ORIGIN ≡ 0

W. Stein

2-Way ANOVA as Linear Models - Unbalanced Example

Standard Two-way ANOVA involves decomposition of Treatment Sum of Squares (SSTR) into Sum of Squares for two factors A & B (SSA & SSB) plus Sums of Squares Interaction (SSI). In balanced two-way ANOVA the decomposition is "orthogonal" in the sense that SSTR=SSA+SSB+SSI. In unbalanced two-way ANOVA (as well as other designs), this simple relationship may not hold, and explicit formulas for calculating the decomposition of SSTR are hard to find. However, treating ANOVA as a Linear Model allows testing through a standard GLM strategy. As with one-way ANOVA, p classification or "indicator" variables are coded into a mimimally specified set "dummy variables" by use of specifically designed contrast matrices. Example below comes from Chapter 23 in Kuter et al. (KNNL) A*pplied Linear Statistical Models* 5th Edition.

Example: **KNNL dataset Table 23.1**

Cell Means ANOVA Model:

< Here the independent variables were explicitly coded into p=6 Indicator variables X1-X6 with each column representing a cell ij.

K := READPRN("c:/2008LinearModelsData/GrowthHormoneCM.txt")

Variable Assignment:

$Y := K^{\langle 0 \rangle}$

$j := 1 .. 6$

$X^{\langle j-1 \rangle} := K^{\langle j \rangle}$ **< design matrix**

$N := \text{length}(Y) \quad N = 14$ **< total number of cases**

$r := \text{cols}(X) \quad r = 6$

$p := r$ **< p used previously, r = p to conform with KNNL**

$i := 0 .. N-1$

$ii := 0 .. N-1$

$I := \text{identity}(N) \quad J_{i,ii} := 1$ **< Identity & One Matrix for matrix calculations**

$n := \begin{pmatrix} 3 & 2 & 2 \\ 1 & 3 & 3 \end{pmatrix}$ **< number of objects in each block**

$$K = \begin{pmatrix} 1.4 & 1 & 0 & 0 & 0 & 0 & 0 \\ 2.4 & 1 & 0 & 0 & 0 & 0 & 0 \\ 2.2 & 1 & 0 & 0 & 0 & 0 & 0 \\ 2.1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1.7 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0.7 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1.1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 2.4 & 0 & 0 & 0 & 1 & 0 & 0 \\ 2.5 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1.8 & 0 & 0 & 0 & 0 & 1 & 0 \\ 2 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0.5 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0.9 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1.3 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \quad X = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\mu := \begin{pmatrix} \text{mean}(1.4,2.4,2.2) & \text{mean}(2.1,1.7) & \text{mean}(0.7,1.1) \\ \text{mean}(2.4) & \text{mean}(2.5,1.8,2) & \text{mean}(0.5,0.9,1.3) \end{pmatrix} \quad \mu = \begin{pmatrix} 2 & 1.9 & 0.9 \\ 2.4 & 2.1 & 0.9 \end{pmatrix}$$ **< means for each block**

Least Squares Estimation of the Regression Parameters:

$$\beta := \left(X^T \cdot X\right)^{-1} \cdot X^T \cdot Y \qquad \beta = \begin{pmatrix} 2 \\ 1.9 \\ 0.9 \\ 2.4 \\ 2.1 \\ 0.9 \end{pmatrix} \quad \begin{pmatrix} \beta1(\mu11) \\ \beta2(\mu12) \\ \beta3(\mu13) \\ \beta4(\mu21) \\ \beta5(\mu22) \\ \beta6(\mu23) \end{pmatrix}$$ **< KNNL Cell Means Model**

Fitted Values & Hat Matrix H:

$Y_h := X \cdot \beta$ **< fitted values Yh**

$H := X \cdot \left(X^T \cdot X\right)^{-1} \cdot X^T$ **< nXn Hat matrix**

Residuals:

$e := Y - Y_h$ **< residuals**

Full Model ANOVA Table for Cell Means Model:

Sum of Squares:		Degrees of Freedom:		Mean Squares:	
$SSTR := Y^T \cdot \left[H - \left(\frac{1}{N}\right) \cdot J\right] \cdot Y$	$SSTR = (4.4743)$	$df_R := p - 1$	$df_R = 5$	$MSTR := \frac{SSTR}{df_R}$	$MSTR = (0.8949)$
$SSE := Y^T \cdot (I - H) \cdot Y$	$SSE = (1.3)$	$df_E := N - p$	$df_E = 8$	$MSE := \frac{SSE}{df_E}$	$MSE = (0.1625)$
$SSTO := Y^T \cdot \left[I - \left(\frac{1}{N}\right) \cdot J\right] \cdot Y$	$SSTO = (5.7743)$	$df_T := N - 1$	$df_T = 13$	$MSTO := \frac{SSTO}{df_T}$	$MSTO = (0.4442)$

^ values verified KNNL p. 957

Explicit Cell Means Decomposition of SSTR:

$\mu c := \mu - \text{mean}(Y)$ $\mu c = \begin{pmatrix} 0.35714 & 0.25714 & -0.74286 \\ 0.75714 & 0.45714 & -0.74286 \end{pmatrix}$ **< centered means for each block**

Note: Vector calculations in MathCad below treat each position in a matrix as as separate scalar calculation.

$sq\mu c := \overrightarrow{(\mu c \cdot \mu c)}$ $sq\mu c = \begin{pmatrix} 0.12755 & 0.06612 & 0.55184 \\ 0.57327 & 0.20898 & 0.55184 \end{pmatrix}$ **< squared means for each block, i.e., square of each item in μc**

$SSCM := \overrightarrow{(n \cdot sq\mu c)}$ $SSCM = \begin{pmatrix} 0.38265 & 0.13224 & 1.10367 \\ 0.57327 & 0.62694 & 1.65551 \end{pmatrix}$ **< each square multiplied by number of each block (n)**

$SS := \text{stack}\left[\left(SSCM^T\right)^{\langle 0 \rangle}, \left(SSCM^T\right)^{\langle 1 \rangle}\right]$ **^ same as anova{} & Anova{car} reports in R**

$SS = \begin{pmatrix} 0.3827 \\ 0.1322 \\ 1.1037 \\ 0.5733 \\ 0.6269 \\ 1.6555 \end{pmatrix}$ $\begin{pmatrix} SS(\mu 11) \\ SS(\mu 12) \\ SS(\mu 13) \\ SS(\mu 21) \\ SS(\mu 22) \\ SS(\mu 23) \end{pmatrix}$ $\sum SS = 4.4743$ **< sum to SSTR above**

"Extra" Sums of Squares are determined by subtraction of SSE R - SSE F (for Reduced model versus Full model respectively). Reduced models are obtained by deletion of one X variable at a time from the desing matrix for FMCs.

GLM Decomposition of SSTR:

> Anova(FMCs) #TYPE 3 ANOVA SS KNNL CELL MEANS MODEL ITEMIZED...

```
Anova Table (Type II tests)
Response: Y
           Sum Sq Df F value  Pr(>F)
X1        0.38265  1  2.3548 0.16344
X2        0.13224  1  0.8138 0.39334
X3        1.10367  1  6.7918 0.03132 *
X4        0.57327  1  3.5278 0.09717 .
X5        0.62694  1  3.8581 0.08510 .
X6        1.65551  1 10.1878 0.01277 *
Residuals 1.30000  8
```

> drop1(FMCs)

```
Single term deletions
Model:
Y ~ X1 + X2 + X3 + X4 + X5 + X6 - 1
       Df Sum of Sq    RSS     AIC
<none>              1.3000 -21.2737
X1      1    0.3827 1.6827 -19.6616
X2      1    0.1322 1.4322 -21.9174
X3      1    1.1037 2.4037 -14.6688
X4      1    0.5733 1.8733 -18.1592
X5      1    0.6269 1.9269 -17.7637
X6      1    1.6555 2.9555 -11.7754
```

Prototype in R:

```
#TWO-WAY ANOVA LINEAR MODELS
require(car)

#CELL MEANS MODEL IN R
#READ STRUCTURED DATA TABLE WITH CELL MEANS CODING
K=read.table("c:/2008LinearModelsData/KNNL23.8R.txt")
K
attach(K)
Y=Y
X=cbind(K$X1,K$X2,K$X3,K$X4,K$X5,K$X6)
X

FMc=lm(Y~X-1)
FMC=lm(Y~X1+X2+X3+X4+X5+X6-1)
summary(FMc)  #CELL MEANS AS ESTIMATES
summary(FMC)  #CELL MEANS AS ESTIMATES
```

```
> K
    i j k   Y X1 X2 X3 X4 X5 X6 Z1 Z2 Z3 Z4
1   1 1 1 1.4  1  0  0  0  0  0  1  0  0  0
2   1 1 2 2.4  1  0  0  0  0  0  1  0  0  0
3   1 1 3 2.2  1  0  0  0  0  0  1  0  0  0
4   1 2 1 2.1  0  1  0  0  0  0  0  1  0  0
5   1 2 2 1.7  0  1  0  0  0  0  0  1  0  0
6   1 3 1 0.7  0  0  1  0  0  0  0  0  1  0
7   1 3 2 1.1  0  0  1  0  0  0  0  0  1  0
8   2 1 1 2.4  0  0  0  1  0  0 -2  2  0  1
9   2 2 1 2.5  0  0  0  0  1  0  0  0  0  1
10  2 2 2 1.8  0  0  0  0  1  0  0  0  0  1
11  2 2 3 2.0  0  0  0  0  1  0  0  0  0  1
12  2 3 1 0.5  0  0  0  0  0  1  0  2 -2  1
13  2 3 2 0.9  0  0  0  0  0  1  0  2 -2  1
14  2 3 3 1.3  0  0  0  0  0  1  0  2 -2  1
```

```
> summary(FMc)  #CELL MEANS AS ESTIMATES
Call:
lm(formula = Y ~ X - 1)
Residuals:
       Min         1Q     Median         3Q        Max
-6.000e-01 -2.000e-01 -4.272e-17  2.000e-01  4.000e-01
Coefficients:
   Estimate Std. Error t value Pr(>|t|)
X1   2.0000     0.2327   8.593 2.60e-05 ***
X2   1.9000     0.2850   6.666 0.000158 ***
X3   0.9000     0.2850   3.157 0.013447 *
X4   2.4000     0.4031   5.954 0.000341 ***
X5   2.1000     0.2327   9.023 1.82e-05 ***
X6   0.9000     0.2327   3.867 0.004761 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Residual standard error: 0.4031 on 8 degrees of freedom
Multiple R-squared: 0.9702,     Adjusted R-squared: 0.9478
F-statistic: 43.34 on 6 and 8 DF,  p-value: 1.134e-05
```

$$\beta = \begin{pmatrix} 2 \\ 1.9 \\ 0.9 \\ 2.4 \\ 2.1 \\ 0.9 \end{pmatrix}$$

^ regression coefficients are the cell means

```
> X
      [,1] [,2] [,3] [,4] [,5] [,6]
 [1,]    1    0    0    0    0    0
 [2,]    1    0    0    0    0    0
 [3,]    1    0    0    0    0    0
 [4,]    0    1    0    0    0    0
 [5,]    0    1    0    0    0    0
 [6,]    0    0    1    0    0    0
 [7,]    0    0    1    0    0    0
 [8,]    0    0    0    1    0    0
 [9,]    0    0    0    0    1    0
[10,]    0    0    0    0    1    0
[11,]    0    0    0    0    1    0
[12,]    0    0    0    0    0    1
[13,]    0    0    0    0    0    1
[14,]    0    0    0    0    0    1
```

^ Cell Means design matrix for two factors

```
#CENTERING CELL MEANS MODEL
Y=scale(Y,scale=F)
FMcs=lm(Y~X-1)
FMCs=lm(Y~X1+X2+X3+X4+X5+X6-1)

anova(FMcs) #TYPE 1 ANOVA SS KNNL CELL MEANS MODEL FOR LUMPED X VARIABLE
anova(FMCs) #TYPE 1 ANOVA SS KNNL CELL MEANS MODEL ITEMIZED FOR EACH X VARIABLE
Anova(FMCs) #TYPE 2 ANOVA SS KNNL CELL MEANS MODEL ITEMIZED...
Anova(FMCs) #TYPE 3 ANOVA SS KNNL CELL MEANS MODEL ITEMIZED...
detach(K)
```

> anova(FMCs) #TYPE 1 ANOVA SS KNNL CELL MEANS MODEL ITEMIZED FOR EACH X VARIABLE

```
Analysis of Variance Table
Response: Y
          Df  Sum Sq Mean Sq F value  Pr(>F)
X1         1 0.38265 0.38265  2.3548 0.16344
X2         1 0.13224 0.13224  0.8138 0.39334
X3         1 1.10367 1.10367  6.7918 0.03132 *
X4         1 0.57327 0.57327  3.5278 0.09717 .
X5         1 0.62694 0.62694  3.8581 0.08510 .
X6         1 1.65551 1.65551 10.1878 0.01277 *
Residuals  8 1.30000 0.16250
```

> Anova(FMCs) #TYPE 2 ANOVA SS KNNL CELL MEANS MODEL ITEMIZED...

```
Anova Table (Type II tests)
Response: Y
           Sum Sq Df F value  Pr(>F)
X1        0.38265  1  2.3548 0.16344
X2        0.13224  1  0.8138 0.39334
X3        1.10367  1  6.7918 0.03132 *
X4        0.57327  1  3.5278 0.09717 .
X5        0.62694  1  3.8581 0.08510 .
X6        1.65551  1 10.1878 0.01277 *
Residuals 1.30000  8
```

Results identical for all ANOVA > Type I-III SS

> Anova(FMCs,type="3") #TYPE 3 ANOVA SS KNNL CELL MEANS MODEL ITEMIZED...

```
Anova Table (Type III tests)
Response: Y
           Sum Sq Df F value  Pr(>F)
X1        0.38265  1  2.3548 0.16344
X2        0.13224  1  0.8138 0.39334
X3        1.10367  1  6.7918 0.03132 *
X4        0.57327  1  3.5278 0.09717 .
X5        0.62694  1  3.8581 0.08510 .
X6        1.65551  1 10.1878 0.01277 *
Residuals 1.30000  8
```

$$\text{SS} = \begin{pmatrix} 0.38265 \\ 0.13224 \\ 1.10367 \\ 0.57327 \\ 0.62694 \\ 1.65551 \end{pmatrix}$$

< Same results as above for all ANOVA functions

Treatment ANOVA Model in R:

< dummy variables derived from default use of factor() function in R

K := READPRN("c:/2008LinearModelsData/GrowthHormoneTR.txt")

Variable Assignment:

$Y := K^{\langle 0 \rangle}$

$j := 1 .. \text{cols}(K) - 1$

$X^{\langle j-1 \rangle} := K^{\langle j \rangle}$ **< design matrix**

$N := \text{length}(Y) \quad N = 14$ **< total number of cases**

$r := \text{cols}(X) \quad r = 6$

$p := r$ **< p used previously, r = p to conform with KNNL**

$i := 0 .. N - 1$

$ii := 0 .. N - 1$

$I := \text{identity}(N) \quad J_{i,ii} := 1$ **< Identity & One Matrix for matrix calculations**

$$K = \begin{pmatrix} 1.4 & 1 & 0 & 0 & 0 & 0 & 0 \\ 2.4 & 1 & 0 & 0 & 0 & 0 & 0 \\ 2.2 & 1 & 0 & 0 & 0 & 0 & 0 \\ 2.1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1.7 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0.7 & 1 & 0 & 0 & 1 & 0 & 0 \\ 1.1 & 1 & 0 & 0 & 1 & 0 & 0 \\ 2.4 & 1 & 1 & 0 & 0 & 0 & 0 \\ 2.5 & 1 & 1 & 1 & 0 & 1 & 0 \\ 1.8 & 1 & 1 & 1 & 0 & 1 & 0 \\ 2 & 1 & 1 & 1 & 0 & 1 & 0 \\ 0.5 & 1 & 1 & 0 & 1 & 0 & 1 \\ 0.9 & 1 & 1 & 0 & 1 & 0 & 1 \\ 1.3 & 1 & 1 & 0 & 1 & 0 & 1 \end{pmatrix} \quad X = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 & 1 \end{pmatrix}$$

Least Squares Estimation of the Regression Parameters:

$$\beta := \left(X^T \cdot X\right)^{-1} \cdot X^T \cdot Y$$

$$\beta = \begin{pmatrix} 2 \\ 0.4 \\ -0.1 \\ -1.1 \\ -0.2 \\ -0.4 \end{pmatrix} \quad \begin{pmatrix} \text{Intercept}(\mu_{11}) \\ \beta A2(\mu_{21} - \mu_{11}) \\ \beta B2(\mu_{12} - \mu_{11}) \\ \beta B3(\mu_{13} - \mu_{11}) \\ \beta A2B2(\mu_{22} - \mu_{21} - \mu_{12} + \mu_{11}) \\ \beta A2B3(\mu_{23} - \mu_{21} - \mu_{13} + \mu_{11}) \end{pmatrix}$$

< regression parameters of Treatment model in R.

$$\mu = \begin{pmatrix} 2 & 1.9 & 0.9 \\ 2.4 & 2.1 & 0.9 \end{pmatrix}$$

^ cell means from above

$$\begin{pmatrix} 2 \\ 2.4 - 2 \\ 1.9 - 2 \\ 0.9 - 2 \\ 2.1 - 2.4 - 1.9 + 2 \\ 0.9 - 2.4 - 0.9 + 2 \end{pmatrix} = \begin{pmatrix} 2 \\ 0.4 \\ -0.1 \\ -1.1 \\ -0.2 \\ -0.4 \end{pmatrix}$$

$$\mu T := \begin{pmatrix} \beta_0 & \beta_2 & \beta_3 \\ \beta_1 & \beta_4 & \beta_5 \end{pmatrix} \quad \mu T = \begin{pmatrix} 2 & -0.1 & -1.1 \\ 0.4 & -0.2 & -0.4 \end{pmatrix}$$

^ calculation of Treatment model in R regression coefficients from μ

Fitted Values & Hat Matrix H:

$Y_h := X \cdot \beta$ **< fitted values Yh**

$H := X \cdot \left(X^T \cdot X\right)^{-1} \cdot X^T$ **< nXn Hat matrix**

Residuals:

$e := Y - Y_h$ **< residuals**

Full Model ANOVA Table for Treatments Model in R:

Sum of Squares:		Degrees of Freedom:		Mean Squares:	
$SSTR := Y^T \cdot \left[H - \left(\frac{1}{N}\right) \cdot J\right] \cdot Y$	$SSTR = (4.4743)$	$df_R := p - 1$	$df_R = 5$	$MSTR := \frac{SSTR}{df_R}$	$MSTR = (0.8949)$
$SSE := Y^T \cdot (I - H) \cdot Y$	$SSE = (1.3)$	$df_E := N - p$	$df_E = 8$	$MSE := \frac{SSE}{df_E}$	$MSE = (0.1625)$
$SSTO := Y^T \cdot \left[I - \left(\frac{1}{N}\right) \cdot J\right] \cdot Y$	$SSTO = (5.7743)$	$df_T := N - 1$	$df_T = 13$	$MSTO := \frac{SSTO}{df_T}$	$MSTO = (0.4442)$

^ values verified KNNL p. 957

GLM Decomposition of SSTR for Treatment Model in R:

```
> drop1(FM2)
Single term deletions

Model:
Y ~ A2 + B2 + B3 + A2.B2 + A2.B3
       Df Sum of Sq     RSS      AIC
<none>                1.3000 -21.2737
A2      1    0.1200   1.4200 -22.0376
B2      1    0.0120   1.3120 -23.1451
B3      1    1.4520   2.7520 -12.7742
A2.B2   1    0.0185   1.3185 -23.0763
A2.B3   1    0.0738   1.3738 -22.5002
```

```
> Anova(FM2,type="III")
Anova Table (Type III tests)

Response: Y
             Sum Sq Df F value  Pr(>F)
(Intercept) 12.0000  1 73.8462 2.6e-05 ***
A2           0.1200  1  0.7385 0.41516
B2           0.0120  1  0.0738 0.79270
B3           1.4520  1  8.9354 0.01735 *
A2.B2        0.0185  1  0.1136 0.74474
A2.B3        0.0738  1  0.4544 0.51923
Residuals    1.3000  8
```

Extra SS for single factors in the design matrix are determined by subtraction: SSE R - SSE F. This produces R's so-called "Type III" (or "Type 3") ANOVA *Marginal* Extra SS. Type III SS include "higher order" interaction terms such as A2.B2 & A2.B3 in FM even when a RM *excludes* a first order factor such as A2 or B2 or B3. As stated in documentation for Anova{car}, Type III ANOVA SS therefore are not marginal in an important sense.

```
> Anova(FM,type="III") #TYPE 3 ANOVA SS TREATMENTS MODEL IN R
Anova Table (Type III tests)

Response: Y
             Sum Sq Df F value  Pr(>F)
(Intercept) 12.0000  1 73.8462 2.6e-05 ***
A            0.1200  1  0.7385 0.41516
B            1.6171  2  4.9758 0.03944 *
A:B          0.0754  2  0.2321 0.79803
Residuals    1.3000  8
```

B= Extra SS value for Reduced Model excluding B2 & B3 from design matrix. A:B=Extra SS value for Reduced Model excluding both interaction terms for Reduced Model

Extra SS for factors represented by more than one column in the design matrix are similarly calculated by subtraction: SSE R - SSE F. They *may not* be calculated, in general, by addition of Extra SS for single columns in the design matrix, such as above.

Prototype in R:

```
#TREATMENTS MODEL IN R
#READ STRUCTURED DATA TABLE WITH NUMERIC CODED FACTOR
K=read.table("c:/2008LinearModelsData/GrowthHormoneR.txt")
K
attach(K)
Y=Rate
A=factor(GenderA)
B=factor(BoneDevB)
detach(K)

#CROSS TABULATIONS VERIFYING CELL MEANS
XX=xtabs(Y~A+B)
XX   #SUM OF Y VALUES IN EACH BLOCK
n=table(A,B)
n    #COUNTS FOR EACH BLOCK
N=length(Y)
N    #TOTAL NUMBER OF OBJECTS
mu=XX/n
mu   #MEANS FOR EACH BLOCK

FM=lm(Y~A*B)
summary(FM)
MM=data.frame(model.matrix(FM))
MM
```

```
> K
   Rate GenderA BoneDevB Replicate
1   1.4       1        1         1
2   2.4       1        1         2
3   2.2       1        1         3
4   2.1       1        2         1
5   1.7       1        2         2
6   0.7       1        3         1
7   1.1       1        3         2
8   2.4       2        1         1
9   2.5       2        2         1
10  1.8       2        2         2
11  2.0       2        2         3
12  0.5       2        3         1
13  0.9       2        3         2
14  1.3       2        3         3
```

```
> MM
   X.Intercept. A2 B2 B3 A2.B2 A2.B3
1             1  0  0  0     0     0
2             1  0  0  0     0     0
3             1  0  0  0     0     0
4             1  0  1  0     0     0
5             1  0  1  0     0     0
6             1  0  0  1     0     0
7             1  0  0  1     0     0
8             1  1  0  0     0     0
9             1  1  1  0     1     0
10            1  1  1  0     1     0
11            1  1  1  0     1     0
12            1  1  0  1     0     1
13            1  1  0  1     0     1
14            1  1  0  1     0     1
```

```
> XX   #SUM OF Y VALUES IN EACH BLOCK
   B
A     1   2   3
  1 6.0 3.8 1.8
  2 2.4 6.3 2.7

> n    #COUNTS FOR EACH BLOCK
   B
A   1 2 3
  1 3 2 2
  2 1 3 3

> mu   #MEANS FOR EACH BLOCK
   B
A     1   2   3
  1 2.0 1.9 0.9
  2 2.4 2.1 0.9
```

```
> summary(FM)

Call:
lm(formula = Y ~ A * B)

Residuals:
       Min         1Q     Median         3Q        Max 
-6.000e-01 -2.000e-01 -1.288e-16  2.000e-01  4.000e-01 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)   2.0000     0.2327   8.593  2.6e-05 ***
A2            0.4000     0.4655   0.859   0.4152    
B2           -0.1000     0.3680  -0.272   0.7927    
B3           -1.1000     0.3680  -2.989   0.0174 *  
A2:B2        -0.2000     0.5934  -0.337   0.7447    
A2:B3        -0.4000     0.5934  -0.674   0.5192    
```

```
anova(FM)  #TYPE 1 ANOVA SS TREATMENTS MODEL IN R
Anova(FM)  #TYPE 2 ANOVA SS TREATMENTS MODEL IN R
Anova(FM,type="III") #TYPE 3 ANOVA SS TREATMENTS MODEL IN R
```

Note differences in Extra SS reported here!

Serial SS with Extra SS calculated in step-wise fashion from botton to top. >

> anova(FM) #TYPE 1 ANOVA SS TREATMENTS MODEL IN R

```
Analysis of Variance Table
Response: Y
          Df Sum Sq Mean Sq F value   Pr(>F)
A          1 0.0029  0.0029  0.0176 0.897785
B          2 4.3960  2.1980 13.5262 0.002713 **
A:B        2 0.0754  0.0377  0.2321 0.798034
Residuals  8 1.3000  0.1625
```

marginal SS excluding interaction terms in FM if the base factors are deleted in RM >

> Anova(FM) #TYPE 2 ANOVA SS TREATMENTS MODEL IN R

```
Anova Table (Type II tests)
Response: Y
          Sum Sq Df F value   Pr(>F)
A         0.0926  1  0.5697 0.472022
B         4.3960  2 13.5262 0.002713 **
A:B       0.0754  2  0.2321 0.798034
Residuals 1.3000  8
```

> Anova(FM,type="III") #TYPE 3 ANOVA SS TREATMENTS MODEL IN R

```
Anova Table (Type III tests)
Response: Y
             Sum Sq Df F value  Pr(>F)
(Intercept) 12.0000  1 73.8462 2.6e-05 ***
A            0.1200  1  0.7385 0.41516
B            1.6171  2  4.9758 0.03944 *
A:B          0.0754  2  0.2321 0.79803
Residuals    1.3000  8
```

incompletely marginal SS which include interaction "higher order" terms in FM even if base factors are deleted from RM >

#GLM TESTS FOR TREATMENTS MODEL IN R

```
attach(MM)
RMi=lm(Y~A2+B2+B3)
RMa=lm(Y~B2+B3+A2.B2+A2.B3)
RMb=lm(Y~A2+A2.B2+A2.B3)
anova(RMi,FM) #FOR INTERACTIONS
anova(RMa,FM) #FOR FACTOR A
anova(RMb,FM) #FOR FACTOR B   #NOTE DIFFERENT RESULT HERE vs KNNL
```

> anova(RMi,FM) #FOR INTERACTIONS

```
Analysis of Variance Table
Model 1: Y ~ A2 + B2 + B3
Model 2: Y ~ A * B
  Res.Df     RSS Df Sum of Sq      F Pr(>F)
1     10 1.37543
2      8 1.30000  2   0.07543 0.2321  0.798
```

> anova(RMa,FM) #FOR FACTOR A

```
Analysis of Variance Table
Model 1: Y ~ B2 + B3 + A2.B2 + A2.B3
Model 2: Y ~ A * B
  Res.Df  RSS Df Sum of Sq      F Pr(>F)
1      9 1.42
2      8 1.30  1      0.12 0.7385 0.4152
```

> anova(RMb,FM) #FOR FACTOR B #NOTE DIFFERENT RESULT HERE vs KNNL

```
Analysis of Variance Table

Model 1: Y ~ A2 + A2.B2 + A2.B3
Model 2: Y ~ A * B
  Res.Df    RSS Df Sum of Sq      F  Pr(>F)
1     10 2.9171
2      8 1.3000  2    1.6171 4.9758 0.03944 *
```

This Extra SS differs from KNNL results calculated with "Treatment Effects" model below. The hypothesis tested is different: Ho: (μ12-μ11) = 0 = (μ13-μ11) >

Treatment Effects Model = contr.sum ANOVA Model in R: **KNNL p.955**

$K := \text{READPRN}(\text{"c:/2008LinearModelsData/GrowthHormone2.txt"})$

< dummy variables derived from function contrasts()=contr.sum in R. In balanced ANOVA, this produces the Treatment Effects model of KNNL called contr.sum in R.

Variable Assignment:

$Y := K^{\langle 3 \rangle}$ $\quad N := \text{length}(Y) \quad N = 14$

$j := 4 .. \text{cols}(K) - 1$

$Xs^{\langle j-4 \rangle} := K^{\langle j \rangle}$

$i := 0 .. N - 1 \qquad OV_i := 1$

$X := \text{augment}(OV, Xs)$ **< design matrix**

$X_F := X$ **< saving X for future reference**

$r := \text{cols}(X) \qquad r = 6$

$p := r$

$ii := 0 .. N - 1$

$I := \text{identity}(N) \qquad J_{i,ii} := 1$

$$K = \begin{pmatrix} 1 & 1 & 1 & 1.4 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 2 & 2.4 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 3 & 2.2 & 1 & 1 & 0 & 1 & 0 \\ 1 & 2 & 1 & 2.1 & 1 & 0 & 1 & 0 & 1 \\ 1 & 2 & 2 & 1.7 & 1 & 0 & 1 & 0 & 1 \\ 1 & 3 & 1 & 0.7 & 1 & -1 & -1 & -1 & -1 \\ 1 & 3 & 2 & 1.1 & 1 & -1 & -1 & -1 & -1 \\ 2 & 1 & 1 & 2.4 & -1 & 1 & 0 & -1 & 0 \\ 2 & 2 & 1 & 2.5 & -1 & 0 & 1 & 0 & -1 \\ 2 & 2 & 2 & 1.8 & -1 & 0 & 1 & 0 & -1 \\ 2 & 2 & 3 & 2 & -1 & 0 & 1 & 0 & -1 \\ 2 & 3 & 1 & 0.5 & -1 & -1 & -1 & 1 & 1 \\ 2 & 3 & 2 & 0.9 & -1 & -1 & -1 & 1 & 1 \\ 2 & 3 & 3 & 1.3 & -1 & -1 & -1 & 1 & 1 \end{pmatrix} \qquad X = \begin{pmatrix} 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & -1 & -1 & -1 & -1 \\ 1 & 1 & -1 & -1 & -1 & -1 \\ 1 & -1 & 1 & 0 & -1 & 0 \\ 1 & -1 & 0 & 1 & 0 & -1 \\ 1 & -1 & 0 & 1 & 0 & -1 \\ 1 & -1 & 0 & 1 & 0 & -1 \\ 1 & -1 & -1 & -1 & 1 & 1 \\ 1 & -1 & -1 & -1 & 1 & 1 \\ 1 & -1 & -1 & -1 & 1 & 1 \end{pmatrix}$$

Least Squares Estimation of the Regression Parameters:

$\beta := \left(X^T \cdot X\right)^{-1} \cdot X^T \cdot Y$

$$\beta = \begin{pmatrix} 1.7 \\ -0.1 \\ 0.5 \\ 0.3 \\ -0.1 \\ 0 \end{pmatrix} \quad \begin{pmatrix} \text{Intercept}(\mu_{..}) \\ \beta A1(\mu_{1.} - \mu_{..}) \\ \beta B1(\mu_{.1} - \mu_{..}) \\ \beta B2(\mu_{.2} - \mu_{..}) \\ \beta A1B1(\mu_{11} - \mu_{1.} - \mu_{.1} + \mu_{..}) \\ \beta A1B2(\mu_{12} - \mu_{1.} - \mu_{.2} + \mu_{..}) \end{pmatrix}$$

< regression parameters from contr.sums model in R, KNNL p. 956.

$\mu = \begin{pmatrix} 2 & 1.9 & 0.9 \\ 2.4 & 2.1 & 0.9 \end{pmatrix}$ **< cell means from above**

$\mu 11 := \mu_{0,0} \qquad \mu 12 := \mu_{0,1}$

$\mu_{1dot} := \text{mean}\left[\left(\mu^T\right)^{\langle 0 \rangle}\right] \qquad \mu_{1dot} = 1.6$

$\mu_{2dot} := \text{mean}\left[\left(\mu^T\right)^{\langle 1 \rangle}\right] \qquad \mu_{2dot} = 1.8$

$\mu_{dot1} := \text{mean}\left(\mu^{\langle 0 \rangle}\right) \qquad \mu_{dot1} = 2.2$

$\mu_{dot2} := \text{mean}\left(\mu^{\langle 1 \rangle}\right) \qquad \mu_{dot2} = 2$

$\mu_{dot3} := \text{mean}\left(\mu^{\langle 2 \rangle}\right) \qquad \mu_{dot3} = 0.9$

Explicit calculation of regression parameters.

$\mu_{dotdot} := \text{mean}(\mu_{1dot}, \mu_{2dot}) \qquad \mu_{dotdot} = 1.7 \qquad \text{mean}(\mu_{dot1}, \mu_{dot2}, \mu_{dot3}) = 1.7 \qquad \text{mean}(2.0, 1.9, 0.9, 2.4, 2.1, 0.9) = 1.7$

^ μ.. mean of means in KNNL p. 954, Table 23.1 >

$$\begin{pmatrix} \mu_{dotdot} \\ \mu_{1dot} - \mu_{dotdot} \\ \mu_{dot1} - \mu_{dotdot} \\ \mu_{dot2} - \mu_{dotdot} \\ \mu 11 - \mu_{1dot} - \mu_{dot1} + \mu_{dotdot} \\ \mu 12 - \mu_{1dot} - \mu_{dot2} + \mu_{dotdot} \end{pmatrix} = \begin{pmatrix} 1.7 \\ -0.1 \\ 0.5 \\ 0.3 \\ -0.1 \\ 0 \end{pmatrix}$$

Fitted Values & Hat Matrix H:

$Y_h := X \cdot \beta$ **< fitted values Yh**

$H := X \cdot \left(X^T \cdot X\right)^{-1} \cdot X^T$ **< nXn Hat matrix**

Residuals:

$e := Y - Y_h$ **< residuals**

Full Model ANOVA Table for Treatments Effects Model = contr.sum in R:

Sum of Squares: **Degrees of Freedom:** **Mean Squares:**

$SSTR := Y^T \cdot \left[H - \left(\frac{1}{N}\right) \cdot J \right] \cdot Y$ $SSTR = (4.4743)$ $df_R := p - 1$ $df_R = 5$ $MSTR := \frac{SSTR}{df_R}$ $MSTR = (0.8949)$

$SSE := Y^T \cdot (I - H) \cdot Y$ $SSE = (1.3)$ $df_E := N - p$ $df_E = 8$ $MSE := \frac{SSE}{df_E}$ $MSE = (0.1625)$

$SSTO := Y^T \cdot \left[I - \left(\frac{1}{N}\right) \cdot J \right] \cdot Y$ $SSTO = (5.7743)$ $df_T := N - 1$ $df_T = 13$ $MSTO := \frac{SSTO}{df_T}$ $MSTO = (0.4442)$

$SSE_F := SSE_0$ **< assigning Full Model SSE for GLM Tests below**

$df_F := df_E$ **< assigning Full Model df_E for GLM Tests below**

GLM Decomposition of SSTR for Treatment Effects = contr.sum in R:

```
> FM2=lm(Y~A1+B1+B2+A1.B1+A1.B2)
> drop1(FM2)
Single term deletions

Model:
Y ~ A1 + B1 + B2 + A1.B1 + A1.B2
       Df  Sum of Sq       RSS      AIC
<none>                  1.3000 -21.2737
A1      1     0.1200    1.4200 -22.0376
B1      1     1.2857    2.5857 -13.6468
B2      1     0.5891    1.8891 -18.0415
A1.B1   1     0.0514    1.3514 -22.7305
A1.B2   1 -1.332e-15    1.3000 -23.2737
```

```
> Anova(FM2,type="3")
Anova Table (Type III tests)
Response: Y
            Sum Sq Df   F value    Pr(>F)
(Intercept) 34.680  1  213.4154 4.729e-07 ***
A1           0.120  1    0.7385   0.41516
B1           1.286  1    7.9121   0.02274 *
B2           0.589  1    3.6252   0.09339 .
A1.B1        0.051  1    0.3165   0.58914
A1.B2        0.000  1 1.259e-32   1.00000
Residuals    1.300  8
```

Extra SS for single factors in the design matrix are determined by subtraction: SSE R - SSE F. This produces R's so-called "Type III" (or "Type 3") ANOVA *Marginal* Extra SS. Type III SS include "higher order" interaction terms such as A1.B1 & A1.B2 in FM even when a RM *excludes* a first order factor such as A1 or B1 or B2. As stated in documentation for Anova{car}, Type III ANOVA SS therefore are not marginal in an important sense.

```
> Anova(FM,type="III") #TYPE 3 ANOVA SS KNNL TREATMENT EFFECTS MODEL
Anova Table (Type III tests)
Response: Y
            Sum Sq Df  F value    Pr(>F)
(Intercept) 34.680  1 213.4154 4.729e-07 ***
A            0.120  1   0.7385  0.415160
B            4.190  2  12.8914  0.003145 **
A:B          0.075  2   0.2321  0.798034
Residuals    1.300  8
```

B= Extra SS value for Reduced Model excluding B1 & B2 from design matrix. A:B=Extra SS value for Reduced Model excluding both interaction terms for Reduced Model

Extra SS for factors represented by more than one column in the design matrix are similarly calculated by subtraction: SSE R - SSE F. They *may not* be calculated, in general, by addition of Extra SS for single columns in the design matrix, such as above.

Example Calculation of Reduced Models for Treatment Effects Model = contr.sum in R:

Reduced Model Excluding Interactions:

$X := \text{augment}\left(X_F^{\langle 0\rangle}, X_F^{\langle 1\rangle}, X_F^{\langle 2\rangle}, X_F^{\langle 3\rangle}\right)$

$$X = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & 0 \\ 1 & -1 & 0 & 1 \\ 1 & -1 & 0 & 1 \\ 1 & -1 & 0 & 1 \\ 1 & -1 & -1 & -1 \\ 1 & -1 & -1 & -1 \\ 1 & -1 & -1 & -1 \end{pmatrix}$$

Least Squares Estimation of the Regression Parameters:

$\beta_I := \left(X^T \cdot X\right)^{-1} \cdot X^T \cdot Y$

$$\beta_I = \begin{pmatrix} 1.6761905 \\ -0.0857143 \\ 0.4666667 \\ 0.3266667 \end{pmatrix} \quad \begin{pmatrix} \text{Intercept} \\ \beta A1 \\ \beta B1 \\ \beta B2 \end{pmatrix}$$

$p := \text{cols}(X) \qquad p = 4$

Fitted Values & Hat Matrix H:

$Y_h := X \cdot \beta_I$ **< fitted values Yh**

$H := X \cdot \left(X^T \cdot X\right)^{-1} \cdot X^T$ **< nXn Hat matrix**

Residuals:

$e := Y - Y_h$ **< residuals**

Reduced Model SSE for Interactions:

Sum of Squares:		Degrees of Freedom:	Mean Squares:	
$SSE := Y^T \cdot (I - H) \cdot Y$	$SSE = (1.3754)$	$df_E := N - p \quad df_E = 10$	$MSE := \frac{SSE}{df_E}$	$MSE = (0.1375)$

$SSE_{Ri} := SSE_0$ **< assigning Reduced Model SSE for Interactions**

$df_{Ri} := df_E$ **< assigning Reduced Model df_E for Interactions**

Reduced Model Excluding Factor A:

$X := \text{augment}\left(X_F^{\langle 0\rangle}, X_F^{\langle 2\rangle}, X_F^{\langle 3\rangle}, X_F^{\langle 4\rangle}, X_F^{\langle 5\rangle}\right)$

$$X = \begin{pmatrix} 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 \\ 1 & -1 & -1 & -1 & -1 \\ 1 & -1 & -1 & -1 & -1 \\ 1 & 1 & 0 & -1 & 0 \\ 1 & 0 & 1 & 0 & -1 \\ 1 & 0 & 1 & 0 & -1 \\ 1 & 0 & 1 & 0 & -1 \\ 1 & -1 & -1 & 1 & 1 \\ 1 & -1 & -1 & 1 & 1 \\ 1 & -1 & -1 & 1 & 1 \end{pmatrix}$$

Least Squares Estimation of the Regression Parameters:

$\beta_A := \left(X^T \cdot X\right)^{-1} \cdot X^T \cdot Y$

$$\beta_A = \begin{pmatrix} 1.6888889 \\ 0.4444444 \\ 0.3277778 \\ -0.0666667 \\ -0.0166667 \end{pmatrix} \quad \begin{pmatrix} \text{Intercept} \\ \beta B1 \\ \beta B2 \\ \beta A1B1 \\ \beta A1B2 \end{pmatrix}$$

$p := \text{cols}(X) \qquad p = 5$

Fitted Values & Hat Matrix H:

$Y_h := X \cdot \beta_A$ **< fitted values Yh**

$H := X \cdot \left(X^T \cdot X\right)^{-1} \cdot X^T$ **< nXn Hat matrix**

Residuals:

$e := Y - Y_h$ **< residuals**

Reduced Model SSE for Factor A:

Sum of Squares:		Degrees of Freedom:	Mean Squares:	
$SSE := Y^T \cdot (I - H) \cdot Y$	$SSE = (1.42)$	$df_E := N - p \quad df_E = 9$	$MSE := \frac{SSE}{df_E}$	$MSE = (0.1578)$

$SSE_{Ra} := SSE_0$ **< assigning Reduced Model SSE for Factor A**

$df_{Ra} := df_E$ **< assigning Reduced Model df_E for Factor A**

Reduced Model Excluding Factor B:

$X := \text{augment}\left(X_F^{\langle 0\rangle}, X_F^{\langle 1\rangle}, X_F^{\langle 4\rangle}, X_F^{\langle 5\rangle}\right)$

$$X = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 0 \\ 1 & -1 & 0 & -1 \\ 1 & -1 & 0 & -1 \\ 1 & -1 & 0 & -1 \\ 1 & -1 & 1 & 1 \\ 1 & -1 & 1 & 1 \\ 1 & -1 & 1 & 1 \end{pmatrix}$$

Least Squares Estimation of the Regression Parameters:

$\beta_B := \left(X^T \cdot X\right)^{-1} \cdot X^T \cdot Y$

$$\beta_B = \begin{pmatrix} 1.6285714 \\ 0.0190476 \\ 0.0666667 \\ -0.1933333 \end{pmatrix} \begin{pmatrix} \text{Intercept} \\ \beta A1 \\ \beta A1B1 \\ \beta A1B2 \end{pmatrix}$$

$p := \text{cols}(X) \qquad p = 4$

Fitted Values & Hat Matrix H:

$Y_h := X \cdot \beta_B$ **< fitted values Yh**

$H := X \cdot \left(X^T \cdot X\right)^{-1} \cdot X^T$ **< nXn Hat matrix**

Residuals:

$e := Y - Y_h$ **< residuals**

Reduced Model SSE for Factor B:

Sum of Squares:		Degrees of Freedom:	Mean Squares:	
$SSE := Y^T \cdot (I - H) \cdot Y$	$SSE = (5.4897)$	$df_E := N - p \quad df_E = 10$	$MSE := \dfrac{SSE}{df_E}$	$MSE = (0.549)$

$SSE_{Rb} := SSE_0$ **< assigning Reduced Model SSE for Factor B**

$df_{Rb} := df_E$ **< assigning Reduced Model df_E for Factor B**

GLM Tests for Effects:

Note: the GLM tests proceed in exactly the same way as for regressions. In addition, "extra" Sums of Squares may be determined by subtracting SSE for Full versus Reduced models. In unbalanced ANOVA, these Sums of Squares are typically not orthogonal.

Interactions:

Hypotheses:

H_0 : all interaction terms = 0

H_1 : at least one interaction term <> 0

GLM Test Statistic:

$$F := \frac{\dfrac{SSE_{Ri} - SSE_F}{df_{Ri} - df_F}}{\dfrac{SSE_F}{df_F}} \qquad F = 0.2321$$

Critical Value of the Test:

$\alpha := 0.05$ **< Probability of Type I error must be explicitly set**

$CV := qF\left(1 - \alpha, df_{Ri} - df_F, df_F\right) \qquad CV = 4.459$

Decision Rule:

IF F > CV, THEN REJECT H_0 OTHERWISE ACCEPT H_0

$F = 0.2321 \qquad CV = 4.459$

Probability Value:

$P := 1 - pF\left(F, df_{Ri} - df_F, df_F\right) \qquad P = 0.798$

Factor A Main Effect:

Hypotheses:

H_0 : all Factor A Effect terms = 0
H_1 : at least one Factor A term <> 0

GLM Test Statistic:

$$F := \frac{\frac{SSE_{Ra} - SSE_F}{df_{Ra} - df_F}}{\frac{SSE_F}{df_F}} \qquad F = 0.7385$$

Critical Value of the Test:

$\alpha := 0.05$ **< Probability of Type I error must be explicitly set**

$$CV := qF(1 - \alpha, df_{Ra} - df_F, df_F) \qquad CV = 5.3177$$

Decision Rule:

IF F > CV, THEN REJECT H_0 OTHERWISE ACCEPT H_0

$F = 0.7385$ $CV = 5.3177$

Probability Value:

$$P := 1 - pF(F, df_{Ra} - df_F, df_F) \qquad P = 0.4152$$

Factor B Main Effect:

Hypotheses:

H_0 : all Factor B Effect terms = 0
H_1 : at least one Factor B term <> 0

GLM Test Statistic:

$$F := \frac{\frac{SSE_{Rb} - SSE_F}{df_{Rb} - df_F}}{\frac{SSE_F}{df_F}} \qquad F = 12.8914$$

Critical Value of the Test:

$\alpha := 0.05$ **< Probability of Type I error must be explicitly set**

$$CV := qF(1 - \alpha, df_{Rb} - df_F, df_F) \qquad CV = 4.459$$

Decision Rule:

IF F > CV, THEN REJECT H_0 OTHERWISE ACCEPT H_0

$F = 12.8914$ $CV = 4.459$

Probability Value:

$$P := 1 - pF(F, df_{Rb} - df_F, df_F) \qquad P = 0.0031$$

Prototype in R:

```
#CONVERTING TO KNNL TREATMENT EFFECTS MODEL = contr.sum MODEL IN R
#FULL MODEL FROM contr.sum
contrasts(A)=contr.sum
contrasts(B)=contr.sum
FM=lm(Y~A*B)
summary(FM)
MM=data.frame(model.matrix(FM))
MM
anova(FM)  #TYPE 1 ANOVA SS KNNL TREATMENT EFFECTS MODEL
Anova(FM)  #TYPE 2 ANOVA SS KNNL TREATMENT EFFECTS MODEL
Anova(FM,type="III") #TYPE 3 ANOVA SS KNNL TREATMENT EFFECTS MODEL
```

```
> summary(FM)
Call:
lm(formula = Y ~ A * B)
Residuals:
       Min         1Q     Median         3Q        Max 
-6.000e-01 -2.000e-01 -8.816e-17  2.000e-01  4.000e-01 
Coefficients:
              Estimate Std. Error   t value Pr(>|t|)    
(Intercept)  1.700e+00  1.164e-01    14.609 4.73e-07 ***
A1          -1.000e-01  1.164e-01    -0.859   0.4152    
B1           5.000e-01  1.778e-01     2.813   0.0227 *  
B2           3.000e-01  1.576e-01     1.904   0.0934 .  
A1:B1       -1.000e-01  1.778e-01    -0.563   0.5891    
A1:B2       -1.768e-17  1.576e-01 -1.12e-16   1.0000    
```

```
> MM
   X.Intercept. A1 B1 B2 A1.B1 A1.B2
1             1  1  1  0     1     0
2             1  1  1  0     1     0
3             1  1  1  0     1     0
4             1  1  0  1     0     1
5             1  1  0  1     0     1
6             1  1 -1 -1    -1    -1
7             1  1 -1 -1    -1    -1
8             1 -1  1  0    -1     0
9             1 -1  0  1     0    -1
10            1 -1  0  1     0    -1
11            1 -1  0  1     0    -1
12            1 -1 -1 -1     1     1
13            1 -1 -1 -1     1     1
14            1 -1 -1 -1     1     1
```

$$\beta = \begin{pmatrix} 1.7 \\ -0.1 \\ 0.5 \\ 0.3 \\ -0.1 \\ 0 \end{pmatrix}$$

^ regression coefficients are treatments and interactions as described above

anova(FM) #TYPE 1 ANOVA SS KNNL TREATMENT EFFECTS MODEL
Anova(FM) #TYPE 2 ANOVA SS KNNL TREATMENT EFFECTS MODEL
Anova(FM,type="III") #TYPE 3 ANOVA SS KNNL TREATMENT EFFECTS MODEL

Note differences in Extra SS reported here!

Serial SS with Extra SS calculated in step-wise fashion from botton to top. >

```
> anova(FM)  #TYPE 1 ANOVA SS KNNL TREATMENT EFFECTS MODEL
Analysis of Variance Table
Response: Y
          Df Sum Sq Mean Sq F value   Pr(>F)   
A          1 0.0029  0.0029  0.0176 0.897785   
B          2 4.3960  2.1980 13.5262 0.002713 **
A:B        2 0.0754  0.0377  0.2321 0.798034   
Residuals  8 1.3000  0.1625                    
```

marginal SS excluding interaction terms in FM > if the base factors are deleted in RM

```
> Anova(FM)  #TYPE 2 ANOVA SS KNNL TREATMENT EFFECTS MODEL
Anova Table (Type II tests)
Response: Y
          Sum Sq Df F value   Pr(>F)
A         0.0926  1  0.5697 0.472022
B         4.3960  2 13.5262 0.002713 **
A:B       0.0754  2  0.2321 0.798034
Residuals 1.3000  8
```

incompletely marginal SS which include interaction > "higher order" terms in FM even if base factors are deleted from RM

```
> Anova(FM,type="III") #TYPE 3 ANOVA SS KNNL TREATMENT EFFECTS MODEL
Anova Table (Type III tests)
Response: Y
            Sum Sq Df  F value    Pr(>F)
(Intercept) 34.680  1 213.4154 4.729e-07 ***
A            0.120  1   0.7385  0.415160
B            4.190  2  12.8914  0.003145 **
A:B          0.075  2   0.2321  0.798034
Residuals    1.300  8
```

```
#GLM TESTS FOR KNNL TREATMENTS EFFECTS MODEL = contr.sum in R

attach(MM)
RMi=lm(Y~A1+B1+B2)
RMa=lm(Y~B1+B2+A1.B1+A1.B2)
RMb=lm(Y~A1+A1.B1+A1.B2)
anova(RMi,FM) #FOR INTERACTIONS KNNL TREATMENT EFFECTS MODEL
anova(RMa,FM) #FOR FACTOR A KNNL TREATMENT EFFECTS MODEL
anova(RMb,FM) #FOR FACTOR B KNNL TREATMENT EFFECTS MODEL = KNNL RESULTS
```

```
> anova(RMi,FM) #FOR INTERACTIONS KNNL TREATMENT EFFECTS MODEL
Analysis of Variance Table
Model 1: Y ~ A1 + B1 + B2
Model 2: Y ~ A * B
  Res.Df     RSS Df Sum of Sq      F Pr(>F)
1     10 1.37543
2      8 1.30000  2   0.07543 0.2321  0.798

> anova(RMa,FM) #FOR FACTOR A KNNL TREATMENT EFFECTS MODEL
Analysis of Variance Table
Model 1: Y ~ B1 + B2 + A1.B1 + A1.B2
Model 2: Y ~ A * B
  Res.Df  RSS Df Sum of Sq      F Pr(>F)
1      9 1.42
2      8 1.30  1      0.12 0.7385 0.4152

> anova(RMb,FM) #FOR FACTOR B KNNL TREATMENT EFFECTS MODEL = KNNL RESULTS
Analysis of Variance Table
Model 1: Y ~ A1 + A1.B1 + A1.B2
Model 2: Y ~ A * B
  Res.Df    RSS Df Sum of Sq      F   Pr(>F)
1     10 5.4897
2      8 1.3000  2    4.1897 12.891 0.003145 **
```

^ Extra SS for each are same as in Type III Anova(FM) above

```
#EXPLICIT STATEMENT OF FM TO ITEMIZE FACTORS
FM2=lm(Y~A1+B1+B2+A1.B1+A1.B2)
Anova(RMi,type="3")
anova(RMi,FM2)
Anova(RMa,FM2)
anova(RMa,FM2)
Anova(FM2,type="3")
anova(RMb,FM2)
detach(MM)
```

```
> Anova(RMi,type="3")
Anova Table (Type III tests)
Response: Y
            Sum Sq Df  F value    Pr(>F)
(Intercept) 38.855  1 282.4931 1.166e-08 ***
A1           0.093  1   0.6730   0.43112
B1           1.260  1   9.1608   0.01275 *
B2           0.745  1   5.4175   0.04223 *
Residuals    1.375 10
```

```
> anova(RMi,FM2)
Analysis of Variance Table
Model 1: Y ~ A1 + B1 + B2
Model 2: Y ~ A1 + B1 + B2 + A1.B1 + A1.B2
  Res.Df     RSS Df Sum of Sq      F Pr(>F)
1     10 1.37543
2      8 1.30000  2   0.07543 0.2321  0.798
```

```
> Anova(RMa,type="3")
Anova Table (Type III tests)
Response: Y
            Sum Sq Df  F value    Pr(>F)
(Intercept) 34.656  1 219.6507 1.252e-07 ***
B1           1.171  1   7.4201   0.02345 *
B2           0.734  1   4.6529   0.05935 .
A1.B1        0.024  1   0.1521   0.70559
A1.B2        0.002  1   0.0117   0.91623
Residuals    1.420  9
```

```
> anova(RMa,FM2)
Analysis of Variance Table
Model 1: Y ~ B1 + B2 + A1.B1 + A1.B2
Model 2: Y ~ A1 + B1 + B2 + A1.B1 + A1.B2
  Res.Df  RSS Df Sum of Sq      F Pr(>F)
1      9 1.42
2      8 1.30  1      0.12 0.7385 0.4152
```

```
> Anova(FM2,type="3")
Anova Table (Type III tests)
Response: Y
            Sum Sq Df   F value    Pr(>F)
(Intercept) 34.680  1  213.4154 4.729e-07 ***
A1           0.120  1    0.7385   0.41516
B1           1.286  1    7.9121   0.02274 *
B2           0.589  1    3.6252   0.09339 .
A1.B1        0.051  1    0.3165   0.58914
A1.B2        0.000  1 1.259e-32   1.00000
Residuals    1.300  8
```

```
> anova(RMb,FM2)
Analysis of Variance Table
Model 1: Y ~ A1 + A1.B1 + A1.B2
Model 2: Y ~ A1 + B1 + B2 + A1.B1 + A1.B2
  Res.Df    RSS Df Sum of Sq      F   Pr(>F)
1     10 5.4897
2      8 1.3000  2    4.1897 12.891 0.003145 **
```