Homework 6 solutions: Gaussian random vectors

EE 325: Probability and Random Processes, Autumn 2019
Instructor: Animesh Kumar, EE, IIT Bombay

Instructions: These problems are a part of the syllabus for the final exam. They are not to be submitted. *If you have queries, then meet the instructor or the TA during office hours.*

1. Let X be a sign random variable with the distribution:

$$\mathbb{P}(X = 1) = 0.5; \quad \mathbb{P}(X = -1) = 0.5. \tag{1}$$

Let

$$\vec{X}(i) := \frac{1}{\sqrt{n}} [X(i, 1), \ldots, X(i, n)]^T \quad \text{for } i = 1, 2, 3. \tag{2}$$

The elements $\sigma(i, j), 1 \leq i \leq 3, 1 \leq j \leq n$ are i.i.d. with the same distribution as X. Let $Y_1, Y_2, Y_3, \ldots$ be an i.i.d. sequence of random variables with mean zero and variance 1. Let $\vec{Y} = (Y_1, \ldots, Y_n)^T$ be a random vector. Answer the following:

(a) What will be the distribution of $Z_1 := \vec{Y}^T\vec{X}(1)$ as n becomes large?

(b) What will be the distribution of the vector $[Z_1, Z_2, Z_3]^T$ as n becomes large? Here $Z_i = \vec{Y}^T\vec{X}(i)$.

Solution:

(a) Consider the random variable $Y_1 = \vec{X}^T\vec{\sigma}_1 = \frac{1}{\sqrt{n}} [\sigma_{11}X_1 + \sigma_{12}X_2 + \ldots + \sigma_{1n}X_n]$. Let $Z_1 = \sigma_{11}X_1, Z_2 = \sigma_{12}X_2, \ldots, Z_n = \sigma_{1n}X_n$. The variables $Z_1, Z_2, \ldots, Z_n$ are i.i.d., since $\sigma_{1,i}$ and X_i for $i \in \{1, 2, \ldots, n\}$ are chosen from separate iid distributions. By the *central limit theorem*,

$$\frac{\sum_{i=1}^n Z_i - n\mathbb{E}(Z_1)}{\sqrt{n\text{var}(Z_1)}} \xrightarrow{d} \mathcal{N}(0, 1). \tag{3}$$

It is noted that $\mathbb{E}(Z_1) = \mathbb{E}(\sigma_{11}X_1) = \mathbb{E}(\sigma_{11})\mathbb{E}(X_1) = 0$ and $\text{var}(Z_1) = \mathbb{E}(\sigma_{11}^2X_1^2) = \mathbb{E}(X_1^2) = 1$. From (3),

$$\frac{\sum_{i=1}^n Z_i}{\sqrt{n}} = \frac{\sigma_{11}X_1 + \sigma_{12}X_2 + \ldots + \sigma_{1n}X_n}{\sqrt{n}} = Y_1 \xrightarrow{d} \mathcal{N}(0, 1).$$

That is $Y_1 \sim \mathcal{N}(0, 1)$ as $n \to \infty$.

(b) To keep the notation simple, vector notation will be used. First observe that $\vec{\sigma}_1$ and $\vec{\sigma}_2$ have length 1, i.e., $\|\vec{\sigma}_i\|_2^2 = 1$ for $i = 1, 2, 3$. From the previous part, we know that Y_1 and Y_2 individually converge to a Gaussian random variable with mean zero and variance one. Next, note that conditioned on $\vec{\sigma}_1$ and $\vec{\sigma}_2$,

$$\begin{aligned} \mathbb{E}(Y_1Y_2|\vec{\sigma}_1, \vec{\sigma}_2) &= \mathbb{E}(\sigma_1^T\vec{X}\vec{X}^T\vec{\sigma}_2|\sigma_1, \sigma_2) && (4)\\ &= \vec{\sigma}_1^T\mathbb{E}(\vec{X}\vec{X}^T|\vec{\sigma}_1, \vec{\sigma}_2)\vec{\sigma}_2 && (5)\\ &= \vec{\sigma}_1^T\vec{\sigma}_2 && (6) \end{aligned}$$

where the last step uses independence of $\vec{\sigma}_1, \vec{\sigma}_2$ from $\vec{X}$, and the covariance matrix of $\vec{X}$ is identity. With high probability, as n becomes large, $\vec{\sigma}_1$ and $\vec{\sigma}_2$ will have a near-zero inner product. This follows by weak law of large numbers, for instance. (The details for the weak law convergence must be worked out by you.) So, conditioned on $\vec{\sigma}_1, \vec{\sigma}_2, \vec{\sigma}_3$, we would expect Y_1, Y_2, Y_3 to converge to a normalized white Gaussian random vector.
Note that this is a stronger claim than just showing $\mathbb{E}(Y_1Y_2) = \mathbb{E}(Y_2Y_3) = \mathbb{E}(Y_1Y_3) = 0$.

2. Assume that $m < n$. Show that if $\vec{Z}$ is an n-dimensional jointly Gaussian random vector and B is a rectangular $m \times n$ matrix, then $B\vec{Z}$ is jointly Gaussian.

Solution:
Let, $B = [b_{ij}]$, and $\vec{Z} = [Z_1\ Z_2\ \dots Z_n]^T$. Then,

$$\vec{X} = B\vec{Z} = \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \dots & b_{mn} \end{bmatrix} \begin{bmatrix} Z_1 \\ Z_2 \\ \vdots \\ Z_n \end{bmatrix} \tag{7}$$

It can be seen that each component of $\vec{X}$ is a linear combination of the components of $\vec{Z}$. Thus when we take linear combination of components of $\vec{X}$, we get linear combination of components of $\vec{Z}$. Since any linear combination of $\vec{Z}$ is a Gaussian random variable, therefore, any linear combination of $\vec{X}$ is also a Gaussian random variable; hence, $\vec{X} = B\vec{Z}$ is jointly Gaussian.

3. If two jointly Gaussian random vectors $\vec{X}$ and $\vec{Y}$ are uncorrelated, show that they are also independent. (BONUS) Will this be true if $\vec{X}$ and $\vec{Y}$ are not jointly Gaussian but marginally Gaussian?

Solution:
Given, $\vec{X}$ and $\vec{Y}$ are jointly Gaussian and are uncorrelated. Therefore, $\mathbb{E}\left((\vec{X} - \mu_{\vec{X}})(\vec{Y} - \mu_{\vec{Y}})^T\right) = 0$.

The covariance matrix of $\begin{bmatrix} \vec{X} \\ \vec{Y} \end{bmatrix}$ is given as,

$$\begin{aligned} K_{\vec{X}\vec{Y}} &= \mathbb{E}\left(\begin{pmatrix} \vec{X} - \mu_{\vec{X}} \\ \vec{Y} - \mu_{\vec{Y}} \end{pmatrix} \left((\vec{X} - \mu_{\vec{X}})^T \mid (\vec{Y} - \mu_{\vec{Y}})^T\right)\right) \\ &= \begin{pmatrix} \mathbb{E}((\vec{X} - \mu_{\vec{X}})(\vec{X} - \mu_{\vec{X}})^T) & 0 \\ 0 & \mathbb{E}((\vec{Y} - \mu_{\vec{Y}})(\vec{Y} - \mu_{\vec{Y}})^T) \end{pmatrix} = \begin{pmatrix} K_{\vec{X}} & 0 \\ 0 & K_{\vec{Y}} \end{pmatrix} \end{aligned}$$

Thus, $\det(K_{\vec{X}\vec{Y}}) = \det(K_{\vec{X}})\det(K_{\vec{Y}})$. If $\vec{X}$ and $\vec{Y}$ are of length n and m,the joint distribution of $\begin{bmatrix} \vec{X} \\ \vec{Y} \end{bmatrix}$ is given by,

$$\begin{aligned} f_{\vec{X}\vec{Y}}(\vec{x}, \vec{y}) &= \frac{1}{(2\pi)^{(n+m)/2}|K_{\vec{X}\vec{Y}}|^{\frac{1}{2}}} \exp\left(-\frac{\begin{pmatrix} \vec{x} - \mu_{\vec{X}} \\ \vec{y} - \mu_{\vec{Y}} \end{pmatrix}^T K_{\vec{X}\vec{Y}}^{-1} \begin{pmatrix} \vec{x} - \mu_{\vec{X}} \\ \vec{y} - \mu_{\vec{Y}} \end{pmatrix}}{2}\right) \\ &= \frac{1}{(2\pi)^{(m+n)/2}|K_{\vec{X}}|^{\frac{1}{2}}|K_{\vec{Y}}|^{\frac{1}{2}}} \exp\left(-\frac{(\vec{x} - \mu_{\vec{X}})^T K_{\vec{X}}^{-1}(\vec{x} - \mu_{\vec{X}})}{2}\right) \exp\left(-\frac{(\vec{y} - \mu_{\vec{Y}})^T K_{\vec{Y}}^{-1}(\vec{y} - \mu_{\vec{Y}})}{2}\right) \\ &= f_{\vec{X}}(\vec{x}) f_{\vec{Y}}(\vec{y}) \end{aligned}$$

Thus, $\vec{X}$ and $\vec{Y}$ are independent.

If $\vec{X}$ and $\vec{Y}$ were not jointly Gaussian, then being uncorrelated need not imply independence. Consider two r.v.'s, X and Z which are independent. X is zero mean, unit variance Gaussian r.v, and Z takes values $+1$ and -1 with equal probability. And let $Y = Z|X|$ (verify that Z is $\mathcal{N}(0, 1)$). Clearly, X and Y are not independent and it can be observed by writing down the joint pdf of (X, Y). Now,

$$\begin{aligned} \mathbb{E}(XY) &= \mathbb{E}(Z|X|X) \\ &= \mathbb{E}\left(Z|X|^2 \mathrm{sgn}(X))\right) \\ &= \mathbb{E}(Z)\mathbb{E}\left(|X|^2 \mathrm{sgn}(X)\right) \qquad \text{[since } X \text{ and } Z \text{ are independent]} \\ &= 0 \times \mathbb{E}\left(|X|^2 \mathrm{sgn}(X)\right) = 0. \end{aligned}$$

Thus X and Y are marginally Gaussian and uncorrelated, but they are dependent.

4. Let $U^T = (\vec{X}^T, \vec{Y}^T)$ be a jointly Gaussian random vector of size $(n+m)$. Show that if $K_{\vec{U}}$ is non-singular, then both $K_{\vec{X}}$ and $K_{\vec{Y}}$ are non-singular. Further, show that if K_U is non-singular and if $K_U^{-1} = \begin{bmatrix} B & C \\ C^T & D \end{bmatrix}$, then B and D are also non-singular and positive definite.

Solution:

(a) The covariance matrix $K_{\vec{U}}$ can be expressed in terms of covariance matrix of $\vec{X}$ and $\vec{Y}$ in the following way,

$$\begin{aligned} K_{\vec{U}} &= \mathbb{E}\left(\begin{pmatrix} \vec{X} \\ \vec{Y} \end{pmatrix} \begin{pmatrix} \vec{X}^T & \vec{Y}^T \end{pmatrix}\right) \\ &= \begin{pmatrix} K_{\vec{X}} & K_{\vec{X}\vec{Y}} \\ K_{\vec{Y}\vec{X}} & K_{\vec{Y}} \end{pmatrix} \end{aligned}$$

Since $K_{\vec{U}}$ is non-singular it is positive definite, $\vec{x}^T K_{\vec{U}} \vec{x} > 0$ for all $\vec{x}$. Now if $\vec{x} = \begin{pmatrix} \vec{a} \\ 0 \end{pmatrix}$, then

$$\begin{pmatrix} \vec{a} \\ 0 \end{pmatrix}^T K_{\vec{U}} \begin{pmatrix} \vec{a} \\ 0 \end{pmatrix} = \vec{a}^T K_{\vec{X}} \vec{a} > 0.$$

Thus, $K_{\vec{X}}$ is positive definite and hence non-singular. Similarly we can take $\vec{x} = \begin{pmatrix} 0 \\ \vec{a} \end{pmatrix}$ and show that $K_{\vec{Y}}$ is positive definite. Also we know that positive definite matices are non-singular.[1]

(b) In this part, one has to show that the sub-matrix B and D are positive definite. Since $K_{\vec{U}}$ is positive definite it can be expressed as $K_{\vec{U}} = Q\Lambda Q^T$, where $\Lambda = \mathrm{diag}\,[\lambda_1, \lambda_2, \ldots, \lambda_{m+n}]$ and $\lambda_i > 0$. Thus $K_{\vec{U}}^{-1} = Q\Lambda^{-1}Q^T$ is also positive definite. Thus by the argument in part (a), B and D are positive definite and hence non-singular.

5. We have seen earlier that if $X \sim \mathcal{N}(0, \sigma_X^2)$ and $Y \sim \mathcal{N}(0, \sigma_Y^2)$ are independent Gaussian random variables, then $X + Y$ is a Gaussian random variable as well. Using induction, show that any linear combination of the components of an IID normalized Gaussian random vector $\vec{W} \sim \mathcal{N}(\vec{0}, I_n)$ is also a Gaussian random variable. (This exercise confirms that $\vec{W}$ is jointly Gaussian.)

Solution:

Since $X \sim \mathcal{N}(0, \sigma_X^2)$ and $Y \sim \mathcal{N}(0, \sigma_Y^2)$, $X + Y \sim \mathcal{N}(0, \sigma_X^2 + \sigma_Y^2)$

Let $S \subset \{1, 2, 3 \ldots n\}$

for any $\{a_1, a_2 \ldots a_n\} \in \mathbb{R}^n$ define .

$$Z_S = \sum_{i \in S} a_i W_i$$

where W_i are i.i.d $\sim \mathcal{N}(0,1)$ Gaussian r.vs MGF of W_i is $g_{W_i}(t) = [\exp(t^2/2)]$

MGF of Z_S is

$$\begin{aligned} g_{Z_S}(t) &= \mathbb{E}[\exp(tZ_S)] \\ &= \mathbb{E}[\exp(\sum_{i \in S} a_i t W_i)] \\ &= \Pi_{i \in S} \mathbb{E}[\exp(a_i t W_i))] \\ &= \Pi_{i \in S} \exp((a_i t)^2/2) \\ &= \exp(\frac{t^2}{2} \sum_{i \in S} a_i^2) \end{aligned}$$

[1]See properties of covariance matrices in the notes. Also try to prove the property yourself.

$$\text{We got } Z_S \sim \mathcal{N}(0, \sum_{i \in S} a_i^2)$$

Hence any linear combination of the components of an IID normalized Gaussian random vector is also a Gaussian random variable

6. Let X and Y be zero-mean jointly Gaussian random variables with $\mathbb{E}(X^2) = \sigma_X^2$, $\mathbb{E}(Y^2) = \sigma_Y^2$, and $\mathbb{E}(XY) = \rho\sigma_X\sigma_Y$.

   (a) Find the conditional probability density function $f_{X|Y}(x|y)$.

   (b) Let $V = Y^3$. Find the conditional probability density function $f_{X|V}(x|v)$. (Hint: think carefully before calculations.)

   (c) Let $Z = Y^2$. Find the conditional probability density function $f_{X|Z}(x|z)$. (Hint: first understand why this is more difficult than (b).)

   Solution:

   (a) Since X, Y are zero-mean and jointly Gaussian, therefore we know that $X = aY + V$, where a is a constant V is a zero mean Gaussian random variable, and Y, V are independent. Further,

   $$a = K_{XY}K_Y^{-1} = \rho\sigma_X\sigma_Y/\sigma_Y^2 = \rho\sigma_X/\sigma_Y.$$

   And,

   $$\sigma_V^2 = K_X - K_{XY}K_Y^{-1}K_{YX} = \sigma_X^2 - \rho\sigma_X\sigma_Y \cdot \sigma_Y^{-2} \cdot \rho\sigma_X\sigma_Y = (1-\rho^2)\sigma_X^2. \tag{8}$$

   Finally, $X|Y = y \sim \mathcal{N}(ay, \sigma_V^2)$. Therefore,

   $$f_{X|Y}(x|y) = \frac{1}{\sqrt{2\pi(1-\rho^2)\sigma_X^2}} \exp\left(-\frac{(x - \rho\sigma_X y/\sigma_Y)^2}{2(1-\rho^2)\sigma_X^2}\right).$$

   (b) Note that $V = Y^3$ is one to one function of Y. Thus, $V = v$ is equivalent to $Y = v^{1/3}$. Therefore, the pdf is given by,

   $$f_{X|V}(x|v) = \frac{1}{\sqrt{2\pi(1-\rho^2)\sigma_X^2}} \exp\left(-\frac{(x - \rho\sigma_X v^{1/3}/\sigma_Y)^2}{2(1-\rho^2)\sigma_X^2}\right).$$

   (c) Now $V = Y^2$. Thus, $V = v$ is equivalent to $Y = \sqrt{v}$ or $Y = -\sqrt{v}$. Thus, the pdf is given by,

   $$f_{X|V}(x|v) = \frac{1}{\sqrt{2\pi(1-\rho^2)\sigma_X^2}} \left[\exp\left(-\frac{(x - \rho\sigma_X\sqrt{v}/\sigma_Y)^2}{2(1-\rho^2)\sigma_X^2}\right) + \exp\left(-\frac{(x + \rho\sigma_X\sqrt{v}/\sigma_Y)^2}{2(1-\rho^2)\sigma_X^2}\right)\right].$$

7. Let X and Y be zero-mean and jointly Gaussian random variables with variances σ_X^2, σ_Y^2 and covariance $\rho\sigma_X\sigma_Y$. Find a 2×2 transformation matrix A such that $\vec{V} = A[X, Y]^T$ has independent components V_1 and V_2.

   Solution: Let K be the covariance matrix of $\vec{Z} = (X, Y)^T$, i.e.,

   $$K = \begin{bmatrix} \sigma_X^2 & \rho\sigma_X\sigma_Y \\ \rho\sigma_X\sigma_Y & \sigma_Y^2 \end{bmatrix}. \tag{9}$$

   Then we know that $Q^T\vec{Z}$ has independent components. Note that $(Q\Lambda^{-1/2})$ will be used as the transformation matrix to make $\vec{Z}$ white (or i.i.d.). Right now we just have to make $\vec{V}$ as *independent*.

   Since K is a covariance matrix, let $K = Q\Lambda Q^T$ be its spectral representation. Then $\vec{V} = Q^T\vec{Z}$ is the desired answer.

8. Let K be the following matrix,

$$K = \begin{bmatrix} 3 & 1 \\ 1 & 2 \end{bmatrix}. \tag{10}$$

(a) Find the eigenpairs of K.

(b) Find Q and Λ such that $K = Q\Lambda Q^T$, and $QQ^T = I_2$.

(c) Find the eigenpairs of K^n, where n is a natural number.

(d) What will be the eigenpairs of K^{-1}?

Solution:

(a) The characteristic equation for finding the eigenvalues of K is given by $\det(K - \lambda I) = 0$. That is,

$$\begin{vmatrix} 3-\lambda & 1 \\ 1 & 2-\lambda \end{vmatrix} = 0$$

$$\text{which results in} \quad \lambda_1 = \frac{5+\sqrt{5}}{2}, \lambda_2 = \frac{5+\sqrt{5}}{2}.$$

The corresponding eigenvectors can be obtained by examining the nullspace of $K - \lambda_1 I$ and $K - \lambda_2 I$ respectively. The corresponding eigenvectors are

$$\begin{bmatrix} 1 \\ \frac{\sqrt{5}-1}{2} \end{bmatrix} \text{ and } \begin{bmatrix} 1 \\ -\frac{\sqrt{5}+1}{2} \end{bmatrix}.$$

These vectors are orthogonal. Upon orthonormalization they become

$$\begin{bmatrix} 0.8507 \\ 0.5257 \end{bmatrix} \text{ and } \begin{bmatrix} 0.5257 \\ -0.8507 \end{bmatrix}.$$

(b) From spectral decomposition theory of positive definite matrices, we know that a choice for the matrix Q can be constructed from these orthonormal eigenvectors. So, $Q = \begin{bmatrix} 0.8507 & 0.5257 \\ 0.5257 & -0.8507 \end{bmatrix}$.

The corresponding matrix of eigenvalues is $\Lambda = \begin{bmatrix} \frac{5+\sqrt{5}}{2} & 0 \\ 0 & \frac{5+\sqrt{5}}{2} \end{bmatrix}$.

(c) Recall that $QQ^T = I_2$. From $K = Q\Lambda Q^T$ by recursive computation it can be shown that $K^n = Q\Lambda^n Q^T \Rightarrow$. Therefore, the eigenvectors of K^n and K are the same. The eigenvalues of K^n are λ_1^n and λ_2^n.

(d) Similar to the previous part, $K^{-1} = Q\Lambda^{-1}Q^T \Rightarrow$. This can be verified since $KQ\Lambda^{-1}Q^T = Q\Lambda Q^T Q\Lambda^{-1}Q^T = I_2$. As a result, (from this directly obtained spectral decomposition), the eigenvectors of K^{-1} and K are the same. The eigenvalues of K^{-1} are $\frac{1}{\lambda_1}$ and $\frac{1}{\lambda_2}$.