grmbayes

```r
library(grmbayes)
```

Introduction

`grmbayes` provides a suite of functions to fit Bayesian geostatistical regression models, originally designed for use with air quality data.

There are three primary functions in the package: - `grm`: fits a Bayesian geostatistical regression model using MCMC - `grm_cv`: fits a Bayesian geostatistical regression model using MCMC with cross-validation - `grm_pred`: predicts values at new locations using a fitted Bayesian geostatistical regression model

Examples of each of these functions are provided in this vignette using a dataset of $PM_{2.5}$ and Chemical Transport Model (CTM) readings collected in the Atlanta Metropolin Area from 2003 to 2005. Run `?cmaq_aqs_matched` for more info. A dataset of $PM_{2.5}$ and satellite-collected Aerosol Optical Depth (AOD) over the same area is also included in this package (`modis_aqs_matched`) but is not used in this vignette.

grm

`grm` fits a Bayesian geostatistical regression model using MCMC, detailed as follows:

$$Y(s,t) = \zeta_0(s,t) + \zeta_1(s,t)X(s,t) + \epsilon(s,t)$$

where

$$\zeta_0(s,t) = \alpha_0(s) + \beta_0(t) + \gamma_0 L + \delta M \zeta_1(s,t) = \alpha_1(s) + \beta_1(t) + \gamma_1 L + \delta M$$

The model is specified in `grmbayes` using the following arguments: - `Y`: a vector of response values (e.g., $PM_{2.5}$) - `X`: a vector of primary covariate values (e.g., CTM) - `L`: a matrix of secondary temporal covariate values (e.g., elevation, forest cover, etc.) - `M`: a matrix of secondary spatio-temporal covariate values (e.g., temperature, wind speed, etc.) - `nngp`: a boolean indicating whether to use a non-negative Gaussian process (NNGP) prior for the spatial component of the model, as opposed to a regular Gaussian process - `covariance`: the covariance function to use for the spatial component of the model - `matern.nu`: the smoothness parameter for the Matern covariance function (if `covariance = "matern"`)

`space.id`, `time.id`, and `spacetime.id` are also required to specify the model. These are vectors of unique identifiers for each location, time, and spatio-temporal combination, respectively. The unique identifiers must be integers starting at 1 and increasing by 1 for each additional location, time, or spatio-temporal combination

‘

```r
cmaq_fit <- grm(Y = cmaq_aqs_matched$pm25,
                X = cmaq_aqs_matched$ctm,
```

```
L = cmaq_aqs_matched[, c("elevation", "forestcover",
                 "hwy_length", "lim_hwy_length",
                 "local_rd_length", "point_emi_any")],
M = cmaq_aqs_matched[, c("tmp", "wind")],
n.iter = 100,
burn = 20,
thin = 1,
nngp = T,
covariance = "matern",
matern.nu = 1.5,
coords = cmaq_aqs_matched[, c("x", "y")],
space.id = cmaq_aqs_matched$space_id,
time.id = cmaq_aqs_matched$time_id,
spacetime.id = cmaq_aqs_matched$spacetime_id,
verbose.iter = 10)
```