Multi-fidelity regression with Gaussian processes

Setup: $\mathcal{D} := \{ (X_L, y_L),\ (X_H, y_H) \}$, $X_L \in \mathbb{R}^{n_L \times d}$, $y_L \in \mathbb{R}^{n_L \times 1}$

$n_H \ll n_L$ $\qquad X_H \in \mathbb{R}^{n_H \times d}$, $y_H \in \mathbb{R}^{n_H \times 1}$

1.) **Prior**:

$$y_L = f_L(x_L) + \varepsilon_L \qquad f_L(x) \sim GP(0, K_L(x, x'; \theta_L$$

$$y_H = f_H(x_H) + \varepsilon_H \qquad f_H(x) = \rho f_L(x) + \delta(x$$

$$\varepsilon_L \sim \mathcal{N}(0, \sigma^2_{n_L} I) \qquad \delta(x) \sim GP(0, K_H(x, x'; \theta_H$$

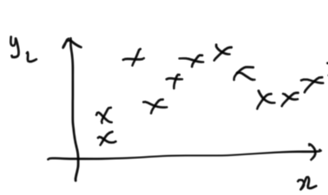
$$\varepsilon_H \sim \mathcal{N}(0, \sigma^2_{n_H} I) \qquad f_L(x) \perp \delta(x)$$

(RBF kernel)

Trainable parameters: $\Theta = \{ \overbrace{\boxed{\sigma^2_{f_L}, \ell_{1_L}, \ldots, \ell_{d_L}}}^{\theta_L}, \overbrace{\boxed{\sigma^2_{f_H}, \ell_{1_H}, \ldots, \ell_{d_H}}}^{\theta_H}, \rho, \sigma^2_{n_L}$

$n_L \times d \quad n_L \times 1 \qquad n_H \times d \quad n_H \times 1$

2.) **Training**: Data $\{X_L, y_L\}$, $\{X_H, y_H\}$



$$p(y_L, y_H \mid X_L, X_H) = \mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \overbrace{\begin{bmatrix} K_{LL} & K_{LH} \\ K_{LH}^T & K_{HH} \end{bmatrix}}^{K = (n_L + n_H) \times (n_L + n} \right)$$

$$K_{LL} = K_L(X_L, X_L; \theta_L) + \sigma^2_{n_L} I \ , \quad \text{size}: n_L \times n_L$$

$$K_{LH} = \text{Cov}[y_L, y_H] = \rho K_L(X_L, X_H; \theta_L) \ , \quad \text{size}: n_L \times n_H$$

$$K_{HH} = \rho^2 K_L(X_H, X_H; \theta_L) + K_H(X_H, X_H, \theta_H) + \sigma^2_{n_H} I \ , \quad \text{size}: n_H \times$$

$$\Rightarrow -\log p(y_L, y_H \mid X_L, X_H) = \frac{1}{2}\log|\mathbb{K}| + \frac{1}{2} y^T \mathbb{K}^{-1} y + \frac{n_L + n_H}{2}\log$$

$\hookrightarrow \mathcal{L}(\theta)$ $\qquad\qquad\qquad\qquad \hookrightarrow y = \begin{bmatrix} y_L \\ y_H \end{bmatrix}$

3. Prediction:

LF pred.:

$$\begin{bmatrix} f_L(x^*) \\ y_L \\ y_H \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \overbrace{k_L(x^*, x^*)}^{k(x^*,x^*)} & \overbrace{k_L(x^*, X_L) \quad \rho k_L(x^*, X_H)}^{k(x^*, X)\ \downarrow} \\ & \mathbb{K} \end{bmatrix}\right.$$

$$p(f_L(x^*) \mid y_L, y_H) \sim \mathcal{N}(\mu_L(x^*), \Sigma_L(x^*))$$

$$\left\{ \begin{aligned} \mu_L(x^*) &= k(x^*, X)\, \mathbb{K}^{-1} y, \quad y = \begin{bmatrix} y_L \\ y_H \end{bmatrix} \\ \Sigma_L(x^*) &= k(x^*, x^*) - k(x^*, X)\, \mathbb{K}^{-1} k(x^*, X)^T \end{aligned} \right\}$$

HF pred.:

$$\begin{bmatrix} f_H(x^*) \\ \begin{bmatrix} y_L \\ y_H \end{bmatrix} \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \overbrace{\boxed{\rho^2 k_L(x^*, x^*) + k_H(x^*, x^*)}}^{k(x^*,x^*)} & \overbrace{\rho k_L(x^*, X_L) \quad \rho^2 k_L(x^*, X_H) + k_H(x^*, X_H)}^{k(x^*, X)\ \downarrow} \\ & \mathbb{K} \end{bmatrix}\right.$$

$$p(f_H(x^*) \mid y_L, y_H) \sim \mathcal{N}(\mu_H(x^*), \Sigma(x^*))$$

$$\mu_H(x^*) = k(x^*, X)\, \mathbb{K}^{-1} y$$

$$\Sigma\ (x^*) = k(x^*, x^*) - k(x^*, X)\, \mathbb{K}^{-1} k(x^*, X)^T$$