MULTIVARIATE STAT. HOMEWORK I.

1. (*2 point*) Let $\mathbf{X} \sim \mathcal{N}_n(\mathbf{m}, \mathbf{C})$ be a random vector. Let us separate the components of $\mathbf{X}$ into two parts: let $\mathbf{X}_1$ and $\mathbf{X}_2$ denote the random vectors consisting of the first p and the other $n-p$ components, respectively. Prove that the random vectors $\mathbf{X}_1$ and $\mathbf{X}_2$ are independent if and only if, the covariance matrix $\mathbf{C}$ can be partitioned in the following way:

$$\begin{pmatrix} \mathbf{C}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{C}_2 \end{pmatrix},$$

where $\mathbf{C}_1$ and $\mathbf{C}_2$ are covariance matrices of $\mathbf{X}_1$ and $\mathbf{X}_2$, respectively, and $\mathbf{0}$ is the zero matrix.

2. (*2 point*) Let $\mathbf{Y} \sim \mathcal{N}_p(\mathbf{m}, \mathbf{C})$ be a random vector and $\mathbf{B}$ be a given non-singular $p \times p$ matrix. Find the distribution of the random vector $\mathbf{X} = \mathbf{BY}$.

3. (*3 point*) Let $(Y, X_1, \ldots, X_p)^T \sim \mathcal{N}_{p+1}(\mathbf{0}, \mathbf{C})$. Prove that the regression plane, i.e., $\mathbb{E}(Y|X_1, \ldots, X_p)$ also solves the

$$\operatorname{Corr}(Y, t(X_1, \ldots, X_p)) \to \max. \qquad \text{(over linear functions } t)$$

problem! This maximum is the maximum correlation between Y and X_i's.

4. (*2 point*) Let the n-dimensional random vector $\mathbf{X}$ follow a so-called symmetric multivariate normal distribution: its components are identically distributed, and the correlation between any two of them is the same number ρ. Find the spectral decomposition of the correlation matrix (suppose that $|\rho| < 1$).

5. (*3 point*) Let $\mathbf{X} \sim \mathcal{N}_2(\mathbf{m}, \mathbf{C})$ be a random vector. Find the density function of $\mathbf{X}$ by means of the entries of $\mathbf{m}$ and $\mathbf{C}$. Determine the distribution of the components, further the conditional distribution of the components conditioned on the given value of the other component. Find the correlation coefficient of the two components and the distribution of an appropriate linear combination of them ($aX_1 + bX_2 + c$). Give the linear transformation that takes $\mathbf{X}$ into a standard normal vector $\mathbf{Y} \sim \mathcal{N}_2(\mathbf{0}, \mathbf{I})$.

6. (*4 points*) Let $(X_1, Y_1)^T$ and $(X_2, Y_2)^T$ be i.i.d. 2-dimensional normally distributed random vectors with expectation $\mathbf{0}$ and covariance matrix

$$\mathbf{C} = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}, \qquad (|\rho| < 1).$$

Define

$$X = \begin{cases} 1, & \text{if } X_1 < X_2, \\ -1, & \text{otherwise} \end{cases} \qquad Y = \begin{cases} 1, & \text{if } Y_1 < Y_2, \\ -1, & \text{otherwise.} \end{cases}$$

Prove that $\operatorname{Cov}(X, Y) = (2/\pi)\sin^{-1}(\rho) = (2/\pi) arcsin(\rho)$!

7. (*3 points*) Let $X_1, \ldots, X_n$ be i.i.d. sample from the $\mathcal{N}(\mu, \sigma^2)$-distribution. We know that the statistic $T(X_1, \ldots, X_n) = (\bar{X}, S_n^{*2})$ provides an unbiased estimate for the parameter vector $\underline{\theta} = (\mu, \sigma^2)$. Prove that the covariance matrix of T does not reach the information bound. (For this, find the information matrix of Fisher!)

8. (*2 point*) We measured the hight and body weight of 100 young men. The average hight was 172.6 cm, while the average weight was 72.3 kg. Suppose that the hight and weight follow 2-dimensional normal distribution with covariance matrix

$$\mathbf{C} = \begin{pmatrix} 40 & 50 \\ 50 & 90 \end{pmatrix}.$$

   Test the zero hipothesis that the expectation of the hight is 170 cm, while that of the weight is 70 kg (choose 95% for the level of confidence)!

9. (*2 point*) We also measured the hight and body weight of 120 young women. The average hight was 165.3 cm, while the average weight was 65.4 kg. Suppose that the hight and weight follow 2-dimensional normal distribution with the above covariance matrix $\mathbf{C}$. Test the zero hipothesis that the hight and weight of young men and women do not differ significantly (choose 95% for the level of confidence)!

10. (*2 point*) Let $\mathbf{W} \sim \mathcal{W}_p(n, \mathbf{C})$ be a random Wishart-matrix and $\mathbf{B}$ be a given $p \times p$ non-singular matrix. Prove that $\mathbf{BWB}^T \sim \mathcal{W}_p(n, \mathbf{BCB}^T)$.