

Support Vector Machines (SVM)

Francisco Caramelo

Tuesday, December 09, 2014

Neste exemplo vamos explorar a utilização de *SupportVectorMachines* para classificação. A base de dados a utilizar é a *Glass* e o que se pretende é a previsão do tipo de vidro que se obtém a partir de características da reacção química.

```
options(warn=-1)

library(e1071)
library(rpart)
library(mlbench)

data(Glass, package = 'mlbench')

head(Glass)
```

##		RI	Na	Mg	Al	Si	K	Ca	Ba	Fe	Type
##	1	1.52101	13.64	4.49	1.10	71.78	0.06	8.75	0	0.00	1
##	2	1.51761	13.89	3.60	1.36	72.73	0.48	7.83	0	0.00	1
##	3	1.51618	13.53	3.55	1.54	72.99	0.39	7.78	0	0.00	1
##	4	1.51766	13.21	3.69	1.29	72.61	0.57	8.22	0	0.00	1
##	5	1.51742	13.27	3.62	1.24	73.08	0.55	8.07	0	0.00	1
##	6	1.51596	12.79	3.61	1.62	72.97	0.64	8.07	0	0.26	1

Vamos dividir aleatoriamente os dados em dois conjuntos: um de treino e outro de teste

```
options(warn=-1)
ind <- 1:nrow(Glass)

testind <- sample(ind, trunc(length(ind)*2/3))
testSet <- Glass[testind,]
trainSet <- Glass[-testind,]
```

Usamos o conjunto de treino para ajustar o hiperplano, definindo o valor C (relativo à função de custo) e o valor γ (relativo ao kernel).

```
options(warn=-1)

svm.model <- svm(Type ~ ., data = trainSet, cost = 100, gamma = 1)
```

Determinamos os valores previstos no conjunto de teste pelo modelo ajustado e comparamos com os valores reais usando uma matrix confusão.

```
options(warn=-1)
library(knitr)
svm.pred <- predict(svm.model, testSet[, -10])
tb <- table(pred = svm.pred, true = testSet[, 10])

kable(tb)
```

	1	2	3	5	6	7
1	29	4	6	0	0	0
2	16	49	3	8	4	4
3	0	1	3	0	0	0
5	0	0	0	1	0	0
6	0	0	0	0	1	0
7	0	1	0	0	0	12

Exercício

1. Usando a base de dados *BreastCancer* da package *mlbench* ajuste um modelo SVM para prever o tipo de tumor classificado na coluna *Class*.
2. Determine a exactidão, a sensibilidade e especificidade do modelo no grupo de teste