Simulation Based Methods for Network Inference

Marthyna Luiza WEBER

Grenoble INP - Ensimag

2023-05-15

Initial definitions

Initial definitions

1. Partial Correlation

Correlation coefficient between X_1 and X_2 after removing the influence of Y, accounting for the scaling of the variables.

- Partial covariance: The partial covariance between X_1 and X_2 with reference to Y is calculated with:

$$Cov(X_1, X_2 \cdot Y) = \mathbb{E}[(X_1 - \hat{X}1(Y))(X_2 - \hat{X}2(Y))]$$

- $\hat{X}_l(Y)$ is the projection of X_l on Y, which is the expected value of X_l given Y:

$$\hat{X}_l(Y) = \mathbb{E}(X_l) + \frac{Cov(X_l, Y)}{Var(Y)}(Y - \mathbb{E}(Y))$$

1. Partial correlation

- Lemma:

$$Cov(X_l, X_m \cdot Y) = Cov(X_l, X_m) - Cov(X_l, Y)Var(Y)^{-1}Cov(X_m Y)^T$$

Given a vector $\mathbf{X} = (X_1, \ldots, X_p)$ of p random variables. We have:

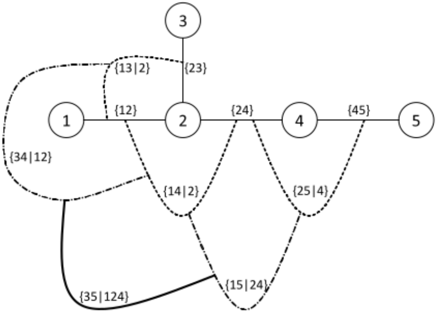
$$Cov^{\text{partial}}(\mathbf{X}_{l,m}) = Cov(X_l, X_m \cdot X_{V\setminus\{l,m\}})$$

- Partial correlation:

$$\rho^{\text{partial}}(\mathbf{X}_{l,m}) = \frac{Cov(X_l, X_m \cdot X_{V\setminus\{l,m\}})}{\sqrt{Var(X_l \cdot X_{V\setminus\{l,m\}})Var(X_m \cdot X_{V\setminus\{l,m\}})}}$$

2. Regular vines

A vine in which two edges in tree T_i are joined by an edge in tree T_{i+1} only if these edges share a common node.

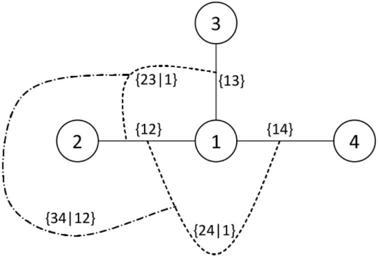


- Constraint set U_e: variables reachable from edge e
- Conditioning set D_e: variables shared between U_e and edge adjacent to e in the next tree
- Conditioned set $\{C_{1e}, C_{2e}\}$: symmetric difference of U_e and D_e

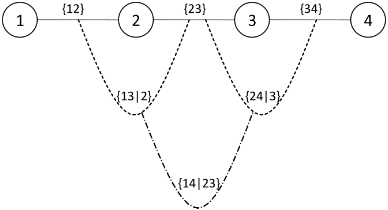
- $\{L|K\}$: the constraint set, with conditioned set L and conditioning set K.

4. C-vines

- C-vines: vines where each tree T_i has a unique node of degree $d - i$
- D-vines: vines where each node in T_i has a degree at most 2.



A C-vine



A D-vine

5. Generatating random correlation matrices with C-vines

Algorithm to generate a random correlation matrix $\boldsymbol{R}$ with density proportional to $det(\boldsymbol{R})^{\eta-1}$, with $\eta > 1$:

1. Initialize $\beta = \eta + \frac{d-1}{2}$
2. Loop for $k = 1, \ldots, d-1$:
   - 2.1 $\beta = \beta - \frac{1}{2}$
   - 2.2 Loop for $i = k+1, \ldots, d$:
     - 2.2.1 generate $p_{k,i;1,\ldots,k-1} \sim \text{Beta}(\beta, \beta)$ on $(-1, 1)$
     - 2.2.2 use the recursive formula for partial correlations calculation
3. Return $\boldsymbol{R}$, a $d \times d$ correlation matrix