ML HW4

Created	@November 18, 2022 4:46 PM
Tags	

Part 1:

Q1:

- k=10

```
kfold_data = cross_validation(x_train, y_train, k=10)
for i, (train_index, val_index) in enumerate(kfold_data):
    print("Split: %s, Training index: %s, Validation index: %s" % (i+
assert len(kfold_data) == 10 # should contain 10 fold of data
assert len(kfold_data[0]) == 2 # each element should contain train fo
assert kfold_data[0][1].shape[0] == 700 # The number of data in each
```

```
0.8s
```

```
(7000,)
Split: 1, Training index: 6300, Validation index: 700
Split: 2, Training index: 6300, Validation index: 700
Split: 3, Training index: 6300, Validation index: 700
Split: 4, Training index: 6300, Validation index: 700
Split: 5, Training index: 6300, Validation index: 700
Split: 6, Training index: 6300, Validation index: 700
Split: 7, Training index: 6300, Validation index: 700
Split: 8, Training index: 6300, Validation index: 700
Split: 9, Training index: 6300, Validation index: 700
Split: 10, Training index: 6300, Validation index: 700
```

- k=13

```
kfold_data = cross_validation(x_train, y_train, k=13)
for i, (train_index, val_index) in enumerate(kfold_data):
    print("Split: %s, Training index: %s, Validation index: %s" % (i+1,
assert len(kfold_data) == 10 # should contain 10 fold of data
assert len(kfold_data[0]) == 2 # each element should contain train fold
assert kfold_data[0][1].shape[0] == 700 # The number of data in each va
```

[22] 0.1s

```
(7000,)
Split: 1, Training index: 6461, Validation index: 539
Split: 2, Training index: 6461, Validation index: 539
Split: 3, Training index: 6461, Validation index: 539
Split: 4, Training index: 6461, Validation index: 539
Split: 5, Training index: 6461, Validation index: 539
Split: 6, Training index: 6461, Validation index: 539
Split: 7, Training index: 6462, Validation index: 538
Split: 8, Training index: 6462, Validation index: 538
Split: 9, Training index: 6462, Validation index: 538
Split: 10, Training index: 6462, Validation index: 538
Split: 11, Training index: 6462, Validation index: 538
Split: 12, Training index: 6462, Validation index: 538
Split: 13, Training index: 6462, Validation index: 538
```

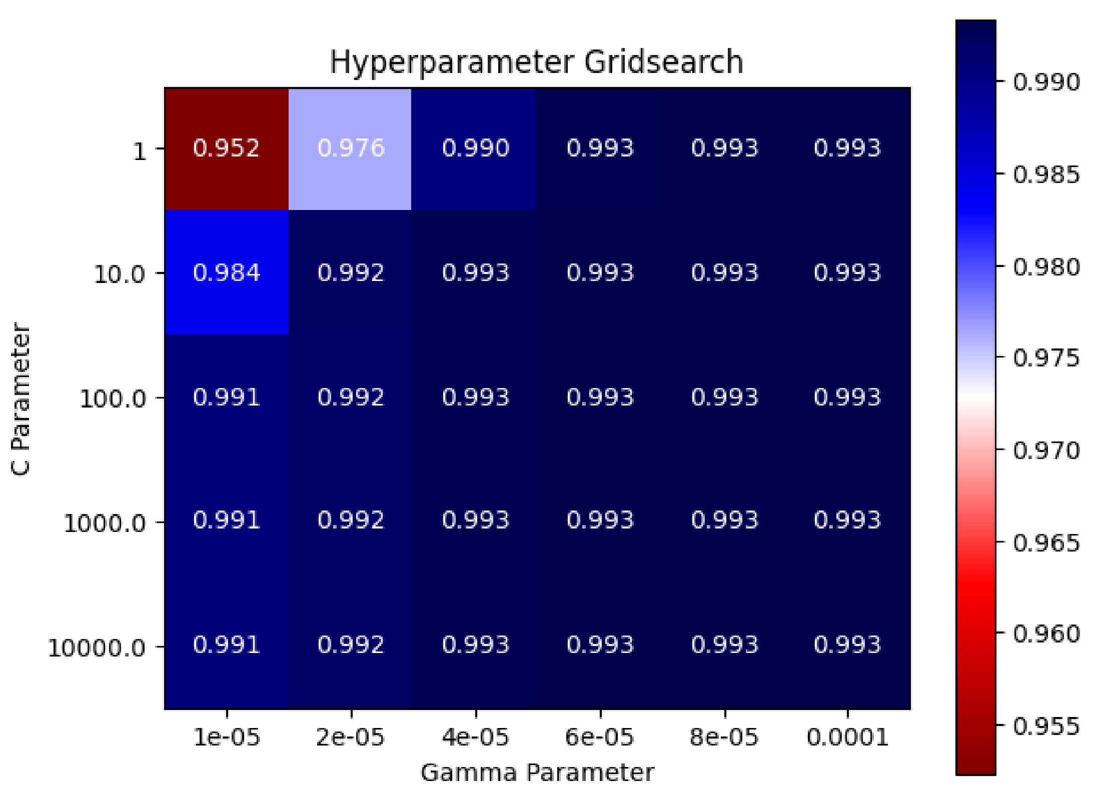
Q2(K=5):

```
print(len(kfold_data))
print(best_parameters)
# print(stat_table)
```

9] ✓ 0.7s

```
5
{'C': 10000.0, 'Gamma': 0.0001, 'kernel': 'rbf'}
```

Q3:


Hyperparameter Gridsearch
C Parameter
1
10.0
100.0
1000.0
10000.0
Gamma Parameter
1e-05
2e-05
4e-05
6e-05
8e-05
0.0001
0.952
0.976
0.990
0.993
0.993
0.993
0.984
0.992
0.993
0.993
0.993
0.993
0.991
0.992
0.993
0.993
0.993
0.993
0.991
0.992
0.993
0.993
0.993
0.993
0.991
0.992
0.993
0.993
0.993
0.993
0.990
0.985
0.980
0.975
0.970
0.965
0.960
0.955

Part 2:

Q1 :

$$\text{The kernel function of } \mathbf{x}_n \text{ and } \mathbf{x}_m,$$
$$k(\mathbf{x}_n,\ \mathbf{x}_m) = \phi(\mathbf{x}_n)^T\phi(\mathbf{x}_m),\ \text{is symmetric.}$$
$$\text{so we have } k(\mathbf{x}_n,\ \mathbf{x}_m) = k(\mathbf{x}_m,\ \mathbf{x}_n)$$
$$\text{Therefore the kernel matrix K is symmetric.}$$
$$\text{We can write K in the form } K = V\Lambda V^T,$$
$$\text{where V is an orthogonal matrix } \mathbf{v}_t \text{ and } \Lambda \text{ contains the eigenvalues of K}$$
$$\text{if K is positive semidefinite, all eigenvalues are nonnegative}$$
$$\text{Consider the mapping } \phi : \mathbf{x}_i \mapsto (\sqrt{\lambda_t}\mathbf{v}_{ti})_{t=1}^n$$
$$\text{We have } k(\mathbf{x}_i,\ \mathbf{x}_j) = \sum_{t=1}^{n}\lambda_t\mathbf{v}_{ti}\mathbf{v}_{tj} = (\mathbf{V\Lambda V^T})_{ij} = K_{ij} = \phi(\mathbf{x}_i)^T\phi(\mathbf{x}_j)...(1)$$
$$\text{Since } \lambda_t \text{ is in the square root, it has to be nonnegative for the mapping to be valid..(2)}$$
$$\text{Therefore K should be positive semidefinite}$$

Q2 :

$$\text{By Taylor Series, we have } e^x = \sum_{n=0}^{\infty}\frac{x^n}{n!}$$
$$\text{We have } k(\mathbf{x},\ \mathbf{x'}) \times k(\mathbf{x},\ \mathbf{x'}),$$
$$k(\mathbf{x},\ \mathbf{x'}) + k(\mathbf{x},\ \mathbf{x'}),$$
$$\text{and } ck(\mathbf{x},\ \mathbf{x'}) \text{ are also valid kernels when } k(\mathbf{x},\ \mathbf{x'}) \text{ is valid,}$$
$$\text{so } e^{k(\mathbf{x},\ \mathbf{x'})} \text{ is also valid.}$$

Q3 :

$\because k(x, x')$ is valid $\therefore K = \begin{array}{c c} & N \\ N & \begin{bmatrix} k(x_1,x_1) & \cdots & k(x_N,x_1) \\ \vdots & \ddots & \vdots \\ k(x_N,x_1) & \cdots & k(x_N,x_N) \end{bmatrix} \end{array}$ is Positive Semidefinite (P.S.D.)

a.) if $k_1(\underline{x}, \underline{x}')$ use the feature map $\phi_1(\underline{x})$

we can use another mapping $\phi : x \mapsto [\phi_1(\underline{x}), 1]^T$

$$\Rightarrow \langle \phi(\underline{x}), \phi(\underline{x}') \rangle = \langle \phi_1(\underline{x}), \phi_1(\underline{x}') \rangle + 1 = k_1(\underline{x}, \underline{x}') + 1 = k(\underline{x}, \underline{x}')$$

$\Rightarrow \underline{k(\underline{x}, \underline{x}') \text{ is valid}}$

c.) $K_2(\underline{x}, \underline{x}') = \exp(\|\underline{x}\|^2) \exp(\|\underline{x}'\|^2) = \exp(\|\underline{x}\|^2 + \|\underline{x}'\|^2)$

$$\Rightarrow G_2 = \begin{bmatrix} \exp(2\|\underline{x}_1\|^2) & \exp(\|\underline{x}_1\|^2+\|\underline{x}_2\|^2) & \exp(\|x_1\|^2+\|\underline{x}_3\|^2) & \cdots \\ \vdots & \exp(2\|\underline{x}_2\|^2) & \exp(\|x_2\|^2+\|\underline{x}_3\|^2) & \cdots \\ & & \exp(2\|\underline{x}_3\|^2) & \vdots \\ \exp(\|\underline{x}_1\|^2+\|\underline{x}_N\|^2) & \cdots & & \ddots \end{bmatrix}$$

$\exp(2\|\underline{x}_1\|^2) > 0$ (first principal minor)

$\exp(2\|\underline{x}_1\|^2 + 2\|\underline{x}_2\|^2) - \exp(\|\underline{x}_1\|^2+\|\underline{x}_2\|^2)\exp(\|\underline{x}_1\|^2+\|\underline{x}_2\|^2) = 0$ (2nd principal minor)

$$\exp(2\|\underline{x}_1\|^2+2\|\underline{x}_2\|^2+2\|\underline{x}_3\|^2) + \exp(\|\underline{x}_1\|^2+\|\underline{x}_2\|^2+\|x_3\|^2+\|\underline{x}_2\|^2+\|\underline{x}_1\|^2+\|\underline{x}_3\|^2)$$
$$+\exp(\|\underline{x}_1\|^2+\|x_3\|^2+\|\underline{x}_1\|^2+\|\underline{x}_2\|^2+\|x_3\|^2+\|\underline{x}_2\|^2) - \exp(2\|\underline{x}_1\|^2+2\|\underline{x}_2\|^2+2\|\underline{x}_3\|^2)$$
$$-\exp(\|\underline{x}_2\|^2+\|x_3\|^2+\|\underline{x}_3\|^2+\|\underline{x}_2\|^2+2\|\underline{x}_1\|^2) - \exp(2\|x_3\|^2+\|x_2\|^2+\|x_1\|^2+\|x_1\|^2+\|x_2\|^2) = 0$$

$\because$ 對稱的 $\therefore$ 後面都會剛好消掉

(用 cofactor 算的話)

$$\begin{array}{l} \cancel{2x_1 \quad 12 \quad 13 \quad 14} \to 0 \\ \cancel{21 \quad 2x_2 \quad 23 \quad 24} \to 0 \\ \cancel{31 \quad 32 \quad 2x_3} \quad 34 \to 2\cancel{x_1} + 2\cancel{x_2} + \cancel{x_4} + \cancel{x_3} + \cancel{x_2} + \cancel{x_1} + \cancel{x_4} + \cancel{x_2} + \cancel{x_1} + \cancel{x_3} + \cancel{41} + \cancel{12} + \cancel{23} \\ \cancel{41 \quad 42 \quad 43} \quad 2x_4 \to 0 \qquad - \cancel{13} - \cancel{2x_2} - \cancel{14} - \cancel{23} - \cancel{42} - \cancel{2x_1} - \cancel{43} - \cancel{21} - \cancel{12} = 0 \end{array}$$

$\Rightarrow$ by Sylvester's criterion, G_2 is P.S.D.

$\Rightarrow k_2(\underline{x}, \underline{x}')$ and $k_1(\underline{x}, \underline{x}')^2$ are valid

$\Rightarrow \underline{k(\underline{x}, \underline{x}') \text{ is valid.}}$

b.) if $k_1(\underline{x}, \underline{x}') = \begin{cases} 1 & , \underline{x} = \underline{x}' \\ 0 & , \sim \end{cases} \Rightarrow G_1 = \underline{I}_{N \times N} \Rightarrow$ P.S.D. , valid.

$\Rightarrow k(\underline{x}, \underline{x}') = k_1 - 1 = \begin{cases} 0 & , \underline{x} = \underline{x}' \\ -1 & , \sim \end{cases} \Rightarrow G = \begin{bmatrix} 0 & -1 & \cdots & -1 \\ -1 & 0 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \\ -1 & \vdots & \cdots & 0 \end{bmatrix}$ is not P.S.D.

e.g. $G = \begin{bmatrix} 0 & -1 & -1 \\ -1 & 0 & -1 \\ -1 & -1 & 0 \end{bmatrix}$,

< 0

$\det(\underline{G} - \lambda \underline{I}) = -\lambda^3 - 2 = 0 \Rightarrow \lambda = -\sqrt[3]{2}, \frac{\sqrt[3]{2} \pm \sqrt{2^{\frac{2}{3}} - 4 \cdot 2^{\frac{2}{3}}}}{2}$

$(\lambda + \sqrt[3]{2})(\lambda^2 - \sqrt[3]{2}\lambda + 2^{\frac{2}{3}})$ $\Rightarrow$ not valid.

d.) $k_1(x, x')$ is valid , $\underline{K}_1$ has eigenvalues ≥ 0

$\Rightarrow$ Sylvester's law of Inertia $\Rightarrow$ $\underline{K}_1$ has pivots ≥ 0

$\Rightarrow \exp(\underbrace{k_1(x, x')}_{\geq 0}) \geq 1 \Rightarrow \exp(k_1(\underline{x}, \underline{x}')) - 1 \geq 0$

$k_1(\underline{x}, \underline{x}')^2 \geq 0$ (only pivots are ensured)

$\Rightarrow k(\underline{x}, \underline{x}') = k_1(\underline{x}, \underline{x}')^2 + \exp(k_1(\underline{x}, \underline{x}')) - 1 \geq 0$

$\Rightarrow \underline{K}$ has all pivots ≥ 0 and $\underline{K}$ is symmetric

$\Rightarrow$ again, $\underline{K}$ has all eigenvalues ≥ 0.

Q4 :

To minimize $f(x)$ subject to an inequality constraint $\underline{g(x) \ge 0}$

Lagrangian function: $L(x, \lambda) = f(x) - \lambda g(x)$

$f(x) = (x-2)^2$, $\underline{g(x) = 3 - (x+3)(x-1)}$

with constraints : $\frac{\partial L(x,\lambda)}{\partial x} = 0 = 2(x-2) - \lambda(-2x-2)$

$$\Rightarrow 2x - 4 + 2\lambda x + 2\lambda = 0$$

$$\Rightarrow x = \frac{4-2\lambda}{2+2\lambda} \quad \cdots ①$$

$$\lambda \ge 0 \quad \cdots ②$$

dual representation :

$$L(\lambda) = \left(\frac{-\lambda+2}{\lambda+1} - 2\right)^2 - \lambda\left[3 - \left(\frac{-\lambda+2}{\lambda+1}\right)^2 - 2\left(\frac{-\lambda+2}{\lambda+1}\right) + 3\right]$$

subject to : $\lambda \ge 0$

$$= t^2 - 4t + 4 - 6\lambda + \lambda t^2 + 2\lambda t$$

$$= (\lambda+1)t^2 + (2\lambda-4)t + 4 - 6\lambda$$

$$= \frac{(-\lambda+2)^2}{\lambda+1} + \frac{(2\lambda-4)(-\lambda+2)}{\lambda+1} + \frac{(4-6\lambda)(\lambda+1)}{\lambda+1}$$

$$= \frac{(1-4+4)+(-2+8-8)+(-6-2+4)}{\lambda+1}$$

$$= \frac{-7\lambda^2+2\lambda}{\lambda+1} = 0$$

$$\frac{(-14\lambda+2)(\lambda+1)+(-7\lambda^2+2\lambda)}{\cancel{(\lambda+1)^2}}$$

$$x = \frac{4-2\lambda}{2+2\lambda} \approx 1.6\cdots \quad ✓$$

$$= -14\lambda^2 - 12\lambda + 2$$

$$\Rightarrow -21\lambda^2 - 10\lambda + 2 = 0$$

$$\frac{-10 \pm \sqrt{100 + \overset{268}{168}}}{42} \approx 0.15$$

- (橫線下面是驗算)，x和KKT算出來($\sqrt{7}-1$)差不多