24. R Code for Repeated Measures

- These slides illustrate a few example R commands for fitting generalized linear models to repeated measures data.

- We focus on the experiment designed to compare the effectiveness of three strength training programs.

- We will fit models that allows for a distinct mean for each of the $3 \times 7 = 21$ combinations of training program and time.

- We assume independence between subjects.
- The models differ in the choice for $\boldsymbol{W}$, which is the variance-covariance structure assumed for the 7 observations from each subject.

```
#Read the data

d=read.delim(
  "http://dnett.github.io/S510/RepeatedMeasures.txt")

#Create Factors

d$Program=as.factor(d$Program)
d$Subj=as.factor(d$Subj)
d$Timef=as.factor(d$Time)

#Load the nlme package

library(nlme)
```

Compound Symmetry Structure for $\boldsymbol{W}$

$$\begin{bmatrix} \sigma_e^2+\sigma_s^2 & \sigma_s^2 & \sigma_s^2 & \sigma_s^2 & \sigma_s^2 & \sigma_s^2 & \sigma_s^2 \\ \sigma_s^2 & \sigma_e^2+\sigma_s^2 & \sigma_s^2 & \sigma_s^2 & \sigma_s^2 & \sigma_s^2 & \sigma_s^2 \\ \sigma_s^2 & \sigma_s^2 & \sigma_e^2+\sigma_s^2 & \sigma_s^2 & \sigma_s^2 & \sigma_s^2 & \sigma_s^2 \\ \sigma_s^2 & \sigma_s^2 & \sigma_s^2 & \sigma_e^2+\sigma_s^2 & \sigma_s^2 & \sigma_s^2 & \sigma_s^2 \\ \sigma_s^2 & \sigma_s^2 & \sigma_s^2 & \sigma_s^2 & \sigma_e^2+\sigma_s^2 & \sigma_s^2 & \sigma_s^2 \\ \sigma_s^2 & \sigma_s^2 & \sigma_s^2 & \sigma_s^2 & \sigma_s^2 & \sigma_e^2+\sigma_s^2 & \sigma_s^2 \\ \sigma_s^2 & \sigma_s^2 & \sigma_s^2 & \sigma_s^2 & \sigma_s^2 & \sigma_s^2 & \sigma_e^2+\sigma_s^2 \end{bmatrix}$$

```
lme(Strength ~ Program * Timef, data = d,
               random = ~ 1 | Subj)
```

Alternative Parameterizaton for Compound Symmetry

$$\sigma^2 \begin{bmatrix} 1 & \rho & \rho & \rho & \rho & \rho & \rho \\ \rho & 1 & \rho & \rho & \rho & \rho & \rho \\ \rho & \rho & 1 & \rho & \rho & \rho & \rho \\ \rho & \rho & \rho & 1 & \rho & \rho & \rho \\ \rho & \rho & \rho & \rho & 1 & \rho & \rho \\ \rho & \rho & \rho & \rho & \rho & 1 & \rho \\ \rho & \rho & \rho & \rho & \rho & \rho & 1 \end{bmatrix}$$

```
gls(Strength ~ Program * Timef, data = d,
    correlation = corCompSymm(form = ~ 1 | Subj))
```

AR(1) Structure for $\boldsymbol{W}$

$$\sigma^2 \begin{bmatrix} 1 & \rho & \rho^2 & \rho^3 & \rho^4 & \rho^5 & \rho^6 \\ \rho & 1 & \rho & \rho^2 & \rho^3 & \rho^4 & \rho^5 \\ \rho^2 & \rho & 1 & \rho & \rho^2 & \rho^3 & \rho^4 \\ \rho^3 & \rho^2 & \rho & 1 & \rho & \rho^2 & \rho^3 \\ \rho^4 & \rho^3 & \rho^2 & \rho & 1 & \rho & \rho^2 \\ \rho^5 & \rho^4 & \rho^3 & \rho^2 & \rho & 1 & \rho \\ \rho^6 & \rho^5 & \rho^4 & \rho^3 & \rho^2 & \rho & 1 \end{bmatrix}$$

```
gls(Strength ~ Program * Timef, data = d,
    correlation = corAR1(form = ~ 1 | Subj))
```

General Positive Definite Structure for $\boldsymbol{W}$

With δ_1 set equal to 1 for identifiablility purposes, a general 7×7 positive definite variance-covariance matrix is parameterized by R as follows:

$$\sigma^2 \text{ diag}(\delta_1, ..., \delta_7) \begin{bmatrix} 1 & \rho_{12} & \rho_{13} & \rho_{14} & \rho_{15} & \rho_{16} & \rho_{17} \\ \rho_{12} & 1 & \rho_{23} & \rho_{24} & \rho_{25} & \rho_{26} & \rho_{27} \\ \rho_{13} & \rho_{23} & 1 & \rho_{34} & \rho_{35} & \rho_{36} & \rho_{37} \\ \rho_{14} & \rho_{24} & \rho_{34} & 1 & \rho_{45} & \rho_{46} & \rho_{47} \\ \rho_{15} & \rho_{25} & \rho_{35} & \rho_{45} & 1 & \rho_{56} & \rho_{57} \\ \rho_{16} & \rho_{26} & \rho_{36} & \rho_{46} & \rho_{56} & 1 & \rho_{67} \\ \rho_{17} & \rho_{27} & \rho_{37} & \rho_{47} & \rho_{57} & \rho_{67} & 1 \end{bmatrix} \text{ diag}(\delta_1, ..., \delta_7)$$

The 7×7 case doesn't fit on one slide, but here is the 5×5 case.

$$\begin{bmatrix} \sigma^2\delta_1^2 & \sigma^2\rho_{12}\delta_1\delta_2 & \sigma^2\rho_{13}\delta_1\delta_3 & \sigma^2\rho_{14}\delta_1\delta_4 & \sigma^2\rho_{15}\delta_1\delta_5 \\ \sigma^2\rho_{12}\delta_1\delta_2 & \sigma^2\delta_2^2 & \sigma^2\rho_{23}\delta_2\delta_3 & \sigma^2\rho_{24}\delta_2\delta_4 & \sigma^2\rho_{25}\delta_2\delta_5 \\ \sigma^2\rho_{13}\delta_1\delta_3 & \sigma^2\rho_{23}\delta_2\delta_3 & \sigma^2\delta_3^2 & \sigma^2\rho_{34}\delta_3\delta_4 & \sigma^2\rho_{35}\delta_3\delta_5 \\ \sigma^2\rho_{14}\delta_1\delta_4 & \sigma^2\rho_{24}\delta_2\delta_4 & \sigma^2\rho_{34}\delta_3\delta_4 & \sigma^2\delta_4^2 & \sigma^2\rho_{45}\delta_4\delta_5 \\ \sigma^2\rho_{15}\delta_1\delta_5 & \sigma^2\rho_{25}\delta_2\delta_5 & \sigma^2\rho_{35}\delta_3\delta_5 & \sigma^2\rho_{45}\delta_4\delta_5 & \sigma^2\delta_5^2 \end{bmatrix}$$

```
gls(Strength ~ Program * Timef, data = d,
    correlation = corSymm(form = ~ 1 | Subj),
    weight = varIdent(form = ~ 1 | Timef))
```

- To understand the reason for an identifiability constraint, notice that an arbitrary positive definite 7×7 covariance matrix depends on only

$$7 + 6 + 5 + 4 + 3 + 2 + 1 = \frac{7(7+1)}{2} = 28$$

parameters. However, we have $\sigma^2,\ 6 + 5 + 4 + 3 + 2 + 1 = 21\ \rho_{ij}$ parameters, and $\delta_1, \ldots, \delta_7$.

- That's 29 parameters for a symmetric positive definite matrix that depends on at most 28 parameters.

- Thus, R chooses to set δ_1 to 1.
- Without such a constraint, it is easy to use different values of the parameters to define the same matrix. For example,

$$\begin{bmatrix} 3 & -1 \\ -1 & 7 \end{bmatrix} = 3\begin{bmatrix} 1 & -\frac{1}{3} \\ -\frac{1}{3} & \frac{7}{3} \end{bmatrix} = 1\begin{bmatrix} 3 & -1 \\ -1 & 7 \end{bmatrix}$$

σ^2	3	1
δ_1	1	$\sqrt{3}$
δ_2	$\sqrt{\frac{7}{3}}$	$\sqrt{7}$
ρ_{12}	$\frac{-1}{3\sqrt{\frac{7}{3}}}$	$\frac{-1}{\sqrt{21}}$

Other Variance-Covariance Structures in R

If you are interested in learning about how to fit other variance-covariance structures in R, the following help commands may be useful.

```
?corClasses
?varClasses
?pdClasses
```

AIC and BIC for Repeated Measures in R

- $AIC = -2\ell(\hat{\boldsymbol{\theta}}) + 2k$
- $BIC = -2\ell(\hat{\boldsymbol{\theta}}) + k\ln(n)$
- $k =$ number of mean parameters (rank of $\boldsymbol{X}$)
  $+$ number of variance parameters
- $n =$ total number of observations $-$ rank of $\boldsymbol{X}$