EE7101: Introduction to Information and Coding Theory

Handout 7 – Arimoto-Blahut

Tay Wee Peng

1 Overview

In this lecture we present a method to compute the capacity of any given channel.

2 Arimoto-Blahut

Given $p(x)$, $p(y|x)$, we want to find $r(x|y)$ to maximize

$$\sum_{x,y} p(x)p(y|x) \log \frac{r(x|y)}{p(x)}.$$

The maximizer is given by

$$r^*(x|y) = \frac{p(x)p(y|x))}{\sum_{x'} p(x')p(y|x')}.$$

Proof.

$$\begin{aligned}
&\sum_{x,y} p(x)p(y|x) \log \frac{r^*(x|y)}{p(x)} - \sum_{x,y} p(x)p(y|x) \log \frac{r(x|y)}{p(x)} \\
&= \sum_{x,y} p(x)p(y|x) \log \frac{r^*(x|y)}{r(x|y)} \\
&= \sum_{x,y} p(y)r^*(x|y) \log \frac{r^*(x|y)}{r(x|y)} \\
&= \sum_{y} p(y)\mathrm{D}(r^*(\cdot|y)||r(\cdot|y)) \\
&\geq 0.
\end{aligned}$$

Equality if and only if $r^*(x|y) = r(x|y)$, $\forall x, y$. □

$$\begin{aligned}
I(X;Y) &= \sum_{x,y} p(x)p(y|x) \log \frac{p(x|y)}{p(x)} \\
&= \max_{r(x|y)} \sum_{x,y} p(x)p(y|x) \log \frac{r(x|y)}{p(x)}
\end{aligned}$$

$$
\begin{aligned}
C &= \max_{p(x)} I(X;Y) \\
&= \max_{p(x)} \max_{r(x|y)} \sum_{x,y} p(x)p(y|x) \log \frac{r(x|y)}{p(x)}
\end{aligned}
$$

Use alternating maximization procedure:

- fix $p(x)$, find optimal $r(x|y)$;
- fix $r(x|y)$, find optimal $p(x)$.

Fixed $p(x)$, optimal

$$
r(x|y) = \frac{p(x)p(y|x)}{\sum_{x'} p(x')p(y|x')}
$$

Fixed $r(x|y)$, use Lagrange multiplier:

$$
J = \sum_{x,y} p(x)p(y|x) \log \frac{r(x|y)}{p(x)} + \lambda \sum_{x} p(x)
$$

$$
\begin{aligned}
\frac{\partial J}{\partial p(x)} = \sum_{y} (-p(y|x) \log p(x) - p(y|x) + p(y|x) \log r(x|y)) + \lambda = 0. \\
-\log p(x) - 1 + \sum_{y} p(y|x) \log r(x|y) + \lambda = 0
\end{aligned}
$$

$$
p(x) = \frac{\exp\left(\sum_y p(y|x) \log r(x|y)\right)}{\sum_{x'} \exp\left(\sum_y p(y|x') \log r(x'|y)\right)}
$$

At kth iteration:

$$
\begin{aligned}
r^k(x|y) &= \frac{p^k(x)p(y|x)}{\sum_{x'} p^k(x')p(y|x')} \\
p^{k+1}(x) &= \frac{\exp\left(\sum_y p(y|x) \log r^k(x|y)\right)}{\sum_{x'} \exp\left(\sum_y p(y|x') \log r^k(x'|y)\right)} \\
&= p^k(x) \frac{\tilde{q}^k(x)}{\sum_{x'} q^k(x')}
\end{aligned}
$$

where $\tilde{q}^k(x)$ and $q^k(x)$ are defined as

$$\tilde{q}^k(x) = \exp\left(\sum_y p(y|x) \log \frac{p(y|x)}{\sum_{x'} p^k(x')p(y|x')}\right)$$
$$q^k(x) = \exp\left(\sum_y p(y|x) \log r^k(x|y)\right).$$

Then,

$$I^k = \mathrm{D}(p^k(x)p(y|x)||p^k(x)p^k(y))$$

where $p^k(y)$ is defined as

$$p^k(y) = \sum_x p^k(x)p(y|x).$$

And

$$\begin{aligned} J^k &= \sum_{x,y} p^{k+1}(x)p(y|x) \log \frac{r^k(x|y)}{p^{k+1}(x)} \\ &= \log\left(\sum_x q^k(x)\right). \end{aligned}$$

Then, we have $I^k \leq J^k \leq I^{k+1} \leq \ldots \leq C$. And I^k and J^k both converge as $k \to \infty$. Next, we want to show that $J^k \to C$ as $k \to \infty$.

Proof. Let $p^*(x)$ be the maximizer for C. Consider

$$
\begin{aligned}
&\sum_x p^*(x) \log\left(\frac{p^{k+1}(x)}{p^k(x)}\right) \\
&= \sum_x p^*(x) \log\left(\frac{\tilde{q}^k(x)}{\sum_{x'} q^k(x')}\right) \\
&= \underbrace{\sum_x p^*(x) \log \tilde{q}^k(x)} - J^k \\
&= \sum_x p^*(x) \sum_y p(y|x) \log \frac{p(y|x)}{\sum_{x'} p^k(x') p(y|x')} \\
&= \sum_y \sum_x p^*(x) p(y|x) \log \frac{p(y|x)}{\sum_{x'} p^k(x') p(y|x')} \\
&\geq \sum_y \sum_x p^*(x) p(y|x) \log \frac{p(y|x)}{\sum_{x'} p^*(x') p(y|x')}, \qquad \because \mathrm{D}(p^*(y)||p^k(y)) \geq 0, \\
&= C.
\end{aligned}
$$

$$
\begin{aligned}
\implies \sum_{k=0}^{m} (C - J^k) &\leq \sum_x \sum_{k=0}^{m} p^*(x) \log\left(\frac{p^{k+1}(x)}{p^k(x)}\right) \qquad \forall m. \\
&= \sum_x p^*(x) \log\left(\frac{p^{m+1}(x)}{p^0(x)}\right) \\
&\leq \sum_x p^*(x) \log\left(\frac{p^*(x)}{p^0(x)}\right) < \infty \\
&\implies J^k \to C.
\end{aligned}
$$

□

But the convergence can be very slow.