

信州大学工学部

学士論文

ランダム符号化を用いない一般情報源に対する
レート・歪み理論の順定理

指導教員 西新 幹彦 准教授

学科 電気電子工学科
学籍番号 15T2006A
氏名 伊藤 佑樹

2019 年 2 月 26 日

目次

1	序論	1
1.1	はじめに	1
1.2	本論文の構成	2
2	数学的準備	2
3	一般のレート・歪み理論	2
3.1	歪み測度による固定長最大歪み符号化	3
3.2	確率的歪みによる固定長最大歪み符号化	10
4	歪み超過を許すレート・歪み理論	13
4.1	$D_\delta(\mathbf{X}, \mathbf{Y})$: 歪み測度による許容	13
4.2	$\overline{D}_\delta(\mathbf{X}, \mathbf{Y})$: 確率的歪みによる許容	16
4.3	$\delta_D(\mathbf{X}, \mathbf{Y})$: 歪み測度の超過確率による許容	19
4.4	$\overline{\delta}_D(\mathbf{X}, \mathbf{Y})$: 確率的歪みの超過確率による許容	21
4.5	歪み超過ゼロとしたときの δ -レート・歪み関数	24
4.6	歪み超過方法まとめ	27
5	まとめ	28
	謝辞	29
	参考文献	29
	付録 A ランダム符号化を用いるレート・歪み理論	30
A.1	一般のレート・歪み理論	30
A.2	定常無記憶情報源のレート・歪み理論	33
	付録 B ε -通信路符号化 (Feinstein の補題)	40
	付録 C 不等式の証明	49
C.1	$(1-x)^y \leq e^{-xy}$ の証明	49
C.2	$e^{-xy} \leq 1-y+e^{-x}$ の証明	49

1 序論

1.1 はじめに

1948 年 Shannon により情報源符号化定理と通信路符号化定理の二つの符号化定理が発表された。情報源符号化とは情報源の記号からアルファベットへの写像、つまり情報源圧縮のことである。誤りなく圧縮する際の必要条件として、平均符号語長は情報源のエントロピーよりも小さくなくてはいけない。これを情報源符号化定理という。そして通信路符号化定理とは、通信路容量よりも低いレートで送れば限りなく誤りが小さくできるというものである。

一般的に情報源符号化問題が意味を持つのは情報源アルファベットが有限また可算無限のときに限られる。これに対して、実数値などを出力する連続情報源の場合、出力された元のデータ $\mathbf{x}$ と符号化のあと復号されたデータ $\mathbf{y}$ とがほとんど一致ようにすることは有限の符号化レートをもつ符号化のもとでは不可能である。この場合にも意味のある符号化を行うためには、 $\mathbf{x}$ と $\mathbf{y}$ との間にある許容範囲内の有限歪みが生じることを許さざるを得ない。したがって予め与えられた歪み基準の範囲内で符号化レートをどこまで小さくできるかというレート・歪み理論が必要となる [1]。本論文では一般情報源に対するレート・歪み問題のうち、固定長最大歪み符号化問題を主に考える。

情報理論において、符号化定理の順定理の証明方法は次のようなものがある。定常無記憶通信路、一般通信路ともに通信路符号化定理ではランダム符号化を用いる手法と非ランダム符号化で行う手法の 2 つが知られている [1]。レート・歪み理論の定常無記憶情報源でもランダム符号化を用いる手法と非ランダム符号化を用いる 2 通りがある [2]。しかし一般情報源に対してはランダム符号化を用いる手法は存在する [1] が、非ランダム符号化の方法は見当たらない。

通信路符号化定理での非ランダム符号化を用いる証明には符号構成に Feinstein の補題が用いられている。それは通信路の出力領域を互いに素な復号化領域で分割していきながら符号を作成していくという方法である。この方法は cookie cutting という名称で知られている。さらに非ランダム符号での定常無記憶情報源におけるレート・歪み理論の証明には、Csiszár and Körner により、逆向き通信路に対して復号語を作成していくという方法でされている [2]。

本研究では、非ランダムな方法で符号の存在を示す補題を導出した。これは Feinstein の補題に似ており cookie cutting を用いて通信路の入力領域を分割するという方法で作成した。この補題は一般情報源に適用可能であり、さらにレート・歪み理論の定義や達成可能性の定義とは完全に独立している。本論文では、よく知られたレート・歪み問題の順定理をこの補題を用いて証明する。さらに同じ補題を用いて、微小な歪み超過を許すことによって達成可能性を緩和した場合の符号化定理を証明する。

1.2 本論文の構成

以降、第2章では情報源符号化について、数学的に準備をする。第3章では本研究の主な成果である非ランダムで符号を構成する補題を導き、一般情報源のレート・歪み理論を証明する。第4章では補題の活用例として、一定の歪み超過を許すレート・歪み問題に関しての例をあげる。その際、歪みの定義を4種類用意し、3章との関係について考察する。

また付録では本研究の背景、参考になった定理をまとめる。

2 数学的準備

本論文では一般情報源 [1] を考える。情報源アルファベットを $\mathcal{X}$ とする。一般情報源 $\mathbf{X}$ は $\mathcal{X}^n$ 上の任意の確率分布 P_{X^n} をもつ確率変数 X^n の列 $\{X^n\}_{n=1}^{\infty}$ として定義される。一般に、各 n に対して X^n は個別の確率空間で定義される。一般情報源は極めて一般的であり、すべての非定常や非エルゴードな情報源を含んでいる。

さらにここで実数値確率変数列 $\{Z_n\}_{n=1}^{\infty}$ に対して次を定義する。[1]

定義 1 確率的上極限

$$p\text{-}\limsup_{n \rightarrow \infty} Z_n \triangleq \inf \left\{ \alpha \mid \lim_{n \rightarrow \infty} \Pr\{Z_n > \alpha\} = 0 \right\} \quad (1)$$

定義 2 確率的下極限

$$p\text{-}\liminf_{n \rightarrow \infty} Z_n \triangleq \sup \left\{ \beta \mid \lim_{n \rightarrow \infty} \Pr\{Z_n < \beta\} = 0 \right\} \quad (2)$$

さらに二つの確率変数 $\mathbf{X}, \mathbf{Y}$ について次を定義する。

定義 3 確率変数に対して

$$\bar{I}(\mathbf{X}; \mathbf{Y}) \triangleq p\text{-}\limsup_{n \rightarrow \infty} \frac{1}{n} \log \frac{P_{Y^n|X^n}(Y^n|X^n)}{P_{Y^n}(Y^n)} \quad (3)$$

とし、 $(\mathbf{X}, \mathbf{Y})$ の相互情報量スペクトル上限と呼ぶ。ここで $P_{Y|X}(y|x)$ は $X = x$ が与えられたときの $Y = y$ の条件付確率を表す。

3 一般のレート・歪み理論

まず歪みについて定義をし、一般情報源に対するレート・歪み理論を確認する。情報源アルファベットを $\mathcal{X}$ 、復号アルファベットを $\mathcal{Y}$ とする。歪み測度 d を各 n に対し、関数 $d_n : \mathcal{X}^n \times \mathcal{Y}^n \rightarrow [0, +\infty)$ と定め、 $d_n(\mathbf{x}, \mathbf{y})$ ($\mathbf{x} \in \mathcal{X}^n, \mathbf{y} \in \mathcal{Y}^n$) を $\mathbf{x}$ と $\mathbf{y}$ の間の歪みと呼ぶ。

さらに符号化は次のように行う． M_n 個の符号語の集合を $\mathcal{M}_n = \{1, 2, \dots, M_n\}$ とする．情報源系列 $\mathbf{x} \in \mathcal{X}^n$ は符号器 $\varphi_n : \mathcal{X}^n \rightarrow \mathcal{M}_n$ で符号化される．さらに復号化は復号器 $\psi_n : \mathcal{M}_n \rightarrow \mathcal{Y}^n$ で行われる．ここで $\varphi_n(\mathbf{x})$ を $\mathbf{x}$ の符号語， $\psi_n(m)$ を m の復号語という．

3.1 歪み測度による固定長最大歪み符号化

定義 4 $(\mathbf{X}, \mathbf{Y})$ の最大歪みを

$$D(\mathbf{X}, \mathbf{Y}) \triangleq p\text{-}\limsup_{n \rightarrow \infty} \frac{1}{n} d_n(X^n, Y^n) \quad (4)$$

と定義し， $(\mathbf{X}, \mathbf{Y})$ の歪みスペクトル上限ともいう．

達成可能性を次のように定める．

定義 5 レート R が D -達成可能であるとは

$$D(\mathbf{X}, \mathbf{Y}) \leq D \quad (5)$$

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log M_n \leq R \quad (6)$$

を満たすような符号の列 $\{(\varphi_n, \psi_n)\}_{n=1}^{\infty}$ が存在するということである．ただしここで $Y^n = \psi_n(\varphi_n(X^n))$ とする．

定義 6 レート・歪み関数は以下のように定義される．

$$R(D) = \inf\{R \mid R \text{ が } D\text{-達成可能}\} \quad (7)$$

ここで本論文の主たる結果を導く補題を示す．

補題 1 任意の一般情報源 $\mathbf{X} = \{X^n\}_{n=1}^{\infty}$ に対し，それと相関のある $\mathbf{Y} = \{Y^n\}_{n=1}^{\infty}$ を考える． R_n ， γ を任意に与えられた正の定数としたとき，逆向き通信路 $\{\hat{\mathbf{W}} : \mathcal{Y} \rightarrow \mathcal{X}\}$ に対して

$$\hat{W}_n(\mathbf{x}|\mathbf{y}) \triangleq \Pr\{X^n = \mathbf{x} \mid Y^n = \mathbf{y}\} \quad (8)$$

$$\mathcal{B}_n(\mathbf{y}) \triangleq \left\{ \mathbf{x} \mid \frac{1}{n} \log \frac{\hat{W}_n(\mathbf{x}|\mathbf{y})}{P_{X^n}(\mathbf{x})} \leq R_n - \gamma \right\} \quad (9)$$

とし， $\mathcal{S}_n(\mathbf{y})$ を $\mathbf{y} \in \mathcal{Y}^n$ を引数とする $\mathcal{X}^n$ の任意の部分集合とすると

$$\frac{1}{n} \log M_n \leq R_n \quad (10)$$

$$\Pr\{X^n \notin \mathcal{S}_n(\psi_n(\varphi_n(X^n)))\} \leq \Pr\{X^n \notin \mathcal{B}_n(Y^n)\} + \Pr\{X^n \notin \mathcal{S}_n(Y^n)\} + e^{-n\gamma} \quad (11)$$

を満たす符号 (φ_n, ψ_n) が存在する．

(証明)

$\mathcal{T}_n(\mathbf{y}) \triangleq \mathcal{B}_n(\mathbf{y}) \cap \mathcal{S}_n(\mathbf{y})$ とおく．復号語を以下のように順次選んでいく．最初に条件

$$\hat{W}^n(\mathcal{T}_n(\mathbf{y})|\mathbf{y}) \geq e^{-n\gamma} \quad (12)$$

を満たす $\mathbf{y} \in \mathcal{Y}^n$ を任意に選び $\psi_n(1)$ とし, $\varphi_n^{-1}(1) \triangleq \mathcal{T}_n(\psi_n(1))$ とおく．このような操作を繰り返して行い, m 回目には

$$\hat{W}^n(\mathcal{T}_n(\mathbf{y}) \setminus \bigcup_{m' < m} \varphi_n^{-1}(m')|\mathbf{y}) \geq e^{-n\gamma} \quad (13)$$

となる $\mathbf{y} \in \mathcal{Y}^n$ を任意に選び, $\psi_n(m)$ とし復号語にする．さらに

$$\varphi_n^{-1}(m) \triangleq \mathcal{T}_n(\psi_n(m)) \setminus \bigcup_{m' < m} \varphi_n^{-1}(m') \quad (14)$$

とする．可能な限りこの操作を続け, M_n 個目まで復号語をとれたとする．ここで $\varphi_n^{-1}(m)$, $m = 1, \dots, M_n$ が互いに素であることを注意し

$$\mathcal{D} \triangleq \bigcup_m \varphi_n^{-1}(m) \quad (15)$$

とおき, 符号化領域と呼ぶ．すると M_n の定義より任意の $\mathbf{y} \in \mathcal{Y}^n$ に対して

$$\hat{W}^n(\mathcal{T}_n(\mathbf{y}) \setminus \mathcal{D}|\mathbf{y}) < e^{-n\gamma} \quad (16)$$

が成り立つ．

以上のように定義された (φ_n, ψ_n) 符号の性能を評価する．まず符号化レートを調べる．ここで $\mathbf{x} \in \varphi_n^{-1}(m)$ ならば $\mathbf{x} \in \mathcal{T}_n(\psi_n(m))$ なので $\mathcal{T}_n(\mathbf{y})$ の定義より $\mathbf{x} \in \mathcal{B}_n(\psi_n(m))$ となって式 (9) より

$$\frac{1}{n} \log \frac{\hat{W}^n(\mathbf{x}|\psi_n(m))}{P_{X^n}(\mathbf{x})} \leq R_n - \gamma \quad (17)$$

が得られる．すなわち, $\mathbf{x} \in \varphi_n^{-1}(m)$ ならば

$$P_{X^n}(\mathbf{x}) \geq \hat{W}^n(\mathbf{x}|\psi_n(m))e^{-n(R_n - \gamma)} \quad (18)$$

が成り立つことに注意しておく．ここで式 (13), (15), (18) より

$$1 \geq P_{X^n}(\mathcal{D}) \quad (19)$$

$$= P_{X^n}(\bigcup_m \varphi_n^{-1}(m)) \quad (20)$$

$$= \sum_m P_{X^n}(\varphi_n^{-1}(m)) \quad (21)$$

$$= \sum_m \sum_{\mathbf{x} \in \varphi_n^{-1}(m)} P_{X^n}(\mathbf{x}) \quad (22)$$

$$\geq \sum_m \sum_{\mathbf{x} \in \varphi_n^{-1}(m)} \hat{W}^n(\mathbf{x} | \psi_n(m)) e^{-n(R_n - \gamma)} \quad (23)$$

$$= \sum_m \hat{W}^n(\varphi_n^{-1}(m) | \psi_n(m)) e^{-n(R_n - \gamma)} \quad (24)$$

$$\geq \sum_m e^{-nR_n} \quad (25)$$

$$= M_n e^{-nR_n} \quad (26)$$

となり

$$\frac{1}{n} \log M_n \leq R_n \quad (27)$$

が得られる．

次に式 (11) を導く． $\mathbf{x} \in \mathcal{D}$ とすると，ある m に対して $\mathbf{x} \in \varphi_n^{-1}(m)$ となる．したがって $\mathbf{x} \in \mathcal{T}_n(\psi_n(m))$ となるので， $\mathcal{T}_n(\mathbf{y})$ の定義より $\mathbf{x} \in \mathcal{S}_n(\psi_n(m))$ となる．一方 $\varphi_n(\mathbf{x}) = m$ なので， $\mathbf{x} \in \mathcal{S}_n(\psi_n(\varphi_n(\mathbf{x})))$ となる．したがって

$$\Pr\{X^n \in \mathcal{D}\} \leq \Pr\{X^n \in \mathcal{S}_n(\psi_n(\varphi_n(\mathbf{x})))\} \quad (28)$$

といえる．一方，式 (16) により任意の $\mathbf{y}$ について

$$\hat{W}^n(\mathcal{T}_n(\mathbf{y}) | \mathbf{y}) \quad (29)$$

$$= \hat{W}^n(\mathcal{T}_n(\mathbf{y}) \cap \mathcal{D} | \mathbf{y}) + \hat{W}^n(\mathcal{T}_n(\mathbf{y}) \setminus \mathcal{D} | \mathbf{y}) \quad (30)$$

$$\leq \hat{W}^n(\mathcal{D} | \mathbf{y}) + e^{-n\gamma} \quad (31)$$

となる．すると式 (28) の左辺は式 (31) を用いることにより

$$\Pr\{X^n \in \mathcal{D}\} \quad (32)$$

$$= P_{X^n}(\mathcal{D}) \quad (33)$$

$$= \sum_{\mathbf{y}} P_{X^n Y^n}(\mathcal{D}, \mathbf{y}) \quad (34)$$

$$= \sum_{\mathbf{y}} P_{Y^n}(\mathbf{y}) \hat{W}_n(\mathcal{D}|\mathbf{y}) \quad (35)$$

$$\geq \sum_{\mathbf{y}} P_{Y^n}(\mathbf{y}) (\hat{W}_n(\mathcal{T}_n(\mathbf{y})|\mathbf{y}) - e^{-n\gamma}) \quad (36)$$

$$= \sum_{\mathbf{y}} \Pr\{X^n \in \mathcal{T}_n(Y^n), Y^n = \mathbf{y}\} - e^{-n\gamma} \quad (37)$$

$$= \Pr\{X^n \in \mathcal{T}_n(Y^n)\} - e^{-n\gamma} \quad (38)$$

$$= \Pr\{X^n \in \mathcal{B}_n(Y^n)\} \quad (39)$$

$$= 1 - \Pr\{X^n \notin \mathcal{B}_n(Y^n)\} \quad (40)$$

$$= 1 - \Pr\{X^n \notin \mathcal{S}_n(Y^n)\} - e^{-n\gamma} \quad (41)$$

と評価でき，式 (28) と式 (41) より

$$\Pr\{X^n \in \mathcal{S}_n(\psi_n(\varphi_n(\mathbf{x})))\} \geq 1 - \Pr\{X^n \notin \mathcal{B}_n(Y^n)\} - \Pr\{X^n \notin \mathcal{S}_n(Y^n)\} - e^{-n\gamma} \quad (42)$$

となるので

$$\Pr\{X^n \notin \mathcal{S}_n(\psi_n(\varphi_n(X^n)))\} \leq \Pr\{X^n \notin \mathcal{B}_n(Y^n)\} + \Pr\{X^n \notin \mathcal{S}_n(Y^n)\} + e^{-n\gamma} \quad (43)$$

が導かれる．□

この補題は任意の集合に対してのものであり，一般情報源にも適用可能である．さらにレート・歪み問題に関わらず用いることが可能である．本論文ではレート・歪み理論の様々な問題に関してこの補題を適用し，有用性を示す．

このあと述べる定理 1 の順定理は補題 1 によって示される．逆定理を示すために次の既存の補題が使われる．

補題 2 ([1] 補題 2.4) 情報源 $\mathbf{X} = \{X^n\}_{n=1}^{\infty}$ が正数 M_n に対して

$$|\{\mathbf{x} \in \mathcal{X} | P_{X^n}(\mathbf{x}) > 0\}| \leq M_n \quad (44)$$

を満たせば、任意の定数 $\gamma > 0$ に対して

$$\Pr \left\{ \frac{1}{n} \log \frac{1}{P_{X^n}(X^n)} \geq \frac{1}{n} \log M_n + \gamma \right\} \leq e^{-n\gamma} \quad (45)$$

が成立する。

(証明)[1]

$$\mathcal{T}_n \triangleq \left\{ \mathbf{x} \in \mathcal{X}^n \mid P_{X^n}(\mathbf{x}) > 0 \text{ かつ } \frac{1}{n} \log \frac{1}{P_{X^n}(\mathbf{x})} \geq \frac{1}{n} \log M_n + \gamma \right\} \quad (46)$$

とおけば、 $\mathbf{x} \in \mathcal{T}_n$ のとき $P_{X^n}(\mathbf{x}) \leq \frac{1}{M_n e^{n\gamma}}$ である。仮定した条件式 (44) より $|\mathcal{T}_n| \leq M_n$ となる。したがって

$$\Pr \left\{ \frac{1}{n} \log \frac{1}{P_{X^n}(X^n)} \geq \frac{1}{n} \log M_n + \gamma \right\} \quad (47)$$

$$= \sum_{\mathbf{x} \in \mathcal{T}_n} P_{X^n}(\mathbf{x}) \quad (48)$$

$$\leq \sum_{\mathbf{x} \in \mathcal{T}_n} \frac{1}{M_n e^{n\gamma}} \quad (49)$$

$$\leq e^{-n\gamma} \quad (50)$$

が得られる。□

ではランダム符号化を用いしないで作成した補題で一般のレート・歪み関数を導く。なおランダム符号化を用いた場合の定常無記憶情報源と一般情報源の証明も付録 A に示す。

定理 1 一般情報源 $\mathbf{X}$ のレート・歪み関数は

$$R(D) = \inf_{Y: D(\mathbf{X}, \mathbf{Y}) \leq D} \bar{I}(\mathbf{X}; \mathbf{Y}) \quad (51)$$

で与えられる。

(証明)

順定理

$D(\mathbf{X}, \mathbf{Y}) \leq D$ を満たす任意の $\mathbf{Y}$ を考える。任意の自然数 k に対して

$$\lim_{n \rightarrow \infty} \Pr \left\{ \frac{1}{n} d_n(X^n, Y^n) > D + \frac{1}{k} \right\} = 0 \quad (52)$$

である。ゆえに十分大きなすべての n で

$$\Pr \left\{ \frac{1}{n} d_n(X^n, Y^n) > D + \frac{1}{k} \right\} < \frac{1}{k} \quad (53)$$

が満たされる．このとき， n に対して，式 (53) を満たす最小の k が決まる．その k に対して， $\gamma_n \triangleq \frac{1}{k}$ とおく．すると $n \rightarrow \infty$ のとき $\gamma_n \rightarrow 0$ かつ

$$\lim_{n \rightarrow \infty} \Pr \left\{ \frac{1}{n} d_n(X^n, Y^n) > D + \gamma_n \right\} = 0 \quad (54)$$

が成り立つ（対角線論法）．

ここで γ を任意の正の定数として

$$R_n \triangleq \bar{I}(\mathbf{X}; \mathbf{Y}) + 2\gamma \quad (55)$$

$$\mathcal{S}_n(\mathbf{y}) \triangleq \left\{ \mathbf{x} \mid \frac{1}{n} d_n(\mathbf{x}, \mathbf{y}) \leq D + \gamma_n \right\} \quad (56)$$

として補題 1 を用いる．すると，ある (φ_n, ψ_n) 符号が存在し，

$$\frac{1}{n} \log M_n \leq \bar{I}(\mathbf{X}; \mathbf{Y}) + 2\gamma \quad (57)$$

$$\begin{aligned} & \Pr \left\{ \frac{1}{n} d_n(X^n, \psi_n(\varphi_n(X^n))) > D + \gamma_n \right\} \\ & \leq \Pr \left\{ \frac{1}{n} \log \frac{\hat{W}_n(X^n|Y^n)}{P_{X^n}(X^n)} > \bar{I}(\mathbf{X}; \mathbf{Y}) + \gamma \right\} + \Pr \left\{ \frac{1}{n} d_n(X^n, Y^n) > D + \gamma_n \right\} + e^{-n\gamma} \end{aligned} \quad (58)$$

が成り立つ．つまり式 (57) より

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log M_n \leq \bar{I}(\mathbf{X}; \mathbf{Y}) + 2\gamma \quad (59)$$

となり，定義 3 と式 (54) と式 (58) より

$$\lim_{n \rightarrow \infty} \Pr \left\{ \frac{1}{n} d_n(X^n, \psi_n(\varphi_n(X^n))) > D + \gamma_n \right\} = 0 \quad (60)$$

が導かれる．したがって確率的上極限の定義より

$$\text{p-lim sup}_{n \rightarrow \infty} \left(\frac{1}{n} d_n(X^n, \psi_n(\varphi_n(X^n))) - \gamma_n \right) \leq D \quad (61)$$

となるが $\gamma_n \rightarrow 0$ より

$$\text{p-lim sup}_{n \rightarrow \infty} \frac{1}{n} d_n(X^n, \psi_n(\varphi_n(X^n))) \leq D \quad (62)$$

を得る．よって式 (59) と式 (62) によって $\bar{I}(\mathbf{X}; \mathbf{Y}) + 2\gamma$ は D -達成可能である．つまり

$$R(D) \leq \bar{I}(\mathbf{X}; \mathbf{Y}) + 2\gamma \quad (63)$$

であるが, $\gamma > 0$ は任意なので

$$R(D) \leq \bar{I}(\mathbf{X}; \mathbf{Y}) \quad (64)$$

となる. さらに $\mathbf{Y}$ の任意性より

$$R(D) \leq \inf_{\mathbf{Y}: D(\mathbf{X}; \mathbf{Y}) \leq D} \bar{I}(\mathbf{X}; \mathbf{Y}) \quad (65)$$

が成り立つ. これで定理の順定理は導かれた.

逆定理 ([1] 定理 5.3)

レート R が D -達成可能であると仮定する. つまり

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log M_n \leq R \quad (66)$$

$$D(\mathbf{X}, \mathbf{Y}) \leq D \quad (67)$$

$$\mathbf{Y} = \{\psi_n(\varphi_n(X^n))\}_{n=1}^{\infty} \quad (68)$$

となる符号の列 $\{(\varphi_n, \psi_n)\}_{n=1}^{\infty}$ が存在する. ここで $\mathbf{Y}$ は符号器 φ_n がとる値の数より少ないことに注意して, 補題 2 より任意の正の定数 γ に対し

$$\Pr \left\{ \frac{1}{n} \log \frac{1}{P_{Y^n}(Y^n)} \geq \frac{1}{n} \log M_n + \gamma \right\} \leq e^{-n\gamma} \quad (69)$$

が成り立つ. $1 \geq P_{Y^n|X^n}(Y^n|X^n)$ より

$$\Pr \left\{ \frac{1}{n} \log \frac{P_{Y^n|X^n}(Y^n|X^n)}{P_{Y^n}(Y^n)} \geq \frac{1}{n} \log M_n + \gamma \right\} \leq e^{-n\gamma} \quad (70)$$

となる. 式 (66) より十分大きなすべての n で

$$\frac{1}{n} \log M_n \leq R + \gamma \quad (71)$$

となるので式 (70) は

$$\Pr \left\{ \frac{1}{n} \log \frac{P_{Y^n|X^n}(Y^n|X^n)}{P_{Y^n}(Y^n)} \geq R + 2\gamma \right\} \leq e^{-n\gamma} \quad (72)$$

となる. よって両辺極限をとると

$$\lim_{n \rightarrow \infty} \Pr \left\{ \frac{1}{n} \log \frac{P_{Y^n|X^n}(Y^n|X^n)}{P_{Y^n}(Y^n)} \geq R + 2\gamma \right\} = 0 \quad (73)$$

となることから

$$\inf \left\{ \alpha \left| \lim_{n \rightarrow \infty} \Pr \left\{ \frac{1}{n} \log \frac{P_{Y^n|X^n}(Y^n|X^n)}{P_{Y^n}(Y^n)} \geq \alpha \right\} = 0 \right. \right\} \leq R + 2\gamma \quad (74)$$

となるので

$$\bar{I}(\mathbf{X}; \mathbf{Y}) \leq R + 2\gamma \quad (75)$$

を得る． γ は任意だったので

$$\bar{I}(\mathbf{X}; \mathbf{Y}) \leq R \quad (76)$$

となり，したがってレート・歪み関数は

$$R(D) \geq \bar{I}(\mathbf{X}; \mathbf{Y}) \quad (77)$$

を満たさなければいけない．ここで $\mathbf{Y}$ は $D(\mathbf{X}, \mathbf{Y}) \leq D$ を満たすことから

$$R(D) \geq \inf_{\mathbf{Y}: D(\mathbf{X}, \mathbf{Y}) \leq D} \bar{I}(\mathbf{X}; \mathbf{Y}) \quad (78)$$

が成り立つ．これで逆定理も導かれた．

順定理・逆定理より

$$R(D) = \inf_{\mathbf{Y}: D(\mathbf{X}, \mathbf{Y}) \leq D} \bar{I}(\mathbf{X}; \mathbf{Y}) \quad (79)$$

を満たす．□

以上がランダム符号を用いずにレート・歪み理論を導く方法である．補題 1 により非ランダム符号で証明することができた．ここからは補題 1 を用いてさらにレート・歪み問題について考える．

3.2 確率的歪みによる固定長最大歪み符号化

前節では歪みの定義に定義 4 の最大歪みを用いた．しかし歪みには別のものも考えられる．その例として，この節では確率変数による最大歪みを考える．

定義 7 分布 $V_n(\cdot | \mathbf{x}, \mathbf{y})$ に従う確率的歪み $D_n(\mathbf{x}, \mathbf{y})$ を考える． $D_n(\mathbf{x}, \mathbf{y})$ に対して $(\mathbf{X}, \mathbf{Y})$ の確率変数による最大歪みを

$$\bar{D}(\mathbf{X}, \mathbf{Y}) \triangleq p\text{-}\limsup_{n \rightarrow \infty} \frac{1}{n} D_n(X^n, Y^n) \quad (80)$$

と定義する．

達成可能性を次のように定める．

定義 8 レート R が $\overline{D}$ の意味で D -達成可能であるとは

$$\overline{D}(\mathbf{X}, \mathbf{Y}) \leq D \quad (81)$$

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log M_n \leq R \quad (82)$$

を満たすような符号の列 $\{(\varphi_n, \psi_n)\}_{n=1}^{\infty}$ が存在するということである。ただしここで $Y^n = \psi_n(\varphi_n(X^n))$ とする。

定義 9 レート・歪み関数は以下のように定義される。

$$R_{\overline{D}}(D) = \inf\{R | R \text{ が } \overline{D} \text{ の意味で } D\text{-達成可能}\} \quad (83)$$

定理 2 一般情報源 $\mathbf{X}$ のレート・歪み関数は

$$R_{\overline{D}}(D) = \inf_{Y: \overline{D}(\mathbf{X}, \mathbf{Y}) \leq D} \overline{I}(\mathbf{X}; \mathbf{Y}) \quad (84)$$

で与えられる。

(証明) 順定理

$\overline{D}(\mathbf{X}, \mathbf{Y}) \leq D$ を満たす任意の $\mathbf{Y}$ を考える。任意の自然数 k に対して

$$\lim_{n \rightarrow \infty} \Pr \left\{ \frac{1}{n} D_n(X^n, Y^n) > D + \frac{1}{k} \right\} = 0 \quad (85)$$

である。ゆえに十分大きなすべての n で

$$\Pr \left\{ \frac{1}{n} D_n(X^n, Y^n) > D + \frac{1}{k} \right\} < \frac{1}{k} \quad (86)$$

が満たされる。このとき、 n に対して、式 (86) を満たす最小の k が決まる。その k に対して、 $\gamma_n \triangleq \frac{1}{k}$ とおく。すると $n \rightarrow \infty$ のとき $\gamma_n \rightarrow 0$ かつ

$$\lim_{n \rightarrow \infty} \Pr \left\{ \frac{1}{n} D_n(X^n, Y^n) > D + \gamma_n \right\} = 0 \quad (87)$$

が成り立つ（対角線論法）。

ここで γ を任意の正の定数として

$$R_n \triangleq \overline{I}(\mathbf{X}; \mathbf{Y}) + 2\gamma \quad (88)$$

$$\mathcal{S}_n(\mathbf{y}) \triangleq \left\{ \mathbf{x} \mid \frac{1}{n} D_n(\mathbf{x}, \mathbf{y}) \leq D + \gamma_n \right\} \quad (89)$$

として補題 1 を用いる。すると、ある (φ_n, ψ_n) 符号が存在し、

$$\frac{1}{n} \log M_n \leq \overline{I}(\mathbf{X}; \mathbf{Y}) + 2\gamma \quad (90)$$

$$\begin{aligned}
& \Pr \left\{ \frac{1}{n} D_n(X^n, \psi_n(\varphi_n(X^n))) > D + \gamma_n \right\} \\
& \leq \Pr \left\{ \frac{1}{n} \log \frac{\hat{W}_n(X^n|Y^n)}{P_{X^n}(X^n)} > \bar{I}(\mathbf{X}; \mathbf{Y}) + \gamma \right\} + \Pr \left\{ \frac{1}{n} D_n(X^n, Y^n) > D + \gamma_n \right\} + e^{-n\gamma}
\end{aligned} \tag{91}$$

が成り立つ。つまり式 (90) より

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log M_n \leq \bar{I}(\mathbf{X}; \mathbf{Y}) + 2\gamma \tag{92}$$

となり、定義 3 と式 (87) と式 (91) より

$$\lim_{n \rightarrow \infty} \Pr \left\{ \frac{1}{n} D_n(X^n, \psi_n(\varphi_n(X^n))) > D + \gamma_n \right\} = 0 \tag{93}$$

が導かれる。したがって確率的上極限の定義より

$$\text{p-lim sup}_{n \rightarrow \infty} \left(\frac{1}{n} D_n(X^n, \psi_n(\varphi_n(X^n))) - \gamma_n \right) \leq D \tag{94}$$

となるが $\gamma_n \rightarrow 0$ より

$$\text{p-lim sup}_{n \rightarrow \infty} \frac{1}{n} D_n(X^n, \psi_n(\varphi_n(X^n))) \leq D \tag{95}$$

を得る。よって式 (92) と式 (95) によって $\bar{I}(\mathbf{X}; \mathbf{Y}) + 2\gamma$ は $\bar{D}$ の意味で D -達成可能である。つまり

$$R_{\bar{D}}(D) \leq \bar{I}(\mathbf{X}; \mathbf{Y}) + 2\gamma \tag{96}$$

であるが、 $\gamma > 0$ は任意なので

$$R_{\bar{D}}(D) \leq \bar{I}(\mathbf{X}; \mathbf{Y}) \tag{97}$$

となる。さらに $\mathbf{Y}$ の任意性より

$$R_{\bar{D}}(D) \leq \inf_{\mathbf{Y}: \bar{D}(\mathbf{X}, \mathbf{Y}) \leq D} \bar{I}(\mathbf{X}; \mathbf{Y}) \tag{98}$$

が成り立つ。これで定理の順定理は導かれた。

逆定理

レート R が $\bar{D}$ の意味で D -達成可能であると仮定する。つまり

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log M_n \leq R \tag{99}$$

$$\bar{D}(\mathbf{X}, \mathbf{Y}) \leq D \tag{100}$$

$$Y^n = \{\psi_n(\varphi_n(X^n))\}_{n=1}^{\infty} \tag{101}$$

となる符号の列 $\{(\varphi_n, \psi_n)\}_{n=1}^{\infty}$ が存在する．定理 1 の逆定理と同様に補題 2 を用いれば式 (99) から

$$\bar{I}(\mathbf{X}; \mathbf{Y}) \leq R + 2\gamma \quad (102)$$

を得る． γ は任意だったので

$$\bar{I}(\mathbf{X}; \mathbf{Y}) \leq R \quad (103)$$

となり，したがってレート・歪み関数は

$$R_{\bar{D}}(D) \geq \bar{I}(\mathbf{X}; \mathbf{Y}) \quad (104)$$

を満たさなければいけない．ここで $\mathbf{Y}$ は $D(\mathbf{X}, \mathbf{Y}) \leq D$ を満たすことから

$$R_{\bar{D}}(D) \geq \inf_{\mathbf{Y}: D(\mathbf{X}, \mathbf{Y}) \leq D} \bar{I}(\mathbf{X}; \mathbf{Y}) \quad (105)$$

が成り立つ．これで逆定理も導かれた．

順定理・逆定理より

$$R_{\bar{D}}(D) = \inf_{\mathbf{Y}: D(\mathbf{X}, \mathbf{Y}) \leq D} \bar{I}(\mathbf{X}; \mathbf{Y}) \quad (106)$$

を満たす．□

4 歪み超過を許すレート・歪み理論

ここまである一定の歪み D を全く超えない場合を考えた．次に歪みが基準を超過する確率 δ がわずかにあっても達成可能とする場合を考える．つまり歪み超過を許すことによりさらにレートを下げる．そして超過確率 δ を許容する方法を 4 種類紹介し，それぞれの場合のレート・歪み関数を定義する．

4.1 $D_{\delta}(\mathbf{X}, \mathbf{Y})$: 歪み測度による許容

定義 10 ([4] 式 (20)) 歪みを

$$D_{\delta}(\mathbf{X}, \mathbf{Y}) \triangleq \inf \left\{ \alpha \mid \limsup_{n \rightarrow \infty} \Pr \left\{ \frac{1}{n} d_n(\mathbf{X}, \mathbf{Y}) > \alpha \right\} \leq \delta \right\} \quad (107)$$

と定義する．

定義 11 R が D_{δ} の意味で (δ, D) -達成可能とは

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log M_n \leq R \quad (108)$$

$$D_{\delta}(\mathbf{X}, \mathbf{Y}) \leq D \quad (109)$$

を満たす符号の列 $\{(\varphi_n, \psi_n)\}_{n=1}^{\infty}$ が存在することである。ただしここで $Y^n = \psi_n(\varphi_n(X^n))$ とする。

定義 12 D_δ における δ -レート・歪み関数は以下のように定義される。

$$R_{D_\delta}(\delta, D) \triangleq \inf\{R | R \text{ は } D_\delta \text{ の意味で } (\delta, D)\text{-達成可能}\} \quad (110)$$

この条件で歪みを許容したときの δ -レート・歪み関数を以下に示す。

定理 3 一般情報源 $\mathbf{X}$ の δ -レート・歪み関数は

$$R_{D_\delta}(\delta, D) = \inf_{Y: D_\delta(\mathbf{X}, \mathbf{Y}) \leq D} \bar{I}(\mathbf{X}; \mathbf{Y}) \quad (111)$$

で与えられる。

(証明)

初めに

$$R_0 \triangleq \inf_{Y: D_\delta(\mathbf{X}, \mathbf{Y}) \leq D} \bar{I}(\mathbf{X}; \mathbf{Y}) \quad (112)$$

とおき任意の $\gamma > 0$ に対して $R = R_0 + 2\gamma$ となるような R を考える。すると

$$R - \gamma = R_0 + \gamma > \bar{I}(\mathbf{X}; \mathbf{Y}) \quad (113)$$

$$D_\delta(\mathbf{X}, \mathbf{Y}) \leq D \quad (114)$$

を満たすような $\mathbf{Y}$ が存在する。つまり

$$\inf \left\{ \alpha \left| \limsup_{n \rightarrow \infty} \Pr \left\{ \frac{1}{n} d_n(\mathbf{X}, \mathbf{Y}) > \alpha \right\} \leq \delta \right\} \leq D \quad (115)$$

を満たされる。ここで上記より、任意の自然数 k に対して

$$\limsup_{n \rightarrow \infty} \Pr \left\{ \frac{1}{n} d_n(\mathbf{X}, \mathbf{Y}) > D + \frac{1}{k} \right\} \leq \delta \quad (116)$$

が成り立つことが分かる。すると十分大きなすべての n で

$$\Pr \left\{ \frac{1}{n} d_n(\mathbf{X}, \mathbf{Y}) > D + \frac{1}{k} \right\} \leq \delta + \frac{1}{k} \quad (117)$$

が成り立つような k が存在する。このとき、 n に対して、式 (117) を満たす最小の k が決まる。その k に対して、 $\gamma_n \triangleq \frac{1}{k}$ とおく。すると $n \rightarrow \infty$ のとき $\gamma_n \rightarrow 0$ となり

$$\Pr \left\{ \frac{1}{n} d_n(\mathbf{X}, \mathbf{Y}) > D + \gamma_n \right\} \leq \delta + \gamma_n \quad (118)$$

と書ける．（対角線論法）ここで

$$\mathcal{S}_n(\mathbf{y}) \triangleq \left\{ \mathbf{x} \mid \frac{1}{n} d_n(\mathbf{x}, \mathbf{y}) \leq D + \gamma_n \right\} \quad (119)$$

$$R_n = R \quad (120)$$

とおき，補題 1 を用いると

$$\frac{1}{n} \log M_n \leq R \quad (121)$$

$$\begin{aligned} & \Pr \left\{ \frac{1}{n} d_n(X^n, \psi_n(\varphi_n(X^n))) > D + \gamma_n \right\} \\ & \leq \Pr \left\{ \frac{1}{n} \log \frac{\hat{W}_n(X^n | Y^n)}{P_{X^n}(X^n)} \leq R - \gamma \right\} + \Pr \left\{ \frac{1}{n} d_n(X^n, Y^n) > D + \gamma_n \right\} + e^{-n\gamma} \end{aligned} \quad (122)$$

が満たされる．式 (122) の両辺の上極限をとると，式 (118) より

$$\limsup_{n \rightarrow \infty} \Pr \left\{ \frac{1}{n} d_n(X^n, \psi_n(\varphi_n(X^n))) > D \right\} \leq \delta \quad (123)$$

となる．したがって式 (123) により

$$\inf \left\{ \alpha \mid \limsup_{n \rightarrow \infty} \Pr \left\{ \frac{1}{n} d_n(X^n, \psi_n(\varphi_n(X^n))) > \alpha \right\} \leq \delta \right\} \leq D \quad (124)$$

と書ける．つまり定義より

$$D_\delta(X^n, \psi_n(\varphi_n(X^n))) \leq D \quad (125)$$

よって式 (121) と (125) より R は D_δ の意味で (δ, D) -達成可能である．つまり δ -レート・歪み関数は

$$R_{D_\delta}(\delta, D) \leq R_0 + 2\gamma \quad (126)$$

となる．ここで $\gamma > 0$ は任意だったので

$$R_{D_\delta}(\delta, D) \leq R_0 \quad (127)$$

である．したがって

$$R_{D_\delta}(\delta, D) \leq \inf_{Y: D_\delta(\mathbf{X}, \mathbf{Y}) \leq D} \bar{I}(\mathbf{X}; \mathbf{Y}) \quad (128)$$

が導かれる．

逆定理 レート R が D_δ の意味で (δ, D) -達成可能であると仮定すると

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log M_n \leq R \quad (129)$$

$$D_\delta(\mathbf{X}, \mathbf{Y}) \leq D \quad (130)$$

$$\mathbf{Y} = \{\psi_n(\varphi_n(X^n))\}_{n=1}^\infty \quad (131)$$

となる符号の列 $\{(\varphi_n, \psi_n)\}_{n=1}^\infty$ が存在する．定理 1 の逆定理と同様に補題 2 を用いれば式 (129) から

$$\bar{I}(\mathbf{X}; \mathbf{Y}) \leq R \quad (132)$$

が導かれる．したがって δ -レート・歪み関数は

$$R_{D_\delta}(\delta, D) \geq \bar{I}(\mathbf{X}; \mathbf{Y}) \quad (133)$$

を満たさなければいけない．ここで $\mathbf{Y}$ は $D_\delta(\mathbf{X}, \mathbf{Y}) \leq D$ を満たすことから

$$R_{D_\delta}(\delta, D) \geq \inf_{\mathbf{Y}: D_\delta(\mathbf{X}, \mathbf{Y}) \leq D} \bar{I}(\mathbf{X}; \mathbf{Y}) \quad (134)$$

が成り立つ．これで逆定理も導かれた．

順定理，逆定理より

$$R_{D_\delta}(\delta, D) = \inf_{\mathbf{Y}: D_\delta(\mathbf{X}, \mathbf{Y}) \leq D} \bar{I}(\mathbf{X}; \mathbf{Y}) \quad (135)$$

が分かる．□

4.2 $\bar{D}_\delta(\mathbf{X}, \mathbf{Y})$: 確率的歪みによる許容

定義 13 歪み $D_n(\mathbf{x}, \mathbf{y})$ は確率変数であり，分布 $V_n(\cdot | \mathbf{x}, \mathbf{y})$ に従う．さらに

$$\bar{D}_\delta(\mathbf{X}, \mathbf{Y}) \triangleq \inf_{n \rightarrow \infty} \{\alpha | p\text{-}\limsup V_n(\{d | d > \alpha\} | X^n Y^n) \leq \delta\} \quad (136)$$

とする．

次に達成可能性について定める．

定義 14 レート R が $\bar{D}_\delta$ の意味で (δ, D) -達成可能とは

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log M_n \leq R \quad (137)$$

$$\bar{D}_\delta(\mathbf{X}, \mathbf{Y}) \leq D \quad (138)$$

を満たす符号の列 $\{(\varphi_n, \psi_n)\}_{n=1}^\infty$ が存在することである．ただしここで $Y^n = \psi_n(\varphi_n(X^n))$ とする．

定義 15 $\overline{D}_\delta$ における δ -レート・歪み関数は以下のように定義される.

$$R_{\overline{D}_\delta}(\delta, D) \triangleq \inf\{R \mid R \text{ は } \overline{D}_\delta \text{ の意味で } (\delta, D)\text{-達成可能}\} \quad (139)$$

定理 4 一般情報源 $\mathbf{X}$ の δ -レート・歪み関数は

$$R_{\overline{D}_\delta}(\delta, D) = \inf_{\mathbf{Y}: \overline{D}_\delta(\mathbf{X}, \mathbf{Y}) \leq D} \overline{I}(\mathbf{X}; \mathbf{Y}) \quad (140)$$

で与えられる.

(証明) 順定理

$$R_0 \triangleq \inf_{\mathbf{Y}: \overline{D}_\delta(\mathbf{X}, \mathbf{Y}) \leq D} \overline{I}(\mathbf{X}, \mathbf{Y}) \quad (141)$$

$$R \triangleq R_0 + 2\gamma \quad (142)$$

とおく. すると

$$\overline{D}_\delta(\mathbf{X}, \mathbf{Y}) \leq D \quad (143)$$

$$R - \gamma = R_0 + \gamma > \overline{I}(\mathbf{X}; \mathbf{Y}) \quad (144)$$

を満たす $\mathbf{Y}$ が存在する. 式 (143) より任意の自然数 k に対して

$$\text{p-}\limsup_{n \rightarrow \infty} V_n \left(\left\{ d \mid d > D + \frac{1}{k} \right\} \middle| X^n Y^n \right) \leq \delta \quad (145)$$

となるが, これは確率的上極限の定義より

$$\inf \left\{ \alpha \mid \lim_{n \rightarrow \infty} \Pr \left\{ V_n \left(\left\{ d \mid d > D + \frac{1}{k} \right\} \middle| X^n Y^n \right) > \alpha \right\} = 0 \right\} \leq \delta \quad (146)$$

と書ける. さらに先ほどの k を用いると

$$\lim_{n \rightarrow \infty} \Pr \left\{ V_n \left(\left\{ d \mid d > D + \frac{1}{k} \right\} \middle| X^n Y^n \right) > \delta + \frac{1}{k} \right\} = 0 \quad (147)$$

となり, 十分大きなすべての n で

$$\Pr \left\{ V_n \left(\left\{ d \mid d > D + \frac{1}{k} \right\} \middle| X^n Y^n \right) > \delta + \frac{1}{k} \right\} < \frac{1}{k} \quad (148)$$

が満たされる. このとき, n に対して, 式 (148) を満たす最小の k が決まる. その k に対して, $\gamma_n \triangleq \frac{1}{k}$ とおく. すると $n \rightarrow \infty$ のとき $\gamma_n \rightarrow 0$ となり

$$\Pr \{ V_n(\{d \mid d > D + \gamma_n\} \mid X^n Y^n) > \delta + \gamma_n \} < \gamma_n \quad (149)$$

と書き換えられる.

ここで

$$R_n \triangleq R \quad (150)$$

$$\mathcal{S}_n(\mathbf{y}) \triangleq \{\mathbf{x} | V_n(\{d | d > D + \gamma_n\} | \mathbf{x}, \mathbf{y}) \leq \delta + \gamma_n\} \quad (151)$$

として補題 1 を用いると

$$\Pr \{X^n \notin \mathcal{S}_n(\psi_n(\varphi_n(X^n)))\} \leq \Pr \{X^n \notin \mathcal{B}_n(Y^n)\} + \Pr \{X^n \notin \mathcal{S}_n(Y^n)\} + e^{-n\gamma} \quad (152)$$

$$\frac{1}{n} \log M_n \leq R \quad (153)$$

となる符号 (φ_n, ψ_n) が存在する. 式 (149) と式 (152) より

$$\lim_{n \rightarrow \infty} \Pr \{V_n(\{d | d > D + \gamma_n\} | X^n \psi_n(\varphi_n(X^n))) > \delta + \gamma_n\} = 0 \quad (154)$$

つまり

$$\mathbf{p}\text{-}\limsup_{n \rightarrow \infty} V_n(\{d | d > D + \gamma_n\} | X^n \psi_n(\varphi_n(X^n))) \leq \delta \quad (155)$$

が成り立つ. 任意の $\gamma' > 0$ を考える. 十分大きなすべての n で $\gamma' > \gamma_n$ となることから

$$\{d | d > D + \gamma'\} \subset \{d | d > D + \gamma_n\} \quad (156)$$

という関係になる. 式 (155) と式 (156) より

$$\mathbf{p}\text{-}\limsup_{n \rightarrow \infty} V_n(\{d | d > D + \gamma'\} | X^n \psi_n(\varphi_n(X^n))) \leq \delta \quad (157)$$

となり,

$$\inf \{\alpha | \mathbf{p}\text{-}\limsup_{n \rightarrow \infty} V_n(\{d | d > \alpha\} | X^n \psi_n(\varphi_n(X^n))) \leq \delta\} \leq D + \gamma' \quad (158)$$

なので定義 13 より

$$\overline{D}_\delta(\mathbf{X}, \{\psi_n(\varphi_n(X^n))\}_{n=1}^\infty) \leq D + \gamma' \quad (159)$$

が導かれる. ここで γ' は任意だったので,

$$\overline{D}_\delta(\mathbf{X}, \{\psi_n(\varphi_n(X^n))\}_{n=1}^\infty) \leq D \quad (160)$$

となる. 式 (153) と式 (160) よりレート $R = R_0 + 2\gamma$ は (δ, D) -達成可能となる. したがって δ -レート・歪み関数は

$$R_{\overline{D}_\delta}(\delta, D) \leq R_0 + 2\gamma \quad (161)$$

を満たす．ところが γ は任意だったので

$$R_{\overline{D}_\delta}(\delta, D) \leq R_0 \quad (162)$$

となる．したがって

$$R_{\overline{D}_\delta}(\delta, D) \leq \inf_{\mathbf{Y}: \overline{D}_\delta(\mathbf{X}, \mathbf{Y}) \leq D} \bar{I}(\mathbf{X}; \mathbf{Y}) \quad (163)$$

となり，順定理が示された．

逆定理 レート R が $\overline{D}_\delta$ の意味で (δ, D) -達成可能であると仮定すると

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log M_n \leq R \quad (164)$$

$$\overline{D}_\delta(\mathbf{X}, \mathbf{Y}) \leq D \quad (165)$$

$$\mathbf{Y} = \{\psi_n(\varphi_n(X^n))\}_{n=1}^\infty \quad (166)$$

となる符号の列 $\{(\varphi_n, \psi_n)\}_{n=1}^\infty$ が存在する．定理 1 の逆定理と同様に補題 2 を用いれば式 (164) から

$$\bar{I}(\mathbf{X}; \mathbf{Y}) \leq R \quad (167)$$

が導かれる．したがって δ -レート・歪み関数は

$$R_{\overline{D}_\delta}(\delta, D) \geq \bar{I}(\mathbf{X}; \mathbf{Y}) \quad (168)$$

を満たさなければいけない．ここで $\mathbf{Y}$ は $\overline{D}_\delta(\mathbf{X}, \mathbf{Y}) \leq D$ を満たすことから

$$R_{\overline{D}_\delta}(\delta, D) \geq \inf_{\mathbf{Y}: \overline{D}_\delta(\mathbf{X}, \mathbf{Y}) \leq D} \bar{I}(\mathbf{X}; \mathbf{Y}) \quad (169)$$

が成り立つ．これで逆定理も導かれた．

順定理，逆定理より

$$R_{\overline{D}_\delta}(\delta, D) = \inf_{\mathbf{Y}: \overline{D}_\delta(\mathbf{X}, \mathbf{Y}) \leq D} \bar{I}(\mathbf{X}; \mathbf{Y}) \quad (170)$$

が満たされる．□

4.3 $\delta_D(\mathbf{X}, \mathbf{Y})$: 歪み測度の超過確率による許容

定義 16 ([4] 定義 2.3) 歪み超過確率を

$$\delta_D(\mathbf{X}, \mathbf{Y}) \triangleq \limsup_{n \rightarrow \infty} \Pr \left\{ \frac{1}{n} d_n(\mathbf{X}, \mathbf{Y}) > D \right\} \quad (171)$$

と定義する．

定義 17 R が δ_D の意味で (δ, D) -達成可能とは

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log M_n \leq R \quad (172)$$

$$\delta_D(\mathbf{X}, \mathbf{Y}) \leq \delta \quad (173)$$

を満たす符号の列 $\{(\varphi_n, \psi_n)\}_{n=1}^{\infty}$ が存在することである．ただしここで $Y^n = \psi_n(\varphi_n(X^n))$ とする．

定義 18 δ_D における δ -レート・歪み関数は以下のように定義される．

$$R_{\delta_D}(\delta, D) \triangleq \inf\{R | R \text{ は } \delta_D \text{ の意味で } (\delta, D) \text{-達成可能}\} \quad (174)$$

これらの条件の下で δ -レート・歪み関数を考えると以下のような定理が導かれる．

定理 5 一般情報源 $\mathbf{X}$ の δ -レート・歪み関数は

$$R_{\delta_D}(\delta, D) = \inf_{Y: \delta_D(\mathbf{X}, \mathbf{Y}) \leq \delta} \bar{I}(\mathbf{X}; \mathbf{Y}) \quad (175)$$

で与えられる．

順定理のみを証明する．逆定理は定理 3 と同様の方法でできる．

(証明)

初めに

$$R_0 \triangleq \inf_{Y: \delta_D(\mathbf{X}, \mathbf{Y}) \leq \delta} \bar{I}(\mathbf{X}; \mathbf{Y}) \quad (176)$$

とおき任意の $\gamma > 0$ に対して $R = R_0 + 2\gamma$ となるような R を考える．すると

$$R - \gamma = R_0 + \gamma > \bar{I}(\mathbf{X}; \mathbf{Y}) \quad (177)$$

$$\limsup_{n \rightarrow \infty} \Pr \left\{ \frac{1}{n} d_n(\mathbf{X}, \mathbf{Y}) > D \right\} \leq \delta \quad (178)$$

をみたすような $\mathbf{Y}$ が存在する．さらに

$$\mathcal{S}_n(\mathbf{y}) = \left\{ \mathbf{x} \mid \frac{1}{n} d_n(\mathbf{x}, \mathbf{y}) \leq D \right\} \quad (179)$$

$$R_n = R \quad (180)$$

とおき，補題 1 を用いると

$$\frac{1}{n} \log M_n \leq R \quad (181)$$

$$\begin{aligned} & \Pr \left\{ \frac{1}{n} d_n(X^n, \psi_n(\varphi_n(X^n))) > D \right\} \\ & \leq \Pr \left\{ \frac{1}{n} \log \frac{\hat{W}_n(X^n|Y^n)}{P_{X^n}(X^n)} \leq R - \gamma \right\} + \Pr \left\{ \frac{1}{n} d_n(X^n, Y^n) > D \right\} + e^{-n\gamma} \end{aligned} \quad (182)$$

が導かれる。したがって式 (178) と (182) より

$$\limsup_{n \rightarrow \infty} \Pr \left\{ \frac{1}{n} d_n(X^n, \psi_n(\varphi_n(X^n))) > D \right\} \leq \delta \quad (183)$$

となる。式 (181) と (183) より $R = R_0 + 2\gamma$ は (δ, D) -達成可能。よって δ -レート・歪み関数は

$$R_{\delta_D}(\delta, D) \leq R_0 + 2\gamma \quad (184)$$

となる。ここで $\gamma > 0$ は任意だったので

$$R_{\delta_D}(\delta, D) \leq R_0 \quad (185)$$

である。したがって

$$R_{\delta_D}(\delta, D) \leq \inf_{Y: \delta_D(\mathbf{X}, \mathbf{Y}) \leq \delta} \bar{I}(\mathbf{X}; \mathbf{Y}) \quad (186)$$

となる。□

4.4 $\bar{\delta}_D(\mathbf{X}, \mathbf{Y})$: 確率的歪みの超過確率による許容

定義 19 歪み超過確率を

$$\bar{\delta}_D(\mathbf{X}, \mathbf{Y}) \triangleq p\text{-}\limsup_{n \rightarrow \infty} V_n(\{d|d > D\} | X^n Y^n) \quad (187)$$

と定義する。

定義 20 R が $\bar{\delta}_D$ の意味で (δ, D) -達成可能とは

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log M_n \leq R \quad (188)$$

$$\bar{\delta}_D(\mathbf{X}, \mathbf{Y}) \leq \delta \quad (189)$$

を満たす符号の列 $\{(\varphi_n, \psi_n)\}_{n=1}^{\infty}$ が存在することである。ただしここで $Y^n = \psi_n(\varphi_n(X^n))$ とする。

定義 21 $\bar{\delta}_D$ における δ -レート・歪み関数は以下のように定義される。

$$R_{\bar{\delta}_D}(\delta, D) \triangleq \inf \{ R | R \text{ は } \bar{\delta}_D \text{ の意味で } (\delta, D)\text{-達成可能} \} \quad (190)$$

これらの条件の下で δ -レート・歪み関数を考えると以下のような定理が導かれる。

定理 6 $\bar{\delta}_D$ における δ -レート・歪み関数は

$$R_{\bar{\delta}_D}(\delta, D) = \inf_{Y: \bar{\delta}_D(\mathbf{X}, \mathbf{Y}) \leq \delta} \bar{I}(\mathbf{X}; \mathbf{Y}) \quad (191)$$

で与えられる。

順定理のみを証明する。逆定理は定理 4 と同様の方法でできる。
(証明) まず

$$R_0 \triangleq \inf_{Y: \bar{\delta}_D(\mathbf{X}, \mathbf{Y}) \leq \delta} \bar{I}(\mathbf{X}; \mathbf{Y}) \quad (192)$$

とおき任意の $\gamma > 0$ に対して $R = R_0 + 2\gamma$ となるような R を考える。すると

$$R - \gamma = R_0 + \gamma > \bar{I}(\mathbf{X}; \mathbf{Y}) \quad (193)$$

$$\bar{\delta}_D(\mathbf{X}, \mathbf{Y}) \leq \delta \quad (194)$$

を満たすような $\mathbf{Y}$ が存在する。つまり

$$\text{p-lim sup}_{n \rightarrow \infty} V_n(\{d|d > D\} | X^n Y^n) \leq \delta \quad (195)$$

を満たす。確率的上極限の定義より

$$\inf \left\{ \alpha \mid \lim_{n \rightarrow \infty} \Pr \{ V_n(\{d|d > D\} | X^n Y^n) > \alpha \} = 0 \right\} \leq \delta \quad (196)$$

と変形できる。さらに任意の自然数 k に対して

$$\lim_{n \rightarrow \infty} \Pr \left\{ V_n(\{d|d > D\} | X^n Y^n) > \delta + \frac{1}{k} \right\} = 0 \quad (197)$$

と書ける。そして十分大きなすべての n に対して

$$\Pr \left\{ V_n(\{d|d > D\} | X^n Y^n) > \delta + \frac{1}{k} \right\} < \frac{1}{k} \quad (198)$$

が成り立つような k が存在する。このとき、 n に対して、式 (198) を満たす最小の k が決まる。その k に対して、 $\gamma_n \triangleq \frac{1}{k}$ とおく。すると $n \rightarrow \infty$ のとき $\gamma_n \rightarrow 0$ となり

$$\Pr \{ V_n(\{d|d > D\} | X^n Y^n) > \delta + \gamma_n \} < \gamma_n \quad (199)$$

となる。(対角線論法) ここで

$$\mathcal{S}_n(\mathbf{y}) \triangleq \{ \mathbf{x} \mid V_n(\{d|d > D\} | \mathbf{x}\mathbf{y}) \leq \delta + \gamma_n \} \quad (200)$$

$$R_n = R \quad (201)$$

とおき，補題 1 を用いると

$$\frac{1}{n} \log M_n \leq R \quad (202)$$

$$\begin{aligned} & \Pr \{V_n(\{d|d > D\}|X^n \psi_n(\varphi_n(X^n))) > \delta + \gamma_n\} \\ & \leq \Pr \left\{ \frac{1}{n} \log \frac{\hat{W}_n(X^n|Y^n)}{P_{X^n}(X^n)} \leq R - \gamma \right\} + \Pr \{V_n(\{d|d > D\}|X^n Y^n) > \delta + \gamma_n\} + e^{-n\gamma} \end{aligned} \quad (203)$$

が満たされる．式 (203) の両辺の極限をとると式 (199) より

$$\lim_{n \rightarrow \infty} \Pr \{V_n(\{d|d > D\}|X^n \psi_n(\varphi_n(X^n))) > \delta\} = 0 \quad (204)$$

となる．したがって

$$\inf \left\{ \alpha \mid \lim_{n \rightarrow \infty} \Pr \{V_n(\{d|d > D\}|X^n \psi_n(\varphi_n(X^n))) > \alpha\} = 0 \right\} \leq \delta \quad (205)$$

である．確率的上極限の定義より

$$\text{p-}\limsup_{n \rightarrow \infty} V_n(\{d|d > D\}|X^n \psi_n(\varphi_n(X^n))) \leq \delta \quad (206)$$

なので

$$\bar{\delta}_D(\mathbf{X}, \psi_n(\varphi_n(X^n))) \leq \delta \quad (207)$$

が導かれ， R が (δ, D) -達成可能といえる．つまり δ -レート・歪み関数は

$$R_{\bar{\delta}_D}(\delta, D) \leq R_0 + 2\gamma \quad (208)$$

となる．ここで $\gamma > 0$ は任意だったので

$$R_{\bar{\delta}_D}(\delta, D) \leq R_0 \quad (209)$$

である．したがって

$$R_{\bar{\delta}_D}(\delta, D) \leq \inf_{Y: \bar{\delta}_D(\mathbf{X}, \mathbf{Y}) \leq \delta} \bar{I}(\mathbf{X}; \mathbf{Y}) \quad (210)$$

が導かれる．□

4.5 歪み超過ゼロとしたときの δ -レート・歪み関数

ここで $\delta = 0$ としたとき 3 章で示したレート・歪み関数とどのような関係があるのかを確認する．また次節でそれぞれの許容方法の関係も考える．まず $\delta = 0$ のときの δ -レート・歪み関数を見ていく．すると以下の定理が成り立つ．

定理 7 レート・歪み関数について

$$R_{D_\delta}(0, D) = R_D(D) \quad (211)$$

が成り立つ．

(証明) $\delta = 0$ を代入すると，定義 1 より

$$D_0(\mathbf{X}, \mathbf{Y}) = \inf \left\{ \alpha \mid \limsup_{n \rightarrow \infty} \Pr \left\{ \frac{1}{n} d_n(X^n, Y^n) > \alpha \right\} \leq 0 \right\} \quad (212)$$

$$= \text{p-lim sup}_{n \rightarrow \infty} \frac{1}{n} d_n(X^n, Y^n) \quad (213)$$

$$= D(\mathbf{X}, \mathbf{Y}) \quad (214)$$

となる．したがって定理が成り立つ．□

定理 8 レート・歪み関数について

$$R_{\overline{D}_\delta}(0, D) = R_{\overline{D}}(D) \quad (215)$$

が成り立つ．

(証明) $\delta = 0$ のときの $(\mathbf{X}, \mathbf{Y})$ の δ -歪みスペクトル上限を考える．すなわち

$$\inf \left\{ \alpha \mid \text{p-lim sup}_{n \rightarrow \infty} V_n(\{d|d > \alpha\} | X^n Y^n) = 0 \right\} \quad (216)$$

の値を考える．条件式の部分は確率的上極限の定義より

$$\inf \left\{ \alpha \mid \lim_{n \rightarrow \infty} \Pr \{ V_n(\{d|d > \alpha\} | X^n Y^n) > \alpha \} = 0 \right\} = 0 \quad (217)$$

である．ここで任意の $\varepsilon > 0$ に対して

$$\lim_{n \rightarrow \infty} \Pr \{ V_n(\{d|d > \alpha\} | X^n Y^n) > \varepsilon \} = 0 \quad (218)$$

と書ける．一方， $V_n(\{d|d > \alpha\} | X^n Y^n) \geq 0$ なので

$$\lim_{n \rightarrow \infty} \Pr \{ V_n(\{d|d > \alpha\} | X^n Y^n) < -\varepsilon \} = 0 \quad (219)$$

となることは明らかであり，式 (218) 式 (219) より

$$\lim_{n \rightarrow \infty} \Pr \{ |V_n(\{d|d > \alpha\}|X^n Y^n)| > \varepsilon \} = 0 \quad (220)$$

である．つまり $V_n(\{d|d > \alpha\}|X^n Y^n)$ は確率収束し

$$V_n(\{d|d > \alpha\}|X^n Y^n) \rightarrow 0 \quad \text{in prob} \quad (221)$$

となる．つまり

$$\lim_{n \rightarrow \infty} \mathbf{E}[V_n(\{d|d > \alpha\}|X^n Y^n)] = 0 \quad (222)$$

となる．さらに期待値の部分は

$$\mathbf{E}[V_n(\{d|d > \alpha\}|X^n, Y^n)] \quad (223)$$

$$= \sum_{\mathbf{x}, \mathbf{y}} P_{X^n Y^n}(\mathbf{x}, \mathbf{y}) V_n(\{d|d > \alpha\}|\mathbf{x}, \mathbf{y}) \quad (224)$$

$$= \sum_{\mathbf{x}, \mathbf{y}} P_{X^n Y^n D_n}(\mathbf{x}, \mathbf{y}, \{d|d > \alpha\}) \quad (225)$$

$$= \sum_{\mathbf{x}, \mathbf{y}} \Pr\{X^n = \mathbf{x}, Y^n = \mathbf{y}, D_n(X^n, Y^n) > \alpha\} \quad (226)$$

$$= \Pr\{D_n(X^n, Y^n) > \alpha\} \quad (227)$$

となる．式 (216) に戻すと

$$\inf\{\alpha | \lim_{n \rightarrow \infty} \Pr\{D_n(X^n, Y^n) > \alpha\} = 0\} \quad (228)$$

$$= \text{p-lim sup}_{n \rightarrow \infty} D_n(X^n, Y^n) \quad (229)$$

$$= \overline{D}(X^n, Y^n) \quad (230)$$

が導かれる．よって

$$\overline{D}_0(\mathbf{X}, \mathbf{Y}) = \overline{D}(\mathbf{X}, \mathbf{Y}) \quad (231)$$

となるので，定義より定理が得られる．□

これらのことから定義 11 は定義 4 の，定義 14 は定義 8 の歪み超過を許容した一般化となっていることが分かる．

さらに超過確率による許容も同様に超過確率ゼロとしたときのことを考える．初めに確率速度による許容 δ_D をゼロとすると

$$\limsup_{n \rightarrow \infty} \Pr \left\{ \frac{1}{n} d_n(X^n, Y^n) > D \right\} = 0 \quad (232)$$

となる．

この歪みの定義を用いることにより，新たな定義でレート・歪み関数が導ける．

定義 22 R が δ_D の意味で D -達成可能とは

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log M_n \leq R \quad (233)$$

$$\delta_D(\mathbf{X}, \mathbf{Y}) = 0 \quad (234)$$

を満たす符号の列 $\{(\varphi_n, \psi_n)\}_{n=1}^{\infty}$ が存在することである．ただしここで $Y^n = \psi_n(\varphi_n(X^n))$ とする．

定義 23 δ_D におけるレート・歪み関数は以下のように定義される．

$$R_{\delta_D}(D) \triangleq \inf\{R | R \text{ は } \delta_D \text{ の意味で } D\text{-達成可能}\} \quad (235)$$

これら条件の下でレート・歪み関数を考えると以下のような系が導かれる．

系 1 δ_D における-レート・歪み関数は

$$R_{\delta_D}(D) = \inf_{Y: \delta_D(\mathbf{X}, \mathbf{Y})=0} \bar{I}(\mathbf{X}; \mathbf{Y}) \quad (236)$$

で与えられる．

(証明)

定理 5 を $\delta = 0$ とすることにより明らか．□

さらに確率的歪みでも同様に， $\delta_D = 0$ のとき

$$\text{p-}\limsup_{n \rightarrow \infty} V_n(\{d | d > D\} | X^n Y^n) = 0 \quad (237)$$

となるので定理 8 の証明と同様に式変形をすることにより

$$\limsup_{n \rightarrow \infty} \Pr \left\{ \frac{1}{n} D_n(X^n, Y^n) > D \right\} = 0 \quad (238)$$

となる．

こちらも達成可能を定義し，レート・歪み問題として考える．

定義 24 R が $\bar{\delta}_D$ の意味で D -達成可能とは

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log M_n \leq R \quad (239)$$

$$\bar{\delta}_D(\mathbf{X}, \mathbf{Y}) = 0 \quad (240)$$

を満たす符号の列 $\{(\varphi_n, \psi_n)\}_{n=1}^{\infty}$ が存在することである．ただしここで $Y^n = \psi_n(\varphi_n(X^n))$ とする．

定義 25 $\bar{\delta}_D$ におけるレート・歪み関数は以下のように定義される。

$$R_{\bar{\delta}_D}(D) \triangleq \inf \{ R | R \text{ は } \bar{\delta}_D \text{ の意味で } D\text{-達成可能} \} \quad (241)$$

これら条件の下でレート・歪み関数を考えると以下のような系が導かれる。

系 2 $\bar{\delta}_D$ における-レート・歪み関数は

$$R_{\bar{\delta}_D}(D) = \inf_{Y: \bar{\delta}_D(\mathbf{X}, \mathbf{Y})=0} \bar{I}(\mathbf{X}; \mathbf{Y}) \quad (242)$$

で与えられる。

(証明)

定理 6 を $\delta = 0$ とすることにより明らか。□

以上のように歪み許容をしたレート・歪み理論は歪みなしの場合のものの拡張となっていることが示せた。次に各歪み許容の関係をまとめる。

4.6 歪み超過方法まとめ

前節までで述べた歪み超過の種類をまとめる。

$$1. D_\delta(\mathbf{X}, \mathbf{Y}) \triangleq \inf \left\{ \alpha \left| \limsup_{n \rightarrow \infty} \Pr \left\{ \frac{1}{n} d_n(X^n, Y^n) > \alpha \right\} \leq \delta \right\} \leq D \quad (\text{定義 11})$$

$$2. \bar{D}_\delta(\mathbf{X}, \mathbf{Y}) \triangleq \inf \left\{ \alpha \left| \text{p-lim sup}_{n \rightarrow \infty} V_n(\{d|d > \alpha\} | X^n Y^n) \leq \delta \right\} \leq D \quad (\text{定義 14})$$

$$3. \delta_D(\mathbf{X}, \mathbf{Y}) \triangleq \limsup_{n \rightarrow \infty} \Pr \left\{ \frac{1}{n} d_n(X^n, Y^n) > D \right\} \leq \delta \quad (\text{定義 17})$$

$$4. \bar{\delta}_D(\mathbf{X}, \mathbf{Y}) \triangleq \text{p-lim sup}_{n \rightarrow \infty} V_n(\{d|d > D\} | X^n Y^n) \leq \delta \quad (\text{定義 20})$$

さらにそれぞれの式において歪み超過をゼロとすることによりレート・歪み問題となる。

$$5. D(\mathbf{X}, \mathbf{Y}) \triangleq \text{p-lim sup}_{n \rightarrow \infty} \frac{1}{n} d_n(X^n, Y^n) \leq D \quad (\text{定義 5})$$

$$6. \bar{D}(\mathbf{X}, \mathbf{Y}) \triangleq \text{p-lim sup}_{n \rightarrow \infty} \frac{1}{n} D_n(X^n, Y^n) \leq D \quad (\text{定義 8})$$

$$3. \delta_D(\mathbf{X}, \mathbf{Y}) \triangleq \limsup_{n \rightarrow \infty} \Pr \left\{ \frac{1}{n} d_n(X^n, Y^n) > D \right\} \leq 0 \quad (\text{定義 22})$$

$$4. \bar{\delta}_D(\mathbf{X}, \mathbf{Y}) \triangleq \limsup_{n \rightarrow \infty} \Pr \left\{ \frac{1}{n} D_n(X^n, Y^n) > D \right\} \leq 0 \quad (\text{定義 24})$$

以上の歪み超過方法すべてで補題 1 を用いることによりレート・歪み関数、 δ -レート・歪み関数の導出ができた。

ここまで 4 種類の歪みと、それらの歪み超過の許容方法を示した。これらの歪みの関係性を整理する。本論文で扱った歪みには歪み測度の任意の関数を用いたものと、確率変数によるものの 2 種類に分けられる。さらにそれらの中でも歪み超過確率を直接表したものと、超過許容の歪みとして表したものの 2 種類があり、計 4 種類の歪みを紹介した。それぞれの定義より次のような包含関係があることが分かる。

$$\begin{aligned} \text{(歪み測度)} \quad \delta_D \leq \delta &\Rightarrow D_\delta \leq D \\ \text{(確率変数歪み)} \quad \bar{\delta}_D \leq \delta &\Rightarrow \bar{D}_\delta \leq D \end{aligned}$$

同様にして歪みを認めない場合も $\delta = 0$ とすることにより、下記のような関係があることが分かる。

$$\begin{aligned} \text{(歪み測度)} \quad \delta_D \leq 0 &\Rightarrow D_0 \leq D \\ \text{(確率変数歪み)} \quad \bar{\delta}_D \leq 0 &\Rightarrow \bar{D}_0 \leq D \end{aligned}$$

つまり 4 章での議論は 3 章の一般化であることも分かる。さらに歪みの関係より、 δ -レート・歪み関数、レート・歪み関数の関係も明らかにできる。

δ -レート・歪み関数

$$\begin{aligned} \text{(歪み測度)} \quad R_{\delta_D}(\delta, D) &\geq R_{D_\delta}(\delta, D) \\ \text{(確率変数歪み)} \quad R_{\bar{\delta}_D}(\delta, D) &\geq R_{\bar{D}_\delta}(\delta, D) \end{aligned}$$

レート・歪み関数

$$\begin{aligned} \text{(歪み測度)} \quad R_{\delta_D}(D) &\geq R_{D_\delta}(D) \\ \text{(確率変数歪み)} \quad R_{\bar{\delta}_D}(D) &\geq R_{\bar{D}_\delta}(D) \end{aligned}$$

以上のように既存の問題設定とは違う様々なレート・歪み問題について考えた。

5 まとめ

本論文ではランダム符号化を用いずに一般情報源のレート・歪み関数を導いた。導出には Feinstein の補題の符号化方法を参考にした補題を発見し適用した。その補題は、復号器からの出力が任意の集合に入らない確率を上から抑えるというものであり、さらに一般情報源に適用可能なので様々な応用が期待される。事例とし本論文では式 (107) や式 (136)などを定義し、その補題を用いることにより、歪み超過を許すレート・歪み関数も明らかにした。この場合の式も歪み超過を許さない場合の式と同様な形を導くことができた。さらにそのほかのレート・歪み問題について考察を広げあらゆる形での定式化をした。

謝辞

本研究を進めるにあたり，多くの助言，細かな指導をしてくださった指導教員の西新幹彦准教授，学会発表の際にコメントをいただいた東京工業大学の植松友彦教授，電気通信大学の韓太舜名誉教授，早稲田大学の齋藤翔太先生，また手助けや協力をしてくださった研究室の皆様に感謝の意を表する。

参考文献

- [1] 韓 太舜, 情報理論における情報スペクトル的方法, 培風館, 1998.
- [2] I.Csiszar, J.Korner, Information Theory Coding Theorems for Discrete Memoryless Systems 2nd ed., Cambridge, 2011.
- [3] S.Saito, H.Yagi, T.Matsushima, "Variable-Length Lossy Compression Allowing Positive Overflow and Excess Distortion Probabilities" The 39th Symposium on Information Theory and its Application (SITA2016), pp164–169, Dec. 2016
- [4] 野村亮, 八木秀樹, 「微小な歪み超過を許す固定長有歪み符号化定理」, 信学技報, IEICE Technical Report, pp71–76, July. 2014
- [5] 八木秀樹, 野村亮, 「符号長または歪みの超過確率を許容する忠実度規範付き楽観的符号化と強逆性」, 第 37 回情報理論とその応用シンポジウム (SITA2014), pp207–212, Dec. 2014
- [6] 原寛貴, 「一般情報源に対する最大歪み測度を用いた忠実度規範付き情報源符号化定理」, 信州大学大学院工学系研究科修士論文 (指導教員: 西新幹彦), 2013 年 2 月

付録 A ランダム符号化を用いるレート・歪み理論

本稿ではレート歪み関数をランダムでない方法で示した．ここでは既存のランダムな方法での証明方法を紹介する．

A.1 一般のレート・歪み理論

以下の証明が本稿の議論をランダム符号を用いてする方法である．なお順定理のみとする．(定理 1 の証明)([1] 定理 5.3)

初めに，条件である $D(\mathbf{X}, \mathbf{Y}) \leq D$ を満たすような $\mathbf{Y}$ を考える．さらに任意の γ に対して

$$\mathcal{T}_n^{(1)} = \left\{ (\mathbf{x}, \mathbf{y}) \in \mathcal{X}^n \times \mathcal{Y}^n \mid \frac{1}{n} \log \frac{P_{Y^n|X^n}(\mathbf{x}|\mathbf{y})}{P_{Y^n}(\mathbf{y})} < \bar{I}(\mathbf{X}; \mathbf{Y}) + \gamma \right\} \quad (243)$$

$$\mathcal{T}_n^{(2)} = \left\{ (\mathbf{x}, \mathbf{y}) \in \mathcal{X}^n \times \mathcal{Y}^n \mid \frac{1}{n} d_n(\mathbf{x}, \mathbf{y}) < D(\mathbf{X}, \mathbf{Y}) + \gamma \right\} \quad (244)$$

とし， $\mathcal{T}_n = \mathcal{T}_n^{(1)} \cap \mathcal{T}_n^{(2)}$ とおく．ここで，

そこで

$$M_n \triangleq e^{n(\bar{I}(\mathbf{X}; \mathbf{Y}) + 2\gamma)} \quad (245)$$

として， $v_1, v_2, \dots, v_{M_n} \in \mathcal{Y}^n$ を互いに独立に分布 P_{Y^n} にしたがって発生させ $C = \{v_1, v_2, \dots, v_{M_n}\}$ とおく．さらに各 $\mathbf{x}$ に対して

$$d_n(\mathbf{x}, v_{i_0}) = \min_{1 \leq i \leq M_n} d_n(\mathbf{x}, v_i) \quad (246)$$

となるような i_0 を定め，符号器 φ_n を $\varphi_n(\mathbf{x}) = i_0$ と定義する．一方，復号器 ψ_n は符号器から番号 i を受け取ったとき $\psi_n(i) = v_i$ と定める．このように定義されたランダムな符号器と復号器の組 (φ, ψ) に対して

$$\bar{P}_e^{(n)} \triangleq \Pr \left\{ \frac{1}{n} d_n(X^n, \psi(\varphi(X^n))) > D(\mathbf{X}, \mathbf{Y}) + \gamma \right\} \quad (247)$$

と定義する．さらに $\bar{P}_e^{(n)}$ について評価する． $v_1, v_2, \dots, v_{M_n}$ は互いに独立なので，

$$\bar{P}_e^{(n)} = \sum_{\mathbf{x} \in \mathcal{X}^n} P_{X^n}(\mathbf{x}) \prod_{i=1}^{M_n} \Pr \left\{ \frac{1}{n} d_n(\mathbf{x}, v_i) > D(\mathbf{X}, \mathbf{Y}) + \gamma \right\} \quad (248)$$

となる．さらに $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{M_n}$ は同一の分布 P_{Y^n} に従うことより

$$\bar{P}_e^{(n)} = \sum_{\mathbf{x} \in \mathcal{X}^n} P_{X^n}(\mathbf{x}) \left(\Pr \left\{ \frac{1}{n} d_n(\mathbf{x}, Y^n) > \bar{D}(\mathbf{X}, \mathbf{Y}) + \gamma \right\} \right)^{M_n} \quad (249)$$

$$= \sum_{\mathbf{x} \in \mathcal{X}^n} P_{X^n}(\mathbf{x}) \left(1 - \Pr \left\{ \frac{1}{n} d_n(\mathbf{x}, Y^n) \leq \bar{D}(\mathbf{X}, \mathbf{Y}) + \gamma \right\} \right)^{M_n} \quad (250)$$

と変形できる．そして

$$\bar{P}_e^{(n)} = \sum_{\mathbf{x} \in \mathcal{X}^n} P_{X^n}(\mathbf{x}) \left(1 - \sum_{\mathbf{y} \in \mathcal{Y}^n} P_{Y^n}(\mathbf{y}) \mathbf{1} \left[\frac{1}{n} d_n(\mathbf{x}, \mathbf{y}) \leq \bar{D}(\mathbf{X}, \mathbf{Y}) + \gamma \right] \right)^{M_n} \quad (251)$$

と書き換えられる．ここで $\mathbf{1}[Z]$ は指示関数 (indicator function) と呼ばれ、条件式 Z において Z が成立するなら 1 を、そうでないなら 0 を出力する関数である．次に

$$J(\mathbf{x}, \mathbf{y}) = \begin{cases} 1 & ((\mathbf{x}, \mathbf{y}) \in \mathcal{T}_n) \\ 0 & (\text{others}) \end{cases} \quad (252)$$

を定義する．すると

$$J(\mathbf{x}, \mathbf{y}) \leq \mathbf{1} \left[\frac{1}{n} d_n(\mathbf{x}, \mathbf{y}) \leq \bar{D}(\mathbf{X}, \mathbf{Y}) \right] \quad (253)$$

が成り立つ．したがって式 (251) は

$$\bar{P}_e^{(n)} \leq \sum_{\mathbf{x} \in \mathcal{X}^n} P_{X^n}(\mathbf{x}) \left(1 - \sum_{\mathbf{y} \in \mathcal{Y}^n} P_{Y^n}(\mathbf{y}) J(\mathbf{x}, \mathbf{y}) \right)^{M_n} \quad (254)$$

となる．ここで、式 (243) より $(\mathbf{x}, \mathbf{y}) \in \mathcal{T}_n$ ならば

$$P_{Y^n}(\mathbf{y}) \geq e^{-n(\bar{I}(X; Y) + \gamma)} P_{Y^n|X^n}(\mathbf{y}|\mathbf{x}) \quad (255)$$

が満たされるので、任意の $(\mathbf{x}, \mathbf{y}) \in \mathcal{X}^n \times \mathcal{Y}^n$ に対して

$$P_{Y^n}(\mathbf{y}) J(\mathbf{x}, \mathbf{y}) \geq e^{-n(\bar{I}(X; Y) + \gamma)} P_{Y^n|X^n}(\mathbf{y}|\mathbf{x}) J(\mathbf{x}, \mathbf{y}) \quad (256)$$

が成り立つ．つまり式 (254) に式 (256) を代入することにより

$$\bar{P}_e^{(n)} \leq \sum_{\mathbf{x} \in \mathcal{X}^n} P_{X^n}(\mathbf{x}) \left(1 - \sum_{\mathbf{y} \in \mathcal{Y}^n} e^{-n(\bar{I}(X; Y) + \gamma)} P_{Y^n|X^n}(\mathbf{y}|\mathbf{x}) J(\mathbf{x}, \mathbf{y}) \right)^{M_n} \quad (257)$$

が得られる．ここで不等式

$$(1 - x)^y \leq e^{-xy} \quad (0 \leq x \leq 1, y \geq 0) \quad (258)$$

を用いて $M_n = e^{n(\bar{I}(X;Y)+2\gamma)}$ を代入すると

$$\bar{P}_e^{(n)} \leq \sum_{\mathbf{x} \in \mathcal{X}^n} P_{X^n}(\mathbf{x}) e^{-(e^{n\gamma} \sum_{\mathbf{y} \in \mathcal{Y}^n} P_{Y^n|X^n}(\mathbf{y}|\mathbf{x}) J(\mathbf{x}, \mathbf{y}))} \quad (259)$$

と評価できる．さらに不等式

$$e^{-xy} \leq 1 - y + e^{-x} \quad (x \geq 0, 0 \leq y \leq 1) \quad (260)$$

を用いると

$$\bar{P}_e^{(n)} \leq 1 - \sum_{\mathbf{x} \in \mathcal{X}^n} P_{X^n}(\mathbf{x}) \sum_{\mathbf{y} \in \mathcal{Y}^n} P_{Y^n|X^n}(\mathbf{y}|\mathbf{x}) J(\mathbf{x}, \mathbf{y}) + e^{-e^{n\gamma}} \quad (261)$$

$$= 1 - \Pr\{X^n Y^n \in \mathcal{T}_n\} + e^{-e^{n\gamma}} \quad (262)$$

$$= \Pr\{X^n Y^n \notin \mathcal{T}_n\} + e^{-e^{n\gamma}} \quad (263)$$

となる．なお上で用いた不等式は付録 C で証明する．一方， $\bar{I}(\mathbf{X}; \mathbf{Y}), \bar{D}(\mathbf{X}, \mathbf{Y})$ の定義より確率的上極限を用いているので

$$\Pr\{X^n Y^n \notin \mathcal{T}_n^{(1)}\} \rightarrow 0, \quad \Pr\{X^n Y^n \notin \mathcal{T}_n^{(2)}\} \rightarrow 0 \quad (n \rightarrow \infty) \quad (264)$$

となることから

$$\Pr\{X^n Y^n \notin \mathcal{T}_n\} \rightarrow 0 \quad (n \rightarrow \infty) \quad (265)$$

である．よって式 (263) より

$$\bar{P}_e^{(n)} \rightarrow 0 \quad (n \rightarrow \infty) \quad (266)$$

を得る．つまり

$$\Pr\left\{\frac{1}{n} d_n(X^n, \psi(\varphi(X^n))) > D(\mathbf{X}, \mathbf{Y}) + \gamma\right\} \rightarrow 0 \quad (n \rightarrow \infty) \quad (267)$$

となる．ランダム符号は平均的な結果を述べているので，少なくとも一つはランダムではないある符号 (φ_n, ψ_n) が存在して式 (270) を満たすことがわかる．ここで γ は任意なので $k \in \mathbf{N}$ に対し $\gamma \triangleq \frac{1}{k}$ とすると

$$\lim_{n \rightarrow \infty} \Pr\left\{\frac{1}{n} d_n(X^n, \psi(\varphi(X^n))) > D(\mathbf{X}, \mathbf{Y}) + \frac{1}{k}\right\} = 0 \quad (268)$$

となり，さらに十分大きなすべての n で

$$\Pr\left\{\frac{1}{n} d_n(X^n, \psi(\varphi(X^n))) > D(\mathbf{X}, \mathbf{Y}) + \frac{1}{k}\right\} < \frac{1}{k} \quad (269)$$

が満たされる．よって $n \rightarrow \infty$ のとき $\gamma_n \rightarrow 0$ となるような γ_n を用い，対角線論法より

$$\Pr \left\{ \frac{1}{n} d_n(X^n, \psi(\varphi(X^n))) > D(\mathbf{X}, \mathbf{Y}) + \gamma_n \right\} = 0 \quad (270)$$

と変形できる．式 (270) を確率的上極限を用いて表すと

$$\text{p-lim sup}_{n \rightarrow \infty} \left\{ \frac{1}{n} d_n(X^n, \psi(\varphi(X^n))) - \gamma_n \right\} \leq D(\mathbf{X}, \mathbf{Y}) \quad (271)$$

と書ける． $n \rightarrow \infty$ のとき $\gamma_n \rightarrow 0$ なので

$$\text{p-lim sup}_{n \rightarrow \infty} \frac{1}{n} d_n(X^n, \psi(\varphi(X^n))) \leq D(\mathbf{X}, \mathbf{Y}) \quad (272)$$

となる．さらに， $M_n \triangleq e^{n(\bar{I}(\mathbf{X}; \mathbf{Y}) + 2\gamma)}$ なので

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log M_n \leq \bar{I}(\mathbf{X}; \mathbf{Y}) + 2\gamma \quad (273)$$

である．式 (272) と式 (273) より $\bar{I}(\mathbf{X}; \mathbf{Y}) + 2\gamma$ は D -達成可能である．つまりレート・歪み関数は定義より

$$R(D) \leq \bar{I}(\mathbf{X}; \mathbf{Y}) + 2\gamma \quad (274)$$

であるが， $\gamma > 0$ は任意なので

$$R(D) \leq \bar{I}(\mathbf{X}; \mathbf{Y}) \quad (275)$$

となる．さらに $\mathbf{Y}$ の任意性より

$$R(D) \leq \inf_{\mathbf{Y}: D(\mathbf{X}, \mathbf{Y}) \leq D} \bar{I}(\mathbf{X}; \mathbf{Y}) \quad (276)$$

となり順定理が示せた．

A.2 定常無記憶情報源のレート・歪み理論

本稿では一般情報源について扱ってきたが，最も基本的な情報源である定常無記憶情報源のレート・歪み理論もランダム符号化を用いる方法で証明されている．本節に限り $\mathbf{X} = \{X^n = (X_1, X_2, \dots, X_n)\}_{n=1}^{\infty}$ を定常無記憶情報源とし，情報源アルファベット $\mathcal{X}$ を有限集合とする．そして歪み測度は加法的なものを仮定する．さらに以下の準備をする．

(Markov の不等式)

Z を非負値の任意確率変数として，任意の $a > 0$ に対して

$$\Pr\{Z \geq a\} \leq \frac{E[Z]}{a} \quad (277)$$

が成り立つ.

(証明)

Z の実現値を z とすると期待値は

$$E[Z] = \sum_z z P_Z(z) \quad (278)$$

$$= \sum_{z>a} z P_Z(z) + \sum_{z\leq a} z P_Z(z) \quad (279)$$

$$\geq \sum_{z>a} a P_Z(z) \quad (280)$$

$$= a \Pr\{Z > a\} \quad (281)$$

と変形でき,

$$\Pr\{Z > a\} \leq \frac{E[Z]}{a} \quad (282)$$

が導かれる.

(Chebyshev の不等式)

Z を任意の確率変数とする. その平均を μ , 分散を σ^2 とするとき, 任意の $a > 0$ に対して

$$\Pr\{|Z - \mu| > a\sigma\} \leq \frac{1}{a^2} \quad (283)$$

が成り立つ.

(証明)

Markov の不等式より, Z を $(Z - \mu)^2$ に, a を $a^2\sigma^2$ に置き換えると

$$\Pr\{(Z - \mu)^2 > a^2\sigma^2\} \leq \frac{E[(Z - \mu)^2]}{a^2\sigma^2} \quad (284)$$

$E[(Z - \mu)^2] = \sigma^2$ なので

$$\Pr\{|Z - \mu| > a\sigma\} \leq \frac{1}{a^2} \quad (285)$$

が導かれる.

では定常無記憶情報源の達成可能性を定義する.

定義 26 ([1] 定義 5.1) 平均歪みを

$$\rho_n \triangleq \frac{1}{n} E d_n(X^n, \psi_n(\varphi_n(X^n))) \quad (286)$$

としたときレート R が D -達成可能とは

$$\limsup_{n \rightarrow \infty} \rho_n \leq D \quad (287)$$

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log M_n \leq R \quad (288)$$

を満たす (n, M_n, ρ_n) 符号が存在することである。

一般情報源のときと同様以下の定理がいえる。

定理 9 ([1] 定理 5.1) 分布 P_X に従う定常無記憶情報源 $\mathbf{X}$ のレート・歪み関数は

$$R(D) = \min_{\mathbf{Y}: \mathbf{Ed}(\mathbf{X}, \mathbf{Y}) \leq D} I(\mathbf{X}; \mathbf{Y}) \quad (289)$$

である。ここで

$$I(\mathbf{X}; \mathbf{Y}) \triangleq \sum_{(x, y) \in \mathcal{X} \times \mathcal{Y}} P_{XY}(x, y) \log \frac{P_{Y|X}(y|x)}{P_Y(y)} \quad (290)$$

とし、これを相互情報量という。

(証明)[1]

任意の $\gamma > 0$ に対し、条件 $\mathbf{Ed}(\mathbf{X}, \mathbf{Y}) \leq D - \gamma$ を満たすような任意の $\mathbf{Y}$ を考える。このとき $R = I(\mathbf{X}; \mathbf{Y}) + 2\gamma$ が D -達成可能であることを示す。 $X^n Y^n = (X_1 Y_1, X_2 Y_2, \dots, X_n Y_n)$ を XY の定常独立系列として $X^n = (X_1, X_2, \dots, X_n)$, $Y^n = (Y_1, Y_2, \dots, Y_n)$ とおく。そこで

$$\mathcal{T}_n^{(1)} = \left\{ (\mathbf{x}, \mathbf{y}) \in \mathcal{X}^n \times \mathcal{Y}^n \left| \frac{1}{n} \log \frac{P_{Y^n|X^n}(\mathbf{y}|\mathbf{x})}{P_{Y^n}(\mathbf{y})} - I(\mathbf{X}; \mathbf{Y}) \right| < \gamma \right\} \quad (291)$$

$$\mathcal{T}_n^{(2)} = \left\{ (\mathbf{x}, \mathbf{y}) \in \mathcal{X}^n \times \mathcal{Y}^n \left| \frac{1}{n} d_n(\mathbf{x}, \mathbf{y}) - \mathbf{Ed}_n(\mathbf{X}, \mathbf{Y}) \right| < \gamma \right\} \quad (292)$$

とおき $\mathcal{T}_n = \mathcal{T}_n^{(1)} \cap \mathcal{T}_n^{(2)}$ とする。ここで仮定した条件 $\mathbf{Ed}(\mathbf{X}, \mathbf{Y}) \leq D - \gamma$ と式 (292) より $(\mathbf{x}, \mathbf{y}) \in \mathcal{T}_n$ なら $\frac{1}{n} d_n(\mathbf{x}, \mathbf{y}) \leq D$ となることに注意しておく。さらに $X^n Y^n$ は定常独立系列なので

$$\frac{1}{n} \log \frac{P_{Y^n|X^n}(Y^n|X^n)}{P_{Y^n}(Y^n)} = \frac{1}{n} \log \prod_{i=1}^n \frac{P_{Y|X}(Y_i|X_i)}{P_Y(Y_i)} \quad (293)$$

$$= \frac{1}{n} \sum_{i=1}^n \log \frac{P_{Y|X}(Y_i|X_i)}{P_Y(Y_i)}, \quad (i = 1, 2, \dots) \quad (294)$$

$$\frac{1}{n} d_n(X^n, Y^n) = \frac{1}{n} \sum_{i=1}^n d_n(X_i, Y_i) \quad (295)$$

が成り立つ．上式の各々の和の中の各項は独立であるので，これらの期待値は

$$\frac{1}{n} \sum_{i=1}^n \mathbf{E} \left[\log \frac{P_{Y|X}(Y_i|X_i)}{P_Y(Y_i)} \right] = \frac{1}{n} \sum_{i=1}^n \sum_{\mathbf{x}, \mathbf{y}} P_{X_i Y_i}(\mathbf{x}, \mathbf{y}) \log \frac{P_{Y_i|X_i}(\mathbf{y}|\mathbf{x})}{P_{Y_i}(\mathbf{y})} \quad (296)$$

$$= \frac{1}{n} \sum_{i=1}^n I(X_i; Y_i) \quad (297)$$

$$= \frac{1}{n} \sum_{i=1}^n I(X; Y) \quad (298)$$

$$= I(X; Y) \quad (299)$$

$$\frac{1}{n} \sum_{i=1}^n \mathbf{E} d_n(X_i, Y_i) = \frac{1}{n} \sum_{i=1}^n \sum_{\mathbf{x}, \mathbf{y}} P_{X_i Y_i}(\mathbf{x}, \mathbf{y}) d_n(\mathbf{x}, \mathbf{y}) \quad (300)$$

$$= \frac{1}{n} \sum_{i=1}^n \sum_{\mathbf{x}, \mathbf{y}} P_{XY}(\mathbf{x}, \mathbf{y}) d_n(\mathbf{x}, \mathbf{y}) \quad (301)$$

$$= \frac{1}{n} \sum_{i=1}^n \mathbf{E} d_n(X, Y) \quad (302)$$

$$= \mathbf{E} d_n(X, Y) \quad (303)$$

となる．さらに式 (294) と式 (295) の和の中の分散をそれぞれ $\sigma_{T1}^2, \sigma_{T2}^2$ とする．ここでそれらは有界であるとし，証明は後述する．すると式 (294) と式 (295) の分散は

$$\sigma_1^2 = \text{Var} \left[\frac{1}{n} \sum_{i=1}^n \log \frac{P_{Y|X}(Y_i|X_i)}{P_Y(Y_i)} \right] = \frac{1}{n^2} \sum_{i=1}^n \text{Var} \left[\log \frac{P_{Y|X}(Y_i|X_i)}{P_Y(Y_i)} \right] \quad (304)$$

$$= \frac{\sigma_{T1}^2}{n} \quad (305)$$

$$\sigma_2^2 = \text{Var} \left[\frac{1}{n} \sum_{i=1}^n d_n(X_i, Y_i) \right] = \frac{1}{n^2} \sum_{i=1}^n \text{Var} [d_n(X_i, Y_i)] \quad (306)$$

$$= \frac{\sigma_{T2}^2}{n} \quad (307)$$

つまり Chebyshev の不等式より

$$\Pr\{X^n Y^n \notin \mathcal{T}_n^{(1)}\} = \Pr \left\{ \left| \frac{1}{n} \sum_{i=1}^n \log \frac{P_{Y|X}(Y_i|X_i)}{P_Y(Y_i)} - I(X; Y) \right| > \gamma \right\} \quad (308)$$

$$\leq \frac{\sigma_{T1}^2}{n\gamma^2} \quad (309)$$

$$\rightarrow 0 \quad (n \rightarrow \infty) \quad (310)$$

$$\Pr\{X^n Y^n \notin \mathcal{T}_n^{(2)}\} = \Pr\left\{\left|\frac{1}{n}d_n(\mathbf{x}, \mathbf{y}) - \mathbf{E}d_n(X, Y)\right| > \gamma\right\} \quad (311)$$

$$\leq \frac{\sigma_{T2}^2}{n\gamma^2} \quad (312)$$

$$\rightarrow 0 \quad (n \rightarrow \infty) \quad (313)$$

よって

$$\Pr\{X^n Y^n \notin \mathcal{T}_n\} \rightarrow 0 \quad (314)$$

が成立する．ここで

$$M_n = e^{-nR} = e^{-n(I(X;Y)+2\gamma)} \quad (315)$$

として, $v_1, v_2, \dots, v_{M_n} \in \mathcal{Y}^n$ を互いに独立に分布 P_{Y^n} にしたがって発生させ $C = \{v_1, v_2, \dots, v_{M_n}\}$ とおく．さらに各 $\mathbf{x}$ に対して

$$d_n(\mathbf{x}, v_{i0}) = \min_{1 \leq i \leq M_n} d_n(\mathbf{x}, v_i) \quad (316)$$

となるような i_0 を定め, 符号器 φ_n を $\varphi_n(x) = i_0$ と定義する．一方, 復号器 ψ_n は符号器から番号 i を受け取ったとき $\psi_n(i) = v_i$ と定める．さらに

$$\overline{P}_e^{(n)} \triangleq \Pr\left\{\frac{1}{n}d_n(X^n, \psi(\varphi(X^n))) > D\right\} \quad (317)$$

と定義する．ここで任意の $a > 0$ を用いて確率変数 Z とその実現値 z に対し

$$\mathbf{E}[Z] = \sum_z z P_Z(z) \quad (318)$$

$$= \sum_{z \leq a} z P_Z(z) + \sum_{z > a} z P_Z(z) \quad (319)$$

$$\leq a \sum_{z \leq a} P_Z(z) + \max Z \sum_{z > a} P_Z(z) \quad (320)$$

$$\leq a + \max Z \sum_{z > a} P_Z(z) \quad (321)$$

となることに注意して, $d_{max} = \max_{x \in \mathcal{X}, y \in \mathcal{Y}} d(x, y)$ とおくと

$$\overline{\rho}_n \triangleq \frac{1}{n} \mathbf{E}_C \mathbf{E}d_n(X^n, \psi(\varphi(X^n))) \leq D + \overline{P}_e^{(n)} d_{max} \quad (322)$$

となる．ここで E, E_C はそれぞれ X^n と符号 C の期待値を表す．ここで前節の定理 1 のランダム符号の証明と同様の方法で $\overline{P}_e^{(n)}$ を上から抑えられる．すなわち

$$\overline{P}_e^{(n)} \leq \Pr\{X^n Y^n \notin \mathcal{T}_n\} + e^{-e^{n\gamma}} \quad (323)$$

である．ゆえに $\Pr\{X^n Y^n \notin \mathcal{T}_n\} \rightarrow 0$ より $\overline{P}_e^{(n)} \rightarrow 0 (n \rightarrow \infty)$ が得られる．式 (322) により，少なくとも一つのランダムでない符号が存在して

$$\bar{\rho}_n \triangleq \frac{1}{n} \mathbf{E}_C \mathbf{E} d_n(X^n, \psi(\varphi(X^n))) \leq D + \overline{P}_e^{(n)} d_{max} \quad (324)$$

を満たさなくてはならない．ゆえに

$$\limsup_{n \rightarrow \infty} \bar{\rho}_n \leq D \quad (325)$$

さらに $M_n = e^{-nR} = e^{-n(I(X;Y)+2\gamma)}$ なので

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log M_n \leq I(X;Y) + 2\gamma \quad (326)$$

である．よって $R = I(X;Y) + 2\gamma$ は D -達成可能である．また，条件 $\mathbf{E} d(X,Y) \leq D - \gamma$ を満たすように Y を選んだので

$$R(D) \leq \min_{\mathbf{Y}: \mathbf{E} d(X,Y) \leq D-\gamma} I(X;Y) + 2\gamma \quad (327)$$

となり， $\gamma > 0$ は任意だったので $\gamma \rightarrow 0$ とすれば

$$R(D) \leq \min_{\mathbf{Y}: \mathbf{E} d(X,Y) \leq D} I(X;Y) \quad (328)$$

が導け，順定理が示せた．□

証明中に用いた分散の一樣有界性を示す．[[1] 注意 3.1]

入出力アルファベット $\mathcal{X}, \mathcal{Y}$ のいずれか一方が有界であるとする．式 (294) の和の中を

$$Z = \log \frac{P_{Y|X}(Y|X)}{P_Y(Y)} \quad (329)$$

とおく．まず Z の期待値 $E(Z)$ は相互情報量 $I(X;Y) \triangleq \sum_{(x,y) \in \mathcal{X} \times \mathcal{Y}} P_{XY}(x,y) \log \frac{P_{Y|X}(y|x)}{P_Y(y)}$ となるので

$$\mathbf{E}[Z] = I(X;Y) \quad (330)$$

$$= H(Y) - H(Y|X) \quad (331)$$

と書かれる．ここで一般性を失うことなく，出力アルファベット $\mathcal{Y}$ が有限であるとする．すると， $H(Y|X) \geq 0$ に注意すると

$$0 \leq \mathbf{E}[Z] \leq H(Y) \leq \log |\mathcal{Y}| \quad (332)$$

が得られる．すなわち $E[Z]$ は一樣に有界である．次に Z の分散 $\text{Var}[Z]$ は

$$\text{Var}[Z] = \mathbf{E}[Z^2] - (\mathbf{E}[Z])^2 \quad (333)$$

となるので Z が有界であることを言うには $\mathbf{E}[Z^2]$ が有界であることを示せばよい．そのために $0 \geq P_Y(y) \geq 1$, $0 \geq P_{Y|X}(y, x) \geq 1$ に注意して, $\mathbf{E}[Z^2]$ を

$$\mathbf{E}[Z^2] = \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} P_{XY}(x, y) \left(\log \frac{P_{Y|X}(y|x)}{P_Y(y)} \right)^2 \quad (334)$$

$$= \sum_{(x, y) \in \mathcal{X} \times \mathcal{Y}: P_{Y|X}(y|x) \leq P_Y(y)} P_{XY}(x, y) \left(\log \frac{P_{Y|X}(y|x)}{P_Y(y)} \right)^2 \\ + \sum_{(x, y) \in \mathcal{X} \times \mathcal{Y}: P_{Y|X}(y|x) \geq P_Y(y)} P_{XY}(x, y) \left(\log \frac{P_{Y|X}(y|x)}{P_Y(y)} \right)^2 \quad (335)$$

$$= \sum_{(x, y) \in \mathcal{X} \times \mathcal{Y}: P_{Y|X}(y|x) \leq P_Y(y)} P_{XY}(x, y) \left(\log \frac{P_Y(y)}{P_{Y|X}(y|x)} \right)^2 \\ + \sum_{(x, y) \in \mathcal{X} \times \mathcal{Y}: P_{Y|X}(y|x) \geq P_Y(y)} P_{XY}(x, y) \left(\log \frac{P_{Y|X}(y|x)}{P_Y(y)} \right)^2 \quad (336)$$

$$\leq \sum_{(x, y) \in \mathcal{X} \times \mathcal{Y}: P_{Y|X}(y|x) \leq P_Y(y)} P_{XY}(x, y) \left(\log \frac{1}{P_{Y|X}(y|x)} \right)^2 \\ + \sum_{(x, y) \in \mathcal{X} \times \mathcal{Y}: P_{Y|X}(y|x) \geq P_Y(y)} P_{XY}(x, y) \left(\log \frac{1}{P_Y(y)} \right)^2 \quad (337)$$

$$\leq \sum_{(x, y) \in \mathcal{X} \times \mathcal{Y}} P_{XY}(x, y) \left(\log \frac{1}{P_{Y|X}(y|x)} \right)^2 \\ + \sum_{(x, y) \in \mathcal{X} \times \mathcal{Y}} P_{XY}(x, y) \left(\log \frac{1}{P_Y(y)} \right)^2 \quad (338)$$

$$= \sum_{x \in \mathcal{X}} P_X(x) \sum_{y \in \mathcal{Y}} P_{Y|X}(y|x) \left(\log \frac{1}{P_{Y|X}(y|x)} \right)^2 \\ + \sum_{y \in \mathcal{Y}} P_Y(y) \left(\log \frac{1}{P_Y(y)} \right)^2 \quad (339)$$

と変形する．そこで, 関数

$$g(u) = u \left(\log \frac{1}{u} \right)^2 \quad (0 \leq u \leq 1) \quad (340)$$

を考える． $t = \log \frac{1}{u}$ とおけば, $g(u)$ は

$$h(t) = t^2 e^{-t} \quad (0 \leq t \leq +\infty) \quad (341)$$

と書ける．すると， $h(t)$ は $t = 2$ で最大値をとり

$$\max_{0 \leq u \leq 1} g(u) = \max_{0 \leq t \leq +\infty} h(t) = \frac{4}{e^2} \quad (342)$$

となる．したがって，式 (342) を式 (339) に適用すると

$$\mathbf{E}[Z^2] \leq \sum_{x \in \mathcal{X}} P_X(x) \sum_{y \in \mathcal{Y}} \frac{4}{e^2} + \sum_{y \in \mathcal{Y}} P_Y(y) \frac{4}{e^2} \quad (343)$$

$$\leq \sum_{x \in \mathcal{X}} P_X(x) \frac{4|\mathcal{Y}|}{e^2} + \frac{4|\mathcal{Y}|}{e^2} \quad (344)$$

$$\leq \frac{8|\mathcal{Y}|}{e^2} \quad (345)$$

を得る．よって Z の分散 $\text{Var}(Z)$ は一様に有界である．

付録 B ε -通信路符号化 (Feinstein の補題)

通信路符号化定理は誤り確率をある与えられた基準値以下にするような符号器復号器の組において，符号化レートを最大限どこまで大きくできるかという形で定式化される．ここでは誤り確率の基準を ε とし，通信路符号化を考える．本章では $\mathcal{X}$ を入力アルファベット， $\mathcal{Y}$ を出力アルファベットとして通信路の遷移確率と呼ばれる条件付確率の組 $W = \{W(y|x)\}_{x \in \mathcal{X}, y \in \mathcal{Y}}$ を条件 $\sum_{y \in \mathcal{Y}} W(y|x) = 1$ ($\forall x \in \mathcal{X}$) を満たすように定める．そして伝送すべきメッセージの集合を $\mathcal{M}_n = \{1, 2, \dots, M_n\}$ としたとき，写像 $\varphi_n: \mathcal{M}_n \rightarrow \mathcal{X}^n$ を符号器とする．ここで $\mathbf{u}_i = \varphi_n(i)$ を符号語と呼び， $\mathcal{C}_n = \{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_{M_n}\}$ を符号と呼ぶ．一方受信側ではあらかじめ定めておいた $\mathcal{Y}^n$ の排反な分割 $\mathcal{Y}^n = \{\mathcal{D}_1 \cup \mathcal{D}_2\} \cup \dots \cup \mathcal{D}_{M_n}$ ($\mathcal{D}_i \cap \mathcal{D}_j = \emptyset$ for $i \neq j$) に基づいて，受け取った出力 $\mathbf{y}$ が $\mathbf{y} \in \mathcal{D}_i$ ならば $i \in \mathcal{M}_n$ が送信されてきたと判定する．このような操作を復号といい，これを表す写像 $\psi_n: \mathcal{Y}^n \rightarrow \mathcal{M}_n$ を復号器と呼ぶ．また通信路符号化における復号誤り確率 ε_n は

$$\varepsilon_n = \frac{1}{M_n} \sum_{i=1}^{M_n} W^n(\mathcal{D}_i^c | \mathbf{u}_i) \quad (346)$$

と定義する．[1] 本章では通信路を一般通信路とする．

まず達成可能性を定義する．

定義 27 レート R が ε -達成可能であるとは

$$\limsup_{n \rightarrow \infty} \varepsilon_n < \varepsilon \quad (347)$$

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \log M_n \geq R \quad (348)$$

となる (n, M_n, ε_n) 符号が存在することである．

定義 28 ε -通信路容量は次のように定められる.

$$C(\varepsilon | \mathbf{W}) = \sup \{ R | R \text{ が } \varepsilon\text{-達成可能} \} \quad (349)$$

ここで通信路 $\mathbf{W}$ を次のように定義する. $\mathbf{X} = \{X^n\}_{n=1}^\infty$ を任意の一般情報源として, 入力を $\mathbf{X}$, 出力を $\mathbf{Y}$ とする. すると

$$P_{X^n Y^n}(\mathbf{x}, \mathbf{y}) = P_{X^n}(\mathbf{x}) W^n(\mathbf{y} | \mathbf{x}) \quad (350)$$

となる.

次の補題が本論文で参考にした Feinstein の補題である.

補題 3 ([1] 補題 2.4) $\mathbf{X} = \{X^n\}_{n=1}^\infty$ を任意の通信路入力としそれに対する通信路 $\mathbf{W} = \{W^n\}_{n=1}^\infty$ からの出力を $\mathbf{Y} = \{Y^n\}_{n=1}^\infty$ とする. M_n を任意に与えられた正の整数とすると, すべての $n = 1, 2, \dots$ に対して

$$\varepsilon_n \leq \Pr \left\{ \frac{1}{n} \log \frac{W^n(Y^n | X^n)}{P_{Y^n}(Y^n)} \leq \frac{1}{n} \log M_n + \gamma \right\} + e^{-n\gamma} \quad (351)$$

を満たす (n, M_n, ε_n) 符号が存在する. ただし, $\gamma > 0$ は任意の定数である.

(証明)

補題を変形する. 初めに

$$i(\mathbf{x}; \mathbf{y}) \triangleq \log \frac{W^n(\mathbf{y} | \mathbf{x})}{P_{Y^n}(\mathbf{y})} \quad (\mathbf{x} \in \mathcal{X}^n, \mathbf{y} \in \mathcal{Y}^n) \quad (352)$$

とおき, さらに $\alpha_n = M_n e^{n\gamma}$ とすると, 補題の右辺は次のように変形ができ, それを λ_n とおく.

$$\lambda_n \triangleq \Pr\{i(X^n; Y^n) \leq \log \alpha_n\} + \frac{M_n}{\alpha_n} \quad (353)$$

となり, 補題は $\varepsilon_n \leq \lambda_n$ と書き換えられる. つまり補題の主張を示すには, ある (n, M_n, ε_n) 符号が存在して, $\varepsilon_n \leq \lambda_n$ を満たせばよい. このとき $\lambda_n \geq 1$ のときは自明なので, 以下では $\lambda_n < 1$ を仮定する. まず, $\mathbf{x} \in \mathcal{X}^n$ に対し

$$B(\mathbf{x}) = \{\mathbf{y} \in \mathcal{Y} | i(\mathbf{x}; \mathbf{y}) > \log \alpha_n\} \quad (354)$$

とおく.

通信路に対する符号 C_n を次のように順次構成していく. 初めに条件

$$W^n(B(\mathbf{u}_1) | \mathbf{u}_1) \geq 1 - \lambda_n \quad (355)$$

を満たす $\mathbf{u}_1 \in \mathcal{X}^n$ を任意に選び $\mathcal{D}_1 = B(\mathbf{u}_1)$ とおく．次に

$$W^n(B(\mathbf{u}_2) - \mathbf{u}_1 | \mathbf{u}_2) \geq 1 - \lambda_n \quad (356)$$

となるような $\mathbf{u}_2 \in \mathcal{X}^n$ を任意に選んで $\mathcal{D}_2 = B(\mathbf{u}_2) - B(\mathbf{u}_1)$ とおく．このような操作を可能な限り続けて行い， L_n 回目までで停止したとする．すると $k = 1, 2, \dots, L_n$ に対して

$$W^n(B(\mathbf{u}_k) - \sum_{i=1}^{k-1} B(\mathbf{u}_i) | \mathbf{u}_k) \geq 1 - \lambda_n \quad (357)$$

が成立する．さらに L_n の定義より，すべての $\mathbf{x} \in \mathcal{X}^n$ に対して

$$W^n(B(\mathbf{x}) - \sum_{i=1}^{L_n} B(\mathbf{u}_i) | \mathbf{x}) < 1 - \lambda_n \quad (358)$$

となる．ここで

$$D_k = B(\mathbf{u}_k) - \sum_{i=1}^{k-1} B(\mathbf{u}_i) \quad (359)$$

とおくことにより，式 (357) は

$$W^n(D_k | \mathbf{u}_k) \geq 1 - \lambda_n \quad (360)$$

と書かれる． D_k を $\mathbf{u}_k$ に対する復号領域とし，符号を $C_n = \{\mathbf{u}_1, \dots, \mathbf{u}_{L_n}\}$ とする．この符号の誤り確率 ε'_n は

$$\varepsilon'_n \triangleq \frac{1}{L_n} \sum_{k=1}^{L_n} W^n(\mathcal{D}_k^c | \mathbf{u}_k) \leq \lambda_n \quad (361)$$

となる．ここで L_n と M_n の関係を考える． $D = \bigcup_{k=1}^{L_n} D_k$ とする．式 (354) について以下の式が成り立つ．

$$\Pr\{i(X^n; Y^n) \leq \log \alpha_n\} \quad (362)$$

$$= \sum_{\mathbf{x} \in \mathcal{X}^n} \Pr\{i(X^n; Y^n) \leq \log \alpha_n, X^n = \mathbf{x}\} \quad (363)$$

$$= \sum_{\mathbf{x} \in \mathcal{X}^n} \Pr\{i(\mathbf{x}; Y^n) \leq \log \alpha_n | X^n = \mathbf{x}\} P_{X^n}(\mathbf{x}) \quad (364)$$

$$= \sum_{\mathbf{x} \in \mathcal{X}^n} W^n(B(\mathbf{x}) | \mathbf{x}) P_{X^n}(\mathbf{x}) \quad (365)$$

$$= \sum_{\mathbf{x} \in \mathcal{X}^n} W^n(B(\mathbf{x}) \cap \mathcal{D} | \mathbf{x}) P_{X^n}(\mathbf{x}) + \sum_{\mathbf{x} \in \mathcal{X}^n} W^n(B(\mathbf{x}) \cap \mathcal{D}^c | \mathbf{x}) P_{X^n}(\mathbf{x}) \quad (366)$$

となる．さらに式 (366) の右辺の第 1 項は

$$\sum_{\mathbf{x} \in \mathcal{X}^n} W^n(B(\mathbf{x}) \cap \mathcal{D}|\mathbf{x})P_{X^n}(\mathbf{x}) \quad (367)$$

$$\leq \sum_{\mathbf{x} \in \mathcal{X}^n} W^n(\mathcal{D}|\mathbf{x})P_{X^n}(\mathbf{x}) \quad (368)$$

$$= P_{Y^n}(\mathcal{D}) = \sum_{k=1}^{L_n} P_{Y^n}(\mathcal{D}_k) \quad (369)$$

$$\leq \sum_{k=1}^{L_n} P_{Y^n}(B(\mathbf{u}_k)) \quad (370)$$

$$= \sum_{k=1}^{L_n} \Pr\{Y^n \in B(\mathbf{u}_k)\} \quad (371)$$

と変形できる．ここで任意の $\mathbf{x} \in \mathcal{X}$ に対して，式 (352) と式 (354) より $\mathbf{y} \in B(\mathbf{x})$ なら $W^n(\mathbf{y}|\mathbf{x}) \geq \alpha_n P_{Y^n}(\mathbf{y})$ なので

$$1 \geq W^n(B(\mathbf{x})|\mathbf{x}) \quad (372)$$

$$= \sum_{\mathbf{y} \in B(\mathbf{x})} W^n(\mathbf{y}|\mathbf{x}) \quad (373)$$

$$\geq \alpha_n \sum_{\mathbf{y} \in B(\mathbf{x})} P_{Y^n}(\mathbf{y}) \quad (374)$$

$$= \alpha_n \Pr\{Y^n \in B(\mathbf{u}_k)\} \quad (375)$$

となることから

$$\Pr\{Y^n \in B(\mathbf{u}_k)\} \leq \frac{1}{\alpha_n} \quad (376)$$

なので，式 (371) に式 (376) を代入すると

$$\sum_{\mathbf{x} \in \mathcal{X}^n} W^n(B(\mathbf{x}) \cap \mathcal{D}|\mathbf{x})P_{X^n}(\mathbf{x}) \leq \frac{L_n}{\alpha_n} \quad (377)$$

が得られる．そして式 (366) の右辺の第 2 項は

$$\sum_{\mathbf{x} \in \mathcal{X}^n} W^n(B(\mathbf{x}) \cap \mathcal{D}^c | \mathbf{x}) P_{X^n}(\mathbf{x}) \quad (378)$$

$$= \sum_{\mathbf{x} \in C_n, \mathbf{x} \in \mathcal{X}^n} W^n(B(\mathbf{x}) \cap \mathcal{D}^c | \mathbf{x}) P_{X^n}(\mathbf{x}) + \sum_{\mathbf{x} \notin C_n, \mathbf{x} \in \mathcal{X}^n} W^n(B(\mathbf{x}) \cap \mathcal{D}^c | \mathbf{x}) P_{X^n}(\mathbf{x}) \quad (379)$$

$$= 0 + \sum_{\mathbf{x} \notin C_n, \mathbf{x} \in \mathcal{X}^n} W^n(B(\mathbf{x}) \cap \mathcal{D}^c | \mathbf{x}) P_{X^n}(\mathbf{x}) \quad (380)$$

$$< 1 - \lambda_n \quad (381)$$

となる．式 (380) の第 1 項は， $\mathbf{x} \in \{\mathbf{u}_1, \dots, \mathbf{u}_{L_n}\}$ のときは $B(\mathbf{x}) \subset \mathcal{D}$ なので 0 となり，第 2 項は式 (358) を用いた．式 (377) と式 (381) を式 (366) に代入すると，

$$\Pr\{i(X^n; Y^n) \leq \log \alpha_n\} < \frac{L_n}{\alpha_n} + 1 - \lambda_n \quad (382)$$

したがって

$$\lambda_n < \frac{L_n}{\alpha_n} + 1 - \Pr\{i(X^n; Y^n) \leq \log \alpha_n\} \quad (383)$$

が導かれ，さらに

$$\lambda_n < \frac{L_n}{\alpha_n} + \Pr\{i(X^n; Y^n) > \log \alpha_n\} \quad (384)$$

と変形できる．これにより λ_n の定義式 (353) より， $M_n < L_n$ が得られる．つまり，符号 $C_n = \{\mathbf{u}_1, \dots, \mathbf{u}_{L_n}\}$ から部分符号 $C_n^* = \{\mathbf{u}_1, \dots, \mathbf{u}_{M_n}\}$ が取り出すことができる．式 (361) と同様に

$$\varepsilon_n \triangleq \frac{1}{M_n} \sum_{k=1}^{M_n} W^n(\mathcal{D}_k^c | \mathbf{u}_k) \leq \lambda_n \quad (385)$$

が得られ， $\varepsilon_n \leq \lambda_n$ が導けた．よって

$$\varepsilon_n \leq \Pr \left\{ \frac{1}{n} \log \frac{W^n(Y^n | X^n)}{P_{Y^n}(Y^n)} \leq \frac{1}{n} \log M_n + \gamma \right\} + e^{-n\gamma} \quad (386)$$

を満たす (n, M_n, ε_n) 符号が存在することがいえた．□

また逆定理を導くために次の補題も使われる．

補題 4 ([1] 補題 3.2) (n, M_n, ε_n) 符号 C_n 上の一様分布に従う確率変数を X^n とし, X^n を入力とする通信路 $\mathbf{W} = \{W^n\}_{n=1}^\infty$ の出力を Y^n とすると, 各 $n = 1, 2, \dots$ に対して

$$\varepsilon_n \geq \Pr \left\{ \frac{1}{n} \log \frac{W^n(Y^n|X^n)}{P_{Y^n}(Y^n)} \leq \frac{1}{n} \log M_n - \gamma \right\} - e^{-n\gamma} \quad (387)$$

が成立する. ($\gamma > 0$ は任意の定数)

(証明)[1]

$$\frac{1}{n} \log \frac{W^n(Y^n|X^n)}{P_{Y^n}(Y^n)} = \frac{1}{n} \log \frac{P_{X^n|Y^n}(X^n|Y^n)}{P_{X^n}(X^n)} \quad (388)$$

と書き換え, $P_{X^n}(\mathbf{x}) = \frac{1}{M_n}$ であることに注意すると

$$\frac{1}{n} \log \frac{P_{X^n|Y^n}(X^n|Y^n)}{P_{X^n}(X^n)} = \frac{1}{n} \log P_{X^n|Y^n}(X^n|Y^n) + \frac{1}{n} \log \frac{1}{P_{X^n}(X^n)} \quad (389)$$

$$= \frac{1}{n} \log P_{X^n|Y^n}(X^n|Y^n) + \frac{1}{n} \log M_n \quad (390)$$

であるので補題は次のように書き換えられ

$$\varepsilon_n \geq \Pr \left\{ \frac{1}{n} \log P_{X^n|Y^n}(X^n|Y^n) + \frac{1}{n} \log M_n \leq \frac{1}{n} \log M_n - \gamma \right\} - e^{-n\gamma} \quad (391)$$

$$\Leftrightarrow \varepsilon_n \geq \Pr \{ P_{X^n|Y^n}(X^n|Y^n) \leq e^{-n\gamma} \} - e^{-n\gamma} \quad (392)$$

となる. ここで符号語数 M_n , 誤り確率 ε_n の任意の符号を考える. メッセージ i ($i = 1, 2, \dots, M_n$) に対して

$$\mathcal{D}_i \triangleq \{ \mathbf{y} \in \mathcal{Y}^n | \psi_n(\mathbf{y}) = i \} = \psi_n^{-1}(i) \quad (393)$$

$$\mathcal{B}_i \triangleq \{ \mathbf{y} \in \mathcal{Y}^n | P_{X^n|Y^n}(\varphi_n(i)|\mathbf{y}) \leq e^{-n\gamma} \} \quad (394)$$

と領域を定義する. これらを用い, 補題の右辺第 1 項の式を変形していくと

$$\Pr \{ P_{X^n|Y^n}(X^n|Y^n) \leq e^{-n\gamma} \} \quad (395)$$

$$= \sum_{\mathbf{x}} \Pr \{ X^n = \mathbf{x}, P_{X^n|Y^n}(X^n|Y^n) \leq e^{-n\gamma} \} \quad (396)$$

$$= \sum_{i=1}^{M_n} \Pr \{ X^n = \varphi_n(i), P_{X^n|Y^n}(X^n|Y^n) \leq e^{-n\gamma} \} \quad (397)$$

$$= \sum_{i=1}^{M_n} \Pr \{ X^n = \varphi_n(i), Y^n \in \mathcal{B}_i \} \quad (398)$$

$$\begin{aligned} &= \sum_{i=1}^{M_n} \Pr \{ X^n = \varphi_n(i), Y^n \in \mathcal{B}_i \cup \mathcal{D}_i^c \} \\ &\quad + \sum_{i=1}^{M_n} \Pr \{ X^n = \varphi_n(i), Y^n \in \mathcal{B}_i \cup \mathcal{D}_i \} \end{aligned} \quad (399)$$

となる．まず式 (399) の第 1 項を考えると

$$\sum_{i=1}^{M_n} \Pr \{X^n = \varphi_n(i), Y^n \in \mathcal{B}_i \cup \mathcal{D}_i^c\} \quad (400)$$

$$\leq \sum_{i=1}^{M_n} \Pr \{X^n = \varphi_n(i), Y^n \in \mathcal{D}_i^c\} \quad (401)$$

$$= \sum_{i=1}^{M_n} P_{X^n}(\varphi_n(i)) P_{Y^n|X^n}(\mathcal{D}_i^c | \varphi_n(i)) \quad (402)$$

$$= \sum_{i=1}^{M_n} \frac{1}{M_n} P_{Y^n|X^n}(\mathcal{D}_i^c | \varphi_n(i)) \quad (403)$$

$$= \frac{1}{M_n} \sum_{i=1}^{M_n} W^n(\mathcal{D}_i^c | \varphi_n(i)) \quad (404)$$

$$= \varepsilon_n \quad (405)$$

であり，第 2 項は

$$\sum_{i=1}^{M_n} \Pr \{X^n = \varphi_n(i), Y^n \in \mathcal{B}_i \cup \mathcal{D}_i\} \quad (406)$$

$$= \sum_{i=1}^{M_n} \sum_{Y^n \in \mathcal{B}_i \cup \mathcal{D}_i} P_{Y^n}(\mathbf{y}) P_{X^n|Y^n}(\varphi_n(i) | \mathbf{y}) \quad (407)$$

$$\leq \sum_{i=1}^{M_n} \sum_{Y^n \in \mathcal{B}_i \cup \mathcal{D}_i} P_{Y^n}(\mathbf{y}) e^{-n\gamma} \quad (408)$$

$$\leq e^{-n\gamma} \sum_{i=1}^{M_n} P_{Y^n}(\mathcal{D}_i) \quad (409)$$

$$= e^{-n\gamma} \quad (410)$$

と変形できるので，式 (399) は

$$\Pr \{P_{X^n|Y^n}(X^n|Y^n) \leq e^{-n\gamma}\} \leq \varepsilon_n + e^{-n\gamma} \quad (411)$$

と評価できる．よって

$$\varepsilon_n \geq \Pr \{P_{X^n|Y^n}(X^n|Y^n) \leq e^{-n\gamma}\} - e^{-n\gamma} \quad (412)$$

が満たされる．□

これらを用いることにより，次の通信路符号化定理が導かれる．

定理 10 ([1] 定理 3.11) 通信路 $\mathbf{W}$ の ε -通信路容量は

$$C(\varepsilon|\mathbf{W}) = \sup_{\mathbf{X}} \underline{I}_{\varepsilon}(\mathbf{X}; \mathbf{Y}) \quad (413)$$

となる。ただしここで

$$\underline{I}_{\varepsilon}(\mathbf{X}; \mathbf{Y}) \triangleq \sup \left\{ \alpha \mid \limsup_{n \rightarrow \infty} \Pr \left\{ \frac{1}{n} \log \frac{P_{Y^n|X^n}(Y^n|X^n)}{P_{Y^n}(Y^n)} < \alpha \right\} \leq \varepsilon \right\} \quad (414)$$

である。

(証明)[1]

順定理

$R = \sup_{\mathbf{X}} \underline{I}_{\varepsilon}(\mathbf{X}; \mathbf{Y}) - 2\gamma$ が ε -達成可能であることを示す。初めに

$$R_0 \triangleq \sup_{\mathbf{X}} \underline{I}_{\varepsilon}(\mathbf{X}; \mathbf{Y}) \quad (415)$$

とおき任意の $\gamma > 0$ に対して $R = R_0 - 2\gamma$ となるような R が ε -達成可能であることを示す。まず R は

$$R + \gamma = R_0 - \gamma < \underline{I}_{\varepsilon}(\mathbf{X}; \mathbf{Y}) \quad (416)$$

を満たす。つまり

$$\limsup_{n \rightarrow \infty} \Pr \left\{ \frac{1}{n} \log \frac{P_{Y^n|X^n}(Y^n|X^n)}{P_{Y^n}(Y^n)} < R + \gamma \right\} \leq \varepsilon \quad (417)$$

となることが分かる。ここで $M_n = e^{nR}$ とおき、補題 3 を用いると

$$\varepsilon_n \leq \Pr \left\{ \log \frac{W^n(Y^n|X^n)}{P_{Y^n}(Y^n)} \leq R + n\gamma \right\} + e^{-n\gamma} \quad (418)$$

となり、両辺上限をとると

$$\limsup_{n \rightarrow \infty} \varepsilon_n \leq \varepsilon \quad (419)$$

が導かれる。そして $M_n = e^{nR}$ なので

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \log M_n \geq R \quad (420)$$

であることから、 R は ε -達成可能である。したがって

$$C(\varepsilon|\mathbf{W}) \geq \sup_{\mathbf{X}} \underline{I}_{\varepsilon}(\mathbf{X}; \mathbf{Y}) - 2\gamma \quad (421)$$

となり、 $\gamma > 0$ は任意だったので

$$C(\varepsilon|\mathbf{W}) \geq \sup_{\mathbf{X}} \underline{I}_{\varepsilon}(\mathbf{X}; \mathbf{Y}) \quad (422)$$

が満たされ，順定理が導けた．

逆定理

R が達成可能であるとする．つまり

$$\limsup_{n \rightarrow \infty} \varepsilon_n < \varepsilon \quad (423)$$

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \log M_n \geq R \quad (424)$$

となる (n, M_n, ε_n) 符号が存在する．ここで補題 4 より任意の $\gamma > 0$ において

$$\varepsilon_n \geq \Pr \left\{ \frac{1}{n} \log \frac{W^n(Y^n|X^n)}{P_{Y^n}(Y^n)} \leq \frac{1}{n} \log M_n - \gamma \right\} - e^{-n\gamma} \quad (425)$$

が成り立つ．また式 (424) は十分大きなすべての n で

$$\frac{1}{n} \log M_n \geq R - \gamma \quad (426)$$

である．式 (425) に式 (426) を代入して

$$\varepsilon_n \geq \Pr \left\{ \frac{1}{n} \log \frac{W^n(Y^n|X^n)}{P_{Y^n}(Y^n)} \leq R - 2\gamma \right\} - e^{-n\gamma} \quad (427)$$

となる．両辺上極限をとると式 (423) より

$$\varepsilon \geq \limsup_{n \rightarrow \infty} \varepsilon_n \geq \limsup_{n \rightarrow \infty} \Pr \left\{ \frac{1}{n} \log \frac{W^n(Y^n|X^n)}{P_{Y^n}(Y^n)} \leq R - 2\gamma \right\} - e^{-n\gamma} \quad (428)$$

が導かれ

$$\underline{I}_\varepsilon(\mathbf{X}; \mathbf{Y}) \leq R - 2\gamma \quad (429)$$

が得られる． $\gamma > 0$ は任意だったので

$$\underline{I}_\varepsilon(\mathbf{X}; \mathbf{Y}) \leq R \quad (430)$$

であるので

$$\sup_{\mathbf{X}} \underline{I}_\varepsilon(\mathbf{X}; \mathbf{Y}) \leq R \quad (431)$$

が成り立つ．したがって ε -通信路容量は

$$C(\varepsilon|\mathbf{W}) \leq \sup_{\mathbf{X}} \underline{I}_\varepsilon(\mathbf{X}; \mathbf{Y}) \quad (432)$$

が得られ，逆定理が示せた．

よって順定理，逆定理より

$$C(\varepsilon|\mathbf{W}) = \sup_{\mathbf{X}} \underline{I}_\varepsilon(\mathbf{X}; \mathbf{Y}) \quad (433)$$

となる．□

付録 C 不等式の証明

付録 A 中の証明で用いた不等式の証明を示す.

C.1 $(1-x)^y \leq e^{-xy}$ の証明

$0 \leq x \leq 1, y \geq 0$ のとき

$$(1-x)^y \leq (e^{-x})^y \quad (434)$$

とみることにより, 不等式は

$$1-x \leq e^{-x} \quad (435)$$

と変形できるので, この式を示す. ここで

$$f(x) = e^{-x} - (1-x) \quad (436)$$

とおくと, 条件の範囲において

$$f'(x) = -e^{-x} + 1 > 0 \quad (437)$$

となるので単調増加といえる. $f(0) = 0$ より

$$f(x) \geq 0 \quad (438)$$

となるので式 (435) が示せた. したがって $0 \leq x \leq 1, y \geq 0$ のとき

$$(1-x)^y \leq e^{-xy} \quad (439)$$

となる. $\square$

C.2 $e^{-xy} \leq 1 - y + e^{-x}$ の証明

$x \geq 0, 0 \leq y \leq 1$ のときを考える. はじめに

$$g(x, y) = e^{-xy} \quad (440)$$

とおく. すると

$$\frac{\partial^2}{\partial y^2} g(x, y) = x^2 e^{-xy} \geq 0 \quad (441)$$

なので下に凸の曲線となる．さらに

$$g(x, 0) = 1, \quad g(x, 1) = e^{-x} \quad (442)$$

であるので，下に凸の定義より 2 点を通る直線は

$$g(x, y) \leq (e^{-x} - 1)y + 1 \quad (443)$$

$$= 1 - y - ye^{-x} \quad (444)$$

となる． $0 \leq y \leq 1$ なので $ye^{-x} \leq e^{-x}$ であることから

$$g(x, y) \leq 1 - y - e^{-x} \quad (445)$$

が導かれ

$$e^{-xy} \leq 1 - y - e^{-x} \quad (446)$$

が得られる．□