In this case,

$$\boldsymbol{G} = \mathrm{Var}(\boldsymbol{u}) \quad = \quad \sigma_s^2 \boldsymbol{I}_{n.\times n.} \text{ and}$$

$$\boldsymbol{R} = \mathrm{Var}(\boldsymbol{e}) \quad = \quad \sigma_e^2 \boldsymbol{I}_{(7n.)\times(7n.)},$$

where $n. = n_1 + n_2 + n_3$ is the total number of subjects.

$$\boldsymbol{\Sigma} = \mathrm{Var}(\boldsymbol{y}) \quad = \quad \boldsymbol{Z}\boldsymbol{G}\boldsymbol{Z}^\top + \boldsymbol{R}$$

is a block diagonal matrix with one block of the form

$$\sigma_s^2 \mathbf{1}\mathbf{1}^\top_{7\times 7} + \sigma_e^2 \boldsymbol{I}_{7\times 7}$$

for each subject.

$$y = X\beta + Zu + e$$

$$G = \mathrm{Var}(\boldsymbol{u}) = \sigma_s^2 \boldsymbol{I}, \;\; \boldsymbol{R} = \mathrm{Var}(\boldsymbol{e}) = \sigma_e^2 \boldsymbol{I}, \text{ and}$$

$$\begin{aligned}
\boldsymbol{\Sigma} &= \mathrm{Var}(\boldsymbol{y}) = \boldsymbol{Z}\boldsymbol{G}\boldsymbol{Z}^\top + \boldsymbol{R} \\
&= \sigma_s^2 \begin{bmatrix} \mathbf{1}\mathbf{1}^\top & & & \\ & \mathbf{1}\mathbf{1}^\top & & \\ & & \ddots & \\ & & & \mathbf{1}\mathbf{1}^\top \end{bmatrix} + \sigma_e^2 \boldsymbol{I} \\
&= \begin{bmatrix} \underline{\sigma_e^2 \boldsymbol{I} + \sigma_s^2 \mathbf{1}\mathbf{1}^\top} & & \\ & \ddots & \\ & & \sigma_e^2 \boldsymbol{I} + \sigma_s^2 \mathbf{1}\mathbf{1}^\top \end{bmatrix}.
\end{aligned}$$

If predicting subject effects is not of interest and random subject effects are included only to introduce correlation among repeated measures on the same subject, we can work with an alternative expression of the same model by using the general linear model $\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$, where

$$\mathrm{Var}(\boldsymbol{y}) = \mathrm{Var}(\boldsymbol{\epsilon}) = \begin{bmatrix} \sigma_e^2\boldsymbol{I} + \sigma_s^2\boldsymbol{1}\boldsymbol{1}^\top & & \\ & \ddots & \\ & & \sigma_e^2\boldsymbol{I} + \sigma_s^2\boldsymbol{1}\boldsymbol{1}^\top \end{bmatrix}.$$

More generally, we can replace the mixed model

$$\boldsymbol{y} = \boldsymbol{X\beta} + \underbrace{\boldsymbol{Zu} + \boldsymbol{e}}$$

with the model

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{\epsilon}$$

where

$$\mathrm{Var}(\boldsymbol{y}) = \mathrm{Var}(\boldsymbol{\epsilon}) = \begin{bmatrix} \boldsymbol{W} & \boldsymbol{0} & \cdots & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{W} & \cdots & \boldsymbol{0} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{0} & \boldsymbol{0} & \cdots & \boldsymbol{W} \end{bmatrix}.$$

generally we assume W are identical across all subjects

- We can choose a structure for $\boldsymbol{W}$ that seems appropriate based on the design and the data.

requires estimation of σ_s^2 & σ_e^2

- One choice for $\boldsymbol{W}$ is a compound symmetric matrix like we have considered previously.

requires a much larger # of parameters to be estimated.

- Another choice for $\boldsymbol{W}$ is an unstructured positive definite matrix.

- A common choice for $\boldsymbol{W}$ when repeated measures are equally spaced in time is the first-order autoregressive structure known as AR(1).

AR(1): First-Order Autoregressive Covariance Structure

same σ^2 across time

$$\boldsymbol{W} = \sigma^2 \begin{bmatrix} 1 & \phi & \phi^2 & \phi^3 & \phi^4 & \phi^5 & \phi^6 \\ \phi & 1 & \phi & \phi^2 & \phi^3 & \phi^4 & \phi^5 \\ \phi^2 & \phi & 1 & \phi & \phi^2 & \phi^3 & \phi^4 \\ \phi^3 & \phi^2 & \phi & 1 & \phi & \phi^2 & \phi^3 \\ \phi^4 & \phi^3 & \phi^2 & \phi & 1 & \phi & \phi^2 \\ \phi^5 & \phi^4 & \phi^3 & \phi^2 & \phi & 1 & \phi \\ \phi^6 & \phi^5 & \phi^4 & \phi^3 & \phi^2 & \phi & 1 \end{bmatrix}$$

measurements further apart are less correlated

where $\sigma^2 \in (0, \infty)$ and $\phi \in (-1, 1)$ are unknown parameters.

- In the next slides, we will see how to fit a variety of general linear models that might be appropriate for repeated measures experiments.
- These slides illustrate a few example R commands for fitting general 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$Program = factor(d$Program)
d$Subj = factor(d$Subj)
d$Timef = factor(d$Time)

#Load the nlme package

library(nlme)
```

Compound Symmetry Structure for $\boldsymbol{W}$

$$\underset{\sim}{y} = X\underset{\sim}{\beta} + Z\underset{\sim}{u} + \underset{\sim}{e}$$

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

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

```
> summary(o.lme)
Linear mixed-effects model fit by REML
 Data: d
      AIC      BIC    logLik
  1466.82 1557.323 -710.4101
```

on main diagonal of W

$(3.099)^2 + (1.094)^2$

≈ 10.8

```
Random effects:
 Formula: ~1 | Subj
        (Intercept) Residual
StdDev:    3.098924 1.094017
```

$\hat{\sigma}_s$ $\hat{\sigma}_e$

$\hat{\sigma}_s^2 = (3.099)^2 \approx 9.6$

.
.
.

```
> #   Examine the estimated variance-covariance
> #   matrix for the subvector of responses
> #   from a single subject.
>
> getVarCov(o.lme, individuals = 1, type = "marginal")
Subj 1
Marginal variance covariance matrix
        1       2       3       4       5       6       7
1 10.8000  9.6033  9.6033  9.6033  9.6033  9.6033  9.6033
2  9.6033 10.8000  9.6033  9.6033  9.6033  9.6033  9.6033
3  9.6033  9.6033 10.8000  9.6033  9.6033  9.6033  9.6033
4  9.6033  9.6033  9.6033 10.8000  9.6033  9.6033  9.6033
5  9.6033  9.6033  9.6033  9.6033 10.8000  9.6033  9.6033
6  9.6033  9.6033  9.6033  9.6033  9.6033 10.8000  9.6033
7  9.6033  9.6033  9.6033  9.6033  9.6033  9.6033 10.8000
```

Alternative Parameterizaton for Compound Symmetry

$\underset{\sim}{y} = X\beta + \underset{\sim}{\epsilon}$

$\sigma^2 = Var(\underset{\sim}{y})$
$= Var(\underset{\sim}{\epsilon})$

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

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

```
> summary(o.cs)
Generalized least squares fit by REML
  Model: Strength ~ Program * Timef
  Data: d
      AIC      BIC    logLik
  1466.82 1557.323 -710.4101

Correlation Structure: Compound symmetry
 Formula: ~1 | Subj
 Parameter estimate(s):
      Rho
0.8891805
.
.
.
Residual standard error: 3.286366
Degrees of freedom: 399 total; 378 residual
```

$\hat{\rho}$

$\hat{\sigma}_\epsilon$

```
> getVarCov(o.cs)
Marginal variance covariance matrix
        [,1]    [,2]    [,3]    [,4]    [,5]    [,6]    [,7]
[1,] 10.8000  9.6033  9.6033  9.6033  9.6033  9.6033  9.6033
[2,]  9.6033 10.8000  9.6033  9.6033  9.6033  9.6033  9.6033
[3,]  9.6033  9.6033 10.8000  9.6033  9.6033  9.6033  9.6033
[4,]  9.6033  9.6033  9.6033 10.8000  9.6033  9.6033  9.6033
[5,]  9.6033  9.6033  9.6033  9.6033 10.8000  9.6033  9.6033
[6,]  9.6033  9.6033  9.6033  9.6033  9.6033 10.8000  9.6033
[7,]  9.6033  9.6033  9.6033  9.6033  9.6033  9.6033 10.8000
```

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

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

```
o.ar1 = gls(Strength ~ Program * Timef, data = d,
        correlation = corAR1(form = ~ 1 | Subj))
```

```
> summary(o.ar1)
Generalized least squares fit by REML
  Model: Strength ~ Program * Timef
  Data: d
       AIC      BIC    logLik
  1312.804 1403.306 -633.4018

Correlation Structure: AR(1)
 Formula: ~1 | Subj
 Parameter estimate(s):
      Phi
0.9517769   Φ̂
.
.
.
Residual standard error: 3.280242
Degrees of freedom: 399 total; 378 residual
```

```
> getVarCov(o.ar1, individual = 3)
Marginal variance covariance matrix
        [,1]    [,2]    [,3]    [,4]    [,5]    [,6]    [,7]
[1,] 10.7600 10.2410  9.7473  9.2772  8.8298  8.4040  7.9988
[2,] 10.2410 10.7600 10.2410  9.7473  9.2772  8.8298  8.4040
[3,]  9.7473 10.2410 10.7600 10.2410  9.7473  9.2772  8.8298
[4,]  9.2772  9.7473 10.2410 10.7600 10.2410  9.7473  9.2772
[5,]  8.8298  9.2772  9.7473 10.2410 10.7600 10.2410  9.7473
[6,]  8.4040  8.8298  9.2772  9.7473 10.2410 10.7600 10.2410
[7,]  7.9988  8.4040  8.8298  9.2772  9.7473 10.2410 10.7600
```

getting smaller

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:

most flexibility in terms correlation between any observations over time within the same individual

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

allow for different Var(y) across time

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

$\sigma^2 \rho_{ij} \delta_i \delta_j$

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

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

allows variance to change over time

```
> summary(o.un)
Generalized least squares fit by REML
  Model: Strength ~ Program * Timef
  Data: d
       AIC      BIC    logLik
  1332.896 1525.706 -617.4479
```

```
Correlation Structure: General
 Formula: ~1 | Subj
 Parameter estimate(s):
 Correlation:
  1     2     3     4     5     6
2 0.960
3 0.925 0.940
4 0.872 0.877 0.956
5 0.842 0.860 0.937 0.960
6 0.809 0.827 0.898 0.909 0.951
7 0.797 0.792 0.876 0.887 0.917 0.953
```

lower triangular part of W

similarly to AR(1) observations further apart in time are less correlated!

```
Variance function:
 Structure: Different standard deviations per stratum
 Formula: ~1 | Timef
 Parameter estimates:
    2     4     6     8    10    12    14
1.000 1.039 1.104 1.071 1.174 1.157 1.203
.
.
.
Residual standard error: 2.963129
Degrees of freedom: 399 total; 378 residual
```

δ_1 $\hat{\delta}_2$ $\hat{\delta}_3$ . . . . . . $\hat{\delta}_7$

```
> getVarCov(o.un, individual = 3)
Marginal variance covariance matrix
        [,1]   [,2]    [,3]    [,4]    [,5]    [,6]    [,7]
[1,] 8.7801 8.7571  8.9656  8.1984  8.6781  8.2203  8.4169
[2,] 8.7571 9.4730  9.4631  8.5686  9.2012  8.7307  8.6875
[3,] 8.9656 9.4631 10.7080  9.9266 10.6660 10.0700 10.2140
[4,] 8.1984 8.5686  9.9266 10.0770 10.6000  9.8987 10.0430
[5,] 8.6781 9.2012 10.6660 10.6000 12.0950 11.3440 11.3640
[6,] 8.2203 8.7307 10.0700  9.8987 11.3440 11.7560 11.6500
[7,] 8.4169 8.6875 10.2140 10.0430 11.3640 11.6500 12.7100
```

- 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 σ^2, $6 + 5 + 4 + 3 + 2 + 1 = 21$ ρ parameters, and $\delta_1, \ldots, \delta_7$.

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

to ensure identifiability R set $\delta_1 = 1$

- 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,

2 time points

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

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

end lecture 05-02-25

39