CIE6020/MAT3350
Selected Topics in Information Theory

Lecture 8: Typical Set

22 February 2019

The Chinese University of Hong Kong, Shenzhen

Weak Typical Set

Lemma

Consider a DMS $\mathcal{A}$ with distribution Q. The probability that an n-length sequence $\mathbf{x}$ is generated is

$$\begin{aligned}
Q^n(\mathbf{x}) &= \prod_{a\in\mathcal{A}} Q(a)^{N(a|\mathbf{x})} \\
&= \prod_{a\in\mathcal{A}} Q(a)^{nP_{\mathbf{x}}(a)} \\
&= 2^{\sum_{a\in\mathcal{A}} nP_{\mathbf{x}(a)} \log Q(a)} \\
&= 2^{-nH(P_{\mathbf{x}})-nD(P_{\mathbf{x}}||Q)}.
\end{aligned}$$

Lemma
For any type P of sequences in $\mathcal{A}^n$ and distribution Q on $\mathcal{A}$,

$$(n+1)^{-|\mathcal{A}|} 2^{-nD(P||Q)} \leq Q^n(T_P^n) \leq 2^{-nD(P||Q)}. \tag{1}$$

Proof.
By Lemma 1

$$Q^n(T_P) = |T_P| 2^{-nH(P)-nD(P||Q)}. \tag{2}$$

The proof is compeled by the bound on $|T_P|$. $\square$

Convergence in Probability

- Let $X^n = (X_1, \ldots, X_n)$ be the i.i.d. sequence with distribution p sampled from $\mathcal{A}$.
- By the weak law of large numbers, P_{X^n} converges in probability to p when $n \to \infty$, i.e.,

$$\lim_{n\to\infty} \Pr\{|P_{X^n}(x) - p(x)| > \delta\} = 0, \forall x \in \mathcal{A}.$$

- Hence, $D(P_{X^n}||p) + H(P_{X^n}) = -\sum_{x\in\mathcal{A}} P_{X^n}(x) \log p(x) \to H(p)$ in probability as $n \to \infty$.

Weak Typical Set

- Fix $\delta > 0$.
- Let

$$\begin{aligned} W_\delta^{(n)} &= \{\mathbf{x} \in \mathcal{A}^n : |D(P_\mathbf{x}||p) + H(P_\mathbf{x}) - H(p)| \le \delta\} \\ &= \bigcup_{\text{type } P \text{ of } \mathcal{A}^n : |D(P||p)+H(P)-H(p)| \le \delta} T_P^n. \end{aligned}$$

Lemma

1. *For any* $\delta > 0$, $\lim_{n\to\infty} \Pr\{X^n \in W_\delta^{(n)}\} = 1$.
2. *For any* $\delta > 0$ *and sufficiently large* n, $|W_\delta^{(n)}| \le 2^{n(H(p)+\delta)}$.

Block Source Coding Theorem

Theorem (Block Source Coding Theorem)

For a discrete memoryless source with distribution p,

$$\lim_{n\to\infty} \frac{\log M^*(n,\epsilon)}{n} = H(X), \text{ for every } \epsilon \in (0,1).$$

Code Construction

- Let $\mathcal{C}_n = W_\delta^{(n)}$.
- For all sufficiently large n, (by Property 1)

$$P_e = \Pr\{X^n \notin \mathcal{W}_\delta^{(n)}\} \leq \epsilon.$$

- So for any $\epsilon > 0$ and all sufficiently large n, $M^*(n, \epsilon) \leq |W_\delta^{(n)}|$.
- Moreover, (by Property 2)

$$\lim_{n\to\infty} \frac{M^*(n,\epsilon)}{n} \leq \lim_{n\to\infty} \frac{\log |W_\delta^{(n)}|}{n} \leq H(p) + \delta.$$

Converse

- Consider a sequence of code $\mathcal{C}_n \subset \mathcal{A}^n$ with $\Pr\{X^n \in \mathcal{C}_n\} \geq 1 - \epsilon$.
- As $\Pr\{X^n \notin W_\delta^{(n)}\} + \Pr\{X^n \in W_\delta^{(n)} \cap \mathcal{C}_n\} \geq P(\mathcal{C}_n) \geq 1 - \epsilon$ and $\Pr\{X^n \notin W_\delta^{(n)}\} \to 0$ (Property 1), for sufficiently large n, $\Pr\{X^n \in W_\delta^{(n)} \cap \mathcal{C}_n\} \geq \frac{1-\epsilon}{2}$.
- Hence, for sufficiently large n

$$\begin{aligned}\frac{1-\epsilon}{2} &\leq \Pr\{X^n \in W_\delta^{(n)} \cap \mathcal{C}_n\} \\ &\leq |\mathcal{C}_n \cap W_\delta^{(n)}| 2^{-n(H(p)-\delta)} \\ &\leq |\mathcal{C}_n| 2^{-n(H(p)-\delta)}.\end{aligned}$$

- So for every $\delta > 0$,

$$\lim_{n\to\infty} \frac{M^*(n,\epsilon)}{n} = \lim_{n\to\infty} \min_{A \subset \mathcal{X}^n : \Pr\{X^n \in \mathcal{C}_n\} \geq 1-\epsilon} \frac{\log |\mathcal{C}_n|}{n} \geq H(p) - \delta.$$

Universal Block Source Coding

Theorem
There exists a sequence of rate R codes such that $P_e \to 0$ for every DMS Q over $\mathcal{A}$ with $H(Q) < R$.