深蓝学院
shenlanxueyuan.com

图论基础

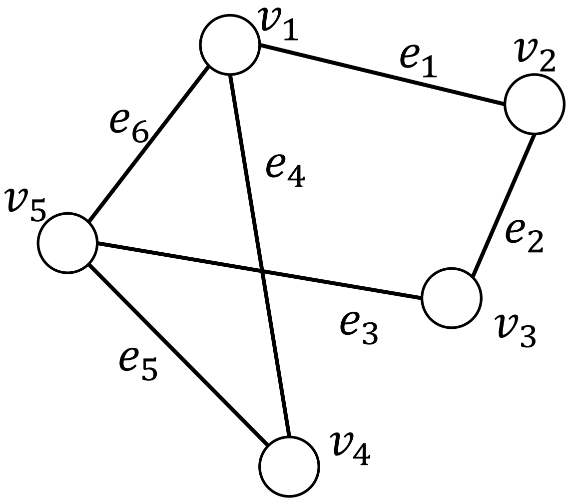
汤继良、王怡琦、金卫：密西根州立大学

马耀：新泽西理工

目录

目录

图的简介



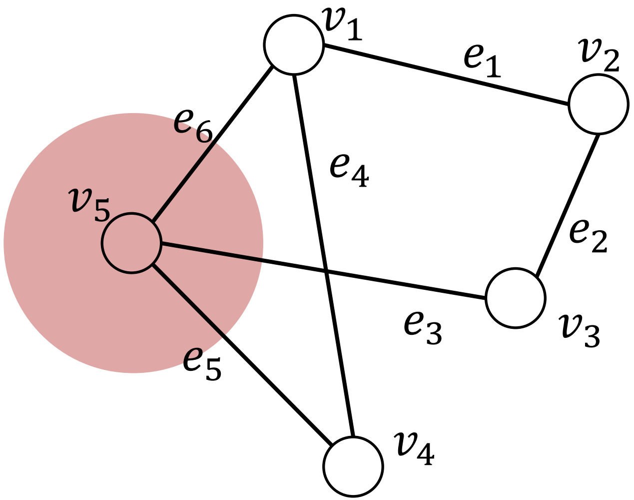
$$\mathbf{A} = \begin{pmatrix} 0 & 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 \end{pmatrix}$$

$$\mathcal{V} = \{v_1, v_2, v_3, v_4, v_5\}$$

$$\mathcal{E} = \{e_1, e_2, e_3, e_4, e_5, e_6\}$$

目录

度



$$\mathcal{V} = \{v_1, v_2, v_3, v_4, v_5\}$$

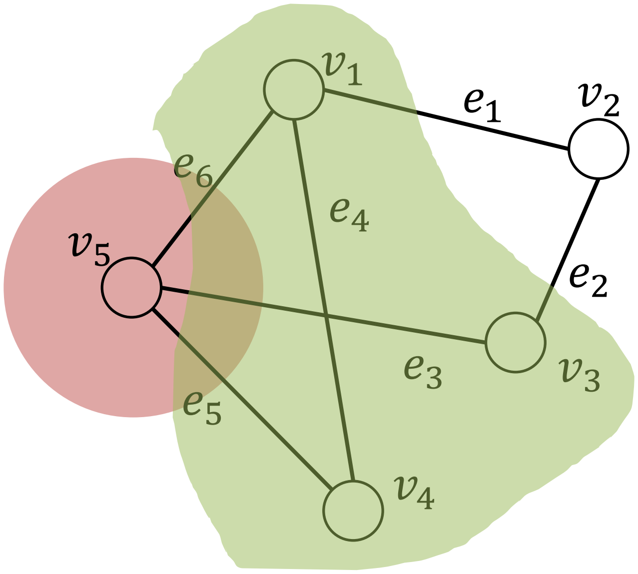
$$\mathcal{E} = \{e_1, e_2, e_3, e_4, e_5, e_6\}$$

$$d(v_i)$$

与v_i节点相连的边的数量

$$d(v_5) = 3$$

节点的邻域



$$\mathcal{V} = \{v_1, v_2, v_3, v_4, v_5\}$$

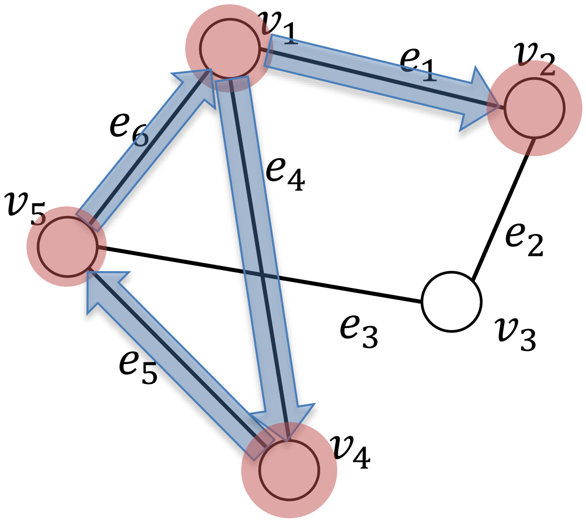
$$\mathcal{E} = \{e_1, e_2, e_3, e_4, e_5, e_6\}$$

$$\mathcal{N}(v_i)$$

与v_i节点相连的点的集合

$$\mathcal{N}(v_5) = \{v_1, v_3, v_4\}$$

途径(walk)

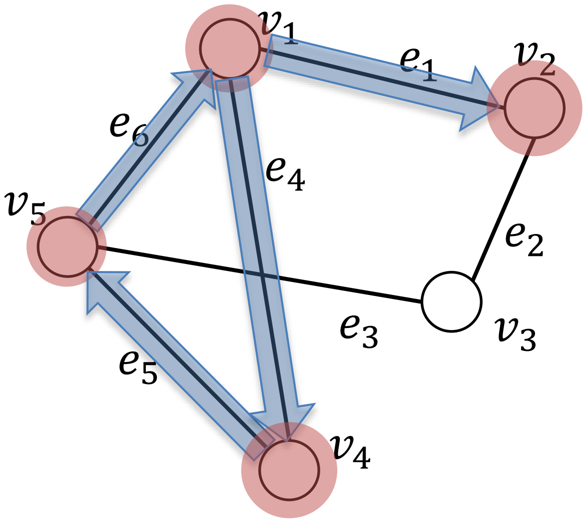


途径是**节点和边的交替序列**，从一个节点开始，以一个节点结束，其中每条边与紧邻的节点相连

$$(v_1, e_4, v_4, e_5, v_5, e_6, v_1, e_1, v_2)$$

途径的长度是途径中包含的边的数量

迹(trail)和路(path)

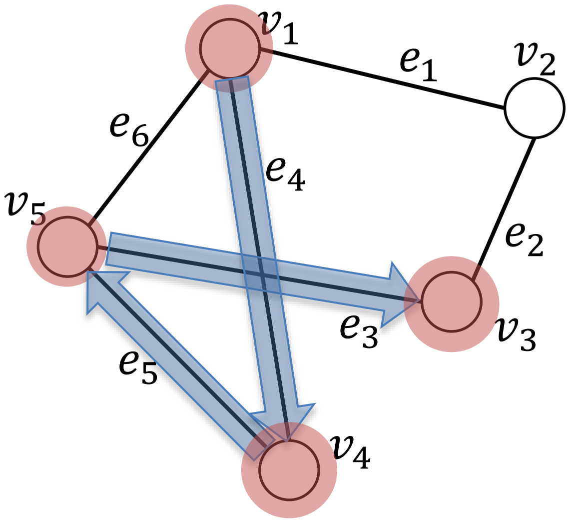


- ❑迹是边各不相同的途径
- ❑路是节点各不相同的途径

$$(v_1, e_4, v_4, e_5, v_5, e_6, v_1, e_1, v_2)$$

是一条迹，但不是路

迹(trail)和路(path)

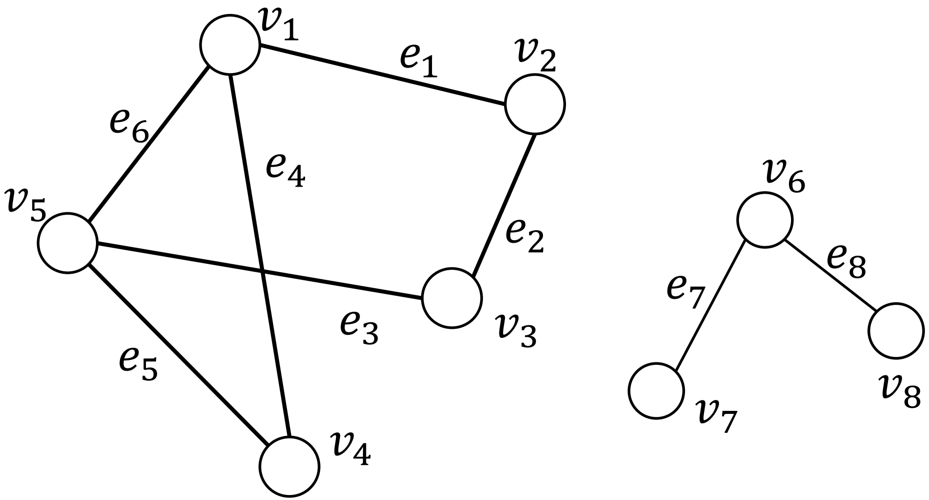


❑迹是边各不相同的途径

❑路是节点各不相同的途径

$$(v_1, e_4, v_4, e_5, v_5, e_3, v_3)$$

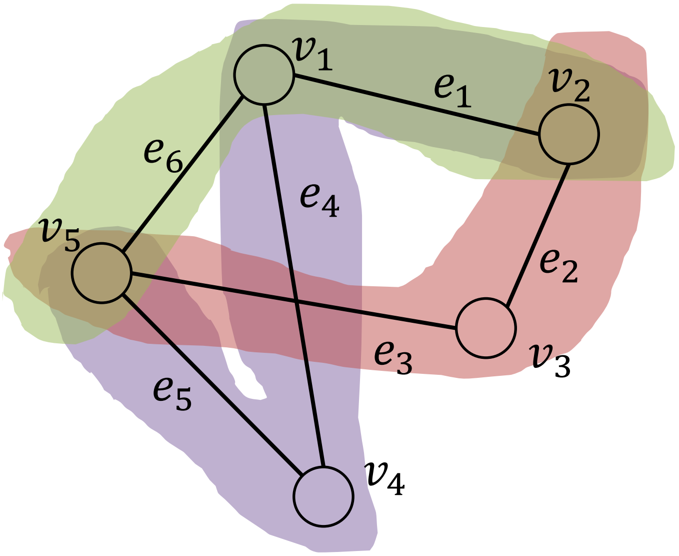
图的连通性



不是连通图

给定一个图，如果图中的**任意两个节点之间都至少存在一条路**，则这个图是一个连通图。

最短路



v_5 和v_2之间的最短路

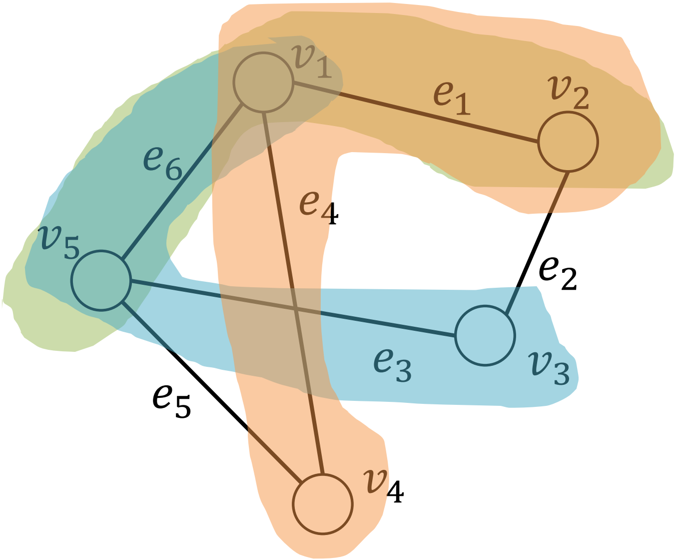
$(v_5, e_6, v_1, e_1, v_2)$

$(v_5, e_3, v_3, e_2, v_2)$

$(v_5, e_5, v_4, e_4, v_1, e_1, v_2)$

给定连通图中的任意两点，连接这两点的长度最小的路被称为这两点之间的最短路。最短路的长度被称为两点间的距离。

图的直径



$$(v_5, e_6, v_1, e_1, v_2)$$

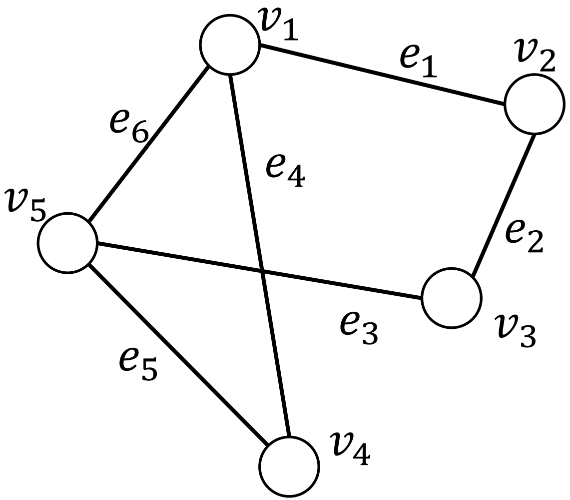
$$(v_1, e_6, v_5, e_3, v_3)$$

$$\vdots$$

直径为2

图的直径是指图中最远的两点间的距离（即最长的最短路的长度）

节点的中心性

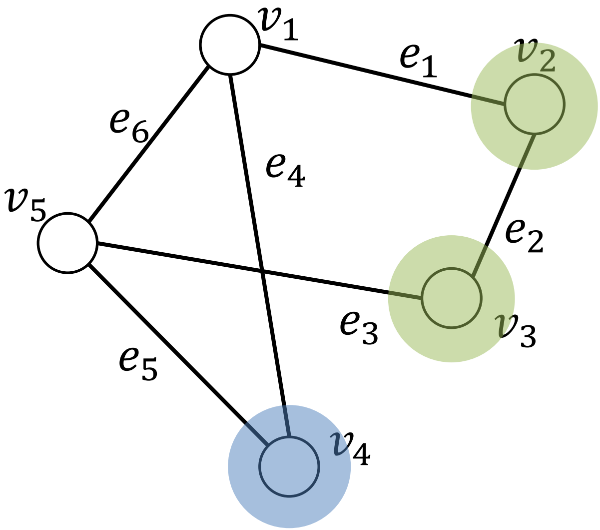


$$\mathcal{V} \rightarrow \mathbb{R}^N$$

将每个节点映射到一个标量

节点的中心性用来衡量节点在图上的重要程度

度中心性

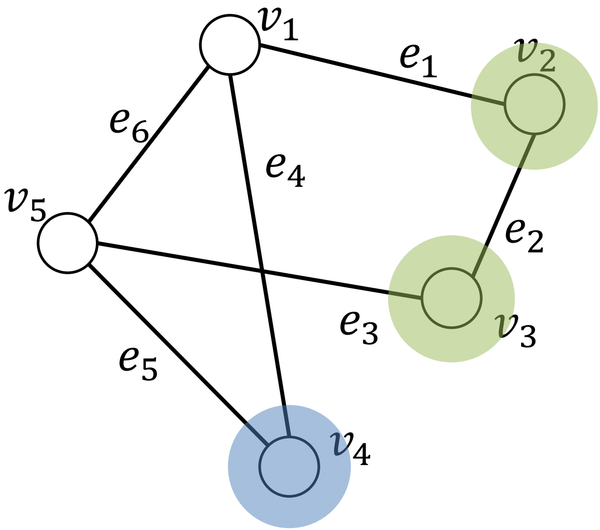


$$c_d(v_1) = 3$$
$$c_d(v_2) = 2$$
$$c_d(v_3) = 2$$
$$c_d(v_4) = 2$$
$$c_d(v_5) = 3$$

利用节点的度来衡量节点的中心性

$$c_d(v_i) = d(v_i) = \sum_{j=1}^{N} \mathbf{A}_{i,j}$$

特征向量中心性



$$c_e(v_1) = 1$$

$$c_e(v_2) = 0.68$$

$$c_e(v_3) = 0.68$$

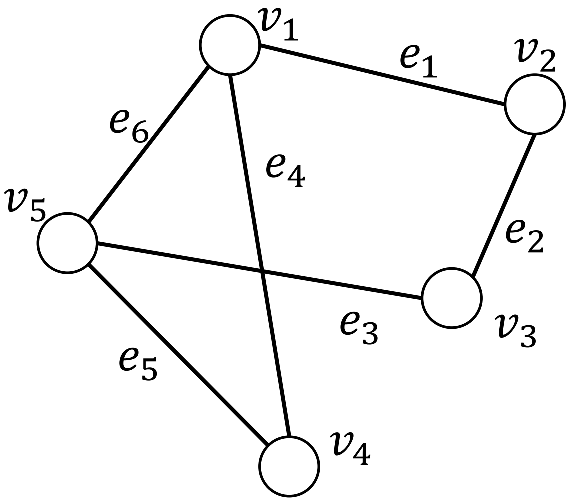
$$c_e(v_4) = 0.81$$

$$c_e(v_5) = 1$$

衡量节点的中心性时同时考虑邻居节点的中心性

$$c_e(v_i) = \frac{1}{\lambda}\sum_{j=1}^{N} \mathbf{A}_{i,j} \cdot c_e(v_j) \Rightarrow \lambda \cdot \mathbf{c}_e = \mathbf{A} \cdot \mathbf{c}_e$$

Katz中心性



$$\alpha = \frac{1}{2}$$

$$\beta = 1$$

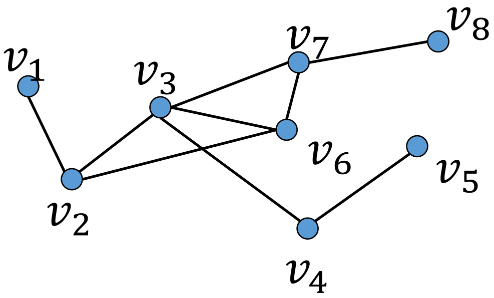
$$c_k(v_1) = 2.2$$
$$c_k(v_2) = 1.8$$
$$c_k(v_3) = 1.8$$
$$c_k(v_4) = 1.9$$
$$c_k(v_5) = 2.2$$

Katz是特征向量中心性的一个变种 $(0, \frac{1}{\lambda})$

$$c_k(v_i) = \alpha \sum_{j=1}^{N} \mathbf{A}_{i,j} c_k(v_j) + \beta \qquad \mathbf{c}_k = (\mathbf{I} - \alpha \cdot \mathbf{A})^{-1} \boldsymbol{\beta}$$

目录

图和图信号

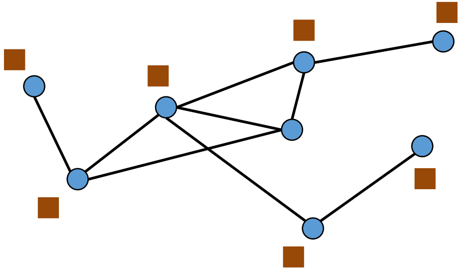


$$\mathcal{V} = \{v_1, \ldots, v_N\}$$

$$\mathcal{E} = \{e_1, \ldots, e_M\}$$

$$\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$$

图和图信号



图信号: $f : \mathcal{V} \rightarrow \mathbb{R}^N$

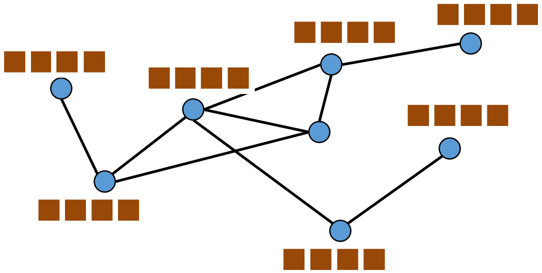
$$\mathcal{V} = \{v_1, \ldots, v_N\}$$

$$\mathcal{E} = \{e_1, \ldots, e_M\}$$

$$\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$$

$$\mathcal{V} \longrightarrow \begin{bmatrix} f(1) \\ f(2) \\ f(3) \\ f(4) \\ f(5) \\ f(6) \\ f(7) \\ f(8) \end{bmatrix}$$

图和图信号



图信号： $f: \mathcal{V} \rightarrow \mathbb{R}^{N \times d}$

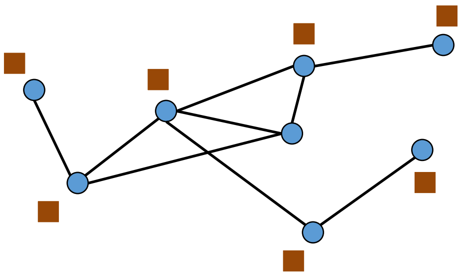
$$\mathcal{V} = \{v_1, \ldots, v_N\}$$

$$\mathcal{E} = \{e_1, \ldots, e_M\}$$

$$\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$$

$$\mathcal{V} \longrightarrow \begin{bmatrix} f(1) \\ f(2) \\ f(3) \\ f(4) \\ f(5) \\ f(6) \\ f(7) \\ f(8) \end{bmatrix}$$

图和图信号



$$\mathcal{V} = \{v_1, \ldots, v_N\}$$

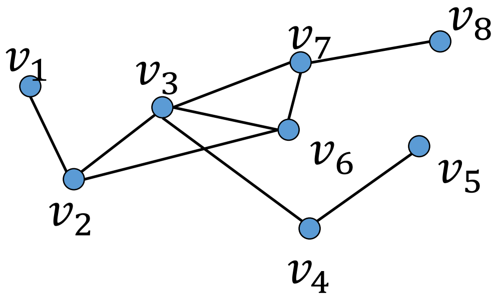
$$\mathcal{E} = \{e_1, \ldots, e_M\}$$

$$\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$$

图信号： $f : \mathcal{V} \rightarrow \mathbb{R}^N$

$$\mathcal{V} \longrightarrow \begin{bmatrix} f(1) \\ f(2) \\ f(3) \\ f(4) \\ f(5) \\ f(6) \\ f(7) \\ f(8) \end{bmatrix}$$

图的矩阵表示



度矩阵： $\mathbf{D} = \mathrm{diag}(degree(v_1), \ldots, degree(v_N))$

度矩阵　　邻接矩阵　　拉普拉斯矩阵

$$\underbrace{\begin{pmatrix} 1&0&0&0&0&0&0&0\\ 0&3&0&0&0&0&0&0\\ 0&0&4&0&0&0&0&0\\ 0&0&0&2&0&0&0&0\\ 0&0&0&0&2&0&0&0\\ 0&0&0&0&0&4&0&0\\ 0&0&0&0&0&0&3&0\\ 0&0&0&0&0&0&0&1 \end{pmatrix}}_{\boldsymbol{D}} - \underbrace{\begin{pmatrix} 0&1&0&0&0&0&0&0\\ 1&0&1&0&0&1&0&0\\ 0&1&0&1&0&1&1&0\\ 0&0&1&0&1&0&0&0\\ 0&0&0&1&0&1&0&0\\ 0&1&1&0&1&0&1&0\\ 0&0&1&0&0&1&0&1\\ 0&0&0&0&0&0&1&0 \end{pmatrix}}_{\boldsymbol{A}} = \underbrace{\begin{pmatrix} 1&-1&0&0&0&0&0&0\\ -1&3&-1&0&0&-1&0&0\\ 0&-1&4&-1&0&-1&-1&0\\ 0&0&-1&2&-1&0&0&0\\ 0&0&0&-1&2&-1&0&0\\ 0&-1&-1&0&-1&4&-1&0\\ 0&0&-1&0&0&-1&3&-1\\ 0&0&0&0&0&0&-1&1 \end{pmatrix}}_{\boldsymbol{L}}$$

拉普拉斯矩阵作为算子

拉普拉斯矩阵可以作为一个差分算子

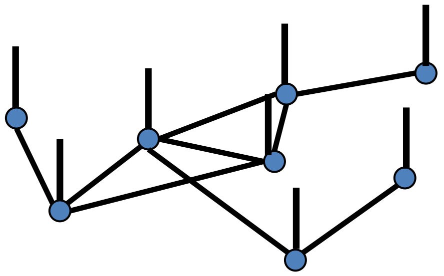
$$\mathbf{h} = \mathbf{L}\mathbf{f} = (\mathbf{D} - \mathbf{A})\mathbf{f} = \mathbf{D}\mathbf{f} - \mathbf{A}\mathbf{f}$$

$$\mathbf{h}(i) = \sum_{v_j \in \mathcal{N}(v_i)} (\mathbf{f}(i) - \mathbf{f}(j))$$

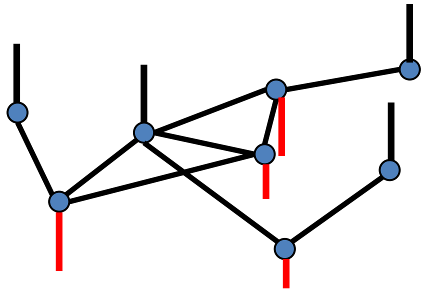
拉普拉斯二次型

$$\mathbf{f}^T\mathbf{L}\mathbf{f} = \frac{1}{2}\sum_{i,j=1}^{N} \mathbf{A}[i,j](\mathbf{f}(i) - \mathbf{f}(j))^2$$

- 衡量了图信号的“光滑度”或者“频率”
- 拉普拉斯矩阵为半正定矩阵



“光滑”的“低频”信号



“不光滑”的“高频”信号

拉普拉斯矩阵的特征分解

拉普拉斯矩阵有一套完整的标准正交的特征向量

$$\mathbf{L} = \begin{bmatrix} | & & | \\ \mathbf{u}_0 & \cdots & \mathbf{u}_{N-1} \\ | & & | \end{bmatrix} \begin{bmatrix} \lambda_0 & & 0 \\ & \ddots & \\ 0 & & \lambda_{N-1} \end{bmatrix} \begin{bmatrix} — & \mathbf{u}_0 & — \\ & \vdots & \\ — & \mathbf{u}_{N-1} & — \end{bmatrix}$$

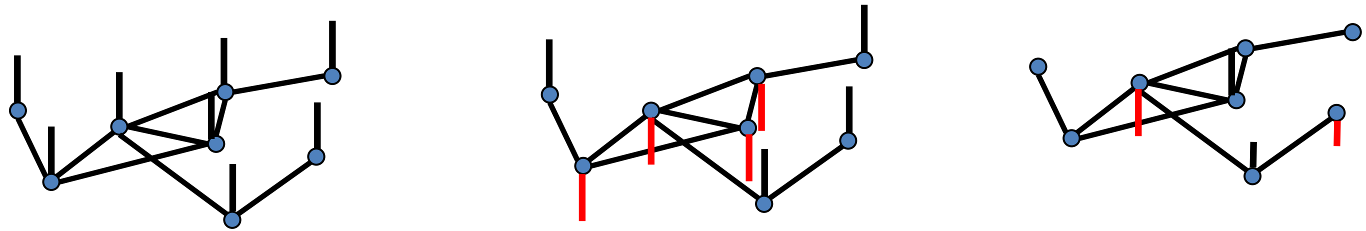
$$\boldsymbol{U} \qquad \boldsymbol{\Lambda} \qquad \boldsymbol{U}^{\boldsymbol{T}}$$

通常我们将这些特征向量按照特征值从小到大排列

$$0 = \lambda_0 < \lambda_1 \leq \cdots \lambda_{N-1}$$

特征向量作为图上的信号

这些特征向量是图信号的一组基



$\boldsymbol{u}_0$ $\boldsymbol{u}_1$ $\boldsymbol{u}_7$

低频 高频

$\boldsymbol{u}_0^T \boldsymbol{L} \boldsymbol{u}_0 = \boldsymbol{\lambda}_0 = \mathbf{0}$ $\boldsymbol{u}_1^T \boldsymbol{L} \boldsymbol{u}_1 = \boldsymbol{\lambda}_1$ $\boldsymbol{u}_7^T \boldsymbol{L} \boldsymbol{u}_7 = \boldsymbol{\lambda}_7$

频率: $\mathbf{u}_i^T \mathbf{L} \mathbf{u}_i$

图傅立叶变换（GFT）

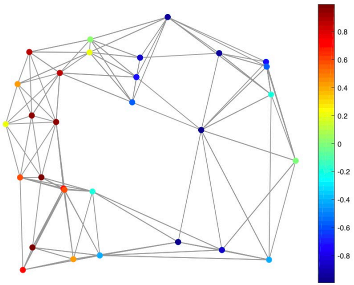
任意的的图信号f可以用图傅立叶级数表示

$$\mathbf{f} = \sum_{i=0}^{N-1} \boxed{\hat{f}_i} \cdot \mathbf{u}_i$$

$$\hat{f}_i = \boldsymbol{f}^\top \boldsymbol{u}_i$$

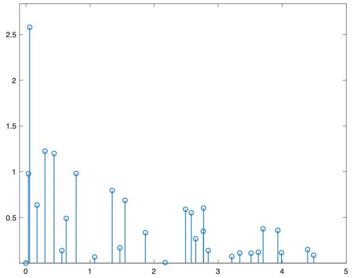
$\boldsymbol{u}_i$:基（特征向量）　　λ_i: 基的频率（特征值）

$\hat{f}_i$:傅立叶系数



$$\hat{\boldsymbol{f}} = \boldsymbol{U}^T \boldsymbol{f}$$

分解图信号



空间域: $\boldsymbol{f}$

谱域: $\hat{\boldsymbol{f}}$

图傅立叶逆变换（IGFT）

任意的的图信号f可以用图傅立叶级数表示

$$\mathbf{f} = \sum_{i=0}^{N-1} \hat{f}_i \cdot \mathbf{u}_i$$

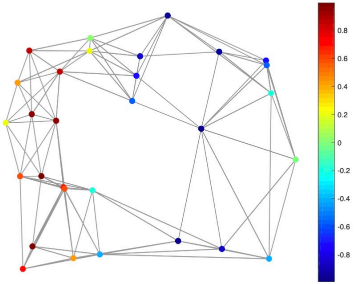
$$\hat{f}_i = \boldsymbol{f}^\top \boldsymbol{u}_i$$

$\boldsymbol{u}_i$:基（特征向量）　λ_i: 基的频率（特征值）

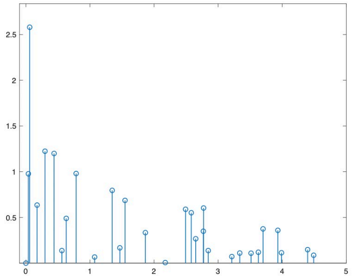
$\hat{f}_i$:傅立叶系数

矩阵形式

$$\boldsymbol{f} = \boldsymbol{U}\,\hat{\boldsymbol{f}}$$



重建图信号

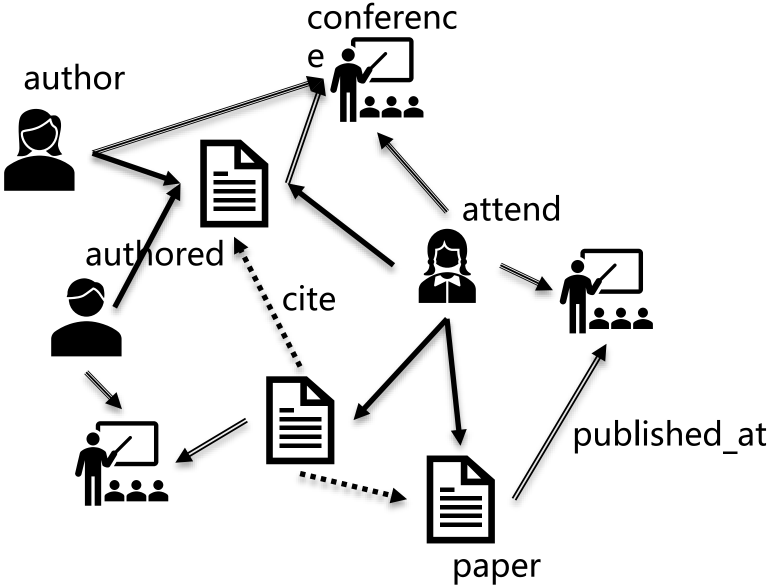


空间域: $\boldsymbol{f}$

谱域: $\hat{\boldsymbol{f}}$

目录

异质图



$$\mathcal{V}, \mathcal{E}$$

包含**不同种类的节点和边**

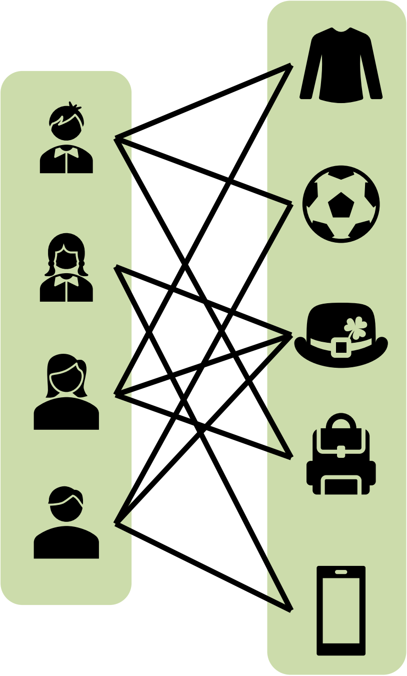
$$\phi_n : \mathcal{V} \to \mathcal{T}_n$$

$$\phi_e : \mathcal{V} \to \mathcal{T}_e$$

$$\mathcal{T}_n = \{\text{author}, \text{paper}, \text{conference}\}$$

$$\mathcal{T}_e = \{\text{authored}, \text{cite}, \text{published\_at}\}$$

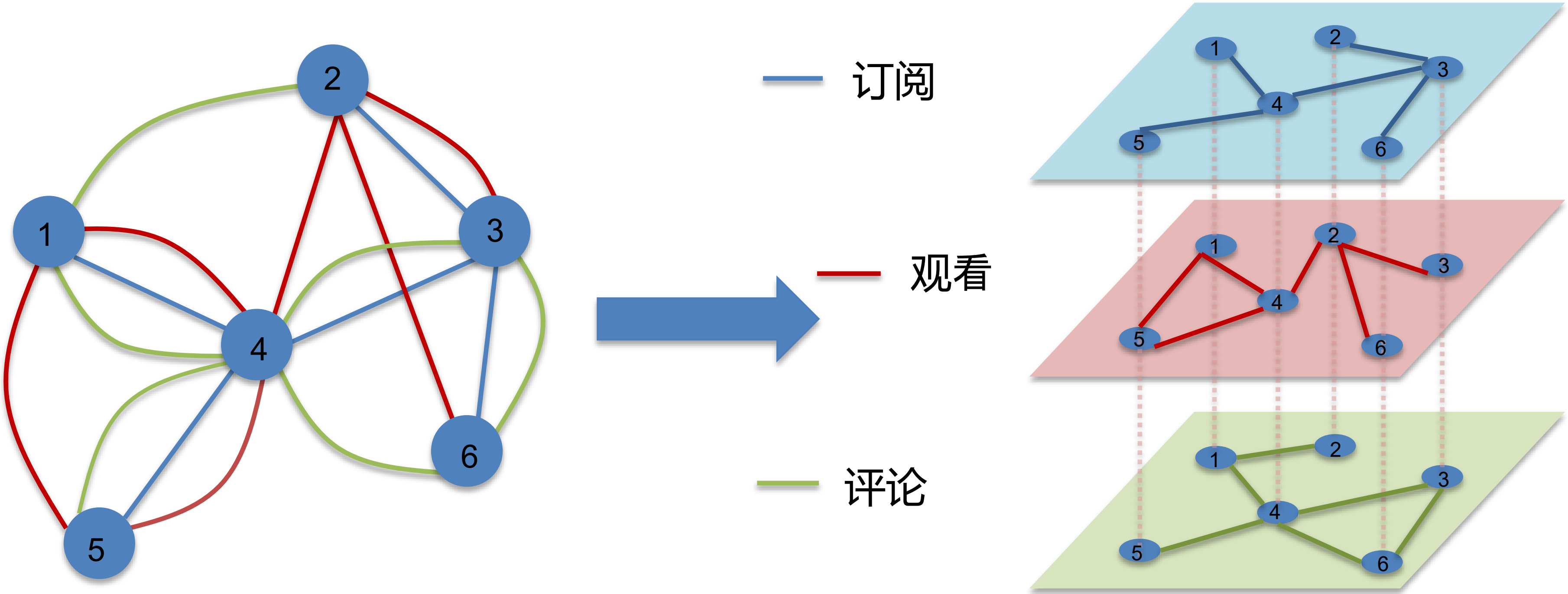
二分图



$\mathcal{V}, \mathcal{E}$

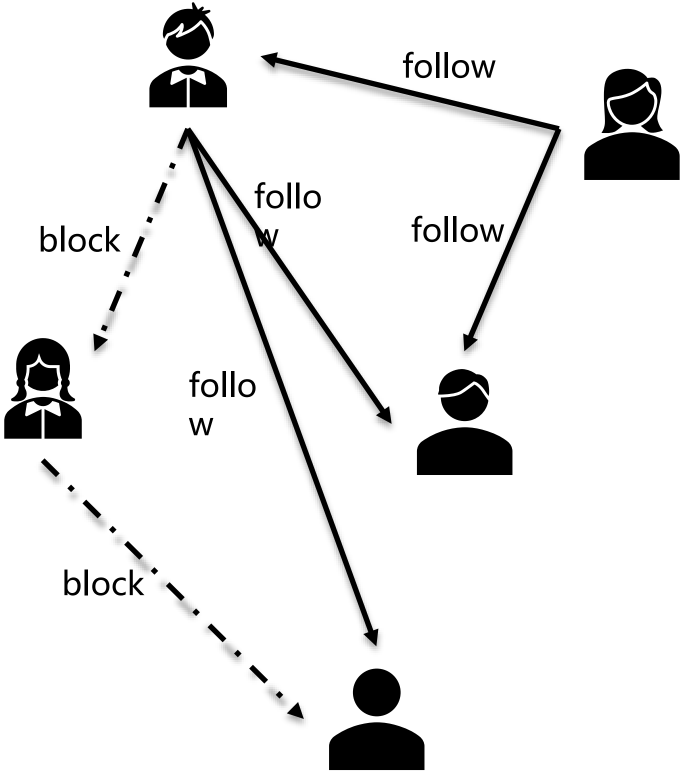
- ❑$\mathcal{V}$可以被分为两个不相交的子集
- ❑每一子集内的任意点之间没有边相连接

多维图



同一对节点之间可以同时存在多种不同的关系

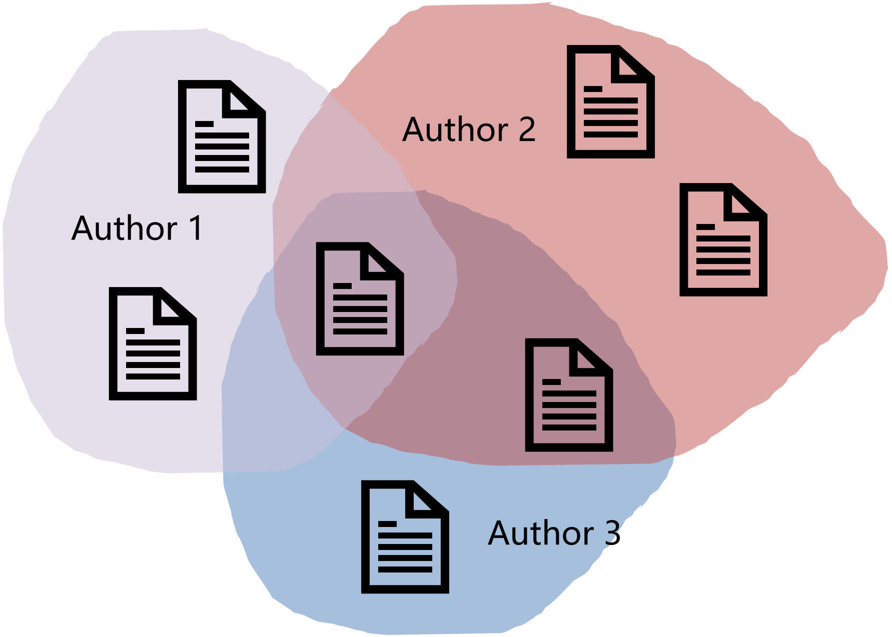
符号图



$\mathcal{V}, \mathcal{E}$

- $\mathcal{E}$可以被分为两个不相交的子集
- 这两个子集分别包含正边和负边

超图

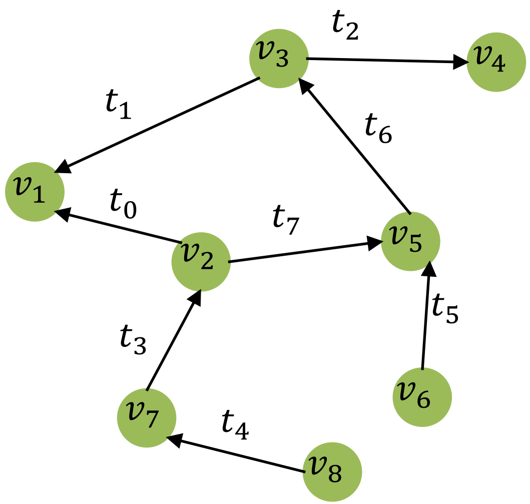


$\mathcal{V}, \mathcal{E}$

- $\mathcal{E}$为超边的集合

- 超边可以描述超越两两之间的关系的更高维度的关系

动态图



$$\mathcal{V}, \mathcal{E}$$

节点和边有对应的时间信息

$$\phi_n : \mathcal{V} \rightarrow \mathcal{T}$$

$$\phi_e : \mathcal{V} \rightarrow \mathcal{T}$$

$\mathcal{T}$表示时间的集合

目录

深蓝学院
shenlanxueyuan.com

感谢聆听！

Thanks for Listening