Gaussian type observation error

February 20, 2024

1 Recap how likelihood works in UQlab

We assume the discrepancy model follow Gaussian distribution with unknown residuals $\boldsymbol{\epsilon}$. Likelihood function can be expressed as:

$$\begin{aligned}\mathcal{L}(\boldsymbol{\theta}|\mathcal{Y}) &= \prod_{i=1}^{N} N(\boldsymbol{y_i}|\mathcal{M}(\boldsymbol{\theta}), \boldsymbol{\Sigma}) \\ &= \prod_{i=1}^{N} \frac{1}{\sqrt{(2\pi)^{N_{\text{out}}}\det(\boldsymbol{\Sigma})}} \exp\left(-\frac{1}{2}\left(\boldsymbol{y_i} - \mathcal{M}(\boldsymbol{\theta})\right)^{\intercal} \boldsymbol{\Sigma}^{-1} \left(\boldsymbol{y_i} - \mathcal{M}(\boldsymbol{\theta})\right)\right)\end{aligned} \tag{1}$$

Because $\boldsymbol{\epsilon}$ is unknown, it is considered in the prior $\vec{\theta}$ with predefined distribution $\boldsymbol{\epsilon} \sim \mathcal{N}(0, \sigma^2)$. $\vec{\theta}$ is a vector of parameters of interests (soil parameters and $\boldsymbol{\epsilon}$).

Note: All measurement data points or covariance matrix $\boldsymbol{\Sigma}^{-1}$ default in UQLAB share i.i.d. assumption, i.e., all discrepancy models of the multiple points hold the same prior and σ^2.

$\boldsymbol{\epsilon}$ is non-informative, thus, its prior for the variance follows uniform distribution $\sigma \sim \mathcal{U}(0,\sigma_0^2)$, in which σ_0^2 should be related to the observation values.

UQLAB suggests (as I see) adopting its mean value of the monitored data points along the simple beam problem. Also UQLAB assume the the covariance matrix at different locations has no influence on the monitored data error (i.e., iid assumption mentioned above).

Tricky part is how to relax the iid assumption in UQLAB's custom likelihood. Wagner(2020) recommands hyperparameters $\boldsymbol{w}$.

2 How Wagner handle non-diagonal covariance matrix

Relax iid:

$$\Sigma_{ij} = \sigma_i \sigma_j R \tag{2}$$

$$\boldsymbol{\sigma} = (\sigma_1, ..., \sigma_N) \tag{3}$$

$$\sigma_i = \sum_{k=0}^{6} \omega_k \times \psi_k \tag{4}$$

In this way, $\boldsymbol{w}_k$ should be related with obeservation. Therefore, Wagner adopts prior as Figure 1:

X_ε:	X_7	$\mathcal{U}(0, 20)$	$1.00 \cdot 10^1$	2.40	$2.40 \cdot 10^{-1}$	°C
	X_8	$\mathcal{U}(-20, 20)$	0	3.40	–	°C
	⋮	⋮	⋮	⋮	⋮	⋮
	X_{13}	$\mathcal{U}(-20, 20)$	0	3.40	–	°C
	X_{14}	$\mathcal{U}(0, 50)$	$2.50 \cdot 10^1$	3.80	$1.52 \cdot 10^{-1}$	s

Figure 1: Priors for $\boldsymbol{w}_k$

The prior's in Figure 1 values is related with observation and it is set as 20. It is reasonable and easy to implement becasue Wagner is not doing sequential updating and need the prior only once. But when it comes to multiple stages, manually selecting the priors related with observations is troublesome. In another word, hyperparameters in Wagner's paper have units.

3 How we handle this

Based on Wagner's paper, we make the hyper parameters $\boldsymbol{w}_k$ as a normalized value to get rid of the units.

$$\Sigma_{ij} = \sigma_i \sigma_j R \tag{5}$$

$$\boldsymbol{\sigma} = (\sigma_1, ..., \sigma_N) \tag{6}$$

$$y = max(\mathcal{Y}_{obs}) \tag{7}$$

$$\sigma_i = \sum_{k=0}^{6} \omega_k \times \psi_k \times y \tag{8}$$

in which, $\mathcal{Y}_{obs}$ is observation in different stages. In this way, the only prior we need to control is now only normalized $\boldsymbol{w}_k$ without units.