日本神経回路学会誌　Vol. 11, No. 1（2004），1–8

解　　説

物理学者でない人にとって平衡統計力学とは

渡辺澄夫

東京工業大学*

Invitation of All People to Statistical Physics

Sumio Watanabe

Tokyo Institute of Technology*

概要

この解説では、人工知能や情報工学などの物理学とは直接に関係していない分野の研究者や技術者のために、平衡統計力学の入門的な紹介を行う。その中でも特に大切な概念のひとつである無限次元空間上の確率分布の基礎的な性質について、物理学の知識・自然への関心・数学的な感受性などを仮定せずに説明を行う。

1. まえがき

最近、人工知能や情報工学など自然を直接の研究対象としない学問体系の中でも統計力学と共通の数理が大切な役割を果たすことが知られるようになってきた。このため、学生時代に物理学を履修しなかった人たちからも統計力学を学びたいという声が高まっている。しかしながら、統計力学についての教科書は古典力学、熱力学、解析力学、量子力学などを学んでいることを前提としているものが多く、また、読者が自然現象についての豊富な知識や旺盛な好奇心を持っていることを仮定して書かれているために、物理学に触れたことのない人や自然に関心を持てない人にとっては統計力学の本を読むことは決して容易なことではない[X*X]。そこで、この解説記事では、物理学を学んだことのない人や自然に関心を持たない人に統計力学を学び始めるよりも前に知っておくと良い基礎的な事柄を紹介する。統計力学においては非常に多くの重要な課題があるが、この解説では主に平衡統計力学について考えることにし、「**無限であることは有限であることと何が違うか**」という問題に的をしぼって述べる。なぜなら「無限自由度を持つ系が有限自由度を持つ系と、どのような質的な違いを持つか」という問いかけは、統計力学において第一義的に重要である一方で、人工知能や情報工学においても大切な役割を果たすと思われるからである。

注意. この解説は、物理学に初めて出会う人のために一切の前提知識なしで理解できるものを目標としている。ストーリー展開の中で自然科学に関心のある人のための説明を「注意」の形で掲載するが、全体を理解する上で「注意」の項は必要ではないから、「注意」の中にわからないことが書いてあったとしても気にしないで、どんどん飛ばして読んでいただければ幸いである。「注意」の項は、まったく理解できなくてもこの解説は理解できるはずである。また、物理学を学ばれた方がこの解説を読まれてもよいが、良くご存知のことが復習されているだけであることを予め申し上げておきたい。

2. 空間の次元

人工知能や情報工学を専門としている研究者や技術者であって、初めて統計力学という言葉に出会う人は「**平衡統計力学とは無限次元空間の上の確率分布論の**

* 〒226-8503 横浜市緑区長津田 4259

X*X 言うまでもなく、物理学は自然の美しさを描き出す学問であり、人工知能や情報工学に役立つかどうかということとはまったく独立に、そのものを学ぶこと自体に至上の価値がある。この解説を読んで、本当の統計力学を知りたいと思ってくださった読者は、基礎から十分な時間と労力をかけてゆっくりと学ぶことを推奨する。

ことである」と考えても、第0近似としては誤っていないであろう。すなわち統計力学においては無限次元の空間の上で確率分布を考えると何が起こるか、という数理について考えることが必要になる。そのために、まずは有限次元の場合から考えてみよう。

注意. 実際の物理現象の中の平衡状態は、確率分布ではないというのが現代の物理学者の理解のようである。自然の平衡状態は確率分布ではないにもかかわらず、確率分布だと考えて計算することで自然現象における平衡状態を正しく予言できるという点に平衡統計力学の凄さがある。従って自然科学としての物理学を学ぶ人に対しては「確率分布でない現象が、なぜ確率分布で厳密に正しく予言されてしまうのか、その秘密」について解説するべきところである。しかしながら本論は物理学を学ぶ人のためのものではなく、物理学において現れてくる数理について、わかりやすく解説するものであるから、本当の自然の平衡状態については考えず確率分布を直接考えることにしよう。

2.1 有限次元の空間

実数全体の集合 $\mathbf{R}$ の n 個の直積から作られる集合 $\mathbf{R}^n$ を考えよう。

$$\mathbf{R}^n = \{(x_1, x_2, ..., x_n)\ ;\ \ |x_i| < \infty \quad (\forall i)\}$$

これは、日常生活の中で私たちが「n 次元実空間」と呼んでいるものである。この集合の要素

$$\mathbf{x} = (x_1, x_2, ..., x_n)$$

のノルム（長さ）を

$$\|\mathbf{x}\| = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$$

と定義する。n 次元実空間の中で、ノルムが有限であるもの全体の集合を考えて $\mathbf{E}^n$ と書くことにする。つまり

$$\mathbf{E}^n = \{\mathbf{x} \in \mathbf{R}^n ; \|\mathbf{x}\| < \infty\}$$

であるが、これは実は n 次元実空間と一致する。

$$\mathbf{R}^n = \mathbf{E}^n$$

この等式は、当たり前のことのように感じる読者があるかもしれないが、実は有限次元の空間が持つ極めて特殊な性質である。物理学における実世界は普通は無限次元なので、このように特殊な性質は、自然の世界では、めったに成立しない。

注意. この解説は、数理について解説するものであるにも関わらず、数学的にきちんと整備されたものではない。例えば「考えている集合に位相は定義されているか」とか、「確率分布が作る数列があったとき、それが収束するとは、どのようなことか」というようなことがらは、読者に適切に考えていただくことにしたい。本論は、数学的な定義がなくても直感的に最後までストーリーがつながるようにすることを目的とする。

2.2 無限次元の空間

有限次元の空間においては、ベクトルの長さが有限であることと、ベクトルの各要素が有限であることが等価であるという特別な性質が成り立つことを上で確認したのであるが、無限次元では、どうなるだろうか。n を∞にしてみよう。上の定義に合わせて、二つの集合を考えよう。

$$\mathbf{R}^\infty = \{(x_1, x_2, ..., x_n, ...)\ ;\ \ |x_i| < \infty (\forall i)\}$$

と定義して、この要素の元を

$$\mathbf{x} = (x_1, x_2, ..., x_n, ...)$$

と書き、そのノルムを

$$\|\mathbf{x}\| = \left(\sum_{i=1}^{\infty} |x_i|^2\right)^{1/2}$$

と定義する。ただし、この式が有限でない $\mathbf{x}$ については、$\|\mathbf{x}\| = \infty$ と定義する。集合 $\mathbf{E}^\infty$ を

$$\mathbf{E}^\infty = \{\mathbf{x} \in \mathbf{R}^\infty\ ;\ \ \|\mathbf{x}\| < \infty\}$$

とする。定義から $\mathbf{E}^\infty \subset \mathbf{R}^\infty$ であるが、$\mathbf{E}^\infty \neq \mathbf{R}^\infty$ である。実際、

$$\mathbf{1} = (1, 1, 1, ...) \in \mathbf{R}^\infty$$

であるが、$\|\mathbf{1}\| = \infty$ であるから

$$\mathbf{1} \notin \mathbf{E}^\infty$$

である。$\epsilon > 0$ が非常に小さな値であるとき、点 $(\epsilon, \epsilon, \epsilon, ...)$ は原点 $\mathbf{0} = (0, 0, 0, ...)$ の近くにあるような気がするかもしれないが、実は原点から無限の遠くにある。

注意. 無限次元の空間では、有限次元の空間では起こらないことがたくさん起こる。

1次元空間では、0 と 1 との距離は 1 であり、
2次元空間では、(0,0) と (1,1) との距離は $\sqrt{2}$ であり、
n 次元空間では、(0,0,...,0) と (1,1,...,1) との距離は $\sqrt{n}$ であり、

無限次元空間では、(0,0,0,...,) と (1,1,1,...) とは、無限に離れている。

である。さらに、$\mathbf{R}^\infty$ の元 $\mathbf{x} = (x_1, x_2, ..., x_n, ...)$ を任意に取ってきたとき、数列 x_n は、収束することもあるし発散することもあるし振動することもあるが、$\mathbf{x}$ が $\mathbf{E}^\infty$ の元であるときには、

$$\lim_{n\to\infty} x_n = 0$$

が成り立つ。

無限次元の空間について、もう少し準備を述べる。$\mathbf{x} = (x_1, x_2, x_3, ...)$ に対して $|\mathbf{x}|_*$ を

$$|\mathbf{x}|_* = \mathrm{limsup}_{n\to\infty}\left|\frac{\sum_{i=1}^n x_i}{(n\log\log n)^{1/2}}\right|$$

と定める（後で重複対数の法則を使うからである）。また、

$$\mathbf{a} = (a_1, a_2, ..., a_n, ...) \in \mathbf{R}^\infty$$

に対して、

$$\mathbf{H}(\mathbf{a}) = \{\mathbf{x} \in \mathbf{R}^\infty\ ;\ |\mathbf{x} - \mathbf{a}|_* \le 1\}$$

と定義しよう。$\mathbf{H}(\mathbf{a})$ は線形空間ではないが、凸集合である。

$$|\mathbf{a} - \mathbf{a}'|_* > 2$$

のときには

$$\mathbf{H}(\mathbf{a}) \cap \mathbf{H}(\mathbf{a}') = \text{空集合}$$

であるから、$\mathbf{R}^\infty$ の中には、互いに共通部分を持たない無限個の $\mathbf{H}(\mathbf{a})$ が存在している。

注意. 古典力学に従う n 個の質点の運動は、すべての質点の位置と運動量の次元の空間、すなわち $\mathbf{R}^{2n}$ の中の軌跡として表現できる。この空間の次元は有限であるから、$\mathbf{R}^{2n}$ だと思っても良いし、$\mathbf{E}^{2n}$ であると思っても良い。有限次元の空間を考えているときには、このようなことを心配しなくても良いのである。しかしながら、無限個の質点の運動を考えたいときには事情が異なってくる。物理現象を正しく記述するためには、$\mathbf{R}^\infty$ か $\mathbf{E}^\infty$ のどちらの空間で考えるべきなのだろうか。あるいはもっと別の空間が必要なのだろうか。$\mathbf{R}^\infty$ の方が集合としての制限がないので便利であると感じる読者があるかも知れないが、この集合では一般にベクトルのノルムや内積が有限の値にならないので、自然科学を記述する上で困った問題が生じることがある。一方 $\mathbf{E}^\infty$ は、無限個の質点の運動を考えるには、集合として小さすぎる。非常に優れた物理学者がしばしば「ベクトルの長さが有限か無限かということは本質的な問題ではない」という主張をしている場合があるが、それは、その物理学者が、無限次元空間論を自由自在に使いこなすことができる高いレベルにあることを主張しているだけであって、物理学者ではない私たちは、その主張を言葉通りに受け取ってはならない。なぜなら、物理学者にとって大切な概念である「相転移」や「普遍性」は、無限次元の空間の中で初めて現れてくるものだからである。

3. 確　率　分　布

3.1　有限と無限

ある集合 Ω の上に定義された確率分布のことを**平衡状態** と呼ぶ。集合 Ω の上に定義された確率分布 P があるとき、全体集合の確率は 1 であるから $P(\Omega) = 1$ である。Ω よりも小さい集合 A でも $P(A) = 1$ となることはありえる。そこで、$P(A) = 1$ となる最小の閉集合 A を、確率分布 P の**サポート**と呼ぶことにする。確率分布のサポートとは、直観的には、確率分布が 0 にならないところ全体のことである。

さて、有限次元の空間 $\mathbf{R}^n$ の上に定義された正規分布（平均が 0 で、分散共分散行列が $n \times n$ の単位行列）は、

$$p_n(\mathbf{x}) = p_n(x_1, x_2, ..., x_n) \propto \exp(-\frac{\|x\|^2}{2})$$

である。この確率分布のサポートは $\mathbf{R}^n$ 全体である。つまり、すべての元の近傍が、正の確率を持っている。

次に、この確率分布を用いて無限次元の集合 $\mathbf{R}^\infty$ の上の確率分布を作ってみよう。

$$\mathbf{x} = (x_1, x_2, x_3, ...) \in \mathbf{R}^\infty$$

について確率分布 $p_n^*(\mathbf{x})$ を、

$$p_n^*(\mathbf{x}) = p_n(x_1, x_2, ..., x_n)\ \varphi(x_{n+1}, x_{n+2}, ...,)$$

と定める。ここで、$\varphi(\cdot)$ は、$(x_{n+1}, x_{n+2}, ...,)$ を要素とする無限次元空間上のある確率分布であるとしよう。実際は、無限次元の上の確率分布は確率密度関数を使って書けるわけではないが厳密に述べようとすると手間がかかるし、数学者でない人には返ってわかりにくいように思われるので、ここは直観的に感じとって納得してもらいたい。さて、このとき無限次元空間上の確率分布が作る列

$$p_1^*(\mathbf{x}),\ p_2^*(\mathbf{x}),\ p_3^*(\mathbf{x}),\ \cdots$$

は、無限次元の集合 $\mathbf{R}^\infty$ の上のある確率分布 $p^*(\mathbf{x})$ に収束して、収束先の確率分布は、

$$p^*(\mathbf{x}) \propto \exp(-\frac{\|\mathbf{x}\|^2}{2})$$

となる。この収束先は確率分布 φ には依存しない。この確率分布は、無限次元空間上の正規分布である。この確率分布は $\mathbf{R}^\infty$ の上に定義されてはいるものの、そのサポートは $\mathbf{H}(\mathbf{0})$ に含まれている。(実は $\mathbf{H}(\mathbf{0})$ の中の一部がサポートになっている)。従って $\mathbf{R}^\infty$ の中の、ごく一部分だけが正の確率を持っている。例えば、$(1,1,1,...)$ の近傍の確率は 0 である。また重複対数の法則を思い出せば $(0,0,..,0)$ の近傍の確率も 0 である。$p^*(\mathbf{x})$ を確率密度関数のようなものと考えたとき、この関数が最大になる点は $(0,0,...,0)$ であるにもかかわらず、その近傍が実現される確率は 0 である。「確率密度関数は、あくまでも確率の密度であるにすぎず確率そのものではない」という事実は、無限次元空間を考えるときにはいつでも思い出さなくてはならない。なぜなら、密度関数を最大にする点と最も起こりやすい点とは全然違うからである。この事実が説明を要しないこととして感受できたとき、平衡統計力学が読者の心に正しく美しく感じ取られたのだと考えてよい。

注意. 正規分布のことを、統計力学では「調和振動子の平衡状態」と呼ぶ。互いに独立に伸びたり縮んだりしているバネ振り子が無限個あるとき、その平衡状態が無限次元の正規分布になるからである。空間の中を自由に飛び回る無限個の分子の速度の分布が正規分布であるときも運動量については正規分布になる (位置は正規分布にならないが)。正規分布なら無限次元の空間においても、たいていの量の平均計算を行うことができるし、理論の上では最も基礎的な分布であると言ってよい。なお、場の量子論の用語を借りて、正規分布のことを「自由場」と呼ぶ物理学者も多い。

注意. 無限次元の空間の中に、任意の有限次元の空間の上への射影の意味で無矛盾な確率分布が定義されたとき、無限次元の空間上に対応する確率分布が存在するか、という問いかけは極めて重要である。答えは比較的自然な条件のもとで Yes であり、**コルモゴルフの拡張定理**と呼ばれている。この定理は物理学において極めて重要な定理ものである。

注意. 無限次元空間の上に、なんらかの平行移動のような変換(群)が準備されていて、それに対して不変な確率分布が作れるか、という問題が物理学においてはしばしば大切な問いかけになる。例えば、ミンコフスキー空間の不変性を満たし、かつ、自由場でない量子場は存在するか、という問題が場の量子論において 30 年以上前から問われている(構成論的場の量子論)。数学的にはこの問題は未解決であり、相互作用を持つ量子場は存在しないのではないかという意見もあるという(場の自明性)†。無限次元空間上の確率分布は、人間が好き勝手に作ることができるわけではなく、むしろ、無限次元であることにより極めて強い制約を受けている。正規分布が無限次元空間で定義できるといっても、そこから、わずかにずれただけで、その確率分布は存在しているとは限らないのである。

3.2 無限次元極限の非唯一性

さて、定数ベクトルを

$$\mathbf{1} = (1,1,..,1) \in \mathbf{R}^n$$

として、有限次元の空間 $\mathbf{R}^n$ 上の確率分布

$$q_n(\mathbf{x}) \propto \exp(-\frac{\|\mathbf{x}\|^2}{2}) + \exp(-\frac{\|\mathbf{x}-\mathbf{1}\|^2}{2})$$

を考えよう。この確率分布から、無限次元の空間の上の確率分布を作ってみよう。

$$\mathbf{x} = (x_1, x_2, x_3, ...) \in \mathbf{R}^\infty$$

に対して、

$$q_n^*(\mathbf{x}) = q_n(x_1, x_2, ..., x_n)\ \varphi(x_{n+1}, x_{n+2}, ...)$$

と定義する。$\varphi(\cdot)$ は、さきほどと同じく無限次元空間上のある確率分布である。このとき、確率分布の列

$$q_1^*(\mathbf{x}),\ q_2^*(\mathbf{x}),\ q_3^*(\mathbf{x}),\ \cdots$$

の収束は、どのようになるだろうか。

読者は、ここで以下を読む前に立ち止まって、この確率分布が作る数列の収束先がどのようになるかを、納得いくまで考えてもらいたい。この問題を考えることは答えを聞くより面白いし、この問題と数理的に等価な問題が、統計力学においてはたいへん重要な意味を持っているからである。

† 場の量子論において相互作用を持つ場が存在しないとすると、私たちは実在していないことになるので大変に困ったことである。このような重要な課題が未解決なのである。

さて、この確率分布の列の収束は、$\varphi(\cdot)$ に依存して異なるものになる。以下では、次の 3 通りのケースについて考える。

(1) $q_n^*(\mathbf{x})$ のサポートが $\mathbf{R}^\infty$ 全体であるときの収束。
(2) $q_n^*(\mathbf{x})$ のサポートが $\mathbf{H}(\mathbf{0})$ に含まれているときの収束。
(3) $q_n^*(\mathbf{x})$ のサポートが $\mathbf{H}(\mathbf{1})$ に含まれているときの収束。

上記の３つのケースにおいて収束先が異なる。条件が異なるのであるから、収束先が異なるのは数学的にはまったく無矛盾なことである。しかしながら、このような場合、物理学的に正しく自然現象を記述するためには、どのように考えるべきなのだろうか。また、上の３つの異なるものはすべて自然においても存在する平衡状態と対応するものなのだろうか。この物理学的な問題については後から述べることにして、まず、上にあげた３つのケースについて、（非厳密であるが数学的な気持ちで）それぞれの収束先を調べよう。ケース (2),(3) が、ケース (1) に比べて易しいので、そちらから順に述べる。

ケース (2) 確率分布 $q_n^*(\mathbf{x})$ のサポートが $\mathbf{H}(\mathbf{0})$ に入るように φ を選んだとする。すると $\mathbf{H}(\mathbf{0})$ の中だけで確率分布の収束を考えることができて、

$$q^*(\mathbf{x}) \propto \exp(-\frac{\|\mathbf{x}\|^2}{2})$$

となる。これは、一個の正規分布であり、有限次元において 2 個の正規分布であったにも関わらず無限次元においては一個になってしまった。

ケース (3) 確率分布 $q_n^*(\mathbf{x})$ のサポートが $\mathbf{H}(\mathbf{1})$ に入るように φ を選んだとする。すると $\mathbf{H}(\mathbf{1})$ の中だけで確率分布の収束を考えることができて、

$$q^*(\mathbf{x}) \propto \exp(-\frac{\|\mathbf{x}-\mathbf{1}\|^2}{2})$$

となる。これはケース (2) と同じように 1 個の正規分布であるが、ケース (2) とは違う正規分布になっている。

ケース (1) 確率分布 $q_n^*(\mathbf{x})$ において、集合 $\mathbf{H}(\mathbf{0})$ の要素が選ばれる確率と集合 $\mathbf{H}(\mathbf{1})$ の要素が選ばれる確率が、$a:b$ になるように φ を選んだとする。このとき、

$$q^*(\mathbf{x}) \propto a\ \exp(-\frac{\|\mathbf{x}\|^2}{2}) + b\ \exp(-\frac{\|\mathbf{x}-\mathbf{1}\|^2}{2})$$

となる。これは、二つの正規分布の混合であるが、それぞれの正規分布のサポートは共通部分を持たないことに注意が必要である。これは、限りなく離れた場所にある二つの正規分布の混合になっている。また、混合比 $a:b$ $(a,b>0)$ は、$1:1$ に限定はされていない。

注意. ここで述べた例において、確率分布 φ は、有限次元の外側が、どのようになっているかを表現するものであり、いわば、境界条件のようなものであると考えてよい。$n\to\infty$ において、外側の状況は消えていくように感じる読者があるかもしれないが、無限次元の空間の上の確率分布を考えるとき、外側の影響は一般には消えないことが普通である。物理学的には、例えば、粒子の個数がアボガドロ数くらいに多くても、その外側からの影響は消えないということに対応しているのではないかと思われる。あるいは、スピン系が磁石になるか、鉄になるか、という問題を考えるとき、無限個のスピン系の外側にある磁場の影響が無視できない（無限次元極限を取るときに、その違いによって結果が変わる）ということに対応している。ところで、内分点にならない確率分布のことを、純粋状態 (Pure State) と呼ぶことにしよう。上の例で言えば、ケース (2) とケース (3) が純粋状態である。ケース (1) のような分布は混合状態 (Mixed State) と呼ばれている。混合状態とは、例えば、水と氷とが混ざった状態や、部分的な磁石と部分的な鉄が混ざった状態に対応しているのかも知れない（量子力学的な混合はさらに不思議なものである）。このような微妙な問題では「数理的に導出された状態が対応する自然現象を持たないこともありうるのではないか」と感じるのが人情というものである。しかしながら、実際は、数理的に導出された平衡状態は必ずそれに対応する物理現象が存在している（と思われる）。物理学における数学の有効性は奇跡というべきものである。物理学が数学によって表現されているのではなくて、物理学とは数学そのもののことではないかと感じられるくらいである。実際、どれほど純粋な数学でも自然と関係を持たない数学は現在までひとつも知られていない。また深い数学的構造を持たない自然の法則も現在までひとつも知られていない。おそらく数学と自然はぴったりと一致しているのであり、その一致について浅い理解しか持たない私たち人間が物理学と数学は異なる学問であると思っているだけなのであろう。

注意. 人工知能や情報工学では、普通は有限個の変数

しか考えないから、無限次元の問題は無関係であると思う読者があるかもしれない。それはそうではない。確率分布を扱う問題においては、無限次元のほうから有限次元を近似したほうが、その挙動を正確に捉えられることが少なくないからである。例えば、混合正規分布は統計学や情報理論などでしばしば考察されるが、考えている変数の次元が 10 次元程度以上なら、平均が違う正規分布のサポートは、ほとんど共通部分を持っていないということが十分に配慮されるべきであろう。人工知能や情報工学の研究者にとっては、有限次元は有限次元であり、無限次元ではない、と思ってしまいやすいところであるが、確率分布を考えるとき、ある程度以上に次元が高ければ、それはむしろ無限次元であると考えた方が良い場合が少なくない。

注意. 計算機を用いて無限次元空間上の確率分布を実現することは、統計力学においては非常に重要な問題である。後で述べる相転移は、無限次元において初めて見ることができるものであるが、有限次元のシミュレーションに基づいて無限次元の場合を実質的に作り出す方法についてたいへん良く研究が行われている。

4. 分 配 関 数

以上で、有限次元の空間の上の確率分布から無限次元の空間の上の確率分布に移行するとき、サポートがどうなっているのかが重要であることを述べた。以下では、サポートは $\mathbf{R}^\infty$ 全体である場合を考えることにする。もちろん、収束先の確率分布のサポートはその中の一部分になることが普通である。

空間 $\mathbf{R}^n$ 上の確率分布

$$\begin{aligned} q_n(\mathbf{x}) \;=\; \frac{1}{Z(n)}\Big\{ & \; n\;\exp(-\frac{\|\mathbf{x}\|^2}{2}) \\ & +\exp(-\frac{\|\mathbf{x}-\mathbf{1}\|^2}{4})\Big\} \end{aligned}$$

を考えよう。この確率分布は、二つの正規分布の混合で作られた分布であるが、混合比が $n:1$ であり、また二つの正規分布は分散が異なっている。$Z(n)$ は、$q_n(\mathbf{x})$ の $\mathbf{x}$ についての積分が 1 になるような正規化定数である。

$$\begin{aligned} Z(n) \;=\; \int\Big\{ & \; n\;\exp(-\frac{\|\mathbf{x}\|^2}{2}) \\ & +\exp(-\frac{\|\mathbf{x}-\mathbf{1}\|^2}{4})\Big\}d\mathbf{x} \end{aligned}$$

この $Z(n)$ のことを**分配関数**という。この確率分布を最大にする $\mathbf{x}$ すなわち、

$$\mathrm{argmax}\; q_n(\mathbf{x})$$

は n が大きくなると原点 $\mathbf{0}$ に近づくが、この $q_n(\mathbf{x})$ から作られる $q_n^*(\mathbf{x})$ も $\mathbf{H}(\mathbf{0})$ にサポートが含まれる確率分布に近づくと考えてよいのだろうか。

分配関数 $Z(n)$ を計算してみよう。すると

$$Z(n) = n\;(2\pi)^{n/2} + (4\pi)^{n/2}$$

となる。この右辺の二つの項は、混合されている二つの確率分布のそれぞれが実現される確率に比例している。すなわち $Z(n)$ は、確率が割り振られる割合を表している。$n\to\infty$ となるとき、右辺の第 1 項は、第 2 項と比べて圧倒的に小さいから

$$Z(n) \cong (4\pi)^{n/2}$$

であり、混合された二つの分布のうち、第 2 項が確率的にほとんどを占めている。$q_n(\mathbf{x})$ を最大にする $\mathbf{x}$ は第一項の側に存在するが、第一項は確率的には消えてしまう項である。すなわち $q_n^*(\mathbf{x})$ は無限次元空間の上で

$$q^*(\mathbf{x}) \propto \exp(-\frac{\|\mathbf{x}-\mathbf{1}\|^2}{4})$$

に収束する。分配関数を計算することで、無限次元上の確率分布において、主要な項になる部分が何かを調べることができるのである。

このことは、次のように言い換えることもできる。

$$F(n) = -\log Z(n)$$

を自由エネルギーと呼ぶことにする。また、部分的な自由エネルギーをそれぞれ

$$\begin{aligned} f_1(n) \;&=\; -\log\int\; n\;\exp(-\frac{\|\mathbf{x}\|^2}{2})d\mathbf{x} \\ f_2(n) \;&=\; -\log\int\exp(-\frac{\|\mathbf{x}\|^2}{4})d\mathbf{x} \end{aligned}$$

と定義すれば、

$$F(n) = -\log\Big(e^{-f_1(n)} + e^{-f_2(n)}\Big)$$

が成り立つが、この式は実質的には

$$F(n) = \min\Big(f_1(n), f_2(n)\Big)$$

と同じであることに注意しよう。そこで

$$f_1(n) >> f_2(n)$$

であるから $n \to \infty$ において、$F(n)$ の主要項を与えているのは $f_2(n)$ の方である。このことを、「物理学的に実現されるのは、自由エネルギーを最も小さくする状態である」と言う。これに対して $\arg\max\ q_n(\mathbf{x})$ は、「エネルギーが最も小さい状態」に対応している。無限次元の確率分布を考えるときには、一般に、**「最もエネルギーが小さい状態」と「最も自由エネルギーが小さい状態」とはまったく異なる。** 両者は互いに無限に離れた距離にあって、少しも似ていない。自然において観測されるのは、自由エネルギーが小さい状態である。

注意. エネルギーと自由エネルギーの本質的な差異は、物理学者にとっては言葉にするまでもない極めて当然で合理的なことである。しかしながら、一方で、情報学や統計学を専門とする人にとっては、決して理解できない魔術のように感じられるところでもある。人類の歴史は、統計物理学と統計情報学とを、まったく別々の分野として発展させてしまった。このため、大学においても物理学科で習う統計物理学と、情報学科で習う統計情報学とは正反対の学問であるかのように認識されている。これは本当に残念なことである。地球以外の星の生物が確率論の勉強をしているとしたら、統計情報学と統計物理学は同じひとつの科目で教えられているのではないだろうか。

5. 相　転　移

今度は、次のような確率分布を考えてみよう。$\alpha > 0$ をパラメータとして、

$$q_n(\mathbf{x}) \propto \exp(-\frac{\|\mathbf{x}\|^2}{2}) + \exp(-\frac{\|\mathbf{x}-\mathbf{1}\|^2}{2\alpha})$$

を考える。これまで述べてきたこととまったく同様にして、この確率分布から無限次元の空間上の確率分布の列 $q_n^*(\mathbf{x})$ $(n = 1, 2, 3, ...)$ を作ることができる。その収束先を $q^*(\mathbf{x})$ としよう。すると、これまで述べてきたことからわかるようにパラメータ α に依存して収束先は異なっていて、

(1) $0 < \alpha < 1$ のとき

$$q^*(\mathbf{x}) = \exp(-\frac{\|\mathbf{x}\|^2}{2})$$

(2) $\alpha = 1$ のとき

$$q^*(\mathbf{x}) = a\exp(-\frac{\|\mathbf{x}\|^2}{2}) + b\exp(-\frac{\|\mathbf{x}-\mathbf{1}\|^2}{2\alpha})$$

(3) $\alpha > 1$ のとき

$$q^*(\mathbf{x}) = \exp(-\frac{\|\mathbf{x}-\mathbf{1}\|^2}{2\alpha})$$

となる。特に、$\alpha = 1$ の前後で $q^*(\mathbf{x})$ は、ガラッと変わってしまうだけでなく、そのサポートが $\mathbf{H}(\mathbf{0})$ に含まれるものから $\mathbf{H}(\mathbf{1})$ に含まれるものへと、無限の距離を移動してしまうことに注意しよう。このような現象のことを**相転移**という。有限次元の空間の上の確率分布 $q_n(\mathbf{x})$ では、相転移は見られないが、無限次元の空間に移行すると、パラメータに対する依存性が解析性を失ってしまうのである。このような現象が、決して特殊なものではなく、むしろ、極めてしばしば見られるということが大切な点である。

物理学においては、様々な物質が温度の変化とともに質的な変化を見せるようなケースが実験的に観察可能であって、その変化が理論的な観点から正しく予測できるかどうかという問題が研究されている。情報学や統計学においては、符号の復号問題などで、ある程度以上に雑音が多くなると復号が急激にできなくなるようなケースや、あるいは、ベイズ推測においてサンプル数が増えるときに、事後分布の実質的なサポートが急激に変化するケースなどが相転移現象である。

注意. 相転移を考えるときのパラメータとして最もよく現れるのは温度である。これは確率分布 $q(\mathbf{x})$ に対して $q(\mathbf{x})^\alpha$ のようにコントロールを行ったときの応答を考える問題に相当する。もちろん、温度以外のパラメータの変化によって相転移が起こることも多い。

注意. 自然においては様々な原子や分子が存在するのであるが、それらが数多く集まってできる無限次元の空間の上の確率分布やその確率分布が持つ相転移は、個々の原子や分子が何であるかには依存しないものであることが知られている。すなわち、系を構成している個々のものの性質よりも無限次元の空間の上の確率分布が持つ数理の方が自然の現象を定めているのである。このことを**普遍性** (universality) という。無限次元の上の確率分布を考えるとき、有限次元の世界よりも遥かに強い数理的な制約条件によって、存在できる確率分布は強く限定されているのである。情報学や統計学においても、数多くの変数が相互作用することによって引き起こされる現象を考察するときには、普遍性の概念を念頭におくことによって気づかれる事柄がありうるのではないだろうか。

注意. 自然における現象が、個々の物質ではなく、無限次元の空間の上の確率分布の数理によって定められているという事実は、同時に、その数理では説明のできない現象を見つけたいという気持ちを物理学者に抱かせるものであったかもしれない。1970 年頃、物理学者は「複雑なものを複雑なまま描き出したい」という願いに基づく研究も行っている。普遍性という強力なテーゼと、それに対するアンチテーゼという試練を経て現代の物理学への発展がなされてきたのである。

6. 平均場近似

先に、「分配関数を計算すると、無限次元空間の上の確率分布が解析できる」と述べたが、残念なことに分配関数を数学的に厳密に計算できるのは「可解模型」という特別に不思議な性質を持つものに限られている。分配関数が計算できない場合に「正しい結果を与える保証はないが、とにかく、まずは計算してみるべき方法」として**平均場近似**がある。

ある確率分布 $p(\mathbf{x})$ を作って、これによって様々な量の平均を計算したいのであるが、例え有限次元の空間であっても確率分布を作ることは決して容易なことではない。そこで、実現可能であると思われる確率分布の集合

$$Q = \{q(\mathbf{x})\}$$

を考えて、この集合の中で最も $p(\mathbf{x})$ に近いものを見つけて代理をさせることを考えてみよう。例えば、相対エントロピー

$$\int q(\mathbf{x}) \log \frac{q(\mathbf{x})}{p(\mathbf{x})} d\mathbf{x}$$

は、どんな確率分布 $p(\mathbf{x})$, $q(\mathbf{x})$ に対しても非負であり、これが 0 になるのは二つの確率分布が等しいときに限られるから、与えられた確率分布 $p(\mathbf{x})$ に対して、確率分布の集合 Q の中から相対エントロピーを最小にするものを見つけ出してくればよい。一般に、上の相対エントロピーを用いて最適化を行うと、$q(\mathbf{x})$ のサポートは、$p(\mathbf{x})$ のサポートに対して局所化されてしまう。このため、無限次元空間の上で平均場近似を考えると、平均場近似のサポートは正しい確率分布よりも真に小さい集合になることが原因で、正しい確率分布では相転移を持たない場合でも平均場近似では相転移を持つことがしばしば観察されている。

注意. 物理学において、たとえばスピン系の解析における平均場近似は、確率分布の集合としてすべてのスピンが独立であるもの

$$q(\mathbf{x}) = q_1(x_1)\, q_2(x_2)\, q_3(x_3) \cdots$$

が用いられている。近年、ベイズ事後分布に平均場近似を適用すると、EM アルゴリズムを改良した学習則を与えることが提案され、変分ベイズ法という名称で知られるようになった。

7. おわりに

情報の符号化・情報圧縮・情報の学習や予測などの分野において、実質的に無限次元の確率分布を扱う必要性が高くなってきた。そのような分野においても、関数空間上の確率分布はよく研究されているが、統計力学で見られるような無限次元の荒々しさが現れてくるケースは、まだ十分には研究されてこなかったようである。これから未来に向かって、非常に多くの変数を扱う情報システムの構築において、無限次元確率論が役立つ場面が増えていくのではないだろうか。

参考文献

1) 荒木不二洋、“統計物理の数理,” 岩波書店, 2004.
2) 荒木不二洋, “量子場の数理,” 岩波書店, 2001.
3) 黒田耕嗣、樋口保成、“統計力学-相転移の数理-,” 培風館,2006.
4) 服部哲弥、“ランダムウォークとくりこみ群,” 共立出版,2004.
5) 田崎晴明, “統計力学,” 近刊, 2007.
6) 原隆, “ 相転移と臨界現象の解明 −パイエルス，オンサガーからウィルソンを経て,” 別冊数理科学, “現代の数理物理学,”サイエンス社, 1998.
7) Ola Bratteli , D. W. Robinson,“ Operator Algebras and Quantum Statistical Mechanics 1: C*- And W*- Algebras Symmetry Groups Decomposition of States,” Springer, 2002.
8) 久保亮五、“統計力学,” 共立出版, 1952.
9) ランダウ、リフシッツ、“統計物理学,” (訳) 岩波書店, 1980.