Likelihood Encoder for Source Coding

Paul Cuff, Eva C. Song
Likelihood Encoder

• Given:
  - Joint Distribution: \( P_{X,U} \)
  - Codebook: \( \{u^n(m)\}_{m=1}^{2^{nR}} \)

• Encoder:

\[
P_{M|X^n}(m|x^n) \propto \prod_{t=1}^{n} P_{X|U}(x_t|u_t(m))
\]

• Encoder (given a codebook and joint distribution) chooses messages stochastically, with probability proportional to the likelihood of the observation \( X^n \) given each codeword.
Previous uses of likelihood encoder

- Our Research Group
  - Schieler, Satpathy, Song
- Purposes:
  - Channel synthesis
  - Rate distortion theory for secrecy systems

- Likelihood Decoder proposed for channel decoding
  - [Yassaee, Aref, Gohari 13], ISIT Student Paper Award
Soft Covering Lemma

Lemma: If the codebook is randomly constructed and $R > I(X; U)$, then

$$E_C \left\| P_{X^n} - \prod P_X \right\|_{TV} \to 0$$

See

- [Wyner 75]: Common Information
- [Han-Verdu 93]: Channel Resolvability
- [Cuff 13]: Distributed Channel Synthesis
Comment

• Analysis of likelihood encoder is simple. If rates are high enough, induced distribution behaves as a uniform distribution over the codebook connected to the channel.
i.i.d. source sequence $X^n$ distributed according to $X_i \sim \mathcal{P}_X$

Encoder $f_n : \mathcal{X}^n \to \mathcal{M}$ (possibly stochastic).
Decoder $g_n : \mathcal{M} \to \mathcal{Y}^n$ (possibly stochastic).
Compression rate: $R$, i.e. $|\mathcal{M}| = 2^{nR}$.
Fidelity requirement: $\mathbb{E} \ d(X^n, Y^n) \leq D$, where $d(x^n, y^n) = \frac{1}{n} \sum_{i=1}^{n} d(x_i, y_i)$.

$$R = \min_{P_{Y|X}: \mathbb{E}[d(X,Y)] \leq D} I(X;Y).$$
Proof of Achievability

• Select $\bar{P}_{X,Y} = \bar{P}_X \bar{P}_{Y|X}$ such that
  • $R > I(X;Y)$
  • $\mathbb{E}[d(X,Y)] < D$

• Use a random codebook and the likelihood encoder
Proof Continued

- Induced distribution: $P_{X^n,Y^n}^{(0)}$
- Idealized distribution: $P_{X^n,Y^n}^{(1)}(x^n, y^n)$

\[
\Delta \left( \prod_{t=1}^{n} \overline{P}_{X|Y}(x_t|y_t) \right) \frac{1}{2^{nR}} \sum_{m=1}^{2^{nR}} \mathbb{1}\{Y^n(m) = y^n\}.
\]

- Notice:
  - $E_C P_{X^n,Y^n}^{(1)}(x^n, y^n) = \prod_{t=1}^{n} \overline{P}_{X,Y}(x_t, y_t)$
  - $P_{Y^n|X^n}^{(0)} = P_{Y^n|X^n}^{(1)}$
Proof Continued

• By the soft covering lemma
  
  \[ E_C \left\| P_{X^n}^{(0)} - P_{X^n}^{(1)} \right\|_{TV} \to 0 \]

• Therefore
  
  \[ E_C \left\| P_{X^n,Y^n}^{(0)} - P_{X^n,Y^n}^{(1)} \right\|_{TV} \to 0 \]

\[
E_C E_{P^{(0)}} d(X^n, Y^n) \leq E_C \left[ E_{P^{(1)}} d(X^n, Y^n) + d_{\max} \left\| P^{(0)} - P^{(1)} \right\|_{TV} \right]
\]

\[
= E_{P} d(X^n, Y^n) + d_{\max} E_C \left\| P^{(0)} - P^{(1)} \right\|_{TV}
\]

\[
< D \quad \text{for } n \text{ large enough.}
\]
Properties of Total Variation

- $\|P_X - Q_X\|_{TV} \leq \|P_{X,Y} - Q_{X,Y}\|_{TV}$
- $\|P_X - Q_X\|_{TV} = \|P_X P_{Y|X} - Q_X P_{Y|X}\|_{TV}$
- $|E_P f(X) - E_Q f(X)| \leq 2d_{max} \|P - Q\|_{TV}$
- If $P(X \neq \hat{X}) \leq \varepsilon$, then $\|P_{X,Y} - P_{\hat{X},Y}\|_{TV} \leq \varepsilon$. 
Source Coding with Decoder Side Information

- Sources $X^n$ and $Y^n$ i.i.d. according to $P_{X,Y}$
- Encoder observes $X^n$
- Decoder observes $Y^n$
- Message from encoder to decoder at rate $R$
- Decoder produces $\hat{X}^n$
- Distortion constraint: $E d(X^n, \hat{X}^n) < D$

$R > I(X; U|Y)$
where $U - X - Y$ form a Markov chain
and there exists an $f$ such that $E d(X, f(U, Y)) < D$
Proof of Achievability

• Select $\bar{P}_{U,X,Y} = \bar{P}_{X,Y} \bar{P}_{U|X}$ and $f$ such that
  • $R > I(X; U|Y)$
  • $E[d(X, f(U, Y))] < D$

• Use a random codebook with two indices:
  • $M_a$ at rate $R_a > I(X; U|Y)$
  • $M_b$ at rate $R_b < I(Y; U)$
  • $R_a + R_b > I(X, Y; U)$

• Likelihood encoder

• Decoder:
  • Channel Decoder to decode $M_b$
  • Apply function $f$
Proof Continued

- \( P^{(0)} = \bar{P}_{X^n,Y^n}P_{LE}(m_a,m_b|x^n)P_D(m'_b|m_a,y^n)P_f(\hat{x}^n|m_a,m'_b,y^n) \)

- \( P^{(1)} = Q_{X^n,Y^n,M_a,M_b}P_D(m'_b|m_a,y^n)P_f(\hat{x}^n|m_a,m'_b,y^n) \)
  - \( Q \) is the idealized distribution
  - Soft covering lemma says \( P^{(0)} \approx P^{(1)} \)
  - Under \( P^{(1)} \), \( P(M_b \neq M'_b) \to 0 \) by channel decoding proof.

- \( P^{(2)} = Q_{X^n,Y^n,M_a,M_b}P_D(m'_b|m_a,y^n)P_f(\hat{x}^n|m_a,m_b,y^n) \)
  \[ E_C P^{(2)} = \prod \bar{P}_{X,\hat{x}} \]
Proof Intuition

\[ X^n, Y^n \]

\[ u^n(m_a, m_b) \]
Proof Intuition

Proof Intuition

$X^n, Y^n$

Channel

$u^n(m_a, m_b)$

Total Variation Error:

$\varepsilon_{SC}$

Soft Covering Lemma
Proof Intuition

\[ X^n, Y^n \]

\[ u^n(m_a, m_b) \]

Channel

Total Variation Error:

\[ \varepsilon_{SC} \]

\[ \varepsilon_{Decode} \]

Channel Decoding
Proof Intuition

Expectation over codebooks

Channel

$X^n, Y^n$

$U^n$

Total Variation Error:

$\varepsilon_{SC}$

$\varepsilon_{Decode}$
Summary

• Simple Proofs of classical (and new) results using the likelihood encoder.

• Proof techniques easily extend to other problems:
  • E.g. Multi-terminal source coding [Berger-Tung]