Kernel parameters

$$\widehat{K}(x_1 - x_2) = \sigma^2 \exp\left(-\frac{(x_1 - x_2)^2}{2l^2}\right) + s^2 \mathbb{I}[x_i = x_j]$$

$$p(f(x_1), f(x_2), \ldots, f(x_n) | \sigma^2, l, s^2) \to \max_{\sigma^2, l, s^2}$$

$$\mathcal{N}(f(x_1), f(x_2), \ldots, f(x_n) | 0, C) \to \max_{\sigma^2, l, s^2}$$

Optimize with gradient ascent