

# DAG structure learning

G. Marchetti

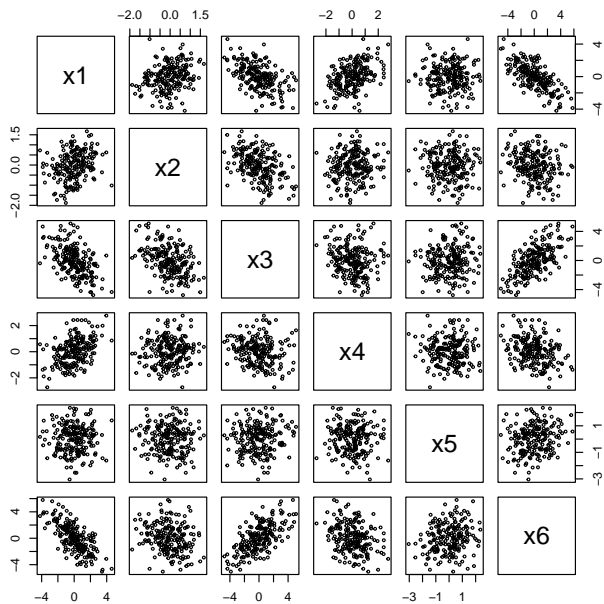
2024-10-21

## Data on six simulated variables

The true generating process is unknown

### The data set

```
source("simdat.R")  
  
pairs(simdat, pch = 1, cex = .5)
```



```
su <- apply(simdat, 2, function(x) rbind(mean(x), sd(x)))
rownames(su) <- c("mean", "sd")
print(round(su, 3))
```

```
      x1      x2      x3      x4      x5      x6
mean -0.006 -0.036 -0.049 0.041 0.029 0.030
sd    1.576  0.699  2.010 1.004 1.046 2.137
```

```
round(cor(simdat), 3)
```

```
      x1      x2      x3      x4      x5      x6
x1  1.000  0.265 -0.514  0.440  0.002 -0.717
x2  0.265  1.000 -0.423  0.075 -0.032 -0.226
x3 -0.514 -0.423  1.000 -0.158  0.082  0.655
x4  0.440  0.075 -0.158  1.000 -0.045 -0.301
x5  0.002 -0.032  0.082 -0.045  1.000  0.187
x6 -0.717 -0.226  0.655 -0.301  0.187  1.000
```

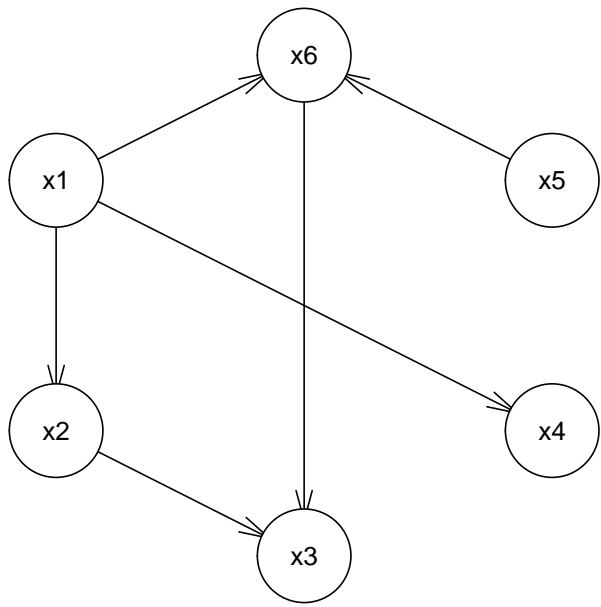
## Structure learning

Hill climbing

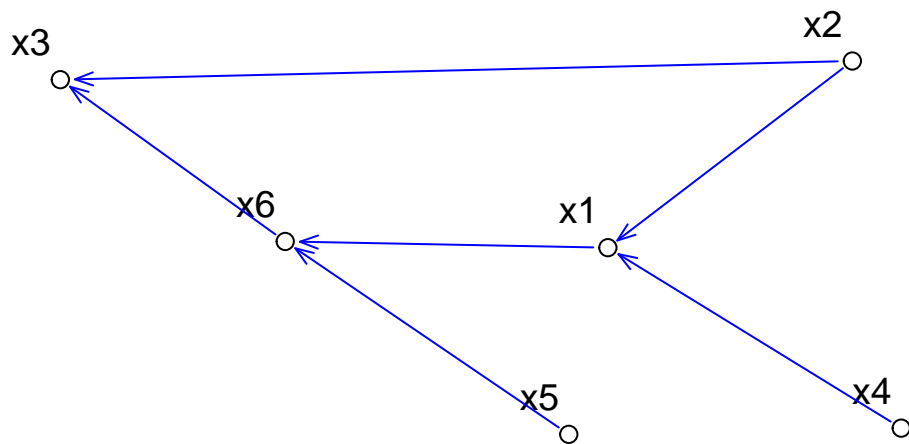
```
out <- hc(X, score = "bic-g", debug = FALSE)
f <- modelstring(out)
f
```

```
[1] "[x1] [x5] [x2|x1] [x4|x1] [x6|x1:x5] [x3|x2:x6]"
```

```
plot(model2network(f))
```



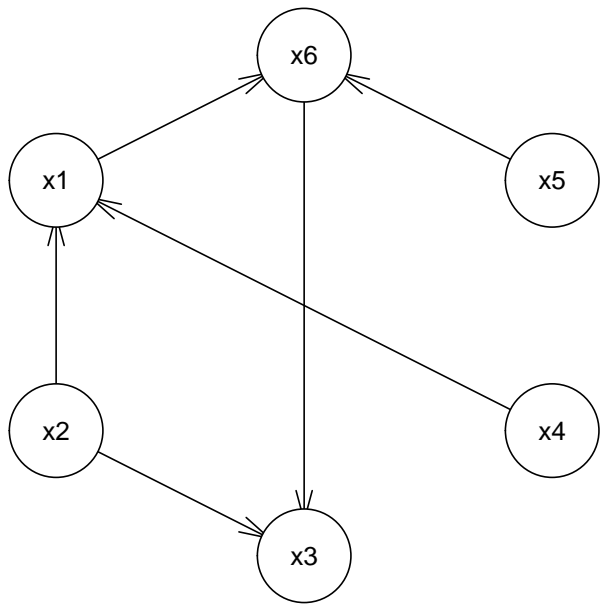
**True DAG**



```

pc <- pc.stable(X, test = "cor", undirected = FALSE)
plot(pc)

```



**Find just an undirected graph**

```
round(correlations(cov(X)), 3)
```

	x1	x2	x3	x4	x5	x6
x1	1.000	0.190	-0.034	0.261	0.232	-0.593
x2	0.290	1.000	-0.321	-0.080	0.007	0.139
x3	-0.509	-0.386	1.000	0.016	0.013	0.461
x4	0.363	0.025	-0.166	1.000	-0.059	-0.021
x5	0.003	0.006	0.130	-0.061	1.000	0.271
x6	-0.713	-0.211	0.642	-0.282	0.218	1.000