Home Assignment - 1

Santhosh Nadig

November 14, 2016

1 Introduction

2 Theory

Let $\mathbf{Y}_k$ represent the n known observations of spatial data and $\mathbf{Y}_u$ represent the m data points that needs to be estimated. For ordinary Kriging ($\mu = \mathbf{1}\beta$), the optimal predictions are given by

$$\begin{aligned}
\hat{\mathbf{Y}}_u &= \mathbf{1}_u\hat{\beta} + \Sigma_{uk}\Sigma_{kk}^{-1}\left(\mathbf{Y}_k - \mathbf{1}_k\hat{\beta}\right) \\
\hat{\beta} &= \left(\mathbf{1}_k^T\Sigma_{kk}^{-1}\mathbf{1}_k\right)^{-1}\mathbf{1}_k^T\Sigma_{kk}^{-1}\mathbf{Y}_k,
\end{aligned}$$

where

$$\begin{bmatrix}\mathbf{Y}_k \\ \mathbf{Y}_u\end{bmatrix} \in \mathbf{N}\left(\begin{bmatrix}\mathbf{1}_k\beta \\ \mathbf{1}_u\beta\end{bmatrix}, \begin{bmatrix}\Sigma_{kk} & \Sigma_{ku} \\ \Sigma_{uk} & \Sigma_{uu}\end{bmatrix}\right).$$

Firstly, we show that the predictions are linear in the observations, i.e., $\hat{\mathbf{Y}}_u = \boldsymbol{\lambda}^T\mathbf{Y}_k$ for some $\boldsymbol{\lambda}$. We may re-write $\hat{\beta}$ as

$$\begin{aligned}
\hat{\beta} &= \boldsymbol{\gamma}^T\mathbf{Y}_k \\
\boldsymbol{\gamma} &= \left(\left(\mathbf{1}_k^T\Sigma_{kk}^{-1}\mathbf{1}_k\right)^{-1}\mathbf{1}_k^T\Sigma_{kk}^{-1}\right)^T.
\end{aligned}$$

We note that the vector $\boldsymbol{\gamma}^T$ is of dimensions $1 \times n$, and therefore, $\hat{\beta}$ is a scalar. Substituting the above in the expression for $\hat{\mathbf{Y}}_u$, we obtain

$$\begin{aligned}
\hat{\mathbf{Y}}_u &= \mathbf{1}_u\boldsymbol{\gamma}^T\mathbf{Y}_k + \Sigma_{uk}\Sigma_{kk}^{-1}\mathbf{Y}_k - \Sigma_{uk}\Sigma_{kk}^{-1}\mathbf{1}_k\boldsymbol{\gamma}^T\mathbf{Y}_k \\
&= \left(\mathbf{1}_u\boldsymbol{\gamma}^T + \Sigma_{uk}\Sigma_{kk}^{-1} - \Sigma_{uk}\Sigma_{kk}^{-1}\mathbf{1}_k\boldsymbol{\gamma}^T\right)\mathbf{Y}_k \\
&= \boldsymbol{\lambda}^T\mathbf{Y}_k,
\end{aligned}$$

where $\boldsymbol{\lambda} = \left(\mathbf{1}_u\boldsymbol{\gamma}^T + \Sigma_{uk}\Sigma_{kk}^{-1} - \Sigma_{uk}\Sigma_{kk}^{-1}\mathbf{1}_k\boldsymbol{\gamma}^T\right)^T$.

Secondly, we show that the predictions are unbiased, i.e., $\mathbf{E}(\hat{\mathbf{Y}}_u) = \mathbf{E}(\mathbf{Y}_u)$. The expected value of $\hat{\beta}$ is given by

$$\begin{aligned}
\mathbf{E}\left(\hat{\beta}\right) &= \left(\mathbf{1}_k^T\Sigma_{kk}^{-1}\mathbf{1}_k\right)^{-1}\mathbf{1}_k^T\Sigma_{kk}^{-1}\ \mathbf{E}(\mathbf{Y}_k) \\
&= \left(\mathbf{1}_k^T\Sigma_{kk}^{-1}\mathbf{1}_k\right)^{-1}\mathbf{1}_k^T\Sigma_{kk}^{-1}\ \mathbf{1}_k\beta \\
&= \beta.
\end{aligned}$$

Therefore,

$$\begin{aligned}
\mathbf{E}(\hat{\mathbf{Y}}_u) &= \mathbf{1}_u \mathbf{E}\left(\hat{\beta}\right) + \Sigma_{uk}\Sigma_{kk}^{-1}\left(\mathbf{E}(\mathbf{Y}_k) - \mathbf{1}_k \mathbf{E}\left(\hat{\beta}\right)\right) \\
&= \mathbf{1}_u \beta + \Sigma_{uk}\Sigma_{kk}^{-1}\left(\mathbf{1}_k \beta - \mathbf{1}_k \beta\right) \\
&= \mathbf{1}_u \beta.
\end{aligned}$$

Finally, consider a second unbiased predictor of $\mathbf{Y}_u$, given by $\tilde{\mathbf{Y}}_u = (\boldsymbol{\lambda} + \boldsymbol{\nu})^T \mathbf{Y}_k$. From the unbiasedness criterion, we have that $\mathbf{E}(\tilde{\mathbf{Y}}_u) = \mathbf{E}(\hat{\mathbf{Y}}_u)$. Therefore, $(\boldsymbol{\lambda} + \boldsymbol{\nu})^T \mathbf{E}(\mathbf{Y}_k) = \boldsymbol{\lambda}^T \mathbf{E}(\mathbf{Y}_k)$, implying $\boldsymbol{\nu}^T \mathbf{1}_k = 0$. Furthermore, consider the variance of the second predictor

$$\begin{aligned}
\mathbf{V}(\tilde{\mathbf{Y}}_u - \mathbf{Y}_u) &= \mathbf{V}(\boldsymbol{\nu}^T \mathbf{Y}_k + (\boldsymbol{\lambda}^T \mathbf{Y}_k - \mathbf{Y}_u)) \\
&= \mathbf{V}(\boldsymbol{\nu}^T \mathbf{Y}_k) + \mathbf{V}(\hat{\mathbf{Y}}_u - \mathbf{Y}_u) + 2\mathbf{C}(\boldsymbol{\nu}^T \mathbf{Y}_k, \hat{\mathbf{Y}}_u - \mathbf{Y}_u) \\
&\geq \mathbf{V}(\hat{\mathbf{Y}}_u - \mathbf{Y}_u) + 2\mathbf{C}(\boldsymbol{\nu}^T \mathbf{Y}_k, \hat{\mathbf{Y}}_u - \mathbf{Y}_u) \\
&= \mathbf{V}(\hat{\mathbf{Y}}_u - \mathbf{Y}_u) + 2\boldsymbol{\nu}^T \mathbf{C}(\mathbf{Y}_k, \hat{\mathbf{Y}}_u) - 2\boldsymbol{\nu}^T \mathbf{C}(\mathbf{Y}_k, \mathbf{Y}_u) \\
&= \mathbf{V}(\hat{\mathbf{Y}}_u - \mathbf{Y}_u) + 2\boldsymbol{\nu}^T \left(\Sigma_{kk}\boldsymbol{\lambda} - \Sigma_{uk}\right)
\end{aligned}$$

Thus, if we can choose $\boldsymbol{\lambda}$ such that $\boldsymbol{\nu}^T \left(\Sigma_{kk}\boldsymbol{\lambda} - \Sigma_{uk}\right) = 0$, then, we have that $\mathbf{V}(\tilde{\mathbf{Y}}_u - \mathbf{Y}_u) \geq \mathbf{V}(\hat{\mathbf{Y}}_u - \mathbf{Y}_u)$ for any unbiased predictor $\tilde{\mathbf{Y}}_u$.

3 Data

4 Least-Squares

5 Universal Kriging

6 Conclusions