Causality: Mathematical Language.

Source: (a) Elements of Causal Inference, Peters et al., Chap 6.
(b) Causality, J. Pearl, 2009. (aka "text")

Structural Causal Model (SCM):

$(X_1, X_2, \dots, X_d)$ a collection of random variables, defined by the SCM $\mathfrak{C}$:

structural assignment

$$X_j := f_j(PA_j, N_j) \quad , \; j = 1, 2, \dots, d.$$

where $PA_j \equiv$ Parents of $X_j \equiv$ subset of $\{X_1, .., X_d\} \setminus X_j$.

$f_j \equiv$ deterministic function.
$N_j \equiv$ random variable, with
$\{N_1, N_2, \cdots, N_d\}$ mutually independent

We can represent the structure of the above SCM through a directed graph $\mathcal{G} = (V, E)$, with vertices corresponding to the random variables, and directed edges incident on node j (corresponds to X_j) from each of its parents, i.e., PA_j. We will assume throughout that the directed graph is acyclic. (For cyclic graphs, see text)

Further, we denote the prob. measure induced on $(X_1, \dots, X_d)$ by $\mathbb{P}^{\mathfrak{C}}(\cdot)$, and the joint dist. on them by $p^{\mathfrak{C}}(\cdot)$. We will abuse notation to use the same for marginals as well.

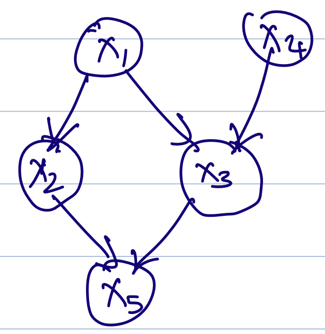
e.g: $p^{\mathfrak{C}}(x_1, \dots, x_d)$, $p^{\mathfrak{C}}(x_1)$, $p^{\mathfrak{C}}(x_3, x_4)$ are the joint and appropriate marginals.

$\boxed{\text{Example}}$

SCM — Structural Causal Model

$$X_1 := f_1(N_1)$$
$$X_2 := f_2(X_1, N_2)$$
$$X_3 := f_3(X_1, X_4, N_3)$$
$$X_4 := f_4(N_4)$$
$$X_5 := f_5(X_2, X_3, N_5)$$

DAG → Directed Acyclic Graph.



$\{N_1, N_2, N_3, N_4, N_5\}$ mutually independent

Remarks:

1. Parents of X_j sometimes called "direct causes" of X_j
2. Assignment vs Equation: The SCM consists of assignments, meaning that if we have the ability to

replace one or more of the assignments, i.e.,

$$X_j := f_j(PA_j, N_j)$$

$\downarrow$ replaced by (aka intervention).

$$X_j := \tilde{f}_j(\tilde{PA}_j, \tilde{N}_j),$$

with $\tilde{N}_j$ indep. of $(N_1, N_2, \dots, N_d)$.

3. SCM $\mathfrak{C}$ often denoted by $\mathfrak{C} = (S, P_N)$, where S denotes the structural equations, and P_N is the (product) measure over $(N_1, \dots, N_d)$.

Prop. (6.3 in text): An SCM $\mathfrak{C}$ induces a unique joint dist. over $(X_1, \dots, X_d)$. (This is often called the Entailed Distribution)

Pf: follows immediately from acyclicity.

Prop (6.31) (SCM $\Rightarrow$ Markov Property)

$p^{\mathfrak{C}}(\cdot)$ a joint dist. induced by an SCM $\mathfrak{C}$, with graph G. Then, $p^{\mathfrak{C}}(\cdot)$ is Globally Markov wrt. G.