

1 Analysis

Analysis of this project will mainly consist of comparison to the implementation of the algorithms in R. It will compare the functionality of the algorithms on data and also the run speed of the code.

1.1 Skeleton Generation

The skeleton generation code has almost identical functionality as the R implementation. Even the order in which edges are tested is the same. However the R implementation has a non-order dependent mode which is not currently implemented in the python version however could be added as an option in the future.

The Alarm_1000, Asia_1000, and Diabetes_1000 datasets found in all produce identical skeletons as seen in Figure .

The χ^2 test for independence was not implemented in the pcalg library in R. Other independence test for discrete data, such as the G^2 test, were implemented. However, the bnlearn library has an implementation of the χ^2 test to use for comparison.

Both tests return identical results on all tested data as seen in Table .

1.2 PC Algorithm Edge Alignment

1.3 FCI Algorithm Edge Alignment

1.4 Run Speed

1.4.1 Algorithms

1.5 Skeleton Generation

1.5.1 Independence Tests

The independence test was the main bottleneck of the algorithms, and was significantly slower than the R implementation. This is because the R implementation actually runs C code behind the scenes, c code is compiled rather than interpreted like python and R. This makes the code much faster.

However the independence test in the algorithm is a parameter and thus the current slow test could be replaced by a faster version which calls C code like the R implementation.