

信州大学工学部

学士論文

一人のエージェントを雇った CEO 問題における
達成可能領域の内界と外界について

指導教員 西新 幹彦 准教授

学科 電気電子工学科
学籍番号 13T2043H
氏名 丞原 篤司

2017 年 2 月 27 日

目次

1	はじめに	1
2	CEO 問題の概要	1
2.1	情報源, ノイズ, 出力シンボルの定義	1
2.2	歪みの定義	2
2.3	符号器と復号器の定義	3
2.4	達成可能性の定義	3
2.5	相互情報量の定義	3
3	CEO 問題	4
3.1	CEO 問題 逆定理	4
3.2	CEO 問題 順定理	7
4	考察	10
5	まとめ	12
付録 A	本文中で用いる定理等	14
A-1	レート歪み 逆定理	14
A-2	レート歪み 順定理	18
A-3	チェイン則	25
A-4	データ処理不等式	26
A-5	相互情報量の凸性	27
A-6	分散の有界性	28
A-7	不等式の証明	30
A-7-1	$(1-x)^y \leq e^{-xy}$ ($0 \leq x \leq 1, y \geq 0$) の導出	30
A-7-2	$e^{-xy} \leq 1-x+e^{-y}$ ($0 \leq x \leq 1, y \geq 0$) の導出	30
付録 B	記号一覧	30

1 はじめに

CEO 問題とはノイズを含んだ情報を観測した者がそれぞれ独立に符号を送信する際に通信が成立しうるレートと歪みの関係を導くことを目的とした問題であり, 実用的には多数のセンサを用いたセンサネットワークへの応用が期待できる. 本研究ではレート歪み理論を応用し, 単一の正規分布入力に対し正規分布のノイズが加えられたときのレートと入出力間の二乗誤差歪みの関係を考えた. その結果, 特殊な条件を与えたもてレートと歪みの関係について達成可能領域の内界と外界を導いた.

2 CEO 問題の概要

まず CEO 問題の設定について述べる. CEO がエージェントを用いてある情報を得ようとする. エージェントはそれぞれ独立に情報を観測するが, 正確には観測ができずいくらかノイズを含んだ形で得られる. そして得た情報は CEO へ向けてそれぞれ独立に送信されその情報の中から CEO が情報を得る.

CEO 問題は図 1 に示すようなモデルで表すことができる. 情報源 X に対しそれぞれ独立なノイズ N が加えられた L 個の信号 $\tilde{X} = X + N$ を $\|\varphi_n\|$ 個の符号語を持つ符号化しレート R で伝送しあるシンボル \hat{X} を復号する. このとき, レート R に対し通信を実現しうる X と \hat{X} の歪み D の関係について考える. CEO 問題は 1996 年に Berger ら [1] によって提起され, 離散無記憶情報源に対し無数の通信路で無歪み伝送を行うような有限のレートが存在しないことと, 無数の通信路と無限のレートにおける無歪み伝送の漸近的ふるまいを決定した. その後, 連続情報源に対し同様の結果を Viswanathan ら [2] が求めた. 本研究では図 2 に示す通り正規分布に従い発生する情報源に対し正規分布に従うようなノイズが含まれる単一の観測を行った場合について考える. 以下にて論文中で用いる設定や記号について定義する. 用いた記号については付録 B にまとめる.

2.1 情報源, ノイズ, 出力シンボルの定義

情報源 X は正規分布 $\mathcal{N}(0, \sigma_X^2)$ に従って生じる任意の実数値とする. また, 長さ n の系列を $X^n = X_1 X_2 \dots X_n$ と表す.

観測する際に生じるノイズ N は情報源とは独立であり正規分布 $\mathcal{N}(0, \sigma_N^2)$ に従うとする. さらに観測した値を $\tilde{X} = X + N \sim \mathcal{N}(0, \sigma_X^2 + \sigma_N^2)$, $\tilde{X}_1 = X_1 + N_1$, $\tilde{X}^n = \tilde{X}_1 \tilde{X}_2 \dots \tilde{X}_n$ と表す.

観測した値 \tilde{X} は符号器 φ_n にて符号化し復号器 ψ_n で復号される. この復号されたシンボル

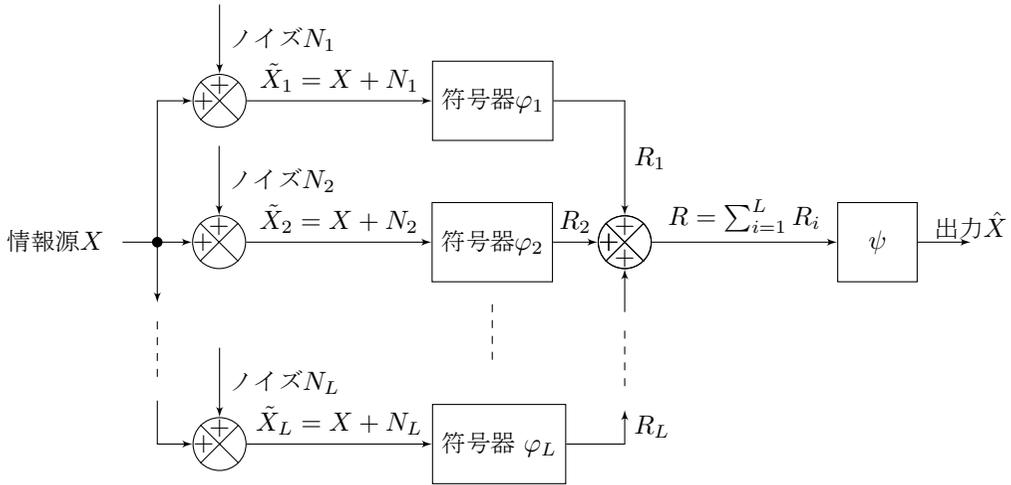


図1 一般の場合の CEO 問題の概要図

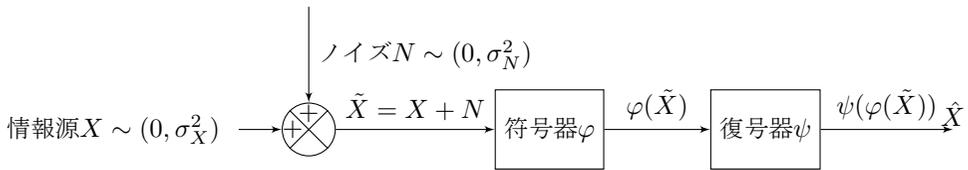


図2 X, N が正規分布に従い $L = 1$ と考えた場合の CEO 問題の概要図

を出力シンボル $\hat{X} = \psi(\varphi(\tilde{X}))$, $\hat{X}^n = \psi(\varphi(\tilde{X}^n))$ とする.

2.2 歪みの定義

この論文において歪みはすべて二乗誤差歪みを考える. 入力値を X 出力値を \hat{X} としたとき, その歪みを $d(X, \hat{X}) = (X - \hat{X})^2$ とする.

また 2 つの系列 $X^n = X_1 X_2 \dots X_n$, $\hat{X}^n = \hat{X}_1 \hat{X}_2 \dots \hat{X}_n$ の間の平均歪みを

$$d_n(X^n, \hat{X}^n) = \frac{1}{n} \sum_{i=1}^n d(X_i, \hat{X}_i) \quad (2.1)$$

と定義する.

2.3 符号器と復号器の定義

符号器 φ_n は入力されたシンボル $\tilde{X}^n \in \mathbb{R}^n$ を $\|\varphi_n\|$ 個の符号語を持つ符号で符号化する。復号器 ψ_n は符号器より受け取ったシンボルをあらかじめ定められたコードブック \mathcal{C} に従いそれぞれに対応する $\hat{X}^n \in \mathbb{R}^n$ へ復号する。

2.4 達成可能性の定義

符号器 φ_n , 復号器 ψ_n に対し

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log \|\varphi_n\| \leq R \quad (2.2)$$

$$\limsup_{n \rightarrow \infty} \mathbb{E} \left[d_n(X^n, \hat{X}^n) \right] \leq D \quad (2.3)$$

をそれぞれ満たすようなレート R と歪み D が存在するとき (R, D) は達成可能であるという。ただし, $\hat{X}^n = \psi_n(\varphi_n(\tilde{X}^n))$ である。また, 達成可能な (R, D) の組からなる領域を達成可能領域と定義する。

2.5 相互情報量の定義

この論文中で扱う対象は入力が連続値で出力が離散値であるようなシステムであることから入出力間の相互情報量を以下のように定義する。

X を連続確率変数, Y を離散確率変数とし X と Y の同時分布を $P_{XY}(x, y)$ とする。このとき相互情報量を

$$I(X; Y) \equiv \int_x \sum_y P_{XY}(x, y) \log \frac{P_{XY}(x, y)}{P_X(x)P_Y(y)} dx \quad (2.4)$$

と定義する。

また,

$$P_X(x) \equiv \sum_y P_{XY}(x, y) \quad (2.5)$$

$$P_Y(y) \equiv \int_x P_{XY}(x, y) dx \quad (2.6)$$

とする。ここで $P_X(x)$ は連続値, $P_Y(y)$ は離散値である。

さらに X を連続確率変数, Y, Z を離散確率変数とした時の相互情報量を

$$I(X; Y, Z) \equiv \int_x \sum_y \sum_z P_{XYZ}(x, y, z) \log \frac{P_{XYZ}(x, y, z)}{P_X(x)P_Y(y, z)} dx \quad (2.7)$$

と定義する.

相互情報量の定義を改めたことによって論文中で用いるエントロピーや確率に関する諸定理が成立することを付録 A-3,A-4,A-5 に示す.

3 CEO 問題

本研究の成果として $0 \leq \mathbb{E}[N\hat{X}] \leq \sigma_N^2$ という条件を付加することによりレート歪み関数 $R(D)$ は

$$\frac{1}{2} \log \frac{\sigma_X^2 + \sigma_N^2}{D + \sigma_N^2} \leq R(D) \leq \frac{1}{2} \log \frac{\sigma_X^2 + \sigma_N^2}{D - \sigma_N^2} \quad (3.1)$$

で表す領域に存在するということが示された. 以下より逆定理, 順定理の順でこの結果を導く.

3.1 CEO 問題 逆定理

定理 1 (CEO 問題 逆定理) 入力ノイズ N 出力 \hat{X} に対し, $0 \leq \mathbb{E}[N\hat{X}]$ が成り立つならば

$$R(D) \geq \inf_{\tilde{X}: \mathbb{E}[d(\tilde{X}, \hat{X})] \leq D + \sigma_N^2} I(\tilde{X}; \hat{X}) \quad (3.2)$$

が成り立つ.

(証明) 基本的な証明の流れは通常のレート歪み逆定理 (付録 A-1) と同様である. 式 (A.2) に対し入力情報源を \tilde{X}^n ($\tilde{X}_i = X_i + N_i$) と考えると式 (A.9) と同様にして

$$\frac{1}{n} \log \|\varphi_n\| \geq I(\tilde{X}_{Q_n}; \hat{X}_{Q_n}) \quad (3.3)$$

が得られる. ここで $Q_n \in \{1, \dots, n\}$, $P_{Q_n}(i) = \frac{1}{n}$ である.

次に歪みについて考える. 同様に式 (A.10) の入力を \tilde{X} と考えると

$$\mathbb{E} \left[d(\tilde{X}_{Q_n}, \hat{X}_{Q_n}) \right] = \mathbb{E} \left[d_n(\tilde{X}^n, \hat{X}^n) \right] \quad (3.4)$$

であり, 歪み $d(\tilde{X}_i, \hat{X}_i)$ は

$$\begin{aligned} d(\tilde{X}_i, \hat{X}_i) &= (\tilde{X}_i - \hat{X}_i)^2 \\ &= (X_i + N_i - \hat{X}_i)^2 \\ &= (X_i - \hat{X}_i)^2 + 2N_i(X_i - \hat{X}_i) + N_i^2 \end{aligned} \quad (3.5)$$

であるから

$$\begin{aligned}
\mathbb{E} \left[d_n(\tilde{X}^n, \hat{X}^n) \right] &= \frac{1}{n} \sum_{i=1}^n \mathbb{E} \left[d(\tilde{X}_i, \hat{X}_i) \right] \\
&= \frac{1}{n} \sum_{i=1}^n \mathbb{E} \left[(X_i - \hat{X}_i)^2 + 2N_i(X_i - \hat{X}_i) + N_i^2 \right] \\
&= \frac{1}{n} \sum_{i=1}^n \left\{ \mathbb{E} \left[(X_i - \hat{X}_i)^2 \right] + 2\mathbb{E} [N_i X_i] - 2\mathbb{E} [N_i \hat{X}_i] + \mathbb{E} [N_i^2] \right\}
\end{aligned} \tag{3.6}$$

となる. 入力 X_i と入力ノイズ N_i はそれぞれ独立であることから $\mathbb{E} [N_i X_i] = \mathbb{E} [N_i] \mathbb{E} [X_i] = 0$ であり $0 \leq \mathbb{E} [N \hat{X}]$ を用いると

$$\mathbb{E} \left[d_n(\tilde{X}^n, \hat{X}^n) \right] \leq \mathbb{E} \left[d_n(X^n, \hat{X}^n) \right] + \sigma_N^2 \tag{3.7}$$

であるから

$$\begin{aligned}
\limsup_{n \rightarrow \infty} \mathbb{E} \left[d_n(\tilde{X}^n, \hat{X}^n) \right] &\leq \limsup_{n \rightarrow \infty} \mathbb{E} \left[d_n(X^n, \hat{X}^n) \right] + \sigma_N^2 \\
&\leq D + \sigma_N^2
\end{aligned} \tag{3.8}$$

となり, 式 (3.4) とあわせると

$$\limsup_{n \rightarrow \infty} \mathbb{E} \left[d_n(\tilde{X}_{Q_n}, \hat{X}_{Q_n}) \right] \leq D + \sigma_N^2 \tag{3.9}$$

を得る. これは式 (A.11) に対応する. 以降は付録 A-1 と同様にして

$$R(D) \geq \inf_{\hat{X}: \mathbb{E}[d(\tilde{X}, \hat{X})] \leq D + \sigma_N^2} I(\tilde{X}; \hat{X}) \tag{3.10}$$

が得られる. □

ここで, 情報源 X とノイズ N がともに正規分布であることを使って, 式 (3.8) の右辺の閉じた形を求める. まず, $\mathbb{E}[d(\tilde{X}, Y)] \leq D + \sigma_N^2$ を満たす任意の連続確率変数 Y を考える. すると

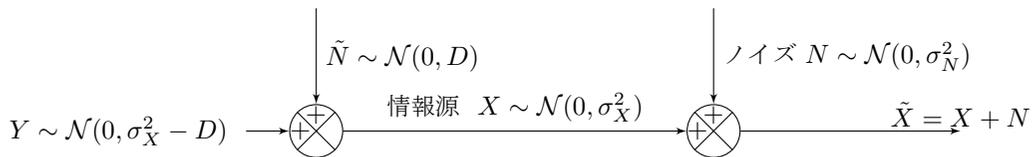


図3 $Y \sim \mathcal{N}(0, \sigma_X^2 - D)$, $\tilde{N} \sim \mathcal{N}(0, D)$ とし $X = Y + \tilde{N}$ としたときの Y と \tilde{X} の関係

相互情報量 $I(\tilde{X}; Y)$ は

$$\begin{aligned}
I(\tilde{X}; Y) &= \bar{h}(\tilde{X}) - h(\tilde{X}|Y) \\
&= \frac{1}{2} \log 2\pi e(\sigma_X^2 + \sigma_N^2) - h(\tilde{X} - Y|Y) \\
&\geq \frac{1}{2} \log 2\pi e(\sigma_X^2 + \sigma_N^2) - h(\tilde{X} - Y) \\
&\geq \frac{1}{2} \log 2\pi e(\sigma_X^2 + \sigma_N^2) - h(\mathcal{N}(0, \mathbb{E}[d(\tilde{X}, Y)])) \\
&\quad (\text{正規分布がエントロピーを最大にすることから}) \\
&= \frac{1}{2} \log 2\pi e(\sigma_X^2 + \sigma_N^2) - \frac{1}{2} \log 2\pi e(\mathbb{E}[d(\tilde{X}, Y)]) \\
&\geq \frac{1}{2} \log 2\pi e(\sigma_X^2 + \sigma_N^2) - \frac{1}{2} \log 2\pi e(D + \sigma_N^2) \\
&= \frac{1}{2} \log \frac{\sigma_X^2 + \sigma_N^2}{D + \sigma_N^2}
\end{aligned} \tag{3.11}$$

と評価されるので

$$\begin{aligned}
\inf_{Y: \mathbb{E}[d(\tilde{X}, Y)] \leq D + \sigma_N^2} I(\tilde{X}, Y) &\geq \inf_{Y: \mathbb{E}[d(\tilde{X}, Y)] \leq D + \sigma_N^2} \frac{1}{2} \log \frac{\sigma_X^2 + \sigma_N^2}{D + \sigma_N^2} \\
&= \frac{1}{2} \log \frac{\sigma_X^2 + \sigma_N^2}{D + \sigma_N^2}
\end{aligned} \tag{3.12}$$

となる.

次に $X = Y + \tilde{N} \sim \mathcal{N}(0, \sigma_X^2)$ となるよう互いに独立な確率変数 $Y \sim \mathcal{N}(0, \sigma_X^2 - D)$, $\tilde{N} \sim \mathcal{N}(0, D)$ を考える. このときのモデルを図3に示す.

まず, この Y について

$$\begin{aligned}
\mathbb{E}[d(\tilde{X}, Y)] &= \mathbb{E}[d(\tilde{X}, Y)] \\
&= \mathbb{E}[d(Y + \tilde{N} + N, Y)] \\
&= \mathbb{E}[(Y + \tilde{N} + N - Y)^2] \\
&= \mathbb{E}[(\tilde{N} + N)^2] \\
&= D + \sigma_N^2
\end{aligned} \tag{3.13}$$

が成り立つことから Y は $\mathbb{E}[d(\tilde{X}, Y)] \leq D + \sigma_N^2$ を満たす. この Y について相互情報量 $I(\tilde{X}; Y)$ を計算すると

$$\begin{aligned}
I(\tilde{X}; Y) &= h(\tilde{X}) - h(\tilde{X}|Y) \\
&= h(\tilde{X}) - h(\tilde{X} - Y|Y) \\
&= h(\tilde{X}) - h(\tilde{N} + N|Y) \\
&= h(\tilde{X}) - h(\tilde{N} + N) \\
&\quad (Y \text{ は } N, \tilde{N} \text{ それぞれに対し独立であるから}) \\
&= \frac{1}{2} \log 2\pi e(\sigma_X^2 + \sigma_N^2) - \frac{1}{2} \log 2\pi e(D + \sigma_N^2) \\
&= \frac{1}{2} \log \frac{\sigma_X^2 + \sigma_N^2}{D + \sigma_N^2}
\end{aligned} \tag{3.14}$$

となる. したがって

$$\inf_{Y: \mathbb{E}[d(\tilde{X}, Y)] \leq D + \sigma_N^2} I(\tilde{X}, Y) \leq \frac{1}{2} \log \frac{\sigma_X^2 + \sigma_N^2}{D + \sigma_N^2} \tag{3.15}$$

が成り立つ. 式 (3.12), (3.15) より, 式 (3.10) の右辺の閉じた形として

$$\inf_{Y: \mathbb{E}[d(\tilde{X}, Y)] \leq D + \sigma_N^2} I(\tilde{X}, Y) = \frac{1}{2} \log \frac{\sigma_X^2 + \sigma_N^2}{D + \sigma_N^2} \tag{3.16}$$

が得られる. さらに, 式 (3.16) を式 (3.10) に代入して

$$\frac{1}{2} \log \frac{\sigma_X^2 + \sigma_N^2}{D + \sigma_N^2} \leq R(D) \tag{3.17}$$

を得る. したがって, 式 (3.1) の左側の不等式が示された.

3.2 CEO 問題 順定理

定理 2 (CEO 問題 順定理) 入力ノイズ N 出力 \hat{X} に対し $\mathbb{E}[N\hat{X}] \leq \sigma_N^2$ が成り立つならば

$$R(D) \leq \inf_{\hat{X}: \mathbb{E}[d(\tilde{X}, \hat{X})] \leq D - \sigma_N^2} I(\tilde{X}; \hat{X}) \tag{3.18}$$

が成り立つ.

(証明) 基本的な証明の流れは通常のレート歪み理論の順定理 (付録 A-2) と同様である. 任意の $\gamma > 0$ を考え $\mathbb{E}[d(\tilde{X}, Y)] \leq D - \sigma_N^2$ となる任意の Y に対して $R \equiv I(\tilde{X}, Y) + 2\gamma$ とする. さらに復号語数を $M_n \equiv e^{nR} = e^{n(I(\tilde{X}; Y) + 2\gamma)}$ とし Y^n の分布に従いランダム符号化によって復号語を生成する.

生成した符号器 φ_n と復号器 ψ_n の性能を評価する. 式 (A.27) と同じく $\hat{X}^n = \psi_n(\varphi_n(\tilde{X}^n))$ とおいて $\bar{P}_e^{(n)}(u)$ を

$$\bar{P}_e^{(n)}(u) = \Pr \left\{ d_n(\tilde{X}^n, \hat{X}^n) > u \right\} \quad (3.19)$$

定義すると式 (A.60) と同様に

$$\limsup_{n \rightarrow \infty} \mathbb{E} \left[d_n(\tilde{X}^n, \hat{X}^n) \right] \leq D - \sigma_N^2 \quad (3.20)$$

が得られる.

一方, 式 (3.5) を変形すると

$$d(X_i, \hat{X}_i) = d(\tilde{X}_i, \hat{X}_i) - 2N_i(X_i - \hat{X}_i) - N_i^2 \quad (3.21)$$

であるから

$$\begin{aligned} \mathbb{E} \left[d_n(X^n, \hat{X}^n) \right] &= \frac{1}{n} \sum_{i=1}^n \mathbb{E} \left[d(X_i, \hat{X}_i) \right] \\ &= \frac{1}{n} \sum_{i=1}^n \mathbb{E} \left[d(\tilde{X}_i, \hat{X}_i) - 2N_i(X_i - \hat{X}_i) - N_i^2 \right] \\ &= \frac{1}{n} \sum_{i=1}^n \left\{ \mathbb{E} \left[(\tilde{X}_i - \hat{X}_i)^2 \right] - 2\mathbb{E} [N_i X_i] + 2\mathbb{E} [N_i \hat{X}_i] - \mathbb{E} [N_i^2] \right\} \end{aligned} \quad (3.22)$$

となる. 入力 X_i と入力ノイズ N_i はそれぞれ独立であることから $\mathbb{E} [N_i X_i] = \mathbb{E} [N_i] \mathbb{E} [X_i] = 0$ であり $\mathbb{E} [N \hat{X}] \leq \sigma_N^2$ を用いると

$$\begin{aligned} \mathbb{E} \left[d_n(X^n, \hat{X}^n) \right] &\leq \frac{1}{n} \sum_{i=1}^n \left\{ \mathbb{E} \left[d(\tilde{X}_i, \hat{X}_i) + 2\sigma_N^2 - \sigma_N^2 \right] \right\} \\ &= \frac{1}{n} \sum_{i=1}^n \mathbb{E} \left[d(\tilde{X}_i, \hat{X}_i) \right] + \sigma_N^2 \\ &= \mathbb{E} \left[d_n(\tilde{X}^n, \hat{X}^n) \right] + \sigma_N^2 \end{aligned} \quad (3.23)$$

が得られる. 式 (3.20), (3.23) より符号の歪みに関する評価として

$$\limsup_{n \rightarrow \infty} \mathbb{E} \left[d_n(X^n, \hat{X}^n) \right] \leq D \quad (3.24)$$

が得られる.

次に符号の符号化レートに関しては式 (A.61) と同様に

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log \|\varphi_n\| \leq R \quad (3.25)$$

が成り立つので (R, D) は達成可能である. $R = I(\tilde{X}, Y) + 2\gamma$ だったことから

$$R(D) \leq I(\tilde{X}, Y) + 2\gamma \quad (3.26)$$

を得る. Y は $\mathbb{E}[d(\tilde{X}, Y)] \leq D - \sigma_N^2$ を満たす任意の確率変数だったので

$$R(D) \leq \inf_{Y: \mathbb{E}[d(\tilde{X}, Y)] \leq D - \sigma_N^2} I(\tilde{X}; Y) + 2\gamma \quad (3.27)$$

が成り立ち, $\gamma > 0$ は任意であったことから

$$R(D) \leq \inf_{Y: \mathbb{E}[d(\tilde{X}, Y)] \leq D - \sigma_N^2} I(\tilde{X}; Y) \quad (3.28)$$

が得られる. □

ここから, 情報源 X とノイズ N がともに正規分布であることを使って, 式 (3.18) の右辺の閉じた形を求める. まず, $\mathbb{E}[d(\tilde{X}, Y)] \leq D - \sigma_N^2$ を満たす任意の連続確率変数 Y を考える. 相互情報量 $I(\tilde{X}; Y)$ が達成可能であることを逆定理の場合と同様に評価すると

$$\begin{aligned} I(\tilde{X}; Y) &\geq \frac{1}{2} \log 2\pi e(\sigma_X^2 + \sigma_N^2) - \frac{1}{2} \log 2\pi e(\mathbb{E}[d(\tilde{X}, Y)]) \\ &\geq \frac{1}{2} \log 2\pi e(\sigma_X^2 + \sigma_N^2) - \frac{1}{2} \log 2\pi e(D - \sigma_N^2) \\ &= \frac{1}{2} \log \frac{\sigma_X^2 + \sigma_N^2}{D - \sigma_N^2} \end{aligned} \quad (3.29)$$

が言えるので

$$\begin{aligned} \inf_{Y: \mathbb{E}[d(\tilde{X}, Y)] \leq D - \sigma_N^2} I(\tilde{X}, Y) &\geq \inf_{Y: \mathbb{E}[d(\tilde{X}, Y)] \leq D - \sigma_N^2} \frac{1}{2} \log \frac{\sigma_X^2 + \sigma_N^2}{D - \sigma_N^2} \\ &= \frac{1}{2} \log \frac{\sigma_X^2 + \sigma_N^2}{D - \sigma_N^2} \end{aligned} \quad (3.30)$$

となる.

次に $X = Y + \tilde{N} \sim \mathcal{N}(0, \sigma_N^2)$ となるよう互いに独立な確率変数 $Y \sim \mathcal{N}(0, \sigma_X^2 - D + 2\sigma_N^2)$, $\tilde{N} \sim \mathcal{N}(0, D - 2\sigma_N^2)$ を考える. このときのモデルを図 4 に示す.

まずこの Y について

$$\begin{aligned} \mathbb{E}[d(\tilde{X}, Y)] &= \mathbb{E}[d(X + N, Y)] \\ &= \mathbb{E}[d(Y + \tilde{N} + N, Y)] \\ &= \mathbb{E}[(Y + \tilde{N} + N - Y)^2] \\ &= \mathbb{E}[(\tilde{N} + N)^2] \\ &= D - \sigma_N^2 \end{aligned} \quad (3.31)$$

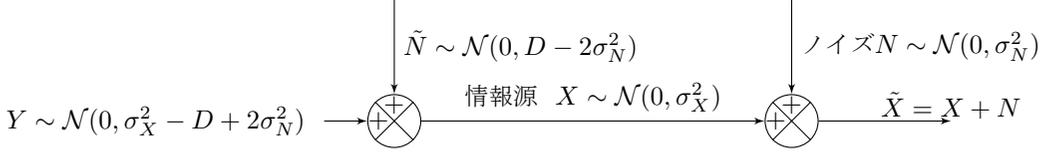


図4 $Y \sim \mathcal{N}(0, \sigma_X^2 - D + 2\sigma_N^2)$, $\tilde{N} \sim \mathcal{N}(0, D - 2\sigma_N^2)$ とし
 $X = Y + \tilde{N}$ としたときの Y と \tilde{X} の関係

であるから Y は $\mathbb{E}[d(\tilde{X}, Y)] \leq D - \sigma_N^2$ を満たす. この Y について相互情報量 $I(\tilde{X}; Y)$ を計算すると

$$\begin{aligned}
 I(\tilde{X}; Y) &= h(\tilde{X}) - h(\tilde{X}|Y) \\
 &= h(\tilde{X}) - h(\tilde{X} - Y|Y) \\
 &= h(\tilde{X}) - h(\tilde{N} + N|Y) \\
 &= h(\tilde{X}) - h(\tilde{N} + N) \\
 &\quad (Y \text{ は } N, \tilde{N} \text{ それぞれに対し独立であるから}) \\
 &= \frac{1}{2} \log 2\pi e(\sigma_X^2 + \sigma_N^2) - \frac{1}{2} \log 2\pi e(D - \sigma_N^2) \\
 &= \frac{1}{2} \log \frac{\sigma_X^2 + \sigma_N^2}{D - \sigma_N^2}
 \end{aligned} \tag{3.32}$$

が得られる. したがって

$$\inf_{Y: \mathbb{E}[d(\tilde{X}, Y)] \leq D - \sigma_N^2} I(\tilde{X}, Y) \leq \frac{1}{2} \log \frac{\sigma_X^2 + \sigma_N^2}{D - \sigma_N^2} \tag{3.33}$$

が成り立つ. 式 (3.30), (3.33) より, 式 (3.28) の右辺の閉じた形として

$$\inf_{Y: \mathbb{E}[d(\tilde{X}, Y)] \leq D - \sigma_N^2} I(\tilde{X}, Y) = \frac{1}{2} \log \frac{\sigma_X^2 + \sigma_N^2}{D - \sigma_N^2} \tag{3.34}$$

が得られる. 式 (3.34) を式 (3.28) に代入して

$$R(D) \leq \frac{1}{2} \log \frac{\sigma_X^2 + \sigma_N^2}{D - \sigma_N^2} \tag{3.35}$$

を得る. したがって, 式 (3.1) の右側の不等式が示された.

4 考察

順定理, 逆定理によって得た関数に対しそれぞれ $\sigma_X^2 = 1, \sigma_N^2 = 0.09$ を代入したときの $D, R(D)$ グラフを図5に示す. 図中実線が順定理によって得た結果であり実線より右上の領域

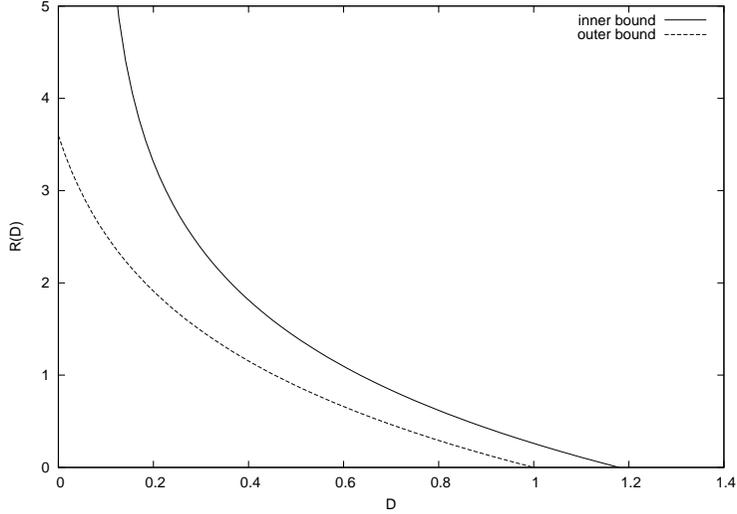


図5 $\sigma_X^2 = 1, \sigma_N^2 = 0.09$ としたときのレート歪み達成領域

が達成可能領域, すなわち内界である. 破線が逆定理によって得た結果であり破線より左下の領域が達成不可能な領域, すなわち外界を表す. 実線と破線の間領域は達成可能であるかどうかかわからない. 得た結果からは判別不能な領域である.

さらに, 式 (3.1) からわかることについて考察する. 式 (3.1) を変形すると

$$\frac{\sigma_X^2 + \sigma_N^2}{D + \sigma_N^2} \leq 2^R \leq \frac{\sigma_X^2 + \sigma_N^2}{D - \sigma_N^2} \quad (4.1)$$

$$\therefore 2^{-R}(\sigma_X^2 + \sigma_N^2) - \sigma_N^2 \leq D \leq 2^{-R}(\sigma_X^2 + \sigma_N^2) + \sigma_N^2$$

が得られる. レート R について極限 $R \rightarrow \infty$ を考えると,

$$-\sigma_N^2 \leq D \leq \sigma_N^2 \quad (4.2)$$

となり, 二乗誤差歪みであることから $D \geq 0$ であることを考慮すると,

$$0 \leq D \leq \sigma_N^2 \quad (4.3)$$

である. $R \rightarrow \infty$ という条件は符号器が受け取った情報をそのまま伝送したとみなすことができ, その場合に外部からのノイズ分だけ歪みが生じるという直観的な見解と矛盾しない.

次に, 関数を決定できなかった理由について考える. 関数を決定することができなかったのは式 (3.5), (3.23) において歪みに関する不等式を $0 \leq \mathbb{E}[N\hat{X}] \leq \sigma_N^2$ という条件なしに導出できなかったことに由来する. 歪みについて解くことができなかった一番の理由は二乗誤差歪みが優加法的な性質を持っていたことである. エージェント一人に対する CEO 問題をレート歪

み理論に適用するとレート歪み理論における符号器復号器間の歪みを $\mathbb{E}[d(\tilde{X}, \hat{X})]$ で表さなければならない。達成可能の条件式 (2.3) を用いて CEO 問題における符号器復号器間の歪み $\mathbb{E}[d(\tilde{X}, \hat{X})]$ を表すことを考える。歪み測度が優加法的であることを考慮すると考えられる歪みの式の例として

$$\mathbb{E}[d(\tilde{X}, \hat{X})] \geq \mathbb{E}[d(X, \hat{X})] + \mathbb{E}[d(X, \tilde{X})] \quad (4.4)$$

となる場合がある。このとき符号器復号器間の歪みを抑える形で式を表すことができず、式 (3.2), (3.18) などが成り立たなかった。しかし、 $0 \leq \mathbb{E}[N\hat{X}] \leq \sigma_N^2$ という条件を加えることによってレート歪み関数を決定することまではできなかったがある範囲内まで絞り込むことができた。

また $\mathbb{E}[N\hat{X}]$ について触れておく。達成可能性を表す数式を導くためにこの値について条件を加えた。数式の上ではノイズと出力の積の期待値であるがこの値が工学的にどのような意味を持つかわからず、今回用いたような $0 \leq \mathbb{E}[N\hat{X}] \leq \sigma_N^2$ という条件を満たす符号器復号器の組が存在するかどうかはわからない。今後同様の問題を考えるとすると $\mathbb{E}[N\hat{X}]$ を決定する必要があるが現在の問題設定下では Y は N に対し独立でないこと以外に Y に対する条件をほぼ与えていないことから数値を特定することは難しいと考えられる。

最後に今回の結果をより一般の場合、すなわちエージェントが一般の L 人の場合に拡張することを少し考える。エージェント $L \geq 2$ の場合では符号器からそれぞれ独立に送られてくる情報を復号器で一つのシンボルに変換しなければならない。そのため現在の問題設定と異なり制限されたレート下で複数送られてくる情報に対しどのような復号を行うと歪みを抑えることができるかという問題が出てくる。そしてこの問題こそが CEO 問題の本質であり難しい問題としているのではないかと考える。このことから本研究で得た結果をそのまま一般の場合に拡張することはできないと考える。

5 まとめ

本研究では正規分布情報源に対しエージェント 1 人が正規分布に従うノイズを含む観測を行った場合の達成可能なレートと歪みの関係について考えた。その結果、ノイズ N と出力 \hat{X} に対し $0 \leq \mathbb{E}[N\hat{X}] \leq \sigma_N^2$ という条件を与えたもとの、レートと歪みの達成可能な領域の内界と外界を絞り込むことができた。条件を与えなければ領域を特定することができなかった原因の一つとして、歪み測度として用いた二乗関数が優加法的な性質を持っていたことがあげられる。また、 $\mathbb{E}[N\hat{X}]$ という式が工学的に何を表しているのかは分からず、現在の問題設定下では値を特定することができないと考えられる。さらに、一般の場合に拡張する場合は現在の問題設定にはない新たな問題を考える必要があり結果をそのまま拡張することは難しいと考えられる。

参考文献

- [1] T.Berger, Z.Zhang, H.Viswanathan, “The CEO Problem,” *IEEE Trans. Inform. Theory*, vol.42, pp. 887–902, May. 1996.
- [2] H.Viswanathan, T.Berger, “The quadratic Gaussian CEO problem,” *IEEE Trans. Inform. Theory*, vol.43, pp. 1549–1559, Sep. 1997.
- [3] 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 年.
- [4] T.M.Cover, J.A.Thomas, 山本 博資ほか訳, 情報理論-基礎と広がり, 共立出版, 2013 年.

付録 A 本文中で用いる定理等

本文中にて引用した連続入力離散出力に対するレート歪み理論の逆定理と順定理をそれぞれ (A-1),(A-2) に示す. またレート歪み逆定理, 順定理それぞれにおいて用いる補助定理を (A-3) から (A-7) にそれぞれ示す.

A-1 レート歪み 逆定理

(A-1,A-2 はいずれも [3][4] を参考に導出した.)

定理 3 (レート歪み 逆定理) レート歪み関数 $R(D)$ は

$$R(D) \geq \inf_{\hat{X}: \mathbb{E}[d(X, \hat{X})] \leq D} I(X; \hat{X}) \quad (\text{A.1})$$

を満たす.

(証明) レート R が D -達成可能であるとする. よってある符号の列 $\{(\varphi_n, \psi_n)\}_{n=0}^{\infty}$ が存在し

$$\begin{aligned} \limsup_{n \rightarrow \infty} \frac{1}{n} \log \|\varphi_n\| &\leq R \\ \limsup_{n \rightarrow \infty} \mathbb{E} \left[d_n(X^n, \hat{X}^n) \right] &\leq D \end{aligned}$$

が成り立つ. このとき

$$\begin{aligned} \log \|\varphi_n\| &\geq H(\varphi_n(X^n)) \\ &\geq H(\varphi_n(X^n)) - H(\varphi_n(X^n) | X^n) \\ &= I(X^n; \varphi_n(X^n)) \\ &\geq I(X^n; \hat{X}^n) \quad (\text{A-4: データ処理不等式より}) \\ &= h(X^n) - h(X^n | \hat{X}^n) \\ &= \sum_{i=1}^n h(X_i) - h(X^n | \hat{X}^n) \\ &= \sum_{i=1}^n h(X_i) - \sum_{i=1}^n h(X_i | \hat{X}^n, X^{i-1}) \\ &\quad (\text{A-3: 微分エントロピーのチェイン則より}) \\ &\geq \sum_{i=1}^n h(X_i) - \sum_{i=1}^n h(X_i | \hat{X}_i) \\ &= \sum_{i=1}^n I(X_i; \hat{X}_i) \end{aligned} \quad (\text{A.2})$$

が成り立つ. ここで X^n と独立である確率変数 $Q_n \in \{1, \dots, n\}$, $P_{Q_n}(i) = \frac{1}{n}$ を導入する. このとき

$$\begin{aligned}
I(X_{Q_n}; \hat{X}_{Q_n} | Q_n) &= \mathbb{E}_{X_{Q_n} \hat{X}_{Q_n} Q_n} \left[\log \frac{P_{X_{Q_n} \hat{X}_{Q_n} | Q_n}(X_{Q_n}, \hat{X}_{Q_n} | Q_n)}{P_{X_{Q_n} | Q_n}(X_{Q_n} | Q_n) P_{\hat{X}_{Q_n} | Q_n}(\hat{X}_{Q_n} | Q_n)} \right] \\
&= \sum_{i=1}^n P_{Q_n}(i) \mathbb{E}_{X_{Q_n} \hat{X}_{Q_n}} \left[\log \frac{P_{X_{Q_n} \hat{X}_{Q_n} | Q_n}(X_{Q_n}, \hat{X}_{Q_n} | i)}{P_{X_{Q_n} | Q_n}(X_{Q_n} | i) P_{\hat{X}_{Q_n} | Q_n}(\hat{X}_{Q_n} | i)} \right] \\
&= \frac{1}{n} \sum_{i=1}^n \mathbb{E} \left[\log \frac{P_{X_i \hat{X}_i}(X_i, \hat{X}_i)}{P_{X_i}(X_i) P_{\hat{X}_i}(\hat{X}_i)} \right] \\
&= \frac{1}{n} \sum_{i=1}^n I(X_i; \hat{X}_i)
\end{aligned} \tag{A.3}$$

であるから

$$\log \|\varphi_n\| \geq nI(X_{Q_n}; \hat{X}_{Q_n} | Q_n) \tag{A.4}$$

となる.

また, Q_n と X が独立であることを示す.

$$\begin{aligned}
P_X(x) &= \sum_{j=1}^n P_{X_{Q_n}}(x, j) \\
&= \sum_{j=1}^n P_{X_j Q_n}(x, j) \\
&= \sum_{j=1}^n P_{X_j}(x) P_{Q_n}(j)
\end{aligned} \tag{A.5}$$

であり X^n の定常性から

$$\begin{aligned}
P_X(x) &= P_{X_i}(x) \sum_{j=1}^n P_{Q_n}(j) \\
&= P_{X_i}(x)
\end{aligned} \tag{A.6}$$

であるためそれぞれ入力確率密度が情報源の確率密度と等しいことがわかる. ゆえに

$$\begin{aligned}
P_{X_{Q_n}}(x, i) &= P_{X_i Q_n}(x, i) \\
&= P_{X_i}(x) P_{Q_n}(i) \\
&= P_X(x) P_{Q_n}(i)
\end{aligned} \tag{A.7}$$

であることから次式を示すことができる.

$$\begin{aligned}
I(X_{Q_n}; \hat{X}_{Q_n} | Q_n) &= \mathbb{E} \left[\log \frac{P_{X_{Q_n} \hat{X}_{Q_n} | Q_n}(X_{Q_n}, \hat{X}_{Q_n} | Q_n) P_{Q_n}(Q^n)}{P_{X_{Q_n} | Q_n}(X_{Q_n} | Q_n) P_{\hat{X}_{Q_n} | Q_n}(\hat{X}_{Q_n} | Q_n) P_{Q_n}(Q^n)} \right] \\
&= \mathbb{E} \left[\log \frac{P_{X_{Q_n} \hat{X}_{Q_n} | Q_n}(X_{Q_n}, \hat{X}_{Q_n} | Q_n) P_{Q_n}(Q^n)}{P_{X_{Q_n}}(X_{Q_n}) P_{\hat{X}_{Q_n} | Q_n}(\hat{X}_{Q_n} | Q_n) P_{Q_n}(Q^n)} \right] \\
&= \mathbb{E} \left[\log \frac{P_{X_{Q_n} \hat{X}_{Q_n} Q_n}(X_{Q_n}, \hat{X}_{Q_n}, Q_n)}{P_{X_{Q_n}}(X_{Q_n}) P_{\hat{X}_{Q_n} Q_n}(\hat{X}_{Q_n}, Q_n)} \right] \tag{A.8} \\
&= I(X_{Q_n}; \hat{X}_{Q_n} Q_n) \\
&= h(X_{Q_n}) - h(X_{Q_n} | \hat{X}_{Q_n} Q_n) \\
&\geq h(X_{Q_n}) - h(X_{Q_n} | \hat{X}_{Q_n}) \\
&= I(X_{Q_n}; \hat{X}_{Q_n})
\end{aligned}$$

よって

$$\frac{1}{n} \log \|\varphi_n\| \geq I(X_{Q_n}; \hat{X}_{Q_n}) \tag{A.9}$$

が示される.

また歪みについて考えると

$$\begin{aligned}
\mathbb{E} \left[d_n(X^n, \hat{X}^n) \right] &= \mathbb{E} \left[\frac{1}{n} \sum_{i=1}^n d(X_i, \hat{X}_i) \right] \\
&= \sum_{i=1}^n \frac{1}{n} \mathbb{E} \left[d(X_i, \hat{X}_i) \right] \\
&= \sum_{i=1}^n P_{Q_n}(i) \mathbb{E} \left[d(X_{Q_n}, \hat{X}_{Q_n}) | Q_n = i \right] \\
&= \sum_{i=1}^n P_{Q_n}(i) \int_x \sum_{\hat{x}} P_{X_{Q_n} \hat{X}_{Q_n} | Q_n}(x, \hat{x} | i) d(x, \hat{x}) dx \tag{A.10} \\
&= \int_x \sum_{\hat{x}} \sum_{i=1}^n P_{Q_n}(i) P_{X_{Q_n} \hat{X}_{Q_n} | Q_n}(x, \hat{x} | i) d(x, \hat{x}) dx \\
&= \int_x \sum_{\hat{x}} P_{X_{Q_n} \hat{X}_{Q_n}}(x, \hat{x}) d(x, \hat{x}) dx \\
&= \mathbb{E} \left[d(X_{Q_n}, \hat{X}_{Q_n}) \right]
\end{aligned}$$

$$\therefore \limsup_{n \rightarrow \infty} \mathbb{E} \left[d(X_{Q_n}, \hat{X}_{Q_n}) \right] \leq D \tag{A.11}$$

であるから $\forall \gamma > 0, \exists n_0 \in \mathbb{N}, \forall n > n_0$ に対し

$$\mathbb{E} \left[d(X_{Q_n}, \hat{X}_{Q_n}) \right] \leq D + \gamma \tag{A.12}$$

を満たす。式 (A.4) より $\exists n_0 \in \mathbb{N}, \forall n > n_0$ に対し

$$\begin{aligned} \frac{1}{n} \log \|\varphi_n\| &\geq I(X_{Q_n}; \hat{X}_{Q_n}) \geq \inf_{Y: \mathbb{E}[d(X_{Q_n}, \hat{X}_{Q_n})] \leq D+\gamma} I(X_{Q_n}; \hat{X}_{Q_n}) \\ &= \inf_{Y: \mathbb{E}[d(X, \hat{X})] \leq D+\gamma} I(X; \hat{X}) \end{aligned} \quad (\text{A.13})$$

となる。さらに達成可能であることから

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log \|\varphi_n\| \leq R \quad (\text{A.14})$$

であり $\forall \delta > 0, \exists n_0 \in \mathbb{N}, \forall n > n_0$ に対して

$$\frac{1}{n} \log \|\varphi_n\| < R + \delta \quad (\text{A.15})$$

となるから式 (A.13), (A.15) より $\exists n_0 \in \mathbb{N}, \forall n > n_0$ に対し

$$R + \delta > \inf_{Y: \mathbb{E}[d(X, \hat{X})] \leq D+\gamma} I(X; \hat{X}) \quad (\text{A.16})$$

である。ここで

$$f(D) \equiv \inf_{Y: \mathbb{E}[d(X, \hat{X})] \leq D} I(X; \hat{X}) \quad (\text{A.17})$$

とおき, $f(D)$ が D に関して下に凸な関数であることを示すことで連続であることを示す。

$\alpha_1 + \alpha_2 = 1, \alpha_1 \geq 0, \alpha_2 \geq 0$ を満たす α_1, α_2 と $\forall D_1 \geq 0, \forall D_2 \geq 0$ を考え $D = \alpha_1 D_1 + \alpha_2 D_2$ とする。また $\mathbb{E}[d(X, \hat{X}_1)] \leq D_1, \mathbb{E}[d(X, \hat{X}_2)] \leq D_2$ をそれぞれ満たす \hat{X}_1, \hat{X}_2 に対し $P_Q(i) \equiv \alpha_i$ なる Q を導入して $\hat{X} \equiv \hat{X}_Q$ とすると

$$\begin{aligned} \mathbb{E}[d(X, \hat{X})] &= \Pr\{Q = 1\} \mathbb{E}[d(X, \hat{X}_1)|Q = 1] + \Pr\{Q = 2\} \mathbb{E}[d(X, \hat{X}_2)|Q = 2] \\ &= \alpha_1 \mathbb{E}[d(X, \hat{X}_1)] + \alpha_2 \mathbb{E}[d(X, \hat{X}_2)] \\ &\leq \alpha_1 D_1 + \alpha_2 D_2 \\ &= D \end{aligned} \quad (\text{A.18})$$

$$\begin{aligned} \alpha_1 f(D_1) + \alpha_2 f(D_2) &= \inf_{\hat{X}_1: \mathbb{E}[d(X, \hat{X}_1)] \leq D_1} \alpha_1 I(X; \hat{X}_1) + \inf_{\hat{X}_2: \mathbb{E}[d(X, \hat{X}_2)] \leq D_2} \alpha_2 I(X; \hat{X}_2) \\ &= \inf_{\substack{\hat{X}_1: \mathbb{E}[d(X, \hat{X}_1)] \leq D_1 \\ \hat{X}_2: \mathbb{E}[d(X, \hat{X}_2)] \leq D_2}} \left\{ \alpha_1 I(X; \hat{X}_1) + \alpha_2 I(X; \hat{X}_2) \right\} \end{aligned} \quad (\text{A.19})$$

(A-5) より相互情報量は凸性を持つことから

$$\begin{aligned} \inf_{\substack{\hat{X}_1: \mathbb{E}[d(X, \hat{X}_1)] \leq D_1 \\ \hat{X}_2: \mathbb{E}[d(X, \hat{X}_2)] \leq D_2}} \left\{ \alpha_1 I(X; \hat{X}_1) + \alpha_2 I(X; \hat{X}_2) \right\} &\geq \inf_{\substack{\hat{X}_1: \mathbb{E}[d(X, \hat{X}_1)] \leq D_1 \\ \hat{X}_2: \mathbb{E}[d(X, \hat{X}_2)] \leq D_2}} I(X; \hat{X}) \\ &\geq \inf_{\hat{X}: \mathbb{E}[d(X, \hat{X})] \leq D} I(X; \hat{X}) \\ &= f(D) \end{aligned} \quad (\text{A.20})$$

より $f(D)$ が D に対して下に凸な関数であることが示され連続であることが分かった。よって

$$R(D) \geq \inf_{\hat{X}: \mathbb{E}[d(X, \hat{X})] \leq D} I(X; \hat{X}) \quad (\text{A.21})$$

が示された。□

A-2 レート歪み 順定理

定理 4 (レート歪み 順定理) レート歪み関数 $R(D)$ は

$$R(D) \leq \inf_{Y: \mathbb{E}[d(X, Y)] \leq D} I(X; Y) \quad (\text{A.22})$$

を満たす。

(証明) Y は $\mathbb{E}[d(X, Y)] \leq D$ を満たす任意の確率変数とする。このとき任意の $\gamma > 0$ に対して $R \equiv I(X, Y) + 2\gamma$ が D -達成可能であることを示す。

$X^n Y^n = (X_1 Y_1, X_2 Y_2, \dots, X_n Y_n)$ を XY の定常かつ独立な系列であると考え

$$T_n^{(1)} \equiv \left\{ (\mathbf{x}, \mathbf{y}) \in \mathcal{X}^n \times \mathcal{Y}^n \mid \left| \frac{1}{n} \log \frac{P_{Y^n|X^n}(\mathbf{y}|\mathbf{x})}{P_{Y^n}(\mathbf{y})} - I(X; Y) \right| < \gamma \right\} \quad (\text{A.23})$$

$$T_n^{(2)}(u) \equiv \{(\mathbf{x}, \mathbf{y}) \in \mathcal{X}^n \times \mathcal{Y}^n \mid d_n(\mathbf{x}, \mathbf{y}) \leq u\} \quad (\text{A.24})$$

と定める。入出力系列が定常かつ独立であることから

$$\frac{1}{n} \log \frac{P_{Y^n|X^n}(Y^n|X^n)}{P_{Y^n}(Y^n)} = \frac{1}{n} \sum_{i=1}^n \log \frac{P_{Y_i|X_i}(Y_i|X_i)}{P_{Y_i}(Y_i)} \quad (\text{A.25})$$

となりその期待値は

$$\begin{aligned} \mathbb{E} \left[\sum_{i=1}^n \log \frac{P_{Y_i|X_i}(Y_i|X_i)}{P_{Y_i}(Y_i)} \right] &= \sum_{i=1}^n \mathbb{E} \left[\log \frac{P_{Y_i|X_i}(Y_i|X_i)}{P_{Y_i}(Y_i)} \right] \\ &= \sum_{i=1}^n I(X_i; Y_i) \\ &= nI(X; Y) \end{aligned} \quad (\text{A.26})$$

となる。また (付録 A-6) より分散は有界であることから $T_n^{(1)}$ に対しチェビシエフの不等式を用いて n を限りなく大きくすると

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

が成り立つ. 次に符号器 φ_n を作る. 復号語の数を $M_n \equiv e^{nR} = e^{n(I(X;Y)+2\gamma)}$ とし Y^n の分布に従い M_n 個の復号語 $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{M_n} \in \mathcal{Y}^n$ を決める. 入力 $\mathbf{x} \in \mathcal{X}^n$ に対して

$$d_n(\mathbf{x}, \mathbf{0}) \geq \min_{1 \leq i \leq M_n} d_n(\mathbf{x}, \mathbf{v}_i)$$

を満たす場合, すなわち基準語 $\mathbf{0}$ よりも近い復号語が存在するとき最も近い復号語のインデックス i_0 を用いて $\varphi_n(\mathbf{x}) \equiv i_0$ としどの復号語よりも基準語に近いとき $\varphi_n(\mathbf{x}) = M_n + 1$ とする. また復号器 ψ_n は i を受け取ったとき $\psi_n(i) = \mathbf{v}_i$ s.t. $i = 1, \dots, M_n + 1$ とし $M_n + 1$ を受け取った場合は $\mathbf{v}_{M_n+1} \equiv \mathbf{0}$ とする. 符号器と復号器の性能を調べるため

$$\bar{P}_e^{(n)}(u) \equiv \Pr \{d_n(X^n, \psi(\varphi(X^n))) > u\} \quad (\text{A.27})$$

$$A_n(\mathbf{x}, u) \equiv P_{X^n}(\mathbf{x}) \mathbf{1}[d_n(\mathbf{x}, \mathbf{0}) > u] \quad (\text{A.28})$$

とおくと

$$\begin{aligned} \bar{P}_e^{(n)} &= \int P_{X^n}(\mathbf{x}) \Pr \{d_n(X^n, \psi(\varphi(X^n))) > u\} d\mathbf{x} \\ &= \int A_n(\mathbf{x}, u) \Pr \{d_n(X^n, \psi(\varphi(X^n))) > u\} d\mathbf{x} \\ &= \int A_n(\mathbf{x}, u) \prod_{i=1}^{M_n} \Pr \{d_n(\mathbf{x}, \mathbf{v}_i) > u\} d\mathbf{x} \\ &= \int A_n(\mathbf{x}, u) (\Pr \{d_n(\mathbf{x}, Y^n) > u\})^{M_n} d\mathbf{x} \\ &= \int A_n(\mathbf{x}, u) (1 - \Pr \{d_n(\mathbf{x}, Y^n) \leq u\})^{M_n} d\mathbf{x} \\ &= \int A_n(\mathbf{x}, u) \left(1 - \sum_{\mathbf{y} \in \mathcal{Y}^n} P_{Y^n}(\mathbf{y}) \mathbf{1}[d_n(\mathbf{x}, \mathbf{y}) \leq u]\right)^{M_n} d\mathbf{x} \end{aligned} \quad (\text{A.29})$$

である. ここで

$$J_u(\mathbf{x}, \mathbf{y}) = \mathbf{1} \left[(\mathbf{x}, \mathbf{y}) \in T_n^{(1)} \cap T_n^{(2)} \right] \quad (\text{A.30})$$

とすると

$$\begin{aligned} J_u(\mathbf{x}, \mathbf{y}) &\leq \mathbf{1} \left[(\mathbf{x}, \mathbf{y}) \in T_n^{(2)} \right] \\ &= \mathbf{1} [d_n(\mathbf{x}, \mathbf{y}) < u] \end{aligned} \quad (\text{A.31})$$

であるから

$$\bar{P}_e^{(n)} \leq \left(1 - \sum_{\mathbf{y} \in \mathcal{Y}^n} P_{Y^n}(\mathbf{y}) J_u(\mathbf{x}, \mathbf{y})\right)^{M_n} \quad (\text{A.32})$$

となり $J_u(\mathbf{x}, \mathbf{y}) = 1$ のとき $(\mathbf{x}, \mathbf{y}) \in T_n^{(1)}$ を満たすため

$$\begin{aligned}
& \frac{1}{n} \log \frac{P_{Y^n|X^n}(\mathbf{y}|\mathbf{x})}{P_{Y^n}(\mathbf{y})} - I(X; Y) < \gamma \\
& \therefore \log \frac{P_{Y^n|X^n}(\mathbf{y}|\mathbf{x})}{P_{Y^n}(\mathbf{y})} < n(I(X; Y) + \gamma) \\
& \quad \therefore \frac{P_{Y^n|X^n}(\mathbf{y}|\mathbf{x})}{P_{Y^n}(\mathbf{y})} < e^{n(I(X; Y) + \gamma)} \\
& \therefore P_{Y^n}(\mathbf{y}) > e^{-n(I(X; Y) + \gamma)} P_{Y^n|X^n}(\mathbf{y}|\mathbf{x})
\end{aligned} \tag{A.33}$$

であり

$$\bar{P}_e^{(n)} \leq \left(1 - \sum_{\mathbf{y} \in \mathcal{Y}^n} e^{-n(I(X; Y) + \gamma)} P_{Y^n|X^n}(\mathbf{y}|\mathbf{x}) J_u(\mathbf{x}, \mathbf{y}) \right)^{M_n} \tag{A.34}$$

が得られる. ここで

$$\begin{aligned}
0 & \leq \sum_{\mathbf{y} \in \mathcal{Y}^n} e^{-n(I(X; Y) + \gamma)} P_{Y^n|X^n}(\mathbf{y}|\mathbf{x}) J_u(\mathbf{x}, \mathbf{y}) \leq 1 \\
0 & \leq M_n
\end{aligned}$$

を満たすことから (付録 A-7-1) より不等式 $(1-x)^y \leq e^{-xy}$ ($0 \leq x \leq 1, y \geq 0$) を用いると

$$\begin{aligned}
\bar{P}_e^{(n)} & \leq \exp \left(-M_n \sum_{\mathbf{y} \in \mathcal{Y}^n} e^{-n(I(X; Y) + \gamma)} P_{Y^n|X^n}(\mathbf{y}|\mathbf{x}) J_u(\mathbf{x}, \mathbf{y}) \right) \\
& = \exp \left(-e^{-n(I(X; Y) + 2\gamma)} e^{-n(I(X; Y) + \gamma)} \sum_{\mathbf{y} \in \mathcal{Y}^n} P_{Y^n|X^n}(\mathbf{y}|\mathbf{x}) J_u(\mathbf{x}, \mathbf{y}) \right) \\
& = \exp \left(-e^{n\gamma} \sum_{\mathbf{y} \in \mathcal{Y}^n} P_{Y^n|X^n}(\mathbf{y}|\mathbf{x}) J_u(\mathbf{x}, \mathbf{y}) \right)
\end{aligned} \tag{A.35}$$

さらに

$$\begin{aligned}
e^{n\gamma} & \geq 0 \\
0 & \leq \sum_{\mathbf{y} \in \mathcal{Y}^n} P_{Y^n|X^n}(\mathbf{y}|\mathbf{x}) J_u(\mathbf{x}, \mathbf{y}) \leq 1
\end{aligned}$$

を満たすことから (付録 A-7-2) より不等式 $e^{-xy} \leq 1 - x + e^{-y}$ ($0 \leq x \leq 1, y \geq 0$) を用いると

$$\bar{P}_e^{(n)} \leq 1 - \sum_{\mathbf{y} \in \mathcal{Y}^n} P_{Y^n|X^n}(\mathbf{y}|\mathbf{x}) J_u(\mathbf{x}, \mathbf{y}) + e^{-e^{n\gamma}} \tag{A.36}$$

ゆえに

$$\begin{aligned}
\bar{P}_e^{(n)} &\leq \int A_n(\mathbf{x}, u) \left(1 - \sum_{\mathbf{y} \in \mathcal{Y}^n} P_{Y^n|X^n}(\mathbf{y}|\mathbf{x}) J_u(\mathbf{x}, \mathbf{y}) + e^{-e^{n\gamma}} \right) d\mathbf{x} \\
&= \int A_n(\mathbf{x}, u) d\mathbf{x} - \int A_n(\mathbf{x}, u) \sum_{\mathbf{y} \in \mathcal{Y}^n} P_{Y^n|X^n}(\mathbf{y}|\mathbf{x}) J_u(\mathbf{x}, \mathbf{y}) d\mathbf{x} \\
&\quad + \int A_n(\mathbf{x}, u) e^{-e^{n\gamma}} d\mathbf{x}
\end{aligned} \tag{A.37}$$

が得られる。また

$$\begin{aligned}
\int A_n(\mathbf{x}, u) d\mathbf{x} &= \int P_{X^n}(\mathbf{x}) \mathbf{1}[d_n(\mathbf{x}, \mathbf{0}) > u] d\mathbf{x} \\
&= \Pr\{d_n(\mathbf{x}, \mathbf{0}) > u\}
\end{aligned} \tag{A.38}$$

$$\begin{aligned}
&\int A_n(\mathbf{x}, u) \sum_{\mathbf{y} \in \mathcal{Y}^n} P_{Y^n|X^n}(\mathbf{y}|\mathbf{x}) J_u(\mathbf{x}, \mathbf{y}) d\mathbf{x} \\
&= \int \sum_{\mathbf{y} \in \mathcal{Y}^n} \mathbf{1}[d_n(\mathbf{x}, \mathbf{0}) > u] \mathbf{1}[(\mathbf{x}, \mathbf{y}) \in T_n^{(1)} \cap T_n^{(2)}] d\mathbf{x} \\
&= \Pr\left[\{d_n(X^n, \mathbf{0}) > u\} \cap \{X^n Y^n \in T_n^{(1)} \cap T_n^{(2)}\}\right]
\end{aligned} \tag{A.39}$$

であるので式 (A.38), (A.39) より

$$\begin{aligned}
&\int A_n(\mathbf{x}, u) d\mathbf{x} - \int A_n(\mathbf{x}, u) \sum_{\mathbf{y} \in \mathcal{Y}^n} P_{Y^n|X^n}(\mathbf{y}|\mathbf{x}) J_u(\mathbf{x}, \mathbf{y}) d\mathbf{x} \\
&= \Pr\{d_n(\mathbf{x}, \mathbf{0}) > u\} - \Pr\left[\{d_n(X^n, \mathbf{0}) > u\} \cap \{X^n Y^n \in T_n^{(1)} \cap T_n^{(2)}\}\right] \\
&= \Pr\left[\{d_n(X^n, \mathbf{0}) > u\} \cap \{X^n Y^n \notin T_n^{(1)} \cap T_n^{(2)}\}\right] \\
&= \Pr\left[\{d_n(X^n, \mathbf{0}) > u\} \cap \{X^n Y^n \notin T_n^{(1)} \cup X^n Y^n \notin T_n^{(2)}\}\right] \\
&\quad (\text{ド・モルガン則より}) \\
&\leq \Pr\left[\{d_n(X^n, \mathbf{0}) > u\} \cap \{X^n Y^n \notin T_n^{(1)}\}\right] \\
&\quad + \Pr\left[\{d_n(X^n, \mathbf{0}) > u\} \cap \{X^n Y^n \notin T_n^{(2)}\}\right] \\
&= \Pr\left[\{d_n(X^n, \mathbf{0}) > u\} \cap \{X^n Y^n \notin T_n^{(1)}\}\right] \\
&\quad + \Pr\left[\{d_n(X^n, \mathbf{0}) > u\} \cap \{d_n(X^n, Y^n) \geq u\}\right] \\
&\leq \Pr\left[\{d_n(X^n, \mathbf{0}) > u\} \cap \{X^n Y^n \notin T_n^{(1)}\}\right] \\
&\quad + \Pr[d_n(X^n, Y^n) \geq u]
\end{aligned} \tag{A.40}$$

となる. さらに式 (A.38) を式 (A.37) 中 3 項目に適用すると

$$\int A_n(\mathbf{x}, u) e^{-\epsilon^{n\gamma}} d\mathbf{x} = \Pr \{d_n(\mathbf{x}, \mathbf{0}) > u\} e^{-\epsilon^{n\gamma}} \quad (\text{A.41})$$

であるから式 (A.40), (A.41) を式 (A.37) に代入すると

$$\begin{aligned} \bar{P}_e^{(n)} \leq & \Pr \left[\{d_n(X^n, \mathbf{0}) > u\} \cap \{X^n Y^n \notin T_n^{(1)}\} \right] \\ & + \Pr [d_n(X^n, Y^n) \geq u] + \Pr [d_n(\mathbf{x}, \mathbf{0}) > u] e^{-\epsilon^{n\gamma}} \end{aligned} \quad (\text{A.42})$$

が得られる.

また $\bar{P}_e^{(n)} \equiv \Pr \{d_n(X^n, \psi(\varphi(X^n))) > u\}$ であるから部分積分を用いて

$$\mathbb{E}_{\varphi_n \phi_n} \mathbb{E}_{X^n} [d_n(X^n, \psi(\varphi(X^n)))] = \int_0^\infty \bar{P}_e^{(n)}(u) du - \left[u \bar{P}_e^{(n)}(u) \right]_0^\infty \quad (\text{A.43})$$

となる. 式 (A.43) の右辺第 1 項目に着目すると

$$\begin{aligned} \mathbb{E}[d_n(X^n, Y^n)] &= \int_0^\infty t \Pr \{d_n(X^n, Y^n) = t\} dt \\ &= \int_0^u t \Pr \{d_n(X^n, Y^n) = t\} dt \\ &\quad + \int_u^\infty t \Pr \{d_n(X^n, Y^n) = t\} dt \end{aligned} \quad (\text{A.44})$$

と分けられる. まず式 (A.44) の右辺第 2 項について着目する. 歪み測度の有界性 $\mathbb{E}[d(X, Y)] < \infty$ より

$$\begin{aligned} \mathbb{E}[d_n(X^n, Y^n)] &= \mathbb{E} \left[\frac{1}{n} \sum_i d(X_i, Y_i) \right] \\ &= \frac{1}{n} \sum_i \mathbb{E}[d(X_i, Y_i)] \\ &= \mathbb{E}[d(X, Y)] < \infty \end{aligned} \quad (\text{A.45})$$

であり u についての極限を考えると

$$\lim_{u \rightarrow \infty} \int_u^\infty t \Pr \{d_n(X^n, Y^n) = t\} dt = \lim_{u \rightarrow \infty} \mathbb{E}[d_n(X^n, Y^n) \mathbf{1}[d_n(X^n, Y^n) > u]] = 0 \quad (\text{A.46})$$

であるから

$$\begin{aligned} \mathbb{E}[d_n(X^n, Y^n) \mathbf{1}[d_n(X^n, Y^n) > u]] &> \mathbb{E}[u \cdot \mathbf{1}[d_n(X^n, Y^n) > u]] \\ &= u \Pr \{d_n(X^n, Y^n) > u\} \end{aligned} \quad (\text{A.47})$$

となる. 式 (A.46), (A.47) より

$$\lim_{u \rightarrow \infty} u \Pr \{d_n(X^n, Y^n) > u\} = 0 \quad (\text{A.48})$$

である. 次に式 (A.44) 中右辺第 1 項目については

$$\lim_{u \rightarrow \infty} \int_0^u t \Pr\{d_n(X^n, Y^n) = t\} dt = \int_0^\infty t \Pr\{d_n(X^n, Y^n) = t\} dt = \mathbb{E}[d_n(X^n, Y^n)] \quad (\text{A.49})$$

となる.

ここで式 (A.42) の両辺を u 倍すると

$$u\bar{P}_e^{(n)}(u) \leq u \Pr\{d_n(X^n, Y^n) > u\} + u \Pr\{d_n(X^n, \mathbf{0}) > u\} + e^{-e^{n\gamma}} u \Pr\{d_n(X^n, \mathbf{0}) > u\} \quad (\text{A.50})$$

両辺に対し u についての極限をとると式 (A.48) より

$$\lim_{u \rightarrow \infty} u\bar{P}_e^{(n)}(u) = 0 \quad (\text{A.51})$$

であるから式 (A.51) を式 (A.43) に代入すると

$$\mathbb{E}_{\varphi_n \phi_n} \mathbb{E}_{X^n} [d_n(X^n, \psi(\varphi(X^n)))] = \int_0^\infty \bar{P}_e^{(n)}(u) du \quad (\text{A.52})$$

であることがわかる. また, 同様に部分積分より

$$\begin{aligned} \int_0^\infty \Pr\{d_n(X^n, Y^n) > u\} du &= \mathbb{E}[d_n(X^n, Y^n)] \\ &\quad + [u \Pr\{d_n(X^n, Y^n) > u\}]_0^\infty \\ &= \mathbb{E}[d_n(X^n, Y^n)] \end{aligned} \quad (\text{A.53})$$

$$\int_0^\infty \Pr\{d_n(X^n, \mathbf{0}) > u\} du = \mathbb{E}[d_n(X^n, \mathbf{0})] \quad (\text{A.54})$$

ここで $B_n \equiv \int_0^\infty \Pr\{X^n Y^n \notin T_n \cap d_n(X^n, \mathbf{0}) > u\} du$ とすると

$$\begin{aligned} B_n &= \int_0^\infty \int_{(\mathbf{x}, \mathbf{y}) \notin T_n} P_{X^n Y^n}(\mathbf{x}, \mathbf{y}) \mathbf{1}[d_n(\mathbf{x}, \mathbf{0}) > u] d(\mathbf{x}, \mathbf{y}) du \\ &= \int_{(\mathbf{x}, \mathbf{y}) \notin T_n} P_{X^n Y^n}(\mathbf{x}, \mathbf{y}) \left\{ \int_0^\infty \mathbf{1}[d_n(\mathbf{x}, \mathbf{0}) > u] du \right\} d(\mathbf{x}, \mathbf{y}) \\ &= \int_{(\mathbf{x}, \mathbf{y}) \notin T_n} P_{X^n Y^n}(\mathbf{x}, \mathbf{y}) d_n(\mathbf{x}, \mathbf{0}) d(\mathbf{x}, \mathbf{y}) \end{aligned} \quad (\text{A.55})$$

歪み測度 d_n は正規分布の情報源に対して一様可積分性を満たすことから

$$\lim_{n \rightarrow \infty} B_n = 0 \quad (\text{A.56})$$

となる. ここで式 (A.42) と式 (A.52) より

$$\begin{aligned} \mathbb{E}_{\varphi_n \phi_n} \mathbb{E}_{X^n} [d_n(X^n, \psi(\varphi(X^n)))] &\leq \int_0^\infty \left(\Pr \left[\{d_n(X^n, \mathbf{0}) > u\} \cap \{X^n Y^n \notin T_n^{(1)}\} \right] \right. \\ &\quad \left. + \Pr [d_n(X^n, Y^n) \geq u] + \Pr [d_n(\mathbf{x}, \mathbf{0}) > u] e^{-e^{n\gamma}} \right) du \end{aligned} \quad (\text{A.57})$$

であるから, 式 (A.57) の右辺に式 (A.53),(A.54) と B_n をそれぞれ代入すると

$$\mathbb{E}_{\varphi_n \phi_n} \mathbb{E}_{X^n} [d_n(X^n, \psi(\varphi(X^n)))] \leq B_n + \mathbb{E}[d_n(X^n, Y^n)] + e^{-e^{n\gamma}} (\mathbb{E}[d_n(X^n, \mathbf{0})]) \quad (\text{A.58})$$

となる. ここで, 歪み測度の一様可積分性から

$$\sup_{n \geq 1} \mathbb{E}[d_n(X^n, \mathbf{0})] < \infty \quad (\text{A.59})$$

である. $n \rightarrow \infty$ の時 $e^{-e^{n\gamma}} = 0$ であることと式 (A.56),(A.59) より式 (A.58) の両辺の上極限を考えると

$$\begin{aligned} \limsup_{n \rightarrow \infty} \mathbb{E} [d_n(X^n, \psi(\varphi(X^n)))] &\leq \limsup_{n \rightarrow \infty} \mathbb{E} [d_n(X^n, Y^n)] \\ &= \limsup_{n \rightarrow \infty} \mathbb{E} [d(X, Y)] \\ &= \mathbb{E} [d(X, Y)] \\ &\leq D \end{aligned} \quad (\text{A.60})$$

が得られる.

一方で $M_n \equiv e^{n(I(X;Y)+2\gamma)}$ であるから

$$\begin{aligned} \limsup_{n \rightarrow \infty} \frac{1}{n} \log \|\varphi_n\| &\leq \limsup_{n \rightarrow \infty} \frac{1}{n} \log(M_n + 1) \\ &= \limsup_{n \rightarrow \infty} \left(\frac{1}{n} \log M_n + \frac{1}{n} \log \frac{M_n + 1}{M_n} \right) \\ &= \limsup_{n \rightarrow \infty} (I(X;Y) + 2\gamma) \\ &= I(X;Y) + 2\gamma \\ &= R \end{aligned} \quad (\text{A.61})$$

である. また任意の入力 \mathbf{X} に対して

$$R(D) \leq I(X;Y) + 2\gamma \quad (\text{A.62})$$

であり, Y は $\mathbb{E}[d_n(X^n, Y^n)] \leq D$ を満たす任意の値であることから

$$R(D) \leq \inf_{Y: \mathbb{E}[d_n(X, Y)] \leq D} I(X;Y) + 2\gamma \quad (\text{A.63})$$

である. $\gamma > 0$ が任意であることから

$$R(D) \leq \inf_{Y: \mathbb{E}[d(X, Y)] \leq D} I(X;Y) \quad (\text{A.64})$$

であり達成可能を満たすことがわかる. \square

A-3 チェイン則

定理 5 (チェイン則) $X^n = X_1, X_2, \dots, X_n$ が $P_{X^n}(x_1, x_2, \dots, x_n)$, $Z^n = Z_1, Z_2, \dots, Z_n$ が $P_{Z^n}(z_1, z_2, \dots, z_n)$ にそれぞれ従うとしたとき

チェイン則

$$h(X^n|Z^n) = \sum_{i=1}^n h(X_i|Z^n, X^{i-1}) \quad (\text{A.65})$$

が成り立つ.

(証明)

$$\begin{aligned} h(X^n|Z^n) &= - \int_{x_1} \cdots \int_{x_n} \sum_{Z^n} P_{X^n Z^n}(X^n, Z^n) \log P_{X^n|Z^n}(X^n|Z^n) dX^n \\ &= - \int_{x_1} \cdots \int_{x_n} \sum_{Z^n} P_{X^n Z^n}(X^n, Z^n) \log \prod_{i=1}^n P_{X_i|Z^n X^{i-1}}(X_i|Z^n, X^{i-1}) dX^n \\ &= - \int_{x_1} \cdots \int_{x_n} \sum_{Z^n} P_{X^n Z^n}(X^n, Z^n) \sum_{i=1}^n \log P_{X_i|Z^n X^{i-1}}(X_i|Z^n, X^{i-1}) dX^n \\ &= - \sum_{i=1}^n \int_{x_1} \cdots \int_{x_i} \sum_{Z^n} \left\{ \int_{X_{i+1}} \cdots \int_{X_n} P_{X^n Z^n}(X^n, Z^n) dX_n \cdots dX_{i+1} \right\} \\ &\quad \times \log P_{X_i|Z^n X^{i-1}}(X_i|Z^n, X^{i-1}) dX^i \\ &= - \sum_{i=1}^n \int_{x_1} \cdots \int_{x_i} \sum_{Z^n} P_{X^n Z^n}(X^i, Z^n) \log P_{X_i|Z^n X^{i-1}}(X_i|Z^n, X^{i-1}) dX^i \\ &= \sum_{i=1}^n h(X_i|Z^n, X^{i-1}) \end{aligned} \quad (\text{A.66})$$

となる. □

次に相互情報量についてのチェイン則を導く

定理 6 (相互情報量に関するチェイン則) $I(X; Y, Z) = I(X; Z) + I(X; Y|Z)$ が成り立つ.

(証明)

$$\begin{aligned}
I(X; Y, Z) &= \int_x \sum_y \sum_z P_{XYZ}(x, y, z) \log \frac{P_{XYZ}(x, y, z)}{P_X(x)P_{YZ}(y, z)} dx \\
&= \int_x \sum_y \sum_z P_{XYZ}(x, y, z) \frac{1}{P_{YZ}(y, z)} dx \\
&\quad + \int_x \sum_y \sum_z P_{XYZ}(x, y, z) \log P_{YZ|X}(y, z|x) dx \\
&= \sum_y \sum_z P_{YZ}(y, z) \frac{1}{P_{YZ}(y, z)} \\
&\quad + \int_x P_X(x) \left(\sum_y \sum_z P_{YZ|X}(y, z|x) \log P_{YZ|X}(y, z|x) \right) dx \\
&= H(YZ) - \int_x P_X(x) H(YZ|X = x) dx \\
&= H(YZ) - H(YZ|X) \\
&= H(Z) + H(Y|Z) - H(Z|X) - H(Y|XZ) \\
&= H(Z) - H(Z|X) + H(Y|Z) - H(Y|XZ) \\
&= I(X; Z) + I(X; Y|Z)
\end{aligned} \tag{A.67}$$

□

同様に $I(X; Y, Z) = I(X; Y) + I(X; Z|Y)$ についても示すことができる。

A-4 データ処理不等式

定理 7 (データ処理不等式) 確率変数 X, Y, Z がマルコフ連鎖をなすとき $I(X; Y) \geq I(X; Z)$ となる。

(証明) 確率変数 X, Y, Z について Z の条件付き分布が Y に依存し、 X に対して条件付き独立であるとする。すると X, Y, Z は同時確率分布 $P_{XYZ}(x, y, z) = P_X(x)P_{Y|X}(y|x)P_{Z|Y}(z|y)$ を満たす。このとき相互情報量におけるチェイン則 (A.67) より

$$I(X; Y, Z) = I(X; Z) + I(X; Y|Z) = I(X; Y) + I(X; Z|Y) \tag{A.68}$$

を満たす。

まずマルコフ性より $P_{XZ|Y}(x, z|y) = P_{X|Y}(x|y)P_{Z|Y}(z|y)$ であるから $I(X; Z|Y) = 0$ である。

次に $I(X; Y|Z) \geq 0$ を示す. 相互情報量の定義式より

$$\begin{aligned}
I(X; Y|Z) &= \mathbb{E}_{XYZ} \left[\log \frac{P_{XY|Z}(XY|Z)}{P_{X|Z}(X|Z)P_{Y|Z}(Y|Z)} \right] \\
&= -\mathbb{E}_{XYZ} \left[\log \frac{P_{X|Z}(X|Z)P_{Y|Z}(Y|Z)}{P_{XY|Z}(XY|Z)} \right] \\
&\geq -\log \left(\mathbb{E}_{XYZ} \left[\frac{P_{X|Z}(X|Z)P_{Y|Z}(Y|Z)}{P_{XY|Z}(XY|Z)} \right] \right) \quad (\text{イエンセンの不等式より}) \\
&= -\log \int_x \sum_y \sum_z P_{XYZ}(x, y, z) \frac{P_{X|Z}(x|z)P_{Y|Z}(y|z)P_Z(z)}{P_{XY|Z}(xy|z)P_Z(z)} dx \\
&= -\log \int_x \sum_y \sum_z P_{X|Z}(x|z)P_{Y|Z}(y|z)P_Z(z) dx \\
&= -\log \sum_z \left(\int_x P_{X|Z}(x|z) dx \right) \left(\sum_y P_{Y|Z}(y|z) \right) P_Z(z) \\
&= -\log 1 \\
&= 0
\end{aligned} \tag{A.69}$$

より $I(X; Y) \geq I(X; Z)$ が示される. □

A-5 相互情報量の凸性

定理 8 (相互情報量の凸性) 相互情報量 $I(X; Y)$ は (X, Y) の組に対し下に凸な関数である.

(証明) まず, $f(x) = x \log x$ とおき $a_i > 0, b_i > 0$ を仮定する. さらに $\lambda_i = \frac{b_i}{\sum_{j=1}^l b_j}$ とおくとイエンゼンの不等式より

$$\begin{aligned}
\sum_{i=1}^l \lambda_i f\left(\frac{a_i}{b_i}\right) &\geq f\left(\sum_{i=1}^l \lambda_i \frac{a_i}{b_i}\right) \\
&= f\left(\sum_{i=1}^l \frac{b_i}{\sum_{j=1}^l b_j} \frac{a_i}{b_i}\right) \\
&= f\left(\sum_{i=1}^l \frac{1}{\sum_{j=1}^l b_j} a_i\right) \\
&= f\left(\frac{\sum_{i=1}^l a_i}{\sum_{j=1}^l b_j}\right) \\
&= \frac{\sum_{i=1}^l a_i}{\sum_{j=1}^l b_j} \log \frac{\sum_{i=1}^l a_i}{\sum_{j=1}^l b_j}
\end{aligned} \tag{A.70}$$

であり, $l = 2$ とすると

$$\frac{b_1}{b_1 + b_2} \frac{a_1}{b_1} \log \frac{a_1}{b_1} + \frac{b_2}{b_1 + b_2} \frac{a_2}{b_2} \log \frac{a_2}{b_2} \geq \frac{a_1 + a_2}{b_1 + b_2} \log \frac{a_1 + a_2}{b_1 + b_2} \quad (\text{A.71})$$

となることから

$$a_1 \log \frac{a_1}{b_1} + a_2 \log \frac{a_2}{b_2} \geq (a_1 + a_2) \log \frac{a_1 + a_2}{b_1 + b_2} \quad (\text{A.72})$$

が得られる. ここで $0 \leq \alpha \leq 1$ としたとき

$$a_1 = \alpha p_1(x, y) \quad a_2 = (1 - \alpha) p_2(x, y) \quad b_1 = \alpha q_1(x, y) \quad b_2 = (1 - \alpha) q_2(x, y) \quad (\text{A.73})$$

とする. 先の式に代入し両辺を x について積分し y について総和をとると

$$\begin{aligned} & \int_x \sum_y \left\{ \alpha p_1(x, y) \log \frac{p_1(x, y)}{q_1(x, y)} + (1 - \alpha) p_2(x, y) \log \frac{p_2(x, y)}{q_2(x, y)} \right\} dx \\ & \geq \int_x \sum_y \left\{ \{ \alpha p_1(x, y) + (1 - \alpha) p_2(x, y) \} \log \frac{\alpha p_1(x, y) + (1 - \alpha) p_2(x, y)}{\alpha q_1(x, y) + (1 - \alpha) q_2(x, y)} \right\} dx \end{aligned} \quad (\text{A.74})$$

となり, 新たに定義した相互情報量の式に当てはめると

$$\alpha I_1(X; Y) + (1 - \alpha) I_2(X; Y) \geq I_\alpha(X; Y) \quad (\text{A.75})$$

が得られる. □

A-6 分散の有界性

定理 9 (分散の有界性) 連続である入力アルファベットを \mathcal{X} , 離散で有限な出力アルファベット \mathcal{Y} を離散値としたときの $Z = \log \frac{P_{Y|X}(y|x)}{P_Y(y)}$ に対する分散 $V(Z)$ は有界である

(証明) 期待値を用いて分散を表すと

$$V(Z) = \mathbb{E}[Z^2] - (\mathbb{E}[Z])^2 \quad (\text{A.76})$$

である. まず, $\mathbb{E}(Z)$ について考える.

$$\begin{aligned} \mathbb{E}(Z) &= \int_x \sum_y P_{XY}(x, y) \left(\log \frac{P_{Y|X}(y|x)}{P_Y(y)} \right) dx \\ &= \int_x \sum_y P_{XY}(x, y) \left(\log \frac{P_{XY}(x, y)}{P_X(x) P_Y(y)} \right) dx \\ &= I(X, Y) \\ &= H(Y) - H(Y|X) \\ &\leq H(Y) \\ &\leq \log |\mathcal{Y}| \end{aligned} \quad (\text{A.77})$$

であり相互情報量の非負性と出力アルファベットが有限であることから $\mathbb{E}(Z)$ は有界であることがわかる. 次に $\mathbb{E}[Z^2]$ について考える.

$$\begin{aligned}
\mathbb{E}[Z^2] &= \int_x \sum_y P_{XY}(x, y) \left(\log \frac{P_{Y|X}(y|x)}{P_Y(y)} \right)^2 dx \\
&= \int_{x: P_{Y|X}(y|x) < P_Y(y)} \sum_y P_{XY}(x, y) \left(\log \frac{P_{Y|X}(y|x)}{P_Y(y)} \right)^2 dx \\
&\quad + \int_{x: P_{Y|X}(y|x) \geq P_Y(y)} \sum_y P_{XY}(x, y) \left(\log \frac{P_{Y|X}(y|x)}{P_Y(y)} \right)^2 dx \\
&= \int_{x: P_{Y|X}(y|x) < P_Y(y)} \sum_y P_{XY}(x, y) \left(\log \frac{P_{Y|X}(y|x)}{P_Y(y)} \right)^2 dx \\
&\quad + \int_{x: P_{Y|X}(y|x) \geq P_Y(y)} \sum_y P_{XY}(x, y) \left(\log \frac{P_{Y|X}(y|x)}{P_Y(y)} \right)^2 dx \\
&\leq \int_x \sum_y P_{XY}(x, y) \left(\log \frac{1}{P_{Y|X}(y|x)} \right)^2 dx + \int_x \sum_y P_{XY}(x, y) \left(\log \frac{1}{P_Y(y)} \right)^2 dx \\
&= \int_x P_X(x) \sum_y P_{Y|X}(y|x) \left(\log \frac{1}{P_{Y|X}(y|x)} \right)^2 dx + \sum_y P_Y(y) \left(\log \frac{1}{P_Y(y)} \right)^2 dx
\end{aligned} \tag{A.78}$$

ここで $0 \leq u \leq 1$ に対する関数 $g(u) = u \log \frac{1}{u}$ を考える. さらに $t = \log \frac{1}{u}$ とおくと $g(u) = t^2 e^{-t} = h(t)$ とできる. $h'(t) = 2te^{-t} + (-t^2 e^{-t}) = te^t(2-t)$ であることから関数 $h(t)$ は t が正の領域において $t = 2$ の場合に最大値をとる.

$$\max_{0 \leq u \leq 1} g(u) = \max_{0 \leq t < \infty} h(t) = h(2) = \frac{4}{e^2} \tag{A.79}$$

これを式 (A.78) に当てはめると

$$\begin{aligned}
\mathbb{E}[Z^2] &\leq \int_x P_X(x) \sum_y \frac{4}{e^2} dx + \sum_y \frac{4}{e^2} \\
&= \frac{4}{e^2} |\mathcal{Y}| + \frac{4}{e^2} |\mathcal{Y}| \\
&= \frac{8}{e^2} |\mathcal{Y}|
\end{aligned} \tag{A.80}$$

より出力アルファベットが有限な場合は $\mathbb{E}[Z^2]$ が有界であることが示され, 分散 $V(Z)$ も有界となることがわかる. \square

A-7 不等式の証明

A-7-1 $(1-x)^y \leq e^{-xy}$ ($0 \leq x \leq 1, y \geq 0$) の導出

$f(x) = e^{-x} + x - 1$ とおくと $f'(x) = -e^{-x} + 1$ であり $x \geq 0$ に対して $f'(x) \geq 0$ となることがわかる. $f(0) = 0$ より $0 \leq x \leq 1$ において

$$f(x) = e^{-x} + x - 1 \geq 0 \quad \therefore e^{-x} \geq 1 - x \quad (\text{A.81})$$

$y \geq 0$ であるから両辺を y 乗すると

$$e^{-xy} \geq (1-x)^y \quad (0 \leq x \leq 1) \quad (\text{A.82})$$

が得られる.

A-7-2 $e^{-xy} \leq 1 - x + e^{-y}$ ($0 \leq x \leq 1, y \geq 0$) の導出

$g(x, y) = e^{-xy}$ とおくと $\frac{\partial^2}{\partial x^2} g(x, y) = y^2 e^{-xy}$ であり $0 \leq x \leq 1, y \geq 0$ において $\frac{\partial^2}{\partial x^2} g(x, y) \geq 0$ であることから $g(x, y)$ は x に対して下に凸な関数であることがわかる.

また $g(0, y) = 1, g(1, y) = e^{-y}$ であることから, この 2 点を通る直線の方程式は

$$g(x, y) \leq (e^{-y} - 1)x + 1 \quad (0 \leq x \leq 1) \quad (\text{A.83})$$

であり, $0 \leq x \leq 1$ において $xe^{-y} \leq e^{-y}$ であるから

$$e^{-xy} \leq 1 - x + e^{-y} \quad (0 \leq x \leq 1) \quad (\text{A.84})$$

が得られる.

付録 B 記号一覧

論文中で用いた主な記号の一覧を表にまとめる. ただし A は確率変数, a は変数, a_n は数列を表す.

表 1 2章で用いる記号一覧

記号	意味
X	情報源, 観測したい対象の情報 (2.5 では任意の確率変数)
N	ノイズ, CEO 問題においてエージェントが 情報源を観測する際に乗るノイズ
\tilde{X}	情報源に対しノイズが加えられた情報 エージェントが観測する情報
\hat{X}	CEO が受け取る復号された情報
\mathbb{R}	実数の集合
\mathcal{C}	コードブック, 復号器が受け取った符号から 出力を生成するための変換表
$\mathbb{E}[A], \mathbb{E}[a]$	それぞれ A, a の期待値
φ	符号器, $\varphi(\tilde{X})$ は \tilde{X} を符号化した情報を表す
ψ	復号器, $\psi(\varphi(\tilde{X}))$ は \tilde{X} を符号化し復号した情報を表す
$\ \varphi_n\ $	符号器の符号語数, $\log \ \varphi_n\ $ は符号語ビット数を表す
$\mathcal{N}(\mu, \sigma^2)$	平均値 μ , 分散 σ^2 の正規分布
L	CEO 問題におけるエージェントの人数
σ_X^2	$L = 1$ の CEO 問題に対する情報源 X の分散
σ_N^2	$L = 1$ の CEO 問題に対する外部ノイズ N の分散
$\limsup_{n \rightarrow \infty} a_n = \alpha$	上極限, $\forall \epsilon > 0, \exists N \in \mathbb{N} \text{ s.t. } n \geq N \Rightarrow a_n < \alpha + \epsilon$ かつ $\forall \epsilon > 0$ に対し $a_n > \alpha - \epsilon$ を満たす要素が無数に存在する
n	系列の長さ, 送受信する情報の数
X^n	入力の文字列
R	レート, 符号器復号器間で送る情報量の上限値
D	歪み, 入出力の値の違いを表す尺度
Y, Z	任意の離散確率変数
$I(X, Y)$	確率変数 X, Y に関する相互情報量

表 2 3章で用いる記号一覧

記号	意味
\inf	下限
Y	逆定理, 順定理それぞれにおいて歪みの達成可能性を確かめるために作った任意の確率変数
$H(A)$	離散確率変数 A に対するエントロピー
$h(A)$	連続確率変数 A に対する微分エントロピー

表 3 付録 A で用いる記号一覧

記号	意味
Q_n	$Q_n \in 1, \dots, n, P_{Q_n} = \frac{1}{n}$ を満たす確率変数
X_{Q_n}	入力文字列 X^n のうち Q_n 番目の要素
\mathbb{N}	自然数の集合
ϵ, δ, γ	それぞれ任意の正の数

表 4 付録 B で用いる記号一覧

記号	意味
Y	ランダム符号化によって生成した出力を表す確率変数
\mathbf{x}	入力の系列の値
\mathbf{y}	ランダム符号化によって得られる出力系列の値
$T_n^{(1)}$	式 (A.25) において相互情報量が収束する (\mathbf{x}, \mathbf{y}) の組み合わせ
$T_n^{(2)}(u)$	歪みが u 以下を満たす (\mathbf{x}, \mathbf{y}) の組み合わせ
M_n	ランダム符号化によって生成する復号語の数
$\mathbf{0}$	基準語, 符号器への入力 \mathbf{x} に対し どの復号語よりも歪みが小さくなる時復号語とする
$\bar{P}_e^{(n)}$	歪みが u 以上となる確率
$A_n(\mathbf{x}, u)$	基準語 $\mathbf{0}$ からの平均歪みが u 以上となる確率
$\mathbf{1}[(\text{条件式})]$	指示関数, 条件式を満たす場合 1, 満たさない場合 0 とする
$J_u(\mathbf{x}, \mathbf{y})$	$T_n^{(1)}$ と $T_n^{(2)}(u)$ をそれぞれ満たす場合 1, 満たさない場合 0 とする