Problem 1

$$
\begin{aligned}
Var(Y_{ij}) &= var(\mu + b_i + e_{ij}) \\
&= var(b_i) + Var(e_{ij}) \quad (b_i \perp e_{ij}) \\
&= \sigma_b^2 + \sigma_e^2
\end{aligned}
$$

$$
\begin{aligned}
Cov(Y_{ij}, Y_{ik}) &= E(Y_{ij} - EY_{ij})(Y_{ik} - EY_{ik}) \\
&= E(Y_{ij} - \mu)(Y_{ik} - \mu) \\
&= E(b_i + e_{ij})(b_i + e_{ik}) \\
&= E(b_i^2 + b_i e_{ik} + b_i e_{ij} + e_{ij} e_{ik}) \\
&= E b_i^2 + E b_i E e_{ik} + E b_i E e_{ij} + E e_{ij} E e_{ik} \quad (e_{ij} \perp e_{ik},\ b_i \perp e_{i\cdot}) \\
&= E b_i^2 \\
&= Var(b_i) + (E b_i)^2 \\
&= \sigma_b^2
\end{aligned}
$$

$$
\begin{aligned}
corr(Y_{ij}, Y_{ik}) &= \frac{cov(Y_{ij}, Y_{ik})}{\sqrt{var(Y_{ij}) \cdot var(Y_{ik})}} \\
&= \frac{\sigma_b^2}{\sigma_b^2 + \sigma_e^2}
\end{aligned}
$$

Both the variance and correlation is constant across visits, so this corresponds to compound symmetry pattern.

Problem 2 Question 2

Note: Answers to Questions 1 & 3 of Problem 2 are attached after this.

$$Y_{ij} = \beta_0 + a_i + b_0 * I_{(sex_i=0)} + b_1 * I_{(sex_i=1)} + \beta_1 * age_{ij} + e_{ij}$$

$$E(Y_{ij}) = E(\beta_0 + a_i + b_0 I(sex_i=0) + b_1 I(sex_i=1) + \beta_1 \cdot age_{ij} + e_{ij})$$

$$= \beta_0 + \beta_1 \cdot age_{ij}$$

$$var(Y_{ij}) = var(a_i + b_k + e_{ij}) = var(a_i) + var(b_k) + var(e_{ij}) = \sigma_a^2 + \sigma_b^2 + \sigma_e^2$$

Variance-covariance matrices:

① Within-subject covariance:

$$cov(Y_{ij}, Y_{ik}) = E[(Y_{ij} - EY_{ij})(Y_{ik} - EY_{ik})]$$

$$= E[(a_i + b_k + e_{ij})(a_i + b_k + e_{ik})]$$

$$= E[a_i^2 + b_k^2 + e_{ij}e_{ik}]$$

$$= \sigma_a^2 + \sigma_b^2 \qquad (\text{since } e_{ij} \perp e_{ik})$$

② Between-subject covariance:

(i) if the subjects are of the same gender:

$$cov(Y_{mj}, Y_{nk}) = E[(Y_{mj} - E_{mj})(Y_{nk} - E_{nk})]$$

$$= E[(a_m + b_l + e_{mj})(a_n + b_l + e_{nk})] \quad (l=0,1)$$

$$= E\, b_l^2$$

$$= \sigma_b^2$$

(ii) if the subjects are of different genders:

$$cov(Y_{mj}, Y_{nk}) = E[(Y_{mj} - E_{mj})(Y_{nk} - E_{nk})]$$

$$= E[(a_m + b_1 + e_{mj})(a_n + b_0 + e_{nk})]$$

$$= 0 \qquad (\text{since } b_1 \perp b_0)$$

Based on the abve calculation, We construct matrices A and B:

$$A=\begin{bmatrix} \sigma_a^2+\sigma_b^2+\sigma_e^2 & \sigma_a^2+\sigma_b^2 & \sigma_a^2+\sigma_b^2 & \sigma_a^2+\sigma_b^2 \\ \sigma_a^2+\sigma_b^2 & \sigma_a^2+\sigma_b^2+\sigma_e^2 & \sigma_a^2+\sigma_b^2 & \sigma_a^2+\sigma_b^2 \\ \sigma_a^2+\sigma_b^2 & \sigma_a^2+\sigma_b^2 & \sigma_a^2+\sigma_b^2+\sigma_e^2 & \sigma_a^2+\sigma_b^2 \\ \sigma_a^2+\sigma_b^2 & \sigma_a^2+\sigma_b^2 & \sigma_a^2+\sigma_b^2 & \sigma_a^2+\sigma_b^2+\sigma_e^2 \end{bmatrix}_{4\times4}, \quad B=\begin{bmatrix} \sigma_b^2 & \sigma_b^2 & \sigma_b^2 & \sigma_b^2 \\ \sigma_b^2 & \sigma_b^2 & \sigma_b^2 & \sigma_b^2 \\ \sigma_b^2 & \sigma_b^2 & \sigma_b^2 & \sigma_b^2 \\ \sigma_b^2 & \sigma_b^2 & \sigma_b^2 & \sigma_b^2 \end{bmatrix}_{4\times4},$$

where $A_{4\times4}$ is the covariance within a subject, and B is the covariance between measurements from 2 subjects of the same gender.

Therefore, we get

$$C_{boy}=\begin{bmatrix} A & B & \cdots & B \\ B & A & \cdots & B \\ \vdots & \vdots & \ddots & \vdots \\ B & B & B & A \end{bmatrix}_{16\times16}, \quad D_{girl}=\begin{bmatrix} A & B & \cdots & B \\ B & A & \cdots & B \\ \vdots & \vdots & \ddots & \vdots \\ B & B & B & A \end{bmatrix}_{11\times11},$$

and finally

$$\mathrm{var}(\underset{\sim}{Y})=\begin{bmatrix} C_{boy} & 0 \\ 0 & D_{girl} \end{bmatrix}_{108\times108} \qquad [108=(16+11)\times4]$$

p8131_hw6_xy2395

Jack Yan

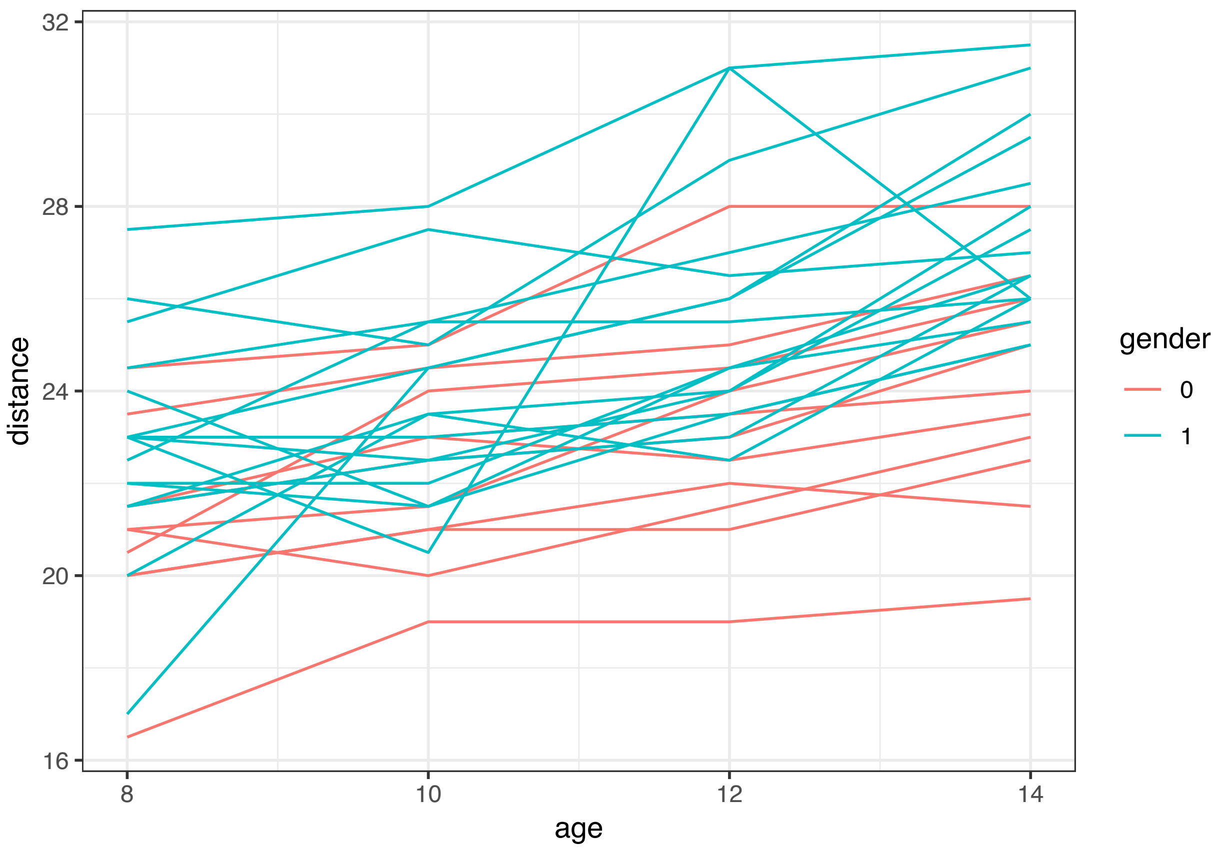
3/31/2019

Problem 2

2.1 Spaghetti Plot

```r
# Import data
dental <-
  read.table('HW6-dental.txt', header = TRUE) %>%
  as.tibble() %>% janitor::clean_names() %>%
  mutate(gender = as.factor(gender))
```

```r
# Spaghetti plot
dental %>%
ggplot(aes(x = age, y = distance, group = child, color = gender)) +
  geom_line() +
  theme_bw()
```



2.2 Marginal Form

Answer to this question is hand-written above.

2.3 Comparing models with different covariance patterns

For the following 3 models, we assume equal variance across measurements at different ages.

```
# Compound Symmetry covariance
compsym = gls(distance ~ gender + age,
              data = dental,
              correlation = corCompSymm(form = ~1 | child),
              method="REML")
# Exponential covariance
expo = gls(distance ~ gender + age,
           data = dental,
           correlation = corExp(form = ~1 | child),
           method = 'REML')
# Autoregressive covariance
auto1 = gls(distance ~ gender + age,
            data = dental,
            correlation = corAR1(form = ~1 | child),
            method = 'REML')
```

```
# Compare coefficient parameter estimates
bind_rows(
  compsym$coefficients,
  expo$coefficients,
  auto1$coefficients,
) %>%
  mutate(CovType = c('CompSym', 'Exp', 'Auto')) %>%
  select(CovType, everything()) %>%
  knitr::kable()
```

CovType	(Intercept)	gender1	age
CompSym	15.38569	2.321023	0.6601852
Exp	15.45999	2.418714	0.6529597
Auto	15.45999	2.418714	0.6529597

The coefficient parameter estimates are similar across the 3 covariance patterns.

```
# Compare covariance estiamtes
# Compound Symmetry
compsym$sigma^2 * corMatrix(compsym$modelStruct$corStruct)[[1]]
```

```
##          [,1]     [,2]     [,3]     [,4]
## [1,] 5.316240 3.266784 3.266784 3.266784
## [2,] 3.266784 5.316240 3.266784 3.266784
## [3,] 3.266784 3.266784 5.316240 3.266784
## [4,] 3.266784 3.266784 3.266784 5.316240
```

```
# Exponential covariance
expo$sigma^2 * corMatrix(expo$modelStruct$corStruct)[[1]]
```

```
##          [,1]     [,2]     [,3]     [,4]
## [1,] 5.296881 3.315144 2.074839 1.298574
## [2,] 3.315144 5.296881 3.315144 2.074839
## [3,] 2.074839 3.315144 5.296881 3.315144
## [4,] 1.298574 2.074839 3.315144 5.296881
```

```r
# Autoregressive covariance
auto1$sigma^2 * corMatrix(auto1$modelStruct$corStruct)[[1]]
```

```
##          [,1]     [,2]     [,3]     [,4]
## [1,] 5.296881 3.315144 2.074840 1.298574
## [2,] 3.315144 5.296881 3.315144 2.074840
## [3,] 2.074840 3.315144 5.296881 3.315144
## [4,] 1.298574 2.074840 3.315144 5.296881
```

The covariance estiamtes are similar across the 3 covariance patterns.