Reliable Communication over noisy channel

$$W \longrightarrow \boxed{\text{enc}} \xrightarrow{X^n} \boxed{p(y|x)} \xrightarrow{Y^n} \boxed{\text{dec}} \longrightarrow \hat{W}$$

message

Discrete memoryless channel with input alphabet $\mathcal{X}$, output alphabet $\mathcal{Y}$, and transition probability assignment $p(y|x)$.

The channel capacity (informational) of DMC is

$$C = \max_{p(x)} I(X;Y).$$

if there is average cost constraint $E[b(X)] \le B$, then

capacity-cost function

$$C(B) = \max_{p(x):\, E[b(X)] \le B} I(X;Y).$$

There is a duality between rate-distortion and capacity-cost

Noisy Channel Coding thm

An (M,n) code for channel $(\mathcal{X}, p(y|x), \mathcal{Y})$ has

1. index set $\{1,2,\ldots,M\}$
2. encoding function $\{1,2,\ldots,M\} \mapsto \mathcal{X}^n$
3. decoding function $\cancel{\{1,2,\ldots,M\}}$ $g: \mathcal{Y}^n \mapsto \{1,2,\ldots,M\}$.

error probability. $\lambda_i = \Pr[g(Y^n) \neq i \mid X^n = X^n(i)]$

max error $\lambda^{(n)}$ for (M,n) code is $\lambda^{(n)} = \max_{i \in \{1,2,\ldots,M\}} \lambda_i$

A rate $R = \frac{\log M}{n}$ is achievable if there is a sequence of $(\lceil 2^{Rn} \rceil, n)$ codes such that $\lambda^{(n)} \to 0$ as $n \to \infty$.

The capacity is supremum of achievable rates.

Thm: All rates below capacity C are achievable. That is for every rate $R < C$, there is a sequence of $(2^{nR}, n)$ codes with $\lambda^{(n)} \to 0$. Conversely, any sequence of $(2^{nR}, n)$ codes with $\lambda^{(n)} \to 0$ must have $R \le C$.

Proof idea:

For each typical input n-sequence, there are approximately $2^{nH(Y|X)}$ possible Y sequences, all of them equally likely.

Want to make sure no two X sequences produce the same Y output sequence. Otherwise we'll have a lot of errors.

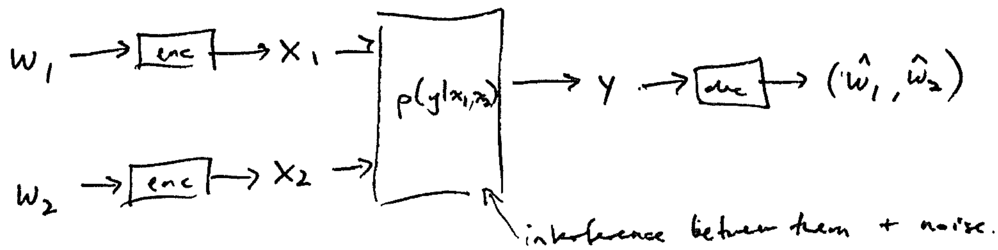
Divide $\approx 2^{nH(Y)}$ typical Y sequences into sets of size $\approx 2^{nH(Y|X)}$ corresponding to different input X sequences. Total number of disjoint sets is less than or equal to $2^{n(H(Y) - H(Y|X))} = 2^{nI(X;Y)}$

so can send at most $\approx 2^{nI(X;Y)}$ distinguishable sequences.

Converse is Fano's inequality-based.

Multiple Access Channel

A discrete multiple-access channel (memoryless) consists of three alphabets $\mathcal{X}_1, \mathcal{X}_2$, and $\mathcal{Y}$ and a probability transition matrix $p(y|x_1, x_2)$



def: An $((2^{nR_1}, 2^{nR_2}), n)$ code for a MAC consists of two message sets $\mathcal{W}_1 = \{1, 2, \ldots, 2^{nR_1}\}$ and $\mathcal{W}_2 = \{1, 2, \ldots, 2^{nR_2}\}$; two encoding functions

$$X_1 : \mathcal{W}_1 \to \mathcal{X}_1^n$$
$$X_2 : \mathcal{W}_2 \to \mathcal{X}_2^n$$

and decoder $g : \mathcal{Y}^n \mapsto \mathcal{W}_1 \times \mathcal{W}_2$.

average error probability

$$P_e^{(n)} = \frac{1}{2^{n(R_1+R_2)}} \sum_{(w_1,w_2) \in \mathcal{W}_1 \times \mathcal{W}_2} \Pr\left[g(Y^n) \neq (w_1, w_2) \mid (w_1, w_2) \text{ sent} \right].$$

def: A rate pair (R_1, R_2) is achievable for the MAC if there exist a sequence of $((2^{nR_1}, 2^{nR_2}), n)$ codes with $P_e^{(n)} \to 0$ as $n \to \infty$.

def: The capacity region of the MAC is closure of set of achievable (R_1, R_2) rate pairs.

MAC capacity

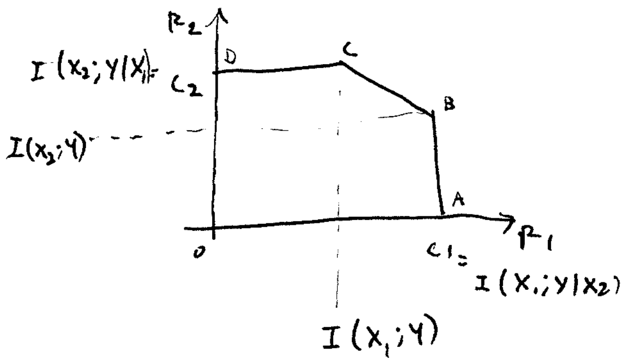
The capacity of a MAC $(\mathcal{X}_1 \times \mathcal{X}_2, p(y|x_1, x_2), \mathcal{Y})$ is closure of convex hull of all (R_1, R_2) satisfying

$$R_1 < I(X_1; Y | X_2)$$

$$R_2 < I(X_2; Y | X_1)$$

$$R_1 + R_2 < I(X_1, X_2; Y)$$

for some product distribution $p_1(x_1)\, p_2(x_2)$ on $\mathcal{X}_1 \times \mathcal{X}_2$.



B: High-level idea is to consider other signals as part of noise, decoding one signal, then "subtracting" it from the received signal.

A: maximum rate achievable from sender 1 when sender 2 not sending anything

Discrete-time AWGN channel

with input power constraint P

noise power N.

$$Y_i = X_i + Z_i \quad , i = 1, 2, \ldots$$

where Z_i are iid Gaussian r.v. with mean 0 and variance N.

signal $X = (X_1, \ldots, X_n)$ has power constraint

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

Capacity C is obtained by maximizing $I(X;Y)$ over all r.v. X

such that $E[X^2] \le P$ and is

$$C = \frac{1}{2} \log\left(1 + \frac{P}{N}\right) \quad \text{bits per transmission}$$

<u>Gaussian MAC</u>

Consider m transmitters, each with power P

Let $Y = \sum_{i=1}^{m} X_i + Z$

Let $C\left(\frac{P}{N}\right) = \frac{1}{2} \log\left(1 + \frac{P}{N}\right)$ be capacity of single-user Gaussian

channel with signal-to-noise ratio P/N.

The achievable rate region for Gaussian MAC is governed by following equations:

$$R_i < C\left(\frac{P}{N}\right)$$

$$R_i + R_j < C\left(\frac{2P}{N}\right)$$

$$R_i + R_j + R_k < C\left(\frac{3P}{N}\right)$$

$$\vdots$$

$$\sum_{i=1}^{m} R_i < C\left(\frac{mP}{N}\right)$$

note that when rates are same; last ineq dominates.

note that the sum of rates of users, $C\left(\frac{mP}{N}\right) \to \infty$ as $m \to \infty$

This is great for wireless communication:

the increasing interference as number of senders increases does not limit total received information

How does this help us in machine learning / AI applications

Is there something like interference that we want to deal with?

transfer / meta learning / etc.?