A New Entropy for Hypergraphs

Isabelle Bloch[1], Alain Bretto[2]
1. LTCI, Télécom ParisTech, Université Paris Saclay, Paris, France
2. GREYC CNRS UMR 6072, NormandieUnicaen, Caen, France
isabelle.bloch@telecom-paristech.fr, alain.bretto@unicaen.fr

Motivation

- Representation of structured information as hypergraphs.
- Entropy measures.
- Fine grained analysis of the structure and complexity of hypergraphs.
- $\Rightarrow$ Entropy vector: entropy values of all partial hypergraphs.

Notations and background

- Hypergraph $H = (V, E = \{e_i, i = 1...m, e_i \subseteq V\}), |V| = n, |E| = m$.
- Incidence matrix I, $L(H) = I(H)I(H)^t = ((|e_i \cap e_j|))_{i,j\in\{1...m\}}$.
- Normalized eigenvalues of $L(H)$: $\mu_i, i = 1...m$.
- Entropy $S(H) = -\sum_{i=1}^{m} \mu_i \log_2(\mu_i)$.
- Partial hypergraph $H' = (V', \{e_j, j \in J\}), J \subseteq \{1...m\}, \cup_{j\in J} e_j \subseteq V' \subseteq V$ (here $V' = V$). Notation: $H' \leq H$.

Main definition: entropy vector

For $i \leq m$:

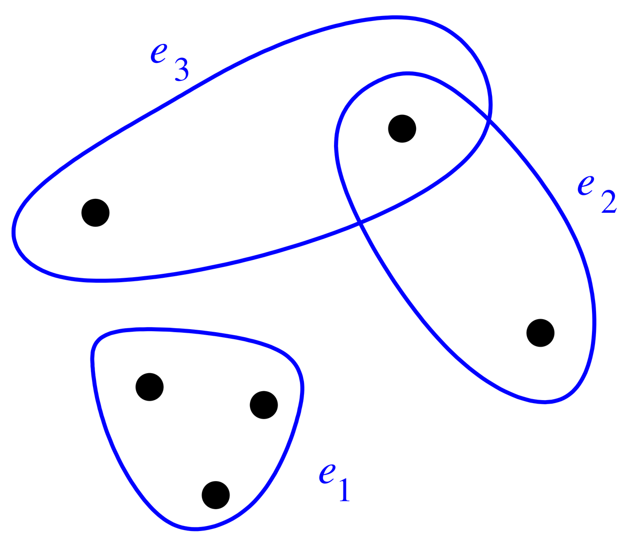
$$SE_i(H) = \{S(H_i) \mid H_i = (V, E_i), H_i \leq H, |E_i| = i\}$$

= set of entropy values of all partial hypergraphs of H whose set of hyperedges has cardinality i, arranged in increasing order.
Entropy vector of the hypergraph H:

$$\boxed{SE(H) = (SE_1(H), SE_2(H), \dots SE_m(H))}$$

with $2^m - 1$ coordinates.

A simple example



- SE_1: three partial hypergraphs containing one hyperedge (e_1, e_2 and e_3, respectivly).

  $$SE_1 = (0, 0, 0)$$

- SE_2: three partial hypergraphs containing two hyperedges.
  (e_1, e_2): $L = \begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix}$, eigenvalues = 2 and 3, $s_1 = -\frac{2}{5}\log_2\frac{2}{5} - \frac{3}{5}\log_2\frac{3}{5} \simeq 0.97$.
  (e_1, e_3): same reasoning.
  (e_2, e_3): $L = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$, eigenvalues = 1 and 3, $s_2 = -\frac{1}{4}\log_2\frac{1}{4} - \frac{3}{4}\log_2\frac{3}{4} \simeq 0.81$.

  $$SE_2 = (s_2, s_1, s_1) \simeq (0.81, 0.97, 0.97)$$

- SE_3: one partial hypergraph containing three hyperedges, i.e. H. $L = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 1 & 2 \end{pmatrix}$, eigenvalues = 1, 3 and 3, $s_3 = -\frac{1}{7}\log_2\frac{1}{7} - 2\frac{3}{7}\log_2\frac{3}{7} \simeq 1.45$.

  $$SE_3 = (s_3) \simeq (1.45)$$

- Entropy vector:

  $$SE(H) = (0, 0, 0, s_2, s_1, s_1, s_3) \simeq (0, 0, 0, 0.81, 0.97, 0.97, 1.45)$$

Some properties

- $S(H) = 0$ if and only if $|E| = 1$.
- $S(H) = \log_2(n) - \log_2(r(H)) = \log_2(m)$, where $r(H) = \frac{n}{m}$ (rank of H), if and only if H is uniform (i.e. $\forall e \in E$, $|e| = r(H)$) and the intersection of any two distinct hyperedges is empty (i.e. for all e, e' in E such that $e \neq e'$, $|e \cap e'| = 0$).
- Two isomorphic hypergraphs have the same entropy vectors.
- Lattice structures:
  - on $\mathcal{H}$ (isomorphim classes of hypergraphs) for the partial ordering defined by the subhypergraph relation $\leq_f$,
  - on $SE_{\mathcal{H}} = \{SE(H) \mid H \in \mathcal{H}\}$ for Pareto partial ordering on vectors.
- $H' \leq_f H \Rightarrow SE(H') \leq SE(H)$.

On going work

- Reducing the complexity ($|SE(H)| = 2^m - 1$)
  - by discarding two small or two large partial hypergraphs;
  - by approximating the computation of entropy;
  - by considering only the leading principal matrices ($m - 1$ instead of $2^m - 1$) after sorting the hyperedges by increasing cardinality.
- Relation between entropy and Zeta function:

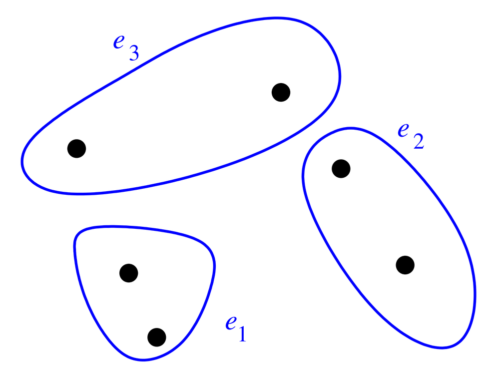
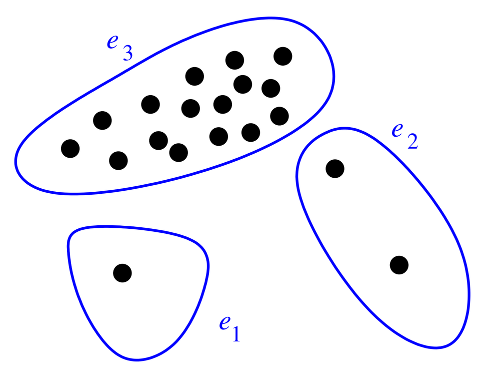
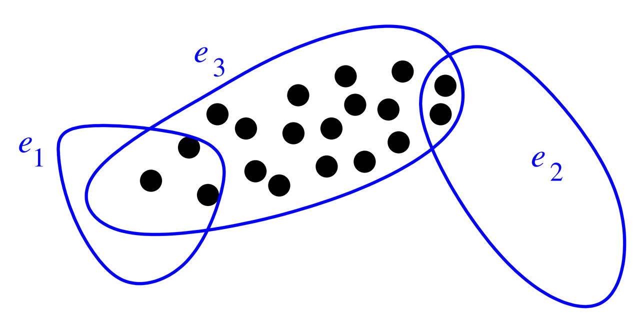
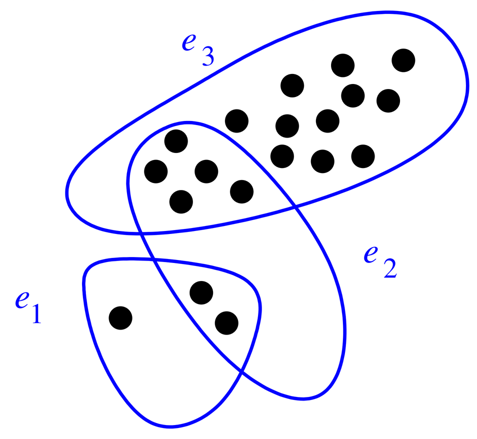
  $$\zeta_H(s) = Tr(\mathcal{L}(H)^{-s}) = \sum_{i=1, \mu_i \neq 0}^{m} \mu_i^{-s}$$

  where $\mathcal{L}(H) = \frac{L(H)}{Tr(L(H))}$.
  First results:

  $$\zeta'_H(-1) = \ln(2)S(H),\ \zeta'_H(0) = -\ln(\det(\mathcal{L}(H))),\ \zeta_H(-s) = e^{(1-s)R_s(H)}$$

  where $R_s(H) = \frac{1}{1-s}\ln(\sum_{i=1}^{m}\mu_i^s)$ (Renyi entropy).
- Illustrations and examples.

A few illustrations

H		
$SE(H)$	$(0, 0, 0, 1, 1, 1, 1.58 = \log_2(3))$	$(0, 0, 0, 0.32, 0.50, 0.92, 0.77)$
H		
$SE(H)$	$(0, 0, 0, 0.42, 0.50, 0.85, 0.83)$	$(0, 0, 0, 0.63, 0.74, 0.75, 1.11)$

References

Bai, L., Escolano, F., Hancock, E.R.: Depth-based hypergraph complexity traces from directed line graphs. PR (2016).
Berge, C.: Hypergraphs. Elsevier Science Publisher (1989).
Bloch, I., Bretto, A.: Mathematical morphology on hypergraphs, application to similarity and positive kernel. CVIU (2013).
Bretto, A.: Hypergraph Theory: an Introduction. Springer-Verlag (2013).
Shannon, C.E.: A mathematical theory of communication. Bell System Technical Journal 27 (1948).