空间广义线性模型

代码实现

黄湘云

2018 年 3 月

目录

1 模型形式

1.1 贝叶斯框架

分层多水平模型

1.1.1 泊松

1.1.2 二项

2 程序实现

2.1 PrevMap 包

(Giorgi and Diggle, 2017) 将 MCML 和 MCMC 方法应用于空间广义线性混合效应模型的参数估计和预测，

2.2 geoR 与 geoRglm 包

2.3 Stan 框架

Stan[1] 是一种概率编程语言 (Carpenter et al., 2017)，可以替代 BUGS (**B**ayesian inference **U**sing **G**ibbs **S**ampling) (Lunn et al., 2009) 作为 MCMC 的高效实现，可用于贝叶斯框架下，标准地统计模型的参数估计，Stan 提供多种语言的接口实现，方便起见，本文采用它提供的 R 语言接口 – rstan 包 (Stan Development Team, 2018)。

```
suppressPackageStartupMessages(library(rstan))
```

2.4 PyMC 框架

This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see http://rmarkdown.rstudio.com.

[1] http://mc-stan.org/

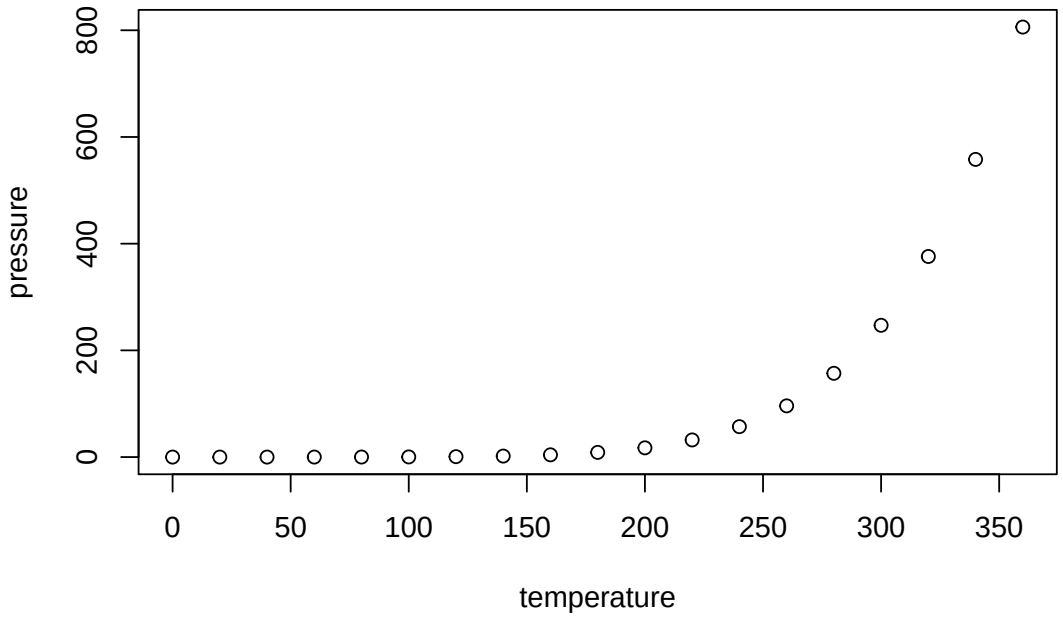
When you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this:

```
summary(cars)
```

```
##      speed           dist
##  Min.   : 4.0   Min.   :  2.00
##  1st Qu.:12.0   1st Qu.: 26.00
##  Median :15.0   Median : 36.00
##  Mean   :15.4   Mean   : 42.98
##  3rd Qu.:19.0   3rd Qu.: 56.00
##  Max.   :25.0   Max.   :120.00
```

2.5 CUDA 框架

基于 GPU 加速是一个不错的选择，Stan 开发者也把 GPU 加速列入开发日程。scikit-cuda (Givon et al., 2015) ArrayFire (Yalamanchili et al., 2015) 等基于 CUDA 开发的通用加速框架获得越来越多的关注。



Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that generated the plot.

2.6 R 进程信息

```
sessionInfo()
```

```
## R version 3.4.3 (2017-11-30)
## Platform: x86_64-w64-mingw32/x64 (64-bit)
## Running under: Windows 8.1 x64 (build 9600)
##
## Matrix products: default
##
## locale:
## [1] LC_COLLATE=Chinese (Simplified)_China.936
## [2] LC_CTYPE=Chinese (Simplified)_China.936
## [3] LC_MONETARY=Chinese (Simplified)_China.936
## [4] LC_NUMERIC=C
## [5] LC_TIME=Chinese (Simplified)_China.936
##
## attached base packages:
## [1] stats     graphics  grDevices utils     datasets  methods   base
##
## other attached packages:
##  [1] rstan_2.17.3         StanHeaders_2.17.2   ggplot2_2.2.1
##  [4] RevoUtilsMath_10.0.1 RevoUtils_10.0.7     RevoMods_11.0.0
##  [7] MicrosoftML_9.3.0    mrsdeploy_1.1.3      RevoScaleR_9.3.0
## [10] lattice_0.20-35      rpart_4.1-12
##
## loaded via a namespace (and not attached):
##  [1] Rcpp_0.12.15           compiler_3.4.3         pillar_1.1.0
##  [4] plyr_1.8.4             rticles_0.4.1          iterators_1.0.9
```

```
##  [7] tools_3.4.3          digest_0.6.15        jsonlite_1.5
## [10] evaluate_0.10.1      tibble_1.4.2         gtable_0.2.0
## [13] rlang_0.1.6.9003     foreach_1.4.4        CompatibilityAPI_1.1.0
## [16] curl_3.1             yaml_2.1.16          gridExtra_2.3
## [19] stringr_1.2.0        knitr_1.19           stats4_3.4.3
## [22] rprojroot_1.3-2      grid_3.4.3           inline_0.3.14
## [25] R6_2.2.2             rmarkdown_1.8        magrittr_1.5
## [28] backports_1.1.2      scales_0.5.0         codetools_0.2-15
## [31] htmltools_0.3.6      colorspace_1.3-2     stringi_1.1.6
## [34] lazyeval_0.2.1       munsell_0.4.3
```

3 算法细节

主要基于 Stan 框架实现

```
library(rstan)
```

3.1 两个证明

渐进正态性和相合性

3.2 符号约定

斜体用于扩展包和框架，如 *knitr*、*PrevMap*、*CUDA*、*Stan* 等，粗体用于软件，如 **R**、**Python** 等，等宽体用于代码和代码块。

> 可能由于 **Pandoc** 转化不当，*knitr* 出来的 PDF 文档，其目录中参考文献是英文 Bibliography，因此需要手动修改 tex 文件的倒数第二行，将 bibname 改为 refname，然后在 **R** 控制台执行 `tinytex::xelatex(file = 'draft4report.tex')` 。

参考文献

Carpenter, B., Gelman, A., Hoffman, M., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M., Guo, J., Li, P., and Riddell, A. (2017). Stan: A probabilistic programming language. *Journal of Statistical Software*, 76(1):1–32.

Giorgi, E. and Diggle, P. J. (2017). PrevMap: An R package for prevalence mapping. *Journal of Statistical Software*, 78(8):1–29.

Givon, L. E., Unterthiner, T., Erichson, N. B., Chiang, D. W., Larson, E., Pfister, L., Dieleman, S., Lee, G. R., van der Walt, S., Menn, B., Moldovan, T. M., Bastien, F., Shi, X., Schlüter, J., Thomas, B., Capdevila, C., Rubinsteyn, A., Forbes, M. M., Frelinger, J., Klein, T., Merry, B., Merill, N., Pastewka, L., Clarkson, S., Rader, M., Taylor, S., Bergeron, A., Ukani, N. H., Wang, F., and Zhou, Y. (2015). scikit-cuda 0.5.1: a Python interface to GPU-powered libraries.

Lunn, D., Spiegelhalter, D., Thomas, A., and Best, N. (2009). The bugs project: Evolution, critique and future directions. *Statistics in Medicine*, 28(25):3049–3067.

Stan Development Team (2018). RStan: the R interface to Stan. R package version 2.17.3.

Yalamanchili, P., Arshad, U., Mohammed, Z., Garigipati, P., Entschev, P., Kloppenborg, B., Malcolm, J., and Melonakos, J. (2015). ArrayFire - A high performance software library for parallel computing with an easy-to-use API.