Relations between matroids and mutual information with application to MAC polar codes

Emmanuel Abbe, EPFL

December 4, 2009

Information...

Information Measures: one terminal

$$P \in M_1(\mathcal{X})$$

$$\rightarrow \quad H(P) = \mathbb{E}_P \log \frac{1}{P}$$

Information Measures: one terminal

$$P \in M_1(\mathcal{X}) \qquad \leftarrow \qquad \text{distribution of } X$$

$$\rightarrow \qquad H(P) = \mathbb{E}_P \log \frac{1}{P} \qquad = H(X)$$

Information Measures: one terminal

$$P \in M_1(\mathcal{X}) \qquad \leftarrow \qquad \text{distribution of } X$$

$$\rightarrow \quad H(P) = \mathbb{E}_P \log \frac{1}{P} \quad = \; H(X)$$

$\rightarrow$ min. avg. nb. of bits to describe X

$\rightarrow$ lower bound on the compression of X

$$X \sim P$$

$$W \in M_1(\mathcal{Y}|\mathcal{X})$$

Information Measures: one terminal

$X \sim P$ input $X \xrightarrow{W} Y$ output

$W \in M_1(\mathcal{Y}|\mathcal{X})$ $\longleftarrow$ channel (distribution)

Information Measures: one terminal

$X \sim P$ input $X \xrightarrow{W} Y$ output

$W \in M_1(\mathcal{Y}|\mathcal{X}) \quad \leftarrow \quad$ channel (distribution)

$$\begin{aligned} \rightarrow \quad & I(P,W) = \mathbb{E}_\mu \log \frac{\mu}{\mu_\mathcal{X} \times \mu_\mathcal{Y}}, \quad \mu = P \circ W \\ = \quad & I(X;Y) = H(X) - H(X|Y) = H(Y) - H(Y|X) \end{aligned}$$

Information Measures: one terminal

$X \sim P$ input $X \xrightarrow{W} Y$ output

$W \in M_1(\mathcal{Y}|\mathcal{X})$ $\leftarrow$ channel (distribution)

$$\rightarrow \quad I(P,W) = \mathbb{E}_\mu \log \frac{\mu}{\mu_\mathcal{X} \times \mu_\mathcal{Y}}, \quad \mu = P \circ W$$
$$= \quad I(X;Y) = H(X) - H(X|Y) = H(Y) - H(Y|X)$$

$\rightarrow$ mutual information

Information Measures: one terminal

Define the uniform mutual information of a channel $W \in M(\mathcal{Y}|\mathcal{X})$ by $I_U(W) := I(U_{\mathcal{X}} \circ W)$, where $U_{\mathcal{X}}$ is the uniform distribution on $\mathcal{X}$.

$$I_U(W) = 1 \quad \Leftrightarrow \quad W = \text{ perfect channel}$$
$$I_U(W) = 0 \quad \Leftrightarrow \quad W = \text{ pure noise channel}$$

Information Measures: one terminal

Define the uniform mutual information of a channel $W \in M(\mathcal{Y}|\mathcal{X})$ by $I_U(W) := I(U_{\mathcal{X}} \circ W)$, where $U_{\mathcal{X}}$ is the uniform distribution on $\mathcal{X}$.

$$\begin{aligned} I_U(W) = 1 \quad &\Leftrightarrow \quad W = \text{ perfect channel} \\ I_U(W) = 0 \quad &\Leftrightarrow \quad W = \text{ pure noise channel} \end{aligned}$$

$\rightarrow$ We know how to transfer information with low complexity on these extremal channels, and polarization bring them

Information Measures: multiple terminals

A multiple access channel (MAC) with m users is an element of $M_1(\mathcal{Y}|\mathcal{X}^m)$,

$$(X_1, \ldots, X_m) \xrightarrow{W} Y$$

Definition

The mutual information collection of a MAC $W \in M(\mathcal{Y}|\mathcal{X}^m)$ is

$$\{I(X[S]; Y X[S^c]), \quad S \subseteq \{1, \ldots, m\}\}$$

where $(X[1 \ldots m], Y)$ has joint distribution $(P_1 \times \ldots \times P_m) \circ W$.

$X[S] = \{X_i\}_{i \in S}$

Information Measures: multiple terminals

A multiple access channel (MAC) with m users is an element of $M_1(\mathcal{Y}|\mathcal{X}^m)$,

$$(X_1, \ldots, X_m) \xrightarrow{W} Y$$

Definition

The mutual information collection of a MAC $W \in M(\mathcal{Y}|\mathcal{X}^m)$ is

$$\{I(X[S]; Y X[S^c]), \quad S \subseteq \{1, \ldots, m\}\}$$

where $(X[1 \ldots m], Y)$ has joint distribution $(P_1 \times \ldots \times P_m) \circ W$.

$X[S] = \{X_i\}_{i \in S}$

Define $\rho: \quad 2^m \to \mathbb{R}$

$$S \mapsto I(X[S]; Y X[S^c])$$

Information Measures: multiple terminals

Operational meaning:

$$0 \leq \sum_{s \in S} R_s \leq \rho(S), \quad \forall S \in 2^m$$

leads to the capacity region of non cooperating users over a memoryless MAC.

Example: $m = 2$:

$$\begin{aligned} R_1 &\leq I(X[1]; Y X[2]) \\ R_2 &\leq I(X[2]; Y X[1]) \\ R_1 + R_2 &\leq I(X[1] X[2]; Y) \end{aligned}$$

Operational meaning:

$$0 \leq \sum_{s \in S} R_s \leq \rho(S), \quad \forall S \in 2^m$$

leads to the capacity region of non cooperating users over a memoryless MAC.

Example: $m = 2$:

$$R_1 \leq I(X[1]; Y X[2])$$
$$R_2 \leq I(X[2]; Y X[1])$$
$$R_1 + R_2 \leq I(X[1] X[2]; Y)$$

Question: take uniform input distributions, what would be an extremal MAC?

matroids...

Matroids: Independence

Definition

A matroid M is an ordered pair $(E, \mathcal{I})$, where E is a finite set called ground set and $\mathcal{I}$ is a collection of a subsets of E called the independent sets, which satisfies:

$(I1)$ $\emptyset \in \mathcal{I}$.

$(I2)$ If $I \in \mathcal{I}$ and $I' \subseteq I$, then $I' \in \mathcal{I}$.

$(I3)$ If $I_1, I_2 \in \mathcal{I}$ and $|I_1| < |I_2|$, then there exists an element $e \in I_2 - I_1$ such that $I_1 \cup e \in \mathcal{I}$.

Matroids: Independence

Definition

A matroid M is an ordered pair $(E, \mathcal{I})$, where E is a finite set called ground set and $\mathcal{I}$ is a collection of a subsets of E called the independent sets, which satisfies:

(I1) $\emptyset \in \mathcal{I}$.

(I2) If $I \in \mathcal{I}$ and $I' \subseteq I$, then $I' \in \mathcal{I}$.

(I3) If $I_1, I_2 \in \mathcal{I}$ and $|I_1| < |I_2|$, then there exists an element $e \in I_2 - I_1$ such that $I_1 \cup e \in \mathcal{I}$.

Examples:

1. Vector matroids: E is the column set of a matrix (over a field), and independent sets defined by *linearly* independent columns.
2. Graphic matroids: E is the set of edges of an undirected graph, and independent sets are collections of edges containing no *cycle*.

Matroids: other definitions

Definition

Let $M = (E, \mathcal{I})$. Define

- $\mathcal{D} = \mathcal{I}^c$, the collection of dependent sets
- $\mathcal{B}$, the collection of bases, i.e., maximal subsets of E which are independent
- $\mathcal{C}$, the collection of circuits, i.e., minimal subsets of E which are dependent.

Definition

We define a rank function $r : 2^m \to \mathbb{Z}_+$ such that for any $S \subseteq E$, $r(S)$ is given by the cardinality of a maximal independent set contained in S.

Matroids: other definitions

The rank function satisfies the following properties.

$(R1)$ If $X \subseteq E$, then $r(X) \leq |X|$.

$(R2)$ If $X_1 \subseteq X_2 \subseteq E$, then $r(X_1) \leq r(X_2)$.

$(R3)$ If $X_1, X_2 \subseteq E$, then

$$r(X_1 \cup X_2) + r(X_1 \cap X_2) \leq r(X_1) + r(X_2).$$

Claim: this can also be used to define a matroid:
an independent set is then a set with $r(X) = |X|$.

Matroid duality

Theorem

Let M be a matroid on E with a set of bases $\mathcal{B}$. Let $\mathcal{B}^* = \{E - B : B \in B\}$. Then $\mathcal{B}^*$ is the set of bases of a matroid on E. We denote this matroid by M^* and call it the dual of M.

Lemma

If r is the rank function of M, then the rank function of M^* is given by

$$r^*(S) = r(S^c) + |S| - |E|.$$

Matroid representation

Definition

A matroid M is representable over a field F if it is isomorphic to a vector matroid over the field F.
If A is a matrix representing M, we denote $M \cong M[A]$.
A $\mathbb{F}_2$ representable matroid is called a binary matroid.

- The restriction of M to S, is denoted by $M|S$ and means...
- The contraction of M by S, is given by $M^*|S^c$
- A matroid N that is obtained from M by a sequence of restrictions and contractions is called a minor of M.

Matroid representation

Definition

A matroid M is representable over a field F if it is isomorphic to a vector matroid over the field F.
If A is a matrix representing M, we denote $M \cong M[A]$.
A $\mathbb{F}_2$ representable matroid is called a binary matroid.

- The restriction of M to S, is denoted by $M|S$ and means...
- The contraction of M by S, is given by $M^*|S^c$
- A matroid N that is obtained from M by a sequence of restrictions and contractions is called a minor of M.

Theorem (Tutte)

A matroid is binary if and only if it has no minor that is $U_{4,2}$.

$U_{4,2} =$ 4 el. ground set and bases are the 2 el. sets

Polymatroid

A polymatroid is a finite set E equipped with a function $f : 2^m \to \mathbb{R}$, such that

$$\begin{array}{ll} (F1) & f(\emptyset) = 0. \\ (F2) & \text{If } X_1 \subseteq X_2 \subseteq E\text{, then } f(X_1) \leq f(X_2). \\ (F3) & \text{If } X_1, X_2 \subseteq E\text{, then} \\ & f(X_1 \cup X_2) + f(X_1 \cap X_2) \leq f(X_1) + f(X_2). \end{array}$$

Such a f is called a β-rank function.

A matroid is a polymatroid for which f is integer valued and bounded links...

Entropic matroids

Let E a finite set and $X[E] = \{X_i\}_{i \in E}$ be a random vector with distribution P_E.
Let $h(I) := h(X[I])$.

Theorem (Lovász '82, ...)

$h(\cdot)$ is a β-rank function.
Hence, (E, h) is a polymatroid.

Definition

A (poly)matroid M is entropic if $M \cong M[h]$.

Some ref.: Han, Fujishige, Zhang, Matús and Yeung
If $|E| \leq 3$, all matroids are entropic, but ottherwise...

Mutual Information matroids

Let E be a finite set, and $X[E] \stackrel{W}{\to} Y$.

Theorem (Hanly et al. '94, ...)

$\rho(S) = I(X[S]; Y X[S^c])$ *is a* β*-rank function on* E*.*
Hence, (E, ρ) *is a polymatroid.*

Definition

A (poly)matroid M is MAC if $M \cong M[\rho]$
A (poly)matroid M is BUMAC if it is MAC and if $P_1, \ldots, P_m$ are the uniform distributions on $\mathcal{X} = \mathbb{F}^2$.

If $|E| \leq 3$, all matroids are BUMAC, but otherwise...

Polar codes application

Single-user setting:
turn n independent channel uses
into n successive extremal channels,
$\rightarrow$ either perfect $I_U(W) = 1$ or pure noise $I_U(W) = 0$.

Polar codes application

Single-user setting:
turn n independent channel uses
into n successive extremal channels,
$\rightarrow$ either perfect $I_U(W) = 1$ or pure noise $I_U(W) = 0$.

Multi-user MAC setting:
turn n independent channel uses
into n successive extremal MACs ???

Polar codes application

Single-user setting:
turn n independent channel uses
into n successive extremal channels,
$\rightarrow$ either perfect $I_U(W) = 1$ or pure noise $I_U(W) = 0$.

Multi-user MAC setting:
turn n independent channel uses
into n successive extremal MACs ???
$\rightarrow \rho(S) \in \mathbb{Z}_+ \Leftrightarrow$ BUMAC matroids [EA and Telatar '09]

Results

Theorem

A matroid is BUMAC if and only if it is binary.

Theorem

A matroid is BUMAC if and only if it is binary.

Theorem

A BUMAC matroid is "equivalent" to a linear deterministic channel:
if $M = M[W]$, and A represents M, we have

$$I(AX[E];Y) = \mathrm{rank} A,$$

where Y is the output through W.

Example

Let a BUMAC with 2 users: $(X[1], X[2]) \xrightarrow{W} Y$ s.t. $I(X[1]; YX[2]) = I(X[2]; YX[1]) = I(X[1]X[2]; Y) = 1$.

This defines a BUMAC matroid M on $E = \{1, 2\}$ given by the ranks

$$(\emptyset, 1, 2, 12) \xrightarrow{r} (0, 1, 1, 1).$$

Then M is binary (thm 1), and in this case represented by $A = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$

Moreover, we have (thm 2)

$I(A \begin{bmatrix} X[1] \\ X[2] \end{bmatrix}; Y) = 1$, i.e., $I(X[1] + X[2]; Y) = 1$.

proofs