無限：確率解析における無限次元と無限大の歴史

飛田　武幸

平成19年10月27日

1　はじめに

どの分野についても言えることであるが、とくに比較的新しい数学の分野である確率論では、理論が急速に進歩するに従って、その基礎や出発点における準備事項もまた絶えず見直しをもとめられる。例えば、理解を助けるために用いられた通常語などは、より深く広い数学用語でその意義を明らかにしなければならなかったりする。

ここで取り上げたいことは確率論における無限という形容詞のつく種々の用語と、派生して生じる超汎関数である。前者は確率の歴史とともに登場し、しかも主題を定めていることである。無限大に対比しては無限小もあり、もし、そのとらえ方に誤解があれば超汎関数の基本的なアイデイアが損なわれるであろう。

確率論は歴史的にも、無限の事象を扱い、しかもそれが"先細りの無限"でなく、毎日太陽が東から昇るように、いつまで経っても変わりはないような無限である。そんなものを正しく過不足なく認識するためには工夫が必要で、当面の研究課題にも係わってくる。さらに連続無限にしたときは、余計な注意も必要となる。

本稿では、このような問題の解決ではなくて、その理解について問題提起をしてみようと思うのである。

2　確率変数系による無限次元の表現

確率論においては、大数の法則、中心極限定理など、無限個の確率変数を扱った歴史は古い。また定常過程のスペクトル分解、Lévy 過程の Lévy-Itô 分解、ホワイトノイズなどの扱いで、連続無限個のランダム量を考えなければならないが、その都度、それぞれの立場から適切な説明がなされてきた。それらを比較対照し、できるだけ統一的な立場で理解しようとすることは自然な成り行きであった。

１０年以上も前の話になる。「無限次元」について、より深く形而上学的な立場からの扱いについての批判を得たいと考え、哲学者の吉田民人先生の門を敲いた。会話の中から教えられたところが多く、大変有難く思い感謝している。

以下に述べるのは、無限について以前から気になっていたが、なかなか言い出せなかったことである。ランダム系の **Reduction** のことも併せて考えてみたい。

対象とする集合の各点に「独立確率変数」を対応させて、その**集合を表現**して、無限集合、あるいは無限所威厳が潜在的にもつ特徴を確率的に考えたい。このとき集合は確率変数列のパラメータとして具現化される。

(1) パラメータの濃度で分類。

i) 可算集合の場合 濃度 **a**.

Z で代表する。すなわち

$$Y_n, n \in Z.$$

ここで、n を平等に扱うとすれば、i.i.d.(independent, identically distributed) 変数の系によって表現される。

このとき、$\{Y_n\}$ の確率分布が対応する。それは無限次元の分布、すなわち $R^{\mathbf{a}}$ 上の確率分布である。(Kolmogoroff の定理).

極限定理、すなわち 0-1 law, 大数の法則、domain of attraction (中心極限定理)、などはこの分布を基礎にする。それは、次元について先細りしていない。すべての変数は平等である。

i.i.d. 変数列の関数　$F(X_n, n \in Z)$ の解析を考えるのには確率分布 m が必要となる。それは１次元分布の可算無限個の直積である。その 台 (support) は当然 l^2 ではない。そこで

$$E \subset l^2 \subset E^*$$

なる Gel'fant triple をつくり、m は E^* の上の測度とみられるから、この空間の上で定義された関数の扱いは無限次元、しかも本質的（要説明）に無限次元解析である。

極限定理は確率論の華とされた時代もあった。それは可算確率論であった。

他に離散無限の特徴を考えたアプローチとして、時系列の *innovation* がある。これは与えられた時系列から、時刻の推移に応じて同じ情報をもつ独立確率変数列を構成する問題である。Innovation という言葉は Wiener, Masani, Kallianpur、Rosenblatt や engineer 達が用いているが、実質的な innovation の意義や構成法は Lévy [8] (1937) Chapt 8. に見られる。条件付確率に制限がついたりして、ガウス過程など特別な場合を除き、あまり良い結果は得られていない。innovation の条件を」弱めて、独立のかわりに直交にすれば、一定の成果はみられた。しかし、我々は直交よりも独立にこだわりたい。

ii) 連続無限集合。**c**.

R で代表される。確率論では連続時間のパラメータとして用いられる。物理学では単に実数というよりは、time operator とか Hamiltonian と conjugate なものとして、付加的な意味も考えられている。我々は独立確率変数系で表現したい。

当面二つの場合を考える。

イ）各 t で独立。i.i.d. 系とすれば、$Y_t, t \in R$ の分布は R^R 上の測度で、任意の座標の置換で不変である。可分性はない。真に連続無限次元の測度による表現となる。

[注意] 真に連続無限次元は本質的な無限次元とは異なる。

ロ）各 dt で独立（象徴的に)。このときは、Gel'fand の意味での**各点独立超過程**により表現される。その特性汎関数 $C(\xi)$ は

$$\xi_1(t)\xi_2(t) = 0 \ \rightarrow \ C(\xi_1 + \xi_2) = C(\xi_1)C(\xi_2)$$

をみたす。

このとき、 i.i.d. に相当するものは、**各点独立定常超過程**となる。

このように、可分性があるときを **可分連続無限** の場合と呼ぼう。

最も基本的な例は Lévy 過程の時間微分である。ここで、微分は見本関数ごとにとるものとする。

次の段階として、「可分性」がある。、関数の解析に進むための大事な要請である。

例 1． $\dot{B}(t)$, ただし、$B(t)$ はパラメータ空間を R とするブラウン運動である。

例 2. $\dot{P}(t)$, ただし、$P(t)$ は ポアソン過程である。

各点独立の定義にあるように (例えば Gel'fand 1955)、t を時間とみて、時間区間が overlap しなければ、二つのランダム量は独立となる。時間区間はいくら狭くてもよいが ξ をきめるために内部がなければならない。象徴的に 各 dt 毎に独立といったのである。このことと、$\dot{B}(t)$ がそれ自身意味をもって独立になることとは区別しなければならない。そこには、区間が 1 点に近づく位相の配慮が必要である。その意味で イ）とは厳格に区別される。$dB(t)$ と $\dot{B}(t)dt$ との違いにも注意したい。

ここで問題は二つある。

（1）各 $\dot{B}(t)$ や $\dot{P}(t)$ などに確固とした定義を与えなければならない。従来の確率解析では、$\dot{B}(t)$ を smear する関数を δ-関数に近づければよいように見えるかもしれないが、その収束の仕方（すなわち topology) は、存在が保障された $\dot{B}(t)$ を含む空間の位相でなければならない。従来の空間では狭すぎる。δ-関数が L^2-関数ではその位相で近似できないように。

（2）　$\dot{B}(t)$ や $\dot{P}(t)$ に identity を与えれば、あとはそれらを変数とする関数（勿論非線形）を考えることになり、それらの超汎関数の導入が続く。

この場合、Fock space のように超汎関数空間の直交分解を考えればよいのではなくて、いわゆる renormalization が必要となる。たとえば $(dB(t))^2 = dt$ とする近似ではなく、両辺の差として残る randomness を拾い上げたい。そこで差を magnify した

$$: \dot{B}(t)^2 := (\frac{dB(t)}{dt})^2 - \frac{1}{dt}$$

が $(\dot{B}(t))^2$ の renormalized function である。

同様にして $(\dot{B}(t))$'s の多項式の renormalization も定義され、それらを含めた**超汎関数空間** が構成される。

超汎関数空間は整域であることも証明されている。(Ojima-Saigo-Hasebe.)

iii) 関数濃度。**f**.

関数空間を $\mathbf{f}$ とする。任意の $f \in \mathbf{f}$ に対して Y_f が対応して系 $\{Y_f, f \in \mathbf{f}\}$ は独立確率変数系をなす。定常性にあたる性質を仮定すれば、その確率分布は R^{R^R} 上の直積測度である。

artificial なもの以外に有意義なものがあるのだだろうか？

3　可分で無限次元の場合

表題のケースで、代表的に、ホワイトノイズ $B(t)$ とポアソン・ノイズ $\dot{P}(t)$ をとりあげて、比較対照しながら、ランダムな無限次元を考察する。

i) 情報量の比較

二つのノイズを情報量の立場から比較する。定常であるため、dt 時間の情報を比較すればよい。

イ）分散（エネルギー）一定として。ガウス分布が最大のエントロピーをもつ。すなわちホワイトノイズである。

ロ）平均を一定（たとえば $\frac{1}{\lambda}$）とすれば、密度関数が $\lambda e^{-\lambda t}$ の指数分布が最大エントロピーをもつ。ところがポアソン過程のとき $dP(t)$ の分布は dt の高次の項を除き $P(dP(t)=0)$ で近似される。それは $e^{-\lambda dt}$ である。密度関数は指数分布のそれと一致する。

ハ）エントロピー有限な Lévy 過程。

指数 α の安定分布は long tail をもつが、緩増加関数の factor を除き遠くで $x^{-(\alpha+1)}$ のオーダーであり、エントロピーは有限である。

ii) 構成法について。

時間的に一様な Lévy 過程により可分な連続無限集合を表現する。時間変数は実数で、線形順序構造が入る。

Interporation の方法にこだわる理由：

1. 逐次（前の近似を活かして、射影的に）近似するのが適当.

2. 時間的に一様に近似していきたい。

3. 可分連続無限の意味を活かした構成をしたい。

イ）ホワイトノイズ（ブラウン運動）　Lévy の interporation　による構成法（近似法でもある）。

T_n を 2 進数 $k/2^{n-1}, k = 0, 1, 2, \cdots, 2^{n-1}$ の集合とする。 $T_0 = \cup_{nge1} T_n$ とおく。

構成 $\{Y_n = Y_n(\omega), n \geq 1\}$ は 確率空間 $(\Omega, \mathbf{B}, P)$ 上の標準ガウス分布に従う独立確率変数列とする。 $X_1(t)$ を

$$X_1(t) = tY_1.$$

で定める。以下帰納法による。

確率過程列 $X_j(t) = X_j(t, x), j \leq n$, が構成されたとする。このとき, $X_{n+1}(t)$ をつぎのように定義する：

$$X_{n+1}(t) \quad = \begin{cases} X_n(t), \quad t \in T_n, \\ \dfrac{1}{2}(X_n(t + 2^{-n}) + X_n(t - 2^{-n})) + 2^{(n+1)/2} Y_k, \\ t \in T_{n+1} - T_n, \ k = k(t) = 2^{n-1} + (2^n t + 1)/2, \\ (k + 1 - 2^n t) X_{n+1}(k2^{-n}) + (2^n t - k) X_{n+1}((k+1)2^{-n}), \\ t \in [k2^{-n}, (k+1)2^{-n}], \end{cases}$$

ここでは ω は省略した。

注意したいことは、確率変数列 $X_n(t), n \geq 1,$ は consistent であり、かつ逐次近似が t について一様だということである。

この極限を $\tilde{X}(t)$ と書く. それは、平均 0 で、独立増分をもち

$$E(|\tilde{X}(t) - \tilde{X}(s)|^2) \leq |t - s|,$$

$$Cov(\tilde{X}(t), \tilde{X}(s)) = t \wedge s.$$

がわかる。

さらに、殆どすべての ω に ついて

$$\lim_{n \to \infty} X_n(t, \omega) = X(t, \omega),$$

が存在し、それは $\tilde{X}(t)$, a.e. (μ) に等しい。

以上をまとめて

定理 極限の過程 $X(t, \omega), t \in [0, 1]$, は確率空間 $(\Omega, \mathbf{B}, P)$ 上のブラウン運動である。.

この interporation による近似は、上記の特性１.〜３ を象徴するのみならず、無限次元回転群の要素の実例を見つけることができて、ホワイトノイズの一つの latent trait をみることができる。

[註] ブラウン運動の構成には Random Fourier series による N. Wiener の方法もあるが、今の我々の目的には沿わないので省略する。

ロ) ポアソン・ノイズを指数分布をもつ確率変数列により構成すること。

これも確率空間 $(\Omega, \mathbf{B}, P)$ 上で行う。

$X_n(\omega), n \geq 0$ を、密度関数が $\lambda e^{-\lambda t}, t \geq 0,$ の、$[0, \infty)$ の指数分布に従う独立確率変数列とする。ここで、λ は正定数である。$P(t,\omega)$ を

$$(\omega; P(t,\omega) \leq n) = (\omega; \sum_{0}^{n} X_k > t)$$

によって定義される。

これを t で微分してポアソン・ノイズが構成される。

定理 ポアソン・ノイズの configuration space は独立に散らばった δ-関数の集合である。この可算無限の状態は projective limit により構成される。

文献 [6] 参照。可分なことは、これからもわかる。

逆に $P(t,\omega)$ から $X_n(\omega)$ も定まる。

構成方法のアイデイアはブラウン運動の Lévy による方法とにているが、双対性を示唆するところもあることに注意したい。

ハ) 複合ポアソン過程 (ノイズ) の構成。

やはり、連続無限個の独立なポアソン過程の複合であり、連続無限に拘らざるをえない。separability を見るために u-軸の分割の細分を可算個ですます構成方法を考えなければなるまい。いろいろなアプローチも考えられようが、節を改めて、安定過程に制限して、一つの試みを述べたい。

一般的な注意であるが、複合ポアソン過程から、見本関数を見て、素な構成要素である各ジャンプ u のポアソン過程を取り出すというが、その intensity (今の場合 Lévy measure $dn(u)$) までをきめるのは瞬間的にはできない。例えば dt 時間の行動を見なければならない。

iii) Innovation による方法。

連続パラメータの確率過程 $X(t)$ の場合にも、可分を仮定して、innovation が定義され、離散的な場合よりも重要な役割を演ずる。そのアイディアは Lévy の 1953 年の Berkeley Lecture にみられる。確率過程 $X(t)$ に対して、stochastic infinitesimal equation

$$\delta X(t) = \Phi(X(\tau), \tau \le t, Y(t), t, dt)$$

を提唱している。ここで $Y(t)$ が innovation である。説明は要しないと思うが、我々は次のことを学び取ることができる。

1) non-random なときの変分と較べて、$Y(t)$ が余分に出てくる。

2) 再び formal に

$$X(t) = \Psi(Y(s), s \le t, t), \quad t \ge 0, \quad X(0) = 0,$$

が期待される。

上の 1), 2) は innovation approach である。これで、$X(t)$ が与えられたとき innovatiopn $Y(t)$ を構成することにより 実数 t の表現が得られた。

4 Calculus

以下ガウス型ノイズを中心にして議論を進める。

$\dot{B}$ の見本関数を x と書く。それは超関数である。超汎関数を $\varphi(x)$ とし、Fourier 変換の類似として、T 変換

$$(T\varphi)(\xi) = \int e^{i<x,\xi>}\varphi(x)d\mu(x),$$

ただし μ はホワイトノイズの確率分布、すなわち標準ガウス分布で ξ はテスト関数である。

また S 変換は

$$(S\varphi)(\xi) = \int \varphi(x+\xi)d\mu(x),$$

である。いずれも超汎関数に対して定義される。とくに $\dot{B}(t)$ の renormalized polynomial には $\xi(t)$ の同じ polynomial が対応する（定数 i などを除く）。

両変換は familiar なヒルベルト空間の同型を与える変換とはことなり、image は再生核ヒルベルト空間である。そこでは、おおくのメリットが見られる。例えば $\{\dot{B}(t)\}$ の分布であるホワイトノイズ測度 μ が shift に関して

エルゴード的であるので、もとの ensemble 平均が時間平均になり、我々にとっては computable なものとなる。また収束は見やすいものとなる等である。

最も重要なことは、$\dot{B}(t)$ による偏微分が S 変換で **Fréchet derivative** に置き換わることである。(Gâteaux derivative ではない。)

$$\partial_t \varphi (= \frac{\partial}{\partial \dot{B}(t)} \varphi)$$

$$= S^{-1} \frac{\delta}{\delta \xi(t)} (S\varphi)(\xi).$$

ここですべての t が一様に出てくることに注意する。

微分（実は annihilation) ∂_t の adjont は ∂_t^* である。それは creation として作用し確率積分が 被積分関数が non-anticipating などの付加条件なしに定義できる。

Lapalcian も同様で

$$\Delta_V = \int \partial_t^2 dt,$$

$$\Delta_L = \int \partial_t^2 (dt)^2.$$

amd so on.

こうして、実数を $\dot{B}(t)$ で表現したことの advantage が見られる。

5 無限と近似

近似と極限とは区別しなければならない。

1) δ 関数は L^2-関数で近似される。しかし、近似の位相は超関数のものである。極限の δ 関数は勿論 L^2-関数ではない。

2) （正規化された）二項分布の極限はガウス分布であるという。離散分布をも含む分布の集合に位相を入れてからの話である。二項分布もガウス分布で近似される。

3) $\dot{B}(t)$ の近似。$\frac{\Delta B}{\Delta}$ で近似されるという。それを主張するなら $\dot{B}(t)$ の住む空間（実は $H_1^{(-1)}$ である）をあらかじめ作っておかなければならない。それはブラウン運動の関数からなるヒルベルト空間 (L^2) ではない。$\dot{B}(t)$ の複雑な関数を扱うときは formal な扱いではすまない。単純な非線形関数 $\dot{B}(t)^2$ でさえ然りである。

補足

理論的には、$\dot{B}(t)$ **に市民権を与える** ことがホワイトノイズ解析のスタートである。

まず isomorphism

$$H_1 = \{\dot{B}(f), f \in L^2(R)\} \;\cong\; L^2(R)$$

を拡張し, order -1 のソボレフ空間 $K(R)^{-1}$ をとり

$$H_1^{(-1)} \;\cong\; K(R)^{-1}$$

を達成し、$\dot{B}(t)$ 達を $H_1^{(-1)}$ の正規のメンバーとすることであった。この H_1 と $H_1^{(-1)}$ との区別を明瞭にしなければ怪我をすることになる。$\frac{\Delta B}{\Delta}$ は H_1 の位相では　$\dot{B}(t)$ に近づけないのである。

$\dot{B}(t)$ 達の非線形関数については、各次数の homogeneous chaos, multiple Wiener integrals 毎に、対称ソボレフ空間 の助けを借りて, 定数 c_n を除き

$$H_n^{(-n)} \;\cong\; \hat{K}(R^n)^{-(n+1)/2}$$

ホワイトノイズの n 次超汎関数空間 $H_n^{(-n)}$ を定義するが、$n \geq 2$ では、いわゆる *renormalization* の技法を必要とする。

以上は ホワイトノイズ解析の一つの重要なポイントである。

6 Things to be done.

1. ブラウン運動の軌跡について.

1次元でも、その法則の細かいことが知られている。注意したいことは、そこに unharmonic ratio がしばしば現れていることである。このまま、そして多次元パラメータの Lévy のブラウン運動において,projective invariance あるいは conformal invariance が示される筈である。

２．多次元値ブラウン運動については、より興味深い invariance が出てくるであろう。ホワイトノイズの言葉で設定したい。ポアソンの場合もノイズにした方が扱いやすい。

３．双対性への指向。

$\dot{B}(t)$ と $\dot{P}(t)$ を比較したとき、それらの性質について、著しい similarity と dissimilarity がみられる。この両者について、それらを繋ぐものの一つと

して duality がある。実数の表現という立場から、両者の間の duality を見ることは興味ある問題であろう。[14] 参照。

いづれも無限の profound な理解を避けては通れまい。

参考文献

[1] W. Feller, An introduction to probability theory and its application. vol.1 (1950) (1967), vol.2 (1966). 邦訳あり。

[2] T. Hida, Stationary stochastic processes. Princeton Univ. Press. 1970.

[3] 飛田武幸、ブラウン運動。(1975) 岩波書店、英訳 (1980), Springer-Verlag. 第3刷 2007.

[4] T. Hida, Analysis of Brownian functionals. Carleton Math. Notes no. 13, 1975

[5] 飛田武幸、ホワイトノイズ解析入門。東京理科大学 講義ノート。June, 2007.

[6] T. Hida and Si Si, Lectures on white noise functionals. Worlde Scientific Pub. Co. 2008.

[7] P. Lévy, Les loi de probabilité dans les ensembles abstraits. Revue de Métaphysique et de Morale. exposés au Collége de France en 1919 et 1924. 21 page article.

[8] P. Lévy, Théorie de l'addition des variables aléatoires. Gauthier-Villars, 1937.

[9] P. Lévy, Processus stochastiques et mouvement brownien. Gauthier-Villars, 1948. 2éme ed. 1965.

[10] P. Lévy, Problèmes concrets d'analyse fonctionnelle. Gauthier-Villars, 1951.

[11] J. Mikusinski, On the square of the Dirac Delta-distribution. Bull. de l'Academie Polonaise des Sciences. VIV, (1966), 511-513.

[12] V.A. Rohlin, On the fundamental ideas of measure theory. Mat. Sbornik 25 (1949), 107-150 :english trans. AMS Translations Ser 1 vol. 10 (1962) 1-54.

[13] Si Si, Effective determination of Poisson noise. IDAQP 2003.

[14] Si Si, A duality between Gaussian noise and Poisson noise, and quadratic Hida distributions. to appear

[15] N. Wiener, Differential space. J. Math. and Phys. 2 (1923), 131-174.