

# Homework 1: Manipulation of multivariate probability distributions, and the Posterior distribution

Lecturer: Georgios Karagiannis

georgios.karagiannis@durham.ac.uk

For Formative assessment, submit the solutions of the parts 1 and 2 from the Exercise 1, and the solution of the Exercise 2.

**Exercise 1. (★★)**

Let  $x \sim T_d(\mu, \Sigma, \nu)$ . Recall that  $x \sim T_d(\mu, \Sigma, \nu)$  is the marginal distribution  $f_x(x) = \int f_{x|\xi}(x|\xi)f_\xi(\xi)d\xi$  of  $(x, \xi)$  where

$$\begin{aligned} x|\xi &\sim N_d(\mu, \Sigma\xi v) \\ \xi &\sim \text{IG}\left(\frac{v}{2}, \frac{1}{2}\right) \end{aligned}$$

Consider partition such that

$$x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}; \quad \mu = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}; \quad \Sigma = \begin{bmatrix} \Sigma_1 & \Sigma_{21}^\top \\ \Sigma_{21} & \Sigma_2 \end{bmatrix},$$

where  $x_1 \in \mathbb{R}^{d_1}$  and  $x_2 \in \mathbb{R}^{d_2}$ .

Address the following:

1. Show that the marginal distribution of  $x_1$  is such that

$$x_1 \sim T_{d_1}(\mu_1, \Sigma_1, \nu)$$

**Hint:** Try to use the form  $f_x(x) = \int f_{x|\xi}(x|\xi)f_\xi(\xi)d\xi$ .

2. Show that

$$\xi|x_1 \sim \text{IG}\left(\frac{1}{2}(d_1 + v), \frac{1}{2}\frac{Q + v}{v}\right)$$

where  $Q = (\mu_1 - x_1)^\top \Sigma_1^{-1}(\mu_1 - x_1)$ .

**Hint:** The PDF of  $y \sim N_d(\mu, \Sigma)$  is

$$f(y) = (2\pi)^{-\frac{d}{2}} \det(\Sigma)^{-\frac{1}{2}} \exp\left(-\frac{1}{2}(y - \mu)^\top \Sigma^{-1}(y - \mu)\right)$$

**Hint:** The PDF of  $y \sim \text{IG}(a, b)$  is

$$f_{\text{IG}(a,b)}(y) = \frac{b^a}{\Gamma(a)} y^{-a-1} \exp\left(-\frac{b}{y}\right) 1_{(0,+\infty)}(y)$$

3. Let  $\xi' = \xi \frac{v}{Q+v}$ , with  $Q = (\mu_1 - x_1)^\top \Sigma_1^{-1}(\mu_1 - x_1)$ , show that

$$\xi'|x_1 \sim \text{IG}\left(\frac{v + d_1}{2}, \frac{1}{2}\right)$$

4. Show that the conditional distribution of  $x_2|x_1$  is such that

$$x_2|x_1 \sim T_{d_2}(\mu_{2|1}, \dot{\Sigma}_{2|1}, \nu_{2|1})$$

where

$$\begin{aligned}\mu_{2|1} &= \mu_2 + \Sigma_{21}\Sigma_{11}^{-1}(x_1 - \mu_1) \\ \dot{\Sigma}_{2|1} &= \frac{\nu + (\mu_1 - x_1)^\top \Sigma_1^{-1}(\mu_1 - x_1)}{\nu + d_1} \Sigma_{2|1} \\ \Sigma_{2|1} &= \Sigma_{22} - \Sigma_{21}\Sigma_1^{-1}\Sigma_{21}^\top \\ \nu_{2|1} &= \nu + d_1\end{aligned}$$

**Hint:** You can use the Example [Marginalization & conditioning] from the Lecture Handout

**Solution.**

1. From what is given, it is  $x|\xi \sim N_d(\mu, \Sigma\xi v)$  and  $\xi \sim \text{IG}(\frac{\nu}{2}, \frac{1}{2})$  namely,

$$f_x(x) = \int f_{x_1, x_2|\xi}(x_1, x_2|\xi) f_\xi(\xi) d\xi = \int f_{x_2|\xi, x_1}(x_2|\xi, x_1) f_{x_1|\xi}(x_1|\xi) f_\xi(\xi) d\xi$$

It is

$$\begin{aligned}f_{x_1}(x_1) &= \int \int f_{x_1, x_2|\xi}(x_1, x_2|\xi) f_\xi(\xi) d\xi dx_2 = \int \int f_{x_2|\xi, x_1}(x_2|\xi, x_1) f_{x_1|\xi}(x_1|\xi) f_\xi(\xi) d\xi dx_2 \\ &= \int \left( \int f_{x_2|\xi, x_1}(x_2|\xi, x_1) dx_2 \right) f_{x_1|\xi}(x_1|\xi) f_\xi(\xi) d\xi = \int f_{x_1|\xi}(x_1|\xi) f_\xi(\xi) d\xi\end{aligned}$$

Because  $x_1|\xi \sim N_{d_1}(\mu_1, \Sigma_1\xi v)$ , and  $\xi \sim \text{IG}(\frac{\nu}{2}, \frac{1}{2})$ , it is  $x_1 \sim T_{d_1}(\mu_1, \Sigma_1, \nu)$  from the statement of the question.

2. From what is given, it is  $x|\xi \sim N_d(\mu, \Sigma\xi v)$ , and hence  $x_1|\xi \sim N_d(\mu_1, \Sigma_1\xi v)$  as marginal of a Normal distribution. From the Bayes Theorem, it is

$$\begin{aligned}f_{\xi|x_1}(\xi|x_1) &\propto f_{x_1|\xi}(x_1|\xi) f(\xi) \\ &\propto \xi^{-\frac{d_1}{2}} \exp\left(-\frac{1}{2}(x_1 - \mu_1)^\top (\Sigma_1\xi v)^{-1} (x_1 - \mu_1)\right) \times \xi^{-\frac{d_1+v}{2}-1} \exp\left(-\frac{1}{\xi} \frac{1}{2}\right) \\ &\propto \xi^{-\frac{d_1+v}{2}-1} \exp\left(-\frac{1}{\xi} \frac{1}{2} \left[(x_1 - \mu_1)^\top \Sigma_1^{-1} (x_1 - \mu_1) \frac{1}{v} + 1\right]\right) \\ &\propto \xi^{-\frac{d_1+v}{2}-1} \exp\left(-\frac{1}{\xi} \frac{1}{2} \frac{Q+v}{v}\right)\end{aligned}$$

This is the kernel of the Inverse Gamma distribution, and hence I can recognize that

$$\xi|x_1 \sim \text{IG}\left(\frac{1}{2}(d_1 + v), \frac{1}{2} \frac{Q+v}{v}\right).$$

3. Let  $\xi' = \xi \frac{v}{Q+v}$ , with  $Q = (\mu_1 - x_1)^\top \Sigma_1^{-1}(\mu_1 - x_1)$ . Then it is

$$\begin{aligned}f(\xi'|x_1) &= f_{\text{IG}(\frac{1}{2}(d_1+v), \frac{1}{2} \frac{Q+v}{v})}(\xi|x_1) \left| \frac{d\xi}{d\xi'} \right| \propto (Q\xi')^{-\frac{d_1+v}{2}-1} \exp\left(-\frac{1}{2} \frac{Q+v}{v} \frac{1}{\frac{Q+v}{v}\xi'}\right) 1_{(0,+\infty)}\left(\frac{Q+v}{v}\xi'\right) \frac{Q+v}{v} \\ &\propto (\xi')^{-\frac{d_1+v}{2}-1} \exp\left(-\frac{1}{2} \frac{1}{\xi'}\right) 1_{(0,+\infty)}(\xi') = f_{\text{IG}(\frac{v+d_1}{2}, \frac{1}{2})}(\xi')\end{aligned}$$

So

$$\xi'|x_1 \sim \text{IG}\left(\frac{v+d_1}{2}, \frac{1}{2}\right)$$

4. I will try to show that

$$\begin{aligned} x_2|\xi', x_1 &\sim \text{N}_{d_2}\left(\mu_{2|1}, (v+d_1)\dot{\Sigma}_{2|1}\xi'\right) \\ \xi'|x_1 &\sim \text{IG}\left(\frac{v+d_1}{2}, \frac{1}{2}\right) \end{aligned}$$

which leads to

$$x_2|x_1 \sim \text{T}_{d_2}(\mu_{2|1}, \dot{\Sigma}_{2|1}, \nu_{2|1})$$

since because

$$f_{x_2|x_1}(x_2|x_1) = \int f_{x_2|\xi, x_1}(x_2|\xi, x_1) f_{\xi}(\xi|x_1) d\xi$$

- I have calculated that

$$\xi'|x_1 \sim \text{IG}\left(\frac{v+d_1}{2}, \frac{1}{2}\right)$$

where  $\xi' = \xi \frac{v}{Q+v}$  with  $Q = (\mu_1 - x_1)^\top \Sigma_1^{-1} (\mu_1 - x_1)$ .

- It is (from multivariate Normal properties of the Example in the Hint)

$$x_2|\xi, x_1 \sim \text{N}_{d_2}\left(\mu_{2|1}, \underbrace{(\Sigma_{22} - \Sigma_{21}\Sigma_1^{-1}\Sigma_{21}^\top)}_{=\Sigma_{2|1}}\xi v\right) \equiv \text{N}_{d_2}(\mu_{2|1}, \Sigma_{2|1}v\xi)$$

where  $\mu_{2|1} = \mu_2 + \Sigma_{21}\Sigma_{11}^{-1}(x_1 - \mu_1)$ . If I rearrange the parameters in order to appear  $\xi' = \xi \frac{v}{Q+v}$  in the covariance I get

$$x_2|\xi, x_1 \sim \text{N}_{d_2}\left(\mu_{2|1}, \Sigma_{2|1}v\xi' \frac{v+Q}{v} \frac{v+d_1}{v+d_1}\right)$$

By setting

$$\dot{\Sigma}_{2|1} = \Sigma_{2|1} \frac{v+Q}{v+d_1}$$

I get

$$x_2|\xi', x_1 \sim \text{N}_{d_2}\left(\mu_{2|1}, (v+d_1)\dot{\Sigma}_{2|1}\xi'\right)$$

So I have

$$\begin{aligned} x_2|\xi', x_1 &\sim \text{N}_{d_2}\left(\mu_{2|1}, (v+d_1)\dot{\Sigma}_{2|1}\xi'\right) \\ \xi'|x_1 &\sim \text{IG}\left(\frac{v+d_1}{2}, \frac{1}{2}\right) \end{aligned}$$

which gives that  $x_2|x_1 \sim \text{T}_{d_2}(\mu_{2|1}, \dot{\Sigma}_{2|1}, \nu_{2|1})$  with  $\nu_{2|1} = v + d_1$ . So the distribution of  $x_2|x_1$  is  $x_2|x_1 \sim \text{T}_{d_2}(\mu_{2|1}, \dot{\Sigma}_{2|1}, \nu_{2|1})$ .

**Note:** Alternatively, one could prove sub-questions (2) and (4) by performing several pages of Matrix calculations to show that

$$\begin{aligned}
 f_X(x|\mu, \Sigma) &= \frac{\Gamma(\frac{\nu+d}{2})}{\Gamma(\frac{\nu}{2})\nu^{\frac{d}{2}}\pi^{\frac{d}{2}}\det(\Sigma)^{\frac{1}{2}}} \left(1 + \frac{1}{\nu}(x - \mu)^T \Sigma^{-1}(x - \mu)\right)^{-\frac{\nu+d}{2}} \\
 &= \dots \\
 &= \frac{\Gamma(\frac{\nu+d_1}{2})}{\Gamma(\frac{\nu}{2})\nu^{\frac{d_1}{2}}\pi^{\frac{d_1}{2}}\det(\Sigma_1)^{\frac{1}{2}}} \left(1 + \frac{1}{\nu}(x_1 - \mu_1)^T \Sigma_1^{-1}(x_1 - \mu_1)\right)^{-\frac{\nu+d_1}{2}} \\
 &\quad \times \frac{\Gamma(\frac{\nu_{2|1}+d_2}{2})}{\Gamma(\frac{\nu_{2|1}}{2})\nu_{2|1}^{\frac{d_2}{2}}\pi^{\frac{d_2}{2}}\det(\Sigma_{2|1})^{\frac{1}{2}}} \left(1 + \frac{1}{\nu_{2|1}}(x_2 - \mu_{2|1})^T \Sigma_{2|1}^{-1}(x_2 - \mu_{2|1})\right)^{-\frac{\nu_{2|1}+d_2}{2}}
 \end{aligned}$$

see Raiffa, H., & Schlaifer, R. (1961; Section 8.3). Applied statistical decision theory. This requires a lot of vector and matrix calculus.

**Exercise 2.** (★★) Let  $x$  be an observation. Consider the Bayesian model

$$\begin{cases} x|\theta & \sim \text{Pn}(\theta) \\ \theta & \sim \Pi(\theta) \end{cases}$$

where  $\text{Pn}(\theta)$  is the Poisson distribution with expected value  $\theta$ . Consider a prior  $\Pi(\theta)$  with density such as  $\pi(\theta) \propto \frac{1}{\theta}$ . Show that the posterior distribution is not always defined.

**Hint-1:** It suffices to show that the posterior is not defined in the case that you collect only one observation  $x = 0$ .

**Hint-2:** Poisson distribution:  $x \sim \text{Pn}(\theta)$  has PMF

$$\text{Pn}(x|\theta) = \frac{\theta^x \exp(-\theta)}{x!} 1(x \in \mathbb{N})$$

**Solution.** The prior with  $\pi(\theta) \propto \frac{1}{\theta}$  is improper because

$$\int \pi(\theta) d\theta \propto \int \frac{1}{\theta} d\theta = \infty$$

So I need to check the properness condition,

$$\underbrace{\int_{\mathbb{R}_+} \text{Pn}(x|\theta) \pi(\theta) d\theta}_{\propto f(x)} \begin{cases} < \infty & \text{posterior distribution is defined} \\ = \infty & \text{posterior distribution is not defined} \end{cases}$$

I will show that the posterior distribution is not defined given that I have collected a single observation  $x = 0$ . So I need to show that

$$\underbrace{\int_{\mathbb{R}_+} \text{Pn}(x=0|\theta) \pi(\theta) d\theta}_{\propto f(x=0)} = \infty$$

It is

$$f(x) \propto \int_{\mathbb{R}_+} \mathbf{Pn}(x|\theta) \frac{1}{\theta} \theta = \int_0^\infty \exp(-\theta) \frac{\theta^x}{x!} \frac{1}{\theta} \theta$$

$$f(x=0) \propto \int_{\mathbb{R}_+} \exp(-\theta) \frac{\theta^0}{0!} \frac{1}{\theta} \theta = \int_0^\infty \exp(-\theta) \frac{1}{\theta} \theta$$

We will use a convergence criteria in order to check if  $\int_0^\infty \exp(-\theta) \frac{1}{\theta} \theta = \infty$ .

Consider  $h(\theta) = \exp(-\theta) \frac{1}{\theta}$ . The function  $h(\theta)$  has an improper behavior at 0, as it is not bounded there. Let  $g(\theta) = \frac{1}{\theta}$ . According to the Limit Comparison Test, it is

$$\lim_{\theta \rightarrow 0^+} \frac{h(\theta)}{g(\theta)} = \lim_{\theta \rightarrow 0^+} \frac{\frac{1}{\theta} \exp(-\theta)}{\frac{1}{\theta}} = 1 \neq 0$$

and

$$\int_0^\infty g(\theta) \theta = \int_0^\infty \frac{1}{\theta} \theta = \infty.$$

Therefore, it will be

$$\underbrace{\int_0^\infty h(\theta) \theta}_{=f(x=0)} = \infty$$

as well.

---