EXAMINATION PAPER

Examination Session:	Year:	Exam Code:
May/June	2024	MATH4337-WE01

Title:
Uncertainty Quantification IV

Time:	2 hours	
Additional Material provided:		
Materials Permitted:		
Calculators Permitted:	No	Models Permitted: Use of electronic calculators is forbidden.

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 We wish to emulate an expensive 1D computer model $f(x)$ over the input range $x \in [0,1]$. We set up a standard Bayes Linear emulator, specifying a squared exponential prior covariance structure, with $\sigma^2 = 1$, $\theta = 1/3$, and prior expectation $\mathrm{E}[f(x)] = -1$. A single run is performed at the location $x^{(1)} = 0.3$, yielding the output $D = f(x^{(1)}) = -2$.

1.1 Find an expression for the emulator expectation $\mathrm{E}_D[f(x)]$.

1.2 Find an expression for the emulator variance $\mathrm{Var}_D[f(x)]$.

1.3 There is interest in determining the lowest value that $f(x)$ could possibly take. Find the input location x that minimises the lower end of the 3-sigma emulator prediction interval.

Q2 We wish to emulate an expensive computer model $f(x)$ with 2D input $x \in \mathcal{X}_0 = [0,1]^2$ and scalar output $f \in \mathbb{R}$. We use a standard Bayes Linear emulator with prior expectation $\mathrm{E}[f(x)] = 0$, and specify covariance structure

$$\mathrm{Cov}[f(x), f(x')] \;=\; \sigma^2 c(x - x')$$

where $c(x - x')$ represents the correlation function. Two runs have been performed and these are at locations $x^{(1)}$ yielding $D_1 = f(x^{(1)})$ and at $x^{(2)}$ yielding $D_2 = f(x^{(2)})$, where $x^{(1)} \neq x^{(2)}$. The vector of run outputs is denoted $D = (D_1, D_2)^T$.

2.1 By explicitly constructing $\mathrm{Var}[D]$, show that its inverse is given by

$$\mathrm{Var}[D]^{-1} \;=\; \frac{1}{\sigma^2(1 - v^2)} \begin{pmatrix} 1 & -v \\ -v & 1 \end{pmatrix}$$

where $v = c(x^{(1)} - x^{(2)})$.

2.2 Hence show that the emulator expectation $\mathrm{E}_D[f(x)]$ is given by:

$$\mathrm{E}_D[f(x)] \;=\; \frac{1}{1 - v^2}\left\{ \Big[w_1(x) - v\, w_2(x)\Big] D_1 \;+\; \Big[w_2(x) - v\, w_1(x)\Big] D_2 \right\}$$

where $w_1(x) = c(x - x^{(1)})$ and $w_2(x) = c(x - x^{(2)})$.

2.3 Demonstrate that the emulator expectation will interpolate the run data. State a constraint on $c(x - x')$ that is required for this interpolation to be continuous.

2.4 Show that the emulator variance $\mathrm{Var}_D[f(x)]$ is given by:

$$\mathrm{Var}_D[f(x)] \;=\; \frac{\sigma^2}{1 - v^2}\left\{ 1 - v^2 - w_1(x)^2 - w_2(x)^2 \;+\; 2v\, w_1(x) w_2(x) \right\}$$

SECTION B

Q3 A Bayes linear emulator with basis functions $g_j(x)$, unknown coefficients β_j and weakly stationary process $u(x)$, takes the form:

$$f(x) \;=\; \sum_{j=1}^{p} \beta_j g_j(x) \;+\; u(x) \;=\; g(x)^T \beta \;+\; u(x)$$

Prior specifications for $\beta \equiv (\beta_1, \ldots, \beta_p)^T$ and $u(x)$ are given by:

$$\mathrm{E}[\beta] = \mu_\beta, \qquad \mathrm{Var}[\beta] = \Sigma_\beta, \qquad \mathrm{E}[u(x)] = 0,$$

$$\mathrm{Cov}[u(x), u(x')] = \sigma^2 c(x - x'), \qquad \mathrm{Cov}[\beta, u(x)] = 0,$$

where $c(x - x')$ represents the usual squared exponential covariance structure and where Σ_β is of full rank. We define the model output from n runs to be $D = \left(f(x^{(1)}), \ldots, f(x^{(n)})\right)^T$ and also $U = \left(u(x^{(1)}), \ldots, u(x^{(n)})\right)^T$ with $\mathrm{Var}[U] \equiv \Omega$ where Ω is an $n \times n$ covariance matrix with individual elements $\Omega_{ij} = \sigma^2 c(x^{(i)} - x^{(j)})$. We also define the $n \times p$ design matrix X with elements $X_{ij} = g_j(x^{(i)})$.

3.1 With the above specifications, show that we can write $D \;=\; X\beta + U$.

3.2 Show that the expectation of the regression coefficients β adjusted by the run data D is given by:

$$\mathrm{E}_D[\beta] \;=\; (X^T \Omega^{-1} X + \Sigma_\beta^{-1})^{-1} \left[\Sigma_\beta^{-1} \mu_\beta \;+\; (X^T \Omega^{-1} X)\, \hat{\beta}_{GLS} \right]$$

where your answer should also contain a definition of the Generalised Least Squares estimate $\hat{\beta}_{GLS}$. *Hint*: you can use the following Matrix Identity which states that for matrices A, B, C, G of appropriate dimension:

$$AB\,(GAB + C)^{-1} \;=\; (BC^{-1}G + A^{-1})^{-1} BC^{-1}$$

3.3 Comment on the form of $\mathrm{E}_D[\beta]$ derived in **Q3.2**.

3.4 Show that the variance of the regression coefficients β adjusted by the run data D is given by:

$$\mathrm{Var}_D[\beta] \;=\; (X^T \Omega^{-1} X + \Sigma_\beta^{-1})^{-1}$$

3.5 Comment on the form of $\mathrm{Var}_D[\beta]$ derived in **Q3.4**.

3.6 Examine the behaviour of $\mathrm{E}_D[\beta]$ and $\mathrm{Var}_D[\beta]$ in the vague β prior limit.

3.7 In which limiting situation would $\mathrm{E}_D[\beta] \to \hat{\beta}_{OLS}$ and $\mathrm{Var}_D[\beta] \to \mathrm{Var}[\hat{\beta}_{OLS}]$, where $\hat{\beta}_{OLS}$ is the Ordinary Least Squares estimate for β? Justify your answer.

Q4 We wish to optimise an expensive 1D function $f(x)$ over the interval $x \in [0, 1]$ using a Bayesian Optimisation approach. We have performed a set of n runs giving model output values $D = (f(x^{(1)}), \ldots, f(x^{(n)}))^T$ and intend to use a Gaussian Process emulator. We denote the highest run output so far found as $f^+ = f(x^+)$ with $x^+ \in \{x^{(1)}, \ldots, x^{(n)}\}$ the corresponding best input so far.

4.1 Define the Probability of Improvement $PI(x)$ and the Expected Improvement $EI(x)$ acquisition functions.

4.2 Say that only a single run has been performed at $x^{(1)} = 0.7$ yielding $D = f(x^{(1)}) = 2$. We employ a Gaussian Process emulator with a squared exponential covariance structure, with $\sigma = 1$ and $\theta = 0.5$ and with constant prior expectation $\mathrm{E}[f(x)] = 0$. By first evaluating the emulator expectation and variance, show that in this case $PI(x)$ is given by:

$$PI(x) \;\; = \;\; \Phi\big[-2\tanh\{2(x-0.7)^2\}^{1/2}\big], \quad \text{for } x \neq 0.7$$

where $\Phi[.]$ is the c.d.f. of the standard normal distribution.

4.3 By examining the derivative of $PI(x)$, determine whether $PI(x)$ attains a maximum in $[0, 1]$ and find $\lim_{x \to 0.7} d[PI(x)]/dx$.

4.4 If we used the $PI(x)$ acquisition function, where would it suggest we place the second run? Discuss the weaknesses of the $PI(x)$ acquisition function, with reference to this example.

4.5 Show that when using a general Gaussian Process emulator, the Expected Improvement acquisition function becomes:

$$EI(x) \;\; = \;\; \sigma_D(x)\phi(z^*) \; + \; \big(\mu_D(x) - f^+\big)\Phi(z^*),$$

where $\mu_D(x)$ and $\sigma_D(x)$ denote the emulator mean and standard deviation updated by the runs D, $z^* = (\mu_D(x) - f^+)/\sigma_D(x)$ and $\phi(.)$ denotes the p.d.f. of the standard normal distribution.

4.6 By re-expressing both acquisition functions in terms of utilities, discuss why $EI(x)$ is generally considered superior to $PI(x)$. However, what fundamental weakness do both acquisition functions possess?