Formula Sheet: Richard L. Smith

$$\begin{pmatrix} Y \\ Y_0 \end{pmatrix} \sim N\left[\begin{pmatrix} X\beta \\ x_0^T\beta \end{pmatrix}, \begin{pmatrix} V & w^T \\ w & v_0 \end{pmatrix}\right]$$

where Y is an n-dimensional vector of observations, Y_0 is some unobserved quantity we want to predict, X and x_0 are known regressors, and β is a p-dimensional regression coefficient vector.

REML Estimation: $\hat{\beta}$ by GLS regression, $\hat{\theta}$ chosen to maximize

$$\ell_n(\theta) = -\frac{1}{2}\log|V(\theta)| - \frac{1}{2}\log|X^T V(\theta)^{-1}X| - \frac{G^2(\theta)}{2}$$

where $G^2 = Y^T W Y$, $W = V^{-1} - V^{-1}X(X^T V^{-1}X)^{-1}X^T V^{-1}$.

Universal kriging: Predictor $\hat{Y}_0 = \lambda^T Y$, $\sigma_0^2 = E\left\{(Y_0 - \hat{Y}_0)^2\right\}$, given by

$$\begin{aligned} \lambda &= w^T V^{-1} + (x_0 - X^T V^{-1}w)^T (X^T V^{-1}X)^{-1}X^T V^{-1}, \\ \sigma_0^2 &= v_0 - w^T V^{-1}w + (x_0 - X^T V^{-1}w)^T (X^T V^{-1}X)^{-1}(x_0 - X^T V^{-1}w). \end{aligned}$$

Bayesian Reformulation of Universal Kriging:

$$\tilde{\psi}(z, Y) = \frac{\int e^{\ell_n(\theta)+Q(\theta)}\psi(z, Y; \theta)d\theta}{\int e^{\ell_n(\theta)+Q(\theta)}d\theta}$$

where $e^{\ell_n(\theta)}$ is the restricted likelihood of θ, $Q(\theta) = \log\pi(\theta)$ and $\psi(z, Y; \theta) = \Phi\left(\frac{z-\lambda^T Y}{\sigma_0}\right)$ or $\frac{1}{\sigma_0}\phi\left(\frac{z-\lambda^T Y}{\sigma_0}\right)$; $\Phi(\cdot)$ is the standard normal c.d.f. and $\phi(\cdot)$ the standard normal density.

Design Criteria

$$V_1 = \sigma_0^2 + \mathrm{tr}\left\{\mathcal{I}^{-1}\left(\frac{\partial\lambda}{\partial\theta}\right)^T V \left(\frac{\partial\lambda}{\partial\theta}\right)\right\}, \quad V_2 = \left(\frac{\partial\sigma_0^2}{\partial\theta}\right)^T \mathcal{I}^{-1}\left(\frac{\partial\sigma_0^2}{\partial\theta}\right),$$

where $\mathcal{I}$ is the observed information matrix for θ.

Zhu-Stein criterion for design: minimize

$$V_3 = V_1 + \frac{V_2}{2\sigma_0^2}.$$

Quantiles of the Predictive Distribution:

$$z_P(Y;\theta) = \lambda^T Y + \sigma_0\Phi^{-1}(P), \quad \hat{z}_P(Y) = \hat{\lambda}^T Y + \hat{\sigma}_0\Phi^{-1}(P), \quad \tilde{z}_P(Y) = \tilde{\psi}^{-1}(P, Y).$$

Likelihood Notation: Define $U_i = \frac{\partial\ell_n(\theta)}{\partial\theta^i}$, $U_{ij} = \frac{\partial^2\ell_n(\theta)}{\partial\theta^i\partial\theta^j}$, $U_{ijk} = \frac{\partial^3\ell_n(\theta)}{\partial\theta^i\partial\theta^j\partial\theta^k}$. Other quantities $Q(\theta)$, $\lambda(\theta)$, $\sigma_0(\theta)$. Suffixes denote partial differentiation, for example $Q_i = \partial Q/\partial\theta^i$, $\sigma_{0ij} = \partial^2\sigma_0/\partial\theta^i\partial\theta^j$. Let $\kappa_{i,j} = n^{-1}E\{U_i U_j\}$, $\kappa_{ijk} = n^{-1}E\{U_{ijk}\}$, $\kappa_{i,jk} = n^{-1}E\{U_i U_{jk}\}$. Suppose inverse of $\{\kappa_{i,j}\}$ matrix has entries $\{\kappa^{i,j}\}$. We assume all these quantities are of $O(1)$ as $n \to \infty$ and we employ the summation convention. Then

$$
\begin{aligned}
nE\left\{\psi(\hat{z}_P(Y)\ ;\ Y,\theta) - \psi(z_P(Y)\ ;\ Y,\theta)\right\} \quad &\sim \quad \phi(\Phi^{-1}(P))\Phi^{-1}(P)\Bigg[-\frac{1}{2}\Phi^{-1}(P)^2\kappa^{i,j}\frac{\sigma_{0i}\sigma_{0j}}{\sigma_0^2} \\
&+\kappa^{i,j}\kappa^{k,\ell}\left(\kappa_{jk,\ell}+\frac{1}{2}\kappa_{jk\ell}\right)\frac{\sigma_{0i}}{\sigma_0}+\frac{1}{2}\kappa^{i,j}\left\{\frac{\sigma_{0ij}}{\sigma_0}-\frac{\lambda_i^TV\lambda_j}{\sigma_0^2}\right\} \\
&-\frac{1}{2}\kappa^{i,k}\kappa^{j,\ell}\cdot\frac{1}{n\sigma_0^2}\left(\lambda_i^TV\frac{\partial W}{\partial\theta^k}V\frac{\partial W}{\partial\theta^\ell}V\lambda_j+\lambda_i^TV\frac{\partial W}{\partial\theta^\ell}V\frac{\partial W}{\partial\theta^k}V\lambda_j\right)\Bigg], \\
nE\left\{\psi(\tilde{z}_P(Y)\ ;\ Y,\theta) - \psi(z_P(Y)\ ;\ Y,\theta)\right\} \quad &\sim \quad \phi(\Phi^{-1}(P))\Phi^{-1}(P)\Bigg[\kappa^{i,j}\kappa^{k,\ell}\left(\kappa_{jk,\ell}+\kappa_{jk\ell}\right)\frac{\sigma_{0i}}{\sigma_0} \\
&-\kappa^{i,j}\left(\frac{\sigma_{0i}\sigma_{0j}}{\sigma_0^2}-\frac{\sigma_{0ij}}{\sigma_0}\right)+\kappa^{i,j}\frac{\sigma_{0i}}{\sigma_0}Q_j \\
&-\frac{1}{2}\kappa^{i,k}\kappa^{j,\ell}\cdot\frac{1}{n\sigma_0^2}\left(\lambda_i^TV\frac{\partial W}{\partial\theta^k}V\frac{\partial W}{\partial\theta^\ell}V\lambda_j+\lambda_i^TV\frac{\partial W}{\partial\theta^\ell}V\frac{\partial W}{\partial\theta^k}V\lambda_j\right)\Bigg],
\end{aligned}
$$

$$
\begin{aligned}
nE\left\{\hat{z}_P - z_P\right\} \quad &\approx \quad \Phi^{-1}(P)\left\{\kappa^{i,j}\kappa^{k,\ell}\sigma_{0\ell}\left(\kappa_{ik,j}+\frac{1}{2}\kappa_{ijk}\right)+\frac{1}{2}\kappa^{i,j}\sigma_{0ij}\right\} \\
nE\left\{\tilde{z}_P - z_P\right\} \quad &\approx \quad \Phi^{-1}(P)\Bigg\{\kappa^{i,j}\kappa^{k,\ell}\sigma_{0\ell}(\kappa_{ik,j}+\kappa_{ijk})+\kappa^{i,j}\left(\sigma_{0ij}-\frac{\sigma_{0i}\sigma_{0j}}{\sigma_0}\right) \\
&\qquad +\kappa^{i,j}Q_j\sigma_{0i}+\frac{1}{2}\Phi^{-1}(P)^2\kappa^{i,j}\frac{\sigma_{0i}\sigma_{0j}}{\sigma_0}+\frac{1}{2}\kappa^{i,j}\frac{\lambda_i^TV\lambda_j}{\sigma_0}\Bigg\}.
\end{aligned}
$$

An improved estimator for z_P is

$$
\begin{aligned}
z_P^\dagger \quad = \quad & \hat{z}_P - n^{-1}\Phi^{-1}(P)\Bigg\{\hat{\kappa}^{i,j}\hat{\kappa}^{k,\ell}\hat{\sigma}_{0\ell}\left(\hat{\kappa}_{ik,j}+\frac{1}{2}\hat{\kappa}_{ijk}\right)+\frac{1}{2}\hat{\kappa}^{i,j}\left(\hat{\sigma}_{0ij}-\frac{\hat{\sigma}_{0i}\hat{\sigma}_{0j}}{\hat{\sigma}_0}\Phi^{-1}(P)^2\right)-\frac{1}{2\hat{\sigma}_0}\hat{\kappa}^{i,j}\hat{\lambda}_i^T\hat{V}\hat{\lambda}_j \\
&-\frac{1}{2n\hat{\sigma}_0}\hat{\kappa}^{i,j}\hat{\kappa}^{k,\ell}\left(\hat{\lambda}_j^T\hat{V}\frac{\partial\hat{W}}{\partial\theta^i}\hat{V}\frac{\partial\hat{W}}{\partial\theta^k}\hat{V}\hat{\lambda}_\ell+\hat{\lambda}_j^T\hat{V}\frac{\partial\hat{W}}{\partial\theta^k}\hat{V}\frac{\partial\hat{W}}{\partial\theta^i}\hat{V}\hat{\lambda}_\ell\right)\Bigg\}.
\end{aligned}
$$

References

Smith, R.L. (2004), Asymptotic theory for kriging with estimated parameters and its application to network design. Preliminary version, available from http://www.stat.unc.edu/postscript/rs/supp5.pdf

Zhu, Z. and Stein, M.L. (2004), Two-step spatial sampling design for prediction with estimated parameters. Preprint, University of North Carolina and University of Chicago.