Gaussian channel

Information theory 2013, lecture 6

Jens Sjölund

8 May 2013

Outline

Introduction

Definition

A Gaussian channel is a time-discrete channel with output Y_i, input X_i and noise Z_i at time i such that

$$Y_i = X_i + Z_i, \quad Z_i \sim \mathcal{N}(0, N)$$

where Z_i is i.i.d and independent of X_i.

- The Gaussian channel is the most important continuous alphabet channel, modeling a wide range of communication channels.
- The validity of this follows from the central limit theorem

The power constraint

- If the noise variance is zero or the input is unconstrained, the capacity of the channel is infinite.
- The most common limitation on the input is a power constraint

$$EX_i^2 \leq P$$

Information capacity

Theorem

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

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

Proof.

Use that $EY^2 = P + N \implies h(Y) \le \frac{1}{2}\log 2\pi e(P+N)$ (Theorem 8.6.5)

$$\begin{aligned} I(X;Y) &= h(Y) - h(Y|X) = h(Y) - h(X+Z|X) = h(Y) - h(Z|X) \\ &= h(Y) - h(Z) \le \frac{1}{2}\log 2\pi e(P+N) - \frac{1}{2}\log 2\pi e N \\ &= \frac{1}{2}\log\left(\frac{P+N}{N}\right) = \frac{1}{2}\log\left(1 + \frac{P}{N}\right) \end{aligned}$$

which holds with equality iff $X \sim \mathcal{N}(0, P)$ □

Outline

Statement of the theorem

Theorem (9.1.1)

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

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

Proof.

See the book. Based on random codes and joint typicality decoding (similar to the proof for the discrete case). □

Handwaving motivation

Volume of n-dimensional sphere $\propto r^n$

Radius of received vectors $\sqrt{\sum_{i=1}^{n} Y_i^2} \leq \sqrt{n(P+N)}$

Using decoding spheres of radius $\sqrt{nN}$ the maximum number of nonintersecting decoding spheres is therefore no more than

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

(note the typo in eq. 9.22 in the book). This corresponds to a rate

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

Outline

Bandlimited channels

A bandlimited channel with white noise can be described as

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

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 all frequencies greater than W.

Theorem (Nyquist-Shannon sampling theorem)

Suppose that a function $f(t)$ is bandlimited to W. Then, the function is completely determined by samples of the function spaced $\frac{1}{2W}$ apart.

Bandlimited channels

Sampling at $2W$ Hertz, the energy per sample is $P/2W$. Assuming a noise spectral density $N_0/2$ the capacity is

$$
\begin{aligned}
C &= \frac{1}{2}\log\left(1+\frac{P/2W}{N_0/2}\right) = \frac{1}{2}\log\left(1+\frac{P}{N_0 W}\right) \quad \text{bits per sample} \\
&= W\log\left(1+\frac{P}{N_0 W}\right) \quad \text{bits per second}
\end{aligned}
$$

which is arguably one of the most famous formulas of information theory.

Outline

Parallel Gaussian channels

Consider k independent Gaussian channels in parallel with a common power constraint

$$\begin{aligned} Y_j &= X_j + Z_j, \quad j = 1, 2, \ldots, k \\ Z_j &\sim \mathcal{N}(0, N_j) \end{aligned}$$

$$E \sum_{j=1}^{k} X_j^2 \leq P$$

The objective is to distribute the total power among channels so as to maximize the capacity.

Information capacity of parallel Gaussian channel

Following the same lines of reasoning as in the single channel case one may show

$$I(X_1,\ldots,X_k;Y_1,\ldots,Y_k) \le \sum_j \Big(h(Y_j) - h(Z_j)\Big) \le \sum_j \frac{1}{2}\log\left(1 + \frac{P_j}{N_j}\right)$$

where $P_j = EX_j^2$ and $\sum P_j = P$. Equality is achieved by

$$(X_1,\ldots,X_k) \sim \mathcal{N}\left(0, \begin{bmatrix} P_1 & 0 & \ldots & 0 \\ 0 & P_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & P_k \end{bmatrix}\right)$$

Maximizing the information capacity

The objective is thus to maximize the capacity subject to the constraint $\sum P_j = P$, which can be done a Lagrange multiplier λ

$$J(P_1, \ldots, P_k) = \sum_j \frac{1}{2} \log\left(1 + \frac{P_j}{N_j}\right) + \lambda\left(\sum_j P_j - P\right)$$

$$\frac{\partial J}{\partial P_i} = \frac{1}{2(P_i + N_i)} + \lambda = 0$$

$$P_i = -\frac{1}{2\lambda} - N_i \triangleq \nu - N_i$$

however $P_i \geq 0$, so the solution needs to modified to

$$P_i = (\nu - N_i)^+ = \begin{cases} \nu - N_i & \text{if } \nu \geq N_i \\ 0 & \text{if } \nu < N_i \end{cases}$$

Water-filling

The gives the solution

$$C = \sum_{j=1}^{k} \frac{1}{2} \log\left(1 + \frac{(\nu - N_j)^+}{N_j}\right)$$

where ν is chosen so that $\sum_{j=1}^{k} (\nu - N_j)^+ = P$ (typo in the summary).

As the signal power increases, power is alloted to the channels where the sum of noise and power is the lowest. By analogy to the way water distributes itself in a vessel, this is referred to as water-filling.

Outline

Channels with colored Gaussian noise

Consider parallel channels with dependent noise.

$$Y_i = X_i + Z_i, \quad Z^n \sim \mathcal{N}(0, K_Z)$$

This can also represent the case of a single channel with Gaussian noise with memory. For a channel with memory, a block of n consecutive uses can be considered as n channels in parallel with dependent noise.

Maximizing the information capacity (1/3)

Let K_X be the input covariance matrix. The power constraint can then be written

$$\frac{1}{n}\sum_i EX_i^2 = \frac{1}{n}\text{tr}(K_X) \leq P$$

As for the independent channels,

$$I(X_1,\ldots,X_k;Y_1,\ldots,Y_k) = h(Y_1,\ldots,Y_k) - h(Z_1,\ldots,Z_k)$$

The entropy of the output is maximal when Y is normally distributed, and since the input and noise are independent

$$K_Y = K_X + K_Z \implies h(Y_1,\ldots,Y_k) = \frac{1}{2}\log\left((2\pi e)^n |K_X + K_Z|\right)$$

Maximizing the information capacity (2/3)

Thus, maximizing the capacity amounts to maximizing $|K_X + K_Z|$ subject to the constraint $\text{tr}(K_X) \leq nP$. The eigenvalue decomposition of K_Z gives

$$K_Z = Q\Lambda Q^t, \quad \text{where } QQ^t = I$$

then,

$$|K_X + K_Z| = |Q||Q^t K_X Q + \Lambda||Q^t| = |Q^t K_X Q + \Lambda| \stackrel{\Delta}{=} |A + \Lambda|$$

and

$$\text{tr}(A) = \text{tr}(Q^t K_X Q) = \text{tr}(Q^t Q K_X) = \text{tr}(K_X) \leq nP$$

Maximizing the information capacity (3/3)

From Hadamard's inequality (Theorem 17.9.2)

$$|A + \Lambda| \leq \prod_i (A_{ii} + \lambda_i)$$

with equality iff A is diagonal. Since

$$\sum_i A_{ii} \leq nP, \quad A_{ii} \geq 0$$

it can be shown using the KKT conditions that the optimal solution is given by

$$A_{ii} = (\nu - \lambda_i)^+$$
$$\sum_i (\nu - \lambda_i)^+ = nP$$

Resulting information capacity

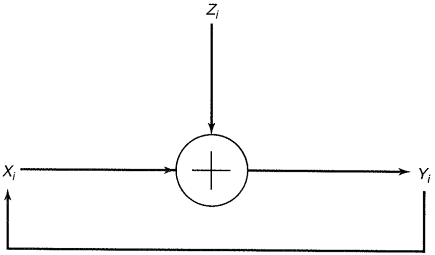
For a channel with colored Gaussian noise

$$C = \frac{1}{n} \sum_{i=1}^{n} \frac{1}{2} \log\left(1 + \frac{(\nu - \lambda_i)^+}{\lambda_i}\right)$$

where λ_i are the eigenvalues of K_Z and ν is chosen so that $\sum_{i=1}^{n} (\nu - \lambda_i)^+ = nP$ (typo in the summary).

Outline

Gaussian channels with feedback



$$Y_i = X_i + Z_i, \quad Z_i \sim \mathcal{N}(0, K_Z^{(n)})$$

Because of the feedback, X^n and Z^n are not independent; X_i is casually dependent on the past values of Z.

Capacity with and without feedback

For memoryless Gaussian channels, feedback does not increase capacity. However, for channels with memory, it does.

Capacity without feedback

$$C_n = \max_{tr(K_X)\leq nP} \frac{1}{2n} \log\left(\frac{|K_X + K_Z|}{|K_Z|}\right) = \frac{1}{2n} \sum_i \log\left(1 + \frac{(\lambda - \lambda_i^{(n)})^+}{\lambda_i^{(n)}}\right)$$

Capacity with feedback

$$C_{n,\text{FB}} = \max_{tr(K_X)\leq nP} \frac{1}{2n} \log\left(\frac{|K_{X+Z}|}{|K_Z|}\right)$$

Bounds on the capacity with feedback

Theorem

$$C_{n,FB} \leq C_n + \frac{1}{2}$$

Theorem (Pinsker's statement)

$$C_{n,FB} \leq 2C_n$$

In words, Gaussian channel capacity is not increased by more than half a bit or by more than a factor 2 when we have feedback.