EXERCISES SEMINAR 1

Laura Júlia Melis
laujul03
11/15/2019

EXERCISE 3.5.

Calculate the generalized sample variance for:

a) $X = \begin{bmatrix} 9 & 1 \\ 5 & 3 \\ 1 & 2 \end{bmatrix}$ " Generalized sample variance $= |S| = \det(S)$ where $S =$ sample variance-covariance matrix.

(i) First, we calculate the vector of means:

$$\bar{x} = \frac{1}{n} X' \cdot 1 \quad " \quad \bar{x} = \frac{1}{3} \begin{bmatrix} 9 & 5 & 1 \\ 1 & 3 & 2 \end{bmatrix} \cdot \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 5 & 5 & 5 \\ 2 & 2 & 2 \end{bmatrix}$$

(ii) We calculate now the sample variance-covariance matrix

$$S = \frac{1}{n-1} \cdot \sum_{j=1}^{n} (x_j - \bar{x})(x_j - \bar{x})' \quad " \quad S = \frac{1}{2} \begin{bmatrix} \overset{9-5}{4} & \overset{5-5}{0} & -4 \\ -1 & 1 & 0 \end{bmatrix} \begin{bmatrix} 4 & -1 \\ 0 & 1 \\ -4 & 0 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 16+0+16 & -4+0+0 \\ -4+0+0 & 1+1+0 \end{bmatrix}$$

$$S = \frac{1}{2} \begin{bmatrix} 32 & -4 \\ -4 & 2 \end{bmatrix} = \begin{bmatrix} 16 & -2 \\ -2 & 1 \end{bmatrix}$$

(iii) Now we can finally calculate the determinant of S.

$$|S| = 16 \cdot 1 - (-2) \cdot (-2) = 16 - 4 = \boxed{12}$$

b) $X = \begin{bmatrix} 3 & 4 \\ 6 & -2 \\ 3 & 1 \end{bmatrix}$ $n = 3$

(i) $\bar{x} = \frac{1}{n} \cdot X' \cdot 1$ " $\bar{x} = \frac{1}{3} \cdot \begin{bmatrix} 3 & 6 & 3 \\ 4 & -2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 4 & 4 & 4 \\ 1 & 1 & 1 \end{bmatrix}$

(ii) $$S = \frac{1}{3-1} \cdot \begin{bmatrix} 3-4 & 6-4 & 3-4 \\ 4-1 & -2-1 & 1-1 \end{bmatrix} \begin{bmatrix} 3-4 & 4-1 \\ 6-4 & -2-1 \\ 3-4 & 1-1 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} -1 & 2 & -1 \\ 3 & -3 & 0 \end{bmatrix} \begin{bmatrix} -1 & 3 \\ 2 & -3 \\ -1 & 0 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 6 & -9 \\ -9 & 18 \end{bmatrix} = \begin{bmatrix} 3 & -9/2 \\ -9/2 & 9 \end{bmatrix}$$

(iii) $|S| = 3 \cdot 9 - \left(-\frac{9}{2}\right)^2 = 27 - \frac{81}{4} = \boxed{\frac{27}{4}}$

EXERCISE 3.6.

Consider the matrix $X = \begin{bmatrix} -1 & 3 & -2 \\ 2 & 4 & 2 \\ 5 & 2 & 3 \end{bmatrix}$

a) Calculate the matrix of deviations, $X - 1\bar{x}'$. Is this matrix of full rank?

(i) We calculate the col. means:

$$\left.\begin{aligned} \bar{x}_1 &= (-1+2+5)/3 = 6/3 = 2 \\ \bar{x}_2 &= (3+4+2)/3 = 9/3 = 3 \\ \bar{x}_3 &= (-2+2+3)/3 = 3/3 = 1 \end{aligned}\right\} 1\bar{x}' = \begin{bmatrix} 2 & 3 & 1 \\ 2 & 3 & 1 \\ 2 & 3 & 1 \end{bmatrix}$$

(ii) Matrix of deviations:

$$X - 1\bar{x}' = \begin{bmatrix} -1 & 3 & -2 \\ 2 & 4 & 2 \\ 5 & 2 & 3 \end{bmatrix} - \begin{bmatrix} 2 & 3 & 1 \\ 2 & 3 & 1 \\ 2 & 3 & 1 \end{bmatrix} = \begin{array}{c} \begin{matrix} d_1 & \ \ d_2 & \ \ d_3 \end{matrix} \\ \begin{bmatrix} -3 & 0 & -3 \\ 0 & 1 & 1 \\ 3 & -1 & 2 \end{bmatrix} \end{array}$$

(iii) A squared matrix (n×n) has full rank either when its rows or its columns are linearly independent

In our case, we can se that $d_2 = d_1 - d_3$, so the matrix of deviations is not of full rank.

b) Determine S and calculate $|S|$. Interpret the latter geometrically.

$$S = \frac{1}{n-1}(X - 1\bar{x}')' \cdot (X - 1\bar{x}') = \frac{1}{3-1}\begin{bmatrix} -3 & 0 & 3 \\ 0 & 1 & -1 \\ -3 & 1 & 2 \end{bmatrix} \cdot \begin{bmatrix} -3 & 0 & -3 \\ 0 & 1 & 1 \\ 3 & -1 & 2 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 9+0+9 & 0+0-3 & 9+0+6 \\ 0+0-3 & 0+1+1 & 0+1-2 \\ 9+0+6 & 0+1-2 & 9+1+4 \end{bmatrix} =$$

$$S = \frac{1}{2}\begin{bmatrix} 18 & -3 & 15 \\ -3 & 2 & -1 \\ 15 & -1 & 14 \end{bmatrix} = \begin{bmatrix} 9 & -3/2 & \frac{15}{2} \\ -3/2 & 1 & -1/2 \\ 15/2 & -1/2 & 7 \end{bmatrix}$$

$$|S| = 9 \cdot 1 \cdot 7 + \frac{15}{2} \cdot \left(-\frac{3}{2}\right)\left(-\frac{1}{2}\right) + \frac{15}{2}\left(-\frac{3}{2}\right)\left(-\frac{1}{2}\right) - \left(\frac{15}{2}\right)^2 \cdot 1 - 9 \cdot \left(-\frac{1}{2}\right)^2 - 7\left(-\frac{3}{2}\right)^2 =$$

$$= 63 + \frac{45}{8} + \frac{45}{8} - \frac{225}{4} - \frac{9}{4} - \frac{63}{4} = 0$$

- Geometrically, the 3 vectors of deviations can be represented in only 2 dimensions. As $\det(S) = 0$, the ellipse collaps to a straight line so the volume that the sample occupies is zero.

c) Calculate the total sample variance.

Total sample variance $= s_{11} + s_{22} + s_{33} = 9 + 1 + 7 = \boxed{17}$

EXERCISE 3.8.

Given $S_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ and $S_2 = \begin{bmatrix} 1 & -1/2 & -1/2 \\ -1/2 & 1 & -1/2 \\ -1/2 & -1/2 & 1 \end{bmatrix}$

a) Calculate the total sample variance for each S. Compare the results.

$S_1 = 1 + 1 + 1 = 3$

$S_2 = 1 + 1 + 1 = 3$

Both cases have a total sample variance of 3 because the variances (the diagonal elements) of each variable are the same although the covariances of each pair of variables are different. In S_1, the covariences are zero, meaning that each pair of variables are linearly independent.

b) Calculate the generalized sample variance and compare.

$|S_1| = 1 \cdot 1 \cdot 1 + 0 + 0 - 0 - 0 - 0 = 1$

$|S_2| = 1 \cdot 1 \cdot 1 + \left(-\frac{1}{2}\right)^3 + \left(-\frac{1}{2}\right)^3 - 1 \cdot \left(-\frac{1}{2}\right)^2 - 1 \cdot \left(-\frac{1}{2}\right)^2 - 1 \cdot \left(-\frac{1}{2}\right)^2 = 1 - \frac{1}{8} - \frac{1}{8} - \frac{1}{4} - \frac{1}{4} - \frac{1}{4} = 0$

The determinant of S_2 equals 0 because there exists linear dependency between its rows so its volume equals 0 (a line, geometrically) while S_1 can be represented as an ellipse.

EXERCISE 3.9

Data matrix about test scores: $X = \begin{bmatrix} 12 & 17 & 29 \\ 18 & 20 & 38 \\ 14 & 16 & 30 \\ 20 & 18 & 38 \\ 16 & 19 & 35 \end{bmatrix}$

a) Obtain the mean corrected data matrix, and verify that the columns are linearly dependant. Specify an $a' = [a_1, a_2, a_3]$ vector that establishes the linear dependance.

The deviation or mean corrected data matrix is calculated by $X - 1\bar{x}'$

(i) $\bar{x}_1 = (12+18+14+20+16)/5 = 80/5 = 16$

$\bar{x}_2 = (17+20+16+18+19)/5 = 90/5 = 18$

$\bar{x}_3 = (29+38+30+38+35)/5 = 170/5 = 34$

$$1\bar{x}' = \begin{bmatrix} 16 & 18 & 34 \\ 16 & 18 & 34 \\ 16 & 18 & 34 \\ 16 & 18 & 34 \\ 16 & 18 & 34 \end{bmatrix}$$

(ii) $$X - 1\bar{x}' = \begin{bmatrix} 12-16 & 17-18 & 29-34 \\ 18-16 & 20-18 & 38-34 \\ 14-16 & 16-18 & 30-34 \\ 20-16 & 18-18 & 38-34 \\ 16-16 & 19-18 & 35-34 \end{bmatrix} = \begin{bmatrix} -4 & -1 & -5 \\ 2 & 2 & 4 \\ -2 & -2 & -4 \\ 4 & 0 & 4 \\ 0 & 1 & 1 \end{bmatrix}$$

$$\begin{array}{ccc} \parallel & \parallel & \parallel \\ d_1 & d_2 & d_3 \end{array}$$

(iii) We can see that columns are linearly dependend because $d_3 = d_1 + d_2$.

If we specify $a = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$, then $(X - 1\bar{x}') \cdot a = 0$.

b) Obtain the sample covariance matrix S, and verify that $|S| = 0$. Also, show that $Sa = 0$, so a can be rescaled to be an eigenvector corresponding to eigenvalue zero.

(i) $$S = \frac{1}{n-1}(X - 1\bar{x}')' \cdot (X - 1\bar{x}') = \frac{1}{4} \begin{bmatrix} -4 & 2 & -2 & 4 & 0 \\ -1 & 2 & -2 & 0 & 1 \\ -5 & 4 & -4 & 4 & 1 \end{bmatrix} \cdot \begin{bmatrix} -4 & -1 & -5 \\ 2 & 2 & 4 \\ -2 & -2 & -4 \\ 4 & 0 & 4 \\ 0 & 1 & 1 \end{bmatrix} = \frac{1}{4} =$$

$$S = \begin{bmatrix} 16+4+4+16+0 & 4+4+4+0+0 & 20+8+8+16+0 \\ 4+4+4+0+0 & 1+4+4+0+1 & 5+8+8+0+1 \\ 20+8+8+16+0 & 5+8+8+0+1 & 25+16+16+16+1 \end{bmatrix} \cdot \frac{1}{4} = \frac{1}{4} \begin{bmatrix} 40 & 12 & 52 \\ 12 & 10 & 22 \\ 52 & 22 & 74 \end{bmatrix} = \begin{bmatrix} 10 & 3 & 13 \\ 3 & 5/2 & 11/2 \\ 13 & 11/2 & 37/2 \end{bmatrix}$$

(ii) $$|S| = 10 \cdot \frac{5}{2} \cdot \frac{37}{2} + 3 \cdot \frac{11}{2} \cdot 13 + 3 \cdot \frac{11}{2} \cdot 13 - 13^2 \cdot \frac{5}{2} - 3^2 \cdot \frac{37}{2} - \left(\frac{11}{2}\right)^2 \cdot 10 = \boxed{0}$$

(iii) $$S \cdot a = \begin{bmatrix} 10 & 3 & 13 \\ 3 & 5/2 & 11/2 \\ 13 & 11/2 & 37/2 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 10+3-13 \\ 3+5/2-11/2 \\ 13+11/2-\frac{37}{2} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

c) Verify that the third column of the data matrix is the sum of the first two columns. That is, show that there is linear dependence with $a_1 = 1$, $a_2 = 1$, and $a_3 = -1$.

$$X_3 = X_1 + X_2 = \begin{bmatrix} 12+17 \\ 18+20 \\ 14+16 \\ 20+18 \\ 16+19 \end{bmatrix} = \begin{bmatrix} 29 \\ 38 \\ 30 \\ 38 \\ 35 \end{bmatrix}$$

EXERCISE 2.26.

Use $\Sigma = \begin{bmatrix} 25 & -2 & 4 \\ -2 & 4 & 1 \\ 4 & 1 & 9 \end{bmatrix}$

a) Find ρ_{13}.

The correlation coefficient ρ_{ik} is defined as : $\rho_{ik} = \dfrac{\sigma_{ik}}{\sqrt{\sigma_{ii}} \cdot \sqrt{\sigma_{kk}}}$

So $\rho_{13} = \dfrac{\sigma_{13}}{\sqrt{\sigma_{11}} \cdot \sqrt{\sigma_{33}}} = \dfrac{4}{\sqrt{25} \cdot \sqrt{9}} = \dfrac{4}{5 \cdot 3} = \dfrac{4}{15} = \boxed{0'26}$

b) Find the correlation between X_1 and $\frac{1}{2}X_2 + \frac{1}{2}X_3$

$$Corr\left(X_1, \tfrac{1}{2}X_2 + \tfrac{1}{2}X_3\right) = \frac{cov\left(X_1, \frac{1}{2}X_2 + \frac{1}{2}X_3\right)}{\sqrt{Var(X_1)} \cdot \sqrt{Var\left(\frac{1}{2}X_2 + \frac{1}{2}X_3\right)}} = \frac{1}{\sqrt{25} \cdot \sqrt{3'75}} = \boxed{0'1033}$$

- $Var(X_1) = \sigma_{11} = 25$
- $Var\left(\frac{1}{2}X_2 + \frac{1}{2}X_3\right) = \left(\frac{1}{2}\right)^2 \cdot Var(X_2) + \left(\frac{1}{2}\right)^2 \cdot Var(X_3) + 2\left(\frac{1}{2}\right)^2 \cdot \sigma_{23} = \frac{1}{4}\sigma_{22} + \frac{1}{4}\sigma_{33} + \frac{1}{2}\sigma_{23} =$

  $= 1 + \frac{9}{4} + \frac{1}{2} = \frac{15}{4} = 3'75$

- $Cov\left(X_1, \frac{1}{2}X_2 + \frac{1}{2}X_3\right) = 1 \cdot \frac{1}{2} Cov(X_1, X_2) + 1 \cdot \frac{1}{2} Cov(X_1, X_3) = \frac{1}{2}\sigma_{12} + \frac{1}{2}\sigma_{13} = \frac{1}{2} \cdot (-2) + \frac{1}{2}(4) = -1 + 2 = 1$

  ↓

  by bilinearity : $cov(aX, bY + cW) = ab \cdot cov(X, Y) + ac \cdot cov(X, W)$.

EXERCISE 2.27.

Derive expressions for the mean and variances of the following linear combinations in terms of the means and covariances of the random variables X_1, X_2 and X_3.

a) $X_1 - 2X_2$: (i) $E(X_1 - 2X_2) = E(X_1) - 2E(X_2) = \boxed{\mu_1 - 2\mu_2}$

(ii) $V(X_1 - 2X_2) = \boxed{\sigma_{11} + 4\sigma_{22} + 4\sigma_{12}}$

b) $-X_1 + 3X_2$: (i) $E(-X_1 + 3X_2) = \boxed{-\mu_1 + 3\mu_2}$

(ii) $V(-X_1 + 3X_2) = \boxed{\sigma_{11} + 9\sigma_{22} - 6\sigma_{12}}$

c) $X_1 + X_2 + X_3$: (i) $E(X_1 + X_2 + X_3) = \boxed{\mu_1 + \mu_2 + \mu_3}$

(ii) $V(X_1 + X_2 + X_3) = \boxed{\sigma_{11} + \sigma_{22} + \sigma_{33} + 2\sigma_{12} + 2\sigma_{23} + 2\sigma_{13}}$

↓

$$Var\left(\sum_{i=1}^{n} a_i X_i\right) = \sum_{i=1}^{n} a_i^2 Var(X_i) + 2\sum_{i=1}^{n}\sum_{j: j>i}^{n} a_i a_j Cov(X_i, X_j)$$

d) $X_1 + 2X_2 - X_3$: (i) $E(X_1 + 2X_2 - X_3) = \boxed{\mu_1 + 2\mu_2 - \mu_3}$

(ii) $V(X_1 + 2X_2 - X_3) = \boxed{\sigma_{11} + 4\sigma_{22} + \sigma_{33} + 4\sigma_{12} - 2\sigma_{13} - 4\sigma_{23}}$

e) $8X_1 - 4X_2$ if $X_1 \perp X_2$: (i) $E(3X_1 - 4X_2) = \boxed{3\mu_1 - 4\mu_2}$

$\sigma_{12} = 0$

(ii) $V(3X_1 - 4X_2) = 9\sigma_{11} + 16\sigma_{22} + \cancel{2 \cdot 3(-4)\sigma_{12}} = \boxed{9\sigma_{11} + 16\sigma_{22}}$

EXERCISE 2.35.

Using the vectors $b' = [-4 \;\; 3]$ and $d' = [1 \;\; 1]$ verify $(b'd)^2 \leq (b'Bb)(d'B^{-1}d)$ if $B = \begin{bmatrix} 2 & -2 \\ -2 & 5 \end{bmatrix}$

- $b' \cdot d = [-4 \;\; 3]\begin{bmatrix} 1 \\ 1 \end{bmatrix} = -4 + 3 = -1 \implies (b'd)^2 = (-1)^2 = +1$
- $b' \cdot B \cdot b = [-4 \;\; 3] \cdot \begin{bmatrix} 2 & -2 \\ -2 & 5 \end{bmatrix}\begin{bmatrix} -4 \\ 3 \end{bmatrix} = [-8 - 6 \quad +8 + 15]\begin{bmatrix} -4 \\ 3 \end{bmatrix} = [-14 \quad 23]\begin{bmatrix} -4 \\ 3 \end{bmatrix} = (-14)\cdot(-4) + 23 \cdot 3 = 125$
- $B^{-1} = \dfrac{(\mathrm{Adj}(B))^T}{|B|} = \dfrac{\left(\begin{bmatrix} 5 & 2 \\ 2 & 2 \end{bmatrix}\right)^T}{2 \cdot 5 - (-2)\cdot(-2)} = \dfrac{\begin{bmatrix} 5 & 2 \\ 2 & 2 \end{bmatrix}}{6} = \begin{bmatrix} 5/6 & 2/6 \\ 2/6 & 2/6 \end{bmatrix}$
- $d' \cdot B^{-1} \cdot d = [1 \;\; 1]\begin{bmatrix} 5/6 & 2/6 \\ 2/6 & 2/6 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \left[\frac{5}{6} + \frac{2}{6} \quad \frac{2}{6} + \frac{2}{6}\right]\begin{bmatrix} 1 \\ 1 \end{bmatrix} = [7/6 \quad 4/6]\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \frac{7}{6} + \frac{4}{6} = \frac{11}{6}$

So: $(b' \cdot d)^2 = 1$

$(b'Bb) \cdot (d'B^{-1}d) = 125 \cdot \frac{11}{6} = \frac{1375}{6} = 229'1\widehat{6}$

$1 < 229'1\widehat{6}$ ✓