Foundations of Machine Learning — Homework Assignment 1

Anirudhan J Rajagopalan
N18824115
ajr619

October 11, 2015

C. Support Vector Machines

1

Installed the software from[2]. The installed version of software is also checked into github at[1].

2

See the following command:

```
$ ./svm-scale -s splice_noise_train.txt.range \
> splice_noise_train.txt > splice_noise_train.txt.scale
$ ./svm-scale -r splice_noise_train.txt.range \
> splice_noise_test.txt > splice_noise_test.txt.scale
```

3

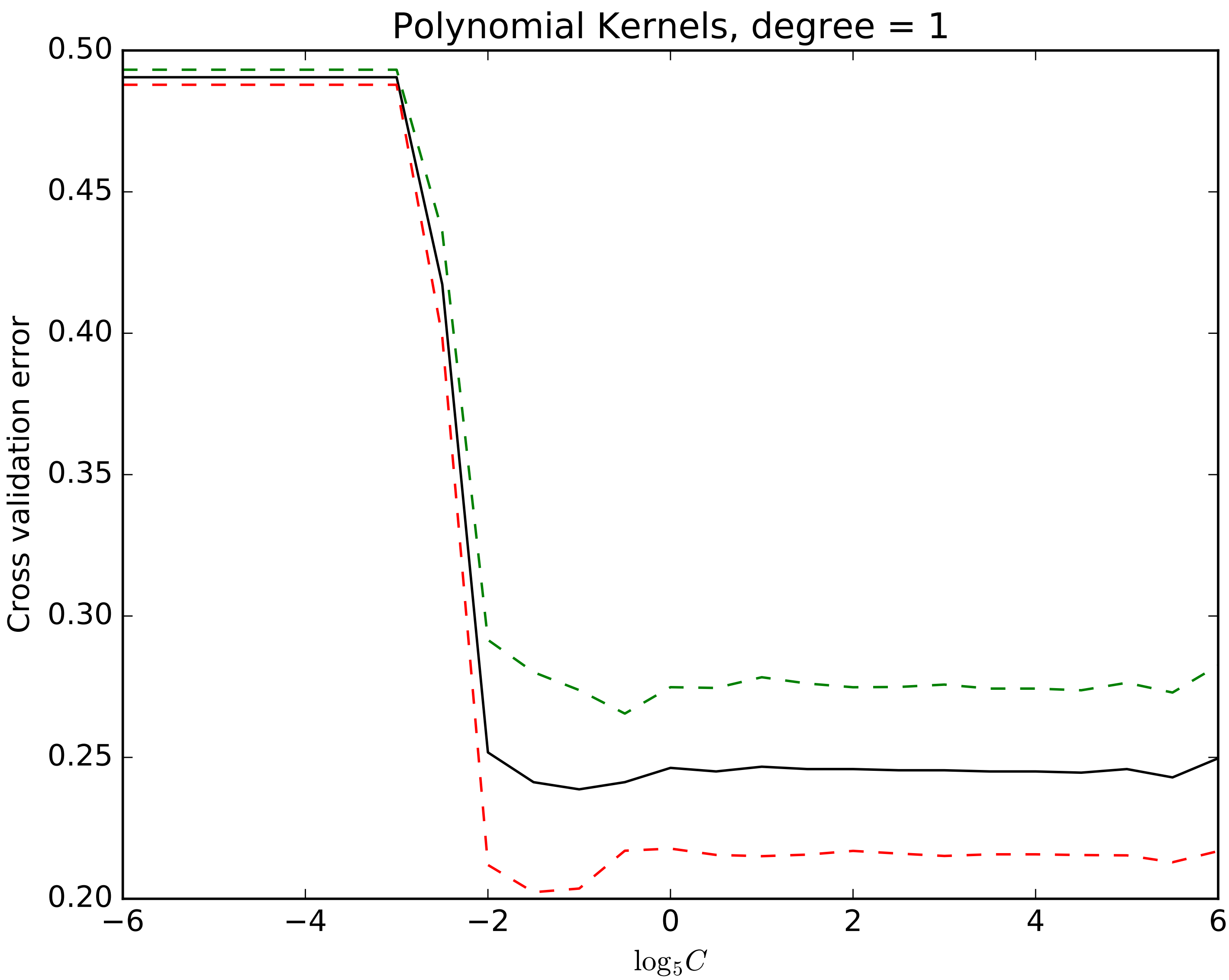
Run training and test script[3] by editing the KERNEL_DEGREE parameter for each value of d = 1, 3, 5.

```
$ python cross_validation.py > deg1.out # KERNEL_DEGREE = 1
$ python cross_validation.py > deg3.out # KERNEL_DEGREE = 3
$ python cross_validation.py > deg5.out # KERNEL_DEGREE = 5
```

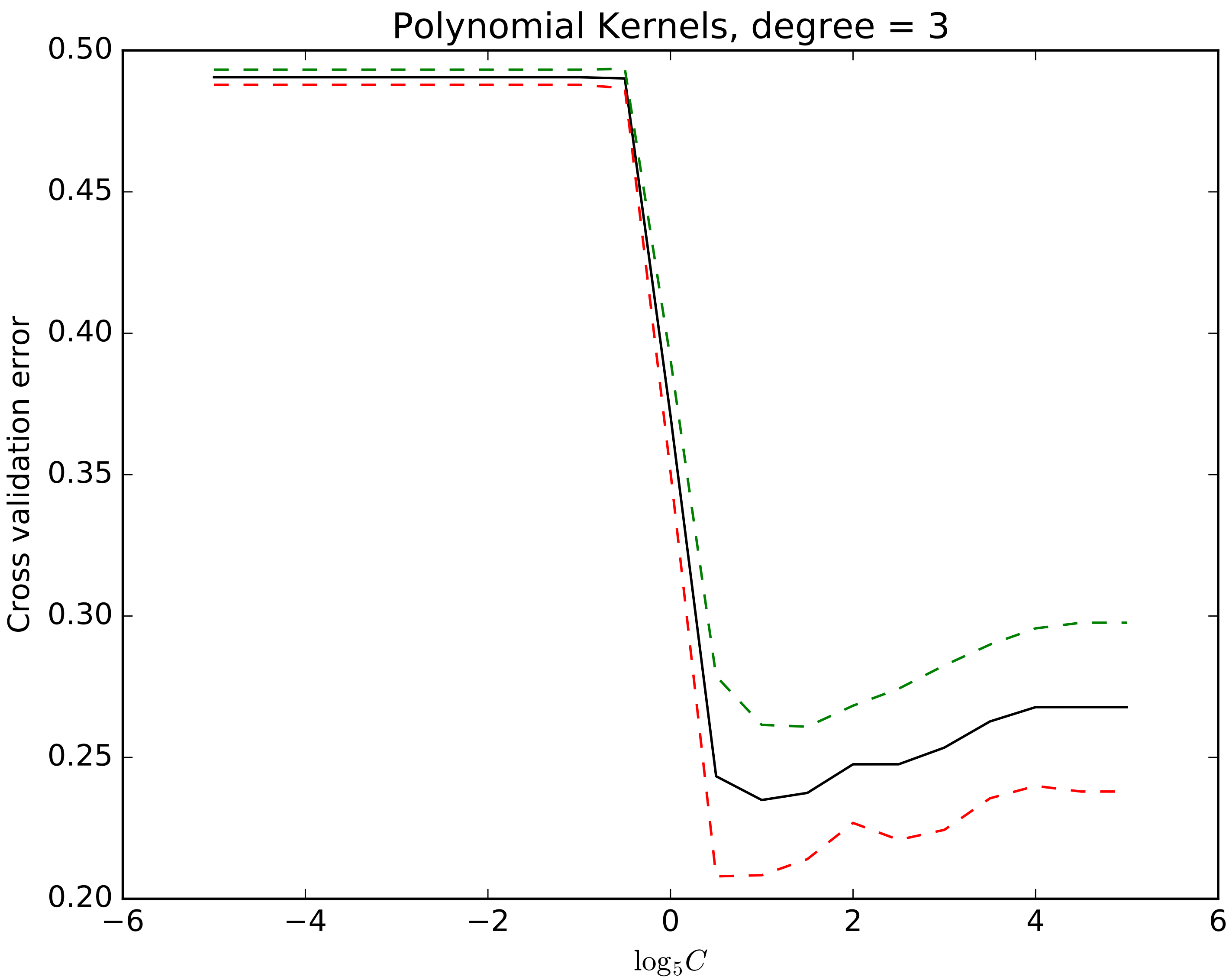
Filter the parameter and accuracy information from the run logs (deg1.out, deg3.out and deg5.out) by the command below.

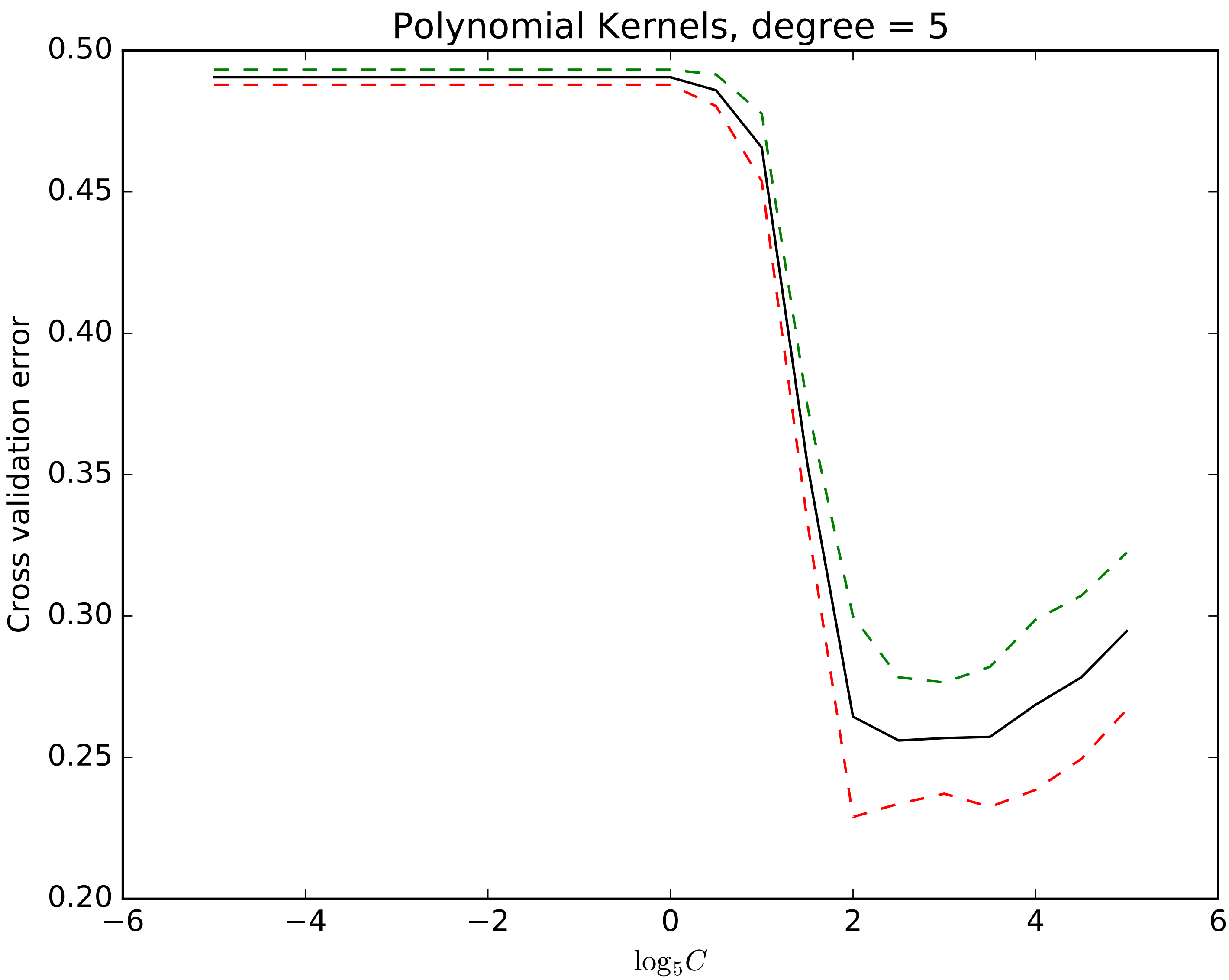
```
$ cat deg1.out | grep OUR | cut -d' ' \
> -f2,3,4,5,6 > deg1.out.filtered
$ cat deg3.out | grep OUR | cut -d' ' \
> -f2,3,4,5,6 > deg3.out.filtered
$ cat deg5.out | grep OUR | cut -d' ' \
> -f2,3,4,5,6 > deg5.out.filtered
```

Use plotter.py[4] to create plots from the output values for KERNEL_DEGREE values 1, 3, 5. The output will be saved as deg1.pdf, deg3.pdf and deg5.pdf. All the three plots are embedded below one after the other.



Polynomial Kernels, degree = 1
Cross validation error
0.50
0.45
0.40
0.35
0.30
0.25
0.20
−6
−4
−2
0
2
4
6
$\log_5 C$


Polynomial Kernels, degree = 3
Cross validation error
0.50
0.45
0.40
0.35
0.30
0.25
0.20
−6
−4
−2
0
2
4
6
$\log_5 C$


Polynomial Kernels, degree = 5
Cross validation error
0.50
0.45
0.40
0.35
0.30
0.25
0.20
−6
−4
−2
0
2
4
6
$\log_5 C$

Best values of c for polynomial kernels 1, 3, 5 are:

$$\begin{aligned} d &= 1 & c* &= 5^{-1.0} = 0.2 & \mathrm{cv-err} &= 23.87\% \pm 3.5\% \\ d &= 3 & c &= 5^{1.0} = 5 & \mathrm{cv-err} &= 23.49\% \pm 2.65\% \\ d &= 5 & c &= 5^{2.5} = 55.9017 & \mathrm{cv-err} &= 25.59\% \pm 2.23\% \end{aligned}$$

The best C measured in the cross-validation set is $C^* = 5^{1.0}$ with degree $d^* = 3$ which gives an average error of $23.49\% \pm 2.65\%$

4

5

6

D. Kernels

1

Given: Kernel, K is defined by $K(x, y) = \sum_{i=1}^{N} \cos^n(x_i^2 - y_i^2)$ for all $(X, Y) \in \mathbb{R}^N \times \mathbb{R}^N$

Solution: We know that

$$\cos(x_i^2 - y_i^2) = \sin(x_i^2).\sin(y_i^2) + \cos(x_i^2).\cos(y_i^2) \tag{1}$$

This can be written as a dot product of two vectors

$$\phi(x_i) = \begin{bmatrix} \cos(x_i^2) \\ \sin(x_i^2) \end{bmatrix} \qquad \text{and} \qquad \phi(y_i) = \begin{bmatrix} \cos(y_i^2) \\ \sin(y_i^2) \end{bmatrix} \tag{2}$$

We know that if K can be written as $\langle \phi(x_i), \phi(y_i) \rangle$, then it is a PDS@.

Also, $\langle \phi(x_i), \phi(y_i) \rangle$ is a scalar. When a scalar is raised to a positive power (n in our case) and summed with N other positive scalar, we get a positive scalar as our answer. Hence

$$K(x, y) = \sum_{i=1}^{N} \cos^n(x_i^2 - y_i^2) \mathrm{isPDS}.$$

References

[1] `http://git.io/v80yn`

[2] `http://www.csie.ntu.edu.tw/~cjlin/libsvm/`

[3] `http://git.io/v80yY`

[4] `http://git.io/v80yk`