Key Methods of Hypergraph Analysis
Day 1:
Basics of Hypergraphs

Moses Boudourides

Moses.Boudourides@gmail.com

instats Seminar

Sunday, March 9, 2025, from 5:00 PM to 6:30 PM UTC

Table of Contents

Hypergraphs

Hypergraphs are generalizations of graphs, in which edges consist of any number of vertices.

Conventions in Terminology and Notation:

The following terms are often used interchangeably due to common conventions, making them effectively synonymous for our purposes:

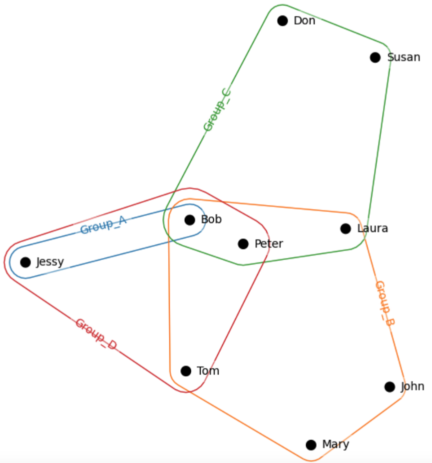
- ***graph–network*** and ***hypergraph–hypernetwork***
- ***vertex–node***
- ***edge–link*** and ***hyperedge–hyperlink***
- ***All graphs, bipartite graphs and hypergraphs are denoted by G, vertices by v, V and edges or hyperedges by e, E.***

Formal Definition of (Undirected) Hypergraphs

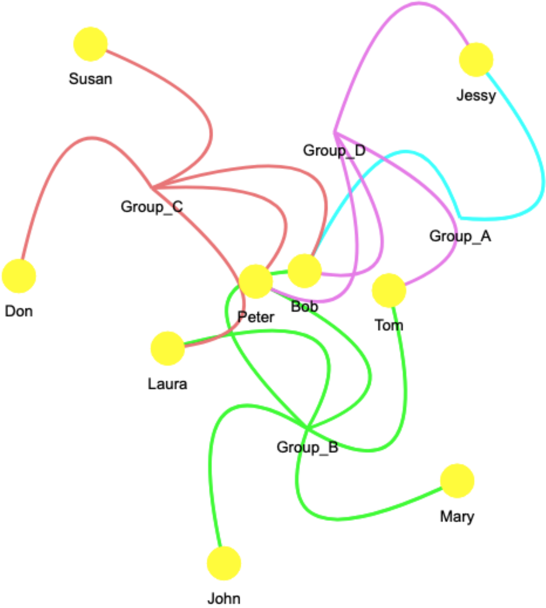
A **hypergraph** $G = (V, E)$ is a pair of two finite sets V and E. The elements of V are called *vertices* and the elements of E, called *hyperedges*, are subsets of vertices (i.e., for all $e \in E$, $e \subseteq V$).

Table and Euler Plot of a Hypergraph

hyperedge	vertices
Group_A	Bob, Jessy
Group_B	Bob, John, Laura, Mary, Peter, Tom
Group_C	Bob, Don, Laura, Peter, Susan
Group_D	Bob, Jessy, Peter, Tom



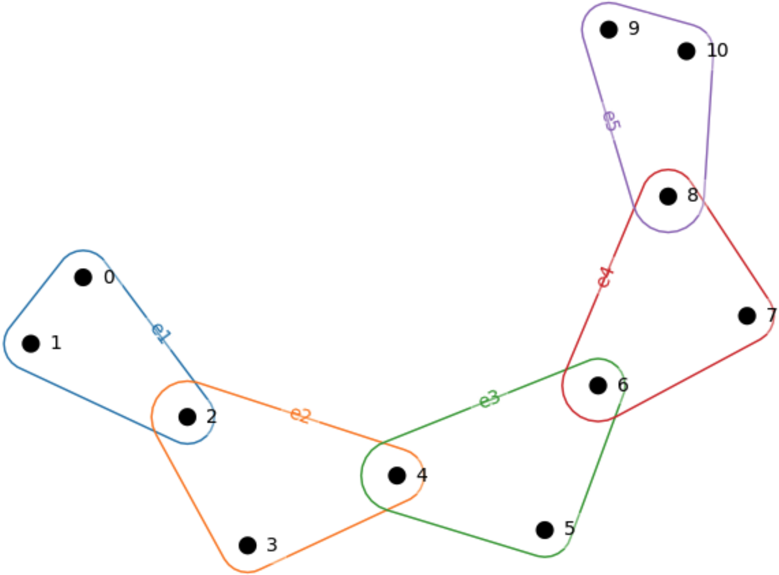
Bipartite Plot of a Hypergraph



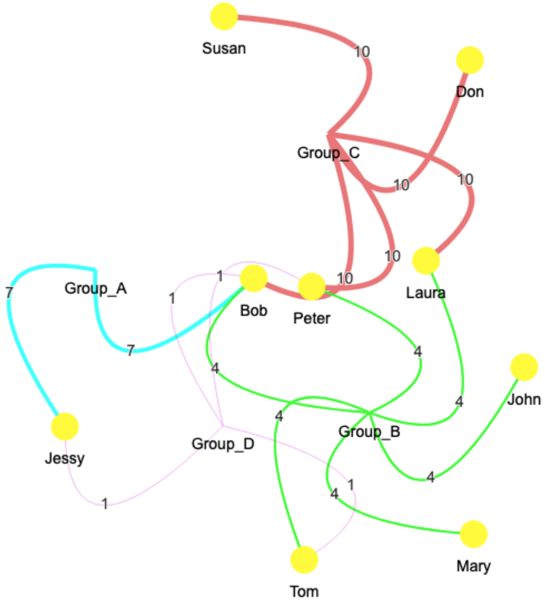
Particular Hypergraphs

- **Empty hypergraph**: $V = \varnothing, E = \varnothing$.
- **Trivial hypergraph**: $V \neq \varnothing, E = \varnothing$.
- **Multiple hypergraph**: E is a multiset (i.e., hyperedges are allowed to repeat).
- **Simple hypergraph**: Hyperedges are not allowed to repeat, i.e., for all all $e_i, e_j \in E$, if $e_i \subseteq e_j$, then $i = j$.
- **Weighted hypergraph**: each hyperedge $e \in E$ is assigned with a weight $w(e)$, i.e., $G = (V, E, \omega)$, where (V, E) is a hypergraph and $\omega : E \to \mathbb{R}$ is the edge weight function.
- **Hypergraph with edge-dependent vertex weights (EDVWs)**: each hyperedge $e \in E$ is assigned with a weight $w(e)$ and each vertex $v \in e$ in the hyperedge is assigned with a weight $\gamma_e(v)$, which depends on that hyperedge, i.e., $G = (V, E, \omega, \gamma)$, where (V, E) is a hypergraph, $\omega : E \to \mathbb{R}$ is the edge weight function and $\gamma : V \times E \to \mathbb{R}$ is the edge-dependent vertex weight function.
- **Linear hypergraph**: A simple hypergraph such that all pairs of hyperedges share at most one vertex.

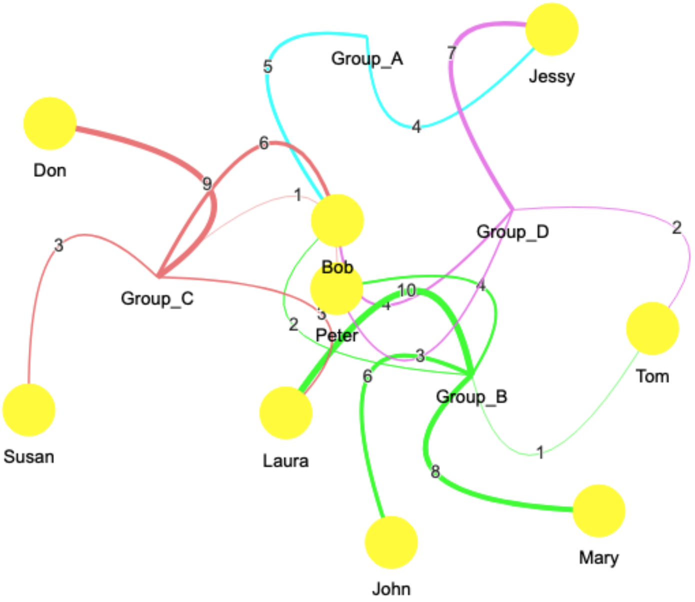
Euler Plot of a Linear Hypergraph



Bipartite Plot of a Multiple (Weighted) Hypergraph



Bipartite Plot of a Hypergraph with EDVWs



Incidence and Adjacency in Hypergraphs

In what follows, unless stated otherwise, hypergraphs are assumed to have non-empty set of vertices, non-empty set of hyperedges, no empty hyperedges, and to be simple hypergraphs.

Let $G = (V, E)$ a hypergraph, where V is a finite set of vertices and E is a finite collection of subsets of V.

- Given a hyperedge $e \in E$, the set of vertices in e is still denoted as $e \subseteq V$ (notice that we are using the same symbol for e as an element of E and as a subset of V).
- Two vertices are called **adjacent** if they belong to (i.e., are contained in) the same hyperedge(s).
- Two hyperedges are called **incident** if their intersection is not empty.
- Sometimes, a hyperedge is said to be incident to any node contained in the hyperedge.

Difference between Incidence and Adjacency in Graphs and Hypergraphs

- In a graph, two distinct vertices u and v can only be either adjacent in a single edge ($(u, v) \in E$) or not ($(u, v) \notin E$); dually, two distinct edges e and f can only be either incident at a single vertex ($e \cap f = \{v\} \neq \varnothing$) or not ($e \cap f = \varnothing$).
- In a hypergraph, both adjacency and incidence are applicaple to sets of vertices and edges and their values can be any nonnegative integers (i.e., they are non-binary). More formally, in a hypergraph $G = (V, E)$, the mappings of adjacency $\mathrm{adj} : 2^V \to \mathbb{Z}_{\geq 0}$ and incidence $\mathrm{inc} : 2^E \to \mathbb{Z}_{\geq 0}$ are defined in the following way:
$$\mathrm{adj}(U) = |\{e \in E : e \supseteq U\}| \text{ and } \mathrm{inc}(F) = |\bigcap_{e \in F} e|,$$
where $U \subseteq V$ and $F \subseteq E$.* In particular, for singletons,
$$\mathrm{adj}(\{v\}) = |\{e \in E : e \ni v\}| \text{ and } \mathrm{inc}(\{e\}) = |e|,$$
and, for pairs of vertices and hyperedges,
$$\mathrm{adj}(\{u, v\}) = |\{e \in E : e \supseteq \{u, v\}\}| \text{ and } \mathrm{inc}(\{e, f\}) = |e \cap f|.$$

---

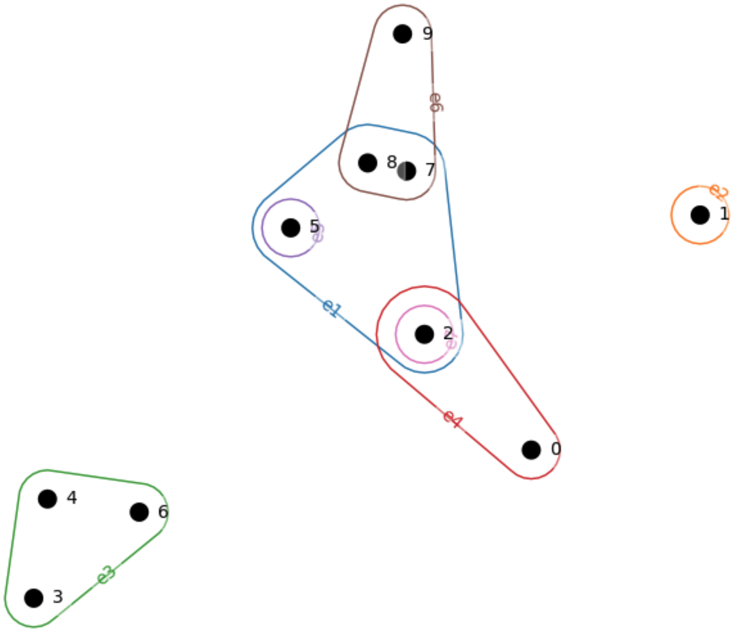
*The *power set* of an arbitrary set S, denoted by 2^S or $\mathscr{P}(S)$, is the set of all subsets of S, i.e., $2^S = \{T : T \subseteq S\}$.
For finite sets, the cardinality of the power set is $|2^S| = 2^{|S|}$.

Isolation and Looping

Let $G = (V, E)$ a hypergraph.

- If $\bigcup_{e \in E} e \subsetneq V$, a vertex $v \in V \setminus \bigcup_{e \in E} e$ is called an **isolated vertex**; in other words, vertex v is isolated if and only if
$$\mathrm{adj}(\{v\}) = 0.$$
- Clearly, if $\bigcup_{e \in E} e = V$, the hypergraph contains no isolated vertices.
- A hyperedge e is an **isolated hyperedge** if there are no other hyperedges incident to it; in other words, hyperedge e is isolated if and only if
$$\mathrm{inc}(\{e, f\}) = 0, \quad \text{for all} \quad f \in E \setminus \{e\}.$$
- A **loop** in a hypergraph is a hyperedge e incident to a single vertex, i.e., $\mathrm{inc}(\{e\}) = |e| = 1$ (e is a singleton).

Euler Plot of a Disconnected Hypergraph



Order, Size, Rank, and Co-Rank

Let $G = (V, E)$ a hypergraph. Given a finite set S, we denote by $|S|$ its **cardinality** (i.e., the number of elements of S).

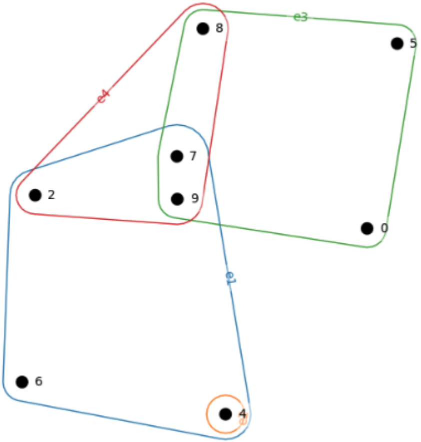
- The **order** of the hypergraph G, denoted as $o(G)$, is defined as the cardinality $|V|$ of its set of vertices.
- The **size** of the hypergraph G is defined as the cardinality $|E|$ of the set of hyperedges.
- The **rank** of the hypergraph G, denoted as $r(G)$, is defined as the maximum of the cardinalities of the hyperedges, i.e., $r(G) = \max_{e \in E} |e|$.
- The **co-rank** of the hypergraph G, denoted as $cr(G)$, is defined as the minimum of the cardinalities of the hyperedges, i.e., $cr(G) = \min_{e \in E} |e|$.

Degrees of Hyperedges and Vertices

Let $G = (V, E)$ a hypergraph.

- The **degree of hyperedge** e, denoted as $\delta(e)$, is defined as $|e|$, i.e., as the cardinality of (i.e., the number of incident vertices to) e, i.e., $\delta(e) = |e| = \mathrm{inc}(\{e\})$.
- If each hyperedge has the same degree k, i.e., if $r(G) = cr(G) = k$, the hypergraph is called **uniform** or k-**uniform**.
- The **star centered in a vertex** $v \in V$ of the hypergraph G, denoted as $G(v)$, is the family of hyperedges incident to v.
- The **the degree of vertex** v, denoted as $\deg(v)$, is defined as $|G(v)|$, i.e., as the number of hyperedges incident to v:
$$\deg(v) = |\{e \in E : v \in e\}| = \mathrm{adj}(\{v\}).$$
- If each vertex has the same degree, the hypergraph is called **regular** or k-**regular**, i.e., if, for every $v \in V, \deg(v) = k$.
- Degree equality:
$$\sum_{v \in V} \deg(v) = \sum_{e \in E} \delta(e).$$

Hypergraph Vertex and Hyperedge Degrees



Hypergraph Data:
{'e1': [2, 4, 6, 7, 9], 'e2': [4], 'e3': [0, 5, 7, 8, 9], 'e4': [2, 7, 8, 9]}

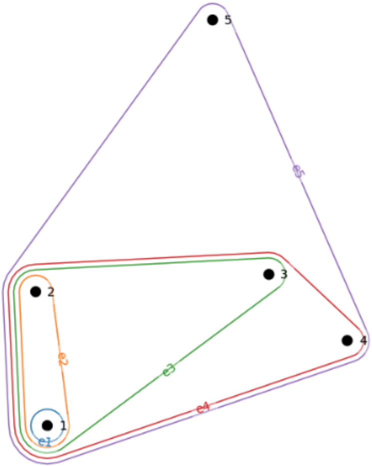
Incidence Matrix:

vertex	e1	e2	e3	e4
0	0	0	1	0
2	1	0	0	1
4	1	1	0	0
5	0	0	1	0
6	1	0	0	0
7	1	0	1	1
8	0	0	1	1
9	1	0	1	1

Node Degrees:
Degree of node 2: 2
Degree of node 4: 2
Degree of node 6: 1
Degree of node 7: 3
Degree of node 9: 3
Degree of node 0: 1
Degree of node 5: 1
Degree of node 8: 2

Hyperedge Degrees:
Degree of hyperedge e1: 5
Degree of hyperedge e2: 1
Degree of hyperedge e3: 5
Degree of hyperedge e4: 4

A Hypergraph with Distinct Vertex and Hyperedge Degrees



Hypergraph Data:
{'e1': {1}, 'e2': {1, 2}, 'e3': {1, 2, 3}, 'e4': {1, 2, 3, 4}, 'e5': {1, 2, 3, 4, 5}}

Incidence Matrix:

vertex	e1	e2	e3	e4	e5
1	1	1	1	1	1
2	0	1	1	1	1
3	0	0	1	1	1
4	0	0	0	1	1
5	0	0	0	0	1

Node Degrees:
Degree of node 1: 5
Degree of node 2: 4
Degree of node 3: 3
Degree of node 4: 2
Degree of node 5: 1

Hyperedge Degrees:
Degree of hyperedge e1: 1
Degree of hyperedge e2: 2
Degree of hyperedge e3: 3
Degree of hyperedge e4: 4
Degree of hyperedge e5: 5

Dual Hypergraphs

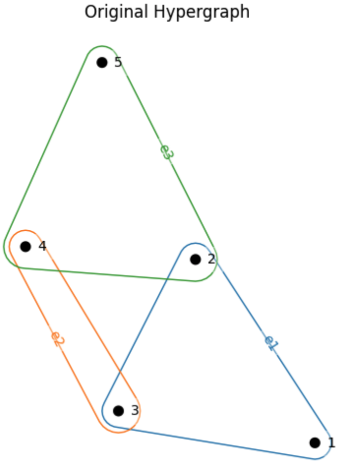
- The dual of a hypergraph G is the hypergraph G^*, where the vertices of G^* correspond to the hyperedges of H, and the hyperedges of G^* correspond to the vertices of G. The incidence relation in G^* links each vertex to the hyperedges incident to this vertex. More formally:

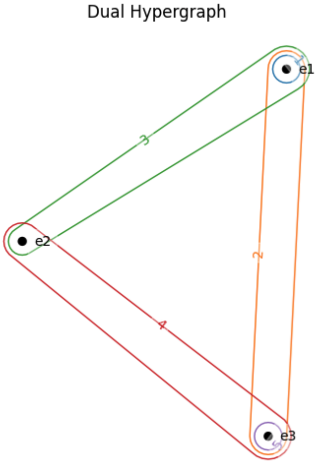
Definition

Let $G = (V, E)$ be a hypergraph with $V = \{v_1, v_2, \ldots, v_n\}$ and $E = \{e_1, e_2, \ldots, e_m\}$. The **dual hypergraph** of G, denoted $G^* = (E^*, V^*)$, is a hypergraph with set of vertices $E^* = \{e_1^*, \ldots, e_m^*\}$ and family of hyperedges $V^* = \{v_1^*, \ldots, v_n^*)$, where $v_i^* = \{e_k^* : v_i \in e_k\}$.

- If two or more vertices in G are incident with the same collection of hyperedges, each of these vertices will produce an identical hyperedge in the dual hypergraph G^*. As a result, the dual hypergraph G^* may, in general, be a multi-hypergraph.

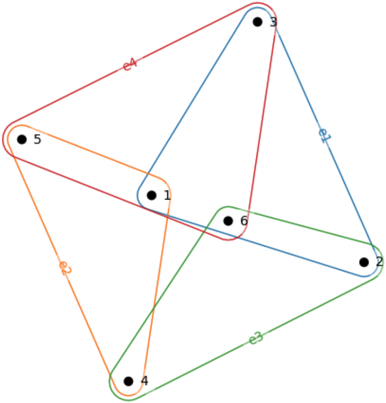
Euler Plots of a Hypergraph and its Dual



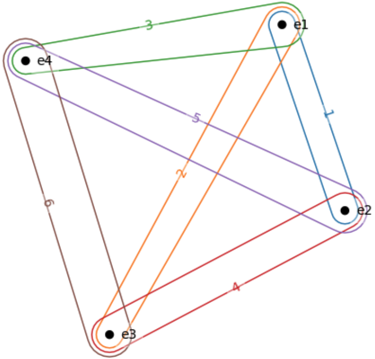


A Uniform Hypergraph with a Uniform Dual

Original Hypergraph



Dual Hypergraph



Hypergraph Matrices

Let $G = (V, E)$ a hypergraph of order $|V| = n$ and size $|E| = m$, which means that $V = \{v_1, v_2, \dots, v_n\}$ and $E = \{e_1, e_2, \dots, e_m\}$.

- The **incidence matrix** of the hypergraph G is a $n \times m$ matrix $H = \{h_{ij}\}$, where

$$h_{ij} = \begin{cases} 1, & \text{if } v_i \in e_j, \\ 0, & \text{if } v_i \notin e_j. \end{cases}$$

- The **adjacency matrix** of the hypergraph G is a square $n \times n$ matrix $A = \{a_{ij}\}$, where

$$a_{ij} = \begin{cases} |\{e \in E : v_i \text{ and } v_j \in e\}|, & \text{if } i \neq j, \\ 0, & \text{if } i = j. \end{cases}$$

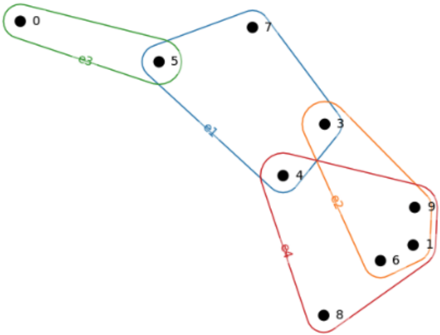
- The degree of a vertex is the sum of its row in the incidence matrix:

$$\deg(v_i) = \sum_{j=1}^{m} h_{ij}, \text{ for } i = 1, \dots, n.$$

- The degree of a hyperedge is the sum of its columns in the incidence matrix:

$$\delta(e_j) = \sum_{i=1}^{n} h_{ij}, \text{ for } j = 1, \dots, m.$$

Hypergraph Incidence and Adjacency Matrices



Hypergraph Data:

{'e1': [3, 4, 5, 7], 'e2': [1, 3, 6, 9], 'e3': [0, 5], 'e4': [1, 4, 6, 8, 9]}

Incidence Matrix:

vertex	e1	e2	e3	e4
0	0	0	1	0
1	0	1	0	1
3	1	1	0	0
4	1	0	0	1
5	1	0	1	0
6	0	1	0	1
7	1	0	0	0
8	0	0	0	1
9	0	1	0	1

Adjacency Matrix:

```
0 0 0 0 1 0 0 0 0
0 0 1 1 0 2 0 1 2
0 1 0 1 1 1 1 0 1
0 1 1 0 1 1 1 1 1
1 0 1 1 0 0 1 0 0
0 2 1 1 0 0 0 1 2
0 0 1 1 1 0 0 0 0
0 1 0 1 0 1 0 0 1
0 2 1 1 0 2 0 1 0
```

Incidence and Adjacency Matrices Relationships

- If H, H^* are the incidence matrices of a hypergaph G and its dual G^*, respectively, it can be shown:
$$(G^*)^* = G, H^* = H^\top.$$
- Denoting by D_v the $n \times n$ diagonal matrix of vertex degrees, it holds:
$$A = HH^\top - D_v.$$
- Note that while H is a binary matrix (with entries of 0 or 1), the matrix A may not be, as its entries can be any nonnegative integers, due to the product $H^\top H$.
- Denoting by A_{G^*} the adjacency matrix of the dual hypergraph G^* and by D_e the $m \times m$ diagonal matrix of hyperedge degrees, it can be shown:
$$A_{G^*} = H^\top H - D_e.$$

Clique Expansion of a Hypergraph

- In the clique expansion of a hypergraph, the resulting graph is constructed on the set of vertices of the hypergraph by treating each hyperedge as a clique, thus connecting every pair of vertices within each hyperedge. Since hyperedges may overlap (if they are incident), the resulting cliques can also overlap, which makes the clique-expanded graph a multi-graph—specifically, a weighted graph. More formally:

Definition

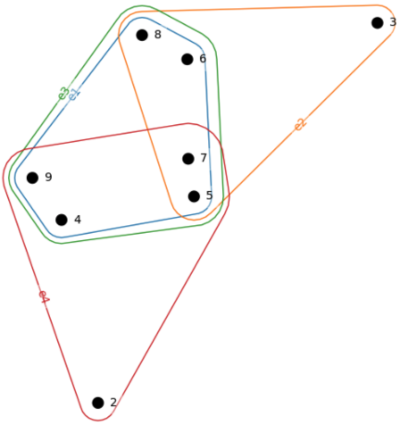
Let $G = (V, E)$ be a hypergraph. The **clique-expanded graph** (or **2–section**) of G is the graph, denoted $\text{Clique}(G) = (V, E')$, where V is the set of vertices of the hypergraph and

$$E' = \{(v_i, v_j) : v_i, v_j \in e, e \in E\}.$$

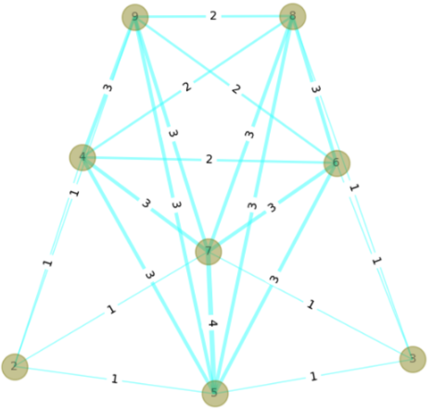
- $\text{Clique}(G)$ is (in general) a weighted graph with edge weights $\omega(v_i, v_j) = |\{e \in E : v_i, v_j \in E\}|$.
- The adjacency matrix of the clique expansion of a hypergraph is the same as the adjacency matrix of the hypergraph (as defined previously).

Hypergraph and its Clique Expansion

Hypergraph



2-Section Graph



Line Expansion of a Hypergraph

- In the line expansion of a hypergraph, the resulting line graph is constructed by representing each hyperedge as a vertex and connecting two vertices with an edge if their corresponding hyperedges intersect. More formally:

Definition

Let $G = (V, E)$ be a hypergraph. The **line graph** (or **intersection graph**) of G, denoted $\mathrm{Line}(G)$, is the graph $\mathrm{Line}(G) = (E, E')$, where E is the set of hyperedges of the hypergraph and

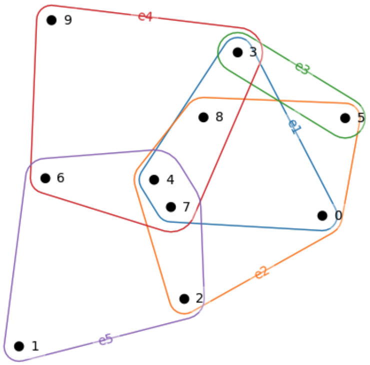
$$E' = \{(e_i, e_j) : e_i, e_j \in E, e_i \cap e_j \neq \varnothing\}.$$

- $\mathrm{Line}(G)$ is (in general) a weighted graph with edge weights $\omega(e_i, e_j) = |e_i \cap e_j|$.
- Similarly, the line graph of the dual hypergraph G^*, denoted $\mathrm{Line}(G^*)$, can be defined analogously.
- In graph theory, the line graphs $\mathrm{Line}(G)$ and $\mathrm{Line}(G^*)$ are commonly referred to as the "top" and "bottom projections", respectively, of the bipartite graph associated with the hypergraph G.

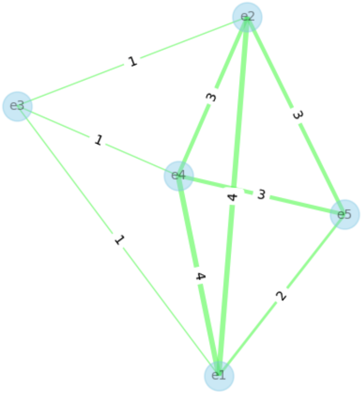
Hypergraph and Line Graph

```
Hypergraph Data:
{'e1': [0, 3, 4, 7, 8], 'e2': [0, 2, 4, 5, 7, 8], 'e3': [3, 5], 'e4': [3, 4, 6, 7, 8, 9], 'e5': [1, 2, 4, 6, 7]}
```

Hypergraph



Line Graph



Matrix Relationships in Hypergraphs, Their Duals, and Line Graphs

The clique expansion of a hypergraph (or its dual) and the line graph of its dual (or the hypergraph, respectively) are isomorphic weighted graphs. They have the same edges and weights, leading to identical adjacency matrices, but differ in their vertex labeling.

Structure	**Matrix**	**Incidence**	**Formula**
Hypergraph G	Incidence matrix H_G	$(v, e), v \in e$	H_G
Dual hypergraph G^*	Incidence matrix H_{G^*}	$(e^*, v^*), v \in e$	$H_{G^*} = H_G^\top$
Hypergraph G	Adjacency matrix (binary) A_G	$(v_i, v_j), v_i, v_j \in e$	$A_G = H_G H_G^\top - D_v$
Dual hypergraph G^*	Adjacency matrix (binary) A_{G^*}	$(e_i^*, e_j^*), e_i \cap e_j \neq \varnothing$	$A_{G^*} = H_G^\top H_G - D_e$
Line graph $L(G)$	Adjacency matrix $A_{L(G)}$	$(e_i, e_j), e_i \cap e_j \neq \varnothing$	$A_{L(G)} = H_G^\top H_G - \mathrm{diag}(H_G^\top H_G)$
Line graph $L(G^*)$	Adjacency matrix $A_{L(G^*)}$	$(v_i^*, v_j^*), v_i, v_j \in e$	$A_{L(G^*)} = H_G H_G^\top - \mathrm{diag}(H_G H_G^\top)$

An Example of Matrix Relationships in Hypergraphs, their Duals, and Line Graphs

Hypergraph: $G = (V, E)$, $V = \{v_1, v_2, v_3, v_4\}$, $E = \{e_1, e_2, e_3\}$, where $e_1 = \{v_2, v_4, v_5\}$, $e_2 = \{v_4, v_5\}$, $e_3 = \{v_2, v_3, v_5\}$, $e_4 = \{v_1, v_2\}$.

Incidence Matrices $H(G)$ and $H(G^*)$:

$$H(G) = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \end{bmatrix}, \quad H(G^*) = \begin{bmatrix} 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 \end{bmatrix}$$

Products $H(G)H(G)^T$ and $H(G)^T H(G)$:

$$H(G)H(G)^T = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 1 & 3 & 1 & 1 & 2 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 2 & 2 \\ 0 & 2 & 1 & 2 & 3 \end{bmatrix}, \quad H(G)^T H(G) = \begin{bmatrix} 3 & 2 & 2 & 1 \\ 2 & 2 & 1 & 0 \\ 2 & 1 & 3 & 1 \\ 1 & 0 & 1 & 2 \end{bmatrix}$$

Adjacency Matrix of G, $A(G)$, and of the Dual Hypergraph, $A(G^*)$:

$$A(G) = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 2 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 2 \\ 0 & 2 & 1 & 2 & 0 \end{bmatrix}, \quad A(G^*) = \begin{bmatrix} 0 & 2 & 2 & 1 \\ 2 & 0 & 1 & 0 \\ 2 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{bmatrix}$$

Adjacency Matrix of the Line Graph of G, $A(L(G))$, and of the Line Graph of the Dual Hypergraph, $A(L(G^*))$:

$$A(L(G)) = \begin{bmatrix} 0 & 2 & 2 & 1 \\ 2 & 0 & 1 & 0 \\ 2 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{bmatrix}, \quad A(L(G^*)) = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 2 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 2 \\ 0 & 2 & 1 & 2 & 0 \end{bmatrix}$$

Vertex Degree Preserving Adjacency (VDPA) Matrix of Hypergraph

As one can easily observe, the hypergraph adjacency matrix (i.e., the adjacency matrix of the clique-expanded graph) overestimates the degree of a vertex v by a factor of $(\delta(e) - 1)$ for each hyperedge containing v. To maintain the correct vertex degree, Kumar et. al (2019)* demonstrated that this can be adjusted by scaling down using the following formula for the adjacency matrix:

$$A_{vpd} = H(D_e - I_m)^{-1} H^\top - D_v.$$

*Kumar, Vaidyanathan, Ananthapadmanabhan, Parthasarathy, and Ravindran (2019). A new measure of modularity in hypergraphs: Theoretical insights and implications for effective clustering. In *International Conference on Complex Networks and Their Applications*

Adjacency and VDPA Matrices

```
{'e1': [0, 1, 3, 9], 'e2': [7], 'e3': [0, 2, 4, 7, 8, 9], 'e4': [0, 2, 3, 5, 6, 7]}
```

Vertex Degrees:
Degree of node 0: 3
Degree of node 1: 1
Degree of node 2: 2
Degree of node 3: 2
Degree of node 4: 1
Degree of node 5: 1
Degree of node 6: 1
Degree of node 7: 3
Degree of node 8: 1
Degree of node 9: 2

Adjacency Matrix:

```
0 1 2 2 1 1 1 2 1 2
1 0 0 1 0 0 0 0 0 1
2 0 0 1 1 1 1 2 1 1
2 1 1 0 0 1 1 1 0 1
1 0 1 0 0 0 0 1 1 1
1 0 1 1 0 0 1 1 0 0
1 0 1 1 0 1 0 1 0 0
2 0 2 1 1 1 1 0 1 1
1 0 1 0 1 0 0 1 0 1
2 1 1 1 1 0 0 1 1 0
```

VDPA Matrix:

0	1	2	3	4	5	6	7	8	9
0.00	0.33	0.4	0.53	0.2	0.2	0.2	0.4	0.2	0.53
0.33	0.00	0.0	0.33	0.0	0.0	0.0	0.0	0.0	0.33
0.40	0.00	0.0	0.20	0.2	0.2	0.2	0.4	0.2	0.20
0.53	0.33	0.2	0.00	0.0	0.2	0.2	0.2	0.0	0.33
0.20	0.00	0.2	0.00	0.0	0.0	0.0	0.2	0.2	0.20
0.20	0.00	0.2	0.20	0.0	0.0	0.2	0.2	0.0	0.00
0.20	0.00	0.2	0.20	0.0	0.2	0.0	0.2	0.0	0.00
0.40	0.00	0.4	0.20	0.2	0.2	0.2	1.0	0.2	0.20
0.20	0.00	0.2	0.00	0.2	0.0	0.0	0.2	0.0	0.20
0.53	0.33	0.2	0.33	0.2	0.0	0.0	0.2	0.2	0.00