67. *Gaussian Measure on the Projective Limit Space of Spheres*

By Takeyuki HIDA and Hisao NOMOTO

(Comm. by Kinjirô KUNUGI, M.J.A., May 9, 1964)

1. Let $S^n(r_n)$ be the sphere with center 0 and radius r_n in the $(n+1)$-dimensional Euclidean space. $S^n(r_n)$ can be expressed as follows:

$$(1)\qquad \begin{cases} x_1 = r_n \prod_{i=1}^{n} \sin\theta_i, \\ x_k = r_n \cos\theta_{k-1} \prod_{i=k}^{n} \sin\theta_i, \qquad 2\leqq k \leqq n, \\ x_{n+1} = r_n \cos\theta_n, \end{cases}$$

where $0\leq\theta_1<2\pi$ and $0\leqq\theta_i\leqq\pi$, $i\geqq 2$.

Consider the probability space $(S^n(r_n), \mathcal{B}_n, P_n)$ with topological Borel field $\mathcal{B}_n$ and uniform probability measure P_n given by

$$(2)\qquad dP_n = \frac{1}{2}\Gamma\left(\frac{n+1}{2}\right)\pi^{-\frac{n+1}{2}}\left[\prod_{i=2}^{n}\sin^{i-1}\theta_i\right]d\theta_1\cdots d\theta_n.$$

Then for $a=(a_1,\cdots,a_{n+1})\in S^n(R)$, we can define a random variable X_n by

$$X_n = X_n(x) = \frac{1}{r_n}\sum_{i=1}^{n+1} a_i x_i, \qquad x\in S^n(r_n),$$

the characteristic function of which is

$$\varphi_n(\lambda) \equiv \int_{S^n(r_n)} \exp\left[i\lambda X_n\right] dP_n(x) = \Gamma\left(\frac{n+1}{2}\right)\left(\frac{|\lambda|\ \|a\|}{2}\right)^{-\frac{n-1}{2}} J_{\frac{n-1}{2}}(|\lambda|\ \|a\|).$$

In particular, if $R=\|a\|=\sqrt{n+1}$, we have

$$(3)\qquad \lim_{n\to\infty}\varphi_n(\lambda) = \exp\left[-\lambda^2/2\right],$$

which shows that the distribution of X_n converges to the standard normal distribution.

The purposes of the present note are to construct the Gaussian measure on a certain infinite dimensional space, with respect to which all the X_n's are measurable and to define a typical family of functions analogous to spherical harmonics. In the last section we shall give an information theoretic characterization of our construction of Gaussian measure.

2. Let $\prod^n=\{(\theta_1,\cdots,\theta_n);\ 0<\theta_1<2\pi, 0<\theta_i<\pi, 2\leqq i\leqq n\}$ and let Ω_n be the open subset of $S^n(\sqrt{n+1})$, which is homeomorphic to $\prod^n$ by (1). The mapping $f_{m,n}(m<n):\Omega_n\to\Omega_m$ induced by the projection $\prod^n\ni(\theta_1,\cdots,\theta_n)\to(\theta_1,\cdots,\theta_m)\in\prod^m$ is continuous. Obviously the following relations hold.

$$f_{l,n}=f_{l,m}\circ f_{m,n},\ l<m<n,$$
$$P_m(A)=P_n(f_{m,n}^{-1}(A))\ (m<n),\ A\in\boldsymbol{B}_m\equiv\mathcal{B}_m[\Omega_m].$$

We have therefore a topological stochastic family $\{(\Omega_n, \boldsymbol{B}_n, P_n); n\geqq 1\}$ in the sense of S. Bochner. Hence we can form a probability space $(\Omega, \boldsymbol{B}, P)$ such that

i) Ω is the subset of the weak product $\prod_{n=1}^{\infty}\Omega_n$ having the property that

$$f_{m,n}(x^{(n+1)})=x^{(m+1)} \text{ with } x=(x^{(2)}, x^{(3)},\cdots)\in\Omega,\ x^{(n+1)}\in\Omega_n,$$

ii) putting $f_{n,\infty}(x)=x^{(n+1)}$, $\boldsymbol{B}$ is the σ-algebra generated by $\bigcup_{n=1}^{\infty}(f_{n,\infty}^{-1}(\boldsymbol{B}_n))$,

iii) the measure induced by $f_{n,\infty}$ from P is identical with P_n on Ω_n (S. Bochner [1], Theorem 5.1.1).

3. With respect to this probability measure $P, \theta_1, \theta_2, \cdots$ are mutually independent and each θ_k is also a random variable on $(\Omega_n, \boldsymbol{B}_n, P_n)$ with $n\geqq k$, namely $\theta_k(x)=\theta_k(x^{(n+1)})$, $x^{(n+1)}\in\Omega_n$.

Concerning the random variables $x_j^{(n)}=x_j^{(n)}(x)=j$-th coordinate of $f_{n-1,\infty}(x)$ $j\leqq n$, $n=1,2,\cdots$, we have

$$\int_{\Omega} x_j^{(n)}dP(x)=0 \text{ for every } n \text{ and } j\ (\leqq n),$$

$$\int_{\Omega} x_i^{(n)}x_j^{(n')}dP(x)=\int_{\Omega_{n'-1}} x_i^{(n)}x_j^{(n')}dP_{n'-1} \quad (n'\geqq n)$$

$$=\begin{cases} \sqrt{\dfrac{n'}{n}}B\Big(\dfrac{n+1}{2},\dfrac{n'}{2}\Big)\Big/B\Big(\dfrac{n}{2},\dfrac{n'+1}{2}\Big), & \text{if } n\leqq n' \text{ and } i=j,\\ 0 \quad \text{if } i\neq j. \end{cases}$$

Let $a=(a_1, a_2,\cdots)$ be in l^2, i.e., $||a||^2=\sum_{i=1}^{\infty}a_i^2<\infty$, and put $a^{(n)}=(a_1,\cdots,a_n)$. Define $X_n=X_n(x)=\sum_{i=1}^{n}a_ix_i^{(n)}$. Then $\{X_n\}$ forms a Cauchy sequence in $L^2(\Omega, P)$. Hence there exists a random variable $X(a)=\underset{n\to\infty}{\text{l.i.m.}}\, X_n$, which is uniquely determined by a. We can easily see by (3) that it subjects to a normal distribution with mean 0 and variance $||a||^2$.

The mapping $\sigma: l^2\ni a\to X(a)=X(a,x)\in L^2(\Omega, P)$, is linear and isometric transformation from l^2 to $L^2(\Omega, P)$. For any complete orthonormal system $\{e_k\}$ of l^2, we have a sequence of independent Gaussian random variables $\{\xi_k\}$, each of which has mean 0 and variance 1, where $\xi_k(x)=\sigma(e_k)$. Every $X(a)$, $a\in l^2$, can therefore be developed in the form (orthogonal development):

$$X(a)=\sum_{k=1}^{\infty}a_k\xi_k, \qquad a_k=(a, e_k) \text{ in } l^2.$$

This series indeed converges almost surely.

Next we consider a series

$$(4) \qquad \sum_{q=1}^{\infty}\xi_q/q^p, \quad p\,(\text{integer})\geqq 1.$$

Obviously it converges almost certainly. Let Ω_p be the set of x's for which (4) converges, and let $\Omega^*=\bigcap_p \Omega_p$. Ω^* has probability one. Now we have

THEOREM 1. *For every $x\in\Omega^*$, the functional on $(\mathfrak{s})$ defined by*

$$F(a)=X(a, x) \qquad a\in(\mathfrak{s})$$

belongs to $(\mathfrak{s})'$. Here $(\mathfrak{s})$ denotes the space of rapidly decreasing sequences and $(\mathfrak{s})'$ is the dual space of $(\mathfrak{s})$.

Define

$$\zeta_n=\sqrt{n+1}\prod_{i=2}^{n}\sin\theta_i \quad \text{and} \quad \eta_n(a)=\sum_{i=2}^{n}[a_{i+1}\cos\theta_i/\sin\theta_2\cdots\sin\theta_i]$$

$n\geqq 2$, $a\in l^2$. Then $\{\zeta_n, \boldsymbol{B}_n;\ n\geqq 2\}$ and $\{\eta_n(a), \boldsymbol{B}_n;\ n\geqq 2\}$ are lower semi-martingale and martingale respectively. Thus we can prove that ζ_n and $\eta_n(a)$ converge to certain random variables ζ and $\eta(a)$ with probability one respectively. Thus $X_n(a)$ can be expressed in the form

$$X_n(a)=(a_1\sin\theta_1+a_2\cos\theta_1+\eta_n(a))\cdot\zeta_n$$

and it converges to

$$X(a)=(a_1\sin\theta_1+a_2\cos\theta_1+\eta(a))\cdot\zeta$$

almost certainly. Combining this with the results obtained above, we have

THEOREM 2. *$X_n(a)$ converges to $X(a)$ both in the mean and with probability one.*

4. The homogeneous harmonic polynomials of p-th degree on S^n can be expressed in the form

$$(5) \qquad Y(p, m_1,\cdots, m_n;\ \theta_2,\cdots,\theta_n, \pm\theta_1)$$
$$=\exp[\pm i m_1\theta_1]\prod_{k=1}^{n-1}(\sin\theta_{k+1})^{m_k}C_{m_{k+1}-m_k}^{m_k+\frac{k}{2}}(\cos\theta_{k+1})$$

where C denotes the Gegenbauer polynomial and $p\geqq m_n\geqq m_{n-1}\geqq\cdots\geqq m_1\geqq 0$. These are also functions on Ω and belongs to $L^2(\Omega, P)$.

Keeping m_n's constant for large n, let us consider the following limit

$$\lim_{n\to\infty}(\sqrt{n})^{m_n}\exp[\pm i m_1\theta_1]\prod_{k=1}^{n-1}(\sin\theta_{k+1})^{m_k}C_{m_{k+1}-m_k}^{m_k+\frac{k}{2}}(\cos\theta_{k+1}).$$

This limit exists almost surely by the same reason as ζ_n converges and it determines a function $Y_p(\{m_k\};\ \{\theta_i\}, \pm\theta_1)$ on Ω.

THEOREM 3. *The family of functions $\{c_p Y_p(\{m_k\};\ \{\theta_i\}, \pm\theta_1);\ p, m_1, m_2,\cdots=0, 1, 2,\cdots\}$ forms a complete orthonormal system of $L^2(\Omega, P)$, where c_p is the normalized constant.*

For these Y's in (5) with $m_n=m_{n-1}=\cdots=m_m$, letting $\theta_n=\theta_{n-1}=\cdots=\theta_m=\dfrac{\pi}{2}$, we have homogeneous harmonic polynomials on S^m. Besides, from $\{Y_p(\{m_k\}, \{\theta_i\}, \pm\theta_1)\}$ on Ω with $m_n=m_{n+1}=\cdots$, we get all the homogeneous harmonic polynomials on S^n if ζ is replaced with ζ_n. In other words, all the harmonic polynomials on finite dimensional sphere can be obtained from $\{Y_p(\{m_k\}, \{\theta_i\}, \pm\theta_1)\}$ by the mapping $f_{n,\infty}$.

5. In our course of constructing Gaussian measure, the uniform measure on S^n and projections $\{f_{m,n}\}$ have played essential roles, which can also be illustrated information-theoretically.

Let us find the probability distribution of $(\theta_1,\cdots,\theta_n)$ which has *maximum entropy* among all distributions satisfying the following three conditions:

i) $(x_1,\cdots,x_{n+1})$ defined by (1) with $r_n=\sqrt{n+1}$ has absolutely continuous distribution on S^n,

ii) the mean value of $\log\left(1-\frac{x_{n+1}^2}{n+1}\right)=\psi\left(\frac{n}{2}\right)-\psi\left(\frac{n+1}{2}\right)$, where ψ is the digamma function,

iii) the (marginal) distribution of $(\theta_1,\theta_2,\cdots,\theta_{n-1})$ gives the uniform measure on $S^{n-1}(1)$.

To solve the problem in question, the next lemma is usefull.

LEMMA. *If $f(x)$ and $g(x)$ are both probability density functions on the N-dimensional Euclidean space R^N, we have*

$$-\int_{R^N} f(x)\log f(x)\,dx \leqq -\int_{R^N} f(x)\log g(x)\,dx.$$

Using this lemma, we can conclude that the distribution having maximum entropy under the above conditions must be the very P_n given by (2).

Now we can illustrate the reason why we have used the uniform measure on the sphere from the point of view of maximum entropy. We first take the uniform probability measure on the circle ($n=1$), the entropy of which is larger than any other possible distributions. Next, suppose that θ_1 subjects to the uniform distribution. Then, from the above discussions, the distribution P_2 of (θ_1,θ_2) has maximum entropy under the restrictions i)-iii). Hence we take P_2 as the distribution of $(\theta_1,\theta_2),\cdots$ and so on.

The above discussions about the determination of the distribution of $(\theta_1,\cdots,\theta_n)$ show that the measure P contains maximum information quantity in some sense, from which we can see the propriety of using uniform measure and projections $\{f_{m,n}\}$ as the entropy maximum preserving mappings.

References

[1] S. Bochner: Harmonic Analysis and the Theory of Probability. Univ. of Calif. Press (1955).

[2] P. Lévy: Problèmes concrets d'analyse fonctionelle. Gauthier-Villars (1951).