Line H. Clemmensen
Section of Statistics and Data Analysis
DTU Compute

Exercises 02582
Module 5
Spring 2019

March 6, 2019

Topics: Optimal separating hyperplanes, convex optimizaion, support vector machines

Exercises:

1 Write up the analytical kernel matrix for a radial basis function and an input X with three observations (rows) and five features (columns). How big is the resulting Kernel matrix?

2 Apply and interpret Optimal Separating Hyperplanes (OSH):

(a) Compute OSH solutions for the non-overlapping 2D data in `synthetic_2D_nool_1.csv`. Play around with different choices of the kernel type `(linear, polynomial, or gaussian)` and the kernel parameter (the standard deviation for the Gaussian kernel or the degree for the polynomial kernel).

(b) Compute OSH solutions for the overlapping 2D data in `synthetic_2D_ol_1.csv`. Using a Gaussian kernel, try to find a value of the kernel parameter that leads to an acceptable separation between the two classes. What do you think about the generalization ability of your solution?

3 Apply and interpret Support Vector Machines (SVM):

(a) Given the data in `Ex3Data.csv`, pick the kernel type as well as values of λ and σ, which gave the best visual results. Explain your decision.

4 Use Lagrange multipliers to derive the optimization problem:

(a) Ridge regression is the solution to the minimization problem $\arg\min_{\boldsymbol{\beta}} \|\mathbf{y}-\mathbf{X}\boldsymbol{\beta}\|^2+\lambda\|\boldsymbol{\beta}\|^2$. We found the solution during the second exercise. An equivalent formulation is $\arg\min_{\boldsymbol{\beta}} \|\mathbf{y}-\mathbf{X}\boldsymbol{\beta}\|^2$ subject to $\|\boldsymbol{\beta}\|^2 \leq t$, a quadratic problem with a quadratic constraint. Solve this convex constrained minimization problem via thefollowing steps:

   (i) Rewrite the constraint in the form $g(x) \leq 0$

   (ii) Move the constraint into the minimization problem using a single Lagrange multiplier.

   (iii) Differentiate w.r.t. $\boldsymbol{\beta}$ and set to zero.

   (iv) Solve for β.

   Does the result look familiar to you?

5 Compute a suitable Support Vector Machine (SVM) classifier for the acute coronary syndrome data in `ACS.csv`. Use the matrix $\mathbf{X}_{truncated}$ as the data matrix and $\mathbf{y}$ as the response variable. Find suitable parameters for a kernel SVM using 10-fold cross-validation and misclassification rate (MCR) as criterion.

(a) You will need to preprocess data. Response should be 1 or -1 and it is a good idea to normalize data.

(b) Calculate the misclassification rate for the test data.

(c) With logistic regression you get a classification rate of 0.84. Can you do better with SVM?

Line H. Clemmensen
Section of Statistics and Data Analysis
DTU Compute

Exercises 02582
Module 5
Spring 2019

March 6, 2019

Topics: Optimal separating hyperplanes, convex optimizaion, support vector machines

Resources for this exercise:

Listing 1: Resources in Matlab

```
svm = fitcsvm(X,Y, ...) % fit SVM
contour(X1,X2,...) % plot contours
eval_svm(X_train, y_train, X_test, y_test, ...) % calculate
% misclassification errors
```

Listing 2: Resources in R

```
library(e1071) # SVM library
svm.model <- svm(Y ~ ., data=Tb, degree=deg, kernel=ker, gamma=gamm)
# fit SVM
```

Listing 3: Resources in Python

```
from sklearn.svm import SVC # Support Vector Classifier in sklearn
```

End of exercise