EXAMINATION PAPER

Examination Session:	Year:	Exam Code:
May/June	2024	MATH4341-WE01

Title:
Spatio-Temporal Statistics

Time:	3 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. Students must use the mathematics specific answer book.

Revision:	

ED01/2024

SECTION A

Q1 (a) Suppose that $\{Z(s) : s \in \mathcal{S}\}$ is a stationary isotropic stochastic process with zero mean and a covariance function given by

$$c(h) = \begin{cases} \sigma^2 \exp\left(-\left|\frac{h}{\phi}\right|^r\right) & \text{if } h > 0 \\ \sigma^2 + \tau^2 & \text{if } h = 0 \end{cases}$$

where $0 < r < 2$.

(i) What are the range, sill, partial sill, and nugget for this covariance model? Justify your answers.

(ii) In terms of the covariance function, derive an expression for the semi-variogram, simplifying your result as much as possible.

(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

$$c(s,t) = \frac{1}{2}\left(\cos(s) + \cos(t) - (t-s)^{\frac{2}{3}}\right)$$

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.

Q2 (a) Consider the Exponential covariance function $c(h) = \sigma^2 \exp\left(-\beta \|h\|_1^1\right)$ for $\sigma^2, \beta > 0$ and $h \in \mathbb{R}^d$. Compute the spectral density from Bochner's theorem.

(b) Consider the Gaussian CAR model with local characteristics $\{\Pr_i(z_i | z_{\mathcal{S}-i})\}$ being Gaussian distributions with mean $\mathrm{E}(Z_i | Z_{\mathcal{S}-i}) = \sum_{j \neq i} b_{i,j} Z_j$ and variance $\mathrm{Var}(Z_i | Z_{\mathcal{S}-i}) = \kappa_i$ for $i \in \mathcal{S}$. Assume that $\{\Pr_i(z_i | z_{\mathcal{S}-i})\}$ are fully compatible with the joint distribution $\Pr_Z(z)$ of $Z = \{Z_i\}$. By using Besag's factorization theorem, derive the joint distribution $\Pr_Z(z)$ of Z as

$$Z \sim \mathrm{N}\left(\mu, (I - B)^{-1} K\right)$$

for some matrices B and K that you will specify. State any assumptions made.

Q3 Consider the discrete-time auto-regressive stochastic model

$$X_t = \phi X_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim N(0, \sigma^2), \quad t = 1, 2, \ldots,$$

given parameters $\phi \in \mathbb{R}$ and $\sigma > 0$.

3.1 Suppose first that the process is initialised deterministically with $X_0 = x_0$ for some specific $x_0 \in \mathbb{R}$.

(i) Show that the (conditional) distribution of the process at time $t > 0$ is given by

$$(X_t | X_0 = x_0) \sim N\left(\phi^t x_0, \frac{1 - \phi^{2t}}{1 - \phi^2}\sigma^2\right).$$

(ii) What values of ϕ will lead to a stable limiting distribution as $t \to \infty$?

(iii) Identify this limiting distribution.

3.2 Suppose now that a random initialisation is used,

$$X_0 \sim N(\mu_0, \sigma_0^2),$$

for some given $\mu_0 \in \mathbb{R}$ and $\sigma_0 > 0$.

(i) Deduce the form of the marginal distribution of the process at time t, X_t.

(ii) Use this to confirm that if the process is initialised with the limiting distribution previously identified, the process is stationary.

Q4 Consider the discrete time order p auto-regressive model

$$X_t = \sum_{j=1}^{p} \phi_j X_{t-j} + \varepsilon_t, \quad \varepsilon_t \sim N(0, \sigma^2), \quad t \in \mathbb{Z},$$

for auto-regressive parameters $\boldsymbol{\phi} = (\phi_1, \phi_2, \ldots, \phi_p)^\mathsf{T} \in \mathbb{R}^p$, and innovation variance $\sigma^2 > 0$. An observed stationary time series $x_1, x_2, \ldots, x_n$ is assumed to be consistent with a model of this form.

4.1 Briefly describe how *moment matching* can be used to estimate the auto-regressive parameters most consistent with the observed time series. Include an appropriate expression for the parameter estimates, $\hat{\boldsymbol{\phi}}$.

4.2 Briefly describe how *least squares fitting* can be used to estimate the auto-regressive parameters most consistent with the observed time series. Include an appropriate expression for the parameter estimates, $\hat{\boldsymbol{\phi}}$.

4.3 Explain why moment matching and least squares fitting are asymptotically equivalent for fitting auto-regressive models.

SECTION B

Q5 Assume a random field Z on $\mathcal{S} = \{1, ..., n\}$ with values in $\mathcal{Z} = \{0, ..., N\}$, where $n \in \mathbb{N} - \{0\}$. Here, $Z(s)$ represents the number of successes at location $s \in \mathcal{S}$. Consider a family of conditional distributions

$$z_i | z_{-i} \sim \text{Binomial}\left(N, \theta_i\left(z_{-i}\right)\right), \quad i \in \mathcal{S}$$

where

$$\theta_i\left(z_{-i}\right) = \left(1 + \exp\left(-\alpha_i - \sum_{j:j\sim i} \beta_{i,j} z_j\right)\right)^{-1}$$

for some $\{\alpha_i\}$ and $\{\beta_{i,j}\}$. We use notation $z_{-i} = \left(z_1, ..., z_{i-1}, z_{i+1}, ..., z_n\right)^\top$.

(a) Show that the conditionals $z_i | z_{-i}$ are compatible as a Besag's auto-model when $\{\alpha_i\}$ and $\{\beta_{i,j}\}$ satisfy certain conditions, and specify these conditions.

Hint: The probability mass function of a random variable $x \sim \text{Binomial}(n, p)$ is $\Pr(x) = \binom{n}{x} p^x (1-p)^{n-x} 1(x \in \{0, ..., n\})$.

(b) Write down the marginal distribution of the associated random field.

(c) What would be the sign of $\{\beta_{i,j}\}$ if you wish to introduce competition between neighbouring sites? What would be the sign of $\{\beta_{i,j}\}$ if you wish to introduce similarity between neighbouring sites? What does α_i represent when $\beta_{i,j} = 0$?

Q6 Consider a statistical model which is a stochastic process $(Z(s))_{s \in \mathcal{S}}$ with $\mathcal{S} = \left(0, \frac{1}{2}\pi\right)$, where $Z(\cdot) \sim \text{GP}(\mu(\cdot), c(\cdot,\cdot))$ with mean function $\mu(s) = \frac{1}{2}s$ and covariance function $c(s,t) = \cos(s-t)$ for any $s \in \mathcal{S}$ and $t \in \mathcal{S}$. Assume there is available a dataset $\{(Z_i, s_i)\}_{i=1}^n$ where $Z_i = Z(s_i)$ and $s_i \in \mathcal{S}$ are point sites.

(a) Imagine you are interested in locations in the domain $\mathcal{S}$. Is the above covariance function $c(\cdot,\cdot)$ a reasonable choice according to the 'First Law of Geography'? Explain your answer.

(b) Compute the block mean $\mu(v)$ for some block $v = [a,b] \subset \mathcal{S}$.

(c) Compute the block covariance function $c(v,s)$ for some block $v = [a,b] \subset \mathcal{S}$ and point $s \in \mathcal{S}$.

(d) Compute the block covariance function $c(v,v')$ for some blocks $v = [a,b] \subset \mathcal{S}$ and $v' = [a',b'] \subset \mathcal{S}$.

(e) Consider a set of sites $\mathfrak{S} = \{s_1, ..., s_n\}$. Consider observations $Z = (Z_1, ..., Z_n)^\top$ where $Z_i = Z(s_i)$ for $i = 1, ..., n$. Derive the predictive stochastic process $[Z(v)|Z]$ at any block $v = [a,b] \subset \mathcal{S}$ with $|v| > 0$.

Hint: Let $x_1 \in \mathbb{R}^{d_1}$, and $x_2 \in \mathbb{R}^{d_2}$. If

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \sim \text{N}_{d_1+d_2}\left(\begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \begin{bmatrix} \Sigma_1 & \Sigma_{21}^\top \\ \Sigma_{21} & \Sigma_2 \end{bmatrix}\right)$$

then

$$x_2 | x_1 \sim \text{N}_{d_2}\left(\mu_{2|1}, \Sigma_{2|1}\right)$$

where

$$\mu_{2|1} = \mu_2 + \Sigma_{21}\Sigma_1^{-1}(x_1 - \mu_1) \text{ and } \Sigma_{2|1} = \Sigma_2 - \Sigma_{21}\Sigma_1^{-1}\Sigma_{21}^\top$$

Q7 Consider the second-order auto-regressive model

$$X_t = X_{t-1} - \frac{1}{2}X_{t-2} + \varepsilon_t, \quad \varepsilon_t \sim N(0,1), \quad t \in \mathbb{Z}.$$

7.1 (i) Confirm that this model is stationary.

(ii) Compute the auto-correlation function, ρ_t, of this process. Your final expression should be an explicit function of t and should not involve the imaginary unit, i.

7.2 (i) Obtain the stationary variance of the process.

(ii) Without further computation, using previous working, write down the partial auto-correlation function for this process.

7.3 (i) Compute the spectral density function for this process as a function of frequency $\omega \in [0, \frac{1}{2}]$. Your final expression should not involve the imaginary unit, i.

(ii) Without further computation, using previous working, at roughly what value of ω would you expect a peak in the spectral density to occur? Justify your answer.

Q8 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

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

is desired (often referred to as the Kalman filter).

8.1 (i) Consider the problem at time t, where we have already computed $\mathbf{m}_{t-1}$, C_{t-1}. Show that the predictive distribution for $\mathbf{X}_t$ can be written

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

for appropriate definitions of $\tilde{\mathbf{m}}_t$, $\tilde{\mathsf{C}}_t$, which you should deduce.

(ii) Construct the joint distribution of $\mathbf{X}_t$ and $\mathbf{Y}_t$ given all of the data up to time $t-1$.

(iii) Use multivariate normal conditioning to deduce expressions for $\mathbf{m}_t$ and C_t.

8.2 Assume now that we have run our Kalman filter to obtain the final filtered moments, $\mathbf{m}_n$, C_n. Deduce the forecast distribution (mean and variance) for one and two steps ahead, $\mathbf{Y}_{n+1}$ and $\mathbf{Y}_{n+2}$, given the data $\mathbf{y}_{1:n}$.

8.3 (i) Suppose that we wish to model a quarterly (period 4) univariate time series using a dynamic linear model consisting of a locally linear trend and a seasonal effect. 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.