Deconvoluting causal networks

Gibran Hemani

2017-01-03

Introduction

Growth in summary data for GWASs on phenotypes is on a steep trajectory, which is asymptoting twoards a situation where we can use two-sample MR to test 'everything against everything'. This promise, however, requires methodological advancements to address several issues including:

- Multiple testing
- Decomposing many indirect effects into a terse set of direct effects

The latter can be summarised as follows. Suppose there are five variables of interest, 1-5, and the causal relationships are

```
1 -> 2
2 -> 3
3 -> 4
4 -> 5
```

This can be depicted in graph form as in Figure 1.

If, however, we performed MR of `1 -> 3`, `1 -> 4`, etc, we would identify associations because they exist indirectly. Hence, after testing everything against everything our graph would look like Figure 2.

The task is to decompose the complete set of associations into the direct effects only. This has the following advantages:

- Identify direct pathways through which a particular exposure influences an outcome
- Identify instances of partial mediation, which might suggest that there are unknown variables that remain to be uncovered that mediate the path from exposure to outcome

MR for mediation has already been developed, especially for the case where there are only three phenotypic variables. Here, to obtain the direct effect of trait 1 on trait 2, $\beta_{1\Rightarrow 2}$ it is obtained as a function of all possible indirect paths:

$$\beta_{1\Rightarrow 2} = \beta_{1\rightarrow 2} - \beta_{1\rightarrow 3}\beta_{2\rightarrow 3}$$

Problems:

- Cycles
- Impossible associations
- Non-gaussian effects

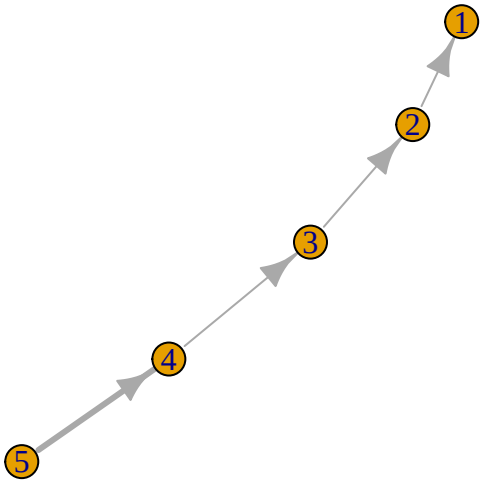

Figure 1: Simulated causal relationships

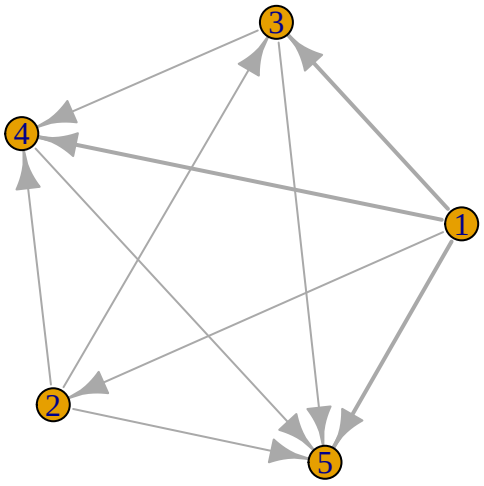

Figure 2: Empirical associations