0816170 Homework 4

Part 1

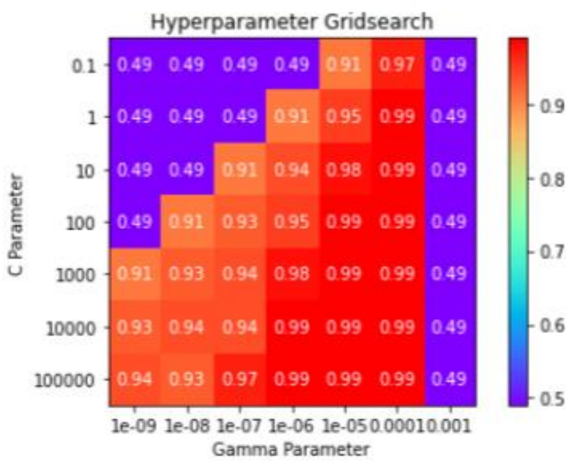
1. K-fold data partition: passed

```
[11]
kfold_data = cross_validation(x_train, y_train, k=10)
assert len(kfold_data) == 10 # should contain 10 fold of data
assert len(kfold_data[0]) == 2 # each element should contain train fold and validation fold
assert kfold_data[0][1].shape[0] == 700 # The number of data in each validation fold should equal to tr
# print(kfold_data)
```

2. Grid Search & Cross-validation

```
{'gamma': 0.0001, 'C': 1}
With accuracy score: 0.9934285714285714
```

3. Plot the grid search results



4. evaluate the performance on the test set

```
[34]
# best_parameters = {'gamma': 0.0001, 'C': 1}
best_C = best_parameters["C"]
best_Gamma = best_parameters["gamma"]
best_model = SVC(C = best_C, kernel = 'rbf',gamma = best_Gamma)
best_model.fit(x_train, y_train)
y_pred = best_model.predict(x_test)
```

```
y_pred = best_model.predict(x_test)
print("Accuracy score: ", accuracy_score(y_pred, y_test))
```

Part 2

1

"$\rightarrow$": A matrix is positive semidifinite means that all of its eigenvalues are non-negative

K is symmetric $\Rightarrow K = V \Lambda V^T$

orthonormal matrix ↲ ↳ contains eigenvalues λ

K is positive semidefinit $\Rightarrow$ all eigenvalues are non-negative

feature map: $\phi: x_i \mapsto (\sqrt{\lambda_t} V_{ti})_{t=1}^{n} \in \mathbb{R}^n$

We have $k(x_i, x_j) = \sum_{t=1}^{n} \lambda_t V_{ti} V_{tj} = (V \Lambda V^T)_{ij} = K_{ij} = \phi(x_i)^T \phi(x_j)$

"$\leftarrow$": λ_t is in the square root

$\Rightarrow \lambda_t$ must be nonnegative

$\Rightarrow K$ must be positive semidefinite

2

According to Taylor expansion around 0:

$$\exp(k_1(x,x')) = \exp(0) + \exp(0) k_1(x,x') + \frac{\exp(0)}{2!} k_1(x,x')^2 + \frac{\exp(0)}{3!} k_1(x,x')^3 + \frac{\exp(0)}{4!} k_1(x,x')^4 + \dots$$

$\Rightarrow$ the exponential of $k_1(x,x')$ is an infinite series of multiplications and additions of $k_1(x,x')$

we know that $k_3(x,x') = k_1(x,x') \cdot k_2(x,x')$

$k_4(x,x') = \alpha k_1(x,x') + \beta k_2(x,x')$

k_3, k_4 are valid kernels

$\Rightarrow$ $\exp(k_1(x,x'))$ is a valid kernel

3

a

Consider a positive coefficients polynomial function: $f(y) = y+1$, let $y = k_1(x,x')$

$\Rightarrow k(x,x') = k_1(x,x')+1 = f(k_1(x,x'))$

According to the texebook, we can know that

$k_1(x,x')+1$ is a valid kernel

b give an example:

let $k(x,x') = -\frac{2}{3}$, $K = \begin{bmatrix} -\frac{2}{3} & -\frac{2}{3} \\ -\frac{2}{3} & -\frac{2}{3} \end{bmatrix}$

K is not positive semidefinite

$\Uparrow$

negative

$\uparrow$

$K - \lambda I = \begin{bmatrix} -\frac{2}{3} - \lambda & -\frac{2}{3} \\ -\frac{2}{3} & -\frac{2}{3} - \lambda \end{bmatrix}$

$\|K - \lambda I\| = (\lambda + \frac{2}{3})^2 - \frac{4}{9} = \lambda(\lambda + \frac{4}{3}) \Rightarrow \lambda = 0 \text{ or } -\frac{4}{3}$

$\Rightarrow$ invalid

c $\exp(\|X\|^2) \cdot \exp(\|X'\|^2)$

$$= \left(1 + \|X\|^2 + \frac{(\|X\|^2)^2}{2!} + \frac{(\|X\|^2)^3}{3!} + \cdots\right)\left(1 + \|X'\|^2 + \frac{(\|X'\|^2)^2}{2!} + \frac{(\|X'\|^2)^3}{3!} + \cdots\right)$$

$\|X\|^2, \|X'\|^2 \geq 0 \Rightarrow \exp(\|X\|^2) \cdot \exp(\|X'\|^2) \geq 0$

$\Rightarrow \exp(\|X\|^2) \cdot \exp(\|X'\|^2)$ is a valid kernel

we know that $k_4(X, X') = k_2(X, X') \cdot k_3(X, X')$ if $k_2(x, x')$ $k_3(x, x')$ are valid, $k_4(x, x')$ is also valid

$\Rightarrow k_1(X, X')^2$ is a valid kernel

The sum of two valid kernel is also valid kernel

$\Rightarrow k_1(x, x')^2 + \exp(\|X\|^2) \cdot \exp(\|X'\|^2) = k(X, X')$ is a valid kernel.

d According to Taylor expansion around 0:

$$\exp(k_1(x, x')) = \exp(0) + \exp(0) k_1(x, x') + \frac{\exp(0)}{2!} k_1(x, x')^2 + \frac{\exp(0)}{3!} k_1(x, x')^3 + \frac{\exp(0)}{4!} k_1(x, x')^4 + \cdots$$

$$\Rightarrow k(x, x') = k_1(x, x')^2 \cancel{-1} \cancel{+1} + k_1(x, x') + \frac{1}{2!} k_1(x, x')^2 + \cdots$$

$k(X, X')$ is an infinite series of multiplications and additions of $k_1(x, x')$

we know that $k_3(X, X') = k_1(X, X') \cdot k_2(X, X')$

$k_4(X, X') = \alpha k_1(X, X') + \beta k_2(x, x')$

if $k_1(x, x')$ $k_2(X, X')$ are valid, both $k_3(x, x')$ $k_4(x, x')$ are valid kernels

$\Rightarrow k(X, X')$ is a valid kernel

4 minimize x^2-4x+4

subject to $x^2+2x-6 \leq 0$

Lagrangian function:

$$L(x,\lambda) = x^2-4x+4+\lambda(x^2+2x-6)$$
$$= (1+\lambda)x^2+(2\lambda-4)x+(4-6\lambda)$$

minimize over x:

$$\frac{\partial L}{\partial x} = 0 \Rightarrow (1+\lambda)x+2\lambda-4=0$$

For $\lambda \neq -1$, the Lagrangian reaches its inimum at $\tilde{x} = \frac{2-\lambda}{1+\lambda}$

$\Rightarrow$ the dual problem:

$$\frac{\lambda^2-4\lambda+4}{1+\lambda}+\frac{-2\lambda^2+8\lambda-8}{1+\lambda}+4-6\lambda$$

$$\underset{\lambda}{\text{maximize}} : L(\lambda) = \frac{-(\lambda-2)^2}{1+\lambda}-6\lambda+4 = \frac{-7\lambda^2+2\lambda}{\lambda+1}$$

subject to $\lambda \geq 0$