Uma introdução matricial à TEORIA DOS GRAFOS

IPRJ/UERJ

25 de junho de 2013

GRAFO

n nós (ou vértices) conectados (ou não) por *m* arcos (ou arestas).

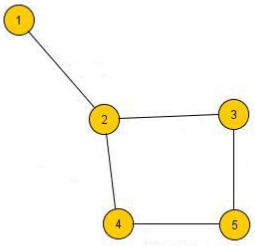


Figura : Um exemplo de grafo:
$G = (V, E), V = \{1, 2, 3, 4, 5\}, E = \{\{1, 2\}, \{2, 3\}, \{2, 4\}, \{3, 5\}, \{4, 5\}\}$

GRAFOS SIMPLES $\rightarrow$ grafo sem laços e sem arestas múltiplas.

GRAFO ORIENTADO $\rightarrow$ cada aresta tem um ponto inicial e um ponto final.

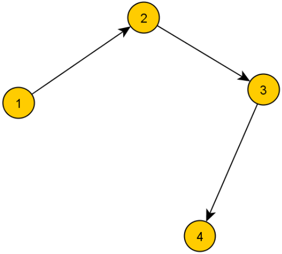


Figura : Um exemplo de grafo orientado:
$G = (V, E), V = \{1, 2, 3, 4\}, E = \{(1, 2), (2, 3), (3, 4)\}$

MATRIZ DE INCIDÊNCIA de um grafo orientado (A):

linha $i \rightarrow$ aresta i / coluna $j \rightarrow$ nó j

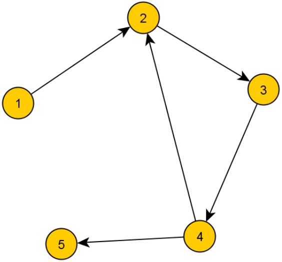
Aresta i:

Sai do nó j_1 : linha i de A terá -1 na coluna j_1
Chega no nó j_2 : linha i de A terá $+1$ na coluna j_2
Todas as outras posições serão completadas com zero.

Exemplo:

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

Qual a matriz de INCIDÊNCIA do grafo abaixo?



MATRIZ DE ADJACÊNCIA (W):

matriz $n \times n \to$ linhas e colunas $\Leftrightarrow$ nós do grafo

$$(W)_{ij} = w_{ij}$$

nós i e j conectados $\Leftrightarrow$ UM
nós i e j desconectados $\Leftrightarrow$ ZERO

$$W = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix}$$

GRAU DO NÓ i $\to$ número de arestas incidentes nesse nó.
MATRIZ GRAU DE UM GRAFO $\to$ cada $(D)_{ii}$ guarda o grau do nó i.

$$D = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}$$

MATRIZ LAPLACIANA (L) de um grafo $\rightarrow L = A^\top A$, onde A é a matriz de incidência de um grafo com n nós.

$$L = \begin{pmatrix} 1 & -1 & 0 & 0 \\ -1 & 2 & 0 & -1 \\ 0 & 0 & 1 & -1 \\ 0 & -1 & -1 & 2 \end{pmatrix}$$

OBS.: Com esta definição podemos provar que $L = D - W$, onde W é a matriz de adjacência e D é a matriz grau do grafo.

Definições para o caso mais geral

CASO GERAL

A cada aresta é atribuído um número real não negativo que será o seu peso.
Na matriz de adjacência W o número 1 será substituído pelo peso correspondente w_{ij} da aresta ligando os nós i e j.
Cada $(D)_{ii} = d_i$ será a soma dos pesos de todas as arestas incidentes ao nó i.

$$(D)_{ii} = d_i = \sum_{j=1}^{j=n} w_{ij}$$

Assim,
$W \rightarrow$ peso das arestas e
$D \rightarrow$ peso dos nós

MATRIZ LAPLACIANA COM PESOS (L)

Para arestas com pesos iguais tínhamos: $L = A^\top A \Rightarrow L = D - W$
Para arestas com pesos quaisquer queremos:
$L = A^\top CA \Rightarrow L = D - W$
$C \rightarrow$ matriz diagonal de ordem $m \times m$ (m é o número de arestas do grafo!) tal que cada número positivo $c_{kk} > 0$ guarda o peso w_{ij} da k -ésima aresta que liga os nós i e j.
Note que antes a matriz C nada mais era que a matriz identidade!

Quatro matrizes especiais

- $W_{n\times n}$ (*Weight*)$\rightarrow$ adjacência (peso das arestas) - linhas e colunas são os nós do grafo
- $D_{n\times n}$ (*Degree*)$\rightarrow$ grau - matriz diagonal, $d_i = (D)_{ii} = \sum_{j=1}^{n} w_{ij}$. representa o grau do nó i (soma dos pesos das arestas ligadas a ele).
- $L_{n\times n} \rightarrow$ laplaciana, $L = D - W$.
- $P_{n\times n} \rightarrow$ normalização de W, $P = D^{-1}W$

$$W \longrightarrow D \begin{array}{l} \nearrow \quad L = D - W \\ \\ \searrow \quad P = D^{-1}W \end{array}$$

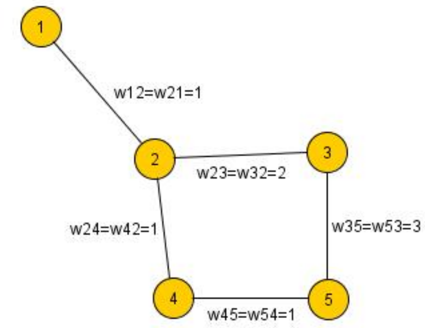
Vendo P de mais perto

$$P = D^{-1}W = \begin{pmatrix} \frac{w_{11}}{d_1} & \frac{w_{12}}{d_1} & \cdots & \frac{w_{1n}}{d_1} \\ \frac{w_{21}}{d_2} & \frac{w_{22}}{d_2} & \cdots & \frac{w_{2n}}{d_2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{w_{n1}}{d_n} & \frac{w_{n2}}{d_n} & \cdots & \frac{w_{nn}}{d_n} \end{pmatrix}$$

Para cada linha i de P : $\quad \sum_{j=1}^{n} p_{ij} = \sum_{j=1}^{n} \frac{w_{ij}}{d_i} = \frac{d_i}{d_i} = 1 \Rightarrow P\mathbb{1} = 1.\mathbb{1}.$

Como $P\mathbf{v} = \lambda_P \mathbf{v} \Rightarrow L\mathbf{v} = \lambda_L D\mathbf{v}$, onde $\lambda_L = 1 - \lambda_P$ então $L\mathbb{1} = 0.D\mathbb{1}$.

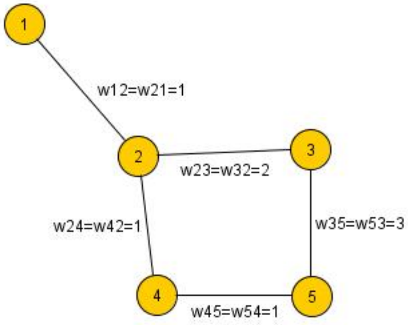
Um exemplo com grafo simples e conexo



$$W = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 2 & 1 & 0 \\ 0 & 2 & 0 & 0 & 3 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 3 & 1 & 0 \end{pmatrix}$$

$$D = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 & 0 \\ 0 & 0 & 5 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 4 \end{pmatrix}$$

L e P para o exemplo anterior



$$L = \begin{pmatrix} 1 & -1 & 0 & 0 & 0 \\ -1 & 4 & -2 & -1 & 0 \\ 0 & -2 & 5 & 0 & -3 \\ 0 & -1 & 0 & 2 & -1 \\ 0 & 0 & -3 & -1 & 4 \end{pmatrix}$$

$$P = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 1/4 & 0 & 2/4 & 1/4 & 0 \\ 0 & 2/5 & 0 & 0 & 3/5 \\ 0 & 1/2 & 0 & 0 & 1/2 \\ 0 & 0 & 3/4 & 1/4 & 0 \end{pmatrix}$$

Cadeia de Markov

Matriz de transição de uma cadeia de Markov $\rightarrow$
entradas não negativas e soma das linhas igual a um.

A matriz $P = D^{-1}W$ é uma matriz de Markov.

- $w_{ij} \rightarrow$ peso da aresta ligando X_i com X_j (afinidade).
- Núcleo gaussiano: simetria, positividade.
- Todo autovalor de uma matriz de Markov é $\mid \lambda \mid \leq 1$.
- A matriz de transição de passo t é igual a matriz de transição de passo 1 elevada a t-ésima potência.

$W \rightarrow$ processo de Markov $\rightarrow$ espectro $\rightarrow \phi$.

Processo de Markov

Estados $\rightarrow$ nós do grafo.

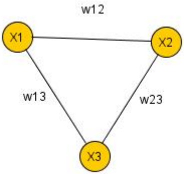
Dinâmica $\rightarrow$ sair de qualquer estado X_i para outro X_j.

Probabilidade de transição $\rightarrow$ elementos da matriz $P = D^{-1}W$.

Maior a afinidade (w_{ij}) entre os nós $\rightarrow$ maior a probabilidade de transição entre eles.

$$P = D^{-1}W = \begin{array}{c|cccc} & X_1 & X_2 & \dots & X_n \\ \hline X_1 & \frac{w_{11}}{d_1} & \frac{w_{12}}{d_1} & \dots & \frac{w_{1n}}{d_1} \\ X_2 & \frac{w_{21}}{d_2} & \frac{w_{22}}{d_2} & \dots & \frac{w_{2n}}{d_2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ X_n & \frac{w_{n1}}{d_n} & \frac{w_{n2}}{d_n} & \dots & \frac{w_{nn}}{d_n} \end{array}$$

Um exemplo



$$w_{12} = exp(\frac{\|X1-X2\|^2}{\sigma^2}); \quad w_{13} = exp(\frac{\|X1-X3\|^2}{\sigma^2}); \quad w_{23} = exp(\frac{\|X2-X3\|^2}{\sigma^2})$$

$$X_1 = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

$$X_2 = \begin{pmatrix} 0.9673 & 0.9511 \\ 0.9285 & 0.9414 \end{pmatrix}$$

$$X_3 = \begin{pmatrix} 0.9611 & 0.9709 \\ 0.9286 & 0.9869 \end{pmatrix}$$

Escolhi $\sigma^2 = 0.012$. Assim temos:

$$W = \begin{pmatrix} 1.0000 & 0.3679 & 0.5294 \\ 0.3679 & 1.0000 & 0.8116 \\ 0.5294 & 0.8116 & 1.0000 \end{pmatrix}$$

$$D = \begin{pmatrix} 1.8973 & 0 & 0 \\ 0 & 2.1795 & 0 \\ 0 & 0 & 2.3410 \end{pmatrix}$$

$$P = D^{-1}W = \begin{pmatrix} 0.5271 & 0.1939 & 0.2790 \\ 0.1688 & 0.4588 & 0.3724 \\ 0.2261 & 0.3467 & 0.4272 \end{pmatrix}$$

$P^t = (D^{-1}W)^t$ representa o estado do sistema no parâmetro t.

$$P^1 = \begin{pmatrix} 0.5271 & 0.1939 & 0.2790 \\ 0.1688 & 0.4588 & 0.3724 \\ 0.2261 & 0.3467 & 0.4272 \end{pmatrix}$$

$$P^2 = \begin{pmatrix} 0.3736 & 0.2879 & 0.3385 \\ 0.2506 & 0.3724 & 0.3770 \\ 0.2743 & 0.3510 & 0.3747 \end{pmatrix}$$

$$P^3 = \begin{pmatrix} 0.3221 & 0.3219 & 0.3560 \\ 0.2802 & 0.3502 & 0.3696 \\ 0.2886 & 0.3441 & 0.3673 \end{pmatrix}$$

$$P^4 = \begin{pmatrix} 0.3046 & 0.3336 & 0.3618 \\ 0.2904 & 0.3431 & 0.3665 \\ 0.2932 & 0.3412 & 0.3656 \end{pmatrix}$$

$$P^5 = \begin{pmatrix} 0.2987 & 0.3376 & 0.3638 \\ 0.2938 & 0.3408 & 0.3654 \\ 0.2948 & 0.3401 & 0.3650 \end{pmatrix}$$

$$P^6 = \begin{pmatrix} 0.2967 & 0.3389 & 0.3644 \\ 0.2950 & 0.3400 & 0.3650 \\ 0.2953 & 0.3398 & 0.3649 \end{pmatrix}$$

$$P^7 = \begin{pmatrix} 0.2960 & 0.3394 & 0.3647 \\ 0.2954 & 0.3397 & 0.3648 \\ 0.2955 & 0.3397 & 0.3648 \end{pmatrix}$$

$$P^8 = \begin{pmatrix} 0.2957 & 0.3395 & 0.3647 \\ 0.2956 & 0.3397 & 0.3648 \\ 0.2956 & 0.3396 & 0.3648 \end{pmatrix}$$

$$P^9 = \begin{pmatrix} 0.2957 & 0.3396 & 0.3648 \\ 0.2956 & 0.3396 & 0.3648 \\ 0.2956 & 0.3396 & 0.3648 \end{pmatrix}$$

$$P^{10} = \begin{pmatrix} 0.2956 & 0.3396 & 0.3648 \\ 0.2956 & 0.3396 & 0.3648 \\ 0.2956 & 0.3396 & 0.3648 \end{pmatrix}$$

Processo de Markov e distância de difusão

$p_{ij} = \frac{w_{ij}}{d_i} \rightarrow$ caminho aleatório - X_i para X_j - um passo.

$P^t = (D^{-1}W)^t \rightarrow$ aumenta para t passos (diferentes escalas).

$p^t_{ij} \rightarrow$ probabilidade de transição de X_i para X_j em t passos.

Quanto maior t mais o processo incorpora a geometria dos dados.

$D_t(X_i, X_j) \rightarrow$ taxa de conectividade entre X_i e X_j por caminhos de comprimento t. Muitos caminhos $\rightarrow$ "distância" pequena

$$D_t(X_i, X_j) = \sqrt{\sum_{X_r \in X} \frac{(p^t_{ir} - p^t_{jr})^2}{\sigma_r}},$$

onde $\sigma_r = \frac{d_r}{\sum_{i=1}^{n} d_i} \rightarrow$ densidade de grau do r-ésimo nó.

A chave do método

A aplicação de difusão ϕ deve preservar distâncias:

distância de difusão (X_i, X_j) $=$ distância euclidiana $(\phi(X_i), \phi(X_j))$.

Como fazer?

$\widetilde{P} = D^{1/2} P D^{-1/2}$ similar a $P = D^{-1} W \rightarrow$ mesmos autovalores.

Vantagem: $\widetilde{P}$ é simétrica $\rightarrow$ decomposição espectral.

Decomposição de $P^t = (D^{-1}W)^t \rightarrow p_{ij}^t = \sum_k \lambda_k^t \mathbf{v}_k(i) \mathbf{e}_k^\top(j)$

- $\lambda, \mathbf{v} \rightarrow$ autovalores e autovetores à direita de P^t
- $\mathbf{e}^\top \rightarrow$ autovetores à esquerda de P^t
- $\lambda_1 = 1 \geq \lambda_2 \geq \ldots \geq \lambda_n$.

Distância e aplicação de difusão

Reescrevendo a distância de difusão entre X_i e X_j :

$$D_t(X_i, X_j) = \sqrt{\sum_{k=2}^{n} \lambda_k^{2t}(\mathbf{v}_k(i) - \mathbf{v}_k(j))^2}.$$

Família de aplicações de difusão que preservam distâncias:

$$\phi_t : X_i \mapsto \begin{pmatrix} \lambda_2^t \mathbf{v}_2(i) \\ \lambda_3^t \mathbf{v}_3(i) \\ \vdots \\ \lambda_n^t \mathbf{v}_n(i) \end{pmatrix}$$

Por quê ϕ?

$$D_t(X_i, X_j) = (\sum_{r=2}^{n} \lambda_r^{2t}(\mathbf{v}_r(i) - \mathbf{v}_r(j))^2)^{1/2} = \| \phi_t(X_i) - \phi_t(X_j) \| .$$

Como $|\lambda| \leq 1$, se t grande $\Rightarrow \lambda_r^t$ insignificantes $\Rightarrow$ *k(t)* pequeno.

$$D_t(X_i, X_j) \cong (\sum_{r=2}^{k(t)} \lambda_r^{2t}(\mathbf{v}_r(i) - \mathbf{v}_r(j))^2)^{1/2} = \| \phi_t(X_i) - \phi_t(X_j) \|$$

com

$$\phi_t : X_i \mapsto \begin{pmatrix} \lambda_2^t \mathbf{v}_2(i) \\ \lambda_3^t \mathbf{v}_3(i) \\ \vdots \\ \lambda_{k(t)}^t \mathbf{v}_{k(t)}(i) \end{pmatrix}$$