Multivariate Analysis 2020/2021
Applied 1st Test - 4 Feb 2021

Group I - A

1. $X = \underset{\sim}{A} Z + \underset{\sim}{b}$ where $\underset{\sim}{Z} \sim N_2(\underset{\sim}{0}, \underset{\sim}{I})$

$$E(\underset{\sim}{X}) = \underset{\sim}{A} E(Z) + b = \underset{\sim}{b} = \underset{\sim}{\mu} =$$

$$Var(X) = \underset{\sim}{A} Var(\underset{\sim}{Z}) \underset{\sim}{A}^t = \underset{\sim}{A}\underset{\sim}{A}^t = \underset{\sim}{\Sigma} = (\underset{\sim}{\Sigma}^{1/2})(\underset{\sim}{\Sigma}^{1/2})^t$$

where $\underset{\sim}{\Sigma} = \underset{\sim}{\Gamma} \underset{\sim}{\Lambda} \underset{\sim}{\Gamma}^t$ the spectral decomposition of $\underset{\sim}{\Sigma}$ then $\underset{\sim}{\Sigma}^{1/2} = \underset{\sim}{\Gamma} \underset{\sim}{\Lambda}^{1/2} \underset{\sim}{\Gamma}^t$

$$\det(\Sigma - \lambda I) = 0 \Leftrightarrow \det \begin{bmatrix} \sigma_{11} - \lambda & \sigma_{12} \\ \sigma_{12} & \sigma_{22} - \lambda \end{bmatrix} = 0$$

$$\Leftrightarrow (\sigma_{11} - \lambda)(\sigma_{22} - \lambda) - \sigma_{12}^2 = 0 \Leftrightarrow \sigma_{11}\sigma_{22} - \sigma_{11}\lambda - \sigma_{22}\lambda + \lambda^2 - \sigma_{12}^2 = 0$$

$$\lambda = \frac{(\sigma_{11} + \sigma_{22}) \pm \sqrt{(\sigma_{11} + \sigma_{22})^2 - 4(\sigma_{11}\sigma_{22} - \sigma_{12}^2)}}{2}$$

$$\lambda = \frac{(\sigma_{11} + \sigma_{22}) \pm \sqrt{(\sigma_{11} - \sigma_{22})^2 + 4\sigma_{12}^2}}{2}$$

Version A:

$$\lambda = \frac{6 \pm \sqrt{4 + 4(4)}}{2} = \frac{6 \pm \sqrt{20}}{2} = 3 \pm \sqrt{5}$$

$$\Leftrightarrow \lambda_1 = 3 + \sqrt{5} \simeq 5.2361 \quad \wedge \quad \lambda_2 = 3 - \sqrt{5} \simeq 0.7639$$

Version B:

$$d = \frac{6 \pm \sqrt{4 + 4(1)^2}}{2} = \frac{6 \pm 2\sqrt{2}}{2} = 3 \pm \sqrt{2}$$

$$d_1 = 4.4142 \quad \vee \quad d_2 = 1.5858$$

Eigenvalues:

Version A:

$$\begin{bmatrix} 2 - d_i & 2 \\ 2 & 4 - d_i \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \underset{\sim}{0} \iff \begin{cases} (2-d_1)x + 2y = 0 \\ \text{—} \end{cases}$$

$$\iff \begin{cases} y = -\frac{(2-d_1)}{2} x \\ \end{cases} \qquad \underset{\sim}{x} = \begin{bmatrix} 1 \\ \frac{d_1 - 2}{2} \end{bmatrix}$$

$$\underset{\sim}{v}_i = \frac{1}{\sqrt{1^2 + \left(\frac{d_i - 2}{2}\right)^2}} \begin{bmatrix} 1 \\ \frac{d_i - 2}{2} \end{bmatrix}$$

$$d_1 = 5.2361 \Rightarrow \underset{\sim}{v}_1 = \frac{1}{\sqrt{1 + \left(\frac{3.2361}{4}\right)^2}} \begin{bmatrix} 1 \\ 1.61805 \end{bmatrix} = \frac{1}{\sqrt{3.6180}} \begin{bmatrix} 1 \\ 1.61805 \end{bmatrix}$$

$$\underset{\sim}{v}_1 = \begin{bmatrix} 0.5257 \\ 0.85065 \end{bmatrix} \qquad \underset{\sim}{v}_2 = \begin{bmatrix} -0.85065 \\ 0.5257 \end{bmatrix}$$

Similarly: (Version B)

$$\underset{\sim}{v}_1 = \begin{bmatrix} 0.9239 \\ 0.3827 \end{bmatrix} \quad \wedge \quad \underset{\sim}{v}_2 = \begin{bmatrix} -0.3827 \\ 0.9239 \end{bmatrix}$$

Thus, $\underset{\sim}{A} = \Gamma \Lambda^{-1/2} \Gamma^t =$

And:

$$\underset{\sim}{A}(\text{Version } A) = \begin{bmatrix} 1.2649 & 0.6325 \\ & 1.8974 \end{bmatrix}$$

$$\underset{\sim}{A}(\text{Version } B) = \begin{bmatrix} 1.9777 & 0.2976 \\ & 1.3825 \end{bmatrix}$$

Thus, $\underset{\sim}{X} = A(\text{Version } i)\, \underset{\sim}{Z} + \underset{\sim}{b}(\text{Version } i)$

$$\underset{\sim}{b}(\text{Version } i) = \begin{cases} (3,2)^t, & i = A \\ (2,3)^t, & i = B \end{cases}$$

As an alternative:

$$\underset{\sim}{A}\underset{\sim}{A}^t = \underset{\sim}{\Sigma} \iff \begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix} \begin{bmatrix} a_1 & a_3 \\ a_2 & a_4 \end{bmatrix} = \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ & \sigma_{22} \end{bmatrix}$$

$$\iff \begin{bmatrix} a_1^2 + a_2^2 & a_1 a_3 + a_2 a_4 \\ & a_3^2 + a_4^2 \end{bmatrix} = \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ & \sigma_{22} \end{bmatrix}$$

$$\begin{cases} a_1^2 + a_2^2 = \sigma_{11} \\ a_1 a_3 + a_2 a_4 = \sigma_{12} \\ a_3^2 + a_4^2 = \sigma_{22} \end{cases} \iff \begin{cases} a_1 = \dfrac{\sigma_{12} - a_2 a_4}{a_3} \end{cases}$$

Session B:

$\sigma_{11} = 4 \quad , \quad \sigma_{12} = 1 \quad , \quad \sigma_{22} = 2$

$$\begin{cases} a_1^2 + a_2^2 = 4 \\ a_1 a_3 + a_2 a_4 = 1 \\ a_3^2 + a_4^2 = 2 \end{cases}$$

A possible solution is:

$$\begin{cases} a_1 = 0 \quad \wedge \quad a_4 = \frac{1}{2} \\ a_2 = 2 \\ a_3 = \frac{\sqrt{7}}{2} \end{cases}$$

Leading to:

$$\underset{\sim}{X} = \begin{bmatrix} 0 & 2 \\ \sqrt{7}/2 & \frac{1}{2} \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} + \begin{bmatrix} 2 \\ 3 \end{bmatrix}$$

Version A:

$\sigma_{11} = 2 \quad , \quad \sigma_{12} = 2 \quad , \quad \sigma_{22} = 4$

$$\begin{cases} a_1^2 + a_2^2 = 2 \\ a_1 a_3 + a_2 a_4 = 2 \\ a_3 + a_4^2 = 4 \end{cases}$$

A possible solution is:

$$\begin{cases} a_1 = 1 \quad \wedge \quad a_4 = 0 \\ a_2 = +1 \\ a_3 = 2 \end{cases}$$

Leading to:

$$\underset{\sim}{X} = \begin{bmatrix} 1 & 1 \\ 2 & 0 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} + \begin{bmatrix} 3 \\ 2 \end{bmatrix}$$

2. $\underset{\sim}{X} \sim N_2(\underset{\sim}{\mu}, \underset{=}{\Sigma})$

(a) $\operatorname{Cov}(\underline{\alpha}^t \underset{\sim}{X}, \beta^t \underset{\sim}{X}) = 0 \iff \alpha^t \underset{=}{\Sigma} \underline{\beta} = 0$

$$\iff (\alpha_1 \; \alpha_2)\begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{bmatrix}\begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} = 0$$

$$\iff (\sigma_{11}\alpha_1 + \alpha_2\sigma_{12} \quad \alpha_1\sigma_{12} + \sigma_{22}\alpha_2)\begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} = 0$$

$$\iff \sigma_{11}\alpha_1\beta_1 + \sigma_{12}\alpha_2\beta_1 + \sigma_{12}\alpha_1\beta_2 + \sigma_{22}\alpha_2\beta_2 = 0$$

$$\iff \sigma_{11}\alpha_1\beta_1 + \sigma_{12}\alpha_2\beta_1 + \sigma_{12}\alpha_1\beta_2 + \sigma_{22}\alpha_2\beta_2 = 0 \quad (*)$$

Thus any $\underline{\alpha}, \underline{\beta}$ verifying $(*)$ is a solution of our case. An example can be:

$$\begin{cases} \alpha_1 = \beta_1 = 1 \\ \alpha_2 = 0 \end{cases}$$

then $\beta_2 : \sigma_{11} + \sigma_{12}\beta_2 = 0 \iff \beta_2 = -\sigma_{11}/\sigma_{12}$

and $\underset{\sim}{\alpha} = (1, 0)^t$, $\underline{\beta} = (1, -\sigma_{11}/\sigma_{12})$

Thus, since $\underset{\sim}{Y} = \begin{bmatrix} \underset{\sim}{\alpha}^t \\ \underset{\sim}{\beta}^t \end{bmatrix} \underset{\sim}{X}$ and $\underset{\sim}{X} \sim N_2(\underset{\sim}{\mu}, \underset{=}{\Sigma})$ then

$$\underset{\sim}{Y} \sim N_2(E(\underset{\sim}{Y}), \operatorname{Var}(\underset{\sim}{Y}))$$

2(b) $\{ \underset{\sim}{x} \in \mathbb{R}^2 : P(MD^2(\underset{\sim}{x}, \underset{\sim}{\mu}) \leq \xi) = 0.90$

Given that $MD^2(\underset{\sim}{x}, \underset{\sim}{\mu}) \sim \chi^2_{(2)}$

then using the tables of the $\chi^2_{(2)}$

$$F_{\chi^2_{(2)}}(\xi) = 0.90 \iff \xi = 4.6052$$

thus, $R = \{ \underset{\sim}{x} \in \mathbb{R}^2 : MD^2(\underset{\sim}{x}, \underset{\sim}{\mu}) \leq 4.6052 \}$

$$(\underset{\sim}{x} - \underset{\sim}{\mu})^t \underset{\sim}{\Sigma}^{-1} (\underset{\sim}{x} - \underset{\sim}{\mu}) \leq \xi$$

$$\underset{\sim}{\Sigma}^{-1} = \frac{1}{\sigma_{11}\sigma_{22} - \sigma_{12}^2} \begin{bmatrix} \sigma_{22} & -\sigma_{12} \\ & \sigma_{11} \end{bmatrix}$$

being so,

$$(x_1 - \mu_1 \quad x_2 - \mu_2) \begin{bmatrix} \sigma_{22} & -\sigma_{12} \\ -\sigma_{12} & \sigma_{11} \end{bmatrix} \begin{bmatrix} x_1 - \mu_1 \\ x_2 - \mu_2 \end{bmatrix} =$$

$$= \begin{bmatrix} -\sigma_{12}(x_2 - \mu_2) + (x_1 - \mu_1)\sigma_{22} & -\sigma_{12}(x_1 - \mu_1) + \sigma_{11}(x_2 - \mu_2) \end{bmatrix} \begin{bmatrix} x_1 - \mu_1 \\ x_2 - \mu_2 \end{bmatrix}$$

$$= -\sigma_{12}(x_2 - \mu_2)(x_1 - \mu_1) + (x_1 - \mu_1)^2\sigma_{22} - \sigma_{12}(x_1 - \mu_1)(x_2 - \mu_2) + \sigma_{11}(x_2 - \mu_2)^2$$

$$= (x_1 - \mu_1)^2\sigma_{22} + \sigma_{11}(x_2 - \mu_2)^2 - 2\sigma_{12}(x_1 - \mu_1)(x_2 - \mu_2)$$

thus, the region R is defined as $\underset{\sim}{x} \in \mathbb{R}^2$ such that:

$$(x_1-\mu_1)^2\sigma_{22} + \sigma_{11}(x_2-\mu_2)^2 - 2\sigma_{12}(x_1-\mu_1)(x_2-\mu_2) \leq$$

$$\leq \xi(\sigma_{11}\sigma_{22}-\sigma_{12}^2)$$

Version A:

$$(x_1-2)^2 2 + 2(x_2-1)^2 - 2(x_1-2)(x_2-1) \leq 4.6052 \cdot (4-1)$$

$$(x_1-2)^2 + (x_2-1)^2 - (x_1-2)(x_2-1) \leq 13.8155$$

Version B:

$$2(x_1-2)^2 + 4(x_2-1)^2 - 2(x_1-2)(x_2-4) \leq 41.94025$$

$$\xi = F^{-1}_{\chi^2(2)}(0.95) = 5.991465$$

3. (a) $E(Y) = E(E(Y|X)) = E\left[\begin{pmatrix} X_1 \\ X_1+X_2 \end{pmatrix}\right]$

$$= \begin{bmatrix} E(X_1) \\ E(X_1)+E(X_2) \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_1+\mu_2 \end{bmatrix}$$

3(b) $Var(\underset{\sim}{X}-\underset{\sim}{Y}) = Var(X) + Var(Y) - Cov(X,Y)$

$$- Cov(\underset{\sim}{Y},\underset{\sim}{X}) = \underset{\sim}{\Sigma}_{11} + \underset{\sim}{\Sigma}_{22} - \underset{\sim}{\Sigma}_{12} - \underset{\sim}{\Sigma}_{21}$$

Given that $E(\underset{\sim}{Y}|\underset{\sim}{X}=\underset{\sim}{x}) = \begin{bmatrix} x_1 \\ x_1+x_2 \end{bmatrix} = \underset{\sim}{\mu}_2 + \Sigma_{YX}\Sigma_{XX}^{-1}(x_1-\mu_1)$

$$\Leftrightarrow \begin{bmatrix} 1 \\ 2+1 \end{bmatrix} + \Sigma_{21} \frac{1}{4-1} \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} x_1-1 \\ x_2-2 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_1+x_2 \end{bmatrix}$$

$$\Leftrightarrow \Sigma_{21} \frac{1}{3} \begin{bmatrix} 2(x_1-1)-(x_2-2) \\ -(x_1-1)+2(x_2-2) \end{bmatrix} = \begin{bmatrix} x_1-1 \\ x_1+x_2-3 \end{bmatrix}$$

$$\Leftrightarrow \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 2(x_1-1)-(x_2-2) \\ -(x_1-1)+2(x_2-2) \end{bmatrix} = 3 \begin{bmatrix} x_1-1 \\ x_1+x_2-3 \end{bmatrix}$$

$$\Leftrightarrow \begin{bmatrix} 2a(x_1-1) - a(x_2-2) - b(x_1-1) + 2b(x_2-2) \\ 2c(x_1-1) - c(x_2-2) - d(x_1-1) + 2d(x_2-2) \end{bmatrix} = \begin{bmatrix} 3(x_1-1) \\ 3(x_1+x_2-3) \end{bmatrix}$$

$$\Leftrightarrow \begin{bmatrix} (2a-b)(x_1-1)+(2b-a)(x_2-2) \\ (2c-d)(x_1-1)+(2d-c)(x_2-2) \end{bmatrix} = \begin{bmatrix} 3(x_1-1) \\ 3(x_1-1)+3(x_2-2) \end{bmatrix}$$

$$\Leftrightarrow \begin{cases} 2a-b=3 \\ 2b-a=0 \\ 2c-d=3 \\ 2d-c=3 \end{cases} \Leftrightarrow \begin{cases} 4b-b=3 \\ a=2b \\ d=2c-3 \\ 4c-6-c=3 \end{cases} \Leftrightarrow \begin{cases} 3b=3 \\ — \\ — \\ 3c=9 \end{cases} \Leftrightarrow \begin{cases} b=1 \\ a=2 \\ d=6-3=3 \\ c=3 \end{cases}$$

$$\Sigma_{21} = \begin{bmatrix} 2 & 1 \\ 3 & 3 \end{bmatrix}$$

thus, $Var(X-Y) = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} + \begin{bmatrix} 3 & 3 \\ 3 & 7 \end{bmatrix} - \begin{bmatrix} 2 & 1 \\ 3 & 3 \end{bmatrix} - \begin{bmatrix} 2 & 3 \\ 1 & 3 \end{bmatrix}$

$$= \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}$$

$Var(X+Y) = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} + \begin{bmatrix} 3 & 3 \\ 3 & 7 \end{bmatrix} + \begin{bmatrix} 2 & 1 \\ 3 & 3 \end{bmatrix} + \begin{bmatrix} 2 & 3 \\ 1 & 3 \end{bmatrix}$

$$= \begin{bmatrix} 9 & 8 \\ 8 & 15 \end{bmatrix}$$

$$\boxed{\text{Group II - Session A}}$$

1. $$\underset{\sim}{\Sigma}(\underset{\sim}{X}) = \begin{bmatrix} a_1^2 \operatorname{Var}(U) & a_1 a_2 \operatorname{Var}(U) & a_1 a_3 \operatorname{Var}(U) \\ & a_2^2 \operatorname{Var}(U) & a_2 a_2 \operatorname{Var}(U) \\ & & a_3^2 \operatorname{Var}(U) \end{bmatrix}$$

$$= \operatorname{Var}(U) \begin{bmatrix} a_1^2 & a_1 a_2 & a_1 a_3 \\ & a_2^2 & a_2 a_3 \\ & & a_3^2 \end{bmatrix}$$

$$\operatorname{Cor}(X) = D^{-1/2} \underset{=}{\Sigma} D^{-1/2}, \quad D^{-1/2} = \operatorname{diag}\left\{ \frac{\sqrt{\operatorname{Var}(U)}}{a_1}, \frac{\sqrt{\operatorname{Var}(U)}}{a_2}, \frac{\sqrt{\operatorname{Var}(U)}}{a_3} \right\}$$

$$= \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

$$\det \begin{bmatrix} 1-d_1 & 1 & 1 \\ 1 & 1-d_1 & 1 \\ 1 & 1 & 1-d_1 \end{bmatrix} = (1-d_1)\left[(1-d_1)^2 - 1\right] - (1-d_1-1) + (1-1+d_1) = 0$$

$$\Leftrightarrow (1-d_1)\left[(1-d_1)^2-1\right] + d_1 + d_1 = 0$$

$$\Leftrightarrow (1-d_1)(1-2d_1+d_1^2-1) + 2d_1 = 0$$

$$\Leftrightarrow (1-d_1)\, d_1(d_1-2) + 2d_1 = 0$$

$$\Leftrightarrow d_1\left[(1-d_1)(d_1-2)+2\right] = 0$$

$$\Leftrightarrow d_1 = 0 \ \vee \ d_1 - \cancel{2} - d_1^2 + 2d_1 + \cancel{2} = 0$$

$$\Leftrightarrow d_1 = 0 \ \vee \ d_1(d_1-3) = 0$$

$$\Leftrightarrow d_1 = 0 \ \vee \ d_1 = 0 \ \vee \ d_1 = 3$$

$$\begin{bmatrix} 1-3 & 1 & 1 \\ 1 & 1-3 & 1 \\ 1 & 1 & 1-3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \underset{\sim}{0} \iff \begin{cases} -2x+y+z=0 \\ x-2y+z=0 \\ x+y-2z=0 \end{cases}$$

$$\begin{cases} \text{---} \\ x = 2y-z \\ 2y-z+y-2z=0 \end{cases} \qquad \begin{cases} - \\ x = z \\ y = z \end{cases} \qquad \underline{x} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

e $\underset{\sim}{\gamma}_1 = \frac{\sqrt{3}}{3} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ since $\gamma_1^t \underline{\gamma}_1 = 1$

Version A: $PC_1 = \frac{\sqrt{3}}{2}(X_1 + X_2 + X_3)$

Version B: $PC1 = \frac{\sqrt{3}}{2}\left[\left(\frac{X_1 - 1/2}{1/\sqrt{12}}\right) + \left(\frac{X_2 - 1/2}{1/2\sqrt{3}}\right) + 2\sqrt{3}\left(X - \frac{1}{2}\right)\right]$

$= 3(X_1 + X_2 + X_3 - 1.5)$

$\underset{\sim}{\gamma}_2$ e $\underline{\gamma}_3$ podem ser escolhidos tal que $\underline{\gamma}_i^t \gamma_j = 0$ $\forall i \neq j$ e $\|\underline{\gamma}_i\| = 1$.

Version - A

2(a) $\sum_{i=1}^{7} \hat{\lambda}_i = tr(\underset{\sim}{S})$

thus, $a = \hat{\lambda}_2 = tr(\underset{\sim}{S}) - \sum_{i \neq 2} \hat{\lambda}_i =$

$= 8379.54 - 8190.32$

$= 189.22$

2(b)

$\hat{\lambda}_i / \sum_{i=1}^{7} \hat{\lambda}_i$:	$\sum_{i=1}^{j} \hat{\lambda}_i / \sum_{i=1}^{7} \hat{\lambda}_i$:
0.9630	0.9630
0.0226	0.9856
0.0091	0.9946
0.0030	0.9977
0.0012	0.9989
0.0007	0.9996
0.0004	1.0000

and $\bar{\lambda} = 1197.077$, $\hat{\lambda}_1 > \bar{\lambda}$

also the 1^{st} PC explains 96.30% of the total variability we choose $(k=1)$ one PC.

2(c)

$\hat{PC}_1 = -0.06X_1 + 0.94X_2 + 0.34X_3 + 0.03X_4 + 0.02X_5 +$
$+ 0.03X_6 + 0.01X_7$

$\hat{PC}_2 = -0.34X_1 - 0.34X_2 + 0.86X_3 + 0.01X_4 - 0.11X_5 +$
$+ 0.09X_6 + 0.11X_7$

The most important weights in PC_1 are the ones associated to:

(+) cardiovascular

(+) cancer

Thus, PC_1 quantifies an weighted mean between the number of death due to cardiovascular diseases and cancer disease. So, states with high (low) scores had many (very low number) of deaths due to these two causes.

In the second PC, we can consider the most expressive weights as the ones associated with:

(−) accidents

(−) cardiovascular

and (+) cancer

and can be interpreted as a contrast between accidents and cardio-diseases and cancer as cause of deaths. Being so, a state with heigh positive (negative) score is a state with many (few) number of deaths due to cancer and few (many) number of death due to accidents and cardio-diseases.

Session - B

2 (a)

$\hat{\lambda}_i/p$	$\sum_{i=1}^{j}\hat{\lambda}_i/p$ $\quad (p=7)$
0.4857	0.4857
0.1757	0.6614
0.1514	0.8129
0.0871	0.9000
0.0614	0.9614
0.0314	0.9929
0.0071	1.0000

and $\bar{\lambda} = 1$, $\hat{\lambda}_i > \bar{\lambda}$, $i=1,2,3$

also the first 3 PCs explains 81.29% of the total variability we choose $(k=3)$ three PCs.

2 (b)

(i) $\underline{\gamma}_1^t \underline{\gamma}_2 = 0 \iff 0.28\,\gamma_{11} = 0.1014$

$\iff \gamma_{11} = a = 0.1014/0.28 = 0.3625$

(ii)
$$\hat{PC}_1 = 0.36\left(\frac{x_1 - 44.31}{s_1}\right) - 0.50\left(\frac{x_2 - 398.53}{s_2}\right)$$
$$- 0.52\left(\frac{x_3 - 178.41}{s_3}\right) - 0.30\left(\frac{x_4 - 26.49}{s_4}\right)$$
$$- 0.27\left(\frac{x_5 - 21.04}{s_5}\right) - 0.40\left(\frac{x_5 - 14.84}{s_6}\right)$$

$$-0.18\,\frac{(x_7 - 10.55)}{s_7}$$

where s_i is the sample standard deviation of the i-th variable.

$$\hat{PC}_2 = 0.28\,\frac{(x_1 - 44.31)}{s_1} + 0.13\,\frac{(x_2 - 398.53)}{s_2}$$

$$-0.10\,\frac{(x_3 - 178.41)}{s_3} + 0.38\,\frac{(x_4 - 26.49)}{s_4}$$

$$+ 0.70\,\frac{(x_5 - 21.04)}{s_5} - 0.37\,\frac{(x_5 - 14.84)}{s_6}$$

$$-0.37\,\frac{(x_7 - 10.55)}{s_7}$$

Interpretation:

$\hat{PC}_1 \to$ contrast between nr of deaths due to accidents versus nr of deaths due to a disease

Scores: high (low) scores means a state with many (few) nr of deaths due to accidents and low (many) nr of deaths due to any disease.

$\hat{PC}_2$ = contrast between pulmonary diseases (pulmunary + pneumonia flu) and diabetes and liver

the scores means:

$-$ ———————————————————— $+ \rightarrow \hat{PC}_1$

low nr pulmon.
high nr diabetes + liver

high nr pulmonary diseases
low nr diabetes + liver

(c)