

信州大学工学部

学士論文

観測値に雑音のあるシステムの
レート・歪み関数の数値計算と基礎的考察

指導教員 西新 幹彦 准教授

学科 電子情報システム工学科
学籍番号 16T2079K
氏名 澤田 真子

2020 年 2 月 20 日

目次

1	はじめに	1
2	定常無記憶情報源のレート・歪み理論	1
3	数値計算のアルゴリズム	3
4	観測値に雑音のあるシステムのレート・歪み理論	4
5	観測値に雑音のあるシステムのレート・歪み関数の数値計算	8
6	まとめ	13
	謝辞	13
	参考文献	13
付録 A	定理 1 の証明	14
A.1	数学的準備	14
A.2	補題 1	15
A.3	定理 1 の証明	17
A.4	分散の一樣有界性	23
付録 B	5 章の数値計算のソースコード	25
付録 C	雑音がある場合のレート・歪み関数の計算	27

1 はじめに

実数値などを出力する連続情報源の場合、有限長のビットで値を表現すると、復元された値が元の値と完全に一致するようには表現できない。例えば携帯電話などで通話をするときに人間の声を完璧に再現して相手に届けることは不可能である。そこで、本人の声に近く聞こえる音声を作り、人間が判断できない周囲の小さな音などを削除する。このとき、削除する部分を大きくするとより効率よく送ることができるが、削除する範囲が多ければ元のデータとの差異は大きくなる。削除された部分を元の人間の声との歪みという。よって、元のデータ $\mathbf{x}$ と復号されたデータ $\mathbf{y}$ の間に歪みが生じることを許容しなければならない。したがって、あらかじめ与えられた歪み基準の範囲内で符号化レートをどれだけ小さくできるかという問題になる。これがレート・歪み問題である [1]。本論文では定常無記憶情報源に対するレート・歪み問題を考える。

近年、IoT の進化がモノとインターネットをつなげ、今では家庭やオフィスビル、工場、さらには都市まで、ありとあらゆるものがネットワークにつながるようになった。データを収集し活用する上で、重要な役割を担っているのがセンサである。センサによって、あらゆる場でデータが収集されさまざまな領域に利用され、離れたモノの状態を計測したり、モノと受信機との電波強度を測ることにより大まかな距離を知ることができる。しかし、センサの性能や感度などにより情報を得るときに、正確には観測できずいくらかノイズが含まれた形で観測される場合がある。また、データが届けられるときにノイズが入る場合もある。そこで本研究では、情報源に雑音を加えられた場合と受信者にデータが届く前に雑音を加えられた場合のレートと歪みの関係について考察した。

本論文は次のような章から構成される。2 章では 1959 年に Shannon により証明された定常無記憶情報源のレート・歪み理論について述べる [2]。3 章では Blahut が提案したレート・歪み関数を求めるアルゴリズムを述べる [3]。4 章では 1971 年に Berger が発表した雑音がある場合のレート・歪みを証明する [4]。5 章では 3 章で示したアルゴリズムを用いて、4 章で述べた観測値に雑音がある場合のレート・歪み問題を数値計算し、グラフを描画することによって雑音の影響を検討する。最後に 6 章で本論文をまとめる。

2 定常無記憶情報源のレート・歪み理論

$\mathbf{X} = \{X^n = (X_1, X_2, \dots, X_n)\}_{n=1}^{\infty}$ を定常無記憶情報源とし、情報源アルファベットを $\mathcal{X}$ 、復号アルファベットを $\mathcal{Y}$ とする。歪み測度 d を関数 $d: \mathcal{X} \times \mathcal{Y} \rightarrow [0, +\infty)$ と定め、 $d(x, y) (x \in \mathcal{X}, y \in \mathcal{Y})$ を x と y の間の歪みと呼ぶ。長さ n のデータ $\mathbf{x} = (x_1, x_2, \dots, x_n)$

と $\mathbf{y} = (y_1, y_2, \dots, y_n)$ の歪み $d_n(\mathbf{x}, \mathbf{y})$ を

$$d_n(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^n d(x_i, y_i) \quad (1)$$

と定義する．すると、情報源文字 1 個あたりの歪みは

$$\frac{1}{n}d_n(\mathbf{x}, \mathbf{y}) \geq 0 \quad (2)$$

で与えられる．

歪み測度 d_n が与えられたときの情報源 $\mathbf{X}$ の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．そして、符号化を符号器 $\psi_n : \mathcal{X}^n \rightarrow \mathcal{M}_n$ によって定め、復号化を復号器 $\varphi_n : \mathcal{M}_n \rightarrow \mathcal{Y}^n$ によって定める．情報源の出力 $\mathbf{x} \in \mathcal{X}$ は符号器 ψ_n によって符号化されたあと復号器 φ_n によって復元される．

では、定常無記憶情報源の達成可能性を定義する．

定義 1 符号 (φ_n, ψ_n) のもとでの情報源シンボル 1 個あたりの平均歪みを $\frac{1}{n}\mathbb{E}[d_n(X^n, \psi_n(\varphi_n(X^n)))]$ としたとき、レート R が歪み測度 d のもとで D -達成可能とは

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \mathbb{E}[d_n(X^n, \psi_n(\varphi_n(X^n)))] \leq D \quad (3)$$

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log M_n \leq R \quad (4)$$

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．

定義 2 レート・歪み関数は以下のように定義される．

$$R_0(D) \triangleq \inf \{ R | R \text{ が } D \text{-達成可能} \} \quad (5)$$

X と Y の間の相互情報量 $I(X; Y)$ を

$$I(X; Y) \triangleq \sum_{(x, y) \in \mathcal{X} \times \mathcal{Y}} P_{XY}(x, y) \log \frac{P_{Y|X}(y|x)}{P_Y(y)} \quad (6)$$

と定義すると、次の定理が成立する．

定理 1 ([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

$$R_0(D) = \min_{Y: \mathbb{E}[d(X, Y)] \leq D} I(X; Y) \quad (7)$$

で与えられる．

定理 1 の証明は付録に記載する．

3 数値計算のアルゴリズム

レート・歪み関数を数値計算によって求めるアルゴリズムが知られている [3]. 初期値に s, q_k^0 を与える. パラメータ s はパラメータ化する点でのレート・歪み関数の接線の傾きに等しくなり, $s < 0$ である. 任意の出力確率ベクトルを q_k で表し, 特に初期値で与えた q_k を q_k^0 とする. j 番目の文字が確率 p_j で生成された場合, 生成された j 番目の文字が k 番目の文字に復号されたときの歪みを d_{jk} で表す. $j = k$ のとき $d_{jk} = 0$ である. $A_{jk} = \exp(sd_{jk})$, $c_k = \sum_j p_j \frac{A_{jk}}{\sum_k A_{jk} q_k}$ を与えると, ある点 $D = \sum_j \sum_k p_j \frac{A_{jk} q_k}{\sum_k A_{jk} q_k} d_{jk}$ で,

$$R(D) \leq sD - \sum_j p_j \log \sum_k A_{jk} q_k - \sum_k q_k \log c_k \quad (8)$$

$$R(D) \geq sD - \sum_j p_j \log \sum_k A_{jk} q_k - \max_k \log c_k \quad (9)$$

が得られる. ここで, $T_U = \sum_k q_k \log c_k$, $T_L = \max_k \log c_k$ とし, q_k, c_k の更新を繰り返すことで, 差 $T_U - T_L$ がある値 ε より小さくなったときの D と $R(D)$ の値を求める. このアルゴリズムをまとめたものを Algorithm1 として示す.

Algorithm 1 レート・歪み関数を求めるアルゴリズム

Input: s, q_k^0

Output: $D, R(D)$

$A_{jk} = e^{sd_{jk}}$

while $T_U - T_L < \varepsilon$ **do**

$c_k = \sum_j p_j \frac{A_{jk}}{\sum_k q_k A_{jk}}$

$q_k = q_k c_k$

$T_U = \sum_k q_k \log c_k$

$T_L = \max_k \log c_k$

end while

$Q_{k|j} = \frac{A_{jk} q_k}{\sum_k A_{jk} q_k}$

$D = \sum_j \sum_k p_j Q_{k|j} d_{jk}$

$R(D) = sD - \sum_j p_j \log \sum_k A_{jk} q_k - \sum_k q_k \log c_k$

4 観測値に雑音のあるシステムのレート・歪み理論

ある情報を得るときに正確には観測できずいくらかノイズを含んだ形で得られる場合と、復号器から受信者に届く前にデータが破損する状況を考える [4]. $X^n = (X_1, X_2, \dots, X_n)$ を定常無記憶情報源とし、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. 情報源から符号器に届くときの X から X^* への変換が確率 $P_{X^*|X}(x^*|x)$ で行われる. X^* は符号器 φ_n で符号化され、通信路を通り、復号器 φ_n で復号される. 復号器から受信者に届くときの Y^* から Y への変換が確率 $P_{Y|Y^*}(y|y^*)$ で行われ、受信者に届いたものを Y とする. 長さ n のとき Y^n とする. 通信路では正確に情報が届けられるとし、この問題をモデル化したものを図 1 に示す. このとき x と y の間の歪みを $d(x, y) (x \in \mathcal{X}, y \in \mathcal{Y})$ とし、長さ n のデータ $\mathbf{x} = (x_1, x_2, \dots, x_n \in \mathcal{X}^n)$ と $\mathbf{y} = (y_1, y_2, \dots, y_n \in \mathcal{Y}^n)$ の歪み $d_n(\mathbf{x}, \mathbf{y})$ を

$$d_n(\mathbf{x}, \mathbf{y}) \triangleq \sum_{i=1}^n d(x_i, y_i) \quad (10)$$

と定義する.

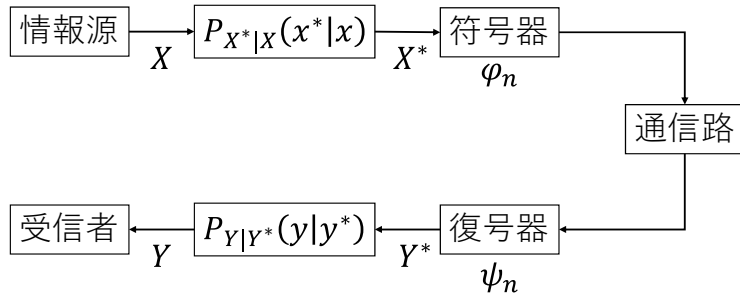


図 1 雑音がある場合の通信システムのモデル

レート・歪み関数は次のように定義される.

定義 3 歪み測度 d のもとでレート R が D -達成可能とは

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \mathbb{E}[d_n(X^n, Y^n)] \leq D \quad (11)$$

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log M_n \leq R \quad (12)$$

となるような符号の列 $\{(\varphi_n, \psi_n)\}_{n=1}^{\infty}$ が存在することである. ただし M_n は符号語数である.

定義 4 雑音がある場合のレート・歪み関数は以下のように定義される.

$$R(D) \triangleq \inf\{R \mid \text{歪み測度 } d \text{ のもとでレート } R \text{ は } D\text{-達成可能} \} \quad (13)$$

$d^*(x^*, y^*) (x^* \in \mathcal{X}, y^* \in \mathcal{Y})$ を

$$d^*(x^*, y^*) \triangleq \frac{1}{P_{X^*}(x^*)} \sum_{x, y} P_X(x) P_{X^*|X}(x^*|x) P_{Y|Y^*}(y|y^*) d(x, y) \quad (14)$$

と定義する.

定理 2 ([4]) 雑音がある場合のレート・歪み関数は

$$R(D) = \min_{Y^*: \mathbb{E}[d^*(X^*, Y^*)] \leq D} I(X^*; Y^*) \quad (15)$$

で与えられる. ここで, X^* は $P_{X^*}(x^*) = \sum_x P_X(x) P_{X^*|X}(x^*|x)$ に従い, Y^* は X^* と相関のある確率変数である.

証明 $\mathbf{x}^* = (x_1^*, x_2^*, \dots, x_n^*)$ と $\mathbf{y}^* = (y_1^*, y_2^*, \dots, y_n^*)$ の歪み $d_n^*(\mathbf{x}^*, \mathbf{y}^*)$ を

$$d_n^*(\mathbf{x}^*, \mathbf{y}^*) \triangleq \sum_{i=1}^n d^*(x_i^*, y_i^*) \quad (16)$$

とする. d_n^* の期待値は

$$\mathbb{E}[d_n^*(X^{*n}, Y^{*n})] = \sum_{\mathbf{x}^*, \mathbf{y}^*} P_{X^*}(\mathbf{x}^*) P_{Y^*|X^*}(\mathbf{y}^*|\mathbf{x}^*) d_n^*(\mathbf{x}^*, \mathbf{y}^*) \quad (17)$$

$$= \sum_{\mathbf{x}^*, \mathbf{y}^*} \left\{ \prod_{j=1}^n P_{X^*}(x_j^*) \right\} P_{Y^*|X^*}(\mathbf{y}^*|\mathbf{x}^*) \sum_{k=1}^n d(x_k^*, y_k^*) \quad (18)$$

$$= \sum_{k=1}^n \sum_{\mathbf{x}^*, \mathbf{y}^*} d(x_k^*, y_k^*) \left\{ \prod_{j=1}^n P_{X^*}(x_j^*) \right\} P_{Y^*|X^*}(\mathbf{y}^*|\mathbf{x}^*) \quad (19)$$

と書ける．一方， d_n の期待値は

$$\mathbb{E}[d_n(X^n, Y^n)] \quad (20)$$

$$= \sum_{\mathbf{x}, \mathbf{x}^*, \mathbf{y}^*, \mathbf{y}} P_X(\mathbf{x}) P_{X^*|X}(\mathbf{x}^*|\mathbf{x}) P_{Y^*|X^*}(\mathbf{y}^*|\mathbf{x}^*) P_{Y|Y^*}(\mathbf{y}|\mathbf{y}^*) d_n(\mathbf{x}, \mathbf{y}) \quad (21)$$

$$= \sum_{k=1}^n \sum_{\mathbf{x}, \mathbf{x}^*, \mathbf{y}^*, \mathbf{y}} d(x_k, y_k) \left\{ \prod_{j=1}^n P_X(x_j) P_{X^*|X}(x_j^*|x_j) P_{Y|Y^*}(y_j|y_j^*) \right\} P_{Y^*|X^*}(\mathbf{y}^*|\mathbf{x}^*) \quad (22)$$

$$= \sum_{k=1}^n \sum_{\mathbf{x}^*, \mathbf{y}^*} P_{Y^*|X^*}(\mathbf{y}^*|\mathbf{x}^*) \sum_{\mathbf{x}, \mathbf{y}} d(x_k, y_k) \prod_{j=1}^n P_X(x_j) P_{X^*|X}(x_j^*|x_j) P_{Y|Y^*}(y_j|y_j^*) \quad (23)$$

$$= \sum_{k=1}^n \sum_{\mathbf{x}^*, \mathbf{y}^*} P_{Y^*|X^*}(\mathbf{y}^*|\mathbf{x}^*) \sum_{x_k, y_k} d(x_k, y_k) P_X(x_k) P_{X^*|X}(x_k^*|x_k) P_{Y|Y^*}(y_k|y_k^*) \\ \times \sum_{\substack{x_1, \dots, x_n \\ y_1, \dots, y_n \\ x_k, y_k \text{ 除く}}} \prod_{\substack{j=1 \\ j \neq k}}^n P_X(x_j) P_{X^*|X}(x_j^*|x_j) P_{Y|Y^*}(y_j|y_j^*) \quad (24)$$

であり， $j \neq k$ の部分について考えると

$$\sum_{\substack{x_1, \dots, x_n \\ y_1, \dots, y_n \\ x_k, y_k \text{ 除く}}} \prod_{\substack{j=1 \\ j \neq k}}^n P_X(x_j) P_{X^*|X}(x_j^*|x_j) P_{Y|Y^*}(y_j|y_j^*) \quad (25)$$

$$= \sum_{x_1, y_1} \cdots \sum_{x_n, y_n} \{P_X(x_1) P_{X^*|X}(x_1^*|x_1) P_{Y|Y^*}(y_1|y_1^*) \\ \cdots P_X(x_n) P_{X^*|X}(x_n^*|x_n) P_{Y|Y^*}(y_n|y_n^*)\} \quad (k \text{ 除く}) \quad (26)$$

$$= \sum_{x_1} \cdots \sum_{x_n} P_X(x_1) P_{X^*|X}(x_1^*|x_1) \cdots P_X(x_n) P_{X^*|X}(x_n^*|x_n) \quad (k \text{ 除く}) \quad (27)$$

$$= \left(\sum_{x_1} P_X(x_1) P_{X^*|X}(x_1^*|x_1) \right) \cdots \left(\sum_{x_n} P_X(x_n) P_{X^*|X}(x_n^*|x_n) \right) \quad (k \text{ 除く}) \quad (28)$$

$$= \prod_{\substack{j=1 \\ j \neq k}}^n P_{X^*}(x_j^*) \quad (29)$$

となり，式 (14)，式 (24)，式 (29) より

$$\mathbb{E}[d_n(X^n, Y^n)] \quad (30)$$

$$= \sum_{k=1}^n \sum_{\mathbf{x}^*, \mathbf{y}^*} P_{Y^*|X^*}(\mathbf{y}^*|\mathbf{x}^*) \left\{ \sum_{x_k, y_k} d(x_k, y_k) P_X(x_k) P_{X^*|X}(x_k^*|x_k) P_{Y|Y^*}(y_k|y_k^*) \right\} \prod_{j \neq k} P_{X^*}(x_j^*) \quad (31)$$

$$= \sum_{k=1}^n \sum_{\mathbf{x}^*, \mathbf{y}^*} d(x_k^*, y_k^*) \left\{ \prod_{j=1}^n P_{X^*}(x_j^*) \right\} P_{Y^*|X^*}(\mathbf{y}^*|\mathbf{x}^*) \quad (32)$$

となる。したがって、式 (19) より

$$\mathbb{E}[d_n(X^n, Y^n)] = \mathbb{E}[d_n^*(X^{*n}, Y^{*n})] \quad (33)$$

が導かれる。すると、定義 3 の式 (11) は

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \mathbb{E}[d_n^*(X^{*n}, Y^{*n})] \leq D \quad (34)$$

と同値である。すなわち、歪み測度 d のもとで D -達成可能であるとは、歪み測度 d^* のもとで D -達成可能である、ということである。ただし、歪み測度 d は情報源 X に対しての歪みであり、歪み測度 d^* は情報源 X^* に対しての歪みである。したがって、

$$R(D) = \inf\{R \mid \text{歪み測度 } d \text{ のもとでレート } R \text{ は } D\text{-達成可能}\} \quad (35)$$

$$= \inf\{R \mid \text{歪み測度 } d^* \text{ のもとでレート } R \text{ は } D\text{-達成可能}\} \quad (36)$$

$$= \min_{Y^*: \mathbb{E}[d^*(X^*, Y^*)] \leq D} I(X^*; Y^*) \quad (37)$$

となる。よって定理 2 が導かれた。 $\square$

定理 2 より雑音がある場合のレート・歪みの問題は図 2 に示すような定常無記憶情報源のレート・歪み理論に帰着することができる。ただし情報源 X^* は $P_{X^*}(x^*) = \sum_x P_X(x) P_{X^*|X}(x^*|x)$ に従う確率変数である。

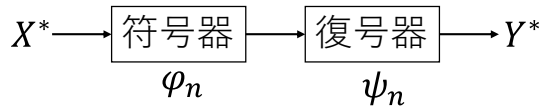


図 2 定常無記憶情報源の通信システム

5 観測値に雑音のあるシステムのレート・歪み関数の数値計算

3章で説明したアルゴリズムを用いて観測値に雑音がある場合のレート・歪み関数を求める。情報源から符号器に届くときに雑音が入る確率 $P_{X^*|X}(x^*|x)$ と、復号器から受信者に届くときに雑音が入る確率 $P_{Y|Y^*}(y|y^*)$ を変化させ、レート・歪み関数のグラフを描き、雑音による影響を考察する。簡単化のため、ハミング歪みのもとで2元情報源に対するレート・歪みを考える。ハミング歪みとは $d(x, y)(x \in \mathcal{X}, y \in \mathcal{Y})$ を x と y の間の歪みとしたとき

$$d(x, y) \triangleq \begin{cases} 0 & (x = y \text{ のとき}) \\ 1 & (x \neq y \text{ のとき}) \end{cases} \quad (38)$$

で与えられる。2元情報源は $\mathcal{X} = \{0, 1\}$ とする。確率分布は

$$P_X(0) = 0.5, P_X(1) = 0.5 \quad (39)$$

$$P_X(0) = 0.4, P_X(1) = 0.6 \quad (40)$$

$$P_X(0) = 0.3, P_X(1) = 0.7 \quad (41)$$

の3つの場合を考え、雑音の入り方は2元対称通信路と同じものとする。したがって、情報源から符号器に届くときに雑音が入り、反転する確率を p_x とすると、

$$P_{X^*|X}(x^*|x) = \begin{cases} 1 - p_x & (x^* = x \text{ のとき}) \\ p_x & (x^* \neq x \text{ のとき}) \end{cases} \quad (42)$$

となり、復号器から受信者に届くときの反転確率を p_y とおくと

$$P_{Y|Y^*}(y|y^*) = \begin{cases} 1 - p_y & (y = y^* \text{ のとき}) \\ p_y & (y \neq y^* \text{ のとき}) \end{cases} \quad (43)$$

となる。なお、 $p_x, p_y < 0.5$ とする。

最初に式 (39) から式 (41) のそれぞれの確率分布に対し、雑音が $P_{X^*|X}(x^*|x)$ のみ、 $P_{Y|Y^*}(y|y^*)$ のみの場合を考える。 $P_X(0) = 0.5$ のもとで反転確率 p_x を 0 から 0.4 まで 0.1 刻みで変化させた場合のレート・歪み関数を図 3 に示す。このとき復号器から受信者に届くときの反転確率 p_y は 0 とし、正確にデータが届くとする。左のグラフから順に $p_x = 0, 0.1, 0.2, 0.3, 0.4$ である。同様に p_y を 0 から 0.4 まで 0.1 刻みで変化させたときのレート・歪み関数を図 4 に示す。このとき反転確率 p_x は 0 とする。左のグラフから順に $p_y = 0, 0.1, 0.2, 0.3, 0.4$ である。また、 $P_X(0) = 0.4$ のもとで p_x を 0, 0.1, 0.2, 0.3 と変化したときのレート・歪み関数を図 5 に、 $p_y = 0, 0.1, 0.2, 0.3, 0.4$ に変化させたときのレート・歪み関数を図 6 に示す。 $P_X(0) = 0.3$ のもとで確率 p_x を 0, 0.1, 0.2 と変化したときのレート・歪み関数を図 7 に、 $p_y = 0, 0.1, 0.2, 0.3$ に変化させたときのレート・歪み関数を図 8 に示す。

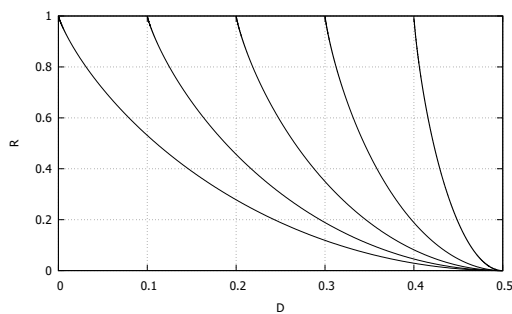


図3 $P_X(0) = 0.5$ のとき p_x を変化

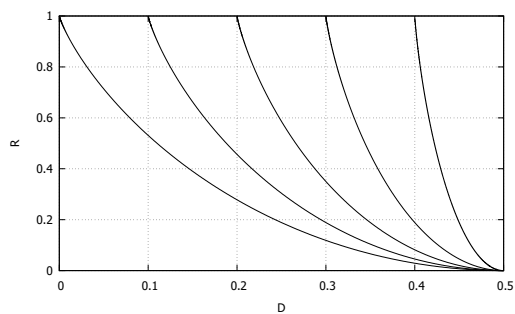


図4 $P_X(0) = 0.5$ のとき p_y を変化

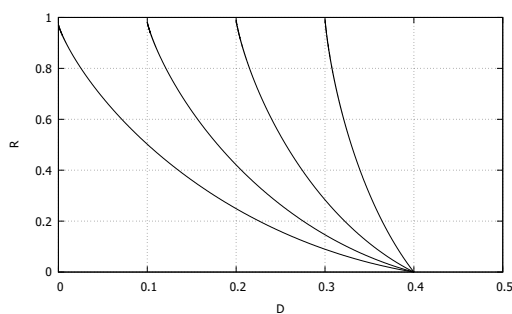


図5 $P_X(0) = 0.4$ のとき p_x を変化

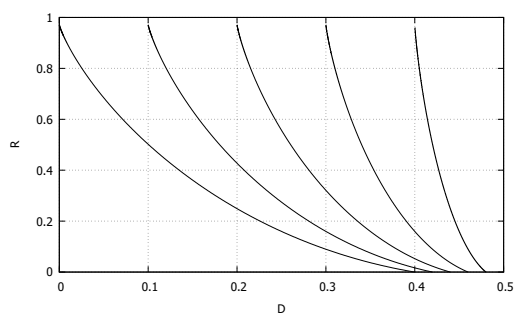


図6 $P_X(0) = 0.4$ のとき p_y を変化

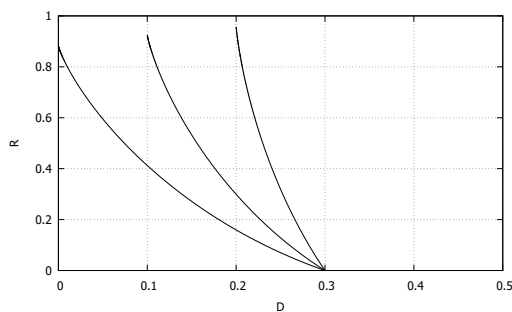


図7 $P_X(0) = 0.3$ のとき p_x を変化

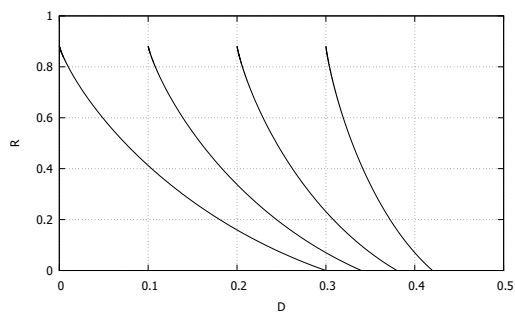


図8 $P_X(0) = 0.3$ のとき p_y を変化

全体は単調減少で下に凸のグラフとなった。これは、歪みを0に近づけるとレートは大きくなり、歪みを許すことでレートが小さくできるというレートと歪みの関係を示している。 $P_X(0) = 0.4, 0.3$ の場合、 $P_X(0) = p_x$ のときはグラフが描けなかった。 $P_X(0) = p_x$ のときのグラフは垂直に線が立つと予想される。しかし、3章のアルゴリズムはレート・歪み関数の

接線の傾き s を与えると、その傾きを持つ点が得られる。垂直になる場合、傾きは $-\infty$ になるので計算ができなかったと考えられる。

反転確率 p_x, p_y を 0 から変化させた場合、図 3 から図 8 に示したようにいずれの確率分布でもレート R が最大値をとるときの D の値は反転確率より小さくできない。 p_x, p_y が正のとき、レートを上げて符号器と復号器の間で正確に通信されても、情報源がある一定の確率 p_x, p_y で反転しているため、情報源 X と受信者が受け取る Y には必ず誤りが生じるからと考えられる。付録 C に $P_X(0) = 0.4$, 反転確率 $p_x = 0.5$ の場合、 $D < 0.4$ の範囲で任意のレートは達成可能でないことを証明する。

図 5 と図 6, 図 7 と図 8 から $R = 0$ のときの D の値に違いがみられた。 p_y が変化すると D の値も変化することがわかる。 $R = 0$ とは、図 1 の通信路で通信をしないことと同じである。ただし、歪みは小さくしたいので $P_X(0) < P_X(1)$ より、復号器では 1 を出し続ける。 $P_{X^*|X}(x^*|x)$ のみの場合、確率 $P_X(1)$ で x と y は一致し、 $P_X(0)$ が歪みとなるが、 $P_{Y|Y^*}(y|y^*)$ がある場合、復号器から受信者に届く間に p_y で反転するため、データは正確に届かなくなり、さらに歪む。よって D の値が大きくなったと考えられる。しかし、 $P_X(0) = P_X(1) = 0.5$ のとき、 $P_{Y|Y^*}(y|y^*)$ が存在するにも関わらず図 3 と図 4 は同様の図が得られた。 $R = 0$ のときの歪み D を考える。 $R = 0$ なので通信路では通信を行わず、復号器では 1 を出力し続けるとする。元のデータと受け取るデータが異なるのは、情報源で 0 を出力し復号器で出力された 1 が受信者に確率 $(1 - p_y)$ で正しく届き 1 を受け取る場合と、情報源は 1 を出し復号器で 1 を出力したが確率 p_y で反転し 0 が届く場合の 2 つの状況が考えられる。このときの歪み D は

$$D = P_X(0)(1 - p_y) + P_X(1)p_y \quad (44)$$

とかける。 $P_X(0) = P_X(1) = 0.5$ であるので、式 (44) は

$$D = P_X(0) \quad (45)$$

$$= 0.5 \quad (46)$$

となる。よって情報源が等確率で $P_X(0) = P_X(1) = 0.5$ のとき $P_{Y|Y^*}(y|y^*)$ が存在しても、 $R = 0$ のときの D の値は $D = 0.5$ となる。よって、 $P_X(0) = P_X(1) = 0.5$ のとき、図 3 と図 4 は同様の図が得られた。

次に、 $P_{X^*|X}(x^*|x)$, $P_{Y|Y^*}(y|y^*)$ がともに存在する場合について考える。 $p_y = 0.1$ に固定し、 p_x を 0 から 0.25 まで 0.05 刻みで変化させたグラフを図 9 に示す。左から順に $p_x = 0, 0.05, 0.10, 0.15, 0.20, 0.25$ である。そして $p_x = 0.1$ に固定し、 p_y を 0 から 0.3 まで 0.05 刻みで変化させたグラフを図 10 に示す。

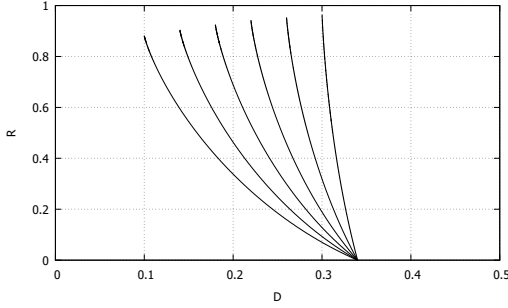


図9 $P_X(0) = 0.3$ のとき $p_y = 0.1$ に固定， p_x を変化

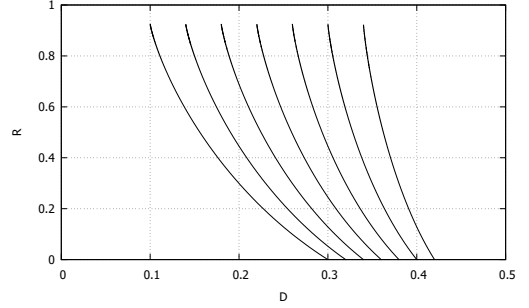


図10 $P_X(0) = 0.3$ のとき $p_x = 0.1$ に固定， p_y を変化

$P_{X^*|X}(x^*|x)$ のみのときと同様に $P_X(0) = p_x$ のときはグラフが描けなかった．先述のようにアルゴリズムが関係していると考えられる．レート R が最大になるときの D の値は図9のように， $P_{Y|Y^*}(y|y^*)$ の雑音の影響に $P_{X^*|X}(x^*|x)$ の雑音の影響が加わった形となった．図10も同様に $P_{X^*|X}(x^*|x)$ の雑音の影響に $P_{Y|Y^*}(y|y^*)$ の雑音の影響が加わったグラフとなった． p_x, p_y がともに存在するとき，情報源 X は $P_{X^*|X}(x^*|x)$ で歪み，レートを上げて通信路で確実にデータが届けられたとしても， $P_{X^*|X}(x^*|x)$ で歪んだデータが届けられる．そこからさらに $P_{Y|Y^*}(y|y^*)$ で歪むため， $P_{X^*|X}(x^*|x)$ と $P_{Y|Y^*}(y|y^*)$ の雑音の影響が合わさったと考えられる． $R = 0$ となる D の値は図9の場合1点に集まったが，図10の場合， p_y が大きくなるにつれて $R = 0$ となる D の値も大きくなった． $P_{X^*|X}(x^*|x)$ の雑音の影響より $P_{Y|Y^*}(y|y^*)$ の雑音の影響の方が大きいと予想される．

p_y が変化するとグラフも右にずれ， D の値が大きくなる理由を考察する．単純化のため反転確率 $p_x = 0$ とすると，式(42)より $P_{X^*|X}(x^*|x) = 1$ である．また，式(14)より，このときの d^* は

$$d^*(x^*, y^*) = \frac{1}{P_{X^*}(x^*)} \sum_{x,y} P_X(x) P_{Y|Y^*}(y|y^*) d(x, y) \quad (47)$$

となる．求めるのは $R(D) = 0$ のときの D の値であるので式(15)より

$$\min_{Y^*: \mathbb{E}[d^*(X^*, Y^*)] \leq D} I(X^*; Y^*) = 0 \quad (48)$$

となる D の値を求めたい．常に $I(X^*; Y^*) \geq 0$ であるので， $R(D) = 0$ を達成する Y^* は

$$I(X^*; Y^*) = 0 \quad (49)$$

を満たす．すなわち， X^* と Y^* は独立である．そのような Y^* の中で最小の $\mathbb{E}[d^*(X^*, Y^*)]$

を求める。よって、式 (43)，式 (47) より

$$\begin{aligned} \mathbb{E}[d^*(X^*, Y^*)] &= \sum_{x^*, y^*} P_{X^*}(x^*) P_{Y^*}(y^*) d^*(x^*, y^*) \end{aligned} \quad (50)$$

$$= \sum_{x^*, y^*} P_{X^*}(x^*) P_{Y^*}(y^*) \frac{1}{P_{X^*}(x^*)} \sum_{x, y} P_X(x) P_{Y|Y^*}(y|y^*) d(x, y) \quad (51)$$

$$= \sum_{y^*} P_{Y^*}(y^*) \sum_{x, y} P_X(x) P_{Y|Y^*}(y|y^*) d(x, y) \quad (52)$$

$$= P_{Y^*}(0) \sum_{x, y} P_X(x) P_{Y|Y^*}(y|0) d(x, y) + P_{Y^*}(1) \sum_{x, y} P_X(x) P_{Y|Y^*}(y|1) d(x, y) \quad (53)$$

$$= P_{Y^*}(0) \{P_X(0) P_{Y|Y^*}(1|0) + P_X(1) P_{Y|Y^*}(0|0)\} \\ + P_{Y^*}(1) \{P_X(0) P_{Y|Y^*}(1|1) + P_X(1) P_{Y|Y^*}(0|1)\} \quad (54)$$

$$= P_{Y^*}(0) \{P_X(0) p_y + P_X(1)(1 - p_y)\} + P_{Y^*}(1) \{P_X(0)(1 - p_y) + P_X(1) p_y\} \quad (55)$$

となる。ここで、 $\mathbb{E}[d^*(X^*, Y^*)]$ 最小化するため、 $P_{Y^*}(0)$ の項と $P_{Y^*}(1)$ の項の大小関係を比較する。2 元情報源であるので、 $P_X(1) = 1 - P_X(0)$ である。 $P_{Y^*}(0)$ の項は

$$P_X(0) p_y + P_X(1)(1 - p_y) = P_X(0) p_y + (1 - P_X(0))(1 - p_y) \quad (56)$$

$$= P_X(0) p_y + 1 - p_y - P_X(0) + P_X(0) p_y \quad (57)$$

$$= (2P_X(0) - 1) p_y + 1 - P_X(0) \quad (58)$$

となる。一方、 $P_{Y^*}(1)$ の項は

$$P_X(0)(1 - p_y) + P_X(1) p_y = P_X(0) - P_X(0) p_y + (1 - P_X(0)) p_y \quad (59)$$

$$= (-2P_X(0) + 1) p_y + P_X(0) \quad (60)$$

となる。式 (58)，式 (60) の差をとると

$$\begin{aligned} &\{(2P_X(0) - 1) p_y + 1 - P_X(0)\} - \{(-2P_X(0) + 1) p_y + P_X(0)\} \\ &= 4P_X(0) p_y - 2p_y - 2P_X(0) + 1 \end{aligned} \quad (61)$$

$$(62)$$

$P_X(0) \leq 0.5$, $p_y < 0.5$ のため、

$$4P_X(0) p_y - 2p_y - 2P_X(0) + 1 > 0 \quad (63)$$

である。したがって、

$$(2P_X(0) - 1) p_y + 1 - P_X(0) > (-2P_X(0) + 1) p_y + P_X(0) \quad (64)$$

となるので、

$$P_{Y^*}(0) = 0, P_{Y^*}(1) = 1 \quad (65)$$

のとき最小値をとり，実際最小値は

$$\mathbb{E}[d^*(X^*, Y^*)] = P_X(0)(1 - p_y) + P_X(1)p_y \quad (66)$$

$$= P_X(0) + \{P_X(1) - P_X(0)\}p_y \quad (67)$$

となる．したがって，確率 p_y が変化するとレートが 0 のときの歪み D の値も変化することが示された． D の値は確率分布 $P_X(x)$ と反転確率 p_y によって決まる．

6 まとめ

本研究では，雑音によりデータが破損する場合のレート・歪み問題を考えた．雑音がある場合でも雑音のないレート・歪み理論に帰着することができる．それゆえ，レート・歪み関数を求めるアルゴリズムを用いて 2 元情報源，ハミング歪みのもとで数値計算を行った．結果は，情報源から符号器に届くときの雑音が入る確率 $P_{X^*|X}(x^*|x)$ がある場合と，復号器から受信者に届くときの雑音が入る確率 $P_{Y|Y^*}(y|y^*)$ がある場合でグラフが異なることが分かった．そして， $P_{Y|Y^*}(y|y^*)$ がある場合， $R = 0$ となる D の値は情報源の確率分布 $P_X(x)$ と反転確率 p_y によって定まることが確認できた．

謝辞

本研究を行うにあたり，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] Claude E. Shannon, Coding Theorems for a Discrete Source With a Fidelity Criterion, IRE Nat. Cony. Rec., 1959.
- [3] Richard E. Blahut, “Computation of Channel Capacity and Rate-Distortion Functions,” IEEE TRANSACTIONS ON INFORMATION THEORY, VOL. IT-18 NO. 4, pp.460–473, JULY. 1972.
- [4] Toby Berger, *RATE DISTORTION THEORY A Mathematical Basis For Data Compression*, PRENTICE-HALL, Inc. Englewood Cliffs, New Jersey, 1971.
- [5] 伊藤祐樹，「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」，信州大学工学部卒業論文（指導教員：西新幹彦），2019 年 3 月．

付録 A 定理 1 の証明

定理 1 の証明の前に以下の準備をする.

A.1 数学的準備

(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

$$\Pr\{Z \geq a\} \leq \frac{\mathbb{E}[Z]}{a} \quad (68)$$

が成り立つ.

証明 平均を μ , Z の実現値を z とすると期待値は

$$\mathbb{E}[Z] = \sum_z z P_Z(z) \quad (69)$$

$$= \sum_{z>a} z P_Z(z) + \sum_{z\leq a} z P_Z(z) \quad (70)$$

$$\geq \sum_{z>a} P_Z(z) \quad (71)$$

$$= a \Pr\{Z > a\} \quad (72)$$

と変形でき,

$$\Pr\{Z \geq a\} \leq \frac{\mathbb{E}[Z]}{a} \quad (73)$$

が導かれる. $\square$

(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

$$\Pr\{|Z - \mu| > a\sigma\} \leq \frac{1}{a^2} \quad (74)$$

が成り立つ.

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

$$\Pr\{(Z - \mu)^2 > a^2\sigma^2\} \leq \frac{\mathbb{E}[(Z - \mu)^2]}{a^2\sigma^2} \quad (75)$$

$\mathbb{E}[(Z - \mu)^2] = \sigma^2$ なので

$$\Pr\{|Z - \mu| > a\sigma\} \leq \frac{1}{a^2} \quad (76)$$

が導かれる. $\square$

A.2 補題 1

補題 1 [5] 情報源 X^n に対し, それと相関のある Y^n 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 $\{\hat{W}^n: \mathcal{Y}^n \rightarrow \mathcal{X}^n\}$ に対して

$$\hat{W}^n(\mathbf{x}|\mathbf{y}) \triangleq \Pr\{X^n = \mathbf{x} | Y^n = \mathbf{y}\} \quad (77)$$

$$\mathcal{B}_n(\mathbf{y}) \triangleq \left\{ \mathbf{x} \left| \frac{1}{n} \log \frac{\hat{W}^n(\mathbf{x}|\mathbf{y})}{P_{X^n}(\mathbf{x})} \leq R_n - \gamma \right. \right\} \quad (78)$$

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

$$\frac{1}{n} \log M_n \leq R_n \quad (79)$$

$$\Pr\{X^n \notin \mathcal{S}_n(\psi_n(\varphi_n(X^n)))\} \leq \Pr\{X^n \notin \mathcal{B}_n(Y^n)\} + \Pr\{X^n \notin \mathcal{S}(Y^n)\} + e^{-n\gamma} \quad (80)$$

を満たす (φ_n, ψ_n) 符号が存在する.

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

$$\hat{W}^n(\mathcal{T}_n(\mathbf{y})|\mathbf{y}) \geq e^{-n\gamma} \quad (81)$$

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. 以降, $m \geq 2$ に対し, m 番目の復号語として

$$\hat{W}^n(\mathcal{T}_n(\mathbf{y}) \setminus \bigcup_{m' < m} \varphi_n^{-1}(m') | \mathbf{y}) \geq e^{-n\gamma} \quad (82)$$

となる $\mathbf{y} \in \mathcal{Y}^n$ を任意に選び, これを復号語 $\psi_n(m)$ とする. さらに

$$\varphi_n^{-1}(m) \triangleq \mathcal{T}_n(\psi_n(m)) \setminus \bigcup_{m' < m} \varphi_n^{-1}(m') \quad (83)$$

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

$$\mathcal{D} \triangleq \bigcup_m \varphi_n^{-1}(m) \quad (84)$$

とおく. すると M_n の定義より任意の $\mathbf{y} \in \mathcal{Y}^n$ に対して

$$\hat{W}^n(\mathcal{T}_n(\mathbf{y}) \setminus \mathcal{D} | \mathbf{y}) < e^{-n\gamma} \quad (85)$$

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 (78) より

$$\frac{1}{n} \log \frac{\hat{W}^n(\mathbf{x}|\psi_n(m))}{P_{X^n}(\mathbf{x})} \leq R_n - \gamma \quad (86)$$

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

$$P_{X^n}(\mathbf{x}) \geq \hat{W}^n(\mathbf{x}|\psi_n(m))e^{-n(R_n - \gamma)} \quad (87)$$

が成り立つことに注意しておく．ここで式 (82), (84), (87) より

$$1 \geq P_{X^n}(\mathcal{D}) \quad (88)$$

$$= P_{X^n}\left(\bigcup_m \varphi_n^{-1}(m)\right) \quad (89)$$

$$= \sum_m P_{X^n}(\varphi_n^{-1}(m)) \quad (90)$$

$$= \sum_m \sum_{\mathbf{x} \in \varphi_n^{-1}(m)} P_{X^n}(\mathbf{x}) \quad (91)$$

$$\geq \sum_m \sum_{\mathbf{x} \in \varphi_n^{-1}(m)} \hat{W}^n(\mathbf{x}|\psi_n(m))e^{-n(R_n - \gamma)} \quad (92)$$

$$= \sum_m \hat{W}^n(\varphi_n^{-1}(m)|\psi_n(m))e^{-n(R_n - \gamma)} \quad (93)$$

$$\geq \sum_m e^{-nR_n} \quad (94)$$

$$= M_n e^{-nR_n} \quad (95)$$

となり

$$\frac{1}{n} \log M_n \leq R_n \quad (96)$$

が得られる．

次に式 (80) 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

$$\Pr\{X^n \in \mathcal{D}\} \leq \Pr\{X^n \in \mathcal{S}_n(\psi_n(\varphi_n(X^n)))\} \quad (97)$$

といえる．一方，式 (85) により任意の $\mathbf{y}$ について

$$\hat{W}^n(\mathcal{T}_n(\mathbf{y})|\mathbf{y}) = \hat{W}^n(\mathcal{T}_n(\mathbf{y}) \cap \mathcal{D}|\mathbf{y}) + \hat{W}^n(\mathcal{T}_n(\mathbf{y}) \setminus \mathcal{D}|\mathbf{y}) \quad (98)$$

$$\leq \hat{W}^n(\mathcal{D}|\mathbf{y}) + e^{-n\gamma} \quad (99)$$

となるので、式 (97) の左辺は

$$\Pr\{X^n \in \mathcal{D}\} = P_{X^n}(\mathcal{D}) \quad (100)$$

$$= \sum_{\mathbf{y}} P_{X^n Y^n}(\mathcal{D}, \mathbf{y}) \quad (101)$$

$$= \sum_{\mathbf{y}} P_{Y^n}(\mathbf{y}) \hat{W}^n(\mathcal{D}|\mathbf{y}) \quad (102)$$

$$\geq \sum_{\mathbf{y}} (\hat{W}^n(\mathcal{T}_n(\mathbf{y})|\mathbf{y}) - e^{-n\gamma}) \quad (103)$$

$$= \sum_{\mathbf{y}} \Pr\{X^n \in \mathcal{T}_n(Y^n), Y^n = \mathbf{y}\} - e^{-n\gamma} \quad (104)$$

$$= \Pr\{X^n \in \mathcal{T}_n(Y^n)\} - e^{-n\gamma} \quad (105)$$

$$= \Pr\{X^n \in \mathcal{B}_n(Y^n) \text{ か } X^n \in \mathcal{S}_n(Y^n)\} - e^{-n\gamma} \quad (106)$$

$$= 1 - \Pr\{X^n \notin \mathcal{B}_n(Y^n) \text{ または } X^n \notin \mathcal{S}_n(Y^n)\} - e^{-n\gamma} \quad (107)$$

$$\geq 1 - \Pr\{X^n \notin \mathcal{B}_n(Y^n)\} - \{X^n \notin \mathcal{S}_n(Y^n)\} - e^{-n\gamma} \quad (108)$$

式 (97) と式 (108) より

$$\Pr\{X^n \in \mathcal{S}_n(\psi_n(\varphi_n(X^n)))\} \geq 1 - \Pr\{X^n \notin \mathcal{B}_n(Y^n)\} - \{X^n \notin \mathcal{S}_n(Y^n)\} - e^{-n\gamma} \quad (109)$$

となるので

$$\Pr\{X^n \notin \mathcal{S}_n(\psi_n(\varphi_n(X^n)))\} \leq \Pr\{X^n \notin \mathcal{B}_n(Y^n)\} + \Pr\{X^n \notin \mathcal{S}_n(Y^n)\} + e^{-n\gamma} \quad (110)$$

が導かれる。 $\square$

A.3 定理 1 の証明

定理 1 の順定理の証明を行う。

証明 初めに、情報源 X に対し、 $\mathbb{E}[d(X, Y)] \leq D - \gamma$ を満たす任意の Y を考える。ただし、 $\gamma > 0$ は任意の定数である。 $X^n Y^n = (X_1 Y_1, X_2 Y_2, \dots, X_n Y_n)$ 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、

$$T_n^{(1)} = \left\{ (\mathbf{x}, \mathbf{y}) \in \mathcal{X}^n \times \mathcal{Y}^n \left| \frac{1}{n} \log \frac{P_{Y^n|X^n}(\mathbf{y}|\mathbf{x})}{P_{Y^n}(\mathbf{y})} - I(X; Y) \right| < \gamma \right\} \quad (111)$$

$$T_n^{(2)} = \left\{ (\mathbf{x}, \mathbf{y}) \in \mathcal{X}^n \times \mathcal{Y}^n \left| \frac{1}{n} d_n(\mathbf{x}, \mathbf{y}) - \mathbb{E}[d(X, Y)] \right| < \gamma \right\} \quad (112)$$

として, $T_n = T_n^{(1)} \cap T_n^{(2)}$ とおく. $X^n Y^n$ が定常独立系列なので,

$$\frac{1}{n} \log \frac{P_{Y^n|X^n}(Y^n|X^n)}{P_{Y^n}(Y^n)} = \frac{1}{n} \log \frac{\prod_{i=1}^n P_{Y|X}(Y_i|X_i)}{\prod_{i=1}^n P_Y(Y_i)} \quad (113)$$

$$= \frac{1}{n} \sum_{i=1}^n \log \frac{P_{Y|X}(Y_i|X_i)}{P_Y(Y_i)} \quad (114)$$

となる. また,

$$d_n(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^n d(x_i, y_i) \quad (115)$$

とおいたので,

$$\frac{1}{n} d_n(X^n, Y^n) = \frac{1}{n} \sum_{i=1}^n d(X_i, Y_i) \quad (116)$$

が成り立つ. 一方, 式 (114), (116) の右辺の和の中の各項は独立でその期待値は

$$\mathbb{E} \left[\log \frac{P_{Y|X}(Y_i|X_i)}{P_Y(Y_i)} \right] = \sum_{x,y} P_{XY}(x, y) \log \frac{P_{Y|X}(y|x)}{P_Y(y)} \quad (117)$$

$$= I(X; Y), \quad (i = 1, 2, \dots) \quad (118)$$

$$\mathbb{E}[d(X_i, Y_i)] = \mathbb{E}[d(X, Y)], \quad (i = 1, 2, \dots) \quad (119)$$

であり, しかもその分散は有界である. 証明は後述する. $T_n^{(1)}$, $T_n^{(2)}$ の定義 (111), (112) に注意して Chebyshev の不等式を適用すれば

$$\Pr\{X^n Y^n \in T_n^{(1)}\} \rightarrow 1 \quad (n \rightarrow \infty) \quad (120)$$

$$\Pr\{X^n Y^n \in T_n^{(2)}\} \rightarrow 1 \quad (n \rightarrow \infty) \quad (121)$$

を得る. したがって,

$$\Pr\{X^n Y^n \in T_n\} = 1 - \Pr\{X^n Y^n \notin T_n\} \quad (122)$$

$$\geq 1 - \Pr\{X^n Y^n \notin T_n^{(1)}\} - \Pr\{X^n Y^n \notin T_n^{(2)}\} \quad (123)$$

$$= 1 - \{1 - \Pr\{X^n Y^n \in T_n^{(1)}\}\} - \{1 - \Pr\{X^n Y^n \in T_n^{(2)}\}\} \quad (124)$$

$$\rightarrow 1 \quad (n \rightarrow \infty) \quad (125)$$

が成立する. ここで

$$\hat{W}^n(\mathbf{x}|\mathbf{y}) \triangleq \Pr\{X^n = \mathbf{x} | Y^n = \mathbf{y}\} \quad (126)$$

$$\mathcal{B}_n(\mathbf{y}) \triangleq \left\{ \mathbf{x} \left| \frac{1}{n} \log \frac{\hat{W}^n(\mathbf{x}|\mathbf{y})}{P_{X^n}(\mathbf{x})} \leq R_n - \gamma \right. \right\} \quad (127)$$

$$\mathcal{S}_n(\mathbf{y}) \triangleq \{\mathbf{x} \in \mathcal{X}^n | (\mathbf{x}, \mathbf{y}) \in T_n^{(2)}\} \quad (128)$$

とおく． $\hat{W}_n(\mathbf{x}|\mathbf{y}) = \prod_{i=1}^n P_{X|Y}(x_i|y_i)$ となることに注意すると補題 1 より

$$\frac{1}{n} \log M_n \leq R_n \quad (129)$$

$$\Pr\{X^n \notin \mathcal{S}_n(\psi_n(\varphi_n(X^n)))\} \leq \Pr\{X^n \notin \mathcal{B}_n(Y^n)\} + \Pr\{X^n \notin \mathcal{S}(Y^n)\} + e^{-n\gamma} \quad (130)$$

を満たす (φ_n, ψ_n) 符号が存在する． 式 (130) の左辺は式 (112) と式 (128) より

$$\Pr\{X^n \notin \mathcal{S}_n(\psi_n(\varphi_n(X^n)))\} = \Pr\{(X^n, \psi_n(\varphi_n(X^n))) \notin T_n^{(2)}\} \quad (131)$$

$$= \Pr\left\{\left|\frac{1}{n}d_n(X^n, \psi_n(\varphi_n(X^n))) - \mathbb{E}[d(X, Y)]\right| \geq \gamma\right\} \quad (132)$$

となる． 右辺の各項は， 式 (111)， 式 (78)， 式 (77) より，

$$\Pr\{X^n \in \mathcal{B}_n(Y^n)\} = \Pr\left\{\frac{1}{n} \log \frac{\hat{W}_n(X^n|Y^n)}{P_{X^n}(X^n)} \leq I(X; Y) + \gamma\right\} \quad (133)$$

$$= \Pr\left\{\frac{1}{n} \log \frac{P_{X^n|Y^n}(X^n|Y^n)}{P_{X^n}(X^n)} \leq I(X; Y) + \gamma\right\} \quad (134)$$

$$= \Pr\left\{\frac{1}{n} \log \frac{P_{Y^n|X^n}(Y^n|X^n)}{P_{Y^n}(Y^n)} \leq I(X; Y) + \gamma\right\} \quad (135)$$

$$= \Pr\left\{\frac{1}{n} \log \frac{P_{Y^n|X^n}(Y^n|X^n)}{P_{Y^n}(Y^n)} - I(X; Y) \leq \gamma\right\} \quad (136)$$

$$\geq \Pr\left\{\left|\frac{1}{n} \log \frac{P_{Y^n|X^n}(Y^n|X^n)}{P_{Y^n}(Y^n)} - I(X; Y)\right| < \gamma\right\} \quad (137)$$

$$= \Pr\{X^n Y^n \in T_n^{(1)}\} \quad (138)$$

$$\rightarrow 1 \quad (n \rightarrow \infty) \quad (139)$$

$$\therefore \Pr\{X^n \notin \mathcal{B}_n(Y^n)\} \rightarrow 0 \quad (140)$$

となる． ただし式 (136) から式 (137) への不等式では

$$\frac{1}{n} \log \frac{P_{Y^n|X^n}(Y^n|X^n)}{P_{Y^n}(Y^n)} - I(X; Y) \leq \left|\frac{1}{n} \log \frac{P_{Y^n|X^n}(Y^n|X^n)}{P_{Y^n}(Y^n)} - I(X; Y)\right| \quad (141)$$

を用いた． また， 式 (128) と式 (121) であるから，

$$\Pr\{X^n \in \mathcal{S}_n(Y^n)\} = \Pr\{X^n Y^n \in T_n^{(2)}\} \quad (142)$$

$$\rightarrow 1 \quad (n \rightarrow \infty) \quad (143)$$

$$\therefore \Pr\{X^n \notin \mathcal{S}_n(Y^n)\} \rightarrow 0 \quad (n \rightarrow \infty) \quad (144)$$

となる． 以上より，

$$\Pr\{X^n \notin \mathcal{S}_n(\psi_n(\varphi_n(X^n)))\} \rightarrow 0 \quad (n \rightarrow \infty) \quad (145)$$

が導かれる．式 (131), 式 (132) より

$$\begin{aligned} & \Pr \left\{ \left| \frac{1}{n} d_n(X^n, \psi_n(\varphi_n(X^n))) - \mathbb{E}[d(X, Y)] \right| < \gamma \right\} \\ & \leq \Pr \left\{ \frac{1}{n} d_n(X^n, \psi_n(\varphi_n(X^n))) < \mathbb{E}[d(X, Y)] + \gamma \right\} \end{aligned} \quad (146)$$

である．ここで,

$$\mathcal{U}_n \triangleq \left\{ \mathbf{x} \in \mathcal{X}^n \mid \frac{1}{n} d_n(\mathbf{x}, \psi_n(\varphi_n(\mathbf{x}))) < D \right\} \quad (147)$$

と定義すると d_n の期待値は

$$\frac{1}{n} \mathbb{E}[d_n(X^n, \psi_n(\varphi_n(X^n)))] \quad (148)$$

$$= \frac{1}{n} \sum_{\mathbf{x}} P_{X^n}(\mathbf{x}) d_n(\mathbf{x}, \psi_n(\varphi_n(\mathbf{x}))) \quad (149)$$

$$= \frac{1}{n} \sum_{\mathbf{x} \in \mathcal{U}^n} P_{X^n}(\mathbf{x}) d_n(\mathbf{x}, \psi_n(\varphi_n(\mathbf{x}))) + \frac{1}{n} \sum_{\mathbf{x} \notin \mathcal{U}^n} P_{X^n}(\mathbf{x}) d_n(\mathbf{x}, \psi_n(\varphi_n(\mathbf{x}))) \quad (150)$$

となる． $d(x, y)$ の最大値を d_{max} とすると $d(x, y) \leq d_{max}$ であるので, $\frac{1}{n} d_n(\mathbf{x}, \mathbf{y}) \leq d_{max}$ となる．よって式 (150) の各項は

$$\frac{1}{n} \sum_{\mathbf{x} \in \mathcal{U}^n} P_{X^n}(\mathbf{x}) d_n(\mathbf{x}, \psi_n(\varphi_n(\mathbf{x}))) < \sum_{\mathbf{x} \in \mathcal{U}^n} P_{X^n}(\mathbf{x}) D \quad (151)$$

$$\leq D \quad (152)$$

$$\frac{1}{n} \sum_{\mathbf{x} \notin \mathcal{U}^n} P_{X^n}(\mathbf{x}) d_n(\mathbf{x}, \psi_n(\varphi_n(\mathbf{x}))) \leq \sum_{\mathbf{x} \notin \mathcal{U}^n} P_{X^n}(\mathbf{x}) d_{max} \quad (153)$$

$$= d_{max} \Pr \left\{ \frac{1}{n} d_n(X^n, \psi_n(\varphi_n(X^n))) \leq D \right\} \quad (154)$$

$$\rightarrow 0 \quad (n \rightarrow \infty) \quad (155)$$

したがって, $R = I(X; Y) + 2\gamma$ は D -達成可能である．また, 条件 $\mathbb{E}[d(X, Y)] \leq D - \gamma$ を満たすような任意の Y を選んだので

$$R(D) \leq \inf_{Y: \mathbb{E}[d(X, Y)] \leq D - \gamma} I(X; Y) + 2\gamma \quad (156)$$

となる．さらに, $\gamma > 0$ は任意だったので $\gamma \rightarrow 0$ とすれば

$$\text{さらに } R(D) \leq \inf_{Y: \mathbb{E}[d(X, Y)] \leq D} I(X; Y) \quad (157)$$

$I(X; Y)$ は D に関して連続で, Y の条件 $\mathbb{E}[d(X, Y)]$ は閉じているので

$$R \leq \min_{Y: \mathbb{E}[d(X, Y)] \leq D} I(X; Y) \quad (158)$$

が導け, 順定理が示せた. □

続いて逆定理の証明を行う。

証明 R が D -達成可能であるとすれば

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \mathbb{E}[d_n(X^n, \psi_n(\varphi_n(X^n)))] \leq D \quad (159)$$

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log M_n \leq R \quad (160)$$

なる符号器と復号器の組 (φ_n, ψ_n) が存在する。そこで、

$$Y^n \triangleq (Y_1^{(n)}, Y_2^{(n)}, \dots, Y_n^{(n)}) \quad (161)$$

$$= \psi_n(\varphi_n(X^n)) \quad (162)$$

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

$$\log M_n \geq H(Y^n) \quad (163)$$

$$\geq I(X^n; Y^n) \quad (164)$$

$$= H(X^n) - H(X^n | Y^n) \quad (165)$$

$$= \sum_{i=1}^n H(X_i) - \sum_{i=1}^n H(X_i | Y^n X^{i-1}) \quad (166)$$

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、 $X^{i-1} = (X_1, X_2, \dots, X_{i-1})$ とおいた。そこで、

$$H(X_i | Y^n X^{i-1}) \leq H(X_i | Y_i^{(n)}) \quad (167)$$

に注意すれば

$$\log M_n \geq \sum_{i=1}^n H(X_i) - \sum_{i=1}^n H(X_i | Y_i^{(n)}) \quad (168)$$

$$= \sum_{i=1}^n I(X_i; Y_i^{(n)}) \quad (169)$$

を得る。ゆえに、

$$\frac{1}{n} \log M_n \geq \frac{1}{n} \sum_{i=1}^n I(X_i; Y_i^{(n)}) \quad (170)$$

である。そこで、確率変数 $Q_n \in \{1, 2, \dots, n\}$ 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 $XY^{(n)} = X_i Y_i^{(n)}$ と定義すると、式 (170) は

$$\frac{1}{n} \log M_n \geq \frac{1}{n} \sum_{i=1}^n I(X_i; Y_i^{(n)}) \quad (171)$$

$$= \sum_{i=1}^n P_{Q_n}(i) I(X_i; Y_i^{(n)}) \quad (172)$$

$$= \sum_{i=1}^n P_{Q_n}(i) I(X; Y^{(n)} | Q_n = i) \quad (173)$$

$$= I(X; Y^{(n)} | Q_n) \quad (174)$$

と書ける．ところが、情報源の定常性により Q_n と X は独立なので

$$I(X; Y^{(n)} | Q_n) = \sum_{x,y,i} P_{XY^{(n)}Q_n}(x, y, i) \log \frac{P_{XY^{(n)}|Q_n}(x, y|i)}{P_{X|Q_n}(x|i)P_{Y^{(n)}|Q_n}(y|i)} \quad (175)$$

$$= \sum_{x,y,i} P_{XY^{(n)}Q_n}(x, y, i) \log \frac{P_{XY^{(n)}|Q_n}(x, y|i)}{P_X(x)P_{Y^{(n)}|Q_n}(y|i)} \quad (176)$$

$$= I(X; Q_n Y^{(n)}) \quad (177)$$

$$\geq I(X; Y^{(n)}) \quad (178)$$

となり、したがって、式 (174) は

$$\frac{1}{n} \log M_n \geq I(X; Y^{(n)}) \quad (179)$$

となる．一方、

$$\frac{1}{n} \mathbb{E}[d_n(X^n, \psi_n(\varphi_n(X^n)))] = \frac{1}{n} \mathbb{E}[d_n(X^n, Y^n)] \quad (180)$$

$$= \frac{1}{n} \sum_{i=1}^n \mathbb{E}[d(X_i, Y_i^{(n)})] \quad (181)$$

$$= \sum_{i=1}^n P_{Q_n}(i) \mathbb{E}[d(X_i, Y_i^{(n)})] \quad (182)$$

$$= \sum_{i=1}^n P_{Q_n}(i) \mathbb{E}[d(X, Y^{(n)}) | Q_n = i] \quad (183)$$

$$= \sum_{i=1}^n P_{Q_n}(i) \sum_{x,y} d(x, y) P_{XY^{(n)}|Q_n}(x, y|i) \quad (184)$$

$$= \sum_{x,y} d(x, y) P_{XY^{(n)}}(x, y) \quad (185)$$

$$= \mathbb{E}[d(X, Y^{(n)})] \quad (186)$$

なので、式 (159) から

$$D \geq \limsup_{n \rightarrow \infty} \mathbb{E}[d(X, Y^{(n)})] \quad (187)$$

したがって、任意に小さい定数 $\gamma > 0$ に対して

$$D + \gamma \geq \mathbb{E}[d(X, Y^{(n)})] \quad (\forall n \geq n_0) \quad (188)$$

すると、式 (179) により

$$\frac{1}{n} \log M_n \geq \min_{Y: \mathbb{E}[d(X, Y)] \leq D + \gamma} I(X; Y) \quad (189)$$

となる。よって、式 (160) から

$$R \geq \limsup_{n \rightarrow \infty} \frac{1}{n} \log M_n \quad (190)$$

$$\geq \limsup_{n \rightarrow \infty} \min_{Y: \mathbb{E}[d(X, Y)] \leq D + \gamma} I(X; Y) \quad (191)$$

$$= \min_{Y: \mathbb{E}[d(X, Y)] \leq D + \gamma} I(X; Y) \quad (192)$$

が得られる。ところが、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

$$R \geq \min_{Y: \mathbb{E}[d(X, Y)] \leq D} I(X; Y) \quad (193)$$

が結論される。よって、順定理と逆定理より

$$R(D) = \min_{Y: \mathbb{E}[d(X, Y)] \leq D} I(X; Y) \quad (194)$$

を満たす。 $\square$

A.4 分散の一様有界性

証明中に用いた分散の一様有界性を示す。 [[1] 注意 3.1]

証明 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 (114) の和の中を

$$Z = \log \frac{P_{Y|X}(Y|X)}{P_Y(Y)} \quad (195)$$

とおく。まず Z の期待値 $\mathbb{E}[Z]$ 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

$$\mathbb{E}[Z] = I(X; Y) \quad (196)$$

$$= H(Y) - H(Y|X) \quad (197)$$

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 Y が有限であるとする．すると， $H(Y|X) \geq 0$ に注意すると

$$0 \leq \mathbb{E}[Z] \leq H(Y) \leq \log|\mathcal{Y}| \quad (198)$$

が得られる．すなわち $\mathbb{E}[Z]$ は一様に有界である．次に Z の分散 $\text{Var}[Z]$ は

$$\text{Var}[Z] = \mathbb{E}[Z^2] - (\mathbb{E}[Z])^2 \quad (199)$$

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

$$\mathbb{E}[Z^2] = \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} P_{XY}(x, y) \left(\log \frac{P_{Y|X}(y|x)}{P_Y(y)} \right)^2 \quad (200)$$

$$\begin{aligned} &= \sum_{(x, y) \in \mathcal{X} \times \mathcal{Y}: P_{Y|X}(y, x) \leq P_Y(y)} P_{XY}(x, y) \left(\log \frac{P_{Y|X}(y|x)}{P_Y(y)} \right)^2 \\ &\quad + \sum_{(x, y) \in \mathcal{X} \times \mathcal{Y}: P_{Y|X}(y, x) \geq P_Y(y)} P_{XY}(x, y) \left(\log \frac{P_{Y|X}(y|x)}{P_Y(y)} \right)^2 \end{aligned} \quad (201)$$

$$\begin{aligned} &= \sum_{(x, y) \in \mathcal{X} \times \mathcal{Y}: P_{Y|X}(y|x) \leq P_Y(y)} P_{XY}(x, y) \left(\log \frac{P_Y(y)}{P_{Y|X}(y|x)} \right)^2 \\ &\quad + \sum_{(x, y) \in \mathcal{X} \times \mathcal{Y}: P_{Y|X}(y, x) \geq P_Y(y)} P_{XY}(x, y) \left(\log \frac{P_{Y|X}(y|x)}{P_Y(y)} \right)^2 \end{aligned} \quad (202)$$

$$\begin{aligned} &\leq \sum_{(x, y) \in \mathcal{X} \times \mathcal{Y}: P_{Y|X}(y|x) \leq P_Y(y)} P_{XY}(x, y) \left(\log \frac{1}{P_{Y|X}(y|x)} \right)^2 \\ &\quad + \sum_{(x, y) \in \mathcal{X} \times \mathcal{Y}: P_{Y|X}(y, x) \geq P_Y(y)} P_{XY}(x, y) \left(\log \frac{1}{P_Y(y)} \right)^2 \end{aligned} \quad (203)$$

$$\begin{aligned} &\leq \sum_{(x, y) \in \mathcal{X} \times \mathcal{Y}} P_{XY}(x, y) \left(\log \frac{1}{P_{Y|X}(y|x)} \right)^2 \\ &\quad + \sum_{(x, y) \in \mathcal{X} \times \mathcal{Y}} P_{XY}(x, y) \left(\log \frac{1}{P_Y(y)} \right)^2 \end{aligned} \quad (204)$$

$$\begin{aligned} &= \sum_{x \in \mathcal{X}} P_X(x) \sum_{y \in \mathcal{Y}} P_{Y|X}(y|x) \left(\log \frac{1}{P_{Y|X}(y|x)} \right)^2 \\ &\quad + \sum_{y \in \mathcal{Y}} P_Y(y) \left(\log \frac{1}{P_Y(y)} \right)^2 \end{aligned} \quad (205)$$

と変形する．そこで，関数

$$g(u) = u \left(\log \frac{1}{u} \right)^2 \quad (0 \leq u \leq 1) \quad (206)$$

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

$$h(t) = t^2 e^{-t} \quad (0 \leq t \leq +\infty) \quad (207)$$

と書ける．すると， $h(t)$ は $t = 2$ で最大値をとり

$$\max_{0 \leq u \leq 1} g(u) = \max_{0 \leq t \leq +\infty} h(t) = \frac{4}{e^2} \quad (208)$$

となる．したがって，式 (208) を式 (205) に適用すると

$$\mathbb{E}[Z^2] \leq \sum_{x \in \mathcal{X}} P_X(x) \sum_{y \in \mathcal{Y}} \frac{4}{e^2} + \sum_{y \in \mathcal{Y}} P_Y(y) \frac{4}{e^2} \quad (209)$$

$$\leq \sum_{x \in \mathcal{X}} P_X(x) \frac{4|\mathcal{Y}|}{e^2} + \frac{4|\mathcal{Y}|}{e^2} \quad (210)$$

$$\leq \frac{8|\mathcal{Y}|}{e^2} \quad (211)$$

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 5 章の数値計算のソースコード

```
#include <stdio.h>
#include <math.h>

double s;
double d[2][2] = { {0, 1}, {1, 0} };

int ratedistortion(double s, double p, double e, double e2) {

    double P[2] = { p, 1.0 - p };
    double Q1[2][2] = { {1.0 - e, e}, {e, 1.0 - e} };
    double Q2[2][2] = { {1.0 - e2, e2}, {e2, 1.0 - e2} };
    double b4[2] = { 0, 0 };

    double d2[2][2] = { {0, 0}, {0, 0} };

    //d*の分母 (P[x*])
    for (int i = 0; i < 2; i++) {
        for (int j = 0; j < 2; j++) {
            b4[i] += P[j] * Q1[i][j];
        }
    }

    // d*(x*, y*) の計算
    for (int i = 0; i < 2; i++) {
        for (int j = 0; j < 2; j++) {
            for (int k = 0; k < 2; k++) {
                for (int l = 0; l < 2; l++) {
```

```

                d2[i][j] += P[k] * Q1[i][k] * Q2[1][j] * d[k][1] / b4[i];
            }
        }
    }

double a[2][2] = {{pow(2.0, s*d2[0][0]), pow(2.0, s*d2[0][1])}, {pow(2.0, s*d2[1][0]), pow(2.0, s*d2[1][1])}};
double q[2] = { 0.5, 0.5 };
double c[2];
double tu = 0, t1 = 1;

while ((t1 - tu) > (1e-6)) {
    double b1[2] = { 0, 0 };

    // c[k] の分母
    for (int j = 0; j < 2; j++) {
        for (int k = 0; k < 2; k++) {
            b1[j] += q[k] * a[j][k];
        }
    }

    // c[k] の計算
    c[0] = c[1] = 0;
    for (int k = 0; k < 2; k++) {
        for (int j = 0; j < 2; j++) {
            c[k] += b4[j] * a[j][k] / b1[j];
        }
    }

    // q[k] と Tu, T1 の計算
    tu = 0;
    for (int k = 0; k < 2; k++) {
        q[k] *= c[k];
        tu += q[k] * log2(c[k]);
    }
    if (c[0] > c[1]) {
        t1 = log2(c[0]);
    }
    else {
        t1 = log2(c[1]);
    }
}

double Q[2][2] = { {0, 0}, {0, 0} };
double D = 0;
double R = 0;

// Q[k][j] の分母
double b2[2] = { 0, 0 };
for (int j = 0; j < 2; j++) {
    for (int k = 0; k < 2; k++) {
        b2[j] += a[j][k] * q[k];
    }
}

// Q[k][j] の計算
for (int k = 0; k < 2; k++) {
    for (int j = 0; j < 2; j++) {
        Q[k][j] = a[j][k] * q[k] / b2[j];
    }
}

// D の計算
for (int j = 0; j < 2; j++) {
    for (int k = 0; k < 2; k++) {
        D += b4[j] * Q[k][j] * d2[j][k];
    }
}

```

```

// R(D) の計算
double b3 = 0;
for (int j = 0; j < 2; j++) {
    b3 += b4[j] * log2(b2[j]);
}
R = s * D - b3 - tu;

printf("D = %f, R(D) = %f\n", D, R);
return 0;
}

int main(void) {
    printf("input P: ");
    double p;
    scanf_s("%lf", &p);

    printf("input error of Q1: ");
    double e;
    scanf_s("%lf", &e);

    printf("input error of Q2: ");
    double e2;
    scanf_s("%lf", &e2);

    for (double s = -50; s < 0; s += 0.01) {
        ratedistortion(s, p, e, e2);
    }

    return 0;
}

```

付録 C 雑音がある場合のレート・歪み関数の計算

5 章の雑音がある場合のレート・歪みの問題設定で例をあげて計算する。歪みは 5 章と同様にハミング歪みである。与えられる 2 元情報源は $\mathcal{X} = \{0, 1\}$ とし、

$$\Pr\{X = 0\} = 0.4, \Pr\{X = 1\} = 0.6 \quad (212)$$

とする。このもとで、 X から X^* への確率を

$$P_{X^*|X}(x^*|x) = 0.5 \quad (213)$$

とする。以上の条件で $D < 0.4$ のとき任意のレートは達成可能ではないことを示す。ただし、 Y^* から Y への確率は誤りがないものとする。したがって、

$$P_{Y|Y^*}(y|y^*) = \begin{cases} 1 & (y = y^* \text{ のとき}) \\ 0 & (y \neq y^* \text{ のとき}) \end{cases} \quad (214)$$

である。このとき d^* は

$$d^*(x^*, y) = \frac{1}{P_{X^*}(x^*)} \sum_{x,y} P_X(x) P_{X^*|X}(x^*|x) d(x, y) \quad (215)$$

となる．では，任意の Y に対して $\mathbb{E}[d^*(X^*, Y)]$ を求める．

$$\mathbb{E}[d^*(X^*, Y)] \quad (216)$$

$$= \sum_{x^*, y} P_{X^*}(x^*) P_{Y|X^*}(y|x^*) d^*(x^*, y) \quad (217)$$

$$= \sum_{x^*, y} P_{Y|X^*}(y|x^*) \sum_x P_X(x) P_{X^*|X}(x^*|x) d(x, y) \quad (218)$$

$$= \sum_{x^*} (P_{Y|X^*}(0|x^*) P_X(1) P_{X^*|X}(x^*|1) + P_{Y|X^*}(1|x^*) P_X(0) P_{X^*|X}(x^*|0)) \quad (219)$$

$$= \frac{1}{2} \sum_{x^*} (P_{Y|X^*}(0|x^*) P_X(1) + P_{Y|X^*}(1|x^*) P_X(0)) \quad (220)$$

$$= \frac{1}{2} \sum_{x^*} (0.6 \cdot P_{Y|X^*}(0|x^*) + 0.4 \cdot P_{Y|X^*}(1|x^*)) \quad (221)$$

$$= \frac{1}{2} \sum_{x^*} (0.6 \cdot P_{Y|X^*}(0|x^*) + 0.4 \cdot P_{Y|X^*}(1|x^*)) \quad (222)$$

$$= \frac{1}{2} \sum_{x^*} (0.2 \cdot P_{Y|X^*}(0|x^*) + 0.4 \cdot P_{Y|X^*}(0|x^*) + 0.4 \cdot P_{Y|X^*}(1|x^*)) \quad (223)$$

$$= \frac{1}{2} \sum_{x^*} (0.2 \cdot P_{Y|X^*}(0|x^*) + 0.4) \quad (224)$$

$$\geq \frac{1}{2} \sum_{x^*} 0.4 \quad (225)$$

$$= 0.4 \quad (226)$$

が導かれた．よって， $\mathbb{E}[d^*(X^*, Y)] \geq 0.4$ が示されたので，雑音が入る確率 $P_{X^*|X}(x^*|x) = 0.5$ のもとで $D < 0.4$ のとき任意のレートは達成可能でない． $\square$