# From Data to Insights - Exercise sheet 2

discussed in week after 2nd lecture

April 26, 2024

## 1 Linear transformations of random vectors

Let $\boldsymbol{X} = (X_1\ ,\ \dots\ ,\ X_n)^T$ be a random vector with covariance matrix $\mathbf{C}_X$, and let $\mathbf{M}$ be a constant (i.e. non-random) $m \times n$-matrix. We can define a new random vector as

$$\boldsymbol{Y} = \mathbf{M} \cdot \boldsymbol{X}\ . \tag{1}$$

Show that the covariance matrix of $\boldsymbol{Y}$ is given by

$$\mathbf{C}_Y = \mathbf{M} \cdot \mathbf{C}_X \cdot \mathbf{M}^T\ . \tag{2}$$

## 2 Fisher matrix for Gaussian random vectors

Let $\boldsymbol{X} = (X_1\ ,\ \dots\ ,\ X_n)^T$ have a multivariate Gaussian distribution, whose covariance matrix $\mathbf{C}$ is know. We do not know the expectation value $\boldsymbol{\mu} = \langle \boldsymbol{X} \rangle$ , but we know that is given by the function

$$\boldsymbol{\mu} = \boldsymbol{\mu}(\boldsymbol{\theta}) \tag{3}$$

for some unknown parameters $\boldsymbol{\theta} = (\theta_1\ ,\ \dots\ ,\ \theta_m)^T$ . I.e. we know that $\boldsymbol{X}$ has the PDF

$$p(\boldsymbol{x}|\boldsymbol{\theta}_{\text{true}}) = \frac{1}{|2\pi\mathbf{C}|^{n/2}}\ \exp\left(-\frac{1}{2}\left[\boldsymbol{x} - \boldsymbol{\mu}(\boldsymbol{\theta}_{\text{true}})\right]^T \cdot \mathbf{C}^{-1} \cdot \left[\boldsymbol{x} - \boldsymbol{\mu}(\boldsymbol{\theta}_{\text{true}})\right]\right)\ , \tag{4}$$

but we do not know the true parameters $\boldsymbol{\theta}_{\text{true}}$ .

In the lecture we defined the $ij$-element of the Fisher matrix $\mathbf{F}$ for the parameters $\boldsymbol{\theta}$ as

$$F_{ij} \equiv -\left\langle \frac{\partial^2}{\partial\theta_i \partial\theta_j} \ln p(\boldsymbol{X}|\boldsymbol{\theta})\Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_{\text{true}}} \right\rangle\ . \tag{5}$$

(To understand the meaning of this expression, recall how we had defined expectation values like $\langle f(X) \rangle$ in the lecture.) Show that in the above situation the Fisher matrix elements can also be written as

$$F_{ij} = \left(\frac{\partial\mu(\boldsymbol{\theta})}{\partial\theta_i}\Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_{\text{true}}}\right)^T \mathbf{C}^{-1} \left(\frac{\partial\mu(\boldsymbol{\theta})}{\partial\theta_j}\Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_{\text{true}}}\right)\ . \tag{6}$$

Now consider the one-dimensional case $n = 1 = m$ , and give an interpretation of Equation 6.

## 3 Linear regression with Gaussian random vectors

Let $\boldsymbol{X} = (X_1\ ,\ \dots\ ,\ X_n)^T$ have a multivariate Gaussian distribution, whose covariance matrix $\mathbf{C}$ is know. We do not know the expectation value $\boldsymbol{\mu} = \langle \boldsymbol{X} \rangle$ , but we know that is of the form

$$\boldsymbol{\mu} = a\boldsymbol{v} \tag{7}$$

for some known vector $\boldsymbol{v}$ but an unknown amplitude $a$ .

We now perform a measurement of $\boldsymbol{X}$ and obtain the result $\boldsymbol{x}$ . We want to use this measurement to estimate the amplitude $a$ . To do so, find the value $a_{\text{best}}$ that minimizes

$$\chi^2(a) = \left[\boldsymbol{x} - a\boldsymbol{v}\right]^T \mathbf{C}^{-1} \left[\boldsymbol{x} - a\boldsymbol{v}\right]\ . \tag{8}$$

In you expression for $a_{\text{best}}$, replace $\boldsymbol{x}$ with $\boldsymbol{X}$ to obtain the random variable $A_{\text{best}}$ . What is the variance of $A_{\text{best}}$ ? Compare this to the inverse Fisher information! In the light of the Cramer-Rao-bound, what do you conclude from that comparison for the estimator $A_{\text{best}}$ ?

# 4 Use a computer to simulate exercise 3

On your computer (or using `jupyter.physik.uni-muenchen.de`), setup a numerical experiment that implements the situation of exercise 3. I.e.

- Choose some values for $\mathbf{C}$ and $\boldsymbol{v}$. You also need to choose a "true" value for $a$, $a_{\boldsymbol{true}}$.
- Draw several random draws from the random variable $\boldsymbol{X}$.
- For each of these draws, calculate $a_{\text{best}}$.
- Is the variance of these $a_{\text{best}}$ values consistent with your findings from Exercise 3?