Home Assignment - 1

Santhosh Nadig

November 15, 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 Swedish Temparature Reconstruction

The data consists of average temperature during June 2005 measured at 250 stations across Sweden. In addition, the latitude, longitude, elevation and distance to both the Swedish coast and to any coastline are provided too. Figure 1 shows the plot of average temperature, elevation and distance to any coast. The average temperatures are higher in the South (lower latitudes) and towards the coast, as expected.

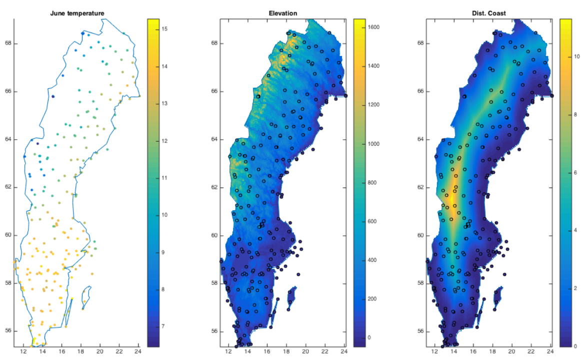

Figure 1: Raw Data Plot...

3.1 Least-Squares

The mean of the spatial data is modelled as a linear function of the covariates. Figure 2 shows a plot of the mean temperature (observations) as a function of the available covariates. The covariates that have a linear relationship are the

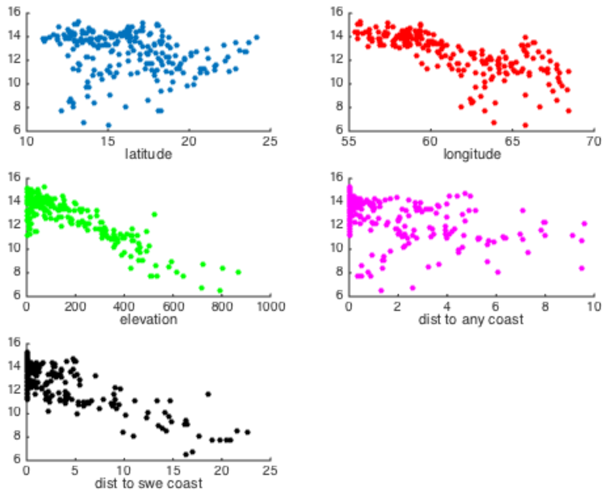

Figure 2: Observation vs covariates

latitude ($\mathbf{x}_1$), elevation ($\mathbf{x}_2$) and distance to Swedish coast ($\mathbf{x}_3$). Thus the mean function is expressed as

$$\boldsymbol{\mu} = \beta_0 + \beta_1\mathbf{x}_1 + \beta_2\mathbf{x}_2 + \beta_3\mathbf{x}_3 \tag{1}$$

where β_0 is the intercept and $\beta_1 \ldots \beta_3$ are the regressors. With $N = 250$ observations from the known locations, the estimated parameters (and their standard deviations) are tabulated below.

Parameter	Estimate	SE
β_0	24.553445	0.928379
β_1	-0.177836	0.015600
β_2	-0.004178	0.000455
β_3	-0.092380	0.015663

The variance of the residuals ${\sigma_\epsilon}^2 = 0.550609$.

3.2 Universal Kriging

4 Conclusions