Durham University

EXAMINATION PAPER

Examination Session:	Year:	Exam Code:
May/June	2025	MATH4341-WE01

Title: Spatio-Temporal Statistics

Time:	2 hours	
Additional Material provided:		
Materials Permitted:		
Calculators Permitted:	Yes	Models Permitted: Casio FX83 series or FX85 series.

Instructions to Candidates:	Answer all questions. Section A is worth 40% and Section B is worth 60%. Within each section, all questions carry equal marks. Write your answer in the white-covered answer booklet with barcodes. Begin your answer to each question on a new page.

Revision:	

ED01/2025

SECTION A

Q1 (a) Consider the semi-variogram

$$\gamma(h) = \begin{cases} 0 & \text{if } h = 0 \\ \alpha + \beta\left(\frac{3|h|}{2} - \frac{|h|^3}{2}\right) & \text{if } |h| \in (0,1) \\ \alpha + \beta & \text{if } |h| \geq 1 \end{cases}$$

for $\alpha > 0$ and $\beta > 0$. What are the sill, range, nugget, and partial sill for this covariance model? Justify your answers.

(b) Let $\{Z(s) : s \in \mathcal{S}\}$ with $\mathcal{S} = (-\pi, \pi)$ be a Gaussian process with mean function $\mu(s) = m$ where $m \in \mathbb{R}$ is a constant and covariance function $c(s,t) = \frac{1}{2}(|s| + |t| - |t - s|)$ for $s, t \in \mathcal{S}$. Report whether or not the stochastic process $Z(\cdot)$ is weakly stationary, intrinsically stationary, continuous, and everywhere differentiable. Justify your answer.

(c) Let B be a $n \times n$ symmetric matrix with zero-valued diagonal elements (namely $[B]_{s,s} = 0$ for $s = 1, ..., n$) and such that $(I - B)$ is positive definite, where I denotes the identity matrix. Consider the conditional autoregression Gaussian model on a finite family $\mathcal{S} = \{1, ..., n\}$, $n > 1$, of sites defined by Gaussian local characteristics, with

$$\mathrm{E}\left(Z_t | Z_{\mathcal{S}\setminus t}\right) = \mu + \sum_{s \neq t} [B]_{s,t}(Z_s - \mu)$$

and $\mathrm{Var}\left(Z_t | Z_{\mathcal{S}\setminus t}\right) = 1$ for $t \in \mathcal{S}$ for some unknown parameter $\mu \in \mathbb{R}$. Show that the joint distribution of $Z = (Z_1, ..., Z_n)^\top$ is Gaussian with mean $\mu\underline{1}$ and covariance matrix $(I - B)^{-1}$. We denote the column vector of ones as $\underline{1} = (1, ..., 1)^\top$.

Q2 Consider the model

$$X_t = \frac{5}{6}X_{t-1} - \frac{1}{6}X_{t-2} + \varepsilon_t - \frac{1}{2}\varepsilon_{t-1} - \frac{1}{4}\varepsilon_{t-2} + \frac{1}{8}\varepsilon_{t-3},\ \varepsilon_t \sim N(0, \sigma^2),\ t \in \mathbb{Z},$$

of mixed auto-regressive moving average form of order ARMA(2, 3).

(a) Show that this model contains a parameter-redundancy allowing simplification to a model of reduced order, and that this simplified model can be written in the form

$$X_t = \frac{1}{3}X_{t-1} + \varepsilon_t - \frac{1}{4}\varepsilon_{t-2}.$$

Identify the order of this reduced model.

(b) Check that the simplified model is stationary, and then compute the stationary variance of the process (in terms of $\sigma > 0$).

(c) Compute the spectral density function, $S(\nu)$, of this process as a function of $\nu \in [0, \frac{1}{2}]$ and σ. Your final expression should not involve the imaginary unit, i.

SECTION B

Q3 Consider a set of random fields $\left\{\left(Z_j^{(u)}\left(s\right):s\in\mathcal{S}\right);j=1,...,k;u=1,...,k\right\}$ with

$$Z_j^{(u)}\left(s\right)=\sum_{p=1}^{k}a_{j,p}^{(u)}w_p^{(u)}\left(s\right),$$

where $\left\{w_p^{(u)}\left(s\right)\right\}$ are intrinsic random fields and $\left\{a_{j,p}^{(u)}\right\}$ are known constants. Let $\tilde{\gamma}_{i,j}^{(u)}\left(h\right)$ be the cross-variogram function of $Z_i^{(u)}\left(s\right)$ and $Z_j^{(u)}\left(s\right)$ for $u=1,...,k$.

(a) Write the definition of the cross-variogram function $\tilde{\gamma}_{i,j}^{(u)}\left(h\right)$ of $Z_i^{(u)}\left(s\right)$ and $Z_j^{(u)}\left(s\right)$ for $u=1,...,k$

(b) Assume that

$$\mathrm{E}\left(w_p^{(u)}\left(s\right)\right)=0$$

$$\mathrm{Cov}\left(w_p^{(u)}\left(s\right),w_q^{(v)}\left(s+h\right)\right)=\begin{cases}\gamma_{p,q}^{(u)}\left(h\right), & u=v\\ 0 & u\neq v\end{cases}$$

for $u=1,...,k$, $p=1,...,k$ and $q=1,...,k$. Show that

$$\tilde{\gamma}_{i,j}^{(u)}\left(h\right)=\sum_{p=1}^{k}a_{i,p}^{(u)}\sum_{q=1}^{k}a_{j,q}^{(u)}\gamma_{p,q}^{(u)}\left(h\right)$$

(c) Assume that

$$\mathrm{E}\left(w_p^{(u)}\left(s\right)\right)=0$$

$$\mathrm{Cov}\left(w_p^{(u)}\left(s\right),w_q^{(v)}\left(s+h\right)\right)=\begin{cases}\gamma^{(u)}\left(h\right), & u=v\,\text{and}\,p=q\\ 0 & u\neq v\,\text{or}\,p\neq q\end{cases}$$

for $u=1,...,k$. For $u=1,...,k$, compute the cross-variogram matrix $\tilde{\Gamma}^{(u)}\left(h\right)$ of vector

$$\left(Z_1^{(u)}\left(s\right),...,Z_k^{(u)}\left(s\right)\right)^\top$$

in the form

$$\tilde{\Gamma}^{(u)}\left(h\right)=B^{(u)}\gamma^{(u)}\left(h\right)$$

and express quantities $B^{(u)}$ as functions of matrix $A^{(u)}$ with $\left[A^{(u)}\right]_{i,p}=a_{i,p}^{(u)}$.

(d) Consider the assumptions in the previous part. Let $\left\{\left(Z_j\left(s\right):s\in\mathcal{S}\right);j=1,...,k\right\}$ be a set of random fields on $s\in\mathcal{S}$. Let

$$Z_j\left(s\right)=\mu_j\left(s\right)+\sum_{u=0}^{m}Z_j^{(u)}\left(s\right)$$

Show that the cross-variogram matrix of $\left(Z\left(s\right);s\in\mathcal{S}\right)$
where $Z\left(s\right)=\left(Z_1\left(s\right),...,Z_k\left(s\right)\right)^\top$ is

$$\Gamma\left(h\right)=\sum_{u=0}^{m}B^{(u)}\gamma^{(u)}\left(h\right)$$

Q4 Consider the constant (time-invariant) dynamic linear model

$$\mathbf{Y}_t = \mathsf{F}\mathbf{X}_t + \boldsymbol{\nu}_t, \quad \boldsymbol{\nu}_t \sim N(\mathbf{0}, \mathsf{V}),$$
$$\mathbf{X}_t = \mathsf{G}\mathbf{X}_{t-1} + \boldsymbol{\omega}_t, \quad \boldsymbol{\omega}_t \sim N(\mathbf{0}, \mathsf{W}),$$

for m-dimensional observation vectors, $\mathbf{Y}_t$, p-dimensional hidden states $\mathbf{X}_t$, and fully specified matrices $\mathsf{F}, \mathsf{G}, \mathsf{V}, \mathsf{W}$ (of appropriate dimensions). The model is considered for $t = 1, 2, \ldots$, and initialised with $\mathbf{X}_0 \sim N(\mathbf{m}_0, \mathsf{C}_0)$. Given n observations $\mathbf{y}_{1:n} = (\mathbf{y}_1, \ldots, \mathbf{y}_n)$, sequential computation of the filtered distributions

$$(\mathbf{X}_t|\mathbf{Y}_{1:t} = \mathbf{y}_{1:t}) \sim N(\mathbf{m}_t, \mathsf{C}_t), \quad t = 1, \ldots, n,$$

has been carried out using the Kalman filter. Interest now focuses on computing the *smoothing* distributions

$$(\mathbf{X}_t|\mathbf{Y}_{1:n} = \mathbf{y}_{1:n}) \sim N(\mathbf{s}_t, \mathsf{S}_t), \quad t = n, \ldots, 1.$$

We begin by noting that $\mathbf{s}_n = \mathbf{m}_n$ and $\mathsf{S}_n = \mathsf{C}_n$.

(a) (i) Consider the problem at time $t < n$, where we have already computed $\mathbf{s}_{t+1}$, S_{t+1}. Write down the form of the joint distribution of

$$\left(\begin{matrix} \mathbf{X}_t \\ \mathbf{X}_{t+1} \end{matrix} \middle| \mathbf{y}_{1:t} \right)$$

(ii) Use multivariate normal conditioning to show that

$$(\mathbf{X}_t|\mathbf{X}_{t+1}, \mathbf{y}_{1:t}) \sim N\left(\mathbf{m}_t + \mathsf{L}_t[\mathbf{X}_{t+1} - \tilde{\mathbf{m}}_{t+1}],\ \mathsf{C}_t - \mathsf{L}_t\tilde{\mathsf{C}}_{t+1}\mathsf{L}_t^{\mathsf{T}}\right),$$

where $\tilde{\mathbf{m}}_{t+1} = \mathsf{G}\mathbf{m}_t$, $\tilde{\mathsf{C}}_{t+1} = \mathsf{G}\mathsf{C}_t\mathsf{G}^{\mathsf{T}} + \mathsf{W}$, and $\mathsf{L}_t = \mathsf{C}_t\mathsf{G}^{\mathsf{T}}\tilde{\mathsf{C}}_{t+1}^{-1}$.

(iii) Explain why this is also the distribution of $(\mathbf{X}_t|\mathbf{X}_{t+1}, \mathbf{y}_{1:n})$, and then marginalise out $\mathbf{X}_{t+1}$ to obtain expressions for $\mathbf{s}_t$ and S_t.

(b) Explain how you would modify the above backward smoothing procedure to instead generate an exact sample from the conditional distribution $(\mathbf{X}_{1:n}|\mathbf{y}_{1:n})$.

(c) (i) Suppose that we wish to model a monthly (period 12) univariate time series using a dynamic linear model consisting of a locally constant trend and a Fourier-based seasonal effect using two harmonics. How would you structure this model? Give explicit forms for F and G, and suggest an appropriate structural form for W (though this may contain unspecified parameters).

(ii) Discuss briefly one approach that could be used to estimate any unspecified parameters in the model. Detailed formulas are not required.