Research Software for Information Theory

Ido Gattegno
Supervised by: Haim H. Permuter

Ben-Gurion University

June 4, 2015
Automated Tasks

Which research tasks can automated?
How?
Overview

We discuss the following topics:

- Shannon-type information inequalities (Yeung, 1997).
- Fourier-Motzkin elimination and simplification.
- Converse research tool - capacity regions.
Linear information inequalities:

- Hold for every PMF on the random variables:
  \[ I(X; Y) \leq I(X; Y, Z) \]

- Larger and simpler (linear) search domain:

Figure: $\Gamma^*_n$ - All inequalities, $\Gamma_n$ - Shannon-type inequalities
Shannon-type information inequalities

- Linear combination of information measures:

  \[
  \begin{array}{c|c|c|c}
  & H(X_1, X_2) & H(X_1) & H(X_2 | X_1) \\
  \hline
  H(X_1) & H(X_2) & & \\
  H(X_1 | X_2) & & H(X_2) & \\
  \hline
  H(X_1 | X_2) & I(X_1; X_2) & H(X_2 | X_1) & \\
  \end{array}
  \]

  **Elemental**

  Figure: Relations between information measures

- Non-negativity of *elemental information measures* imply others!

  \[
  G_{h\ell} \geq 0
  \]
Shannon’s Information Measures

- Reading an expression:

\[ S_1 \cap S_2 \]

\[ H(X_1|X_2) \]

\[ I(X_1;X_2) \]

\[ H(X_2|X_1) \]

Figure: Commas and symbols → set operations

- Constraints on PMF induce constraints on information measures:

Markov: \( X - Y - Z \) ⇔ \( I(X;Z|Y) = 0 \)

- General constraints

\[ \mathcal{Q}h_\ell = 0 \]
Solving a linear programming (LP) problem:

\[ \rho^* = \min_{h} f^T h \]
\[ \text{subject to:} \]
\[ G h \geq 0 \]
\[ Q h = 0 \]

If \( \rho^* = 0 \) then \( f^T h \geq 0 \).

How can we use this mechanism?
Fourier-Motzkin elimination:

- Eliminate variables from linear constraints system:
  \[
  \begin{align*}
  x &< 1 \\
  x + y &> 2 \\
  \end{align*}
  \Rightarrow \quad y > 1
  \]

  Results in a linear constraints system:
  \[Ax \geq b\]

- Identification of a redundant constraint:
  \[
  \rho^* = \min_{x: A^{(i)}x \geq b^{(i)}} a_i^\top x
  \]

  if \(\rho^* \geq b_i\) then \(a_i^\top \geq b_i\) is a redundant constraint.
Fourier-Motzkin elimination

- The outcome system may contain redundant constraints.
  - Implied by other constraints.
  - Implied by Shannon’s type inequalities - **Add matrix \( G \) as constraints!**

- Software available at:
  - http://www.ee.bgu.ac.il/~fmeit/.

- What more can we do?
Converse prover

Capacity of communication channels

Figure: Discrete memoryless channel - Point-to-Point

\[ M \rightarrow X^n, \quad M \in [1 : 2^{nR}] \]
\[ Y^n \rightarrow \hat{M}, \quad P_{error} \xrightarrow{n \to \infty} 0 \]
\[ C_{PTP} = \max_{P(x)} I(X; Y) \]
Achievable transmission rate (√).

Upper bound for all rates ← GOAL.

We can use:

Joint PMF:

\[ P(m) \prod_{i=1}^{n} 1(x_i|m)P(y_i|x_i), \quad (x_i = f(m)) \]

induce constraints on information measures!

Inductive proof for upper bounds.
**Converse prover**

**Inductive proof**

- For $k = \{2, 3, \ldots \}$,

$$
P(m, x^k, y^k) = P(m) \underbrace{P(x^{k-1} | m)}_{H(X^{k-1} | M) = 0} \underbrace{P(y^{k-1} | x^{k-1})}_{I(Y^{k-1}; M | X^{k-1}) = 0} \underbrace{P(y_k | x_k)}_{\ldots} \underbrace{1(x_k | m)}_{P(m, x^k, y^k) = P(m)}
$$

- Shannon-type inequalities:

$$
I(M; Y^{k-1}, Y_k) \leq I(X^{k-1}; Y^{k-1}) + I(X_k; Y_k)
$$

$$
I(X^{k-1}, X_k; Y^{k-1}, Y_k) \leq I(X^{k-1}; Y^{k-1}) + I(X_k; Y_k)
$$

- Conclude that

$$
I(M; Y^n) \leq \sum_{i=1}^{n} I(X_i; Y_i), \quad \forall n = \{2, 3, \ldots \}
$$
Example - Gelfand-Pinsker (GP)

Joint PMF:

\[ P(m, s^n, x^n, y^n) = p(m) \prod_{i=1}^{n} p(s_i) \mathbf{1}(x_i|m, s_i) p(y_i|s_i, x_i) \]

Channel capacity:

\[ C_{SI-E} = \max_{P(u|s), x(u,s)} (I(U; Y) - I(U; S)) \]
Using STIs, we show that:

$$I(M; Y^n) \leq \sum_{i=1}^{n} I(M, S_{i+1}^n, Y^{i-1}; Y_i) - I(M, S_{i+1}^n, Y^{i-1}; S_i).$$

Equivalent to:

$$I(M; Y^n) \leq \sum_{i=1}^{n} I(U_i; Y_i) - I(U_i; S_i).$$

where $U_i = (M, S_{i+1}^n, Y^{i-1})$. 
Conclusions and further research

- Shannon-type inequalities are used in research software (√).
- Finding auxiliary random variables ← How?.
- Extend the result to multi-user channels.

Thank you!