BIOS6643. L8 Covariance Structures R_i

1. Dental study

Dental study

- The orthodontic study data of Potthoff and Roy (1964).

- World famous data set that is used to introduce features of longitudinal data modeling and analysis

- A study was conducted involving 27 children, 16 boys and 11 girls

- For each child, the distance (mm) from the center of the pituitary to the pterygomaxillary fissure was measured at ages 8, 10, 12, and 14 years of age

- The pterygomaxillary fissure is a vertical opening in the human skull.

- **Objectives** of the study included:

    - Determine if distances over time are larger on average for boys than for girls
    - Determine if the rate of change of distance over time is different for boys and girls.

```
# Read in the data
dat.den <- read.table("/Users/juarezce/Documents/OneDrive - The University of Colorado Denver/BIOS6643/
dat.den <- dat.den[,2:5]      # remove the first column
colnames(dat.den) <- c("id","age","distance","gender")

# Total number of individuals
m <- max(dat.den$id)

head(dat.den, 2)

##   id age distance gender
## 1  1   8       21      0
## 2  1  10       20      0

table(dat.den$gender)

##
##  0  1
## 44 64
```
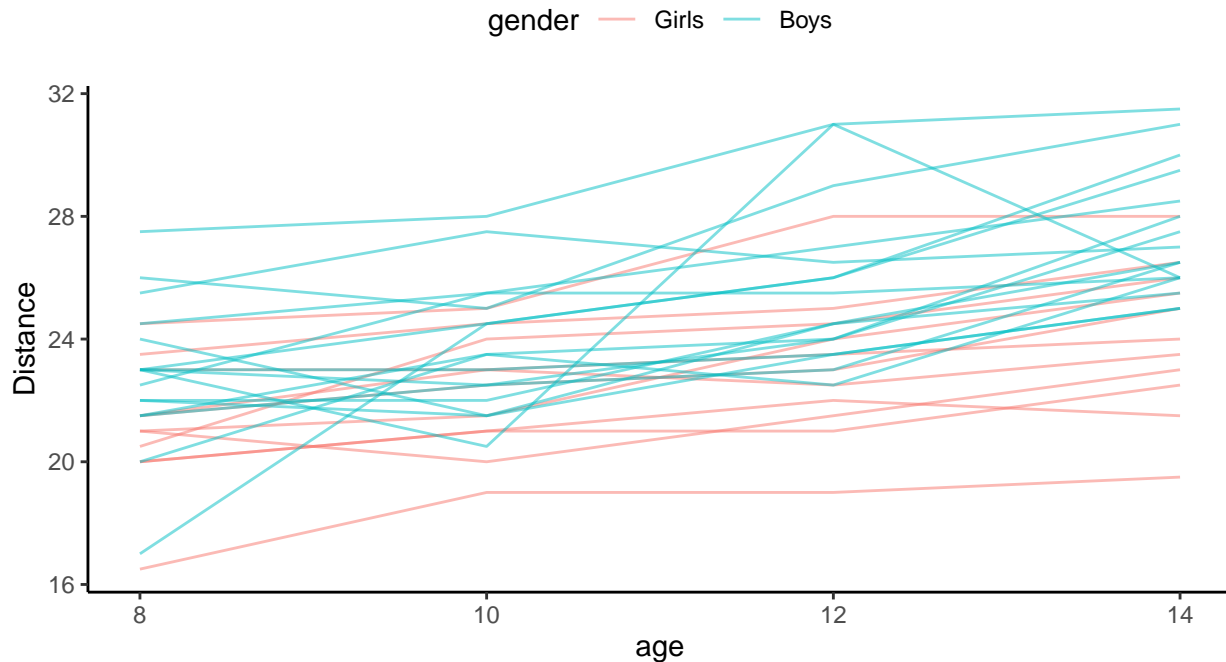
Some descriptives statistics:


```
##   id gender distance.8 distance.10 distance.12 distance.14
## 1  1  Girls       21.0        20.0        21.5        23.0
## 5  2  Girls       21.0        21.5        24.0        25.5
## 9  3  Girls       20.5        24.0        24.5        26.0

##  distance.8 distance.10 distance.12 distance.14
##    21.18182    22.22727    23.09091    24.09091

##  distance.8 distance.10 distance.12 distance.14
##    22.87500    23.81250    25.71875    27.46875

##             distance.8 distance.10 distance.12 distance.14
## distance.8    4.513636    3.354545    4.331818    4.356818
## distance.10   3.354545    3.618182    4.027273    4.077273
## distance.12   4.331818    4.027273    5.590909    5.465909
## distance.14   4.356818    4.077273    5.465909    5.940909

##             distance.8 distance.10 distance.12 distance.14
## distance.8    6.016667    2.291667    3.629167    1.612500
## distance.10   2.291667    4.562500    2.193750    2.810417
## distance.12   3.629167    2.193750    7.032292    3.240625
## distance.14   1.612500    2.810417    3.240625    4.348958

##             distance.8 distance.10 distance.12 distance.14
## distance.8   1.0000000   0.8300900   0.8623146   0.8413558
## distance.10  0.8300900   1.0000000   0.8954156   0.8794236
## distance.12  0.8623146   0.8954156   1.0000000   0.9484070
## distance.14  0.8413558   0.8794236   0.9484070   1.0000000

##             distance.8 distance.10 distance.12 distance.14
## distance.8   1.0000000   0.4373932   0.5579310   0.3152311
## distance.10  0.4373932   1.0000000   0.3872909   0.6309234
## distance.12  0.5579310   0.3872909   1.0000000   0.5859866
```
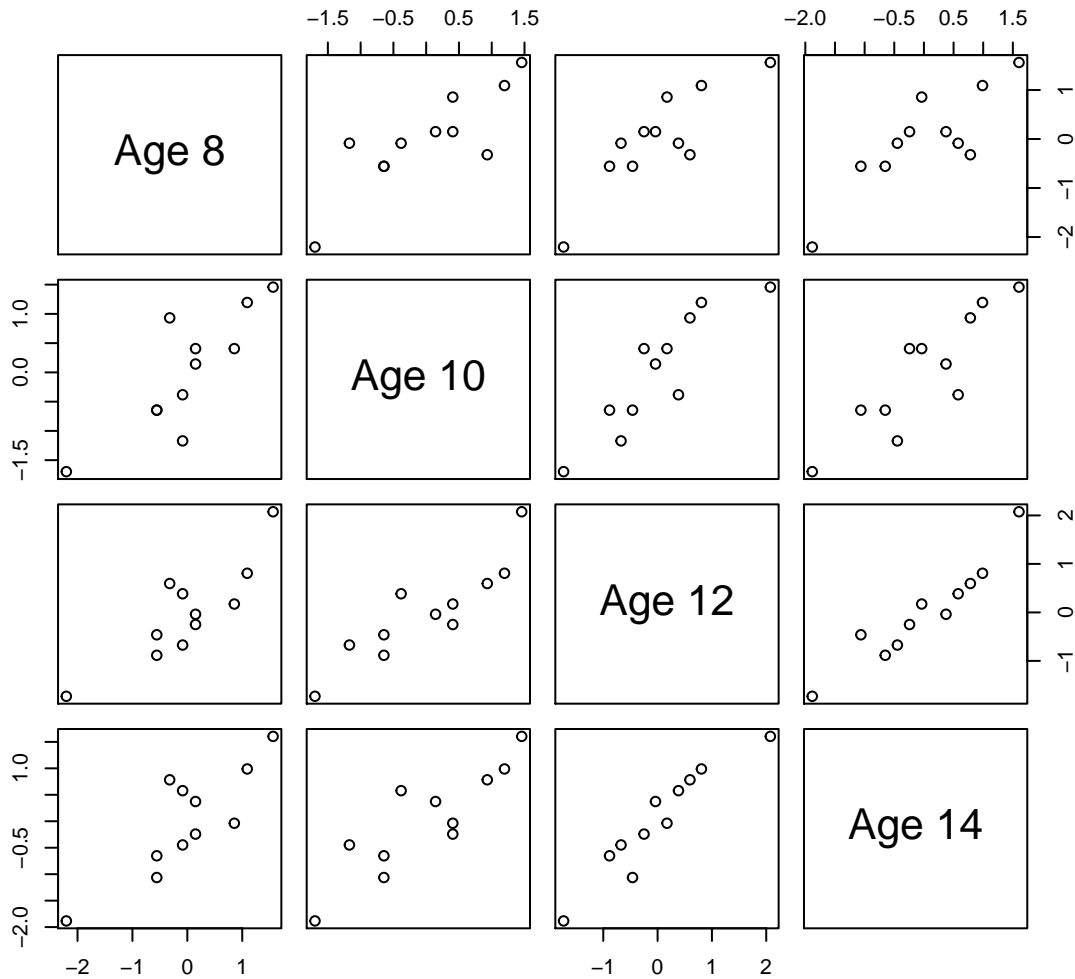
```
## distance.14   0.3152311    0.6309234    0.5859866    1.0000000
```

Calculating autocorrelation:

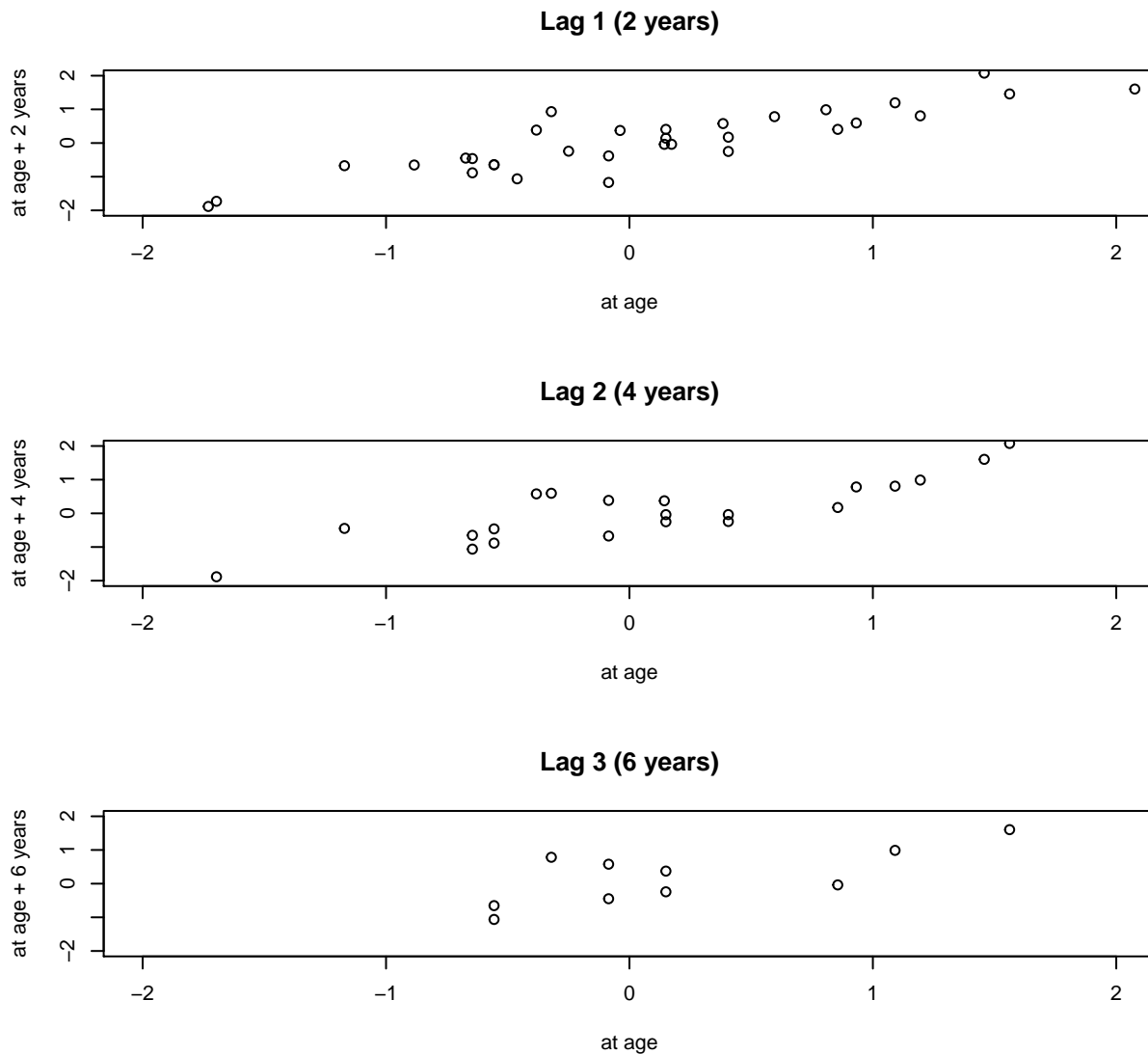**Girls**

```
##              distance.8 distance.10 distance.12 distance.14
## distance.8     5.415455    2.716818    3.910227    2.710227
## distance.10    2.716818    4.184773    2.927159    3.317159
## distance.12    3.910227    2.927159    6.455739    4.130739
## distance.14    2.710227    3.317159    4.130739    4.985739

##              distance.8 distance.10 distance.12 distance.14
## distance.8    1.0000000   0.5706991   0.6613199   0.5215831
## distance.10   0.5706991   1.0000000   0.5631673   0.7262157
## distance.12   0.6613199   0.5631673   1.0000000   0.7280983
## distance.14   0.5215831   0.7262157   0.7280983   1.0000000

## [1] 0.891 0.871 0.841
```

Lag 1 (2 years)

Lag 2 (4 years)

Lag 3 (6 years)


```
## [1] 0.470 0.594 0.315
```

1. Unstructured correlation

```r
# Common unstructured correlation with variances changing over time
# for both genders -- we extract components of the fit.

head(dat.den, 2)
```

```
##   id age distance gender
## 1  1   8       21  Girls
## 2  1  10       20  Girls
```

```r
dental.un <- gls(distance ~ -1 + gender + age:gender, data=dat.den,
                 correlation=corSymm(form =  ~ 1 | id),
                 weights = varIdent(form = ~ 1 | age), method="ML")
beta.un <- coef(dental.un)
sebeta.un <- summary(dental.un)$tTable[,"Std.Error"]
```

```
V.un <- getVarCov(dental.un)  ## V_i matrix
Gamma.un <- cov2cor(V.un) ## Gamma_i=correlation matrix
vars.un <- diag(V.un) ## Lambda_i= diagonal matrix of variances
V.un
```

```
## Marginal variance covariance matrix
##         [,1]   [,2]   [,3]   [,4]
## [1,] 5.1192 2.4409 3.6105 2.5223
## [2,] 2.4409 3.9279 2.7175 3.0624
## [3,] 3.6105 2.7175 5.9798 3.8235
## [4,] 2.5223 3.0624 3.8235 4.6180
##    Standard Deviations: 2.2626 1.9819 2.4454 2.149
```

```
Gamma.un
```

```
## Marginal variance covariance matrix
##          [,1]    [,2]    [,3]    [,4]
## [1,] 1.00000 0.54433 0.65256 0.51876
## [2,] 0.54433 1.00000 0.56072 0.71903
## [3,] 0.65256 0.56072 1.00000 0.72759
## [4,] 0.51876 0.71903 0.72759 1.00000
##    Standard Deviations: 1 1 1 1
```

```
vars.un
```

```
## [1] 5.119154 3.927933 5.979800 4.618016
```

2. Common Compound symmetry

```
dental.cs <- gls(distance ~ -1 + gender + age:gender, data=dat.den,
                 correlation=corCompSymm(form = ~ 1 | id),method="ML")
summary(dental.cs)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: distance ~ -1 + gender + age:gender
##   Data: dat.den
##        AIC      BIC    logLik
##   440.6391 456.7318 -214.3195
##
## Correlation Structure: Compound symmetry
##  Formula: ~1 | id
##  Parameter estimate(s):
##       Rho
## 0.6178309
##
## Coefficients:
##                   Value Std.Error   t-value p-value
## genderGirls    17.372727 1.1836504 14.677246       0
## genderBoys     16.340625 0.9814310 16.649794       0
## genderGirls:age  0.479545 0.0940671  5.097906       0
## genderBoys:age   0.784375 0.0779963 10.056560       0
##
##   Correlation:
##               gndrGr gndrBy gndrG:
## genderBoys       0.000
```

```
## genderGirls:age -0.874  0.000
## genderBoys:age   0.000 -0.874  0.000
##
## Standardized residuals:
##        Min          Q1         Med          Q3         Max
## -2.53554804 -0.59684854 -0.07593689  0.60047221  2.36905129
##
## Residual standard error: 2.214758
## Degrees of freedom: 108 total; 104 residual
```

```
beta.cs <- coef(dental.cs)
sebeta.cs <- summary(dental.cs)$tTable[,"Std.Error"]
V.cs <- getVarCov(dental.cs)
Gamma.cs <- cov2cor(V.cs)
vars.cs <- diag(V.cs)
V.cs
```

```
## Marginal variance covariance matrix
##        [,1]   [,2]   [,3]   [,4]
## [1,] 4.9052 3.0306 3.0306 3.0306
## [2,] 3.0306 4.9052 3.0306 3.0306
## [3,] 3.0306 3.0306 4.9052 3.0306
## [4,] 3.0306 3.0306 3.0306 4.9052
##   Standard Deviations: 2.2148 2.2148 2.2148 2.2148
```

```
Gamma.cs
```

```
## Marginal variance covariance matrix
##         [,1]    [,2]    [,3]    [,4]
## [1,] 1.00000 0.61783 0.61783 0.61783
## [2,] 0.61783 1.00000 0.61783 0.61783
## [3,] 0.61783 0.61783 1.00000 0.61783
## [4,] 0.61783 0.61783 0.61783 1.00000
##   Standard Deviations: 1 1 1 1
```

```
vars.cs
```

```
## [1] 4.905152 4.905152 4.905152 4.905152
```

3. Compound symmetry by gender

```
dental.cs2 <- gls(distance ~ -1 + gender + gender:age, data=dat.den,
                  correlation=corCompSymm(form = ~ 1 | id),
                  weights = varIdent(form = ~ 1 | gender), method="ML")
beta.cs2 <- coef(dental.cs2)
sebeta.cs2 <- summary(dental.cs2)$tTable[,"Std.Error"]
V.cs2.girl <-  getVarCov(dental.cs2, individual=1)
V.cs2.boy <-  getVarCov(dental.cs2, individual=12)
Gamma.cs2.girl <- cov2cor(V.cs2.girl)
Gamma.cs2.boy <- cov2cor(V.cs2.boy)
V.cs2.girl
```

```
## Marginal variance covariance matrix
##        [,1]   [,2]   [,3]   [,4]
## [1,] 2.8677 2.0728 2.0728 2.0728
## [2,] 2.0728 2.8677 2.0728 2.0728
```

```
## [3,] 2.0728 2.0728 2.8677 2.0728
## [4,] 2.0728 2.0728 2.0728 2.8677
##    Standard Deviations: 1.6934 1.6934 1.6934 1.6934
```

`Gamma.cs2.girl`

```
## Marginal variance covariance matrix
##         [,1]    [,2]    [,3]    [,4]
## [1,] 1.00000 0.72281 0.72281 0.72281
## [2,] 0.72281 1.00000 0.72281 0.72281
## [3,] 0.72281 0.72281 1.00000 0.72281
## [4,] 0.72281 0.72281 0.72281 1.00000
##    Standard Deviations: 1 1 1 1
```

`V.cs2.boy`

```
## Marginal variance covariance matrix
##        [,1]   [,2]   [,3]   [,4]
## [1,] 8.4514 6.1088 6.1088 6.1088
## [2,] 6.1088 8.4514 6.1088 6.1088
## [3,] 6.1088 6.1088 8.4514 6.1088
## [4,] 6.1088 6.1088 6.1088 8.4514
##    Standard Deviations: 2.9071 2.9071 2.9071 2.9071
```

`Gamma.cs2.boy`

```
## Marginal variance covariance matrix
##         [,1]    [,2]    [,3]    [,4]
## [1,] 1.00000 0.72281 0.72281 0.72281
## [2,] 0.72281 1.00000 0.72281 0.72281
## [3,] 0.72281 0.72281 1.00000 0.72281
## [4,] 0.72281 0.72281 0.72281 1.00000
##    Standard Deviations: 1 1 1 1
```

4. AR(1)

```
dental.ar1 <- gls(distance ~ -1 + gender + age:gender, data=dat.den,
                correlation=corAR1(form = ~ 1 | id), method="ML")
beta.ar1 <- coef(dental.ar1)
sebeta.ar1 <- summary(dental.ar1)$tTable[,"Std.Error"]
V.ar1 <- getVarCov(dental.ar1)  # or corMatrix(dental.un$modelStruct$corStruct)[[1]]
Gamma.ar1 <- cov2cor(V.ar1)
vars.ar1 <- diag(V.ar1)
V.ar1
```

```
## Marginal variance covariance matrix
##        [,1]   [,2]   [,3]   [,4]
## [1,] 4.8908 2.9693 1.8027 1.0944
## [2,] 2.9693 4.8908 2.9693 1.8027
## [3,] 1.8027 2.9693 4.8908 2.9693
## [4,] 1.0944 1.8027 2.9693 4.8908
##    Standard Deviations: 2.2115 2.2115 2.2115 2.2115
```

`Gamma.ar1`

```
## Marginal variance covariance matrix
```

```
##         [,1]    [,2]    [,3]    [,4]
## [1,] 1.00000 0.60712 0.36859 0.22378
## [2,] 0.60712 1.00000 0.60712 0.36859
## [3,] 0.36859 0.60712 1.00000 0.60712
## [4,] 0.22378 0.36859 0.60712 1.00000
##   Standard Deviations: 1 1 1 1
```

```
vars.ar1
```

```
## [1] 4.890787 4.890787 4.890787 4.890787
```

5. One-dependent (banded)

```
## autoregressive moving average process, with arbitrary orders for the autoregressive and moving avera

dental.1d <- gls(distance ~ -1 + gender +  age:gender, data=dat.den,
               correlation=corARMA(form = ~ 1 | id, q=1),method="ML")
beta.1d <- coef(dental.1d)
sebeta.1d <- summary(dental.1d)$tTable[,"Std.Error"]
V.1d <- getVarCov(dental.1d)  #  or corMatrix(dental.un$modelStruct$corStruct)[[1]]
Gamma.1d <- cov2cor(V.1d)
vars.1d <- diag(V.1d)
V.1d
```

```
## Marginal variance covariance matrix
##        [,1]   [,2]   [,3]   [,4]
## [1,] 4.5294 1.6120 0.0000 0.0000
## [2,] 1.6120 4.5294 1.6120 0.0000
## [3,] 0.0000 1.6120 4.5294 1.6120
## [4,] 0.0000 0.0000 1.6120 4.5294
##   Standard Deviations: 2.1282 2.1282 2.1282 2.1282
```

```
Gamma.1d
```

```
## Marginal variance covariance matrix
##        [,1]   [,2]   [,3]   [,4]
## [1,] 1.0000 0.3559 0.0000 0.0000
## [2,] 0.3559 1.0000 0.3559 0.0000
## [3,] 0.0000 0.3559 1.0000 0.3559
## [4,] 0.0000 0.0000 0.3559 1.0000
##   Standard Deviations: 1 1 1 1
```

```
vars.1d
```

```
## [1] 4.52935 4.52935 4.52935 4.52935
```

6. Exponential spatial correlation

```
## exponential spatial correlation
dental.exp <- gls(distance ~ -1 + gender +  age:gender, data=dat.den,
               correlation=corExp(form = ~ 1 | id),method="ML")
V.exp <- getVarCov(dental.exp)  #  or corMatrix(dental.un$modelStruct$corStruct)[[1]]
Gamma.exp <- cov2cor(V.exp)
vars.exp <- diag(V.exp)
V.exp
```

```
## Marginal variance covariance matrix
##         [,1]   [,2]   [,3]   [,4]
## [1,] 4.8908 2.9693 1.8027 1.0944
## [2,] 2.9693 4.8908 2.9693 1.8027
## [3,] 1.8027 2.9693 4.8908 2.9693
## [4,] 1.0944 1.8027 2.9693 4.8908
##   Standard Deviations: 2.2115 2.2115 2.2115 2.2115
```

```
Gamma.exp
```

```
## Marginal variance covariance matrix
##          [,1]    [,2]    [,3]    [,4]
## [1,] 1.00000 0.60712 0.36859 0.22378
## [2,] 0.60712 1.00000 0.60712 0.36859
## [3,] 0.36859 0.60712 1.00000 0.60712
## [4,] 0.22378 0.36859 0.60712 1.00000
##   Standard Deviations: 1 1 1 1
```

```
vars.exp
```

```
## [1] 4.890786 4.890786 4.890786 4.890786
```

Comparing models

```
anova(dental.un,dental.cs,dental.cs2,dental.ar1, dental.1d, dental.exp)
```

```
##            Model df      AIC      BIC    logLik   Test  L.Ratio p-value
## dental.un      1 14 447.4770 485.0269 -209.7385
## dental.cs      2  6 440.6391 456.7318 -214.3195 1 vs 2  9.16201  0.3288
## dental.cs2     3  7 430.6521 449.4270 -208.3261 2 vs 3 11.98695  0.0005
## dental.ar1     4  6 452.6810 468.7738 -220.3405 3 vs 4 24.02890  <.0001
## dental.1d      5  6 469.4166 485.5094 -228.7083
## dental.exp     6  6 452.6810 468.7738 -220.3405
```

```
#  We now fit full and reduced models with this covariance structure
#  and compare via likelihood ratio test
```

```
full <- gls(distance ~ -1 + gender + gender:age, data = dat.den,
          correlation=corCompSymm(form = ~ 1 | id),
              weights = varIdent(form = ~ 1 | gender),method="ML")
reduced <- gls(distance ~ -1 + gender + age, data = dat.den,
            correlation=corCompSymm(form = ~ 1 | id),
            weights = varIdent(form = ~ 1 | gender),method="ML")
```

```
anova(full,reduced)
```

```
##          Model df      AIC      BIC    logLik   Test  L.Ratio p-value
## full         1  7 430.6521 449.4270 -208.3261
## reduced      2  6 436.7324 452.8252 -212.3662 1 vs 2 8.080332  0.0045
```

Using Wald test to compare models

```
##  Note we can test the difference in slopes directly using a
##  Wald-type test
```

```
full.alt <- gls(distance ~ gender + age + gender:age, data=dat.den,
                correlation=corCompSymm(form = ~ 1 | id),
                weights = varIdent(form = ~ 1 | gender),method="ML")

summary(full.alt)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: distance ~ gender + age + gender:age
##   Data: dat.den
##        AIC      BIC    logLik
##   430.6521 449.427 -208.3261
##
## Correlation Structure: Compound symmetry
##  Formula: ~1 | id
##  Parameter estimate(s):
##       Rho
## 0.7228109
## Variance function:
##  Structure: Different standard deviations per stratum
##  Formula: ~1 | gender
##  Parameter estimates:
##    Girls     Boys
## 1.00000 1.71672
##
## Coefficients:
##                    Value Std.Error    t-value p-value
## (Intercept)     17.372727 0.8175862 21.248802  0.0000
## genderBoys      -1.032102 1.4222595 -0.725678  0.4697
## age              0.479545 0.0612544  7.828755  0.0000
## genderBoys:age   0.304830 0.1065571  2.860716  0.0051
##
##  Correlation:
##                (Intr) gndrBy age
## genderBoys     -0.575
## age            -0.824  0.474
## genderBoys:age  0.474 -0.824 -0.575
##
## Standardized residuals:
##         Min          Q1         Med          Q3         Max
## -2.78081373 -0.63039864 -0.08111374  0.52140257  2.87744432
##
## Residual standard error: 1.693422
## Degrees of freedom: 108 total; 104 residual
```

2. Models for the Stepped Care data

Exercise: Consider the models used here for the Stepped-Care trial and select a model that fits the data well.

Objective of the study was to determine if there is an intervention effect.

```
# Read in data
dat.step <- read.csv("/Users/juarezce/Documents/OneDrive - The University of Colorado Denver/BIOS6643/B

head(dat.step, 3)
```

```
##   id time    treat time6 time12 time24      cops
## 1  1    0  control     0      0      0  83.26686
## 2  1    6  control     1      0      0  81.52480
## 3  1   12  control     0      1      0  88.36082
```

```
```
```