原著論文

エントロピーを導入した ファジィ c-平均法の統計力学的解釈†

安田 真*1・古橋 武*2・大熊 繁*3

本論文の目的は、これまでファジィクラスタリングと統計力学との対応関係に関して得られた知見について総括するとともに、新たに得られた Bose-Einstein 分布型の帰属度関数を用いた正規化の制約のない可能性クラスタリングの枠組みを示し、そのクラスタリングメカニズムの解釈を行うことである。

データのクラスターへの帰属度の評価尺度として Shannon エントロピーを導入したファジィ c-平均 (FCM) 法は帰属度関数として Maxwell-Boltzmann (あるいは Gibbs) 分布を、また、ファジィエントロピーを導入した FCM 法は Fermi-Dirac 分布と同じ帰属度関数を与える。この対応関係で、ファジィクラスタリングにおける正規化の制約は、統計力学的な全粒子数一定の制約に相当する。

さらに、物理的には粒子数が保存されない状態も存在し、この状態を記述する分布関数が Bose-Einstein 分布として知られている。そこで、統計力学的とのアナロジーから、この Bose-Einstein 分布型の帰属度関数を求めてそのクラスタリングアルゴリズムを提案することで、正規化の制約のないファジィクラスタリングの可能性を示した。

キーワード：ファジィ c-平均法, ファジィエントロピー, 可能性クラスタリング, 統計力学, Maxwell-Boltzmann 分布, Fermi-Dirac 分布, Bose-Einstein 分布

1. はじめに

統計力学は非常に多数の要素からなるマクロな物理系の性質を調べる理論であるが、近年、情報処理の分野に統計力学的な方法を導入する試みが盛んとなり、情報統計力学と呼ばれる一分野を形成している [1, 2]。情報統計力学では主としてイジングスピン系のモデル [3, 4] や緩和過程 [5, 6] が扱われている。一方、ファジィ理論においては、特にファジィクラスタリングと統計力学的な分布関数との間に深い関係が見られる。本論文の目的は、ファジィクラスタリングと統計力学との対応関係に関してこれまでに得られた知見について総括し、さらに、統計力学とのアナロジーから新たに得られた Bose-Einstein 分布型の帰属度関数に着目した正規化の制約のない可能性クラスタリングの枠組みを示し、そのクラスタリングメカニズムの解釈を行うことである。

クラスタリングはデータに内在する構造を明らかにして分類する手法であり、データ解析の重要なツールである。実世界における多くのデータ分布ではクラスターの境界は必ずしも明確ではなく、クリスプなクラスタリングよりもファジィクラスタリングの方が適していると考えられる。ファジィクラスタリング手法としては、ファジィ c-平均 (FCM) 法 [7] が標準的な手法として知られている。FCM 法では個々のデータの各クラスターへの帰属度の総和を 1 に規格化する確率的な制約が課される。

Li, Mukaidono と Miyamoto [8, 9] はデータのクラスターへの帰属度の評価尺度として Shannon エントロピーを導入した FCM 法の改良手法 (エントロピー正則化法) を提案したが、彼らとは独立に Rose ら [10] が示していた通り、これは古典統計の Maxwell-Boltzmann (あるいは Gibbs) 分布に他ならなかった。さらに、筆者ら [11] は、FCM 法にファジィエントロピー [12] を導入することで、帰属度関数として量子統計の 1 つである Fermi-Dirac 分布と同じ関数形が得られることを明らかにした。このとき、ファジィクラスタリングにおける正規化の制約は、統計力学的な全粒子数一定の制約に対応すると考えられる [13]。

しかしながら、物理的には粒子の生成・消滅が随時起こって粒子数が保存されない状態も存在し、この状態を記述する分布関数が、もう 1 つの量子統計である Bose-Einstein 分布として知られている。そこで、この

† Entropy based Fuzzy c-means Clustering: Analogy with Statistical Mechanics

Makoto YASUDA, Takeshi FURUHASHI and Shigeru OKUMA

*1 岐阜工業高等専門学校電気情報工学科

Dept. of Electrical and Computer Engineering, Gifu National College of Technology

*2 名古屋大学工学研究科計算理工学専攻

Dept. of Computational Science and Engineering, Nagoya University

*3 名古屋大学工学研究科電気工学専攻

Dept. of Electrical Engineering, Nagoya University

Bose-Einstein 分布型の帰属度関数を求めることで、正規化の制約のないファジィクラスタリングが実現できる。FCM 法に確率的な制約を課さないファジィクラスタリングは可能性クラスタリング[14]として知られており、可能性クラスタリングの範疇でも様々な計算手法が提案されている[15]。Bose-Einstein 分布型の帰属度関数もその一つであるが、このように統計力学的な分布関数とファジィクラスタリングとの対応を取る利点は、(1)統計力学的なエントロピーからファジィクラスタリングのエントロピーを算出可能なこと、(2)統計力学的な観点からクラスタリングのメカニズムの理解が可能なことと、(3)統計力学的な手法に基づく解析や新たなクラスタリング手法の開発が可能となると期待されることである。この中、本論文では(1)と(2)について述べる。

2. FCM 法

p 次元の実空間 R^p のデータ集合 $X = \{x_1, \dots, x_n\}$ ($x_k = (x_k^1, \dots, x_k^p)$) を c 個のクラスターに分類する。 $V = \{v_1, \dots, v_c\}$ ($v_i = (v_i^1, \dots, v_i^p)$) をクラスター中心、 $u_{ik} \in [0, 1]$ ($i = 1, \dots, c; k = 1, \dots, n$) を帰属度関数とする。

標準的な FCM 法ではクラスタリングの目的関数を

$$J = \sum_{k=1}^n \sum_{i=1}^c (u_{ik})^m d_{ik} \quad (1)$$

$$(m > 1, d_{ik} = \|x_k - v_i\|^2)$$

とする。正規化の制約

$$\sum_{i=1}^c u_{ik} = 1 \quad \forall k \quad (2)$$

のもとで u_{ik} , v_i の最適解を求めるために、 η_k を未定乗数とする Lagrange 関数

$$L = J - \sum_{k=1}^n \eta_k \left(\sum_{i=1}^c u_{ik} - 1 \right) \quad (3)$$

を考えると、 $\partial L / \partial u_{ik} = 0$ から

$$u_{ik} = \left[\sum_{j=1}^c \left(\frac{d_{ik}}{d_{jk}} \right)^{\frac{1}{m-1}} \right]^{-1}$$

$\partial L / \partial v_i = 0$ から

$$v_i = \frac{\sum_{k=1}^n (u_{ik})^m x_k}{\sum_{k=1}^n (u_{ik})^m}$$

を得る[7, 16]。

3. Maxwell-Boltzmann 分布

3.1 ファジィクラスタリングにおける Maxwell-Boltz-

mann 分布

エントロピー正則化法では FCM 法の目的関数(1)式で $m=1$ とし、FCM 法の Lagrange 関数(3)式に Shannon (あるいは Maxwell-Boltzmann) エントロピー

$$S^{MB} = - \sum_{k=1}^n \sum_{i=1}^c u_{ik} \log u_{ik} \quad (4)$$

を導入する。新たな Lagrange 関数は

$$L^{MB} = J - \lambda S^{MB} - \sum_{k=1}^n \eta_k \left(\sum_{i=1}^c u_{ik} - 1 \right) \quad (5)$$

となり、 $\partial L^{MB} / \partial u_{ik} = 0$ は

$$u_{ik}^{MB} = \frac{e^{-\beta d_{ik}}}{\sum_{j=1}^c e^{-\beta d_{jk}}} \quad \left(\beta = -\frac{1}{\lambda} \right) \quad (6)$$

$\partial L^{MB} / \partial v_i = 0$ は

$$v_i = \frac{\sum_{k=1}^n u_{ik} x_k}{\sum_{k=1}^n u_{ik}} \quad (7)$$

を与える[10]。

3.2 Maxwell-Boltzmann 分布の統計力学的解釈

(6)式で β^{-1} を系の仮想的な温度 T とみなせば、これは統計力学における Maxwell-Boltzmann 分布に一致する¹

すなわち、前節の内容は統計力学的なアナロジーで捉え直すことができる。

粒子系の取り得る状態を l , そのエネルギーを ϵ_l とするとき、

$$Z^{MB} = \sum_l e^{-\beta \epsilon_l} \quad (8)$$

は状態和、あるいは分配関数と呼ばれる。また、Helmholtz の自由エネルギーはエネルギー $\mathcal{E}^{MB}$, エントロピー $\mathcal{S}^{MB}$ と温度 T から

$$\mathcal{F}^{MB} = \mathcal{E}^{MB} - T \mathcal{S}^{MB} \quad (9)$$

で定義される。Helmholtz の自由エネルギーは物質の熱力学的性質を規定する特性関数の一種でエネルギーの次元を持ち、分配関数から

$$\mathcal{F}^{MB} = -T \log Z^{MB}$$

と計算される。同様に $\mathcal{E}^{MB}$ は

¹ 統計力学における β と T の関係は $\beta^{-1} = k_B T$ (k_B は Boltzmann 定数)である。しかし、ここでは簡単のため、 $k_B = 1$ とする。

$$\mathcal{E}^{\text{MB}} = \frac{1}{Z^{\text{MB}}} \sum_l \epsilon_l e^{-\beta \epsilon_l}$$

と求まる. 平均粒子数 $\langle n_l^{\text{MB}} \rangle$ は

$$\mathcal{E}^{\text{MB}} = \sum_l \epsilon_l \langle n_l^{\text{MB}} \rangle$$

の関係から

$$\langle n_l^{\text{MB}} \rangle = \frac{e^{-\beta \epsilon_l}}{\sum_k e^{-\beta \epsilon_k}}$$

となる [17]. このように, 分配関数が分かれば他の物理量も計算できる.

一方, (8) 式に対応するファジィクラスタリングの分配関数は, d_{ik} が ϵ_l に対応し, かつ, データ x_k はすべてのクラスターに属し得るので

$$Z^{\text{MB}} = \prod_{k=1}^n \sum_{i=1}^c e^{-\beta d_{ik}}$$

と書ける. すると, ファジィクラスタリングの Helmholtz の自由エネルギーは

$$\begin{aligned} F^{\text{MB}} &= -T \log Z^{\text{MB}} \\ &= -\frac{1}{\beta} \sum_{k=1}^n \log \left(\sum_{i=1}^c e^{-\beta d_{ik}} \right) \end{aligned}$$

となる. 熱平衡状態では Helmholtz の自由エネルギーは極小値をとる [18] ので, クラスター中心は $\partial F^{\text{MB}} / \partial v_i = 0$ として求められ, 結果は前出の (7) 式と一致する. また, エネルギーは

$$E^{\text{MB}} = \sum_{k=1}^n \sum_{i=1}^c \frac{d_{ik} e^{-\beta d_{ik}}}{\sum_{j=1}^c e^{-\beta d_{jk}}}$$

となる.

Maxwell-Boltzmann 型の統計力学とファジィクラスタリングの主要な変数の対応関係を表 1 にまとめる.

4. Fermi-Dirac 分布

4.1 ファジィクラスタリングにおける Fermi-Dirac 分布

(5) 式で Shannon エントロピーをファジィエントロピー

$$S^{\text{FD}} = - \sum_{k=1}^n \sum_{i=1}^c \{ u_{ik} \log u_{ik} + (1-u_{ik}) \log (1-u_{ik}) \} \quad (10)$$

に置き換えたものを L^{FD} とすると, $\partial L^{\text{FD}} / \partial u_{ik} = 0$ は

$$u_{ik}^{\text{FD}} = \frac{1}{e^{\alpha_k + \beta d_{ik}} + 1} \quad \left(\beta = \frac{1}{\lambda}, \alpha_k = -\frac{\eta_k}{\lambda} \right)$$

という量子統計の 1 つである Fermi-Dirac 分布を与える [11].

4.2 Fermi-Dirac 分布の統計力学的解釈

Fermi-Dirac 統計に従う粒子系の分配関数は, エネルギーと全粒子数が一定

$$\sum_l \epsilon_l \langle n_l^{\text{FD}} \rangle = \mathcal{E}, \quad \sum_l \langle n_l^{\text{FD}} \rangle = \mathcal{N} \quad (11)$$

という制約下で,

$$Z_{\mathcal{N}}^{\text{FD}} = \sum_{\epsilon_l \langle n_l^{\text{FD}} \rangle = \mathcal{E}, \sum_l n_l = \mathcal{N}} e^{-\beta \sum_l \epsilon_l \langle n_l^{\text{FD}} \rangle}$$

と書ける. ここで, 平均粒子数 $\langle n_l^{\text{FD}} \rangle$ を可変として結果的に全粒子数が $\mathcal{N}$ となるように新たなパラメータ α で調節すると, この新たな分配関数 (大分配関数) は

$$\Xi^{\text{FD}} = \sum_{N=0}^{\infty} (e^{-\alpha})^N Z_N^{\text{FD}} = \prod_l \sum_{\langle n_l^{\text{FD}} \rangle=0}^{\infty} (e^{-\alpha - \beta \epsilon_l})^{\langle n_l^{\text{FD}} \rangle}$$

から

$$\Xi^{\text{FD}} = \prod_l (1 + e^{-\alpha - \beta \epsilon_l})$$

表 1 ・ Maxwell-Boltzmann 型の統計力学とファジィクラスタリング

	統計力学	ファジィクラスタリング
自由エネルギー	$\mathcal{F}^{\text{MB}} = -\frac{1}{\beta} \log \left(\sum_l e^{-\beta \epsilon_l} \right)$	$F^{\text{MB}} = -\frac{1}{\beta} \sum_{k=1}^n \log \left(\sum_{i=1}^c e^{-\beta d_{ik}} \right)$
分布・帰属度関数	$\langle n_l^{\text{MB}} \rangle = \frac{e^{-\beta \epsilon_l}}{\sum_k e^{-\beta \epsilon_k}}$	$u_{ik}^{\text{MB}} = \frac{e^{-\beta d_{ik}}}{\sum_{j=1}^c e^{-\beta d_{jk}}}$
エントロピー	$S^{\text{MB}} = - \sum_l \langle n_l^{\text{MB}} \rangle \log \langle n_l^{\text{MB}} \rangle$	$S^{\text{MB}} = - \sum_{k=1}^n \sum_{i=1}^c u_{ik}^{\text{MB}} \log u_{ik}^{\text{MB}}$

と計算される[17]. Helmholtz の自由エネルギーは大分配関数から

$$\mathcal{F}^{\text{FD}} = -\mathcal{T} \left(\log \Xi^{\text{FD}} - \alpha \frac{\partial}{\partial \alpha} \log \Xi^{\text{FD}} \right)$$

と求められる[19]ので

$$\mathcal{F}^{\text{FD}} = -\frac{1}{\beta} \left\{ \sum_l \log(1 + e^{-\alpha - \beta \epsilon_l}) + \alpha \mathcal{N} \right\}$$

となる.

一方, ファジィクラスタリングではデータ x_k はすべてのクラスターに属し得ることに注意すると, ファジィクラスタリングの大分配関数と自由エネルギーはそれぞれ,

$$\Xi^{\text{FD}} = \prod_{k=1}^n \prod_{i=1}^c (1 + e^{-\alpha_k - \beta d_{ik}})$$

$$\mathcal{F}^{\text{FD}} = -\frac{1}{\beta} \sum_{k=1}^n \left\{ \sum_{i=1}^c \log(1 + e^{-\alpha_k - \beta d_{ik}}) + \alpha_k \right\}$$

となる[11].

Fermi-Dirac 型の統計力学とファジィクラスタリングの対応関係を表2にまとめる.

5. Bose-Einstein 分布

5.1 ファジィクラスタリングにおける正規化の制約

3, 4 章で見たように, FCM 法にエントロピーを導入したファジィクラスタリングに現れる分布関数には統計力学的に対応する分布関数がある. 前章で導入したパラメータ α と α_k は, 統計力学では(11)式, ファジィクラスタリングでは(2)式を満たすように定められる. したがって, 全粒子数一定の制約に確率的な正規化の制約を対応させて考えることが可能である.

5.2 統計力学における Bose-Einstein 分布

しかし, 物理的には全粒子数一定の制約のない状態もあり得る. そこで, 本章では(11)式の制約のない統計力学を参考に, (2)式の制約を課さないファジィクラスタリングを導出する. 全粒子数一定の制約のない状態とは全粒子数が不定の系に相当し, 箱の中の輻射のように, 閉じた系であっても常に光子が生成・消滅している状態を表す. このとき, 大分配関数は

$$\Xi^{\text{BE}} = \prod_l \frac{1}{1 - e^{-\beta \epsilon_l}}$$

となるため, $\mathcal{F}^{\text{BE}}$ として

$$\mathcal{F}^{\text{BE}} = \frac{1}{\beta} \sum_l \log(1 - e^{-\beta \epsilon_l})$$

を得る. また, 平均粒子数は

$$\langle n_l^{\text{BE}} \rangle = \frac{1}{e^{\beta \epsilon_l} - 1}$$

となる. これは, もう1つの量子統計である Bose-Einstein 統計で全粒子数が不定の場合に相当し, 全粒子数が一定の場合にはそれぞれ

$$\Xi^{\text{BE}} = \prod_l \frac{1}{1 - e^{-\alpha - \beta \epsilon_l}} \quad (12)$$

$$\mathcal{F}^{\text{BE}} = \frac{1}{\beta} \left\{ \sum_l \log(1 - e^{-\alpha - \beta \epsilon_l}) - \alpha \mathcal{N} \right\}$$

$$\langle n_l^{\text{BE}} \rangle = \frac{1}{e^{\alpha + \beta \epsilon_l} - 1}$$

となる[17].

5.3 ファジィクラスタリングにおける Bose-Einstein 分布

次に, 前節の内容をファジィクラスタリングの言葉に置き換える. 制約(2)を除いた大分配関数は

表2 Fermi-Dirac 型の統計力学とファジィクラスタリング*

	統計力学	ファジィクラスタリング
自由エネルギー	$\mathcal{F}^{\text{FD}} = -\frac{1}{\beta} \left\{ \sum_l \log(1 + e^{-\alpha - \beta \epsilon_l}) + \alpha \mathcal{N} \right\}$	$\mathcal{F}^{\text{FD}} = -\frac{1}{\beta} \sum_{k=1}^n \left\{ \sum_{i=1}^c \log(1 + e^{-\alpha_k - \beta d_{ik}}) + \alpha_k \right\}$
分布・帰属度関数	$\langle n_l^{\text{FD}} \rangle = \frac{1}{e^{\alpha + \beta \epsilon_l} + 1}$	$u_{ik}^{\text{FD}} = \frac{1}{e^{\alpha_k + \beta d_{ik}} + 1}$
エントロピー	$\mathcal{S}^{\text{FD}} = - \sum_l \{ \langle n_l^{\text{FD}} \rangle \log \langle n_l^{\text{FD}} \rangle + (1 - \langle n_l^{\text{FD}} \rangle) \log(1 - \langle n_l^{\text{FD}} \rangle) \}$	$\mathcal{S}^{\text{FD}} = - \sum_{k=1}^n \sum_{i=1}^c \{ u_{ik}^{\text{FD}} \log u_{ik}^{\text{FD}} + (1 - u_{ik}^{\text{FD}}) \log(1 - u_{ik}^{\text{FD}}) \}$

*) α , α_k は $\sum_l \langle n_l^{\text{FD}} \rangle = \mathcal{N}$, $\sum_{i=1}^c u_{ik}^{\text{FD}} = 1$ を満たすように定める.

$$\Xi^{\text{BE}} = \prod_{k=1}^n \prod_{i=1}^c \frac{1}{1 - e^{-\beta d_{ik}}}$$

であるが、これは $d_{ik}=0$ で意味を持たなくなる。そこで、(12)式で全粒子数 $\mathcal{N}$ が常に一定値を取らずに変動する場合を考え、これに対応した大分配関数

$$\Xi^{\text{BE}} = \prod_{k=1}^n \prod_{i=1}^c \frac{1}{1 - e^{-\alpha_k - \beta d_{ik}}} \quad (13)$$

を用いる。(13)式から自由エネルギーと帰属度関数は

$$F^{\text{BE}} = \frac{1}{\beta} \sum_{k=1}^n \left\{ \sum_{i=1}^c \log(1 - e^{-\alpha_k - \beta d_{ik}}) - \alpha_k m_k \right\}$$

$$u_{ik}^{\text{BE}} = \frac{1}{e^{\alpha_k + \beta d_{ik}} - 1} \quad (14)$$

と計算され、 u_{ik} の i に関する和は実数 m_k となる。 m_k が 1 以外の値を取ることを許すことで、正規化の制約のないクラスタリングが実現できる。クラスター中心 v_i は、 $\partial F^{\text{BE}} / \partial v_i = 0$ から

$$v_i = \frac{\sum_{k=1}^n u_{ik}^{\text{BE}} x_k}{\sum_{k=1}^n u_{ik}^{\text{BE}}} \quad (15)$$

となる。以上のように、統計力学に則った確率的でないファジィクラスタリングの定式化が可能である。

Bose-Einstein 型の統計力学とファジィクラスタリングの対応関係を表 3 にまとめる。

5.4 Bose-Einstein 分布のエントロピー

(9)式に示すように、エントロピー $\mathcal{S}$ は自由エネルギー $\mathcal{F}$ とエネルギー $\mathcal{E}$ から

$$\mathcal{S} = \beta(\mathcal{E} - \mathcal{F})$$

と導かれ、ファジィクラスタリングの場合も

$$\mathcal{S}^{\text{MB}} = \beta(E^{\text{MB}} - F^{\text{MB}})$$

$$\mathcal{S}^{\text{FD}} = \beta(E^{\text{FD}} - F^{\text{FD}})$$

が満たされている。そこで、この関係を利用してファジィクラスタリングの Bose-Einstein 分布のエントロピー $\mathcal{S}^{\text{BE}}$ が求められる。すなわち、

$$E^{\text{BE}} = \sum_{k=1}^n \sum_{i=1}^c d_{ik} u_{ik}^{\text{BE}}$$

であるから、 $\mathcal{S}^{\text{BE}} = \beta(E^{\text{BE}} - F^{\text{BE}})$ として

$$\begin{aligned} \mathcal{S}^{\text{BE}} &= - \sum_{k=1}^n \sum_{i=1}^c \{ u_{ik}^{\text{BE}} \log u_{ik}^{\text{BE}} \\ &\quad - (1 + u_{ik}^{\text{BE}}) \log(1 + u_{ik}^{\text{BE}}) \} \end{aligned} \quad (16)$$

を得る。

以上 3 種類の帰属度関数とエントロピー関数の大まかな形状を図 1 と 2 に示す。図では i と k を固定した状態で、横軸を $|d_{ik}|$ として表示している。

6. 統計力学とファジィクラスタリングの対応

6.1 エントロピー最大化法

これまでの議論をまとめると、統計力学的には次式のように、孤立系で全粒子数と全エネルギーが一定の制約下でエントロピー $\mathcal{S}$ を最大化することで熱平衡状態における最も確らしい粒子分布が決まる。

$$\delta \mathcal{S} + \alpha \sum_l \delta(\mathcal{N} - \langle n_l \rangle) + \beta \sum_l \delta(\mathcal{E} - \epsilon_l \langle n_l \rangle) = 0$$

一方、ファジィクラスタリングで用いた(4)、(10)、(16)式のエントロピーは、

$$\begin{aligned} -\mathcal{S} &= \sum_{k=1}^n \sum_{i=1}^c \{ u_{ik} \log \left(\frac{1 - Xu_{ik}}{u_{ik}} \right) \\ &\quad - \frac{1}{X} \log(1 - Xu_{ik}) \} \end{aligned}$$

のようにまとめられる。 $X=0$ が Maxwell-Boltzmann 型、 $X=1$ が Fermi-Dirac 型、 $X=-1$ が Bose-Einstein

表 3 Bose-Einstein 型の統計力学とファジィクラスタリング

	統計力学*	ファジィクラスタリング
自由エネルギー	$\mathcal{F}^{\text{BE}} = \frac{1}{\beta} \left\{ \sum_l \log(1 - e^{-\alpha - \beta \epsilon_l}) - \alpha \mathcal{N} \right\}$	$F^{\text{BE}} = \frac{1}{\beta} \sum_{k=1}^n \left\{ \sum_{i=1}^c \log(1 - e^{-\alpha_k - \beta d_{ik}}) - \alpha_k m_k \right\}$
分布・帰属度関数	$\langle n_l^{\text{BE}} \rangle = \frac{1}{e^{\alpha + \beta \epsilon_l} - 1}$	$u_{ik}^{\text{BE}} = \frac{1}{e^{\alpha_k + \beta d_{ik}} - 1}$
エントロピー	$\mathcal{S}^{\text{BE}} = - \sum_l \{ \langle n_l^{\text{BE}} \rangle \log \langle n_l^{\text{BE}} \rangle \\ - (1 + \langle n_l^{\text{BE}} \rangle) \log(1 + \langle n_l^{\text{BE}} \rangle) \}$	$\mathcal{S}^{\text{BE}} = - \sum_{k=1}^n \sum_{i=1}^c \{ u_{ik}^{\text{BE}} \log u_{ik}^{\text{BE}} \\ - (1 + u_{ik}^{\text{BE}}) \log(1 + u_{ik}^{\text{BE}}) \}$

*) 全粒子数 $\mathcal{N}$ は可変であるとみなす。

型のエントロピーである. FCM 法に基づくファジィクラスタリングでは, $-S$ を正規化と目的関数最小の制約下で最小化する.

$$-\delta S + \alpha_k \sum_{k=1}^n \sum_{i=1}^c \delta(Y - u_{ik}) - \beta \sum_{k=1}^n \sum_{i=1}^c \delta(d_{ik} u_{ik}) = 0$$

$$\left(Y = \frac{(m_k - 1)(X^2 - X) + 2}{2} \right)$$

この結果, 帰属度関数として

$$u_{ik} = \frac{1}{e^{\alpha_k + \beta d_{ik}} + X}$$

を得る. ファジィクラスタリングは, それぞれのエントロピーを最大化する最も確からしいデータ配置を求めることに相当する.

FCM 法にエントロピーを導入したファジィクラスタリングでは, これまで述べたように, 分布関数, エネルギー, 自由エネルギー, エントロピーや分配関数等の統計力学的な状態量に対応する情報量的な量を計算可能である. 一方, 標準的な FCM 法にはエントロピー項 $-S$ が存在しない. このため, 分配関数や自由エネルギーを定めることができず, 統計力学的に解釈することはできない.

6.2 粒子分布とデータ分布の対応

6.1で, エネルギー準位を $\epsilon_l (l=0, 1, \dots, \Lambda)$, ϵ_l に含まれる準位数を g_l とする. さらに, $g_l \langle n_l \rangle = \mathcal{N}_l$ とすると, (11) 式より $\langle n_l \rangle / \mathcal{N}$ は全粒子中で準位 l に属するものの割合を表し, l に関する総和は 1 となる. そこで, 1 個のクラスターをデータが格納可能な複数個の仮想的な箱の集合とみなすと, ファジィクラスタリングにおけるデータ分布を表す諸量は

$$\frac{\langle n_l \rangle}{\mathcal{N}} \Leftrightarrow u_{ik}$$

$$l \Leftrightarrow i$$

$$\Lambda \Leftrightarrow c$$

$$g_l \Leftrightarrow \text{クラスターを構成する箱の数}$$

$$\mathcal{N}_l \Leftrightarrow \text{クラスターに属するデータ数}$$

のように対応しているものと考えられる. このとき, 統計力学とファジィクラスタリングとの本質的な差異は, 異なる準位に属する個々の粒子が区別不可能なのに対して, データは区別可能かつ複数クラスターに同時に帰属可能なことである. さらに, クラスタリングをクラスター間でのデータの (帰属度の大小の上での) 遷移とみなす [11] と, Fermi-Dirac 型の帰属度関数ではデータは空箱にしか入れない.

しかし, Bose-Einstein 型ではクラスターに属し得るデータ数に上限がなく, データ密度が高い領域にデータが集中するように遷移確率はデータ密度に比例すると解釈できる.

最後に, 統計力学との間に対応関係を定めることが可能なファジィクラスタリング方法を表 4 に示す. 形式的には表 4 以外のファジィクラスタリング方法 (正規化の制約を除いた FCM 法と Shannon エントロピーやファジィエントロピーの組み合わせなど) を考えることは可能であるが, それらは統計力学の範疇にはない.

7. ファジィクラスタリングアルゴリズム

Bose-Einstein 型の帰属度関数 (14) 式は, 温度 T が下がる (β が大きくなる) と分布の広がり小さくなる. そこで, 徐々に温度を下げながら最適なクラスター形状と配置を求めるアニーリング法を用いる. アルゴリズムの有効性を示すために簡単な例題を使った実験を行い, [11] の Fermi-Dirac 型の帰属度関数のアルゴリズムの結果と比較する.

図1 3種類の帰属度関数

図2 3種類の帰属度関数のエントロピー

7.1 アルゴリズム

(14)式はまた、 α_k の増大に対しても分布の広がり小さくする。そこで、 α_k にもアニーリング法を適用する。Bose-Einstein 型の帰属度関数に対するアルゴリズム(以下 BE とする)を以下にまとめる。

step 1. (条件設定) クラスタ数 c 、アニーリングの最高温度 T_{high} 、最低温度 T_{low} 、温度の更新割合 T_{rate} 、 α_k の初期値、アルゴリズムの収束判定のしきい値 e_0 、 e_1 を与える。

step 2. (初期化) ランダムな位置に初期クラスターを生成する。 $T = T_{high}$ とする。アルゴリズムの収束判定のしきい値の現在値を $e_c = e_0$ とする。

step 3. 現在の温度 T_c に対して $T = T_c e^{-T_c/100T_{high}}$ と更新し、 α_k を $\alpha_k = \alpha_k \cdot T/T_c$ とする。(14)式の帰属度関数 u_{ik}^{BE} を計算する。

step 4. (15)式のクラスター中心 v_i を計算する。

step 5. (収束テスト) 1ステップ前の中心値 $\hat{v}_i$ との差を $e = \max_{1 \leq i \leq c} \|v_i - \hat{v}_i\|$ と求め、 $e < e_1$ を満たせば処理を終了する。また、 $e_c \leq e$ を満たせば6.へ、満たさなければ $e_c = e$ として3.へ戻る。

step 6. (冷却処理) $T < T_{low}$ を満たせば処理を終了する。満たさなければ $T = T \cdot T_{rate}$ と温度を更新して $e_c = e_0$ として3.へ戻る。

Fermi-Dirac 型の帰属度関数に対するアルゴリズム(以下 FD とする)の場合は、

step 1.で、 βd_{ik} をその i に対する平均値 $\overline{\beta d_{ik}}$ で置き換ええると $\alpha_k = \log(c-1) - \overline{\beta d_{ik}}$ を得ることから、これを α_k の初期値とする。

step 3.で、

$$\alpha_k = \log \left\{ (e^{\beta d_{ik}} + 1) \sum_{j=1}^c \frac{1}{e^{\beta d_{jk}} + 1} - 1 \right\} - \beta d_{ik}$$

表 4 統計力学とファジィクラスタリングの対応

統計力学	ファジィクラスタリング
粒子数一定の Maxwell-Boltzmann 分布	FCM+Shannon エントロピー+正規化
粒子数一定の Fermi-Dirac 分布	FCM+ファジィエントロピー+正規化
粒子数一定の Bose-Einstein 分布	FCM+Bose-Einstein エントロピー+正規化
粒子数不定の Bose-Einstein 分布	FCM+Bose-Einstein エントロピー

(a) BE (Thigh)

(a) FD (Thigh)

(b) BE (convergent state)

(b) FD (convergent state)

図 3 ファジィクラスタリング結果

表5 収束までのループ回数(10回の計算の平均値)

初期クラスター数	5	10	20
BE	14	15	15
FD	29	76	670

と選ぶと正規化の制約が満たされるため、 α_k を上式で求める。 u_{ik}^{BE} のかわりに u_{ik}^{FD} を用いる。

step 5. で、 $e < e_0$ を満たせば6.へ、満たさなければ3.へ戻る。

とする点が異なる。

7.2 実験結果

実験には図3の[11]と同じ約2300個のデータ(・印)からなる5個のクラスターを用い、BEとFDのアルゴリズムを適用した。

条件設定は、 $c=5, 10, 20$, $T_{\text{high}}=1.0 \times 10^{28}$, $T_{\text{low}}=1.0 \times 10^{20}$, $T_{\text{rate}}=0.5$ である。また、BEの場合は $\alpha_k = -1 \times 10^{-3}$ 、(図のスケールに対して) $e_0=500$, $e_1=10$, FDの場合は $e_0=20$ とした。また、FDでは最も e が小さくなったときを収束状態とした。

初期クラスター数が10の場合について、(a) T_{high} 近傍、(b)収束状態、で得たクラスター中心(○印)を図3に示す。BE、FDいずれの場合も、(a)の高温状態では β 値が相対的に小さく帰属度関数がデータ領域全体に渡って広がるためクラスターはデータの重心付近にできるが、(b)では分布関数の幅が狭まり、実際の中心に近い位置にクラスター中心は移動した。収束温度はBE、FDとも $1 \sim 5 \times 10^{26}$ 程度であった。また、収束時の e の値は、全般にFDの方が大きくなった。

収束までに要したアルゴリズムのループ回数は、表5に示す通り、BEの方がFDよりも少なくなっている。これには、収束条件の違い以外に、(1)図1に見られる通り、BEの関数形状における遷移領域[11]がFDよりも急峻であることで収束が早まった可能性、(2)FDでは k ごとに α_k を求めていることが、逆に、収束性に悪影響を及ぼしてしまった可能性、などの要因が考えられ、さらなる検証が必要である。

なお、本実験で用いた α_k の初期値や e_0 , e_1 の値等は最適化したものではなく、今後、このようなパラメータの設定方法や指針が求められる。

8. おわりに

本論文では、統計力学とファジィクラスタリングとの関係についてまとめた。ファジィ c 平均法にエントロピーを導入したファジィクラスタリングは、Shannon エントロピーの場合はMaxwell-Boltzmann型、ファジィエントロピーの場合はFermi-Dirac型、Bose-

Einstein エントロピーの場合はBose-Einstein型の帰属度関数を与えることを示した。

今後の課題は、これらの帰属度関数を用いたファジィクラスタリングのアルゴリズムの検証や改良と、統計力学的な手法に基づく解析の検討である。また、Bose-Einstein型の帰属度関数は統計力学とのアナロジーから形式的に導いたに過ぎないので、パラメータの設定法に関する検討以外にも、従来の確率的な方法とのより詳細な比較に基づく可能性クラスタリングとしての評価が必要となる。

参考文献

- [1] 西森他, “岩波講座 物理の世界 物理と情報シリーズ,” 岩波書店, 2003~.
- [2] 篠本滋, “情報の統計力学,” 丸善出版株式会社, 1992.
- [3] M. Mezard, G. Parisi and M. A. Virasoro, “*Spin Glass Theory and Beyond*,” Singapore: World Scientific Publishing, 1987.
- [4] H. Nishimori, “*Statistical Physics of Spin Glasses and Information Processing: An Introduction*,” New York: Oxford University Press, 2001.
- [5] E. Aarts and J. Korst, “*Simulated Annealing and Boltzmann Machines*,” John Wiley & Sons, Chichester, 1990.
- [6] S. Kirkpatrick, C. D. Gelatt and M. P. Vecchi, “Optimization by simulated annealing,” *Science*, Vol.220, pp.671-680, 1983.
- [7] J. C. Bezdek, “*Pattern Recognition with Fuzzy Objective Function Algorithms*,” New York: Prentice Hall, 1981.
- [8] R.-P. Li and M. Mukaidono, “A Maximum entropy approach to fuzzy clustering,” *Proc. of the 4th IEEE Int. Conf. Fuzzy Systems (FUZZ-IEEE/IFES'95)*, pp.2227-2232, 1995.
- [9] S. Miyamoto and M. Mukaidono, “Fuzzy c-means as a regularization and maximum entropy approach,” *Proc. of the 7th Int. Fuzzy Systems Association World Congress*, Vol.II, pp.86-92, 1997.
- [10] K. Rose, E. Gurewitz and B. C. Fox, “A deterministic annealing approach to clustering,” *Pattern Recognition Letters*, Vol.11, No.9, pp.589-594, 1990.
- [11] 安田真, 古橋武, 松崎元昭, 大熊繁, “確定的アニーリング法を用いたファジィクラスタリングとその統計力学的性質,” 日本ファジィ学会誌, Vol.13, No.5, pp.522-529, 2001.
- [12] A. DeLuca and S. Termini, “A definition of a nonprobabilistic entropy in the setting of fuzzy sets theory,” *Information and Control*, Vol.20, pp.301-312, 1972.
- [13] 安田真, 古橋武, 大熊繁, “ファジィクラスタリングと統計力学,” 第19回ファジィシステムシンポジウム講演論文集, pp.589-592, 2003.
- [14] R. Krishnapuram and J. M. Keller, “A possibilistic approach to clustering,” *IEEE Trans. Fuzzy Sys-*

- tems, Vol.1, No.2, pp.98-110, 1993.
- [15] 渋谷和宏, 宮本定明, 高田治, 馬屋原一孝, “ファジィ c-平均クラスタリングにおける制約条件と正則化および可能性クラスタリング,” 日本ファジィ学会誌, Vol.13, No.6, pp.701-715, 2001.
- [16] 宮本定明, “クラスター分析入門,” 森北出版, 1999.
- [17] 戸田盛和, 久保亮五, 斎藤信彦, 橋爪夏樹, “(岩波講座現代物理学の基礎 5) 統計物理学,” 岩波書店, 1978.
- [18] L. E. Reichl, “A Modern Course in Statistical

Physics,” New York: John Wiley & Sons, 1998.

- [19] 高橋康, “統計力学入門,” 講談社, 1984.

(2004年10月4日 受付)

(2005年5月17日 採録)

[問い合わせ先]

〒501-0495 岐阜県本巣市上真桑

岐阜工業高等専門学校電気情報工学科

安田 真

TEL : 058-320-1363

E-mail : yasuda@gifu-nct.ac.jp

著者紹介

やすだ まこと
安田 真 [非会員]

1985年 関西学院大学理学部物理学科卒業。2002年より岐阜工業高等専門学校電気情報工学科助教授。博士(工学)。主としてエージェント技術とソフトウェアエンジニアリングに関する研究に従事。計測自動制御学会, 情報処理学会, 人工知能学会, 日本ソフトウェア科学会各会員。

ふるはし たけし
古橋 武 [正会員]

名古屋大学大学院計算理工学専攻。1985年3月 名古屋大学大学院博士課程修了。1987年3月 名古屋大学工学部助手。1990年12月 名古屋大学工学部助教授。2001年1月 三重大学工学部教授。2004年4月 名古屋大学教授。現在に至る。工学博士。主としてソフトウェアエンジニアリングに関する研究に従事。

おおくま しげる
大熊 繁 [非会員]

1977年 名古屋大学大学院博士課程満了。1977年 名古屋大学助手, 1979年 同講師, 1984年 同助教授, 1990年より同教授。工学博士。主として, ロボティクス, パワーエレクトロニクス, 創発型ソフトウェアに関する研究に従事。電気学会, ロボット学会, IEEE, パワーエレクトロニクス研究会各会員。

Entropy based Fuzzy c-means Clustering: Analogy with Statistical Mechanics

by

Makoto YASUDA, Takeshi FURUHASHI and Shigeru OKUMA

Abstract :

In this paper, we summarize the statistical mechanical representation of fuzzy clustering. Then, we give a framework of a possibilistic clustering based on a Bose-Einstein type membership function, and examine its clustering mechanisms.

The fuzzy c-means clustering (FCM) method regularized with Shannon entropy gives the Maxwell-Boltzmann (or Gibbs) distribution function as a membership function. Similarly, by introducing fuzzy entropy to the FCM, we obtain the Fermi-Dirac type membership function. In these cases, the constraint that the sum of all particles is fixed is correspondent with the normalization constraint in fuzzy clustering.

Furthermore, it is known that the state in which the total number of particles is not conserved exists and written by the Bose-Einstein distribution function. Thus, by the analogy of statistical mechanics, we obtain the Bose-Einstein type membership function without the constraint of normalization and propose a new fuzzy clustering algorithms.

Keywords : fuzzy c-means clustering, fuzzy entropy, possibilistic clustering, statistical mechanics, Maxwell-Boltzmann distribution, Fermi-Dirac distribution, Bose-Einstein distribution

Contact Address : **Makoto YASUDA**

Dept. of Electrical and Computer Engineering, Gifu National College of Technology

Kamimakuwa, Motosu-shi, Gifu 501-0495, Japan

TEL : 058-320-1363

E-mail : yasuda@gifu-nct.ac.jp