2. Consider again the random sample $\mathbf{X}$ from $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ and $\mathbf{x}$ and $\mathbf{S}$ are the sample mean and the sample variance-covariance matrix, respectively.

a. Show that $\bar{\mathbf{x}}$ and $\mathbf{S}$ are jointly complete and sufficient statistics for $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$

b. Show that $\bar{\mathbf{x}}$ and $\mathbf{S}$ are independently distributed

3. Let the column vectors $\mathbf{x}_i : p \times 1 \sim N_p(\mathbf{0}, \sigma^2 \mathbf{I}_p)$, where $i = 1, 2, ..., n$. Show that

$$\sum_{i=1}^{n} \frac{\mathbf{x}_i' \mathbf{x}_i}{\sigma^2} \sim \chi^2(pn),$$

where $\chi^2(n)$ represents the chi-square distribution with n degrees of freedom (**2 points**)

2.

(a)

$$f(\mathbf{x}; \theta) = \left(\frac{1}{2\pi^{p/2} |\boldsymbol{\Sigma}|^{1/2}}\right)^n \exp\left(-\frac{1}{2}\left(\sum_{i=1}^{n} \mathbf{x}_i' \boldsymbol{\Sigma}^{-1} \mathbf{x}_i - 2\sum_{i=1}^{n} \mathbf{x}_i' \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu} + n\boldsymbol{\mu}'\boldsymbol{\mu}\right)\right)$$

$$= \left(\frac{1}{2\pi^{p/2} |\boldsymbol{\Sigma}|^{1/2}}\right)^n \exp\left(-\frac{1}{2}\left(\sum_{i=1}^{n} \mathbf{x}_i' \boldsymbol{\Sigma}^{-1} \mathbf{x}_i\right)\right) \cdot \exp\left(\sum_{i=1}^{n} \mathbf{x}_i' \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu} - \frac{n}{2}\boldsymbol{\mu}'\boldsymbol{\mu}\right)$$

So, $\bar{\mathbf{x}}$ and $\mathbf{S}$ are sufficient statistics.

(b)

Let $\mathbf{z}_i = \mathbf{x}_i - \boldsymbol{\mu}$,for i =1,2, …, n.

Let $\bar{\mathbf{z}} = \frac{1}{n} \cdot \sum_{i=1}^{n} \mathbf{z}_i = \frac{1}{n} \cdot \sum_{i=1}^{n} \mathbf{x}_i - \frac{1}{n} \cdot \sum_{i=1}^{n} \boldsymbol{\mu} = \bar{\mathbf{x}} - \boldsymbol{\mu}$.

We have, $\mathbf{z}_i \sim N_p(\mathbf{0}, \boldsymbol{\Sigma})$ and

$$\text{Cov}(\mathbf{z}_i, \mathbf{z}_j) = \begin{cases} 0\ , if\ i = j \\ \boldsymbol{\Sigma}\ , if\ i \neq j \end{cases}$$

$$\begin{aligned}(n-1)\mathbf{S} &= \sum_{i=1}^{n} (\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})' \\ &= \sum_{i=1}^{n} ((\mathbf{x}_i - \boldsymbol{\mu}) - (\bar{\mathbf{x}} - \boldsymbol{\mu}))((\mathbf{x}_i - \boldsymbol{\mu}) - (\bar{\mathbf{x}} - \boldsymbol{\mu}))' \\ &= \sum_{i=1}^{n} (\mathbf{z}_i - \bar{\mathbf{z}})(\mathbf{z}_i - \bar{\mathbf{z}})' \\ &= \sum_{i=1}^{n} (\mathbf{z}_i \cdot \mathbf{z}_i{}' - \bar{\mathbf{z}} \cdot \mathbf{z}_i{}' - \mathbf{z}_i \cdot \bar{\mathbf{z}}' + \bar{\mathbf{z}} \cdot \bar{\mathbf{z}}') \\ &= \sum_{i=1}^{n} \mathbf{z}_i \cdot \mathbf{z}_i{}' - \bar{\mathbf{z}} \cdot \sum_{i=1}^{n} \mathbf{z}_i{}' - \left(\sum_{i=1}^{n} \mathbf{z}_i\right) \bar{\mathbf{z}}' + n \cdot \bar{\mathbf{z}} \cdot \bar{\mathbf{z}}' \\ &= \sum_{i=1}^{n} \mathbf{z}_i \cdot \mathbf{z}_i{}' - n \cdot \bar{\mathbf{z}} \cdot \bar{\mathbf{z}}'\end{aligned}$$

Take an n-dimensional orthogonal matrix A=(a_{ij}) with all elements of the first column are $\frac{1}{\sqrt{n}}$, Let

$$(\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_n) = (\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_n)\text{A}$$

That is, $\mathbf{y}_i = \sum_{j=1}^{n} a_{ji} \mathbf{z}_j$, thus we have

$E(\mathbf{y}_i) = \sum_{j=1}^{n} a_{ji} E(\mathbf{z}_j) = \mathbf{0}$.

$$\mathrm{Cov}(\mathbf{y}_i,\mathbf{y}_k) = \mathrm{Cov}\left(\sum_{j=1}^{n} a_{ji}\mathbf{z}_j , \sum_{l=1}^{n} a_{lk}\mathbf{z}_l\right) = \sum_{j=1}^{n}\sum_{l=1}^{n} a_{ji}a_{lk}Cov(\mathbf{z}_j,\mathbf{z}_l)$$

$$= \sum_{j=1}^{n} a_{ji}a_{jk} = \begin{cases} 0 , if\ i = k \\ \boldsymbol{\Sigma} , if\ i \neq k \end{cases}$$

Therefore, $\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_n$ are independently and identically distributed normal distribution random variables and $\mathbf{y}_i \sim N_p(\mathbf{0}, \boldsymbol{\Sigma})$.

$$\mathbf{y}_1 = \sum_{j=1}^{n} a_{j1}\,\mathbf{z}_j = \sum_{j=1}^{n} \frac{1}{\sqrt{n}}\mathbf{z}_j = \sqrt{n}\bar{\mathbf{z}}$$

$$\mathbf{y}_1 \cdot \mathbf{y}_1' = n \cdot \bar{\mathbf{z}} \cdot \bar{\mathbf{z}}'$$

$$\sum_{i=1}^{n} \mathbf{y}_i \cdot \mathbf{y}_i{}' = (\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_n)(\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_n)'$$

$$= (\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_n)\mathrm{A}\big((\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_n)\mathrm{A}\big)'$$

$$= (\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_n)\mathrm{AA}'^{(\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_n)'} = \sum_{i=1}^{n} \mathbf{z}_i \cdot \mathbf{z}_i{}'$$

$$(n-1)\mathbf{S} = \sum_{i=1}^{n} \mathbf{z}_i \cdot \mathbf{z}_i{}' - n \cdot \bar{\mathbf{z}} \cdot \bar{\mathbf{z}}' = \sum_{i=1}^{n} \mathbf{y}_i \cdot \mathbf{y}_i{}' - \mathbf{y}_1 \cdot \mathbf{y}_1' = \sum_{i=2}^{n} \mathbf{y}_i \cdot \mathbf{y}_i{}'$$

Because

$$\bar{\mathbf{x}} = \bar{\mathbf{z}} + \boldsymbol{\mu} = \frac{\mathbf{y}_1}{\sqrt{n}} + \boldsymbol{\mu}$$

and

$$\mathbf{S} = \frac{1}{n-1} \cdot \sum_{i=2}^{n} \mathbf{y}_i \cdot \mathbf{y}_i{}'$$

$\bar{\mathbf{x}}$ only depends on $\mathbf{y}_1$, while $\mathbf{S}$ only depends on $\mathbf{y}_2, \mathbf{y}_3, \dots, \mathbf{y}_n$. We can conclude that $\bar{\mathbf{x}}$ and $\mathbf{S}$ are independently distributed.

3.

Let $\mathbf{x}_i = \begin{pmatrix} \mathbf{x}_i^1 \\ \mathbf{x}_i^2 \\ \dots \\ \mathbf{x}_i^p \end{pmatrix}$. Since $\mathbf{x}_i \sim N_p(\mathbf{0}, \sigma^2 \mathbf{I}_p)$, $\mathbf{x}_i^1$, $\mathbf{x}_i^2$, …, $\mathbf{x}_i^p$ are independently and identically distributed normal distribution random variables. $\mathbf{x}_i^j \sim N(0, \sigma^2)$ for j = 1, 2, …, p. So $\frac{\mathbf{x}_i^j}{\sigma} \sim N(0, 1)$

Therefore,

$$\sum_{i=1}^{n} \frac{\mathbf{x}_i{}' \mathbf{x}_i}{\sigma^2} = \sum_{i=1}^{n} \sum_{j=1}^{p} \frac{\mathbf{x}_i^{j\,2}}{\sigma^2}$$

According to the definition of chi-square distribution, we have

$$\sum_{i=1}^{n} \frac{\mathbf{x}_i{}' \mathbf{x}_i}{\sigma^2} \sim \chi^2(pn)$$