Lecture 4: Some Interesting Applications of Gaussian Process Regression

Topics
- GP quadrature
- Statistical inverse problem with GP surrogates
- Physics-informed spatial modeling

Gaussian Process Quadrates

GP quadrature treats numerical integration as a statistical inference problem.

Goal: estimate a definite integral $I = \int_a^b g(t)\,dt$ where $g(t)$ has a known functional form.

Numerical Quadrature: uses quadrature approximation

$$\int_a^b g(t)\,dt \approx \sum_{i=0}^{n} w_i\, g(t_i)$$

- t_i : locations on $[a,b]$ where g is evaluated
- w_i : numerical coefficients assigned to each function value according to a quadrature rule

It can be shown that different quadrature rules can integrate certain polynomials exactly. But, for other functions g there is approximation error.

Gaussia Process Quadrature ("uncertainty-aware approximation")

First, recognize that even though we know functional form of g, we do not know its integral I. A Bayesian would first put a prior on I. But since g is easier to model and I is a linear transformation of g, we define the prior on g first.

$$g \sim GP\left(m_0(t),\, C_0(t,s)\right) \qquad C_0 : [a,b]\times[a,b] \to \mathbb{R}$$

This implies the following joint prior

$$\begin{bmatrix} g \\ \cdots \end{bmatrix} \sim GP\left(\begin{bmatrix} m_0(t) \\ \cdots \end{bmatrix}, \begin{bmatrix} C_0(t,s) & \int_a^b C_0(t,s)\,ds \\ \cdots & \cdots \end{bmatrix} \right)$$

$$\begin{bmatrix} g \\ I \end{bmatrix} \sim GP\left(\begin{bmatrix} m_0(t) \\ \int_a^b m_0(t)\,dt \end{bmatrix}, \begin{bmatrix} C_0(t,s) & \int_a^b C_0(t,s)\,ds \\ \int_a^b C_0(t,s)\,dt & \int_a^b\int_a^b C_0(t,s)\,dt\,ds \end{bmatrix} \right)$$

Just like in the numerical case, we evaluate g at the nodes $t_{obs} = (t_1, \dots, t_n)^T$. Note, there is no observation error in this case. $\left(g(t_{obs}) = (g(t_1), \dots, g(t_n))^T \right)$

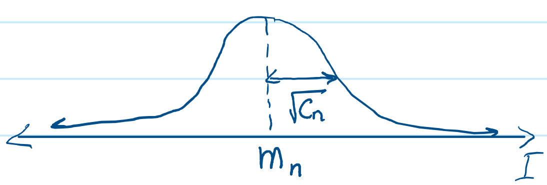
$$\begin{bmatrix} g(t_{obs}) \\ I \end{bmatrix} \sim N\left(\begin{bmatrix} m_0(t_{obs}) \\ \int_a^b m_0(t)\,dt \end{bmatrix}, \begin{bmatrix} C_0(t_{obs}, t_{obs}) & \int_a^b C_0(t_{obs}, s)\,ds \\ \int_a^b C_0(t, t_{obs})\,dt & \int_a^b\int_a^b C_0(t,s)\,dt\,ds \end{bmatrix} \right)$$

By properties of MVN random vectors, we have

$$I \mid g(t_{obs}) \sim N_1(m_n, C_n) \qquad m_n \in \mathbb{R},\ C_n \in \mathbb{R}^+$$

- $m_n = \int_a^b C_0(t, t_{obs})\,dt \left[C_0(t_{obs}, t_{obs}) \right]^{-1} g(t_{obs})$
- $C_n = \int_a^b\int_a^b C_0(t,s)\,dt\,ds - \int_a^b C_0(t, t_{obs})\,dt \left[C_0(t_{obs}, t_{obs}) \right]^{-1} \int_a^b C_0(t_{obs}, s)\,ds$

- posterior mean provides an approximation of I
- posterior variance of I provides some notion of uncertainty that is related to prior choices and the number of nodes.



- Probabilisticnumerics.org

GP Surrogates within the Statistical Inverse Problem

Forward problem: – given some inputs (e.g. physical parameter, inital states), predict the behaviour of a system

Forward problem: - given some inputs (e.g. physical parameters, initial states), predict the behaviour of a system
- Typically involves time-intensive computation
- Example: nanoindentation displacement

Statistical inverse problem: Given some measurements of the system, recover the unknown inputs and quantify uncertainty from all sources
- Example: GP regression where we recover an unknown function based on some observations

Bayesian Hierarchical Model (BHM)

BHM is a multi-level statistical model connected conditionally

Example: Top layer might be a likelihood

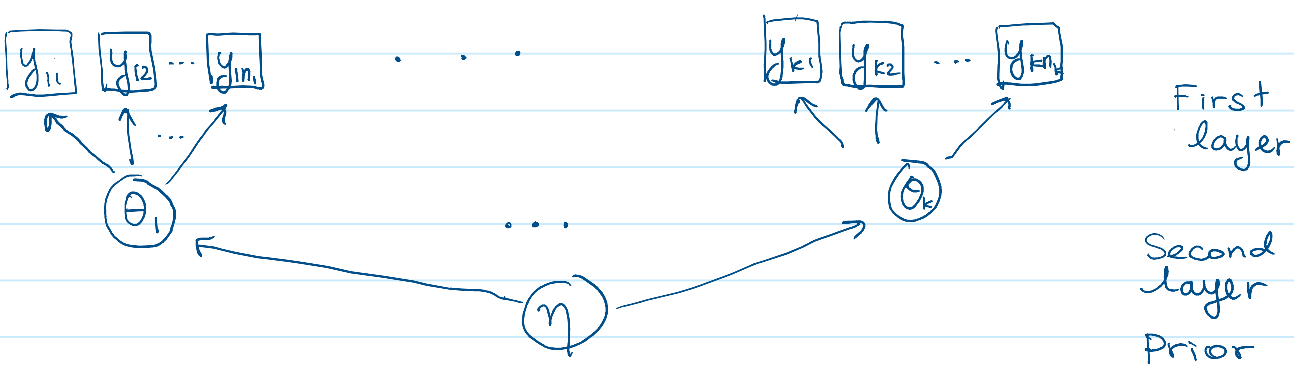
$$Y_{ij} \overset{ind}{\sim} f(y_{ij} \mid \theta_i) \qquad \underbrace{i = 1, \ldots, k}_{\text{group index}}, \quad \underbrace{j = 1, \ldots, n_i}_{\text{individual index}}$$

Second layer might be a model for $\theta = (\theta_1, \ldots, \theta_k)^T$

$$\theta_i \overset{ind}{\sim} f(\theta_i \mid \eta) \qquad i = 1, \ldots, k$$

Final layer might be a prior on η

$$\eta \sim f(\eta)$$



Posterior distribution has density

$$f\left(\eta, \theta_1, \ldots, \theta_k \mid \{y_{ij}\}_{\substack{i=1,\ldots,k \\ j=1,\ldots,n_i}}\right)$$

$$f\left(\eta, \theta_1, \ldots, \theta_k \,\middle|\, \{y_{ij}\}_{\substack{i=1,\ldots,k\\ j=1,\ldots,n_i}}\right)$$

$$\propto \underbrace{f\left(\{y_{ij}\}_{\substack{i=1,\ldots,k\\ j=1,\ldots,n_i}} \,\middle|\, \theta_1, \ldots, \theta_k\right)}_{\text{likelihood}} \underbrace{f(\theta_1, \ldots, \theta_k \mid \eta)}_{\text{middle layer}} \underbrace{f(\eta)}_{\text{prior}}$$

Cannot evaluate functionals of posterior in closed form, so we use MCMC [Gelman et.al. Bayesian Data Analysis]

Surrogate models for inference

Consider the following data-generating mechanism

true, unknown parameters

$$Y_i = m(t_i; \theta) + \varepsilon_i, \qquad \varepsilon_i \overset{\text{ind}}{\sim} N(0, \sigma^2) \quad i = 1, \ldots, n$$

constitutive model
(expensive forward model)

We need to interrogate m repeatedly for different θ, we create a surrogate model based on a finite number of input-output pairs

$$\{\theta_j, m(\theta_j)\}_{j=1,\ldots,m}$$

Construct statistical model of the expensive model m that keep track of uncertainty between training locations

$$f(m(\cdot) \mid m(\theta_1), \ldots, m(\theta_m))$$

We achieve this using GP regression by putting a GP prior on m

$$m \sim GP(m_0, C_0)$$

$m_0 : \Theta \to \mathbb{R}$
$C_0 : \Theta \times \Theta \to \mathbb{R}^+$
positive semi-definite

The posterior

$$m \mid m(\theta_1), \ldots, m(\theta_m) \sim GP(m_n, C_n)$$

$$m \mid m(\theta_1), \dots, m(\theta_m) \sim GP(m_n, C_n)$$

where for $\theta_{obs} = (\theta_1, \dots, \theta_m)^T$

$$m_n(\theta) = C_0(\theta, \theta_{obs}) [C_0(\theta_{obs}, \theta_{obs})]^{-1} m(\theta_{obs})$$

$$C_n(\theta, \theta^*) = C_0(\theta, \theta^*) - C_0(\theta, \theta_{obs}) [C_0(\theta_{obs}, \theta_{obs})]^{-1} C_0(\theta_{obs}, \theta^*)$$

This gives the middle layer of our BHM

$$f(\theta, m(\theta) \mid y_1, \dots, y_n, m(\theta_1), \dots, m(\theta_m))$$

$$\propto \underbrace{f(y_1, \dots, y_n \mid m(\theta))}_{\text{likelihood}} \underbrace{f(m(\theta) \mid m(\theta_1), \dots, m(\theta_m))}_{\text{GP posterior}} \underbrace{f(\theta)}_{\text{prior}}$$

To solve the statistical inverse problem, we use MCMC to sample from this distribution $\Rightarrow$ Bayesian inference

<u>Physics-informed GPs</u>

arxiv.org/abs/2511.22868