Shannon's Noisy Channel Coding Theorem

Version 2.10: May 19, 2004 11:33 am GMT-5

Behnaam Aazhang

Abstract

A statement of Shannon's noisy channel coding theorem.

Shannon's Noisy Channel Coding Theorem

It is highly recommended that the information presented in Mutual Information[1] and in Typical Sequences[2] be reviewed before proceeding with this document. An introductory module on the theorem is available at Noisy Channel Theorems [3].

Theorem 1: Shannon's Noisy Channel Coding

The capacity of a discrete-memoryless channel is given by

$$C = \max_{p_X(x)} \{\mathscr{I}(X;Y) | p_X(x)\} \tag{1}$$

where $\mathscr{I}(X;Y)$ is the mutual information between the channel input X and the output Y. If the transmission rate R is less than C, then for any $\epsilon > 0$ there exists a code with block length n large enough whose error probability is less than ϵ. If $R > C$, the error probability of any code with any block length is bounded away from zero.

Example 1:

If we have a binary symmetric channel with cross over probability 0.1, then the capacity $C \approx 0.5$ bits per transmission. Therefore, it is possible to send 0.4 bits per channel through the channel reliably. This means that we can take 400 information bits and map them into a code of length 1000 bits. Then the whole code can be transmitted over the channels. One hundred of those bits may be detected incorrectly but the 400 information bits may be decoded correctly.

[1] http://cnx.rice.edu/content/m10178/latest/
[2] http://cnx.rice.edu/content/m10179/latest/
[3] http://cnx.rice.edu/content/m0073/latest/

Before we consider continuous-time additive white Gaussian channels, let's concentrate on discrete-time Gaussian channels

$$Y_i = X_i + \eta_i \tag{2}$$

where the X_i's are information bearing random variables and η_i is a Gaussian random variable with variance σ_η^2. The input X_i's are constrained to have power less than P

$$\frac{1}{n}\sum_{i=1}^{n}\left(X_i^{\ 2}\right) \leq P \tag{3}$$

Consider an output block of size n

$$\mathbf{Y} = \mathbf{X} + \eta \tag{4}$$

For large n, by the Law of Large Numbers,

$$\frac{1}{n}\sum_{i=1}^{n}\left(\eta_i^{\ 2}\right) = \frac{1}{n}\sum_{i=1}^{n}\left(\left(|y_i - x_i|\right)^2\right) \leq {\sigma_\eta}^2 \tag{5}$$

This indicates that with large probability as n approaches infinity, $\mathbf{Y}$ will be located in an n-dimensional sphere of radius $\sqrt{n{\sigma_\eta}^2}$ centered about $\mathbf{X}$ since $\left(|\mathbf{y} - \mathbf{x}|\right)^2 \leq n{\sigma_\eta}^2$

On the other hand since X_i's are power constrained and η_i and X_i's are independent

$$\frac{1}{n}\sum_{i=1}^{n}\left(y_i^{\ 2}\right) \leq P + {\sigma_\eta}^2 \tag{6}$$

$$|\mathbf{Y}| \leq n\left(P + {\sigma_\eta}^2\right) \tag{7}$$

This mean $\mathbf{Y}$ is in a sphere of radius $\sqrt{n\left(P + {\sigma_\eta}^2\right)}$ centered around the origin.

How many $\mathbf{X}$'s can we transmit to have nonoverlapping $\mathbf{Y}$ spheres in the output domain? The question is how many spheres of radius $\sqrt{n{\sigma_\eta}^2}$ fit in a sphere of radius $\sqrt{n\left(P + {\sigma_\eta}^2\right)}$.

$$\begin{aligned} M &= \frac{\left(\sqrt{n\left({\sigma_\eta}^2 + P\right)}\right)^n}{\left(\sqrt{n{\sigma_\eta}^2}\right)^n} \\ &= \left(1 + \frac{P}{{\sigma_\eta}^2}\right)^{\frac{n}{2}} \end{aligned} \tag{8}$$

Exercise 1:

How many bits of information can one send in n uses of the channel?

Solution:

$$\log_2\left(\left(1 + \frac{P}{{\sigma_\eta}^2}\right)^{\frac{n}{2}}\right) \tag{9}$$

The capacity of a discrete-time Gaussian channel $C = \frac{1}{2}\log_2\left(1 + \frac{P}{{\sigma_\eta}^2}\right)$ bits per channel use.

When the channel is a continuous-time, bandlimited, additive white Gaussian with noise power spectral density $\frac{N_0}{2}$ and input power constraint P and bandwidth W. The system can be sampled at the Nyquist rate to provide power per sample P and noise power

$$\begin{aligned} {\sigma_\eta}^2 &= \int_{-W}^{W}\frac{N_0}{2}\,df \\ &= WN_0 \end{aligned} \tag{10}$$

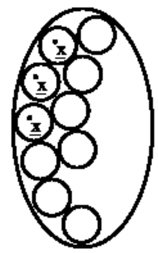

Figure 1

The channel capacity $\frac{1}{2}\log_2\left(1+\frac{P}{N_0W}\right)$ bits per transmission. Since the sampling rate is $2W$, then

$$C=\frac{2W}{2}\log_2\left(1+\frac{P}{N_0W}\right)\text{ bits/trans. x trans./sec} \quad (11)$$

$$C=W\log_2\left(1+\frac{P}{N_0W}\right)\frac{\text{bits}}{\text{sec}} \quad (12)$$

Example 2:

The capacity of the voice band of a telephone channel can be determined using the Gaussian model. The bandwidth is 3000 Hz and the signal to noise ratio is often 30 dB. Therefore,

$$C=3000\log_2\left(1+1000\right)\approx 30000\frac{\text{bits}}{\text{sec}} \quad (13)$$

One should not expect to design modems faster than 30 Kbs using this model of telephone channels. It is also interesting to note that since the signal to noise ratio is large, we are expecting to transmit 10 bits/second/Hertz across telephone channels.