数理解析研究所講究録
第 1916 巻 2014 年 156-171

一般化エントロピに基づく統計力学の拡張とその情報幾何構造

茨城大学・工学部・電気電子工学科　和田 達明 (Tatsuaki WADA)
Department of Electrical and Electronic Engineering, Ibaraki University

はじめに

統計力学の拡張へのひとつのアプローチとして、従来の Boltzmann-Gibbs-Shannon(BGS) エントロピーをパラメータ拡張した一般化エントロピに基づいた拡張について紹介し、関連する熱力学的ポテンシャル関数の間に成立する Legendre 変換構造と情報幾何構造について解説する。

1　一般化エントロピに基づく統計力学の拡張

ここで紹介する一般化熱統計学は、Callen による熱統計学 [1] へ更に一般化エントロピを導入し、その最大エントロピ原理に基づく統計力学の拡張である。そのメリットとして、一般化エントロピを最大化する状態は、従来の平衡統計力学における熱平衡状態（指数型分布）に限らず、より広いクラスの熱的状態（非指数型分布）を取り扱うことを可能とする。

ここでは、歴史的な発展順序とは逆になるが、より広い一般的な概念である Naudts によって創られた一般化熱統計学 [2] を先に説明し、そのひとつの具体例として κ-エントロピに基づく熱統計学について説明する。

1.1　一般化熱統計学 (generalized thermostatistics)

Callen はその著書 [1] において、熱力学 (thermodynamics) に対比して、熱統計学 (thermostatistics) という枠組みを導入した。この熱統計学は力学や電磁気学のように特定の力に対する動的応答などを扱う理論ではなく、当時の物理理論の範疇外であった情報理論におけるエントロピを最大化する状態を" 熱平衡状態"[*1] として特徴付ける、エントロピ最大原

[*1] 一般化エントロピを（適切な条件の下で）最大化する状態が、ここで言う"熱平衡状態" であり、必ずしも通常の平衡統計力学における熱平衡状態とは限らず、非平衡定常や準熱平衡状態である場合をも含んでいる。

理に基づく熱力学的理論体系である。

Naudts が導入した一般化熱統計学 [2] は、Callen による熱統計学を、パラメータ拡張したエントロピ（一般化エントロピ）に基づいて一般化した熱統計学である。与えられた正の実増加関数 $\phi(s)$ により、対数関数の拡張（ϕ-対数と呼ばれる）を

$$\ln_\phi(x) \equiv \int_1^x \frac{1}{\phi(s)} ds, \quad \underset{\phi(s)\to s}{\longrightarrow} \quad \ln(x). \tag{1}$$

で導入する。これより、$\phi(s) = s$ とすれば、ϕ-対数関数は通常の対数関数に帰着することが分かる。一般化熱統計学では、この ϕ-対数関数を用いて従来の BGS-エントロピの表式を模した

$$S_\phi = -\int dx\, p(x)\, \ln_\phi p(x) \quad \underset{\phi(s)\to s}{\longrightarrow} \quad S^{\mathrm{BGS}}, \tag{2}$$

という形の一般化エントロピに基づいた熱統計学である。

一般化熱統計学の具体例としては、Tsallis[3] と Kaniadakis[4] がそれぞれ独立に提案した、互いに異なるパラメータ拡張である一般化エントロピに基づいた枠組みが知られている。Tsallis 熱統計学 [3] については、幾つかの解説や須鎗氏による書籍 [5] もあるので、ここでは簡単に概略を説明する。この場合の ϕ-対数関数としては、

$$\phi_q(s) = s^q, \quad \ln_q(x) = \frac{x^{1-q} - 1}{1-q}, \quad q > 0. \tag{3}$$

という、Tsallis の q-対数関数である。一般化エントロピとしては、次の Tsallis エントロピを導入する。

$$S_q \equiv \int dx \left(\frac{p(x)^q - p(x)}{1-q} \right) \quad \underset{q\to 1}{\longrightarrow} \quad S^{\mathrm{BGS}} = -\int dx p(x) \ln p(x), \tag{4}$$

ここで q は実数パラメータで、S_q を適切な条件の下で最大化する状態は、q-対数関数の逆関数である次の q-指数関数

$$\exp_q(x) \equiv [1 + (1-q)x]^{\frac{1}{1-q}} \tag{5}$$

により表される漸近的ベキ型分布となる。長距離相互作用が働いている自己重力系 [6] や純電子プラズマにおける熱的分布 [7] や、光学格子中の原子の運動量分布 [8] を代表とする異常拡散におけるいわゆるロングテールを持つ分布などを Tsallis エントロピ最大原理に基づいて説明できると考えられている。

1.2 κ-エントロピに基づく熱統計学

イタリアのトリノ工科大学の Giorgio Kaniadakis[4] が提案した κ エントロピ S_κ は、先の Tsallis エントロピ S_q とは異なる一般化エントロピであり、実数パラメータ κ を用いて

$(-1 < \kappa < 1)$

$$S_\kappa[p] \equiv -\int dx\, p(x) \ln_\kappa p(x), \quad \underset{\kappa \to 0}{\longrightarrow} S^{\mathrm{BGS}}, \tag{6}$$

と表せる。ここで κ-対数関数とその逆関数である κ-指数関数は、それぞれ

$$\ln_\kappa(x) \equiv \frac{x^\kappa - x^{-\kappa}}{2\kappa}, \quad \underset{\kappa \to 0}{\longrightarrow} \ln(x), \tag{7}$$

$$\exp_\kappa(x) \equiv \left(\kappa x + \sqrt{1 + \kappa^2 x^2}\right)^{\frac{1}{\kappa}}, \quad \underset{\kappa \to 0}{\longrightarrow} \exp(x). \tag{8}$$

と定義され、$\kappa \to 0$ の極限において、それぞれ通常の対数関数と指数関数に帰着する。先の一般化熱統計学における ϕ-対数関数 (1) の定義式の $\phi(s)$ として

$$\phi_\kappa(s) = \frac{2s}{s^\kappa + s^{-\kappa}}, \tag{9}$$

とすれば、κ-対数関数 (7) となる。

次に、エネルギー $V(x)$ の期待値を

$$U[p] \equiv \int dx\, p(x) V(x), \tag{10}$$

と記すと、S_κ に関する最大エントロピ法

$$\frac{\delta}{\delta p}\left(S_\kappa[p] - \beta U[p] - \gamma \int p(x) dx\right) = 0, \tag{11}$$

の解である最適分布は、

$$p_\kappa^{\mathrm{ME}}(x) = \alpha_\kappa \exp_\kappa\left[-\frac{1}{\lambda_\kappa}\left(\gamma + \beta V(x)\right)\right], \quad \underset{V(x)>>1}{\longrightarrow} V(x)^{\frac{1}{\kappa}}, \tag{12}$$

と求まり、漸近的ベキ則に従う確率分布である。β と γ はそれぞれ、エネルギー期待値および確率分布の規格化条件

$$\int dx\, p(x) = 1, \tag{13}$$

に関する Lagrange 未定乗数である。また、α_κ と λ_κ は κ に依存する定数で、それぞれ

$$\alpha_\kappa = \left(\frac{1-\kappa}{1+\kappa}\right)^{\frac{1}{2\kappa}}, \quad \lambda_\kappa = \sqrt{1-\kappa^2}. \tag{14}$$

であり、両者の間には

$$\alpha_\kappa = \exp_\kappa\left(-\frac{1}{\lambda_\kappa}\right), \tag{15}$$

の関係がある。

κ-エントロピに基づく拡張の場合には、κ-対数関数 (7) に加えて、

$$u_\kappa(x) \equiv \frac{x^\kappa + x^{-\kappa}}{2} \quad \underset{\kappa \to 0}{\longrightarrow} \quad 1. \tag{16}$$

で定義される κ-ユニット関数が重要である。両者は相補的な変形関数であり、双曲線三角関数を用いてそれぞれ

$$\ln_\kappa x = \frac{1}{\kappa} \sinh\Big(\kappa \ln x\Big), \quad u_\kappa(x) = \cosh\Big(\kappa \ln x\Big), \tag{17}$$

と表すことができる。また、以下の関係式

$$\frac{d}{dx}\Big(x \ln_\kappa(x)\Big) = \lambda_\kappa \ln_\kappa \left(\frac{x}{\alpha_\kappa}\right), \tag{18}$$

$$\frac{d}{dx}\Big(x\, u_\kappa(x)\Big) = \lambda_\kappa u_\kappa \left(\frac{x}{\alpha_\alpha}\right), \tag{19}$$

をそれぞれ満たす。前出の 2 つの定数 α_κ と λ_κ は、これらの関係式を満たすように定めてある。

この枠組みにおける Legendre 構造は文献 [9] において調べられ、例えば、パラメータ κ で拡張された Massieu ポテンシャル Φ_κ

$$\Phi_\kappa(\beta) \equiv \mathcal{I}_\kappa + \gamma, \quad \text{with} \quad \mathcal{I}_\kappa \equiv \int_{-\infty}^{\infty} dx\, \frac{1}{2}\Big[\, \big(p_\kappa^{\mathrm{ME}}(x)\big)^{1+\kappa} + \big(p_\kappa^{\mathrm{ME}}(x)\big)^{1-\kappa} \Big]. \tag{20}$$

に対して、

$$\Phi_\kappa = S_\kappa - \beta U, \quad \frac{d}{d\beta}\Phi_\kappa = -U, \tag{21}$$

という熱力学的関係式 (Legendre 関係式) が成立する。また、κ-拡張された Massieu ポテンシャルと自由エネルギー F_κ とは、

$$\Phi_\kappa = -\beta F_\kappa, \tag{22}$$

の関係で結ばれている。以下、これらの Legendre 関係式の導出を順次みていこう。

先ず、最適分布 (12) より、

$$-\lambda_\kappa \ln_\kappa \left(\frac{p_\kappa^{\mathrm{ME}}}{\alpha_\kappa}\right) = \gamma + \beta, \tag{23}$$

と変形できて、関係式 (18) を利用すると、

$$\begin{aligned}
\frac{dS_\kappa}{d\beta} &= -\int dx\, \frac{dp_\kappa^{\mathrm{ME}}}{d\beta} \frac{d}{dp_\kappa^{\mathrm{ME}}}\Big(p_\kappa^{\mathrm{ME}} \ln_\kappa p_\kappa^{\mathrm{ME}}\Big) = -\int dx\, \frac{dp_\kappa^{\mathrm{ME}}}{d\beta} \lambda_\kappa \ln_\kappa \left(\frac{p_\kappa^{\mathrm{ME}}}{\alpha_\kappa}\right) \\
&= \int dx\, \frac{dp_\kappa^{\mathrm{ME}}}{d\beta}\,(\gamma + \beta V(x)) = \gamma \frac{d}{d\beta}\int dx\, p_\kappa^{\mathrm{ME}} + \beta \frac{d}{d\beta}\int dx\, p_\kappa^{\mathrm{ME}} V(x) \\
&= \beta \frac{d}{d\beta}\int dx\, p_\kappa^{\mathrm{ME}} V(x) = \beta\, \frac{dU}{d\beta}.
\end{aligned} \tag{24}$$

となることが分かるので、関係式

$$\frac{dS_\kappa}{dU} = \beta, \tag{25}$$

が確かめられる。次に、恒等式

$$\lambda_\kappa \ln_\kappa \left(\frac{x}{\alpha_\kappa}\right) = \ln_\kappa x + u_\kappa(x), \tag{26}$$

と、式 (23) とを組み合わせて

$$\ln_\kappa p_\kappa^{\mathrm{ME}} + u_\kappa \left(p_\kappa^{\mathrm{ME}}\right) + \gamma + \beta V(x) = 0, \tag{27}$$

を得る。この両辺に p_κ^{ME} を乗じて積分すると、

$$\Phi_\kappa = \mathcal{I}_\kappa + \gamma = S_\kappa - \beta U, \tag{28}$$

が成立することが分かる。κ-拡張された Legendre 構造をまとめると、

$$\Phi_\kappa(\beta) = S_\kappa(U) - \beta\, U, \quad \frac{d}{d\beta}\Phi_\kappa(\beta) = -U, \quad \frac{d}{dU}S_\kappa(U) = \beta. \tag{29}$$

のようになる。これらは、標準的な場合 ($\kappa \to 0$ の極限に相当) と同じ関係式を満たしていることが分かる。

1.3 Fokker-Planck 方程式の κ-拡張版

熱平衡状態を記述する従来の熱統計学や統計力学を支える根本原理に関して、系の動的な振る舞い記述する拡散方程式や Fokker-Planck(FP) 方程式

$$\frac{\partial}{\partial t} p(x, t) = \frac{\partial}{\partial x}\left\{p(x,t)\frac{\partial}{\partial x}\left[V(x) + D\,\ln\left(p(x,t)\right)\right]\right\}, \tag{30}$$

の果たす役割が重要であるように、κ という実数パラメータで拡張された κ-熱統計学に対して重要な役割を果たす非線型拡散方程式や非線型 FP 方程式 [10] が存在する。

FP 方程式の κ-拡張版 [11, 12] は、以下のような非線形拡散項

$$\frac{\partial}{\partial x}\left[p(x,t)\frac{\partial}{\partial x}\lambda_\kappa \ln_\kappa\left(\frac{p(x,t)}{\alpha}\right)\right] \underset{\kappa\to 0}{\longrightarrow} \frac{\partial}{\partial x}\left[p(x,t)\,\frac{\partial}{\partial x}\ln p(x,t)\right] = \frac{\partial^2}{\partial x^2}\,p(x,t) \tag{31}$$

を導入することで得られる。また、対応する κ-拡張された FP 方程式は、

$$\frac{\partial}{\partial t} p(x, t) = \frac{\partial}{\partial x}\left\{p(x,t)\frac{\partial}{\partial x}\left[V(x) + D\,\lambda_\kappa\,\ln_\kappa\left(\frac{p(x,t)}{\alpha}\right)\right]\right\}, \tag{32}$$

と表せる。ここで D は拡散定数であり、$\kappa \to 0$ の極限において、通常の FP 方程式 (30) に帰着する。

この非線型 FP 方程式の停留状態

$$\frac{\partial}{\partial t} p_\kappa^{\mathrm{st}}(x) = 0 \tag{33}$$

を求めるために、式 (32) を上式に代入し、両辺を積分して、積分定数を C とすると、

$$V(x) + D\lambda_\kappa \ln_\kappa\left(\frac{p_\kappa^{\mathrm{st}}(x)}{\alpha_\kappa}\right) = C, \tag{34}$$

を得る。これより

$$p_\kappa^{\mathrm{st}}(x) = \alpha_\kappa \exp_\kappa\left[-\frac{1}{\lambda_\kappa}\Big(-\frac{C}{D} + \frac{1}{D}V(x)\Big)\right] \tag{35}$$

であることが判る。ところで、先に求めた最大エントロピ法による解 (12) と比較すると、

$$C = -\frac{\gamma}{\beta} \quad \text{and} \quad \frac{1}{D} = \beta \tag{36}$$

という対応が成立する。つまり停留解 $p^{\mathrm{st}}(x)$ は、κ エントロピを最大化する最適解 $p_\kappa^{\mathrm{ME}}(x)$ において $\beta = 1/D$ と置いたものに等しいのである。

先の κ-拡張版の非線型 FP 方程式 (32) に従って $p(x,t)$ が時間発展する時に、Lyapunov 関数

$$\mathcal{L}_\kappa(t) \equiv U[p] - D\, S_\kappa[p]\ , \tag{37}$$

の時間依存性を計算してみると、

$$\begin{aligned}
\frac{d\mathcal{L}_\kappa(t)}{dt} &= \int_{-\infty}^{\infty} dx\, \frac{\partial p}{\partial t} \cdot \frac{\delta}{\delta p}\Big(U[p] - DS_\kappa[p]\Big) \\
&= \int_{-\infty}^{\infty} dx \left\{\frac{\partial}{\partial x}\, p\, \frac{\partial}{\partial x}\left[V(x) + D\,\lambda_\kappa\, \ln_\kappa\left(\frac{p}{\alpha_\kappa}\right)\right]\right\} \times \left[V(x) + D\lambda_\kappa \ln_\kappa\left(\frac{p}{\alpha_\kappa}\right)\right] \\
&= -\int_{-\infty}^{\infty} dx\, p \left[\frac{\partial}{\partial x}\Big(V(x) + D\lambda_\kappa \ln_\kappa\left(\frac{p}{\alpha_\kappa}\right)\Big)\right]^2 \le 0,
\end{aligned} \tag{38}$$

となるので、$\mathcal{L}_\kappa(t)$ は時間発展に伴って非増加関数であることが判る。

ここまでをまとめると [12]

- κ-拡張版の非線型 FP 方程式の解の時間発展に伴い、常に非増加となる Lyapunovu 関数 $\mathcal{L}_\kappa(t)$ が存在する。
- 時刻無限大の極限である停留状態は、κ-エントロピ S_κ を最大とする最適分布 (12) に等しい。

$$p^{\mathrm{st}}(x) = \lim_{t\to\infty} p(x,t) = p_\kappa^{\mathrm{ME}}(x) \tag{39}$$

- 停留状態において Lyapunov 関数は最小となり、その点で κ-拡張版の自由エネルギ $F_\kappa = -\Phi_\kappa/\beta$ も最小値となり、対応する一般化エントロピの最大値と一致する。

$$\min \mathcal{L}_\kappa(t) = \mathcal{L}_\kappa^{\min} = U[p_\kappa^{\mathrm{ME}}] - D\, S_\kappa[p_\kappa^{\mathrm{ME}}] = F_\kappa[p_\kappa^{\mathrm{ME}}] \quad \Leftrightarrow \quad \max S_\kappa \tag{40}$$

となる。これらのことは、$\kappa \to 0$ の極限においてそれぞれ通常の統計力学の場合に帰着し、時間発展を記述するのは線型 FP(30) で、その停留分布は指数型分布であり、通常のエントロピ S^{BGS} を（適切な条件下において）最大化する状態である。

1.4 Bregman divergence による κ-エントロピ最大法の表現

確率論や情報理論における divergence とは、一般にある確率分布関数 p と参照分布関数 r との間の異なる度合いを表す非負の量として定義され、

$$\begin{cases} \text{非負性：} & D(p||r) \geq 0 \\ D(p||r) = 0 & \text{の時のみ}\ p = r \end{cases} \tag{41}$$

を満たす $D(p||r)$ を divergence と呼ぶ。一般に $D(p||r)$ は、確率分布関数 p と r の交換に対して非対称である。良く知られている divergence は、Kullback-Leibler(KL) divergence

$$D^{\mathrm{KL}}(p||r) = \int dx\, p(x) \ln\left(\frac{p(x)}{r(x)}\right), \tag{42}$$

や、その凸関数 f による一般化である Csiszár f-divergence

$$D^{\mathrm{C}}(p||r) = \int dx\, p(x)\, f\left(\frac{p(x)}{r(x)}\right), \tag{43}$$

で、情報理論や統計力学を含め様々な分野で多くの応用がなされている。しかしながら、通常の対数関数の性質 $\ln(x/y) = \ln(x) - \ln(y)$ が q-対数関数や κ-対数関数等の拡張された ϕ-対数関数では一般に成立しないことから、これらの変形関数に基づく一般化熱統計学の枠組みと良く整合するのは、次の Bregman divergence である。

確率分布関数 $p(x)$ と $r(x)$ に対する Bregmann divergence(BD)[13] は、

$$D^{\mathrm{B}}(p||r) \equiv \int_{-\infty}^{\infty} dx \Big[f\big(p(x)\big) - f\big(r(x)\big) - \Big(p(x) - r(x)\Big) f'\big(r(x)\big) \Big], \tag{44}$$

と定義される。ここで $f(x)$ は凸関数で、$f'(x)$ は f の微分を表す。

次に、先の κ-拡張版の自由エネルギ最小原理を、この BD の枠組みで捉え直してみる。凸関数 f と、参照分布 r として、それぞれ

$$f(p) = p \ln_\kappa(p), \quad r(x) = p_\kappa^{\mathrm{ME}} = \alpha_\kappa \exp_\kappa[-\frac{1}{\lambda_\kappa}\Big(\gamma + \beta V(x)\Big)] \tag{45}$$

を採用して、以下の時間依存性を持つ BD

$$D_\kappa^{\mathrm{B}}\Big(p(t)\,||\,p_\kappa^{\mathrm{ME}}\Big) \equiv \int_{-\infty}^{\infty} dx \Big[p(x,t)\ln_\kappa\left(p(x,t)\right) - p_\kappa^{\mathrm{ME}}(x)\ln_\kappa(p_\kappa^{\mathrm{ME}}(x)) + \Big(p(x,t) - p_\kappa^{\mathrm{ME}}(x)\Big)\Big(\beta V(x) + \gamma\Big)\Big] \tag{46}$$

を導入する。関係式 (18) より、

$$f'\left(p^{\mathrm{ME}}(x)\right) = \lambda_\kappa \ln_\kappa\left(\frac{p^{\mathrm{ME}}(x)}{\alpha_\kappa}\right) = -\gamma - \beta V(x), \tag{47}$$

であることに注意すると、

$$\begin{aligned} D_\kappa^{\mathrm{B}}\Big(p(t)\,||\,p_\kappa^{\mathrm{ME}}\Big) &= \beta U[p(t)] - S_\kappa[p(t)] - \beta U[p_\kappa^{\mathrm{ME}}] + S_\kappa[p_\kappa^{\mathrm{ME}}] \\ &= \beta\Big(F_\kappa[p(t)] - F_\kappa[p_\kappa^{\mathrm{ME}}]\Big) = \beta\Big(\mathcal{L}_\kappa(t) - \mathcal{L}_\kappa^{\min}\Big), \end{aligned} \tag{48}$$

となる。この表式から、divergence の非負性により、Lyapunov 関数の最小値は $\mathcal{L}_\kappa^{\min}$ であり、そこで自由エネルギ $F_\kappa(t)$ が最小となることが良く判る。

2 情報幾何構造

情報幾何 [14] は、確率分布属のなす多様体 (統計多様体) 上の微分幾何構造を基に創られ、不変な Rieman 計量である Fisher 計量と、互いに双対である Affine 接続からなる幾何構造を持つ。ここでは、情報幾何について後に必要となる幾つかの基本概念について説明する。

簡単のため、離散確率分布に対する次の統計モデル：

$$\mathcal{S}^n \equiv \left\{\mathbf{p} = (p_\mu) | p_\mu > 0, \sum_{\mu=0}^{n} p_\mu = 1\right\}. \tag{49}$$

を考える。次式で定義される自然基底接ベクトル：

$$\partial_i \equiv \frac{\partial}{\partial p_i} - \frac{\partial}{\partial p_0}, \quad i = 1, \cdots, n. \tag{50}$$

を導入し、$\mathcal{S}^n$ 上の Riemannian 計量として、 Fisher 計量:

$$\left(g^F(\mathbf{p})\right)_{ij} \equiv \frac{1}{p_i}\delta_{ij} + \frac{1}{p_0} = \sum_{\mu=0}^{n} p_\mu \left(\partial_i \ln p_\mu\right)\left(\partial_j \ln p_\mu\right), \quad i,j,k = 1, \cdots, n. \tag{51}$$

を採用する。この設定において、α-接続は、

$$\nabla_{\partial_i}^{(\alpha)} \partial_j = \sum_{k=1}^{n} \Gamma_{ij}^{(\alpha)k}, \quad i,j = 1, \cdots, n, \tag{52}$$

$$\Gamma_{ij}^{(\alpha)k}(\mathbf{p}) = \frac{1+\alpha}{2}\left(-\frac{1}{p_k}\delta_{ij}^k + p_k\,(g^F)_{ij}\right). \tag{53}$$

と表される。$\alpha = \pm 1$ の場合は、**指数型接続** $\nabla^{(e)} = \nabla^{(1)}$、**混合型接続** $\nabla^{(m)} = \nabla^{(-1)}$ とそれぞれ呼ばれている。

また、α-divergence は、

$$D^{\alpha}[\mathbf{p}:\mathbf{r}] = \frac{4}{1-\alpha^2}\sum_{\mu=0}^{n}\left\{\left(\frac{1-\alpha}{2}\right)p_{\mu} + \left(\frac{1+\alpha}{2}\right)r_{\mu} - p_{\mu}^{\frac{1-\alpha}{2}} r_{\mu}^{\frac{1+\alpha}{2}}\right\}. \tag{54}$$

と表される。

統計モデル、計量、接続の 3 つ組 $(\mathcal{S}, g^{\mathrm{F}}, \nabla^{(\alpha)})$ を**統計多様体**と呼び、良く調べられているのは次の指数型分布族

$$\mathcal{S}_{exp} \equiv \left\{ p(x;\boldsymbol{\theta}) \middle| p(x;\boldsymbol{\theta}) = \exp\left[\sum_{i=0}^{n}\theta^i c_i(x) - \Psi(\boldsymbol{\theta})\right], \right\}, \tag{55}$$

に対する統計多様体 $(\mathcal{S}_{exp}, g^{\mathrm{F}}, \nabla^{(e)}, \nabla^{(m)})$ である。指数型分布族 (55) におけるパラメータ $\{\theta^i\}$ は、この統計多様体上の $\nabla^{(e)}$-affine 座標系を構成し、これに双対な $\nabla^{(m)}$-affine 座標系 $\{\eta_i\}$ の各成分は、

$$\eta_i = \int dx p(x;\boldsymbol{\theta}) c_i(x), \tag{56}$$

であり、$c_i(x)$ の期待値に等しい。また、これら互いに双対な affine 座標系にはポテンシャル関数が存在し、

$$\eta_i = \frac{\partial \Psi(\boldsymbol{\theta})}{\partial \theta^i}, \qquad \theta^i = \frac{\partial \Psi^{\star}(\boldsymbol{\eta})}{\partial \eta_i}, \quad i = 1, \cdots, n, \tag{57}$$

とそれぞれ表され、互いに Legendre 双対なポテンシャル

$$\Psi(\boldsymbol{\theta}) = \boldsymbol{\theta}\cdot\boldsymbol{\eta} - \Psi^{\star}(\boldsymbol{\eta}), \tag{58}$$

である。Fisher 計量は、これらの互いに双対なポテンシャルと、

$$\left(g^{\mathrm{F}}(\boldsymbol{\theta})\right)_{ij} = \frac{\partial^2}{\partial\theta^i\partial\theta^j}\Psi(\boldsymbol{\theta}), \quad \left(g^{\mathrm{F}}(\boldsymbol{\eta})\right)^{ij} = \frac{\partial^2}{\partial\eta_i\partial\eta_j}\Psi^{\star}(\boldsymbol{\eta}), \tag{59}$$

の関係で結ばれている。

このようにポテンシャル関数 $\Psi(\boldsymbol{\theta})$ から affine 座標ベクトル $\boldsymbol{\eta}$ が導出されるということは、$\boldsymbol{\eta}$ が $\boldsymbol{\theta}$ に関する微分に関して渦なし場であるということである。

$$\nabla_{\theta} \times \boldsymbol{\eta} = 0. \tag{60}$$

同様にして、双対 affine 座標ベクトル $\boldsymbol{\theta}$ も渦なし場

$$\nabla_{\eta} \times \boldsymbol{\theta} = 0, \tag{61}$$

であるので、ポテンシャル関数 $\Psi^{*}(\boldsymbol{\eta})$ が存在する。

2.1 q-指数型分布族の統計多様体と α-表現

一般化エントロピの中でも基本的であり比較的良く研究されているのは、Tsallis エントロピに関する分野である。Tsallis エントロピに基づく統計力学の拡張と、情報幾何学との関連性が甘利、小原、松添 [15, 16, 17] によって示された。そこで重要な鍵となったのは、情報幾何学における α-表現を用いて、q-指数型分布族

$$\mathcal{S}_{\exp_q} \equiv \left\{ p(x;\boldsymbol{\theta}) \middle| p(x;\boldsymbol{\theta}) = \exp_q \left[\sum_{i=0}^{n} \theta^i c_i(x) - \Psi_q(\boldsymbol{\theta}) \right], \right\}, \tag{62}$$

に対する統計多様体を構成することができたことであろう。これは、例えば $\alpha = 1 - 2q$ と対応付けると、

$$\begin{aligned} p_\mu &\mapsto \ell^{\alpha}(p_\mu) = \frac{2}{1-\alpha} p_\mu^{\frac{1-\alpha}{2}} = \frac{p_\mu^q}{q} \\ p_\mu &\mapsto \ell^{-\alpha}(p_\mu) = \frac{2}{1+\alpha} p_\mu^{\frac{1+\alpha}{2}} = \frac{p_\mu^{1-q}}{1-q} = \ln_q p_\mu + \frac{1}{1-q} \end{aligned} \tag{63}$$

のように、α-表現と q-対数関数とが旨く整合していることに起因している。それは、両者ともに単一のベキ指数で特徴付けられた関数（or 写像）であるからである。

しかしながら、式 (7) の κ-対数関数 $\ln_\kappa p_\mu$ は 2 つのベキ指数 $\pm\kappa$ で特徴付けられた関数であり、単一のベキ指数しか持たない α-表現と結びつけるのは恐らく不可能であろう。よって、κ-指数型分布族から統計多様体を構成するためには、α-表現とは異なるタイプの表現が適していると思われる。

Zhang [18] は、以下に示す**共役表現**を導入した。先ず、ある単調関数 $\rho(p_\mu)$ に対して、確率 p_μ の ρ-表現を写像 $p_\mu \mapsto \rho(p_\mu)$ で定める。ある滑らかで凸な関数 f が存在して、

$$\begin{aligned} \tau(p_\mu) &= f'(\rho(p_\mu)) = ((f^\star)')^{-1}(\rho(p_\mu)) \\ \rho(p_\mu) &= (f')^{-1}(\tau(p_\mu)) = (f^\star)'(\tau(p_\mu)), \end{aligned} \tag{64}$$

が成立する時に、τ-表現 $p_\mu \mapsto \tau(p_\mu)$ がこの f に関して ρ-表現と共役であると言う。f' は f の引数に関する微分を表し、 f^{-1} は f の逆関数を表す。凸関数 f と $f^\star$ は、互いに Legendre 双対である。

$$f(\rho) = \rho\tau(\rho) - f^\star(\tau(\rho)) \quad \Leftrightarrow \quad f^\star(\tau) = \rho(\tau)\tau - f(\rho(\tau)). \tag{65}$$

この共役表現に関する Bregman ダイバージェンスは、

$$D_{f,\rho}[\mathbf{p}:\mathbf{r}] = \sum_\mu \Big[f(\rho(p_\mu)) - f(\rho(r_\mu)) - f'(\rho(r_\mu))\big(\rho(p_\mu) - \rho(r_\mu)\big) \Big], \tag{66}$$

$$D_{f^*,\tau}[\mathbf{p}:\mathbf{r}] = \sum_\mu \Big[f^*(\tau(p_\mu)) - f^*(\tau(r_\mu)) - (f^*)'(\tau(r_\mu))\big(\tau(p_\mu) - \tau(r_\mu)\big) \Big], \tag{67}$$

である。

α-表現は、以下に示すように共役表現の一例である。

$$\rho(p_\mu) = \ell^{\alpha}(p_\mu), \qquad \tau(p_\mu) = \ell^{-\alpha}(p_\mu), \tag{68}$$

$$f_\alpha(\rho) = \frac{2}{1+\alpha}\left(\frac{1-\alpha}{2}\rho\right)^{\frac{2}{1-\alpha}}, \quad f^{\star}_\alpha(\tau) = \frac{2}{1-\alpha}\left(\frac{1+\alpha}{2}\tau\right)^{\frac{2}{1+\alpha}}. \tag{69}$$

これらの写像と凸関数 f_α により、α-ダイバージェンス (54) は $D_{f_\alpha,\rho}[\mathbf{p}:\mathbf{r}]$ と表される。

さて、この共役表現を用いて、統計力学における確率分布関数に対する統計多様体とその情報幾何構造をみていこう。先ず $\mathcal{S}^n$ を n-次元の単体上において、標準的な統計力学における正準分布は、

$$p_j = \frac{1}{Z}\exp(-\beta E_j) = \exp(-\beta E_j - \ln Z), \quad j = 1, 2, \cdots, n, \tag{70}$$

と書き直せる。 $p_0 = 1 - \sum_{j=1}^n p_j$ は確率の規格化条件を満たすための従属変数であり、分配関数と

$$\ln p_0 = -\ln Z = \ln\left(\frac{1}{1 + \sum_{j=1}^n \exp(-\beta E_j)}\right), \tag{71}$$

の関係がある。この場合の共役表現は、

$$\rho(p_\mu) = \ln p_\mu, \qquad f(\rho) = \exp(\rho), \tag{72}$$

$$\tau(p_\mu) = f'\left(\rho(p)\right) = p_\mu, \qquad f^{\star}(\tau) = \tau\ln\tau - \tau. \tag{73}$$

である。パラメータ θ^i とポテンシャル関数 $\Psi(\boldsymbol{\theta})$ をそれぞれ、

$$\theta^i \equiv \rho(p_i) - \rho(p_0) = \ln p_i - \ln p_0 = -\beta E_i, \quad i = 1, 2, \cdots, n \tag{74}$$

$$\Psi(\boldsymbol{\theta}) \equiv -\ln p_0 = \ln Z. \tag{75}$$

で導入すると、正準分布 (70) は

$$p_i = \exp\left(\theta^i - \Psi(\boldsymbol{\theta})\right), \quad i = 1, 2, \cdots, n, \tag{76}$$

$$p_0 = 1 - \sum_{i=1}^n p_i = \frac{1}{Z}, \tag{77}$$

と表される。確率 p_μ の規格化条件より、

$$\begin{aligned} 0 &= \frac{\partial}{\partial\theta^i}\sum_{\mu=0}^n p_\mu = \frac{\partial}{\partial\theta^i}\sum_{j=1}^n p_j + \frac{\partial p_0}{\partial\theta^i} = \sum_{j=1}^n \left(\delta_{ij} - \frac{\partial\Psi(\boldsymbol{\theta})}{\partial\theta^i}\right)p_j - \frac{\partial\Psi(\boldsymbol{\theta})}{\partial\theta^i}p_0 \\ &= p_i - \frac{\partial\Psi(\boldsymbol{\theta})}{\partial\theta^i}\sum_{\mu=0}^n p_\mu = p_i - \frac{\partial\Psi(\boldsymbol{\theta})}{\partial\theta^i}. \end{aligned} \tag{78}$$

よって、

$$\eta_i = \frac{\partial\Psi(\boldsymbol{\theta})}{\partial\theta^i} = p_i = \tau(p_i). \tag{79}$$

を得る。双対ポテンシャル $\Psi^\star(\boldsymbol{\eta})$ は、Legendre 変換によって求められて、

$$\begin{aligned}\Psi^\star(\boldsymbol{\eta}) &= \sum_{i=1}^{n}\theta^i\eta_i - \Psi(\boldsymbol{\theta}) = \sum_{i=1}^{n}(\ln p_i - \ln p_0)\,p_i + \ln p_0\\ &= \sum_{\mu=0}^{n}(\ln p_\mu - \ln p_0)\,p_\mu + \ln p_0 = \sum_{\mu=0}^{n} p_\mu \ln p_\mu = -S.\end{aligned} \tag{80}$$

と (負の) エントロピである。

次に Legendre 構造を調べよう。互いに Legendre 双対な関数の組が 3 つある。先ず、共役表現における $f(\rho)$ と $f^*(\tau)$; 熱統計学におけるエントロピ S と分配関数 $\ln Z$; そして、情報幾何におけるポテンシャル関数である $\Psi(\boldsymbol{\theta})$ と $\Psi^\star(\boldsymbol{\eta})$ である。これら 3 つの Legendre 双対関数の組は、互いにどのような関係にあるのだろうか？既に良く調べられている標準の場合には、

$$\Psi(\boldsymbol{\theta}) = \ln Z = -\ln p_0, \tag{81}$$

$$\Psi^\star(\boldsymbol{\eta}) = -S = \sum_\mu \tau(p_\mu)\ln\tau(p_\mu) = \sum_\mu f\Big(\tau(p_\mu)\Big) + 1. \tag{82}$$

ところで、式 (65) より

$$\sum_\mu f\big(\rho(p_\mu)\big) = \sum_\mu \rho(p_\mu)\tau(p_\mu) - \sum_\mu f^\star\big(\tau(p_\mu)\big). \tag{83}$$

この両辺から $\rho(p_0)\sum_\mu \tau(p_\mu)$ を引いて、(74) の関係を利用すると、

$$\begin{aligned}\sum_\mu f\big(\rho(p_i)\big) - \rho(p_0)\sum_\mu \tau(p_\mu) &= \sum_\mu \big(\rho(p_\mu) - \rho(p_0)\big)\tau(p_\mu) - \sum_\mu f^\star\big(\tau(p_i)\big)\\ &= \sum_{i=1}^{n}\theta^i\,\tau(p_i) - \sum_\mu f^\star\big(\tau(p_\mu)\big).\end{aligned} \tag{84}$$

更に、$\sum_\mu \tau(p_\mu) = \sum_\mu p_\mu = 1$ であるので、

$$\sum_i \theta^i\,\eta_i = \sum_i f\big(\rho(p_i)\big) - \rho(p_0) + \sum_\mu f^\star\big(\tau(p_\mu)\big), \tag{85}$$

を得る。この関係式と式 (80) の最初の関係式を比較すると、結局

$$\Psi(\boldsymbol{\theta}) + \Psi^\star(\boldsymbol{\eta}) = \sum_\mu f\big(\rho(p_\mu)\big) - \rho(p_0) + \sum_\mu f^\star\big(\tau(p_\mu)\big), \tag{86}$$

であることが分かる。この関係式から、Ψ と $\Psi^\star$ の和しか定まらないことに注意しよう。従って、右辺の 3 項をどのように Ψ と $\Psi^\star$ へ割り当てるかの自由度が存在する。先の通常の場合には、(75) に示したように $\Psi = -\rho(p_0) = -\ln p_0$ と選んだので、残りの 2 項の和が $\Psi^\star$ に等しいはずであり、実際

$$\sum_\mu f\left(\rho(p_\mu)\right) + \sum_\mu f^\star\left(\tau(p_\mu)\right) = \sum_\mu p_\mu + \sum_\mu \left(p_\mu \ln p_\mu - p_\mu\right) = -S = \Psi^\star(\boldsymbol{\eta}), \qquad (87)$$

である。

2.2 κ-表現

次の κ-指数型分布族

$$\mathcal{S}_{\exp_\kappa} \equiv \left\{ p(x;\boldsymbol{\theta}) \middle| p(x;\boldsymbol{\theta}) = \alpha_\kappa \exp_\kappa \left[\frac{1}{\lambda_\kappa} \left(\sum_{i=0}^{n} \theta^i c_i(x) - \gamma(\boldsymbol{\theta}) \right) \right], \right\}. \qquad (88)$$

に対して統計多様体を構成するするのに適切な以下の表現を、ここでは κ-表現 [19] と呼ぶことにする。

$$\tau(p_\mu) = p_\mu, \qquad f^\star_\kappa(\tau) = \tau \ln_\kappa \tau. \qquad (89)$$

$$\rho(p_\mu) = \lambda_\kappa \ln_\kappa \left(\frac{p_\mu}{\alpha_\kappa}\right), \quad f_\kappa(\rho) = \alpha_\kappa \exp_\kappa \left(\frac{\rho}{\lambda_\kappa}\right) u_\kappa \left[\alpha_\kappa \exp_\kappa \left(\frac{\rho}{\lambda_\kappa}\right)\right]. \qquad (90)$$

ポテンシャル関数としては、

$$\Psi^\star_\kappa = \sum_\mu f^\star_\kappa(\tau(p_\mu)) = \sum_\mu p_\mu \ln_\kappa p_\mu = -S_\kappa. \qquad (91)$$

と選び、

$$\theta^i(\boldsymbol{p}) = \rho(p_i) - \rho(p_0) = -\beta E_i, \quad i = 1, 2, \cdots, n, \qquad (92)$$

とする。これより、

$$p_i = \alpha_\kappa \exp_\kappa \left[\frac{1}{\lambda_\kappa}(\theta^i - \gamma)\right], \quad p_0 = \alpha_\kappa \exp_\kappa \left[-\frac{\gamma}{\lambda_\kappa}\right], \qquad (93)$$

である。

恒等式

$$\frac{d}{dx} \exp_\kappa(x) = \frac{\exp_\kappa(x)}{u_\kappa\left(\exp_\kappa(x)\right)}, \qquad (94)$$

と、確率分布 p_μ の規格化条件より、

$$0 = \frac{\partial}{\partial \theta^i} \sum_{\mu=0}^{n} p_\mu \frac{\partial}{\partial \theta^i} \sum_{j=1}^{n} p_j + \frac{\partial p_0}{\partial \theta^i} = \frac{1}{\lambda_\kappa} \sum_{\mu=0}^{n} \left(\delta_{i\mu} - \frac{\partial \gamma}{\partial \theta^i}\right) \frac{p_\mu}{u_\kappa\left(\frac{p_\mu}{\alpha_\kappa}\right)}, \qquad (95)$$

となるので、

$$\frac{\partial \gamma}{\partial \theta^i} = \frac{\frac{p_i}{u_\kappa\left(\frac{p_i}{\alpha_\kappa}\right)}}{\sum_{\mu=0}^{n} \frac{p_\mu}{u_\kappa\left(\frac{p_\mu}{\alpha_\kappa}\right)}}, \quad i = 1, 2, \cdots, n, \tag{96}$$

を得る。次の関係式

$$\frac{\partial p_j}{\partial \theta^i} = \frac{p_i}{\lambda_\kappa u_\kappa\left(\frac{p_i}{\alpha_\kappa}\right)}\left(\delta_{ij} - \frac{\partial \gamma}{\partial \theta^i}\right), \quad i, j = 1, 2, \cdots, n, \tag{97}$$

$$\frac{\partial p_0}{\partial \theta^i} = -\frac{p_0}{\lambda_\kappa u_\kappa\left(\frac{p_0}{\alpha_\kappa}\right)}\frac{\partial \gamma}{\partial \theta^i}, \quad i = 1, 2, \cdots, n, \tag{98}$$

を用いると、

$$\begin{aligned}\frac{\partial}{\partial \theta^i}\sum_{\mu=0}^{n} p_j u_\kappa(p_\mu) &= \sum_{j=1}^{n} \frac{\partial p_j}{\partial \theta^i}\lambda_\kappa u_\kappa\left(\frac{p_j}{\alpha_\kappa}\right) + \frac{\partial p_0}{\partial \theta^i}\lambda_\kappa u_\kappa\left(\frac{p_0}{\alpha_\kappa}\right) \\ &= \sum_{j=1}^{n} p_j\left(\delta_{ij} - \frac{\partial \gamma}{\partial \theta^i}\right) - p_0\frac{\partial \gamma}{\partial \theta^i} = p_i - \frac{\partial \gamma}{\partial \theta^i},\end{aligned} \tag{99}$$

が分かるので、結局ポテンシャル関数

$$\Psi(\boldsymbol{\theta}) = \sum_{\mu=0}^{n} p_\mu u_\kappa(p_\mu) + \gamma = \mathcal{I}_\kappa + \gamma, \tag{100}$$

と、

$$\eta_i = \frac{\partial}{\partial \theta^i}\Psi(\boldsymbol{\theta}) = p_i, \quad i = 1, 2, \cdots, n, \tag{101}$$

が確かめられる。この双対となる関係式も以下のように容易に確かめられて、

$$\frac{\partial}{\partial \eta_i}\Psi^\star(\boldsymbol{\eta}) = \frac{\partial}{\partial \eta_i}\sum_{\mu=0}^{n} p_\mu \ln_\kappa p_\mu \tag{102}$$

$$\begin{aligned}&= \sum_{j=1}^{n} \frac{\partial p_j}{\partial p_i}\lambda_\kappa \ln_\kappa\left(\frac{p_j}{\alpha_\kappa}\right) + \frac{\partial p_0}{\partial p_i}\lambda_\kappa \ln_\kappa\left(\frac{p_0}{\alpha_\kappa}\right) \\ &= \rho(p_i) - \rho(p_0) = \theta^i, \quad i = 1, 2, \cdots, n,\end{aligned} \tag{103}$$

を得る。

この κ-表現に対応する Bregman ダイバージェンスは、

$$D_\kappa[\boldsymbol{p} : \boldsymbol{r}] = \sum_\mu \left[p_\mu u_\kappa(p_\mu) - r_\mu u_\kappa(r_\mu) - r_\mu\left(\lambda_\kappa \ln_\kappa\left(\frac{p_\mu}{\alpha_\kappa}\right) - \lambda_\kappa \ln_\kappa\left(\frac{r_\mu}{\alpha_\kappa}\right)\right)\right], \tag{104}$$

$$D_\kappa^*[\boldsymbol{p} : \boldsymbol{r}] = \sum_\mu \left[p_\mu \ln_\kappa p_\mu - r_\mu \ln_\kappa r_\mu - \lambda_\kappa \ln_\kappa\left(\frac{r_\mu}{\alpha_\kappa}\right)(p_\mu - r_\mu)\right]. \tag{105}$$

となる。これと式 (46)、式 (47) とを比較すると、先に紹介した (48) と等価な関係であることが分かる。

おわりに

一般化エントロピに基づく統計力学の拡張として、Naudts による一般化熱統計学の考え方について先ず説明し、その具体例として Kaniadakis 等が提案した κ-エントロピに基づく拡張である κ-熱統計学の基礎と、関連する Legendre 構造、非線型 Fokker-Planck 方程式について説明した後に、κ-指数型確率分布族に対する統計多様体を構成し、その情報幾何構造について論じた。

謝辞

本研究を遂行する上で有益な助言や議論をして頂いた、名古屋工業大学の松添 博氏、およびトリノ工科大学の Antinio M. Scarfone 博士に感謝します。本研究は JSPS 科研費 25400188 の助成を受けたものです。

参考文献

[1] H.B. Callen, *Thermodynamics and an Introduction to Thermostatics* 2nd ed. (Wiley New York 1985)

[2] J. Naudts, *Generalized Thermostatistics*, (Springer 2011).

[3] C. Tsallis, *Introduction to Nonextensive Statistical Mechanics: Approaching a Complex World* (Springer-Verlag New York 2009).

[4] G. Kaniadakis, Physica A **296** (2001) 405; G. Kaniadakis, A.M. Scarfone, Physica A **305** (2002) 69; G. Kaniadakis, Phys. Rev. E **66** (2002) 056125; G. Kaniadakis, PRE **72** (2005) 036108.

[5] 須鎗弘樹: 複雑系のための基礎数理 – べき乗則とツァリスエントロピーの数理 – (牧野書店 2010).

[6] A. Taruya, M. Sakagami, Physica A **307** 185-206 (2002).

[7] B. M. Boghosian, Phys. Rev. E **53**, 4754 (1996).

[8] E. Lutz, Phys. Rev. A **67**, 051402 (2003).

[9] A.M. Scarfone, T. Wada, Prog. Theor. Phys. Suppl. **162** (2006) 45.

[10] T.D. Frank, *Nonlinear Fokker-Planck Equations*, Springer (2005).

[11] T. Wada, A.M. Scarfone, AIP Conference Proceedings **965** (2007) 177-180.

[12] T. Wada, A.M. Scarfone, Eur. Phys. J. B **70**, 65-71 (2009).

[13] L.M. Bregman, USSR Comp. Math. Phys., **7** (1967) 200.

[14] S-I. Amari, H. Nagaoka, *Methods of Information Geometry,* AMS Mathematical Monographs (Oxford Univ. Press 2000).

[15] H. Matsuzoe, A. Ohara, Recent Progress in Differential Geometry and Its Related Fields, World Sci. Publ., (2011), 55-71.

[16] S-I. Amari, A. Ohara and H. Matsuzoe, Physica A., **391**, 4308-4319, (2012).

[17] A. Ohara, H. Matsuzoe and S-I. Amari, Mod. Phys. Let. B, **26** No. 1250063, (2012).

[18] J. Zhang, Neural Comp. **16**, 159, (2004).

[19] A.M. Scarfone, T. Wada, submitted to J. Phys. A: Math. Theor.