# 4M24 CW - High-Dimensional MCMC

Candidate Number: 5488A

December 2024

## 1 Simulation

### a Gaussian Process Prior

Our prior is a Gaussian Process has zero mean and a squared exponential covariance kernel, $k(\boldsymbol{x}, \boldsymbol{x'})$, with length scale $\ell$. The coordinates, $\{\boldsymbol{x_n}\}_{n=1}^{N}$, of our samples at placed on a regular $D \times D$ grid in $[0, 1]^2$.

$$k(\boldsymbol{x}, \boldsymbol{x'}) = \exp\left(-\frac{\|\boldsymbol{x} - \boldsymbol{x'}\|^2}{2\ell^2}\right) \tag{1}$$

Our samples, $\boldsymbol{u}$, collected into an $N \times 1$ vector and is distributed $\boldsymbol{u} \sim \mathcal{N}(\boldsymbol{0}, C)$, where $C$ is the $N \times N$ covariance matrix with entries $C_{ij} = k(\boldsymbol{x_i}, \boldsymbol{x_j})$.

Samples from this prior are shown in Figure 1 for 3 values of $\ell$. Larger values result in a smoother surface with more correlation between nearby points.

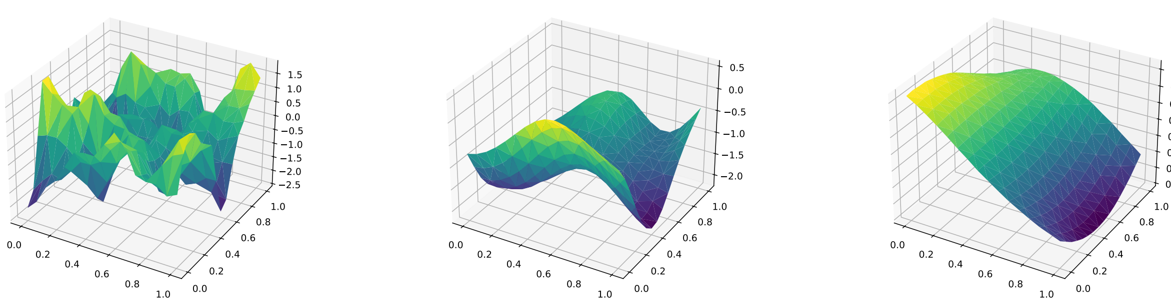
We subsample the grid with $M$ uniform random draws and apply independent Gaussian measurement noise, $\boldsymbol{\epsilon}$, to the observations. This subsampling can be captured by the matrix $M \times N$ matrix $G$ with entries $G_{ij} = 1$ if the $i$th observation is at the $j$th grid point and 0 otherwise. The observations, $\boldsymbol{v}$. We also define the subsampling factor $f := N/M$.

$$\boldsymbol{v} = G\boldsymbol{u} + \boldsymbol{\epsilon} \quad \boldsymbol{\epsilon} \sim \mathcal{N}(\boldsymbol{0}, I) \tag{2}$$

One sample is produced from this model with $D = 16$, $f = 4$ and $\ell = 0.3$ to be used as our dataset for the analysis within this section. Figure 2 shows the latent surface, $\boldsymbol{u}$, and $M = \frac{N}{f} = 64$ noisy observations, $\boldsymbol{v}$.

### b Likelihoods and MCMC

We now proceed to infer the latent surface, $\boldsymbol{u}$, from the noisy observations, $\boldsymbol{v}$, using MCMC. To compute our posterior we need to evaluate the likelihood, $p(\boldsymbol{v}|\boldsymbol{u})$, and the prior, $p(\boldsymbol{u})$. The form of the prior was given previously



(a) $\ell = 0.1$ (b) $\ell = 0.3$ (c) $\ell = 1$

Figure 1: Samples from the Gaussian Process Prior

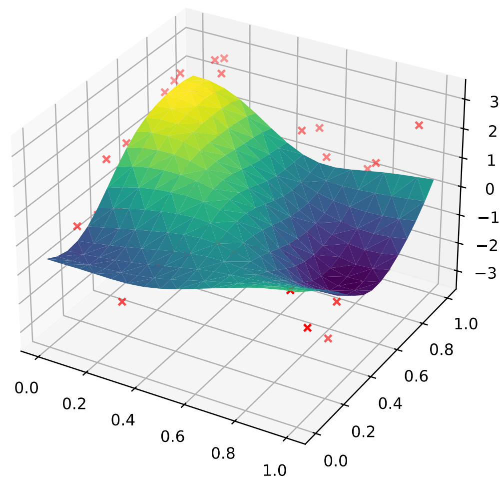



Figure 2: Simulated dataset: $\boldsymbol{v}$ - red crosses, $\boldsymbol{u}$ - surface

but is repeated below and its logarithm can be computed with simple algegbraic manimpulation.

$$\begin{aligned} \boldsymbol{u} &\sim \mathcal{N}(\mathbf{0}, K) \\ \ln p(\boldsymbol{u}) &= -\frac{N}{2}\ln(2\pi) - \frac{1}{2}\ln(|K|) - \frac{1}{2}\boldsymbol{v}^T K^{-1}\boldsymbol{v} \\ &= -\frac{1}{2}\boldsymbol{v}^T K^{-1}\boldsymbol{v} + \text{const} \end{aligned} \tag{3}$$

Likewise the likelihood is given below.

$$\begin{aligned} \boldsymbol{v}|\boldsymbol{u} &\sim \mathcal{N}(G\boldsymbol{u}, I) \\ \ln p(\boldsymbol{v}|\boldsymbol{u}) &= -\frac{M}{2}\ln(2\pi) - \frac{1}{2}\ln(|I|) - \frac{1}{2}(\boldsymbol{v} - G\boldsymbol{u})^T(\boldsymbol{v} - G\boldsymbol{u}) \\ &= -\frac{1}{2}(\boldsymbol{v} - G\boldsymbol{u})^T(\boldsymbol{v} - G\boldsymbol{u}) + \text{const} \end{aligned} \tag{4}$$

Computation of the posterior is straightforward using Baye's rule. Note that we only need to compute the log-prior and log-likelihood up to a constant which greatly saves on computation.

$$p(\boldsymbol{u}|\boldsymbol{v}) \propto p(\boldsymbol{v}|\boldsymbol{u})p(\boldsymbol{u}) \therefore \ln p(\boldsymbol{u}|\boldsymbol{v}) = \ln p(\boldsymbol{v}|\boldsymbol{u}) + \ln p(\boldsymbol{u}) + \text{const} \tag{5}$$

We now consider two MCMC algorithms for generating samples from the posterior.

### b.1 Gaussian random walk Metropolis-Hastings