

信州大学工学部

学士論文

観測値に雑音のあるシステムにおける
ガウス型情報源のレート・歪み関数

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 17T2089A
氏名 征矢 愛耶乃

2021 年 2 月 16 日

目次

1	はじめに	1
2	連続な情報源のレート・歪み理論	1
2.1	観測値に雑音のないシステムの一般的な連続値情報源のレート・歪み理論 . . .	2
2.2	観測値に雑音のないシステムのガウス型情報源のレート・歪み理論	3
2.3	観測値に雑音のあるシステムの一般的な連続値情報源のレート・歪み理論 . . .	5
3	観測値に雑音のあるシステムのガウス型情報源のレート・歪み理論	9
3.1	レート・歪み関数の計算	9
3.2	得られた結果と考察	13
4	まとめ	17
	謝辞	17
	参考文献	17
	付録 A 定理 1 の証明	18
A.1	数学的準備	18
A.2	補題 1	19
A.3	定理 1 の証明	21
	付録 B 微分エントロピーの最大化定理	28
	付録 C 通常の相互情報量と定数倍された変数を持つ相互情報量の関係	29

1 はじめに

実数値などを出力する連続値情報源の場合、1 個の任意の実数を記述するためには無限長のビットが必要であり、有限長のビットでは完全には表現できない。例えば携帯電話のメールなどで画像を送信するとき、ファイルサイズを小さくするために意図的にデータの一部を削除して送信する。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。この差異を元の画像との歪みという。よって、効率よくデータ通信を行うためには元のデータ x と復号されたデータ 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ことが問題になる。これがレート・歪み問題である [1][7]。

一方、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、モノと受信機間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。

本研究では、上記で述べた 2 つの問題に着目し、観測値に雑音のあるシステムのレート・歪み問題を考える。離散値情報源の場合の結果はすでに知られているが、本研究では連続値情報源を対象とする。問題設定を一般化するために観測値だけでなく、復号器と受信者の間にも雑音が存在すると仮定する。そのもとで、連続値情報源に対する一般的なレート・歪み関数の定理を証明し、さらに重要な具体例としてガウス型情報源に対するレート・歪み関数を導出する。

本論文は次のような章から構成される。2 章では 観測値に雑音のないシステムの一般的な連続値情報源のレート・歪み理論と、ガウス型情報源のレート・歪み理論、観測値に雑音のあるシステムの一般的な連続値情報源のレート・歪み理論について述べる [2][6][7]。3 章では観測値に雑音のあるシステムのガウス型情報源のレート・歪み関数を計算し、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 4 章で本論文をまとめる。

2 連続な情報源のレート・歪み理論

本章では連続値情報源の観測値に雑音のない場合と雑音がある場合のレート・歪み理論について述べる。

2.1 観測値に雑音のないシステムの一般的な連続値情報源のレート・歪み理論

観測値に雑音のない通常のレート・歪み理論について従来研究の結果 [1] を確認する．特に，情報源は連続値を取ると仮定する． $\mathbf{X} = \{X^n = (X_1, X_2, \dots, X_n)\}_{n=1}^{\infty}$ 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 $\mathbf{y} = (y_1, y_2, \dots, y_n)$ の間の歪み $d_n(\mathbf{x}, \mathbf{y})$ 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，シンボル 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，レートと歪みの達成可能性を次のように定義する．

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．

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

ここで，確率密度 μ をもつ確率変数 X, Y に対して以下の量を定義する．連続値情報源のエントロピー $h(X)$ と $h(X|Y)$ を

$$h(X) = - \int_{\mathcal{X}} \mu_X(x) \log \mu_X(x) dx \quad (6)$$

$$h(X|Y) = - \iint_{\mathcal{X}, \mathcal{Y}} \mu_X(x) \mu_{X|Y}(x|y) \log \mu_{X|Y}(x|y) dx dy \quad (7)$$

と定める．一方， X と Y の間の相互情報量 $I(X; Y)$ を

$$I(X; Y) \triangleq \iint_{x, y} \mu_{XY}(x, y) \log \frac{\mu_{Y|X}(y|x)}{\mu_Y(y)} dx dy \quad (8)$$

と定義する．すると，相互情報量 $I(X; Y)$ はエントロピー h を用いて

$$I(X; Y) \triangleq h(X) - h(X|Y) \quad (9)$$

と書ける．上記の定義のもとで次の定理が成立する．

定理 1 ([1]) 分布 μ_X に従う連続値の場合の定常無記憶情報源 X のレート・歪み関数は

$$R_0(D) = \min_{Y: \mathbb{E}[d(X, Y)] \leq D} I(X; Y) \quad (10)$$

で与えられる．

付録 A に証明を掲載する．

2.2 観測値に雑音のないシステムのガウス型情報源のレート・歪み理論

定理 1 の特別な場合として，ガウス型情報源のレート・歪み関数を確認する．シンボル単位
の歪みとして二乗誤差歪みを用いる．すなわち

$$d(x, y) \triangleq (x - y)^2 \quad (11)$$

と定義する．このとき，レート・歪み関数は次のように表される．

定理 2 ([7]) 二乗誤差歪み $\mathcal{N}(0, \sigma^2)$ に基づく情報源に対するレート・歪み関数は

$$R(D) = \begin{cases} \frac{1}{2} \log \frac{\sigma^2}{D} & (0 \leq D \leq \sigma^2) \\ 0 & (D > \sigma^2) \end{cases} \quad (12)$$

で与えられる．

証明 定理 1 より，連続値におけるレート・歪み関数は

$$R_0(D) = \min_{Y: \mathbb{E}[d(X, Y)] \leq D} I(X; Y) \quad (13)$$

で与えられる． $\mathbb{E}(X - Y)^2 \leq D$ であるような，任意の Y を考える．ここでガウス型情報源

であったことに注意すると、 X と Y の間の相互情報量の定義より、

$$I(X; Y) = h(X) - h(X|Y) \quad (14)$$

$$= \frac{1}{2} \log(2\pi e) \sigma^2 - h(X - Y|Y) \quad (15)$$

$$= \frac{1}{2} \log(2\pi e) \sigma^2 - h(X - Y) \quad (16)$$

$$\geq \frac{1}{2} \log(2\pi e) \sigma^2 - \frac{1}{2} \log(2\pi e) - h(\mathcal{N}(0, \mathbb{E}[(X - Y)^2])) \quad (17)$$

$$= \frac{1}{2} \log(2\pi e) \sigma^2 - \frac{1}{2} \log(2\pi e) \mathbb{E}[(X - Y)^2] \quad (18)$$

$$\geq \frac{1}{2} \log(2\pi e) \sigma^2 - \frac{1}{2} \log(2\pi e) D \quad (19)$$

$$= \frac{1}{2} \log \frac{\sigma^2}{D} \quad (20)$$

となる．ここで、式 (16) は条件付けることによってエントロピーが減ることから、式 (17) では与えられた 2 次モーメントのもとで正規分布がエントロピーを最大化している．これについては付録 B にて証明する．したがって、

$$R(D) \geq \frac{1}{2} \log \frac{\sigma^2}{D} \quad (21)$$

が成り立つ．

次に式 (21) の下界が達成されることを確かめる．図 1 に示すような X, Y, Z を考える． Y と Z は独立である．このとき $\mathbb{E}[d(X, Y)] = \mathbb{E}[Z^2] = D$ であることに注意する．

$D \leq \sigma^2$ の場合を考える．このとき X と Y の相互情報量 $I(X; Y)$ は

$$I(X; Y) = h(X) - h(X|Y) \quad (22)$$

$$= \frac{1}{2} \log(2\pi e) \sigma^2 - h(X - Y|Y) \quad (23)$$

$$= \frac{1}{2} \log(2\pi e) \sigma^2 - h(Z|Y) \quad (24)$$

$$= \frac{1}{2} \log(2\pi e) \sigma^2 - h(Z) \quad (25)$$

$$= \frac{1}{2} \log(2\pi e) \sigma^2 - \frac{1}{2} \log(2\pi e) D \quad (26)$$

$$= \frac{1}{2} \log \frac{\sigma^2}{D} \quad (27)$$

となる． $I(X; Y)$ の最小値を考えると

$$\frac{1}{2} \log \frac{\sigma^2}{D} \geq \min_{Y: \mathbb{E}[d(X, Y)] \leq D} I(X; Y) \quad (28)$$

$$= R(D) \quad (29)$$

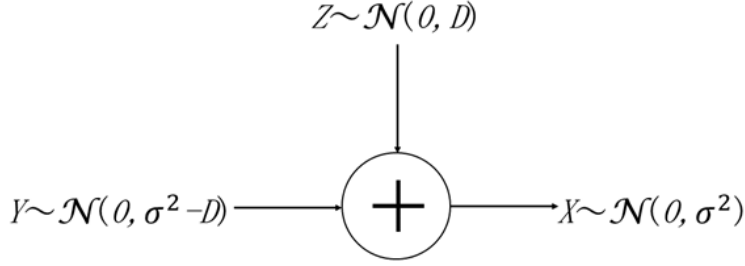


図1 ガウス型情報源に対する同時分布

が成り立つ． $\sigma^2 < D$ の場合を考える． Y を定数 0 とおくと

$$\mathbb{E}[d(X, Y)] = \mathbb{E}[(X - Y)^2] \quad (30)$$

$$= \mathbb{E}[X^2] \quad (31)$$

$$= \sigma^2 < D \quad (32)$$

となり，また X と Y は互いに独立なので

$$I(X; Y) = 0 \quad (33)$$

となる．また $I(X; Y)$ は常に非負であるから

$$\min_{Y: \mathbb{E}[d(X, Y)] \leq D} I(X; Y) = 0 \quad (34)$$

になる．よって

$$R(D) = 0 \quad (35)$$

が成り立つ．式 (29)，式 (35) より

$$R(D) = \begin{cases} \frac{1}{2} \log \frac{\sigma^2}{D} & (0 \leq D \leq \sigma^2) \\ 0 & (D > \sigma^2) \end{cases} \quad (36)$$

となる． $\square$

2.3 観測値に雑音のあるシステムの一般的な連続値情報源のレート・歪み理論

情報源を観測するときに正確には観測できずいくらかノイズを含むと仮定する．さらに，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．離散的な情報源の場合の結果は知られているが [4]，本論文では連続値情報源の場合でも同様の結果が得られることを示す．

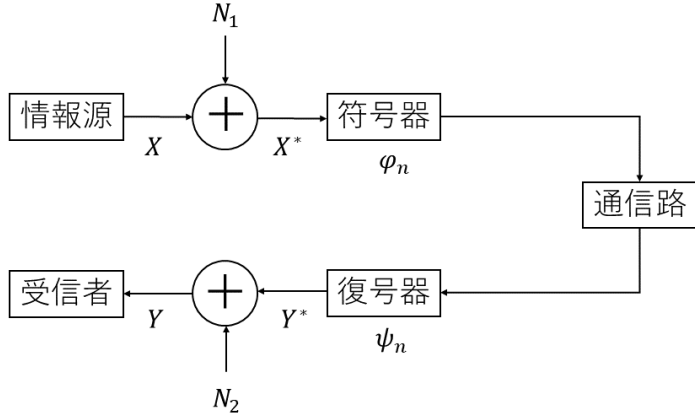


図2 雑音がある場合の通信システムのモデル

$X^n = (X_1, X_2, \dots, X_n)$ 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 $\mu_{X^*|X}(x^*|x)$ 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 ψ_n で復号され、 Y^* となる。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 $\mu_{Y|Y^*}(y|y^*)$ 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2に示す。このとき x と y の間の歪みを $d(x, y) (x \in \mathcal{X}, y \in \mathcal{Y})$ とする。長さ n のデータの歪み d_n は式(1)によって定義する。

達成可能性は定義1にしたがって定義する。

定義3 ([6]) 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。

$$R(D) \triangleq \inf \{ R \mid \text{歪み測度 } d \text{ のもとでレート } R \text{ は } D \text{ -達成可能} \} \quad (37)$$

定理を述べる準備として $d^*(x^*, y^*) (x^* \in \mathcal{X}, y^* \in \mathcal{Y})$ を

$$d^*(x^*, y^*) \triangleq \frac{1}{\mu_{X^*}(x^*)} \iint_{x, y} \mu_X(x) \mu_{X^*|X}(x^*|x) \mu_{Y|Y^*}(y|y^*) d(x, y) dx dy \quad (38)$$

と定義する。

定理3 雑音がある場合のレート・歪み関数は

$$R(D) = \min_{Y^*: \mathbb{E}[d^*(X^*, Y^*)] \leq D} I(X^*; Y^*) \quad (39)$$

で与えられる。ここで、 X^* は $\mu_{X^*}(x^*) = \int_x \mu_X(x) \mu_{X^*|X}(x^*|x) dx$ に従い、 Y^* は X^* と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

$$d_n^*(\mathbf{x}^*, \mathbf{y}^*) \triangleq \sum_{i=1}^n d^*(x_i^*, y_i^*) \quad (40)$$

とする． d_n^* の期待値は

$$\mathbb{E}[d_n^*(X^{*n}, Y^{*n})] = \iint_{\mathbf{x}^*, \mathbf{y}^*} \mu_{X^*}(\mathbf{x}^*) \mu_{Y^*|X^*}(\mathbf{y}^*|\mathbf{x}^*) d_n^*(\mathbf{x}^*, \mathbf{y}^*) d\mathbf{x}^* d\mathbf{y}^* \quad (41)$$

$$= \iint_{\mathbf{x}^*, \mathbf{y}^*} \left\{ \prod_{j=1}^n \mu_{X^*}(x_j^*) \right\} \mu_{Y^*|X^*}(\mathbf{y}^*|\mathbf{x}^*) d\mathbf{x}^* d\mathbf{y}^* \bullet \sum_{k=1}^n d(x_k^*, y_k^*) \quad (42)$$

$$= \sum_{k=1}^n d(x_k^*, y_k^*) \iint_{\mathbf{x}^*, \mathbf{y}^*} \left\{ \prod_{j=1}^n \mu_{X^*}(x_j^*) \right\} \mu_{Y^*|X^*}(\mathbf{y}^*|\mathbf{x}^*) d\mathbf{x}^* d\mathbf{y}^* \quad (43)$$

と書ける．一方， d_n の期待値は

$$\mathbb{E}[d_n(X^n, Y^n)] \quad (44)$$

$$= \iiint_{\mathbf{x}, \mathbf{x}^*, \mathbf{y}^*, \mathbf{y}} \mu_X(\mathbf{x}) \mu_{X^*|X}(\mathbf{x}^*|\mathbf{x}) \mu_{Y^*|X^*}(\mathbf{y}^*|\mathbf{x}^*) \mu_{Y|Y^*}(\mathbf{y}|\mathbf{y}^*) d_n(\mathbf{x}, \mathbf{y}) d\mathbf{x} d\mathbf{x}^* d\mathbf{y}^* d\mathbf{y} \quad (45)$$

$$= \sum_{k=1}^n d(x_k, y_k) \times \iiint_{\mathbf{x}, \mathbf{x}^*, \mathbf{y}^*, \mathbf{y}} \left\{ \prod_{j=1}^n \mu_X(x_j) \mu_{X^*|X}(x_j^*|x_j) \mu_{Y|Y^*}(y_j|y_j^*) \right\} \mu_{Y^*|X^*}(\mathbf{y}^*|\mathbf{x}^*) d\mathbf{x} d\mathbf{x}^* d\mathbf{y}^* d\mathbf{y} \quad (46)$$

$$= \sum_{k=1}^n \iint_{\mathbf{x}^*, \mathbf{y}^*} \mu_{Y^*|X^*}(\mathbf{y}^*|\mathbf{x}^*) d\mathbf{x}^* d\mathbf{y}^* \times \iint_{\mathbf{x}, \mathbf{y}} d(x_k, y_k) \prod_{j=1}^n \mu_X(x_j) \mu_{X^*|X}(x_j^*|x_j) \mu_{Y|Y^*}(y_j|y_j^*) d\mathbf{x} d\mathbf{y} \quad (47)$$

$$= \sum_{k=1}^n \iint_{\mathbf{x}^*, \mathbf{y}^*} \mu_{Y^*|X^*}(\mathbf{y}^*|\mathbf{x}^*) d\mathbf{x}^* d\mathbf{y}^* \times \iint_{x_k, y_k} d(x_k, y_k) \mu_X(x_k) \mu_{X^*|X}(x_k^*|x_k) \mu_{Y|Y^*}(y_k|y_k^*) d\mathbf{x} d\mathbf{y} \times \iint_{\substack{x_1, \dots, x_n \\ y_1, \dots, y_n \\ x_k, y_k \text{ 除く } j \neq k}} \prod_{j=1}^n \mu_X(x_j) \mu_{X^*|X}(x_j^*|x_j) \mu_{Y|Y^*}(y_j|y_j^*) d\mathbf{x} d\mathbf{y} \quad (48)$$

であり, $j \neq k$ の部分について考えると

$$\iint_{\substack{x_1, \dots, x_n \\ y_1, \dots, y_n \\ x_k, y_k \text{ 除く}}} \prod_{\substack{j=1 \\ j \neq k}}^n \mu_X(x_j) \mu_{X^*|X}(x_j^*|x_j) \mu_{Y|Y^*}(y_j|y_j^*) d\mathbf{x} d\mathbf{y} \quad (49)$$

$$= \iint_{x_1, y_1} \cdots \iint_{x_n, y_n} \{ \mu_X(x_1) \mu_{X^*|X}(x_1^*|x_1) \mu_{Y|Y^*}(y_1|y_1^*) \cdots \mu_X(x_n) \mu_{X^*|X}(x_n^*|x_n) \mu_{Y|Y^*}(y_n|y_n^*) \} d\mathbf{x} d\mathbf{y} \quad (k \text{ 除く}) \quad (50)$$

$$= \int_{x_1} \cdots \int_{x_n} \mu_X(x_1) \mu_{X^*|X}(x_1^*|x_1) \cdots \mu_X(x_n) \mu_{X^*|X}(x_n^*|x_n) d\mathbf{x} \quad (k \text{ 除く}) \quad (51)$$

$$= \left(\int_{x_1} \mu_X(x_1) \mu_{X^*|X}(x_1^*|x_1) d\mathbf{x} \right) \cdots \left(\int_{x_n} \mu_X(x_n) \mu_{X^*|X}(x_n^*|x_n) d\mathbf{x} \right) \quad (k \text{ 除く}) \quad (52)$$

$$= \prod_{\substack{j=1 \\ j \neq k}}^n \mu_{X^*}(x_j^*) \quad (53)$$

となり, 式 (38), 式 (48), 式 (53) より

$$\mathbb{E}[d_n(X^n, Y^n)] \quad (54)$$

$$= \sum_{k=1}^n \iint_{\mathbf{x}^*, \mathbf{y}^*} \mu_{Y^*|X^*}(\mathbf{y}^*|\mathbf{x}^*) d\mathbf{x}^* d\mathbf{y}^* \times \left\{ \iint_{x_k, y_k} d(x_k, y_k) \mu_X(x_k) \mu_{X^*|X}(x_k^*|x_k) \mu_{Y|Y^*}(y_k|y_k^*) d\mathbf{x} d\mathbf{y} \right\} \prod_{j \neq k} \mu_{X^*}(x_j^*) \quad (55)$$

$$= \sum_{k=1}^n \iint_{\mathbf{x}^*, \mathbf{y}^*} d(x_k^*, y_k^*) \left\{ \prod_{j=1}^n \mu_{X^*}(x_j^*) \right\} \mu_{Y^*|X^*}(\mathbf{y}^*|\mathbf{x}^*) d\mathbf{x}^* d\mathbf{y}^* \quad (56)$$

となる. したがって, 式 (43) より

$$\mathbb{E}[d_n(X^n, Y^n)] = \mathbb{E}[d_n^*(X^{*n}, Y^{*n})] \quad (57)$$

が導かれる. すると, 定義 1 の式 (3) は

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \mathbb{E}[d_n^*(X^{*n}, Y^{*n})] \leq D \quad (58)$$

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,

$$R(D) = \inf \{ R | \text{歪み測度 } d \text{ のもとでレート } R \text{ は } D\text{-達成可能} \} \quad (59)$$

$$= \inf \{ R | \text{歪み測度 } d^* \text{ のもとでレート } R \text{ は } D\text{-達成可能} \} \quad (60)$$

$$= \min_{Y^*: \mathbb{E}[d^*(X^*, Y^*)] \leq D} I(X^*; Y^*) \quad (61)$$

となる. よって定理 3 が導かれた. $\square$

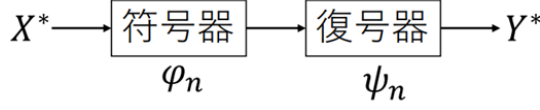


図 3 定常無記憶情報源の通信システム

定理 3 より連続値における雑音がある場合のレート・歪みの問題は図 3 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 $\mu_{X^*}(x^*) = \int_x \mu_X(x) \mu_{X^*|X}(x^*|x) dx$ に従う確率変数である．

3 観測値に雑音のあるシステムのガウス型情報源のレート・歪み理論

3.1 レート・歪み関数の計算

定理 3 の特別な場合として，ガウス型情報源に対し，図 2 のような雑音のあるシステムのレート・歪み関数を計算する．情報源 X は $X \sim \mathcal{N}(0, \sigma_x^2)$ ，符号器に入る前に付加される雑音 N_1 は $N_1 \sim \mathcal{N}(0, \sigma_1^2)$ ，受信者に届く前に付加される雑音 N_2 は $N_2 \sim \mathcal{N}(0, \sigma_2^2)$ に従うとする．歪みとして式 (11) で定義される二乗誤差歪みを用いる．このとき X はガウス型情報源であるので確率密度関数は

$$\mu_X(x) = f_{\sigma_x^2}(x) \triangleq \frac{1}{\sqrt{2\pi\sigma_x^2}} \exp\left(-\frac{x^2}{2\sigma_x^2}\right) \quad (62)$$

となる．

Y^* で条件付けられた Y の密度関数 $\mu_{Y|Y^*}(y|y^*)$ について考える． $\mu_{Y|Y^*}(y|y^*)$ を式変形していくと

$$\mu_{Y|Y^*}(y|y^*) = \Pr\{Y = y | Y^* = y^*\} \quad (63)$$

$$= \Pr\{Y^* + N_2 = y | Y^* = y^*\} \quad (64)$$

$$= \Pr\{N_2 = y - y^* | Y^* = y^*\} \quad (65)$$

$$= \Pr\{N_2 = y - y^*\} \quad (66)$$

$$= \mu_{N_2}(y - y^*) \quad (67)$$

$$= f_{\sigma_2^2}(y - y^*) \quad (68)$$

となる．これと同様にして X で条件付けられた X^* の密度関数 $\mu_{X^*|X}(x^*|x)$ は

$$\mu_{X^*|X}(x^*|x) = f_{\sigma_1^2}(x^* - x) \quad (69)$$

となる.

X^* の密度関数 $\mu_{X^*}(x^*)$ を求めると,

$$\mu_{X^*}(x^*) = \int_x \mu_X(x) \mu_{X^*|X}(x^*|x) dx \quad (70)$$

$$= \int_x f_{\sigma_x^2}(x) f_{\sigma_1^2}(x^* - x) dx \quad (71)$$

$$= \int_x \frac{1}{\sqrt{2\pi\sigma_x^2}} \exp\left(-\frac{x^2}{2\sigma_x^2}\right) \frac{1}{\sqrt{2\pi\sigma_1^2}} \exp\left(-\frac{(x^* - x)^2}{2\sigma_1^2}\right) dx \quad (72)$$

$$= \frac{1}{2\pi\sqrt{\sigma_x^2\sigma_1^2}} \int_x \exp\left(-\frac{x^2}{2\sigma_x^2} - \frac{(x^* - x)^2}{2\sigma_1^2}\right) dx \quad (73)$$

となる. ここで, 式 (73) の指数関数の中について考えると

$$-\frac{x^2}{2\sigma_x^2} - \frac{(x^* - x)^2}{2\sigma_1^2} = -\frac{\sigma_1^2 x^2 + \sigma_x^2 (x^* - x)^2}{2\sigma_x^2 \sigma_1^2} \quad (74)$$

$$(75)$$

となり, さらに分子について考えると

$$\sigma_1^2 x^2 + \sigma_x^2 (x^* - x)^2 \quad (76)$$

$$= \sigma_1^2 x^2 + \sigma_x^2 (x^{*2} - 2x^*x + x^2) \quad (77)$$

$$= (\sigma_1^2 + \sigma_x^2) x^2 + 2\sigma_x^2 x^*x + \sigma_x^2 x^{*2} \quad (78)$$

$$= (\sigma_1^2 + \sigma_x^2) \left(x^2 - \frac{2\sigma_x^2 x^*}{\sigma_1^2 + \sigma_x^2} x + \left(\frac{\sigma_x^2 x^{*2}}{\sigma_1^2 + \sigma_x^2} \right)^2 - \left(\frac{\sigma_x^2 x^{*2}}{\sigma_1^2 + \sigma_x^2} \right)^2 \right) + \sigma_x^2 x^{*2} \quad (79)$$

$$= (\sigma_1^2 + \sigma_x^2) \left(x - \frac{\sigma_x^2 x^*}{\sigma_1^2 + \sigma_x^2} \right)^2 - \frac{(\sigma_x^2 x^*)^2}{\sigma_1^2 + \sigma_x^2} + \sigma_x^2 x^{*2} \quad (80)$$

$$= (\sigma_1^2 + \sigma_x^2) \left(x - \frac{\sigma_x^2 x^*}{\sigma_1^2 + \sigma_x^2} \right)^2 + \frac{\sigma_1^2 \sigma_x^2 x^{*2}}{\sigma_1^2 + \sigma_x^2} \quad (81)$$

となる. よって式 (73) の指数関数の中は

$$-\frac{\left(x - \frac{\sigma_x^2 x^*}{\sigma_1^2 + \sigma_x^2} \right)^2}{2 \left(\frac{\sigma_1^2 \sigma_x^2}{\sigma_1^2 + \sigma_x^2} \right)} - \frac{x^*}{2(\sigma_1^2 + \sigma_x^2)} \quad (82)$$

と変形できる．よって式 (73) は

$$\mu_{X^*}(x^*) = \frac{1}{2\pi\sqrt{\sigma_x^2\sigma_1^2}} \exp\left(-\frac{x^{*2}}{2(\sigma_1^2 + \sigma_x^2)}\right) \int_x \exp\left(\frac{-\left(x - \frac{\sigma_x^2 x^*}{\sigma_1^2 + \sigma_x^2}\right)^2}{2\left(\frac{\sigma_1^2\sigma_x^2}{\sigma_1^2 + \sigma_x^2}\right)}\right) dx \quad (83)$$

$$= \frac{1}{2\pi\sqrt{\sigma_x^2\sigma_1^2}} \exp\left(-\frac{x^{*2}}{2(\sigma_1^2 + \sigma_x^2)}\right) \sqrt{2\pi\left(\frac{\sigma_1^2\sigma_x^2}{\sigma_1^2 + \sigma_x^2}\right)} \quad (84)$$

$$= \frac{1}{\sqrt{2\pi(\sigma_1^2 + \sigma_x^2)}} \exp\left(-\frac{x^{*2}}{2(\sigma_1^2 + \sigma_x^2)}\right) \quad (85)$$

$$= f_{\sigma_1^2 + \sigma_x^2}(x^*) \quad (86)$$

となる．つまり， X^* は $X^* \sim \mathcal{N}(0, \sigma_x^2 + \sigma_1^2)$ の二乗誤差歪みをもつ．

式 (11)，式 (62)，式 (68)，式 (69)，式 (86) を式 (38) に代入すると

$$d^*(x^*, y^*) = \frac{1}{f_{\sigma_1^2 + \sigma_x^2}(x^*)} \iint_{x,y} f_{\sigma_x^2}(x) f_{\sigma_1^2}(x^* - x) f_{\sigma_2^2}(y - y^*) (x - y)^2 dx dy \quad (87)$$

$$= \frac{1}{f_{\sigma_1^2 + \sigma_x^2}(x^*)} \iint_{x,y} f_{\sigma_x^2}(x) f_{\sigma_1^2}(x^* - x) f_{\sigma_2^2}(y - y^*) (x^2 - 2xy + y^2) dx dy \quad (88)$$

となる．分配法則を適用し，まず $\int_x f_{\sigma_x^2}(x) f_{\sigma_1^2}(x^* - x) x^2 dx$ について考える．ここで

$$f_{\sigma_x^2}(x) f_{\sigma_1^2}(x^* - x) \quad (89)$$

$$= \frac{1}{\sqrt{2\pi\sigma_x^2}} \exp\left(-\frac{x^2}{2\sigma_x^2}\right) \frac{1}{\sqrt{2\pi\sigma_1^2}} \exp\left(-\frac{(x^* - x)^2}{2\sigma_1^2}\right) \quad (90)$$

$$= \frac{1}{2\pi\sigma_x\sigma_1} \exp\left(-\frac{x^2}{2\sigma_x^2} - \frac{(x^* - x)^2}{2\sigma_1^2}\right) \quad (91)$$

$$= \frac{1}{2\pi\sigma_x\sigma_1} \exp\left(-\frac{\left(x - \frac{\sigma_x^2 x^*}{\sigma_1^2 + \sigma_x^2}\right)^2}{2\left(\frac{\sigma_x^2\sigma_1^2}{\sigma_1^2 + \sigma_x^2}\right)} - \frac{x^{*2}}{2(\sigma_1^2 + \sigma_x^2)}\right) \quad (92)$$

$$= \frac{1}{2\pi\sigma_x\sigma_1} \exp\left(-\frac{x^{*2}}{2(\sigma_1^2 + \sigma_x^2)}\right) \exp\left(-\frac{\left(x - \frac{\sigma_x^2 x^*}{\sigma_1^2 + \sigma_x^2}\right)^2}{2\left(\frac{\sigma_x^2\sigma_1^2}{\sigma_1^2 + \sigma_x^2}\right)}\right) \quad (93)$$

$$= \frac{1}{\sqrt{2\pi(\sigma_1^2 + \sigma_x^2)}} \exp\left(-\frac{x^{*2}}{2(\sigma_1^2 + \sigma_x^2)}\right) \frac{1}{\sqrt{2\pi\frac{\sigma_x^2\sigma_1^2}{\sigma_1^2 + \sigma_x^2}}} \exp\left(-\frac{\left(x - \frac{\sigma_x^2 x^*}{\sigma_1^2 + \sigma_x^2}\right)^2}{2\left(\frac{\sigma_x^2\sigma_1^2}{\sigma_1^2 + \sigma_x^2}\right)}\right) \quad (94)$$

$$= f_{\sigma_1^2 + \sigma_x^2}(x^*) f_{\frac{\sigma_x^2\sigma_1^2}{\sigma_1^2 + \sigma_x^2}}\left(x - \frac{\sigma_x^2 x^*}{\sigma_1^2 + \sigma_x^2}\right) \quad (95)$$

となるので,

$$\int_x f_{\sigma_x^2}(x) f_{\sigma_1^2}(x^* - x) x^2 dx = f_{\sigma_1^2 + \sigma_x^2}(x^*) \int_x f_{\sigma_1^2 + \sigma_x^2}(x^*) f_{\frac{\sigma_x^2 \sigma_1^2}{\sigma_1^2 + \sigma_x^2}} \left(x - \frac{\sigma_x^2 x^*}{\sigma_1^2 + \sigma_x^2} \right) x^2 dx \quad (96)$$

と変形できる. ここで

$$\hat{x} = x - \frac{\sigma_x^2 x^*}{\sigma_1^2 + \sigma_x^2} \quad (97)$$

とおくと, 式 (96) は

$$f_{\sigma_1^2 + \sigma_x^2}(x^*) \int_x f_{\frac{\sigma_x^2 \sigma_1^2}{\sigma_1^2 + \sigma_x^2}} \left(x - \frac{\sigma_x^2 x^*}{\sigma_1^2 + \sigma_x^2} \right) x^2 dx \quad (98)$$

$$= f_{\sigma_1^2 + \sigma_x^2}(x^*) \int_{\hat{x}} f_{\frac{\sigma_x^2 \sigma_1^2}{\sigma_1^2 + \sigma_x^2}}(\hat{x}) \left(\hat{x} + \frac{\sigma_x^2 x^*}{\sigma_1^2 + \sigma_x^2} \right)^2 d\hat{x} \quad (99)$$

$$= f_{\sigma_1^2 + \sigma_x^2}(x^*) \int_{\hat{x}} f_{\frac{\sigma_x^2 \sigma_1^2}{\sigma_1^2 + \sigma_x^2}}(\hat{x}) \left(\hat{x}^2 + 2 \left(\frac{\sigma_x^2 x^*}{\sigma_1^2 + \sigma_x^2} \right) \hat{x} + \left(\frac{\sigma_x^2 x^*}{\sigma_1^2 + \sigma_x^2} \right)^2 \right) d\hat{x} \quad (100)$$

$$= f_{\sigma_1^2 + \sigma_x^2}(x^*) \left(\frac{\sigma_x^2 \sigma_1^2}{\sigma_1^2 + \sigma_x^2} + \left(\frac{\sigma_x^2 x^*}{\sigma_1^2 + \sigma_x^2} \right)^2 \right) \quad (101)$$

となる. さらに $\int_x f_{\sigma_x^2}(x) f_{\sigma_1^2}(x^* - x) x dx$ について考えると

$$\int_x f_{\sigma_x^2}^2(x) f_{\sigma_1^2}(x^* - x) x dx = f_{\sigma_1^2 + \sigma_x^2}(x^*) \int_x f_{\frac{\sigma_x^2 \sigma_1^2}{\sigma_1^2 + \sigma_x^2}} \left(x - \frac{\sigma_x^2 x^*}{\sigma_1^2 + \sigma_x^2} \right) x dx \quad (102)$$

$$= f_{\sigma_1^2 + \sigma_x^2}(x^*) \left(\frac{\sigma_x^2 x^*}{\sigma_1^2 + \sigma_x^2} \right) \quad (103)$$

となる. 次に y についての積分 $\int_y f_{\sigma_2^2}(y - y^*) y dy$ を考えると

$$\int_y f_{\sigma_2^2}(y - y^*) y dy = y^* \quad (104)$$

となり, さらに $\int_y f_{\sigma_2^2}(y - y^*) y^2 dy$ は $\hat{y} = y - y^*$ を用いれば

$$\int_y f_{\sigma_2^2}(y - y^*) y^2 dy = \int_{\hat{y}} f_{\sigma_2^2}(\hat{y}) (\hat{y} + y^*)^2 d\hat{y} \quad (105)$$

$$= \int_{\hat{y}} f_{\sigma_2^2}(\hat{y}) (\hat{y}^2 + 2\hat{y}y^* + y^{*2}) d\hat{y} \quad (106)$$

$$= \sigma_2^2 + y^{*2} \quad (107)$$

と表せる．以上の結果を用いて式 (101), 式 (103), 式 (104), 式 (107) を式 (38) に代入すれば

$$d^*(x^*, y^*) = \left(\frac{\sigma_x^2 \sigma_1^2}{\sigma_1^2 + \sigma_x^2} \right) + \left(\frac{\sigma_x^2 x^*}{\sigma_1^2 + \sigma_x^2} \right)^2 - 2 \left(\frac{\sigma_x^2 x^* y^*}{\sigma_1^2 + \sigma_x^2} \right) + \sigma_2^2 + y^{*2} \quad (108)$$

$$= \left(\frac{\sigma_x^2 \sigma_1^2}{\sigma_1^2 + \sigma_x^2} \right) + \sigma_2^2 + \left(\left(\frac{\sigma_x^2 x^*}{\sigma_1^2 + \sigma_x^2} \right) - y^* \right)^2 \quad (109)$$

と求められる． d^* の期待値は

$$\mathbb{E}[d^*(X^*, Y^*)] = \mathbb{E} \left[d \left(\left(\frac{\sigma_x^2}{\sigma_1^2 + \sigma_x^2} \right) X^*, Y^* \right) + \frac{\sigma_x^2 \sigma_1^2}{\sigma_1^2 + \sigma_x^2} + \sigma_2^2 \right] \quad (110)$$

$$= \mathbb{E} \left[d \left(\left(\frac{\sigma_x^2}{\sigma_1^2 + \sigma_x^2} \right) X^*, Y^* \right) \right] + \frac{\sigma_x^2 \sigma_1^2}{\sigma_1^2 + \sigma_x^2} + \sigma_2^2 \quad (111)$$

となる．よって

$$\mathbb{E}[d^*(X^*, Y^*)] \leq D \quad (112)$$

は

$$\mathbb{E} \left[d \left(\left(\frac{\sigma_x^2}{\sigma_1^2 + \sigma_x^2} \right) X^*, Y^* \right) \right] \leq D - \frac{\sigma_x^2 \sigma_1^2}{\sigma_1^2 + \sigma_x^2} - \sigma_2^2 \quad (113)$$

となり，定理 3 より雑音のあるシステムでのガウス型情報源 $\mathbf{X}$ のレート・歪み関数は

$$R(D) = \min_{Y^*: \mathbb{E} \left[d \left(\left(\frac{\sigma_x^2}{\sigma_1^2 + \sigma_x^2} \right) X^*, Y^* \right) \right] \leq D - \frac{\sigma_x^2 \sigma_1^2}{\sigma_1^2 + \sigma_x^2} - \sigma_2^2} I(X^*; Y^*) \quad (114)$$

$$= \frac{1}{\left(\frac{\sigma_x^2}{\sigma_1^2 + \sigma_x^2} \right)} \min_{Y^*: \mathbb{E} \left[d \left(\left(\frac{\sigma_x^2}{\sigma_1^2 + \sigma_x^2} \right) X^*, Y^* \right) \right] \leq D - \frac{\sigma_x^2 \sigma_1^2}{\sigma_1^2 + \sigma_x^2} - \sigma_2^2} I \left(\left(\frac{\sigma_x^2}{\sigma_1^2 + \sigma_x^2} \right) X^*; Y^* \right) \quad (115)$$

$$= \begin{cases} \frac{\sigma_1^2 + \sigma_x^2}{2\sigma_x^2} \log \frac{\sigma_x^4}{\left(D - \left(\frac{\sigma_1^2 \sigma_x^2}{\sigma_1^2 + \sigma_x^2} \right) - \sigma_2^2 \right) (\sigma_1^2 + \sigma_x^2)} & (D \leq \sigma_x^2 + \sigma_2^2) \\ 0 & (D > \sigma_x^2 + \sigma_2^2) \end{cases} \quad (116)$$

と計算できる．式 (114), 式 (115) の変形には

$$I(X^*; Y^*) = \frac{1}{\left(\frac{\sigma_x^2}{\sigma_1^2 + \sigma_x^2} \right)} I \left(\left(\frac{\sigma_x^2}{\sigma_1^2 + \sigma_x^2} \right) X^*; Y^* \right) \quad (117)$$

を用いた．証明は付録 C に示す．

3.2 得られた結果と考察

3.1 節の式 (116) で求めた式を使い，符号化前と復号後の付加された雑音の分散を変化させたレート・歪み関数を計算した．ガウス型情報源 X は $X \sim \mathcal{N}(0, 1)$ とし，雑音は

$N_1 \sim \mathcal{N}(0, \sigma_1^2)$, $N_2 \sim \mathcal{N}(0, \sigma_2^2)$ のみの場合を考える． $\sigma_x^2 = 1$ のもとで σ_1^2 を 0 から 1 まで 0.2 刻みで変化させた場合のレート・歪み関数を図 4 に示す．このとき復号器から受信者に届くときの雑音 σ_2^2 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 $\sigma_1^2 = 0, 0.2, 0.4, 0.6, 0.8, 1$ である．同様に σ_2^2 を 0 から 1 まで 0.2 刻みで変化させたときのレート・歪み関数を図 5 に示す．このとき雑音 σ_1^2 は 0 とする．左のグラフから順に $\sigma_2^2 = 0, 0.2, 0.4, 0.6, 0.8, 1$ である．また， $\sigma_x^2 = 0.8$ のもとで σ_1^2 を 0, 0.2, 0.4, 0.6, 0.8, 1 と変化したときのレート・歪み関数を図 6 に， $\sigma_2^2 = 0, 0.2, 0.4, 0.6, 0.8, 1$ に変化させたときのレート・歪み関数を図 7 に示す． $\sigma_x^2 = 0.6$ のもとで確率 σ_1^2 を 0, 0.2, 0.4, 0.6, 0.8, 1 と変化したときのレート・歪み関数を図 8 に， $\sigma_2^2 = 0, 0.2, 0.4, 0.6, 0.8, 1$ に変化させたときのレート・歪み関数を図 9 に示す．

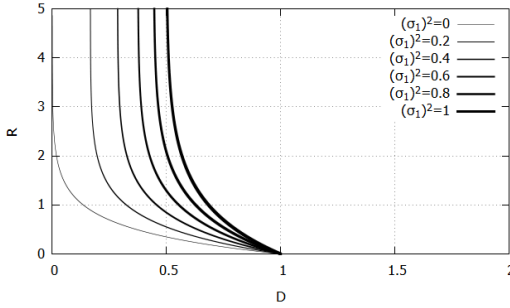


図 4 $\sigma_x^2 = 1.0$ のとき σ_1^2 を変化

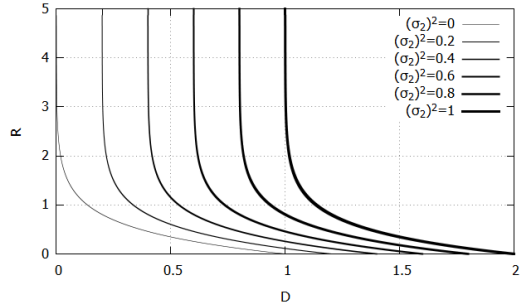


図 5 $\sigma_x^2 = 1.0$ のとき σ_2^2 を変化

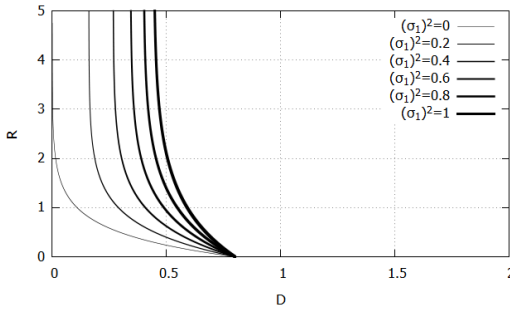


図 6 $\sigma_x^2 = 0.8$ のとき σ_1^2 を変化

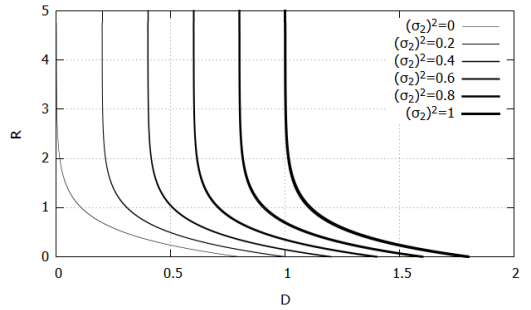


図 7 $\sigma_x^2 = 0.8$ のとき σ_2^2 を変化

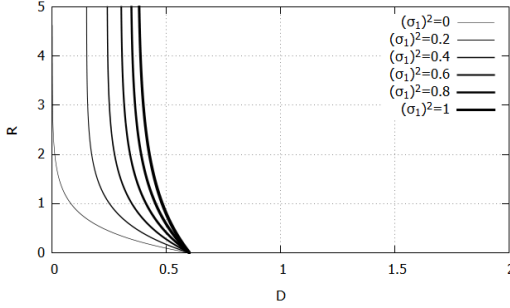


図 8 $\sigma_x^2 = 0.6$ のとき σ_1^2 を変化

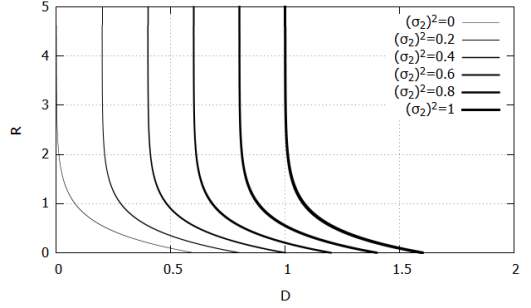


図 9 $\sigma_x^2 = 0.6$ のとき σ_2^2 を変化

横軸は歪み，縦軸はレートを表している．プロットを省略したが， $D > \sigma_x^2$ のとき R の値は 0 となった．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なるというレートと歪みの関係を表している．

雑音 N_1, N_2 のパワー σ_1^2, σ_2^2 を 0 から変化させたとき，図 4 から図 9 に示したようにいずれの場合でもレート R が最大値をとるとき D の値は情報源の歪み σ_x^2 より小さくできない．これは，情報源のパワーより大きい歪みを許した場合，デコーダーで適当な値を出力し続けられ良いということになり，通信の意味がなくなってしまうことを意味している．

図 4, 図 6, 図 8 から， σ_1^2 の値を変化させると曲線の発散する場所に違いがみられた． σ_1^2 の値を大きくすると，グラフの発散する場所は σ_x^2 に近くなった．これは式 (114), 式 (116) より， $D \leq \sigma_x^2 + \sigma_2^2$ のときレート・歪み関数 $R(D)$ が

$$R(D) = \min_{Y^*: \mathbb{E}\left[d\left(\left(\frac{\sigma_x^2}{\sigma_1^2 + \sigma_x^2}\right)X^*, Y^*\right)\right] \leq D - \frac{\sigma_x^2 \sigma_1^2}{\sigma_1^2 + \sigma_x^2} - \sigma_2^2} I(X^*; Y^*) \quad (118)$$

$$= \frac{\sigma_1^2 + \sigma_x^2}{2\sigma_x^2} \log \frac{\sigma_x^4}{\left(D - \left(\frac{\sigma_1^2 \sigma_x^2}{\sigma_1^2 + \sigma_x^2}\right) - \sigma_2^2\right) (\sigma_1^2 + \sigma_x^2)} \quad (119)$$

となり，ここで $\sigma_x^2 = 1, \sigma_2^2 = 0$ を代入すると

$$R(D) = \frac{\sigma_1^2 + 1}{2} \log \frac{1}{\left(D - \left(\frac{\sigma_1^2}{\sigma_1^2 + 1}\right)\right) (\sigma_1^2 + 1)} \quad (120)$$

$$= \frac{\sigma_1^2 + 1}{2} \log \frac{1}{D(\sigma_1^2 + 1) - \sigma_1^2} \quad (121)$$

$$= \frac{\sigma_1^2 + 1}{2} \log \frac{1}{\sigma_1^2(D - 1) + D} \quad (122)$$

$$(123)$$

となる．このとき $D \leq 1$ である．ここで σ_1^2 を ∞ に近づけると

$$\lim_{\sigma_1^2 \rightarrow \infty} \frac{\sigma_1^2 + 1}{2} \log \frac{1}{\sigma_1^2 (D - 1) + D} \quad (124)$$

$$= \infty \times -\infty \quad (125)$$

$$= -\infty \quad (126)$$

となるので， σ_1^2 を ∞ に近づけるとグラフの傾きは $-\infty$ になり，発散する場所が σ_x^2 の値に近づいていくと考えられる．

図 4 と図 5，図 6 と図 7，図 8 と図 9 から $R = 0$ のときの D の値に違いがみられた．グラフは σ_2^2 を変化させた分だけ，正の方向に平行移動している．これは， σ_2^2 が復号器を通ったあとに付加される雑音のパワーであり，式 (116) を見ると σ_2^2 が D に対して引き算をしているため，元の形を保ったまま正の方向に平行移動したのだと考えられる．

次に， N_1, N_2 がともに存在する場合について考える． $\sigma_2^2 = 0.2$ に固定し， σ_1^2 を 0 から 0.5 まで 0.1 刻みで変化させたグラフを図 10 に示す．左から順に $\sigma_1^2 = 0, 0.1, 0.2, 0.3, 0.4, 0.5$ である．そして $\sigma_1^2 = 0.2$ に固定し， σ_2^2 を 0 から 0.5 まで 0.1 刻みで変化させたグラフを図 11 に示す．

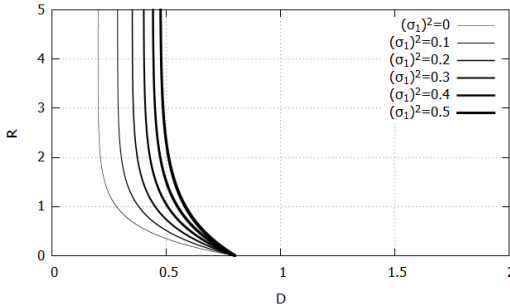


図 10 $\sigma_x^2 = 0.6$ のとき $\sigma_2^2 = 0.2$ に固定， σ_1^2 を変化

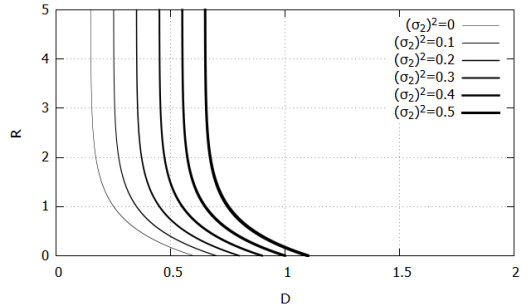


図 11 $\sigma_x^2 = 0.6$ のとき $\sigma_1^2 = 0.2$ に固定， σ_2^2 を変化

レート R が最大になるときの D の値は図 10 のように， σ_2^2 の雑音の影響を受け 0.2 正の方向に平行移動した後に， σ_1^2 の雑音の影響を受け曲線の発散する場所が変化した形となった．図 10 も同様に σ_1^2 の雑音の影響を受け曲線の発散する場所が変化した後に， σ_2^2 の雑音の影響を受け正の方向に 0.1 ずつ平行移動が加わったグラフとなった． N_1, N_2 がともに存在するとき，情報源 X は σ_x^2 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， σ_1^2 で歪んだデータが届けられる．そこからさらに符号化されたあと σ_2^2 で歪むため， σ_1^2 と σ_2^2 の雑音の影響が合わさったと考えられる．

4 まとめ

本研究では、雑音によりデータに歪みが生じるシステム内での連続値におけるレート・歪み問題を考えた。連続値の場合でも、離散値のレート・歪み理論と同じ定理が成り立つことを確認できた。また、ガウス型情報源においても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、レート・歪み関数を計算によって求めることができた。今後の課題としては、ガウス型情報源以外の情報源や通信路自体に雑音がある場合にどのような結果が出るのか検討してみたい。

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.
- [6] 澤田真子, 「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」, 信州大学工学部卒業論文 (指導教員: 西新幹彦), 2020 年 3 月.
- [7] T. Cover, J. Thomas (山本博資, 古賀弘樹, 有村光晴, 岩本貢訳), 情報理論 基礎と広がり, 共立出版株式会社, 2012.

付録 A 定理 1 の証明

([6]) 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

$$\Pr\{Z \geq a\} \leq \frac{\mathbb{E}[Z]}{a} \quad (127)$$

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

$$\mathbb{E}[Z] = \sum_z z P_Z(z) \quad (128)$$

$$= \sum_{z>a} z P_Z(z) + \sum_{z \leq a} z P_Z(z) \quad (129)$$

$$\geq \sum_{z>a} P_Z(z) \quad (130)$$

$$= a \Pr\{Z > a\mu\} \quad (131)$$

と変形でき,

$$\Pr\{Z \geq a\} \leq \frac{\mathbb{E}[Z]}{a} \quad (132)$$

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

$$\Pr\{|Z - \mu| > a\sigma\} \leq \frac{1}{a^2} \quad (133)$$

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

$$\Pr\{(Z - \mu)^2 > a^2\sigma^2\} \leq \frac{\mathbb{E}[(Z - \mu)^2]}{a^2\sigma^2} \quad (134)$$

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

$$\Pr\{|Z - \mu| > a\sigma\} \leq \frac{1}{a^2} \quad (135)$$

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

$$\hat{W}^n(\mathbf{x}|\mathbf{y}) \triangleq \mu\{X^n = \mathbf{x} | Y^n = \mathbf{y}\} \quad (136)$$

$$\mathcal{B}_n(\mathbf{y}) \triangleq \left\{ \mathbf{x} \left| \frac{1}{n} \log \frac{\hat{W}^n(\mathbf{x}|\mathbf{y})}{\mu_{X^n}(\mathbf{x})} \leq R_n - \gamma \right. \right\} \quad (137)$$

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

$$\frac{1}{n} \log M_n \leq R_n \quad (138)$$

$$\Pr\{X^n \notin \mathcal{S}_n(\psi_n(\varphi_n(X^n)))\} \leq \Pr\{X^n \notin \mathcal{B}_n(Y^n)\} + \Pr\{X^n \notin \mathcal{S}(Y^n)\} + e^{-n\gamma} \quad (139)$$

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

$$\hat{W}^n(\mathcal{T}_n(\mathbf{y})|\mathbf{y}) = \int_{\mathcal{T}_n(\mathbf{y})} \hat{W}^n(\mathbf{x}|\mathbf{y}) d\mathbf{x} \geq e^{-n\gamma} \quad (140)$$

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

$$\hat{W}^n(\mathcal{T}_n(\mathbf{y}) \setminus \bigcup_{m' < m} \varphi_n^{-1}(m') | \mathbf{y}) \geq e^{-n\gamma} \quad (141)$$

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

$$\varphi_n^{-1}(m) \triangleq \mathcal{T}_n(\psi_n(m)) \setminus \bigcup_{m' < m} \varphi_n^{-1}(m') \quad (142)$$

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

$$\mathcal{D} \triangleq \bigcup_m \varphi_n^{-1}(m) \quad (143)$$

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

$$\hat{W}^n(\mathcal{T}_n(\mathbf{y}) \setminus \mathcal{D} | \mathbf{y}) < e^{-n\gamma} \quad (144)$$

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 (137) より

$$\frac{1}{n} \log \frac{\hat{W}^n(\mathbf{x}|\psi_n(m))}{\mu_{X^n}(\mathbf{x})} \leq R_n - \gamma \quad (145)$$

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

$$\mu_{X^n}(\mathbf{x}) \geq \hat{W}^n(\mathbf{x}|\psi_n(m))e^{-n(R_n-\gamma)} \quad (146)$$

が成り立つことに注意しておく．ここで式 (141), (143), (146) より

$$1 \geq \mu_{X^n}(\mathcal{D}) \quad (147)$$

$$= \mu_{X^n}\left(\bigcup_m \varphi_n^{-1}(m)\right) \quad (148)$$

$$= \sum_m \mu_{X^n}(\varphi_n^{-1}(m)) \quad (149)$$

$$= \sum_m \int_{\varphi_n^{-1}(m)} \mu_{X^n}(\mathbf{x}) d\mathbf{x} \quad (150)$$

$$\geq \sum_m \int_{\varphi_n^{-1}(m)} \hat{W}^n(\mathbf{x}|\psi_n(m))e^{-n(R_n-\gamma)} d\mathbf{x} \quad (151)$$

$$= \sum_m \hat{W}^n(\varphi_n^{-1}(m)|\psi_n(m))e^{-n(R_n-\gamma)} \quad (152)$$

$$\geq \sum_m e^{-nR_n} \quad (153)$$

$$= M_n e^{-nR_n} \quad (154)$$

となり

$$\frac{1}{n} \log M_n \leq R_n \quad (155)$$

が得られる．

次に式 (139) 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

$$\Pr\{X^n \in \mathcal{D}\} \leq \Pr\{X^n \in \mathcal{S}_n(\psi_n(\varphi_n(X^n)))\} \quad (156)$$

といえる．一方，式 (144) により任意の $\mathbf{y}$ について

$$\hat{W}^n(\mathcal{T}_n(\mathbf{y})|\mathbf{y}) = \hat{W}^n(\mathcal{T}_n(\mathbf{y}) \cap \mathcal{D}|\mathbf{y}) + \hat{W}^n(\mathcal{T}_n(\mathbf{y}) \setminus \mathcal{D}|\mathbf{y}) \quad (157)$$

$$\leq \hat{W}^n(\mathcal{D}|\mathbf{y}) + e^{-n\gamma} \quad (158)$$

となるので、式 (156) の左辺は

$$\Pr\{X^n \in \mathcal{D}\} = \mu_{X^n}(\mathcal{D}) \quad (159)$$

$$= \int_{\mathbf{y}} \mu_{X^n Y^n}(\mathcal{D}, \mathbf{y}) d\mathbf{y} \quad (160)$$

$$= \int_{\mathbf{y}} \mu_{Y^n}(\mathbf{y}) \hat{W}^n(\mathcal{D}|\mathbf{y}) d\mathbf{y} \quad (161)$$

$$\geq \int_{\mathbf{y}} \{\mu_{Y^n}(\mathbf{y}) (\hat{W}^n(\mathcal{T}_n(\mathbf{y})|\mathbf{y}) - e^{-n\gamma})\} d\mathbf{y} \quad (162)$$

$$= \int_{\mathbf{y}} \hat{W}^n(\mathcal{T}_n(\mathbf{y})|\mathbf{y}) d\mathbf{y} - \int_{\mathbf{y}} \{\mu_{Y^n}(\mathbf{y}) e^{-n\gamma}\} d\mathbf{y} \quad (163)$$

$$= \int_{X^n Y^n} ((\mathcal{T}_n(\mathbf{y}), \mathbf{y}) d\mathbf{y} - e^{-n\gamma} \quad (164)$$

$$= \Pr\{X^n \in \mathcal{T}_n(Y^n)\} - e^{-n\gamma} \quad (165)$$

$$= \Pr\{X^n \in \mathcal{B}_n(Y^n) \text{ かつ } X^n \in \mathcal{S}_n(Y^n)\} - e^{-n\gamma} \quad (166)$$

$$= 1 - \Pr\{X^n \notin \mathcal{B}_n(Y^n) \text{ または } X^n \notin \mathcal{S}_n(Y^n)\} - e^{-n\gamma} \quad (167)$$

$$\geq 1 - \Pr\{X^n \notin \mathcal{B}_n(Y^n)\} - \{X^n \notin \mathcal{S}_n(Y^n)\} - e^{-n\gamma} \quad (168)$$

式 (156) と式 (168) より

$$\Pr\{X^n \in \mathcal{S}_n(\psi_n(\varphi_n(X^n)))\} \geq 1 - \Pr\{X^n \notin \mathcal{B}_n(Y^n)\} - \{X^n \notin \mathcal{S}_n(Y^n)\} - e^{-n\gamma} \quad (169)$$

となるので

$$\Pr\{X^n \notin \mathcal{S}_n(\psi_n(\varphi_n(X^n)))\} \leq \Pr\{X^n \notin \mathcal{B}_n(Y^n)\} + \Pr\{X^n \notin \mathcal{S}_n(Y^n)\} + e^{-n\gamma} \quad (170)$$

が導かれる。 $\square$

A.3 定理 1 の証明

[6] 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、

$$T_n^{(1)} = \left\{ (\mathbf{x}, \mathbf{y}) \in \mathcal{X}^n \times \mathcal{Y}^n \mid \left| \frac{1}{n} \log \frac{\mu_{Y^n|X^n}(\mathbf{y}|\mathbf{x})}{\mu_{Y^n}(\mathbf{y})} - I(X; Y) \right| < \gamma \right\} \quad (171)$$

$$T_n^{(2)} = \left\{ (\mathbf{x}, \mathbf{y}) \in \mathcal{X}^n \times \mathcal{Y}^n \mid \left| \frac{1}{n} d_n(\mathbf{x}, \mathbf{y}) - \mathbb{E}[d(X, Y)] \right| < \gamma \right\} \quad (172)$$

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,

$$\frac{1}{n} \log \frac{\mu_{Y^n|X^n}(Y^n|X^n)}{\mu_{Y^n}(Y^n)} = \frac{1}{n} \log \frac{\prod_{i=1}^n \mu_{Y|X}(Y_i|X_i)}{\prod_{i=1}^n \mu_Y(Y_i)} \quad (173)$$

$$= \frac{1}{n} \sum_{i=1}^n \log \frac{\mu_{Y|X}(Y_i|X_i)}{\mu_Y(Y_i)} \quad (174)$$

となる. また,

$$d_n(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^n d(x_i, y_i) \quad (175)$$

とおいたので,

$$\frac{1}{n} d_n(X^n, Y^n) = \frac{1}{n} \sum_{i=1}^n d(X_i, Y_i) \quad (176)$$

が成り立つ. 一方, 式 (174), (176) の右辺の和の中の各項は独立でその期待値は

$$\mathbb{E} \left[\log \frac{\mu_{Y|X}(Y_i|X_i)}{\mu_Y(Y_i)} \right] = \int_{x,y} \mu_{XY}(x, y) \log \frac{\mu_{Y|X}(y|x)}{\mu_Y(y)} dx dy \quad (177)$$

$$= I(X; Y), \quad (i = 1, 2, \dots) \quad (178)$$

$$\mathbb{E}[d(X_i, Y_i)] = \mathbb{E}[d(X, Y)], \quad (i = 1, 2, \dots) \quad (179)$$

であり, 期待値は有界で, 分散も有界と仮定する. $T_n^{(1)}$, $T_n^{(2)}$ の定義 (171), (172) に注意して Chebyshev の不等式を適用すれば

$$\Pr\{X^n Y^n \in T_n^{(1)}\} \rightarrow 1 \quad (n \rightarrow \infty) \quad (180)$$

$$\Pr\{X^n Y^n \in T_n^{(2)}\} \rightarrow 1 \quad (n \rightarrow \infty) \quad (181)$$

を得る. したがって,

$$\Pr\{X^n Y^n \in T_n\} = 1 - \Pr\{X^n Y^n \notin T_n\} \quad (182)$$

$$\geq 1 - \Pr\{X^n Y^n \notin T_n^{(1)}\} - \Pr\{X^n Y^n \notin T_n^{(2)}\} \quad (183)$$

$$= 1 - \{1 - \Pr\{X^n Y^n \in T_n^{(1)}\}\} - \{1 - \Pr\{X^n Y^n \in T_n^{(2)}\}\} \quad (184)$$

$$\rightarrow 1 \quad (n \rightarrow \infty) \quad (185)$$

が成立する. ここで

$$\hat{W}^n(\mathbf{x}|\mathbf{y}) \triangleq \mu_{X^n|Y^n}(\mathbf{x}|\mathbf{y}) \quad (186)$$

$$\mathcal{B}_n(\mathbf{y}) \triangleq \left\{ \mathbf{x} \left| \frac{1}{n} \log \frac{\hat{W}^n(\mathbf{x}|\mathbf{y})}{\mu_{X^n}(\mathbf{x})} \leq R_n - \gamma \right. \right\} \quad (187)$$

$$\mathcal{S}_n(\mathbf{y}) \triangleq \{\mathbf{x} \in \mathcal{X}^n | (\mathbf{x}, \mathbf{y}) \in T_n^{(2)}\} \quad (188)$$

任意の γ について考える. $R_n = R \triangleq I(X; Y) + 2\gamma$ とおくと, 補題 1 より

$$\frac{1}{n} \log M_n \leq R_n \quad (189)$$

$$\Pr\{X^n \notin \mathcal{S}_n(\psi_n(\varphi_n(X^n)))\} \leq \Pr\{X^n \notin \mathcal{B}_n(Y^n)\} + \Pr\{X^n \notin \mathcal{S}(Y^n)\} + e^{-n\gamma} \quad (190)$$

を満たす (φ_n, ψ_n) 符号が存在する. 式 (190) の左辺は式 (172) と式 (188) より

$$\Pr\{X^n \notin \mathcal{S}_n(\psi_n(\varphi_n(X^n)))\} = \Pr\{(X^n, \psi_n(\varphi_n(X^n))) \notin T_n^{(2)}\} \quad (191)$$

$$= \Pr\left\{\left|\frac{1}{n}d_n(X^n, \psi_n(\varphi_n(X^n))) - \mathbb{E}[d(X, Y)]\right| \geq \gamma\right\} \quad (192)$$

となる. 右辺の各項は, 式 (171), 式 (137), 式 (136) より,

$$\Pr\{X^n \in \mathcal{B}_n(Y^n)\} = \Pr\left\{\frac{1}{n} \log \frac{\hat{W}^n(X^n|Y^n)}{\mu_{X^n}(X^n)} \leq I(X; Y) + \gamma\right\} \quad (193)$$

$$= \Pr\left\{\frac{1}{n} \log \frac{\mu_{X^n|Y^n}(X^n|Y^n)}{\mu_{X^n}(X^n)} \leq I(X; Y) + \gamma\right\} \quad (194)$$

$$= \Pr\left\{\frac{1}{n} \log \frac{\mu_{Y^n|X^n}(Y^n|X^n)}{\mu_{Y^n}(Y^n)} \leq I(X; Y) + \gamma\right\} \quad (195)$$

$$= \Pr\left\{\frac{1}{n} \log \frac{\mu_{Y^n|X^n}(Y^n|X^n)}{\mu_{Y^n}(Y^n)} - I(X; Y) \leq \gamma\right\} \quad (196)$$

$$\geq \Pr\left\{\left|\frac{1}{n} \log \frac{\mu_{Y^n|X^n}(Y^n|X^n)}{\mu_{Y^n}(Y^n)} - I(X; Y)\right| < \gamma\right\} \quad (197)$$

$$= \Pr\{X^n Y^n \in T_n^{(1)}\} \quad (198)$$

$$\rightarrow 1 \quad (n \rightarrow \infty) \quad (199)$$

$$\therefore \Pr\{X^n \notin \mathcal{B}_n(Y^n)\} \rightarrow 0 \quad (200)$$

となる. ただし式 (196) から式 (197) への不等式では

$$\frac{1}{n} \log \frac{\mu_{Y^n|X^n}(Y^n|X^n)}{\mu_{Y^n}(Y^n)} - I(X; Y) \leq \left|\frac{1}{n} \log \frac{\mu_{Y^n|X^n}(Y^n|X^n)}{\mu_{Y^n}(Y^n)} - I(X; Y)\right| \quad (201)$$

を用いた. また, 式 (188) と式 (181) であるから,

$$\Pr\{X^n \in \mathcal{S}_n(Y^n)\} = \Pr\{X^n Y^n \in T_n^{(2)}\} \quad (202)$$

$$\rightarrow 1 \quad (n \rightarrow \infty) \quad (203)$$

$$\therefore \Pr\{X^n \notin \mathcal{S}_n(Y^n)\} \rightarrow 0 \quad (n \rightarrow \infty) \quad (204)$$

となる. 以上より,

$$\Pr\{X^n \notin \mathcal{S}_n(\psi_n(\varphi_n(X^n)))\} \rightarrow 0 \quad (n \rightarrow \infty) \quad (205)$$

が導かれる．ここで

$$R(D) \triangleq \inf_{Y: \mathbb{E}[d(X, Y)] \leq D} I(X; Y) \quad (206)$$

が D に関して下に凸かどうか考える． Y_1, Y_2 を

$$\mathbb{E}[d(X, Y_1)] \leq D_1 \quad (207)$$

$$\mathbb{E}[d(X, Y_2)] \leq D_2 \quad (208)$$

を満たす確率変数とする．また， $\mu_Q(i) = \alpha_i, i = 1, 2$ となるような確率変数を Q とし，

$$Y' \triangleq \begin{cases} Y_1 & (Q = 1) \\ Y_2 & (Q = 2) \end{cases} \quad (209)$$

となるような確率変数 Y' を定義する．このとき X と Y' の期待値は

$$\mathbb{E}[d(X, Y')] = \iint_{x, y} d(x, y) \mu'_{XY}(x, y) dx dy \quad (210)$$

$$= \iint_{x, y} \sum_i d(x, y) \mu_{XY'Q}(x, y, i) dx dy \quad (211)$$

$$= \sum_i \mu_Q(i) \iint_{x, y} d(x, y) \mu_{XY_iQ}(x, y | i) dx dy \quad (212)$$

$$= \sum_i \mu_Q(i) \mathbb{E}[d(X, Y_i) | Q = i] \quad (213)$$

$$= \sum_i \mu_Q(i) \mathbb{E}[d(X, Y_i)] \quad (214)$$

$$= \alpha_1 \mathbb{E}[d(X, Y_1)] + \alpha_2 \mathbb{E}[d(X, Y_2)] \quad (215)$$

$$\leq \alpha_1 D_1 + \alpha_2 D_2 \quad (216)$$

となる． Q と X は独立なので

$$\alpha_1 I(X; Y_1) + \alpha_2 I(X; Y_2) = \sum_i \mu_Q(i) I(X; Y_i) \quad (217)$$

$$= \sum_i \mu_Q(i) I(X; Y' | Q = i) \quad (218)$$

$$= I(X; Y' | Q) \quad (219)$$

$$\geq I(X; Y') \quad (220)$$

となる． よって

$$\alpha_1 R(D_1) + \alpha_2 R(D_2) = \inf_{Y_1: \mathbb{E}[d(X, Y_1)] \leq D_1} \alpha_1 I(X; Y_1) + \inf_{Y_2: \mathbb{E}[d(X, Y_2)] \leq D_2} \alpha_2 I(X; Y_2) \quad (221)$$

$$= \inf_{\begin{cases} Y_1: \mathbb{E}[d(X, Y_1)] \leq D_1 \\ Y_2: \mathbb{E}[d(X, Y_2)] \leq D_2 \\ Y' = Y_Q \end{cases}} (\alpha_1 I(X; Y_1) + \alpha_2 I(X; Y_2)) \quad (222)$$

$$\geq \inf_{\begin{cases} Y_1: \mathbb{E}[d(X, Y_1)] \leq D_1 \\ Y_2: \mathbb{E}[d(X, Y_2)] \leq D_2 \\ Y' = Y_Q \end{cases}} I(X; Y') \quad (223)$$

$$\geq \inf_{Y: \mathbb{E}[d(X, Y)] \leq \alpha_1 D_1 + \alpha_2 D_2} I(X; Y) \quad (224)$$

$$= R(\alpha_1 D_1 + \alpha_2 D_2) \quad (225)$$

となる． このことから R は下に凸で連続である． 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

$$R(D) \leq \inf_{Y: \mathbb{E}[d(X, Y)] \leq D - \gamma} I(X; Y) + 2\gamma \quad (226)$$

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

$$\text{さらに } R(D) \leq \inf_{Y: \mathbb{E}[d(X, Y)] \leq D} I(X; Y) \quad (227)$$

$I(X; Y)$ 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

$$R \leq \min_{Y: \mathbb{E}[d(X, Y)] \leq D} I(X; Y) \quad (228)$$

が導け， 順定理が示せた． $\square$

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

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \mathbb{E}[d_n(X^n, \psi_n(\varphi_n(X^n)))] \leq D \quad (229)$$

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log M_n \leq R \quad (230)$$

なる符号器と復号器の組 (φ_n, ψ_n) が存在する． そこで，

$$Y^n \triangleq (Y_1^{(n)}, Y_2^{(n)}, \dots, Y_n^{(n)}) \quad (231)$$

$$= \psi_n(\varphi_n(X^n)) \quad (232)$$

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

$$\log M_n \geq H(Y^n) \quad (233)$$

$$\geq I(X^n; Y^n) \quad (234)$$

$$= H(X^n) - H(X^n | Y^n) \quad (235)$$

$$= \sum_{i=1}^n H(X_i) - \sum_{i=1}^n H(X_i | Y^n X^{i-1}) \quad (236)$$

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、

$$H(X_i | Y^n X^{i-1}) \leq H(X_i | Y_i^{(n)}) \quad (237)$$

に注意すれば

$$\log M_n \geq \sum_{i=1}^n H(X_i) - \sum_{i=1}^n H(X_i | Y_i^{(n)}) \quad (238)$$

$$= \sum_{i=1}^n I(X_i; Y_i^{(n)}) \quad (239)$$

を得る。ゆえに、

$$\frac{1}{n} \log M_n \geq \frac{1}{n} \sum_{i=1}^n I(X_i; Y_i^{(n)}) \quad (240)$$

である。そこで、確率変数 $Q_n \in \{1, 2, \dots, n\}$ を $\mu_{Q_n}(i) = 1/n (i = 1, 2, \dots, n)$ 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 (240) は

$$\frac{1}{n} \log M_n \geq \frac{1}{n} \sum_{i=1}^n I(X_i; Y_i^{(n)}) \quad (241)$$

$$= \sum_{i=1}^n \mu_{Q_n}(i) I(X_i; Y_i^{(n)}) \quad (242)$$

$$= \sum_{i=1}^n \mu_{Q_n}(i) I(X; Y^{(n)} | Q_n = i) \quad (243)$$

$$= I(X; Y^{(n)} | Q_n) \quad (244)$$

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

$$I(X; Y^{(n)} | Q_n) = \iint_{x,y} \sum_i \mu_{XY^{(n)}Q_n}(x, y, i) \log \frac{\mu_{XY^{(n)}|Q_n}(x, y|i)}{\mu_{X|Q_n}(x|i)\mu_{Y^{(n)}|Q_n}(y|i)} dx dy \quad (245)$$

$$= \iint_{x,y} \sum_i \mu_{XY^{(n)}Q_n}(x, y, i) \log \frac{\mu_{XY^{(n)}|Q_n}(x, y|i)}{\mu_X(x)\mu_{Y^{(n)}|Q_n}(y|i)} dx dy \quad (246)$$

$$= I(X; Q_n Y^{(n)}) \quad (247)$$

$$\geq I(X; Y^{(n)}) \quad (248)$$

となり，したがって，式 (244) は

$$\frac{1}{n} \log M_n \geq I(X; Y^{(n)}) \quad (249)$$

となる．一方，

$$\frac{1}{n} \mathbb{E}[d_n(X^n, \psi_n(\varphi_n(X^n)))] = \frac{1}{n} \mathbb{E}[d_n(X^n, Y^n)] \quad (250)$$

$$= \frac{1}{n} \sum_{i=1}^n \mathbb{E}[d(X_i, Y_i^{(n)})] \quad (251)$$

$$= \sum_{i=1}^n \mu_{Q_n}(i) \mathbb{E}[d(X_i, Y_i^{(n)})] \quad (252)$$

$$= \sum_{i=1}^n \mu_{Q_n}(i) \mathbb{E}[d(X, Y^{(n)}) | Q_n = i] \quad (253)$$

$$= \sum_{i=1}^n \mu_{Q_n}(i) \iint_{x,y} d(x, y) \mu_{XY^{(n)}|Q_n}(x, y|i) dx dy \quad (254)$$

$$= \iint_{x,y} d(x, y) \mu_{XY^{(n)}}(x, y) dx dy \quad (255)$$

$$= \mathbb{E}[d(X, Y^{(n)})] \quad (256)$$

なので，式 (229) から

$$D \geq \limsup_{n \rightarrow \infty} \mathbb{E}[d(X, Y^{(n)})] \quad (257)$$

したがって，任意に小さい定数 $\gamma > 0$ に対して

$$D + \gamma \geq \mathbb{E}[d(X, Y^{(n)})] \quad (\forall n \geq n_0) \quad (258)$$

すると，式 (249) により

$$\frac{1}{n} \log M_n \geq \min_{Y: \mathbb{E}[d(X, Y)] \leq D + \gamma} I(X; Y) \quad (259)$$

となる．よって，式 (230) から

$$R \geq \limsup_{n \rightarrow \infty} \frac{1}{n} \log M_n \quad (260)$$

$$\geq \limsup_{n \rightarrow \infty} \min_{Y: \mathbb{E}[d(X, Y)] \leq D + \gamma} I(X; Y) \quad (261)$$

$$= \min_{Y: \mathbb{E}[d(X, Y)] \leq D + \gamma} I(X; Y) \quad (262)$$

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

$$R \geq \min_{Y: \mathbb{E}[d(X, Y)] \leq D} I(X; Y) \quad (263)$$

が結論される．よって，順定理と逆定理より

$$R(D) = \min_{Y: \mathbb{E}[d(X, Y)] \leq D} I(X; Y) \quad (264)$$

を満たす． $\square$

付録 B 微分エントロピーの最大化定理

本文中で用いた微分エントロピーの最大化定理について証明する．

証明 [7] $g(x)$ を任意の確率密度関数とする．分布 g の期待値をそれぞれ μ , σ^2 とおく． $\phi_K(x)$ を平均 μ , 分散 σ^2 の正規分布の確率密度関数とする．2 分布間のカルバック・ライブラー情報量を求めると

$$0 \leq D(g||\phi) \quad (265)$$

$$= \int_{-\infty}^{\infty} g(x) \log \left(\frac{g(x)}{\phi_K(x)} \right) dx \quad (266)$$

$$= -h(g) - \int_{-\infty}^{\infty} g(x) \log(\phi_K(x)) dx \quad (267)$$

$$(268)$$

となる．ここで

$$\int_{-\infty}^{\infty} g(x) \log(\phi_K(x)) dx \quad (269)$$

$$= \int_{-\infty}^{\infty} g(x) \log \left(\frac{1}{\sqrt{2\pi\sigma^2}} \exp \left(-\frac{(x-\mu)^2}{2\sigma^2} \right) \right) dx \quad (270)$$

$$= \log \left(\frac{1}{\sqrt{2\pi\sigma^2}} \right) \int_{-\infty}^{\infty} g(x) dx - \frac{1}{2\sigma^2} \log(e) \int_{-\infty}^{\infty} g(x) (x-\mu)^2 dx \quad (271)$$

$$= -\frac{1}{2} \log(2\pi\sigma^2) - \frac{1}{2} \log(e) \quad (272)$$

$$= -\frac{1}{2} \{ \log(2\pi\sigma^2) + \log(e) \} \quad (273)$$

$$= -\frac{1}{2} \log(2\pi e\sigma^2) \quad (274)$$

$$= -h(\phi_K) \quad (275)$$

となるので，式 (288) は

$$0 \leq -h(g) + h(\phi_K) \quad (276)$$

と置き換えられる．ここで期待値 0，分散 σ^2 の正規分布を ϕ_K とおけば，微分エントロピーは平均値を平行移動させても不変であるため，

$$h(\phi) = h(\phi_K) \quad (277)$$

となる．よって

$$h(g) \leq h(\phi_K) \quad (278)$$

を得る．ただし等号成立は g が正規分布のときである．以上より分散が σ^2 であるような連続分布の中で微分エントロピーを最大化するのは正規分布である． $\square$

付録 C 通常の相互情報量と定数倍された変数を持つ相互情報量の関係

式 (114)，式 (115) で示した式変形について証明する．

証明 まず $I(X; Y)$ と $I(aX; Y)$ について考える．このとき $a > 0$ とする．いま， $X \sim \mathcal{N}(\mu, \sigma^2)$ とすると $a > 0$ であるから， $aX \sim \mathcal{N}(a\mu, a^2\sigma^2)$ となる．相互情報量の定義より

$$I(aX; Y) = h(aX) - h(aX|Y) \quad (279)$$

となる．ここで

$$h(aX) = - \int_{-\infty}^{\infty} \mu_{aX}(x) \log \mu_{aX}(x) dx \quad (280)$$

$$= - \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} \exp\left(-\frac{x^2}{2a^2\sigma^2}\right) \ln \left\{ \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{x^2}{2a^2\sigma^2}\right) \right\} dx \quad (281)$$

$$= - \frac{a}{\sqrt{2\pi\sigma^2}a^2} \int_{-\infty}^{\infty} \exp\left(-\frac{x^2}{2a^2\sigma^2}\right) \left\{ -\frac{1}{2} \ln(2\pi\sigma^2) - \frac{x^2}{2(a\sigma)^2} \right\} dx \quad (282)$$

$$= \frac{a}{2} \ln(2\pi\sigma^2) + \frac{a}{\sqrt{2\pi\sigma^2}a^2} \int_{-\infty}^{\infty} \exp\left(-\frac{x^2}{2a^2\sigma^2}\right) \left(\frac{x^2}{2a^2\sigma^2} \right) dx \quad (283)$$

$$= \frac{a}{2} \ln(2\pi\sigma^2) + \frac{a}{2} \quad (284)$$

$$= \frac{a}{2} \ln(2\pi e\sigma^2) \quad (285)$$

$$= \frac{a}{2} \log(2\pi e\sigma^2) \quad (286)$$

$$= ah(x) \quad (287)$$

となる．次に

$$h(aX|Y) = - \iint \mu_{aX,Y}(x,y) \log \mu_{aX|Y}(x|y) dx dy \quad (288)$$

となる．ここで， $\mu_{aX,Y}(x,y), \mu_{aX|Y}(x|y)$ は

$$\mu_{aX,Y}(x,y) = Pr\{aX = x, Y = y\} \quad (289)$$

$$= Pr\{X = \frac{x}{a}, Y = y\} \quad (290)$$

$$= \mu_{XY}\left(\frac{x}{a}, y\right) \quad (291)$$

$$\mu_{aX|Y}(x|y) = Pr\{aX = x|Y = y\} \quad (292)$$

$$= \mu_{X|Y}\left(\frac{x}{a}|y\right) \quad (293)$$

となるので，式 (288) は

$$h(aX|Y) = - \iint \mu_{XY}\left(\frac{x}{a}, y\right) \mu_{X|Y}\left(\frac{x}{a}|y\right) dx dy \quad (294)$$

となる． $x' = \frac{x}{a}$ とおくと， $dx' = \frac{1}{a} dx$ となるから

$$- \iint \mu_{XY}\left(\frac{x}{a}, y\right) \mu_{X|Y}\left(\frac{x}{a}|y\right) dx dy \quad (295)$$

$$= -a \iint \mu_{XY}(x', y) \mu_{X|Y}(x'|y) dx' dy \quad (296)$$

$$= ah(X|Y) \quad (297)$$

と変形できる．したがって

$$I(aX; Y) = ah(X) - ah(X|Y) \quad (298)$$

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

となる．以上のことから式 (114) から式 (115) のように変形できる． $\square$