PROBABILITIES AND THE SHAPE OF THE UNIT BALL IN $\mathbb{R}^n$

ADI BEN-ISRAEL

Abstract. For a random variable $\mathbf{x}$ uniformly distributed on the unit sphere $\mathbf{S}_n$ in $\mathbb{R}^n$, or in the unit ball $\mathbf{B}_n$, we compute the probabilities $\text{Prob}\{\langle \mathbf{u}, \mathbf{x} \rangle \le \alpha\}$ for $-1 \le \alpha \le 1$ and some $\mathbf{u} \in \mathbf{S}_n$ (the results are independent of $\mathbf{u}$.) As $n \to \infty$ these probabilities converge to $\text{Prob}\{z \le \alpha \sqrt{n}\}$, where z has the standard normal distribution $N(0,1)$.

Consequently most of the area of $\mathbf{S}_n$ [or the volume of $\mathbf{B}_n$] is in "narrow" belts [or slabs] around the "equator". For example, 99.7% of the sphere area [or the ball volume] is in such a belt [or slab] of width $6/\sqrt{n}$. This result is counter–intuitive, at least contrary to intuition in 3 dimensions.

1. Introduction

Our geometric intuition was acquired in 2–dimensional and 3–dimensional spaces, and is not a reliable guide in spaces of higher dimension.

A case in point is $\mathbb{R}^n$ (with Euclidean norm), its unit ball $\mathbf{B}_n$, and unit sphere $\mathbf{S}_n$. For large n, a narrow belt [or slab] around an equator accounts for most of the area of $\mathbf{S}_n$ [or volume of $\mathbf{B}_n$.] For example, 68.2% [99.7%] of the area of $\mathbf{S}_n$ is in an equatorial belt of width $2/\sqrt{n}$ [resp. $6/\sqrt{n}$]. In particular, the sum of the areas of any two equatorial belts of width $2/\sqrt{n}$ gives 136.4% of the area of the sphere. The overlap of the two belts accounts for the "surplus".

This result is weird since there is no preferred direction in $\mathbf{S}_n$, and any great circle is an "equator".

We have thus to abandon the mental image of the unit ball as a "round" object. Perhaps the word to describe it is "oblate" (flattened at the "poles"), but then it is oblate from every direction one "looks" at it.

A probabilistic model is used to explain these results. The computation of the areas of equatorial belts is given in Section 3, and the results are confirmed for large n using the Central Limit Theorem in Section 4. The corresponding results for the volumes of slabs of the unit ball are given in Section 5.

2. A probability problem

Let $\|\cdot\|$ denote the Euclidean norm, and let $\mathbf{B}_n(r) := \{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x}\| \le r\}$ and $\mathbf{S}_n(r) := \{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x}\| = r\}$ be, respectively, the **ball** and **sphere** of radius $r > 0$. The unit ball $\mathbf{B}_n(1)$ and unit sphere $\mathbf{S}_n(1)$ are denoted by $\mathbf{B}_n$ and $\mathbf{S}_n$, respectively.

Let $v_n(r)$ and v_n denote the volumes of $\mathbf{B}_n(r)$ and $\mathbf{B}_n$, and $a_n(r)$ and a_n the areas of $\mathbf{S}_n(r)$ and $\mathbf{S}_n$, respectively. Then, for $n = 2, 3, \cdots$,

$$a_n = \frac{2\pi^{n/2}}{\Gamma(n/2)} \ , \ a_n(r) = a_n \, r^{n-1} \ , \tag{1}$$

$$v_n = \frac{a_n}{n} \ , \ v_n(r) = v_n \, r^n \ , \tag{2}$$

$$dv_n(r) = v_n'(r) \, dr = a_n(r) \, dr \ , \tag{3}$$

where $\Gamma(\cdot)$ is the Gamma function, see, e.g., [1].

Let $\mathbf{x}$ be a random vector uniformly distributed on $\mathbf{S}_n$, and let $\mathbf{u} = (u_1, \cdots, u_n) \in \mathbf{S}_n$ be fixed. For $-1 \le \alpha \le 1$, consider the probability[1]

$$\text{Prob}\,\{\langle \mathbf{u}, \mathbf{x} \rangle \le \alpha\} \ , \tag{4}$$

Date: June 26, 2011.

[1]Thanks to Professor A. Ben–Tal for suggesting this problem.

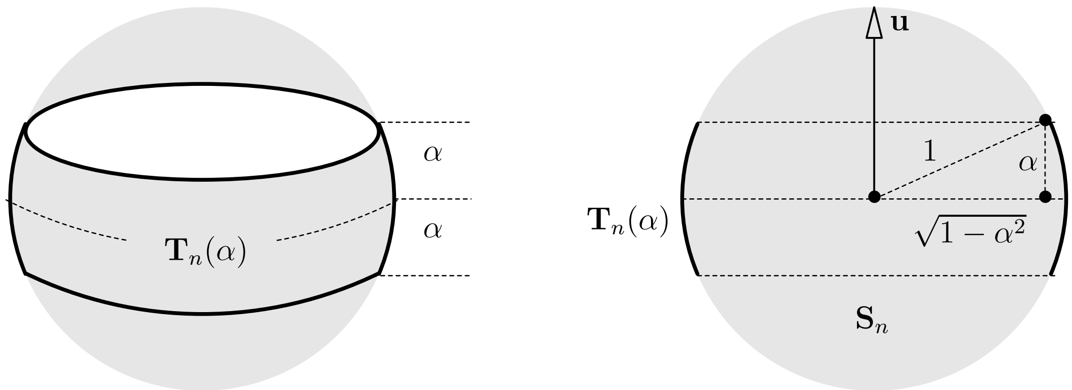

FIGURE 1. The equatorial belt $\mathbf{T}_n(\alpha)$

which, by symmetry, is independent of $\mathbf{u}$. For convenience, consider a symmetric version, for $0 \le \alpha \le 1$,

$$p_n(\alpha) := \text{Prob}\,\{|\langle\, \mathbf{u}, \mathbf{x}\,\rangle| \le \alpha\}\ . \tag{5}$$

The probabilities (4) then follow from

$$\text{Prob}\,\{\langle\, \mathbf{u}, \mathbf{x}\,\rangle \le \alpha\} = \tfrac{1}{2}\begin{cases} 1 - p_n(|\alpha|), & \alpha \le 0; \\ 1 + p_n(\alpha), & \alpha > 0, \end{cases} \tag{6}$$

where $p_n(0) = 0$, and $p_n(1) = 1$.

$$\tag{7}$$

Fixing $\mathbf{u} = (0, 0, \cdots, 0, 1)$, the vectors $\mathbf{x} = (x_1, x_2, \cdots, x_n)$ satisfying

$$|\langle\, \mathbf{u}, \mathbf{x}\,\rangle| \le \alpha \tag{8}$$

belong to the set

$$\mathbf{T}_n(\alpha) = \{\mathbf{x} \in \mathbf{S}_n : -\alpha \le x_n \le \alpha\}\ . \tag{9}$$

We can visualize $\mathbf{u}$ as the "north pole" of $\mathbf{S}_n$, and $\mathbf{T}_n(\alpha)$ as a belt around the "equator", of height 2α, see Fig. 1, left side. Let $t_n(\alpha)$ be the area of $\mathbf{T}_n(\alpha)$. The probability $p_n(\alpha)$ is therefore the ratio,

$$p_n(\alpha) = \frac{t_n(\alpha)}{a_n}\ . \tag{10}$$

3. The area computation

The projection of $\mathbf{T}_n(\alpha)$ on the "horizontal" hyperplane $u_n = 0$ is the ring

$$(1 - \alpha^2)^{1/2} \le r \le 1\ ,$$

n	$k=1$	$k=2$	$k=3$
2	.5		
3	.5773502693		
4	.6089977810	1	
5	.6260990336	.9838699099	
10	.6565636037	.9632125020	.9999914613
100	.6802515257	.9550652747	.9976960345
500	.6822048238	.9546087218	.9973798784
1000	.6824473395	.9545539777	.9973400661

TABLE 1. Some probabilities $p_n(\alpha)$ for $\alpha = k/\sqrt{n}$

where the lower radius is computed by Pythagoras' Theorem, see Fig. 1, right side. As in [1, Appendix C] we use spherical shells of radius r and volume $dv_{n-1}(r)$ to compute

$$\begin{aligned} t_n(\alpha) &= 2 \int_{(1-\alpha^2)^{1/2}}^{1} \frac{dv_{n-1}(r)}{\sqrt{1-r^2}} \\ &= 2\, a_{n-1} \int_{(1-\alpha^2)^{1/2}}^{1} \frac{r^{n-2}}{\sqrt{1-r^2}}\, dr \ , \ \text{by (3)} \ , \\ &= a_{n-1} \int_{1-\alpha^2}^{1} x^{(n-3)/2}\, (1-x)^{-1/2}\, dx \ , \ \text{for } x = r^2 \ . \end{aligned}$$

$$\begin{aligned} \therefore\ p_n(\alpha) = \frac{t_n(\alpha)}{a_n} &= \frac{a_{n-1}}{a_n} \int_{1-\alpha^2}^{1} x^{(n-3)/2}\, (1-x)^{-1/2}\, dx \ , \\ &= \frac{\Gamma(\frac{n}{2})}{\Gamma(\frac{n-1}{2})\, \Gamma(\frac{1}{2})} \int_{1-\alpha^2}^{1} x^{(n-3)/2}\, (1-x)^{-1/2}\, dx \ , \end{aligned} \tag{11}$$

see, e.g., [1, Equation (C.3)]. For the special cases $\alpha = 0$ and $\alpha = 1$, (11) gives the correct results $p_n(0) = 0$ and $p_n(1) = 1$.

Table 1 lists values of (11) for some n, and three values of $\alpha = \frac{k}{\sqrt{n}}$, $k = 1, 2, 3$. In particular, the last column shows that for large n, 99.7% of the area of $\mathbf{S}_n$ is in an equatorial belt of width $6/\sqrt{n}$.

4. A PROBABILISTIC ARGUMENT

The limit of the probability $p_n(\alpha)$ as $n \to \infty$ can be computed from the Central Limit Theorem (see, e.g., [2, Chapter X]).

To this end, consider the random (vector) variable $\mathbf{y} = (y_1, y_2, \cdots, y_n)$ in $\mathbf{S}_n$, where the components y_i are independent and identically distributed, assuming the values $\pm 1/\sqrt{n}$ with probabilities 0.5.

The vector $\mathbf{y}$ assumes the 2^n vector values $\frac{1}{\sqrt{n}}(\pm 1, \pm 1, \cdots, \pm 1)$ with equal probabilities $1/2^n$, and is "uniformly distributed" in $\mathbf{S}_n$ in the sense that its values are equally spaced: any two adjacent values (differing by one sign change) have between them the same distance $2/\sqrt{n}$. The random variable $\mathbf{y}$ can thus be considered a discretization of the continuous random variable $\mathbf{x}$ of § 2.

Let the vector $\mathbf{u} = \frac{1}{\sqrt{n}}(1, 1, \cdots, 1)$ be fixed. As $n \to \infty$, the random variable $\langle \mathbf{u}, \mathbf{y} \rangle$ tends, by the Central Limit Theorem, to be normally distributed with mean 0 and standard deviation $1/\sqrt{n}$. The probability that $|\langle \mathbf{u}, \mathbf{y} \rangle| \le k/\sqrt{n}$ can thus be approximated by

$$F(k) = \frac{1}{\sqrt{2\,\pi}} \int_{-k}^{k} e^{-x^2/2}\, dx \ . \tag{12}$$

For large n, the areas of equatorial belts $\mathbf{T}_n(k/\sqrt{n})$ are thus proportional to the symmetric areas (12) under the standard normal distribution.

The following values of $F(k)$ show good agreement with the limiting values of Table 1.

k	1	2	3
$F(k)$	.6826894920	.9544997360	.9973002039

5. Volumes of slabs

Consider a central ("horizontal") slab of the unit ball defined by

$$\mathbf{U}_n(\alpha) := \{\mathbf{x} \in \mathbf{B}_n : -\alpha \leq x_n \leq \alpha\} \tag{13}$$

see Fig. 1, and let $u_n(\alpha)$ denote its volume. Then,

$$u_n(\alpha) = 2\, v_{n-1} \int_0^{\alpha} \sqrt{1-x^2}^{\;n-1}\, dx\ .$$

$$\therefore\ \frac{u_n(\alpha)}{v_n} = 2\,\frac{v_{n-1}}{v_n} \int_0^{\alpha} \left(1-x^2\right)^{\frac{n-1}{2}}\, dx$$

$$= 2\,\frac{\Gamma\left(\frac{n}{2}+1\right)}{\Gamma\left(\frac{1}{2}\right)\,\Gamma\left(\frac{n+1}{2}\right)} \int_0^{\alpha} \left(1-x^2\right)^{\frac{n-1}{2}}\, dx\ . \tag{14}$$

Table 2 lists values of (14) for some n, and three values of $\alpha = \frac{k}{\sqrt{n}}$, $k = 1, 2, 3$. These values also tend to the corresponding values (12) in the standard normal distribution. The reason here is that for $x = k/\sqrt{n}$, the integrand in (14) converges, as $n \to \infty$, to the integrand of (12),

$$\lim_{x=k/\sqrt{n},\, n\to\infty} \left(1-x^2\right)^{\frac{n-1}{2}} = \lim_{n\to\infty} \left(1-\frac{k^2}{n}\right)^{\frac{n-1}{2}} = e^{-k^2/2}\ .$$

6. Summary

Let $\mathbf{x}$ be a random vector variable $\mathbf{x}$ uniformly distributed on the unit sphere $\mathbf{S}_n$, or in the unit ball $\mathbf{B}_n$. Let $\mathbf{u}$ be a fixed vector in $\mathbf{S}_n$, and $-1 \leq \alpha \leq 1$. Then for large n, the probability

$$\text{Prob}\{\langle \mathbf{u}, \mathbf{x}\rangle \leq \alpha\} \approx \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\alpha\sqrt{n}} e^{-x^2/2}\, dx\ .$$

n	$k=1$	$k=2$	$k=3$
2	.8183098858		
3	.7698003590		
4	.7468300050	.9999999999	
5	.7334302965	.9972863177	
10	.7074815444	.9796369263	.9999992221
100	.6851152384	.9571859870	.9979213001
500	.6831736751	.9550391285	.9974314161
1000	.6829315232	.9547695624	.9973662462

Table 2. The ratios of the volumes $u_n(\alpha)/v_n$ for $\alpha = k/\sqrt{n}$

References

[1] A. Ben–Israel, *An application of the matrix volume in probability*, Lin. Algeb. and Appl. **321**(2001), 9-25.
[2] W. Feller, *An Introduction to Probability Theory and Its Applications* (2nd edition), Vol. 1, J. Wiley, 1957.

Adi Ben-Israel, RUTCOR–Rutgers Center for Operations Research, Rutgers University, 640 Bartholomew Rd., Piscataway, NJ 08854-8003, USA
E-mail address: adi.benisrael@gmail.com