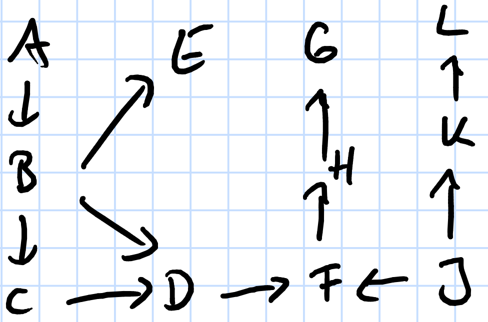
[1]

27.4.2022

Graphs



Def: A graph $G = (V, E)$ consists of finitely many nodes V and edges $E \subseteq V \times V$ with $\forall v \in V$ $(v,v) \notin E$. We mostly work with directed graphs, i.e. $(v,w) \in E \Rightarrow (w,v) \notin E$

Def: (for details, see Sec. 6.1)

- path: $D \rightarrow F \leftarrow J \rightarrow K$
- collider on a path
- dir. path: $B \rightarrow D \rightarrow F$
- v-structure: $D \rightarrow F \leftarrow J$ and D and J not dir. connected $\Rightarrow$ "v-structure"
- dir. cycle $X \rightarrow Y \rightarrow Z \rightarrow W$ and $W \rightarrow X$
- dir. acyclic graph (DAG): no dir. cycles
- PA
- CH
- DE · ND: non-desc.
- AN
- $V = DE_D \,\dot\cup\, ND_D \,\dot\cup\, \{D\}$

disjoint union

- Slight abuse of notation: denote sets & vectors by the same letter.

- d-separation

Let $\underline{X}, \underline{Y}, \underline{Z}$ be disjoint.

(ii) $\underline{X}, \underline{Y}$ d-connected given/by $\underline{Z}$ iff $\exists\ X \in \underline{X},\ Y \in \underline{Y}$ s.t. $\exists$ path from X to Y that is not blocked by $\underline{Z}$.

(i) A path $X = i_1, \dots, i_m = Y$ is blocked by $\underline{Z}$ iff

$\exists$ node i_k s.t. $i_{k-1} \to i_k \to i_{k+1}$ and $i_k \in \underline{Z}$

OR $\exists$ —— " —— $i_{k-1} \leftarrow i_k \leftarrow i_{k+1}$ —— " ——

OR $\exists$ —— " —— $i_{k-1} \leftarrow i_k \to i_{k+1}$ —— " ——

OR $\exists$ collider on the path, i.e., a node i_k s.t. $i_{k-1} \to i_k \leftarrow i_{k+1}$ s.t.

$$i_k \notin \underline{Z} \quad \text{and} \quad DE_{i_k} \cap \underline{Z} = \emptyset.$$

(iii) if not d-connected: d-separated.

We sometimes write

$$\underline{X} \text{ d-sep } \underline{Y} \mid \underline{Z} \quad \text{or} \quad \underline{X} \perp\!\!\!\perp_G \underline{Y} \mid \underline{Z}.$$

$\boxed{2}$

Q: Given a DAG G over A, B, C, D. Assume that all (!) d-sep. statements are

- A d-sep C
- $\{A, C\}$ d-sep $D \mid B$
- C d-sep $D \mid \{A, B\}$
- A d-sep $D \mid \{B, C\}$

+ symmetrics & trivially impl. statements.

What is G?

$$\begin{tikzcd} A \arrow[dr] & & C \arrow[dl] \\ & B \arrow[r] & D \end{tikzcd}$$

Rem: In a DAG G, there is an edge between X and Y $\Leftrightarrow$ $\nexists\, \underline{Z}$ s.t. X d-sep. $Y \mid \underline{Z}$.

Proof: "$\Rightarrow$" trivial, "$\Leftarrow$", see A1.

$\boxed{3}$

Rem: We can repr. $G = (V, E)$ by a $d \times d$ 0-1-matrix A. $A_{jk} = 1 \Leftrightarrow (j,k) \in E$. (adjacency matrix). We have

(i) A^2 encodes number of dir. paths of length 2. E.g.,

$$A^2_{j,k} = 1 \Leftrightarrow \exists!\, \ell : (j,\ell), (\ell,k) \in E$$

(ii) In general: $A^m \cong$ # dir. paths of length m.

(iii)

(iv) Think about using sparse matrices.

Thm: (McKay 2004, not difficult)

There is a bijection between DAGs over d nodes and $d \times d$ 0-1 matrices w/ eigenvalues > 0:

A adj. matrix of DAG $\Leftrightarrow$ $A + \mathrm{id}$ 0-1 matrix w/ eigenvalues > 0.

Structural Causal Models

goal:

counterfactuals ← SCM → obs. distribution

SCM → causal graph

SCM → int. distribution

Def: An SCM $\mathfrak{C}$ (sometimes SEM) over d RVs $\underline{X} = (X_1,\dots,X_d)$ consists of d assignments

$$(*) \qquad X_j := f_j(PA_j, N_j), \qquad j\in\{1,\dots,d\}$$

and a noise distr. $\mathbb{P}_{\underline{N}} = \mathbb{P}_{N_1}\otimes\dots\otimes\mathbb{P}_{N_d}$. Here $PA_j \subseteq \{X_1,\dots,X_d\}$. We define a corresponding graph over $X_1,\dots,X_d$ by drawing edges from var's on the rhs to var's on the lhs. This graph is assumed to be acyclic.

Random vector $(\underline{X},\underline{N})$ is a solution to the SCM if $\mathbb{P}_{\underline{N}}$ is the distr. of $\underline{N}$ (this implies that $N_1,\dots,N_d$ are jointly indep.) and (*) hold a.s.

For us: "SCM" $\hat{=}$ SCM + solution.

<u>Ex:</u> $X := N_X$ altitude

$Y := 15 - 6X + N_Y$ temp. $\quad X \to Y$

$N_X, N_Y \overset{iid}{\sim} \mathcal{N}(0,1)$

$$\begin{pmatrix} X \\ Y \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} 0 \\ 15 \end{pmatrix}, \begin{pmatrix} 1 & -6 \\ -6 & 37 \end{pmatrix} \right)$$

The SCM entails a unique distr. $\mathbb{P}_{\underline{X}}^{\mathfrak{C}}$ (or $\mathbb{P}_{\underline{X}}$) over $X_1, \dots, X_d$. In other words: all sol'n have the same marginal distr over $\underline{X}$. Applying the assign. iteratively yields

$$\begin{pmatrix} X_1 \\ \vdots \\ X_d \end{pmatrix} = \begin{pmatrix} \tilde{f}_1(N_{\{k : \pi(k) < \pi(1)\}}, N_1) \\ \vdots \\ \tilde{f}_d(N_{\{k : \pi(k) < \pi(d)\}}, N_d) \end{pmatrix}$$

noise var's appearing "before" X_d.

<u>Rem:</u> (i) Formal treatment

Bongers et al arXiv: 1611.06221v6

(ii) We do not distinguish between

$Y := 0 \cdot X + Z + N_Y$ and

$Y := Z + N_Y$

We can define an equiv. relationship and choose the latter as representative.

(iii) cycles: AZ.

Interventions

Consider an SCM $\mathfrak{C}$ over $\underline{X}$ w/ $\mathbb{P}_{\underline{N}}$. We can obtain a new SCM $\tilde{\mathfrak{C}}$ by replacing some of the assignments. E.g., replace

$$X_5 := 4X_7^2 + \sin(X_9) + N_5 \quad \text{by}$$
$$X_5 := 4X_7^2 - X_1 \quad + \tilde{N}_5$$

⚠ Do not introduce cycles & the new noise distr. (here: $N_1, \dots, N_4, \tilde{N}_5, N_6, \dots$) need to be indep.

We write: $\tilde{\mathfrak{C}} = \mathfrak{C};\, do(X_5 := 4X_7^2 - X_1 + \tilde{N}_5)$

and $\mathbb{P}_{X_8}^{\tilde{\mathfrak{C}}} = \mathbb{P}_{X_8}^{\mathfrak{C};\, do(X_5 := \underline{\qquad\qquad}, \underline{\qquad})}$

interventional distribution.

<u>Ex</u>: $\mathfrak{C} := \begin{array}{l} X := N_X \\ Y := 15 - 6X + N_Y \end{array}$ $\quad N_X, N_Y \overset{iid}{\sim} \mathcal{N}(0,1)$

$$\Rightarrow \begin{pmatrix} X \\ Y \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} 0 \\ 15 \end{pmatrix}, \begin{pmatrix} 1 & -6 \\ -6 & 37 \end{pmatrix} \right)$$

$$\mathbb{P}_Y^{\mathfrak{C};\, do(X:=3)} = \mathcal{N}(-3, 1) \neq \mathbb{P}_Y^{\mathfrak{C}}$$

$$\mathbb{P}_Y^{\mathfrak{C};\, do(X:=4)} = \mathcal{N}(-9, 1)$$

$$\mathbb{P}_X^{\mathfrak{C};\, do(Y:=-10)} = \mathcal{N}(0,1) = \mathbb{P}_X^{\mathfrak{C}}$$ "intervening on the effect".