Homework4

Laha Ale

March 17, 2019

```
library(spdep)
```

```
## Warning: package 'spdep' was built under R version 3.5.3
## Loading required package: sp
## Loading required package: Matrix
## Loading required package: spData
## Warning: package 'spData' was built under R version 3.5.3
## To access larger datasets in this package, install the spDataLarge
## package with: `install.packages('spDataLarge',
## repos='https://nowosad.github.io/drat/', type='source')`
## Loading required package: sf
## Warning: package 'sf' was built under R version 3.5.3
## Linking to GEOS 3.6.1, GDAL 2.2.3, PROJ 4.9.3
```

```
library(maps)
library(maptools)
```

```
## Warning: package 'maptools' was built under R version 3.5.3
## Checking rgeos availability: FALSE
##      Note: when rgeos is not available, polygon geometry     computations in maptools depend on gpcl
##      which has a restricted licence. It is disabled by default;
##      to enable gpclib, type gpclibPermit()
```

```
library(classInt)
library(RColorBrewer)
```

Excercise 11

(a)

```
state.sat.scores = read.table("state-sat.dat", header=F)
colnames(state.sat.scores) <- c("STATE","VERBAL","MATH","ELIGIBLE")
head(state.sat.scores)
```

```
##    STATE VERBAL MATH ELIGIBLE
## 1    ala    561  555        9
## 2 alaska    516  514       50
## 3   ariz    524  525       34
## 4    ark    563  556        6
## 5  calif    497  514       49
## 6   colo    536  540       32
```

```
# create listw
usa.state <- map(database="state", fill=TRUE, plot=FALSE)
state.ID <- sapply(strsplit(usa.state$names, ":"),
```

```r
                   function(x) x[1])

usa.poly <- map2SpatialPolygons(usa.state,
                               IDs=state.ID)
usa.nb <- poly2nb(usa.poly)
usa.listb <- nb2listw(usa.nb, style="B")
usa.listw <- nb2listw(usa.nb, style="W")

# train SAR model

x = ((state.sat.scores$STATE=="alaska") |
       (state.sat.scores$STATE=="hawaii") |
       (state.sat.scores$STATE=="us"))

index = c(1:nrow(state.sat.scores))[x]
state.sat.scores.contig = state.sat.scores[-index,]

# binnary weights
stat.sat.sar.b = spautolm(ELIGIBLE~ VERBAL+MATH,
                        data=state.sat.scores.contig,
                        family="SAR",
                        listw=usa.listb,
                        zero.policy=TRUE)
summary(stat.sat.sar.b)
```

```
##
## Call:
## spautolm(formula = ELIGIBLE ~ VERBAL + MATH, data = state.sat.scores.contig,
##     listw = usa.listb, family = "SAR", zero.policy = TRUE)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -24.58922  -7.52705   0.86387   5.96918  16.47887
##
## Coefficients:
##               Estimate Std. Error z value  Pr(>|z|)
## (Intercept) 390.464541  30.188641 12.9342 < 2.2e-16
## VERBAL       -0.675203   0.202635 -3.3321 0.0008619
## MATH          0.022018   0.196717  0.1119 0.9108791
##
## Lambda: 0.15957 LR test value: 26.091 p-value: 3.2565e-07
## Numerical Hessian standard error of lambda: 0.016181
##
## Log likelihood: -179.7922
## ML residual variance (sigma squared): 75.416, (sigma: 8.6842)
## Number of observations: 49
## Number of parameters estimated: 5
## AIC: 369.58
```

```r
# row-normalized weights
stat.sat.sar.w = spautolm(ELIGIBLE~ VERBAL+MATH,
                        data=state.sat.scores.contig,
                        family="SAR",
                        listw=usa.listw,
```

```
                          zero.policy=TRUE)
summary(stat.sat.sar.w)
```

```
##
## Call:
## spautolm(formula = ELIGIBLE ~ VERBAL + MATH, data = state.sat.scores.contig,
##     listw = usa.listw, family = "SAR", zero.policy = TRUE)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -21.8003  -5.4151  -0.5949   6.4708  17.6367
##
## Coefficients:
##              Estimate Std. Error z value  Pr(>|z|)
## (Intercept) 384.67658   30.09101 12.7838 < 2.2e-16
## VERBAL       -0.79214    0.19162 -4.1338 3.567e-05
## MATH          0.13957    0.19012  0.7341    0.4629
##
## Lambda: 0.74799 LR test value: 31.918 p-value: 1.6082e-08
## Numerical Hessian standard error of lambda: 0.087975
##
## Log likelihood: -176.8789
## ML residual variance (sigma squared): 66.477, (sigma: 8.1533)
## Number of observations: 49
## Number of parameters estimated: 5
## AIC: 363.76
```

As we can see from the above model that both VERBAL and MATH are significant since their P-Value is small enough; therefore, we shall keep both. The binnary weights model is $ELIGIBLE = 390.464541 - 0.675203 * VERBAL + 0.022018 * MATH$,and The row-normalized weights model is $ELIGIBLE = 384.67658 - 0.79214 * VERBAL + 0.1395 * MATH$. The ELIGIBLE has a negative relation with VERBAL(not make sense in the real world) and a positive relationship with MATH score. A possible reason VERBAL score has a negative relative with ELIGIBLE is that if a student spends too much time on the VERBAL test, he/she may not has time to work on MATH; further, it may lower down the total score of the SAT.