Statistical analysis of network data lecture 11

Sofia Olhede

EPFL

November 27, 2024

1 Hypergraphs

Hypergraphs I

- The basic idea that lets us define a hypergraph is to generalize a graph.
- Any subset of nodes may form an edge, not only pairs.
- Let $V = \{1, \ldots, n\}$ be the vertex set, and let $\mathcal{D} = \{D_1, D_2, \ldots, D_m\}$ be a family of subsets of V.
- The pair $\mathcal{H} = \{V, \mathcal{D}\}$ is a hypergraph with vertex set V where we write $V(\mathcal{H})$.
- The edge set $\mathcal{D}$ is sometimes written by $\mathcal{D}(\mathcal{H})$.
- $n = |V|$ is the order of the hypergraphs.
- The number of edges are usually written as $m(\mathcal{H})$.

Hypergraphs II

- We include the cases V and $\mathcal{D}$ are empty.
- If the hypergraph has no multiple or included edges then it is simple.
- An included edge is an edge that is a subset of other edges.
- Two vertices are adjacent if there is an edge that connects them, or an hyperedge.
- Two adjacent vertices are neighbours.
- All the adjacent vertices of i are called the neighbourhood of i.

Hypergraphs III

- If a vertex i belongs to an edge $\mathcal{D}_j \in \mathcal{D}$ then i and $\mathcal{D}_j$ are **incident**.
- We write $\mathscr{D}(i)$ for $i \in V$ for all edges containing vertex i.
- Thus the number $|\mathscr{D}(i)|$ is the degree of vertex i.
- In contrast, $|\mathcal{D}_j|$ is the cardinality of edge $\mathcal{D}_j$ (this was always 2 before).
- We define the **maximum degree** of hypergraph $\mathcal{H}$:

$$\Delta(\mathcal{H}) = \max_{i \in V} |\mathscr{D}(i)|$$

- A hypergraph in which all vertices have degree $k \geq 0$ is known as a k-**regular hypergraph**.

Hypergraphs IV

- A hypergraph in which all edges have the same cardinality $l \geq 0$ is known as a *l*-uniform hypergraph.
- The rank of a hypergraph $\mathcal{H}$ is

$$r(\mathcal{H}) = \max_{D \in \mathcal{D}} |D|.$$

- An edge of a hypergraph which has no vertices is called an empty edge.
- The cardinality of an empty edge is zero.
- A vertex of a hypergraph which is incident to no edges is an isolated vertex.
- The degree of an isolated vertex is trivially zero.

Hypergraphs V

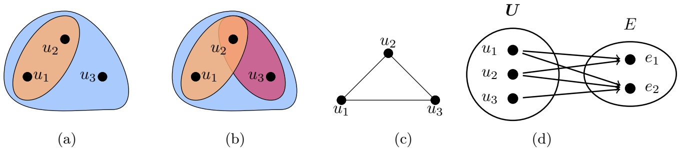


Figure 1: Figures 1a and 1b depict two possible hypergraphs, where a node belongs to a hyperedge if it lies within the associated shaded region. Figure 1c is the graph obtained by replacing hyperedges in Figure 1a, or equivalently in Figure 1b, by cliques. The hypergraphs in 1a and 1b cannot be recovered from 1c. Figure 1d presents the hypergraph relationships in Figure 1a as a bipartite graph, where an edge from a population node to a hyperedge node indicates membership of a hyperedge.

From Turnball et al. (2019).

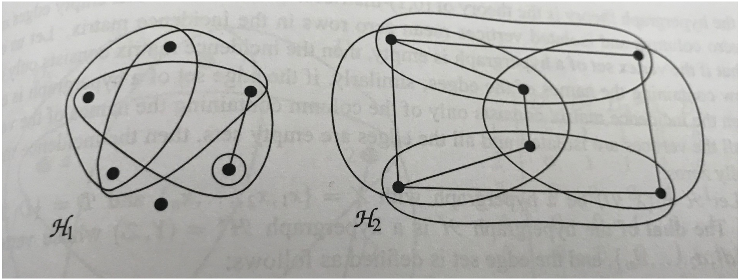
Hypergraphs VI

- An edge of cardinality one is a **singleton**. A vertex of degree one is a **pendant vertex**.
- Thus a simple hypergraph $\mathcal{H}$ with $|D_i| = 2$, for every $D_i \in \mathcal{D}$ is a simple graph.
- The simple hypergraphs $\mathcal{H}_1$ and $\mathcal{H}_2$ are **isomorphic** if there is a one-to-one correspondence between vertex sets such that any subset of vertices form an edge in $\mathcal{H}_1$ if and only if the corresponding subset of vertices form an edge in $\mathcal{H}_2$.

Hypergraphs VII

- We can imagine examples from:
- **Computer Science:** vertices are computers in a network. Edges are subsets of computers of the same manufacturers, there is one subset for each manufacturer.
- **Computer Science:** vertices are records in a relational data base, the edges are the subsets of records for which chosen attributes are true. A subset is defined for each attribute.
- **Genetics:** the vertices are species, the edges are sets of species with common hereditary property,
- **Sociology:** vertices are employees working for a given company, edges are given interests of the employees,
- **Healthcare:** vertices are different illnesses, edges are the groups of illnesses that are treated by the same medication, one hyperedge for each medication like aspirin.

Hypergraphs VIII



From Voloshin (2009).

- a) Find order & number of edges.
- b) Find included edges (if any).
- c) Find multiple edges (if any).
- d) Is the hypergraph simple?
- e) For any pair of vertices, determine if they are adjacent.
- f) For every vertex, determine degree and neighbourhood.

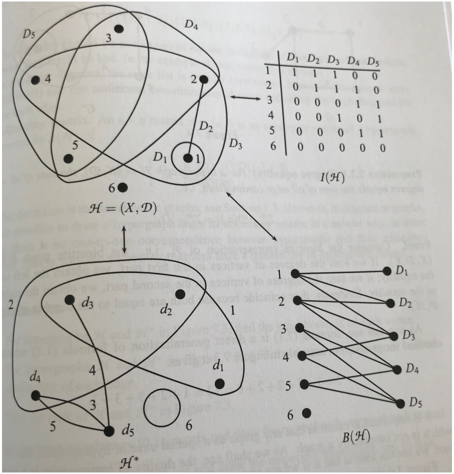
Hypergraphs IX

- The **incidence matrix** of a hypergraph $\mathcal{H} = (V, \mathcal{D})$ is a matrix $I(\mathcal{H})$ with n rows and m columns for each each of the edges.
- The incidence matrix takes the form

$$I_{ij}(\mathcal{H}) = \begin{cases} 1 & \text{if} \quad i \in \mathcal{D}_j \\ 0 & \text{if} \quad i \notin \mathcal{D}_j \end{cases}$$

- Thus, just like simple graphs, is the theory of symmetric $(0,1)$-matrices, hypergraphs are also the theory of $(0,1)$-matrices
- By convention, if the vertex is empty then then the incidence matrix only consistent of a row with the names of edges.
- If the edge set of hypergraphs is empty then the incidence matrix is only a column containing the names of vertices.

Hypergraphs X



From Voloshin (2009).

Hypergraphs XI

- The dual of a hypergraph $\mathcal{H}$ is a hypergraph $\mathcal{H}^*$ whose incidence matrix is its transpose, e.g.

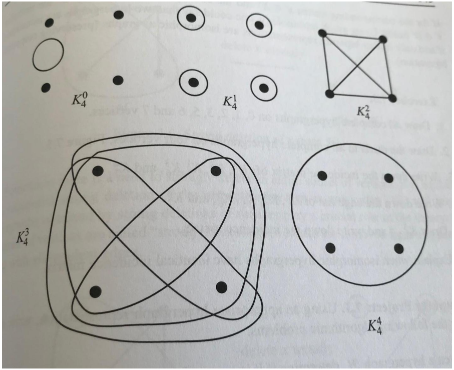
$$I(\mathcal{H}^*) = I^T(\mathcal{H})$$

- This means it has m vertices, and n edges.
- For a hypergraph $\mathcal{H}$ we define the **bipartite representation** of $\mathcal{H}$. This is the bipartite graph

$$\mathcal{B}(\mathcal{H}) = (V, \mathcal{D}; E)$$

  The vertex set in the bipartite graph is $V \cup \mathcal{D}$ where V us the left part, $\mathcal{D}$ is the right part and E is the edge set.
- A vertex $i \in V$ is adjacent to vertex $D \in \mathcal{D}$ in $\mathcal{B}(\mathcal{H})$ if and only if vertex $i \in V$ is incident to edge $D \in \mathcal{D}$.
- In this way any bipartite graph is a bipartite representation of a hypergraph.

Hypergraphs XII



- The complete hypergraphs. The complete r-uniform hypergraph (for $0 \leq r \leq n$) is the simple hypergraph $K_n^r = (V, \mathcal{D})$ such that $|V| = n$ and $\mathcal{D}(K_n^r)$ coincides with all the r-subsets of V.

Hypergraphs XIII

- This means a complete graph on n vertices is a complete 2-uniform hypergraph K_n^2 also denoted as K_n.
- Among all complete hypergraphs on four vertices it follows only K_4^2 is a simple graph.
- All complete hypergraphs are r-uniform and $\begin{pmatrix} n-1 \\ r-1 \end{pmatrix}$-regular hypergraphs.
- **Paths and Cycles.** In a hypergraph $\mathcal{H} = (V, \mathcal{D})$ an alternating sequence

$$\mu = v_0 D_0 v_1 D_1 v_2 \dots v_{t-1} D_{t-1} v_t$$

of distinct nodes $v_0, v_1, v_2, \dots, v_{t-1}$ and distinct edges $D_0, D_1, D_2, \dots, D_{t-1}$ satisfying $v_i, v_{i+1} \in D_i, i = 0, 1, \dots, t-1$ called a **path**. This is connecting v_0 and v_t a (v_0, v_t)-path, which corresponds to a cycle if $v_t = v_0$.

Hypergraphs XIV

- **Connected hypergraphs.** The hypergraph $\mathcal{H} = (V, \mathcal{D})$ is referred to as **connected** if for any pair of vertices there is a path between them.
- If $\mathcal{H}$ is not connected then it consists of two or more connected components each are a connected hypergraph.
- **Bipartite hypergraphs.** A hypergraph $\mathcal{H} = \{V, \mathcal{D}\}$ is bipartite if its vertex set can be split into two parts that are disjoint, V_1 and V_2 say, in such a way that each hyperedge of cardinality equal or bigger than 2 contains vertices from both V_1 and V_2.
- The consequence is that there is no hyperedge inside V_1 or fully inside V_2.
- A **complete** r**-partite hypergraph** is an r-uniform hypergraph $\mathcal{H} = \{V, \mathcal{D}\}$ such that V can be partitioned into r non-empty parts.
- In this case each edge contains precisely one vertex from each part, and all such subsets form $\mathcal{D}$.

Hypergraphs XV

- Two hypergraphs $\mathcal{H}_1$ and $\mathcal{H}_2$ are isomorphic written as $\mathcal{H}_1 \equiv \mathcal{H}_2$ if there is a one-to-one correspondence between V_1 and V_2 and a one-to-one correspondence between $\mathcal{D}_1$ and $\mathcal{D}_2$ such that for each vertex $i \in V$ and for every edge $D \in \mathcal{D}$ we have that $x \in D \Leftrightarrow i' \in V'$ and the corresponding edge $D' \in \mathcal{D}_2$ the inclusion $i' \in D'$ holds.
- There are some basic hypergraph operations that can be used on hypergraphs.
- **Strong deletion of a vertex.** A strong deletion of i from $\mathcal{H}$ is removing all the edges that contain i and then taking $V\backslash\{i\}$.
- **Weak deletion of a vertex.** A weak deletion of i from $\mathcal{H}$ is taking $V\backslash\{i\}$ and from each hyperedge of $\mathscr{D}(i)$.
- Similarly edges can be strongly and weakly deleted.

Hypergraphs XVI

- By using strong and weak deletions we create subhypergraphs.
- **Subhypergraphs.** Assume we have a hypergraph $\mathcal{H} = \{V, \mathcal{D}\}$. Any hypergraph $\mathcal{H}' = \{V', \mathcal{D}'\}$ such that $V' \subset V$ and $\mathcal{D}' \subset \mathcal{D}$ is a subhypergraph of $\mathcal{H}$.
- $\mathcal{H}'$ can be obtained from $\mathcal{H}$ by strong deletion.
- **Induced subhypergraphs.** A hypergraph $\mathcal{H}' = \{V', \mathcal{D}'\}$ is an induced subhypergraph of a hypergraph $\mathcal{H} = \{V, \mathcal{D}\}$ if $V' \subset V$ and all edges of $\mathcal{H}$ completely contained in V' form the family $\mathcal{D}'$. Subhypergraph $\mathcal{H}'$ may be obtained by deleting $V \backslash V'$.
- **Partial subhypergraphs.** For a hypergraph $\mathcal{H} = \{V, \mathcal{D}\}$ any subhypergraph $\mathcal{H}' \subset \mathcal{H}$ such that $\mathcal{H}' = (V, \mathcal{D}')$ is called a partial subhypergraphs. Partial subhypergraphs have the same vertex set as the hypergraph itself and may be obtained only by weak deletion.

Hypergraphs XVII

- **Strongly independent sets.** For a hypergraph $\mathcal{H} = \{V, \mathcal{D}\}$ a subset of vertices $S \subset V$ is a strongly independent set if $|S \cap D| \leq 1$ for every hyperedge D.
- **Transversals.** A set $T \subset V$ is a transversal of a hypergraph $\mathcal{H} = \{V, \mathcal{D}\}$ if $|T \cap D| \geq 1$ for every D. The cardinality of a minimum transversal is denoted by $\tau(\mathcal{H})$.