

TMA4267 – Linear Statistical Models

Solutions to Exercise 3 – V2015

February 12, 2015

Last change: A and B swapped for dfs in Problem 2.

Previous changes: 5c missing - in exp corrected, corrected typo in 5f 2.15 (had used 2.25) for constant c, and subsequent length of axis. added lf to problem 5, problem 3 and 6 has now also R script in this file (and separately in www).

Problem 1: Treatment of tennis elbow

Hypotheses:

Let μ_A , μ_B and μ_C be the expected pain-free grip force for each of the the three treatments.

$$H_0 : \mu_A = \mu_B = \mu_C \text{ vs. } H_1 : \text{at least one pair differs}$$

This hypothesis can be tested using one-way analysis of variance. We need to fill in the ANOVA table (SS, MS, df, F), which can be calculated from the summary statistics.

Let $\bar{x}_A$ denote the average and s_A the standard deviation of treatment A. Ditto for treatments B and C. Let $\bar{x}$ denote the grand mean.

$$\begin{aligned}SSA &= n_A(\bar{x}_A - \bar{x})^2 + n_B(\bar{x}_B - \bar{x})^2 + n_C(\bar{x}_C - \bar{x})^2 \\&= 63 \cdot (70.2 - 69.0)^2 + 65 \cdot (83.6 - 69.0)^2 + 60 \cdot (51.8 - 69.0)^2 \\&= 31697 \\SSE &= (n_A - 1)s_A^2 + (n_B - 1)s_B^2 + (n_C - 1)s_C^2 \\&= 62 \cdot 25.4^2 + 64 \cdot 22.9^2 + 59 \cdot 23.0^2 \\&= 104773\end{aligned}$$

Source	SS	df	MS	F
Treatment	31697	2	15849	28
Error	104773	185	566	
Total	136470	187		

The F statistic, here observed to be 28, should be compared with the critical value $f_{0.05,2,185} = 3.0$, and we thus reject the null hypothesis.

Assumptions:

The one-way ANOVA model is

$$Y_{ij} = \mu_i + \varepsilon_{ij}$$

where the error terms are independent and normally distributed with the same variance across treatment groups.

Conclusion:

There is reason to believe that the expected pain-free grip force is not the same for all the treatments.

Problem 2: Gene activity

The two-way ANOVA model can be written as

$$y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + \varepsilon_{ijk}$$

where α_i are the cell population levels, β_j are the gene levels and γ_{ij} the cell population and gene interaction. In the ANOVA output the sums of squares of the total variability in the data is decomposed into variability between the cell population (A), between the genes (B) and due to interaction between cell population and gene (AB), and the remaining variability is called the residual variability, $SS_{tot} = SSA + SSB + SSAB + SSE_{error}$. The sum of squares give these decomposed numerical values. A degree of freedom is associated with each sum, reflecting the amount of information in the sum - and technically associating the scaled sum with a χ^2 distribution with this number of degrees of freedom. The Mean squares are the Sum of squares divided by the degrees of freedom. The F-value is the ratio between the Mean Sq for the given factor (A, B or AB) and the Mean Sq for the residual. The p -value related to each of three tests based on the F-value and the F-distribution. The null hypotheses testes are wrt the parameters α_i begin equal (or the same for the other parameters).

Missing entries:

All Dfs: B: number of genes ($r=3$) -1=3-2=2, A: number of cell populations ($n=3$) -1=3-1=2, AB: $(r-1)(n-1)=4$, Error: $nr(m-1)$ where m is number of repeated observations, here $m=5$, so $3 \cdot 4=36$, Total: number of observation -1=45-1=44, or as the sum of the other dfs.

Sum Sq for AB: Mean Sq for AB $\cdot$ df for AB = $30.645 \cdot 4 = 122.98$.

Mean Sq for A: Sum Sq for A / df for A = $3.240/2 = 1.62$.

p -value for A: Tail in the Fisher distr with 2 and 36 df for $F=5.935$. Fom the table for 0.05 this critical value is around 3.3, which means that the p -value is below 0.05. For the 0.01 table the critical value is arond 5.4, which means that the p -value is below 0.01. Computer software (which is not available at the exam) would give 0.006 as the p -value.

The interaction term has a p -value below $2 \cdot 10^{-16}$ and is thus significant. We could proceed to look at differences between genes for specific cell populations or vice versa.

Problem 3: Simple matrix calculations

TMA4267: Exercise 3, Problem 3: Simple matrix calculations

#a construct A

A <- matrix(c(9,-2,-2,6),ncol=2)

A

#b symmetric?

t(A)

```

# yes  $t(A)=A$ 
t(A)==A

# c positive definite
#  $t(x) \% \% A \% \% x > 0$  for all  $x$ 
# just showing how this is calculated
x <- matrix(rnorm(2,0,1),ncol=1)
t(x) \% \% A \% \% x

# d
# but we may also use the fact that
# a symmetric postitive definite matrix
# has only positive eigenvalues

ev <- eigen(A)
names(ev)
ev$values
# yes, positive eigenvalues
# normalized eigenvectors?
ev$vectors

# first eigenvector, length
sum(ev$vectors[,1]^2)
# or
t(ev$vectors[,1]) \% \% ev$vectors[,1]
# second
t(ev$vectors[,2]) \% \% ev$vectors[,2]

# e spectral theorem

P <- ev$vectors
lambda <- diag(ev$values)

P \% \% lambda \% \% t(P)

# f inverse
Ainv <- solve(A)
Ainv
# or using the spectral decom
lambdainv <- diag(1/ev$values)
P \% \% lambdainv \% \% t(P)

# g, ups, used the fact that the eigenvalues of Ainv are
# the inverse of the eigenvalues of A already ...
eigen(Ainv)$values
diag(lambdainv)

# h since A is SPD it may ba a covariance matrix

```

```

#i correlation matrix
varvec <- diag(A)
invsdmat <- diag(1/sqrt(varvec))

corrmat <- invsdmat%%A%%invsdmat
corrmat
# builtin
cov2cor(A)

# j
# X has mean mu and covariance matrix A
mu <- matrix(c(3,1),ncol=1)
B <- matrix(c(1,1,1,2),ncol=2)
d <- matrix(c(1,2),ncol=1)

# E and Cov for s=BX
# mean is s
B%%mu
# cov(s) is B A B^T
B%%A%%t(B)

# E and Cov for t=t(d)X
# mean is
t(d)%%mu
# cov(t) is
t(d)%%A%%d

# E and Cov for v rbind X and 3X
# mean of 3X is 3mu
# cov of 3X is 9 covX
# mean
rbind(mu,3*mu)
# cov v is a matrix with four blocks
# block1 is cov of X
block1 <- A
# block 2 is cov of X and 3X=3 covX
block2 <- 3*A
# block 3 is block 2 transposed
block3 <- t(block2)
# block4 is Cov(3X)=9A
block4 <- 9*A

covv <- cbind(rbind(block1,block2),rbind(block3,block4))
covv

```

Problem 4: Covariance

Let $\mathbf{V}$ be a random vector with mean $\boldsymbol{\mu}$ and covariance matrix $E((\mathbf{V} - \boldsymbol{\mu})(\mathbf{V} - \boldsymbol{\mu})^T) = \boldsymbol{\Sigma}$. Show that $E(\mathbf{V}\mathbf{V}^T) = \boldsymbol{\Sigma} + \boldsymbol{\mu}\boldsymbol{\mu}^T$.

$$\begin{aligned}
\boldsymbol{\Sigma} &= \mathbb{E}((\mathbf{V} - \boldsymbol{\mu})(\mathbf{V} - \boldsymbol{\mu})^T) \\
&= \mathbb{E}(\mathbf{V}\mathbf{V}^T - \mathbf{V}\boldsymbol{\mu}^T - \boldsymbol{\mu}\mathbf{V}^T + \boldsymbol{\mu}\boldsymbol{\mu}^T) \\
&= \mathbb{E}(\mathbf{V}\mathbf{V}^T) - \mathbb{E}(\mathbf{V})\boldsymbol{\mu}^T - \boldsymbol{\mu}\mathbb{E}(\mathbf{V}^T) + \boldsymbol{\mu}\boldsymbol{\mu}^T \\
&= \mathbb{E}(\mathbf{V}\mathbf{V}^T) - \boldsymbol{\mu}\boldsymbol{\mu}^T - \boldsymbol{\mu}\boldsymbol{\mu}^T + \boldsymbol{\mu}\boldsymbol{\mu}^T \\
&= \mathbb{E}(\mathbf{V}\mathbf{V}^T) - \boldsymbol{\mu}\boldsymbol{\mu}^T \\
\mathbb{E}(\mathbf{V}\mathbf{V}^T) &= \boldsymbol{\Sigma} + \boldsymbol{\mu}\boldsymbol{\mu}^T
\end{aligned}$$

Problem 5: Symmetric projection matrices

Let the dimension of $\mathbf{A}$ be $n \times n$.

a) Prove that the eigenvalues of a projection matrix are 0 and 1.

$$\begin{aligned}
\mathbf{A}\mathbf{x} &= \lambda\mathbf{x} \\
\mathbf{A}^2\mathbf{x} &= \mathbf{A}\lambda\mathbf{x} = \lambda(\mathbf{A}\mathbf{x}) = \lambda^2\mathbf{x}
\end{aligned}$$

λ^2 is an eigenvalue of $\mathbf{A}^2$, but $\mathbf{A}^2 = \mathbf{A}$ so

$$\begin{aligned}
\mathbf{A}\mathbf{x} &= \mathbf{A}^2\mathbf{x} \\
\lambda\mathbf{x} &= \lambda^2\mathbf{x}
\end{aligned}$$

Since $\mathbf{x} \neq \mathbf{0}$

$$\begin{aligned}
\lambda &= \lambda^2 \\
\lambda(\lambda - 1) &= 0 \\
\lambda &= 0 \text{ or } \lambda = 1
\end{aligned}$$

b) The projection is orthogonal if and only if $\mathbf{0} = (\mathbf{A}\mathbf{y})^T(\mathbf{y} - \mathbf{A}\mathbf{y}) = \mathbf{y}^T\mathbf{A}^T(\mathbf{I} - \mathbf{A})\mathbf{y}$ for all $\mathbf{y}$, that is, if and only if $\mathbf{A}^T(\mathbf{I} - \mathbf{A}) = \mathbf{0}$. If $\mathbf{A}$ is symmetric, $\mathbf{A}^T(\mathbf{I} - \mathbf{A}) = \mathbf{A}(\mathbf{I} - \mathbf{A}) = \mathbf{A} - \mathbf{A}^2 = \mathbf{A} - \mathbf{A} = \mathbf{0}$. Conversely, if $\mathbf{A}^T(\mathbf{I} - \mathbf{A}) = \mathbf{0}$, $\mathbf{A}^T = \mathbf{A}^T\mathbf{A}$, so that $\mathbf{A} = (\mathbf{A}^T)^T = (\mathbf{A}^T\mathbf{A})^T = \mathbf{A}^T\mathbf{A} = \mathbf{A}^T$, showing that $\mathbf{A}$ is symmetric.

Note that $\mathbf{A}^T(\mathbf{I} - \mathbf{A}) = \mathbf{0}$ means not only that $\mathbf{A}\mathbf{y}$ and $\mathbf{y} - \mathbf{A}\mathbf{y}$ are orthogonal, but that any vector $\mathbf{A}\mathbf{y}$ in the column space (image, range) of $\mathbf{A}$ is orthogonal to any vector $(\mathbf{I} - \mathbf{A})\mathbf{z}$ in the null space (kernel) of $\mathbf{A}$ (which is the column space of $\mathbf{I} - \mathbf{A}$).

c) This should be relatively clear directly. We know that $\mathbf{A}$ is symmetric and idempotent, and we know that the rank of a symmetric matrix equals the number of non-zero eigenvalues. Then, since $\mathbf{A}$ has only $\lambda = 0$ and $\lambda = 1$ as eigenvalues and $\text{rank}(\mathbf{A}) = r$, then r eigenvalues must be 1 and the remaining $(n - r)$ must be 0.

d) What is the relationship between the trace and rank of a symmetric projection matrix?

$$\begin{aligned}\text{rank}(\mathbf{A}) &= r \\ \text{tr}(\mathbf{A}) &= \text{tr}(\mathbf{P}\mathbf{\Lambda}\mathbf{P}^T) = \text{tr}(\mathbf{P}^T\mathbf{P}\mathbf{\Lambda}) \\ &= \text{tr}(\mathbf{\Lambda}) = \sum_{i=1}^r \lambda_i = \sum_{i=1}^r 1 + \sum_{i=r+1}^n 0 = r\end{aligned}$$

So $\text{rank}(\mathbf{A}) = \text{tr}(\mathbf{A})$.

e)

$$\begin{aligned}\mathbf{H} &= \mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T \\ \mathbf{J} = \mathbf{1}\mathbf{1}^T &= \begin{bmatrix} 1 & 1 & \cdots & 1 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & 1 & \cdots & 1 \end{bmatrix}\end{aligned}$$

Show that each of the following matrices are symmetric and idempotent, and also find the rank (or trace) of the matrices.

$$\begin{aligned}
\mathbf{H} &= \mathbf{X}(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \\
\mathbf{H}^T &= (\mathbf{X}(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T)^T = \mathbf{X}(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T = \mathbf{H} \\
\mathbf{H}^2 &= (\mathbf{X}(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T) \mathbf{X}(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T = \mathbf{X}(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T = \mathbf{H} \\
\text{tr}(\mathbf{H}) &= \text{tr}(\mathbf{X}(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T) = \text{tr}(\mathbf{X}^T \mathbf{X}(\mathbf{X}^T \mathbf{X})^{-1}) = \text{tr}(\mathbf{I}_{p \times p}) = p \\
(\mathbf{I} - \mathbf{H})^T &= \mathbf{I} - \mathbf{H}^T = \mathbf{I} - \mathbf{H} \\
(\mathbf{I} - \mathbf{H})^2 &= (\mathbf{I} - \mathbf{H})(\mathbf{I} - \mathbf{H}) = \mathbf{I} - 2\mathbf{H} + \mathbf{H}^2 = \mathbf{I} - \mathbf{H} \\
\text{tr}(\mathbf{I} - \mathbf{H}) &= \text{tr}(\mathbf{I}_{n \times n}) - \text{tr}(\mathbf{H}) = n - p \\
\frac{1}{n} \mathbf{J}^T &= \frac{1}{n} \mathbf{J} \\
\left(\frac{1}{n} \mathbf{J}\right)^2 &= \frac{1}{n^2} \mathbf{J} \mathbf{J} = \frac{1}{n^2} \begin{bmatrix} 1 & 1 & \cdots & 1 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & 1 & \cdots & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & \cdots & 1 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & 1 & \cdots & 1 \end{bmatrix} \\
&= \frac{1}{n^2} \begin{bmatrix} n & n & \cdots & n \\ \vdots & \vdots & \vdots & \vdots \\ n & n & \cdots & n \end{bmatrix} = \frac{1}{n} \begin{bmatrix} 1 & 1 & \cdots & 1 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & 1 & \cdots & 1 \end{bmatrix} = \frac{1}{n} \mathbf{J} \\
\text{tr}\left(\frac{1}{n} \mathbf{J}\right) &= \frac{1}{n} \text{tr}(\mathbf{J}) = \frac{1}{n} n = 1 \\
\left(\mathbf{I} - \frac{1}{n} \mathbf{J}\right)^T &= \mathbf{I} - \frac{1}{n} \mathbf{J}^T = \left(\mathbf{I} - \frac{1}{n} \mathbf{J}\right) \\
\left(\mathbf{I} - \frac{1}{n} \mathbf{J}\right)^2 &= \mathbf{I} - 2\frac{1}{n} \mathbf{J} + \frac{1}{n^2} \mathbf{J}^2 = \mathbf{I} - 2\frac{1}{n} \mathbf{J} + \frac{1}{n} \mathbf{J} = \left(\mathbf{I} - \frac{1}{n} \mathbf{J}\right) \\
\text{tr}\left(\mathbf{I} - \frac{1}{n} \mathbf{J}\right) &= \text{tr}(\mathbf{I}) - \text{tr}\left(\frac{1}{n} \mathbf{J}\right) = n - 1 \\
\left(\mathbf{H} - \frac{1}{n} \mathbf{J}\right)^T &= \mathbf{H}^T - \frac{1}{n} \mathbf{J}^T = \left(\mathbf{H} - \frac{1}{n} \mathbf{J}\right) \\
\left(\mathbf{H} - \frac{1}{n} \mathbf{J}\right)^2 &= \mathbf{H}^2 - \frac{1}{n} \mathbf{H} \mathbf{J} - \frac{1}{n} \mathbf{J} \mathbf{H} + \frac{1}{n^2} \mathbf{J}^2
\end{aligned}$$

Since $\mathbf{H} \mathbf{X} = \mathbf{X}$ and the first column of $\mathbf{X}$ is a column of 1s, then $\mathbf{H} \mathbf{J} = \mathbf{J}$. We also need a result for $\mathbf{J} \mathbf{H}$, and $\mathbf{J} \mathbf{H} = ((\mathbf{J} \mathbf{H}^T)^T = (\mathbf{H}^T \mathbf{J}^T)^T = (\mathbf{H} \mathbf{J})^T = \mathbf{J}^T = \mathbf{J}$. Inserting this above gives.

$$\begin{aligned}
\left(\mathbf{H} - \frac{1}{n} \mathbf{J}\right)^2 &= \mathbf{H}^2 - \frac{1}{n} \mathbf{J} - \frac{1}{n} \mathbf{J} + \frac{1}{n} \mathbf{J} = \left(\mathbf{H} - \frac{1}{n} \mathbf{J}\right) \\
\text{tr}\left(\mathbf{H} - \frac{1}{n} \mathbf{J}\right) &= \text{tr}(\mathbf{H}) - \text{tr}\left(\frac{1}{n} \mathbf{J}\right) = p - 1
\end{aligned}$$