STAT 3690 Homework 1

zhiyanggeezhou.github.io

Zhiyang Zhou (zhiyang.zhou@umanitoba.ca)

Due at Feb 28 11:59 am (Central Time)

Answers must be submitted electronically via Crowdmark. Please enclose your R source code (if applicable) as well.

1. The function $\mathrm{cov}(\cdot,\cdot)$ is bilinear, i.e., for random vectors $\mathbf{W}$, $\mathbf{X}$, $\mathbf{Y}$ and $\mathbf{Z}$ and fixed matrices $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ and $\mathbf{D}$, one has $\mathrm{cov}(\mathbf{AW}+\mathbf{BX},\mathbf{Y}) = \mathbf{A}\boldsymbol{\Sigma}_{\mathbf{WY}}+\mathbf{B}\boldsymbol{\Sigma}_{\mathbf{XY}}$ and $\mathrm{cov}(\mathbf{W},\mathbf{CY}+\mathbf{DZ}) = \boldsymbol{\Sigma}_{\mathbf{WY}}\mathbf{C}^\top+\boldsymbol{\Sigma}_{\mathbf{WZ}}\mathbf{D}^\top$, where $\mathbf{AW}+\mathbf{BX}$ and $\mathbf{CY}+\mathbf{DZ}$ both make sense.
   a. Prove this bilinearity.
   b. Rephrase $\mathrm{cov}(\mathbf{AW}+\mathbf{BX},\mathbf{CY}+\mathbf{DZ})$ in terms of matrices $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, $\mathbf{D}$, $\boldsymbol{\Sigma}_{\mathbf{WY}}$, $\boldsymbol{\Sigma}_{\mathbf{WZ}}$, $\boldsymbol{\Sigma}_{\mathbf{XY}}$ and $\boldsymbol{\Sigma}_{\mathbf{XZ}}$.

1a. By definitions, let $\mu_{AW} = E(AW)$ & $\mu_{BX} = E(BX)$, then

$$
\begin{aligned}
\mathrm{cov}(AW+BX, Y) &= E\{(AW+BX-\mu_{AW}-\mu_{BX})(Y-\mu_Y)^T\} \\
&= E\{(AW-\mu_{AW})(Y-\mu_Y)^T\} + E\{(BX-\mu_{BX})(Y-\mu_Y)^T\} \\
&= AE\{(W-\mu_W)(Y-\mu_Y)^T\} + BE\{(X-\mu_X)(Y-\mu_Y)^T\} \\
&= A\Sigma_{WY} + B\Sigma_{XY}
\end{aligned}
$$

So,

$$
\begin{aligned}
\mathrm{cov}(W, CY+DZ) &= \{\mathrm{cov}(CY+DZ, W)\}^T \\
&= (C\Sigma_{YW} + D\Sigma_{ZW})^T \\
&= \Sigma_{WY}C^T + \Sigma_{WZ}D^T
\end{aligned}
$$

1b. According to 1a.,

$$
\begin{aligned}
\mathrm{cov}(AW+BX, CY+DZ) &= \mathrm{cov}(AW, CY+DZ) + \mathrm{cov}(BX, CY+DZ) \\
&= \mathrm{cov}(AW, CY) + \mathrm{cov}(AW, DZ) + \mathrm{cov}(BX, CY) + \mathrm{cov}(BX, DZ) \\
&= A\Sigma_{WY}C^T + A\Sigma_{WZ}D^T + B\Sigma_{XY}C^T + B\Sigma_{XZ}D^T
\end{aligned}
$$

2. Let $\mathbf{A}$ be a square matrix with eigendecomposition $\mathbf{A} = \mathbf{U}\Lambda\mathbf{U}^{-1}$. Given a real number c ($\neq$ any eigenvalue of $\mathbf{A}$), express the eigendecomposition of $(\mathbf{A}-c\mathbf{I})^{-1}$ in terms of $\mathbf{U}$, Λ, $\mathbf{I}$ and c. (Hint: $\mathbf{I} = \mathbf{UU}^{-1}$)

2. $A - cI = U\Lambda U^{-1} - cUU^{-1}$

$= U(\Lambda - cI)U^{-1}$

$\because$ c is not an eigenvalue of A

$\therefore$ $\Lambda - cI$ is invertible and

$$(\Lambda - cI)^{-1} = \begin{bmatrix} (\lambda_1 - c)^{-1} & & \\ & \ddots & \\ & & (\lambda_p - c)^{-1} \end{bmatrix}$$, where $\lambda_1, \dots, \lambda_p$ are eigenvalues of $p \times p$ matrix A

Finally, $(A-cI)^{-1} = U(\Lambda - cI)^{-1}U^{-1}$ because $U(\Lambda-cI)^{-1}U^{-1}(A-cI) = I$

3. Let W be a discrete random variable such that $\Pr(W = 1) = \Pr(W = -1) = 1/2$. Define $Y = WX$ with $X \sim N(0, 1)$ and $X \perp\!\!\!\perp W$. Prove the following identities.
   a. $Y \sim N(0, 1)$.
   b. X and Y are uncorrelated with each other.
   c. X is not independent of Y.

3a. $Pr(Y \le y) = Pr(WX \le y)$

$= Pr(WX \le y \mid W=1)\, Pr(W=1) + Pr(WX \le y \mid W=-1)\, Pr(W=-1)$

$= Pr(X \le y \mid W=1) \cdot \frac{1}{2} + Pr(-X \le y \mid W=-1) \cdot \frac{1}{2}$

$= \frac{1}{2} Pr(X \le y) + \frac{1}{2} Pr(-X \le y)$ $\quad (\because X \perp\!\!\!\perp W)$

$= \frac{1}{2} Pr(X \le y) + \frac{1}{2} Pr(X \le y)$ $\quad (\because$ symmetry of pdf of $N(0,1))$

$= Pr(X \le y)$

$\therefore$ cdf of Y is identical to cdf of X

3b. $cov(X, Y) = E(XY) - E(X)E(Y)$

$= E(WX^2) - E(X)\, E(WX)$

$= E(W)E(X^2) - E(X)\, E(W)\, E(X)$ $\quad (\because X \perp\!\!\!\perp W)$

$= 0$ $\quad (\because E(W) = 0)$

3c. Suppose $X \perp\!\!\!\perp Y$. Then we may reach a contradiction. E.g.,

① $X \perp\!\!\!\perp Y$ and $X, Y \sim N(0,1)$

$\Rightarrow (X, Y)^T \sim MVN_2(0, I)$

$\Rightarrow$ support $(X,Y) = \mathbb{R}^2$ contradicting the identity that support $(X,Y) = \{(x,y) : x = \pm y,\ x \in \mathbb{R}\}$

② $X \perp\!\!\!\perp Y \Rightarrow X^2 \perp\!\!\!\perp Y^2$ contradicting the identity the $Pr(X^2 = Y^2) = 1$.

Since the assumption leads to contradictions, we conclude that $X \not\perp\!\!\!\perp Y$.

4. Let $\mathbf{X} = [X_1, X_2, X_3]^\top \sim MVN_3(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ with

$$\boldsymbol{\mu} = [6, 1, 4]^\top, \quad \boldsymbol{\Sigma} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 5 & -1 \\ 1 & -1 & 3 \end{bmatrix}.$$

a. Find the conditional distribution of X_2 given $X_1 = 2$ and $X_3 = 1$.
b. Find the distribution of random 2-vector $\mathbf{Y} = [3X_1 - 2X_2 + X_3, X_2 - X_3]^\top$.
c. Find $w_1, w_2 \in \mathbb{R}$ such that $W = w_1 X_1 + w_2 X_2 + X_3$ is independent of $\mathbf{Y}$. (Hint: don't forget to verify the normality of random 3-vector $[W, \mathbf{Y}^\top]^\top$ after figuring out values of w_1 and w_2.)

4a. Let $Z_1 = X_2$, $Z_2 = [X_1, X_3]^T$. Then.

$$\begin{bmatrix} Z_1 \\ Z_2 \end{bmatrix} \sim MVN_3\left(\begin{bmatrix} 1 \\ 6 \\ 4 \end{bmatrix}, \begin{bmatrix} 5 & 1 & -1 \\ 1 & 1 & 1 \\ -1 & 1 & 3 \end{bmatrix}\right)$$

$$\Rightarrow E(X_2 \mid X_1=2, X_3=1) = E(Z_1 \mid Z_2 = [2, 1]^T)$$

$$= 1 + [1, -1]\begin{bmatrix} 1 & 1 \\ 1 & 3 \end{bmatrix}^{-1}\left(\begin{bmatrix} 2 \\ 1 \end{bmatrix} - \begin{bmatrix} 6 \\ 4 \end{bmatrix}\right)$$

$$= -4$$

$$\mathrm{cov}(X_2 \mid X_1=2, X_3=1) = \mathrm{cov}(Z_1 \mid Z_2 = [2, 1]^T)$$

$$= 5 - [1, -1]\begin{bmatrix} 1 & 1 \\ 1 & 3 \end{bmatrix}^{-1}\begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

$$= 2$$

$$\Rightarrow X_2 \mid X_1=2, X_3=1 \sim N(-4, 2)$$

b. Let $Y = AX$. Then $A = \begin{bmatrix} 3 & -2 & 1 \\ 0 & 1 & -1 \end{bmatrix}$.

So, $Y \sim MVN_2(A\mu, A\Sigma A^T) = MVN_2\left(\begin{bmatrix} 20 \\ -3 \end{bmatrix}, \begin{bmatrix} 30 & -16 \\ -16 & 10 \end{bmatrix}\right)$

c. Let $b = [w_1\ w_2\ 1]^T$. Then $W = bX$

Let $W \perp\!\!\!\perp Y$. Then $\mathrm{cov}(W, Y) = b\Sigma A^T = O_{1\times 2}$

Solve $b\Sigma A^T = [2w_1 - 8w_2 + 8,\ 6w_2 - 4] = O_{1\times 2}$ with respect to w_1, w_2.

Obtain $w_1 = -4/3$, $w_2 = 2/3$

Finally, verify the normality of $[W, Y^T]^T$. Actually,

$\because [W, Y^T]^T = \begin{bmatrix} b^T \\ A \end{bmatrix} X$

and $\det\left(\begin{bmatrix} b^T \\ A \end{bmatrix} \Sigma [b\ A^T]\right) = 53.78 > 0$.

$\therefore [W, Y^T]^T \sim MVN_3\left(\begin{bmatrix} b^T \\ A \end{bmatrix}\mu, \begin{bmatrix} b^T \\ A \end{bmatrix}\Sigma[b\ A^T]\right)$

i.e., $[W, Y^T]^T$ is of MVN

```r
options(digits = 4)
## Q4a.
(Mu.condi = 1+t(matrix(c(1,-1))) %*%
solve(matrix(
  c(1, 1,
    1, 3), ncol = 2, nrow = 2
))%*%
(matrix(c(2,1))-matrix(c(6,4))))
(Sigma.condi = 5-t(matrix(c(1,-1))) %*%
solve(matrix(
  c(1, 1,
    1, 3), ncol = 2, nrow = 2
))%*%
matrix(c(1,-1)))

## Q4b.
Mu = matrix(c(6, 1, 4))
Sigma = matrix(
  c(1, 1, 1,
    1, 5, -1,
    1, -1, 3), ncol = 3, nrow = 3
)
A = matrix(
  c(3, -2, 1,
    0, 1, -1),
  ncol = 3, nrow = 2, byrow = T
)
```

```r
(Mu.Y = A %*% Mu)
(Sigma.Y = A %*% Sigma %*% t(A))

## Q4c.
b = matrix(c(-4/3, 2/3, 1))
det(rbind(t(b), A) %*% Sigma %*% t(rbind(t(b), A))) # check the positive definiteness
```