Description

1 Method

Denote the design space as $\mathcal{X}$. Given M models $\{\mathcal{M}_i\}_{i=1}^M$, each with parameters $\theta_i \subseteq \Theta_i$ and prior distribution $p(\mathcal{M}_i)$, we first give each θ_i a (multivariate Gaussian) prior $p(\theta_i|\mathcal{M}_i)$. (For simplicity, we assume that one of $\{\mathcal{M}_i\}_{i=1}^M$ is the ground-truth, i.e., $\mathcal{M}_{\text{true}}$.) Then we find several local maxima of the density $p(\theta_i|\mathcal{M}_i)$, denoted by $\{\theta_{i,s}^{\text{MAP}}\}_{s=1}^{K_i}$, using HMC, and estimate the response $y_{i,s}(x) \triangleq \mathcal{M}_i(x;\theta_{i,s}^{\text{MAP}}) + \epsilon_{i,s}$, where $\{\epsilon_{i,s}\}_{i\in[M],s\in[K_i]} \overset{iid}{\sim} \mathcal{N}(0,\sigma_{\text{n}}^2)$. Thus $y_{i,s}(x) \sim \mathcal{N}(\mathcal{M}_i(x;\theta_{i,s}^{\text{MAP}}),\sigma_{\text{n}}^2)$. For any $(i,s) \in [M]\times[K_i]$ and $(j,t) \in [M]\times[K_j]$, compute

$$
\begin{aligned}
D_{(i,s),(j,t)}(x) &\triangleq D_{\text{KL}}\left(\mathcal{N}(\mathcal{M}_i(x;\theta_{i,s}^{\text{MAP}}),\sigma_{\text{n}}^2),\mathcal{N}(\mathcal{M}_j(x;\theta_{j,t}^{\text{MAP}}),\sigma_{\text{n}}^2)\right)\\
&= \frac{\left(\mathcal{M}_i(x;\theta_{i,s}^{\text{MAP}}),\sigma_{\text{n}}^2) - \mathcal{M}_j(x;\theta_{j,t}^{\text{MAP}}),\sigma_{\text{n}}^2)\right)^2}{2\sigma_{\text{n}}^2}.
\end{aligned}
$$

We choose the design point x^* to be a local minimum of

$$S(x) \triangleq -\log\det D(x).$$

Then we simulate the response at x^*, i.e., $y(x^*)$, by

$$y(x^*) \triangleq \mathcal{M}_{\text{true}}(x^*) + \epsilon.$$

With the data pair $(x^*, y(x^*))$, we obtain the log-likelihood (up to some constants)

$$\log p\left((x^*,y(x^*))|\theta_i,\mathcal{M}_i\right) = -\frac{(y(x^*) - \mathcal{M}_i(x^*;\theta_i))^2}{2\sigma_{\text{n}}^2}.$$

Then the log-posterior (again, up to some constants) is given by

$$\log p\left(\theta_i|(x^*,y(x^*)),\mathcal{M}_i\right) = \log p(\theta_i|\mathcal{M}_i) + \log p\left((x^*,y(x^*))|\theta_i,\mathcal{M}_i\right).$$

2 Test Model

We take equation (I.24.6) from Feynman's lecture notes, which is

$$E = cm^{e_1}(\omega^{e_2} + \omega_0^{e_3})z^{e_4}, \tag{1}$$

where $c = 1/4$, $e_1 = 1$ and $e_2 = e_3 = e_4 = 2$. This model has four inputs $x \triangleq (m, \omega, \omega_0, z)$ and five parameters $\theta \triangleq (c, e_1, e_2, e_3, e_4)$. We use three candidate models, the first of which is the ground-truth model in (1). The other two models are

$$E = cm^{e_1}\omega^{e_2}\omega_0^{e_3}z^{e_4}, \tag{2}$$

$$E = cm^{e_1}(\omega^{e_2} + z^{e_4})\omega_0^{e_3}. \tag{3}$$

We can encode the initial values of the parameters of each model, say θ_i in $\mathcal{M}_i$, in the prior distribution $p(\theta_i|\mathcal{M}_i)$.