SINGLE NUGGET KRIGING

Minyong R. Lee and Art B. Owen

Stanford University

Supplementary Material

This supplementary material consists of: (1) derivation of the Conditional Maximum Likelihood Estimator (CMLE); (2) proof of Proposition 4; (3) proof of Proposition 6; (4) details on the test functions used in numerical experiments. Equations that are specific to this supplementary file are labeled as (S*). All other referenced equations correspond to those in the main manuscript.

S1 Derivation of the CMLE

By the Woodbury formula,

$$\begin{aligned}\tilde{K}^{-1} &= (K - \mathbf{k}(\mathbf{x}_0)k(\mathbf{x}_0,\mathbf{x}_0)^{-1}\mathbf{k}(\mathbf{x}_0)^T)^{-1} \\ &= K^{-1} + \frac{K^{-1}\mathbf{k}(\mathbf{x}_0)\mathbf{k}(\mathbf{x}_0)^T K^{-1}}{k(\mathbf{x}_0,\mathbf{x}_0) - \mathbf{k}(\mathbf{x}_0)^T K^{-1}\mathbf{k}(\mathbf{x}_0)}.\end{aligned}$$

Therefore,

$$\mathbf{k}(\mathbf{x}_0)^T\tilde{K}^{-1} = \mathbf{k}(\mathbf{x}_0)^T K^{-1} + \frac{\mathbf{k}(\mathbf{x}_0)^T K^{-1}\mathbf{k}(\mathbf{x}_0)\mathbf{k}(\mathbf{x}_0)^T K^{-1}}{k(\mathbf{x}_0,\mathbf{x}_0) - \mathbf{k}(\mathbf{x}_0)^T K^{-1}\mathbf{k}(\mathbf{x}_0)} = \frac{1}{1-\rho(\mathbf{x}_0)^2}\mathbf{k}(\mathbf{x}_0)^T K^{-1}. \tag{S1.1}$$

Thus, by differentiating the conditional log likelihood (3.2) with respect to y_0, we get

$$\frac{\partial l(y_0)}{\partial y_0} = \frac{1}{k(\mathbf{x}_0, \mathbf{x}_0)}(\mathbf{y} - \tilde{m})^T \tilde{K}^{-1}\mathbf{k}(\mathbf{x}_0) = \frac{1}{(1-\rho(\mathbf{x}_0)^2)k(\mathbf{x}_0, \mathbf{x}_0)}(\mathbf{y} - \tilde{m})^T K^{-1}\mathbf{k}(\mathbf{x}_0)$$

from (S1.1). Thus, when $\mathbf{k}(\mathbf{x}_0) \neq 0$, solving $\partial l(y_0)/\partial y_0 = 0$ leads to

$$\hat{y_0} = \beta + \frac{1}{\rho(\mathbf{x}_0)^2}\mathbf{k}(\mathbf{x}_0)^T K^{-1}(\mathbf{y} - \beta\mathbf{1}).$$

S2 Proof of Proposition 4

Proof. From the conditional distribution of $\mathbf{y}$ given $Y(\mathbf{x}_0) = y_0$ ((3.1)),

$$\begin{aligned}
&\mathbb{E}[\,(\hat{Y}_{\mathrm{K}}(\mathbf{x}_0) - Y(\mathbf{x}_0))^2 \,\big|\, Y(\mathbf{x}_0) = y_0\,] \\
&= \mathbf{k}(\mathbf{x}_0)^T K^{-1}\mathbf{k}(\mathbf{x}_0) - \frac{(\mathbf{k}(\mathbf{x}_0)^T K^{-1}\mathbf{k}(\mathbf{x}_0))^2}{k(\mathbf{x}_0, \mathbf{x}_0)} + (y_0 - \beta)^2\left(1 - \frac{\mathbf{k}(\mathbf{x}_0)^T K^{-1}\mathbf{k}(\mathbf{x}_0)}{k(\mathbf{x}_0, \mathbf{x}_0)}\right)^2
\end{aligned}$$

and

$$\begin{aligned}
&\mathbb{E}[\,(\hat{Y}_{\mathrm{SiNK}}(\mathbf{x}_0) - Y(\mathbf{x}_0))^2 \,\big|\, Y(\mathbf{x}_0) = y_0\,] \\
&= k(\mathbf{x}_0, \mathbf{x}_0) - \mathbf{k}(\mathbf{x}_0)^T K^{-1}\mathbf{k}(\mathbf{x}_0) + (y_0 - \beta)^2\left(1 - \sqrt{\frac{\mathbf{k}(\mathbf{x}_0)^T K^{-1}\mathbf{k}(\mathbf{x}_0)}{k(\mathbf{x}_0, \mathbf{x}_0)}}\right)^2.
\end{aligned} \tag{S2.1}$$

Now since $y_0 \sim N(\beta, k(\mathbf{x}_0, \mathbf{x}_0))$,

$$\mathbb{E}[(\hat{Y}_{\mathrm{K}}(\mathbf{x}_0) - Y(\mathbf{x}_0))^2] = k(\mathbf{x}_0, \mathbf{x}_0) - \mathbf{k}(\mathbf{x}_0)^T K^{-1}\mathbf{k}(\mathbf{x}_0) \quad \text{and}$$

$$\begin{aligned}\mathbb{E}[(\hat{Y}_{\text{SiNK}}(\mathbf{x}_0) - Y(\mathbf{x}_0))^2] &= 2k(\mathbf{x}_0, \mathbf{x}_0) - 2\sqrt{k(\mathbf{x}_0, \mathbf{x}_0)\mathbf{k}(\mathbf{x}_0)^T K^{-1}\mathbf{k}(\mathbf{x}_0)} \\ &= \frac{2}{1+\rho(\mathbf{x}_0)}\mathbb{E}[(\hat{Y}_{\text{K}}(\mathbf{x}_0) - Y(\mathbf{x}_0))^2].\end{aligned}$$

by the definition of $\rho(\mathbf{x}_0)$. □

S3 Proof of Proposition 6

Proof. Let $\rho = \rho(\mathbf{x}_0)$. From (S2.1),

$$\mathbb{E}[(\hat{Y}_{\text{K}}(\mathbf{x}_0) - Y(\mathbf{x}_0))^2 \big| S(\mathbf{x}_0)] = k(\mathbf{x}_0, \mathbf{x}_0)(\rho^2 - \rho^4 + S(\mathbf{x}_0)^2(1-\rho^2)^2) \text{ and}$$

$$\mathbb{E}[(\hat{Y}_{\text{SiNK}}(\mathbf{x}_0) - Y(\mathbf{x}_0))^2 \big| S(\mathbf{x}_0)] = k(\mathbf{x}_0, \mathbf{x}_0)(1 - \rho^2 + S(\mathbf{x}_0)^2(1-\rho)^2).$$

Using

$$\mathbb{E}[S(\mathbf{x}_0)^2 \big| S(\mathbf{x}_0) > M] = \frac{1}{1-\Phi(M)}\int_M^\infty s^2\phi(s)ds = \frac{M\phi(M) + 1 - \Phi(M)}{1-\Phi(M)},$$

we get the inequality for $\rho \geq -1 + \sqrt{1 + (1-\Phi(M))/(M\phi(M))}$. □

S4 Test Functions

S4.1 Borehole Function (Morris, Mitchell, and Ylvisaker (1993))

$$f(\mathbf{x}) = \frac{2\pi T_u(H_u - H_l)}{\log(r/r_w)\left(1.5 + \frac{2LT_u}{\log(r/r_w)r_w^2 K_w} + \frac{T_u}{T_l}\right)}$$

The ranges of the eight variables are r_w : (0.05, 0.15), r = (100, 50000), T_u = (63070, 115600), H_u = (990, 1110), T_l = (63.1, 116), H_l = (700, 820), L = (1120, 1680), and K_w = (9855,12045).

S4.2 Welch (Welch et al. (1992))

$$f(\mathbf{x}) = \frac{5x_{12}}{1+x_1} + 5(x_4 - x_{20})^2 + x_5 + 40x_{19}^3 - 5x_{19}$$
$$+ 0.05x_2 + 0.08x_3 - 0.03x_6 + 0.03x_7 - 0.09x_9 - 0.01x_{10} - 0.07x_{11} + 0.25x_{13}^2$$
$$- 0.04x_{14} + 0.06x_{15} - 0.01x_{17} - 0.03x_{18},\ \mathbf{x} \in [-0.5, 0.5]^{20}.$$

S4.3 Friedman (Friedman, Grosse, and Stuetzle (1983))

$$f(\mathbf{x}) = 10\sin(\pi x_1 x_2) + 20(x_3 - 0.5)^2 + 10x_4 + 5x_5,\ \mathbf{x} \in [0,1]^5.$$

S4.4 Robot Arm (An and Owen (2001))

$$f(\mathbf{x}) = (u^2 + v^2)^{0.5}$$

where $u = \sum_{i=1}^{4} L_i \cos\Big(\sum_{j=1}^{i} \theta_j\Big)$, $v = \sum_{i=1}^{4} L_i \sin\Big(\sum_{j=1}^{i} \theta_j\Big)$, and
$\mathbf{x} = (\theta_1, \dots, \theta_4, L_1, \dots, L_4) \in [0, 2\pi]^4 \times [0,1]^4$.

References

An, J. and Owen, A. B. (2001). Quasi-regression. *Journal of Complexity* **17**.4, pp. 588–607.

Friedman, J. H., Grosse, E., and Stuetzle, W. (1983). Multidimensional additive spline approximation. *SIAM Journal on Scientific and Statistical Computing* **4**.2, pp. 291–301.

Morris, M. D., Mitchell, T. J., and Ylvisaker, D. (1993). Bayesian design and analysis of computer experiments: Use of derivatives in surface prediction. *Technometrics* **35**.3, pp. 243–255.

Welch, W. J., Buck, R. J., Sacks, J., Wynn, H. P., Mitchell, T. J., and Morris, M. D. (1992). Screening, predicting, and computer experiments. *Technometrics* **34**.1, pp. 15–25.