

修士論文の和文要旨

研究科・専攻	大学院 情報理工 学研究科 情報・ネットワーク工学 専攻 博士前期課程		
氏 名	田口 慎吾	学籍番号	1731098
論文題目	自己相似分布上のレート歪関数とシャノン下界の漸近的最適性		
要 旨	<p>有歪情報源符号化では情報源アルファベットと再生アルファベット間に定義された歪尺度に基づいて、平均歪に対する達成可能な情報レートの最小値をレート歪関数として表現することを目指す。</p> <p>まず基本的仮定として情報源は定常無記憶であるとする。</p> <p>情報源アルファベットと再生アルファベットをいずれもユークリッド空間にとり、歪尺度を情報源記号と再生記号の差にのみ依存する差歪尺度の加法により定義されるものとする。</p> <p>このときレート歪関数には、シャノン下界式が知られている。歪がごく小さいときのレート歪関数とシャノン下界の差が0になる性質を漸近的最適性 (Asymptotic Tightness) というが、情報源の確率分布が密度関数を持つ場合を対象に Koch らによる先行研究がある。</p> <p>一方、情報源の確率分布が密度関数を持たない特異分布に関しては未解明である。</p> <p>本研究では、特異分布のひとつであり、自己相似性という再帰的表現と相性のいい性質を持つ自己相似分布を対象として、レート歪関数に対するシャノン下界の振る舞いを理論的に解析することを目標とする。</p> <p>結論として、自己相似分布におけるレート歪関数に対するシャノン下界の漸近的最適性は一般に不成立であることが示された。</p>		

平成31年度 修士学位論文

自己相似分布上のレート歪関数と
シャノン下界の漸近的最適性

電気通信大学 大学院 情報理工学研究科
博士前期課程 情報・ネットワーク工学専攻

1731098 田口 慎吾

指導教員 川端 勉 教授 八木 秀樹 准教授

提出 平成31年1月28日



目次

1	はじめに	2
2	自己相似集合	5
2.1	自己相似	5
2.2	相似次元	6
2.3	自己相似集合の表現方法	6
2.4	歪測度関数についての仮定	8
2.5	自己相似集合の離散集合近似	8
3	既存研究	11
3.1	シャノン下界について	11
3.2	シャノン下界の漸近的最適性に関する既存研究の結果	13
3.2.1	漸近的最適性についての研究の流れ	13
3.2.2	定理 4([3]) の証明の概要	14
3.3	これまでの本研究の要旨と訂正内容について	16
3.3.1	結果までの流れ	16
3.3.2	修正前の研究において正しくなかった箇所	19
4	自己相似分布に対するレート歪関数とシャノン下界について	22
4.1	自己相似分布に対するレート歪関数とシャノン下界の漸近的性質	22
4.2	漸近的最適性が不成立となる反例	22
5	まとめ	29
A	系 1 の証明	31

第 1 章

はじめに

データ圧縮においてまず考えられるべき重要なことのひとつは、圧縮（すなわち符号化）したデータを解凍（すなわち復号化、または再構成）したものが、もとのデータと実用上一致することである。有歪のデータ圧縮を考えたとき、一般に期待歪と情報レートはトレードオフの関係にあることが知られており、符号化と復号化の間に加わる期待歪が小さいほど、データ圧縮に要する情報レートは小さくて済む。この意味で、期待歪と情報レートを知ること、データ圧縮の性能を知ることには等しいといえる。そして、期待歪と情報レートとの関係は、一般にレート歪関数によって表される。

レート歪関数は、非可逆圧縮に関する圧縮達成の尺度のひとつとして有用な関数である。情報源が i.i.d. に従い d 次元、実数値を持つ確率ベクトル $\{X_k, k \in \mathbb{Z}\}$ を生み出し、それらが分布 P_X に従う場合を考える。もしすべてのブロック長 n と歪み制約 D について、情報源ベクトルの系列 X_1, \dots, X_n を再構築ベクトルの系列 $\hat{X}_1, \dots, \hat{X}_n$ へと量子化するなら、最小のレート $R(D)$ （単位は nats/記号）は、[1], [2] より

$$R(D) = \inf_{P_{\hat{X}|X}: E[\|X - \hat{X}\|^r] \leq D} I(X; \hat{X}) \quad (1.1)$$

となる。これがレート歪関数である。レート歪関数は正規分布 $\mathcal{N}(\mu, \sigma)$ に従う情報源をもつなど一部の特殊な場合を除いて、一般に解析的に解を求めることが難しいため、上界や下界による評価は可能である。特に、 $\rho(X, \hat{X})$ が X と \hat{X} の差の関数の場合には、下界（Shannon 下界）は評価しやすいことが多い。

離散値を持つ確率変数ベクトル X とその再構成確率変数ベクトル Y があり、それぞれ実現値として x_j, y_k をもつとする。これらの確率は以下のように条件を考える。

$$\sum_{x_j} P(x_j) = 1 \quad (1.2)$$

$$\sum_{y_k} Q(y_k) = 1 \quad (1.3)$$

$$P(x_j) > 0 \quad \text{for } j = 1, 2, \dots, M \quad (1.4)$$

$$Q(y_k) \geq 0 \quad \text{for } k = 1, 2, \dots, N \quad (1.5)$$

さらに，表記の簡単のために以下のように書くことがある．

$$P(\mathbf{x}_j) = P_j \quad (1.6)$$

$$Q(\mathbf{y}_k | \mathbf{x}_j) = Q_{k|j} \quad (1.7)$$

$$Q(\mathbf{y}_k) = \sum_j P_j Q_{k|j} \quad (1.8)$$

$$\rho(\mathbf{x}_j, \mathbf{y}_k) = \rho_{jk} \quad (1.9)$$

$R(D)$ のシャノン下界 $R_{SLB}(D)$ は [2] より次式のように表される．

$$R_{SLB}(D) = H(\mathbf{X}) - H(\{z_j\}) \quad (1.10)$$

ここで $\{z_j\}$ は以下のようになる．

$$\{z_j\} = \frac{e^{s\rho_j k_0}}{C} \quad (1.11)$$

ただし，

$$k_0 = \arg \max_k \sum_j e^{s\rho_{jk}}, \quad (1.12)$$

また C は

$$C \triangleq \sum_j e^{s\rho_j k_0} \quad (1.13)$$

とした．また，有限の整数 M, N に対して

$$j \in \mathcal{A}_M, \quad k \in \mathcal{A}_N \quad (1.14)$$

かつ

$$\mathcal{A}_M := \{1, 2, \dots, M\} \quad (1.15)$$

$$\mathcal{A}_N := \{1, 2, \dots, N\} \quad (1.16)$$

とする．

本稿の議論の大きな参考になっている T. Koch の論文 [3] では，連続型の分布を持ち，許容された歪みを仮定したレート歪関数とシャノン下界との差が，整数部分の微分エントロピーが有限であるような確率ベクトルを持ったすべての情報源に対して漸近的にタイトであるということが説明されている．対照的に，もし情報源の確率ベクトルの整数部分が無限のエントロピーを持つとき，そのレート歪関数はすべての有限な歪みに対して発散することもまた説明されている．

以上のような研究がなされてきたわけであるが、ここで疑問が生じる。それは、扱う情報源が確率密度関数 (pdf) を持たない場合、レート歪関数とシャノン下界との差はどのように振る舞うのかという疑問である。

pdf を持たない情報源の確率分布には、離散型の確率分布と特異型の確率分布がある。特に特異型の分布において、レート歪関数のシャノン下界に対する漸近的達成がいえるかどうかは分かっていない。この特異型分布に従う情報源に対しては [9] などの結果が有界性を示すにとどまっている。特に、特異型分布のひとつにフラクタル集合上の分布があり、このフラクタル集合は微細構造が繰り返し現れる特徴的な構造を持つことから、何らかの数学的表現によって理論的評価を行うことができると期待される。一方でフラクタル集合上の分布に対するレート歪関数について、シャノン下界の有用性は確認されていないため、これを考察することを本研究の課題とした。より具体的には、フラクタル集合の中でも扱いが容易な自己相似集合上の確率分布のレート歪関数に対する、シャノン下界の漸近的な振る舞いについて詳しく考察する。

本稿では、フラクタル集合の中でも自己相似集合を取り扱い、自己相似集合上の確率分布に対するレート歪関数 $R(D)$ について漸近的達成がいえるかどうかについて考察をしていく。そのため本稿の構成の概要は以下ようになる。第 2 章では、本稿で主に扱う自己相似集合に関する概要を説明している。第 3 章では、議論の対象であるレート歪関数について、既存研究を用いて概要を説明している。第 4 章では、1~3 章で導入した性質や定理などの手段を用いて、自己相似集合上の確率測度に対するレート歪関数とシャノン下界の漸近的性質を考察していく。第 5 章では、本稿で得られた結果をまとめている。

第 2 章

自己相似集合

この章では本稿で特異分布を持った確率ベクトルの一例として扱うことになる自己相似集合について、基本的な性質やのちに利用する性質、あるいは具体的な例などを紹介することによって、これ以降の章で行う近似などの説明に対する 根拠を示す狙いがある。

2.1 自己相似

ふたつの集合 S と $T (\subset \mathbb{R}^d)$ について、 $W : \mathbb{R}^d \rightarrow \mathbb{R}^d$ という関数 $W(x)$ が以下のように定義されるとする。

$$W(x) = \alpha O x + \beta \quad (2.1)$$

ただし、 $\alpha \in (0, 1)$ 、 O は直交行列、 $\beta \in \mathbb{R}^d$ であり、以下が成立するとする。

$$S = WT \quad (2.2)$$

このとき、ふたつの集合 T と S は相似であるという。

集合 S が自己相似であるとは、 $m < \infty$ なる m について、 S と相似な集合 $\{S_i\}_{i=1}^m$ の非交和（互素な要素の和）が存在することである。あるいは、 $S_i = W_i S$ となる前述の形の相似変換 W_i ($i = 1, \dots, m$) が存在することであり、それは次のように表される。

$$S = \uplus_{i=1}^m S_i \quad (2.3)$$

特に基本的な自己相似集合に、 α -カントール集合がある。この集合を具体的な場合として以下に示した。

例 1 : α -カントール集合

任意の $0 < \alpha < 1$ に対し、以下の関係を満たす集合 $S \subset [0, 1]$ が存在する。

$$S = \alpha S \cup \{\alpha S + (1 - \alpha)\} \quad (2.4)$$

これは $m = 2, W_0(x) = \alpha x, W_1(x) = \alpha x + (1 - \alpha)$ の自己相似集合である．このような自己相似集合を α -カントール集合という．また，一般にカントール集合と呼ばれるのは， $\alpha = \frac{1}{3}$ のときである．

2.2 相似次元

相似変換 W_i の縮小パラメータ α_i が开区間 $(0, 1)$ 上にあるとき，以下を定義する．

$$f(d) \triangleq \sum_{i=1}^m \alpha_i^d \quad (2.5)$$

α_i は $0 < \alpha_i < 1$ を常に満たすため，関数 $f(d)$ は d について単調減少である．さらに， $f(0) = m > 1$ であることと $\lim_{d \rightarrow \infty} f(d) = 0$ であることから，関数 $f(d)$ は $f(d) = 1$ の解をただひとつ持つことが分かる．また，その解 $d = d^*$ は正の実数で，有限である．この d^* が，集合 S の Housdorff-Besicovitch 次元であり，以下に示した方程式の唯一の正の解であることはよく知られている．

$$\sum_{i=1}^m \alpha_i^{d^*} = 1 \quad (2.6)$$

例えば，以前に示した自己相似集合のひとつである α -カントール集合の Housdorff-Besicovitch 次元は $d^* = \log_{1/\alpha} 2$ である．本稿では，これ以降相似次元として Housdorff-Besicovitch 次元を用いる．

2.3 自己相似集合の表現方法

どんな自己相似集合 S と $M = \{1, \dots, m\}$ 上の全ての無限過程の集合 M^* との間にも，一対一写像が存在する．どんな点 $x \in S$ も，以下のようにして無限過程 $b(x) = \{b_1, b_2, \dots, b_j, \dots\} \in M^*$ に写像される．

$$b_1 = i \Leftrightarrow x \in S_i \quad (2.7)$$

また，すべての $j = 2, 3, \dots$ について再帰的に，

$$b_j = i \Leftrightarrow (W_{b_{j-1}}^{-1}, W_{b_{j-2}}^{-1}, \dots, W_{b_1}^{-1}x) \in S \quad (2.8)$$

を定義する．

要素が M ，長さが n の系列集合を M^n とする．また，任意の $X_1 \subset R^d$ に対して

$$X_n = \bigsqcup_{b_1, \dots, b_n \in M^n} W_{b_1} W_{b_2} \cdots W_{b_n}(X_0) \quad (2.9)$$

と定義する．このとき， $X_\infty = \lim_{n \rightarrow \infty} X_n$ は自己相似集合 $S = \bigsqcup_{i=1}^m W_i S_i$ となる．

例：カントール集合の構成概要

以下にカントール集合の構成のイメージを図として示した．実際には無限回の相似変換を施すため，カントール集合を正確に図示することは不可能である．

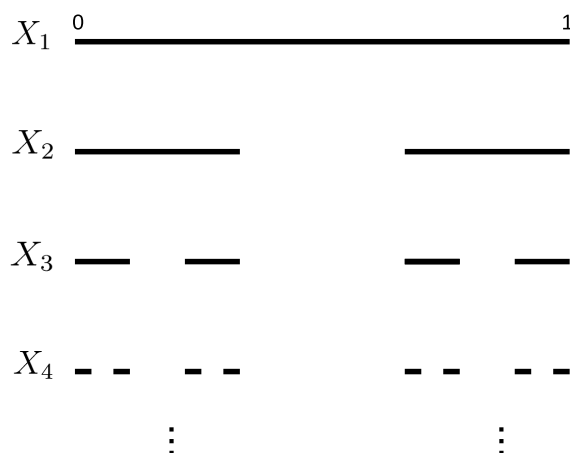


図 2.1: カントール集合の構成概要 (1/3-カントール集合)

すべての i について $\alpha_i < 1$ であるため， $j > 1$ であるような b_j が小さなスケール上にある x の値を決めている間，比較的小さな j についてインデックス b_j は x の大きなスケールの値を決める．

自己相似集合 S を支持する任意の確率測度 μ について，写像 $x \mapsto b(x)$ のもとで，無限系列の集合 M^* 上の確率測度 μ^* との対応が存在する．集合 M^* に属する系列 $b(x)$ は，有限なアルファベット M 上の確率過程の標本点として解釈される．特に，系列の測度 μ^* は μ がスケールに関して不変のとき定常であるといい，また集合 S 上のラフな地点（すなわち大きいスケールでの x の値）が，ごく小さいスケール上の μ の詳細な振る舞いに対する影響を僅かにしか持たないときエルゴード性を持つという．すべての確率のベクトル p_i , ($i = 1, \dots, m$) は任意の測度集合 $A \subset \mathbb{R}^d$ に対して，以下のように自己相似集合 S に対応するただ一つの自己相似確率測度 μ をもつ．

$$\mu(A) = \sum_{i=1}^m p_i \mu((W_i^{-1}A)) \tag{2.10}$$

ここで， p_i は標本点 $\{x \in S_i\} = \{x : b_1(x) = i\}$ の確率であることに注意する．さらに， μ^* のもとでシンボル b_i は互いに同一で独立な分布に従う (=i.i.d.) M 上の確率変数であり，

周辺確率 p_i , ($i = 1, \dots, m$) をもつ．すなわち自己相似測度は，自己相似集合上でエルゴード性と定常性を持つ測度の特別な場合だと分かる．

例 2 : α -カントール測度

α -カントール測度は $p_1 = p_2 = \frac{1}{2}$ での自己相似測度 μ である．ただし，集合 S は例 1 で挙げた α -カントール集合である．特に $\alpha = \frac{1}{3}$ のとき， α -カントール測度のことを単にカントール測度と呼ぶ．

2.4 歪測度関数についての仮定

後の計算を簡単にするために，歪関数 $\rho(\cdot, \cdot)$ について 3 つの仮定をおく．

まず自己相似集合 S の表現を特徴づける相似変換 W_i のそれぞれについて， $r \geq 1$ の次数をもつ斉次性を仮定する．具体的には，任意の点 $\mathbf{x}, \mathbf{y} \in \mathbb{R}^d$ について

$$\rho(W_i \mathbf{x}, W_i \mathbf{y}) = \alpha_i^r \rho(\mathbf{x}, \mathbf{y}) \quad i = 1, \dots, m \quad (2.11)$$

が成立する．ここで α_i は相似変換 W_i の縮小パラメータである．

また，歪は三角不等式に従っていることが望ましい．すなわち

$$|\rho(\mathbf{x}, \mathbf{z})^{1/r} - \rho(\mathbf{x}, \mathbf{y})^{1/r}| \leq \rho(\mathbf{y}, \mathbf{z})^{1/r} \quad (2.12)$$

が (2.11) で与えられた $r \geq 1$ について成立する．

さらに自己相似集合 S は歪関数に対して以下のような上界を持つ．

$$D_r^*(1) \triangleq \min_{\mathbf{y}} \max_{\mathbf{x} \in S} \rho(\mathbf{x}, \mathbf{y}) < \infty \quad (2.13)$$

この性質は特に， $D \geq D_r^*(1)$ ならば常に $R(D) = 0$ であることを示している．

2.5 自己相似集合の離散集合近似

いま，歪関数 $\rho(\cdot)$ が相似変換の斉次性，三角不等式，および有界性をそれぞれ満たすと仮定し，集合 S から M^* への写像について考える．一連の仮定より， M^* の 2 つの系列が十分に長い語頭を持つとすれば，系列が表す 2 つの点の距離は相対的に小さいことを意味する．よって，離散測度 μ_δ は， M^* 内の系列の語頭を用いた近似に対応している．すなわち， $\forall \delta \in (0, 1)$ に対して，集合 S の無限系列 $b(\mathbf{x}) = \{b_1, \dots, b_j, \dots\}$ を次のようなインデックス $n_\delta(\mathbf{x})$ で切り取る．

$$n_\delta(\mathbf{x}) \triangleq \inf \left\{ n : \sum_{j=1}^n \ln \frac{1}{\alpha_{b_j}} \geq \ln \frac{1}{\delta} \right\} \quad (2.14)$$

ここで, α_{b_j} は点 $x \in S$ の系列表現における, j 番目の相似変換の縮小パラメータである. 有限木 T_δ を以下のように構成する.

- 木の根は長さ 0 の系列である.
- 系列 $b(x)$ が m -ary の系列だとした場合, その根は b_1 に対応する子を m 個持つ.
- 深さ j のそれぞれの中間節点は b_{j+1} に対応する子を m 個持つ.
- 系列 $\{b_1, b_2, \dots, b_{n_\delta}\}$ は木 T_δ の葉に対応する.

木 T_δ を用いて, 以下のような量子化を考える.

(a) \hat{x} を以下を満たす \mathbb{R}^d 上の点とする.

$$D_r^*(1) = \max_{x \in S} \rho(x, \hat{x}) \quad (2.15)$$

(b) 全ての $x \in S$ に対して切り取られた系列 $\{b_1, b_2, \dots, b_{n_\delta(x)}\}$ を決める.

(c) 相似変換 W_i を, 以下の対応を形成するために使用する.

$$x_\delta \triangleq W_{b_1} W_{b_2} \cdots W_{b_{n_\delta(x)}}(\hat{x}) \quad (2.16)$$

T_δ は有限木なので, 上記 (a), (b), (c) により切り取った相異なる系列は有限である. つまり, x_δ は有限なアルファベット Λ_δ (その要素数は $\delta \downarrow 0$ で無限大に近づく) に属する. ここで, x_δ と x は自己相似集合 S の規模が大きいき, ほとんど同一であることに注意する. 特に相似変換の斉次性より, x と x_δ の離散近似誤差は以下のように抑えられる.

$$\rho(x, x_\delta) \leq \prod_{j=1}^{n_\delta(x)} \alpha_{b_j}^r \sup_{z \in S} \rho(z, \hat{x}) \leq \delta^r D_r^*(1) \quad (2.17)$$

ただし $z \triangleq W_{b_{n_\delta}}^{-1} W_{b_{n_\delta-1}}^{-1} \cdots W_{b_1}^{-1}(x)$ である.

以下では, [8] に倣って “Cantor like” 集合を導入する. “Cantor like” 集合に対して, 量子化写像 $x_\delta = Q_\delta(x)$ は次の性質を持つ.

“Cantor like” 集合の定義 ([8]);

十分大きな K に対して以下が成立する自己相似集合 S を “Cantor like” 集合であるという.

$$C(K) \triangleq \sup_{\delta > 0} \sup_{y \in \mathbb{R}^d} \sum_{x_\delta \in \Lambda_\delta} e^{-K\delta^{-r}\rho(x_\delta, y)} < \infty \quad (2.18)$$

補題 1.

$$\inf_{\delta > 0} \min_{i \neq j} \left[\min_{\mathbf{x} \in \Lambda_{\delta/\alpha_i}, \mathbf{x}' \in \Lambda_{\delta/\alpha_j}} \rho(W_i(\mathbf{x}), W_j(\mathbf{x}'))^{1/r} \right] = \epsilon > 0 \quad (2.19)$$

が成り立つとき，自己相似集合 S は “Cantor like” 集合である．特に，任意の $a^* < \frac{\epsilon}{2}$ および $K > \frac{d^*}{r(a^*)^r}$ に対して，

$$C(K) \leq 1 + \frac{(\min_i \alpha_i)^{-d^*} e^{-K(a^*)^r}}{1 - \phi^{(rK(a^*)^r - d^*)}} < \infty \quad (2.20)$$

である．ただし， d^* は S のハウスドルフ-ベシコヴィッチ次元 (Hausdorff-Besicovitch dimension) であり， $\phi := \frac{2a^*}{\epsilon} < 1$ である．

Proof. [8] を参照． □

また，以下の定理を用いた．

定理 1 ([8], Proposition 4.1). 自己相似集合のレート歪関数を $R(D)$ ，その離散近似集合のレート歪関数を $R_\delta(D)$ とするとき

$$R_\delta((D^{1/r} + \delta D_r^*(1)^{1/r})^r) \leq R(D) \leq R_\delta((D^{1/r} - \delta D_r^*(1)^{1/r})^r) \quad (2.21)$$

が成立する．ただし， $D < 0$ に対して， $R_\delta(D) = \infty$ とする．

Proof. [9] を参照． □

第 3 章

既存研究

この章では、主にみつつのことを確認する。ひとつは、本稿の議論の対象であるシャノン下界 (Shannon Lower Bound, SLB) についての確認である。レート歪理論について詳細にまとめられている T. Berger の著書 [2] を参考にして、シャノン下界の基本的な性質をまとめた。ふたつめの確認は、本稿の先行研究である T. Koch の“連続型の分布のレート歪関数に対するシャノン下界の漸近的達成” [3] について触れることで、既存研究の簡単なまとめや手法の紹介を行うことである。そしてみつつめの確認は、本研究の途中経過として情報理論研究会などで報告を行った内容について触れることである。この内容は、発表時点では訂正が必要な表現だったこともあり、本章で詳しく述べる。

3.1 シャノン下界について

[2] を参考に離散型確率変数の場合を考える。所与の離散分布 P に従う確率変数 X および歪尺度が差の関数 (difference distortion measure) のレート歪関数 $R(D)$ について、以下のようにシャノン下界が導出される。

まず [2] より以下の定理を出発点に用いる。

定理 2

レート歪関数の代替的定義は次のように表せる。

$$R(D) = \max_{s \leq 0, \lambda \in \Lambda_s} \left[sD + \sum_j P_j \log \lambda_j \right] \quad (3.1)$$

ただし、 Λ_s は任意の $k \in \mathcal{A}_N$ に関して以下の制約条件を満たす全ての $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_M)$ の集合であるとする。

$$\sum_j \lambda_j P_j e^{s\rho_{jk}} \leq 1 \quad (3.2)$$

ここで ρ_{jk} は歪測度関数である。

いま, $C \neq 0$ をある k の値に対する定数とし, かつ任意の j に対して $P_j > 0$ と制約を課す. このとき,

$$\lambda_j = \frac{1}{CP_j} \quad (3.3)$$

とおくことで, (3.2) の制約から以下が得られる.

$$\sum_j P_j \lambda_j e^{s\rho_{jk}} = \frac{1}{C} \sum_j e^{s\rho_{jk}} \leq 1 \quad (3.4)$$

すなわち,

$$C \geq \sum_j e^{s\rho_{jk}} \quad (3.5)$$

である. (3.5) の右辺を最大化するような k が存在し, これを k_0 とすると, 等号が成立しているような C を選ぶことができ,

$$C = \sum_j e^{s\rho_{jk_0}} \quad (3.6)$$

が成り立つ. (3.1) より

$$\begin{aligned} R(D) &\geq sD + \sum_j P_j \log \lambda_j \\ &= sD + \sum_j P_j \log \frac{1}{CP_j} \\ &= sD + H(P) - \sum_j P_j \log C \\ &= H(P) + sD - \log C \end{aligned} \quad (3.7)$$

この (3.7) で得られた下界を s について 1 階ずつ微分すると,

$$\begin{aligned} &\frac{d}{ds} [H(P) + sD - \log C] \\ &= D - \frac{\sum_{j=1}^M \rho_{jk_0} e^{s\rho_{jk_0}}}{\sum_{j'=1}^{M'} e^{s\rho_{j'k_0}}} \\ &= D - \sum_{j=1}^M \rho_{jk_0} z_j \end{aligned} \quad (3.8)$$

$$\begin{aligned} &\frac{d^2}{ds^2} [H(P) + sD - \log C] \\ &= -\sum_{j=1}^M (\rho_{jk_0})^2 z_j + \left\{ \sum_{j=1}^M \rho_{jk_0} z_j \right\}^2 \end{aligned} \quad (3.9)$$

のように計算できる．ここで $\{z_j\}$ は以下のように表せる．

$$\begin{aligned} z_j &= \frac{e^{s\rho_{jk_0}}}{\sum_{j'=1}^{M'} e^{s\rho_{j'k_0}}} \\ &= \frac{e^{s\rho_{jk_0}}}{C} \end{aligned} \quad (3.10)$$

(3.9) は常に負であるため，関数 $H(P) + sD - \log C$ は s について上に凸であり，その最大となる s の値は (3.8) が 0 になる s の値である．このため $H(P) + sD - \log C$ が最大のときの D を D_s と表すとき，

$$D_s = \sum_j \rho_{jk_0} z_j \quad (3.11)$$

である．以上のことから，

$$\begin{aligned} R(D) &\geq H(P) + sD_s - \log C \\ &= H(P) + s \sum_j \rho_{jk_0} z_{jk_0} - \log C \\ &= H(P) + \sum_j \frac{e^{s\rho_{jk_0}}}{C} \log \frac{e^{s\rho_{jk_0}}}{C} \\ &= H(P) - H(\{z_j\}) \end{aligned} \quad (3.12)$$

となる．

3.2 シヤノン下界の漸近的最適性に関する既存研究の結果

3.2.1 漸近的最適性についての研究の流れ

T. Koch の研究 [3] は，T. Linder と R. Zamir の研究 [4] によって得られた定理の条件を緩和したものだといえる．まずは T. Linder と R. Zamir の研究について紹介する．

定理 3(Linder and Zamir [4], Corollary 1)

X は pdf を持ち， $h(X)$ が有限であるとする．さらに $E[\|X\|^\alpha] < \infty$ を満たす $\alpha > 0$ が存在することを仮定する．このときシヤノン下界は漸近的にタイトである．すなわち，

$$\lim_{D \downarrow 0} \{R(D) - R_{SLB}(D)\} = 0$$

が成立する．

Proof. [4] を参照． □

T. Koch は定理 1 から証明の手法を見直すことによって，シヤノン下界の漸近的最適性をさらに本質的に追究した．その結果，以下のような定理を導いた．

定理4(Koch [3], Theorem 2)

d 次元, 実数値の情報源からなる確率変数ベクトル \mathbf{X} が pdf を持つとする. さらに $H(\lfloor \mathbf{X} \rfloor) < \infty$ と $|h(\mathbf{X})| < \infty$ を仮定する. このときシャノン下界は漸近的にタイトである. すなわち,

$$\lim_{D \downarrow 0} \{R(D) - R_{SLB}(D)\} = 0$$

が成立する.

Proof. [3] を参照. □

ただし, $\lfloor \mathbf{a} \rfloor$, $\mathbf{a} = (a_1, \dots, a_d) \in \mathbb{R}^d$ はベクトル成分ごとの床関数を意味する. つまり, $\lfloor \mathbf{a} \rfloor = (\lfloor a_1 \rfloor, \dots, \lfloor a_d \rfloor)$ である.

定理5(Koch [3], Theorem 3)

すべての期待歪 $D > 0$ に対して, d 次元, 実数値の情報源からなる確率変数ベクトル \mathbf{X} のレート歪関数 $R(D)$ が有限であることは, $H(\lfloor \mathbf{X} \rfloor) < \infty$ であることと同値である.

Proof. [3] を参照. □

3.2.2 定理4([3]) の証明の概要

以下では, Koch の論文 [3] 中の証明を一部抜粋して紹介する. ただし, 内容は [3] の議論のため, 扱う確率ベクトル \mathbf{X} は pdf を持ち, $h(\mathbf{X})$ は \mathbf{X} の微分エントロピーである.

$R_{SLB}(D)$ の漸近的最適性を示すために, $D \downarrow 0$ とすることで 0 に収束するような “ $R(D)$ と $R_{SLB}(D)$ との差の上界” を導き出す. (1.1) の観点から, $R(D)$ の上界は $\hat{X} = \mathbf{X} + \mathbf{Z}_D$ を選ぶことによって得られる. ここで \mathbf{Z}_D は d 次元で, 実数値をとる確率変数ベクトルであり, \mathbf{X} と独立で以下の pdf をもつ.

$$f_{\mathbf{Z}_D}(\mathbf{z}) = \left(\frac{d}{r}\right)^{\frac{d}{r}-1} \frac{1}{V_d \Gamma(d/r) D^{\frac{d}{r}}} e^{-\frac{d}{rD} \|\mathbf{z}\|^r}, \quad \mathbf{z} \in \mathbb{R}^d \quad (3.13)$$

(1.1) かつ $E[\|\mathbf{Z}_D\|^r] = D$ より

$$\begin{aligned} R(D) &\leq I(\mathbf{X}; \mathbf{X} + \mathbf{Z}_D) \\ &= h(\mathbf{X} + \mathbf{Z}_D) - h(\mathbf{Z}_D) \end{aligned} \quad (3.14)$$

(3.13) は連続な分布に対するレート歪関数のシャノン下界の pdf であるため,

$$R_{SLB}(D) = h(\mathbf{X}) - h(\mathbf{Z}_D) \quad (3.15)$$

(3.14) と (3.15) を組み合わせて、以下を得る .

$$0 \leq R(D) - R_{SLB}(D) \leq h(\mathbf{X} + \mathbf{Z}_D) - h(\mathbf{X}) \quad (3.16)$$

(3.16) より、シャノン下界 $R_{SLB}(D)$ の漸近的最適性は以下を示すことで示される .

$$\overline{\lim}_{D \downarrow 0} h(\mathbf{X} + \mathbf{Z}_D) \leq h(\mathbf{X}) \quad (3.17)$$

(3.17) を達成するため、以下の pdf を持つ確率変数ベクトル \mathbf{Y}_D と \mathbf{Y}_0 を導入する .

$$f_{\mathbf{Y}_D}(\mathbf{y}) = \sum_{\mathbf{i} \in \mathbb{Z}^d} Pr(\lfloor \mathbf{X} + \mathbf{Z}_D \rfloor = \mathbf{i}) \mathbb{1}\{\mathbf{y} = \mathbf{i}\}, \mathbf{y} \in \mathbb{R}^d$$

$$f_{\mathbf{Y}_0}(\mathbf{y}) = \sum_{\mathbf{i} \in \mathbb{Z}^d} Pr(\lfloor \mathbf{X} \rfloor = \mathbf{i}) \mathbb{1}\{\mathbf{y} = \mathbf{i}\}, \mathbf{y} \in \mathbb{R}^d$$

ただし、 $\mathbb{1}\{\cdot\}$ は指示関数を表す . すなわち

$$\mathbb{1}\{\text{命題}\} = \begin{cases} 1 & (\text{命題が真}) \\ 0 & (\text{命題が偽}) \end{cases}$$

相対エントロピーの性質から以下が得られる .

$$D(f_{\mathbf{X}+\mathbf{Z}_D} \| f_{\mathbf{Y}_D}) = H(\lfloor \mathbf{X} + \mathbf{Z}_D \rfloor) - h(\mathbf{X} + \mathbf{Z}_D) \quad (3.18)$$

$$D(f_{\mathbf{X}} \| f_{\mathbf{Y}_0}) = H(\lfloor \mathbf{X} \rfloor) - h(\mathbf{X}) \quad (3.19)$$

ここで $D(f \| g)$ は f と g 、ふたつの pdf の間の相対エントロピー (KL ダイバージェンス) を表す ([5], (2.26)–(2.27) を参照) . 確率変数ベクトル \mathbf{Z}_D は、 $D^{1/r} \mathbf{Z}_1$ と同じ pdf をもつ . ただし、 \mathbf{Z}_1 は $D = 1$ のときの \mathbf{Z}_D を表す . \mathbf{X} と \mathbf{Z}_D は独立より、 D が 0 に近づくとき、 $\mathbf{X} + \mathbf{Z}_D \Rightarrow \mathbf{X}$ (分布収束) が成立する . さらに、 \mathbf{X} の分布がルベーグ測度に関して絶対連続で、集合 \mathbb{Z}^d が可算なので、確率 $Pr(\mathbf{X} \in \mathbb{Z}^d)$ は 0 である . そのため、以下が成立する .

$$\lim_{D \downarrow 0} Pr(\lfloor \mathbf{X} + \mathbf{Z}_D \rfloor = \mathbf{i}) = Pr(\lfloor \mathbf{X} \rfloor = \mathbf{i}), \mathbf{i} \in \mathbb{R}^d \quad (3.20)$$

つまり、 $f_{\mathbf{Y}_D}$ が点別に $f_{\mathbf{Y}_0}$ へと収束していることがいえ、このことは Scheffé の補題 ([6], Th.16.12 を参照) によって、 D が 0 に近づくとき、 $\mathbf{Y}_D \Rightarrow \mathbf{Y}_0$ であることを示している .

相対エントロピーの下半連続性 ([7] の補題 4 の証明) より

$$\underline{\lim}_{D \downarrow 0} D(f_{\mathbf{X}+\mathbf{Z}_D} \| f_{\mathbf{Y}_D}) \geq D(f_{\mathbf{X}} \| f_{\mathbf{Y}_0}) \quad (3.21)$$

これと (3.18)、(3.19) を合わせて

$$\underline{\lim}_{D \downarrow 0} \{H(\lfloor \mathbf{X} + \mathbf{Z}_D \rfloor) - h(\mathbf{X} + \mathbf{Z}_D)\} \geq H(\lfloor \mathbf{X} \rfloor) - h(\mathbf{X}) \quad (3.22)$$

さらに、[3] 中で証明されている次の補題を用いる .

補題2(Koch [3], Lemma 1)

X と Z は互いに独立で d 次元の確率変数ベクトルであるとする． $E[\|Z\|^r] < \infty$ と仮定すると，以下がいえる．

(i) $H(\lfloor X \rfloor) = \infty$ ならば，すべての $\epsilon > 0$ に対して $H(\lfloor X + \epsilon Z \rfloor) = \infty$

(ii) $H(\lfloor X \rfloor) < \infty$ かつ $Pr(\mathbf{X} \in \mathbb{Z}^d) = 0$ ならば，

$$\lim_{\epsilon \downarrow 0} H(\lfloor X + \epsilon Z \rfloor) = H(\lfloor X \rfloor) \quad (3.23)$$

Proof. [3] を参照． □

この補題2の(ii)より，

$$\begin{aligned} & \lim_{D \downarrow 0} H(\lfloor X + Z_D \rfloor) \\ &= \lim_{D \downarrow 0} H(\lfloor X + D^{1/r} Z_1 \rfloor) = H(\lfloor X \rfloor) \end{aligned} \quad (3.24)$$

(3.24) に (3.22) を組み合わせることで (3.17) が得られる．これより， $H(\lfloor X \rfloor) < \infty$ かつ $|h(\mathbf{X})| < \infty$ ならばシャノン下界が漸近的に達成されることが示された．

3.3 これまでの本研究の要旨と訂正内容について

以下では，これまでに考察してきた内容の要旨をまとめるとともに，その内容に一部正しくない仮定を用いた箇所があったため，その点の説明および訂正を行うものとする．なお，この節の内容は第41回情報理論とその応用シンポジウムで発表した内容にそれ以前の研究の経過を含めて修正をしたものになる．

3.3.1 結果までの流れ

自己相似分布に従う確率変数 X の量子化ベクトル $X_\delta = Q_\delta(X)$ が量子化写像 $Q_\delta(\cdot)$ によって離散確率分布 $P(x_\delta)$ およびレート歪関数 $R_\delta(D)$ を持つと仮定すると， $R_{SLB,\delta}(D)$ を以下のように評価することができる．ただしレート歪関数とシャノン下界の差の上界を準備するために， X_δ の再構成ベクトルは $\hat{X}_\delta = X_\delta + Z_{D,\delta}$ となるようにする．確率変数 $Z_{D,\delta}$ は関数 z_j を pmf に持つものとする．このとき任意の $\delta \in (0, 1)$ に関して自己相似分布の離散近似ができ，それぞれの離散点（代表点）に適当に番号を割り振ることができる．

こうして割り当てた番号を，もとの自己相似分布の台を離散近似した集合については $j = (1, 2, \dots, M)$ ，再構成された要素の集合については $k = (1, 2, \dots, N)$ とすれば，前述の離散分布に関するシャノン下界が利用できる．

$$R(D) = \lim_{\delta \downarrow 0} R_\delta(D) \quad (3.25)$$

\mathbf{X}_δ は離散確率分布 $P(\mathbf{x}_\delta)$ をもつ確率変数であるため

$$R_\delta(D) \leq I(\mathbf{X}_\delta; \mathbf{X}_\delta + \mathbf{Z}_{D,\delta}) \leq H(\mathbf{X}_\delta) \quad (3.26)$$

がいえる．一方，自己相似分布に従うレート歪関数のシャノン下界は

$$\begin{aligned} R_{SLB}(D) &\triangleq \lim_{\delta \downarrow 0} R_{SLB,\delta}(D) \\ &= \lim_{\delta \downarrow 0} \{H(\mathbf{X}_\delta) - H(\mathbf{Z}_{D,\delta})\} \end{aligned} \quad (3.27)$$

と定義することができる．このとき (3.25) および (3.27) より，以下の事実は直ちに確かめられる．

$\epsilon_1 > \epsilon_2 > \dots > 0$ となる $\{\epsilon_i\}_{i=1}^\infty$ を任意に固定する．各 $i = 1, 2, \dots, \infty$ について $\delta_i < \epsilon_i$ なら，ごく小さい任意の δ_i に対して

$$R(D) \leq \{R_{\delta_i}(D) + \frac{\epsilon_i}{2}\}$$

であり， $D_i := \delta_i^r D_r^*(1)$ のとき [9] より以下がいえる．

$$R_\delta(D) \leq \{H(\mathbf{X}_{\delta_i}) + \frac{\epsilon_i}{2}\} \quad (3.28)$$

またシャノン下界についても

$$\begin{aligned} R_{SLB}(D) &\geq \{R_{SLB,\delta_i}(D) - \frac{\epsilon_i}{2}\} \\ &= \{H(\mathbf{X}_{\delta_i}) - H(\mathbf{Z}_{D,\delta_i}) - \frac{\epsilon_i}{2}\} \end{aligned} \quad (3.29)$$

が成立し， $D \geq D_i$ のとき

$$R_{\delta_i}(D) \leq R_\delta(D_i) \leq H(\mathbf{X}_{\delta_i}) \quad (3.30)$$

より，十分大きな i について (3.28) から (3.29) を両辺とも引くことで，以下が成立する．

$$\begin{aligned} R_{\delta_i}(D) - R_{SLB,\delta_i}(D) &\leq H(\mathbf{X}_{\delta_i}) - H(\mathbf{X}_{\delta_i}) + H(\mathbf{Z}_{D,\delta_i}) + \epsilon_i \\ &= H(\mathbf{Z}_{D,\delta_i}) + \epsilon_i \end{aligned} \quad (3.31)$$

(3.31) より求める上界は $H(\mathbf{Z}_{D,\delta})$ ，すなわち $H(\{z_j\})$ を具体的に計算することで得られる．

これについては,

$$\begin{aligned}
H(\{z_j\}) &= -\sum_j \frac{e^{s\rho_{jk_0}}}{C} \log \frac{e^{s\rho_{jk_0}}}{C} \\
&= \log C - \frac{s}{C} \sum_j \rho_{jk_0} e^{s\rho_{jk_0}} \\
&= \log C - \frac{s}{C} \sum_{j:j \neq k_0} \rho_{jk_0} e^{s\rho_{jk_0}} \tag{3.32}
\end{aligned}$$

が成り立つ．ここで右辺第二項の総和の対象が j から $j \neq k_0$ となる j へと変わっているのは, $j = k_0$ のとき $\rho_{jk_0} = 0$ となるからである．

いま, (3.32) の右辺に現れる各項について評価をする．

まず第一項の $\log C$ については, [9] の結果が利用できる．“Cantor like” な集合を台に持つ自己相似分布を仮定すると, (2.20) より以下がいえる．

$$\log C < \infty \tag{3.33}$$

次に第二項については, 補題 1 を少し改変することで示せる以下の系を用いる．

系 1. (2.19) が成り立つとき, 任意の $0 < a^* < \frac{\varepsilon}{2}$ および $K > \frac{r+d^*}{r(a^*)^r}$ に対して

$$\begin{aligned}
& -\sup_{\delta > 0} s \sum_{j:j \neq k_0} \rho_{jk_0} e^{s\rho_{jk_0}} \\
& \leq 1 + \frac{D_r^*(1)(\min_l \alpha_l)^{-d^*} K e^{-K(a^*)^r}}{1 - \phi(rK(a^*)^r - r - d^*)} \\
& < \infty \tag{3.34}
\end{aligned}$$

が成り立つ．

Proof. 付録 A を参照． □

系 1 の結果より以下を示すことができる．

$$\begin{aligned}
-\frac{s}{C} \sum_{j:j \neq k_0} \rho_{jk_0} e^{s\rho_{jk_0}} & \leq -s \sum_{j:j \neq k_0} \rho_{jk_0} e^{s\rho_{jk_0}} \\
& < \infty \tag{3.35}
\end{aligned}$$

以上より,

$$\sup_{\delta_i > 0} H(\mathbf{Z}_{D, \delta_i}) < \infty \tag{3.36}$$

がいえる．

この結果は, 導入部分でも述べたように [9] などによって分かっていた結果と同様であり, エントロピー表現を用いた過程を経ても, 自己相似分布上のレート歪関数とシャノン下界の差が有界であるということを示せることが分かった．

3.3.2 修正前の研究において正しくなかった箇所

前節の内容は見直しをしたものであり、それまでは以下のような計算および結果を考えていた。

(3.32) 以降、一般の場合の具体的な計算は難しいので、以下のように特殊な場合に絞って考える。

まず1次元における全てのパスの長さが同一なコントロール集合を台に持つコントロール分布を考える。また歪関数として $\rho_{jk} = |x_j - y_k|^r$ を用いる。このとき、 $j = k$ ならば $\rho_{jk} = 0$ となるように対応関係をつくるものとする。はじめに、ある k に対する C について考える。

$$\begin{aligned} C &= \sum_j e^{s\rho_{jk_0}} \\ &= \sum_{j:j=k_0} e^{s\rho_{jk_0}} + \sum_{j:j \neq k_0} e^{s\rho_{jk_0}} \\ &= 1 + \sum_{j:j \neq k_0} e^{s\rho_{jk_0}} \end{aligned} \quad (3.37)$$

$$\leq 1 + 2 \sum_{n=1}^{\infty} e^{s|\ln \delta_i|^r} \quad (3.38)$$

$$= 1 + 2 \sum_{n=1}^{\infty} e^{-K_i n^r} \quad (3.39)$$

ただし、表記を簡潔にする目的で

$$K_i := -s(D_i)\delta_i^r \quad (3.40)$$

とおいた。ここまで単に s と表していたパラメータを $s(D_i)$ と表現したのは、パラメータ s が D_i の関数であることを強調するためである。また (3.37) から直ちに

$$C = 1 + \sum_{j:j \neq k_0} e^{s\rho_{jk_0}} > 1 \quad (3.41)$$

が得られる。式中に現れる1は、再構成アルファベットが生成アルファベットと同一の場合、すなわち $\rho_{jk} = 0$ の場合の結果として出現している。 $\rho_{jk} = 0$ 以外の場合については、1次元コントロール集合の離散近似を不均一な感覚で並ぶ点列ととらえて、均一な点列の間隔で級数の ρ_{jk} の部分を具体的に計算できる形にすることによって、不等式で評価できるようにした。この考え方は、(3.37) から (3.38) への変形などで用いていることを述べておく。

これらより ,

$$\begin{aligned}
 H(\{z_j\}) &= -\sum_j \frac{e^{s\rho_{jk_0}}}{C} \log \frac{e^{s\rho_{jk_0}}}{C} \\
 &= \log C - \frac{s}{C} \sum_j \rho_{jk_0} e^{s\rho_{jk_0}} \\
 &\leq \log C - \frac{s}{C} \sum_{j:j \neq k_0} \rho_{jk_0} e^{s\rho_{jk_0}} \\
 &\leq \log \left\{ 1 + 2 \sum_{n=1}^{\infty} e^{-K_i n^r} \right\} + 2K_i \delta_i^{-r} \sum_{n=1}^{\infty} |n\delta_i|^r e^{-K_i \delta_i^{-r} |n\delta_i|^r} \\
 &= \log \left\{ 1 + 2 \sum_{n=1}^{\infty} e^{-K_i n^r} \right\} + 2 \sum_{n=1}^{\infty} K_i n^r e^{-K_i n^r} \tag{3.42}
 \end{aligned}$$

と表せる . いま , (3.42) の右辺に現れる総和に関して以下の不等式がいえる . 任意の $r \geq 1$ および $K_i > 0$ について ,

$$\sum_{n=1}^{\infty} e^{-K_i n^r} \leq \sum_{n=1}^{\infty} e^{-K_i n} \tag{3.43}$$

$$\sum_{n=1}^{\infty} K_i n^r e^{-K_i n^r} \leq \sum_{n=1}^{\infty} K_i n e^{-K_i n} \tag{3.44}$$

が成り立つ . (3.43) および (3.44) を用いることで無限級数の和を計算することができ , (3.42) は以下のように表せる .

$$\begin{aligned}
 H(\{z_j\}) &\leq \log \left\{ 1 + 2 \sum_{n=1}^{\infty} e^{-K_i n} \right\} + 2 \sum_{n=1}^{\infty} K_i n e^{-K_i n} \\
 &= \log \left\{ 1 + \frac{2e^{-K_i}}{1 - e^{-K_i}} \right\} + \frac{2K_i e^{-K_i}}{(1 - e^{-K_i})^2} \tag{3.45}
 \end{aligned}$$

この部分までは表現として正しくない箇所はないが , この後に導入した仮定が適していなかった . その仮定についても以下で説明している .

さらに ,

$$\lim_{i \rightarrow \infty} \{-s(D_i)\delta_i^r\} = \lim_{i \rightarrow \infty} K_i = \infty \tag{3.46}$$

を満たす $\{(\delta_i, D_i)\}_{i=1}^{\infty}$ が存在すると仮定すると , (3.45) より

$$\begin{aligned}
 \lim_{i \rightarrow \infty} H(\{z_j\}) &\leq \lim_{i \rightarrow \infty} \left[\log \left\{ 1 + \frac{2e^{-K_i}}{1 - e^{-K_i}} \right\} + \frac{2K_i e^{-K_i}}{(1 - e^{-K_i})^2} \right] \\
 &= 0 \tag{3.47}
 \end{aligned}$$

である。(3.47) より, $D_i < \epsilon_i$ なる減少点列 $\{D_i\}_{i=1}^{\infty}$ において

$$\lim_{i \rightarrow \infty} \{R(D_i) - R_{SLB}(D_i)\} = 0 \quad (3.48)$$

が得られる。

この仮定の問題点は, そもそも工学的観点から見たときに δ_i が量子化変数であり, 自己相似分布を対象として評価を行っている以上, 変数としての δ_i の値はほとんど 0 に漸近していると考えて扱うべきところを, パラメータ s のほうがはるかに早く負の無限大に発散するとしている点にある。つまり, どの場合でも $-s \ll 1/\delta_i$ であるため, $-s(D_i)\delta_i^r$ が正の無限大に発散する状況は考えられないのである。

この仮定が存在しない場合についての条件だったため, (3.47) を利用して漸近的最適性を主張することが誤りにつながった。

以上のような経緯で, 自己相似分布上のレート歪関数とシャノン下界が漸近的にタイトであることの証明および条件の解明はできなかった。そのため次章にあたる本稿の結果では, 今回使用したシャノン下界がレート歪関数に対してタイトにならない条件について研究した結果を紹介し, レート歪関数とシャノン下界との間にギャップがあることを示した。

第 4 章

自己相似分布に対するレート歪関数とシャノン下界について

4.1 自己相似分布に対するレート歪関数とシャノン下界の漸近的性質

まず前章で述べたように，自己相似分布に対するレート歪関数とシャノン下界に関して，以下のような結果が得られた．

定理 4

自己相似分布に従う情報源が存在し，この生成アルファベット $\{x\}$ と再構成アルファベット $\{y\}$ の間の歪測度を $\rho(x, y)$ とする．“Cantor like” 集合を台に持つ 1 次元自己相似分布を仮定する．このとき，レート歪関数 $R(D)$ とシャノン下界 $R_{SLB}(D)$ の差の有界性，すなわちある正数 B について

$$R(D) - R_{SLB}(D) < B \quad (4.1)$$

であることが成立する．

この結果は有界性を述べたもので，自己相似分布に対するレート歪関数の解析を行った先行研究 [9] や [8] の結果と同様である．本研究では先に自己相似分布上のレート歪関数に対するシャノン下界を定義し，離散分布に近似されたそれを計算することで有界性を述べた．これは D が 0 に漸近しても同じ結果がいえる．

4.2 漸近的最適性が不成立となる反例

前節の結果を踏まえ，今度はレート歪関数とシャノン下界との間にある程度の差があると示すことを考える．以下では，自己相似分布の中でも特別な例として 1 次元の (1/3)

カントール分布を考える．十分大きな正の数 m' について，1次元 (1/3) カントール分布に従う確率変数 X の離散近似は，量子化関数 $Q'(\cdot)$ を用いて以下のように表せる．

$$Q'(X) = X_b = \sum_{i=1}^{m'} (-1)^{b_i} 3^{-i} \quad (4.2)$$

ただし， $b_i \in \{0, 1\}$ であるとする．1次元 (1/3) カントール分布は，名前の通り1次元 (1/3) カントール集合を台に持つ特異型確率分布であるため，その微細構造は2分木に対応させることができ，ある深さ l までの離散近似における標本点は深さ分の長さ l を持つ系列 $\mathbf{b}^{(l)} = b_1 b_2 \cdots b_l$ と一対一対応の関係にある．ただし， $l \ll m'$ であることに注意する．

一般のシャノン下界は (3.7) より

$$R_{SLB,\delta}(D) = H(P_{X_b}) + sD - \log C$$

と表せるが，この下界は導出の過程で $1/P_{X_b} \lambda_{X_b}$ を，パラメータ s を固定した元での定数 C において計算しているため，この部分を定数以上に自由に値を設定できるように改良することで，さらにタイトな下界を得ることができると考えられる．具体的には，まず

$$\lambda_{X_b} = \frac{1}{C_{\omega(\mathbf{b})} P_{X_b}}$$

を満たす $C_{\omega(\mathbf{b})}^{(l)}$ を定義し，これをシャノン下界における C のようにして整理していく．ただし，

$$C_{\omega(\mathbf{b})}^{(l)} = \begin{cases} a & (\omega(\mathbf{b}^{(l)}) = 0) \\ b & (\omega(\mathbf{b}^{(l)}) = 1) \end{cases}$$

かつ

$$\omega(\mathbf{b}^{(l)}) = \begin{cases} 0 & (\mathbf{b}^{(l)} \text{ is even parity}) \\ 1 & (\mathbf{b}^{(l)} \text{ is odd parity}) \end{cases}$$

であるとした．これらより $\lambda_{X_{b^{(l)}}}$ の制約は

$$\begin{aligned} & \sum_{\mathbf{b}^{(l)}} P_{X_{b^{(l)}}} \lambda_{b^{(l)}} e^{s\rho(X_{b^{(l)}}, Y)} \\ &= \frac{1}{a} \sum_{\omega(\mathbf{b}^{(l)})=0} e^{s\rho(X_{b^{(l)}}, Y)} + \frac{1}{b} \sum_{\omega(\mathbf{b}^{(l)})=1} e^{s\rho(X_{b^{(l)}}, Y)} \\ &\leq 1 \end{aligned} \quad (4.3)$$

と表せる．この制約式の右辺を最大化する Y_1 を以下のように定義する．

$$Y_1 \triangleq \arg \max_Y \left[\frac{1}{a} \sum_{\omega(\mathbf{b}^{(l)})=0} e^{s\rho(X_{b^{(l)}}, Y)} + \frac{1}{b} \sum_{\omega(\mathbf{b}^{(l)})=1} e^{s\rho(X_{b^{(l)}}, Y)} \right] \quad (4.4)$$

この Y_1 を用いて, (4.3) は等式として表すことができるようになる.

$$\begin{aligned}
& \sum_{X_{b^{(l)}}} P_{X_{b^{(l)}}} \lambda_{b^{(l)}} e^{s\rho(X_{b^{(l)}}, Y_1)} \\
&= \frac{1}{a} \sum_{\omega(b^{(l)})=0} e^{s\rho(X_{b^{(l)}}, Y_1)} + \frac{1}{b} \sum_{\omega(b^{(l)})} e^{s\rho(X_{b^{(l)}}, Y_1)} \\
&= \frac{C_0^{(l)}}{a} + \frac{C_1^{(l)}}{b} = 1
\end{aligned} \tag{4.5}$$

今, 一般のシャノン下界の導出と同様に (3.1) から計算して,

$$\begin{aligned}
R(D) &\geq sD + \sum_{X_{b^{(l)}}} P_{X_{b^{(l)}}} \log \lambda_{X_{b^{(l)}}} \\
&= H(P_{X_{b^{(l)}}}) + sD + \sum_{X_{b^{(l)}}} P_{X_{b^{(l)}}} \log P_{X_{b^{(l)}}} \lambda_{X_{b^{(l)}}} \\
&= H(P_{X_{b^{(l)}}}) + sD - \sum_{\omega(b^{(l)})=0} P_{X_{b^{(l)}}} \log a - \sum_{\omega(b^{(l)})=1} P_{X_{b^{(l)}}} \log b \\
&= H(P_{X_{b^{(l)}}}) + sD - P_0 \log a - P_1 \log b \\
&=: R_{SLB,2}(D)
\end{aligned} \tag{4.6}$$

となる. ここで, 簡単のために

$$P_{\omega(b^{(l)})} := \sum_{X_{b^{(l)}}} P_{X_{b^{(l)}}} \tag{4.7}$$

とおいた.

一般のシャノン下界においても第一項である所与の分布のシャノンエントロピー $H(P_{X_{b^{(l)}}})$ および第二項である sD は共通なので, 一般のシャノン下界 $R_{SLB,\delta}(D)$ とよりタイトな下界 $R_{SLB,2}(D)$ の大小関係は, 残りの部分の比較によって評価できる. これを式で表すと, 以下のようなになる.

$$\begin{aligned}
R_{SLB,2}(D) - R_{SLB,\delta}(D) &= -P_0 \log a - P_1 \log b + \log C \\
&= -P_0 \log a - P_1 \log b + (P_0 + P_1) \log(C_0^{(l)} + C_1^{(l)}) \\
&= -P_0 \log \frac{a}{C_0^{(l)} + C_1^{(l)}} - P_1 \log \frac{b}{C_0^{(l)} + C_1^{(l)}}
\end{aligned} \tag{4.8}$$

ここで,

$$\begin{aligned}
Q_0 &:= \frac{C_0^{(l)}}{a} \\
Q_1 &:= \frac{C_1^{(l)}}{b}
\end{aligned}$$

とおく． $Q_0 + Q_1 = 1$ であることに注意する．いま定義した Q_0 および Q_1 を用いて (4.8) から続きを計算していくと，

$$\begin{aligned}
R_{SLB,2}(D) - R_{SLB,\delta}(D) &= -P_0 \log \frac{C_0^{(l)}/Q_0}{C_0^{(l)} + C_1^{(l)}} - P_1 \log \frac{C_0^{(l)}/Q_0}{C_0^{(l)} + C_1^{(l)}} \\
&= -P_0 \log \left(\frac{C_0^{(l)}/Q_0}{C_0^{(l)} + C_1^{(l)}} \cdot \frac{P_0}{P_0} \right) - P_1 \log \left(\frac{C_0^{(l)}/Q_0}{C_0^{(l)} + C_1^{(l)}} \cdot \frac{P_1}{P_1} \right) \\
&= -P_0 \log \left(\frac{P_0}{Q_0} \right) - P_0 \log \left(\frac{C_0^{(l)}/(C_0^{(l)} + C_1^{(l)})}{P_0} \right) \\
&\quad - P_1 \log \left(\frac{P_1}{Q_1} \right) - P_1 \log \left(\frac{C_1^{(l)}/(C_0^{(l)} + C_1^{(l)})}{P_1} \right) \\
&= -D({}^dP \parallel {}^dQ) + D({}^dP \parallel \{C_0^{(l)}/(C_0^{(l)} + C_1^{(l)}), C_1^{(l)}/(C_0^{(l)} + C_1^{(l)})\})
\end{aligned} \tag{4.9}$$

のように整理できる．ただし式中の関数 $D(\cdot \parallel \cdot)$ は KL ダイバージェンスを表しており，その定義は以下のようにになっている．

$$D(P_A \parallel P_B) \triangleq \sum_x p_A(x) \log \frac{p_A(x)}{p_B(x)} \tag{4.10}$$

定義中の P_A および P_B は任意の離散確率分布を， $p_A(x)$ および $p_B(x)$ はそれぞれの確率を表している．ここで，(4.9) 中の dP や dQ は大文字の変数が確率分布であることを強調する目的で採用した．今，十分大きな有限長 l のビット列によって離散近似される 1 次元の (1/3) カントール分布において，ビット列 (パリティ) の偶奇は半数ずつ存在するため，確率分布 dP は ${}^dP \in \{1/2, 1/2\}$ だと考えてよい．また dQ は a と b の値を任意の正の実数値に決めてもよいため， ${}^dQ \in \{1/2, 1/2\}$ を実現する最適な a, b を設定できる．このことから ${}^dQ \in \{1/2, 1/2\}$ と考えてもよい．これらのことから (4.9) の第一項について $-D({}^dP \parallel {}^dQ) = 0$ とみることができ， $R_{SLB,2}(D) - R_{SLB,\delta}(D)$ の評価は $D({}^dP \parallel \{C_0^{(l)}/(C_0^{(l)} + C_1^{(l)}), C_1^{(l)}/(C_0^{(l)} + C_1^{(l)})\})$ に依ることが分かる．

以降では， $D({}^dP \parallel \{C_0^{(l)}/(C_0^{(l)} + C_1^{(l)}), C_1^{(l)}/(C_0^{(l)} + C_1^{(l)})\})$ を具体的に評価していく．まず，これ以降の計算上の表記が分かりやすくなるように以下のように定義しなおす．

$$C_0^{(l)}(s, 1/4) := C_0^{(l)} = \sum_{\omega(b^{(l)})=0} e^{s|X_{b^{(l)}} - \frac{1}{4}|} \tag{4.11}$$

$$C_1^{(l)}(s, 1/4) := C_1^{(l)} = \sum_{\omega(b^{(l)})=1} e^{s|X_{b^{(l)}} - \frac{1}{4}|} \tag{4.12}$$

ただし，計算上の都合で歪関数 $\rho(X_{b^{(l)}}, Y_1)$ を差の絶対値 $|X_{b^{(l)}} - Y_1|$ と仮定した．また上の定義において Y_1 が $\frac{1}{4}$ となっているのは，1 次元 (1/3) カントール集合における (4.4) の

左辺 [] の内部を最大化する点が $1/4$ だと考えられるためである。1次元 ($1/3$) カントール集合はユークリッド距離 1 の閉区間を対象として、再帰的 (かつ無限) に写像を作用させることでつくられるが、2分木の枝をジグザグにおりていった先を選ぶ方が $\exp(\rho(\cdot))$ が大きくなる。そしてその点は1次元の場合2点存在し、もとの閉区間を $[-1/2, 1/2]$ とすると、初項 $\pm 1/3$ 、公比 $-1/3$ の数列の収束する点を見ることに等しいので、上述の2点は $\pm 1/4$ となる。

以上のような定義をすることで、 $C_0^{(l)}(s, \pm 1/4) := C_0^{(l)}$ および $C_1^{(l)}(s, \pm 1/4) := C_1^{(l)}$ を評価することができる。具体的には、以下のような自己相似性を利用した再帰的計算が行える。

$$\begin{aligned}
C_0^{(l)}(s, 1/4) &= \sum_{\omega(\mathbf{b}^{(l)})=0} e^{s|X_{\mathbf{b}^{(l)}} - \frac{1}{4}|} \\
&= \sum_{b_1=0, \omega(\mathbf{b}^{(l-1)})=0} e^{s|3^{-1} + 3^{-1}X_{\mathbf{b}^{(l-1)}} - \frac{1}{4}|} + \sum_{b_1=1, \omega(\mathbf{b}^{(l-1)})=1} e^{s| -3^{-1} + 3^{-1}X_{\mathbf{b}^{(l-1)}} - \frac{1}{4}|} \\
&= \sum_{\omega(\mathbf{b}^{(l-1)})=0} e^{3^{-1}s|X_{\mathbf{b}^{(l-1)}} + \frac{1}{4}|} + \sum_{\omega(\mathbf{b}^{(l-1)})=1} e^{3^{-1}s|X_{\mathbf{b}^{(l-1)}} - \frac{7}{4}|} \\
&= C_0^{(l-1)}(3^{-1}s, -1/4) + C_1^{(l-1)}(3^{-1}s, 7/4) \tag{4.13}
\end{aligned}$$

ここで、式中にて

$$X_{\mathbf{b}^{(l)}} = (-1)^{b_1} \cdot 3^{-1} + 3^{-1} \cdot X_{\mathbf{b}^{(l-1)}} \tag{4.14}$$

の関係を用いたことに注意する。(4.13)と同様にして、 $C_0^{(l)}(s, -1/4)$, $C_1^{(l)}(s, 1/4)$, $C_1^{(l)}(s, -1/4)$ も以下のように計算できる。

$$C_0^{(l)}(s, -1/4) = C_0^{(l-1)}(3^{-1}s, -7/4) + C_1^{(l-1)}(3^{-1}s, 1/4) \tag{4.15}$$

$$C_1^{(l)}(s, 1/4) = C_0^{(l-1)}(3^{-1}s, 7/4) + C_1^{(l-1)}(3^{-1}s, -1/4) \tag{4.16}$$

$$C_1^{(l)}(s, -1/4) = C_0^{(l-1)}(3^{-1}s, 1/4) + C_1^{(l-1)}(3^{-1}s, -7/4) \tag{4.17}$$

(4.13), (4.15) から (4.17) の4つの式を各項ごとに都度使っていくことで、 $C_0^{(l)}(s, 1/4)$ および $C_1^{(l)}(s, 1/4)$ は以下のように表すことができる。

$$\begin{aligned}
C_0^{(l)}(s, 1/4) &= C_1^{(l-1)}(3^{-1}s, 7/4) + C_0^{(l-2)}(3^{-2}s, -7/4) \\
&\quad + C_0^{(l-3)}(3^{-3}s, 7/4) + C_1^{(l-4)}(3^{-4}s, -7/4) + \dots \\
&\quad + C_*^{(1)}(3^{-(l-1)}s, \pm 7/4) + C_*^{(1)}(3^{-(l-1)}s, \mp 1/4) \tag{4.18}
\end{aligned}$$

$$\begin{aligned}
C_1^{(l)}(s, 1/4) &= C_0^{(l-1)}(3^{-1}s, 7/4) + C_1^{(l-2)}(3^{-2}s, -7/4) \\
&\quad + C_1^{(l-3)}(3^{-3}s, 7/4) + C_0^{(l-4)}(3^{-4}s, -7/4) + \dots \\
&\quad + C_*^{(1)}(3^{-(l-1)}s, \pm 7/4) + C_*^{(1)}(3^{-(l-1)}s, \mp 1/4)
\end{aligned} \tag{4.19}$$

ただし，十分大きな正の数 l の偶奇によって最後の方の項は C_0 か C_1 かが変化するため，0 または 1 のどちらか 1 ビットを表す表現 $*$ ，および $*$ と反対の 1 ビットを表す表現 $\bar{*}$ を導入した．また第二引数の符号もこれに伴って正負が決まるため，最後の二項間で正負の関係性が整合するように表した．

上の 2 式は多くの項からなるが，右辺最後の一項以外は第二引数の絶対値が $7/4$ である．そのため C_0 および C_1 が最大となる $\pm 1/4$ から距離があり，全体からすると十分小さい値をとる．従って， $C_0^{(l)}(s, 1/4)$ および $C_1^{(l)}(s, 1/4)$ の支配的な項は最後の項であるといえる．このことから， $C_0^{(l)}(s, 1/4)$ と $C_1^{(l)}(s, 1/4)$ との比は

$$\begin{aligned}
C_0^{(l)} : C_1^{(l)} &\triangleq C_0^{(l)}(s, 1/4) : C_1^{(l)}(s, 1/4) \\
&\simeq C_*^{(1)}(3^{-(l-1)}s, \mp 1/4) : C_*^{(1)}(3^{-(l-1)}s, \mp 1/4)
\end{aligned} \tag{4.20}$$

と表せる．

ここで (4.20) の最右辺より， C_0 と C_1 との比にほとんど等しい関数 $C_0^{(1)}(\cdot, \cdot)$ および $C_1^{(1)}(\cdot, \cdot)$ の第一，第二引数はどちらも共通であることが分かる．この関数 $C_0^{(1)}(\cdot, \cdot)$ および $C_1^{(1)}(\cdot, \cdot)$ は長さ 1 のビットのアルファベット 0 と 1 のどちらかについて和をとるので，1 次元 $(1/3)$ カントール集合においては区間 $[-1/2, -1/6]$ と $[1/6, 1/2]$ のどちらにある点 $Y_1 = -1/4$ または $Y_1 = 1/4$ があるのかで大小関係が変化する．より具体的には，ある点 Y_1 は必ず区間 $[-1/2, -1/6]$ と $[1/6, 1/2]$ のどちらかに存在しており，その上である点 Y_1 が区間 $[1/6, 1/2]$ に存在すれば $C_0^{(1)}(\cdot, \cdot)$ が，点 Y_1 が区間 $[-1/2, -1/6]$ に存在すれば $C_1^{(1)}(\cdot, \cdot)$ が他方に比べて十分に大きくなる．これは関数形が \exp の肩の上にパラメータ s と歪関数 $\rho()$ の積の負数が乗っている形のため，ある点 Y_1 との距離が関数全体に大きく寄与するからである．

どちらが大きいかはある点 Y_1 にどちらを選ぶのか，また l 回の再帰的な展開操作の最後で関数 $C_0^{(1)}(\cdot, \cdot)$ および $C_1^{(1)}(\cdot, \cdot)$ の第二引数の符号がどちらになるのかに左右される．しかし，どちらが大きいにせよ，この 2 つの量は一方がもう一方に比べて十分大きくなることが分かった．この事実より，KL ダイバージェンス $D(dP || \{C_0^{(l)}/(C_0^{(l)} + C_1^{(l)}), C_1^{(l)}/(C_0^{(l)} + C_1^{(l)})\})$ は第二引数である分布 $\{C_0^{(l)}/(C_0^{(l)} + C_1^{(l)}), C_1^{(l)}/(C_0^{(l)} + C_1^{(l)})\}$ が $\{1 \cdot 2, 1/2\}$ でないことより 0 より大きな値を持つといえる．すなわち，

$$D(dP || \{C_0^{(l)}/(C_0^{(l)} + C_1^{(l)}), C_1^{(l)}/(C_0^{(l)} + C_1^{(l)})\}) > 0 \tag{4.21}$$

である．

以上のことから, 1次元 (1/3) カントール分布と絶対値歪に対する一般のシャノン下界 $R_{SLB,\delta}(D)$ とよりタイトな下界 $R_{SLB,2}(D)$ の差が

$$R_{SLB,2}(D) - R_{SLB,\delta}(D) \simeq D(dP||\{C_0^{(l)}/(C_0^{(l)} + C_1^{(l)}), C_1^{(l)}/(C_0^{(l)} + C_1^{(l)})\}) > 0 \quad (4.22)$$

である反例が示された.

第 5 章

まとめ

本稿では密度関数を持たない特異型分布確率分布のひとつである自己相似分布に関して、変数 δ を用いた離散近似を行い $\delta \downarrow 0$ の極限をとることで、非可算無限個の離散点を持つ離散分布の問題として解析を行った。この手法は川端-Dembo ら [9] によって行われた先行研究のものであり、それを利用してレート歪関数とシャノン下界の差をエントロピー表現で上界することはできた。しかし工学的観点からこの上界表現が 0 に収束する条件が存在しないことが分かり、本稿の結果では有界性を述べるに止まった。そこで視点を変えて、離散分布のレート歪関数の下界表現を一般のシャノン下界よりもタイトにしたときに、これと一般のシャノン下界とのギャップが生じるような場合と条件を導いた。これにより、計算に利用していたシャノン下界を工夫し、もっと別の表現や適した形にすることで生じていたギャップが解消され、漸近的最適性がいえるようになる可能性がみえた。あるいはこのギャップがどのようなときにも解消されないようであれば、離散近似を行うことによって漸近的最適性が示せなくなっている可能性を議論することができるようになるだろう。本研究の今後の課題として、うまく自己相似分布上のレート歪関数の下界をデザインすることが考えられる。また、自己相似分布特有の微細な構造に起因する再帰的な評価に関しても、事前に見通したほどうまく利用することができなかったため、この部分でもさらに改善が期待できると考える。

謝辞

最後に，本研究を行うにあたり大変なご多忙のなかお時間をいただき，著者が研究の課題や解決の糸口を自力で掴めるよう最後まで根気強くご指導やご指摘をしていただきました川端勉教授に心より感謝いたします．また，至るところでお世話になりました大濱靖匡教授，八木秀樹准教授，Santoso Bagus 助教に深く感謝いたします．

付録 A

系 1 の証明

ほとんどの前提条件を [8] から変えることなく，系 1 は証明することができる．以下では証明について詳しい内容を述べる．

Proof.

$$\begin{aligned}\hat{C}(K) &\triangleq - \sup_{\delta > 0} s \sum_{\mathbf{x}_\delta \in \Lambda_\delta} \rho(\mathbf{x}_\delta, \mathbf{y}) e^{s\rho(\mathbf{x}_\delta, \mathbf{y})} \\ &= \sup_{\delta > 0} \left[K \delta^{-r} \sum_{\mathbf{x}_\delta \in \Lambda_\delta} \rho(\mathbf{x}_\delta, \mathbf{y}) e^{-K\delta^{-r}\rho(\mathbf{x}_\delta, \mathbf{y})} \right]\end{aligned}$$

を定義する．ただし K は十分大きな正の定数である．上式の上界が有界であることを示すのが目標である．

任意の K および $a > 0$ において

$$B_{\delta, a, \mathbf{y}} \triangleq \{\mathbf{x}_\delta \in \Lambda_\delta \mid \rho(\mathbf{x}_\delta, \mathbf{y}) \leq a^r\} \quad (\text{A.1})$$

およびその Λ_δ 内の補集合 $B_{\delta, a, \mathbf{y}}^c$ を定義する． $a < \frac{\epsilon}{2}$ ならば $B_{\delta, a, \mathbf{y}}$ 内の全ての点の b_1 はひとつの値しかとらないことが，[9] によって示されている．ここで

$$\hat{f}(\delta, \mathbf{y}) \triangleq K \delta^{-r} \sum_{\mathbf{x}_\delta \in \Lambda_\delta} \rho(\mathbf{x}_\delta, \mathbf{y}) e^{-K\delta^{-r}\rho(\mathbf{x}_\delta, \mathbf{y})} \quad (\text{A.2})$$

$$\triangleq \hat{g}_1(\delta, a, \mathbf{y}) + \hat{g}_2(\delta, a, \mathbf{y}) \quad (\text{A.3})$$

とする．ただし，

$$\hat{g}_1(\delta, a, \mathbf{y}) \triangleq K \delta^{-r} \sum_{\mathbf{x}_\delta \in B_{\delta, a, \mathbf{y}}} \rho(\mathbf{x}_\delta, \mathbf{y}) e^{-K\delta^{-r}\rho(\mathbf{x}_\delta, \mathbf{y})} \quad (\text{A.4})$$

$$\hat{g}_2(\delta, a, \mathbf{y}) \triangleq K \delta^{-r} \sum_{\mathbf{x}_\delta \in B_{\delta, a, \mathbf{y}}^c} \rho(\mathbf{x}_\delta, \mathbf{y}) e^{-K\delta^{-r}\rho(\mathbf{x}_\delta, \mathbf{y})} \quad (\text{A.5})$$

である． $a^* < \frac{\epsilon}{2}$ を固定すると，一般性を失うことなく， $B_{\delta, a^*, \mathbf{y}}$ 内の全ての点が $b_1 = i$ と対応していると仮定できる．また $\mathbf{x} \in B_{\delta, a^*, \mathbf{y}}$ ならば $W_i^{-1}(\mathbf{x}) \in B_{\delta/\alpha_i, a^*/\alpha_i, \mathbf{y}/\alpha_i}$ である．すなわち，

$$\hat{g}_1(\delta, a^*, \mathbf{y}) = \hat{g}_1(\delta/\alpha_i, a^*/\alpha_i, W_i^{-1}\mathbf{y}) \quad (\text{A.6})$$

である．これについて $\delta' = \delta/\alpha_i$ ， $a' = a^*/\alpha_i$ ， $\mathbf{y}' = W_i^{-1}\mathbf{y}$ とおいて $g_1(\delta', a', \mathbf{y}')$ に再び同様の操作を繰り返すことができる．以上の操作を a' にあたる量が $\frac{\epsilon}{2}$ を下回るまで操作を繰り返すと，最終的に任意の δ および \mathbf{y} について $\psi \leq \frac{2a^*}{\epsilon}$ を満たす ψ が存在し，かつ

$$\begin{aligned} \hat{g}_1(\delta, a^*, \mathbf{y}) &= \hat{g}_1(\delta/\psi, a^*/\psi, \tilde{\mathbf{y}}) \\ &\leq \hat{f}(\delta/\psi, \tilde{\mathbf{y}}) \end{aligned} \quad (\text{A.7})$$

を満たすような点 $\tilde{\mathbf{y}}$ が存在するとわかる．また，

$$\hat{g}_2(\delta, a^*, \mathbf{y}) \leq |\Lambda_\delta| K \delta^{-r} D_r^*(1) e^{-K(a^*)^r \delta^{-r}} \quad (\text{A.8})$$

がいえる．ただし $|\Lambda_\delta|$ は集合 Λ_δ の要素数を表すものとする．以上より，

$$\begin{aligned} \check{h}(\delta) &\triangleq \sup_{\mathbf{y}} \hat{f}(\delta, \mathbf{y}) \\ &\leq \sup_{\psi \leq \frac{2a^*}{\epsilon}} [\check{h}(\delta/\psi)] + |\Lambda_\delta| K \delta^{-r} D_r^*(1) e^{-K(a^*)^r \delta^{-r}} \end{aligned} \quad (\text{A.9})$$

がいえる．ここで， $|\Lambda_\delta|$ については [9] 中で以下の上界が用意されたため，本証明もそれに倣うものとする．

$$|\Lambda_\delta| \leq (\min_l \{\alpha_l\} \delta)^{-d^*} \quad (\text{A.10})$$

いま， $\phi := \frac{2a^*}{\epsilon} < 1$ および $\hat{h}(z) \triangleq \check{h}(\phi^z)$ とおく．こうしたとき $\hat{C}(K) = \sup_{\delta > 0} [\check{h}(\delta)] = \sup_z [\hat{h}(z)]$ であることに注意する． $z > 0$ において (A.9) と (A.10) より，

$$\hat{h}(z) \leq (\min_l \{\alpha_l\} \phi^z)^{-d^*} K \phi^{-zr} D_r^*(1) e^{-K(a^*)^r \phi^{-zr}} + \sup_{z' \leq z-1} [\hat{h}(z')] \quad (\text{A.11})$$

が得られる．ここで (A.11) 右辺の第二項の \sup の条件がどのように得られたかについて説明する． $\phi^z/\psi =: \phi^{z'}$ とおく．一方 $\phi^{z-1} = \phi^z/\phi$ となるので $\phi := \frac{2a^*}{\epsilon} \geq \psi$ より

$$\phi^{z'} := \phi^z/\psi \geq \phi^z/\phi = \phi^{z-1} \quad (\text{A.12})$$

がいえ， $\phi < 1$ なので最終的に $z' \leq z-1$ となる．これらの関係式から $\psi = \phi^{z-z'}$ がいえることにも注意する．(A.11) は右辺の項の中に左辺の関数と同じものが引数を縮小して再帰的に登場するため，再帰的に上界を課し続けることを考えることができる．すると

z' の値が負になるが、これは $\delta > 1$ となることと等しい。 $\delta > 1$ のとき ($z < 0$ のとき) 集合 Λ_δ はひとつの点のみを要素として持つので $h(z) \leq 1$ である。 よって以下のように表せる。

$$\sup_z [\hat{h}(z)] \leq 1 + K(\min_l \alpha_l)^{-d^*} D_r^*(1) \sup_{z_0 \geq 0, z_n \geq z_{n-1} + 1} \sum_{n=0}^{\infty} \phi^{-z_n(r+d^*)} e^{-K(a^*)^r \phi^{-z_n r}} \quad (\text{A.13})$$

ここで $K > \frac{r+d^*}{r(a^*)^r}$ の条件下において、関数 $\phi^{-z(r+d^*)} e^{-K(a^*)^r \phi^{-zr}}$ は z に対して単調減少なので (A.13) の上界は系列 $z_n = n$ に対して得られ、

$$\begin{aligned} \hat{C}(K) &= \sup_z \hat{h}(z) \\ &\leq 1 + K(\min_l \alpha_l)^{-d^*} D_r^*(1) \sum_{n=0}^{\infty} \phi^{-n(r+d^*)} e^{-K(a^*)^r \phi^{-nr}} \end{aligned} \quad (\text{A.14})$$

と表せる。 また $\phi^{-nr} \geq -nr \ln \phi + 1$ が一般に成り立つため、最終的に以下のように書ける。

$$\begin{aligned} \hat{C}(K) &\leq 1 + K(\min_l \alpha_l)^{-d^*} D_r^*(1) \sum_{n=0}^{\infty} \phi^{-n(r+d^*)} e^{-K(a^*)^r (-nr \ln \phi + 1)} \\ &= 1 + \frac{D_r^*(1)(\min_l \alpha_l)^{-d^*} K e^{-K(a^*)^r}}{1 - \phi^{(rK(a^*)^r - r - d^*)}} \end{aligned} \quad (\text{A.15})$$

(A.15) の最右辺はいずれも有限の定数で構成されることから、

$$\hat{C}(K) < \infty \quad (\text{A.16})$$

であることが示された。

□

参考文献

- [1] C. E. Shannon, “Coding theorem for a discrete source with a fidelity criterion,” *IRE International Convention Record*, vol. 7, pp. 142–163, 1959.
- [2] T. Berger, *Rate Distortion Theory: Mathematical Basis for Data Compression*, ser. Electrical Engineering Series. Prentice Hall, 1971.
- [3] T. Koch, “The Shannon Lower Bound Is Asymptotically Tight”, *IEEE Trans. Inform. Theory*, vol. 62, No. 11, pp. 6155–6161, Nov. 2016.
- [4] T. Linder and R. Zamir, “On the asymptotic tightness of the Shannon lower bound,” *IEEE Trans. Inform. Theory*, vol. 40, No. 6, pp.2026–2031, Nov. 1994.
- [5] T. M. Cover and J. A. Thomas, “*Elements of Information Theory 2nd ed*”, Willey, 2006.
- [6] P. Billingsley, *Probability and Measure*, (Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics), 3rd ed. New York, NY, USA: Wiley, 1995.
- [7] I. Csiszár, “Arbitrarily varying channel with general alphabets and states, ” *IEEE Trans. Inform. Theory*, vol. 38, no. 6, pp. 1725–1742, Nov. 1992.
- [8] T. Kawabata, “Geometric and Stochastic Theories of Source Coding, ” Ph. D dissertation submitted to Univ. Tokyo, Jul. 1993.
- [9] T. Kawabata and A. Dembo, “The Rate-Distortion Dimension of Sets and Measures, ” *IEEE Trans. Inform. Theory*, vol. 40, No. 5, pp. 1564–1572, Sept. 1994.