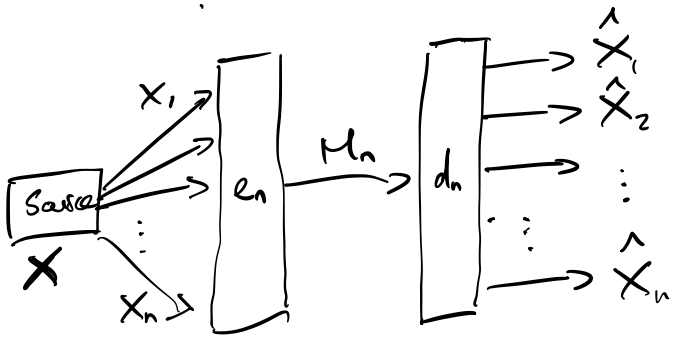
Review of Week 4: Block coding



the whole string is denoted X^n

$$R^*(X) = \inf\Big\{ R : \exists \{e_n, d_n, M_n\}_{n\in\mathbb{N}} \text{ where } M_n \in \lceil 2^{nR} \rceil \text{ s.t. } \lim_{n\to\infty} P[X^n \neq \hat{X}^n] = 0 \Big\}$$

We have shown: $R^*(X) \geq H(X)$ for any DMS

For variable length codes $\ell^*(X) \geq H(X)$

Use variable length code to encode X^n

$\Rightarrow n\ell^*(X)$ in expectation

due to law of large numbers, we have

$$\lim_{n\to\infty} P\left[\left|\frac{1}{n}\ell(X^n) - \ell^*(X)\right| > \delta\right] = 0 \quad \forall \delta > 0$$

$\Rightarrow$ we can encode using $n\ell^*(X) + n\delta$ bits.

(corresponds to a rate $R = \ell^*(x) + \delta$)

↳ guarantees that $P[X^n \neq \hat{X}^n] \xrightarrow[n\to\infty]{} 0$.

$\Rightarrow R^*(x) \leq H(x) + 1$

Section 2.3.3 : Proof achievability and typical sets

typical sequences $\quad P(0) = \frac{1}{3}, \; P(1) = \frac{2}{3}$

00101110111100 } typical : empirical frequency ≈ probability
11010111101 }
0000001000 } atypical
↖ too many 0's!

Def. Let $\varepsilon \in (0,1)$, X a DMS. The **ε-typical set of X**, for every $n \in \mathbb{N}$, is

$$A_\varepsilon^{(n)} := \left\{ x^n \in \mathcal{X}^n : \left| \frac{1}{n} \log \frac{1}{P_{X^n}(x^n)} - H(X) \right| \leq \varepsilon \right\}$$

$$= \frac{1}{n} \sum_{i=1}^{n} \log \frac{1}{P_X(x_i)}$$

An element of the ε-typical set is called a ε-typical sequence.

Prop. (Asymptotic Equipartition Property). Let $\varepsilon \in (0,1)$.
The sequence of typical sets $A_\varepsilon^{(n)}(X)$ for $n \in \mathbb{N}$ satisfies

1) $H(X) - \varepsilon \leq \frac{1}{n} \log \frac{1}{P_{X^n}(x^n)} \leq H(X) + \varepsilon$
for any $x^n \in A_\varepsilon^{(n)}(X)$ ✓ by definition

2) $\lim_{n\to\infty} P[X^n \in A_\varepsilon^{(n)}(X)] = 1$

3) The size of the set satisfies
$$|A_\varepsilon^{(n)}(X)| \leq 2^{n(H(X)+\varepsilon)}$$

Proof of 3) observe that $P_{X^n}(x^n) \geq 2^{-n(H(X)+\varepsilon)}$
$P_{X^n}(x^n) \leq 2^{-n(H(X)-\varepsilon)}$

$$1 \geq Pr[X^n \in A_\varepsilon^{(n)}] = \sum_{x^n \in A_\varepsilon^{(n)}} P_{X^n}(x^n)$$
$$\geq \sum_{x^n \in A_\varepsilon^{(n)}} 2^{-n(H(X)+\varepsilon)}$$
$$= |A_\varepsilon^{(n)}| \, 2^{-n(H(X)+\varepsilon)} \qquad \square$$

Proof of 2)
$$P[X^n \in A_\varepsilon^{(n)}] = P\left[\left|\frac{1}{n} \log \frac{1}{P_{X^n}(x^n)} - H(X)\right| \leq \varepsilon\right]$$

$$= P\left[\left|\frac{1}{n}\sum_{i=1}^{n}\underbrace{\left(\log\frac{1}{P_X(X_i)} - H(X)\right)}_{Z_i}\right| \le \varepsilon\right]$$

$$= 1 - P\left[\left|\frac{1}{n}\sum_{i=1}^{n} Z_i\right| > \varepsilon\right]$$

$$Z_i = \log\frac{1}{P_X(X_i)} - H(X)$$

$E[Z_i] = 0$, $Z_1, Z_2, \ldots, Z_n$ are i.i.d.

by law of large numbers : $\lim_{n\to\infty} P\left[\left|\frac{1}{n}\sum_{i=1}^{n} Z_i\right| > \varepsilon\right] = 0$ □

Proof of achievability ($R^*(X) \le H(X)$)

First fix $\varepsilon > 0$ and set $R = H(X) + 2\varepsilon$.

Define $idx : A_\varepsilon^{(n)} \to [\,|A_\varepsilon^{(n)}|\,]$

L_n : number of bits for the message $\le n(R-\varepsilon)+1$

$L_n = \lceil nH(X) + \varepsilon \rceil \le n(H(X)+\varepsilon) + 1 \le nR - 1 \le \lfloor nR \rfloor$

because $\varepsilon \ge \frac{2}{n}$ for n sufficiently large

↑ this is a rate R code!

$$e_n(x^n) = \begin{cases} idx(x^n) & \text{if } x^n \in A_\varepsilon^{(n)} \\ 0^{L_n} & \text{else} \end{cases}$$

$d_n(m)$ outputs x^n s.t. $\text{idx}(x^n) = m$

$$P[\hat{X}^n \neq X^n] =$$

$$P[\hat{X}^n \neq X^n \mid X^n \in A_\varepsilon^{(n)}] \cdot P[X^n \in A_\varepsilon^{(n)}] \quad (=0)$$

$$+ P[\hat{X}^n \neq X^n \mid X^n \notin A_\varepsilon^{(n)}] \cdot P[X^n \notin A_\varepsilon^{(n)}]$$

$$\leq 1 \qquad \xrightarrow{n\to\infty} 0$$

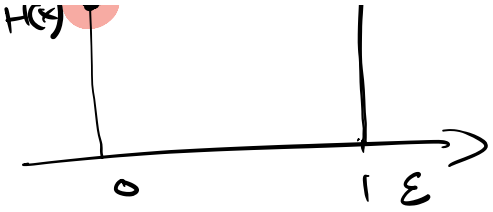
$$\Rightarrow \lim_{n\to\infty} P[\hat{X}^n \neq X^n] = 0.$$

$\Rightarrow R = H(X) + 2\varepsilon$ is achievable, where $\varepsilon > 0$ is arbitrary.

$$R^*(X) \leq H(X) \qquad \square$$

2.3.4 Strong converse

Define $R_\varepsilon^*(X) = \inf \{ R : \exists \{e_n, d_n, M_n\}_{n\in\mathbb{N}}$ where $M_n \in [2^{\lfloor nR \rfloor}]$ s.t. $\lim_{n\to\infty} P[X^n \neq \hat{X}^n] \leq \varepsilon \}$



← strong converse property

Prop.: Let $\varepsilon, \mu \in (0,1)$. Then, there exists a N_0 s.t. for all $n \geq N_0$

4) $P[X^n \in A_\varepsilon^{(n)}(X)] \geq 1-\mu$ ← by definition of the limit and property 2)

5) $|A_\varepsilon^{(n)}(X)| \geq (1-\mu) 2^{n(H(X)-\varepsilon)}$.

Theorem: For any sequence of $(\varepsilon, 2^{\lfloor nR \rfloor})$-code with $R < H(X)$, we have

$$\lim_{n\to\infty} Pr[\hat{X}^n \neq X^n] = 1.$$

Proof: Fix $R < H(X)$. Fix an encoder e_n (arbitrary)

$$e_n : \mathcal{X}^n \to [2^{\lfloor nR \rfloor}]$$

We can split $\mathcal{X}^n$ into subsets

$$D_m = \{x^n : e_n(x^n) = m\}, \quad m \in [2^{\lfloor nR \rfloor}]$$

The optimal decoder: maximum likelihood decoder

$$d_n(m) = \underset{x^n \in D_m}{\operatorname{argmax}} P_{X^n}(x^n)$$

As usual $\hat{X}^n = d_n(e_n(X^n))$

Fix $\mu > 0$, and $\varepsilon > 0$ s.t. $R < H(X) - 2\varepsilon$

$$P[\hat{X}^n = X^n] = \sum_{m=1}^{2^{\lfloor nR \rfloor}} \sum_{x^n \in D_m} P_{X^n}(x^n) \cdot \underbrace{P[\hat{X}^n = X^n \mid X^n = x^n]}_{\delta_{x^n, d_n(m)}}$$

$$= \sum_{m=1}^{2^{\lfloor nR \rfloor}} P_{X^n}(d_n(m))$$

$$= \sum_{m : d_n(m) \in A_\varepsilon^{(n)}} \underbrace{P_{X^n}(d_n(m))}_{\leq 2^{-n(H(X)-\varepsilon)}} + \sum_{m : d_n(m) \notin A_\varepsilon^{(n)}} \underbrace{P_{X^n}(d_n(m))}_{\leq \mu}$$

$$\leq \sum_{m : d_n(m) \in A_\varepsilon^{(n)}} 2^{-n(H(X)-\varepsilon)} + \mu$$

$$\leq 2^{nR} \cdot 2^{-n(H(X)-\varepsilon)} + \mu$$

$$\leq 2^{n(H(X)-2\varepsilon)} 2^{-n(H(X)-\varepsilon)} + \mu$$

$$= 2^{-n\varepsilon} + \mu$$

$$\Rightarrow \lim_{n \to \infty} P[\hat{X}^n \neq X^n] \geq 1 - \mu \quad \text{for any } \mu > 0$$

□