

## Homework#3 Problem-4

Arinjay Jain

```
mod.ls <- lm(type ~ .-1, spam_df)
mod.ridge <- lm.ridge(type ~ ., spam_df)
mod.pcr <- pcr(formula = type ~ ., data = spam_df, validation = "CV")
mod.plsr <- plsr(formula = type ~ ., data = spam_df, validation = "CV")

mod.lasso <- lars( as.matrix(spam_df[,1:ncol(spam_df) - 1]) ,
spam_df[, ncol(spam_df)], type = "lasso")

mods.coeffs <- data.frame(ls = mod.ls$coef,
ridge = mod.ridge$coef,
lasso = mod.lasso$beta[10, ],
pcr = mod.pcr$coef[ , ,10],
plsr = mod.plsr$coef [ , ,10]
)
print(mods.coeffs)
```

	ls	ridge	lasso	pcr
## make	-3.320614e-02	-0.015210964	0.000000e+00	6.852457e-03
## address	1.445165e-03	-0.015544883	0.000000e+00	-7.179303e-03
## all	6.886632e-02	0.019799864	0.000000e+00	1.002829e-02
## num3d	1.488479e-02	0.016624613	0.000000e+00	1.607165e-02
## our	1.044914e-01	0.056624615	0.000000e+00	1.637237e-02
## over	1.556968e-01	0.032537865	0.000000e+00	4.027027e-03
## remove	2.361170e-01	0.083344676	1.288261e-01	8.388154e-03
## internet	1.148364e-01	0.037692167	0.000000e+00	7.743828e-03
## order	9.539755e-02	0.020190220	0.000000e+00	5.172981e-03
## mail	2.441980e-02	0.009713729	0.000000e+00	9.728780e-03
## receive	4.259627e-02	0.011457923	0.000000e+00	7.064517e-03
## will	-2.967378e-03	-0.024003875	0.000000e+00	2.059066e-02
## people	4.461982e-02	0.003583153	0.000000e+00	1.571913e-03
## report	1.782142e-02	0.001628810	0.000000e+00	6.847997e-04
## addresses	1.423489e-02	0.004794431	0.000000e+00	2.377096e-03
## free	8.738923e-02	0.061978312	1.811450e-02	2.070649e-02
## business	5.607731e-02	0.022962142	0.000000e+00	1.372234e-02
## email	6.586632e-02	0.029419879	0.000000e+00	1.028947e-02
## you	3.509620e-02	0.025090798	8.243847e-05	3.725894e-02
## credit	6.934306e-02	0.031460401	0.000000e+00	1.047542e-02
## your	6.644707e-02	0.063268269	6.072533e-02	1.208755e-01
## font	5.395411e-02	0.045915010	0.000000e+00	3.840729e-02
## num000	1.815880e-01	0.061223563	1.250773e-01	6.105218e-03
## money	8.676865e-02	0.040227122	0.000000e+00	9.082325e-03
## hp	-1.114529e-02	-0.038729258	-3.787290e-03	-4.212542e-02
## hpl	-1.132212e-02	-0.019181830	0.000000e+00	-1.561358e-02
## george	-2.894929e-03	-0.041081890	0.000000e+00	-1.364055e-02

## num650	6.756574e-03	0.002147293	0.000000e+00	-6.090028e-03
## lab	-8.492401e-03	-0.004419532	0.000000e+00	-3.257424e-03
## labs	-4.897516e-02	-0.023721288	0.000000e+00	-4.782031e-03
## telnet	-2.659938e-02	-0.009395594	0.000000e+00	-2.935618e-03
## num857	-9.721247e-02	0.002080130	0.000000e+00	-2.255951e-03
## data	-1.699704e-02	-0.023336347	0.000000e+00	-6.296787e-03
## num415	8.467819e-02	0.016846585	0.000000e+00	-2.288171e-03
## num85	-2.114293e-02	-0.016588055	0.000000e+00	-4.966276e-03
## technology	6.252226e-02	0.010660327	0.000000e+00	-4.296559e-03
## num1999	-6.526106e-03	-0.014062357	0.000000e+00	-9.551978e-03
## parts	-4.929705e-02	-0.011789989	0.000000e+00	7.146520e-06
## pm	-1.784907e-02	-0.008585837	0.000000e+00	-6.273293e-03
## direct	4.302282e-02	0.014261336	0.000000e+00	-6.161637e-04
## cs	8.134337e-04	-0.003020819	0.000000e+00	-6.413008e-03
## meeting	-2.260310e-02	-0.028312897	0.000000e+00	-6.677286e-03
## original	-4.790158e-02	-0.014152090	0.000000e+00	-2.938567e-03
## project	-1.349720e-02	-0.020137399	0.000000e+00	-4.080641e-03
## re	-2.208665e-02	-0.035661420	0.000000e+00	-5.493849e-02
## edu	-2.013634e-02	-0.034449413	0.000000e+00	-3.365203e-02
## table	-1.527856e-01	-0.014885396	0.000000e+00	8.938687e-05
## conference	-3.052703e-02	-0.016634503	0.000000e+00	-1.596912e-03
## charSemicolon	-1.182385e-01	-0.034107231	0.000000e+00	2.269575e-03
## charRoundbracket	5.610857e-02	-0.016208572	0.000000e+00	-4.638150e-03
## charSquarebracket	9.956557e-03	-0.006459196	0.000000e+00	-7.150786e-04
## charExclamation	7.633179e-02	0.055502858	5.673513e-03	5.091158e-03
## charDollar	2.381562e-01	0.057328041	1.309207e-01	3.895747e-03
## charHash	3.028379e-02	0.011888664	0.000000e+00	3.746403e-03
## capitalAve	5.830736e-04	0.007381713	0.000000e+00	3.925910e-04
## capitalLong	-4.652227e-05	0.013008326	0.000000e+00	1.737404e-04
## capitalTotal	1.123789e-04	0.048419005	1.580004e-05	1.169891e-04
##	plsr			
## make	-1.306140e-02			
## address	-1.153633e-02			
## all	4.446061e-02			
## num3d	6.722095e-03			
## our	9.698674e-02			
## over	9.236715e-02			
## remove	2.026510e-01			
## internet	1.014336e-01			
## order	5.973678e-02			
## mail	1.946713e-02			
## receive	3.302884e-02			
## will	-2.310075e-02			
## people	3.429487e-02			
## report	6.946563e-03			
## addresses	5.425777e-02			
## free	6.543878e-02			
## business	8.330377e-02			
## email	7.102060e-02			
## you	1.338095e-02			
## credit	7.700125e-02			
## your	5.404990e-02			
## font	3.429337e-02			
## num000	1.746184e-01			

```
## money          9.245503e-02
## hp             -2.470157e-02
## hpl            -2.229266e-02
## george         -1.140491e-02
## num650         -8.773347e-03
## lab            3.101716e-03
## labs           -2.837491e-02
## telnet         -2.995848e-03
## num857         1.745671e-02
## data           -5.346435e-02
## num415         1.766427e-02
## num85          -2.964026e-02
## technology     1.602170e-02
## num1999        -4.842552e-02
## parts          -2.283924e-02
## pm             -3.392767e-02
## direct         3.069870e-02
## cs             -1.679095e-02
## meeting        -4.197962e-02
## original       -3.304476e-02
## project        -3.934925e-02
## re             -3.343689e-02
## edu            -3.242758e-02
## table          -1.254560e-02
## conference     -4.159883e-02
## charSemicolon  -6.021949e-02
## charRoundbracket -3.403010e-02
## charSquarebracket -9.055554e-03
## charExclamation 6.446419e-02
## charDollar     1.282678e-01
## charHash       2.892184e-02
## capitalAve     3.150601e-04
## capitalLong    3.996749e-05
## capitalTotal   8.533364e-05
```

```
mods.coefs$xs = row.names(mods.coefs)
plot.data <- melt(mods.coefs, id ="xs")

ggplot (data = plot.data,
aes (x = factor(xs),y= value,group = variable,
colour = variable)) +
geom_line() +
geom_point() +
xlab("Factor") +
ylab("Regression Coefficient") +
scale_colour_hue(name ="Regression Method ",
labels = c("OLS",
"Ridge",
"Lasso",
"PCR",
"PLS")
)
```

