IMPORTANT !!! Research idea!!

Gaussian process regression with vague input locations.

Given. known location $x \mapsto y$.

unknown location $\tilde{x} \mapsto \tilde{y}$. $p(\tilde{x})$ prior.

Gaussian process prior for the field.

Inference of the location $\tilde{x}$:

$$p(\tilde{x} \mid \tilde{y}, y, x) = \frac{p(\tilde{x})\, p(\tilde{y}, y \mid \tilde{x}, x)}{\int p(\tilde{x})\, p(\tilde{y}, y \mid \tilde{x}, x)\, d\tilde{x}} \quad \begin{aligned} &= p(\tilde{x}, \tilde{y}, y \mid x) \\ &= p(\tilde{y}, y \mid x) \end{aligned}$$

Prediction for the new y^* at an unseen location x^*:

$$\underbrace{\quad}_{= p(y^* \mid x^*, \tilde{y}, \tilde{x}, y, x) \text{ Regular GPR posterior}}$$

$$p(y^* \mid x^*, \tilde{y}, y, x) = \int \boxed{\frac{p(y^*, \tilde{y}, y \mid x^*, \tilde{x}, x)}{p(\tilde{y}, y \mid \tilde{x}, x)}}\; \underbrace{p(\tilde{x} \mid \tilde{y}, y, x)}_{\text{posterior } \tilde{x}}\, d\tilde{x}$$

$$= \frac{\int p(\tilde{x})\, p(y^*, \tilde{y}, y \mid x^*, \tilde{x}, x)\, d\tilde{x}}{\int p(\tilde{x})\, p(\tilde{y}, y \mid \tilde{x}, x)\, d\tilde{x}} \quad \begin{aligned} &= p(y^*, \tilde{y}, y \mid x^*, x) \\ &= p(\tilde{y}, y \mid x) \end{aligned}$$

MAP approximation: $\quad GP(0, k(\cdot,\cdot))$

$$\tilde{x}_0 = \max_{\tilde{x}} p(\tilde{x})\, p(\tilde{y}, y \mid \tilde{x}, x)$$

$$= \max_{\tilde{x}} \exp\left(-\frac{1}{2} \|\tilde{x} - x_0\|^2_{A^{-1}}\right) \exp\left[-\frac{1}{2} \begin{Bmatrix} \tilde{y} \\ y \end{Bmatrix}^T \underbrace{k\left(\begin{Bmatrix} \tilde{x} \\ x \end{Bmatrix}, \begin{Bmatrix} \tilde{x} \\ x \end{Bmatrix}\right)}_{(+\sigma_n^2 I)}^{-1} \begin{Bmatrix} \tilde{y} \\ y \end{Bmatrix}\right]$$

in this case: $p(\tilde{x} \mid \tilde{y}, y, x) \approx \delta(\tilde{x} - \tilde{x}_0)$.

$$p(y^* \mid x^*, \tilde{y}, y, x) = \int \frac{p(y^*, \tilde{y}, y \mid x^*, \tilde{x}, x)}{p(\tilde{y}, y \mid \tilde{x}, x)}\, \delta(\tilde{x} - \tilde{x}_0)\, d\tilde{x} = \underbrace{\frac{p(y^*, \tilde{y}, y \mid x^*, \tilde{x}_0, x)}{p(\tilde{y}, y \mid \tilde{x}_0, x)}}_{\text{GPR posterior}}$$

Marginal likelihood.

$$p(\tilde{y}, y \mid x) = \int p(\tilde{x})\, p(\tilde{y}, y \mid \tilde{x}, x)\, d\tilde{x}$$ intractable.

Variational inference for $p(\tilde{x} \mid \tilde{y}, y, x)$. approximated by normal distribution

Hermite quadrature for the Bayesian prediction

$$y^* \mid x^*, \tilde{y}, y, x \sim \int GP\left(m_{post}(x^*; \tilde{x}, x, \tilde{y}, y), k_{post}(x^*, x^{*\prime}; \tilde{x}, x, \tilde{y}, y)\right) \underbrace{p(\tilde{x} \mid \tilde{y}, y, x)}_{\text{approximated as } \mathcal{N}} d\tilde{x}$$

transformed into independent standard ξ. $x = Q\eta + p$. Standarize.

Hermite quadrature points (W_r, ξ_r)

$$y^* \mid x^*, \tilde{y}, y, x \,\dot{\sim}\, \sum_{r=1}^{R} W_r \, GP\left(m_{post}(x^*; Q\eta_r + p, x, \tilde{y}, y), k_{post}(x^*, x^{*\prime}; Q\eta_r + p, x, \tilde{y}, y)\right).$$

Hermite quadrature for marginal likelihood

$$p(\tilde{y}, y \mid x) \approx \sum_{r=1}^{R} W_r \, p(\tilde{y}, y \mid Q\eta_r + p, x)$$ ok for maxmization

Gaussian mixture ?!

Sparse grid !!

\# Strategy 1: MAP point estimate for $\tilde{x}$.

$\longrightarrow$ Regular GPR with $\tilde{x} = \tilde{x}_0$.

$\boxed{\text{\# Stategy 2:}}$ Variational inference to approximate $\tilde{x} \mid \tilde{y}, y, x$ as $\mathcal{N}$.

BETTER. $\longrightarrow$ Quadrature (Hermite, sparse grid). for $y^* \mid x^*, y, y, x$ marginalizing the posterior $\tilde{x}$.

For both: Quadrature rule to approximate the marginal likelihood $p(\tilde{y}, y \mid x)$.