Uniform Law on the Unit Sphere of a Banach Space

by Bernard Beauzamy
Société de Calcul Mathématique SA
111 Faubourg Saint Honoré
75008 Paris – France

September 2008

Abstract

We investigate the construction of a uniform probability law on the unit sphere of a finite-dimensional, real, rearrangement-invariant Banach space. We show that this construction is possible under a geometrical condition, and it comes from a probability on $\mathbb{R}$: there is a density f , on $\mathbb{R}$, such that if $X_1,...,X_K$ are independent variables following this law, the law of $\frac{(X_1,...,X_K)}{\|(X_1,...,X_K)\|}$ is the uniform law on the unit sphere of $\mathbb{R}^K$, equipped with this norm.

I. Introduction

For practical purposes, one may want to generate a sample of points on the unit sphere of a Banach space. Here are two examples :

– Choosing a direction at random is useful for the search of solutions of PDE's : one starts at a given point and then moves to some direction. In this case, the norm is usually the Euclidean norm. See the acknowledgements at the end of the paper for our original motivation, in the framework of two contracts with IRSN (France).

– Choosing a proportion at random is useful, for instance in economics : a budget may depend on some goods, and one wants to study the variations of the budget, depending on various weights on the goods. In this case, the norm is usually the sum of absolute values of the coefficients, that is the l_1 norm.

In all cases, one deals with a finite-dimensional real Banach space, that is the space $\mathbb{R}^K$ equipped with some norm.

In general, one does not have any predefined wish or preference about the points which must be chosen : they need to be on the unit sphere, but no direction is privileged. This means that the law we want for our sample is the uniform law on the unit sphere.

Our final observation, about the motives, is that usually one wants to use a law on the real numbers (building a law directly on the unit sphere of a Banach space is very complicated) : one wants some density function f, on the real numbers, such that if we draw a sample $X_1,...,X_K$ using this law, and then normalize it correctly, that is replace it by $\frac{(X_1,...,X_K)}{\|(X_1,...,X_K)\|}$, then the corresponding points on the unit sphere will follow a uniform law.

II. A preliminary negative remark

First, we observe that such a law cannot be obtained from a uniform law on each component separately.

Indeed, take the Euclidean norm. First draw a sample for a random vector $(x_1,...,x_K)$; each x_k following a uniform law, for instance on the interval $[0,1]$, and then normalize, that is replace each x_k by :

$$y_k = \frac{x_k}{\sqrt{\sum_{i=1}^{K} x_i^2}}, \text{ so that } \sum_{k=1}^{K} y_k^2 = 1.$$

The vector $(y_1,...,y_K)$ is indeed on the unit sphere of l_2, but does not follow a uniform law. Let us see this on a very simple example, in dimension 2.

We take two random numbers x_1 and x_2 with uniform law between 0 and 1. This means that we have a uniform law in the unit square $[0,1]^2$. For each x_1 and x_2, we consider the normalized vector :

$$y_1 = \frac{x_1}{\sqrt{x_1^2+x_2^2}}, \; y_2 = \frac{x_2}{\sqrt{x_1^2+x_2^2}}$$

The law of this vector has density :

If $0 \le \theta \le \frac{\pi}{4}$, $f(\theta) = \frac{1}{2\cos^2(\theta)}$

If $\frac{\pi}{4} \le \theta \le \frac{\pi}{2}$, $f(\theta) = \frac{1}{2\sin^2(\theta)}$

This is quite different from a uniform law.

This remark is quite general : a uniform law in the square does not project to a uniform law on the unit sphere, no matter what the norm is.

III. Uniform law on a sphere

What we mean here by "uniform law" means "uniform with respect to Lebesgue measure". It means that the probability to fall in some set depends only on the Lebesgue measure of this set, and not on the position on the set on the unit sphere.

Since Lebesgue measure is invariant under permutation of the coordinates, we investigate only norms which have the same property. More precisely, we assume that our Banach space is equipped with a norm $\| . \|$ which satisfies :

$$\left\|\left(x_{\sigma(1)},\dots,x_{\sigma(K)}\right)\right\| = \left\|\left(x_1,\dots,x_K\right)\right\|,$$

for any permutation σ of the set $\{1,\dots,K\}$. This is also necessary since we want our law to derive from a law on the real numbers : we draw a sample and then normalize. This implies that all coordinates have the same law.

Let $h(x_1,\dots,x_K)$ be the density of probability we want to build. Then, the fact that we have a uniform law in a rearrangement-invariant space means that this function h will be constant on its domain of definition : $h(x_1,\dots,x_K)$ does not depend on $(x_1,\dots,x_K)$, provided $\|(x_1,\dots,x_K)\| = 1$. This is the fundamental property we will use.

We now show how to build a uniform distribution on the unit sphere, starting from a probability density on the real line. We observe that it is enough to construct our probability density for positive variables. Indeed, there are 2^K sets of equal measure, depending on the sign of each x_k, and the construction of the density in one of them transfers to all the others.

The general construction we give is inspired (with corrections) by the one given in the book by Luc Devroye [2], chapter 5, theorem 2.2, for the l_1 norm.

Let :

$$C_K = \left\{\left(x_1,\dots,x_K\right);\, x_k \geq 0\ \forall k, \left\|\left(x_1,\dots,x_K\right)\right\| = 1\right\}$$

be the positive part of the unit sphere.

We need to introduce a restriction for the norm, which is of geometrical nature. In order to explicit this condition, we investigate the condition :

$$\left\|\left(x_1,\dots,x_{K-1},x_K\right)\right\| \leq s \qquad (1)$$

for given $s\,; x_1,\dots,x_{K-1}$, that is as a condition on x_K alone.

First, we observe that (1) may never be satisfied, if s is too small or $x_1,\dots,x_{K-1}$ too large.

Let :

$$A_{K-1}(s)=\left\{\left(x_1,\ldots,x_{K-1}\right);\exists x_K\ ,\left\|\left(x_1,\ldots,x_K\right)\right\|\le s\right\}$$

be the set on which (1) can be satisfied.

If there is a point x_K for which (1) holds, then there is a unique x_K with

$$\left\|\left(x_1,\ldots,x_{K-1},x_K\right)\right\|=s$$

We denote this point by $\varphi\left(s\,;x_1,\ldots,x_{K-1}\right)$. In other words,

$$\varphi\left(s\,;x_1,\ldots,x_{K-1}\right)=\max\left\{x_K\ ;\left\|\left(x_1,\ldots,x_{K-1},x_K\right)\right\|\le s\right\}$$

and we have

$$x_K\le\varphi\left(s\,;x_1,\ldots,x_{K-1}\right)$$

if $x_K\ge 0$ satisfies :

$$\left\|\left(x_1,\ldots,x_{K-1},x_K\right)\right\|\le s\,.$$

For instance, for the Euclidean norm,

$$\varphi\left(s\,;x_1,\ldots,x_{K-1}\right)=\left(s^2-x_1^2-\ldots-x_{K-1}^2\right)^{1/2}$$

and

$$A_{K-1}(s)=\left\{\left(x_1,\ldots,x_{K-1}\right);x_1^2+\cdots+x_{K-1}^2\le s^2\right\}$$

is the ball of radius s in the Euclidean $K-1$ dimensional space.

We can now state our result :

Theorem. - *Let $\|\ \|$ be a norm on $\mathbb{R}^K$, invariant under permutation of the variables. Then, there is a probability density on $\mathbb{R}$ such that, if $X_1,\ldots,X_K$ are independent variables following this law, the law of $\dfrac{\left(X_1,\ldots,X_K\right)}{\left\|\left(X_1,\ldots,X_K\right)\right\|}$ is the uniform law on the unit sphere of $\mathbb{R}^K$, equipped with this norm, if and only if the values of $\dfrac{\partial\varphi}{\partial x_1}\left(s\,;z\right)$ depend only on the values of x_1 and of $\varphi\left(s\,;z\right)$. In other words, the answer is positive if and only if there is a function G such that :*

$$\frac{\partial\varphi}{\partial x_1}\left(s\,;z\right)=G\left(x_1,\varphi\left(s\,;z\right)\right)\tag{2}$$

Proof of the Theorem.

We set :

$$S = \left\| \left(X_1, ..., X_K \right) \right\| .$$

This is a positive random variable. We also define :

$$Z = \left(X_1, ..., X_{K-1} \right), \; z = \left(x_1, ..., x_{K-1} \right).$$

Let $f_{Z,S}(z,s)$ be the density of the joint law of the $K-$uple $\left(Z,S\right)$. This density can be computed as a conditional probability, knowing Z , that is :

$$f_{Z,S}(z,s) = f_{S|Z}(s) \times f_Z(z) \qquad (3)$$

where $f_{S|Z}(s)$ denotes the conditional probability density of S knowing Z and $f_Z(z)$ is the density of Z .

The proof of the Theorem will be divided into three steps.

Step 1

Under the above assumption, the density $f_{Z,S}(z,s)$ *is independent of* $z = (x_1, ..., x_{K-1})$.

So we are looking for an identity of the following form :

$$f_{S|Z}(s) \times f_Z(z) = c(s) \qquad (4)$$

where the right-hand side is independent of $z = (x_1, ..., x_{K-1})$.

We have, since the variables $X_1, ..., X_{K-1}$ are independent :

$$f_Z(z) = f(x_1) \times \cdots \times f(x_{K-1})$$

and

$$F_{S|Z}(s) = P\left\{ \left\| \left(X_1, ..., X_K \right) \right\| \leq s \mid X_1 = x_1, ..., X_{K-1} = x_{k-1} \right\}$$

that is :

$$F_{S|Z}(s) = P\left\{ \left\| \left(x_1, ..., x_{K-1}, X_K \right) \right\| \leq s \right\}$$

This probability is entirely determined by the law on X_K , since the norm is known and the values of $x_1, ..., x_{K-1}$ are fixed.

The condition :

$$\left\|\left(x_1,\ldots,x_{K-1},X_K\right)\right\|\le s$$

is equivalent to the fact that X_K is in the interval :

$$0\le X_K\le\varphi\left(s\,;x_1,\ldots,x_{K-1}\right),$$

by definition of the function φ, provided that $\left(x_1,\ldots,x_{K-1}\right)\in A_{K-1}(s)$.

For instance, for the Euclidean norm, the condition

$$\left\|\left(x_1,\ldots,x_{K-1},X_K\right)\right\|\le s$$

is equivalent to :

$$0\le X_K\le\left(s^2-x_1^2-\ldots-x_{K-1}^2\right)^{1/2}$$

provided that :

$$x_1^2+\cdots+x_{K-1}^2\le s^2.$$

Let f be the density of probability we are looking for. The identity (4) is equivalent to :

$$f\left(\varphi\left(s\,;x_1,\ldots,x_{K-1}\right)\right)\times f(x_1)\times\cdots\times f(x_{K-1})=c(s) \qquad (5)$$

is independent of $(x_1,\ldots,x_{K-1})$, provided that $\left(x_1,\ldots,x_{K-1}\right)\in A_{K-1}(s)$.

Since all variables are equivalent, it is enough to ensure this for x_1. Computing the derivative with respect to this variable, we get :

$$\begin{aligned}f'\left(\varphi\left(s\,;x_1,\ldots,x_{K-1}\right)\right)\frac{\partial\varphi}{\partial x_1}\left(s\,;x_1,\ldots,x_{K-1}\right)\times f(x_1)\times\cdots\times f(x_{K-1})+\\+f\left(\varphi\left(s\,;x_1,\ldots,x_{K-1}\right)\right)\times f'(x_1)\times f(x_2)\times\cdots\times f(x_{K-1})=0\end{aligned}$$

that is,

$$f'\left(\varphi\left(s\,;x_1,\ldots,x_{K-1}\right)\right)\frac{\partial\varphi}{\partial x_1}\left(s\,;x_1,\ldots,x_{K-1}\right)\times f(x_1)=-f\left(\varphi\left(s\,;x_1,\ldots,x_{K-1}\right)\right)\times f'(x_1) \qquad (6)$$

This can be written :

$$\frac{f'\left(\varphi\left(s\,;z\right)\right)}{f\left(\varphi\left(s\,;z\right)\right)}=-\frac{1}{\dfrac{\partial\varphi}{\partial x_1}\left(s\,;z\right)}\frac{f'(x_1)}{f(x_1)} \qquad (7)$$

We are looking for a density function f satisfying equation (7). Set :

$$g(x) = \frac{f'(x)}{f(x)}$$

This is the logarithmic derivative of the function f. From the identity (7), we deduce :

$$g\big(\varphi(s\,;z)\big) = -\frac{1}{\dfrac{\partial\varphi}{\partial x_1}(s\,;z)}\, g(x_1) \tag{8}$$

In order to understand what this equation means, let us come back for a while to the case of the Euclidean norm. In this case, we have :

$$\frac{\partial\varphi}{\partial x_1} = -\frac{x_1}{\sqrt{s^2 - x_1^2 - \ldots - x_{K-1}^2}}$$

and the identity (8) becomes :

$$g\left(\left(s^2 - x_1^2 - \ldots - x_{K-1}^2\right)^{1/2}\right) = \frac{\sqrt{s^2 - x_1^2 - \ldots - x_{K-1}^2}}{x_1}\, g(x_1) \tag{9}$$

Taking $x_1 = 1$, $x_2 = \cdots = x_{K-1} = 0$, we get, for $s \geq 1$:

$$g\left(\left(s^2 - 1\right)^{1/2}\right) = c\sqrt{s^2 - 1}$$

for some constant c. This means, for all $u \geq 0$:

$$g(u) = cu$$

which gives

$$Log\,|\,f\,| = c\frac{u^2}{2}$$

and :

$$f(u) = \exp\left(c\frac{u^2}{2}\right).$$

Now, the value of c is determined by normalization : the integral of f must be 1, and we find :

$$f(u) = \exp\left(-\frac{\pi u^2}{4}\right), \text{ for } u \geq 0\,.$$

The same argument holds in general. The identity (8) allows the construction of the function g: fix any value for x_1, for instance $x_1 = 1$. Let c be the unknown value of $g(x_1)$. Fix $u \geq 0$ and find $s, x_2, ..., x_{K-1}$ such that $\varphi(s\,;1, x_2, ..., x_{K-1}) = u$ (this is always possible) : then the identity (8) defines $g(u)$.

From the definition of g, one deduces f, up to a multiplicative constant, which is computed by normalization (the integral of f must be 1). This concludes our construction.

However, there is a compatibility condition ; the identity :

$$g\left(\varphi(s\,;z)\right) = -\frac{1}{\dfrac{\partial\varphi}{\partial x_1}(s\,;z)}\, g(x_1)$$

must give the same value to g when the value of $\varphi(s, z)$ is fixed. This means that the value of the derivative $\dfrac{\partial\varphi}{\partial x_1}(s\,;z)$, for fixed x_1, must depend only on the value of $\varphi(s\,;z)$. In other words, there must exist a function G such that :

$$\frac{\partial\varphi}{\partial x_1}(s\,;z) = G\left(x_1, \varphi(s\,;z)\right) \tag{10}$$

This is the case for the Euclidean norm, since :

$$\frac{\partial\varphi}{\partial x_1}(s\,;z) = \frac{-x_1}{\varphi(s\,;z)}$$

and the same holds for all the l_p norms, $1 \leq p \leq +\infty$.

We note that condition (10) does not need to be satisfied for any norm on the space $\mathbb{R}^K$. Indeed, a norm on $\mathbb{R}^K$ is defined by a symmetric convex body in this space (see for instance B. Beauzamy [1]), and condition (10) means that the value of the derivative $\dfrac{\partial\varphi}{\partial x_1}(s\,;z)$, for fixed x_1, depends only on the value of $\varphi(s\,;z)$. Let us see what this means geometrically, in 3 dimensions :

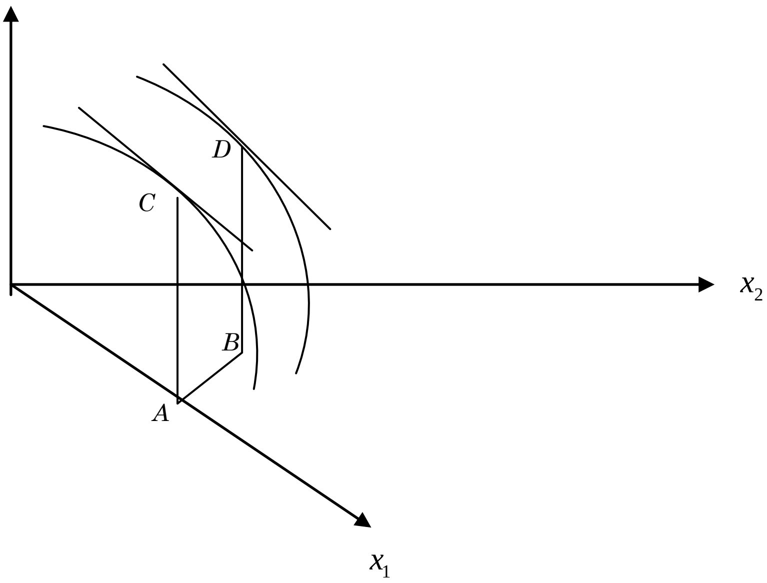

Figure : slopes of the tangent to the unit sphere at two different points

On this picture, the two points A and B have the same value for the first coordinate x_1 and the same value for $\varphi(1; x_1, x_2)$, which is the length $AC = BD$. But the tangent to the unit sphere does not need to be have the same slope in C and in D.

This conclues Step 1 of the proof.

Step 2

We just showed that $f_{Z,S}(z,s)$ was independent of $z = (x_1, ..., x_{K-1})$, for z in the set $A_{K-1}(s)$.

Let $U = \left(\frac{X_1}{S}, ..., \frac{X_{K-1}}{S} \right)$ and let $f_{U,S}(u,s)$ be the joint density of (U, S) ; this density is independent of u, in the set $A_{K-1}(1)$, since :

$$f_{U,S}(u,s) = f_{Z,S}(us,s),$$

for all $s > 0$.

The density of U itself can be obtained, integrating with respect to s that is :

$$f_U(u) = \int_0^{+\infty} f_{U,S}(u,s) ds \qquad (11)$$

and we see that this value is constant on the whole set $A_{K-1}(1)$.

But now, since the density of $\frac{X_1}{S},...,\frac{X_{K-1}}{S}$ is constant, so is the density of

$$\left(\frac{X_1}{S},...,\frac{X_{K-1}}{S},\varphi\left(1;\frac{X_1}{S},...,\frac{X_{K-1}}{S}\right)\right)=\left(\frac{X_1}{S},...,\frac{X_{K-1}}{S},\frac{X_K}{S}\right) \tag{12}$$

This concludes the proof of the Theorem.

Corollaries.

We now show how this general result applies to specific norms :

a) The case of the Euclidean norm

This case has been treated along with the proof of the Theorem. We obtain :

Theorem 2. *-Let $K \geq 1$ and let $X_1,...,X_K$ be independent normal variables (mean 0, variance 1). Then the vector :*

$$\frac{X_1}{\sqrt{X_1^2+\cdots+X_K^2}},\dots,\frac{X_K}{\sqrt{X_1^2+\cdots+X_K^2}}$$

follows a uniform law on the unit sphere of the $K-$dimensional Euclidean space.

In other words, in order to obtain a uniform law on the sphere, one should not start with uniform variables, but with Gaussian.

A proof of this result can be found in the book by R. J. Muirhead [3], p. 37 (communicated by Paul Deheuvels). This proof uses the fact that the normal law on $\mathbb{R}^K$ is invariant under all orthogonal transformations of the space into itself. So, its projection upon the unit sphere is itself invariant under the projections of the transformations. The unit sphere is a locally compact group for these transformations, and a general result says that there is only one invariant Haar measure, which must be the uniform law.

b) The case of the l_1 norm

This is treated in the book by Luc Devroye [2] ; we get :

Theorem 3. *-Let $K \geq 1$ and let $X_1,...,X_K$ be independent variables following an exponential law. Then the vector :*

$$\frac{X_1}{X_1+\cdots+X_K},\dots,\frac{X_K}{X_1+\cdots+X_K}$$

follows a uniform law on the positive part of the unit sphere of the $K-$dimensional l_1 space.

c) The case of the l_p norms, $1 \le p \le \infty$

This is a generalization of the previous ones. The case $p = \infty$ is obvious (this is just the uniform law in the unit ball). For the case $1 \le p < +\infty$, we get :

Theorem 4. *-Let $K \ge 1$ and let $X_1, \ldots, X_K$ be independent variables following a law with density $\exp\left\{-t^p\right\}$ $t \ge 0$. Then the vector :*

$$\frac{X_1}{\left(X_1^p + \cdots + X_K^p\right)^{1/p}}, \ldots, \frac{X_K}{\left(X_1^p + \cdots + X_K^p\right)^{1/p}}$$

follows a uniform law on the positive part of the unit sphere of the $K-$dimensional l_p space.

Proof of Theorem 4.

We indicate briefly how this density is obtained. In this case, we have :

$$\varphi\left(s\,;x_1, \ldots, x_{K-1}\right) = \left(s^p - x_1^p - \ldots - x_{K-1}^p\right)^{1/p}$$

We have :

$$\frac{\partial \varphi}{\partial x_1} = -\left(\frac{x_1}{\varphi}\right)^{p-1}$$

so condition (10) is satisfied.

The identity (8) becomes :

$$g\left(\left(s^p - x_1^p - \ldots - x_{K-1}^p\right)^{1/p}\right) = \left(\frac{\left(s^p - x_1^p - \ldots - x_{K-1}^p\right)^{1/p}}{x_1}\right)^{p-1} g(x_1) \qquad (13)$$

Taking $x_1 = 1$, $x_2 = \cdots = x_{K-1} = 0$, we get, for $s \ge 1$:

$$g\left(\left(s^p - 1\right)^{1/p}\right) = c\left(s^p - 1\right)^{\frac{p-1}{p}}$$

for some constant c. This means, for all $u \ge 0$:

$$g(u) = cu^{p-1}$$

which gives

$$Log\left|f\right| = c\frac{u^p}{p}$$

and :

$$f(u) = \exp\left(c\frac{u^p}{p}\right).$$

The value of c should be determined by normalization ; of course $c<0$, but we may just as well take :

$$f(u) = \exp\left(-u^p\right), \text{ for } u \geq 0,$$

since we divide at the end by the quantity $\left(X_1^p + \cdots + X_K^p\right)^{1/p}$. This concludes the proof of Theorem 4.

References

[1] B. Beauzamy : Introduction to Banach Spaces and their Geometry. North Holland, Collection « Notas de Matematica », vol. 68, second edition,1985.

[2] Luc Devroye : Non Uniform Random Variate Generation, Springer-Verlag, New York, 1986.

[3] R. J. Muirhead : Aspects of Multivariate Statistical Theory. Wiley, New York, 1982.

Acknowledgements

We thank Paul Deheuvels for indicating the book by Luc Devroye, and Luc Devroye for his comments about the proofs.

Originally, our preoccupation about "random directions" in a Banach spaces comes from two contracts we had with the "Institut de Radioprotection et de Sûreté Nucléaire" (IRSN, France) : R50/11026029 (17/11/2006) and DA 79160/CA 11032655 (11/03/2008) ; they deal with the safety of nuclear reactors, and the question was : in a computational code depending upon many parameters, find the configuration which might lead to the highest temperature. The main tool was the "Experimental Probabilistic Hypersurface", which provides a probability law for the temperature, at each point of the configuration space. See http://www.scmsa.com/RMM/RMM_EPH.htm for more details.

bernard.beauzamy@scmsa.com