HW14

Zach White

November 3, 2016

```
climate = read.table("http://www.stat.duke.edu/courses/Fall10/sta290/datasets/climate.dat", header = TRU
climate$T.M = factor(climate$T.M, labels = c("T","M"))
titles = c("Mg/Ca", "Alkenone", "Faunal", "Sr/Ca", "Del180", "IceCore", "Pollen", "Noble Gas")
climate$proxy = factor(climate$proxy, labels = titles)

X = model.matrix(lm(deltaT~.,data = climate))

# I'm losing the Mg/Ca. Just do it manually
climate =read.table("http://www.stat.duke.edu/courses/Fall10/sta290/datasets/climate.dat", header = TRUE
climate$T.M = droplevels(factor(climate$T.M, labels = c("T","M")))
climate$proxy = droplevels(factor(climate$proxy, labels = titles))

dummies.prox = NULL
for(i in 1:length(titles)){
  dummies.prox =cbind(dummies.prox,as.numeric(climate$proxy == titles[i]))
}

colnames(dummies.prox) = titles

interaction1 = dummies.prox * climate$latitude
interaction2 = dummies.prox * climate$lat^2

interact.labels = c("Mg/Ca:lat1","Alkenoe:lat1","Faunal:lat1","Sr/Ca:lat1", "Del180:lat1","IceCore:lat1
interact2.labels = c("Mg/Ca:lat2","Alkenoe:lat2","Faunal:lat2","Sr/Ca:lat2", "Del180:lat2","IceCore:lat2
colnames(interaction1) = interact.labels
colnames(interaction2) = interact2.labels

Y= climate$deltaT
## Lasso
# Interactions first
X.int.first = as.matrix(cbind(interaction1,interaction2,dummies.prox,poly(climate$latitude,2)))
first.prelim = cv.glmnet(X.int.first,climate$deltaT,alpha = 1)
best.lam.first = first.prelim$lambda.min
interaction.first = glmnet(X.int.first,Y,alpha = 1, lambda = best.lam.first)
int.first = coef(interaction.first)

# Interactions flipped
X.int.flipped = as.matrix(cbind(interaction2,interaction1,dummies.prox,poly(climate$latitude,2)))
flipped.prelim = cv.glmnet(X.int.flipped,Y,alpha = 1)
best.lam.flipped = flipped.prelim$lambda.min
interaction.flipped = glmnet(X.int.flipped,Y,alpha = 1, lambda = best.lam.flipped)
int.flipped = coef(interaction.flipped)

# Poly first
X.poly.first = as.matrix(cbind(poly(climate$latitude,2),dummies.prox,interaction1,interaction2))
poly.prelim = cv.glmnet(X.poly.first,Y,alpha = 1)
best.lam.poly = poly.prelim$lambda.min
```

```r
poly.first = glmnet(X.poly.first,Y,alpha = 1, lambda = best.lam.poly)
poly.flipped = coef(poly.first)

round(int.first,3)
```

```
## 27 x 1 sparse Matrix of class "dgCMatrix"
##                        s0
## (Intercept)    -2.874
## Mg/Ca:lat1      .
## Alkenoe:lat1    .
## Faunal:lat1     0.006
## Sr/Ca:lat1      .
## Del180:lat1     0.011
## IceCore:lat1    0.141
## Pollen:lat1     .
## Noble Gas:lat1  .
## Mg/Ca:lat2      .
## Alkenoe:lat2    .
## Faunal:lat2     .
## Sr/Ca:lat2      .
## Del180:lat2     0.001
## IceCore:lat2    .
## Pollen:lat2    -0.002
## Noble Gas:lat2  .
## Mg/Ca           .
## Alkenone        0.072
## Faunal          0.574
## Sr/Ca          -1.666
## Del180          .
## IceCore        -0.120
## Pollen         -1.286
## Noble Gas      -1.977
## 1               .
## 2              -0.894
```

```r
round(int.flipped,3)
```

```
## 27 x 1 sparse Matrix of class "dgCMatrix"
##                        s0
## (Intercept)    -2.849
## Mg/Ca:lat2      .
## Alkenoe:lat2    .
## Faunal:lat2     .
## Sr/Ca:lat2      .
## Del180:lat2     0.001
## IceCore:lat2   -0.015
## Pollen:lat2    -0.002
## Noble Gas:lat2  .
## Mg/Ca:lat1      .
## Alkenoe:lat1    .
## Faunal:lat1     0.003
## Sr/Ca:lat1      .
## Del180:lat1     0.009
## IceCore:lat1    0.005
```

```
## Pollen:lat1     .
## Noble Gas:lat1  .
## Mg/Ca           .
## Alkenone        0.027
## Faunal          0.529
## Sr/Ca          -1.640
## Del180          .
## IceCore         .
## Pollen         -1.270
## Noble Gas      -1.962
## 1               .
## 2              -0.720
```

```r
round(poly.flipped,3)
```

```
## 27 x 1 sparse Matrix of class "dgCMatrix"
##                     s0
## (Intercept)    -2.684
## 1               .
## 2              -1.175
## Mg/Ca           .
## Alkenone        0.383
## Faunal          0.656
## Sr/Ca          -2.405
## Del180         -1.121
## IceCore        -2.294
## Pollen         -1.535
## Noble Gas      -2.699
## Mg/Ca:lat1      0.006
## Alkenoe:lat1   -0.004
## Faunal:lat1     0.038
## Sr/Ca:lat1      .
## Del180:lat1     0.013
## IceCore:lat1    0.031
## Pollen:lat1    -0.044
## Noble Gas:lat1  .
## Mg/Ca:lat2      .
## Alkenoe:lat2   -0.002
## Faunal:lat2    -0.001
## Sr/Ca:lat2      .
## Del180:lat2     0.004
## IceCore:lat2    0.000
## Pollen:lat2    -0.005
## Noble Gas:lat2  .
```

Discussion? Is this due to me just changing the λ

```r
# Bayesian Horseshoe
# Interactions first
X.int.first = as.matrix(cbind(interaction1,interaction2,dummies.prox,poly(climate$latitude,2)))
int.first.bhs = bhs(X.int.first,Y,T=10000)
int.first.bhs.coef = colMeans(int.first.bhs$beta)
names(int.first.bhs.coef) = c(interact.labels,interact2.labels,titles,"lat1","lat2")

# Interactions flipped
X.int.flipped = as.matrix(cbind(interaction2,interaction1,dummies.prox,poly(climate$latitude,2)))
```

```r
int.flipped.bhs = bhs(X.int.flipped,Y,T=10000)
int.flipped.bhs.coef = colMeans(int.flipped.bhs$beta)
names(int.flipped.bhs.coef) = c(interact2.labels,interact.labels,titles,"lat1","lat2")

# Poly first
X.poly.first = as.matrix(cbind(poly(climate$latitude,2),dummies.prox,interaction1,interaction2))
poly.first.bhs = bhs(X.poly.first,Y,T=10000)
poly.first.bhs.coef = colMeans(poly.first.bhs$beta)
names(poly.first.bhs.coef) = c("lat1","lat2",titles,interact.labels,interact2.labels)
```

```r
round(int.first.bhs.coef,3)
```

```
##     Mg/Ca:lat1   Alkenoe:lat1    Faunal:lat1     Sr/Ca:lat1    Del180:lat1
##          0.004         -0.002          0.010          0.002          0.010
##   IceCore:lat1    Pollen:lat1 Noble Gas:lat1     Mg/Ca:lat2   Alkenoe:lat2
##          0.062         -0.007         -0.005          0.000         -0.001
##    Faunal:lat2     Sr/Ca:lat2    Del180:lat2   IceCore:lat2    Pollen:lat2
##          0.000         -0.002          0.002         -0.006         -0.002
## Noble Gas:lat2          Mg/Ca       Alkenone         Faunal          Sr/Ca
##          0.000          0.382          0.501          0.822         -1.188
##         Del180        IceCore         Pollen      Noble Gas           lat1
##         -0.132         -0.535         -1.262         -1.757          0.116
##           lat2
##         -0.822
```

```r
round(int.flipped.bhs.coef,3)
```

```
##     Mg/Ca:lat2   Alkenoe:lat2    Faunal:lat2     Sr/Ca:lat2    Del180:lat2
##          0.000         -0.001          0.000         -0.002          0.002
##   IceCore:lat2    Pollen:lat2 Noble Gas:lat2     Mg/Ca:lat1   Alkenoe:lat1
##         -0.006         -0.002          0.000          0.004         -0.002
##    Faunal:lat1     Sr/Ca:lat1    Del180:lat1   IceCore:lat1    Pollen:lat1
##          0.010          0.002          0.010          0.060         -0.006
## Noble Gas:lat1          Mg/Ca       Alkenone         Faunal          Sr/Ca
##         -0.004          0.418          0.546          0.872         -1.134
##         Del180        IceCore         Pollen      Noble Gas           lat1
##         -0.131         -0.542         -1.221         -1.721          0.099
##           lat2
##         -0.858
```

```r
round(poly.first.bhs.coef,3)
```

```
##           lat1           lat2          Mg/Ca       Alkenone         Faunal
##          0.104         -0.812          0.374          0.493          0.805
##          Sr/Ca         Del180        IceCore         Pollen      Noble Gas
##         -1.157         -0.142         -0.535         -1.252         -1.781
##     Mg/Ca:lat1   Alkenoe:lat1    Faunal:lat1     Sr/Ca:lat1    Del180:lat1
##          0.004         -0.002          0.010          0.002          0.010
##   IceCore:lat1    Pollen:lat1 Noble Gas:lat1     Mg/Ca:lat2   Alkenoe:lat2
##          0.060         -0.007         -0.005          0.000         -0.001
##    Faunal:lat2     Sr/Ca:lat2    Del180:lat2   IceCore:lat2    Pollen:lat2
##          0.000         -0.002          0.002         -0.007         -0.002
## Noble Gas:lat2
##          0.000
```

It is interesting to note that in our first two lasso models where the interactions come first but are flipped, there isn't that much change. Many of the same variables that are shrunk to zero are still zero, and many of the coefficients are the same. However, when we put our poly term first, there are more terms included, which is suprising because in the lasso the order shouldn't matter.

Also, under our bayesian horseshoe models, it seems like there are consistently more coefficients not zero. As far as measures of uncertainty are concerned, it seems clear that lasso won't really have measures of uncertainty, but a clear benefit of doing the bayesian horseshe are the measures of uncertainty where we could calculate credible intervals quite easily