# Multivariate Analysis of Variance

## Within-Subjects Designs

Andrew Zieffler

# Differences between RM-ANOVA and RM-MANOVA

*Within-Subjects Mixed Model ANOVA; RM-ANOVA;* or *Univariate Mixed Model ANOVA*

- There is only **one response variable**—the values repeatedly observed over time
- There is only **one sum of squares** of the treatment effect (effect of time)

*Multivariate Analysis of Variance* or *RM-MANOVA*

- There are **multiple response variables**
- There are multiple **sum of squares** for the treatment effect (effect of time)

Multivariate approach to repeated measures is based on **difference score(s)**

$$D_i = Y_{\mathrm{T2}} - Y_{\mathrm{T1}}$$

Order of subtraction is arbitrary

$$\boxed{D_1 = 194 - 172 = +22}$$

Difference score indicates nature of change for given *i*

- Student 1's reading score increased over time
- Student 16's reading score decreased over time

Interpretation based on order of subtraction

| | Time | | |
|---|---|---|---|
| Student | T1 | T2 | *D* |
| 1 | 172 | 194 | +22 |
| 3 | 191 | 215 | +24 |
| 5 | 207 | 213 | +6 |
| 6 | 191 | 195 | +4 |
| 7 | 199 | 218 | +19 |
| 9 | 149 | 177 | +28 |
| 11 | 218 | 239 | +21 |
| 12 | 228 | 246 | +18 |
| 13 | 228 | 239 | +11 |
| 14 | 199 | 235 | +36 |
| 16 | 228 | 227 | –1 |
| 17 | 201 | 219 | +18 |
| 18 | 218 | 221 | +3 |
| 20 | 204 | 214 | +10 |
| *M:* | *202.4* | *218* | *+15.64* |

# Mean Difference Scores are the Key

Applied researchers are generally more interested in the mean change over time

$$\bar{D} = \frac{\sum D_i}{N}$$

Difference of the two treatment means is equal to the mean of the difference scores between the treatments

$$\bar{D} = \bar{Y}_1 - \bar{Y}_2$$

**Important Implication**

$$\text{If } \bar{Y}_1 = \bar{Y}_2, \text{ then } \bar{D} = 0$$

$$H_0 : \mu_{.1} - \mu_{.2} = 0$$

is equivalent to

$$H_0 : \delta = 0$$

Multivariate approach to repeated measures uses a **model comparison method** for hypothesis testing, similar to that for comparing nested ANOVA models.

Specify two competing models:

- Full model
- Reduced model (fewer parameters)

# Competing Models in MANOVA

**Statistical Model**

$$D_i = \delta + \epsilon_i$$

Population mean difference

*i*th subject's deviation from the mean difference

Hypothesis 1

$$\delta \neq 0$$

To obtain the competing models, we make two assumptions about delta

Hypothesis 2

$$\delta = 0$$

$$D_i = \delta + \epsilon_i$$

**Full model**

The $D_i$ are influenced by an effect of time (delta) *and* sampling (random) error

$$D_i = 0 + \epsilon_i$$
$$D_i = \epsilon_i$$

**Reduced model**

The $D_i$ are only influenced by sampling (random) error

# Examining the Residuals

Estimate the residuals for both models from the sample data

**Full model**

$$\hat{\epsilon}_i = D_i - \hat{\delta}$$

$$\hat{\epsilon}_i = D_i - 15.6$$

**Reduced model**

$$\hat{\epsilon}_i = D_i$$

| | Time | | | Errors | |
|---|---|---|---|---|---|
| Subject | T1 | T2 | *D* | Full | Red. |
| 1 | 172 | 194 | +22 | +6.4 | +22 |
| 3 | 191 | 215 | +24 | +8.4 | +24 |
| 5 | 207 | 213 | +6 | –9.6 | +6 |
| 6 | 191 | 195 | +4 | –11.6 | +4 |
| 7 | 199 | 218 | +19 | +3.4 | +19 |
| 9 | 149 | 177 | +28 | +12.4 | +28 |
| 11 | 218 | 239 | +21 | +5.4 | +21 |
| 12 | 228 | 246 | +18 | +2.4 | +18 |
| 13 | 228 | 239 | +11 | –4.6 | +11 |
| 14 | 199 | 235 | +36 | +20.4 | +36 |
| 16 | 228 | 227 | –1 | –16.6 | –1 |
| 17 | 201 | 219 | +18 | +2.4 | +18 |
| 18 | 218 | 221 | +3 | –12.6 | +3 |
| 20 | 204 | 214 | +10 | –5.6 | +10 |
| *M:* | *202.4* | *218* | *+15.6* | | |

| | Time | | | Errors | |
|---|---|---|---|---|---|
| Subject | T1 | T2 | *D* | Full | Red. |
| 1 | 172 | 194 | +22 | +6.4 | +22 |
| 3 | 191 | 215 | +24 | +8.4 | +24 |
| 5 | 207 | 213 | +6 | –9.6 | +6 |
| 6 | 191 | 195 | +4 | –11.6 | +4 |
| 7 | 199 | 218 | +19 | +3.4 | +19 |
| 9 | 149 | 177 | +28 | +12.4 | +28 |
| 11 | 218 | 239 | +21 | +5.4 | +21 |
| 12 | 228 | 246 | +18 | +2.4 | +18 |
| 13 | 228 | 239 | +11 | –4.6 | +11 |
| 14 | 199 | 235 | +36 | +20.4 | +36 |
| 16 | 228 | 227 | –1 | –16.6 | –1 |
| 17 | 201 | 219 | +18 | +2.4 | +18 |
| 18 | 218 | 221 | +3 | –12.6 | +3 |
| 20 | 204 | 214 | +10 | –5.6 | +10 |
| *M:* | *202.4* | *218* | *+15.6* | | |

$$\mathrm{SSE} = \sum \epsilon_i^2$$

**Full model**

$$\begin{aligned}\mathrm{SSE} &= (6.4)^2 + (8.4)^2 + \ldots + (-5.6)^2 \\ &= 1447.214\end{aligned}$$

**Reduced model**

$$\begin{aligned}\mathrm{SSE} &= (22)^2 + (24)^2 + \ldots + (10)^2 \\ &= 4873\end{aligned}$$

The full model (based on the hypothesis that delta ≠ 0) fits better than the reduced model (based on the hypothesis that delta = 0). Is superiority of fit just due to sampling error?

# ΔF-Test

The $\Delta F$-test is test used to compare two nested models. It examines whether the change in the SSE between the models is within what would be expected due to sampling error.

Testing requires examination of SS relative to the *df*.

*k* is the number of parameters in the reduced model (intercept and delta)

$$\Delta F = \frac{\left(\frac{\text{SS}_{\text{Reduced}} - \text{SS}_{\text{Full}}}{k-1}\right)}{\left(\frac{\text{SS}_{\text{Full}}}{N-k+1}\right)}$$

$$\Delta F = \frac{\text{SS}_{\text{Reduced}} - \text{SS}_{\text{Full}}}{\text{SS}_{\text{Full}}} \cdot \frac{N-k+1}{k-1}$$

**Our Data**

$$\Delta F = \frac{\left(\frac{4873-1447.2}{2-1}\right)}{\left(\frac{1447.2}{14-2+1}\right)}$$

$$\Delta F(1,\ 13) = 30.8$$

```
> 1 - pf(30.8, 1, 13)

[1] 9.388262e-05
```

Reject the null hypothesis, $F(1, 13) = 30.8$, $p < 0.001$.

# MANOVA using R

```
> dvm = cbind(mpls3$read.5, mpls3$read.8)
```

Use the wide-formatted data to create a matrix of the repeated measures

```
> mlm.1 = lm(dvm ~ 1)
> mlm.1

Call:
lm(formula = dvm ~ 1)

Coefficients:
             [,1]   [,2]
(Intercept)  202.4  218.0
```

Fit a multivariate linear model

Note that the coefficients are equal to the means of each time point.

```
> my.design = factor(c("read.5", "read.8"))
```

Create a variable that defines the intra-subject design based on the levels of the repeated measures

```
> library(car)

> Anova(
    mlm.1,
    idata = data.frame(my.design),
    idesign = ~ my.design
    )
```

Use the `Anova()` function from the **car** library

The `idata=` argument takes a data frame that has defined the intra-subject design of the repeated measures

The `idesign=` argument takes a one-sided model describes the intra-subject design of the repeated measures

```
Note: model has only an intercept; equivalent type-III tests substituted.

Type III Repeated Measures MANOVA Tests: Pillai test statistic
            Df test stat approx F num Df den Df    Pr(>F)    
(Intercept)  1   0.99143  1503.39      1     13 8.009e-15 ***
my.design    1   0.70301    30.77      1     13 9.427e-05 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

# Many Different Multivariate Test Statistics

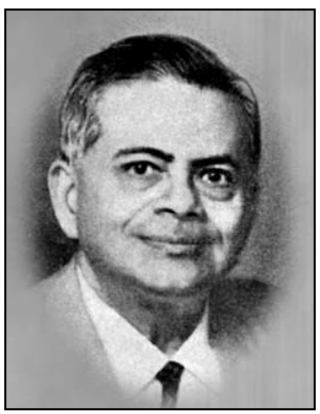
Maurice Stevenson Bartlett

K. C. Sreedharan Pillai

Pillai–Bartlett Trace
(Pillai's Trace)

Wilk's $\Lambda$ (Lambda)

Samuel Stanley Wilks



Samarendra Nath Roy

Roy's Largest Root

Harold Hotelling

Hotelling–Lawley Trace
(Hotelling's Trace)


Derrick Norman Lawley

```
## Wilk's lambda
> Anova(mlm.1, idata = data.frame(my.design), idesign = ~ my.design, test = "Wilks")

Type III Repeated Measures MANOVA Tests: Wilks test statistic
            Df test stat approx F num Df den Df    Pr(>F)
(Intercept)  1  0.008573  1503.39      1     13 8.009e-15 ***
my.design    1  0.296986    30.77      1     13 9.427e-05 ***
---


## Hotelling-Lawley trace
> Anova(mlm.1, idata = data.frame(my.design), idesign = ~ my.design,
    test = "Hotelling-Lawley")

Type III Repeated Measures MANOVA Tests: Hotelling-Lawley test statistic
            Df test stat approx F num Df den Df    Pr(>F)
(Intercept)  1   115.646  1503.39      1     13 8.009e-15 ***
my.design    1     2.367    30.77      1     13 9.427e-05 ***
---


## Roy's largest root
> Anova(mlm.1, idata = data.frame(my.design), idesign = ~ my.design, test = "Roy")

Type III Repeated Measures MANOVA Tests: Roy test statistic
            Df test stat approx F num Df den Df    Pr(>F)
(Intercept)  1   115.646  1503.39      1     13 8.009e-15 ***
my.design    1     2.367    30.77      1     13 9.427e-05 ***
---
```

- The *F*-value based on Wilk's Lambda has been found to be the best choice for a number of reasons (e.g., Davis, 2002)
  - ✓ Wilk's Lamda = 0.297, $F(1, 3) = 30.77$, $p < 0.001$

- With only two time points all of the multivariate tests are equivalent
  - ✓ The RM-ANOVA (univariate) approach gives the exact same results

```
## RM-ANOVA
> aov.2 <- ezANOVA(data = mplsLong, dv = read, wid = studentID, within = .(grade),
    detailed = TRUE)
> aov.2

$ANOVA
       Effect DFn DFd         SSn        SSd          F            p p<.05       ges
1 (Intercept)   1  13 1236900.893 10695.6071 1503.39400 8.009359e-15     * 0.9908523
2       grade   1  13    1712.893   723.6071   30.77306 9.426897e-05     * 0.1304355
```

# Three Time Points

"Multivariate" refers to that when there are more than two time points, there are **multiple difference scores** (i.e., multiple response variables).

$$D_1 = Y_{\text{Grade 6}} - Y_{\text{Grade 5}}$$
$$D_2 = Y_{\text{Grade 7}} - Y_{\text{Grade 6}}$$
$$D_3 = Y_{\text{Grade 8}} - Y_{\text{Grade 7}}$$

Only need ($k$ – 1) difference scores, where $k$ is the number of time points.

| | Grade | | | | | | |
|---|---|---|---|---|---|---|---|
| Student | 5 | 6 | 7 | 8 | *D* | *D* | *D* |
| 1 | 172 | 185 | 179 | 194 | +13 | –6 | +15 |
| 3 | 191 | 199 | 203 | 215 | +8 | +4 | +12 |
| 5 | 207 | 213 | 212 | 213 | +6 | –1 | +1 |
| 6 | 191 | 189 | 206 | 195 | –2 | +17 | –11 |
| 7 | 199 | 208 | 213 | 218 | +9 | +5 | +5 |
| 9 | 149 | 154 | 174 | 177 | +5 | +20 | +3 |
| 11 | 218 | 231 | 233 | 239 | +13 | +2 | +6 |
| 12 | 228 | 232 | 248 | 246 | +4 | +16 | –2 |
| 13 | 228 | 236 | 228 | 239 | +8 | –8 | +11 |
| 14 | 199 | 210 | 225 | 235 | +11 | +15 | +10 |
| 16 | 228 | 226 | 234 | 227 | –2 | +8 | –7 |
| 17 | 201 | 210 | 208 | 219 | +9 | –2 | +11 |
| 18 | 218 | 220 | 217 | 221 | +2 | –3 | +4 |
| 20 | 204 | 215 | 219 | 214 | +11 | +4 | –5 |
| ***M:*** | ***202.4*** | ***209.1*** | ***214.2*** | ***218*** | ***+6.8*** | ***+5.1*** | ***+3.8*** |

Is there a main effect of treatment?

$$H_0 : \mu_{\cdot 1} = \mu_{\cdot 2} = \mu_{\cdot 3}$$

Under the null hypothesis, all of the treatment means are equal. This implies that the differences between treatment means are all zero.

The null hypothesis can be re-stated as

$$H_0 : \delta_1 = \delta_2 = 0$$

Multivariate approach to repeated measures tests the two (or more) difference scores simultaneously

Similar to univariate approach we will compute the residuals for both the full and reduced model, and then compare the fit. In MANOVA, the $\Delta F$-test is an extension of that for two time points.

In the **two time point** case

- One difference score ($D_{i1}$)
- $SS_{Reduced}$ is the sum of the squared $D_1$ scores
- $SS_{Full}$ is the sum of the squared residuals based on the deviations between $D_i$ and $\bar{D}$

In the **three time point** case

- **Two** difference scores ($D_{i1}$ and $D_{i2}$)
- **Two** sets of equations/residuals for the reduced model
- **Two** sets of equations/residuals for the full model

# Competing Models

Model assumed under hypothesis delta ≠ 0

The $D_i$ are influenced by the treatment effect and sampling (random) error

**Full model**

$$D_{i1} = \delta_1 + e_{i1}$$
$$D_{i2} = \delta_2 + e_{i2}$$
$$D_{i3} = \delta_3 + e_{i3}$$

Residuals

$$e_{i1} = D_{i1} - \delta_1$$
$$e_{i2} = D_{i2} - \delta_2$$
$$e_{i3} = D_{i3} - \delta_3$$

Model assumed under hypothesis delta = 0

The $D_i$ are only influenced by sampling (random) error

**Reduced model**

$$D_{i1} = e_{i1}$$
$$D_{i2} = e_{i2}$$
$$D_{i3} = e_{i3}$$

Residuals

$$e_{i1} = D_{i1}$$
$$e_{i2} = D_{i2}$$
$$e_{i3} = D_{i3}$$

<u>Assumption</u>

$\delta_1 = \delta_2 = 0$

| | Grade | | | | | | | Reduced model | | | Full model | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Student | 5 | 6 | 7 | 8 | *D* | *D* | *D* | *e* | *e* | *e* | *e* | *e* | *e* |
| 1 | 172 | 185 | 179 | 194 | +13 | –6 | +15 | +13 | –6 | +15 | +6.2 | –11.1 | +11.2 |
| 3 | 191 | 199 | 203 | 215 | +8 | +4 | +12 | +8 | +4 | +12 | +1.2 | +8.2 | +8.2 |
| 5 | 207 | 213 | 212 | 213 | +6 | –1 | +1 | +6 | –1 | +1 | –0.8 | –2.8 | –2.8 |
| 6 | 191 | 189 | 206 | 195 | –2 | +17 | –11 | –2 | +17 | –11 | –8.8 | –14.8 | –14.8 |
| 7 | 199 | 208 | 213 | 218 | +9 | +5 | +5 | +9 | +5 | +5 | +2.2 | +1.2 | +1.2 |
| 9 | 149 | 154 | 174 | 177 | +5 | +20 | +3 | +5 | +20 | +3 | –1.8 | –0.8 | –0.8 |
| 11 | 218 | 231 | 233 | 239 | +13 | +2 | +6 | +13 | +2 | +6 | +6.2 | +2.2 | +2.2 |
| 12 | 228 | 232 | 248 | 246 | +4 | +16 | –2 | +4 | +16 | –2 | –2.8 | –5.8 | –5.8 |
| 13 | 228 | 236 | 228 | 239 | +8 | –8 | +11 | +8 | –8 | +11 | +1.2 | +7.2 | +7.2 |
| 14 | 199 | 210 | 225 | 235 | +11 | +15 | +10 | +11 | +15 | +10 | +4.2 | +6.2 | +6.2 |
| 16 | 228 | 226 | 234 | 227 | –2 | +8 | –7 | –2 | +8 | –7 | –8.8 | –10.8 | –10.8 |
| 17 | 201 | 210 | 208 | 219 | +9 | –2 | +11 | +9 | –2 | +11 | +2.2 | +7.2 | +7.2 |
| 18 | 218 | 220 | 217 | 221 | +2 | –3 | +4 | +2 | –3 | +4 | –4.8 | +0.2 | +0.2 |
| 20 | 204 | 215 | 219 | 214 | +11 | +4 | –5 | +11 | +4 | –5 | +4.2 | –8.8 | –8.8 |
| ***M:*** | ***202.4*** | ***209.1*** | ***214.2*** | ***218*** | ***+6.8*** | ***+5.1*** | ***+3.8*** | | | | | | |

| Student | Reduced model | | | Full model | | |
|---|---|---|---|---|---|---|
| | $e$ | $e$ | $e$ | $e$ | $e$ | $e$ |
| 1 | +13 | –6 | +15 | +6.2 | –11.1 | +11.2 |
| 3 | +8 | +4 | +12 | +1.2 | +8.2 | +8.2 |
| 5 | +6 | –1 | +1 | –0.8 | –2.8 | –2.8 |
| 6 | –2 | +17 | –11 | –8.8 | –14.8 | –14.8 |
| 7 | +9 | +5 | +5 | +2.2 | +1.2 | +1.2 |
| 9 | +5 | +20 | +3 | –1.8 | –0.8 | –0.8 |
| 11 | +13 | +2 | +6 | +6.2 | +2.2 | +2.2 |
| 12 | +4 | +16 | –2 | –2.8 | –5.8 | –5.8 |
| 13 | +8 | –8 | +11 | +1.2 | +7.2 | +7.2 |
| 14 | +11 | +15 | +10 | +4.2 | +6.2 | +6.2 |
| 16 | –2 | +8 | –7 | –8.8 | –10.8 | –10.8 |
| 17 | +9 | –2 | +11 | +2.2 | +7.2 | +7.2 |
| 18 | +2 | –3 | +4 | –4.8 | +0.2 | +0.2 |
| 20 | +11 | +4 | –5 | +4.2 | –8.8 | –8.8 |

**Reduced Model**

$$\mathrm{SSE}_1 = 959$$
$$\mathrm{SSE}_2 = 1409$$
$$\mathrm{SSE}_3 = 997$$

**Full Model**

$$\mathrm{SSE}_1 = 314.36$$
$$\mathrm{SSE}_2 = 1048.94$$
$$\mathrm{SSE}_3 = 796.36$$

# Sum of Cross-Products

In order to compute the $\Delta F$ statistic, we also need to incorporate information about the correlation between the residuals (e.g., $r_{e_{i1},e_{i2}}$ for both the full and reduced models)

To do this we compute the cross-products of the residuals and then find the **sum of the cross-products**

The numerator of a covariance is actually a sum of cross-products.

$$SCP_{(12)} = \sum (e_{i1} \times e_{i2})$$

This sum is computed for both the full and reduced models.

| | Reduced model | | | | | | Full model | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Student | *e* | *e* | *e* | *e* | *e* | *e* | *e* | *e* | *e* | *e* | *e* | *e* |
| 1 | +13 | –6 | +15 | –78 | +195 | –90 | +6.2 | –11.1 | +11.2 | –68.8 | +69.4 | –124.3 |
| 3 | +8 | +4 | +12 | +32 | +96 | +48 | +1.2 | +8.2 | +8.2 | –1.3 | +9.8 | –9.0 |
| 5 | +6 | –1 | +1 | –6 | +6 | –1 | –0.8 | –2.8 | –2.8 | +4.9 | +2.2 | +17.1 |
| 6 | –2 | +17 | –11 | –34 | +22 | –187 | –8.8 | –14.8 | –14.8 | –104.7 | +130.2 | –176.1 |
| 7 | +9 | +5 | +5 | +45 | +45 | +25 | +2.2 | +1.2 | +1.2 | –0.2 | +2.6 | –0.1 |
| 9 | +5 | +20 | +3 | +100 | +15 | +60 | –1.8 | –0.8 | –0.8 | –26.8 | +1.4 | –11.9 |
| 11 | +13 | +2 | +6 | +26 | +78 | +12 | +6.2 | +2.2 | +2.2 | –19.2 | +13.6 | –6.8 |
| 12 | +4 | +16 | –2 | +64 | –8 | –32 | –2.8 | –5.8 | –5.8 | –30.5 | +16.2 | –63.2 |
| 13 | +8 | –8 | +11 | –64 | +88 | –88 | +1.2 | +7.2 | +7.2 | –15.7 | +8.6 | –94.3 |
| 14 | +11 | +15 | +10 | +165 | +110 | +150 | +4.2 | +6.2 | +6.2 | +41.6 | +26.0 | +61.3 |
| 16 | –2 | +8 | –7 | –16 | +14 | –56 | –8.8 | –10.8 | –10.8 | –25.5 | +95.0 | –31.3 |
| 17 | +9 | –2 | +11 | –18 | +99 | –22 | +2.2 | +7.2 | +7.2 | –15.6 | +15.8 | –51.1 |
| 18 | +2 | –3 | +4 | –6 | +8 | –12 | –4.8 | +0.2 | +0.2 | +38.9 | –1.0 | –1.6 |
| 20 | +11 | +4 | –5 | +44 | –55 | –20 | +4.2 | –8.8 | –8.8 | –4.6 | –37.0 | +9.7 |
| **SSE** | **959** | **1409** | **997** | | | | **314.5** | **1048.9** | **796.4** | | | |
| **SCP** | | | | **254** | **713** | **–213** | | | | **–227.8** | **353.4** | **–481.8** |

# SSCP Matrices

We have six pieces of quantitative information that we must include in the omnibus $\Delta F$ test:

1. Sum of squares based on $e_{i1}$
2. Sum of squares based on $e_{i2}$
3. Sum of squares based on $e_{i3}$
4. Information about the correlation between $e_{i1}$ and $e_{i2}$ contained in the sum of the cross-products
5. Information about the correlation between $e_{i1}$ and $e_{i3}$ contained in the sum of the cross-products
6. Information about the correlation between $e_{i2}$ and $e_{i3}$ contained in the sum of the cross-products

We will organize these three pieces of quantitative information in a matrix called the **Sum of squares and cross-products matrix (SSCP)**.

$$\text{SSCP} = \begin{bmatrix} \text{SS}_1 & \text{SCP}_{1,2} & \text{SCP}_{1,3} \\ \text{SCP}_{2,1} & \text{SS}_2 & \text{SCP}_{2,3} \\ \text{SCP}_{3,1} & \text{SCP}_{3,2} & \text{SS}_3 \end{bmatrix}$$

We compute a SSCP matrix for both the reduced and full models.

# SSCP Matrices

**Reduced Model**

$$\mathrm{SSCP}_{\mathrm{Reduced}} = \begin{bmatrix} 959 & 254 & 713 \\ 254 & 1409 & -213 \\ 713 & -213 & 997 \end{bmatrix}$$

**Full Model**

$$\mathrm{SSCP}_{\mathrm{Full}} = \begin{bmatrix} 314.5 & -227.8 & 353.4 \\ -227.8 & 1048.9 & -481.8 \\ 353.4 & -481.8 & 796.4 \end{bmatrix}$$

```
## Create the SSCP for the reduced model
> X = as.matrix(mpls2[c("d1", "d2", "d3")])
> sscp_r = t(X) %*% X
> sscp_r

    d1   d2   d3
d1 959  254  713
d2 254 1409 -213
d3 713 -213  997
```

```
## Create the SSCP for the full model
> Y = as.matrix(mpls2[c("e1", "e2", "e3")])
> sscp_f = t(Y) %*% Y
> sscp_f

          e1        e2        e3
e1  314.3571 -227.7857  353.3571
e2 -227.7857 1048.9286 -481.7857
e3  353.3571 -481.7857  796.3571
```

# Determinants of the SSCP Matrices

The $\Delta F$ test only accommodates scalars (single numbers). We **summarize the information in the matrix** into a single value by computing the **determinant**.

Determinants of matrices that are larger than 2x2 have much more complex formulas and should be computed using technology.

```
## Compute the determinants
> det(sscp_r)
[1] 445904611

> det(sscp_f)
[1] 94888555
```

**Smaller determinant** indicates better fit. The full model seems to fit better.

# Compute the $\Delta F$ Statistic

$$\Delta F = \frac{\frac{det(\mathrm{SSCP}_{\mathrm{Reduced}}) - det(\mathrm{SSCP}_{\mathrm{Full}})}{k-1}}{\frac{det(\mathrm{SSCP}_{\mathrm{Full}})}{N-k+1}}$$

$N$ = sample size
$k$ = Number of measurement waves

$$\Delta F = \frac{\frac{445904611 - 94888555}{4-1}}{\frac{94888555}{14-4+1}} = 13.56$$

This would be evaluated in the *F*-distribution having 3 and 11 *df*.

# Using R

```
## Fit the MANOVA
> dvm = cbind(mpls2$read.5, mpls2$read.6, mpls2$read.7, mpls2$read.8)
> mlm.1 <- lm(dvm ~ 1)
> mlm.1

Call:
lm(formula = dvm ~ 1)

Coefficients:
             [,1]   [,2]   [,3]   [,4]
(Intercept)  202.4  209.1  214.2  218.0

## Specify within-subjects design
> my.design = factor(c(“read.5”, “read.6”, “read.7”, “read.8”))

## Obtain MANOVA results
> Anova(mlm.1, idata = data.frame(my.design), idesign = ~ my.design, test = ”Wilks”)
```

# Results

```
Note: model has only an intercept; equivalent type-III tests substituted.

Type III Repeated Measures MANOVA Tests: Wilks test statistic
            Df test stat approx F num Df den Df    Pr(>F)
(Intercept)  1  0.008633  1492.86      1     13 8.381e-15 ***
my.design    1  0.212800    13.56      3     11 0.0005143 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Note: If the result would have been statistically significant, we would have had to do post hoc tests to find out which difference was significantly difference from 0.

There are different ideas about what the correct plan of attack should be when this occurs (univariate analysis for each difference, contrast tests, trend analysis, etc.)

Rather than dealing with this here, we will instead forge on to the linear mixed-effects model.