Answers Session 1: Introduction

Exercise 2.1

a) Load packages and data:

```
library(lavaan)
```

```
## This is lavaan 0.6-1
```

```
## lavaan is BETA software! Please report any bugs.
```

```
library(semPlot)
Math.data <- read.table("MathHmwk.txt", header = TRUE)
```

b) Fit a linear regression model:

```
Math.lm <- lm(MathAchievement ~ MathHomework, data = Math.data)
Math.lm
```

```
##
## Call:
## lm(formula = MathAchievement ~ MathHomework, data = Math.data)
##
## Coefficients:
##  (Intercept)  MathHomework
##        47.03          1.99
```

c) To standardize, remember that in bivariate regression: $r = b\frac{s_x}{sy}$:

```
coef(Math.lm)[2] * sd(Math.data$MathHomework) / sd(Math.data$MathAchievement)
```

```
## MathHomework
##    0.3199936
```

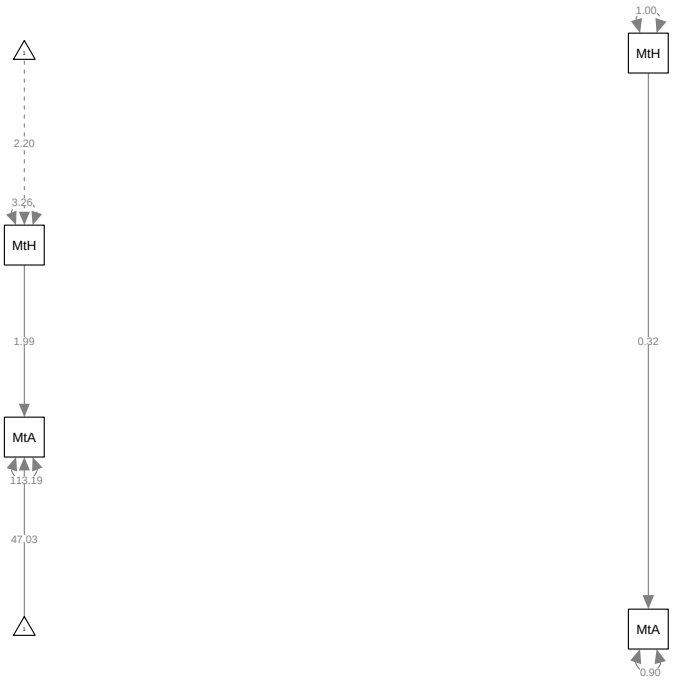
d) Fit the SEM model and let the computer make the path models:

```
Math.model <- '
  MathAchievement ~ MathHomework
'
Math.fit <- sem(Math.model, data = Math.data, meanstructure = TRUE)
summary(Math.fit, standardized = TRUE)
```

```
## lavaan (0.6-1) converged normally after  17 iterations
##
##   Number of observations                           100
##
##   Estimator                                         ML
##   Model Fit Test Statistic                       0.000
##   Degrees of freedom                                 0
##
## Parameter Estimates:
##
##   Information                                 Expected
##   Information saturated (h1) model          Structured
##   Standard Errors                             Standard
##
## Regressions:
##                    Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
```

```
##   MathAchievement ~
##     MathHomework       1.990    0.589    3.378    0.001    1.990    0.320
##
## Intercepts:
##                    Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
##    .MathAchievemnt   47.032    1.677   28.045    0.000   47.032    4.188
##
## Variances:
##                    Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
##    .MathAchievemnt  113.190   16.007    7.071    0.000  113.190    0.898
```

```
par(mfrow = c(1,2))
semPaths(Math.fit, whatLabels = "par")
semPaths(Math.fit, whatLabels = "std", intercepts = FALSE)
```



e) The solutions are the same.

Exercise 2.2

a) Input the correlation matrix:

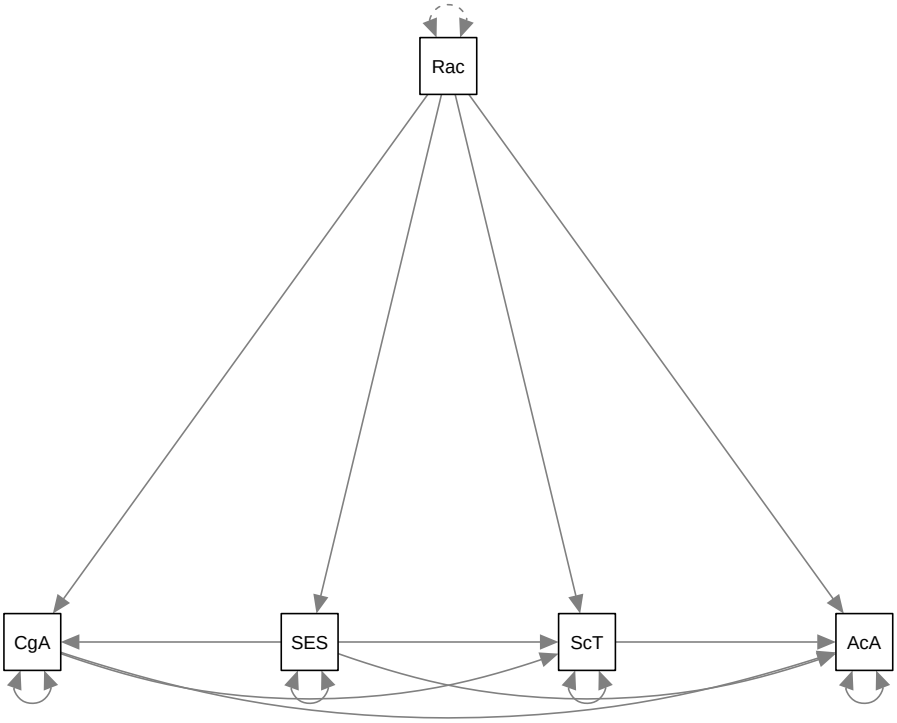
```
HighSchool.cor <- lav_matrix_lower2full(c(
  1.000,
  0.178, 1.000,
  0.230, 0.327, 1.000,
  0.106, 0.245, 0.183, 1.000,
  0.195, 0.356, 0.721, 0.178, 1.000
))
rownames(HighSchool.cor) <- colnames(HighSchool.cor) <-
  c("Race", "SES", "CognAb", "SchoolTy", "AcAch")
```

b) Create the path model in lavaan:

```
HighSchool.mod <- '
  CognAb ~ Race + SES
  SES ~ Race
  SchoolTy ~ SES + Race + CognAb
  AcAch ~ Race + SES + SchoolTy + CognAb
'
HighSchool.fit <- sem(HighSchool.mod, sample.cov = HighSchool.cor,
                      sample.nobs = 18058)
summary(HighSchool.fit, standardized = TRUE)
```

```
## lavaan (0.6-1) converged normally after  16 iterations
##
##   Number of observations                         18058
##
##   Estimator                                         ML
##   Model Fit Test Statistic                       0.000
##   Degrees of freedom                                 0
##
## Parameter Estimates:
##
##   Information                                 Expected
##   Information saturated (h1) model          Structured
##   Standard Errors                             Standard
##
## Regressions:
##                    Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
##   CognAb ~
##     Race              0.177    0.007   25.260    0.000    0.177    0.177
##     SES               0.295    0.007   42.061    0.000    0.295    0.295
##   SES ~
##     Race              0.178    0.007   24.308    0.000    0.178    0.178
##   SchoolTy ~
##     SES               0.202    0.008   26.499    0.000    0.202    0.202
##     Race              0.046    0.007    6.151    0.000    0.046    0.046
##     CognAb            0.106    0.008   13.804    0.000    0.106    0.106
##   AcAch ~
##     Race              0.015    0.005    2.935    0.003    0.015    0.015
##     SES               0.128    0.005   23.336    0.000    0.128    0.128
##     SchoolTy          0.022    0.005    4.191    0.000    0.022    0.022
##     CognAb            0.671    0.005  122.482    0.000    0.671    0.671
##
## Variances:
##                    Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
##    .CognAb            0.863    0.009   95.021    0.000    0.863    0.863
##    .SES               0.968    0.010   95.021    0.000    0.968    0.968
##    .SchoolTy          0.926    0.010   95.021    0.000    0.926    0.926
##    .AcAch             0.463    0.005   95.021    0.000    0.463    0.463
```

```
semPaths(HighSchool.fit)
```




```
HighSchool.cor
```

```
##            Race   SES CognAb SchoolTy AcAch
## Race      1.000 0.178  0.230    0.106 0.195
## SES       0.178 1.000  0.327    0.245 0.356
## CognAb    0.230 0.327  1.000    0.183 0.721
## SchoolTy 0.106 0.245  0.183    1.000 0.178
## AcAch     0.195 0.356  0.721    0.178 1.000
```

d) The first-order, bivariate correlation between school type and academic achievement is .178. When controlling for the linear effects of Race, SES, and Cognitive Abilities, the correlation is much lower: .022 (but with $> 18,000$ observations, still significant).

Exercise 2.3

a) Input the covariances:

```
MacKinnon.cov <- lav_matrix_lower2full(c(
  84.85,
  71.28, 140.34,
  18.83,  -6.25, 72.92,
  60.05,  84.54, 37.18, 139.48
))
rownames(MacKinnon.cov) <- colnames(MacKinnon.cov) <-
  c("TeachExp", "SocCli", "MatCov", "StudAch")
```

b) Write the syntax:

```
MacKinnon.mod <- '
  SocCli ~ a1*TeachExp
  MatCov ~ a2*TeachExp
  StudAch ~ TeachExp + b1*SocCli + b2*MatCov
```

```
  a1b1 := a1*b1
  a2b2 := a2*b2
'
MacKinnon.fit <- sem(MacKinnon.mod, sample.cov = MacKinnon.cov, sample.nobs = 40)
summary(MacKinnon.fit, standardized = TRUE)
```

```
## lavaan (0.6-1) converged normally after  17 iterations
##
##   Number of observations                            40
##
##   Estimator                                         ML
##   Model Fit Test Statistic                       3.687
##   Degrees of freedom                                 1
##   P-value (Chi-square)                           0.055
##
## Parameter Estimates:
##
##   Information                                 Expected
##   Information saturated (h1) model          Structured
##   Standard Errors                             Standard
##
## Regressions:
##                    Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
##   SocCli ~
##     TeachExp  (a1)    0.840    0.154    5.456    0.000    0.840    0.653
##   MatCov ~
##     TeachExp  (a2)    0.222    0.142    1.559    0.119    0.222    0.239
##   StudAch ~
##     TeachExp          0.112    0.186    0.603    0.546    0.112    0.084
##     SocCli    (b1)    0.569    0.142    4.006    0.000    0.569    0.545
##     MatCov    (b2)    0.530    0.154    3.446    0.001    0.530    0.366
##
## Variances:
##                    Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
##    .SocCli           78.448   17.542    4.472    0.000   78.448    0.573
##    .MatCov           67.023   14.987    4.472    0.000   67.023    0.943
##    .StudAch          63.323   14.159    4.472    0.000   63.323    0.425
##
## Defined Parameters:
##                    Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
##     a1b1              0.478    0.148    3.229    0.001    0.478    0.356
##     a2b2              0.118    0.083    1.421    0.155    0.118    0.088
```

c) The indirect effects are the effect of Teacher Expectancies on Student Achievement, both via Social Climate and via Material Covered. Only the indirect effect via Social Climate is significant, and explains over 10 percent of variance in Student Achievement (standardized estimate a1b1 = .356 and a2b2 = .088; if we square those values, we get the proportion of variance explained).