Control Bayesian network arcs

Aaron Smith

6/23/2020

Let’s build some Bayesian networks where we decide to include or exclude arcs before learning the model.

# Load the data

We will use the Groceries data set from the arules package. We will only use the ten most frequently purchased items.

options(digits = 1)  
data(Groceries,package = "arules")  
Groceries

## Loading required package: arules

## Loading required package: Matrix

##   
## Attaching package: 'arules'

## The following objects are masked from 'package:base':  
##   
## abbreviate, write

## transactions in sparse format with  
## 9835 transactions (rows) and  
## 169 items (columns)

M\_Groceries <- as.data.frame(as(  
 object = Groceries,  
 Class = "matrix"  
))  
colnames(M\_Groceries) <- make.names(colnames(M\_Groceries))  
M\_Groceries <- M\_Groceries[  
 order(rowSums(M\_Groceries),decreasing = TRUE),  
 order(colSums(M\_Groceries),decreasing = TRUE)  
]  
M\_Groceries <- M\_Groceries[,1:10]  
M\_Groceries <- M\_Groceries[  
 order(rowSums(M\_Groceries),decreasing = TRUE),  
 order(colSums(M\_Groceries),decreasing = TRUE)  
]  
for(j in 1:ncol(M\_Groceries)) M\_Groceries[,j] <- factor(as.numeric(M\_Groceries[,j]))

# Learn the models

## A model with no predetermined arcs

require(bnlearn)

## Loading required package: bnlearn

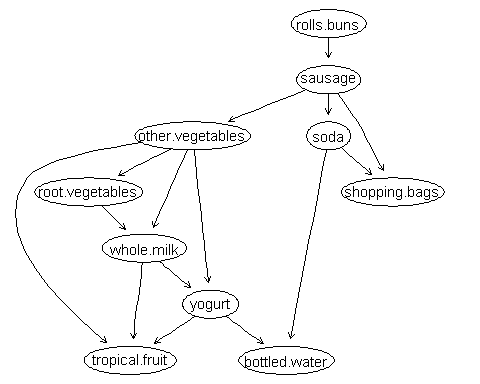
##   
## Attaching package: 'bnlearn'

## The following object is masked from 'package:arules':  
##   
## discretize

## The following object is masked from 'package:stats':  
##   
## sigma

hc\_Groceries\_1 <- hc(  
 x = M\_Groceries  
)  
graphviz.plot(  
 x = hc\_Groceries\_1,  
 shape = "ellipse"  
)

## Loading required namespace: Rgraphviz

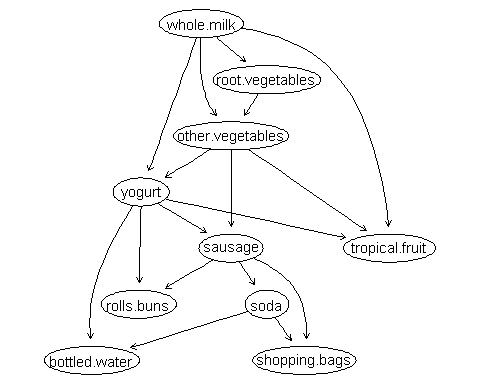


## A model with predetermined arcs to be included

We use a white list data frame to tell R which arcs to include in the model. The white list arcs are added to the nodes, then the remaining possible arcs are evaluated for being worthy of inclusion.

The white list data frame has to fields, from and to.

whitelist <- data.frame(  
 from = c("sausage","root.vegetables"),  
 to = c("rolls.buns","other.vegetables")  
)  
hc\_Groceries\_2 <- hc(  
 x = M\_Groceries,  
 whitelist = whitelist  
)  
graphviz.plot(  
 x = hc\_Groceries\_2,  
 shape = "ellipse"  
)

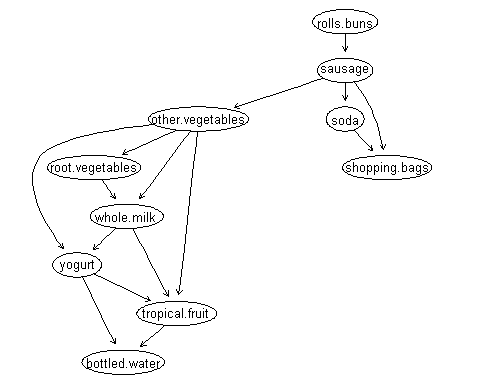


## A model with predetermined arcs to be excluded

We use a black list data frame to exclude arcs from our model. These combinations of variables are skipped over for evaluation of being worth of inclusion in the model.

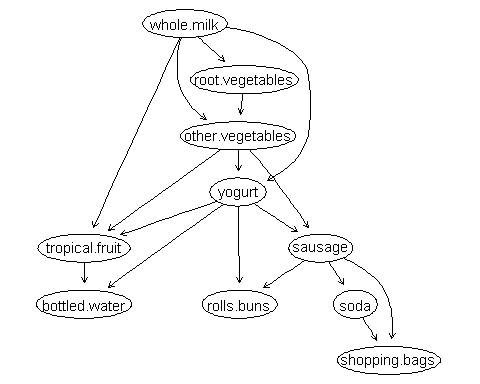
The black list data frame has to fields, from and to.

blacklist <- data.frame(  
 from = c("bottled.water","soda"),  
 to = c("soda","bottled.water")  
)  
hc\_Groceries\_3 <- hc(  
 x = M\_Groceries,  
 blacklist = blacklist  
)  
graphviz.plot(  
 x = hc\_Groceries\_3,  
 shape = "ellipse"  
)



## A model with predetermined arcs to be included and arcs to be excluded

hc\_Groceries\_4 <- hc(  
 x = M\_Groceries,  
 whitelist = whitelist,  
 blacklist = blacklist  
)  
graphviz.plot(  
 x = hc\_Groceries\_4,  
 shape = "ellipse"  
)



# Adjusting arcs after the model was built

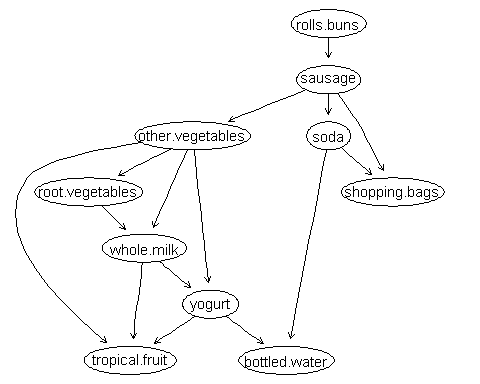
## Letting data pick the direction

The choose.direction() function enables us to add an arc when we know there should be one, but we do not know the direction, or double check if an arc is pointed in the correct manner. Note: the function will not switch the direction if it would result in a cycle in the graph.

arcs(hc\_Groceries\_1)

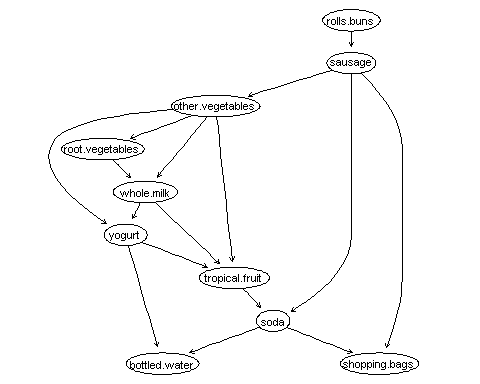
## from to   
## [1,] "other.vegetables" "root.vegetables"   
## [2,] "root.vegetables" "whole.milk"   
## [3,] "whole.milk" "yogurt"   
## [4,] "yogurt" "tropical.fruit"   
## [5,] "other.vegetables" "whole.milk"   
## [6,] "rolls.buns" "sausage"   
## [7,] "other.vegetables" "tropical.fruit"   
## [8,] "other.vegetables" "yogurt"   
## [9,] "soda" "bottled.water"   
## [10,] "sausage" "other.vegetables"  
## [11,] "sausage" "shopping.bags"   
## [12,] "sausage" "soda"   
## [13,] "whole.milk" "tropical.fruit"   
## [14,] "yogurt" "bottled.water"   
## [15,] "soda" "shopping.bags"

hc\_Groceries\_1 <- choose.direction(  
 x = hc\_Groceries\_1,  
 arc = c("soda","shopping.bags"),  
 data = M\_Groceries  
)  
graphviz.plot(  
 x = hc\_Groceries\_1,  
 shape = "ellipse"  
)



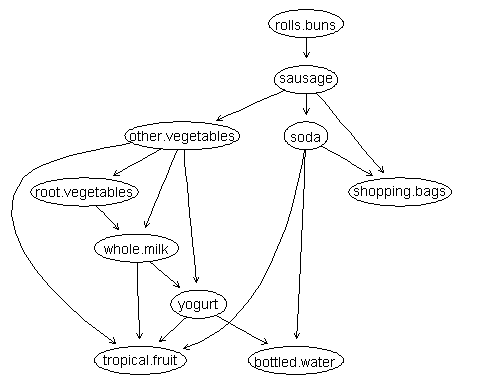
## Adding an arc after the model was built

hc\_Groceries\_1 <- set.arc(  
 x = hc\_Groceries\_1,  
 from = "tropical.fruit",  
 to = "soda",  
 check.cycles = FALSE  
)  
graphviz.plot(  
 x = hc\_Groceries\_1,  
 shape = "ellipse"  
)



## Reversing an arc’s direction

hc\_Groceries\_1 <- reverse.arc(  
 x = hc\_Groceries\_1,  
 from = "tropical.fruit",  
 to = "soda",  
 check.cycles = FALSE  
)  
graphviz.plot(  
 x = hc\_Groceries\_1,  
 shape = "ellipse"  
)



## Dropping an arc after the model was built

hc\_Groceries\_1 <- drop.arc(  
 x = hc\_Groceries\_1,  
 from = "tropical.fruit",  
 to = "soda"  
)  
graphviz.plot(  
 x = hc\_Groceries\_1,  
 shape = "ellipse"  
)

