Project 2 in course STAT340 – 2020

November 10, 2020

- **Deadline: Monday 7th Dec. The report (with code and all the derivations) should be submitted in blackboard, preferable as a R-markdown document.**
- You may hand in a joint report with maximum two in each group.

Consider the Tokyo-rainfall example in Chapter 4 in the GMRF book. Download the data as follows

```
> tfile=tempfile()
> download.file("http://hrue.r-inla-download.org/Tokyo.RData", destfile=tfile)
trying URL 'http://hrue.r-inla-download.org/Tokyo.RData'
Content type 'unknown' length 991 bytes
==================================================
downloaded 991 bytes
> load(tfile)
> str(Tokyo)
'data.frame':   366 obs. of  3 variables:
 $ y   : int  0 0 1 1 0 1 1 0 0 0 ...
 $ n   : int  2 2 2 2 2 2 2 2 2 2 ...
 $ time: int  1 2 3 4 5 6 7 8 9 10 ...
> unlink(tfile)
```

Remove the line with `n=1` using (it's not worth the 'pain')

```
> Tokyo = Tokyo[Tokyo$n==2,]
> Tokyo$time = 1:nrow(Tokyo)
```

and use the probit-link. This makes the implementation of the MCMC algorithms easier, and all days have two observations.

The task is to implement and compare the "properties" of the following MCMC algorithms. (Many of the details are spelled out in Chapter 4 in the GMRF book.)

1. Single site sampler with a simple proposal for the spline
2. Single site sampler with a Gaussian-approximation proposal for the spline
3. Block update samples, with two blocks: the precision parameter and the spline (using the Gaussian approximation)
4. One-block sampler, where you sample the precision parameter and the spline (using the Gaussian approximation) in one block.

5. The auxiliary variable approach, where you update in the three blocks: the precision, the spline and the auxiliary variables.

6. The auxiliary variable approach, where you update in the two blocks: the precision and the spline jointly, and the auxiliary variables.

Notes

- Scale the (circulant) precision matrix with 67500 and use an exponential prior for $1/\sqrt{\kappa}$ with rate 1.55, where κ is the precision parameter for the RW2 model. The structure matrix for the circulant spline would then be

  ```
  > inla.as.sparse(toeplitz(c(6,-4,1,rep(0,5),1,-4)))
  10 x 10 sparse Matrix of class "dgTMatrix"

   [1,]  6 -4  1  .  .  .  .  .  1 -4
   [2,] -4  6 -4  1  .  .  .  .  .  1
   [3,]  1 -4  6 -4  1  .  .  .  .  .
   [4,]  .  1 -4  6 -4  1  .  .  .  .
   [5,]  .  .  1 -4  6 -4  1  .  .  .
   [6,]  .  .  .  1 -4  6 -4  1  .  .
   [7,]  .  .  .  .  1 -4  6 -4  1  .
   [8,]  .  .  .  .  .  1 -4  6 -4  1
   [9,]  1  .  .  .  .  .  1 -4  6 -4
  [10,] -4  1  .  .  .  .  .  1 -4  6
  ```

  if n was 10. In this project $n = 365$. Scale the matrix with the above mentioned number, since

  ```
  > R = 67500 * toeplitz(c(6,-4,1,rep(0,360),1,-4))
  > exp(mean(log(diag(INLA:::inla.ginv(R, rankdef=1)))))
  [1] 1.000633234
  ```

- Make sure you take advantage of the sparsity of the precision matrices using the algorithm(s) from Chapter 2 in the GMRF-book. All the posterior marginals should be the same for all algorithms, after a long run, of'course.

- Be aware that for some of the samplers, the convergence will be extremely slow. Maybe it will not even converge within the time you have.