1. DEFINITIONS

All definitions given are for 2-dimensional inputs.

1.1. Assume.

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

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

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.

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

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

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

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

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

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

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.**

$$\left[l, \sigma^2, y^*, y_1^\delta, y_2^\delta \,|y\right] \propto \left[y, y^*, y_1^\delta, y_2^\delta \,|\sigma^2, l\right] \left[\sigma^2, l\right] \tag{9}$$

$$\propto \left[y^*, y_1^\delta, y_2^\delta \,|y, \sigma^2, l\right] \left[\left[y \,|\sigma^2, l\right]\right] \left[\sigma^2, l\right] \tag{10}$$

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

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

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:

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

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

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

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

Then,

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

$$S = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21} \tag{16}$$

2.2. **Example.** Given

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

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

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

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 $\left[y^*, y_1^\delta, y_2^\delta \left| y, \sigma^2, l\right.\right]$.

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

3.1. **First derivative.**

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

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

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

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

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

3.2. **Second derivative.**

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

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

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

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

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

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

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