Geometry of the entropy region - II

Laszlo Csirmaz

Central European University, Budapest

IHP, Paris, February 16, 2016

Outline

Polymatroids

Polymatroids are the abstracted view of submodular functions, they are closely related to entropy (*S. Fujishige*, 1978)

Definition

The pair $\langle f, N \rangle$ is a **polymatroid**, if f assigns non-negative real numbers to the non-empty subsets of N such that

1. non-negative: $f(I) \ge 0$;
2. monotonic: if $I \subseteq J \subseteq N$ then $f(I) \le f(J)$;
3. submodular: for subsets $I, J \subseteq N$,

$$f(I) + f(J) \ge f(I \cup J) + f(I \cap J).$$

N is the **ground set**, and f is the **rank function**.

Polymatroids are identified with the **rank function** which is a sequence of length $2^N - 1$ indexed by the non-empty subsets of N.

Entropy expressions

For $I \subseteq N$, δ_I is the vector of length $2^N - 1$ which is 0 everywhere, except at the index I, where it is 1.

Definition (special entropy expressions)

For disjoint non-empty subsets $I, J, K \subseteq N$ define

- $(I \mid J) \stackrel{\text{def}}{=} \delta_{I \cup J} - \delta_I$;
- $(I, J) \stackrel{\text{def}}{=} \delta_I + \delta_J - \delta_{I \cup J}$;
- $(I, J \mid K) \stackrel{\text{def}}{=} \delta_{I \cup K} + \delta_{J \cup K} - \delta_{I \cup J \cup K} - \delta_K$.

Definition

$\delta_I \cdot f$ is the **scalar product** of these two vectors, the value is $f(I)$.

Monotonicity can be expressed as $(I \mid J)f \geq 0$;

Submodularity can be expressed as $(I, J \mid K)f \geq 0$.

Geometry of polymatroids

Claim

Polymatroids on the ground set N form a pointed convex polyhedral cone Γ_N *in the* 2^N-1*-dimensional Euclidean space.*

Proof.

The collection of polymatroids is the intersection of finitely many closed half-spaces corresponding to the *Shannon* inequalities

$$(I\,|\,J)f \ge 0, \quad \text{and} \quad (I,J\,|\,K)f \ge 0.$$

The all-zero point is on the bounding hyperplanes, thus the intersection is a pointed cone. □

Actually, the following set defines all **facets** of the cone:

$$(a\,|\,N-\{a\})\,f \ge 0 \quad \text{for all } a \in N, \text{ and}$$

$$(a,b\,|\,K)\,f \ge 0 \quad \text{for all } a \ne b \in N-K,\ K \subset N.$$

Outline

Entropic polymatroids

Let $\langle \xi_i : i \in N \rangle$ be random variables with some joint distribution. Define f on the non-empty subsets of N by

$$f(J) \stackrel{\text{def}}{=} \boldsymbol{H}(\langle \xi_i : i \in J \rangle),$$

where $\boldsymbol{H}(\cdot)$ is the entropy function. Then f is a polymatroid.

Definition

A polymatroid is **entropic** if it can be written as the entropy of some collection of random variables.
The polymatroid is **almost entropic** (aent) if it is in the closure (in the usual Euclidean topology) of entropic polymatroids.

Remark

The collection of entropic polymatroids is **not** closed when $|N| \geq 3$.

A special entropic polymatroid

Fix the ground set N. For non-empty $S \subseteq N$ define $\langle \boldsymbol{r}_S, N \rangle$ as

$$\boldsymbol{r}_S(I) = 1 \text{ if } I \cap S \neq \emptyset, \text{ and } 0 \text{ otherwise.}$$

Claim

The polymatroid $\boldsymbol{r}_S$ is entropic.

Proof.

Define a probability table with two lines, each with probability 1/2. In the top line all entries are 0; in the bottom line entries are 1 for columns in S (blue cells), and 0 for columns not in S. If $I \cap S = \emptyset$ then only one row remains ($\boldsymbol{H} = 0$), otherwise both rows remain ($\boldsymbol{H} = 1$).

1	2	3	4	...	Prob
0	0	0	0	...	1/2
0	1	1	0	...	1/2

□

More entropic polymatroids

Claim

Let $\lambda > 0$. Then the polymatroid $\lambda \boldsymbol{r}_S$ is entropic.

Proof.

Let ξ be a single random variable with $\boldsymbol{H}(\xi) = \lambda$. Columns in $i \in S$ contain identical copies of ξ, columns not in S contain a fixed value. □

Claim

If f and g are entropic, then so is $f + g$.

Proof.

If $\vec{\xi}$ represents f, and $\vec{\eta}$ represents g, then take *independent* copies of $\vec{\xi}$ and $\vec{\eta}$ (the number of rows will multiply). In this case entropies add up. □

A useful corollary

Claim

Every point in the cone $\mathcal{C} = \left\{ \sum_{S\subseteq N} \lambda_s \boldsymbol{r}_S \;:\; \lambda_S \ge 0 \right\}$ *is entropic. The cone* $\mathcal{C}$ *is full dimensional.*

Proof.

We need to show that the $2^N - 1$ vectors $\boldsymbol{r}_S \in \mathbb{R}^{2^N-1}$ are linearly independent. Do it by induction on N. Fix $a \in N$; J, S are non-empty subsets of $N{-}a$; $\mathrm{rank}(\mathcal{M}) = 2^{N-1} - 1$ by induction, so

	J	a	aJ
$\boldsymbol{r}_S$	$\mathcal{M}$	0	$\mathcal{M}$
$\boldsymbol{r}_a$	0	1	$1\ldots1$
$\boldsymbol{r}_{aS}$	$\mathcal{M}$	1	$1\ldots1$

$$\Rightarrow \begin{array}{|ccc|} \hline \mathcal{M} & 0 & \mathcal{M} \\ 0 & 1 & 1 \\ \mathcal{M} & 0 & 0 \\ \hline \end{array} \Rightarrow \begin{array}{|ccc|} \hline 0 & 0 & \mathcal{M} \\ 0 & 1 & 0 \\ \mathcal{M} & 0 & 0 \\ \hline \end{array}$$

□

Outline

The set of entropic polymatroids is **almost** closed

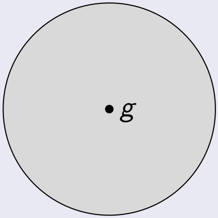
Definition

Γ^*_N is the collection of entropic polymatroids, and $\bar{\Gamma}^*_N$ is the closure: the pointwise limits of entropic polymatroids.

Theorem (Yeung, Matus)

Internal points of $\bar{\Gamma}^*_N$ are entropic: $int(\bar{\Gamma}^*_N) \subset \Gamma^*_N$.

Proof.



Let g be an internal point of $\bar{\Gamma}^*_N$.
$\mathcal{C}$ is the cone from the previous page.

□

The set of entropic polymatroids is **almost** closed

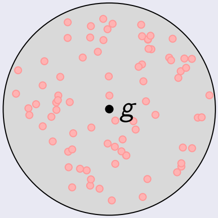
Definition

Γ_N^* is the collection of entropic polymatroids, and $\bar{\Gamma}_N^*$ is the closure: the pointwise limits of entropic polymatroids.

Theorem (Yeung, Matus)

Internal points of $\bar{\Gamma}_N^*$ *are entropic:* $int(\bar{\Gamma}_N^*) \subset \Gamma_N^*$.

Proof.



Let g be an internal point of $\bar{\Gamma}_N^*$.
$\mathcal{C}$ is the cone from the previous page.
Entropic points are **dense** around g.

□

The set of entropic polymatroids is **almost** closed

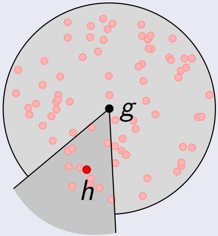
Definition

Γ_N^* is the collection of entropic polymatroids, and $\bar{\Gamma}_N^*$ is the closure: the pointwise limits of entropic polymatroids.

Theorem (Yeung, Matus)

Internal points of $\bar{\Gamma}_N^*$ *are entropic:* $int(\bar{\Gamma}_N^*) \subset \Gamma_N^*$.

Proof.



Let g be an internal point of $\bar{\Gamma}_N^*$.
$\mathcal{C}$ is the cone from the previous page.
Entropic points are **dense** around g.
Pick an entropic point h inside $g - \mathcal{C}$.

□

The set of entropic polymatroids is **almost** closed

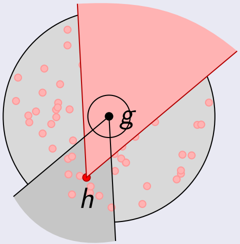
Definition

Γ_N^* is the collection of entropic polymatroids, and $\bar{\Gamma}_N^*$ is the closure: the pointwise limits of entropic polymatroids.

Theorem (Yeung, Matus)

Internal points of $\bar{\Gamma}_N^*$ are entropic: $int(\bar{\Gamma}_N^*) \subset \Gamma_N^*$.

Proof.



Let g be an internal point of $\bar{\Gamma}_N^*$.
$\mathcal{C}$ is the cone from the previous page.
Entropic points are **dense** around g.
Pick an entropic point h inside $g - \mathcal{C}$.
Every point in $h + \mathcal{C}$ is entropic,
g is an internal point of $h + \mathcal{C}$ $\Rightarrow$
a neighbor of g (and g) is entropic. □

Outline

Methods

Presently known methods to get new entropy inequalities are:

1. Zhang–Yeung method (1998)
2. Makarychev *et al.* technique (2002)
3. Matúš' polymatroid convolution (2007)
4. Maximum entropy extension (2014)

Equivalence of #1 and #2 for **balanced** inequalities (see later) was shown by Tarik Kaced (2013).

Research problem

Show that methods #3 and #4 are actually **stronger** than the other two.

Zhang–Yeung method

Zhang – Yeung, 1998

$\vec{X}$, $\vec{Y}$, Z are (collections of) random variables.

Copy-variable method

(A) If we have an information inequality of the form

$$u(\vec{X},\vec{Y}) + v(\vec{Y},Z) + \lambda \boldsymbol{I}(Z,\vec{X}\,|\,\vec{Y}) \ge 0$$

for some $\lambda \ge 0$,

(B) then the following stronger inequality also holds:

$$u(\vec{X},\vec{Y}) + v(\vec{Y},Z) \ge 0$$

Example: This is a **Shannon** inequality (checked by itip*):

$$\boldsymbol{I}(a,b) \le \boldsymbol{I}(a,b\,|\,c) + \boldsymbol{I}(a,b\,|\,d) + \boldsymbol{I}(c,d) + \\ + \boldsymbol{I}(a,b\,|\,z) + \boldsymbol{I}(z,a\,|\,b) + \boldsymbol{I}(z,b\,|\,a) + 3\boldsymbol{I}(z,cd\,|\,ab)$$

*http://xitip.epfl.ch, or https://github.com/lcsirmaz/minitip

Zhang–Yeung method

Zhang – Yeung, 1998

$\vec{X}$, $\vec{Y}$, Z are (collections of) random variables.

Copy-variable method

(A) If we have an information inequality of the form

$$u(\vec{X},\vec{Y}) + v(\vec{Y},Z) + \lambda \boldsymbol{I}(Z,\vec{X} \mid \vec{Y}) \geq 0$$

for some $\lambda \geq 0$,

(B) then the following stronger inequality also holds:

$$u(\vec{X},\vec{Y}) + v(\vec{Y},Z) \geq 0$$

Example: This is a **new** entropy inequality:

$$\boldsymbol{I}(a,b) \leq \boldsymbol{I}(a,b \mid c) + \boldsymbol{I}(a,b \mid d) + \boldsymbol{I}(c,d) + \\ + \boldsymbol{I}(a,b \mid z) + \boldsymbol{I}(z,a \mid b) + \boldsymbol{I}(z,b \mid a)$$

*`http://xitip.epfl.ch`, or `https://github.com/lcsirmaz/minitip`

Makarychev *et al.* technique

Makarychev – Makarychev – Romashchenko – Vereschagin, 2002

$\vec{X}$, $\vec{Y}$, Z are (collections of) random variables.

Makarychev technique

(A) If we have an information inequality of the form

$$u(\vec{X}, \vec{Y}) + v(\vec{Y}, Z) \geq 0,$$

where $v()$ is a linear combination of entropies,

(B) then the following inequality also holds:

$$u(\vec{X}, \vec{Y}) + v(\vec{Y}, Z) - \lambda \boldsymbol{H}(Z \,|\, \vec{Y}) \geq 0,$$

where λ is the sum of coefficients in v involving Z.

Example: This is a **Shannon** inequality:

$$\boldsymbol{H}(z) \leq 2\boldsymbol{H}(z \,|\, a) + 2\boldsymbol{H}(z \,|\, b) + \boldsymbol{I}(a, b \,|\, c) + \boldsymbol{I}(a, b \,|\, d) + \boldsymbol{I}(c, d).$$

Makarychev *et al.* technique

Makarychev – Makarychev – Romashchenko – Vereschagin, 2002

$\vec{X}$, $\vec{Y}$, Z are (collections of) random variables.

Makarychev technique

(A) If we have an information inequality of the form

$$u(\vec{X},\vec{Y}) + v(\vec{Y},Z) \ge 0,$$

where $v()$ is a linear combination of entropies,

(B) then the following inequality also holds:

$$u(\vec{X},\vec{Y}) + v(\vec{Y},Z) - \lambda \mathbf{H}(Z \mid \vec{Y}) \ge 0,$$

where λ is the sum of coefficients in v involving Z.

Example: This is a **new** entropy inequality:

$$\mathbf{H}(z) \le 2\mathbf{H}(z \mid a) + 2\mathbf{H}(z \mid b) - 3\mathbf{H}(z \mid ab) + \mathbf{I}(a,b \mid c) + \mathbf{I}(a,b \mid d) + \mathbf{I}(c,d).$$

Frantisek Matúš' convolution method

Matúš, 2007

Let $\langle g, N \rangle$ be a polymatroid, $a \in N$, and $0 < t$. Define the polymatroids $g{\downarrow}^a_t$ and $g{\uparrow}^a_t$ as

$$\begin{aligned} g{\downarrow}^a_t(J) &= \min\{g(aJ) - t, g(J)\}, \\ g{\uparrow}^a_t(J) &= \min\{g(aJ), g(J) + t\}, \end{aligned} \qquad \text{for all } J \subseteq N.$$

Matúš method

(A) If we have an entropy inequality $u(g) \geq 0$

(B) then for all $a \in N$ and $0 \leq t \leq g(a)$, $u(g{\downarrow}^a_t) \geq 0$ is also an information inequality;

(B') for all $a \in N$, $u(g{\uparrow}^a_t) \geq 0$ is also an information inequality.

Example: This is a **Shannon** inequality:

$$\begin{aligned} \boldsymbol{H}(z) \leq{} & 2\boldsymbol{H}(z \mid a) + 2\boldsymbol{H}(z \mid b) \\ & + \boldsymbol{I}(a, b \mid c) + \boldsymbol{I}(a, b \mid d) + \boldsymbol{I}(c, d). \end{aligned}$$

Frantisek Matúš' convolution method

Matúš, 2007

Let $\langle g, N\rangle$ be a polymatroid, $a \in N$, and $0 < t$. Define the polymatroids $g{\downarrow}_t^a$ and $g{\uparrow}_t^a$ as

$$\begin{aligned} g{\downarrow}_t^a(J) &= \min\{g(aJ) - t, g(J)\}, \\ g{\uparrow}_t^a(J) &= \min\{g(aJ), g(J) + t\}, \end{aligned} \qquad \text{for all } J \subseteq N.$$

Matúš method

(A) If we have an entropy inequality $u(g) \geq 0$

(B) then for all $a \in N$ and $0 \leq t \leq g(a)$, $u(g{\downarrow}_t^a) \geq 0$ is also an information inequality;

(B') for all $a \in N$, $u(g{\uparrow}_t^a) \geq 0$ is also an information inequality.

Example: This is a **new** inequality using $g{\downarrow}_t^z$ with $t \leq \boldsymbol{H}(z \mid ab)$:

$$\begin{aligned} \boldsymbol{H}(z) - t &\leq 2\boldsymbol{H}(z \mid a) - 2t + 2\boldsymbol{H}(z \mid b) - 2t \\ &\quad + \boldsymbol{I}(a, b \mid c) + \boldsymbol{I}(a, b \mid d) + \boldsymbol{I}(c, d). \end{aligned}$$

Maximum entropy method

Csirmaz – Matus, 2013

$\vec{U}$ is a collection of random variables, $\vec{X}_k$, $\vec{Y}_k$ $\vec{Z}_k$ are subsets of $\vec{U}$.

Maximum entropy method

(A) Suppose we have an information inequality of the form

$$u(\vec{U}) + v(\,(\vec{X}_1, \vec{Y}_1 \,|\, \vec{Z}_1), (\vec{X}_2, \vec{Y}_2 \,|\, \vec{Z}_2) \dots) \geq 0,$$

where no term in $u()$ intersects both X_k and Y_k at the same time.

(B) Then the following stronger inequality also holds:

$$u(\vec{U}) + v(0, 0, \dots) \geq 0.$$

Example: This is a **Shannon** inequality:

$$\begin{aligned}(u,v) \leq &(u,v|y) + (u,v|t) + (y,v|x) + (t,v|z) + \\ &+ (u,y|v) + (u,z|v) + (v,z|u) + (x,u) + 4(xy,zt|uv)\end{aligned}$$

Maximum entropy method

Csirmaz – Matus, 2013

$\vec{U}$ is a collection of random variables, $\vec{X}_k$, $\vec{Y}_k$ $\vec{Z}_k$ are subsets of $\vec{U}$.

Maximum entropy method

(A) Suppose we have an information inequality of the form

$$u(\vec{U}) + v(\,(\vec{X}_1, \vec{Y}_1 \,|\, \vec{Z}_1), (\vec{X}_2, \vec{Y}_2 \,|\, \vec{Z}_2) \dots) \ge 0,$$

where no term in $u()$ intersects both X_k and Y_k at the same time.

(B) Then the following stronger inequality also holds:

$$u(\vec{U}) + v(0, 0, \dots) \ge 0.$$

Example: This is a **new** entropy inequality:

$$\begin{aligned}(u, v) \le &(u, v|y) + (u, v|t) + (y, v|x) + (t, v|z) + \\ &+ (u, y|v) + (u, z|v) + (v, z|u) + (x, u)\end{aligned}$$