1. DEFINITIONS

All definitions given are for 2-dimensional inputs.

1.1. **Full likelihood.**

(1) $$\left[y^*, y_1^\delta, y_2^\delta, y \,|\, \sigma^2, l \right] \sim N\left(0, K\right)$$

(2) $$K = \begin{bmatrix} K^{00}\left(x^*, x^*\right) & K^{10}\left(x_1^\delta, x^*\right)^T & K^{10}\left(x_2^\delta, x^*\right)^T & K^{00}\left(x, x^*\right)^T \\ K^{10}\left(x_1^\delta, x^*\right) & K^{11}\left(x_1^\delta, x_1^\delta\right) & K^{11}\left(x_2^\delta, x_1^\delta\right)^T & K^{01}\left(x, x_1^\delta\right)^T \\ K^{10}\left(x_2^\delta, x^*\right) & K^{11}\left(x_2^\delta, x_1^\delta\right) & K^{11}\left(x_2^\delta, x_2^\delta\right) & K^{01}\left(x, x_2^\delta\right)^T \\ K^{00}\left(x, x^*\right) & K^{01}\left(x, x_1^\delta\right) & K^{01}\left(x, x_2^\delta\right) & K^{00}\left(x, x\right) \end{bmatrix}$$

where

- $y = y(x)$: observed target function values (training set)
- $y^* = y(x^*)$: unobserved function values
- $y_k^\delta = \frac{\partial}{\partial x_k} y\left(x^\delta\right)$: vectors of partial derivatives in the kth input dimension
- σ^2: constant variance parameter
- l: length-scale parameter
- $K^{00}(\cdot, \cdot), K^{10}(\cdot, \cdot), K^{11}(\cdot, \cdot)$: covariance functions. The superscripts indicate the argument where the derivative is taken. Note that $K^{10}\left(x, x'\right) = K^{01}(x', x)^T$.

1.2. **Covariance functions.**

(3) $$K^{00}\left(x, x'\right) = Cov\left(y\left(x\right), y\left(x'\right)\right) = \sigma^2 \prod_{k=1}^{d} g\left(x_k, x_k'; l_k\right)$$

(4) $$K^{10}\left(x, x'\right) = \frac{\partial}{\partial x_i} Cov\left(y\left(x\right), y\left(x'\right)\right) = \sigma^2 \frac{\partial}{\partial x_i} \prod_{k=1}^{d} g\left(x_k, x_k'; l_k\right)$$

(5) $$K^{11}\left(x, x'\right) = \frac{\partial^2}{\partial x_j' \partial x_i} Cov\left(y\left(x\right), y\left(x'\right)\right) = \sigma^2 \frac{\partial^2}{\partial x_j' \partial x_i} \prod_{k=1}^{d} g\left(x_k, x_k'; l_k\right)$$

(6) $$g\left(x, x'\right) = \left[1 + \theta\left|x - x'\right| + \frac{1}{3}\theta^2 \left|x - x'\right|^2\right] \exp\left\{-\theta\left|x - x'\right|\right\}$$

(7) $$\frac{\partial}{\partial x} g\left(x, x'\right) = -\frac{1}{3}\theta^2\left(x - x'\right)\left[1 + \theta\left|x - x'\right|\right] \exp\left\{-\theta\left|x - x'\right|\right\}$$

(8) $$\frac{\partial^2}{\partial x' \partial x} g\left(x, x'\right) = \frac{1}{3}\theta^2 \left[1 + \theta\left|x - x'\right| - \theta^2(x - x')^2\right] \exp\left\{-\theta\left|x - x'\right|\right\}$$

where

- $d = 2$ for 2-dimensional inputs. i.e. x is $n \times 2$ matrix.
- $i, j \in \{1, 2, ..., d\}$
- $g(\cdot, \cdot)$ is from the Matern class of covariance functions with parameter $5/2$; the subscripts for the inputs in $g(\cdot, \cdot)$ are dropped for legibility

- $\theta = \sqrt{5}/l$

See derivatives of $g(\cdot, \cdot)$ in section 3.

2. SETUP

2.1. The posterior density.

$$(9) \qquad \left[l, \sigma^2, y^*, y_1^\delta, y_2^\delta \,|\, y \right] \propto \left[y, y^*, y_1^\delta, y_2^\delta \,|\, \sigma^2, l \right] \left[\sigma^2, l \right]$$

$$(10) \qquad \propto \left[y^*, y_1^\delta, y_2^\delta \,|\, y, \sigma^2, l \right] \left[\left[y \,|\, \sigma^2, l \right] \right] \left[\sigma^2, l \right]$$

The goal is to evaluate densities on the RHS. Specifically,

$$\left[y^*, y_1^\delta, y_2^\delta \,|\, y, \sigma^2, l \right] \sim N(m, S)$$

due to the properties of the multivariate normal distribution (https://en.wikipedia.org/wiki/Multivariate_normal_distribution#Conditional_distributions). That is, the formulas from Kriging follow immediately.

To calculate m, S, start by dividing the covariance matrix K into the appropriate blocks:

$$(11) \qquad K = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}, \text{ where}$$

$$(12) \qquad \Sigma_{11} = \begin{bmatrix} K^{00}(x^*, x^*) & K^{10}(x_1^\delta, x^*)^T & K^{10}(x_2^\delta, x^*)^T \\ K^{10}(x_1^\delta, x^*) & K^{11}(x_1^\delta, x_1^\delta) & K^{11}(x_2^\delta, x_1^\delta)^T \\ K^{10}(x_2^\delta, x^*) & K^{11}(x_2^\delta, x_1^\delta) & K^{11}(x_2^\delta, x_2^\delta) \end{bmatrix}$$

$$(13) \qquad \Sigma_{12} = \Sigma_{21}^T = \begin{bmatrix} K^{00}(x, x^*)^T \\ K^{01}(x, x_1^\delta)^T \\ K^{01}(x, x_2^\delta)^T \end{bmatrix} = \begin{bmatrix} K^{00}(x^*, x) \\ K^{10}(x_1^\delta, x) \\ K^{10}(x_2^\delta, x) \end{bmatrix}$$

$$(14) \qquad \Sigma_{22} = K^{00}(x, x).$$

Then,

$$(15) \qquad m = 0 + \Sigma_{12}\Sigma_{22}^{-1}(y - 0) = \begin{bmatrix} K^{00}(x^*, x) \\ K^{10}(x_1^\delta, x) \\ K^{10}(x_2^\delta, x) \end{bmatrix} K^{00}(x, x)^{-1}y$$

$$(16) \qquad S = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$$

2.2. **Example.** Given

(17)
$$y(x) = 11x_1^{10} + 9x_1^8 + 7x_1^6 + 10x_2^9 + 8x_2^7$$

(18)
$$\frac{\partial}{\partial x_1} y(x) = 110x_1^9 + 72x_1^7 + 42x_1^5$$

(19)
$$\frac{\partial}{\partial x_2} y(x) = 90x_2^8 + 56x_2^6$$

where

- y: 10×1
- x^*: 2×2
- x_1^δ: 8×2
- x_2^δ: 8×2

Then including the GP parameters σ^2 and l, there are 20 unknowns. Fixing GP parameters, we can calculate the covariance matrix S required for $\left[y^*, y_1^\delta, y_2^\delta \, | y, \sigma^2, l \right]$.

2.3. **Problem.** Assuming reasonable values for σ^2 and l, following the formula for S given above, the resulting matrix is negative definite. Adding small nuggets does not fix this.

It seems that if I use rather small values of l (all entries less than approximately 0.3), then this problem disappears.

3. Derivatives for the Matern

Recall: $g(x, x') = \left[1 + \theta \, |x - x'| + \frac{1}{3}\theta^2 |x - x'|^2 \right] \exp\{-\theta \, |x - x'|\}$

3.1. **First derivative.**

(20)
$$\frac{\partial}{\partial x} g(x, x') = \left[\theta sign(x - x') + \frac{2}{3}\theta^2 (x - x') \right] \exp\{-\theta \, |x - x'|\}$$

(21)
$$+ \left[1 + \theta \, |x - x'| + \frac{1}{3}\theta^2 (x - x')^2 \right] \exp\{-\theta \, |x - x'|\} \left(-\theta sign(x - x') \right)$$

(22)
$$= \exp\{-\theta \, |x - x'|\} \left[-\frac{1}{3}\theta^2 (x - x') - \frac{1}{3}\theta^3 (x - x')^2 sign(x - x') \right]$$

(23)
$$= \exp\{-\theta \, |x - x'|\} \left(-\frac{1}{3}\theta^2 (x - x') \right) \left[1 + \theta (x - x') sign(x - x') \right]$$

(24)
$$= -\frac{1}{3}\theta^2 (x - x') \left[1 + \theta \, |x - x'| \right] \exp\{-\theta \, |x - x'|\}$$

3.2. Second derivative.

$$(25) \quad \frac{\partial^2}{\partial x' \partial x} g\left(x, x'\right) = \frac{\partial}{\partial x'}\left[-\frac{1}{3}\theta^2\left(x - x'\right)\left[1 + \theta\left|x - x'\right|\right]\exp\left\{-\theta\left|x - x'\right|\right\}\right]$$

$$(26) \quad = \frac{\partial}{\partial x'}\left[\left[-\frac{1}{3}\theta^2\left(x - x'\right) - \frac{1}{3}\theta^3(x - x')^2 sign\left(x - x'\right)\right]\exp\left\{-\theta\left|x - x'\right|\right\}\right]$$

$$(27) \quad = \left[\frac{1}{3}\theta^2 - \frac{2}{3}\theta^3\left(x - x'\right)\left(-1\right)sign\left(x - x'\right)\right]\exp\left\{-\theta\left|x - x'\right|\right\}$$

$$(28) \quad + \left[-\frac{1}{3}\theta^2\left(x - x'\right) - \frac{1}{3}\theta^3(x - x')^2 sign\left(x - x'\right)\right]\exp\left\{-\theta\left|x - x'\right|\right\}\left(\theta sign\left(x - x'\right)\right)$$

$$(29) \quad = \exp\left\{-\theta\left|x - x'\right|\right\}\left[\frac{1}{3}\theta^2 + \frac{2}{3}\theta^3\left|x - x'\right| - \frac{1}{3}\theta^3\left|x - x'\right| - \frac{1}{3}\theta^4(x - x')^2\right]$$

$$(30) \quad = \exp\left\{-\theta\left|x - x'\right|\right\}\left[\frac{1}{3}\theta^2 + \frac{1}{3}\theta^3\left|x - x'\right| - \frac{1}{3}\theta^4(x - x')^2\right]$$

$$(31) \quad = \frac{1}{3}\theta^2\left[1 + \theta\left|x - x'\right| - \theta^2(x - x')^2\right]\exp\left\{-\theta\left|x - x'\right|\right\}$$