

AstroStatistics and Cosmology Homework

Jacopo Tissino

2020-10-22

Contents

1	Exercise 4	1
2	Exercise 5	4
3	Exercise 6	5

Exercises 1–3 and 7 are in Jupyter notebooks in the folder `astrostat_homework`. They can be most easily accessed through the following links:

1. https://nbviewer.jupyter.org/github/jacopok/notes/blob/master/ap_third_semester/astrostat_homework/exercises_123.ipynb
2. https://nbviewer.jupyter.org/github/jacopok/notes/blob/master/ap_third_semester/astrostat_homework/exercise_7.ipynb.

1 Exercise 4

After being given a probability distribution $\mathbb{P}(x)$, we define the *characteristic function* ϕ as its Fourier transform, which can also be expressed as the expectation value of $\exp(-i\vec{k} \cdot \vec{x})$:

$$\phi(\vec{k}) = \int d^n x \exp(-i\vec{k} \cdot \vec{x}) \mathbb{P}(x) = \mathbb{E} \left[\exp(-i\vec{k} \cdot \vec{x}) \right]. \quad (1.1)$$

Claim 1.1. *A multivariate normal distribution*

$$\mathcal{N}(\vec{x}|\vec{\mu}, C) = \frac{1}{(2\pi)^{n/2} \sqrt{\det C}} \exp\left(-\frac{1}{2} \vec{y}^\top C^{-1} \vec{y}\right) \Big|_{\vec{y}=\vec{x}-\vec{\mu}}, \quad (1.2)$$

has a characteristic function equal to

$$\phi(\vec{k}) = \exp\left(-i\vec{\mu} \cdot \vec{k} - \frac{1}{2} \vec{k}^\top C \vec{k}\right). \quad (1.3)$$

Proof: completing the square. The integral we need to compute is given, absorbing the normalization into a factor N , by

$$\phi(\vec{k}) = N \int d^n x \exp\left(-i\vec{k} \cdot \vec{x} - \frac{1}{2}\vec{y}^\top C^{-1}\vec{y}\right) \Big|_{\vec{y}=\vec{x}-\vec{\mu}}. \quad (1.4)$$

The only integrals we really know how to do are Gaussian ones, so we want to rewrite the argument of the exponential so that it is a quadratic form. The manipulation goes as follows, considering the opposite of the argument the exponential in order to have less minus signs and defining the symmetric matrix $V = C^{-1}$:

$$i\vec{k} \cdot \vec{x} + \frac{1}{2}\vec{y}^\top V\vec{y} = i\vec{k} \cdot \vec{x} + \frac{1}{2}\vec{x}^\top V\vec{x} - \vec{x}^\top V\vec{\mu} + \frac{1}{2}\vec{\mu}^\top V\vec{\mu} \quad (1.5)$$

$$= \frac{1}{2}\vec{x}^\top V\vec{x} + \vec{x}^\top (i\vec{k} - V\vec{\mu}) + \frac{1}{2}\vec{\mu}^\top V\vec{\mu} \quad (1.6)$$

$$= \underbrace{\frac{1}{2}(\vec{x} + V^{-1}(i\vec{k} - V\vec{\mu}))^\top V(\vec{x} + V^{-1}(i\vec{k} - V\vec{\mu}))}_{\textcircled{1}} + \underbrace{-\frac{1}{2}(i\vec{k} - V\vec{\mu})^\top V^{-1}(i\vec{k} - V\vec{\mu}) + \frac{1}{2}\vec{\mu}^\top V\vec{\mu}}_{\textcircled{2}}, \quad (1.7)$$

which we can now integrate, since it is now a quadratic form in terms of a shifted variable, $\vec{x} + \vec{p}$, where the constant (with respect to \vec{x}) vector \vec{p} is given by $V^{-1}(i\vec{k} - V\vec{\mu})$.¹

Now, shifting the integral from one in $d^n x$ to one in $d^n(x + p)$ does not change the measure, since the Jacobian of a shift is the identity. Then, we have

$$\phi(\vec{k}) = N \int d^n(x + p) \exp(-\textcircled{1} - \textcircled{2}) \quad (1.12)$$

$$= N \sqrt{\frac{(2\pi)^n}{\det V}} \exp(-\textcircled{2}) \quad (1.13)$$

$$= \underbrace{\frac{1}{\sqrt{\det V \det C}}}_{=1} \exp(-\textcircled{2}), \quad (1.14)$$

since the determinant of the inverse is the inverse of the determinant.

¹ In the last step we applied the matrix square completion formula: for a symmetric matrix A and vectors \vec{x} , \vec{b} we have

$$\frac{1}{2}(\vec{x} + A^{-1}\vec{b})^\top A(\vec{x} + A^{-1}\vec{b}) - \frac{1}{2}\vec{b}^\top A^{-1}\vec{b} = \quad (1.8)$$

$$= \frac{1}{2}[\vec{x}^\top A\vec{x} + \vec{x}^\top A A^{-1}\vec{b} + (A^{-1}\vec{b})^\top A\vec{x} + (A^{-1}\vec{b})^\top A A^{-1}\vec{b} - \vec{b}^\top A^{-1}\vec{b}] \quad (1.9)$$

$$= \frac{1}{2}[\vec{x}^\top A\vec{x} + \vec{x}^\top \vec{b} + \vec{b}^\top (A^{-1})^\top A\vec{x} + \vec{b}^\top (A^{-1})^\top \vec{b} - \vec{b}^\top A^{-1}\vec{b}] \quad (1.10)$$

$$= \frac{1}{2}\vec{x}^\top A\vec{x} + \vec{b}^\top \vec{x}, \quad (1.11)$$

which we used with $\vec{b} = i\vec{k} - V\vec{\mu}$.

Now, we only need to simplify ②:

$$\textcircled{2} = -\frac{1}{2} \left[-\vec{k}^\top V^{-1} \vec{k} - 2i\vec{\mu}^\top V V^{-1} \vec{k} + \vec{\mu}^\top V V^{-1} V \vec{\mu} \right] + \frac{1}{2} \vec{\mu}^\top V \vec{\mu} \quad (1.15)$$

$$= \frac{1}{2} \vec{k}^\top C \vec{k} + i\vec{\mu}^\top \vec{k}, \quad (1.16)$$

inserting which into the exponent yields the desired result. \square

Proof: by diagonalization. We now follow a different approach: the covariance matrix C is symmetric, so we will always be able to find an orthogonal matrix O (satisfying $O^\top = O^{-1}$) such that $C = O^\top D O$, where D is diagonal. We will then also have $V = C^{-1} = O^\top D^{-1} O$. Let us denote the eigenvalues of D as λ_i , and the eigenvalues of D^{-1} as $d_i = \lambda_i^{-1}$.

Defining $\vec{z} = O\vec{x}$, $\vec{m} = O\vec{\mu}$, $\vec{u} = O\vec{k}$ the negative of the argument of the integral becomes:

$$i\vec{k} \cdot \vec{x} + \frac{1}{2} (\vec{x} - \vec{\mu})^\top C^{-1} (\vec{x} - \vec{\mu}) = i\vec{u} \cdot \vec{z} + \frac{1}{2} (\vec{z} - \vec{m})^\top D^{-1} (\vec{z} - \vec{m}) \quad (1.17)$$

$$= i\vec{u} \cdot \vec{z} + \frac{1}{2} \sum_i d_i (z_i - m_i)^2 \quad (1.18)$$

$$= \sum_i \left[iu_i z_i + \frac{d_i}{2} (z_i^2 + m_i^2 - 2m_i z_i) \right] \quad (1.19)$$

$$= \sum_i \left[z_i^2 \frac{d_i}{2} + z_i (iu_i - m_i d_i) + \frac{d_i}{2} m_i^2 \right]. \quad (1.20)$$

With this, and since by $\det O = 1$ we have $d^n z = d^n x$, we can decompose our Gaussian integral into a product of Gaussian integrals:

$$\phi(\vec{k}) = N \int d^n x \exp \left(-i\vec{k} \cdot \vec{x} - \frac{1}{2} (\vec{x} - \vec{\mu})^\top C^{-1} (\vec{x} - \vec{\mu}) \right) \quad (1.21)$$

$$= N \int d^n z \exp \left(-\sum_i \left[z_i^2 \frac{d_i}{2} + z_i (iu_i - m_i d_i) + \frac{d_i}{2} m_i^2 \right] \right) \quad (1.22)$$

$$= N \prod_i \int dz_i \exp \left(-z_i^2 \frac{d_i}{2} - z_i (iu_i - m_i d_i) - \frac{d_i}{2} m_i^2 \right) \quad (1.23)$$

$$= N \prod_i \sqrt{\frac{2\pi}{d_i}} \exp \left(\frac{(iu_i - m_i d_i)^2}{2d_i} - \frac{d_i m_i^2}{2} \right) \quad (1.24)$$

$$= \frac{1}{\sqrt{\det C \det V}} \prod_i \exp \left(\frac{-u_i^2 + m_i^2 d_i^2 - 2iu_i m_i d_i}{2d_i} - \frac{d_i m_i^2}{2} \right) \quad (1.25)$$

$$= \exp \left(\sum_i \left[-\frac{u_i^2}{2d_i} - iu_i m_i \right] \right) \quad (1.26)$$

$$= \exp \left(-\frac{1}{2} \vec{u}^\top C \vec{u} - i\vec{u} \cdot \vec{m} \right) \quad (1.27)$$

$$= \exp \left(-\frac{1}{2} \vec{k}^\top C \vec{k} - i\vec{k} \cdot \vec{\mu} \right), \quad (1.28)$$

where we have used the expression for the single-variable Gaussian integral:

$$\int dz \exp(-az^2 + bz + c) = \sqrt{\frac{\pi}{a}} \exp\left(\frac{b^2}{4a} + c\right), \quad (1.29)$$

which comes from the one-variable completion of the square:

$$-az^2 + bz + c = -a\left(z - \frac{b}{2a}\right)^2 + \frac{b^2}{4a} + c. \quad (1.30)$$

Also, we used the fact that orthogonal transformation do not change fully-contracted objects, such as scalar products or bilinear forms. \square

2 Exercise 5

We can calculate the moments of a distribution through its characteristic function:

$$\mathbb{E}\left[x_\alpha^{n_\alpha} \dots x_\beta^{n_\beta}\right] = \frac{\partial^{n_\alpha \dots n_\beta} \phi(\vec{k})}{\partial(-ik_\alpha)^{n_\alpha} \dots \partial(-ik_\beta)^{n_\beta}} \Big|_{\vec{k}=0}. \quad (2.1)$$

In the multivariate Gaussian case we can then calculate the mean (component by component) as

$$\mathbb{E}(x_\alpha) = \frac{\partial \phi(\vec{k})}{\partial(-ik_\alpha)} \Big|_{\vec{k}=0} \quad (2.2)$$

$$= \frac{\partial}{\partial(-ik_\alpha)} \Big|_{\vec{k}=0} \exp\left(-\frac{1}{2} \vec{k}^\top C \vec{k} - i \vec{k} \cdot \vec{\mu}\right) \quad (2.3)$$

$$= \left[-i \sum_{\beta} k_\beta C_{\beta\alpha} + \mu_\alpha\right] \exp\left(-\frac{1}{2} \vec{k}^\top C \vec{k} - i \vec{k} \cdot \vec{\mu}\right) \Big|_{\vec{k}=0} \quad (2.4)$$

$$= \mu_\alpha, \quad (2.5)$$

where we used the fact that the differentiation of a symmetric bilinear form is as follows:

$$\frac{\partial}{\partial k_\alpha} \left(\sum_{\beta\gamma} k_\beta k_\gamma C_{\beta\gamma} \right) = 2 \sum_{\beta\gamma} \delta_{\beta\alpha} k_\gamma C_{\beta\gamma} = 2 \sum_{\gamma} k_\gamma C_{\alpha\gamma}. \quad (2.6)$$

The covariance matrix can be computed by linearity as

$$\tilde{C}_{\alpha\beta} = \mathbb{E}\left[(x_\alpha - \mathbb{E}(x_\alpha))(x_\beta - \mathbb{E}(x_\beta))\right] = \mathbb{E}[x_\alpha x_\beta] - \mu_\alpha \mu_\beta, \quad (2.7)$$

the first term of which reads as follows:

$$\mathbb{E}[x_\alpha x_\beta] = \frac{\partial^2 \phi(\vec{k})}{\partial(-ik_\beta) \partial(-ik_\alpha)} \Big|_{\vec{k}=0} \quad (2.8)$$

$$= \frac{\partial}{\partial(-ik_\beta)} \Big|_{\vec{k}=0} \left[-i \sum_{\beta} k_\beta C_{\beta\alpha} + \mu_\alpha \right] \exp \left(-\frac{1}{2} \vec{k}^\top C \vec{k} - i \vec{k} \cdot \vec{\mu} \right) \quad (2.9)$$

$$= C_{\alpha\beta} + \mu_\alpha \mu_\beta, \quad (2.10)$$

therefore, as expected, $\tilde{C}_{\alpha\beta}$ is indeed $C_{\alpha\beta}$.

3 Exercise 6

Claim 3.1. *The characteristic function of a multivariate Gaussian is, up to normalization, a multivariate Gaussian.*

Proof. The characteristic function is the exponential of (minus)

$$\frac{1}{2} \vec{k}^\top C \vec{k} + i \vec{k} \cdot \vec{\mu} = \frac{1}{2} \left(\vec{k} + i C^{-1} \vec{\mu} \right)^\top C \left(\vec{k} + i C^{-1} \vec{\mu} \right) + \frac{1}{2} \vec{\mu}^\top C^{-1} \vec{\mu}, \quad (3.1)$$

which means that the characteristic function is in the form

$$\phi(\vec{k}) = \text{const} \times \exp \left(-\frac{1}{2} (\vec{k} - \vec{m})^\top C (\vec{k} - \vec{m}) \right), \quad (3.2)$$

a multivariate normal with mean $\vec{m} = -i C^{-1} \vec{\mu}$ and covariance matrix C^{-1} , the inverse of the covariance matrix of the corresponding MVN. \square