Information Theory with Applications, Math6397
Lecture Notes from November 06, 2014

taken by Carlos Ortiz

Last Time

- Channel coding for continuous channels
- Additive Gaussian white noise (AGWN)

Capacity of AGWN, continued

Recall that for a AGWN channel with $E[X^2] \leq S$,

$$C(S) = \frac{1}{2}\ln(1 + \frac{S}{\sigma^2})$$

Now it turns out that this channel is the worst among all channels with $E[N_j] = 0$ and $E[N_j^2] = \sigma^2$.

5.6.29 Theorem. *If γ is an additive memoryless channel with noise $\{N_j\}_{j=1}^{\infty}$ of zero mean and variance σ^2 then,*

$$\frac{1}{2}\ln(1 + \frac{S}{\sigma^2}) \leq C(S)$$

Proof. Let $\widetilde{\gamma}$ be a AWGN channel with $\widetilde{N}_j$ having mean zero and variance σ^2. Also, consider a Gaussian sequence $\{\widetilde{X}_j\}_{j=1}^{\infty}$ with $E[\widetilde{X}_j^2] = S$. Now let us compare the mutual information between the Gaussian and the (possibly) non-Gaussian Channel.

$$I(\widetilde{X}; \gamma(\widetilde{X})) - I(\widetilde{X}; \widetilde{\gamma}(\widetilde{X}))$$

$$= \int\int p_{\widetilde{X}}(x) p_N(y-x) \ln(\frac{p_N(y-x)}{p_Y(y)}) dydx - \int\int p_{\widetilde{X}}(x) p_{\widetilde{N}}(y-x) \ln(\frac{p_{\widetilde{N}}(y-x)}{p_{\widetilde{Y}}(y)}) dydx$$

By equality of second moment we can replace $\widetilde{N}$ with N and get,

$$= \int\int p_{\widetilde{X}}(x) p_N(y-x) \ln(\frac{p_N(y-x)}{p_Y(y)}) dydx - \int\int p_{\widetilde{X}}(x) p_N(y-x) \ln(\frac{p_N(y-x)}{p_{\widetilde{Y}}(y)}) dydx$$

$$= \int\int p_{\widetilde{X}}(x)p_N(y-x)\ln(\frac{p_N(y-x)p_{\widetilde{Y}}(y)}{p_Y(y)p_{\widetilde{N}}(y-x)})dydx \geq \int\int p_{\widetilde{X}}(x)p_N(y-x)(1-\frac{p_Y(y)p_{\widetilde{N}}(y-x)}{p_N(y-x)p_{\widetilde{Y}}(y)})dydx$$

$$= 1 - \int\int \frac{p_{\widetilde{X}}(x)p_{\widetilde{N}}(y-x)p_Y(y)}{p_{\widetilde{Y}}(y)} = 1 - \int \frac{p_Y(y)}{p_{\widetilde{Y}}(y)}[\int p_{\widetilde{X}}(x)p_{\widetilde{N}}(y-x)dx]dy$$

$$= 1 - \int \frac{p_Y(y)}{p_{\widetilde{Y}}(y)}[p_{\widetilde{X}+\widetilde{N}=\widetilde{Y}}(y)]dy = 1 - \int p_Y(y)dy = 0$$

and equality holds if and only if

$$\frac{p_Y(y)}{p_{\widetilde{Y}}(y)} = \frac{p_N(y-x)}{p_{\widetilde{N}}(y-x)}$$

Now if we pick a y such that this equality holds, then the left hand side is fixed and the right hand side is a constant for almost every x, so $p_N(y-x) = c \cdot p_{\widetilde{N}}(y-x)$ and normalization forces $c = 1$. This shows equality holds if and only if $\{N_j\}_{j=1}^{\infty}$ is Gaussian. Next lets compare the Gaussian and the (possibly) non-Gaussian Capacity,

$$\max_{\substack{P_X \\ E[X^2]\leq S}} I(X;\widetilde{\gamma}(\widetilde{X})) = I(\widetilde{X};\widetilde{\gamma}(\widetilde{X})) \leq I(\widetilde{X};\gamma(\widetilde{X})) \leq \max_{\substack{P_X \\ E[X^2]\leq S}} I(X;\gamma(X)) = C(S)$$

□

5.7 Partially Noisy Channel

Suppose you have $Y_j = X_j + N_j$ where

$$N_j = \begin{cases} \widetilde{N}_j & \text{with probability 0.1} \\ 0 & \text{with probability 0.9} \end{cases}$$

where $\widetilde{N}_j$ are Gaussian with mean zero and variance σ^2.
What is the Capacity of this Channel?
We have that

$$p_N(x) = 0.1\delta(x) + \frac{0.9}{\sqrt{2\pi\sigma^2}}e^{-\frac{1}{2\sigma^2}x^2}$$

hence,

$$h(N) = \lim_{\Delta\to 0}(H(N^{\Delta}) + \ln(\Delta)) = -\infty$$

Now if we choose X so that $h(Y)$ is finite we get,

$$C(S) \geq h(Y) - h(N) = \infty$$

How is this possible?
Pick X uniformly distributed between $-\sqrt{S}$ and $\sqrt{S}$ and transmit rational numbers repeatedly. If the receiver gets a rational number as an output, the $\widetilde{N}_j = 0$ almost surely. As a result, using the channel n times gives decoding error probability of $(0.9)^n \overset{n\to\infty}{\longrightarrow} 0$.

5.8 Capacity for parallel AGWN channels

5.8.30 Theorem. *Suppose we have k Channels with Gaussian white noisy variables having variance $\sigma_1, ..., \sigma_k$ and an overall power constraint $\sum\limits_{i=1}^{k} E[X_i^2] \leq S$, then*

$$C(S) = \sum_{i=1}^{k} \frac{1}{2} \ln(1 + \frac{S_i}{\sigma_i^2})$$

where $S_i = \max\{0, \theta - \sigma_i^2\}$ and θ is chosen such that $\sum\limits_{i=1}^{k} S_i = S$. The process of choosing this θ if often know as "water filling algorithm".

Proof. By definition,

$$\max_{\substack{P_{X^{\otimes k}} \\ \sum_{i=1}^{k} E[X_i^2] \leq S}} I(X;Y)$$

Now since noise is independent of X,

$$I(X,Y) = h(Y) - h(Y|X) = h(Y) - h(X + N|X) = h(Y) - h(N|X)$$

$$= h(Y) - h(N) \leq \sum_{i=1}^{k} h(Y_i) - h(N_i) = \sum_{i=1}^{k} I(Y_i; N_i) \leq \sum_{i=1}^{k} \frac{1}{2} \ln(1 + \frac{S_i}{\sigma_i^2})$$

Now if we maximize the right hand side subject to $\sum\limits_{i=1}^{k} S_i = S$ we get our desired result.

To this end, we note that the sum of the logarithms is concave in $\{(S_i + \sigma_i^2)/\sigma_i^2\}$, thus averaging among $S_i + \sigma_i^2$ for indices with $S_i > 0$ increases the right hand side. For a given θ, then $S_i + \sigma_i^2 = \theta$ when $S_i > 0$ achieves the maximum. Using the monotonicity of the logarithm, we can choose θ so that $\sum_{i=1}^{k} S_i = S$. □