$$
$$V_2(R) = \pi R^2 \tag{38}$$

one gets for even and odd values of n:

$$V_{2p}(R) = \frac{\pi^p \cdot R^{2p}}{p!} \tag{39}$$
$$V_{2p+1}(R) = \frac{2^{p+1} \cdot \pi^p \cdot R^{2p+1}}{1 \cdot 3 \cdot \ldots \cdot (2p+1)} \tag{40}$$

A.3 Hypersimplex

For obtaining the volume V_n of a n-dimensional hypersimplex, the generalized CAVALIER principle can be applied directly. Figure 16, which shows the 3-dimensional case, should stand for the general case. The $(n+1)$th corner A of a n-dimensional hypersimplex with borderlength a is placed at the height h_n above the base hypersimplex, which is composed of the remaining n points. Therefore, the volume is

$$V_n = \frac{1}{n}\, h_n V_{n-1}. \tag{41}$$

Since $V_1 = a$ is known, one basically has to get a formula for h_n in order to obtain an expression for V_n. Two relations will be used, which can be verified by using the two triangles $\triangle_1$ and $\triangle_2$ in figure 16. Be r_n the radius of the hypersphere surrounding the hypersimplex. Triangle $\triangle_1 = \triangle_{AMN}$ is generated from the following three points: corner A; the perpendicular M from A onto the base hypersimplex (of dimension $(n-1)$); and the center N of a hypersimplex of dimension $(n-2)$, which is disjoint to A (in the figure the line $\overline{BD}$). Triangle $\triangle_2 = \triangle_{AMB}$ is composed from the same A and M and a further corner B, which also borders the hypersimplex containing N.

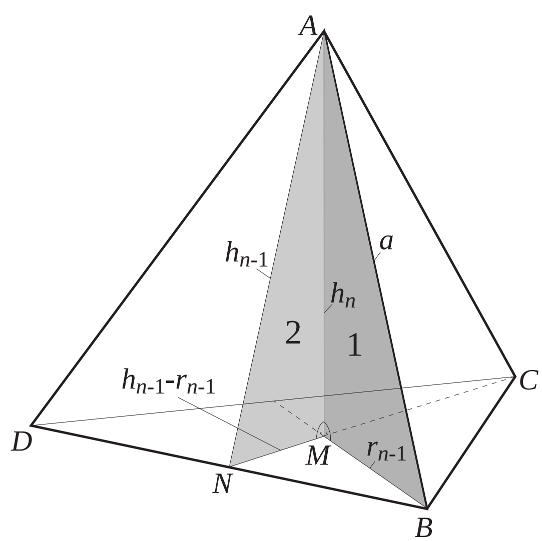


Figure 16: Derivations for a hypersimplex.

From $\triangle_1$ one gets

$$\triangle_1 : \; h_n^2 + r_{n-1}^2 = a^2 \tag{42}$$

and from $\triangle_2$ respectively

$$\triangle_2 : \; h_{n-1}^2 = h_{n^2} + (h_{n-1} - r_{n-1})^2. \tag{43}$$

This can be combined:

$$a^2 - r_{n-1}^2 = h_{n-1}^2 - (h_{n-1} - r_{n-1})^2 \tag{44}$$

$$a^2 = 2h_{n-1}r_{n-1} \tag{45}$$

and therefore, by replacing $n-1$ by n:

$$a^2 = 2h_n r_n. \tag{46}$$

This means: if h_{n-1} and r_{n-1} are known, h_n and r_n can be computed by

$$\begin{aligned} h_n^2 &= a^2 - r_{n-1}^2 && (47) \\ r_n &= \frac{a^2}{2h_n}. && (48) \end{aligned}$$

For the regular triangle the values of h_2 and r_2 are known. It is $h_2 = \sqrt{3}/2\,a$ and $r_2 = a/\sqrt{3}$. So, the values of h_n and r_n can be determined recursively.

$$\begin{aligned} h_n &= \sqrt{\frac{n+1}{2n}}\,a && (49) \\ r_n &= \sqrt{\frac{n}{2(n+1)}}\,a. && (50) \end{aligned}$$

For the volume of the hypersimplex follows:

$$V_n = \frac{1}{n}\sqrt{\frac{n+1}{2n}}\,a V_{n-1} \tag{51}$$

which gives

$$V_n = \frac{1}{n!}\sqrt{\frac{n+1}{2^n}}\,a^n. \tag{52}$$

For $n = 1$ this holds by $V_1 = a$. Be equation (52) true for $n = k$. Then, for $n = k+1$ it follows:

$$\begin{aligned} V_{k+1} &= \frac{1}{(k+1)!}\sqrt{\frac{k+2}{2^{k+1}}}\,a^{k+1} = \frac{1}{k+1}\cdot\sqrt{\frac{k+2}{2(k+1)}}\,a\cdot\frac{1}{k!}\sqrt{\frac{k+1}{2^k}}\,a^k \\ &= \frac{1}{k+1}\,h_{k+1}V_k && (53) \end{aligned}$$

and the equation holds for $n = k+1$ as well. This verifies equation (52).
For the inner angle θ of a triangle with the sides r_n, r_n and a and by using the cosine theorem one gets

$$a^2 = 2r_n^2(1-\cos\theta), \tag{54}$$

and, if using equation (50)

$$\theta = \arccos\left(-\frac{1}{n}\right). \tag{55}$$

References

[1] E. A. ABBOTT, *Flatland*, Penguin Books, 1987.

[2] T. F. BANCHOFF, *Beyond the Third Dimension*, Scientific American Library, 1996.

[3] S. J. GOULD, *Bully for Brontosaurus*, Penguin Books, 1991.

[4] R. HECHT-NIELSEN, *Neurocomputing*, Addison–Wesley Publishing Company, Reading, MA, u.a., 1991.

[5] D. MARR, *Vision*, MIT Press, 1981.

[6] H. POINCARÉ, *Mathematics and Science: Last Essays*, Dover Publications, Inc., New York, 1963.

[7] I. RECHENBERG, *Evolutionsstrategie'94*, frommann–holzboog, 1994.

[8] R. RUCKER, *The Fourth Dimension*, Houghton Mifflin Company, Boston, 1984.