TMA4267 Linear Statistical models L27

Exam August 2011

Problem 1

Let $\underline{X} = \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix} \sim N_3(\mu, \Sigma)$

$\mu = \begin{bmatrix} 4 \\ -3 \\ 1 \end{bmatrix}$, $\Sigma = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & -\frac{3}{2} \\ 0 & -\frac{3}{2} & 5 \end{bmatrix}$

a) Find $f(x_1, x_2)$.

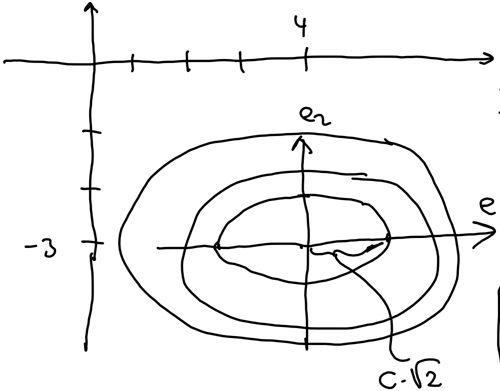
Subsets of mvN vectors are mvN

$(X_1, X_2) \sim N_2\left(\begin{bmatrix} 4 \\ -3 \end{bmatrix}, \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}\right)$

1

$\begin{bmatrix} 6 \\ 1 \end{bmatrix}$

Contours of this pdf: $(X-\mu)^T \Sigma^{-1} (X-\mu) = c^2$



$$\Sigma_{12} = \underbrace{\begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}}$$

$\lambda_1 = 2 \quad \lambda_2 =$

$e_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad e_2 =$

$$\left[\begin{matrix} \det(\Sigma_{12} - \lambda I) = 0 \\ \Sigma e = \lambda e \end{matrix} \right]$$

(see E1)

halflengths $\sqrt{\lambda_i} \cdot c$

$\sqrt{2}$ in x-direction

1 in y

$$CX = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -3 \\ 0 & 3 & 1 \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix}$$

Linear combinations of mvN variables are mvN.

$$CX \sim N_3$$

$$E(CX) = C\mu = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -3 \\ 0 & 3 & 1 \end{bmatrix} \begin{bmatrix} 4 \\ -3 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ -6 \\ -8 \end{bmatrix}$$

$$Cov(CX) = C \Sigma C^T = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -3 \\ 0 & 3 & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & -\frac{3}{2} \\ 0 & -\frac{3}{2} & 5 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 3 \\ 0 & -3 & 1 \end{bmatrix}$$

$$= \dots = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 55 & 0 \\ 0 & 0 & 5 \end{bmatrix}$$

AX and BX

$A \Sigma B^T = 0 \iff AX$ and BX independent

Problem 2: MLR (and DOE)

a) $Y = \beta_0 + \beta_1 X_1 + \dots + \beta_7 X_7 + \varepsilon$

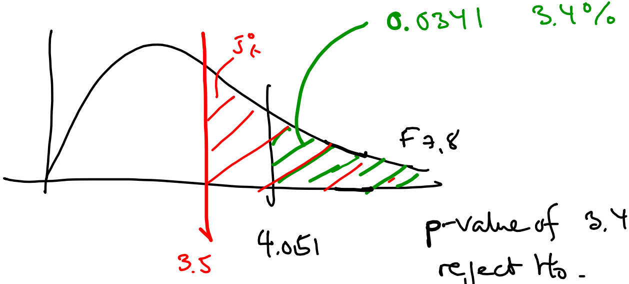
F-statistic:

$H_0: \beta_1 = \beta_2 = \dots = \beta_7 = 0$

vs

H_1: at least one $\beta_i \neq 0$ for $i = 1, .., 7$

$$F = \frac{\overbrace{SSR/(\overset{7}{p-1})}^{MSR}}{\underbrace{SSE/(\underset{8}{n-p})}_{MSE}} = 4{,}051$$



p-value of 3.4% $\Rightarrow$ reject H_0.

Problem 2:

start with explaining the output

X

Disperse Phase Volume x_1 (%)	Salinity x_2 (%)	Temperature x_3 ($^\circ$C)	Time delay x_4 (hours)	Surfactant concentration x_5 (%)	Span: Triton x_6	Solid Particles x_7 (%)	Voltage Y (kw/cm)
40	1	4	0.25	2	0.25	0.5	0.64
80	1	4	0.25	4	0.25	2	0.80
40	4	4	0.25	4	0.75	0.5	3.20
80	4	4	0.25	2	0.75	2	0.48
40	1	23	0.25	4	0.75	2	1.72
80	1	23	0.25	2	0.75	0.5	0.32
40	4	23	0.25	2	0.25	2	0.64
80	4	23	0.25	4	0.25	0.5	0.68
40	1	4	24	2	0.75	2	0.12
80	1	4	24	4	0.75	0.5	0.88
40	4	4	24	4	0.25	2	2.32
80	4	4	24	2	0.25	0.5	0.40
40	1	23	24	4	0.25	0.5	1.04
80	1	23	24	2	0.25	2	0.12
40	4	23	24	2	0.75	0.5	1.28
80	4	23	24	4	0.75	2	0.72

1
1
1

A regression analysis was performed and some output from the computer package R is given below:

element [i,i] in $(X^TX)^{-1}$ matrix

$\sqrt{(X^TX)^{-1}_{[i,i]}}\, s$

$\hat{\beta} = (X^TX)^{-1}X^TY$

$Cov(\hat{\beta}) = (X^TX)^{-1}\sigma^2$

```
> lmv=lm(y~x1+x2+x3+x4+x5+x6+x7)
> summary(lmv)

Call:
lm(formula = y ~ x1 + x2 + x3 + x4 + x5 + x6 + x7)
```

$\hat{\beta}_i$ $\quad \hat{se}(\hat{\beta}_i)$ $\quad \hat{\beta}_i / \hat{se}(\hat{\beta}_i)$

$H_0: \beta_i = 0$ vs $\beta_i \neq 0$

-0.79 0.79

```
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.591491   0.747886   0.791  0.45182
x1          -0.020500   0.006650  -3.083  0.01506 *
x2           0.170000   0.088671   1.917  0.09151 .
x3          -0.015263   0.014001  -1.090  0.30738
x4          -0.008421   0.011201  -0.752  0.47368
x5           0.460000   0.133006   3.458  0.00859 **
x6           0.520000   0.532024   0.977  0.35700
x7          -0.126667   0.177341  -0.714  0.49538
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

n-p

```
Residual standard error: 0.532 on 8 degrees of freedom
Multiple R-squared:  0.78, Adjusted R-squared: 0.5874
F-statistic: 4.051 on 7 and 8 DF,  p-value: 0.03401
```

$s = \sqrt{\frac{SSE}{n-p}}$

$n = 16,\ p = 8,\ n-p = 8$

Verify the values for R^2 and R^2_{adj}.

$$R^2 = \frac{SSR}{SST} = 1 - \frac{SSE}{SST}$$

$$\underset{\uparrow \text{ total}}{SST} = \sum_{i=1}^{n} (y_i - \bar{y})^2 = SSR + SSE$$

$$SSE = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$$

s in printout

$$S^2 = \frac{SSE}{n-p}, \quad \underline{SSE} = \underline{s^2 \cdot (n-p)} = 0.532^2 \cdot 8$$

$$= \underline{2.26}$$

$$F = \frac{SSR/7}{SSE/8} = 4.051$$

in printout

$$\underline{SSR} = F \cdot \frac{SSE}{8} \cdot 7 = 4.051 \cdot 2.26 \cdot \frac{7}{8} = \underline{8.0257}$$

$$R^2 = \frac{SSR}{\underset{SSR+SSE}{SST}} = \frac{8.0257}{8.0257 + 2.26} = \frac{8.03}{10.29} = \underline{\underline{0.78}}$$

$$R^2_{adj} = 1 - \frac{SSE / \overset{n-p}{8}}{SST / \underset{n-1}{15}} = 1 - \frac{2.26 / 8}{10.26 / 15} = \underline{\underline{0.5874}}$$

SS	df
SSR	$p-1$
SSE	$n-p$
SST	$n-1$

Backwards elimination: not in this years reading list.
We instead use R^2_{adj} (or AIC, BIC, Mallow's C_p) to select a model. That is, not hypothesis tests.

b)

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_5 X_5 + \varepsilon$$

$$\hat{Y} = 0.385 - 0.02 \cdot X_1 + 0.17 X_2 + 0.46 \cdot X_5$$

X_2 increase by one (all others kept constant):
y increase by 0.17.

Model A: all $\beta_1 - \beta_7$ in model $\rightarrow$ a) SSR_A, SSE_A (p)

Model B: only $\beta_1, \beta_2, \beta_5$ in model b) SSR_B, SSE_B (r)

$H_0: \beta_3 = \beta_4 = \beta_6 = \beta_7 = 0$ vs

H_1: at least one $\neq 0$

$$F = \frac{(SSR_A - SSR_B)/\underbrace{(8-4)}_{p-r}}{SSE_A / \underbrace{(16-8)}_{n-p}} \sim F_{4,8}$$

partial F test

$SSE_A = 2.26$ from a

$SSR_A = 8.0257$ —

SSR_B : same as in a, or small shortcut

$SSE_B = (0.5144)^2 \cdot 12 = 3.175$

$SST_B = SST_A = 10.29$

$SSR_B = SST_B - SSE_B = 10.29 - 3.175$

$= \underline{7.1147}$

$$F = \frac{(8.0257 - 7.1147)/4}{2.26/8} = 0.8047$$

$f_{0.05,\,4,8} = 3.84$

If $F_{obs} > 3.84 \Rightarrow$ reject H_0. Here don't reject.

c)

$$X_1 : \begin{bmatrix} 40 \\ 80 \end{bmatrix} \rightarrow \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

$\frac{1}{2}(40+80)$

$$X_{1\,new} = \frac{X_{1\,org} - 60}{20}$$

$\frac{1}{2}(80-40)$

A B C D	4
AB, AC, ...,	6
ABC, ABD, ACD, BCD	4
ABCD	1

Using simple maths : we find

$$\left.\begin{aligned} E &= ABC \\ F &= BCD \\ G &= ACD \end{aligned}\right\} \text{generators}$$

$$I = ABCE = BCDF = ACDG$$

defining relations

Resolutions: IV $\qquad 2^{7-3}_{IV}$

d)

Compare regression in c with b wrt x_1, factor A.

b): $\hat{\beta}_1 = -0.0205$

$\hat{A} = -0.82$

$$\hat{Y} = \hat{\beta}_0 + \overset{-0.0205}{\hat{\beta}_{1\,old}}\, \underbrace{X_{1\,old}}_{40,\,80} + \cdots$$

$\hat{A} = 2 \cdot \hat{\beta}_1$

$$\hat{Y} = \hat{\beta}_0 + \hat{\beta}_{1\,new}\, \underbrace{X_{1\,new}}$$

$$\frac{\hat{A}}{2} \qquad \underbrace{-\frac{0.82}{2}\left[\frac{X_{1\,old} - 60}{20}\right]}$$

See lecture 18 for this change of variables.

$$-\frac{0.82}{2}\,\frac{1}{20} = \underline{\underline{-0.0205}}$$

2b: $\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 x_{1avg} + \hat{\beta}_2 x_{2avg} + \hat{\beta}_5 x_{5avg}$

$\hat{\beta}_1 \downarrow A \qquad \hat{\beta}_2 \downarrow B \qquad \hat{\beta}_5 \downarrow E = ABC$

2d) 5 largest effects:
A
B
E
~~AB~~
~~AE~~

∴ A and E found to be significant

Two different models--- SSE will differ, and we get different results with respect to significance.

That was how far we got in this last lecture. We decided that the tentative solutions to Problem 3 be scanned and appended to the class notes - see next page.

Please read the remark in the end of the tentative solutions to Problem 3.

Problem 3: one-way ANOVA

BF2010 = our Bingham & Fry 2010 textbook (Ch 2.6)

$$Y_{ij} = \mu + \alpha_i + \varepsilon_{ij} \qquad i = 1, \dots, k$$

$j = 1, \dots, n \leftarrow$ equal group sizes

ε_{ij} i.i.d $N(0, \sigma^2)$

[BF2010: use r instead of k and $\mu + \alpha_i = \mu_i$]

$\bar{Y}_{\cdot\cdot} = \frac{1}{k \cdot n} \sum_i \sum_j Y_{ij}$ ← in BF2010 $Y_{\bullet\bullet} = \frac{1}{kn} \sum\sum Y_{ij}$

$\bar{Y}_{i\cdot} = \frac{1}{n} \sum_{j=1}^{n} Y_{ij}$ ($Y_{i\cdot}$ in BF2010)

Define two new matrices – useful for centering (= subtracting means)

$J_n = 1_{n \times n}$ ← $n \times n$ matrix with every entry equal to 1

and later J is J_{nk} ← $nk \times nk$

J^* is a block diagonal matrix with $\frac{1}{n} J_n$ on the diagonal, of k blocks (one for each treatment group).

$$J^* = \begin{bmatrix} \frac{1}{n} J_n & 0 & \cdots & 0 \\ 0 & \frac{1}{n} J_n & & 0 \\ \vdots & & \ddots & \\ 0 & 0 & & \frac{1}{n} J_n \end{bmatrix}_{nk \times nk}$$

Remark: we have worked with $\frac{1}{n} J_n$ in E4 Problem 4 and have there found that

$\frac{1}{n} J_n$ is symmetric and idempotent, with trace $\mathrm{tr}(\frac{1}{n} J_n) = 1$.

a) $\alpha_i = 0 \quad \forall i = 1, \ldots, k$

Then $Y_{ij} = \mu + \varepsilon_{ij}$ and all treatment means are equal.

An estimator for α_i:

$\hat{\alpha}_i = \bar{Y}_{i.} - \bar{Y}_{..}$ (the group mean minus the overall mean)

(since $\hat{\mu} = \bar{Y}_{..}$)

b) $A = I - J^*$ and $B = J^* - \frac{1}{nk} J$

Remark: A and B can be used to produce SSA and SSE in a one-way ANOVA (will do in c), and then it is important that

1) A and B are symmetric

2) A and B are idempotent $\Rightarrow$ projection matrices

3) A and B are orthogonal, $AB = 0$.

If this is true we may find quadratic forms Y^TAY and Y^TBY that $\sim \sigma^2 \chi^2_{df}$ (where $df = \text{rank}(A)$ or $\text{rank}(B)$), and we may form a F-distributed ratio

1) First we look at I, J^* and $\frac{1}{nk} J$: all $nk \times nk$ matrices.

Symmetric: all are obviously symmetric

Idempotent: $I \cdot I = I$ we should know.

J^* has $\frac{1}{n} J$ on blocks on the diagonal. From E4P4 we know $\frac{1}{n} J$ is idempotent, and thus:

$$J^* \cdot J^* = \begin{bmatrix} (\frac{1}{n}J)^2 & 0 & \cdots & 0 \\ 0 & (\frac{1}{n}J)^2 & & \\ \vdots & & \ddots & \\ 0 & & & (\frac{1}{n}J)^2 \end{bmatrix} = \begin{bmatrix} \frac{1}{n}J & 0 & \cdots & 0 \\ 0 & \frac{1}{n}J & & \\ & & \ddots & \\ 0 & & & \frac{1}{n}J \end{bmatrix} = J^*$$

$\frac{1}{nk}J$ is also idempotent, since $\frac{1}{n}J_n$ is idempotent

2) Now to A and B

$$A = I - J^*$$

symmetric: $(I - J^*)^T = I^T - J^{*T} = \underline{I - J^*}$

idempotent: $(I - J^*)(I - J^*) = I^2 - 2IJ^* + J^{*2}$

$$= I - 2J^* + J^* = \underline{I - J^*}$$

so, A is symmetric and idempotent

$$B = J^* - \frac{1}{nk}J$$

symmetric: $(J^* - \frac{1}{nk}J)^T = J^{*T} - \frac{1}{nk}J^T = \underline{J^* - \frac{1}{nk}J}$

Idempotent: $(J^* - \frac{1}{nk}J)(J^* - \frac{1}{nk}J)$

$$= J^{*2} - 2\underbrace{J^* \cdot \frac{1}{nk}J}_{\overset{\shortparallel}{\frac{1}{nk}J \text{ since:}}} + \left(\frac{1}{nk}J\right)^2 = J^* - 2\frac{1}{nk}J + \frac{1}{nk}J$$

$$= \underline{J^* - \frac{1}{nk}J}$$

$$J^* \cdot \frac{1}{nk}J = \begin{bmatrix} \frac{1}{n}J_n & 0 & \cdots & 0 \\ 0 & \frac{1}{n}J_n & & \\ \vdots & & \ddots & \\ 0 & & & \frac{1}{n}J_n \end{bmatrix} \frac{1}{k}\begin{bmatrix} \frac{1}{n}J_n & \frac{1}{n}J_n & \cdots & \frac{1}{n}J_n \\ \frac{1}{n}J_n & & & \\ \vdots & & & \\ \frac{1}{n}J_n & & & \frac{1}{n}J_n \end{bmatrix}$$

$$= \frac{1}{k}\begin{bmatrix} (\frac{1}{n}J_n)^2 & (\frac{1}{n}J_n)^2 & \cdots & (\frac{1}{n}J_n)^2 \\ (\frac{1}{n}J_n)^2 & \ddots & & \vdots \\ \vdots & & & \\ (\frac{1}{n}J_n)^2 & & & (\frac{1}{n}J_n)^2 \end{bmatrix} = \frac{1}{k}\begin{bmatrix} \frac{1}{n}J_n & \frac{1}{n}J_n & \cdots & \frac{1}{n}J_n \\ \vdots & & & \\ \frac{1}{n}J_n & & \cdots & \frac{1}{n}J_n \end{bmatrix} = \frac{1}{nk}J$$

so, B is symmetric and idempotent

Finally: $AB = 0$

$$(I - J^*)\left(J^* - \frac{1}{nk}J\right) = \underbrace{IJ^*}_{J^*} - \frac{1}{nk}J - \underbrace{J^{*2}}_{\substack{J^* \\ \text{(idemp.)}}} + \underbrace{\frac{1}{nk}J\,J^*}_{\frac{1}{nk}J \text{ (prev. page)}}$$

$$= J^* - J^* - \frac{1}{nk}J + \frac{1}{nk}J = \underline{\underline{0}}$$

c) Y is now a vector with n observations from each of the k treatments – in this order, so our model in a) may be written with matrices → as a standard MLR.

$$Y = \begin{bmatrix} Y_{11} \\ \vdots \\ Y_{1n} \\ Y_{21} \\ \vdots \\ Y_2 \\ \vdots \\ Y_{kn} \end{bmatrix} = \underbrace{\begin{bmatrix} \mu \\ \mu \\ \vdots \\ \\ \vdots \\ \\ \\ \mu \end{bmatrix} + \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_2 \\ \alpha \\ \alpha_k \end{bmatrix}}_{\mu^*} + \begin{bmatrix} \varepsilon_{11} \\ \vdots \\ \varepsilon_{1n} \\ \\ \\ \\ \\ \varepsilon_{kn} \end{bmatrix}$$

The aim here is to use the results about quadratic forms and projection matrices in one-way ANOVA.
In V2014 we use results about $\bar{X}$ and S^2 in a $X \sim N(\mu, \sigma^2)$ setting to find an F-statistic for testing $H_0: \alpha_1 = \alpha_2 = \dots = \alpha_k = 0$ vs H_1: at least one $\alpha_i \neq 0$, that is – we avoided projection matrices – which we took later in the course.

Theory that we need:

1) $Y \sim N_n(0, \sigma^2 I)$ and A symmetric and idempotent projection matrix with rank(A):

$$Y^T A Y \sim \sigma^2 \chi^2_{\text{rank}(A)}$$ Theorem 3.26 p 82 BF2010

2) AY and BY with $Y \sim N_n(0, \sigma^2 I)$ are independent iff $AB = 0$. Theorem 4.15, p 110, BF2010

3) $\text{rank}(A) = \text{tr}(A)$ for symmetric, idempotent matrices
e.g. E4P4, Prop 3.24 p 81 BF2010

We are told that $SSA = Y^T B Y$ and $SSE = Y^T A Y$ (but not asked to verify this), but how may we use this?

We should know that $H_0: \alpha_1 = \alpha_2 = \cdots = \alpha_n = 0$ can be tested by comparing MSA and MSE, where

$MSA = \frac{SSA}{df_1}$ with $SSA \sim \sigma^2 \chi^2_{df_1}$ and

$MSE = \frac{SSE}{df_2}$ with $SSE \sim \sigma^2 \chi^2_{df_2}$, so that

$$F_{obs} = \frac{MSA}{MSE} \sim F_{df_1, df_2}$$

and we reject H_0 when $F_0 > f_{\alpha, df_A, df_E}$.

[This should be known from part 2, chapter 2.6 of BF2010].

There are just some small steps left.

i) $df_2 = \text{rank}(A) = \text{tr}(A) = \text{tr}(\underset{nk \times nk}{I} - J^*) = nk - (nk \cdot \frac{1}{n})$

$= \underline{nk - k}$ for SSE

$df_1 = \text{rank}(B) = \text{tr}(B) = \text{tr}(J^* - \frac{1}{nk} J)$

$= \text{tr}(J^*) - \text{tr}(\frac{1}{nk} J) = \underline{k - 1}$ for SSA

ii) In 1) Y has mean 0, but we have "$\mu+\alpha_i$".

First for $SSE = Y^T A Y$. What if we instead look at $(Y-\mu^*)^T A (Y-\mu^*)$, where $(Y-\mu^*)$ has mean 0?

$$(Y-\mu^*)^T A (Y-\mu^*) = Y^T A Y - 2Y^T A\mu^* + \mu^{*T} A \mu^*$$

If $A\mu^* = 0$, this is ok.

$$A\mu^* = \underbrace{(I - J^*)}\begin{bmatrix} \mu+\alpha_1 \\ \vdots \\ \mu+\alpha_1 \\ \mu+\alpha_2 \\ \vdots \\ \mu+\alpha_k \end{bmatrix}$$

$$= \begin{bmatrix} \begin{matrix} 1-\frac{1}{n} & -\frac{1}{n} & -\frac{1}{n} & \cdots & -\frac{1}{n} \\ -\frac{1}{n} & 1-\frac{1}{n} & & & \\ & & \ddots & & \\ -\frac{1}{n} & & & & 1-\frac{1}{n} \end{matrix} & \begin{matrix} 0 & \cdots & 0 \end{matrix} & \\ 0 & \begin{matrix} 1-\frac{1}{n} & -\frac{1}{n} & \cdots & -\frac{1}{n} \\ & 1-\frac{1}{n} & & \\ & & \ddots & \\ -\frac{1}{n} & & & 1-\frac{1}{n} \end{matrix} & \\ & & \ddots & \\ & & & \boxed{} \end{bmatrix}$$

$1-\frac{1}{n}$ on diag
$-\frac{1}{n}$ off diag

row 1: $(\mu+\alpha_1)\,[\,1-\frac{1}{n}\underbrace{-\frac{1}{n}-\frac{1}{n}\cdots-\frac{1}{n}}_{n}\,] = 0$

row 2 — " —

row n — " —

row $n+1$ $(\mu+\alpha_2)\,[\,1-\frac{1}{n}-\frac{1}{n}\cdots-\frac{1}{n}\,] = 0$

row nk $(\mu+\alpha_k)\,[\,1-\frac{1}{n}-\frac{1}{n}\cdots-\frac{1}{n}\,] = 0$

[redacted]

$\Rightarrow B\mu^* = 0$. Then $(Y-\mu^*)^T A (Y-\mu^*) = Y^T B Y$

and $\underline{\underline{SSE = Y^T A Y \sim \sigma^2 \cdot \chi^2_{(nk-k)}}}$

Then for $SSA = Y^T B Y$ we also need to do the same, but here $B\mu^* = 0$ only when H_0 is true so that $\mu^* = \begin{bmatrix} \mu \\ \mu \\ \vdots \\ \mu \end{bmatrix}$.

$$(Y-\mu^*)^T B (Y-\mu^*) = Y^T B Y - 2 Y^T B \mu^* + \mu^T B \mu$$

$$B\mu^* = \left(J^* - \frac{1}{nk} J\right) \begin{bmatrix} \mu \\ \mu \\ \vdots \\ \mu \end{bmatrix} =$$

$$\begin{bmatrix} \frac{1}{n}J_n - \frac{1}{nk}J_n & -\frac{1}{nk}J_n & \cdots & -\frac{1}{nk}J_n \\ -\frac{1}{nk}J_n & \frac{1}{n}J_n - \frac{1}{nk}J_n & & \vdots \\ \vdots & & \ddots & \\ -\frac{1}{nk}J_n & \cdots & & \frac{1}{n}J_n - \frac{1}{nk}J_n \end{bmatrix} \begin{bmatrix} \mu \\ \hline \\ \hline \vdots \\ \hline \mu \end{bmatrix} \begin{matrix} \}k \\ \}k \\ \\ \\ \end{matrix}$$

to match the blocks

$-\frac{1}{nk}J_n$ off block diagonal

$\frac{1}{n}J_n - \frac{1}{nk}J_n$ on block diagonal

rows $1-k$: $\left(\frac{1}{n}\underbrace{J_n - \frac{1}{nk}J_n - \frac{1}{nk}J_n}_{0} \cdots\right) \cdot \begin{bmatrix} \mu \\ \mu \end{bmatrix} = 0$

and the same for the rest.

$B\mu^* = 0$ when H_0 is true, else not.

Then $(Y-\mu^*)^T B (Y-\mu^*) = Y^T B Y$ under H_0, and

$$SSA = Y^T B Y \sim \sigma^2 \chi^2_{(k-1)}$$

Finally, since $AB = 0$ then SSE and SSA are independent.

Thus, a test for H_0 is

$$F_0 = \frac{Y^T B Y / (k-1)}{Y^T A Y / (n \quad)} \sim F_{k-1, nk-k}$$

Reject H_0 when $f_0 > f_{\alpha, k-1, nk-k}$.

↑ observed value of F_0

Then we are done.

Do you now understand why we did not use projection matrices to do the one-way ANOVA?

This is NOT a course in idempotent matrices, but they may be useful in proving distributions of quadratic forms. However, there are — for one way ANOVA — other alternative proofs. See Lecture 7 for alternative proof for F_0 above in one-way ANOVA.

One final comment. If we look at one-way ANOVA using $Y = X\beta + \varepsilon$, with a dummy variable coding X of the treatments (see 4.2 in BF2010) then the MLR results (L15) says

$SSE = Y^T (I - H) Y$, so that $H = X(X^T X)^{-1} X^T = J^*$.

This is easy to verify e.g in R.