Electrical Engineering 229A Lecture 14 Notes

Daniel Raban

October 12, 2021

1 Joint ε-Weak Typicality and the Slepian-Wolf Theorem

1.1 Properties of joint ε-weak typicality

Suppose $(X_1, Y_1), (X_2, Y_2), \dots$ are i.i.d. with $(X_i, Y_i) \in \mathscr{X} \times \mathscr{Y}$ finite and $(X_i, Y_i) \sim (p(x, y), x \in \mathscr{X}, y \in \mathscr{Y})$. We think of the X_is as being seen by Alice and the Y_is as being seen by Bob.

Definition 1.1 (Joint ε-weak typicality)**.** Define the set $A_\varepsilon^{(n)} \subseteq \mathscr{X}^n \times \mathscr{Y}^n$ to be the set of $(x_1^n, y_1^n) \in \mathscr{X}^n \times \mathscr{Y}^n$ such that

1. $|-\frac{1}{n} \log p(x_1^n) - H(X)| < \varepsilon$,
2. $|-\frac{1}{n} \log p(y_1^n) - H(Y)| < \varepsilon$,
3. $|-\frac{1}{n} \log p(x_1^n, y_1^n) - H(X, Y)| < \varepsilon$.

Here are some properties of this:

Theorem 1.1.

1.

$$\mathbb{P}((X_1^n, Y_1^n) \in A_\varepsilon^{(n)}) \xrightarrow{n \to \infty} 1.$$

Proof. Use the weak law of large numbers. □

2.

$$|A_\varepsilon^{(n)}| \leq 2^{nH(X,Y)} 2^{n\varepsilon}.$$

Proof. For all $(x_1^n, y_1^n) \in A_\varepsilon^{(n)}$,

$$p(x_1^n, y_1^n) \geq 2^{-nH(X,Y)} 2^{-n\varepsilon}$$

and

$$1 \geq \sum_{(x_1^n, y_1^n) \in A_\varepsilon^{(n)}} p(x_1^n, y_1^n).$$ □

3. *For all large enough n,*

$$|A_\varepsilon^{(n)}| \geq (1-\delta)2^{nH(X,Y)}2^{-n\varepsilon}.$$

Proof. For all $(x_1^n, y_1^n) \in A_\varepsilon^{(n)}$,

$$p(x_1^n, y_1^n) \leq 2^{-nH(X,Y)}2^{n\varepsilon}$$

and, for all large enough n,

$$\sum_{(x_1^n, y_1^n) \in A_\varepsilon^{(n)}} p(x_1^n, y_1^n) \geq 1 - \delta. \qquad \square$$

4. *If $(\widetilde{X}_1^n, \widetilde{Y}_1^n) \sim p(x_1^n)p(y_1^n)$, then*

   (a)

   $$\mathbb{P}((\widetilde{X}_1^n, \widetilde{Y}_1^n) \in A_\varepsilon^{(n)}) \leq 2^{-nI(X;Y)}2^{3n\varepsilon}.$$

   Proof. The left hand side is

   $$\begin{aligned}\sum_{(x_1^n, y_1^n)} p(x_1^n)p(y_1^n) &\leq |A_\varepsilon^{(n)}|2^{-nH(X)}2^{n\varepsilon}2^{-nH(Y)}2^{n\varepsilon} \\ &\leq 2^{nH(X,Y)}2^{-nH(X)}2^{-nH(Y)}2^{3n\varepsilon}.\end{aligned} \qquad \square$$

   (b) For all $\delta > 0$,

   $$\mathbb{P}((\widetilde{X}_1^n, \widetilde{Y}_1^n) \in A_\varepsilon^{(n)}) \geq (1-\delta)2^{-nI(X;Y)}2^{-3n\varepsilon}.$$

   Proof. The left hand side is

   $$\begin{aligned}\sum_{(x_1^n, y_1^n)} p(x_1^n)p(y_1^n) &\geq |A_\varepsilon^{(n)}|2^{-nH(X)}2^{-n\varepsilon}2^{-nH(Y)}2^{-n\varepsilon} \\ &\geq (1-\delta)2^{nH(X,Y)}2^{-nH(X)}2^{-nH(Y)}2^{-3n\varepsilon}.\end{aligned} \qquad \square$$

1.2 The Slepian-Wolf theorem on distributed lossless compression

In this section, lossless is interpreted in the sense of asymptotically vanishing error probability. The scenario is that Alice sees $X_1, \ldots, X_n$ and Bob sees $Y_1, \ldots, Y_n$. The pairs (X_i, Y_i) with $i = 1, \ldots, n$ are iid and $(X_i, Y_i) \sim (p(x,y), x \in \mathscr{X}, y \in \mathscr{Y})$. Alice compresses X_1^n, and Bob compresses Y_1^n. A fusion centor sees the compressed representations and wants to recover (X_1^n, Y_1^n) with small probability of error (going to 0 as $n \to \infty$). The problem is: What region of (Alice's bits/symbol, Bob's bits/symbol) is achievable?

Definition 1.2. We say that the pair of rates (R_1, R_2) is **achievable** if there is a sequence $((e_n^{(1)}, e_n^{(2)}, d_n), n \geq 1)$ where

$$e_n^{(1)} : \mathscr{X}^n \to [M_n^{(1)}] = \{1, \ldots, M_n^{(1)}\}, \quad \text{with} \quad \limsup_{n\to\infty} \frac{1}{n} \log M_n^{(1)} \leq R_1,$$

$$e_n^{(2)} : \mathscr{X}^n \to [M_n^{(2)}], \quad \text{with} \quad \limsup_{n\to\infty} \frac{1}{n} \log M_n^{(2)} \leq R_2,$$

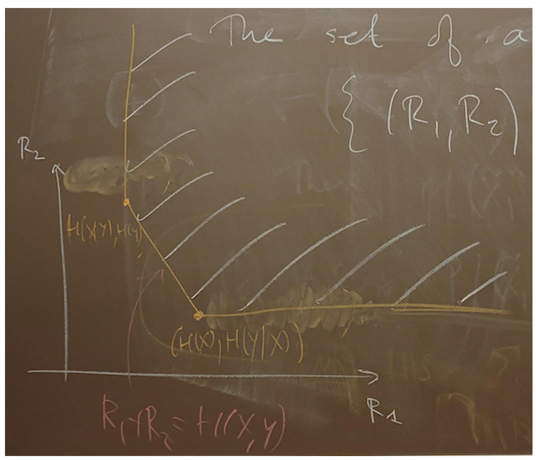
$$d_n : [M_n^{(1)}] \times [M_n^{(2)}] \to \mathscr{X}^n \times \mathscr{Y}^n,$$

such that

$$\mathbb{P}(d_n(e_n^{(1)}(X_1^n), e_n^{(2)}(Y_1^n)) \neq (X_1^n, Y_1^n)) \xrightarrow{n\to\infty} 0.$$

Theorem 1.2 (Slepian-Wolf)**.** *The set of achievable rate pairs is*

$$\{(R_1, R_2) : R_1 \geq H(X \mid Y), R_2 \geq H(Y \mid X), R_1 + R_2 \geq H(X, Y)\}.$$



We will prove the achievability using the *probabilistic method*; i.e. we will show that a suitable $((e_n^{(1)}, e_n^{(2)}, d_n), n \geq 1)$ exists without explicitly demonstrating it. Here is an example of the probabilistic method.

Example 1.1. Suppose that $f : [0, 1] \to \mathbb{R}_+$. To show "there exists some x such that $f(x) > 10$," it's enough to show that $\mathbb{E}[f(Z)] > 10$ where $Z \sim \text{Unif}([0, 1])$.

Proof. Achievability: It is enough to show that for all $\varepsilon > 0$, if (R_1, R_2) is such that $R_1 \geq H(X \mid Y) + \varepsilon$, $R_2 \geq H(Y \mid X) + \varepsilon$, and $R_1 + R_2 \geq H(X, Y) + \varepsilon$, then (R_1, R_2) is achievable. We use a "random binning" argument: $(e_n^{(1)}, e_n^{(2)}, d_n)$ will be random variables with

- $e_n^{(1)}$: randomly assign each $x_1^n \in \mathscr{X}^n$ to one of $M_n^{(1)}$ bins uniformly, independently over x_1^n,
- $e_n^{(2)}$: randomly assign each $y_1^n \in \mathscr{Y}^n$ to one of $M_n^{(2)}$ bins uniformly, independently over y_1^n
- $d_n(m_n^{(1)}, m_n^{(2)}) = (\widehat{x}_1^n, \widehat{x}_2^n)$ if there is exactly one $(\widehat{x}_1^n, \widehat{y}_1^n)$ with $e_n^{(1)}(\widehat{x}_1^n) = m_n^{(1)}$ and $e_n^{(2)}(\widehat{y}_1^n) = m_n^{(2)}$. Otherwise, $d_n(m_n^{(1)}, m_n^{(2)})$ can take any value.

Now we upper bound $\mathbb{P}(d_n(e_n^{(1)}(X_1^n), e_n^{(2)}(Y_1^n)) \neq (X_1^n, Y_1^n))$, where the randomness is in both (X_1^n, Y_1^n) and $(e_n^{(1)}, e_n^{(2)}, d_n)$. We have

$$\mathbb{P}(d_n(e_n^{(1)}(X_1^n), e_n^{(2)}(Y_1^n)) \neq (X_1^n, Y_1^n)) \leq \underbrace{\mathbb{P}(E_{0,n})}_{\xrightarrow{n\to\infty} 0} + \mathbb{P}(E_{1,n}) + \mathbb{P}(E_{2,n}) + \mathbb{P}(E_{12,n}).$$

Here,

- $E_{0,n} = \{(X_1^n, Y_1^n) \notin A_n^{(\delta)}\}$ for some $\delta > 0$, and the corresponding probability goes to 0 as $n \to \infty$.
- $E_{1,n} = \{\exists\, \widetilde{x}_1^n \neq X_1^n \text{ with } e_n^{(1)}(\widetilde{x}_1^n) = e_n^{(1)}(X_1^n) \text{ and } (\widetilde{x}_1^n, y_1^n) \in A_n^{(\delta)}\}$. Here,
$$\mathbb{P}(E_{1,n}) \leq \sum_{(x_1^n, y_1^n)} p(x_1^n, y_1^n) \sum_{\substack{\widetilde{x}_1^n \neq x_1^n \\ (\widetilde{x}_1^n, y_1^n) \in A_n^{(\delta)}}} \underbrace{\mathbb{P}(e_n^{(1)}(\widetilde{x}_1^n) = e_n^{(1)}(x_1^n))}_{=1/M_n^{(1)}}.$$
Now $|\{\widetilde{x}_1^n : (\widetilde{x}_1^n, y_1^n) \in A_n^{(\delta)}\}| \leq 2^{nH(X|Y)} 2^{2n\delta}$ because
$$\begin{aligned} 1 &\geq \sum_{\widetilde{x}_1^n : (\widetilde{x}_1^n, y_1^n) \in A_n^{(\delta)}} p(\widetilde{x}_1^n \mid y_1^n) \\ &= \sum_{\widetilde{x}_1^n : (\widetilde{x}_1^n, y_1^n) \in A_n^{(\delta)}} \frac{p(\widetilde{x}_1^n, y_1^n)}{p(y_1^n)} \\ &\geq |\{\widetilde{x}_1^n : (\widetilde{x}_1^n, y_1^n) \in A_n^{(\delta)}\}| 2^{-nH(X|Y)} 2^{-2n\delta}. \end{aligned}$$
So
$$\mathbb{P}(E_{1,n}) \leq \sum_{(x_1^n, y_1^n)} p(x_1^n, y_1^n) 2^{nH(X|Y)} 2^{2n\delta} 2^{-nR_1}.$$
But $R_1 > H(X \mid Y) + \varepsilon$ by assumption, so if $2\delta < \varepsilon$, the right hand side goes to 0 as $n \to \infty$.
- $E_{2,n}$ is defined similarly to $E_{1,n}$, and $\mathbb{P}(E_{2,n}) \to 0$ as $n \to \infty$.

We are now left with $\mathbb{P}(E_{12,n})$, which we will examine next time. $\square$