Partial Least Squares Regression for Generalized Linear Models

`plsRglm`

BIGslu Code Club

2021.09.16

Overview

- Intro to partial least squares (PLS)
- Intro to plsRglm
- Discussion uses and applications

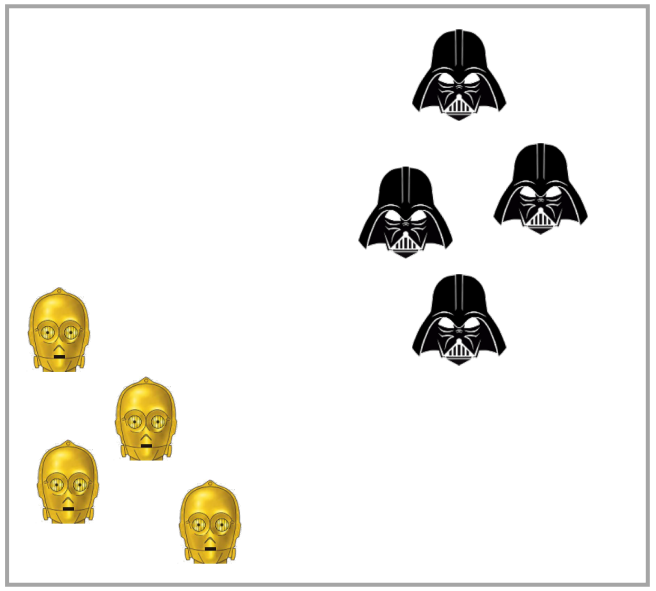
Principal component analysis (PCA)

	var1	var2	…	varN
obs1				
obs2				
…				
obsM				

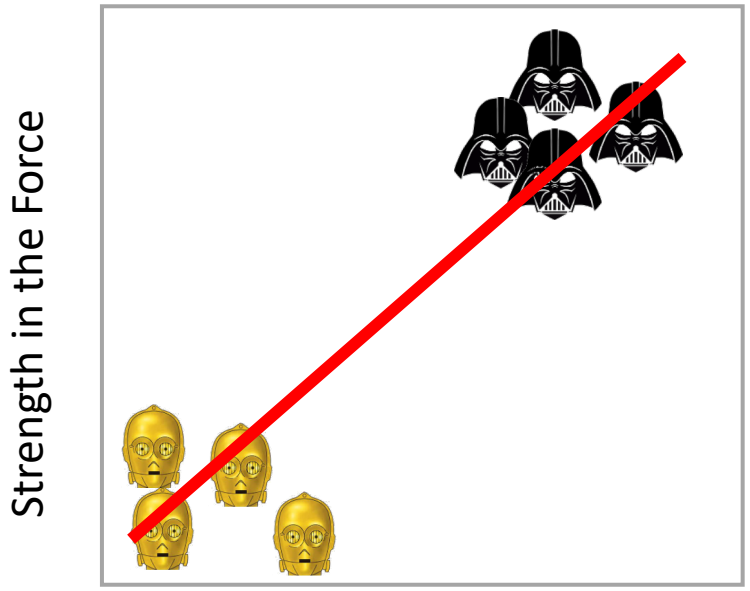
M x N

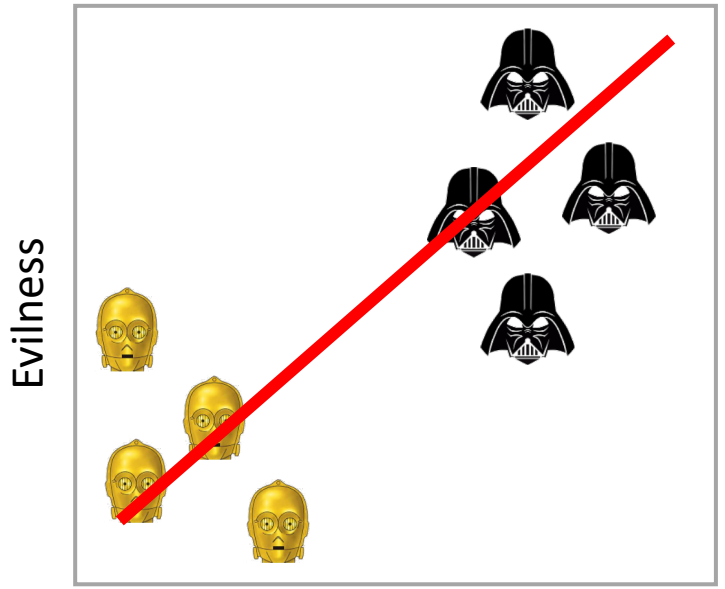
	PC1	PC2
obs1		
obs2		
…		
obsM		

M x 2
Explains 75%



Partial least squares (PLS) regression





...

Partial least squares (PLS) regression

Different versions can:

- Take one or more predictor matrices
- Take one or more outcome matrices
- Account for co-variance and correlation within predictors

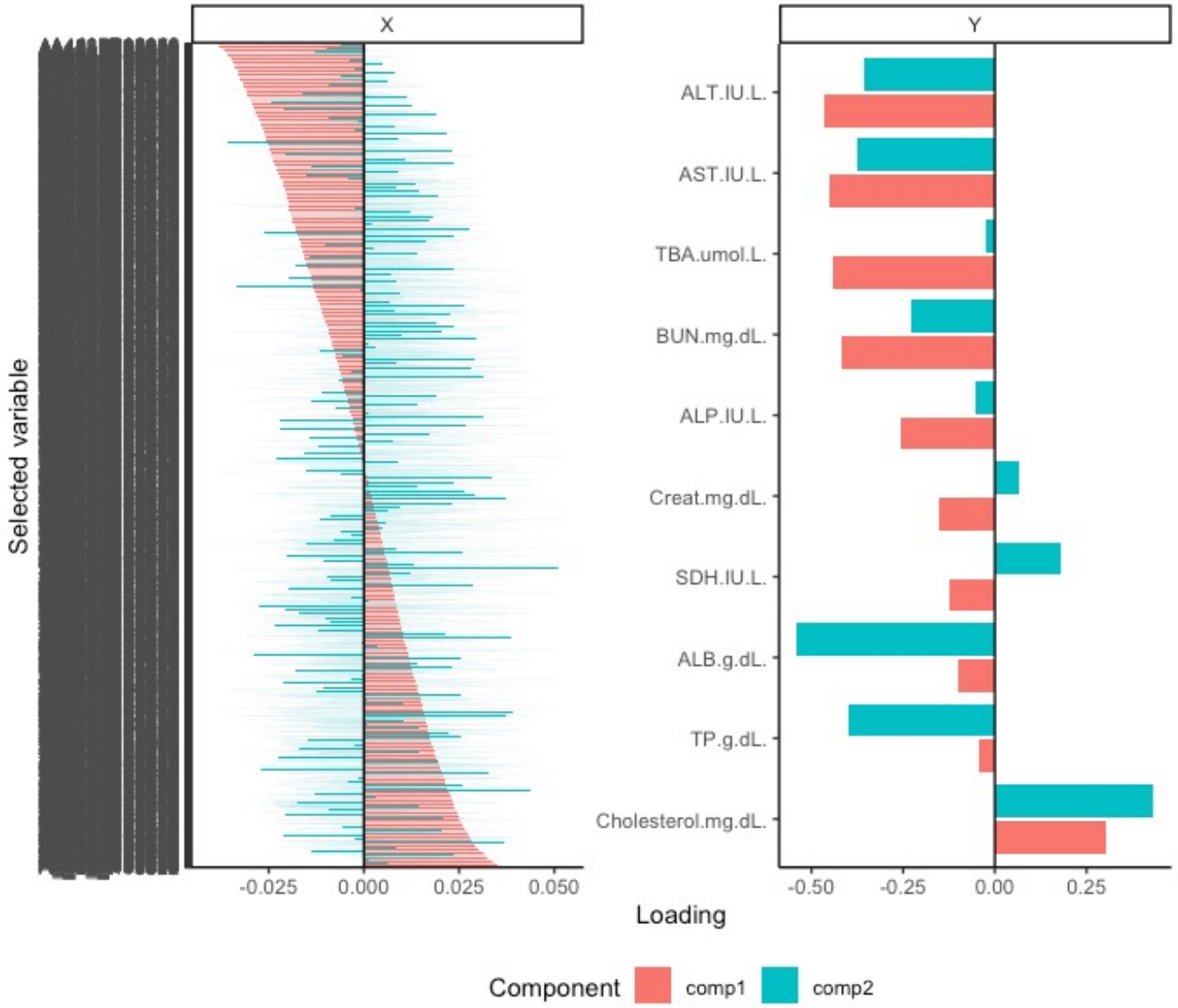
PLS in mixOmics

```
library(mixOmics)

data(liver.toxicity)
X <- liver.toxicity$gene
Y <- liver.toxicity$clinic

#PLS
result <- pls(X, Y, ncomp = 2)
tune.pls <- perf(result, validation = 'loo',
                 progressBar = FALSE)
```

PLS in mixOmics




```
loadX <- as.data.frame(result$loadings$X) %>%
  rownames_to_column(var) %>%
  mutate(space=X)
loadY <- as.data.frame(result$loadings$Y) %>%
  rownames_to_column(var)%>%
  mutate(space=Y)

loadXY <- bind_rows(loadY, loadX) %>%
  pivot_longer(comp1:comp2, names_to = comp) %>%
  filter(value != 0) %>%
  arrange(desc(value))

plot.dat %>%
  ggplot(aes(x=var, y=value, fill=comp)) +
  geom_bar(position = position_dodge2(width = 0.9,
                preserve = single),
                stat=identity) +
  coord_flip() +
  theme_classic() +
  facet_wrap(~space, scales=free, ncol=2) +
  labs(x=Selected variable, y=Loading,
                fill=Component) +
  geom_hline(yintercept=0) +
  theme(legend.position = bottom)
```

What if you model isn’t a simple regression of Y ~ X?

PLS in `plsRglm`

```
library(plsRglm)
data(Cornell)

cv.Modpls <- cv.plsR(Y~., data=Cornell, nt=2)
```

Input formula or give model type

- pls
- pls-glm-gaussian
- pls-glm-logistic
- pls-glm-poisson …

https://cran.r-project.org/web/packages/plsRglm/vignettes/plsRglm.pdf