STAT 517: Stochastic Modeling II **Winter 2023**

Homework 01 - Partial solutions

Instructor: Abel Rodriguez *TA: Apara Venkat*

Note: This is only a high-level sketch of the solutions. If you catch any typos or have questions, post them in the discussion board.

1. Start by ordering nodes on the lattice from left to right and top to bottom. So node at location (1, 1) would be the first node in the sequence, node at location (2, 1) would be the $L+1$-st node in the sequence, and so on. Then, for nodes not on the left or top boundary,

$$P(X_t \mid X_1, \ldots, X_{t-1}) = P(X_t \mid X_{t-1}, X_{t-L})$$

Nodes at the left (top) boundary have dependence only on X_{t-L} (X_{t-1}). Thus, we have an L-th order Markov chain.

We can use backpropagation to evaluate the partition function. Recursively compute starting from $t = L^2 - L$ to 1:

$$\begin{aligned} q(X_{t+L-1} = x_{t+L-1}, \ldots, X_t = x_t) = \sum_{i \in \{-1,1\}} & P(X_{t+L} = i \mid X_{t+L-1} = x_{t+L-1}, \ldots, X_t = x_t) \times \\ & q(X_{t+L} = i, \ldots, X_{t+1} = x_{t+1}) \end{aligned}$$

for all possible $x_i \in \{-1, 1\}$. For $t = L^2 - L$ to L^2, set $q(\cdot) = 1$. Finally, evaluate,

$$Z = \sum_i P(X_1 = i) \sum_{x_t} q(X_1 = x_1, \ldots, X_L = x_L)$$

Each $q(\cdot)$ takes $\mathcal{O}(2^L)$ operations. There are $\mathcal{O}(L^2)$ of them. Thus, the complexity is $\mathcal{O}(L^2 2^L)$ which is still exponential.

2. Define $b_{ij} = \theta_1(\mathbb{I}(j = N(i)) + \mathbb{I}(j = S(i))) + \theta_2(\mathbb{I}(j = W(i)) + \mathbb{I}(j = E(i)))$. Define a vector $b_i \in \mathbb{R}^L$ with each element defined as described. Then,

$$Z_i = \mu_i + b_i^\top (Z - \mu) + \epsilon_i$$

Vectorizing this, we get $Z = \mu + B(Z - \mu) + \epsilon$ where B is just composed of the b_i vectors. Then, results from multivariate gaussians give the desired result. And it immediately follows that this is a CAR model.

3. See code on Canvas.

4. See [Preisler, 1993] for a spatial model and discussion on estimating standard errors. Key idea here is that any regular GLM implementation will give consistent estimates for the parameters of the model but the naive GLM errors are incorrect because of spatial dependence. We need to bootstrap to obtain confidence intervals. Code is available on Canvas.

5. Need to implement a noisy exchange algorithm. See code on Canvas.

References

[Preisler, 1993] Preisler, H. K. (1993). Modelling spatial patterns of trees attacked by bark-beetles. *Journal of the Royal Statistical Society Series C: Applied Statistics*, 42(3):501–514.