[招待講演] 新たな情報量による情報理論の再定式化

植松友彦

東京工業大学大学院

電子情報通信学会 情報理論研究会
13 July 2015

背景

- Han-Verdú (1993): 非定常あるいは非エルゴードの一般情報源と一般通信路に対する情報理論の創始。
- Han (1997): 最も一般的な情報理論である「情報スペクトルの理論」を構築。
- Renner-Wolf (2004): 情報スペクトルの代わりに Smooth エントロピーを用いた一般情報源の固定長符号化や乱数生成問題の再定式化。
- Wang-Colbeck-Renner (2009): Smooth Rényi ダイバージェンスによる通信路符号化定理の再定式化。

背景

- Uyematsu (2009): 一般情報源に関する固定長符号化定理の smooth エントロピーを用いた統一的な証明手法の確立。
- Uyematsu (2010): 一般情報源に関する乱数生成問題の smooth エントロピーを用いた統一的な証明手法の確立。
- Uyematsu-Matsuta (2014): 一般情報源のレート・歪み関数の smooth Rényi ダイバージェンスを用いた再定式化。

本発表

情報理論の各種問題の smooth エントロピーや smooth Rényi ダイバージェンスによる再定式化と証明手法について解説する。

特に、これらの情報量の特徴として

- その極限がエントロピー・スペクトルあるいは相互情報量スペクトルの上限／下限と一致すること
- 本質的な操作的意味をもつこと
- 符号化逆定理が容易に導けること

について説明する。

Smooth エントロピー

定義 1 (Smooth エントロピー)

任意の $\delta \in (0,1)$ に対し、$\mathcal{Z}$ 上の確率分布 P の smooth 最大エントロピー $H_0^\delta(P)$ ならびに smooth 最小エントロピー $H_\infty^\delta(P)$ は、

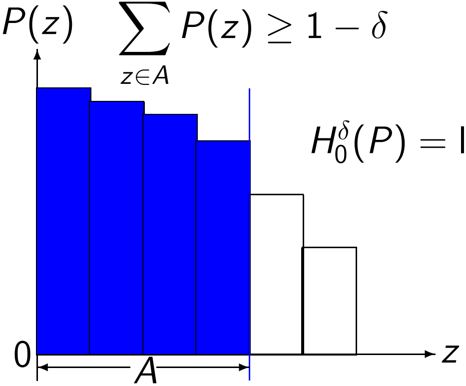
$$H_0^\delta(P) \triangleq \min_{\substack{A \subset \mathcal{Z} \\ P(A) \geq 1-\delta}} \log |A| \tag{1}$$

$$H_\infty^\delta(P) \triangleq \sup_{\substack{\beta \geq 1/|\mathcal{Z}| \\ \sum_{z \in \mathcal{Z}} |P(z)-\beta|^+ \leq \delta}} -\log \beta \tag{2}$$

によって定義される。但し、$|t|^+ \triangleq \max\{0, t\}$ である。

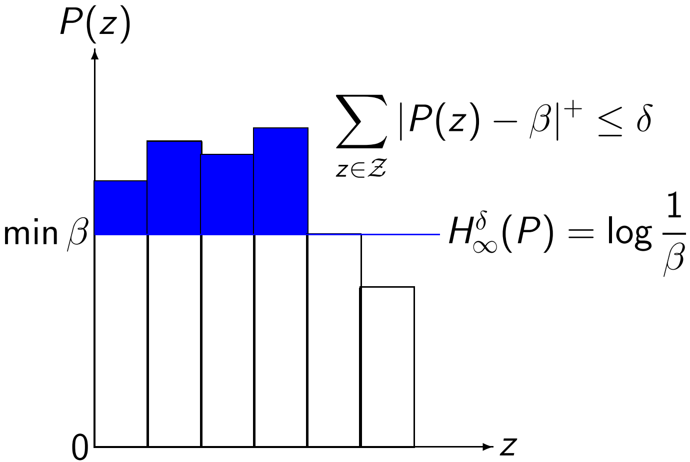
smooth 最大エントロピーの計算法

$$\sum_{z\in A} P(z) \geq 1-\delta$$



$$H_0^{\delta}(P) = \log_2 |A|$$

smooth 最小エントロピーの計算法



Smooth エントロピーの性質

- $H_0^\delta(P)$ は δ に対する単調減少関数である。
- $H_\infty^\delta(P)$ は δ に対する単調増加関数である。
- 無記憶情報源や有限マルコフ情報源などの多くの情報源について、有限の系列長に対する $H_0^\delta(P)$ や $H_\infty^\delta(P)$ は容易に計算可能である。

Smooth Rényi ダイバージェンス

定義 2 (Smooth Rényi ダイバージェンス)

任意の $\delta \in (0,1)$ に対し、$\mathcal{Z}$ 上の 2 つの確率分布 P と Q の smooth 最小 Rényi ダイバージェンス $D_0^\delta(P\|Q)$ と smooth 最大 Rényi ダイバージェンス $D_\infty^\delta(P\|Q)$ は、

$$D_0^\delta(P\|Q) \triangleq \sup_{\substack{\Psi:\mathcal{Z}\to[0,1]:\\ \sum_{z\in\mathcal{Z}}\Psi(z)P(z)\geq 1-\delta}} -\log \sum_{z\in\mathcal{Z}} \Psi(z)Q(z) \quad (3)$$

$$D_\infty^\delta(P\|Q) \triangleq \inf_{\substack{\Psi:\mathcal{Z}\to[0,1]:\\ \sum_{z\in\mathcal{Z}}\Psi(z)P(z)\geq 1-\delta}} \log \sup_{z\in\mathcal{Z}} \frac{\Psi(z)P(z)}{Q(z)} \quad (4)$$

によって定義される。

Smooth Rényi ダイバージェンスの性質

- $D_0^\delta(P\|Q)$ は δ に対する単調増加関数である。
- $D_\infty^\delta(P\|Q)$ は δ に対する単調減少関数である。
- smooth エントロピーと同様に、無記憶情報源や有限マルコフ情報源などの多くの情報源について、有限の系列長に対する $D_0^\delta(P\|Q)$ や $D_\infty^\delta(P\|Q)$ は容易に計算可能である。

一般情報源

一般情報源を

$$\mathbf{X} = \left\{ X^n = (X_1^{(n)}, X_2^{(n)}, \ldots, X_n^{(n)}) \right\}_{n=1}^{\infty}$$

によって定まる確率変数列 X^n として定義する。

- X^n: 情報源アルファベット $\mathcal{X}$ の直積 $\mathcal{X}^n$ 上の確率分布 P_{X^n} に従う確率変数
- $X_j^{(n)}$ $(j = 1, 2, \ldots, n)$: $\mathcal{X}$ 上に値を取る確率変数

以下、一般情報源 $\mathbf{X}$ の従う確率分布の列を $\{P_{X^n}\}$ によって表す。

一般情報源の例

情報源アルファベット: $\mathcal{X} = \{0, 1\}$

1. 情報源 1 (系列長 n に応じて分布が変わる)
   - 系列長 n が偶数ならば、$X^n = 11\ldots1$ に確定する。
   - 系列長 n が奇数ならば、X^n は一様分布に従う。

$$P_{X^n}(x^n) = 2^{-n} \quad \forall x^n \in \mathcal{X}^n.$$

2. 情報源 2 (最後の記号は必ず 1)

$$P_{X^n}(x^n) = \begin{cases} 2^{-n+1} & \text{if } x_n = 1 \\ 0 & \text{if } x_n = 0 \end{cases}$$

for $x^n = x_1x_2\ldots x_n$.

エントロピー・スペクトル上限と下限

定義3（エントロピー・スペクトル上限と下限）

一般情報源 $\mathbf{X}$ のエントロピー・スペクトル上限 $\overline{H}(\mathbf{X})$ と下限 $\underline{H}(\mathbf{X})$ は、

$$
\begin{aligned}
\overline{H}(\mathbf{X}) &\triangleq \inf\left\{\alpha : \lim_{n\to\infty} \Pr\left\{\frac{1}{n}\log\frac{1}{P_{X^n}(X^n)} > \alpha\right\} = 0\right\} \\
\underline{H}(\mathbf{X}) &\triangleq \sup\left\{\alpha : \lim_{n\to\infty} \Pr\left\{\frac{1}{n}\log\frac{1}{P_{X^n}(X^n)} < \alpha\right\} = 0\right\}
\end{aligned}
$$

によって定義される。

相互情報量スペクトル上限と下限

定義 3 (相互情報量スペクトル上限と下限)

相関を有する 2 つの一般情報源 $(\mathbf{X}, \mathbf{Y})$ に対する相互情報量スペクトル上限 $\overline{I}(\mathbf{X};\mathbf{Y})$ と下限 $\underline{I}(\mathbf{X};\mathbf{Y})$ は、

$$\overline{I}(\mathbf{X};\mathbf{Y}) \triangleq \inf\left\{\alpha : \lim_{n\to\infty} \Pr\left\{\frac{1}{n}\log\frac{P_{Y^n|X^n}(Y^n|X^n)}{P_{Y^n}(Y^n)} > \alpha\right\} = 0\right\}$$

$$\underline{I}(\mathbf{X};\mathbf{Y}) \triangleq \sup\left\{\alpha : \lim_{n\to\infty} \Pr\left\{\frac{1}{n}\log\frac{P_{Y^n|X^n}(Y^n|X^n)}{P_{Y^n}(Y^n)} < \alpha\right\} = 0\right\}$$

によって定義される。

smooth エントロピーの極限

定理 1 ($\overline{H}(\mathbf{X}), \underline{H}(\mathbf{X}), H_0^\delta(P_{X^n}), H_\infty^\delta(P_{X^n})$ の関係)

一般情報源 $\mathbf{X}$ について

$$\lim_{\delta\downarrow 0} \limsup_{n\to\infty} \frac{1}{n} H_0^\delta(P_{X^n}) = \overline{H}(\mathbf{X}) \tag{5}$$

$$\lim_{\delta\downarrow 0} \liminf_{n\to\infty} \frac{1}{n} H_\infty^\delta(P_{X^n}) = \underline{H}(\mathbf{X}) \tag{6}$$

式 (5) と式 (6) は、量子情報理論の枠組みにおいて Datta と Renner によって、一般情報源の枠組みにおいては著者によって示された。

smooth Rényi ダイバージェンスの極限

定理 2 ($\overline{I}(\mathbf{X};\mathbf{Y}), \underline{I}(\mathbf{X};\mathbf{Y})$ と $D_\infty^\delta(P_{X^nY^n}\|P_{X^n}\times P_{Y^n}), D_0^\delta(P_{X^nY^n}\|P_{X^n}\times P_{Y^n})$ の関係)

相関を有する 2 つの一般情報源 $(\mathbf{X},\mathbf{Y})$ について

$$\lim_{\delta\downarrow 0}\limsup_{n\to\infty}\frac{1}{n}D_\infty^\delta(P_{X^nY^n}\|P_{X^n}\times P_{Y^n}) = \overline{I}(\mathbf{X};\mathbf{Y}) \quad (7)$$

$$\lim_{\delta\downarrow 0}\liminf_{n\to\infty}\frac{1}{n}D_0^\delta(P_{X^nY^n}\|P_{X^n}\times P_{Y^n}) = \underline{I}(\mathbf{X};\mathbf{Y}) \quad (8)$$

式 (7) は Warsi によって、式 (8) は Wang らによって示された。

考察

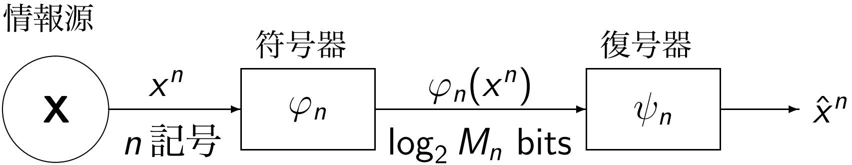
系列長 n が有限の場合において、smooth エントロピーや smooth Rényi ダイバージェンスを尺度として符号化レート等を評価できれば、系列長 n を無限に長くした場合に情報スペクトルの理論によって得られた結果に一致することが期待できる。

以下では、情報理論の各種問題について、smooth エントロピーや smooth Rényi ダイバージェンスを用いて、有限長の系列に対して符号化レートの限界を評価する。

情報源の固定長符号化

- 系列長: $n = 1, 2, \dots$
- 整数の集合: $\mathcal{M}_n = \{1, 2, \dots, M_n\}$
- 符号器: $\varphi_n : \mathcal{X}^n \to \mathcal{M}_n$
- 復号器: $\psi_n : \mathcal{M}_n \to \mathcal{X}^n$
- 一般情報源 **X** からの出力系列を表す確率変数: X^n
- 誤り確率: $\epsilon_n \triangleq \Pr\{X^n \neq \psi_n(\varphi_n(X^n))\}$
- 符号化レート: $r_n \triangleq \frac{1}{n} \log M_n$

情報源の固定長符号化



誤り確率 $\epsilon_n = \Pr\{\hat{X}^n \neq X^n\}$ を ϵ 以下にしたまま、
符号化レート $r = \frac{1}{n} \log M_n$ を最小化したい。

固定長符号化の例

アルファベット: $\{0, 1\}$
系列長: $n = 3$
符号語の集合: $\mathcal{M}_3 = \{1, 2, 3, 4\}$
符号器 $\varphi_3 : \{0, 1\}^3 \to \{0, 1, 2, 3\}$
復号器 $\psi_3 : \{0, 1, 2, 3\} \to \{0, 1\}^3$

x^3	$\varphi_3(x^3)$	$\psi_3(\varphi_3(x^3))$
000	1	000
100	2	100
010	3	010
001	4	001
otherwise	1	000

固定長符号化の例

x^3	$\varphi_3(x^3)$	$\psi_3(\varphi_3(x^3))$
000	1 (00)	000
100	2 (01)	100
010	3 (10)	010
001	4 (11)	001
otherwise	1 (00)	000

$$P(0) = 0.9, P(1) = 1 - P(0) = 0.1$$

$$\begin{aligned} \epsilon_3 &= 1 - 0.9^3 - 3 \times 0.9^2 \times 0.1 = 0.028 \\ r_3 &= \frac{1}{3} \log 4 \end{aligned}$$

ϵ-固定長符号化問題

定義 4 (ϵ-最小許容符号化レート)

一般情報源 $\mathbf{X}$ に対し、レート R が ϵ-達成可能であるとは、ある整数 n_0 が存在して、

$$\epsilon_n \le \epsilon \ (\forall n \ge n_0) \text{ かつ } \limsup_{n\to\infty} r_n \le R$$

を満足する (n, M_n, ϵ_n) 符号の列が存在することである。
更に、ϵ-最小許容固定長符号化レートを

$$R_f(\epsilon|\mathbf{X}) \triangleq \inf\{R : R \text{ が } \epsilon\text{-達成可能}\}$$

によって定義する。

固定長符号化定理と逆定理

定理 3

一般情報源 $\mathbf{X}$ と任意の (n, M_n, ϵ) 符号について

$$H_0^{\epsilon}(P_{X^n}) \le \log M_n \tag{9}$$

が成り立つ。また逆に、任意の $\epsilon \in (0,1)$ に対して、

$$\log M_n = H_0^{\epsilon}(P_{X^n}) \tag{10}$$

を満たす (n, M_n, ϵ) 符号が存在する。

式 (9) は Renner と Wolf によって、また式 (10) は著者によって示された。

定理3の証明

式 (9) の証明: 任意の (n, M_n, ϵ) 符号 (φ_n, ψ_n) において、正しく復号される $\mathcal{X}^n$ の系列の集合を

$$C_n = \{x^n \in \mathcal{X}^n : \psi_n(\varphi_n(x^n)) = x^n\}$$

とすれば、$P_{X^n}(C_n) \geq 1 - \epsilon$ が成り立つ。これと smooth 最大エントロピーの定義から直ちに

$$\log M_n \geq \log |C_n| \geq H_0^{\epsilon}(P_{X^n})$$

が導かれる。

定理3の証明

式 (10) の証明: smooth 最大エントロピーの定義の右辺の最小値を与える $\mathcal{X}^n$ の部分集合を A^* とする。すなわち、

$$H_0^{\epsilon}(P_{X^n}) = \min_{\substack{A\subset\mathcal{Z} \\ P(A)\geq 1-\epsilon}} \log|A| = \log|A^*|$$

とする。

A^* に含まれる記号と整数 $\{1, 2, \ldots, |A^*|\}$ を一対一対応させる符号器と復号器を考える。

この符号は、誤り確率 ϵ 以下で符号化レートが最小の $(n, |A^*|, \epsilon)$ 符号である。 Q.E.D.

ϵ-固定長符号化定理

定理 3 と ϵ-最小許容符号化レートの定義を組み合せると、次の定理を得る。

定理 4

一般情報源 $\mathbf{X}$ と任意の $\epsilon \in (0,1)$ について、

$$R_f(\epsilon|\mathbf{X}) = \limsup_{n\to\infty} \frac{1}{n} H_0^{\epsilon}(P_{X^n})$$

が成り立つ。

情報源を用いた一様乱数生成問題

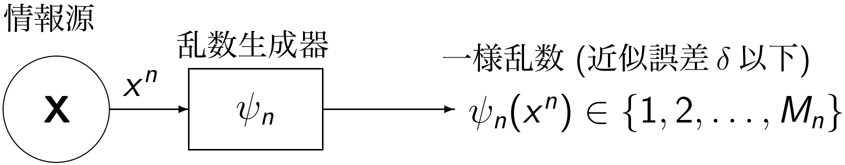
- 整数の集合: $\mathcal{U}_{M_n} \triangleq \{1, 2, \ldots, M_n\}$
- 乱数生成写像: $\psi_n : \mathcal{X}^n \to \mathcal{U}_{M_n}$
- 近似の評価尺度:

$$d_n \triangleq \sum_{i=1}^{M_n} \left| \sum_{\substack{x^n \in \mathcal{X}^n: \\ \psi_n(x^n)=i}} P_{X^n}(x^n) - \frac{1}{M_n} \right|^+$$

- 写像のレート:

$$r_n \triangleq \frac{1}{n} \log M_n$$

情報源を用いた一様乱数生成問題



出力される乱数の近似誤差を δ 以下にしたまま、
写像のレート $r_n = \frac{1}{n} \log M_n$ を最大化したい。

δ-intrinsic randomness

定義 5(δ-intrinsic randomness)

一般情報源 $\mathbf{X}$ に対し、レート R が δ-達成可能であるとは、ある整数 n_0 が存在して、

$$d_n \leq \delta \ (\forall n \geq n_0) \text{ かつ } \liminf_{n\to\infty} r_n \geq R$$

を満足する (n, M_n, d_n) 写像の列が存在することである。更に、δ-intrinsic randomness を

$$S_i(\delta|\mathbf{X}) \triangleq \sup\{R : R \text{ が } \delta\text{-達成可能}\}$$

によって定義する。

Intrinsic randomnessの基本定理

定理 5

一般情報源 $\mathbf{X}$ と任意の (n, M_n, δ) 写像について

$$\log M_n \le H_\infty^\delta(P_{X^n}) \tag{11}$$

が成り立つ。また逆に、任意の $\delta \in (0,1)$、任意の $\gamma > 0$ ならびに $e^{-n\gamma} \le \delta$ を満足する n について、M_n が

$$\log M_n \le H_\infty^\delta(P_{X^n}) - n\gamma \tag{12}$$

を満たすならば、(n, M_n, δ) 写像が存在する。

尚、式 (11) は Renner と Wolf によって示された。

定理5の証明

式(11)のみを示す。X^n に対して (n, M_n, δ) 写像 ψ_n が存在するとき、$\mathcal{X}^n$ の部分集合を

$$A_n \triangleq \{x^n \in \mathcal{X}^n : P_{X^n}(x^n) \geq 1/M_n\}$$

によって定める。写像 ψ_n による A_n の像を $\psi_n(A_n)$ とすれば、

$$\sum_{\substack{x^n \in \mathcal{X}^n: \\ \psi_n(x^n)=i}} P_{X^n}(x^n) \geq \frac{1}{M_n}, \quad \forall i \in \psi_n(A_n)$$

が成り立つ。

これから直ちに、

$$
\begin{aligned}
\delta &\geq \sum_{i=1}^{M_n} \left| \sum_{\substack{x^n \in \mathcal{X}^n: \\ \psi_n(x^n)=i}} P_{X^n}(x^n) - \frac{1}{M_n} \right|^+ \\
&\geq \sum_{i \in \psi_n(A_n)} \left(\sum_{\substack{x^n \in \mathcal{X}^n: \\ \psi_n(x^n)=i}} P_{X^n}(x^n) - \frac{1}{M_n} \right) \\
&\geq \sum_{x^n \in A_n} P_{X^n}(x^n) - \frac{|\psi_n(A_n)|}{M_n} \\
&\geq \sum_{x^n \in A_n} \left(P_{X^n}(x^n) - \frac{1}{M_n} \right)
\end{aligned}
$$

$$
\begin{aligned}
\delta \;&\geq\; \sum_{x^n \in A_n} \left(P_{X^n}(x^n) - \frac{1}{M_n} \right) \\
&=\; \sum_{\color{red}{x^n \in \mathcal{X}^n}} \left| P_{X^n}(x^n) - \frac{1}{M_n} \right|^{+}
\end{aligned}
$$

が得られ、これと smooth 最小エントロピーの定義から直ちに、

$$
\log M_n \leq H_\infty^\delta(P_{X^n})
$$

が導かれる。 Q.E.D.

δ-intrinsic randomness

定理 5 と δ-intrinsic の定義を組み合せると、δ-intrinsic randomness について次の定理を得る。

定理 6

一般情報源 $\mathbf{X}$ と任意の $\delta \in (0,1)$ について、

$$S_i(\delta|\mathbf{X}) = \liminf_{n\to\infty} \frac{1}{n} H_\infty^\delta(P_{X^n})$$

が成り立つ。

一般通信路

記号列の組 $(x^n, y^n) \in \mathcal{X}^n \times \mathcal{Y}^n$ に対して、

$$\sum_{y^n \in \mathcal{Y}^n} W^n(y^n|x^n) = 1, \quad \forall x^n \in \mathcal{X}^n$$

を満たす任意の条件付き確率分布 W^n の系列
$\mathbf{W} = \{W^n\}_{n=1}^{\infty}$ として、一般通信路を定義する。

一般の入力過程 $\mathbf{X} = \{X^n\}_{n=1}^{\infty}$ に対応する通信路 $\mathbf{W}$ の出力過程 $\mathbf{Y} = \{Y^n\}_{n=1}^{\infty}$ を

$$P_{X^nY^n}(x^n, y^n) = P_{X^n}(x^n)W^n(y^n|x^n)$$

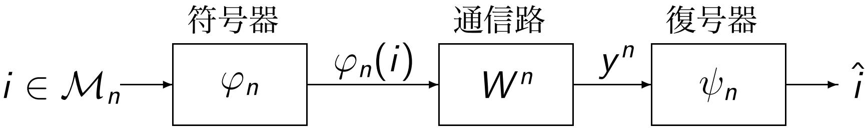
によって定める。

通信路符号化問題

- 系列長: $n = 1, 2, \ldots$
- 通報の集合: $\mathcal{M}_n = \{1, 2, \cdots, M_n\}$
- 符号器: $\varphi_n : \mathcal{M}_n \to \mathcal{X}^n$
- 復号器: $\psi_n : \mathcal{Y}^n \to \mathcal{M}_n$
- 符号化レート: $r_n = \frac{1}{n} \log M_n$
- (平均) 誤り率:

$$\epsilon_n = \frac{1}{M_n} \sum_{i=1}^{M_n} \underbrace{\sum_{\substack{y^n \in \mathcal{Y}^n: \\ \psi_n(y^n) \neq i}} W^n(y|\varphi_n(i))}_{\text{メッセージ } i \text{ を送信して正しく復号できない確率}}$$

通信路符号化



誤り確率 $\epsilon_n = \Pr\{\hat{I}^n \neq I^n\}$ を ϵ 以下にしたまま、
符号化レート $r = \frac{1}{n} \log M_n$ を最大化したい。

ϵ-通信路容量

定義 6（ϵ-通信路容量)

一般通信路 $\mathbf{W}$ に対し、レート R が ϵ-達成可能であるとは、

$$\limsup_{n\to\infty} \epsilon_n \le \epsilon \text{ かつ } \liminf_{n\to\infty} r_n \ge R$$

を満足する (n, M_n, ϵ_n) 符号の列が存在することである。更に、ϵ-達成可能なレートの上限として、ϵ-通信路容量を

$$C(\epsilon|\mathbf{W}) \triangleq \sup\{R : R \text{ が } \epsilon\text{-達成可能}\}$$

によって定義する。

通信路符号化定理と逆定理

定理 7 (通信路符号化定理と逆定理)

通信路 W^n と任意の (n, M_n, ϵ) 符号について

$$\begin{aligned}\log M_n &\leq \sup_{P_{X^n}} \sup_{\Psi \in \mathcal{S}(\epsilon)} - \log \sum_{(x^n, y^n) \in \mathcal{X}^n \times \mathcal{Y}^n} \Psi(x^n, y^n) \\ &\quad \times P_{X^n}(x^n) P_{Y^n}(y^n) \qquad (13) \\ &\leq \sup_{P_{X^n}} D_0^{\epsilon}(P_{X^n Y^n} \| P_{X^n} \times P_{Y^n}) \qquad (14)\end{aligned}$$

が成り立つ。但し、$\mathcal{S}(\epsilon)$ は次の 2 条件を満足する関数 $\Psi : \mathcal{X}^n \times \mathcal{Y}^n \to \{0, 1\}$ の集合である。

定理 7 (続き)

$$\sum_{(x^n, y^n) \in \mathcal{X}^n \times \mathcal{Y}^n} \Psi(x^n, y^n) W(y^n | x^n) P_{X^n}(x^n) \geq 1 - \epsilon$$

$$\sum_{x^n \in \mathcal{X}^n : P_{X^n}(x^n) > 0} \Psi(x^n, y^n) = 1, \quad \forall y^n \in \mathcal{Y}^n$$

また、任意の $\epsilon_1 \in [0, 1)$ と $\epsilon_2 \in (0, 1 - \epsilon_1]$ に対して、

$$\log M_n \geq \sup_{P_{X^n}} D_0^{\epsilon_1}(P_{X^n Y^n} \| P_{X^n} \times P_{Y^n}) + \log \epsilon_2 \tag{15}$$

を満たす $(n, M_n, \epsilon_1 + \epsilon_2)$ 符号が存在する。

通信路符号化定理と逆定理

式 (13) は著者によって示された上限であり、式 (14) による上限は Wang らによって示された。尚、著者らの証明は情報処理不等式を使っていないという点で、Wang らの証明よりも単純である。

式 (15) は確率的復号器に対して Wang らによって示されたが、この不等式自体は復号器を確定的復号器に限定しても成立することが証明できる。

定理 7 の証明

式 (13) の成立のみを示す。

与えられた (n, M_n, ϵ) 符号 (φ_n, ψ_n) に対し、関数 $\Psi_c : \mathcal{X}^n \times \mathcal{Y}^n \to \{0, 1\}$ を

$$\Psi_c(x^n, y^n) = \begin{cases} 1 & \text{if } \psi_n(y^n) = x^n \\ 0 & \text{otherwise} \end{cases}$$

によって定める。

定理７の証明

P_{X^n} を $\{\varphi_n(m) : m \in \mathcal{M}_n\}(\subset \mathcal{X}^n)$ 上の一様分布によって定めると、誤り率の定義から

$$\epsilon = 1 - \sum_{(x^n,y^n)\in\mathcal{X}^n\times\mathcal{Y}^n} \Psi_c(x^n, y^n)W^n(y^n|x^n)P_{X^n}(x^n)$$

が成り立つ。従って、$\Psi_c \in \mathcal{S}(\epsilon)$ である。一方、

$$\log M_n = -\log \sum_{(x^n,y^n)\in\mathcal{X}^n\times\mathcal{Y}^n} \Psi_c(x^n, y^n)P_{X^n}(x^n)P_{Y^n}(y^n)$$

が成り立つ。

定理7の証明

これから直ちに、

$$
\begin{aligned}
\log M_n &\le \sup_{\Psi \in \mathcal{S}(\epsilon)} -\log \sum_{(x^n, y^n) \in \mathcal{X}^n \times \mathcal{Y}^n} \Psi(x^n, y^n) \\
&\quad \times P_{X^n}(x^n) P_{Y^n}(y^n) \\
&\le \sup_{P_{X^n}} \sup_{\Psi \in \mathcal{S}(\epsilon)} -\log \sum_{(x^n, y^n) \in \mathcal{X}^n \times \mathcal{Y}^n} \Psi(x^n, y^n) \\
&\quad \times P_{X^n}(x^n) P_{Y^n}(y^n)
\end{aligned}
$$

が得られる。尚、式 (14) は smooth 最小 Rényi ダイバージェンスの定義から直ちに導かれる。 Q.E.D.

ϵ-通信路容量

定理 8

一般通信路 $\mathbf{W}$ が与えられたとき、任意の入力過程 $\mathbf{X}$ と任意の $\epsilon \in [0, 1)$ について、

$$C(\epsilon|\mathbf{W}) = \lim_{\delta \downarrow 0} \liminf_{n \to \infty} \frac{1}{n} D_0^{\epsilon+\delta}(P_{X^n Y^n} \| P_{X^n} \times P_{Y^n})$$

が成り立つ。

尚、Wang らは $C(\epsilon|\mathbf{W})$ の上界と下界を示しただけであり、この定理そのものは示していない。

歪み測度

$\mathcal{Y}$: 復元アルファベット

歪み測度

$$d_n : \mathcal{X}^n \times \mathcal{Y}^n \to [0, +\infty), \quad n = 1, 2, \ldots$$

を歪み測度と呼ぶ。

$d_n(x^n, y^n)$: $x^n (\in \mathcal{X}^n)$ と $y^n (\in \mathcal{Y}^n)$ の間の歪み

$d_n(x^n, y^n)/n$: 情報源記号 1 個あたりの歪み

一様可積分性

定義 7 (一様可積分性)

一般情報源 $\mathbf{X}$ と歪み測度 d_n に対し、ある基準語 $r^n \in \mathcal{Y}^n$ $(n = 1, 2, \ldots)$ と呼ばれる系列が存在して、

$$\lim_{u\to\infty} \sup_{n\geq 1} \sum_{x^n\in\mathcal{X}^n : d_n(x^n, r^n)/n \geq u} P_{X^n}(x^n) \left(\frac{1}{n} d_n(x^n, r^n)\right) = 0 \tag{16}$$

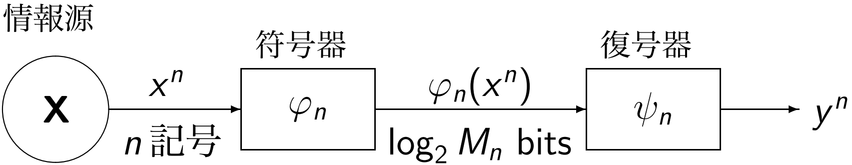
が成り立つとき、歪み測度 d_n は r^n $(n = 1, 2, \ldots)$ を基準語として一般情報源 $\mathbf{X}$ に関し一様可積分性を満たすという。

情報源の歪みを許した固定長符号化

- 系列長: $n = 1, 2, \ldots$
- 歪み測度: $d_n : \mathcal{X}^n \times \mathcal{Y}^n \to [0, +\infty)$
- 整数の集合: $\mathcal{M}_n = \{1, 2, \ldots, M_n\}$
- 符号器: $\varphi_n : \mathcal{X}^n \to \mathcal{M}_n$
- 復号器: $\psi_n : \mathcal{M}_n \to \mathcal{Y}^n$
- 一般情報源 $\mathbf{X}$ からの出力系列を表す確率変数: X^n
- 符号化レート:

$$r_n \triangleq \frac{1}{n} \log M_n$$

情報源の歪みを許した固定長符号化



1 記号あたりの歪み $d(x^n, y^n)/n$ を一定値 D 以下にしたまま、符号化レート $r = \frac{1}{n} \log M_n$ を最小化したい。

(a) 最大歪み基準

符号 (φ_n, ψ_n) が与えられたとき、

$$\operatorname*{p\text{-}\limsup}_{n\to\infty} \frac{1}{n} d_n(X^n, \psi_n(\varphi_n(X^n)))$$

を最大歪みと呼ぶ。但し、確率変数の列 $\{Z_n\}_{n=1}^{\infty}$ の確率的上極限 p-lim sup は

$$\operatorname*{p\text{-}\limsup}_{n\to\infty} Z_n \triangleq \inf\{\alpha : \lim_{n\to\infty} \Pr\{Z_n > \alpha\} = 0\}$$

によって定義される。

(a) 最大歪み基準

定義 8

一般情報源 $\mathbf{X}$ に対し、(R,D) が *fm-達成可能*であるとは、

$$\text{p-}\limsup_{n\to\infty} \frac{1}{n} d_n(X^n, \psi_n(\varphi_n(X^n))) \le D$$

かつ

$$\limsup_{n\to\infty} \frac{1}{n} \log M_n \le R$$

を満足する符号列 $\{(\varphi_n, \psi_n)\}_{n=1}^{\infty}$ が存在することである。

(a) 最大歪み基準

定義 8(続き)

更に、最大歪み D が定められたとき、レート・歪み関数 $R_{fm}(D|\mathbf{X})$ を

$$R_{fm}(D|\mathbf{X}) \triangleq \inf\{R : (R, D) \text{ が } fm\text{-達成可能}\}$$

によって定義する。

(b) 平均歪み基準

符号 (φ_n, ψ_n) が与えられたとき、

$$E\left[\frac{1}{n} d_n(X^n, \psi_n(\varphi_n(X^n)))\right]$$

を平均歪みと呼ぶ。

(b) 平均歪み基準

定義 9

一般情報源 $\mathbf{X}$ に対し、(R,D) が *fa-達成可能*であるとは、

$$\limsup_{n\to\infty} E\left[\frac{1}{n} d_n(X^n, \psi(\varphi_n(X^n)))\right] \le D$$

かつ

$$\limsup_{n\to\infty} \frac{1}{n} \log M_n \le R$$

を満足する符号列 $\{(\varphi_n, \psi_n)\}_{n=1}^{\infty}$ が存在することである。

(b) 平均歪み基準

定義 9(続き)

更に、平均歪み D が定められたとき、**レート・歪み関数 $R_{fa}(D|\mathbf{X})$** を

$$R_{fa}(D|\mathbf{X}) \triangleq \inf\{R : (R, D) \text{ が } fa\text{-達成可能}\}$$

によって定義する。

符号化逆定理

定理 9 (符号化逆定理)

任意の関数 $f_n : \mathcal{X}^n \to \mathcal{Y}^n$ に対して、$\hat{Y}^n = f_n(X^n)$ とおく。このとき、任意の $\delta \in (0,1)$ に対して、

$$
\begin{aligned}
\log \|f_n\| &\geq \sup_{Q_{X^n}} D_\infty^\delta(Q_{X^n} \times P_{\hat{Y}^n|X^n} \| Q_{X^n} \times Q_{\hat{Y}^n}) + \log \delta \\
&\geq D_\infty^\delta(P_{X^n\hat{Y}^n} \| P_{X^n} \times P_{\hat{Y}^n}) + \log \delta \qquad (17)
\end{aligned}
$$

が成り立つ。

定理 9 の証明

部分集合 $B_\delta^n(\subset \mathcal{Y}^n)$ と関数 $\Phi_o : \mathcal{X}^n \times \mathcal{Y}^n \to [0,1]$ を

$$\begin{aligned} B_\delta^n &\triangleq \left\{ y^n \in \mathcal{Y}^n : Q_{\hat{Y}^n}(y^n) \geq \frac{\delta}{\|f_n\|} \right\}, \\ \Phi_o(x^n, y^n) &\triangleq 1[(x^n, y^n) \in \mathcal{X}^n \times B_\delta^n] \end{aligned} \tag{18}$$

によって定義する。このとき、

$$\sum_{(x^n, y^n) \in \mathcal{X}^n \times \mathcal{Y}^n} \Phi_o(x^n, y^n) Q_{X^n}(x) P_{\hat{Y}^n | X^n}(y^n | x^n) \;>\; 1 - \delta$$

が成り立つ。

定理 9 の証明

これから、

$$
\begin{aligned}
&D_\infty^\delta(Q_{X^n} \times P_{\hat{Y}^n|X^n} \| Q_{X^n} \times Q_{\hat{Y}^n}) \\
&\le \log \sup_{(x^n,y^n)\in\mathcal{X}^n\times\mathcal{Y}^n} \frac{\Phi_o(x^n,y^n) Q_{X^n}(x^n) P_{\hat{Y}^n|X^n}(y^n|x^n)}{Q_{X^n}(x^n) Q_{\hat{Y}^n}(y^n)} \\
&= \log \sup_{(x^n,y^n)\in\mathcal{X}^n\times B_\delta^n} \frac{P_{\hat{Y}^n|X^n}(y^n|x^n)}{Q_{\hat{Y}^n}(y^n)} \\
&\le \log \sup_{x^n\in\mathcal{X}^n} \sum_{y^n\in\mathcal{Y}^n} \frac{P_{\hat{Y}^n|X^n}(y^n|x^n)}{\delta/\|f_n\|}
\end{aligned}
$$

定理 9 の証明

$$
\begin{aligned}
& D_{\infty}^{\delta}(Q_{X^n} \times P_{\hat{Y}^n|X^n} \| Q_{X^n} \times Q_{\hat{Y}^n}) \\
\leq \; & \log \sup_{x^n \in \mathcal{X}^n} \sum_{y^n \in \mathcal{Y}^n} \frac{P_{\hat{Y}^n|X^n}(y^n|x^n)}{\delta / \|f_n\|} \\
= \; & \log \sup_{x^n \in \mathcal{X}^n} \frac{\|f_n\|}{\delta} \\
= \; & \log \|f_n\| - \log \delta
\end{aligned}
$$

が得られる。左辺において Q_{X^n} に関して sup を取ることで、最初の不等式が導ける。2 番目の不等式は自明である。 Q.E.D.

レート・歪み関数 $R_{fm}(D|\mathbf{X})$

定理 11 (最大歪み基準に対するレート・歪み関数)

一般情報源 $\mathbf{X}$ と任意の歪み測度 d_n が与えられたとき、任意の $D > 0$ に対し、レート・歪み関数 $R_{fm}(D|\mathbf{X})$ は、

$$
\begin{aligned}
R_{fm}(D|\mathbf{X}) = \lim_{\epsilon \downarrow 0} \lim_{\delta \downarrow 0} \limsup_{n \to \infty} \\
\inf_{\substack{P_{Y^n|X^n}: \\ \Pr\{(1/n) d_n(X^n, Y^n) > D\} \le \epsilon}} \frac{1}{n} D_\infty^\delta (P_{X^n Y^n} \| P_{X^n} \times P_{Y^n}) \qquad (19)
\end{aligned}
$$

によって定まる。

レート・歪み関数 $R_{fa}(D|\mathbf{X})$

定理 13 (平均歪み基準に対するレート・歪み関数)

一般情報源 $\mathbf{X}$ が与えられたとき、歪み測度 d_n は、$r^n \in \mathcal{Y}^n$ を基準語として一般情報源 $\mathbf{X}$ に関して一様可積分性を満たすとする。このとき、レート・歪み関数 $R_{fa}(D|\mathbf{X})$ は、

$$R_{fa}(D|\mathbf{X}) = \lim_{\epsilon \downarrow 0} \lim_{\delta \downarrow 0} \limsup_{n \to \infty} \inf_{\substack{P_{Y^n|X^n}: \\ (1/n)E[d_n(X^n, Y^n)] \le D + \epsilon}} \frac{1}{n} D_{\infty}^{\delta}(P_{X^nY^n} \| P_{X^n} \times P_{Y^n}) \tag{20}$$

によって定まる。

むすび

情報理論の下記の問題について smooth エントロピーや smooth Rényi ダイバージェンスによる再定式化と証明手法について解説した。

- 情報源の固定長符号化
- 情報源を用いた一般乱数生成問題
- 通信路符号化問題
- レート・歪み理論

今後は、多端子情報理論への応用等を検討する予定である。

むすび

お先に勉強させて頂きました。