

Exercício 6 - Aplicação LSSVM

Rodrigo Machado Fonseca - 2017002253

December 4, 2021

1 Introdução

Neste exercício utilizaremos o *Least Square Support Vector Machine* (LSSVM) para resolver o problema de classificação de tipos de vidros do banco de dados Glass (base nativa da Linguagem), a partir de suas características químicas.

2 Least Square SVM

Least Square SVM (LSSVM) são versões de mínimos quadrados de máquinas de vetores de suporte (SVM). Nesta versão, encontra-se a solução resolvendo um conjunto de equações lineares em vez de um problema de programação quadrática convexa para SVMs clássicos. Algoritmos eficientes e escaláveis, como aqueles baseados em gradiente, podem ser aplicados para resolver LSSVM.

A diferença entre o SVM e o LSSVM baseia-se, principalmente, na definição do erro. Neste caso definimos o erro como uma igualdade, conforme a equação a seguir:

$$\min_{w,v,e} J_p(w, v, e) = 0.5w^T w + \gamma 0.5 \sum_{i=1}^N e_i^2 \quad (1)$$

sujeito a

$$y_i[w^T \varphi(X_i) + b] = 1 - e_i, \quad i = 1, \dots, N \quad (2)$$

O γ é o parâmetro de regularização e estar relacionado as variáveis de folga.

O problema está sujeito as seguintes restrições:

$$w = \sum_i^N \alpha_i y_i \varphi(x_i) \quad (3)$$

$$\sum_i^N \alpha_i y_i = 0 \quad (4)$$

$$\alpha_i = \gamma e_i \quad \forall_i \quad (5)$$

$$y_i[w^T \varphi(X_i) + b] - 1 + e_i = 0 \quad \forall_i \quad (6)$$

A condição 3 possibilita que a LSSVM não resulte, necessariamente, em valores de α nulos, como em SVM. O parâmetro γ é uma proporcionalidade, não limitando assim o valor superior de α_i .

3 Metodologia

A priori, será carregada a base de dados Glass do R. A seguir pode-se ver uma pequena amostra da base de dados:

	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	1
2	1.51761	13.89	3.60	1.36	72.73	0.48	7.83	0	0	1
3	1.51618	13.53	3.55	1.54	72.99	0.39	7.78	0	0	1
4	1.51766	13.21	3.69	1.29	72.61	0.57	8.22	0	0	1
5	1.51742	13.27	3.62	1.24	73.08	0.55	8.07	0	0	1

Em sequência, iremos utilizar a normalização para que todas variáveis possam ter o mesmo peso no modelo, com a seguinte equação:

$$x'_i = \frac{x_i - \min(x)}{\max(x) - \min(x)} \quad (7)$$

Neste problema, as classes estão desbalanceadas, como pode ser visto a seguir, onde cada coluna representa o *label* e a linha representa a quantidade de dados de cada classe:

y	1	2	3	5	6	7
	70	76	17	13	9	29

Por fim, iremos separar 70% do conjunto para treino e 30% para teste.

4 Resultados

Neste experimento iremos para cada conjunto de parâmetros "kpar" e "tau"(γ) faremos ele 10 vezes o experimento e por fim analisaremos a média e o desvio padrão.

```
> set.seed(1)
> accuracy_final <- matrix(nrow=10, ncol=9)
> for(i in 1:10){
+   index <- createDataPartition(y,
+                                 p = 0.7,
+                                 list = FALSE)
+   x_train <- x[index,]
+   y_train <- y[index]
+   x_test <- x[-index,]
+   y_test <- y[-index]
+   r <- c(0.1, 0.1, 0.1, 0.2, 0.3, 0.4, 0.1, 0.1, 0.1)
+   tau <- c(0.3, 0.45, 0.5, 0.5, 0.5, 0.9, 0.2, 0.1, 0.8)
+   results <- training_lssvm(x_train, y_train, x_test, y_test, r, tau)
+   svm <- results[[1]]
+   accuracy_list <- results[[2]]
+   accuracy_final[i,] <- as.matrix(accuracy_list, nrow=1)
+ }
```

Nas tabelas a seguir estão representadas as médias e os desvios padrão para cada experimento.

	colMeans.accuracy_final.
1	0.5730159
2	0.5809524
3	0.5777778
4	0.5666667
5	0.5761905
6	0.5492063
7	0.5777778
8	0.5825397
9	0.5730159

	sd
1	0.03214041

2 0.03604103
3 0.03604103
4 0.02487335
5 0.03265883
6 0.02504161
7 0.02718568
8 0.02804732
9 0.03774827

5 Discussão

Neste exercício, assim como no anterior, os parâmetros da função utilizada para treinar foram ajustadas, a fim de obter o melhor resultado possível. Após o problema chegar a 58% de acurácia, ele estagnou e não foi possível aumentar a acurácia. Tentamos deixar o parâmetro "kpar" automático, porém isso não se motrou mais vantajoso.

É possível notar que a acurácia média obtida tem valor menor do que aquela obtida no exercício anterior, que tratava do mesmo problema, porém utilizando o classificador SVM. Este resultado nos leva a crer que as SVMs, apesar de apresentarem carga computacional maior, parecem apresentar desempenho um pouco melhor. O único ponto que devemos tomar cuidado é o fato que o as parâmetros foram definidos de forma arbitrária e sem nenhum rigor matemático.

Logo, com o experimento foi possível compreender melhor o comportamento do algoritmo LSSVM, além de que conseguimos construir e validar um classificador LSSVM. É possível afirmar que a prática foi bem executada e o classificador construído apresentou um resultado razoável.