March 3, 2020

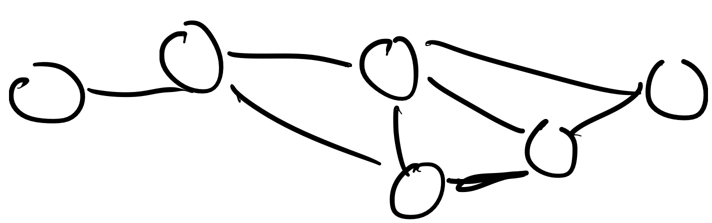
Class so far:

1. least squares → objective functions
   subroutine
2. optimization → SGD
3. dimensionality redux / matrix fact.
   PCA ⟷ spectral clustering
   latent factors ⟷ node representation learning

Next: "network science"

methods for graph data

Graph: set of nodes and edges

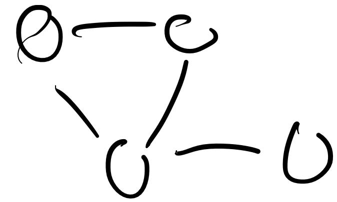


Example: FB social network

$$i \overset{\text{undir.}}{\text{———}} j \qquad\qquad i \xrightarrow{\text{dir.}} j$$

also network

network	node	edge	undir / dir
FB	users	friendship	undir
FB	users	message	dir
Cora	papers	citation	dir
Email	email addresses	sent email	dir
Coauthorship	scientists	coauthored paper	undir
Co-purchasing			
Co-viewing			

$\boxed{A}$ vs. 

Use relational structure:

1. Which nodes are important or central
2. How a network is organized (clustering)
3. How networks evolve?
4. How stuff spreads?
5. Predict properties about nodes?

Today:

1. matrices associated with graphs
2. common network structure

Adjacency matrix A

$$A_{ij} = \begin{cases} 1 & \text{if } (i) \longrightarrow (j) \text{ in graph} \\ 0 & \text{otherwise} \end{cases} \qquad \begin{array}{cc} & n \\ n & \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & \end{bmatrix} \end{array}$$

undirected: $A = A^T$

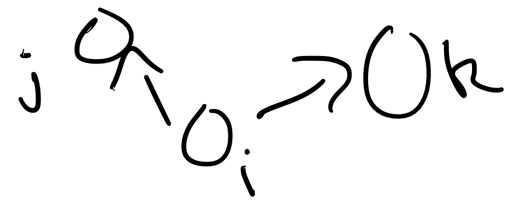
$$A^k_{ij} = \#\text{ of length-}k\text{ paths from } i \text{ to } j$$

$$\left[A^{k-1} \cdot A\right]_{ij} = \sum_{\ell} A^{k-1}_{i\ell} A_{\ell j}$$

weighted: $W_{ij} = \#$ of emails sent from i to j

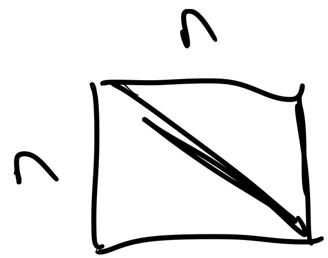
Diagonal degree matrix D

$d = A\mathbf{1}$ $\quad d_i = \sum_j A_{ij} 1$ = out-degree of i



$d_i = 2$

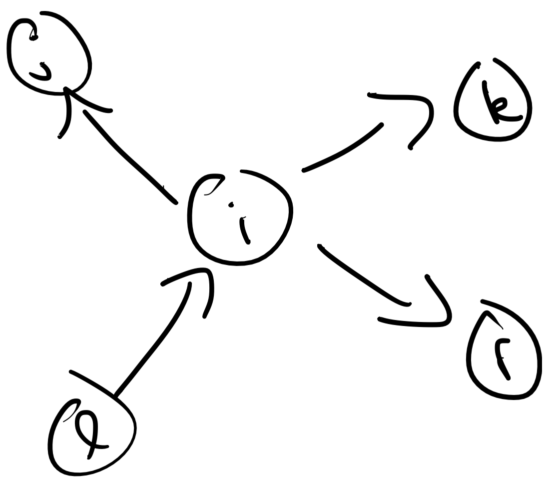
$D_{ii} = d_i$



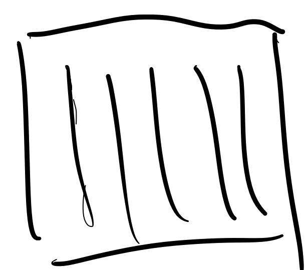
D

Random walk matrix P

$P = A^T D^{-1}$ $\quad P_{ji} = A^T_{ji} D^{-1}_{ii} = A_{ij}/d_i$



$P_{ri} = P_{ji} = P_{ki} = 1/3$



columns stochastic

Graph Laplacian L

(undir. graph) $G=(V,E)$

$L = D - A$

Claim: $x^T L x = \sum_{(i,j)\in E} (x_i - x_j)^2$

Cor: L symm. pos. semidef

$$= \sum_{(i,j)\in E} x_i^2 + x_j^2 - 2x_i x_j$$

$$= \sum_{i\in V} d_i x_i^2 - \sum_{(i,j)\in E} 2x_i x_j$$

$$= x^T D x - \sum_{1\le i,j\le n} A_{ij} x_i x_j$$

$$= x^T D x - x^T A x$$

$$= x^T L x$$

A, D, P, L

1. Matrix-vector matrix-matrix multiplication

$$z = P^T y = D^{-1} A y \quad z_i = \frac{1}{d_i} \sum_j A_{ij} y_j$$

2. Solve systems of equations:

PageRank: $(I - \alpha P) x = (1-\alpha) \mathbb{1}$

3. Compute eigenvals/eigenvecs

Spectral clustering: $Lx = \lambda x$

Basic spectral graph theory

Claim: number of connected components

= number of zero eigenvalues in L

Proof: $c_i^{(S)} = \begin{cases} 1 & i \in S \\ 0 & \text{o/w} \end{cases}$

$$c^T L c = \sum_{(i,j) \in E} (c_i - c_j)^2 = \sum_{i,j \in S} (1-1)^2 + \sum_{i,j \notin S} (0-0)^2 = 0$$

$$c^{(1)}, \ldots, c^{(k)} \qquad c^{(i)T} c^{(j)} = 0$$

$$0 = x^T L x = \underbrace{\sum_{(i,j) \in E} (x_i - x_j)^2}_{\text{constant if } i,j \text{ in same component}} \qquad x = \sum_j \alpha_j c^{(j)}$$

$$c^T L c = 0 \implies Lc = 0c$$

$$L = V \Lambda V^T$$

$$x^T L x = x^T V \Lambda V^T x \qquad y = V^T x$$

$$= y^T \Lambda y = \sum_i \lambda_i y_i^2 = 0$$

$$\implies y_i \neq 0 \Rightarrow \lambda_i = 0$$

$$V \begin{pmatrix} 0 & \\ & \ddots \end{pmatrix} V^T x$$

$$\frac{x^T L x}{x^T x} \qquad x \in \operatorname{span}\{v_1, \dots, v_k\}$$

$$L v_i = 0$$

Watts + Strogatz (98)

Milgram (1969)

296 "random" people

target person in Boston

64 successful

6 degrees of separation

diameter ≤ 6