TMA4267 - Linear Statistical Models
Solutions to Exercise 1 - V2017

January 6, 2017

Problem 1: Simple matrix calculations

```
# 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 2: Mean and covariance of linear combinations

Here $\boldsymbol{X} = \begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix}$ is a trivariate random vector with mean $\boldsymbol{\mu} = \mathrm{E}(\boldsymbol{X}) = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ and covariance matrix $\boldsymbol{\Sigma} = \mathrm{Cov}(\boldsymbol{X}) = \boldsymbol{I} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, and $\boldsymbol{A} = \begin{pmatrix} \frac{2}{3} & -\frac{1}{3} & -\frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} & -\frac{1}{3} \\ -\frac{1}{3} & -\frac{1}{3} & \frac{2}{3} \end{pmatrix}$. Further, $\boldsymbol{Y} = \begin{pmatrix} Y_1 \\ Y_2 \\ Y_3 \end{pmatrix} = \boldsymbol{AX}$, and we are asked to find $\mathrm{E}(\boldsymbol{AX})$ and $\mathrm{Cov}(\boldsymbol{AX})$.

$$
\begin{aligned}
\mathrm{E}(\boldsymbol{Y}) = \boldsymbol{A\mu} &= \begin{pmatrix} \frac{2}{3} & -\frac{1}{3} & -\frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} & -\frac{1}{3} \\ -\frac{1}{3} & -\frac{1}{3} & \frac{2}{3} \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \\
\mathrm{Cov}(\boldsymbol{Y}) &= \boldsymbol{A}\,\mathrm{Cov}(\boldsymbol{X})\boldsymbol{A}^T = \boldsymbol{AIA}^T = \boldsymbol{AA}^T \\
&= \begin{pmatrix} \frac{2}{3} & -\frac{1}{3} & -\frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} & -\frac{1}{3} \\ -\frac{1}{3} & -\frac{1}{3} & \frac{2}{3} \end{pmatrix} \begin{pmatrix} \frac{2}{3} & -\frac{1}{3} & -\frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} & -\frac{1}{3} \\ -\frac{1}{3} & -\frac{1}{3} & \frac{2}{3} \end{pmatrix} \\
&= \begin{pmatrix} \frac{2}{3} & -\frac{1}{3} & -\frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} & -\frac{1}{3} \\ -\frac{1}{3} & -\frac{1}{3} & \frac{2}{3} \end{pmatrix} = \boldsymbol{A}
\end{aligned}
$$

Observe that $\boldsymbol{A}$ is idempotent.

Problem 3: Covariance

Starting with the given definition, and expanding, gives

$$
\begin{aligned}
\mathrm{Cov}(V, W) &= \mathrm{E}((V - \mu)(W - \eta)^T) \\
&= \mathrm{E}(VW^T - \mu W^T - V\eta^T + \mu\eta^T) \\
&= \mathrm{E}(VW^T) - \mu\,\mathrm{E}(W)^T - \mathrm{E}(V)\eta^T + \mu\eta^T \\
&= \mathrm{E}(VW^T) - \mu\eta^T - \mu\eta^T + \mu\eta^T \\
&= \mathrm{E}(VW^T) - \mu\eta^T,
\end{aligned}
$$

which is what we were asked to show.

Problem 4: The square root matrix and the Mahalanobis transform

a) Assume $\boldsymbol{\Sigma}$ is symmetric and positive definite, and $(\lambda_i, \boldsymbol{e}_i)$, $i = 1, ..., p$ are the eigenvalues and eigenvectors of $\boldsymbol{\Sigma}$.

A symmetric matrix has real eigenvalues. A positive definite matrix fulfills

$$\boldsymbol{x}^T \boldsymbol{\Sigma} \boldsymbol{x} > 0 \text{ for all } \boldsymbol{x} \neq 0$$

Let $\boldsymbol{x} = \boldsymbol{e_i}$ be the ith eigenvector of $\boldsymbol{\Sigma}$.

$$\begin{aligned}
\boldsymbol{e}_i^T \boldsymbol{\Sigma} \boldsymbol{e}_i &= \boldsymbol{e}_i^T \boldsymbol{P} \boldsymbol{\Lambda} \boldsymbol{P}^T \boldsymbol{e}_i \\
&= \boldsymbol{e}_i^T (\sum_{i=1}^{p} \lambda_i \boldsymbol{e}_i \boldsymbol{e}_i^T) \boldsymbol{e}_i \\
&= \boldsymbol{e}_i^T (\lambda_i \boldsymbol{e}_i \boldsymbol{e}_i^T) \boldsymbol{e}_i \\
&= \lambda_i > 0
\end{aligned}$$

so all eigenvalue of a SPD must be positive.

What about the eigenvalues and eigenvectors of the inverse matrix of $\boldsymbol{\Sigma}$?

First, $\boldsymbol{\Sigma}^{-1}$ exists since $\boldsymbol{\Sigma}$ is SPD. Next, consider the eigenvalue-eigenvector pair $(\lambda_i, \boldsymbol{e}_i)$ of $\boldsymbol{\Sigma}$

$$\begin{aligned}
\boldsymbol{\Sigma} \boldsymbol{e}_i &= \lambda \boldsymbol{e}_i \\
\boldsymbol{\Sigma}^{-1} \boldsymbol{\Sigma} \boldsymbol{e}_i &= \lambda \boldsymbol{\Sigma}^{-1} \boldsymbol{e}_i \\
\boldsymbol{I} \boldsymbol{e}_i &= \lambda \boldsymbol{\Sigma}^{-1} \boldsymbol{e}_i \\
\frac{1}{\lambda_i} \boldsymbol{e}_i &= \boldsymbol{\Sigma}^{-1} \boldsymbol{e}_i,
\end{aligned}$$

meaning $(1/\lambda_i, \boldsymbol{e}_i)$ is an eigenvalue-eigenvector pair of $\boldsymbol{\Sigma}^{-1}$.

So, $\boldsymbol{\Sigma}$ and $\boldsymbol{\Sigma}^{-1}$ have the same eigenvectors. And, if λ_i is an eigenvalue for $\boldsymbol{\Sigma}$ then $1/\lambda_i$ is an eigenvalue for $\boldsymbol{\Sigma}^{-1}$. Hence, all eigenvalues are positive.

b) Show that $\boldsymbol{\Sigma}^{\frac{1}{2}}$ and $\boldsymbol{\Sigma}^{-\frac{1}{2}}$ are symmetric:

$$\begin{aligned}
\boldsymbol{\Sigma}^{\frac{1}{2}} &= \boldsymbol{P} \boldsymbol{\Lambda}^{\frac{1}{2}} \boldsymbol{P}^T \\
(\boldsymbol{\Sigma}^{\frac{1}{2}})^T &= (\boldsymbol{P} \boldsymbol{\Lambda}^{\frac{1}{2}} \boldsymbol{P}^T)^T = \boldsymbol{P} \boldsymbol{\Lambda}^{\frac{1}{2}} \boldsymbol{P}^T = \boldsymbol{\Sigma}^{\frac{1}{2}}
\end{aligned}$$

Since transposing a diagonal matrix leaves the matrix unchanges. To prove that $\boldsymbol{\Sigma}^{-\frac{1}{2}}$ is symmetric, just replace $\boldsymbol{\Sigma}^{\frac{1}{2}}$ by $\boldsymbol{\Sigma}^{-\frac{1}{2}}$ in the above equations.

We show the three given identities as follows:

$$\begin{aligned}
\boldsymbol{\Sigma}^{\frac{1}{2}}\boldsymbol{\Sigma}^{\frac{1}{2}} &= \boldsymbol{P}\boldsymbol{\Lambda}^{\frac{1}{2}}\boldsymbol{P}^T\boldsymbol{P}\boldsymbol{\Lambda}^{\frac{1}{2}}\boldsymbol{P}^T \\
&= \boldsymbol{P}\boldsymbol{\Lambda}^{\frac{1}{2}}\boldsymbol{\Lambda}^{\frac{1}{2}}\boldsymbol{P}^T \\
&= \boldsymbol{P}\boldsymbol{\Lambda}\boldsymbol{P}^T = \boldsymbol{\Sigma} \\
\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{\Sigma}^{-\frac{1}{2}} &= \boldsymbol{P}\boldsymbol{\Lambda}^{-\frac{1}{2}}\boldsymbol{P}^T\boldsymbol{P}\boldsymbol{\Lambda}^{-\frac{1}{2}}\boldsymbol{P}^T \\
&= \boldsymbol{P}\boldsymbol{\Lambda}^{-\frac{1}{2}}\boldsymbol{\Lambda}^{-\frac{1}{2}}\boldsymbol{P}^T \\
&= \boldsymbol{P}\boldsymbol{\Lambda}^{-1}\boldsymbol{P}^T = \boldsymbol{\Sigma}^{-1} \\
\boldsymbol{\Sigma}^{\frac{1}{2}}\boldsymbol{\Sigma}^{-\frac{1}{2}} &= \boldsymbol{P}\boldsymbol{\Lambda}^{\frac{1}{2}}\boldsymbol{P}^T\boldsymbol{P}\boldsymbol{\Lambda}^{-\frac{1}{2}}\boldsymbol{P}^T \\
&= \boldsymbol{P}\boldsymbol{\Lambda}^{\frac{1}{2}}\boldsymbol{\Lambda}^{-\frac{1}{2}}\boldsymbol{P}^T \\
&= \boldsymbol{P}\boldsymbol{I}\boldsymbol{P}^T = \boldsymbol{I}
\end{aligned}$$

where $\boldsymbol{I}$ is the identity matrix.

c)

$$\boldsymbol{Y} = \boldsymbol{\Sigma}^{-\frac{1}{2}}(\boldsymbol{X} - \boldsymbol{\mu})$$

$$\begin{aligned}
\mathrm{E}(Y) &= \mathrm{E}(\boldsymbol{\Sigma}^{-\frac{1}{2}}(\boldsymbol{X} - \boldsymbol{\mu})) = \boldsymbol{\Sigma}^{-\frac{1}{2}}(\mathrm{E}(\boldsymbol{X}) - \boldsymbol{\mu}) = \boldsymbol{0} \\
\mathrm{Cov}(Y) &= \mathrm{Cov}(\boldsymbol{\Sigma}^{-\frac{1}{2}}(\boldsymbol{X} - \boldsymbol{\mu})) = \boldsymbol{\Sigma}^{-\frac{1}{2}}\,\mathrm{Cov}(\boldsymbol{X})(\boldsymbol{\Sigma}^{-\frac{1}{2}})^T \\
&= \boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{\Sigma}\boldsymbol{\Sigma}^{-\frac{1}{2}} = \boldsymbol{P}\boldsymbol{\Lambda}^{-\frac{1}{2}}\boldsymbol{P}^T\boldsymbol{P}\boldsymbol{\Lambda}\boldsymbol{P}^T\boldsymbol{P}\boldsymbol{\Lambda}^{\frac{1}{2}}\boldsymbol{P}^T \\
&= \boldsymbol{P}\boldsymbol{\Lambda}^{-\frac{1}{2}}\boldsymbol{\Lambda}\boldsymbol{\Lambda}^{-\frac{1}{2}}\boldsymbol{P}^T \\
&= \boldsymbol{P}\boldsymbol{I}\boldsymbol{P}^T = \boldsymbol{I}
\end{aligned}$$

Problem 5: The normal and chi-square distribution

a) $U \sim N(0,1)$. Find pdf and MGF of $X = U^2$.
Denote by ϕ the pdf of the standard Normal distribution.
Let $X = U^2$ and $U = \sqrt{X}$.

$$\begin{aligned}
F_X(x) &= P(U^2 \le x) = P(-\sqrt{x} \le U \le \sqrt{x}) = F_U(\sqrt{x}) - F_U(-\sqrt{x}) \\
f_X &= \frac{d}{dx}F_X(x) = f_U(\sqrt{x})\frac{d}{dx}\sqrt{x} - f_U(-\sqrt{x})\frac{d}{dx}(-\sqrt{x}) \\
&= f_U(\sqrt{x})\frac{1}{2\sqrt{x}} + f_U(-\sqrt{x})\frac{1}{2\sqrt{x}} \\
&= \frac{1}{\sqrt{2\pi}}e^{-x/2}\frac{1}{2\sqrt{x}} + \frac{1}{\sqrt{2\pi}}e^{-x/2}\frac{1}{2\sqrt{x}} \\
&= \frac{1}{\sqrt{2\pi}}e^{-x/2}x^{-1/2} \\
&= \frac{1}{\sqrt{2}\Gamma(1/2)}e^{-x/2}x^{1/2-1}
\end{aligned}$$

MGF:

$$
\begin{aligned}
M_{U^2}(t) &= \int_{-\infty}^{\infty} e^{tu^2}\phi(u)du = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{tu^2}e^{-u^2/2}du \\
&= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-u^2(1-2t)/2}du \text{ using } u = v(1-2t), du = (1-2t)dv \\
&= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-v^2/2}\frac{1}{\sqrt{1-2t}}dv \\
&= \frac{1}{\sqrt{1-2t}} \text{ for } t < \frac{1}{2}
\end{aligned}
$$

b) $V \sim \chi^2_p$.

First we use the result from a) to find the MGF for the χ^2_p. Since V can be formed by a sum of p independent χ^2_1 variables, then the MGF of V is the product of the MGF of p χ^2_1 variables.

$$M_V(t) = [M_U^2(t)]^p = \frac{1}{(1-2t)^{p/2}}$$

Then we find the MGF of V directly from $f_V(v)$.

$$
\begin{aligned}
M_V(t) &= \int_{-\infty}^{\infty} e^{tv}\frac{1}{\Gamma(p/2)2^{p/2}}v^{(p/2)-1}e^{-v/2}dv \\
&= \frac{1}{\Gamma(p/2)2^{p/2}}\int_{-\infty}^{\infty} e^{-v/2(1-2t)}v^{p/2-1}dv, \text{ let } u = v(1-2t), du = (1-2t)dv \\
&= \frac{1}{\Gamma(p/2)2^{p/2}}\int_{-\infty}^{\infty} e^{-u/2}\frac{u^{p/2-1}}{(1-2t)^{p/2-1}}\frac{du}{(1-2t)} \\
&= \frac{1}{(1-2t)^{p/2}}\frac{1}{\Gamma(p/2)2^{p/2}}\int_{-\infty}^{\infty} e^{-u/2}u^{p/2-1}du \\
&= \frac{1}{(1-2t)^{p/2}}
\end{aligned}
$$

(The last integral equals 1 since the integrand is the χ^2-distribution.) We see that the two calculations of $M_V(t)$ are equal, and thus conclude that the given $f_V(v)$ is for the χ^2_p-distribution.

Problem 6: N and Chi-square by simulation - in R

```
B <- 10000
n <- 10

# a
rnorm(B,0,1) # draw B standard normal variates
dchisq(1,1) # density at x=1 for chi-square df=1
pt(0,n-1) # cdf at x=0 for t-distr with df=n-1
qf(0.05,1,2) #critical value with area 0.05 to the left
qf(0.05,1,2,lower.tail=FALSE) # critical value with area 0.05 to the right
qf(0.95,1,2) # same as above
```

```
# b
?curve
# how far out? 4 sds ok?
curve(dnorm,-4,4,type="l")
abline(v=qnorm(0.05),col=2)
abline(v=qnorm(0.95),col=2)
# for the fun of it, adding shades to tails
tt <- seq(from = -4, to=qnorm(0.05), length = 50)
dtt <- dnorm(tt)
polygon(x = c(-4, tt, qnorm(0.05)), y = c(0, dtt, 0), col = "gray")
tt <- seq(from = qnorm(0.95), to=4, length = 50)
dtt <- dnorm(tt)
polygon(x = c(qnorm(0.95),tt,4), y = c(0, dtt, 0), col = "gray")

# c
x <- rnorm(B,0,1)
y <- x^2
range(y)
hist(y,nclass=100,prob=TRUE)
dchisq1 <- function(x) return(dchisq(x,df=1))
curve(dchisq1,min(y),max(y),add=TRUE,col=2)
# curve only takes a function with ONE argument, needed to make a df=1 version of dchisq
abline(v=qchisq(0.1,1),col=3)
abline(v=qchisq(0.9,1),col=3)
```

Problem 7: Symmetric idempotent matrices

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

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

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

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

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

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

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

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

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

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

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

d)

$$\boldsymbol{J} = \boldsymbol{1}\boldsymbol{1}^T = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & 1 & \cdots & 1 \end{bmatrix}$$

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

$$\begin{aligned}
\frac{1}{n}\boldsymbol{J}^T &= \frac{1}{n}\boldsymbol{J} \\
(\frac{1}{n}\boldsymbol{J})^2 &= \frac{1}{n^2}\boldsymbol{JJ} = \frac{1}{n^2}\begin{bmatrix} 1 & 1 & \cdots & 1 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & 1 & \cdots & 1 \end{bmatrix}\begin{bmatrix} 1 & 1 & \cdots & 1 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & 1 & \cdots & 1 \end{bmatrix} \\
&= \frac{1}{n^2}\begin{bmatrix} n & n & \cdots & n \\ \vdots & \vdots & \vdots & \vdots \\ n & n & \cdots & n \end{bmatrix} = \frac{1}{n}\begin{bmatrix} 1 & 1 & \cdots & 1 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & 1 & \cdots & 1 \end{bmatrix} = \frac{1}{n}\boldsymbol{J} \\
\text{tr}(\frac{1}{n}\boldsymbol{J}) &= \frac{1}{n}\,\text{tr}(\boldsymbol{J}) = \frac{1}{n}n = 1 \\
(\boldsymbol{I} - \frac{1}{n}\boldsymbol{J})^T &= \boldsymbol{I} - \frac{1}{n}\boldsymbol{J}^T = (\boldsymbol{I} - \frac{1}{n}\boldsymbol{J}) \\
(\boldsymbol{I} - \frac{1}{n}\boldsymbol{J})^2 &= \boldsymbol{I} - 2\frac{1}{n}\boldsymbol{J} + \frac{1}{n^2}\boldsymbol{J}^2 = \boldsymbol{I} - 2\frac{1}{n}\boldsymbol{J} + \frac{1}{n}\boldsymbol{J} = (\boldsymbol{I} - \frac{1}{n}\boldsymbol{J}) \\
\text{tr}(\boldsymbol{I} - \frac{1}{n}\boldsymbol{J}) &= \text{tr}(\boldsymbol{I}) - \text{tr}(\frac{1}{n}\boldsymbol{J}) = n - 1
\end{aligned}$$