LECTURE 11

Last time:

- The channel coding theorem overview
- Upper bound on the error probability
- Bound on not being typical
- Bound on too many elements being typical
- Coding theorem (weak)

Lecture outline

- Strong coding theorem
- Revisiting channel and codes
- Bound on probability of error
- Error exponent

Revisiting channel and codes

Consider a DMC with transition probabilities $P_{Y|X}(y|x)$

For any block length N, let

$P_{\underline{Y}^N|\underline{X}^N}(\underline{y}^N|\underline{x}^N) = \prod_{i=1}^{N} P_{Y|X}(y_i|x_i)$

$P_{\underline{X}^N}(\underline{x}^N) = \prod_{i=1}^{N} P_X(x_i)$

$P_{\underline{Y}^N}(\underline{y}^N) = \prod_{i=1}^{n} P_Y(y_i)$

in particular we can select the input probability to be capacity-achieving, since IID inputs yield capacity for a DMC

the output alphabet $\mathcal{Y}$ and the input alphabet $\mathcal{X}$ may be different

Revisiting channel and codes

The code is a block code with bit blocks of length L being mapped onto code sequences of length N

For binary sequences, the block code maps all the possible $M = 2^L$ binary sequences onto sequences $\underline{x}^N$

The rate of a block code is $R = \frac{\log M}{N}$

Let τ_c be the duration of an output symbol y from $\mathcal{Y}$, the data rate in bits is $\frac{R}{\tau_c}$

For any positive integer and $R > 0$, a (N, R) block code is a code of length N which has $\lceil 2^{NR} \rceil$ codewords

Upper bound on probability

Recall that for the weak coding theorem we performed a typicality-based decoding

That decoding led to a WLLN type of argument, which was the source of the poor handle we have on the behavior of error probability with N

Let us then consider another criterion for decoding: maximum likelihood (ML)

select m for which probability of receiving $\underline{y}^N$ is maximum

$$P_{\underline{Y}^N,\underline{X}^N}\left(\underline{y}^N|\underline{x}^N(m)\right) \geq P_{\underline{Y}^N,\underline{X}^N}\left(\underline{y}^N|\underline{x}^N(m')\right)$$

$\forall m' \neq m$

Let $\mathcal{Y}_m$ be the set of output vectors $\underline{y}^N$ whose decoding is the message m

The probability of error when the message m was transmitted is:

$$P_{e,m} = \sum_{\underline{y}^N \in \mathcal{Y}_m^C} P_{\underline{Y}^N|\underline{X}^N}\left(\underline{y}^N|\underline{x}^N(m)\right)$$

Upper bound on probability

Theorem:

The average probability of decoding error given that the message m was sent, averaged over all possible block codes, is bounded, for any choice of $\rho \in [0,1]$ by

$$E_{codebooks}[P_{e,m}] \leq (M-1)^{\rho}$$
$$\sum_{\underline{y}^N} \left[\sum_{\underline{x}^N} P_{\underline{X}^N}\left(\underline{x}^N\right) P_{\underline{Y}^N|\underline{X}^N}\left(\underline{y}^N|\underline{x}^N\right)^{\frac{1}{1+\rho}} \right]^{1+\rho}$$

Proof:

The probability of error given that m was transmitted averaged over all possible codes is:

$$E_{codebooks}[P_{e,m}] = \sum_{\underline{x}^N(m)} \sum_{\underline{y}^N}$$
$$P_{\underline{X}^N}\left(\underline{x}^N(m)\right) P_{\underline{Y}^N|\underline{X}^N}\left(\underline{y}^N|\underline{x}^N(m)\right)$$
$$Pr[error|m, \underline{X}^N = \underline{x}^N(m), \underline{Y}^N = \underline{y}^N]$$

Upper bound on probability

Proof continued

For a given $m, \underline{x}^N(m), \underline{y}^N$, let $A\left(m', \underline{x}^N(m), \underline{y}^N\right)$ be the event that

$$P_{\underline{Y}^N|\underline{X}^N}\left(\underline{y}^N|\underline{x}^N(m)\right) \leq P_{\underline{Y}^N|\underline{X}^N}\left(\underline{y}^N|\underline{x}^N(m')\right)$$

an error occurs when at least one of the events $A\left(m', \underline{x}^N(m), \underline{y}^N\right)$, $m \neq m'$ takes place

therefore

$$\begin{aligned} & Pr[error|m, \underline{X}^N = \underline{x}^N(m), \underline{Y}^N = \underline{y}^N] \\ = \; & Pr\left(\bigcup_{m \neq m'} A\left(m', \underline{x}^N(m), \underline{y}^N\right)\right) \\ \leq \; & \left[\sum_{m \neq m'} Pr\left(A\left(m', \underline{x}^N(m), \underline{y}^N\right)\right)\right]^\rho \end{aligned}$$

Upper bound on probability

Proof continued

Why not just use the union bound

$$Pr\left(\bigcup_{m\neq m'} A\left(m', \underline{x}^N(m), \underline{y}^N\right)\right)$$
$$\leq \left[\sum_{m\neq m'} Pr\left(A\left(m', \underline{x}^N(m), \underline{y}^N\right)\right)\right]$$

if RHS is ≥ 1, then it remains so even after being raised to a power

if RHS if ≤ 1, then it increases when raised to a power in $[0, 1]$

Let us now compute

$Pr\left(A\left(m', \underline{x}^N(m), \underline{y}^N\right)\right)$

as a sum over the possible encodings of m'

Upper bound on probability

Proof continued

$$
\begin{aligned}
& Pr\left(A\left(m', \underline{x}^N(m), \underline{y}^N\right)\right) \\
= & \sum_{\underline{x}^N(m'): P_{\underline{Y}^N|\underline{X}^N}\left(\underline{y}^N|\underline{x}^N(m)\right) \leq P_{\underline{Y}^N|\underline{X}^N}\left(\underline{y}^N|\underline{x}^N(m')\right)} \\
& P_{\underline{X}^N}\left(\underline{x}^N(m')\right) \\
\leq & \sum_{\underline{x}^N(m')} P_{\underline{X}^N}\left(\underline{x}^N(m')\right) \\
& \frac{P_{\underline{Y}^N|\underline{X}^N}\left(\underline{y}^N|\underline{x}^N(m')\right)^r}{P_{\underline{Y}^N|\underline{X}^N}\left(\underline{y}^N|\underline{x}^N(m)\right)^r}
\end{aligned}
$$

for any $r > 0$

note that the last expression does not depend on m' because we sum over all the possible codes for m'

Upper bound on probability

Proof continued

Combining results, we obtain that

$$
\begin{aligned}
& Pr[error|m, \underline{X}^N = \underline{x}^N(m), \underline{Y}^N = \underline{y}^N] \\
\leq \; & \left[\sum_{m \neq m'} Pr\left(A\left(m', \underline{x}^N(m), \underline{y}^N \right) \right) \right]^{\rho} \\
\leq \; & \left[(M-1) \sum_{\underline{x}^N(m')} P_{\underline{X}^N}\left(\underline{x}^N(m') \right) \right. \\
& \left. \frac{P_{\underline{Y}^N|\underline{X}^N}\left(\underline{y}^N | \underline{x}^N(m') \right)^r}{P_{\underline{Y}^N|\underline{X}^N}\left(\underline{y}^N | \underline{x}^N(m) \right)^r} \right]^{\rho}
\end{aligned}
$$

Upper bound on probability

Proof continued

Averaging the error over all possible codes:

$$E_{codebooks}[P_{e,m}] \leq$$

$$(M-1)^{\rho} \sum_{\underline{y}^N} \left[\sum_{\underline{x}^N(m)} P_{\underline{X}^N}\left(\underline{x}^N(m)\right) \right.$$

$$\left. P_{\underline{Y}^N|\underline{X}^N}\left(\underline{y}^N|\underline{x}^N(m)\right)^{1-r\rho} \right]$$

$$\left[\sum_{\underline{x}^N(m')} P_{\underline{X}^N}\left(\underline{x}^N(m')\right) \right.$$

$$\left. P_{\underline{Y}^N|\underline{X}^N}\left(\underline{y}^N|\underline{x}^N(m')\right)^{r} \right]^{\rho}$$

Upper bound on probability

Proof continued

Picking $r = \frac{1}{1+\rho}$ implies $1 - r\rho = r$ so

Averaging the error over all possible codes:

$$
\begin{aligned}
&E_{codebooks}[P_{e,m}] \leq \\
&(M-1)^{\rho} \sum_{\underline{y}^N} \left[\sum_{\underline{x}^N} P_{\underline{X}^N}\left(\underline{x}^N\right) \right. \\
&\left. P_{\underline{Y}^N|\underline{X}^N}\left(\underline{y}^N|\underline{x}^N\right)^{\frac{1}{1+\rho}} \right]^{1+\rho}
\end{aligned}
$$

QED!

Upper bound on probability

Have we used the DMC nature of the channel? Only insofar as it provides block-by-block memorylessness. Let us now make greater use of the DMC assumption

We assume $P_{\underline{X}^n}(\underline{x}^n) = \prod_{i=1}^N P_X(x_i)$ so

$$\begin{aligned}
& E_{codebooks}[P_{e,m}] \leq \\
& (M-1)^\rho \sum_{y_1} \dots \sum_{y_N} \left[\sum_{x_1} \dots \sum_{x_N} \prod_{i=1}^N P_X(x_i) \right. \\
& \left. P_{Y|X}(y_i|x_i)^{\frac{1}{1+\rho}} \right]^{1+\rho} \\
= \; & (M-1)^\rho \sum_{y_1} \dots \sum_{y_N} \left[\prod_{i=1}^N \sum_x P_X(x) \right. \\
& \left. P_{Y|X}(y_i|x)^{\frac{1}{1+\rho}} \right]^{1+\rho} \\
= \; & (M-1)^\rho \prod_{i=1}^N \sum_y \left[\sum_x P_X(x) \right. \\
& \left. P_{Y|X}(y_i|x)^{\frac{1}{1+\rho}} \right]^{1+\rho} \\
= \; & (M-1)^\rho \{ \sum_y \left[\sum_{x_N} P_X(x) P_{Y|X}(y_i|x)^{\frac{1}{1+\rho}} \right]^{1+\rho} \}^N
\end{aligned}$$

Upper bound on probability

From our definition of M and R, $M - 1 \leq 2^{NR}$

Hence

$$E_{codebooks}[P_{e,m}] \leq 2^{-N(E_0(\rho,P_X(x)))-\rho R}$$

for

$$\begin{aligned} & E_0(\rho, P_X(x)) \\ = \; & -\log\left(\sum_y \left[\sum_{x_N} P_X(x) P_{Y|X}(y_i|x)^{\frac{1}{1+\rho}}\right]^{1+\rho}\right) \end{aligned}$$

This looks exponential, but we need to make sure that

$$E_0(\rho, P_X(x))) - \rho R > 0$$

Upper bound on probability

What we have done:

related the probability of error to some exponential function of the input and transition PMFs

What needs to be done:

- get rid of the expectation over codes by throwing out the worst half of the codes
- Show that the bound behaves well (exponent is $-N\alpha$ for some $\alpha > 0$)
- Relate the bound to capacity - this was immediate in the weak coding theorem because we were using the WLLN and therefore capacity was related to the sample mean, which we used to perform typical set decoding