Regularization and Variable Selection for Parametric Models (1)

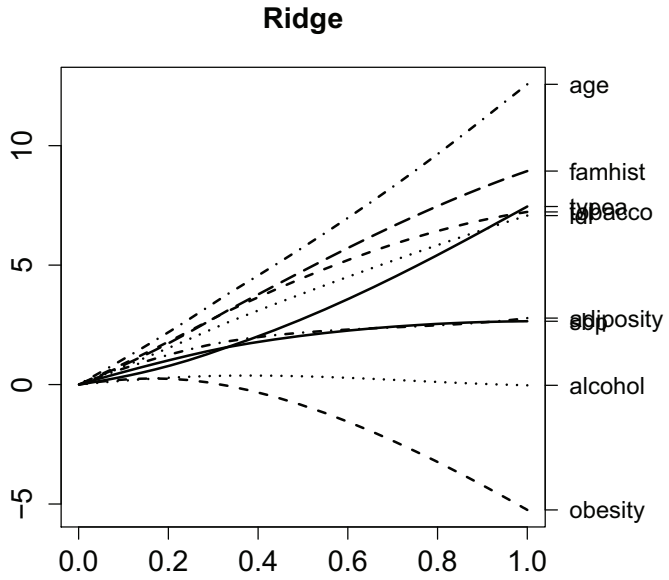
February 8, 2012

```
> library(catdata)
> data(heart, package="catdata")
> X<-heart[,-1]
> y<-heart[,1]
> X.std<-scale(X)
> p<-ncol(X)
> n<-length(y)
> family <- binomial()
> n.fold<-10
> ylab.text<-""
> xlab.text<-""
> Width = 6
> Height = 6
> oma.vec<-c(1,1,1,3)
> size.axis=1.4
> size.lab=1.4
> size.main=1.4
> size.right=1.2
> size.width=2.0
> colour=1

> library(lqa)
> library(glmnet)
```

RIDGE

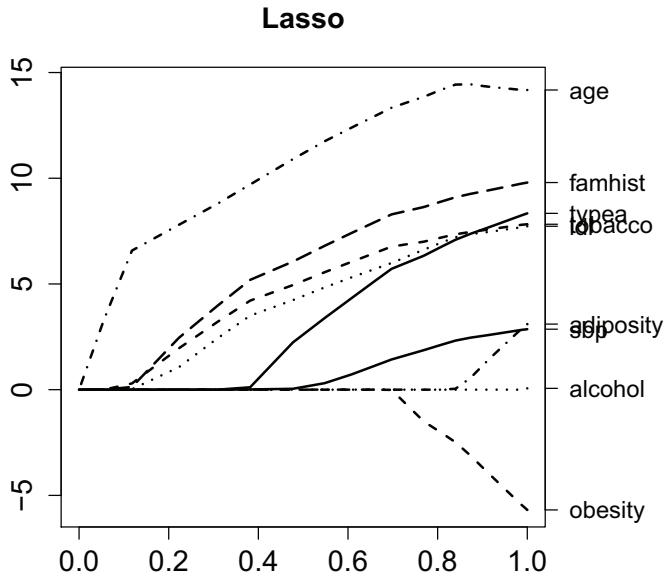
COEF BUILD-UPS

```
> main.text<-"Ridge"
> penalty.family<-ridge
> Plot.mat<-plot.lqa (y = y, x = X, family=family, penalty.family=penalty.family,
+ add.MLE = FALSE, ret.true=TRUE,really.plot = FALSE,show.standardized=TRUE,gamma=0.01)

> par(oma=oma.vec,cex.axis=size.axis,cex.lab=size.axis,cex.main=size.main)
> matplot(Plot.mat$s1,Plot.mat$beta.mat,type="l",ylab=ylab.text,xlab=xlab.text,
+ main=main.text,lwd=size.width, col=colour)
> axis(4, at = Plot.mat$beta.mat[1, ], labels = colnames(X), adj = 0, las = 1,
+ cex.axis=size.right)
```



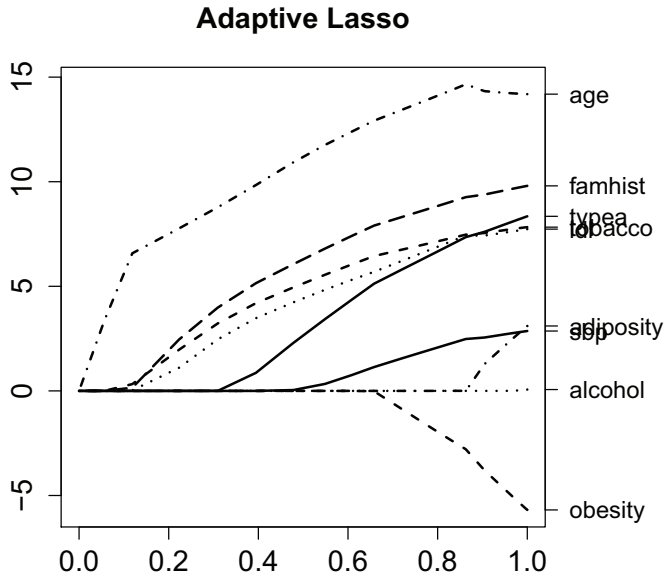
LASSO
COEF BUILD-UPS

```
> main.text<-"Lasso"
> penalty.family<-lasso
> Plot.mat<-plot.lqa (y = y, x = X, family=family, penalty.family=penalty.family,
+ add.MLE = TRUE, ret.true=TRUE,really.plot = FALSE,show.standardized=TRUE,gamma=0.5)

> par(oma=oma.vec,cex.axis=size.axis,cex.lab=size.lab,cex.main=size.main)
> matplot(Plot.mat$s1,Plot.mat$beta.mat,type="l",ylab=ylab.text,xlab=xlab.text,
+ main=main.text,lwd=size.width, col=colour)
> axis(4, at = Plot.mat$beta.mat[1, ], labels = colnames(X), adj = 0, las = 1,
+ cex.axis=size.right)
```

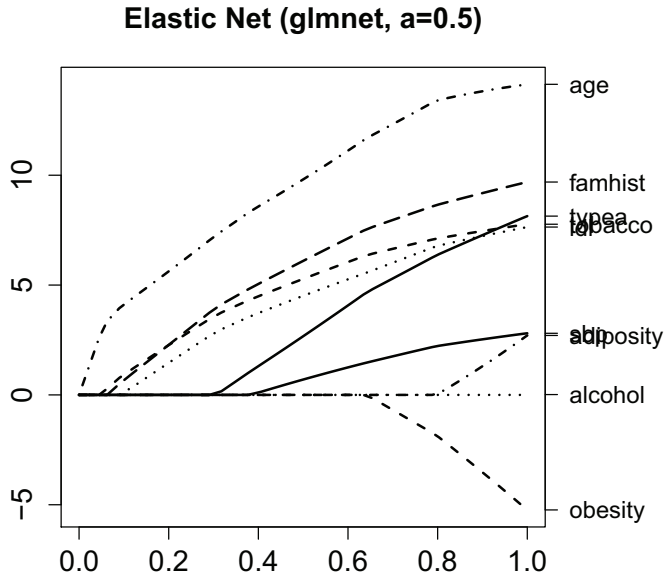


ADAPTIVE LASSO

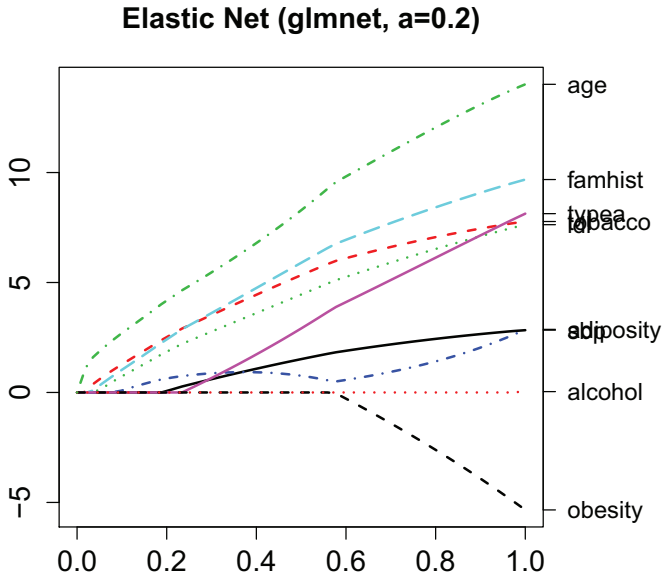
```
> #Tuning parameter
> g<-1
> ### Weights
>
> mle.std<-glm(y~X.std,family=binomial(),control=glm.control(maxit=1000))$coef[-1]
> w<-1/(abs(mle.std)^g)
> ### COEF BUILD-UPS
>
> main.text<-"Adaptive Lasso"
> penalty.family<-adaptive.lasso
> Plot.mat<-plot.lqa (y = y, x = X, family=family, penalty.family=penalty.family,
+ add.MLE = FALSE, ret.true=TRUE,really.plot = FALSE,show.standardized=TRUE,weigths=w)

> par(oma=oma.vec,cex.axis=size.axis,cex.lab=size.lab,cex.main=size.main)
> matplot(Plot.mat$s1,Plot.mat$beta.mat,type="l",ylab=ylab.text,xlab=xlab.text,
+ main=main.text,lwd=size.width, col=colour)
> axis(4, at = Plot.mat$beta.mat[1, ], labels = colnames(X), adj = 0, las = 1,
+ cex.axis=size.right)
```




```
> opt.min<-Inf
> ##### glmnet for fixed choice
> alpha<-0.5
> Path<-glmnet(X.std, y,family="binomial", alpha = alpha, nlambda = 100)

> par(oma=oma.vec,cex.axis=size.axis,cex.lab=size.axis,cex.main=size.main)
> matplot(colSums(abs(Path$beta))/max(colSums(abs(Path$beta))),t(Path$beta)*sqrt(n)
+ ,type="l",ylab=ylab.text,xlab=xlab.text,main="Elastic Net (glmnet, a=0.5)",
+ lwd=size.width, col=1)
> axis(4, at = Path$beta[,ncol(Path$beta)]*sqrt(n), labels = colnames(X), adj = 0,
+  las = 1,cex.axis=size.right)
```

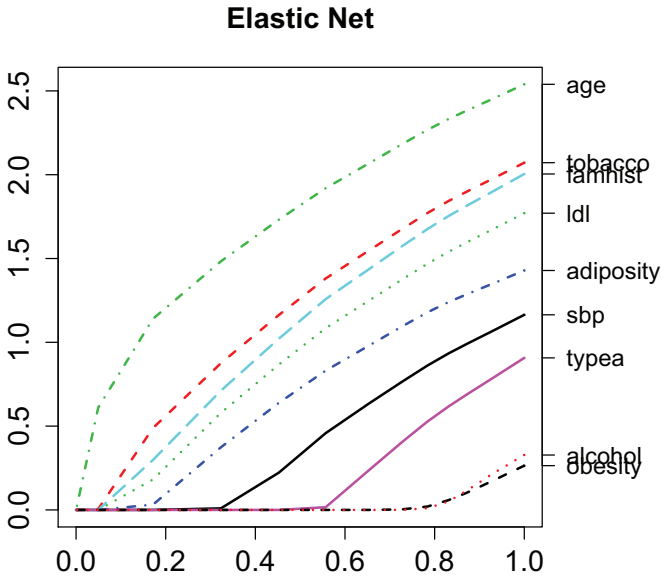



```
> alpha<-0.2
> Path<-glmnet(X.std, y,family="binomial", alpha = alpha, nlambda = 100)

> par(oma=oma.vec,cex.axis=size.axis,cex.lab=size.axis,cex.main=size.main)
> matplot(colSums(abs(Path$beta))/max(colSums(abs(Path$beta))),t(Path$beta)*sqrt(n)
+ ,type="l",ylab=ylab.text,xlab=xlab.text,main="Elastic Net (glmnet, a=0.2)",
+ lwd=size.width)
> axis(4, at = Path$beta[,ncol(Path$beta)]*sqrt(n), labels = colnames(X), adj = 0,
+ las = 1,cex.axis=size.right)
>
```



ELASTIC NET lqa with λ_1, λ_2
Fixed Tuning parameter

```
> lambda2 <- 1
> ### COEF BUILD-UPS
>
> main.text<-"Elastic Net"
> penalty.family<-enet
> Plot.mat<-plot.lqa (y = y, x = X, family=family, penalty.family=penalty.family,
+ offset.values = c (NA, lambda2),add.MLE = FALSE, ret.true=TRUE,really.plot = FALSE,
+ standardize = TRUE,show.standardized=TRUE)

> par(oma=oma.vec,cex.axis=size.axis,cex.lab=size.lab,cex.main=size.main)
> matplot(Plot.mat$s1,Plot.mat$beta.mat,type="l",ylab=ylab.text,xlab=xlab.text,
+ main=main.text,lwd=size.width)
> axis(4, at = Plot.mat$beta.mat[1, ], labels = colnames(X), adj = 0, las = 1,
+ cex.axis=size.right)
```


Elastic Net
0.0
0.5
1.0
1.5
2.0
2.5
0.0
0.2
0.4
0.6
0.8
1.0
age
tobacco
famhist
ldl
adiposity
sbp
typea
alcohol
obesity