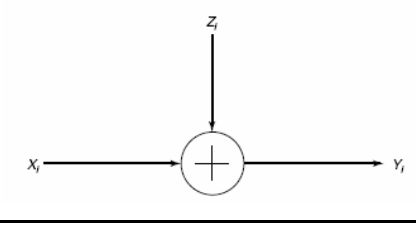
Gaussian Channel

Introduction

The most important continuous alphabet channel is the Gaussian channel depicted in Figure. This is a time-discrete channel with output Y_i at time i, where Y_i is the sum of the input X_i and the noise Z_i. The noise Z_i is drawn i.i.d. from a Gaussian distribution with variance N. Thus,

$$Y_i = X_i + Z_i,\ Z_i \sim N(0,N).$$

The noise Z_i is assumed to be independent of the signal X_i . The continuos alphabet is due to the presence of Z, that is a continuous random variable.



Gaussian Channel

This channel is a model for some common communication channels, such as wired and wireless telephone channels and satellite links.

Without further conditions, the capacity of this channel may be infinite. If the noise variance is zero, the receiver receives the transmitted symbol perfectly. Since X can take on any real value, the channel can transmit an arbitrary real number with no error.

If the noise variance is nonzero and there is no constraint on the input, we can choose an infinite subset of inputs arbitrarily far apart, so that they are distinguishable at the output with arbitrarily small probability of error. Such a scheme has an infinite capacity as well. Thus if the noise variance is zero or the input is unconstrained, the capacity of the channel is infinite.

Power Limitation

The most common limitation on the input is an energy or power constraint.We assume an average power constraint. For any codeword $(x_1, x_2, \dots, x_n)$ transmitted over the channel, we require that:

$$\frac{1}{n}\sum_{i=1}^{n} x_i^2 \le P$$

The additive noise in such channels may be due to a variety of causes. However, by the central limit theorem, the cumulative effect of a large number of small random effects will be approximately normal, so the Gaussian assumption is valid in a large number of situations.

Usage of the Channel

We first analyze a simple suboptimal way to use this channel. Assume that we want to send 1 bit over the channel.

Given the power constraint, the best that we can do is to send one of two levels, $+\sqrt{P}$ or $-\sqrt{P}$. The receiver looks at the corresponding Y received and tries to decide which of the two levels was sent.

Assuming that both levels are equally likely (this would be the case if we wish to send exactly 1 bit of information), the optimum decoding rule is to decide that $+\sqrt{P}$ was sent if $Y > 0$ and decide $-\sqrt{P}$ was sent if $Y < 0$.

Probability of Error

The probability of error with such a decoding scheme can be computed as follows:

$$
\begin{aligned}
P_e &= \frac{1}{2}\Pr(Y < 0 \mid X = +\sqrt{P}) + \frac{1}{2}\Pr(Y > 0 \mid X = -\sqrt{P}) \\
&= \frac{1}{2}\Pr(Z < -\sqrt{P} \mid X = +\sqrt{P}) + \frac{1}{2}\Pr(Z > \sqrt{P} \mid X = -\sqrt{P}) \\
&= \Pr(Z > \sqrt{P}) \\
&= 1 - \Phi(\sqrt{P/N})
\end{aligned}
$$

Where $\Phi(x)$ is the cumulative normal function

$$\Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{\frac{-t^2}{2}}\, dt$$

Using such a scheme, we have converted the Gaussian channel into a discrete binary symmetric channel with crossover probability P_e.

Similarly, by using a four-level input signal, we can convert the Gaussian channel into a discrete four input channel.

In some practical modulation schemes, similar ideas are used to convert the continuous channel into a discrete channel. The main advantage of a discrete channel is ease of processing of the output signal for error correction, but some information is lost in the quantization.

Definitions

We now define the (information) capacity of the channel as the maximum of the mutual information between the input and output over all distributions on the input that satisfy the power constraint.

Definition The *information capacity* of the Gaussian channel with power constraint P is

$$C = \max_{f(x):EX^2 \le P} I(X;Y)$$

We can calculate the information capacity as follows: Expanding $I(X; Y)$, we have:

$I(X; Y) = h(Y) - h(Y|X)$
$= h(Y) - h(X + Z|X)$
$= h(Y) - h(Z|X)$
$= h(Y) - h(Z)$

since Z is independent of X.

Now, h(Z) = 1/2 log 2πeN, and: $EY^2 = E(X + Z)^2 = EX^2 + 2EXEZ + EZ^2 = P + N$, since X and Z are independent and $EZ = 0$.

Given $EY^2 = P + N$, the entropy of Y is bounded by 12 log $2\pi e(P + N)$ because the normal maximizes the entropy for a given variance.

Information Capacity

Applying this result to bound the mutual information, we obtain:

$$I(X;Y) = h(Y) - h(Z)$$
$$\leq \frac{1}{2}\log 2\pi e(P+N) - \frac{1}{2}\log 2\pi eN$$
$$= \frac{1}{2}\log\left(1+\frac{P}{N}\right)$$

Hence, the information capacity of the Gaussian channel is:

$$C = \max_{f(x):EX^2 \leq P} I(X;Y) = \frac{1}{2}\log\left(1+\frac{P}{N}\right)$$

and the maximum is attained when $X \sim N(0, P)$.

We will now show that this capacity is also the supremum of the rates achievable for the channel.

(M,n) Code for the Gaussian Channel

Definition An *(M, n)* code for the Gaussian channel with power constraint P consists of the following:

1. An index set $\{1, 2, \dots, M\}$.
2. An encoding function $x : \{1, 2, \dots, M\} \rightarrow \chi^n$, yielding codewords $x^n(1), x^n(2), \dots, x^n(M)$, satisfying the power constraint P; that is, for every codeword:

$w = 1, 2, \dots, M$.
$$\sum_{i=1}^{n} x_i^2(w) \leq nP$$

3. A decoding function $g : Y^n \rightarrow \{1, 2, \dots, M\}$.

The rate and probability of error of the code are defined as for the discrete case.

Rate for a Gaussian Channel

Definition A rate R is said to be *achievable* for a Gaussian channel with a power constraint P if there exists a sequence of $(2^{nR}, n)$ codes with codewords satisfying the power constraint such that the maximal probability of error $\lambda(n)$ tends to zero.

The capacity of the channel is the supremum of the achievable rates.

Capacity of the Gaussian Channel

Theorem The capacity of a Gaussian channel with power constraint P and noise variance N is

$$C = \frac{1}{2}\log\left(1 + \frac{P}{N}\right)$$

bits per transmission.

Capacity of the Gaussian Channel

We present a plausibility argument as to why we may be able to construct $(2^{nC}, n)$ codes with a low probability of error.

Consider any codeword of length n. The received vector is normally distributed with mean equal to the true codeword and variance equal to the noise variance.

With high probability, the received vector is contained in a sphere of radius $\sqrt{n(N+\varepsilon)}$ around the true codeword.

If we assign everything within this sphere to the given codeword, when this codeword is sent there will be an error only if the received vector falls outside the sphere, which has low probability.

Capacity of the Gaussian Channel

Similarly, we can choose other codewords and their corresponding decoding spheres.

How many such codewords can we choose? The volume of an n-dimensional sphere is of the form $C_n r^n$ where r is the radius of the sphere. In this case, each decoding sphere has radius $\sqrt{nN}$.

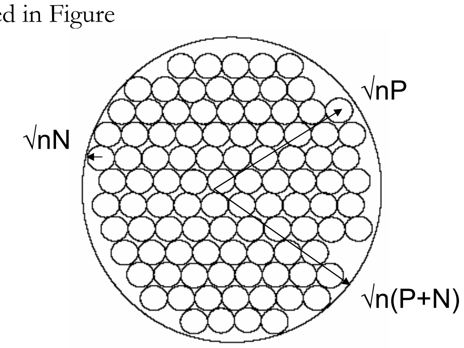
These spheres are scattered throughout the space of received vectors. The received vectors have energy no greater than $n(P + N)$, so they lie in a sphere of radius $\sqrt{n(P + N)}$. The maximum number of nonintersecting decoding spheres in this volume is no more than

$$\frac{C_n(n(P+N))^{n/2}}{C_n(nN)^{n/2}} = 2^{n/2}\log\left(1+\frac{P}{N}\right)$$

Capacity of the Gaussian Channel

Thus, the rate rate of the code is $1/2 \log(1 + P/N)$.

This idea is illustrated in Figure



This sphere-packing argument indicates that we cannot hope to send at rates greater than C with low probability of error. However, we can actually do almost as well as this

Converse to the Coding Theorem for Gaussian Channels

The capacity of a Gaussian channel is $C = 1/2 \log(1 + P/N)$. In fact, rates $R > C$ are not achievable.

The proof parallels the proof for the discrete channel. The main new ingredient is the power constraint.

Bandlimited Channels

A common model for communication over a radio network or a telephone line is a bandlimited channel with white noise. This is a continuous-time channel. The output of such a channel can be described as the convolution:

$Y(t) = (X(t) + Z(t)) * h(t),$

where $X(t)$ is the signal waveform, $Z(t)$ is the waveform of the white Gaussian noise, and $h(t)$ is the impulse response of an ideal bandpass filter, which cuts out all frequencies greater than W.

Bandlimited Channels

We begin with a representation theorem due to Nyquist and Shannon which shows that sampling a bandlimited signal at a sampling rate $1/2W$ is sufficient to reconstruct the signal from the samples.

Intuitively, this is due to the fact that if a signal is bandlimited to W, it cannot change by a substantial amount in a time less than half a cycle of the maximum frequency in the signal, that is, the signal cannot change very much in time intervals less than $1/2W$ seconds.

Nyquist Theorem

Theorem Suppose that a function f (t) is bandlimited to W, namely, the spectrum of the function is 0 for all frequencies greater than W. Then the function is completely determined by samples of the function spaced 1/2W seconds apart.

Capacity of Bandlimited Channels

A general function has an infinite number of degrees of freedom—the value of the function at every point can be chosen independently.

The Nyquist–Shannon sampling theorem shows that a bandlimited function has only $2W$ degrees of freedom per second.

The values of the function at the sample points can be chosen independently, and this specifies the entire function.

If a function is bandlimited, it cannot be limited in time. But we can consider functions that have most of their energy in bandwidth W and have most of their energy in a finite time interval, say $(0, T)$.

Capacity of Bandlimited Channels

Now we return to the problem of communication over a bandlimited channel.

Assuming that the channel has bandwidth W, we can represent both the input and the output by samples taken $1/2W$ seconds apart.

Each of the input samples is corrupted by noise to produce the corresponding output sample. Since the noise is white and Gaussian, it can be shown that each noise sample is an independent, identically distributed Gaussian random variable.

Capacity of Bandlimited Channels

If the noise has power spectral density $N_0/2$ watts/hertz and bandwidth W hertz, the noise has power $N_0/2\ (2W) = N_0W$ and each of the $2WT$ noise samples in time T has variance $N_0WT/2WT = N_0/2$.

Looking at the input as a vector in the $2TW$-dimensional space, we see that the received signal is spherically normally distributed about this point with covariance $N_0/2\ (I)$.

Capacity of Bandlimited Channels

Now we can use the theory derived earlier for discrete-time Gaussian channels, where it was shown that the capacity of such a channel is

$$C = \frac{1}{2}\log\left(1 + \frac{P}{N}\right)$$

Let the channel be used over the time interval $[0, T]$. In this case, the energy per sample is $PT/2WT = P/2W$, the noise variance per sample is N0/2 $(2W)\ T/2WT = N_0/2$, and hence the capacity per sample is

$$C = \frac{1}{2}\log\left(1 + \frac{P/2W}{N_0/2}\right) = \frac{1}{2}\log\left(1 + \frac{P}{N_0W}\right)$$

bits per sample.

Capacity of Bandlimited Channels

Since there are $2W$ samples each second, the capacity of the channel can be rewritten as:

$$C = W \log\left(1 + \frac{P}{N_0 W}\right)$$

This equation is one of the most famous formulas of information theory. It gives the capacity of a bandlimited Gaussian channel with noise spectral density $N_0/2$ watts/Hz and power P watts.

If we let $W \rightarrow \infty$ we obtain:

$$C = \frac{P}{N_0} \log_2 e$$

as the capacity of a channel with an infinite bandwidth, power P, and noise spectral density $N0/2$. Thus, for infinite bandwidth channels, the capacity grows linearly with the power.

Example: Telephone Line

To allow multiplexing of many channels, telephone signals are bandlimited to 3300 Hz.

Using a bandwidth of 3300 Hz and a SNR (signal-to-noise ratio) of 33 dB (i.e., $P/N_0W = 2000$) we find the capacity of the telephone channel to be about 36,000 bits per second.

Practical modems achieve transmission rates up to 33,600 bits per second in both directions over a telephone channel. In real telephone channels, there are other factors, such as crosstalk, interference, echoes, and nonflat channels which must be compensated for to achieve this capacity.

Example: Telephone Line

The V.90 modems that achieve 56 kb/s over the telephone channel achieve this rate in only one direction, taking advantage of a purely digital channel from the server to final telephone switch in the network.

In this case, the only impairments are due to the digital-to-analog conversion at this switch and the noise in the copper link from the switch to the home.

These impairments reduce the maximum bit rate from the 64 kb/s for the digital signal in the network to the 56 kb/s in the best of telephone lines.

Example: Telephone Line

The actual bandwidth available on the copper wire that links a home to a telephone switch is on the order of a few megahertz; it depends on the length of the wire.

The frequency response is far from flat over this band. If the entire bandwidth is used, it is possible to send a few megabits per second through this channel.

Schemes such at DSL (Digital Subscriber Line) achieve this using special equipment at both ends of the telephone line (unlike modems, which do not require modification at the telephone switch).