

信州大学工学部

学士論文

最大歪み測度を用いた
忠実度規範つき情報源符号化に関する研究

指導教員 西新 幹彦 准教授

学科 電気電子工学科
学籍番号 07T2071K
氏名 原 寛貴

2011 年 3 月 31 日

目次

1	序章	1
1.1	研究の背景と目的	1
1.2	構成	1
2	概要	1
2.1	情報源	1
2.2	符号	1
2.3	歪み測度	2
2.4	忠実度規範	2
3	レート・歪み理論	4
3.1	平均歪み測度に対してレート R が平均忠実度規範 D で達成可能	4
3.2	平均歪み測度に対してレート R が CK 最大忠実度規範 D で達成可能	13
3.3	最大歪み測度に対してレート R が CK 最大忠実度規範 D で達成可能	13
4	まとめ	15
	謝辞	15
	参考文献	15

1 序章

1.1 研究の背景と目的

現代の情報社会において扱うデータ量は膨大なものとなってきた。これをそのまま伝送、または圧縮することは非常に効率が悪い。そこで、情報を効率よく符号化し通信時間、通信容量を抑える情報圧縮技術は必要不可欠なものである。また、音声や画像などの本来連続的な情報は、デジタル回線において歪みなく伝送することは不可能である。元の情報源とこれを符号化・復号化したものとの間に一定の歪みを許容し、その上でどれだけ圧縮できるかを考えることが現実的である。本研究では、歪み測度が平均歪みであるか最大歪みであるかの2種類を用意する。さらに予め与えられた歪み測度に対して、受信者が受け取った情報源が元の情報源にどれだけ忠実であるかという忠実度規範を4種類用意する。そして、許容する歪みの範囲内で忠実度規範を達成する最小符号化レートを求めることを目的とする。本論文の主な内容は [2] の定理 5.1 に対してより詳細な証明を与えたことと、[1] の Problem 2.2.12 に対して証明を与えたことである。

1.2 構成

本論文の構成は2章で本研究の概要を述べ、3章では歪み測度に対してレート R が達成可能な最小符号化レートを与え、4章でまとめとしている。

2 概要

2.1 情報源

情報源とは、情報源アルファベットと呼ばれる有限集合 $\mathcal{X}$ から値をとる確率変数列 $\{X_i\}_{i=1}^{\infty}$ である。本研究では定常無記憶情報源

$$X^n = (X_1, X_2, \dots, X_n) \quad (1)$$

のレート・歪み理論について考える。

2.2 符号

情報源 X^n の符号化を以下のように行う。まず、 X^n に対する符号を

$$\mathcal{M}_n = \{1, 2, \dots, M_n\}$$

として符号化を $\varphi_n : \mathcal{X}^n \rightarrow \mathcal{M}_n$ によって定める. $\varphi_n(x)$ は $x \in \mathcal{X}^n$ の符号語と呼ぶ. 復元アルファベットとして有限集合 $\mathcal{Y}$ を用意し, 復号化を $\psi_n : \mathcal{M}_n \rightarrow \mathcal{Y}^n$ によって定める. ここで, 符号化レートを $r_n \triangleq \frac{1}{n} \log |\mathcal{M}_n|$ と定める. 情報源からの出力 $x \in \mathcal{X}^n$ は符号器 φ_n により符号化された後, 復号器 ψ_n により $y \in \mathcal{Y}^n$ と変換される. 本研究では x と y の間の情報源文字 1 個当りの歪み $d_n(x, y)$ を一定の大きさ以下に抑えた上で符号化レート r_n をできるだけ小さくしたい.

2.3 歪み測度

歪み測度と呼ばれる関数 $d : \mathcal{X} \times \mathcal{Y} \rightarrow [0, +\infty)$ を定め, $d(x, y) (x \in \mathcal{X}, y \in \mathcal{Y})$ を x と y の間の歪みと呼ぶ. 長さ n のデータ $\mathbf{x} = (x_1, x_2, \dots, x_n) \in \mathcal{X}^n$ と $\mathbf{y} = (y_1, y_2, \dots, y_n) \in \mathcal{Y}^n$ の間の歪み $d_n(\mathbf{x}, \mathbf{y})$ を平均歪みと最大歪みの 2 種類を定義する. ここで, d_n は系列の歪みであり, d は 1 シンボルの歪みである.

$$d_n(\mathbf{x}, \mathbf{y}) \triangleq \begin{cases} \frac{1}{n} \sum_{i=1}^n d(x_i, y_i) & (\text{平均歪み測度}) \\ \max_{1 \leq i \leq n} d(x_i, y_i) & (\text{最大歪み測度}) \end{cases} \quad (2)$$

と定義する.

2.4 忠実度規範

符号器 φ_n によって, 出力系列 $X^n \in \mathcal{X}^n$ を符号化し, 復号器 ψ_n によって復号化して得られる系列を $Y^n \in \mathcal{Y}^n$ とする. このとき, 系列の歪み $d_n(X^n, Y^n)$ が小さい程, 符号器と復号器の列 (φ_n, ψ_n) は情報源に忠実であるといえる. そこで, 任意の小さな数 $\delta > 0$ に対して, 情報源からの出力系列 X^n が D より大きな歪みを伴って復元される確率が δ 以下であることを意味する規範を最大忠実度規範と呼ぶ. また, 系列の歪みの平均が D 以下であることを意味する規範を平均忠実度規範と呼ぶ. 次に, この忠実度規範を 4 種類定義する.

定義 1 最大忠実度規範 D で達成可能とは

$$\text{p-}\limsup_{n \rightarrow \infty} d_n(X^n, Y^n) \leq D \quad (3)$$

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log |\mathcal{M}_n| \leq R \quad (4)$$

が成り立つことである. ただし, 確率変数列 $\{Z_n\}_n$ に対して

$$\text{p-}\limsup_{n \rightarrow \infty} Z_n \triangleq \inf\{\alpha \mid \lim_{n \rightarrow \infty} \Pr\{Z_n > \alpha\} = 0\}$$

と定める. これにより, (3), (4) は次のように言い換えられる.

任意の $\delta > 0$ に対して, 十分大きなすべての n で

$$\begin{aligned}\Pr\{d_n(X^n, Y^n) > D + \delta\} &< \delta \\ \frac{1}{n} \log |\mathcal{M}_n| &< R + \delta\end{aligned}$$

を満たす符号器と復号器の列 (φ_n, ψ_n) が存在する.

定義 2 平均忠実度規範 D で達成可能とは

$$\begin{aligned}\limsup_{n \rightarrow \infty} E[d_n(X^n, Y^n)] &\leq D \\ \limsup_{n \rightarrow \infty} \frac{1}{n} \log |\mathcal{M}_n| &\leq R\end{aligned}$$

が成り立つことである. これは次のように言い換えられる.

任意の $\delta > 0$ に対して, 十分大きなすべての n で

$$\begin{aligned}E[d_n(X^n, Y^n)] &< D + \delta \\ \frac{1}{n} \log |\mathcal{M}_n| &< R + \delta\end{aligned}$$

を満たす符号器と復号器の列 (φ_n, ψ_n) が存在する.

定義 3 CK 最大忠実度規範 D で達成可能とは, 任意の $\delta > 0$ に対して, 十分大きなすべての n で

$$\begin{aligned}\Pr\{d_n(X^n, Y^n) > D\} &< \delta \\ \frac{1}{n} \log |\mathcal{M}_n| &< R + \delta\end{aligned}$$

を満たす符号器と復号器の列 (φ_n, ψ_n) が存在することである.

定義 4 CK 平均忠実度規範 D で達成可能とは, 任意の $\delta > 0$ に対して, 十分大きなすべての n で

$$\begin{aligned}E[d_n(X^n, Y^n)] &\leq D \\ \frac{1}{n} \log |\mathcal{M}_n| &< R + \delta\end{aligned}$$

を満たす符号器と復号器の列 (φ_n, ψ_n) が存在することである.

3 レート・歪み理論

3.1 平均歪み測度に対してレート R が平均忠実度規範 D で達成可能

定義 5

$R_1(D|X) \triangleq \inf\{R \mid \text{平均歪み測度に対して } R \text{ は平均忠実度規範 } D \text{ で達成可能}\}$

定理 1 定常無記憶情報源 X のレート・歪み関数 $R_1(D|X)$ は

$$R_1(D|X) = \min_{Y: Ed(X,Y) \leq D} I(X;Y) \quad (5)$$

で得られる. また, 右辺の $\min$ は条件 $Ed(X,Y) \leq D$ を満たす集合 $\mathcal{Y}$ の中に値をとるすべての確率変数 Y にわたる最小値を示す. このような確率変数全体は有界閉集合になるので最小値が存在することが保証される.

補題 1 式 (5) の右辺は D の関数として下に凸である. 以下にそれを示す. D の関数を

$$f(D) \triangleq \min_{Y: Ed(X,Y) \leq D} I(X;Y)$$

と定義する. Y_1, Y_2 をそれぞれ $Ed(X, Y_1) \leq D_1, Ed(X, Y_2) \leq D_2$ を満たす任意の確率変数とし, $\alpha_1 \geq 0, \alpha_2 \geq 0$ を $\alpha_1 + \alpha_2 = 1$ を満たす任意の定数とする. まず, X とは独立な確率変数 Q を $P_Q(1) = \alpha_1, P_Q(2) = \alpha_2$ と定め, 確率変数 Y を $Q = i (i = 1, 2)$ のとき $Y = Y_i$ と定めれば,

$$Ed(X, Y) = \alpha_1 Ed(X, Y_1) + \alpha_2 Ed(X, Y_2) \leq \alpha_1 D_1 + \alpha_2 D_2 \quad (6)$$

が成り立つ. また, Q と X は独立ということに注意すると, $I(X; Q) = 0$ であるため,

$$\begin{aligned} \alpha_1 I(X; Y_1) + \alpha_2 I(X; Y_2) &= I(X; Y|Q) \\ &= I(X; QY) \\ &\geq I(X; Y) \end{aligned} \quad (7)$$

を得る. (6) と (7) を合わせれば

$$\begin{aligned} \alpha_1 f(D_1) + \alpha_2 f(D_2) &= \alpha_1 \min_{Y_1: Ed(X, Y_1) \leq D_1} I(X; Y_1) + \alpha_2 \min_{Y_2: Ed(X, Y_2) \leq D_2} I(X; Y_2) \\ &= \min_{Y_1: Ed(X, Y_1) \leq D_1} \alpha_1 I(X; Y_1) + \min_{Y_2: Ed(X, Y_2) \leq D_2} \alpha_2 I(X; Y_2) \\ &= \min_{\substack{Y_1: Ed(X, Y_1) \leq D_1 \\ Y_2: Ed(X, Y_2) \leq D_2}} (\alpha_1 I(X; Y_1) + \alpha_2 I(X; Y_2)) \\ &\geq \min_{Y: Ed(X, Y) \leq \alpha_1 D_1 + \alpha_2 D_2} I(X, Y) \\ &= f(\alpha_1 D_1 + \alpha_2 D_2) \end{aligned}$$

となり, $f(D)$ が D の関数として下に凸であることがわかる. ゆえに, $f(D)$ は D に関して連続である.

(定理 1 の証明) 証明は [2] に負うが, より詳細に記述する.

1) Direct part:

初めに, 条件 $Ed(X, Y) \leq D - \gamma$ を満たす任意の Y を与えたときに, $R \triangleq I(X; Y) + 2\gamma$ が平均忠実度規範 D で達成可能であることを示す. ここで, $\gamma > 0$ は任意の小さい定数である. $X^n Y^n \triangleq (X_1 Y_1, X_2 Y_2, \dots, X_n Y_n)$ を XY の定常独立系列として $X^n \triangleq (X_1, X_2, \dots, X_n)$, $Y^n \triangleq (Y_1, Y_2, \dots, Y_n)$ とおく. そこで,

$$T_n^{(1)} \triangleq \left\{ (\mathbf{x}, \mathbf{y}) \in \mathcal{X}^n \times \mathcal{Y}^n \mid \left| \frac{1}{n} \log \frac{P_{Y^n|X^n}(\mathbf{y}|\mathbf{x})}{P_{Y^n}(\mathbf{y})} - I(X; Y) \right| < \gamma \right\} \quad (8)$$

$$T_n^{(2)} \triangleq \left\{ (\mathbf{x}, \mathbf{y}) \in \mathcal{X}^n \times \mathcal{Y}^n \mid d_n(\mathbf{x}, \mathbf{y}) - Ed(X, Y) < \gamma \right\} \quad (9)$$

として $T_n \triangleq T_n^{(1)} \cap T_n^{(2)}$ とおく. そして, $Ed(X, Y) \leq D - \gamma$ と (9) より

$$(\mathbf{x}, \mathbf{y}) \in T_n \text{ ならば } d_n(\mathbf{x}, \mathbf{y}) \leq D \quad (10)$$

であることに気をつける. また, $X^n Y^n = (X_1 Y_1, X_2 Y_2, \dots, X_n Y_n)$ が定常独立系列なので

$$\frac{1}{n} \log \frac{P_{Y^n|X^n}(Y^n|X^n)}{P_{Y^n}(Y^n)} = \frac{1}{n} \sum_{i=1}^n \log \frac{P_{Y|X}(Y_i|X_i)}{P_Y(Y_i)} \quad (11)$$

$$d_n(X^n, Y^n) = \frac{1}{n} \sum_{i=1}^n d(X_i, Y_i) \quad (12)$$

が成立する. そして, (11), (12) 右辺の和の中の各項は独立でその期待値は

$$\begin{aligned} \mathbf{E} \left[\log \frac{P_{Y|X}(Y_i|X_i)}{P_Y(Y_i)} \right] &= I(X; Y) \quad (i = 1, 2, \dots) \\ Ed(X_i, Y_i) &= Ed(X, Y) \quad (i = 1, 2, \dots) \end{aligned}$$

であり, その分散は有界である. 以下に分散が有界である事を示す. $Z \triangleq \frac{1}{n} \sum_{i=1}^n Z_i$, $Z_i \triangleq \log \frac{P_{Y|X}(Y_i|X_i)}{P_Y(Y_i)}$ とおく. Z の分散 $V(Z)$ は,

$$\begin{aligned} V(Z) &= \mathbf{E}(Z^2) - (\mathbf{E}(Z))^2 \\ &= \mathbf{E} \left\{ \left(\frac{1}{n} \sum_{i=1}^n Z_i \right)^2 \right\} - \left\{ \mathbf{E} \left(\sum_{i=1}^n Z_i \right) \right\}^2 \\ &= \mathbf{E} \left\{ \left(\frac{1}{n} \sum_{i=1}^n Z_i \right)^2 \right\} - \left\{ \frac{1}{n} \sum_{i=1}^n (\mathbf{E}(Z_i)) \right\}^2 \\ &= \mathbf{E} \left\{ \left(\frac{1}{n} \sum_{i=1}^n Z_i \right)^2 \right\} - (\mathbf{E}[Z_1])^2 \end{aligned} \quad (13)$$

ここで, (13) 右辺第 1 項は

$$\begin{aligned}
\mathbf{E}\left\{\left(\frac{1}{n}\sum_{i=1}^n Z_i\right)^2\right\} &= \frac{1}{n^2}\mathbf{E}\left\{\left(\sum_{i=1}^n Z_i\right)^2\right\} \\
&= \frac{1}{n^2}\mathbf{E}\{(Z_1 + Z_2 + \cdots + Z_n)^2\} \\
&= \frac{1}{n^2}\{n\mathbf{E}(Z^2) + (n^2 - n)\mathbf{E}(Z_1)^2\} \\
&= \frac{1}{n}\{\mathbf{E}(Z_1^2) + (n-1)\mathbf{E}(Z_1)^2\}
\end{aligned}$$

ゆえに, 分散 $V(Z)$ は

$$\begin{aligned}
V(Z) &= \frac{1}{n}\{\mathbf{E}(Z_1^2) + (n-1)\mathbf{E}(Z_1)^2\} - \{\mathbf{E}(Z_1)\}^2 \\
&= \frac{1}{n}[\mathbf{E}(Z_1^2) + (n-1)\mathbf{E}(Z_1)^2 - n\{\mathbf{E}(Z_1)\}^2] \\
&= \frac{1}{n}[\mathbf{E}(Z_1) - \{\mathbf{E}(Z_1)\}^2] \\
&= \frac{1}{n}V(Z_1)
\end{aligned}$$

となり, 分散は有界である. よって, $T_n^{(1)}, T_n^{(2)}$ の定義 (8), (9) に注意して Chebyshev の不等式を適用すれば, $n \rightarrow \infty$ によって,

$$\begin{aligned}
\Pr\{X^n Y^n \in T_n^{(1)}\} &= \Pr\left\{\left|\frac{1}{n}\log \frac{P_{Y^n|X^n}(Y^n|X^n)}{P_{Y^n}(Y^n)} - I(X; Y)\right| < \gamma\right\} \\
&= \Pr\left\{\left|\frac{1}{n}\sum_{i=1}^n \log \frac{P_{Y|X}(Y_i|X_i)}{P_Y(Y_i)} - I(X; Y)\right| < \gamma\right\} \\
&\geq 1 - \frac{\frac{1}{n}\sigma^2}{\gamma^2} \\
&\rightarrow 1
\end{aligned}$$

$$\begin{aligned}
\Pr\{X^n Y^n \in T_n^{(2)}\} &= \Pr\left\{\left|d_n(X^n, Y^n) - \mathbf{E}d(X, Y)\right| < \gamma\right\} \\
&= \Pr\left\{\left|\frac{1}{n}\sum_{i=1}^n d(X_i, Y_i) - \mathbf{E}d(Y, Y)\right| < \gamma\right\} \\
&\geq 1 - \frac{\frac{1}{n}\sigma^2}{\gamma^2} \\
&\rightarrow 1
\end{aligned}$$

を得る. よって,

$$\begin{aligned}\Pr\{X^n Y^n \notin T_n\} &= \Pr\{X^n Y^n \notin T_n^{(1)} \cap T_n^{(2)}\} \\ &\leq \Pr\{X^n Y^n \notin T_n^{(1)}\} + \Pr\{X^n Y^n \notin T_n^{(2)}\} \\ &\rightarrow 0\end{aligned}$$

であることから,

$$\Pr\{X^n Y^n \in T_n\} \rightarrow 1 \quad (n \rightarrow \infty) \quad (14)$$

が成立する.

さてここで,

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

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} に従って, 発生させる. さらに, これらを符号語として $\mathcal{C} \triangleq \{v_1, v_2, \dots, v_{M_n}\}$ とおく (ランダム符号化). 次に, 各 $x \in \mathcal{X}^n$ に対して

$$d_n(x, v_{i_0}) = \min_{1 \leq i \leq M_n} d_n(x, v_i) \quad (15)$$

となる番号 i_0 を用いて符号器 φ_n を $\varphi_n(x) = i_0$ と定義する. すなわち

$$\varphi(x) \triangleq \arg \min_{1 \leq i \leq M_n} d_n(x, v_i)$$

である. また, 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
に定義した符号器と復号器の組 (φ_n, ψ_n) に対して

$$\overline{P}_e^{(n)} \triangleq \Pr\left\{d_n(X^n, \psi_n(\varphi_n(X^n))) > D\right\} \quad (16)$$

とおく. ここで, $A \triangleq \{d_n(X^n, \psi_n(\varphi_n(X^n))) > D\}$, $\overline{A} \triangleq \{d_n(X^n, \psi_n(\varphi_n(X^n))) \leq D\}$ とすれば,

$$\begin{aligned}\overline{\rho}_n &\triangleq \mathbf{E}_{\mathcal{C}} \mathbf{E} d_n(X_n, \psi_n(\varphi_n(X^n))) \\ &= \Pr\{A\} \mathbf{E}_{\mathcal{C}} \mathbf{E} \left[d_n(X_n, \psi_n(\varphi_n(X^n))) \middle| A \right] + \Pr\{\overline{A}\} \mathbf{E}_{\mathcal{C}} \mathbf{E} \left[d_n(X_n, \psi_n(\varphi_n(X^n))) \middle| \overline{A} \right] \\ &\leq \overline{P}_e^{(n)} d_{max} + D\end{aligned} \quad (17)$$

を得る. ただし, $d_{max} \triangleq \max_{x \in \mathcal{X}, y \in \mathcal{Y}} d(x, y)$ であり, $\mathbf{E}$ は X^n に関する期待値を示し, $\mathbf{E}_{\mathcal{C}}$ はランダム符号 $\mathcal{C}$ に関する期待値を表す.

ここで、 $\bar{P}_e^{(n)}$ の値を評価する。そのために、 $v_1, v_2, \dots, v_{M_n}$ は互いに独立で同一の分布 P_{Y^n} に従うことを考慮すると、符号 (φ_n, ψ_n) の定義と (16) から

$$\begin{aligned}
\bar{P}_e^{(n)} &= \sum_{\mathbf{x} \in \mathcal{X}^n} \Pr\{X^n = \mathbf{x} \text{ かつ } d_n(\mathbf{x}, \psi_n(\varphi_n(\mathbf{x}))) > D\} \\
&= \sum_{\mathbf{x} \in \mathcal{X}^n} P_{X^n}(\mathbf{x}) \Pr\left\{d_n(\mathbf{x}, \psi_n(\varphi_n(\mathbf{x}))) > D\right\} \\
&= \sum_{\mathbf{x} \in \mathcal{X}^n} P_{X^n}(\mathbf{x}) \Pr\left\{\text{どの } v_i (i = 1, 2, \dots, M_n) \text{ に対しても } d_n(\mathbf{x}, v_i) > D\right\} \\
&= \sum_{\mathbf{x} \in \mathcal{X}^n} P_{X^n}(\mathbf{x}) \prod_{i=1}^{M_n} \Pr\left\{d_n(\mathbf{x}, v_i) > D\right\} \\
&= \sum_{\mathbf{x} \in \mathcal{X}^n} P_{X^n}(\mathbf{x}) \left(\Pr\left\{d_n(\mathbf{x}, Y^n) > D\right\}\right)^{M_n} \\
&= \sum_{\mathbf{x} \in \mathcal{X}^n} P_{X^n}(\mathbf{x}) \left(1 - \Pr\left\{d_n(\mathbf{x}, Y^n) \leq D\right\}\right)^{M_n} \\
&= \sum_{\mathbf{x} \in \mathcal{X}^n} P_{X^n}(\mathbf{x}) \left(1 - \sum_{\mathbf{y} \in \mathcal{Y}^n} P_{Y^n}(\mathbf{y}) \mathbf{1}\left\{d_n(\mathbf{x}, \mathbf{y}) \leq D\right\}\right)^{M_n} \tag{18}
\end{aligned}$$

が導かれる。ここで、

$$J(\mathbf{x}, \mathbf{y}) \triangleq \begin{cases} 1 & \text{for } (\mathbf{x}, \mathbf{y}) \in T_n \\ 0 & \text{otherwise} \end{cases} \tag{19}$$

とおくと、(10) より

$$J(\mathbf{x}, \mathbf{y}) \leq \mathbf{1}\{d_n(\mathbf{x}, \mathbf{y}) \leq D\}$$

が成立する。ゆえに、(18) の $\bar{P}_e^{(n)}$ は

$$\bar{P}_e^{(n)} \leq \sum_{\mathbf{x} \in \mathcal{X}^n} P_{X^n}(\mathbf{x}) \left(1 - \sum_{\mathbf{y} \in \mathcal{Y}^n} P_{Y^n}(\mathbf{y}) J(\mathbf{x}, \mathbf{y})\right)^{M_n} \tag{20}$$

と上から抑えることができる。しかし、(8) より $(\mathbf{x}, \mathbf{y}) \in T_n$ ならば

$$\begin{aligned}
\frac{1}{n} \log \frac{P_{Y^n|X^n}(\mathbf{y}|\mathbf{x})}{P_{Y^n}(\mathbf{y})} - I(X; Y) &< \gamma \\
\log \frac{P_{Y^n|X^n}(\mathbf{y}|\mathbf{x})}{P_{Y^n}(\mathbf{y})} &< n\{I(X; Y) + \gamma\} \\
\frac{P_{Y^n|X^n}(\mathbf{y}|\mathbf{x})}{P_{Y^n}(\mathbf{y})} &< e^{n(I(X; Y) + \gamma)} \\
P_{Y^n}(\mathbf{y}) &> e^{-n(I(X; Y) + \gamma)} P_{Y^n|X^n}(\mathbf{y}|\mathbf{x})
\end{aligned}$$

なので, 任意の $(x, y) \in \mathcal{X}^n \times \mathcal{Y}^n$ に対して

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

が成り立つ. これを (20) の右辺に代入すれば

$$\overline{P}_e^{(n)} \leq \sum_{\mathbf{x} \in \mathcal{X}^n} P_{X^n}(\mathbf{x}) \left(1 - e^{-n(I(X;Y)+\gamma)} \sum_{\mathbf{y} \in \mathcal{Y}^n} P_{Y^n|X^n}(\mathbf{y}|\mathbf{x})J(\mathbf{x}, \mathbf{y}) \right)^{M_n} \quad (21)$$

を得る. ここで, 不等式

$$(1-x)^y \leq e^{-xy} \quad (0 \leq x \leq 1, y \geq 0) \quad (22)$$

を用いて $y = M_n$, $x = e^{-n(I(X;Y)+\gamma)} \sum_{\mathbf{y} \in \mathcal{Y}^n} P_{Y^n|X^n}(\mathbf{y}|\mathbf{x})J(\mathbf{x}, \mathbf{y})$ と代入すれば

$$\begin{aligned} \overline{P}_e^{(n)} &\leq \sum_{\mathbf{x} \in \mathcal{X}^n} P_{X^n}(\mathbf{x}) \exp \left\{ -M_n e^{-n(I(X;Y)+\gamma)} \sum_{\mathbf{y} \in \mathcal{Y}^n} P_{Y^n|X^n}(\mathbf{y}|\mathbf{x})J(\mathbf{x}, \mathbf{y}) \right\} \\ &= \sum_{\mathbf{x} \in \mathcal{X}^n} P_{X^n}(\mathbf{x}) \exp \left\{ -e^{n\gamma} \sum_{\mathbf{y} \in \mathcal{Y}^n} P_{Y^n|X^n}(\mathbf{y}|\mathbf{x})J(\mathbf{x}, \mathbf{y}) \right\} \end{aligned}$$

となる. さらに, 不等式

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

に対して $y = \sum_{\mathbf{y} \in \mathcal{Y}^n} P_{Y^n|X^n}(\mathbf{y}|\mathbf{x})J(\mathbf{x}, \mathbf{y})$, $x = e^{n\gamma}$ を用いると

$$\begin{aligned} \overline{P}_e^{(n)} &\leq \sum_{\mathbf{x} \in \mathcal{X}^n} P_{X^n}(\mathbf{x}) \left\{ 1 - \sum_{\mathbf{y} \in \mathcal{Y}^n} P_{Y^n|X^n}(\mathbf{y}|\mathbf{x})J(\mathbf{x}, \mathbf{y}) + e^{-e^{n\gamma}} \right\} \\ &= 1 - \sum_{\mathbf{x} \in \mathcal{X}^n} P_{X^n}(\mathbf{x}) \sum_{\mathbf{y} \in \mathcal{Y}^n} P_{Y^n|X^n}(\mathbf{y}|\mathbf{x})J(\mathbf{x}, \mathbf{y}) + e^{-e^{n\gamma}} \\ &= 1 - \sum_{(\mathbf{x}, \mathbf{y}) \in \mathcal{X}^n \mathcal{Y}^n} P_{X^n Y^n}(\mathbf{x}, \mathbf{y})J(\mathbf{x}, \mathbf{y}) + e^{-e^{n\gamma}} \\ &= 1 - \sum_{(\mathbf{x}, \mathbf{y}) \in T_n} P_{X^n Y^n}(\mathbf{x}, \mathbf{y}) + e^{-e^{n\gamma}} \\ &= 1 - P(X^n Y^n \in T_n) + e^{-e^{n\gamma}} \\ &= P(X^n Y^n \notin T_n) + e^{-e^{n\gamma}} \\ &\rightarrow 0 \quad (n \rightarrow \infty) \end{aligned} \quad (23)$$

となる. 一方, (17) により少なくとも1つのランダムでない符号器と復号器の組 (φ_n, ψ_n) が存在し

$$\rho_n \triangleq \mathbf{E} d_n(X_n, \psi_n(\varphi_n(X^n))) \leq D + \overline{P}_e^{(n)} d_{max}$$

を満たさなければならない。よって (φ_n, ψ_n) の列を適切に選ぶことによって

$$\begin{aligned} \limsup_{n \rightarrow \infty} \rho_n &= \limsup_{n \rightarrow \infty} \mathbf{Ed}_n(X_n, \psi_n(\varphi_n(X^n))) \\ &\leq D \end{aligned}$$

とできる。ところが, $M_n = e^{n(I(X;Y)+2\gamma)}$ なので,

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

が成立している。よって, $R = I(X;Y) + 2\gamma$ は平均忠実度規範 D で達成可能である。さらに, Y は条件 $\mathbf{Ed}(X, Y) \leq D - \gamma$ を満たす範囲で任意にとつたので,

$$R_1(D|X) \leq \min_{Y: \mathbf{Ed}(X, Y) \leq D - \gamma} I(X;Y) + 2\gamma$$

である。ここで, 補題 1 によりこの不等式の右辺は γ に関して連続である。 $\gamma > 0$ は任意だったため $\gamma \rightarrow 0$ とすれば

$$R_1(D|X) \leq \min_{Y: \mathbf{Ed}(X, Y) \leq D} I(X;Y)$$

を得る。

2) Converse part:

R が平均忠実度規範 D で達成可能であるとすれば,

$$\limsup_{n \rightarrow \infty} \mathbf{Ed}_n(X^n, \psi_n(\varphi_n(X^n))) \leq D \quad (24)$$

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log M_n \leq R \quad (25)$$

なる符号器と復号器の組 (φ_n, ψ_n) が存在する。そこで,

$$Y^n \triangleq (Y_1^{(n)}, Y_2^{(n)}, \dots, Y_n^{(n)}) = \psi_n(\varphi_n(X^n))$$

とおけば, 符号化関数 φ_n のとり得る値の数は高々 M_n , よって Y^n のとり得る値の数も高々 M_n なので

$$\begin{aligned} \log M_n &\geq H(Y^n) \\ &\geq I(X^n; Y^n) \\ &= H(X^n) - H(X^n | Y^n) \\ &= \sum_{i=1}^n H(X_i) - \sum_{i=1}^n H(X_i | Y^n X^{i-1}) \end{aligned}$$

となる。ただし、最後の等式で情報源の無記憶性とエントロピーのチェイン則を用いて、 $X^{i-1} = (X_1, X_2, \dots, X_{i-1})$ とおいた。そこで、

$$H(X_i | Y^n X^{i-1}) \leq H(X_i | Y_i^{(n)})$$

に注意すると

$$\begin{aligned} \log M_n &\geq \sum_{i=1}^n H(X_i) - \sum_{i=1}^n H(X_i | Y_i^{(n)}) \\ &= \sum_{i=1}^n I(X_i; Y_i^{(n)}) \end{aligned}$$

を得る。ゆえに

$$\frac{1}{n} \log M_n \geq \frac{1}{n} \sum_{i=1}^n I(X_i; Y_i^{(n)}) \quad (26)$$

となる。そして、確率変数 $Q_n \in \{1, 2, \dots, n\}$ の分布を $P_{Q_n}(i) = 1/n (i = 1, 2, \dots, n)$ とし、確率変数 $XY^{(n)}$ を $Q_n = i$ のとき $XY^{(n)} \triangleq X_i Y_i^{(n)}$ と定義すると (25) は

$$\frac{1}{n} \log M_n \geq I(X; Y^{(n)} | Q_n) \quad (27)$$

とかける。しかし、情報源の定常性より

$$\begin{aligned} P_{Q_n X}(i, x) &= P_{Q_n}(i) \cdot P_{X|Q_n}(x|i) \\ &= P_{Q_n}(i) \cdot P_{X_i}(x) \\ &= P_{Q_n}(i) \cdot P_X(x) \end{aligned}$$

となり、 Q_n と X は独立となる。よって、

$$I(X; Y^{(n)} | Q_n) = I(X; Q_n Y^{(n)}) \geq I(X; Y^{(n)})$$

となる。したがって、(27) は

$$\frac{1}{n} \log M_n \geq I(X; Y^{(n)}) \quad (28)$$

となる。一方,

$$\begin{aligned}
Ed_n(X^n, \psi_n(\varphi_n(X^n))) &= Ed_n(X^n, Y^n) \\
&= E \left\{ \frac{1}{n} \sum_{i=1}^n d(X_i, Y_i^{(n)}) \right\} \\
&= \sum_{i=1}^n \frac{1}{n} E \{ d(X_i, Y_i) \} \\
&= \sum_{i=1}^n P_{Q_n} E \{ d(X, Y^{(n)}) | Q_n = i \} \\
&= Ed(X, Y^{(n)})
\end{aligned}$$

なので, (24) から

$$\limsup_{n \rightarrow \infty} Ed(X, Y^{(n)}) \leq D$$

ゆえに, 任意に小さい定数 $\gamma > 0$ に対して十分大きなすべての n で

$$Ed(X, Y^{(n)}) \leq D + \gamma$$

となる。そうすると, (28) より

$$\begin{aligned}
\frac{1}{n} \log M_n &\geq I(X; Y^{(n)}) \\
&\geq \min_{Y: Ed(X, Y) \leq D + \gamma} I(X; Y)
\end{aligned}$$

となる。よって, (25) から n が十分大きければ

$$R \geq \min_{Y: Ed(X, Y) \leq D + \gamma} I(X; Y) - \gamma$$

が得られる。しかし, 補題 1 により, この不等式の右辺は γ に関して連続である。そこで, $\gamma > 0$ は任意であったことに注意して $\gamma \rightarrow 0$ とすれば

$$R \geq \min_{Y: Ed(X, Y) \leq D} I(X; Y)$$

が結論される。 R は平均忠実度規範 D で達成可能な任意のレートと仮定されていたので,

$$R_1(D|X) \geq \min_{Y: Ed(X, Y) \leq D} I(X; Y)$$

が成り立つ。 $\square$

3.2 平均歪み測度に対してレート R が CK 最大忠実度規範 D で達成可能

定義 6

$$R_2(D|X) \triangleq \inf\{R \mid \text{平均歪み測度に対して } R \text{ は CK 最大忠実度規範 } D \text{ で達成可能}\}$$

定理 2 定常無記憶情報源 X のレート・歪み関数 $R_2(D|X)$ は,

$$R_2(D|X) = \min_{Y: E[d(X,Y)] \leq D} I(X; Y) \quad (29)$$

で与えられる. 説明は省略する.

3.3 最大歪み測度に対してレート R が CK 最大忠実度規範 D で達成可能

定義 7

$$R_3(D|X) \triangleq \inf\{R \mid \text{最大歪み測度に対して } R \text{ は CK 最大忠実度規範 } D \text{ で達成可能}\}$$

定理 3 定常無記憶情報源 X のレート・歪み関数 $R_3(D|X)$ は,

$$R_3(D|X) = \min_{Y: \Pr\{d(X,Y) \leq D\} = 1} I(X; Y) \quad (30)$$

で与えられる.

(証明) まず,

$$d'(x, y) \triangleq \begin{cases} 0 & \text{if } d(x, y) \leq D \\ 1 & \text{otherwise} \end{cases} \quad (31)$$

と定める. このとき,

$$E[d'(x, y)] = 0 \Leftrightarrow \Pr\{d(x, y) \leq D\} = 1$$

が成り立つ. よって,

$$\min_{Y: E[d'(x, y)] = 0} I(X; Y) = \min_{Y: \Pr\{d(x, y) \leq D\} = 1} I(X; Y)$$

となる. 左辺は「歪み測度 d' による平均歪みに対して符号化レート R は CK 最大忠実度規範 0 で達成可能」の意味でレートの下限である. つまり, 定理 2 より $R_2(0|X)$ に等しい. またこれは, 任意の $\delta > 0$ に対して十分大きなすべての n で次の条件を満たす符号 (φ_n, ψ_n) が存在することを意味する.

$$\begin{aligned} \frac{1}{n} \log \|\varphi_n\| &< R + \delta \\ \Pr\left\{\frac{1}{n} \sum_{i=1}^n d'(X_i, \hat{Y}_i^{(n)}) = 0\right\} &> 1 - \delta \end{aligned} \quad (32)$$

ただし $\hat{Y}^{(n)} \triangleq \hat{Y}_1^{(n)} \cdots \hat{Y}_n^{(n)} = \psi_n(\varphi_n(X^n))$ とおいた. ここで,

$$\begin{aligned} \frac{1}{n} \sum_{i=1}^n d'(X_i, \hat{Y}_i^{(n)}) = 0 &\Leftrightarrow \sum_{i=1}^n d'(X_i, \hat{Y}_i^{(n)}) = 0 \\ &\Leftrightarrow i = 1, \dots, n \text{ に対して } d'(X_i, \hat{Y}_i^{(n)}) = 0 \\ &\Leftrightarrow i = 1, \dots, n \text{ に対して } d(X_i, \hat{Y}_i^{(n)}) \leq D \\ &\Leftrightarrow \max_{1 \leq i \leq n} d(X_i, \hat{Y}_i^{(n)}) \leq D \end{aligned}$$

に注意すると,

$$\Pr \left\{ \frac{1}{n} \sum_{i=1}^n d'(X_i, \hat{Y}_i^{(n)}) = 0 \right\} = \Pr \left\{ \max_{1 \leq i \leq n} d(X_i, \hat{Y}_i^{(n)}) \leq D \right\}$$

が成り立つ. よって (32) は,

$$\Pr \left\{ \max_{1 \leq i \leq n} d(X_i, \hat{Y}_i^{(n)}) \leq D \right\} > 1 - \delta \Leftrightarrow \Pr \left\{ \max_{1 \leq i \leq n} d(X_i, \hat{Y}_i^{(n)}) > D \right\} < \delta$$

と表すことができる. 以上より, 「歪み測度 d' による平均歪みに対して符号化レート R は CK 最大忠実度規範 0 で達成可能」と「歪み測度 d による最大歪みに対して符号化レート R は CK 最大忠実度規範 D で達成可能」は同値である. よって, 達成可能なレートの下限は,

$$R_3(D|X) = \min_{Y: \Pr\{d(X, Y) \leq D\}=1} I(X; Y) \quad (33)$$

である. $\square$

4 まとめ

連続的な情報の伝送は無歪みでは不可能なため、有歪み圧縮を用いることで伝送が可能になる。そして、一定の歪みを許容した上で情報をどれだけ圧縮できるかを目的とした。そのうえで、歪み測度を2通り、忠実度規範を4種類考え、それぞれの歪み測度に対してそれぞれの忠実度規範を適用することで最小符号化レートを求めた。全部で8種類の組み合わせがあり、本研究ではそのうち3種類の組み合わせにおける最小符号化レートを記述した。平均歪み測度に対して平均忠実度規範 D で達成可能な最小符号化レートはより詳細に証明を与えた。平均歪み測度に対して CK 最大忠実度規範 D で達成可能な最小符号化レートは説明を省略した。最大歪み測度に対して CK 最大忠実度規範 D で達成可能な最小符号化レート導出した。

本研究で求めた平均歪み測度に対する最小符号化レートは、同値となった。これは n を十分大きくすると CK 最大忠実度規範の δ が 0 に収束するため最大忠実度規範と同義とみなされ、同値となったと考えられる。また、CK 最大忠実度規範を平均歪み測度と最大歪み測度それぞれに対して適用したときの最小符号化レートは確率変数 Y の条件が異なる値が得られた。最大歪み測度のほうが平均歪み測度に比べ歪みが大きくなるため、確率変数 Y の条件がより厳しくなり、その結果より大きい最小符号化レートが得られた。

謝辞

本研究を終えるにあたり、終始一貫して丁寧なご指導を賜りました指導教員の西新幹彦准教授に心より感謝と敬意の意を申し上げます。

参考文献

- [1] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*, Academic Press, New York, 1981.
- [2] 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.
- [3] 植松 友彦, 現代シャノン理論, 培風館, 1998.