Bayesian approach of the parameter inverse problem under uncertainties

Victor Trappler

Directeurs de Thèse : Arthur VIDARD (Inria)
Élise ARNAUD (UGA)
Laurent DEBREU (Inria)

15 janvier 2020

Table des matières

(Joint) Posterior formulation

Priors

$$
\begin{aligned}
K &\sim \mathcal{U}(\mathbb{K}), \quad p(k) \\
U &\sim \mathcal{U}(\mathbb{U}), \quad p(u)
\end{aligned}
$$

Likelihood model

$$\begin{aligned} p(y \mid k, u, \sigma^2) &= \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{1}{2\sigma^2} SS(k,u)\right] \\ &= \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{1}{2\sigma^2} \|\mathcal{M}(k,u) - y\|_{\Sigma}^2\right] \end{aligned}$$

Now to Bayes' theorem

$$p(k, u \mid y, \sigma^2) = \frac{p(y \mid k, u, \sigma^2)p(k,u)}{\iint_{\mathbb{K}\times\mathbb{U}} p(y \mid k, u, \sigma^2)p(k,u)\,\mathrm{d}(k,u)}$$

Let us assume an hyperprior for σ^2 : $p(\sigma^2)$

GP for sequential adaptative design, and Relative-regret based family of estimators

Random processes

Let us assume that we have a map f from a p dimensional space to $\mathbb{R}$:

$$\begin{array}{rccc} f: & \mathbb{X} \subset \mathbb{R}^p & \longrightarrow & \mathbb{R} \\ & x & \longmapsto & f(x) \end{array} \tag{1}$$

This function is assumed to have been evaluated on a design of n points, $\mathcal{X} \subset \mathbb{X}^n$. We wish to have a probabilistic modelling of this function We introduce random processes as way to have a prior distribution on function

GP modelling of the penalized cost function

GP processes

Let $\Delta_\alpha(\mathbf{k},\mathbf{u}) = J(\mathbf{k},\mathbf{u}) - \alpha J^*(\mathbf{u})$. Furthermore, we assume that we constructed a GP on J on the joint space $\mathbb{K}\times\mathbb{U}$, based on a design $\mathcal{X} = \left\{(\mathbf{k}^{(1)},\mathbf{u}^{(1)}),\dots,(\mathbf{k}^{(n)},\mathbf{u}^{(n)})\right\}$, denoted as $(\mathbf{k},\mathbf{u}) \mapsto Y(\mathbf{k},\mathbf{u})$.

As a GP, Y is described by its mean function m_Y and its covariance function $\kappa_Y(\cdot,\cdot)$:

$$Y(\mathbf{k},\mathbf{u}) \sim \mathcal{N}\left(m_Y(\mathbf{k},\mathbf{u}), \sigma_Y^2(\mathbf{k},\mathbf{u})\right) \tag{2}$$

Analogous to J and J^*, we define Y^* as

$$Y^*(\mathbf{u}) \sim \mathcal{N}\left(m_Y^*(\mathbf{u}), \sigma_Y^{2,*}(\mathbf{u})\right) \tag{3}$$

Then we have

$$\Delta_\alpha(\mathbf{k},\mathbf{u}) \sim \mathcal{N}\left(m_Y(\mathbf{k},\mathbf{u}) - \alpha m_Y^*(\mathbf{u}), \sigma_Y^2(\mathbf{k},\mathbf{u}) + \alpha^2\sigma_Y^{2,*}(\mathbf{u})\right) \tag{4}$$

Definition of m_Y^* ?

$$J^*(\mathbf{u}) = J(\mathbf{k}^*(\mathbf{u}), \mathbf{u}) = \min_{\mathbf{k}\in\mathbb{K}} J(\mathbf{k}, \mathbf{u}) \tag{5}$$

As J^* is unknown, we can use first use a plug-in approach, and define

$$m_Y^*(\mathbf{u}) = \min_{\mathbf{k}\in\mathbb{K}} m_Y(\mathbf{k}, \mathbf{u}) \tag{6}$$

The surrogate conditional minimiser is used in Ginsbourger profiles etc.

Approximation of the objective probability using GP

We are going now to use a different notation for the probabilities, taken with respect to the GP : $\mathcal{P}$, to represent the uncertainty encompassed by the GP.

Defined somewhere else, we have

$$\begin{align}
\Gamma_\alpha(\mathbf{k}) &= \mathbb{P}_{\mathbf{U}}\left[J(\mathbf{k}, \mathbf{U}) \leq \alpha J^*(\mathbf{U})\right] \tag{7}\\
&= \mathbb{E}_{\mathbf{U}}\left[\mathbb{1}_{J(\mathbf{k},\mathbf{U})\leq \alpha J^*(\mathbf{U})}\right] \tag{8}
\end{align}$$

This classification problem can be approached with a plug-in approach, or a probablistic one :

$$\begin{align}
\mathbb{1}_{J(\mathbf{k},\mathbf{u})\leq \alpha J^*(\mathbf{u})} &\approx \mathbb{1}_{m_Y(\mathbf{k},\mathbf{u})\leq \alpha m_Y^*(\mathbf{u})} \tag{9}\\
\mathbb{1}_{J(\mathbf{k},\mathbf{u})\leq \alpha J^*(\mathbf{u})} &\approx \mathcal{P}\left[\Delta_\alpha(\mathbf{k}, \mathbf{u}) \leq 0\right] = \pi(\mathbf{k}, \mathbf{u}) \tag{10}
\end{align}$$

Using the GPs, for a given $\mathbf{k}$, α and $\mathbf{u}$, the probability for our meta model to verify the inequality is given by Based on those two approximation, the approximated probability Γ is

$$\begin{align}
\hat{\Gamma}_{\alpha,n}(\mathbf{k}) &= \mathbb{P}_U\left[m_Y(\mathbf{k}, \mathbf{u}) \leq \alpha m_Y^*(\mathbf{u})\right] && \text{(plug-in)}\\
\hat{\Gamma}_{\alpha,n}(\mathbf{k}) &= \mathbb{E}_U\left[\mathcal{P}\left[\Delta_\alpha(\mathbf{k}, \mathbf{u}) \leq 0\right]\right] && \text{(Probabilistic approx)}
\end{align} \tag{11}$$

Considering the joint distribution of $Y(\mathbf{k}, \mathbf{u})$ and $Y^*(\mathbf{u}) = Y(\mathbf{k}^*(\mathbf{u}), \mathbf{u})$, we have

$$\begin{bmatrix} Y(\mathbf{k}, \mathbf{u}) \\ Y^*(\mathbf{u}) \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} m_Y(\mathbf{k}, \mathbf{u}) \\ m_Y^*(\mathbf{u}) \end{bmatrix}; \begin{bmatrix} C\left((\mathbf{k}, \mathbf{u}), (\mathbf{k}, \mathbf{u})\right) & C\left((\mathbf{k}, \mathbf{u}), (\mathbf{k}^*(\mathbf{u}), \mathbf{u})\right) \\ C\left((\mathbf{k}, \mathbf{u}), (\mathbf{k}^*(\mathbf{u}), \mathbf{u})\right) & C\left((\mathbf{k}^*(\mathbf{u}), \mathbf{u}), (\mathbf{k}^*(\mathbf{u}), \mathbf{u})\right) \end{bmatrix}\right) \tag{12}$$

By multiplying by the matrix $[1 \quad -\alpha]$ yields

$$\begin{align}
\Delta_\alpha(\mathbf{k}, \mathbf{u}) &\sim \mathcal{N}\left(m_\Delta(\mathbf{k}, \mathbf{u}); \sigma^2_\Delta(\mathbf{k}, \mathbf{u})\right) \tag{13}\\
m_\Delta(\mathbf{k}, \mathbf{u}) &= m_Y(\mathbf{k}, \mathbf{u}) - \alpha m_Y^*(\mathbf{u}) \tag{14}\\
\sigma^2_\Delta(\mathbf{k}, \mathbf{u}) &= \sigma^2_Y(\mathbf{k}, \mathbf{u}) + \alpha^2\sigma^2_{Y^*}(\mathbf{k}, \mathbf{u}) - 2\alpha C\left((\mathbf{k}, \mathbf{u}), (\mathbf{k}^*(\mathbf{u}), \mathbf{u})\right) \tag{15}
\end{align}$$

The probability of coverage for the set $\{Y - \alpha Y^*\}$ is π, and can be computed using the CDF of the standard normal distribution Φ :

$$\pi(\mathbf{k}, \mathbf{u}) = \Phi\left(-\frac{m_{\Delta_\alpha}(\mathbf{k}, \mathbf{u})}{\sigma_{\Delta_\alpha}(\mathbf{k}, \mathbf{u})}\right) \tag{16}$$