Bayesian inference formalism

William Fauriat
CEA

June 27, 2025

1 Statistical model

The log-likelihood function for N observations of a d-dimensional multivariate Gaussian distribution is given by:

$$\mathcal{L}(\boldsymbol{\beta}, \boldsymbol{\Sigma}) = -\frac{Nd}{2}\ln(2\pi) - \frac{N}{2}\ln(|\boldsymbol{\Sigma}|) - \frac{1}{2}\sum_{i=1}^{N}(\mathbf{y}_i - f(\mathbf{x}_i, \boldsymbol{\beta}))^T \boldsymbol{\Sigma}^{-1}(\mathbf{y}_i - f(\mathbf{x}_i, \boldsymbol{\beta}))$$

$f(x, \boldsymbol{\beta})$ is the generative model (deterministic here)

Σ is the covariance matrix such that observations $\mathbf{y} \sim \mathcal{N}(f(\mathbf{x}, \boldsymbol{\beta}), \Sigma)$
Σ can be $\Sigma = \sigma\mathbb{I} + \mathrm{Cov}(\mathbf{y}, \mathbf{y}')$
To add model error add $\Sigma_{GP} = \mathrm{Cov}(f(x, \beta), f(x', \beta)) + \eta$

1.1 Note when inference is directly about the multivariate parameters

$$p(\mu, \Sigma | X) \propto p(X | \mu, \Sigma) \times p(\mu, \Sigma)$$

$$\mathcal{L}(\boldsymbol{\mu}, \boldsymbol{\Sigma}) = -\frac{Nd}{2}\ln(2\pi) - \frac{N}{2}\ln(|\boldsymbol{\Sigma}|) - \frac{1}{2}\sum_{i=1}^{N}(\mathbf{x}_i - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\mathbf{x}_i - \boldsymbol{\mu})$$

So here when comparing notations:

$\mu \leftrightarrow f(x, \beta) \in \mathbb{R}^d$ and $x_i \leftrightarrow y_i \in \mathbb{R}^d$??

and generally we have just one observation (one set) $N = 1$ of (potentially correlated) experiences defined by $(x)_{i=1,\dots,d}$ and a descriptive model $f(x, \beta)$ with $\beta \in \mathbb{R}^p$

2 Inference

2.1 Metropolis Hastings ratio

$$\alpha(x, x') = \min\left(1, \frac{\pi(x')q(x|x')}{\pi(x)q(x'|x)}\right)$$