Multi-use Information Theory — December 21st, 2009

Lecture 10

Lecturer: Haim Permuter — *Scribe: Iddo Naiss*

I. Covering lemma

Lemma 1 (Covering lemma) Let X^n be i.i.d random vector $\sim P_X$, and let $\{y^n(i)\}_{i=1}^{2^{nR}}$ be a set of r.v y^n. For a joint distribution $P_{X,Y}$, if $R > I(X;Y)$ then there exists sets of $\{y^n(i)\}_{i=1}^{2^{nR}}$ s.t.-

$$\Pr\{\forall i (x^n, y^n(i) \notin T_\epsilon^n(X,Y))\} \to 0.$$

For the proof of the lemma, we use the following lemma:

Lemma 2 [1, CH 10] If $0 < y < 1$ then $(1-y)^n \leq e^{-yn}$

Proof: Let $f(y) = e^{-y} - 1 + y$. Then $f(0) = 0$ and $f'(y) = -e^{-y} + 1 > 0$ for $y > 0$, and hence $f(y) > 0$ for $y > 0$. Thus for $0 \leq y \leq 1$, we have $1 - y \leq e^{-y}$, and raising this to the nth power, we obtain

$$(1-y)^n \leq e^{-yn}.$$

■

Now we can give the proof for the covering lemma:

Proof: Let us generate y^n by drawing i.i.d $\sim P_Y$.

Let us fix i, and so, from the strong ϵ typicality, we have:

$$\Pr\{(x^n, y^n(i) \in T_\epsilon^n(X,Y))\} \doteq 2^{-n(I(X;Y)\pm\epsilon)},$$

i.e.,

$$2^{-n(I(X;Y)+\epsilon)} \leq \Pr\{(x^n, y^n(i) \in T_\epsilon^n(X,Y))\} \leq 2^{-n(I(X;Y)\pm\epsilon)}.$$

Now, we can conclude:

$$\begin{aligned}
\Pr\{\forall i : (x^n, y^n(i) \notin T_\epsilon^n(X,Y))\} &\overset{(a)}{=} \prod_i \Pr\{(x^n, y^n(i) \notin T_\epsilon^n(X,Y))\} \\
&\overset{(b)}{\leq} \prod_i (1 - 2^{-n(I(X;Y)\pm\epsilon)}) \\
&= (1 - 2^{-n(I(X;Y)\pm\epsilon)})^{2^{nR}} \\
&\overset{(c)}{\leq} e^{-2^{n(R-I(X;Y)-\epsilon)}},
\end{aligned}$$

where

(a) follows from the fact that the probabilities are independent,

(b) follows from the strongly typical set properties,

(c) follows from lemma 2.

And finally, if $R > I(X;Y) + \epsilon$ we have

$$\Pr\{\forall i : (x^n, y^n(i) \notin T_\epsilon^n(X,Y))\} \to 0.$$

■

II. Writing on dirty paper

We consider the Gaussian Gelfand-Pinsker channel, when the state is known at the Encoder:

The name of the setting is due to its qualities-when the state is known at the encoder, and acts like noise

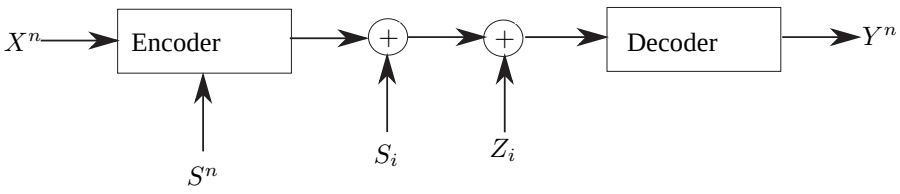

Fig. 1. Writing on dirty paper channel

in the channel itself, thus it can be referred as writing on a dirty paper, a name which was given by T. Cover.

Let $Y = X + S + Z$, where the state S is known to the encoder non-causally, and $S_i \sim N(0, \sigma_s^2)$ i.i.d, $Z_i \sim N(0, \sigma_z^2)$, $\frac{1}{n}\sum_1^n E(X_i^2) \leq P$.

What is the capacity of this setting?

Lemma 3 (Capacity of-writing on dirty paper channel) For the setting above, the capacity is

$$C = \frac{1}{2}\log(1 + \frac{P}{\sigma_z^2}).$$

Proof: Notice, that when the state is known at the decoder and encoder, then

$$C^* = \frac{1}{2}\log(1 + \frac{P}{\sigma_z^2}).$$

We will attempt to show achievability for C^*, and thus show that this is the capacity for this channel as well.

Let us use $U = \alpha S + X$, and let $X \sim N(0, P)$, then:

$$I(U;Y) - I(U;S) \quad = \quad h(U) - h(U|Y) - h(U) + h(U|S)$$

$$= \quad h(U|S) - h(U|Y). \tag{1}$$

Now, let us consider each expression:

Observe that $h(U|S) = h(X) = \frac{1}{2}\log(2\pi eP)$.

Also, $h(U|Y) = h(U,Y) - h(Y)$, when $h(Y) = \frac{1}{2}\log(2\pi e\sigma_y^2)$, when $\sigma_y^2 = P + \sigma_2^2 + \sigma_z^2$.

Further, $h(U,Y) = \frac{1}{2}\log(2\pi e)^2|K_{U,Y}|$, and to compute $|K_{U,Y}|$ we need:

$\sigma_u^2 = \alpha^2\sigma_s^2 + P$, and $cov(U,Y) = 2\sigma_s^2 + P$

Now, we can have:

$$\begin{aligned}
|K_{u,y}| &= \sigma_y^2\sigma_u^2 - cov(u,y)^2 \\
&= (P + \sigma_s^2 + \sigma_z^2)(\alpha^2\sigma_s^2 + P) - (2\sigma_s^2 + P)^2 \\
&= P\alpha^2\sigma_s^2 + P^2 + \alpha^2\sigma_s^4 + \sigma_s^2 P + \alpha^2\sigma_s^2\sigma_z^2 + \sigma_z^2 P - P^2 - 4\sigma_s^2 P - 4\sigma_s^4 \\
&= P\sigma_s^2(\alpha-1)^2 + \sigma_z^2(P + \alpha^2\sigma_s^2)
\end{aligned} \tag{2}$$

Therefore we obtain: $R = \frac{1}{2}\log\frac{P(P+\sigma_z^2+\sigma_s^2)}{P\sigma_s^2(\alpha-1)^2+\sigma_z^2(P+\alpha^2\sigma_s^2)}$. To find the maximum for the expression, we can differentiate over α. We know that $\log$ is a monotone function, and so we look for minimum in the denominator. By differentiating the denominator over α we obtain-

$$2P\sigma_s^2(\alpha-1) - 2\sigma_s^2\sigma_z^2\alpha = 0,$$

and therefore

$$\alpha = \frac{P}{P+\sigma_z^2}.$$

Now, we obtain:

$$\begin{aligned}
R &= \frac{1}{2}\log\frac{P(P+\sigma_z^2+\sigma_s^2)}{P\sigma_s^2\frac{\sigma_s^4}{(P+\sigma_z^2)^2} + \sigma_z^2(P + \frac{P^2\sigma_s^2}{(P+\sigma_z^2)^2})} \\
&= \frac{1}{2}\log\frac{(P+\sigma_z^2+\sigma_s^2)(P+\sigma_z^2)^2}{\sigma_z^2(\sigma_z^2\sigma_s^2 + P\sigma_s^2 + P^2 + 2P\sigma_z^2 + \sigma_z^4)} \\
&= \frac{1}{2}\log\frac{(P+\sigma_z^2+\sigma_s^2)(P+\sigma_z^2)^2}{\sigma_z^2(P+\sigma_z^2)(P+\sigma_s^2+\sigma_z^2)} \\
&= \frac{1}{2}\log(1 + \frac{P}{\sigma_z^2}) = C^*.
\end{aligned}$$

Thus, we get that the capacity of the writing on dirty paper channel, is C^*, the same as if there was no state information at all. ■

III. Bounding cardinality of auxiliary random variable

We say, that the support of a random variable, is the set of values it gets, with probability strictly larger then zero.

The following lemma is used for bounding the support of an auxiliary r.v. in capacity expressions:

Lemma 4 [4, CH 3] Let $\{f_i\}_1^k$ be continues functions, such that $\forall j = 1..k : E[f_j(X)] = A_j$.
Then there exists a random variable X' with finite alphabet $\{x_1..x_k\} \in |\mathcal{X}|$, with probabilities $\{\alpha 1..\alpha k\}$, $\sum_{i=1}^k \alpha_i = 1$, s.t.

$$\forall j = 1..k : \sum \alpha_i f_j(x_i) = E[f_j(X')] = A_j. \tag{3}$$

Example 1 (MAC) We know that a rate in the capacity region is given by:

$$R_1 = I(X_1; Y|X_2, Q) = \sum_q p(q) I(X1; Y|X_2, Q = q), \tag{4}$$

$$R_1 = I(X_2; Y|X_1, Q) = \sum_q p(q) I(X2; Y|X_1, Q = q), \tag{5}$$

$$R_1 + R_2 = I(X_1, X_2; Y|Q) = \sum_q p(q) I(X1, X_2; Y|Q = q). \tag{6}$$

Thus we get 3 conditions, and by the lemma, we can use Q with alphabet of cardinality 3.

Example 2 (Causal state information) We know that the capacity is given by:

$$C = \max_{p(u), x=f(u,s)} I(U;Y) = \max_{p(u), x=f(u,s)} H(Y) - H(Y|U).$$

If so, the first condition will come from:
$(1) : H(Y|U) = \sum_u p(u) H(Y|U = u)$.
Also, we have:
$(2) - (|\mathcal{Y}|) : p(y) = \sum_u p(u) p(y|u)$,
Thus the support of U is with cardinality as $|\mathcal{Y}|$.

REFERENCES

[1] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. New-York: Wiley, 2006.

[2] S. I. Gelfand and M. S. Pinsker, "Coding for channel with random parameters", *Probl. Control Inf. Theory*, vol.9, no.1, pp.19-31, 1980.

[3] M.H.M. Costa, "Writing on dirty paper", *IEEE Trans. Inf. Theory*, vol.IT-29, no.3,pp.439-441, May 1983.

[4] I. Csiszar and J. Korner, "Coding theorems for discrete memoryless systems".