Bayesian Models

Kelechi Akwataghibe

December 2021

Model 1: Gaussian Process Regression

We model the observations y_i as a latent function $f(x_i)$ corrupted by Gaussian noise with variance d_i:

$$y_i = f(x_i) + \epsilon_i, \quad \epsilon_i \sim \mathcal{N}(0, d_i) \tag{1}$$

The latent function is modeled as a Gaussian Process with a zero mean prior and a Squared Exponential (RBF) covariance kernel:

$$f(x) \sim \mathcal{GP}\left(0, k(x, x')\right) \tag{2}$$

$$k(x, x') = \sigma_f^2 \exp\left(-\frac{(x - x')^2}{2\ell^2}\right) \tag{3}$$

where $\sigma_f = 6$ is the signal amplitude and $\ell = 4$ is the length-scale.

Given training data $\mathbf{X}, \mathbf{y}$ and test points $\mathbf{X}_*$, the posterior distribution over functions is:

$$f_* | \mathbf{X}, \mathbf{y}, \mathbf{X}_* \sim \mathcal{N}(\bar{f}_*, \mathsf{cov}(f_*)) \tag{4}$$

Model 2: Bayesian Finite Mixture Model (3D)

We assume the data $D_i = \{x_i, y_i, z_i\}$ arises from one of $K = 3$ latent clusters. Let $s_i \in \{1, 2, 3\}$ be the latent cluster assignment for the i-th observation.

Likelihood

The data dimensions are conditionally independent given the cluster assignment $s_i = k$:

$$x_i | s_i = k \sim \mathcal{N}(\mu_{x,k}, 1) \tag{5}$$

$$y_i | s_i = k \sim \mathcal{N}(\mu_{y,k}, \tau^{-1}) \tag{6}$$

$$z_i | s_i = k \sim \mathsf{Poisson}(\lambda_{z,k}) \tag{7}$$

Priors

We place conjugate priors on the parameters:

$$s_i \sim \mathsf{Categorical}(\pi) \tag{8}$$

$$\pi \sim \mathsf{Dirichlet}(\alpha, \alpha, \alpha) \tag{9}$$

$$\mu_{x,k} \sim \mathcal{N}(0, 1) \tag{10}$$

$$\mu_{y,k} \sim \mathcal{N}(2, \tau^{-1}) \tag{11}$$

$$\lambda_{z,k} \sim \mathsf{Gamma}(\alpha_z, \beta_z) \tag{12}$$

$$\tau \sim \mathsf{Gamma}(\alpha_\tau, \beta_\tau) \tag{13}$$

Model 3: Bayesian Hidden Markov Model (HMM)

Let w_t be the observed data at time t and $s_t \in \{1, 2, 3\}$ be the hidden state.

State Transition

The hidden states follow a first-order Markov chain governed by a transition matrix $\mathbf{P}$:

$$s_t | s_{t-1} = i \sim \text{Categorical}(\mathbf{P}_{i,:}) \tag{14}$$

where $\mathbf{P}_{ij} = \Pr(s_t = j | s_{t-1} = i)$.

Emission Model

The observations are normally distributed with a state-specific mean ϕ_k:

$$w_t | s_t = k \sim \mathcal{N}(\phi_k, 1) \tag{15}$$

Priors

We place priors on the transition rows and emission means:

$$\mathbf{P}_{i,:} \sim \text{Dirichlet}(\beta_1, \beta_2, \beta_3) \tag{16}$$

$$\phi_k \sim \mathcal{N}(\mu_0, \sigma_0^2) \tag{17}$$

Inference

Posterior sampling is performed via Gibbs sampling using the Forward-Filtering Backward-Sampling (FFBS) algorithm to sample the state trajectory $\mathbf{s}_{1:T}$:

$$p(\mathbf{s}_{1:T} | \mathbf{w}_{1:T}) = p(s_T | \mathbf{w}_{1:T}) \prod_{t=T-1}^{1} p(s_t | s_{t+1}, \mathbf{w}_{1:t}) \tag{18}$$