Homework 6: Gaussian random vectors

EE 325 (DD): Probability and Random Processes, Autumn 2016
Instructor: Animesh Kumar, EE, IIT Bombay

Set-A

1. 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,

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

   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.

2. 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.

3. 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 $\Lambda = \mathrm{diag}\left[\lambda_1, \lambda_2, \ldots, \lambda_{m+n}\right]$ 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.

Set-B

1. Consider the coupon collector problem. Let $p_1, p_2, \ldots, p_n$ be the (unequal) probabilities of drawing coupons $1, 2, \ldots, n$. Assume that $\sum_{1 \le i \le n} p_i = 1$ and $p_i > 0$ for all i. How many independent draws of the coupons will be necessary to ensure that all the coupons have been collected with high probability?

Solution: Let $\mathcal{E}_i(m)$ be the event that the coupon i is not collected after m random independent draws. Then,

$$\mathbb{P}(\mathcal{E}_i(m)) = (1 - p_i)^m. \tag{2}$$

By the union bound, we get

$$\mathbb{P}\left(\bigcup_{i=1}^{n} \mathcal{E}_i(m)\right) \le \sum_{i=1}^{n} (1 - p_i)^m. \tag{3}$$

Let $p := \min\{p_1, \ldots, p_n\}$. Then a sufficient condition is $m \ge \frac{1}{p} \log \frac{1}{p}$ to ensure that the above probability goes to zero.

Further, to ensure that the coupon which corresponds to p has been drawn, it is required that $m \ge \omega(p)$.

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

2. Let $c, |c| < 1$ be a constant which is observed via a single bit quantizer after random dithering. The observations are $Y_1, \ldots, Y_n$, where

$$\begin{aligned} Y_i &= 1, \text{ if } c + W_i \geq 0 \\ &= 0, \text{ if } c + W_i < 0. \end{aligned}$$

The random variables $W_1, \ldots, W_n$ are Unif$[-2, 2]$ random variables. Write down an estimate for c from $Y_1, \ldots, Y_n$ such that its mean-squared error decreases proportional to $1/n$. Let $\widehat{C}$ be your estimate. What do you expect the distribution of $\sqrt{n}(\widehat{C} - c)$ to be when n approaches infinity?

Solution: From the lecture, note that $\mathbb{P}(Y_i = 1) = \mathbb{P}(c + W \geq 0) = \mathbb{P}(W \geq -c) = \mathbb{P}(W \leq c) = F_W(c)$ by the symmetry in the distribution of W. For $c \in [-2, 2]$, the cdf of W is given by $F_W(c) = 0.5 + 0.25c$.

By the weak law of large numbers (WLLN) we know that

$$\frac{1}{n} Y_i \xrightarrow{\mathbb{P}} F_W(c). \tag{4}$$

So an estimate $\widehat{C}$ for c is given by

$$\widehat{C} := F_W^{-1}\left(\frac{Y_1 + \ldots + Y_n}{n}\right). \tag{5}$$

By the central limit theorem, we expect

$$\frac{1}{\sqrt{n}}(Y_i - F_W(c)) \xrightarrow{d} \mathcal{N}(0, F_W(c)(1 - F_W(c))) \tag{6}$$

where $F_W(c)(1 - F_W(c))$ is the variance of Y_i. Since $F_W^{-1}(c) = 4(c - 0.5)$, so the estimate $\widehat{C}$ is given by

$$\begin{aligned} \widehat{C} &= F_W^{-1}\left(\frac{Y_1 + \ldots + Y_n}{n}\right) & (7) \\ &= 4\left(\frac{Y_1 + \ldots + Y_n}{n}\right) - 2. & (8) \end{aligned}$$

Therefore, $\sqrt{n}(\widehat{C} - c)$ will be equal to $4\left(\frac{Y_1 - F_W(c) + \ldots + Y_n - F_W(c)}{\sqrt{n}}\right)$. This will converge (by CLT) to $\mathcal{N}(0, 16F_W(c)(1 - F_W(c)))$ distribution.

3. Let X and Z be IID normalized Gaussian random variables. Let $Y = |Z|\text{sign}(X)$, where $\text{sign}(x) = 1$ if $x \geq 0$ and $\text{sign}(x) = -1$ if $x < 0$. Show that Y is a Gaussian random variable, but (Y, Z) are not jointly Gaussian.

Solution: As given, $X \sim \mathcal{N}(0, 1)$ and $Z \sim \mathcal{N}(0, 1)$, and X and Z are independent. Since, $Y = |Z|\text{sign}(X)$,it is a continuous random variable. Using total probability rule, for $y \geq 0$

$$\begin{aligned} f_Y(y) &= f_Z(y)\mathbb{P}(X \geq 0) + f_Z(-y)\mathbb{P}(X < 0) \\ &= f_Z(y). \end{aligned}$$

Similarly,for $y < 0$

$$f_Y(y) = f_Z(y).$$

Hence $Y \sim \mathcal{N}(0, 1)$.

To see that (Z, Y) are not jointly Gaussian, we will show that their one particular linear combination may not Gaussian. Let $W = Y + Z$. Then, it is easy to see that $\mathbb{P}(W = 0) = \frac{1}{2}$. And so W is not Gaussian. Hence (Y, Z) are not jointly Gaussian.

4. 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

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

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

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

5. 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,

$$\sigma_V^2 = K_X - K_{XY} K_Y^{-1} K_{YX} = \sigma_X^2 - \rho\sigma_X\sigma_Y \cdot \sigma_Y^{-2} \cdot \rho\sigma_X\sigma_Y = (1 - \rho^2)\sigma_X^2. \tag{9}$$

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,

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

6. Let $\vec{Z} = A\vec{W}$ for some IID Gaussian random vector $\vec{W} \sim \mathcal{N}(\vec{0}, I_n)$. Assume that $\vec{U} = B\vec{Z}$. Show that $\vec{U}$ is a Gaussian random vector. You can assume that A and B are non-singular $(n \times n)$ matrices of real numbers.

Solution:

At a high level, any linear combination of $\vec{Z}$ will eventually be a linear combination of $\vec{W}$. The vector $\vec{W}$ is jointly Gaussian. Thus, all linear combinations of $\vec{Z}$ are Gaussian r.v. Hence, one can argue that $\vec{Z}$ is also jointly Gaussian.

To prove it formally, we will show that any linear combination of $\vec{Z}$ is a Gaussian r.v. We have,

$$\vec{b}^T\vec{Z} = \vec{b}^T A\vec{W} = (A^T\vec{b})^T\vec{W} = \vec{c}^T\vec{W}, \text{ with } \vec{c} = A^T\vec{b}.$$

Thus, $\vec{b}^T\vec{Z}$ is a Gaussian r.v. since $\vec{c}^T\vec{W}$ will be a Gaussian r.v. This completes the proof.

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.,

$$K = \begin{bmatrix} \sigma_X^2 & \rho\sigma_X\sigma_Y \\ \rho\sigma_X\sigma_Y & \sigma_Y^2 \end{bmatrix}. \tag{10}$$

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,

$$K = \begin{bmatrix} 3 & 1 \\ 1 & 2 \end{bmatrix}. \tag{11}$$

(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)Characteristic equation of K is $|K - \lambda I| = 0$

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

$$\Rightarrow \lambda_1 = \frac{5+\sqrt{5}}{2}, \lambda_2 = \frac{5+\sqrt{5}}{2}$$

Their corresponding eigenvectors are $\begin{bmatrix} 1 \\ \frac{\sqrt{5}-1}{2} \end{bmatrix}$ and $\begin{bmatrix} 1 \\ -\frac{\sqrt{5}+1}{2} \end{bmatrix}$ and after normalization they become $\begin{bmatrix} 0.8507 \\ 0.5257 \end{bmatrix}$ and $\begin{bmatrix} 0.5257 \\ -0.8507 \end{bmatrix}$.

b)So $Q = \begin{bmatrix} 0.8507 & 0.5257 \\ 0.5257 & -0.8507 \end{bmatrix}$ and $\Lambda = \begin{bmatrix} \frac{5+\sqrt{5}}{2}, & 0 \\ 0 & \frac{5+\sqrt{5}}{2} \end{bmatrix}$.

c)$K^n = Q\Lambda^n Q^T \Rightarrow$ eigen vectors remains same whereas eigen values are λ_1^n and λ_2^n.

d)$K^{-1} = Q\Lambda^{-1}Q^T \Rightarrow$ eigen vectors remains same whereas eigen values are $\frac{1}{\lambda_1}$ and $\frac{1}{\lambda_2}$.