

Project 1

Richard Affolter

25 02 2021

Problem 1: Conditional independence and BNs

- We can check for conditional independence $A \perp B \mid C$ if $P(A, B \mid C) = P(A \mid C) P(B \mid C)$.
- We can check for marginal independence by marginalizing with respect to C:
 $A \perp B$ if $P(A, B) = \sum_C P(A, B, C) = P(A) P(B)$

a)

$$P(A, B, C) = P(A \mid C) P(B \mid C) P(C)$$

$$\begin{aligned} P(A, B \mid C) &= \frac{P(A, B, C)}{P(C)} \\ &= P(A \mid C) P(B \mid C) \\ \Rightarrow A \perp B \mid C \end{aligned}$$

$$\sum_C P(A, B, C) = \sum_C P(A \mid C) P(B \mid C) P(C)$$

in general this does not factorize to the product $P(A) P(B)$.

We can make this more visible by using Bayes rules:

$$\begin{aligned} \sum_C P(A, B, C) &= \sum_C P(A \mid C) P(B \mid C) P(C) \\ &= \sum_C \frac{P(C \mid A) P(A)}{P(C)} \frac{P(C \mid B) P(B)}{P(C)} P(C) \\ &= P(A) P(B) \sum_C \frac{P(C \mid A) P(C \mid B)}{P(C)} \end{aligned}$$

Therefore for this network we have conditional independence but no marginal independence.

b)

$$P(A, B, C) = P(A) P(B) P(C \mid A, B)$$

$$\begin{aligned}
P(A, B \mid C) &= \frac{P(A, B, C)}{P(C)} \\
&= \frac{P(A) P(B) P(C \mid A, B)}{P(C)} \\
&\neq P(A \mid C) P(B \mid C)
\end{aligned}$$

$$\begin{aligned}
&\sum_C P(A, B, C) \\
&= \sum_C P(A) P(B) P(C \mid A, B) \\
&= P(A) P(B) \sum_C P(C \mid A, B) \\
&= P(A) P(B) \\
&\Rightarrow A \perp B
\end{aligned}$$

Therefore for this network we have marginal independence but no conditional independence.

Problem 2: Markov blanket

$$\text{MB}(D) = \{B, F, E, G, C\}$$

$$\begin{aligned}
&P(D \mid A, B, C, E, F, G) \\
&= \frac{P(A, B, C, D, E, F, G)}{P(A, B, C, E, F, G)} \\
&= \frac{P(A, B, C, D, E, F, G)}{\sum_D P(A, B, C, D, E, F, G)} \\
&= \frac{P(A)P(B)P(F)P(E \mid A)P(D \mid B, F)P(G \mid E, D)P(C \mid D)}{\sum_D P(A)P(B)P(F)P(E \mid A)P(D \mid B, F)P(G \mid E, D)P(C \mid D)} \\
&= \frac{P(A)P(B)P(F)P(E \mid A)P(D \mid B, F)P(G \mid E, D)P(C \mid D)}{P(A)P(B)P(F)P(E \mid A) \sum_D P(D \mid B, F)P(G \mid E, D)P(C \mid D)} \\
&= \frac{P(D \mid B, F)P(G \mid E, D)P(C \mid D)}{\sum_D P(D \mid B, F)P(G \mid E, D)P(C \mid D)}
\end{aligned}$$

$$\begin{aligned}
P(D \mid \text{MB}(D)) &= P(D \mid B, F, E, G, C) \\
&= \frac{P(B, C, D, E, F, G)}{P(B, C, E, F, G)} \\
&= \frac{\sum_A P(A, B, C, D, E, F, G)}{\sum_A \sum_D P(A, B, C, D, E, F, G)} \\
&= \frac{\sum_A P(A)P(B)P(F)P(E \mid A)P(D \mid B, F)P(G \mid E, D)P(C \mid D)}{\sum_A \sum_D P(A)P(B)P(F)P(E \mid A)P(D \mid B, F)P(G \mid E, D)P(C \mid D)} \\
&= \frac{P(B)P(F)P(D \mid B, F)P(G \mid E, D)P(C \mid D) \sum_A P(A)P(E \mid A)}{P(B)P(F) \sum_A P(A)P(E \mid A) \sum_D P(D \mid B, F)P(G \mid E, D)P(C \mid D)} \\
&= \frac{P(D \mid B, F)P(G \mid E, D)P(C \mid D)}{\sum_D P(D \mid B, F)P(G \mid E, D)P(C \mid D)} \\
&\Rightarrow P(D \mid A, B, C, E, F, G) = P(D \mid \text{MB}(D))
\end{aligned}$$

Data Analysis Problem 3: Learning Bayesian networks from protein data

read in the data provided in sachs.data.txt

```
sachs.data = read.table(file = "sachs.data.txt", header = TRUE)
```

(a) learn structure

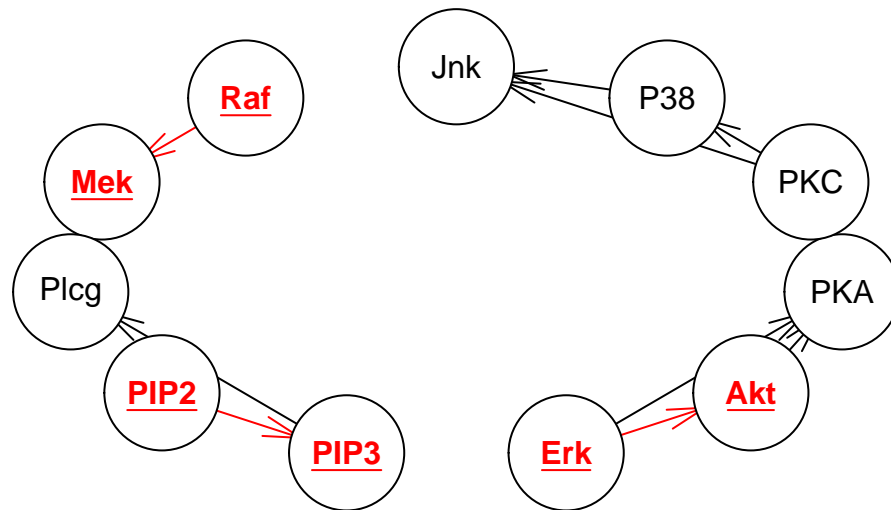
```
library(bnlearn)

# hill-climbing method
bn.hc = hc(sachs.data)

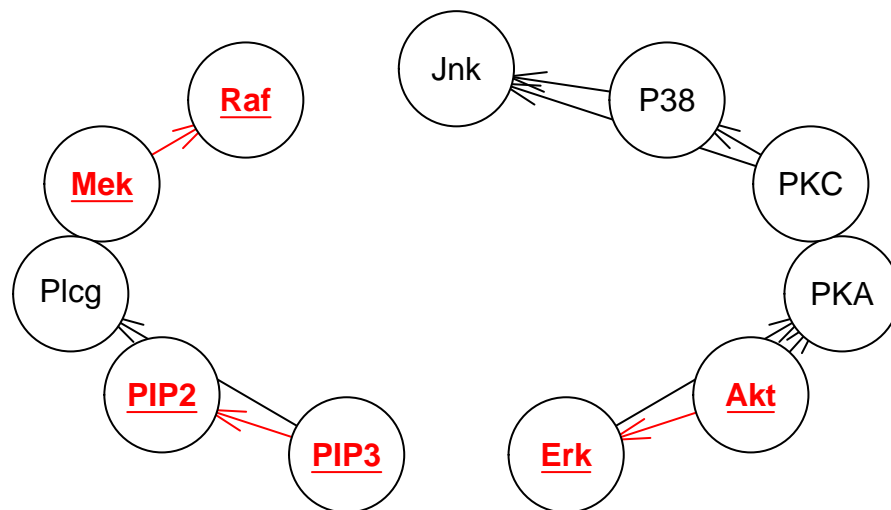
# tabu-climbing method
bn.tabu = tabu(sachs.data)

# plotting and highlighting differences
par(mfrow = c(2, 1))
differences = c("Raf", "Mek", "PIP2", "PIP3", "Erk", "Akt")
plot(bn.hc, main = "Hill-Climbing", highlight = differences)
plot(bn.tabu, main = "Tabu-search", highlight = differences,
      sub = expression(
        "side by side comparison of the Bayesian network structures learned by\n",
        "hill-climbing and tabu-search algorithms.\n",
        "Differences between the networks are highlighted in red"
      )
)
```

Hill-Climbing



Tabu-search



side by side comparison of the Bayesian network structures learned by hill-climbing and tabu-search algorithms.
Differences between the networks are highlighted in red

We can see that there are differences between the networks. The directions of multiple edges are reversed between the upper and the lower graph.

(b) Estimate conditional densities

```
hc.fit = bn.fit(bn.hc, sachs.data)
tabu.fit = bn.fit(bn.tabu, sachs.data)

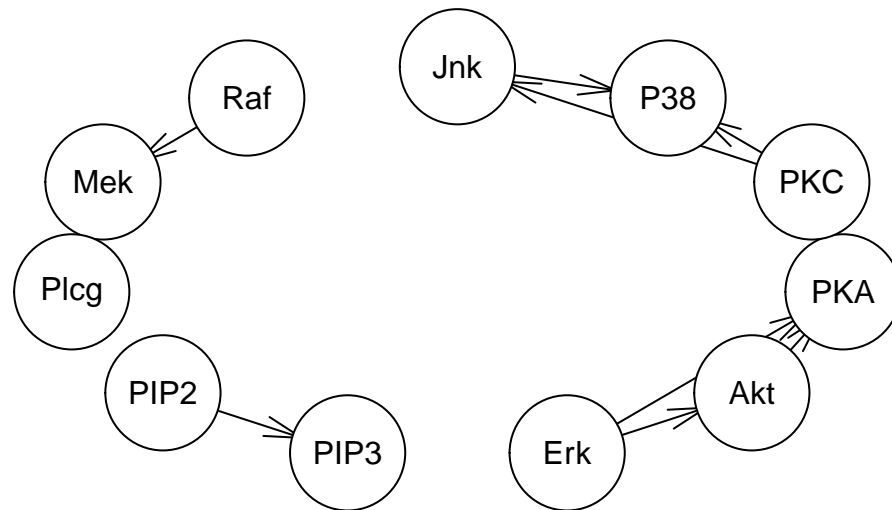
for (x in c("Plcg", "PIP3")){
  cat(hc.fit[[x]]$node, "hill-climbing", "\n")
  print(hc.fit[[x]]$coefficients)
  cat("\n")
  cat(tabu.fit[[x]]$node, "tabu-search", "\n")
  print(tabu.fit[[x]]$coefficients)
  cat("\n")
}
```

```
## Plcg hill-climbing
## (Intercept)      PIP3
## 17.73549627  0.05776208
##
## Plcg tabu-search
## (Intercept)      PIP3
## 17.73549627  0.05776208
##
## PIP3 hill-climbing
## (Intercept)      PIP2
## 22.2957897   0.1000896
##
## PIP3 tabu-search
## (Intercept)
## 30.45849
```

We can see that the densities between of the node “Plcg” is the same for both networks but there is a difference between the densities for node “PIP3”. This can be explained by the different structures of the networks. While “Plcg” has the same Markov Blanket in both networks, “PIP3” has one parent and one child in the hill-climbing graph and no parents and 2 children in the tabu search graph.

```
arcs = boot.strength(sachs.data, R = 400, algorithm = "hc")
bn.av = averaged.network(arcs, threshold = 0.75)
plot(bn.av, main = "Bootstrap network with Hill-climb")
```

Bootstrap network with Hill-climb



```
av.fit = bn.fit(bn.av, sach.s.data)

for (x in c("Plcg", "PIP3")){
  cat(av.fit[[x]]$node, "\n")
  print(av.fit[[x]]$coefficients)
  cat("\n")
}
```

```
## Plcg
## (Intercept)
## 19.49484
##
## PIP3
## (Intercept)      PIP2
## 22.2957897      0.1000896
```