TMA4250 Spatial Statistics Exercise 1, Spring 2019

Group members: Henrik Syversveen Lie, Ã~yvind Auestad (SKRIVE STUDENTNUMMER I STEDET??)

01.02.2019

Contents

Problem 1: Gaussian RF - model characteristics

We consider the continuous spatial variable $\{r(x) : x \in D : [1, 50] \subset \mathbb{R}^1$, and assume that it is modeled as a stationary 1D Gaussian RF with the following model parameters:

$$\mathrm{E}\{r(x)\} = \mu_r = 0$$
$$\mathrm{Var}\{r(x)\} = \sigma_r^2$$
$$\mathrm{Corr}\{r(x), r(x')\} = \rho_r(\tau),$$

where $\rho_r(\tau); \tau = |x-x'|/10$ is the spatial correlation function. Let $D : [1, 50]$ be discretised in $L \in \{1, 2, \ldots, 50\}$ and define the discretised Gaussian RF $\{r(x); x \in L\}$, represented by the n-vector $\mathbf{r} \in \mathbb{R}^n$.

Let the spatial correlation function $\rho_r(\tau)$, be either Powered exponential with parameter $\nu_r \in [1, 1.9]$ or Matern with parameter $\nu_r \in [1, 3]$. Let the variance take the values $\sigma_r^2 \in [1, 5]$.

a)

We restrict the spatial correlation functions to be only positive definite functions. Covariance matrices need to be positive definite, and a positive definite correlation function ensures that all covariances matrices in our stationary Gaussian RF are positive definite for all configurations and dimensions. Further, we define a function $\rho(\tau) : \mathbb{R}^q \to \mathbb{R}$ to be positive definite if

$$\sum_{i=1}^{n}\sum_{j=1}^{n}\alpha_i\alpha_j\rho(\mathbf{x}_i - \mathbf{x}_j) \geq 0,$$

for all configuations $[\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n] \in \mathbb{R}^{q \times n}$,

for all weights $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_n) \in \mathbb{R}^n$,

for all $n \in \mathbb{N}_+ \setminus \{1\}$.

To further investigate different correlation functions, we display two types of correlation functions, the Matern and Powered exponential functions, for different values of ν_r.

Display of two correlation functions for different values of ν
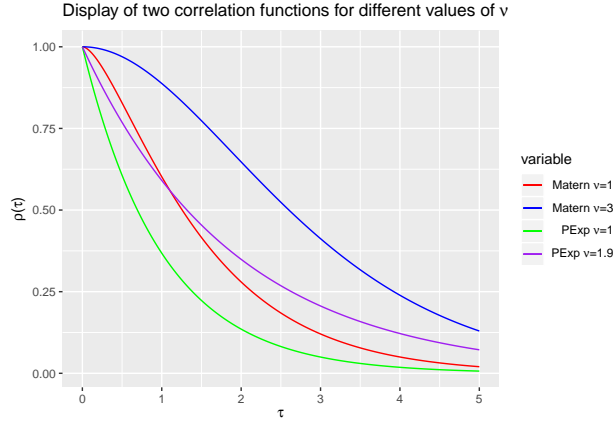
For all correlation functions $\rho_r(0) = 1$, and $\rho_r(\tau) \in [-1, 1]$; $\tau \in \mathbb{R}_+$, because the function represents correlation between two random variables. The correlation functions are continuous everywhere, except at $\tau = 0$ where a step may occur. If the correlation function is continuous at $\tau = 0$, then the random field is continuous almost everywhere. Away from $\tau = 0$, the correlation function must be smooth. From the displayed correlation functions, we see that $\rho_r(\tau) \to 0$ as $\tau \to \infty$. This will be the case for all correlation functions, implying that two points $\mathbf{x}$ and $\mathbf{x}'$ will tend towards being uncorrelated as $|\mathbf{x} - \mathbf{x}'| \to \infty$. Uncorrelated Gaussian random fields means that they will be independent, and by extension asymptotics Gaussion random fields are ergodic.

The correlation functions define correlation between points in our random field. This means that larger values of the correlation function implies a smoother random field. From the plotted correlation function, we see that the function values increases for larger values of ν_r. Consequently, larger values of ν_r implies smoother random fields for both these correlation functions.

We now define the variogram function, which for stationary Gaussian random fields is defined as $\gamma_r(\tau) = 2\sigma_r^2[1 - \rho_r(\tau)]$. Then, we display the variogram functions associated with the previously plotted correlation functions for $\sigma_r^2 = 1$.

Variogram for two correlation functions for different values of ν

b)
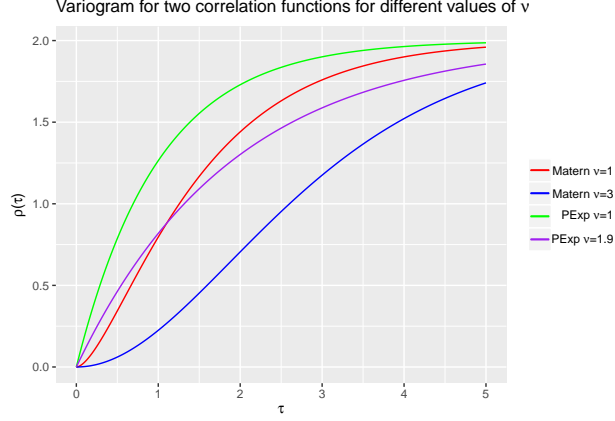
The prior Gaussion random field is defined on the discretized representation $L \in \{1, 2, \ldots, 50\}$ by

$$\mathbf{r} \sim p(\mathbf{r}) = \phi_n(\mathbf{r}; \mu_r \mathbf{i}_n, \sigma_r^2 \mathbf{\Sigma}_r^\rho),$$

which is a discretized stationary Gaussian random field with expectation $\mu(\mathbf{x}) = \mu_r$, variance $\sigma^2(\mathbf{x}) = \sigma_r^2$ and correlation function $\rho(\mathbf{x}, \mathbf{x}') = \rho_r(\mathbf{x} - \mathbf{x}')$.

We now want to simulate ten realizations of the Gaussian random field on L for all the previously displayed correlation functions and $\sigma_r^2 \in [1, 5]$.
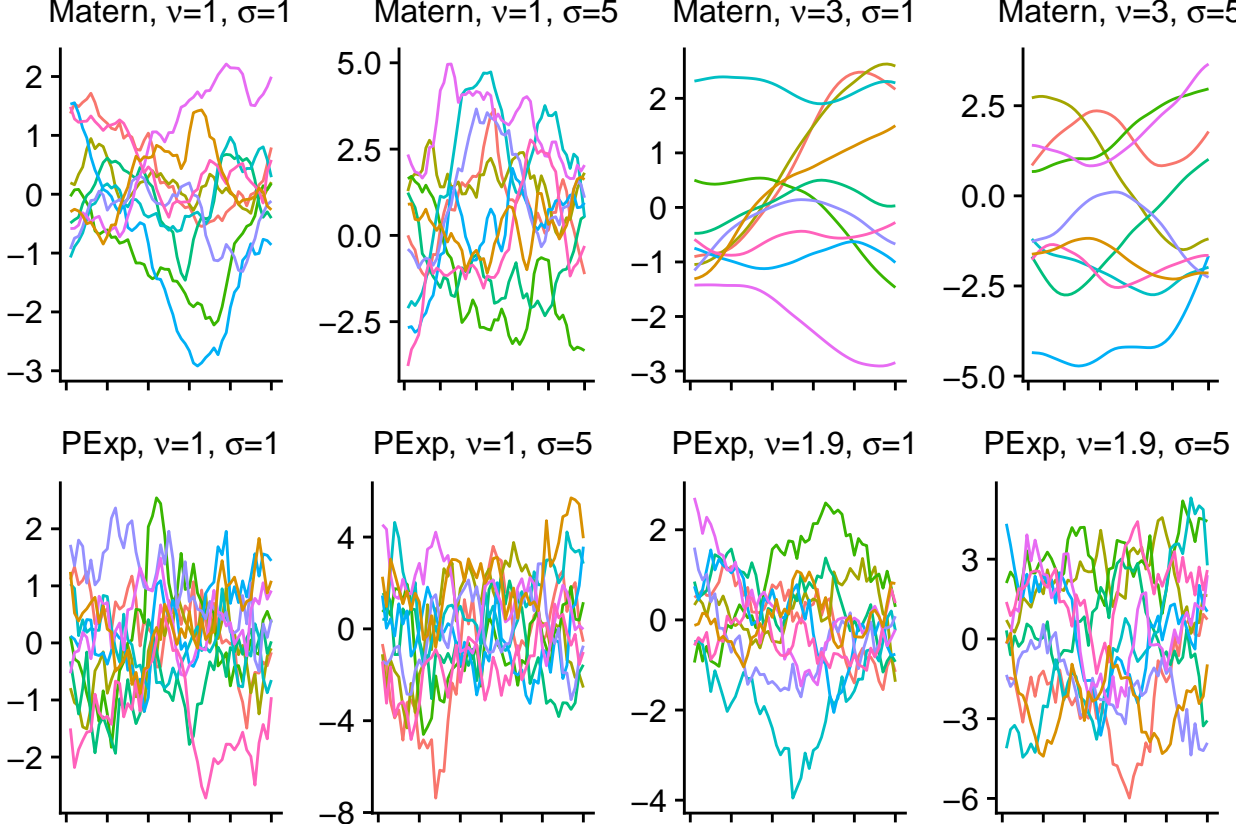
From the display, we see that a greater value of ν_r leads to smoother random fields. Also, a higher value of σ_r^2 leads to larger variations around the expected level $\mu_r = 0$. Also, the Matern correlation function leads to smoother random fields than the Powered exponential function.

c)

We now observe the spatial variable as $\{d(x); x \in [10, 25, 30] \subset L\}$ according to the acquisition model,

$$d(x) = r(x) + \epsilon(x) \qquad x \in [10, 25, 30],$$

with measurement errors $\epsilon(\cdot)$ centred, i.i.d Gaussian with variance σ_ϵ^2. Further, we assume that $r(x)$ and $r(x')$ are independent for all x, x'.

The likelihood model $p(\mathbf{d}|\mathbf{r})$ links the observations to the spatial variable. The observations $\mathbf{d}$ are known, while $\mathbf{r}$ is the spatial variable. This means that the likelihood $p(\mathbf{d}|\mathbf{r})$ is not a pdf w.r.t. $\mathbf{r}$, and we do not need to normalize the likelihood.

We can define a Gauss-linear likelihood model relative to the discretized spatial variable $\mathbf{r}$, and observe $\mathbf{d} \in \mathbb{R}^m$ according to

$$[\mathbf{d}|\mathbf{r}] = \mathbf{H}\mathbf{r} + \boldsymbol{\epsilon}_{d|r} \sim p(\mathbf{d}|\mathbf{r}) = \phi_m(\mathbf{d}; \mathbf{H}\mathbf{r}, \boldsymbol{\Sigma}_{d|r}),$$

where $\mathbf{H}$ is a $(m \times n)$ observation matrix. In our case, because we have three observations, $m = 3$. Also, the observations are assumed to have i.i.d. errors, which implies $\boldsymbol{\Sigma}_{d|r} = \sigma_\epsilon^2 I_{m \times m}$.

d)

The pdf for the discretised posterior Gaussian random field, given the observations is defined in the following way,

$$[\mathbf{r}|\mathbf{d}] \sim p(\mathbf{r}|\mathbf{d}) = \phi_n(\mathbf{r}; \boldsymbol{\mu}_{r|d}, \boldsymbol{\Sigma}_{r|d},)$$

with

$$\boldsymbol{\mu}_{r|d} = \mu_r \mathbf{i}_n + \sigma_r^2 \boldsymbol{\Sigma}_r^\rho \mathbf{H}^T \left[\sigma_r^2 \mathbf{H}\boldsymbol{\Sigma}_r^\rho \mathbf{H} + \boldsymbol{\Sigma}_{d|r}\right]^{-1} [\mathbf{d} - \mu_r \mathbf{H}\mathbf{i}_n]$$

$$\boldsymbol{\Sigma}_{r|d} = \boldsymbol{\Sigma}_{r|d}^\sigma \boldsymbol{\Sigma}_{r|d}^\rho \boldsymbol{\Sigma}_{r|d}^\sigma = \sigma_r^2 \boldsymbol{\Sigma}_r^\rho - \sigma_r^2 \boldsymbol{\Sigma}_r^\rho \mathbf{H}^T \left[\sigma_r^2 \mathbf{H}\boldsymbol{\Sigma}_r^\rho \mathbf{H} + \boldsymbol{\Sigma}_{d|r}\right]^{-1} \sigma_r^2 \mathbf{H}\boldsymbol{\Sigma}_r^\rho$$

We use as prior model one of the realizations with $\sigma_r^2 = 5$, $\nu = 1$ and Matern correlation function. Then, a prediction of the spatial variable $\{\hat{r}(\mathbf{x}); \mathbf{x} \in L\}$ represented by the vector $\hat{\mathbf{r}}$ is taken to minimize squared error, yielding
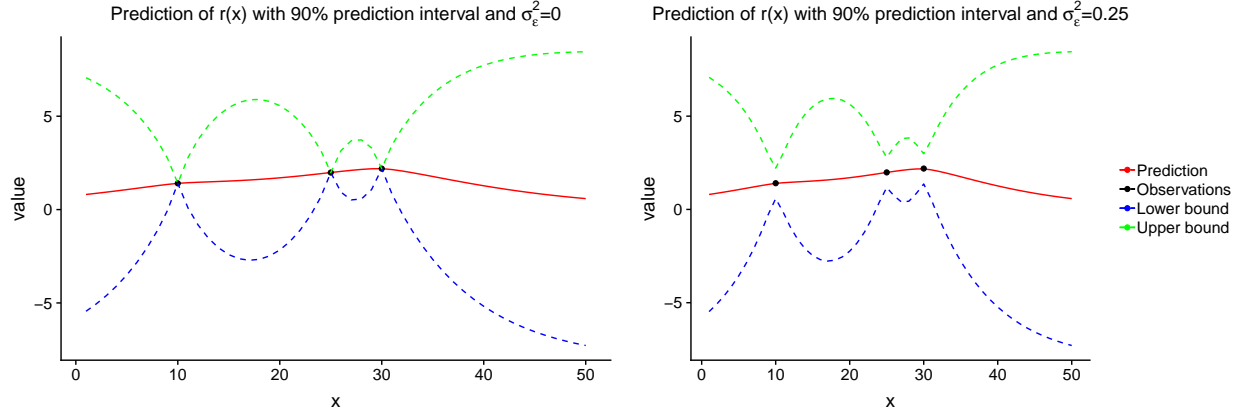
$$\hat{\mathbf{r}} = \mathbf{E}[\mathbf{r}|\mathbf{d}] = \boldsymbol{\mu}_{r|d}.$$

The associated $(1 - \alpha)$ prediction intervals are,

$$PI_\alpha = \boldsymbol{\mu}_{r|d} \pm z_{\alpha/2} \boldsymbol{\sigma}_{r|d},$$

where $\boldsymbol{\sigma}_{r|d}$ is a n-vector containing the diagonal elements of the standard deviation matrix $\boldsymbol{\Sigma}_{r|d}^\sigma$.
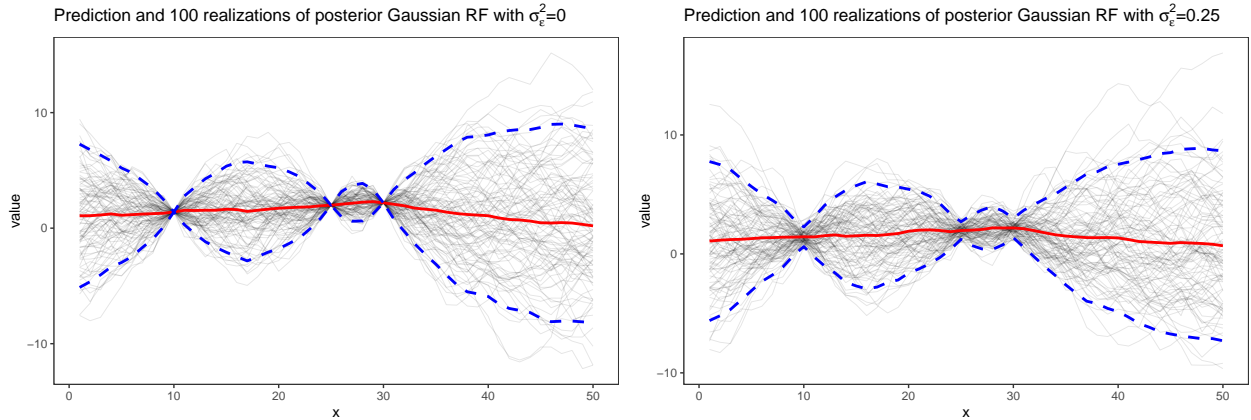
We compute and plot both the prediction and the interval.

Prediction of r(x) with 90% prediction interval and $\sigma_\varepsilon^2=0$ (left), Prediction of r(x) with 90% prediction interval and $\sigma_\varepsilon^2=0.25$ (right)

From the plots, we see that with no observation error, the prediction coincides with the observations in the points where we have an observation. Also, in these points there is no uncertainty, so the lower and upper bounds are also equal the observations. When observation error is present, we get uncertainty in the observed points. In addition, adding an observation error leads to larger uncertainty, and thus variance, over the entire prediction.

e)

We now go on to simulate 100 realizations from the posterior distribution, using the mean and covariance matrix previously computed. Then, we make a prediction based on the 100 samples and create a 90% prediction interval by computing the empirical variance.

Prediction and 100 realizations of posterior Gaussian RF with $\sigma_\varepsilon^2=0$ (left), Prediction and 100 realizations of posterior Gaussian RF with $\sigma_\varepsilon^2=0.25$ (right)

Discuss the relation between the model parameters and the realizations, and discuss the relation between the analytically and empirically obtained predictions with prediction intervals!!!!!!!!!

f)

SPØR OM DENNE!!!!

We use the previously generated 100 realizations with $\sigma_\epsilon^2 = 0$ to provide a prediction $\hat{A}_r$ for the non-linear function on $\{r(x); x \in D\}$,

$$A_r = \int_D I(r(x) > 2)dx.$$

We also compute the prediction variance. An alternative predictor is,

$$\tilde{A}_r = \sum_{x \in L} I(\hat{r}(x) > 2).$$

This is also computed.

A hat: 20.95

A tilde: 9

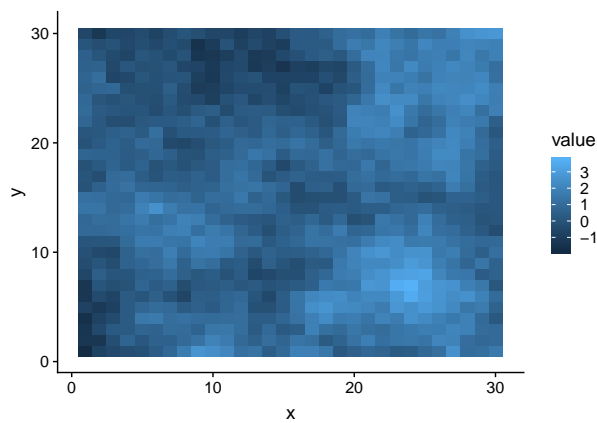## Predicted variance of A hat: 76.12879

Jensen's inequality states that for a random variable X and a convex function ψ,
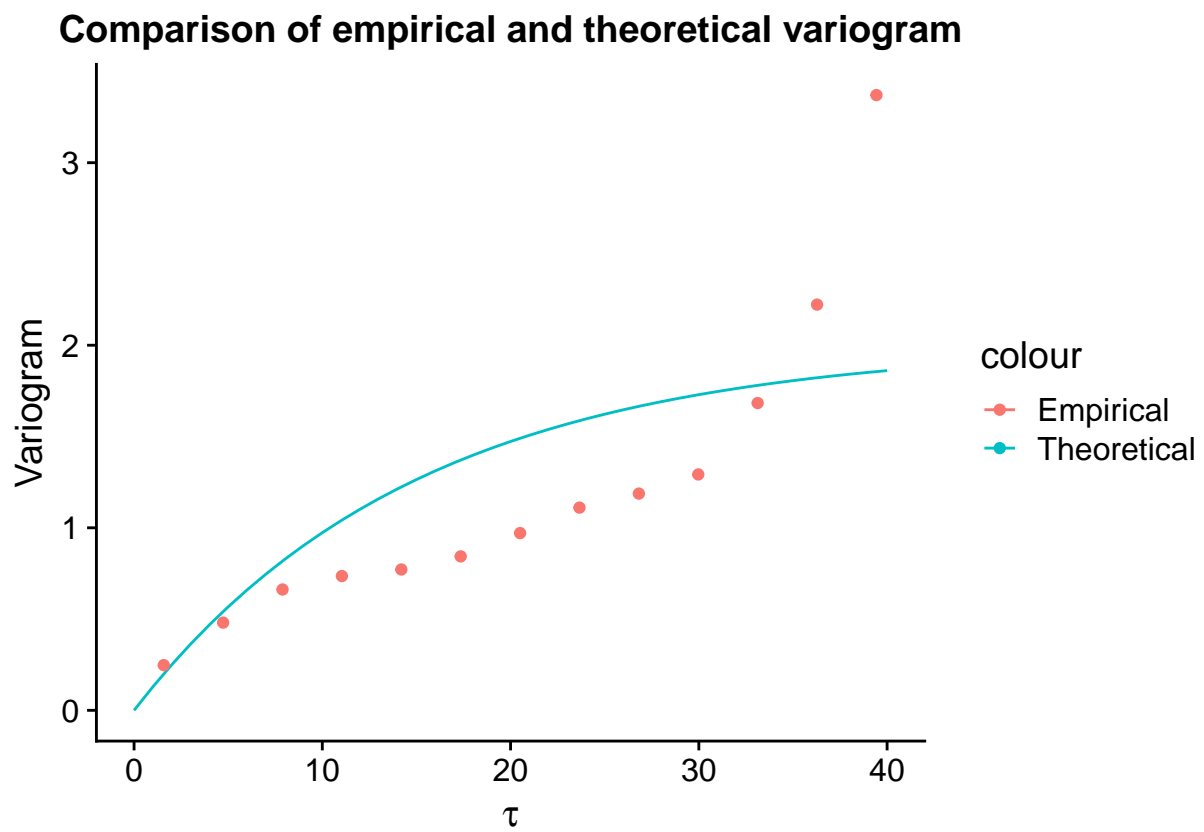
$$\psi(E[X]) \le E[\psi(X)].$$

Taking $X = r(x)$ and $\psi(\xi) = I(\xi > 2)$, this implies,??? ## g)
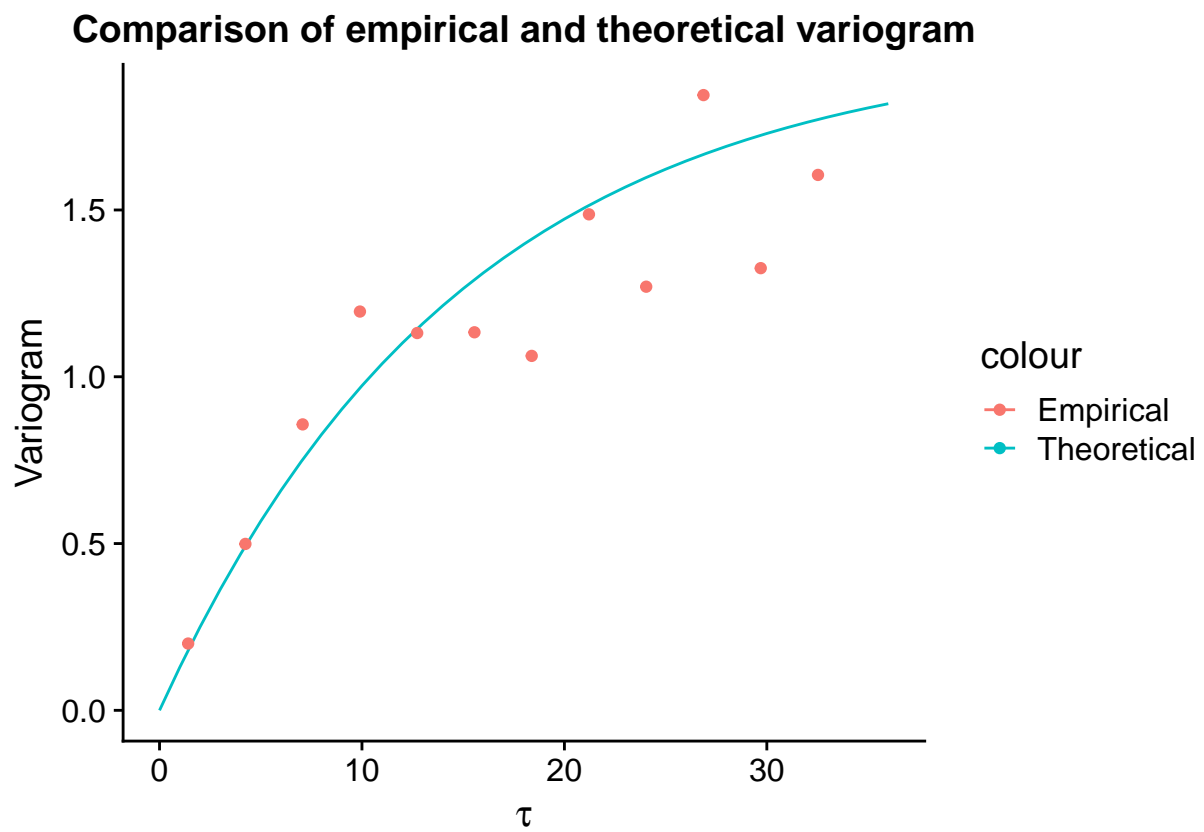
Problem 3: Parameter estimation

a)

b)

Comparison of empirical and theoretical variogram

c)

Comparison of empirical and theoretical variogram

Comparison of maxlik... and theoretical variogram