Exercise Sheet: Bayesian Statistics

Lecturer & author: Georgios P. Karagiannis georgios.karagiannis@durham.ac.uk

Part I

Matrix & vector calculus

The exercises about Matrix & vector calculus are optional and can be skipped.

Exercise 1. ($\star$)Let A, B be $K \times K$ invertible matrices . Show that

$$(A+B)^{-1} = A^{-1}(A^{-1}+B^{-1})^{-1}B^{-1}$$

Solution. It is

$$\begin{aligned}(A+B)^{-1} &= A^{-1}(I+A^{-1}B)^{-1}\\ &= A^{-1}(A^{-1}+B^{-1})^{-1}B^{-1}\end{aligned}$$

Exercise 2. ($\star\star$)[Woodbury matrix identity] Verify that

$$(A+UCV)^{-1} = A^{-1} - A^{-1}U\left(C^{-1}+VA^{-1}U\right)^{-1}VA^{-1}$$

if A and C are non-singular.

Solution.

By checking that $(A+UCV)\,(A+UCV)^{-1} = I$

$$\begin{aligned}(A+UCV) \times \Big[A^{-1} - A^{-1}U\left(C^{-1}+VA^{-1}U\right)^{-1}VA^{-1}\Big]&\\ &= I + UCVA^{-1} - (U+UCVA^{-1}U)(C^{-1}+VA^{-1}U)^{-1}VA^{-1}\\ &= I + UCVA^{-1} - UC(C^{-1}+VA^{-1}U)(C^{-1}+VA^{-1}U)^{-1}VA^{-1}\\ &= I + UCVA^{-1} - UCVA^{-1} = I.\end{aligned}$$

So

$$(A+UCV)^{-1} = A^{-1} - A^{-1}U\left(C^{-1}+VA^{-1}U\right)^{-1}VA^{-1}$$

Exercise 3. ($\star\star$)[Sherman–Morrison formula] Let A be a $K \times K$ invertible matrix and u and v two $K \times 1$ column vectors. Verify that

$$\left(A+uv^{\top}\right)^{-1} = A^{-1} - \frac{1}{1+v^{\top}A^{-1}u}A^{-1}uv^{\top}A^{-1}$$

if $1 + v^\top A^{-1} u \neq 0$, and if A is non-singular.

Solution.

$$
\begin{aligned}
(A + uv^T)(A + uv^T)^{-1} &= (A + uv^T)\left(A^{-1} - \frac{A^{-1}uv^TA^{-1}}{1 + v^TA^{-1}u}\right) \\
&= AA^{-1} + uv^TA^{-1} - \frac{AA^{-1}uv^TA^{-1} + uv^TA^{-1}uv^TA^{-1}}{1 + v^TA^{-1}u} \\
&= I + uv^TA^{-1} - \frac{uv^TA^{-1} + uv^TA^{-1}uv^TA^{-1}}{1 + v^TA^{-1}u} \\
&= I + uv^TA^{-1} - \frac{u\cancel{(1 + v^TA^{-1}u)}v^TA^{-1}}{\cancel{1 + v^TA^{-1}u}} \\
&= I + uv^TA^{-1} - uv^TA^{-1} \\
&= I
\end{aligned}
$$

Exercise 4. ($\star\star\star$)[Block partition matrix inversion] Let A be $K \times K$ invertible matrix, and let $B = A^{-1}$ its inverse. Consider Partition

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix};\ B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}$$

Namely, $B_{11} = \left[A^{-1}\right]_{11}$ is the upper corner of the A^{-1}, etc...
Show that

$$
\begin{aligned}
A_{11}^{-1} = B_{11} &= B_{12}B_{22}^{-1}B_{21} \\
A_{11}^{-1}A_{12} &= -B_{12}B_{22}^{-1}
\end{aligned}
$$

Hint: Start by noticing that

$$AB = I \iff \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}\begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} \iff \begin{cases} A_{11}B_{11} + A_{12}B_{21} & = I \\ A_{11}B_{12} + A_{12}B_{22} & = 0 \end{cases}$$

Solution. It is

$$AB = I \iff \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}\begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} \iff \begin{cases} A_{11}B_{11} + A_{12}B_{21} & = I \\ A_{11}B_{12} + A_{12}B_{22} & = 0 \end{cases}$$

So

$$
\begin{aligned}
A_{11}B_{12} + A_{12}B_{22} = 0 &\iff \\
A_{11}^{-1}\left(A_{11}B_{12} + A_{12}B_{22}\right)B_{22}^{-1} = 0 &\iff \\
B_{12}B_{22}^{-1} + A_{11}^{-1}A_{12} = 0 &
\end{aligned}
$$

So

$$A_{11}^{-1}A_{12} = -B_{12}B_{22}^{-1}$$

Also

$$
\begin{aligned}
A_{11}B_{12} + A_{12}B_{22} = 0 &\iff \\
(A_{11}B_{12} + A_{12}B_{22})B_{22}^{-1}B_{21} = 0 &\iff \\
A_{11}B_{12}B_{22}^{-1}B_{21} + A_{12}B_{21} = 0 & \\
A_{12}B_{21} &= -A_{11}B_{12}B_{22}^{-1}B_{21}
\end{aligned}
$$

Then, we plug in the above in $A_{11}B_{11} + A_{12}B_{21} = I$ we get

$$
\begin{aligned}
A_{11}B_{11} + A_{12}B_{21} = I &\iff \\
A_{11}B_{11} - A_{11}B_{12}B_{22}^{-1}B_{21} = I &\iff \\
B_{11} - B_{12}B_{22}^{-1}B_{21} = A_{11}^{-1} &
\end{aligned}
$$

So

$$A_{11}^{-1} = B_{11} = B_{12}B_{22}^{-1}B_{21}$$

Part II

Random variables

Exercise 5. ($\star$)Let $y \in \mathcal{Y} \subseteq \mathbb{R}$ be a univariate random variable with CDF $F_y(\cdot)$. Consider a bijective function $h : \mathcal{Y} \to \mathcal{Z}$ with $z = h(y)$, and h^{-1} its inverse. The PDF of z is

$$F_z(z) = \begin{cases} F_Y(h^{-1}(z)) & \text{if } h \nearrow \\ \\ 1 - F_Y(h^{-1}(z)) & \text{if } h \searrow \end{cases}$$

Solution. It is $z = h(y) \Leftrightarrow y = h^{-1}(z)$

For if $h \nearrow$ it is

$$F_z(z) = \mathrm{P}(Z \le z) = \mathrm{P}(h^{-1}(Z) \le h^{-1}(z)) = \mathrm{P}(Y \le h^{-1}(z)) = F_Y(h^{-1}(z))$$

For if $h \searrow$ it is

$$F_z(z) = \mathrm{P}(Z \le z) = \mathrm{P}(h^{-1}(Z) \ge h^{-1}(z)) = \mathrm{P}(Y \ge h^{-1}(z)) = 1 - F_Y(h^{-1}(z))$$

Exercise 6. ($\star$)Let $y \in \mathcal{Y} \subseteq \mathbb{R}$ be a univariate random variable with PDF $f_y(\cdot)$. Consider a bijective function $h : \mathcal{Y} \to \mathcal{Z} \subseteq \mathbb{R}$ and let h^{-1} be the inverse function of h. Consider a univariate random variable such that $z = h(y)$. The PDF of z is

$$f_z(z) \ = f_y(y)|\det(\frac{\mathrm{d}y}{\mathrm{d}z})| \ = f_y(h^{-1}(z))|\det(\frac{\mathrm{d}}{\mathrm{d}z}h^{-1}(z))|$$

Solution. It is $z = h(y) \Leftrightarrow y = h^{-1}(z)$

For if $h \nearrow$ it is

$$F_z(z) = \mathrm{P}(Z \le z) = \mathrm{P}(h^{-1}(Z) \le h^{-1}(z)) = \mathrm{P}(Y \le h^{-1}(z)) = F_Y(h^{-1}(z))$$

and

$$f_z(z) = \frac{\mathrm{d}}{\mathrm{d}z}F_z(z) = \frac{\mathrm{d}}{\mathrm{d}z}F_Y(h^{-1}(z)) = \frac{\mathrm{d}}{\mathrm{d}h^{-1}}F_Y(h^{-1})\det(\frac{\mathrm{d}}{\mathrm{d}z}h^{-1}(z))$$

For if $h \searrow$ it is

$$F_z(z) = \mathrm{P}(Z \le z) = \mathrm{P}(h^{-1}(Z) \ge h^{-1}(z)) = \mathrm{P}(Y \ge h^{-1}(z)) = 1 - F_Y(h^{-1}(z))$$

and

$$f_z(z) = \frac{\mathrm{d}}{\mathrm{d}z}F_z(z) = \frac{\mathrm{d}}{\mathrm{d}z}\left[1 - F_Y(h^{-1}(z))\right] = -\frac{\mathrm{d}}{\mathrm{d}h^{-1}}F_Y(h^{-1})\det(\frac{\mathrm{d}}{\mathrm{d}z}h^{-1}(z))$$

but $\det(\frac{\mathrm{d}}{\mathrm{d}z}h^{-1}(z)) < 0$ because $h \searrow$. So in both cases:

$$f_z(z) = f_y(h^{-1}(z))|\det(\frac{\mathrm{d}}{\mathrm{d}z}h^{-1}(z))|$$

Exercise 7. ($\star$)Let $y \sim \mathrm{Ex}(\lambda)$ r.v. with Exponential distribution with rate parameter $\lambda > 0$, and $f_{\mathrm{Ex}(\lambda)}(y) = \lambda\exp(-\lambda y)1(y \ge 0)$. Let $z = 1 - \exp(-\lambda y)$. Calculate the PDF of z, and recognize its distribution.

Solution. It is $z = 1 - \exp(-\lambda y) \Longleftrightarrow y = -\frac{1}{\lambda}\log(1-z)$, and $z \in [0,1]$. So $h^{-1}(z) = -\frac{1}{\lambda}\log(1-z)$. Then

$$f_z(z) = f_{\text{Ex}(\lambda)}(h^{-1}(z)) \times \left|\det\left(\frac{\mathrm{d}}{\mathrm{d}z}h^{-1}(z)\right)\right| = f_{\text{Ex}(\lambda)}\left(-\frac{1}{\lambda}\log(1-z)\right) \times \left|\det\left(\frac{\mathrm{d}}{\mathrm{d}z}\frac{-1}{\lambda}\log(1-z)\right)\right|$$
$$= \exp\left(-\lambda\frac{-1}{\lambda}\log(1-z)\right) 1(-\frac{1}{\lambda}\log(1-z) \geq 0) \times \left|-\frac{1}{\lambda}\frac{1}{1-z}\right| = 1(z \in [0,1])$$

From the density, we recognize that $z \sim \mathrm{U}(0,1)$ follows a uniform distribution.

Exercise 8. ($\star$)Prove the following properties

1. Let matrix $A \in \mathbb{R}^{q\times d}$, $c \in \mathbb{R}^q$, and $z = c + Ay$ then

$$\mathrm{E}(z) = \mathrm{E}(c + Ay) = c + A\mathrm{E}(y)$$

2. Let random variables $z \in \mathcal{Z}$ and $y \in \mathcal{Y}$, and let functions ψ_1 and ψ_2 defined on $\mathcal{Z}$ and $\mathcal{Y}$, then

$$\mathrm{E}(\psi_1(z) + \psi_2(y)) = \mathrm{E}(\psi_1(z)) + \mathrm{E}(\psi_2(y))$$

3. If random variables $z \in \mathcal{Z}$ and $y \in \mathcal{Y}$ are independent then

$$\mathrm{E}(\psi_1(z)\psi_2(y)) = \mathrm{E}(\psi_1(z))\mathrm{E}(\psi_2(y))$$

   for any functions ψ_1 and ψ_2 defined on $\mathcal{Z}$ and $\mathcal{Y}$.

Solution.

1. It is

$$\mathrm{E}(z) = \mathrm{E}(c + Ay) = \int (c + Ay)\,\mathrm{d}F(y) = c + A\int y\mathrm{d}F(y) = c + A\mathrm{E}(y)$$

2. It is

$$\mathrm{E}(\psi_1(z) + \psi_2(y)) = \int (\psi_1(z) + \psi_2(y))\,\mathrm{d}F((z,y)) = \int \psi_1(z)\mathrm{d}F((z,y)) + \int \psi_1(z)\mathrm{d}F((z,y))$$
$$= \int \psi_1(z)\mathrm{d}F(z) + \int \psi_1(z)\mathrm{d}F(z) = \mathrm{E}(\psi_1(z)) + \mathrm{E}(\psi_2(y))$$

3. If random variables $z \in \mathcal{Z}$ and $y \in \mathcal{Y}$ then

$$\mathrm{d}F(z,y) = \mathrm{d}F(z)\mathrm{d}F(y)$$

   It is

$$\mathrm{E}(\psi_1(z)\psi_2(y)) = \int (\psi_1(z)\psi_2(y))\,\mathrm{d}F((z,y)) = \left(\int \psi_1(z)\mathrm{d}F(z)\right)\left(\int \psi_2(y)\mathrm{d}F(y)\right)$$

Exercise 9. ($\star$)Prove the following properties of the covariance matrix

1. $\mathrm{Cov}(z,y) = \mathrm{E}(zy^\top) - \mathrm{E}(z)\,(\mathrm{E}(y))^\top$
2. $\mathrm{Cov}(z,y) = (\mathrm{Cov}(y,z))^\top$
3. $\mathrm{Cov}_\pi(c_1 + A_1 z, c_2 + A_2 y) = A_1\mathrm{Cov}_\pi(x,y)A_2^\top$, for fixed matrices A_1, A_2, and vectors c_1, c_2 with suitable dimensions.

4. If z and y are independent random vectors then $\text{Cov}(z, y) = 0$

Solution.

1. It is

$$\begin{aligned}
\text{Cov}(z, y) &= \text{E}\left((z - \text{E}(z))(y - \text{E}(y))^\top\right) \\
&\text{E}\left(zy^\top - z\text{E}(y)^\top - \text{E}(z)y^\top + \text{E}(z)\text{E}(y)^\top\right) \\
&= \text{E}(zy^\top) - \text{E}(z)\left(\text{E}(y)\right)^\top
\end{aligned}$$

2. It is

$$\begin{aligned}
(\text{Cov}(y, z))^\top = \left(\text{E}\left((z - \text{E}(z))(y - \text{E}(y))^\top\right)\right)^\top &= \text{E}\left(\left((z - \text{E}(z))(y - \text{E}(y))^\top\right)\right)^\top \\
&= \text{E}\left((y - \text{E}(y))(z - \text{E}(z))^\top\right) = \text{Cov}(y, z)
\end{aligned}$$

3. It is

$$\begin{aligned}
\text{Cov}(c_1 + A_1 z, c_2 + A_2 y) &= \text{E}\left((c_1 + A_1 z)(c_2 + A_2 y)^\top\right) - \text{E}(c_1 + A_1 z)\left(\text{E}(c_2 + A_2 y)\right)^\top \\
&= ... = A_1\left(\text{E}(zy^\top) - \text{E}(z)\left(\text{E}(y)\right)^\top\right)A_2^\top = A_1 \text{Cov}(z, y) A_2^\top
\end{aligned}$$

4. Obviously since

$$\text{Cov}(z, y) = 0 \Longleftrightarrow \text{Cov}(z_i, y_j) = \begin{cases} & i = j \\ & i \neq j \end{cases}$$

Exercise 10. ($\star$)Prove that the (i, j)-th element of the covariance matrix between vector z and y is the covariance between their elements z_i and y_j:

$$\left[\text{Cov}(z, y)\right]_{i,j} = \text{Cov}(z_i, y_j)$$

Solution.

It is

$$\begin{aligned}
\left[\text{Cov}(z, y)\right]_{i,j} &= \left[\text{E}(zy^\top) - \text{E}(z)\left(\text{E}(y)\right)^\top\right]_{i,j} = \\
&= \left[\text{E}(zy^\top)\right]_{i,j} - \left[\text{E}(z)\left(\text{E}(y)\right)^\top\right]_{i,j} \\
&= \text{E}(z_i y_j^\top) - \text{E}(z_i)\left(\text{E}(y_j)\right)^\top = \text{Cov}(z_i, y_j)
\end{aligned}$$

Exercise 11. ($\star$)Prove the following properties of $\text{Var}(Y)$ for a random vector $y \in \mathcal{Y} \subseteq \mathbb{R}^d$

1. $\text{Var}(y) = \text{E}(yy^\top) - \text{E}(y)\left(\text{E}(y)\right)^\top$
2. $\text{Var}(c + Ay) = A\text{Var}(y)A^\top$, for fixed matrix A, and vectors c with suitable dimensions.
3. $\text{Var}(y) \geq 0$; (semi-positive definite)

Solution.

1. $\text{Var}(y) = \text{Cov}(y, y) = \text{E}(yy^\top) - \text{E}(y)\left(\text{E}(y)\right)^\top$
2. $\text{Var}(c + Ay) = \text{Cov}(c + Ay, c + Ay) = A\text{Cov}(y, y)A^\top = A\text{Var}(y)A^\top$

3. For any vector $x \in \mathbb{R}^q$

$$t^\top \mathrm{Var}(y)t = t^\top \mathrm{E}\left((y-\mathrm{E}(y))(y-\mathrm{E}(y))^\top\right)t$$
$$= \mathrm{E}\left(\left(t^\top(y-\mathrm{E}(y))\right)\left(t^\top(y-\mathrm{E}(y))\right)^\top\right)$$
$$= \mathrm{E}\left(zz^\top\right) = \mathrm{E}\left(\sum_{j=1}^{d} z_j^2\right) \geq 0$$

for $z = t^\top(y - \mathrm{E}(y))$.

Exercise 12. ($\star$)Prove the following properties of characteristic functions

1. $\varphi_{A+Bx}(t) = e^{it^TA}\varphi_x(B^Tt)$ if $A \in \mathbb{R}^d$ and $B \in \mathbb{R}^{k\times d}$ are constants
2. $\varphi_{x+y}(t) = \varphi_x(t)\varphi_y(t)$ if and only if x and y are independent
3. if $M_x(t) = \mathrm{E}(e^{t^Tx})$ is the moment generating function, then $M_x(t) = \varphi_x(-it)$

Solution.

1. It is
$$\varphi_{A+Bx}(t) = \mathrm{E}(e^{it^T(A+Bx)}) = \mathrm{E}(e^{A+it^TBx}) = \mathrm{E}(e^{it^TA}e^{iB^Ttx}) = e^{it^TA}\mathrm{E}(e^{i(B^Tt)x}) = e^{it^TA}\varphi_x(B^Tt)$$
2. straightforward
3. straightforward

Exercise 13. ($\star$)Show that if $X \sim \mathrm{Ex}(\lambda)$ then $\varphi_X(t) = \frac{\lambda}{\lambda - it}$.

Solution. It is
$$\varphi_X(t) = \int_{-\infty}^{\infty} e^{itX} \underbrace{\lambda e^{-\lambda x} 1(X>0)}_{=f_{\mathrm{Ex}}(x|\lambda)} \mathrm{d}x = \lambda \int_{-\infty}^{\infty} e^{-x(\lambda - itX)} \mathrm{d}x = \frac{\lambda}{\lambda - it}$$

Exercise 14. ($\star$)

1. Find $\varphi_X(t)$ if $X \sim \mathrm{Br}(p)$.
2. Find $\varphi_Y(t)$ if $Y \sim \mathrm{Bin}(n, p)$

Solution.

1. It is
$$\varphi_X(t) = \sum_{x=0,1} e^{itX} P(X = x) = e^{it0}(1-p) + e^{it1}p = (1-p) + pe^{it}$$
2. Because Binomial r.v. results as a summation of n IID Bernoulli r.v., it is $Y = \sum_{i=1}^n X_i$, where $X_i \sim \mathrm{Br}(p)$ $i = 1, ..., n$ and IID. Then
$$\varphi_Y(t) = \varphi_{\sum X_i}(t) = \prod_{i=1}^{n} \varphi_{X_i}(t) = ((1-p) + pe^{it})^n$$

Exercise 15. ($\star\star\star$)Prove the following statement related to the Bayesian theorem:
Assume a probability space $(\Omega, \mathscr{F}, P)$. Let a random variable $y : \Omega \to \mathcal{Y}$ with distribution $F(\cdot)$. Consider a partition $y = (x, \theta)$ with $x \in \mathcal{X}$ and $\theta \in \Theta$. Then the probability density function (PDF), or the probability mass function (PMF) of $\theta|x$ is

$$f(\theta|x) = \frac{f(x|\theta)f(\theta)}{\int f(x|\theta)\mathrm{d}F(\theta)} \tag{1}$$

Hint Consider cases where x is discrete and continuous. In the later case use the mean value theorem :

$$\int_A f(x)g(x)\mathrm{d}x = f(\xi)\int_A g(x)\mathrm{d}x$$

where $\xi \in A$ if A is connected, and $g(x) \geq 0$ for $x \in A$.

Solution. We consider separately two cases.

x is discrete:

Let $\Theta_0 \subseteq \Theta$ be any sub-set of Θ; I need to show that

$$P(\theta \in \Theta_0|x) = \frac{\int_{\Theta_0} f(x|\theta)\mathrm{d}F(\theta)}{\int_{\Theta} f(x|\theta)\mathrm{d}F(\theta)} = \begin{cases} \int_{\Theta_0} \frac{f(x|\theta)f(\theta)}{\int_{\Theta} f(x|\theta)\mathrm{d}F(\theta)}\mathrm{d}\theta & ,\theta \text{ cont.} \\ \sum_{\theta\in\Theta_0} \frac{f(x|\theta)f(\theta)}{\int_{\Theta} f(x|\theta)\mathrm{d}F(\theta)} & ,\theta \text{ discr.} \end{cases}$$

By Bayes theorem it is

$$P(\theta \in \Theta_0|x) = \frac{P(\Theta_0, x)}{P(x)}$$

where $P(x) = \int_{\Theta} f(x|\theta)\mathrm{d}F(\theta)$ and $P(\Theta_0, x) = \int_{\Theta_0} f(x|\theta)\mathrm{d}F(\theta)$.

x is continuous:

Let $\Theta_0 \subseteq \Theta$ be any sub-set of Θ; because the probability $P(x) = 0$, I need to show that

$$\lim_{r\to 0} P(\theta \in \Theta_0|B_r(x)) = \frac{\int_{\Theta_0} f(x|\theta)\mathrm{d}F(\theta)}{\int_{\Theta} f(x|\theta)\mathrm{d}F(\theta)} = \begin{cases} \int_{\Theta_0} \frac{f(x|\theta)f(\theta)}{\int_{\Theta} f(x|\theta)\mathrm{d}F(\theta)}\mathrm{d}\theta & ,\theta \text{ cont.} \\ \sum_{\theta\in\Theta_0} \frac{f(x|\theta)f(\theta)}{\int_{\Theta} f(x|\theta)\mathrm{d}F(\theta)} & ,\theta \text{ discr.} \end{cases}$$

for an open ball $B_r(x) = \{x' \in \mathcal{X} : |x' - x| < r\}$. By Bayes theorem

$$P(\theta \in \Theta_0|B_r(x)) = \frac{P(\Theta_0, B_r(x))}{P(B_r(x))}$$

where

$$P(\Theta_0, B_r(x)) = \int_{\Theta_0} \left[\int_{B_r(x)} f(\zeta|\theta)\mathrm{d}\zeta\right] \mathrm{d}F(\theta)$$

$$P(B_r(x)) = \int_{\Theta} \left[\int_{B_r(x)} f(\zeta|\theta)\mathrm{d}\zeta\right] \mathrm{d}F(\theta)$$

By mean value theorem[1] there exists $\zeta' \in B_r(y)$ such as

$$\int_{B_r(x)} f(\zeta|\theta)\mathrm{d}\zeta = f(\zeta'|\theta)\int_{B_r(x)} \mathrm{d}\zeta = f(\zeta'|\theta)\ \|B_r(x)\|$$

Then

$$P(\theta \in \Theta_0|B_r(x)) = \frac{\int_{\Theta_0}\left[f(\zeta'|\theta)\ \|B_r(x)\|\right]\mathrm{d}F(\theta)}{\int_{\Theta}\left[f(\zeta'|\theta)\ \|B_r(x)\|\right]\mathrm{d}F(\theta)} \quad \xrightarrow{r\to 0} \quad \frac{\int_{\Theta_0} f(\zeta|\theta)\mathrm{d}F(\theta)}{\int_{\Theta} f(\zeta|\theta)\mathrm{d}F(\theta)}$$

Exercise 16. ($\star$)Prove that:

1. if $Z \sim \mathrm{N}(0, I)$ then $\varphi_Z(t) = \exp(-\frac{1}{2}t^T t)$, where $Z \in \mathbb{R}^d$
2. if $X \sim \mathrm{N}(\mu, \Sigma)$ then $\varphi_X(t) = \exp(it^T\mu - \frac{1}{2}t^T\Sigma t)$, where $X \in \mathbb{R}^d$

Hint: Assume as known that if $Z \sim \mathrm{N}(0, 1)$ then $\varphi_Z(t) = \exp(-\frac{1}{2}t^2)$, where $Z \in \mathbb{R}$

Solution.

1. It is

$$\varphi_Z(t) = \mathrm{E}(\exp(it^T Z)) = \mathrm{E}(\exp(i\sum_{j=1}^{d}(t_j Z_j))) = \mathrm{E}(\prod_{j=1}^{d}\exp(it_j Z_j)) = \prod_{j=1}^{d}\mathrm{E}(\exp(it_j Z_j))$$
$$= \prod_{j=1}^{d}\varphi_{Z_j}(t) = \prod_{j=1}^{d}\exp(-\frac{1}{2}t_j^2) = \exp(-\frac{1}{2}\sum_{j=1}^{d}t_j^2) = \exp(-\frac{1}{2}t^T t)$$

2. Assume a matrix L such as $\Sigma = LL^T$. It is $X = \mu + LZ$. Then

$$\varphi_X(t) = \varphi_{\mu+LZ}(t) = e^{it^T\mu}\varphi_Z(L^T t) = e^{it^T\mu}\exp(-\frac{1}{2}(L^T t)^T L^T t)$$
$$= e^{it^T\mu}\exp(-\frac{1}{2}t^T LL^T t) = \exp(it^T\mu - \frac{1}{2}t^T\Sigma t)$$

Exercise 17. ($\star$)Show the following properties of the Characteristic Function

1. $\varphi_x(0) = 1$ and $|\varphi_x(t)| \leq 1$ for all $t \in \mathbb{R}^d$
2. $\varphi_{A+Bx}(t) = e^{it^T A}\varphi_x(B^T t)$ if $A \in \mathbb{R}^d$ and $B \in \mathbb{R}^{k\times d}$ are constants
3. x and y are independent then $\varphi_{x+y}(t) = \varphi_x(t)\varphi_y(t)$ (we do not proov the other way around)
4. if $M_x(t) = \mathrm{E}(e^{t^T x})$ is the moment generating function, then $M_x(t) = \varphi_x(-it)$

Solution.

1. It is $\varphi_x(0) = \mathrm{E}(e^{i0^T x}) = \mathrm{E}(1) = 1$. Also

$$|\varphi_x(t)| = \left|\mathrm{E}(e^{it^T x})\right| = \left|\int\left(\cos(t^T x) + i\sin(t^T x)\right)\mathrm{d}F(x)\right| \leq \int\left|\cos(t^T x) + i\sin(t^T x)\right|\mathrm{d}F(x) \leq \int 1\mathrm{d}F(x) = 1$$

2. It is

$$\varphi_{A+Bx}(t) = \mathrm{E}(e^{it^T(A+Bx)}) = \mathrm{E}(e^{it^T A + Bit^T x}) = \mathrm{E}(e^{Ai}e^{i(B^T t)^\top x}) = e^{it^T A}\varphi_x(B^T t)$$

[1] $\int_A f(x)g(x)\mathrm{d}x = f(\xi)\int_A g(x)\mathrm{d}x$ where $\xi \in A$ if A is connected, and $g(x) \geq 0$ for $x \in A$.

3. It is

$$\varphi_{x+y}(t) = \mathrm{E}(e^{it^T(x+y)}) = \mathrm{E}(e^{it^T x} e^{it^T y}) = \mathrm{E}(e^{it^T x})\mathrm{E}(e^{it^T y}) = \varphi_x(t)\varphi_y(t)$$