Homework 2

Alessandro Bianchetti

November 25, 2021

Note: calculations of this work were cross-checked with the students Mila Racca and Kai Aidan Growcoot. This report is however independent and different from theirs.

1 Exercise 4

The purpose here is to prove that the characteristic function of a MVN is given by

$$\phi(\vec{k}) = \exp\left(-i\mu^T \cdot \vec{k} - \frac{1}{2}\vec{k}^T C^{-1}\vec{k}\right) \tag{1}$$

Let's first remind the definition of characteristic function:

$$\phi(\vec{k}) = \int d^n x \exp\left(-i\vec{k} \cdot \vec{x}\right) p(\vec{x}) \tag{2}$$

where $p(\vec{x})$ is whatever distribution. In our case it's the MVN function

$$\mathcal{N}(\vec{x}|\mu, C) = \frac{1}{(2\pi)^{N/2}\sqrt{\det C}} \exp\left[-\frac{1}{2}(\vec{x}-\mu)^T C^{-1}(\vec{x}-\vec{\mu})\right] \tag{3}$$

So let's plug this definition inside 2 and see what happens. We need to do this in two different ways: first, we're going to apply the square completion technique, thus solving the Fourier Transform. Then we're going to do it by rotating the inverse of the covariance matrix C^{-1} making it diagonal.

1.1 Method 1: square completion

For simplicity, from now on we refer to the normalization factor of the multivariate Gaussian as B.

$$\phi(\vec{k}) = \int d^n x \exp\left(-i\vec{k} \cdot \vec{x}\right) B \exp\left\{-\frac{1}{2}(\vec{x}-\vec{\mu})^T C^{-1}(\vec{x}-\vec{\mu})\right\} =$$
$$= \int d^N u B \exp\left[-\frac{1}{2}\vec{u}^T C^{-1}\vec{u} - i\vec{k}^T(\vec{u}+\vec{\mu})\right]$$

where we replaced $\vec{u} = \vec{x} - \vec{\mu}$. Let's play on the exponent a bit in order to build a new squared quantity at the exponent, this time involving the frequency $\vec{k}$. The formula of the square completion is given by

$$\vec{x}^T A\vec{x} + \vec{b} \cdot \vec{x} + c = (\vec{x} + \vec{m})^T A(\vec{x} + \vec{m}) + n \tag{4}$$

$$\text{where} \qquad \vec{m} = \frac{1}{2}A^{-1}\vec{b} \qquad n = c - \frac{1}{4}\vec{b}^T A^{-1}\vec{b} \tag{5}$$

So, in our case the exponent is

$$-\frac{1}{2}\vec{u}^T C^{-1}\vec{u} - i\vec{k} \cdot (\vec{u} + \vec{\mu}) = -\frac{1}{2}\left[\vec{u}^T C^{-1}\vec{u} + 2i\vec{k} \cdot \vec{u} + 2i\vec{k} \cdot \vec{\mu}\right].$$

Therefore, if we keep the $-\frac{1}{2}$ out for a minute, we can define the following:

$$\begin{cases} A = C^{-1} \\ \vec{b} = 2i\vec{k} \\ c = 2i\vec{k} \cdot \vec{\mu} \end{cases} \quad \text{and} \quad \begin{cases} \vec{m} = iC\vec{k} \\ n = 2i\vec{k} + \vec{k}^T C\vec{k} \end{cases}$$

Now we apply 4 and factor $-\frac{1}{2}$ in again, and the exponent will look like

$$-\frac{1}{2}\vec{u}^T C^{-1}\vec{u} - i\vec{k}\cdot\vec{u} - i\vec{k}\cdot\vec{\mu} == -\frac{1}{2}\left[(\vec{u} + iC\vec{k})^T C^{-1}(\vec{u} + iC\vec{k})\right] - \frac{1}{2}\vec{k}^T C\vec{k} - i\vec{k}\cdot\vec{\mu}$$

We now apply this to the integral we were working on:

$$\phi(\vec{k}) = \exp\left[-\frac{1}{2}\vec{k}^T C\vec{k} - i\vec{k} \cdot \vec{\mu}\right] \int d^N uB \exp\left[-\frac{1}{2}(\vec{u} + iC\vec{k})^T C^{-1}(\vec{u} + iC\vec{k})\right]$$

and we can again change variable $\vec{t} = \vec{u} + iC\vec{k}$, so that

$$\phi(\vec{k}) = \exp\left[-\frac{1}{2}\vec{k}^T C\vec{k} - i\vec{k} \cdot \vec{\mu}\right] \int d^N tB \exp\left[-\frac{1}{2}\vec{t}^T C^{-1}\vec{t}\right]$$

but in this last equation the integral is a perfectly normalised Gaussian integrated over infinity, meaning it corresponds to 1. Finally:

$$\phi(\vec{k}) = \exp\left[-\frac{1}{2}\vec{k}^T C\vec{k} - i\vec{k} \cdot \vec{\mu}\right]$$

Alternative solution It is perhaps possible to try out a different path. In fact, after the first substitution (introducing $\vec{u}$), we can immediately make the following change of variables, switching to $\vec{t}$:

$$\sqrt{C}\vec{t} = \vec{u}$$

where $\sqrt{C}$ is defined as the squared root of the matrix C, meaning a matrix such that $\sqrt{C}\sqrt{C} = C$. This means that, working on the exponent, we have

$$\vec{u}^T C\vec{u} = (\sqrt{C}\vec{t})^T C^{-1}\sqrt{C}\vec{t} = \vec{t} \cdot \vec{t} = t^2.$$

Such a change of coordinates will also introduce the determinant of the transformation matrix $\sqrt{C}$ in the integral. Moreover, $\det \sqrt{C} = \sqrt{\det C}$ due to Binet's Theorem ($\det(A \cdot B) = \det A \det B$ for any couple of matrices A and B).

$$\phi(\vec{k}) = \int d^N u B \exp\left\{-\frac{1}{2}\vec{u}^T C^{-1}\vec{u} - i\vec{k}^T(\vec{u}+\vec{\mu})\right\} =$$
$$= B\sqrt{\det C}\int d^N t e^{-i\vec{k}\cdot\vec{\mu}} e^{-i\vec{k}\sqrt{C}\vec{t}} e^{-\frac{1}{2}t^2} = B\sqrt{\det C} e^{-i\vec{k}\cdot\vec{\mu}} \int d^N t e^{-\frac{1}{2}[t^2+2i\vec{k}\sqrt{C}\vec{t}]}$$

Here we perform a square completion by adding and subtracting the term $\frac{1}{2}\vec{k}^T C\vec{k}$ to the exponent (or applying the formula for square completion 4 if you prefer).

$$\phi(\vec{k}) = B\sqrt{\det C} e^{-i\vec{k}\cdot\vec{\mu}} e^{-\frac{1}{2}\vec{k}^T C\vec{k}} \int d^n t e^{-\frac{1}{2}[\vec{t}+i\sqrt{C}\vec{k}]^2}$$

A final substitution $\vec{y} = \vec{t} + i\sqrt{C}\vec{k}$ leads to

$$\phi(\vec{k}) = B\sqrt{\det C} e^{-i\vec{k}\cdot\vec{\mu}} e^{-\frac{1}{2}\vec{k}^T C\vec{k}} \int d^n y e^{-\frac{1}{2}y^2}$$

But the integral in the latter is equal to $(2\pi)^{N/2}$, and therefore it cancels out with the term $B\sqrt{\det C} = (2\pi)^{-N/2}$. So we're left with

$$\phi(\vec{k}) = e^{-i\vec{k}\cdot\vec{\mu}} e^{-\frac{1}{2}\vec{k}^T C\vec{k}}$$

which is again the expected result.

1.2 Method 2: diagonalization

This time we try introducing a rotation, governed by a rotation matrix.

$$R(\theta) = \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \tag{6}$$

With a change of coordinates imposed by a rotation matrix there is no need to worry about introducing the determinant of the Jacobian as we did in the previous section, since $\det R(\theta) = 1\ \forall\theta$. We choose R so to make the inverse of the covariance matrix diagonal $R^T C^{-1} R = \Lambda$.

So let's start again from the point where we've already made the shift of coordinates, introducing $\vec{u} = \vec{x} - \vec{\mu}$:

$$\phi(\vec{k}) = \int d^N u B \exp\left[-\frac{1}{2}\vec{u}^T C^{-1}\vec{u} - i\vec{k}^T(\vec{u}+\vec{\mu})\right]$$

Now we set $\vec{u} = R\vec{y}$

$$\phi(\vec{k}) = \int d^N y B \exp\left[-\frac{1}{2}\vec{y}^T R^T C^{-1} R\vec{y} - i\vec{k}^T (R\vec{y} + \vec{\mu})\right] =$$
$$= \int d^N y B \exp\left[-\frac{1}{2}\vec{y}^T \Lambda \vec{y} - i\vec{k}^T R\vec{y} - i\vec{k}^T \vec{\mu}\right]$$

Now we better define $\vec{z} = \vec{k}^T R$, so that

$$\phi(\vec{k}) = \int d^N y B \exp\left[\sum_i (\frac{1}{2}\lambda_i y_i^2 - h_i y_i) - i\vec{k}^T \vec{\mu}\right]$$
$$= B \exp -i\vec{k}\cdot\vec{\mu} \int d^n y \Pi_i \exp\left[-\frac{1}{2}\lambda_i y_i^2 - z_i y_i\right]$$

And we proceed again by completing the square:

$$\phi(\vec{k}) = B \exp -i\vec{k}\cdot\vec{\mu}\Pi_i \int d^n y \exp\left[-\left(\sqrt{\frac{\lambda_i}{2}} y_i + \frac{z_i}{\sqrt{2\lambda_i}}\right)^2\right] \exp\frac{z_i^2}{2\lambda_i} =$$
$$= B \exp\left[-i\vec{k}\cdot\vec{\mu}\right] \Pi_i \sqrt{\frac{2}{\lambda_i}} (2\pi)^{1/2} \exp\frac{h_i^2}{2\lambda_i}$$

Remember $B = (2\pi)^{-N/2} (\det C)^{-1/2}$, and $\det C^{-1/2} = \Pi_i \sqrt{\lambda_i}$ using eigenvalues.

$$\phi(\vec{k}) = (2\pi)^{-N/2} \Pi_i \sqrt{\lambda_i} \exp -i\vec{k}\cdot\vec{\mu} (2\pi)^{N/2} \Pi_i \sqrt{\lambda_i} \sqrt{\frac{2}{\lambda_i}} \exp\frac{z_i^2}{2\lambda_i}$$
$$= \exp\left[-i\vec{k}\cdot\vec{\mu}\right] \Pi_i \exp\left[\frac{z_i^2}{2\lambda_i}\right] = \exp\left[-i\vec{k}\cdot\vec{\mu}\right] \exp\left[-\frac{\vec{k}^T R \Lambda^{-1} R^T \vec{k}}{2}\right]$$

but $R\Lambda^{-1}R^T = C$, so finally

$$\phi(\vec{k}) = \exp\left[-i\vec{k}\cdot\vec{\mu} - \frac{1}{2}\vec{k}^T C \vec{k}\right]$$

which is the result we wanted to achieve.

2 Exercise 5

The characteristic function also has the property of generating momenta, if we derive on the generic components $-ik_\alpha$. We want to test such property

$$E[x_\alpha^{n_\alpha} \cdots x_\beta^{n_\beta}] = \frac{\partial^{n_\alpha + \cdots + n_\beta} \phi(\vec{k})}{\partial^{n_\alpha}(-ik_\alpha) \cdots \partial^{n_\beta}(-ik_\beta)}\Big|_{\vec{k}=0} \tag{7}$$

for two simple cases: the mean of a single component x_α and the covariance of two components x_α and x_β.

2.1 Mean

For the first step, we try to compute the derivative in the RHS of 7 for $n_\alpha = 1$ and $n_\gamma = 0 \ \forall \gamma \neq \alpha$. Therefore

$$\frac{\partial \phi(\vec{k})}{\partial(-ik_\alpha)}\Big|_{\vec{k}=0} = i\frac{\partial \phi(\vec{k})}{\partial k_\alpha}\Big|_{\vec{k}=0} = i\left[T_\alpha \phi(\vec{k})\right]\Big|_{\vec{k}=0} \tag{8}$$

where T_α is the internal derivative, meaning the derivative of the exponent with respect to k_α.

$$T_\alpha = \frac{\partial}{\partial k_\alpha}\left[-i\sum_i k_i\mu_i - \frac{1}{2}\sum_i\sum_j k_i C_{ij} k_j\right] = -i\mu_\alpha \sum_j C_{\alpha j} k_j$$

Now let's evaluate the expression at $\vec{k} = 0$. Actually $\phi(\vec{k})|_{\vec{k}=0} = 1$ and $T_\alpha|_{\vec{k}=0} = -i\mu_\alpha$. So 8 becomes

$$\frac{\partial \phi(\vec{k})}{\partial(-ik_\alpha)}\Big|_{\vec{k}=0} = i \cdot (-i\mu_\alpha) = \mu_\alpha$$

but that is exactly $E[x_\alpha]$, so we correctly generated a first-order momentum of expectation, proving the generating equation right.

2.2 Covariance

The model we have to follow is the same as before, but this time we have to perform a second derivative, over the coordinates k_α and k_β. Therefore we're setting $n_\alpha = 1$, $n_\beta = 1$ and $n_\gamma = 0 \ \forall \gamma \neq \alpha, \beta$.

$$\frac{\partial^2 \phi(\vec{k})}{(-i)^2 \partial k_\alpha \partial k_\beta}\Big|_{\vec{k}=0} = -\frac{\partial \phi(\vec{k})}{\partial k_\alpha \partial k_\beta}\Big|_{\vec{k}=0} = -\frac{\partial}{\partial k_\alpha}\left[\frac{\partial \phi(\vec{k})}{\partial k_\beta}\right]\Big|_{\vec{k}=0} =$$
$$= -\frac{\partial}{\partial k_\alpha}\left[T_\beta \phi(\vec{k})\right]\Big|_{\vec{k}=0} = \left[-\frac{\partial T_\beta}{\partial k_\alpha} - T_\alpha T_\beta \phi(\vec{k})\right]\Big|_{\vec{k}=0}$$

where

$$\frac{\partial T_\beta}{\partial k_\alpha} = \frac{\partial}{\partial k_\alpha}\left[-i\mu_\beta - \sum_j C_{\beta j} k_j\right] = -C_{\alpha\beta}.$$

This means that

$$\frac{\partial^2 \phi(\vec{k})}{(-i)^2 \partial k_\alpha \partial k_\beta}\Big|_{\vec{k}=0} = [C_{\alpha\beta} - T_\alpha T_\beta]\phi(\vec{k})\Big|_{\vec{k}=0} = [C_{\alpha\beta} - (-i\mu_\alpha)(-i\mu_\beta)] = C_{\alpha\beta} + \mu_\alpha\mu_\beta$$

This is the result we wanted to reach. In fact

$$Cov[x_\alpha x_\beta] = E[x_\alpha x_\beta] - E[x_\alpha]E[x_\beta] = C_{\alpha\beta} + \mu_\alpha\mu_\beta - \mu_\alpha\mu_\beta = C_{\alpha\beta} \tag{9}$$

which is the correct result.

3 Exercise 6

The idea here is proving that the characteristic function we derived is again Gaussian-shaped. So we will just integrate $\phi(\vec{k})$ over the frequency space and hope we get a simple constant.

$$\int d^n k \exp\left[-i\vec{k}\cdot\vec{\mu} - \frac{1}{2}\vec{k}^T C\vec{k}\right]$$

Again we complete the square of the exponent by applying 4. This time

$$\begin{cases} \vec{m} = -iC^{-1}\vec{\mu} \\ n = \mu^T C^{-1}\mu \end{cases}$$

and the exponent becomes

$$-i\vec{k}\cdot\vec{\mu} - \frac{1}{2}\vec{k}^T C\vec{k} = (\vec{k} - iC^{-1}\vec{\mu})^T C(\vec{k} - iC^{-1}\vec{\mu}) + \mu^T C^{-1}\vec{\mu}$$

Therefore, the integral becomes

$$\int d^n k \exp\left[-i\vec{k}\cdot\vec{\mu} - \frac{1}{2}\vec{k}^T C\vec{k}\right] = \int d^n k \exp\left[-\frac{1}{2}\left(\vec{k}^T C\vec{k} + 2i\vec{k}\cdot\vec{\mu}\right)\right] =$$
$$= \exp\left[-\frac{1}{2}\left(\mu^T C^{-1}\mu\right)\right]\int d^N k \exp\left[-\frac{1}{2}(\vec{k} - iC^{-1}\vec{\mu})^T C(\vec{k} - iC^{-1}\vec{\mu})\right]$$

Now inside the integral there is clearly a Gaussian with mean $iC^{-1}\mu$. So the integral is equal to $(2\pi)^{N/2}(\det C)^{-\frac{1}{2}}$.

$$\int d^n k \exp\left[-i\vec{k}\cdot\vec{\mu} - \frac{1}{2}\vec{k}^T C\vec{k}\right] = \exp\left[-\frac{1}{2}\mu^T C^{-1}\mu\right](2\pi)^{N/2}(\det C)^{-\frac{1}{2}}$$

which is simply a constant quantity. We have thus proved that the characteristic function is an unnormalised Gaussian function.