Computer exercise 4
Covariance estimation and kriging
Spatial statistics and image analysis, TMS016

1 Introduction

The purpose of this computer exercise is to give an introduction to parameter estimation and kriging for Gaussian random field models for spatial data.

Before you begin, download the Matlab files for the exercise from the course homepage. When in doubt about how to use a specific function, use `help` and `doc` to get more information.

2 Data generation

Use the methods from Computer Exercise 3 to simulate a Gaussian random field $X(s)$ on a regular $n \times n$ lattice (let n be at least 50). Use a Matérn covariance function and assume a regression for the mean using the two basis functions $B_1(s) = 1$ and $B_2(s) = x$ (here x is the x-coordinate of $s = (x, y)$). This means that you can construct $X(s)$ as

$$X(s) = B_1(s)\beta_1 + B_2(s)\beta_2 + Z(s)$$

where $Z(s)$ is a mean-zero Gaussian random field with a Matérn covariance function. Based on a simulation of $X(s)$, construct N observations (let N be around 500)

$$Y_i = X(s_i) + \sigma_i, \quad i = 1, \dots, N$$

where $\sigma_i \sim \mathsf{N}(0, \sigma_e^2)$ are independent measurement noise terms and β_1, β_2 are some regression coefficients that you can choose as you want. Choose the observation locations $s_1, \dots, s_N$ by randomly selecting N of the locations on the lattice. This can be done by

```
>> ind = randperm(n*n);
>> ind_o = ind(1:n_obs);
>> loc_o = loc(ind_o,:);
```

Plot the simulated field $X(s)$ using `imagesc` as well as the observations using `scatter`. Using the `hold on` command, you can plot the two in the same figure. You might then have to set the two arguments `x` and `y` in `imagesc` so that the two plots use the same scale on the x and y axes.

3 Parameter estimation

We will now use the simulated data `y` to estimate the model parameters using the classical geostatistical approach.

- Start by estimating the mean μ using least-squares, and compute the residuals `e = y - mu`. Compare the estimated regression parameters to the true ones.
- Use the function `emp_variogram` to compute a binned estimate of the variogram and compare with the true variogram that can be computed using the function `matern_variogram`.
- Use the function `cov_ls_est` to perform least-squares estimation of a Matérn variogram to the binned estimate. Plot it together with the true variogram and the binned estimate.

- Update the estimate of the regression parameters using GLS. Compare with the OLS estimate as well as the true parameters. Hint: To compute the GLS estimate, you will need to compute the covariance matrix for the observations. Take a look at Computer Exercise 2 for how to compute the distance matrix `D` for the locations. The covariance matrix can then be obtained using the function `matern_covariance` with this matrix as the first argument. Remember to add the nugget effect to this matrix.

4 Kriging prediction

We will now use the estimated model parameters to perform kriging prediction. We reconstruct the field at all locations on the grid.

- Compute the needed matrices $\mathbf{\Sigma}_o$, $\mathbf{\Sigma}_p$, $\mathbf{\Sigma}_{op}$, $\mathbf{B}_o$, and $\mathbf{B}_p$. This can be done in a similar way as when you computed the covariance matrix for the GLS estimate: The easiest way is to merge the two set of locations, compute the corresponding covariance matrix, and finally extract the required matrices as blocks of this matrix.
- Compute the kriging predictor $\hat{X}$ and compare with the true field.

5 Likelihood-based parameter estimation

- Redo the estimation of the parameters using maximum-likelihood. This can be done using the `cov_ml_est` function. Compare the results to those obtained using least-squares.
- Re-compute the kriging predictor based on the ML parameter estimates and compare with the previous predictor. Is there a large difference?
- Compute the variances of the kriging predictions based on the two parameter estimates and compare the differences.