Yuxuan Wang, December 2020, solution for final project Q5.

5 Structure Learning with PC algorithm

5.1 Implement the functions(PC1,PC2,PC3)

5.1.1 Implement PC1

To implement PC1, A graph with no edge is initiated. Then for any subset $\mathbf{Z} \subset \mathbf{V}\backslash\{X,Y\}$, CI test is executed. An undirected edge between $\{X,Y\}$ is added if for all $\mathbf{Z}$, the results of CI tests are zeros.

5.1.2 Implement PC2

To implement PC2, A complete graph is initiated. Then the neighbors of $\{X,Y\}$ are checked by executing CI test. An edge between $\{X,Y\}$ is removed if $CI(X,Y|\mathbf{Z}) = 1$, where $\mathbf{Z} \subset N(X)_{\hat{\mathcal{G}}} \cup N(Y)_{\hat{\mathcal{G}}}\backslash\{X,Y\}$.

5.1.3 Implement PC3

PC3 is similar to PC2 but start from the moral graph instead of the complete graph. In moral graph, two variables are connected with an edge if they are in each other's "Markov Blanket (Mb)", i.e. they are neighbors or they share a common child.

5.2 Results of three methods

For D1, there are 5 viarables in the dataset. In order to learn the structure behind these varibles, PC1, PC2 and PC3 are implemented seperately. It can be observed that the three results are same.
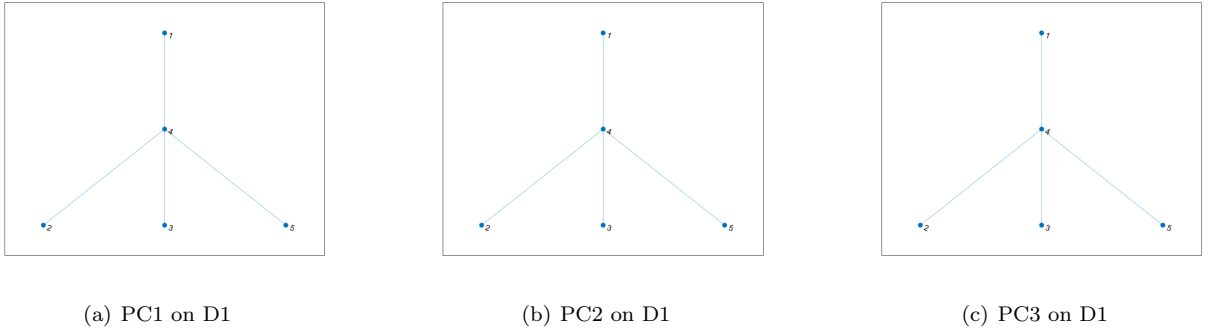

(a) PC1 on D1   (b) PC2 on D1   (c) PC3 on D1

Figure 1: Structure learning on D1

Figure2 shows the results of structure learing on D2.

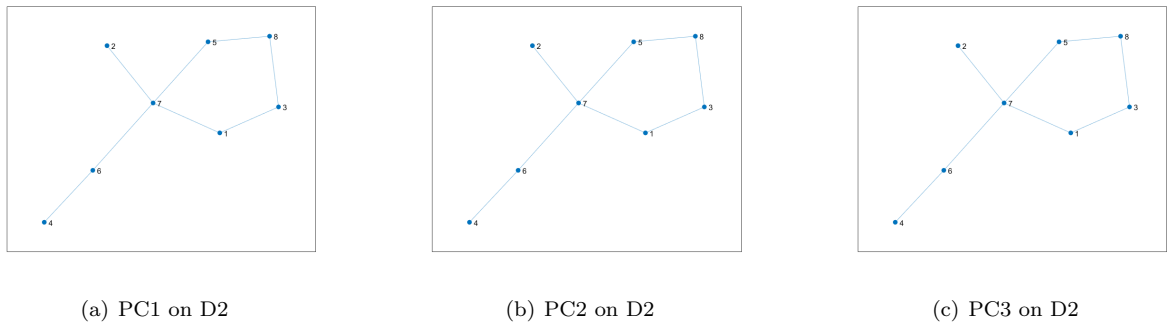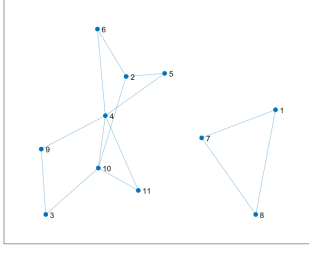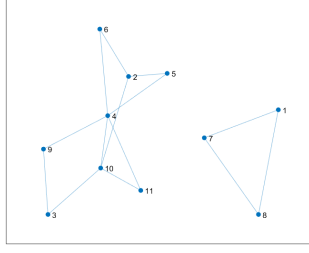(a) PC1 on D2   (b) PC2 on D2   (c) PC3 on D2
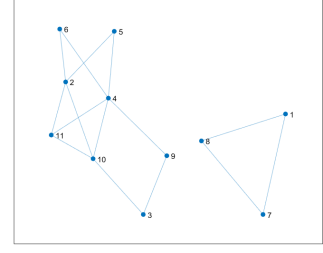
Figure 2: Structure learning on D2

Figure2 shows the results of structure learing on D2. The major difference is that there is an edge between 2 and 11 in the result of PC3.

(a) PC1 on D3

(b) PC2 on D3

(c) PC3 on D3

Figure 3: Structure learning on D2

5.3 Comparison on three methods

To compare the efficiency of the three methods, the number of CI tests have been recorded in the following table. When the data dimension is small, the number of CI tests are quite close. However, as the dimension of data increases, PC3 is the most efficient. PC1 takes most computational resourses.

Table 1: Times to run CI tests

	PC1	PC2	PC3
D1	51	47	45
D2	607	376	113
D3	7501	3079	368