UNIVERSITETET I OSLO

Det matematisk-naturvitenskapelige fakultet

Examination in:	STK4150 — Environmental and spatial statistics - solutions
Day of examination:	Monday 3. june 2015.
Examination hours:	14.30 – 18.00.

This examination set consists of 4 pages.

Appendices:	None
Permitted aids:	Approved calculator, Cressie and Wikle: Statistics for Spatio-temporal data

Make sure that your copy of the examination set is complete before you start solving the problems.

Problem 1.

(a) We have

$$\begin{aligned}
E[Y(\boldsymbol{s})] =&\beta_0 + \beta_1 x(\boldsymbol{s})\\
\mathrm{Cov}[Y(\boldsymbol{s}), Y(\boldsymbol{v})] =&C_\delta(||\boldsymbol{s}-\boldsymbol{v}||)\\
E[Z(\boldsymbol{s})] =&\beta_0 + \beta_1 x(\boldsymbol{s})\\
\mathrm{Cov}[Z(\boldsymbol{s}), Z(\boldsymbol{v})] =&C_\delta(||\boldsymbol{s}-\boldsymbol{v}||) + \sigma_\varepsilon^2 I(\boldsymbol{s}=\boldsymbol{v})
\end{aligned}$$

The processes are Gaussian since they are a linear combination of Gaussian processes.

The nugget effect is equal to σ_ε^2 for the Z-process while it is zero for the Y-process.

(b) In this case the nugget effect becomes $\sigma_1^2 + \sigma_\varepsilon^2$ for the Z-process while it becomes σ_1^2 for the Y-process.

Since we only observe the Z's, we will not be able to distinguish between σ_1^2 and σ_ε^2.

(Continued on page 2.)

(c) Define $\boldsymbol{Z} = (Z(\boldsymbol{s}_1), ..., Z(s_n))^T$. Then the simultaneous distribution for $(\boldsymbol{Z}, Y(\boldsymbol{s}_0))$ is multivariate Gaussian and therefore also the conditional distribution for $Y(\boldsymbol{s}_0)$ given $\boldsymbol{Z}$ is Gaussian with

$$\begin{aligned}
E[Y(\boldsymbol{s}_0)|\boldsymbol{Z}] =& \beta_0 + \beta_1 x(\boldsymbol{s}_0) + \boldsymbol{c}_Y^T \boldsymbol{\Sigma}^{-1}(\boldsymbol{Z} - \beta_0 \mathbf{1} - \beta_1 \boldsymbol{X}) \\
E[Z(\boldsymbol{s}_0)|\boldsymbol{Z}] =& \beta_0 + \beta_1 x(\boldsymbol{s}_0) + \boldsymbol{c}_Z^T \boldsymbol{\Sigma}^{-1}(\boldsymbol{Z} - \beta_0 \mathbf{1} - \beta_1 \boldsymbol{X}) \\
\text{Var}[Y(\boldsymbol{s}_0)|\boldsymbol{Z}] =& C_\delta(0) - \boldsymbol{c}_Y^T \boldsymbol{\Sigma}^{-1} \boldsymbol{c}_Y \\
\text{Var}[Z(\boldsymbol{s}_0)|\boldsymbol{Z}] =& C_\delta(0) - \boldsymbol{c}_Z^T \boldsymbol{\Sigma}^{-1} \boldsymbol{c}_Z + \sigma_\varepsilon^2
\end{aligned}$$

where $\boldsymbol{X} = (x(\boldsymbol{s}_1), ..., x(\boldsymbol{s}_n))^T$, $\boldsymbol{c}_Y$ is the vector of covariances between $Y(\boldsymbol{s}_0)$ and the observations, $\boldsymbol{c}_Z$ is the vector of covariances between $Z(\boldsymbol{s}_0)$ and the observations and $\boldsymbol{\Sigma}$ is the covariance matrix for the observations. Note that $\boldsymbol{c}_Y$ is indetical to $\boldsymbol{c}_Z$ with exception to the case where $\boldsymbol{s}_0$ matches one of the observations.

They are optimal in the sense of minimising expected quadratic loss.

(d) Without the restriction, $\boldsymbol{s}_0$ could be one of the observation points in which case $Z(\boldsymbol{s}_0)$ is observed and then the prediction error is exactly zero.

With the restriction, note first that then $\boldsymbol{c}_Z = \boldsymbol{c}_Y$.

For $n = 1$ we have that (using $h = ||\boldsymbol{s}_0 - \boldsymbol{s}_1||$)

$$\text{Var}[Z(\boldsymbol{s}_0)|\boldsymbol{Z}] = C_\delta(0) + \sigma_\varepsilon^2 - C_\delta(h)^2/(C_\delta(0) + \sigma_\varepsilon^2)$$

For correlation functions we have that: $C_\delta(0) \geq C_\delta(h)$ for all h so using this expression in the equation above we find that.

$$\begin{aligned}
\text{Var}[Z(\boldsymbol{s}_0)|\boldsymbol{Z}] \geq& C_\delta(0) + \sigma_\varepsilon^2 - \tfrac{C_\delta(0)^2}{C_\delta(0)+\sigma_\varepsilon^2} \\
=& \sigma_\varepsilon^2 + \frac{C_\delta(0)\sigma_\varepsilon^2}{C_\delta(0) + \sigma_\varepsilon^2}
\end{aligned}$$

Further, this limit can be obtained arbitrarily close by shrinking h to zero.

(e) We have that

$$\begin{aligned}
E[Y(\boldsymbol{s})] =& \beta_0 + \beta_1 E[x(\boldsymbol{s})] = \beta_0 \\
\text{Cov}[Y(\boldsymbol{s}), Y(\boldsymbol{v})] =& \beta_1^2 C_x(||\boldsymbol{s} - \boldsymbol{v}||) + C_\delta(||\boldsymbol{s} - \boldsymbol{v}||) \\
E[Z(\boldsymbol{s})] =& \beta_0 \\
\text{Cov}[Z(\boldsymbol{s}), Z(\boldsymbol{v})] =& \beta_1^2 C_x(||\boldsymbol{s} - \boldsymbol{v}||) + C_\delta(||\boldsymbol{s} - \boldsymbol{v}||) + \sigma_\varepsilon^2 I(\boldsymbol{s} = \boldsymbol{v})
\end{aligned}$$

One can now utilise that $\{y(\boldsymbol{s}), z(\boldsymbol{s})\}$ still are multivariate Gaussian and use ordinary rules about conditional distributions within MVN's to find optimal predictions and corresponding uncertainties.

(Continued on page 3.)

(f) Define $\boldsymbol{X}$ to be the observed covariates. Since everything is Gaussian, we have that $(\boldsymbol{Z}, \boldsymbol{X}, Y(\boldsymbol{s}_0))$ is multivariate Gaussian. Further, $[\boldsymbol{X}, Y(\boldsymbol{s}_0)|\boldsymbol{Z}]$ is also multivariate Gaussian. From this distribution we obtain the prediction error for $Y(\boldsymbol{s}_0)|\boldsymbol{Z}$. Further, by using the ordinary equations for conditional distributions, we can from this distribution derive $[Y(\boldsymbol{s}_0)|\boldsymbol{Z}, \boldsymbol{X}]$. The variance in this distribution will always be smaller than the variance in the distribution for $[Y(\boldsymbol{s}_0)|\boldsymbol{Z}]$ showing that there is a gain in including the x-observations.

The same argument holds if we want to predict $Z(\boldsymbol{s}_0)$.

Problem 2.

(a) Define C^ν to be the covariance function for $\{\nu_t(\boldsymbol{s})\}$. We have that

$$\begin{aligned}
C^\kappa(t, \boldsymbol{s}, t, \boldsymbol{v}) =&\mathrm{Cov}[\kappa(t, \boldsymbol{s}), \kappa(t, \boldsymbol{v})]\\
=&\mathrm{Cov}[\kappa(t-1, \boldsymbol{s}) + \nu_t(\boldsymbol{s}), \kappa(t-1, \boldsymbol{v}) + \nu_t(\boldsymbol{v})]\\
=&C^\kappa(t-1, \boldsymbol{s}, t-1, \boldsymbol{v}) + C^\nu(\boldsymbol{s}, \boldsymbol{v})\\
=&C^\kappa(t-2, \boldsymbol{s}, t-2, \boldsymbol{v}) + 2C^\nu(\boldsymbol{s}, \boldsymbol{v})\\
&\cdots\\
=&C^\kappa(0, \boldsymbol{s}, 0, \boldsymbol{v}) + tC^\nu(\boldsymbol{s}, \boldsymbol{v})\\
=&(t+1)C^\nu(\boldsymbol{s}, \boldsymbol{v})
\end{aligned}$$

which changes with time, showing that it is not stationary.

Further,

$$\begin{aligned}
C^\kappa(t, \boldsymbol{s}, t+k, \boldsymbol{v}) =&\mathrm{Cov}[\kappa(t, \boldsymbol{s}), \kappa(t, \boldsymbol{v})]\\
=&\mathrm{Cov}[\kappa(t, \boldsymbol{s}), \kappa(t+k-1, \boldsymbol{v}) + \nu_{t+k}(\boldsymbol{v})]\\
=&C^\kappa(t, \boldsymbol{s}, t+k-1, \boldsymbol{v})\\
&\cdots\\
=&C^\kappa(t, \boldsymbol{s}, t, \boldsymbol{v})\\
=&(t+1)C^\nu(\boldsymbol{s}, \boldsymbol{v})
\end{aligned}$$

which is clearly separable.

(b) We have that

$$\begin{aligned}
C_Y(t, r, \boldsymbol{s}, \boldsymbol{v}) \equiv&\mathrm{Cov}[Y(t, \boldsymbol{s}), Y(r, \boldsymbol{v})]\\
=&\mathrm{Cov}[\delta(\boldsymbol{s}) + \gamma(t) + \kappa(t, \boldsymbol{s}), \delta(\boldsymbol{v}) + \gamma(r) + \kappa(r, \boldsymbol{v})]\\
=&\mathrm{Cov}[\delta(\boldsymbol{s}), \delta(\boldsymbol{v})] + \mathrm{Cov}[\gamma(t), \gamma(r)] + \mathrm{Cov}[\kappa(t, \boldsymbol{s}), \kappa(r, \boldsymbol{v})]\\
=&C^\delta(\boldsymbol{s}, \boldsymbol{v}) + C^\gamma(t, r) + C^\kappa(t, \boldsymbol{s}, r, \boldsymbol{v})
\end{aligned}$$

(Continued on page 4.)

Separability requires $C_Y(t, r, \boldsymbol{s}, \boldsymbol{v}) = \widetilde{C}^s(\boldsymbol{s}, \boldsymbol{v})\widetilde{C}^t(t, r)$ for some appropriate functions $\widetilde{C}^s(\boldsymbol{s}, \boldsymbol{v}), \widetilde{C}^t(t, r)$. This will not be the case for the additive structure.

It will inherit the non-stationary property of the κ-process (and even more so since the γ-process is not stationary either).

(c) We see that the most important component in the model is the spatio-temporal part $\kappa_t(\boldsymbol{s})$ which contain both spatial and temporal dependence. In fact it seems like almost all the variance is explained through this part, the others being very small.

(d) EOF are simply principal components, that is linear combinations that shows most variability where here variability means over time. Since these then are directly related to the process and how it changes, it makes sense to include them as "covariates" but with time-dependent coefficients. It is reasonable that the α's corresponding to the first EOF var most since the EOF's are ordered according to their variability. Further, using EOF's, makes it easier to do prediction.

(e) The parameters have actually not changed much, indicating that we are not able to explain very much through the EOF's. It is however reasonable that all the variances decreases since more of the variability now is moved to the expectation structure.

The likelihood value is increasing considerably. Even when taking into account the difference in number of parameters (30 $\alpha_{t,k}$'s) the difference is large. The gain in AIC will be

$$2 * (6840.2 - 4581.7) - 2 * 30 = 4455.9$$

The new model is therefore preferable.

END