1.

2. GAUSSIAN PROCESSES

The file `KernelCode.R` distributed with the exam contains code to construct a `kernlab` function for the Matern covariance function with $\nu = 3/2$:

$$k(\mathbf{x}, \mathbf{x}') = \sigma_f^2 \left(1 + \frac{\sqrt{3}r}{\ell}\right) \exp\left(-\frac{\sqrt{3}r}{\ell}\right).$$

where $r = |\mathbf{x} - \mathbf{x}'|$.

(a) Let $f \sim GP(0, k(\mathbf{x}, \mathbf{x}'))$ a priori and let $\sigma_f^2 = 1$ and $\ell = 0.5$. Plot $k(0, z)$ as a function of z. You can use the grid `zGrid = seq(0.01,1,by=0.01)` for the plotting. Carefully interpret this plot in a way that shows that you understand what it shows. Connect your discussion to the smoothness of f. Finally, repeat this exercise with $\sigma_f^2 = 0.5$ and discuss the effect this change has on the distribution of f (2 p.)

(b) The file lidar.RData contains two variables `logratio` and `distance`. Load the variables into memory with the R command `load("lidar.RData")`. Compute the posterior distribution of f in the model

$$logratio = f(distance) + \varepsilon, \quad \varepsilon \sim N(0, 0.05^2).$$

You should do this for both length scales $\ell = 1$ and $\ell = 5$. Set $\sigma_f = 1$. Your answer should be in the form of a scatter plot of the data overlayed with curves for

i. the posterior mean of f
ii. 95% probability intervals for f
iii. 95% prediction intervals for a new data point y
Use the `gausspr` function in the kernlab package for i), but not for ii) and iii).
Discuss the differences in results from using the two length scales.
Do you think a GP with a Matern($\nu = 3/2$) kernel is a good model for this data? If not, what could be the problem with this model? Can you think of a better model? Discuss.
[Hint: $Cov(f) = K(X_\star, X_\star) - K(X_\star, X)\left[K(X, X) + \sigma_n^2 I\right]^{-1} K(X, X_\star)$ and remember that `%*%` does matrix multiplication and `solve` computes inverses in R] (2 p)

(c) (No need to do any computations here). Discuss how a Bayesian would handle the case where the kernel hyperparameters are unknown. What if the noise variance is unknown? (1 p)