LECTURE 17

Last time:

- Differential entropy
- Entropy rate and Burg's theorem
- AEP

Lecture outline

- White Gaussian noise
- Bandlimited WGN
- Additive White Gaussian Noise (AWGN) channel
- Capacity of AWGN channel
- Application: DS-CDMA systems
- Spreading
- Coding theorem

Reading: Sections 10.1-10.3.

White Gaussian Noise (WGN)

WGN is a good model for a variety of noise processes:

- ambient noise

- dark current

- noise from amplifiers

A sample at any time $N(t)$ is Gaussian

It is defined by its power spectrum: $R(t) = N_0\delta(t)$

Its power spectral density is N_0 over the whole spectrum (flat PSD)

What is its bandwidth?

What is its energy?

Do we ever actually have WGN?

Bandlimited WGN

WGN is always bandlimited to some bandwidth W, thus the noise is passed through a bandpass filter

The power density spectrum is now non-zero only over an interval W of frequency

The energy of the noise is then WN_0

The Nyquist rate is then W and samples are $\frac{1}{W}$ apart in time

Discrete-time samples of the bandlimited noise, which by abuse of notation we still denote N, are IID

$N_i \sim \mathcal{N}(0, \sigma_N^2)$

AWGN Channel

We operate in sampled time, hence already in a bandlimited world

We ignore issues of impossibility of bandlimiting and time-limiting simultaneously (one can obtain bounds for the limiting in time, and we are inherently considering very long times, as per the coding theorem)

$$Y_i = X_i + N_i$$

Nyquist rate of the output is Nyquist rate of the input

The input and the noise are independent

The input has a constraint on its average energy, or average variance

Capacity of AWGN Channel

We have a memoryless channel

Same arguments as for DMC imply the X_is should be IID, yielding IID Y_is

The variance of Y_i is $\sigma_X^2 + \sigma_N^2$

Hence we omit the i subscript in the following

$$
\begin{aligned}
& I(X;Y) = h(Y) - h(Y|X) \\
= \;& h(Y) - E_X[h(Y-x|X=x)] \\
= \;& h(Y) - E_X[h(N|X=x)] \\
= \;& h(Y) - h(N) \\
\leq \;& \frac{1}{2}\ln\left(2\pi e(\sigma_X^2+\sigma_N^2)\right) - \frac{1}{2}\ln\left(2\pi e\sigma_N^2\right)
\end{aligned}
$$

To achieve inequality with equality, select the X_is to be IID $\sim \mathcal{N}(0, \sigma_X^2)$ (could have any mean other than 0, but would affect second moment without changing variance, so not useful in this case), hence the X_is are themselves samples of bandlimited WGN

Capacity of AWGN Channel

Another view of capacity

$$\begin{aligned} I(X;Y) &= h(X) - h(X|Y) \\ &= h(X) - E_Y[h(X-\alpha y|Y=y)] \end{aligned}$$

In particular, we can select α so that αY is the LLSE estimate of X from Y

For X Gaussian, then this estimate is also the MMSE estimate of X from Y

The error of the estimate is $\widetilde{X} = X - \alpha y$

The error is independent of the value y (recall that for any jointly Gaussian random variables, the first rv can be expressed as the weighted sum of the second plus an independent Gaussian rv)

The error is Gaussian

Capacity of AWGN Channel

Note that this maximizes $E_Y[h(X-\alpha y|Y=y)]$

The $\widetilde{X}_i$s are still IID

$$\begin{aligned} & E_Y[h(X-\alpha y|Y=y)] \\ = \; & E_Y[h(\widetilde{X}|Y=y)] \\ \leq \; & h(\widetilde{X}) \\ \leq \; & \frac{1}{2}\ln\left(2\pi e\sigma^2_{\widetilde{X}}\right) \end{aligned}$$

We can clearly see that adding a constant to the input would not help

Note that, if the channel is acting like a jammer to the signal, then the jammer cannot control the $h(X)$ term, only can control the $h(X|Y)$ term

All of the above inequalities are reached with equality if the channel is an AWGN channel

Capacity of AWGN Channel

Thus, the channel is acting as an optimal jammer to the signal

Saddle point:

- if the channel acts as an AWGN channel, then the sender should send bandlimited WGN under a specific variance constraint (energy constraint)

- if the input is bandlimited WGN, then under a variance constraint on the noise, the channel will act as an AWGN channel

The capacity is thus

$$C = \frac{1}{2}\ln\left(\frac{\sigma_Y^2}{\sigma_N^2}\right) = \frac{1}{2}\ln\left(1 + \frac{\sigma_X^2}{\sigma_N^2}\right)$$

for small SNR, roughly proportional to SNR

Application: DS-CDMA systems

Consider a direct-sequence code division multiple access system

Set $\mathcal{U}$ of users sharing the bandwidth

Every user acts as a jammer to every other user

$X_i^j = \sum_{k \in \mathcal{U} \setminus \{j\}} X_i^k + N_i$

The capacity for user j is

$$C_j = \frac{1}{2} \ln \left(1 + \frac{\sigma^2_{X^j}}{\sum_{k \in \mathcal{U} \setminus \{j\}} \sigma^2_{X_i^k} + \sigma^2_N} \right)$$

Spreading

What happens when use more bandwidth (spreading)?

Let us revisit the expression for capacity:

$$C = \frac{1}{2} \ln \left(1 + \frac{\sigma_X^2}{\sigma_N^2}\right)$$

Problem: when we change the bandwidth, we also change the number of samples

Think in terms of number of degrees of freedom

For time T and bandwidth W

The total energy of the input stays the same, but the energy of the noise is proportional to W

Or, alternatively, per degree of freedom, we have the same energy per degree of freedom for the noise, but the energy for the input decreases proportionally to $\frac{1}{W}$

Spreading

For a given T with energy constraint $\mathcal{E}$ over that time in the input, maximum mutual information in terms of nats per second is:

$$\frac{1}{2T} W \ln\left(1 + \frac{\mathcal{E}}{WN_0}\right)$$

limit as $W \to \infty$ is $\frac{\mathcal{E}}{2TN_0}$

spreading is always beneficial, but its effect is bounded

Geometric interpretation in terms of concavity

Coding theorem

We no longer have a discrete input alphabet, but a continuous input alphabet

We now have transition pdfs $f_{Y|X}(y|x)$

For any block length n, let

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

$f_{\underline{X}^n}(\underline{x}^n) = \prod_{i=1}^n f_X(x_i)$

$f_{\underline{Y}^n}(\underline{y}^n) = \prod_{i=1}^n f_Y(y_i)$

Let R be an arbitrary rate $R < C$

For each n consider choosing a code of $M = \lfloor e^{nR} \rfloor$ codewords, where each codeword is chosen independently with pdf assignment $f_{\underline{X}^n}(\underline{x}^n)$

Coding theorem

Let $\epsilon = \frac{C-R}{2}$ and let the set T_ϵ^n be the set of pairs $(\underline{x}^n, \underline{y}^n)$ such that

$$|\frac{1}{n} i\left(\underline{x}^n; \underline{y}^n\right) - C| \leq \epsilon$$

where i is the sample natural mutual information

$$i\left(\underline{x}^n; \underline{y}^n\right) = \ln\left(\frac{f_{\underline{Y}^n|\underline{X}^n}(\underline{y}^n|\underline{x}^n)}{f_{\underline{Y}^n}(\underline{y}^n)}\right)$$

For every n and each code in the ensemble, the decoder, given $\underline{y}^n$, selects the message m for which $\left(\underline{x}^n(m), \underline{y}^n\right) \in T_\epsilon^n$.

We assume an error if there are no such codewords or more than one codeword.

Upper bound on probability

Let λ_m be the event that, given message m enters the system, an error occurs.

The mean probability of error over all ensemble of codes is

$$E[\lambda_m] = P(\lambda_m = 1)$$

(indicator function)

Error occurs when

$$\left(\underline{x}^n(m), \underline{y}^n\right) \notin T_\epsilon^n$$

or

$$\left(\underline{x}^n(m'), \underline{y}^n\right) \in T_\epsilon^n \text{ for } m' \neq m$$

Upper bound on probability

Hence, through the union bound

$$\begin{aligned} E[\lambda_m] &= P\left(\left(\left(\underline{x}^n(m), \underline{y}^n\right) \notin T_\epsilon^n\right) \right. \\ &\qquad \left. \cup \bigcup_{m' \neq m} \left(\underline{x}^n(m'), \underline{y}^n\right) \in T_\epsilon^n | m \right) \\ &\leq P\left(\left(\underline{x}^n(m), \underline{y}^n\right) \notin T_\epsilon^n\right) \\ &+ \sum_{m' \neq m} P\left(\left(\underline{x}^n(m'), \underline{y}^n\right) \in T_\epsilon^n | m\right) \end{aligned}$$

The probability of the pair not being typical approaches 0 as $n \to \infty$

$$\frac{i\left(\underline{x}^n(m); \underline{y}^n\right)}{n} = \frac{1}{n} \sum_{i=1}^{n} \ln\left(\frac{f_{Y|X}(y_i|x_i)}{f_Y(y_i)}\right)$$

Through the WLLN, the above converges in probability to

$$C = \int_{x \in \mathcal{X}, y \in \mathcal{Y}} f_X(x) f_{Y|X}(y|x) \ln\left(\frac{f_{Y|X}(y|x)}{f_Y(y)}\right) dx dy$$

Hence, for any ϵ

$$\lim_{n \to \infty} P\left(\left(\underline{x}^n(m), \underline{y}^n\right) \in T_\epsilon^n | m\right) \to 0$$

Upper bound to probability of several typical sequences

Consider $m' \neq m$

Given how the codebook was chosen, the variables $\underline{X}^n(m'), \underline{Y}^n$ are independent *conditioned on m having been transmitted*

Hence

$$P\left(\underline{X}^n(m'), \underline{Y}^n \in T_\epsilon^n | m\right)$$

$$= \int_{\left(\underline{x}^n(m'), \underline{y}^n\right) \in T_\epsilon^n} f_{\underline{X}^n}(\underline{x}^n(m')) f_{\underline{Y}^n}(\underline{y}^n) d\underline{x}^n(m') d\underline{y}^n$$

Because of the definition of T_ϵ^n, for all pairs in the set

$$f_{\underline{Y}^n}(\underline{y}^n) \leq f_{\underline{Y}^n|\underline{X}^n}\left(\underline{y}^n | \underline{x}^n(m')\right) e^{-n(C-\epsilon)}$$

Upper bound to probability of several typical sequences

$$
\begin{aligned}
& P\left(\underline{X}^n(m'), \underline{Y}^n \in T_\epsilon^n | m\right) \\
\leq \; & \int_{\left(\underline{x}^n(m'), \underline{y}^n\right) \in T_\epsilon^n} f_{\underline{Y}^n|\underline{X}^n}\left(\underline{y}^n | \underline{x}^n(m')\right) e^{-n(C-\epsilon)} \\
& f_{\underline{X}^n}(\underline{x}^n(m')) d\underline{x}^n(m') d\underline{y}^n \\
\leq \; & e^{-n(C-\epsilon)}
\end{aligned}
$$

$E[\lambda_m]$ is upper bounded by two terms that go to 0 as $n \to \infty$, thus the average probability of error given that m was transmitted goes to 0 as $n \to \infty$

This is the average probability of error averaged over the ensemble of codes, therefore $\forall \delta > 0$ and for any rate $R < C$ there must exist a code length n with average probability or error less than δ

Thus, we can create a sequence of codes with maximal probability of error converging to 0 as $n \to \infty$

We can expurgate codebooks as before

6.441 Information Theory
Spring 2010