Lossy Compression, Asymptotic Equipartition Property

Niklas Schmidt, 07.05.2015

In the following X will be a random variable taking values in $A_X = \{a_1, \ldots, a_k\}$, having probabilities $P_X = \{p_1, \ldots, p_k\}$, with $P(X = a_i) = p_i$. We define $X^n := (X_1, \ldots, X_n)$, with i.i.d. random variables $X_1, \ldots, X_n$, all having the same distribution as X.

Definition:

- $\{0,1\}^+ := \bigcup_{n \in \mathbb{N}} \{0,1\}^n$
- A map $C : A_X^n \to \{0,1\}^+$ is called a code. $C(A_X^n)$ is the set of codewords.
- A map $D : \{0,1\}^+ \to A_X^n$ is called a decoder.
- The maximal length of C is defined as $L(C) := \max_{x \in A_X^n} \ell_x$ with $\ell_x = m$ iff $C(x) \in \{0,1\}^m$

Theorem:
Let $C : A_X^n \to \{0,1\}^+$ be a code with $H(X^n) \geq L(C)$, h^{-1} the inverse of the binary entropy function and $P_e = \frac{1}{n} \sum_{i=1}^{n} P(X_i \neq f_i(C(X^n))$ the "average bit error" for a sequence of functions $f_i : \{0,1\}^+ \to A_X$. Then

$$P_e \geq h^{-1}\left(\frac{H(X^n) - L(C)}{n}\right)$$

Definition: Raw bit content:

$$H_0(X) = \log_2 |A_X|$$

Definition:

- S_δ is the smallest subset of A_X satisfying $P(X \in S_\delta) \geq 1 - \delta$
- $H_\delta(X) := \log_2 |S_\delta|$ (Essential bit content)

Definition: Typical set:

$$T_{n\beta} := \left\{ x \in A_X^n : \left| \frac{1}{n} \log_2 \frac{1}{P(x)} - H \right| < \beta \right\}$$

where $H := H(X)$

- $2^{-n(H+\beta)} < P(x) < 2^{-n(H-\beta)} \quad \forall\, x \in T_{n\beta}$
- $2^{n(H-\beta)} < |T_{n\beta}| < 2^{n(H+\beta)}$

Theorem: (Asymptotic equipartition property)
Let $\beta > 0$. Then

$$P(X^n \in T_{n\beta}) \longrightarrow 1 \text{ for } n \to \infty$$

Theorem: (Shannon's source coding theorem)
Let X be a random variable, $H(X) =: H$, $\varepsilon > 0$, $0 < \delta < 1$. Then there exists an $N \in \mathbb{N}$ s.th. for $n > N$

$$\left| \frac{1}{n} H_\delta(X^n) - H \right| < \varepsilon$$

References

[1] MacKay, David: "Information Theory, Inference, and Learning Algorithms", Cambridge University Press; First edition (2003)

[2] Cover, Thomas M.; Thomas, Joy A.: "Elements of Information Theory", Wiley-Interscience; Second edition (2006)

[3] Wolf, Michael: "http://www-m5.ma.tum.de/foswiki/pub/M5/Allgemeines/MA5103_2012W/lecture3.pdf" (2012)

[4] Wolf, Michael: "http://www-m5.ma.tum.de/foswiki/pub/M5/Allgemeines/MA5103_2012W/lecture4.pdf" (2012)

[5] Wolf, Michael: "http://www-m5.ma.tum.de/foswiki/pub/M5/Allgemeines/MA5103_2012W/lecture6.pdf" (2012)