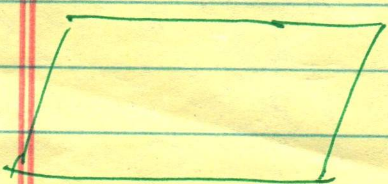
① — Part I

Lecture 16 ~~3~~ — rusty (~~sooo~~ "swarm" Dirt Poles)

Jan 18, 2024

Not mentioned / emphasized last time:

1. Writing
2. network data
3. libraries
4. Mark your calendar. 4/9 Paul's talk



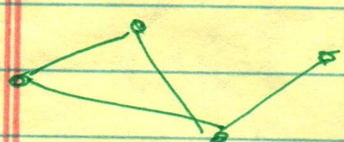
Macroscopic

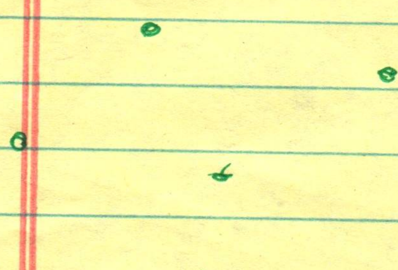
"large-scale"

structure

dynamics — function

how does structure constr. function.





microscopic

"local-level"

(multi-scale
coarse-grained.)

If grant writing !!

if you cannot find "relationships", maybe ~~it~~ is not a good idea

② – Part II (exclude)

"DAG"

Key Messages

– simple v.s. "non-simple" (directed, bipartite, w/ self-loop, multiedges [+ annotations] graph)
"existence of an edge"

– bipartite & it's "one-mode" projection.

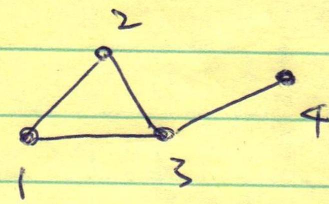
– ~~directed~~

– representation (most common: edge list; adjacency list / matrix (dense / sparse))

– $G = (V, E)$

$$\begin{cases} V = [n] = \{1, 2, \ldots, n\} \\ E \subseteq V \times V = \{(1,2), (2,3), (1,3), (1,4)\} \end{cases}$$

set — we don't care order, just scattered stuff



$$\mathbf{A} = \{A_{ij}\}_{\substack{i=1..n \\ j=1..n}}$$

bold in papers

$$= \{A_{ij}\}_{\substack{1 \le i \le n \\ 1 \le j \le n}} = (A_{ij})$$

meaning "we care about order"

tuple

degrees:

$$k_i = \sum_j A_{ij}$$

in this case

$n = 4$

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

$$m = \frac{1}{2} \sum_{ij} A_{ij} = 4.$$

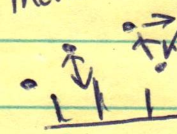
③ – Part III

} Graph-level properties (i.e. how we describe a graph)

"non-simple"

See things encircled ! On page 2.

Exploratory / mention



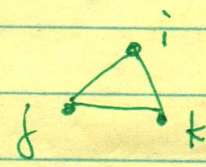
compute case

— ⊕ multiplex (or multilayed)
— ⊕ connected v.s. disconnected
— ⊕ temporal

(maybe interesting to ~~find~~ study "empirical distribution")

— k-cliques ,

$$\# \text{3-cliques} = \frac{1}{6} \sum_{ijk} A_{ij} A_{jk} A_{ki}$$



"Poly-time" for ~~#~~ counting # k-cliques.

"poly-~~sp~~ (memory) space"

communicate

possible ~~cons~~ constraints that makes a prob. hard.

"simultaneous existence of i, j, k"

6 = 3 !

— important : sparse v.s. dense

" ~~Most~~ (almost all) empirical networks are sparse "

perhaps / due to limited resources to maintain interaction, or other reasons.

$\langle d \rangle \sim$ "$\log N$"

Part 4 — ~~Watch slide~~

Intro. " ~~why we~~ learn from a graph "

What to

(how do we describe them?)

$$P(G \mid \underline{\theta}) = \prod_{i<j} P(A_{ij} \mid \underline{\theta}) = \prod_{i<j} P(i \to j \mid \underline{\theta})$$

or $= \prod_{i<j} p_{ij}$

Part 5

$f: G \to \{\theta_i\}$ 3. "explanatory analysis."

Part 4

if we have data G and some known properties $\{\theta_i\}$

2. "predictive analysis"

Ask ① students.

In science,
With data, what do we want to do?

"generalize" { explain / predict

With ~~theory~~

Observation-driven **model** (~~Inspired by observation~~)

---

theory-driven model

if network has x, y, z properties
then we can achieve X, Y, Z (dynamics, structure, function)

Part 5.

generate

(probabilistic) Model → Data

inference

algorithms.

ask: What are important feature that we want to capture.

e.g. "degrees" configuration model.

e.g. just one type in "Bayesian network" CSCI 5822

Finally: ~~Cert~~ 4. Causal ~~analysis~~ modeling

Part 6.
See slides!

Q: Why those 6 networks?
apparently by "tagging" or metadata.

thinking: "attachment styles"
PCA; 4 major types from survey.
OR "~~res~~ Carnegie classification"
R1, R2, R3

"natural way to define distances between graphs?"