

2019年5月20日（川端）

データ圧縮基礎資料

7 レート歪み理論

7.1 歪を許容する符号化とレート歪み定理

本節では, 以下のような, 固定長の情報源符号化スキームを考える. 離散定常無記憶情報源に限定する. 情報源アルファベット $\mathcal{X}$ ($\#\mathcal{X} < \infty$) 上に確率分布 $p_X(x), x \in \mathcal{X}$ が与えられている. $\mathcal{X}$ と再生アルファベット $\mathcal{Y}$ ($\#\mathcal{Y} < \infty$) との間に歪み尺度 $\rho(x, y), x \in \mathcal{X}, y \in \mathcal{Y}$ が定義されている. 確率遷移行列 ($p_{Y|X}(y|x)$) を与えると, $\mathcal{X} \times \mathcal{Y}$ 上に同時確率 $p_X(x)p_{Y|X}(y|x)$ ならびに相互情報量

$$I(p_X, p_{Y|X})(= I(X; Y)) = \sum_{x, y} p_X(x)p_{Y|X}(y|x) \log \frac{p_{Y|X}(y|x)}{\sum_{x'} p_X(x')p_{Y|X}(y|x')}$$

が定まる. $D \geq 0$ に対して,

$$R(D) = \inf_{p_{Y|X}: E[\rho(X, Y)] \leq D} I(p_X, p_{Y|X})$$

を離散定常情報源 $X^n, p_{X^n}(x^n) = \prod_{l=1}^n p_X(x_l)$ のレート歪み関数と呼ぶ. n -ブロック符号器を

$$\varphi_n: \mathcal{X}^n \rightarrow \mathcal{M} = \{1, \dots, M\},$$

n -ブロック復号器を

$$\psi_n: \mathcal{M} = \{1, \dots, M\} \rightarrow \mathcal{Y}^n$$

と定義する. また, ブロック平均歪み $\bar{\rho}: \mathcal{X}^n \times \mathcal{Y}^n \rightarrow [0, \infty)$ を

$$\bar{\rho}(x^n, y^n) = \frac{1}{n} \sum_{l=1}^n \rho(x_l, y_l)$$

と定義する. 以上の設定のもとで, 次の定理が成り立つ.

定理 1 (レート歪み定理, 歪みを許容する情報源符号化定理)

順定理 どんな $\delta > 0$ に対しても, ある $n_0 > 0$ より大きな n をとれば, 不等式

$$\#\varphi_n(\mathcal{X}^n) \leq 2^{n(R(D)+\delta)}$$

および

$$E[\bar{\rho}(X^n, \psi_n \circ \varphi_n(X^n))] \leq D + \delta$$

を同時に満足する φ_n, ψ_n が存在する.

逆定理 どんな $n > 0$ ならびにどんな φ_n, ψ_n に対しても不等式

$$\#\varphi_n(\mathcal{X}^n) \leq 2^{n\tilde{R}}$$

および

$$E[\bar{\rho}(X^n, \psi_n \circ \varphi_n(X^n))] \leq D$$

を満たす $\tilde{R}$ と D は不等式

$$\tilde{R} \leq R(D)$$

を満足していなければならない.

7.2 レート歪み定理（逆定理の証明）

逆定理から先に証明する。まず、

補題 1 関数 $R(D)$ は単調減少かつ凸である。

証明: $R(D)$ の定義より、 $D \uparrow$ なら、 $p_{Y|X}$ に対する条件が緩和、したがって $\inf \downarrow$ 。次に、 $R(D)$ の凸性、すなわち $\bar{\lambda} = 1 - \lambda$ と表すとき、

$$R(\lambda D_1 + \bar{\lambda} D_2) \leq \lambda R(D_1) + \bar{\lambda} R(D_2)$$

をいう。 $p_{Y|X}^{(1)}, p_{Y|X}^{(2)}$ はそれぞれ D_1, D_2 に対する $R(D_1), R(D_2)$ を達成するものとする。

$$p_{Y|X}^0 = \lambda p_{Y|X}^{(1)} + \bar{\lambda} p_{Y|X}^{(2)}$$

とするとき、もし $I(p_X, p_{Y|X})$ が $p_{Y|X}$ について凸であること、

$$I(p_X, p_{Y|X}^0) \leq \lambda I(p_X, p_{Y|X}^{(1)}) + \bar{\lambda} I(p_X, p_{Y|X}^{(2)})$$

がいえたなら、 $R(D)$ を定義する最小化の制約条件は D および $p_{Y|X}$ に関する線形制約であるから、

$$R(\lambda D_1 + \bar{\lambda} D_2) \leq I(p_X, p_{Y|X}^0)$$

となり、証明が終わる。そこで、仮定した相互情報量の凸性であるが、

$$p_Y^{(i)} = \sum_{x \in \mathcal{X}} p_X(x) p_{Y|X}^{(i)}(y|x)$$

とかくと $\log$ 和不等式により、

$$\begin{aligned} & \sum_{x,y} p_X(x) p_{Y|X}^{(0)}(y|x) \log \frac{p_{Y|X}^{(0)}(y|x)}{p_Y^{(0)}(y)} \\ & \leq \sum_{x,y} \lambda p_X(x) p_{Y|X}^{(1)}(y|x) \log \frac{\lambda p_{Y|X}^{(1)}(y|x)}{\lambda p_Y^{(1)}(y)} + \sum_{x,y} \bar{\lambda} p_X(x) p_{Y|X}^{(2)}(y|x) \log \frac{\bar{\lambda} p_{Y|X}^{(2)}(y|x)}{\bar{\lambda} p_Y^{(2)}(y)} \end{aligned} \quad (1)$$

からいえている。証明終

証明:(逆定理) まず $Y^n = \psi_n(\varphi(X^n))$ と定義しよう。

$$\begin{aligned} n\tilde{R} & \geq H(Y^n) & (\sharp \psi_n(\varphi_n(\mathcal{X}^n)) \leq \sharp \varphi_n(\mathcal{X}^n) \leq 2^{n\tilde{R}}) \\ & \geq I(X^n; Y^n) & (*) \\ & = H(X^n) - H(X^n|Y^n) & (*) \\ & = \sum_{l=1}^n H(X_l) - \sum_{l=1}^n H(X_l|X^{l-1}, Y^n) & (\text{無記憶性, チェインルール}) \\ & \geq \sum_{l=1}^n H(X_l) - \sum_{l=1}^n H(X_l|Y_l) & (*) \\ & = \sum_{l=1}^n I(X_l; Y_l) & (*) \\ & \geq \sum_{l=1}^n R(E[\rho(X_l, Y_l)]) & (R(D) \text{ の定義}) \\ & \geq nR(E[\bar{\rho}(X^n, Y^n)]) & (\bar{\rho} \text{ の定義, } R(D) \text{ の凸性}) \\ & \geq nR(D) & (R(D) \text{ は } \downarrow) \end{aligned}$$

上記で $(*)$ はエントロピー、相互情報量の初等的性質を用いたことを示す。証明終

7.3 順定理の証明の準備

証明の前に、定義、といくつか関連した補題を述べる。

定義 1 $\mathcal{X} \times \mathcal{Y}$ 上の確率分布 $p(x, y)$ が与えられたとき、次の4つの不等式

$$\begin{aligned} & \left| -\frac{1}{n} \log p(x^n) - H(X) \right| < \varepsilon, \quad \left| -\frac{1}{n} \log p(y^n) - H(Y) \right| < \varepsilon \\ & \left| -\frac{1}{n} \log p(x^n, y^n) - H(X, Y) \right| < \varepsilon, \quad \left| \bar{\rho}(x^n, y^n) - E[\rho(X, Y)] \right| < \varepsilon \end{aligned}$$

を満足する対を ε -歪標準系列と呼び、その集合を $A_{\varepsilon, \rho}^{(n)}$ とかく。

補題 2 $(X^n, Y^n) \sim \prod_{l=1}^n p(x_l, y_l)$ とするとき、

$$\lim_{n \rightarrow \infty} P((X^n, Y^n) \in A_{\varepsilon, \rho}^{(n)}) = 1.$$

□

証明:

$$\begin{aligned} P\left((X^n, Y^n) \notin A_{\varepsilon, \rho}^{(n)}\right) &\leq \\ &P\left(\left| -\frac{1}{n} \log p(x^n) - H(X) \right| \geq \varepsilon\right) + P\left(\left| -\frac{1}{n} \log p(y^n) - H(Y) \right| \geq \varepsilon\right) \\ &+ P\left(\left| -\frac{1}{n} \log p(x^n, y^n) - H(X, Y) \right| \geq \varepsilon\right) + P\left(\left| \bar{\rho}(x^n, y^n) - E[\rho(X, Y)] \right| \geq \varepsilon\right) \end{aligned}$$

右辺の各確率は大数の弱法則（チェビシェフの不等式）により $\sigma^2/(\varepsilon^2 n)$ と上界できる。ただし、 σ^2 は、各々の場合、

$$-\log p(X_l), \quad -\log p(Y_l), \quad -\log p(X_l, Y_l), \quad \rho(X_l, Y_l)$$

の分散である。 n を十分大にとればそれらの和は任意に小さくできる。証明終

補題 3 $0 \leq \alpha, \beta \leq 1$ および $n \geq 1$ に対し、不等式 $(1 - \alpha\beta)^n \leq 1 - \alpha + e^{-\beta n}$ が成り立つ。□

証明: $(1 - \alpha\beta)^n$ は α について凸関数であるから、 $\alpha = 0, 1, \alpha$ の3点における値について不等式

$$(1 - \alpha\beta)^n \leq (1 - \alpha)(1 - \beta \cdot 0)^n + \alpha(1 - \beta \cdot 1)^n$$

が成立。右辺の第2項については

$$\alpha(1 - \beta)^n \leq \alpha \cdot e^{-n\beta} \leq e^{-n\beta}$$

が成立。証明終

7.4 レート歪み定理（順定理の証明）

符号器 φ_n , 復号器 ψ_n の構成法を与える. $R(D)$ を達成する $p(y|x)$ を用いて, $p(x)$ との同時分布から $\mathcal{Y}^n$ に関する周辺確率分布 $p(y^n) = \sum_{x^n \in \mathcal{X}^n} p(x^n)p(y^n|x^n)$ を求め, これに従って独立に $M := 2^{n\hat{R}}$ 本の再生ベクトル $y^{n(i)}, 1 \leq i \leq M$ を発生させる. その集合を $\mathcal{C} := \{y^{n(i)}, 1 \leq i \leq M\}$ とかく.

$$\mathcal{I}(x^n) := \left\{ i : (x^n, y^{n(i)}) \in A_{\varepsilon, \rho}^{(n)} \right\}$$

とすると,

$$\varphi_n(x^n) := \begin{cases} \min_{i \in \mathcal{I}(x^n)} i & \text{if } \mathcal{I}(x^n) \neq \phi \\ 1 & \text{otherwise} \end{cases}$$

と定義し, さらに $\psi_n(i) := y^{n(i)}$ と定義する. こうすると, $\mathcal{I}(x^n) \neq \phi$ であれば, $\psi_n(\varphi_n(x^n)) \in A_{\varepsilon, \rho}^{(n)}$ だから

$$E \left[\bar{\rho}(X^n, \psi_n(\varphi_n(X^n))) \right] \leq P(\mathcal{I}(X^n) \neq \phi) (E[\rho(X, Y)] + \varepsilon) + P(\mathcal{I}(X^n) = \phi) \max_{x, y} \rho(x, y)$$

が成立する. $P_e := P(\mathcal{I}(X^n) = \phi)$ と置き, 両辺の $\mathcal{C}$ についての期待値をとると, さらに

$$E_{\mathcal{C}} \left[E[\bar{\rho}(X^n, \psi_n(\varphi_n(X^n))) \right] \leq D + \varepsilon + \max_{x, y} \rho(x, y) E_{\mathcal{C}}[P_e]$$

となる. 以下 $E_{\mathcal{C}}[P_e]$ を評価する.

$$\begin{aligned} E_{\mathcal{C}}[P_e] &= E_{\mathcal{C}} \left[\sum_{x^n} p(x^n) \prod_{i=1}^M \mathbf{1} \left\{ (x^n, Y^{n(i)}) \notin A_{\varepsilon, \rho}^{(n)} \right\} \right] \\ &\leq \sum_{x^n} p(x^n) \left(1 - \sum_{y^n} p(y^n) \mathbf{1} \left\{ (x^n, y^n) \in A_{\varepsilon, \rho}^{(n)} \right\} \right)^M \end{aligned} \quad (2)$$

さて, $(x^n, y^n) \in A_{\varepsilon, \rho}^{(n)}$ ならば, 不等式

$$\begin{aligned} p(y^n) &= \frac{p(y^n)p(x^n)p(y^n|x^n)}{p(x^n, y^n)} \\ &\geq 2^{-n(H(X)+\varepsilon)} \cdot 2^{-n(H(Y)+\varepsilon)} \cdot 2^{n(H(X, Y)-\varepsilon)} p(y^n|x^n) \\ &\geq 2^{-n(I(X; Y)+3\varepsilon)} p(y^n|x^n) \\ &= 2^{-n(R(D)+3\varepsilon)} p(y^n|x^n) \end{aligned}$$

が定義より導ける. これを用いると, 式 (2) はさらに

$$\begin{aligned} E_{\mathcal{C}}[P_e] &\leq \sum_{x^n} p(x^n) \left(1 - 2^{-n(R(D)+3\varepsilon)} \sum_{y^n} p(y^n|x^n) \mathbf{1} \left\{ (x^n, y^n) \in A_{\varepsilon, \rho}^{(n)} \right\} \right)^M \\ &\leq \sum_{x^n} p(x^n) \left(1 - \sum_{y^n} p(y^n|x^n) \mathbf{1} \left\{ (x^n, y^n) \in A_{\varepsilon, \rho}^{(n)} \right\} + \exp\{-2^{n(\hat{R}-R(D)-3\varepsilon)}\} \right) \\ &\leq \varepsilon + \exp\{-2^{n(\hat{R}-R(D)-3\varepsilon)}\} \end{aligned}$$

のように上界できる. ただし, 第2の不等式で補題3を用い, 最後の不等式では, $A_{\varepsilon, \rho}^{(n)}$ の定義を用いた. ここで $\hat{R} > R(D) + 3\varepsilon$ ととりさえすれば, $E_{\mathcal{C}}[P_e]$ は十分小にできるので, 定理がほぼ成立することが理解できる. 詳しくは, 形式的に定理を完成すればよい. 証明終