Methods

Marshall Honaker

First, let's import the necessary packages:

```
library(fields)
```

```
## Loading required package: spam
```

```
## Loading required package: dotCall64
```

```
## Loading required package: grid
```

```
## Spam version 2.5-1 (2019-12-12) is loaded.
## Type 'help( Spam)' or 'demo( spam)' for a short introduction
## and overview of this package.
## Help for individual functions is also obtained by adding the
## suffix '.spam' to the function name, e.g. 'help( chol.spam)'.
```

```
##
## Attaching package: 'spam'
```

```
## The following objects are masked from 'package:base':
##
##     backsolve, forwardsolve
```

```
## See https://github.com/NCAR/Fields for
##  an extensive vignette, other supplements and source code
```

```
library(RandomFields)
```

```
## Loading required package: sp
```

```
## Loading required package: RandomFieldsUtils
```

```
##
## Attaching package: 'RandomFields'
```

```
## The following object is masked from 'package:RandomFieldsUtils':
##
##     RFoptions
```

```
library(mvtnorm)
```

Wu and Narisetty (2020) outline a score based likelihood approach to Bayesian multiple quantile regression. Here we provide R code for a modified version of this approach adapted for spatial data. Annotations will be made/provided to explain the ways in which the approach taken by Wu & Narisetty was modified to accomdodate spatially dependent data.

To begin, we will define so of the more basic, necessary functions. First, we will define the score function.

Basic Functions & Methods

In the aforementioned paper, the score function is given by: $s_\tau(\beta) = \sum_{i=1}^{n} x_i \psi_\tau(y_i - x_i^T\beta)$, where $\psi_\tau(u) = \tau - I_{\{u<0\}}(u)$. Note that in this case, $\psi_\tau(.)$ is essentially the check-loss function mentioned earlier in the paper applied top the residuals.

```
get_score <- function(y, X, tau, beta){
  ## outputs px1 vector
  u <- y - X%*%beta
  psi <- ifelse(u < 0, tau - 1, tau)
  return(t(X)%*%psi)
}
```

I'm also going to define a function to obtain the spatial covaraince matrix. I'm going to set it to use the Matérn covariance function with $\nu = 1.5$ and $\phi = 1.5$ but we can change this later or make it more adaptive as need be.

```
get_spatial_covar_mat <- function(locs){  ## outputs nxn spatial covariance matrix
  dist <- rdist(locs)
  return(Matern(dist, range = .2, nu = 1.5))
}
```

Next, we will write a method to obtain a spatially adapted rendition of the proposed working likelihood function. The working likelihood function proposed in the paper is given by $L(Y|X,\beta) = C\exp\left(-\frac{1}{2n}s_\tau(\beta)^T W s_\tau(\beta)\right)$ where W is a p x p weight matrix. Now, to account for the spatial variability/dependence within the data, we will instead use the sample Mahalanobis distance given by $(X-\hat{\mu})^T\Sigma^{-1}(X-\hat{\mu})$ where Σ is the spatial covariance matrix. *Note that while the paper requires W to be positive definite, the sample Mahalanobis matrix is only positive semi-definite. We will need to find a way to make sure that this this change will still lead to valid inference based on the posterior later on.*

```
get_likelihood <- function(y, X, tau, beta, locs, C){
  ## Outputs 1x1 scalar that is the likelihood
  score <- get_score(y, X, tau, beta)
  X_centered <- X - colMeans(X)
  coef <- -1/(2*length(y))
  kernel <- exp(coef*t(score)%*%t(X_centered)%*%solve(get_spatial_covar_mat(locs))%*%X_centered%*%score)
  return(C*kernel[1,1])
}
```

Testing basic methods

Now that we have defined the basic functions, let's test them to make sure they work.

```
n <- 100
locs <- cbind(runif(n, 0, 10), runif(n, 0, 10))
m1 <- RMexp(var = 2, scale = 1.5) +
  RMnugget(var = .1) +
  RMtrend(mean = 1)
sim_vals <- RFsimulate(m1, x = locs[,1], y = locs[,2])
```

```
## New output format of RFsimulate: S4 object of class 'RFsp';
## for a bare, but faster array format use 'RFoptions(spConform=FALSE)'.
```

```
test_X <- cbind(rnorm(n), rnorm(n), rnorm(n), rnorm(n))
test_beta <- c(1, 2, 3, 4)
test_y <- test_beta[2]*locs[,1] + test_beta[3]*locs[,2] + sim_vals$variable1
test_tau <- .5
```

```
get_score(test_y, test_X, test_tau, test_beta)
```

```
##            [,1]
## [1,] -5.0762052
## [2,]  0.5226628
## [3,]  2.0288405
## [4,]  4.8776837
```

```
get_likelihood(test_y, test_X, test_tau, test_beta, locs, 1)
```

```
## [1] 4.635878e-24
```

So it seems like everything is working alright. Now let's move on to the importance sampling.

Ada Importance Sampling for a Single Quantile Level

Wu & Narisetty describe an Importance Sampling (IS) procedure on pgs. 13-14. I've done my best to code it here, but there are still some issues. I will discuss these issues later.

The `get_updated_params()` function simply returns a list whose elements are the parameter values (the mean vector and the covariance matrix) for a given set of simulated observations. This makes the `adaIS_singleQuantile()` function a little more readable when we have to update the parameter values in each iteration.

```
get_updated_params <- function(y, X, tau, beta, Sigma, locs, draw){
  ## outputs list with mu (which is a p-length vector of estimated means) and S (which is the estimated
  n <- length(y)
  p <- ncol(X)

  w <- get_likelihood(y, X, tau, beta, locs, 1)*(1/(2*n)^p)/dmvnorm(draw, beta, Sigma)

  mu_hat <- apply(w*draw, 2, sum)/sum(w)

  S <- matrix(nrow = p, ncol = p)
  for(i in 1:p){
    for(j in 1:p){
      S[i,j] <- sum(w*(draw[,i] - mu_hat[i])*(draw[,j] - mu_hat[j]))/(sum(w)*(n - 1))
    }
  }
  temp <- list(mu_hat, S)
  names(temp) <- c("mu", "S")
  return(temp)
}

adaIS_singleQuantile <- function(y, X, tau, C, locs, M, num_reps){
  ## outputs list with $mu (a p-length vector of estimated means) and $S (the estimated pxp covariance
  p <- ncol(X) ## Step1: Initialize starting values for IS algorithm
  n <- length(y) ## Step1: Initialize starting values for IS algorithm
  beta <- coef(lm(y ~ X))[-1] ## Step1: Initialize starting values for IS algorithm
  S <- cov(X) ## Step1: Initialize starting values for IS algorithm

  params <- list(beta, S)  ## Put out initial parameters into a list called params so we can update it
  names(params) <- c("mu", "S")

  for(i in 1:num_reps){ ## Step 4: Repeat steps 2 and 3 until we achieve the desired effective sample s
```

```
    draws <- rmvnorm(M, params$mu, params$S)  ## Step 2: Simulate M values from the proposal distributi
    params <- get_updated_params(y, X, tau, params$mu, params$S, locs, draws) ## Step 3: Update the par
  }
  return(params)
}
```

Now, to test our IS method.

```
m <- 10000
S_0 <- cov(test_X)
initial_draw <- rmvnorm(m, test_beta, S_0)

get_updated_params(test_y, test_X, test_tau, test_beta, cov(test_X), locs, initial_draw)
```

```
## $mu
## [1] 1.958582 1.087302 1.789942 4.061821
##
## $S
##              [,1]        [,2]         [,3]         [,4]
## [1,]  0.054408605 0.002453104 -0.016339267  0.015565465
## [2,]  0.002453104 0.034673553  0.004946878  0.012238539
## [3,] -0.016339267 0.004946878  0.034706794 -0.008864378
## [4,]  0.015565465 0.012238539 -0.008864378  0.030833007
```

```
adaIS_singleQuantile(test_y, test_X, test_tau, 1, locs, m, 22)
```

```
## $mu
##         X1         X2         X3         X4
## -1.1213664  0.5923757 -1.1332155  0.9113398
##
## $S
##      [,1]          [,2]          [,3]          [,4]
## [1,]    0  0.000000e+00  0.000000e+00  0.000000e+00
## [2,]    0  3.479491e-80  6.958981e-80 -6.958981e-80
## [3,]    0  6.958981e-80  1.391796e-79 -1.391796e-79
## [4,]    0 -6.958981e-80 -1.391796e-79  3.735137e-34
```

So everything seems to be working, thought it does need a little tuning/refinement.

Issues so Far

So it seems like the issue with the `adaIS_singleQuantile()` method is with the number of iterations of the IS, not the number of draws on each iteration. I was able to run the IS with 1000000 draws per iteration for 10 iterations and it ran just fine. (Generally, we seem to be able to make the number of draws as large as we want as long as the number of iterations is relatively low.) However, when we increase the number of iterations anywhere above ~13, we get the following error:

```
Error in eigen(sigma, symmetric = TRUE) : infinite or missing values in 'x'
```

I'm not entirely sure what is causing this, but I have two theories:

One, it could have something to do with the calculations for the posterior covariance matrix. I mentioned earlier, when we swap the weight matrix W for the sample spatial Mahalanobis distance in the working likelihood, the formula for the posterior covariance given in the paper no longer yields a positive definite covariance matrix. I found a work around to get the IS to work in the short term, but it doesn't take into account the importance weights.

Two, this could actually be because we swapped the weight matrix for the sample spatial Mahalanobis distance. The paper specifies that W must be positive definite, but the sample Mahalanobis distance is only positive *semi*-definite. That may explain why it doesn't work for larger iterations (eventually, the IS just randomly encounters the "semi-definite" case in which one of the Eigenvalues is effectively zero and it can't perform the necessary computations).

Because the likelihood function $L(Y|X,\beta)$ is defined with respect to the score function $s_\tau(\beta)$ (that is, $L(Y|X,\beta) = C\exp\left(-\frac{1}{2n}s_\tau(\beta)^T(X-\hat{\mu})^T\Sigma^{-1}(X-\hat{\mu})s_\tau(\beta)\right)$) and the score function is the gradient of the log-likelihood with respect to the parameters, we will need to solve a differential equation to get a functional form for the score function and likelihood.

UPDATE:

I think the use of the term "score function" here may be a little misleading. Though I've never seen it defined in this way, I found this slide deck developed by Koenker himself which sheds a little light on the subject. After going through this (especially slide 15), I don't think $s_\tau(\beta)$ is the traditional "score" function (i.e; the gradient of the log-likelihood). It seems like when they said "score function" they meant "*rank*score function", which is another way of defining quantiles. It seems that the rankscore function actually comes from the quantile regression objective function $\hat{\beta}(\tau) = \arg\min_\beta \sum_{i=1}^n \rho_\tau(y_i - x_i^T\beta)$, there's no need to solve any sort of more complicated differential equation since the likelihood is no longer a function of its own derivative.