Name: Student ID:

GAUSSIAN PROCESSES PROBLEM SHEET

1 Performing integrals over normal distributes takes practice.

(a) Consider the integral

$$I_1 = \int_{-\infty}^{\infty} \mathrm{e}^{-x^2/2} \mathrm{d}x.$$

Directly evaluating this is difficult, but there is a trick. Consider instead

$$I_1^2 = \int_{-\infty}^{\infty} \mathrm{e}^{-x^2/2} \mathrm{d}x \int_{-\infty}^{\infty} \mathrm{e}^{-y^2/2} \mathrm{d}y = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathrm{e}^{-(x^2+y^2)/2} \mathrm{d}x\mathrm{d}y.$$

By making the change of variables to polar coordinates where $r = \sqrt{x^2+y^2}$ and $\theta = \arctan(y/x)$ (so that $x = r\cos(\theta)$, $y = r\sin(\theta)$) then $\mathrm{d}x\mathrm{d}y = r\mathrm{d}r\mathrm{d}\theta$. Note that to integrate over all space we let θ vary from 0 to 2π and r to vary from 0 to ∞. Write down the integral in polar coordinate, make a further the change of variables $u = r^2/2$ to evaluate I_1^2 hence compute I_1 [5 marks]

5

(b) By making a change of variables compute

$$I_2 = \int_{-\infty}^{\infty} \mathrm{e}^{-(x-\mu)^2/(2\sigma^2)} \mathrm{d}x$$

[5 marks]

5

(c) By using the identity $\mathrm{e}^{a+b} = \mathrm{e}^a \mathrm{e}^b$, or more generally

$$\mathrm{e}^{\sum_i a_i} = \prod_i \mathrm{e}^{a_i}$$

compute

$$I_3 = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \mathrm{e}^{-\frac{1}{2}\|\boldsymbol{x}\|_2^2} \mathrm{d}x_1 \cdots \mathrm{d}x_n$$

where $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)^{\mathsf{T}}$. [3 marks]

$\overline{3}$

(d) By using the fact that for a positive semi-definite matrix, $\boldsymbol{\Xi}$, we can use the eigenvector decomposition $\boldsymbol{\Xi}^{-1} = \mathbf{V}\boldsymbol{\Lambda}^{-1}\mathbf{V}^{\mathsf{T}}$ where $\mathbf{V}$ is an orthogonal matrix with determinant $\det(\mathbf{V}) = \pm 1$ and $\boldsymbol{\Lambda}^{-1}$ is a diagonal matrix with elements λ_i^{-1} compute

$$I_4 = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \mathrm{e}^{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^{\mathsf{T}}\boldsymbol{\Xi}^{-1}(\boldsymbol{x}-\mu)} \mathrm{d}x_1 \cdots \mathrm{d}x_n.$$

[6 marks]

$\overline{6}$

(e) Using the facts, that $\Xi = \mathbf{V}\boldsymbol{\Lambda}\mathbf{V}^{\mathsf{T}}$, for any two square matrices $\mathbf{A}$ and $\mathbf{B}$ the determinants satisfy $\det(\mathbf{AB}) = \det(\mathbf{A})\det(\mathbf{B})$, and $\det(\mathbf{V}) = \det(\mathbf{V}^{\mathsf{T}}) = \pm 1$ show that $\det(\Xi) = \prod_i \lambda_i$. [1 mark]

$\overline{}$
1

End of question 1

(a) $\frac{}{5}$	(b) $\frac{}{5}$	(c) $\frac{}{3}$	(d) $\frac{}{6}$	(e) $\frac{}{1}$	Total $\frac{}{20}$

2 Consider a multivariate normal distribution

$$f_{\boldsymbol{X},\boldsymbol{Y}}(\boldsymbol{x},\boldsymbol{y}) = \mathcal{N}\left(\begin{pmatrix} \boldsymbol{x} \\ \boldsymbol{y} \end{pmatrix} \middle| \begin{pmatrix} \boldsymbol{0} \\ \boldsymbol{0} \end{pmatrix}, \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^\mathsf{T} & \mathbf{C} \end{pmatrix} \right)$$

where $\mathbf{A}$ and $\mathbf{C}$ are symmetric (positive definite) matrices. The matrix

$$\Xi = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^\mathsf{T} & \mathbf{C} \end{pmatrix}$$

is the covariance matrix.

We want to compute the conditional probability density function $f_{\boldsymbol{X},\boldsymbol{Y}}(\boldsymbol{x} \mid \boldsymbol{y})$. This is complicated because the normal distribution involve the inverse of the covariance matrix. Let

$$\mathbf{U} = \begin{pmatrix} \mathbf{I} & \mathbf{B}\mathbf{C}^{-1} \\ \mathbf{0} & \mathbf{I} \end{pmatrix}, \qquad \mathbf{D} = \begin{pmatrix} \mathbf{A} - \mathbf{B}\mathbf{C}^{-1}\mathbf{B}^\mathsf{T} & \mathbf{0} \\ \mathbf{0}^\mathsf{T} & \mathbf{C} \end{pmatrix}$$

where $\mathbf{I}$ is the identity matrix.

(a) Compute $\mathbf{U}\mathbf{D}$ [3 marks]

$\overline{}$
3

(b) Using the previous result compute $(\mathbf{U}\mathbf{D})\mathbf{U}^\mathsf{T}$. Hence show $\Xi = \mathbf{U}\mathbf{D}\mathbf{U}^\mathsf{T}$. [3 marks]

$\overline{}$
3

(c) Given that $\Xi = \mathbf{U}\mathbf{D}\mathbf{U}^\mathsf{T}$ write down Ξ^{-1} in terms of $\mathbf{U}$ and $\mathbf{D}$ [1 mark]

$\overline{}$
1

(d) Demonstrate by direct multiplication that

$$\mathbf{U}^{-1} = \begin{pmatrix} \mathbf{I} & -\mathbf{B}\mathbf{C}^{-1} \\ \mathbf{0} & \mathbf{I} \end{pmatrix} \quad \mathbf{D}^{-1} = \begin{pmatrix} (\mathbf{A} - \mathbf{B}\mathbf{C}^{-1}\mathbf{B}^\mathsf{T})^{-1} & \mathbf{0} \\ \mathbf{0}^\mathsf{T} & \mathbf{C}^{-1} \end{pmatrix} \quad (\mathbf{U}^\mathsf{T})^{-1} = \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ -\mathbf{C}^{-1}\mathbf{B}^\mathsf{T} & \mathbf{I} \end{pmatrix}$$

i.e. show $\mathbf{U}^{-1}\mathbf{U} = \mathbf{I}$, $\mathbf{D}^{-1}\mathbf{D} = \mathbf{I}$ and $(\mathbf{U}^\mathsf{T})^{-1}\mathbf{U}^\mathsf{T} = \mathbf{I}$. [6 marks]

i

ii

iii

$\overline{}$
6

(e) Letting $\boldsymbol{z} = \binom{\boldsymbol{x}}{\boldsymbol{y}}$ then we can write

$$f_{\boldsymbol{X},\boldsymbol{Y}}(\boldsymbol{x},\boldsymbol{y}) = f_{\boldsymbol{Z}}(\boldsymbol{z}) = \mathcal{N}(\boldsymbol{z} \mid \boldsymbol{0},\Xi) = \frac{1}{\sqrt{\det(2\pi\Xi)}} \mathrm{e}^{-\frac{1}{2}\boldsymbol{z}^{\mathsf{T}}\Xi^{-1}\boldsymbol{z}}.$$

where $\det(2\pi\Xi)$ is the determinant of the matrix $2\pi\Xi$ and is introduced to ensures that $f_{\boldsymbol{X},\boldsymbol{Y}}(\boldsymbol{x},\boldsymbol{y})$ is normalised. From parts (c) and (d)

$$\Xi^{-1} = \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ -\mathbf{C}^{-1}\mathbf{B}^{\mathsf{T}} & \mathbf{I} \end{pmatrix} \begin{pmatrix} (\mathbf{A} - \mathbf{B}\mathbf{C}^{-1}\mathbf{B}^{\mathsf{T}})^{-1} & \mathbf{0} \\ \mathbf{0}^{\mathsf{T}} & \mathbf{C}^{-1} \end{pmatrix} \begin{pmatrix} \mathbf{I} & -\mathbf{B}\mathbf{C}^{-1} \\ \mathbf{0} & \mathbf{I} \end{pmatrix}.$$

Expand out $\boldsymbol{z}^{\mathsf{T}}\Xi^{-1}\boldsymbol{z} = (\boldsymbol{x}^{\mathsf{T}},\boldsymbol{y}^{\mathsf{T}})\Xi^{-1}\binom{\boldsymbol{x}}{\boldsymbol{y}}$ (start by multiplying the vectors $\boldsymbol{z}^{\mathsf{T}}$ by $(\mathbf{U}^{\mathsf{T}})^{-1}$ and $\boldsymbol{z}$ by $\mathbf{U}^{-1}$) [4 marks]

$\overline{}$
4

In the next question we use the short-hand notation

$$\int f(\boldsymbol{z})\mathrm{d}\boldsymbol{z} = \int \cdots \int f(\boldsymbol{z})\mathrm{d}z_1\mathrm{d}z_2\ldots\mathrm{d}z_n$$

(f) To compute $f_{\boldsymbol{X}|\boldsymbol{Y}}(\boldsymbol{x} \mid \boldsymbol{y}) = f_{\boldsymbol{X},\boldsymbol{Y}}(\boldsymbol{x},\boldsymbol{y})/f_{\boldsymbol{Y}}(\boldsymbol{y})$ we need to find

$$\begin{aligned} f_{\boldsymbol{Y}}(\boldsymbol{y}) &= \int_{-\infty}^{\infty} f(\boldsymbol{x},\boldsymbol{y})\mathrm{d}\boldsymbol{x} \\ &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{\det(2\pi\boldsymbol{\Xi})}} \mathrm{e}^{-\frac{1}{2}(\boldsymbol{x}^\mathsf{T} - \boldsymbol{y}^\mathsf{T}\mathbf{C}^{-1}\mathbf{B}^\mathsf{T})(\mathbf{A}-\mathbf{B}\,\mathbf{C}^{-1}\mathbf{B}^\mathsf{T})^{-1}(\boldsymbol{x}-\mathbf{B}\,\mathbf{C}^{-1}\boldsymbol{y}) - \frac{1}{2}y^\mathsf{T}\mathbf{C}^{-1}\boldsymbol{y}} \mathrm{d}\boldsymbol{x} \\ &= \frac{\mathrm{e}^{-\frac{1}{2}y^\mathsf{T}\mathbf{C}^{-1}\boldsymbol{y}}}{\sqrt{\det(2\pi\boldsymbol{\Xi})}} \int_{-\infty}^{\infty} \mathrm{e}^{-\frac{1}{2}(\boldsymbol{x}^\mathsf{T} - \boldsymbol{y}^\mathsf{T}\mathbf{C}^{-1}\mathbf{B}^\mathsf{T})(\mathbf{A}-\mathbf{B}\,\mathbf{C}^{-1}\mathbf{B}^\mathsf{T})^{-1}(\boldsymbol{x}-\mathbf{B}\,\mathbf{C}^{-1}\boldsymbol{y})} \mathrm{d}\boldsymbol{x} \end{aligned}$$

By making a change of variable from $\boldsymbol{x}$ to $\boldsymbol{u} = \boldsymbol{x} - \mathbf{B}\mathbf{C}^{-1}\boldsymbol{y}$ rewrite the integral

$$I = \int_{-\infty}^{\infty} \mathrm{e}^{-\frac{1}{2}(\boldsymbol{x}^\mathsf{T} - \boldsymbol{y}^\mathsf{T}\mathbf{C}^{-1}\mathbf{B}^\mathsf{T})(\mathbf{A}-\mathbf{B}\,\mathbf{C}^{-1}\mathbf{B}^\mathsf{T})^{-1}(\boldsymbol{x}-\mathbf{B}\,\mathbf{C}^{-1}\boldsymbol{y})} \mathrm{d}\boldsymbol{x}$$

then use

$$\int_{-\infty}^{\infty} \mathrm{e}^{-\frac{1}{2}\boldsymbol{u}^\mathsf{T}\mathbf{M}^{-1}\boldsymbol{u}} \mathrm{d}\boldsymbol{u} = \sqrt{\det(2\pi\mathbf{M})}$$

to evaluate $f_{\boldsymbol{Y}}(\boldsymbol{y})$. [5 marks]

5

(g) Using $f_{\boldsymbol{X}|\boldsymbol{Y}}(\boldsymbol{x} \mid \boldsymbol{y}) = f_{\boldsymbol{X},\boldsymbol{Y}}(\boldsymbol{x},\boldsymbol{y})/f_{\boldsymbol{Y}}(\boldsymbol{y})$ write down $f_{\boldsymbol{X}|\boldsymbol{Y}}(\boldsymbol{x} \mid \boldsymbol{y})$. [3 marks]

$\overline{}$
3

End of question 2

(a) $\frac{}{3}$	(b) $\frac{}{3}$	(c) $\frac{}{1}$	(d) $\frac{}{6}$	(e) $\frac{}{4}$	(f) $\frac{}{5}$	(g) $\frac{}{3}$	Total $\frac{}{25}$