TMA4250 Spatial Statistics
Assignment 3: Mosaic Random Fields

H. Omre IMF/NTNU

April 2006

Introduction

This assignment contains problems related to Mosaic Random Fields or more specifically Markov Random Fields. The R-package should be used in solving the problems and relevant functions can be found in the R library `spatial` which can be loaded by the instruction `library(spatial)`.

Problem 1: Markov RF

This problem is based on observations of seismic data over a domain $\mathcal{D} \subset \mathbb{R}^2$. The objective is to identify the underlying lithology ({sand, shale}) distribution over $\mathcal{D}$.

The data are collected on a regular (75×75) grid $\mathcal{L}_\mathcal{D}$, and the seismic data are denoted $o : \{o_x; x \in \mathcal{L}_\mathcal{D}\}$. The data are available in the R library MASS in the file *seismic.dat*.

Moreover, observations of the lithology distribution ({sand, shale}) in a geologically comparable domain $\mathcal{D}_c \subset \mathbb{R}^2$ is available. The lithology distribution is collected on a regular (66×66) grid $\mathcal{L}_{\mathcal{D}_c}$, over $\mathcal{D}_c$. The observations with code 0 for sand and 1 for shale is available in the R library MASS in the file *complit_new.dat* and are shown in Figure 1.

Assume that the underlying lithology surface can be represented by a mosaic RF $\{L(x); x \in \mathcal{D} \subset \mathbb{R}^2\}$ discretized into $L : \{L_x : x \in \mathcal{L}_\mathcal{D}\}$ with $L_x \in \{0, 1\}$ representing sand and shale respectively.

The seismic data collection procedure defines the likelihood model:

$$[o_x | L = l] = \begin{cases} 0.02 + U_x \text{ if } l_x = \text{sand} \\ 0.08 + U_x \text{ if } l_x = \text{shale} \end{cases} \quad ; x \in \mathcal{L}_\mathcal{D}$$

with $U_x; x \in \mathcal{L}_\mathcal{D}$ iid $\text{Gauss}(0, 0.06^2)$.

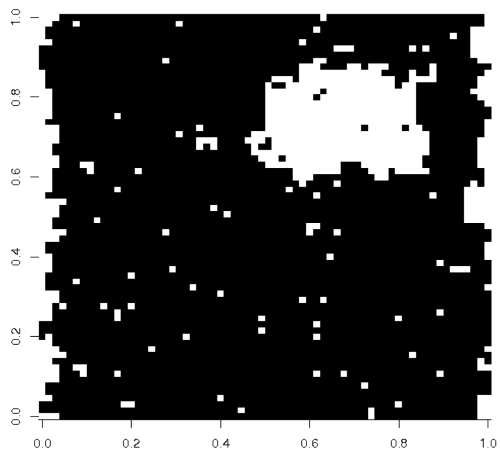

Figure 1: Observed data

a) Consider a uniform prior model on L, i.e.

$$\text{Prob}\{L = l\} = \text{const.}$$

Develop an expression for the posterior model $\text{Prob}\{L = l|o\}$, simulate 10 realizations of the posterior mosaic RF $\{L_x; x \in \mathcal{L}_\mathcal{D}|o\}$, and display them. Develop expressions for the posterior expectation $\text{E}\{L|o\}$ and variance $\text{Var}\{L|o\}$, and display them. Develop expressions for marginal maximum aposteriori $\text{MMAP}\{L|o\}$, and display the results. Comment on the results.

b) Consider a Markov RF (MRF) prior model for L with neighborhood system $\partial : \{\partial_x; x \in \mathcal{L}_\mathcal{D}\}$ consisting of four closest neighbors:

$$\text{Prob}\{L_x = l_x|L_y = l_y; y \in \partial_x\} = \text{const} \times \exp\{\beta \sum_{y \in \partial_x} \text{I}(l_y = l_x)\} \; ; \; \forall x \in \mathcal{L}_\mathcal{D}$$

with $\text{I}(A) = 1$ if A is true and $\text{I}(A) = 0$ else.

Specify the associated Gibbs formulation for the MRF, i.e. $\text{Prob}\{L = l\}$. Develop expressions for the posterior models $\text{Prob}\{L = l|o\}$ and $\text{Prob}\{L_x = l_x|L_{-x} = l_{-x}, o\}; \forall x \in \mathcal{L}_\mathcal{D}$. We are interested in realizations from $\text{Prob}\{L = l|o\}$ and estimates of $\text{E}\{L|o\}$, $\text{Var}\{L|o\}$ and MMAP. Explain how they can be determined.

Use the observations from the geologically comparable domain $\mathcal{D}_c$ to estimate β by a maximum pseudo-likelihood procedure. Use torus boundary conditions and denote the estimate $\hat{\beta}$.

c) Set the parameter $\beta = \hat{\beta}$ and use a Gibbs sampler MCMC-algorithm to simulate from the posterior $\text{Prob}\{L = l|o\}$. Use torus boundary conditions to avoid border problems. Document that the algorithm has converged. Display 10 independent realizations.

Estimate $\text{E}\{L|o\}$, $\text{Var}\{L|o\}$ and $\text{MMAP}\{L|o\}$ and display the results.

Comment on the results.

d) Compare the results in **a)** and **c)** and comment on them.