HW6_answer

Guojing Wu | gw2383

3/31/2019

Problem 1

$$
\begin{aligned}
varY_{i,j} &= var(\mu + b_i + e_{i,j}) \\
&= var(b_i) + var(e_{i,j}) \\
&= \sigma_b^2 + \sigma_e^2 \\
cov(Y_{i,j}, Y_{i,k}) &= E(Y_{i,j}Y_{i,k}) - E(Y_{i,j})E(Y_{i,k}) \\
&= E[(\mu + b_i + e_{i,j})(\mu + b_i + e_{i,k})] - E(Y_{i,j})E(Y_{i,k}) \\
&= E(\mu^2 + b_i^2 + 2\mu b_i + \mu e_{i,j} + \mu e_{i,k} + b_i e_{i,j} + b_i e_{i,k} + e_{i,j}e_{i,k}) - \mu^2
\end{aligned}
$$

Since $b_i \sim N(0, \sigma_b^2)$, $(\frac{b_i}{\sigma_b})^2 \sim \chi_1^2$, so $E(b_i^2) = \sigma_b^2$. And since $e_{i,j}, e_{i,k}$ are independent, $E(e_{i,j}e_{i,k}) = E(e_{i,j})(e_{i,k})$. So,

$$
\begin{aligned}
cov(Y_{i,j}, Y_{i,k}) &= \mu^2 + \sigma_b^2 - \mu^2 \\
&= \sigma_b^2 \\
corr(Y_{i,j}, Y_{i,k}) &= \frac{cov(Y_{i,j}, Y_{i,k})}{\sqrt{varY_{i,j}varY_{i,k}}} \\
&= \frac{\sigma_b^2}{\sigma_b^2 + \sigma_e^2}
\end{aligned}
$$

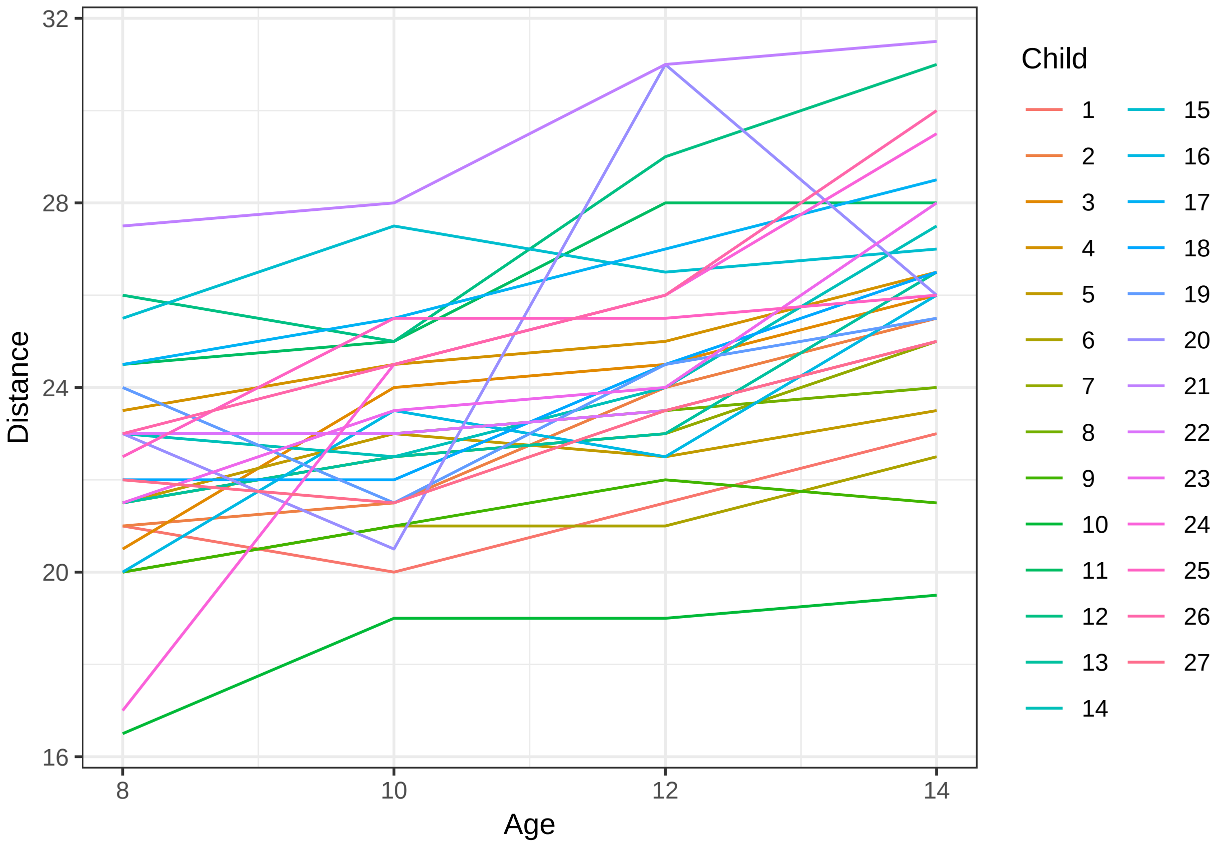
This suggest that the covariance pattern is compound symmetry

Problem 2

```
data_dist = read.table("HW6-dental.txt", header = T) %>%
  mutate(Child = as.factor(Child),
         Gender = as.factor(Gender))
```

1) spaghetti plot

```
data_dist %>%
  ggplot(aes(x = Age, y = Distance, group = Child, col = Child)) +
  geom_line()
```



In general, the older the child is, the longer the distance is.

2)

We assume that $Y_1 \sim Y_{16}$ is male and $Y_{17} \sim Y_{27}$ is female. The covariance within same person is σ_a^2, the covariance within the same gender is σ_b^2:

$$
\begin{aligned}
E(Y_{i,j}) &= \beta_0 + \beta_1 age_{i,j} \\
\Sigma_{Y_i} &= \begin{pmatrix}
\sigma_a^2+\sigma_b^2+\sigma_e^2 & \sigma_a^2+\sigma_b^2 & ... & \sigma_a^2+\sigma_b^2 \\
 & \sigma_a^2+\sigma_b^2+\sigma_e^2 & ... & \sigma_a^2+\sigma_b^2 \\
 & & ... & \\
 & & ... & \sigma_a^2+\sigma_b^2+\sigma_e^2
\end{pmatrix} \\
\mathbf{\Sigma} &= \begin{pmatrix}
\Sigma_{Y_1} & \sigma_b^2 E & ... & \sigma_b^2 E & 0 & 0 & & 0 \\
 & \Sigma_{Y_2} & ... & \sigma_b^2 E & 0 & 0 & & 0 \\
 & & ... & & & & & \\
 & & & \Sigma_{Y_{16}} & 0 & 0 & & 0 \\
 & & & & \Sigma_{Y_{17}} & \sigma_b^2 E & ... & \sigma_b^2 E \\
 & & & & & \Sigma_{Y_{18}} & ... & \sigma_b^2 E \\
 & & & & & & ... & \\
 & & & & & & ... & \Sigma_{Y_{27}}
\end{pmatrix}
\end{aligned}
$$

3)

i)

```r
CompSymm <- gls(Distance ~ Age + Gender, data_dist, correlation = corCompSymm(form = ~ 1 | Child), weig

Expo <- gls(Distance ~ Age + Gender, data_dist, correlation = corExp(form = ~ 1 | Child), weights = var

Auto <- gls(Distance ~ Age + Gender, data_dist, correlation = corAR1(form = ~ 1 | Child), weights = var

coeff = rbind(CompSymm$coefficients, Expo$coefficients, Auto$coefficients)
rownames(coeff) = c("CompSymmetry", "Exponential", "Autoregressive")
coeff %>% knitr::kable()
```

	(Intercept)	Age	Gender1
CompSymmetry	15.39576	0.6639005	2.151163
Exponential	15.35489	0.6593467	2.427299
Autoregressive	15.35488	0.6593470	2.427314

By the table above, the coefficient parameter estimates obtained by three methods are very similar.

ii)

```r
varfunc = rbind(CompSymm$modelStruct$varStruct, Expo$modelStruct$varStruct, Auto$modelStruct$varStruct)
varfunc = cbind(rep(0, nrow(varfunc)), varfunc)
rownames(varfunc) = c("CompSymmetry", "Exponential", "Autoregressive")
colnames(varfunc) = c("8", "10", "12", "14")
varfunc %>% exp() %>% knitr::kable()
```

	8	10	12	14
CompSymmetry	1	0.8745886	1.042036	0.9649285
Exponential	1	0.9033459	1.031869	0.9078637
Autoregressive	1	0.9033473	1.031880	0.9078541

By the table above, the variance function of different age obtained by three methods are very similar.

iii)

```r
getVarCov(CompSymm)
```

```
## Marginal variance covariance matrix
##        [,1]   [,2]   [,3]   [,4]
## [1,] 5.6136 3.0281 3.6078 3.3409
## [2,] 3.0281 4.2938 3.1554 2.9219
## [3,] 3.6078 3.1554 6.0954 3.4813
## [4,] 3.3409 2.9219 3.4813 5.2267
##   Standard Deviations: 2.3693 2.0722 2.4689 2.2862
```

```r
getVarCov(Expo)
```

```
## Marginal variance covariance matrix
##        [,1]   [,2]   [,3]   [,4]
## [1,] 5.7424 3.2866 2.3786 1.3259
## [2,] 3.2866 4.6860 3.3914 1.8905
## [3,] 2.3786 3.3914 6.1143 3.4083
```

```
## [4,] 1.3259 1.8905 3.4083 4.7330
##   Standard Deviations: 2.3963 2.1647 2.4727 2.1755
```

```r
getVarCov(Auto)
```

```
## Marginal variance covariance matrix
##         [,1]   [,2]   [,3]   [,4]
## [1,] 5.7425 3.2867 2.3787 1.3260
## [2,] 3.2867 4.6861 3.3915 1.8905
## [3,] 2.3787 3.3915 6.1145 3.4084
## [4,] 1.3260 1.8905 3.4084 4.7330
##   Standard Deviations: 2.3963 2.1647 2.4727 2.1755
```

By the table above, the covariance matrices matches each pattern. And since it's equal intervals of time, the autoregression and exponential methods are pretty similar.