BTRY 4030 - Fall 2018 - Homework 4 Q1

Greg Benton gwb67

Due Friday, November 9, 2018

Instructions:

You may either respond to the questions below by editing the hw4_2018_q1.Rmd to include your answers and compiling it into a PDF document, or by handwriting your answers and scanning these in.

You may discuss the homework problems and computing issues with other students in the class. However, you must write up your homework solution on your own. In particular, do not share your homework RMarkdown file with other students.

Question 1

We will start of with repeating the analysis from the 2017 midterm, with the idea of getting specific about the interpretation of a sequential ANOVA test.

We know that the sums of squares for each covariate is unchanged when the covariates are orthogonal. When they aren't, we need to ask "What is the null hypothesis for this test?"

We describe the test as being "the additional effect of $\mathbf{x}_j$ after controlling for X_{j-1}", but what does that mean, mathematically?

To do this, we'll break up the covariate matrix $X = [X_{j-1},\ \bar{X}_j]$ where $\bar{X}_j = [\mathbf{x}_j, \ldots, \mathbf{x}_p]$ and similarly, the coefficient vector will be broken into $\beta = (\beta_{j-1}^T, \bar{\beta}_j^T)^T$ so that we can write the linear regression model as

$$\mathbf{y} = X_{j-1}\beta_{j-1} + \bar{X}_j\bar{\beta}_j + \mathbf{e}.$$

We will **not** assume that X_j is orthogonal to $\bar{X}_j$. Note that this can be done for any choice of $j \in \{1, \ldots, p\}$.

a. Consider regressing $\mathbf{y}$ on only X_{j-1}. Give an expression for the estimated $\hat{\beta}_{j-1}$ and the corresponding fitted values.

This is standard linear regression with the design matrix being X_{j-1}. We know the form of the regression coefficients from this:

$$\hat{\beta}_{j-1} = (X_{j-1}^T X_{j-1})^{-1} X_{j-1}^T \mathbf{y}.$$

The corresponding fitted values are just $X_{j-1}\hat{\beta}_{j-1}$ which is

$$\hat{\mathbf{y}} = X_{j-1}(X_{j-1}^T X_{j-1})^{-1} X_{j-1}^T \mathbf{y} = H_{j-1}\mathbf{y}.$$

b. Show the fitted values (written in terms of true coefficients and errors) from the full regression can be re-written using the fitted values from Part 2a. and the matrix of residuals $\bar{R}_j$ obtained from regressing each column of $\bar{X}_j$ on X_{j-1} along with coefficients and errors.

First note that we can write $\mathbf{y}$ as

$$\mathbf{y} = H_{j-1}\mathbf{y} + (I - H_{j-1})\mathbf{y}$$

and then the fitted values from the full regression are

$$\hat{\mathbf{y}} = H\mathbf{y} = HH_{j-1}\mathbf{y} + (H - HH_{j-1})\mathbf{y}.$$

Note that H is a projection matrix that projects onto the column space of $X = [X_{j-1}\bar{X}_j]$, since H_{j-1} is the projection matrix onto the column space of X_{j-1}, its columns must also be in that span, thus $HH_{j-1} = H_{j-1}$. So we have

$$\hat{\mathbf{y}} = H_{j-1}\mathbf{y} + (H - H_{j-1})\mathbf{y} = \hat{\mathbf{y}}_{j-1} + (H - H_{j-1})\mathbf{y}$$

where $(H - \hat{H}_{j-1})\mathbf{y}_{j-1}$ are the fitted values from the model using just X_{j-1}. Now plugging in $\mathbf{y} = X_{j-1}\beta_{j-1} + \bar{X}_j\bar{\beta}_j + \epsilon$ in for the last $\mathbf{y}$ in the expression above we get

$$\hat{\mathbf{y}} = \hat{\mathbf{y}}_{j-1} + (H - H_{j-1})X_{j-1}\beta_{j-1} + (H - H_{j-1})\bar{X}_j\bar{\beta}_j + (H - H_{j-1})\epsilon.$$

Let's examine term by term:

$(H - H_{j-1})X_{j-1} = X_{j-1} - X_{j-1} = 0$ since X_{j-1} is in the column space of both X and X_{j-1}.

$(H - H_{j-1})\bar{X}_j = \bar{R}_j$ Since $H\bar{X}_j = \bar{X}_j$ and $H_{j-1}\bar{X}_j$ is the result of regressing the columns of $\bar{X}_j$ on X_{j-1}, we get the residuals of this regression.

$$\hat{\mathbf{y}} = \hat{\mathbf{y}}_{j-1} + \bar{R}_j\bar{\beta}_j + (H - H_{j-1})\epsilon.$$

c. Show that the sums of squares $\mathbf{y}^T(H - H_{j-1})\mathbf{y}$ for $\bar{X}_j|X_{j-1}$ is the same if you replace $\bar{X}_j$ with $\bar{R}_j$ (Why must $H\mathbf{y}$ be the same in both cases?)

Here we notice that $(H - H_{j-1})\bar{R}_j = (H - H_{j-1})(I - H_{j-1})\bar{X}_j = (I - H_{j-1})\bar{X}_j = \bar{R}_j$ and hence

$$(H - H_{j-1})(X_{j-1}\beta_{j-1} + \bar{X}_j\bar{\beta}_j + \mathbf{e}) = (H - H_{j-1})(X_{j-1}\beta_{j-1} + \bar{R}_j\bar{\beta}_j + \mathbf{e})$$

from which

$$\mathbf{y}^T(H - H_{j-1})\mathbf{y} = \mathbf{y}^T(H - H_{j-1})(H - H_{j-1})\mathbf{y}$$

is the same whether $\mathbf{y}$ is given in terms of $\bar{X}_j$ or $\bar{R}_j$.

As an aside, note that H is unchanged since we can write

$$[X_{j-1}, \bar{R}_j] = [X_{j-1}, \bar{X}_j - X_{j-1}(X_{j-1}^TX_{j-1})^{-1}X_{j-1}^T\bar{X}_j] = [X_{j-1}, \bar{X}_j]\begin{bmatrix} I & -(X_{j-1}^TX_{j-1})^{-1}X_{j-1}^T\bar{X}_j \\ 0 & I \end{bmatrix}$$

and we know that the H calculated with XA is the same as when calculated with X.

d. Within the sequential test for $\mathbf{x}_j|X_{j-1}$ show that the sum of squares $\mathbf{y}^T(H_j - H_{j-1})\mathbf{y}$ the the same whether you use the original X or $X^* = [X_{j-1}, (I - H_{j-1})\mathbf{x}_j, (I - H_j)\bar{X}_{j+1}]$.

If we write out

$$[X_{j-1}, (I - H_{j-1})\mathbf{x}_j] = [X_{j-1}, \mathbf{x}_j]\begin{bmatrix} I & -(X_{j-1}^TX_{j-1})^{-1}X_{j-1}^T\mathbf{x}_j \\ 0 & I \end{bmatrix}$$

We see that neither H_{j-1} nor H_j are affected by this transformation and therefore that $(H_j - H_{j-1})\mathbf{y}$ is also unaffected and so is

$$\mathbf{y}^T(H_j - H_{j-1})\mathbf{y} = \mathbf{y}^T(H_j - H_{j-1})(H_j - H_{j-1})\mathbf{y}$$

e. Show that, when using X^* (with corresponding coefficients β^*, the sum of squares $\mathbf{y}^T(H_j - H_{j-1})\mathbf{y}$ is only affected by the true value of β_j.

Continuing on from part e. we have

$$\begin{aligned}(H_j - H_{j-1})\mathbf{y} &= (H_j - H_{j-1})(X_{j-1}\beta_{j-1} + (I - H_{j-1})x_j\beta_j + (I - H_j)\bar{X}_{j+1}\bar{\beta}_j + \mathbf{e}) \\ &= (H_j - H_{j-1})(I - H_{j-1})\mathbf{x}_j\beta_j + (H_j - H_{j-1})(I - H_j)\bar{X}_{j+1}\bar{\beta}_j + (H_j - H_{j-1})\mathbf{e}) \\ &= (H_j - H_{j-1})\mathbf{x}_j\beta_j + (H_j - H_{j-1})\mathbf{e})\end{aligned}$$

because $(H_j - H_{j-1})(I - H_j) = 0$.

This quantity now only depends on β_j and no other coefficients.

f. Hence give a detailed interpretation of the meaning of rejecting the jth sequential test.

The null hypothesis for the jth sequential test (assuming $\mathbf{x}_j$ is a vector) is $H_0 : \beta_j = 0$. This is tested by examining what portion of the variance can be explained by the jth predictor. This can be seen in the form of the F statistic for the jth test, which is

$$\frac{SSE(\beta_j|\beta_{j-1})}{MSE} \sim F^1_{n-p-1},$$

which we see is a ratio of the drop in SSE when adding the jth predictor and the MSE for the full model.

Here we can see that we are only using the part of $\mathbf{x}_j$ that cannot be predicted by X_{j-1} and of $\bar{X}_{j+1}$ that cannot be predicted from X_j. We could therefore describe the test as saying that **There is a consistent signal in y that is associated with $\mathbf{x}_j$, including through correlations with X_{j+1} that cannot be modeled using only X_{j-1}.**

Question 2

Here we will illustrate the results from Question 1 with a real world data set. We will use a study of mortality in 55 US cities as it is influenced by pollutants NOX (nitrous oxide) and SO2 (sulfur dioxide), while controlling weather (PRECIP) and sociological variables (EDUC and NONWHITE). In this case we will be interested in the sequential test for EDUC with the covariates taken in the order in the data set.

You can find the data in `airpollution.csv` on CMS.

a. Create a new data set (referred to X^* below) in which NONWHITE, NOX and SO2 are replaced with the residuals after regressing each of them on PRECIP and EDUC.

```
air_pol = read.table("airpollution.csv", sep = ",", header = TRUE)
```

First, create a new data set

```
X_ast = air_pol
fit1 = lm(air_pol$NONWHITE ~ air_pol$PRECIP + air_pol$EDUC)
fit2 = lm(air_pol$NOX ~ air_pol$PRECIP + air_pol$EDUC)
fit3 = lm(air_pol$SO2 ~ air_pol$PRECIP + air_pol$EDUC)
X_ast$NONWHITE = fit1$resid
X_ast$NOX = fit2$resid
X_ast$SO2 = fit3$resid
```

b. Show that when producing a model to predict MORT with either the original covariates or the new covariates, you get the same predicted values (use the maximum absolute difference in predictions to show this).

```
fit1 = lm(air_pol$MORT ~ air_pol$PRECIP + air_pol$EDUC + air_pol$NONWHITE
+ air_pol$NOX + air_pol$SO2)

fit2 = lm(X_ast$MORT ~ X_ast$PRECIP + X_ast$EDUC + X_ast$NONWHITE
+ X_ast$NOX + X_ast$SO2)

max(abs(fit1$fitted.values - fit2$fitted.values))
```

```
## [1] 1.136868e-13
```

c. Add SO2 to MORT (this increases the coefficient of SO2 in the model by 1) and obtain a sequential ANOVA table (using the function `anova`) using the new response. Show that this changes the sum of squares for EDUC when using the original data.

```r
fit3 = lm(air_pol$MORT + air_pol$SO2 ~ air_pol$PRECIP + air_pol$EDUC
+ air_pol$NONWHITE + air_pol$NOX + air_pol$SO2)

anova(fit3)[2,2]
```

```
## [1] 85565.69
```

```r
anova(fit1)[2,2]
```

```
## [1] 16288.04
```

This changes the sum of squares.

d. Do the same thing using the new data set X^* and observe that the sum of squares for EDUC does not change.

```r
fit4 = lm(X_ast$MORT + X_ast$SO2 ~ X_ast$PRECIP + X_ast$EDUC
+ X_ast$NONWHITE + X_ast$NOX + X_ast$SO2)

anova(fit4)[2,2]
```

```
## [1] 16288.04
```

```r
anova(fit2)[2,2]
```

```
## [1] 16288.04
```

This does not

e. What happens if you add EDUC to MORT (ie, make its coefficient larger) instead? Are there differences between the two data sets? Why?

```r
fit5 = lm(air_pol$MORT + air_pol$EDUC ~ air_pol$PRECIP + air_pol$EDUC
+ air_pol$NONWHITE + air_pol$NOX + air_pol$SO2)
anova(fit5)
```

```
## Analysis of Variance Table
##
## Response: air_pol$MORT + air_pol$EDUC
##                  Df Sum Sq Mean Sq F value    Pr(>F)
## air_pol$PRECIP    1  13309   13309 10.6815 0.0019817 **
## air_pol$EDUC      1  14892   14892 11.9523 0.0011378 **
## air_pol$NONWHITE  1  44349   44349 35.5946 2.639e-07 ***
## air_pol$NOX       1  22243   22243 17.8521 0.0001035 ***
## air_pol$SO2       1    647     647  0.5191 0.4746491
## Residuals        49  61052    1246
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
fit4 = lm(X_ast$MORT + X_ast$EDUC ~ X_ast$PRECIP + X_ast$EDUC
+ X_ast$NONWHITE + X_ast$NOX + X_ast$SO2)
anova(fit4)
```

```
## Analysis of Variance Table
##
## Response: X_ast$MORT + X_ast$EDUC
##                Df Sum Sq Mean Sq F value    Pr(>F)
## X_ast$PRECIP    1  13309   13309 10.6815 0.0019817 **
## X_ast$EDUC      1  14892   14892 11.9523 0.0011378 **
## X_ast$NONWHITE  1  44349   44349 35.5946 2.639e-07 ***
```

```
## X_ast$NOX        1  22243   22243 17.8521 0.0001035 ***
## X_ast$SO2        1    647     647  0.5191 0.4746491
## Residuals       49  61052    1246
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

These are the same because in the sequential ANOVA we regress on EDUC before the other covariates that are associated with it; so unlike in part d, we don't ascribe any of the extra signal to covariates other than EDUC.

Question 3

Here we will turn to categorical covariates. In class, we saw that we may be interested in combinations of levels of a categorical covariate. We also saw that we can use a t-test to evaluate one combination; here we will examine multiple.

We can generally express a *matrix* of contrasts L. For example, in a one-way design with four levels A, B, C, and D, the matrix

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

provides the contrasts $(\mu_A - \mu_B, \mu_D - \mu_C)$ and we might be interested in testing that $L\mu = 0$ (i.e. $\mu_A = \mu_B$ AND $\mu_C = \mu_D$).

In this question, we will assume that the design is *balanced*, that is that there are an equal number, n of observations in each level.

a) In this question, we will use the *mean model* coding for our analysis (ie, set $\beta_0 = 0$ and use indicators for all four classes) so that $\beta = \mu$. Using this framework, what is the variance of $\hat{\beta}$? Express this in terms of numbers $n_A = \ldots = n_D = n$ of observations from each level.

Here we observe that $\mathbf{x}_i^T \mathbf{x}_j = 0$ *if* $i \neq j$ *and* $\mathbf{x}_i^T \mathbf{x}_i = n$ *so that*

$$\text{var}(\hat{\beta}) = \sigma^2 (X^T X)^{-1} = \frac{\sigma^2}{n} I_4$$

b) From the above result, what is the variance of $L\hat{\beta}$?

$$\text{var}(L\hat{\beta}) = L\text{var}(\hat{\beta})L^T = \frac{\sigma^2}{n} LL^T = \frac{2\sigma^2}{n} I_2$$

c) Using the expression above, show that $(L\hat{\beta})^T \text{var}(L\hat{\beta})^{-1} L\hat{\beta}$ has a χ^2_2 distribution.

Writing the rows of L *as* $\mathbf{l}_1$ *and* $\mathbf{l}_2$ *we observe that*

$$\begin{aligned}(L\hat{\beta})^T \text{var}(L\hat{\beta})^{-1} L\hat{\beta} &= \frac{\left(\mathbf{l}_1\hat{\beta}\right)^2}{2\sigma^2/n} + \frac{\left(\mathbf{l}_2\hat{\beta}\right)^2}{2\sigma^2/n} \\ &= \left(\frac{\mathbf{l}_1\hat{\beta}}{\text{sd}(\mathbf{l}_1\hat{\beta})}\right)^2 + \left(\frac{\mathbf{l}_2\hat{\beta}}{\text{sd}(\mathbf{l}_2\hat{\beta})}\right)^2 \\ &\sim \chi^2_2\end{aligned}$$

because $\mathbf{l}_1\hat{\beta}/sd(\mathbf{l}_1\hat{\beta}) \sim N(0,1)$ *and the two components are independent.*

Note that the argument is somewhat more general. If $\mathbf{y} \sim N(0, \Sigma)$ *and we take the eigen decomposition* $\Sigma = UDU^T$ *then* $\mathbf{y}^T \Sigma^{-1} \mathbf{y} = \mathbf{z}^T \mathbf{z}$ *where* $\mathbf{z} = D^{-1/2} U^T \mathbf{y}$ *and we observe that the variance of* $\mathbf{z}$ *is* $D^{-1/2} U^T U D U^T U D^{-1/2} = I$. *So* $\mathbf{z} \sim N(0, I)$ *and the result follows.*

d) Why is $(L\hat{\beta})^T \text{var}(L\hat{\beta})^{-1} L\hat{\beta}$ independent of MSE?

Because this expression is a function of $\hat{\beta}$ and MSE is a function of $\hat{\mathbf{e}}$ and we know from HW2 that $\hat{\beta}$ and $\hat{\mathbf{e}}$ are independent.

e) Hence, find an F statistic to test $H_0 : L\mu = 0$.

Here the natural thing to do is to replace σ^2 with MSE. Formally we have

$$\frac{\left(\mathbf{l}_1\hat{\beta}\right)^2 + \left(\mathbf{l}_2\hat{\beta}\right)^2}{2MSE/n}\frac{1}{2} = \frac{\left(\mathbf{l}_1\hat{\beta}\right)^2 + \left(\mathbf{l}_2\hat{\beta}\right)^2}{2\sigma^2/n}\frac{1}{2}\frac{\sigma^2}{MSE} \sim \frac{\chi^2_2/2}{\chi^2_{n-4}/(n-4)}$$

f) Many analyses start by assuming that L is orthonormal, $LL^T = I$. Why?

When X is coded in mean model form, the we have that $Var(\hat{\beta}) = cI$ and therefore $var(L\hat{\beta}) = cI$ if $LL^T = I$. This means that each contrast l_j can be tested independently.

Further, obsreving that if $LL^T = I$ then $\hat{\beta}^T L^T (L^T (X^T X)^{-1})^{-1} L\hat{\beta} = \hat{\beta}^T X^T X \hat{\beta}$, simplifying some calculations.

Question 4

Lets apply contrasts to real-world data. The file `NutritionStdy.csv` contains data on 314 patients undergoing elective surgery was collected to look at the relationship between the log-concentration of beta-carotene in the blood (`BetaPlasma`) and a number of personal characteristics and dietary factors. We will consider the following five variables observed in this study as predictors for a MLR regression model with response `log(BetaPlasma)`:

- **Quetelet** the Quetelet index (Weight/Height^2).
- **Vitamin** 1 = regular, 2 = Occasionally, 3 = No
- **NumSmoke** Daily number of cigarettes smoked
- **Fiber** Grams of fiber consumed per day
- **BetaDiet** Dietary beta-carotene consumed per day

Here we will test some hypotheses about vitamin intake.

Lets apply contrasts to real-world data. The file `NutritionStdy.csv` contains data on 314 patients undergoing elective surgery was collected to look at the relationship between the log-concentration of beta-carotene in the blood (`BetaPlasma`) and a number of personal characteristics and dietary factors. We will consider the following five variables observed in this study as predictors for a MLR regression model with response `log(BetaPlasma)`:

- **Quetelet** the Quetelet index (Weight/Height^2).
- **Vitamin** 1 = regular, 2 = Occasionally, 3 = No
- **NumSmoke** Daily number of cigarettes smoked
- **Fiber** Grams of fiber consumed per day
- **BetaDiet** Dietary beta-carotene consumed per day

Here we will test some hypotheses about vitamin intake.

a) Read the data in, making sure to specify that `Vitamin` is a factor and fit a linear model to predict `log(BetaPlasma)` from the remaining covariates. Extract the covariate matrix for this model using the `model.matrix` command in `R` with the output of the `lm` command as an argument.

```r
dat = read.csv("./NutritionStdy.csv", header=T)
dat$Vitamin = as.factor(dat$Vitamin)

mod = lm(log(BetaPlasma) ~ ., data=dat)
design_mat = model.matrix(mod) # covariate matrix
```

b) Write down a matrix of contrasts applied to this covariate matrix to test the hypothesis that i) Levels 1 and 2 of Vitamin are the same, and ii) that Level 3 is the same as the average of Levels 1 and 2. Produce this matrix in `R`; simply producing the `R` code with some explanation will suffice.

We have that β_0 is the coefficient for Vitamin level 1, β_2 is the coefficient for Vitamin level 2, and likewise for β_3. Note that testing $\mu_1 = \mu_2$ is equivalent to $\beta_0 + \beta_2 - \beta_0 = 0$, which is encoded below as `cntr1`. Testing $\mu_3 = \frac{\mu_1+\mu_2}{2}$ is equivalent to testing $\beta_0 + \beta_3 - \frac{1}{2}(\beta_0 + \beta_0 + \beta_2) = \beta_3 - \frac{1}{2}\beta_2 = 0$, which is encoded below as `cntr2`.

```
cntr1 = c(0, 0, 1, 0, 0, 0, 0) # testing if vitamin1 = vitamin2
cntr2 = c(0, 0, -0.5, 1, 0, 0, 0) # testing if 1/2(vitamin1 + vitamin2) = vitamin3

contrast_mat = rbind(cntr1, cntr2)
```

c) What are the estimated values of the two contrasts defined in the previous part? Verify these values against the coefficients supplied by `lm` when you use the additional argument `contrasts = list(Vitamin = "contr.helmert")`. Your answers should be 2 and 3 times the coefficients for `Vitamin1` and `Vitamin2` respectively. (**Bonus** explain why this is the case – you will need to look up Helmert contrasts.)

```
# generate the model
model = lm(log(BetaPlasma) ~ ., data=dat)
# now test the contrasts
contrast_mat %*% model$coefficients
```

```
##              [,1]
## cntr1 -0.05244799
## cntr2 -0.28716767
```

Seen above the tests for the contrasts are -0.05244799 and -0.28716767 respectively.

Now if we run the built in Helmert contrasts we get:

```
helmert_mod = lm(log(BetaPlasma) ~ ., data=dat, contrasts = list(Vitamin="contr.helmert"))
helmert_mod
```

```
##
## Call:
## lm(formula = log(BetaPlasma) ~ ., data = dat, contrasts = list(Vitamin = "contr.helmert"))
##
## Coefficients:
## (Intercept)     Quetelet     Vitamin1     Vitamin2     NumSmoke
##   5.567e+00   -3.376e-02   -2.622e-02   -9.572e-02   -5.117e-02
##       Fiber     BetaDiet
##   1.811e-02    4.907e-05
```

And we see that the generated coefficients are -0.002622 and -0.009572 which are one-half and one-third the values we got for the contrasts manually tested above, as desired.

d) Test this hypothesis using the formuae you derived in Question 3. Why does this not change the F statistic that you get from the `Anova` function?

```
L = contrast_mat
X = model.matrix(model)
L_beta = L %*% model$coefficients
inv_var = solve(L %*% solve(t(X) %*% X) %*% t(L))
numerator = t(L_beta) %*% inv_var %*% L_beta/2
denom = sum(model$residuals^2)/(nrow(X) - 3 - 1)
F_stat = numerator/denom
```

```r
print(F_stat)
```

```
##          [,1]
## [1,] 6.669368
```

Now testing the helmert model

```r
library(car)
```

```
## Warning: package 'car' was built under R version 3.5.1
```

```
## Loading required package: carData
```

```r
Anova(helmert_mod, type="III")
```

```
## Anova Table (Type III tests)
##
## Response: log(BetaPlasma)
##             Sum Sq  Df  F value    Pr(>F)
## (Intercept) 321.65   1 704.2991 < 2.2e-16 ***
## Quetelet     12.48   1  27.3371 3.167e-07 ***
## Vitamin       6.03   2   6.6048  0.001554 **
## NumSmoke      3.46   1   7.5711  0.006284 **
## Fiber         2.18   1   4.7760  0.029614 *
## BetaDiet      1.24   1   2.7257  0.099765 .
## Residuals   140.21 307
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

These manually calculated value and the F statistic read from the above table are not identical, but are within reason given there are a number of inversions of poorly conditioned matrices (generally numerically unstable) taking place above.