CSE522/Math581EB Coding Theory **Fall 2001**

Lecture 2: October 4

Lecturer: Elchanan Mossel *Scribe: Elchanan Mossel*

2.1 Noisy channels

Shannon [CS48] was interested in transmitting information over noisy channels. Mathematically, a *channel* is modeled by an input alphabet Σ_{in}, an output alphabet Σ_{out}, and a matrix $(P_{i,j})_{i\in\Sigma_{in},j\in\sigma_{out}}$, where $P_{i,j} \geq 0$ is the probability of receiving j when transmitting i. For all i, $\sum_j P_{i,j} = 1$. Two important families of channels are

- *Binary symmetric channels*, where $\Sigma_{in} = \Sigma_{out} = \{0,1\}$ and $M = \begin{pmatrix} 1-p & p \\ p & 1-p \end{pmatrix}$.
- *Binary erasure channels* where bits are not flipped, but only erased, so $\Sigma_{in} = \{0,1\}, \Sigma_{out} = \{0,1,?\}$ and $P_{i,i} = 1-p$, $P_{i,?} = p$, for $i = 0, 1$. This channel has some practical interest and is easier to hadnle than the binary symmetric channel.

2.2 Shannon noisy coding theorem

This following theorem of Shannon resulted in a revolution in information theory.

Theorem 2.1. *For every channel $\mathcal{C}$, there exists a constant $Cap(\mathcal{C})$, such that*

- *For all $R < Cap(\mathcal{C})$ there exist a pair of function $E : \{0,1\}^k \to \Sigma_{in}^{k/R}$ and $D : \Sigma_{out}^{k/R} \to \{0,1\}^k$, with the following property. If we pick a message uniformly at random from $\{0,1\}^k$, encode it with E, send the result through the noisy channel and decode the result with D, then we recover the original message with probability $1 - o(1)$,*
$$\mathbf{P}[D(\mathcal{C}(E(x))) = x] = 1 - o(1).$$
- *For all $R > Cap(\mathcal{C})$ and all functions $E : \{0,1\}^k \to \Sigma_{in}^{k/R}$ and $D : \Sigma_{out}^{k/R} \to \{0,1\}^k$, it holds that if we pick a message uniformly at random from $\{0,1\}^k$, encode it with E, send the result through the noisy channel and decode the result with D, then we recover the original message with probability $o(1)$,*
$$\mathbf{P}[D(\mathcal{C}(E(x))) = x] = 1 - o(1).$$
- *For the binary symmetric channel, $Cap(C) = 1 - H(p)$, where*
$$H(p) = -p\log_2 p - (1-p)\log_2 p.$$
- *For the erasure channel $Cap(C) = 1 - p$.*

2.3 The Hamming metric

Definition 2.2.

- *For $x, y \in [q]^n$, the* Hamming distance *between x and y is*

$$\Delta(x, y) = \#\{1 \le i \le n : x_i \neq y_i\}.$$

- *The Hamming ball of radius r centered at y is*

$$B(y, r) = \{x : \Delta(x, y) \le r\}.$$

- *$Vol(r, n)$ denotes the volume of any r-radius ball in $\{0,1\}^n$.*

By Striling formula

$$n! = (1 + o(1))\sqrt{2\pi n}(n/e)^n = 2^{o(1)n}(n/e)^n.$$

Therefore for $0 < p < 1/2$,

$$\begin{aligned} Vol(pn, n) &= |B(pn, r)| = \sum_{i=0}^{pn} \binom{n}{i} \sim \binom{n}{pn} \Big(\text{because } \frac{\binom{n}{j+1}}{\binom{n}{j}} = \frac{n-j}{j+1} \ge \frac{1-p}{p}\Big) \\ &= \frac{n!}{(pn)!((1-p)n)!} = 2^{o(1)n} \frac{(n/e)^n}{(pn/e)^{pn}((1-p)n/e)^{(1-p)n}} \\ &= 2^{o(1)n} \frac{1}{p^{pn}(1-p)^{(1-p)n}} = 2^{(H(p)+o(1))n}. \end{aligned}$$

2.4 Upper bound for binary symmetric channels

We will use Chernoff bound in the proof, see .e.g. Appendix A of [AS92].

Lemma 2.3. *If $X_1, \ldots, X_n$ are i.i.d. random variables taking values 0 and 1, then*

$$\mathbf{P}[|\sum_{i=1}^{n} X_i - \mathbf{E}[\sum_{i=1}^{n} X_i]| > \epsilon n] \le 2\exp(-\epsilon^2 n/2).$$

Proof of Upper bound of noisy coding theorem for binary symmetric channels: We fix $0 < p < 1/2$. Take $E : \{0,1\}^k \to \{0,1\}^n$ to be a random function. The decoding function $D : \{0,1\}^n \to \{0,1\}^k$ is defined as follows: for $y \in \{0,1\}^n$, define $D(y)$ to be $m \in \{0,1\}^k$ such that $\Delta(E(m), y)$ is minimized (if there is more than one such m, choose one of them in an arbitrary manner).

Fix $\epsilon > 0$,m and $E(m)$. Let y be the received vector. Let $r = (p + \epsilon)n$. In order that $D(y) \neq m$ one of the following events should hold.

- $y \notin B(E(m), r)$, i.e. there are at least r errors.
- There exists $m' \neq m$ s.t. $E(m') \in B(y, r)$.

We note that

- By Chernoff bound the probability of the first event is bounded by $2\exp(-\epsilon^2 n/2)$.

- Fix m'. Since $E(m')$ is random,

$$\mathbf{P}[E(m') \in B(y,r)] = \frac{Vol(B(y,(p+\epsilon)n))}{2^n} = 2^{(H(p+\epsilon)-1+o(1))n}.$$

So taking a union bound over all $m \neq m' \in \{0,1\}^k$,

$$\mathbf{P}[\exists m' \neq m \text{ s.t. } E(m') \in B(E(m),r)] \leq 2^{(H(p+\epsilon)-1+o(1))n+k}$$

The total decoding error probability is bounded by

$$2\exp(-\epsilon^2 n/2) + 2^{(H(p+\epsilon)-1+o(1))n+k}$$

This probability is exponentially small if $n = k/R$ and $R < 1 - H(p+\epsilon)$. ■

2.5 Lower bound for binary symmetric channel

Proof of upper bound of noisy coding theorem for binary symmetric channels: Fix $1/2 > p > 0$ and $\epsilon > 0$ s.t. $p + \epsilon < 1/2$. Let $E : \{0,1\}^k \to \{0,1\}^n$ and $D : \{0,1\}^n \to \{0,1\}^k$ be an arbitrary pair of encoding/decoding functions. For a message m, let S_m be the set of all $y \in \{0,1\}^n$ which decode as m. For $m \in \{0,1\}^k$, let A_m be the event that the received message is m given that the transmitted message is $E(m)$. Let A be the event that for a random message m, the received message is m, when the transmitted message is $E(m)$. We will show that

$$\mathbf{P}[A_m] \leq 2\exp(-\epsilon^2 n/2) + \frac{|S_m|}{2^{(H(p-\epsilon)+o(1))n}}. \tag{2.1}$$

Thus

$$\begin{aligned}\mathbf{P}[A] &= 2^{-k}\sum_{m\in\{0,1\}^k}\mathbf{P}[A_m] \leq 2\exp(-\epsilon^2 n/2) + 2^{-k}\sum_m \frac{|S_m|}{2^{H(p-\epsilon)n}} \\ &\leq 2\exp(-\epsilon^2 n/2) + 2^{(1-H(p-\epsilon)+o(1))n-k}.\end{aligned}$$

The right hand side decays exponentially when $R = k/n > 1 - H(p-\epsilon)$, thus proving the theorem.

In order to prove (2.1), let y be the received word, and let η be the total number of errors. By Chernoff bound $\mathbf{P}[|\eta - pn| > \epsilon n] \leq 2\exp(-\epsilon^2 n/2)$. When $(p-\epsilon)n \leq r \leq (p+\epsilon)n$, the number of y's with r errors is

$$\binom{n}{r} \geq \binom{n}{(p-\epsilon)n} = 2^{(H(p-\epsilon)+o(1))n}.$$

All the words with r errors have the same probability of being received given that $E(m)$ was transmitted:

$$\mathbf{P}[y \text{ received when } E(m) \text{ transmitted }] = \frac{1}{\binom{n}{r}} \leq 2^{-(H(p-\epsilon)+o(1))n}.$$

Now

$$\begin{aligned}\mathbf{P}[A_m] &\leq \mathbf{P}[P[|\eta - pn| > \epsilon n] + \sum_{y\in S_m,(p-\epsilon)n\leq D(y,E(m))\leq(p+\epsilon)n}\mathbf{P}[y \text{ received when } E(m) \text{ transmitted }] \\ &\leq 2\exp(-\epsilon^2 n/2) + \frac{|S_m|}{2^{(H(p-\epsilon)+o(1))n}}.\end{aligned}$$

So (2.1) follows. ■

References

[1] N. Alon and J. H. Spencer (1992) *The probabilistic method*. With an appendix by Paul Erds, John Wiley and Sons.

[2] Shannon (1948), C. E. A mathematical theory of communication. *Bell System Tech. J.* **27**, 379–423, 623–656.